ORIGINAL PAPER

Towards discovering dual functional inhibitors against both wild type and K103N mutant HIV-1 reverse transcriptases: molecular docking and QSAR studies on 4,1-benzoxazepinone analogues

Zhenshan Zhang · Mingyue Zheng · Li Du · Jianhua Shen · Xiaomin Luo · Weiliang Zhu · Hualiang Jiang

Abstract To find useful information for discovering dual functional inhibitors against both wild type (WT) and K103N mutant reverse transcriptases (RTs) of HIV-1, molecular docking and 3D-QSAR approaches were applied to a set of twenty-five 4,1-benzoxazepinone analogues of efavirenz (SUSTIVA®), some of them are active against the two RTs. 3D-QSAR models were constructed, based on their binding conformations determined by molecular docking, with r^2_{cv} values ranging from 0.656 to 0.834 for CoMFA and CoMSIA, respectively. The models were then validated to be highly predictive and extrapolative by inhibitors in two test sets with different molecular skeletons. Furthermore, CoMFA models were found to be well matched with the binding sites of both WT and K103N RTs. Finally, a reasonable pharmacophore model of 4,1-benzoxazepinones were established. The application of the model not only successfully differentiated the experimentally determined inhibitors from non-inhibitors, but also discovered two potent inhibitors from the compound database SPECS. On the basis of both the 3D-QSAR and pharmacophore models, new clues for discovering and designing potent dual functional drug leads against HIV-1 were proposed: (i) adopting positively charged aliphatic group at the *cis*-substituent of C3; (ii) reducing the electronic density at the position of O4; (iii) positioning a small branched aliphatic group at position of C5; (iv) using the negatively charged bulky substituents at position of C7.

Introduction

HIV-1 reverse transcriptase (RT), a virally encoded human immunodeficiency virus (HIV) reverse transcriptase, catalyzes the replication of single-stranded viral RNA to yield double-stranded DNA before the viral genome is integrated into the DNA of host [1, 2]. As a milestone in treating AIDS [3, 4], RT has been the target of several antiviral therapeutic agents, including non-nucleoside RT inhibitors (NNRTIs) [5–7]. However, NNRTIs are embarrassed for rapidly inducing drug resistance, resulting from the mutations at the amino acid residues that form the NNRTI-binding pocket of HIV-RT [8–11]. Among the mutations, K103N (from lysine 103 to asparagines) is most frequently observed (>90%) in patients failing therapy, either alone or in combination with additional mutations, which is designated as a "pan-class resistance mutation" [12]. It has been demonstrated that efavirenz, as a NNRTI, is clinically effective in controlling HIV infection, but its binding affinity to the K103N RT decreases about 6-fold [13], while nevirapine (another NNRTI, VIRAMUNE®) losses its efficiency in 40-fold to the mutant [14]. Accordingly, many studies [15–18]

Z. Zhang · M. Zheng · L. Du · J. Shen · X. Luo · W. Zhu (✉) · H. Jiang
Center for Drug Discovery and Design, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 555 Zuchongzhi Road, Shanghai 201203, China
e-mail: wlzhu@mail.shcnc.ac.cn

H. Jiang (✉)
School of Pharmacy, East China University of Science and Technology, Shanghai 200237, China

have been carried out since there is a widely accepted opinion that the development of compounds with the ability to inhibit both the wild type (WT) and K103N RTs is critical to a new effective AIDS therapy [13]. Thus, to understand molecular mechanism of the experimentally observed activity loss of the inhibitors due to K103N mutation is a key step towards discovering and designing potent dual functional inhibitors [19]. The X-ray crystal structures indicate that efavirenz, as many other NNRTIs, binds to the non-nucleoside inhibitor-binding pocket (NNIBP) of RT with a butterfly-like conformation [20–22], resulting in a distortion of the RT catalytic aspartate triad [23, 24]. Interaction features between inhibitors and RTs reveal that the moderate flexibility of a drug structure might help to minimize the drug resistance of the mutant [25]. Accordingly, Cocuzza et al. designed and synthesized a series of 4,1-benzoxazepinone analogues of efavirenz [26], some of them show pretty good property as dual functional inhibitors, e.g., compounds **6** and **17** (Table 1).

Three-dimensional quantitative structure–activity relationship analysis (3D-QSAR) is believed to be able to correlate the physico-chemical properties or structural features of compounds with their biological activities, providing useful clues for discovering and designing new drug leads. While, molecular docking can be applied to explore the binding mode and interaction mechanism of a ligand with its target. Therefore, the combination of the 3D-QSAR study with molecular docking would be an approach to study how 4,1-benzoxazepinones interact with the two RTs at atomic level, and to provide the valuable clues for discovering and designing potent dual functional inhibitors.

To the best of our knowledge, no previous effort has been carried out employing molecular docking combined with 3D-QSAR approaches. In this paper, we studied the binding mode of twenty-five 4,1-benzoxazepinones [26] with both WT and K103N RTs, respectively, using molecular docking approach. Following the docking results, 3D-QSAR models were constructed by using approaches of comparative molecular field analysis (CoMFA) [27] and comparative molecular similarity analysis (CoMSIA) [28]. The aims of the present research are to demonstrate the binding mode of the 4,1-benzoxazepinones to WT and K103N RTs; and to construct QSAR models for the inhibitors. These models cannot only be used in rapidly and accurately predicting the activities of designed inhibitors, but also offer a pharmacophore for virtual screening in identifying new types of dual functional inhibitors from existing compound databases.
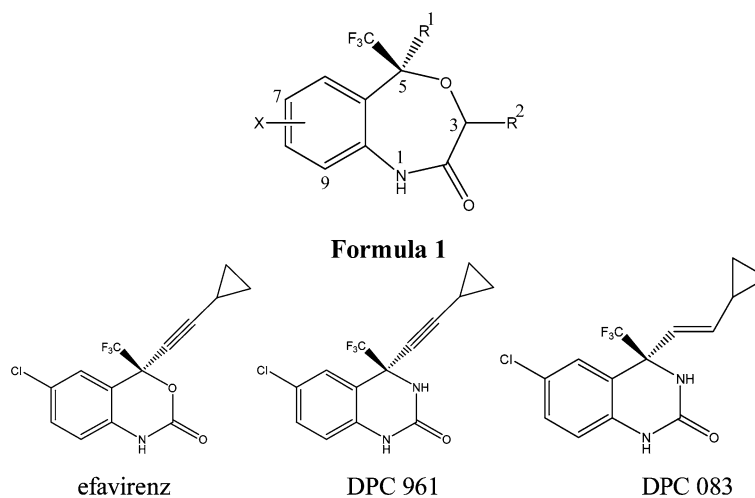
Computational methods

