

# Structural dynamics of native and V260E mutant C-terminal domain of HIV-1 integrase

Balasubramanian Sangeetha · Rajagopalan Muthukumaran ·  
Ramaswamy Amutha

Received: 9 September 2014 / Accepted: 6 January 2015 / Published online: 14 January 2015  
© Springer International Publishing Switzerland 2015

**Abstract** The C-terminal domain (CTD) of HIV-1 integrase is a five stranded  $\beta$ -barrel resembling an SH3 fold. Mutational studies on isolated CTD and full-length IN have reported V260E mutant as either homo-dimerization defective or affecting the stability and folding of CTD. In this study, molecular dynamics simulation techniques were used to unveil the effect of V260E mutation on isolated CTD monomer and dimer. Both monomeric and dimeric forms of wild type and V260E mutant are highly stable during the simulated period. However, the stabilizing  $\pi$ -stacking interaction between Trp243 and Trp243' at the dimer interface is highly disturbed in CTD-V260E ( $>6$  Å apart). The loss in entropy for dimerization is  $-30$  and  $-25$  kcal/mol for CTD-wt and CTD-V260E respectively signifying a weak hydrophobic interaction and its perturbation in CTD-V260E. The mutant Glu260 exhibits strong attraction/repulsion with all the basic/acidic residues of CTD. In addition to this, the dynamics of CTD-wild type and V260E monomers at 498 K was analyzed to elucidate the effect of V260E mutation on CTD folding. Increase in SASA and reduction in the number of contacts in CTD-V260E during simulation highlights the instability caused by the mutation. In general, V260E mutation affects both multimerization and protein folding with a pronounced effect on protein folding rather than multimerization. This study emphasizes the importance of the hydrophobic nature and SH3 fold of CTD in proper functioning of HIV integrase and perturbing this nature would be a rational approach toward designing more selective and potent allosteric anti-HIV inhibitors.

**Keywords** Integrase · C-terminal domain · Hydrophobic · V260E · SH3 · Unfolding

## Introduction

Integrase (IN) is a vital protein in HIV life cycle which catalyzes the integration of reverse transcribed viral DNA into the host DNA. Integration is a two-step reaction, where IN removes GT dinucleotide from the conserved CAGT segment of viral DNA terminals, to expose the 3'-OH groups of CA dinucleotide (3'-processing) and the subsequent nucleophilic attack by the exposed 3'-hydroxyl on the phosphodiester backbone of host DNA (strand transfer) [1–6]. IN tetramer is responsible for the insertion of both the ends of viral DNA into host DNA with 5 base pair separation [7–10]. Biochemical and mutagenesis studies have shown that IN possesses three functionally distinct domains [11–14] such as N-terminal domain (NTD) [15], catalytic core domain (CCD) [12] and C-terminal domain (CTD) [16]. The NTD (amino acids 1–45) has helix-turn-helix and HHCC (His12, His16, Cys40 and Cys43) Zinc binding motifs. NTD binds to a  $Zn^{2+}$  ion and plays important role in multimerization, viral DNA binding and host cell co-factor interaction [17]. CCD (amino acids 55–212) contains the active site formed by the catalytic triad D-D-(35)-E (Asp64, Asp116 and Glu152) which coordinates two  $Mg^{2+}$  ions and catalyzes both 3'-processing and strand transfer reactions [18, 19]. CTD (220–288) is the least conserved domain of IN and is primarily involved in the augmentation of IN multimerization, non-specific DNA binding and host cell co-factor binding [20, 21]. These three functional domains are linked by flexible linkers which renders increased conformational flexibility to integrase.

B. Sangeetha · R. Muthukumaran · R. Amutha (✉)  
Centre for Bioinformatics, School of Life Sciences, Pondicherry University, Puducherry 605014, India  
e-mail: ramutha@bicpu.edu.in; amutha\_ramu@yahoo.com

All three domains of IN are necessary for tetramerization and concerted integration. The conformational flexibility and insoluble nature of IN are the major reasons for the lack of crystal structures of full-length IN [22]. Despite this flexibility and insolubility, structures of individual domains and two-domain constructs (NTD–CCD [23] and CCD–CTD [24]) are determined by X-ray diffraction as well as NMR techniques and are being used by several scientific groups to explore the structural and functional dynamics of IN [23, 25–30]. The recently solved structure of PFV intasome provides clear picture on the organization of IN tetramer during the concerted integration of viral and host DNA [31]. The IN tetramer possesses inner and outer monomeric units in which the NTD of one monomer and CCD of another monomer forms the dimer interface between the inner monomers, and the CCDs of inner and outer monomers form another dimer interface. The CTDs of both the inner monomers involve in viral DNA binding. Though PFV intasome is the only available structure of a complete IN tetramer with DNA, it fails to map the NTDs and CTDs of the two outer monomers. A study by Bojja et al. [32, 33] has reported a different conformation of IN dimer in the absence of DNA, in which, the NTDs and CTDs form the dimer interface with a stabilizing stacking interaction between Trp243 from both monomers while CCDs are reported to involve only in tetramerization. The major difference between dimer of PFV inner monomers and HIV dimer structures is the orientation of CTDs. It has been reported that CTD exhibits high conformational flexibility with major structural rearrangements during DNA binding [34]. Also, the CCD–CTD linker is highly flexible and exists as (1) coil in PFV, ASV and HIV tetramer structures and (2) helix in CCD–CTD two-domain structure. Hence, the spatial arrangements of CTD determined by the flexibility of CCD–CTD linker are highly important in HIV life cycle.

Mutagenesis studies have revealed two types of mutations in integrase, classified under class I and class II [35]. Mutations that affect HIV-1 replication by blocking integration are categorized as class I. On the other hand, class II mutations affect HIV-1 replication by blocking the non-enzymatic activity of IN such as interaction with reverse transcriptase and other cellular co-factors causing defects in nuclear localization of the pre-integration complex, virus assembly and release. Since, mutations at the catalytic triad blocks the catalytic activity of IN, it is considered as class I mutant [36, 37]. The co-ordination of  $Zn^{2+}$  ion by the conserved HHCC motif maintains the structural integrity of NTD [15, 38] and hence mediates its interaction with reverse transcriptase and LEDGF [39–41]. Mutations at this conserved HHCC motif are considered as class II since it affects these interactions and further blocks HIV-1 life-cycle [42, 43]. Several class I and II mutations have been

reported in all three domains of IN and several studies have analyzed the effect of these mutations on the structure and functions of HIV-1 IN [44].

The structure of isolated CTD (220–270) that resembles the topology of a SH3 domain has been reported by NMR studies [16, 45, 46]. The structure of C-terminal 18 residues (271–288) is not yet reported due to its increased flexibility rendered by the presence of charged residues. Truncation studies have highlighted the importance of this C-terminal tail (C-tail) region in IN activity. Mutational study by deletion of 3 amino acids has identified the effect of HIV<sub>1–270</sub> to be similar to a class II mutant phenotype [47]. Report on single amino acid truncation studies on C-terminal 28 residues (260–288) revealed gradual loss of infectivity. IN truncations beyond 269 (specifically up to 260, 265 and 268) were classified as class II mutants since these mutants were unable to produce significant levels of reverse-transcribed DNA and resembles the phenotype of a CTD-deletion mutant (HIV<sub>1–212</sub>) [48]. SH3 domains express *cis-trans* functioning when present in homodimers. This property is inherited by HIV-1 IN CTD which is also functional in both *cis* and *trans* conformations with CCD. It has been highlighted that the SH3 fold is important for proper poly-protein processing and viral assembly [49]. IN truncations beyond 268 disturbs the SH3 fold and hence affects IN functions and thereby the assembly of intasome [48].

CTD is reported to be necessary and sufficient for the binding of integrase with reverse transcriptase (RT) [50, 51]. The residues Arg231, Leu242, Trp243, Gly247, Ala248, Val250, Ile251, Glu252, and Lys258 of CTD are responsible for interaction with the fingers-palm domain (residues 1–242) and connection domain (residues 387–422) of RT [50]. This RT binding interface of CTD is similar to its homo-dimerization interface. Studies on isolated CTD and full-length IN revealed the importance of several residues of CTD domain. The residues Val234, Glu246, Lys258, Arg262, Lys264 and Lys273 have been reported to be involved in DNA binding [52–55] while the residues Leu241, Leu242, Val260, Lys266 and Arg269 involve in multimerization [16, 24, 45, 54, 56]. Trp235 mutant HIV-1 IN is identified as a class I replication defective mutant since it fails to bind reverse transcriptase [57]. V260E mutation is the only mutant protein identified by yeast-two hybrid system while screening for IN mutants defective in homodimerization [56]. The V260E mutant was reported to affect tetramerization of IN rather than the dimerization since minimal non-concerted 3'-processing and DNA strand transfer activities were reported in vitro. The residue V260 is highly conserved and it resides in the second most conserved region (259–VVPRRK–264) of CTD [58]. Also, the mutation of hydrophobic Val residue (buried in a hydrophobic core) to a negatively charged Glu residue is hypothesized to affect the stability of SH3 fold and hence the entire folding [16, 54]. The importance of CTD and the

intriguing effects of class II mutant V260E on the function of IN prompted to study the role of V260E mutation in dimerization and folding of isolated CTD.

Molecular dynamics simulations have been used extensively to explore the structural and functional properties of various biomolecules. In recent years, molecular dynamics studies have been well applied in exploring the structural dynamics of HIV proteins such as protease, reverse transcriptase and integrase [25, 59–63]. In particular, the influence of IN-CCD mutations such as G140A/G149A, Thr66, Ser153, Met154, etc. on IN function and drug resistance have also been explored using molecular dynamics simulations [64, 65]. Studying protein folding by molecular dynamics simulations is a highly challenging one, since it is a random and time-taking process. The study of protein folding by unfolding at high temperatures is well reported in folding pathway analysis of proteins such as Chymotrypsin inhibitor 2 [66], engrailed homeodomain [67], cell-cycle protein p13suc1 [68], hen egg white lysozyme [69], ubiquitin [70], human prion protein domain [71] and bovine pancreatic trypsin inhibitor [72]. In this study, molecular dynamics simulations have been performed on wild type and mutated (V260E) CTDs in both monomeric and dimeric forms to understand the effect of V260E mutant on IN function. The knowledge acquired from this study could lead to the development of novel allosteric inhibition mechanism against IN targeting CTD.

## Materials and methods

The structure of CTD (residues 221–270) was extracted from the CCD–CTD two-domain structure (PDB ID: 1EX4 [24]). As the C-tail (residues 271–288) is necessary for the functional dynamics of full-length IN, CTD with C-tail was modeled in a random coiled geometry as adopted in our previous study [25]. In the previous study, restrained molecular dynamics simulation was performed on IN with C-tail, in which the C-tail was allowed to move freely without any restraints to attain a least energy conformation [25]. The least energy conformation of CTD monomer with C-tail was used further for this study. Using this CTD monomer, the dimer of CTD was modeled as observed by NMR studies (PDB ID: 1QMC [46]). The monomer and dimer conformations of both wild type and V260E mutant CTD were subjected to molecular dynamics simulation using AMBER 11 package [73] with ff99SB protein force field [74] to study the effect of mutation on multimerization/folding of CTD.

