

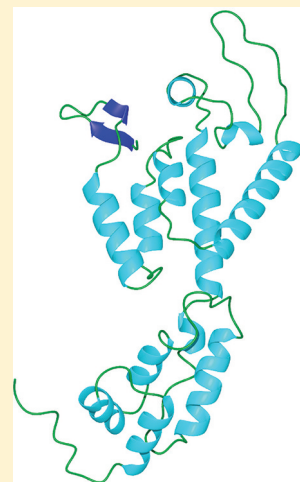
Structure of a Monomeric Mutant of the HIV-1 Capsid Protein

Ronald Shin,[†] Ywh-Min Tzou,[‡] and N. Rama Krishna^{*,†,‡}

[†]Comprehensive Cancer Center and [‡]Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham, Birmingham, Alabama 35294, United States

S Supporting Information

ABSTRACT: The capsid protein (CA) of HIV-1 plays a significant role in the assembly of the immature virion and is the critical building block of its mature capsid. Thus, there has been significant interest in the CA protein as a target in the design of inhibitors of early and late stage events in the HIV-1 replication cycle. However, because of its inherent flexibility from the interdomain linker and the monomer–dimer equilibrium in solution, the HIV-1 wild-type CA monomer has defied structural determinations by X-ray crystallography and nuclear magnetic resonance spectroscopy. Here we report the detailed solution structure of full-length HIV-1 CA using a monomeric mutant that, though noninfective, preserves many of the critical properties of the wild-type protein. The structure shows independently folded N-terminal (NTD) and C-terminal domains (CTD) joined by a flexible linker. The CTD shows some differences from that of the dimeric wild-type CTD structures. This study provides insights into the molecular mechanism of the wild-type CA dimerization critical for capsid assembly. The monomeric mutant allows investigation of interactions of CA with human cellular proteins exploited by HIV-1, directly in solution without the complications associated with the monomer–dimer equilibrium of the wild-type protein. This structure also permits the design of inhibitors directed at a novel target, viz., interdomain flexibility, as well as inhibitors that target multiple interdomain interactions critical for assembly and interactions of CA with host cellular proteins that play significant roles within the replication cycle of HIV-1.



Retroviruses typically consist of a central capsid core particle encapsulating two copies of RNA and the viral enzymes. The capsids are composed of ~1500 copies of a capsid protein (CA) that is initially part of a Gag polypeptide synthesized in the infected host cell.^{1,2} The retroviral capsid proteins are typically ~24–27 kDa in size and are highly α -helical. The Gag proteins capture the viral RNA, assemble either in the cytosol (B- and D-type retroviruses) or at the cell membrane (C-type, HTLV/BLV, and lentiviruses), and bud into the enveloped immature virus particles.² Gag is then proteolytically cleaved by the viral protease into the major structural proteins of the virus,^{2,3} followed by a maturation process in which the capsid proteins condense to form the mature capsid of the virus with a distinct shape characteristic of the genus.³ The HIV-1 capsid is a conical-shaped fullerene structure.⁴

The capsid protein (CA) of HIV-1 plays a significant role in the early stages of the viral life cycle, controlling the virion size, morphology, and Gag assembly.^{5–7} Electron cryotomography images of the immature virions have shown that along with the spacer SP1, CA domains also play an important role in the formation of the hexameric Gag lattice.⁵ Most importantly, intermolecular CTD–CTD interactions appear to be important in the assembly of the hexameric Gag lattice.^{5,7} Electron microscopy studies show that the mature capsid of HIV-1 is a fullerene cone, with its surface composed primarily of hexameric CA rings, with 12 pentameric rings of CA that allow the cone to close at both ends.⁴ In fact, the surface of the mature capsid consists of hexameric (and pentameric) rings of

the N-terminal domains (NTDs) stabilized by NTD–NTD interactions, with each ring linked to neighboring hexamers through the interhexamer dimerization of the C-terminal domains (CTDs). Additional intermolecular NTD–CTD and CTD–CTD interactions further stabilize the mature capsid surface lattice.^{1,3,8,9} Thus, because of the critical role of CA in the assembly of the immature particles and mature capsids, recently there has been rather significant interest in the CA protein as an antiviral therapeutic target for the design of inhibitors of early and late stage events in the HIV-1 replication cycle.^{1,4,10–15} Thus, the availability of the structure of the full-length HIV-1 CA monomer would be of critical importance for efforts in the structure-based design of inhibitors. Such a monomeric structure will also facilitate a structural biological characterization of the interactions of the HIV-1 capsid protein with host cell proteins exploited by HIV-1 in its replication cycle, such as cyclophilin A and lysyl-tRNA synthetase.

However, HIV-1 wild-type full-length CA monomer protein has defied structural determinations by X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy because of the high degree of flexibility of the interdomain linker that made it difficult to crystallize, and the monomer–dimer equilibrium in solution that resulted in exchange broadening and the disappearance of many peaks from the CTD because of

Received: July 26, 2011

Revised: September 26, 2011

Published: October 13, 2011



its reversible CTD–CTD dimerization. Thus, efforts have focused on the structural determinations by crystallography or NMR spectroscopy of isolated domains, viz., the NTD,^{16,17} the dimeric structures of the isolated wild-type CTD,^{8,18} a domain-swapped CTD dimer,¹⁹ and the structure of the isolated monomeric mutants of the CTD.^{20,21} For the full-length wild-type CA protein, three crystallographic studies of CA dimers have been reported: a parallel dimer of CA with the NTD stabilized by complexation with Fab but with a disordered CTD²² and antiparallel head-to-tail dimers of CA stabilized by complexation with Fab²³ or triiodide.²⁴ These dimer structures of the wild-type (wt) CA exhibit intermolecular CA–CA interactions and associated structural perturbations.

In this work, we report the detailed solution structure of the full-length HIV-1 CA protein in a monomeric state. To achieve this goal, we have utilized the strategy of making two critical mutations in the CTD of the wt CA that are known to disrupt the infectivity as well as the weak face-to-face dimerization in solution,¹⁸ viz., W184A/M185A, resulting in a monomeric form of the CA protein. Most importantly, these two mutations preserve many of the critical properties of the wt CA. (i) Unmyristoylated Gag protein containing these two mutations in CA is assembly active and forms spherical viruslike particles in vitro and in mammalian cells,²⁵ though the particles tend to be somewhat irregular in size and shape. (ii) Similarly, CA proteins with W184A and/or M185A mutations are assembly active, though much less efficient than the wild type. For example, the M185A-CA mutant has been shown to be assembly active but assembles ~10 times slower than the wild type.²⁶ The quadruple mutant protein A14C/E45C/W184A/M185A-CA polymerizes under reducing conditions (i.e., in the presence of β -mercaptoethanol to prevent spurious cross-linking of cysteines) and forms long tubes in vitro, though somewhat less efficiently than the wild type¹ (it is only after polymerization that the cysteines were allowed in this study to oxidize so that stable CA hexamers with correct cross-linking could be isolated for crystallization). The CTD–CTD dimer interface is preserved between these mutant CA hexamers despite the W184A and M185A mutations.¹ Following a suggestion that our double mutant might in fact show such assembly activity at high concentrations (private communication, P. Prevelige, Jr., 2011), we have undertaken in vitro assembly reactions. The data are shown in Figure 1, demonstrating again, and in agreement with the findings of Pornillos et al.¹ for the quadruple mutant, that our double mutant W184A/M185A-CA is indeed assembly active (but at higher concentrations than the wild type) and its assembly rate is much slower than that of the wild type (estimated to be ~28 times slower on the basis of the data at 60 μ M). (iii) The secondary structures of the double mutant Gag and the wt Gag²⁵ as well as of the double mutant CA and the wild-type CA²⁰ are also identical, as established by CD on very dilute solutions. (iv) Further, mutations at positions 184 and 185 that inhibit CTD–CTD dimerization in both CA and Gag do not significantly alter the affinity of CA for host cell binding partners; e.g., the binding properties of M185A-CA with prolyl isomerase cyclophilin A²⁷ and that of Gag polyprotein containing the W184A/M185A-CA domain with human lysyl-tRNA synthetase²⁸ are similar to those of the respective wild-type molecules. Taken together, the results of the Gag and capsid assembly reactions described above, CD data, and binding assays with host cell binding partners demonstrate that the monomeric mutant CA, though not infective, retains many

