Original article

Discovery of potential new InhA direct inhibitors based on pharmacophore and 3D-QSAR analysis followed by in silico screening

Xiao-Yun Lu [a], Ya-Dong Chen [a], Yong-Jun Jiang [b], Qi-Dong You [a],*

[a] Department of Medicinal Chemistry, China Pharmaceutical University, 24 Tongjiaxiang, Nanjing 210009, PR China
[b] Ningbo Institute of Technology, Zhejiang University, Ningbo 315104, PR China

ABSTRACT

This study develops an efficient approach for discovering new InhA direct inhibitors in theory. The InhA-bound conformation of a pyrrolidine carboxamide inhibitor was used to build a pharmacophore model. This model with feature-shape query was successfully used to identify and align the bioactive conformations of pyrrolidine carboxamide analogues and screen SPECS database. A statistically valid 3D-QSAR with good results ($r^2_{cv} = 0.660$ and $r^2 = 0.962$) was obtained. From database screening, 30 hits were selected and identified as potential leads, which exhibit good estimated activities by 3D-QSAR model. Docking studies were carried out on two representative hits to analyze their interactions with InhA. Also, the interactions between existing pyrazole inhibitors and InhA were explored based on the pharmacophore model.

© 2009 Elsevier Masson SAS. All rights reserved.

1. Introduction

Tuberculosis (TB) is an ancient infectious disease of global influence, re-emerged with multi-drug resistant strains (MDR-TB) and acquired immune deficiency syndrome (AIDS). According to World Health Organization (WHO), one third of the world's population are infected with *Mycobacterium tuberculosis* (MTB) and 8.2 million new TB cases will occur worldwide in 2020 [1,2]. Furthermore, it has been more than 40 years since a new drug for TB was discovered. Therefore, it is an imperative need to develop novel anti-tubercular drugs that can be equally effective against MTB and MDR-TB, and shorten the duration of therapy.

Mycolic acids, the major components of the cell wall of MTB, are unusually long chain α-alkyl, β-hydroxy of fatty acids of 60–90 carbons [3]. The enzymes involved in the biosynthesis of mycolic acids are attractive targets for designing new anti-tubercular agents. The NADH-dependent enoyl acyl carrier protein reductase, InhA, which is encoded by the mycobacterium gene *inhA*, is a key enzyme in fatty acids' elongation. It catalyzes the NADH-dependent stereospecific reduction of α,β-unsaturated fatty acids bound to the acyl carrier protein [4]. It is now well established that InhA is the primary target of isoniazid (INH) [5], a front-line agent for the treatment of TB. As a prodrug, INH must be first activated by KatG, a catalase–peroxidase enzyme with dual activities of catalase and peroxidase oxidizing INH to an acyl radical binding to the position 4 of nicotinamide adenine dinucleotide (NAD) to form an active INH–NAD adduct [6]. Clinical studies indicate that a significant number of the strains resistant to INH arise from mutations in KatG [7]. Thus, inhibitors targeting InhA directly without this activation requirement would be promising candidates for the development of novel anti-tubercular agents.

More recently, it has been reported that broad-spectrum antibacterial inhibitor triclosan (TCN) also actually targets InhA without KatG activation [8]. Based on the mechanism of action of TCN, a novel class of InhA direct inhibitors, diphenyl ethers, was designed using structure-based drug design [9,10]. The great success engendered by this method prompted scientists to apply high-throughput screening technology to discover other InhA direct inhibitors. Based on this basic viewpoint, some novel classes of InhA direct inhibitors have been identified, such as indole-5-amides [11,12], pyrazole derivatives [11,12], pyrrolidine carboxamides [13], arylamides [14], imidazopiperidines [15], etc. Interestingly, it is found that all of these InhA direct inhibitors have conserved interactions with the active site of InhA, a hydrogen bond network with the active amino acid Tyr158 and the 2'-hydroxyl group of the nicotinamide ribose of the nucleotide and stacking interactions with the NAD$^+$ nicotinamide

Table 1
Structures, pIC$_{50}$ values (experimental and predicted) and residuals of pyrrolidine carboxamide compounds.

Compd	R$_1$	R$_2$	IC$_{50}$ (μM)	pIC$_{50}$	CoMSIA	
					PA[b]	Δ[c]
1[a]	Cyclohexyl		10.66 ± 0.51	4.97	5.08	−0.11
2	Cyclohexyl		34.88 ± 2.06	4.46	4.42	0.04
3[a]	Cyclohexyl		0.89 ± 0.05	6.05	5.36	0.69
4	Cyclohexyl		28.02 ± 4.29	4.55	4.69	−0.14
5	Cyclohexyl		1.35 ± 0.05	5.87	5.89	−0.02
6	Cyclohexyl		14.50 ± 0.79	4.84	4.69	0.15
7[a]	Cyclohexyl		16.79 ± 0.54	4.77	5.27	−0.5
8	Cyclohexyl		3.51 ± 0.09	5.45	5.41	0.04
9	Cyclohexyl		10.59 ± 0.48	4.98	5.15	−0.17
10	Cyclohexyl		5.55 ± 0.21	5.26	5.47	−0.21
11	Cyclohexyl		73.58 ± 9.97	4.13	4.43	−0.3
12	Cyclohexyl		56.02 ± 10.23	4.25	4.38	−0.13
13	Cyclohexyl		0.97 ± 0.03	6.01	6.17	−0.16

(*continued on next page*)

