Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry 13 (2005) 4704–4712
Bioorganic & Medicinal Chemistry

3D-QSAR studies on c-Src kinase inhibitors and docking analyses of a potent dual kinase inhibitor of c-Src and c-Abl kinases☆

Ram Thaimattam,[a,*] Pankaj R. Daga,[a] Rahul Banerjee[a] and Javed Iqbal[b]

[a]*Department of Molecular Modeling and Drug Design, Dr. Reddy's Laboratories Ltd, Bollaram Road, Miyapur, Hyderabad 500 049, India*

[b]*Discovery Chemistry, Dr. Reddy's Laboratories Ltd, Discovery Research, Bollaram Road, Miyapur, Hyderabad 500 049, India*

Received 24 March 2005; revised 26 April 2005; accepted 26 April 2005
Available online 23 May 2005

Abstract—Three-dimensional quantitative structure–activity relationship (3D-QSAR) analyses were carried out on quinazoline, quinoline, and cyanoquinoline derivatives inhibiting c-Src kinase. Comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) 3D-QSAR models were developed. The conventional r^2 values for CoMFA and CoMSIA are 0.93 and 0.89, respectively. In addition, a homology model of c-Src kinase with the activation loop resembling the active conformation was constructed using the crystal structure of the kinase domain of Lck. The ATP binding pocket of the active form of c-Src is similar to that of the c-Abl kinase in which the activation loop resembles that of an active form. One of the potent c-Src and c-Abl dual kinase inhibitors (**77** or SKI-606) was docked inside the active sites of both c-Src and c-Abl. The orientation and hydrogen bonding interactions of **77** are similar in both kinases. The results of 3D-QSAR analyses and structure based studies will be useful for the design of novel c-Src and c-Abl dual kinase inhibitors.

1. Introduction

Kinases constitute almost 2% of the human genome and are critical to the transmission of signals both between and within cells. Aberrant kinase activity has been associated with many diseases, in particular those involving inflammatory or proliferative responses. There are about 500 kinases, and all of them make use of ATP for phosphorylation of a specific Tyr/Ser/Thr on proteins. Therefore, designing small-molecule ATP-competitive inhibitors targeting a specific enzyme is not an easy task. However, the success of drugs targeting kinases, such as imatinib (Novartis), erlotinib (OSI Pharmaceuticals Inc.), and gefitinib (AstraZeneca), have led to intense interest in drug development efforts directed at kinases.[1]

Currently, there is an increased interest in the design of the c-Src and Bcr–Abl tyrosine kinase inhibitors. Increased c-Src expression and activity are associated with tumor malignancy and poor prognosis.[2] On the other hand, abnormal kinase activity of Bcr–Abl kinase, which is a result of the fusion of genes responsible for c-Abl kinase and Bcr protein, is responsible for chronic myeloid leukemia (CML).[3] c-Abl activates Src family kinases in an Abl-kinase-independent manner suggesting that c-Src kinase inhibitors might be useful in the treatment of CML.[4b] Reports also indicate that the inhibition of both c-Src and c-Abl would probably be helpful for the treatment of imatinib-resistant CML.[4]

Several c-Src kinase inhibitors have been identified to date.[5–8] These include various scaffolds: dihyropyrimido-quinolinones,[5] pyrrazolo-pyrimidines,[6] 4-anilinoquinazolines,[7] and others.[8] Recently, Boshelli and co-workers[9] have published a series of 4-anilino-3-quinolinecarbonitriles exhibiting potent c-Src kinase inhibitory activity and few of them inhibited c-Abl kinase activity as well.[9d] In this study, we have developed predictive 3D-QSAR models using quinazoline and quinoline derivatives, which have been reported as c-Src kinase inhibitors.[9] The studies include comparative molecular field analysis (CoMFA)[10] and comparative molecular similarity indices analysis (CoMSIA).[11] A homology model of the active form of c-Src kinase

Keywords: 3D-QSAR; Homology modeling; Docking; c-Src kinase inhibitors; Src–Abl dual kinase inhibitors.
☆ DRL Publication No. 477.
* Corresponding author. Tel.: +91 40 2304 5439; fax: +91 40 2304 5438; e-mail: ramthaimattam@drreddys.com

0968-0896/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2005.04.065

was also constructed using the crystal structure of the Lck kinase domain.[12] In addition, docking studies of **77**, one of the potent c-Src and c-Abl dual kinase (Src–Abl dual) inhibitors, were carried out inside the active sites of both c-Src and c-Abl to gain insight into the structural requirements for Src–Abl dual inhibitory activity for this class of molecules. Compound **77**, which is also known as SKI-606 is currently in the clinical trials.[9d]

2. Methods

2.1. Data set

Series of potent quinazoline and quinoline derivatives, reported to have c-Src kinase inhibitory activities, were chosen in this study (Scheme 1).[9] In vitro c-Src kinase inhibitory activities were converted into the corresponding pIC_{50} ($-\log IC_{50}$) values. The pIC_{50} values were used as dependent variables in the CoMFA and CoMSIA analyses. The total set of c-Src kinase inhibitors (87 compounds) was divided into training (73 compounds) and test (14 compounds) sets in the approximate ratio 5:1 (Tables 1 and 2). Training and test sets were chosen such that structurally diverse molecules exhibiting activities of a wide range were included in both sets. The pIC_{50} values of the molecules spanned a range of 4 log units.

2.2. Molecular modeling

Bioactive conformation and molecular alignment rules are crucial variables in 3D-QSAR. The bioactive conformation of one of the most active cyanoquinolines (**77**) was simulated using FlexX (see Docking studies). The docked conformation was used as the template to construct the 3D models of all the compounds using SYBYL 6.92.[13] All structures were minimized using MMFF94[14] force fields and Powell method with a convergence criterion of 0.001 kcal/mol, and using a distance dependent dielectric constant value of 1. Atomic charges were calculated using the MMFF94 method.

2.3. Alignment

The binding pose of **77** was used as the template for aligning the rest of the molecules to it. Various alignment methods, namely atomfit, multifit, and database, were carried out. In addition, shape-based alignment using FlexS was also explored.

2.4. CoMFA and CoMSIA

The aligned training set molecules were placed in a 3D grid box such that the entire set was included in it. CoMFA fields were generated using sp^3 carbon probe atom carrying +1 charge to generate steric (Lennard–Jones 6–12 potential) and electrostatic (Coulomb potential) fields at each grid point.[10] The steric and electrostatic energy values in CoMFA were truncated at 30 kcal/mol. The CoMFA fields were scaled by the CoMFA-STD method in SYBYL. The CoMSIA method defines explicit hydrophobic (H) and hydrogen bond donor (D), and acceptor (A) descriptors in addition to the steric (S) and electrostatic (E) fields used in CoMFA. The CoMSIA fields were derived according to Klebe et al.[11] Arbitrary definition of cutoff limits is not required in CoMSIA method, wherein a distance dependent Gaussian type functional form will take into account abrupt changes of potential energy near the molecular surface. The default value of 0.3 was used as the attenuation factor.

