

# RSC Advances



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

*Accepted Manuscripts* are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. This *Accepted Manuscript* will be replaced by the edited, formatted and paginated article as soon as this is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

## ARTICLE

# Molecular modeling studies of dihydro-alkyloxy-benzyl-oxopyrimidines (DABOs) as Non-nucleoside Inhibitors of HIV-1 Reverse Transcriptase using 3D-QSAR, Topomer CoMFA and molecular docking simulations

Cite this: DOI: 10.1039/x0xx00000x

Minghui Dong,<sup>a</sup> Yujie Ren\*<sup>a</sup>

The human immunodeficiency virus type 1 (HIV-1) reverse transcriptase (RT) is generally regarded as target for the treatment of Acquired immune deficiency syndrome (AIDS). Non-nucleoside RT inhibitors (commonly known as NNRTIs) are currently being used in the treatment of HIV-1 infections. In this work, a series of NNRTIs were studied using a combination of molecular modeling techniques including three-dimensional quantitative structure-activity relationship (3D-QSAR), Topomer CoMFA and molecular docking simulations. The best optimum comparative molecule field analysis (CoMFA) model yielded a leave-one-out correlation coefficient ( $q^2$ ) and a non-cross-validated correlation coefficient ( $r^2$ ) of 0.636 and 0.993, respectively. The respective  $q^2$  and  $r^2$  of the best comparative molecular similarity indices analysis (CoMSIA) model were 0.655 and 0.998. The models were validated by test sets, and predicted correlation coefficients ( $r^2_{\text{pred}}$ ) of 0.907 and 0.886 obtained from the CoMFA and CoMSIA models, thus judging the robustness of the model. The analysis of Topomer CoMFA, obtained a  $q^2$  of 0.546 and a  $r^2_{\text{pred}}$  value of 0.718 which suggested the model had a good predictive ability ( $q^2 > 0.2$ ). The results indicated the steric, hydrophobic and electrostatic fields play key roles in models. Molecular docking elucidated the conformations of compounds and key amino acid residues at the docking pocket of RT protein.

Received 00th January 2014,  
Accepted 00th January 2014

DOI: 10.1039/x0xx00000x

www.rsc.org/

## 1. Introduction

The HIV-1 infection causes acquired immunodeficiency syndrome that is a fatal human health-threatening disease.<sup>1</sup> The disease presents a serious health care challenge because each year it affects an increasing number of people.<sup>2</sup> Since the time when it was recognized 30 years ago, 30 million people have died of HIV-related causes.<sup>3</sup> An estimated 34 million people are living with HIV/AIDS worldwide.<sup>4,5</sup> Among them about 3.4 million children aged less than 15 years have been infected mainly through mother-to-children transmission.<sup>6</sup> In 2010, 1.8 million people died of AIDS related illness,<sup>7</sup> it has further reaching implications in terms of economic, social and political stability.<sup>8</sup>

In AIDS therapy, the fundamental strategy is to inhibit viral replication. Many studies have demonstrated that the viral reverse transcriptase of HIV was a multifunctional enzyme critical to the viral life cycle, and thus provided an attractive target in the search for anti-HIV therapies.<sup>9</sup> HIV was first isolated in 1983 by Luc Montagnier<sup>10</sup> and accepted as the causative agent of AIDS by 1984,<sup>11</sup> had a single strand RNA containing its genomic material and the viral enzyme reverse transcriptase enables the virus to utilise this single strand RNA genome as a template for the

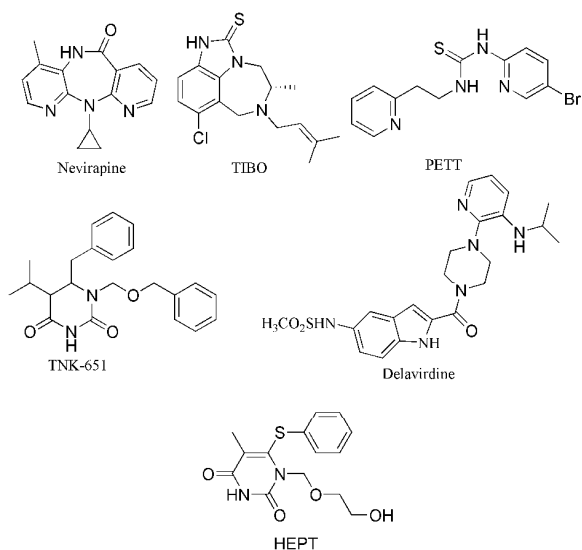
production of an intermediate single DNA strand, and ultimately double-stranded DNA.<sup>12</sup> Among the major types of HIV originated from primate species, HIV-1 which is more easily transmitted and accounts for the vast majority of global HIV infections.<sup>13,14</sup> The HIV-1 RT is an asymmetric heterodimer, comprising a p66 subunit (560 amino acids) and a p51 subunit (440 amino acids)<sup>15</sup> which has been used as an effective target for antiretroviral drugs.