Table 1 (continued)
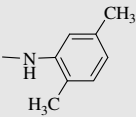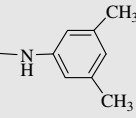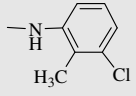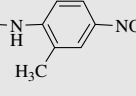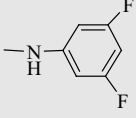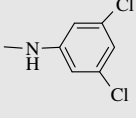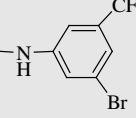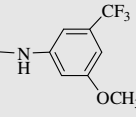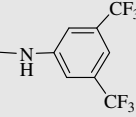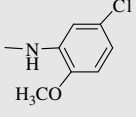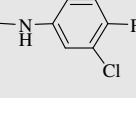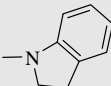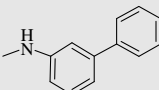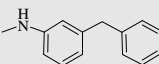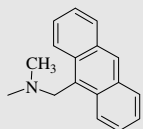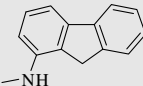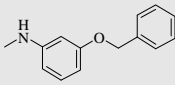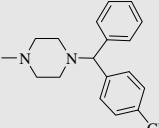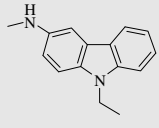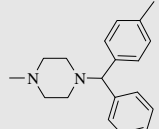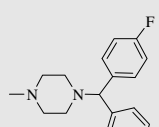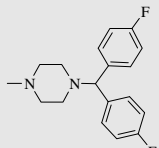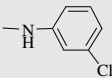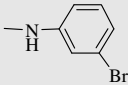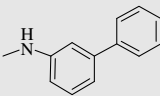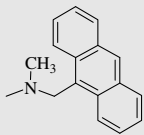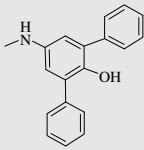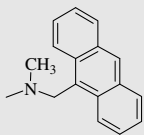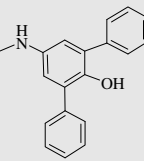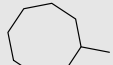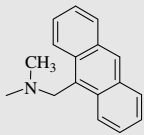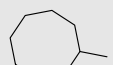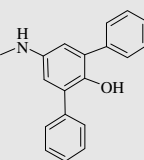
| Compd | R₁ | R₂ | IC₅₀ (µM) | pIC₅₀ | CoMSIA | |
					PA[b]	Δ[c]
14	Cyclohexyl		10.05 ± 0.33	5.00	4.97	0.03
15	Cyclohexyl		3.14 ± 0.12	5.50	5.43	0.07
16	Cyclohexyl		23.12 ± 1.00	4.64	4.62	0.02
17	Cyclohexyl		31.37 ± 1.45	4.50	4.39	0.11
18[a]	Cyclohexyl		1.49 ± 0.05	5.83	6.19	−0.36
19	Cyclohexyl		0.39 ± 0.01	6.41	6.18	0.23
20	Cyclohexyl		0.85 ± 0.05	6.07	6.04	0.03
21	Cyclohexyl		1.30 ± 0.04	5.89	5.93	−0.04
22[a]	Cyclohexyl		3.67 ± 0.17	5.44	5.48	−0.04
23	Cyclohexyl		1.60 ± 0.06	5.80	5.74	0.06
24	Cyclohexyl		14.83 ± 0.98	4.83	4.93	−0.1

Table 1 (continued)

Compd	R$_1$	R$_2$	IC$_{50}$ (μM)	pIC$_{50}$	CoMSIA	
					PA[b]	Δ[c]
25	Cyclohexyl		5.1	5.29	5.31	−0.02
26[a]	Cyclohexyl		0.39 ± 0.01	6.41	6.19	0.22
27[a]	Cyclohexyl		0.41 ± 0.01	6.39	6.40	−0.01
28	Cyclohexyl		0.75 ± 0.04	6.12	5.98	0.14
29	Cyclohexyl		1.39 ± 0.02	5.86	5.79	0.07
30	Cyclohexyl		3.39 ± 0.10	5.47	5.31	0.16
31	Cyclohexyl		4.47 ± 0.28	5.35	5.29	0.06
32	Cyclohexyl		2.57 ± 0.17	5.59	5.52	0.07
33	Cyclohexyl		5.18 ± 0.34	5.29	5.26	0.03
34	Cyclohexyl		6.41 ± 0.12	5.19	5.17	0.02
35	Cyclohexyl		5.51 ± 0.22	5.26	5.14	0.12

Table 1 (continued)

Compd	R_1	R_2	IC_{50} (µM)	pIC_{50}	CoMSIA	
					PA^b	Δ^c
36	(phenyl)	—NH— (phenyl, Cl)	3.94 ± 0.34	5.40	5.31	0.09
37	(phenyl)	—NH— (phenyl, Br)	13.55 ± 0.85	4.87	4.94	−0.07
38	(norbornyl)	—NH— (biphenyl)	0.845 ± 0.05	6.07	6.33	−0.26
39	(norbornyl)	—N(CH_3)— (anthracenylmethyl)	0.46 ± 0.01	6.34	6.46	−0.12
40	Cyclohexyl	—NH— (terphenyl, OH)	0.14 ± 0.01	6.85	6.99	−0.14
41	(cycloheptyl)	—N(CH_3)— (anthracenylmethyl)	0.62 ± 0.05	6.21	5.92	0.29
42	(cycloheptyl)	—NH— (terphenyl, OH)	0.36 ± 0.03	6.44	6.40	0.04
43	(cyclooctyl)	—N(CH_3)— (phenanthrenylmethyl)	0.32 ± 0.02	6.49	6.58	−0.09
44	(cycloheptyl)	—NH— (terphenyl, OH)	1.29 ± 0.10	5.89	5.79	0.10

[a] Test set molecules.
[b] Predicted activity.
[c] Residual of experimental and predicted activities.

ring. These interactions provide a critical role for further development of structurally novel InhA direct inhibitors.

As there are a limited number of successful inhibitors available on the market to MDR-TB, it is an urgent need for developing additional therapeutically useful InhA direct inhibitors. Most of computer-aided efforts directed towards developing new InhA inhibitors were focused on structure-based (docking) techniques [9,10,16] and three-dimensional quantitative structure–activity relationship (3D-QSAR) methodologies [17]. However, despite the excellent potential of current docking methodologies, they lack the ability to predict accurately the potencies of in silico hits due to their inability to evaluate binding free energies correctly. Meanwhile, despite the excellent predictive potential of 3D-QSAR based methodologies, they generally lack the ability to act as effective searches for new hits. Accordingly, we were prompted to develop a structure-based pharmacophore modeling integrated with a predictive 3D-QSAR model. The pharmacophore model can be effectively used as 3D search query to mine 3D compound libraries for new InhA inhibitors, whereas the associates QSAR models can help to predict the selected hits for their binding affinity. In order to incorporate the structural information of the receptor and identify the bioactive conformations of pyrrolidine carboxamide inhibitors of InhA (the crucial step of alignment in 3D-QSAR analysis), a 3D-QSAR analysis was performed based on pharmacophore alignment [18].

Therefore, in the present study we employed the structure-based pharmacophore model built from the crystallographic structure of the InhA–pyrrolidine carboxamide complex (PDB code: 2H7M) by LigandScout 2.0 [19]. Subsequently, 3D-QSAR analysis (comparative molecular field analysis – CoMFA/comparative molecular similarity indices analysis – CoMSIA) was developed using a series of pyrrolidine carboxamide-based InhA inhibitors based on pharmacophore alignment. The reliable pharmacophore model with feature-shape of biologically active conformation was used as 3D search query to screen compounds libraries for retrieving potential hits of inhibitory activity targeting InhA. The hits with good estimated activities by 3D-QSAR model were further refined by molecular docking. In addition, the interactions between the existing pyrazole inhibitors and InhA were also explored based on the pharmacophore model.

