



Original article

3D-QSAR with the aid of pharmacophore search and docking-based alignments for farnesyltransferase inhibitors

Madhura Vaidya, Mathias Weigt, Michael Wiese*

Pharmaceutical Institute, University of Bonn, An der Immenburg 4, 53121 Bonn, Germany

ARTICLE INFO

Article history:

Received 2 December 2008

Received in revised form

26 January 2009

Accepted 29 April 2009

Available online 8 May 2009

Keywords:

Farnesyltransferase inhibitors

3D-QSAR

CoMFA

CoMSIA

ABSTRACT

Farnesyltransferase is a potential drug target for treating various types of cancers. Three-dimensional quantitative structure–activity relationships (3D-QSAR) for a series of farnesyltransferase inhibitors were investigated using comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) techniques. Pharmacophore search and molecular docking methods were used for construction of the molecular alignments. While the 3D-QSAR models were created for a training set of 33 compounds, their external predictivity was proven using a test set of 12 compounds. The results provided a comprehensive insight into the relationship between the structural features and the activities of farnesyltransferase inhibitors. This investigation will facilitate optimization of the design of new potential farnesyltransferase inhibitors.

© 2009 Elsevier Masson SAS. All rights reserved.

1. Introduction

Farnesyltransferase (Ftase) inhibitors (FTIs) are potential therapeutic agents for the treatment of a variety of cancers. Ftase catalyzes the transfer of farnesyl moiety from farnesyl pyrophosphate to a cysteine (C) residue found in the polypeptide CAAX motif (A, aliphatic amino acid, X, Met or Ser) in the carboxy terminal of a group of membrane-bound small G proteins such as Ras, RhoB, RhoE, lamins A and B, and transducin. Ras farnesylation is a post-translational modification step required for association with the plasma membrane where the Ras proteins function. The latter play a pivotal role in the cell surface growth receptor signal transduction pathways that regulate cell differentiation, proliferation, migration, and survival [1]. Although the farnesylation assists in the subcellular localization and transformation of the oncogenic Ras variants, the efficacy of FTIs in preclinical testing does not correlate with the presence or absence of activating mutations of Ras [2]. The antineoplastic effects of FTIs are due to malignant transformation blockage by the inhibition of farnesylation of proteins other than Ras (e.g. centromere-binding proteins CENP-E and CENP-F) [3]. Thus FTIs may stop cancer growth by interfering with bipolar spindle formation during transition from prophase to metaphase in mitosis. While FTIs are active against the tumors driven by oncogenic H-Ras, they are inefficient against those by N-Ras and K-Ras.

The latter are possible substrates for alternative prenylation by geranylgeranyl-transferase when Ftase is blocked by FTIs [4].

A wide range of structural classes have been identified as FTIs. Novel inhibitors can be classified either as molecules designed on the CAAX motif (peptidomimetics) that interact with the Ftase residues involved in Ras protein binding or the Farnesyl moiety (FPP mimetics) that interact with the FPP binding site or bisubstrate inhibitors that incorporate structural motifs of both FPP and the CAAX tetrapeptide [5–7]. In the past few years non-thiol, non-peptidic and imidazole-containing chemical entities have emerged as FTIs [8]. Tipifarnib (R 115777), a selective non-thiol FTI, has been extensively used as a template in designing novel classes of FTIs. Abbott Laboratory has reported different series of compounds containing a cyanobenzyl group on the imidazole ring [9–13]. These inhibitors have shown promising results in the preclinical testing in cell culture and animal models of cancer. However, only a few QSAR studies have been reported by now for this class of compounds [14,15].

An important challenge for the compounds treated in the present study is to deal with their flexibility since the applications of CoMFA and CoMSIA require the optimized 3D conformations of all molecules. As crystal structures of Ftase with some FTIs are available in the published literature [16–22], new FTIs could be designed by virtual screening and docking. However, unlike QSAR methods, these methods do not readily yield information regarding the importance of molecular substructures for activity. In order to facilitate the design of selective Ftase inhibitors, a three-dimensional quantitative structure–activity relationship (3D-QSAR) study was conducted using comparative molecular field

* Corresponding author.

E-mail address: mwiese@uni-bonn.de (M. Wiese).

analysis (CoMFA) [23] and comparative molecular similarity indices analysis (CoMSIA) [24] based on pharmacophore elucidation and docking. Alignments of ligands that are known to bind to the Ftase protein were generated by superimposition on pharmacophore points and docking into the protein binding site. Subsequently, these alignments were used to derive 3D-QSAR models. The 3D-QSAR models may be useful in the design of novel FTIs.

2. Results and discussion

A dataset of 46 FTIs containing an imidazole and cyanophenyl group as structural key elements was used to perform 3D-QSAR studies. In such investigations, the molecular alignment and conformation determination are very important for the reliability and validity of the resulting model. Due to the flexibility of the compounds, it is difficult to choose a suitable conformation that achieves a meaningful superimposition. In an ideal alignment the biologically active conformations should be aligned taking into account the orientations that the ligands adopt at the binding site of the protein. Receptor-guided alignment has been shown to produce models with better statistics than those from the ligand-based approach, presumably because the alignment using receptor information is more realistic [25,32]. The availability of X-ray data from crystallized protein ligand complexes enabled the inclusion of additional information from the receptor site. Therefore, we applied a strategy of combining conformations obtained from

docking and pharmacophoric alignment. For the 3D-QSAR methods, CoMFA and CoMSIA, this provides a reasonable solution.

2.1. Pharmacophore elucidation and statistics of CoMFA and CoMSIA

The crystal structures of Ftase were aligned using MOE and ligands from these aligned structures were taken as the pharmacophore template. The essential pharmacophore features of the FTIs for interaction with Ftase are the two aromatic rings and the imidazole nitrogen (hydrogen-bond acceptor). The cyano group representing another hydrogen-bond acceptor and the other aromatic/hydrophobic ring served as optional features. The cyanophenyl and imidazole ring atoms remained as the main pharmacophore features (Fig. 1). The constructed pharmacophore model was comparable to that reported by Equbal et al. [26]. The pharmacophore served as a helpful tool in constructing an alignment of a set of 46 compounds. The alignment of the compounds under study is illustrated in Fig. 2.

The PLS results of the CoMFA which are summarized in Table 2 show that all statistical indices are reasonably high. As depicted in Table 2, the used CoMFA model leads to a q^2 value of 0.761 and an estimated standard error of prediction of 0.348 ($r^2_{\text{conv}} = 0.997$, $F = 2016$, $n = 5$). These values indicate that the CoMFA model has a good conventional statistical correlation and it allows good predictions of the biological activity data of the FTIs also for the test