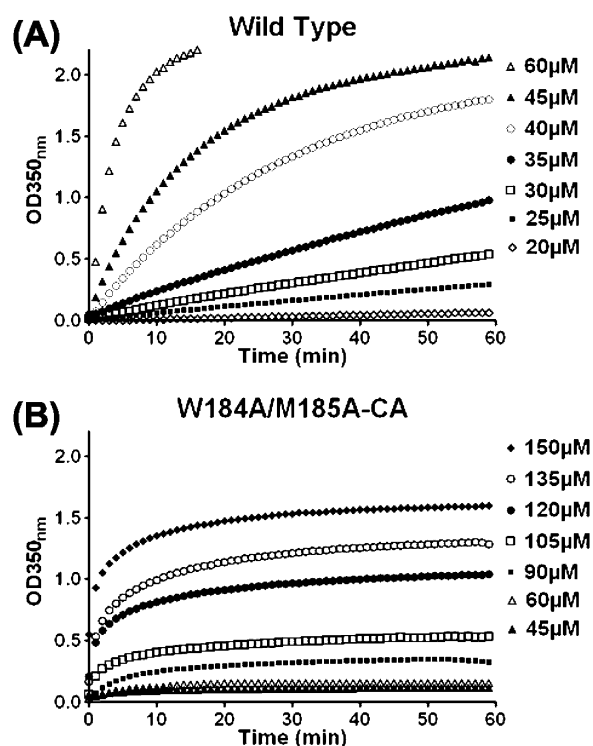


Figure 1. In vitro assembly of the HIV-1 capsid protein: (A) wild-type CA and (B) double mutant W184A/M185A-CA at a final NaCl concentration of 2.25 M. The concentrations of the proteins are indicated at the right.

of the critical properties of the wild-type CA even though it exists as a monomer in solution. Residues W184 and M185 in HIV-1 CA appear to be important for infectivity,¹⁸ dimerization in solution, improving the efficiency of assembly, and providing uniformity in the size and shapes of the assembled capsids. However, as demonstrated by others previously¹ and by us in this work (Figure 1), the capsid protein is assembly active though much less efficient than the wt protein when these two residues are replaced with alanine. We note that these two residues are not conserved in other retroviruses.¹⁹

MATERIALS AND METHODS

To construct the pET20bHIV-1CA vector for expression of HIV-1 CA with a C-terminal His tag, plasmid WISP98-85 (from P. Prevelige, Jr.) encoding 184A/185A doubly mutated HIV-1 CA was used as a template. The coding sequence of double mutant HIV-1 CA was amplified via polymerase chain reaction to introduce *Nde*I and *Sal*I cutting sites at the 5' and 3' ends, respectively, and ligated with a *Nde*I and *Xho*I cut pET20B (EMD). This new construct, pET20bHIV-1CA (Figure S5 of the Supporting Information), was transformed into *Escherichia coli* Rosetta2 (DE3) cells (EMD) for expression. Autoinduction medium P5052N25 was used to produce HIV-1 [¹⁵N]CA (Supporting Information). For ¹⁵N and ¹³C labeling of HIV-1 CA, bacteria were grown in P040 medium (Supporting Information), with 25 mM ¹⁵NH₄Cl and 0.4% [¹³C]glucose. Overexpression was induced with 0.8 mM IPTG at an OD₆₀₀ of 1.2–1.8 for 8 h at 37 °C.

For triple labeling, *E. coli* cells in 100% D₂O were transferred to P040 containing 75% D₂O for ²H, ¹⁵N, and ¹³C labeling of HIV-1 CA. Protein was induced as described above, for 12 h. Ammonium sulfate precipitation and Q-column absorption

to single exponentials using NMRView to compute the R_1 and R_2 values for each residue.

In vitro CA assembly reactions were performed using the turbidity assay as previously described by other laboratories.²⁶ Purified wild-type (WT) and W184A/M185A double mutant (DM) capsid proteins were dialyzed in 10 mM β -mercaptoethanol and 50 mM Tris-HCl (pH 8.0) (buffer D). Proteins were concentrated with an Amicon Ultra 10K apparatus (Millipore, Billerica, MA) to 1 mM for WT and 2 mM for DM. Various concentrations of proteins were diluted, and the volumes of solutions were adjusted with buffer D to 100 μ L. Within 5 min of preparation of the solution, 900 μ L of 2.5 M NaCl in buffer D was added to trigger the polymerization (final salt concentration of 2.25 M). Capsid assembly was monitored every minute for 1 h at an OD of 350 nm on a Varian Cary 50 UV–visible spectrophotometer controlled by Kinetics software in the WinUV package (Agilent, Santa Clara, CA). The delay between adding the salt and the first OD measurement was between 9 and 15 s. The in vitro assembly reaction data are shown in Figure 1.

RESULTS AND DISCUSSION

The monomeric mutant protein with the W184A and M185A critical mutations was cloned and expressed using the procedures described in Materials and Methods. The detailed sequence-specific assignments for the backbone nuclei (NH, ^{15}N , $^{13}\text{C}\alpha$, αH , $^{13}\text{C}\beta$, βH , and ^{13}CO) were completed by following one-bond and two-bond correlations in 3D NMR spectra recorded using the methods described above. Figure 2 shows the ^{15}N HSQC spectrum of the monomeric mutant with all the peaks identified. Side chain assignments were completed by a combination of HCCH-TOCSY, HCCH-COSY, ^{13}C NOESY-HSQC, and ^{15}N NOESY-HSQC measurements on appropriate samples. The distance, torsion angle, and hydrogen bond constraints were used to generate the final set of 500 structures, and the 20 lowest-energy structures were retained and minimized using CYANA 2.1, until no further reductions in the target function were evident. The structural statistics for the 20 best structures are given in Table 1. The sequential and short-range NOE connectivities identified in the monomeric mutant CA are shown in Figure S1 of the Supporting Information. Representative panels from 3D NMR data are shown in Figures S2 and S3 of the Supporting Information.

