J Mol Model (2013) 19:4349–4368
DOI 10.1007/s00894-013-1943-4

ORIGINAL PAPER

Computational design of a full-length model of HIV-1 integrase: modeling of new inhibitors and comparison of their calculated binding energies with those previously studied

Selami Ercan · Necmettin Pirinccioglu

Received: 21 April 2013 / Accepted: 11 July 2013 / Published online: 2 August 2013

Abstract A full-length model of integrase (IN) of the human immunodeficiency virus type 1 (HIV-1) was constructed based on the distinctly resolved X-ray crystal structures of its three domains, named N-terminal, catalytic core and C-terminal. Thirty-one already known inhibitors with varieties of structural differences as well as nine newly tested ones were docked into the catalytic core. The molecular dynamic (MD) and binding properties of these complexes were obtained by MD calculations. The binding energies calculated by molecular mechanic/Poisson Boltzmann solvation area were significantly correlationed with available IC_{50}. Four inhibitors including two newly designed were also docked into the full-length model and their MD behaviors and binding properties were calculated. It was found that one of the newly designed compounds forms a better complex with HIV-1 IN compared to the rest including raltegravir. MD calculations were performed with AMBER suite of programs using ff99SB force field for the proteins and the general Amber force field for the ligands. In conclusion, the results have produced a promising standpoint not only in the construction of the full-length model but also in development of new drugs against it. However, the role of multimer formation and the involvement of DNAs, and their subsequent effect on the complexation and inhibition, are required to arrive at a conclusive decision.

Keywords Binding energies · Computational modeling · Docking · HIV-1 integrase · Full-length · MM/PBSA · Molecular dynamics · New inhibitors

Introduction

Retroviruses possess ability to reverse transcribe a single-stranded RNA genome into a linear double stranded DNA. One of their main proteins is integrase (IN) which functions to insert the linear viral DNA into the genome of the target cell to establish a stable infection, which catalyses the integration process in two distinct steps [1]. There are three functionally distinct domains of integrases, characterized by biochemical and mutational analyses as the N-terminal domain (NTD; residues 1 to 49) containing a zinc binding HHCC motif and contributing to multimer formation [2, 3] as the C-terminal domain (CTD; identified as residues 212 to 288 in deletion studies) non-specifically binds DNA [4–7] and as the catalytic core domain (CCD; residues 50 to 211) containing the catalytic triad DD35E motif that is well conserved among the retroviral integrase superfamily [2, 8–10].

Inhibiting integration process has occupied a considerable place in targeting antiretroviral drugs. Some of the clinically relevant IN inhibitors, which are named as IN strand transfer inhibitors, have been proved to show selectivity for the strand transfer reaction and only weakly inhibit 3' processing [11].

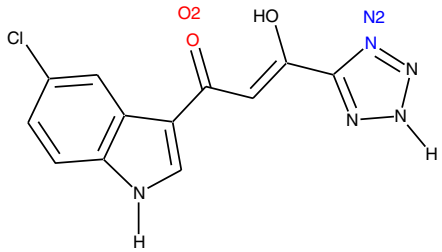
This work was partially presented as a poster (P156) at Chemical Physics Congress X TOBB University of Economics and Technology Ankara, Turkey, 10–12 October 2012

Electronic supplementary material The online version of this article (doi:10.1007/s00894-013-1943-4) contains supplementary material, which is available to authorized users.

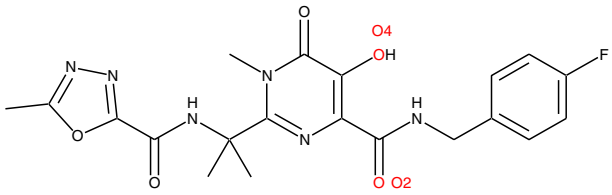
S. Ercan · N. Pirinccioglu (✉)
Department of Chemistry, University of Dicle, Faculty of Science, 21280 Diyarbakir, Turkey
e-mail: pirincn@dicle.edu.tr

Present Address:
S. Ercan
Department of Chemistry, University of Batman, Faculty of Arts & Science, Batman, Turkey

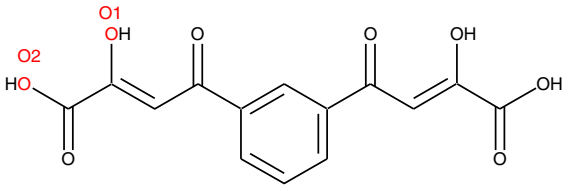
Table 1 Structures of ligands included in this study



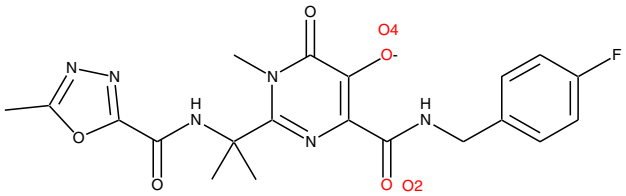
L01 [21]



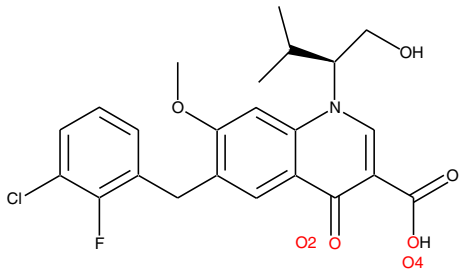
L02 (Raltegravir) [51]



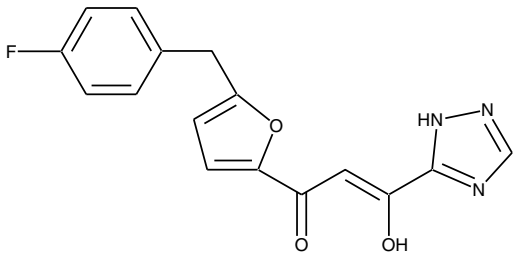
L03 [53]



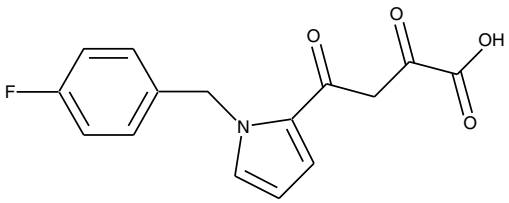
L02t [52]



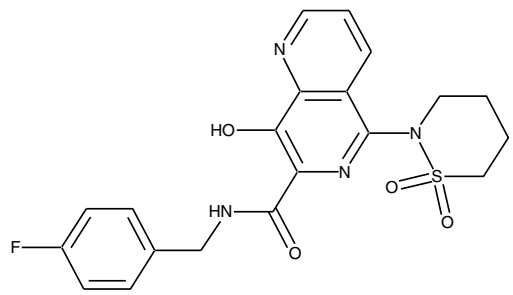
L04 (Elvitegravir) [54]



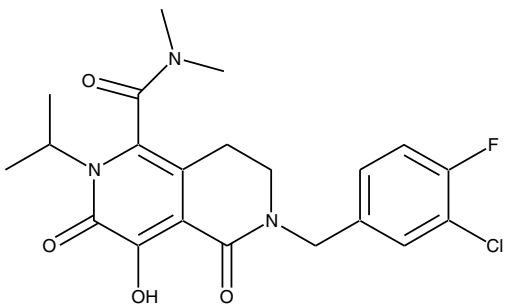
L05 [55,56]



L06 [57]



L07 [54,58]



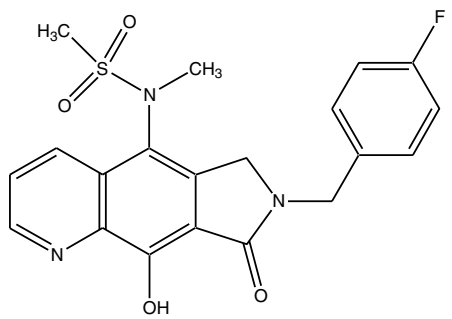
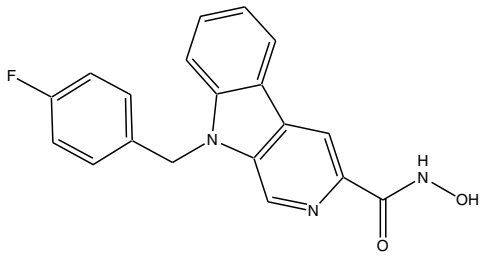
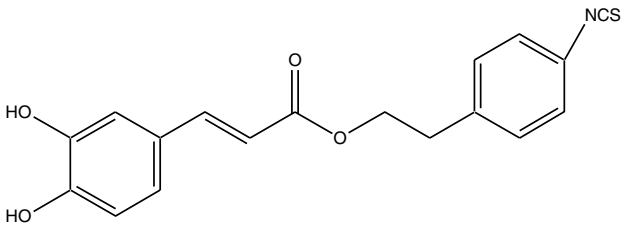


Table 1 (continued)

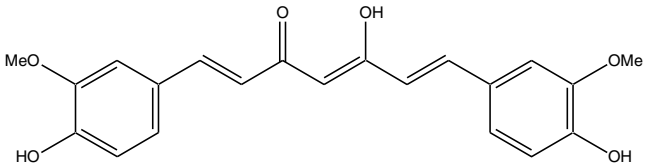
L08 [59]



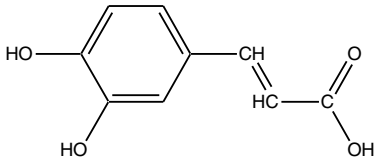
L09 [60]



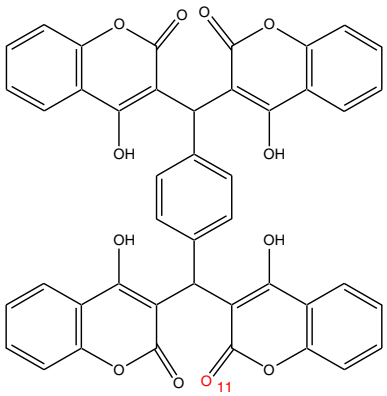
L10 [58]



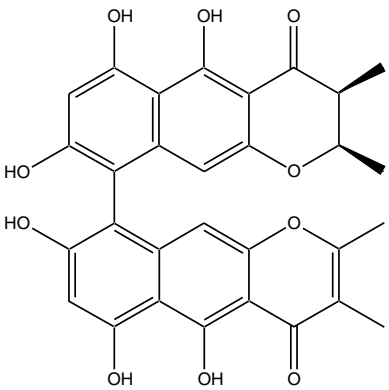
L11 [61]



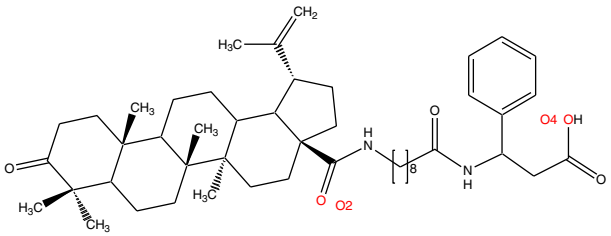
L12 [62]



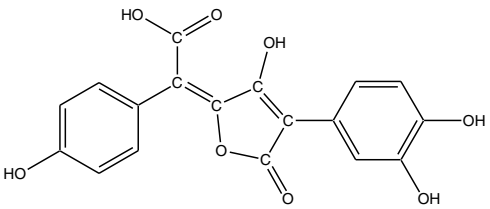
L15 [65]



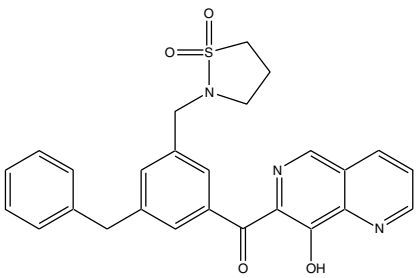
L13 [63]



L14 [64]



L17 [66]



L16 [66]

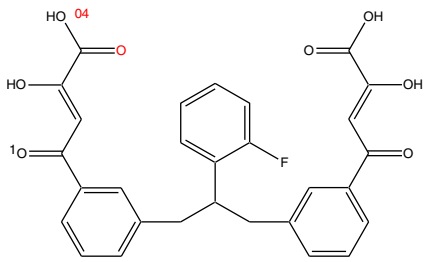
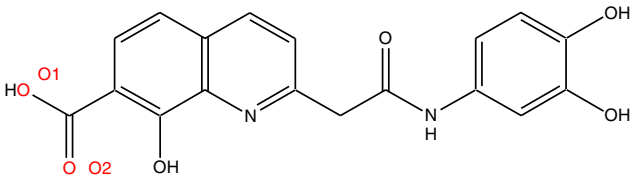
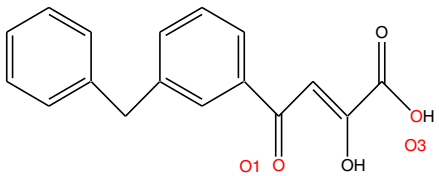