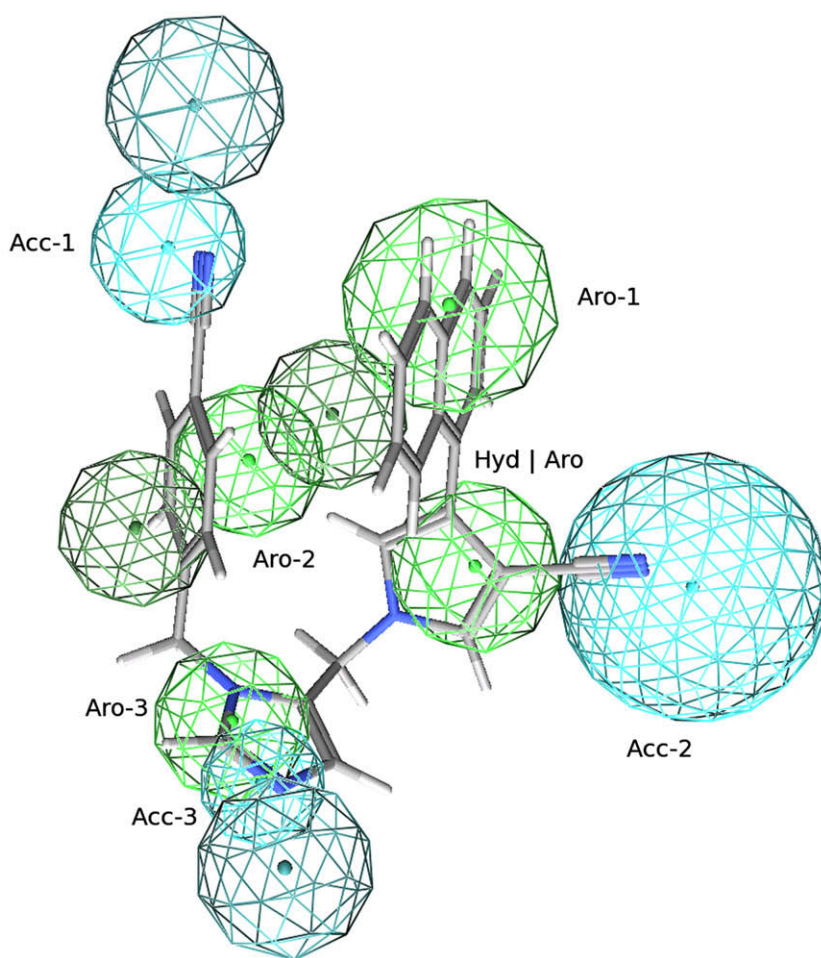


Fig. 1. Pharmacophore mapping of compound **43**. Pharmacophore features are color coded with cyan and green contours representing the hydrogen-bond acceptor feature (HA), hydrophobic feature (HY) and/or ring aromatic features (RA), respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

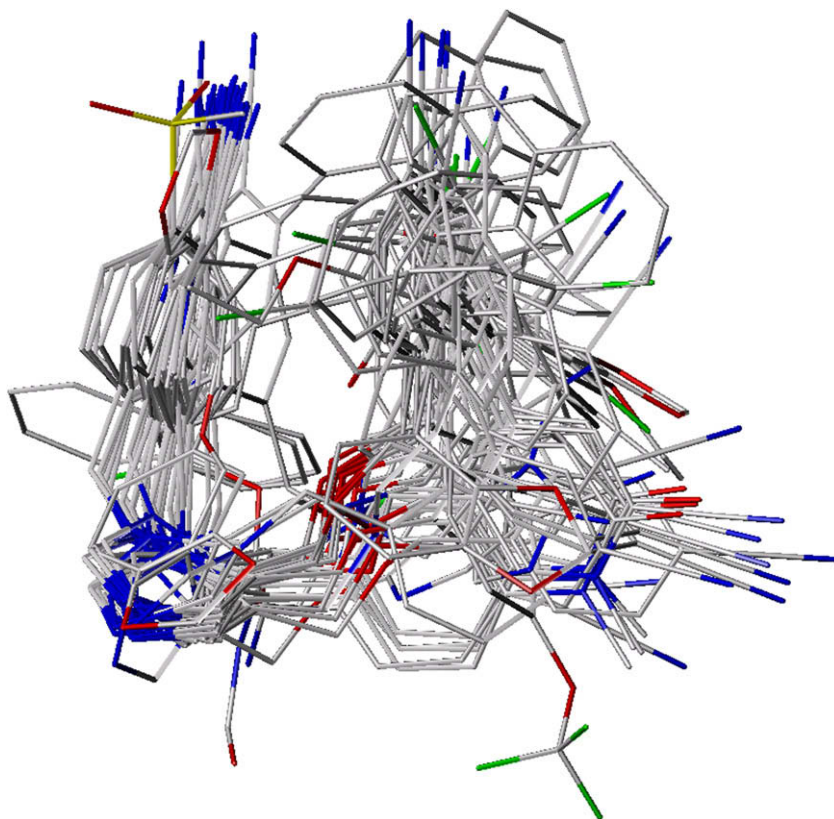


Fig. 2. Alignment of training set compounds based on pharmacophore model.

set. The CoMSIA was performed using steric, electrostatic, hydrophobic and hydrogen-bond donor and acceptor descriptors. Several combinations of descriptors were used, which are complimentary to previously generated CoMFA models (Table 3). CoMSIA models exhibited lower predictive properties than the CoMFA models. A combination of steric and electrostatic fields only resulted in a CoMSIA model with a higher q^2 value (0.700), a good r^2_{conv} (0.991) and a lower SEE (0.070), than other CoMSIA models. Therefore, it was selected as the best model (Table 2) to generate contour maps and explain the SAR.

The high value of q^2 appears to be necessary but not the sufficient condition for the model to have a high predictive power. The external validation is important to establish a reliable QSAR model. To validate the stability and predictivity of the 3D-QSAR models, 13 compounds that were not included in the construction of the CoMFA and CoMSIA models were selected as the test set. Both the CoMFA and CoMSIA models showed reasonable external predictivity, yielding p^2 of 0.725 and 0.585 respectively. Compound **29** was removed as an outlier as it was not fitting in either of the alignments and was over-predicted in both models. The data for calculated and experimental activities of training and test sets are listed in Tables 6 and 7 respectively. The relationships between experimental inhibitory activities (pIC_{50}) and calculated activities by the CoMFA and CoMSIA models are presented in Figs. 4 and 5. The leave-many-out cross-validation with 5 groups was carried out 100 times and the mean r^2_{cv} was 0.669 for CoMFA and 0.594 for CoMSIA.

2.2. Docking and statistics of CoMFA and CoMSIA

All the compounds were docked into the active site of Ftase and the resulting relative conformations of all compounds were used for the following CoMFA and CoMSIA calculations. The resulting

alignment of compounds is shown in Fig. 3. The distance between the atom N1 of the imidazolyl group of the most active compound **46** and the zinc ion was 2.71 Å. The indole ring was oriented in a hydrophobic pocket defined by Tyr361 and Trp106. The cyano group on the phenyl ring was found to interact with Arg202).

CoMFA models (Table 4) generated from docked conformations showed q^2 of 0.750 with 5 components, r^2_{conv} of 0.989, F value of 508 and a good external predictivity ($p^2 = 0.747$). Using steric, electrostatic, hydrophobic, and hydrogen-bond donor and acceptor properties as descriptors, CoMSIA was performed. The results are presented in Table 5. As with the pharmacophore alignment, the best q^2 was obtained using only steric and electrostatic descriptor variables. This demonstrates that these variables are essential for the description of the interaction of FTIs with Ftase. A CoMSIA model with a cross-validated q^2 of 0.737 for 6 components, a conventional r^2 of 0.987 and also good external predictivity ($p^2 = 0.698$) was obtained (Table 4). These data demonstrate that the CoMSIA model is satisfactorily predictive too. The best CoMFA and CoMSIA models were subjected to rigorous statistical cross-validation using five groups to ascertain the true internal predictivity of the models. The mean r^2_{cv} values from 100 runs for CoMFA and CoMSIA are 0.653 and 0.643 respectively. The actual versus calculated biological activities (Tables 6 and 7) using the CoMFA and CoMSIA models of the training and test sets are plotted in Figs. 6 and 7 respectively.

2.3. Interpretation of QSAR models

To visualize the information content of the derived 3D-QSAR models, CoMFA and CoMSIA contour maps were generated. The field energies at each lattice point were calculated as the scalar results of the coefficient and the standard deviation associated with

Table 1

Structures and activities of the molecules in the training and test sets. Compounds forming the test set are marked by an asterisk.

