Journal of Computer-Aided Molecular Design **17:** 475–493, 2003.
© 2003 *Kluwer Academic Publishers. Printed in the Netherlands.*

3D-QSAR and docking studies on 4-anilinoquinazoline and 4-anilinoquinoline epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors

Haregewein Assefa, Shantaram Kamath & John K. Buolamwini*
Department of Pharmaceutical Sciences, College of Pharmacy, University of Tennessee Health Sciences Center, 847 Monroe Avenue Suite 327, Memphis, TN 38163, U.S.A.

Received 12 February 2003; Accepted in revised form 2 June 2003

Key words: 3D-QSAR, CoMFA, CoMSIA, EGFR, receptor-guided consensus dynamics, tyrosine kinase

Summary

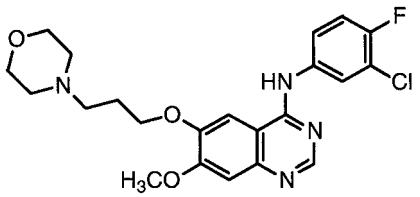
The overexpression and/or mutation of the epidermal growth factor receptor (EGFR) tyrosine kinase has been observed in many human solid tumors, and is under intense investigation as a novel anticancer molecular target. Comparative 3D-QSAR analyses using different alignments were undertaken employing comparative molecular field analysis (CoMFA) and comparative molecular similarity analysis (CoMSIA) for 122 anilinoquinazoline and 50 anilinoquinoline inhibitors of EGFR kinase. The SYBYL multifit alignment rule was applied to three different conformational templates, two obtained from a MacroModel Monte Carlo conformational search, and one from the bound conformation of erlotinib in complex with EGFR in the X-ray crystal structure. In addition, a flexible ligand docking alignment obtained with the GOLD docking program, and a novel flexible receptor-guided consensus dynamics alignment obtained with the DISCOVER program in the INSIGHTII modeling package were also investigated. 3D-QSAR models with q^2 values up to 0.70 and r^2 values up to 0.97 were obtained. Among the 4-anilinoquinazoline set, the q^2 values were similar, but the ability of the different conformational models to predict the activities of an external test set varied considerably. In this regard, the model derived using the X-ray crystallographically determined bioactive conformation of erlotinib afforded the best predictive model. Electrostatic, hydrophobic and H-bond donor descriptors contributed the most to the QSAR models of the 4-anilinoquinazolines, whereas electrostatic, hydrophobic and H-bond acceptor descriptors contributed the most to the 4-anilinoquinoline QSAR, particularly the H-bond acceptor descriptor. A novel receptor-guided consensus dynamics alignment has also been introduced for 3D-QSAR studies. This new alignment method may incorporate to some extent ligand-receptor induced fit effects into 3D-QSAR models.

Introduction

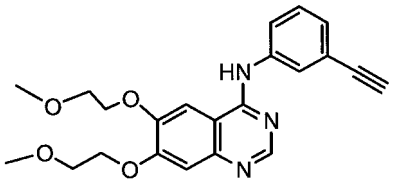
The epidermal growth factor receptor (EGFR) or ErbB1 is the prototype of the Type 1 family of integral membrane receptor tyrosine kinases (the others being HER-2 or ErbB2, HER-3 or ErbB3 and HER-4 or ErbB4), which is involved in mitogenic signal transduction [1]. Like most growth factor receptor kinases, EGFR is activated by binding of its cognate ligands, epidermal growth factor (EGF) or transforming growth factor alpha (TGF-α) to the extracellular domain causing receptor dimerization, which leads to activation of the intracellular kinase domain. EGFR is frequently overexpressed, amplified and/or mutated in many human solid tumors including breast, ovarian, non-small cell lung, and squamous cell cancers [1, 2], and has been shown to affect proliferation, angiogenesis and cancer metastasis. As a result, it is being pursued intensely as a novel molecular target for cancer selective therapy [3–5]. Success in small molecule drug discovery against EGFR as an anticancer target has come from selective inhibition of its

*To whom correspondence should be addressed. E-mail: jbuolamwini@utmem.edu

kinase activity using compounds that compete against ATP binding at the catalytic site [3]. Many compounds, particularly 4-anilinoquinazolines [5–9] have been shown to selectively inhibit EGFR kinase activity with high potency. The cancer therapeutic potential of the 4-anilinoquinazolines is underscored by the advancement of ZD 1839 (**1**, Iressa) [10, 11] and erlotinib (**2**, Tarceva) [12, 13] to Phase III clinical trials [14].



1



2

To date, there has been very little application of 3D-QSAR to kinase inhibitors. Although the X-ray structure of EGFR in complex with an inhibitor has been published [13], which will facilitate structure-based design, 3D-QSAR is important to allow more accurate prediction of activities and binding affinities, as current scoring functions have not attained the level of accuracy to allow their general applicability in predicting biological activity or binding affinity from docking energies. In this study, we have used the X-ray crystal structure in conjunction with three-dimensional (3D) quantitative structure–activity relationship (QSAR) studies, comparative molecular field analysis (CoMFA) [14] and comparative molecular similarity indices analysis (CoMSIA) [15] 3D-QSAR methods to analyze two related series of EGFR tyrosine kinase inhibitors, 4-anilinoquinazolines and 4-anilinoquinolines to gain quantitative insights into the influence of steric, electrostatic, hydrophobic and H-bonding effects on their interaction with the kinase, and to establish predictive models for drug design.

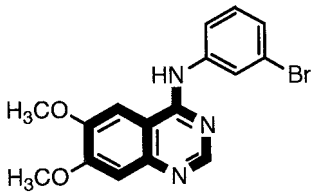


Figure 1. Atoms used in 'Multifit' alignments; shown in bold face.

Methods

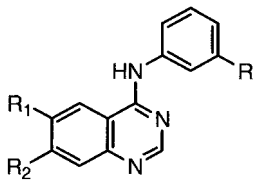
Dataset

The 4-anilinoquinazoline set and their biological data (Tables 1 and 3) were taken from literature reports [6–9], while the 4-anilinoquinolines were taken from the report by Wissner et al. [16] (Table 2). The structures were built in SYBYL 6.8 (Tripos Associates Inc.) on a Silicon Graphics Octane (R12000) workstation. Energy minimizations were performed using the Tripos force field [17] with a distance-dependent dielectric and the Powell conjugate gradient algorithm, with a convergence criterion of 0.01 Kcal/(mole*A).

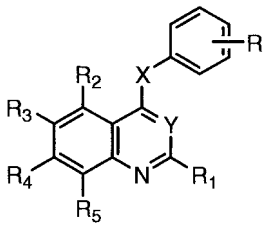
Multifit alignments

Two different molecular alignments based on MacroModel Monte Carlo conformational search-derived templates, as well as a third alignment based on the X-ray crystal structure of compound **2** (erlotinib) were derived using the 'Multifit' alignment rule in SYBYL. The two conformational search templates were derived as follows. A conformation of compound **3**, one of the very active quinazoline compounds, was constructed in SYBYL, its structure minimized, and then used as a starting point for a Monte Carlo conformational search employing MacroModel version 7.0 (Schrodinger, Inc.) running on a Silicon Graphics Octane 2 workstation. The conformational search was carried out using the MMFF94s force field in MacroModel for 5000 iterations using water as solvent. The global minimum conformation, designated Conf1, and the next lowest energy conformation from the global minimum, 2 kJ/mole energy difference, designated Conf2 were used as templates to construct the other molecules. The third set of molecules was constructed based on the recently reported X-ray crystal structure of erlotinib (**2**) bound to the ATP site of the kinase domain of EGFR (PDB code: 1M17). This conformational set is designated Conf3. The reference atoms used for the 'Multifit' alignments were the 10 heavy atoms of the quinazoline or quinoline ring system

Table 1. Structures and inhibitory activities of 4-anilinoquinazoline derivatives.



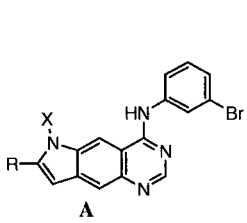
Compd	R_1	R_2	R	pIC_{50}	compd	R_1	R_2	R	pIC_{50}
3	OMe	OMe	Br	10.60	**27**	H	NO2	Br	6.00
4	H	H	H	6.46	**28**	H	NO_2	I	6.27
5	H	H	F	7.25	**29**	OMe	OMe	H	7.54
6	H	H	Cl	7.64	**30**	OMe	OMe	F	8.42
7	H	H	Br	7.57	**31**	OMe	OMe	Cl	9.51
8	H	H	I	7.10	**32**	OMe	OMe	I	9.05
9	H	H	CF_3	6.24	**33**	OMe	OMe	CF_3	9.62
10	OMe	H	H	7.26	**34**	NMe_2	H	Br	7.08
11	OMe	H	Br	7.52	**35**	$NHCO_2Me$	H	Br	7.92
12	NH_2	H	H	6.11	**36**	H	OH	Br	8.33
13	NH_2	H	CF_3	6.24	**37**	H	NHAc	Br	7.40
14	NH_2	H	Br	9.11	**38**	H	NHMe	Br	8.15
15	NO_2	H	H	5.30	**39**	H	NHEt	Br	7.92
16	H	OMe	H	6.92	**40**	H	NMe_2	Br	7.96
17	H	OMe	Br	8.00	**41**	NH_2	NHMe	Br	9.16
18	H	NH_2	H	7.00	**42**	NH_2	OMe	Br	8.42
19	H	NH_2	F	8.70	**43**	NH_2	Cl	Br	8.19
20	H	NH_2	Cl	9.60	**44**	NO_2	NH_2	Br	7.28
21	H	NH_2	Br	10.00	**45**	NO_2	NHAc	Br	7.55
22	H	NH_2	I	9.46	**46**	NO_2	OMe	Br	7.82
23	H	NH_2	CF3	8.48	**47**	NO_2	Cl	Br	7.60
24	H	NO_2	H	4.92	**48**	OH	OH	Br	9.77
25	H	NO_2	F	5.21	**49**	OEt	OEt	Br	11.22
26	H	NO_2	Cl	6.09	**50**	OPr	OPr	Br	9.77



