Mutational Analysis and Allosteric Effects in the HIV-1 Capsid Protein Carboxyl-Terminal Dimerization Domain

Xiang Yu,†‡ Qiuming Wang,†‡ Jui-Chen Yang,‡ Idit Buch,§ Chung-Jung Tsai,‖ Buyong Ma,‖ Stephen Z. D. Cheng,∥ Ruth Nussinov,*,§‖ and Jie Zheng*

Departments of Chemical and Biomolecular Engineering and Polymer Science, The University of Akron, Akron, Ohio 44325, Basic Research Program, SAIC-Frederick, Inc., Center for Cancer Research Nanobiology Program, NCI-Frederick, Frederick, Maryland 21702, and Sackler Institute of Molecular Medicine, Department of Human Genetics and Molecular Medicine, Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel

Received October 10, 2008; Revised Manuscript Received December 8, 2008

The carboxyl-terminal domain (CTD, residues 146–231) of the HIV-1 capsid (CA) protein plays an important role in the CA–CA dimerization and viral assembly of the human immunodeficiency virus type 1. Disrupting the native conformation of the CA is essential for blocking viral capsid formation and viral replication. Thus, it is important to identify the exact nature of the structural changes and driving forces of the CTD dimerization that take place in mutant forms. Here, we compare the structural stability, conformational dynamics, and association force of the CTD dimers for both wild-type and mutated sequences using all-atom explicit-solvent molecular dynamics (MD). The simulations show that Q155N and E159D at the major homology region (MHR) and W184A and M185A at the helix 2 region are energetically less favorable than the wild-type, imposing profound negative effects on intermolecular CA–CA dimerization. Detailed structural analysis shows that three mutants (Q155N, E159D, and W184A) display much more flexible local structures and weaker CA–CA association than the wild-type, primarily due to the loss of interactions (hydrogen bonds, side chain hydrophobic contacts, and $\pi$-stacking) with their neighboring residues. Most interestingly, the MHR that is far from the interacting dimeric interface is more sensitive to the mutations than the helix 2 region that is located at the CA–CA dimeric interface, indicating that structural changes in the distinct motif of the CA could similarly allosterically prevent the CA capsid formation. In addition, the structural and free energy comparison of the five residues shorter CA (151–231) dimer further indicates that hydrophobic interactions, side chain packing, and hydrogen bonds are the major, dominant forces driving in stabilizing the CA interface.

Introduction

Human immunodeficiency retroviruses type 1 (HIV-1) can initially self-assemble into immature, noninfectious, spherical particles consisting of approximately 5000 Gag molecules.1–3 Initiated by the viral protease, the Gag proteins are then cleaved into different new proteins including matrix (MA), capsid (CA), nucleocapsid (NC), and three smaller peptides.4,5 During the subsequent virus maturation, MA remains attached to the viral membrane, whereas hundreds of chemically identical CA proteins spontaneously assemble into large and complex capsid structures with different morphologies that enclose the NC–RNA complex. Genetic analyses have revealed that CA assembly is essential to the understanding of the mechanism of viral activity, infectivity, and assembly6 and also important in the development of new drugs.7 CA is composed of two folded domains, an N-terminal domain (NTD) and a C-terminal domain (CTD), in which these two domains are connected by a flexible linker. The NTD (residues 1–146) consisting of seven $\alpha$-helices and an extended proline-loop5,8–10 is responsible for viral capsid formation,8 while the CTD (residues 148–231) consisting of four $\alpha$-helices is responsible for assembly of immature Gag shell11 and virion formation.12

