

3D QSAR Models Built on Structure-Based Alignments of Abl Tyrosine Kinase Inhibitors

Federico Falchi,^[a] Fabrizio Manetti,^[a] Fabio Carraro,^[b] Antonella Naldini,^[b] Giovanni Maga,^[c] Emmanuele Crespan,^[c] Silvia Schenone,^{*,[d]} Olga Bruno,^[d] Chiara Brullo,^[d] and Maurizio Botta^{*,[a]}

The X-ray crystallographic coordinates of the Abl tyrosine kinase domain in its active, inactive, and Src-like inactive conformations were used as targets to simulate the binding mode of a large series of pyrazolo[3,4-d]pyrimidines (known Abl inhibitors) by means of GOLD software. Receptor-based alignments provided by molecular docking calculations were submitted to a GRID-GOLPE protocol to generate 3D QSAR models. Analysis of the results showed that the models based on the inactive and Src-like inactive conformations had very poor statistical parameters, whereas the sole model based on the active conformation of Abl was characterized by significant

internal and external predictive ability. Subsequent analysis of GOLPE PLS pseudo-coefficient contour plots of this model gave us a better understanding of the relationships between structure and affinity, providing suggestions for the next optimization process. On the basis of these results, new compounds were designed according to the hydrophobic and hydrogen bond donor and acceptor contours, and were found to have improved enzymatic and cellular activity with respect to parent compounds. Additional biological assays confirmed the important role of the selected compounds as inhibitors of cell proliferation in leukemia cells.

Introduction

Chronic myelogenous leukemia (CML) is a myeloproliferative disorder that is usually characterized (in the leukemic cells of more than 95% of patients) by the presence of the Philadelphia chromosome, resulting from a translocation between chromosomes 9 and 22. This translocation results in the transfer of the Abelson (*abl*) oncogene to the breakpoint cluster region (*bcr*) of chromosome 22, leading to the fused *bcr-abl* gene and to the expression of a fusion protein (Bcr-Abl) with deregulated tyrosine kinase activity.^[1]

The kinase domains (highly conserved structures of about 300 residues) of the fusion protein (Bcr-Abl) and *c-Abl* are identical in sequence and catalyze the transfer of the γ -phosphate group of ATP to the hydroxy group of a tyrosine residue of the peptide substrate. The structure of the kinase domain can be divided into two lobes. At their interface, a number of highly conserved residues form the ATP binding pocket and the catalytic machinery.

Several conformations of the Abl tyrosine kinase domain have been identified and are referred to as the inactive, active, and Src-like inactive conformations.^[2] The activation loop, a centrally located regulatory element of the protein kinase, is the region of the kinase domain that differs markedly between the active and inactive states. In the active state, the activation loop is in an extended or open conformation with an aspartic residue (of the strictly conserved Asp-Phe-Gly motif at the N-terminal base of the activation loop) positioned so as to interact properly with the magnesium ion coordinated by ATP. The remaining residues of the loop are positioned away from the catalytic center so that the C-terminal portion of this loop provides a support for substrate binding. On the other hand, the

activation loop of the inactive state switches to a more closed conformation.^[3]

Imatinib, the first line therapy for CML treatment, binds the inactive conformation of the kinase domain.^[4] Nevertheless, resistance to this drug is frequently observed mainly because of mutations in the kinase domain of Bcr-Abl.^[5] For this reason, new Abl inhibitors able to bind both the active and inactive conformations, or that preferentially bind the active form of the protein, were discovered.^[6] In this context, our previous search of novel multitarget kinase inhibitors led to the design, synthesis, biochemical evaluation, and preliminary molecular modeling simulations of a large series of pyrazolo[3,4-d]pyrimi-

[a] Dr. F. Falchi, Dr. F. Manetti, Prof. Dr. M. Botta
Dipartimento Farmaco Chimico Tecnologico
Università degli Studi di Siena
Via Alcide de Gasperi 2, 53100 Siena (Italy)
Fax: (+39) 0577-234333
E-mail: botta@unisi.it

[b] Prof. Dr. F. Carraro, Prof. Dr. A. Naldini
Dipartimento di Fisiologia, Sezione di Neuroimmunofisiologia
Università degli Studi di Siena
Via Aldo Moro 2, 53100 Siena (Italy)

[c] Dr. G. Maga, Dr. E. Crespan
IGM-CNR, Via Abbiategrasso 207, 27100 Pavia (Italy)

[d] Prof. Dr. S. Schenone, Prof. Dr. O. Bruno, Dr. C. Brullo
Dipartimento di Scienze Farmaceutiche
Università degli Studi di Genova
Viale Benedetto XV 3, 16132 Genova (Italy)
Fax: (+39) 010-3538358
E-mail: schensil@unige.it

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/cmdc.200800441>.

dines.^[7] These compounds were used to generate 3D QSAR models of Abl inhibitors as described in this paper, which were then used to rationalize relationships between the structural properties of compounds and their affinity, and to guide us toward the synthesis of next generation compounds with inhibitory activity toward Abl better than that of parent compounds.

In particular, the availability of Abl crystal structures in the active, inactive, and Src-like inactive conformations (PDB^[8] entries 2GQG,^[9] 1EP,^[10] and 2G2I,^[2] respectively) in complex with various inhibitors allowed us to apply a structure-based approach to generate alignments of Abl inhibitors within the enzyme binding pocket. The three-dimensional structure of the target protein and a docking protocol based on GOLD software^[11] were used to guide the selection of inhibitor alignment. Aligned inhibitors were in turn used to build and validate three structure-based 3D QSAR models (one for each Abl conformation) for Abl inhibitors by means of the programs GRID^[12] and GOLPE.^[13] The application of this approach was based on the fact that recent studies described the combination of receptor-based methods and 3D QSAR analysis as a useful tool for the design and prediction of bioactive compounds.^[14]

Results and Discussion

Chemistry

Compounds **86–88** and **92–96** (Table 1) were prepared as shown in Scheme 1. The 1-(2-hydroxy-2-phenylethyl)-6-thioxo-1,5,6,7-tetrahydro-4H-pyrazolo-[3,4-d]pyrimidin-4-one **97**, prepared according to our procedure,^[15] was alkylated at the C6 sulfur atom using the appropri-

Table 1. Structure of pyrazolo-pyrimidine derivatives and their affinity toward Abl.