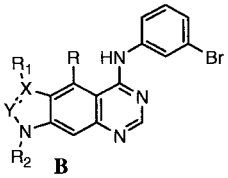
Compd	R_1	R_2	R_3	R_4	R_5	R	X	Y	pIC_{50}
51	H	OMe	OMe	H	H	3′-Br	NH	N	5.86
52	NH_2	H	OMe	OMe	H	3′-Br	NH	N	6.33
53	H	H	OMe	OMe	H	3′-Br	NMe	N	6.82
54	H	OMe	OMe	OMe	H	3′-Br	NH	N	9.17
55	H	H	OMe	OMe	H	2′-Br	NH	N	6.89
56	H	H	OMe	OMe	H	3′,5′-diBr	NH	N	6.95
57	H	H	H	H	H	H	$NH(CH2)_2$	N	5.39
58	H	H	H	H	H	H	$NHCH_2$	N	6.49

Table 1. Continued.

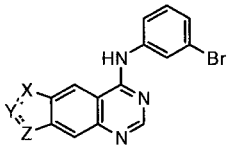
Compd	R_1	R_2	R_3	R_4	R_5	R	X	Y	pIC_{50}
59	H	H	H	H	H	4′-OMe	$NHCH_2$	N	5.00
60	H	H	H	H	H	H	NMe	N	4.00
61	H	H	H	H	H	3′-OMe	NH	N	6.07
62	H	H	H	H	H	3′-Me	NH	N	6.04
63	H	H	H	H	H	3′-Br	NH	CH	5.26
64	H	NO_2	H	H	H	H	$NHCH_2$	N	5.10
65	H	OMe	H	H	H	H	$NHCH_2$	N	5.31
66	H	H	OMe	H	H	H	$NHCH_2$	N	6.70
67	H	H	H	NO_2	H	H	$NHCH_2$	N	5.23
68	H	H	H	OMe	H	H	$NHCH_2$	N	7.24
69	H	H	OMe	OH	H	H	$NHCH_2$	N	6.23
70	H	H	OH	OMe	H	H	$NHCH_2$	N	7.25
71	H	NO_2	H	H	H	3′-Br	NH	N	6.45
72	H	OMe	H	H	H	3′-Br	NH	N	6.86
73	H	H	H	H	NH_2	3′-Br	NH	N	6.98
74	H	H	H	H	OMe	3′-Br	NH	N	6.01



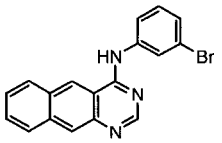
A



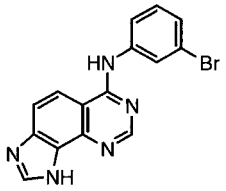
B



C



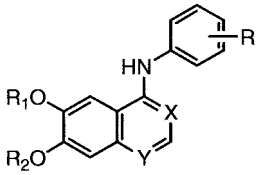
97



98

Compd	Type	R	R_1	R_2	X	Y	Z	pIC_{50}
75	A	Me			H			9.54
76	A	H			Me			11.00
77	A	H			$(CH_2)_2NMe_2$			8.88
78	B	H	-	Me	N	CH		10.60
79	B	H	-	$(CH_2)_2NMe_2$	N	CH		7.66
80	B	H	H	$CH_2CH(OH)CH_2OH$	CH	N		7.92
81	B	H	H	$(CH_2)_2NMe_2$	CH	N		7.40
82	B	H	H	$(CH_2)_2$Nmorpholide	CH	N		8.43
83	B	H	H	CH_2COOH	CH	N		7.28
84	B	H	H	Me	CH	CH		9.10
85	B	H	H	$CH_2CH(OH)CH_2OH$	CH	CH		8.80
86	B	H	H	$(CH_2)_2$Nmorpholide	CH	CH		8.43
87	B	H	H	$(CH_2)_3$Nmorpholide	CH	CH		8.06
88	B	H	H	CH_2COOH	CH	CH		8.29
89	B	H	$CH_2N(CH_2CH_2OH)_2$	H	CH	CH		8.46
90	B	H	$(CH_2)_2$Nmorpholide	H	CH	CH		8.32
91	B	H	$CH_2N(Me)CH_2COOMe$	H	CH	CH		8.47
92	B	H	$CH_2N(Me)CH_2COOH$	H	CH	CH		9.14
93	B	H	H	H	C	N		9.36
94	C				NH	N	N	8.39
95	C				S	CH	N	7.36
96	C				NH	N	CH	9.47
97								11.52
98								6.57

Table 2. Structures and inhibitory activities of 4-anilinoquinoline derivatives.

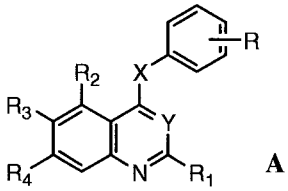


Compd	X	Y	R_1	R_2	R	pIC_{50}
99	C-CO_2Et	N	Me	Me	3′-Br	4.09
100	C-CN	N	Et	Et	3′-Br	6.38
101	C-CN	N	Et	Et	3′-Cl, 4′-F	6.01
102	C-CN	N	Me	Me	3′-Br	6.72
103	C-CN	N	Me	Me	3′-Cl	6.74
104	C-CN	N	Me	Me	3′-CF_3	5.84
105	C-CN	N	Me	Me	3′,4′-di-OMe	4.30
106	C-CN	N	Me	Me	3′-F	6.13
107	C-CN	N	Me	Me	3′-CN	4.85
108	C-CN	N	Me	Me	4′-F	4.32
109	C-CN	N	Me	Me	3′-COMe	5.13
110	C-CN	N	Me	Me	3′-NH_2	6.09
111	C-CN	N	Me	Me	3′-$N(Me)_2$	6.07
112	C-CN	N	Me	Me	3′-NO_2	6.06
113	C-CN	N	Me	Me	3′-Cl, 4′-F	6.27
114	C-CN	N	Me	Me	4′-Cl, 2′-F	4.86
115	C-CN	N	Me	Me	3′-OH	5.20
116	C-CN	N	Me	Me	4′-Me	5.56
117	C-CN	N	Me	Me	3′-$CONH_2$	5.11
118	C-CN	N	Me	Me	3′-Br, 4′-Me	5.75
119	C-CN	N	Me	Me	3′-Cl, 4′-OH	5.04
120	C-CN	N	Me	Me	3′,5′-di-Cl, 4′-OH	5.50
121	C-CN	N	Me	Me	5′-Cl, 2′-OH	4.77
122	C-CN	N	Me	Me	3′-SMe	6.12
123	C-CN	N	Me	Me	3′-Cl, 4′-Me	5.83
124	C-CN	N	Me	Me	2′-F, 4′-Br	4.04
125	C-CN	N	Me	Me	4′-$CH(Me)_2$	6.64
126	C-CN	N	Me	Me	2′-$CH(Me)_2$	6.44
127	C-CN	N	Me	Me	3′-$CH(Me)_2$	4.21
128	C-CN	N	Me	Me	4′-Br	5.18
129	C-CN	N	Me	Me	2′-Br	5.38
130	C-CN	N	Me	Me	3′-CF_3, 4′-F	6.80
131	C-CN	N	Me	Me	2′-Me, 3′-Br	6.08
132	C-CN	N	Me	Me	3′-Me, 4′-Br	6.21
133	C-CN	N	Me	Me	3′-N_3	5.66
134	C-CN	N	Me	Me	3′-NHCOMe	4.16
135	C-CN	N	Et	Me	3′-Br	4.67
136	C-CN	N	Me	Et	3′-Br	4.48
137	C-CN	N	-CH_2CH_2-		3′-Br	3.98
138	C-CH_2OH	N	Me	Me	3′-Br	3.99
139	C-CHO	N	Me	Me	3′-Br	5.47
140	C-CO_2H	N	Me	Me	3′-Br	4.47
141	C-$CONH_2$	N	Me	Me	3′-Br	4.82
142	C-CN	N	H	H	3′-Br	6.46
143	C-CN	N	$MeOCH_2$	$MeOCH_2$	3′-Br	5.17

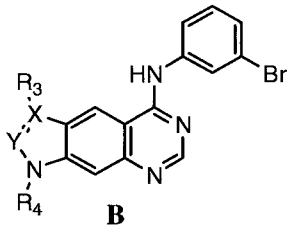
Table 2. Continued.

Compd	X	Y	R_1	R_2	R	pIC_{50}
144	C-CN	N	$MeO(CH_2)_2$	$MeO(CH_2)_2$	3′-Br	5.96
145	C-CN	N	$-CH_2-$		3′-Br	5.17
146	C-CN	N	$-(CH_2)_3-$		3′-Br	5.29
147	N	C-CN	Me	Me	3′-Br	4.15
148	N	C-CN	Me	Me	3′-Cl	4.05

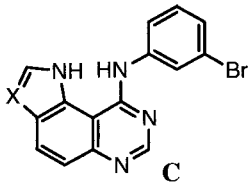
Table 3. Structures and EGFR kinase inhibitory activity of compounds in the test set.



