

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 π -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 with the CA (146–231) dimer further indicates that hydrophobic interactions, side chain packing, and hydrogen bonds are the major, dominant driving forces 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.^{1,4–6} 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 distinct 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 α -helices and an extended proline-loop^{5,8–10} is responsible for viral capsid formation,⁸ while the CTD (residues 148–231) consisting of

four α -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-workers¹³ 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

* To whom correspondence should be addressed. E-mail: ruthn@ncifcrf.gov (R.N.); zhengj@uakron.edu (J.Z.).

[†] These authors contributed equally to this work.

[‡] Department of Chemical and Biomolecular Engineering, The University of Akron.

[§] Tel Aviv University.

^{||} NCI-Frederick.

[⊥] Department of Polymer Science, The University of Akron.

Table 1. General Properties of Simulation Systems

system	mutation region	RMSD ^a			Rg ^a		
		dimer	monomer	monomer	dimer	monomer	monomer
wild-type (146–231)		2.8 ± 0.4	2.3 ± 0.3	2.4 ± 0.3	15.8 ± 0.3	11.4 ± 0.2	11.3 ± 0.1
wild-type (151–231)		4.5 ± 0.3	2.8 ± 0.5	3.8 ± 0.3	15.6 ± 0.3	11.2 ± 0.1	11.3 ± 0.1
Trp184Ala (146–231)	helix 2	6.7 ± 0.5	2.5 ± 0.2	3.4 ± 0.3	16.5 ± 0.3	11.7 ± 0.1	11.7 ± 0.2
Met185Ala (146–231)	helix 2	3.1 ± 0.4	1.9 ± 0.2	2.0 ± 0.2	15.4 ± 0.2	11.5 ± 0.1	11.6 ± 0.2
Gln155Asn (146–231)	MHR	5.9 ± 0.8	5.0 ± 0.6	3.6 ± 0.3	16.8 ± 0.5	11.2 ± 0.1	11.5 ± 0.2
Glu159Asp (146–231)	MHR	5.4 ± 0.5	2.6 ± 0.2	3.8 ± 0.6	16.7 ± 0.2	11.7 ± 0.1	11.3 ± 0.3

^a RMSD and Rg values were averaged for the last 10 ns.

variety of approaches including nuclear magnetic resonance (NMR),⁶ X-ray diffraction,^{8,16} and cryo-electron microscopy (EM),^{5,9,13,17} as well as theoretical approaches including molecular dynamics (MD) simulation and normal-mode analysis (NMA).^{7,18} 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.⁵ tested 21 mutations in the NTDs and 6 mutations at the CTDs to examine the assembly ability of CA proteins into helical capsid in vitro. They found that only M39A and A42D at the NTDs and W184A at the CTDs severely attenuated the formation of cylindrical capsid, thus reducing virus infectivity. Mammano et al.¹⁹ reported that conservative mutation of two invariant residues at the major homology region (MHR) of the CTDs (Q155N and E159D) significantly abolished viral replication. Gamble et al.⁸ also confirmed the importance of residues Trp184 and Met185 in stabilizing the CTDs of the CA proteins. Thus, the disruption of either hexameric organization formed by six NTDs or dimeric interface associated by two CTDs correlates with the prevention of HIV capsid formation. However, detailed atomic-resolution structures of these mutants and the corresponding structural correlation of the substituted residue with its neighboring residues and the whole virus capsid have not been reported. In this work, we have performed multiple all-atom, explicit-solvent MD simulations to examine the structural stability and dynamics of CTDs of the CA protein for the wild-type and mutated sequences, particularly focusing on the impact of mutations on the interactions at the dimeric interface. Four mutations at the dimeric interface (W184A and M185A at the helix 2 region and Q155N and E159D at the MHR) were carefully selected to examine the effect of specific single point mutation on the association force, structural stability, and folding of CTDs of the CA protein (Table 1). These simulations allow us to obtain detailed atomic-level information about the dynamic behavior of the C-terminal domain interface of HIV-1 and its underlying driving force, which is critical for stabilizing the dimeric structure; in turn, these could allow us to identify pharmacological targets to block the capsid formation.

