Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Molecular Graphics and Modelling 23 (2005) 317–328
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Cluster analysis and three-dimensional QSAR studies of HIV-1 integrase inhibitors

Hongbin Yuan, Abby Parrill*

Chemistry Department, The University of Memphis, Memphis, TN 38152, USA

Received 30 December 2003; received in revised form 14 October 2004; accepted 14 October 2004
Available online 21 November 2004

Abstract

Three-dimensional quantitative structure-activity relationship (3D QSAR) and cluster analysis were applied to a variety of HIV-1 integrase inhibitors. One structure was chosen from each of 11 classes of inhibitors to represent the whole class in descriptor-based cluster analysis. The 11 classes of inhibitors were classified into two groups. The molecular field analysis (MFA) models for these two clusters had r^2 values of 0.90 and 0.95 and q^2 values of 0.85 and 0.91 that were noticeably enhanced from those of conventional QSAR models. The five test compounds, which were proposed to have a common binding site near the metal in HIV-1 integrase based on docking studies by Sotriffer et al., were utilized to compare the predictive capability of MFA and conventional QSAR models. Among these five compounds, only L-chicoric acid belongs to cluster 1 and the other four belong to cluster 2. MFA models give better overall predictions and more importantly the activity of these test compounds is better predicted by the MFA model derived from the cluster each test compound belongs to. The necessity of dividing the inhibitors into two groups to obtain predictive QSAR models supports the likelihood of two separate binding sites.
© 2004 Elsevier Inc. All rights reserved.

Keywords: Human immunodeficiency virus; Integrase; Inhibition; Cluster analysis; Quantitative structure-activity relationship

1. Introduction

Human immunodeficiency virus type 1 (HIV-1) integrase is an enzyme required for viral replication [1]. In two reactions IN catalyzes, 3′-processing and strand transfer, the last two nucleotides of the viral DNA 3′-end are cleaved and the remaining viral fragment is inserted into the host DNA [2,3]. HIV-1 integrase is currently recognized as an attractive target to develop effective treatments against HIV/AIDS.[4,5] A large number of HIV-1 integrase inhibitors have been discovered [6]. However, the mechanism of inhibition is incompletely understood [7].

Computational chemistry has developed into an important contributor to rational drug design. Quantitative structure-activity relationship (QSAR) modeling results in a quantitative correlation between chemical structure and biological activity [8]. Conventional QSAR relies on descriptors to characterize structure [9,10]. In contrast, three-dimensional QSAR uses the steric and electrostatic fields surrounding small molecules to reflect interactions between ligands and target molecules. Comparative molecular field analysis (CoMFA) is a well known three-dimensional QSAR method in which partial least squares (PLS) is the method of data analysis and a field fit technique is required for optimal alignment by minimizing the RMS field differences between studied molecules [11]. Molecular field energies can also be used as structural descriptors with other regression methods if coupled with variable selection techniques, such as the genetic function approximation (GFA) [27].

Several lines of evidence suggest that the known HIV-1 integrase inhibitors interact with integrase at more than one binding site. First, competitive binding studies indicate that some inhibitors display competitive binding, while others do not [12]. Second, crystallographic studies on HIV-1 and ASV integrase inhibitor complexes showed three distinct inhibitor binding sites in the two structurally similar enzymes [13–15]. Third, in our previous computational studies, a number of classes of HIV-1 integrase inhibitors

* Corresponding author. Tel.: +1 90 1678 2638; fax: +1 90 1678 3447.
E-mail address: aparrill@memphis.edu (A. Parrill).

1093-3263/$ – see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2004.10.003

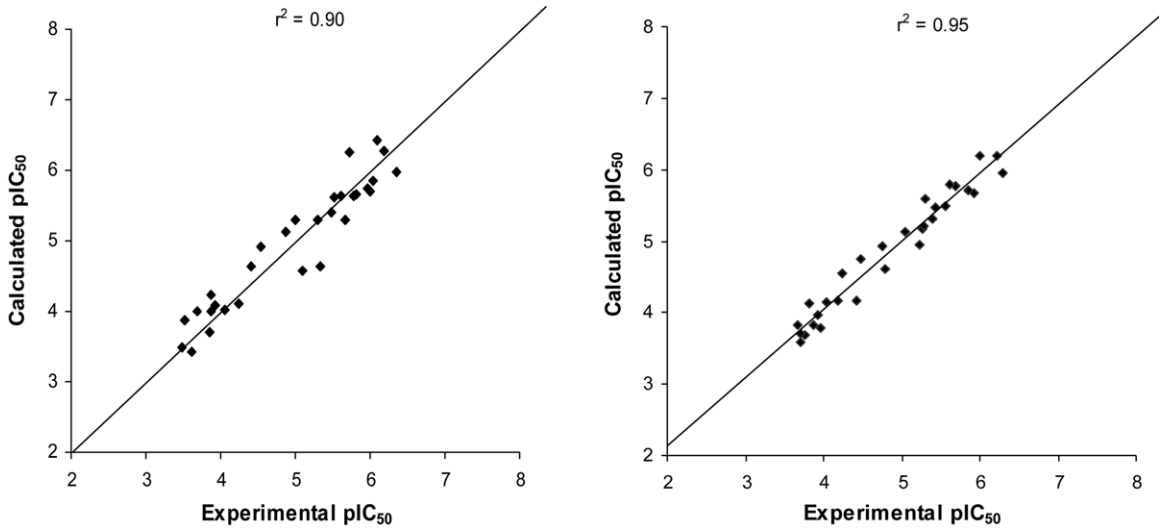
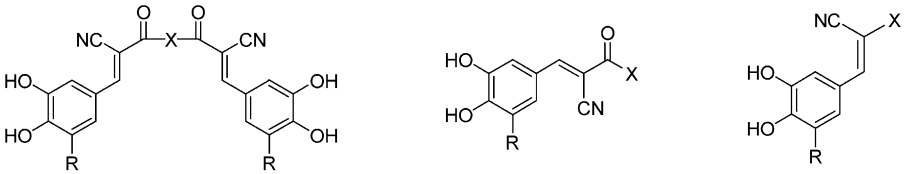


Fig. 1. (A) Calculated pIC_{50} by Model 1 vs. experimental pIC_{50} of compounds in cluster 1. (B) Calculated pIC_{50} by Model 2 vs. experimental pIC_{50} of compounds in cluster 2.

have been clustered into two groups that show different structure-activity relationships and are likely to act at two independent binding sites [16]. The biological activities (IC_{50}s) of inhibitors in each group can be calculated correctly by two conventional QSAR models separately developed from the two groups. Likewise, Makhija and Kulkarni [17] also found that development of QSAR equations to describe integrase inhibition was facilitated by separating the inhibitors into two groups.