2. Materials and methods

2.1. Pharmacophore study

2.1.1. Pharmacophore model generation

The crystal structure of InhA with a pyrrolidine carboxamide inhibitor (PDB code: 2H7M) was used as starting structure for the generation of pharmacophore model in the present study. The software LigandScout 2.0 [20,21] was applied for the detection and interpretation of crucial interaction patterns between InhA and the ligand. LigandScout extracts and interprets the ligand and the macromolecular environment from PDB file, then automatically creates and visualizes an advanced pharmacophore model. The pharmacophore model was exported as a hypoedit script and converted into Discovery Studio 2.1 [22] format with Hypoedit tool. The bound ligand was converted to shape query and was merged with the pharmacophore model to give a combined feature-shape query, which was used for in silico screening.

2.1.2. Conformational model analysis

For the training and test sets molecules, conformational models representing their available conformational space were calculated. All molecules were subjected to Diverse Conformation Generation protocol to produce a maximum of 255 conformations within 20 kcal/mol in energy from the global minimum. All other parameters used were kept at their default settings. The molecules associated with their conformational models were mapped onto the pharmacophore model using "flexible" fitting method and "best mapping only" option to obtain the bioactive conformation of each molecule in Ligand Pharmacophore Mapping protocol in Discovery Studio 2.1.

2.2. 3D-QSAR study

2.2.1. Data sets

A series of pyrrolidine carboxamides reported by He et al. as InhA inhibitors are considered in this study [13]. The biological data were considered comparable and divided into a training set and a testing set as shown in Table 1. The training set selected randomly consists of 37 compounds, and the testing set is comprised of 7 compounds. The IC_{50} values were converted into $pIC_{50}(-\log IC_{50})$ for use in 3D-QSAR analysis. Compounds that did not have activity in the exact number form were ignored for the study.

2.2.2. Alignment rule

Since the major weakness of the traditional 3D-QSAR has been the alignment of the molecules, we applied pharmacophore-based alignment in the present study. The advantage of pharmacophore alignment is that it can assume R conformations of pyrrolidine carboxamides as the biologically active conformations. All the molecules in the training and test sets were mapped simultaneously onto the pharmacophore model using "flexible" fitting method and "best mapping only" option in the Ligand Pharmacophore Mapping protocol in Discovery Studio 2.1. The conformation

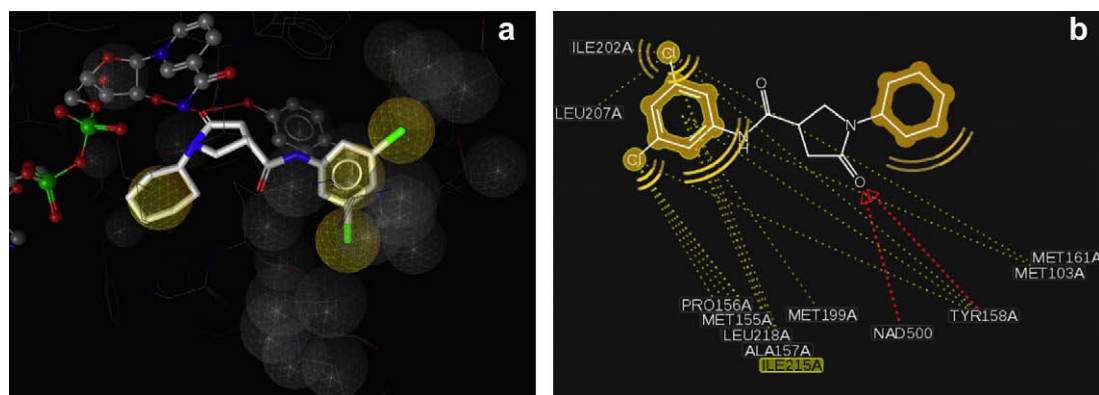**Fig. 1.** (a) LigandScout pharmacophore model generated from InhA–compound **19** complex (red arrows, HBA; green arrow, HBD; yellow spheres, hydrophobic sites; gray spheres, excluded volumes). (b) Schematic 2D molecular interactions' plot of compound **19** with residues of the InhA binding site. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Summary of the RMSD values for five compounds by mapping conformations and the bound conformations in the crystal structures.

Compd.	1	3	13	16	19
PDB code	2H7I	2H7L	2H7N	2H7P	2H7M
RMSD (Å)	1.41	0.79	0.47	1.27	0.40

selected for each compound, which was assumed to be the bioactive conformation, corresponded to the conformation which best fits the pharmacophore model. The final aligned molecules were exported to SYBYL6.9 for CoMFA and CoMSIA.

2.2.3. CoMFA/CoMSIA model

CoMFA was performed using the QSAR option of SYBYL6.9 [23]. The steric and electrostatic field energies were calculated using the Lennard–Jones and the Coulomb potentials, respectively, with a $1/r$ distance-dependent dielectric constant in all intersections of a regularly spaced (0.2 nm) grid. The electrostatic fields were computed using Gasteiger–Huckel charge calculation methods. An sp3 hybridized carbon atom with a radius of 1.53 Å and a charge of +1.0 was used as a probe to calculate the steric and electrostatic energies between the probe and the molecules using the Tripos force field. The standard parameters implemented in SYBYL6.9 were used. The truncation for both the steric and the electrostatic energies was set to 30 kcal/mol.

The CoMSIA method involves a common probe atom and similarity indices calculated at regularly spaced grid intervals for the prealigned molecules. The similarity indices descriptors were derived with the same lattice box used in CoMFA. CoMSIA calculates hydrophobic, H-bond donor and acceptor fields in addition to steric and electrostatic fields. The distance dependence between the grid point and each atom of molecule was determined by Gaussian function through the similarity indices calculated at all grid points, and a default value of 0.3 was used as an attenuation factor.

Partial least squares (PLS) [24] methodology was used for all 3D-QSAR analyses. The cross-validation [25,26] analysis was performed using leave one out (LOO) method in which one compound is removed from the dataset and its activity is predicted using the model derived from the rest of the dataset. The cross-validated r^2_{cv} that resulted in optimum number of components and lowest standard error of prediction was considered for further analysis.

To speed up the analysis and reduce noise, a minimum filter value σ of 2.00 kcal/mol was used. Final analysis was performed to calculate conventional r^2 using the optimum number of components obtained from the cross-validation analysis. The predictive power of the 3D-QSAR models was determined from a set of 7 molecules that were excluded during model development.

2.3. In silico screening

Pharmacophore-based database searching is a kind of ligand-based virtual screening, which can be efficiently used to find novel and potential leads for further development. In this study, the pharmacophore model with feature-shape query was adopted to screen the SPECS database in 3D Database Search protocol [27]. The resulting hits were ranked according to their Best Fit values.

2.4. Molecular docking

Docking procedure aims to generate and score putative protein–ligand complexes according to their calculated binding affinities. Docking studies were carried out using GOLD docking software [28], version 3.1, which uses a powerful genetic algorithm (GA) method for conformation search and docking and being widely regarded as one of the best docking programs [29]. Docking experiments were performed using the default GOLD fitness function (VDW = 4.0, H-bonding = 2.5) and default evolutionary parameters: population size = 100; selection pressure = 1.1; operations = 100,000; islands = 5; niche size = 2; migration = 10; mutation = 95; crossover = 95. The Goldscore function was used to rank different binding pose. The center of the bound ligand was defined as the binding site. Ten docking runs were performed per structure. All poses were output into a single *.sdf file.

