3D-Quantitative Structure−Activity Relationships of HEPT Derivatives as HIV-1 Reverse Transcriptase Inhibitors, Based on Ab Initio Calculations

Supa Hannongbua,*,† Kanda Nivesanond,† Luckhana Lawtrakul,‡ Pornpan Pungpo,† and Peter Wolschann‡

Department of Chemistry, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand, and Institut für Theoretische Chemie und Molekulare Strukturbiologie der Universität Wien, Währinger Strasse 17, A-1090, Wien, Austria

Received August 29, 2000

Comparative molecular field analysis (CoMFA) has been applied to a large set of 1-[(2-hydroxyethoxy)-methyl]-6-(phenylthio)thymine (HEPT) analogues. The starting geometry of HEPT was obtained from crystallographic data of HEPT/HIV-1 reverse transcriptase (RT) complexes. The structures of 101 HEPT derivatives were considered and fully optimized by ab initio molecular orbital calculations at the HF/3-21G level. The best CoMFA model is satisfactory in both statistical significance and predictive ability. It shows excellent, high predictive ability as $r^2_{cv} = 0.858$. The derived model indicates the importance of steric contributions (64.4%) as well as electrostatic interactions for the HIV-1 RT inhibition. In addition, steric and electrostatic contour maps from this analysis agree well with the experimentally observed trend that there are steric interactions between the side chain of HEPT and an aromatic ring of Tyr181. It is concluded that a moderately sized group at C5 enhances contact with Tyr181 enough to push it into a position which renders the protein nonfunctional, but a smaller group has insufficient steric requirements to do this and a larger group renders the ligand too large for the cavity. The mutation-induced resistance of reverse transcriptase is explained by this analysis. The obtained results not only lead to a better understanding of structural requirements of this set of compounds for the inhibition but also enable the suggestions for new and more potent drugs.

INTRODUCTION

HIV-1 reverse transcriptase (HIV-1 RT) is an attractive target for the drug therapy of AIDS because it is essential for HIV replication and it is not required for normal host cell replication. This enzyme is a multifunctional enzyme that copies the RNA genome of HIV-1 into DNA, which is subsequently integrated into the host cell's genome. The non-nucleoside reverse transcriptase inhibitors (NNRTIs), for example, HEPT, Nevirapine, TIBO, and α-APA,[1−4] are specific against HIV-1 and do not inhibit host cell polymerases. These NNRTIs have low cytotoxicity and produce few side effects. The NNRTIs binding site is located in the p66 palm subdomain of RT and constitutes mainly aromatic hydrophobic residues. Recently, the study of complex structures of HIV-1 RT with different NNRTIs[5,6] reveals that there is a significant rearrangement of a three-stranded β-sheet in the p66 subunit with respect to the rest of the polymerase, and NNRTIs inhibit HIV-1 RT by locking the polymerase active site into an inactive conformation. However, a serious problem with the NNRTIs is the emergence of viral strains that have point mutations in the region encoding HIV-1 RT which prevent these drugs from inhibiting RT.

Among the representatives of the NNRTIs, 1-[(2-hydroxyethoxy)methyl]-6-phenylthiothymine (HEPT) constitutes an important inhibitor and has been extensively studied for many years.[7−21] There is considerable interest in the set up of quantitative structure−activity relationships (QSAR) of HEPT derivatives.[8−15] These methods are widely accepted to be useful for the explanation of structural requirements of biologically active compounds.[22] In particular, 3D-QSAR provides a possibility to gain more insight into drug/receptor interactions with respect to steric and electrostatic interactions.[23] However, the mechanisms of inhibition and the mutation of HIV-1 RT have not been clarified until now.

In our previous studies, QSAR and 3D-QSAR, based on comparative molecular field analysis (CoMFA), were applied to various classes of NNRTIs.[8,9,13,15,24,25] As X-ray crystal structure of RT and the structures of the inhibition complexes are available, this information can be applied in various molecular modeling studies to develop new and more potent compounds. Moreover, the combination of X-ray investigations on the complexes with CoMFA studies allows a much more accurate interpretation of the CoMFA contour maps, it leads also to a better knowledge of the enzyme-ligand interaction. Also the decrease or the increase of inhibitor activity against a mutant enzyme contributes to the understanding of the inhibition mechanism.[24,25]

In the present work, a CoMFA study has been performed to determine quantitative structure−activity relationships, in particular to investigate steric and electrostatic interactions

* Corresponding author phone: +66-(2)-9428900; fax: +66-(2)-5793955; e-mail: fscisph@ku.ac.th.
† Kasetsart University.
‡ Institut für Theoretische Chemie und Molekulare Strukturbiologie der Universität Wien.

10.1021/ci0001278 CCC: $20.00
Published on Web 04/03/2001

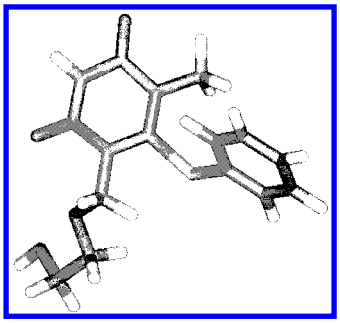


Figure 1. Structure of 1-[(2-hydroxyethoxy)methyl]-6-(phenylthio)thymine (HEPT), obtained from complexed structure between HEPT and HIV-1 reverse transcriptase.

of HEPT analogues, with a much more extended set of HEPT derivatives (Figure 1) with widely different potencies.[18−21] A more accurate method (ab initio (HF/3-21G)), compared to previous published work,[9] has been used for the calculations of the molecular geometries and the charges. The general aim of the work is to obtain more information about the structural requirements underlying the inhibition of NNRTI in comparison to the results of various X-ray investigations.

METHODS OF CALCULATION

Biological Data. The HEPT analogues are separated into two groups; 82 compounds served for the training set and 19 compounds for the test set, sampling from various ranges of logarithm unit of activities. All chemical structures of these two sets are depicted in Tables 1 and 2, respectively. Their biological activities are also presented in the tables expressed as log(1/C), where C is the effective concentration of the compound required to achieve 50% (EC_{50}) protection of MT-4 cell against the cytopathic effect of HIV-1. These inhibitory data were taken from refs 18−21, where they are described to be determined under the same experimental conditions.

Molecular Modeling and Quantum Chemical Calculations. The starting geometry of HEPT was obtained from crystallographic data of the enzyme−inhibitor complex structure.[26] All derivatives were constructed and modified by ALCHEMY2000.[27] Then, full geometrical optimization, based on ab initio (HF/3-21G), of all structures was carried out by automated geometry optimization to either minima or saddle points using redundant internal coordinates. The calculations were performed by the GAUSSIAN94 software[28] on a cluster of Digital Alpha Servers (21004/275).