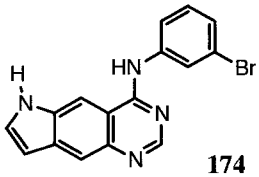
A



B



C



174

Compd	Type	R_1	R_2	R_3	R_4	R	X	Y	pIC_{50}
149	A	H	H	NO_2	H	3′-Br	NH	N	6.05
150	A	H	H	NHMe	H	3′-Br	NH	N	8.40
151	A	H	H	NH_2	NH_2	3′-Br	NH	N	9.92
152	A	H	H	NH_2	NMe_2	3′-Br	NH	N	6.80
153	A	H	H	NO_2	NHMe	3′-Br	NH	N	7.17
154	A	H	H	NO_2	NMe_2	3′-Br	NH	N	5.70
155	A	H	H	$-OCH_2O-$		3′-Br	NH	N	7.82
156	A	H	H	OBu	OBu	3′-Br	NH	N	6.98
157	A	H	H	OMe	OMe	4′-Br	NH	N	9.02
158	A	H	H	OMe	OMe	3′,4′-diBr	NH	N	10.14
159	A	H	H	H	H	4′-Cl	$NHCH_2$	N	5.15
160	A	H	H	H	H	3′-Br	O	N	6.12
161	A	H	H	NO_2	H	H	$NHCH_2$	N	5.00
162	A	H	H	NH_2	H	H	$NHCH_2$	N	5.85
163	A	H	H	OMe	OMe	H	$NHCH_2$	N	8.00
164	A	H	NH_2	H	H	3′-Br	NH	N	6.36
165	B			H	Me		C	N	9.43
166	B			H	$(CH_2)_2NMe_2$		C	CH	7.39
167	B			H	$(CH_2)_3NMe_2$		C	CH	7.68
168	B			$CH_2N(Me)(CH_2)_2NMe_2$	H		C	CH	8.12
169	B			–	$(CH_2)_2NMe_2$		N	CH	6.69
170	B			H	H		C	CH	9.36
171	B			–	–		N	CHCH	8.77
172	C						N		7.54
173	C						CH		8.91
174									11.10
2									8.70

Table 4. PLS statistics of CoMFA and CoMSIA 3D QSAR models of Conf1, Conf2, and Conf3.

Dataset	PLS statistics	Conf1		Conf2		Conf3	
		CoMFA	CoMSIA	CoMFA	CoMSIA	CoMFA	CoMSIA
Quinazolines	q^2	0.622	0.640	0.630	0.651	0.643	0.643
	PRESS	0.954	0.928	0.985	0.923	0.975	0.927
	r^2	0.863	0.812	0.851	0.848	0.880	0.782
	s	0.575	0.670	0.625	0.609	0.565	0.724
	F	88.09	73.616	97.158	78.057	103.143	77.225
	PLS components	6	5	5	6	6	4
	Contribution						
	steric	0.530	0.075	0.531	0.069	0.519	0.106
	electrostatic	0.470	0.341	0.469	0.331	0.481	0.306
	hydrophobic		0.219		0.231		0.222
	donor		0.257		0.244		0.268
	acceptor		0.108		0.126		0.097
Quinolines	q^2	0.517	0.602	0.700	0.587	0.632	0.666
	PRESS	0.656	0.601	0.505	0.584	0.555	0.543
	r^2	0.931	0.937	0.935	0.939	0.968	0.953
	s	0.248	0.239	0.236	0.224	0.163	0.205
	F	67.426	74.273	71.581	77.297	151.785	100.544
	PLS components	5	6	6	6	6	6
	Contribution						
	steric	0.543	0.082	0.382	0.067	0.438	0.079
	electrostatic	0.457	0.191	0.618	0.258	0.562	0.193
	hydrophobic		0.298		0.295		0.249
	donor		0.105		0.082		0.166
	acceptor		0.324		0.298		0.313

and the nitrogen or oxygen atom attached to the 4-position of the quinoline or the quinazoline ring (see Figure 1). Partial atomic charges were calculated using the semi-empirical program MOPAC 6.0 in SYBYL, and applying the AM1 Hamiltonian [18].

Docking and receptor-guided consensus dynamics alignments

The docking of the quinoline derivative inhibitors, compounds **99**–**148** into the ATP site was carried out with the program GOLD (v1.2) [19] from the Cambridge Crystallographic Data Center, UK. Both the free acid and the ionized states of compound **140** were used. The binding pocket was defined as residues lying within 15 Å of Thr 830.

The DISCOVER module in INSIGHTII 2000 (Accelrys, San Diego, CA) was used to carry out the flexible receptor-guided consensus dynamics simulation [20] using the docking modes obtained with GOLD as the starting structures. Consensus restraint was applied to the heavy atoms of the 4-anilinoquinoline moiety of the inhibitors within the receptor. Compounds **108** and **125** were excluded, as we could not obtain a binding mode common to the other compounds. Compound **133** had to be eliminated from the consensus alignment procedure, as the program did not have parameters for the azido group. The side chains of the ATP site amino acid residues within 15 Å of Thr 830 were free to move during the dynamics run while the rest of the atoms of the receptor were held rigid. The intermolecular interaction between the inhibitors was switched off, but each of the inhibitors interacted with the receptor during the dynamics. A short dynamics run of 5 ps was used to achieve the alignments. A temperature of 300 K and an integration time step of 1 fs were used. At the end of the dynamics run, the structures were minimized by 1000 cycles of steepest descents, followed by 2000 cycles of conjugate gradients. The consensus alignment with the lowest

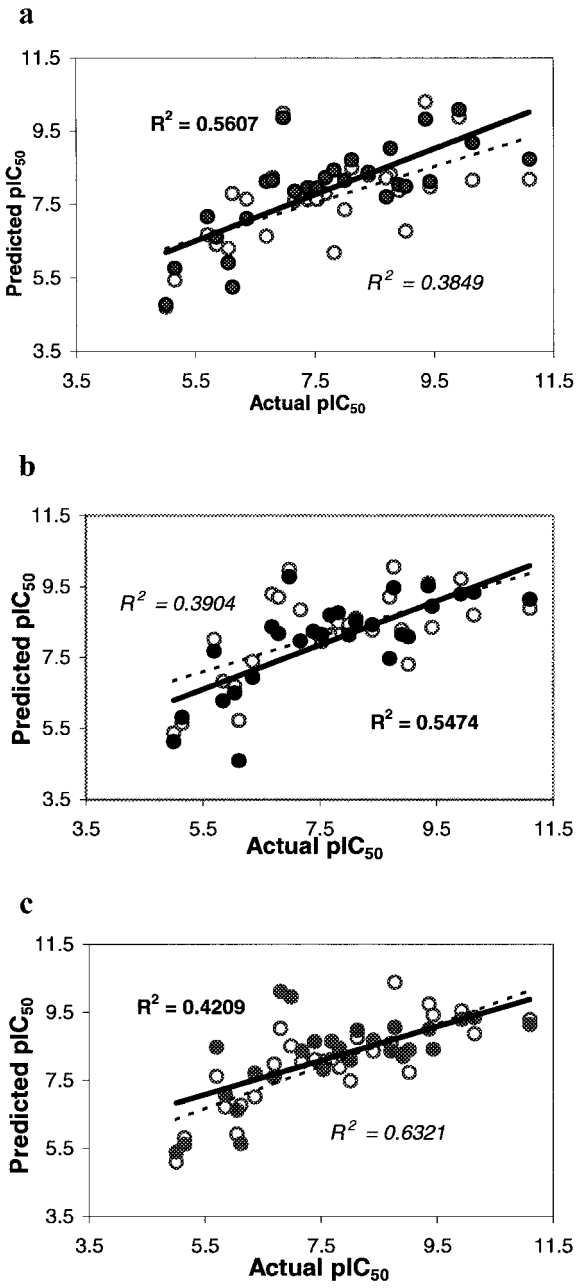


Figure 2. External test set prediction curves of the quinazoline derivatives. Open circles and dotted lines represent predictions by CoMFA models, while predictions by CoMSIA models are shown by filled circles and solid lines. R^2 values for the CoMFA models are italicized while the R^2 values of the CoMSIA models are in bold face lettering. (a) Predictions by Conf1 models; (b) predictions by Conf2 models; (c) predictions by Conf3 models.

total energy was used for CoMFA and CoMSIA model generation.

CoMFA and CoMSIA 3D QSAR Models

In deriving the CoMFA and CoMSIA descriptor fields, a three-dimensional cubic lattice with grid spacing of 2 Å, and extending 4 Å units beyond the aligned molecules in all directions was created to encompass the aligned molecules. CoMFA descriptors were calculated using an sp^3 carbon probe atom with a van der Waals radius of 1.52 Å and a charge of +1.0 to generate steric (Lennard-Jones 6-12 potential) field energies and electrostatic (Coulombic potential) fields with a distance-dependent dielectric at each lattice point. An energy cut-off of 30 kcal/mole was used, and the CoMFA steric and electrostatic fields generated in this way were scaled by the CoMFA-STD method.

