doi:10.1016/j.jmb.2007.07.070 J. Mol. Biol. (2007) 373, 355–366

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Structure of the Antiviral Assembly Inhibitor CAP-1 Complex with the HIV-1 CA Protein

Brian N. Kelly[1]†, Sampson Kyere[2]†, Isaac Kinde[2], Chun Tang[2] Bruce R. Howard[3], Howard Robinson[4], Wesley I. Sundquist[1]* Michael F. Summers[2]* and Christopher P. Hill[1]*

[1]*Department of Biochemistry University of Utah, Salt Lake City, UT 84112-5650, USA*

[2]*Department of Chemistry and Biochemistry, University of Maryland, Baltimore County Baltimore, MD 21250, USA*

[3]*Department of Physical Science, Southern Utah University, Cedar City UT 84720-2470, USA*

[4]*Biology Department Brookhaven National Laboratory, Upton NY 11973-5000, USA*

Received 28 June 2007; accepted 27 July 2007
Available online 15 August 2007

Edited by P. Wright

The CA domain of the human immunodeficiency virus type 1 (HIV-1) Gag polyprotein plays critical roles in both the early and late phases of viral replication and is therefore an attractive antiviral target. Compounds with antiviral activity were recently identified that bind to the N-terminal domain of CA (CA^N) and inhibit capsid assembly during viral maturation. We have determined the structure of the complex between CA^N and the antiviral assembly inhibitor *N*-(3-chloro-4-methylphenyl)-*N′*-{2-[({5-[(dimethylamino)-methyl]-2-furyl}-methyl)-sulfanyl]ethyl}-urea) (CAP-1) using a combination of NMR spectroscopy and X-ray crystallography. The protein undergoes a remarkable conformational change upon CAP-1 binding, in which Phe32 is displaced from its buried position in the protein core to open a deep hydrophobic cavity that serves as the ligand binding site. The aromatic ring of CAP-1 inserts into the cavity, with the urea NH groups forming hydrogen bonds with the backbone oxygen of Val59 and the dimethylamonium group interacting with the side-chains of Glu28 and Glu29. Elements that could be exploited to improve binding affinity are apparent in the structure. The displacement of Phe32 by CAP-1 appears to be facilitated by a strained main-chain conformation, which suggests a potential role for a Phe32 conformational switch during normal capsid assembly.

Keywords: HIV capsid; inhibitor; structure; assembly; CAP-1

Introduction

The AIDS epidemic continues to be a significant international health problem, with approximately 40 million people living with human immunodeficiency virus (HIV) infection world-wide.[1] In 2006 alone, 4.3 million individuals became infected with HIV, and approximately 3 million deaths were attributed to AIDS. Therapeutic agents currently used to treat HIV infection target the viral reverse transcriptase, protease, and envelope proteins, and drugs that target the integrase enzyme are undergoing clinical trials‡. Although sustained reductions in viral load can be achieved for many years with combination drug therapies,[2–4] inadequate suppression due to poor compliance, resistance, and interactions with other drugs or diet can be a significant problem for some patients and can lead to the spread of drug-resistant strains.[2,3,5–8] Inhibition of other viral components may provide the best approach for attacking viral resistance.[2]

**Corresponding authors.* E-mail addresses: wes@biochem.utah.edu; summers@hhmi.umbc.edu; chris@biochem.utah.edu.

† B.N.K. and S.K. contributed equally to this work.

Abbreviations used: CA^N and CA^C, N and C-terminal domains of the CA protein, respectively; CAP-1, *N*-(3-chloro-4-methylphenyl)-*N′*-{2-[({5-[(dimethylamino)-methyl]-2-furyl}-methyl)-sulfanyl]ethyl}-urea); Gag, Gag polyprotein; HS/MQC, heteronuclear single/multiple quantum coherence; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; HIV-1, human immunodeficiency virus type 1; EM, electron microscopy; NNRTI, non-nucleoside reverse transcriptase inhibitor; DMSO, dimethyl sulfoxide; MLV, murine leukemia virus.

‡ http://www.aidsinfo.nih.gov

The CA protein plays critical roles in the early and late phases of replication and has long been considered an attractive potential therapeutic target.[9] CA is originally synthesized as a 231 amino acid domain within the 55 kDa Gag precursor polyprotein. During the late phase of viral replication, the CA domain helps mediate the assembly of ~4000 copies of Gag into the immature virus particle.[10] Subsequent processing by the viral protease triggers conformational changes in CA that promote its assembly into the capsid, a conical protein shell composed of about 1500 CA molecules that encloses two copies of the viral genome and the viral enzymes essential for infectivity. Proper assembly of this core particle is critical for viral replication, and mutations that reduce or increase core stability lead to dramatic reductions in viral infectivity.[11] CA comprises two domains[12] that have distinct roles in stabilizing the viral capsid architecture.[13–15] In cylindrical *in vitro* assemblies that mimic viral capsids, the N-terminal domain (CA^N, residues 1–146) forms hexamers and the C-terminal domain (CA^C, residues 147–231), which is dimeric in solution,[16] links adjacent hexamers.[14] A crystal structure of the murine leukemia virus (MLV) CA^N protein has enabled atomic-level modeling of the HIV-1 CA^N hexamer.[17] Crystal structures are also available for the CA^C dimer,[16,18] although a domain-swapped dimer model has been proposed from analogy with the structurally related SCAN domain.[19,20] Finally, biochemical studies indicate that the N and C-terminal domains form intermolecular contacts in the mature capsid lattice.[21–24]