Alignment Rule and CoMFA Setup. A critical step in the construction of the CoMFA model is the attainment of an alignment rule. All HEPT compounds, corresponding to minimum energy conformers, were used in this analysis. The rigid thymine rings, common for all investigated compounds, were aligned on a HEPT template structure by the atom matching method, available with SYBYL 6.5.[29] To calculate the electrostatic interaction, partial atomic charges were obtained from GAUSSIAN. A grid spacing with 2 Å was used to generate a cubic lattice around all molecules based on the molecular volume of the structures. These dimensions ensured that the grid extended beyond the molecular size by 4.0 Å in all directions. Molecular interactions between probe atoms and aligned molecules were then calculated. Concerning the variation of the probe atoms, three types of different probe atoms, i.e., sp^3 carbon atom with +1 charge (default probe atom in SYBYL), sp^3 oxygen atom with −1 charge, and H atom with +1 charge, were used. The probe atom was placed at each lattice point, and the interactions of the steric and electrostatic fields with each atom in the molecule were calculated with CoMFA standard scaling and then put into a CoMFA QSAR table. The minimum sigma and energy cutoff values were set to 2.0 and 30 kcal/mol, respectively, for both electrostatic and steric fields.

To obtain CoMFA models, a partial least squares (PLS) method was used to perform the correlation between steric and electrostatic properties and the inhibitory activities. The orthogonal latent variables were extracted by the NIPALS-algorithm[30] and subjected to full cross-validation (leave-one-out method). The number of components (noc) at which the difference of the r^2_{cv} value to the next one was less than 0.05[31] was used during the analyses. Subsequently, a non-cross-validated analysis was employed to analyze the CoMFA results, using the optimal number of components previously identified.

Predictive Ability. The overall predictive ability of the analysis was expressed in terms of r^2_{cv}, in CoMFA, respectively, and is defined as

$$r^2_{cv} = (\mathrm{SSY} - \mathrm{PRESS})/\mathrm{SSY}$$

SSY stands for the variance of the biological activities around the mean value, and PRESS is the prediction error sum of squares, derived from the leave-one-out method. The uncertainty of the prediction is defined as

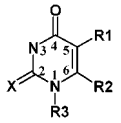
$$S_{\mathrm{PRESS}} = [\mathrm{PRESS}/(n-k-1)]^{1/2}$$

where k is the number of variables in the model and n is the number of compounds used in the study.

RESULTS AND DISCUSSION

CoMFA Results. All obtained analyses included both field types, i.e., steric and electrostatic fields. CoMFA with default setting probe atom (sp^3C) yielded the best CoMFA model as shown in Table 3 (model 1, $r^2_{cv} = 0.837$). Separate analysis of steric or electrostatic field types only were also performed, resulting in models 2 and 3. Both models are acceptable in both statistical criteria and predictive ability as r^2_{cv} is greater than 0.6 (model 2, $r^2_{cv} = 0.771$ and model 3, $r^2_{cv} = 0.726$). However, the predictive abilities are not as good as in model 1. Considering other models, derived by different probe atoms, sp^3O (−1) and H (+1), slightly lower predictive abilities are obtained (models 4−9). Model 1 expresses a higher level of internal consistency, but compound 36 is significantly out of line (Table 1). This compound was eliminated, and then models 10−12 were obtained. However, the residual is still in the same range as calculated for model 1. Then, a consecutive outlier, compound 50, was removed, resulting in models 13−15. The elimination of more compounds than these two outliers cannot improve the model; therefore, the final model, model 13, is satisfactory in both predictive ability ($r^2_{cv} = 0.858$) and the number of outliers. The contribution of the steric interactions is larger, as indicated by 64.4% of the total contribution. Substantially, other statistical results of the best CoMFA model (model 13) are the conventional r^2 (r^2 =

Table 1. Structure of HEPT Derivatives and Experimental and Calculated log(1/C) HIV-1 RT Inhibitory Affinities of the Training Set, Based on CoMFA Analysis Model 1