CoMSIA similarity indices descriptors were derived according to Klebe et al. [15] with the same lattice box as was used for the CoMFA calculations, with a grid spacing of 2 Å employing an sp^3 carbon probe atom with a van der Waals radius of 1.4 Å and a charge of +1.0 as implemented in SYBYL. CoMSIA similarity indices ($A_{F,k}$) between a molecule j with atoms i at a grid point q and the probe atom are calculated by Eq. 1 as follows:

$$A^q_{F,k}(j) = -\Sigma\omega_{\text{probe},k}\omega_{ik}e^{-\alpha r\, iq^2}. \quad (1)$$

The five physicochemical properties or descriptors (k), steric, electrostatic, hydrophobic, and hydrogen bond donor and hydrogen bond acceptor, were evaluated using the probe atom. The descriptor k of atom i is ω_{ik}, while the descriptor of the probe atom is ω_{probe}. A Gaussian-type distance dependence was used between the grid point q and each atom i in the molecule. The value of the attenuation factor (α) was set to 0.3. In CoMSIA, the steric indices relate to the third power of the atomic radii, electrostatic descriptors are derived from atomic partial charges, hydrophobic fields from atom-based parameters developed by Viswanadhan et al. [21], and hydrogen bond donor and acceptors from a rule-based method derived from experimental values [22].

The CoMFA and CoMSIA descriptors were used as independent variables, and pIC_{50} values as dependent variable, in partial least squares (PLS) regression analyses to derive 3D QSAR models. The predictive value of the models was evaluated first by leave-one-out cross validation. The cross-validated coefficient,

q^2, was calculated using Eq. 2:

$$q^2 = 1 - \frac{\Sigma(Y_{\text{predicted}-\text{Y}_{\text{observed}}})^2}{\Sigma(Y_{\text{observed}} - Y_{\text{mean}})^2}, \quad (2)$$

where Y_{pred}, Y_{actual} and Y_{mean} are predicted, actual and mean values of the target property (pIC_{50}), respectively. $\Sigma(\text{Y}_{\text{pred}} - \text{Y}_{\text{actual}})^2$ is the predictive sum of squares (PRESS). To maintain the optimum number of PLS components and minimize the tendency to over fit the data, the number of components corresponding to the lowest PRESS value was used for deriving the final PLS regression models [23]. In addition to the q^2 and the corresponding PRESS and number of components, the conventional correlation coefficient r^2 and its standard error *s* were also computed. CoMFA and CoMSIA coefficient maps were generated by interpolation of the pair-wise products between the PLS coefficients and the standard deviations of the corresponding CoMFA or CoMSIA descriptor values.

Results and discussion

CoMFA and CoMSIA modeling of the Quinazoline dataset

The X-ray crystal structure of EGFR kinase domain in complex with the inhibitor erlotinib (**2**) was released recently, after we had already conducted QSAR studies on these molecules and presented the results at a meeting [24]. Prior to the release of the crystal structure, two MacroModel conformational search-derived alignment sets (Conf1 and Conf2) were used to derive 3D-QSAR models. We have now included a third conformational set, Conf3, based on the crystal structure of erlotinib in the complex with EGFR kinase domain. Initial CoMFA PLS analyses with all 96 anilinoquinazolines in the training set resulted in q^2 values of 0.461–0.487. The high standard error of prediction (PRESS: 1.121–1.162) prompted the omission of five possible outliers in the final models. The CoMFA models obtained with the remaining 91 compounds afforded q^2 values of 0.622, 0.630, 0.643 for Conf1, Conf2 and Conf3, respectively (see Table 4). To make sure that these results were not fortuitous, we performed PLS analysis runs with randomized (scrambled) pIC_{50} values, as well as group cross-validations (see Table 5). In the randomization exercise, the vast majority of the analyses resulted in negative q^2 values, confirming that the QSAR models obtained with the actual data did not arise by chance. The deviations in the q^2 values from the group cross-validation results were minimal, indicating robustness of the models as well. Compound **52** was a common outlier in all models derived with all the three conformational alignment sets for the quinazolines. This compound (**52**) is also the only inhibitor with a 2-position substituent in the data set, which might account for its outlier status.

CoMSIA PLS analysis of the initial 4-anilinoquinazoline training set of 96 compounds resulted in q^2 values of 0.552–0.562. However, as was the case with the CoMFA models, standard errors of prediction were high, necessitating the removal of some outliers. To obtain comparable models to those of the CoMFA analyses, five possible outliers were eliminated here as well, resulting in q^2 values of 0.640, 0.651 and 0.644 for Conf1, Conf2 and Conf3 models, respectively. In the CoMSIA models, compounds **51**, **52**, and **56** were common outliers. It is worth noting that compound **52** was also a common outlier in the CoMFA models.

CoMFA and CoMSIA modeling of the Quinoline dataset

Initial PLS CoMFA and CoMSIA analyses with all 50 4-anilinoquinolines and compound **3**, using the 'Multifit' alignments afforded very low q^2 values. Models with reasonable q^2 values were obtained after removing 14 possible outliers. QSAR analysis using the 4-anilinoquinoline dataset was important, as this series explored the effect of substitution at the 1- and 3-positions that were not explored in the quinazolines set. CoMFA PLS analyses of the training sets afforded QSAR models with q^2 values of 0.517–0.700. The common outliers in the CoMFA models for all conformations were compounds **105**, **107**, **125**, **135**, and **136**. CoMSIA PLS analysis resulted in models with q^2 values of 0.587–0.666. In the CoMSIA models, eight compounds, **3**, **105**, **108**, **112**, **127**, **135**, **136**, and **132** were common outliers.

Model validation with an external test set

The QSAR models obtained with the 4-anilinoquinazoline dataset were further validated using an external test set of 27 compounds including erlotinib. Among the Conf1 QSAR models, the CoMSIA model performed better by predicting the inhibitory activity of 20 compounds of the test set within ± 1 pIC_{50} unit. The CoMFA model on the other hand predicted 17 compounds with a similar accuracy. This difference in

Table 5. Results of group cross-validation (GV) and randomization exercises.

Dataset	PLS statistics	Conf1		Conf2		Crys temp	
		CoMFA	CoMSIA	CoMFA	CoMSIA	CoMFA	CoMSIA
Quinazolines	GVq^2 (STDEV)	0.612	0.636	0.633	0.645	0.640	0.636
		(0.010)	(0.018)	(0.016)	(0.015)	(0.014)	(0.016)
	Rand q^2	−0.139	−0.143	−0.143	−0.133	−0.107	−0.155
Quinolines	GVq^2 (STDEV)	0.500	0.581	0.677	0.564	0.589	0.623
		(0.039)	(0.033)	(0.023)	(0.035)	(0.052)	(0.044)
	Rand q^2	−0.177	−0.240	−0.196	−0.126	−0.267	−0.224

Table 6. Residuals of the predictions of the test set by the CoMFA and CoMSIA models derived from the anilinoquinazolines.

Compd	pIC_{50}	Residuals					
		Conf1		Conf2		Conf3	
		CoMFA	CoMSIA	CoMFA	CoMSIA	CoMFA	CoMSIA
2 (Erlotinib)	8.7	−0.478	−0.989	0.515	−1.226	−0.115	−0.337
149	6.05	0.258	−0.138	0.640	0.444	−0.130	0.574
150	8.40	−0.091	−0.023	−0.125	0.031	−0.042	0.282
151	9.92	−0.039	0.169	−0.192	−0.617	−0.366	−0.627
152	6.80	1.429	1.355	2.407	1.366	2.227	3.311
153	7.17	0.430	0.689	1.678	0.806	0.868	1.197
154	5.70	0.982	1.467	2.311	1.982	1.924	2.777
155	7.82	−1.632	0.619	0.668	0.948	0.064	0.635
156	6.98	3.011	2.885	3.005	2.804	1.535	2.971
157	9.02	−2.244	−1.023	−1.713	−0.937	−1.293	−0.626
158	10.14	−1.973	−0.958	1.438	−0.797	−1.264	−0.794
159	5.15	0.283	0.616	0.493	0.661	0.667	0.469
160	6.12	1.677	−0.871	−0.396	−1.52	0.65	−0.484
161	5.00	−0.293	−0.231	0.368	0.13	0.105	0.396
162	5.85	0.566	0.759	0.980	0.432	0.86	1.238
163	8.00	−0.638	0.156	0.447	0.135	−0.515	0.084
164	6.36	1.278	0.762	1.036	0.586	0.654	1.351
165	9.43	−1.459	−1.307	−1.079	−0.481	−0.013	−1.011
166	7.39	0.236	0.571	0.858	0.845	0.731	1.249
167	7.68	0.118	0.554	0.457	1.023	0.44	0.965
168	8.12	0.380	0.602	0.498	0.393	0.634	0.847
169	6.69	−0.049	1.44	2.608	1.673	1.292	0.908
170	9.36	0.942	0.464	0.247	0.175	0.38	−0.357
171	8.77	−0.445	0.269	1.292	0.701	1.61	0.294
172	7.54	0.092	0.412	0.411	0.601	0.532	0.281
173	8.91	−1.027	−0.874	−0.623	−0.754	−0.629	−0.706
174	11.10	−2.917	−2.364	−2.213	−1.951	−1.813	−1.959

Table 7. PLS statistics of the QSAR models for quinolines using docking and consensus dynamics alignments.