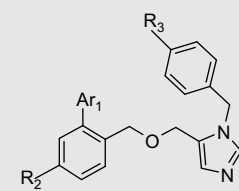
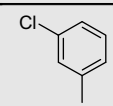
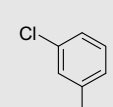
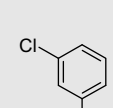
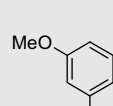
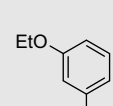
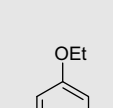
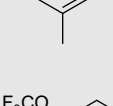
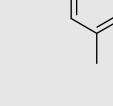
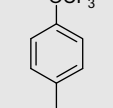
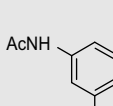
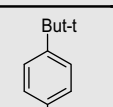
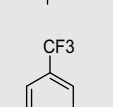
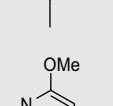
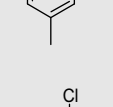
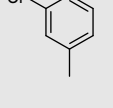
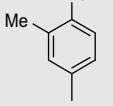
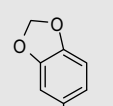
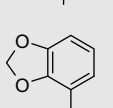
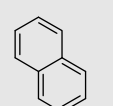
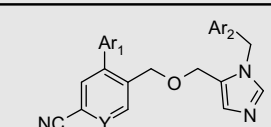
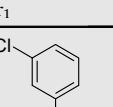
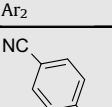
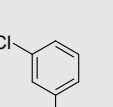
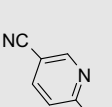


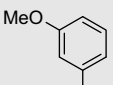

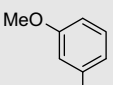

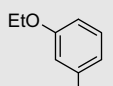

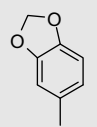
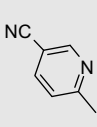
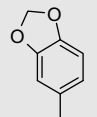
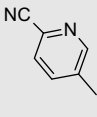
				
Compound	Ar ₁	R ₂	R ₃	pIC ₅₀
1*		Cl	CN	9.21
2		CN	CN	9.43
3		CN	MeSO ₂	7.02
4*		CN	Cl	8.66
5*		CN	Cl	8.85
6		CN	CN	8.90
7		CN	CN	9.31
8*		CN	CN	9.19
9		CN	CN	7.89
10		CN	CN	8.89

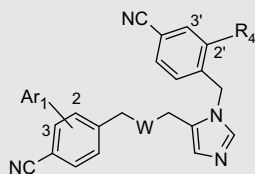
Table 1 (continued)

Compound	Ar ₁	R ₂	R ₃	pIC ₅₀
11*		CN	CN	9.36
12*		CN	CN	8.37
13*		CN	CN	8.12
14*		CN	CN	9.43
15		CN	CN	9.22
16		CN	Cl	8.08
17		CN	CN	9.09
18		CN	CN	9.04
19		CN	CN	8.89
				
Compound	Ar ₁	Y	Ar ₂	pIC ₅₀
20		N		8.37
21		CH		8.77

(continued on next page)

Table 1 (continued)

Compound	Ar ₁	Y	Ar ₂	pIC ₅₀
22		CH		8.52
23*		CH		8.36
24		CH		8.82
25		CH		9.24
26		CH		8.33
27*		CH		8.52



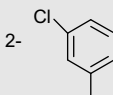

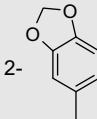
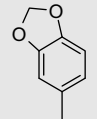
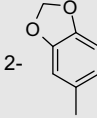




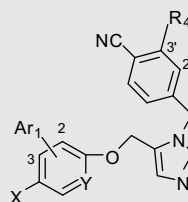




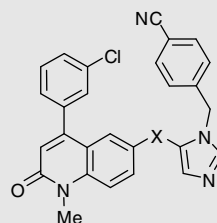
Compound	Ar ₁	W	R ₄	pIC ₅₀
28		O		8.00
29		O		6.46
30	 H	O		8.17
31	H	O		9.02

Table 1 (continued)

Compound	Ar ₁	W	R ₄	pIC ₅₀
32		O	H	8.32
33		NH	H	8.19



Compound	Ar ₁	X	Y	R ₄	pIC ₅₀
34*		Cl	CH	H	8.40
35		CN	CH	H	9.42
36	H	CN	CH		9.49
37	H	CN	N		9.05







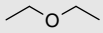
Compound	X	pIC ₅₀
38		7.33
39		7.54
40		8.42
41		8.68
42*		7.05

Table 1 (continued)

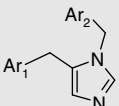
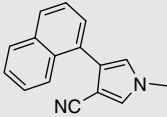
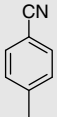
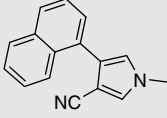
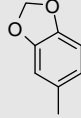
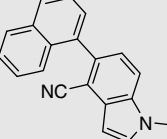
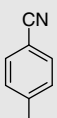
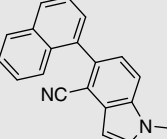
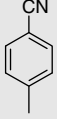
			
Compound	Ar ₁	Ar ₂	pIC ₅₀
43			9.24
44			8.21
45			9.03
46			9.82

Table 2

Best CoMFA and CoMSIA models using pharmacophore-based alignment.

	CoMFA	CoMSIA
q^2	0.761	0.700
SEP	0.348	0.398
Optimal no. of components	5	6
r^2_{conv}	0.997	0.991
SEE	0.037	0.070
F	2016	455
p^2	0.756	0.698
r^2_{cv} (5 groups) ^a	0.669	0.594
Contribution steric	0.575	0.575
Electrostatic	0.425	0.425

^a Average of 100 runs.

a particular column of the data table ("STDEV*COEFF"). This was always plotted as the percentage of the contribution to the CoMFA or CoMSIA equation. The green contours represent the regions of high steric tolerance and the yellow contours represent regions of

unfavorable steric effects. The blue contours describe the favorable interactions with the positively charged groups, whereas red contours describe those with the negatively charged groups.

2.3.1. Pharmacophore-based alignment for CoMFA and CoMSIA

The CoMFA steric and electrostatic contour plots from pharmacophore-based alignment are shown in Fig. 8. Compound **46** is the most active compound while **3** is a less active compound from the training set. The *para*-substituent of the phenyl group at N3 position of imidazole of the compounds is oriented towards the red contours, thereby indicating favorable interactions with negatively charged groups. Besides, a large positively charged blue contour is found near position 4 of the imidazole ring too. The aromatic substituents at 4 position of the imidazole ring were found to orient in the negative charge favorable red contours. These red regions indicate that more negatively charged groups substituted at *para*-position of phenyl ring and 4 position on the imidazole ring will enhance the biological activity. The indole and naphthyl rings of compound **46** are found to be mapped with some negative charge favorable red contour whereas the substituent of compound **3** at the same position is mapped near the positive charge favorable blue contours. The blue contours around the basic nitrogen atom of the imidazole ring clearly indicate that, in this region, a positive charge is mandatory for activity. The green contours around the indole ring of compound **46** suggest that the steric interactions play an important role. These green contours are flanked by sterically unfavorable yellow contours. The orientation of the substituent on 4 position of the imidazole group is very important for activity. The groups substituted at the 4 position on the imidazole ring are sterically preferred to produce good steric and hydrophobic interactions with the receptor (with Tyr361 and Trp106 as observed in docking studies), but these bulky groups must be restricted to some extent in order to avoid bad steric contacts between the inhibitor and the receptor. Most active compound **46** has the lipophilic fragment oriented in these sterically favorable contours, whereas the less active compound **3** is oriented in the sterically unfavorable yellow region. Similar sterically favored green regions appear around the 4 position of the imidazole group of all these compounds. There is a small sterically unfavorable region around the *para*-position of the phenyl ring and this result can be explained by considering the higher inhibitory activity of compound **46** than that of compound **3** which is oriented in the sterically unfavorable yellow CoMFA contours. The Ftase inhibitory activity seems quite sensitive to changes in the electron-withdrawing substituent at *para*-position of phenyl ring.

The CoMSIA contour plots from pharmacophore-based alignment are shown in Fig. 9. The steric and electrostatic contour maps of CoMSIA are similar to those of CoMFA in explaining the effect of substitution on the biological activity. Near the substituent at 4 position of imidazole, a very large green contour compared to that of CoMFA is observed. While the substituents at 4 position of the imidazole ring of both of the compounds are oriented in the same sterically favored green regions, only substituents of compound **3** on the same position are oriented in some sterically unfavorable yellow CoMSIA fields. Similarly a substituent at *para*-position of phenyl ring

Table 3

Results of CoMSIA analyses with pharmacophore-based alignment.

