JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

Impact of Resistance Mutations on Inhibitor Binding to HIV-1 Integrase

Qi Chen,[†] John K. Buolamwini,[‡] Jeremy C. Smith,[§,∥] Aixiu Li,[⊥] Qin Xu,[†] Xiaolin Cheng,*[,§,∥] and Dongqing Wei*[,†]

[†]State Key Laboratory of Microbial Metabolism and College of Life Science and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China

[‡]Department of Pharmaceutical Sciences, College of Pharmacy, University of Tennessee Health Science Center, Memphis, Tennessee 38163, United States

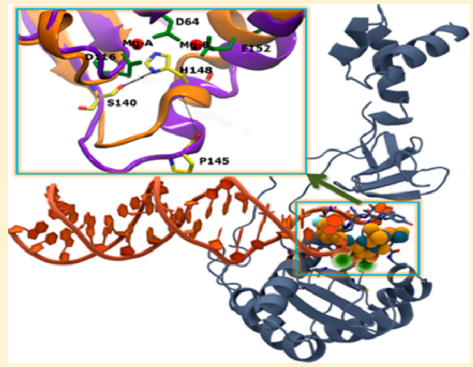
[§]UT/ORNL Center for Molecular Biophysics, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, United States

[∥]Department of Biochemistry and Cellular and Molecular Biology, University of Tennessee, Knoxville, Tennessee 37996, United States

[⊥]Drug Design Laboratory of the Basic Science Department, Logistics College of Chinese People's Armed Police Force, Tianjin 300162, China

Supporting Information

ABSTRACT: HIV-1 integrase (IN) is essential for HIV-1 replication, catalyzing two key reaction steps termed 3′ processing and strand transfer. Therefore, IN has become an important target for antiviral drug discovery. However, mutants have emerged, such as E92Q/N155H and G140S/Q148H, which confer resistance to raltegravir (RAL), the first IN strand transfer inhibitor (INSTI) approved by the FDA, and to the recently approved elvitegravir (EVG). To gain insights into the molecular mechanisms of ligand binding and drug resistance, we performed molecular dynamics (MD) simulations of homology models of the HIV-1 IN and four relevant mutants complexed with viral DNA and RAL. The results show that the structure and dynamics of the 140s' loop, comprising residues 140 to 149, are strongly influenced by the IN mutations. In the simulation of the G140S/Q148H double mutant, we observe spontaneous dissociation of RAL from the active site, followed by an intrahelical swing-back of the 3′-OH group of nucleotide A17, consistent with the experimental observation that the G140S/Q148H mutant exhibits the highest resistance to RAL compared to other IN mutants. An important hydrogen bond between residues 145 and 148 is present in the wild-type IN but not in the G140S/Q148H mutant, accounting for the structural and dynamical differences of the 140s' loop and ultimately impairing RAL binding in the double mutant. End-point free energy calculations that broadly capture the experimentally known RAL binding profiles elucidate the contributions of the 140s' loop to RAL binding free energies and suggest possible approaches to overcoming drug resistance.



INTRODUCTION

Human immunodeficiency virus type 1 (HIV-1) integrase (IN) is essential for HIV-1 replication, which facilitates the insertion of viral DNA into the genome of the host cell. HIV-1 IN is a 32 kDa protein, consisting of three structurally and functionally distinct domains, the N-terminal domain (NTD, residues 1−49), the catalytic core domain (CCD, residues 50−212), and the C-terminal domain (CTD, residues 213−288). The CCD domain contains a highly conserved catalytic triad motif, D64D116E152, which is required for the enzymatic activity.[1]

IN catalyzes two key chemical reactions, termed 3′ processing and strand transfer. During 3′ processing, IN removes two terminal nucleotides from the 3′ ends of both viral DNA strands to expose the invariant CA 3′-OH DNA ends. In the subsequent strand transfer step, IN catalyzes the integration of the viral DNA into the host cell chromatin using its exposed 3′-OHs to attack the phosphodiester backbones of the host DNA.[2]

Given its importance in HIV replication together with the fact that no human counterpart exists for the enzyme, IN has become a promising target for antiviral drug discovery. From the time IN was first considered as a potential drug target, more than a decade elapsed before raltegravir (RAL) was approved by the FDA (in 2007).[3,4] Later, in 2012, elvitegravir (EVG) was also approved by the FDA for HIV treatment.[5] Both RAL and EVG bind to the CCD of IN and specifically block the strand transfer step, and they thus belong to IN strand transfer inhibitors (INSTIs), the latest class of antiretroviral drugs for HIV treatment. Other INSTIs in clinical trials include

Received: September 15, 2013
Published: November 8, 2013

   dx.doi.org/10.1021/ci400537n | J. Chem. Inf. Model. 2013, 53, 3297−3307

dolutegravir (S/GSK1349572) and MK-2048 (Figure S1).[6−8] The common coplanar oxygen atoms of these IN inhibitors are thought to bind to the two catalytically essential Mg^{2+} ions that are also coordinated by the carboxylate groups of the DDE motif. The binding of INSTIs displaces the reactive 3′-OH of the terminal A17 away from the active site, which disrupts the strand transfer process.[9,10]

As for many previous anti-HIV drugs, drug-resistant mutants have emerged in patients for both RAL and EVG, such as the primary mutations Y143R/C, Q148H/R/K, and N155H as well as the secondary mutations E92Q/G, T97A, G140S/A, and G163R.[11−14] A brief summary of the effects of HIV-1 IN mutations on RAL sensitivity is given in Table 1.[4,15] In order to design novel INSTIs that are less prone to the drug resistance problem, it is necessary to obtain a thorough understanding of the molecular mechanisms for drug resistance. Earlier experimental efforts were directed to the characterization of biochemical activities and resistance profiles of several IN mutants for RAL and EVG.[15−17] It was shown that most primary INSTI resistance mutations occur in the catalytic domain of IN.

Table 1. Effects of HIV-1 IN Mutations on RAL Sensitivity[4,15]

mutations	fold change from WT RAL IC_{50}
N155H	10−30
N155H + E92Q	80−150
Q148H	18
Q148H + G140S	521
Q148H + G140A	>150
Q148R	34
Q148R + G140S	405
Q148R + G140A	>150
Y143R	<30
Y143R+97A	>150

Since the first structure of the DNA binding domain of HIV-1 IN, determined by multidimensional NMR spectroscopy in 1995,[18] numerous partial retroviral IN structures have become available,[19−21] and based on these partial HIV-1 IN structures or homology models, several molecular simulation studies have been reported.[22−25] For example, Perryman et al. developed a restrained dynamic model to analyze the flexibility and conformational dynamics of IN and predicted the binding modes of RAL in the wild-type (WT) and in the G140S/Q148H double mutant.[26] Also, Huang et al. conducted comparative molecular dynamics (MD) simulations of IN with several different diketoacid inhibitors (DKAs) to investigate their mechanisms of inhibition.[27] MD simulations have also been employed to study the dynamics of the loop comprising residues 140−149 (the 140s' loop), which moves like a gate, opening and closing over the catalytic site.[26,28,29] There are indications that the conformational flexibility of the 140s' loop may be important for the enzyme's activity,[30−32] as mutations such as G140A and G149A that decrease the loop flexibility also reduce the activity. However, the results from these previous MD studies were inconclusive because the binding positions of the inhibitors and/or viral DNA were either inaccurate or not included in the simulations.

The crystal structures of a full length prototype foamy virus (PFV) IN in complex with viral DNA and inhibitors RAL or EVG have recently been determined,[33,34] revealing hitherto unknown extensive protein−protein and protein−DNA interactions within a complete, functional HIV-1 intasome. INSTIs are known to preferentially bind to IN-DNA complexes over free INs. Therefore, these new IN structures should constitute an improved structural basis for understanding inhibitor binding and drug resistance in HIV-1 IN. Taking advantage of the sequence and structural similarity between HIV-1 and PFV, Krishnan et al. constructed a set of structural models for the HIV-1 intasome, which have shed light on many molecular details important for HIV-1 integration.[35] For instance, the models show a stacking interaction between the oxadiazole ring of RAL and the phenol ring of Y143, which explains why the mutation of Y143 confers primary resistance to RAL but not to EVG. However, for the other mutations that are not in contact with the inhibitors, such as Q148 and N155, the molecular mechanisms of HIV-1 resistance to RAL and EVG remain elusive. These mutations have been suggested to affect the positions of the key drug contacts, but structural rearrangements upon the mutations may not be adequately predicted by homology modeling. The N155H and Q148K/R/H mutations, carried by 68% of subjects treated with RAL, are both located within or near the 140s' loop, but the mechanistic details of how the loop structure and dynamics correlate with the drug resistance remain elusive.