It is well-known that the CA protein can spontaneously assemble into different molecular structures such as cones, spheres, and long helical tubes both in vivo and in vitro.9,13 The different capsid morphologies may represent different assembly stages or pathways to stable virus capsids. Li et al.9 reported image reconstructions of helical assemblies of the HIV-1 CA protein in vitro using cryo-electron microscopy (cryo-EM). They found that NTDs of CA were organized in hexameric rings while CTDs of CA associated via dimeric contacts by interacting with adjacent hexamers. Briggs and co-workers14 also observed similar CA hexamers with local p6 symmetry in vivo. Recently, Ganser-Pornillos et al.14 reported a three-dimensional structure of hexameric arrays of full-length HIV CA at 9 Å resolution using the EM and image analysis.14 All these reports suggest that CA capsid assemblies share a similar lattice organization (i.e., hexamerization of the NTDs and dimerization of the CTDs) in vitro and in vivo, although the detailed lattice parameters of hexameric packing are different from those proposed for HIV CA structures. Mutational analyses revealed that although mutations in both domains of the CA protein can alter or abolish capsid assembly, virus replication and infectivity, and protein stability,5,15 mutations in the CTD are more disruptive to capsid formation than mutations in the NTD. The CA protein has been extensively studied with respect to its molecular structure and biological function using a wide
variety of approaches including nuclear magnetic resonance (NMR), X-ray diffraction, and cryo-electron microscopy (EM), as well as theoretical approaches including molecular dynamics (MD) simulation and normal-mode analysis (NMA). However, the precise nature of the folding pathway, the driving force that determines the overall stability of the capsid, and the packing and binding mode between NTDs and CTDs remain unclear.

From a simulation point of view, it is a very challenging task to directly investigate the self-assembly process of the virus capsid, as well as the dynamic and structural properties of the whole virus capsid, using conventional all-atom MD simulation due to the large number of atoms (more than 1 million) and possible assembly pathways. However, it is well-known that the formation of the virus capsid requires formation of a stable hexameric structure of NTDs and dimeric structure of CTDs and that the dynamic behavior of the domain interface also determines HIV infectivity, capsid assembly, and protein stability. Ganser-Pornillos et al. tested 21 mutations in the hexameric structure of NTDs and dimeric structure of CTDs and confirmed the importance of residues Trp184 and Met185 in the packing along a C2 symmetry (Figure 1). All starting structures of these mutants were built from the wild-type by replacing the W184A and M185A located in helix 2 region, and Q155N and E159D located in helix 3 region, and Q155N and E159D located in helix 4 region. The images are generated by VMD.

**Simulation Methods.** The X-ray crystal structure (resolution 2.6 Å, PDB code 1A43) of the C-terminal dimerization domain of the HIV-1 CA protein (146–231) was used as the starting point for all the MD simulations. In the native HIV capsid, two CTDs form a dimeric structure along a 2-fold screw axis. Each CTD monomer contains four helical structures: helix 1 (161–174; red), helix 2 (179–192; yellow), helix 3 (196–205; green), and helix 4 (211–217; purple). (b) Key residues in the helix 2 with all side chain orientations facing the interface in licorice representation (Glu180, Val181, Trp184, Met185, Thr188, Leu189, and Gin192). The images are generated by VMD.
major homology region (MHR). The structure of the designed mutant was first minimized for 500 steps using the steepest decent algorithm with the backbone of the protein restrained before being subjected to the following system setup and production runs. The N- and C-termini were blocked by acetyl and amine groups, respectively.

**MD Simulations.** All MD simulations were performed using the NAMD program\textsuperscript{20} with the CHARMM27 force field.\textsuperscript{21–23}

Each system was solvated by a TIP3P\textsuperscript{24} water box with a minimum distance of 10 Å from any edge of the box to any protein atom. The periodic boundary condition and minimum image convention were applied to all x, y, and z directions. The standard simulation protocol consisting of initial minimization, heating procedure, equilibrium, and production run was applied to each MD simulation. Each system was initially energy minimized using the conjugate gradient method for 5000 steps with the protein constrained and for additional 5000 steps for the whole system. The system was then subjected to 200 ps of heating procedure, constraining the backbone atoms of the protein to allow relaxation of water molecules. The following 1000 ps equilibrium run was performed without position constraints on the protein. The production simulation used the velocity verlet integrator with an NVT ensemble, a time step of 2 fs, at a temperature of 300 K. The temperature was controlled at 300 K by a Nose-Hoover thermostat. Van der Waals interactions were calculated by the switch function with a twin range cutoff of 10.0 and 12.0 Å and electrostatic interactions were calculated by the force-shift function with a cutoff of 12.0 Å. Nonbonded neighbor lists were updated automatically based on the heuristic algorithm. Each system was simulated for 20 ns and the trajectories were saved at 1.0 ps intervals for later analysis. The details of the setup of the simulations are listed in Table 1.