To test whether the GOLD program is feasible for the ligand binding to InhA, the InhA–compound **19** complex structure (PDB code: 2H7M) was initially chosen and the docking structure of compound **19** was compared with its crystallographic structure. The resulting docked conformations of compound **19** and that of crystallographic structure were very similar (RMSD: 0.96 Å). This result demonstrates that GOLD analysis is suitable for the identification of the binding mode between the inhibitors and InhA.

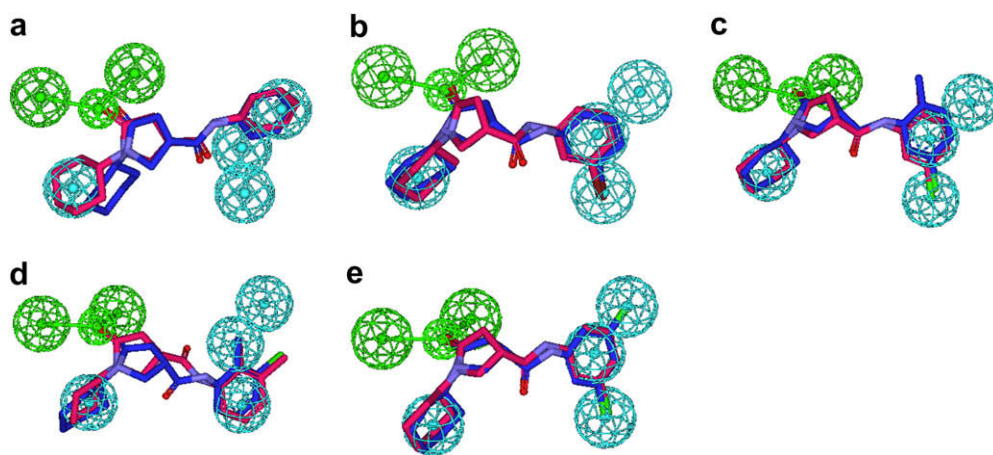**Fig. 2.** The best mapping conformation (red stick) and the bound conformation in the crystal structure (blue stick) are superimposed on the pharmacophore model. (a) For compound **1**. (b) For compound **3**. (c) For compound **13**. (d) For compound **16**. (e) For compound **19**. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
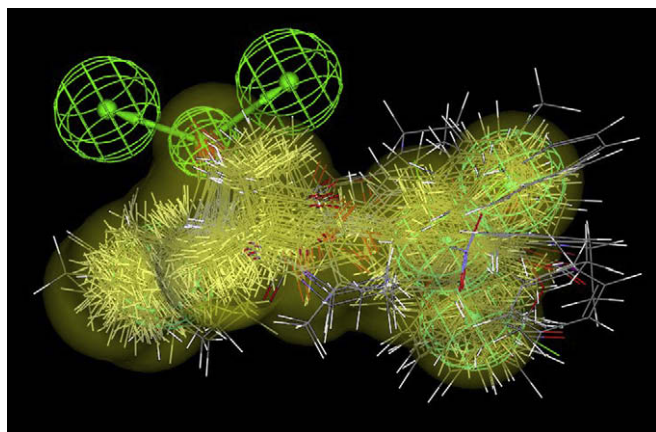
Fig. 3. Molecular alignments used in the present study, obtained from the pharmacophore model alignment.

3. Results and discussion

3.1. Pharmacophore study

3.1.1. Pharmacophore model

As shown in Fig. 1a, the pharmacophore model automatically generated by the LigandScout 2.0 program includes six features: two hydrogen bond acceptors (HBA) and four hydrophobic groups. Besides, the program automatically generated a series of excluded volumes in the model. Both HBA features characterize the carbonyl group of the ligand which forms two hydrogen bonds with Tyr158 and the 2′-hydroxyl group of the nicotinamide ribose of the nucleotide (Fig. 1b). One of the hydrophobic features is located on the cyclohexyl and the other three hydrophobic features are located on the 3- and 5-chloro-substituted phenyl group, respectively. It is clear that the LigandScout program could not interpret the lactam ring as an aromatic ring automatically, which shows π–π interactions between the amide carbonyl group and the NAD^+ nicotinamide ring. So the feature-shape of compound **19** was added manually on the pharmacophore model to obtain appropriate model for in silico screening.

3.1.2. Pharmacophore evaluation

A reliable pharmacophore model may be used to determine the bioactive conformations of the ligands that share the same binding mode. The conformation selected for each compound, assumed as the bioactive conformation, corresponds to the conformation which best fits the pharmacophore. To verify whether the pharmacophore model finds the correct bioactive conformation, we applied the method to five pyrrolidine carboxamide compounds, whose bioactive conformations are known from X-ray structures of the ligand–enzyme complexes (Table 2). The bound conformations of five compounds were respectively mapped onto the pharmacophore model using "flexible" fitting method and "best mapping only" option in the Ligand Pharmacophore Mapping protocol and meanwhile superimposed to the best mapping conformations (Fig. 2). The root mean square distance (RMSD) values between the heavy atom positions of the bound and the best mapping conformation are shown in Table 2. From the results, it is important to note that the best mapping conformations of the ligands are in good agreement with their bioactive conformations. Hence, the results showed that the pharmacophore model is capable of reproducing the bioactive conformation from the Protein Data Bank and is reliable enough to retrieve compounds that fit all the features of the query from chemical databases. These findings also support our choice for the bioactive conformation obtained from the best mapping conformation of the calculated ensemble to the alignment in 3D-QSAR analysis (Fig. 3).

3.2. 3D-QSAR model

3.2.1. CoMFA and CoMSIA results

The stepwise development of CoMFA and CoMSIA models using different fields is presented in Table 3. The predictability of the models is the most important criterion for assessment of both methods. In comparison to CoMFA, CoMSIA methodology has the advantage of exploring more fields. The best CoMSIA model included steric, electrostatic and hydrogen bond donor fields (CoMSIA-SED) and has an r^2_{cv} value of 0.660 using 5 components, with a low standard deviation (SEE = 0.143) and a high Fischer ratio ($F = 158.49$). The predicted values for the 37 compounds in the training set and 7 compounds in the test set using CoMSIA-SED models are shown in Table 1. The correlations between the predicted and experimental values of all compounds are shown in Fig. 4. The PLS statistics of both CoMFA and CoMSIA 3D-QSAR models indicate that CoMSIA-SED produced significantly better results than CoMFA.

As we were interested in using CoMSIA-SED model to predict the activity of hits, it was therefore imperative to judge the predictive ability of the model. The most appropriate method to

Table 3
Summary of 3D-QSAR analysis results obtained using CoMFA and CoMSIA.