Simulation Methods. Model Systems. 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

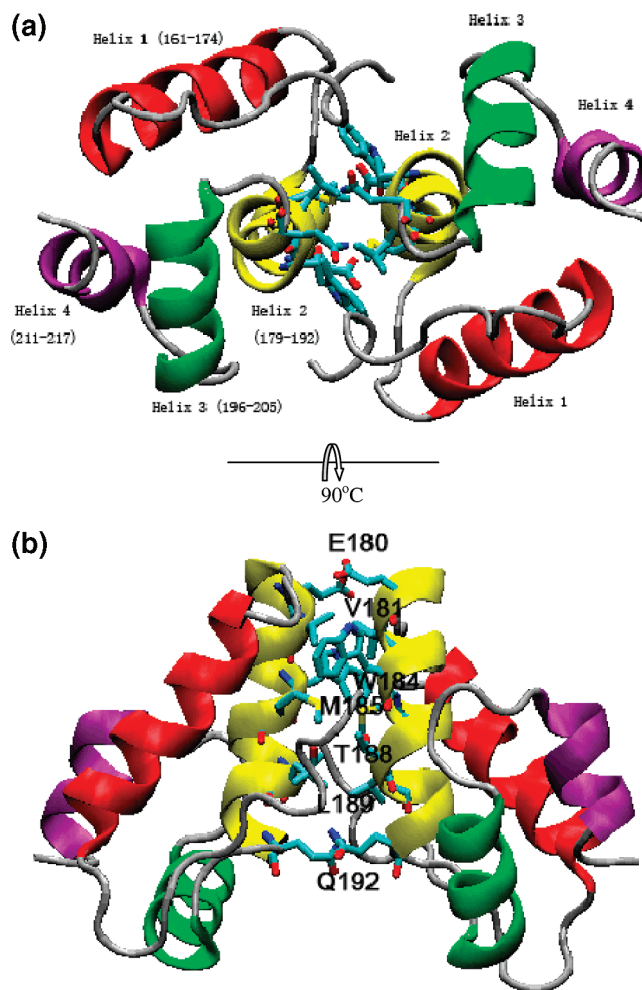


Figure 1. Crystal dimeric structure of the C-terminal domain of CA protein (146–231; PDB code 1A43). (a) Secondary structure of CA dimer in cartoon representation. Each monomer consists of four helices: helix 1 (161–174; red), helix 2 (179–192; yellow), helix 3 (196–205; green), and helix 4 (211–217; 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 Gln192). The images are generated by VMD.⁴⁷

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 (residues 161–174), helix 2 (residues 179–192), helix 3 (residues 196–205), and helix 4 (residues 211–217). Two CTDs form a dimeric interface via helix 2 packing along a C₂ symmetry (Figure 1). All starting structures of the mutants were built from the wild-type by replacing the side chains of the targeted residues but without changing the backbone conformations and side-chain orientations. Four mutations were used in the simulations: W184A and M185A located in helix 2 region, and Q155N and E159D located in

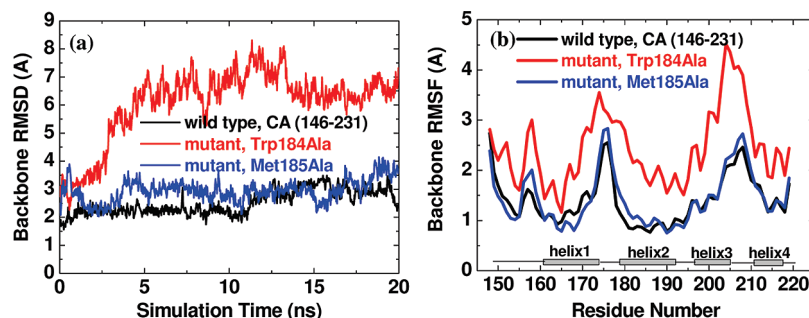


Figure 2. Structural comparison and characterization of wild-type CA (146–231) and its mutants (W184A and M185A) at the helix 2 region. (a) Backbone RMSD relative to the initial energy-minimized structure and (b) residue-based backbone RMSF relative to the average structure for the wild-type and mutants. The rmsd shows structural deviation for the whole structure, while the RMSF shows structural fluctuation for each individual residue.

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 program²⁰ with the CHARMM27 force field.^{21–23} Each system was solvated by a TIP3P²⁴ 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).^{25,26} Thus, the interaction energy between two dimers was calculated by

$$\langle \Delta G_{\text{dimer}} \rangle = E_{\text{dimer}}^{\text{mm}} + G_{\text{dimer}}^{\text{gbsw}} - (E_{\text{monomer1}}^{\text{mm}} + G_{\text{monomer1}}^{\text{gbsw}}) - (E_{\text{monomer2}}^{\text{mm}} + G_{\text{monomer2}}^{\text{gbsw}}) \quad (1)$$

where E^{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 β dimer^{27,28} and von Hippel-Lindau tumor suppressor protein complex.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,^{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 ≤ 3.5 Å and the angle D–H...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.