**Interaction Energy.** Dimeric structures were extracted from explicit MD trajectories by excluding water molecules at a 10 ps time interval. Each structure was first subjected to energy minimization for 1000 conjugate gradient steps. At the minimized state, the solvation energy was then calculated using the Generalized Born with a simple switching method (GBSW).\textsuperscript{25,26}

Thus, the interaction energy between two dimers was calculated by

\[
\langle AG_{\text{dimer}} \rangle = E_{\text{dimer}}^{\text{mm}} + C_{\text{dimer}}^{\text{hw}} - (E_{\text{monomer1}}^{\text{mm}} + C_{\text{monomer1}}^{\text{hw}}) - (E_{\text{monomer2}}^{\text{mm}} + C_{\text{monomer2}}^{\text{hw}}) \tag{1}
\]

where \(E_{\text{mm}}^{\text{mm}}\) is the molecular mechanical energy of the selected system consisting of all bonded and nonbonded energy terms in the CHARMM potential energy function. The monomer trajectory is simply obtained from the dimer trajectory by neglecting other solutes. The method has been used to examine the binding energy of Aβ dimer\textsuperscript{27,28} and von Hippel-Lindau tumor suppressor protein complex.\textsuperscript{29}

**Results and Discussion**

**Impact of Mutations at the Helix 2 Region (W184A and M185A) on the Structure and Dynamics of the CA Dimer.** Figure 2 shows the backbone root-mean-square deviation (RMSD) for the wild-type (146–231) and mutants W184A and M185A. Trp184 and Met185 residues in the helix 2 are deeply buried in the middle region of the dimeric interface. The RMSDs of both the wild-type and M185A mutant were maintained at 2.8 and 3.1 Å for 20 ns, respectively (Figure 2a and Table 1), indicating that (i) both sequences were considerably stable and (ii) replacement of Met185 with Ala has a little influence on the overall stability of dimeric CTDs. In contrast, when the large aromatic Trp184 was replaced by the small Ala in the helix 2 region, the dimeric structure experienced a relatively large structural deviation with a continuous increase in rmsd of 7.3 Å. Consistent with the rmsd analysis, W184A mutant also displayed the largest residue-based root-mean-square fluctuation (RMSF) as compared to the wild-type (Figure 2b), in which the RMSF was used to assess the local dynamical variation of each individual residue. There was no apparent difference in the RMSF between the wild-type and the M185A mutant. The RMSF profile also shows that the all helix structures were consistently stable, with RMSF around 1 Å, while the linkers connecting helices were in general much more flexible. Especially, residues at the helix 2 regions that comprise the dimeric interface have the lowest RMSF values, suggesting that increased additional contacts indeed can reduce the fluctuation of the local secondary structure and of the overall protein structure. This is not surprising: Similar stabilizing behavior was also observed in other protein and peptide assemblies, in which associations through noncovalent interactions between the protein/peptide complexes mutually stabilize each other due to the increased interactions at the interfaces. For example, higher order aggregates of amyloid fibrils generally exhibit more stable molecular structures,\textsuperscript{28,30,31} with the stability increasing with size.

To further identify which interaction could affect the overall stability of the dimeric structure, hydrogen bonds and side chain contacts were calculated to estimate hydrophilic and hydrophobic interactions. A hydrogen bond is identified if the distance between donor D and acceptor A is \(\leq 3.5\) Å and the angle \(D-\text{H}\cdots\text{A}\) is \(\leq 120^\circ\), while a side chain contact is defined if the center-mass-distance between sidechains is less than 6.5 Å.
The side chain of Trp184, which is located in the middle of helix 2, extended into the dimeric interface and formed two hydrogen bonds with Val181 and five side chain contacts with its neighboring Trp184 and Val181 (Figure 3a). Two aromatic rings in Trp184 are closely packed perpendicular to each other to form a T-shape geometry, providing additional π-stacking interactions to stabilize the dimeric interface.32 Substitution of Trp184 by the small alanine not only weakened these interactions considerably (i.e., hydrogen bonds, side chain contacts, and π-stacking), but also disrupted compact geometrical packing, causing a dramatic displacement of one CTD domain shifting ~6 Å relative to the other along the helix 2 axis (Figure 3b).