### Molecular dynamics simulations

The wild type and V260E mutant structures of CTD were solvated in a cubic box of TIP3P [75] water model extending up to 10 Å on all sides of the protein. All the

systems were neutralized by adding Cl<sup>−</sup> counter-ions. The modeled system was relaxed in two phases in which, the solvent molecules along with counter-ions were relaxed first followed by the relaxation of the entire system including the protein. The short-range electrostatic and van der Waals interactions were truncated with a cut-off distance of 10 Å and the long-range electrostatics were treated with a particle-mesh Ewald procedure. All bonds involving hydrogens were constrained using the SHAKE algorithm. The systems were gradually heated to 300 K over 50 ps at constant volume with harmonic position restraints of 10 kcal/mol/Å<sup>2</sup> on the protein. The density of the system was relaxed by maintaining the pressure at 1 atm and the harmonic restraints on the protein were gradually relaxed over 80 ps. Simulation at constant pressure for a period of 500 ps was performed to equilibrate the system followed by a production simulation for 50 ns.

The role of V260E mutation on CTD folding was studied by unfolding the wild type and mutant CTDs by introducing high temperatures. In this study, the readiness for unfolding by wild type and V260E mutant CTD was monitored to elucidate the role of V260E mutation on folding. The relaxed CTD monomer of both wild type and mutant forms were heated to 498 K including harmonic restraints to maintain backbone chiralities [76], followed by the regular protocol as explained above. The unfolding simulations were performed for a period of 20 ns. Since protein folding is a random process, the starting conformation and environment might influence the unfolding behavior. The unfolding behavior of CTD-wt and CTD-V260E structures was further validated using conformational sampling. The structures extracted at an interval of 5 ns from the 50 ns simulation trajectories at 300 K were subjected to multiple independent simulations at 498 K. A total of 12 simulations comprising structures from (1) initial, (2) equilibrated and (3) 10 intermediate structures at an interval of 5 ns, were performed for a period of 20 ns. From the resulting trajectories for both CTD-wt and CTD-V260E, the sampled conformations were clustered together based on distance-RMSD and a minimum cut-off distance of 4 Å between clusters.

### MM/GBSA binding energy analysis

The binding energy of CTD dimers was analyzed by MM/GBSA [77] method as implemented in AMBER 11. Snapshots at 100 ps interval were extracted from the last 15 ns of the 50 ns simulation where the RMSD has been stabilized and the binding free energy was computed as an average of these snapshots using the following expression:

$$\Delta G_{bind} = \Delta G_{complex} - (\Delta G_{receptor} + \Delta G_{ligand})$$

where,

$$\Delta G = \Delta E_{mm} + \Delta G_{sol} - T\Delta S$$

$$\Delta E_{mm} = \Delta E_{int} + \Delta E_{ele} + \Delta E_{vdW}$$

$$\Delta E_{int} = \Delta E_{bond} + \Delta E_{angle} + \Delta E_{dihedral}$$

$$\Delta G_{sol} = \Delta G_{pol} + \Delta G_{nonpol}$$

$$\Delta G_{nonpol} = \gamma SASA + \beta$$

The terms  $\Delta G_{complex}$ ,  $\Delta G_{receptor}$ , and  $\Delta G_{ligand}$  represent the average free energies of the complex, receptor and ligand respectively. In this study, one of the CTD monomers is considered as receptor and the other as ligand for the calculation of binding free energy.  $\Delta E_{mm}$  is the gas phase (vacuum) molecular mechanics energy and is calculated as the sum of internal (bond, angle, dihedral), electrostatic and vdW energies. The solvation energy,  $\Delta G_{sol}$  includes polar  $\Delta G_{pol}$  and non-polar  $\Delta G_{non-pol}$  terms. The polar component is calculated using Generalized Born approximation with a grid spacing of 0.5 Å and the dielectric constants, 1.0 and 80.0 for solute and solvent, respectively. The non-polar component is obtained by calculating the solvent accessible surface area using a probe of radius 1.4 Å and linear combination of pairwise overlaps (LCPO) approximation. The entropy of the system,  $\Delta S$  is calculated by harmonic approximation using MM/GBSA-NMODE program implemented in AMBER 11. The binding energy components of each residue pair was calculated by MMGBSA binding energy decomposition method and the interaction energies of Val260/Glu260 with all other residues were analyzed.

## Results and discussion

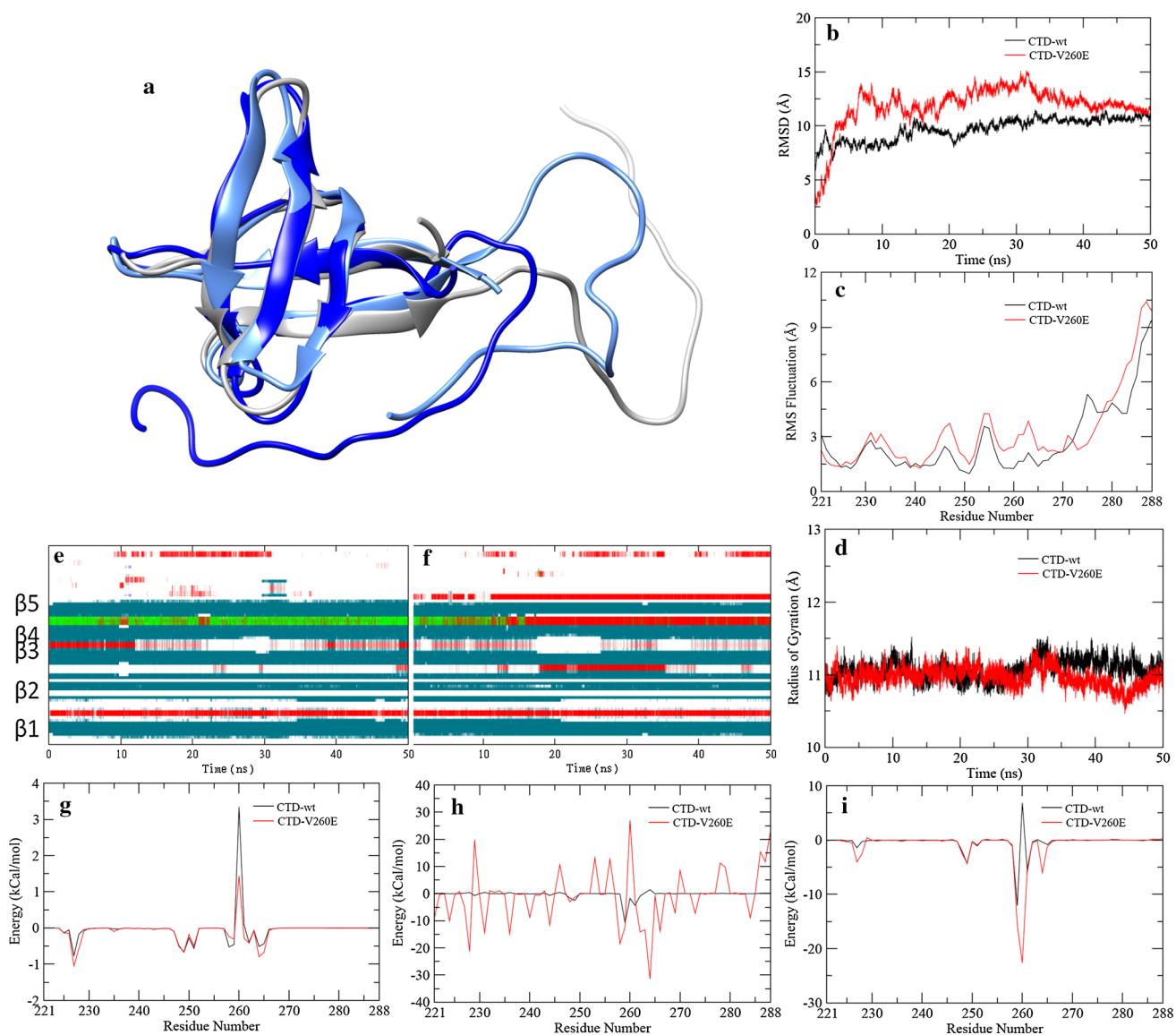
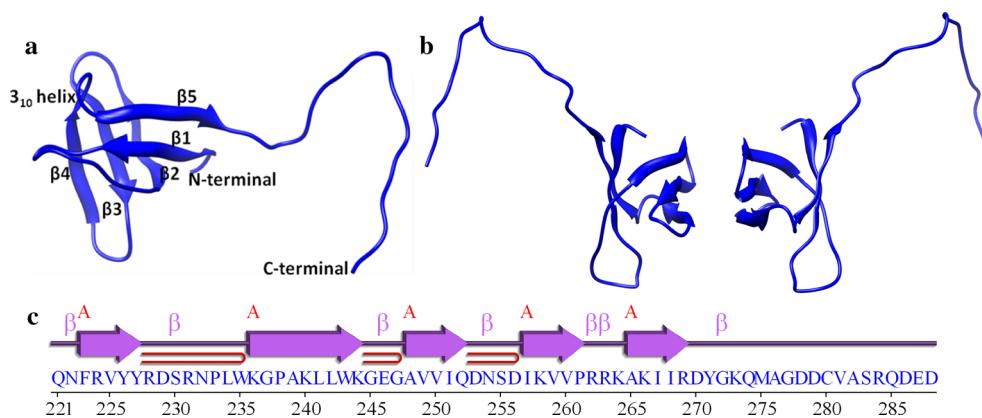
CTD monomer is a five-stranded  $\beta$ -barrel formed by two anti-parallel  $\beta$ -sheets and resembles SH3 fold [16, 46] (Fig. 1a, c). These five  $\beta$ -strands are formed by the residues 223–228 ( $\beta$ 1), 235–244 ( $\beta$ 2), 249–252 ( $\beta$ 3), 257–260 ( $\beta$ 4) and 265–269 ( $\beta$ 5), respectively. In various structures of CTD,  $\beta$ 2 starts anywhere in between residues 232 and 236 and extends up to 244. The  $\beta$ 4 and  $\beta$ 5 sheets are connected by a  $\beta$ -helix formed by residues 262–264. The core of  $\beta$ -barrel is hydrophobic in nature and is held together by Val225, Tyr227, Ala239, Leu241, Val249, Ile251, Val260 and Ala265 residues. The  $\beta$ -sheets 2, 3 and 4 of two CTD monomers orient in an anti-parallel fashion to form the CTD–CTD dimer interface (Fig. 1b) which is also similar to the CTD–CTD interface observed in full length ASV and HIV-1 structures [32, 33]. The hydrophobic surface of the dimer interface comprises the residues Leu242, Trp243, Ala248, Val250, Ile257 and Val259 from both the monomers. At the dimer interface, H-bonding interaction is also reported between the side chains of Glu246 and Gln252'

and vice versa. The residue Val260 is located at the C-terminal end of  $\beta$ 4 (Fig. 1a), which is equivalent to the 2nd ligand binding site in a regular SH3 domain [49]. Though the neighboring residues Lys258, Arg262 and Lys264 are associated in DNA binding, Val260 does not involve in direct interaction with DNA. Since Val260 is a part of the hydrophobic core, it plays a vital role in maintaining the structural integrity of CTD and multimerization. A pair of H-bonds observed between the two conserved residues of CTD- $\beta$ 3 (Val249) and  $\beta$ 4 (Val260) are also reported as important for the structural stability of CTD [58]. In order to study the importance of Val260 in IN function, molecular dynamics simulation studies of wild type and Val to Glu (V260E) mutated CTD in monomer and dimer forms are performed in this study.