To combat disease, several drugs which are currently in clinical use have been developed to inhibit this potential target, including Nevirapine,<sup>16</sup> TIBO,<sup>17</sup> PETT,<sup>18</sup> TNK-651,<sup>19</sup> Delavirdine<sup>20</sup> and HEPT<sup>21</sup> (Fig.1). NNRTIs inhibit RT by binding to the enzyme in a hydrophobic pocket located at a distance of around 10Å from its catalytic site.<sup>22</sup> However, their therapeutic effectiveness are limited by toxicity, unfavorable pharmacokinetics and emergence of resistant viral strains.<sup>23-25</sup> The ability of the virus to circumvent the inhibitors via mutation clearly signifies the need for improved therapeutic agents.<sup>26</sup> Therefore, there is still a need to identify new NNRTIs, eliciting intense efforts in the development of potent inhibitors. Fortunately, the dihydro-alkyloxy-benzyl-oxopyrimidines (DABOs)<sup>27-29</sup> derivatives were disclosed as non-nucleoside reverse transcriptase inhibitor (NNRTI)<sup>30,31</sup> by Artico et al group in 1992,<sup>33-34</sup> endowed with inhibitory potencies in the low

nanomolar range.

For saving resources and expedite the development of NNRTIs, computer simulation techniques used to offer further means or design effective inhibitors about RT and explore the inhibition mechanisms, recently. Among several computational methods, the quantitative structure-activity relationship (QSAR)<sup>35</sup> study is one of the most effective computational approaches in drug design without a doubt. The results of comparative molecular field analysis (CoMFA)<sup>36</sup> and the comparative molecular similarity indices analysis (CoMSIA)<sup>37</sup> studies not only used to predict the activity of newly inhibitors but provide beneficial information in structural modification. Up till now, it has been successfully applied in many biological and medicinal studies.

This method has been applied to facilitate the design of more specific and potent HIV-1 inhibitors. Such as, Horrick Sharma et al.<sup>38</sup> used 3D-QSAR method to study of 3-keto salicylic acid chalcones and related amides as novel HIV-1 integrase inhibitors, Hamid Abedi et al.<sup>39</sup> developed a 3D-QSAR model for new phenyloxazolidinones derivatives as potent HIV-1 protease inhibitors and so on. In this work, 3D-QSAR based on the ligand-based alignment rule, applied to determine the structural factors that affect the inhibitory activity of RT inhibitors. In addition, we used the second generation of CoMFA technique-Topomer CoMFA which generated a QSAR model by splitting the molecule into fragments, aligning each fragment and calculating steric and electrostatic field descriptor values for the topomerically aligned fragments. The contour maps got by models were used to support each other. Molecular docking was carried out to study the possible binding modes of inhibitors at the active site of HIV-1 protein. To the best of our knowledge, this is the first report on 3D-QSAR-CoMFA/CoMSIA/Topomer models for this series of compounds. Also, have not been reported simultaneously in the literature. The models and docking information derived, we hope, will be of great assistance in NNRTIs and accurate activity predictions for newly designed HIV-1 reverse transcriptase inhibitors in future.



**Fig.1** Chemical structures of Well-Known RT Inhibitors

## 2. Results and Discussion

### 2.1 Computational approach

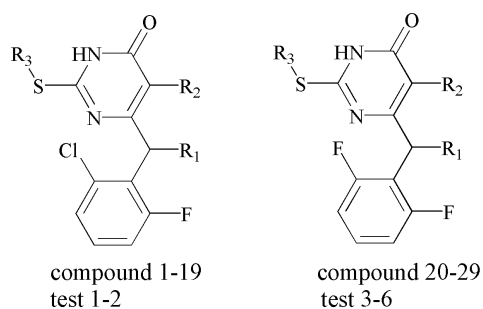
The 3D-QSAR, Topomer CoMFA modeling analyses, calculations, molecular docking simulations were performed using the SYBYL-X 2.0 package (Tripos Inc., St. Louis, USA) running on windows 7 workstation. All compounds (creation of training set, test set and docking calculations) involved in this study were optimized using Tripos force field and Gasteiger-Huckel charges. The structural energy minimization was terminated when using Powell gradient algorithm with a convergence criterion of 0.005 kcal/(mol·Å) reached and a maximum of 10000 iterations.

### 2.2 Data sets

The 35 compounds involved in this study which were reported recently by Dante Rotili.<sup>40</sup> Their capability to inhibit by 50% the HIV-1-induced cytopathic effect in MT-4 cells (HIV-1 Strain:NL4-3) were expressed as micromolar ( $\mu\text{M}$ ) units and converted to corresponding  $\text{pEC}_{50}$  values according to the following formula:

$$\text{pEC}_{50} = -\log \text{EC}_{50}$$

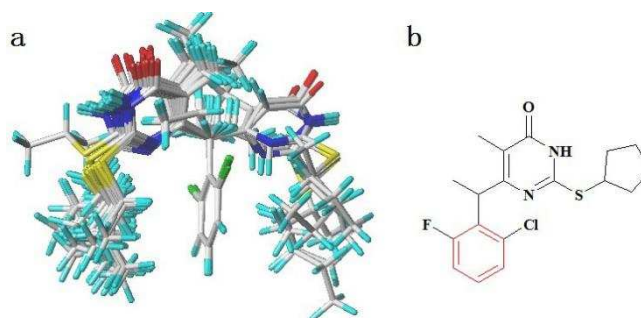
The samples were randomly divided into a training set of 29 compounds for model generation and a test set of 6 compounds for model validation. The data set was selected by considering both the distribution of biological data and structural diversity. The structures and Anti-HIV-1 activities expressed as  $\text{pEC}_{50}$  against HIV-1 are shown in Tables 1. In this study, the highest activity of compound 14 is taken as the template molecule and each molecule has to be superimposed onto it for the analysis. The alignment of training set is shown in Fig. 2a. As shown in Fig. 2b, the red atoms used to automatically position the compounds by using database alignment.