The CoMFA/CoMSIA fields combined with observed biological activities (pIC_{50}) were included in a molecular spreadsheet and partial least square (PLS)[15] methods were applied to generate 3D-QSAR models. The PLS algorithm with the leave-one-out[16] cross-validation method was employed to choose optimum number of components and assess the statistical significance of each model. All cross-validated PLS analyses were performed with a column filter value of 2.0. The cross-validated $r^2(r^2_{cv})$ was calculated using the equation

$$r^2_{cv} = 1 - \frac{\sum(Y_{predicted} - Y_{observed})^2}{\sum(Y_{observed} - Y_{mean})^2}, \quad (1)$$

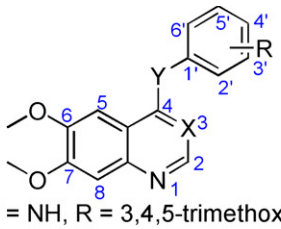
where $Y_{predicted}$, $Y_{observed}$, and Y_{mean} are predicted, actual, and mean values of the target property (pIC_{50}), respectively. The optimum number of components was chosen, which gave less standard error of prediction and high r^2_{cv}. The predictive $r^2(r^2_{pred})$ value was calculated using

$$r^2_{pred} = (SD - PRESS)/SD, \quad (2)$$

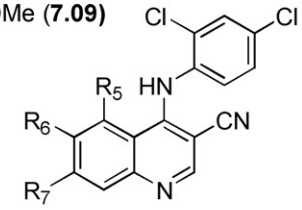
where SD is the sum of squared deviations between the biological activity of the test set and the mean activity of training set molecules, and PRESS is the sum of squared deviations between the actual and the predicted activities of the test set molecules. In addition, the r^2_{cv}, r^2_{pred}, and number of components, the conventional correlation coefficient (r^2), and its standard error were also computed for each model. The CoMFA/CoMSIA results were graphically interpreted by field contribution maps using the ''STDEV*COEFF'' field type.

2.5. Homology modeling

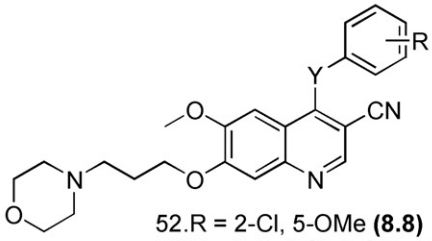
Tertiary structure of c-Src in which the activation loop resembles that of an active form, for which a crystal structure is not available, was constructed by the method of homology modeling using the crystal structure of the human Lck kinase domain. In homology modeling, 3D structure of homologous protein is used as a template to construct an unknown structure. The amino acid sequence identity between the kinase domains of c-Src and Lck is approximately 67%. The homology model of the c-Src kinase domain includes amino acid residues 255–523. The alignment of the human c-Src and Lck kinase domain amino acid sequences is shown in Scheme 2.[17] The 3D structure modeling and refinements were carried out using Swiss-PdbViewer.[18] The model was finally minimized with MMFF94 force fields using SYBYL.



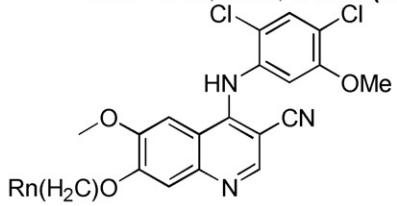
1. X = N, Y = NH, R = 3,4,5-trimethoxy **(6.55)**
2. X = N, Y = NH, R = 3,5-dimethoxy-4-ethoxy **(6.47)**
3. X = N, Y = NH, R = 3,5-dimethoxy-4-n-butoxy **(6.31)**
4. X = N, Y = NH, R = 3,5-dimethoxy-4-t-butoxy **(6.08)**
5. X = N, Y = NH, R = 3,5-dimethoxy-4-isopropoxy **(5.57)**
6. X = C-CN, Y = NH, R = 3,4,5-trimethoxy **(7.46)**
7. X = C-CN, Y = NH, R = 3,5-dimethoxy-4-ethoxy **(7.20)**
8. X = C-CN, Y = NH, R = 3,5-dimethoxy-4-isopropoxy **(6.96)**
9. X = C-CN, Y = NH, R = 3,4-dimethoxy **(6.02)**
10. X = C-CN, Y = NH, R = 2-Bromo-3,4,5-trimethoxy **(6.77)**
11. X = CH, Y = NH, R = 3,4,5-trimethoxy **(6.70)**
12. X = C-CN, Y = O, R = 3,4,5-trimethoxy **(6.54)**
13. X = C-CN, Y = NH, R = 2,4-diCl **(7.52)**
14. X = N, Y = NH, R = 2,4-diCl **(6.6)**
15. X = CH, Y = NH, R = 2,4-diCl **(7.08)**
16. X = C-CN, Y = S, R = 2,4-diCl **(5.66)**
17. X = C-CHO, Y = NH, R = 2,4-diCl **(6.60)**
23. X = C-CN, Y = NH, R = 2-Cl **(7.02)**
24. X = C-CN, Y = NH, R = 4-Cl **(7.28)**
25. X = C-CN, Y = NH, R = 2,3-diCl **(7.05)**
26. X = C-CN, Y = NH, R = 2,5-diCl **(7.40)**
27. X = C-CN, Y = NH, R = 2,6-diCl **(7.21)**
28. X = C-CN, Y = NH, R = 3,5-diCl **(5.96)**
29. X = C-CN, Y = NH, R = 3,4-diCl **(7.24)**
30. X = C-CN, Y = NH, R = 2-F, 4-Cl **(7.28)**
31. X = C-CN, Y = NH, R = 2-Br, 4-Cl **(8.00)**
32. X = C-CN, Y = NH, R = 2-I, 4-Cl **(8.21)**
33. X = C-CN, Y = NH, R = 2-Cl, 4-I **(7.02)**
34. X = C-CN, Y = NH, R = 2-Cl, 4-Br **(7.82)**
40. X = C-CN, Y = NH, R = 2-Cl, 5-OMe **(8.00)**
41. X = C-CN, Y = NH, R = 2,4-di-Cl, 5-OMe **(8.37)**
42. X = C-CN, Y = NH, R = 2,4-di-Cl, 5-OEt **(5.92)**
43. X = C-CN, Y = NH, R = 2,4-di-Cl, 5-OH **(6.72)**
44. X = C-CN, Y = NH, R = 2-F,4-Cl, 5-OMe **(7.18)**
45. X = C-CN, Y = NH, R = 2-Br,4-Cl, 5-OMe **(8.52)**
46. X = C-CN, Y = NH, R = 2-Me,4-Cl, 5-OMe **(8.40)**
47. X = C-CN, Y = NH, R = 4-Cl, 3-OMe **(6.92)**
48. X = C-CN, Y = NH, R = 2-Cl, 4-Me, 5-OMe **(8.54)**
49. X = C-CN, Y = NH, R = 2,4-di-Me, 5-OMe **(8.34)**
50. X = N, Y = NH, R = 2,4-di-Cl, 5-OMe **(6.41)**
51. X = CH, Y = NH, R = 2,4-di-Cl, 5-OMe **(7.09)**