In a published docking study, five compounds, including 1-(5-chloroindol-3-yl)-3-(tetrazolyl)-1,3-propanedione enol (5ClTEP), L-chicoric acid (LCH), quinalizarin, 4-acetylamino-5-hydroxynaphthalene-2,7-disulfonic acid (Y3), 3,-3′,3′-tetramethyl-1,1′-spirobis(indan)-5,5′6,6′-tetrol (TMS), were reported to have a similar binding site near the metal, which is close to the site of 5ClTEP in its crystallographic complex with integrase [18]. However, several lines of experimental evidence indicate that LCH is unlikely to have the same binding mode as that of 5ClTEP. First, the inhibition of HIV-1 integrase by LCH is independent of the concentration of DKA-II, a diketo acid, whereas 5ClTEP and L-731,988 compete with DKA-II [19]. Second, LCH interaction with integrase is independent of viral DNA [20], while the binding of L-731,998 requires the presence of viral

Table 1
Structures and activities of tyrphostins with selected field energies [30]



1-4 5 6

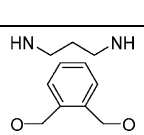
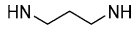
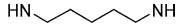
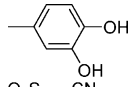
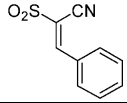
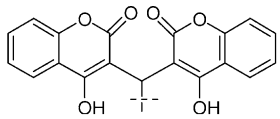
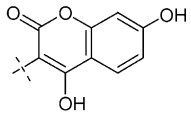
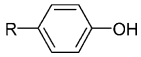
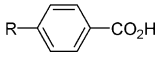
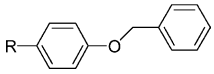
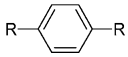
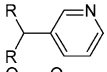
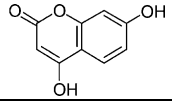
No.	X	R	pIC_{50} Experimental	pIC_{50} Calculated	H^+/447	H^+/682	CH_3/569	HO^-/617	HO^-/1708
1		Br	6.10	6.43	−2.58377	−1.03094	−0.00860107	−0.329263	−1.01074
2		H	5.52	5.63	−5.05219	−2.23952	−0.00903416	3.82683	−0.434793
3		NO_2	5.48	5.40	−5.32686	−0.39736	−0.00779307	4.33585	−0.167574
4		OH	6.35	5.99	−4.98624	−1.07531	−0.00756896	0.665118	−1.06446
5		H	6.00	5.70	−1.11836	0.310261	−0.00697458	0.467678	−0.424470
6		H	5.33	4.63	−3.81798	−3.35653	−0.00654948	0.170703	−1.88378

Table 2
Structures and activities of coumarins with selected field energies [31]

R = 
1-4


5

No.	Structure	pIC_{50}		$H^+/447$	$H^+/682$	$CH_3/569$	$HO^-/617$	$HO^-/1708$
		Experimental	Calculated					
1		3.87	3.99	−3.48471	−7.76964	−0.00873423	3.75764	−0.912846
2		4.24	4.10	−30.00	−30.00	−0.00915313	19.0440	7.57106
3		4.86	5.13	−2.66068	−2.01978	−0.00768280	2.12762	−0.669191
4		5.82	5.65	−3.48471	−5.87255	−0.0130689	4.85467	−1.30899
5		4.05	4.02	−6.03003	−8.32631	−0.00859463	4.37590	−0.970154
6		3.52	3.88	−2.62444	−0.557371	−0.00411248	2.34609	−0.326702

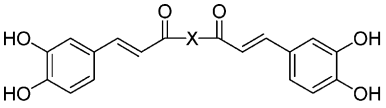
DNA [21]. It should also be noted that Sotriffer et al. [18], restricted their docking region close to the metal and the other nearby binding site, corresponding to the binding site of Y3 in ASV, was not completely included. In addition, the missing solvation parameter of the metal ion might lead to incorrect scoring in docking [22]. In this paper, cluster analysis and three-dimensional QSAR modeling were performed to explore HIV-1 integrase inhibition by 11 classes of inhibitors. The significance of the QSAR models was evaluated using cross-validation, randomization tests and an external test set of five IN inhibitors that have been previously examined in docking studies and for which two crystal structures in complex with integrase enzymes are known.

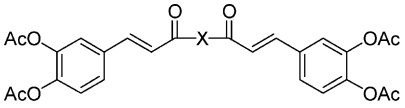
2. Methods

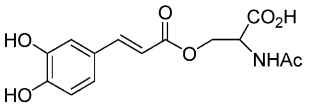
2.1. Cluster analysis

Cluster analysis is a statistical method to partition a data set into classes or categories consisting of similar elements.

Table 3
Structures and activities of chicoric acids with selected field energies [32]


1-3


4-5


6

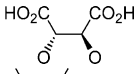
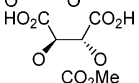
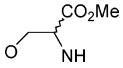
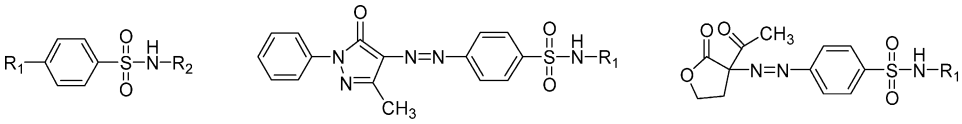
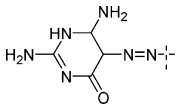
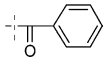
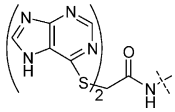
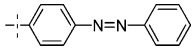
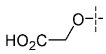
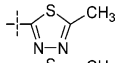
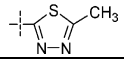
No.	X	pIC_{50}		$H^+/447$	$H^+/682$	$CH_3/569$	$HO^-/617$	$HO^-/1708$
		Experimental	Calculated					
1		5.96	5.74	−30.00	−30.00	−0.00734496	28.9723	18.4908
2		4.42	4.64	−5.26094	−2.43362	−0.00666475	2.75172	−1.00541
3		6.19	6.27	−2.25662	−0.129494	−0.00632954	−2.21429	−0.89538
4		5.00	5.30	−30.00	−30.00	−0.00986004	30.00	16.3790
5		5.60	5.63	−5.21333	−3.59458	−0.00929654	2.25618	−1.30530
6		3.48	3.49	−29.6118	−30.00	−0.00454307	15.0614	7.65278

Table 4
Structures and activities of sulfonamides with selected field energies [33]

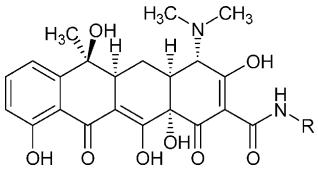


1-4 5 6

No.	R_1	R_2	pIC_{50}		$H^+/447$	$H^+/682$	$CH_3/569$	$HO^-/617$	$HO^-/1708$
			Experimental	Calculated					
1			3.92	4.09	−3.70079	−0.177622	−0.00725496	4.41617	−1.55252
2			4.54	4.91	−1.54010	−5.36377	−0.00577128	−2.83745	−1.40647
3			3.85	3.69	−3.27596	−4.57242	−0.00902152	3.12653	−3.01713
4			5.09	4.57	−2.21685	−3.52806	−0.00548112	0.000872135	−0.694307
5			3.87	4.24	−4.75554	−2.11866	−0.0070374	2.42247	−2.17126
6			3.61	3.42	−6.92364	−0.950375	−0.00507927	3.46164	−2.42085