Here, we perform MD simulations of homology models of the full-length IN and four relevant mutants, apo or complexed with viral DNA and RAL, with the aim of understanding how the IN mutations may affect the structure and dynamics of the 140s' loop and how the induced changes of the 140s' loop may in turn confer drug resistance to mutant HIV-1 INs from a dynamic perspective. The work provides a unified mechanistic view of the mutant resistance to RAL for the three principal genetic pathways, N155H, Q148H/R/K and Y143H/R/C, all directly or indirectly linked to the 140s' loop. We anticipate that the mechanistic insight will help guide future computational and experimental efforts in the development of novel inhibitors against drug-resistant HIV-1 IN mutants.

METHODS

Simulation Structures. The homology models of the full length HIV-1 IN together with viral DNA and RAL were generated using Modeler.[36] In addition to the WT, four different mutant HIV-1 IN models were built, including G140S and Q148H single mutants and G140S/Q148H and E92Q/N155H double mutants. The crystal structure of PFV IN (PDB code: 3L2T, including Zn^{2+}, Mg^{2+} and RAL) was used as the template for homology modeling. The HIV-1 IN sequence was obtained from the work of Hare et al.[34] An ensemble of 20 homology models was generated for each IN-DNA complex. Based on the comparison of several scoring parameters, including molpdf,[36] DOPE[37] and GA341 scores,[38,39] one best-scored structure was chosen as the starting structure for each MD simulation. From the Ramachandran plot, the percentage of residues in the allowed region are 99.2%, 98.9%, 97.8%, 98.4% and 99.5%, respectively, suggesting that the stereochemistry for all the five IN models is reasonable. The overall quality factors evaluated using Errat[40] for the five systems are 50.0, 54.4, 47.5, 54.5, and 50.5, respectively, indicating that no significant atomic clash is present in any of these models.

MD Simulation. The LEaP module in Amber11 was used to add missing hydrogen atoms.[41] The protonation states of protein side chains were determined at pH 7.0 using

PDB2PQR[42] and then verified manually. The Amber FF99SB force field[43] was used for the protein and DNA. The force field parameters for RAL were developed using the Antechamber package, with the partial charges obtained using the RESP method,[44,45] which are provided in Table S1. The five IN-DNA complex systems were each solvated in a TIP3P water box,[46] with a minimum padding of 15.0 Å on each side of the rectangular box. Each of the final simulation systems contained ~140 000 atoms.

NAMD 2.8 was used for running all the simulations.[47] Each system was first minimized for 20 000 steps and then equilibrated for 2 ns, followed by ~150 ns production simulation. In all the simulations, the van der Waals (vdW) interaction was smoothly turned off between 10.5 and 12 Å using a switching function. Long-range electrostatic interactions were treated using the particle-mesh Ewald (PME) method[48] with a 1.0 Å grid spacing. The r-RESPA multiple time-step method was employed, with time steps of 2 fs for bonded, 2 fs for short-range nonbonded, and 4 fs for long-range electrostatic interactions.[49] The bonds between hydrogen and the heavy atoms were constrained using the SHAKE algorithm.[50] In previous MD studies of HIV-1 IN,[26] none of the restraint-free models displayed proper monodentate interactions of DDE with the two Mg^{2+} ions. Therefore, In the present simulations, the distances between Zn^{2+} and its coordination residues (H12, C40, H16 and C43) as well as Mg^{2+} and their coordination residues (D64, D116, and E152) were harmonically restrained to their equilibrium values (1.5−3.5 Å) with a force constant of 10 kcal/mol/Å^2.

We first performed five MD simulations, one for each of the WT and four mutant HIV-1 IN systems, in an NPT ensemble. Langevin dynamics was used to maintain a constant temperature at 300 K, while the Nosé−Hoover Langevin piston algorithm[51,52] was used to maintain a constant pressure at 1 bar. To improve statistics, we conducted an additional set of five ~20 ns simulations with the same MD control parameters, except at a higher temperature (350 K) to enhance the conformational sampling. Results presented below are from the simulations at 300 K unless indicated otherwise.

The molecular mechanics Poisson−Boltzmann/surface area (MM-PB/SA) method was used to compute the binding free energy of the ligand in the WT and different mutant IN-DNA complexes.[53,54] The total binding energy ΔG_{bind} was defined as $\Delta G_{bind} = G_{complex} - G_{receptor} - G_{ligand}$. Each free energy term consisted of the gas phase MM energy (ΔE_{gas}), the solvation free energy (ΔG_{sol}), and the vibrational entropy contributions ($T\Delta S$). ΔG_{sol} was estimated from the PB theory and solvent accessible surface area (SASA) calculations which yielded ΔG_{polar} and $\Delta G_{nonpolar}$. A surface tension coefficient (γ) of 0.0072 kcal/(mol·Å^2) was used to calculate the nonpolar solvation free energy contribution. Due to its prohibitive computational cost and the inherent difficulty in determining accurate absolute entropy for large protein−DNA-ligand systems, the vibrational entropy contribution was not included in our calculation. E_{gas} and ΔG_{sol} were computed for 1250 snapshots extracted evenly from the last 50 ns of the MD trajectories with a sampling interval of 40 ps. To improve the estimation of ΔG_{bind}, the "receptor" contained only the IN core domain, two active site Mg^{2+} ions and three terminal DNA base pairs near RAL.

Trajectory analyses, including the root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), structural clustering, and principal component analysis (PCA), were performed using the Gromacs 4.5.5 program.[55] Structural clustering was done using g_cluster with the gromos algorithm.[56] Cross-correlation analysis was used to quantify correlated motions between any pair of residues in the simulations. The cross-correlation coefficient C_{ij}, between atoms i and j, was computed as follows:

$$C_{ij} = \langle \Delta r_i \times \Delta r_j \rangle / (\langle \Delta r_i \times \Delta r_i \rangle \langle \Delta r_j \times \Delta r_j \rangle)^{1/2}$$

where Δr_i and Δr_j are the displacement vectors for atoms i and j, respectively. The angle brackets denote an ensemble average over the trajectory. The value of C_{ij} ranges from −1 to 1, with the positive (correlated) residue pair moving in the same direction and the negative (anticorrelated) pair moving in the opposite direction. The correlation matrices of Cα atoms were calculated using the R package Bio3D.[57]

RESULTS AND DISCUSSION

Overall Dynamics of the CCD Domain. We first calculated the RMSDs of the CCD domain after aligning only the backbone atoms of the CCD domain during the simulations. As shown in Figure S2, the RMSDs of the CCD domain for all the simulated systems show no major differences and reach a plateau after about 25 ns with a final RMSD value of ~3 Å, which is comparable to those observed in previous studies.[58,59] This indicates that all the simulations have reached equilibrium and can be used for further analysis. The RMSDs of the 140s' loop are displayed in Figure 1, which evidently exhibit much greater variation among the five simulated systems. In particular, the 140s' loop of the Q148H and G140S/Q148H mutants deviates much more from the initial structures than that of the WT, the G140S and E92Q/N155H mutants. We also calculated the RMSFs of the Cα atoms for all the five systems, which are displayed in Figure S3. As expected, most secondary structure elements, such as the α helices and β strands, exhibit small fluctuations. However, in both the WT and mutant INs, several regions exhibit remarkable flexibility, including residues 140−149 (the 140s' loop), 166−171, and 186−194. As the flexibility of the 140s' loop may play a critical role in integration function and drug resistance, the structure and dynamics of this loop will be further analyzed below.

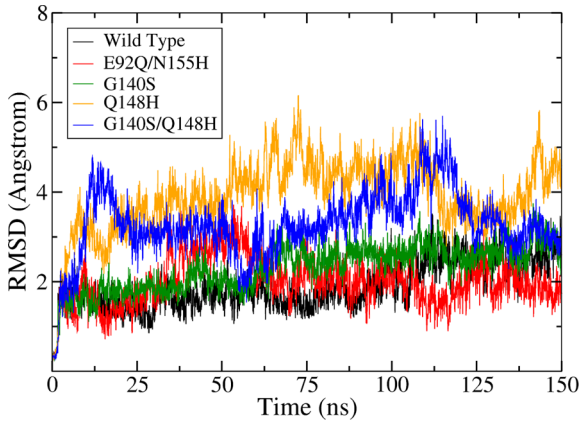


Figure 1. RMSD of the 140s' loop as a function of time during the simulations of the HIV-1 IN-DNA complex and four mutant models. The RMSDs were calculated for the 140s' loop backbone atoms with the entire CCD domains superposed on the corresponding starting structures.