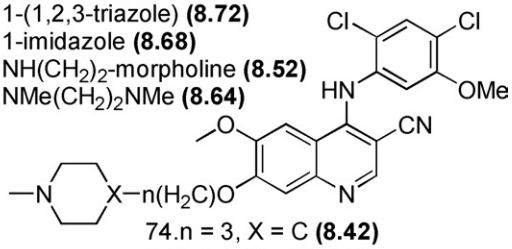
18. R_5 = OMe, R_6 = H, R_7 = H **(5.92)**
19. R_5 = H, R_6 = OMe, R_7 = H **(6.49)**
20. R_5 = H, R_6 = H, R_7 = OMe **(6.70)**
21. R_5 = H, R_6 = OEt, R_7 = OEt **(7.96)**
22. R_5 = H, R_6 = O-n-Bu, R_7 = O-n-Bu **(6.80)**
35. R_5 = H, R_6 = R_7 = $(CH_2)_3$-morpholine **(7.72)**
36. R_5 = H, R_6 = $(CH_2)_3$-morpholine, R_7 = Me **(6.66)**
37. R_5 = H, R_6 = Me, R_7 = $(CH_2)_3$-morpholine **(8.42)**
38. R_5 = H, R_6 = Me, R_7 = $(CH_2)_3$-N-methyl-piperazine **(8.06)**
39. R_5 = H, R_6 = Me, R_7 = $(CH_2)_3$-4-hydroxy-piperidine **(8.46)**



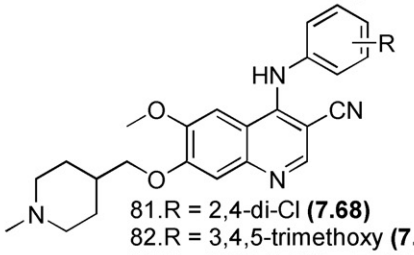
52. R = 2-Cl, 5-OMe **(8.8)**
53. R = 2,4-di-Cl, 5-OMe **(9.10)**
54. R = 2-Br,4-Cl, 5-OMe **(9.02)**
55. R = 2-Me,4-Cl, 5-OMe **(8.92)**
56. R = 2,4-di-Me, 5-OMe **(8.77)**
57. R = 2-Cl, 4-Me, 5-OMe **(8.96)**



58. n = 4, R = Morpholine **(8.89)**
59. n = 5, R = Morpholine **(8.60)**
60. n = 3, R = N-Methylpiperazine **(8.92)**
61. n = 3, R = N-ethylpiperazine **(9.11)**
62. n = 3, R = cis-3,4,5-tri-Me-Piperazine **(8.85)**
63. n = 3, R = 4-n-Pr-piperazine **(8.96)**
64. n = 3, R = piperazine **(8.96)**
65. n = 3, R = 4-$(CH_2)_2$)OH-piperazine **(9.12)**
66. n = 3, R = 4-Me-homopiperazie **(8.85)**
67. n = 3, R = piperidine **(8.92)**
68. n = 3, R = 4-OH-piperidine **(9.19)**
69. n = 3, R = 2-(1,2,3-triazole) **(8.03)**
70. n = 3, R = 1-(1,2,3-triazole) **(8.72)**
71. n = 3, R = 1-imidazole **(8.68)**
72. n = 3, R = NH$(CH_2)_2$-morpholine **(8.52)**
73. n = 3, R = NMe$(CH_2)_2$NMe **(8.64)**



74. n = 3, X = C **(8.42)**
75. n = 1, X = C **(8.15)**
76. n = 2, X = C **(8.11)**
77. n = 3, X = N **(8.57)**
78. n = 0, X = C **(6.64)**
79. R = 2-sub analog of 75 **(7.06)**
80. R = 3-sub analog of 75 **(8.05)**



81. R = 2,4-di-Cl **(7.68)**
82. R = 3,4,5-trimethoxy **(7.46)**
83. R = 2-Cl, 5-OMe **(7.92)**
84. R = 2-Me, 5-OMe **(7.62)**
85. R = 2,4-di-Me **(7.20)**
86. R = 2,4-di-Me, 5-OMe **(7.89)**
87. R = 2,4-di-Cl, 5-OEt **(5.85)**

Scheme 1. Molecules used for 3D-QSAR studies and their actual pIC_{50} values in the parentheses.

Table 1. Experimental pIC_{50} (Exp), predicted pIC_{50} (Pred), and residual values (δ) of molecules used in the training set for CoMFA and CoMSIA

Molecule	Exp	CoMFA		CoMSIA	
		Pred	δ	Pred	δ
2	6.47	6.63	0.16	6.41	−0.06
3	6.31	6.37	0.06	6.06	−0.25
4	6.08	5.89	−0.19	5.87	−0.21
5	5.57	5.54	−0.02	5.92	0.35
6	7.46	7.59	0.13	7.36	−0.09
7	7.20	7.27	0.06	7.07	−0.13
8	6.96	6.89	−0.07	6.79	−0.17
9	6.02	5.63	−0.39	5.83	−0.19
10	6.77	7.02	0.25	7.30	0.53
11	6.70	6.64	−0.06	6.86	0.16
12	6.54	6.47	−0.07	6.67	0.13
13	7.52	7.13	−0.39	7.16	−0.37
14	6.60	6.63	0.03	6.52	−0.08
15	7.08	7.16	0.08	6.87	−0.20
16	5.66	5.70	0.04	5.67	0.01
19	6.49	6.69	0.19	6.78	0.29
20	6.70	6.86	0.16	6.83	0.13
21	7.96	7.94	−0.02	7.54	−0.42
22	6.80	6.91	0.11	7.19	0.40
23	7.02	7.43	0.40	7.65	0.62
24	7.28	6.95	−0.33	7.25	−0.02
25	7.05	6.60	−0.44	6.80	−0.25
26	7.40	7.12	−0.28	7.20	−0.20
27	7.21	7.31	0.10	7.25	0.03
28	5.96	6.13	0.17	6.28	0.32
30	7.28	7.37	0.09	7.14	−0.13
31	8.00	7.68	−0.33	7.86	−0.14
32	8.21	7.83	−0.38	8.39	0.18
37	8.42	8.55	0.13	8.50	0.08
38	8.06	8.01	−0.05	8.19	0.13
39	8.46	8.35	−0.11	9.03	0.57
42	5.92	6.56	0.64	7.03	1.11
44	7.18	7.73	0.55	7.30	0.12
45	8.52	7.86	−0.66	7.87	−0.66
46	8.40	8.19	−0.21	7.83	−0.57
47	6.92	7.10	0.18	7.33	0.41
49	8.34	8.27	−0.07	8.10	−0.23
50	6.41	6.91	0.50	6.62	0.21
52	8.80	8.74	−0.06	8.90	0.10
53	9.10	8.87	−0.23	8.84	−0.26
54	9.02	8.80	−0.22	8.93	−0.09
55	8.92	9.24	0.32	8.99	0.07
56	8.77	8.82	0.05	9.19	0.42
57	8.96	8.91	−0.05	8.93	−0.03
58	8.89	8.64	−0.25	8.63	−0.26
59	8.60	8.39	−0.21	8.55	−0.05
60	8.92	8.77	−0.15	8.74	−0.18
61	9.11	9.13	0.01	8.90	−0.21
62	8.85	8.85	0.00	8.84	−0.02
63	8.96	9.01	0.06	8.93	−0.03
64	8.96	9.00	0.05	9.05	0.09
65	9.12	9.27	0.15	9.14	0.02
66	8.85	8.79	−0.07	8.27	−0.59
67	8.92	8.77	−0.15	8.51	−0.41
68	9.19	9.21	0.02	9.47	0.27
69	8.03	8.48	0.45	8.12	0.09
70	8.72	9.18	0.46	8.54	−0.18
71	8.68	8.88	0.21	8.26	−0.42
72	8.52	8.45	−0.08	8.53	0.01
73	8.64	8.61	−0.03	8.91	0.27
74	8.42	8.78	0.36	8.75	0.32
75	8.15	7.58	−0.58	7.31	−0.85
76	8.11	7.99	−0.12	8.19	0.07
77	8.57	8.53	−0.04	8.77	0.20
78	6.64	6.98	0.34	6.90	0.26
79	7.06	6.86	−0.20	6.69	−0.37
81	7.68	7.27	−0.40	7.18	−0.50
82	7.46	7.29	−0.16	7.15	−0.30
83	7.92	7.61	−0.31	7.60	−0.32
84	7.62	7.67	0.05	7.72	0.10
85	7.20	7.29	0.09	7.54	0.34
86	7.89	7.84	−0.05	7.79	−0.09
87	5.85	6.67	0.81	7.00	1.14