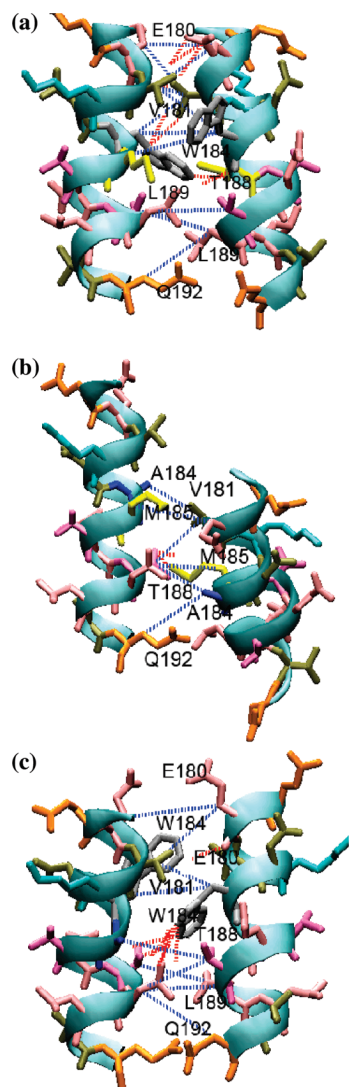


Figure 3. Intermolecular interaction network between the helix2^I–helix2^{II} motifs forming a CA dimeric interface for (a) wild-type (146–231), (b) W184A, and (c) M185A. All interacting residues are labeled by their residue names. Hydrogen bonds are represented by blue dotted lines, while side chain contacts are represented by red dotted lines.

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 consider-

ably 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.⁸ Gln155 and Glu159 are located at the edges of the turn and the interactions between Gln155 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 Gln155, while Asp was selected to replace Glu159; Gln 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, Gln155 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, Gln155 formed nine hydrogen bonds with Asn195 and Ala194, while within the MHR region, Gln155 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 Gln155 and Asp159 become more flexible and free to move as compared to the turn initially restricted by Gln155 and Glu159 so that hydrogen bonds and side chain contacts between Gln155 and its neighboring Asn195 and Ala194 tend to break and form easily. Taken together, the substitution of Gln155 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 Gln155 and Glu159 could be attributed to the loss of hydrogen bonds and van der Waal 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 E180^I–E180^{II}, Q192^I–Q192^{II}, W184^I–W184^{II}, or

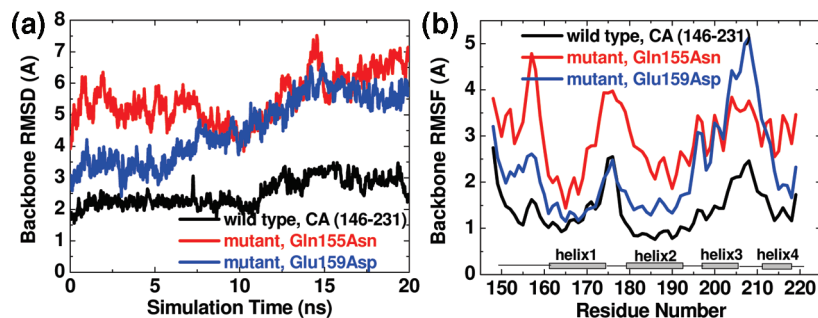


Figure 4. Structural comparison and characterization of wild-type CA (146–231) and its mutants (Q155N and E159D) at the MHR region. (a) Backbone RMSD relative to the initial energy-minimized structure and (b) residue-based backbone RMSF relative to average structure for the wild-type and mutants.

A184^I–A184^{II}, and M185^I–M185^{II} or A185^I–A185^{II}, 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 Q192^I–Q192^{II}, 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 helix2^I–helix2^{II} interface (Figure 6b). These results suggest that the most critical interaction at the dimeric interface is W184^I–W184^{II}; whereas M185^I–M185^{II} 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.