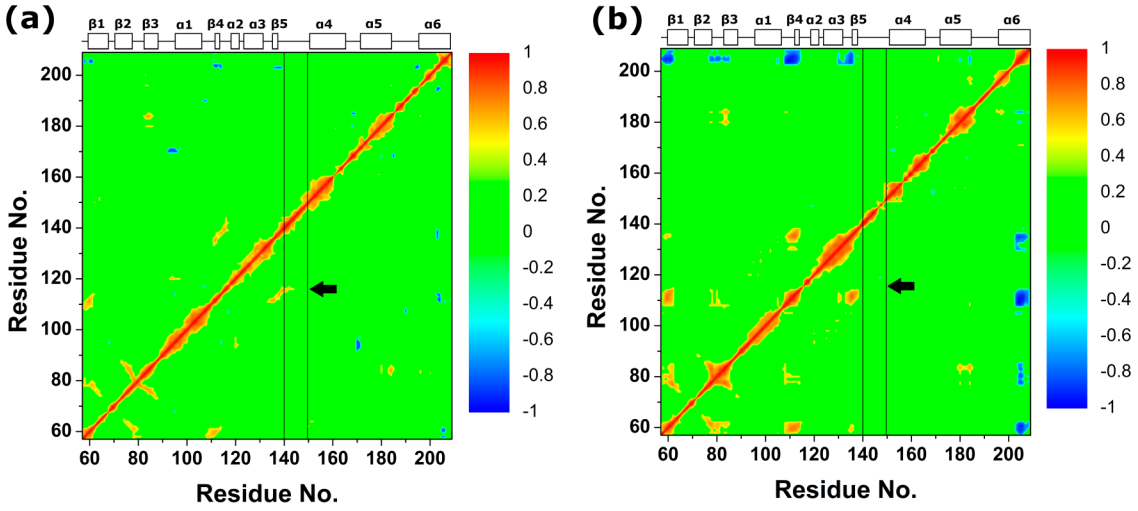


Figure 2. Correlated fluctuations of the Cα atoms in the CCD domain, calculated from the MD simulations of the WT IN (a) and the G140S/Q148H double mutant (b). The color bars on the right indicate the extent of the correlation. Residue pairs with a high level of correlated motion are shown in red and orange, while anticorrelated motions are represented by the blue/cyan regions. Green indicates weak correlation. An absolute cutoff value of 0.5 (values in [−0.5, 0.5] were not shown) was chosen to filter weak correlations.

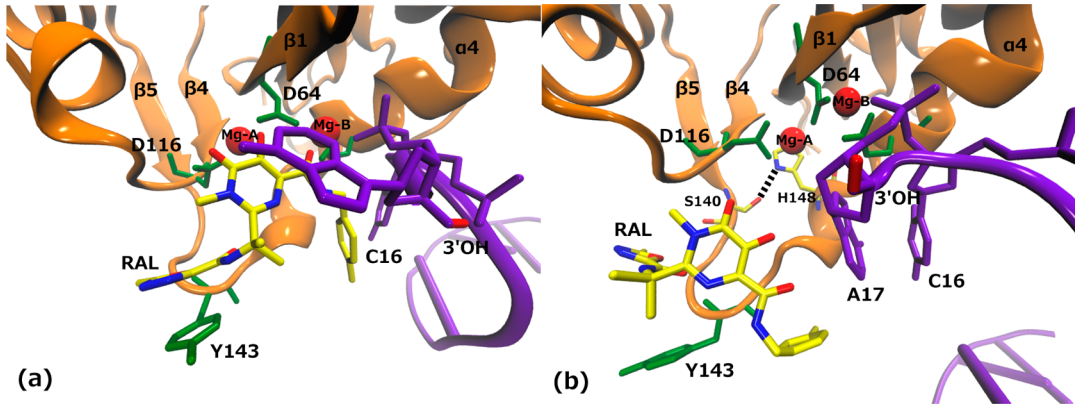


Figure 3. Representative snapshots of binding modes of RAL (taken from the final 5 ns of the respective MD simulations) in the active sites of the WT (a) and the G140S/Q148H double mutant (b) HIV-1 INs. The protein backbone is shown in orange cartoon representation, with invariant retroviral DNA (purple), RAL and side chains of D64, D116, E152, and Y143 (green) as sticks; Mg^{2+} ions (A and B) and 3′-OH of the DNA end are also indicated.

Correlated Fluctuations in the CCD Domain. A cross-correlation analysis can be used to reveal how the atomic fluctuations or movements of different parts of the system are coupled. Similar studies have been reported for rationalizing drug resistance in HIV-1 protease[60,61] and the dynamics of the 140s' loop in HIV-1 IN.[61] Here, we calculated the correlated fluctuations for all the Cα atoms of the CCD domain using the entire 150 ns MD trajectories. Figure 2 shows the cross-correlation maps for the WT and the G140S/Q148H mutant, while the correlation maps for the other three mutants are provided in Figure S4. Corresponding quantitative correlation coefficients between the 140s' loop and the DDE motif are also listed in Table S2. The correlation involving the 140s' loop is much weaker in the two double mutants than in the WT IN. The strong correlation in the WT between the 140s' loop and the active site residue D116 is diminished in the G140S/Q148H double mutant. Instead, an analogous correlation appears between the active site regions (residues 112−116, the β4 sheet) and the β2 (residues 71−78) and β5 (residues 136−139) sheets. A similar situation was found in the E92Q/N155H double mutant; no strong correlation was found between the 140s' loop and the DDE motif (Figure S4). However, the correlation between the 140s' loop and DDE motif is still present in both the G140S and Q148H single mutants. These results coincide with the fact that the two double mutants are the two most resistant of the four studied here.

We further quantified the correlation between the active site residues (D64, D116, and E152) and the 140s' loop by counting the number of residues involved using cross-correlation coefficient C_{ij} of 0.5 as the threshold value. As listed in Table S3, the number of correlations for the 140s' loop is much smaller in the E92Q/N155H and G140S/Q148H double mutants than in the WT, being 80 for the WT, and 69 and 62 in E92Q/N155H and G140S/Q148H, respectively. Error bars listed in Tables S2 and S3 were calculated with block averaging and bootstrapping methods, and most are within 10% of the corresponding average values, suggesting that the sampling is converged and the statistics is significant. Taken together, the above results show that the 140s' loop is less extensively engaged in correlated motion in the two double mutants. In particular, significantly weakened correlation between the 140s' loop and the DDE motif is seen upon the

dx.doi.org/10.1021/ci400537n | J. Chem. Inf. Model. 2013, 53, 3297−3307

mutation of both 140S and 148H, and this in terms of not only the intensity of the correlation but also the number of the correlated residues.

The 140s' loop makes no direct contact with the DDE motif. Rather, the interactions are mediated through the binding of the inhibitors: the common coplanar oxygen atoms on one side of the inhibitors chelate with the divalent metal ions and then interact with the DDE motif, while the functional groups on the opposite side of the inhibitors engage the 140s' loop. The attenuated correlation between the 140s' loop and the DDE motif thus indicates that the G140S/Q148H mutation may have altered the binding mode of RAL in the active site, leading to resistance to RAL of the G140S/Q148H mutant.

Binding Modes of RAL in INs. The crystal structure of the PFV IN in complex with RAL shows direct contact between the proximal oxygen atoms of the ligand with the two Mg^{2+} ions. The halobenzyl group of the ligand stacks against the C16 base at the 3′-end of the viral DNA, causing the displacement of the reactive 3′-OH of the terminal A17 nucleotide away from the active site.[62] These two interactions are considered to be the main structural reason for the INSTI binding and the mechanism of inhibition.[31,63] Figure 3 shows two representative snapshots obtained from the simulations of the WT and the G140S/Q148H mutant HIV-1 INs. In the WT, the three coplanar oxygen atoms of RAL are found to engage the two catalytic Mg^{2+} ions that are also coordinated by the DDE motif. The fluorobenzyl and oxadiazole rings of RAL form $\pi-\pi$ stacking interactions with the C16 of the reactive DNA strand and Y143, respectively. One of the Mg^{2+} ions (Mg-B) near the DNA also interacts with the phosphate group of the DNA's 3′-end A17. These observed interactions can potentially disrupt the 3′ processing and subsequent host DNA binding,[63,64] in good correspondence with the PFV IN crystal structures.