Disruption of capsid reorganization can be an effective approach to viral inhibition, with examples including targeting of capsid assembly in hepatitis B virus (HBV)[25] and blocking of capsid disassembly in picorona viruses.[26] Support for this approach for HIV-1 was provided by CAI, a peptide inhibitor of CA^C–CA^C interactions that inhibits immature and mature particle formation *in vitro*,[27,28] although CAI was unable to inhibit the release of HIV-1 particles when added to virus producing cells in cell culture or by peptide transfection. Similarly, betulinic acid appears to impede HIV maturation by binding to an assembled form of Gag and blocking processing of the C terminus of CA by the viral protease.[29] Small molecule inhibitors were also identified in a search for agents that bind directly to HIV-1 CA.[30] One of these compounds, *N*-(3-chloro-4-methylphenyl)-*N*′-{2-[({5-[(dimethylamino)-methyl]-2-furyl}-methyl)-sulfanyl]ethyl}urea (CAP-1) inhibits capsid assembly *in vitro*, HIV-1 infectivity *in vivo*, and leads to the production of poorly infectious virons with abnormal core morphologies.[30] NMR chemical shift mapping experiments revealed that CAP-1 binds to a site near the C-terminal end of CA^N,[30] and to date, more than two dozen additional compounds have been identified that bind specifically to this site (M. F. S., unpublished results).

In an effort to understand the molecular basis for inhibition by CAP-1, we determined the three-dimensional structure of the complex with CA^N using a combination of X-ray crystallography and NMR spectroscopy. The two approaches provide complementary information that was useful for addressing problems associated with relatively weak binding and poor ligand solubility. CAP-1 binds to a deep hydrophobic cavity that is formed upon extrusion of the Phe32 side-chain from a buried to an exposed position. Intermolecular interactions that appear important for binding, and additional elements that may be exploited to enhance binding affinity, are identified in the structure. The displacement of Phe32 appears to be facilitated by a strained main-chain conformation, which has implications for both the mechanism of capsid assembly and its inhibition by CAP-1.

Results and Discussion

Structure of CA^N crystallized in the presence of CAP-1

The HIV-1 CA^N protein (residues 1–146, wt $HIV\text{-}1_{NL4\text{-}3}$ sequence) was expressed recombinantly in *Escherichia coli*, purified, and crystallized in the presence of CAP-1 at concentrations above the binding constant. Although the structure was determined at 1.5 Å resolution and conformational changes that appear to open the CAP-1 binding site were observed (discussed below), density for CAP-1 was not visible. CAP-1 binds with modest affinity (~800 μM) and is poorly soluble, and our preferred explanation is that the CAP-1 that crystallized with the protein diffused out of the binding site and precipitated during crystal growth, while the ligand-bound CA^N conformation was retained because lattice contacts with a neighboring molecule occur in this region of the structure (Supplementary Data, Figure S1). It is possible that CAP-1 bound in multiple related conformations that reduce the interpretable electron density, although we do not see the residual density that might be expected at medium/low resolution for this situation. A third possibility is that the structure observed arose simply because of favorable lattice forces in the new crystal form, although this seems unlikely because we were unable to grow crystals from these conditions in the absence of CAP-1. Regardless of the reason for the missing CAP-1 density, the functional relevance of the crystal structure as the conformation that binds CAP-1 is indicated by independent NMR data that are discussed below.

Comparison of the crystal structure with other X-ray and NMR determinations of CA^N indicated that the majority of the protein residues were unaffected by the presence of CAP-1.[12,31,32] However, a significant conformational change was observed for residues that earlier NMR heteronuclear single quantum coherence (HSQC) titration experiments had indicated were in the vicinity of the

CAP-1 binding site.[30] In the structure crystallized in the presence of CAP-1, the side chain of Phe32 is repositioned ~6 Å from its buried location in the protein core to a solvent exposed environment Figure 1. The Phe32 side-chain, in turn, displaces the side-chain of Tyr145 from a partly buried position within a hydrophobic hollow in the free protein to a poorly ordered conformation that is not clearly defined in electron density maps of the protein crystallized in the presence of CAP-1. In addition, electron density was not observed for the side-chain of His62, which was well defined in structures determined for the free CA^N protein.

NMR studies of CAP-1:CA^N

CAP-1 contains a urea moiety that can adopt two conformations (1 and 2; Figure 2). Positive nuclear Overhauser effects (NOEs) were observed from the CAP-1 N1-H to both the H-5 and H-6 protons of the free CAP-1 ligand (2 mM, 5% DMSO-d_6/95% H_2O), indicating that the 2-conformation is favored ([1]/[2] ~ 0.5/1) (Figure 2). NMR data were also obtained for CAP-1 in the presence of CA^N (0.100 mM; [CAP-1]/[CA^N] = 20:1). Negative transfer-NOEs were observed under these conditions, and the relative N1-H to H5 and H6 intensities shifted in favor of the 1 conformer ([1]/[2] ~ 2:1) (Figure 2). The residual N1-H to H5 signal results from rapid exchange between free (20-fold excess) and bound CAP-1 coupled with urea bond isomerization. These data indicate that the 1 conformer is preferentially, if not exclusively, bound by CA^N.

2D NOESY spectra were obtained for CA^N as a function of added CAP-1 (Figure 2(d)). Significant chemical shift changes were observed for a number of CA^N residues, including Val27, Ala31, His62, Leu138, Ile141, and Tyr145. In some cases, intramolecular NOE intensities changed significantly upon CAP-1 binding. For example, NOEs between His 62-$H^{\delta 2}$ and both Val 27-γ_1CH_3 and Val 59-γ_2CH_3 decreased upon titration with CAP-1 (Figure 2), whereas His62-$H^{\varepsilon 1}$ exhibited a significant increase in NOE intensity with Ala64-CH_3 in the presence of CAP-1. These changes indicate that, upon CAP-1 binding, the His62 side-chain no longer packs against Val27 and Val59, but instead interacts with the N-terminal end of helix H4. In addition, Phe32 H^{ζ} undergoes a large upfield shift (from 6.81 ppm to 6.57 ppm) and exhibits increased NOE intensities to His62-$H^{\delta 2}$ and -$H^{\varepsilon 1}$ upon CAP-1 binding, indicating that the Phe32 side-chain packs against the side-chain of His62 in the CAP-1:CA^N complex.