### Dynamics of wild type and V260E mutant CTD in monomer form

The wild type and V260E mutant CTD were simulated for a period of 50 ns to understand their conformational stability. Superimposition of the initial and 50 ns simulated structures of CTD-wt and CTD-V260E revealed stable dynamics of the  $\beta$ -barrel and a flexible C-tail (Fig. 2a). The RMSD of wild type CTD (hereafter named as CTD-wt) fluctuates around 9.5 Å while that of V260E mutant CTD (hereafter named as CTD-V260E) shows increased conformational flexibility with RMSD ranging at  $11.8 \pm 2$  Å (Fig. 2b). The RMSD of CTD-V260E increases gradually from 3.6 to 11.7 Å during the course of the simulation. The highly dynamic CTDs are observed mainly due to the flexibility of C-tail and the loop regions connecting the  $\beta$ -sheets (Fig. 2c) where CTD-V260E is comparatively more flexible than CTD-wt. The stability of the hydrophobic core holding the  $\beta$ -barrel was analyzed in terms of the radius of gyration (Fig. 2d) and is observed to be highly stable in both CTD-wt and V260E. The C-tails of both CTD-wt and V260 are highly flexible and express free dynamics. In CTD-V260E, the distance between the N- and C-terminal residues varies considerably around  $38.0 \pm 6.8$  Å. Whereas in CTD-wt the N- and C-terminal residues are closer ( $18.6 \pm 2.8$  Å) with more intra-tail H-bonds than CTD-V260E. In CTD-V260E, the C-tail aligns in an anti-parallel manner to  $\beta$ 5 and buries the DNA binding residues Lys244 and Arg262 by forming H-bonds between Arg262 ( $\beta$ -helix) and Asp278, Asp279 (C-tail). The stability of the secondary structures of CTD-wt and CTD-V260E was analyzed and shown in Fig. 2e, f respectively. In CTD-wt, the  $\beta$ -helix is highly stable whereas it uncoils to form a helical turn in CTD-V260E. Also, the loop region connecting  $\beta$ 3 and  $\beta$ 4 is slightly perturbed in CTD-V260E. In general, irrespective of the mutation, the structural integrity of SH3 fold is maintained in both CTD-wt and V260E

**Fig. 1** Structure of HIV-1 integrase C-terminal domain in monomer (**a**) and dimer (**b**) forms. The residues forming the 5  $\beta$ -strands of the CTD  $\beta$ -barrel is shown in **c**



**Fig. 2** MD simulation analysis of CTD-wt and V260E monomer. Superimposition of initial (gray) and 50 ns simulated structures of CTD-wt (light blue) and CTD-V260E (dark blue) is shown in **a**. RMSD (**b**), RMS fluctuation (**c**), radius of gyration (**d**), pairwise interaction energy of V260/Glu260 decomposed into vdw (**g**),

electrostatic (**h**) and total (**i**) energy of CTD-wt (blue) and CTD-V260E (red) simulations are shown. **e, f** Depicts the variation in the secondary structure (dark cyan for  $\beta$ -sheets, green for  $3_{10}$ -helix and red for turns) of CTD-wt and CTD-V260E respectively

structures. The residues Val260 (in CTD-wt) and Glu260 (in CTD-V260E) express stable H-bonding interaction with Val249 during the course of simulation. An additional H-bond between Glu260 and Arg228 is also observed.

The pairwise interaction energies between Val 260 (Glu260) and the entire CTD were analyzed from the last 15 ns of the simulation and the interaction energy was decomposed into its individual components (Fig. 2g–i). In wild type, only the residues Val249, Val259 and Pro261 express significant favorable electrostatic contributions. In CTD-V260E, all the basic residues significantly contribute for favorable interaction, while the acidic residues involve in unfavorable interactions. This reveals that mutation of a hydrophobic residue into a polar charged residue (V260E) alters the electrostatic behavior of CTD and creates a charged cluster on the DNA binding face of the CTD barrel (by the residues Glu260, Arg228, Lys258, Lys236, Arg263 and Lys264). Even though, CTD  $\beta$ -barrel in SH3 fold is stable during the dynamics, such increased attractive/repulsive nature of CTD is expected to obstruct DNA binding or the initial formation of  $\beta$ -barrel [54].

#### Dynamics of wild type and V260E mutant CTD in dimer form

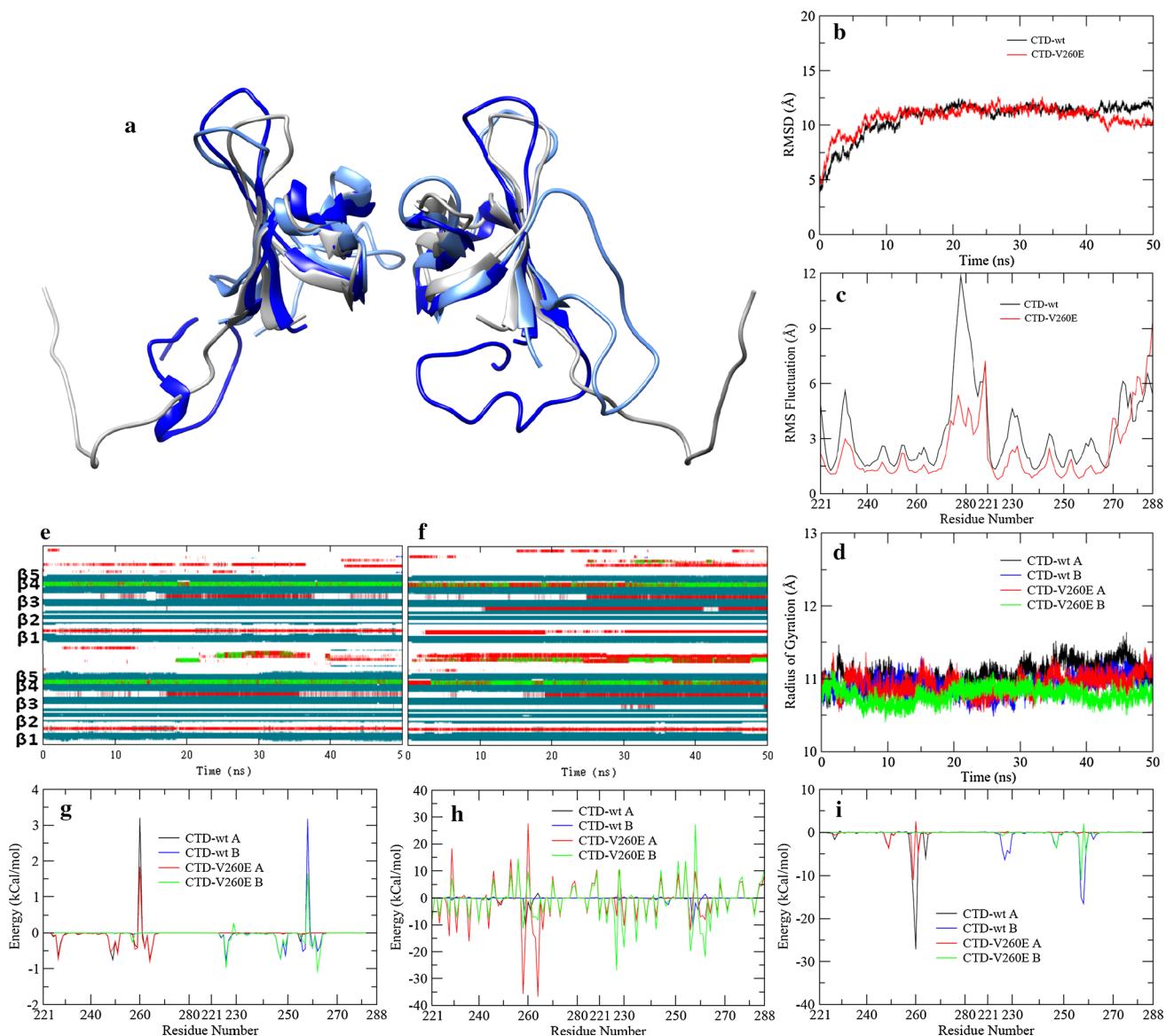
V260E mutation was identified as a multimerization negative mutant by a yeast two-hybrid screening study. Hence, the effect of V260E mutant on multimerization was studied by simulating the wild type and V260E mutant forms of CTD dimer for a period of 50 ns and the superimposition of initial and simulated conformations is shown in Fig. 3a. During the simulation period, the RMSD of both CTD-wt and CTD-V260E dimers increase to  $\sim 12$  Å and remains stable thereafter. The atomic fluctuation of  $C\alpha$  positions from the average conformation is considerably reduced in CTD-V260E than CTD-wt. Also, the radius of gyration of the  $\beta$ -barrel depicting the stability of the hydrophobic core reveals that CTD-V260E dimer is slightly more compact than the CTD-wt dimer. The variation in the secondary structure observed during the simulation of both CTD-wt and V260E was analyzed (Fig. 3e, f respectively). The secondary structures in both CTDs are well maintained during dynamics except the uncoiling of  $3_{10}$ -helix in CTD-V260E. The H-bonding interaction between Val260 (Glu260) and Val249 which plays a key role in maintaining the structural integrity was stable in both CTD-wt and V260E. Additional H-bonds are formed between Glu260' and Arg228', Asp229' and Ser240' during the simulation. Such additional H-bonding interactions are not observed for Val260.

The binding energy and stability of both the systems were analyzed using MM/GBSA and NMODE methods and are tabulated in Table 1. The total entropy of both

CTD-wt and CTD-V260E monomers are similar ( $\sim 842$  kcal/mol). The loss in entropy ( $\Delta S$ ) for the dimerization of CTD-wt and CTD-V260E was calculated as  $-30$  and  $-25$  kcal/mol respectively. These negative entropies signify the existence of weak hydrophobic interaction at the dimer interface. The increase in entropy of CTD-V260E signifies the instability in both the dimer interface as well as the hydrophobic nature. The vdW energy component of the binding energy of both CTD dimers is  $-39$  kcal/mol. The electrostatic energy is more favorable in CTD-wt ( $-16$  kcal/mol) than CTD-V260E ( $+12$  kcal/mol) and is vice versa for the polar component of solvation energy. Hence, the net effect of electrostatic is similar in both CTD-wt and CTD-V260E ( $\sim 32$  kcal/mol). The  $\Delta G$  value of CTD-wt and CTD-V260E are calculated as  $-11.6 \pm 3$  and  $-13.6 \pm 3$  kcal/mol respectively. The high negative  $\Delta G_{\text{Gas}}$  value indicates energetically favorable conformation of CTD-wt in gas phase and the high positive  $\Delta G_{\text{sol}}$  denotes the unfavorable state due to its hydrophobic nature. It is also expected that, the buried state of CTD in the full-length IN tetramer might further decrease the  $\Delta G_{\text{sol}}$  of CTD-wt and hence favor multimerization. Increase in charge based interactions in full-length IN and the presence of DNA may also act as an additional factor in destabilizing the pre integration complex by V260E mutation.

