J Mol Model (2011) 17:1537–1550
DOI 10.1007/s00894-010-0857-7

ORIGINAL PAPER

3,4,5-Trisubstituted-1,2,4-4*H*-triazoles as WT and Y188L mutant HIV-1 non-nucleoside reverse transcriptase inhibitors: docking-based CoMFA and CoMSIA analyses

Elena Cichero · Laura Buffa · Paola Fossa

Received: 30 July 2010 / Accepted: 14 September 2010 / Published online: 5 October 2010

Abstract 3,4,5-Trisubstituted-1,2,4-4*H*-triazoles (TTs) have recently been identified as a new class of potent non-nucleoside HIV-1 reverse transcriptase (RT) inhibitors. Two series of triazoles have been studied, one of which was also screened against the Y188L mutant. A computational strategy based on molecular docking studies followed by comparative molecular fields analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) has been used to elucidate the atomic details of the RT/TT interactions and to identify the most important features impacting the TT antiretroviral activity. Two 3D-QSAR CoMFA and CoMSIA models were derived, using the TT pEC_{50} values measured against wild-type (WT) HIV-1 (model A) and the Y188L mutant form (model B), respectively, as the dependent variable. The final model A CoMSIA (r_{ncv}^2=0.97, r_{cv}^2=0.89, SEE=0.314, and r_{pred}^2=0.82) and model B CoMSIA (r_{ncv}^2=0.91, r_{cv}^2=0.61, SEE=0.236, and r_{pred}^2=0.73) analyses were more predictive. The results allowed us to obtain useful information for the design of new compounds with improved potency towards WT HIV-1 or that are potentially active against the Y188L mutant.

Keywords HIV-1 · Reverse transcriptase · CoMFA · CoMSIA · Y188L HIV-1 mutant · Triazoles

Electronic supplementary material The online version of this article (doi:10.1007/s00894-010-0857-7) contains supplementary material, which is available to authorized users.

E. Cichero (✉) · L. Buffa · P. Fossa
Dipartimento di Scienze Farmaceutiche, Università degli Studi di Genova,
Viale Benedetto XV n.3.,
16132 Genova, Italy
e-mail: cichero@unige.it

Introduction

Reverse transcriptase (RT) is a key enzyme in the HIV replication cycle, and it is one of the main targets in the development of drugs to treat HIV infection and AIDS [1–5]. RT catalyzes the conversion of viral RNA into double-stranded DNA, which is then integrated into the host genome. Non-nucleoside RT inhibitors (NNRTIs) bind to an allosteric hydrophobic pocket (NNRTI binding site: NNIBS) that is located at about 10 Å from the polymerase active site and created upon inhibitor interaction, and lock the enzyme into an inactive form by affecting the geometry of the aspartyl residues at the active site of the polymerase [6]. In the past fifteen years, more than fifty structurally diverse NNRTIs have been described [6–12]. The fact that cross-resistance extends to the whole NNRTI class calls for the development of new agents which are capable of inhibiting clinically relevant NNRTI-resistant mutants [13, 14].

Triazole derivatives, and in particular 3,4,5-trisubstituted-1,2,4-4*H*-triazoles (TTs), have been recently described as a novel class of potent NNRTIs [15, 16]. In 2006, Girardet and co-workers discussed a series of TTs that bear an amide function at the linker group; this amide is positioned between the sulfur atom (at position 3 on the triazole) and the R1 substituent (Fig. 1) [15].

Recently, Kirschberg and co-workers described a series of triazoles (active against WT HIV-1 RT) which bear a methylene linker that is located between the sulfur atom (at position 3 on the triazole) and the R1 substituent (Fig. 1) [16].

The first series of compounds displayed submicromolar or nanomolar potency against WT, and (some of them) were also potent against the Y188L mutant. The second series was characterized by a loss of potency (micromolar log order) toward the WT.

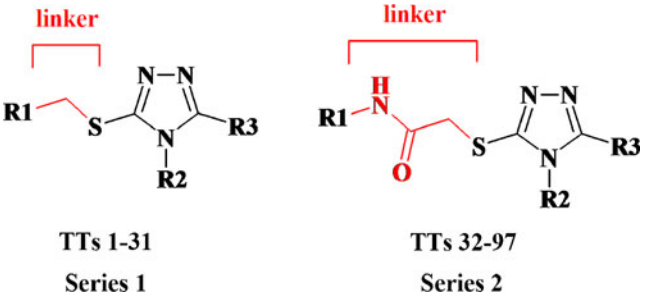


Fig. 1 Chemical structure of TTs **1–31** (series 1) and **32–97** (series 2)

In order to elucidate the most important RT/TT interactions and identify the features that impact most strongly on the TT antiretroviral activity, we performed a computational study of a dataset of 97 TTs selected from the two series of triazoles described above. Our aim was also to elaborate a quantitative structure–activity relationship (QSAR) model that would facilitate TT antiretroviral activity prediction prior to synthesis, and to obtain useful suggestions for the design of new TTs with improved potency against the Y188L mutant too.

Based on the crystal structure of RT in complex with 4-benzyl-3-[(2-chlorobenzyl)sulfanyl]-5-thien-2-yl-4*H*-1,2,4-triazole (TT **18**) [16], a flexible docking simulation was performed on compounds **1–97**. The most probable docking poses were selected and submitted to 3D-QSAR studies involving comparative molecular fields analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA). Two 3D-QSAR CoMFA and CoMSIA models were derived, using the TT pEC_{50} values measured against the wild-type (WT) HIV-1 (model A) and the Y188L mutant form (model B), respectively, as the dependent variable. These results allowed us to obtain useful information for the design of new compounds with improved potency towards WT HIV-1 or that are potentially active against the Y188L mutant.

Materials and methods

Data set

A dataset of 97 triazoles was selected from two sample data sets. TTs **1–31** (Table 1) [16] and TTs **32–97** (Table 2) [15] were screened against WT HIV-1 on MT-2 cells and HeLa-JC53 cells, respectively. Furthermore, some compounds belonging to series 2 (**39–47**, **53–68** and **75–96**) were also screened against the Y188L mutant.

Since the pEC_{50} values for **1–97** are comparable (Efavirenz was used as a reference compound for biological assays with both cell lines, and results within the same order of magnitude were obtained), all of the compounds were submitted to CoMFA and CoMSIA analyses (using the TT pEC_{50} values measured against the WT HIV-1, model A, as the dependent variable).