The computational methods are similar to our previous work [29]. Briefly, the 3D structures of the twenty-five 4,1-benzoxazepinone (Table 1) were constructed using the standard parameters of the molecular modeling software package Sybyl 6.9 (http://www.tripos.com), followed by geometrical optimization with the Powell method [30] to a root-mean-squared (rms) energy gradient of 0.05 kcal · mol^{-1} · Å^{-1}. Tripos force field with Gasteiger–Hückel charges [31, 32] was employed during the compounds minimization. These compounds were then docked into the NNIBPs of the crystal structures of WT RT (PDB entry 1FK9) and K103N RT (PDB entry 1FKO) using AutoDock3.0.3 [33]. The conformation with the lowest binding free energy was selected as binding conformation for structural alignment and 3D-QSAR analyses. CoMFA and CoMSIA studies were then performed based on the aligned conformations in training set. The constructed 3D-QSAR models were validated by the compounds in test set. The 25 compounds were randomly divided into a training set of twenty 4,1-benzoxazepinones (unasterisked molecules in Table 1) and a test set of the rest five compounds (**21** and **22**, efavirenz, DPC 961, and DPC 083 with asterisk in Table 1). For the 3D-QSAR studies against K103N RT, only fourteen 4,1-benzoxazepinones were included in training set as other six 4,1-benzoxazepinones showed no inhibitory activity against the RT mutant, but the test set is the same as that for WT RT. All the calculations were carried out on a Silicon Graphics IRIS ORIGIN3200 workstation with four CPUs.
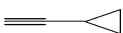
Results and discussion

Interaction modes of inhibitors to WT and K103N RTs

Alignment of the binding conformations of efavirenz to WT and K103N RTs derived from the docking revealed similar conformations to that in crystal structures (Figures not shown). The root mean square deviations (RMSDs) between the docked conformations and the crystal conformations are ~0.34 and ~0.41 Å for WT and K103N RTs, respectively, suggesting that the parameter set for the docking by AutoDock3.0.3 is suitable to reproduce the X-ray structures. Therefore, the docking approach should be acceptable for searching binding conformations of the other inhibitors accordingly. It was found that all inhibitors adopt a butterfly-like conformation within

Table 1 Structures, biological activities pIC_{90} and binding free energies from AutoDock 3.0.3 (ΔG in kcal/mol) of 4,1-benzoxazepinones[a]

Formula 1

efavirenz　　　DPC 961　　　DPC 083

Compound	X	R^1	R^2	pIC_{90}^{b}	ΔG^{b}	pIC_{90}^{c}	ΔG^{c}
1	7-Cl		H	8.04	−11.60	5.75	−6.81
2	7-Cl		H	8.01	−11.28	6.11	−8.28
3	7-Cl		H	7.32	−7.40	ND[d]	−6.90
4	7-Cl		H	8.04	−10.92	5.86	−6.91
5	7-Cl		cis CH_2CF_3	7.77	−9.28	ND	−6.44
6	7-Cl		cis Me	8.66	−11.24	7.54	−10.13
7	7-Cl		trans Me	8.15	−9.64	5.95	−6.05
8	7-Cl		cis nPr	7.76	−10.13	ND	−6.85
9	7-Cl		cis iPr	7.68	−9.57	ND	−6.54
10	7-F		cis Et	8.28	−11.24	6.93	−8.13
11	7-F		trans Me	6.72	−7.11	ND	−6.61
12	7-F		trans Et	7.06	−9.91	ND	−6.18
13	7-F		cis Me	8.28	−11.97	6.48	−8.79
14	6,7-diF		cis Me	8.24	−10.59	7.12	−10.73
15†	7-Cl		H	8.41	−12.31	6.44	−7.41
16†	7-Cl		cis Me	8.60	−11.41	7.46	−10.92
17†	7-Cl		cis Et	8.11	−9.87	7.89	−10.22
18†	7-F		cis Me	8.12	−10.15	6.93	−8.52
19†	7-F		cis Et	8.60	−11.66	7.15	−9.02
20†	6,7-diF		cis Me	8.28	−12.25	7.09	−10.84
21*	7-Cl		H	7.61	−9.95	5.39	−6.95

Table 1 continued

Compound	X	R^1	R^2	pIC_{90}^b	ΔG^b	pIC_{90}^c	ΔG^c
22*	7-Cl	══──◁	*cis* Et	8.08	–11.21	6.86	–8.08
23* (efavirenz)				8.72	–11.46	7.31	–10.94
24* (DPC 961)				8.70	–12.03	8.00	–10.18
25* (DPC 083)				8.68	–12.86	7.57	–10.94

[a] All molecules were taken from literature 26

[b] Result for WT RT

[c] Result for K103N RT

[d] ND = not determined

† R^1 substitution with *trans* geometry

* Compounds of the test set

the binding sites of both the WT and K103N RTs, with similar orientation to efavirenz (result not shown). However, some distinctions were also observed at the atomic level in terms of the binding modes with the two RTs among different inhibitors. Taking compounds **6** and **7** as examples, the only difference between their molecular structures is the *cis-/trans*-substitutes at the R^2 position (Table 1), but compound **6** is a rather good dual functional inhibitor (pIC_{90} is 8.66 and 7.54 for WT and K103N, respectively), while compound **7** losses its efficiency more than 160-fold to the K103N RT (pIC_{90}, 8.15, 5.95). Figure 1a depicts the superposition of the binding sites of WT and K103N RTs with docked compound **6**, which reveals that the compound in the NNIBP of K103N RT (colored in magenta) moves deeper away from the entrance of the binding pocket in comparison with that in WT RT (colored in atomic type). Meanwhile, the positions of the residues Pro 95, Pro 97, Val 106, Tyr 181, Tyr 188, Trp 229 of P66, Glu 138 of P51 are also changed remarkably (labeled in orange). These changes in conformation should result from the K103N mutation [22]. Figure 1b, c show the interaction modes of compound **6** with the two RTs, demonstrating a very similar interaction mode of compound **6** with the NNIBP of WT to that of the K103N RT. The amide group of the ligand forms two hydrogen bonds with the pockets, suggesting that this moiety is essential for inhibiting the RTs activity. On the other hand, the cyclopropyl-ethynyl group and condensed ring of the inhibitor interacts hydrophobically with the binding pockets. Figure 1d, e show the interaction modes of the compound **7** with the NNIBPs of the WT and K103N RTs. Although the compound adopts very similar interaction mode with the WT RT as the compound **6** does through the hydrogen bond and hydrophobic interactions (Fig. 1d), no intermolecular hydrogen bonds are observed between the

compound **7** and the NNIBP of the K103N RT (Fig. 1e). Therefore, it should be the broken hydrogen bond, which plays an essential role in the loss of the inhibitory activity of most 4,1-benzoxazepinones against K103N RT.

Compound **6** is a rather good dual functional inhibitor against the two RTs (Fig. 1b, c). However, it is a little more potent against WT RT than against K103N RT (Table 1). Therefore, to get a deeper insight into the influence of K103N mutation, a comparison was made between the binding details of the compound with WT RT (Fig. 1b) and that with K103N RT (Fig. 1c). The hydrogen bond lengths in Fig. 1b are shorter than that in Fig. 1c, and there are more hydrophobic interaction pairs in Fig. 1b than Fig. 1c, suggesting that both the hydrogen bonding and hydrophobic interactions between compound **6** and WT RT are stronger than the interactions between the compound with K103N. This observation is in agreement with the calculated binding free energy (Table 1). Moreover, the compound may form additional alkane-π interactions [35] with Val106 of WT RT (Fig. 1a). Figure 2 depicts the calculated hydrophobic and electrostatic surfaces for the compound and the NNIBPs of the two proteins. Obviously, the hydrophobic surface of the compound match more complementary to NNIBP of WT RT (Fig. 2a) than to that of K103N RT (Fig. 2c), which could be attributed to the upside-down shift of the orientation of residue 181 upon K103N mutation [22]. On the other hand, the mutation does not bring obvious change in the match of the electrostatic interaction between the inhibitor and the binding pockets of both WT and K103N RTs (Fig. 2b, d). Thus, both Figs. 1, 2 demonstrate that the mutation of K103N leads to the differences in hydrogen bonding, electrostatic and hydrophobic interactions, consequently, different inhibitory abilities of the compound **6** against the two RTs.