The pairwise residue interaction energies between the two monomers were calculated and are shown in Fig. 3g–i. The electrostatic and polar interaction energies are negligible in CTD-wt whereas they play a major role in CTD-V260E. The intra-chain interaction energies are more pronounced than inter-chain interaction energies. In CTD-V260E, a strong attractive interaction is observed between Glu260 and basic residues such as Arg228, Arg231, Lys236, Lys244, Lys258, Val259, Arg262, Arg263, Lys264, Lys266, Arg269, Arg231' and Lys258'. In contrast to this, acidic residues such as Asp229, Glu246, Asp253, Asp256, Asp270, Asp288, Asp229' and Asp256' express strong repulsive interaction with Glu260. The only strong inter-chain interaction is observed between Glu260' and Asp256. Among these strongly interacting residues, Lys244 (near Leu242 and Trp243), Asp256, Lys258, Val259, Lys266 and Arg269 are reported as important for dimerization. These residues express strong intra-monomer interactions and might be a key factor in disturbing the stability of the dimer interface.

The hydrophobic dimer interface of CTD is lined by the residues Leu242, Trp243, Ala248, Val250, Ile257 and Val259 of both CTD monomers (Fig. 4a). The  $R_g$  values of these residues from both monomers were calculated to analyze the stability of the dimer interface and are plotted in Fig. 4b. In the CTD-wt, the  $R_g$  of dimer interface shows a marked increase in the initial phase of simulation (until 7 ns) and thereafter stabilizes to  $\sim 6.5$  Å. On the other hand, the dimer interface in CTD-V260E is highly stable during the



**Fig. 3** MD simulation analysis of CTD-wt and V260E dimer. Superimposition of initial (gray) and 50 ns simulated structures of CTD-wt (light blue) and CTD-V260E (dark blue) is shown in **a**. RMSD (**b**), RMS fluctuation (**c**), radius of gyration (**d**), pairwise interaction energy of V260/Glu260 decomposed into vdw (**g**),

electrostatic (**h**) and total (**i**) energy of CTD-wt (blue/black) and CTD-V260E (red/green) simulations are shown. **e, f** Depicts the variation in the secondary structure of CTD-wt and CTD-V260E respectively

simulated period. The  $\pi$ -stacking interaction between Trp243 of both the CTD monomers is responsible for maintaining the dimer interface and hence its stability was analyzed by calculating the distance between the centroids of both indole side chains. It is observed from Fig. 4c that the  $\pi$ -stacking interaction is absent in both the dimers during the initial phase of simulation and after  $\sim 30$  ns of simulation, the Trp243 residues in CTD-wt express an interaction within  $\sim 4.5$  Å which lies well in the characteristic  $\pi$ -stacking distance. Whereas, in CTD-V260E dimer interface, these two Trp243 residues lie at a distance more than 6 Å

apart. All these analyses reveal less significant effects of V260E mutation on CTD dimerization by (1) increasing intra-domain interactions and (2) perturbing the hydrophobic nature of CTD.

Unfolding of wild type and V260E mutant forms of CTD

Lutzke et al. [54] have hypothesized that V260E mutation in isolated CTD and full-length IN may affect the folding of IN. In our study, V260E mutation does not show

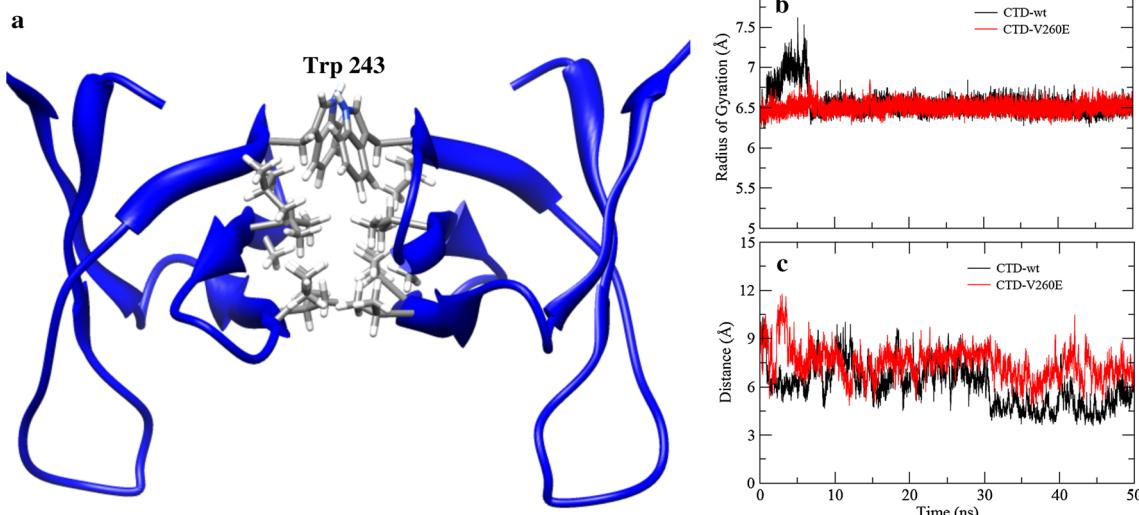
**Table 1** Entropy and binding energy components calculated by NMODE and MMGBSA methods for CTD-wt and CTD-V260E dimers

Energy component	CTD-wt	CTD-V260E
$\Delta E_{vdW}$	−38.4	−39.9
$\Delta E_{elec}$	−16.5	12.2
$\Delta E_{1-4} \text{ vdw}$	0	0
$\Delta E_{1-4} \text{ elec}$	0	0
$\Delta E_{GB}$	48.9	20.2
$\Delta E_{SURF}$	−5.7	−5.9
$\Delta G_{\text{gas}}$	−54.9	−27.8
$\Delta G_{\text{solv}}$	43.4	14.2
$\Delta G_{\text{bind}}$	−11.56	−13.55
Entropy	−30.87	−25.77

pronounced influence on multimerization and hence its role in CTD folding was analyzed by temperature induced unfolding, a procedure well reported in the literature to provide useful insights into protein folding mechanism [66–72]. The wild type and V260E mutant CTD structures were heated to 498 K over a period of 1 ns followed by a production simulation of 20 ns. Figure 5a depicts the superimposition of initial and simulated conformations of CTD-wt and V260E which reveals marked distortions in the CTD-V260E structure. The RMSD of CTD-wt is highly stable at 7 Å as shown in Fig. 5b. The CTD-V260E is highly dynamic and the RMSD varies from ~4 to 10 Å during the initial 3 ns and later fluctuates around 10 Å. The C $\alpha$  atoms of the entire CTD with V260E mutant expresses increased positional fluctuations than the CTD-wt

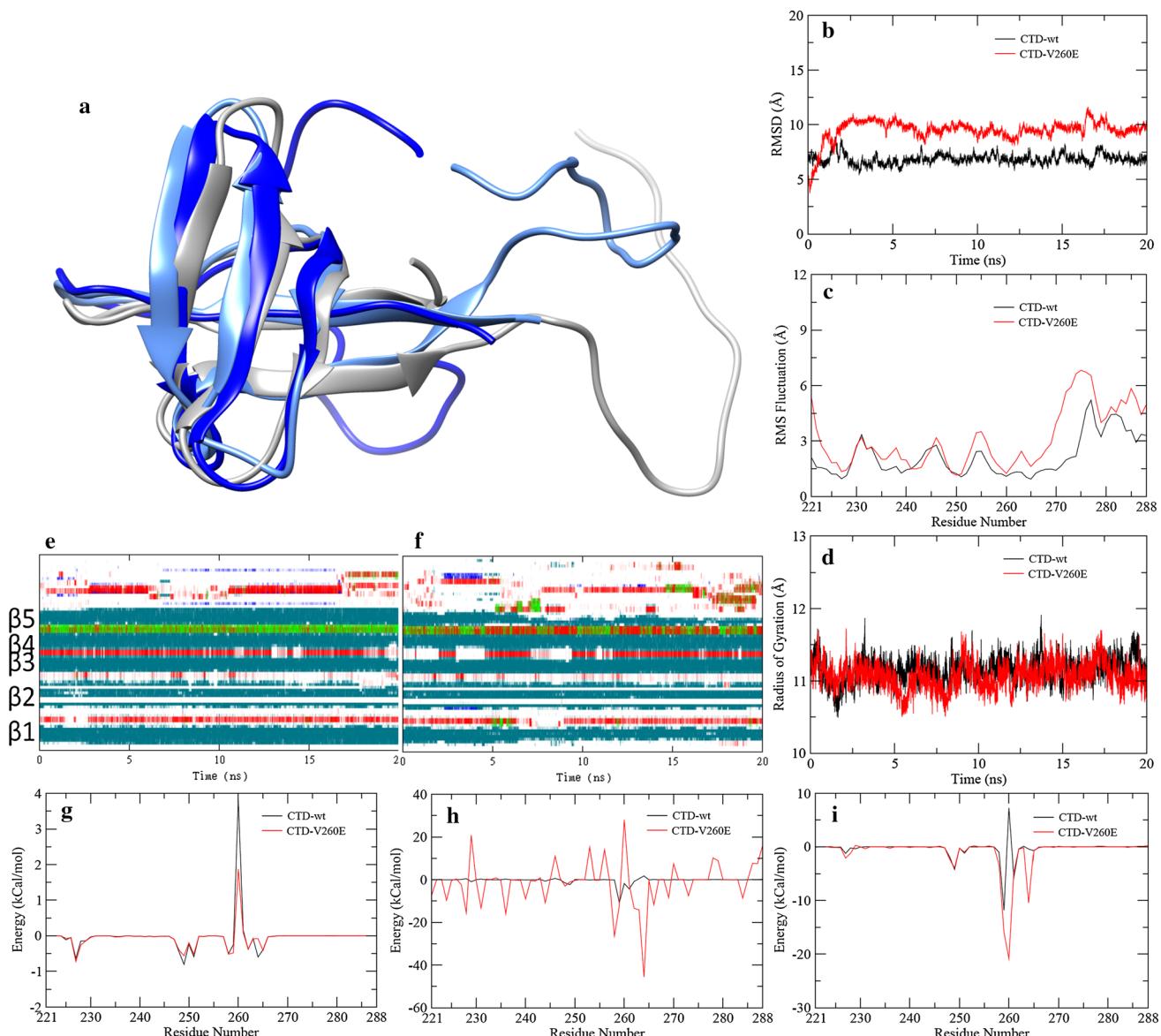
(Fig. 5c). The radius of gyration of the  $\beta$ -barrel is highly stable ( $11 \pm 0.2$  Å) in the CTD-wt whereas it is fluctuating around  $11 \pm 0.8$  Å in the CTD-V260E. The hydrophobic core of  $\beta$ -barrel formed by the residues Val225, Tyr227, Ala239, Leu241, Val249, Ile251, Val260 and Ala265 is however stable in both CTD-wt and CTD-V260E simulations. The C-tail conformation is highly dynamic in CTD-V260E and interacts with the  $\beta$ -barrel. On the other hand, the C-tail in CTD-wt is highly stable and maintains a constant head-to-tail distance of 10 Å.