In addition, CoMFA and CoMSIA models were also derived using the pEC_{50} values for **39–47**, **53–68** and **75–96**, measured against the Y188L mutant form (model B), as the dependent variable.

The molecular structures of compounds **1**–**97** (Tables 1, 2) were built, parameterized (Gasteiger–Hückel method) and energy-minimized within MOE using the MMFF94 force field [17].

Docking-based ligand alignment

To locate the appropriate binding orientations and conformations of triazole derivatives within the NNIBS, a computational searching method was applied. Starting from a database including all 97 compounds, a docking procedure was performed. Thus, based on the three-dimensional structural coordinates of the RT/**18** complex (PDB entry 2RKI) [16] (**18** in the bioactive conformation), each inhibitor was docked into the NNIBS using the flexible docking module implemented in Surflex [18]. Surflex-Dock uses an empirically derived scoring function that is based on the binding affinities of protein–ligand complexes and on their X-ray structures.

Since for all compounds the best-docked geometries were in agreement with the crystallographic data of the RT/**18** complex (and thus already aligned), they were directly submitted to CoMFA [19] and CoMSIA [20] studies using the Sybyl 1.0 software [21].

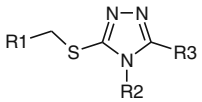
3D-QSAR analysis

Training set and test set. For both models (A and B), the training and the test set compounds were divided up manually according to a representative range of biological activities and structural variations. For QSAR analysis, EC_{50} values were transformed into pEC_{50} values and then used as response variables. Compound binding affinity covered 4 log orders of magnitude.

Triazole derivatives **1–97** *(TTs as WT NNRTIs, model A).* All of the compounds were placed into either a training set (**1**, **3**, **4**, **6**, **7**, **10**, **11**, **13**, **14**, **16**, **19**, **22**, **24–31**, **33**, **36**, **39**, **41**, **45**, **48–51**, **53**, **54**, **57–60**, **64–71**, **77**, **78**, **80**, **81**, **84**, **86–88**, **90**, **91**, **94–96**) for model generation, or a test set (**2**, **5**, **8**, **9**, **12**, **15**, **17**, **18**, **20**, **21**, **23**, **32**, **34**, **35**, **37**, **38**, **40**, **42–44**, **46**, **47**, **52**, **55**, **56**, **61–63**, **72–76**, **79**, **82**, **83**, **85**, **89**, **92**, **93**, **97**) for model validation, containing 56 and 41 compounds, respectively.

Triazole derivatives **39–47**, **53-68** *and* **75–96** *(TTs as Y188L mutant NNRTIs, model B).* Compounds **39–47**,

Table 1 Molecular structures of triazole derivatives **1–31**



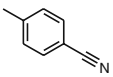
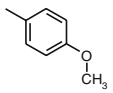
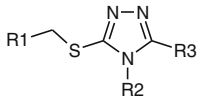
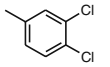
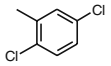
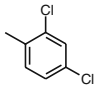
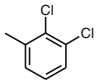
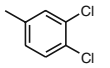
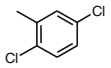
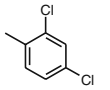
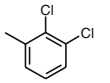
Comp.	R1	R2	R3	WT pEC_{50}
1				4.63
2				5.34
3				4.98
4				5.43
5				5.55
6				4.38
7				4.64
8				5.32
9				5.08
10				5.08
11				4.55
12				4.63
13				5.24
14				5.52
15				6.88
16				5.80
17				4.29
18				6.36
19				5.77
20				5.70
21				6.24
22				5.77

Table 1 (continued)



Comp.	R1	R2	R3	WT pEC_{50}
23				5.33
24				5.05
25				5.48
26				5.89
27				5.62
28				5.46
29				6.00
30				5.96
31				6.14

53–68 and **75-96** were divided into a training set (**39–44, 47, 53–59, 61-64, 67, 68, 75–80, 82–92, 95**) for model generation and into a test set (**45, 46, 60, 65, 66, 81, 93, 94, 96**) for model validation, containing 38 and 9 compounds respectively.

CoMFA and CoMSIA analyses

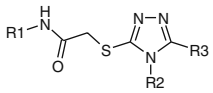
The CoMFA method is a widely used 3D-QSAR technique that relates the biological activities of a series of molecules with their steric and electrostatic fields, which are calculated by placing the aligned molecules, one by one, in a 3D cubic lattice with a grid spacing of 2 Å. The van der Waals potential and coulombic terms, which represent steric and electrostatic fields, respectively, were calculated using the standard Tripos force field method. The column-filtering threshold value was set to 2.0 kcal mol^{-1} in order to improve the signal-to-noise ratio. A methyl probe with a charge of +1 was used to calculate the CoMFA steric and electrostatic fields. An energy cut-off of 30 kcal mol^{-1} was applied to avoid infinite energy values inside the molecule.

The CoMSIA method calculates five descriptors: the steric, electrostatic, and hydrophobic parameters and the H-bond donor and H-bond acceptor properties. The similarity index descriptors were calculated using the same lattice box employed for the CoMFA calculations and an sp^3 carbon as probe atom with a +1 charge, +1 hydrophobicity, +1 H-bond acceptor and +1 H-bond donor properties.

Partial least-squares analysis and model validation

The partial least-squares (PLS) method, an extension of multiple regression analysis, was used to derive the 3D-QSAR models in which the CoMFA and CoMSIA descriptors were used as independent variables and pEC_{50} values were used as dependent variables. Leave-one-out (LOO) cross-validation method was used to check the predictivity of the derived model and to identify the optimal number of components (ONC) that yield the highest cross-validated r^2 (r^2_{cv}). In the LOO methodology, one molecule is omitted from the dataset and a model is derived that involves the other compounds. The activity of the omitted molecule is then predicted using this model.