	substituent				log(1/C)		
no.	X	R1	R2	R3	exptl	calcd	residual
1	O	Me	SC_6H_{11}	$CH_2OCH_2CH_2OH$	5.085	5.569	−0.484
2	O	Me	CH_2Ph	$CH_2OCH_2CH_2OH$	4.637	5.580	−0.943
3	O	Me	C≡CPh	$CH_2OCH_2CH_2OH$	4.853	5.277	−0.424
4	O	I	SPh	$CH_2OCH_2CH_2OH$	5.443	5.272	0.171
5	O	SPh	SPh	$CH_2OCH_2CH_2OH$	4.677	4.507	0.170
6	O	$COCHMe_2$	SPh	$CH_2OCH_2CH_2OH$	4.920	5.704	−0.784
7	O	COPh	SPh	$CH_2OCH_2CH_2OH$	4.885	5.407	−0.522
8	O	CH_2Ph	SPh	$CH_2OCH_2CH_2OH$	4.637	4.808	−0.171
9	O	$CH{=}CPh_2$	SPh	$CH_2OCH_2CH_2OH$	6.075	5.647	0.428
10	O	C=CMe	SPh	$CH_2OCH_2CH_2OH$	4.720	5.592	−0.872
11	O	C=CPh	SPh	$CH_2OCH_2CH_2OH$	5.468	4.752	0.716
12	O	CH=CHPh (cis)	SPh	$CH_2OCH_2CH_2OH$	5.221	5.545	−0.324
13	O	$CH{=}CH_2$	SPh	$CH_2OCH_2CH_2OH$	5.958	4.799	1.159
14	O	Me	SPh(2-Me)	$CH_2OCH_2CH_2OH$	4.148	4.864	−0.716
15	O	Me	SPh(2-OME)	$CH_2OCH_2CH_2OH$	4.720	5.265	−0.545
16	O	Me	SPh(3-Me)	$CH_2OCH_2CH_2OH$	5.584	5.482	0.102
17	O	Me	SPh(3-Et)	$CH_2OCH_2CH_2OH$	5.568	5.384	0.184
18	O	Me	SPh(3-*t*-Bu)	$CH_2OCH_2CH_2OH$	4.920	5.376	−0.456
19	O	Me	SPh(3-CF_3)	$CH_2OCH_2CH_2OH$	3.460	4.667	−1.207
20	O	Me	SPh(3-Cl)	$CH_2OCH_2CH_2OH$	4.885	5.086	−0.201
21	O	Me	SPh(3-Br)	$CH_2OCH_2CH_2OH$	5.243	5.367	−0.124
22	O	Me	SPh(3-I)	$CH_2OCH_2CH_2OH$	4.999	5.552	−0.553
23	O	Me	SPh(3-NO_2)	$CH_2OCH_2CH_2OH$	4.468	5.207	−0.739
24	O	Me	SPh(3-OH)	$CH_2OCH_2CH_2OH$	4.085	4.790	−0.705
25	O	Me	SPh(3-OMe)	$CH_2OCH_2CH_2OH$	4.657	4.830	−0.173
26	O	Me	SPh(3,5-Me_2)	$CH_2OCH_2CH_2OH$	6.584	6.031	0.553
27	O	Me	SPh(3,5-Cl_2)	$CH_2OCH_2CH_2OH$	5.885	5.507	0.378
28	S	Me	SPh(3,5-Me_2)	$CH_2OCH_2CH_2OH$	6.656	6.334	0.322
29	O	Me	SPh(3-COOMe)	$CH_2OCH_2CH_2OH$	5.102	5.389	−0.287
30	O	Me	SPh(3-COMe)	$CH_2OCH_2CH_2OH$	5.136	5.595	−0.459
31	O	Me	SPh(3-CN)	$CH_2OCH_2CH_2OH$	4.999	5.420	−0.421
32	O	$CH_2CH{=}CH_2$	SPh	$CH_2OCH_2CH_2OH$	5.601	6.057	−0.456
33	O	COOMe	SPh	$CH_2OCH_2CH_2OH$	5.180	5.126	0.054
34	O	CONHPh	SPh	$CH_2OCH_2CH_2OH$	4.744	5.063	−0.319
35	S	Et	SPh	$CH_2OCH_2CH_2OH$	6.957	6.157	0.800
36	S	Pr	SPh	$CH_2OCH_2CH_2OH$	4.999	6.430	−1.431
37	S	Et	SPh(3,5-Me_2)	$CH_2OCH_2CH_2OH$	8.106	7.864	0.242
38	S	i-Pr	SPh(3,5-Me_2)	$CH_2OCH_2CH_2OH$	8.300	8.636	−0.336
39	S	Et	SPh(3,5-Cl_2)	$CH_2OCH_2CH_2OH$	7.365	7.503	−0.138
40	O	Et	SPh	$CH_2OCH_2CH_2OH$	6.920	6.113	0.807
41	O	Pr	SPh	$CH_2OCH_2CH_2OH$	5.468	6.090	−0.622
42	O	i-Pr	SPh	$CH_2OCH_2CH_2OH$	7.199	7.133	0.066
43	O	Et	SPh(3,5-Me_2)	$CH_2OCH_2CH_2OH$	7.884	7.696	0.188
44	O	i-Pr	SPh(3,5-Me_2)	$CH_2OCH_2CH_2OH$	8.567	8.615	−0.048
45	O	Et	SPh(3,5-Cl_2)	$CH_2OCH_2CH_2OH$	7.852	7.485	0.367
46	O	Me	SPh	$CH_2OCH_2CH_2OCH_2Ph$	4.698	5.314	−0.616
47	O	Me	SPh	CH_2OMe	5.677	4.928	0.749
48	O	Me	SPh	CH_2OPr	5.443	5.293	0.150
49	O	Me	SPh	CH_2OBu	5.327	5.395	−0.068
50	O	Me	SPh	CH_2OCH_2Ph	7.054	6.145	0.909
51	S	Et	SPh (3,5-Me_2)	CH_2OEt	8.355	7.995	0.360
52	S	Et	SPh (3,5- Cl_2)	CH_2OEt	7.885	7.506	0.379
53	S	Et	SPh	CH_2-*i*-Pr	6.656	6.396	0.260
54	S	Et	SPh	CH_2OCH_2-c-Hex	6.455	7.498	−1.043
55	S	Et	SPh	CH_2OCH_2Ph	8.106	7.643	0.463
56	S	Et	SPh (3,5-Me_2)	CH_2OCH_2Ph	8.160	9.419	−1.259
57	S	Et	SPh	$CH_2OCH_2C_6H_4$(4-Me)	7.107	7.771	−0.664
58	S	Et	SPh	CH_2OCH_2(4-Cl)	7.919	7.593	0.326
59	S	Et	SPh	$CH_2OCH_2CH_2Ph$	7.040	6.755	0.285
60	S	i-Pr	SPh	CH_2OEt	7.852	7.790	0.062
61	S	i-Pr	SPh	CH_2OCH_2Ph	8.179	8.695	−0.516
62	S	c-Pr	SPh	CH_2OEt	7.021	7.086	−0.065
63	O	Et	SPh (3,5- Cl_2)	CH_2OEt	8.129	7.369	0.760
64	O	Et	SPh	CH_2O-*i*-Pr	6.467	6.134	0.333
65	O	Et	SPh	CH_2O−c-Hex	5.397	6.477	−1.080
66	O	Et	SPh	CH_2OCH_2-c-Hex	6.346	6.814	−0.468
67	O	Et	SPh	CH_2OCH_2Ph	8.228	7.423	0.805
68	O	Et	SPh	$CH_2OCH_2CH_2Ph$	7.016	6.766	0.250
69	O	i-Pr	SPh	CH_2OCH_2Ph	8.567	8.864	−0.297
70	O	Me	SPh	Et	5.657	6.063	−0.406
71	O	Me	SPh	Bu	5.920	5.234	0.686
72	O	Et	CH_2Ph	$CH_2OCH_2CH_2OH$	6.455	6.362	0.093
73	O	Et	CH_2Ph(3,5-Me)	$CH_2OCH_2CH_2OH$	7.885	6.921	0.964
74	O	Et	CH_2Ph	CH_2OEt	7.386	7.267	0.119
75	O	i-Pr	CH_2Ph	$CH_2OCH_2CH_2OH$	7.199	7.305	−0.106
76	O	i-Pr	CH_2Ph(3,5-Me)	$CH_2OCH_2CH_2OH$	8.567	7.969	0.598
77	O	i-Pr	CH_2Ph	CH_2OEt	8.375	8.128	0.247
78	O	i-Pr	CH_2Ph(3,5-di-Me)	CH_2OEt	9.220	8.726	0.494
79	O	i-Pr	CH_2Ph	Bu	7.374	7.819	−0.445
80	O	Et	CH_2Ph	CH_2CH_2OMe	6.601	6.945	−0.344
81	O	i-Pr	CH_2Ph	CH_2CH_2OMe	7.282	7.145	0.137
82	O	Me	SPh	$CH_2OCH_2CH_2OH$	5.154	4.816	0.338