**Table 1** The structures, actual and predicted pEC<sub>50</sub> (μM) values of COMFA and CoMSIA.

No	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	Actual pEC <sub>50</sub>	Predicted COMFA	Residual	Predicted COMSIA	Residual
1	H	H	i-Pr	6.42	6.431	-0.011	6.419	0.001
2	H	H	s-Bu	6.80	6.752	0.048	6.776	0.024
3	H	H	n-Bu	6.54	6.583	-0.043	6.562	-0.022
4	H	Me	i-Pr	6.92	6.946	-0.026	6.905	0.015
5	H	Me	n-Pr	7.07	7.091	-0.021	7.149	-0.079
6	H	Me	s-Bu	7.33	7.252	0.078	7.272	0.058
7	H	Me	n-Bu	6.96	7.071	-0.111	7.000	-0.040
8	Me	H	i-Pr	8.70	8.952	-0.252	8.816	-0.116
9	Me	H	n-Pr	8.52	8.469	0.051	8.476	0.044
10	Me	H	s-Bu	9.00	9.025	-0.025	8.941	0.059
11	Me	H	cyclopentyl	9.52	9.278	0.242	9.522	-0.002
12	Me	Me	n-Pr	10.30	10.307	-0.007	10.207	0.093
13	Me	Me	s-Bu	9.52	9.473	0.047	9.560	-0.040
14	Me	Me	cyclopentyl	10.52	10.517	0.003	10.524	-0.004
15	Me	Et	n-Pr	7.55	7.536	0.014	7.527	0.023
16	Me	Et	n-Bu	7.19	7.098	0.092	7.159	0.031
17	Et	Me	n-Pr	7.59	7.497	0.093	7.605	-0.015
18	Et	Me	s-Bu	7.27	7.263	0.007	7.251	0.019
19	Et	Me	n-Bu	6.96	6.905	0.055	6.963	-0.003

20	Me	Me	i-Pr	9.00	8.903	0.097	8.985	0.015
21	Me	Me	n-Pr	7.55	7.453	0.097	7.625	-0.075
22	Me	Me	s-Bu	9.05	9.151	-0.101	9.112	-0.062
23	Me	Me	n-Bu	7.30	7.482	-0.182	7.302	-0.002
24	Me	Et	i-Pr	7.05	7.113	-0.063	7.005	0.045
25	Me	Et	n-Pr	6.62	6.586	0.034	6.592	0.028
26	Et	Me	i-Pr	8.40	8.581	-0.181	8.374	0.026
27	Et	Me	n-Pr	7.92	7.884	0.036	7.916	0.004
28	Et	Me	s-Bu	9.10	9.048	0.052	9.119	-0.019
29	Et	Me	cyclopentyl	7.96	7.984	-0.024	7.966	-0.006
test1 <sup>a</sup>	H	H	cyclopentyl	6.82	6.767	0.053	7.373	0.553
test2	H	Me	cyclopentyl	6.74	6.933	-0.193	6.828	-0.088
test3	Me	Me	cyclopentyl	9.00	8.771	0.229	8.835	0.165
test4	Me	Et	s-Bu	6.85	7.743	-0.893	7.614	-0.764
test5	Me	Et	n-Bu	7.64	6.792	0.848	7.121	0.519
test6	Me	Et	cyclopentyl	6.8	7.213	-0.413	7.577	-0.777

<sup>a</sup>test = test set compounds.

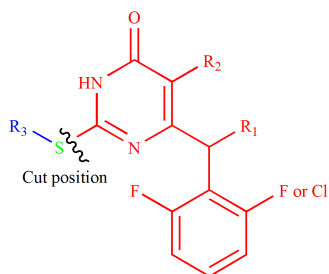


**Fig. 2** a) The alignment of training set compounds. b) The atoms used to automatically position the compounds by using database alignment (red).

Topomer CoMFA is the second generation of CoMFA. The study was carried out with the same training and test sets, each of the training set structure was broken into two sets of fragments shown as R1 (red) and R2 (blue) groups as shown in Fig. 3. The method provides a means for an alignment-independent 3D-QSAR approach, which providing the means for automated search for activity in fragment libraries. Each topomer fragment was applied

with topomer alignment to make a 3D invariant representation.<sup>41</sup> Steric and electrostatic interaction energies were calculated using the carbon  $sp^3$  probe.

## ARTICLE



**Fig. 3** R1 fragment is represented by the red color and R2 fragment is represented by sky blue color.

### 2.3 CoMFA, CoMSIA and Topomer CoMFA statistical results

To generate statistically significant 3D-QSAR models, we used the ligand-based alignment rule. The alignment result is shown in Figure 2. In this study, the regression analysis was carried out using the partial least squares (PLS) method,<sup>42,43</sup> some statistical parameters were used to analysis the stand or fall of these models, including the cross-validated coefficient ( $q^2$ ), the standard deviation

of error prediction ( $r^2$ ), standard error of estimate (SEE) and F-statistic values, a high  $q^2$  and  $r^2$  value ( $q^2 > 0.5$ ,  $r^2 > 0.6$ ) is considered as a proof of high predictive ability of the model.<sup>44</sup> Table 2 lists the statistical parameter results of the CoMFA and CoMSIA analyses. For CoMFA analysis, steric and electrostatic fields were considered, which gave an optimal number of components (ONC) of 8 and  $q^2$  of 0.636. The standard deviation of error prediction  $r^2$ , SEE and F-statistic values are 0.993, 0.118 and 333.5, respectively. The contributions of the steric and electrostatic fields are 58.5% and 41.5%. For CoMSIA analysis, five descriptor fields were considered. The CoMSIA model obtained satisfactory  $q^2$ ,  $r^2$ , SEE and F-statistic values of 0.655, 0.998, 0.069 and 488.2, respectively. The steric, electrostatic, hydrophobic, hydrogen bond donor, and hydrogen bond acceptor fields contribution of 22.1%, 19.9%, 24.5%, 18.7% and 14.8%, respectively. The steric, electrostatic, hydrophobic fields were found to be the important contributions in the optimal CoMSIA model. All the statistical results of different fields summarized in Table 2. The actual and predicted pEC50 value of the training and test set molecules were illustrated in Fig. 4, respectively.