Table 2. Experimental pIC_{50} (Exp), predicted pIC_{50} (Pred), and residual values of molecules used in the test set for CoMFA and CoMSIA

Molecule	Exp	CoMFA		CoMSIA	
		Pred	δ	Pred	δ
1	6.55	6.23	−0.32	6.15	−0.40
17	6.60	7.43	0.83	7.11	0.50
18	5.92	5.30	−0.62	5.58	−0.34
29	7.24	6.83	−0.41	6.74	−0.50
33	7.02	7.70	0.67	7.72	0.70
34	7.82	7.68	−0.15	7.66	−0.16
35	7.72	7.93	0.20	8.34	0.62
36	6.66	7.28	0.62	7.64	0.98
40	8.00	7.52	−0.48	7.77	−0.23
41	8.37	7.78	−0.58	7.54	−0.83
43	6.72	7.52	0.79	7.64	0.92
48	8.54	8.08	−0.46	7.97	−0.56
51	7.09	7.30	0.21	6.99	−0.10
80	8.05	7.39	−0.66	7.69	−0.35

2.6. Docking studies

One of the most potent c-Src kinase inhibitors, **77**, was docked inside the active site of c-Src (PDB accession code: 2SRC)[19] using FlexX, an automated docking program incorporated in SYBYL. Hydrogen atoms were added to the protein while all the amino acid residues were considered in the neutral form. The ligand–receptor complex was then minimized using MMFF94 force fields and Powell method with a gradient convergence value of 0.05 kcal/mol and a distance dependent dielectric constant value of 1. Atomic charges were computed using MMFF94 method. All amino acid residues within an 8 Å radius around the ligand were minimized while treating the rest of the protein as an aggregate. The binding pose of **77** inside the active site of c-Src was used as the reference for the other docking studies (see Docking analysis), and essentially the same minimization procedure was used for the minimizations.

3. Results and discussion

The CoMFA and CoMSIA 3D-QSAR methods were used for deriving 3D-QSAR models for quinazoline and quinoline derivatives exhibiting c-Src kinase inhibitory activities. The chemical structures of molecules and their actual pIC_{50} values are shown in Scheme 1. The

```
2SRC   249   KPQTQGLAKDAWEIPRESLRLEVKLGQGCFGEVWMGTWNGTTRVAIKTLKPGTMSPEAFL
3LCK   231   KP----WWEDEWEVPRETLKLVERLGAGQFGEVWMGYYNGHTKVAVKSLKQGSMSPDAFL
             **      :* **:***:*:*  :** * ******* :** *:**:*:** *:***:***

2SRC   309   QEAQVMKKLRHEKLVQLYAVVSEEPIYIVTEYMSKGSLLDFLKGETGKYLRLPQLVDMAA
3LCK   287   AEANLMKQLQHQRLVRLYAVVTQEPIYIITEYMENGSLVDFLKTPSGIKLTINKLLDMAA
              **:**:*:*::**:*****::*****:****.:***:****  :*  * :  :*:****

2SRC   369   QIASGMAYVERMNYVHRDLRAANILVGENLVCKVADFGLARLIEDNEYTARQGAKFPIKW
3LCK   347   QIAEGMAFIEERNYIHRDLRAANILVSDTLSCKIADFGLARLIEDNE-TAREGAKFPIKW
             ***.***::*. **:***********.:.* **:************** ***:********

2SRC   429   TAPEAALYGRFTIKSDVWSFGILLTELTTKGRVPYPGMVNREVLDQVERGYRMPCPPECP
3LCK   407   TAPEAINYGTFTIKSDVWSFGILLTEIVTHGRIPYPGMTNPEVIQNLERGYRMVRPDNCP
             *****  ** ****************:.*:**:*****.* **:::******  * :**

2SRC   489   ESLHDLMCQCWRKEPEERPTFEYLQAFLEDYFTSTEPQQPGENL
3LCK   467   EELYQLMRLCWKERPEDRPTFDYLRSVLEDFFTAT---------
             *.*::**  **:.:**:****:**:.***:**:*
```

Scheme 2. Alignment of human c-Src and Lck kinase domains. The alignment was performed using ClustalX.[19] Residues that are identical in the two sequences are indicated by asterisk below the sequence; those that are very similar are indicated by colon, while similar residues are indicated by a dot.

data set was divided into training and test sets (Tables 1 and 2). The predictive power of the 3D-QSAR models, derived using the training set, was assessed by predicting biological activities of the test set molecules. The docked conformation of **77** inside the active site of c-Src (PDB code: 2SRC)[19] was used as the template for the superposition of rest of the molecules. Several 3D-QSAR models were generated using four different alignment methods.

CoMFA and CoMSIA 3D-QSAR methods are based on the assumption that the changes in binding affinities of ligands are related to changes in molecular properties represented by fields. In CoMFA and CoMSIA methods, the steric fields are represented by green and yellow-colored contours (green, bulky substitution favored; yellow, bulky substitution disfavored) and the electrostatic fields are indicated by red and blue-colored contours (red, electronegative group favored; blue, electropositive group favored). The explicit hydrogen bond donor (D), hydrogen bond acceptor (A), and hydrophobic (H) fields defined in CoMSIA are represented by cyan- and purple-colored contours (cyan, favored; purple, disfavored), magenta and red contours (magenta, favored; red, disfavored), and yellow- and white-colored contours (yellow, favored; white, disfavored), respectively.