CAP-1 precipitates at concentrations above 3 mM, and the maximum CA^N:CAP-1/CA^N(free) ratio achieved in the 2D NOESY experiments was therefore ~75%/25%. Under these conditions, Phe32 exhibited diminished but detectable NOEs with Trp23, Val36 and Leu138 due to rapid exchange between the CA^N (Phe32 sequestered) and CA^N:CAP-1 (Phe32 exposed) species. Well-resolved intermolecular NOE cross-peaks were also observed upon addition of CAP-1 to CA^N, including NOEs between CAP-1 H6 and the methyl protons of Val27, Val59, Ala65 and Met66; CAP-1 H3 and H5 NOEs to Val27 and Ala31; and CAP-1 C1-methyl NOEs to Trp23, Val59, Ala65, Leu138 and Ile141 (Figure 2). These data are consistent with a unique binding mode, in which the aromatic ring of CAP-1 is

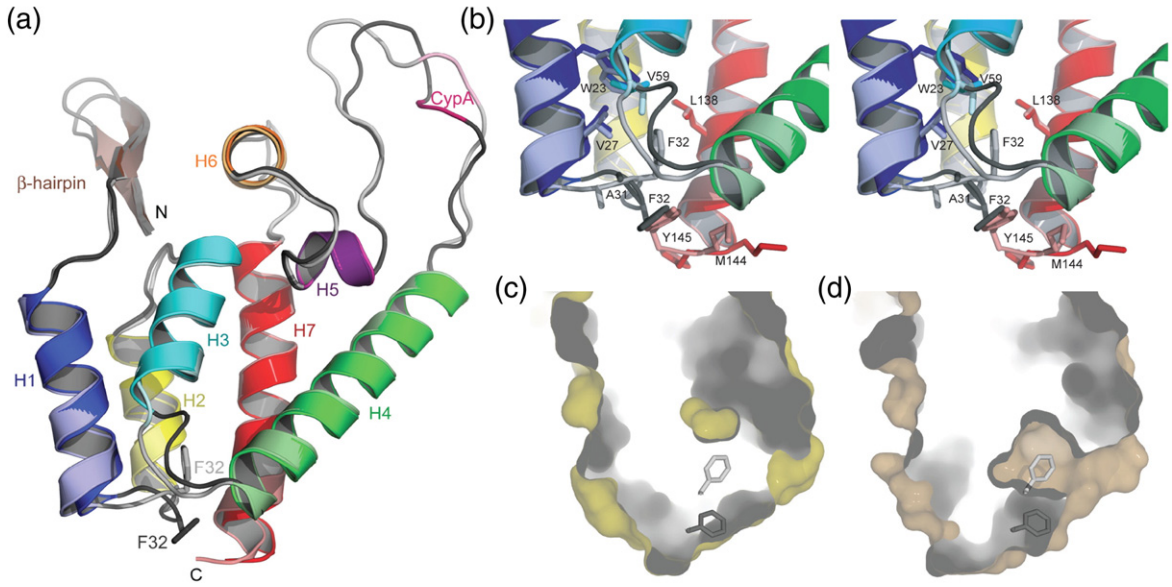


Figure 1. Structural changes induced in CA^N when crystallized in the presence of CAP-1. (a) Ribbon diagram of CA^N crystallized in presence (darker colors) or absence of CAP-1. Phe32 is shown explicitly. N and C termini, secondary structural elements, and the cyclophilin A binding site are labeled. (b) Close up stereo view of the structural changes in the presence of CAP-1. (c) Surface representation of CA crystallized in the absence (left) and presence (right) of CAP-1. Phe32 is shown explicitly in the open and closed conformations.