Parameter	CoMFA	CoMSIA								
	S,E	S	E	H	D	A	S,E	E,H	S,E,D	ALL
r^2_{cv}	0.466	0.242	0.602	0.348	0.419	0.183	0.641	0.450	0.660	0.602
N	5	2	6	1	3	2	7	2	5	4
r^2	0.981	0.627	0.940	0.579	0.681	0.569	0.975	0.811	0.962	0.947
SEE	0.102	0.432	0.184	0.452	0.405	0.464	0.122	0.307	0.143	0.167
F-value	317.73	28.52	78.46	48.14	23.48	22.41	158.71	72.78	158.49	143.24
Contributions										
Steric	0.553						0.324	0.511	0.192	0.115
Electrostatic	0.447						0.676	0.489	0.419	0.205
Hydrophobic										0.176
Donor									0.389	0.330
Acceptor										0.174
r^2_{pred}									0.646	

r^2_{cv} = cross-validated correlation coefficient; N = No. of components; r^2 = conventional correlation coefficient; SEE = standard error of estimate; F-value = F-test value; S = steric field, E = electrostatic field, H = hydrophobic field, D = hydrogen bond donor field, A = hydrogen bond acceptor field; r^2_{pred} = predicted correlation coefficient for test set of compounds.
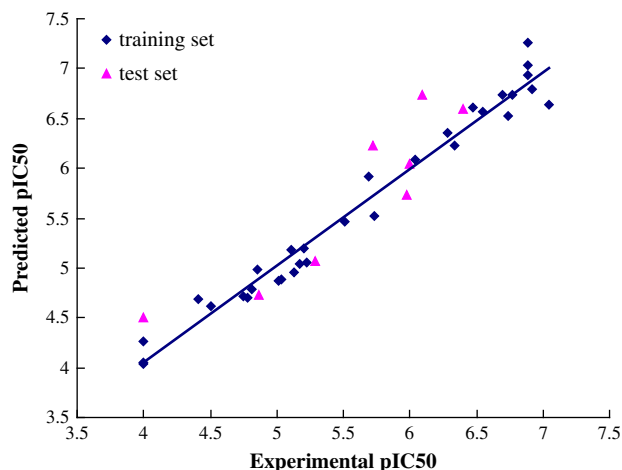
Fig. 4. Graph of experimental values versus predicted values for training and test set compounds.

test the predictive ability is to predict activities of molecules excluded from model building. All compounds within test set, either least active or most active, were predicted close to their experimental activities, thus yielding a good predictive r^2_{pred} of 0.646. The result indicates that the obtained model is robust and predictive.

3.2.2. CoMSIA contour plots analysis

Steric, electrostatic and hydrogen bond donor contour plots obtained using CoMSIA-SED method are a powerful tool to explore the protein–ligand interactions, as presented in Fig. 5. For simplicity, the interaction between only compound **19** and the contour map is shown. Contour plots show the requirements of the basic scaffolds of the pyrrolidine carboxamide for increasing InhA inhibitory activity.

Two large green contours observed near 3- and 5-chloro atoms of compound **19** indicate that bulky substitutions at these positions increase the activity (Fig. 5a). This can be explained by the fact that compounds **38–44** are more potent than molecules with small substituents, such as compounds **1–12**. From the pharmacophore model, the bulky aromatic substituents in this region can interact better with InhA via hydrophobic interactions. A large yellow contour, located at para position of the phenyl ring of compound **19**, suggests that any bulky substitution at this position is likely to decrease the activity. That is why compounds **4**, **6**, **11**, **12**, **17** are less potent as compared to the compounds without substituents at this position like compounds **1**, **3**, **5**, etc. Several small yellow contours near the cyclohexyl ring of compound **19** indicate that steric bulk is

disfavored for the activity. That is why the activity of compound **40** is higher than compounds **41–44**.

The CoMSIA electrostatic contour is shown in Fig. 5b. A big blue contour around the scaffold of pyrrolidine carboxamide suggests that the positive charge in this zone may be necessary for the binding. Two red counters overlapping amide group conform that the negatively charged environment is favorable for the activity.

The CoMSIA hydrogen bond donor contour plot is displayed in Fig. 5c. The cyan contours represent the region favoring hydrogen bond donor substituents. The magenta contours indicate that hydrogen bond acceptor groups in these regions are required for higher activity. It can be explained by the fact that the dual hydrogen bond network involved the oxygen atom on the pyrrolidine carbonyl group, catalytic residue Tyr158, and the 2′-hydroxyl group of the nicotinamide ribose of the nucleotide.

3.3. In silico screening

3.3.1. Hit compounds' analysis

The pharmacophore model with feature-shape query was adopted to screen the SPECS database comprised of about 230,000 molecules. A hit list of 299 compounds ranking by Best Fit values were obtained, which included some compounds structurally similar to the existing InhA inhibitors and novel scaffolds also emerged. Only database hits with a Fit value higher than 2.5 were extracted and put into further analysis.

Hydrogen bonds provide strong interactions between the ligand and the protein catalytic residue as well as its co-substrate. As displayed in Fig. 1b, the carbonyl group of pyrrolidine ring forms the hydrogen bond network with catalytic residue Tyr158 and the 2′-hydroxyl group of the nicotinamide ribose of the nucleotide. Several other reported potent InhA inhibitors have carbonyl group or hydroxyl group which might also form hydrogen bonds with the residue and co-substrate [30,31]. Therefore, more attentions were paid to the two HBA features of the pharmacophore model. In addition, the lactam ring of the ligand shows π–π interactions with the NAD$^+$ nicotinamide ring and does not exhibit in the pharmacophore model. So the hits with two HBA features incorporating an aromatic ring structure were considered firstly. Thus, a small set of 30 hits were ultimately selected after careful observations, analyses and comparisons (Table 4). From the structure information, the hits are mainly divided into two classes, namely, thiazolones and nitrobenzene carboxamides. Besides, some other new structures are also obtained, such as hits ZINC00089683, ZINC00096782, and ZINC00105198. Among the hits, thiazolones are close to the current studied classes of InhA inhibitors, pyrrolidine carboxamides. The results demonstrate that the pharmacophore model was able to retrieve hits with similar feature to existing InhA inhibitors as well as novel scaffolds.
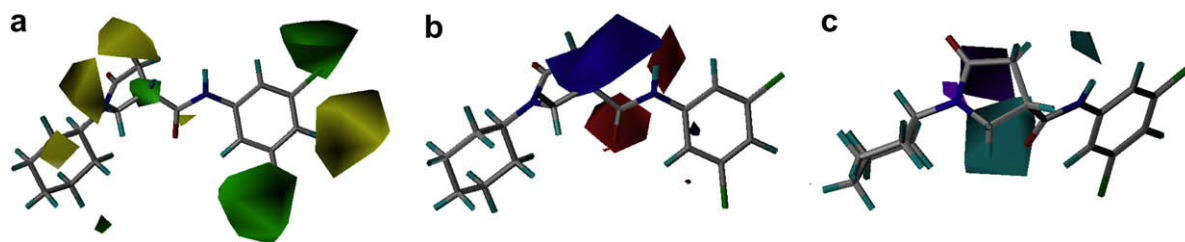

Fig. 5. Contour maps of the pharmacophore-based CoMSIA-SED model. (a) Steric fields: green contours indicate regions where bulky groups increase activity, while yellow contours indicate regions where bulky groups decrease activity. (b) Electrostatic fields: blue contours indicate regions where electropositive groups increase activity, while red contours indicate regions where electronegative groups increase activity. (c) Hydrogen bond donor fields: Cyan contour indicates regions where hydrogen bond donor groups increase activity. Purple is disfavored. Compound **19** is shown inside the fields.