3.1. CoMFA and CoMSIA 3D-QSAR Analyses

CoMFA and CoMSIA 3D-QSAR models were obtained using standard procedures.[10,11] Various 3D-QSAR models were generated using atomfit, multifit, database, and shape-based alignment methods. Statistics of the PLS analyses for CoMFA and CoMSIA models were more or less similar in all these cases (not shown); however, the CoMFA and CoMSIA contour plots were better in case of shape-based alignment. In addition, CoMSIA model that included all fields (S, E, D, A, and H) performed better than other field combinations. Therefore, the CoMFA and CoMSIA models derived using shape-based alignment were considered in the final analysis. The statistical details of these two models are summarized in Table 3. While the cross-validated r^2 value for CoMFA is slightly lower than that of CoMSIA (0.612 and 0.688, respectively), the conventional r^2 is better in case of CoMFA (the respective r^2 values for CoMFA and CoMSIA are 0.931 and 0.887) with an optimum number of six components. The standard error of estimation (SEE) and predictive r^2 values for CoMFA are better (0.284 and 0.591, respectively) than those of CoMSIA (0.363 and 0.539, respectively). The standard correlation coefficient (r^2_{conv}) relates how well the equation or model is able to calculate the actual experimental values of the compounds comprising the training set; while the cross-validated r^2 (called r^2_{cv}) indicates how well each compound in the training set can be estimated by the remainders left. The r^2_{cv} and r^2_{pred} values are always somewhat lower, and often much lower than the conventional r^2_{conv} for the same data, because it is harder to predict values, which are not used in deriving the model than it is to fit the same values while including them in a model (SYBYL, Ligand-Based Design Manual).

Table 3. Summary of 3D-QSAR analyses on c-SRC kinase inhibitors

	CoMFA	CoMSIA
r^2_{cv} [a]	0.612	0.688
N [b]	6	6
r^2_{conv} [c]	0.931	0.887
SEE [d]	0.284	0.363
r^2_{pred} [e]	0.591	0.539
Field contributions [f]	0.549 (S); 0.448 (E)	0.140 (S); 0.274 (E); 0.120 (A); 0.143 (D); 0.319 (H)

[a] Cross-validated correlation coefficient.
[b] Optimum number of components obtained form cross-validated PLS analysis, and the same used in final non-cross-validated analysis.
[c] Non-cross-validated correlation coefficient.
[d] Standard error of estimate.
[e] Predictive r^2.
[f] Field contributions: steric (S) and electrostatic (E) fields from CoMFA and CoMSIA. Hydrophobic (H), hydrogen bond donor (D), and acceptor (A) fields from CoMSIA.

The actual and predicted pIC_{50} values of the training and test sets are shown in Tables 1 and 2, respectively.

3.2. CoMFA and CoMSIA contour plots

The CoMFA steric and electrostatic contour plots obtained from the shape-based alignment are shown in Figure 1. The green contour observed near the 2′-position of the aniline ring (see Scheme 1 for the atom numbering) indicates that any bulky group substitution at this position might increase the activity (Fig. 1a). Molecules **41**, **48**–**49** show similar activities as the volumes of Cl and CH_3 groups are comparable. The yellow regions near the 5′-position of the aniline ring suggest that any bulky group substitution at this position is likely to decrease the activity. Examination of steric fields at 6- and 7-positions indicates that these two positions show different substitutional requirements. While a bulky substituent at the 6-position decreases the activity, the effect is opposite at the 7-position; bulky group substitutions at this position enhance the activity. The CoMFA electrostatic contours are shown in Figure 1b. The blue contour near aniline NH indicates that electropositive group substitution at this position enhances the activity. The red contour near CN group suggests that electronegative group at 3-position is desirable for higher activity. Therefore, cyanoquinolines are more active than corresponding quinazolines or unsubstituted quinolines. The red regions near the 2′-, 4′-, and 5′-positions of the aniline ring indicate that electronegative substituents at these positions may enhance the activity. The CoMSIA steric and electrostatic contour plots (figures not shown) generated using shape-based alignment method are more or less similar to the corresponding CoMFA plots.

The CoMSIA hydrogen bond donor and acceptor contour plots are shown in Figure 2a and b, respectively.

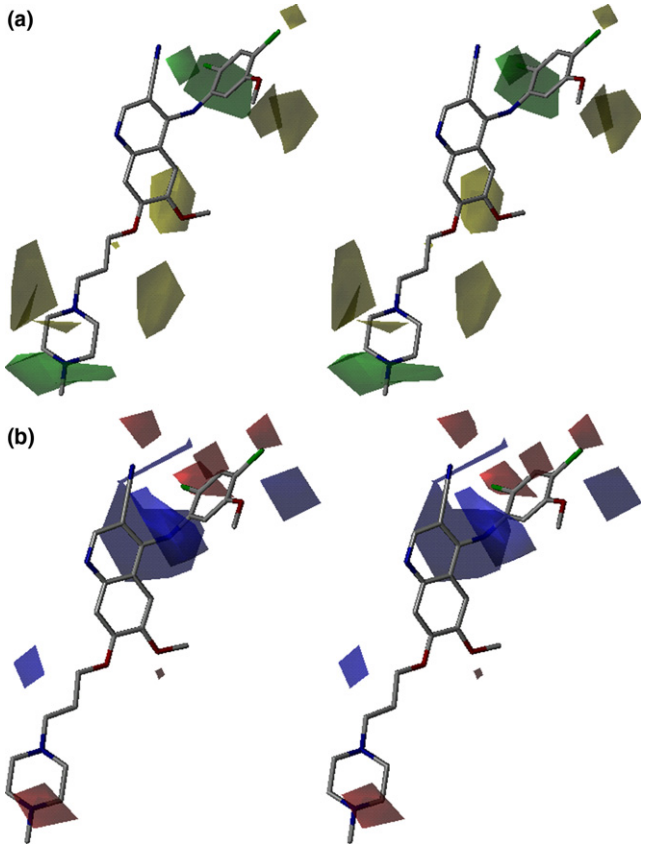


Figure 1. Stereoview of the contour plots (STDEV*COEFF) of the CoMFA (shape-based alignment): (a) steric fields; green contours indicate regions where bulky groups increase activity, while yellow contours indicate regions where bulky groups decrease activity, and (b) electrostatic fields; blue contours indicate regions where electropositive groups increase activity, while red contours indicate regions where electronegative groups increase activity. Potent c-Src inhibitor **77** is displayed in the background for reference.