According to previous experiments (Table 1), the Q148H mutation causes a modest level of drug resistance to RAL (a 10−20 fold change in the susceptibility of mutant IN relative to the WT IN), and the G140S mutation induces less resistance (2−10 fold change) than the Q148H mutation.[16] However, when the G140S mutation is combined with the Q148 mutation, the double mutant shows the strongest resistance (>100 fold change) among all known HIV-1 IN mutants.[11,12] Also, it has been demonstrated that the presence of RAL IN restores its catalytic power only upon the G140S/Q148H double mutation.[65−67] During the simulation of the G140S/Q148H mutant, RAL moves away from the binding pocket and the viral DNA (Figure 3b), no longer chelating with the two catalytic Mg^{2+} ions, and the stacking interaction between RAL and C16 is replaced by a π-cation interaction between the ligand and Arg231. Notably, following the departure of RAL, the displaced A17 in the 3′-end of the viral DNA swings back to interact with C16 and the two Mg^{2+} ions, akin to what is present in the crystal structure of the apo IN-DNA complex.[62] These results suggest a potential restoration of the catalytic ability of HIV-1 IN in the G140S/Q148H double mutant, thus providing an explanation for the resistance mechanism of the double mutant to RAL.

To gain insight into the mobility of RAL in the active sites of the WT and four mutant HIV-1 IN-DNA complexes, we plot in Figure 4 the distributions of the center of mass of RAL during the five simulations. A substantial displacement of RAL is observed only in the G140S/Q148H double mutant. Modest displacements are observed in Q148H and E92Q/N155H, while the WT and G140S mutant HIV-1 IN show very small

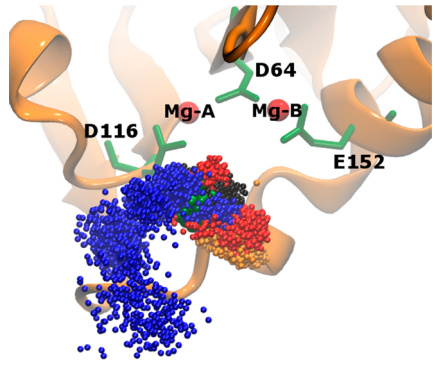


Figure 4. Distributions of the center of mass of RAL during the MD simulations of the WT (black spheres), E92Q/N155H (red spheres), G140S (green spheres), Q148H (orange spheres), and G140S/Q148H (blue spheres) IN-DNA complexes. The protein backbone is shown in orange cartoon representation, with side chains of D64, D116, and E152 as green sticks and two Mg^{2+} ions as red balls.

deviations, indicating that RAL binds very tightly to the latter two models. As shown in Figure S5, the binding mode of RAL in G140S is quite similar to that in the WT with an RMSD of 1.85 Å for the two ligand binding poses, which explains why the G140S mutation induces only insignificant drug resistance. In Q148H, the phosphate group of the 3′-end A17 still makes contact with the Mg-B ion. Yet, the positions of the interacting RAL, DNA and both Mg^{2+} ions have all shifted by ∼2 Å due to a new hydrogen bond formed between the pyrimidine ring of RAL and the imidazole moiety of H148. This structural rearrangement is consistent with the RMSD of the 140s' loop during the simulation (Figure 1). The movement of RAL away from the Mg^{2+} ions could conceivably result in a low binding affinity in the Q148H mutant.

In the E92Q/N155H double mutant, the phosphate group of the 3′-end adenosine moves away from Mg-B and establishes a new interaction with the imidazole moiety of H155 (Figure 5a). Accordingly, the distance between the phosphate group and Mg-B increases to ∼5 Å as compared to ∼3 Å in the other systems (Figure 5b). The ligand-DNA stacking interaction remains in the E92Q/N155H double mutant, but the positions of the stacking pair are both shifted because the DNA 3′-end has moved. As a result of these structural changes, RAL moves away from the two Mg^{2+} ions, potentially impairing the RAL binding in the E92Q/N155H mutant. The resistance mechanism proposed for E92Q/N155H does not directly involve the 140s' loop, consistent with the fact that its loop RMSD is comparable to the WT during the simulation (Figure 1). Furthermore, our simulation of the apo N155H mutant IN-DNA complex revealed a similar DNA 3′ end structural shift from Mg-B to H155 after a few ns of the MD simulation. N155H is one of the three primary mutations that confer drug resistance. Previous analysis of the PFV intasome crystal structure suggested that N155 might induce drug resistance via displacing the DNA backbone and the catalytic residues' carboxylates.[64] Our simulations thus confirmed this hypothesis and revealed a rather dramatic structural rearrangement of the DNA 3′ end upon this mutation, which eventually leads to a different and less optimal binding mode of RAL in the N155 mutant.

MD simulation has been a powerful computational approach to study the structure and dynamics of biomolecules. As computer power continues to increase, so does the length of an

dx.doi.org/10.1021/ci400537n | J. Chem. Inf. Model. 2013, 53, 3297−3307

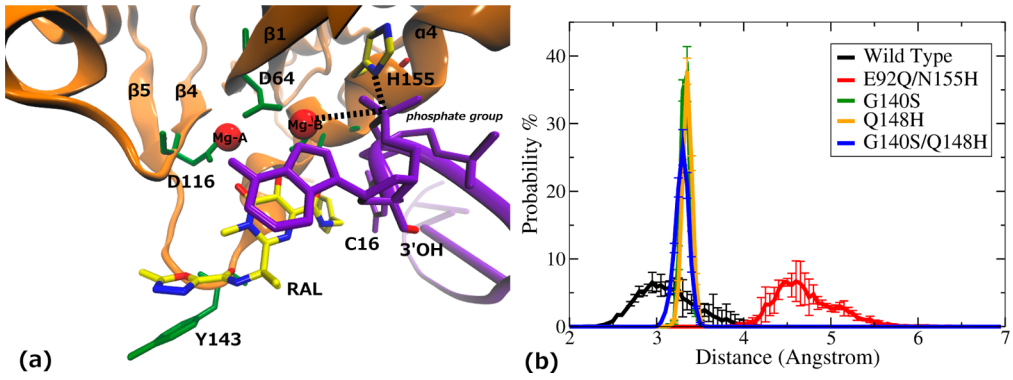


Figure 5. Binding mode of RAL in the active site of the E92Q/N155H double mutant taken from the final 5 ns of the MD simulations (a). Probability distributions of the distance between Mg-B and the DNA phosphate group in the five IN-DNA complexes (b).

MD simulation. However, drawing any major conclusion based on a single MD trajectory remains problematic, largely due to limited sampling in any single MD simulation which could be biased by the initial conditions. To alleviate this problem, we have run a second set of simulations for the five HIV-1 IN-DNA complexes, each for a duration of ∼20 ns but at an elevated temperature of 350 K to enhance the sampling. We find most of the binding mode results observed in the 300 K simulations are also present in the 350 K simulations. At 350 K, as shown in Figure S6, RAL is found to depart furthest from the binding pockets of G140S/Q148H and E92Q/N155H, indicating RAL binds more weakly to these two double mutants than the other three IN-DNA complexes.

Structural Clustering of the 140s' Loop. The Ω-shaped hairpin 140s' loop is situated in the cleft between the CCD and CTD domains of HIV-1 IN, constituting a boundary of the active site. The loop also makes extensive contacts with the viral as well as the host DNAs and has thus been thought to be essential for IN function. A number of studies have been conducted to investigate the role of the 140s' loop in ligand binding.[30−32] In particular, two primary drug resistant mutations are located in this loop, i.e., Y143R/C and Q148H/R/K.[66] Thus, interest in understanding how the 140s' loop may be linked to drug resistance has intensified recently.[68,69] Using reversible digitally filtered MD simulations, Williams and Essex showed that the 140s' loop was more flexible in the G140A/G149A mutant than in the WT HIV-1 IN.[61] However, Perryman et al. concluded that the 140s' loop had less conformational variation in G140S/Q148H than in the WT.[26] We note that the 140s' loop is highly conserved in the PFV and HIV-1 INs, with 6 out of 10 residues being identical. Therefore, using an improved full-length homology model of the HIV-1 IN in complex with RAL and the viral DNA, our analysis should lead to more compelling results on how different mutations may affect the behavior of the 140s' loop as well as its implication for drug resistance.