	S	E	H	A	D	S + E	S + H	S + E + H	S + E + A	S + E + H + A
q^2	0.398	0.647	0.205	−0.057	0.608	0.737	0.342	0.653	0.670	0.614
SEP	0.543	0.431	0.613	0.684	0.455	0.373	0.567	0.427	0.409	0.442
Optimal no. of components	4	5	3	1	6	6	4	6	5	5
r^2_{conv}	0.845	0.972	0.797	0.098	0.979	0.987	0.895	0.988	0.989	0.988
SEE	0.276	0.118	0.310	0.631	0.104	0.081	0.227	0.080	0.079	0.079
F	38	190	37	3	205	339	59	349	466	429

Table 4

Best CoMFA and CoMSIA models using docking-based alignment.

	CoMFA	CoMSIA
q^2	0.750	0.737
SEP	0.357	0.373
Optimal no. of components	5	6
r^2_{conv}	0.989	0.987
SEE	0.073	0.081
F	508	339
$r^2_{\text{prediction}}$	0.746	0.698
p^2	0.747	0.698
r^2_{cv} (5 groups) ^a	0.653	0.643
Contribution Steric	0.595	0.275
Electrostatic	0.405	0.725

^a Average of 100 runs.

of compound **3** was found to be oriented in sterically unfavorable yellow contours. A yellow region around the linker between phenyl and imidazole ring represents an area where no substitution is favored.

2.3.2. Docking-based alignment for CoMFA and CoMSIA

CoMFA and CoMSIA contour maps of steric and electrostatic fields revealed similar results. The CoMFA contour maps are displayed in Fig. 10. While the substituents at 4 position of the imidazole ring of all compounds are oriented in the green CoMFA contours favored by steric interactions, those of **34** and **46** are even better positioned to the same CoMFA contours and that of compound **3** is oriented towards some yellow sterically unfavorable contours. The substituent at *para*-position of phenyl of compound **3** showed bad steric contacts with the receptor, which explains the lower activity of this compound. The N3 substituent of all the compounds is oriented in negatively charged (electron-rich) favorable red regions, thereby indicating that an electron-withdrawing substituent enhances the activity. Strong charge withdrawing groups linked to the phenyl ring will make the charge distributed on the phenyl ring relatively positive. A blue area, which represents a region where electron deficiency is advantageous, can be seen around the phenyl ring. This can be attributed to the ring substituents with a strong electron-withdrawing effect such as the cyano group. The substituents at 4 position of imidazole ring of both compounds are assigned by red contours as regions favored by electron-rich groups only that of compound **3** is positioned to some blue contours.

The CoMSIA contour maps (Fig. 11) show a green contour around the *meta*- and *ortho*-positions of the phenyl ring. These regions support the observation that compounds **31**, **35**, and **37** with electron-rich substitution at the *ortho*-position are among the compounds with high Ftase inhibitory activity. Again the substituents at 4 position of the phenyl ring found to be oriented towards sterically favorable green contoured regions; only that of compound **3** is located in the sterically unfavorable yellow regions.

The results of 3D-QSAR, CoMFA and CoMSIA obtained from both alignments are complimentary to each other. The detailed contour analysis of both CoMFA and CoMSIA models enabled us to point out

Table 6

Actual and calculated Ftase inhibitory activities of training set used in the present study from best CoMFA and CoMSIA models.

Compound	Actual activity (pIC ₅₀)	Calculated activity (pIC ₅₀)			
		CoMFA ^a	CoMSIA ^a	CoMFA ^b	CoMSIA ^b
2	9.43	9.412	9.439	9.350	9.258
3	7.02	7.017	7.043	7.020	7.186
6	8.9	9.000	9.024	8.948	8.870
7	9.31	9.315	9.271	9.374	9.315
9	7.89	7.869	7.852	7.818	7.876
10	8.89	8.840	8.843	8.904	8.834
15	9.22	9.224	9.191	9.186	9.198
16	8.08	8.082	7.998	8.079	8.068
17	9.09	9.076	9.064	9.113	9.119
18	9.04	9.033	9.048	9.051	9.027
19	8.89	8.862	8.861	8.872	8.917
20	8.37	8.374	8.364	8.370	8.285
21	8.77	8.752	8.772	8.873	8.845
22	8.52	8.524	8.474	8.408	8.350
24	8.82	8.809	8.827	8.806	8.920
25	9.24	9.240	9.238	9.231	9.285
26	8.33	8.363	8.339	8.292	8.368
28	8.00	8.003	8.067	8.069	7.991
30	8.17	8.141	8.099	8.094	8.186
31	9.02	9.016	9.002	8.999	9.017
32	8.32	8.336	8.359	8.329	8.308
33	8.19	8.166	8.154	8.249	8.212
35	9.42	9.453	9.400	9.494	9.431
36	9.49	9.503	9.514	9.353	9.480
37	9.05	9.072	9.143	9.051	9.069
38	7.33	7.454	7.568	7.244	7.297
39	7.54	7.511	7.517	7.557	7.597
40	8.42	8.404	8.363	8.479	8.411
41	8.68	8.670	8.599	8.616	8.676
43	9.24	9.219	9.210	9.295	9.376
44	8.21	8.157	8.150	8.225	8.063
45	9.03	9.014	9.078	9.186	9.107
46	9.82	9.850	9.891	9.722	9.815

^a From pharmacophore-based alignment.^b From docking-based alignment.

several structural requirements as mentioned in above discussion for the observed inhibitory activities (Fig. 12). The substituent at the 4 position on the imidazole ring may be replaced by a more hydrophobic aromatic substitution, which will result in the favorable placement of this substituent in conformationally and sterically favored region. Comparing the pharmacophore-based with the receptor-based 3D-QSAR the improved performance of the CoMSIA approach is obvious, while the differences in the case of CoMFA are negligible. Thus the combination of docking for the selection of the alignment followed by 3D-QSAR seems to be advantageous.

For a subset of the dataset considered here a QSAR study based on RMS alignment was reported by Puntambekar et al. [14]. Despite the different alignment their 3D-QSAR contour maps are comparable to those presented here. However, this study provides significant better statistical coefficients and additional information about the steric and electrostatic interactions of the N3 imidazole substituent emphasizing the benefit of combining docking with 3D-QSAR.

Table 5

Results of CoMSIA analyses using docking-based alignment.

	S	E	H	A	D	S + E	S + H	S + E + H	S + E + A	S + E + H + A
q^2	0.675	0.415	0.375	−0.086	−0.099	0.700	0.589	0.613	0.540	0.580
SEP	0.414	0.545	0.534	0.693	0.697	0.398	0.465	0.427	0.540	0.462
Optimal no. of components	6	5	2	1	1	6	6	3	5	5
r^2_{conv}	0.966	0.971	0.806	0.108	0.441	0.991	0.991	0.960	0.989	0.996
SEE	0.134	0.122	0.297	0.628	0.497	0.070	0.067	0.138	0.073	0.045
F	122	179	62	3	24	455	500	229	507	1357

Table 7

Actual and calculated FtsE inhibitory activities of test set used in the present study from best CoMFA and CoMSIA models.

Compound	Actual activity (pIC ₅₀)	Calculated activity (pIC ₅₀)			
		CoMFA ^a	CoMSIA ^a	CoMFA ^b	CoMSIA ^b
1	9.21	9.11	9.17	8.92	8.91
4	8.66	8.61	8.60	8.50	8.47
5	8.85	8.71	8.70	8.84	8.83
8	9.19	9.07	9.01	9.19	9.20
11	9.36	9.09	8.93	8.80	9.09
12	8.37	8.90	9.05	8.41	8.31
13	8.12	8.45	8.39	8.26	8.34
14	9.43	9.46	9.24	9.03	9.05
23	8.36	8.55	8.65	8.39	8.39
27	8.52	8.74	8.76	8.69	8.43
34	8.40	8.84	8.87	8.41	8.37
42	7.05	7.88	7.76	7.82	8.07

^a From pharmacophore-based alignment.

^b From docking-based alignment.