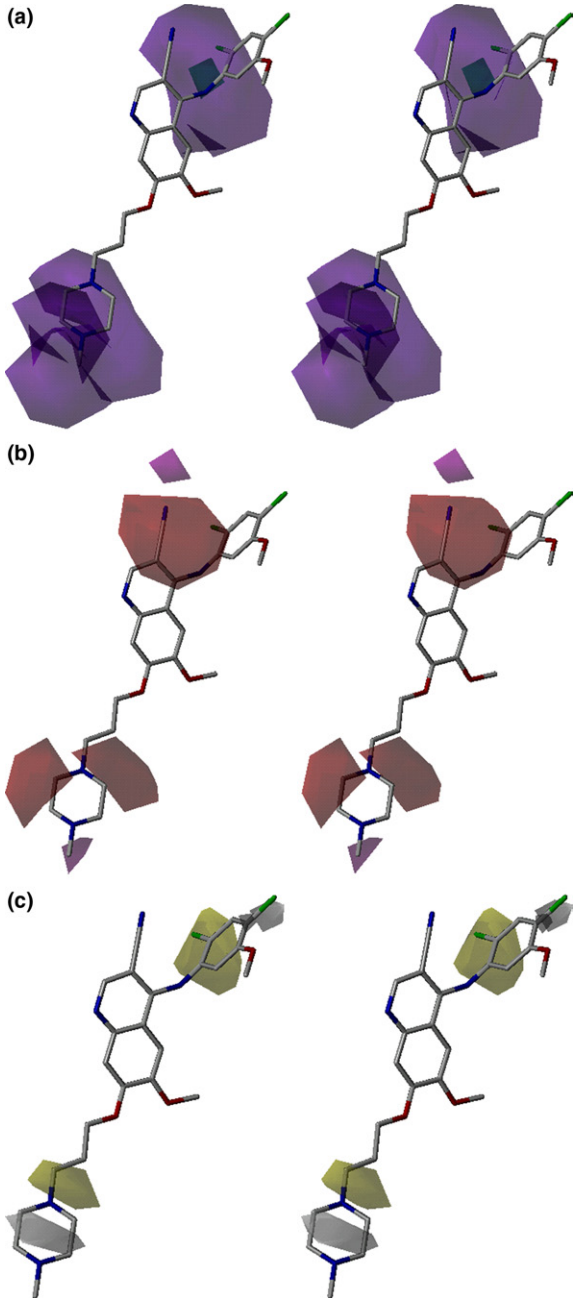


Figure 2. Stereoview of the contour plots of the CoMSIA: (a) hydrogen bond donor fields (cyan, favored; purple, disfavored), (b) hydrogen bond acceptor fields (magenta, favored; red, disfavored), and (c) hydrophobic fields (yellow, favored; white, disfavored). Potent c-Src inhibitor **77** is displayed in the background for reference.

The cyan contour near the aniline NH indicates that hydrogen bond donor group at this position enhances the activity (Fig. 2a). The magenta contour near the cyano group (Fig. 2b), which matches with the red region in the CoMFA electrostatic contour plots (Fig. 1b), corroborates that an acceptor group at this position is required for higher activity. The CoMSIA hydrophobic contour map is shown in Figure 2c. The analysis of CoMSIA steric, electrostatic, and hydrophobic contour plots reveals that substituents that combine both electronegative and lipophillic properties are preferred at the 2′-position.

The predictive power of CoMFA and CoMSIA 3D-QSAR models was evaluated using a test set of 14 molecules. The PLS statistics of both CoMFA and CoMSIA 3D-QSAR models indicate that CoMFA is somewhat better than CoMSIA. In both models, the predicted values fall close to the actual pIC_{50} values, not deviating by more than 1 logarithmic unit (Fig. 3, Tables 1 and 2) except for the two training set molecules. Molecules **42** and **87** are outliers (residual more than 1.0) in CoMSIA.

3.3. Homology model of c-Src (the active form)

The activation loop, a segment of kinase domain, plays a key role in catalytic regulation. Phosphorylation of Tyr/Ser/Thr residue on this loop is required for full activity of many kinases. It adopts distinct conformations in active and inactive forms in a phosphorylation dependent manner; the residue that undergoes phosphorylation is exposed in active form while buried in

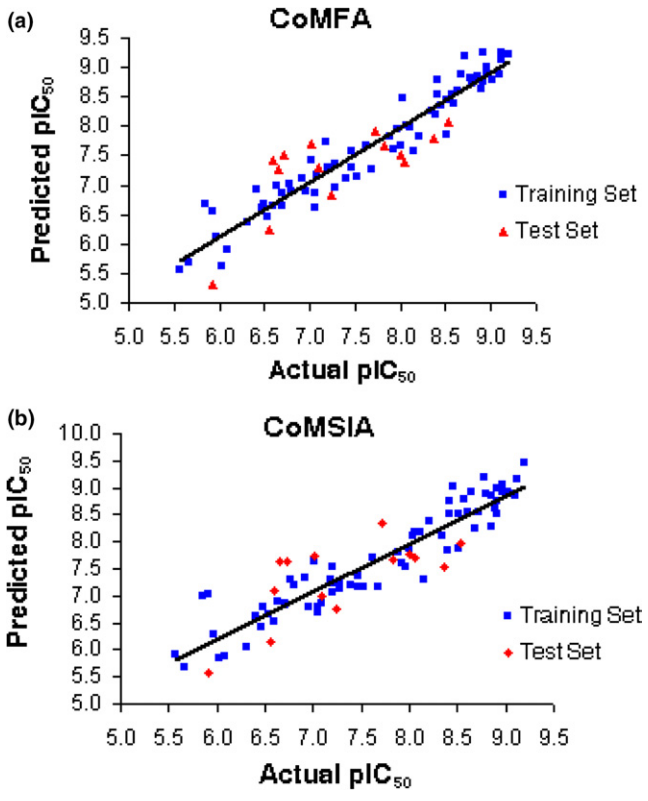


Figure 3. Actual versus predicted pIC_{50} of training (blue) and test (red) set molecules for (a) CoMFA and (b) CoMSIA 3D-QSAR models.

case of inactive conformation. In addition, differences exist in the inter-lobe orientation, the disposition helix C in the N-lobe, and the conformation of N-terminal part of the activation loop (the DFG motif). An ion–pair interaction between conserved amino acids Lys (from the glycine-rich loop) and Glu (located on C helix) is a characteristic feature of active conformation of protein kinases. This ion–pair interaction is absent in inactive conformations of many protein kinases such as c-Src and CDK, but not in the imatinib complex of c-Abl.[20] Imatinib, which is currently used for the treatment of CML, is found to be inactive against 15 clinically relevant Bcr–Abl mutants.[4] All these mutants are not directly involved in the imatinib binding except Thr315Ile; this mutant, which is located near the adenosine binding pocket, might block imatinib binding.[4,20] The specificity of imatinib is connected with its ability to bind to precise inactive conformation of c-Abl. The shapes and sizes of kinase-specific pockets are very distinct in the active and inactive forms.[20] Additionally, structural information suggests that imatinib cannot bind to Src kinases, although most of the residues in the ATP binding pocket are conserved in these kinases. Furthermore, studies indicate that the use of Src–Abl dual inhibitors might be useful for the treatment of imatinib-resistant CML.[4] For example, BMS-354825, a potent Src–Abl dual inhibitor, is found to inhibit 14 of 15 imatinib-resistant Bcr–Abl mutants.[4c] These observations prompted us to study the binding mode of one of the potent Src–Abl dual inhibitors (**77**) in the active sites of both c-Src and c-Abl.

The crystal structure of ANP bound to the inactive form of c-Src (2SRC.pdb) is available,[19] while that of an active form is not determined. The crystal structures of c-Abl in which the activation loop resembles the active and the inactive conformations with small-molecule inhibitors (PD173955 and imatinib, respectively) are also reported.[20] The crystal structure of the active form of Lck is accessible as well.[12] Both c-Src and Lck belong to Src family and share high sequence homology in their kinase domains; the sequence identity is about 67% (Scheme 2). In addition, the activation loop of c-Src in the active form might adopt a conformation very similar to the one observed in the active forms of Lck and IRT kinases.[19] Therefore, the crystal structure of the kinase domain of Lck was used as the template for building a homology model of an active form of c-Src. The overlay of the active and inactive forms of c-Src is shown in Figure 4. In the active form of the c-Src (homology model), the C helix is rotated inward facilitating formation of an ion–pair interaction between conserved Lys295 and Glu310. As a result, the volume of the kinase-specific pocket turns out to be smaller than that observed in the inactive form. On the other hand, the active sites of both c-Src and c-Abl, in which the activation loop resembles that of an active kinase, are almost similar.