Table 2 Molecular structures of triazole derivatives **32–97**



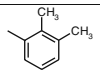
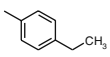
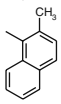
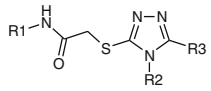
Comp.	R1	R2	R3	WT pEC_{50}	Y188L pEC_{50}
32				7.22	-
33				6.40	-
34				9.15	-
35				9.22	-
36				6.96	-
37				9.70	-
38				10.00	-
39				8.40	6.23
40				8.10	5.24
41				5.08	6.70
42				9.00	7.00
43				9.00	6.80
44				9.00	6.96
45				9.00	7.05
46				9.00	6.62
47				9.00	6.49
48				8.52	-
49				8.15	-
50				7.00	-
51				8.30	-

Table 2 (continued)



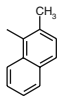
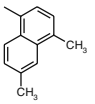
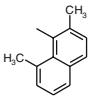
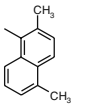
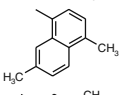
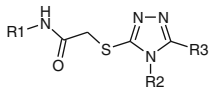
Comp.	R1	R2	R3	WT pEC_{50}	Y188L pEC_{50}
52				9.40	-
53				8.70	6.14
54				8.15	5.38
55				9.00	6.68
56				9.22	7.16
57				9.30	6.82
58				9.00	6.32
59				9.52	6.48
60				9.52	7.92
61				9.30	7.01
62				9.30	7.24
63				9.15	6.34
64				9.10	6.70
65				9.10	6.80
66				8.52	6.70
67				9.00	6.96
68				9.40	6.62
69				9.40	-

Table 2 (continued)



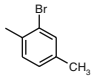
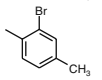
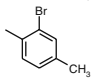
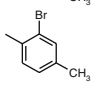
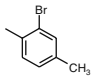
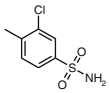
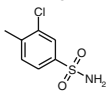
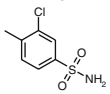
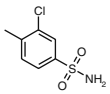
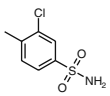
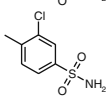
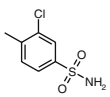
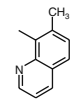
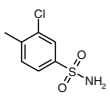
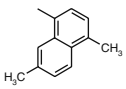
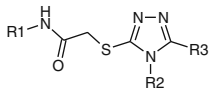
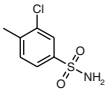
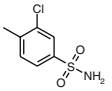
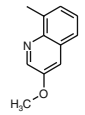
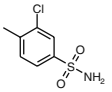
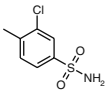
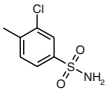
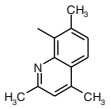
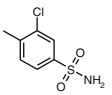
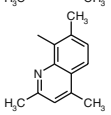
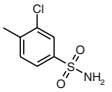
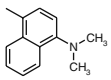
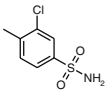
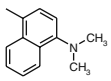
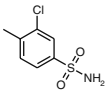
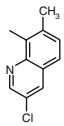
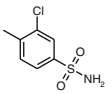
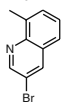
Comp.	R1	R2	R3	WT pEC_{50}	Y188L pEC_{50}
70				7.15	-
71				7.00	-
72				10.00	-
73				9.30	-
74				9.40	-
75				9.30	6.72
76				9.30	7.31
77				8.40	5.62
78				8.70	6.74
79				9.70	7.70
80				9.40	7.60
81				9.00	6.41
82				9.22	8.22
83				10.00	8.40
84				8.52	7.59
85				9.22	7.82
86				9.15	6.96

Table 2 (continued)



Comp.	R1	R2	R3	WT pEC_{50}	Y188L pEC_{50}
87				9.40	7.85v
88				9.10	6.43
89				9.00	7.17
90				8.40	7.39
91				9.15	7.96
92				9.00	8.40
93				9.30	7.06
94				8.70	8.15
95				8.70	7.00
96				9.10	6.72
97				9.30	-

The ONC obtained from the cross-validation methodology was used in the subsequent regression model. The final CoMFA and CoMSIA models were generated using non-cross-validated PLS analysis. To further assess the statistical confidence and robustness of the derived models, a 100-cycle bootstrap analysis was performed. This is a procedure in which *n* random selections are made from among the original set of *n* objects, and this process is repeated multiple times (in this case, 100 times in order to obtain good statistical information). During each run, some objects may not be included in the PLS analysis, whereas others may be included more then once. The mean correlation coefficient is represented as the bootstrap r^2 ($r_{boot}{}^2$). To validate the CoMFA- and CoMSIA-derived models, their predictive abilities for the test set of compounds (expressed as $r_{pred}{}^2$) were determined using the following equation:

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

SD is the sum of the squared deviations between the biological activities of the test set molecules and the mean activity of the training set compounds, while PRESS is the sum of the squared deviation between the observed and the predicted activities of the test set compounds.

All calculations were carried out on a PC running the Windows XP operating system.

Results and discussion

Docking-based ligand alignment

In this work, for the molecular modeling analysis we used the X-ray coordinates of the crystal structure of the TT **18**/RT complex. According to our calculations, all inhibitors display at least an H-bond between the K103 ε-amino group and the N1 or N2 atoms of the triazole ring, while several van der Waals and π–π interactions with two hydrophobic pockets (P_1 and P_2) are detected. As shown in Fig. 2, region P_1 (occupied by the R2 substituent of the TTs) includes the P95, Y181, Y188 and W229 residues, while P_2 (occupied by the R1 substituent of the TTs) consists of the K104, V106, F227, L234, H235, P236 and Y318 residues. All of the compounds orient the R3 group towards the NNIBS residues E138, V179, I180.

In addition, compounds **75–96** (the most active of the series, pEC_{50}=8.40–10.00) display H-bond interactions between the R1 sulfonamide moiety and the NNIBS residues K104 (the backbone carbonyl oxygen atom) and V106 (the backbone nitrogen portion), while the amide linker (located between the R1 group and the sulfur atom at position 3 on the triazole) is involved in a double H-bond with H235 and Y318. On the contrary, compounds **1-31** do not display additional H-bond interactions within the P2 pocket, showing a loss of potency (pEC_{50}=4.29–6.88) toward the WT.

In order to validate the docking procedure, we compared the most probable binding pose selected for each compound with the X-ray structure of **18** in complex with RT, and obtained ligand conformations that were perfectly superposed on the available crystallographic data. Thus, they were aligned together inside the NNIBS for the following 3D-QSAR analysis.

CoMFA and CoMSIA analyses

Model A

CoMFA and CoMSIA studies were developed using, respectively, CoMFA steric and electrostatic fields and CoMSIA steric, electrostatic, hydrophobic, H-bond donor and H-bond acceptor properties as independent variables, and the ligand pEC_{50} as a dependent variable.