3. Methods

3.1. Biological data

A dataset comprising 46 imidazole and cyanophenyl containing FTIs was taken from the literature [11,12]. The compound structures and their biological data are given in Table 1. In this QSAR study, the biological activity of each compound has been expressed as the negative logarithm to the base ten of IC₅₀ (pIC₅₀). Thus, the data positively correlated linearly to the free energy change. The dataset was divided into training and test set based on the conformance with the pharmacophore model where the molecules that generated the best pharmacophore model were selected as training set and the remaining molecules were used as test set. All further studies were performed using the same training and test sets. The training set of 33 compounds (Table 1) was used for the generation of QSAR models and the predictivity of the models was analyzed by using the test set of 13 compounds where one compound had to be removed.

3.2. 3D-QSAR (CoMFA/CoMSIA) studies

All calculations were carried out on an x86-compatible PC running Gentoo Linux. For molecular modeling, SYBYL 7.3 [27], MOE 2006.08 for Linux [28] and GOLD version 3.1.1 [29] were used. Conformational clustering was done using a locally written program [30] in MATLAB R13 for Linux [31]. The compounds were constructed from the fragments in the SYBYL database with standard bond lengths and bond angles and geometry optimization was carried out using the standard Tripos force field with distance dependent-dielectric function and energy gradient of 0.001 kcal/mol. Conformational analyses of all 46 compounds from Table 1 were done using a repeated molecular dynamics-based simulated annealing approach as implemented in SYBYL 7.3. MMFF94 served as the force field with distance dependent electrostatics. A molecule was heated up to 1000 K within 2000 fs, held at this temperature for 2000 fs and annealed to 0 K for 10,000 fs using an exponential annealing function. By applying this procedure, a total of 100 conformations were sampled during 100 cycles to account for conformational flexibility and find the most likely conformations occurring most often in the resulting pool. All conformations were then further minimized with the MMFF94s' force field using Powell's method.

3.2.1. Molecular alignment

In CoMFA the relative interaction energies depend strongly on the relative molecular positions. Hence the proper alignment of molecules is a very important input. The conformations obtained from pharmacophore matching and docking studies were used as an initial alignment for the QSAR study.

The FtsE crystal structures 1LD7 [16], 1LD8 [16], 1MZC [17], 1NI1 [18], 1X81 [19], 2FOY [20] and 1SA4 [21] were aligned using the program MOE. Although the ligands bind to the enzyme active site in different modes, they have some common features such as cyanophenyl, imidazole and the hydrophobic rings lying in the same orientation. Therefore these features were used for the pharmacophore generation. The conformations of each compound under study were overlaid with the resulting pharmacophore. An initial model was built from the best scoring alignments selected

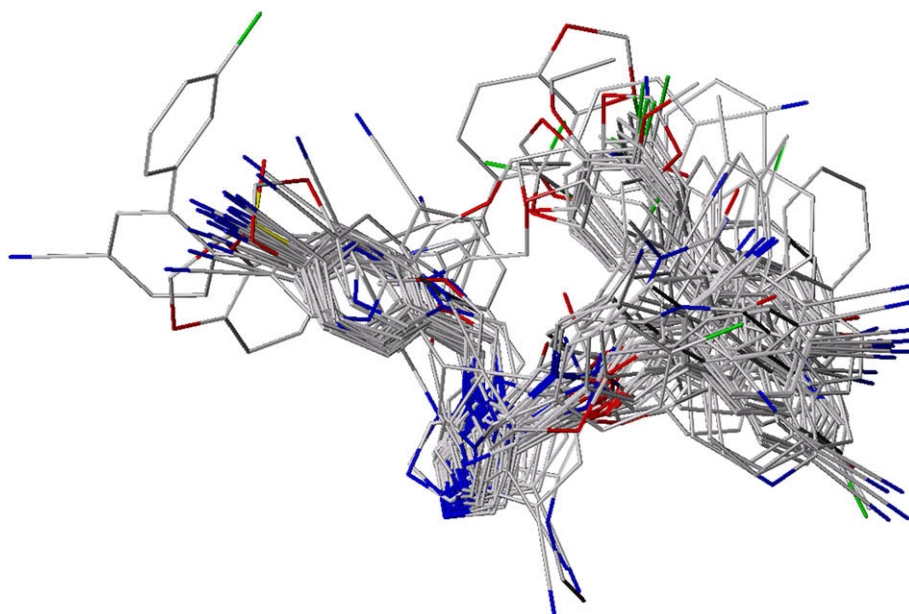


Fig. 3. Docking-based alignment of training set compounds.

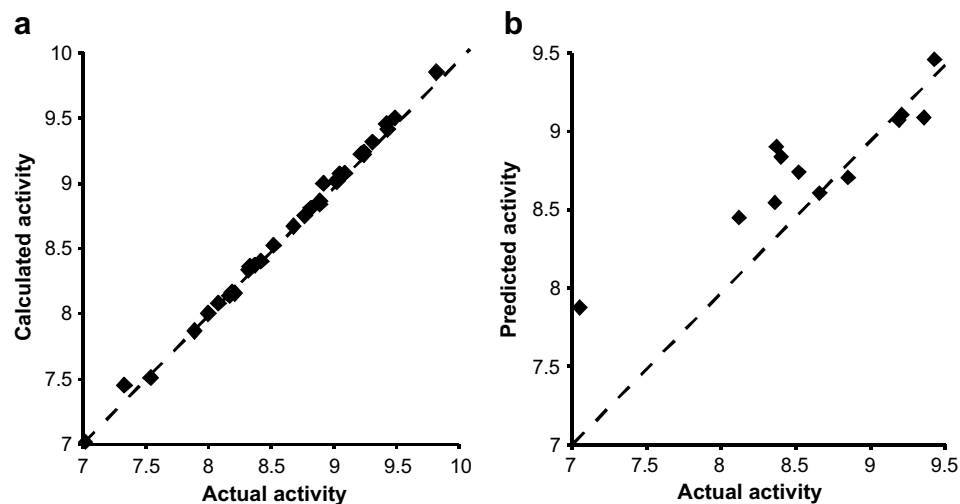


Fig. 4. Graph of observed activity versus calculated activities of training set (a) and test set (b) compounds from CoMFA of pharmacophore-based alignment.

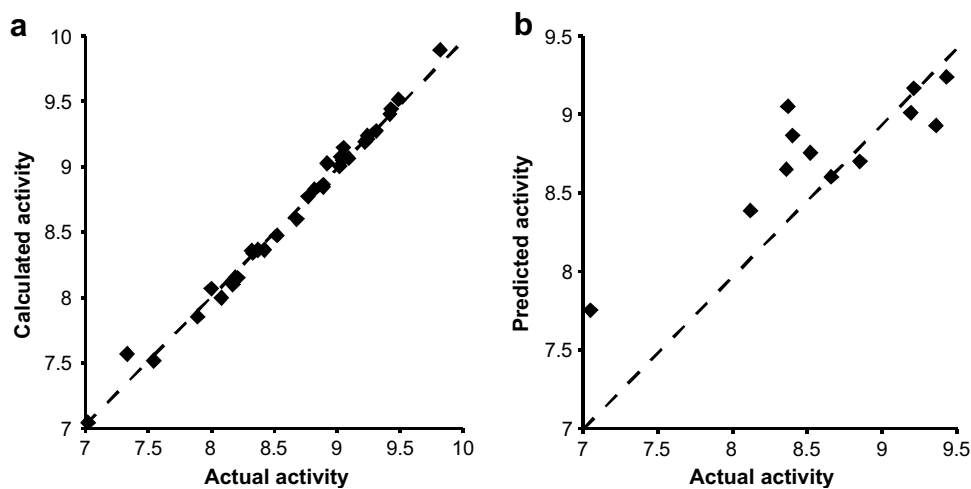


Fig. 5. Graph of observed activity versus calculated activities of training set (a) and test set (b) compounds from CoMSIA of pharmacophore-based alignment.

