J Comput Aided Mol Des (2011) 25:349–369
DOI 10.1007/s10822-011-9425-1

Docking and quantitative structure–activity relationship studies for 3-fluoro-4-(pyrrolo[2,1-*f*][1,2,4]triazin-4-yloxy)aniline, 3-fluoro-4-(1H-pyrrolo[2,3-b]pyridin-4-yloxy)aniline, and 4-(4-amino-2-fluorophenoxy)-2-pyridinylamine derivatives as c-Met kinase inhibitors

Julio Caballero · Miguel Quiliano · Jans H. Alzate-Morales · Mirko Zimic · Eric Deharo

Received: 7 October 2010 / Accepted: 3 April 2011 / Published online: 13 April 2011

Abstract We have performed docking of 3-fluoro-4-(pyrrolo[2,1-*f*][1,2,4]triazin-4-yloxy)aniline (FPTA), 3-fluoro-4-(1H-pyrrolo[2,3-b]pyridin-4-yloxy)aniline (FPPA), and 4-(4-amino-2-fluorophenoxy)-2-pyridinylamine (AFPP) derivatives complexed with c-Met kinase to study the orientations and preferred active conformations of these inhibitors. The study was conducted on a selected set of 103 compounds with variations both in structure and activity. Docking helped to analyze the molecular features which contribute to a high inhibitory activity for the studied compounds. In addition, the predicted biological activities of the c-Met kinase inhibitors, measured as IC_{50} values were obtained by using quantitative structure–activity relationship (QSAR) methods: Comparative molecular similarity analysis (CoMSIA) and multiple linear regression (MLR) with topological vectors. The best CoMSIA model included steric, electrostatic, hydrophobic, and hydrogen bond-donor fields; furthermore, we found a predictive model containing 2D-autocorrelation descriptors, GETAWAY descriptors (GETAWAY: Geometry, Topology and Atom-Weight AssemblY), fragment-based polar surface area (PSA), and MlogP. The statistical parameters: cross-validate correlation coefficient and the fitted correlation coefficient, validated the quality of the obtained predictive models for 76 compounds. Additionally, these models predicted adequately 25 compounds that were not included in the training set.

Keywords c-Met kinase inhibitors · Molecular docking · Quantitative structure–activity relationships · CoMSIA · Topological descriptors

Electronic supplementary material The online version of this article (doi:10.1007/s10822-011-9425-1) contains supplementary material, which is available to authorized users.

J. Caballero (✉) · J. H. Alzate-Morales
Centro de Bioinformática y Simulación Molecular, Facultad de Ingeniería en Bioinformática, Universidad de Talca, 2 Norte 685, Casilla 721, Talca, Chile
e-mail: jcaballero@utalca.cl; jmcr77@yahoo.com

M. Quiliano · M. Zimic
Bioinformatics Unit-Drug Design Group, Laboratorios de Investigación y Desarrollo, Facultad de Ciencias y Filosofía, Universidad Peruana Cayetano Heredia, Lima, Perú

E. Deharo
UPS, UMR 152 (Laboratoire de Pharmacochimie des Substances Naturelles et Pharmacophores Redox), Université de Toulouse, 31062 Toulouse cedex 9, France

E. Deharo
IRD; UMR-152, Mission IRD Casilla 18-1209, Lima, Perú

Introduction

c-Met is a receptor tyrosine kinase that is expressed in endothelial and epithelial cells. In normal cells, c-Met is activated by its ligand hepatocyte growth factor (HGF)/scatter factor [1]. When HGF binds to c-Met [3], causes receptor dimerisation and autophosphorylation of tyrosines 1234 and 1235. Phosphorylated c-Met further triggers the activation of downstream signaling pathways such as the Ras/MAPK, the c-Src, and the PI3 K/Akt pathways [2]. Under normal conditions, the pleiotropic effects of the HGF mediated c-Met activation is essential for normal physiological events, such as placental development and liver regeneration [3]; however, in cancer, both HGF and c-Met have been closely linked to the regulation of the

metastatic process [4]. In this sense, c-Met has recently attracted considerable interest as a therapeutic target based on the discovery that aberrant c-Met activity is related with the occurrence of various cancers including lung, prostate, renal, ovarian, gastric, and liver cancers [5].

According to the role of c-Met signaling in cancer progression and metastases, the c-Met receptor is considered a potential target for cancer therapy. c-Met tyrosine kinase inhibitors are able to block autophosphorylation of the c-Met kinase, thereby interrupting its downstream signaling pathways [6]. To date, a number of small molecule inhibitors targeting c-Met kinase have been reported, and some of them have been tested in preclinical settings and lead compounds have entered clinical trials [7]. Recently, the structures of several c-Met kinase–inhibitor complexes have been reported [8–17]. That structural information can be used to establish the molecular features that provide good inhibitory properties.

Nowadays, computational applications are essential for rational drug design. Computer simulation techniques potentially provide a further means to interpret the structural data of X-ray experiments in new knowledge that can be exploited for designing new drugs. Models that are able to explain the interactions and predict the biological activity of compounds by their structural properties are powerful tools for designing highly active molecules. Computational methods, such as docking [18–20], quantitative structure–activity relationship (QSAR) [21–26], pharmacophore [27, 28], de novo design [29, 30], quantum mechanics/molecular mechanics (QM/MM) [18, 19, 26, 31, 32], and molecular dynamics (MD) [19, 33] have been used for studying kinase inhibitors. These methods have been scarcely applied for studying c-Met kinase inhibitors. One example was reported by Asses et al. [34]. They applied molecular modeling and molecular mechanics to analyze the distribution of ligands interaction energy on c-Met kinase residues, and deduced a new model of the active site allowing for an unambiguous identification of ligand binding modes. They proposed that the binding of known ligands on the c-Met kinase catalytic site involves seven identified structurally-distinct areas, where five areas match the generic kinase ATP binding site, and two others are allosteric regions. In another paper, Larsen et al. [20] studied the orientations of tea catechins as inhibitors of c-Met kinase by using docking experiments.

In the current work, we analyzed the three-dimensional (3D) positioning of 3-fluoro-4-(pyrrolo[2,1-*f*][1,2,4]triazin-4-yloxy)aniline (FPTA), 3-fluoro-4-(1H-pyrrolo[2,3-b]pyridin-4-yloxy)aniline (FPPA), and 4-(4-amino-2-fluorophenoxy)-2-pyridinylamine (AFPP) derivatives [8–12] inside the active site of c-Met kinase. Our approach was to obtain bioactive conformations of the inhibitors by molecular docking, and to analyze the obtained docked positions. This docking analysis can bring to light the most probable binding modes for each ligand. Our second goal was to develop robust and validated QSAR models. QSAR is a method used in drug design, which benefits from the latest bioinformatics databases and in silico tools. With QSAR analysis, it is possible to analyze the probable structural elements affecting the biological activity of compounds, and the model can be used as an artificial tool in designing new bioactive compounds. Here, we report QSAR predictive models by using comparative molecular similarity analysis (CoMSIA) [35] and topological vectors applied to docked structures.

Materials and methods

Docking

In molecular docking, we attempt to predict the structure (or structures) of the intermolecular complex formed between two or more molecules. Docking has been widely used to suggest the binding modes of protein inhibitors [36–40]. Most docking algorithms are able to generate a large number of possible structures, and so they also require a means to score each structure to identify those of most interest. In general, the "docking problem" is thus concerned with the generation and evaluation of plausible structures of intermolecular complexes.

Docking was performed by using ICM methodology [41]. The characteristics of ICM docking method have been described in previous reports [19]. The c-Met kinase structure was used as a template for the docking simulations. The protein coordinates were extracted from the X-ray crystal structure of c-Met kinase in complex with compound **B1** (accession code in Protein Data Bank: 3ce3). The structures and c-Met kinase inhibitory activities (IC_{50}) of 103 compounds (including FPTA, FPPA, and AFPP derivatives) were taken from the literature [8–12]. The structures of the inhibitors were sketched by using Molecular Editor [42] of ICM software [43]. Protein and inhibitor structures were converted into ICM objects. During protein conversion process, hydrogens were added and the modified structure was optimized. Meanwhile, during ligand conversions, two-dimensional (2D) representations were converted in three-dimensional (3D) ones, partial charges were assigned, and rotatable bonds were identified. IcmPocketFinder [44] was used for identifying putative pockets with a tolerance value of 4.6. Initial ligand position and orientation, and box position and size, were kept according to the values suggested by the program. The docking poses for each ligand were analyzed by examining their relative total energy score. The more energetically favorable conformation was selected as the best pose.

QSAR modeling

For QSAR modeling, IC_{50} values were converted in logarithmic activities pIC_{50}. The experimental biological activity values IC_{50} (μM) are measures of inhibitory activity, and they refer to the micromolar concentration of the compounds leading to 50% inhibition of c-Met kinase activity. The chemical structures and experimental activities are shown in Table 1 (ChemDraw structures are available in Supplementary Material). Prior to molecular descriptor calculations, the data set was divided in training and test sets. Compounds **B40** and **B41** were not taken into account because they have a high value of error in their measurement of activity. 25 compounds were randomly chosen as a test set and were used for external validation for the QSAR models. 3D conformations obtained from docking experiments were used for calculating the descriptors.

CoMSIA method was applied by using the protocol described in previous studies [23, 24, 45]. In addition, a multiple linear regression (MLR) model was built by using 2D and 3D topological descriptors. Dragon [46] computer software was employed to calculate Geometry, Topology and Atom-Weight AssemblY (GETAWAY) descriptors (197 descriptors) [47] and 2D-Autocorrelation descriptors (96 descriptors) [48]; in addition, Ghose-Crippen molar refractivity (MR) [49], fragment-based polar surface area (PSA) [50], and Miroguchi octanol–water partition coefficient (MlogP) [51] were calculated. Descriptors that stayed constant or almost constant were eliminated in each set, and pairs of variables with a correlation coefficient greater than 0.95 were classified as intercorrelated, and only one of these was included for succeeding selection steps. After this, 139 descriptors were obtained. Since many molecular descriptors were available for each QSAR analysis and only statistically significant variables were desired, variable selection was addressed. Genetic algorithm (GA) search was carried out in order to build QSAR models. MLR models were developed including 9 or 10 variables for each set of descriptors to obtain predictive structure–activity equations. The quality of each model was proven by the square multiple correlation coefficient (R^2) and the standard deviation (SD). The models with R^2-value above 0.6 were selected and they were tested in cross-validation experiments. GA search was carried out exploring MLR models by using BuildQSAR software [52].

The quality of the fit of the training set of CoMSIA and MLR models was measured by R^2-values. However, a most important measure is the prediction quality. An internal leave-one-out (LOO) cross-validation process was carried out by estimating R^2 of LOO cross-validation (Q^2) and standard deviation (SD_{CV}). A data point was removed (left-out) from the training set, and the model was refitted; the predicted value for that point was then compared to its actual value. This was repeated until each data point was omitted once; the sum of squares of these deletion residuals can then be used to calculate Q^2. In addition, the predictive power of the model was also measured by an external validation process that consists in predicting the activity of unknown compounds forming the test set. In this case, the residuals of predictions were evaluated. Such criteria have been formulated as the requirements for a QSAR model to have highly predictive power.

Results and discussion

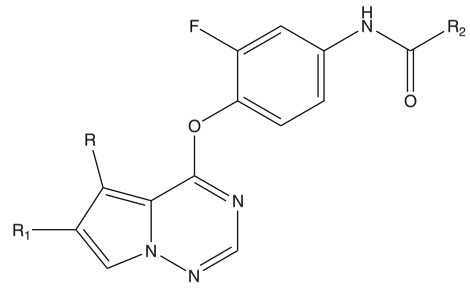
Docking results

The first step in our docking study was the application of docking methodology in order to reproduce the crystal structures of c-Met kinase-inhibitor complexes for compounds **A23**, **B1**, **B23**, **C2**, **C7**, and **C25** (accession codes in PDB: 3c1x, 3ce3, 3ctj, 3l8v, 3cth, and 3f82 respectively). In Fig. 1 is shown the alignment of the cocrystallized structure versus the docked structure for all inhibitors used to validate the docking methodology. According to this figure, the docked structures fitted in an optimal way with the inhibitor X-ray crystal structures. *N,N'*-disubstituted urea group of compounds **A23**, **B23**, and **C7** are in cis,trans conformation in crystal structures; however, they were obtained in trans,trans conformation by docking simulation (Fig. 1a, c, e). Some cis,trans conformations were obtained for compounds **B23** and **C7** (and some of the analogous compounds in the whole dataset), but they were not representative of the docking results and they had poor scores. We ascribe these results to the ICM algorithm, which did not consider the cis,trans conformations as more favourable; however, the inhibitors were adequately oriented. In Table 2 are presented the values of the root mean square deviation (rmsd) for the docked structures with respect to the cocrystal inhibitor structures considering all heavy atoms. Rmsd values were <2.0 Å in all the cases analyzed, therefore, we can state that software ICM found a correct binding mode of the studied ligands.