RMSD-based clustering was first performed for all the 140s' loop structures sampled in the MD simulations (a total of 37 531 snapshots for each of the five HIV-IN systems). Using an RMSD cutoff of 1 Å, the numbers of clusters are 21 for the WT and 120 for the G140S/Q148H double mutant IN, indicating considerably more structural variation in the double mutant. For the other mutants, the numbers of clusters are 21, 53, and 50 for the G140S, Q148H, and E92Q/N155H mutants, respectively. The same clustering procedure was also performed for the 140s' loop structures obtained from the 350 K simulations, leading to numbers consistent with those at 300 K: 25 for the WT, 87 for G140S/Q148H, 29 for G140S, 56 for Q148H, and 58 for E92Q/N155H. The computed number of clusters follows the same trend as the resistance profiles of the four mutants compared to the WT. It has been suggested that the conformation of the loop might depend on the presence of substrates or inhibitors.[64] This may partially explain why the G140S/Q148H mutant exhibits the largest variation, as in this mutant RAL has left the active site within the first 20 ns of the simulation. For the E92Q/N155H mutant, none of the residues in the 140s' loop are mutated. We thus speculate that the increased variation of the loop in E92Q/N155H is due to the shifted and weakened ligand binding.

We further performed PCA of the 140s' loop Cα atom fluctuations during the MD simulations. PCA is a technique that can transform a number of (possibly) correlated variables into a (smaller) number of uncorrelated variables, called principal components (PCs), while retaining those characteristics of the data set that contribute most to its variance. As PCA reduces the dimensionality of the trajectory data, it is able to highlight important dynamical features, with the first few PC modes describing the overall dynamics of the system occurring within the simulation time scale. Figure 6 displays the MD trajectories projected onto the first and the second PC modes. Consistent with the above analysis, the 140s' loop in G140S/Q148H is more widely distributed than in the other four systems. Q148H also shows relatively large variation, especially

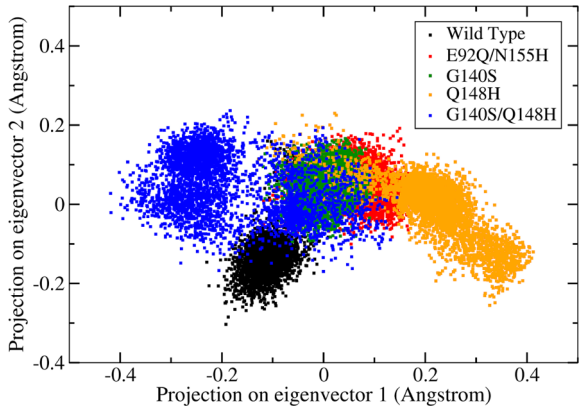


Figure 6. Projection of MD simulations of the five IN-DNA complexes onto the first and the second PC modes from the PCA of the 140s' loop Cα atom fluctuations.

dx.doi.org/10.1021/ci400537n | J. Chem. Inf. Model. 2013, 53, 3297−3307

along the first PC mode, while E92Q/N155H exhibits only modest variation along the two PC modes.

Hydrogen-Bond Network in the 140s' Loop. To understand why and how the G140S/Q148H double mutation causes such a significant conformational change of the 140s' loop, we analyzed detailed hydrogen-bond interactions surrounding the RAL binding site, especially those involving the 140s' loop. Figure 7 shows the probability distributions of the distances between residues 145 and 148 (Figure 7b) and between residues 140 and 148 (Figure 7c) in the five HIV-1 IN complexes. A hydrogen bond is present between Pro145 and Gln148 during the entire simulation times in all the five simulations except for the G140S/Q148H double mutant. In G140S/Q148H, this hydrogen bond breaks within the first few ns of the simulation and never reforms, and instead a new hydrogen bond forms between H148 and S140. As a result of this hydrogen bonding switch, the 140s' loop assumes a more open structure over the ligand binding site, as displayed in Figure 7a. Moreover, the unlocking of the P145-Q148 interaction, together with the rearrangement of the hydrogen bond involving S140, makes the 140s' loop more dynamic than that in the WT. This hydrogen-bond pattern shift and the associated structural and dynamical changes of the 140s' loop still occur in the simulation at 350 K (Figure S7) and that of the apo G140S/Q148H double mutant-DNA complex. Hence, we conclude that the observed structural and dynamic changes of the 140s' loop in the G140S/Q148H mutant can be attributed to the loss of an important hydrogen-bond interaction within the loop, which ultimately leads to much weakened binding of RAL in the mutant enzyme. These results confirm previous notions that the G140S/Q148H double mutation changes the hydrogen-bond network across the 140s' loop that might have stabilized the loop[70] and thus could play a role in the evolution of resistance to RAL and related INSTIs.[64]

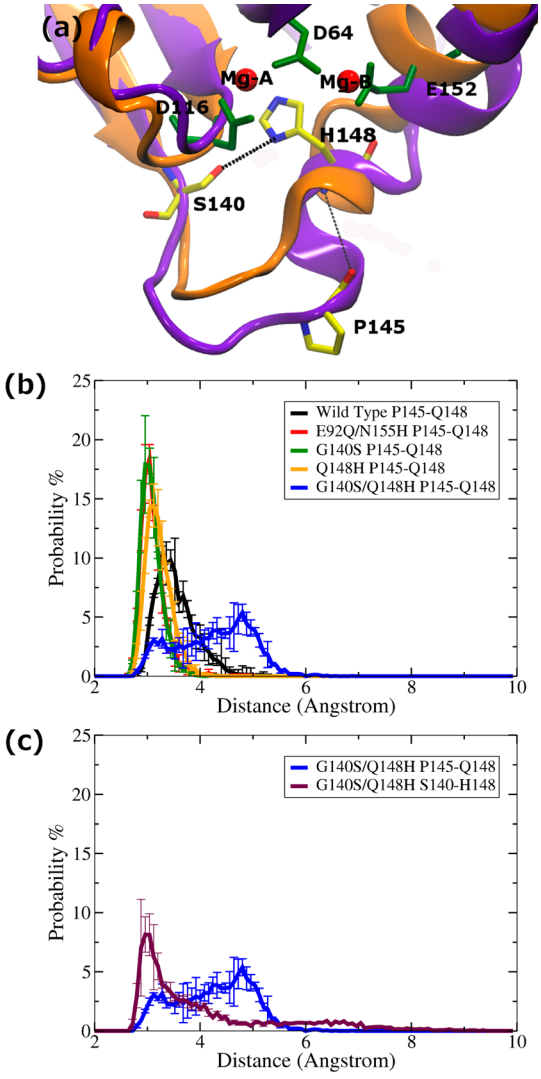


Figure 7. (a) Hydrogen-bond pattern shift associated with the conformational change in the 140s' loop in the G140S/Q148H double mutant (purple) in comparison with that in the WT (orange) IN. Snapshots are taken from the final 5 ns of the respective MD simulations. The protein backbone is shown as cartoons, with side chains of S140, H148, P145, D64, D116, and E152 as sticks; Mg^{2+} ions (A and B) are also indicated. Probability distributions of the distances between residues 145 and 148 (b) and between residues 140 and 148 (c) in the five IN-DNA complexes.

MM-PB/SA Binding Energy Calculation. To obtain a more quantitative understanding of the binding of RAL to different HIV-1 INs, the end-point free energy (MM-PB/SA) calculations were performed. Table 2 lists the MM-PB/SA binding energies of RAL in the five IN-DNA complexes. In the WT, the computed MM-PB/SA binding energy is -16.1 ± 6.8 kcal/mol, which is off by 5−6 kcal/mol from the −10.5 kcal/mol obtained with fluorescence anisotropy measurement[71] and −10.6 kcal/mol obtained from previous MD simulation.[72] The van der Waals interaction (ΔE_{vdW}) contributes more to the total binding energy (ΔG_{bind}) than the electrostatic term ($\Delta G_{PB,elec}$), consistent with RAL making extensive contacts with both the protein and DNA, whereas the Columbic interaction (ΔE_{elec}) is greatly offset by the solvent screening effect (ΔG_{polar}).

Overall, the computed binding free energy trends compare well to the drug resistance profiles of the various IN mutants. In particular, compared with those for the WT and the two single mutants, the binding free energies of -6.3 ± 7.7 kcal/mol for E92Q/N155H and -8.2 ± 2.6 kcal/mol for G140S/Q148H are significantly weaker, confirming the hypothesis that the two double mutations may trigger conformational changes within the catalytic pocket that result in a weakening of the binding of RAL to IN. The binding energy for Q148 lies between those for the WT and the double mutants, whereas the mutation of G140/S has little effect on the binding energy.

Measurements of the inhibitor dissociation rates from various mutant IN-DNA complexes have shown a striking correlation between dissociation rate and level of resistance,[73] implying that the RAL binding is not a gated process. That is, a more dynamic 140s' loop will likely decrease the RAL binding (i.e., increase k_{off}) rather than increase it (i.e., increase k_{on}). Overall, our findings suggest that the resistance to RAL within the Q148H, E92Q/N155H and G140S/Q148H mutants arises from the structural changes of the 140s' loop that weaken the RAL binding, while the dynamical (flexibility) changes of the 140s' loop that seemingly correlate with the drug resistance profiles are actually a consequence of the altered ligand binding in the three mutant INs studied here.