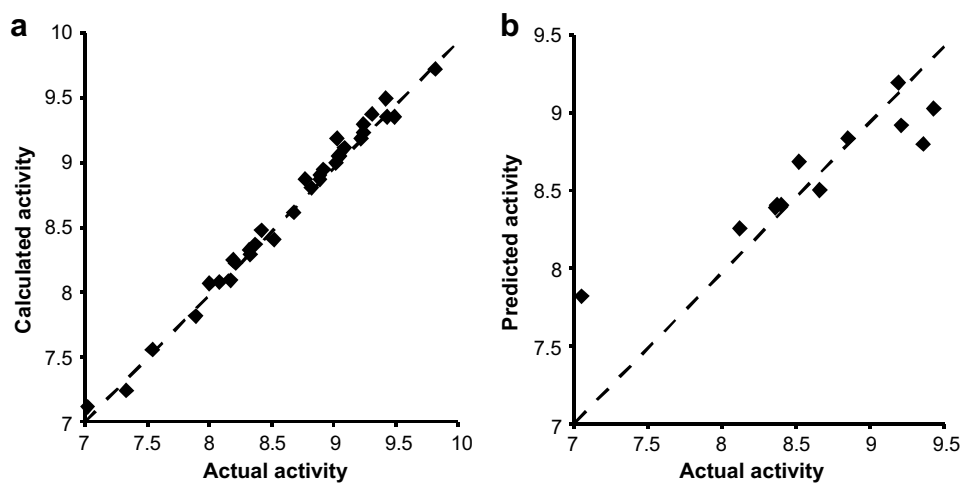


Fig. 6. Graph of observed activity versus predicted activities of training set (a) and test set (b) compounds from CoMFA of docking-based alignment.

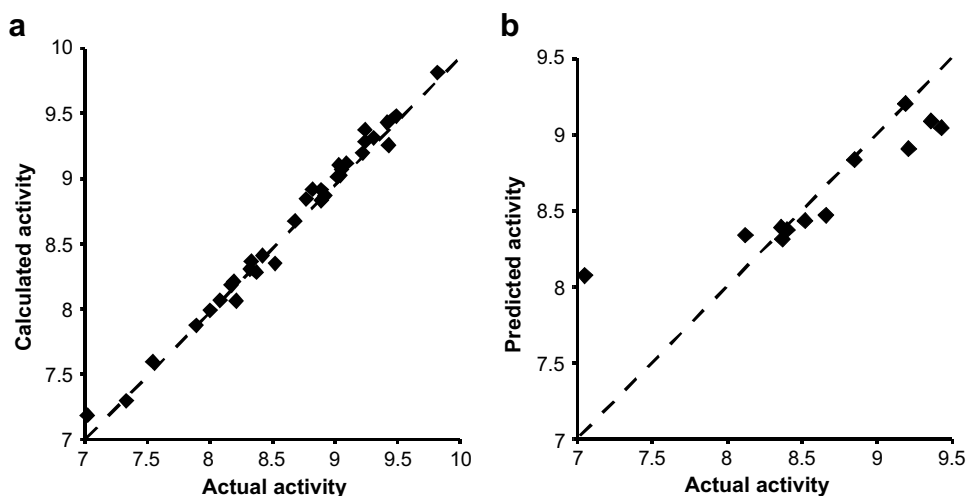


Fig. 7. Graph of observed activity versus predicted activities of training set (a) and test set (b) compounds from CoMSIA of docking-based alignment.

based on minimum root-mean-square distance (RMS) criteria. In the case of poorly predicted activities other alignments, at most 3–10 per compound, were tried.

For the docking, the crystal structure of the human Ftase in complex with its inhibitor U99 (PDB entry 1LD7) [16] was obtained from the Brookhaven Protein Data Bank. Water molecules were discarded from the pdb file and hydrogens were added. Charges were calculated by the Gasteiger–Hückel method and the charge of

zinc was set to +2. The structure was then subjected to energy minimization using the Powell method until an energy gradient of 0.05 kcal/mol was achieved. All the molecules were docked into the enzyme binding site using the GOLD program. The binding site was initially defined as all residues of the target within 10 Å from the zinc metal atom. GOLD score was chosen as the fitness function and the standard default settings were used in all calculations. From 10 independent genetic algorithm runs for each molecule, at most 10

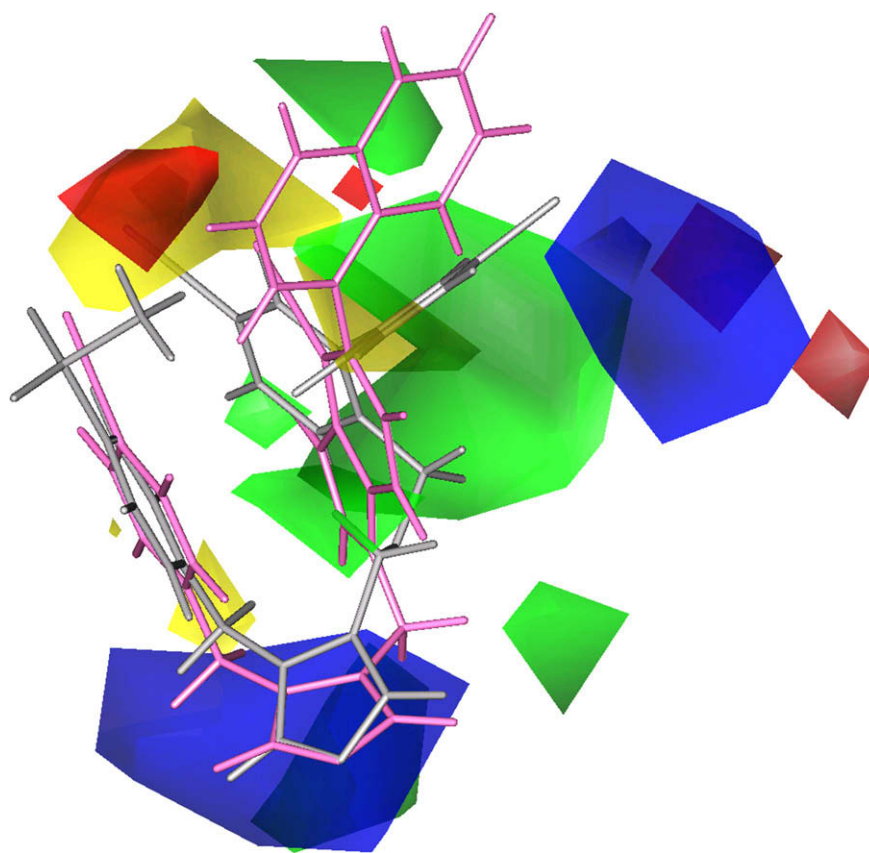


Fig. 8. CoMFA steric and electrostatic STDEV*COEFF contour plots from the pharmacophore-based alignment. Sterically favored areas are represented by green polyhedra. Sterically disfavored regions are represented by yellow polyhedra. Positively charged favored areas are represented by blue polyhedra. Negatively charged favored areas are represented by red polyhedra. The active compound **46** is shown in pink and less active molecule **3** is shown in gray. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