Hierarchical cluster analysis (HCA) was performed in the Cerius2 [23] program based on descriptors that were calculated for all inhibitors in ionization states appropriate for neutral aqueous solution in the MOE [24] program. These descriptors include approximately 200 topological and geometric descriptors representing both shape and charge distribution, and thus, generally reflect structural diversity. In cluster analysis, the central molecule based on

Table 5
Structures and activities of tetracyclines with selected field energies [34]



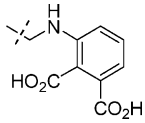
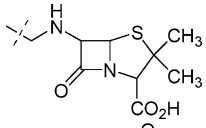
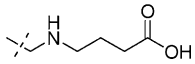
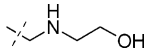
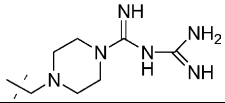
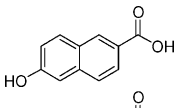
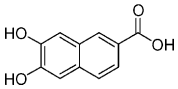
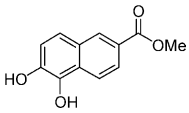
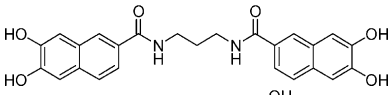
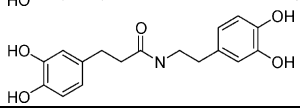
No.	R	pIC_{50}		$H^+/447$	$H^+/682$	$CH_3/569$	$HO^-/617$	$HO^-/1708$
		Experimental	Calculated					
1	H	3.69	4.00	−5.98974	−8.27870	−0.00967383	4.68879	−1.66182
2		5.77	5.65	−30.00	−30.00	−0.011534	30.00	15.6555
3		5.66	5.30	−30.00	−30.00	−0.0112631	16.8792	6.56058
4		5.72	6.26	−30.00	−30.00	−0.0110681	15.1621	7.27514
5		5.29	5.29	0.862480	−11.5565	−0.0121365	−0.995964	−1.91866
6		6.05	5.85	−2.41165	−6.82476	−0.0109239	−0.918494	−2.43139

Table 6
Structures and activities of arylamides and naphthalene-based compounds with selected field energies [35]

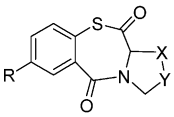
No.	Structure	pIC_{50}		CH_3/122	CH_3/124	CH_3/790	CH_3/875	HO^-/457
		Experimental	Calculated					
1		3.76	3.68	−0.00188494	−0.00764501	−0.00956500	9.14499	20.2327
2		5.27	5.16	−0.00319648	−0.00703669	−0.00895679	30.00	24.6195
3		4.23	4.55	−0.00377226	−0.0114547	−0.00953412	30.00	−0.926916
4		6.01	6.20	−0.00800967	−0.0236869	−0.0217271	30.00	30.00
5		4.48	4.76	−0.00482476	−0.0279633	−0.0125376	30.00	30.00

descriptor values was selected from each of 11 classes of inhibitors to represent the whole class. The 11 class representative HIV-1 integrase inhibitors were clustered into two groups [16]. Clustering was also applied as a technique to select training compounds that maximized structure diversity as reflected by the complete set of structural descriptors. In group 1, six compounds from each of five classes were selected; in group 2, five compounds from each of six classes were selected. This procedure placed 30 structures from each cluster into training sets to develop models.

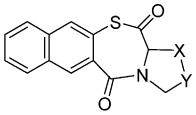
2.2. Molecular field analysis (MFA)

Molecular field analysis (MFA) is a method implemented in the Cerius2 program that is, similar to CoMFA. Although HIV-1 integrase inhibitors in a cluster have a variety of structures, they might have the same binding site. Therefore, it is reasonable to perform a three-dimensional QSAR study for each cluster of HIV-1 integrase inhibitors. Three-dimensional QSAR studies require alignment of all structures as they might be expected to interact with a common binding site. Minimum energy conformations of the training set inhibitors obtained from a random incremental pulse search (RIPS) [25] search in MOE using the MMFF94 [26] force field were used as input structures for compound alignment. Crystallographic structures of the training set inhibitors in complex with HIV IN are not currently publicly available, thus low energy structures are the most reasonable starting point in the absence of additional information regarding the biologically active conformation. Field fit structural alignment was performed in Cerius2 to maximize the overlap between potential fields around inhibitors. A grid spacing of 2.00 Å and three different probes, CH_3, H^+, and HO^- were used to create the fields and simulate van der Waals, electrostatic and hydrogen bonding interactions, respectively. Energies outside the range of −30 to 30 kcal were truncated. Genetic function

Table 7
Structures and activities of thiatholothiazepines with selected field energies [36]



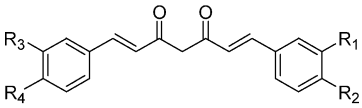
1-4



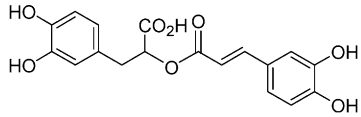
5

No.	R	–X–Y–	pIC_{50}		CH_3/122	CH_3/124	CH_3/790	CH_3/875	HO^-/457
			Experimental	Calculated					
1	H	–S–CH_2–	3.96	3.78	−0.00411189	−0.00985801	−0.0136887	−0.800902	−4.27639
2	Cl	–CH_2–S–	3.87	3.82	−0.00402343	−0.0098176	−0.0136803	3.35465	−3.54594
3	OMe	–CH_2–S–	3.67	3.82	−0.00441444	−0.0121311	−0.0159894	16.1312	−8.31920
4	H	–CH_2–S(O)–	3.72	3.58	−0.00414848	−0.0098983	−0.015648	−0.82828	−5.11926
5		–CH_2–S–	4.04	4.15	−0.00384319	−0.0154405	−0.0173733	18.8841	30.00

Table 8
Structures and activities of curcumins with selected field energies [37]



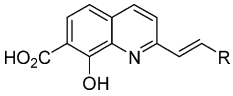
1-4



5

No.	R_1	R_2	R_3	R_4	pIC_{50}		$CH_3/122$	$CH_3/124$	CH3/790	$CH_3/875$	$HO^-/457$
					Experimental	Calculated					
1	H	OH	H	OH	3.92	3.96	−0.00782526	−0.0474237	−0.0110142	0.903056	−1.13112
2	OCH_3	OH	OMe	OH	3.82	4.13	−0.00718164	−0.0610727	−0.0113628	−0.374627	30.00
3	OH	OH	OH	OH	5.22	4.95	−0.00753891	−0.0412673	−0.0106773	0.128889	30.00
4	OMe	OH	OH	OH	4.74	4.93	−0.00685358	−0.0498652	−0.0115998	−0.757776	30.00
5					5.05	5.13	−0.00777328	−0.0461812	−0.0135787	6.99023	30.00

Table 9
Structures and activities of styrylquinolines with selected field energies [38]