Table 2. Structure and Predicted log(1/C) HIV-1 RT Inhibitory Affinities of the Test Set

	substituents				log(1/C)		
no.	X	R1	R2	R3	exptl	calcd[a]	residual
83	O	Me	OPh	$CH_2OCH_2CH_2OH$	4.070	5.807	−1.737
84	O	Me	C≡CH	$CH_2OCH_2CH_2OH$	5.259	4.387	0.872
85	O	C≡CH	SPh	$CH_2OCH_2CH_2OH$	4.744	4.698	0.046
86	O	Me	SPh(3-F)	$CH_2OCH_2CH_2OH$	5.481	4.140	1.341
87	S	*i*-Pr	SPh	$CH_2OCH_2CH_2OH$	7.228	7.514	−0.286
88	O	Me	SPh	$CH_2OCH_2CH_2OMe$	5.060	5.221	−0.161
89	O	Me	SPh	$CH_2OCH_2CH_2O-C_6H_{11}$-*n*	4.259	5.052	−0.793
90	O	Me	SPh	CH_2OEt	6.480	5.098	1.382
91	O	Me	SPh	$CH_2OCH_2CH_2SiMe_3$	4.494	5.062	−0.568
92	S	Et	SPh	CH_2OEt	7.584	6.979	0.605
93	S	Et	SPh	CH_2O−c-Hex	5.795	6.245	−0.450
94	O	Et	SPh	CH_2OEt	7.720	6.653	1.067
95	O	Et	SPh(3,5-Me_2)	CH_2OEt	8.266	8.002	0.264
96	O	Et	SPh(3,5-Me_2)	CH_2OCH_2Ph	8.493	8.935	−0.442
97	O	i-Pr	SPh	CH_2OEt	7.919	7.648	0.271
98	O	c-Pr	SPh	CH_2OEt	6.999	6.952	0.047
99	O	Et	CH_2Ph(3,5-Me_2)	$EtOCH_2$	7.199	7.906	−0.707
100	O	Et	CH_2Ph	Bu	6.677	6.963	−0.286
101	S	Me	SPh	$CH_2OCH_2CH_2OH$	6.008	5.153	0.855

[a] Calculated by CoMFA model 13.

Table 3. Summary of CoMFA Models with 82 HEPT Derivatives with the Different Probe Atoms

model	probe atom	field type	noc	r^2_{cv}	S_{press}	$r^{2\,a}$	*s*	*F*	outliers (residual)	steric contb[b]
1	$sp^3C(+1)$	both	4	0.837	0.567	0.930	0.371	255.352	compd 36(−1.431)	63.9
2		steric	3	0.771	0.667	0.849	0.542	145.838		
3		electrostatic	2	0.726	0.724	0.830	0.571	192.327		
4	$sp^3O(-1)$	both	4	0.830	0.578	0.926	0.382	239.639	compd 36(−1.427)	61.5
5		steric	5	0.782	0.658	0.914	0.414	161.394		
6		electrostatic	2	0.726	0.724	0.830	0.571	192.327		
7	H(+1)	both	4	0.820	0.594	0.922	0.392	226.828	compd 36(−1.577)	59.8
8		steric	3	0.749	0.697	0.833	0.569	130.006		
9		electrostatic	2	0.726	0.724	0.830	0.571	192.327		
10	$sp^3C(+1)$	both	4	0.849	0.545	0.936	0.356	275.922	compd 50(1.255)[c]	63.4
11		steric	2	0.774	0.660	0.838	0.560	201.628		
12		electrostatic	2	0.781	0.651	0.861	0.519	242.022		
13	$sp^3C(+1)$	both	4	0.858	0.531	0.941	0.342	300.134	compd. 67(1.106)[d]	64.4
14		steric	4	0.848	0.544	0.937	0.351	277.997		
15		electrostatic	4	0.860	0.529	0.942	0.340	303.570		

[a] Conventional r^2. [b] Steric contribution in %. [c] Elimination of compound 36 (remaining 81 compounds in the training set). [d] Elimination of compounds 36 and 50 (remaining 80 compounds in the training set).

0.941) and the standard error of estimation (0.342). *F* is 300.134 and the probability (P) of obtaining this value of *F* if r^2 is actually zero (probability of $r^2 = 0$) is lower than 0.001. The plot between predicted and experimental inhibitory affinities of the non-cross-validated analysis of model 13 is presented in Figure 2.

Prediction of the Inhibitory Activity for the Compounds in the Test Set. Consequently, the best CoMFA model (model 13) was used to predict the inhibitory activity of the compounds in the test set. The observed and fitted inhibitory activities of 19 compounds (compounds 83−101) are listed in Table 2. Based on the residual value, it is indicated that model 13 is able to predict the activities of HEPT derivatives which are not included in the training set. In particular, this model predicts well the activities for compounds 85, 87, 88, 93, 95, 97, 98, and 100. These structures are substituted at C5 position by ethyl or propyl groups. Compound 85 is substituted by ethynyl group and compound 98 is substituted by cyclopropyl group. Further, the model can be useful to predict the activities of compounds 84, 89, 91, 92, 96, 99, and 101. The prediction for compounds 83, 86, 90, and 94 shows larger residuals. A comparison of the fitted and experimental activities of the test set is plotted in Figure 3. The numbering of the compounds is used as labeling.

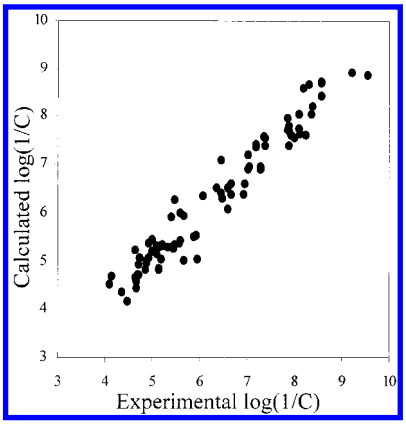


Figure 2. Plot of calculated versus experimental HIV-1 RT inhibitory affinities (log(1/C)), obtained from non-cross-validation of CoMFA model 13 for training set.

Steric and Electrostatic Contributions. The CoMFA steric and electrostatic fields for all HEPT derivatives are

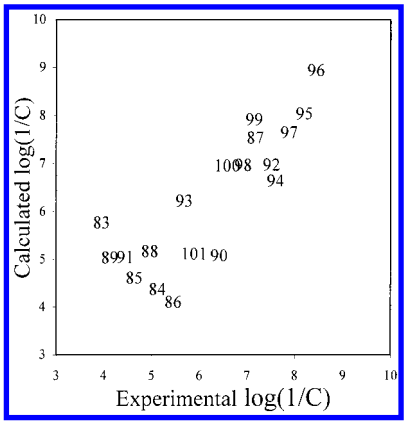


Figure 3. Plot of calculated versus experimental HIV-1 RT inhibitory affinities (log(1/C)), obtained from non-cross-validation of CoMFA model 13 for test set.