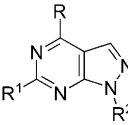
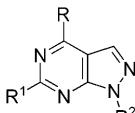
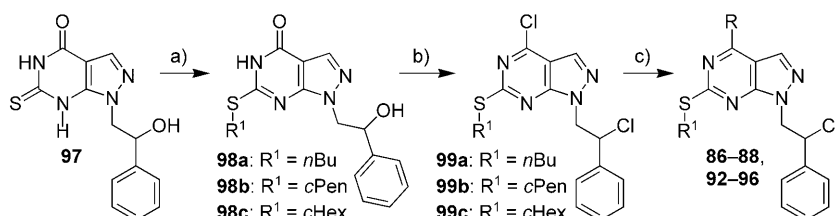
Compd ^[a]	R			K _i [μM] ^[b]
		R ¹	R ²	
1 ^[d]	NHCH ₂ C ₆ H ₅	H	CH ₂ CHClC ₆ H ₅	0.37
2 ^[c]	NHCH ₂ C ₆ H ₄ mF	H	CH ₂ CHClC ₆ H ₅	0.04
3 ^[d]	NHCH ₂ C ₆ H ₄ oF	H	CH ₂ CHClC ₆ H ₅	0.73
4 ^[d]	NHCH ₂ C ₆ H ₅	H	CH ₂ CHClC ₆ H ₄ pF	0.44
5 ^[d]	NHCH ₂ C ₆ H ₄ pF	H	CH ₂ CHClC ₆ H ₄ pF	0.20
6 ^[c]	NHCH ₂ C ₆ H ₄ mF	H	CH ₂ CHClC ₆ H ₄ pF	0.10
7 ^[c]	NHCH ₂ C ₆ H ₄ oF	H	CH ₂ CHClC ₆ H ₄ pF	0.02
8 ^[c]	NHCH ₂ C ₆ H ₄ mCl	H	CH ₂ CHClC ₆ H ₄ pF	0.12
9 ^[d]	NHCH ₂ C ₆ H ₄ oCl	H	CH ₂ CHClC ₆ H ₄ pF	0.20
10 ^[c]	NHCH ₂ C ₆ H ₄ pF	H	CH ₂ CHClC ₆ H ₄ pCl	0.08
11 ^[c]	NHCH ₂ C ₆ H ₄ mCl	H	CH ₂ CHClC ₆ H ₄ pCl	0.08
12 ^[d]	NHC ₃ H ₇	H	CHCHC ₆ H ₅	7.4
13 ^[c]	NH-cyclopropyl	H	CHCHC ₆ H ₅	50
14 ^[c]	NHC ₄ H ₉	H	CHCHC ₆ H ₅	0.40
15 ^[c]	NH(CH ₂) ₂ OC ₂ H ₅	H	CHCHC ₆ H ₅	12
16 ^[c]	1-pyrrolidinyl	H	CHCHC ₆ H ₅	3.9
17 ^[d]	1-piperidinyl	H	CHCHC ₆ H ₅	0.20
18 ^[d]	4-morpholinyl	H	CHCHC ₆ H ₅	7.8
19 ^[c]	NH-cyclohexyl	H	CHCHC ₆ H ₅	0.80
20 ^[c]	1-hexahydroazepinyl	H	CHCHC ₆ H ₅	0.30
21 ^[c]	N(C ₂ H ₅) ₂	SC ₃ H ₇	CH ₂ CHClC ₆ H ₅	0.28
22 ^[c]	NHC ₃ H ₇	SC ₃ H ₇	CH ₂ CHClC ₆ H ₅	0.25
23 ^[d]	NHC ₃ H ₇	N(CH ₃) ₂	CH ₂ CHClC ₆ H ₅	0.32
24 ^[d]	NHCH ₂ C ₆ H ₅	N(CH ₃) ₂	CH ₂ CHClC ₆ H ₅	0.16
25 ^[c]	NHC ₆ H ₅	SCH ₃	CH ₂ CHClC ₆ H ₅	0.40
26 ^[c]	NHC ₃ H ₇	H	CH ₂ CHOHC ₆ H ₅	50
27 ^[d]	1-piperidinyl	H	CH ₂ CHOHC ₆ H ₅	50
28 ^[d]	NHC ₄ H ₉	H	CH ₂ CHOHC ₆ H ₅	50
29 ^[c]	4-morpholinyl	H	CH ₂ CHOHC ₆ H ₅	50
31 ^[d]	NHC ₄ H ₉	SC ₂ H ₅	CH ₂ CHClC ₆ H ₅	0.30
35 ^[d]	NHC ₃ H ₇	SCH ₃	CH ₂ CHClC ₆ H ₅	4.8
36 ^[c]	N(C ₂ H ₅) ₂	SCH ₃	CH ₂ CHClC ₆ H ₅	0.40
37 ^[d]	NH(CH ₂) ₂ OC ₂ H ₅	SCH ₃	CH ₂ CHClC ₆ H ₅	1.5
40 ^[c]	NHCH ₂ C ₆ H ₅	SCH ₃	CH ₂ CHClC ₆ H ₅	0.30
42 ^[d]	4-(2,6-dimethyl)morpholinyl	SCH ₃	CH ₂ CHClC ₆ H ₅	3.5
44 ^[c]	4-morpholinyl	SC ₂ H ₅	CHCHC ₆ H ₅	9.0
48 ^[c]	1-piperidinyl	SCH ₃	CHCHC ₆ H ₅	1.3
49 ^[c]	4-morpholinyl	SCH ₃	CHCHC ₆ H ₅	6.8
50 ^[c]	4-(2,6-dimethyl)morpholinyl	SCH ₃	CHCHC ₆ H ₅	4.7
51 ^[d]	NHC ₃ H ₇	H	CH ₂ CHBrC ₆ H ₅	4.8
52 ^[d]	NH-cyclopropyl	H	CH ₂ CHBrC ₆ H ₅	4.8
53 ^[d]	NHCH(CH ₃) ₂	H	CH ₂ CHBrC ₆ H ₅	2.1
54 ^[c]	NHC ₄ H ₉	H	CH ₂ CHBrC ₆ H ₅	0.70
55 ^[d]	NH(CH ₂) ₂ OC ₂ H ₅	H	CH ₂ CHBrC ₆ H ₅	1.6
56 ^[c]	1-pyrrolidinyl	H	CH ₂ CHBrC ₆ H ₅	0.90
57 ^[d]	1-piperidinyl	H	CH ₂ CHBrC ₆ H ₅	1.2
58 ^[d]	4-morpholinyl	H	CH ₂ CHBrC ₆ H ₅	0.20
59 ^[d]	NH-cyclohexyl	H	CH ₂ CHBrC ₆ H ₅	0.50
60 ^[c]	1-hexahydroazepinyl	H	CH ₂ CHBrC ₆ H ₅	4.9
61 ^[d]	NHC ₃ H ₇	SCH ₃	CH ₂ CHClC ₆ H ₄ pF	0.57
62 ^[d]	NHC ₄ H ₉	SCH ₃	CH ₂ CHClC ₆ H ₄ pF	0.11
63 ^[d]	4-morpholinyl	SCH ₃	CH ₂ CHClC ₆ H ₄ pF	0.97
65 ^[c]	NHC ₃ H ₇	SCH ₃	CH ₂ CHClC ₆ H ₄ pCl	0.40
66 ^[d]	NHC ₄ H ₉	SCH ₃	CH ₂ CHClC ₆ H ₄ pCl	0.39
67 ^[d]	4-morpholinyl	SCH ₃	CH ₂ CHClC ₆ H ₄ pCl	0.50
68 ^[c]	NHC ₃ H ₇	SCH ₃	CH ₂ CHBrC ₆ H ₅	0.88
69 ^[d]	NHC ₄ H ₉	SCH ₃	CH ₂ CHBrC ₆ H ₅	0.32
70 ^[c]	NH(CH ₂) ₂ OC ₂ H ₅	SCH ₃	CH ₂ CHBrC ₆ H ₅	0.55
71 ^[c]	1-pyrrolidinyl	SCH ₃	CH ₂ CHBrC ₆ H ₅	1.4
72 ^[d]	1-piperidinyl	SCH ₃	CH ₂ CHBrC ₆ H ₅	0.88
73 ^[c]	NHCH ₂ C ₆ H ₅	SCH ₃	CH ₂ CHBrC ₆ H ₅	0.19

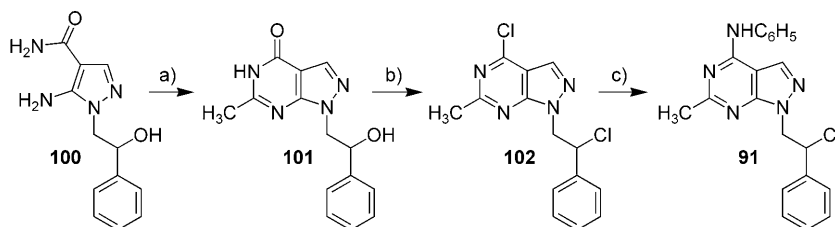
Table 1. (Continued)

Compd ^[a]	R				<i>K_i</i> [μM] ^[b]
		R ¹	R ²		
77 ^[d]	N(C ₂ H ₅) ₂	SCH ₃	CH ₂ CHClC ₆ H ₄ pCl	0.30	
78 ^[d]	NHC ₆ H ₄ <i>m</i> F	SCH ₃	CH ₂ CHClC ₆ H ₄ pF	0.22	
79 ^[c]	NHCH ₂ C ₆ H ₄ pF	SCH ₃	CH ₂ CHClC ₆ H ₄ pF	0.22	
80 ^[d]	NHCH ₂ C ₆ H ₄ oF	SCH ₃	CH ₂ CHClC ₆ H ₄ pF	0.34	
81 ^[c]	N(C ₂ H ₅) ₂	SC ₂ H ₅	CH ₂ CHClC ₆ H ₅	0.33	
82 ^[d]	NHC ₄ H ₉	SC ₃ H ₇	CH ₂ CHClC ₆ H ₅	0.52	
83 ^[c]	NHCH ₂ C ₆ H ₅	SC ₃ H ₇	CH ₂ CHClC ₆ H ₅	0.10	
84 ^[e]	NHCH ₂ C ₆ H ₄ oCl	SCH ₃	CH ₂ CHClC ₆ H ₅	0.09 [31, 6.3] ^[f]	
85 ^[e]	NHC ₆ H ₄ <i>m</i> Br	SCH ₃	CH ₂ CHClC ₆ H ₅	0.06 [30, 21] ^[f]	
86 ^[e]	NHCH ₂ C ₆ H ₅	S-cyclopentyl	CH ₂ CHClC ₆ H ₅	0.06 [19, 8.8] ^[f]	
87 ^[e]	NHCH ₂ CH ₂ C ₆ H ₅	S-cyclopentyl	CH ₂ CHClC ₆ H ₅	0.07 [19, 8.4] ^[f]	
88 ^[e]	NHC ₆ H ₄ <i>m</i> Cl	S-cyclohexyl	CH ₂ CHClC ₆ H ₅	0.07 [21, 25] ^[f]	
89 ^[e]	NHCH ₂ C ₆ H ₄ oF	H	CH ₂ CHClC ₆ H ₄ pBr	0.44 [7.9, 4.2] ^[f]	
90 ^[e]	NHCH ₂ C ₆ H ₄ <i>m</i> F	H	CH ₂ CHClC ₆ H ₄ pBr	0.02 [11, 6.0] ^[f]	
91 ^[e]	NHC ₆ H ₅	CH ₃	CH ₂ CHClC ₆ H ₅	0.02 [32, 25] ^[f]	
92 ^[e]	NHC ₃ H ₇	SC ₄ H ₉	CH ₂ CHClC ₆ H ₅	0.02 [32, 9.0] ^[f]	
93 ^[e]	NHC ₆ H ₄ <i>m</i> F	SC ₄ H ₉	CH ₂ CHClC ₆ H ₅	0.04 [18, 3.8] ^[f]	
94 ^[e]	NHC ₆ H ₄ <i>m</i> Cl	SC ₄ H ₉	CH ₂ CHClC ₆ H ₅	0.09 [19, 8.9] ^[f]	
95 ^[e]	NHCH ₂ C ₆ H ₅	SC ₄ H ₉	CH ₂ CHClC ₆ H ₅	0.17 [12, 15] ^[f]	
96 ^[e]	NHCH ₂ CH ₂ C ₆ H ₅	SC ₄ H ₉	CH ₂ CHClC ₆ H ₅	0.24 [21, 15] ^[f]	

[a] Compounds omitted in this table are reported in table S1 in the Supporting Information. [b] Affinity values were determined as previously described.^[7] [c] Training set compounds. [d] Test set compounds. [e] Synthesized on the basis of QSAR suggestions. [f] Inhibitory activity toward human leukemia KU-812 and K-562 cell lines, expressed as IC₅₀ values, which represent the mean of five experiments determined as previously described,^[7b,c] using PP2 as the reference compound; PP2 showed IC₅₀ values of 15 and 12 μM toward KU-812 and K-562 cells, respectively.



Scheme 1. Synthesis of compounds **86–88** and **92–96**. Reagents and conditions: a) RBr, K₂CO₃, DMF, RT or 80 °C; b) POCl₃/DMF, CHCl₃, reflux, 8 h; c) method A: amines, anhydrous toluene, RT, 48 h; method B: anilines, absolute EtOH, reflux, 4 h.



Scheme 2. Synthesis of compound **91**. Reagents and conditions: a) EtONa, CH₃COOEt, absolute EtOH, reflux, 6 h; b) POCl₃/DMF, CHCl₃, reflux, 12 h; c) method B: aniline, absolute EtOH, reflux, 4 h.