Table 2 Summary of the COMFA/COMSIA PLS statistical results

	$q^2$	ONC	$r^2$	SEE	F	$r_{\text{pred}}^2$	Field contribution(%) <sup>a</sup>				
							S	E	H	D	A
COMFA											
S+E	0.636	8	0.993	0.118	333.5	0.907	58.5	41.5	-	-	-
COMSIA											
H+A	0.604	11	0.992	0.129	202.0	0.964	-	-	65.3	-	34.7
S+A	0.678	11	0.987	0.170	115.6	0.910	63.2	-	-	-	36.8
H+D	0.742	12	0.995	0.107	268.7	0.844	-	-	64.5	35.5	-
S+H	0.814	9	0.991	0.133	232.8	0.864	44.2	-	55.8	-	-
S+E+H	0.697	14	0.998	0.072	517.2	0.820	29.9	35.7	34.4	-	-
S+H+D	0.842	12	0.996	0.100	308.8	0.828	30.6	-	40.8	28.6	-
S+H+A	0.757	11	0.996	0.098	353.7	0.913	33.3	-	42.3	-	24.4
S+E+H+D	0.711	16	0.999	0.060	643.2	0.826	25.2	25.7	29.4	19.8	-
<b>S+E+H+A+D</b>	<b>0.655</b>	16	<b>0.998</b>	0.069	488.2	<b>0.886</b>	22.1	19.9	24.5	18.7	14.8



<sup>a</sup>Abbreviations: COMFA and COMSIA with different field contributions such as *S* (steric); *E* (electrostatic); *H* (hydrophobic); *D* (H-bond donor); *A* (H-bond acceptor).

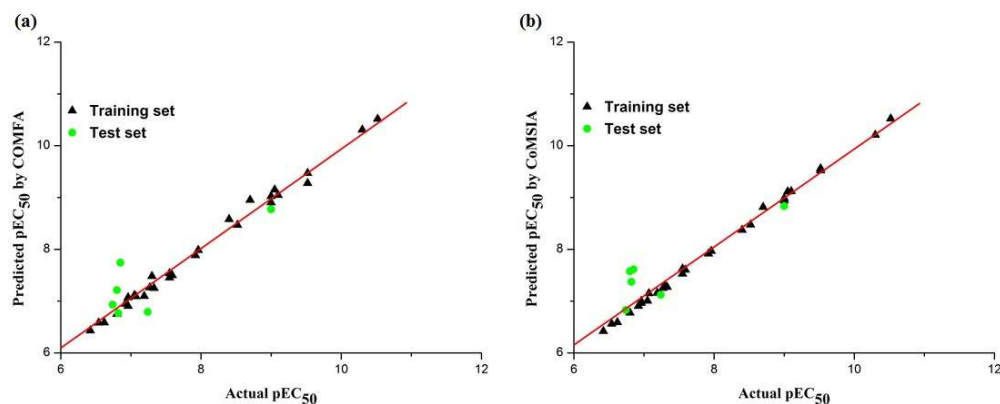


Fig. 4 Plots of predicted versus actual pEC<sub>50</sub> values for all the molecules based on CoMFA (a) and CoMSIA models (b)

In Topomer CoMFA analysis, all the models were investigated using the full cross validated ( $q^2$ ) PLS leave-one-out (LOO) method with CoMFA standard options for scaling of the variables. The model reflects the quantitative relationship between the structure and activity. The  $q^2$  value of 0.546, an optimized component of 6 and  $r^2$  value of 0.798, which suggested the model also has predictive ability ( $q^2 > 0.2$ ). The pEC<sub>50</sub> value of test set was predicted with the  $r^2_{\text{pred}}$  value of 0.718.

#### 2.4 Contour maps of CoMFA, CoMSIA and Topomer CoMFA

To visualize the field effects on the target compound in 3D space, the model results were graphically interpreted by using field type “StDev\*Coeff”. For choosing the best biological activity against HIV-1 cells which makes it easier to explain the contour maps. Hence using compound 14 as a reference structure to illustrate all contour maps of the models, explaining the key structural features required for inhibitory activity.