The final CoMFA model was generated via non-cross-validated PLS analysis with the optimum number of components (ONC)=5 to give a non-cross-validated r^2 (r_{ncv}^2)=0.984, a test set r^2 (r_{pred}^2)=0.80, a standard error of the estimate (SEE)=0.229, a steric contribution=0.700, and an electrostatic contribution=0.300. The reliability of the

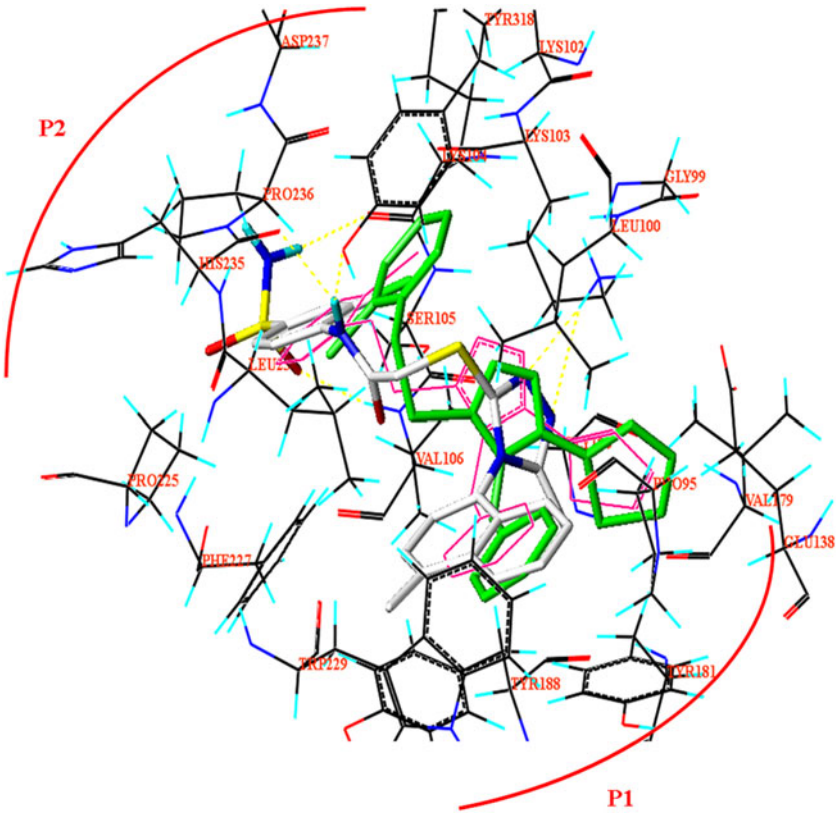


Fig. 2 Analysis of docking into the HIV-1 reverse transcriptase X-ray structure (2RKI). The X-ray pose of compound **18** is shown as *dark pink lines*. The docking poses of **18** (shown in *green*) and **79** (C in *white*, N in *blue*, O in *red*, S in *yellow*) are shown as *sticks*. The P1 and P2 pockets of RT and the most important residues are labeled. Hydrogen bonds are colored in *yellow*

model thus generated was supported by bootstrapping results.

A CoMSIA model consisting of steric, electrostatic, hydrophobic, H-bond donor and H-bond acceptor fields with r_{ncv}^2=0.970, r_{pred}^2=0.82, SEE=0.314, steric contribution=0.178, electrostatic contribution=0.144, hydrophobic contribution=0.257, H-bond acceptor contribution=0.230, and H-bond donor contribution=0.191 was derived. All statistical parameters that support the CoMFA and CoMSIA models are reported in Online Resource 1 of the Electronic supplementary material.

Experimental and predicted binding affinity values for the training set and test set are reported in Online Resource 2 of the Electronic supplementary material.

As shown in Fig. 3a (for simplicity, only the structures of compounds **15** and **83**, which exhibit high WT pEC_{50} values, are depicted as representatives of the two series), for compounds **1–31** (series 1), the steric contour map predicts favorable interaction polyhedra (green areas) near one of the two *ortho* positions of the R1 phenyl ring, while unfavorable areas (reported in yellow) are depicted around the two R1 *meta* positions. Yellow polyhedra also surround the flexible R2 benzyl moiety. Accordingly, compounds **1–31** (pEC_{50}=4.55–6.88, bearing a benzyl moiety at the R2 substituent) show lower pEC_{50} values than compounds **32–97** (pEC_{50}=5.64–10.00, bearing no flexible R2 substituents).

As far as compounds **32–97** (series 2) are concerned, favorable interaction polyhedra (green) are located in the proximity of the 1-naphthyl positions 4 and 8 at the R2 substituent (or around positions 2 and 4 for compounds bearing a phenyl ring instead of a 1-naphthyl one), and around the *para* and one of the two *ortho* positions of the R1 phenyl moiety. The reliability of the steric map calculations is verified by the higher pEC_{50} values of the compounds bearing a 4-substituted-1-naphthyl moiety (or a 4-substituted phenyl ring) at the R2 substituent in comparison with those of TTs bearing the 1-naphthyl or phenyl moiety (unsubstituted) in R2. These results are also supported by the high pEC_{50} values of the triazoles bearing a sulfonyl moiety at the R1 phenyl ring *para* position (instead of a methyl one).

On the contrary, unfavorable polyhedra (yellow) are depicted around the linker amide group oxygen atom and in the proximity of the 1-naphthyl positions 6 and 7 at the R2 substituent (underlining the beneficial introduction of the phenyl moiety, instead of the 1-naphthyl one).

For compounds **1-31** (series 1), the electrostatic field contour map of the CoMFA analysis plotted in Fig. 3b (for simplicity, only the structures of compounds **15** and **83**, which exhibit high WT pEC_{50} values, are depicted as representatives of the two series) predicts favorable electropositive substituents (blue polyhedra) around the *para* position of the R1 (hetero)aryl moiety and in the proximity of the 3-thienyl ring sulfur atom (R3 substituent) of compound **15**. These results are in agreement with the slightly higher pEC_{50} value of **14** (R1=4-methoxyphenyl; pEC_{50}=5.52) compared with that of **12** (R1=4-cyanophenyl; pEC_{50}=4.63), and are also verified by the difference in pEC_{50} among compounds bearing a 2-thienyl moiety at the triazole R3 substituent (instead of a 3-thienyl one).