Solution Structure of the Full-Length Monomeric Mutant CA. The individual NTD and CTD in the family of 20 NMR structures are shown in Figure 3, while Figure 4 shows a representative structure of the full-length protein from this family. Panels a and b of Figure S4 of the Supporting Information show a superposition of the 20 full-length protein structures, with the NTD aligned and CTD aligned, respectively, in separate panels. The NMR structure exhibits the distinct independently folded NTD and CTD joined by a highly flexible five-residue linker. No interdomain NOE contacts could be unequivocally established in the NOESY data. The highly flexible nature of the interdomain linker can be judged from the spatial distribution of domains that are compatible with the NMR data in panels a and b of Figures S4 of the Supporting Information, as well as from the ^{15}N relaxation data in Figure 5. This figure also identifies other flexible regions such as the loops between helices and the C-terminal tail following residue 220. The NTD consists of seven distinct helices spanning residues 17–29, 36–43, 51–57, 63–83, 101–104, 111–118, and 126–144. A short β -hairpin

Table 1. NMR and Refinement Statistics for the W184A/M185A-CA Structures

NMR Distance and Dihedral Constraints				
no. of distance constraints				
total NOE				3052
intraresidue				1005
inter-residue				
sequential ($ i - j = 1$)				907
medium-range ($ i - j < 4$)				720
long-range ($ i - j > 5$)				420
hydrogen bonds				190
total no. of dihedral angle restraints				
ϕ				192
ψ				193
ω				1
Structure Statistics				
violations (mean and standard deviation)				
no. of distance constraints (>0.1 Å)				4 ± 3
no. of dihedral angle constraints ($>5.0^\circ$)				0
maximum dihedral angle violation (deg)				2.70 ± 0.10
maximum distance constraint violation (Å)				0.16 ± 0.05
average target function				1.93
average pairwise rmsd ^a (Å)				
NTD/CTD ^b				
heavy				$1.84 \pm 0.21, 1.53 \pm 0.24$
backbone				$1.57 \pm 0.24, 1.01 \pm 0.22$
NTD/CTD cores ^c				
heavy				$1.17 \pm 0.19, 1.06 \pm 0.16$
backbone				$0.71 \pm 0.20, 0.53 \pm 0.16$
Comparison of HIV1 W184A/M185A-CA with Other Similar Structures (Å)				
	NTD ^d	NTD (all helices)	CTD ^e	CTD (all helices) ^f
HIV-1 CA NTD (NMR ¹⁶)	1.58	1.75	—	—
HIV-1 CA CTD (NMR ²⁰)	—	—	1.17	1.27
HIV-1 CA (cystal ²⁴)	1.81	1.77	1.69	1.83
HIV-1 CA NTD (cystal ¹⁷)	1.85	1.82	—	—
HIV-1 CA CTD (2KOD ⁸)	—	—	2.19	2.21
HIV-1 CA CTD (2J00 ²¹)	—	—	6.53	7.01
Ramachandran Analysis				
residues in favored regions (%)				84.5
residues in additional allowed regions (%)				14.5
residues in generously allowed regions (%)				0.8
residues in disallowed regions (%)				0.1

^aFive hundred structures were calculated within CYANA 2.1. The pairwise rmsd was calculated among 20 refined structures. ^bDomains are defined as residues 1–144 for the NTD and residues 150–221 for the CTD. ^cCores are defined as the major helices for the NTD (helices 1–4 and 7) and CTD (helices 9–12). ^dSuperposition of helices 1–4 and 7. ^eSuperposition of helices 9–12. ^fSuperposition of all helices, including the 3_{10} helix (helix 8).

structure¹⁶ involving residues 2–12 at the N-terminus has also been identified (Figure 4). Proline 122 has been identified as a *cis*-proline³¹ by cisprocheck on the basis of its unique $^{13}\text{C}\beta$ and $^{13}\text{C}\gamma$ chemical shifts. The CTD in the full-length monomeric mutant CA consists of a short 3_{10} helix (helix 8) formed by residues 150–152 followed by four helices (helices 9–12) spanning residues 161–175, 185–193, 196–205, and 211–220, respectively (some of the references in the literature did not

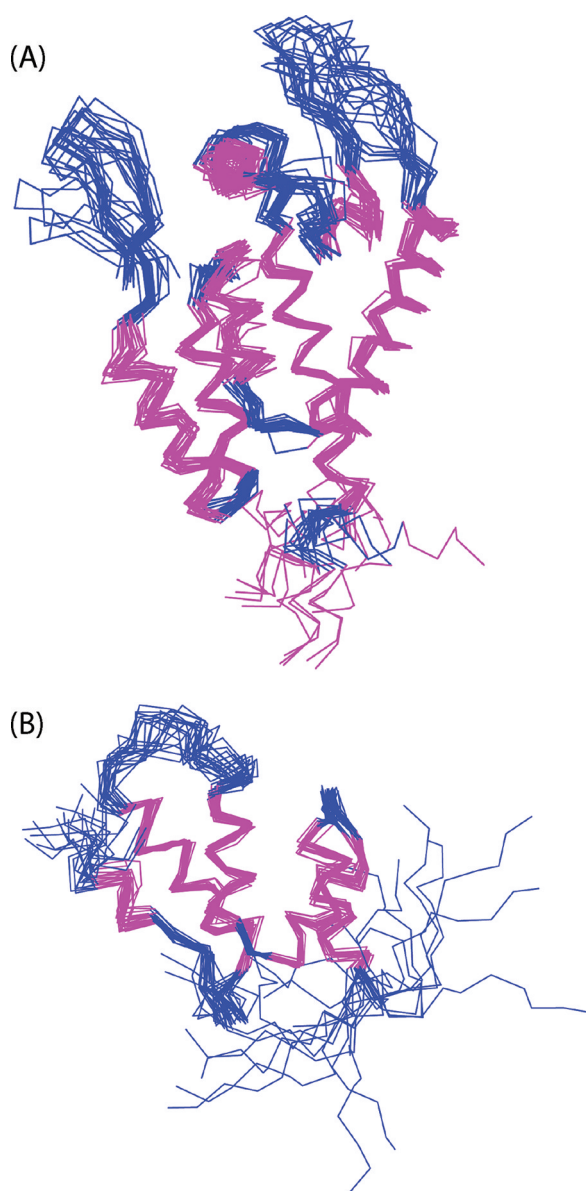


Figure 3. Superposition of the NTD and CTD of the HIV-1 CA monomeric mutant from the 20 best structures. (A) Superposition of the NTDs. The helices are colored magenta. The structures also include the flexible linker at its C-terminus, as well as residues constituting the 3_{10} helix of the CTD, to show the highly disordered nature of the flexible linker. Figure S4a of the Supporting Information shows the distribution of the CTDs when the NTDs are aligned. A comparison of the monomeric mutant CA's NTD with that of the isolated NTD NMR structure¹⁶ shows an rmsd of 1.58 Å for the alignment of the five major helices (H1–H4 and H7). (B) Superposition of the CTDs from the 20 best NMR structures. Figure S4b of the Supporting Information shows the distribution of the NTDs when the CTDs are aligned. A comparison of the monomeric mutant CA's CTD with that of the isolated CTD NMR structure²⁰ shows an rmsd of 1.17 Å for the alignment of the four helices (H9–H12). The structural statistics are listed in Table 1.

label the 3_{10} helix as helix 8, so they report only helices 8–11 in the CTD, which correspond to helices 9–12, respectively, in our numbering scheme).

Table 1 shows the rmsd values for the best fit alignment of the individual NTD and CTD with previously published crystal and NMR structures of isolated domains, as well as with the

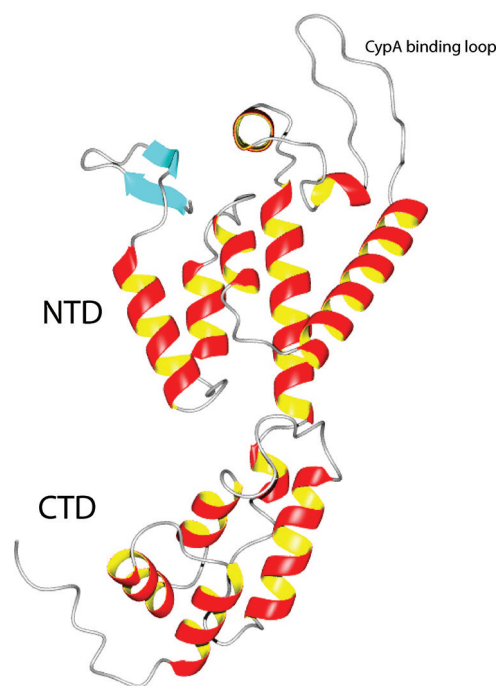


Figure 4. Representative member of the NMR structures of the full-length monomeric mutant CA as a ribbon diagram (MolMol). The secondary structure consists of a short β hairpin (aqua), with helices 1–7 in the NTD at the top and helices 8–12 (including the short 3_{10} helix as helix 8) in the CTD at the bottom. Even though our construct has the start codon residue Met (unassigned) prior to Pro1, the general arrangement of the β -hairpin is very similar to that seen in the isolated NTD structure.¹⁶ The linker between the NTD and CTD is five residues long (residues 145–149 corresponding to YSPTS) and highly flexible in solution (see Figure 5). In the crystal structure of the antiparallel dimer of wt CA, it is two residues long (SP).²⁴ The cyclophilin A (CypA) binding loop¹⁶ is located between helices 4 and 5 at the top right.

