

# Homology Model-Guided 3D-QSAR Studies of HIV-1 Integrase Inhibitors

Horrick Sharma,<sup>†</sup> Xiaolin Cheng,<sup>‡,§</sup> and John K. Buolamwini<sup>\*,†</sup>

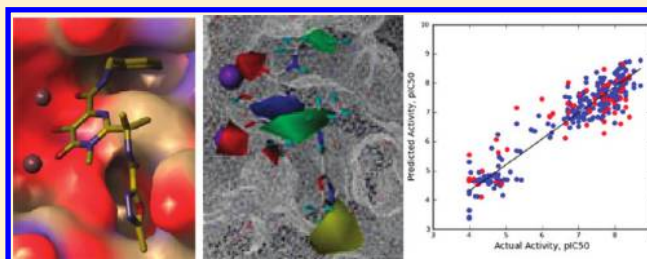
<sup>†</sup>Department of Pharmaceutical Sciences, College of Pharmacy, University of Tennessee Health Science Center, Memphis, Tennessee 38163, United States

<sup>‡</sup>UT/ORNL Center for Molecular Biophysics, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, United States

<sup>§</sup>Department of Biochemistry and Cellular and Molecular Biology, University of Tennessee, Knoxville, Tennessee 37996, United States

## Supporting Information

**ABSTRACT:** In the present study, we report the exploration of binding modes of potent HIV-1 integrase (IN) inhibitors MK-0518 (raltegravir) and GS-9137 (elvitegravir) as well as chalcone and related amide IN inhibitors we recently synthesized and the development of 3D-QSAR models for integrase inhibition. Homology models of DNA-bound HIV-1 IN were constructed on the basis of the X-ray crystal structure of the foamy virus IN-DNA complex (PDB ID: 3L2T) and used for docking. The binding modes of raltegravir and elvitegravir in our homology models are in accordance with those in the foamy virus structure revealing interactions important for inhibitor-IN binding. To gain further insights into the structural requirements for IN inhibition, three-dimensional quantitative structure activity relationship (3D-QSAR) studies were conducted using raltegravir, elvitegravir, and their analogs; our synthesized 3-keto salicylic acid IN inhibitor series; as well as other structurally related HIV-1 IN inhibitors. In the first part of the study with 103 compounds, atom-fit alignments, I and II, and docking-based alignment, III, were used to develop 3D-QSAR models 1, 2, and 3, respectively, each comprising comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) 3D-QSARs. This initial analysis indicated that the docking-based (structure-based) model 3 performed better than the atom-fit (ligand-based) models 1 and 2, in terms of statistical significance and robustness. Thus, the docking-based alignment was then subsequently used with an expanded data set of 296 compounds for building a more comprehensive 3D-QSAR, model 4. Model 4 afforded good  $q^2$  values of 0.70 and 0.75 for CoMFA and CoMSIA 3D-QSARs, respectively, and showed good predictive performance on an external validation test set of 59 compounds with predictive  $r^2$  values up to 0.71. The HIV IN-DNA homology model of biological relevance and the comprehensive 3D-QSAR models developed in the present study provide insights and new predictive tools for structure-based design and optimization of IN inhibitors.



## ■ INTRODUCTION

Highly active antiretroviral therapy (HAART) cocktails consisting of reverse transcriptase<sup>1</sup> and protease inhibitors<sup>2</sup> have been successful in reducing viral load in acquired immunodeficiency disease (AIDS) patients but fail to eradicate the human immunodeficiency virus (HIV).<sup>3,4</sup> Although the approval of the entry inhibitor enfuvirtide<sup>5</sup> and cell surface receptor antagonist maraviroc<sup>6</sup> has offered different therapeutic options, the persistence of the latent virus and the rapid emergence of drug resistance<sup>7,8</sup> underscore an urgent need to target other enzymes central to HIV replication. In this regard, HIV-1 integrase (IN) has emerged as a promising new target with the introduction of raltegravir for the clinical management of AIDS, which is still a devastating pandemic.

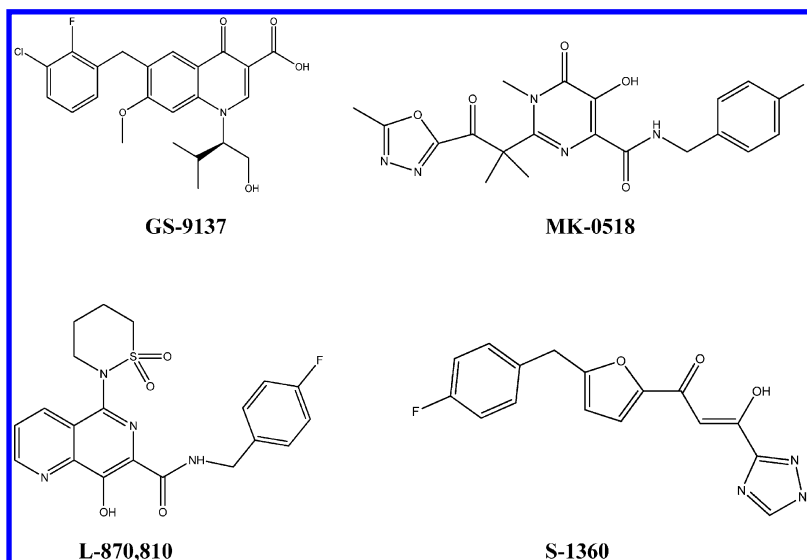
IN is an important enzyme in HIV replication with apparently no human counterpart,<sup>9,10</sup> making it an attractive therapeutic target. IN catalyzes the integration of the viral DNA into the host genome by two catalytic steps, namely, 3'-processing

(3'-P) and strand transfer (ST). In the 3'-P reaction, which occurs in the cytoplasm of an infected cell within a pre-integration complex (PIC), site-specific endonucleolytic activity removes GT dinucleotide from the long terminal repeat region (LTR) of the viral genome. The terminal nucleotides removed are adjacent to the highly conserved CA dinucleotide at each 3' end of the proviral DNA. Chemically, the reaction involves the activation of a water molecule by the Mg<sup>2+</sup> ion, which is involved in chelation with Glu152 and Asp64; the water then acts as a nucleophile and attacks the phosphodiester bond of the viral DNA to generate free 3'-OH ends. The primed viral DNA in association with the viral protein, in a preintegration complex, then translocates into the nucleus of the host cell. Within the nucleus, viral IN catalyzes the next step of integration, viz., the ST. In the

Received: October 12, 2011

Published: January 18, 2012

Chart 1. Representative Structures of Potent HIV-1 IN Inhibitors



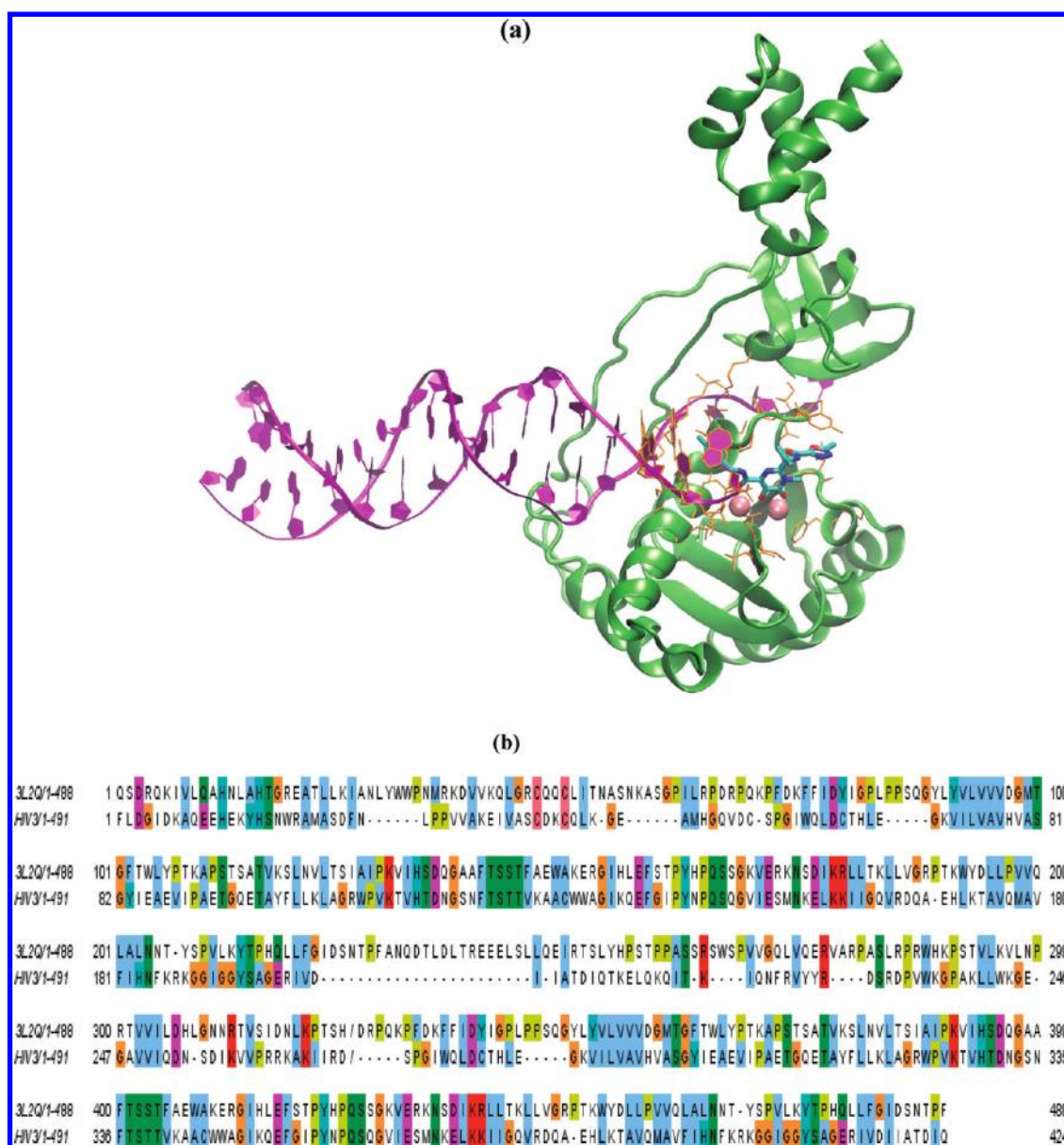
process of ST, IN introduces a staggered cut in the target DNA at sites separated by five base pairs, resulting in overhangs with 5' phosphorylated ends. IN then performs the ST, also called the 3' end joining reaction wherein the primed 3' ends of the viral DNA are joined to the 5' ends of the host/target DNA.

$\beta$ -Diketo acids and bioisosteres thereof are the most advanced class to have shown promising HIV-1 IN inhibitory activity. The approved raltegravir,<sup>11</sup> which belongs to the class of *N*-methyl-4-hydroxypyrimidinone carboxamides,<sup>12,13</sup> diketotriazole S-1360,<sup>14</sup> and naphthyridine carboxamide L-870,810<sup>15</sup> all have three metal-chelating groups in a coplanar conformation mimicking the  $\beta$ -diketo acid motif (Chart 1). Promising IN inhibitors were thus designed as bioisosteres of the diketo acid functionality. However, the advent of monoketo quinolone second-generation HIV-1 IN inhibitors<sup>16,17</sup> showed that the presence of even two coplanar chelating moieties, represented by the  $\beta$ -ketone and a carboxylic acid in elvitegravir (GS-9137, Chart 1), is sufficient for potent clinically relevant IN inhibition. Although it is established that potent IN inhibitors have two common structural features, viz., a hydrophobic benzyl group and a chelating motif capable of binding two  $Mg^{2+}$  ions, a detailed understanding of the structural features essential for HIV-IN inhibitory activity still evades researchers. Therefore, it is important to develop quantitative structure–activity relationships (QSARs) that combine the most advanced first- and second-generation inhibitors to arrive at a comprehensive and more unified model for IN inhibition. Although QSAR studies on IN inhibitors have been reported,<sup>18–23</sup> there are none combining the two classes, viz., the *N*-methyl-4-hydroxypyrimidinone-carboxamide and 4-quinolone-3-carboxylic acid as represented by the approved raltegravir, the clinically advanced elvitegravir, and their analogs.

Despite two decades of tremendous effort leading to promising IN inhibitors,<sup>24–26</sup> their binding modes at the IN active site remain elusive. The absence of a complete DNA-bound HIV IN crystal structure has presented a significant challenge for the structure-based design of potent IN inhibitors. Although X-ray structures of individual human HIV domains and their combinations (core domain and N terminal domain;

core domain and C terminal domain; and core domain with an inhibitor (SCITEP)) have been reported,<sup>27–33</sup> most of these structures have an unresolved flexible loop. Moreover, there are differences in conformations of the active site residues in the two structures that show a completely resolved flexible loop (chain B of 1BIS and chain C of 1BL3).<sup>34,35</sup> The absence of a 3D structure of IN and viral DNA along with the two-divalent metals bound in the active site makes the structural information hitherto available limited and may be far from reality. As shown in the Figure 1a, the inhibitor binding site is located at the interface between the catalytic core domain (CCD) and C-terminal domain (CTD), involving contacts with the 3' terminus of the viral DNA. Therefore, the correct binding mode of drug molecules could not have been predicted on the basis of previous partial IN structures. It has been shown that the full-length IN–DNA–inhibitor complex model would be required to gain a better understanding of how the inhibitors act against IN.

The objective of this study was to obtain new composite 3D-QSAR models incorporating important physicochemical parameters from potent IN inhibitory compounds to guide our lead optimization efforts on a novel series of 3-keto salicylic acid IN inhibitors recently discovered in our laboratory.<sup>23</sup> To do this, we first developed a new homology model of HIV-1 IN in complex with DNA and inhibitors, based on the recently published crystal structure of the foamy virus IN–DNA complex (PDB 3L2T and 3L2U)<sup>36</sup> for use in structure-based 3D-QSAR studies. Although homology models using the foamy virus IN–DNA complex as a template were recently reported by Tang et al.<sup>37</sup> and Krishnan et al.,<sup>38</sup> they were not without shortcomings. The model developed by Tang et al. consisted of only the CCD–DNA complex and may not be a true representation considering the absence of CTD and NTD, both of which have been implicated in viral DNA binding. Although, the model reported by Krishnan et al. involved all three domains and the viral DNA, no direct contacts of raltegravir with mutant residues Gln148 and Asn155, that are involved in viral resistance, are observed. Docking simulations were then performed to explore binding interactions of inhibitors with



**Figure 1.** (a) PFV IN active site. Protein is shown in cartoon representation (green). Inhibitor RLT is shown in licorice representation. DNA is shown in ribbon representation (purple). Two magnesium ions are shown as mauve spheres. The residues within 6 Å of RLT are shown as orange lines. Only one protein chain is shown for clarity. (b) Amino acid sequence alignment of PFV and HIV-1 INs. Residues are colored according to the clustal color scheme.

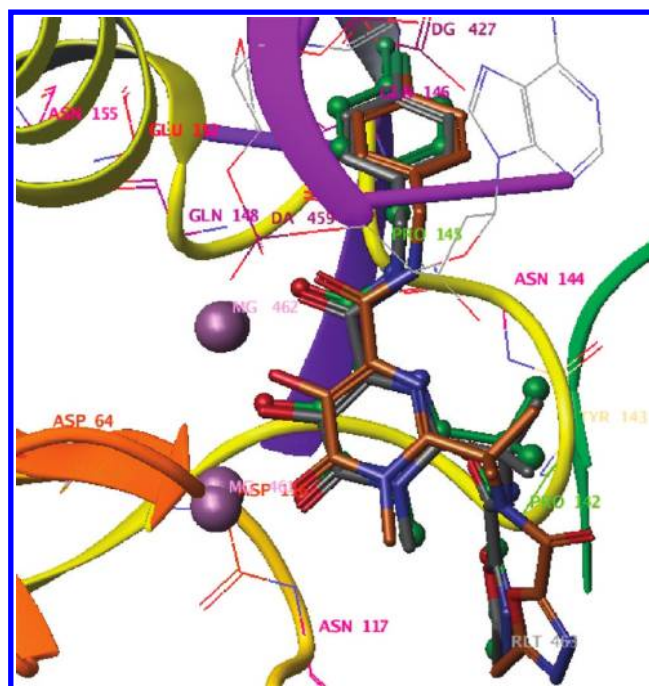
## ■ MATERIALS AND METHODS

**Homology Modeling.** The homology models of the DNA bound HIV-1 IN were constructed with the MODELER program<sup>39,40</sup> (v8.0) based on the foamy virus IN–DNA complex X-ray crystal structures (3L2T and 3L2U).<sup>36</sup> The sequence alignment between HIV-1 and PFV IN was performed using T-Coffee<sup>41</sup> with manual adjustments on two loop regions: loop 1 connecting the N-terminal domain (NTD) and catalytic core domain (CCD) and loop 2 connecting CCD and the C-terminal domain (CTD). The overall alignment is very similar to that reported<sup>36</sup> except for a few residues in the two loop regions (see Figure 1b for the final sequence alignment). As shown in Figure 1a, both loops are long and directly interact with the DNA strands, and thus extra care is needed to build them correctly for dynamical exploration of the



system, whereas their precise structures may not be essential for the docking and QSAR studies presented below, as these loops are located on the backside of the inhibitor binding site. Due to the low homology between these two loops and their counterparts in PFV, the philosophy adopted here is to allow them to interact with the DNA as much (favorably) as possible. Residues in the two loop terminal regions were skipped in alignment to allow the extensive DNA interactions with the modeled loops as in the PFV IN complexes. Furthermore, two basic residues, His51 and Lys211, were made to align with Arg117 and His304, respectively, such that both can interact more favorably with the negatively charged DNA. The loop modeling module as implemented in MODELER was further used to refine the loop coordinates based on a scoring function and optimization schedule. We believe this sequence plus structural alignment represents a more effective way to construct local low-homology regions than a purely sequence-based alignment. For the ligands raltegravir and elvitegravir, we assume that they interact similarly with the target and the template, so we relied on MODELER to extract and satisfy distance restraints automatically during model building. Conserved interactions involving residues Tyr212 and Pro214 and the two ligands were confirmed and restrained during the subsequent structural optimization. The final models were selected mainly in terms of both the Discrete Optimized Protein Energy (DOPE) assessment score and the model's stereochemical parameters, which were evaluated by means of the PROCHECK<sup>42</sup> (v3.3) software. The obtained model has a PROCHECK score greater than 93%; the quality of the CCD domain is slightly better than that of the NTD and CTD domains in terms of both DOPE energies and PROCHECK scores. Overall, our final homology model of HIV-1 IN shows remarkable similarity to that of HPV IN despite the much shorter interdomain linkers in HIV-1 IN and the low sequence conservation. The structures of the protein–DNA complex appear to be well stabilized, with extensive interactions between the protein and the DNA.

**Docking Simulations.** For the prediction of bioactive conformation and binding modes of our earlier synthesized chalcone and amide compounds,<sup>23</sup> we conducted docking simulations using the GLIDE (v5.7) program from Schrödinger Inc. (Portland, Oregon).<sup>43</sup> Both standard-precision (SP) and extra-precision (XP) modes of GLIDE were evaluated using a flexible ligand and rigid receptor routine. The protein preparation wizard of the Maestro (v9.2) interface in the Schrödinger modeling package was used to prepare the protein, applying the OPLS-2005 force field. Compounds were constructed using the 3D-sketcher module in Maestro. The carboxylic acid moiety was constructed as a carboxylate for docking. The inhibitors were energy minimized using the MacroModel (v9.9) molecular mechanics program and prepared for docking using the LigPrep (v2.5) facility in Maestro. Twenty docking poses were obtained for each molecule, the docking being terminated if two consecutive solutions were within a RMSD of 0.5 Å. Overlay of the X-ray derived conformation of raltegravir (Figure 2) over the docked (SP) conformation resulted in a lower RMSD of 0.60 Å between all heavy atoms as compared with a RMSD of 0.64 Å obtained with GLIDE XP. The SP option was therefore selected for subsequent docking studies; moreover, in addition to being slightly better in reproducing the crystal structure conformation, it is also less computationally intensive. To test the ranking accuracy of GLIDE (SP), the docked poses of



**Figure 2.** Overlay of the crystal structure conformation of raltegravir (colored green and represented by ball-and-stick) over the homology model derived conformation (colored in element and represented by tubes) and the GLIDE docking (SP) derived pose (carbons colored orange and shown in tubes). The homology model is shown in ribbons, and  $\text{Mg}^{2+}$  is shown by a purple sphere. Hydrogens are omitted for clarity.

raltegravir and elvitegravir were analyzed. It was found that the poses of elvitegravir have both better G scores and E-model energies than those of raltegravir, consistent with the enzyme inhibitory activities of the two compounds, with elvitegravir (ST  $\text{IC}_{50}$  of 0.0072  $\mu\text{M}$ ) being more potent than raltegravir (ST  $\text{IC}_{50}$  of 0.015  $\mu\text{M}$ ).