The number of residues within 7 Å was calculated over the simulation period to analyze the conformational changes due to temperature induced unfolding of CTD. Prior to simulation, the number of residues within 7 Å was similar in both CTD-wt and V260E (~384). After 20 ns simulations, the number of contacts in CTD-wt increased up to ~457 while in CTD-V260E, it was observed as ~437 (Fig. 6a). The residue–residue contact maps generated for the starting conformation and 20 ns simulated conformation shows clear changes in the structural integrity of the SH3 fold of CTD-V260E [Fig. 6c(i–iii)]. The length of  $\beta$ -sheets 1 and 2 [labeled as region I and II in Fig. 6c(i)] is reduced in CTD-V260E. Region 3 representing the contact between  $\beta$ 1 and  $\beta$ 5 is dense in CTD-wt represent the stable interaction while it is diminished in CTD-V260E. New contacts between  $\beta$ 2 and C-tail are observed to form in CTD-V260E (region IV). Also, intra-tail contacts are increased in both the simulations (marked as region V). The solvent accessibility of CTD-V260E is also increased up to  $4,096 \pm 90$  Å<sup>2</sup> during the simulation period while it is less ( $4,054 \pm 90$  Å<sup>2</sup>) in CTD-wt.



**Fig. 4** The hydrophobic residues forming the dimer interface between two CTD monomers are displayed in stick representation in **a**. **b** Depicts the radius of gyration of the hydrophobic dimer

interface. The stability of the  $\pi$ -stacking interaction between Trp243 of both chains is monitored in **c**



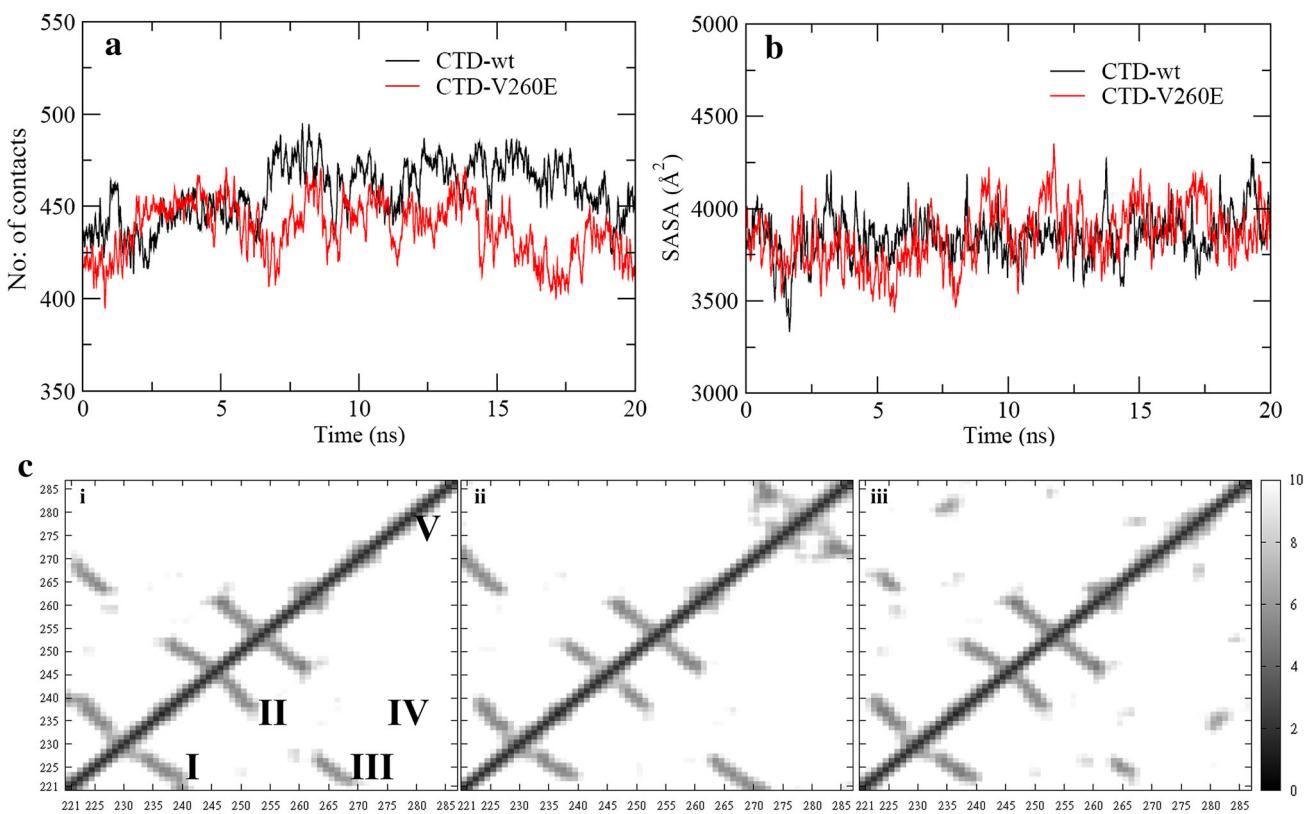
**Fig. 5** MD simulation analysis of CTD-wt and V260E monomers at 498 K. Superimposition of initial (gray) and 50 ns simulated structures of CTD-wt (light blue) and CTD-V260E (dark blue) is shown in **a**. RMSD (**b**), RMS fluctuation (**c**), radius of gyration (**d**),

pairwise interaction energy of V260/Glu260 decomposed into vdw (**g**), electrostatic (**h**) and total (**i**) energy of CTD-wt (blue) and CTD-V260E (red) simulations are shown. **e, f** Depicts the variation in the secondary structure of CTD-wt and CTD-V260E respectively

The secondary structural changes observed during the simulations at 498 K has been plotted using DSSP (Fig. 5e, f). In CTD-wt, the overall secondary structures are well preserved with minor disturbance at  $\beta 2$ . The  $\beta_{10}$ -helix between  $\beta 4$  and  $\beta 5$  is also highly stable. The C-tail expresses a helical turn during the first half of the simulation period and is disturbed at the later stages. On the other hand, the secondary structures of CTD-V260E are highly disturbed. During dynamics, the length of  $\beta$ -sheets 1 and 5 are highly diminished and  $\beta 2$  toggles between anti-parallel and parallel orientation. The loop between  $\beta 3$  and  $\beta 4$  shows vibrant dynamics. Also, the  $\beta_{10}$ -helix between  $\beta 4$  and  $\beta 5$  exists only

up to the simulation of 10 ns and after uncoils to a helical turn. Two observed H-bonds between Val260 and Val249 in CTD-wt are highly conserved and consistent during dynamics. Similar is the case with the equivalent H-bond (between Glu260 and Val249) in CTD-V260E. The pairwise interaction energies between Val260/Glu260 and rest of the CTD residues also show a similar trend as in CTD-wt (Fig. 5g-i). These observations clearly indicate the enhanced dynamics of CTD-V260E to mediate unfolding while the CTD-wt is stable even under unfolding conditions.

The unfolding of CTD-wt and CTD-V260E was further validated by performing 12 individual simulations using

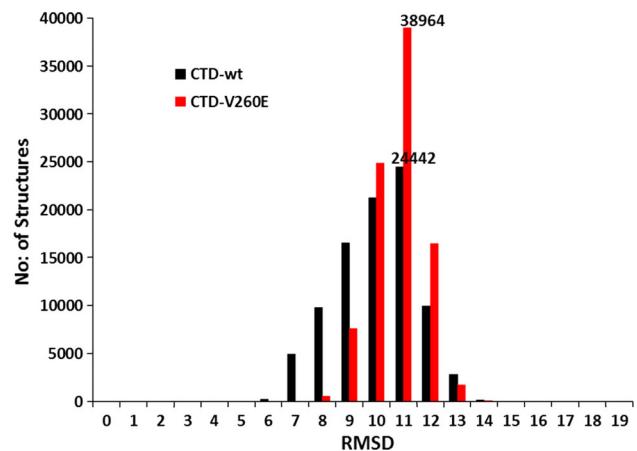


**Fig. 6** Variation in the number of atoms within 7 Å of each atom and the solvent accessible surface area of CTD-wt (black) and CTD-V260E (red) is depicted in **a**, **b** respectively. The contact map of

initial, 20 ns simulated conformations of CTD-wt and CTD-V260E are shown in **c(i)**, **c(ii)** and **c(iii)** respectively

the structures retrieved from 50 ns trajectory simulated at 300 K as starting structures. Each simulation was carried out for 20 ns at 498 K and the trajectories of last 15 ns were accounted for in this study. Overall, about 90,000 structures sampled during the simulations of CTD-wt and CTD-V260E were retrieved for further analysis. The RMSD of all these structures was calculated with respect to the initial structure and the observed population in every 1 Å is shown in Fig. 7. The CTD-wt simulation shows a gradual increase in population from 6 to 11 Å and substantially decreases in the next 2 Å range. In contrast to this, in CTD-V260E simulation, majority of the structures exhibit RMSD in the range of 9–12 Å. A maximum of 24,442 and 38,964 conformations are present in 11–12 Å range in CTD-wt and CTD-V260E, respectively, reinforces the dynamic behavior of CTD-V260E.