Fig. 1 Superposition of the NNIBPs of WT and K103N RTs and schematic representations of hydrogen bonds and hydrophobic interactions of compounds **6** and **7**, respectively, with the WT and K103N RTs. Dashed lines represent hydrogen bonds and spiked residues form hydrophobic contacts with the ligands. (**a**) The superposition of the NNIBPs of WT and K103N RTs via protein backbone with docked compound **6**. The residues within 6.5 Å around compound **6** are colored in black for the WT RT (PDB entry 1FK9, with compound **6** in atomic type color), and magenta for K103N RT (PDB entry: 1FKO, with compound **6** in magenta), respectively. Those residues numbered in orange show remarkable conformational changes after K103N mutation. (**b**) Compound **6** with WT RT. (**c**) Compound **6** with K103N RT. (**d**) Compound **7** with WT RT. (**e**) Compound **7** with K103N RT. Figure 1**a** was generated with WebLab and rendered with POV-ray (http://www.povray.org/). Figure 1**b**–**e** were generated by the program LIGPLOT [34]
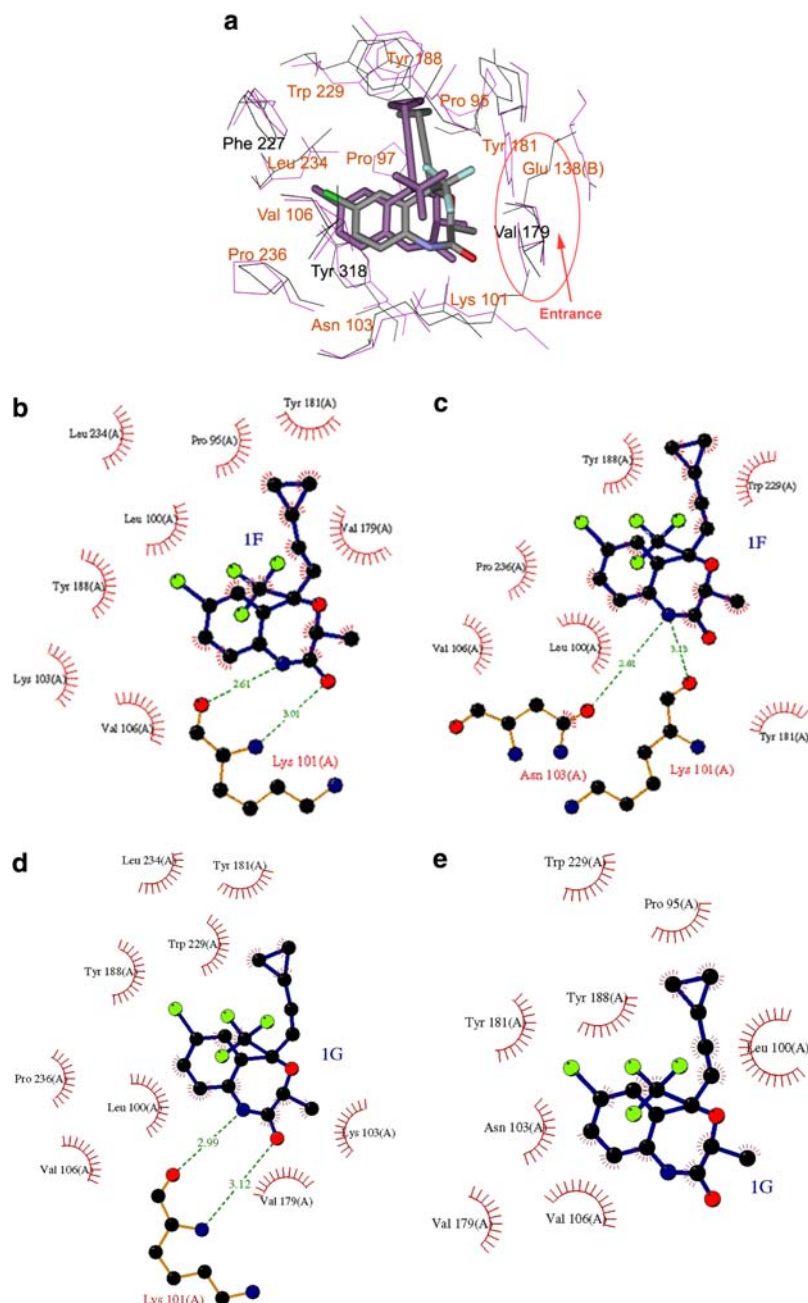
The correlation between binding free energy and inhibitory activity

The predicted binding free energies (ΔGs) of the 4,1-benzoxazepinones with WT and K103N RTs by using the software package AutoDock3.0.3 are shown in Table 1. Rather good correlations (Eq. (1) for WT RT and Eq. (2) for K103N RT) were found between the predicted ΔGs and the experimentally determined inhibitory activities, $-\log IC_{90}$ values, using linear regression analysis method (Fig. 3a, b), with r^2 values of 0.631 and 0.711 for WT RT and K103N RT, respectively, indicating that our predicted ΔG values, thereby, the docking approach and parameters, are reasonable.

$$-\log IC_{90} = 5.021 - 0.285 \times \Delta G_{WT}$$
$$(n = 20, \ r^2 = 0.631, \ F = 33.483, \ s = 0.307) \tag{1}$$

$$-\log IC_{90} = 3.634 - 0.357 \times \Delta G_{K103N}$$
$$(n = 14, \ r^2 = 0.711, \ F = 32.951, \ s = 0.362). \tag{2}$$