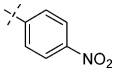
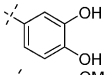
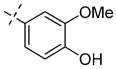
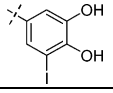
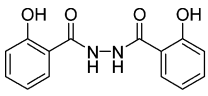
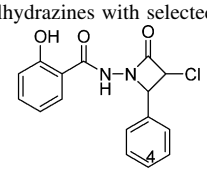
No.	R	pIC_{50}		$CH_3/122$	$CH_3/124$	$CH_3/790$	$CH_3/875$	$HO^-/457$
		Experimental	Calculated					
1		5.28	5.22	−0.0047946	−0.0201365	−0.0124655	30.00	30.00
2		5.92	5.67	−0.00598121	−0.0270987	−0.0117406	30.00	30.00
3		5.62	5.79	−0.00576639	−0.0230287	−0.0115205	30.00	30.00
4		5.55	5.50	−0.00588453	−0.0289423	−0.0116489	30.00	30.00
5		5.40	5.30	−0.00520945	−0.0225424	−0.012913	30.00	30.00

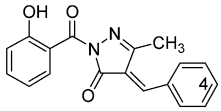
Table 10
Structures and activities of salicylhydrazines with selected field energies [39]



1



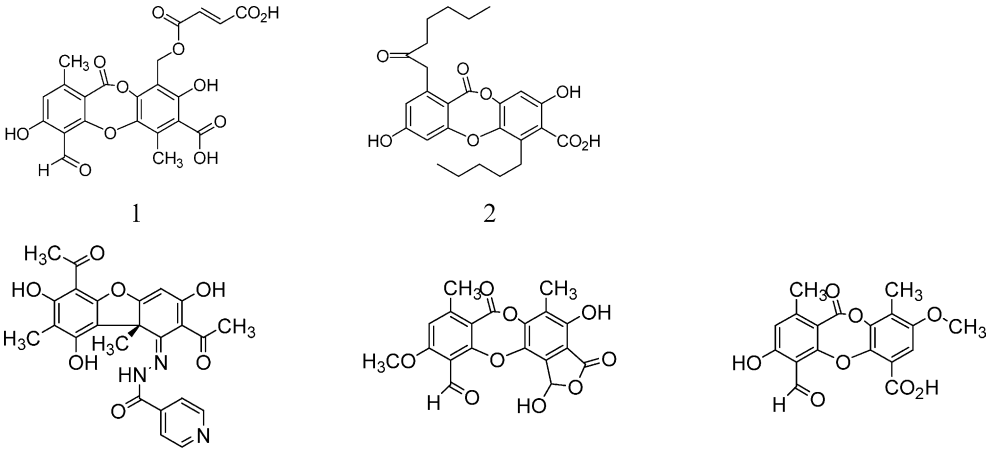
2



3-5

No.	R	pIC_{50}		$CH_3/122$	$CH_3/124$	$CH_3/790$	$CH_3/875$	$HO^-/457$
		Experimental	Calculated					
1		5.68	5.76	−0.00525606	−0.0186915	−0.0110141	30.00	30.00
2	3-OH	3.72	3.70	−0.00475681	−0.0183617	−0.0144746	−0.957104	−1.21848
3	$3\text{-}NO_2$	5.85	5.72	−0.00604808	−0.0234799	−0.0137955	30.00	30.00
4	4-OH	6.22	6.20	−0.00636578	−0.021824	−0.0121627	30.00	30.00
5	$3,4\text{-}(OCH_3)_2$	6.30	5.96	−0.00702488	−0.0320756	−0.0128058	30.00	30.00

Table 11
Structures and activities of depsides with selected field energies [40]



No.	pIC_{50}		CH_3/122	CH_3/124	CH_3/790	CH_3/875	HO^-/457
	Experimental	Calculated					
1	5.31	5.60	−0.00695920	−0.0205828	−0.0225290	30.00	30.00
2	4.41	4.18	−0.0103905	−0.0763859	−0.0259188	30.00	30.00
3	4.17	4.16	−0.00763607	−0.0347257	−0.0250509	30.00	0.246368
4	5.44	5.46	−0.00539255	−0.0150369	−0.0169657	30.00	30.00
5	4.79	4.61	−0.00504124	−0.0167300	−0.0186781	30.00	17.1815

approximation [27] was used for variable selection in model building.

2.3. Model validation

The statistical significance of the models was evaluated using leave-one-out cross-validation and randomization tests in which activity values were randomly associated with a set of molecular field energies prior to GFA. Randomization tests were performed using 99 different random activity assignments in order to demonstrate that the observed correlation coefficient for the non-random activity assignment was meaningful rather than random.

The ability of the models to predict activity is reflected in part by the cross-validation q^2 value, but was also evaluated using an external test set. The five molecules in the external test set include one molecule that has been crystallized in the HIV IN active site (5ClTEP), and one molecule that has been crystallized on the distant side of the ASV IN catalytic loop (Y3). The conformation used in model validation was selected in the same fashion as described for the training set compounds, but was compared to the bound conformation

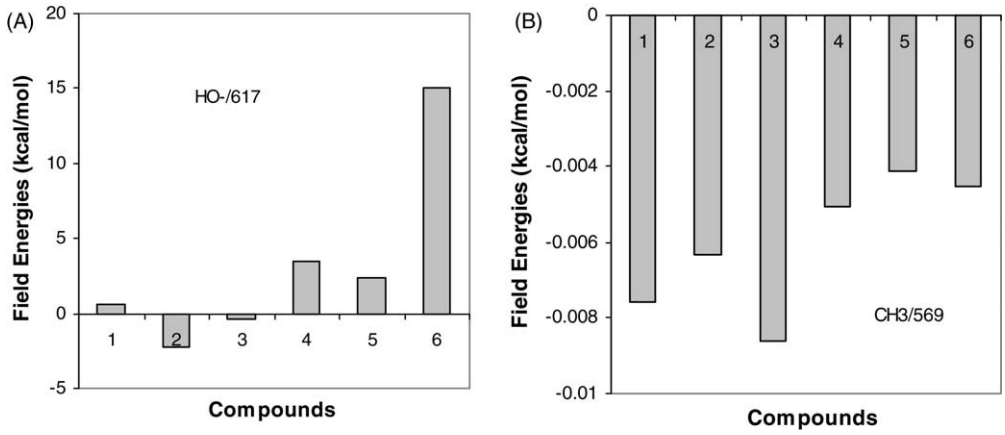


Fig. 2. Comparison of selected field energies of most and least active compounds in cluster 1. Compounds 1–6: Tyr4, Chi3, Tyr1, Sul6, Cou6, Chi6, in order from the most to the least active.