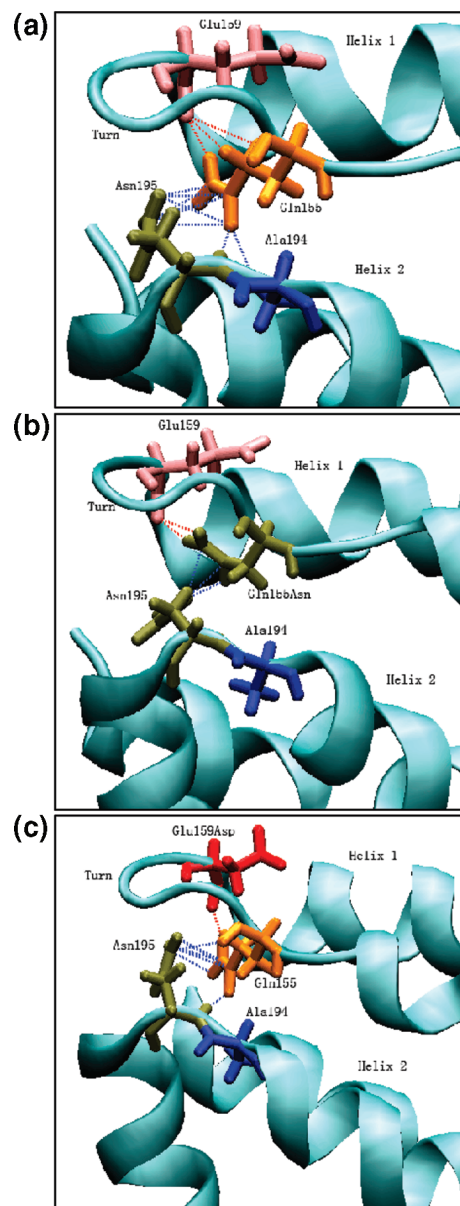


Figure 5. Intramolecular hydrogen bonding network between the turn region (residues 155–159) and helix 2 (residues 179–192) for (a) wild-type (146–231), (b) Q155N, and (c) E159D. Hydrogen bonding interactions between residues 155 and 159 within the MHR region were represented by red dotted lines, while hydrogen bonding interactions between the residue at the position 155 (MHR region) and residues of Ala194 and Asn195 (helix 2 region) were represented by blue dotted lines.

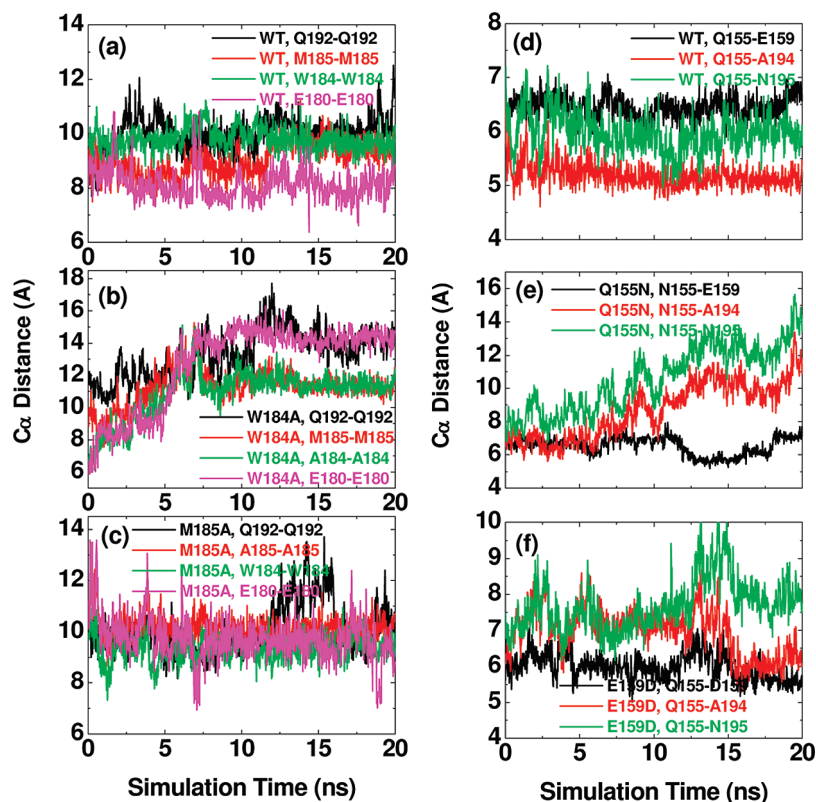


Figure 6. Variations of pairwise interface residue distances at the selected positions of 192^I–192^{II}, 185^I–185^{II}, 184^I–184^{II}, and 180^I–180^{II} between the helix2^I–helix2^{II} motifs for (a) wild-type, (b) W184A, and (c) M185A, where I and II indicate different CTD monomers. Variations of pairwise residue distances at the selected positions of 155–159, 155–194, and 155–195 within the same CTD monomer for (d) wild-type (146–231), (e) Q155N, and (f) E159D.