Fig. 2 The hydrophobic and electrostatic surfaces of NNIBPs of the RTs and compound **6**. (**a**) The complementarities of the lipophilic potential (LP) surfaces between the WT RT NNIBP and compound **6**. (**b**) The complementarities of the electrostatic potential (EP) surfaces between the WT RT NNIBP and compound **6**. (**c**) The LP surface complementarities between K103N RT NNIBP and compound **6**. (**d**) The EP surface complementarities between K103N RT NNIBP and compound **6**. Brown and blue colors represent lipophilic and hydrophilic tendencies, respectively, while red and blue colors represent positive and negative potential, respectively
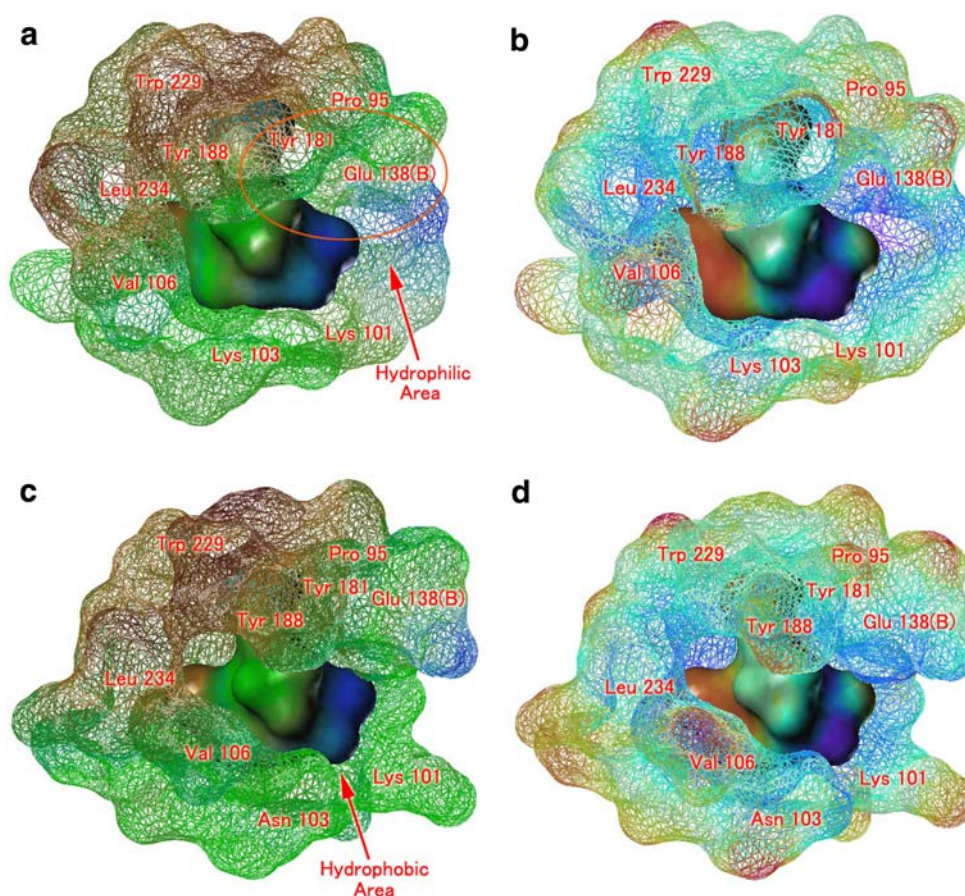
QSAR models for WT RT

Testing on different gird size shows that the default value of 2 Å is better than less grid sizes (1.5 and 1.0 Å, data not shown). Table 2 provides the PLS analysis results for QSAR studies using different probe atoms with the default gird size of 2 Å. The best CoMFA result for WT inhibition is from the probe atom of sp^3 C(+1) (Table 2a, $r^2_{cv} = 0.743$). The non-cross-validated PLS analysis with the number of components (noc) of five revealed a conventional r^2 value of 0.988; a standard error of estimate, s, of 0.064; and an F value of 200.816. Regarding CoMSIA model, the best CoMSIA model for WT RT inhibition with a r^2_{cv} value of 0.656 for five components was achieved at the default lattice step size of 2.0 Å (Table 2b), and the conventional PLS analysis releases r^2, s and F values are 0.931, 0.147, 143.495, respectively. All theses calculated parameters suggest that both the CoMFA and the CoMSIA models are fairly predictive. The steric and electrostatic contributions from the best CoMFA model are 61.0 and 39.0%, respectively; and the steric, electrostatic, hydrophobic, H-bond donor, and H-bond acceptor contributions from the best CoMSIA model are 27.3, 16.7, 30.1, 18.6 and 7.3%, respectively. Thus, both

CoMFA and CoMSIA results demonstrate that steric and hydrophobic interactions are more important than the others between inhibitors and WT RT. This is in good agreement with the observation from Fig. 2 that the mismatch in hydrophobic interaction after K103N mutation results in a drop in inhibitory ability of compound **6** against K103N RT.

The predicted activities of the compounds in the training set by the best CoMFA and CoMSIA models versus their experimental activities (pIC_{90}) are listed in Table 3. Good correlations were observed between the experimental and the predicted activities (Figures not shown), indicating that the two models should be highly predictive. Furthermore, the predicted $-logIC_{90}$ values for those compounds losing their activities against K103N RT were predicted less than 5.59 and 5.83 by CoMFA and CoMSIA, respectively, that are in well agreement with the experimental observations that the compounds are weaker inhibitors against K103N RT.

3D-QSAR models for K103N RT

The PLS statistics of the 3D-QSAR model with different probe atoms for K103N RT were also summa-
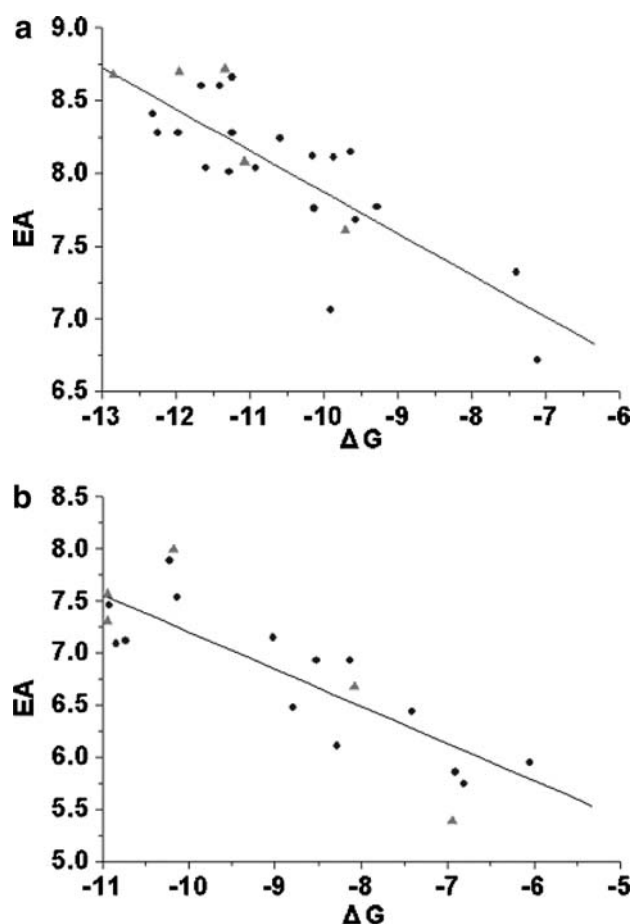
Fig. 3 Correlation between the predicted binding free energies from AutoDock 3.0.3 (ΔGs in kcal/mol) of 4,1-benzoxazepinones to, (**a**) WT RT; (**b**) K103N RT, with the experimental activities ($-\log IC_{90}$, EA). ●, compounds of the training set; ▲, compounds of the test set

Table 2 Statistical indexes of the constructed 3D-QSAR models based on the docked binding conformations using different probe atoms[a]

Probe atom	PLS statistics	WT model	K103N model
(a) CoMFA results			
$sp^3C(+1)$	r^2_{cv}	0.743	0.700
	noc	5	2
	r^2	0.988	0.866
	S	0.064	0.272
	F	200.816	90.101
	Field contribution (%)		
	Steric	61.0	58.5
	Electrostatic	39.0	41.5
$sp^3O(-1)$	r^2_{cv}	0.721	0.681
	noc	6	4
	r^2	0.976	0.864
	S	0.086	0.274
	F	132.334	88.895
H(+1)	r^2_{cv}	0.678	0.676
	noc	6	5
	r^2	0.960	0.824
	S	0.108	0.286
	F	102.835	85.002
(b) CoMSIA results			
$sp^3C(+1)$	r^2_{cv}	0.656	0.834
	noc	5	4
	r^2	0.931	0.974
	S	0.147	0.136
	F	143.495	101.127
	Field contribution (%)		
	Steric	27.3	22.1
	Electrostatic	16.7	24.4
	Hydrophobic	30.1	18.8
	H-donor	18.6	25.6
	H-acceptor	7.3	9.1