The remaining compounds were docked following the same docking protocol. Comparison of the different docking results for all 103 ligands shows that, in principle, all compounds adopt the same binding mode (Fig. 2). This similar binding mode is not surprising since all compounds contain related scaffolds. In general, all compounds kept the hydrophobic interactions between the N-substituent of 3-fluoroaniline moiety (R_2 substituent for series A and series B, and R_1 substituent for series C in Table 1) and the big hydrophobic pocket inside the allosteric site formed by Ala1221, Phe1134, Val1220, His1202, Phe1200, Leu1195,

Table 1 Experimental and predicted c-Met kinase inhibitory activities (IC_{50}) of 3-fluoro-4-(pyrrolo[2,1-*f*][1,2,4]triazin-4-yloxy)aniline (FPTA), 3-fluoro-4-(1H-pyrrolo[2,3-b]pyridin-4-yloxy)aniline (FPPA), and 4-(4-amino-2-fluorophenoxy)-2-pyridinylamine (AFPP) derivatives using QSAR models

A



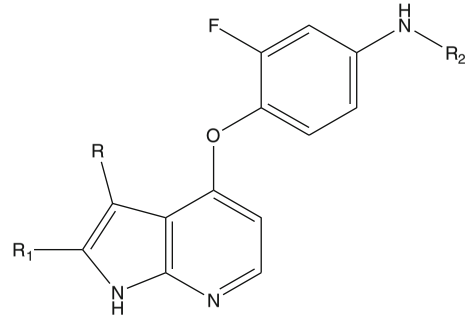
Compound[b]	R	R_1	R_2	IC_{50}(μM)		
				Exp.	CoMSIA Pred.[c]	MLR Pred.[c]
A1	Me	H	2-fluorophenyl	0.76	0.71	0.66
A2	Me	2-(4-methyl-1-piperazinyl)ethoxy	2-fluorophenyl	0.09	0.12	0.16
A3	Me	2-(4-methyl-1-piperazinyl)ethoxy	2,5-difluorophenyl	0.04	0.05	0.28
A4	Me	2-(4-methyl-1-piperazinyl)ethoxy	2-fluoro-5-methylphenyl	0.09	0.10	0.11
A5	Me	2-(4-methyl-1-piperazinyl)ethoxy	3-acetylphenyl	0.09	0.06	0.04
A6	Me	2-(4-methyl-1-piperazinyl)ethoxy	5-chloro-2-fluorophenyl	0.07	0.10	0.06
A7	Me	2-(4-methyl-1-piperazinyl)ethoxy	3-thienyl	0.19	0.14	0.41
A8	Me	2-(4-morpholinyl)ethoxy	2-fluoro-5-methylphenyl	0.51	0.36	0.21
A9	Me	3-(dimethylamino)propoxy	2,6-difluorophenyl	0.03	0.04	0.04
A10	Me	H	2-(4-fluoroanilino)-2-oxoethyl	0.78	0.87	0.56
A11	Me	2-(4-methyl-1-piperazinyl)ethoxy	2-(4-fluoroanilino)-2-oxoethyl	0.04	0.03	0.19
A12	Me	2-(4-morpholinyl)ethoxy	2-(4-fluoroanilino)-2-oxoethyl	0.13	0.18	0.09
A13	Me	3-(dimethylamino)propoxy	2-(4-fluoroanilino)-2-oxoethyl	0.03	0.05	0.02
A14	Me	2-(4-methyl-1-piperazinyl)ethoxy	[(4-fluorophenyl)acetyl]amino	0.05	0.05	0.07
A15	Me	2-(4-morpholinyl)ethoxy	[(4-fluorophenyl)acetyl]amino	0.15	0.11	0.09
A16	Cl	H	[(4-fluorophenyl)acetyl]amino	2.00	0.78	3.89
A17	Ph	H	[(4-fluorophenyl)acetyl]amino	0.26	0.35	0.04
A18	[(3-aminopropyl)amino]methyl	H	[(4-fluorophenyl)acetyl]amino	1.41	1.02	0.22
A19	(4-methyl-1-piperazinyl)methyl	H	[(4-fluorophenyl)acetyl]amino	0.85	1.07	0.30
A20	4-morpholinylmethyl	H	[(4-fluorophenyl)acetyl]amino	1.00	1.20	0.17
A21	[(3*R*)-3-aminopyrrolidinyl]methyl	H	[(4-fluorophenyl)acetyl]amino	0.44	0.55	0.30
A22	[3-(aminomethyl)-1-azetidinyl]methyl	H	[(4-fluorophenyl)acetyl]amino	0.18	0.27	0.20
A23	(4-amino-1-piperidinyl)methyl	H	[(4-fluorophenyl)acetyl]amino	0.045	0.089	0.331
A24	(4-hydroxy-1-piperidinyl)methyl	H	[(4-fluorophenyl)acetyl]amino	0.89	0.36	0.22

Table 1 continued

A

Compound[b]	R	R_1	R_2	IC_{50}(μM)		
				Exp.	CoMSIA Pred.[c]	MLR Pred.[c]
A25	(4-piperidinylamino)methyl	H	[(4-fluorophenyl)acetyl]amino	0.15	0.42	0.24
A26	[4-(aminomethyl)-1-piperidinyl]methyl	H	[(4-fluorophenyl)acetyl]amino	1.00	0.42	0.19
A27	[(3*R*,4*R*)-4-amino-3-hydroxypiperidinyl]methyl	H	[(4-fluorophenyl)acetyl]amino	0.037	0.031	0.372
A28	Me	H	2-(4-fluoroanilino)-2-oxoethyl	0.78	0.87	0.59
A29[a]	Me	2-(4-methyl-1-piperazinyl)ethoxy	2,6-difluorophenyl	0.03	**0.11**	**0.11**
A30[a]	Me	2-(4-methyl-1-piperazinyl)ethoxy	3-pyridinyl	0.16	**0.09**	**0.32**
A31[a]	Me	3-(dimethylamino)propoxy	3-pyridinyl	0.07	**0.05**	**0.05**
A32[a]	Me	2-(dimethylamino) ethoxy	2-(4-fluoroanilino)-2-oxoethyl	0.14	**0.06**	**0.13**
A33[a]	Me	3-(dimethylamino)propoxy	[(4-fluorophenyl)acetyl]amino	0.04	**0.07**	**0.06**
A34[a]	1-piperazinylmethyl	H	[(4-fluorophenyl)acetyl]amino	0.41	**0.43**	**0.54**
A35[a]	[(3*S*)-3-aminopyrrolidinyl]methyl	H	[(4-fluorophenyl)acetyl]amino	0.19	**0.26**	**0.35**
A36[a]	[4-(acetylamino)-1-piperidinyl]methyl	H	[(4-fluorophenyl)acetyl]amino	0.44	**0.32**	**0.04**
A37[a]	(4-aminocyclohexylidene)methyl	H	[(4-fluorophenyl)acetyl]amino	0.0015	**0.2818**	**0.0832**

B



Compound[b]	R	R_1	R_2	IC_{50}(μM)		
				Exp.	CoMSIA Pred.[c]	MLR Pred.[c]
B1	H	H	[1-(4-fluorophenyl)-2-oxo-1,2-dihydro-3-pyridinyl]carbonyl	0.0018	0.0049	0.0093
B2	H	H	[5-(4-fluorophenyl)-4-oxo-1,4-dihydro-3-pyridinyl]carbonyl	0.0028	0.0019	0.0324
B3	H	H	(1-oxide-6-phenyl-2-pyridiniumyl)carbonyl	0.0013	0.0007	0.0008
B4	H	H	(1-*tert*-butyl-2-oxo-1,2-dihydro-3-pyridinyl)carbonyl	1.00	0.36	0.17
B5	H	H	(1-cyclopropyl-2-oxo-1,2-dihydro-3-pyridinyl)carbonyl	1.00	0.36	0.17
B6	H	H	(1-isopentyl-2-oxo-1,2-dihydro-3-pyridinyl)carbonyl	0.030	0.042	0.030

Table 1 continued

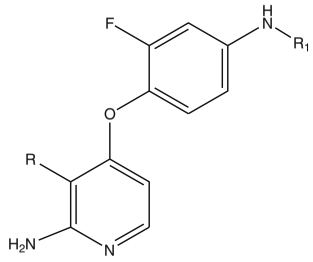
B						
Compound[b]	R	R_1	R_2	$IC_{50}(\mu M)$		
				Exp.	CoMSIA Pred.[c]	MLR Pred.[c]
B7	H	H	[1-(2-pyridinyl)-2-oxo-1,2-dihydro-3-pyridinyl]carbonyl	0.010	0.011	0.126
B8	H	H	[1-(3-pyridinyl)-2-oxo-1,2-dihydro-3-pyridinyl]carbonyl	0.070	0.098	0.269
B9	H	H	{1-[4-(aminocarbonyl)phenyl]-2-oxo-1,2-dihydro-3-pyridinyl}carbonyl	1.00	1.58	0.14
B10	H	H	[5-(4-fluorophenyl)-1-(2-hydroxyethyl)-4-oxo-1,4-dihydro-3-pyridinyl]carbonyl	0.0038	0.0076	0.0009
B11	H	H	[1-cyclopropyl-5-(4-fluorophenyl)-4-oxo-1,4-dihydro-3-pyridinyl]carbonyl	0.0034	0.0020	0.0037
B12	H	H	[1-(4-chloro-3-fluorophenyl)-5-(4-fluorophenyl)-4-oxo-1,4-dihydro-3-pyridinyl]carbonyl	0.005	0.003	0.002
B13	H	H	[6-(4-fluorophenyl)-1-oxide-2-pyridiniumyl]carbonyl	0.0018	0.0009	0.0083
B14	H	H	3-(2-toluidinocarbonyl)-2-pyridinyl	0.006	0.004	0.010
B15	H	H	3-[(2-chloroanilino)carbonyl]-2-pyridinyl	0.024	0.038	0.023
B16	H	H	3-{[4-fluoro(methyl)anilino]carbonyl}-2-pyridinyl	0.19	0.06	0.05
B17	H	H	3-[(benzylamino)carbonyl]-2-pyridinyl	0.021	0.017	0.015
B18	H	H	3-{[2-(2-hydroxyethyl)anilino]carbonyl}-2-pyridinyl	0.009	0.009	0.005
B19	H	H	3-[(2,4-difluoroanilino)carbonyl]-6-methyl-2-pyridinyl	0.018	0.024	0.050
B20	H	H	5-bromo-3-[(2,4-difluoroanilino)carbonyl]-2-pyridinyl	0.009	0.014	0.014
B21	H	H	3-[(2,4-difluoroanilino)carbonyl]-4-pyridinyl	0.27	0.14	0.04
B22	H	H	3-[(2,4-difluoroanilino)carbonyl]-2-pyrazinyl	0.055	0.151	0.032
B23	H	H	{[(4-fluorophenyl)acetyl]amino}carbonyl	0.11	0.05	0.05
B24	(3-hydroxy-1-pyrrolidinyl)acetyl	H	{[(4-fluorophenyl)acetyl]amino}carbonyl	0.019	0.008	0.010
B25	(4-methyl-1-piperazinyl)acetyl	H	{[(4-fluorophenyl)acetyl]amino}carbonyl	0.011	0.007	0.012
B26	[(3-pyridinylmethyl)amino]carbonyl	H	{[(4-fluorophenyl)acetyl]amino}carbonyl	0.011	0.010	0.004
B27	Hydroxymethyl	H	{[(4-fluorophenyl)acetyl]amino}carbonyl	0.004	0.010	0.006
B28	2-amino-1,3-thiazol-4-yl	H	{[(4-fluorophenyl)acetyl]amino}carbonyl	0.007	0.004	0.013
B29	3-pyridinyl	H	{[(4-fluorophenyl)acetyl]amino}carbonyl	0.002	0.004	0.005
B30[a]	H	H	(2-oxo-1-phenyl-1,2-dihydro-3-pyridinyl)carbonyl	0.0019	**0.0048**	**0.0048**
B31[a]	H	H	[1-(4-fluorobenzyl)-2-oxo-1,2-dihydro-3-pyridinyl]carbonyl	0.0082	**0.0525**	**0.0055**
B32[a]	H	H	[5-(4-fluorophenyl)-1-methyl-4-oxo-1,4-dihydro-3-pyridinyl]carbonyl	0.0031	**0.0027**	**0.0115**
B33[a]	H	H	[1-(3-fluoro-4-methoxyphenyl)-5-(4-fluorophenyl)-4-oxo-1,4-dihydro-3-pyridinyl]carbonyl	0.0023	**0.0589**	**0.0028**
B34[a]	H	H	3-[(2,4-difluoroanilino)carbonyl]-2-pyridinyl	0.007	**0.040**	**0.011**
B35[a]	H	H	3-{[(5-fluoro-2-pyridinyl)amino]carbonyl}-2-pyridinyl	0.034	**0.054**	**0.032**
B36[a]	H	H	3-(4-fluoroanilino)-3-oxopropanoyl	0.064	**0.060**	**0.032**
B37[a]	(dimethylamino)acetyl	H	{[(4-fluorophenyl)acetyl]amino}carbonyl	0.044	**0.005**	**0.011**