Table 1 (continued)

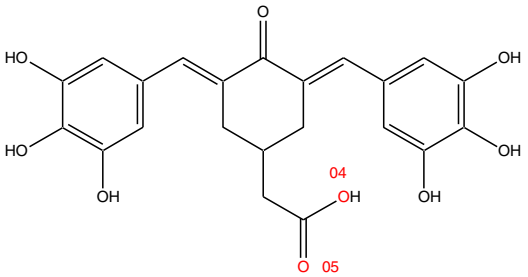
L19 [66]



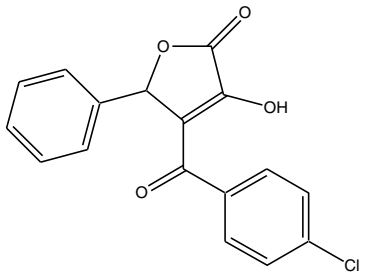
L20 [66]



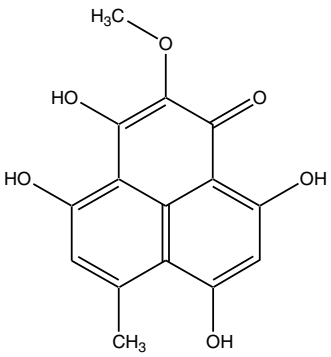
L21 [66]



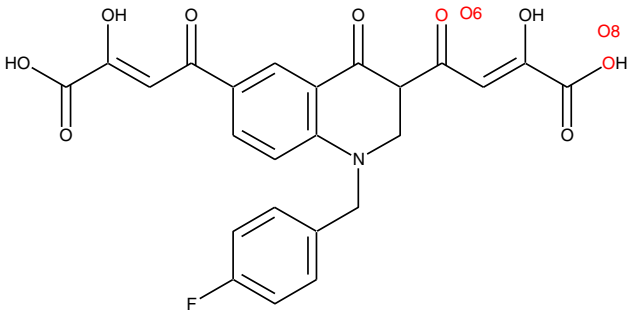
L18 [66]



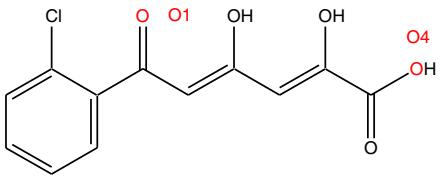
L22 [66]



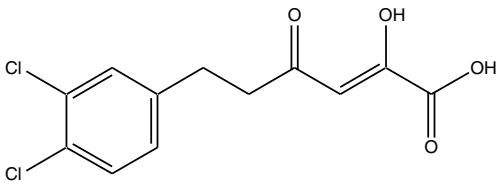
L23 [67]



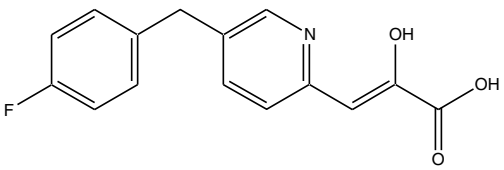
L24 [68]



L27 [69]



L25 [69]



L26 [69]

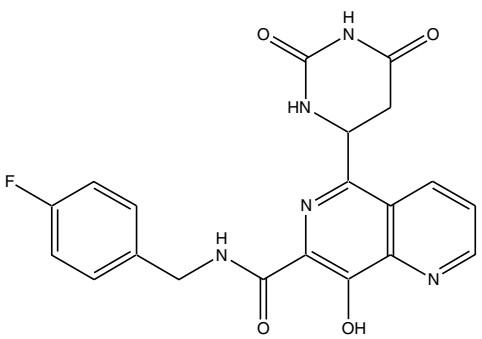
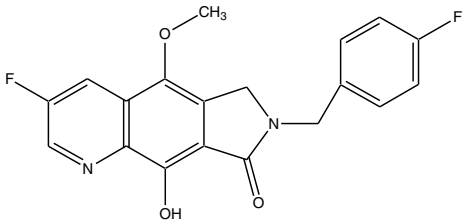
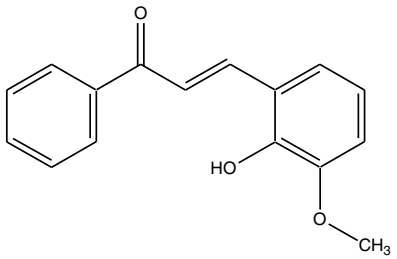


Table 1 (continued)

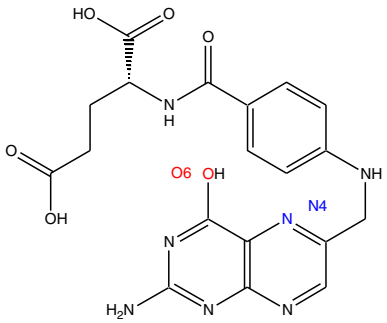
L28 [69]



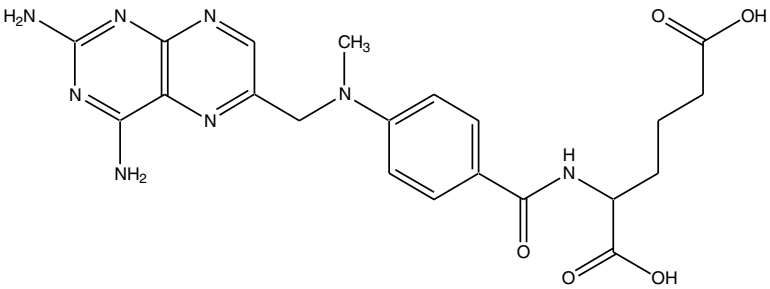
L29 [69]



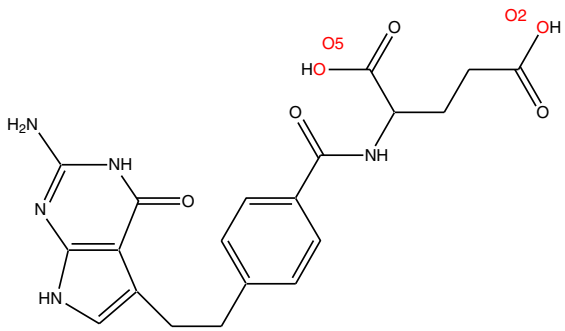
L30 [70]



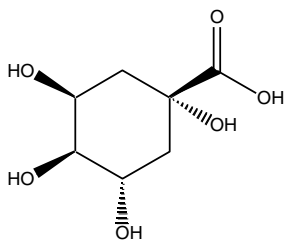
L31 [71]



L32 (folic acid) [72]

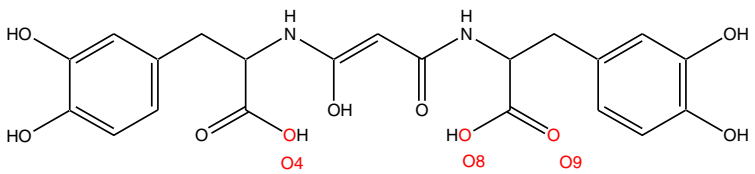


L33 [72]

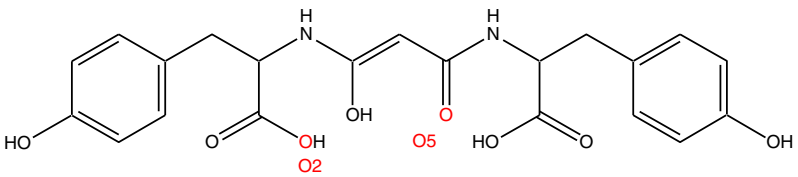


L34 [72]

L35

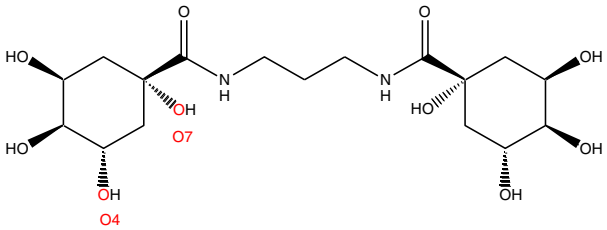


LGA

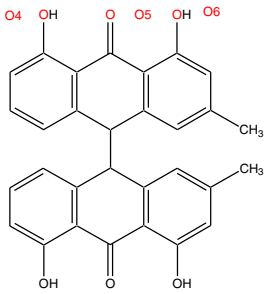


LGB

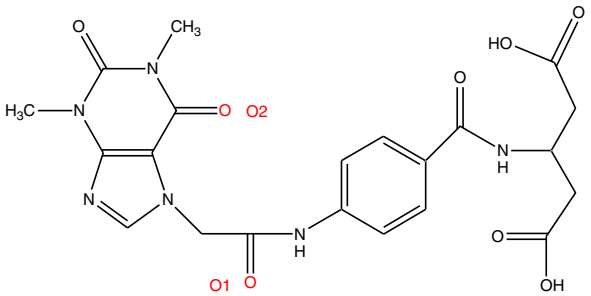
Table 1 (continued)



LGC



LGD [73]



LGE

Among them raltegravir is only approved by the Food and Drug Administration (FDA) while elvitegravir and dolutegravir are in advanced clinical trials [12].

One of the hampers in the development of new and effective candidates of HIV-1 IN is the lack of a refined crystal structure of full-length HIV-1 IN and its complexes with the viral DNA (or at least with the relevant oligonucleotide ends of the viral genome) [1, 13, 14] although individual domains have been solved by nuclear magnetic resonance (NMR) [3, 4, 15, 16] and crystallography [4, 17–25], including several two-domain structures [18, 26]. However, recently, several crystal structures of the full-length integrase of the prototype foamy virus (PFV) complexed with its cognate viral DNA have been reported [27]. The authors argue that these crystal structures provide a plausible inhibition mechanism of DNA strand transfer, which may be relevant for HIV-1 IN. But this does not still provide the required information for anti-HIV drug design because of the relatively low resolution of the crystal structure and the rather low sequence similarity of PFV IN versus HIV-1 IN. Molecular dynamic approaches have been used to estimate binding free energies of some inhibitors to this model complexed with DNA [28].

Two valuable chapters regarding experimental and theoretical studies of HIV-1 IN have recently been reported [29, 30]. Few reports have also issued the construction of a full-length model of the enzyme from the individual X-rays of the domains and even including its complex with DNA [31–36]. We also report the construction of a full-length model of HIV-1 IN from known X-ray structures of three domains with similar procedures. First of all, the missing residues forming individual domains were completed from X-ray structures where available. Forty inhibitors with varieties of structural changes including raltegravir and four newly designed ones were docked into the CCD and their molecular dynamic calculations were studied. The binding free energies of these ligands to the protein were calculated by molecular mechanic/Poisson-Boltzmann surface area (MM/PBSA). Four ligands including two newly proposed ones were also docked into the constructed full-length model, followed by MD and MM/PBSA calculations.

Methods

The amino acid sequence of the full-length protein (288 residues) was taken from literature [37]. Crystal structure information of the different domains was obtained from the coordinates found in the Protein Data Bank (PDB) [38]. The PDB codes of the structural templates used for the study are 1QS4 [21], 1BI4 and 1BL3 (structures of the CCD dimmer) [24], 1WJA (solution structure of the NTD) [15], 1IHV (structure of the CTD) [16], 1EX4 (the combined CCD and CTDs) [18] and 1K6Y (the combined CCD and NTD) [26]. First of all, all the missing residues from two combined domains (1EX4 and 1K6Y) were taken from the individual X-ray structures where available and they were superimposed using the CCD coordinates as a reference. The resulting coordinates were recorded as a separate file in PDB format and the coordinates of one of the CCDs were deleted. As a result a model with residues from 1–270 based on X-ray structures was

completed. The missing residues (271–288) were constructed by a similar protocol as described in the literature [39].