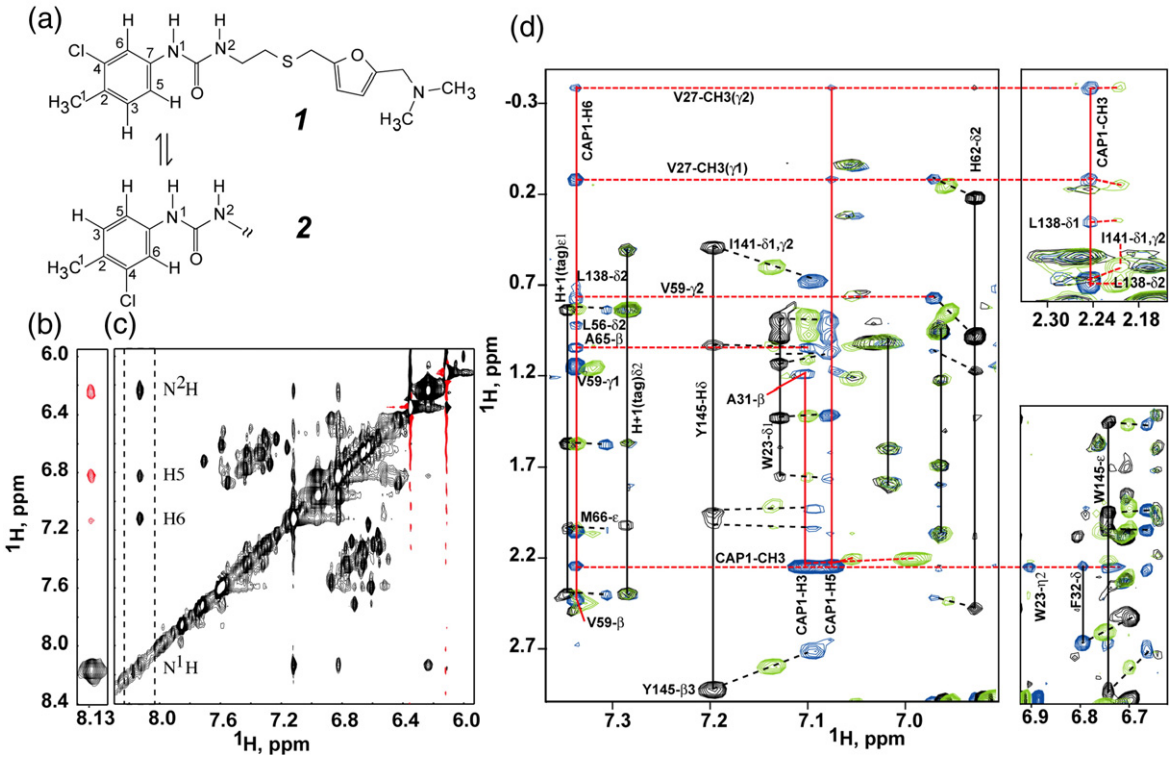


Figure 2. Portions of the 2D NOESY spectra (11 °C, 95% H_2O/5% DMSO-d_6) used to determine the intramolecular orientation of the amide protons in CAP-1. (a) Structures of two possible CAP-1 conformations. (b) Portion of the 2D NOESY spectrum of free CAP-1 showing the positive (red) intramolecular ^{1}H-^{1}H NOEs of HN1 to HN2, H5 and H6. HN1 in the free state is preferentially orientated closer to H5. (c) Row and corresponding 2D NOESY spectrum of CAP-1 (2 mM) in the presence of CA^N (0.1 mM) showing the negative (black) ^{1}H-^{1}H transfer NOEs. (d) Portions of the 2D NOESY data obtained for CA^N (0.7 mM) in the presence of increasing amounts of CAP-1 (CA^N:CAP-1 = 1:0 (black), 1:1 (green), 1:4 (blue)). Chemical shift changes observed upon titration are denoted by broken lines, and intermolecular NOEs are labeled.

sequestered within the hydrophobic pocket vacated by Phe32.

Structure of the CAP-1:CA^N complex

As indicated above, the NMR and X-ray crystallographic data provided complementary information and were therefore used jointly to determine the structure of the CA^N:CAP-1 complex. A starting model for refinement trajectories was built by manually docking CAP-1 into the Phe32 cavity of the CA^N crystal structure obtained in the presence of CAP-1. It was not possible to generate reasonable models consistent with the NOE data using CA^N crystal structures obtained in the absence of CAP-1, in which the Phe32 side-chain was buried. Atoms with well-defined electron density were restrained to the coordinate positions of the crystal structure, and atoms that lacked well-defined density were allowed to move during the calculations. After heating and equilibration at 350 K, a total of 20 structures obtained at 0.2 ps intervals were independently cooled for 2 ps to ~0 K and subjected to energy minimization, which afforded 20 final CAP-1:CA^N structures (Table 1 and Figure 3).

The position and orientation of the CAP-1 phenyl group is well defined by the NMR data within the binding pocket. The C1 methyl group packs against the side-chains of Leu138 and Ile141, the aromatic H6 proton and chlorine atom pack against the side-chain of Val59, and the H3 and H5 protons on the opposite side of the phenyl ring pack against Ala31 and the aromatic ring of Phe32. This specific packing arrangement is consistent with the observation of Val59 and Met66 side chain NOEs with CAP-1 H6 (but not H3 or H5), and Ala31 and Phe32 side-chain NOEs with CAP-1 H3 and H5 (but not H6) (Figure 4). In all the energy-minimized structures, the urea N^1H and N^2H protons are within hydrogen bonding distance of the backbone carbonyl of Val59 (Figure 4). The remaining atoms of CAP-1 are exposed to solvent and appear generally disordered, except that the dimethylamonium group resides near the carboxylate of Glu28 (Figure 4).

Evidence that CAP-1 binding is promoted by Ala31–Phe32 main-chain strain

Because the burial of phenylalanine side-chains is highly favorable, it was surprising to find that the Phe32 side-chain is displaced from the core of the protein upon CAP-1 binding. We therefore inspected Phe32 in unbiased (simulated annealing omit) density for all crystal structures of HIV-1 CA^N crystallized in the absence of CAP-1 that have been published at high resolution.[31,32] Of these 22

Table 1. NMR/X-ray AMBER refinement data

Restraints					
Intramolecular NOEs			62 His $H^{\varepsilon 1}$ 64 - Ala H^{β} 3.8		
Intermolecular NOEs					
138 Ile $H^{\delta 2}$	Cap-1	$C1H_3$ 4.3	23 Trp $H^{\eta 2}$	Cap-1	$C1H_3$ 5.5
141 Ile $H^{\gamma 2}$	Cap-1	$C1H_3$ 4.3	27 Val $H^{\gamma 1}$	Cap-1	H6 5.5
32 Phe H^{δ}	Cap-1	H5 6.1	27 Val $H^{\gamma 1}$	Cap-1	H5 5.5
32 Phe H^{δ}	Cap-1	H3 6.1	31 Val H^{β}	Cap-1	H5 3.8
32 Phe H^{ε}	Cap-1	H5 6.1	31 Val H^{β}	Cap-1	H3 5.0
32 Phe H^{ε}	Cap-1	H3 6.1	59 Val H^{β}	Cap-1	H6 3.8
32 Phe H^{ζ}	Cap-1	H5 5.0	59 Val $H^{\gamma 2}$	Cap-1	H6 5.5
32 Phe H^{ζ}	Cap-1	H3 5.0	65 Ala H^{β}	Cap-1	H6 5.5
66 Met H^{ε}	Cap-1	H6 5.5	65 Ala H^{β}	Cap-1	H5 5.5
			65 Ala H^{β}	Cap-1	H3 5.5
Torsion angle restraints			C6-C7-N1-C8 180		
			C7-N1-C8-N2 180		
			N1-C8-N2-C9 180		
Atoms fixed to X-ray coordinates			All atoms: residues 1-25, 27-59, 63-144		
			Backbone atoms: residues 26,62		
Refined structures (20 total)[a]					
Total energy			−4928.5±1.5		
AMBER energy			−5019.0±1.5		
Restraint energy			90.5±1.4		
Distance penalty			0.000±0.000		
Torsion penalty			0.02±0.01		

[a] Energies (kcal/mol) are reported as the mean±standard deviation for the 20 refined structures.

crystallographically independent views, Phe32 has well defined density in the buried position in 14 cases, but has weak density in the other eight examples. Phe32 might be adopting a range of exposed conformations in the structures with weak density, although in no case is density for an ordered Phe32 conformation visible outside of the usual buried position.