Interestingly, although the W184A mutant tended to disrupt the dimeric interface, it imposes a little influence on the local secondary structures since all helices were still well preserved during the simulations. In contrast, the M185A substitution caused no gross perturbation of the whole dimeric structure, only subtle changes in the interaction sites (Figure 3c). Since Met185 does not interact with neighboring residues at the interface, an Ala substitution with comparable hydrophobicity relative to Met will not introduce new interactions or considerably lose existing interactions. It should be noted that in the M185A mutant, Trp184 still dominated a large number of interactions with Val181, Thr188, and Leu189 at the dimeric interface of the M185A. Thus it is clear that Trp184 plays an important role in orienting and stabilizing the CA dimer.

Impact of Mutations at the MHR Region (Q155N and E159D) on the Structure and Dynamics of the CA Dimer. The MHR region consisting of 20 amino acids (residues 153–172) forms a compact strand-turn-helix motif.3 Gin155 and Glu159 are located at the edges of the turn and the interactions between Gin155 and Glu159 stabilize the turn structure and restrain the turn motion. It is generally accepted that a turn/loop region is very flexible and thus is vulnerable to mutations.33–35 To minimize the effect of the mutation residue size and hydrophobicity at the target position, Asn was selected to replace Gin155, while Asp was selected to replace Glu159; Gin versus Asn and Glu versus Asp have comparable size and hydrophobicity. Figure 4a clearly shows that both mutants occurring at the MHR region experienced much larger structural deviation than the wild-type, with the RMSD quickly rising to a plateau of ~5.5 Å. As expected, the RMSF profile also showed a similar increased trend with three sharp peaks occurring at the MHR region (residues 155–159), the hinge connecting helix 1 and helix 2 (residues 174–178), and the hinge connecting helix 3 and helix 4 (residues 205–210; Figure 4b). As shown in Figure 5, Gin155 was closely packed against Ala194 and Asn195 at the helix 2 region bridging the MHR and helix 2 motifs. Between the MHR and helix 2 regions, Gin155 formed nine hydrogen bonds with Asn195 and Ala194, while within the MHR region, Gin55 formed three hydrogen bonds with Glu159 (Figure 5a). The Q155N substitution significantly disrupted the hydrogen-bond network of residue 155 with Ala194, Asn195, and Glu159, decreasing from nine (wild-type) to three (Q155N; Figure 5b). The loss of hydrogen bonding and side chain contacts between the MHR and helix 2 regions strongly affects the structural fluctuation of the local turn region (residues 155–159), as indicated by RMSF values rising from 1.81 Å (wild-type) to 4.47 Å (Q155N; Figure 4b). Interestingly, although residue Glu159 is far from the helix 2 region and is not involved in any contact with residues in the helix 2, E159D also led to the loss of two hydrogen bonds and potential van der Waals interactions within the MHR region (Figure 4c). As a consequence, the turn restricted by Gin155 and Asp159 become more flexible and free to move as compared to the turn initially restricted by Gin155 and Glu159 so that hydrogen bonds and side chain contacts between Gin155 and its neighboring Asn195 and Ala194 tend to break and form easily. Taken together, the substitution of Gin155 and Glu159 has a large destabilizing effect not only on the local structure and dynamics, but also on the overall organization of the CA dimer. The structural instability of Gin155 and Glu159 could be attributed to the loss of hydrogen bonds and van der Waals interactions. It is reasonable to expect that substitutions at positions of 155 and 159, other than Q155N and E159D, will lead to even larger structural perturbation in the loop due to the large differences of molecular size and hydrophobicity of sidechains.