[a] The default grid size value of 2 Å was the best choice for both CoMFA and CoMSIA models

rized in Table 3. Testing on different probe atoms (sp^3 O(-1) and H($+1$)) and grid spacing size (1.5 and 1.0 Å) also shows that the default probe atom, sp^3 C($+1$), and default size of 2 Å produce the best CoMFA and CoMSIA results for K103N RT with r^2_{cv} values of 0.700 and 0.834, respectively, (Table 2a, b). The conventional r^2, s and F values are 0.866, 0.272 and 90.101 for CoMFA, and 0.974, 0.136 and 101.127 for CoMSIA, respectively. Meanwhile, the steric and electrostatic contributions from CoMFA analysis are 58.5 and 41.5%, respectively. The less contribution from steric than that in CoMFA for WT RT is consistent with the mismatched hydrophobic interaction upon K103N mutation (Fig. 2a, c). In addition, CoMSIA analysis also reveals a significant change in the hydrophobic interaction. The hydrophobic contribution for the binding of the inhibitors to WT RT is ~30.1% that is the most important contribution among all interactions, but it is only ~18.8% for the binding of the

inhibitors to K103N RT. These results are in agreement with the deduction based on Fig. 2a, c. The first three important factors to the binding of K103N mutant are steric, electrostatic and hydrogen bond donor interactions, which is in line with the conclusion of the previous 3D-QSAR analysis on a class of nevirapine derivatives against the HIV-1 Y181C RT [36].

Testing of 3D-QSAR model

The values of the predicted $-\log IC_{90}$ of the five compounds in the test set using our constructed CoMFA and CoMSIA models are consistent with the calculated ΔGs with Autodock3.0.3 (triangle labeled molecules in Fig 3a, b), demonstrating that these models are highly predictive (Predicted activity values in Table 3), therefore, should be helpful in designing new RT inhibitors.

Table 3 Predicted activities (PA, –logIC$_{90}$) vs. experimented activities (EA, –logIC$_{90}$) and residues (δ) of 4,1-benzoxazepinones by CoMFA and CoMSIA[a]

Compd.	EA[a]	EA[b]	CoMFA[b]		CoMSIA[b]		CoMFA[c]		CoMSIA[c]	
			PA	δ	PA	δ	PA	δ	PA	δ
1	8.04	5.75	8.06	–0.02	8.08	–0.04	5.74	0.01	5.87	–0.12
2	8.01	6.11	8.01	–0.00	8.02	–0.01	5.82	0.29	5.97	0.14
3	7.32	ND	7.37	–0.05	7.12	0.20	5.03	ND	5.83	ND
4	8.04	5.86	8.12	–0.08	8.16	–0.12	5.78	0.08	5.91	–0.05
5	7.77	ND	7.71	0.06	7.79	–0.01	4.82	ND	4.75	ND
6	8.66	7.54	8.58	0.08	8.50	0.16	7.16	0.38	7.53	0.01
7	8.15	5.95	8.13	0.02	8.13	0.12	6.43	–0.48	6.12	–0.17
8	7.76	ND	7.77	–0.01	7.81	–0.05	5.59	ND	5.77	ND
9	7.68	ND	7.75	–0.07	7.90	–0.22	5.56	ND	5.31	ND
10	8.28	6.93	8.31	–0.03	8.30	–0.02	6.64	0.29	6.90	–0.06
11	6.72	ND	6.74	–0.02	6.84	–0.12	5.57	ND	5.38	ND
12	7.06	ND	6.99	0.07	7.06	–0.00	4.95	ND	5.42	ND
13	8.28	6.48	8.22	–0.06	8.36	–0.08	6.68	–0.20	6.67	–0.19
14	8.24	7.12	8.16	0.08	8.15	0.09	7.04	–0.08	7.17	–0.05
15	8.41	6.44	8.45	–0.04	8.41	–0.00	6.36	0.08	6.37	0.07
16	8.60	7.46	8.61	–0.01	8.59	0.01	7.68	–0.22	7.51	0.05
17	8.11	7.89	8.07	0.04	8.29	–0.18	7.64	0.25	7.94	–0.05
18	8.12	6.93	8.12	–0.00	8.02	0.10	7.19	–0.26	6.69	0.24
19	8.60	7.15	8.64	–0.04	8.50	0.10	7.21	–0.06	7.14	0.01
20	8.28	7.09	8.22	0.06	8.29	–0.01	7.29	–0.20	7.10	–0.01
21*	7.61	5.39	7.68	–0.07	7.76	–0.15	5.74	–0.35	5.38	0.01
22*	8.08	6.86	7.99	0.09	7.75	0.33	7.16	–0.30	6.68	0.11
23* (efavirenz)	8.72	7.31	8.60	0.12	8.73	–0.01	7.11	0.20	7.18	0.13
24* (DPC 961)	8.70	8.00	8.68	0.02	8.53	0.17	8.09	–0.09	7.93	0.07
25* (DPC 083)	8.68	7.57	8.77	–0.09	8.75	–0.07	7.53	0.04	7.56	0.01

[a] All molecules were taken from literature 26

[b] Results for WT RT

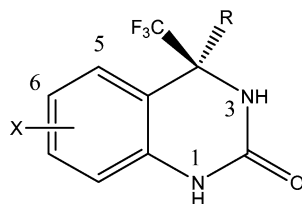[c] Results for K103N RT

*Compounds in the test set

However, the test set looks too little to convincingly validate the constructed 3D-QSAR models, therefore, new validation is necessary. Although no compounds with the same molecular skeleton as that in the training set was found, we noticed a set of 25 quinazolinones that have similar molecular structures as training set with available inhibitory activity tested under the same experimental protocol [37] against the two RTs (Table 4, Formula 2). So, this set of compounds was used to further testify the reliability and extrapolation capability of the constructed 3D-QSAR models. The procedure and parameter used for 3D-QSAR validation by the 25 quinazolinones are the same as that for 4,1-benzoxazepinones. The constructed 3D-QSAR models were employed to predict corresponding inhibitory activities of the new set of compounds.

The correlation coefficients, r^2, of the experimental activities versus predicted activities of the 25 compounds using the CoMFA and the CoMSIA models are 0.795 and 0.676 for WT RT, and 0.736 and 0.807 for K103N RT, respectively. Considering the experimental data in the new test set were from a different batch of experiment with different molecular skeleton, the constructed QSAR models should be highly predictive and have strong extrapolation capability.

CoMFA contour maps correlate with WT and K103N RT topologies

The CoMFA contour polyhedra of the steric and electrostatic contributions for WT and K103N RTs were depicted in Fig. 4a, b, respectively. To aid in visualization, compound **6** is displayed in the maps together with the residues of the NNIBP within 6.5 Å surrounding the active sites. The colored polyhedra in the map indicate those areas in 3D space where changes in the field values for the 4,1-benzoxazepinones correlate strongly with concomitant changes in inhibitory activities. Detrimental and beneficial steric interactions are displayed in yellow and green contours, respectively, while blue and red contours illustrate the regions of desirable positive and negative electrostatic interactions. The distributions of CoMSIA contours are similar to the CoMFA (Figures not shown), therefore, only CoMFA contours are discussed below.