PLS statistics	Docked		Consensus	
	CoMFA	CoMSIA	CoMFA	CoMSIA
q^2	0.588	0.538	0.598	0.583
PRESS	0.557	0.584	0.535	0.586
r^2	0.957	0.967	0.931	0.923
s	0.187	0.159	0.225	0.252
F	99.15	145.71	67.63	57.74
PLS components	4	5	5	6
Group validation				
Average q^2	0.563(0.054)	0.507(0.030)	0.565(0.051)	0.545(0.052)
Bootstrapping				
Average r^2	0.980(0.003)	0.996(0.009)	0.957(0.005)	0.942(0.007)
Average s	0.011(0.002)	0.002(0.001)	0.012(0.003)	0.019(0.004)
Contribution				
steric	0.510	0.071	0.449	0.113
electrostatic	0.490	0.241	0.551	0.222
hydrophobic	–	0.208	–	0.248
donor	–	0.199	–	0.123
acceptor	–	0.281	–	0.293

the predictive abilities is apparent from the prediction curves in Figure 2a ($R^2 = 0.561$ vs. $R^2 = 0.385$). In the Conf2 models also, CoMSIA was better at predicting the external test set ($R^2 = 0.547$ vs. $R^2 = 0.390$, Figure 2b) than CoMFA, predicting the activity of 19 of the 27 compounds within ± 1 pIC_{50} unit. The Conf2 CoMFA model, on the other hand, predicted only 16 of the 27 compounds with similar precision. Unlike the Conf1 and Conf2 models, in the Conf3 models, CoMFA performed better at predicting the test set than CoMSIA, as reflected in the prediction curves in Figure 2c ($R^2 = 0.632$ vs. $R^2 = 0.421$). Three compounds of the test set, compounds **152**, **156** and **174**, were poorly predicted by all the models. The outlier status of these compounds appears to be related to the experimental data rather than the structure related.

PLS contour maps

CoMFA steric and electrostatic PLS coefficient contributions for the quinazoline set of compounds with Conf2 alignment are shown in Figure 3a. A yellow region near the hydrogen on the anilino nitrogen shows that a bulky substituent at this site decreases potency. This is consistent with the SAR indicating lower activities for compounds **53** ($pIC_{50} = 6.82$) and **60** ($pIC_{50} = 4.00$), in which a methyl group is introduced at this position, compared to their non-methylated analogs, compounds **3** ($pIC_{50} = 10.60$) and **4** ($pIC_{50} = 6.46$), respectively. The green contour near the 3′-position of the anilino ring indicates that generally steric bulk is favored at that site. However, the SAR of the anilinoquinazolines indicates that there is a limit as to how large the group can be, as shown by the following relative order of potencies: **29** < **30** < **32** < **31** < **3** (H < F < I < Cl < Br). The green contour around the methoxy group indicates that some degree of bulkiness is again permitted around that region while the appearance of the yellow contour near the green restricts the size of the group. The red contour near the 6- and 7-positions shows that negative electrostatic potential around this region increases potency. Compounds with electronegative groups such as an amine or hydroxyl group at the 6- and/or 7-position have shown higher potency compared to their unsubstituted analogs. The blue contour next to the red contour around the 6- or 7-position shows that electron deficient groups at this site might enhance potency.

The CoMSIA steric and electrostatic contour maps shown in Figure 3b, especially the steric plot, are

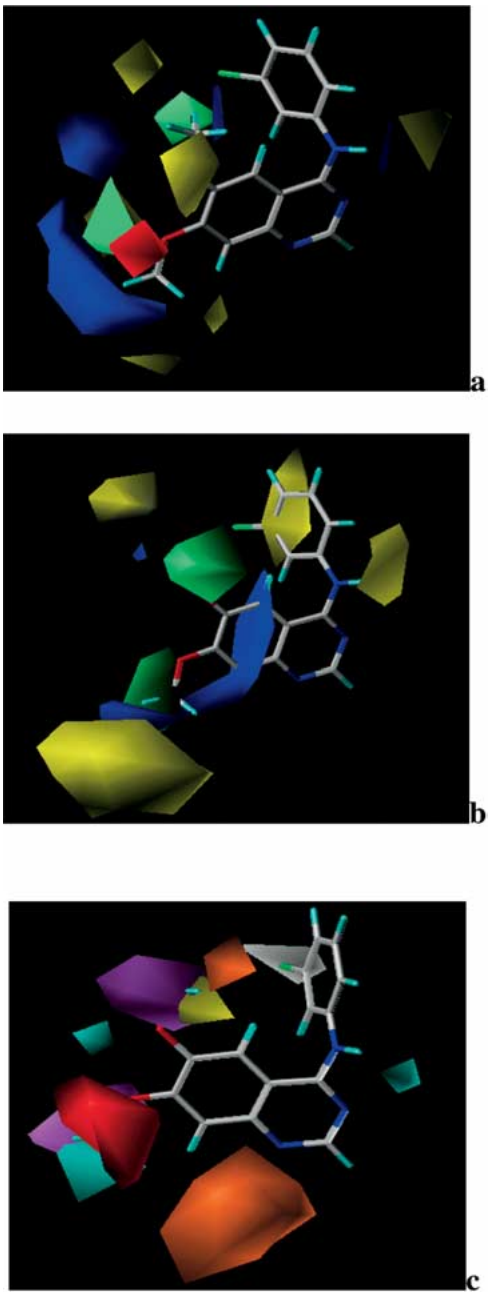


Figure 3. CoMFA and CoMSIA contour maps of the quinazoline training set generated for the Conf2 models. (a) CoMFA steric and electrostatic contour maps: green, bulky substituents favor activity; yellow, bulky substituents are detrimental to activity; red, negative electrostatic potential favors activity; blue, positive electrostatic potential enhances potency. (b) CoMSIA steric and electrostatic contour map: color designation the same as in Figure 3a. (c) CoMSIA hydrophobic and H-bonding contour maps: white, hydrophobic substituent favors activity; yellow, hydrophobic substituent disfavors activity; cyan, H-bond donor group favors activity; orange, H-bond donor disfavors activity; magenta, H-bond acceptor favors activity; red, H-bond acceptor disfavors activity.

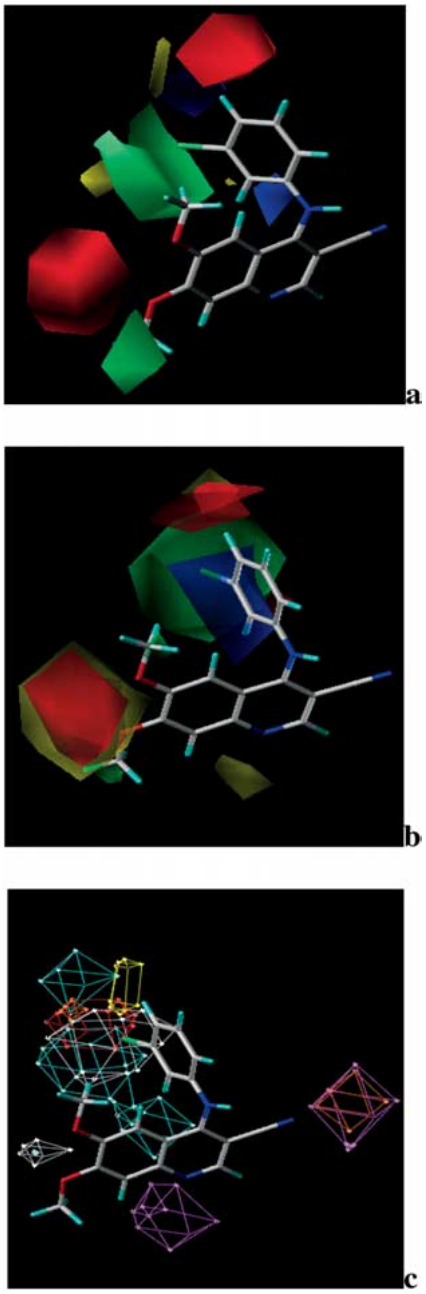


Figure 4. CoMFA and CoMSIA contour maps of the quinoline set with the Conf2 models. Color designations are the same as in Figures 3a–c. (a) CoMFA steric and electrostatic contour maps. (b) CoMSIA steric and electrostatic contour maps. (c) CoMSIA hydrophobic and H-bonding contour maps.

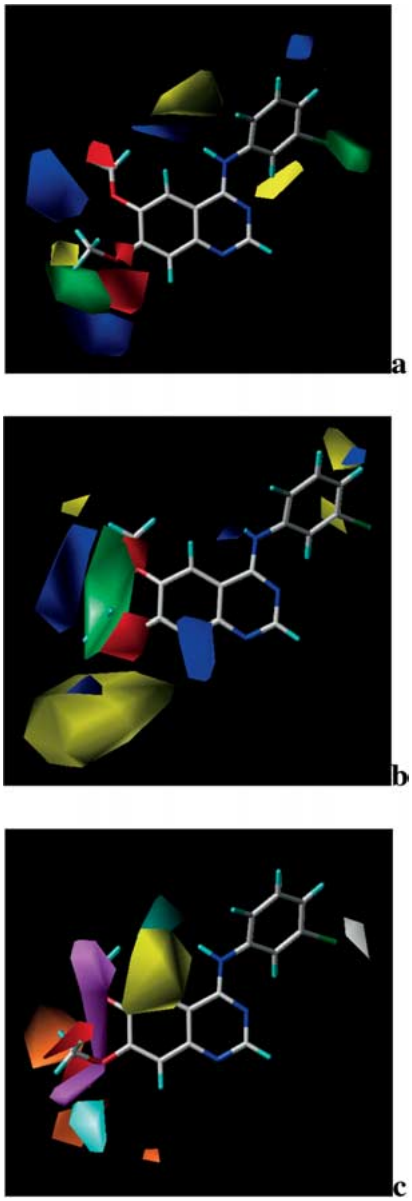


Figure 5. CoMFA and CoMSIA contour plots of the quinazoline compounds based on the conformation of erlotinib (**2**) in the X-ray structure of its complex with EGFR kinase domain (Conf3). Color designations are the same as in Figures 3a–c. (a) CoMFA steric and electrostatic contour plots. (b) CoMSIA steric and electrostatic contour plots. (c) CoMSIA hydrophobic and H-bonding contour plots.