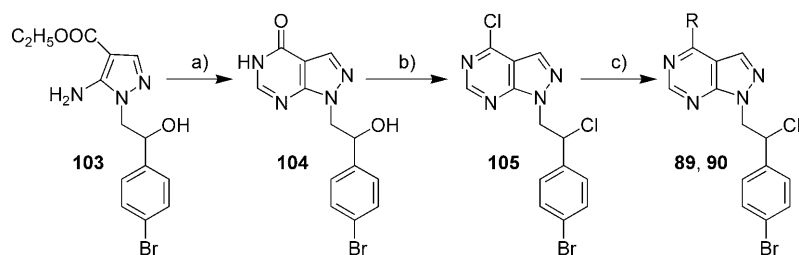
ate alkyl bromide and anhydrous K₂CO₃ in anhydrous *N,N*-dimethylformamide (DMF). The 6-alkylthio derivatives **98a–c** were in turn chlorinated with the Vilsmeier complex (POCl₃/

DMF, 1:1, 10 equiv) in CHCl₃ at reflux for 8 h to obtain the dichloro analogues **99a–c**. The crude products were purified in good yield by chromatography on a Florisil column. Finally, reaction of **99a** and **99b** with an excess of various amines in anhydrous toluene at RT for two days afforded **86**, **87**, **92**, **95**, **96** in yields ranging from 66 to 93 % (method A). The aniline derivatives **88**, **93**, and **94** were obtained by reaction between **99a** and **99c** and the appropriate aniline in ethanol at reflux for 4 h (method B).

Compound **91** was prepared as shown in Scheme 2. The 5-amino-1-(2-hydroxy-2-phenylethyl)-1*H*-pyrazole-4-carboxamide **100**, prepared according to our procedure,^[16] was treated with NaOEt and EtOAc in absolute ethanol at reflux for 6 h to afford the 1-(2-hydroxy-2-phenylethyl)-6-methyl-1,5-dihydro-4*H*-pyrazolo[3,4-*d*]pyrimidin-4-one **101** that was in turn chlorinated with the Vilsmeier complex (POCl₃/DMF, 1:1, 30 equiv) in CHCl₃ at reflux for 12 h to obtain the dichloro derivative **102**. Reaction of the latter with aniline in absolute ethanol at reflux for 4 h gave **91** (method B).

Compounds **89** and **90** were prepared as shown in Scheme 3. Reaction of ethyl 5-amino-1-[2-(4-bromophenyl)-2-hydroxyethyl]-1*H*-pyrazole-4-carboxylate **103**, prepared according to our procedure,^[17] with formamide for 8 h at 190 °C led to the pyrazolo[3,4-*d*]pyrimidinone **104** in high yield. Treatment of the latter with an excess of the Vilsmeier complex (POCl₃/DMF, 1:1, 10 equiv) for 12 h at reflux in CHCl₃ gave the dichloro derivative **105**. Finally, the C4 chlorine atom was substituted with the appropriate benzylamines in anhydrous toluene at RT to afford

the final compounds **89** and **90** (method A). The synthesis of compounds **84** and **85** has been previously reported.^[15]



Scheme 3. Synthesis of compounds **89** and **90**. Reagents and conditions: a) formamide, 190 °C, 8 h; b) POCl₃/DMF, CHCl₃, reflux, 12 h; c) method A: benzylamines, anhydrous toluene, RT, 48 h.

Molecular modeling calculations

Docking simulations

To check if the active site of Abl is the preferential binding site of our compounds, blind docking calculations were performed on the overall kinase domain of the enzyme in its active, inactive, and Src-like inactive forms, using the Abl–dasatinib complex (entry 2GQG of the PDB),^[9] the Abl–gleevec (STI-571) complex (entry 1IEP),^[10] and the Abl–ATP complex (entry 2G2I),^[12] respectively. In particular, for each crystal structure, after removal of the ligand, the kinase domain was relaxed through an energy minimization step to avoid any steric bump which may affect the X-ray structure. The software AutoDock was then used to carry out blind docking simulations of inhibitors into the relaxed structure of the Abl kinase domain. As a result, all the clusters found by the software showed the inhibitor accommodated within the kinase domain of the enzyme, supporting the hypothesis that this is the preferential binding site for pyrazolo-pyrimidines.

In the next step, the binding mode of pyrazolo-pyrimidine derivatives was further investigated by means of docking simulations focused on the Abl kinase domain, following a computational protocol previously described (software GOLD).^[7b] However, to check for the ability of the software to appropriately handle Abl inhibitors, docking simulations were first performed on the co-crystallized inhibitors or substrate. As a result, the best solution or the most populated cluster of ligand conformations contained the experimental pose of each co-crystallized inhibitor or ligand, with a minimal deviation of its atomic coordinates. Having a reliable computational protocol in our hands, it was used to dock pyrazolo-pyrimidine into the Abl kinase domain. For this purpose, and to take into account the structural variability of the binding site, the structures of the three Abl conformations reported above were used.

Two principal orientations were found for pyrazolo-pyrimidines within the ATP pocket of Abl, depending on the presence or not of an alkylthio substituent at C6 position of the heterocyclic nucleus, according to results previously reported.^[7b] Briefly, compounds lacking the substituent at C6 showed an interaction pathway similar to dasatinib, although only a hydrogen bond with Met318 was found, whereas an additional interaction between the hydroxy group of Thr315 and the amino group at C4 of the ligands further stabilized their

complexes with Abl. Several hydrophobic contacts with the pockets referred to as hydrophobic region I (HRI, mainly constituted by Val256, Ala269, Lys271, Met290, Val299, Ile313, Thr315, Leu370, and Phe382) and hydrophobic region II (HRII, mainly constituted by Leu248, Gly249, Tyr253, Gly321, and Phe317) were also found.

In the alternative binding mode, the presence of the alkylthio substituent at the C6 position induced a reorientation of the pyrazolo-pyrimidine scaffold, whereas side chains at N1 and C4 retained lipophilic contacts with the hydrophobic regions of the enzyme binding site. In comparison with the previous binding mode, only the hydrogen bond usually involving the amino group of the C4 side chain and the carbonyl backbone of Met318 was maintained.

Compounds bearing a bulky substituent at C4 showed a hydrogen bond between the N2 of the pyrazolo-pyrimidine ring and the NH backbone of Met318. However, in both cases, the loss of a hydrogen bond was not associated with a decrease in activity. In addition, it is worth noting that all the compounds bearing a tertiary amino group at position 4 showed a common orientation, independent from the presence or not of a C6 substitution. Such an orientation, very similar to that observed for compounds unsubstituted at position 6, resulted in the loss of the hydrogen bonds previously described. These observations further supported our hypothesis that the lack of hydrogen bonds does not seem to have a significant effect on activity. In fact, although hydrophobic interactions appeared to be very important for the binding of pyrazolo-pyrimidine derivatives to Abl kinase, hydrogen bond interactions did not appear to play a fundamental role for the affinity of compounds.^[7b] This is in agreement with the fact that not only compounds establishing a single hydrogen bond with the target but also many of those whose binding is characterized by no hydrogen bonds showed good inhibitory activity.

Docking simulations on inactive and Src-like inactive conformations of Abl led to results which were difficult to interpret. In fact, in both cases it was impossible to identify a preferential binding mode, although contacts typical of STI-571 and ATP were reproduced. Also in these cases, hydrophobic interactions are mainly responsible for complex stabilization, instead of hydrogen bonds.

Generation of 3D QSAR models for Abl inhibitors

The 3D QSAR analysis was performed on the basis of the following steps: 1) data set collection, 2) structure-based alignment of molecules, 3) computation of molecular descriptors and principal component analysis (PCA), 4) training and test set selection, 5) variable selection, QSAR model generation, and validation, and 6) PLS analysis of the final 3D QSAR model.

Data set collection

A series of 83 pyrazolo-pyrimidines were collected from our previous studies on Abl inhibitors, along with their affinity

toward Abl (table S1 in the Supporting Information). Being aware that such compounds may have low solubility,^[14] we have chosen to prune the original dataset of 83 entries by removing compounds whose low solubility may have affected the assessment of the activity toward Abl. Compounds with a QPlog $P_{o/w}$ value (calculated with the software QikProp)^[18] lower than 6.5 (−2.0 and 6.5 define the range of log P values reported for 95% of known drugs)^[18] were removed, leading to a set of 68 entries (Table 1). These compounds share a common structural pattern, consisting of a pyrazolo-pyrimidine moiety with various functional groups at positions 1, 4, and 6, and show affinity values toward Abl (expressed as inhibition constants, K_i) ranging from 0.02 μM to higher than 50 μM (the highest test concentration). Inactive compounds were arbitrarily assigned a K_i value of 50 μM . Affinity values were transformed into the corresponding $-\log K_i$ (p K_i) for subsequent calculations.

Structure-based alignment of molecules

Focused docking calculations (performed following the same protocol as described above) were performed on the 68 compounds to find the most profitable interaction pattern with the binding site of the enzyme. As different binding modes were found among the conformations suggested by GOLD, the lowest energy docked conformation of the most populated cluster was chosen to define the alignment rule for subsequent QSAR analysis.

Computation of molecular descriptors and principal component analysis

Ligands, aligned each other, were imported into GRID and interaction energies between selected probes and each molecule (molecular interaction fields, MIFs) were calculated using a grid spacing of 1 Å. Three probes were chosen to describe all the molecules. The C3 probe, corresponding to a methyl group, was used to account for steric contacts. The O probe (a sp^2 carbonyl oxygen) was used as a hydrogen bond acceptor, and the N1 probe (neutral flat NH, that is, amide) was used as a hydrogen bond donor. Interaction energy values at each grid point were imported into the GOLPE software for a preliminary PCA.

Training and test set selection

The main components were then imported into Cerius2^[19] software. The selection of training set compounds was performed taking into account both the structural features and activity profile of the data set of 68 entries. In detail, we have tried to maximize the structural diversity, described in the orthogonal variables (principal components), keeping the original distribution profile of activity. As a result, 34 compounds were selected as the training set, whereas the remaining 34 compounds constituted the test set (Table 1). Compounds belonging to the training and the test set appeared to be uniformly dispersed in the descriptor space and in the activity distribution profile (Figure 1).