To understand better the conformational changes that occur upon CAP-1 binding, we determined the crystal structure of a CA^N mutant (A92E) to 1.9 Å resolution in the absence of CAP-1. Ala92 resides in a flexible loop that is well removed from the CAP-1 binding site, and mutation of this residue alters HIV-1's dependence on cyclophilin A[33,34] but does not lead to global or local structural perturbations.[35] The CA^N(A92E) protein was crystallized using conditions different from those previously reported for the selenomethionine-substituted variant of this construct.[32] The eight molecules within the asymmetric unit of these crystals are very similar to those observed in previous CA^N and CA^N:cyclophilin X-ray structures, with the Phe32 side-chain buried within the folded core of the protein. Surprisingly, six of the eight molecules in the new CA^N(A92E) structure have well defined electron density for a *cis* Ala31–Phe32 peptide bond, whereas two molecules have density consistent with the previously observed *trans* peptide bonds (Figure 5). This contrasts with all previously reported CA^N structures, which have at least one well-defined *trans* Ala31–Phe32 conformation molecule and in some cases have additional molecules in the asymmetric unit that display ill-defined density but never show a well-defined *cis* Ala31–Phe32 peptide. Note that the cases of unclear densities are not easily modeled simply as a mixture of the well-defined *cis* and *trans* conformations.

The observation of a *cis* Ala31-Phe32 peptide bond in the CA^N crystal structure is surprising because *cis* peptide bonds are not typically seen before residues other than proline due to their relatively high conformational energy. The reason that the *cis* conformation is energetically accessible for Ala31–Phe32 is explained by inspection of the Ala31 main-chain. The Ala31 phi angle of *trans* conformation structures in the absence of CAP-1 is unfavorable (ranging from +43° to +57°) whereas Ala31 phi for the *cis* conformation is favorable (−60° to −75°). Thus, strain in the main-chain is manifest either as a *cis* conformation for the Ala31–Phe32 peptide or as an unfavorable phi angle for Ala31. In contrast, the CA^N structure crystallized in the presence of CAP-1 (i.e. Phe32 "out") displays both a *trans* Ala31–Phe32 peptide and a favorable Ala31 phi angle (−45°). Taken together, these observations suggest that the energetically unfavorable exposure of the Phe32 side-chain in the presence of CAP-1 is partially offset by relief of conformational strain in the main-chain. The NMR data obtained for CA^N do not exhibit signals or NOEs characteristic of a *cis* conformer, and we therefore believe that the Ala31–Phe32 bond exists predominantly as an ensemble of strained *trans* conformations under physiological conditions.

Steric strain involving non-Pro *cis* peptide bonds is rare in protein structures, and when present is usually associated with functional sites.[36,37] In this regard, Phe32 is conserved in 610 of 613 HIV-1 sequences present in the Los Alamos HIV-1 data base, and the remaining three sequences have a Phe side-chain shifted from position 32 by just one residue in the alignment.[38] Phe32 is also conserved in 63 of the 64 available HIV-2 Gag sequences (with the remaining one sequence conservatively substituted by Leu) and in 60 of the 67 available SIV sequences (substituted seven times by Trp). Taken together, these findings suggest that main-chain strain of Ala31–Phe32 may have been evolutionarily selected, perhaps to facilitate capsid assembly (see below).

Mechanism of inhibition of capsid assembly by CAP-1

CA^N functions in the assembly of both immature virions and the cone-shaped capsid that characterizes mature infectious virions. Electron microscopy (EM) studies have revealed that CA^N adopts a hexagonal lattice within the immature virion, but the precise domain orientation is not yet known.[39] The mature capsid adopts a fullerene organization in which the majority of the surface is formed by CA^N hexamers that are linked through CA^C dimers,[16,18] with five pentagonal defects distributed at the narrow end of the conical assembly and seven at the wide end.[13] Moderate resolution models for the hexagonal portion of the mature capsid have