Regarding TTs **32–97** (series 2), less positive moieties are predicted to be favored (red areas) in the proximity of the *para* position of the R1 phenyl ring (sulfonyl oxygen atoms in **83**), near the triazole N1 and N2 nitrogen atoms, and around the linker amide group. Accordingly, compounds **32–97** (pEC_{50}=5.64–10.00, bearing an amide function as a linker between the sulfur atom at the triazole position 3 and the R1 phenyl moiety) show higher pEC_{50}

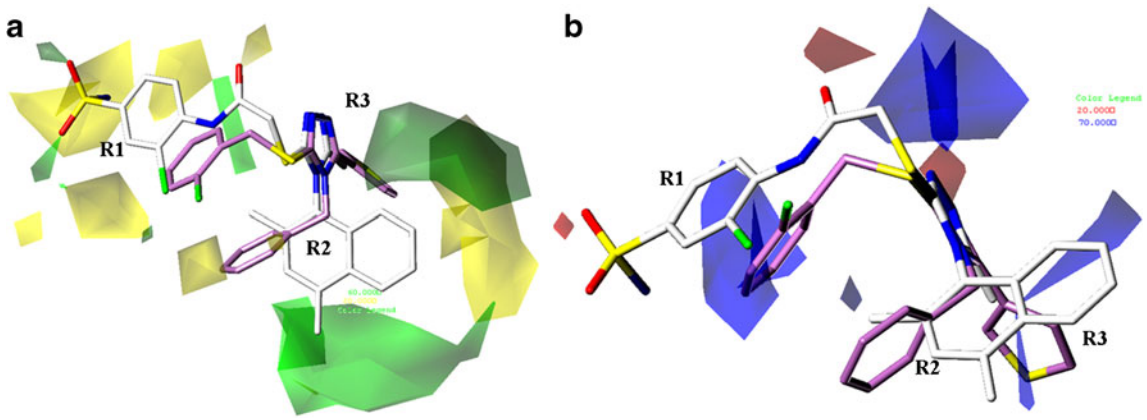


Fig. 3 Contour maps of CoMFA steric regions (**a**) (*green* are favored; *yellow* are disfavored) and CoMFA electrostatic areas (**b**) for model A are displayed around compounds **15** (as the representative of series 1 TTs; C is *pink*) and **83** (as the representative of series 2 TTs; C is *white*). *Blue regions* are favorable for more positively charged groups; *red regions* are favorable for less positively charged groups. Ligands are depicted in stick mode and are colored by atom type

values than compounds **1–31** [pEC_{50}=4.55 – 6.88, bearing a methylene linker between the sulfur atom at the triazole position 3 and the R1 (hetero)aryl moiety].

In addition, more electropositive substituents are predicted to be beneficial (blue area) around the linker methylene, on the R1 phenyl ring plane, and in the proximity of position 2 of the 1-naphthyl ring (R2 substituent). The presence of electropositive groups is also allowed near the 1-naphthyl position 8. These results are supported by the high pEC_{50} values of triazoles bearing a sulfonyl moiety at the R1 phenyl ring's *para* position (instead of a methyl one)], or a 1-naphthyl group (instead of a 8-quinoline moiety) at the triazole R2 substituent.

The CoMSIA steric and electrostatic regions are in agreement with the CoMFA steric and electrostatic areas.

The calculated CoMSIA hydrophobic contours (in Fig. 4, for simplicity, only the structures of compounds **15** and **83**, which exhibit high WT pEC_{50} values, are depicted as representatives of the two series) display favorable hydrophobic substituents (yellow polyhedra) in the proximity of the methylene of the R2 benzyl group (compounds **1–31**, series 1) and the carbon atom of the R2 1-naphthyl group (compounds **32–97**, series 2). Yellow polyhedra also surround the *meta* positions of the R1 (hetero)aryl moiety (compounds **1–31**). On the contrary, hydrophobic moieties are predicted to be disfavored (white areas) in the proximity of the 3-thienyl ring sulfur atom in compound **15** (R3 substituent in compounds **1–31**) and near the *para* position of the R1 phenyl ring, near the R3 substituent, and around the linker amide group in compounds **32–97**.

The CoMSIA H-bond acceptor map (see Online Resource 3 of the Electronic supplementary material) illustrates that H-bond acceptor groups are predicted to be beneficial (magenta areas) near the triazole N1 and N2 nitrogen atoms (for both of the series), in the proximity of the *para* position in the R1 phenyl ring of compounds **32-97**, and near the linker amide group oxygen atoms. The presence of an H-bond acceptor seems to be disfavored near the *meta* positions of the R1 (hetero)aryl moiety (compounds **1–31**) and around the nitrogen atom of the R1 sulfonylamide moiety (compounds **75–96**). Triazoles bearing a sulfonyl moiety at the *para* position of the R1 phenyl ring have higher pEC_{50} values than those with a methyl moiety at this position, and there is a difference in pEC_{50} values between compounds **32–97** (pEC_{50}=5.64–10.00, bearing an amide group as a linker between the sulfur atom at position 3 on the triazole and the R1 phenyl moiety) and compounds **1–31** [pEC_{50}=4.55 – 6.88, bearing a methylene linker between the sulfur atom at position 3 of the triazole and the R1 (hetero)aryl moiety].

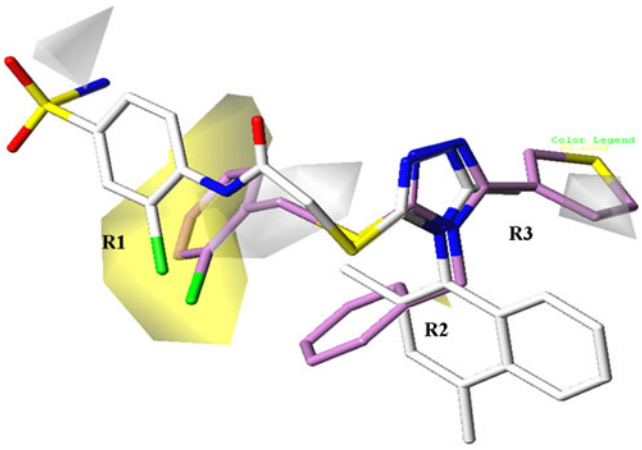


Fig. 4 Contour maps of CoMSIA hydrophobic regions (*yellow* is favored; *white* is disfavored) for model A are depicted around compounds **15** (as the representative of series 1 TTs, C is *pink*) and **83** (as the representative of series 2 TTs, C is *white*), as shown in stick mode and colored by atom type

Accordingly, H-bond donor functions (see Online Resource 3 of the Electronic supplementary material) would be beneficial (cyan areas) and detrimental (purple polyhedra) around the nitrogen and oxygen atoms of the R1 sulfonylamide moiety of compounds **75–96**. Purple polyhedra are also located in the proximity of the triazole N1 and N2 nitrogen atoms (for both series) and near the linker amide function (for compounds **32–97**).