Comparison of the Structure and Dynamics of the CA (151–231) Dimer with the CA (146–231) Dimer. Gamble and co-workers⁸ 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 μ 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 7a). Structural comparison shows that the five residues (Ser146–Ile150) in the CA protein are sitting in the middle of a pocket formed by the helix2^I–helix2^{II}–helix3^{II}. Within 6.5 Å distance from these five residues, Val181, Trp184, Met185, Gln187, Thr188, Lys189, Leu189, Leu190, Lys203, Leu205 from the helix 2 and helix 3 formed a strong hydrophobic contact network stabilizing the CA (146–231) dimer (Table 3). The absence of these hydrophobic interactions reduces the binding affinity of the two CTDs and the binding energy at the dimeric interface increases from –131.8 kcal/mol for CA (146–231) to –87.9 kcal/mol for CA (151–231; Table 2). It can be clearly seen in Figure 7c that as compared to the CA (146–231) dimer, the loss of native contacts at the dimeric interface (hydrogen bonds and side chain contacts, Table 2) leads to the change in the relative orientation of the two CTDs (151–231), in which one of the CTD monomers is rotated away from the interface.

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
Trp184Ala (146–231)	4.2 ± 1.2	19.2 ± 4.1	–85.4 ± 7.3
Met185Ala (146–231)	8.2 ± 2.2	24.9 ± 3.3	–90.9 ± 10.2
Gln155Asn (146–231)	4.0 ± 2.2	15.6 ± 4.7	–62.3 ± 14.0
Glu159Asp (146–231)	4.7 ± 2.1	19.3 ± 3.0	–63.5 ± 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...A is $\leq 120^\circ$. ^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 179–192	helix 2 ^{II} 179–192	helix 3 ^{II} 196–205
T148 ^I	M185 (5.5)	W184 (2.2) Q187 (3.1) T188 (2.8)	K203 (6.1) L205 (2.9)
S149 ^I		W184 (5.8) Q187 (6.3) T188 (5.6)	K203 (5.2) L205 (4.5)
I150 ^I	V181 (6.1) M185 (2.3) L189 (2.9) L190 (5.8)	W184 (4.7) Q187 (5.5) T188 (2.1) L189 (6.3) Q192 (6.1)	K203 (6.4) L205 (6.5)

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

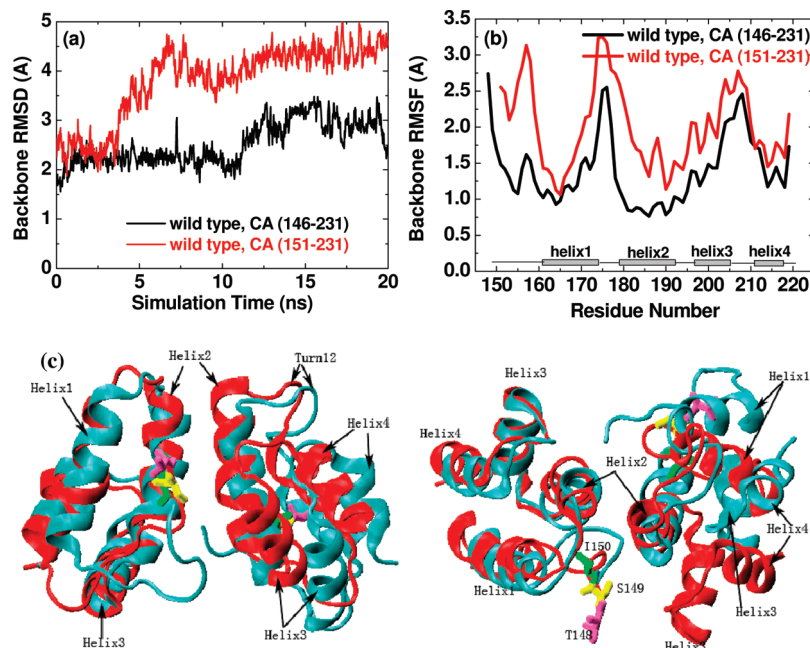


Figure 7. Structural analysis between wild-type CA (146–231) and truncated five-residue short CA (151–231). (a) Backbone RMSD relative to initial energy-minimized structure, (b) residue-based backbone RMSF relative to average structure, and (c) structural superimposition of CA (146–231, red color) with CA (151–231, cyan color). One CA (151–231) monomer tends to rotate away from the other.