Table 1 continued

B

Compound[b]	R	R_1	R_2	$IC_{50}(\mu M)$		
				Exp.	CoMSIA Pred.[c]	MLR Pred.[c]
B38[a]	Methoxycarbonyl	H	{[(4-fluorophenyl)acetyl]amino}carbonyl	0.005	**0.007**	**0.017**
B39[a]	H	3-pyridinyl	{[(4-fluorophenyl)acetyl]amino}carbonyl	0.002	**0.015**	**0.004**
B40	H	H	[5-(4-fluorophenyl)-4-oxo-4*H*-pyran-3-yl]carbonyl	1.40	–	–
B41	H	H	(6-phenyl-2-pyridinyl)carbonyl	0.47	–	–

C



Compound[b]	R	R_1	$IC_{50}(\mu M)$		
			Exp.	CoMSIA Pred.[c]	MLR Pred.[c]
C1	H	3-[(4-chloroanilino)carbonyl]-2-pyridinyl	0.011	0.007	0.012
C2	H	3-[(2,4-difluoroanilino)carbonyl]-2-pyridinyl	0.008	0.014	0.017
C3	H	5-[(2,4-difluoroanilino)carbonyl]-2-(methylsulfanyl)-4-pyrimidinyl	0.009	0.022	0.014
C4	H	5-bromo-3-[(2,4-difluoroanilino)carbonyl]-2-pyridinyl	0.013	0.021	0.020
C5	H	3-[(2,4-difluoroanilino)carbonyl]-2-pyrazinyl	0.069	0.085	0.051
C6	H	3-(4-fluoroanilino)-3-oxopropanoyl	0.10	0.05	0.04
C7	H	{[(4-fluorophenyl)acetyl]amino}carbonyl	0.035	0.034	0.078
C8	3-hydroxy-1-propynyl	{[(4-fluorophenyl)acetyl]amino}carbonyl	0.003	0.007	0.004
C9	3-(dimethylamino)-1-propynyl	{[(4-fluorophenyl)acetyl]amino}carbonyl	0.004	0.007	0.004
C10	2-pyridinylethynyl	{[(4-fluorophenyl)acetyl]amino}carbonyl	0.005	0.006	0.003
C11	3-amino-1-propynyl	{[(4-fluorophenyl)acetyl]amino}carbonyl	0.005	0.005	0.005
C12	(4-methyl-1-piperazinyl)methyl	[1-(4-fluorophenyl)-2-oxo-1,2-dihydro-3-pyridinyl]carbonyl	0.0059	0.0060	0.0087
C13	3-amino-1-propynyl	[1-(4-fluorophenyl)-2-oxo-1,2-dihydro-3-pyridinyl]carbonyl	0.0038	0.0021	0.0029
C14	3-hydroxy-1-propynyl	[1-(4-fluorophenyl)-2-oxo-1,2-dihydro-3-pyridinyl]carbonyl	0.0027	0.0020	0.0025
C15	Ethynyl	[1-(4-fluorophenyl)-2-oxo-1,2-dihydro-3-pyridinyl]carbonyl	0.0010	0.0023	0.0004
C16	Isopropyl	[1-(4-fluorophenyl)-2-oxo-1,2-dihydro-3-pyridinyl]carbonyl	0.0034	0.0042	0.0032
C17	Cl	[1-(4-fluorophenyl)-2-oxo-1,2-dihydro-3-pyridinyl]carbonyl	0.0010	0.0018	0.0055

Table 1 continued

C						
Compound[b]	R	R_1	IC_{50}(μM)			
			Exp.	CoMSIA Pred.[c]	MLR Pred.[c]
C18	Cl	[1-(4-fluorophenyl)-4-isopropoxy-2-oxo-1,2-dihydro-3-pyridinyl]carbonyl	0.0032	0.0019	0.0219
C19	Cl	[1-(4-fluorophenyl)-4-(2-hydroxyethoxy)-2-oxo-1,2-dihydro-3-pyridinyl]carbonyl	0.0026	0.0026	0.0066
C20[a]	H	5-[(2,4-difluoroanilino)carbonyl]-2-(dimethylamino)-4-pyrimidinyl	0.12	**0.01**	**0.02**
C21[a]	H	3-[(2,4-difluoroanilino)carbonyl]-4-pyridinyl	3.50	**0.16**	**0.34**
C22[a]	H	5-[(2,4-difluoroanilino)carbonyl]-4-pyrimidinyl	0.022	**0.089**	**0.041**
C23[a]	Ethynyl	{[(4-fluorophenyl)acetyl]amino}carbonyl	0.002	**0.015**	**0.007**
C24[a]	Hydroxymethyl	[1-(4-fluorophenyl)-2-oxo-1,2-dihydro-3-pyridinyl]carbonyl	0.0034	**0.0016**	**0.0040**
C25[a]	Cl	[4-ethoxy-1-(4-fluorophenyl)-2-oxo-1,2-dihydro-3-pyridinyl]carbonyl	0.0039	**0.0033**	**0.0065**

[a] Test set compounds

[b] Compounds **A1–A27**, **A29–A37** are from Ref. [10]; **B1–B13**, **B30–B33** are from Ref. [11]; **B14–B22**, **B34**, **B35**, **C1–C5**, **C20–C22** are from Ref. [12]; **A28**, **B23–B29**, **B36–B39**, **C6–C11**, **C23** are from Ref. [12], and **C12–C19**, **C24**, **C25** are from Ref. [12], ChemDraw structures are available at Supplementary Material

[c] Test set predictions are in boldface

Leu1140, Met1131, and Val1139. Furthermore, the 3-fluoroaniline moiety was docked near Leu1157 (gatekeeper residue) and forms an edge-face π-stacking interaction with Phe1223 (DFG motif), in all compounds. The cores (pyrrolo[2,1-*f*][1,2,4]triazine for series A, 1*H*-pyrrolo[2,3-*b*]pyridine for series B, and 2-pyridinamine for series C, were placed in the hinge region between residues Met1211 and Ile1084, taking a perpendicular position with respect to the 3-fluoroaniline ring. Some compounds in the dataset also occupied the ribose pocket in the external part of the active site, and others occupied an additional small pocket inside the allosteric site of the c-Met kinase.

According to docking results, five regions were identified in the drug-enzyme complexes under study (see Fig. 3): (i) the scaffold region (SR), where substituent groups which interact with hinge region are located; (ii) external region (ER), where substituent groups that occupy the ribose pocket are located; (iii) central region (CR), where substituent groups that interact with Lys1110 and Asp1222 are located, (iv) allosteric site I (ASI), where big hydrophobic substituent groups are located, and (v) allosteric site II (ASII), where small less hydrophobic groups are located. Each identified region has some substituent group variability in the current dataset. We carefully analyzed the general contribution to the inhibitory potency of different chemical features present in each identified region. For this, we defined several categories for each region, which are shown in top part of Fig. 4.

The substituent group variability of the SR defines the three series of compounds studied (A, B, and C): FPTA, FPPA, and AFPP. Red boxes in Fig. 4 represent the three categories for SR; these chemical features were denoted as 1, 2, and 3 for plotting against inhibitory activity in the bottom part of Fig. 4 (red plot). According to SR chemical feature plot, compounds containing element 1 (pyrrolo[2,1-*f*][1,2,4]triazine scaffold) have less inhibitory activity, in comparison to compounds containing elements 2 and 3 (1*H*-pyrrolo[2,3-*b*]pyridine and 2-pyridinamine scaffolds). In addition, elements 2 and 3 have the same contribution to the potency of inhibitors. As authors previously described, the pyrrolo[2,1-*f*][1,2,4]triazine scaffold forms hydrogen bond (HB) interactions with backbone NH of Met1160 [8]. Instead, 1*H*-pyrrolo[2,3-*b*]pyridine and 2-pyridinamine scaffolds accept HB interactions from backbone NH of Met1160, and donate HB to the carbonyl of Met1160 [11]. Docking experiments reproduced these interactions for the majority of compounds (Fig. 5).

Green boxes in Fig. 4 represent the three chemical categories for ER. This zone can be occupied by large or small substituents, or can be unoccupied; these chemical features were denoted as 1, 2, and 3, respectively, and were plotted against inhibitory activity in the bottom part of Fig. 4 (green plot). According to ER chemical feature plot, for

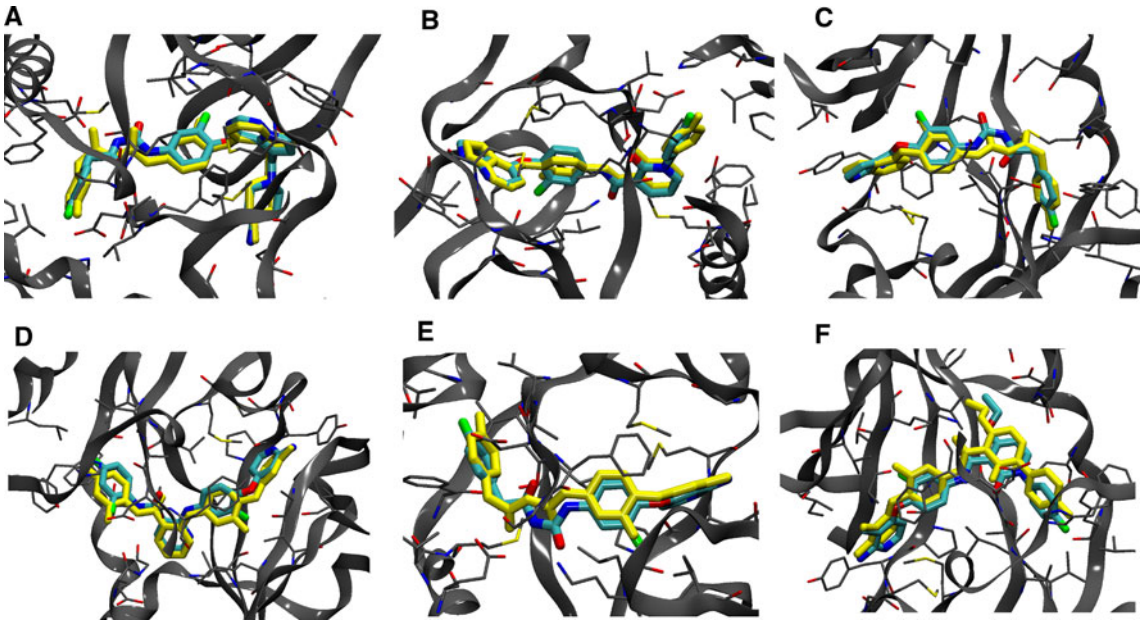


Fig. 1 Alignment of inhibitor docked structures on inhibitor X-ray reference structures, for the c-Met kinase complexes. **a** Compound **A23** (PDB: 3c1x); **b** Compound **B1** (PDB: 3ce3); **c** Compound **B23** (PDB: 3ctj), **d** Compound **C2** (PDB: 3l8v); **e** Compound **C7** (PDB: 3cth); **f** Compound **C25** (PDB: 3f82). Crystal structures are represented in *cyan*, and docking results are represented in *yellow*

Table 2 Docking accuracy

PDB code of c-Met kinase/inhibitor complex	Cocrystallized inhibitor	rmsd values
3c1x	**A23**	0.554
3ce3	**B1**	0.492
3ctj	**B23**	1.882
3l8v	**C2**	0.286
3cth	**C7**	0.406
3f82	**C25**	0.396

Root mean squared deviation (rmsd) values with respect to the reference crystal structure as a measure for the quality of the prediction

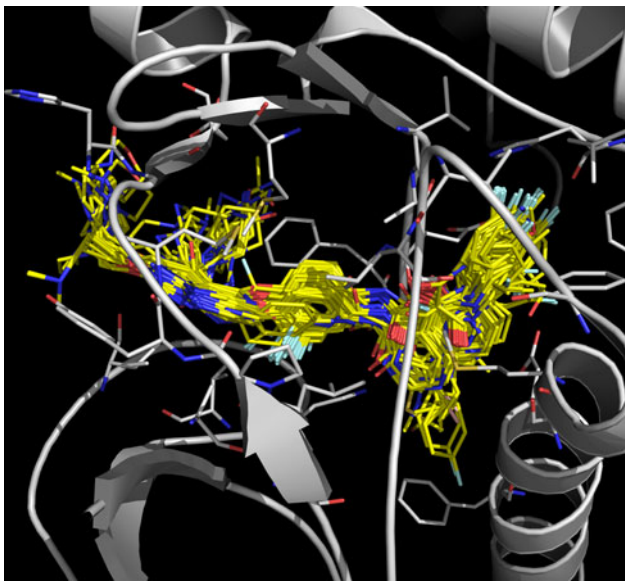


Fig. 2 Alignment of all compound docked structures within the modeled c-Met kinase binding site

FPTA derivatives (red circles), the occupancy of ribose pocket by large groups has a positive effect in the inhibitory activity. Docking poses for compounds containing large groups such as 2-(4-methyl-1-piperazinyl)ethoxy in compound **A2** ($pIC_{50} = 1.05$) or (4-amino-1-piperidinyl)methyl in compound **A23** ($pIC_{50} = 1.35$), showed key ionic interaction with the side chain of Asp1164. Analogous compounds such as **A1** ($pIC_{50} = 0.12$) and **A18** ($pIC_{50} = -0.15$) do not contain any substituent group inside ER and thereby have a poor activity. The absence of a substituent group able to donate an HB to the side chain of Asp1164, also deteriorates the activity of FPTA derivatives (series A). For instance, compound **A8** contains a 2-(4-morpholinyl)ethoxy group (with an HB acceptor oxygen atom facing Asp1164) in ER without establishing an HB interaction with Asp1164; this can explain the poor activity of this compound ($pIC_{50} = 0.29$).