domains in the recent crystal structure of the triiodide-stabilized antiparallel dimer of full-length wild-type CA. The NTD in the full-length monomeric mutant CA shows excellent agreement with the NMR structure of the isolated NTD¹⁶ with an rmsd of 1.58 Å. Similarly, the CTD of the monomeric mutant CA shows excellent agreement with the NMR structure of the isolated W184A/M185A mutant of the CTD²⁰ with an rmsd of 1.17 Å. Figure 6 shows a comparison of the CTDs of the monomeric mutant CA with those from the wt CA in the antiparallel dimer crystal structure²⁴ and the wt CTD dimer NMR structure.⁸ Most noteworthy is the difference in helix 10 between these structures. In the crystal structure and the CTD dimer NMR structure,⁸ helix 10 is longer and shows a bend, whereas in the monomeric mutant, it is shorter and straight.

The NTD and CTD are connected by a five-residue linker formed by residues 145–149 corresponding to the sequence YSPTS. Of these, Y145 and S149 are the partially unwound terminal residues of helices 7 and 8 (the 3_{10} helix), respectively, on either side of the linker. The linker in the NMR structure is highly flexible, as evidenced by the random coil chemical shifts for the central three residues, a lack of sequential $\text{NH}(i)$ – $\text{NH}(i + 1)$ connectivities, and the ^{15}N relaxation data (Figure 5). In contrast, in the triiodide-stabilized antiparallel dimer crystal structure of the wt CA, the linker is only two residues long (residues 146 and 147 corresponding to SP). The flexibility of the linker is thought to have important consequences for the

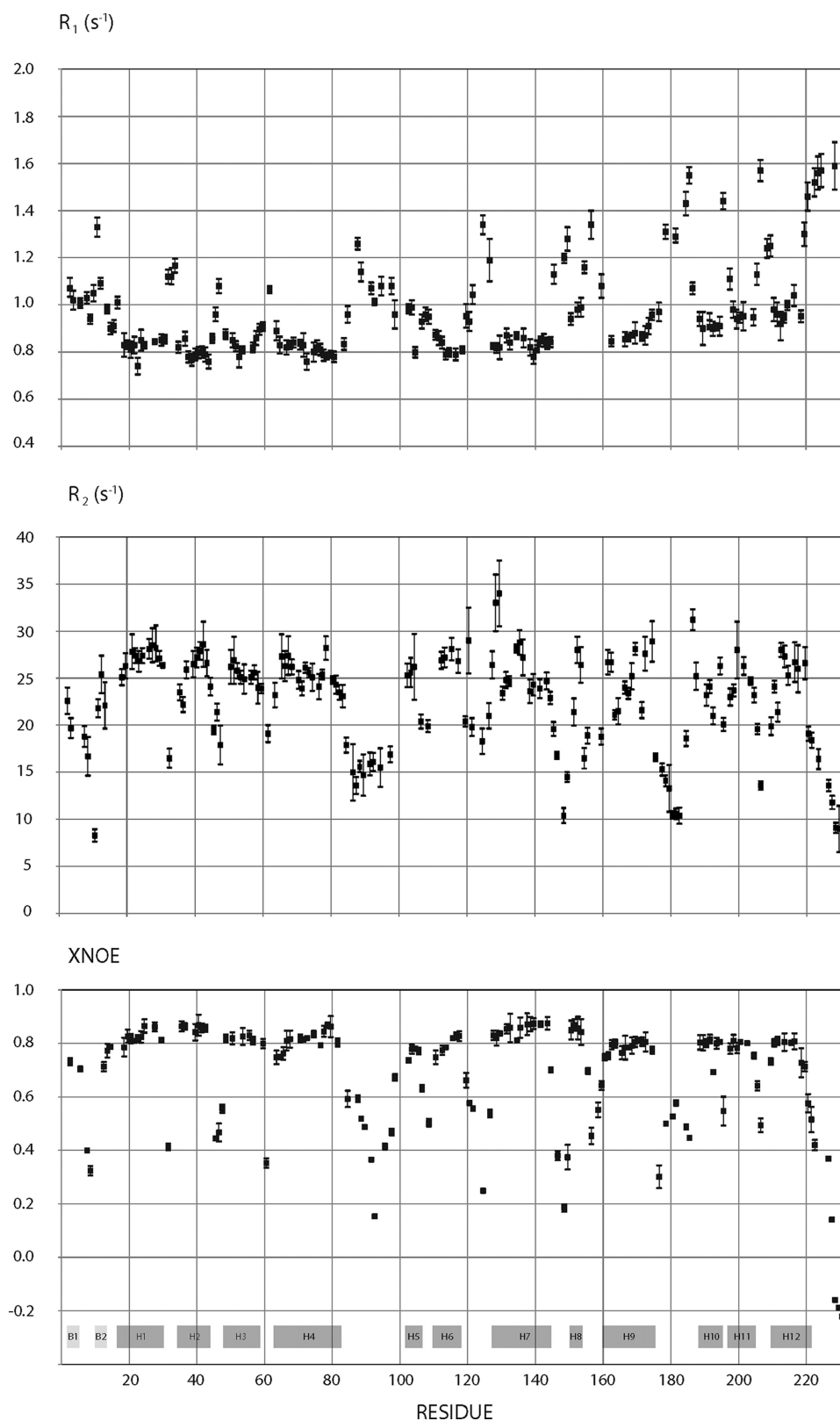


Figure 5. ^{15}N longitudinal (R_1) and transverse (R_2) relaxation rates and NOEs for W184A/M185A-CA measured at 600 MHz. A total of 158 peaks without overlap were selected from the two-dimensional NMR spectra, and their relaxation data were fitted to single exponentials. The locations of the secondary structures are identified at the bottom, with B standing for β strand and H for helices. Helix 8 refers to the short 3_{10} helix. The flexible interdomain linker is located between H7 and H8. The error bars for R_1 and R_2 refer to the deviations of data points with the single-exponential fitted curve. For the NOEs, the error bars refer to the noise in the spectrum. Residues 220 and 221 in helix 12 exhibited biexponential behavior and as a result show larger error bars when fitted to a single exponential.