**MM-GBSA Analysis.** Empirical scoring functions used by various docking programs are designed to include various approximations in order to determine the binding pose and estimate the binding affinities of a large number of molecules in short periods of time. Docking procedures treat proteins as rigid, do not consider protein and ligand desolvation, and ignore the entropic terms in energy calculations. Although docking algorithms have been highly successful in reproducing correct binding poses, the approximate scoring functions often fail to rank order compounds on the basis of their binding affinities. Recent advances in drug discovery have used a combination of molecular mechanics and continuum solvation methods to determine and predict the relative binding free energies with better accuracy; these include Molecular Mechanics/Poisson–Boltzmann Surface Area (MM-PBSA)<sup>44</sup> and Molecular Mechanics/Poisson–Generalized Born Surface Area (MM-GBSA).<sup>45</sup> The MM-PB/GBSA method evaluates binding free energies by minimization and MD simulation of the protein–ligand complexes. Equation 1 calculates the free energy of binding ( $\Delta G_{\text{bind}}$ ) between a ligand and its receptor as

$$\Delta G_{\text{bind}} = \Delta H - T\Delta S = \Delta E_{\text{MM}} + \Delta G_{\text{sol}} - T\Delta S \quad (1)$$

$$\Delta E_{\text{MM}} = \Delta E_{\text{internal}} + \Delta E_{\text{electrostatic}} + \Delta E_{\text{vdW}} \quad (2)$$

$$\Delta G_{\text{sol}} = \Delta G_{\text{PB/GB}} + \Delta G_{\text{SA}} \quad (3)$$

where  $\Delta E_{\text{MM}}$  is the MM energy difference between the protein–ligand complex and the sum of the energies of free ligand and unliganded protein.  $\Delta G_{\text{sol}}$  is the corresponding difference in the solvation energies, and  $-T\Delta S$  is the corresponding difference in conformational entropy upon binding.  $\Delta E_{\text{internal}}$ ,  $\Delta E_{\text{electrostatic}}$ , and  $\Delta E_{\text{vdw}}$  are internal (bond, angle, torsion), electrostatic, and van der Waals (vdW) energies, respectively.  $\Delta G_{\text{PB/GB}}$  and  $\Delta G_{\text{SA}}$  are the electrostatic (polar) and nonelectrostatic (nonpolar) solvation energies, respectively.

The docking-derived conformations of 98 compounds, which have common binding modes and were used to build 3D-QSAR model 3, were first subjected to energy minimization by the local optimization module in the Prime package (v3.0 from Schrödinger Inc.), and binding free energies were calculated using the OPLS-2005 force field and the GBSA continuum model. The entropy penalty upon ligand binding was ignored in estimating the binding free energies, which is calculated in prime-MM-GBSA using eq 4:

$$\Delta G_{\text{bind}} = \Delta E_{\text{MM}} + \Delta G_{\text{GB}} + \Delta G_{\text{SA}} \quad (4)$$

**3D-QSAR Studies.** 3D-QSAR methods like comparative molecular field analysis (CoMFA)<sup>46</sup> and comparative molecular similarity analysis (CoMSIA)<sup>47</sup> are widely used to guide the synthesis of novel compounds by identifying the physiochemical properties that are important for activity, as well as an activity prediction tool. 3D structure building and CoMFA and CoMSIA studies were performed using the SYBYL program (version X1.1) on a Dell computer workstation with the Red Hat Linux operating system. All of the molecules were assigned MOPAC partial atomic charges. The CoMFA descriptors, steric (Lennard-Jones 6–12 potential) and electrostatic (Coulombic potential) field energies, were calculated using the Tripos force field with a distance-dependent dielectric constant, using the SYBYL default parameters: 2 Å grid points spacing, an  $\text{sp}^3$  carbon probe atom with +1 charge, vdW radius of 1.52 Å, and energy cutoff of 30 kcal/mol. The five similarity indices in CoMSIA, i.e., steric, electrostatic, hydrophobic, H-bond donor, and H-bond acceptor descriptors, were calculated using a  $\text{C}^{1+}$  probe atom with a radius of 1.0 Å placed at regular grid spacing of 2 Å. A Gaussian-type distance dependence was used between grid point  $q$  and each atom  $i$  of the molecule. The default value of 0.3 was used as the attenuation factor ( $R$ ). Column filtering was set to 2.0 kcal/mol. CoMSIA steric indices are related to the third power of the atomic radii. Electrostatic descriptors are derived from atomic partial charges. Hydrophobic fields are derived from atom-based parameters, and H-bond donor and acceptor indices are based on experimental results (SYBYL program).

**Partial Least Squares Analysis.** Partial least-squares (PLS) regression analysis that derives a mathematical relationship between dependent and independent variables, particularly when the number of descriptors outnumbers the number of compounds,<sup>48</sup> has been widely used to correlate important structural features with biological activity. For the present study, CoMFA and CoMSIA descriptors were used as independent variables, and ST  $\text{pIC}_{50}$  (biological activity) values were used as dependent variables. The predictive value of the models was evaluated by leave-one-out (LOO) cross-validation. The cross-validated coefficient,  $q^2$ , was calculated using eq 5:

$$q^2 = 1 - \frac{\sum (Y_{\text{predicted}} - Y_{\text{actual}})^2}{\sum (Y_{\text{actual}} - Y_{\text{mean}})^2} \quad (5)$$

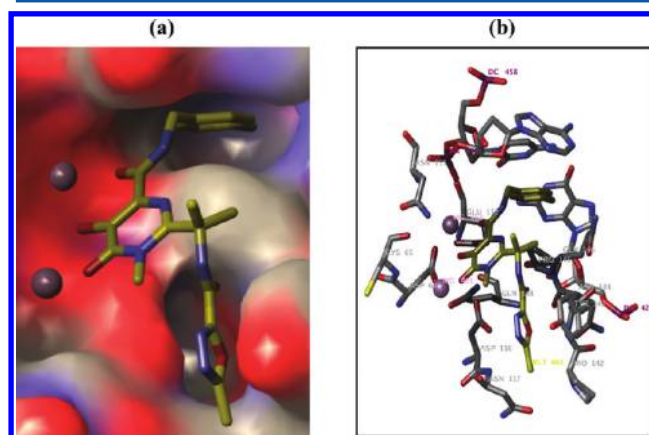
where  $Y_{\text{predicted}}$ ,  $Y_{\text{actual}}$ , and  $Y_{\text{mean}}$  are predicted, actual, and mean values of the target property ( $\text{pIC}_{50}$ ), respectively.  $\sum (Y_{\text{predicted}} - Y_{\text{actual}})^2$  is the predictive sum of squares (PRESS). Conventional correlation coefficient  $r^2$  and its standard error,  $s$ , were also computed for the final PLS models. CoMFA and CoMSIA coefficient maps were generated by interpolation of the pairwise products between the PLS coefficients and the standard deviations of the corresponding CoMFA or CoMSIA descriptor values. External validation of the developed models was performed by determining the predictive  $r^2$  of the test set compounds using eq 6:

$$r^2 = (\text{SD} - \text{PRESS})/\text{SD} \quad (6)$$

where SD is the sum of the squared deviations between the biological activity of the molecules in the test set and the mean biological activity of the training set.

## RESULTS AND DISCUSSION

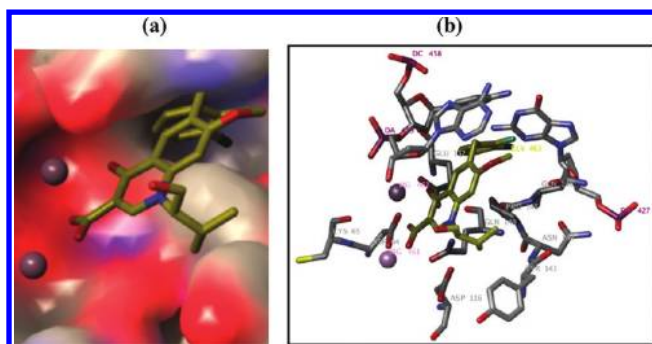
**Homology Model: Analysis of Binding Sites.** Interactions of the Substituted Benzyl Group. Binding modes of raltegravir and elvitegravir are shown in Figures 3 and 4, respectively.



**Figure 3.** (a) Binding mode of raltegravir within the active site of IN-DNA homology model, (b) binding interactions of raltegravir.  $\text{Mg}^{2+}$  ions are represented by purple spheres. Active site residues are represented by tubes and colored by element. Raltegravir is represented in tubes and highlighted with yellow carbons. Other atoms in the ligand are colored as follows: O, red; N, blue; F, bright green. Hydrogens are omitted for clarity.

The docking indicated that the *para*-fluoro benzyl ring of raltegravir is nicely inserted into the cavity (Figure 3a) that is formed by displacement of the 3'-adenosine when viral DNA associates with integrase in a preintegration complex. This unpairing of the terminal 3'-adenosine of the donor DNA duplex is essential to unmasking the scissile phosphodiester bond for 3'-P and subsequent ST reactions. Thus, the interactions of viral DNA with the IN and subsequent unpairing of its 3'-adenosine result in a cavity being formed at the DNA–protein interface. The 3-chloro-2-fluorobenzyl ring of elvitegravir also occupied the same cavity (Figure 4a) formed by the DNA–protein interface as the *para*-fluoro benzyl ring of raltegravir but is tilted slightly upward at about a  $45^\circ$  angle within the cavity relative to raltegravir. The phenyl rings of both raltegravir and elvitegravir are involved in a  $\pi$ – $\pi$  interaction with the cytosine base of the viral DNA (see Figures 3b and 4b). While the phenyl ring of raltegravir interacts with the adenine and guanine bases, it is the methoxy and *meta*-chlorine





**Figure 4.** (a) Binding mode of elvitegravir within the active site of the IN–DNA homology model, (b) binding interactions of elvitegravir.  $\text{Mg}^{2+}$  ions are represented by purple spheres. Active site residues are represented by tubes and colored by element. Elvitegravir is represented in tubes and highlighted with yellow carbons. Other atoms in the ligand are colored as follows: O, red; N, blue; F, bright green; Cl, dark green. Hydrogens are omitted for clarity.

substituents in elvitegravir that make contacts with the adenine and guanine bases of the viral DNA, respectively. Furthermore, the *para*-fluorine substituent in raltegravir and the *ortho*-fluorine substituent in elvitegravir are engaged in electrostatic interactions with the 4-amino group of the cytosine of the DNA. The phenyl rings of both compounds interact with Glu152. Gln148 appears within a 4 Å radius of raltegravir and elvitegravir, in our homology models. Thus, our homology models suggest physical contacts between raltegravir and elvitegravir with Gln148, which might explain the observation that mutation of this residue results in resistance to these drugs. We should note that in the crystal structure of foamy virus IN–DNA in complex with raltegravir or elvitegravir, there appeared to be no physical contact between both drugs and the equivalent residue Ser217.<sup>36</sup> The *p*-fluoro phenyl ring of raltegravir also makes extensive vdW contacts with the side chain methylenes and electrostatic contacts with the amino group of Gln146. In contrast, while the vdW contact with Gln146 is maintained, there is no apparent electrostatic interaction with this residue in the case of elvitegravir, which on the other hand appears to be directed more toward Glu152.

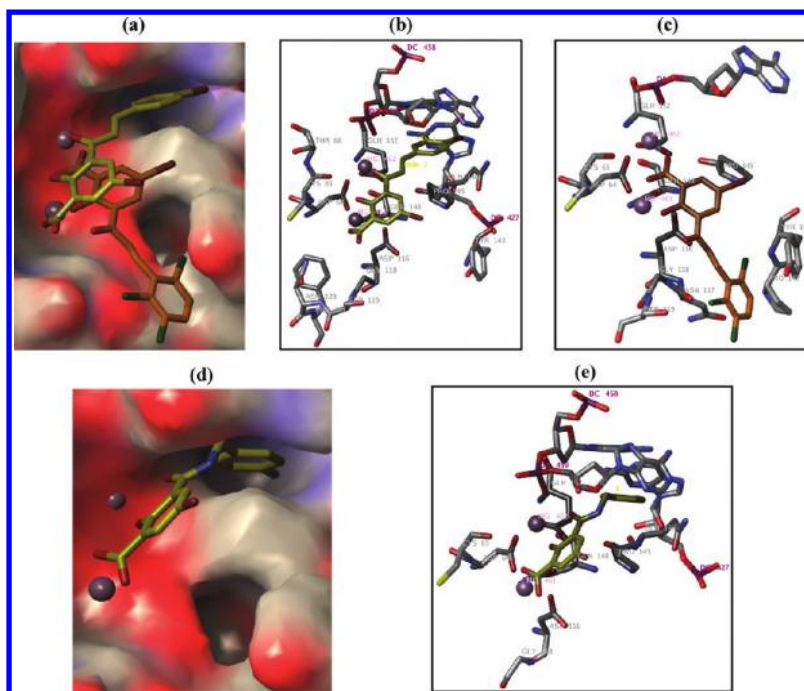
**Interactions of Metal-Chelating Motif.** The three-metal chelating groups in raltegravir are involved in coordination with the two metal atoms ( $\text{Mg}^{2+}$ ) in the IN active site. The pyrimidinone oxo and the acidic hydroxyl oxygen coordinate in a bidentate manner with the  $\text{Mg}^{2+}$  between Asp116 and Asp64. Likewise, the keto oxygen of the carboxamide moiety and the acidic hydroxyl oxygen are chelated with the second  $\text{Mg}^{2+}$  between Asp64 and Glu152. This hydroxyl moiety may also engage in a H-bonding interaction with Gln148, Asp64, and Asp116. The binding pose of the two-metal chelating pharmacophoric elements of elvitegravir shows the carboxylate moiety chelating the  $\text{Mg}^{2+}$  between Asp116 and Asp64. Furthermore, the 4-keto group and one of the oxygens of the carboxylate coordinate with the second  $\text{Mg}^{2+}$  between Asp64 and Glu152.

**Interactions of Side Chain.** The *gem*-dimethyl oxadiazole carboxamide moiety in raltegravir is inserted into the cleft formed by Asp116; Asn117; and loop residues Pro142, Tyr143, Asn144, and Pro145. One of the methyl groups of the *gem*-dimethyl moiety interacts with Pro145. The oxadiazole ring is

involved in a  $\pi$ -stacking interaction with the phenyl ring of Tyr143. The methyl on the oxadiazole is engaged in vdW interactions with the side chain methylene of Asn117. The hydroxyalkyl of the elvitegravir chain is also directed toward this cleft. One of the two methyl groups of the isopropyl substituents makes hydrophobic contacts with Tyr143, while the other makes contact with Pro145. The binding interactions revealed by the model are in accordance with those seen with the foamy virus crystal structure complex with these inhibitors and confirm the importance of this side chain in the binding of raltegravir and elvitegravir. Tyr143 is a conserved residue among all of the classes of polynucleotidyl transferases<sup>49</sup> and is proposed to be important for viral DNA binding. Another key residue to target is Asn117, which is implicated in the binding of the host DNA<sup>50</sup> and may thus be important for potent strand transfer (ST) inhibitory activity. Our homology model has also revealed Arg263 and Lys264 as important residues close to the active site that are involved in viral DNA binding.

**Docking Studies.** The best poses of the two most active compounds of the chalcone series (6 and 16) selected on the basis of GLIDE scores are shown in Figure 5. The topmost pose of compound 6 showed that its *para*-bromobenzyl group is inserted into the cavity formed by the DNA–protein interface (Figure 5a). The phenyl ring interacts with the cytosine and adenine bases of the viral DNA and also makes favorable vdW contacts with Gln146 and Pro145 (Figure 5b). Furthermore, the *para*-bromo substituent interacted with the amino group of Gln146. The 3-keto and the carboxylate group chelated the two divalent ( $\text{Mg}^{2+}$ ) ions in accordance with the proposed mechanism of action of ST inhibitors. The hydroxyl group also weakly coordinated with  $\text{Mg}^{2+}$  and makes H-bonding contact with the carboxylate of Asp116. The docking pose of compound 16 revealed that the 2,3,6-trichloro-substituted phenyl ring does not dock into the cavity formed by the DNA–protein interface but was placed rather on the opposite (180°) side exposed to the solvent. In this orientation, the trichlorophenyl group interacted with Pro142, the methylene of Ser119, and the methine of Gly118 (Figure 5c). The salicylic acid motif coordinates with the two  $\text{Mg}^{2+}$  ions, and the hydroxyl group H-bonds to the carboxylate of Asp116. The bromine substituent on the salicylic acid appears to weakly interact with Tyr143 and Pro145. The docking poses thus suggest the presence of multiple binding modes for the chalcones. Their smaller size gives them the flexibility to adopt different poses within the IN active site as opposed to the relatively bulky elvitegravir and raltegravir, which preferentially docked in their respective “bioactive conformations.”

The amide derivative (compound 51) docked into the IN active site as shown in Figure 5d and e. Similar to the *p*-fluorobenzyl group in raltegravir, the *p*-fluorobenzyl group of compound 51 is also fitted well into the cavity formed at the IN–DNA interface and makes contacts with Pro145 and Glu152. The phenyl ring  $\pi$ -stacks with the pyrimidinone ring of cytosine as well as interacts with the adenine base of the viral DNA. The fluorine atom makes vdW contacts with side chain methylenes of Gln146 and may be involved in H-bonding with the amino group of the cytosine. The 3-keto salicylic acid motif coordinates with the two metals in the IN active site. Furthermore, the 2-hydroxyl group forms a H bond with Glu152. Notably, no interactions with Tyr143 and Asn117 were observed. These compounds could be optimized to interact more extensively with these residues to improve their inhibitory activities. The models developed could thus



**Figure 5.** (a) Comparison of the docked poses of compounds **6** (carbons colored yellow) and **16** (carbons colored orange), (b) binding interactions of compound **6**, (c) binding interactions of compound **16**, (d) binding pose of compound **51** (carbons colored yellow), and (e) binding interactions of compound **51** within the IN–DNA homology model active site.  $\text{Mg}^{2+}$  ions are represented by purple spheres. The IN–DNA complex is colored by element and represented in tubes. Other atoms in the ligands are colored as follows: O, red; Br, dark red; Cl, dark green; F, bright green. Hydrogens are omitted for clarity.

serve as a surrogate platform for the structure-based design of potent IN inhibitors. With a further objective of using the docked conformations to build a docking-based QSAR model, all 103 IN inhibitors of the data set were docked on the IN active site, and 20 poses for each compound were generated. The top-ranked (based on the GLIDE scores) poses of 85 compounds were similar in their binding modes. Of the remaining 18 compounds, 13 of them had lower ranked poses that shared similar binding modes with the top-ranked poses of the above 85 compounds. Markedly different poses were obtained for compounds **63**, **86**, **89**, **92**, and **95**. These compounds may have a different binding mode or require an induced fit. Although it is well established that the GLIDE docking is successful in giving reasonably good ligand poses, it often fails to rank-order compounds on the basis of their binding affinities. We observed that GLIDE score showed no correlation ( $r^2 = 0.05$ ) with the ST inhibitory activities for the 98 similarly docked compounds. Thus, to further validate their “bioactive conformations,” a postdocking MM-GBSA rescoring was performed.