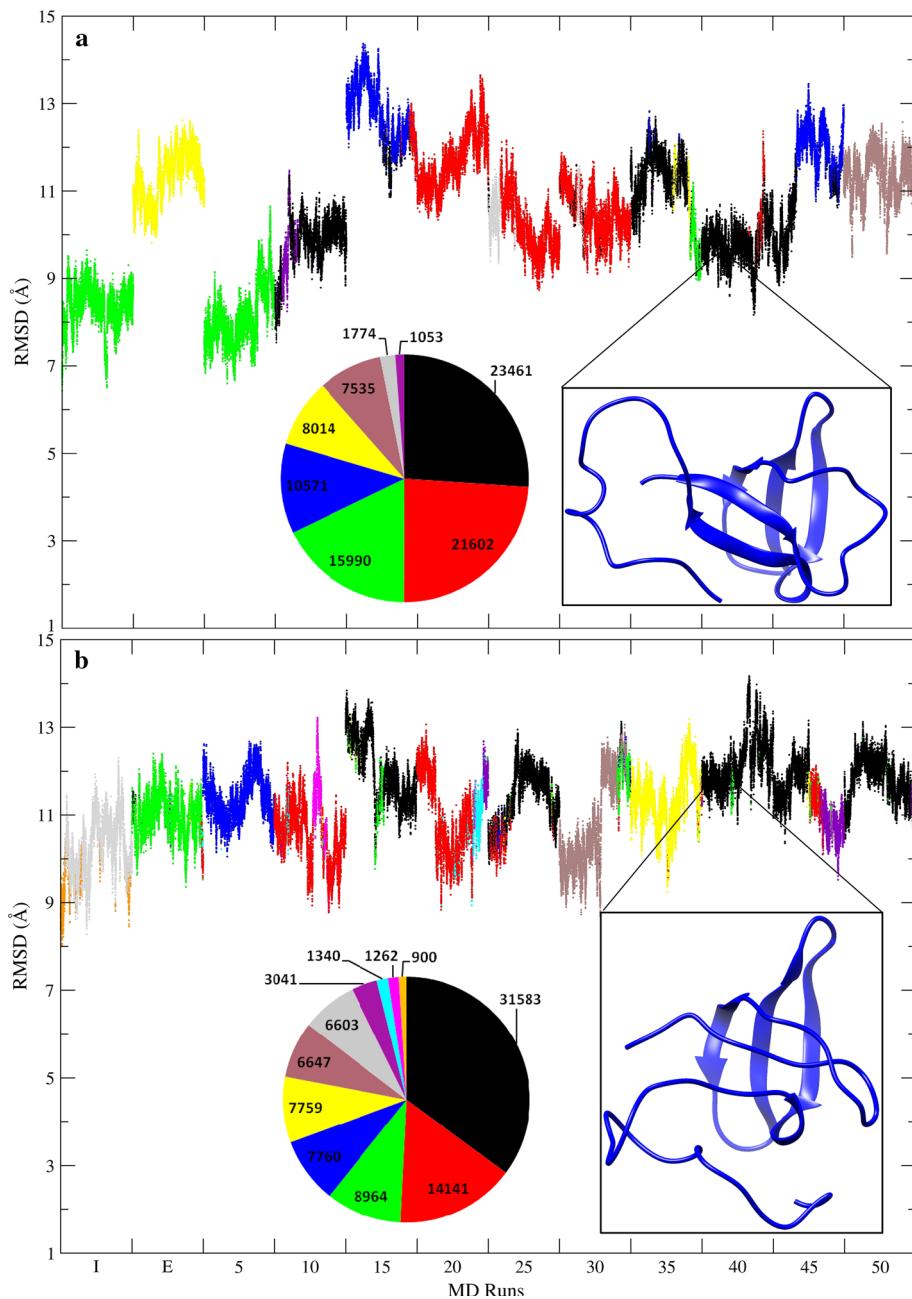
The conformations retrieved from CTD-wt and CTD-V260E simulations were clustered based on the distance-RMSD and a minimum cut-off distance of 4 Å between clusters. CTD-wt resulted in 8 clusters, while the dynamic CTD-V260E resulted in 11 clusters (inset in Fig. 8). The Fig. 8a, b shows the distribution of these clusters over the RMSD trajectory (calculated with respect to the native structure) observed for CTD-wt and CTD-V260E



**Fig. 7** Histogram showing the RMSD of CTD-wt (black) and CTD-V260E (red) from 12 independent simulations performed at 498 K

simulations, respectively. In CTD-wt, the most populated cluster (black) sampling RMSD about ~10 Å represent the folded state of CTD, while the structures expressing higher RMSD (~14 Å) denote the dynamic unfolded state (classified as fourth cluster colored in blue). The intermediate structures between these folded and unfolded states form the second cluster (red). In contrast to this, majority

**Fig. 8** Clustering analysis of CTD-wt (a) and CTD-V260E (b). The pie chart (*inset*) represents the number of structures in each cluster. Distribution of clusters over the complete trajectory is shown by coloring the RMSD according to the color adopted in pie chart. The representative conformation of the maximum populated cluster (*black*) is also shown as an *inset*. The X-axis represents the time at which initial structures were extracted from 300 K simulation (I is the initial and E is the equilibrated structures) and each division represents the trajectory from the last 15 ns of 12 simulations performed at 498 K

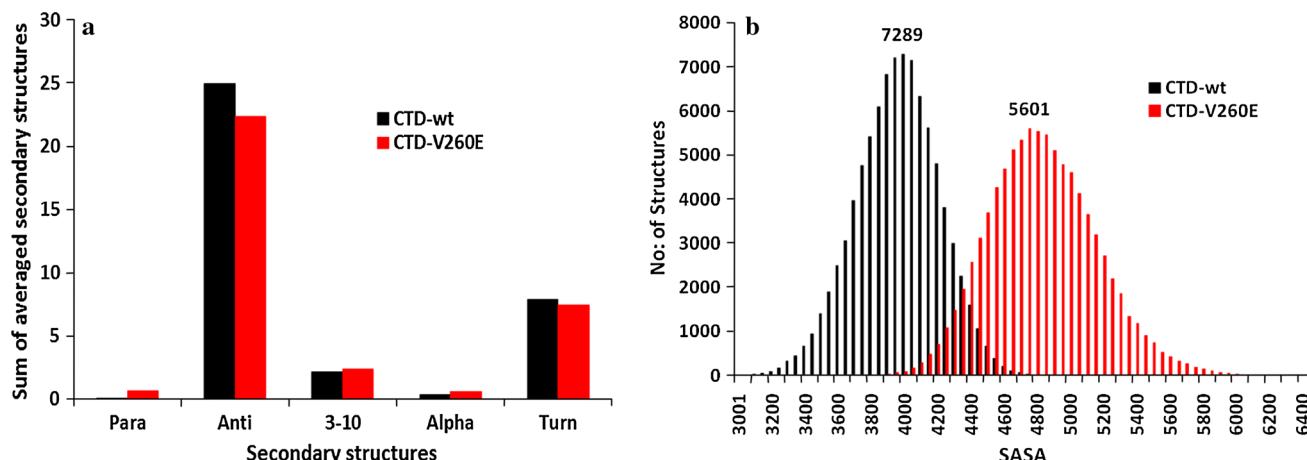


of CTD-V260E conformations exist in highly deviated/unfolded state and hence are sampled as the first cluster (black) while the intermediates span several clusters. The representative conformations (closest to the centroid of cluster) retrieved from first cluster is shown as an inset in Fig. 8a, b. These structures reinforce the perturbed nature of both  $\beta$ 1 and  $\beta$ 5 of CTD-V260E observed in the single 20 ns simulation at 498 K, which remains stable in CTD-wt.

The secondary structure of each residue averaged over a set of conformations could explain the stability of its three dimensional structure and is analyzed in this study. Figure 9a depicts a comparison of the total secondary structures

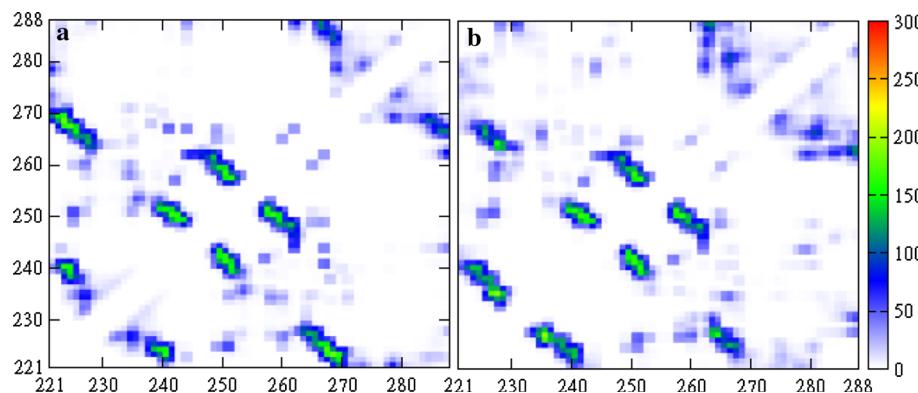
calculated for the conformations belonging to first cluster in both CTD-wt and CTD-V260E. In CTD-V260E, the occurrence of anti-parallel  $\beta$ -sheets is reduced while new parallel  $\beta$ -sheets are formed, and demonstrates a conformational variation from the regular SH3 fold. The SASA of CTD-wt (Fig. 9b) varies between 3,000 and 4,800  $\text{\AA}^2$ , whereas, the SASA of CTD-V260E is drastically increased between 4,000 and 6,000  $\text{\AA}^2$  and indicates the unfolding of CTD-V260E.

The number of native contacts within 7  $\text{\AA}$  was calculated for the most populated cluster by ignoring the contacts between  $i$ th and  $i + 4$ th residues (shown in Fig. 10). The native contacts in each conformation were determined by considering the initial structure as reference to analyze



**Fig. 9** Histogram of total secondary structure (a) and solvent accessible surface area (b) calculated from the conformations in the most populated cluster of CTD-wt (black) and CTD-V260E (red) simulations

**Fig. 10** The native contact map of CTD-wt (a) and CTD-V260E (b) derived from the most populated cluster



the conformational changes due to temperature induced unfolding of CTD. Similar to the contacts observed in Fig. 6c, the  $\beta 1$  interacts strongly with  $\beta 2$ , while the interaction between  $\beta 1$  and  $\beta 5$  is considerably diminished in CTD-V260E. Such strong interaction of  $\beta 1$  and  $\beta 2$  results in a free dynamics of the loops connecting  $\beta 3$ – $\beta 4$ – $\beta 5$  and further diminishes the number of contacts at the  $\beta 3$ – $\beta 4$ – $\beta 5$ -loops interface. The reduction in total anti-parallel  $\beta$ -sheets in Fig. 9a is due to the loss of  $\beta 1$ – $\beta 5$  contacts, while the interactions within C-tail represent the occurrence of parallel  $\beta$ -sheets. All these analyses reinforce the CTD-V260E mediated unfolding behavior.