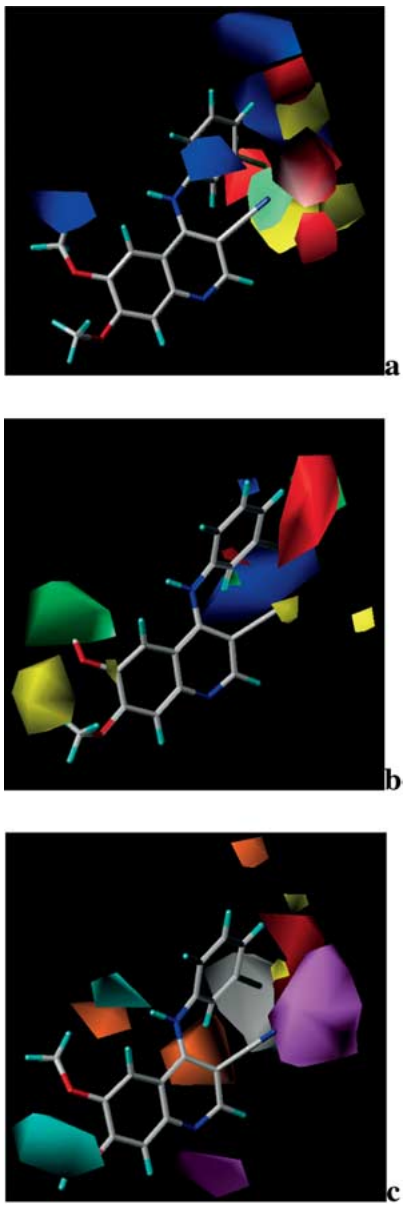


Figure 6. CoMFA and CoMSIA contour plots of the quinoline derivatives with the Conf3 models. Color designations are the same as in Figures 3a–c. (a) CoMFA steric and electrostatic contour plots. (b) CoMSIA steric and electrostatic contour plots. (c) CoMSIA hydrophobic and H-bonding contour plots.

similar to those of the CoMFA. The CoMSIA hydrophobic and H-bond PLS coefficient plots of the quinazolines are shown in Figure 3c. A hydrophobic group is favored at the 3′-position of the aniline-ring as shown by a white contour while the yellow contour near the 6-position of the quinazoline ring system shows that a hydrophobic group is disfavored at this

position. The cyan contour near the anilino NH shows that a H-bond donor is favored at this location. This is consistent with the SAR of the quinazolines, which demonstrates that methylation of the anilino NH results in a dramatic decrease in potency. Thus, the anilino NH might be involved in H-bonding at the binding site. As discussed later in the docking studies of the 4-anilinoquinolines, this NH might actually participate in H-bonding with Asp 831. The cyan contour near the methoxy substituents also shows that a H-bond donor is favored at this position. Examination of the SAR of the quinazoline derivatives shows that substitution of H-bond donor groups such as an amino or hydroxyl group at the 6- and/or 7-positions generally increases potency. The orange contours near the N-1 nitrogen and the 6-methoxy substituent show that a H-bond donor is disfavored at these positions. The orange contour near the N-1 position in Figure 3c appears to be artifactual since we did not have structural variation at this position. It is important to note that a similar contour was absent in the corresponding locality in the Conf3 model (see Figure 5c). CoMSIA hydrogen-bonding contours may not always all be consistently located [25], and there are efforts to refine the CoMSIA-type hydrogen-bond descriptors [26]. The results here show that alignment can influence even the locations of PLS contour maps. The H-bond contour maps of the Conf3 (X-ray structure) template are more consistent with the structural variations in the compounds than the Conf2 maps. This may have to do with the fact that the Conf3 template is better, as indicated by its superior performance with regard to predicting activities of the external test set.

The CoMFA and CoMSIA PLS contour maps for the Conf2 model of the quinoline dataset are depicted in Figure 4a–c. The green contour near the methoxy and around the 3′-position of the anilino group in both Figures 4a and 4b shows that a bulky group may favor inhibitory potency at this location. This is consistent with the information presented in Figures 3a and 3b. Both the CoMFA (Figure 4a) and CoMSIA (Figure 4b) maps are indicating that a negative potential near the 6,7-methoxy groups and the 4′-position of the aniline ring may enhance potency. The magenta near the 3-CN group and N-1 of the quinoline ring predicts that H-bond acceptor groups are favored in these regions. This is interesting in light of the model proposed by Wissner et al. [16]. They proposed that N-1 and N-3 of the quinazoline ring are H-bond acceptor points in the ligand, and attempted to validate this hypothesis by synthesizing the 3-cyano quinoline series. The X-ray crystallographic structure of EGFR in complex with a quinazoline inhibitor indicates that the N-1 nitrogen interacts with the Met769 NH through H-bonding. The N-3 nitrogen is also bridged by a water molecule to H-bond donor groups in the receptor, particularly Thr766. In the conformations of the quinoline series obtained from the GOLD docking exercise (*vide infra*), the 3-CN group occupies the position of this bridging water, which bridges with Thr766 rather than Thr830, which was proposed by the docking model of Wissner et al. [16].

Figures 5a–c show CoMFA and CoMSIA contour maps for the quinazoline training set based on Conf3. These contour maps give similar information to that of the Conf2 conformational set (Figure 3a–c). The red contours near the methoxy groups in Figures 5a and 5b indicate that high electron density favors activity in that region. These contours are flanked by blue contours, suggesting that electron deficient groups might favor activity beyond the red contour regions. This is consistent with the high potency exhibited by 6,7-dihydroxy and 6,7-diamino analogs. The green contour near the 3′-position of the anilino ring as well as the blue and yellow contours near the anilino NH in Figure 5a are similar to those of Figure 3a. Additional information shown by this contour map is the yellow contour near the 2′-position of the anilino ring in Figure 5a, which restricts bulky substituents at this site. The hydrophobic and H-bonding contour maps are shown in Figure 5c. There is a prominent magenta contour near the methoxy substituents, which indicates that a H-bond acceptor is favored at this position similar to the case in Figure 3c. The cyan contour near the anilino NH and the white contour near the 3′-position of the anilino ring are similar to those of 3c. An obvious difference, however, is the absence of an orange contour around the N-1 position, which occurs in Figure 3c. The Conf3 model does not have this apparent artifact.

The CoMFA contour map of the quinoline set with the Conf3 conformation, Figure 6a, gives information regarding substituent effects mainly around the 3-cyano substituent and on the anilino ring. The green contour near the 3′-position of the anilino ring is similar to that of Figure 4a. The red contour near the 3-CN group suggests that a negative electrostatic potential increases potency and the yellow contour adjacent to it prohibits bulky substituents. Unlike the CoMFA maps, the CoMSIA steric and electrostatic maps depicted in Figure 6b show steric influence at the 6- and 7-positions. There are two magenta contours

around the 3-CN and the N-1 in Figure 6c, indicating the preference for H-bond acceptor groups in these positions.

Superimposition of the PLS contour maps unto the binding site surface physicochemical property maps shows complementarities (Figures 7a–d). Figure 7a shows steric contours and receptor topography complementarity, while Figures 7b, c and d show complementarities between electrostatic, hydrophobic and H-bonding contours and receptor surface property maps, respectively.

Docking and consensus dynamics alignment QSAR models

In addition to QSAR analysis with the multifit alignments as described (*vide supra*), we also explored the use of a docking alignment and a novel flexible receptor-guided consensus dynamics alignment. The superimposition of the quinolines **98**–**148** in their docked orientations in the ATP site afforded a docking alignment shown within the binding site in Figure 8a. The notable feature about this docked mode is the placement of the 3-position substituent in the same position as the water molecule that forms the salt bridge-type hydrogen bond between erlotinib and EGFR in the crystallized complex [13], as shown by the superimposition of the cyano group on the water molecule in Figure 8b. It should be noted that without this water molecule we could not reproduce the binding mode of erlotinib in the crystal structure (data not shown). The alignment obtained by the flexible receptor-guided consensus dynamics is shown in Figure 8c, depicting a tighter clustering of the molecules compared to those in the docked alignment.