**MM-GBSA Results.** The docked poses of the pruned data set were subjected to Molecular Mechanics/Generalized Born Surface Area (MM-GBSA)<sup>44</sup> free energy calculations with PRIME (v) of the Schrödinger package. The calculated  $\Delta G_{\text{bind}}$  along with vdW contributions to binding free energy are listed in Supporting Information Table S1. The MM-GBSA calculations revealed an improved correlation, with  $r^2 = 0.34$ , between the binding free energies ( $\Delta G_{\text{bind}}$ ) and the ST inhibitory activities (Figure 6a). A low, albeit better, correlation might be a result of a large and diverse set of chemical structures included in the analysis and a poor estimation of electrostatic binding energies. However, the vdW contribution to the MM-GBSA binding free energy correlated best with

the activities with an  $r^2$  value of 0.57 (Figure 6b), which went up to 0.62 with the removal of one outlier (compound **33**). A better correlation with the vdW energy term is significant since hydrophobic interactions are one of the most important contributors to IN inhibitory activity, and all potent IN inhibitors have a common hydrophobic halogenated benzyl ring system as a pharmacophoric feature. Another important contribution to the IN inhibitory activity comes from the electrostatic interactions, notably the coordination of the metal-chelating heteroatoms in an IN inhibitor with  $\text{Mg}^{2+}$  in the IN active site. The electrostatic contributions to the binding free energy are however not well predicted (data not shown). Previous studies<sup>51,52</sup> have shown that electrostatic energies are difficult to estimate, particularly when the data sets have both charged and neutral molecules, as is the case with the present study. Moreover, the negatively charged residues, Asp64, Asp116, Glu152, and the two  $\text{Mg}^{2+}$  ions are involved in important interactions with inhibitors in the IN active site. This system warrants separate studies to determine the effects of protein charge states, parametrization of  $\text{Mg}^{2+}$ , and local dielectric constant on the calculation of polar solvation energies. Although Prime MM-GBSA did not accurately predict binding affinities, it was successful in ranking the compounds according to relative activity and thus giving support to the use of the docking derived conformations as “bioactive conformations” for 3D-QSAR studies.

**3D-QSAR Studies.** *3D-QSAR Model 1, 2, and 3.* A total of 103 compounds were used for the initial 3D-QSAR studies. These include our earlier synthesized<sup>23</sup> 3-keto salicylic acid chalcones (Table 1) and amide derivatives (Table 2), and the reported *N*-methyl-4-hydroxypyrimidinone carboxamides<sup>12,13</sup> (Table 3a–f) and 4-quinolone-3-carboxylic acid and analogs<sup>16,17</sup>

(Table 4a and b). Although the data set represented several different compound classes, they all share a somewhat similar structural core. These compounds have a common mechanism of action (chelation of the divalent metal and inhibition

of ST) and cover a wide range of  $IC_{50}$  values spanning over 5 log units, thus constituting a suitable set of compounds for QSAR studies. The data set was also widely distributed with regard to activity, comprising potent IN inhibitors having

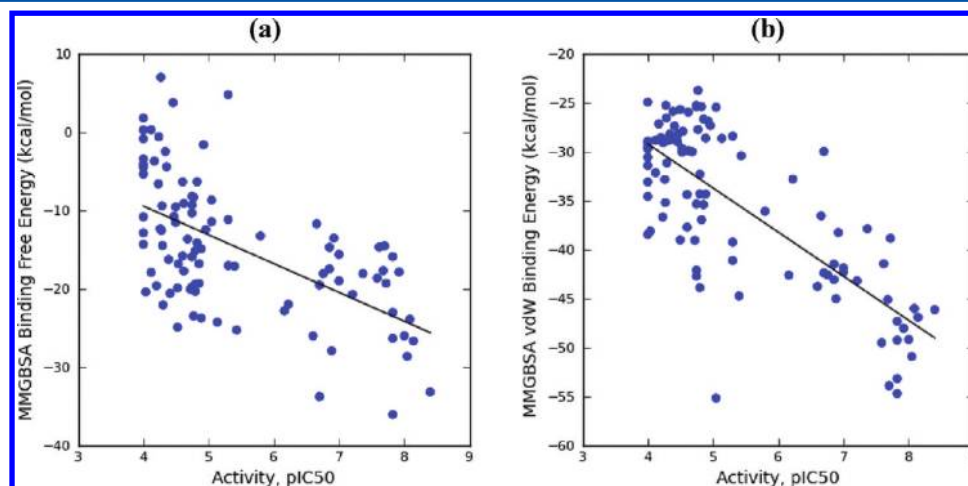
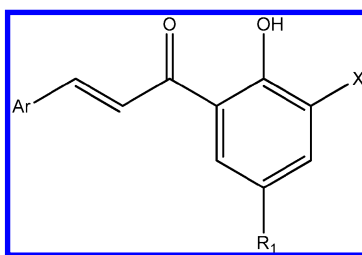


Figure 6. Plot of MM-GBSA (a)  $\Delta G_{\text{bind}}$  and (b)  $\Delta G_{\text{vdW}}$  energy correlations with ST inhibitory activities ( $pIC_{50}$ s).

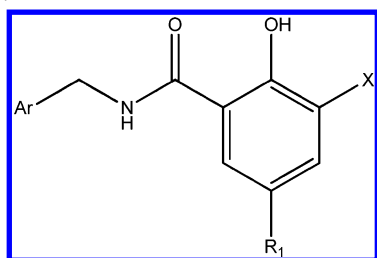
Table 1. Structure and HIV-1 IN Inhibitory Activities of 3-Keto Salicylic Acid Chalcones Used in the 3D-QSAR Study<sup>23</sup>



cpd	Ar	X	R <sub>1</sub>	ST <sup>a</sup> IC <sub>50</sub> (μM)	cpd	Ar	X	R <sub>1</sub>	ST <sup>a</sup> IC <sub>50</sub> (μM)
1	Ph	COOH	Br	25 ± 5	26	2,3-di-Cl-Ph	COOH	F	30 ± 9
2	2-Cl-Ph	COOH	Br	14 ± 4	27	2,5-di-Cl-Ph	COOH	F	29 ± 6
3	2-F-Ph	COOH	Br	32 ± 3	28	2,3,5-tri-Cl-Ph	COOH	F	15 ± 4
4	3-Cl-Ph	COOH	Br	9 ± 3	29	2,3,6-tri-Cl-Ph	COOH	F	39 ± 6
5	4-Cl-Ph	COOH	Br	12 ± 1	30	3,5-di-Me-Ph	COOH	Br	58
6	4-Br-Ph	COOH	Br	5 ± 2	31	2-Cl-3,4-di-MeO-Ph	COOH	Br	76
7	4-I-Ph	COOH	Br	35	32	6-F-3,4-di-MeO-Ph	COOH	Br	76
8	2,4-di-Cl-Ph	COOH	Br	7 ± 4	33	2-BzO-Ph	COOH	Br	9
9	2,3-di-Cl-Ph	COOH	Br	19 ± 11	34	3-(4-Cl-PhO)-Ph	COOH	Br	18
10	2,3,4-di-MeO-Ph	COOH	Br	34 ± 6	35	4-(4-Cl-PhO)-Ph	COOH	Br	16
11	2,5-di-Cl-Ph	COOH	Br	51 ± 16	36	3-thienyl	COOH	Br	18 ± 1
12	2,6-di-Cl-Ph	COOH	Br	18 ± 4	37	3-(cyclopentyloxy)-Ph	COOH	Br	55
13	3,4-di-Cl-Ph	COOH	Br	16 ± 8	38	3-(2-OEt-naphthyl)	COOH	Br	46
14	2,3,5-tri-Cl-Ph	COOH	Br	14 ± 4	39	5-MeO-naphthyl	COOH	Br	54 ± 25
15	2,3,6-tri-Cl-Ph	COOH	Br	13 ± 5	40	5-benzothiophene	COOH	Br	11 ± 1
16	4-Cl-Ph	COOH	Cl	<3.7	41	2-furan-2-yl-Ph	COOH	Br	16 ± 2
17	2,4-di-Cl-Ph	COOH	Cl	53 ± 21	42	3-Cl-4,5-methylenedioxy-Ph	COOH	Br	63
18	2,3-di-Cl-Ph	COOH	Cl	>100	43	3-OCF <sub>3</sub> -Ph	COOH	Br	52
19	2,5-di-Cl-Ph	COOH	Cl	92 ± 14	44	3-NO <sub>2</sub> -Ph	COOH	Br	32 ± 6
20	2,3,5-tri-Cl-Ph	COOH	Cl	25	45	3-Br-Ph	COOH	Br	18 ± 1
21	2,3,6-tri-Cl-Ph	COOH	Cl	17 ± 3	46	3-F-Ph	COOH	Br	59 ± 2
22	3,4-di-Cl-Ph	COOH	Cl	30 ± 5	47	2,3,6-tri-Cl-Ph	CN	Br	44 ± 2
23	4-Cl-Ph	COOH	F	41 ± 9	48	2,3,6-tri-Cl-Ph	COOH	CH <sub>3</sub>	50 ± 9
24	2,4-di-Cl-Ph	COOH	F	17 ± 3	49	2,3,5-tri-Cl-Ph	NO <sub>2</sub>	Br	68 ± 5
25	2,4-di-Cl-Ph	COOH	F	24 ± 9	50	2,3,6-tri-Cl-Ph	NO <sub>2</sub>	Br	21 ± 3

<sup>a</sup>ST represents strand transfer activity.



**Table 2. Structure and HIV-1 IN Inhibitory Activities of 3-Keto Salicylic Acid Amides Used in the 3D-QSAR Study<sup>23</sup>**

cpd	Ar	R <sub>1</sub>	X	ST <sup>a</sup> IC <sub>50</sub> (μM)
51	4-F-Ph	Br	COOH	15 ± 3
52	4-F-Ph	H	CH <sub>3</sub>	>100
53	4-F-Ph	Br	H	>100
54	4-F-Ph	Br	OH	13

<sup>a</sup>ST represents strand transfer activity.

IC<sub>50</sub> values of 7–8 nM, moderately active 3-keto salicylic acid derivatives with IC<sub>50</sub> values of 4–5 μM, and inactive compounds with IC<sub>50</sub> values of 100 μM. To evaluate the predictive ability of models, the data set was divided into a training set of 90 compounds and an external validation test set of 13 compounds. The test set was selected randomly using factor analysis in the SYBYL package (vX1.1, Tripos Associates Inc. St. Louis, MO) and contained compounds representing all four different potency categories.

**Conformation and Alignment of the Data Set.** For 3D-QSAR studies, we compared two atom-fit alignment rules (models 1 and 2), with the docking alignment (model 3). Determination of the “bioactive conformation” is of the utmost importance for the development of a successful 3D-QSAR model of the CoMFA or CoMSIA type. To this end, the bioactive conformation of raltegravir (compound **65** in Table 3c) was obtained from our developed raltegravir–IN–DNA homology model. This conformation of raltegravir was used as a template to build the remaining *N*-methyl-4-hydroxypyrimidinone carboxamides. Similarly, the homology-model-derived conformation of elvitegravir (compound **88** in Table 4a) was used to build the remaining 4-quinolone-3-carboxylic acid analogs. The top-ranked pose of the amide derivative **51** obtained from the docking simulations was used as a template to build the other amide derivatives. As discussed in the docking studies, chalcones appear to have multiple binding modes. The highest-ranked poses of the two most active compounds **6** and **16** were each used as a bioactive conformation template to build the remaining chalcones of the series. To take into account the conformations of the two most active chalcone derivatives, two different alignments rules, alignment I (model 1) and alignment II (model 2), were generated. It was primarily aimed to assess their effect on the overall unified 3D-QSAR model for IN inhibition. However, the conformations of the *N*-methyl-4-hydroxypyrimidinone carboxamides, 4-quinolone-3-carboxylic acids, and 3-keto salicylic acid amides remained identical in both alignments.

**Alignment I.** For alignment I, the atom-fit alignment function in SYBYL was used for a pairwise alignment of elvitegravir, chalcone, and amide derivatives with raltegravir as a template. The conformations of these compounds were judiciously aligned over the raltegravir template with atomic coordinates adjusted to provide the maximum overlap of the molecules in the data set. All of the compounds were then partially energy

minimized by the Simplex procedure to relieve excessive strain. The X-ray and homology-model-derived conformations revealed partial overlap of the pyrimidinone and quinolone moieties of raltegravir and elvitegravir, respectively. The carboxylate of elvitegravir and the 6-oxo group of raltegravir chelate the same Mg<sup>2+</sup> ion (between Asp116 and Asp64). Therefore, for alignment, the pyridone ring of elvitegravir was superimposed over the pyrimidinone ring of raltegravir such that the 3'-carbon of the quinolone system was superimposed over the 6'-carbon of the pyrimidinone ring of raltegravir. In this orientation, the C-5 carbon of the quinolone ring in elvitegravir superimposed over the carbonyl carbon of the carboxamide moiety in raltegravir. This alignment rule resulted in a good overlap of the six-membered rings and matched the crucial metal-chelating pharmacophoric groups. The docked conformation of the top-ranked poses of the amide analog **51** and chalcone derivative **6** revealed that their salicylic acid groups mimicked the pyrimidinone ring of raltegravir. The docking also suggested that the 3-keto oxygen of reference compounds **51** and **6** and the carbonyl oxygen of the carboxamide of raltegravir interacted with the same Mg<sup>2+</sup> ion (between Asp64 and Glu152). Likewise, the carboxylate groups of compounds **51** and **6** and the pyrimidinone oxo of raltegravir coordinated with the Mg<sup>2+</sup> between Asp64 and Asp116. Thus, for alignment of the amide and chalcone derivatives, the salicylic acid rings of compounds **51** and **6** were aligned over the pyrimidinone ring of raltegravir such that 1'-carbon atoms of compounds **51** and **6** were superimposed over the 6'-carbon of the pyrimidinone ring of raltegravir. Furthermore, the carbonyl carbons of compounds **51** and **6** were superimposed over the carbonyl carbon of the carboxamide moiety of raltegravir. This alignment correctly positioned their halogenated phenyl rings into the hydrophobic cavity as depicted in the docking studies. The superimposition of the data set using alignment I is shown in Figure 7a.

**Alignment II.** The pyridone ring of elvitegravir and the salicylic acid group of compound **51** were superimposed (atom-fit) over the pyrimidinone ring of raltegravir similarly to alignment I. However, the docked pose of compound **16** was used as the bioactive conformation for the chalcone derivatives. As discussed in the docking section, the halogenated phenyl ring of compound **16** was placed in the opposite (180°) direction. In this orientation, the carboxylate of compound **16** is directed toward the carbonyl oxygen of the benzyl carboxamide moiety of raltegravir, and both functionalities interact with the Mg<sup>2+</sup> between Asp64 and Glu152. Thus, for this alignment, the salicylic acid group and the carboxylate carbon of compound **16** were respectively superimposed over the pyrimidinone ring and the carbonyl carbon of the carboxamide moiety of raltegravir. The superimposition of the data set using alignment II is shown in Figure 7b.

**Alignment III.** Theoretically, overlaying the poses of the compounds docked at the IN active site should provide a better bioactive conformational alignment than the atom-fitting alignments I and II. Thus, to derive a truly docking based 3D-QSAR, all 103 compounds were docked and their binding poses analyzed. The poses of 98 compounds that bound similarly were validated by MM-GBSA rescoring, which gave better correlations ( $\Delta G_{\text{bind}}$   $r^2 = 0.34$ ;  $\Delta E_{\text{vdW}}$   $r^2 = 0.57$ ) with ST inhibitory activities. These conformations were then overlaid to produce alignment III, as depicted

Table 3. Structure and HIV-1 IN Inhibitory Activities of *N*-Methyl-4-Hydroxypyrimidinone-Carboxamides Used in the 3D-QSAR Models<sup>12,13</sup>

(a)					
Cpd	R <sub>1</sub>	R <sub>2</sub>	ST <sup>a</sup> IC <sub>50</sub> (μM)		
55		H	0.010		
56		H	0.015		
57	OH	H	0.004		
58		Me	0.015		
59		H	0.020		
60		H	0.026		
(b)					
Cpd	R <sub>1</sub>	ST <sup>a</sup> IC <sub>50</sub> (μM)			
61	Me	>5.0			
62		0.021			
63		0.004			
64		0.009			
(c)					
Cpd	R	ST <sup>a</sup> IC <sub>50</sub> (μM)			
65		0.015			
66		0.012			
67		0.015			
(d)					
Cpd	n	ST <sup>a</sup> IC <sub>50</sub> (μM)			
68	1	0.12			
69	2	0.22			
(e)					
Cpd	R	ST <sup>a</sup> IC <sub>50</sub> (μM)			
70		0.062			
71		0.69			
72		0.14			
73		0.10			
74		0.20			
75		0.25			
76		0.13			
77		4.0			
(f)					
Cpd	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	n	ST <sup>a</sup> IC <sub>50</sub> (μM)
78	H	H	Me	1	>5
79	2-F	H	H	1	0.175
80	3-Cl	H	H	1	0.062
81	3-Br	H	H	1	0.019
82	3-OMe	H	H	1	0.140

<sup>a</sup>ST represents strand transfer activity.

**Table 4. Structure and HIV-1 IN Inhibitory Activities of 4-Quinolone-3-Carboxylic Acid–Carboxamides Used in the 3D-QSAR Study<sup>16,17</sup>**

(a)						
Cpd	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	ST <sup>a</sup> IC <sub>50</sub> (μM)		
83		H	H	1.6 ± 0.300		
84		H	H	0.043 ± 0.008		
85			H	0.024 ± 0.011		
86			OCH <sub>3</sub>	0.009 ± 0.002		
87			H	0.008 ± 0.001		
88			OCH <sub>3</sub>	0.007 ± 0.002		

(b)						
Cpd	X	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	ST <sup>a</sup> IC <sub>50</sub> (μM)
89	O	F	Cl	H	(CH <sub>2</sub> ) <sub>2</sub> OH	>100
90	S	H	H	H	(CH <sub>2</sub> ) <sub>2</sub> OH	18.5 ± 3.7
91	CH <sub>2</sub>	Cl	H	Cl	(CH <sub>2</sub> ) <sub>2</sub> OH	0.2 ± 0.0
92	CH <sub>2</sub>	Cl	H	Cl	(CH <sub>2</sub> ) <sub>3</sub> OH	1.3 ± 0.1
93	CH <sub>2</sub>	Cl	H	Cl	(CH <sub>2</sub> ) <sub>4</sub> OH	0.6 ± 0.3
94	CH <sub>2</sub>	Cl	H	Cl	(CH <sub>2</sub> ) <sub>2</sub> N(CH <sub>3</sub> ) <sub>2</sub>	24.1 ± 5.9
95	CH <sub>2</sub>	Cl	H	Cl	(CH <sub>2</sub> ) <sub>2</sub> OCH <sub>3</sub>	16.5 ± 0.7
96	CO	Cl	H	Cl	(CH <sub>2</sub> ) <sub>2</sub> OH	>100
97	CO	Cl	H	Cl	(CH <sub>2</sub> ) <sub>3</sub> OH	>100
98	CO	Cl	H	Cl	(CH <sub>2</sub> ) <sub>4</sub> OH	>100
99	CHCH	H	H	H	(CH <sub>2</sub> ) <sub>2</sub> OH	>100
100	CONH	H	H	H	(CH <sub>2</sub> ) <sub>2</sub> OH	>100
101	NHCO	H	H	H	(CH <sub>2</sub> ) <sub>2</sub> OH	>100
102	NH	H	Cl	F	(CH <sub>2</sub> ) <sub>2</sub> OH	>100
103	NH	H	Cl	F	CH <sub>2</sub> CH <sub>3</sub>	>100

<sup>a</sup>ST represents strand transfer activity.

in Figure 7c, which was used to build a docking-based 3D-QSAR, model 3.