presented as contour maps in Figures 4 and 5. X-ray crystal structures of HIV-1 RT complexes with various inhibitors are available, the amino acid residues surrounding HEPT in the structure of the interaction complex were merged into both figures. In Figure 4, the green and yellow contour maps represent regions whose occupancy by parts of the ligands increases or decreases the receptor binding affinity. It is interesting to note that there is only one favorable steric region corresponding to the location around the C5-side chain (R1 substitution) at the thymine ring. This leads to the idea that an additional bulky group at this position would increase the activity, but the size of this group should not be too large because of the limited volume of the binding pocket. This suggestion is based on the experimental trend that an isopropyl group is the best group for C5-substitution.[20] Furthermore, the positive steric contour region is close to the position of the amino acids Tyr181 and Tyr188. Both amino acids are considered to be very important for the inhibition of the enzymes' activity, due to their changed positions at the inhibition site as a consequence of the association of the inhibitor molecule. This evidence can be supported by the fact that there are substantial steric interactions between the side chains of the Tyr181 and R1 substituent of the thymine ring of the HEPT molecules. A less favorable interaction between the amino acid Tyr181 and C5-substituent, therefore, decreases the affinity of this inhibitor. Several crystal structures of the RT complex with inhibitors, such as HEPT (compound 82), MKC-442 (compound 77), and TNK-651 (R1 = *i*-Pr, R2 = CH_2Ph, R3 = CH_2OCH_2Ph),[26] again indicate that the substituent at position 5 of the thymine ring is essential for HEPT analogues to bind to RT. The 5-methyl group of the lead compound, HEPT, interacts only slightly with Tyr181, whereas the larger substituents of MKC-442 and TNK-651, the 5-isopropyl group, appears to force Tyr181 to change its conformation. Tyr181 is then able to form strong interactions with the 6-phenyl ring of the inhibitor and other aromatic residues (Figure 6). According to the experimental result,[32] the mutation of Tyr181 to Cys apparently eliminates favorable contacts of the aromatic ring of the tyrosine and the bound inhibitor, reducing the NNRTI binding affinity. It should also be mentioned that the contribution of the C5 parameter is considered in our previous QSAR analysis.[9] The obtained results can be concluded that a moderately sized group at C5 enhances contact with Tyr181 enough to push it into a position which renders the protein nonfunctional, but a smaller group has insufficient steric requirements to do this and a larger group renders the ligand too large for the cavity.

The yellow contours, which are located between the R2 and R3 substituents, suggest that the substituents should not have branches because they may have steric interaction which could inhibit the association between the N1 side-chain and the 6-phenyl ring of the inhibitor with the amino acid residues in these area (Val106, Phe227, Trp229, Leu234, His235, and Pro236). These results strongly reveal the importance of the steric feature of the inhibitors contributing to affinity through contour maps. The architecture of the binding pocket undergoes conformational changes to accept each of inhibi-

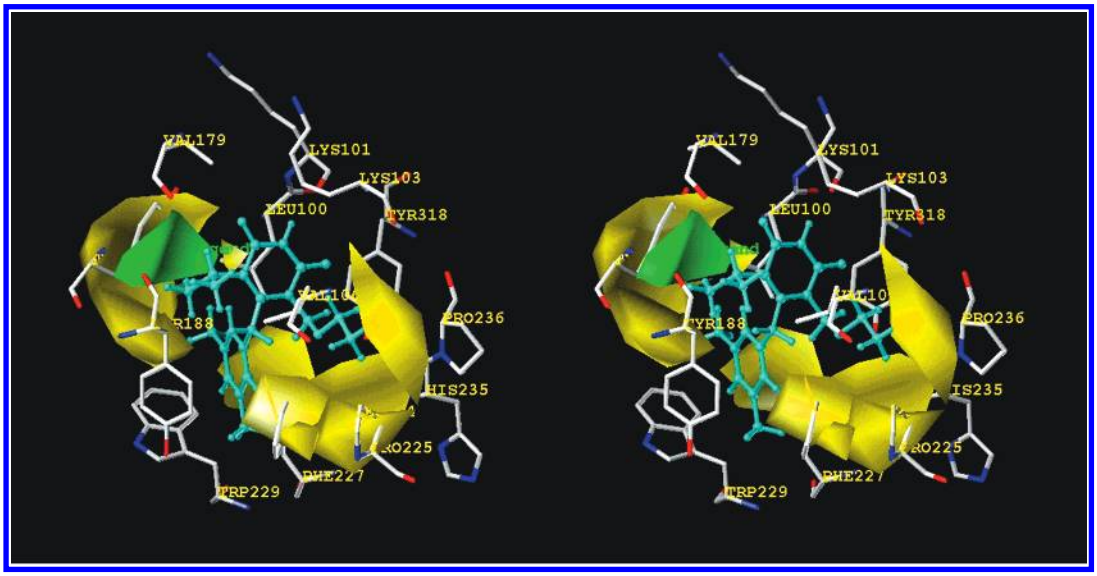


Figure 4. Steric std*coeff contour map, obtained from CoMFA model 13 for HIV-1 inhibition, is shown together with amino acid residues within 6 Å around HIV-1 RT active site. HEPT compound is presented inside the field as light-blue stick structure, with undisplayed hydrogens. Sterically CoMFA favored areas are represented by green region. Sterically CoMFA unfavored areas are represented by yellow region (level of steric contour contribution = 80%).

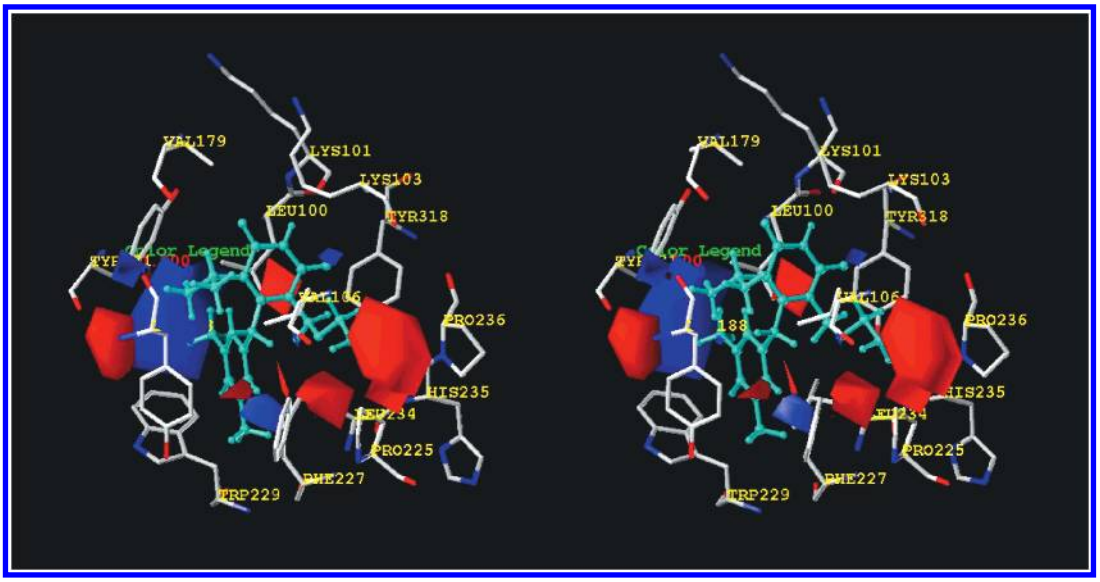


Figure 5. Electrostatic std*coeff contour map, obtained from CoMFA model 13 for HIV-1 inhibition, is shown together with amino acid residues within 6 Å around HIV-1 RT active site. HEPT compound is presented inside the field as a light-blue stick structure, with undisplayed hydrogens. Negative charge CoMFA favored areas are represented by red region. Negative charge CoMFA unfavored areas are represented by blue region (level of electrostatic contour contribution = 80%).