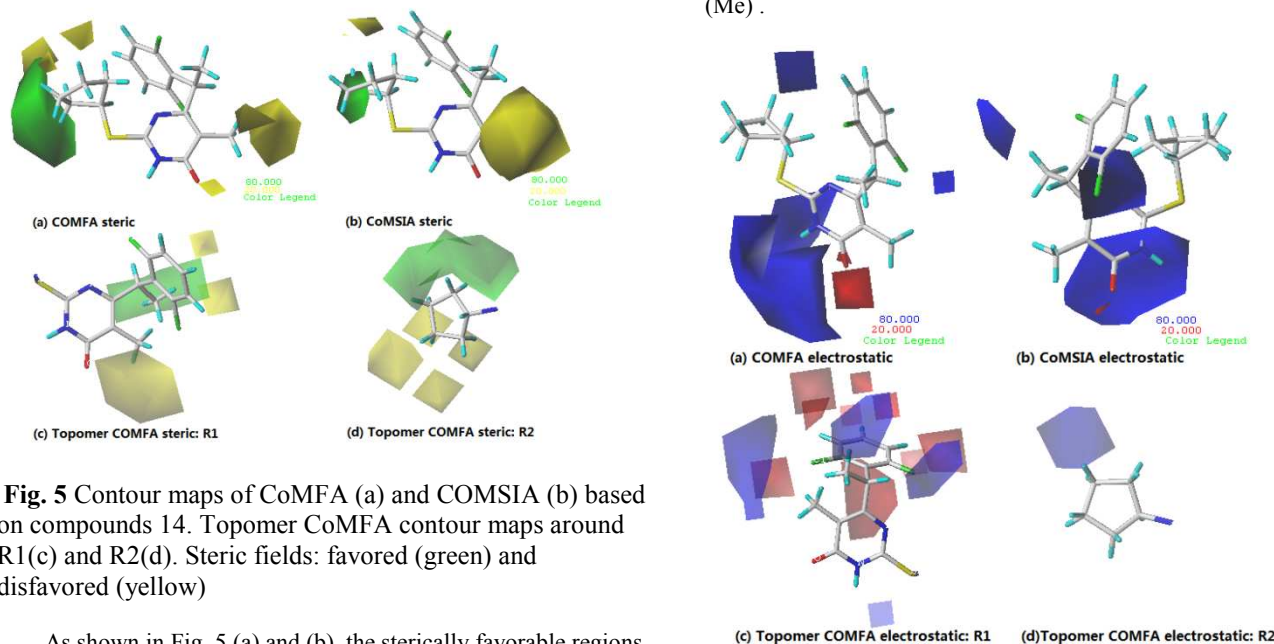


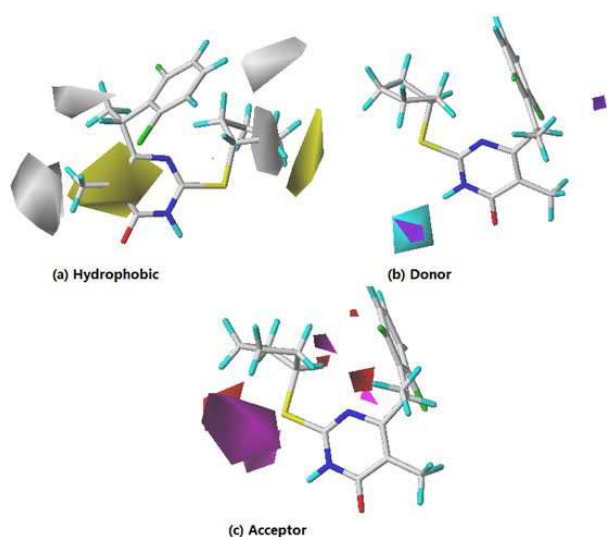
Fig. 5 Contour maps of CoMFA (a) and CoMSIA (b) based on compounds 14. Topomer CoMFA contour maps around R1 (c) and R2 (d). Steric fields: favored (green) and disfavored (yellow)

As shown in Fig. 5 (a) and (b), the sterically favorable regions

are represented in green and yellow contours suggest that the bulky substituents would not be tolerated. The default values of 80% contribution for favored and 20% for disfavored regions were set for visualization of the contour maps. It can be observed that the steric contour map of CoMFA is similar to that of CoMSIA. The topomer CoMFA 3D interaction maps (steric and electrostatic interactions) around R1 and R2 are shown in Figures 5 (c) and (d), respectively. The meaning of colors is the same as the CoMFA's. The steric descriptor is found to be identical with the CoMFA and CoMSIA models, which prove the consistency of the results. A green contour covering the cyclopentyl group links to R<sub>3</sub> indicates the presence of a bulky group for good biological activity. This is in agreement with the experimental data. The order of inhibitory activity is: isomeric (i) > normal (n) and secondary (s) > normal (n). For example, 2 (s-Bu) > 3 (n-Bu), 6 (s-Bu) > 7 (n-Bu), 8 (i-Pr) > 9 (n-Pr), 18 (s-Bu) > 19 (n-Bu), 20 (i-Pr) > 21 (n-Pr), 22 (s-Bu) > 23 (n-Bu), 24 (i-Pr) > 25 (n-Pr), 26 (i-Pr) > 27 (n-Pr). The yellow contour surrounding the R<sub>2</sub> position indicates that compound with bulky substitution could not possess better biological activity as observed in 15 (Et) < 12 (Me), 24 (Et) < 20 (Me) and 25 (Et) < 21 (Me).

**Fig. 6** Contour maps of CoMFA (a) and COMSIA (b) based on compounds 14. Topomer CoMFA contour maps around R1 (c) and R2 (d). Electrostatic fields: electropositive (blue) and electronegative (red)

In figure 6(a), a large blue color contour around the pyrimidine group indicates that the electropositive groups are favorable to the activity, the red region around O atom at the pyrimidine ring suggests the electronegative groups are favorable to the activity. As shown in Fig 6(b), two pieces of medium-sized region of blue contour locate at the pyrimidine group shows the importance of electropositive atoms in imparting better biological activity. The result of Topomer CoMFA (electrostatic contour map) is almost similar to that of the CoMFA and CoMSIA model, the analysis of the electrostatic contour map is neglected herein. Moreover, it is recognized that all the models have a blue contour region around R<sub>3</sub> position, which indicates that with electropositive groups lead to a increase of bioactivity.