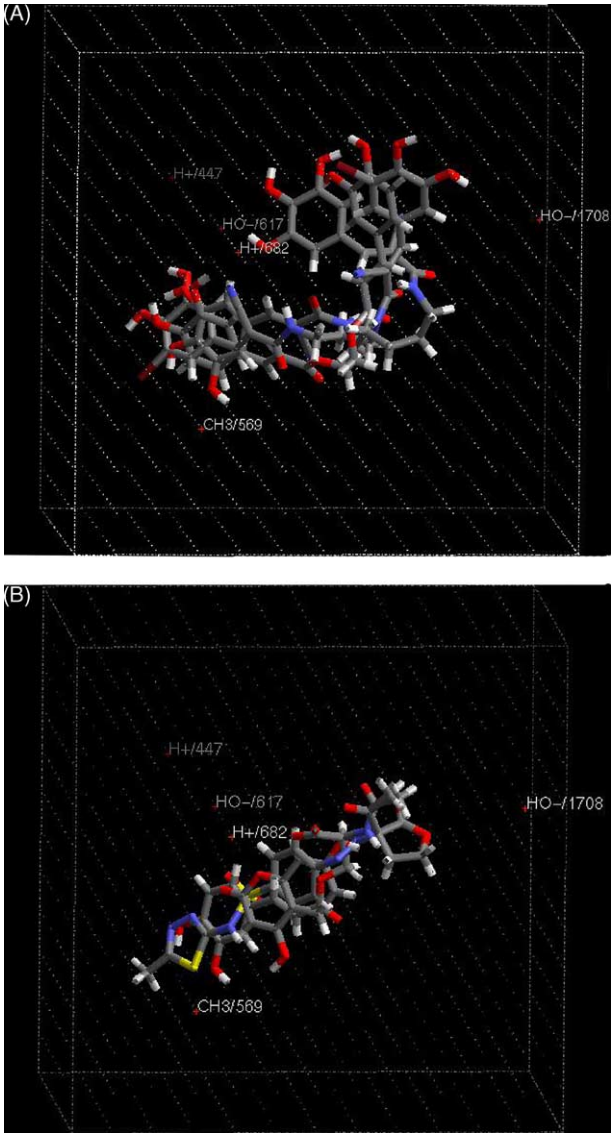


Fig. 3. (A) Superimposed active compounds (Tyr4, Chi3, Tyr1) with selected grid points in cluster 1. (B) Superimposed least active compounds (Sul6, Cou6, Chi6) with selected grid points in cluster 1.

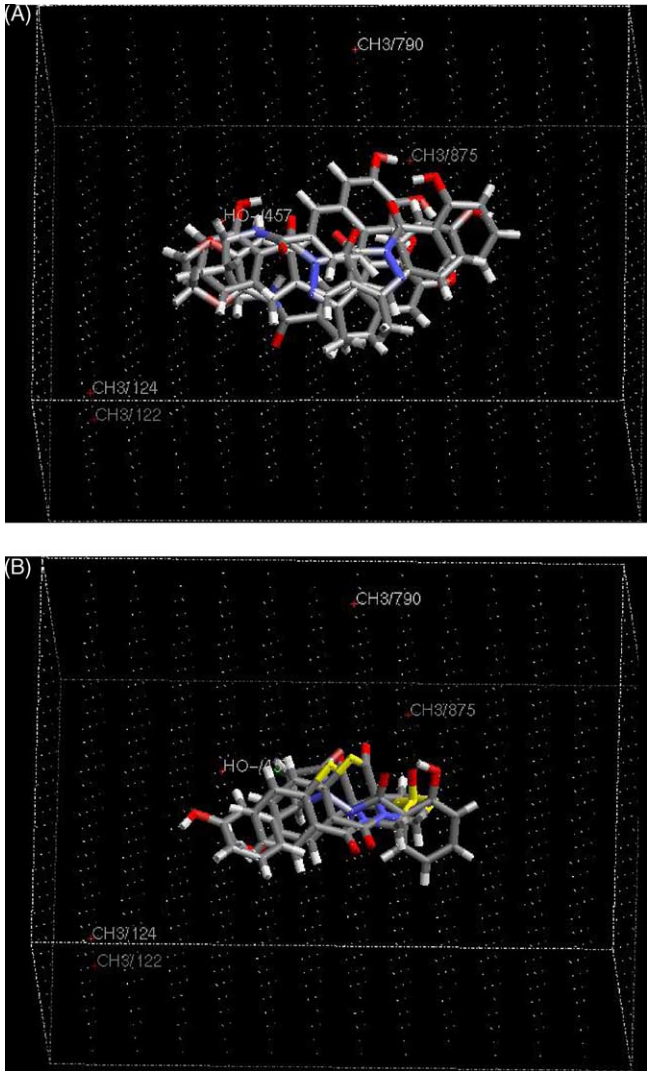


Fig. 5. (A) Superimposed active compounds (Sal5, Sal4, Ary4) with selected grid points in cluster 2. (B) Superimposed inactive compounds (Sal2, Thi4, Thi3) with selected grid points in cluster 2.

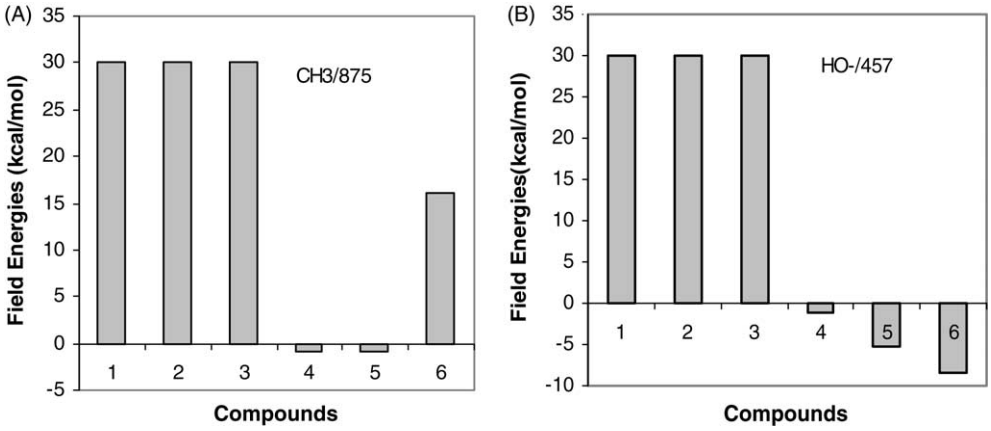
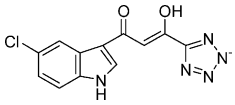
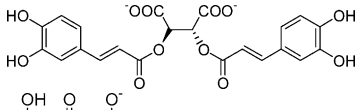
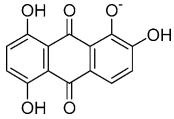
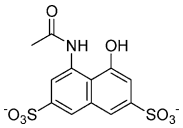
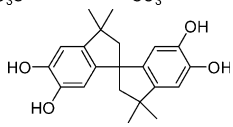


Fig. 4. Comparison of selected field energy of most and least active compounds in cluster 2. Compounds 1–6: Sal5, Sal4, Ary4, Sal2, Thi4, Thi3, in order from the most to the least active.