Interdomain and Intradomain Motions. To quantitatively monitor specific interactions and relative motions of mutated residues and their neighboring residues, four residue-pair distances in the helix 2 region and three residue-pair distances in the MHR region were selected and measured as a function of simulation time. Figures 6a–c show the intermolecular distances of four identical residue pairs at the dimeric interface between E180I–E180I, Q192I–Q192I, W184I–W184I, or...
A184I–A184II, and M185I–M185II or A185I–A185II, where I and II represent different monomeric CTDs. Glu180 and Gln192 are located at the edges of the helix 2, while Trp184 and Met185 are located in the middle of the helix 2 with all sidechains pointing toward the dimeric interface. Positions 180, 184, 185, and 192, which are almost evenly distributed along helix 2, were selected to monitor the opening-closing motions between two CTDs. As can be seen in Figure 6a, all four residue pairs were able to maintain their initial contacts at the dimeric interface with small distance fluctuations, indicating that the dimeric interface is well preserved during the entire simulations. M185A displayed some distance fluctuation between Q192I–Q192II, but no significant difference was observed between the wild-type and the mutant M185A (Figure 6c), primarily due to that short Met185 did not involve many interactions with its neighboring residues. In contrast, for the mutant W184A, all four residue-pair distances experienced large fluctuations, quickly increasing within the first 10 ns, which indicated that they lost their original contacts at the helix2I–helix2II interface (Figure 6b). These results suggest that the most critical interaction at the dimeric interaction is W184I–W184II; whereas M185I–M185II interaction has little effect on stabilization of the dimer interface.

Within the single CTD, the intramolecular distance is calculated by averaging the mass center distance between selected pair residues for the last 10 ns. The averaged distances between Q155–E159, Q155–A194, and Q155–N195 in the wild-type were 6.4 ± 0.2, 5.1 ± 0.2, and 5.9 ± 0.3 Å, respectively. The Gln155 position is unique because Gln155 not only interacts with Glu159 to stabilize the hydrophobic turn (Gly156–Pro157–Lys158) at the MHR region, but also forms hydrogen bonds with Asn195 and Ala194, to associate with the helix 2 and MHR motifs. It is expected that Gln155 will present a large effect on the stability or conformation of the CA dimer. The Q155N mutant quickly lost interactions with Asn195 and Ala194 at the helix 2 region, as indicated by the continuous increase in pairwise residue distances; however, the interactions between Asn155 and Glu159 were well conserved (Figure 6e). Interestingly, although Glu159 is far from the interface between the MHR and helix 2 motifs and does not interact with Ala194 and Asn195, a replacement of Glu159 with Asp is likely to allosterically perturb interactions near the turn region, with the bent turn shifting slightly away from the MHR–helix2 interface. Considering all three pairwise residue distances among the wild-type and mutants at the MHR region, it is clear that both mutants displayed larger fluctuations of the selected distances than the wild-type, indicating that mutants at the MHR region prefer a flexible conformation.

Figure 4. Structural comparison and characterization of wild-type CA (146–231) and its mutants (Q155N and E159D) at the MHR region. (a) Backbone RMSD relative to the initial energy-minimized structure and (b) residue-based backbone RMSF relative to average structure for the wild-type and mutants.
Comparison of the Structure and Dynamics of the CA (151–231) Dimer with the CA (146–231) Dimer. Gamble and co-workers\(^8\) reported that the binding affinity of CA (146–231) protein dimer is almost the same as that of full-length HIV protein, but they also found that a five-residue shorter CA (151–231) can not be dimerized even at the high concentration up to 100 \(\mu\)M, suggesting that the five missing hydrophobic residues (Ser146–Pro147–Thr148–Ser149–Ile150) could play an important role in the CA dimer formation. We therefore performed an additional simulation to investigate the stability and dynamics of the shorter CA (151–231) dimer for comparison. Consistent with experimental observation,\(^8,16\) the simulation shows that the shorter CA (151–231) dimer displayed its instability with a continuously increasing RMSD up to 4.5 Å, as compared to the RMSD of 2.8 Å for the stable CA (146–231) dimer (Figure 6). Structural comparison shows that the five residues (Ser146–Ile150) in the CA protein are sitting in the middle of a pocket formed by the helix2I–helix2II motifs for (a) wild-type, (b) W184A, and (c) M185A, where I and II indicate different CTD monomers. Variations of pairwise residue distances at the selected positions of 192I–192II, 185I–185II, 184I–184II, and 180I–180II between the helix2I–helix2II motifs for (a) wild-type, (b) W184A, and (c) M185A, where I and II indicate different CTD monomers. Variations of pairwise residue distances at the selected positions of 155–159, 155–194, and 156–195 within the same CTD monomer for (d) wild-type (146–231), (e) Q155N, and (f) E159D.