The analysis of ER chemical feature plot also shows that the occupancy of ribose pocket does not have a positive effect in the inhibitory activity for FPPA and AFPP derivatives (blue and green circles in green plot of Fig. 4). However, compounds of these series that occupy ER area are better inhibitors than analogous compounds that do not occupy this region. For instance, FPPA derivatives such as

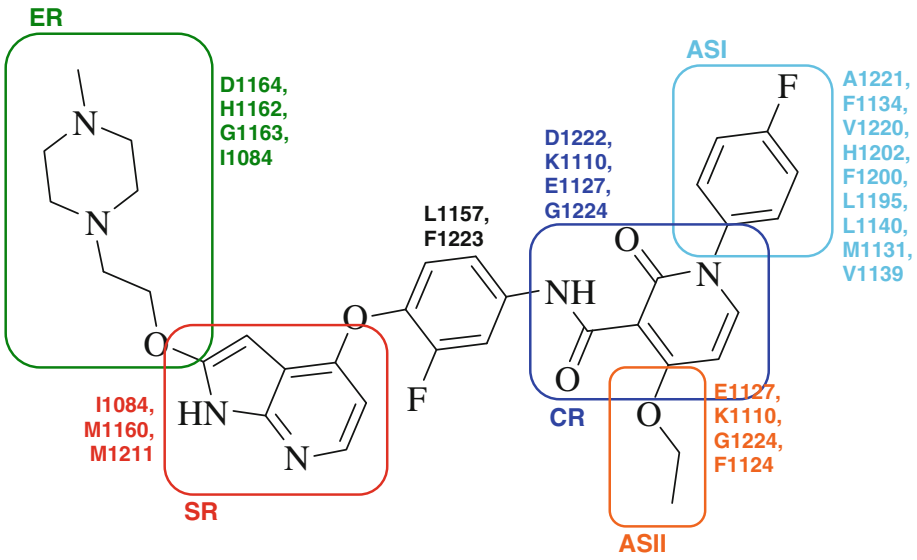


Fig. 3 Regions in which inhibitor-enzyme complexes are divided for discussion purposes. Different regions are represented by different colors

B24–B29 ($pIC_{50} > 1.3$) are more potent inhibitors than compound **B23** ($pIC_{50} = 0.96$), and AFPP derivatives such as **C8–C10** ($pIC_{50} > 2.2$) are more potent inhibitors than compound **C7** ($pIC_{50} = 1.46$). For FPPA derivatives, it is noteworthy that compounds containing small substituent groups are better inhibitors than compounds that contain large substituent groups (chemical features 1 and 2 in green plot, blue circles). For instance, compounds **B27–B29**, **B38**, and **B39** ($pIC_{50} > 2.1$) are more potent inhibitors than compounds **B24–B26**, and **B37** (pIC_{50} between 1.3 and 2). According to our docking experiments, small substituent groups in ER area are located close to side chain of Val1092 and Gly1085. Docking also reveals that none of the large substituent groups that occupy ribose pocket, which were reported in references [11] and [12] on compounds of series B and C, have the interaction with side chain of Arg1164.

Blue boxes in Fig. 4 represent the five chemical categories for CR. This zone can be occupied by groups able to establish HB interactions with Asp1222 and Lys1110 such as 2-oxo-1,2-dihydro-3-pyridinecarboxamide (or similar groups), 2-aminonicotinamide (or similar groups), benzamide, *N*-acetylurea, and malonamide; these chemical features were numbered from 1 to 5, respectively, and they were plotted against inhibitory activity in the bottom part of Fig. 4 (blue plot). According to CR chemical feature plot, the presence of a benzamide group (chemical feature 3) does not contribute to a highly c-Met kinase inhibitory activity. Docked conformations obtained for compounds containing this group establish HBs between the aminocarbonyl NH group of the inhibitor and backbone carbonyl group of Asp1222, and between the aminocarbonyl CO group of the inhibitor and side chain of Lys1110; while the phenyl group is located between side chain of Asp1222, and side chain of Met1131. The majority of compounds containing this group have the allosteric site (ASI region) unoccupied, which can explain their poor activity. However, compound **B41** ($pIC_{50} = 0.33$) has a substituent group inside ASI region and has a poor activity too. The pyridinecarboxamide group (chemical feature 1) is similar to benzamide, but it has an additional carbonyl group. According to docking experiments, this group has the same interactions as benzamide, but it also interacts with backbone NH of Asp1222. This additional HB causes an increase in inhibitory activity of the studied compounds (chemical feature 1 in blue plot of Fig. 4). Change of the ring nitrogen position (compounds **B2** and **B13,** $pIC_{50} = 2.55$ and 2.74, respectively) does not have a negative influence in the activity. Docking of these compounds showed the same position as analogous compound **B1** containing 2-aminonicotinamide group ($pIC_{50} = 2.74$) (Fig. 6). On the other hand, the substitution of the ring nitrogen by an oxygen atom decreases the activity (compound **B40**, $pIC_{50} = -0.15$). Docking of this compound also showed the same position than **B1**; in this sense, we think that the negative effect of 4-oxo-4H-pyran-3-yl group should be explained by considering electronic effects.

According to CR chemical feature plot, compounds containing 2-aminonicotinamide group (chemical feature 2) have $pIC_{50} < 2.3$, in this sense, this group does not contribute to a highly c-Met kinase inhibitory activity. Docked conformations obtained for compounds containing this group form HBs between the aminocarbonyl CO group of the inhibitor and backbone NH of Asp1222, and between the amine group of the inhibitor and backbone CO of Asp1222; while the aromatic ring is located between Gly1224, and side chain of Leu1142. However, this group does not form HB interactions with Lys1110. Instead, the

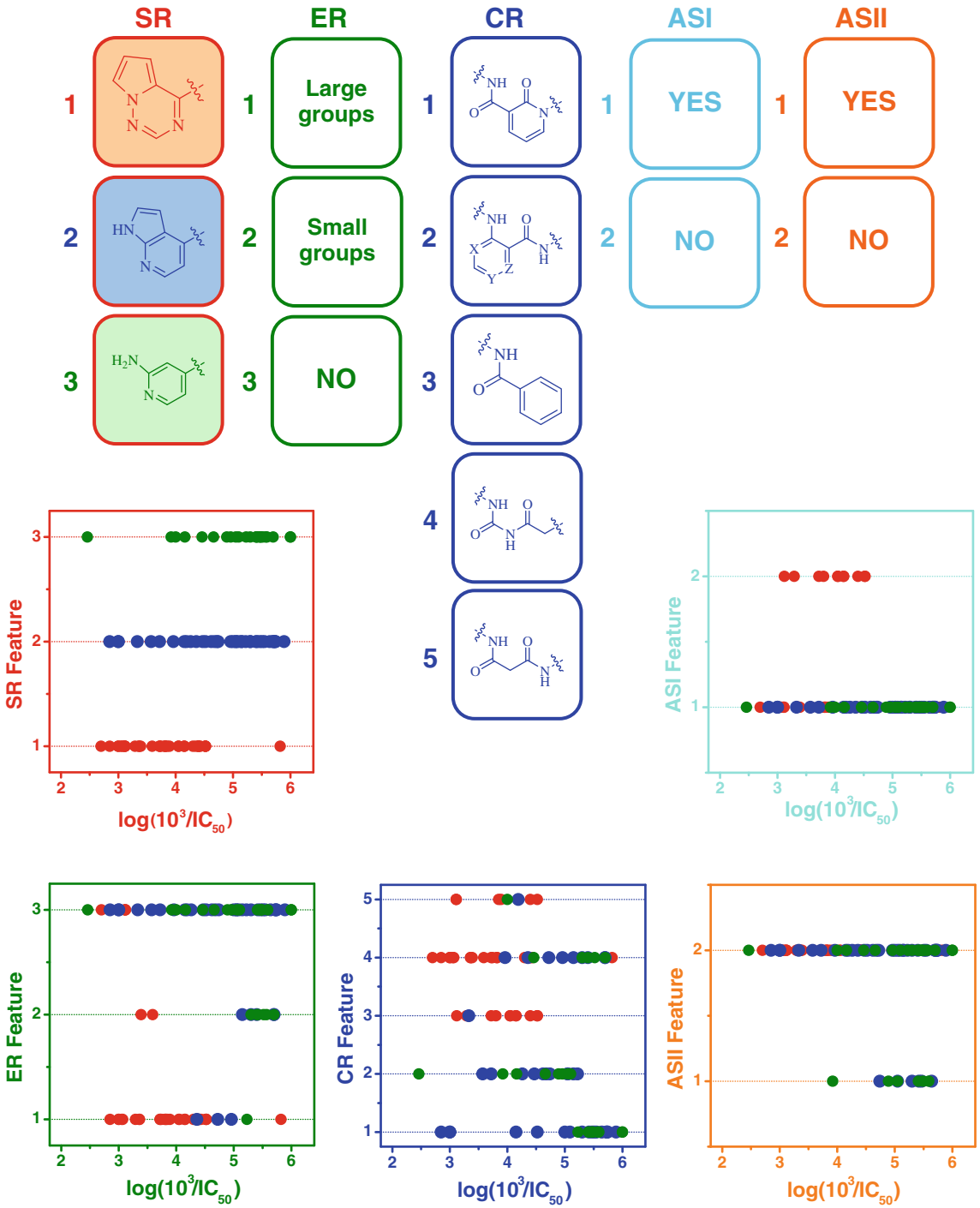


Fig. 4 *Top* Categories according to the chemical features encountered in each region defined in Fig. 3. Each number corresponds to one chemical feature for each region. *Box colours* correspond to those shown in Fig. 3. *Bottom* Chemical feature plots for each region. Each plot shows the chemical features against c-Met inhibitory activity. *Circle colors* in the plots indicate the scaffold of the plotted compound: *red* FPTA, *blue* FPPA, and *green* AFPP. Plot colors correspond to those shown in Fig. 3

aminocarbonyl NH group of the 2-aminonicotinamide group is oriented toward the side chain of Glu1127 (the side chain of Glu1127 has a high mobility when comparing the available X-ray structures of c-Met in PDB). The obtained docking poses suggest that Glu1127 can approach aminocarbonyl NH group and forms an additional HB. Williams et al. [10] realized that methylation of the aminocarbonyl NH group (as in compound **B16**) results in compounds with significantly reduced activity; this could be due to the loss of the interaction with Glu1127. When

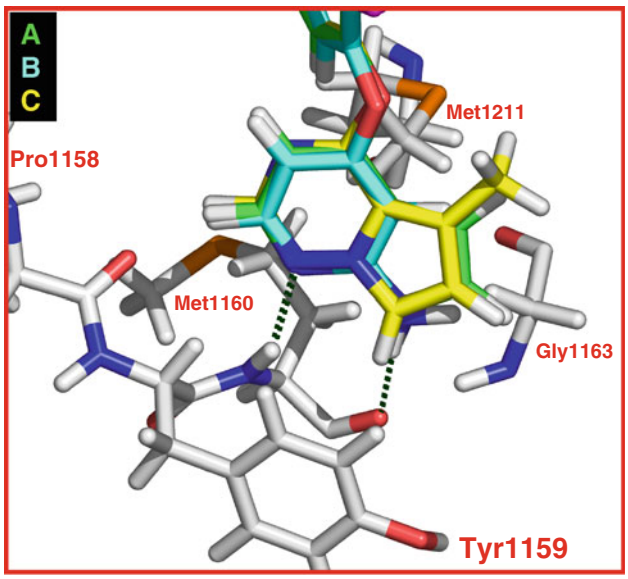


Fig. 5 Orientations of pyrrolo[2,1-*f*][1,2,4]triazine, 1*H*-pyrrolo[2,3-*b*]pyridine, and 2-pyridinamine scaffolds at site SR according to docking experiments

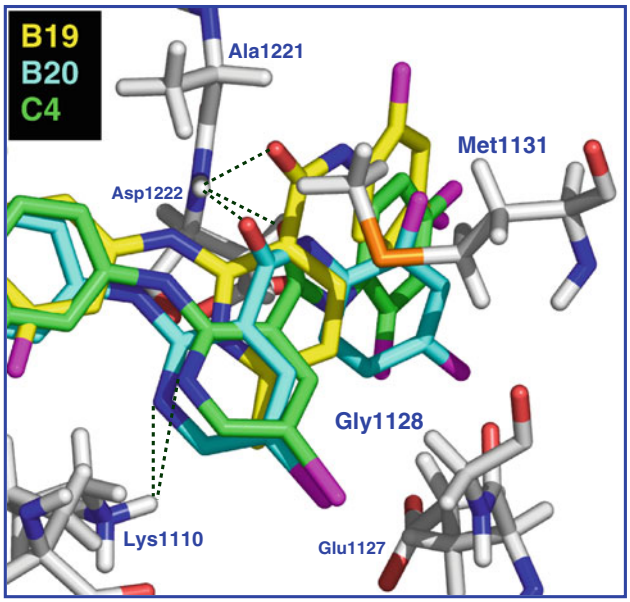


Fig. 7 Orientation of groups at site CR for compounds **B19**, **B20**, and **C4**, according to docking experiments

