



Original article

Virtual screening for Raf-1 kinase inhibitors based on pharmacophore model of substituted ureas

Hui-Fang Li^a, Tao Lu^a, Tian Zhu^a, Yong-Jun Jiang^b, Sha-Sha Rao^c, Li-Ye Hu^a, Bo-Tao Xin^a, Ya-Dong Chen^{a,*}^a Department of Organic Chemistry, China Pharmaceutical University, 24 Tongjiaxiang, Nanjing 210009, PR China^b Ningbo Institute of Technology, Zhejiang University, Ningbo 315104, PR China^c Sansom Institute, School of Pharmacy and Medical Sciences, University of South Australia, Adelaide 5000, Australia

ARTICLE INFO

Article history:

Received 11 June 2008

Received in revised form 2 September 2008

Accepted 4 September 2008

Available online 23 September 2008

Keywords:

Raf-1 kinase inhibitors

Pharmacophore

Virtual screening

Molecular docking

ABSTRACT

A three-dimensional (3D) quantitative pharmacophore model was developed from a training set of structurally diverse substituted ureas against Raf-1 kinase, which was well validated to be highly predictive by two methods, namely, test set prediction and Cat-Scramble method. Then a virtual database searching was performed with the pharmacophore model as a 3D query, and the bioactivities of the retrieved hits were predicted by the pharmacophore. Subsequently, molecular docking was carried out on the selected hits whose estimated IC₅₀ was less than 1 μM. Finally, 29 hits were identified as potential leads against Raf-1 kinase, which exhibited good estimated activities, high docking scores, similar binding mode to experimentally proven compounds and favorable drug-like properties. The study may facilitate the discovery and rational design of novel leads with potent inhibitory activity targeting Raf-1 kinase.

© 2008 Elsevier Masson SAS. All rights reserved.

1. Introduction

Ras/Raf/MEK/ERK signal transduction pathway transmits signals from cell-surface receptor tyrosine kinases to the nucleus, aberrant regulation of which leads to cancer and other human diseases [1–5]. Raf serine/threonine kinase is an essential component of the pathway, which is also the first identified and most characterized downstream effector of the central signal transduction mediator Ras [6–8]. There are three members in the Raf family, Raf-1 (also termed as C-Raf), A-raf, and B-raf [4,9–11]. Raf-1 was first reported to be the only isoform of the Raf family that is expressed ubiquitously, and most of the significant contributions describing the role of Raf in signal transduction actually dealt with Raf-1 [12]. Raf-1 is recruited to the plasma membrane by the Ras GTPase, where it undergoes phosphorylation and becomes activated. Once activated, Raf-1 phosphorylates and activates MEK, which in turn phosphorylates ERK. The activated ERK phosphorylates several transcription factors regulating genes associated with cell proliferation, cell survival and angiogenesis [5,11,13].

Raf-1 kinase is a validated target for the treatment of cancer, which has attracted intense research interest [14]. During the past decade, a variety of compounds have been reported as Raf-1 kinase

inhibitors, which can be divided into several structural classes such as ureas, urea bioisosteres, imidazoles, benzamides, oxindoles, and aza-stilbenes [5,15–20]. Among the different classes of Raf-1 kinase inhibitors, ureas have been investigated to the greatest extent [5,15,16,20]. In particular, ureas have been the focus of the research collaboration between Onyx and Bayer. With the considerable number and variety of ureas identified as inhibitors of Raf-1 kinase, sorafenib (tosylate salt of Bay 43-9006), initially and mainly targeting Raf-1, has finally emerged as a successful representative of ureas, approved by FDA for the treatment of advanced renal cell carcinoma in December 2005 and unresectable hepatocellular carcinoma in December 2007. The success of sorafenib confirms substituted ureas which may contain some essential functionalities for Raf-1 inhibitory activity, which are promising Raf-1 inhibitors worthy of further investigation.

In 2004, the crystal structure of B-Raf in complex with Bay 43-9006 was resolved and revealed a distinctive protein–ligand interaction mode [13]. The distal 4-pyridyl ring of Bay 43-9006 occupies the ATP adenine binding pocket of the kinase domain. The lipophilic trifluoromethyl phenyl ring at the opposite end of the molecule inserts into a hydrophobic pocket. The urea functionality forms two crucial hydrogen bonds, one with the backbone aspartate, and the other with the glutamate side chain [13,21]. As the residues of B-Raf that contact with Bay 43-9006 are conserved in Raf-1, consistent with much of the structure–activity relationship of Raf-1 inhibition by Bay 43-9006 and its derivatives [5,13,20], the

* Corresponding author. Tel.: +86 25 86185180; fax: +86 25 86185179.

E-mail address: nonorgan@cpu.edu.cn (Y.-D. Chen).

interaction mode can provide significant information for the conduction and validation of a ligand-based study and structure-based study, such as pharmacophore and molecular docking.

As the pharmacophore model has been successfully applied to undertake 3D searches on large database [22–26], it is possible to retrieve compounds with similar functional groups to Bay 43-9006 from a database of small molecules based on a pharmacophore model, which may have potential inhibitory activity against Raf-1 kinase. Given the large set of substituted ureas published [15,16,20], it is viable to explore the key chemical features responsible for the bioactivities of substituted ureas by developing a quantitative pharmacophore model (also termed as hypothesis in Catalyst) in the HypoGen module of Catalyst 4.10 software package. And it may be more efficient to perform a virtual database screening based on a pharmacophore model generated from diverse substituted ureas rather than various classes of Raf-1 kinase inhibitors as previously reported [27].

The pharmacophore model, which combines and visualizes critical features responsible for Raf-1 inhibitory activity, can not only act as a powerful search tool for a virtual database screening but also serves the purpose of assessing other compounds for their ability to map the essential features and predicting their activities. Additionally, molecular docking was used as a complementary tool to prioritize the hits from the pharmacophore-based virtual screening. The concurrent use of pharmacophore-based searching and molecular docking is anticipated to make hit discovery more efficiently in the hit selection process [25].

The main aim of the study is to construct a reliable pharmacophore model using available substituted ureas currently known as Raf-1 kinase inhibitors, and to perform a database searching with the resultant pharmacophore to retrieve potential hits of inhibitory activity targeting Raf-1 kinase. The hits with good estimated activity were further refined by molecular docking using the program of GOLD (Genetic Optimisation for Ligand Docking, version 3.0.1) to obtain potent Raf-1 kinase inhibitors.