QSAR modeling with all the molecules in the docked alignment gave q^2 values of -0.054 and -0.032 for CoMFA and CoMSIA, respectively. Corresponding q^2 values obtained using the consensus alignment were 0.020 and 0.037, respectively. The improvement in q^2 is significant but still quite low. The cross-validation prediction residuals were used to eliminate possible outliers, ending up with CoMFA and CoMSIA models that had q^2 values of 0.588 and 0.538, respectively for the docked alignment, after removing compounds **105**, **107**, **109**, **121**, **126**, **134**, **135**, **136**, **138**, **139**, **140** (the ionized form) and **145** for the CoMFA model, and compounds **105**, **107**, **112**, **121**, **124**, **135**, **136**, **137**, **138** and **141** for the CoMSIA model. In the case of the consensus dynamics alignment, a little higher q^2 values of 0.598 and 0.583 were obtained for CoMFA and CoMSIA, respectively, after dropping compounds **104**, **105**, **119**, **121**, **124**, **127**, **135**, **136**, **137** and **138** from the CoMFA analysis, and compounds **105**, **112**, **121**, **127**, **132**, **134**, **135**, **136**, **137** **143** and **145** from the CoMSIA analysis. Group cross-validation and bootstrapping analyses were used to assess the robustness of these models. The PLS statistics are summarized in Table 7. The statistics for these latter alignment models are comparable with those obtained with the multifit alignment (Table 4). The necessity to remove substantial numbers of compounds in the quinoline set for all the alignments suggests there are factors in the assay that the models cannot account for. It should be noted that the experimental system used to test the quinolines was different from that used to test the quinazolines, and this might account for the differences in QSAR results obtained with the two datasets.

The CoMFA and CoMSIA steric contours in both the docked and consensus dynamics alignment models show significant similarities (Figures 9a–d). The electrostatic contours of the CoMFA models have a general similarity (Figures 9a and b), whereas those of the CoMSIA models are quite dissimilar (Figures 9c and d). Similar elements can be seen in the steric and electrostatic maps between these models and the multifit models as well (Figures 4 and 6). The hydrophobic and hydrogen bond contours of the models also do show similarities and differences (Figure 10a–d). The hydrophobic contours appear to be the most consistently located throughout the models and complement with the receptor surface lipophilicity. The hydrogen bond contours do show interesting complementarity with the receptor surface (Figure 10c and d). In the docked alignment CoMSIA, two prominent H-bond acceptor contours are located at the 1-position nitrogen and the 3-position substituent, whereas only the latter was observed in the consensus alignment. Only one H-bond donor contour was observed in both the docking and receptor-guided consensus alignment models, but occurring in different locations. In the docking alignment model, the contour is close to the anilino NH, which is more in line with the models obtained by 'Multifit' alignments described above (see Figures 4c and 6c) whereas in the consensus alignment model, the donor contour is located close to the 2-position hydrogen (see Figure 10b). In the docking and consensus alignment, a H-bond is apparent between the anilino NH and the Asp831 carboxyl group. This suggests that the anilino NH is involved in H-bonding at the binding site. The location of the donor con-

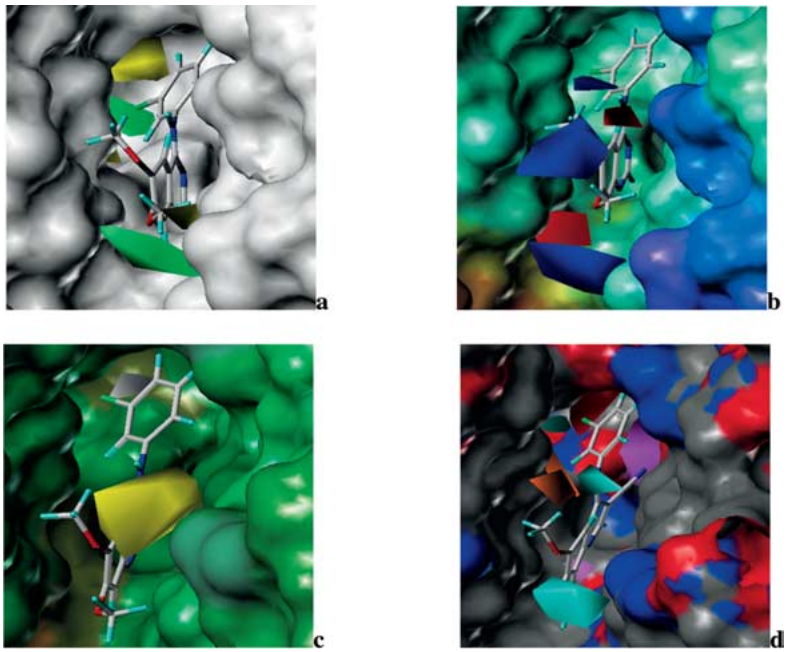


Figure 7. CoMFA and CoMSIA contour plots of the quinazoline and quinoline sets with the Conf3 models superimposed on the surface property maps of the active site of EGFR kinase. (a) CoMFA steric contour map of the quinazoline compounds projected over the water accessible surface. (b) CoMFA electrostatic contour map of the quinazoline set over the electrostatic potential surface map (red, positive potential; blue, negative potential) of the active site. (c) CoMSIA hydrophobic contour map of the quinazoline set over the lipophilic potential surface map (brown, lipophilic; blue, hydrophilic) of the active site. (d) CoMSIA H-bonding contour map of the quinoline dataset over the H-bonding property surface map (red, H-bond donor; blue, H-bond acceptor) of the active site.

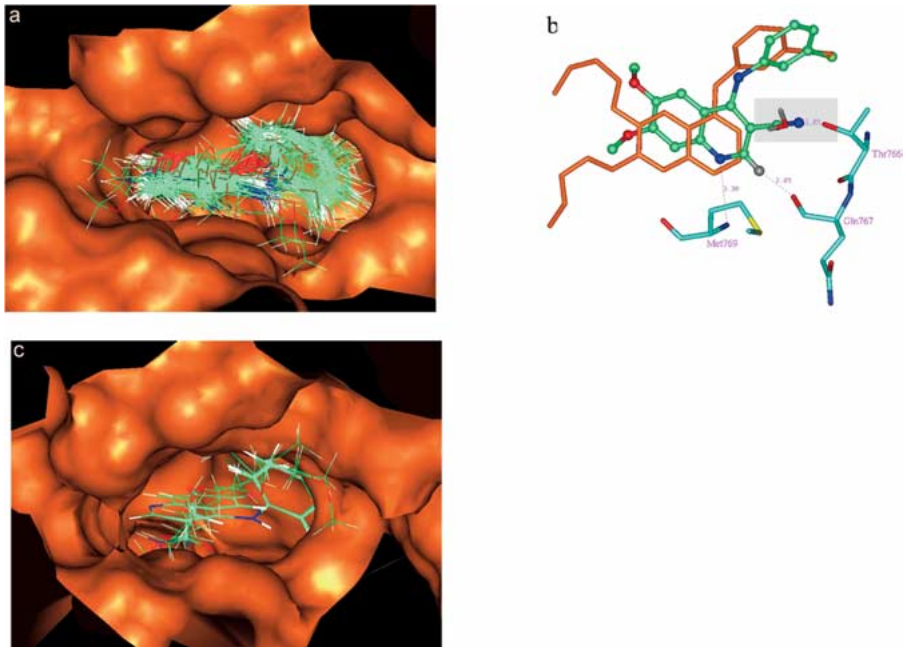


Figure 8. (a) Superimposition of the docked conformation of quinoline derivatives in the ATP site of EGFR. (b) Comparison of the orientation of the docking mode of a 4-anilinoquinoline derivative (colored by atom type) with that of the crystal structure (orange) of erlotinib in the ATP site. The position of the cyano group of the quinoline derivative coincides with that of the bridging water molecule that interacted with erlotinib in the crystal structure of the complex. (c) Receptor-guided consensus dynamics alignment of the 4-anilinoquinoline derivatives within the ATP site.

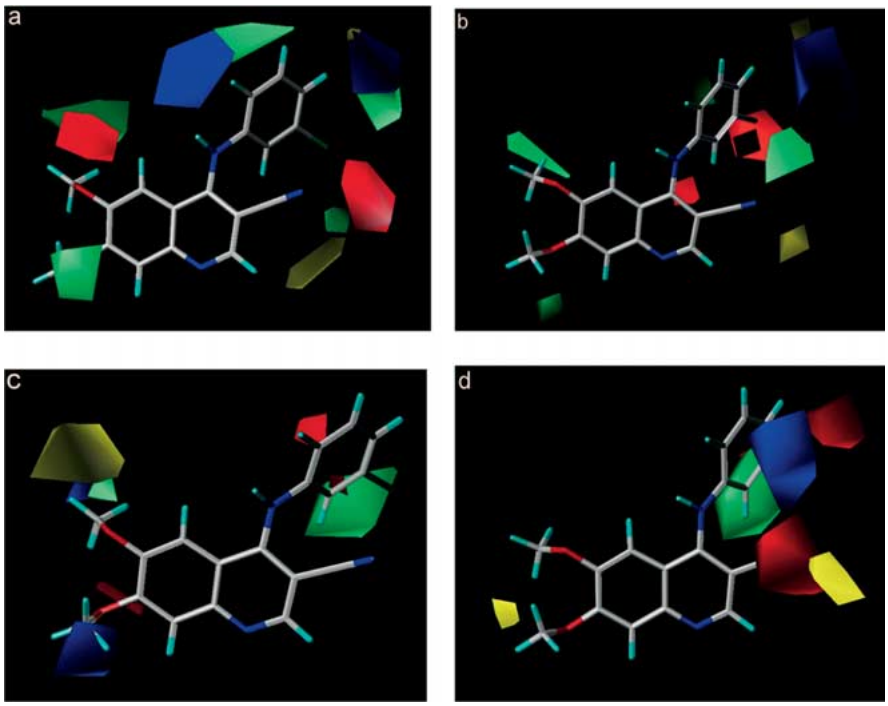


Figure 9. Steric and electrostatic contours obtained from CoMFA and CoMSIA models; (a) and (b) are contours for CoMFA models using docked and consensus alignments, respectively while the CoMSIA contours in (c) and (d) are for docked and consensus alignments, respectively. Color designations are the same as in Figures 3a and 3b.