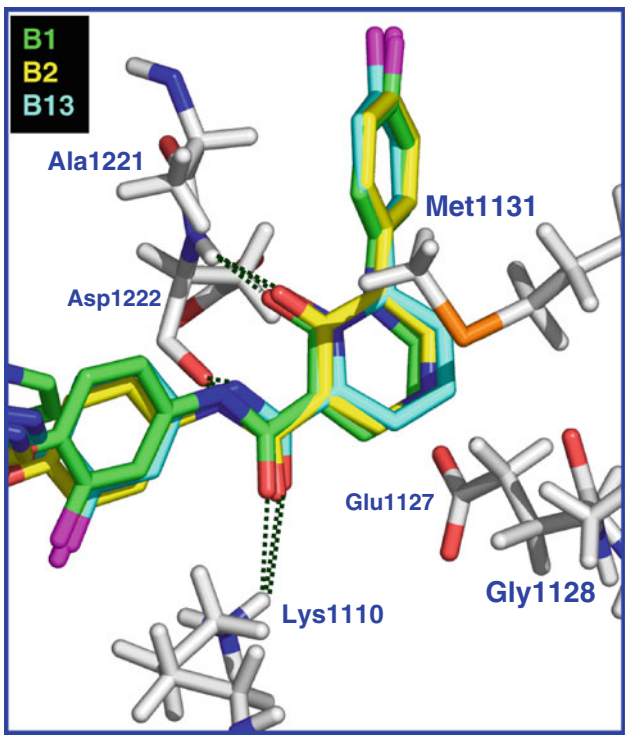


Fig. 6 Orientation of groups at site CR for compounds **B1**, **B2**, and **B13**, according to docking experiments

the pyridine moiety of 2-aminonicotinamide group is changed by other groups (compounds **B22**, **C5** and **C22**), the activity decreases ($pIC_{50} < 1.7$); the same happens when the position of the ring nitrogen is changed (compounds **B21** and **C21**). Docking of the compounds containing these changes at CR showed the same position than analogous compounds containing 2-aminonicotinamide group (compounds **B34** and **C2,** $pIC_{50} = 2.15$ and 2.10, respectively); in this sense, we think that this effect should be explained by considering electronic effects.

Compounds containing Br substituent at position 5 of pyridine moiety of 2-aminonicotinamide group (compounds **B20** and **C4,** $pIC_{50} = 2.05$ and 1.89, respectively) had the same position than analogous compounds since Br group was located at region ASII, between Lys 1110 and Glu1127. On the other hand, compound **B19** ($pIC_{50} = 1.74$) has a methyl substituent at position 6 of pyridine moiety of 2-aminonicotinamide group, and it can not have the same position than analogous compounds. Methyl substituent occupies the position of the pyridine ring due to steric hindrance with Lys1110; but, it does not lose the interactions with Asp1222 (Fig. 7). Though, it changed its mode of interaction in the SR region. According to these results, it is expected that bigger substituent groups at position 6 of pyridine moiety should decrease the activity dramatically, and bigger substituent groups at position 5 should be tolerated.

CR chemical feature plot (blue box in Fig. 4) also shows that compounds containing *N*-acetylurea group (chemical feature 4) are better inhibitors than compounds containing malonamide group (chemical feature 5). X-ray crystal of compounds containing *N*-acetylurea group shows that this group established HB interactions with Asp1222 and Glu1127 [8, 11]. Authors observed that *N*-acetylurea group forms an intramolecular HB, the acetyl carbonyl forms an HB with the backbone NH of Asp1222, and the acetylated

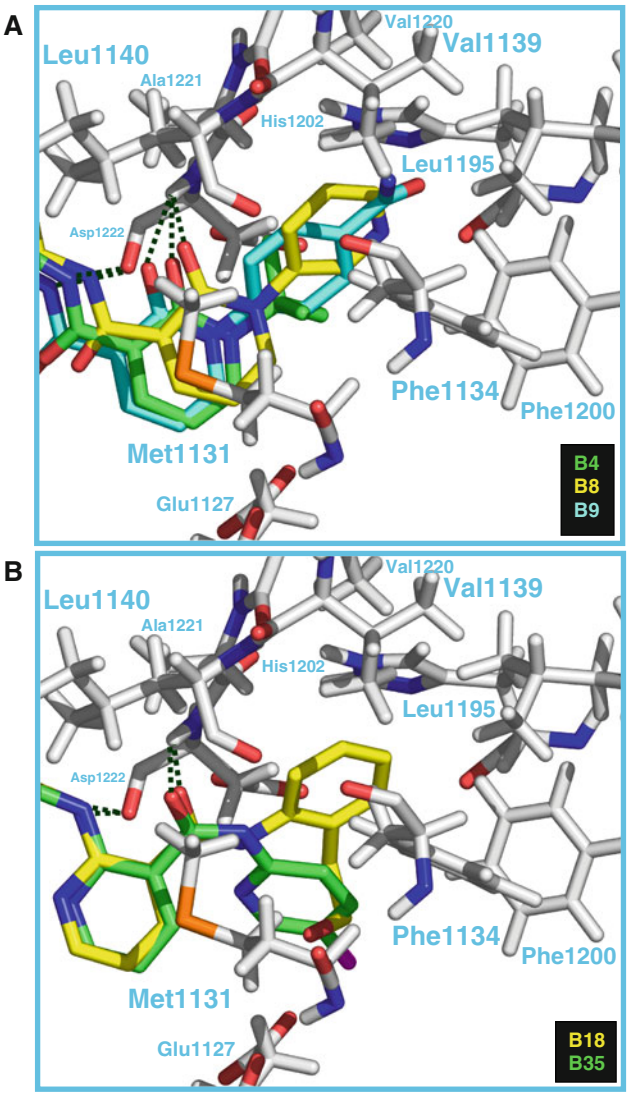


Fig. 8 Orientation of groups at site ASI according to docking experiments: **a** for compounds **B4**, **B8**, and **B9**, **b** for compounds **B18**, and **B35**

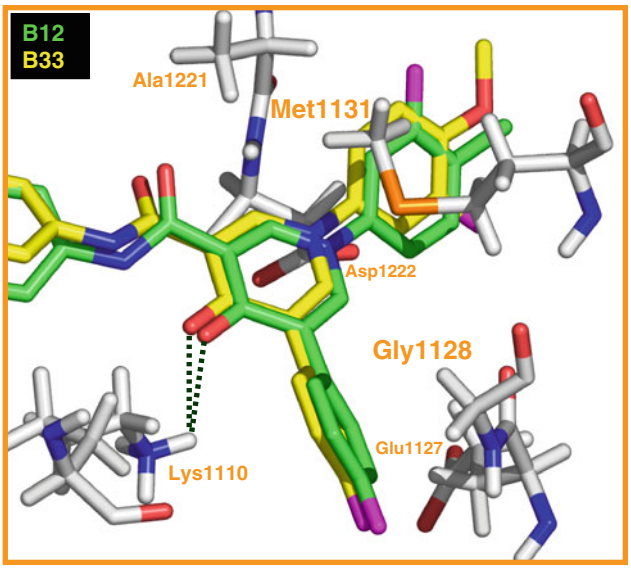


Fig. 9 Orientation of groups at site ASII for compounds **B12** and **B33**, according to docking experiments

NH of urea forms an HB with side chain of Glu1127. The HB interaction with Glu1127 is not possible for compounds containing malonamide group, which explains why compounds containing chemical feature 4 are better inhibitors. Unfortunately, we did not obtain the interaction with Glu1127 by docking experiments for all the compounds containing *N*-acetylurea group. Glu1127 has a high mobility, for this reason we repeated docking experiments for several compounds containing this group (**A23**, **B23** and **C7**), by using other available crystal structures of c-Met kinase (PDB codes: 3c1x, 3cth, 3ctj) that have Glu1127 in most favorable positions for establishing the mentioned HB interaction. However, the HB interaction with Glu1127 was not finally reproduced. This interaction could be studied considering the movement of the system by using molecular dynamics.

Cyan boxes in Fig. 4 represent the two chemical categories for ASI. This zone can be occupied by hydrophobic substituents, or can be unoccupied; these chemical features were denoted as 1 and 2, respectively, and they were plotted against inhibitory activity in the bottom part of Fig. 4 (cyan plot). According to ASI chemical feature plot, the occupancy of ASI has a positive effect in the inhibitory activity. Substituted and unsubstituted phenyl groups seem to be the best option to occupy ASI; the most active compounds, such as **A37**, **B3** and **C15** ($pIC_{50} > 2.8$) have this kind of substituent group. Docking experiments show that these groups occupy the entire pocket conformed by residues Ala1221, Phe1134, Val1220, His1202, Phe1200, Leu1195, Leu1140, Met1131, and Val1139. The presence of substituent groups of a shorter length such as *t*-butyl (compound **B4,** $pIC_{50} = 0.00$), at ASI decreases the inhibitory activity. Docking experiments show that this substituent group occupies the entrance of the site between Met1131, Phe1134, and Asp1222 (Fig. 8a). Less hydrophobic substituent groups such as pyridinyl and 4-aminocarbonylphenyl also have a bad influence in activity (compound **B8** and **B9,** $pIC_{50} = 1.15$ and 0.00, respectively), since they do not establish the proper interactions within the hydrophobic pocket. Several compounds placed the hydrophilic part of the substituent designed to occupy ASI close to Glu1127. For instance, compound **B18** ($pIC_{50} = 2.05$) has the phenyl group inside ASI, and orients its hydroxyethyl moiety towards side chain of Glu1127; while compound **B35** ($pIC_{50} = 1.47$) orients its

pyridine ring between side chains of Glu127 and Asp1222, out of ASI (Fig. 8b).

Orange boxes in Fig. 4 represent the two chemical categories for ASII. This zone can be occupied, or can be unoccupied; these chemical features were denoted as 1 and 2, respectively, and they were plotted against inhibitory activity in the bottom part of Fig. 4 (orange plot). According to ASII chemical feature plot, the occupancy of ASII does not have a negative effect in the inhibitory activity. Several compounds have a hydrophobic group inside ASI, and an additional smaller substituent group inside other pocket between Glu1127, Lys1110, Gly1224, and Phe1124 at the end of the cavity. Compounds **B10**, **B11**, and **B32** ($pIC_{50} = 2.42$, 2.47, and 2.51, respectively), were derived from **B2** ($pIC_{50} = 2.55$), but they have a substituent group inside ASII; the activities suggest that the occupancy of ASII did not contribute to a better activity for these compounds. The same happens for compounds **C18**, **C19**, and **C25** ($pIC_{50} = 2.49$, 2.59, and 2.41 respectively), which were derived from **C17** ($pIC_{50} = 3.00$). Docking also shows that compounds **B12** and **B33** ($pIC_{50} = 2.30$ and 5.64 respectively) have large substituent groups that were accommodated inside ASI, while 4-F-phenyl group was placed inside ASII (Fig. 9). This change causes that carbonyl of pyridinone group at CR loses its interaction with the backbone NH of Asp1222, although it can establish a new interaction with side chain of Lys1110. According to these results, ASII tolerates hydrophobic and hydrophilic substituent groups, when ASI was previously occupied; but it does not modulate ligand-enzyme interaction.

QSAR results

In an attempt to find a structure–activity relationship, we plotted the ICM docked binding energies against the pIC_{50} values. No correlation between these two variables was found ($R^2 < 0.1$); therefore, we carried out QSAR models.

First, we carried out CoMSIA models. The predictability of the models is the most important criterion for assessment of QSAR methods. We found that all CoMSIA models using one field were statistically unacceptable ($Q^2 < 0.5$) (data not shown). Fifteen models including two or more fields that have $Q^2 > 0.5$ were found. Table 3 reports the obtained models that have $Q^2 > 0.6$. The best CoMSIA model (CoMSIA-SEHD) has a Q^2 value of 0.654 using six components and includes steric, electrostatic, hydrophobic, and HB donor fields. The model explains 93.1 of the variance, has a low standard deviation ($s = 0.253$) and a high Fischer ratio ($F = 154.57$). The model showed similar contributions of the electrostatic, hydrophobic, and HB donor fields of 29.3%, 28.6% and 32.7%, repectively; and a steric contribution of 9.3%. The predictions of pIC_{50} values for the 76 studied compounds in the training set using CoMSIA-SEHD model are shown in Table 1. The correlations between the calculated and experimental values of pIC_{50} (from training and LOO crossvalidation) are shown in Fig. 10a.

The contour plots of the CoMSIA fields are presented in Fig. 11 for the modeled c-Met kinase inhibitors. To aid in visualization, compound **C15** is displayed in the maps. In Fig. 11a, the green isopleth at ER close to Ile1084 indicates region where a bulky group favored the activity. Compounds with small substituents at this zone, such as **B27–B29**, are good c-Met kinase inhibitors.

Blue isopleths indicate regions where positive charges favored the activity, or negative charges disfavored the activity (Fig. 11a). A blue isopleth is observed at ER close to side chain of Asp1164 which indicates that positively charged groups at this zone favor the activity. Other blue isopleths are located at SR, and at CR on nitrogen atoms of 2-oxo-3-pyridinecarboxamide group of compound **C15**. These isopleths indicate that positively charged groups at these regions are essential to establish the interactions inside the c-Met kinase pocket.