Table 12
Comparisons of experimental and calculated pIC_{50} of five test structures

Structures	3′-Processing IC_{50} (μM)	Experimental pIC_{50}	MFA prediction (pIC_{50})		Conventional QSAR prediction (pIC_{50})	
			Model 1	Model 2	Model 1	Model 2
5ClTEP* [13] 	2.20	5.66	1.74	**5.38**	**5.84**	5.97
LCH [32] 	1.10	5.96	**6.31**	3.77	5.66	**5.97**
QLZ [41] 	4.00	5.40	1.80	**4.98**	4.19	**5.05**
Y3 [14] 	16.20	4.79	3.00	**4.55**	4.12	**5.03**
TMS [42] 	17.00	4.77	7.68	**5.09**	2.29	**3.71**
Average percent residual			6.1%		7.4%	

*5CITEP, 1-(5-chloroindol-3-yl)-3-(tetrazolyl)-1,3-propanedione enol; LCH, L-chicoric acid; QLZ, quinalizarin; Y3, 4-acetylamino-5-hydroxynaphthalene-2,7-disulfonic acid; TMS, 3,3,3′,3′-tetramethyl-1,1′-spirobis(indan)-5,5′6,6′-tetrol. Residual = (Predicted pIC_{50} − Experimental pIC_{50})/Experimental pIC_{50}. The better predicted pIC_{50}s by two MFA models or two conventional QSAR models were used and highlighted in bold.

from the crystal structure for additional validation of the methods used.

2.4. Results and discussion

Eleven classes of HIV-1 integrase inhibitors were classified into the two clusters from which two MFA models were developed. This division of structures implies that inhibitors from different clusters might have different binding modes or binding sites on the enzyme. Separation of structures into two groups was also necessary to obtain predictive conventional QSAR models [16,17].

Two MFA models were developed, Models 1 and 2 for clusters 1 and 2, respectively (Fig. 1). Compared with our previously reported conventional QSAR models [16], these MFA models have higher r^2 and q^2. In randomization trials, 95 of 99 and 99 of 99 models resulting from GFA had lower r values than the model resulting from non-random activity values. Thus confidence levels for non-random correlations are 95% and 99%. The compounds in the training sets and the energies used to derive the models are shown in Tables 1–11. Five field energies were selected to derive each model. Models 1 and 2 are obviously different. Model 1 is composed of two H^+ energies, two HO^- energies and one CH_3 energy, while Model 2 has four CH_3 energies and only one HO^- energy. The distinction of the two models indicates that the inhibitors in the two groups have different binding modes.

Model 1: $r^2 = 0.90$, $q^2 = 0.85$, non-random confidence = 95%"

$$pIC_{50} = 2.70393 - 0.180054 \times \text{“}H^+/447\text{”} + 0.176222 \times \text{“}H^+/682\text{”} - 452.298 \times \text{“}CH_3/569\text{”} - 0.374541 \times \text{“}HO^-/617\text{”} + 0.565036 \times \text{“}HO^-/1708\text{”}$$

Model 2: $r^2 = 0.95$, $q^2 = 0.91$, non-random confidence = 99%"

$$pIC_{50} = 3.16779 - 669.632 \times \text{“}CH_3/122\text{”} + 60.451 \times \text{“}CH_3/124\text{”} + 103.347 \times \text{“}CH_3/790\text{”} + 0.018706 \times \text{“}CH_3/875\text{”} + 0.026153 \times \text{“}HO^-/457\text{”}$$

In order to visualize the models effectively, the three most and three least active compounds in each group are

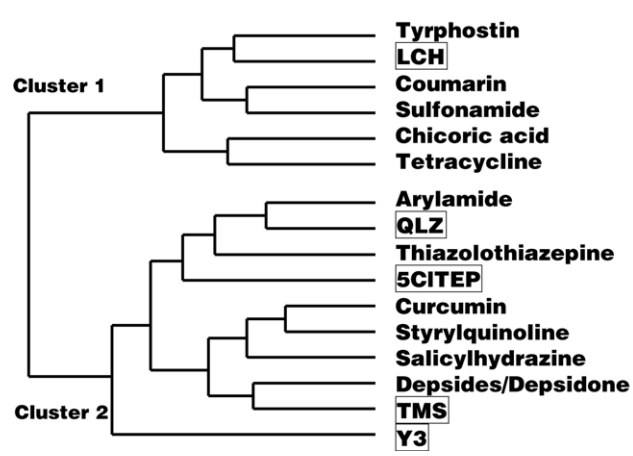


Fig. 6. Cluster analysis of 11 classes of inhibitors plus 5 test compounds. Test compounds are boxed.

superimposed and the selected field energies are highlighted (Figs. 2–5). Through the comparison of these field energies of active and inactive compounds, there are several obvious differences. In Model 1, the unfavorable HO^-/617 has small or negative values for active compounds (compounds Tyr4, Chi3, Tyr1) and large positive values for inactive compounds (compounds Sul6, Cou6, Chi6). CH_3/569 is apparently more negative for the three active compounds than the same energy for the three inactive compounds (Fig. 2).

In Model 2, both CH_3/875 and HO^-/457 have large positive values for active compounds. For inactive compounds, one structure (Thi3) has moderate energies and the other two have small negative values (Fig. 4). These dissimilarities can be rationalized by different volumes occupied by active and inactive compounds as shown in Fig. 5. It suggests that these two positions represent regions that must be occupied in order to have high activities. The other three van der Waals interactions in this model have less obvious trends relative to activities.

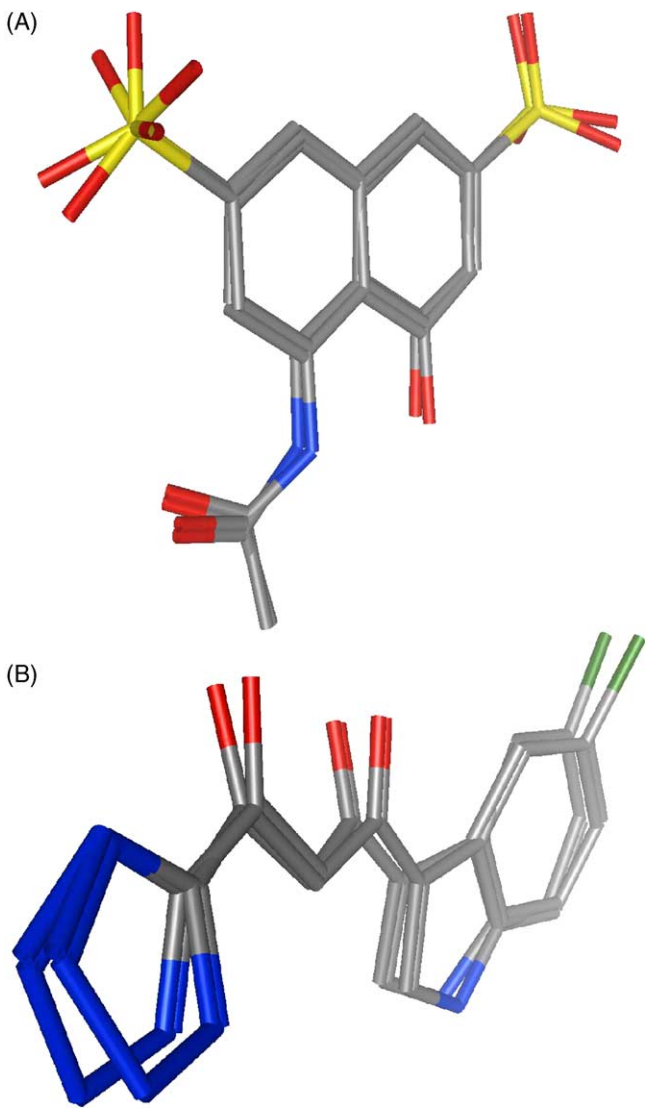


Fig. 7. Comparison of conformations used in model validation and from crystallographic studies of IN complexes for (A) 5ClTEP and (B) Y3.