Computational modeling

Forty ligands as presented in Table 1 are included in the study. Nine of them (**L32**, **L33**, **L34**, **L35**, **LGA**, **LGB**, **LGC**, **LGD** and **LGE**) are newly employed as HIV-1 IN inhibitors. **L35** and **LGD** are naturally occurring while**LGA**, **LGB**, **LGC** and **LGE** were newly designed for the other purposes in the group. They were all optimized with Gaussian 03 using semi-empirical AM1 method [40]. Conjugated keto acid were computed at B3LYP/6-31+(d) level to see the course of tautomerism and the conformer with a lower energy was chosen for the calculations. All molecular dynamic (MD) simulations were conducted by AssistedModelBuilding with Energy Refinement (AMBER version 11) [41] suite of programs at TR-GRID clusters (TÜBİTAK).Austian model with bond and charge correction (AM1-Bcc) atomic partial charges for the ligands were determined by antechamber module of AMBER package [42]. Xleap as implemented in AMBER was employed to prepare parameter/topology and coordinate files and it was also used to solvate and to neutralize the system for MD simulations. All the proteins and the complexes were solvated in a TIP3P [43] water box with dimensions of 10 Å from the solute having a space of 0.4 Å, initially generated at the boundary of the complex and the solvent molecules during the solvation process. ff99SB [44] force field was employed for the protein while for the ligands, the general Amber force field (GAFF) [45] was adopted in simulation because it handles small organic molecules. A SHAKE algorithm was applied to constrain all bonds containing hydrogen atoms [46]. The non-bonded cut off was kept at 10 Å, and long range electrostatic interactions were treated by the particle mesh Ewald (PME) [47] method with fast Fourier transform grid having approximately 0.1 nm space. Trajectory snapshots were taken at each 1 ps, which were finally used for analysis. Parameters for residues associated with zinc are taken from the literature and employed without modification [48, 49]. Ptrajmodul of AMBER was used to obtain energy and RMSD changes as wall as bond analyses involving the interactions between the ligands and the protein during the molecular dynamic simulations. They are presented in GraphPad Prism 4. Cluster analyses were carried out using the MMTSB Toolset. Three dimensional structures were displayed using Chimera (UCSF) [50].

Molecular dynamic simulations

The proteins

For all the structures concerned, the following procedure was applied. The introduced missing residues in each structure including their neigboring amino acids were energy-minimized, keeping the rest of the protein restrained with a force of 500 kcal mol^{-1} A^{-2}, using 2500 steps of steepest descent, followed by 2500 steps of conjugate gradient (igb=0) with a cutoff of 999. Then these parts were heated from 0 K to 400 K with a restrain of 500 kcal mol^{-1} A^{-2} on the rest of the structure without bondry conditions (igb=0) with a cutoff of 999 for a period of 3.2 ps in four steps. The final structure for each protein was solvated with TIP3P model as mentioned above and the solvent was energy-minimized keeping all the solutes restrained with a force of 500 kcal mol^{-1} A^{-2}, using 2500 steps of steepest descent, followed by 2500 steps of conjugate gradient. Then the whole system was energy-minimized using 2500 steps of steepest descent, followed by 2500 steps of conjugate gradient without any restraint. The solvent was heated from 0 K to 300 K for a period of 200 ps, keeping all the solutes restrained with a force constant of 50 kcal mol^{-1} A^{-2}, followed by equilibration for a period of 1 ns at 300 K, with a restrain of 1 kcal mol^{-1} A^{-2} on the solutes. Final simulations, the production phase, were performed for 10 ns in the canonical ensemble at 300 K temperature and 1 atm pressure without any restraint for the individual domains while the simulation periods were 30 ns for the protiens with 270 and 288 residues and 15 ns for the matel free protein. Step size was 2 fs for the entire simulation. A Langevin thermostat and barostat were used for coupling the temperature and pressure.

The complexes

Systems were minimized in two steps; in the first step, the solvent was energy-minimized keeping the protein, metal(s)/counter ions and ligands restrained with a force of 500 kcal mol^{-1} A^{-2}, using 2500 steps of steepest descent, followed by 2500 steps of conjugate gradient. In the second step, the whole

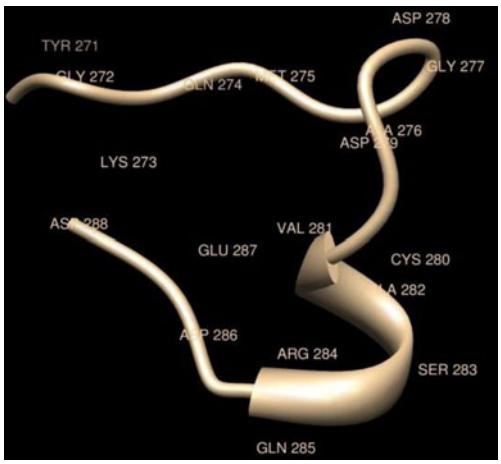
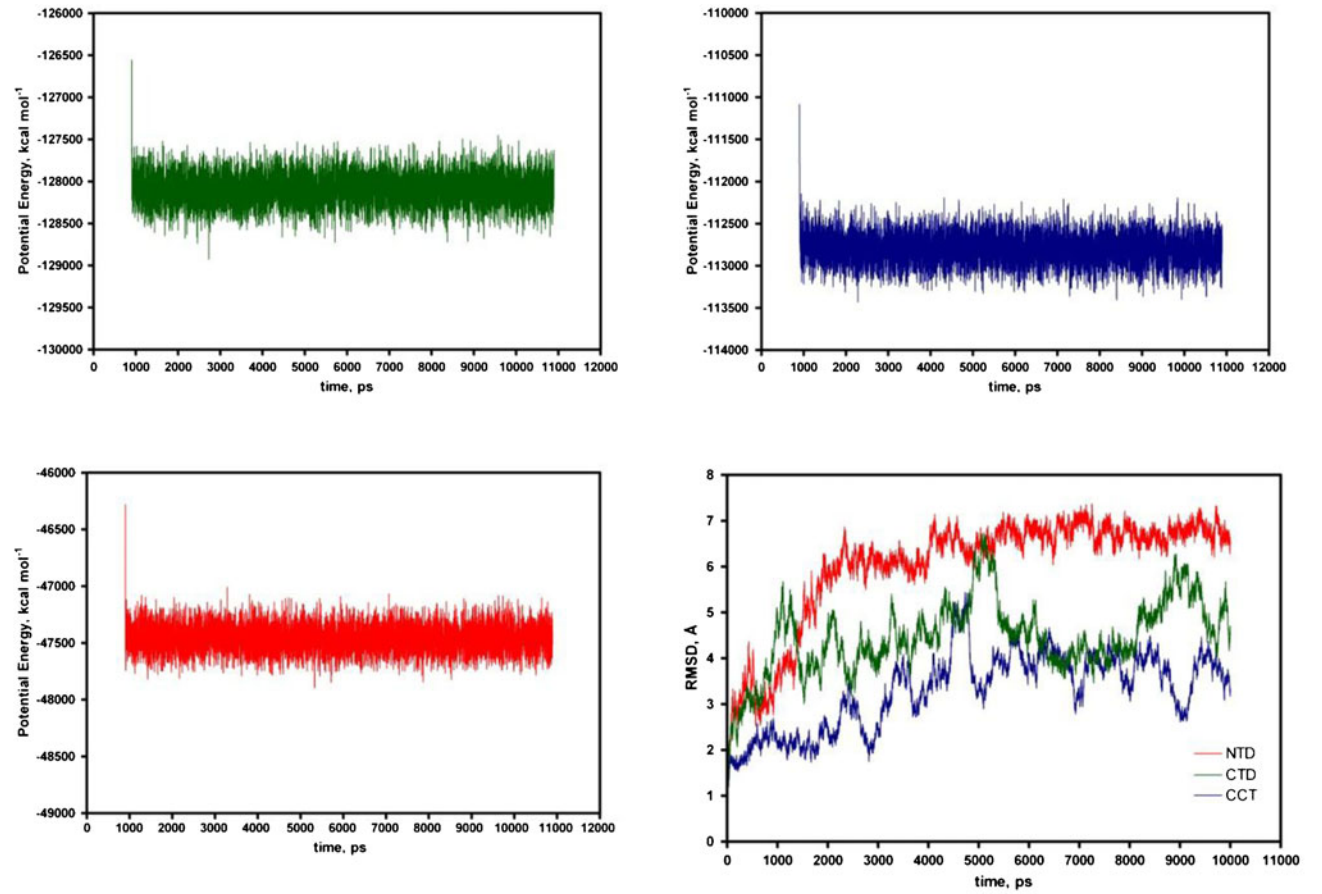


Fig. 1 The structure of the peptide prepared to complete the missing sequence of integrase extracted from MD trajectories

system was energy-minimized using 2500 steps of steepest descent, followed by 2500 steps of conjugate gradient without any restraint. The solvent was heated from 0 K to 300 K for a period of 200 ps, keeping all the solutes restrained with a force constant of 50 kcal mol^{-1} A^{-2}, followed by equilibration for a period of 1 ns at 300 K, with a restrain of 1 kcal mol^{-1} A^{-2} on the protein-lignad complex. Final simulations, the production phase, were performed for 10 ns in the canonical ensemble at 300 K temperature and 1 atm pressure without any restraint. Step size was 2 fs for the entire simulation. A Langevin thermostat and barostat were used for coupling the temperature and pressure.

Docking study

In this study, docking studies were performed by Dock 6.5 [74], with default settings to obtain a population of possible conformations and orientations for the guests in the binding



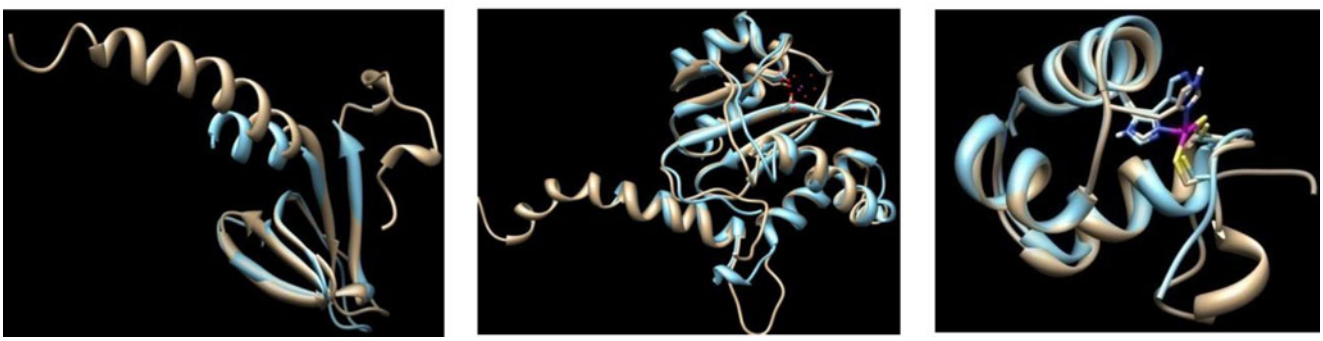


Fig. 2 Potential energy of three domains [CTD (*green*), CCD (*blue*) and NTD (*red*)] as a function of time during MD for a period of 10 ns at 300 K and their backbone RMSD during the same MD, compared to their starting coordinates (*upper*). The structures of three domains (*open brown*) from the frames with the largest population obtained from the cluster analyses (supplementary data) of the trajectories superimposed with their corresponding original X-ray structures (*open blue*) (*lower*)