2. Materials and methods

2.1. Pharmacophore study

2.1.1. Dataset

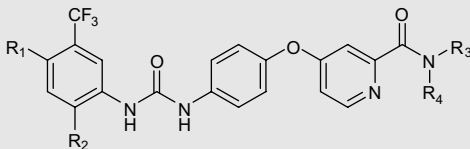
The compounds in this study were collected from a series of substituted ureas as Raf-1 kinase inhibitors published in recent years [16,20], considering both structural diversity and wide coverage of the activities. As all the biological data was obtained from the publications of Bayer Research Center, the assay method used should be the same and the IC_{50} values need no normalization. The dataset was divided into a training set and a test set (Table 1). The training set consisted of 22 structurally diverse compounds, the bioactivities of which ranked from 6 to 17 000 nM spanning over four orders of magnitude, including the most active, several moderately active, and some poorly active compounds. To validate of the pharmacophore model, a test set of another 28 substituted ureas were selected in a similar way to the training set to guarantee maximal 3D diversity and continuous bioactivity magnitude.

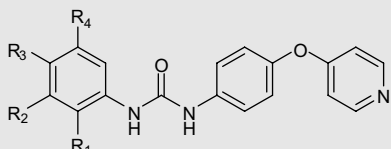
2.1.2. Conformational analysis

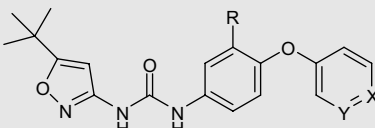
Molecular flexibility was explored by considering each compound as a collection of conformers representing different areas of the conformational space accessible to the molecule within a given energy range. The conformational space of each compound was extensively sampled utilizing the poling algorithm within the ConForm module of Catalyst software package. Conformers were generated by “Best quality” conformational search option based on the CHARMM force field, with an energy threshold of 20 kcal/mol from the lowest energy level and a maximum of 250 conformers for

Table 1

2D structures and experimental IC_{50} values of training and test set compounds

					
No.	R ₁	R ₂	R ₃	R ₄	IC_{50} (nM)
1	Br	H	H	CH ₃	6
2	Cl	H	H	CH ₃	12
23	Cl	H	H	CH ₂ CH ₃	26
3	Cl	H	H	3-Pyridyl	44
24	H	–OCH ₃	H	CH ₃	53
25	Cl	H	H	4-Morpholinyl-(CH ₂) ₂	73
4	Cl	OCH ₃	H	CH ₂ CH ₃	88
27	Cl	H	H	(4-Morpholinyl)phenyl	160
28	Cl	H	H	4-(Me ₂ N)-phenyl	230

					
No.	R ₁	R ₂	R ₃	R ₄	IC_{50} (nM)
29	H	H	H	–OCH ₃	440
10	H	H	H	<i>tert</i> -Bu	870
13	OCH ₃	Cl	H	H	2700
14	OCH ₃	H	NO ₂	H	4400

				
No.	X	Y	R	IC_{50} (nM)
5	CH	C(CONHMe)	H	120
7	C(OCH ₃)	CH	H	500
31	CH	C(OCH ₃)	H	700
32	C(COCH ₃)	CH	H	820
11	C(CH ₃)	CH	H	1000
12	C(CONHMe)	CH	H	1100
14	CH	C(CO ₂ H)	H	4400
43	C(SO ₂ NHMe)	CH	H	7500
18	C(NHAc)	CH	CF ₃	10 000

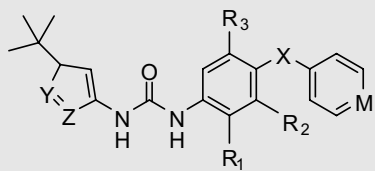
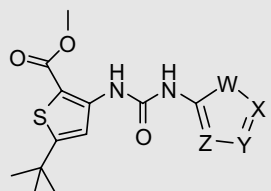
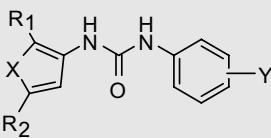
							
No.	Y	Z	X	M	R ₁	R ₂	IC_{50} (nM)
26	N	O	O	N	H	H	120
6	O	N	O	N	H	H	230
8	O	N	O	H	CH ₃	CH ₃	650
9	N	NCH ₃	O	N	H	H	800
33	O	N	O	CH	H	H	1100
17	O	N	O	CH	H	CF ₃	9500
34	O	N	O	CH	CH ₃	H	1400
30	O	N	CH ₂	N	H	H	590
35	O	N	CHOH	N	H	H	1500

Table 1 (continued)

					
No.	W	X	Y	Z	IC ₅₀ (nM)
37	S	C-Et	N	N	2000
42	N	N-CH ₃	CH	CH	6700
15	CH	N-Et	CH	CH	7600
16	N	N-Et	CH	CH	8000
44	CH	N-iPr	CH	CH	8100
45	S	C-tert-Bu	N	N	11 000
48	N	N-iPr	CH	CH	16 000

					
No.	X	Y	R ₁	R ₂	IC ₅₀ (nM)
36	S	4-CH ₃	COOCH ₃	tert-Bu	1700
38	NH	4-CH ₃	COOCH ₃	tert-Bu	3000
39	S	4-CH ₃	COOCH ₃	iPr	4000
40	NHCH ₃	4-CH ₃	COOCH ₃	tert-Bu	5000
41	S	4-CH ₃	COOCH ₃	CH ₂ Br	6000
19	NH	3,4-Cl ₂	COOCH ₃	tert-Bu	12 000
46	S	4-CH ₃	CH ₂ NH ₂	tert-Bu	13 000
47	S	4-CH ₃	CON(CH ₃) ₂	tert-Bu	15 000
20	S	H	COOCH ₃	tert-Bu	17 000
22	S	3-Cl	COOCH ₃	tert-Bu	22 000
49	S	4-OH	COOCH ₃	tert-Bu	18 000
21	S	3-CH ₃	COOCH ₃	tert-Bu	20 000
50	S	3,4-Cl ₂	COOCH ₃	tert-Bu	20 000

each compound. Default settings were kept for the other parameters.

2.1.3. Generation of pharmacophore models

Based on the previous studies on the chemical features of substituted ureas and an initial analysis of the training set with the feature dictionary in Catalyst [5,16,20,28], three chemical features were selected for the pharmacophore generation, namely hydrogen-bond acceptor (HBA), hydrogen-bond donor (HBD), and hydrophobic (HY). The uncertainty was set as 2.0 instead of the default value 3.0, as was reported in other literatures [22,29]. Top 10 pharmacophore models were generated using HypoGen module implemented in Catalyst. The relationship between the structures of the training set compounds and their experimentally determined inhibitory activities against Raf-1 kinase was investigated for the generation of pharmacophore models.