3.4. Docking analysis

Molecule **77** was docked inside the active site of the inactive form of c-Src (2SRC.pdb) using FlexX. The simulated binding pose is consistent with the crystal

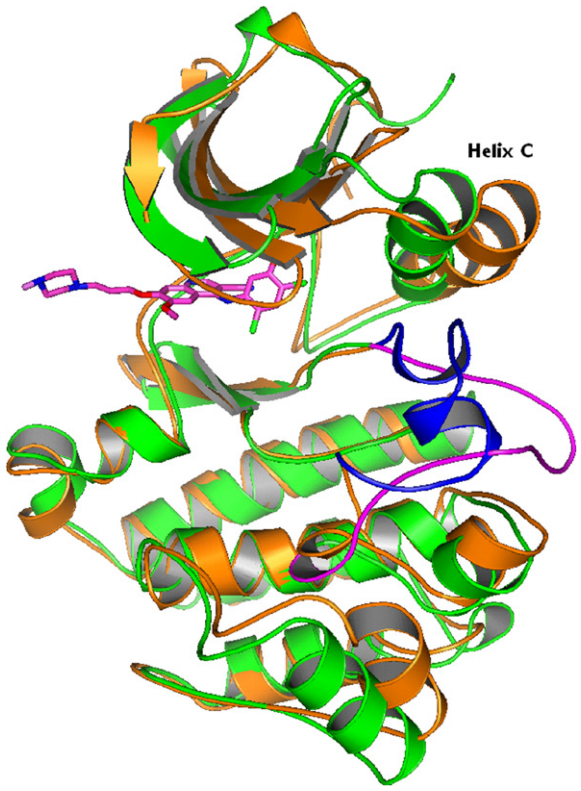


Figure 4. The overlay of ribbon representations of the active (homology model, green) and inactive (crystal structure, orange, 2SRC.pdb) forms of c-Src. The activation loops are colored magenta (in active) and blue (in inactive) for clarity. The simulated binding pose of **77** inside the active site is also shown (in magenta C atoms).

structure bound conformations of quinazolines inside the active sites of EGFR, P38, and CDK;[21] and similar to the docked conformations of cyanoquinolines in the active site of EGFR.[22] The binding pose of **77** inside the active site of c-Src is shown in Figure 5. Hydrogen bonding interactions with the backbone NH of Met341 and the side chain OH group of Thr338 are observed. In addition, a C–H $\cdots$ O hydrogen bonding interaction between the quinoline C2 and the carbonyl O of Glu339 is present. The cyano group probably mimics the role of structural water present in the crystal structures of quinazolines bound to EGFR and P38. The aniline group is positioned in the kinase-specific pocket, while the 4-methylpiperazine is oriented toward the solvent region. Molecule **77** adopts virtually similar binding pose in the active form as well (c-Src homology model, Fig. 4). In addition, the orientation and the hydrogen bonding interactions of **77** inside the ATP binding pockets of c-Abl kinase (1M52.pdb, 1IEP.pdb, respectively; figures not shown) are virtually similar to those present in the c-Src active site, indicating that the binding of **77** would be insensitive to whether the activation loop corresponds to the active or inactive conformation.

3.5. QSAR versus docking

QSAR analyses provide quantitative estimation of activity and insight into key structural elements to design

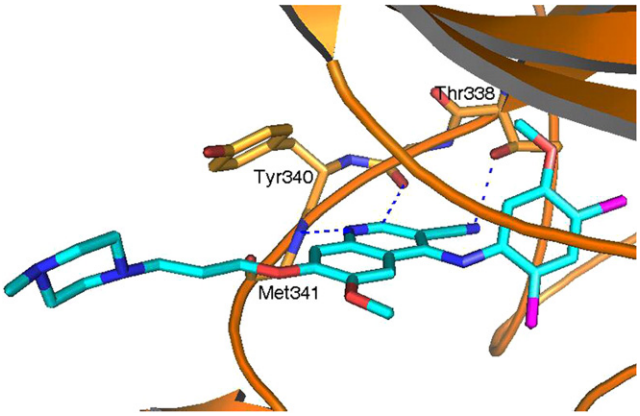


Figure 5. The docked conformation of **77** inside the active site of c-Src (2SRC.pdb, inactive form). The hydrogen bonds are shown in blue broken lines. The ligand is shown in cyan, which is located in the adenosine binding pocket by forming hydrogen bonding interactions with the backbone NH of Met341, the carbonyl O of Glu339, and the side chain OH group of Thr338.

potential lead candidates, while docking studies will be useful for planning focused SAR studies. In the absence of crystal structure data, docking methods can complement X-ray structure determination. All these tools aid in lead optimization. Lead optimization is an iterative process in which structural modifications are introduced to improve activity and potency using SAR studies, X-ray structural information of potent ligand–target receptor complexes, and/or docking studies. In this connection, the crystal structure determination of **77** with c-Src or c-Abl might be worth investigating to see if the CN group binds covalently to the conserved Thr in the active site similar to that observed in the crystal structure of 2-cyanopyrrolidine derivative bound to the active site Ser of dipeptidyl peptidase IV (DPP-IV).[23]

The results of 3D-QSAR and docking studies validate each other. For instance, the substituents that combine both electronegative and lipophillic properties at the 2′-, 4′-, and 5′-positions on the aniline ring are located in the hydrophobic pocket formed by Met314, Val323, Ile336, Leu393, Ala403, and Phe405. The structure based studies, however, provided additional insights into inhibitor–enzyme interactions as well. The aniline NH is not involved in any direct hydrogen bonding interaction with the receptor. The docking studies provide a feasible explanation why hydrophilic groups are desirable at the 7-position, and also suggest that the size of the substituent at the 4′-position might be crucial for both kinase selectivity and Src–Abl dual inhibitory activity.

4. Conclusions

CoMFA and CoMSIA 3D-QSAR models were developed for quinazoline and quinoline derivatives as c-Src kinase inhibitors. The docked conformation of **77** was used as the template for aligning the rest of the molecules to it. The CoMFA model demonstrated better predictive ability than the CoMSIA. A high degree of

sequence identity between c-Src and Lck kinase domains enabled us to construct a reasonably good homology model of the active form of c-Src. The kinase-specific pocket-size of the active form is smaller than that of the inactive form. In addition, the ATP binding pockets of the active forms of both c-Src and c-Abl are almost similar. The binding pose and hydrogen bonding interactions of one of the potent Src–Abl dual inhibitors (**77**) are similar in both c-Src and c-Abl. The 3D-QSAR results are consistent with the docking studies. This information will be useful in the design of novel anticancer drug candidates.