Table 2. Average Hydrogen Bonds, Sidechain Contacts, and Interaction Energies between Two Monomeric CTDs for the last 10 ns

| System       | Hydrogen Bonds\(^a\) | Sidechain Contact\(^b\) | Interaction Energy (kcal/mol) |
|--------------|----------------------|-------------------------|-------------------------------|
| Wild-type (146–231) | 8.7 ± 1.9 | 26.9 ± 5.9 | -131.8 ± 10.2 |
| Wild-type (151–231) | 7.0 ± 1.5 | 22.0 ± 3.0 | -87.9 ± 9.1 |
| W184A (146–231) | 4.2 ± 1.2 | 19.2 ± 4.1 | -85.4 ± 7.3 |
| M185A (146–231) | 8.2 ± 2.2 | 24.9 ± 3.3 | -90.9 ± 10.2 |
| Glu155Asp (146–231) | 4.0 ± 2.2 | 15.6 ± 4.7 | -62.3 ± 14.0 |
| Gln155Asn (146–231) | 4.7 ± 2.1 | 19.3 ± 3.0 | -63.8 ± 8.4 |

\(^a\) Hydrogen bond is identified if the distance between donor D and acceptor A is < 3.5 Å and the angle \(D-H\cdots A\) is < 120°. \(^b\) Side-chain contact is defined if the center of mass distance between two side chains is less than 6.5 Å.

Table 3. Interacting Contact Network within 6.5 Å Radius of Three Residues (T148–S149–I150) with Helix 2 and Helix 3

| Distance (Å) | Helix 2\(^i\) | Helix 2\(^ii\) | Helix 3\(^i\) |
|-------------|-------------|-------------|-------------|
| T148\(^i\) | M185 (5.5) | W184 (2.2) | K203 (6.1) |
| S149\(^i\) | W184 (5.8) | Q187 (3.1) | L205 (2.9) |
| I150\(^i\) | V181 (6.1) | W184 (4.7) | K203 (6.4) |

Overall CTD–CTD Interactions and Allosteric Effects. The stability of the CTD dimeric complex is determined by direct interactions between the proteins like hydrogen bonds, salt bridge, hydrophobic effects, and van der Waals contacts.\(^31\)
To provide a more quantitative identification of the driving force underlying the association of CTDs, the interactions between the two CTDs were calculated and compared among the wild-type and four mutants (Figure 8). The average interaction energies for the last 20 ns were $-131.8 \pm 10.2$, $-87.9 \pm 9.1$, $-85.4 \pm 7.3$, $-90.9 \pm 10.2$, $-62.3 \pm 14.0$, and $-63.5 \pm 8.4$ kcal/mol for the wild-type (146–231), wild-type (151–231), W184A, M185A, Q155N, and E159D, respectively (Table 2 and Figure 8). The wild-type (146–231) has the most favorable, lowest interaction energy consistent with previous structural analysis, whereas the Q155N and E159D mutants have comparable unfavorable free energies. Comparing these interactions among stable wild-type and unstable mutants suggests that the overall CTD–CTD interactions are mainly controlled by van der Waals interactions through side chain contacts, although hydrogen bonds at the interface also contribute to the stabilizing force to some extent. It is clear that the disruption of the specific interactions by Trp184, Gln155, and Glu159 is responsible for the loss of the total favorable interactions between the CA dimer, leading to a tendency to dissociate. Subtle change in the MHR region will result in large unfavorable interactions for the whole dimer, implying that the MHR region is an allosteric potential mutation target for destabilizing the dimeric CA structure and thus for preventing HIV replication. Comparison of the interaction energy among unstable mutants implies a more significant loss of interactions in the MHR than in the helix 2 region, consistent with the structural analysis described above.