The quality of HypoGen pharmacophore models is best described in terms of fixed cost, null cost, total cost and other statistical parameters [30]. Fixed cost represents a simple model that fits all data perfectly, while null cost presumes that there is no relationship in the data and that the experimental activities are normally distributed around their average value. And total cost sums over error cost, weight cost and configuration cost. For an expected pharmacophore model, the total cost should be close to the fixed cost, and there should be a significant difference between null and total cost. Further, a value of 40–60 bits for the unit of cost difference implies a 75–90% probability of the correlation between experimental and predicted activities.

2.1.4. Pharmacophore model validation

The main purpose of validating a quantitative pharmacophore model is to determine whether it is capable of identifying active

compounds, predicting their activities accurately and further performing a database searching [31]. Two validation procedures were followed, namely, test set prediction method and Cat-Scramble method. A test set of 28 diverse compounds with Raf-1 kinase inhibitory activity was selected to validate the best pharmacophore model. The test set covers the similar structural diversity to the training set in order to ascertain the broadness of the pharmacophore predictability. The Cat-Scramble validation procedure is a cross-validation based on Fischer's randomization test. The goal of this type of validation is to check whether there is a strong correlation between the chemical structures and the biological activity [22,31]. This is done by randomizing the activity data associated with the training set compounds, generating pharmacophore hypotheses using the same features and parameters to develop the original pharmacophore model.

2.2. Virtual screening

Virtual screening has been investigated as an *in silico* tool for drug discovery, which has been widely used for lead identification in drug discovery programs. Virtual screening methods are generally divided into ligand-based virtual screening (LBVS) and structure-based virtual screening (SBVS). LBVS techniques outperform SBVS in measuring compound similarity to known potent molecules, while SBVS techniques have a better potential to identify compounds with a novel core scaffold. So the combined use of LBVS both and SBVS should increase the chances of finding novel hits [32].

2.3. Database searching

Pharmacophore-based database searching is a kind of LBVS, which can be efficiently used to find novel, potential leads for further development from a virtual database. A well-validated pharmacophore model includes the chemical functionalities responsible for the bioactivities, therefore, it can perform a database searching as a 3D query. Fast or Best flexible search databases method can be applied for database searching to retrieve putative compounds, which are defined as compounds having their chemical moieties spatially map with corresponding features in the pharmacophoric query.

2.4. Molecular docking

Molecular docking is a computationally intensive SBVS technique that generates and scores putative protein–ligand complexes according to their calculated binding affinities. Given the crystal structure of the target, molecular docking automatically samples ligand conformations and protein–ligand interactions with a specified region of the protein surface. It has been successfully used for identifying active compounds by filtering out those that do not fit into the binding sites [33–35]. In absence of the structural information of the target, a homology model can be constructed and used for molecular docking. In this study, molecular docking was performed with a homology model of Raf-1 kinase by the program GOLD.

3. Results and discussion

3.1. Pharmacophore study

3.1.1. Generation of pharmacophore models

Ten pharmacophore models were generated, the cost values, correlation coefficients, RMSD, and pharmacophore features were listed in Table 2. OutHypo-8599.1 (Hypo1 for simplicity), consisting of spatial arrangement of five chemical features (Fig. 1), including

Table 2

Ten pharmacophore models generated in HypoGen module

Hypothesis	Total cost	Cost difference ^a	RMSD	Correlation	Features ^b
1	85.0133	107.1097	0.580	0.985	HBA HBD HY HY HY
2	87.0235	105.0995	0.721	0.977	HBA HBD HY HY HY
3	90.3371	101.7859	0.904	0.964	HBA HBD HY HY HY
4	90.9304	101.1926	0.936	0.962	HBA HBD HY HY HY
5	91.0668	101.0562	0.943	0.961	HBA HBA HY HY HY
6	93.0812	99.0418	1.032	0.953	HBA HY HY HY HY
7	93.0947	99.0283	1.027	0.954	HBA HY HY HY HY
8	93.3147	98.8083	0.994	0.957	HBA HBD HY HY HY
9	93.4722	98.6508	1.044	0.952	HA HA HY HY
10	93.4861	98.6369	1.052	0.951	HA HA HY HY HY

^a The difference between null and total cost; null cost is 192.123 bits; fixed cost is 81.2951 bits; configuration cost is 16.3032 bits.

^b Abbreviations used for features: HBA, H-bond acceptor; HBD, H-bond donor; HY, hydrophobic.

one hydrogen-bond acceptor (HBA), one hydrogen-bond donor (HBD), and three hydrophobic (HY) features, was identified as the best model. It had the lowest total cost (85.0133), the least difference between total and fixed cost (3.7182), the highest cost difference between null cost and total cost (107.1097), the least RMSD (0.57951), and a strong correlation coefficient (0.985498) between experimental and estimated activities. And the configuration cost of Hypo1 is 16.3032 bits, which did not exceed the recommended value of 17 and could guarantee the entire conformation space sampled during the pharmacophore generation.

One method of judging a pharmacophore's merit is its ability to predict the activities of individual compounds within the training set [22]. As is shown in Table 3, almost all of the compounds in the training set were properly predicted for their bioactivities with quite low error values between the experimental and estimated IC₅₀ values. Furthermore, the estimated IC₅₀ value of Bay 43-9006 (compound 2 of the training set) exactly equaled to the experimental IC₅₀ value of 12 nM, which showed a perfect correlation.

3.1.2. Pharmacophore model validation

A test set of 28 compounds with Raf-1 kinase inhibitory activity were used to validate the best pharmacophore model Hypo1, yielding a correlation coefficient of 0.901289 (Fig. 2), which suggested a good correlation between the experimental and estimated activities. As is shown in Table 4, we can see that most IC₅₀ values of the compounds in the test set were estimated properly, with the error values mainly around 3.

For the Cat-Scramble validation of Hypo1, a confidence level of 95% was selected, and thus 19 spreadsheets (Table 5) were generated. The data from this validation clearly indicated that all values generated after randomization produced models with no predictive values similar or near to that of Hypo1. Out of the 19 runs, the correlation coefficients were all relatively low, and the RMSD values were high and the difference between total costs and the null cost were also high, which were not desirable for a good model. Thus these two validation procedures provided strong confidence on the best pharmacophore model, Hypo1.

3.1.3. Evaluation of the pharmacophore with the crystal complex

As the residues of B-Raf that contact with Bay 43-9006 were conserved in Raf-1 kinase [5], the crystal structure of B-Raf in complex with Bay 43-9006 was utilized to further evaluate the best pharmacophore Hypo1. When mapped onto Hypo1 by choosing "Best fit" option, the bound conformation of Bay 43-9006 from the crystal complex fit quite well with all the features of the pharmacophore. The mapping of Bay 43-9006 with Hypo1 was shown in Fig. 3, from which we can see that the hydrogen-bond donor (HBD) was mapped onto the -NH of the urea group and the hydrogen-bond acceptor (HBA) mapped onto N-atom of pyridyl ring. Concerning the three hydrophobic (HY) features, one was mapped onto trifluoromethyl group, another one onto the middle phenyl and the third onto the amide NH group.