Acknowledgments

We acknowledge Dr. N. Selvakumar and Dr. R. Rajagopalan for their constant support and encouragement.

References and notes

1. (a) Capdeville, R.; Buchdunger, E.; Zimmermann, J.; Matter, A. *Nat. Rev. Drug Discovery* **2002**, *1*, 493; (b) Dowell, J.; Minna, J. D.; Kirkpatrick, P. *Nat. Rev. Drug Discovery* **2005**, *4*, 13; (c) Muhsin, M.; Graham, J.; Kirkpatrick, P. *Nat. Rev. Drug Discovery* **2003**, *2*, 515.
2. Biscardi, J. S.; Tice, D. A.; Parsons, S. J. *Adv. Cancer Res.* **1999**, *76*, 61.
3. Pendergast, A. M.; Gishizky, M. L.; Havlik, M. H.; Witte, O. N. *Mol. Cell. Biol.* **1993**, *13*, 1728.
4. (a) Tatton, L.; Morley, G. M.; Chopra, R.; Khwaja, A. *J. Biol. Chem.* **2003**, *278*, 4847; (b) Warmuth, M.; Simon, N.; Mitina, O.; Mathes, R.; Fabbro, D.; Manley, P. W.; Buchdunger, E.; Forster, K.; Moarefi, I.; Hallek, M. *Blood* **2003**, *101*, 664; (c) Shah, N. P.; Tran, C.; Lee, F. Y.; Chen, P.; Norris, D.; Sawyers, C. L. *Science* **2004**, *305*, 399; (d) Daub, H.; Specht, K.; Ullrich, A. *Nat. Rev. Drug Discovery* **2004**, *3*, 1001.
5. Dow, R. L.; Bechle, B. M.; Chou, T. T.; Goddard, C.; Larson, E. R. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 1007.
6. Hanke, J. H.; Gardner, J. P.; Dow, R. L.; Changelian, P. S.; Brissette, W. H.; Weringer, E. J.; Pollok, B. A.; Connelly, P. A. *J. Biol. Chem.* **1996**, *271*, 695.
7. Myers, M. R.; Setzer, N. N.; Spada, A. P.; Zulli, A. L.; Hsu, C.-Y. J.; Zilberstein, A.; Johnson, S. E.; Hook, L. E.; Jacoski, M. V. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 417.
8. (a) Thompson, A. M.; Rewcastle, G. W.; Boushelle, S. L.; Hartl, B. G.; Kraker, A. J.; Lu, G. H.; Batley, B. L.; Panek, R. L.; Showalter, H. D. H.; Denny, W. A. *J. Med. Chem.* **2000**, *43*, 3134; (b) Missbach, M.; Altmann, E.; Widler, L.; Susa, M.; Buchdunger, E.; Mett, H.; Meyer, T.; Green, J. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 945; (c) Klutchko, S. R.; Hamby, J. M.; Boschelli, D. H.; Wu, Z. P.; Kraker, A. J.; Amar, A. M.; Hartl, B. G.; Shen, C.; Klohs, W. D.; Steinkampf, R. W.; Driscoll, D. L.; Nelson, J. M.; Elliott, W. L.; Roberts, B. J.; Stoner, C. L.; Vincent, P. W.; Dykes, D. J.; Panek, R. L.; Lu, G. H.; Major, T. C.; Dahring, T. K.; Hallak, H.; Bradford, L. A.; Showalter, H. D. H.; Doherty, A. M. *J. Med. Chem.* **1998**, *41*, 3276.
9. (a) Wang, Y. D.; Miller, K.; Boschelli, D. H.; Ye, F.; Wu, B.; Floyd, M. B.; Powell, D. W.; Wissner, A.; Weber, J. M.; Boschelli, F. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 2477; (b) Boschelli, D. H.; Wang, Y. D.; Ye, F.; Wu, B.; Zhang, N.; Dutia, M.; Powell, D. W.; Wissner, A.; Arndt, K.; Weber, J. M.; Boschelli, F. *J. Med. Chem.* **2001**, *44*, 822; (c) Boschelli, D. H.; Ye, F.; Wang, Y. D.; Dutia, M.; Johnson, S. L.; Wu, B.; Miller, K.; Powell, D. W.; Yaczko, D.; Young, M.; Tischler, M.; Arndt, K.; Discafani, C.; Etienne, C.; Gibbons, J.; Grod, J.; Lucas, J.; Weber, J. M.; Boschelli, F. *J. Med. Chem.* **2001**, *44*, 3965; (d) Boschelli, D. H.; Wang, Y. D.; Johnson, S.; Wu, B.; Ye, F.; Barrios Sosa, A. C.; Golas, J. M.; Boschelli, F. *J. Med. Chem.* **2004**, *47*, 1599.
10. Cramer, R. D., III; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959.
11. Klebe, G.; Abraham, U.; Mietzner, T. *J. Med. Chem.* **1994**, *37*, 4130.
12. Yamaguchi, H.; Hendrickson, W. A. *Nature* **1996**, *384*, 484.
13. SYBYL 6.92 Molecular Modeling software, Tripos Associates, Inc.: 1669, South Hanley Road, Suite 303, St. Louis, Missouri, MO 63144-2913, USA.
14. Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 490.
15. (a) Dunn, W. J., III; Wold, S.; Edlund, U.; Hellberg, S.; Gasteiger, J. *Quant. Struct.-Act. Relat.* **1984**, *3*, 131; (b) Geladi, P. *J. Chemom.* **1988**, *2*, 231.
16. Cramer, R. D., III; Bunce, J. D.; Patterson, D. E. *Quant. Struct.-Act. Relat.* **1988**, *7*, 18.
17. Higgins, D. G.; Sharpe, P. M. *Gene* **1988**, *73*, 237.
18. Guex, N.; Peitsch, M. C. *Electrophoresis* **1997**, *18*, 2714.
19. Xu, W.; Doshi, A.; Lei, M.; Eck, M. J.; Harrison, S. C. *Mol. Cell* **1999**, *3*, 629.
20. Nagar, B.; Bornmann, W. G.; Pellicena, P.; Schindler, T.; Veach, D. R.; Miller, W. T.; Clarkson, B.; Kuriyan, J. *Cancer Res.* **2002**, *62*, 4236.
21. (a) Stamos, J.; Sliwkowski, M. X.; Eigenbrot, C. *J. Biol. Chem.* **2002**, *277*, 46265; (b) Shewchuk, L.; Hassell, A.; Wisely, B.; Rocque, W.; Holmes, W.; Veal, J.; Kuyper, L. F. *J. Med. Chem.* **2000**, *43*, 133.
22. (a) Wissner, A.; Berger, D. M.; Boschelli, D. H.; Floyd, M. B., Jr.; Greenberger, L. M.; Gruber, 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.; Zhang, N. *J. Med. Chem.* **2000**, *43*, 3244; (b) Akula, N.; Bhalla, J.; Sridhar, J.; Pattabiraman, N. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 3397.
23. Oefner, C.; D'Arcy, A.; Mac Sweeney, A.; Pierau, S.; Gardiner, R.; Dale, G. E. *Acta Crystallogr., Sect. D* **2003**, *59*, 1206.