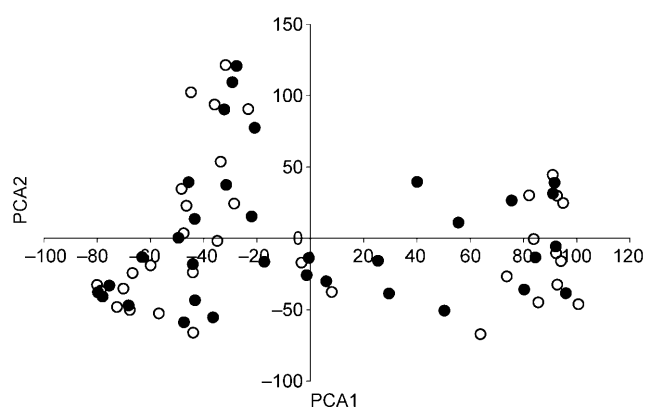


Figure 1. Distribution of training (●) and test set (○) compounds in the descriptor space.

Variable selection, QSAR model generation, and validation

The descriptor set of training set compounds was imported into GOLPE. As many of the variables derived from the GRID analysis do not contribute to the correlation between chemical structure and biological activity (that is, they could be considered noise that decreases the quality of the model), they were removed using the pretreatment tool of GOLPE (details on the various steps of variable selection are reported in table S2 in the Supporting Information).

Accordingly, starting from 39 168 variables describing the training set compounds in the active form of the Abl binding site, application of the preliminary pretreatment (absolute values lower than 0.01 kcal mol^{-1} were set to zero, whereas variables that exhibited only two values were removed) decreased the number of variables to 17 770. Moreover, the D-optimal pre-selection procedure was also applied once (using 50% as reduction level and five components) to obtain the most informative variables correlated with biological activity. Each selection of variables was preceded by calculation of a PLS model based only on the last selected variables. After D-optimal variable pre-selection, the number of variables was further decreased to 8885.

A smart region definition (SRD) algorithm was applied to select and group the regions of variables of highest importance for the model. Groups were then used in the fractional factorial design (FFD) procedure replacing the original variables. FFD selection was applied until no further statistical improvement was observed (two runs). As a result, the model obtained from the reduced set of variables (1229) was characterized by a significant enhancement of the predictive power with respect to the parent model generated on the entire set of variables (Table 2). In fact, the validation procedure showed an improved internal predictive ability ($q^2=0.90$) for the final model with five components, in comparison with the model based on all the 39 168 variables ($q^2=0.52$). This result allowed us to conclude that many of the original variables generated noise, instead of contributing to the robustness of the model. Moreover, the predictive power of the model was assessed by calculating affinity values for the test set compounds. As a result, a good correlation was found (0.51 as standard deviation).

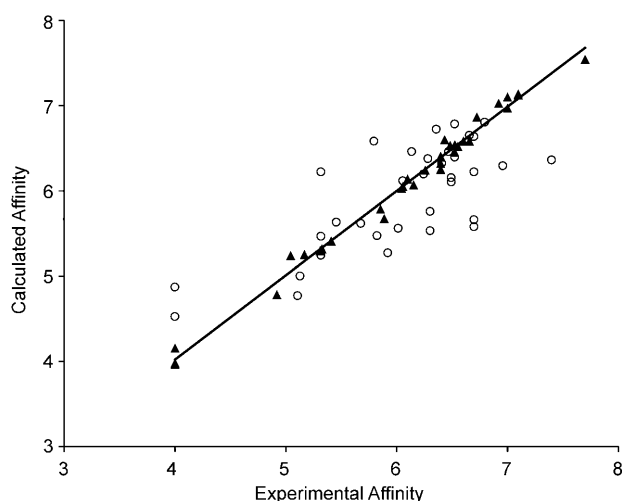
Table 2. Statistical parameters of calculations leading to the final 3D QSAR model for Abl inhibitors based on the active conformation of the enzyme.

Training and Test Set Selection ^[a] PCA ^[b]				First PLS ^[c]		GOLPE Software Last PLS ^[d]			Prediction on Test Set
Components	XVarExp	Xaccum	R ²	Validation (LOO)		Validation (LOO)		SDEP Ext	
				SDEP	q ²	R ²	SDEP	q ²	
0			0.0000	0.9243	−0.0615	0.0000	0.9243	−0.0615	0.7491
1	18.0533	18.0533	0.6882	0.8263	0.1516	0.7193	0.7243	0.3482	0.6279
2	10.6689	28.7221	0.8956	0.6367	0.4963	0.9420	0.4211	0.7797	0.5639
3	8.4715	37.1936	0.9397	0.6541	0.4683	0.9712	0.3527	0.8455	0.5035
4	6.7509	43.9445	0.9694	0.6124	0.5340	0.9810	0.3142	0.8773	0.5073
5	4.6377	48.5821	0.9732	0.623	0.5178	0.9894	0.2850	0.8991	0.5142

[a] Cerius2 software. [b] 68 objects, 39 168 variables. [c] 34 objects, 39 168 variables. [d] 34 objects, 1229 variables.

[a] Cerius2 software. [b] 68 objects, 39 168 variables. [c] 34 objects, 39 168 variables. [d] 34 objects, 1229 variables.

tion of error of predictions, SDEP) between experimental and predicted activity values of the external set of compounds (Table 2 and Figure 2).

**Figure 2.** Calculated (estimated and predicted) versus experimental affinity of training (●) and test set (▲) compounds.

The same computational protocol applied to compounds aligned into the inactive and Src-like inactive conformations of Abl did not result in any statistically significant 3D QSAR model. For this reason, subsequent calculations were performed only on the basis of Abl inhibitors aligned into the active form of the kinase domain.

PLS analysis of the final 3D QSAR model

One important feature of 3D QSAR analysis is the graphical representation of the model, which makes its interpretation easier. Several options are available in GOLPE to display the final model. Among these, PLS pseudo-coefficients are very useful because they allow the visualization of favorable and unfavorable interactions between the probes and the molecules under study. For this reason, GRID plots of PLS pseudo-coefficients were analyzed by means of GOLPE with the purpose of getting a better understanding of the structure–activi-

ty relationships. Compounds **7** and **83** (showing the best activity toward Abl among the C6-unsubstituted and C6-substituted pyrazolo-pyrimidines, respectively) were chosen as representative examples of the entire set of compounds to analyze plots of PLS coefficients.

Contour maps for the C3 probe indicated sterically disfavored and favored regions of space (magenta and yellow, respectively). In particular, regions with higher positive values corresponded to a profitable interaction between a molecular substituent and the methyl probe, contributing to an improvement in affinity. In contrast, the principal regions with negative values corresponded to unfavorable interactions between a substituent of the molecule and the methyl probe, contributing to a decrease in activity. Analysis of **7** embedded into the C3 contour map showed HRI mainly contoured by favorable regions (yellow dihedrals, Figure 3 A), whereas HRII was characterized by favorable regions corresponding to the para position of the terminal phenyl ring at the N1 side chain. In an attempt to better fill such a region, compounds **89** and **90**, bearing a bromine atom at the para position of the N1 phenyl ring, were synthesized and found to have a significantly different affinity. In particular, the high affinity of **90** (0.02 μM) was accounted for by the fact that both the N1 and C4 substituents occupied favorable regions, with the bromine matching HRII in a region corresponding to the side chain of Tyr253 and the *m*-F phenyl group lying within the yellow spheres close to HRI. In contrast, the *o*-F substituent of **89** was found in proximity to a forbidden part of HRI, near the Thr315 carbonyl group (upper right portion of Figure 3 A), probably accounting for the lower affinity (0.44 μM). Moreover, it is important to note that an unfavorable region (magenta dihedrals, corresponding to the carbonyl group of Gly254) was found in front of the C6 position, accounting for the fact that C6-substituted pyrazolo-pyrimidines adopted an alternative binding mode with respect to the corresponding C6-unsubstituted analogues.

Accordingly, **83**, characterized by a good affinity toward isolate Abl (0.10 μM), showed the alternative binding mode imposed by the presence of a C6 substituent, leading the N1 side chain to be located within HRI, while the alkyl substituent at C6 was accommodated within HRII (Figure 3 B). Analysis of the methyl contour map showed two different favorable zones in front of the C6 portion of the heterocyclic scaffold of C6-substituted compounds. One of them (corresponding to the side