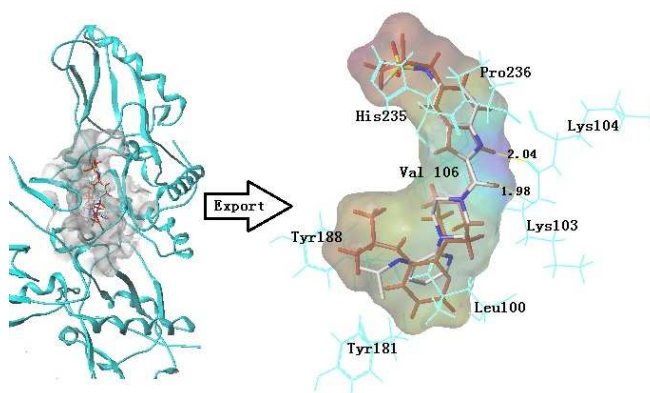
**Fig. 7** Contour maps of COMSIA based on compounds 14. (a) Hydrophobic field: the yellow color shows the favored hydrophobic area, the white color shows the disfavored hydrophobic area. (b) Hydrogen bond donor field: the cyan color shows the favored H-donor area, and the purple color represents the disfavored H-donor area. (c) Hydrogen bond acceptor field: the magenta color shows the favored H-acceptor area, the red color shows the disfavored H-acceptor area,

As shown in Table 2, the hydrophobic field found to be the most important contributions in the optimal CoMSIA model. It plays important role in bioactivity than other fields. From figure 7a, one large yellow (hydrophobic favorable) contour map cross the R<sub>2</sub> position, suggesting introduction of hydrophobic substituents into the ring pyrimidine will be benefit for inhibitory activity. This is in agreement with the experimental data: 4(-CH<sub>3</sub>) > 1(H), 6(-CH<sub>3</sub>) > 2(H), 7(-CH<sub>3</sub>) > 3(H), 12(-CH<sub>3</sub>) > 9(H), 13(-CH<sub>3</sub>) > 10(H) and 14(-CH<sub>3</sub>) > 11(H). Besides, a small yellow contour map observed farther from R<sub>3</sub>, indicating the hydrophobic field has little effect on inhibitory activity.

One small white (hydrophilic favorable) contour map around

the ring cyclopentyl (R<sub>3</sub>), suggesting that introduction of hydrophilic substituents into the ring cyclopentyl may be benefit for inhibitory activity. Since the other white contour maps farther from compound 14, the analysis of these contour maps are neglected herein. As shown in Fig 7b and 7c, the purple contour represents where hydrogen bond donor disfavors the biological activity and the cyan contour shows the favored hydrogen bond donor area, in which the magenta color shows the hydrogen bond acceptor group would be beneficial to the bioactivity, whereas the red color shows the disfavored H-acceptor area.

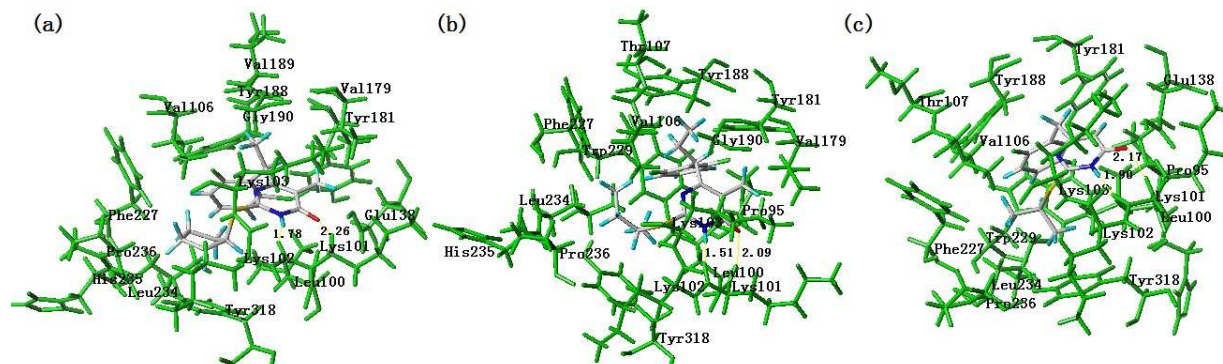
## 2.5 Docking Analysis



**Fig. 8** Re-docking result of the bis (heteroaryl) piperazine (BHAP) U-90152 ( Red color ligand) into the binding site of HIV-1 reverse transcriptase protease and the cognate ligand ( Atom type color ) from the protein ( PDB:1KTS). Hydrogen bonds are shown as yellow lines, with distance unit of Å. The inhibitor and the important residues are shown as stick model.

To validate the 3D-QSAR results, docking simulation was performed to study the binding environment. Here, the Surflex program (SYBYL-X 2.0, Tripos Inc.St. Louis, USA) was used to explore the probable binding conformation. Initially, the performance of docking software was tested by re-docking experiment. For this purpose, crystal structure of protein (PDB code:1KLM) with the cognate ligand was re-docked. As is shown in Fig.8, it can be recognized that the re-docking result and the cognate ligand are almost completely superimposed together. Their rotational tendency are basically similar. Meanwhile, it shows interacting modes of U-90152 in the binding site of RT receptor, some key amino acids (Lys104, Tyr188, Tyr181, His235, Val106, Leu100 and Pro236) interact with the inhibitor by hydrophobic interactions at the binding site. The key residue Lys103 interact with the inhibitor by hydrogen bond. The hydrogen bond distances observed are 2.04 Å(Lys103-HN-H...O-), 1.98 Å (Lys103-O...H-NH-), respectively. Two of these interactions are similar to the results discussed by Robert et al.<sup>45</sup> Both results can be inferred that the rationality of the program, and the docking result is reliable.