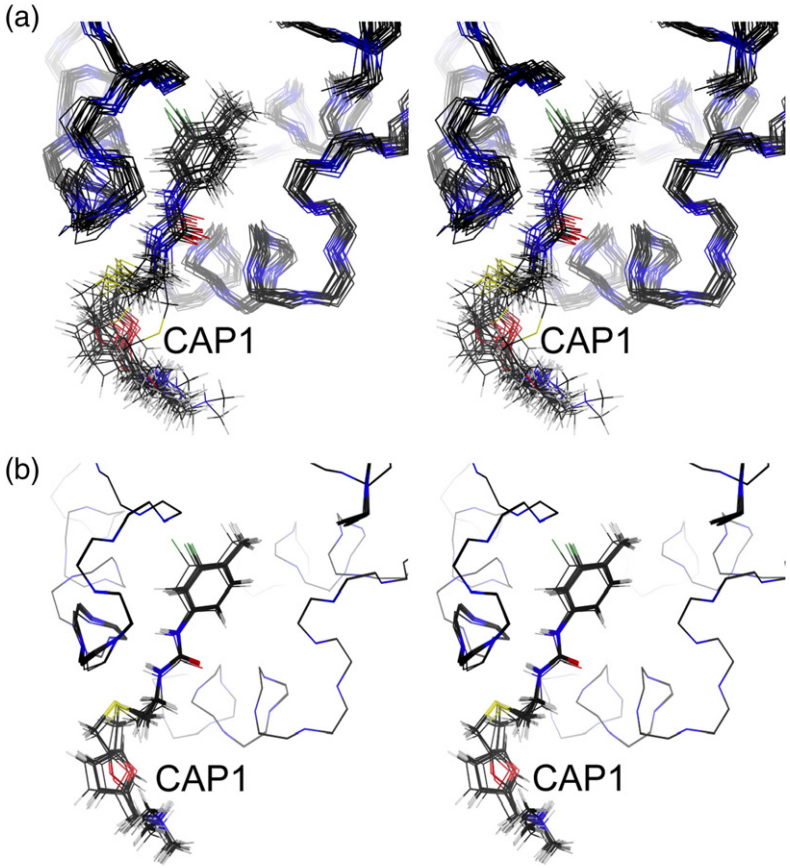


Figure 3. CAP-1:CA^N structures calculated by restrained molecular dynamics with AMBER using the hybrid X-ray/NMR approach. (a) Ensemble of 20 structures calculated after equilibration at 350 K using NOESY NMR-derived distance restraints. The positions of CA^N atoms with well-defined electron density were restrained to coordinates of the crystal structure. (b) Ensemble of 20 refined models obtained by energy minimization after cooling to 0 K.

been derived by docking high resolution CA^N and CA^C structures into EM reconstructions of *in vitro* assembled CA tubes[14] and 2D crystalline sheets (M. Yeager, personal communication). A crystal structure of Moloney murine leukaemia virus (MLV) CA^N also revealed a hexameric assembly[17] that allows construction of a HIV-1 CA^N hexamer by homology modelling (Figure 6). The MLV and HIV-1 CA^N hexamers are stabilized by intermolecular packing interactions between helices H1, H2, and H3, and perhaps also by weak interactions between the six N-terminal β-hairpins at the top of the

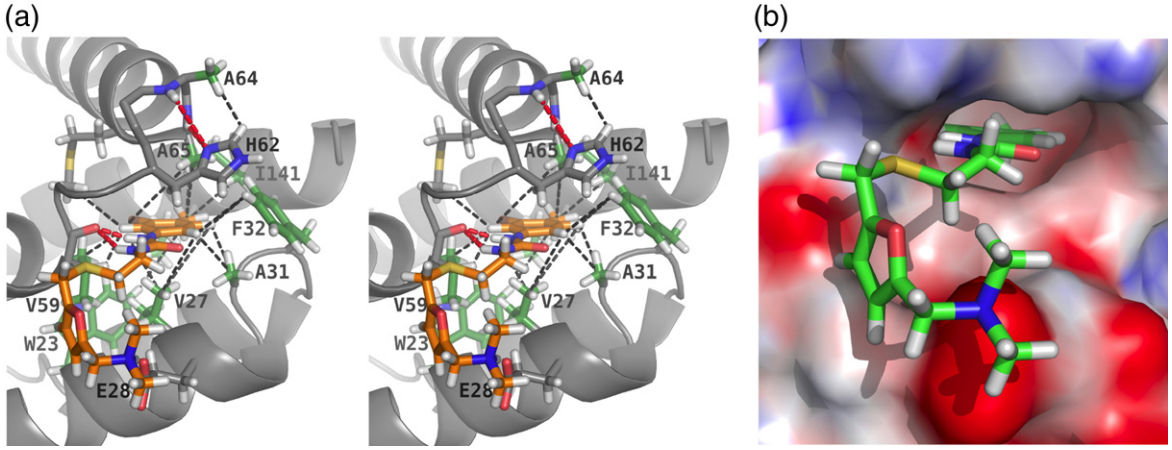


Figure 4. Representative CAP-1:CA^N structure calculated by restrained molecular dynamics using the hybrid X-ray/NMR approach. (a) Stereo view of the CAP-1 binding with observed NOEs (broken black lines) and potential hydrogen bonds (broken red lines) labeled. The side-chain of Phe32, which is displaced from the core upon CAP-1 binding, is also shown. (b) Electrostatic surface representation of the CAP-1 binding site showing the insertion of the CAP-1 aromatic ring into the pocket vacated by Phe32.

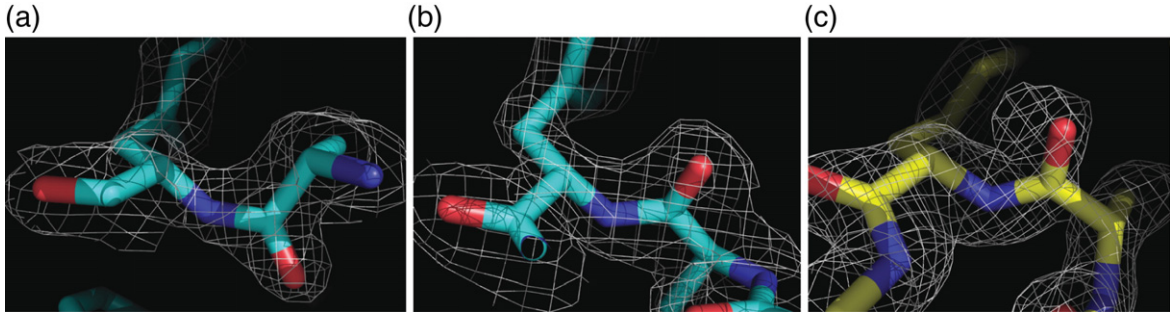


Figure 5. Buried Phe32 conformation seen in the absence of CAP-1 is associated with backbone strain. (a) Example of a *cis* Ala31–Phe32 peptide. (b) Example of a *trans* Ala31–Phe32 peptide. (c) The Ala31–Phe32 in the structure crystallized in the presence of CAP-1 (*trans*). Density is shown for simulated annealing omit maps.[61]

hexamer. The six C termini, which lead to the CA^C dimerization domains, are located at the bottom of the hexamer model (Figure 6(a)).