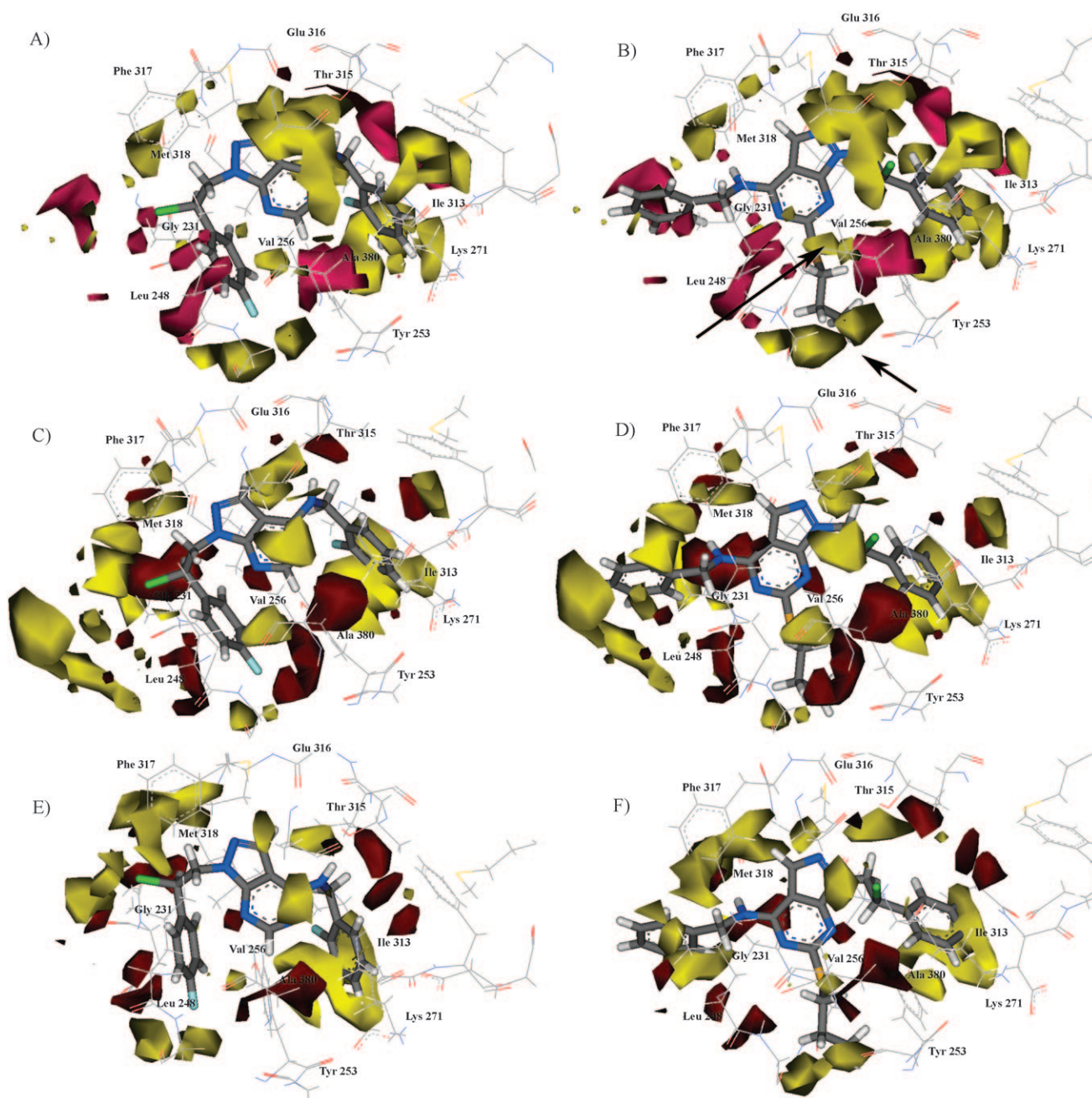


Figure 3. PLS coefficient plots obtained with the C3 [panels A) and B)], N1 [panels C) and D)], and O [panels E) and F)] probes, projected in the Abl-7 and Abl-83 complexes, respectively. Favorable interactions (that is, an increase in pK) between a substituent and the probe occur in yellow regions, whereas unfavorable interactions (a decrease in pK) between a substituent and the probe occur in magenta regions.

chains of Val256 and Leu248 and indicated with a long arrow in Figure 3B) was very close to the central core and suggested that a small hydrophobic group in this position could give active compounds. As an example, the C6 methyl derivative **91**, whose methyl group perfectly filled the favorable contour, had an affinity of $0.02 \mu\text{M}$. The second region, (mainly corresponding to the side chain of Tyr253 and indicated with a short arrow in Figure 3B) allowing for the presence of hydrophobic groups, is located in such a way to be matched by the alkyl moiety of a thioalkyl substituent at C6. Among thioalkyl substituents, cyclic moieties were suggested to have the opti-

mal shape and size properties to fit such a region, the cyclopentyl (**86** and **87**) and cyclohexyl (**88**) derivatives being characterized by good affinity values (0.06 , 0.07 , and $0.07 \mu\text{M}$, respectively). Also compounds bearing smaller substituents, such as **84** and **85**, had similar affinity values (0.09 and $0.06 \mu\text{M}$, respectively). On the other hand, an *n*-butyl chain causes a translation of the ligand within the binding site, even if the affinity was maintained at low micromolar values in many cases (**92–96**).

Moreover, the good affinity found for **93–94** (0.04 and $0.09 \mu\text{M}$, respectively) seemed to depend on the fact that the

meta position of the anilino moiety at C4 is contoured by a favorable region. Among compounds bearing an unsubstituted phenyl ring at C4, lengthening the benzylamino side chain of **95** to **96** did not significantly influence affinity. Similarly, **86** and **87** showed comparable affinity. On the basis of these results, the size and shape of the amino side chain at position 4 was worthy of further consideration. In fact, although the linear portion of the chain (corresponding to the ethylenamino or methylenamino spacer) was accommodated in a region of space without any contact with methyl contours, the phenyl ring was embedded in a narrow groove between forbidden regions defined by carbonyl moieties of Thr319 and Tyr320. As a result, an interesting affinity was found, in general, for the phenylethylamino, benzylamino, and anilino derivatives with respect to compounds with bulkier (five- and six-membered heterocyclic rings, such as pyrrolidine, morpholine, and piperidine) substituents at C4. This finding suggested that optimal substituents at position 4 should have a planar shape, whereas cyclic amino substituents showed steric hindrance, interacting with the magenta portion of the methyl probe contour.

Contour maps generated with the N1 probe mimic the hydrogen bond donor interactions. It is interesting to note that many of the positive maps (yellow) are positioned in the same portion of space as found for the C3 probe, indicating that the major effect of the N1 probe in these regions could be considered to be of steric nature. Analysis of the N1 contour map calculated for **7** showed a favorable region between C3 and C4, corresponding to the hydroxy terminus of Thr315 (Figure 3C). Such an interaction, common to many of the pyrazolo-pyrimidines, could account for the lower affinity showed by compounds unable to make a hydrogen bond contact with Thr315 (that is, **57**). On the other hand, the C4 amino group of **83**, located in the proximity of a profitable region where the carbonyl group of Met318 is embedded, is found to interact by hydrogen bonds with the ligands (Figure 3D).

The contribution of the O probe to the PLS model represents both the ability to accept hydrogen bond contacts and to make steric interactions. In fact, similar to that found for the N1 probe, some portions of the O probe maps occupy regions of space similar to those previously identified with the C3 probe, suggesting that the O probe is also involved in steric interactions. This probe does not discriminate for activity, even if a zone of favorable interaction is in front of the N2 of the pyrazole ring, as was found for **7** (Figure 3E). The same favorable interaction pattern was found for **83**, although with a different orientation of the pyrazolo-pyrimidine scaffold (Figure 3F).

The new synthesized molecules were then tested to assess their effects on the proliferation of two Bcr-Abl-positive leuke-

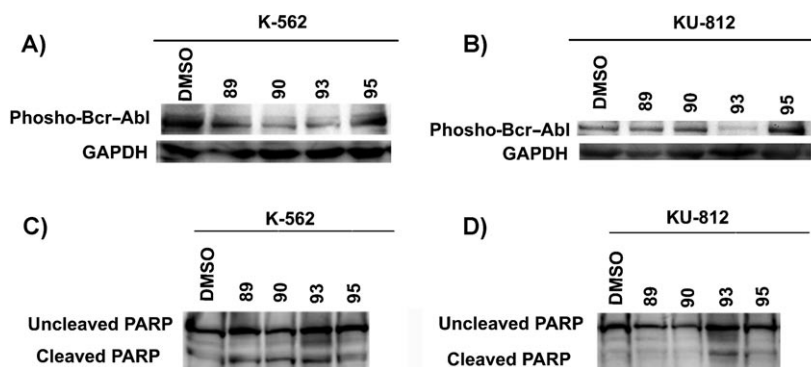


Figure 4. Inhibition of Bcr-Abl phosphorylation [panels A) and B)] and pro-apoptotic effect [panels C) and D)] of selected compounds on K-562 and KU-812 cells, respectively, measured after 72 h incubation. Values shown in panels A) and B) indicate the percentage of phospho-Bcr-Abl over GAPDH, whereas values shown in panels C) and D) indicate the percentage of cleaved over uncleaved PARP.

mia cell lines (namely, K-562 and KU-812), in comparison with PP2, chosen as the reference compound. As several of the new compounds showed an improved activity with respect to parent compounds previously described^[7b,c] (Table 1), they were submitted to additional biological assays in the same cell lines. In particular, phosphorylation of Bcr-Abl was studied to check if the antiproliferative effect of the selected compounds toward each cell line was dependent on the inhibition of Bcr-Abl activity. Compounds significantly decreased Bcr-Abl phosphorylation in K-562 cells (Figure 4A), whereas the effect on Bcr-Abl phosphorylation was lower in KU-812 cells relative to K-562 cells (Figure 4B). These results suggested that the effects mediated by selected compounds on the proliferation of leukemia cells were a consequence of the decrease in Bcr-Abl kinase activity.

Selected compounds were also evaluated for their pro-apoptotic activity on both cell lines by means of a poly-ADP-ribose polymerase (PARP) assay. Compounds **89**, **90**, and **93** potentially induced apoptosis in K-562 cells (Figure 4C), whereas **95** showed lower pro-apoptotic activity, even if higher than that found for the control (DMSO). Analysis of pro-apoptotic activity in KU-812 cells showed that **93** was the best apoptotic inducer (Figure 4D). On the basis of the ability to induce apoptosis in both cell lines, the ratio between Bax mRNA and Bcl-xL mRNA expression after 72 h incubation was also investigated (the pro-apoptotic action of Bax is antagonized by Bcl-xL and this ratio could help explain the molecular mechanism of apoptosis induction). The compounds were able to increase the Bax/Bcl-xL ratio in K-562 cells, confirming their induction, and **90** was the best inducer (Figure 5A). Moreover, the compounds also caused an increase in the Bax/Bcl-xL ratio in KU-812 cells, with respect to control (Figure 5B).

Finally, the expression of the cyclin-dependent kinase inhibitor p21 was also investigated because p21 expression is associated with cell growth inhibition. Exposure of both K-562 and KU-812 cells to compounds for 72 h affected p21 mRNA expression. However, although all the test compounds were good inducers of p21 in K-562 cells (Figure 5C), **89** and **90** were the most active compounds in KU-812 cells (Figure 5D).