Red isopleths indicate regions where negative charges favored the activity, or positive charges disfavored the activity (Fig. 10b). A red isopleth close to Ile1084 indicates that negatively charged groups are tolerated in this

Table 3 CoMSIA analysis results using several different field combinations

Model	NC	R^2	S	F	Q^2	Scv	Fraction				
							Steric	Electrostatic	Hydrophobic	H-bond donor	H-bond acceptor
CoMSIA-SHD	6	0.926	0.261	144.84	0.640	0.577	0.150		0.406	0.444	
CoMSIA-EHD	6	0.921	0.270	134.66	0.623	0.590		0.333	0.315	0.352	
CoMSIA-SEHD	**6**	**0.931**	**0.253**	**154.57**	**0.654**	**0.565**	**0.093**	**0.293**	**0.286**	**0.327**	
CoMSIA-SHDA	6	0.927	0.259	146.98	0.605	0.604	0.125		0.321	0.342	0.212
CoMSIA-ALL	6	0.922	0.268	136.84	0.613	0.598	0.082	0.243	0.244	0.269	0.162

Only models with $Q^2 > 0.6$ are included. The best model is in bold letters. NC is the number of components from partial least square analysis, R^2 is the square of correlation coefficient, S is the standard deviation of the regression, F is the Fisher ratio, Q^2 and Scv are the correlation coefficient and standard deviation of the leave-one-out (LOO) cross-validation, respectively

zone. In fact, compounds with 2-(4-methyl-1-piperazinyl)ethoxy group have the ethoxy oxygen atom at this position. Another red isopleth close to Val1092 suggests that positively charged groups are not desired for a high c-Met kinase inhibitory activity. In fact, compounds with amine-containing substituents at position 5 of pyrrolo[2,1-f][1,2,4]triazine moiety of FPTA derivatives, such as **A18**–**A21**, have the amine nitrogen at this position, and are less active ($pIC_{50} \sim 0$). Other red isopleth located near Lys1110 suggests that this zone can be occupied by negatively charged groups. According to docking results, the oxygen atom of 2-fluorophenoxy moiety of some active compounds can occupy this zone (for instance: compounds **B2**, **B3**, **B7**, and **B13**). Finally, red isopleths located at SR near Pro1158 and at site ASI indicate that positively charged positive groups are not tolerated in these zones.

In Fig. 11b, yellow contours represent areas where substitution of hydrophobic groups would be favorable for activity, while white contours represent regions where placement of hydrophobic groups diminishes the c-Met kinase inhibitory activity. There are yellow isopleths at ER between Phe1223 and Asp1164, and at site ASI. In these regions hydrophobic groups are suggested by the model to contribute to a high c-Met kinase inhibitory activity. In the other hand, there are white isopleths at SR close to Ile1084, at CR close to backbone of Asp1222, and in the surroundings of 2-fluorophenoxy moiety. In these regions, hydrophobic groups do not favour the activity. The white isopleth at SR indicates that the model identified that 2-pyridinylamine is better scaffold than pyrrolo[2,1-*f*][1,2,4]triazine and 1H-pyrrolo[2,3-b]pyridine. The white isopleth at CR indicates that hydrophobic groups such as the benzamide group (compounds **A1**–**A3**) are not desired in this zone for a high activity.

Cyan isopleths represent regions where HB donor groups favor activity (Fig. 11b). There is a cyan isopleth at SR close to the backbone of Met1160, which indicates the requirement of an HB donor group to establish HB interactions in this zone. Another cyan isopleth is located at the entrance of site ASI, close to Ala1221, which indicates that HB donor groups in this zone are desired. According to this, the presence of *N*-acetylurea group favors the activity. Finally, two purple isopleths are shown in Fig. 11b. The first one close to backbone of Asp1222 indicates region

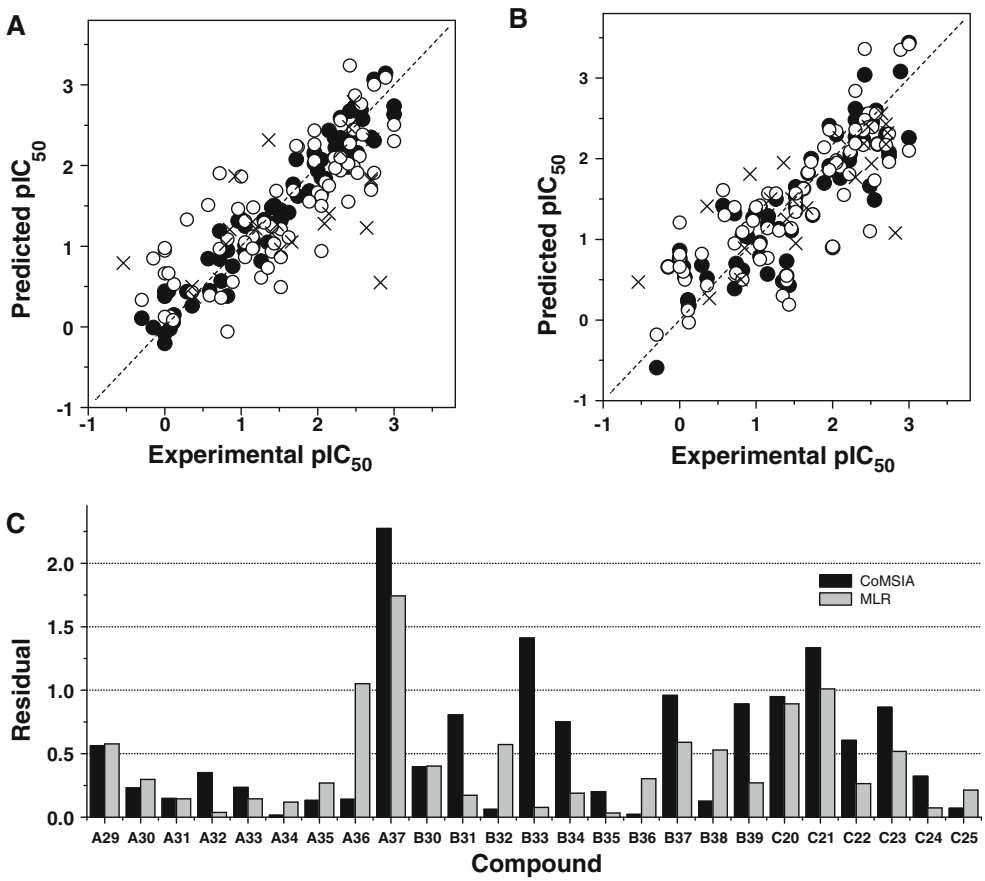


Fig. 10 Scatter plot of the experimental activities versus predicted activities for CoMSIA-SEHD (**a**) and MLR (**b**) models: (*filled circle*) training set predictions (*open circle*) LOO cross-validated predictions (*multiple symbol*) test set predictions. **c** Residuals of the predictions for test set compounds by using both models

where an HB donor groups are disfavored for activity. The second one is located near Gly1128 and is related to the presence of 2-aminonicotinamide group, which is not favorable for a high activity.

Models developed by using 3D non-interpretable vectors are limited when rational interpretation of them is achieved; however, they can be employed as artificial predictive tools. With this in mind, we developed models including topological molecular descriptors. A set of descriptors was used for modeling c-Met inhibition of FPTA, FPPA, and AFPP derivatives. 3D conformations obtained by docking experiments were used for calculating all descriptors. MLR analysis combined with a GA feature selection was performed on the training set (Table 1). The best model was chosen as the most reliable relationship between the calculated descriptors and the binding affinity pIC_{50}. The obtained linear model is shown in Eq. 1:

$$\begin{aligned} pIC_{50} &= 13.602 \times MATS1v - 14.767 \times MATS2e + 3.118 \\ &\quad \times MATS7e - 15.833 \times I_{SH} - 57.368 \times R8u^{+} - 5.088 \\ &\quad \times R2m - 9.295 \times R3m^{+} + 29.528 \times R1v^{+} + 0.032 \\ &\quad \times PSA + 0.943 \times M\log P + 16.905 \\ N &= 76 \quad R^2 = 0.733 \quad SD = 0.489 \quad F = 20.186 \\ Q^2 &= 0.668 \quad SD_{CV} = 0.532 \end{aligned} \tag{1}$$

where N is the number of compounds included in the training set, R^2 is the square of correlation coefficient, SD is the standard deviation of the regression, F is the Fisher ratio, Q^2 and SD_{CV} are the correlation coefficient and standard deviation of the leave-one-out (LOO) cross-validation, respectively. The symbols and definition of the descriptors showed in Eq. 1 are given in Table 4. The linear model in Eq. 1 is able to explain about 73.3% of c-Met kinase inhibitory activity variance. Since the Q^2 value was about 0.668, the model was considered to be a good predictive one. The training-set predictions (log ($10^3/IC_{50}$)) for all compounds in the training set are presented in Table 1. In turn, plots of training set and LOO cross-validation predictions versus experimental pIC_{50} values are shown in Fig. 10b.

The obtained model includes three Moran's 2D autocorrelation descriptors (*MATS1v*, *MATS2e*, and *MATS7e*), five GETAWAY descriptors (I_{SH}, $R8u^{+}$, *R2 m*, *R3* m^{+}, and $R1v^{+}$), fragment-based polar surface area (PSA), and Miroguchi octanol–water partition coefficient (MlogP). There is no significant intercorrelation between the linear GA selected descriptors, as shown in Table 5. The variables and statistics of the eleven better models according to Q^2 values are reported in Supplementary material. From the analysis of these models, we can observe that descriptors *MATS1v*, *MATS2e*, *MATS7e*, $R1v^{+}$, PSA, and MlogP appear in five or more than five models with similar but slightly lesser quality. This suggests that these variables are more relevant for the studied activity in relation to the other variables in Eq. 1.

Moran's 2D autocorrelation descriptors test whether the value of an atomic property at one atom in the molecular structure is independent of the values of the property at neighboring atoms. If dependence exists, the property is said to exhibit spatial autocorrelation. The autocorrelation vectors capture the degree of similarity between molecules from the 2D topological graph

$$MATS(l, p_k) = \frac{N}{2L} \frac{\sum_{ij} \delta_{ij}(p_{ki} - \bar{p}_k)(p_{kj} - \bar{p}_k)}{\sum_{i} (p_{ki} - \bar{p}_k)} \tag{2}$$

where $MATS(l,\ p_k)$ is Moran's index at spatial lag l weighted by property p_k. Besides, p_{ki} and p_{kj} are the values of property k of atom i and j respectively; $\bar{p}_k$ is the average value of property k, L is the number of nonzero values in the sum, N is the number of atoms in the molecule, and $\delta(l,\ d_{ij})$ is a Dirac-delta function defined as

$$\delta(l, d_{ij}) = \left\{ \begin{array}{ll} 1 & \text{if } d_{ij} = l \\ 0 & \text{if } d_{ij} \neq l \end{array} \right\} \tag{3}$$

where d_{ij} is the topological distance or spatial lag between atoms i and j. Equation 1 indicates that *MATS1v*, an atomic

Table 4 Symbols of the descriptors used in the reported model and their definitions

Symbols	Definitions
MATS1v	Moran autocorrelation of lag 1 weighted by atomic van der Waals volumes
MATS2e	Moran autocorrelation of lag 2 weighted by atomic Sanderson electronegativities
MATS7e	Moran autocorrelation of lag 7 weighted by atomic Sanderson electronegativities
I_{SH}	Standarized information content on the leverage equality
R8u+	R maximal autocorrelation of lag 8 unweighted
R2m	R autocorrelation of lag 2 weighted by atomic masses
R3m+	R maximal autocorrelation of lag 3 weighted by atomic masses
R1v+	R maximal autocorrelation of lag 1 weighted by atomic van der Waals volumes
PSA	Fragment-based polar surface area
MlogP	Miroguchi octanol–water partition coefficient

Table 5 Correlation matrix of the descriptors selected by GA

	MATS1v	MATS2e	MATS7e	ISH	R8u+	R2m	R3m+	R1v+	PSA	MLOGP
MATS1v	1									
MATS2e	0.197	1								
MATS7e	0.133	0.006	1							
ISH	0.081	0.003	0.008	1						
R8u+	0.012	0.004	0.024	0.014	1					
R2m	0.048	0.006	0.145	0.006	0.023	1				
R3m+	0.097	0.006	0.032	0.006	0.056	0.174	1			
R1v+	0.195	0.059	0.122	0	0.147	0.063	0.139	1		
PSA	0.106	0.014	0.073	0.053	0.106	0.128	0.014	0.002	1	
MLOGP	0.329	0.418	0.093	0.009	0.002	0.011	0.024	0.129	0.181	1

van der Waals volume weighted term, influences positively the activity; and *MATS2e* and *MATS7e*, atomic Sanderson electronegativity weighted terms, reflect a complex relationship between electronegativity and inhibitory activity.