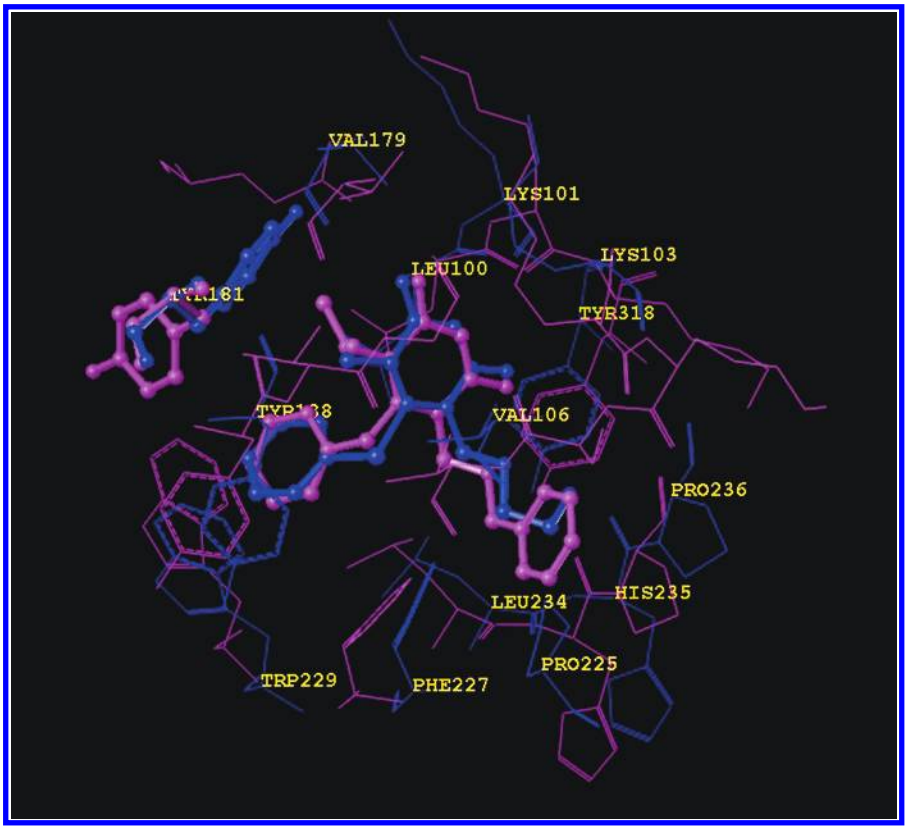


Figure 6. Superposition the crystal structures of the binding pockets of HEPT series, HEPT/HIV-1 RT complex shown in blue color, and TNK-651 complex shown in magenta color. The structures of inhibitors and amino acid residue Tyr181 are displayed in stick.

tors. However, the volume of the pocket has some limitation; there is not enough space to accommodate molecules which are too large.

The electrostatic contribution contour map is depicted in Figure 5. A blue positive electrostatic contour region located near the substituent, which is attached to C6 position and a larger one close to the C5 position of the thymine ring show the favorable positive charges. These results complement with the QSAR model[8] that atomic charge at C5 position and molar refractivity, representing steric interaction of substituents of the compounds, play an important role in the model, particularly the relationships between the atomic charge and the nature of the substituent on the atom. A red contour region, close to carbonyl group of the thymine ring and the N1-side chain, suggests that high negative charges in these areas enhance the affinity.

According to the contour maps obtained from the CoMFA model, it turns out that one of the important binding sites of the HEPT-derived inhibitors is the interactions of an alkyl side chain positioned at C5 with Tyr 181 and Tyr 188

residues in the binding pocket of RT. Numerous studies have indicated that Tyr181Cys HIV-1 RT mutation is frequently found in the presence of NNRTIs.[33−37] This mutation leads to a resistance against the NNRTIs due to the loss of the hydrophobic interactions of the inhibitors with the surrounding amino acids for steric reasons. In the case of nevirapine analogues this resistance is proven experimentally,[32] and moreover, it is also possible to explain the difference of the inhibitor interactions between wild type and mutant type enzymes from the CoMFA analysis.[25]

In HEPT derivatives the substituents on the phenylthio side chain at C6 of the thymine ring increase the hydrophobicity of the molecules which results in a favorable inhibition by the improvement of the interaction to the hydrophobic binding pocket of HIV-1 RT, which is rather similar (i.e. steric and electrostatic contour maps show good consistence with inhibitor-enzyme complex structure derived by experimental data) to nevirapine and to another important NNRTI.[24,25] Therefore, to suggest a substance to be a high potent inhibitor of RT against either wild-type or mutant HIV-1 RT, it must bind tightly to the specific amino acid residues that yield the structural changes at the active site of the enzyme and result in low reverse transcriptase activity.

CONCLUSIONS

A 3D-QSAR investigation on HIV-1 RT inhibitors was performed, using structures calculated by ab initio methods. For a large set of HEPT derivatives, these CoMFA studies successfully explain steric and electrostatic interactions of the molecules with the amino acids of the inhibition pocket of the HIV-1 RT. The CoMFA model is of satisfactory quality, having $r^2_{cv} = 0.858$. Steric fields are the primary contributor (64.4%) to the CoMFA model of RT inhibition. The steric contours correspond well with the geometrical position and charge requirements of amino acid residues of the active site and are consistent with the experimentally observed trend that there are steric interactions between the side chain of HEPT and an aromatic ring of Tyr181. The obtained results can be concluded that a moderately sized group at C5 enhances contact with Tyr181 enough to push it into a position which renders the protein nonfunctional, but a smaller group has insufficient steric requirements to do this and a larger group renders the ligand too large for the cavity. It also indicates clearly that losing favorable interactions between aromatic ring of Tyr181 and HEPT may decrease the affinity of this inhibitor. Furthermore, this study provides the common basic information and structural requirements in the class of NNRTIs that are useful for the design of new and high potent compounds.

ACKNOWLEDGMENT

This work was supported by grants from the National Research Council of Thailand and Fonds zur Foerderung der Wissenschaftlichen Forschung (P12257-CHE) under the Austria-Thailand Cooperative Science Program (NRCT-FWF) and partially supported by a Grant-in-Aid for thesis from the Graduate School, Kasetsart University. S. Hannongbua is grateful to the Thailand Research Fund and KURDI for a research fellowship (RSA4480001). L. Lawtrakul is grateful to the DSPT project for scholarships. P. Pungpo is grateful to the RGJPhD project for the scholarships (3.C.KU/41/B.1). The generous supply of computing time and Sybyl 6.5 installation by the high performance-computing center of the National Electronics and Computer Technology (NECTEC) and Cluster computer of Universität Wien are also gratefully acknowledged. We thank A. Karpfen and G. Ecker for helpful comments.