Regarding the CoMFA contour for WT RT, there is a big region of yellow contour near the cyclopropyl group of compound **6** (Fig. 4a), which shows that increasing steric bulk at this position is unfavorable for the inhibitory activities. Indeed, it was found that there

Table 4 Structure, predicted activities (PA, –logIC$_{90}$) vs. experimental activities (EA, –logIC$_{90}$) and residues (δ) of quinazolinones by CoMFA and CoMSIA[a]

Formula 2

Compd.	X	R	EA[b]	EA[c]	CoMFA[b]		CoMSIA[b]		CoMFA[c]		CoMSIA[c]	
					PA	δ	PA	δ	PA	δ	PA	δ
1	6-Cl		8.47	6.80	7.94	0.53	7.82	0.65	6.15	0.65	6.74	0.06
2	6-Cl		8.15	6.60	7.86	0.29	7.89	0.26	6.23	0.37	6.48	0.12
3	6-F		8.60	7.15	8.71	–0.11	8.43	0.17	6.34	0.81	6.97	0.18
4	6–F		8.30	6.32	7.74	0.56	7.78	0.52	6.19	0.13	6.69	–0.37
5	6-F		8.18	6.49	7.93	0.25	7.65	0.53	6.21	0.28	6.51	–0.02
6	5-Cl, 6-F		8.63	7.14	8.29	0.34	8.74	–0.11	7.20	–0.06	6.74	0.40
7	5,6-diF		8.21	6.72	7.70	0.51	7.58	0.63	6.68	0.04	6.64	0.08
8	6-MeO		8.09	6.47	7.66	0.43	7.70	0.39	6.25	0.22	6.75	–0.28
9	6-MeO		8.49	6.55	8.32	0.17	8.61	–0.12	6.44	0.11	6.65	–0.10
10	5,6-diF		8.82	7.85	8.85	–0.03	8.63	0.19	7.68	0.17	7.46	0.39
11	5-F		8.85	7.05	8.77	0.08	8.62	0.23	7.41	0.64	7.30	–0.25
12	5-Cl, 6-F		8.52	7.82	8.42	0.10	8.56	–0.04	7.45	0.37	7.40	0.42
13	5-Cl		8.60	7.20	8.83	–0.23	8.13	0.47	7.39	–0.19	7.31	–0.11
14	5,6-diF		8.67	7.88	8.73	–0.06	8.58	0.09	7.33	0.55	7.33	0.55
15	5,6-diF		8.67	7.85	8.71	–0.04	8.49	0.18	7.47	0.38	7.43	0.42
16	6-F		8.70	7.32	8.76	–0.06	8.61	0.09	7.37	–0.05	7.33	–0.01
17	5,6-diF		8.70	6.96	8.67	0.03	8.69	0.01	7.11	–0.15	7.08	–0.12
18	5-Cl, 6-F		8.57	7.74	8.78	–0.21	8.60	–0.03	7.32	0.42	7.30	0.44
19	6-MeO		8.54	7.40	8.62	–0.08	8.59	–0.05	7.48	–0.08	7.28	0.12
20	5-F, 6-Cl		8.32	7.74	8.16	0.16	8.21	0.11	7.34	0.40	7.38	0.36
21	6-MeO		8.42	7.25	8.21	0.21	8.30	0.12	7.37	–0.08	7.16	0.09
22	6-Cl		8.48	7.59	8.35	0.13	8.40	0.08	7.42	0.17	7.43	0.16
23	6-F		8.59	7.57	8.78	–0.19	8.57	0.02	7.49	0.08	7.44	0.13

Table 4 continued

Compd.	X	R	EA[b]	EA[c]	CoMFA[b]		CoMSIA[b]		CoMFA[c]		CoMSIA[c]	
					PA	δ	PA	δ	PA	δ	PA	δ
24	5,6-diCl	═══──◁	8.10	7.74	7.95	0.15	7.98	0.12	7.36	0.38	7.37	0.37
25	6-Cl	═══──<	8.52	7.66	8.48	0.04	8.71	−0.19	7.49	0.17	7.50	0.16

[a] All molecules were taken from literature 37

[b] Results for WT RT

[c] Results for K103N RT

are three aromatic sidechains from residues Tyr181, Tyr188 and Trp229 in the sub-pocket of NNIBP within a distance of less than 3.0 Å to the cyclopropyl group, therefore, any larger substitute may collide with these residues, leading to dropping in bioactivity. The inhibitory activity order of **1** >**2** > **3** is in line with this explanation since substitution sizes increase from ethyl, isopropyl to furanyl. This is also in good agreement with the previous 3D-QSAR work on a set of nevirapine derivatives [36]. Another big region of yellow contour near the *trans*-substitute of C3 of the compounds (Formula 1 in Table 1) indicates that bulky substituents in this position could decrease the molecular inhibitory activity. Indeed, compounds **11** and **12** with *trans*-methyl and ethyl groups exhibit weaker binding affinities than most of others. Green polyhedra around *cis*-substitute of C3 of the compound show that more bulky substituent is needed in this position to improve molecular inhibitory activities. Experimental data demonstrate that the moderately increasing in the volume of *cis*-substitute at C3 results in an increase in inhibitory activity, such as **6** > **4**, **16** > **15**, and **19** > **18**. A big region of green contour is opposite to the C7 of the compound, indicating that more bulky substituents in this position are favored as stronger interaction might be achieved through the interaction with Leu234. It might be the reason that compounds with chlorine-substitute at this position perform better inhibition than those with fluorine-substitute: **7** > **11**, **6** > **13**, **16** > **18**.

Regarding electrostatic contours, there is a blue polyhedra near the amide of the compound indicates that the positively charged substituents around this position may increase inhibitory activity. This could be attributed to the strong hydrogen bond between the main chain carbonyl oxygen of Lys101 and the amide NH of compound. Another blue polyhedra was found close to C7 of compound, showing that less negative substituents should strengthen the binding of the inhibitors to the WT RT. The electronegativity of Cl

atom is less than that of F atom. This could be another reason that the inhibitory activities of the compounds with chlorine-substitute at this position are stronger than those with fluorine-substitute. A red polyhedron around carbonyl oxygen of compound demonstrates
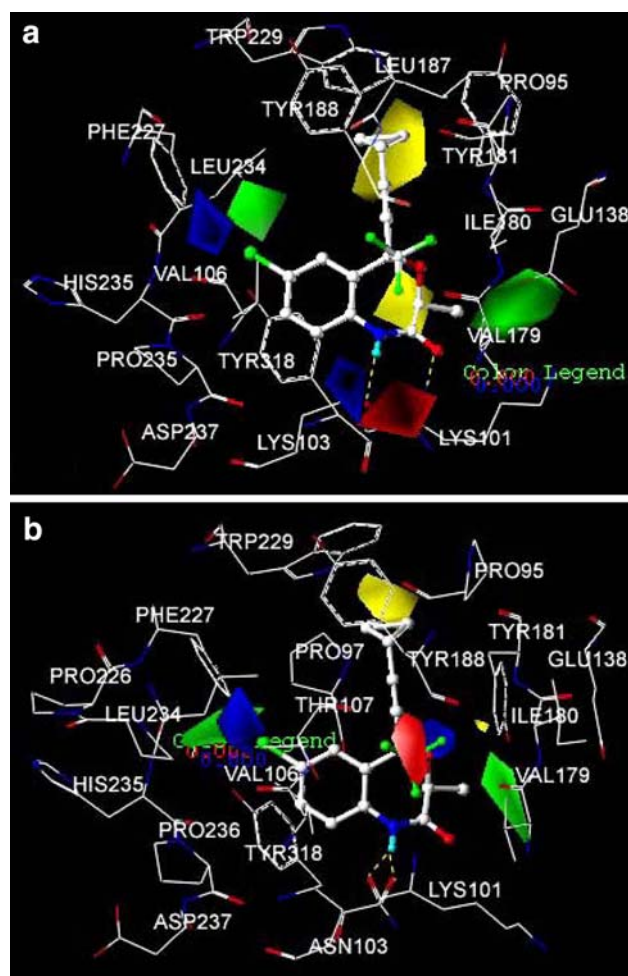

Fig. 4 CoMFA contour maps displayed with compound **6** and key residues of the NNIBPs within 6.5 Å surrounding the compound. (**a**) Steric and electrostatic contour maps for WT RT; (**b**) steric and electrostatic contour maps for K103N RT

that more negative substituent in this region is favorable to improving the inhibitory activity of a compound. This is in agreement with another hydrogen bond between the carbonyl oxygen of the compound and the NH in the main chain of Lys101.