In order to verify the reliability of the 3D-QSAR analyses, CoMFA and CoMSIA maps for model A were compared with the results of the docking analysis. Briefly, the CoMFA steric model is found to match with the NNIBS 3D topology, suggesting a bulky substitution in the proximity of the P_1 pocket occupied by the R2 substituent (involved in π–π stacking interactions with Y181, Y188 and W229). The CoMFA electrostatic map points out the beneficial presence of small electropositive substituents at the R3 substituent, which establish hydrophobic contacts with V179 and I180.

The CoMSIA H-bond acceptor and H-bond donor maps confirm the importance of the formation of hydrogen-bond interactions to stabilize the RT/TT complex. Accordingly, compounds **75–96** (the most active TTs of the two series, pEC_{50} range=8.40–10.00) bear the following three H-bond acceptor centers: the *para* position of the R1 phenyl ring (H bonds with H235 and Y318), the linker portion (H bonds with K104 and V106), and the triazole ring (K103). On the contrary, compounds **1–31** show lower pEC_{50} values (pEC_{50} range=4.29–6.88) because of the absence of H-bond acceptor functions at the TT R1 and linker groups.

Model B

CoMFA and CoMSIA studies were developed using, respectively, CoMFA steric and electrostatic fields and CoMSIA steric, electrostatic, hydrophobic, H-bond donor and H-bond acceptor properties as independent variables, and the ligand pEC_{50} as a dependent variable.

The final CoMFA model was generated via non-cross-validated PLS analysis with the optimum number of components (ONC)=6 to give a non-cross-validated r^2 (r_{ncv}^2)= 0.967, a test set r^2 (r_{pred}^2)=0.64, a standard error of the estimate (SEE)=0.149, a steric contribution=0.684, and an electrostatic contribution=0.316. The reliability of the model thus generated was supported by the bootstrapping results.

A CoMSIA model consisting of steric, electrostatic, hydrophobic, H-bond donor and H-bond acceptor fields with an r_{ncv}^2=0.913, an r_{pred}^2=0.73, a SEE=0.236, a steric contribution=0.172, an electrostatic contribution=0.153, a hydrophobic contribution=0.231, an H-bond acceptor contribution=0.164, and an H-bond donor contribution=0.281 was derived. All of the statistical parameters supporting the CoMFA and CoMSIA models are reported in Online Resource 4 of the Electronic supplementary material.

Experimental and predicted binding affinities for the training set and the test set are reported in Online Resource 5 of the Electronic supplementary material. The distributions of experimental and predicted pEC_{50} values for the training set and the test set compounds according to the CoMFA and CoMSIA analyses for model B are reported in Fig. 5

As shown in Fig. 6a (for simplicity, only the structure of compound **92** is depicted as a representative), the steric contour map predicts favorable interaction polyhedra (green) around positions 2 and 5 (or positions 4 and 7) of the 8-quinoline (or 1-naphthyl) moiety of the R2 substituent. The introduction of a small group at positions 4 and 6 of the 8-quinoline moiety is also allowed (or around positions 3 and 5 for TTs bearing a 1-naphthyl ring in R2). The reliability of the steric map calculations is verified by the higher pEC_{50} values of compounds bearing a 4-substituted-1-naphthyl moiety (or a 4-substituted phenyl ring) at the R2 substituent in comparison with those of TTs bearing the 1-naphthyl or phenyl moiety (unsubstituted) in R2.

Unfavorable polyhedra (yellow) surround the oxygen atom of the linker carbonyl group, one of the *ortho* positions, and both of the *meta* positions of the *para*-substituted phenyl ring (R1 substituent). Yellow areas are also located around the trifluoromethyl substituent (R3 substituent) and around positions 1 and 3 (or positions 6 and 8) of the 8-quinoline (or 1-naphthyl) moiety.

Based on the electrostatic field contour map from the CoMFA analysis plotted in Fig. 6b (for simplicity, only the