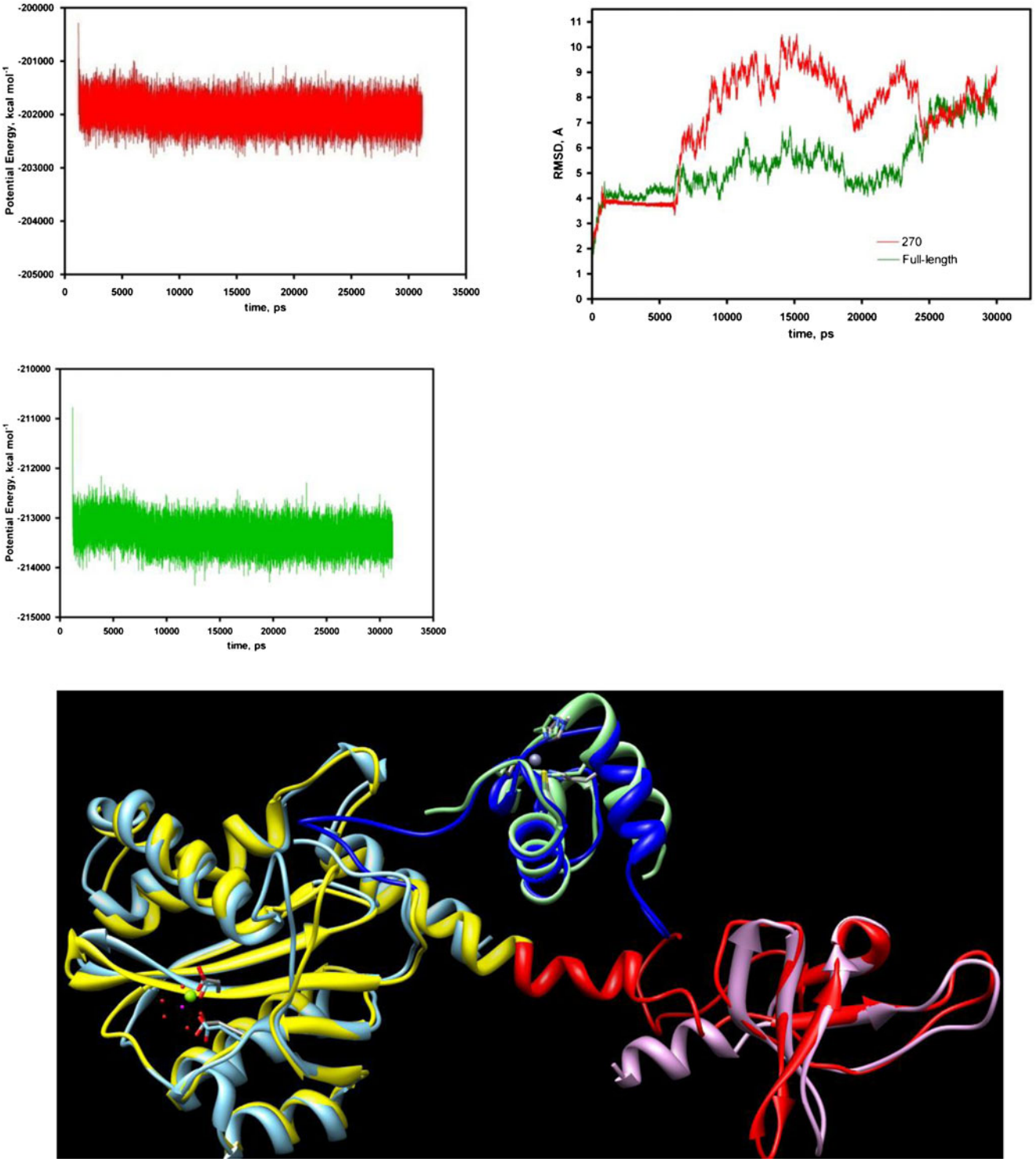


Fig. 3 Potential energy of two models [the full-length (*green*) and 270 (*red*)] as a function of time during MD for a period of 30 ns at 300 K for the former model and 30 ns at 300 K for the later and their corresponding backbone RMSD during the same MD, compared to the starting coordinates (*upper*). The superimposed structure of the constructed full-length model (*open brown*) with X-ray structures of its corresponding domains (*lower*)

site. A sphere around the center of the binding pocket was formed to define as binding pocket for the docking studies. All torsion angles in each compound were allowed to rotate freely.

MM-PBSA

The MM-PBSA module of AMBER (v11) was applied to compute the binding free energy (ΔG_{bind}) of each complex

[75, 76]. For each complex, a total number of 100 snapshots were extracted from the trajectories of each complex. The interaction energy was calculated according to the following equation:

$$\Delta G = \Delta E_{MM} + \Delta G_{sol}^{polar} + \Delta G_{sol}^{nonpolar} - T\Delta S_{solute}, \quad (1)$$

where ΔE_{MM} is the gas-phase energy, denoting the sum of molecular mechanical (MM) energies of molecules from internal (ΔE_{int}), electrostatic (ΔE_{ele}), and van der Waals energies (ΔE_{vdw}). The solvation free energy (ΔG_{sol}) is composed of polar ($\Delta G_{solpolar}$) and nonpolar ($\Delta G_{solnonpolar}$) parts. $T\Delta S$ is the contribution of conformational entropy to the binding. Here, the polar solvation free energy was calculated by solving the Poisson-Boltzmann equation using the program Delphi II [77]. The dielectric boundary was defined using a 1.4 Å probe on the atomic surface. The values of the interior dielectric constant and the exterior dielectric constant were set to 1 and 80, respectively. The non-polar solvation free energy was calculated from the solvent-accessible surface area (*SASA*) algorithm [78]:

$$\Delta G_{nonpolar} = \gamma SASA + b, \quad (2)$$

where γ is the surface tension proportionality constant (the value is 0.00542 kcal mol^{-1} Å^{-2}). The free energy of nonpolar solvation for a point solute (b) is set to 0.92 kcal mol^{-1}.

During conformational searching and the evaluation of configuration integrals, *W*elec is computed with a simplified but fast generalized Born model. The electrostatic solvation energy of each energy-well is then corrected toward a more accurate but time-consuming finite-difference solution of the Poisson equation. The dielectric cavity radius of each atom is set to the mean of the solvent probe radius 1.4 Å for water and the atom's van der Waals radius, and the dielectric boundary between the molecule and the solvent is the solvent-accessible molecular surface. The solvation calculations use a water dielectric constant of 80. The method produces the final estimated binding free energy using both Poisson-Boltzmann and generalized Born solvation models. The change of the entropy upon binding, $-T\Delta S$, was estimated using the NMODE module of AMBER for only three snapshots.

Results

The structure of the peptide prepared to complete the missing sequence of integrase extracted from MD trajectories is presented in Fig. 1. This peptide was linked to the CTD obtained from 1EX4 by xleap and the structure was subjected to MD simulations. The completed CCD and NTD were also subjected to MD calculations. Potential energy of these three domains as a function of time during MD for a period of 10 ns at 300 K and their backbone RMSD during the same MD, compared to their starting coordinates, are presented in Fig. 2, along with the structures of one of the frame with the largest population obtained from the cluster analysis (supplementary data) of the trajectories superimposed with their original X-ray structures.

Two models were constructed based on the X-ray coordinates of two combined structures as mentioned in the method part, one with 270 residues and the other with 288 residues, so called the full-length model to see the effect of the missing 18 residues on the molecular dynamic behaviors as well as on conformational changes in HIV-1 IN. These structures were subjected to MD calculations. Potential energy of the proteins as a function of time during MD for a period of 30 ns and their corresponding backbone RMSD during the same MD, compared to the starting coordinates, are presented in Fig. 3. The structures of one of the frames with the largest population obtained from the cluster analysis (supplementary data) of the trajectories superimposed with the corresponding starting coordinates are also displayed in Fig. 3. For comparison, the same structures were also superimposed with a previously reported model [32] as well as with the foamy prototype [27] as shown in Fig. 4. Metal ions were removed from the full-length model and this metal free structure was

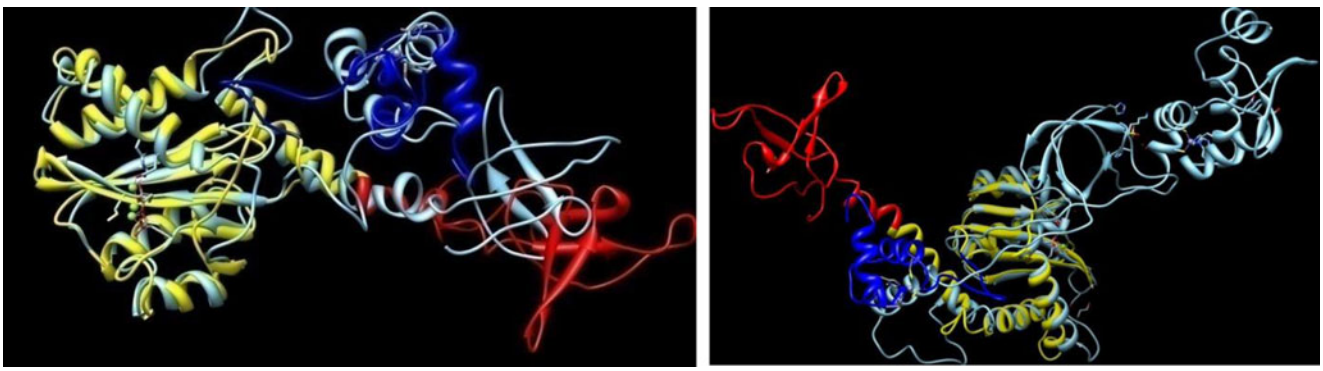


Fig. 4 The superimposed structure of the full-length model with the previously reported model (*left*) [32] as well as with the foamy prototype (*right*) [27]. For the sake of clarity, DNA in both models is omitted

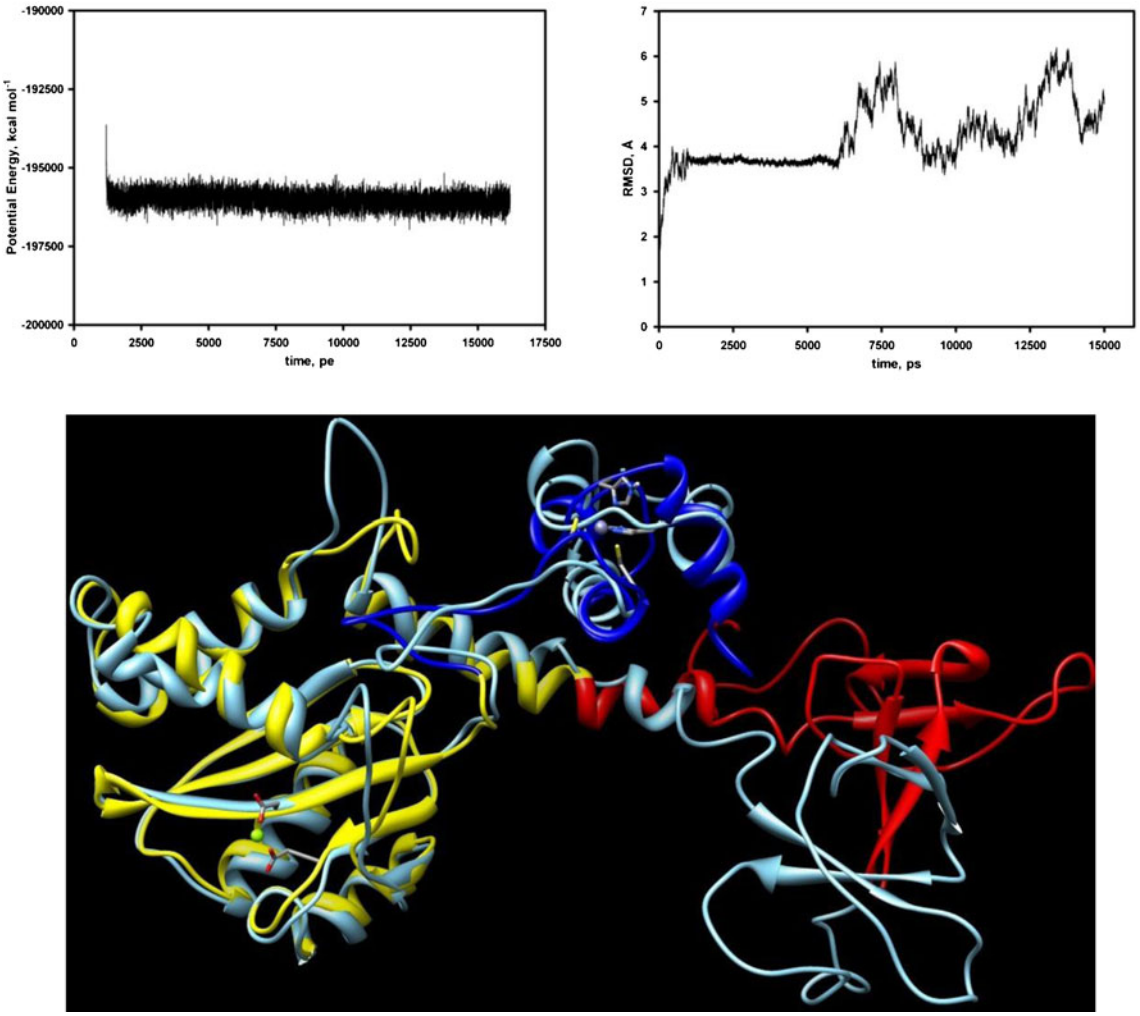


Fig. 5 Potential energy of the full-length model free of metals as a function of time during MD for a period of 15 ns at 300 K (*upper left*) and its corresponding backbone RMSD during the same MD, compared to the starting coordinates (*upper right*). The superimposed structures of the full-length free of metals with the one including metals (*lower*)

subjected to MD calculations to see their effects on the structural characteristics of the protein. The results are summarized in Fig. 5.