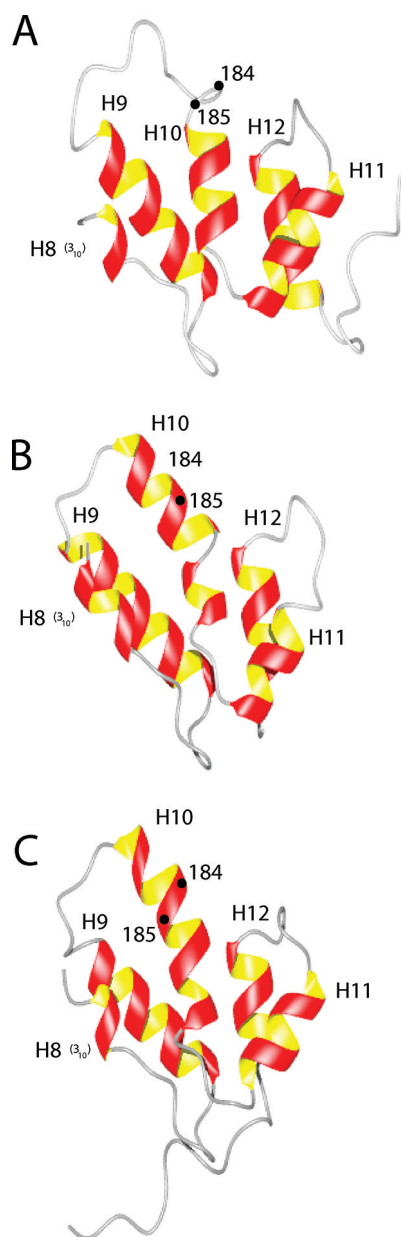


Figure 6. Comparison of the CTDs. (A) CTD of W184A/M185A-CA. (B) CTD of the wild-type antiparallel dimer CA crystal structure.²⁴ (C) CTD of the wild-type CTD dimer NMR structure.⁸ The location of residues 184 and 185 is also shown. In the crystal structure and in the wild-type CTD dimer NMR structure, helix 10 is longer and shows a bend. The hydrophobic groove formed by helices 9 and 10 has been a target for inhibitors such as CAI¹⁴ and NYAD-13.¹⁵ The CTD–CTD dimerization interface involving helix 10 has been targeted by inhibitors of dimerization.^{12,38} Further, helix 12 in the CTD has been suggested to be the binding site for human lysyl-tRNA synthetase,²⁸ which is thought to play a critical role in targeting tRNA^{Lys} for viral packaging.

capsid shape and morphology of the mature capsids (vide infra).

Insight into the Detailed Molecular Mechanism of CA CTD Dimerization during Capsid Assembly and the Critical Role of Residues W184 and M185. Our structure of the monomeric mutant CA displays a helix 10 that is straight and shorter (residues 185–193) than that seen in crystallographic and NMR structures of wt CTD dimers.^{8,18} On the

basis of this observation, we can suggest the following detailed molecular mechanism for the dimerization of the wt CA during the assembly of the capsid: we propose that just like the monomeric mutant CA, the wt CA in a monomeric state also exhibits a straight and short helix 10 with residues 179–184 forming part of the loop between helices 9 and 10. This is reasonable because the circular dichroism data are virtually identical for the wt CA and the monomeric mutant CA.²⁰ During the assembly process, two wt CA molecules that diffuse toward each other may form an initial encounter complex in which helix 10 still is short as in the monomer, but the complex involves a loose association of hydrophobic surfaces on each CTD. This initial encounter complex then transitions to a bound complex in which the hydrophobic residue-rich surfaces involving helix 10 and the 3₁₀ helix form a hydrophobic pocket, with helix 10 experiencing a conformational switch involving an elongation at its N-terminus with residues 179–185 switching to a helical conformation. This seems reasonable because residues in this region such as Q, E, V, K, W, and M undergo increases in their helical propensity versus that of A in a hydrophobic environment.³³ We further suggest that this helix 10 elongation is accompanied by a kink in the helix to optimally accommodate the packing of residues across the hydrophobic dimer interface in the bound complex. The hydrophobic interactions identified in the wt CTD dimer NMR structure⁸ include residues Y145, T148, L151, V181, W184, M185, and V191. Among these, W184 and M185 from each CTD participate in a pair of critical intermolecular W–M aromatic ring–methyl group hydrophobic interactions. With these additional intermolecular interactions in place to stabilize the dimeric interface, the wt CTD dimer has much less variability in the helix 10 crossing angle across the CTD–CTD dimeric interface, thus resulting in the more uniform formation of mature capsids. The self-association affinity of the wt CA with a K_d of 18 μM ¹⁸ also facilitates the formation of encounter complexes, thus leading to higher efficiency in assembly. Mutations such as W184A and M185A (where A has a shorter side chain than W or M) result in a loss of some of these critical intermolecular hydrophobic interactions needed to stabilize the dimer interface in the bound complex under assembly conditions, thus leading to a wider variability in the helix 10 crossing angles for each CTD–CTD dimeric interface, which in turn might contribute to the assembly of abnormal or defective capsids. Further, these mutations also significantly decrease the self-association affinity,¹⁸ thereby leading to fewer encounter complexes. This in turn can result in inefficient assembly. Thus, the loss of infectivity of HIV-1 CA due to W184A and M185A mutations³⁴ might be the combined result of both factors mentioned above, viz., a decreased level of CTD dimerization and formation of defective capsids. Molecular dynamics simulations³⁵ show that the W184A mutation is much more detrimental to the stability of the dimeric interface than the M185A mutation. We also note that W184 and M185 are unique for HIV-1 and are not conserved in other retroviruses (e.g., RSV and HTLV); as such, the fidelity of CTD–CTD dimerization during assembly of these other retroviruses is likely to be governed by other residues in the dimeric interface.

Because of the helical propensity³³ of residues 179–185 near the N-terminus [indeed, in the monomeric mutant we see a few weak sequential NH(*i*)–NH(*i* + 1) contacts among these residues (Figure S1 of the Supporting Information)], helix 10 is dynamic and exhibits a high degree of plasticity,^{1,8,20,21} with these residues easily switching to helical conformations in

response to hydrophobic associations. This plasticity can be easily appreciated in the crystal structures [Protein Data Bank (PDB) entry 3H4E] of the A14C/E45C/W184A/M185A-CA hexamers¹ in the hexamer sheet where helix 10 exhibits considerable variations among the different CAs in the hexameric rings, ranging from an elongated 3.5-turn bent helix (residues 179–189) to a 1-turn short helix (residues 189–192), presumably reflecting variations in the contacts with the neighboring hexamers in the lattice. Interestingly, the Ccmk4-templated CA with W184A and M185A mutations (PDB entry 3GV2) exhibits a longer helix with a kink (despite the W184A and M185A mutations) in all the CAs in the hexamer sheet; we speculate that this might be due to the Ccmk4 protein [to which the CA (residues 1–226) is attached through the C-terminus of CTD via a two-residue linker], which forms rigid hexamers itself,¹ thus stabilizing the CTD orientations in the interhexamer interactions, resulting in a greater hydrophobic contact surface in the CTD–CTD pairs. This suggestion is supported by the observation that the elongation and bend in helix 10 are also seen in a monomeric CTD–peptide complex¹⁵ in which the peptide is found to associate with helices 9 and 10 together.

Our proposed molecular mechanism for CTD dimerization described above, viz., that the elongation accompanied by a kink for helix 10 in the wt CTD might be the result of a hydrophobic environment from the association of this helix with an interacting partner molecule, is supported by a number of recent studies in which helix 10 generally exhibits elongation and bending (or kink) in the wt CTD dimer in solution,⁸ in wt CTD–peptide complexes,¹⁴ in crystal structures of the full-length wt CA head-to-tail antiparallel dimers in which helix 10 of the CA is interacting with a helix from the second CA,^{23,24} and in the crystal structures of hexameric assemblies where the wt CTD–CTD interface has been observed with sufficient resolution.^{1,3,8}

Implications for New Directions in the Design of Inhibitors. Previous attempts at structure-based design of CA inhibitors were based primarily on the structures of the isolated NTD and CTD because the full-length HIV-1 CA monomer structure was not available. Accordingly, these efforts focused on developing inhibitors of interdomain interactions. Our structural determination of the full-length monomer as well as the availability of chemical shifts will facilitate inhibitor design for a totally new target, viz., interdomain flexibility, because such interdomain flexibility is a critical requirement for the productive assembly of fullerene cone-shaped mature capsids (vide infra). For example, (i) it may be possible to identify or design small molecules that tightly bind directly to the flexible linker area and minimize the degree of interdomain flexibility. (ii) Alternately, by fragment-based design of high-affinity inhibitors by SAR-by-NMR,³⁶ one can design inhibitors that simultaneously bind to both the NTD and the CTD (i.e., with one functional group binding to the NTD and the second functional group binding to the CTD), thereby tethering the CTD to the NTD, and significantly diminish interdomain flexibility. This approach might indeed be feasible because the free monomeric mutant exhibits considerable flexibility across the interdomain linker (Figure 5 and Figure S4a,b of the Supporting Information). Both approaches will benefit from a structural study of the complexes of the inhibitors (and initial leads) with the full-length monomeric mutant CA. Such a reduction in interdomain flexibility is likely to have profound detrimental effects on the assembly of mature capsids and

immature virions. Thus, our structural study in this work provides a firm foundation for undertaking these new directions that target interdomain flexibility.