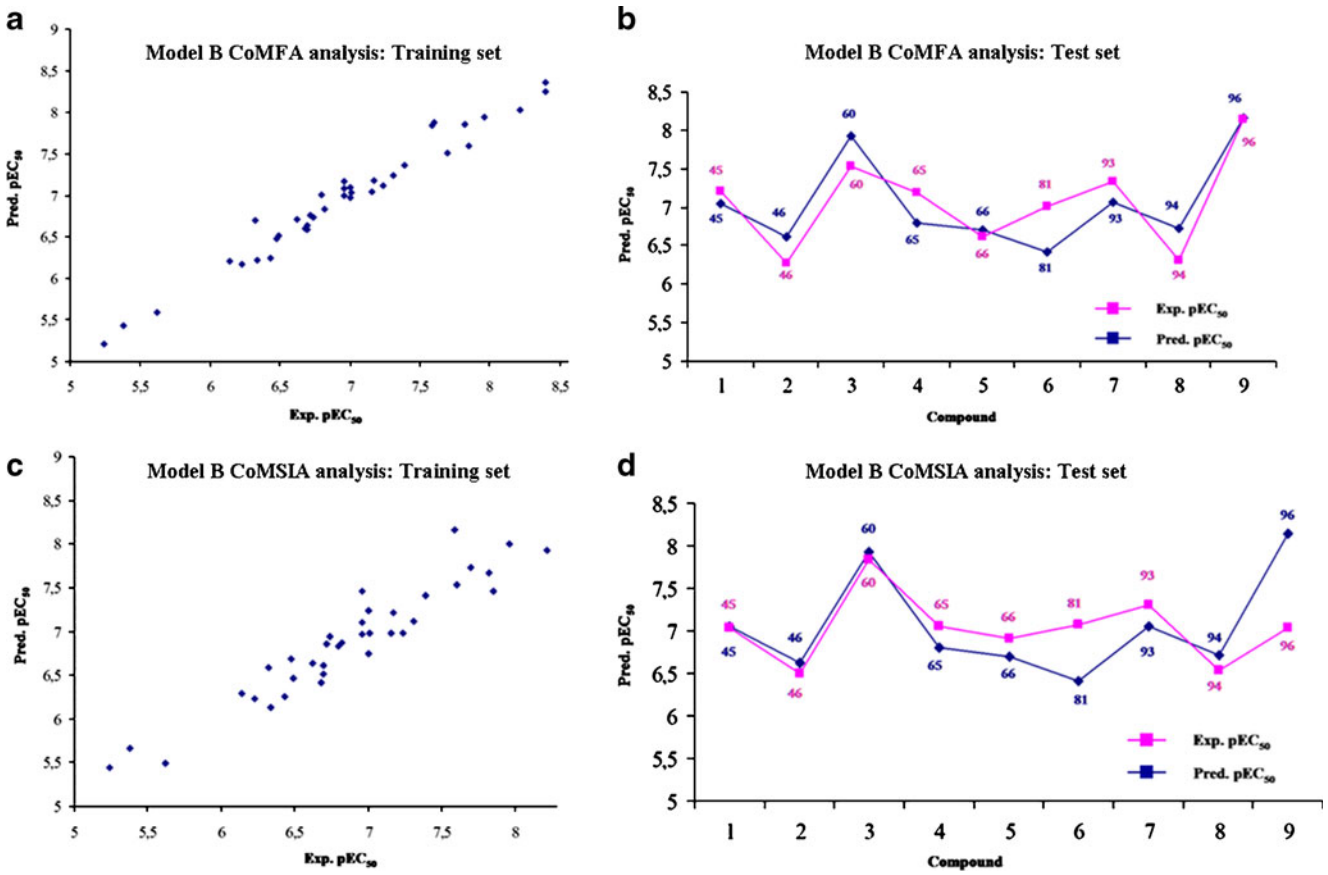


Fig. 5 Distributions of experimental and predicted pEC_{50} values for the training set compounds according to CoMFA analysis for model B (**a**), the test set compounds according to CoMFA analysis for model B (**b**), the training set compounds according to CoMSIA analysis for model B (**c**), and the test set compounds according to CoMSIA analysis for model B (**d**)

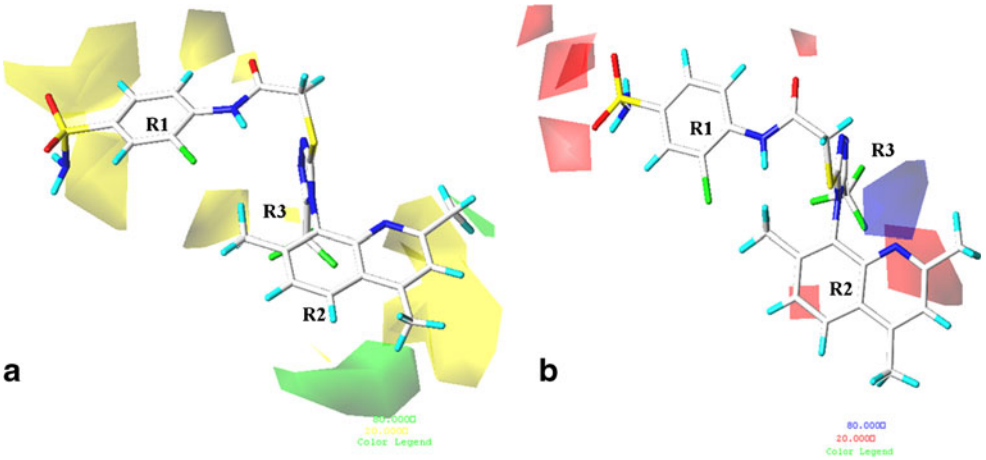


Fig. 6 Contour maps of CoMFA steric regions for model B (**a**) (*green* is favored; *yellow* is disfavored) and CoMFA electrostatic areas for model B (**b**) are shown for compound **92** (depicted as sticks, colored by atom type). *Blue regions* are favorable for more positively charged groups; *red regions* are favorable for less positively charged groups

structure of compound **92** is depicted as a representative), less positive moieties would be favored (red polyhedra) in the proximity of amide and sulfonyl oxygen atoms (see linker and R1 substituent for **32–97**), near position 6 (or position 3) of the 8-quinolinyl (or 1-naphthyl) at the R2 substituent, and around the trifluoromethyl group (R3 substituent). On the other hand, more electropositive substituents are predicted to be beneficial (blue area) around position 1 (or position 8) of the 8-quinolinyl (or 1-naphthyl) at the R2 substituent.

These results are supported by the high pEC_{50} values of triazoles bearing a sulfonyl moiety at the *para*-position of the phenyl ring of R1 (instead of a methyl one), a trifluoromethyl group at the R3 substituent (instead of a methyl one), or a 1-naphthyl ring (instead of a 8-quinoline moiety) at the R2 substituent.

The CoMSIA steric and electrostatic regions are in agreement with the CoMFA steric and electrostatic areas.

As shown in Fig. 7 (for simplicity, only the structure of compound **92** is depicted as a representative), the introduction of small hydrophobic substituents (yellow polyhedra) in the proximity of positions 1 and 5 (or positions 4 and 8) of the 8-quinolinyl (or 1-naphthyl) moiety at the R2 substituent, around the trifluoromethyl substituent (R3 substituent), and around one of the *ortho* positions on the phenyl ring of R1 is allowed. The reliability of the hydrophobic map calculation is verified by the higher pEC_{50} values of TTs bearing a 4-alkyl-1-naphthyl moiety at the R2 substituent in comparison with TTs bearing a 4-methoxy-1-naphthyl moiety in R2. These results are also supported by the high pEC_{50} values of TTs bearing a 1-naphthyl group (instead of a 8-quinoline moiety) at the R2 substituent, or a trifluoromethyl group (instead of a methyl one) at the R3 substituent.

The model B CoMSIA H-bond acceptor map (see Online Resource 5 of the Electronic supplementary material) shows that the introduction of groups bearing H-bond acceptor functions at the *para* position on the R1 phenyl ring would be beneficial (magenta polyhedra). On the contrary, H-bond acceptor functions are predicted to be disfavored (green areas) in the proximity of position 1 on the 8-quinolinyl (or position 8 on the 1-naphthyl) at the R2 substituent. Accordingly, H-bond donors are predicted to be favorable (cyan region) and unfavorable (purple region) near the nitrogen atom and oxygen atoms of the sulfonylamide on R1, respectively (see Online Resource 5 of the Electronic supplementary material). In addition, purple polyhedra are also located around position 7 (or position 2) of the 8-quinolinyl (or 1-naphthyl) on R2. These results are supported by evidence of high affinities of compounds bearing a sulfonyl moiety at the *para* position on the phenyl ring of R1 (instead of a methyl one).