Table 4
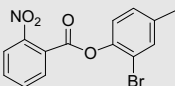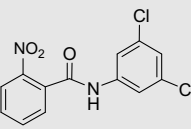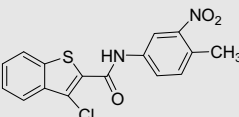Structures of final 30 hits retrieved from the SPECS database by in silico screening.
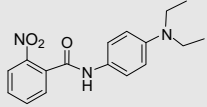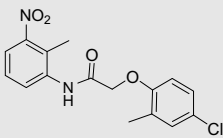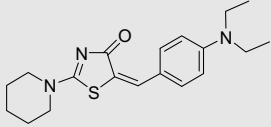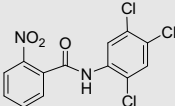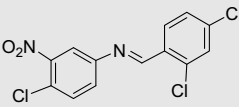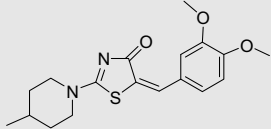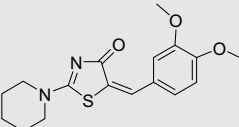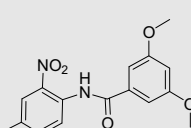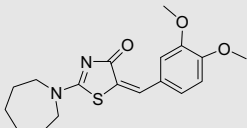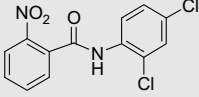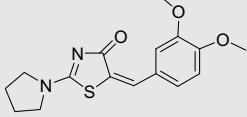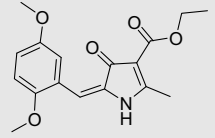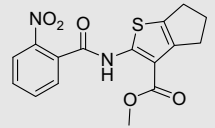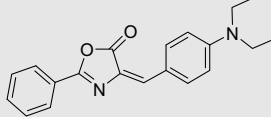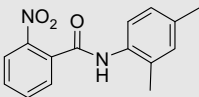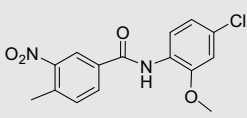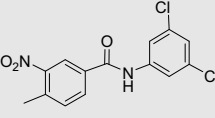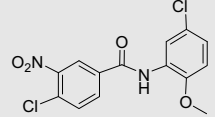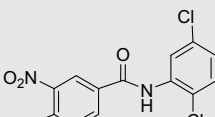
Hits	Structure	pIC$_{50}$[a]	EIC$_{50}$[b] (μM)
ZINC00038499		5.02	9.55
ZINC00038684		5.21	6.17
ZINC00038812		5.67	2.14
ZINC00038831		5.36	4.37
ZINC00038887		5.46	3.47
ZINC00050620		5.21	6.17
ZINC00054641		4.92	12.02
ZINC00010109		5.33	4.68
ZINC00063052		4.97	10.72
ZINC00063055		4.94	11.48
ZINC00063977		5.87	1.35

Table 4 (continued)

Hits	Structure	pIC$_{50}$[a]	EIC$_{50}$[b] (μM)
ZINC00067596		4.96	10.96
ZINC00069333		4.70	19.95
ZINC00083852		5.14	7.24
ZINC00089683		5.44	3.63
ZINC00096438		5.74	1.82
ZINC00096782		5.48	3.31
ZINC00101882		4.98	10.47
ZINC00101991		5.30	5.01
ZINC00102008		5.06	8.71
ZINC00102780		5.39	4.07
ZINC00102805		4.70	19.95

Table 4 (continued)

Hits	Structure	pIC$_{50}$[a]	EIC$_{50}$[b] (μM)
ZINC00103139		4.28	52.48
ZINC00103336		5.30	5.01
ZINC00103353		5.15	7.08
ZINC00103357		4.86	13.80
ZINC00105198		5.29	5.13
ZINC00106692		4.75	17.78
ZINC00115740		4.98	10.47
ZINC00115754		5.00	10.00

[a] Predicted activity by 3D-QSAR model.
[b] Estimated IC$_{50}$.

3.3.2. Pharmacophore mapping

As displayed in Fig. 6, the representative hits were mapped onto the combined feature-shape query (for simplicity, the excluded volume was not shown). In hit ZINC00050620, the carbonyl group of thiazolone serves as the two HBA features of the pharmacophore model. The cyclohexyl group occupies one of the hydrophobic features and the N,N-diethylbenzenamine locates in the other three hydrophobic features. In hit ZINC00038812, the nitro group fits the two HBA features of the pharmacophore model. The methyl group maps one of the hydrophobic features while the chloro-substituted benzothiophene maps the other three hydrophobic features. In hit ZINC0096782 and hit ZINC00105198, the carbonyl group of furanone and sulfonamide group fit the two HBA features, respectively.

3.3.3. Explanation of key interactions between existing pyrazole inhibitors and InhA

The scaffold of the nitrobenzene carboxamides hits is analyzed. As shown in Table 4, the scaffold of nitrobenzene is present in 19 compounds evidently. In fact, this scaffold is already reported in an earlier important InhA inhibitor, pyrazole derivatives [11,12]. The exact nature of the pyrazole derivatives' interaction with the active site of the enzyme has not been fully elucidated. Here, the representative pyrazole derivatives, Genz-7585 and Genz-5542, were mapped onto the pharmacophore model in the Ligand Pharmacophore Mapping protocol. As we can see in Fig. 7, Genz-7585 and Genz-5542 fit five of the six features and miss one of the hydrophobic features. The nitro substitutes of the phenyl groups serve as the two HBA features of the pharmacophore model. The phenyl rings of the pyrazole derivative map one of the hydrophobic features while the others are mapped by the trifluoromethyl-substituted pyrimidine moieties. So the nitrobenzene carboxamide hits and the pyrazole inhibitors may have the similar binding modes with InhA.