Forty ligands were docked into the active site of the structure of the CCD with dock scores ranging from −94.41 to −33.79 kcal mol^{-1}. For the validity of the docking algorithm, **L01** was chosen as a reference because of its available X-ray structure with the CCD as illustrated in Fig. 6. The complexes of 22 ligands with good dock scores (roughly lower than −50 kcal mol^{-1}) were subjected to MD calculations. Potential energy of each complex as a function of time during MD for a period of 13–15 ns at 300 K and its backbone RMSD during the same MD, compared to the starting coordinates, are presented in the supplementary data. The RMSD of only each ligand is also analyzed as shown in the supplementary data. Four ligands (**L01**, **L02**, **LGA** and **LGB**) were also docked into one of the structures of the modeled full-length protein with a higher population (mentioned earlier as in Fig. 3b) as presented in Fig. 7. These complexes were also subjected to MD simulations. The results of energy and RMSD changes as well as structural comparison for each complex are displayed in Figs. 8 and 9. Changes in RMSD corresponding to the coordinates of only ligands are displayed in

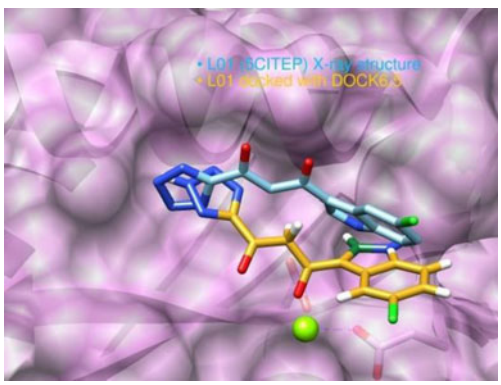


Fig. 6 The location of **L01** docked into the active site of the CCD produced by Dock 6.5., superimposed with the X-ray structure [21]

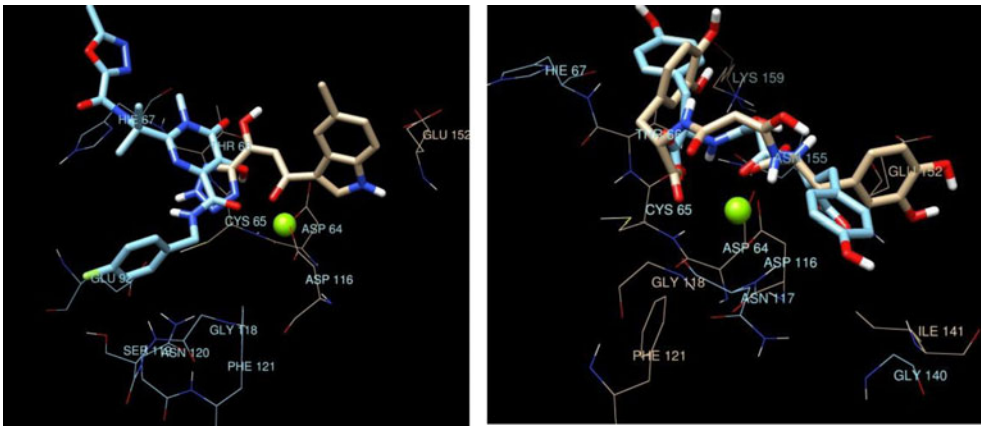


Fig. 7 The locations of four ligands (**L01**, **L02**, **LGA** and **LGB**) docked into the active site of the full-length model produced by Dock 6.5. Left: **L01** (*open brown*) and **L02** (*open blue*); right: **LGA** (*open brown*) and **LGB** (*open blue*)

the supplementary data. The binding free energies calculated by MM/PBSA for the complexes of the CCD with 22 ligands are listed in Table 2 and those for the complexes of the full-length model with four ligands are listed in Table 3. The final estimated binding free energies calculated by generalized Born solvation model for the complexes of the CCD with ligands are significantly correlated with available IC_{50} values as shown in Eq. 3 where **L02** is excluded from the regression (Fig. 10). A poor correlation is somehow obtained for the binding eneries calculated by Poisson-Boltzmann solvation model.

$$\Delta G_{GB} = 30.20 \pm 9.179\ IC_{50} - 88.16 \pm 8.512\ \left(r^2 = 0.5198\right) \text{ where } \mathbf{L02} \text{ is excluded} \quad (3)$$

$$\Delta G_{GB} = 24.48 \pm 11.25\ IC_{50} - 79.95 \pm 10.02\ \left(r^2 = 0.3009\right) \text{ where } \mathbf{L02} \text{ is included} \quad (4)$$

$$\Delta G_{PB} = 8.925 \pm 4.394 IC_{50} - 36.33 \pm 4.256\ \left(r^2 = 0.3143\right) \text{ where } \mathbf{L02} \text{ is excluded} \quad (5)$$

$$\Delta G_{PB} = 6.528 \pm 4.953 IC_{50} - 32.89 \pm 4.593\ \left(r^2 = 0.3009\right) \text{ where } \mathbf{L02} \text{ is included} \quad (6)$$

Discussion

As mentioned earlier, the lack of a 3D structure of HIV-1 IN and its complex with DNAs (either host or native) form a major problem in the development of HIV-1 IN oriented drugs despite many attempts using mainly core domains [79–83] and some employing modeled full-length structure of HIV-1 IN. However, current study involves the individual construction of the domains before building the full-length model of HIV-1 IN, to see the effect of their explicit contributions to the 3D structure of the model. MD calculations indicate that the CTD undergoes dynamical behaviors, mostly caused by conformational changes in the peptide prepared to complete the missing sequence as well as those in the loop composed of the residues 228–239 (Fig. 2a). On the other hand, the rest of the domain does not experience considerable conformetional changes during MD calculations (Fig. 2a). Likewise, the

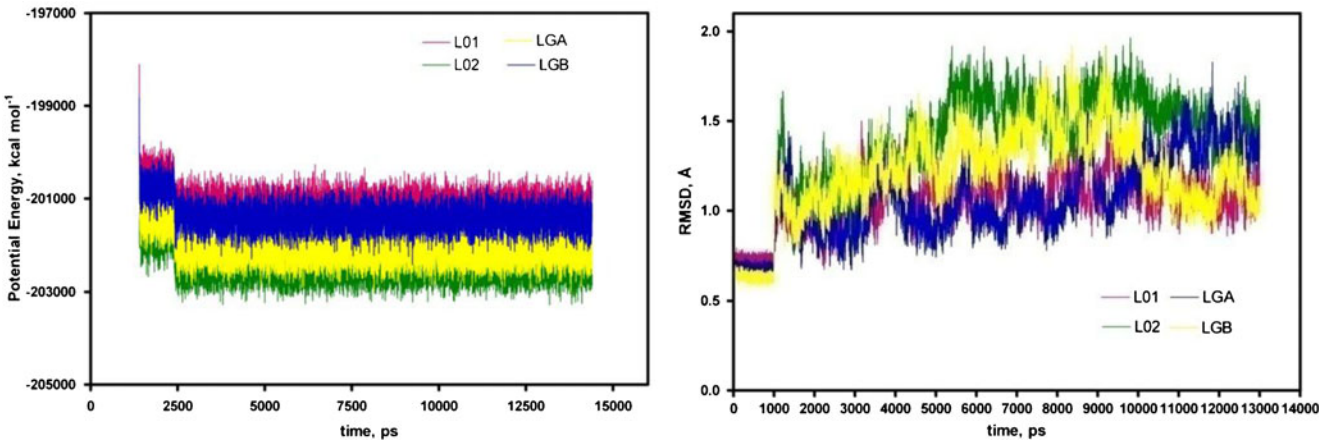


Fig. 8 Potential energy of the complexes of the full-length model with **L01, L02**, **LGA** and **LGB** as a function of time during MD for a period of 13 ns at 300 K and their corresponding backbone RMSD during the same MD, compared to the starting coordinates

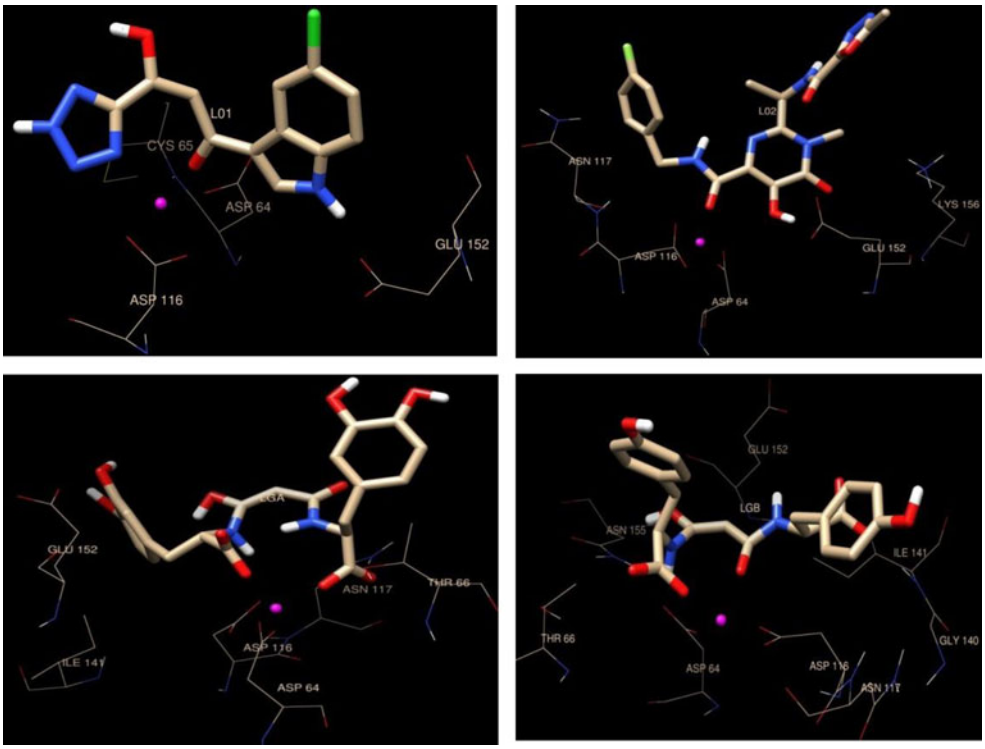


Fig. 9 The structures of the complexes of the full-length model with **L01** (*upper left*), **L02** (*upper right*), **LGA** (*lower left*) and **LGB** (*lower right*) obtained from the frame with the largest population superimposed with their corresponding starting coordinates. For each picture, open brown represents the starting coordinates whereas open blue represents the complex subjected to MD calculations

calculations performed on the CCD point out that the helix and sheet parts of this domain does not undergo conformational changes while the loops, particularly those between 139 and 150, and 186 and 195 residues seem to experience significant conformational changes during the MD (Fig. 2b). They also shows that magnesium ion forms slightly shorter bonds with the residues involved and shifts about 1.33 Å compared with X-ray structure as illustrated in Fig. 11. The results produced from the MD calculations of NTD point out that parameters employed for zinc and associated residues are acceptable since this part of the protein maintains its structural characteristics (Fig. 2c). They also signify that the domain does not experience significant conformational changes except some changes in the loop with residues 39–60.

The superimposed structure of the constructed full-length model with X-ray structures of its corresponding domains (Fig. 3a) provides a promising posture to start with. Potential energy of this model as a function of time during MD indicates that the structure has converged to a reasonable level (Fig. 3b). The backbone RMSD during the same MD (Fig. 3) demonstrates that most of the fluctuation is due to the conformational changes in the last 18 residues. It seems that there is not much change in the rest of the protein as observed in the individual domains (Fig. 3c). It was observed that the full-length model lacking metal ions experiences significant conformational changes, especially in the regions where ions are held and the HHCC motif in NTD and DDE motif in CCD compared to the one involving metal ions as seen in Fig. 5.