**Biological Implications.** Because HIV capsid assembly depends on both CTD–CTD dimerization and NTD–NTD hexamerization, the introduction of site mutations to lower the binding affinity/structural stability of the C-terminal dimerization and N-terminal hexamerization is essential for preventing virus replication. 36 The CA dimeric interface is mainly associated through hydrophobic helix2I–helix2II interactions, consisting of a cluster of four hydrophobic residues (Leu189, Met185, Trp184, and Val181) in the middle of helix 2, and one polar residue (Gln192) and one acidic residue (Glu180) at both ends of helix 2. Hydrophobic Trp184 and Met185 are deeply buried in the dimer interface. Simulations show that the large aromatic Trp184 imposes strong stability by restricting the motions of nearby residues. When the large Trp was replaced by the Ala at position 184, the smaller residue can not attain sufficient packing interactions with its neighboring residues, leading to the increase of the structural flexibility of the dimeric interface due to the loss of interactions. Alcaraz et al. 15 and Ganser-Pornillos et al. 5 also observed that HIV capsid assembly is abolished when the sole Trp184 is mutated to Ala. Unlike Trp184, structural analysis of M185A does not show a significant difference from the wild-type 146–231, but the binding free energy of the dimeric complex lost $\sim 40.9$ kcal/mol as compared to the wild-type, implying that some interactions could be lost upon the mutation. In vitro experiments by Sitch et al. 1 reported that the binding affinity of a capsid assembly inhibitor (CAI) for the M185A mutant was reduced but not completely abolished in ELISAs. Alamo et al. also found that Met185 did not contribute to stability, 37 in agreement with our simulation results. Furthermore, Alanine scanning mutagenesis at the helix 2 region showed that Ala substitutions at Gln180, Glu187, and Glu192 unexpectedly increase the affinity of the CA dimer interface, leading to more stable dimers. 15,37,38 This is not surprising: when charged Glu residues are replaced by hydrophobic Ala, unfavorable electrostatic repulsions between Glu180–Glu180 and Glu187–Glu187 pairs are eliminated and replaced by more
favorable hydrophobic interactions. The Gln192 residues are located at the end of the helix 2 with the Gln192 side chains substantially exposed to water. The hydrophilic Gln192–Gln192 interactions are weak, attacked by water molecules and easily broken. This fact also indicates that hydrophobic interactions are the major, dominant driving force in stabilizing the CA interface.

Interestingly, the simulation results show that helix 2 is not the only important motif that can be used to disrupt the CA dimeric interface. It appears that the MHR is even more venerable to mutation than the helix 2, via long-range allosteric effects. Similar allosteric effects on the HIV-1 protease flap that have also been observed.39–42 Substitutions of conserved Gln155 by Asn and Glu159 by Asp only result in subtle changes in the initial structures at the MHR because Gln vs Asn and Glu vs Asp have similar molecular size and hydrophobicity; thus, it is expected that these mutations would neither introduce large new interactions, nor alter backbone conformation. However, the simulations reveal the unexpected consequence of dramatically reducing the binding affinity of CA dimerization, although the MHR is far from the CA dimer interface. The MHR motif contains a five-residue loop (Gln155–Gly156–Pro157–Lys158–Glu159) that covalently links the helix 1 and helix 2 motifs. The motion and conformation of this loop is mainly restricted by the four most highly conserved residues of Gln155, Gly156, Glu159, and Arg167.46 Especially, Gln155 is a key residue that bridges Glu159 at the helix 1 with Asn195 and Ala194 at the helix 2 via hydrogen bonding (Figure 5). The disruption of this hydrogen bonding network (Figure 5) could permit the motion of the loop that exerts a stretching force to pull the helix 2 away from the CA interface, destabilizing the binding of the CA dimer. Gamble and co-workers8 found that Q155N and E159D mutations can block HIV-1 replication,43 indicating that the distinct motifs in the CA could perform the same function.