3.3.4. Prediction of the activities of hits based on QSAR model

The identified hits were aligned to the training set in QSAR study to predict their bioactivities based on 3D-QSAR model. The pIC$_{50}$ values and EIC$_{50}$ are shown in Table 4. Based on our QSAR methodology, we were able to identify some molecules as potential hits. According to COMSIA-SED model, hit ZINC00063977 is the most active hit while hit ZINC00103139 was the least active. In general, the thiazolone series are less potent than the nitrobenzene carboxamide series. From the 3D-QSAR model, the introduction of larger hydrophobic substituents to the hits might provide more hydrophobic contacts with the active site of InhA and improve the inhibitory activities.

3.4. Molecular docking

Among the hits, ZINC00050620 and ZINC00063977 serve as promising new leads against InhA with the highest estimated IC$_{50}$ value in thiazolone series and nitrobenzene carboxamide series, respectively (Table 4). To visualize their favorable interactions with the active site of InhA, the molecular docking by GOLD program was performed. The best docking poses of two hits are shown in Fig. 8. In the binding mode of the hit ZINC0050620, the carbonyl group of thiazolone forms hydrogen bond network with catalytic residue Tyr158 and the 2′-hydroxyl group of the nicotinamide ribose of the nucleotide (Fig. 8a). The cyclohexyl ring of the hit ZINC0050620 occupies the hydrophobic pocket and interacts with the side chains of Gly96, Phe97, and the nicotinamide ribose. The π–π interactions between the thiazolone ring and the NAD$^+$ nicotinamide ring are also observed. The N,N-diethylbenzenamine group is surrounded by Ala157, Met155, Pro156, Ile215, and Leu218 with hydrophobic contacts. An additional hydrogen bond is observed between thiazolone nitrogen and the 2′-hydroxyl group of the nicotinamide ribose of the nucleotide. The docking hit ZINC00063977 shows that the nitro group that maps the two HBA features in the pharmacophore model forms a hydrogen bond network with catalytic residue Tyr158 and the 2′-hydroxyl group of the nicotinamide ribose of the nucleotide (Fig. 8b). The methyl group has van der Waals interactions with the side chains of Gly96, Phe97, and the nicotinamide ribose. The dimethoxybenzene group of hit ZINC00063977 occupies the large hydrophobic pocket and interacts with the lipophilic side chains of Ala157, Met155, Pro156, Ile215, and Leu218.
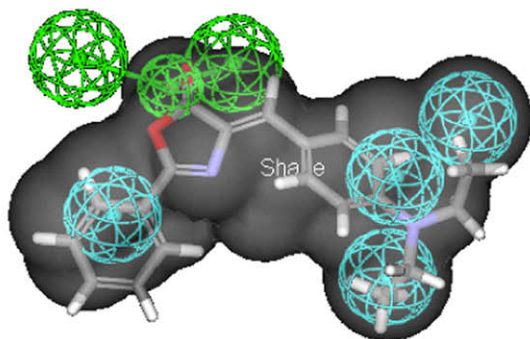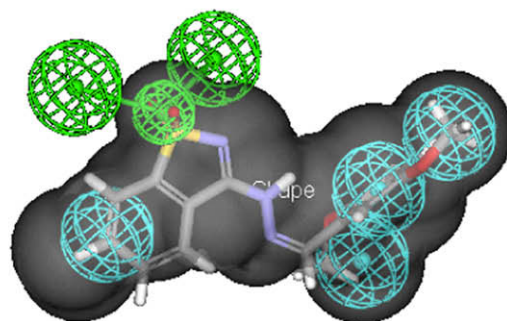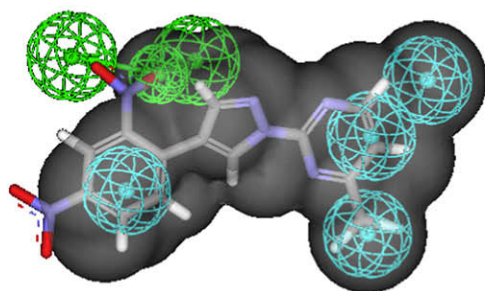
ZINC00050620
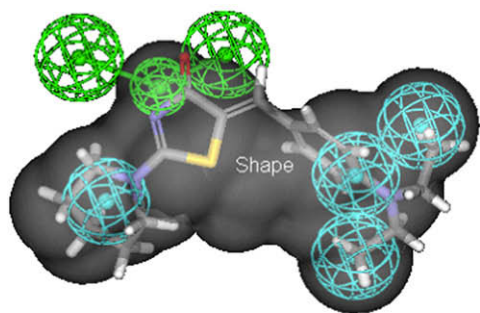
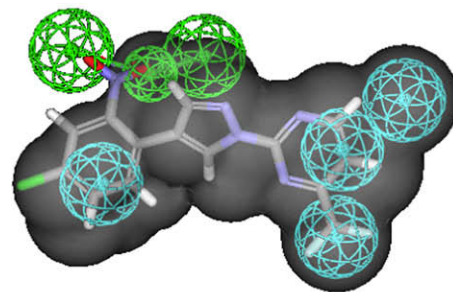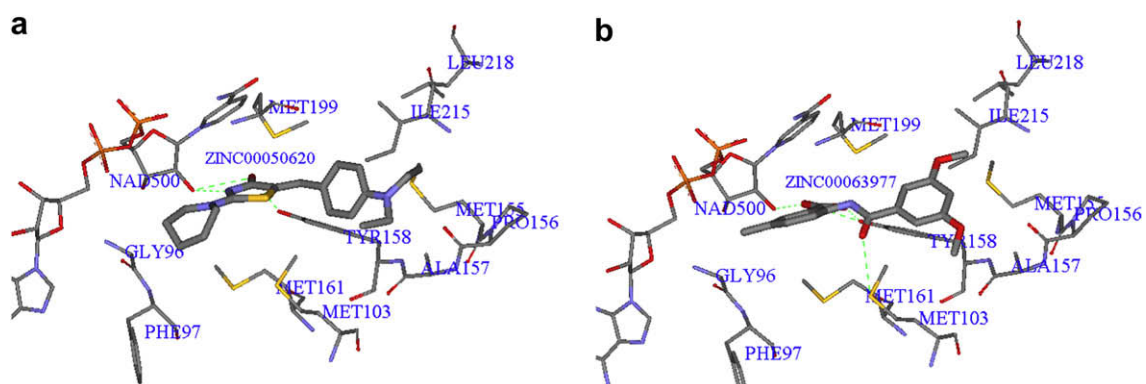ZINC00038812
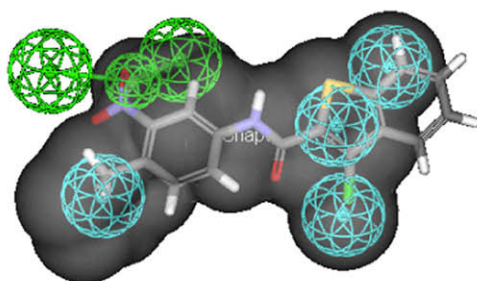
ZINC00096782

ZINC00105198

Fig. 6. The representative hits shown here as mapped to the pharmacophore model.