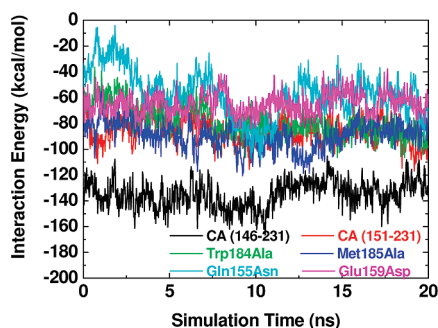


Figure 8. Interaction energies between two facing CA monomers for all wild-types and mutants.

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 ± 10.2 , -87.9 ± 9.1 , -85.4 ± 7.3 , -90.9 ± 10.2 , -62.3 ± 14.0 , and -63.5 ± 8.4 cal/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 helix2^I–helix2^{II} 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 ~ 40.9 kcal/mol as compared to the wild-type, implying that some interactions could be lost upon the mutation. In vitro experiments by Sicht 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 Glu180, Glu187, and Gln192 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

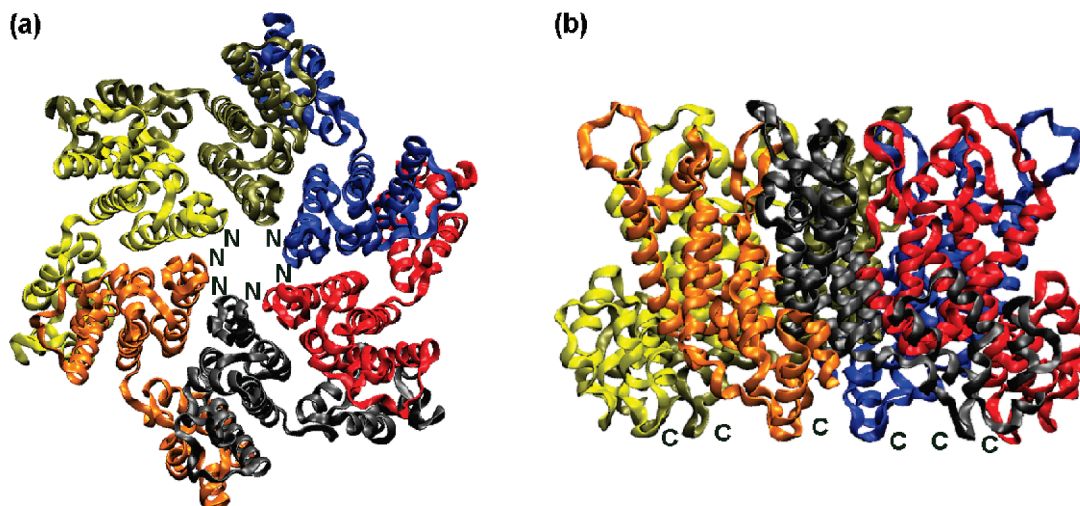


Figure 9. Structure of the HIV-1 CA hexamer. The six monomers are highlighted in different colors. Each monomer has 219 residues and includes both the NTD (1–145) and the CTD (146–219). Two perspectives are shown: (a) as viewed from the exterior of the viral capsid. The CTD of each monomer hides behind the NTD of the neighboring monomer. Thus, only the NTDs are seen clearly. The N-termini are located in the central cavity, close to the outer surface of the capsid. (b) The same hexamer with the six-fold axis vertical. Both the NTDs and the CTDs of the monomers are presented.

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 is a target for drug design 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 Arg167.¹⁶ 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 mutants 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, physico-chemical 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 \rightarrow 3D wrapping algorithm^{44–46} 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

Research Summer Fellowship, and Firestone Research Program at the University of Akron. This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under Contract N01-CO-12400. This research was supported (in part) by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