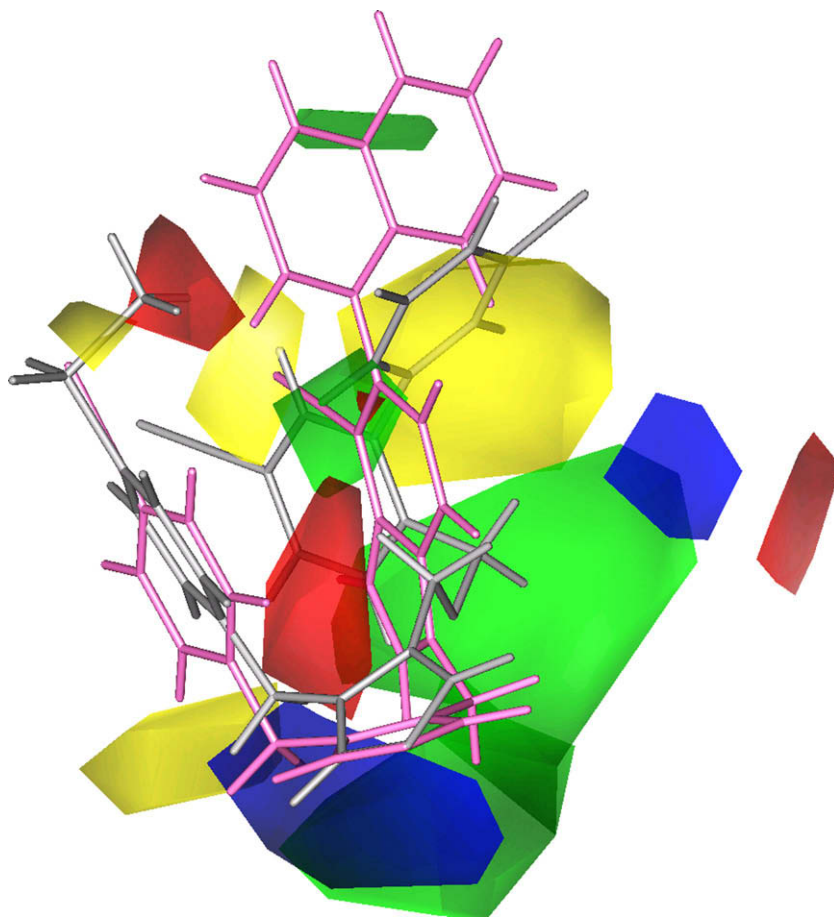


Fig. 9. CoMSIA steric and electrostatic fields pharmacophore-based alignment. Yellow indicates the regions where the steric interactions hinder binding and green areas where steric interaction will enhance binding. Red areas indicating negatively charged groups will enhance activity and blue areas where positively charged groups will enhance activity. The active compound **46** is shown in pink and less active molecule **3** is shown in gray. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

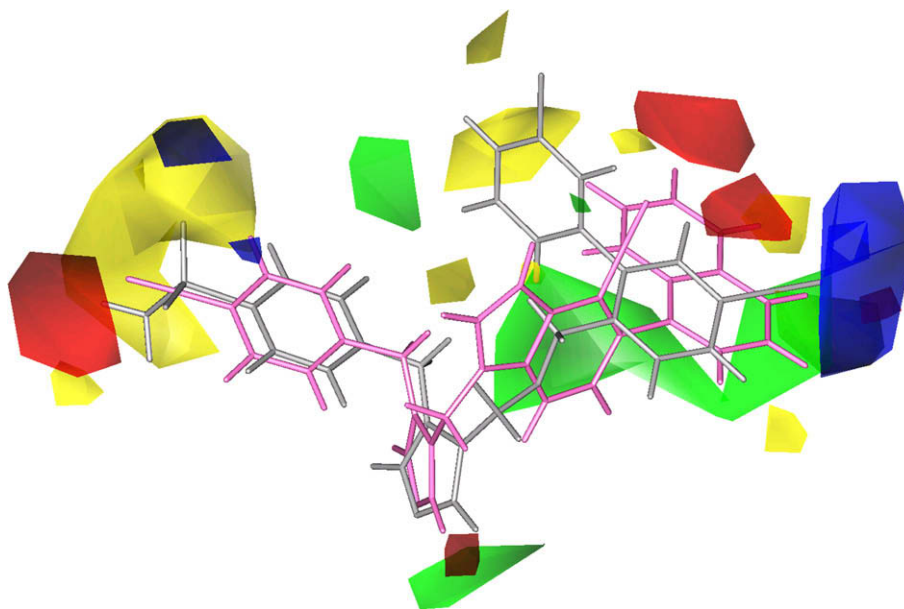


Fig. 10. CoMFA steric and electrostatic STDEV*COEFF contour plots from docking-based alignment. Sterically favored areas are represented by green polyhedra. Sterically disfavored regions are represented by yellow polyhedra. Positively charged favored areas are represented by blue polyhedra. Negatively charged favored areas are represented by red polyhedra. The active compound **46** is shown in pink and less active molecule **3** is shown in gray. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

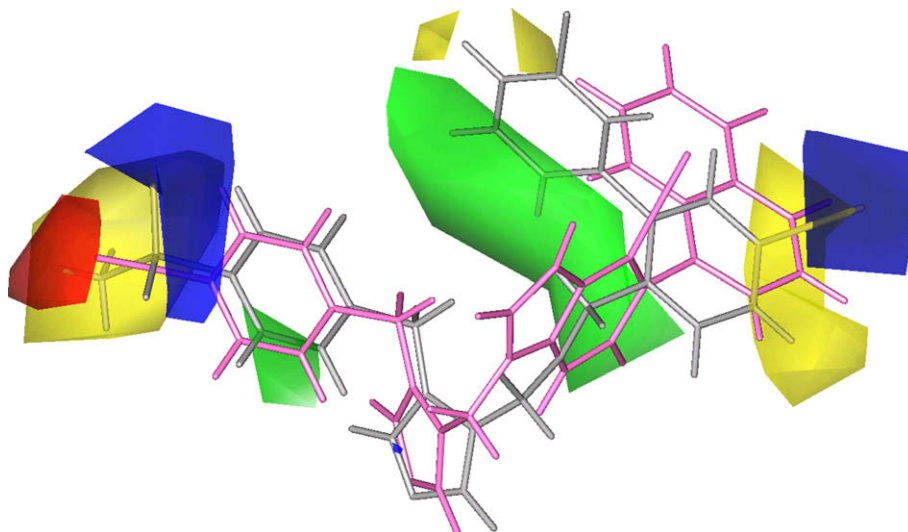


Fig. 11. CoMSIA steric and electrostatic fields from docking-based alignment. The active compound **46** is shown in pink and less active molecule **3** is shown in gray. Yellow indicates the regions where the steric interactions hinder binding and green areas where steric interaction will enhance binding. Red areas indicating negatively charged groups will enhance activity and blue areas where positively charged groups will enhance activity. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

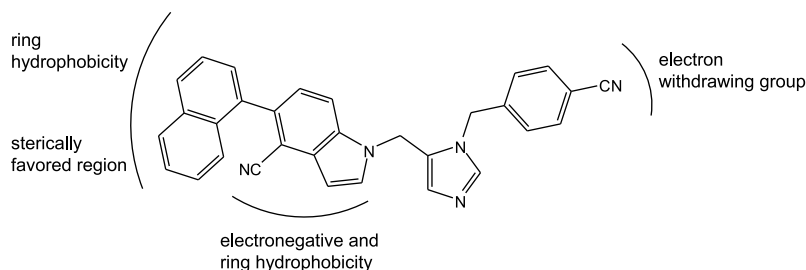


Fig. 12. Structural requirements for non-thiol based Ftase inhibitors.

best ranked docking poses were obtained. These were transferred to a SYBYL database and used for the CoMFA/CoMSIA study. The initial models were built using the best ranked pose of each compound. In the case of poorly predicted activity values other combinations were elaborated to find the best QSAR model.

3.2.2. CoMFA and CoMSIA settings

The CoMFA descriptors, steric (Lennard-Jones 6–12 potential) and electrostatic (Coulombic potential) field energies, were calculated by using the following standard parameters: an sp³ carbon probe atom with +1 charge and a van der Waals radius of 1.52 Å and energy cutoff of 30 kcal/mol. CoMSIA similarity indices descriptors (steric, electrostatic, hydrophobic, H-bond donor and H-bond acceptor fields) were calculated using a probe with a radius of 1.0 Å and a default value of 0.3 as the attenuation factor. A grid spacing of 2 Å was used for both CoMFA and CoMSIA. The CoMFA and CoMSIA descriptors were used as independent variables and pIC₅₀ as the dependent variable in the partial least squares (PLS) analysis to derive 3D-QSAR models. In subsequent PLS analyses, the alignment was refined and the CoMFA/CoMSIA models were optimized. To speed up the analysis and reduce noise, a minimum filter value σ of 2.0 kcal/mol was used. Final analysis was performed to calculate the conventional r^2 without cross-validation using the optimum number of components. To prove that these models were not a result of a chance correlation, a stability test was performed using the random-groups PLS method. Within this method, cross-validation was done with groups of compounds, which were

excluded from the model-building regression. Because of the random selection of the group members, this cross-validation was repeated hundred times. The final r^2_{cv} value was calculated by taking the mean of 100 runs. To further test the predictive power of the derived CoMFA and CoMSIA models, biological activities of the test set molecules were predicted using models derived from training set and p^2 values which are unbiased by the mean of the test set and representing the external predictivity were calculated.