When the mapping of Hypo1 with the bound conformation of Bay 43-9006 was superimposed to the active site of the crystal structure of B-Raf, it showed that the pharmacophore model accommodated quite well into the cavity of B-Raf, where the important hydrogen bonds interact with key residues at the active site. As is shown in Fig. 4, the urea group formed a hydrogen bond with Asp593, and the nitrogen of pyridyl ring and the amide with Cys531 formed two hydrogen bonds. These analyses revealed that Hypo1 included all the essential features the crystallographic studies covered [13], which were consistent with the interaction mechanism of B-Raf and Bay 43-9006 [28]. And the results also suggested Hypo1 was reliable enough to retrieve compounds that fit all the features of the query from a chemical database.

3.2. Virtual screening

In this paper, a pharmacophore-based virtual screening was conducted to find potential Raf-1 kinase inhibitors. A stepwise

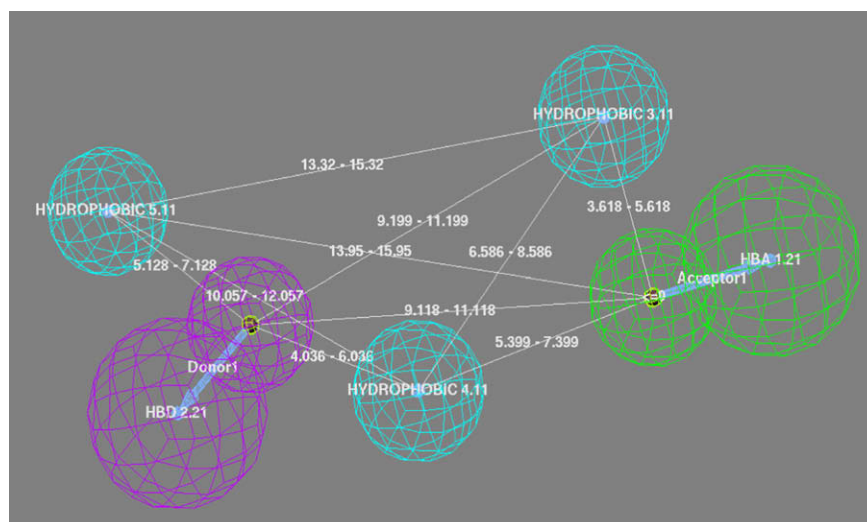


Fig. 1. Features of the best pharmacophore (Hypo1) are color-coded as follows: hydrogen-bond acceptor, green; hydrophobic, blue; hydrogen-bond donor, violet. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Experimental IC₅₀ vs. estimated IC₅₀ values of training set based on Hypo1

Compounds	Experimental IC ₅₀ (nM)	Estimated IC ₅₀ (nM)	Error	Reference
1	6	7.4	1.2	[20]
2	12	12	1	[20]
3	44	34	−1.3	[16]
4	88	60	−1.5	[20]
5	120	160	1.3	[20]
6	230	270	1.2	[20]
7	500	670	1.3	[20]
8	650	900	1.4	[20]
9	800	370	−2.2	[20]
10	870	1300	1.5	[20]
11	1000	660	−1.5	[20]
12	1100	2700	2.5	[20]
13	2700	3000	1.1	[20]
14	4400	7900	1.8	[20]
15	7600	9200	1.2	[20]
16	8000	7900	−1	[20]
17	9500	9900	1	[20]
18	10 000	9700	−1	[20]
19	12 000	10 000	−1.1	[20]
20	17 000	12 000	−1.5	[20]
21	20 000	12 000	−1.7	[20]
22	22 000	12 000	−1.8	[20]

virtual screening procedure was applied, wherein the pharmacophore-based virtual screening was followed by estimated activity prefiltration, molecular docking studies and drug-likeness evaluation. The sequential virtual screening performed in this study was schematically represented in a flowchart in Fig. 5, from which we can see the number of hits reduced for each screening step.

3.2.1. Database searching

The validated pharmacophore model, Hypo1, was used as a query to search the NCI database comprised of 238 819 molecules by the option of “Best flexible search databases/spreadsheet” in Catalyst. A hit list of 4958 compounds matching the pharmacophore model was obtained, which included some compounds structurally similar to the existing Raf-1 kinase inhibitors and novel scaffolds also emerged. Then the 3D quantitative pharmacophore

Table 4
Experimental IC₅₀ vs. estimated IC₅₀ values of test set based on Hypo1

Compounds	Experimental IC ₅₀ (nM)	Estimated IC ₅₀ (nM)	Error	Reference
23	26	55	2.1	[20]
24	53	110	−2	[16]
25	73	180	2.5	[16]
26	120	200	1.7	[20]
27	160	85	−1.9	[16]
28	230	120	−2	[16]
29	440	480	1.1	[20]
30	590	450	−1.3	[20]
31	700	190	−3.7	[20]
32	820	2900	3.5	[20]
33	1100	5800	5.3	[20]
34	1400	620	−2.3	[20]
35	1500	1900	1.3	[20]
36	1700	5500	3.2	[20]
37	2000	5600	2.8	[20]
38	3000	6300	2.1	[20]
39	4000	6100	1.5	[20]
40	5000	7400	1.5	[20]
41	6000	5700	−1.1	[20]
42	6700	8200	1.2	[20]
43	7500	2800	−2.7	[20]
44	8100	5700	−1.4	[20]
45	11000	5300	−2.1	[20]
46	13000	5400	−2.4	[20]
47	15000	7800	−1.9	[20]
48	16000	6500	−2.5	[20]
49	18000	6900	−2.6	[20]
50	20000	7900	−2.5	[20]

Hypo1 was used to estimate the bioactivities of the initial hits, and it was found that a total of 328 hits had estimated IC₅₀ values below 0.5 μ M, and 360 hits with estimated IC₅₀ values between 0.5 and 1 μ M. To sample a sufficient chemical space and increase hit-rate, hits with estimated IC₅₀ below 1 μ M were considered as active new hits. A set of 688 hits satisfied the specified cutoff value and hence proceeded for further evaluation by molecular docking.