REFERENCES AND NOTES

(1) De Clercq, E. Antiviral Therapy of Human Immunodeficiency Virus Infections. *Clin. Microbiol. Rev*. **1995**, *8*, 200−239.
(2) Young, S. D. Non-Nucleoside Inhibitors of HIV-1 Reverse Transcriptase. *Perspect. Drug Descov. Des.* **1993**, *1*, 181−192.
(3) De Clercq, E. Toward Improved Anti-HIV Chemotherapy: Therapeutic Strategies For Intervention With HIV-1 Infections. *J. Med. Chem*. **1995**, *38*, 2491−2517.
(4) Arnold, E.; Das, K.; Ding, J.; Yadav, P. N. S.; Hsiou, Y.; Boyer, P. L.; Hughes, S. H. Targeting HIV Reverse Transcriptase For Anti-AIDS Drug Design. *Drug Des. Discov*. **1996**, *13*, 29−47.
(5) Esnouf, R.; Re, J.; Ross, C.; Jones, Y.; Stammers, D.; Stuart, D. Mechanism of Inhibition of HIV-1 Reverse Transcriptase by Non-Nucleoside Inhibitors. *Struct. Biol*. **1995**, *2*, 303−308.
(6) Sarafianos, S. G.; Das, K.; Ding, J.; Boyer, P. L.; Hughes, S. H.; Arnold, E. Touching the Hear of HIV-1 Drug Resistance: The Fingers Close Down on the dNTP at the Polymerase Active Site. *Chem. Biol*. **1999**, *6*, R137−R146.
(7) Miyasaka, T.; Tanaka, H.; Baba, M.; Hayakawa, H.; Walker, R. T.; Balzarini, J.; De Clercq, E. A Novel Lead for Specific Anti-HIV-1 Agents: 1-[(2-Hydroxyethoxy)methyl]-6-(phenylthio) thymine. *J. Med. Chem*. **1989**, *32*, 2507−2509.
(8) Hannongbua, S.; Lawtrakul, L.; Limtrakul, J. Structure−Activity Correlation Study of HIV-1 Inhibitors: Electronic and Molecular Parameters. *J. Comput.-Aided Mol. Design*. **1996**, *10*, 145−152.
(9) Hannongbua, S.; Lawtrakul, L.; Sotriffer, C. A.; Rode, B. M. Comparative Molecular Field Analysis of HIV-1 Reverse Transcriptase Inhibitors in the Class of 1-[(2-Hydroxyethoxy)-methyl]-6-(phenylthio)thymine. *Quant. Struct.-Act. Relat*. **1996**, *15*, 389−394.
(10) Kireev, D. B.; Chretien, J. R.; Grierson, D. S.; Monneret, C. A 3D QSAR Study of a Series of HEPT Analogues: The Influence of Conformational Mobility on HIV-1 Reverse Transcriptase Inhibition. *J. Med. Chem*. **1997**, *40*, 4257−4264.
(11) Luco, J. M.; Ferretti, F. H. QSAR Based on Multiple Linear Regression and PLS Methods for the Anti-HIV Activity of a Large Group of HEPT Derivatives. *J. Chem. Inf. Comput. Sci*. **1997**, *37*, 392−401.
(12) Tronchet, J. M. J.; Grigorov, M.; Dolatshahi, N.; Moriaud, F.; Weber, J. A QSAR Study Confirming the Heterogeneity of the HEPT Derivative Series Regarding Their Interaction with HIV Reverse Transcriptase. *Eur. J. Med. Chem*. **1997**, *32*, 279−299.
(13) Lawtrakul, L.; Hannongbua, S. Quantitative Structure−Activity Relationships of HIV-1 RT Inhibitors in the Class of 1-[(2-Hydroxyethoxy)methyl]-5,6-substituted Thymine [HEPT] Analogues. *Sci. Pharm*. **1999**, *67*, 43−56.
(14) Jalali-Heravi, M.; Parastar, F. Use of Artificial Neural Networks in a QSAR Study of Anti-HIV Activity for a Large Group of HEPT Derivatives. *J. Chem. Inf. Comput. Sci*. **2000**, *40*, 147−154.
(15) Klein, C. T.; Lawtrakul, L.; Hannongbua, S.; Wolschann, P. Accessible Charges in Structure−Activity Relationships. A Study on HEPT-based HIV-1 RT Inhibitors. *Sci. Pharm*. **2000**, *68*, 25−40.
(16) Lawtrakul, L.; Hannongbua, S.; Beyer, A.; Wolschann, P. Conformational Study of the HIV-1 Reverse Transcriptase Inhibitor 1-[(2-Hydroxyethoxy)methyl]-6-(phenylthio)thymine (HEPT). *Biol. Chem*. **1999**, *380*, 265−267.
(17) Lawtrakul, L.; Hannongbua, S.; Beyer, A.; Wolschann, P. Molecular Calculations on the Conformation of the HIV-1 Reverse Transcriptase Inhibitor 1-[(2-Hydroxyethoxy)methyl]-6-(phenylthio)thymine (HEPT). *Monatsh. Chem*. **1999**, *130*, 1347−1363.
(18) Tanaka, H.; Baba, M.; Kayakawa, H.; Sakamaki, T.; Miyasaka, T.; Ubasawa, M.; Takashima, H.; Sekiya, K.; Nitta, I.; Shigeta, S.; Walker, R. T.; Balzarini, J.; De Clercq, E. A New Class of HIV-1-Specific 6-substituted Acyclouridine Derivatives: Synthesis and Anti-HIV-1 Activity of 5- or 6-substituted Analogues of 1-[(2-Hydroxyethoxy)-methyl]-6-(phenylthio)thymine (HEPT). *J. Med. Chem*. **1991**, *34*, 349−357.
(19) Tanaka, H.; Takashima, H.; Ubasawa, M.; Sekiya, K.; Nitta, I.; Baba, M.; Shigeta, S.; Walker, R. T.; De Clercq, E.; Miyasaka, T. Structure−Activity Relationships of 1-[(2-Hydroxyethoxy)methyl]-6-(phenylthio)-thymine Analogues: Effect of Substitutions at the C-6 Phenyl Ring and at the C-5 Position on Anti-HIV-1 Activity. *J. Med. Chem*. **1992**, *35*, 337−345.
(20) Tanaka, H.; Takashima, H.; Ubasawa, M.; Sekiya, K.; Nitta, I.; Baba, M.; Shigeta, S.; Walker, R. T.; De Clercq, E.; Miyasaka, T. Synthesis

and Antiviral Activity of Deoxy Analogues of 1-[(2-Hydroxyethoxy)-methyl]-6-(phenylthio)thymine (HEPT) as Potent and Selective Anti-HIV-1 Agents. *J. Med. Chem*. **1992**, *35*, 4713−4719.