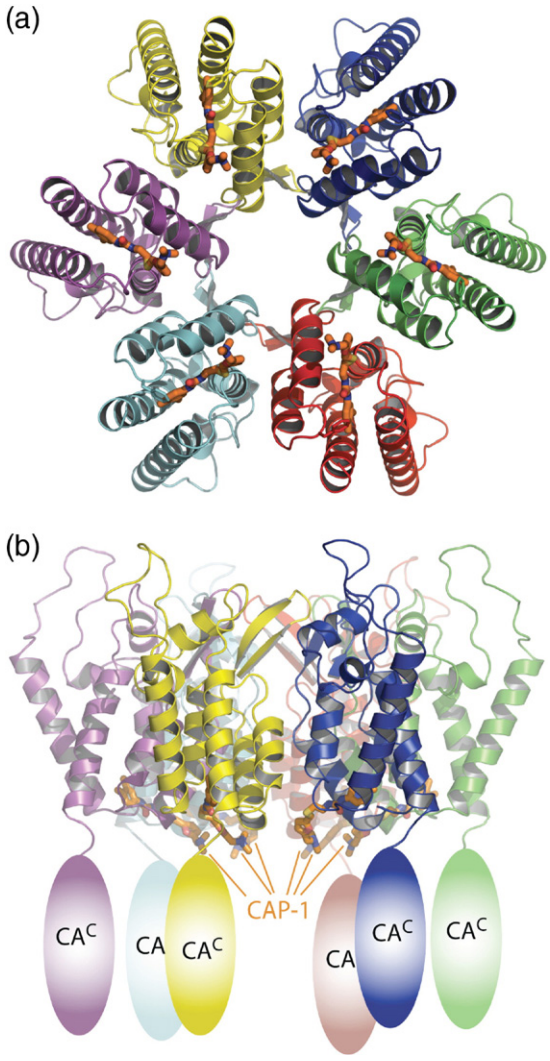


Figure 6. (a) Model of the HIV-1 CA^N hexamer of the mature capsid lattice modeled on the MLV CA^N structure.[17] CAP-1 is shown in stick representation. (b) Orthogonal view. The approximate location of CA^C is indicated.

The CAP-1 binding site and Phe32 are located on the bottom (inner) surface of the CA^N hexamer model. Binding of CAP-1 would not obviously impact CA^N hexamer formation because the nearest approach of a modelled CAP-1 atom to the adjacent CA^N subunit is ~5 Å, and occurs between a flexible CA^N side-chain and the exposed and presumably mobile CAP-1 furyl ring. In addition, CAP-1 does not appear to induce significant conformational changes away from its binding site. Similarly, it is difficult to envision how CAP-1 binding to CA^N would alter CA^C dimerization. It therefore seems most likely that CAP-1 binding inhibits formation of the third intermolecular interface, between C^N and CA^C domains of adjacent molecules within the hexamer. The importance of CA^N:CA^C interactions was initially suggested by genetic analyses of the RSV CA protein[24] and by hydrogen exchange experiments,[21,22,40] which indicated that several of the conserved residues adjacent to HIV-1 CA Phe32 participate in an intermolecular CA^N-CA^C interface upon capsid assembly.[21] Furthermore, the flexible side-chain of CA^N Lys70 is susceptible to chemical cross-linking with Lys182 on the CA^C domain of a second CA molecule during mild alkylation of assembled tubes,[21] and mutation of Lys70 inhibits CA tube assembly *in vitro*.[41]

Very recently, Ganser-Pornillos & Yeager (unpublished results) have visualized this third CA^N-CA^C interface at moderate (9 Å) resolution in cryo-EM reconstructions of 2D crystals of hexagonal CA arrays. Their studies show that CA^C domains pack in a groove between the N-terminal end of CA^N helix 4 and the C-terminal end of CA^N helix 7, i.e. in the region distorted by binding of CAP-1. Although side-chain detail is not available at the current resolution, it seems likely that the conformational changes that we observe upon CAP-1 binding would inhibit this packing interaction. One possibility is that exposure of the Phe32 side-chain, displacement of the CA^N helix 7 Tyr145 side-chain and repacking of the His62 side-chain against the N-terminal end of CA^N helix 4 might alter the surface of CA^N in a manner that inhibits its interactions with CA^C. An alternative possibility is that, during capsid assembly, CA^C binding to CA^N normally triggers a conformational change similar

to that observed upon CAP-1 binding, thereby allowing residues of CA^C to occupy the CAP-1 binding pocket. This mechanism would provide a biological explanation for the strained Phe32 backbone and the observed conformational switch, and would assume that CAP-1 functions as a direct competitor for the Phe32 pocket. Additional studies will be required to evaluate these or other potential CA^N-CA^C binding modes.

Potential for future drug development

Although CAP-1 exhibits antiviral activity in cell cultures at non-toxic doses, its affinity for CA^N (0.8 mM) is significantly below the levels needed for therapeutic use. The structure reported here of the CAP-1:CA^N complex provides details that may be useful for developing new assembly inhibitors with improved affinities. The cavity vacated by Phe32 encloses a volume of 264 $Å^3$ and presents a total of 249 $Å^2$ of Connoly molecular surface to its inner walls.[42] The buried portion of CAP-1 occupies only 194 $Å^3$ (73%) of the cavity volume in our joint NMR/X-ray structure, and modifications that improve the fit to the binding site should greatly improve inhibitor binding. The carbonyl oxygen of Ala31 is located within the largely hydrophobic pocket, and modifications that enable hydrogen bonding with this buried oxygen atom should similarly enhance binding. In addition, the backbone oxygen atom of Val59 and the NH groups of Gly61 and His62 are available for hydrogen bonding at the mouth of the cavity, and these groups might also be exploited to enhance binding affinity.

The discovery and development the non-nucleoside reverse transcriptase inhibitors (NNRTIs) followed a similar pathway. The NNRTIs bind to a pocket that forms only in the presence of inhibitors[43–45] and involves reorientation of aromatic sidechains from native buried positions.[44,45] In addition, the initially discovered NNRTIs exhibited relatively poor affinities for RT, but the affinities were substantially improved by structure-based drug design (for example, see Artico *et al.*[46]). These similarities provide grounds for optimism that useful inhibitors that target the CAP-1 binding site can be developed following a similar strategy. An unfortunate problem with the NNRTIs is that their binding site can readily accommodate mutations, and this has led to the development of resistance to this class of inhibitors. Because Phe32 is highly conserved, it is conceivable that the CAP-1 binding site may be less susceptible to drug-induced evolutionary pressure.