The two MFA models were employed to calculate the activities of five test inhibitors, which were not used to derive the models. Similar to the conventional QSAR [16], neither MFA Model 1 nor MFA Model 2 can accurately calculate activities for all test structures alone. Model 2 gives correct pIC_{50}s for four structures, but not for LCH. However, Model 1 provides a reasonable prediction for LCH (Table 12). This correlates with the result of cluster analysis (Fig. 6). The five test inhibitors were added to cluster analysis with the 11 representative inhibitors. LCH belongs to cluster 1, from which Model 1 was derived, while the other four structures are members of cluster 2, from which Model 2 was derived.

The conformations resulting from conformational search and subsequent field fit procedures can be compared with bound crystallographic conformations in only two cases, 5ClTEP (PDB entry 1QS4 [13]) and Y3 (PDB entry 1A5V [14]). Fig. 7 shows overlays of the conformation used in model validation on the crystallographic conformation for 5ClTEP (A) and Y3 (B). Root mean square deviations over all heavy atoms were 0.42 and 0.40, respectively. Thus, the conformations used in model validation are very similar to the bound crystallographic conformations.

Table 12 shows the comparison of MFA and the conventional QSAR models for the five test compounds. Unlike MFA results, both conventional QSAR models can predict 5ClTEP and LCH reasonably well; Conventional QSAR Model 2 has good predictions for QLZ and Y3; TMS cannot be well predicted by either conventional QSAR model. MFA models not only give better overall predictions but also can predict activity for test compounds from the cluster they belong to.

3. Conclusions

Several former studies suggested that CoMFA would fail when molecules were very dissimilar [11,28,29], The results from MFA performed in Cerius2 indicate that inhibitors with a diverse set of structures can be analyzed to get predictive models when these inhibitors are classified into appropriate groups through cluster analysis. Both of the cluster analysis and MFA models support the hypothesis of two separate binding sites. These MFA models have higher structure-activity correlations than our previously published conventional QSAR models. The addition of three-dimensional QSAR modeling of HIV-1 integrase inhibitors provides a stronger capability to predict activities of ligands.

Acknowledgments

Support from NIH/NIAID (Grant R15 AI 45984-01) and NSF (STI-9602656, CHE-9708517) are gratefully acknowl-

edged. We thank the Chemical Computing Group for their donation of the MOE program. This work was also supported by the funds from the University of Memphis.