(21) Tanaka, H.; Takashima, H.; Ubasawa, M.; Sekiya, K.; Inouye, N.; Baba, Masanori; Shigeta, S.; Walker, R. T.; De Clercq, E.; Miyasaka, T. Synthesis and Antiviral Activity of 6-Benzyl Analogues of 1-[(2-Hydroxyethoxy)methyl]-6-(phenylthio)thymine (HEPT) as Potent and Selective Anti-HIV-1 Agents. *J. Med. Chem*. **1995**, *38*, 2860−2865.

(22) Kubinyi, H. QSAR: *Hansch Analysis and Related Approaches*; VCH: Weinheim, 1993.

(23) Kubinyi, H. *3D-QSAR in Drug Design, Theory, Methods and Applications*; ESCOM: Leiden, 1993.

(24) Hannongbua, S.; Pungpo, P.; Limtrakul, J.; Wolschann, P. Quantitative Structure−Activity Relationships and Comparative Molecular Field Analysis of TIBO derivatived HIV-1 Reverse Transcriptase Inhibitors. *J. Comput.-Aided Mol. Design*. **1999**, *13*, 563−577.

(25) Pungpo, P.; Hannongbua, S. Three-Dimensional Quantitative Structure−Activity Relationships Study on HIV-1 Reverse Transcriptase Inhibitors in the Class of Dipyridodiazepinone Derivatives, Using Comparative Molecular Field Analysis. *J. Mol. Graph. Mod*. **2000**, *18*, 581−590.

(26) Hopkins, A. L.; Ren, J.; Esnouf, R. M.; Willcox, B. E.; Jones, Y.; Ross, C.; Miyasaka, T.; Walker, R. T.; Tanaka, H.; Stammers, D. K.; Stuart, D. I. Complexes of HIV-1 Reverse Transcriptase with Inhibitors of the HEPT Series Reveal Conformational Changes Relevant to the Design of Potent Non-Nucleoside Inhibitors. *J. Med. Chem*. **1996**, *39*, 1589−1600.

(27) ALCHEMY2000; Tripos Associates Inc.: St. Louis, MO, 1996.

(28) GAUSSIAN94; Carnegie Office Park, Building 6, Pittsburgh, PA, 1995.

(29) SYBYL Molecular Modelling Softwares, Version 6.5; Tripos Associates, Inc.: St. Louis, MO, 1998.

(30) Wold, S.; Johansson, E.; Cochi, M.; In *3D QSAR in Drug Design, Theory, Methods and Applications*; Kubinyi, H., Ed.; ESCOM: Leiden, 1993; pp 523−549.

(31) SYBYL Molecular Modelling Softwares, Version 6.3; Tripos Associates, Inc.: St. Louis, MO, 1996; p 229.

(32) Proudfoot, J. R.; Hargrave, K. D.; Kapadia, S. R.; Patel, U. R.; Grozinger, K. G.; Mcneil, D. W.; Cullen, E.; Cardozo, M.; Tong, L.; Kelly, T. A.; Rose, J.; David, E.; Mauldin, S. C.; Fuchs, V. U.; Vitous, J.; Hoermann, M.; Klunder, J. M.; Raghavan, P.; Skiles, J. W.; Mui, P.; Richman, D. D.; Sullivan, J. L.; Shih, C.; Grob, P. M.; Adams, J. Novel Non-Nucleoside Inhibitors of Human Imunodeficiency Virus Type 1 (HIV-1) Reverse Transcriptase. 4. 2-Substituted Dipyridodiazepinones as Potent Inhibitors of Both Wild-Type and Cysteine-181 HIV-1 Reverse Transcriptase Enzyme. *J. Med. Chem*. **1995**, *38*, 4830−4838.

(33) Pauwels, R.; Andries, K.; Debyser, Z.; Daele, P. V.; Schols, D.; Stoffels, P.; De Vreese, K.; Woestenborghs, R.; Vandamme, A.; Janssen, C. G. M.; Anne, J.; Cauwenbergh, G.; Desmyter, J.; Heykants, J.; Janssen, M. A. C.; De Clercq, E.; Janssen, P. A. J. Potent and Highly Selective Human Immunodeficiency Virus Type 1 (HIV-1) Inhibition by a Series of Alpha-Anilinophenylacetamide Derivatives Targeted at HIV-1 Reverse Transcriptase. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1711−1715.

(34) Buckheit, R. W., Jr.; Fliakas-Boltz, V.; Yeagy-Bargo, S.; Weislow, O.; Mayers, D. L.; Boyer, P. L.; Hughes, S. H.; Pan, B.; Chu, S.; Bader, J. P. Resistance to 1-[(2-Hydroxyethoxy)methyl]-6-(phenylthio)-thymine Derivatives is Generated by Mutations at Multiple Sites in the HIV-1 Reverse Transcriptase. *Virology* **1995**, *210*, 186−193.

(35) Balyarini, J.; Perez-Perez, M.-J.; Velazquez, S.; San-Felix, A.; Camarasa, M.-J.; De Clercq, E.; Karlsson, A. Suppression of the Breakthrough of Human Immunodeficiency Virus Type 1 (HIV-1) in Cell Culture by Thiocarboxanilide Derivatives when Used Individually or in Combination with Other HIV-1-specific Inhibitors (i.e., TSAO Derivatives). *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 5470−5474.

(36) Das, K.; Ding, J.; Hsiou, Y.; Clark, A. D., Jr.; Moereels, H.; Koymans, L.; Andries, K.; Pauwels, R.; Janssen, P. A. J.; Boyer, P. L.; Clark, P.; Smith, R. H., Jr.; Smith, M. B. K.; Michejda, C. J.; Hughes, S. H.; Arnold, E. Crystal Structures of 8-Cl and 9-Cl TIBO Complexed with Wild-Type HIV-1 RT and 8-Cl TIBO Complexed with the Tyr181Cys HIV-1 RT Drug-Resistant Mutant. *J. Mol. Biol*. **1996**, *264*, 1085−1100.

(37) Yu, H.; Das, K.; Ding, J.; Clark, A. D., Jr.; Rösner, M.; Winkler, I.; Riess, G.; Hughes, S. H.; Arnold, E. Structures of Tyr188Leu Mutant and Wild-Type HIV-1 Reverse Transcriptase Complexed with the Non-Nucleoside Inhibitor HBY 097: Inhibitor Flexibility is a Useful Design Feature for Reducing Drug Resistance. *J. Mol. Biol*. **1998**, *284*, 313−323.

CI0001278