3.2.2. Molecular docking

As the crystal structure of Raf-1 kinase is not available, a homology model of Raf-1 kinase was constructed on Swiss-Model, using the crystal structure of the kinase domain of human

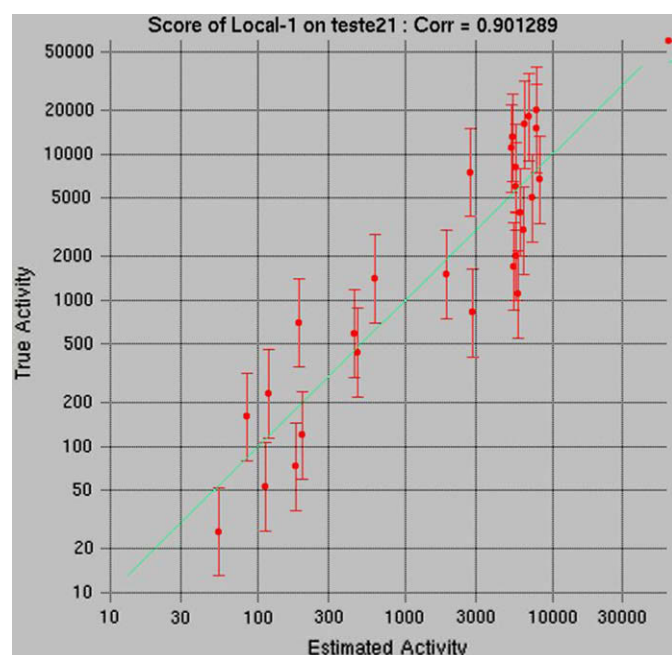


Fig. 2. The bound conformation of Bay 43-9006 mapped to Hypo1.

Table 5
Results from cross-validation using Cat-Scramble in catalyst^a

	Total cost	Fixed cost	RMSD	Correlation
Results for unscrambled	85.0133	81.2951	0.580	0.985
Results for scrambled				
1	147.549	82.3253	2.42555	0.704197
2	114.826	78.9151	1.80681	0.8485535
3	135.877	78.1064	2.23178	0.759368
4	125.965	81.4723	1.88413	0.838276
5	130.891	78.71	1.99138	0.821811
6	151.128	78.6179	2.53078	0.673603
7	138.685	79.9648	2.09485	0.804412
8	132.873	81.6896	2.12669	0.783227
9	142.75	77.8552	2.4205	0.705496
10	148.514	78.8956	2.47882	0.689814
11	168.441	80.1312	2.81348	0.568714
12	147.994	77.7839	2.50142	0.681994
13	138.626	80.6867	2.23303	0.759182
14	129.913	78.7357	2.07409	0.797458
15	129.504	81.6896	2.06563	0.79672
16	148.672	78.7861	2.43158	0.708675
17	137.497	80.2101	2.24351	0.755279
18	123.989	79.7813	1.9222	0.828913
19	145.119	78.1742	2.4078	0.712494

^a Null cost is 192.123.

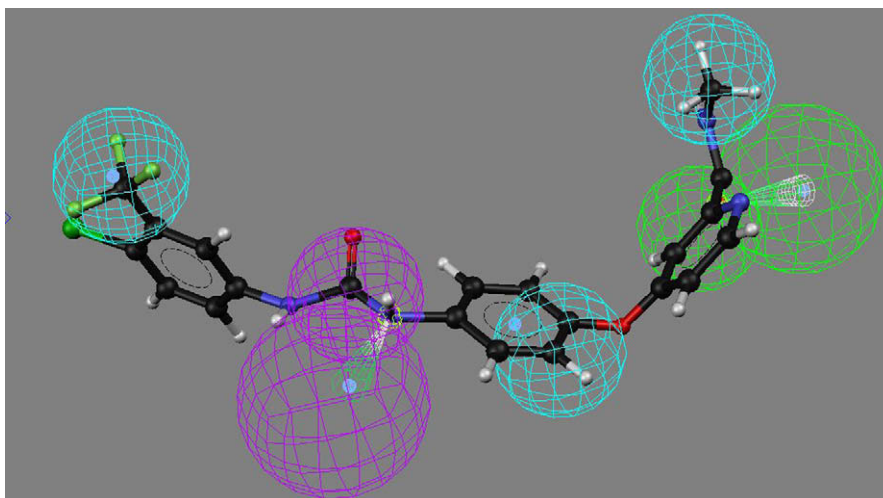


Fig. 3. Hypo1 mapped with Bay 43-9006 at the active site of B-Raf crystal structure.

B-Raf isoform (Protein Data Bank, PDB code: 1UWH) as the template protein [13]. The amino acid sequence identity between Raf-1 and B-Raf kinase domains is about 81% [19,28]. The homology model of Raf-1 kinase includes amino acid residues 342–615. The 3D structure modeling and refinements were implemented on Swiss-Model. The derived homology model of Raf-1 was finally minimized with MMFF94 force fields in Sybyl 6.9.

Molecular docking was carried out with Raf-1 kinase homology model by the program of GOLD. This program from Cambridge Crystallographic Data Center UK uses genetic algorithm for docking flexible ligands into protein binding sites to explore the full range of ligand conformational flexibility with partial flexibility of the protein [31]. The results of the docking were demonstrated mainly by GOLD fitness score. The X-ray coordinates of Bay 43-9006 bound to the active site of B-Raf were used to define active site region with an active site radius of 10 Å. The annealing parameters of van der Waals and hydrogen-bond interactions were considered within 4.0 and 2.5 Å, respectively, and other parameters were kept at the default setting. A superimposition of docked Bay 43-9006 onto the crystallographic geometry yielded RMSD of 0.55 Å, which revealed that GOLD program performed well in reproducing experimentally observed binding conformation of Bay 43-9006, and therefore, GOLD is suitable for a molecular docking study.

Then molecular docking by GOLD was performed on those 688 hits incorporated with Bay 43-9006 as a comparison to prioritize the hits, on the basis of their ability to form favorable interactions within the active site of Raf-1 kinase. As the number of hits to be docked was a bit large, the parameters of library screening were

applied followed by the standard default setting to perform a high performance computing. GOLDScore was selected as the fitness function, and docking results were reported as the highest scoring pose for each compound.

The GOLD fitness scores were presented in descending order to make the results more evident. Bay 43-9006 had a score of

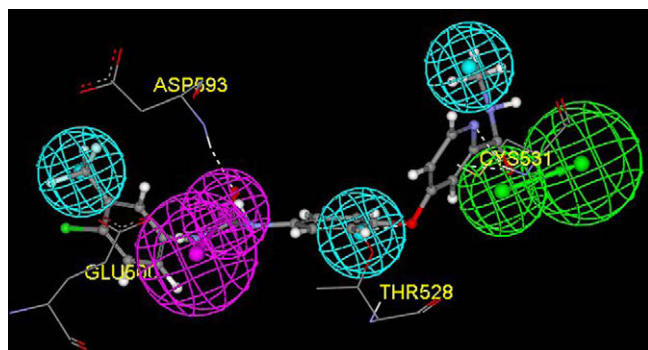


Fig. 4. Correlation between experimental and predicted activities for the 28 test set molecules against Hypo1.