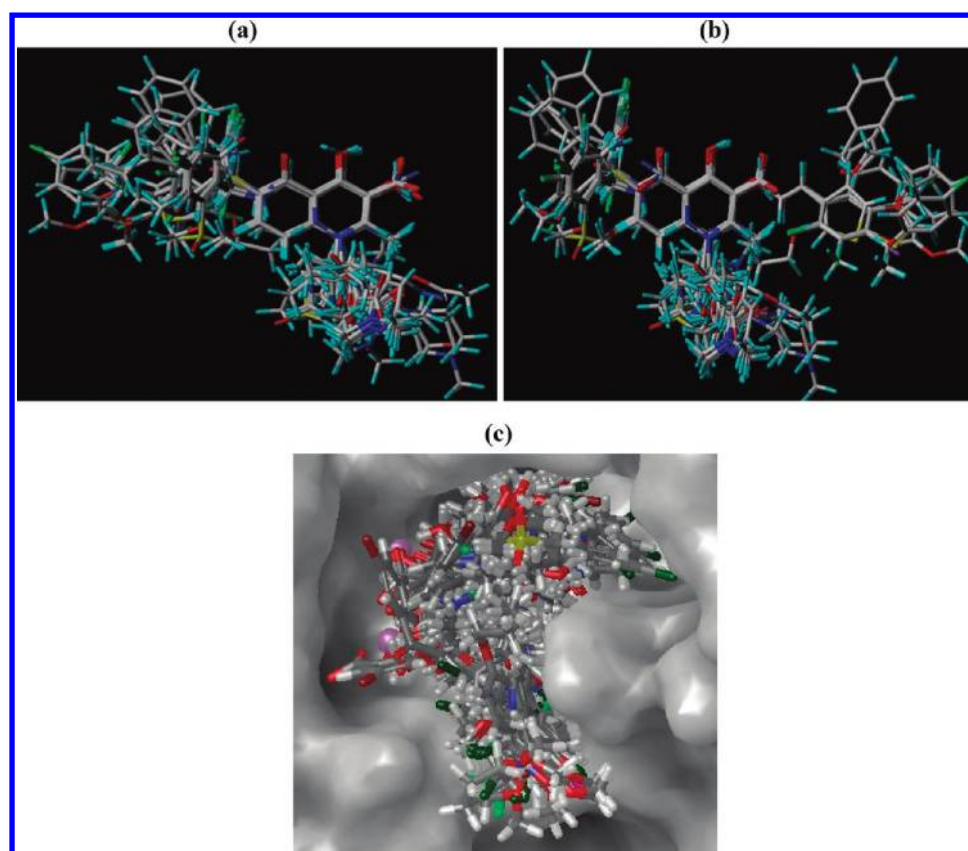
**PLS Results.** PLS analysis along with the leave-one-out (LOO) cross-validation method was used to develop 3D-QSAR models and determine the predictive ability (using the cross-

validated coefficient  $q^2$ ) and the optimal number of components to use in the final regression models. The optimal number of components used (five for both models 1 and 2) was determined by the number of components that yielded the smallest standard error of prediction. Atom-fit alignments I and II led to two 3D-QSAR models: model 1 and model 2, respectively. Alignment I may be more plausible since placement of the hydrophobic phenyl rings of the chalcones into the cavity should have more extensive contact of the inhibitors with IN than alignment II, in which part of the phenyl ring of the inhibitor is solvent-exposed. However, significant  $q^2$  and  $r^2$  values were obtained for both CoMFA and CoMSIA 3D-QSARs models I and II (Table 5). In terms of PLS statistics, the CoMFA of both models performed better than their CoMSIA counterparts. The CoMFA 3D-QSAR model 1 performed slightly better than the CoMFA model 2, with  $q^2$  and  $r^2$  values of 0.82 and 0.94, respectively, compared to 0.79 and 0.93 for CoMFA model 2. Similar  $q^2$  values around 0.71 were obtained for both CoMSIA models 1 and 2. Due to its extreme residual value, compound 61 was identified as an outlier in model 1. It was the only compound in the series which did not have a hydrophobic *p*-fluoro benzyl ring, and thus its structural uniqueness might be a factor for it being an outlier.

To test the robustness of the predictive capabilities of the models, they were used to predict the activity of an external test set of 13 compounds which were excluded from the training set. The residuals of the training and test set predictions for model 1 are listed in Supporting Information Tables S2 and S3, respectively. The corresponding residual values for model 2 are given in Supporting Information Tables S4 and S5. For model 1, the CoMFA prediction curves for the training and the test sets are shown in Figure 8a and the corresponding CoMSIA predictions in Figure 8b. The corresponding CoMFA and CoMSIA predictions for model 2 are shown in Figure 9a and b, respectively. As is evident, the present study gave high correlations between actual and predicted values for the external test set. CoMFA and CoMSIA gave comparable predictive  $r^2$  values of 0.93 and 0.83 for model 1 and 0.90 and 0.82 for model 2. In addition, more stringent group cross-validation (20 times) was applied to further test the reliability of models. As shown in Table 5, a  $r_{cv}^2$  value of 0.82 was obtained for CoMFA and one of 0.71 for CoMSIA of model 1. The corresponding values for model 2 were 0.78 and 0.69, respectively. These results demonstrate that both models are stable and statistically robust. Moreover, low CoMFA and CoMSIA scrambling average  $q^2$  values of 0.17 and 0.18, respectively, for model 1 and 0.31 and 0.25 for model 2 also validate the statistical significance of both models, showing that they were not obtained by chance correlation. The result of bootstrapping further corroborates the robustness of the derived QSAR models.

PLS results of model 3, which was developed using the docking derived alignment III, are also shown in Table 5. The models, which could be developed with only three components, gave  $q^2$  and  $r^2$  values of 0.61 and 0.93 for CoMFA and the corresponding values of 0.60 and 0.94 for CoMSIA. Both the CoMFA and the CoMSIA models gave a low standard error of estimate (SEE) and large  $F$  values of 0.36 and 221.03 and of 0.33 and 263.92, respectively. Cross-validation and bootstrapping results indicated that both CoMFA and CoMSIA models were statistically robust. For assessing predictivity, the





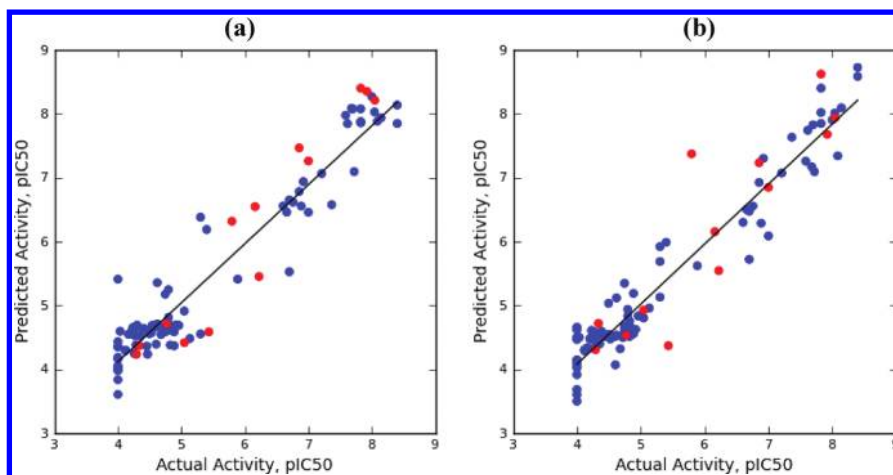
**Figure 7.** Superimposition of the data set using (a) alignment I and (b) alignment II. (c) Docked conformations used for alignment III. Atoms are represented as O, red; Br, dark green; Cl, green; F, bright green; H, cyan; C, white; S, yellow; N, blue.

**Table 5.** PLS Statistics of CoMFA and CoMSIA 3D-QSAR Models 1-3

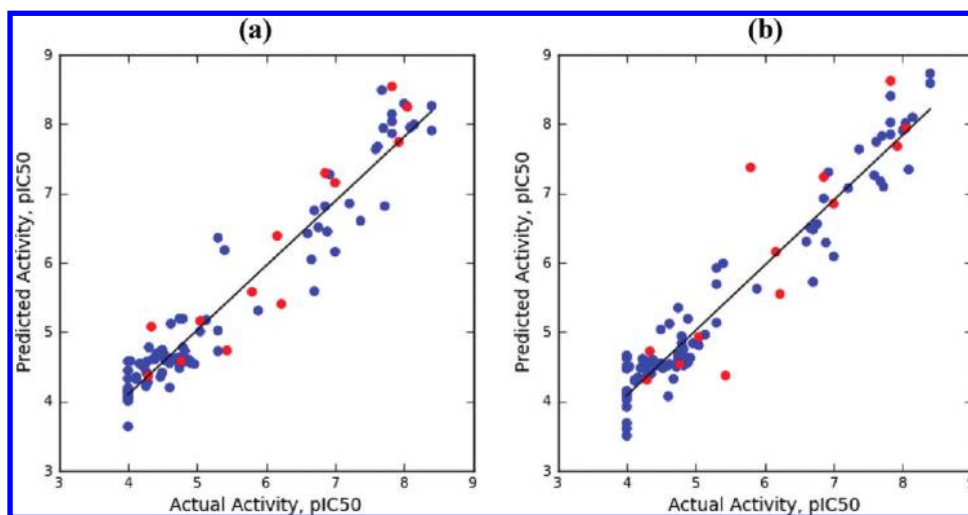
PLS statistics	model 1		model 2		model 3	
	CoMFA	CoMSIA	CoMFA	CoMSIA	CoMFA	CoMSIA
$q^2$	0.82	0.71	0.79	0.71	0.61	0.60
$r^2$	0.94	0.93	0.93	0.94	0.93	0.94
$s$	0.34	0.36	0.38	0.35	0.36	0.33
$F$	284.69	240.51	222.65	261.51	221.03	263.92
PLS component	5	5	5	5	3	3
field contribution						
steric	0.526	0.105	0.528	0.107	0.453	0.122
electrostatic	0.474	0.288	0.472	0.292	0.547	0.242
hydrophobic		0.242		0.223		0.193
H-donor		0.224		0.214		0.232
H-acceptor		0.141		0.164		0.211
bootstrapping $r^2$ (20 runs)	$0.94 \pm 0.012$	$0.95 \pm 0.016$	$0.90 \pm 0.01$	$0.95 \pm 0.01$	$0.95 \pm 0.011$	$0.96 \pm 0.009$
cross-validation $q^2$ (20 runs)	0.82	0.71	0.78	0.69	0.60	0.59

same 13 compounds were used as an external test set. The CoMFA model performed significantly better for the test set with a predictive  $r^2$  value of 0.82 as compared with a predictive  $r^2$  value of 0.62 obtained with the CoMSIA model. The residuals of the training and test set predictions of model 3 are summarized in Supporting Information Tables S6 and S7, respectively. Plots of CoMFA and CoMSIA predictions for the training and the test set are shown in Figures 10a and 10b. Comparison of the QSAR models suggests that the atom-fit alignment models 1 and 2 gave higher  $q^2$  values when compared with the docking-based alignment model 3. However, models 1 and 2 were derived using five components, and the higher  $q^2$  values might result from overfitting of the data

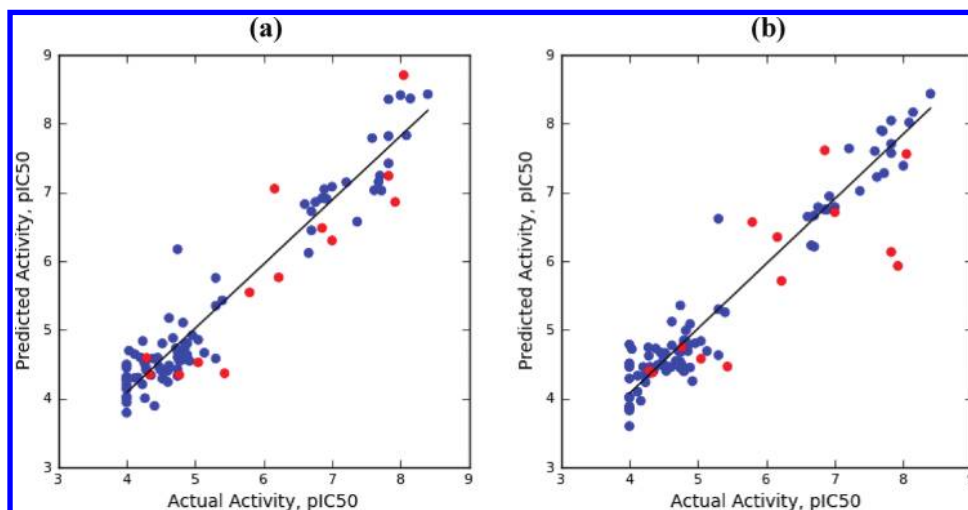
because of the greater number of components used. Thus, to check for any overfitting, models 1 and 2 were revisited using three components. Rederived model 1 gave  $q^2$  and  $r^2$  values of 0.67 and 0.85 for CoMFA and 0.60 and 0.84 for CoMSIA, respectively. However, the model gave higher SEE and smaller  $F$  values (0.54 and 173.69 for CoMFA and 0.57 and 151.09 for CoMSIA) compared to the docking-based model 3. Rederived models 2 with 3 components also gave lower  $q^2$  and  $r^2$  values ( $q^2 = 0.58$  and  $r^2 = 0.75$  for CoMFA and  $q^2 = 0.55$  and  $r^2 = 0.76$  for CoMSIA) and higher SEE and smaller  $F$  values when compared to model 3. Thus, in terms of the robustness and ability to generalize, the docking-derived 3D-QSAR models are better than the ligand-based alignments I and II.



**Figure 8.** Prediction curves for model 1 (a) CoMFA predictions of the training and the test sets and (b) CoMSIA predictions of the training and the test sets. The training and test sets are represented by blue and red spheres, respectively.



**Figure 9.** Prediction curves for model 2 (a) CoMFA predictions of the training and the test sets and (b) CoMSIA predictions of the training and the test sets. The training and the test sets are represented by blue and red spheres, respectively.



**Figure 10.** Prediction curves for model 3 (a) CoMFA predictions of the training and the test sets and (b) CoMSIA predictions of the training and the test sets. The training and test sets are represented by blue and red spheres, respectively.

Also, compared to our previous pharmacophore-based CoMFA and CoMSIA 3D-QSAR modeling of the salicylic acid derived

IN inhibitors, this structure-based and composite 3D-QSAR approach has generated better models, with best  $q^2$  and

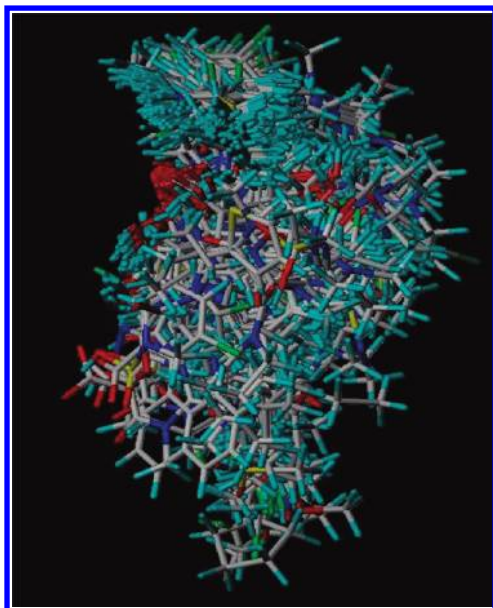


Figure 11. Alignment used for 3D-QSAR model 4.

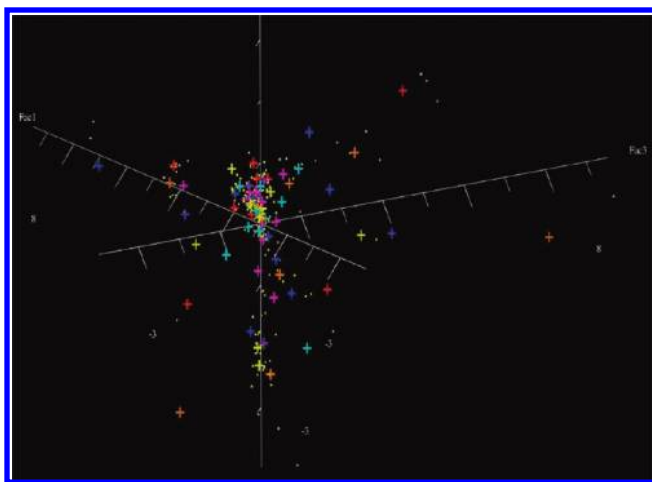


Figure 12. Selection of the model 4 test set by factor analysis. The selected test set compounds are highlighted with a plus sign.

predictive  $r^2$  values of 0.82 and 0.93, respectively, compared to  $q^2$  and predictive  $r^2$  values of 0.54 and 0.57, respectively.<sup>23</sup> Taken together, the results suggest that the docking-based composite QSAR models are better predictive tools than our previous QSAR models<sup>23</sup> for lead optimization of the 3-keto salicylic acid series.

**3D-QSAR Model 4.** A docking-based alignment was then subsequently chosen to build a comprehensive 3D-QSAR model 4. Remaining compounds of the N-methyl pyrimidones class<sup>12,13</sup> (Table 6a–c) were added to the data set. Furthermore, diverse, but related, classes of IN inhibitors including dihydroxypyrimidine (Tables 7a–c and 13) carboxamides,<sup>53–56</sup> bicyclic pyrimidinones<sup>57</sup> (Table 8), *tert*-butyl-*N*-methyl pyrimidones<sup>58</sup> (Table 9a–c), hydroxypyrrolinone carboxamides<sup>59</sup> (Table 10), pyridopyrimidinone carboxamides<sup>60</sup> (Table 11a–c), pyrazinopyrimidine carboxamides<sup>61</sup> (Table 12), and 2-pyrroli-dinyl *N*-methyl pyrimidones<sup>62</sup> (Table 14a,b) reported in the literature were also incorporated into the data set. With regard to homogeneity of the data, most of the compounds (225 in total) are from one laboratory,<sup>12,13,53–62</sup> the next largest set (54 compounds) came from our laboratory,<sup>23</sup> and the rest came from two other laboratories.<sup>16,17</sup> Looking at the very low concentrations of substrate DNA used relative to the  $K_m$  of integrase, there is no significant difference between the  $IC_{50}$  and  $K_i$  values according to the Cheng–Prusoff equation, eq 7:

$$IC_{50} = \left( 1 + \frac{[S]}{K_m} \right) K_i \quad (7)$$

where  $[S]$  is the substrate DNA concentration and  $K_m$  is the Michealis–Menton constant of integrase.<sup>16,63–66</sup> The  $IC_{50}$  values were thus used in the 3D-QSAR analysis as reported. Literature compounds were selected to exclude racemic mixtures, compounds with undefined stereochemistry, and compounds with  $IC_{50}$  values greater than 20  $\mu M$ . Thus, 199 additional inhibitors other than the 103 inhibitors considered for models 1–3 were selected for further analysis, all of which were then docked (GLIDE SP) into the IN–DNA homology model. Analysis of the docking results revealed that the topmost poses of all but one compound bound in a similar binding mode to those of the 98 compounds used in the docking-based 3D-QSAR model 3. This compound (132b, Table 7a) had a bulky *N*-methyl naphthyl carboxamide as the hydrophobic aryl moiety rather than the usual fluorobenzyl group. It appeared that the lack of flexibility

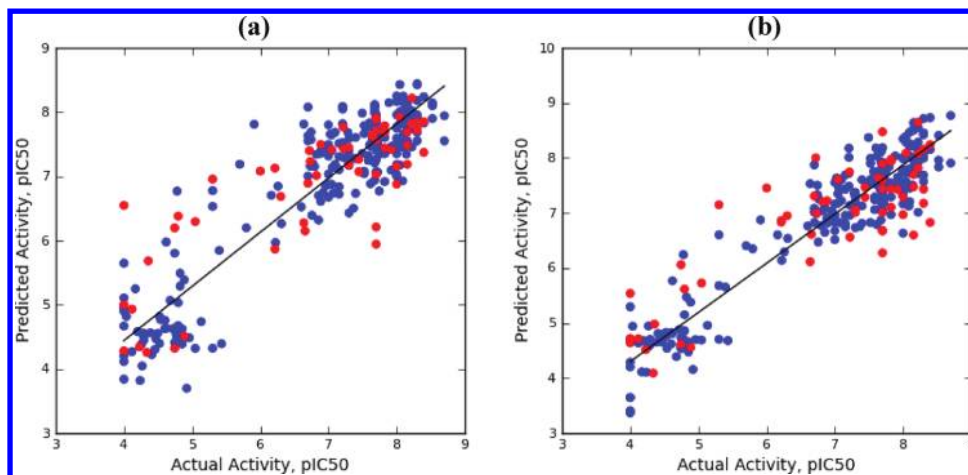
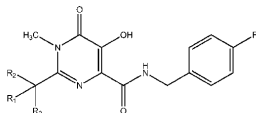
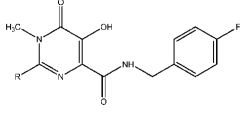
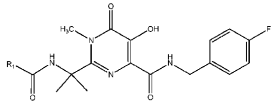


Figure 13. Prediction curves for model 4 (a) CoMFA predictions of the training and the test sets and (b) CoMSIA predictions of the training and the test sets. The training and test sets are represented by blue and red spheres, respectively.