References and Notes

- Sticht, J.; Humbert, M.; Findlow, S.; Bodem, J.; Muller, B.; Dietrich, U.; Werner, J.; Krausslich, H.-G. A peptide inhibitor of HIV-1 assembly in vitro. *Nat. Struct. Mol. Biol.* **2005**, *12* (8), 671–677.
- Sundquist, W. I.; Hill, C. P. How to assemble a capsid. *Cell* **2007**, *131* (1), 17–19.
- Ganser-Pornillos, B. K.; Yeager, M.; Sundquist, W. I. The structural biology of HIV assembly. *Curr. Opin. Struct. Biol.* **2008**, *18* (2), 203–217.
- Douglas, C. C.; Thomas, D.; Lanman, J.; Prevelige, P. E. Investigation of N-terminal domain charged residues on the assembly and stability of HIV-1CA. *Biochemistry* **2004**, *43* (32), 10435–10441.
- Ganser-Pornillos, B. K.; von Schwedler, U. K.; Stray, K. M.; Aiken, C.; Sundquist, W. I. Assembly properties of the human immunodeficiency virus type 1 CA protein. *J. Virol.* **2004**, *78* (5), 2545–2552.
- Gitti, R. K.; Lee, B. M.; Walker, J.; Summers, M. F.; Yoo, S.; Sundquist, W. I. Structure of the amino-terminal core domain of the HIV-1 capsid protein. *Science* **1996**, *273* (5272), 231–235.
- Tama, F.; Brooks Iii, C. L. Diversity and identity of mechanical properties of icosahedral viral capsids studied with elastic network normal mode analysis. *J. Mol. Biol.* **2005**, *345* (2), 299–314.
- Gamble, T. R.; Yoo, S.; Vajdos, F. F.; von Schwedler, U. K.; Worthylake, D. K.; Wang, H.; McCutcheon, J. P.; Sundquist, W. I.; Hill, C. P. Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. *Science* **1997**, *278* (5339), 849–853.
- Li, S.; Hill, C. P.; Sundquist, W. I.; Finch, J. T. Image reconstructions of helical assemblies of the HIV-1 CA protein. *Nature* **2000**, *407* (6802), 409–413.
- Ternois, F.; Sticht, J.; Duquerroy, S.; Krausslich, H.-G.; Rey, F. A. The HIV-1 capsid protein C-terminal domain in complex with a virus assembly inhibitor. *Nat. Struct. Mol. Biol.* **2005**, *12* (8), 678–682.
- Franke, E. K.; Yuan, H. E.; Bossolt, K. L.; Goff, S. P.; Luban, J. Specificity and sequence requirements for interactions between various retroviral Gag proteins. *J. Virol.* **1994**, *68* (8), 5300–5305.
- Reicin, A. S.; Ohagen, A.; Yin, L.; Hoglund, S.; Goff, S. P. The role of Gag in human immunodeficiency virus type 1 virion morphogenesis and early steps of the viral life cycle. *J. Virol.* **1996**, *70* (12), 8645–8652.
- Briggs, J. A. G.; Wilk, T.; Welker, R.; Krüsslich, H.-G.; Fuller, S. D. Structural organization of authentic, mature HIV-1 virions and cores. *EMBO J.* **2003**, *22* (7), 1707–1715.
- Ganser-Pornillos, B. K.; Cheng, A.; Yeager, M. Structure of full-length HIV-1 CA: A model for the mature capsid lattice. *Cell* **2007**, *131* (1), 70–79.
- Alcaraz, L. A.; Alamo, M. d.; Mateu, M. G.; Neira, J. L. Structural mobility of the monomeric C-terminal domain of the HIV-1 capsid protein. *FEBS J.* **2008**, *275* (13), 3299–3311.
- Gamble, T. R.; Vajdos, F. F.; Yoo, S.; Worthylake, D. K.; Houseweart, M.; Sundquist, W. I.; Hill, C. P. Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. *Cell* **1996**, *87* (7), 1285–1294.
- Briggs, J. A. G.; Grunewald, K.; Glass, B.; Forster, F.; Krausslich, H.-G.; Fuller, S. D. The mechanism of HIV-1 core assembly: insights from three-dimensional reconstructions of authentic virions. *Structure* **2006**, *14* (1), 15–20.
- Verkhivker, G.; Tiana, G.; Camilloni, C.; Provasi, D.; Broglia, R. A. Atomistic simulations of the HIV-1 protease folding inhibition. *Biophys. J.* **2008**, *95* (2), 550–562.
- Mammano, F.; Ohagen, A.; Höglund, S.; Göttlinger, H. G. Role of the major homology region of human immunodeficiency virus type 1 in virion morphogenesis. *J. Virol.* **1994**, *68* (8), 4927–4936.
- Kale, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. NAMD2: Greater scalability for parallel molecular dynamics. *J. Comput. Phys.* **1999**, *151* (1), 283–312.
- Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM—A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **1983**, *4* (2), 187–217.
- MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998**, *102* (18), 3586–3616.