The superimposed structure of the model with the previously reported one [31] illustrates that deformations occurrs in the previously constructed model in the one of the helices between the residues 19 and 5 in the NTD region and also significant changes in the sheets in the CTD region (Fig. 4c). It is also obvious to see that a tendency of folding at the

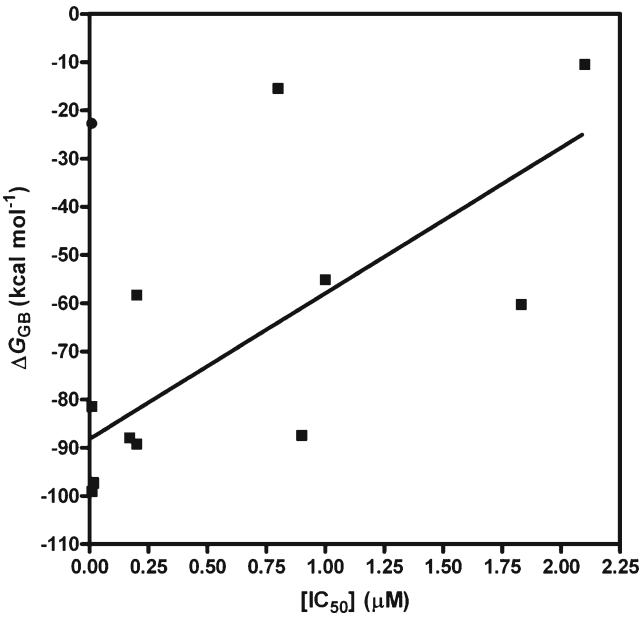


Fig. 10 The correlation of binding energies for the complexes of the ligands with the CCD calculated by MM/PBSA with available IC_{50} of ligands. The value for **L02** is excluded from the correlation

Table 2 Dock scores and binding energies calculated by MM/PBSA for the complexes of CCD with the ligands

Ligands	IC_{50}(μM)[a]	DS (kcal mol^{-1})[b]	ΔG_{GB}(kcal mol^{-1})[c]	ΔG_{PB} (kcal mol^{-1})[d]
L01	2.1	−41.19	−10.45±4.07	−6.87±4.00
L02	0.009	−50.57	−22.68±6.40	−12.28±3.74
L02t	0.009	−51.85	−81.42±6.90	−29.16±4.12
L03	1.83	−58.65	−60.24 ±5.50	−30.64±3.02
L04	0.015	−56.38	−97.11±11.26	−22.21±4.20
L05	0.02	−43.01		–
L06	0.05	−49.16		–
L07	0.007	−46.22		
L08		−46.58		
L09		−43.89		
L10		−45.43		
L11		−48.45		
L12		−49.14		
L13	0.8	−70.93	−15.46±8.04	−18.26±4.44
L14	4.0	−37.53		–
L15		−37.52		
L16	1.1	−43.56		–
L17	0.17	−94.41	−87.88±9.81	−46.23±7.14
L18	0.01	−51.50	−99.04±7.78	−40.38±3.44
L19	0.01	−47.76		–
L20	0.2	−59.07	−89.21±6.08	−45.49±3.86
L21	0.9	−57.62	−87.38±7.53	−37.74±4.88
L22	0.2	−56.61	−58.30±6.48	−30.50±5.15
L23		−34.38		
L24		−33.79		
L25	1.00	−51.74	−55.13±5.53	−27.61±3.60
L26	0.01	−47.98		–
L27	0.017	−59.85	−97.29±8.84	−36.52±5.71
L28	0.095	−43.98		–
L29	0.01	−44.46		–
L30		−40.37		
L31		−41.31		
L32	–	−60.77	−14.83±8.67	−6.24±6.04
L33	–	−60.70	−102.47±11.66	−49.60±4.95
L34	–	−59.67	−5.23±9.09	−1.61±4.81
L35	–	−42.45	−50.09±5.08	−22.42±2.99
LGA	–	−67.21	−156.20±8.14	−84.05±5.32
LGB	–	−67.81	−138.86±9.56	−45.24±6.35
LGC	–	−58.74	−1.34±11.86	−8.88±5.60
LGD	–	−41.23	−13.98±6.88	−0.37±10.37
LGE	–	−63.01	−89.15±10.68	−28.60±5.65

[a]) The half maximal inhibitory concentration, which is a measure of the effectiveness of a compound in inhibiting biological or biochemical function. They are taken from the corresponding literature listed in Table 1

[b]) Dock scores obtained by Dock 6.5

[c]) ΔG_{GB} is the final estimated binding free energy using generalized Born solvation model

[d]) ΔG_{PB} is the final estimated binding free energy using Poisson-Boltzmann solvation model

connecting point (residues 209–210) between the CCD and the CTD occurs. This change was also observed in the complex of our model with the ligands as it will be discussed later (Figs. 8–9). As to the comparasion with the struture of the FPV, apart from the CCDs, there are quite large structural differences, particularly in the orientations of the domains (CTD and NTD) (Fig. 4).

Dock 6.5 procedure successfully located 40 ligands on the surface of magnesium ion binding site with scores −94.41 to −33.79 kcal mol^{-1}. The docking algorithm is proved to be reliable since it locates **L01** with a similar conformation to that of X-ray structure (Fig. 6). The complexes of 22 ligands including five newly designed ones with good scores were chosen for MD calculations. Binding free energies calculated

Table 3 Binding energies and energy contribution to binding energies calculated by MM/PBSA for the complexes of the full-length model with the ligands

Ligands	EE[a]	vdW[b]	ΔG_{GB}, kcal mol^{-1}	ΔG_{PB}, kcal mol^{-1}
L01	−57.94	−9.18	−5.66±5.88	2.04±4.13
L02	−78.42	−12.60	−12.64±4.88	−12.29±5.98
LGA	−366.71	−4.49	−113.75±11.70	−50.06±9.58
LGB	−280.44	−7.87	−126.43±9.38	−60.24±4.38

[a]) ELE is non-bonded electrostatic energy+1,4-electrostatic energy

[b]) vdW is non-bonded van der Waals energy+1,4-van der Waals energy

for the ligands by MM/MPBSA demonstrate that raltegravir **L02** has a lower value of binding energy (−22.68 kcal mol^{-1}) compared to its salt **L02t** (−81.42 kcal mol^{-1}) while **L22**, **L33** and **LGB** (newly designed) possess a binding energy around −45 kcal mol^{-1} compared to **LGA**, also newly designed, which has the largest negative energy of binding (around −156 kcal mol^{-1}) compared to the rest in the list (Table 2). It is quite interesting to see that calculated binding energies by MM/PBSA are significantly correlated with available IC_{50} determined experimentally (Fig. 10).

Calculations indicate that **L01**, **L02, L02t** and **L32** do not undergo significant conformational changes during MD calculations while the rest shows rather dynamic behaviors**, L35** with little conformational changes. Two water molecules accompany **L01**, **L13**, **L17**, **L34** and **L35** to interact with the metal while only one interacts with the metal in the case of the rest of the ligands. However, the complexes of **LGA** and **LGD** do not involve water coordination. The results from MD calculations reveal that **L01** interacts with the metal ion via its O2 and N2 donor atoms with average bond lengths of 2.19 and 3.37 Å and forms hydrogen bonds with ASN117 by hydroxyl and with GLU152 by HN of the indole ring.

It is interesting to see that **L02** (raltegravir) and its salt **L02t** coordinate to the ion in a similar mode through O1 and O4 donors with average bond lengths of 1.91 and 3.40 Å for the former and 1.90 and 2.09 Å as indicated in Fig. 12. THR66 and LYS159 interact with oxadiazole ring while SER119 interacts with florine through hydrogen bonding in the complex of the protein with **L02t**. These two interactions were not observed for the complex of **L02**. However, the difference in the calculated binding energies between raltegravir and its salt, which is almost fourfold, is largely due to a better chelation capacity of the salt compared with its neutral form, obviously bacause of a better donor capacity of O4 in its ionized (O^-) form compared with that in its neutral form (OH). On the other hand, experimantal calculations indicated that they have similar IC_{50} values [51, 52]. It is apparent that **L02** would be in its ionized form at physiological pH since it has a pK_a of 6.7 [84] and therefore it would interact with integrase with this form, which is a key point in the interaction of integrase with these types of ligands.

Magnesium is chelated by **L04** through O2 and 04 donors with the average bond lenths of 1.85 and 3.52 Å. Its phenolic hydroxyl forms hydrogen bond with LYS159, which also provides electrostatic interactions with the flour, while the other hydroxyl forms hydrogen bond with CYS65. The results also confirm that **L13** interacts with the metal ion via O11 donor with an average bond length of 1.99 Å. One of the hydroxyl groups on the chromene ring forms a hydrogen bond with ILE141 while the other two rings which are not involved in the chelation form pi-pi interactions with TYR143 and PHE139 residues. **L17** exhibits a very peculiar interaction mode with the active site of HIV-1 IN. It is folded to chelate magnesium through the carboxylate and the amide carbonyl attached to the chrysene ring. The other amide group forms a hydrogen bond with HIE67 and ILE141 is involved in van der Waals interactions with the chrysene ring. **L18** chelates magnesium via O1 and O2 donors with average bond lengths of 1.93 and 1.90 Å and also form van der Waals interactions with ILE141 as in L17 through biphenyl rings. **L20** shows similar chelating mode to that of **L18** and the flouroaryl group is sandwiched between TYR99 and ILE97. **L21** only interacts with the magnesium ion with the carboxylate group and forms

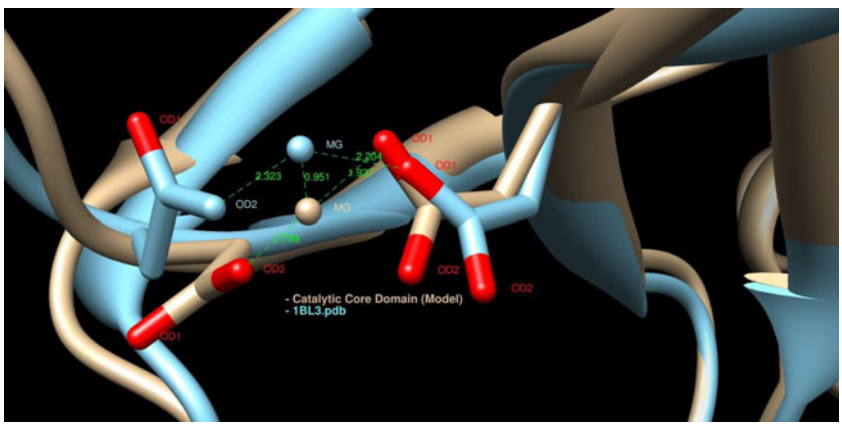


Fig. 11 The position of the magnesium ion in the protein subjected to MD calculations superimposed on the original X-ray structure [24]

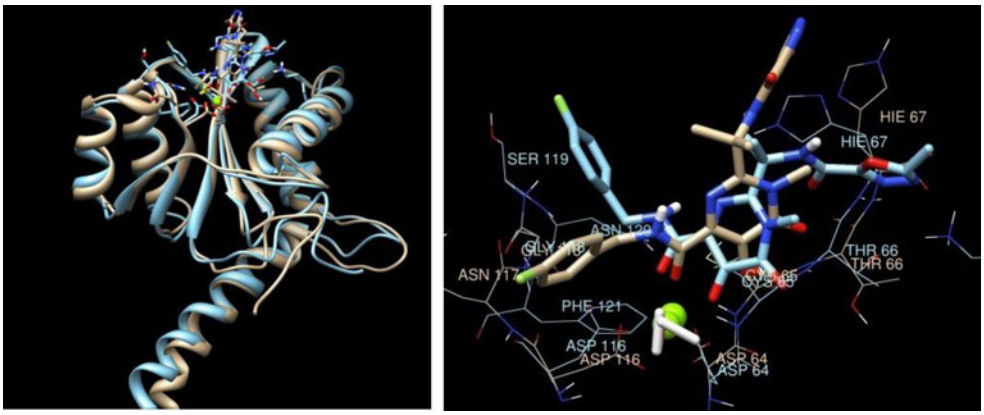


Fig. 12 The position of **L02** (in *open brown*) and its salt **L02t** (in *open blue*) in the active site of the CCD obtained from MD calculations

hydrogen bonds with GLU152, ASN155 and LYS156 via aryl hydroxyls while with THR66 via the amide carbonyl.**L22** is also coordinated to the metal ion via only one carboxyl group and one of its aryl hydroxyls forms a hydrogen bond with SER119 whereas the metal is chelated by the carboxyl and keto functions of **L25** (O1 and O4 donor atoms) with average bond lengths of 1.96 and 1.89 Å. **L27** also chelates the ion via the carboxyl and keto functions (O6 and O8 atoms) with average bond lengths of 2.00 and 2.13 Å and forms hydrogen bonds with LYS156 and 159, and THR66 via its other head bearing keto carboxyl group. **L32** is bound to the metal by N4 and O6 donors with average bond lenths of 2.19 and 2.20 Å although the coordination with the carboxyl groups is expected. Instead, these groups form hydrogen bonds with LYS156 and LYS159 as in **L27**. Magnesium is chelated by the terminal carboxylate of **L33** whose pteridine part is surrounded by three LYS residues (156, 159 and 160) through hydrogen bonds. **L34** interacts with the metal via pyrimidin ring (N4 donor with 2.11 Å) rather than via two carboxyl groups, one of which forms a hydrogen bond with HIE67. **L35** coordinates to magnesium via carboxylate with 1.89 Å and forms hydrogen bonds with THR66 and ASN155. However, it is the one with poorest binding energy among the ligands studied (Table 2). It is quite noticeable to find out that the binding energies calculated by MM/PBSA for the ligands are significantly correlated with the available IC_{50} values as illustrated in Fig. 10.