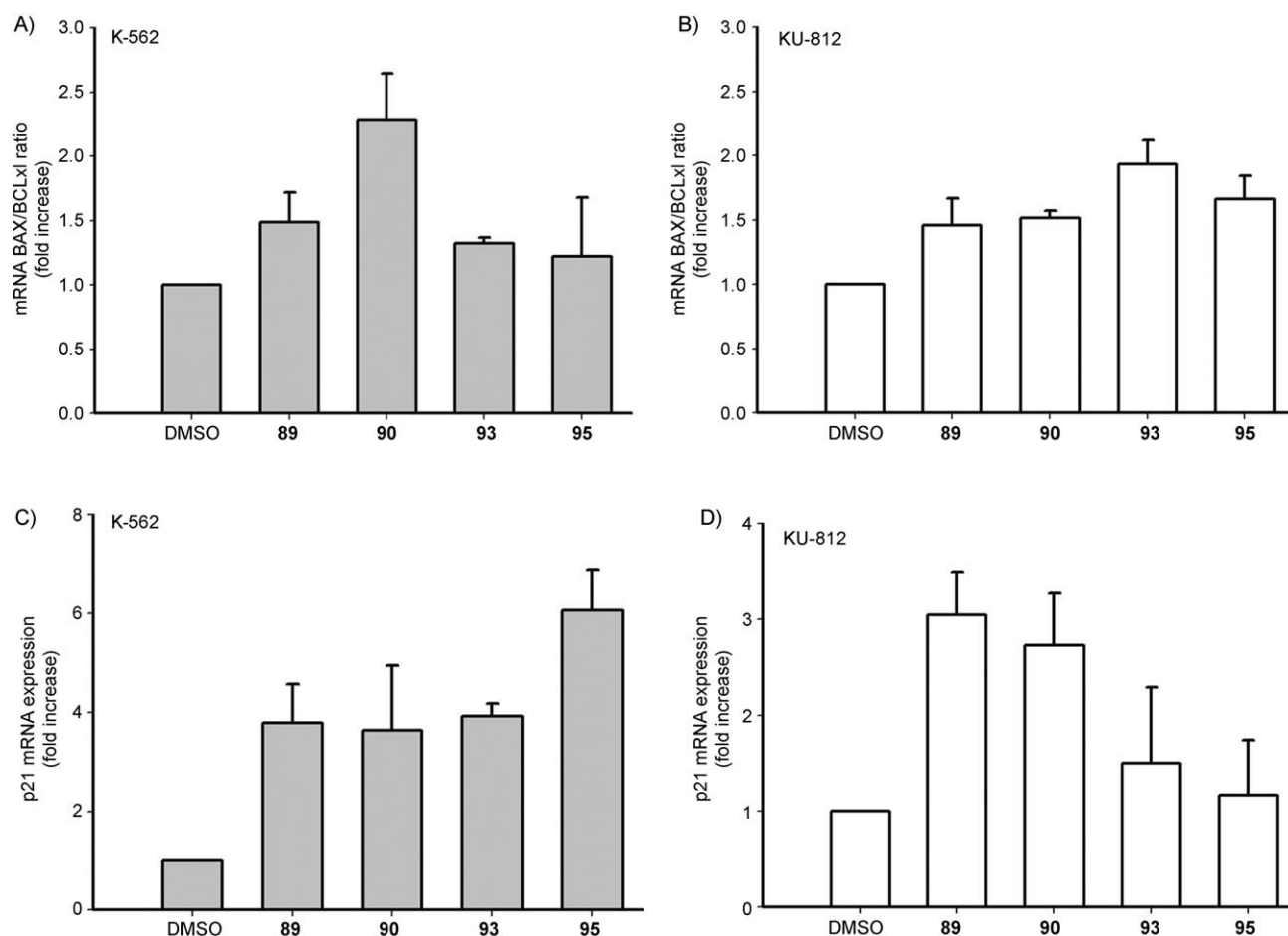


Figure 5. The pro-apoptotic response of A) K-562 and B) KU-812 cells after 72 h incubation. The expression of p21 mRNA in C) K-562 and D) KU-812 cells after 72 h exposure to compounds. The expression of Bax, Bcl-xL, and p21 mRNA was determined by qRT-PCR. Results are expressed as the mean \pm SEM of three independent experiments. All values are expressed as fold increase relative to the expression of β -actin.

These results confirmed the important role of the selected compounds as inhibitors of cell proliferation.

Conclusions

A set of pyrazolo[3,4-*d*]pyrimidine derivatives, along with their affinity values toward isolated Abl tyrosine kinase, were used to build 3D QSAR models based on structure-based alignments of compounds into the enzyme active site. Alternative binding modes for C6-substituted and unsubstituted compounds were confirmed, as well as the importance of hydrophobic interactions with respect to hydrogen bond contacts, as hypothesized in our previous calculations.^[7a] The sole model constructed on the active conformation of Abl gave significant results in terms of both internal and external predictivity. In fact, it was able to estimate/predict the affinity of a very large number of inhibitors differently decorated at positions N1, C4, and C6 of the central core. Suggestions derived from analysis of PLS pseudo-coefficients led to the synthesis of new derivatives, which were found to have good affinity values toward the enzyme and further validated the model itself. Additional biological assays confirmed the important role of the selected compounds as inhibitors of cell proliferation in leukemia cells.

Experimental Section

Chemistry

Starting materials were purchased from Aldrich-Italia (Milan, Italy). Melting points were determined with a Büchi 530 apparatus and are uncorrected. IR spectra were measured with a PerkinElmer 398 spectrophotometer. ¹H NMR spectra were recorded on a Varian Gemini 200 (200 MHz) instrument. Chemical shifts are reported as δ (ppm) relative to (CH₃)₄Si as internal standard; *J* values are given in Hz. ¹H patterns are described with the following abbreviations: s = singlet, d = doublet, t = triplet, q = quartet, quint = quintet, sx = sextet, m = multiplet, br = broad. All compounds were tested for purity by TLC (Merck, silica gel 60 F₂₅₄, CHCl₃ as the eluent). Analyses for C, H, N, S were within $\pm 0.3\%$ of the theoretical value. Synthesis and analytical data of compounds **84** and **85** have been already reported.^[15]

General procedure for the synthesis of 6-(alkylthio)-1-(2-hydroxy-2-phenylethyl)-1,5-dihydro-4H-pyrazolo[3,4-*d*]pyrimidin-4-ones **98a–c.** A mixture of 1-(2-hydroxy-2-phenylethyl)-6-thioxo-1,5,6,7-tetrahydro-4H-pyrazolo[3,4-*d*]pyrimidin-4-one **97** (2.88 g, 10 mmol), the appropriate alkyl bromide (10.14 mmol) and anhydrous K₂CO₃ (1.38 g, 10 mmol) in anhydrous DMF (10 mL) was stirred at RT for 24 h for **98a**, 8 h for **98b**, or at 80 °C for 15 h for **98c**. The mixture was poured into cold H₂O, the white solid was fil-

tered, washed with H₂O, and recrystallized from absolute EtOH to afford **98a–c** as white solids.

98a. White powder (1.89 g, 55%); mp: 173–174 °C; ¹H NMR (200 MHz, (CD₃)₂SO): δ = 12.08 (brs, 1H, NH disappears with D₂O), 7.87 (s, 1H, H-3), 7.26–7.02 (m, 5H, Ar), 5.55 (d, 1H, OH, disappears with D₂O), 5.04–4.90 (m, 1H, CHOH), 4.38–4.14 (m, 2H, CH₂N), 3.04 (t, *J* = 7.2 Hz, 2H, CH₂S), 1.58 (quint, *J* = 7.2 Hz, 2H, CH₂CH₂CH₃), 1.35 (sx, *J* = 7.2 Hz, 2H, CH₂CH₃), 0.85 ppm (t, *J* = 7.2, 3H, CH₃); IR (KBr): $\tilde{\nu}$ = 3500–3310 (NH+OH), 1697 cm^{−1} (CO); Anal. calcd for C₁₇H₂₀N₄O₂S: C 59.28, H 5.85, N 16.27, S 9.31, found: C 59.22, H 6.05, N 16.31, S 9.15.

98b. White powder (1.96 g, 55%); mp: 213–214 °C; ¹H NMR (200 MHz, (CD₃)₂SO): δ = 12.19 (brs, 1H, NH disappears with D₂O), 7.87 (s, 1H, H-3), 7.28–7.03 (m, 5H, Ar), 5.57 (d, 1H, OH disappears with D₂O), 5.12–4.92 (m, 1H, CHOH), 4.42–4.17 (m, 2H, CH₂N), 3.97–3.76 (m, 1H, CHS), 2.27–1.94 and 1.74–1.42 ppm (2×m, 8H, 4CH₂ cyclopentyl); IR (KBr): $\tilde{\nu}$ = 3150–2850 (NH+OH), 1703 cm^{−1} (CO); Anal. calcd for C₁₈H₂₀N₄O₂S: C 60.65, H 5.66, N 15.72, S 9.00, found: C 60.31, H 5.82, N 15.70, S 8.90.

98c. White powder (1.78 g, 48%); mp: 226–227 °C; ¹H NMR (200 MHz, (CD₃)₂SO): δ = 12.22 (brs, 1H, NH disappears with D₂O), 7.90 (s, 1H, H-3), 7.40–7.10 (m, 5H, Ar), 5.60 (d, 1H, OH disappears with D₂O), 5.22–5.00 (m, 1H, CHOH), 4.39–4.20 (m, 2H, CH₂N), 3.99–3.89 (m, 1H, CHS), 1.85–1.30 ppm (m, 10H, 5CH₂ cyclohexyl); IR (KBr): $\tilde{\nu}$ = 3315–2930 (NH+OH), 1704 cm^{−1} (CO); Anal. calcd for C₁₉H₂₂N₄O₂S: C 61.60, H 5.99, N 15.12, S 8.66, found: C 61.56, H 6.05, N 15.11, S 8.36.

General procedure for the synthesis of 6-(alkylthio)-4-chloro-1-(2-chloro-2-phenylethyl)-1H-pyrazolo[3,4-*d*]pyrimidines **99a–c.** The Vilsmeier complex, previously prepared from POCl₃ (1.53 g, 10 mmol) and anhydrous DMF (0.73 g, 10 mmol), was added to a suspension of the appropriate compound **98a–c** (1 mmol) in CHCl₃ (50 mL). The mixture was held at reflux for 8 h. The solution was washed with H₂O (2×20 mL), dried (MgSO₄), filtered, and concentrated under reduced pressure. The crude oil was purified by column chromatography (Florisil, 100–200 mesh) using Et₂O as the eluent, to afford the pure products **99a–c** as white solids.