This study highlights the role of Val260 in maintaining the structural integrity of CTD by providing a stable hydrophobic environment. Molecular dynamics simulation studies on wild type and V260E mutant CTDs revealed useful insights into the mechanism of inhibition of HIV-IN activity by the V260E mutation. The CTD monomer and dimer are stable in both wild type and mutant forms even after 50 ns molecular dynamics simulations. Their RMS deviation from the initial structure is majorly influenced by

the conformation of the C-tail alone. Also, the hydrophobic core of the CTD  $\beta$ -barrel is essentially stable. The C-tails are highly flexible and show random behavior since they lack in an interacting partner. The hydrophobic nature of CTD is disturbed due to V260E mutation and hence an electrostatic behavior is induced. The unfavorable interaction induced by the mutation of a hydrophobic residue (Val) to a polar charged residue (Glu) is compensated by the electrostatic interaction of Glu260 with the acidic residues of CTD and confers an internal stability to the CTD monomer and dimer. A strong charged cluster including the residues Glu260, Arg228, Lys258, Lys236, Arg263 and Lys264 is formed on the neutral face of the CTD monomer and thus disturbs the chemical nature of CTD. Though MD analysis revealed a stable CTD dimer, the residues reported as important for dimerization express strong intra-monomer electrostatic attraction in CTD-V260E. The CTD-V260E dimer expresses increased conformational entropy than the CTD-wt dimer and the difference in their binding energy is 2 kcal/mol. The important  $\pi$ -stacking interaction between Trp243 is also disturbed in CTD-V260E. Overall,

the dimerization of CTD is not highly affected by the V260E mutant, however, the nature of the CTD-V260E dimer is predominated by electrostatic interaction rather than hydrophobic as in CTD-wt. Glu260 is also observed to form intra-monomer H-bonds in all simulated systems whereas such interactions are absent in the wild type. The spatial arrangement of full-length ASV and PFV IN dimers are similar except the orientation of CTD [32]. The ASV dimer is stabilized by CTD–CTD interface and in particular by the  $\pi$ -stacking interaction of Trp259 (Trp243 in HIV) from both the CTDs. When DNA binds, the Trp259–Trp259' stacking interaction is disturbed along with a CTD rotation at the CCD–CTD linker region. This unstacking and rotation of CTD changes its orientation in the intasome and facilitates binding and stabilization of the DNA by hydrophobic interactions. In the CTD-V260E mutant dimer, the stacking between Trp243 and Trp243' is disturbed and hence it is clear that, the mutation affects dimerization in the absence of DNA substrate. Also, the change in the nature of CTD from hydrophobic to electrostatic might influence its interaction within the dimer interface or with the DNA substrate. It is also reported that the activity of V260E mutant is comparable to wild type activity and is supported by our present observation that the CTD-V260E mutation does not significantly affect the dimerization. As this mutation is also reported to influence the folding mechanism of CTD, role of V260E on CTD folding was studied. High temperature induced unfolding study on CTD-wt and CTD-V260E mutant revealed that the CTD-V260E mutant unfolds more easily than CTD-wt along with major structural changes in  $\beta$ 1,  $\beta$ 2 and  $\beta$ 5 sheets and the 3<sub>10</sub>-helix. The CTD-wt expresses stable dynamics with conserved SH3 fold. As the V260E mutation disturbs the hydrophobic nature, the subsequent change in CTD topology results in major effects on protein folding rather than multimerization.

## Conclusion

Val260 does not directly involve in multimerization or DNA binding activities of IN. However, mutation of a hydrophobic residue (Val) to a polar residue (Glu) disturbs the function of IN by altering the chemical nature of CTD. SH3 domains are common protein–protein interaction sites majorly mediated by hydrophobic interactions. It is also clear from our study that the hydrophobic nature of CTD highly influences its activity such as IN multimerization and folding. Targeting HIV-1 IN has been majorly focused on the inhibition of catalytic activity by IN strand transfer inhibitors (INSTIs). Zinc ejectors that act on IN-NTD are also been studied as an emerging class of inhibitors that affect IN activity allosterically. Previous studies on

designing inhibitors targeting CTD have identified monoclonal antibodies such as 2–19, 8–22 and 33, and small molecule pyridoxal 5'-phosphate (PLP) and its derivatives that block IN-RT interaction and DNA binding. In addition to these inhibition methods, the major finding from our study that perturbing the hydrophobic nature of CTD inactivates IN might pave the way for identifying novel anti-HIV small molecule allosteric inhibitors designed specifically to target CTD.

**Acknowledgments** One of the Authors, Balasubramanian Sangeetha acknowledges the Department of Science and Technology, India for providing financial support through INSPIRE fellowship. Ramaswamy Amutha acknowledges Science and Engineering Research Board, India for providing computational facilities in the form of Fast Track Research Project for Young Scientists.

## References

- Chen H, Engelman A (2001) Asymmetric processing of human immunodeficiency virus type 1 cDNA in vivo: implications for functional end coupling during the chemical steps of DNA transposition. *Mol Cell Biol* 21(20):6758–6767. doi:[10.1128/MCB.21.20.6758-6767.2001](https://doi.org/10.1128/MCB.21.20.6758-6767.2001)
- Guiot E, Carayon K, Delelis O, Simon F, Tauc P, Zubin E, Gottikh M, Mouscadet JF, Brochon JC, Deprez E (2006) Relationship between the oligomeric status of HIV-1 integrase on DNA and enzymatic activity. *J Biol Chem* 281(32):22707–22719. doi:[10.1074/jbc.M602198200](https://doi.org/10.1074/jbc.M602198200)
- Sherman PA, Fyfe JA (1990) Human immunodeficiency virus integration protein expressed in *Escherichia coli* possesses selective DNA cleaving activity. *Proc Natl Acad Sci USA* 87(13):5119–5123
- Vink C, Yeheskiely E, van der Marel GA, van Boom JH, Plasterk RH (1991) Site-specific hydrolysis and alcoholysis of human immunodeficiency virus DNA termini mediated by the viral integrase protein. *Nucleic Acids Res* 19(24):6691–6698
- Wei SQ, Mizuuchi K, Craigie R (1997) A large nucleoprotein assembly at the ends of the viral DNA mediates retroviral DNA integration. *EMBO J* 16(24):7511–7520. doi:[10.1093/emboj/16.24.7511](https://doi.org/10.1093/emboj/16.24.7511)
- Engelman A, Mizuuchi K, Craigie R (1991) HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer. *Cell* 67(6):1211–1221
- Bushman FD, Craigie R (1991) Activities of human immunodeficiency virus (HIV) integration protein in vitro: specific cleavage and integration of HIV DNA. *Proc Natl Acad Sci USA* 88(4):1339–1343
- Bushman FD, Fujiwara T, Craigie R (1990) Retroviral DNA integration directed by HIV integration protein in vitro. *Science* 249(4976):1555–1558
- Delelis O, Carayon K, Saib A, Deprez E, Mouscadet JF (2008) Integrase and integration: biochemical activities of HIV-1 integrase. *Retrovirology* 5:114. doi:[10.1186/1742-4690-5-114](https://doi.org/10.1186/1742-4690-5-114)
- Muller HP, Varmus HE (1994) DNA bending creates favored sites for retroviral integration: an explanation for preferred insertion sites in nucleosomes. *EMBO J* 13(19):4704–4714
- 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 USA* 90(8):3428–3432