From the ligands, newly designed, **LGA** is held within the active site by coordinating to the metal through O4, O8 and O9 donor atoms with 1.88, 1.90 and 1.84 Å average bond lengths and also by interacting with THR66 and LYS159 through hydrogen bonds. It is interesting to see that the ligand caps the magnesium ion so that the coordination of water molecules are completely blocked, which are thought to be involved in the cleavage of the viral DNA (3'-processing) [85]. **LGB** is structurally similar to **LGA** but it has a different binding mode compared to **LGA**, probably due to the existence of two phenolic hydroxyl groups in **LGA. LGB** forms coordination with the ion through O2 and O5 atoms with 1.86 and 1.91 Å, and has hydrogen bonds through THR66 GLU152 and ASN156. **LGC** exhibits to interact with magnesium via O4 and O7 donors with average bond lengths of 2.03 and 2.03 Å, and forms hydrogen bonds with CYS65, GLU152, ASN155 and LYS156. **LGD** interacts with the metal ion to form a

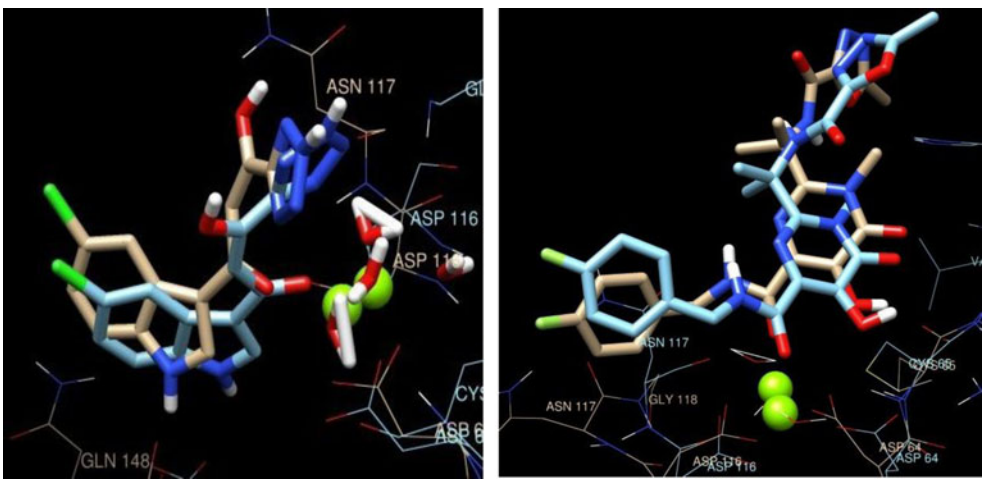


Fig. 13 The conformation of **L01** (*left*) and **L02** (*right*) in the proteins: the full-length (*open blue*) and CCD (*open brown*)

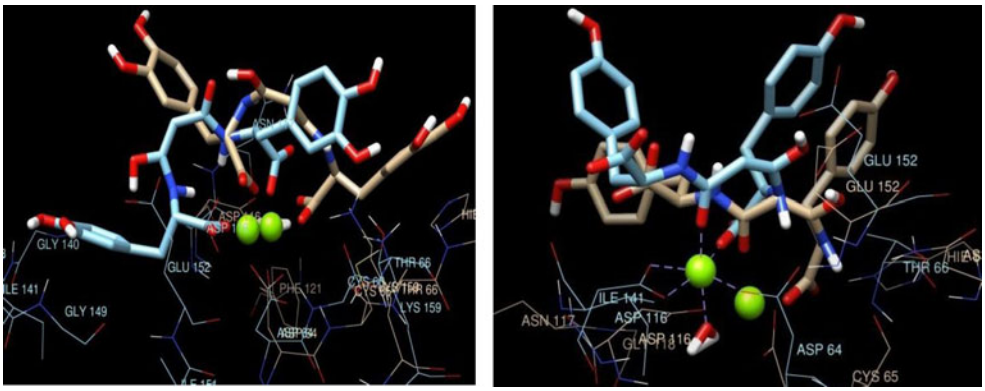


Fig. 14 The conformation of **LGA** (*left*) and **LGB** (*right*) in the proteins: the full-length (*open blue*) and CCD (*open brown*)

complex with CCD through one of the phenolic hydroxyls and carbonyl donors with 3.231 and 2.090 Å average bond lengths, and forms hydrogen bonds TYR143, GLN148 and GLU152. **LGE** interacts with one carbonyl oxygen in the caffeine ring and the peptide carbonyl attached to it with average bond lengths of 2.004 and 2.008 Å, instead of two carboxylates, which form hydrogen bonds with THR66, LYS159 and ASN155. The other carbonyl group in the caffeine ring also forms a hydrogen bond with GLN148.

In summary, calculations indicate that all ligands interact with the protein through the coordination of at least one of the donor atoms in ligands with the metal ion, regardless of structural differences. It seems that the chelation and the interaction of the other parts of ligands with the periphery of the active site of the protein via non-covalent interactions enhance the binding energy, particularly in **LGA** as mentioned above.

The complexation mode of the full-length model with **L01** and **L02** is quite similar to those with CCD (Fig. 13), though there is some conformational changes in the residues between 199–217 in the CCD complexed with **L01** and a slight change in the same area of the full-length model complexed with **L02**, leading connection to the NTD. However, the complexation mode of **LGA** with the full-length model is slightly different than that with the CCD, with almost changes in the protein backbone (a tendency of folding, thus deformation in the connecting point between the CCD and CTD) observed in the complexes with **L01** and **L02** (Figs. 8–9). The later occurs via the coordination of O4, O8 and O9 atoms with average bond lengths of 1.87, 1.93 and 1.91 Å while these average bond lengths are 1.86, 1.86 and 3.73 Å in the complex with the full-length model as exhibited in Fig. 14. This means that O9 does not strongly participate in the complexation with the model as in the CCD. Whereas **LGB** as shown in **L01** and **L02** gives similar binding mode to interact with the full-length model compared to the CCD. The average bond-lengths for the interacting atoms (O2, O5 and N2) with magnesium ion for the complex of **LGB** with the full-length model are 1.86, 1.91 and 3.53 Å and those with the CDD are 1.87, 1.88 and 3.30 Å.

Conclusions

A full-length structure of HIV-1 IN was successfully modeled based on available X-ray structures of its individual domains. Docking and molecular dynamic calculations produced for the complexes of the CCD with a large number of accessible inhibitors presented consistent results with experimental ones. The calculations obtained for this model complexed with two newly designed compounds offered significant results that may lead to the development of new anti-HIV drugs. The involvement of DNAs and two metal ions in the complextion as well as the experimental calculation of the binding energies of newly designed compounds to HIV-IN will be the focus of further perspectives.

Acknowledgments We are grateful to TUBITAK (Scientific and Technological Council of Turkey) for computational facilities and to DA Case (University of California, San Francisco) for a waiver licence of AMBER.

References

1. Pommier Y, Johnson A, Marchand C (2005) Integrase inhibitors to treat HIV/AIDS. Nat Rev Drug Discov 4:236–248
2. Bushman FD, Engelman A, Palmer I, Wingfield P, Craigie R (1993) Domains of the integrase protein of human immunodeficiency virus type 1 responsible for polynucleotidyl transfer and zinc binding. Proc Natl Acad Sci U S A 90:3428–3432
3. Cai M, Huang Y, Caffrey M, Zheng R, Craigie R, Clore GM, Gronenborn AM (1998) Solution structure of the His12 ->Cys mutant of the N-terminal zinc binding domain of HIV-1 integrase complexed to cadmium. Protein Sci 7:2669–2674
4. Eijkelenboom AP, Sprangers R, Hård K, Puras Lutzke RA, Plasterk RH, Boelens R, Kaptein R (1999) Refined solution structure of the C-terminal DNA-binding domain of human immunovirus-1 integrase. Proteins 36:556–564
5. Engelman A, Hickman AB, Craigie R (1994) The core and carboxylterminal domains of the integrase protein of human immunodeficiency virus type 1 each contribute to nonspecific DNA binding. J Virol 68:5911–5917
6. Jenkins TM, Esposito D, Engelman A, Craigie R (1997) Critical contacts between HIV-1 integrase and viral DNA identified by structure-based analysis and photo-crosslinking. EMBO J 16:6849–6859

7. Lu R, Ghory HZ, Engelman A (2005) Genetic analyses of conserved residues in the carboxyl-terminal domain of human immunodeficiency virus type 1 integrase. J Virol 79:10356–10368
8. Chen H, Wei SQ, Engelman A (1999) Multiple integrase functions are required to form the native structure of the human immunodeficiency virus type I intasome. J Biol Chem 274: 17358–17364
9. Engelman A, Bushman FD, Craigie R (1993) Identification of discrete functional domains of HIV-1 integrase and their organization within an active multimeric complex. EMBO J 12:3269–3275
10. Hickman AB, Palmer I, Engelman A, Craigie R, Wingfield P (1994) Biophysical and enzymatic properties of the catalytic domain of HIV-1 integrase. J BiolChem 269:29279–29287
11. Marchand C, Maddali K, Métifiot M, Pommier Y (2009) HIV-1 IN inhibitors: 2010 update and perspectives. Curr Top Med Chem 9:1016–1037
12. Métifiot M, Marchand C, Maddali K, Pommier Y (2010) Resistance to integrase inhibitors. Viruses 2:1347–1366
13. Neamati N (2001) Structure-based HIV-1 integrase inhibitor design: a future perspective. Expert Opin Investig Drugs 10:281–296
14. Chen IJ, Neamati N, MacKerell AD Jr (2002) Structure based inhibitor design targeting HIV-1 integrase. Curr Drug Targets Infect Disord 2:217–234
15. Cai M, Zheng R, Caffrey M, Craigie R, Clore GM, Gronenborn AM (1997) Solution structure of the N-terminal zinc binding domain of HIV-1 integrase. Nat Struct Biol 4:567–577
16. Lodi PJ, Ernst JA, Kuszewski J, Hickman AB, Engelman A, Craigie R, Clore GM, Gronenborn AM (1995) Solution structure of the DNA binding domain of HIV-1 integrase. Biochemistry 34:9826–9833
17. Bujacz G, Alexandratos J, Qing ZL, Clement-Mella C, Wlodawer A (1996) The catalytic domain of human immunodeficiency virus integrase: ordered active site in the F185H mutant. FEBS Lett 398:175–178
18. Chen JC, Krucinski J, Miercke LJ, Finer-Moore JS, Tang AH, Leavitt AD, StroudR M (2000) Crystal structure of the HIV-1 integrase catalytic core and C-terminal domains: a model for viral DNA binding. Proc Natl Acad Sci USA 97:8233–8238
19. Cherepanov P, Ambrosio AL, Rahman S, Ellenberger T, Engelman A (2005) Structural basis for the recognition between HIV-1 integrase and transcriptional coactivator p75. Proc Natl Acad Sci USA 102:17308–17313
20. Dyda F, Hickman AB, Jenkins TM, Engelman A, Craigie R, Davies DR (1994) Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyltransferases. Science 266:1981–1986
21. Goldgur Y, Craigie R, Cohen GH, Fujiwara T, Yoshinaga T, Fujishita T, Sugimoto H, Endo T, Murai H, Davies DR (1999) Structure of the HIV-1 integrase catalytic domain complexed with an inhibitor: a platform for antiviral drug design. Proc Natl Aca dSci USA 96:13040–13043
22. Goldgur Y, Dyda F, Hickman AB, Jenkins TM, Craigie R, Davies DR (1998) Three new structures of the core domain of HIV-1 integrase: an active site that binds magnesium. Proc Natl Acad Sci USA 95:9150–9154
23. Greenwald J, Le V, Butler SL, Bushman FD, Choe S (1999) The mobility of an HIV-1 integrase active site loop is correlated with catalytic activity. Biochemistry 38:8892–8898
24. Maignan S, Guilloteau JP, Zhou-Liu Q, Clement-Mella C, Mikol V (1998) 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 282:359–368
25. Wielens J, Headey SJ, Jeevarajah D, Rhodes DI, Deadman J, Chalmers DK, Scanlon MI, Parker MW (2010) Crystal structure of the HIV-1 integrase core domain in complex with sucrose reveals details of an allosteric inhibitory binding site. FEBS Lett 584:1455–1462
26. Wang JY, Ling H, Yang W, Craigie R (2001) Structure of a two-domain fragment of HIV-1 integrase: implications for domain organization in the intact protein. EMBO J 20:7333–7343
27. Hare S, Gupta SS, Valkov E, Engelman A, Cherepanov P (2010) Retroviral intasome assembly and inhibition of DNA strand transfer. Nature 464:232–236
28. Johnson BC, Métifiot M, Pommier Y, Hughes SH (2012) Molecular dynamics approaches estimate the binding energy of HIV-1 ınıtegrase ınhibitors and correlate with in vitro activity. Antimicrob Agents Chemother 56:411–419
29. Jalkoski M, Alexandratos JN, Wlodawer A, Bujacz G (2011) Structural studies of retroviral integrases. In: Neamati N (ed) HIV-1 integrase: mechanism and inhibitor design, 1st edn. Wiley, New York, pp 35–49
30. Liao C, Nicklaus MC (2011) HIV-1 integrase-DNA models. In: Neamati N (ed) HIV-1 integrase: mechanism and inhibitor design, 1st edn. Wiley, New York, pp 429–455
31. Rajeshri Karkia G, Tanga Y, Burke TR Jr, Nicklau MC (2004) Model of full-length HIV-1 integrasecomplexed with viral DNA as template for anti-HIV drug design. J Comput Aided Mol Des 18:739–760
32. De Luca L, Vistoli G, Pedretti A, Barreca ML, Chimirri A (2005) Molecular dynamics studies of the full-length integrase-DNA complex. Biochem Biophys Res Commun 336:1010–1016
33. Wang L, Liu C, Chen W, Wang C (2005) Constructing HIV-1 integrase tetramer and exploring influences of metal ions on forming integrase-DNA complex. BiochemBiophys Res Commun 337:313–319
34. Hu J, Wang C (2010) Molecular dynamics simulation of HIV-1 ınıtegrase dimer complexed with viral DNA. Chin J Chem 28:33–40
35. Balasubramanian S, Rajagopalan M, Ramaswamy A (2012) Structural dynamics of full-length retroviral integrase: a molecular dynamics analysis. J Biomol Struct Dyn 29:1163–1174
36. Xue W, Liu H, Yao X (2012) Molecular mechanism of HIV-1 integrase–vDNA interactions and strand transfer inhibitor action: a molecular modeling perspective. J ComputChem 33:527–536
37. Holler TP, Foltin SK, Ye QZ, Hupe DJ (1993) HIV1 integrase expressed in Escherichia coli from a synthetic gene. Gene 136:323–328
38. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. Nucl Acids Res 8:235–242
39. Simmerling C, Strockbine B, Roitberg AE (2002) All-atom structure prediction and folding simulations of a stable protein. J Am Chem Soc 124:11258–11259
40. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, 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 JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA (2004) Gaussian 03, Revision C02. Gaussian Inc, Wallingford CT
41. Case DA, Darden TA, Cheatham TE III, Simmerling CL, Wang J, Duke RE, Luo R, Walker RC, Zhang W, Merz KM, Roberts B, Wang B, Hayik S, Roitberg A, Seabra G, Kolossváry I, Wong KF, Paesani F, Vanicek J, Liu J, Wu X, Brozell SR, Steinbrecher T, Gohlke H, Cai Q, Ye X, Wang J, Hsieh M-J, Cui G, Roe DR, Mathews DH, Seetin MG, Sagui C, Babin V, Luchko T, Gusarov S, Kovalenko A,