The application of the MM-PB/SA approach also allowed us to estimate individual energy components that contribute to ligand binding. To do this, we analyzed the pair interaction energies between RAL and (1) the 140s' loop, (2) the binding pocket residues with the two divalent Mg^{2+} but excluding the

dx.doi.org/10.1021/ci400537n | J. Chem. Inf. Model. 2013, 53, 3297−3307

Table 2. Binding Energies of RAL to Various IN-DNA Complexes Computed Using the MM-PB/SA Method[a]

energy (kcal/mol)	WT	E92Q/N155H	G140S	Q148H	G140S/Q148H
ΔE_{elec}	−53.1 ± 8.4	−22.9 ± 3.6	−59.3 ± 8.1	−23.1 ± 4.1	−7.7 ± 4.1
ΔE_{vdW}	−33.4 ± 3.5	−53.0 ± 2.2	−31.5 ± 4.7	−54.3 ± 2.4	−14.7 ± 2.7
ΔG_{polar}	75.6 ± 11.6	75.7 ± 8.9	79.1 ± 7.9	70.1 ± 7.4	17.0 ± 3.4
$\Delta G_{nonpolar}$	−5.1 ± 0.3	−6.1 ± 0.1	−5.0 ± 0.3	−6.3 ± 0.2	−2.8 ± 0.3
$\Delta G_{PB,elec}$	22.5 ± 8.3	52.8 ± 8.4	19.8 ± 7.6	47.1 ± 6.7	9.3 ± 2.5
ΔG_{bind}	−16.1 ± 6.8	−6.3 ± 7.7	−16.7 ± 6.3	−13.5 ± 6.9	−8.2 ± 2.6

[a] ΔE_{elec} and ΔE_{vdW} are the molecular mechanics electrostatic and van der Waals contributions, respectively. ΔG_{polar} is the PB solvation free energy, whereas $\Delta G_{nonpolar}$ is the nonpolar component of the solvation free energy estimated by surface area. $\Delta G_{PB,elec}$ is the electrostatic component of the binding energy ΔG_{bind}.

140s' loop, and (3) the three terminal DNA base pairs. Pair interaction analysis was performed using the final 50 ns of the MD trajectories. The results listed in Table S4 indicate that the interactions with the binding pocket residues dominate the total pair interactions, although we recognize this dominance would be less pronounced when combined with the solvation effect. RAL makes more favorable interactions with the binding pocket residues in the WT than in the Q148H, E92Q/N155H, and G140S/Q148H mutants. In particular, the computed interaction energy is −77.4 kcal/mol in the WT compared to −13.9 kcal/mol in the G140S/Q148H mutant. The pair interaction energy analysis shows no obvious correlation between the interaction energies contributed by the 140s' loop and the total binding energies, indicating that the 140s' loop plays an indirect role in the RAL binding, which may be modulated by the structure and dynamics of the loop.

While the metal binding is one of the key interactions for RAL binding to INs, a quantitative assessment of this interaction remains a significant challenge with classical force field calculations. Several issues remain to be addressed, including partial charges, coordination geometry, polarization, and even charge-transfer effects of the Mg ion, which could adversely affect our comparison of computed interaction energies. For example, it has been shown that divalent Mg ions in MD simulations have led to considerable overestimation of the binding of their coordinating ligands.[74,75] Therefore, the interaction energies between RAL and the IN-DNA complexes presented above were not meant to provide quantitative predictions of the binding affinities. These values were most likely overestimated due to the inadequacy of the Mg^{2+} parameters as well as lack of entropy−enthalpy compensation. Nonetheless, our computed binding free energy trends seem to compare well to the drug resistance profiles of the various IN mutants, suggesting that the method could be of some value if used judiciously while bearing in mind the limitations of the approach.

Cross drug resistance has emerged in clinical trials for first-generation INSTIs, RAL and EVG, that share two common mutation pathways.[14] The discovery of next generation INSTIs, such as the recent development of dolutegravir (S/GSK1349572, DTG) and MK-2048,[76,77] is thus becoming an urgent need. DTG can tolerate the common mutations that lead to drug resistance to RAL and EVG.[8,78] Although the essential binding mode is similar for the first- and second-generation INSTIs,[62] an aromatic ring of RAL that stacks against Y143 (located in the 140s' loop) is not present in DTG and other second-generation INSTIs. The lack of this stacking interaction possibly makes the binding of second-generation INSTIs less dependent on the structure and dynamics of the 140s' loop, which provides a plausible explanation for their reduced sensitivity to those mutations that can affect the 140s' loop. As listed in Table S4, the 140s' loop contributes to RAL binding mainly through vdW interactions, which only account for a small fraction of the total interaction energy relative to that with the binding pocket residues. Therefore, this, together with the fact that the 140s' loop is very prone to mutation, suggests a possible new pathway toward overcoming the drug resistance problem: to reduce the dependence of the inhibitor binding on the 140s' loop while maximally exploring the interactions with other regions of the binding pocket, e.g., the β4-α2 region across the DNA binding site. This strategy could also be combined with the design of compounds with a higher binding affinity for the active site Mg^{2+} divalent metals.[72]

CONCLUSION

HIV-1 IN is an important target for developing antiretroviral drugs. However, the atomic structure of the full length HIV-1 IN-DNA complex is not yet available. Moreover, while two of the three main IN mutation pathways, Q148H/R/K and Y143H/R/C, that confer primary resistance to RAL are located in the same loop comprising residues 140−149, the structure of this flexible 140s' loop has not been completely resolved, and its dynamic behavior remains poorly characterized. In this study, we built and performed MD simulations of homology models for the WT and several RAL resistant mutant HIV-1 INs in complex with viral DNA and RAL, based on the recently determined crystal structure of the PFV intasome in its potentially functional state. Our MD simulations show that the structure and dynamics of the 140s' loop are strongly correlated with the IN mutations and that this could affect the binding of first-generation INSTIs and thus play an important role in drug resistance. Spontaneous departure of RAL from the binding site was observed to occur in G140S/Q148H, one of the mutants conferring highest drug resistance. We find that the breaking of the P145-Q148 hydrogen bond along with the formation of an H148-S140 hydrogen bond resulting from the G140S/Q148H double mutation may be responsible for the observed structural and dynamical changes of the 140s' loop, leading to a more 'open' structure of the loop and eventually impairing RAL binding to the double mutant IN. Following the departure of RAL from the active site, A17 at the 3′ end of viral DNA swings back to interact with the active site carboxylates and Mg^{2+} ions, such that the reactive 3′-OH is restored to a proper position for the 3′-processing and/or strand transfer reaction. Future antiviral development efforts against multidrug resistant HIV IN mutants may benefit from exploring the interactions of potential drugs with other regions of the catalytic pocket.

■ ASSOCIATED CONTENT

Supporting Information

Force field parameters for the ligand are provided in Table S1. The quantitative correlation results are given in Table S2 and Table S3. Pair interaction energies between the ligand and the IN-DNA complexes are detailed in Table S4. 2-D diagrams of the four HIV-IN inhibitors disused in this study are shown in Figure S1. The dynamic changes of the catalytic core domain during the simulations are shown in Figure S2 and Figure S3. Figure S4 shows the cross-correlation maps for the E92Q/N155H, G140S and Q148H mutants. Figure S5 shows the ligand binding modes in the G140S and Q148H mutants. The distributions of the ligand and the hydrogen bond pattern shift during the 350K simulations are displayed in Figure S6 and Figure S7, respectively. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: chengx@ornl.gov

*E-mail: dqwei@sjtu.edu.cn

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was partially supported by grants from the National Basic Research Program of China (973 Program, contract no. 2012CB721000), the National High-Tech R&D Program (863) Program contract no. 2012AA020307), and Shanghai Science and Technology Commission (contract no. 11JC1406400) awarded to D.Q.W. This work is also supported by grants from the National Natural Science Foundation of China (nos. 30472166 and 81241114) awarded to A.L. J.C.S. acknowledges support from the National Institute of Health.

■ REFERENCES

(1) Zheng, R.; Jenkins, T. M.; Craigie, R. Zinc folds the N-terminal domain of HIV-1 integrase, promotes multimerization, and enhances catalytic activity. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93* (24), 13659−13664.

(2) Pommier, Y.; Johnson, A. A.; Marchand, C. Integrase inhibitors to treat HIV/AIDS. *Nat. Rev. Drug Discovery* **2005**, *4* (3), 236−248.