Genz-7585

Genz-5542

Fig. 7. Pyrazole derivatives (Genz-7585 and Genz-5542) shown here as mapped to the pharmacophore model.

Fig. 8. Best binding poses of hits resulting from docking. (a) For ZINC00050620. (b) For ZINC00063977.

In addition, the amide carbonyl group oxygen is hydrogen bonded to the residue Met103.

4. Conclusions

In our study, an efficient virtual screening model based on the bioactive conformation of a potent active compound bound to InhA has been described. Firstly, structure-based pharmacophore model was built and applied to identify the bioactive conformations and align the pyrrolidine carboxamide compounds to build 3D-QSAR models. The 3D-QSAR model was statistically validated and its predictive ability was evaluated with an external dataset. The steric, electrostatic, and hydrogen bond donor requirements for the inhibitory activity are well described from the contour maps of the studies. Subsequently, the pharmacophore model with feature-shape query was used as 3D search query to screen the SPECS database for discovery of new InhA direct inhibitors. The resulting hits were selected based on the conserved interactions between InhA and inhibitors, mapped onto the pharmacophore model, and evaluated using the 3D-QSAR model. The screening hits were mainly divided into two series, namely, thiazolones and nitrobenzene carboxamides. The two representative hits with highest estimated IC_{50} values were docked into the active site of InhA to analyze the interactions. The identified possible lead compounds will be valuable for further synthesis. In summary, the whole procedure of pharmacophore modeling, 3D-QSAR study, in silico screening, and molecular docking resulted in the identification of some putative new direct inhibitors of InhA for further investigation.

Acknowledgement

References

[1] World Health Organization, Tuberculosis Facts (2008).http://www.who.int/tb.
[2] P. Nunn, B. Williams, K. Floyd, C. Dye, G. Elzinga, M. Raviglione, Nat. Rev. Immunol. 5 (2005) 819–826.
[3] K. Takayama, C. Wang, G.S. Besra, Clin. Microbiol. Rev. 18 (2005) 81–101.
[4] D.A. Rozwarski, C. Vilcheze, M. Sugantino, R. Bittman, J.C. Sacchettini, J. Biol. Chem. 274 (1999) 15582–15589.
[5] A. Banerjee, E. Dubnau, A. Quemard, V. Balasubramanian, K.S. Um, T. Wilson, D. Collins, G. de Lisle, W.R. Jacobs Jr., Science 263 (1994) 227–230.
[6] Y. Zhang, B. Heym, B. Allen, D. Young, S. Cole, Nature 358 (1992) 591–593.
[7] X. Zhao, H. Yu, S. Yu, F. Wang, J.C. Sacchettini, R.S. Magliozzo, Biochemistry 45 (2006) 4131–4140.
[8] S.L. Parikh, G. Xiao, P.J. Tonge, Biochemistry 39 (2000) 7645–7650.
[9] T.J. Sullivan, J.J. Truglio, M.E. Boyne, P. Novichenok, X. Zhang, C.F. Stratton, H.J. Li, T. Kaur, A. Amin, F. Johnson, R.A. Slayden, C. Kisker, P.J. Tonge, ACS Chem. Biol. 1 (2006) 43–53.
[10] C.W. am Ende, S.E. Kundson, N. Liu, J. Childs, T.J. Sullivan, M. Boyne, H. Xu, D.L. Knudson, F. Johnson, C.A. Peloquin, R.A. Slayden, P.J. Tonge, Bioorg. Med. Chem. Lett. 18 (2008) 3029–3033.
[11] M.R. Kuo, H.R. Morbidoni, D. Alland, S.F. Sneddon, B.B. Gourlie, M.M. Staveski, M. Leonard, J.S. Gregory, A.D. Janjigian, C. Yee, J.M. Musser, B. Kreiswirth, H. Iwamoto, R. Perozzo, W.R. Jacobs Jr., J.C. Sacchettini, D.A. Fidock, J. Biol. Chem. 278 (2003) 20851–20859.
[12] M.M. Staveski, S.F. Sneddon, C. Yee, A.D. Janjigian. Inha inhibitors and methods of use thereof, WO01/056974, 2001.
[13] X. He, A. Alian, R. Stroud, P.R. Ortiz de Montellano, J. Med. Chem. 49 (2006) 6308–6323.
[14] X. He, A. Alian, P.R. Ortiz de Montellano, Bioorg. Med. Chem. 15 (2007) 6649–6658.
[15] M.D. Wall, M. Oshin, G.A. Chung, T. Parkhouse, A. Gore, E. Herreros, B. Cox, K. Duncan, B. Evans, M. Everett, A. Mendoza, Bioorg. Med. Chem. Lett. 17 (2007) 2740–2744.
[16] G.S. Rao, R. Vijayakrishnan, M. Kumar, Chem. Biol. Drug. Des. 72 (2008) 444–449.
[17] A. Kumar, M.I. Siddiqi, J. Mol. Model. 14 (2008) 923–935.
[18] Y.D. Chen, H.F. Li, W.Q. Tang, C.C. Zhu, Y.J. Jiang, J.W. Zou, Q.S. Yu, Q.D. You, 3D-QSAR studies of HDACs inhibitors using Pharmacophore based alignment, Eur. J. Med. Chem. (2009). doi:10.1016/j.ejmech.2008.12.008.
[19] The software is from inte:ligand (www.inteligand.com).
[20] G. Wolber, T. Langer, J. Chem. Inf. Comput. Sci. 45 (2005) 160–169.
[21] G. Wolber, A. Dornhofer, T. Langer, J. Comput. Aided Mol. Des. 20 (2006) 773–788.
[22] Accelrys Inc., Discovery Studio 2.1., 2008 http://www.accelrys.com.
[23] Sybyl Molecular Modeling Software Packages, Ver. 6.9, TRIPOS, Associates, Inc., St Louis, 2001.
[24] L. Stahle, S. Wold, J. Chemom. 1 (1987) 185–196.
[25] R.D. Cramer, J.D. Bunce, D.E. Patterson, Quant. Struct.-Act. Relat. 7 (1988) 18–25.
[26] B.L. Podlogar, D.M. Ferguson, Drug. Des. Discov. 17 (2000) 4–12.
[27] http://www.specs.net.
[28] G. Jones, P. Willett, R.C. Glen, A.R. Leach, R. Taylor, J. Mol. Biol. 267 (1997) 727–748.
[29] M. Kontoyianni, L.M. McClellan, G.S. Sokol, J. Med. Chem. 47 (2004) 558–565.
[30] P.J. Tonge1, C. Kisker, R.A. Slayden, Curr. Top. Med. Chem. 7 (2007) 489–498.
[31] H. Lu, P.J. Tong, Acc. Chem. Res. 41 (2008) 11–20.