Kollman PA (2010) AMBER 11. University of California, San Francisco
42. Jakalian A, Jack DB, Bayly CI (2002) Fast, efficient generation of high-quality atomic charges AM1-BCC model: II Parameterization and validation. J Comput Chem 23:1623–1641
43. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Lee M, Duan TY, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA (2000) Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc Chem Res 33:889–897
44. Lee MC, Duan Y (2004) Distinguish protein decoys by using a scoring function based on a new AMBER force field, short molecular dynamics simulations, and the generalized born solvent model. Proteins 55:620–634
45. Wang JM, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general amber force field. J Comput Chem 25:1157–1174
46. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferin TE (2004) UCSF Chimera–a visualization system for exploratory research and analysis. J Comput Chem 25:1605–1612
47. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellneyer DC, Fox T, Caldwell JW, Kollman PA (1995) A 2nd Generation force-field for the simulation of proteins, nucleic acids, and organic-molecules. J Am Chem Soc 117:5179–5197
48. Darden T, York D, Pedersen L (1993) Particle Mesh Ewald - an NLOG (N) Method for ewald sums in large system. J Chem Phys 98:10089–10092
49. Walker R (2003) The Development of a QM/MM Based Linear Response Method and its Application to Proteins. PhD Thesis, Imperial College Department of Chemistry, London
50. Lin F, Wang R (2010) Systematic Derivation of AMBER Force Field Parameters Applicable to Zinc-Containing Systems. J Chem Theory Comput 6:1852–1870
51. Cahn P, Sued O (2007) Raltegravir: a new antiretroviral class for salvage therapy. Lancet 369:1235–1236
52. Cocohoba J, Dong BJ (2007) Raltegravir: the frist HIV integrase inhibitor. ClinTher 30:1747–1765
53. Marchand C, Johnson AA, Karki RG, Pais GC, Zhang X, Cowansage K, PatelT A, Nicklaus MC, Burke TR, Pommier Y (2003) Metal-dependent inhibition of HIV-1 integrase by β-diketo acids and resistance of the soluble double-mutant (F185K/C280S). Mol Pharmacol 64:600–609
54. Hombrouck A, Van Remoortel B, Michiels M, Noppe W, Christ F, Eneroth A, Sahlberg BL, Benkestock K, Vrang V, Johansson NG, Barreca ML, De Luca L, Ferro S, Chimirri A, Debyser Z, Witvrouw M (2008) Preclinical evaluation of 1H-benzylindole derivatives as novel human immunodeficiency virus integrase strand transfer inhibitors. Antimicrob Agents Chemother 52:2861–2869
55. Alves CN, Marti S, Castillo R, Andres J, Moliner V, Tunon I, Silla E (2007) Calculation of binding energy using BLYP/MM for the HIV-1 integrase complexed witht he S-1360 and two analogues. Bioorg Med Chem 15:3818–3824
56. Di Santo R, Costi R, Roux A, Miele G, Crucitti GC, Iacovo A, Rosi F, Lavecchia A, Marinelli L, Di Giovanni C, Novellino E, Palmisano L, Andreotti M, Amici R, Galluzzo CM, Nencioni L, Palamara AT, Pommier Y, Marchand C (2008) Novel quinolinonyl diketo acid derivatives as HIV-1 integrase inhibitors: design, synthesis, and biological activities. J Med Chem 51:4744–4750
57. Nair V, Uchil V, Chi G, Dams I, Cox A (2007) Biologically-validated HIV integrase inhibitors with nucleobase scaffolds: structure, synthesis, chemical biology, molecular modeling, and antiviral activity. Nucleosides Nucleotides Nucleic Acids 26:665–668
58. Goethals O, Vos A, Von Ginderen M, Geluykens P, Smits V, Schols D, Hertogs K, Clayton R (2010) Primary mutations selected in vitro with raltegravir confer large fold changes in susceptibility to first-generation integrase inhibitors, but minor fold changes to inhibitors with second-generation resistance profiles. Virology 402:338–346
59. Zhao H, Neamati N, Hong H, Mazumder A, Wang S, Sunder S, Milne GWA, Pommier Y, Burke TR Jr (1997) Coumarin-based inhibitors of HIV integrase. J Med Chem 40:242–249
60. Hare S, Vos AM, Clayton RF, Thuring JW, Cummings MD, Cherepanov P (2010) Molecular mechanisms of retroviral integrase inhibition and the evolution of viral resistance. Proc Natl Acad Sci USA 107:20057–20062
61. Zhang X, Neamati N, Lee YK, Orr A, Brown RD, Whitaker N, Pommier Y, Burke TR Jr (2001) Arylisothiocyanate-containing esters of caffeic acid designed as affinity ligands for HIV-1 integrase. Bioorg Med Chem 9:1649–1657
62. Yuan H, Parrill A (2005) Cluster analysis and three-dimensional QSAR studies of HIV-1 integrase inhibitors. J Mol Graph Model 23:317–328
63. Mazumder A, Wang S, Neamati N, Nicklaus M, Sunder S, Chen J, Milne GWA, Rice WG, Burke TR Jr, Pommier Y (1996) Antiretroviral agents as inhibitors of both human immunodeficiency virus type 1 integrase and protease. J Med Chem 39:2472–2481
64. Singh SB, Zink DL, Bills GF, Teran A, Silverman KC, Lingham RB, Felock P, Hazuda DJ (2003) Four Novel Bis-(naphtho-γ-pyrones) Isolated from Fusarium species as inhibitors of HIV-1 integrase. Bioorg Med Chem Lett 13:713–717
65. Singh SB, Jayasuriya H, Dewey R, Polishook JD, Dombrowski AW, Zink DL, Guan Z, Collado J, Platas G, Pelaez F, Felock PJ, Hazuda DJ (2003) Isolation, structure, and HIV-1-integrase inhibitory activity of structurally diverse fungal metabolites. J Ind Microbiol Biotechnol 721–731
66. Sechi M, Carta F, Sannia L, Dallochio R, Dessi A, Al-Safi RI, Neamati N (2009) Design, synthesis, molecularmodeling, and anti-HIV-1 integrase activity of a series of photoactivatable diketo acid-containingi nhibitors as affinity probes. Antivir Res 81:267–276
67. Barreca ML, Rao A, De Luca L, Zappalà M, Gurnari C, Monforte P, De Clercq E, Van Maele B, Debyser Z, Witvrouw M, BriggsJ M, Chimirri A (2004) Efficient 3D data bases creening for novel HIV-1 IN inhibitors. J Chem Inf Comput Sci 44:1450–1455
68. Shiomi K, Matsui R, Isozaki M, Chiba H, Sugai T, Yamaguchi Y, Masuma R, Tomoda H, Chiba T, Yan H, Kitamura Y, Sugiura W, Omura S, Tanaka H (2005) Fungal phenalenones inhibit HIV-1 integrase. J Antibiot (Tokyo) 58:65–68
69. Dubey S, Satyanarayana YD, Lavania H (2007) Development of integrase inhibitors for treatment of AIDS: an overview. Eur J Med Chem 42:1159–1168
70. Jin H, Cai RZ, Schacherer L, Jabri S, Tsiang M, Fardis M, Chen X, Chen JM, Kim CU (2006) Design, synthesis, and SAR studies of novel and highly active tri-cyclic HIV integraseinhibitors. Bioorg Med Chem Lett 16:3989–3992
71. Deng J, Kelley JA, Barchi JJ, Sanchez T, Dayam R, Pommier Y, Neamati N (2006) Miningthe NCI antiviralcompoundsfor HIV-1 integrase inhibitors. Bioorg Med Chem 14:3785–3792
72. Kamal MA, Christopherson RI (2004) Accumulation of 5-phosphoribosyl-1-pyrophosphate in human CCRF-CEM leukaemia cells treated with antifolates. Int J Biochem Cell Biol 36:545–551
73. Mai LP, Guéritte F, Dumontet V, Tri MV, Hill B, Thoison O, Guénard D, Sévenet T (2001) Cytotoxicity of Rhamnosylanthraquinones and Rhamnosylanthrones from *Rhamnus nepalensis*. J Nat Prod 64:1162–1168
74. Lang PT, Moustakas D, Brozell S, Carrascal N, Mukherjee S, Pegg S, Raha K, Shivakumar D, Rizzo R, Case DA, Shoichet B, Kuntz I (2007) DOCK 61. University of California, San Francisco, http://dockcompbioucsfedu
75. Kollman PA, 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 DA,

Cheatham TE III (2000) Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc Chem Res 33:889–897
76. Srinivasan J, Cheatham TE III, Cieplak P, Kollman PA, Case DA (1998) Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices. J Am Chem Soc 120:9401–9409
77. Honig B, Nicholls A (1995) Classical electrostatics in biology and chemistry. Science 268:1144–1149
78. Sitkoff D, Sharp KA, Honig B (1994) Accurate calculation of hydration free energies using macroscopic solvent models. J Phys Chem 98:1978–1988
79. Lins RD, Briggs JM, Straatsma TP, Carlson HA, Greenwald J, Choe S, McCammon JA (1999) Molecular dynamics studies on the HIV-1 integrase catalytic domain. Biophys J 76:2999–3011
80. Vajragupta O, Boonchoong P, Morris GM, Olson AJ (2005) Active site binding modes of curcumin in HIV-1 protease and integrase. Bioorg Med Chem Lett 15:3364–3368
81. Healy EF, Sanders J, King PJ, Robinson WE Jr (2009) A docking study of L-chicoric acid with HIV-1 integrase. J Mol Graph Model 27:584–589
82. Sippel M, Sotriffer CA (2010) Molecular dynamics simulations of the HIV-1 ıntegrase dimerization ınterface: guidelines for the design of a novel class of ıntegrase ınhibitors. J Chem Inf Model 50:604–614
83. Huanga M, Grant GH, Richards WG (2011) Binding modes of diketo-acid inhibitors of HIV-1 integrase: a comparative molecular dynamics simulation study. J Mol Graph Model 29:956–964
84. Moss DM, Siccardi M, Murphy M, Piperakis MM, Khoo SH, Back DJ, Owen A (2012) Divalent metals and pH alter raltegravir disposition ın vitro. Antimicrob Agents Chemother 56:3020–3026
85. Engelman A, Mizuuchi K, Craigie R (1991) HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer. Cell 67:1211–1221