Conclusions
The CTD dimer interface is required for HIV capsid assembly. Here, we have performed a series of MD simulations to study the dynamics and stability of the HIV CA C-terminal domain (146–231) dimer and its atomic-scale structural changes upon single-point mutagenesis at the helix 2 and MHR regions. Four mutation sites are carefully selected, Trp184 and M185 which are deeply buried in the helix 2 region and Gln155 and Glu159 which regulate the loop motion at the MHR region. MD simulations show that the disruption of the wild-type CA C-terminal domain can destabilize the CA–CA dimerization and thus inhibit HIV viral assembly. The Q155N and E159D mutations at the MHR greatly destabilize the binding affinity of the CA dimer by disrupting the hydrogen bonding network between the loop and the helix 1 and 2 motifs. This loop at the MHR is a major target leading to the conformational changes and unfolding of the CTD. At the helix 2 region, the W184A mutation is known to abolish HIV CA protein assembly in vitro. Substitution of the large, hydrophobic, and aromatic Trp184, which is deeply buried in the helix 2 region by the small Ala, not only weakens the hydrogen bonds, side chain contacts, and π stacking, but also disrupts compact geometrical packing, causing a dramatic destabilization of the CA dimeric interface as well. It appears that structural changes in the distinct motifs of the CA could similarly prevent the CA capsid formation, although the allosteric mutations at the MHR region affect the structural stability and in particular the binding affinity of the CA dimer more than those occurring at the helix 2. Examination of the relationships between the unstable mutants and the stable wild-type and between structure and function suggest that the hydrophobic interaction, side-chain packing, and hydrogen bonds are important driving forces in stabilizing the CA dimer.

It is desirable to directly address the energetics, physicochemical stability, and folding of whole HIV capsid assembly with both N- and C-terminal domains. We are currently applying a new generic computational method based on symmetrical assembly and 2D → 3D wrapping algorithm44–46 to construct the HIV protein tube without end or edge effects, and comparing our structures with the full-length structure of HIV-1 CA hexamer (Figure 9) as recently determined by Ganser-Pornillos and co-workers.14 This ongoing study will provide more detailed insights into the mechanism of the HIV CA capsid formation.

Acknowledgment. This work is supported by American Chemical Society Petroleum Research Fund (48188-G5), Faculty
(42) Carlsson, J.; Boukharta, L.; Aqvist, J. Combining docking, molecular dynamics and the linear interaction energy method to predict binding modes and affinities for non-nucleoside inhibitors to HIV-1 reverse transcriptase. *J. Med. Chem.* 2008, 51 (9), 2648–2656.

(43) Thali, M.; Bukovsky, A.; Kondo, E.; Rosenwirth, B.; Walsh, C. T.; Sodroski, J.; Gottlinger, H. G. Functional association of cyclophilin A with HIV-1 virions. *Nature* 1994, 372 (6504), 363–365.

(44) Tsai, C.-J.; Zheng, J.; Nussinov, R. Designing a nanotube using naturally occurring protein building blocks. *PLoS Comput. Biol.* 2006, 2 (4), e42.

(45) Tsai, C.-J.; Zheng, J.; Aleman, C.; Nussinov, R. Structure by design: from single proteins and their building blocks to nanostructures. *Trends Biotechnol.* 2006, 24 (10), 449–454.

(46) Tsai, C.-J.; Zheng, J.; Zanuy, D.; Haspel, N.; Wolfson, H.; Aleman, C.; Nussinov, R. Principles of nanostructure design with protein building blocks. *Proteins: Struct., Funct., Bioinf.* 2007, 68 (1), 1–12.

(47) Humphrey, W.; Dalke, A.; Schulten, K. VMD—Visual molecular dynamics. *J. Mol. Graphics* 1996, 14 (1), 33–38.

BM801151R