**Fig. 9** Docking result of the compound 14 (a) , 28 (b) and 1 (c) into the binding site of HIV-1 reverse transcriptase protease. Hydrogen bonds are shown as yellow lines, with distance unit of Å. The inhibitor and the important residues are shown as stick model.

After validating the docking reliability, Surflex-Dock was used for docking. Herein the compounds were originally reported to be the most potent inhibitors toward HIV-1 and least active compound were selected for more detailed analysis. Fig. 9 shows the interacting mode of compounds 14, 28 and 1 in the binding site of RT receptor. As shown in Fig. 9a, compound 14 was docked into the binding cavity with the carboxyl directing towards the hydrophobic group of Leu100, Lys102, Val106, Val179, Tyr181, Tyr188, Val189, Gly190, Phe227, Leu234, His235, Pro236 and Tyr318. One hydrogen bond was formed between the pyrimidine ring and the Lys101 residue. The hydrogen bond distances observed are 2.26 Å (Lys101-HN-H...O-), 1.78 Å (Lys101-O...H-NH-), respectively. Robert et al. revealed that hydrogen bonding to the main chain of Lys-103 and strong hydrophobic interactions are particularly important for any successful inhibitors. But many NNIs, including HEPT and TIBO analogues,<sup>45</sup> also forming a hydrogen bond to the carbonyl oxygen of Lys101. The Glu138 and Lys103 residues interact with 14 by electrostatic interaction. The presence of these interactions affect the nearby protein structure.

As shown in Fig.9b, compound 28 was docked into the binding cavity with the carboxyl directing towards the hydrophobic group of Pro95, Leu100, Val106, Thr107, Val179, Tyr181, Tyr188, Gly190, Phe227, Trp29, Leu234, His235, Pro236 and Tyr318. One hydrogen bond was formed between the pyrimidine ring and the Lys101 residue. The distances observed are 1.51 Å (Lys101-HN-H...O-), 2.09 Å (Lys101-O...H-NH-), respectively. The residues Lys102 and Lys103 interact with 28 by electrostatic interaction. Fig. 9c depicts the interactions between the compound 1 and the active site. The ligand was anchored into the binding site via one hydrogen bond to the key residues of Lys101. The distances observed are 2.17 Å (Lys101-HN-H...O-), 1.90 Å (Lys101-O...H-NH-), respectively. Moreover, the compound was stabilized by the hydrophobic interactions offered by Leu100, Pro95, Val106, Lys102, Val179, Tyr181, Tyr188, Trp229, Leu234, Phe227 and Tyr318. Compared with the compounds 14 and 28, the number of hydrophobic group around compound 1 obviously decreased. The key residues Pro236 and His235 can not be found in the binding pocket, simultaneously.

### 3. Conclusion

In the present work, a series of DABOs derivatives as Non-nucleoside inhibitors were studied by computer-aided drug design processes, such as 3D-QSAR/Topomer CoMFA studies and molecular docking simulations. The built models are favored by internal and external predictions and the statics are convincing and

comparable. The models can not only be extrapolated to predict novel and more potent inhibitors, but the contour maps obtained from CoMFA, CoMSIA and Topomer CoMFA analyses provide a useful insight for structure-based design for designing new chemical entities with high HIV-1 inhibitory activity. For a better understanding of the binding modes of inhibitors at the active site of HIV-1 protein, molecular docking analyses of the representative compounds were performed. Some key residues such as Leu100, Lys102, Val106, Tyr181, Tyr188, His235, Pro236 hydrophobic interactions, as well as hydrogen bonds (Lys101, Lys103) between inhibitors and the active site were observed. The inhibitors interact with the main residues (Lys101 and 103), which may generate binding affinity that is less susceptible to point mutations. This study could serve as a basis for the development of HIV-1 inhibitors.

### Acknowledgements

This work is supported by the Science and Technology Commission of Shanghai Municipality (No. 13142201001, 13DZ1930402, 13DZ1930403).

### Notes and references

<sup>a</sup> School of Chemical and Environmental Engineering, Shanghai Institute of Technology, Shanghai, P. R. China. E-mail: clab@sit.edu.cn

- Xu Y, Liu H, Niu C.Y, Luo C, Luo X.M, Shen J.H, Chen K.X, Jiang H.L, Bioorgan. Med. Chem, 2004, 12, 6193–6208.
- Joseph R, Florent B, Catia T, Francois M, J Comput Aided Mol. Des, 2008, 22, 831–841.
- www.unaids.org. Accessed 3 February 2012
- Soliman, M.E.S, Drug Dev. Res, 2013, 74, 283–295.
- Johnson B.C, Pauly G.T, Rai G, Patel D, Bauman J.D, Baker H.L, Das K, Schneider J.P, Maloney D.J, Arnold E.; et al., Retrovirology, 2012, 9, 1–23
- Hamid Abedi, Homeira Ebrahimzadeh, Jahan B. Ghasemi, Struct Chem, 2013, 24, 433–444. DOI 10.1007/s11224-012-0092-1
- http://www.who.int/hiv/mediacentre/news64/. Accessed 3 February 2012
- R. C. Gallo and L. Montagnier, Science, 2002, 298, 1730–1731.
- Robert M. Esnouf, Jingshan Ren, Andrew L. Hopkins, Carl K. Ross, E. Yvonne Jones, David K. Stammers, David I. Stuart, Proc. Natl. Acad. Sci, 1997, 94, 3984–3989.
- R. C. Gallo and L. Montagnier, N. Engl. J. Med., 2003, 349, 2283–2285.