99a. White powder (0.31 g, 80%); mp: 81–82 °C; ¹H NMR (200 MHz, CDCl₃): δ = 8.00 (s, 1H, H-3), 7.32–7.10 (m, 5H, Ar), 5.19–5.05 (m, 1H, CHCl), 4.95–4.80 and 4.72–4.47 (2×m, 2H, CH₂N), 3.09 (t, *J* = 7.2 Hz, 2H, CH₂S), 1.61 (quint, *J* = 7.2 Hz, 2H, CH₂CH₂CH₃), 1.40 (sx, *J* = 7.2 Hz, 2H, CH₂CH₃), 0.87 ppm (t, *J* = 7.2 Hz, 3H, CH₃); Anal. calcd for C₁₇H₁₈N₄Cl₂S: C 53.55, H 4.76, N 14.69, S 8.41, found: C 53.74, H 4.90, N 14.61, S 8.33.

99b. White powder (0.25 g, 63%); mp: 70–71 °C; ¹H NMR (200 MHz, CDCl₃): δ = 7.94 (s, 1H, H-3), 7.42–7.19 (m, 5H, Ar), 5.50–5.38 (m, 1H, CHCl), 4.97–4.82 and 4.80–4.67 (2×m, 2H, CH₂N), 4.07–3.93 (m, 1H, CHS), 2.32–2.10 and 1.85–1.54 ppm (2×m, 8H, 4CH₂ cyclopentyl); Anal. calcd for C₁₈H₁₈N₄Cl₂S: C 54.96, H 4.61, N 14.24, S 8.15, found: C 54.83, H 4.60, N 14.20, S 8.00.

99c. White powder (0.35 g, 86%); mp: 77–78 °C; ¹H NMR (200 MHz, CDCl₃): δ = 7.72 (s, 1H, H-3); 7.38–7.21 (m, 5H, Ar), 5.35–5.25 (m, 1H, CHCl), 4.84–4.64 and 4.62–4.58 (2×m, 2H, CH₂N), 3.95–3.87 (m, 1H, CHS), 1.82–1.32 ppm (m, 10H, 5CH₂ cyclohexyl); Anal. calcd for C₁₉H₂₀N₄Cl₂S: C 56.02, H 4.95, N 13.75, S 7.87, found: C 56.31, H 4.90, N 13.90, S 7.85.

Synthesis of 1-(2-hydroxy-2-phenylethyl)-6-methyl-1,5-dihydro-4H-pyrazolo[3,4-*d*]pyrimidin-4-one **101.** A solution of NaOEt [prepared from Na (1.84 g, 80 mmol) and absolute EtOH (30 mL)] and

EtOAc (7.93 g, 90 mmol) were added to a solution of 5-amino-1-(2-hydroxy-2-phenylethyl)-1H-pyrazole-4-carboxamide **100** (2.46 g, 10 mmol) in absolute EtOH (90 mL). The mixture was held at reflux for 6 h. After cooling, ice-water was added, and the solution was acidified with 3% acetic acid. A solid precipitated that was filtered, washed with H₂O, and recrystallized from absolute EtOH to afford compound **101** as a white solid (1.62 g, 60%); mp: 253–254 °C; ¹H NMR (200 MHz, (CD₃)₂SO): δ = 11.91 (brs, 1H, NH disappears with D₂O), 7.90 (s, 1H, H-3), 7.30–7.10 (m, 5H, Ar), 5.52 (brs, 1H, OH disappears with D₂O), 5.06–4.92 (m, 1H, CHOH), 4.42–4.27 and 4.20–4.07 (2×m, 2H, CH₂N), 2.25 ppm (s, 3H, CH₃); IR (KBr): $\tilde{\nu}$ = 3400–3150 (NH+OH), 1660 cm^{−1} (CO); Anal. calcd for C₁₄H₁₄N₄O₂: C 62.21, H 5.22, N 20.73, found: C 62.20, H 5.26, N 20.99.

Synthesis of 4-chloro-(2-chloro-2-phenylethyl)-6-methyl-1H-pyrazolo[3,4-*d*]pyrimidine **102.** The Vilsmeier complex, previously prepared from POCl₃ (4.60 g, 30 mmol) and anhydrous DMF (2.20 g, 30 mmol), was added to a suspension of 1-(2-hydroxy-2-phenylethyl)-6-methyl-1,5-dihydro-4H-pyrazolo[3,4-*d*]pyrimidin-4-one **101** (0.27 g, 1 mmol) in CHCl₃ (10 mL). The mixture was held at reflux for 12 h. The solution was washed with H₂O (2×20 mL), dried (MgSO₄), filtered, and concentrated under reduced pressure. The crude oil was purified by column chromatography (Florisil, 100–200 mesh), using Et₂O as the eluent, to afford the product as a yellow oil, which crystallized, standing in a refrigerator upon adding a 1:1 mixture of Et₂O/petroleum ether (PE) (bp: 40–60 °C), as a white solid (0.21 g, 67%); mp: 96–97 °C; ¹H NMR (200 MHz, CDCl₃): δ = 8.01 (s, 1H, H-3), 7.41–7.16 (m, 5H, Ar), 5.51–5.39 (m, 1H, CHCl), 5.04–4.90 and 4.81–4.68 (2×m, 2H, CH₂N), 2.69 ppm (s, 3H, CH₃); Anal. calcd for C₁₄H₁₂N₄Cl₂: C 54.74, H 3.94, N 18.24, found: C 54.54, H 4.15, N 18.06.

Synthesis of 1-[2-(4-bromophenyl)-2-hydroxyethyl]-1,5-dihydro-4H-pyrazolo[3,4-*d*]pyrimidin-4-one **104.** A suspension of ethyl-5-amino-1-[2-(4-bromophenyl)-2-hydroxyethyl]-1H-pyrazole-4-carboxylate **103** (3.54 g, 10 mmol) in formamide (10 g, 333 mmol) was heated at 190 °C for 8 h and then poured into H₂O (300 mL). The crude product was filtered and purified by dissolving in 2 M NaOH (100 mL) and boiling with charcoal, followed by precipitation with glacial acetic acid. The solid was filtered and recrystallized from absolute EtOH to give compound **104** as a white solid (2.35 g, 70%); mp: 269–270 °C; ¹H NMR (200 MHz, (CD₃)₂SO): δ = 12.04 (brs, 1H, NH, disappears with D₂O), 7.98 (s, 1H, H-6), 7.94 (s, 1H, H-3), 7.44–7.31 and 7.21–7.07 (2×m, 4H, Ar), 5.68 (d, 1H, OH, disappears with D₂O), 5.07–4.90 (m, 1H, CHOH), 4.43–4.14 ppm (m, 2H, CH₂N); IR (KBr): $\tilde{\nu}$ = 3274 (NH), 2900–3100 (OH), 1694 cm^{−1} (CO); Anal. calcd for C₁₃H₁₁N₄O₂Br: C 46.59, H 3.31, N 16.72, found: C 46.71, H 3.13, N 16.85.

Synthesis of 4-chloro-1-[2-chloro-2-(4-bromophenyl)ethyl]-1H-pyrazolo[3,4-*d*]pyrimidine **105.** The Vilsmeier complex, previously prepared from POCl₃ (1.53 g, 10 mmol) and anhydrous DMF (0.73 g, 10 mmol), was added to a suspension of 1-[2-(4-bromophenyl)-2-hydroxyethyl]-1,5-dihydro-4H-pyrazolo[3,4-*d*]pyrimidin-4-one **104** (335 mg, 1 mmol) in CHCl₃ (50 mL). The mixture was held at reflux for 12 h. The solution was washed with H₂O (2×20 mL), dried (MgSO₄), filtered, and concentrated under reduced pressure. The crude oil was purified by column chromatography (Florisil, 100–200 mesh) using Et₂O as the eluent, to afford the product as a yellow oil, which crystallized, standing in a refrigerator upon adding a 1:1 mixture of Et₂O/PE (bp: 40–60 °C), as a white solid (0.26 g, 70%); mp: 108–109 °C; ¹H NMR (200 MHz, CDCl₃): δ = 8.74 (s, 1H, H-6), 8.15 (s, 1H, H-3), 7.46–7.35 and 7.30–7.24 (2×m, 4H, Ar), 5.50–5.46 (m, 1H, CHCl), 5.05–4.90 and 4.87–4.82 ppm (2×m,

2H, CH₂N); Anal. calcd for C₁₃H₉N₄Cl₂Br: C 41.97, H 2.44, N 15.06, found: C 41.78, H 2.32, N 15.02.

General procedure for the synthesis of final compounds 86–96. *Method A:* (86, 87, 89, 90, 92, 95, 96). The appropriate amine (40 mmol) was added to a solution of the appropriate 4-chloro derivative 99a, 99b, and 105 (10 mmol) in anhydrous toluene (10 mL), and the mixture was stirred at RT for 48 h. The organic phase was washed with H₂O (2 × 10 mL), dried (MgSO₄), filtered, and concentrated under reduced pressure. The crude oil was crystallized by adding a 1:1 mixture of Et₂O/PE (bp: 40–60 °C).

86. White powder (4.31 g, 93%); mp: 159–160 °C; ¹H NMR (200 MHz, CDCl₃): δ = 7.65 (s, 1H, H-3), 7.44–7.14 (m, 10H, Ar), 5.52–5.41 (m, 1H, CHCl), 4.88–4.58 (m, 4H, CH₂N+CH₂Ar), 4.10–3.92 (m, 1H, SCH), 2.26–2.05 and 1.82–1.51 ppm (2 × m, 8H, 4CH₂ cyclopentyl); IR (KBr): $\tilde{\nu}$ = 3240 cm⁻¹ (NH); Anal. calcd for C₂₅H₂₆N₅ClS: C 64.71, H 5.65, N 15.09, S 6.91, found: C 64.63, H 5.81, N 15.13, S 7.00.