4. Conclusions

Spatial alignment of the compounds is very important to derive meaningful 3D-QSAR models. The conformations of all compounds, obtained from the pharmacophore-based alignment and docking into the active site of Ftase, were used for the CoMFA and CoMSIA calculations. Both alignment procedures led to statistically robust and predictive 3D-QSAR models with superior r^2 and q^2 values compared to those of Puntambekar et al. [14]. Despite the different alignments, the contour maps from both models are similar in explaining the influence of substitution on activity. The predictive ability of the models was confirmed by predicting the activity of 12 compounds used as external test set. Therefore, the application of these models for quantitative prediction of Ftase inhibitory activities is possible within the covered structural space. The robust QSAR model and its three-dimensional contour map provide guidelines to design compounds with new scaffolds and optimize current molecules.

References

- [1] F.L. Zhang, P.J. Casey, *Annu. Rev. Biochem.* 65 (1996) 241–269.
- [2] W. Du, P.F. Lebowitz, G.C. Prendergast, *Mol. Cell. Biol.* 19 (1999) 1831–1840.
- [3] A. Oliff, *Biochim. Biophys. Acta* 1423 (1999) C19–C30.
- [4] A.D. Basso, P. Kirschmeier, W.R. Bishop, *J. Lipid Res.* 47 (2006) 15–31.
- [5] J.B. Gibbs, D.L. Pompliano, S.D. Mosser, E. Rands, R.B. Lingham, S.B. Singh, E.M. Scolnick, N.E. Kohl, A. Oliff, *J. Biol. Chem.* 268 (1993) 7617–7620.
- [6] J.L. Goldstein, M.S. Brown, S.J. Stradley, Y. Reiss, L.M. Gierasch, *J. Biol. Chem.* 266 (1991) 15575–15578.
- [7] D.M. Leonard, *J. Med. Chem.* 40 (1997) 2971–2990.
- [8] S. Halazy, J.P. Gotteland, M. Lamothe, D. Perrin, B.T. Hill, *Drugs Future* 20 (1997) 1133–1146.
- [9] N.-H. Lin, L. Wang, X. Wang, G.T. Wang, J. Cohen, W. Gu, H. Zhang, S.H. Rosenberg, H.L. Sham, *Bioorg. Med. Chem. Lett.* 14 (2004) 5057–5062.
- [10] L. Wang, G.T. Wang, X. Wang, Y. Tong, G. Sullivan, D. Park, N.M. Leonard, Q. Li, J. Cohen, W. Gu, H. Zhang, J.L. Bauch, C.G. Jakob, C.W. Hutchins, V.S. Stoll, K. Marsh, S.H. Rosenberg, H.L. Sham, N.-H. Lin, *J. Med. Chem.* 47 (2004) 612–626.
- [11] Q. Li, K.W. Woods, W. Wang, N.-H. Lin, A. Claiborne, W. Gu, J. Cohen, V.S. Stoll, C.W. Hutchins, D. Frost, S.H. Rosenberg, H.L. Sham, *Bioorg. Med. Chem. Lett.* 15 (2005) 2033–2039.
- [12] Q. Li, G.T. Wang, T. Li, S.L. Gwaltney, K.W. Woods, A. Claiborne, X. Wang, W. Gu, J. Cohen, V.S. Stoll, C.W. Hutchins, D. Frost, S.H. Rosenberg, H.L. Sham, *Bioorg. Med. Chem. Lett.* 14 (2004) 5057–5062.
- [13] N.-H. Lin, L. Wang, J. Cohen, W. Gu, D. Frost, H. Zhang, S.H. Rosenberg, H.L. Sham, *Bioorg. Med. Chem. Lett.* 13 (2003) 3821–3825.
- [14] D.S. Puntambekar, R. Giridhar, M.R. Yadav, *Bioorg. Med. Chem. Lett.* 16 (2006) 1821–1827.
- [15] D.S. Puntambekar, R. Giridhar, M.R. Yadav, *Eur. J. Med. Chem.* 43 (2008) 142–154.
- [16] I.M. Bell, S.N. Gallicchio, M. Abrams, L.S. Beese, D.C. Beshore, H. Bhimnathwala, M.J. Bogusky, C.A. Buser, J.C. Culberson, J. Davide, M. Ellis-Hutchings, C. Fernandes, J.B. Gibbs, S.L. Graham, K.A. Hamilton, G.D. Hartman, D.C. Heimbrook, C.F. Homnick, H.E. Huber, J.R. Huff, K. Kassahun, K.S. Koblan, N.E. Kohl, R.B. Lobell, J.J. Lynch Jr., R. Robinson, A.D. Rodrigues, J.S. Taylor, E.S. Walsh, T.M. Williams, C.B. Zartman, *J. Med. Chem.* 45 (2002) 2388–2409.
- [17] E.S. Polovnikova, M.J. McLeish, E.A. Sergienko, J.T. Burgner, N.L. Anderson, A.K. Bera, F. Jordan, G.L. Kenyon, M.S. Hasson, *Biochemistry* 42 (2003) 1820–1830.
- [18] Y. Tong, N.H. Lin, L. Wang, L. Hasvold, W. Wang, N. Leonard, T. Li, Q. Li, J. Cohen, W.Z. Gu, H. Zhang, V. Stoll, J. Bauch, K. Marsh, S.H. Rosenberg, H.L. Sham, *Bioorg. Med. Chem. Lett.* 13 (2003) 1571–1574.
- [19] Q. Li, A. Claiborne, T. Li, L. Hasvold, V.S. Stoll, S. Muchmore, C.G. Jakob, W. Gu, J. Cohen, C. Hutchins, D. Frost, S.H. Rosenbergand, H.L. Sham, *Bioorg. Med. Chem. Lett.* 14 (2004) 5367–5370.
- [20] K.M. Jude, A.L. Banerjee, M.K. Haldar, S. Manokaran, B. Roy, S. Mallik, D.K. Srivastava, D.W. Christianson, *J. Am. Chem. Soc.* 128 (2006) 3011–3018.
- [21] T.S. Reid, L.S. Beese, *Biochemistry* 43 (2004) 6877–6884.
- [22] C.L. Strickland, P.C. Weber, W.T. Windsor, Z. Wu, H.V. Le, M.M. Albanese, C.S. Alvarez, D. Cesarz, J. del Rosario, J. Deskus, A.K. Mallams, F.G. Njoroge, J.J. Piwinski, S. Remiszewski, R.R. Rossman, A.G. Taveras, B. Vibulbhan, R.J. Doll, V.M. Girijavallabhan, A.K. Ganguly, *J. Med. Chem.* 42 (1999) 2125–2135.
- [23] R.D. Cramer III, J.D. Bunce, D.E. Patterson, *Prog. Clin. Biol. Res.* 291 (1989) 161–165.
- [24] G. Klebe, *Perspect. Drug Discov. Des.* 12/13/14 (1998) 87–104.
- [25] K.H. Hyun, D.Y. Lee, B.-S. Lee, C.K. Kim, *QSAR Comb. Sci.* 23 (2004) 637–649.
- [26] T. Equbal, O. Silakari, G. Rambabu, M. Ravikumar, *Bioorg. Med. Chem. Lett.* 17 (2007) 1594–1600.
- [27] SYBYL 7.3, Tripos Inc., St. Louis, Missouri, USA.
- [28] MOE 2006.08 for Linux, Chemical Computing Group Inc., Montreal, Quebec, Canada.
- [29] GOLD version 3.1.1 is distributed by the Cambridge Crystallographic Data Center, UK.
- [30] M. Weigt, Ph.D. Thesis, University of Bonn, 2006.
- [31] MATLAB R13 for Linux, The MathWorks Inc., Natick, MA, USA.
- [32] C. Kunick, K. Lauenroth, K. Wieking, X. Xie, C. Schultz, R. Gussio, D. Zaharevitz, M. Leost, L. Meijer, A. Weber, F.S. Jørgensen, T. Lemcke, *J. Med. Chem.* 47 (2004) 22–36.