Comparing these results with those obtained from model A, we have identified the following key features that enhance TT activity towards the Y188L mutant (instead of against the WT): (i) the introduction of a small substituent at position 3 of the 1-naphthyl is allowed (compare the

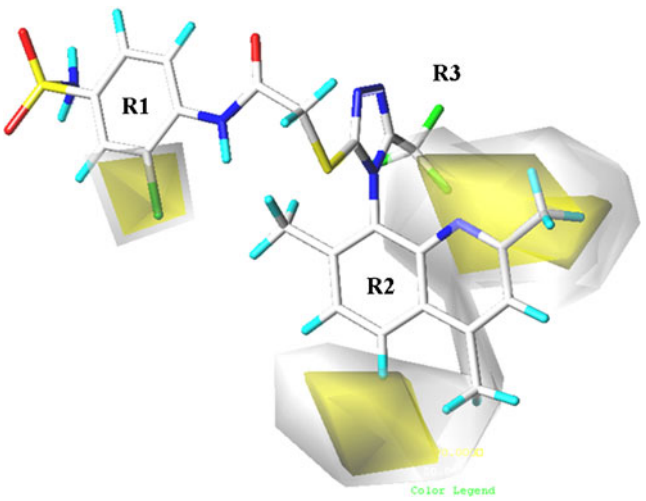


Fig. 7 Contour map of CoMSIA hydrophobic regions for model B (*yellow* is favored; *white* is disfavored) is shown for compound **92** (depicted in stick mode and colored by atom type)

CoMFA steric contour map around R2 reported in Fig.7 with Fig. 3a), (ii) electronegative groups around position 3 of the 1-naphthyl (compare the CoMFA electrostatic contour map around R2 reported in Fig.7b with Fig. 3b), and (iii) a hydrophobic group in the proximity of R3 (compare the CoMSIA hydrophobic contour map reported in Fig. 4 with Fig. 7).

Based on these results, the introduction of a 3-quinolinyl moiety at substituent R2 and an alkyl (isopropyl, cyclopropyl or *t*-butyl) group at substituent R3 seem to be particularly favorable routes to increasing TT potency against the Y188L mutant. In addition, a 3,5-dimethyl-phenyl or a 2-naphthyl moiety could prove effective as an R2 substituent.

Conclusions

The computational studies presented here analyzed the main interactions responsible for TT activity and gave useful suggestions for the synthesis of new analogs with improved potency against WT HIV-1 and also against the Y188L mutant.

The models elaborated here will be exploited to design new TTs and predict their activities prior to synthesis.

Acknowledgments This work was financially supported by the University of Genoa. Dr. B. Domenichini is gratefully acknowledged. E.C. was financially supported by a post-doc fellowship from Area Chimica, University of Genoa.

References

1. Jonckheere H, Anne J, De Clercq E (2000) The HIV-1 reverse transcription (RT) process as target for RT inhibitors. Med Res Rev 20:129–154
2. De Clercq E (2001) New developments in anti-HIV chemotherapy. Farmaco 56:3–12
3. De Clercq E (2005) Emerging anti-HIV drugs. Expert Opin Emerg Drugs 10:241–273
4. De Clercq E (2005) New approaches toward anti-HIV chemotherapy. E J Med Chem 48:1297–1313
5. Barbaro G, Scozzafava A, Mastrolorenzo A, Supuran CT (2005) Highly active antiretroviral therapy: current state of the art, new agents and their pharmacological interactions useful for improving therapeutic outcome. Curr Pharm Des 11:1805–1843
6. Balzarini J (2004) Current status of the non-nucleoside reverse transcriptase inhibitors of human immunodeficiency virus type 1. Curr Top Med Chem 4:921–944
7. De Clercq E (1998) The role of non-nucleoside reverse transcriptase inhibitors (NNRTIs) in the therapy of HIV-1 infection. Antiviral Res 38:153–179
8. Pedersen OS, Pedersen EB (1999) Non-nucleoside reverse transcriptase inhibitors, the NNRTI boom. Antivir Chem Chemother 10:285–314
9. Pedersen OS, Pedersen EB (2000) The flourishing syntheses of non-nucleoside reverse transcriptase inhibitors. Synthesis 4:479–495
10. Campiani G, Ramunno A, Maga G, Nacci V, Fattorusso C, Catalanotti B, Morelli E, Novellino E (2002) Non-nucleoside HIV-1 reverse transcriptase (RT) inhibitors:past, present, and future perspectives. Curr Pharm Des 8:615–657
11. De Clercq E (2004) Non-nucleoside reverse transcriptase inhibitors (NNRTIs): past, present, and future. Chem Biodiversity 1:44–64
12. Pauwels R (2004) New non-nucleoside reverse transcriptase inhibitors (NNRTIs) in development for the treatment of HIV infections. Curr Opin Pharmacol 4:437–446
13. Leigh Brown AJ, Frost SD, Mathews WC, Dawson K, Hellmann NS, Daar ES, Richman DD, Little SJ (2003) Transmission fitness of drug-resistant human immunodeficiency virus and the prevalence of resistance in the antiretroviral-treated population. J Infect Dis 187:683–686
14. Richman DD, Morton SC, Wrin T, Hellmann N, Berry S, Shapiro MF, Bozzette SA (2004) The prevalence of antiretroviral drug resistance in the United States. AIDS 18:1393–1401
15. De La Rosa M, Kim HW, Gunic E, Jenket C, Boyle U, Koh YH, Korboukh I, Allan M, Zhang W, Chen H, Xu W, Nilar S, Yao N, Hamatake R, Lang SA, Hong Z, Zhang Z, Girardet JL (2006) Tri-substituted triazoles as potent non-nucleoside inhibitors of the HIV-1 reverse transcriptase. Bioorg Med Chem Lett 16:4444–4449
16. Kirschberg TA, Balakrishnan M, Huang W, Hluhanich R, Kutty N, Liclican AC, McColl DJ, Squires NH, Lansdon EB (2007) Triazole derivatives as non-nucleoside inhibitors of HIV-1 reverse transcriptase—structure–activity relationships and crystallographic analysis. Bioorg Med Chem Lett 18:1131–1134
17. Chemical Computing Group Inc. (2010) MOE. Chemical Computing Group Inc., Montreal (see http://www.chemcomp.com)
18. Jain AN (1996) Scoring noncovalent protein-ligand interactions: a continuous differentiable function tuned to compute binding affinities. J Comput Aided-Mol Des 10:427–440
19. Cramer RD III, Patterson DE, Bunce JD (1989) Recent advances in comparative molecular field analysis (CoMFA). Prog Clin Biol Res 291:161–165
20. 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 37:4130–4146
21. Tripos Inc. (2010) Sybyl. Tripos Inc., St. Louis (see tripos.com/sybyl)