Method B: (88, 91, 93, 94). The appropriate aniline (20 mmol) was added to a solution of the appropriate 4-chloro derivative 99a, 99c, and 102 (10 mmol) in absolute EtOH (10 mL), and the mixture was held at reflux for 4 h. After cooling, the white solid was filtered, washed with H₂O, and recrystallized from absolute EtOH.

91. White powder (1.96 g, 54%); mp: 220–221 °C; ¹H NMR (200 MHz, CDCl₃): δ = 11.06 (brs, 1H, NH disappears with D₂O), 7.62–7.19 (m, 10H, Ar), 6.98 (s, 1H, H-3), 5.39–5.26 (m, 1H, CHCl), 4.97–4.79 and 4.70–4.47 (2 × m, 2H, CH₂N), 2.71 ppm (s, 3H, CH₃); IR (KBr): $\tilde{\nu}$ = 3150 cm⁻¹ (NH); Anal. calcd for C₂₀H₁₈N₅Cl: C 66.01, H 4.99, N 19.25, found: C 65.92, H 5.24, N 18.98.

¹H NMR and IR spectral data and elemental analysis data of final compounds (87–90 and 92–96) are reported in the Supporting Information.

Computational protocol

The three-dimensional structures of protein and ligands were prepared for computational studies according to procedures previously described.^[7a,b]

Blind docking calculations. Blind docking simulations were performed by means of AutoDock 3.0.5.^[20] Mass-centered grid maps were generated with 0.49 Å spacing by the AutoGrid program covering the whole protein target. The genetic algorithm (GA) was applied using default parameters. The number of generations was set to 1 × 10⁶, and the stopping criterion was defined by the total number of energy evaluations. Starting structures of the selected compounds were randomly defined to obtain totally unbiased results. The population of the GA was set to 500, the number of trials was 300, and the number of energy evaluations was 10 × 10⁶.

Focused docking calculations. Docking simulations and alignment were performed according to procedures previously described^[7b] using GOLD 3.2 with the ChemScore scoring function. CHO_TYPE parameter was set to SPECIAL.

Computation of molecular descriptors. The interaction energies (molecular interaction fields, MIFs) were calculated with the C3, O and N1 probes for each molecule of the dataset, using a grid sufficiently large to accommodate all the aligned ligands in all directions (along x, y, and z axes), and a grid spacing of 1 Å. The C3 probe, corresponding to a methyl group, was used to account for steric contacts. The O probe (a sp² carbonyl oxygen) was used as a

hydrogen bond acceptor, and the N1 probe (neutral flat NH, that is, amide) was used as a hydrogen bond donor.

Training set selection. The selection of training set compounds was performed taking into account both the structural features and activity profile of the entire data set.^[21] In detail, to maximize the structural diversity and to maintain the original distribution profile of affinity data, the training set was selected on the basis of the simultaneous optimization of multiple properties of a subset of compounds, which is a common concept in the field of library design.^[21] Selection of the training set was carried out by means of the C2.LibProfile and C2.Diversity modules, implemented in Cerius2.

Variable selection, QSAR model generation, and validation. Details on the various steps conducted with GOLPE for variable selection are reported in table S2 in the Supporting Information.

Acknowledgements

This study was partially supported by funds from PRIN 2007N7KYCY and the Fondazione Monte dei Paschi di Siena. E.C. is supported by a FIRC Fellowship (2008–2010). We thank Prof. Gabriele Cruciani for the use of GOLPE and GRID software.

Keywords: 3D QSAR analysis • Abl inhibitors • ligand design • medicinal chemistry • molecular modeling

- [1] a) J. D. Rowley, *Nature* **1973**, 243, 290–293; b) C. L. Sawyers, *N. Engl. J. Med.* **1999**, 340, 1330–1340; c) T. G. Lugo, A. M. Pendergast, A. J. Muller, O. N. Witte, *Science* **1990**, 247, 1079–1082.
- [2] N. M. Levinson, O. Kuchment, K. Shen, M. A. Young, M. Koldobskiy, M. Karplus, P. A. Cole, J. Kuriyan, *PLoS Biol.* **2006**, 4, 753–767.
- [3] T. Schindler, W. Bornmann, W. T. Miller, B. Clarkson, J. Kuriyan, *Science* **2000**, 289, 1938–1942.
- [4] M. D. Moen, K. McKeage, G. L. Plosker, M. A. Siddiqui, *Drugs* **2007**, 67, 299–320.
- [5] S. Padmanabhan, S. Ravella, T. Curiel, F. Giles, *Future Oncol.* **2008**, 4, 359–377.
- [6] a) G. Noronha, J. Cao, C. P. Chow, E. Dneprovskaya, R. M. Fine, J. Hood, X. Kang, B. Klebansky, D. Lohse, C. C. Mak, A. McPherson, M. S. S. Palanki, V. P. Pathak, J. Renick, R. Soll, B. Zeng, *Curr. Top. Med. Chem.* **2008**, 8, 905–921; b) S. Schenone, F. Manetti, M. Botta, *Mini Rev. Med. Chem.* **2007**, 7, 191–201.
- [7] a) F. Manetti, A. Pucci, M. Magnani, G. A. Locatelli, C. Brullo, A. Naldini, S. Schenone, G. Maga, F. Carraro, M. Botta, *ChemMedChem* **2007**, 2, 343–353; b) F. Manetti, C. Brullo, M. Magnani, F. Mosci, B. Chelli, E. Crespan, S. Schenone, A. Naldini, O. Bruno, M. L. Trincavelli, G. Maga, F. Carraro, C. Martini, F. Bondavalli, M. Botta, *J. Med. Chem.* **2008**, 51, 1252–1259; c) S. Schenone, C. Brullo, O. Bruno, F. Bondavalli, L. Mosti, G. Maga, E. Crespan, F. Carraro, F. Manetti, C. Tintori, M. Botta, *Eur. J. Med. Chem.* **2008**, 43, 2665–2676.
- [8] Brookhaven Protein Data Bank, <http://www.pdb.org> (accessed March 2, 2009).
- [9] J. S. Tokarski, J. Newitt, C. Y. J. Chang, J. D. Cheng, M. Wittekind, S. E. Kiefer, K. Kish, F. Y. F. Lee, R. Borzilleri, L. J. Lombardo, D. Xie, Y. Zhang, H. E. Klei, *Cancer Res.* **2006**, 66, 5790–5797.
- [10] B. Nagar, W. Bornmann, P. Pellicena, T. Schindler, D. R. Veatch, W. T. Miller, B. Clarkson, J. Kuriyan, *Cancer Res.* **2002**, 62, 4236–4243.
- [11] GOLD software is distributed by Cambridge Crystallographic Data Centre: http://www.ccdc.cam.ac.uk/products/life_sciences/gold/ (accessed March 2, 2009).
- [12] GRID v22, Molecular Discovery Ltd., Pinner, Middlesex (UK), <http://www.moldiscovery.com/> (accessed March 2, 2009).

- [13] *GOLPE v4.5*, Multivariate Infometric Analyses Srl, Viale dei Castagni 16, Perugia (Italy), <http://www.miasrl.com/golpe.htm> (accessed March 2, 2009).
- [14] C. Tintori, M. Magnani, S. Schenone, M. Botta, *Eur. J. Med. Chem.* **2009**, *44*, 990–1000.
- [15] F. Manetti, A. Santucci, G. A. Locatelli, G. Maga, A. Spreafico, T. Serchi, M. Orlandini, G. Bernardini, N. P. Caradonna, A. Spallarossa, C. Brullo, S. Schenone, O. Bruno, A. Ranise, F. Bondavalli, O. Hoffmann, M. Bologna, A. Angelucci, M. Botta, *J. Med. Chem.* **2007**, *50*, 5579–5588.
- [16] O. Bruno, C. Brullo, F. Bondavalli, S. Schenone, A. Ranise, N. Arduino, M. B. Bertolotto, F. Montecucco, L. Ottonello, F. Dallegri, M. Tognolini, V. Ballabeni, S. Bertoni, E. Barocelli, *J. Med. Chem.* **2007**, *50*, 3618–3626.
- [17] T. Tuccinardi, S. Schenone, F. Bondavalli, C. Brullo, O. Bruno, L. Mosti, A. T. Zizzari, C. Tintori, F. Manetti, O. Ciampi, M. L. Trincavelli, C. Martini, A. Martinelli, M. Botta, *ChemMedChem* **2008**, *3*, 898–913.
- [18] *QikProp v3.1*, Schrödinger, LLC, New York, NY (USA), <http://www.schrodinger.com/> (accessed March 2, 2009).
- [19] *Cerius2 v4.8* (currently implemented in Discovery Studio 2.0), Accelrys Inc., San Diego, CA (USA), <http://accelrys.com/> (accessed March 2, 2009).
- [20] a) G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, A. J. Olson, *J. Comput. Chem.* **1998**, *19*, 1639–1662; b) C. Hetényi, D. Van Der Spoel, *Protein Sci.* **2002**, *11*, 1729–1739.
- [21] V. J. Gillet, W. Khatib, P. Willet, P. J. Fleming, D. V. S. Green, *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 375–385.

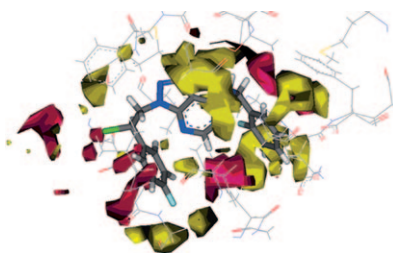
Received: December 17, 2008

Revised: February 6, 2009

Published online on ■ ■ ■, 2009

FULL PAPERS

Quality QSAR: A combination of docking calculations and a statistical approach toward Abl inhibitors resulted in a 3D QSAR model, the analysis of which led to the identification of ligand portions important for affinity. New compounds designed on the basis of the model were found to have very good affinity for the target, providing further validation of the model itself.



F. Falchi, F. Manetti, F. Carraro, A. Naldini, G. Maga, E. Crespan, S. Schenone, O. Bruno, C. Brullo, M. Botta**

■■■ – ■■■

3D QSAR Models Built on Structure-Based Alignments of Abl Tyrosine Kinase Inhibitors