Table 6. Structure and HIV-1 IN Inhibitory Activities of *N*-Methyl-4-Hydroxypyrimidinone Carboxamides Included in the 3D-QSAR Model 4<sup>12,13</sup>

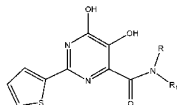
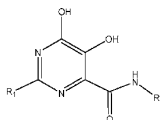

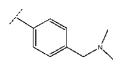

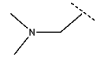
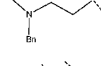
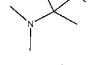
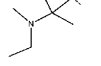
(a)				(c)		
						
Cpd	R <sub>1</sub>	R <sub>2</sub>	ST <sup>a</sup> IC <sub>50</sub> (μM)	Cpd	R	ST <sup>a</sup> IC <sub>50</sub> (μM)
104		Me	0.23	116		0.01
105		Me	0.003	117		0.19
106			0.002	118		0.10
107		Me	0.007	119		0.18
108		Me	0.008	120		0.20
109		Me	0.018	121		0.13
(b)				122		0.03
				123		0.061
Cpd	R <sub>1</sub>	ST <sup>a</sup> IC <sub>50</sub> (μM)		124		0.02
110		0.020		125		0.02
111		0.007		126		0.03
112		0.007		127		0.021
113		0.008		128		0.062
114		0.006				
115		0.004				

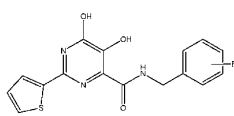
<sup>a</sup>ST represents strand transfer activity.

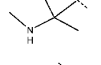
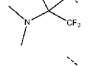
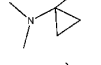
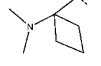
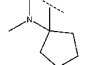
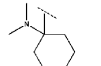
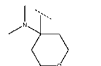
and the presence of the *N*-methyl carboxamide substituent further increased the steric bulk and prevented it from docking into the cavity formed by the IN–DNA interface. The docked alignment of the remaining 198 additional inhibitors was combined with the 98 compounds used in model 3, to obtain a new aligned molecule set of 296 compounds (Figure 11). Out of these, 60 compounds were selected as an external validation test set using Factor Analysis (in the SYBYL program) before developing 3D-QSAR model 4 with the remaining 236 compounds (training set). To select a varied test set, compounds were randomly chosen from each quadrant (Figure 12) to

prevent bias toward any functionality. It was observed that the compounds selected covered a wide range of activities and were quite representative of the structural classes making up the training set. PLS analysis of model 4 gave good  $q^2$  and  $r^2$  values of 0.70 and 0.84 for CoMFA and corresponding values of 0.75 and 0.89 for CoMSIA, respectively, for four PLS components (Table 15). The residuals of the training and the test set are listed in Tables 16 and 17, respectively. The models were further validated by group cross-validation, which gave  $q^2$  values of 0.70 and 0.75 for CoMSIA and CoMFA, respectively. The results of boot strapping and activity scrambling analysis (Table 15)

Table 7. Structure and HIV-1 IN Inhibitory Activities of Dihydroxypyrimidine Carboxamides Included in the 3D-QSAR Model 4<sup>53–55</sup>

(a)				(c)			
							
Cpd	R	R <sub>1</sub>	ST <sup>a</sup> IC <sub>50</sub> (μM)	Cpd	R <sub>1</sub>	R <sub>2</sub>	ST <sup>a</sup> IC <sub>50</sub> (μM)
129	CH <sub>2</sub> Ph	H	0.085	158	H	CH <sub>2</sub> -4F-Ph	0.06
130	CH <sub>2</sub> CH <sub>2</sub> Ph	H	0.02	159	CH <sub>3</sub>	CH <sub>2</sub> -4F-Ph	0.06
131	Ph	H	1.0	160	Ph	CH <sub>2</sub> -Ph	0.07
132	1-Naphthalene	H	0.03	161	CH <sub>2</sub> -Ph	CH <sub>2</sub> -4F-Ph	0.05
132b	1-Naphthalene	CH <sub>3</sub>	0.21	162	CH <sub>2</sub> NMe <sub>2</sub>	CH <sub>2</sub> -4F-Ph	0.20
133	( <i>S</i> )-CH(CH <sub>3</sub> )-2-naphthyl	H	0.04	163	C(Me) <sub>2</sub> NMe <sub>2</sub>	CH <sub>2</sub> -4F-Ph	0.05
134	( <i>R</i> )-CH(CH <sub>3</sub> )-2-naphthyl	H	0.61	164	<i>i</i> Pr	CH <sub>2</sub> -4F-Ph	0.10
135	( <i>S</i> )-1-Indane	H	0.02	165	<i>p</i> -tolyl	CH <sub>2</sub> -4F-Ph	0.02
136	( <i>R</i> )-1-Indane	H	0.61	166	2-thiazolyl	CH <sub>2</sub> -4F-Ph	0.02
137	CH <sub>2</sub> -4-pyridine	H	16.6	167	2-pyridyl	CH <sub>2</sub> -4F-Ph	0.05
138	CH <sub>2</sub> -2-thiophene	H	0.10	168		CH <sub>2</sub> -4F-Ph	0.15
139	CH <sub>2</sub> -2-(1,3-Thiazole)	H	0.20	169		CH <sub>2</sub> -4F-Ph	0.15
140	CH <sub>2</sub> -2-benzothiophene	H	0.01	170		CH <sub>2</sub> -4F-Ph	1.23
141	CH <sub>2</sub> -3-benzothiophene	H	0.05	171		CH <sub>2</sub> -4F-Ph	0.20
142	CH <sub>2</sub> -3-1 <i>H</i> -indole	H	0.02	172		CH <sub>2</sub> -4F-Ph	2.00
143	( <i>S</i> )CH(CH <sub>3</sub> )Ph	H	0.50	173		CH <sub>2</sub> -4F-Ph	0.05
144	( <i>R</i> )CH(CH <sub>3</sub> )Ph	H	5.00	174		CH <sub>2</sub> -4F-Ph	0.06

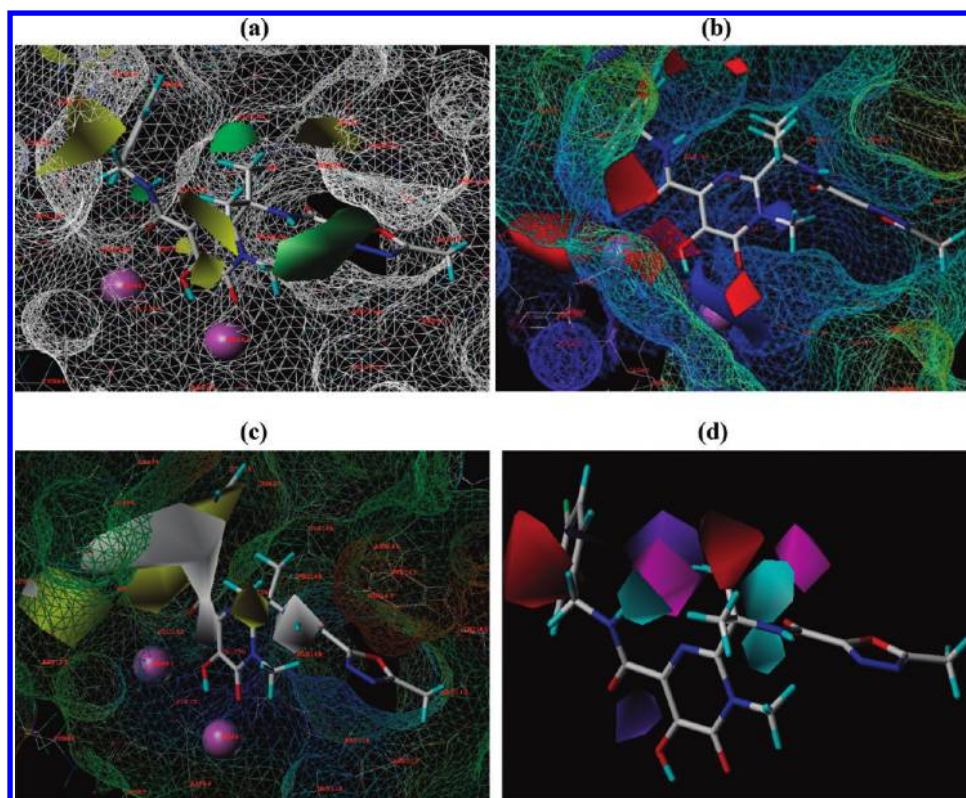
(b)		
		
Cpd	R	ST <sup>a</sup> IC <sub>50</sub> (μM)
145	2-OMe	0.09
146	3-OMe	0.05
147	4-OMe	0.55
148	2-Me	0.11
149	3-Me	0.04
150	4-Me	0.07
151	2-Ph	0.13
152	3-Ph	0.05
153	4-Cl	0.02
154	2-F	0.16
155	3-F	0.07
156	4-F	0.01
157	3,4-Cl	0.01

175		CH <sub>2</sub> -4F-Ph	0.04
176		CH <sub>2</sub> -4F-Ph	0.008
177		CH <sub>2</sub> -4F-Ph	0.01
178		CH <sub>2</sub> -4F-Ph	0.052
179		CH <sub>2</sub> -4F-Ph	0.088
180		CH <sub>2</sub> -4F-Ph	0.08
181		CH <sub>2</sub> -4F-Ph	0.026

<sup>a</sup>ST represents strand transfer activity.

further demonstrated the stability and robustness of the models. Finally, the models were validated by the external test set of 60

compounds, which were predicted by CoMFA and CoMSIA models with good predictive  $r^2$  values of 0.64 and 0.66, respectively.



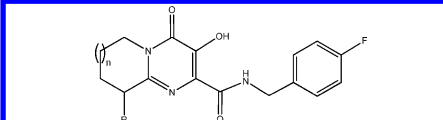
**Figure 14.** Contour maps of model 1 superimposed on raltegravir. (a) CoMFA steric contour, green and yellow contours illustrate regions where steric bulk has favorable and unfavorable effects on activity, respectively. (b) CoMFA electrostatic contour, blue and red isopleths indicate regions where an increase of positive and negative charge is required for activity. (c) CoMSIA hydrophobic contour, yellow and white contours enclose areas favorable for hydrophobic and hydrophilic groups, respectively. (d) CoMSIA H-bond donor and acceptor contour, purple contours indicate regions where the H-bond donor increases activity, whereas a cyan contour illustrates regions where the H-bond donor decreases activity. Magenta isopleths indicate areas where the H-bond acceptor increases affinity, while red isopleths indicate areas where the H-bond acceptor decreases affinity. The Connolly surface of IN–DNA is shown by lines.  $\text{Mg}^{2+}$  ions are shown by magenta spheres; raltegravir is colored by element and represented in tubes.

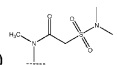
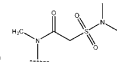
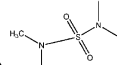
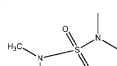
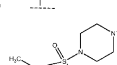
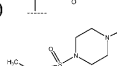
A common outlier, compound **91**, was poorly predicted by both the models, and its removal improved the CoMFA and CoMSIA predictive  $r^2$  values to 0.69 and 0.71, respectively. The CoMFA and CoMSIA predictive curves for the training and the test sets (after removal of compound **91**) are shown in Figure 13a and b, respectively. The models developed might be useful in the design of new related HIV-1 IN inhibitors.

**Contour Maps. Model 1: CoMFA Contour Maps.** The PLS coefficients derived from CoMFA steric and electrostatic contours mapped around raltegravir are displayed in Figure 14a and b, respectively. The IN–DNA binding site is represented as a Connolly surface, which was generated by the MOLCAD program in the SYBYL modeling package. There are green and yellow contours in the region close to the *gem*-dimethyl side chain. This is the region where substantial modifications were made that eventually led to the optimized *gem*-dimethyl oxadiazole carboxamide moiety. This side chain is positioned such that the oxadiazole ring occupies a pocket, which is indicated by a green contour in the cleft formed by residues Asn117, Pro142, Tyr143, Asn144, and Pro145. The *N*-hydroxy-alkyl side chain of the quinolone ring of elvitegravir also occupies the green contour and thus shows nearly 6 times improved activity relative to the unsubstituted analog (**84** vs **87**). The two yellow contours above the pyrimidinone ring indicate that steric bulk is unfavorable in this region. The *N*-acetyl substituted piperidine ring of compound **77** fits into a

yellow contour, which might explain the decrease in activity of this compound and other structural analogs of 4-quinolone-3-carboxylic acids with a bulky phenyl ring, such as compounds **100** and **101**. Moreover, their hydroxyethyl side chains also point toward the steric disfavored yellow contour region, further supporting a decrease in activity of these analogs. There are several yellow regions around the hydrophobic phenyl ring, suggesting that steric bulk is not favored at these positions possibly because of clashing with the DNA. This might be one of the reasons why fluorine that has a similar size to hydrogen is the optimal substituent on this ring. The electrostatic map of CoMFA shows a blue contour close to the pyrimidinone ring, implying that groups that have electropositive character are favored for activity. This is supported by the presence of electron-deficient nitrogen-containing ring systems in many different classes of IN inhibitors including the pyrimidinone ring of raltegravir and the quinolone ring of elvitegravir, as well as the acidic phenolic hydrogen in this region. There are three red contours close to the metal-chelating pharmacophore of raltegravir, indicating that the presence of electron-rich groups is favored for activity. There is also a red contour at the *ortho* position on the fluorobenzyl group. This is explained by the contour superimposed on elvitegravir, which reveals the *ortho*-fluoro substituent occupying these red contours. A distal red contour which is close to the *meta* position on the benzyl group of elvitegravir again suggests that negative charge is favorable



**Table 8. Structure and HIV-1 IN Inhibitory Activities of Bicyclic Pyrimidinones Included in the 3D-QSAR Model 4<sup>57</sup>**


Cpd	R	n	ST <sup>a</sup> IC <sub>50</sub> (μM)
182	H	0	0.008
183	H	1	0.005
184	H	2	0.019
185		1	0.007
186		1	0.007
187		1	0.007
188		1	0.005
189		1	0.010
190		1	0.007

<sup>a</sup>ST represents strand transfer activity.

in this region. This is reflected by the presence of an electronegative chlorine substituent at this position in elvitegravir and related analogs. The CoMFA contours help explain the loss in activity of compounds **37**, **38**, and **39**, which have cyclopentyl and substituted polycyclic naphthalene rings placed in steric bulk disfavored yellow-contoured regions.

**Model 1: CoMSIA Contour Maps.** CoMSIA steric and electrostatic contours mapped onto raltegravir are similar to those of the CoMFA model in having two yellow contours which result from the steric clash with the amide backbone of Tyr143 and the solvent accessible surface. Like in the CoMFA map, there is a green contour which is occupied by the carboxamide moiety of the oxadiazole side chain. However, in compound **78**, the side chain is oriented toward the yellow contours and explains the loss in activity. Another steric bulk-favored green region is occupied by the *p*-fluorobenzyl group, which has a positive influence on potency. This is confirmed by the significant loss of activity of compound **61**, which lacks a phenyl at this position. The green contour is flanked by a large yellow contour, suggesting that bulk tolerance within the cavity is quite limited. This could be the reason behind the reduced potency of compounds **32** and **37** in the chalcone series whose methoxy and bulky cyclopentyl groups fit into the disfavored yellow region. Similarly for compounds **38** and **39**, a loss in activity might result from the naphthyl rings placed in the bulk disallowed yellow region. The phenyl ring in compounds **89**, **96**, and **99–101** fits into the disfavored yellow region, and that might account for the significant loss in activity of these analogs.

The CoMSIA electrostatic map shows two red contours close together near the oxadiazole ring which is involved in  $\pi$ -stacking interactions with the phenyl ring of the Tyr143.

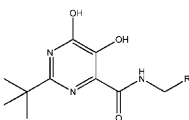
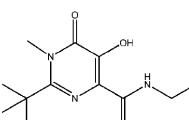








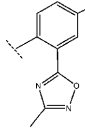
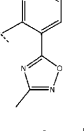


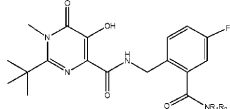


Similar to the CoMFA electrostatic contour map, there is a blue contour over the hydroxyl group of raltegravir. This possibly reflects the ionizable acidic OH group of raltegravir. Lastly, there are two small red contours around the oxygen of the methoxy and the hydroxyl groups in elvitegravir, correctly predicting the requirement of electron rich groups at this location.

CoMSIA hydrophobic contours superimposed on raltegravir (Figure 14c) show the presence of two large yellow regions, indicating that hydrophobic groups at these positions are favored for activity. While one of the yellow contours is occupied by the phenyl ring in raltegravir, the phenyl ring of elvitegravir is oriented toward the second yellow contour. There is also a small yellow region at the *para* position on the phenyl ring. Indeed, the presence of hydrophobic halogens like bromine in compound **6** or chlorine in compound **8** suggest the presence of hydrophobic interactions at this position. A hydrophilic contour occurs around the ribose and the phosphate regions of the viral DNA. This voluminous white contour encloses the carboxamide moiety and extends up to the central hydroxyl substituent in the pyrimidinone ring of raltegravir. The reduced potency of compound **78** may be explained by having the hydrophobic methyl group of its *N*-methyl carboxamide moiety located in the white region, which is counter to the preference of a hydrophilic moiety. A yellow contour fits over the methyl, and a white contour fits over the carbonyl group of the oxadiazole carboxamide side chain in raltegravir in agreement with their hydrophobic and hydrophilic characters, respectively.

The H-bond donor–acceptor contour plots mapped onto raltegravir are shown in Figure 14d. The magenta contour signifies that acceptor groups on the inhibitor are favored at that location, whereas the red contour encompasses a region where H-bond acceptors are disfavored, probably because of the presence of the carbonyl oxygen of the amide backbone of Tyr143 in the receptor. Another magenta contour is merged within the purple contour, suggesting an H-bond acceptor is favored but a donor is disfavored, which relate to groups that interact with the backbone NH of Pro145. It is close to the 8' position of the quinolone ring in elvitegravir, indicating that acceptor groups are well tolerated at this position. There is another red contour above the phenyl ring at the *ortho* and *meta* positions in raltegravir, implying that H-bond acceptors are disfavored at these positions. One of the purple regions is merged with and complements the magenta contour at the 8' position of the quinolone ring. The other purple region appears to complement the presence of the amino group of Gln148 in the receptor site, signifying that the presence of a H-bond donor at this region is detrimental to activity. Cyan contours indicate regions where H-bond donor groups increase activity. One of the cyan contours fits the NH group of the pyrimidinone carboxamide functionality in raltegravir, while the other is located in the region close to the hydroxyl of the alkyl hydroxyl group of elvitegravir.