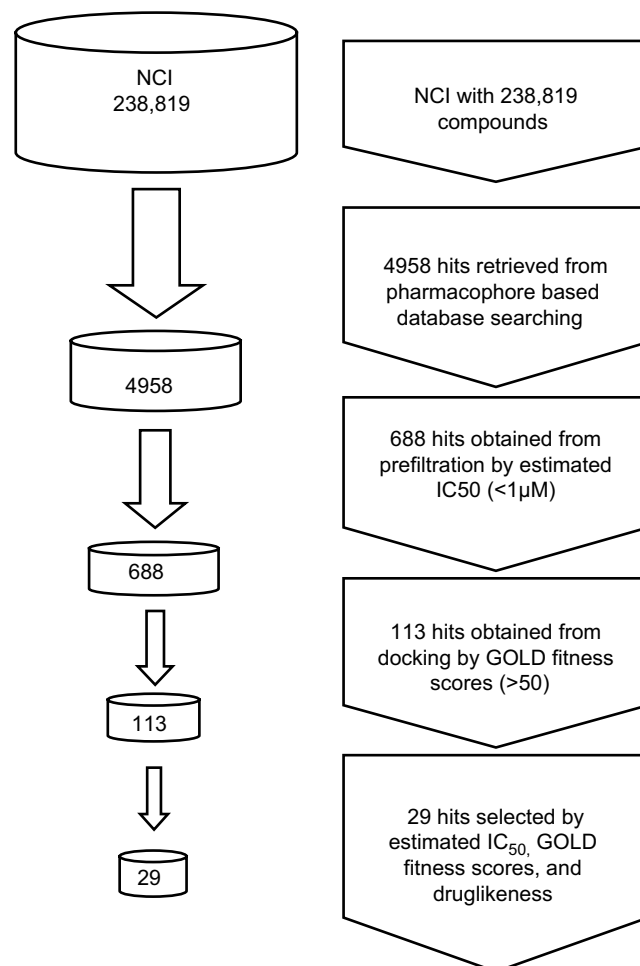


Fig. 5. Flowchart of virtual screening procedure applied in this paper.

Table 6
Examples of final hits retrieved from the NCI database after the whole virtual screening procedure

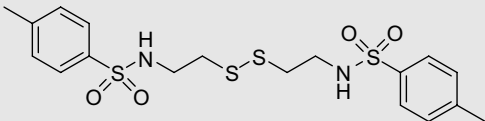
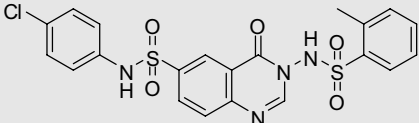
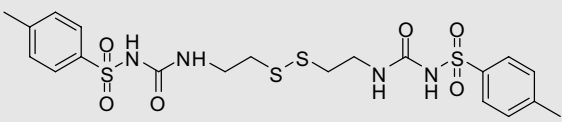
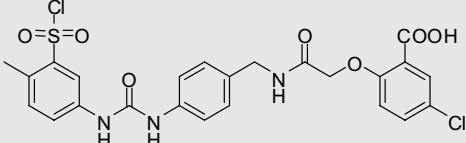
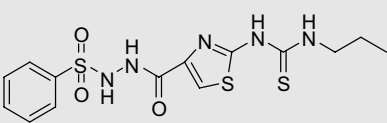
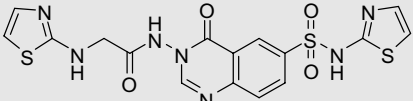
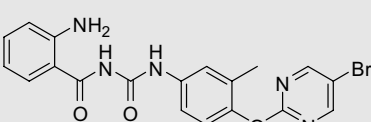
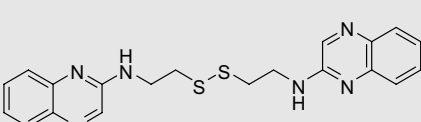
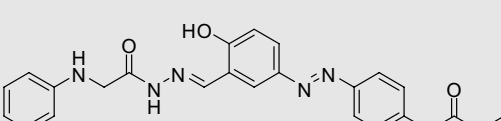
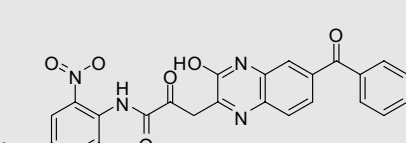
Hits	Structure	EC ₅₀ ^a (μM)	GFS ^b
NCI0342021		0.029	53.79
NCI0687776		0.088	52.19
NCI0120649		0.11	59.95
NCI0212040		0.26	52.18
NCI0681620		0.30	54.53
NCI0687770		0.36	56.12
NCI0647884		0.39	51.28
NCI0149855		0.4	60.6
NCI0341662		0.55	53.79
NCI0648594		0.6	54.88

Table 6 (continued)

Hits	Structure	EC ₅₀ ^a (μM)	GFS ^b
NCI0672592		0.66	61.28
NCI0647884		0.39	51.28
NCI0687769		0.78	55.71
NCI0697654		0.83	59.52
NCI0062052		0.86	51.49

^a Estimated IC₅₀.^b GOLD fitness score.

52.18, lower than some highly ranked hits, which indicated that these highly ranked hits might have a better or comparable activities with Bay 43-9006. The hits with GOLD fitness scores higher than 50 were selected for detailed investigation, and 113 hits were obtained. The highest scoring poses of those highly ranked compounds were superimposed to the Raf-1 homology model for the investigation of the interactions between the protein and the ligands. For most compounds with high scores, hydrogen bonds between ligands and Glu393 and/or Asp486, or Cys424 involving pyridyl N-atom and amide –NH group were observed from the superimposition. It revealed that hydrogen bond played major roles in deciding the fitness scores of the compounds.

In addition to estimated activities and GOLD fitness scores, Lipinski's rule-of-five, widely accepted as a guideline for its conceptual simplicity and ease of calculation, was also used as a criterion for the assessment of possible hits on drug-like properties. A small set of 29 hits with high estimated IC₅₀ values, good GOLD fitness scores, and favorable drug-likeness were ultimately selected after careful observations, analyses and comparisons. Some representative hits are shown in Table 6, from which the current structural classes of Raf-1 kinase inhibitors can be seen evidently, including ureas (e.g., NCI0212285, NCI0120649, NCI0212040), urea bioisosteres (e.g., NCI0687770, NCI0341662, NCI0648594, NCI0647884), oxindoles (e.g., NCI0697654) and relevant structures. Besides, some novel structures were also obtained. The diversity of the hits demonstrated the pharmacophore model was able to retrieve

hits with similar features to existing Raf-1 kinase inhibitors as well as novel scaffolds.