(3) 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* (18), 5843−5855.

(4) Okeke, N. L.; Hicks, C. Role of raltegravir in the management of HIV-1 infection. *HIV/AIDS* 3, 81−92.

(5) Marchand, C. The elvitegravir Quad pill: the first once-daily dual-target anti-HIV tablet. *Expert Opin. Invest. Drugs* 21, (7), 901−904.

(6) Marchand, C.; Maddali, K.; Metifiot, M.; Pommier, Y. HIV-1 IN inhibitors: 2010 update and perspectives. *Curr. Top. Med. Chem.* **2009**, *9* (11), 1016−1037.

(7) Shimura, K.; Kodama, E. N. Elvitegravir: a new HIV integrase inhibitor. *Antiviral Chem. Chemother.* **2009**, *20* (2), 79−85.

(8) Lenz, J. C.; Rockstroh, J. K. S/GSK1349572, a new integrase inhibitor for the treatment of HIV: promises and challenges. *Expert Opin. Invest. Drugs* **2011**, *20* (4), 537−548.

(9) Nowotny, M. Retroviral integrase superfamily: the structural perspective. *EMBO Rep.* **2009**, *10* (2), 144−151.

(10) Ceccherini-Silberstein, F.; Malet, I.; D'Arrigo, R.; Antinori, A.; Marcelin, A. G.; Perno, C. F. Characterization and structural analysis of HIV-1 integrase conservation. *AIDS Rev.* **2009**, *11* (1), 17−29.

(11) Fransen, S.; Gupta, S.; Danovich, R.; Hazuda, D.; Miller, M.; Witmer, M.; Petropoulos, C. J.; Huang, W. Loss of raltegravir susceptibility by human immunodeficiency virus type 1 is conferred via multiple nonoverlapping genetic pathways. *J. Virol.* **2009**, *83* (22), 11440−11446.

(12) Delelis, O.; Malet, I.; Na, L.; Tchertanov, L.; Calvez, V.; Marcelin, A. G.; Subra, F.; Deprez, E.; Mouscadet, J. F. The G140S mutation in HIV integrases from raltegravir-resistant patients rescues catalytic defect due to the resistance Q148H mutation. *Nucleic Acids Res.* **2009**, *37* (4), 1193−1201.

(13) Delelis, O.; Thierry, S.; Subra, F.; Simon, F.; Malet, I.; Alloui, C.; Sayon, S.; Calvez, V.; Deprez, E.; Marcelin, A. G.; Tchertanov, L.; Mouscadet, J. F. Impact of Y143 HIV-1 integrase mutations on resistance to raltegravir in vitro and in vivo. *Antimicrob. Agents Chemother.* **2010**, *54* (1), 491−501.

(14) Mbisa, J. L.; Martin, S. A.; Cane, P. A. Patterns of resistance development with integrase inhibitors in HIV. *Infect. Drug Resist.* **2011**, *4*, 65−76.

(15) Blanco, J. L.; Varghese, V.; Rhee, S. Y.; Gatell, J. M.; Shafer, R. W. HIV-1 integrase inhibitor resistance and its clinical implications. *J. Infect. Dis.* **2011**, *203* (9), 1204−1214.

(16) Fransen, S.; Gupta, S.; Frantzell, A.; Petropoulos, C. J.; Huang, W. Substitutions at amino acid positions 143, 148, and 155 of HIV-1 integrase define distinct genetic barriers to raltegravir resistance in vivo. *J. Virol.* **2012**, *86* (13), 7249−7255.

(17) Marinello, J.; Marchand, C.; Mott, B. T.; Bain, A.; Thomas, C. J.; Pommier, Y. Comparison of raltegravir and elvitegravir on HIV-1 integrase catalytic reactions and on a series of drug-resistant integrase mutants. *Biochemistry* **2008**, *47* (36), 9345−9354.

(18) Lodi, P. J.; Ernst, J. A.; Kuszewski, J.; Hickman, A. B.; Engelman, A.; Craigie, R.; Clore, G. M.; Gronenborn, A. M. Solution structure of the DNA binding domain of HIV-1 integrase. *Biochemistry* **1995**, *34* (31), 9826−9833.

(19) Al-Mawsawi, L. Q.; Fikkert, V.; Dayam, R.; Witvrouw, M.; Burke, T. R., Jr.; Borchers, C. H.; Neamati, N. Discovery of a small-molecule HIV-1 integrase inhibitor-binding site. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103* (26), 10080−10085.

(20) Shkriabai, N.; Patil, S. S.; Hess, S.; Budihas, S. R.; Craigie, R.; Burke, T. R., Jr.; Le Grice, S. F.; Kvaratskhelia, M. Identification of an inhibitor-binding site to HIV-1 integrase with affinity acetylation and mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (18), 6894−6899.

(21) Lee, D. J.; Robinson, W. E., Jr. Preliminary mapping of a putative inhibitor-binding pocket for human immunodeficiency virus type 1 integrase inhibitors. *Antimicrob. Agents Chemother.* **2006**, *50* (1), 134−142.

(22) Kirkpatrick, D. L.; Watson, S.; Ulhaq, S. Structure-based drug design: combinatorial chemistry and molecular modeling. *Comb. Chem. High Throughput Screening* **1999**, *2* (4), 211−221.

(23) Brigo, A.; Lee, K. W.; Fogolari, F.; Mustata, G. I.; Briggs, J. M. Comparative molecular dynamics simulations of HIV-1 integrase and the T66I/M154I mutant: binding modes and drug resistance to a diketo acid inhibitor. *Proteins* **2005**, *59* (4), 723−741.

(24) Carlson, H. A.; Masukawa, K. M.; Rubins, K.; Bushman, F. D.; Jorgensen, W. L.; Lins, R. D.; Briggs, J. M.; McCammon, J. A. Developing a dynamic pharmacophore model for HIV-1 integrase. *J. Med. Chem.* **2000**, *43* (11), 2100−2114.

(25) Zhao, Y.; Li, W.; Zeng, J.; Liu, G.; Tang, Y. Insights into the interactions between HIV-1 integrase and human LEDGF/p75 by molecular dynamics simulation and free energy calculation. *Proteins* **2008**, *72* (2), 635−645.

(26) Perryman, A. L.; Forli, S.; Morris, G. M.; Burt, C.; Cheng, Y.; Palmer, M. J.; Whitby, K.; McCammon, J. A.; Phillips, C.; Olson, A. J. A dynamic model of HIV integrase inhibition and drug resistance. *J. Mol. Biol.* **2010**, *397* (2), 600−615.