**Model 2 Contours.** The CoMFA steric (Figure 15a) and electrostatic (Figure 15b) contours on raltegravir arising from model 2 have similarities to the contours obtained in model 1. One green contour again fits into the oxadiazole ring, suggesting that bulk is favored. Nearby yellow regions similar to those in the CoMFA contour of model 1 indicate sterically disallowed regions. Likewise, there are several red and one blue contour around the metal-chelating pharmacophore. Similar to CoMFA of model 1, the contours also explain the loss of

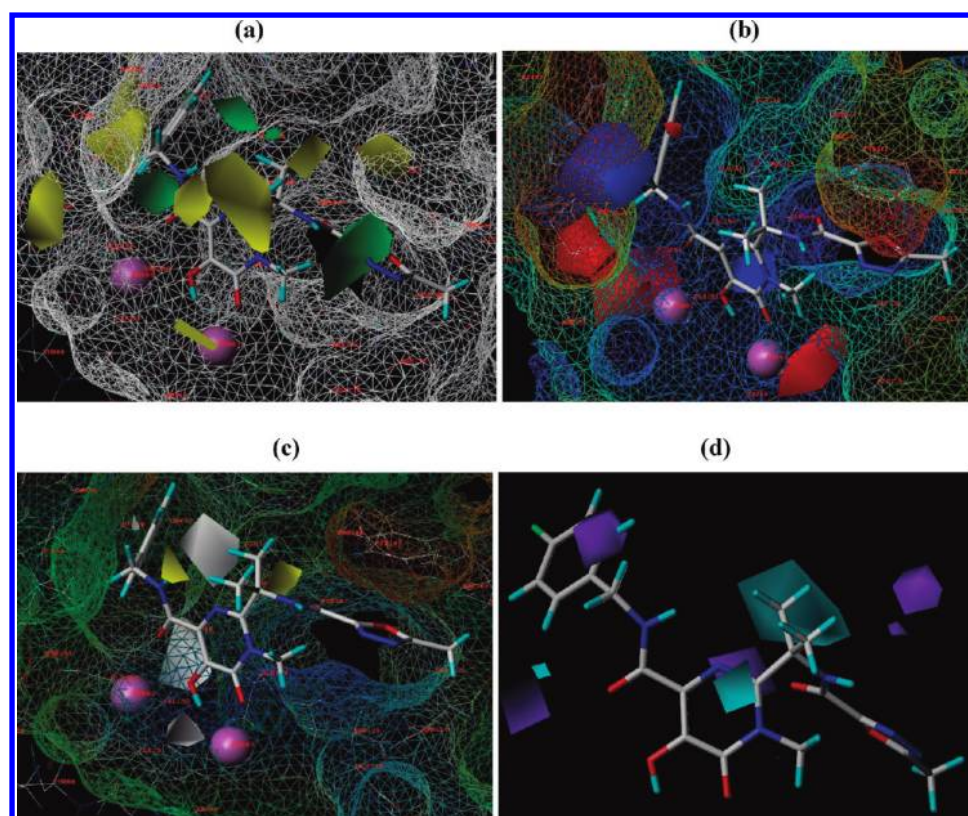
Table 9. Structure and HIV-1 IN Inhibitory Activities of *tert*-Butyl-*N*-Methyl Pyrimidones Included in the 3D-QSAR Model 4<sup>58</sup>

(a)			(b)		
					
Cpd	R	ST <sup>a</sup> IC <sub>50</sub> (μM)	Cpd	R	ST <sup>a</sup> IC <sub>50</sub> (μM)
191		0.01	197		0.01
192		0.01	198		0.02
193		0.01	199		0.02
194		0.02	200		0.02
195		0.02	201		0.02
196		0.06	202		0.02
(c)					
					
Cpd	NR <sub>1</sub> R <sub>2</sub>	ST <sup>a</sup> IC <sub>50</sub> (μM)			
203	NH <sub>2</sub>	0.02			
204	NHCH <sub>2</sub> CH <sub>3</sub>	0.07			
205	NH(CH <sub>2</sub> ) <sub>2</sub> CH <sub>3</sub>	0.07			
206	NHCH(CH <sub>3</sub> ) <sub>2</sub>	0.09			
207	NHCH <sub>2</sub> CF <sub>3</sub>	0.09			
208	N(CH <sub>3</sub> ) <sub>2</sub>	0.18			
209		0.23			
210		0.49			

<sup>a</sup>ST represents strand transfer activity.

inhibitory activity of compounds **61**, **78**, **89**, **96**, **97**, and **99–101**. In addition, the CoMFA contour map further reveals green regions around the two *meta* positions on the phenyl ring, suggesting that bulky groups could be added at these positions. One of the green contours is occupied by the methyl of the methoxy group in elvitegravir, corroborating the increase in activity seen with the presence of methoxy substituents in compounds **85** vs **86** and **87** vs **88**.

The CoMSIA steric and electrostatic contours projected onto raltegravir reveal SAR information similar to CoMSIA isopleths of model 1. However, there is a distant yellow contour which fits around the phenyl ring of the chalcone derivatives. In alignment II, the phenyl rings of the chalcones were oriented in the opposite (180°) direction to those in alignment I and not within the cavity formed by the displacement of 3'-adenosine of the viral DNA.

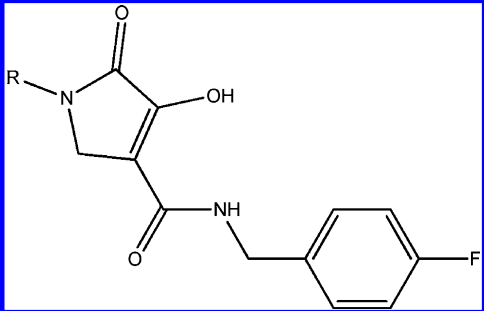


**Figure 15.** Contour maps of steric and electrostatic fields of model 2 projected on raltegravir. (a) CoMFA steric contour: green and yellow contours illustrate regions where steric bulk has favorable and unfavorable effects on activity, respectively. (b) CoMFA electrostatic contour: blue and red isopleths indicate regions where an increase of positive and negative charge is required for activity. (c) CoMSIA hydrophobic contour: yellow and white contours enclose areas favorable for hydrophobic and hydrophilic groups, respectively. (d) CoMSIA H-bond donor and acceptor contour: purple contours indicate regions where H-bond donor decreases activity, whereas cyan contour illustrates regions where H-bond donor increases activity. Connolly surface of IN–DNA is shown by lines.  $Mg^{2+}$  ions are shown by magenta spheres; raltegravir is colored by element and represented in tubes.

The hydrophobic contour of CoMSIA projected on raltegravir (Figure 15c) shows two white regions around the metal-chelating pharmacophore and another similar white region below the NH of the carboxamide moiety in raltegravir and oxygen of the methoxy substituent in elvitegravir. Again as in model 1, a yellow contour close to the *gem*-dimethyl substituent is observed, suggesting a hydrophobic favored region. The H-bond donor and acceptor contours (Figure 15d) show a large cyan isopleth located on the hydroxyl group of the hydroxalkyl side chain in elvitegravir. This appears to complement the red polyhedra observed in the corresponding contour maps of model 1. There are small cyan and magenta contours close to the pyrimidinone ring of raltegravir showing the positive role of a H-bond donor. There are two purple contours around the benzyl groups of raltegravir and elvitegravir implying that H-bond donors are not tolerated on the phenyl ring. One of the purple contours fits over the *meta*-chloro substituent of elvitegravir, suggesting the preference for a H-bond acceptor at this position.

**Model 3 Contours.** Contours resulting from docking-based alignment III are mostly similar to the contours that resulted from models 1 and 2. Since details of these contours have been described above, only the contours that differ markedly are further discussed. CoMFA and CoMSIA steric and electrostatic contours are shown in Figure 16a and b, respectively. A large green contour observed in the CoMFA map is directed toward the *gem*-dimethyl substituent, indicating the importance of

**Table 10.** Structure and HIV-1 IN Inhibitory Activities of Hydroxypyrrrolinone Carboxamides Included in the 3D-QSAR Model 4<sup>59</sup>

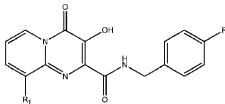
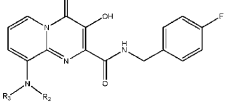









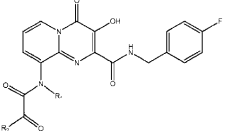



		
cpd	R	ST <sup>a</sup> IC <sub>50</sub> (μM)
211	CH <sub>2</sub> -Ph	0.1
212	CH <sub>2</sub> -3-Me-Ph	0.006
213	(CH <sub>2</sub> ) <sub>2</sub> -2-Me-Ph	0.004
214	(CH <sub>2</sub> ) <sub>3</sub> -Ph	0.005
215	CH <sub>2</sub> -2-pyridine	0.094

<sup>a</sup>ST represents strand transfer activity.

this sterically favored group. Large yellow contours near the oxadiazole ring are evident in both CoMFA and CoMSIA maps and may have resulted from the lack of space beyond the oxadiazole shown by the surface protrusion on the right upper



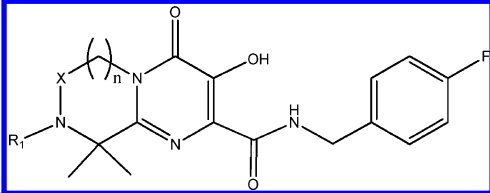
Table 11. Structure and HIV-1 IN Inhibitory Activities of Pyridopyrimidinone Carboxamides Included in the 3D-QSAR Model 4<sup>60</sup>
























(a)			(b)			
						
Cpd	R <sub>1</sub>	ST <sup>a</sup> IC <sub>50</sub> (μM)	Cpd	R <sub>2</sub>	R <sub>3</sub>	ST <sup>a</sup> IC <sub>50</sub> (μM)
216	H	0.022	223	Et		0.185
217	NHMe	0.13	224	Et		0.023
218	NHEt	0.036	225	Et		0.05
219	NHCOPh	0.018	226	Et		0.021
220	Me	0.009	227	Me		0.05
221		0.17	228	Et		0.021
222		0.185	229	Et		0.018
(c)						
						
Cpd	R <sub>1</sub>	R <sub>2</sub>	ST <sup>a</sup> IC <sub>50</sub> (μM)			
230	Me	NMe <sub>2</sub>	0.026			
231	Et	NMe <sub>2</sub>	0.036			
232	nPr	NMe <sub>2</sub>	0.018			
233	iPr	NMe <sub>2</sub>	0.021			
234	Et	NEt <sub>2</sub>	0.019			
235	Et		0.025			
236	Et		0.021			
237	Et		0.021			

<sup>a</sup>ST represents strand transfer activity.

side of the oxadiazole ring of raltegravir. Similarly, a green common contour is mapped onto the phenyl ring, suggesting that bulk is favored inside the cavity. In the CoMFA map, a yellow contour is observed close to the benzyl moiety, implying the possibility of steric clash with the surrounding DNA. Electrostatic contours show several red contours close to the metal-chelating oxygen atoms, implying that electronegative groups increase activity. A blue contour common in both CoMFA and CoMSIA maps is also present near the metal-

chelating pharmacophore. These contours, which suggest electropositive or H-donor groups, are favored for activity resulted from the acidic phenolic OH residues. A large blue contour (near the green region) observed in CoMFA is indicative of positive groups interacting with the surrounding deoxyribose phosphate backbone in the DNA, an insight that could not have been obtained without a complete HIV IN–DNA model. Both CoMFA and CoMSIA maps have blue contours mapped onto the pyrimidinone ring, suggesting that the electropositive ring

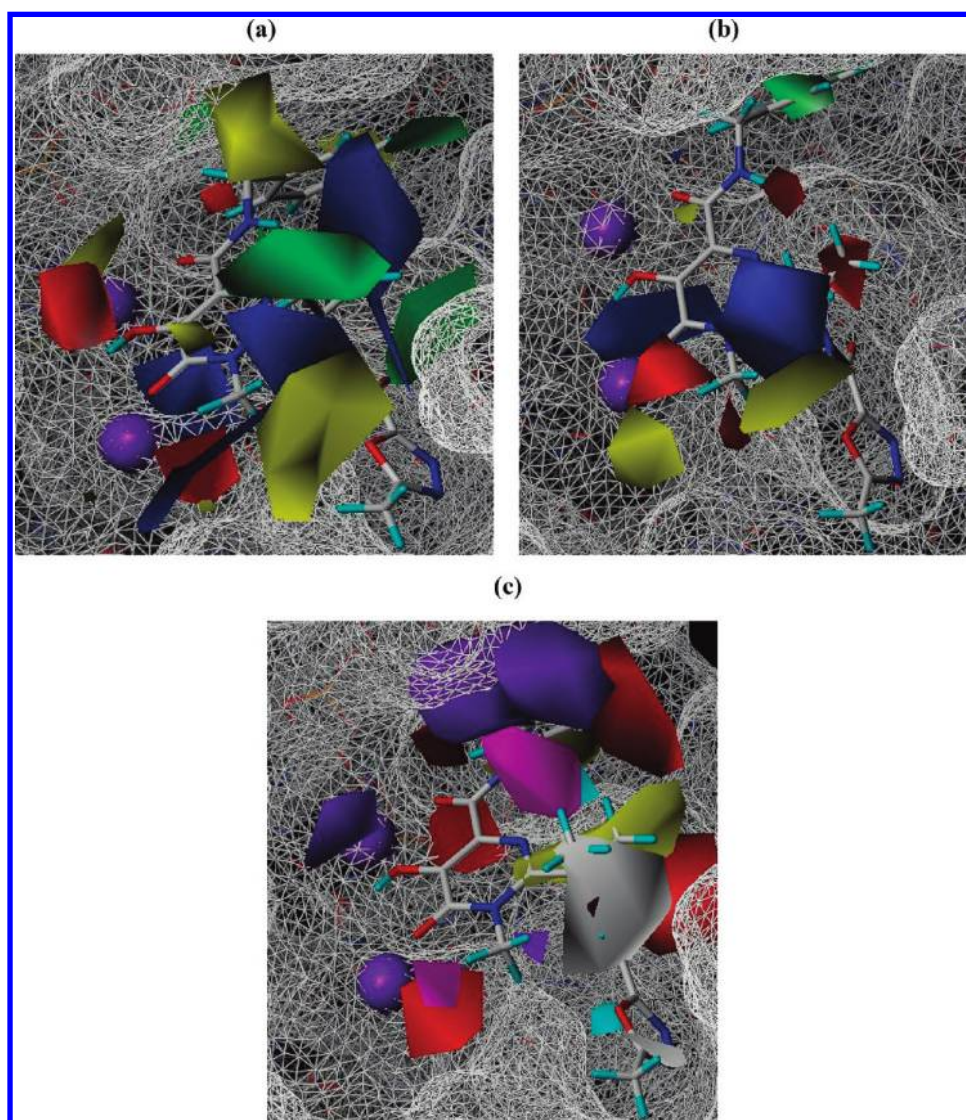
Table 12. Structure and HIV-1 IN Inhibitory Activities of Pyrazinopyrimidine Carboxamides Included in the 3D-QSAR Model 4<sup>61</sup>


Cpd	n	X	R <sub>1</sub>	ST <sup>a</sup> IC <sub>50</sub> (μM)	Cpd	n	X	R <sub>1</sub>	ST <sup>a</sup> IC <sub>50</sub> (μM)
238	0	CH <sub>2</sub>	CH <sub>3</sub>	0.03	267	1	CH <sub>2</sub>		0.006
239	0	CO	H	0.032	268	1	CH <sub>2</sub>		0.007
240	1	CO	H	0.005	269	1	CH <sub>2</sub>		0.003
241	1	SO <sub>2</sub>	H	0.007	270	1	CH <sub>2</sub>		0.009
242	1	CH <sub>2</sub>	H	0.016	271	1	CH <sub>2</sub>		0.004
243	1	CH <sub>2</sub>	CH <sub>3</sub>	0.006	272	1	CH <sub>2</sub>		0.025
244	2	CH <sub>2</sub>	H	0.02	273	1	CH <sub>2</sub>		0.021
245	1	CH <sub>2</sub>	COCH <sub>3</sub>	0.005	274	1	CH <sub>2</sub>		0.005
246	1	CH <sub>2</sub>	COCON(Me) <sub>2</sub>	0.006	275	1	CH <sub>2</sub>		0.008
247	1	CH <sub>2</sub>	SO <sub>2</sub> CH <sub>3</sub>	0.01	276	1	CH <sub>2</sub>		0.062
248	1	CH <sub>2</sub>	CONHCH <sub>2</sub> CH <sub>3</sub>	0.015	277	1	CH <sub>2</sub>		0.013
249	1	CH <sub>2</sub>	SO <sub>2</sub> N(CH <sub>3</sub> ) <sub>2</sub>	0.008	278	1	CH <sub>2</sub>		0.006
250	1	CH <sub>2</sub>	(CH <sub>2</sub> ) <sub>2</sub> N(CH <sub>3</sub> ) <sub>2</sub>	0.062	279	1	CH <sub>2</sub>		0.002
251	1	CH <sub>2</sub>	(CH <sub>2</sub> ) <sub>2</sub> -morpholine	0.003	280	1	CH <sub>2</sub>		0.006
252	1	CH <sub>2</sub>	CH <sub>2</sub> CONH <sub>2</sub>	0.005	281	1	CH <sub>2</sub>		0.002
253	1	CH <sub>2</sub>	CH <sub>2</sub> CONHCH <sub>3</sub>	0.021	282	1	CH <sub>2</sub>		0.005
254	1	CH <sub>2</sub>	CH <sub>2</sub> CON(CH <sub>3</sub> ) <sub>2</sub>	0.008	283	1	CH <sub>2</sub>		0.006
255	1	CH <sub>2</sub>	CH <sub>2</sub> CO-morpholine	0.007					
256	1	CH <sub>2</sub>	CH <sub>2</sub> CO-pyrrolidine	0.007					
257	1	CH <sub>2</sub>	CH <sub>2</sub> CO-(4-Me-piperazine)	0.021					
258	1	CH <sub>2</sub>	CH <sub>2</sub> COOH	0.008					
259	1	CH <sub>2</sub>	CH <sub>2</sub> CN	0.005					
260	1	CH <sub>2</sub>	Ph	0.005					
261	1	CH <sub>2</sub>		0.009					
262	1	CH <sub>2</sub>		0.012					
263	1	CH <sub>2</sub>		0.006					
264	1	CH <sub>2</sub>		0.005					
265	1	CH <sub>2</sub>		0.007					
266	1	CH <sub>2</sub>		0.009					

<sup>a</sup>ST represents strand transfer activity.

system promotes IN inhibitory activity. The CoMSIA hydrophobic, H-bond donor, and H-bond acceptor contours are shown in Figure 16c. There are yellow contours that map onto the hydrophobic phenyl ring and *gem*-dimethyl substituent, suggesting that hydrophobic groups are favored at these positions. There are white hydrophilic regions which map on

the carboxamide moiety and on the oxadiazole ring. The H-bonding contours show two purple contours close to the *o* position on the phenyl ring and a red contour close to the *m* position, indicating that H-bond donors and acceptors are disfavored at these positions, respectively. There is a purple contour close to the metal-chelating atoms, implying that H-bond donors



**Figure 16.** Contour maps of model 3 projected on raltegravir. (a) CoMFA steric and electrostatic contours. (b) CoMSIA steric and electrostatic contours: green and yellow contours illustrate regions where steric bulk has favorable and unfavorable effects on activity, respectively, and blue and red isopleths indicate regions where an increase of positive and negative charge is required for activity. (c) CoMSIA hydrophobic, H-bond donor and acceptor contours: yellow and white contours enclose areas favorable for hydrophobic and hydrophilic groups, respectively, and purple contours indicate regions where the H-bond donor decreases activity, whereas the cyan contour illustrates regions where the H-bond donor increases activity. Magenta isopleths indicate areas where the H-bond acceptor increases affinity, while red isopleths indicate areas where the H-bond acceptor decreases affinity. Connolly surface of IN–DNA is shown by lines.  $\text{Mg}^{2+}$  ions are shown by purple spheres; raltegravir is colored by element and represented in tubes.

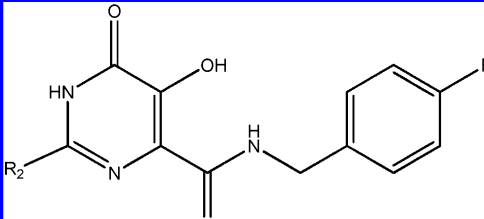
are disfavored. Overlapping magenta and red contours suggest that the H-bond acceptor is favored, while beside it the H-bond donor is disfavored. Overlapping cyan and magenta contours occur around the NH of the benzyl carboxamide moiety, indicating that NH can act as both a H-bond donor and acceptor at this position. Thus, the contour maps of model 3 reinforced the SAR features explained by contour maps of 3D-QSAR models 1 and 2. This is not surprising since the docking-based alignment III was in a way a composite of alignments I and II.

**Model 4 Contours.** There is a general concordance between the contour maps of model 3 and model 4, although some contours show differences in size or shifts in location. CoMFA contours of model 4 are depicted in Figure 17a, and the CoMSIA contours are in Figures 17b and c. CoMFA steric contours show a yellow region close to the entrance of the

cavity, suggesting that any bulky group directed toward this region may sterically clash with the viral DNA similar to that of model 3 (Figure 16a). There is another yellow contour beyond the oxadiazole ring, indicating a clash with the protein around Tyr 143. There are green contours that map onto the *N*-methyl and *gem*-dimethyl groups, implying that bulky moieties are favored in this region. CoMSIA electrostatic contours show a blue region which fits the central pyrimidinone ring, again implying the importance of electropositive groups around this side of the ring. Finally, electron-rich favored red regions occur close to the metal chelating oxygen atoms. The CoMSIA steric contours are mostly similar to those of CoMFA, except the presence of a more prominent green contour inside the cavity that accommodates the fluorobenzyl group and the absence of the adjacent yellow contour. The CoMSIA hydrophobic and



**Table 13.** Structure and HIV-1 IN Inhibitory Activities of N-Methyl Pyrimidones Analogs Included in the 3D-QSAR Model 4<sup>56</sup>



Cpd	R <sub>2</sub>	ST <sup>a</sup> IC <sub>50</sub> (μM)
284		0.08
285		0.09
286		0.02

<sup>a</sup>ST represents strand transfer activity.

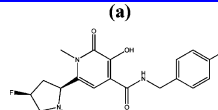
H-bonding contours show a large yellow isopleth inside the cavity indicating that hydrophobic groups are favored within the cavity. Similarly, yellow contours are also observed closed to the *gem*-dimethyl substituent. There is a white, hydrophilic-favored region which arises due the presence of the oxo-functionality of the pyrimidinone ring. The presence of cyan and magenta contours close together may be a result of the NH of the benzyl carboxamide moiety and the N-1 of the pyrimidine ring, which are H-bond donor and acceptor moieties, respectively. There are several purple and red contours indicating regions where H-bond donors and acceptors are both disfavored. Finally, the two magenta contours close to the purple isopleths suggest that H-bond acceptors are favored at these regions.