GETAWAY descriptors capture the 3D features determined by previous docking experiments. I_{SH} is the standarized information content on the leverage equality defined as:

$$I_{SH} = 1 - \frac{\sum_{g=1}^{G} N_g \times \log_2 N_g}{A_0 \times \log_2 A_0} \quad (4)$$

where A_0 represents the number of non-hydrogen atoms in the molecule, Ng is the number of non-hydrogen atoms with the same leverage value and G is the number of equivalence classes into which the atoms are partitioned according to the leverage equality [47]. If all the atoms have different leverage values, i.e., the molecule does not show any element of symmetry, $I_{SH} = 1$; otherwise, if all the atoms have equal leverage values (a perfectly symmetric theoretical case), $I_{SH} = 0$. In this sense, this index encodes information on the molecule entropy (thermodynamic entropy) and hence should be useful in modeling physicochemical properties related to entropy and symmetry. In the current model I_{SH} had a negative influence; this means that molecular symmetry has a positive influence in c-Met kinase inhibitory activity.

$R8u^+$, *R2 m*, $R3\ m^+$, and $R1v^+$ belong to R-GETAWAY descriptors which combine the information provided by the molecular influence matrix (MIM) with geometry matrix G. This junction leads to the influence/distance matrix R whose elements resemble the single terms in the sums of gravitational indices [47]. The autocorrelation R-GETAWAY descriptor ($R(l,p_k)$) at lag l weighted by property p_k is defined as:

$$R(l, p_k) = \sum_{i=1}^{A-1} \sum\nolimits_{j>1} \frac{\sqrt{h_{ii} \times h_{jj}}}{r_{ij}} \times w_i \times w_j \times \delta(l; d_{ij}) \quad (5)$$

where h_{ii} and h_{jj} are the leverages of the two considered atoms and r_{ij} is their geometric distance. Besides, d_{ij} is the topological distance between atoms i and j, while w_i and w_j are the property vectors of the atoms i and j. The function $\delta(l; d_{ij})$ is a Dirac-delta function defined in Eq. 3.

R maximal autocorrelation descriptors represent the maximal contribution to the autocorrelation at lag l weighted by property p_k. These descriptors take into account local aspects of the molecule, and are defined as:

$$R(l, p_k)^+ = \max_{ij} \left(\frac{\sqrt{h_{ii} \times h_{jj}}}{r_{ij}} \times w_i \times w_j \times \delta(l; d_{ij}) \right). \quad (6)$$

The information provided by R-GETAWAY descriptors in MLR model (Eq. 1) is weighted by atomic masses, having a negative influence, and van der Waals volumes, having a positive influence.

Ideally, QSAR models should be simple, transparent, and mechanistically comprehensible [53]. A QSAR model should be interpretable in terms of the parameters included. In general, the information contained in 2D autocorrelation and GETAWAY selected descriptors is related to distributions of atomic properties across the entire molecules. This information is not explicitly interpretable by medicinal chemists; however, 2D autocorrelation and GETAWAY descriptors are provided by molecular structures, and have been widely used for establishing structure–activity correlations [22, 23, 48, 54–56]. The elucidation of the molecular key features of the modeled data set from our proposed space is a difficult task. However, we can conclude that the structural information obtained here shows that a proper distribution of atomic van der Waals volumes, electronegativities, and atomic masses has a great influence

on the c-Met kinase inhibitory activities of the studied inhibitors.

According to MLR model fragment-based polar surface area, (*PSA*) and Miroguchi logP (MlogP) have positive influences in the c-Met inhibitory activity. The presence of *PSA* in the model reflects that a high number of polar groups are important to establish the interactions described above at sites ER, SR and CR. The presence of MlogP in the model reflects that hydrophobic groups filling site ASI should be present for a high activity. These variables include the effects previously described by docking experiments.

High values of Q^2 are a necessary, but not a sufficient, condition for a model to possess significant predictive power. For a stronger evaluation of model applicability for prediction on new chemicals, the external validation of the models is also recommended. We used CoMSIA-SEHD and MLR models to predict the inhibitory activities of an external test set of c-Met kinase inhibitors that was not used in the training process. The results of the predictions are shown in Table 1. The plots of the predicted versus experimental values are shown in Fig. 10a, b, and the residuals of the predictions are shown in Fig. 10c. This analysis reveals that both models have acceptable predictive capabilities. Both models fail in the prediction of compound **A37** (residuals above 1.5). The MLR procedure improves the CoMSIA method according to regression. The increment in the quality of the prediction of the external test set is showy, from $R^2_{test} = 0.496$ (R^2_{test} is R^2 of test set fitting) for the CoMSIA-SEHD model to $R^2_{test} = 0.725$ for the MLR one (these values did not consider the prediction of compound **A37**).

Further tests on external validation were carried out according to Golbraikh and Tropsha [57]. These authors formulate the following criteria for a QSAR model to have high predictive power: (i) high value of Q^2 ($Q^2 > 0.5$), (ii) correlation coefficient R_{test} between the predicted and observed activities of compounds from an external test set close to 1 ($R^2_{test} > 0.6$), (iii) at least one of the correlation

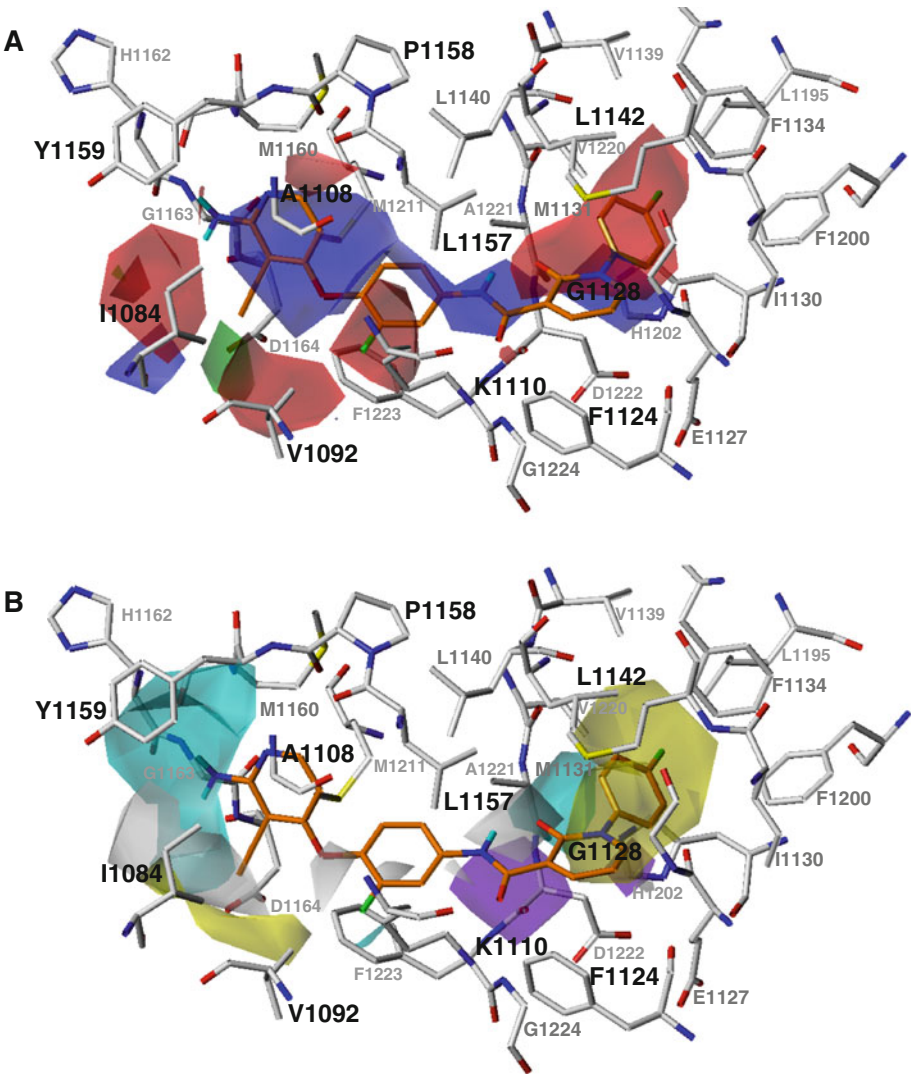


Fig. 11 CoMSIA contour maps for c-Met kinase deriving from model CoMSIA-SEHD. Compound **C15** is shown inside the fields. The amino acid residues located close to the binding pocket of c-Met kinase are represented for comparing their position with the position of isopleths. **a** steric and electrostatic fields: *green* isopleth indicates region where a bulky group favored the activity; *blue* isopleths indicate regions where positive charges favored the activity, or negative charges disfavored the activity; *red* isopleths indicate regions where negative charges favored the activity, or positive charges disfavored the activity. **b** Hydrophobic and h-bond donor fields: *yellow* isopleths represent areas where hydrophobic groups are favorable for activity, *white* isopleths represent regions where hydrophobic groups diminishes the activity; *cyan* contours represent regions where H-bond donor fields favor activity; purple contours indicate regions where H-bond donor fields are disfavored for activity

coefficients for regressions through the origin (predicted versus observed activities, or observed versus predicted activities), i.e. R_0^2 or R'^2_0 should be close to R^2_{test} ($[(R^2_{test}-R_0^2)/R^2_{test}]$ or $[(R^2_{test}-R'^2_0)/R^2_{test}] < 0.1$), (iv) at least one slope (k or k′) of regression lines through the origin should be close to 1. (It corresponds to R_0^2 or R'^2_0 that is closer to R^2_{test}); k or k′ should satisfy: $0.85 \leq k \leq 1.15$, or $0.85 \leq k' \leq 1.15$.

The results of this additional test, when the outlier compound **A37** was not considered in the analysis, are presented in Table 6. Model CoMSIA-SEHD satisfied some of the above mentioned criteria. However, the selected MLR model satisfied all the requirements. This test shows that MLR model is more reliable as a predictive tool.

According to the above mentioned results, CoMSIA-SEHD model is mechanistically interpretable and can orient in chemical synthesis of new candidates; but MLR model leads to a more precise prediction; in this sense, it is possible to combine both models for studying the structure–activity relationship of FPTA, FPPA, and AFPP derivatives as c-Met kinase inhibitors.

A main concern of investigators interested in chemical synthesis and biological evaluations is how they can use QSAR models for predicting new active compounds. A simple way to predict the activity of a new compound is by using linear Eq. 1. 2D and 3D topological descriptors, PSA, and MlogP of the new compound can be calculated by using Dragon software (http://michem.disat.unimib.it/chm/links/qsarlinks.htm#softwares). On the other hand, the use of CoMSIA-SEHD model as a predictive tool requires the use of Sybyl software (http://tripos.com/index.php?family=modules,SimplePage&page=SYBYL-X). The user should reproduce our model and predict new compounds by aligning them in the same grid.

Conclusions

In this paper, we report a theoretical study on drug design area for c-Met kinase inhibitors. We selected a data set of 103 compounds including FPTA, FPPA, and AFPP derivatives and studied the 3D positioning of them inside the active site of c-Met kinase by docking experiments. Our approach reproduced the previously reported position of compounds **A23**, **B1**, **B23**, **C2**, **C7**, and **C25** in the c-Met kinase's active site. The remaining compounds were oriented in a similar manner. The interactions established for all compounds within c-Met kinase active site were carefully described.

Additionally, predictive QSAR models were built by using CoMSIA and a MLR method including 2D autocorrelation descriptors, GETAWAY descriptors, and some physico-chemical properties calculated for the conformations obtained by docking experiments. 76 compounds were used for training these models. The best CoMSIA model included steric, electrostatic, hydrophobic, and HB donor fields; meanwhile, the best MLR model contains three Moran's 2D autocorrelation descriptors, five GETAWAY descriptors, *PSA*, and MlogP. Both models were validated by LOO cross-validation and then were used in the prediction of an external set which contains compounds that were not included during the determination of the models.

On the basis of the present results, this theoretical study could bring useful molecular information about the ligand specificity for improving interaction with c-Met kinase binding pocket. Docking experiments allowed the characterization of the specific interactions between the studied inhibitors and c-Met kinase's active site, while QSAR allowed two models able to predict new potent candidates. Moreover, the interpretation of the CoMSIA fields makes it possible to draw conclusions concerning the most appropriate features for the analogues.

Table 6 Tests on external validation for best CoMSIA and MLR models

	Model CoMSIA-SEHD without compound **A37**	Model MLR (Eq. 1) without compound **A37**
R^2_{test}	0.496	0.725
$(R^2_{test} - R_0^2)/R^2_{test}$	0.425	0.195
k	0.842	0.914
$(R^2_{test} - R'^2_0)/R^2_{test}$	0.020	0.006
k′	1.044	1.022

Acknowledgments J.C. thanks "Becas Universidad de Talca" for financial support through a doctoral fellowship. M.Q. and E.D. gratefully acknowledge the Institut de Recherche pour le Développement (UMR 152 IRD-UPS) for financial support. J.H.A.M. acknowledges the financial support through project FONDECYT No 11100177.