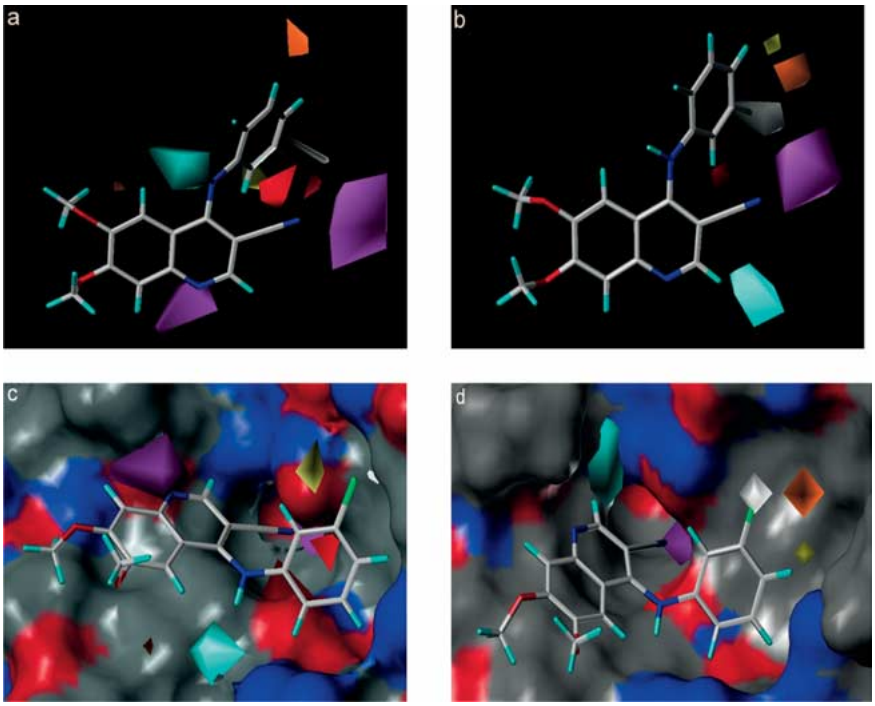


Figure 10. CoMSIA contours for hydrophobic, H-bond donor and H-bond acceptor descriptors. Contours in (a) and (b) were obtained from docked and consensus alignments, respectively (color designations are the same as in Figure 3c) while corresponding projections on the H-bonding receptor surface map are shown in (c) and (d) for docked and consensus alignments, respectively. The colors of the contours and receptor surface property map have the same definitions as given in Figure 7d.

tour on the NH is not surprising, but the location of the donor contour at the H-2 position in the receptor-guided consensus dynamics alignment model seems fortuitous, as was seen for the donor disfavored contour at the vicinity of the N-1 in the Conf2 model (Figure 3c). The hydrogen bond contours in CoMSIA have sometimes been inconsistent, as was demonstrated in a comparative study reported by Bohm et al. [25]. It is not surprising therefore that new hydrogen-bond descriptors are being developed for such use [26]. The fortuitous appearance of the H-bond donor contour in the consensus dynamics alignment model is interesting, as it coincides with the relatively new realization that the C-2 hydrogen in these quinazoline kinase inhibitors does form a hydrogen bond with the backbone carbonyl of Gln 767 or an equivalent residue in the ATP site [27]. The observation of the different location of the contour maps relative to the positions in the molecules again emphasizes the critical importance of alignment rule in deriving CoMFA and CoMSIA 3D-QSAR models.

Conclusions

3D-QSAR models were generated from 'Multifit' alignments based on three different conformational templates, as well as ATP site docked and receptor-guided consensus dynamics alignments of the 4-anilinoquinazoline and 4-anilinoquinoline derivative EGFR kinase inhibitors. The 3D-QSAR models derived from the quinazoline set were validated by an external test set. In regard to prediction of the activities of the test set of quinazolines, the Conf3 alignment models performed best, showing the superiority of the bound conformation over conformational search derived conformations. The use of such a large number of compounds in the quinazoline training dataset and a significant number of external test set compounds in validation enables us to conclude that the X-ray crystallographically determined bound conformation should be sought for 3D-QSAR modeling of ATP-competitive tyrosine kinase inhibitors. This should be kept in mind as investigators apply CoMFA and CoMSIA analyses to receptor tyrosine kinase inhibitors, which analyses have not yet been extensively explored.

The CoMSIA studies undertaken with the quinoline set strongly indicate that a H-bond acceptor is required in the 3-position to achieve high EGFR kinase inhibitory activity. This is an interesting finding in light of the model proposed by Wissner et al. [16] which purported that N-1 and N-3 are H-bond acceptor points in the quinazoline inhibitors even before the X-ray crystallographic determination of the 3D structure of EGFR in complex with a quinazoline inhibitor was accomplished.

We have also introduced, for the first time, a novel receptor-guided consensus dynamics alignment for 3D-QSAR studies. This alignment incorporates to some extent ligand-receptor induced fit effects into the QSAR model [28]. After we had finished our studies, a report appeared describing the use of consensus dynamics without receptor guidance for obtaining alignments for CoMFA and CoMSIA studies [29].

References

1. Hong, W.K. and Ullrich, A., Oncol. Biotherap. 1 (2000) 1.
2. Salomon, D.S., Brandt, R., Cardiello, F. and Normanno, N., Crit. Rev. Oncol. Hematol. 19 (1995) 183.
3. Woodburn, J.R., Pharmacol. Ther., 82 (1999) 241.
4. Gosh, S., Liu, X.-P., Zheng, Y. and Uckun, F.M., Curr. Cancer Drug Targets, 1 (2001) 129.
5. Mendelsohn, J. and Baselga, J., Oncogene, 19 (2000) 6550.
6. Rewcastle, G.W., Palmer, B.D., Bridges, A.J., Showalter, H.D.H., Sun, L., Nelson, J., McMichael, A., Kraker, A.J., Fry, D.W. and Denny, W.A., J. Med. Chem., 39 (1996) 918.
7. Bridges, A.J., Zhou, H., Cody, D.R., Rewcastle, G.W., McMichael, A., Showalter, H.D.H., Fry, D.W., Kraker, A.J. and Denny, W.A., J. Med. Chem., 39 (1996) 267.
8. Rewcastle, G.W., Denny, W.A., Bridges, A.J., Zhou, H., Cody, D.R., McMichael, A. and Fry, D.W., J. Med. Chem., 38 (1995) 3482.
9. Palmer, B.D., Trumpp-Kallmeyer, S., Fry, D.W., Nelson, J.M., Showalter, H.D. and Denny, W.A., J. Med. Chem., 40 (1997) 1519.
10. Baselga, J. and Averbuch, S.D., Drugs, 60 (Suppl. 1) (2000) 33.
11. Ciardiello, F., Caputo, R., Bianco, R. et al., Clin. Cancer Res., 6 (2000) 2053.
12. de Bono, J.S. and Rowinsky, E.K., Trend Mol. Med., 8 (2002) S19.
13. Stamos, J., Sliwkowski, M.X. and Eigenbrot, C.J., Biol. Chem., 277 (2002), 46265.
14. Cramer III, R.D., Patterson, D.E. and Bunce, J.D., J. Am. Chem. Soc., 110 (1988) 5959.
15. Klebe, G., Abraham, U. and Mietzner, T.J., J. Med. Chem., 37 (1994) 4130.
16. Wissner, A., Berger, D.M., Boschelli, D.H., Floyd, M.B. Jr., Greenberger, L.M., Graber, B.C., Johnson, B.D., Mamuya, N., Nilakantan, R., Reich, M.F., Shen, R., Tsou, H.R., Upeslacis, E., Wang, Y.F., Wu, B., Ye, F. and Zhang, N., J. Med. Chem., 43 (2000) 3244.
17. Clark, M., Cramer III, R.D. and Van Opdenbosch, N., J. Comput. Chem., 10 (1989) 982.
18. Stewart, J.J., J. Comput.-Aided Mol. Design, 4 (1990) 1.
19. Jones, G., Willet, P., Glen, R.C., Leach, A.R. and Taylor, R., J. Mol. Biol., 267 (1997) 727.

20. Kamath, S. and Coutinho, E., J. Biosci., 22 (1997) 315.
21. Viswanadhan, V.N., Ghoise, A.K., Revenkar, G.R. and Robins, R., J. Chem. Inf. Comput. Sci., 29 (1989) 163.
22. Klebe, G., J. Mol. Biol., 237 (1994) 212.
23. Buolamwini, J.K., Raghavan, K., Fesen, M.R., Pommier, Y., Kohn, K.W. and Weinstein J.N., Pharm. Res., 13 (1996) 1891.
24. Buolamwini, J.K. and Assefa, H., CoMFA and CoMSIA 3D-QSAR Studies of Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitors. 221st ACS National Meeting, San Diego, CA, April 1–5, 2001.
25. Bohm, M., Sturzebecher, J. and Klebe, G., J. Med. Chem., 42 (1999) 458.
26. Bohm, M. and Klebe, G., J. Med. Chem., 45 (2002) 1585.
27. Pierce, A.C., Sandretto, K.L. and Bemis G.W., Proteins, 49 (2001) 567.
28. Vedani, A. and Dobbler, M., J. Med. Chem., 45 (2002) 2139.
29. Datar, P., Desai, P., Coutinho, E. and Iyer, K.J., J. Mol. Model., 8 (2002) 290.