Figure 4b depicts the CoMFA contour of K103N inhibition, which shows four distinctive contour regions in comparison with Fig. 4a. These changes should be the result of mutation K103N that leads to remarkable shifts in positions of several key residues in NNIBP: Pro 95, Pro 97, Val 106, Tyr 181, Tyr 188, Trp 229 of P66, Glu 138 of P51 (Figs. 1a, 2) [22], consequently the different inhibitory activities of same ligand against WT RT and K103N RT.

In comparison with Fig. 4a, two regions of colored polyhedra disappeared in Fig. 4b. One is the big region of yellow contour in Fig. 4a near the *trans*-substitute of C3 of the compound. This is in agreement with the fact that most of *trans*-substitute compounds, e.g., compounds **11** and **12** (Table 1), lost the inhibitory activities for K103N mutation. As the mutated Asn 103 could push the compound deeper into the NNIBP, the *trans*-substitute at R^2 position of ligand might reinforce this movement and then breaks the hydrogen-bonding connection between ligand and receptor which ultimately leads to the loss of inhibitory activities against K103N RT of these compounds. The other is the blue and red polyhedra near the amide of compound (Fig. 4a), indicating that the electrostatic interactions in this area, which should be essential to the binding of the compounds to the binding pocket of WT RT, is not an essential factor affecting inhibitory activity due to the K103N mutation. Obviously, it's reasonable since most of this series of compounds lose the effective hydrogen bonding at this position after K103N mutation (Fig. 1d, e, represented by compound **7**). Therefore, a good inhibitor with dual function against both WT and K103N RT should have at least good electrostatic or hydrogen bond interactions with residues in the NNIBP near to amide group, and avoid bulky functional group interacting with the residues near to *trans*-substitute of C3.

On the other hand, there are two new regions of colored electrostatic polyhedras appeared in Fig. 4b in comparison with Fig. 4a. One is a red polyhedron near the trifluoromethyl of compound **6**, which is in agreement with the crystal structure that there are positively charged backbone atoms of polar residue at the corresponding position, namely, Tyr188 (Fig. 2d). The other is a blue polyhedron close to oxygen atom at O4 of the compound (Formula 1 in Table 1), indicating that low electron density in this region would increase the inhibitory activity. This is consistent with the sug-

gestion that an appropriate polar or charged linker group at this position might stabilize the binding of inhibitors to the mutated RT [21]. Moreover, the replacement of the O4 atom with N atom which has a lower electron density leads to a potent dual functional inhibitor, DPC 961 [38], against the two RTs.

Pharmacophore elements and drug design implication

Considering the interaction mechanism of 4,1-benzoxazepinones to the RTs, the 3D-pharmacophore elements of these compounds were figured out by running the software CATALYST 4.6 (http://www.accelrys.com/). As shown in Fig. 5, the left cyan hydrophobic center element (HYP1) is to contact with Val 106 through alkane-π interactions; magenta hydrogen bond donor element (HBD) is corresponding to the main chain carbonyl oxygen of Lys101 or Asn103; the green hydrogen bond acceptor element (HBA) is opposite to the main chain NH of Lys101; the upper cyan hydrophobic center element (HYP2) could interact with Tyr181, Tyr188 and Trp229 via hydrophobic interactions (refer to Fig. 2a for the residues). With this
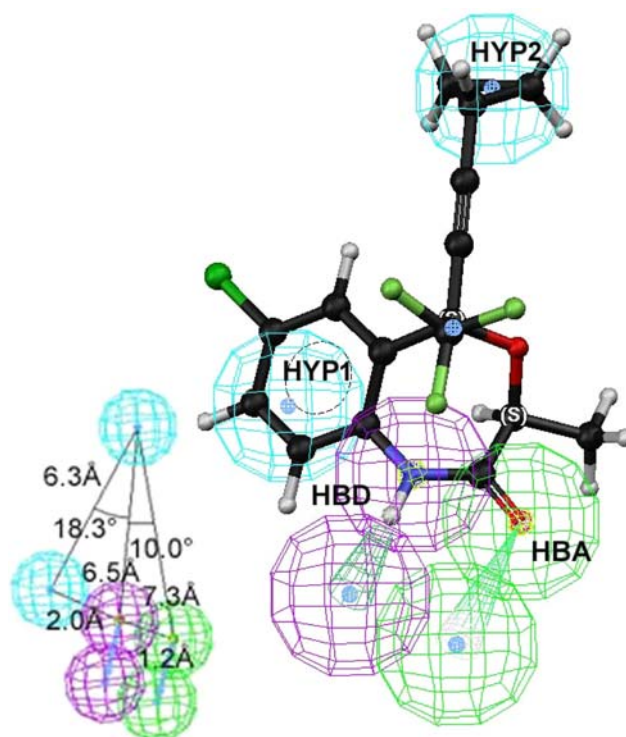

Fig. 5 Pharmacophore elements of 4,1-benzoxazepinones. Compound **6** is mapped with the pharmacophore hypotheses. The inter-bond angles and the distance between elements are also annotated. This picture was drawn by CATALYST software (http://www.accelrys.com/)

model, virtual screening against compound databases could be performed for discovering potential compounds as dual functional inhibitors against both WT and K103N RTs.

To further validate this pharmacophore model, two tests were carried out. Test one is about the success rate of identifying the 50 inhibitors in Tables 1, 4 from the non-inhibitors. Five hundreds compounds that were randomly selected from SPECS database (http://www.specs.net) were mixed with the 50 inhibitors, then, a virtual screening was performed based on the pharmacophore model using CATALYST 4.6. Exactly, the 50 inhibitors and only the 50 inhibitors were successfully differentiated from the 500 compounds from SPECS. It was found that each of the pharmacophore elements is mapped correctly onto the corresponding functional group of each inhibitor with highly consistent fit values ranging from 5.05 to 5.99, indicating that these hypothetic pharmacophore elements align with the functional group very well.

Test two is on discovering potent inhibitors by applying the pharmacophore model on virtual screening in the database SPECS. Fifty compounds were identified through the virtual screening as potential dual inhibitors against the two RTs. The preliminary bioassay on 12 compounds of them shows that two of them have good inhibitory activity against WT RT with the IC_{50} value less than 10 μM measured by the same method as reference [39] (structures not shown). Therefore, these two tests demonstrate that the pharmacophore model could specifically represent the common features of dual functional inhibitors against both the HIV WT RT and K103N RT.