Additionally, using fragment-based design of high-affinity inhibitors by SAR-by-NMR, it may be feasible to simultaneously inhibit two or more separate interactions such as the intermolecular interdomain interactions (i.e., NTD–NTD, NTD–CTD, and CTD–CTD) that play a role during the assembly of the mature capsid as well as the interactions of CA with host cell proteins such as cyclophilin A and lysyl-tRNA synthetase. With NMR assignments and structures available from this study, they may facilitate a combined STD-NMR/CORCEMA-based³⁷ development of new leads for inhibitors.

Previous work from other laboratories has identified the assembly inhibitor CAP-1 that binds to the NTD.¹³ It is thought CAP-1 functions by inhibiting the intermolecular NTD–CTD interactions during the assembly of the mature capsid. The binding site for CAP-1 is at the bottom end of the hydrophobic cavity formed by helices 1, 2, 4, and 7 in the NTD (Figure 7A). In our NMR structures, though there is some distribution among their side chain orientations, the aromatic rings of Phe32, His62, and Tyr145 consistently tend to point inward toward the bottom of the hydrophobic cavity formed by these four helices (Figure 7A). A second inhibitor, PF-3450074,¹⁰ that binds to the NTD at a different site, viz., to a hydrophobic cavity formed by helices 3–5 and 7 (Figure 7B), and affects both early and late stage events in the HIV-1 replication cycle also has been identified. It thus appears to act from a mechanism different from that of CAP-1.

On the CTD, the hydrophobic pocket constituted by helix 10 and its neighboring helix 9 (Figure 6) have been the focus of assembly inhibitors. These inhibitors include CAI¹⁴ and the cell-penetrating hydrocarbon-stapled peptide NYAD-13 and its analogues.¹⁵ Interestingly, even though both these peptides bind essentially in the same site in the nonpolar groove involving helix 10 (that contributes to CTD–CTD dimerization), they do not appear to seriously alter the monomer–dimer equilibrium of the CTD in solution. Thus, their ability to inhibit the formation of both immature and mature particles is thought to be due to the formation of nonfunctional dimers upon peptide binding.^{14,15} The CTD–CTD dimerization interface involving helix 10 has also been targeted by designing peptide-based and dendrimer-based inhibitors that directly interfere with dimerization.^{12,38}

Implications for Studying Interactions with Host Cell Binding Partners. HIV-1 exploits a large number of host cell proteins in its replication cycle.^{39,40} For example, the capsid protein binds to human cyclophilin A (CypA)⁴¹ and to lysyl-tRNA synthetase.²⁸ The precise role of human CypA within the HIV-1 life cycle remains poorly understood, though it has been known for some time that CypA is required for the optimal infectivity of the virus and for viral replication.^{41,42} The binding site for cyclophilin A⁴¹ is located in the loop between helices 4 and 5 on the NTD (Figure 4). Even though the crystal structure of the NTD–CypA complex has been published,⁴¹ the structure of the complex of CypA with full-length CA is not yet available. Such a structural study with full-length CA is important for addressing the effects of CypA on the CTD.⁴³ Thus, the monomeric mutant CA study described here will serve as a basis for the study of the CA–CypA complex in solution to probe the effects of CypA binding on the CTD under different solvent conditions. Human lysyl-tRNA synthetase (LysRS) is thought to play a critical role in targeting

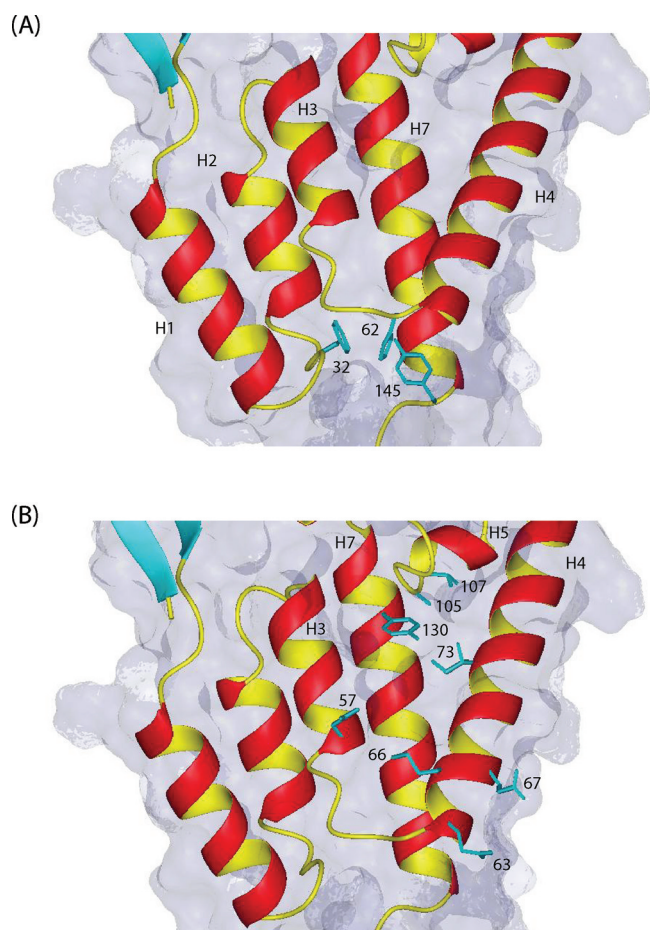


Figure 7. (A) Close-up view of the hydrophobic binding pocket on the NTD among helices 1, 2, 4, and 7 for the assembly inhibitor CAP-1.¹³ Though there is some distribution of side chain orientations among the 20 structures, the aromatic rings of Phe32, His62, and Tyr145 tend to point inward toward the bottom of the cavity. The binding of the inhibitor is accompanied by a displacement of these aromatic side chains. (B) Binding pocket for antiviral compound PF-3450074 formed by helices 3–5 and 7 on the NTD.¹⁰ Some of the binding pocket residues involved in interactions with the inhibitor are also shown (from the crystal structure of the inhibitor with a mutant NTD¹⁰). This inhibitor targets both early and late events in the replication cycle. The cyclophilin-dependent mechanism of antiviral activity of PF-3450074 is different from that of CAP-1.

tRNA^{Lys} for viral packaging.²⁸ For lysyl-tRNA synthetase, helix 12 on the CTD has been suggested²⁸ as the binding site (Figure 6). Thus, these capsid protein–human cellular protein interactions are also attractive targets for inhibitor design. The restriction factor rhTrim5 α from rhesus macaque monkeys restricts HIV-1 infection by a mechanism that was thought to involve binding to the conical capsid and cause premature uncoating.⁴⁴ Thus, the design of rhTrim5 α mimetics that bind to HIV-1 CA is an attractive strategy in drug design.⁴⁴ With the availability of the NMR structure of a full-length monomeric mutant CA, it is now possible to study the interaction of CA with these host cellular proteins directly in solution by NMR without the complications associated with monomer–dimer equilibrium (it is anticipated that crystallization of these complexes for X-ray crystallographic studies could still likely pose a problem because of the flexibility of the interdomain linker). The ability, provided by our NMR data and structure, to develop high-affinity inhibitors by SAR-by-NMR³⁶ that

simultaneously inhibit two or more of the events described above within the HIV-1 replication cycle, including various interdomain interactions during the assembly of both immature virions and mature capsids, as well as interdomain flexibility, and interactions with host cell proteins exploited by the virus, adds yet another exciting dimension to future efforts in the development of antivirals for HIV-1.