References

[1] E. De Clercq, Toward improved anti-HIV chemotherapy: therapeutic strategies for intervention with HIV infections, J. Med. Chem. 38 (1995) 2491–2517.
[2] W.E.J. Robinson, HIV integrase: the next target? Infect. Med. 15 (1998) 129–137.
[3] N. Neamati, Structure-based HIV-1 integrase inhibitor design: a future perspective, Expert Opin. Invest. Drugs 10 (2001) 281–296.
[4] S.D. Young, Inhibition of HIV-1 integrase by small molecules: the potential for a new class of AIDS chemotherapeutics, Curr. Opin. Drug Discov. Dev. 4 (2001) 402–410.
[5] I.J. Chen, N. Neamati, A.D. MacKerell Jr., Structure-based inhibitor design targeting HIV-1 integrase, Curr. Drug Targets Infect. Disord. 2 (2002) 217–234.
[6] N. Neamati, Patented small molecule inhibitors of HIV-1 integrase: a 10-year saga, Expert Opin. Ther. Patents 12 (2002) 709–724.
[7] A.L. Parrill, HIV-1 integrase inhibition: binding sites, structure activity relationships and future perspectives, Curr. Med. Chem. 10 (2003) 1811–1824.
[8] C. Hansch, A quantitative approach to biochemical structure-activity relationships, Acc. Chem. Res. 2 (1969) 232–239.
[9] A.R. Katritzky, E.V. Gordeeva, Traditional topological indices vs. electronic geometrical and combined molecular descriptors in QSAR/QSPR research, J. Chem. Inf. Comput. Sci. 33 (1993) 835–857.
[10] H. Van de Waterbeemd, Recent progress in QSAR-technology, Drug Des. Discov. 9 (1993) 277–285.
[11] R.D. Cramer III, D.E. Patterson, J.D. Bunce, Comparative molecular field analysis (CoMFA). Part 1. Effect of shape on binding of steroids to carrier proteins, J. Am. Chem. Soc. 110 (1998) 5959–5967.
[12] C. Marchand, X. Zhang, G.C. Pais, K. Cowansage, N. Neamati, T.R. Burke Jr., Y. Pommier, Structural determinants for HIV-1 integrase inhibition by beta-diketo acids, J. Biol. Chem. 277 (2002) 12596–12603.
[13] Y. Goldgur, R. Craigie, G.H. Cohen, T. Fujiwara, T. Yoshinaga, T. Fujishita, H. Sugimoto, T. Endo, H. Murai, D.R. Davies, Structure of the HIV-1 integrase catalytic domain complexed with an inhibitor: a platform for antiviral drug design, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 13040–13043.
[14] J. Lubkowski, F. Yang, J. Alexandratos, A. Wlodawer, H. Zhao, T.R. Burke Jr., N. Neamati, Y. Pommier, G. Merkel, A.M. Skalka, Structure of the catalytic domain of avian sarcoma virus integrase with a bound HIV-1 integrase-targeted inhibitor, Proc. Natl. Acad. Sci. U.S.A. 95 (1998) 4831–4836.
[15] V. Molteni, J. Greenwald, D. Rhodes, Y. Hwang, W. Kwiatkowski, F.D. Bushman, J.S. Siegel, S. Choe, Identification of a small-molecule binding site at the dimer interface of the HIV integrase catalytic domain, Acta Crystallogr. D Biol. Crystallogr. 57 (2001) 536–544.
[16] H. Yuan, A.L. Parrill, QSAR studies of HIV-1 integrase inhibition, Bioorg. Med. Chem. 10 (2002) 4169–4183.
[17] M.T. Makhija, V.M. Kulkarni, QSAR of HIV-1 integrase inhibitors by genetic function approximation method, Bioorg. Med. Chem. 10 (2002) 1483–1497.
[18] C.A. Sotriffer, H. Ni, J.A. McCammon, Active site binding modes of HIV-1 integrase inhibitors, J. Med. Chem. 43 (2000) 4109–4117.
[19] J.A. Grobler, K. Stillmock, B. Hu, M. Witmer, P. Felock, A.S. Espeseth, A. Wolfe, M. Egbertson, M. Bourgeois, J. Melamed, J.S. Wai, S. Young, J. Vacca, D.J. Hazuda, Diketo acid inhibitor mechanism and HIV-1 integrase: implications for metal binding in the active site of phosphotransferase enzymes, Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 6661–6666.
[20] K. Zhu, M.L. Cordeiro, J. Atienza, W.E. Robinson Jr., S.A. Chow, Irreversible inhibition of human immunodeficiency virus type 1 integrase by dicaffeoylquinic acids, J. Virol. 73 (1999) 3309–3316.
[21] A.S. Espeseth, P. Felock, A. Wolfe, M. Witmer, J. Grobler, N. Anthony, M. Egbertson, J.Y. Melamed, S. Young, T. Hamill, J.L. Cole, D.J. Hazuda, HIV-1 integrase inhibitors that compete with the target DNA substrate define a unique strand transfer conformation for integrase, Proc. Natl. Acad. Sci. U.S.A. 97 (2000) 11244–11249.
[22] G.M. Morris, D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, A.J. Olson, Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function, J. Comp. Chem. 19 (1998) 1639–1662.
[23] Cerius2 Program, Molecular Simulations Inc., 1999.
[24] MOE Program, Chemical Computing Group Inc., 2002.
[25] D.M. Ferguson, D.J. Raber, A new approach to probing conformational space with molecular mechanics: random incremental pulse search, J. Am. Chem. Soc. 111 (1989) 4371–4378.
[26] T.A. Halgren, Merck molecular force field. Part I. Basis, form, scope, parameterization, and performance of MMFF94*, J. Comp. Chem. 17 (1996) 490–519.
[27] D. Rogers, A.J. Hopfinger, Application of genetic function approximation to quantitative structure-activity relationships and quantitative structure-property relationships, J. Chem. Inf. Comput. Sci. 34 (1994) 854–866.
[28] K. Raghavan, J.K. Buolamwini, M.R. Fesen, Y. Pommier, K.W. Kohn, J.N. Weinstein, Three-dimensional quantitative structure-activity relationship (QSAR) of HIV integrase inhibitors: a comparative molecular field analysis (CoMFA) study, J. Med. Chem. 38 (1995) 890–897.
[29] J.K. Buolamwini, H. Assefa, CoMFA and CoMSIA 3D QSAR and docking studies on conformationally-restrained cinnamoyl HIV-1 integrase inhibitors: exploration of a binding mode at the active site, J. Med. Chem. 45 (2002) 841–852.
[30] A. Mazumder, A. Gazit, A. Levitzki, M. Nicklaus, J. Yung, G. Kohlhagen, Y. Pommier, Effect of tyrphostins, protein kinase inhibitors, on Human immunodeficiency virus type 1 integrase, Biochemistry 35 (1995) 15111–15122.
[31] H. Zhao, N. Neamati, H. Hong, A. Mazumder, S. Wang, S. Sunder, G.W. Milne, Y. Pommier, T.R. Burke Jr., Coumarin-based inhibitors of HIV integrase, J. Med. Chem. 40 (1997) 242–249.
[32] Z. Lin, N. Neamati, H. Zhao, Y. Kiryu, J.A. Turpin, C. Aberham, K. Strebel, K. Kohn, M. Witvrouw, C. Pannecouque, Z. Debyser, E. De Clercq, W.G. Rice, Y. Pommier, T.R. Burke Jr., Chicoric acid analogues as HIV-1 integrase inhibitors, J. Med. Chem. 42 (1999) 1401–1414.
[33] M.C. Nicklaus, N. Neamati, H. Hong, A. Mazumder, S. Sunder, J. Chen, G.W. Milne, Y. Pommier, HIV-1 integrase pharmacophore: discovery of inhibitors through three-dimensional database searching, J. Med. Chem. 40 (1997) 920–929.
[34] N. Neamati, H. Hong, S. Sunder, G.W.A. Milne, Y. Pommier, Potent inhibitors of human immunodeficiency virus type 1 integrase: identification of a novel four-point pharmacophore and tetracyclines as novel inhibitors, Mol. Pharmacol. 52 (1997) 1041–1055.
[35] H. Zhao, N. Neamati, A. Mazumder, S. Sunder, Y. Pommier, T.R. Burke Jr., Arylamide inhibitors of HIV-1 integrase, J. Med. Chem. 40 (1997) 1186–1194.
[36] N. Neamati, J.A. Turpin, H.E. Winslow, J.L. Christensen, K. Williamson, A. Orr, W.G. Rice, Y. Pommier, A. Garofalo, A. Brizzi, G. Campiani, I. Fiorini, V. Nacci, Thiazolothiazepine inhibitors of HIV-1 integrase, J. Med. Chem. 42 (1999) 3334–3341.
[37] A. Mazumder, N. Neamati, S. Sunder, J. Schulz, H. Pertz, E. Eich, Y. Pommier, Curcumin analogs with altered potencies against HIV-1 integrase as probes for biochemical mechanisms of drug action, J. Med. Chem. 40 (1997) 3057–3063.
[38] F. Zouhiri, J.F. Mouscadet, K. Mekouar, D. Desmaele, D. Savoure, H. Leh, F. Subra, M. Le Bret, C. Auclair, J. d'Angelo, Structure-activity

relationships and binding mode of styrylquinolines as potent inhibitors of HIV-1 integrase and replication of HIV-1 in cell culture, J. Med. Chem. 43 (2000) 1533–1540.

[39] N. Neamati, H. Hong, J.M. Owen, S. Sunder, H.E. Winslow, J.L. Christensen, H. Zhao, T.R. Burke Jr., G.W. Milne, Y. Pommier, Salicylhydrazine-containing inhibitors of HIV-1 integrase: implication for a selective chelation in the integrase active site, J. Med. Chem. 41 (1998) 3202–3209.

[40] N. Neamati, H. Hong, A. Mazumder, S. Wang, S. Sunder, M.C. Nicklaus, G.W. Milne, B. Proksa, Y. Pommier, Depsides and depsidones as inhibitors of HIV-1 integrase: discovery of novel inhibitors through 3D database searching, J. Med. Chem. 40 (1997) 942–951.

[41] C.M. Farnet, B. Wang, J.R. Lipford, F.D. Bushman, Differential inhibition of HIV-1 preintegration complexes and purified integrase protein by small molecules, Proc. Natl. Acad. Sci. U.S.A. 93 (1996) 9742–9747.

[42] V. Molteni, D. Rhodes, K. Rubins, M. Hansen, F.D. Bushman, J.S. Siegel, A new class of HIV-1 integrase inhibitors: the 3,3,3′ 3′-tetramethyl-11′-spirobi(indan)-5,5′,6,6′-tetrol family, J. Med. Chem. 43 (2000) 2031–2039.