Among these hits, compound NCI0342021 serves as a promising novel lead against Raf-1 kinase with the highest estimated IC₅₀ value of 0.029 μM and a high GOLD fitness score above that of Bay 43-9006, which suggests it may be a promising novel lead against Raf-1 kinase. It mapped all the features of Hypo1 by choosing “Best fit” (Fig. 6(a)), and the docked conformation formed several hydrogen bonds with key residues in the active site of Raf-1 homology model (Fig. 6(b)). And the analysis by ligplot 4.22 showed similar interactions on hydrogen bond and hydrophobic interactions consistent with the pharmacophore mapping and docking result (Fig. 6(c)).

4. Conclusion

Pharmacophore identification of substituted ureas as Raf-1 inhibitors indicated that the best pharmacophore model consisted of one hydrogen-bond acceptor (HBA), one hydrogen-bond donor (HBD), and three hydrophobic (HY) features. The pharmacophore model was well validated to be of high predictability for estimating the activities over a variety of compounds, and evaluating how well diverse compounds can be mapped onto the pharmacophore before conducting any further experimental study. Furthermore, the pharmacophore could be used as a powerful search tool for consequent database searching to retrieve potential compounds with inhibitory activities against Raf-1 kinase. Molecular docking acted as an additional tool for pharmacophore-based virtual

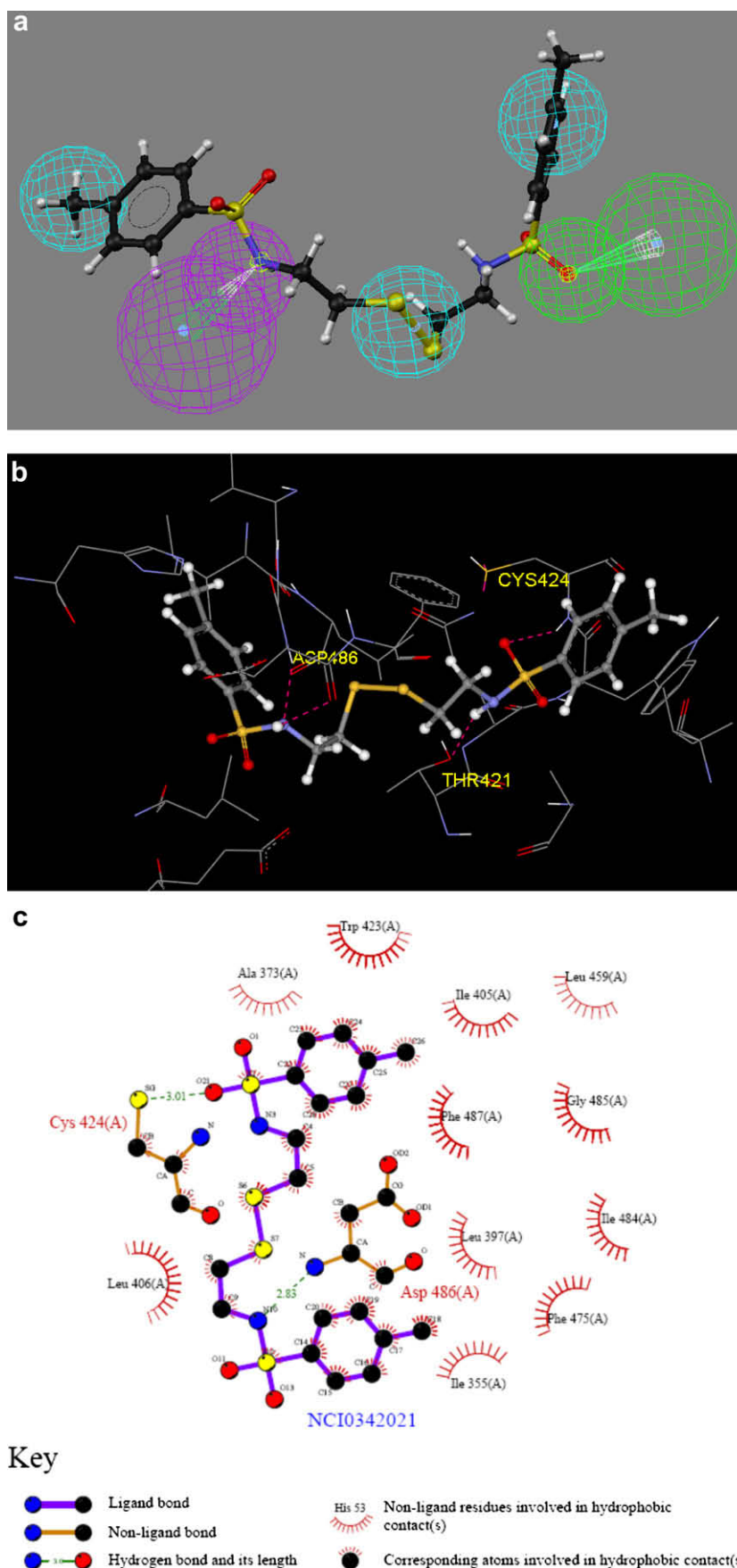


Fig. 6. (a) Pharmacophore mapping of NCI0342021; (b) docking result with Raf-1 homology model of NCI0342021; (c) the docking result was analyzed using the ligplot 4.22 to identify some specific contacts between atoms of NCI0342021 and Raf-1 homology model.

screening, the concurrent use of which is believed to make the discovery of potent Raf-1 kinase inhibitors more efficient. The whole procedures of pharmacophore modeling, database searching, and molecular docking resulted in the identification of some putative novel inhibitors of the Raf-1 kinase for further investigation.

Acknowledgement

This work was supported by grants from Natural Science Foundation of Jiangsu Province (BK2006151), National Found for Fostering Talents of Basic Science (NFFTB, No. J0630858) and Creation Programme for Graduates.