- 11 R. C. Gallo, *Science*, 2002, 298, 1728–1730
- 12 Chevonne Reynolds, Charles B. de Koning, Stephen C. Pelly, William A. L. van Otterlo and Moira L. Bode, *Chem. Soc. Rev.*, 2012, 41, 465–467
- 13 R. C. Gallo, *Retrovirology*, 2006, 3, 72–78.
- 14 M. Ndour, P. S. Sow, A. M. Coll-Seck, S. Badiane, C. T. Ndour, N. Diakhate, B. Diop, M. Faye, M. Soumare, G. Diouf and R. Colebunders, *Trop. Med. Int. Health*, 2000, 5, 687–691.
- 15 Kohlstaedt, L.A., Wang, J., Friedman, J.M., Rice, P.A., Steitz, T.A., *Science*, 1992, 256, 1783–1790.
- 16 Merluzzi V. J, Hargrave K. D, Labadia M, Grozinger K, Skoog M, Wu J. C, Shih C.-K, Eckner K, Hattox S, Adams J, Rosenthal A. S, Faanes R, Eckner R. J and Koup R. A, et al., *Science* 1990, 250, 1411–1413.
- 17 Pauwels R, Andries K, Desmyter J, et al., *Nature* 1990, 343, 470–474.
- 18 Cantrell A. S, Engelhardt P, Hogberg M, et al., *J. Med. Chem*, 1996, 39, 4261–4274.
- 19 Hopkins A. L, Ren J, Eshouf R. M, et al., *J. Med. Chem*, 1996, 39, 1589–1600.
- 20 Dueweke T.J., Poppe S.M., Romero D.L., et al., *Antimicrob Agents Chemother*. 1993, 37, 1127–1131.
- 21 Baba M., Tanaka H., De Clercq E., et al., *Biochem. Biophys. Res. Commun*, 1989, 165, 1375–1381.
- 22 Marie-Pierre de Béthune, *Antiviral Research*, 2010, 85, 75–90.
- 23 Richman D. D., Fischl M. A., Grieco M. H., et al., *N. Engl. J. Med*, 1987, 317, 192–197.
- 24 Richman D. D, *Antimicrob. Agents Chemother*, 1993, 37, 1207–1221.
- 25 Terasaki T., Pardridge W. M, *J. Infect. Dis*, 1988, 158, 630–632.
- 26 Vijay M, Khedkar Premlata K, Ambre Jitender Verma Mushtaque S, et al., *J Mol Model*, 2010, 16, 1251–1268.
- 27 Artico M, *Drugs Future*, 2002, 27, 159–175.
- 28 Yang S., Chen F. E., De Clercq E, *Curr. Med. Chem*, 2012, 19, 152–162.
- 29 Novakov I. A., Artico M., Mai A., et al., *Russ. Chem. Bull*, 2012, 61, 1399–1418.
- 30 Jayaweera D., Dilanchian P, *Expert Opin. Pharmacother*, 2012, 13, 2601–2612.
- 31 Usach I., Melis V., Peris J. E., *J. Int. AIDS Soc*, 2013, 16, 1–14.
- 32 Botta M., Artico M., Massa S., et al., *Eur. J. Med. Chem*, 1992, 27, 251–257.
- 33 Artico M., Massa S., Mai A., et al., *Antivir. Chem. Chemother*, 1993, 4, 361–368.
- 34 Massa S., Mai A., Artico M., et al., *Antivir. Chem. Chemother*, 1995, 6, 1–8.
- 35 Van de Waterbeemd H, *Chemometric Methods in Drug Design*. VCH, Weinheim, 1995, chapter 1.
- 36 Cramer RD, Patterson DE, Bunce JD, *J Am Chem Soc*, 1988, 110, 5959–5967
- 37 Klebe G, Abraham U, Mietzner T, *J Med Chem*, 1994, 37:4130–4146
- 38 Horrick Sharma, Shivaputra Patil, Tino W. Sanchez, Nouri Neamati, Raymond F. Schinazi, John K. Buolamwini, *Bioorganic Medicinal Chemistry*, 2011, 6, 2030–2045.
- 39 Hamid Abedi, Homeira Ebrahimzadeh, Jahan B. Ghasemi, *Structural Chemistry*, 2013, 2, 433–444.
- 40 Dante Rotili, Domenico Tarantino, Maxim B. Nawrozkij, et al., *J. Med. Chem*, 2014, 57, 5212–5225.
- 41 Cramer RD III, Clark RD, Patterson DE, Ferguson AM, *J Med Chem*, 1996, 39, 3060–3069
- 42 Stähle L, Wold S, *J Chemom*, 1987, 1, 185.
- 43 Wold S., Albano C., Dunn W. J., et al., *Chemometrics*, 1984, 138, 17.
- 44 P.P. Roy, J.T. Leonard, K. Roy, *Chemom. Intell. Lab. Syst*, 2008, 90, 31–42.
- 45 Robert M. Eshouf, Jingshan Ren, Andrew L. Hopkins, Carl K. Ross, et al., *Proc. Natl. Acad. Sci*, 1997, 94, 3984–3989.