Capsid Shapes and Polymorphism. Retroviral capsids such as HIV-1 and RSV exhibit considerable polymorphism in shape, and it has been suggested that this polymorphism might be a result of the flexible linker^{1,45} that permits some degree of play in the relative orientations of two domains across the NTD–CTD and CTD–CTD interdomain interfaces during the assembly. The curvature of the fullerene cone-shaped HIV-1 capsid has been proposed¹ to result from CA hexamers consisting of a ring of NTDs surrounded by a mobile belt of CTDs (with mobility arising from the flexible linker) that interact through CTD–CTD dimerization with neighboring hexamers. Our NMR structural and relaxation data lend strong support to this notion; viz., a highly flexible linker can lead to variations in the various CA–CA interfaces that contribute to the variable curvature of the cone-shaped mature HIV-1 capsid as well as the polymorphism in mature capsid shapes.^{1,45}

■ ASSOCIATED CONTENT

Supporting Information

A description of the media used in protein expression and Figures S1–S7. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Accession Codes

The coordinates for the family of 20 lowest-energy structures (Figure 3), including the representative NMR structure (Figure 4) of the HIV-1 monomeric mutant CA, have been deposited in the Protein Data Bank (entry 2LF4). The detailed chemical shifts for the W184A/M185A-CA have been deposited in the BioMagResBank at the University of Wisconsin (Madison, WI) (entry 17738) and have been reported elsewhere.⁴⁶

■ AUTHOR INFORMATION

Corresponding Author

*MCLM-490, Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham, Birmingham, AL 35294-0005. Phone: (205) 934-5695. Fax: (205) 934-6475. E-mail: nrk@uab.edu.

Funding

Support of this work by National Institute of Allergy and Infectious Diseases Grants 1R21AI081591 and 3R21AI081591-02S1, National Cancer Institute Grant 1P30 CA13148 that supported the NMR Facility, and National Center for Research Resources Grant 1S10RR021064-01A1 for the 600 MHz CryoProbe is gratefully acknowledged.

■ ACKNOWLEDGMENTS

We thank Prof. Peter Prevelige, Jr., for the clone of the non-His-tagged double mutant CA and for discussions about the assembly properties of the wild-type and mutant CA proteins of HIV-1, and Dr. Clemens Anklin of Bruker Biospin Inc. for the SOFAST-HMQC sequence.

■ ABBREVIATIONS

HIV-1, human immunodeficiency virus-1; CA, capsid protein; NTD, N-terminal domain; CTD, C-terminal domain; rmsd, root-mean-square deviation.