12. 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(8):3269–3275
13. Engelman A, Craigie R (1992) Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro. *J Virol* 66(11):6361–6369
14. van Gent DC, Vink C, Groeneger AA, Plasterk RH (1993) Complementation between HIV integrase proteins mutated in different domains. *EMBO J* 12(8):3261–3267
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(7):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(31):9826–9833
17. Zheng R, Jenkins TM, Craigie R (1996) Zinc folds the N-terminal domain of HIV-1 integrase, promotes multimerization, and enhances catalytic activity. *Proc Natl Acad Sci USA* 93(24):13659–13664
18. 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(16):9150–9154
19. 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(2):359–368. doi:[10.1006/jmbi.1998.2002](https://doi.org/10.1006/jmbi.1998.2002)
20. Engelman A, Hickman AB, Craigie R (1994) The core and carboxyl-terminal domains of the integrase protein of human immunodeficiency virus type 1 each contribute to nonspecific DNA binding. *J Virol* 68(9):5911–5917
21. Woerner AM, Klutchnik M, Levin JG, Marcus-Sekura CJ (1992) Localization of DNA binding activity of HIV-1 integrase to the C-terminal half of the protein. *AIDS Res Hum Retroviruses* 8(2):297–304
22. Jaskolski M, Alexandratos JN, Bujacz G, Wlodawer A (2009) Piecing together the structure of retroviral integrase, an important target in AIDS therapy. *FEBS J* 276(11):2926–2946. doi:[10.1111/j.1742-4658.2009.07009.x](https://doi.org/10.1111/j.1742-4658.2009.07009.x)
23. 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(24):7333–7343. doi:[10.1093/emboj/20.24.7333](https://doi.org/10.1093/emboj/20.24.7333)
24. Chen JC, Krucinski J, Miercke LJ, Finer-Moore JS, Tang AH, Leavitt AD, Stroud RM (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(15):8233–8238. doi:[10.1073/pnas.150220297](https://doi.org/10.1073/pnas.150220297)
25. Balasubramanian S, Rajagopalan M, Ramaswamy A (2012) Structural dynamics of full-length retroviral integrase: a molecular dynamics analysis. *J Biomol Struct Dyn* 29(6):659–670. doi:[10.1080/07391102.2011.672630](https://doi.org/10.1080/07391102.2011.672630)
26. Heuer TS, Brown PO (1998) Photo-cross-linking studies suggest a model for the architecture of an active human immunodeficiency virus type 1 integrase-DNA complex. *Biochemistry* 37(19):6667–6678. doi:[10.1021/bi972949c](https://doi.org/10.1021/bi972949c)
27. Karki RG, Tang Y, Burke TR Jr, Nicklaus MC (2004) Model of full-length HIV-1 integrase complexed with viral DNA as template for anti-HIV drug design. *J Comput Aided Mol Des* 18(12):739–760. doi:[10.1007/s10822-005-0365-5](https://doi.org/10.1007/s10822-005-0365-5)
28. Krishnan L, Li X, Naraharisetti HL, Hare S, Cherepanov P, Engelman A (2010) Structure-based modeling of the functional HIV-1 intasome and its inhibition. *Proc Natl Acad Sci USA* 107(36):15910–15915. doi:[10.1073/pnas.1002346107](https://doi.org/10.1073/pnas.1002346107)
29. Podtelezhnikov AA, Gao K, Bushman FD, McCammon JA (2003) Modeling HIV-1 integrase complexes based on their hydrodynamic properties. *Biopolymers* 68(1):110–120. doi:[10.1002/bip.10217](https://doi.org/10.1002/bip.10217)
30. Wielens J, Crosby IT, Chalmers DK (2005) A three-dimensional model of the human immunodeficiency virus type 1 integration complex. *J Comput Aided Mol Des* 19(5):301–317. doi:[10.1007/s10822-005-5256-2](https://doi.org/10.1007/s10822-005-5256-2)
31. Hare S, Gupta SS, Valkov E, Engelman A, Cherepanov P (2010) Retroviral intasome assembly and inhibition of DNA strand transfer. *Nature* 464(7286):232–236. doi:[10.1038/nature08784](https://doi.org/10.1038/nature08784)
32. Bojja RS, Andrade MD, Weigand S, Merkel G, Yarychkivska O, Henderson A, Kummerling M, Skalka AM (2011) Architecture of a full-length retroviral integrase monomer and dimer, revealed by small angle X-ray scattering and chemical cross-linking. *J Biol Chem* 286(19):17047–17059. doi:[10.1074/jbc.M110.212571](https://doi.org/10.1074/jbc.M110.212571)
33. Bojja RS, Andrade MD, Merkel G, Weigand S, Dunbrack RL Jr, Skalka AM (2013) Architecture and assembly of HIV integrase multimers in the absence of DNA substrates. *J Biol Chem* 288(10):7373–7386. doi:[10.1074/jbc.M112.434431](https://doi.org/10.1074/jbc.M112.434431)
34. Zhao Z, McKee CJ, Kessl JJ, Santos WL, Daigle JE, Engelman A, Verdine G, Kvaratskhelia M (2008) Subunit-specific protein footprinting reveals significant structural rearrangements and a role for N-terminal Lys-14 of HIV-1 integrase during viral DNA binding. *J Biol Chem* 283(9):5632–5641. doi:[10.1074/jbc.M705241200](https://doi.org/10.1074/jbc.M705241200)
35. Engelman A (1999) In vivo analysis of retroviral integrase structure and function. *Adv Virus Res* 52:411–426
36. Shin CG, Taddeo B, Haseltine WA, Farnet CM (1994) Genetic analysis of the human immunodeficiency virus type 1 integrase protein. *J Virol* 68(3):1633–1642
37. Taddeo B, Haseltine WA, Farnet CM (1994) Integrase mutants of human immunodeficiency virus type 1 with a specific defect in integration. *J Virol* 68(12):8401–8405
38. Lee SP, Xiao J, Knutson JR, Lewis MS, Han MK (1997)  $Zn^{2+}$  promotes the self-association of human immunodeficiency virus type-1 integrase in vitro. *Biochemistry* 36(1):173–180. doi:[10.1021/bi961849o](https://doi.org/10.1021/bi961849o)
39. Engelman A, Englund G, Orenstein JM, Martin MA, Craigie R (1995) Multiple effects of mutations in human immunodeficiency virus type 1 integrase on viral replication. *J Virol* 69(5):2729–2736
40. Nakamura T, Masuda T, Goto T, Sano K, Nakai M, Harada S (1997) Lack of infectivity of HIV-1 integrase zinc finger-like domain mutant with morphologically normal maturation. *Biochem Biophys Res Commun* 239(3):715–722. doi:[10.1006/bbrc.1997.7541](https://doi.org/10.1006/bbrc.1997.7541)
41. Wu XY, Liu HM, Xiao HL, Conway JA, Hehl E, Kalpana GV, Prasad V, Kappes JC (1999) Human immunodeficiency virus type 1 integrase protein promotes reverse transcription through specific interactions with the nucleoprotein reverse transcription complex. *J Virol* 73(3):2126–2135
42. Leavitt AD, Robles G, Alessandro N, Varmus HE (1996) Human immunodeficiency virus type 1 integrase mutants retain in vitro integrase activity yet fail to integrate viral DNA efficiently during infection. *J Virol* 70(2):721–728
43. Masuda T, Planelles V, Krogstad P, Chen IS (1995) Genetic analysis of human immunodeficiency virus type 1 integrase and the U3 att site: unusual phenotype of mutants in the zinc finger-like domain. *J Virol* 69(11):6687–6696
44. Engelman A (2011) Pleiotropic nature of HIV-1 integrase mutations. In: Neamati N (ed) HIV-1 INTEGRASE: mechanism and inhibitor design. Wiley, Hoboken
45. Eijkelkamp AP, Lutzke RA, Boelens R, Plasterk RH, Kaptein R, Hard K (1995) The DNA-binding domain of HIV-1 integrase has an SH3-like fold. *Nat Struct Biol* 2(9):807–810
46. Eijkelkamp AP, Sprangers R, Hard 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(4):556–564. doi:[10.1002/\(SICI\)1097-0134\(19990901\)36:4<556::AID-PROT18>3.0.CO;2-6](https://doi.org/10.1002/(SICI)1097-0134(19990901)36:4<556::AID-PROT18>3.0.CO;2-6)
47. Dar MJ, Monel B, Krishnan L, Shun MC, Di Nunzio F, Helland DE, Engelman A (2009) Biochemical and virological analysis of the 18-residue C-terminal tail of HIV-1 integrase. *Retrovirology* 6:94. doi:[10.1186/1742-4690-6-94](https://doi.org/10.1186/1742-4690-6-94)
  48. Mohammed KD, Topper MB, Muesing MA (2011) Sequential deletion of the integrase (Gag-Pol) carboxyl terminus reveals distinct phenotypic classes of defective HIV-1. *J Virol* 85(10): 4654–4666. doi:[10.1128/JVI.02374-10](https://doi.org/10.1128/JVI.02374-10)
  49. Mayer BJ, Eck MJ (1995) SH3 domains. Minding your p's and q's. *Curr Biol* 5(4):364–367
  50. Hehl EA, Joshi P, Kalpana GV, Prasad VR (2004) Interaction between human immunodeficiency virus type 1 reverse transcriptase and integrase proteins. *J Virol* 78(10):5056–5067
  51. Zhu K, Dobard C, Chow SA (2004) Requirement for integrase during reverse transcription of human immunodeficiency virus type 1 and the effect of cysteine mutations of integrase on its interactions with reverse transcriptase. *J Virol* 78(10):5045–5055
  52. Dirac AM, Kjems J (2001) Mapping DNA-binding sites of HIV-1 integrase by protein footprinting. *Eur J Biochem* 268(3):743–751
  53. Gao K, Butler SL, Bushman F (2001) Human immunodeficiency virus type 1 integrase: arrangement of protein domains in active cDNA complexes. *EMBO J* 20(13):3565–3576. doi:[10.1093/emboj/20.13.3565](https://doi.org/10.1093/emboj/20.13.3565)
  54. Lutzke RA, Plasterk RH (1998) Structure-based mutational analysis of the C-terminal DNA-binding domain of human immunodeficiency virus type 1 integrase: critical residues for protein oligomerization and DNA binding. *J Virol* 72(6):4841–4848
  55. Lutzke RA, Vink C, Plasterk RH (1994) Characterization of the minimal DNA-binding domain of the HIV integrase protein. *Nucleic Acids Res* 22(20):4125–4131
  56. Kalpana GV, Reicin A, Cheng GS, Sorin M, Paik S, Goff SP (1999) Isolation and characterization of an oligomerization-negative mutant of HIV-1 integrase. *Virology* 259(2):274–285. doi:[10.1006/viro.1999.9767](https://doi.org/10.1006/viro.1999.9767)
  57. Ishikawa T, Okui N, Kobayashi N, Sakuma R, Kitamura T, Kitamura Y (1999) Monoclonal antibodies against the minimal DNA-binding domain in the carboxyl-terminal region of human immunodeficiency virus type 1 integrase. *J Virol* 73(5):4475–4480
  58. Ceccherini-Silberstein F, Malet I, D'Arrigo R, Antinori A, Marcelin AG, Perno CF (2009) Characterization and structural analysis of HIV-1 integrase conservation. *AIDS Rev* 11(1):17–29
  59. Hornak V, Okur A, Rizzo RC, Simmerling C (2006) HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations. *Proc Natl Acad Sci USA* 103(4):915–920. doi:[10.1073/pnas.0508452103](https://doi.org/10.1073/pnas.0508452103)
  60. Kim JH, Hartley TL, Curran AR, Engelman DM (2009) Molecular dynamics studies of the transmembrane domain of gp41 from HIV-1. *Biochim Biophys Acta* 1788(9):1804–1812. doi:[10.1016/j.bbamem.2009.06.011](https://doi.org/10.1016/j.bbamem.2009.06.011)
  61. 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(6):2999–3011
  62. Madrid M, Jacobo-Molina A, Ding J, Arnold E (1999) Major subdomain rearrangement in HIV-1 reverse transcriptase simulated by molecular dynamics. *Proteins* 35(3):332–337. doi:[10.1002/\(SICI\)1097-0134\(19990515\)35:3<332::AID-PROT7>3.0.CO;2-R](https://doi.org/10.1002/(SICI)1097-0134(19990515)35:3<332::AID-PROT7>3.0.CO;2-R)
  63. Sangeetha B, Muthukumaran R, Amutha R (2014) The dynamics of interconverting D- and E-forms of the HIV-1 integrase N-terminal domain. *Eur Biophys J* 43(10–11):485–498. doi:[10.1007/s00249-014-0979-4](https://doi.org/10.1007/s00249-014-0979-4)
  64. Barreca ML, Lee KW, Chimirri A, Briggs JM (2003) Molecular dynamics studies of the wild-type and double mutant HIV-1 integrase complexed with the 5CITEP inhibitor: mechanism for inhibition and drug resistance. *Biophys J* 84(3):1450–1463. doi:[10.1016/S0006-3495\(03\)74958-3](https://doi.org/10.1016/S0006-3495(03)74958-3)
  65. Brigo A, Lee KW, Fogolari F, Mustata GI, Briggs JM (2005) Comparative molecular dynamics simulations of HIV-1 integrase and the T66I/M154I mutant: binding modes and drug resistance to a diketo acid inhibitor. *Proteins* 59(4):723–741. doi:[10.1002/prot.20447](https://doi.org/10.1002/prot.20447)
  66. Li A, Daggett V (1994) Characterization of the transition state of protein unfolding by use of molecular dynamics: chymotrypsin inhibitor 2. *Proc Natl Acad Sci USA* 91(22):10430–10434
  67. Mayor U, Guydosh NR, Johnson CM, Grossmann JG, Sato S, Jas GS, Freund SM, Alonso DO, Daggett V, Fersht AR (2003) The complete folding pathway of a protein from nanoseconds to microseconds. *Nature* 421(6925):863–867. doi:[10.1038/nature01428](https://doi.org/10.1038/nature01428)
  68. Alonso DO, Alm E, Daggett V (2000) Characterization of the unfolding pathway of the cell-cycle protein p13suc1 by molecular dynamics simulations: implications for domain swapping. *Structure* 8(1):101–110
  69. Kazmirski SL, Daggett V (1998) Non-native interactions in protein folding intermediates: molecular dynamics simulations of hen lysozyme. *J Mol Biol* 284(3):793–806. doi:[10.1006/jmbi.1998.2192](https://doi.org/10.1006/jmbi.1998.2192)
  70. Alonso DO, Daggett V (1998) Molecular dynamics simulations of hydrophobic collapse of ubiquitin. *Protein Sci* 7(4):860–874. doi:[10.1002/pro.5560070404](https://doi.org/10.1002/pro.5560070404)
  71. Gu W, Wang T, Zhu J, Shi Y, Liu H (2003) Molecular dynamics simulation of the unfolding of the human prion protein domain under low pH and high temperature conditions. *Biophys Chem* 104(1):79–94
  72. Kazmirski SL, Daggett V (1998) Simulations of the structural and dynamical properties of denatured proteins: the “molten coil” state of bovine pancreatic trypsin inhibitor. *J Mol Biol* 277(2):487–506. doi:[10.1006/jmbi.1998.1634](https://doi.org/10.1006/jmbi.1998.1634)
  73. Case DA, Darden TA, Cheatham TE, 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
  74. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* 65(3):712–725. doi:[10.1002/prot.21123](https://doi.org/10.1002/prot.21123)
  75. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79(2):926–935. doi:[10.1063/1.445869](https://doi.org/10.1063/1.445869)
  76. Schreiner E, Trabuco LG, Freddolino PL, Schulten K (2011) Stereochemical errors and their implications for molecular dynamics simulations. *BMC Bioinformatics* 12:190. doi:[10.1186/1471-2105-12-190](https://doi.org/10.1186/1471-2105-12-190)
  77. Miller BR, McGee TD, Swails JM, Homeyer N, Gohlke H, Roitberg AE (2012) MMPBSA.py: an efficient program for end-state free energy calculations. *J Chem Theory Comput* 8(9): 3314–3321. doi:[10.1021/ct300418h](https://doi.org/10.1021/ct300418h)