(27) Huang, M.; Grant, G. H.; Richards, W. G. Binding modes of diketo-acid inhibitors of HIV-1 integrase: a comparative molecular dynamics simulation study. *J. Mol. Graphics Modell.* **2011**, *29* (7), 956−964.
(28) Weber, W.; Demirdjian, H.; Lins, R. D.; Briggs, J. M.; Ferreira, R.; McCammon, J. A. Brownian and essential dynamics studies of the HIV-1 integrase catalytic domain. *J. Biomol. Struct. Dyn.* **1998**, *16* (3), 733−745.
(29) Xue, W.; Liu, H.; Yao, X. Molecular mechanism of HIV-1 integrase-vDNA interactions and strand transfer inhibitor action: a molecular modeling perspective. *J. Comput. Chem.* **2012**, *33* (5), 527−536.
(30) 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* (28), 8892−8898.
(31) Mouscadet, J. F.; Delelis, O.; Marcelin, A. G.; Tchertanov, L. Resistance to HIV-1 integrase inhibitors: A structural perspective. *Drug Resist. Updates* **2010**, *13* (4−5), 139−150.
(32) Lee, M. C.; Deng, J.; Briggs, J. M.; Duan, Y. Large-scale conformational dynamics of the HIV-1 integrase core domain and its catalytic loop mutants. *Biophys. J.* **2005**, *88* (5), 3133−46.
(33) Maertens, G. N.; Hare, S.; Cherepanov, P. The mechanism of retroviral integration from X-ray structures of its key intermediates. *Nature* **2010**, *468* (7321), 326−329.
(34) Hare, S.; Gupta, S. S.; Valkov, E.; Engelman, A.; Cherepanov, P. Retroviral intasome assembly and inhibition of DNA strand transfer. *Nature* **2010**, *464* (7286), 232−236.
(35) 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* (36), 15910−15915.
(36) Sali, A.; Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234* (3), 779−815.
(37) Shen, M. Y.; Sali, A. Statistical potential for assessment and prediction of protein structures. *Protein Sci.* **2006**, *15* (11), 2507−24.
(38) Melo, F.; Sanchez, R.; Sali, A. Statistical potentials for fold assessment. *Protein Sci.* **2002**, *11* (2), 430−48.
(39) John, B.; Sali, A. Comparative protein structure modeling by iterative alignment, model building and model assessment. *Nucleic Acids Res.* **2003**, *31* (14), 3982−92.
(40) Colovos, C.; Yeates, T. O. Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci.* **1993**, *2* (9), 1511−1519.
(41) Case, D. A.; Cheatham, T. E., 3rd; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26* (16), 1668−88.
(42) Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* **2004**, *32* (Web Server issue), W665−W667.
(43) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **2006**, *65* (3), 712−25.
(44) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Antechamber: an accessory software package for molecular mechanical calculations. *J. Am. Chem. Soc.* **2001**, *222*, U403−U403.
(45) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford, CT, 2008.
(46) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79* (2), 926.
(47) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26* (16), 1781−802.
(48) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N-log (N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089.
(49) Tuckerman, M.; Berne, B. J.; Martyna, G. J. Reversible multiple time scale molecular dynamics. *J. Chem. Phys.* **1992**, *97*, 1990.
(50) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of< i> n</i>-alkanes. *J. Comput. Phys.* **1977**, *23* (3), 327−341.
(51) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant-pressure molecular-dynamics simulation-the Langevin piston method. *J. Chem. Phys.* **1995**, *103* (11), 4613−4621.
(52) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* **1994**, *101*, 4177.
(53) Wang, W.; Donini, O.; Reyes, C. M.; Kollman, P. A. Biomolecular simulations: recent developments in force fields, simulations of enzyme catalysis, protein-ligand, protein-protein, and protein-nucleic acid noncovalent interactions. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 211−43.
(54) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., 3rd Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, *33* (12), 889−97.
(55) Hess, B.; Kutzner, C.; Van Der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008**, *4* (3), 435−447.
(56) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; Van Gunsteren, W. F.; Mark, A. E. Peptide folding: when simulation meets experiment. *Angew. Chem., Int. Ed.* **1999**, *38* (1−2), 236−240.
(57) Grant, B. J.; Rodrigues, A. P.; ElSawy, K. M.; McCammon, J. A.; Caves, L. S. Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics* **2006**, *22* (21), 2695−6.
(58) Xue, W.; Jin, X.; Ning, L.; Wang, M.; Liu, H.; Yao, X. Exploring the Molecular Mechanism of Cross-Resistance to HIV-1 Integrase Strand Transfer Inhibitors by Molecular Dynamics Simulation and Residue Interaction Network Analysis. *J. Chem. Inf. Model.* **2013**, *53* (1), 210−222.
(59) Xue, W.; Liu, H.; Yao, X. Molecular mechanism of HIV-1 integrase−vDNA interactions and strand transfer inhibitor action: A molecular modeling perspective. *J. Comput. Chem.* **2012**, *33* (5), 527−536.
(60) Yi, M., Dynamical Basis for Drug Resistance of HIV-1 Protease. *BMC Struct. Biol.* 11.
(61) Williams, S. L.; Essex, J. W. Study of the conformational dynamics of the catalytic loop of WT and G140A/G149A HIV-1 integrase core domain using reversible digitally filtered molecular dynamics. *J. Chem. Theory Comput.* **2009**, *5* (2), 411−421.
(62) Cherepanov, P.; Maertens, G. N.; Hare, S. Structural insights into the retroviral DNA integration apparatus. *Curr. Opin. Struct. Biol.* **2011**, *21* (2), 249−256.

(63) Li, X.; Krishnan, L.; Cherepanov, P.; Engelman, A. Structural biology of retroviral DNA integration. *Virology* **2011**, *411* (2), 194−205.
(64) Hare, S.; Vos, A. M.; Clayton, R. F.; Thuring, J. W.; Cummings, M. D.; Cherepanov, P. Molecular mechanisms of retroviral integrase inhibition and the evolution of viral resistance. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107* (46), 20057−20062.
(65) Metifiot, M.; Maddali, K.; Naumova, A.; Zhang, X.; Marchand, C.; Pommier, Y. Biochemical and pharmacological analyses of HIV-1 integrase flexible loop mutants resistant to raltegravir. *Biochemistry* **2010**, *49* (17), 3715−3722.
(66) Reigadas, S.; Masquelier, B.; Calmels, C.; Laguerre, M.; Lazaro, E.; Vandenhende, M.; Neau, D.; Fleury, H.; Andreola, M. L. Structure-analysis of the HIV-1 integrase Y143C/R raltegravir resistance mutation in association with the secondary mutation T97A. *Antimicrob. Agents Chemother.* **2011**, *55* (7), 3187−3194.
(67) Mesplede, T.; Quashie, P. K.; Wainberg, M. A., Resistance to HIV integrase inhibitors. *Curr Opin HIV AIDS*.
(68) Malet, I.; Delelis, O.; Valantin, M. A.; Montes, B.; Soulie, C.; Wirden, M.; Tchertanov, L.; Peytavin, G.; Reynes, J.; Mouscadet, J. F.; Katlama, C.; Calvez, V.; Marcelin, A. G. Mutations associated with failure of raltegravir treatment affect integrase sensitivity to the inhibitor in vitro. *Antimicrob. Agents Chemother.* **2008**, *52* (4), 1351−8.
(69) Myers, R. E.; Pillay, D. Analysis of natural sequence variation and covariation in human immunodeficiency virus type 1 integrase. *J. Virol.* **2008**, *82* (18), 9228−35.
(70) Metifiot, M.; Johnson, B.; Smith, S.; Zhao, X. Z.; Marchand, C.; Burke, T.; Hughes, S.; Pommier, Y. MK-0536 inhibits HIV-1 integrases resistant to raltegravir. *Antimicrob. Agents Chemother.* **2011**, *55* (11), 5127−5133.
(71) Ammar, F. F.; Abdel-Azeim, S.; Zargarian, L.; Hobaika, Z.; Maroun, R. G.; Fermandjian, S. Unprocessed Viral DNA Could Be the Primary Target of the HIV-1 Integrase Inhibitor Raltegravir. *PLoS One* **2012**, *7* (7), e40223.
(72) Johnson, B. C.; Metifiot, M.; Pommier, Y.; Hughes, S. H. Molecular dynamics approaches estimate the binding energy of HIV-1 integrase inhibitors and correlate with in vitro activity. *Antimicrob. Agents Chemother.* **2012**, *56* (1), 411−419.
(73) Wainberg, M. A.; Mesplede, T.; Quashie, P. K. The development of novel HIV integrase inhibitors and the problem of drug resistance. *Curr. Opin. Virol.* **2012**, *2* (5), 656−662.
(74) Piquemal, J.-P.; Perera, L.; Cisneros, G. A.; Ren, P.; Pedersen, L. G.; Darden, T. A. Towards accurate solvation dynamics of divalent cations in water using the polarizable amoeba force field: From energetics to structure. *J. Chem. Phys.* **2006**, *125* (5), 054511−054511−7.
(75) Saxena, A.; Sept, D. Multisite ion models that improve coordination and free energy calculations in molecular dynamics simulations. *J. Chem. Theory Comput.* **2013**, *9* (8), 3538−3542.
(76) Bar-Magen, T.; Sloan, R. D.; Donahue, D. A.; Kuhl, B. D.; Zabeida, A.; Xu, H.; Oliveira, M.; Hazuda, D. J.; Wainberg, M. A. Identification of novel mutations responsible for resistance to MK-2048, a second-generation HIV-1 integrase inhibitor. *J. Virol.* **2010**, *84* (18), 9210−9216.
(77) Hare, S.; Smith, S. J.; Metifiot, M.; Jaxa-Chamiec, A.; Pommier, Y.; Hughes, S. H.; Cherepanov, P. Structural and functional analyses of the second-generation integrase strand transfer inhibitor dolutegravir (S/GSK1349572). *Mol. Pharmacol.* **2011**, *80* (4), 565−572.
(78) Kobayashi, M.; Yoshinaga, T.; Seki, T.; Wakasa-Morimoto, C.; Brown, K. W.; Ferris, R.; Foster, S. A.; Hazen, R. J.; Miki, S.; Suyama-Kagitani, A.; Kawauchi-Miki, S.; Taishi, T.; Kawasuji, T.; Johns, B. A.; Underwood, M. R.; Garvey, E. P.; Sato, A.; Fujiwara, T. In Vitro antiretroviral properties of S/GSK1349572, a next-generation HIV integrase inhibitor. *Antimicrob. Agents Chemother.* **2011**, *55* (2), 813−821.