■ REFERENCES

- (1) Pornillos, O., Ganster-Pornillos, B. K., Kelly, B. N., Hua, Y., Whitby, F. G., Stout, C. D., Sundquist, W. I., Hill, C. P., and Yeager, M. (2009) X-ray structures of the hexameric building block of the HIV capsid. *Cell* 137, 1282–1292.
- (2) Campos-Olivas, R., Newman, J. L., and Summers, M. F. (2000) Solution structure and dynamics of the Rous sarcoma virus capsid protein and comparison with capsid proteins of other retroviruses. *J. Mol. Biol.* 296, 633–649.
- (3) Ganster-Pornillos, B. K., Cheng, A., and Yeager, M. (2007) Structure of full-length HIV-1 CA: A model for the mature capsid lattice. *Cell* 131, 70–79.
- (4) Sundquist, W. I., and Hill, C. P. (2007) How to assemble a capsid. *Cell* 131, 17–19.
- (5) Wright, E. R., Schooler, J. B., Ding, H. J., Kieffer, C., Fillmore, C., Sundquist, W. I., and Jensen, G. J. (2007) Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells. *EMBO J.* 26, 2218–2226.
- (6) Ako-Adjei, D., Johnson, M. C., and Vogt, V. M. (2005) The retroviral capsid domain dictates virion size, morphology, and coassembly of gag into virus-like particles. *J. Virol.* 79, 13463–13472.
- (7) Briggs, J. A., Riche, J. D., Glass, B., Bartonov, V., Zanetti, G., and Kräusslich, H. G. (2009) Structure and assembly of immature HIV. *Proc. Natl. Acad. Sci. U.S.A.* 106, 11090–11095.
- (8) Byeon, I. J., Meng, X., Jung, J., Zhao, G., Yang, R., Ahn, J., Shi, J., Concel, J., Aiken, C., Zhang, P., and Gronenborn, A. M. (2009) Structural convergence between Cryo-EM and NMR reveals intersubunit interactions critical for HIV-1 capsid function. *Cell* 139, 780–790.
- (9) Lanman, J., Lam, T. T., Barnes, S., Sakalian, M., Emmett, M. R., Marshall, A. G., and Prevelige, P. E. Jr. (2003) Identification of novel interactions in HIV-1 capsid protein assembly by, high-resolution mass spectrometry. *J. Mol. Biol.* 325, 759–772.
- (10) Blair, W. S., Pickford, C., Irving, S. L., Brown, D. G., Anderson, M., Bazin, R., Cao, J., Ciaramella, G., Isaacson, J., Jackson, L., Hunt, R., Kjerrstrom, A., Nieman, J. A., Patick, A. K., Perros, M., Scott, A. D., Whitby, K., Wu, H., and Butler, S. L. (2010) HIV capsid is a tractable target for small molecule therapeutic intervention. *PLoS Pathog.* 6, 1–10.
- (11) Adamson, C. S., and Freed, E. O. (2010) Novel approaches to inhibiting HIV-1 replication. *Antiviral Res.* 85, 119–141.
- (12) Neira, J. L. (2009) The capsid protein of human immunodeficiency virus: Designing inhibitors of capsid assembly. *FEBS J.* 276, 6110–6117.
- (13) Kelly, B. N., Kyere, S., Kinde, I., Tang, C., Howard, B. R., Robinson, H., Sundquist, W. I., Summers, M. F., and Hill, C. P. (2007) Structure of the antiviral assembly inhibitor CAP-1 complex with the HIV-1 CA protein. *J. Mol. Biol.* 373, 355–366.
- (14) Ternois, F., Sticht, J., Duquerooy, S., Kräusslich, H. G., and Rey, F. A. (2005) The HIV-1 capsid protein C-terminal domain in complex with a virus assembly inhibitor. *Nat. Struct. Mol. Biol.* 12, 678–682.
- (15) Bhattacharya, S., Zhang, H., Debnath, A. K., and Cowburn, D. (2008) Solution structure of a hydrocarbon stapled peptide inhibitor in complex with monomeric C-terminal domain of HIV-1 capsid. *J. Biol. Chem.* 283, 16274–16278.
- (16) Gitti, R. K., Lee, B. M., Walker, J., Summers, M. F., Yoo, S., and Sundquist, W. I. (1996) Structure of the amino-terminal core domain of the HIV-1 capsid protein. *Science* 273, 231–235.
- (17) Kelly, B. N., Howard, B. R., Wang, H., Robinson, H., Sundquist, W. I., and Hill, C. P. (2006) Implications for viral capsid assembly from crystal structures of HIV-1 Gag(1–278) and CA(N)(133–278). *Biochemistry* 45, 11257–11266.
- (18) Gamble, T. R., Yoo, S., Vajdos, F. F., von Schwedler, U. K., Worthylake, D. K., Wang, H., McCutcheon, J. P., Sundquist, W. I., and Hill, C. P. (1997) Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. *Science* 278, 849–853.
- (19) Ivanov, D., Tsodikov, O. V., Kasanov, J., Ellenberger, T., Wagner, G., and Collins, T. (2007) Domain-swapped dimerization of the HIV-1 capsid C-terminal Domain. *Proc. Natl. Acad. Sci. U.S.A.* 104, 4353–4358.
- (20) Wong, H. C., Shin, R., and Krishna, N. R. (2008) Solution structure of a double mutant of the carboxy-terminal dimerization domain of the HIV-1 capsid protein. *Biochemistry* 47, 2289–2297.
- (21) Alcaraz, L. A., del Alamo, M., Barrera, F. N., Mateu, M. G., and Neira, J. L. (2007) Flexibility in HIV-1 assembly subunits: Solution structure of the monomeric C-terminal domain of the capsid protein. *Biophys. J.* 93, 1264–1276.
- (22) Momany, C., Kovari, L. C., Prongay, A. J., Keller, W., Gitti, R. K., Lee, B. M., Gorbalenya, A. E., Tong, L., McClure, J., Ehrlich, L. S., Summers, M. F., Carter, C., and Rossmann, M. G. (1996) Crystal structure of dimeric HIV-1 capsid protein. *Nat. Struct. Biol.* 3, 763–770.
- (23) Berthet-Colominas, C., Monaco, S., Novelli, A., Sibai, G., Mallet, F., and Cusack, S. (1999) Head-to-tail dimers and interdomain flexibility revealed by the crystal structure of HIV-1 capsid protein (p24) complexed with a monoclonal antibody Fab. *EMBO J.* 18, 1124–1136.
- (24) Du, S., Betts, L., Yang, R., Shi, H., Concel, J., Ahn, J., Aiken, C., Zhang, P., and Yeh, J. I. (2011) Structure of the HIV-1 full-length capsid protein in a conformationally trapped unassembled state induced by small-molecule binding. *J. Mol. Biol.* 406, 371–386.
- (25) Datta, S. A., Curtis, J. E., Ratcliff, W., Clark, P. K., Crist, R. M., Lebowitz, J., Krueger, S., and Rein, A. (2007) Conformation of HIV-1 Gag Protein in Solution. *J. Mol. Biol.* 365, 812–824.
- (26) Lanman, J., Sexton, J., Sakalian, M., and Prevelige, P. E. Jr. (2002) Kinetic analysis of the role of intersubunit interactions in human immunodeficiency virus type 1 capsid protein assembly in vitro. *J. Virol.* 76, 6900–6908.
- (27) Bosco, D. A., and Kern, D. (2004) Catalysis and binding of cyclophilin A with different HIV-1 capsid constructs. *Biochemistry* 43, 6110–6119.
- (28) Kovaleski, B. J., Kennedy, R., Hong, M. K., Datta, S. A., Kleiman, L., Rein, A., and Musier-Forsyth, K. (2006) In vitro characterization of the interaction between HIV-1 Gag and human lysyl-tRNA synthetase. *J. Biol. Chem.* 281, 19449–19456.
- (29) Delaglio, F., Grzesiek, S., Vuister, G. W., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* 6, 277–293.
- (30) Bartels, C., Xia, T., Billeter, M., Guentert, P., and Wuethrich, K. (1995) The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J. Biomol. NMR* 6, 1–10.
- (31) Güntert, P. (2004) Automated NMR structure calculation with CYANA. *Methods Mol. Biol.* 278, 353–378.
- (32) Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* 44, 213–223.
- (33) Krittana, C., and Johnson, W. C. (2000) Proteins, The relative order of helical propensity of amino acids changes with solvent environment. *Proteins: Struct., Funct., Genet.* 39, 132–141.
- (34) Von Schwedler, U. K., Stray, K. M., Garrus, J. E., and Sundquist, W. I. (2003) Functional surfaces of the human immunodeficiency virus type 1 capsid protein. *J. Virol.* 77, 5439–5450.
- (35) Yu, X., Wang, Q., Yang, J. C., Buch, I., Tsai, C. J., Ma, B., Cheng, S. Z., Nussinov, R., and Zheng, J. (2009) Mutational analysis and allosteric effects in the HIV-1 capsid protein carboxyl-terminal dimerization domain. *Biomacromolecules* 10, 390–399.
- (36) Shuker, S. B., Hajduk, P. J., Meadows, R. P., and Fesik, S. W. (1996) Discovering high-affinity ligands for proteins: SAR by NMR. *Science* 274, 1531–1534.

- (37) Krishna, N. R., and Jayalakshmi, V. (2006) Complete relaxation and conformational exchange matrix analysis of STD-NMR spectra of ligand-receptor complexes. *Prog. NMR Spectrosc.* 49, 1–25.
- (38) Doménech, R., Abian, O., Bocanegra, R., Correa, J., Sousa-Herves, A., Riguera, R., Mateu, M. G., Fernandez-Megia, E., Velázquez-Campoy, A., and Neira, J. L. (2010) Dendrimers as potential inhibitors of the dimerization of the capsid protein of HIV-1. *Biomacromolecules* 11, 2069–2078.
- (39) Freed, E. O. (2004) HIV-1 and the host cell: An intimate association. *Trends Microbiol.* 12, 170–177.
- (40) Ott, D. E. (2002) Potential roles of cellular proteins in HIV-1. *Rev. Med. Virol.* 12, 359–74.
- (41) Gamble, T. R., Vajdos, F. F., Yoo, S., Worthylake, D. K., Houseweart, M., Sundquist, W. I., and Hill, C. P. (1996) Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. *Cell* 87, 1285–1294.
- (42) Thali, M., Bukovsky, A., Kondo, E., Rosenwirth, B., Walsh, C. T., Sodroski, J., and Göttinger, H. G. (1994) Functional association of cyclophilin A with HIV-1 virions. *Nature* 372, 363–365.
- (43) Bon Homme, M., Carter, C., and Scarlata, S. (2005) The cysteine residues of HIV-1 capsid regulate oligomerization and cyclophilin A-induced changes. *Biophys. J.* 88, 2078–2088.
- (44) Mascarenhas, A. P., and Musier-Forsyth, K. (2009) The capsid protein of human immunodeficiency virus: Interactions of HIV-1 capsid with host protein factors. *FEBS J.* 276, 6118–6127.
- (45) Cardone, G., Purdy, J. G., Cheng, N., Craven, R. C., and Steven, A. C. (2009) Visualization of a missing link in retrovirus capsid assembly. *Nature* 457, 694–698.
- (46) Shin, R., Tzou, Y.-M., Wong, H. C., and Krishna, N. R. (2011) ¹H, ¹⁵N and ¹³C resonance assignments for a monomeric mutant of the HIV-1 capsid protein. *Biomol. NMR Assignments*, in press.
- (47) Han, Y., Ahn, J., Concel, J., Byeon, I. J., Gronenborn, A. M., Yang, J., and Polenova, T. (2010) Solid-state NMR studies of HIV-1 capsid protein assemblies. *J. Am. Chem. Soc.* 132, 1976–1987.
- (48) Chen, B., and Tycko, R. (2010) Structural and dynamical characterization of tubular HIV-1 capsid protein assemblies by solid state nuclear magnetic resonance and electron microscopy. *Protein Sci.* 19, 716–730.

■ NOTE ADDED AFTER ASAP PUBLICATION

Errors in the caption to Figure 3 have been corrected. This paper was published on October 13, 2011; the corrected version reposted on October 14, 2011.