- MacKerell Jr, A. D.; Banavali, N.; Foloppe, N. Development and current status of the CHARMM force field for nucleic acids. *Biopolymers* **2000**, *56* (4), 257–265.
- William, L. J.; Jayaraman, C.; Jeffrey, D. M.; Roger, W. I.; Michael, L. K. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79* (2), 926–935.
- Im, W.; Lee, M. S.; Brooks, C. L., III. Generalized born model with a simple smoothing function. *J. Comput. Chem.* **2003**, *24* (14), 1691–1702.
- Im, W.; Feig, M.; Brooks, C. L. An implicit membrane generalized born theory for the study of structure, stability, and interactions of membrane proteins. *Biophys. J.* **2003**, *85* (5), 2900–2918.
- Zheng, J.; Jang, H.; Ma, B.; Tsai, C.-J.; Nussinov, R. Modeling the Alzheimer a{beta}17–42 fibril architecture: Tight intermolecular sheet-sheet association and intramolecular hydrated cavities. *Biophys. J.* **2007**, *93* (9), 3046–3057.
- Zheng, J.; Jang, H.; Ma, B.; Nussinov, R. Annular structures as intermediates in fibril formation of alzheimer A[beta]17–42. *J. Phys. Chem. B* **2008**, *112* (22), 6856–6865.
- Liu, J.; Nussinov, R. Allosteric effects in the marginally stable von Hippel-Lindau tumor suppressor protein and allostery-based rescue mutant design. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *105* (3), 901–906.
- Zheng, J.; Jang, H.; Nussinov, R. [beta]2-microglobulin amyloid fragment organization and morphology and its comparison to A-[beta] suggests that amyloid aggregation pathways are sequence-specific. *Biochemistry* **2008**, *47* (8), 2497–2509.
- Zheng, J.; Ma, B.; Tsai, C.-J.; Nussinov, R. Structural stability and dynamics of an amyloid-forming peptide G2NQQNY from the yeast prion sup-35. *Biophys. J.* **2006**, *91* (3), 824–833.
- Gazit, E. A possible role for pi-stacking in the self-assembly of amyloid fibrils. *FASEB J.* **2002**, *16* (1), 77–83.
- Zheng, J.; Zanuy, D.; Haspel, N.; Tsai, C. J.; Aleman, C.; Nussinov, R. Nanostructure design using protein building blocks enhanced by conformationally constrained synthetic residues. *Biochemistry* **2007**, *46* (5), 1205–1218.
- Haspel, N.; Zanuy, D.; Zheng, J.; Aleman, C.; Wolfson, H.; Nussinov, R. Changing the charge distribution of {beta}-helical-based nanostructures can provide the conditions for charge transfer. *Biophys. J.* **2007**, *93* (1), 245–253.
- Zheng, J.; Ma, B.; Chang, Y.; Nussinov, R. Molecular dynamics simulations of alzheimer's peptide A[beta]40 elongation and lateral association. *Front. Biosci.* **2008**, *13* (1), 3919–3930.
- Lidon-Moya, M. C.; Barrera, F. N.; Bueno, M.; Perez-Jimenez, R.; Sancho, J.; Mateu, M. G.; Neira, J. L. An extensive thermodynamic characterization of the dimerization domain of the HIV-1 capsid protein. *Protein Sci.* **2005**, *14* (9), 2387–2404.
- del Alamo, M.; Neira, J. L.; Mateu, M. G. Thermodynamic dissection of a low affinity protein-protein interface involved in human immunodeficiency virus assembly. *J. Biol. Chem.* **2003**, *278* (30), 27923–27929.
- Alcaraz, L. A.; del Alamo, M.; Barrera, F. N.; Mateu, M. G.; Neira, J. L. Flexibility in HIV-1 assembly subunits: solution structure of the monomeric C-terminal domain of the capsid protein. *Biophys. J.* **2007**, *93* (4), 1264–1276.
- Damm, K. L.; Ung, P. M. U.; Quintero, J. J.; Gestwicki, J. E.; Carlson, H. A. A poke in the eye: Inhibiting HIV-1 protease through its flap-recognition pocket. *Biopolymers* **2008**, *89* (8), 643–652.
- Ding, F.; Layten, M.; Simmerling, C. Solution structure of HIV-1 protease flaps probed by comparison of molecular dynamics simulation ensembles and EPR experiments. *J. Am. Chem. Soc.* **2008**, *130* (23), 7184–7185.
- Hornak, V.; Okur, A.; Rizzo, R. C.; Simmerling, C. HIV-1 protease flaps spontaneously close to the correct structure in simulations following manual placement of an inhibitor into the open state. *J. Am. Chem. Soc.* **2006**, *128* (9), 2812–2813.

- (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.; Rosenwlrth, 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