References

- [1] P.J. Roberts, C.J. Der, *Oncogene* 26 (2007) 3291–3310.
- [2] J.A. McCubrey, L.S. Steelman, W.H. Chappell, S.L. Abrams, E.W. Wong, F. Chang, B. Lehmann, D.M. Terrian, M. Milella, A. Tafuri, F. Stivala, M. Libra, J. Basecke, C. Evangelisti, A.M. Martelli, R.A. Franklin, *Biochim. Biophys. Acta* 1773 (2007) 1263–1284.
- [3] M.J. Robinson, M.H. Cobb, *Curr. Opin. Cell Biol.* 9 (1997) 180–186.
- [4] W. Kolch, *Biochem. J.* 351 (Pt 2) (2000) 289–305.
- [5] R.A. Smith, J. Dumas, L. Adnane, S.M. Wilhelm, *Curr. Top. Med. Chem.* 6 (2006) 1071–1089.
- [6] J. Avruch, A. Khokhlatchev, J.M. Kyriakis, Z. Luo, G. Tzivion, D. Vavvas, X.F. Zhang, *Recent Prog. Horm. Res.* 56 (2001) 127–155.
- [7] S.L. Campbell, R. Khosravi-Far, K.L. Rossman, G.J. Clark, C.J. Der, *Oncogene* 17 (1998) 1395–1413.
- [8] C.J. Marshall, *Curr. Opin. Cell Biol.* 8 (1996) 197–204.
- [9] H. Chong, H.G. Vikis, K.L. Guan, *Cell. Signal.* 15 (2003) 463–469.
- [10] C. Hagemann, U.R. Rapp, *Exp. Cell Res.* 253 (1999) 34–46.
- [11] S.S. Sridhar, D. Hedley, L.L. Siu, *Mol. Cancer Ther.* 4 (2005) 677–685.
- [12] S.M. Storm, J.L. Cleveland, U.R. Rapp, *Oncogene* 5 (1990) 345–351.
- [13] P.T. Wan, M.J. Garnett, S.M. Roe, S. Lee, D. Niculescu-Duvaz, V.M. Good, C.M. Jones, C.J. Marshall, C.J. Springer, D. Barford, R. Marais, *Cell* 116 (2004) 855–867.
- [14] M. Baccarini, *FEBS Lett.* 579 (2005) 3271–3277.
- [15] R.A. Smith, J. Barbosa, C.L. Blum, M.A. Bobko, Y.V. Caringal, R. Dally, J.S. Johnson, M.E. Katz, N. Kennure, J. Kingery-Wood, W. Lee, T.B. Lowinger, J. Lyons, V. Marsh, D.H. Rogers, S. Swartz, T. Walling, H. Wild, *Bioorg. Med. Chem. Lett.* 11 (2001) 2775–2778.
- [16] U.R. Khire, D. Bankston, J. Barbosa, D.R. Brittelli, Y. Caringal, R. Carlson, J. Dumas, T. Gane, S.L. Heald, B. Hibner, J.S. Johnson, M.E. Katz, N. Kennure, J. Kingery-Wood, W. Lee, X.G. Liu, T.B. Lowinger, I. McAlexander, M.K. Monahan, R. Natero, J. Renick, B. Riedl, H. Rong, R.N. Sibley, R.A. Smith, D. Wolanin, *Bioorg. Med. Chem. Lett.* 14 (2004) 783–786.
- [17] N.J. Liverton, J.W. Butcher, C.F. Claiborne, D.A. Claremon, B.E. Libby, K.T. Nguyen, S.M. Pitzzenberger, H.G. Selnick, G.R. Smith, A. Tebben, J.P. Vacca, S.L. Varga, L. Agarwal, K. Dancheck, A.J. Forsyth, D.S. Fletcher, B. Frantz, W.A. Hanlon, C.F. Harper, S.J. Hofsess, M. Kostura, J. Lin, S. Luell, E.A. O'Neill, S.J. O'Keefe, et al., *J. Med. Chem.* 42 (1999) 2180–2190.
- [18] K. Lackey, M. Cory, R. Davis, S.V. Frye, P.A. Harris, R.N. Hunter, D.K. Jung, O.B. McDonald, R.W. McNutt, M.R. Peel, R.D. Rutkowske, J.M. Veal, E.R. Wood, *Bioorg. Med. Chem. Lett.* 10 (2000) 223–226.
- [19] K.L. Oterloney McDonald, Ronda Davis-Ward, Edgar Wood, P.M. Vicente Samano, Felix Deanda, Robert Hunter, *Bioorg. Med. Chem. Lett.* 16 (2006) 5378–5383.
- [20] T.B. Lowinger, B. Riedl, J. Dumas, R.A. Smith, *Curr. Pharm. Des.* 8 (2002) 2269–2278.
- [21] S. Wilhelm, C. Carter, M. Lynch, T. Lowinger, J. Dumas, R.A. Smith, B. Schwartz, R. Simantov, S. Kelley, *Nat. Rev. Drug Discov.* 5 (2006) 835–844.
- [22] I.J. Macdougall, R. Griffith, *J. Mol. Graph. Model.* 26 (2008) 1113–1124.
- [23] K.B. Nagakumar Bharatham, Keun Woo Lee, *J. Mol. Graph. Model.* 25 (2007) 813–823.
- [24] S. Vadivelan, B.N. Sinha, S.J. Irudayam, S.A. Jagarlapudi, *J. Chem. Inf. Model.* 47 (2007) 1526–1535.
- [25] R. Dayam, T. Sanchez, O. Clement, R. Shoemaker, S. Sei, N. Neamati, *J. Med. Chem.* 48 (2005) 111–120.
- [26] S. Vadivelan, B.N. Sinha, G. Rambabu, K. Boppana, S.A. Jagarlapudi, *J. Mol. Graph. Model.* 26 (2008) 935–946.
- [27] T. Zhu, Y. Jiao, Y.D. Chen, X. Wang, H.F. Li, L.Y. Zhang, T. Lu, *Bioorg. Med. Chem. Lett.* 18 (2008) 2346–2350.
- [28] R. Thaimattam, P. Daga, S.A. Rajjak, R. Banerjee, J. Iqbal, *Bioorg. Med. Chem.* 12 (2004) 6415–6425.
- [29] Y.D. Chen, Y.J. Jiang, J.W. Zhou, Q.S. Yu, Q.D. You, *J. Mol. Graph. Model.* 26 (2008) 1160–1168.
- [30] A.K. Debnath, *J. Med. Chem.* 45 (2002) 41–53.
- [31] N. Bharatham, K. Bharatham, K.W. Lee, *J. Mol. Graph. Model.* 25 (2007) 813–823.
- [32] S.O. Ingo Muegge, *Drug Discov. Today: Technol.* 3 (2006) 405–411.
- [33] E. Vangrevelinghe, K. Zimmermann, J. Schoepfer, R. Portmann, D. Fabbro, P. Furet, *J. Med. Chem.* 46 (2003) 2656–2662.
- [34] I.G. Tikhonova, C.S. Sum, S. Neumann, S. Engel, B.M. Raaka, S. Costanzi, M.C. Gershengorn, *J. Med. Chem.* 51 (2008) 625–633.
- [35] J.M. Rollinger, T.M. Steindl, D. Schuster, J. Kirchmair, K. Anrain, E.P. Ellmerer, T. Langer, H. Stuppner, P. Wutzler, M. Schmidtke, *J. Med. Chem.* 51 (2008) 842–851.