## CONCLUSION

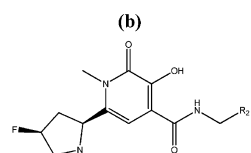
In the present study, we have explored the binding modes of potent HIV IN inhibitors by homology modeling, docking, and 3D-QSAR studies of the HIV IN–DNA complex. The homology models were constructed using the recently reported foamy virus IN–DNA X-ray crystal structure as a template. The binding modes of raltegravir and elvitegravir are in accordance with the foamy virus structure, revealing interactions important for inhibitor–IN binding. The *para*-fluorobenzyl group in raltegravir and the halogen-substituted benzyl group of elvitegravir insert into the cavity formed by displacement of the 3' end of the viral DNA and engage in strong electrostatic and vdW interactions with the DNA–IN interface. The metal-chelating pharmacophore of raltegravir and elvitegravir interacted with the two metal ions (Mg<sup>2+</sup>) in the IN active site.

Further, 3D-QSAR models were developed from several different classes of IN inhibitors using different alignment rules based on atom fitting and docking with the homology model. After initial docking and pose validation using MM-GBSA rescoring, a first docking-based QSAR model was developed that performed better than the atom-based alignment models, 1 and 2. The docking-based alignment rule was then used to

**Table 14.** Structure and HIV-1 IN Inhibitory Activities of 2-Pyrrolidinyl N-Methyl Pyrimidones Included in the 3D-QSAR Model 4<sup>62</sup>



Cpd	R	ST <sup>a</sup> IC <sub>50</sub> (μM)
287	Me	0.023
288	c-Pr	0.02
289	CH <sub>2</sub> (3-isoxazole)	0.01
290	COMe	0.015
291	CO <sub>2</sub> Me	0.01
292	CONMe <sub>2</sub>	0.01
293	SO <sub>2</sub> Me	0.015
294	SO <sub>2</sub> NMe <sub>2</sub>	0.009
295	CO-pyrazine	0.01
296	COCONMe <sub>2</sub>	0.021

Cpd	R <sub>1</sub>	R <sub>2</sub>	ST <sup>a</sup> IC <sub>50</sub> (μM)
297	COMe		0.024
298	COMe		0.027
299	COMe		0.032
300	CO <sub>2</sub> Me		0.020
301	CO <sub>2</sub> Me		0.019

<sup>a</sup>ST represents strand transfer activity.

**Table 15.** PLS Statistics of CoMFA and CoMSIA 3D-QSAR Model 4

PLS statistics	model 4	
	CoMFA	CoMSIA
q <sup>2</sup>	0.70	0.75
r <sup>2</sup>	0.84	0.89
s	0.55	0.46
F	244.98	376.09
PLS component	4	4
field contribution		
steric	0.358	0.095
electrostatic	0.642	0.312
hydrophobic		0.241
H donor		0.176
H acceptor		0.177
bootstrapping r <sup>2</sup> (20 runs)	0.88 ± 0.01	0.92 ± 0.01
cross-validation q <sup>2</sup> (20 runs)	0.70	0.75
scrambling (20 runs)	0.23	0.21

Table 16. Residuals of the Predictions of the Training Set by CoMFA and CoMSIA of Model 4

cpd	actual ST <sup>a</sup>	residuals		cpd	actual ST <sup>a</sup>	residuals	
	pIC <sub>50</sub>	CoMFA	CoMSIA		pIC <sub>50</sub>	CoMFA	CoMSIA
1	4.6	−0.15	−0.22	154	6.79	−0.35	0.02
2	4.85	0.46	0.19	155	7.15	0.43	0.19
3	4.49	0.03	0.02	157	8.00	0.99	0.43
4	5.04	0.72	0.35	158	7.22	−0.22	0.08
5	4.92	1.22	0.76	159	7.22	−0.07	0.26
6	5.3	0.97	0.59	160	7.15	−0.15	0.34
7	4.45	0.15	−0.26	161	7.3	−0.28	0.15
8	5.13	0.39	0.17	163	7.3	0.13	−0.03
9	4.72	0.09	−0.16	164	7.00	0.12	0.04
10	4.46	−0.003	−0.18	165	7.7	−0.17	0.64
11	4.29	−0.15	−0.41	166	7.7	−0.13	−0.002
12	4.74	0.33	0.2	169	6.82	−0.51	−0.21
13	4.79	0.11	−0.35	170	5.91	−1.9	−0.96
14	4.85	0.22	0.28	171	6.7	−1.37	−0.31
16	5.43	1.03	0.74	172	5.7	−1.48	−0.7
17	4.27	−0.26	−0.39	173	7.3	−0.03	−0.07
19	4.03	−0.78	−0.9	174	7.22	0.01	−0.11
20	4.6	−0.02	−0.01	175	7.39	−0.2	−0.2
21	4.76	0.24	0.13	176	8.09	0.56	0.89
22	4.52	0.12	−0.23	177	7.00	−0.03	−0.16
23	4.38	−0.17	−0.26	178	7.28	−0.4	−0.07
24	4.77	0.24	0.04	179	7.05	−0.07	−0.4
25	4.61	−0.01	−0.19	180	7.09	−0.58	−0.12
26	4.52	−0.29	−0.3	181	7.58	0.33	0.29
27	4.53	−0.22	−0.003	182	8.09	0.44	0.64
28	4.82	−0.46	0.29	183	8.3	0.74	0.72
29	4.4	0.18	−0.4	184	7.72	0.22	0.01
30	4.23	0.41	0.12	187	8.15	0.4	0.42
32	4.11	−0.15	−0.51	189	8.00	−0.16	−0.008
34	4.74	−1.05	0.01	190	8.15	0.19	−0.07
37	4.25	−0.12	−0.3	191	8.00	0.65	0.63
39	4.26	0.22	−0.67	192	8.00	0.97	1.36
40	4.95	0.47	0.25	193	8.00	0.81	0.44
41	4.79	−0.23	−0.06	194	7.7	−0.21	0.68
42	4.2	−0.38	−0.47	195	7.7	−0.26	−0.05
43	4.28	−0.27	−0.3	197	8.00	0.55	0.21
44	4.49	−0.05	−0.25	199	7.7	0.01	−0.005
45	4.74	0.14	0.15	202	7.7	0.29	−0.08
48	4.3	−0.22	−0.38	204	7.15	0.05	−0.27
49	4.16	−1.08	0.05	205	7.15	−0.36	−0.53
50	4.67	−0.38	0.28	207	7.04	−0.59	−0.52
51	4.82	−0.66	−0.64	208	6.74	−0.82	−0.76
53	4.00	−1.107	−0.687	210	6.31	0.04	−0.22
54	4.88	−0.5	−0.49	211	7.00	−0.8	0.03
55	8.00	−0.17	0.01	212	8.22	0.76	1.2
56	7.82	0.22	0.53	215	7.02	−0.62	−0.6
57	8.39	0.15	0.24	216	7.65	0.19	0.2
60	7.58	0.26	0.34	219	7.74	0.97	0.27
62	7.67	−0.34	0.8	221	6.77	−0.42	−0.47
64	8.04	−0.38	−0.23	223	6.73	−0.42	−0.25
65	7.82	0.9	0.85	225	7.3	0.1	0.02
68	6.92	0.17	0.07	226	7.67	−0.4	−0.12
69	6.65	0.5	0.04	227	7.3	0.08	−0.03
70	7.00	0.32	0.32	228	7.67	−0.04	0.59

Table 16. continued

cpd	actual ST <sup>a</sup>	residuals		cpd	actual ST <sup>a</sup>	residuals	
	pIC <sub>50</sub>	CoMFA	CoMSIA		pIC <sub>50</sub>	CoMFA	CoMSIA
71	6.16	−0.54	−0.44	229	7.74	0.05	−0.003
72	6.85	0.53	−0.24	230	7.58	0.33	0.52
73	7.00	−0.31	0.16	232	7.74	0.35	0.16
74	6.69	−0.3	−0.38	233	7.67	0.47	0.47
75	6.6	0.07	−0.16	234	7.72	0.25	0.03
76	6.88	0.27	0.07	235	7.6	0.11	0.07
77	5.39	−0.44	−0.25	236	7.67	0.03	0.03
78	5.3	−1.22	−0.37	237	7.67	0.15	0.26
79	6.75	0.36	0.28	238	7.52	−0.14	−0.08
80	7.2	−0.31	−0.27	239	7.49	−0.48	−0.87
81	7.72	0.65	0.53	240	8.3	−0.14	−0.15
82	6.85	0.15	0.2	241	8.15	0.6	0.3
83	5.79	−0.4	−0.55	242	7.79	0.07	−0.04
84	7.36	0.86	0.61	243	8.22	0.38	0.51
85	7.61	0.73	0.82	244	7.7	0.44	−0.21
87	8.08	1.15	0.96	245	8.3	0.6	0.61
88	8.14	0.8	0.63	246	8.22	0.3	0.3
93	6.22	0.25	0.08	247	8.00	−0.25	−0.34
94	4.62	−1.35	−1.14	248	7.82	0.47	0.27
96	4.00	−0.9	−0.2	249	8.09	0.65	0.05
97	4.00	−1.64	−1.29	250	7.2	−0.61	−0.51
99	4.00	0.16	0.63	251	8.52	0.42	0.11
100	4.00	−0.2	0.59	252	8.3	0.34	−0.12
101	4.00	−0.11	0.35	253	7.67	−0.41	−0.43
102	4.00	−0.66	0.35	254	8.09	0.07	−0.11
103	4.00	−0.95	−0.28	255	8.15	0.16	−0.34
104	6.63	−1.17	−0.89	256	8.15	0.2	−0.05
105	8.52	0.37	0.68	257	7.67	0.74	−0.25
106	8.69	0.45	0.39	258	8.09	−0.05	0.36
108	8.09	0.26	0.29	259	8.3	0.38	0.26
109	7.74	0.4	−0.006	261	8.04	0.23	0.31
110	7.7	−0.23	−0.48	262	7.92	0.82	0.27
111	8.15	0.39	−0.37	263	8.22	0.2	−0.06
112	8.15	0.13	−0.48	264	8.3	0.12	−0.41
113	8.09	0.18	−0.16	265	8.15	0.01	0.2
115	8.39	0.57	−0.33	267	8.22	0.3	−0.33
116	8.00	0.25	0.83	268	8.15	0.08	0.15
118	7.00	−0.49	−0.55	269	8.52	0.62	0.56
119	6.74	−0.94	−0.58	270	8.04	0.23	0.26
120	6.69	−0.93	−1.17	272	7.6	−0.23	−0.57
121	6.88	−0.47	−0.02	273	7.67	−0.18	−0.46
122	7.52	−0.26	−0.38	274	8.3	0.61	0.3
124	7.7	0.1	0.16	275	8.09	−0.15	−0.37
126	7.52	0.48	−0.46	276	7.2	−0.83	−0.97
127	7.67	−0.44	0.16	277	7.88	−0.26	−0.15
128	7.2	−0.88	−0.8	278	8.22	0.47	0.08
129	7.07	0.33	0.24	279	8.7	1.14	0.79
130	7.7	0.36	0.22	281	8.7	0.75	−0.07
132	7.52	0.44	0.69	282	8.3	−0.12	0.18
133	7.39	0.12	0.27	284	7.09	−0.08	−0.5
135	7.7	0.43	0.38	285	7.04	−0.35	−0.3
137	4.78	−1.98	−1.46	287	7.63	0.07	−0.08
138	7.00	−0.43	0.48	288	7.84	−0.14	0.41
139	6.7	−0.85	−0.12	289	8.00	0.04	0.56



Table 16. continued

cpd	actual ST <sup>a</sup>	residuals		cpd	actual ST <sup>a</sup>	residuals	
	pIC <sub>50</sub>	CoMFA	CoMSIA		pIC <sub>50</sub>	CoMFA	CoMSIA
144	6.77	−1.47	−1.3	291	8.00	0.4	0.25
145	7.04	0.16	−0.13	292	8.00	0.47	0.68
146	7.3	−0.1	0.56	294	8.04	0.03	−0.35
147	6.26	−0.58	−0.03	295	8.00	0.23	0.24
148	6.95	0.11	−0.07	296	7.67	−0.3	0.15
149	7.39	0.66	0.58	297	7.61	0.15	−0.41
150	7.15	0.31	0.4	298	7.56	−0.19	−0.18
151	6.88	−0.5	0.04	299	7.49	0.14	−0.03
152	7.3	0.87	0.08	300	7.7	0.05	0.2
153	7.7	0.62	0.34	301	7.72	0.19	−0.18

<sup>a</sup>ST represents strand transfer activity.

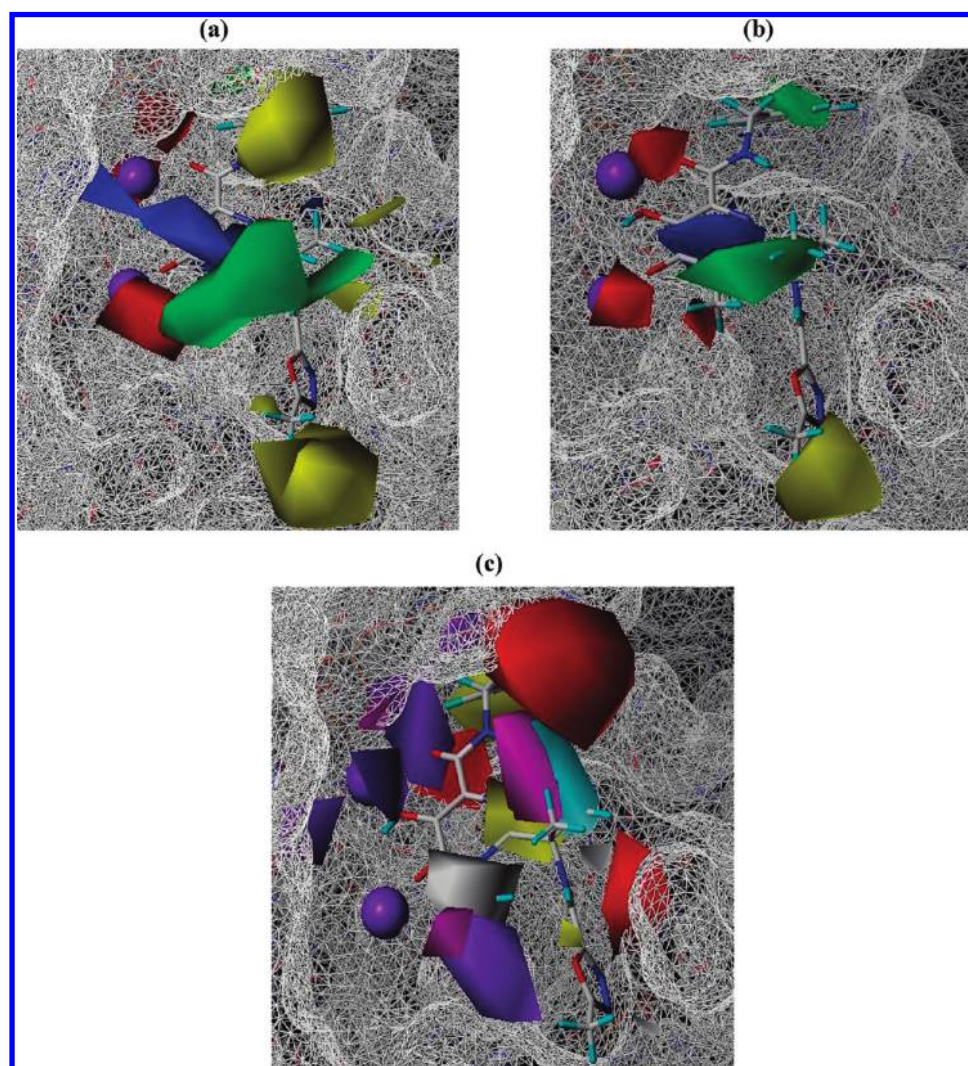
Table 17. Residuals of the Predictions of the Test Set by CoMFA and CoMSIA of Model 4

cpd	actual ST <sup>a</sup> pIC <sub>50</sub>	residuals		cpd	actual ST <sup>a</sup> pIC <sub>50</sub>	residuals	
		CoMFA	CoMSIA			CoMFA	CoMSIA
15	4.88	0.37	0.32	156	8.00	1.12	1.03
18	4.00	−0.28	−0.64	162	6.7	−0.19	−0.6
31	4.12	−0.81	−0.59	167	7.44	−0.14	0.29
33	5.04	−1.24	−0.68	168	7.01	−0.18	−0.38
35	4.79	−1.58	−0.82	185	8.15	0.67	1.55
36	4.74	0.42	0.12	186	8.15	0.96	0.67
38	4.33	0.08	0.24	188	8.3	0.57	0.14
46	4.23	−0.11	−0.28	196	7.22	−0.21	0.66
47	4.35	−1.32	−0.62	198	7.7	−0.02	−0.19
52	4.00	−1	−0.72	200	7.7	−0.21	0.68
58	7.82	0.13	0.71	201	7.7	0.13	0.09
59	7.7	1.48	1.42	203	7.7	1.75	1.01
61	5.3	−1.65	−1.84	206	7.04	−0.37	−0.56
66	7.92	0.5	0.49	209	6.63	0.36	0.52
67	7.82	0.13	−0.15	213	8.39	1.02	1.57
90	4.74	−1.45	−1.31	214	8.3	0.58	1.12
91 <sup>b</sup>	6.7	2.19	2.33	217	6.88	−0.61	−0.33
98	4.00	−2.54	−1.53	218	7.44	0.37	−0.02
107	8.15	0.46	0.45	220	8.04	0.86	0.47
114	8.22	0.41	−0.42	222	6.73	−0.49	−0.26
117	6.72	−0.67	−1.28	224	7.63	0.003	−0.001
123	7.21	−0.55	−0.53	231	7.44	0.17	0.15
125	7.7	−0.2	0.26	260	8.3	0.45	0.86
131	6.00	−1.08	−1.45	266	8.04	0.12	−0.04
134	6.21	−0.91	−0.61	271	8.39	0.54	0.15
136	6.21	0.34	−0.64	280	8.22	0.004	0.4
140	8.00	0.84	0.7	283	7.82	0.4	0.14
141	7.3	0.12	0.23	286	7.7	−0.18	−0.77
142	7.7	0.7	1.02	290	7.82	0.03	0.39
143	6.3	−0.38	−0.64	293	7.82	0.38	−0.11

<sup>a</sup>ST represents strand transfer activity. <sup>b</sup>Represents outlier.

build and validate a more extensive 3D-QSAR model, model 4, with good predictive abilities for CoMFA and CoMSIA. The CoMFA and CoMSIA contour maps highly complement the topology of the IN binding site and thus have the potential for interactive use to design novel HIV IN inhibitors. To the best of our knowledge, this is the largest 3D-QSAR study on HIV-1

IN inhibitors, and the first study that has combined two of the most advanced classes of IN inhibitors to derive a common model for IN inhibition. In the absence of a complete HIV–IN–DNA crystal structure, this structure-based 3D-QSAR approach is being used in attempts to improve the IN inhibitory activity of our recently discovered 3-ketosalicyclic acid series, as



**Figure 17.** Contour maps of model 4 projected on raltegravir. (a) CoMFA steric and electrostatic contours. (b) CoMSIA steric and electrostatic contours: green and yellow contours illustrate regions where steric bulk has favorable and unfavorable effects on activity, respectively, and blue and red isopleths indicate regions where an increase of positive and negative charge is required for activity. (c) CoMSIA hydrophobic, H-bond donor, and acceptor contours: yellow and white contours enclose areas favorable for hydrophobic and hydrophilic groups, respectively, and purple contours indicate regions where the H-bond donor decreases activity, whereas the cyan contour illustrates regions where H-bond donor increases activity. Magenta isopleths indicate areas where the H-bond acceptor increases affinity, while red isopleths indicate areas where the H-bond acceptor decreases affinity. Connolly surface of IN–DNA is shown by lines. Mg<sup>2+</sup> ions are shown by purple spheres; raltegravir is colored by element and represented in tubes.

well as to derive pharmacophores for 3D database searching, and to conduct virtual screening campaigns to discover novel HIV IN inhibitors.