On the basis of our constructed 3D-QSAR models and pharmacophore elements, a potent dual inhibitor that is capable of inhibiting both WT and K103N RTs should include following features (refer to Formula 1 in Table 1 for atom numbering): (i) adopting the positively charged aliphatic group at the *cis*-substituent of C3, e.g., methanamine, ethanamine, which are supposed to interact with negatively charged residue Glu138 of P51, and hydrophobic residues Val179, Ile180 of P66; (ii) reducing the electronic density at the position of O4 via changing O atom to N atom or even carbon atom; (iii) positioning a small branched aliphatic group at position of C5 to fit the space enclosed by residues Tyr181, Tyr188 due to K103N mutation; (iv) using the negatively charged bulky substituents at position of C7 to interact with its surrounding residues, such as Leu234. These features should be suitable not only for designing 4,1-benzoxazepinone analogues, but also for the new chemical entity.

Conclusions

Molecular docking and 3D-QSAR studies were carried out to explore the inhibition mechanism of 4,1-benzoxazepinones against HIV-1 RTs of both wild-type (WT) and K103N mutant-type (K103N). The CoMFA and CoMSIA models have reasonably high r^2_{cv} values ranging from 0.656 to 0.834 for WT and K103N RTs. Validation by two sets of compounds show that the constructed models are not only highly predictive, but also extrapolative. The contour maps of these models match the structural topology of the NNIBPs of the two RTs very well, which clearly shows the different binding mechanism of the inhibitors for these two RTs. The CoMFA analysis reveals that the steric interactions are the major factor for inhibition to WT, whilst CoMSIA analysis shows that steric, electrostatic and hydrogen bond donor effects have approximately the identical contributions to the binding of the inhibitors to K103N. These results lead to a better understanding of the interactions of 4,1-benzoxazepinones with WT and K103N RTs. In addition, a reasonable pharmacophore model for the compounds was established. Based on the 3D-QSAR models for the two RTs, and taking into account the pharmacophore studies, some useful clues were figured out for designing new potent dual functional inhibitors against both the wild type and K103N RTs of HIV-1. Indeed, the application of the model led to a discovery of two potent HIV WT RT inhibitors.

References

1. Brian GT, Summers MF (1999) J Mol Biol 285:1
2. Emerman M, Malim MH (1998) Science 280:1880
3. Mitsuya H, Broder S (1987) Nature 325:773
4. Esnouf R, Ren J, Ross C, Jones Y, Stammers D, Stuart D (1995) Nat Struct Biol 2:303
5. Jones PS (1998) Antiviral Chem Chemother 9:283
6. Pedersen OS, Pedersen EB (1999) Antiviral Chem Chemother 10:285
7. Vella S, Palmisano L (2000) Antiviral Res 45:1
8. Bardsley-Elliot A, Perry CM (2000) Pediatric Drugs 2:373
9. Corbett JW, Ko SS, Rodgers JD, Gearhart LA, Magnus NA, Bacheler LT, Diamond S, Jeffrey S, Klabe RM, Cordova BC, Garber S, Logue K, Trainor GL, Anderson PS, Erickson-Viitanen SK (2000) J Med Chem 43:2019
10. Moyle G (2001) Drugs 61:19
11. Lee K, Gulick RM (2001) Curr Infec Dis Rep 3:193
12. Adkins JC, Nobel S (1998) Drugs 56:1055
13. Robert WB Jr (2001) Expert Opin Investig Drugs 10:1423

14. Young SD, Britcher SF, Tran LO, Payne LS, Lumma WC, Lyle TA, Huff JR, Anderson PS, Olsen DB, Carroll SS (1995) Antimicrob Agents Chemother 39:2602
15. Lindberg J, Sigurdsson S, Lowgren S, Andersson HO, Sahlberg C, Noreen R, Fridborg K, Zhang H, Unge T (2002) Eur J Biochem 269:1670
16. Ren J, Nichols C, Bird L, Chamberlain P, Weaver K, Short S, Stuart DI, Stammers DK (2001) J Mol Biol 312:795
17. Jay AM, David DC, Abdul M, Beverly CC, Ronald MK, Steven PS (2001) Bioorg Med Chem Lett 11:619
18. Chamberlain PP, Ren J, Nichols CE, Douglas L, Lennerstrand J, Larder BA, Stuart DI, Stammers DK (2002) J Virol 76:10015
19. Medina-Franco JL, Rodríguez-Morales S, Juárez-Gordiano C, Hernández-Campos A, Castillo R (2004) J Comput Aided Mol Des 18:345
20. De Clercq E (2001) Curr Med Chem 8:1543
21. Ding J, Das K, Moereels H, Koymans L, Andries K, Paul AJ, Atephen J, Huges H, Arnold E (1995) Struct Biol 2:407
22. Ren J, Milton J, Weaver KL, Short SA, Stuart DI, Stammers DK (2000) Structure 8:1089
23. Kohlstaedt LA, Wang J, Friedman JM, Rice PA, Steitz TA (1992) Science 256:1783
24. Ding J, Das K, Tantillo C, Zhang W, Clark ADJ, Jessen S, Lu X, Hsiou Y, Jacobo-Molina A, Andries K (1995) Structure 3:365
25. Mao C, Sudbeck EA, Venkatachalam TK, Uckun FM (1999b) Antiviral Chem Chemother 10:233
26. Cocuzza AJ, Chidester DR, Cordova BC, Klabe RM, Jeffrey S, Diamond S, Weigelt CA, Ko SS, Bacheler LT, Erickson-Viitanen SK, Rodgers JD (2001) Bioorg Med Chem Lett 11:1389
27. Cramer RD, Patterson DE, Bunce JD (1988) J Am Chem Soc 110:5959
28. Klebe G, Abraham U (1999) J Comput Aided Mol Des 13:1
29. Liu G, Zhang Z, Luo X, Shen J, Liu H, Shen X, Chen K, Jiang H (2004) Bioorg Med Chem 12:4147
30. Powell MJD (1977) Math Programming 12:241
31. Marsili M, Gasteiger J (1980) Croat Chem Acta 53:601
32. Gasteiger J, Marsili M (1980) Tetrahedron 36:3219
33. Morris GM, Goodsell DS, Huey R, Hart WE, Halliday S, Belew R, Olson AJ (1999) AutoDock, Version 3.0.3. The Scripps Research Institute, Molecular Graphics Laboratory, Department of Molecular Biology
34. Wallace AC, Laskowski RA, Thornton JM (1995) Protein Eng 8:127
35. Kim KS, Tarakeshwar P, Lee JY (2000) Chem Rev 100:4145
36. Pungpo P, Hannongbua S (2000) J Mol Graphics Mod 18:581
37. Corbett JW, Ko SS, Rodgers JD, Gearhart LA, Magnus NA, Bacheler LT, Diamond S, Jeffrey S, Klabe RM, Cordova BC, Garber S, Logue K, Trainor GL, Anderson PS, Erickson-viitanen SK (2000) J Med Chem 43:2019
38. Campiani G, Ramunno A, Maga G, Nacci V, Fattorusso C, Novellino E (2002) Curr Pharma Des 8:615
39. Urabe T, Sano K, Tanno M, Mizoguchi J, Otani M, Lee MH, Takasaki T, Kusakabe H, Imagawa DT, Nakai M (1992) J Virol Methods 40:145