References

1. Liu X, Yao W, Newton RC, Scherle PA (2008) Targeting the c-MET signaling pathway for cancer therapy. Expert Opin Investig Drugs 17(7):997–1011
2. Peruzzi B, Bottaro DP (2006) Targeting the c-Met signaling pathway in cancer. Clin Cancer Res 12(12):3657–3660
3. Huh C-G, Factor VM, Sánchez A, Uchida K, Conner EA, Thorgeirsson SS (2004) Hepatocyte growth factor/c-met signaling pathway is required for efficient liver regeneration and repair. Proc Natl Acad Sci USA 101(13):4477–4482
4. Lesko E, Majka M (2008) The biological role of HGF-MET axis in tumor growth and development of metastasis. Front Biosci 13(13):1271–1280
5. Comoglio PM, Giordano S, Trusolino L (2008) Drug development of MET inhibitors: targeting oncogene addiction and expedience. Nat Rev Drug Discov 7(6):504–516

6. Porter J (2010) Small molecule c-Met kinase inhibitors: a review of recent patents. Expert Opin Ther Pat 20(2):159–177
7. Eder JP, Vande Woude GF, Boerner SA, LoRusso PM (2009) Novel therapeutic inhibitors of the c-Met signaling pathway in cancer. Clin Cancer Res 15(7):2207–2214
8. Schroeder GM, Chen X-T, Williams DK, Nirschl DS, Cai Z-W, Wei D, Tokarski JS, An Y, Sack J et al (2008) Identification of pyrrolo[2, 1-f][1, 2, 4]triazine-based inhibitors of Met kinase. Bioorg Med Chem Lett 18(6):1945–1951
9. Kim KS, Zhang L, Schmidt R, Cai Z-W, Wei D, Williams DK, Lombardo LJ, Trainor GL, Xie D et al (2008) Discovery of pyrrolopyridine–pyridone based inhibitors of met kinase: synthesis, X-ray crystallographic analysis, and biological activities. J Med Chem 51(17):5330–5341
10. Williams DK, Chen X-T, Tarby C, Kaltenbach R, Cai Z-W, Tokarski JS, An Y, Sack JS, Wautlet B et al (2010) Design, synthesis and structure-activity relationships of novel biarylamine-based Met kinase inhibitors. Bioorg Med Chem Lett 20(9): 2998–3002
11. Cai Z-W, Wei D, Schroeder GM, Cornelius LAM, Kim K, Chen X-T, Schmidt RJ, Williams DK, Tokarski JS et al (2008) Discovery of orally active pyrrolopyridine- and aminopyridine-based Met kinase inhibitors. Bioorg Med Chem Lett 18(11):3224–3229
12. Schroeder GM, An Y, Cai Z-W, Chen X-T, Clark C, Cornelius LAM, Dai J, Gullo-Brown J, Gupta A et al (2009) Discovery of N-(4-(2-Amino-3-chloropyridin-4-yloxy)-3-fluorophenyl)-4-ethoxy-1-(4-fluorophenyl)-2-oxo-1, 2-dihydropyridine-3-carboxamide (BMS-777607), a selective and orally efficacious inhibitor of the met kinase superfamily. J Med Chem 52(5):1251–1254
13. Albrecht BK, Harmange J-C, Bauer D, Berry L, Bode C, Boezio AA, Chen A, Choquette D, Dussault I et al (2008) Discovery and optimization of triazolopyridazines as potent and selective inhibitors of the c-Met Kinase†. J Med Chem 51(10):2879–2882
14. Boezio AA, Berry L, Albrecht BK, Bauer D, Bellon SF, Bode C, Chen A, Choquette D, Dussault I et al (2009) Discovery and optimization of potent and selective triazolopyridazine series of c-Met inhibitors. Bioorg Med Chem Lett 19(22):6307–6312
15. D'Angelo ND, Bellon SF, Booker SK, Cheng Y, Coxon A, Dominguez C, Fellows I, Hoffman D, Hungate R et al (2008) Design, synthesis, and biological evaluation of potent c-Met inhibitors. J Med Chem 51(18):5766–5779
16. Porter J, Lumb S, Lecomte F, Reuberson J, Foley A, Calmiano M, le Riche K, Edwards H, Delgado J et al (2009) Discovery of a novel series of quinoxalines as inhibitors of c-Met kinase. Bioorg Med Chem Lett 19(2):397–400
17. Porter J, Lumb S, Franklin RJ, Gascon-Simorte JM, Calmiano M, Riche KL, Lallemand B, Keyaerts J, Edwards H et al (2009) Discovery of 4-azaindoles as novel inhibitors of c-Met kinase. Bioorg Med Chem Lett 19(10):2780–2784
18. Alzate-Morales JH, Caballero J, Vergara-Jaque A, González-Nilo FD (2009) Insights into the structural basis of N2 and O6 substituted guanine derivatives as cyclin-dependent kinase 2 (CDK2) inhibitors: prediction of the binding modes and potency of the inhibitors by docking and ONIOM calculations. J Chem Inf Model 49(4):886–899
19. Alzate-Morales JH, Vergara-Jaque A, Caballero J (2010) Computational study on the interaction of N1 substituted pyrazole derivatives with B-Raf kinase: an unusual water wire hydrogen-bond network and novel interactions at the entrance of the active site. J Chem Inf Model 50(6):1101–1112
20. Larsen CA, Bisson WH, Dashwood RH (2009) Tea catechins inhibit hepatocyte growth factor receptor (MET Kinase) activity in human colon cancer cells: kinetic and molecular docking studies. J Med Chem 52(21):6543–6545
21. Fernandez M, Tundidor-Camba A, Caballero J (2005) Modeling of cyclin-dependent kinase inhibition by 1H-Pyrazolo[3, 4-d]Pyrimidine derivatives using artificial neural network ensembles. J Chem Inf Model 45(6):1884–1895
22. González M, Caballero J, Helguera A, Garriga M, González G, Fernández M (2006) 2D autocorrelation modelling of the inhibitory activity of cytokinin-derived cyclin-dependent kinase inhibitors. Bull Math Biol 68(4):735–751
23. Caballero J, Fernández M, Saavedra M, González-Nilo FD (2008) 2D Autocorrelation, CoMFA, and CoMSIA modeling of protein tyrosine kinases' inhibition by substituted pyrido[2, 3-d]pyrimidine derivatives. Bioorg Med Chem 16(2):810–821
24. Caballero J, Fernández M, González-Nilo FD (2008) Structural requirements of pyrido[2, 3-d]pyrimidin-7-one as CDK4/D inhibitors: 2D autocorrelation, CoMFA and CoMSIA analyses. Bioorg Med Chem 16(11):6103–6115
25. Gueto C, Ruiz JL, Torres JE, Méndez J, Vivas-Reyes R (2008) Three-dimensional quantitative structure-activity relationship studies on novel series of benzotriazine based compounds acting as Src inhibitors using CoMFA and CoMSIA. Bioorg Med Chem 16(5):2439–2447
26. Alzate-Morales J, Caballero J (2010) Computational study of the interactions between guanine derivatives and cyclin-dependent kinase 2 (CDK2) by CoMFA and QM/MM. J Chem Inf Model 50(1):110–122
27. Muthas D, Sabnis YA, Lundborg M, Karlén A (2008) Is it possible to increase hit rates in structure-based virtual screening by pharmacophore filtering? An investigation of the advantages and pitfalls of post-filtering. J Mol Graph Model 26(8):1237–1251
28. Xie H-Z, Li L-L, Ren J-X, Zou J, Yang L, Wei Y-Q, Yang S-Y (2009) Pharmacophore modeling study based on known spleen tyrosine kinase inhibitors together with virtual screening for identifying novel inhibitors. Bioorg Med Chem Lett 19(7):1944–1949
29. Uno M, Ban HS, Nabeyama W, Nakamura H (2008) de novo design and synthesis of N-benzylanilines as new candidates for VEGFR tyrosine kinase inhibitors. Org Biomol Chem 6(6): 979–981
30. Vieth M, Erickson J, Wang J, Webster Y, Mader M, Higgs R, Watson I (2009) Kinase inhibitor data modeling and de novo inhibitor design with fragment approaches. J Med Chem 52(20): 6456–6466
31. Alzate-Morales JH, Contreras R, Soriano A, Tuñon I, Silla E (2007) A computational study of the protein-ligand interactions in CDK2 inhibitors: using quantum mechanics/molecular mechanics interaction energy as a predictor of the biological activity. Biophys J 92(2):430–439
32. Alzate-Morales JH, Caballero J, Gonzalez-Nilo FD, Contreras R (2009) A computational ONIOM model for the description of the H-bond interactions between NU2058 analogues and CDK2 active site. Chem Phys Lett 479(1–3):149–155
33. Villacañas O, Pérez JJ, Rubio-Martínez J (2002) Structural analysis of the inhibition of Cdk4 and Cdk6 by p16INK4a through molecular dynamics simulations. J Biomol Struct Dyn 20:347–358
34. Asses Y, Leroux V, Tairi-Kellou S, Dono R, Maina F, Maigret B (2009) Analysis of c-Met kinase domain complexes: a new specific catalytic site receptor model for defining binding modes of ATP-competitive ligands. Chem Biol Drug Des 74(6):560–570
35. Klebe G, Abraham U, Mietzner T (1994) Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. J Med Chem 37(24):4130–4146
36. Wellenzohn B, Liedl KR, Rode BM, Zaheer-ul-haq (2003) Molecular docking studies of natural cholinesterase-inhibiting steroidal alkaloids from sarcococca saligna. J Med Chem 46(23): 5087–5090
37. Rojo LE, Alzate-Morales J, Saavedra IN, Davies P, Maccioni RB (2010) Selective interaction of lansoprazole and astemizole with

tau polymers: potential new clinical use in diagnosis of alzheimer's disease. J Alzheimer Dis 19(2):573–589
38. Hanessian S, Moitessier N, Therrien E (2001) A comparative docking study and the design of potentially selective MMP inhibitors. J Comput Aided Mol Des 15(10):873–881
39. Caballero J, Vergara-Jaque A, Fernández M, Coll D (2009) Docking and quantitative structure–activity relationship studies for sulfonyl hydrazides as inhibitors of cytosolic human branched-chain amino acid aminotransferase. Mol Divers 13(4): 493–500
40. Lagos CF, Caballero J, Gonzalez-Nilo FD, David Pessoa-Mahana C, Perez-Acle T (2008) Docking and quantitative structure-activity relationship studies for the bisphenylbenzimidazole family of non-nucleoside inhibitors of HIV-1 reverse transcriptase. Chem Biol Drug Des 72(5):360–369
41. Abagyan R, Totrov M, Kuznetsov D (1994) ICM–A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation. J Comput Chem 15(5):488–506
42. Molecular Editor, version 2.5, La Jolla, CA, Molsoft LLC, 2006
43. ICM, version 3.4-8, La Jolla, CA, Molsoft LLC, 2006
44. An J, Totrov M, Abagyan R (2005) Pocketome via comprehensive identification and classification of ligand binding envelopes. Mol Cell Proteomics 4(6):752–761
45. Caballero J, Fernández M (2008) Artificial neural networks from MATLAB in medicinal chemistry. Bayesian-regularized genetic neural networks (BRGNN): application to the prediction of the antagonistic activity against human platelet thrombin receptor (PAR-1). Curr Top Med Chem 8(18):1580–1605
46. DRAGON, version 3.0, Milano, Italy, Milano Chemometrics, 2003
47. Consonni V, Todeschini R, Pavan M (2002) Structure/response correlations and similarity/diversity analysis by GETAWAY descriptors. 1. Theory of the novel 3D molecular descriptors. J Chem Inf Comput Sci 42(3):682–692
48. Fernández M, Caballero J (2007) QSAR modeling of matrix metalloproteinase inhibition by N-hydroxy-[alpha]-phenylsulfonylacetamide derivatives. Bioorg Med Chem 15(18):6298–6310
49. Viswanadhan VN, Ghose AK, Revankar GR, Robins RK (1989) Atomic physicochemical parameters for three dimensional structure directed quantitative structure-activity relationships. 4. Additional parameters for hydrophobic and dispersive interactions and their application for an automated superposition of certain naturally occurring nucleoside antibiotics. J Chem Inf Comput Sci 29(3):163–172
50. Ertl P, Rohde B, Selzer P (2000) Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. J Med Chem 43(20):3714–3717
51. Miroguchi I, Hirono S, Liu Q, Nakagome I, Matsushita Y (1992) Simple method of calculating octanol/water partition coefficient. Chem Pharm Bull 40:127–130
52. de Oliveira DB, Gaudio AC (2000) BuildQSAR: a new computer program for QSAR analysis. Quant Struct Act Relat 19(6): 599–601
53. Cronin MTD, Schultz TW (2003) Pitfalls in QSAR. J Mol Struct 622(1–2):39–51
54. Fernandez M, Carreiras MC, Marco JL, Caballero J (2006) Modeling of acetylcholinesterase inhibition by tacrine analogues using Bayesian-regularized Genetic Neural Networks and ensemble averaging. J Enzyme Inhib Med Chem 21(6):647–661
55. Fatemi MH, Gharaghani S (2007) A novel QSAR model for prediction of apoptosis-inducing activity of 4-aryl-4-H-chromenes based on support vector machine. Bioorg Med Chem 15(24):7746–7754
56. González MP, Terán C, Teijeira M, González-Moa MJ (2005) GETAWAY descriptors to predicting A2A adenosine receptors agonists. Eur J Med Chem 40(11):1080–1086
57. Golbraikh A, Tropsha A (2002) Beware of q2!. J Mol Graph Model 20(4):269–276