## ■ ASSOCIATED CONTENT

### ● Supporting Information

Results of free energy calculations by Prime MM-GBSA, residuals of the predictions of the training set by CoMFA and CoMSIA of model 1, residuals of the predictions of the test set by CoMFA and CoMSIA of model 1, residuals of the predictions of the training set by CoMFA and CoMSIA of model 2, residuals of the predictions of the test set by CoMFA and CoMSIA of model 2, residuals of the predictions of the training set by CoMFA and CoMSIA of model 3, and residuals of the predictions of the test set by CoMFA and CoMSIA of model 3. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*Phone: (901) 448-7533. Fax: (901) 448-6828. E-mail: [jbuolamwini@uthsc.edu](mailto:jbuolamwini@uthsc.edu).

## ■ ACKNOWLEDGMENTS

Financial support from the National Institute of Allergy and Infectious Diseases (NIAID), NIH Grant No. AI084710, and the Department of Pharmaceutical Sciences is gratefully acknowledged.

## ■ REFERENCES

- (1) Sweeney, Z. K.; Klumpp, K. Improving non-nucleoside reverse transcriptase inhibitors for first-line treatment of HIV infection: the development pipeline and recent clinical data. *Curr. Opin. Drug Discovery Dev.* **2008**, *11*, 458–470.
- (2) von Hentig, N. Atazanavir/ritonavir: a review of its use in HIV therapy. *Drugs Today (Barc.)* **2008**, *44*, 103–132.

- (3) Pomerantz, R. J.; Horn, D. L. Twenty years of therapy for HIV-1 infection. *Nat. Med.* **2003**, *9*, 867–873.
- (4) Richman, D. D. HIV chemotherapy. *Nature* **2001**, *410*, 995–1001.
- (5) Manfredi, R.; Sabbatani, S. A novel antiretroviral class (fusion inhibitors) in the management of HIV infection. Present features and future perspectives of enfuvirtide (T-20). *Curr. Med. Chem.* **2006**, *13*, 2369–2384.
- (6) MacArthur, R. D.; Novak, R. M. Reviews of anti-infective agents: maraviroc: the first of a new class of antiretroviral agents. *Clin. Infect. Dis.* **2008**, *47*, 236–241.
- (7) Zhang, L.; Ramratnam, B.; Tenner-Racz, K.; He, Y.; Vesanen, M.; Lewin, S.; Talal, A.; Racz, P.; Perelson, A. S.; Korber, B. T.; Markowitz, M.; Ho, D. D. Quantifying residual HIV-1 replication in patients receiving combination antiretroviral therapy. *N. Engl. J. Med.* **1999**, *340*, 1605–1613.
- (8) Richman, D. D. HIV chemotherapy. *Nature* **2001**, *410*, 995–1001.
- (9) Esposito, D.; Craigie, R. HIV integrase structure and function. *Adv. Virus Res.* **1999**, *52*, 319–333.
- (10) Sakai, H.; Kawamura, M.; Sakuragi, J.; Sakuragi, S.; Shibata, R.; Ishimoto, A.; Ono, N.; Ueda, S.; Adachi, A. Integration is essential for efficient gene expression of human immunodeficiency virus type 1. *J. Virol.* **1993**, *67*, 1169–1174.
- (11) U.S. Food and Drug Administration. <http://www.accessdata.fda.gov/scripts/cder/drugsatfda/index.cfm?fuseaction=SearchDrugDetails> (accessed December 22, 2011).
- (12) Summa, V.; Petrocchi, A.; Bonelli, F.; Crescenzi, B.; Donghi, M.; Ferrara, M.; Fiore, F.; Gardelli, C.; Gonzalez Paz, O.; Hazuda, D. J.; Jones, P.; Kinzel, O.; Laufer, R.; Monteagudo, E.; Muraglia, E.; Nizi, E.; Orvieto, F.; Pace, P.; Pescatore, G.; Scarpelli, R.; Stillmock, K.; Witmer, M. V.; Rowley, M. Discovery of raltegravir, a potent, selective orally bioavailable HIV-integrase inhibitor for the treatment of HIV-AIDS infection. *J. Med. Chem.* **2008**, *51*, 5843–5855.
- (13) Gardelli, C.; Nizi, E.; Muraglia, E.; Crescenzi, B.; Ferrara, M.; Orvieto, F.; Pace, P.; Pescatore, G.; Poma, M.; Ferreira Mdel, R.; Scarpelli, R.; Homnick, C. F.; Ikemoto, N.; Alfieri, A.; Verdirame, M.; Bonelli, F.; Paz, O. G.; Taliani, M.; Monteagudo, E.; Pesci, S.; Laufer, R.; Felock, P.; Stillmock, K. A.; Hazuda, D.; Rowley, M.; Summa, V. Discovery and synthesis of HIV integrase inhibitors: development of potent and orally bioavailable N-methyl pyrimidones. *J. Med. Chem.* **2007**, *50*, 4953–4975.
- (14) Billich, A. S-1360 Shionogi-GlaxoSmithKline. *Curr. Opin. Invest. Drugs* **2003**, *4*, 206–209.
- (15) Hazuda, D. J.; Anthony, N. J.; Gomez, R. P.; Jolly, S. M.; Wai, J. S.; Zhuang, L.; Fisher, T. E.; Embrey, M.; Guare, J. P. Jr.; Egbertson, M. S.; Vacca, J. P.; Huff, J. R.; Felock, P. J.; Witmer, M. V.; Stillmock, K. A.; Danovich, R.; Grobler, J.; Miller, M. D.; Espeseth, A. S.; Jin, L.; Chen, I. W.; Lin, J. H.; Kassahun, K.; Ellis, J. D.; Wong, B. K.; Xu, W.; Pearson, P. G.; Schleif, W. A.; Cortese, R.; Emini, E.; Summa, V.; Holloway, M. K.; Young, S. D. A naphthyridine carboxamide provides evidence for discordant resistance between mechanistically identical inhibitors of HIV-1 integrase. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 11233–11238.
- (16) Sato, M.; Motomura, T.; Aramaki, H.; Matsuda, T.; Yamashita, M.; Ito, Y.; Kawakami, H.; Matsuzaki, Y.; Watanabe, W.; Yamataka, K.; Ikeda, S.; Kodama, E.; Matsuoka, M.; Shinkai, H. Novel HIV-1 integrase inhibitors derived from quinolone antibiotics. *J. Med. Chem.* **2006**, *49*, 1506–1508.
- (17) Pasquini, S.; Mugnaini, C.; Tintori, C.; Botta, M.; Trejos, A.; Arvela, R. K.; Larhed, M.; Witvrouw, M.; Michiels, M.; Christ, F.; Debyser, Z.; Corelli, F. Investigations on the 4-quinolone-3-carboxylic acid motif. 1. Synthesis and structure-activity relationship of a class of human immunodeficiency virus type 1 integrase inhibitors. *J. Med. Chem.* **2008**, *51*, 5125–5129.
- (18) Yuan, H.; Parrill, A. L. QSAR studies of HIV-1 integrase inhibition. *Bioorg. Med. Chem.* **2002**, *10*, 4169–4183.
- (19) Gupta, P.; Roy, N.; Garg, P. Docking-based 3D-QSAR study of HIV-1 integrase inhibitors. *Eur. J. Med. Chem.* **2009**, *44*, 4276–4287.
- (20) Lu, P.; Wei, X.; Zhang, R. CoMFA and CoMSIA 3D-QSAR studies on quinolone carboxylic acid derivatives inhibitors of HIV-1 integrase. *Eur. J. Med. Chem.* **2010**, *45*, 3413–3419.
- (21) Kuo, C. L.; Assefa, H.; Kamath, S.; Brzozowski, Z.; Slawinski, J.; Saczewski, F.; Buolamwini, J. K.; Neamati, N. Application of CoMFA and CoMSIA 3D-QSAR and docking studies in optimization of mercaptobenzenesulfonamides as HIV-1 integrase inhibitors. *J. Med. Chem.* **2004**, *47*, 385–399.
- (22) Buolamwini, J. K.; Assefa, H. 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.* **2002**, *45*, 841–852.
- (23) Sharma, H.; Patil, S.; Sanchez, T. W.; Neamati, N.; Schinazi, R. F.; Buolamwini, J. K. Synthesis, biological evaluation and 3D-QSAR studies of 3-keto salicylic acid chalcones and related amides as novel HIV-1 integrase inhibitors. *Bioorg. Med. Chem.* **2011**, *19*, 2030–2045.
- (24) Sato, M.; Kawakami, H.; Motomura, T.; Aramaki, H.; Matsuda, T.; Yamashita, M.; Ito, Y.; Matsuzaki, Y.; Yamataka, K.; Ikeda, S.; Shinkai, H. Quinolone carboxylic acids as a novel monoketo acid class of human immunodeficiency virus type 1 integrase inhibitors. *J. Med. Chem.* **2009**, *52*, 4869–4882.
- (25) Boros, E. E.; Edwards, C. E.; Foster, S. A.; Fuji, M.; Fujiwara, T.; Garvey, E. P.; Golden, P. L.; Hazen, R. J.; Jeffrey, J. L.; Johns, B. A.; Kawasui, T.; Kiyama, R.; Koble, C. S.; Kurose, N.; Miller, W. H.; Mote, A. L.; Murai, H.; Sato, A.; Thompson, J. B.; Woodward, M. C.; Yoshinaga, T. Synthesis and antiviral activity of 7-benzyl-4-hydroxy-1,5-naphthyridin-2(1H)-one HIV integrase inhibitors. *J. Med. Chem.* **2009**, *52*, 2754–2761.
- (26) Dubois, M.; Bailly, F.; Mbemba, G.; Mouscadet, J. F.; Debyser, Z.; Witvrouw, M.; Cotellet, P. Reaction of rosmarinic acid with nitrite ions in acidic conditions: discovery of nitro- and dinitro-rosmarinic acids as new anti-HIV-1 agents. *J. Med. Chem.* **2008**, *51*, 2575–2579.
- (27) Wang, J. Y.; Ling, H.; Yang, W.; Craigie, R. Structure of a two-domain fragment of HIV-1 integrase: implications for domain organization in the intact protein. *EMBO J.* **2001**, *20*, 7333–7343.
- (28) Chen, J. C.; Krucinski, J.; Miercke, L. J.; Finer-Moore, J. S.; Tang, A. H.; Leavitt, A. D.; Stroud, R. M. Crystal structure of the HIV-1 integrase catalytic core and C-terminal domains: a model for viral DNA binding. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 8233–8238.
- (29) Dyda, F.; Hickman, A. B.; Jenkins, T. M.; Engelman, A.; Craigie, R.; Davies, D. R. Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science* **1994**, *266*, 1981–1986.
- (30) Bujacz, G.; Alexandratos, J.; Qing, Z. L.; Clement-Mella, C.; Wlodawer, A. The catalytic domain of human immunodeficiency virus integrase: ordered active site in the F185H mutant. *FEBS Lett.* **1996**, *398*, 175–178.
- (31) Greenwald, J.; Le, V.; Butler, S. L.; Bushman, F. D.; Choe, S. The mobility of an HIV-1 integrase active site loop is correlated with catalytic activity. *Biochemistry* **1999**, *38*, 8892–8898.
- (32) Wielens, J.; Headey, S. J.; Jeevarajah, D.; Rhodes, D. I.; Deadman, J.; Chalmers, D. K.; Scanlon, M. J.; Parker, M. W. Crystal structure of the HIV-1 integrase core domain in complex with sucrose reveals details of an allosteric inhibitory binding site. *FEBS Lett.* **2010**, *584*, 1455–1462.
- (33) Goldgur, Y.; Craigie, R.; Cohen, G. H.; Fujiwara, T.; Yoshinaga, T.; Fujishita, T.; Sugimoto, H.; Endo, T.; Murai, H.; Davies, D. R. 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.* **1999**, *96*, 13040–13043.
- (34) Goldgur, Y.; Dyda, F.; Hickman, A. B.; Jenkins, T. M.; Craigie, R.; Davies, D. R. Three new structures of the core domain of HIV-1 integrase: an active site that binds magnesium. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 9150–9154.
- (35) Maignan, S.; Guilloteau, J. P.; Zhou-Liu, Q.; Clement-Mella, C.; Mikol, V. Crystal structures of the catalytic domain of HIV-1 integrase free and complexed with its metal cofactor: high level of similarity of the active site with other viral integrases. *J. Mol. Biol.* **1998**, *282*, 359–368.



- (36) Hare, S.; Gupta, S. S.; Valkov, E.; Engelman, A.; Cherepanov, P. Retroviral intasome assembly and inhibition of DNA strand transfer. *Nature* **2010**, *464*, 232–236.
- (37) Tang, J.; Maddali, K.; Pommier, Y.; Sham, Y. Y.; Wang, Z. Scaffold rearrangement of dihydroxypyrimidine inhibitors of HIV integrase: Docking model revisited. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 3275–3279.
- (38) Krishnan, L.; Li, X.; Naraharisetty, H. L.; Hare, S.; Cherepanov, P.; Engelman, A. Structure-based modeling of the functional HIV-1 intasome and its inhibition. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 15910–15915.
- (39) Sali, A.; Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779–815.
- (40) Sali, A.; Potterton, L.; Yuan, F.; van Vlijmen, H.; Karplus, M. Evaluation of comparative protein modeling by MODELLER. *Proteins* **1995**, *23*, 318–326.
- (41) Notredame, C.; Higgins, D. G.; Heringa, J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **2000**, *302*, 205–217.
- (42) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **1993**, *26*, 283–291.
- (43) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **2004**, *47*, 1750–1759.
- (44) Kuhn, B.; Gerber, P.; Schulz-Gasch, T.; Stahl, M. Validation and use of the MM-PBSA approach for drug discovery. *J. Med. Chem.* **2005**, *48*, 4040–4048.
- (45) Huang, N.; Kalyanaraman, C.; Irwin, J. J.; Jacobson, M. P. Physics-based scoring of protein-ligand complexes: enrichment of known inhibitors in large-scale virtual screening. *J. Chem. Inf. Model.* **2006**, *46*, 243–253.
- (46) Cramer, R. D. 3rd; Patterson, D. E.; Bunce, J. D. Recent advances in comparative molecular field analysis (CoMFA). *Prog. Clin. Biol. Res.* **1989**, *291*, 161–165.
- (47) Klebe, G.; Abraham, U.; Mietzner, T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. *J. Med. Chem.* **1994**, *37*, 4130–4146.
- (48) Clark, M.; Cramer, R. D. III. The Probability of Chance Correlation Using Partial Least Squares (PLS). *Quant. Struct.-Act. Relat.* **1993**, *12*, 137–145.
- (49) Engelman, A.; Craigie, R. Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro. *J. Virol.* **1992**, *66*, 6361–6369.
- (50) Lins, R. D.; Adesokan, A.; Soares, T. A.; Briggs, J. M. Investigations on human immunodeficiency virus type 1 integrase/DNA binding interactions via molecular dynamics and electrostatics calculations. *Pharmacol. Ther.* **2000**, *85*, 123–131.
- (51) Tounge, B. A.; Reynolds, C. H. Calculation of the binding affinity of beta-secretase inhibitors using the linear interaction energy method. *J. Med. Chem.* **2003**, *46*, 2074–2082.
- (52) Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J. Chem. Inf. Model.* **2011**, *51*, 69–82.
- (53) Petrocchi, A.; Koch, U.; Matassa, V. G.; Pacini, B.; Stillmock, K. A.; Summa, V. From dihydroxypyrimidine carboxylic acids to carboxamide HIV-1 integrase inhibitors: SAR around the amide moiety. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 350–353.
- (54) Summa, V.; Petrocchi, A.; Matassa, V. G.; Gardelli, C.; Muraglia, E.; Rowley, M.; Paz, O. G.; Laufer, R.; Monteagudo, E.; Pace, P. 4,5-dihydroxypyrimidine carboxamides and N-alkyl-5-hydroxypyrimidinone carboxamides are potent, selective HIV integrase inhibitors with good pharmacokinetic profiles in preclinical species. *J. Med. Chem.* **2006**, *49*, 6646–6649.
- (55) Pace, P.; Di Francesco, M. E.; Gardelli, C.; Harper, S.; Muraglia, E.; Nizi, E.; Orvieto, F.; Petrocchi, A.; Poma, M.; Rowley, M.; Scarpelli, R.; Laufer, R.; Gonzalez Paz, O.; Monteagudo, E.; Bonelli, F.; Hazuda, D.; Stillmock, K. A.; Summa, V. Dihydroxypyrimidine-4-carboxamides as novel potent and selective HIV integrase inhibitors. *J. Med. Chem.* **2007**, *50*, 2225–2239.
- (56) Nizi, E.; Orsale, M. V.; Crescenzi, B.; Pescatore, G.; Muraglia, E.; Alfieri, A.; Gardelli, C.; Spieser, S. A.; Summa, V. Dihydroxypyrimidine and N-methylpyrimidinone HIV-integrase inhibitors: improving cell based activity by the quaternarization of a chiral center. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 4617–4621.
- (57) Muraglia, E.; Kinzel, O.; Gardelli, C.; Crescenzi, B.; Donghi, M.; Ferrara, M.; Nizi, E.; Orvieto, F.; Pescatore, G.; Laufer, R.; Gonzalez-Paz, O.; Di Marco, A.; Fiore, F.; Monteagudo, E.; Fonsi, M.; Felock, P. J.; Rowley, M.; Summa, V. Design and synthesis of bicyclic pyrimidinones as potent and orally bioavailable HIV-1 integrase inhibitors. *J. Med. Chem.* **2008**, *51*, 861–874.
- (58) Di Francesco, M. E.; Pace, P.; Fiore, F.; Naimo, F.; Bonelli, F.; Rowley, M.; Summa, V. Development of 2-t butyl-N-methyl pyrimidinones as potent inhibitors of HIV integrase. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 2709–2713.
- (59) Pace, P.; Spieser, S. A.; Summa, V. 4-Hydroxy-5-pyrrolinone-3-carboxamide HIV-1 integrase inhibitors. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 3865–3869.
- (60) Donghi, M.; Kinzel, O. D.; Summa, V. 3-Hydroxy-4-oxo-4H-pyrido[1,2-a]pyrimidine-2-carboxylates—a new class of HIV-1 integrase inhibitors. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 1930–1934.
- (61) Petrocchi, A.; Jones, P.; Rowley, M.; Fiore, F.; Summa, V. N-(4-Fluorobenzyl)-3-hydroxy-9,9-dimethyl-4-oxo-6,7,8,9-tetrahydro-4H-pyrazino[1,2-a]pyrimidine-2-carboxamides a novel class of potent HIV-1 integrase inhibitors. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 4245–4249.
- (62) Ferrara, M.; Fiore, F.; Summa, V.; Gardelli, C. Development of 2-pyrrolidinyl-N-methyl pyrimidinones as potent and orally bioavailable HIV integrase inhibitors. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 5031–5034.
- (63) Zhuang, L.; Wai, J. S.; Embrey, M. W.; Fisher, T. E.; Egbertson, M. S.; Payne, L. S.; Guare, J. P. Jr.; Vacca, J. P.; Hazuda, D. J.; Felock, P. J.; Wolfe, A. L.; Stillmock, K. A.; Witmer, M. V.; Moyer, G.; Schleif, W. A.; Gabryelski, L. J.; Leonard, Y. M.; Lynch, J. J. Jr.; Michelson, S. R.; Young, S. D. Design and synthesis of 8-hydroxy-[1,6]-naphthyridines as novel inhibitors of HIV-1 integrase in vitro and in infected cells. *J. Med. Chem.* **2003**, *46*, 453–456.
- (64) Hazuda, D. J.; Felock, P. J.; Hastings, J. C.; Pramanik, B.; Wolfe, A. L. Differential divalent cation requirements uncouple the assembly and catalytic reactions of human immunodeficiency virus type 1 integrase. *J. Virol.* **1997**, *71*, 7005–7011.
- (65) Debyser, Z.; Cherepanov, P.; Pluymers, W.; De Clercq, E. Assays for the evaluation of HIV-1 integrase inhibitors. *Methods Mol. Biol.* **2001**, *160*, 139–155.
- (66) Yu, F.; Jones, G. S.; Hung, M.; Wagner, A. H.; MacArthur, H. L.; Liu, X.; Leavitt, S.; McDermott, M. J.; Tsiang, M. HIV-1 integrase preassembled on donor DNA is refractory to activity stimulation by LEDGF/p75. *Biochemistry* **2007**, *46*, 2899–908.