In summary, we have determined the structure of the CA^N:CAP-1 complex and identified structural features that may be exploited to enhance binding affinity. Binding involves a major reorientation of Phe32, which appears to be promoted by main-chain strain. We speculate that this strain may be evolutionarily conserved to allow structural changes associated with CA^N-CA^C interactions during capsid assembly, and that CAP-1 binding interferes with these interactions. Efforts to develop new inhibitors with improved efficacy using the CA^N:CAP-1 structure as a guide are underway.

Materials and Methods

Structure determination of CA^N crystallized in the presence of CAP-1

CA^N was prepared using a published procedure,[47] except that a final step was added in which protein was dialyzed into 10 mM Tris (pH 8.0), 50 mM NaCl, 5 mM β-mercaptoethanol and run on a S75 (Pharmacia) sizing column. Crystals were obtained overnight at 21 °C, 13 °C and 4 °C conducted by mixing solutions of CA^N (0.84 mM, with 10 mM Tris (pH 8.0), 50 mM NaCl, 5 mM β-mercaptoethanol) and CAP-1 (Maybridge Chemicals, Cornwall, England, 15.7 mM in DMSO-d_6) ([CA^N] = 0.72 mM; CAP-1:CA^N ratio = 6.1:1). Crystals grew in sitting drops that were prepared using a Hydra 96+1 crystallization robot (Robbins Scientific, Sunnyvale, California). The drop was a mixture of 0.5 μl CA^N:CAP-1 solution and 0.5 μl reservoir (100 mM Tris (pH 8.5), 5% (w/v) PEG 8000, 20% PEG 300, and 10% (v/v) glycerol). Data were collected from several crystals grown at each of these temperatures; in all cases the diffraction and map quality were similar, although the best data (reported here) was from a crystal grown at 13 °C. Crystals were mounted in a nylon loop and flash-cooled in liquid nitrogen without use of an additional cryo-protectant. X-ray diffraction data were processed with MOSFLM.[48] The single CA^N molecule in the asymmetric unit was located by molecular replacement using MOLREP.[49] Refinement was performed using REFMAC5[50] and map fitting was with XTALVIEW.[51] Figures were made with PyMol§.

Crystallization and structure determination of CA^N A92E

CA^N(A92E) was prepared and purified as described.[12] The sequence (Gag residues 133–278; CA residues 1–146) corresponds to wild-type HIV-1_{NL4-3} CA^N, except that Ala92 was replaced by Glu. The purified protein was dialyzed against 10 mM Tris (pH 8.0), 50 mM NaCl, 2 mM β-mercaptoethanol, run on a S75 gel filtration column (Pharmacia), and concentrated to 0.9 mM. Crystals grew after several weeks at 21 °C in sitting drops with a reservoir solution of 24% PEG 4500, 0.60 M $MgCl_2$, and 100 mM Tris–HCl (pH 8.5), and a drop of two parts protein solution to one part reservoir solution.[52] Crystals were briefly transferred to a cryo-protectant consisting of well solution supplemented with 10% glycerol, then suspended in a nylon loop and flash-cooled in liquid nitrogen. HKL and SCALEPACK[53] were used for data processing. The structure was determined *via* molecular replacement with PHASER[54] using a previously solved HIV-1 CA^N as the search model. Refinement used REFMAC5[50] in the CCP4 suite of programs.[55] Model building was done with Coot.[56]

Inspection of the diffraction pattern and indexing trials revealed that half of the reflections are systematically weak, indicating the presence of translational pseudo-

§ http://pymol.sourceforge.net/

symmetry. The diffraction pattern could therefore be indexed either in the true space group or in the pseudo-space group in which the systematically weak reflections are ignored and the *c* axis length is halved. Both the true and pseudo-cells belong to space group *P*1. The structure was initially solved by molecular replacement in the pseudo-space group and the four molecules of that asymmetric unit were refined before confirming the true cell structure by molecular replacement and independent refinement of all eight molecules. As expected for large numbers of systematically weak reflections, some statistics for the pseudo-cell (R_{free} = 21.7%; resolution = 1.45 Å) appear better than for the true cell (R_{free} = 26.8%; resolution = 1.90 Å). Resolution is defined as the Bragg spacing at which half of the measured reflections are less than twice their estimated standard deviation. Nevertheless, refinement statistics for the true cell indicate that this structure is of good quality (Table 2).

NMR data collection and analysis

Samples of CA^N containing a C-terminal His-6 tag were prepared using a modification of a described protocol.[30] The gene encoding the N-terminal domain (NTD) of HIV-1 CA (residues 1 through 151) with a C-terminal His-6 tag was PCR-amplified from pNL4-3, sub-cloned into pET-11a (Novagen) and subsequently transformed into BL21 competent cells (Stratagene). Cells were grown in LB medium or M9 minimal medium containing $^{15}NH_4Cl$ and/or[^{13}C]glucose (Isotec) as its sole nitrogen and/or carbon source. Protein expression was induced in shake flasks at an A_{600} of 0.6 with 1 mM IPTG. The cells were harvested and lysed with a microfluidizer (Microfluidics) and the protein was purified to homogeneity by cobalt affinity (Talon) and cation exchange (Amersham) chromatographies.

NMR data were collected on Bruker 600 and 800 MHz instruments equipped with cryo-probes at 11 °C or 35 °C using protein samples of 100 μM –700 μM CA^N in 25 mM sodium phosphate (pH 7.0), 5 mM DTT, 10% ^2H_2O, and 5% DMSO-d_6. NMR signals of CAP-1 were assigned by standard 2D NMR methods.[57] Resonances of the protons attached to carbons in free CAP-1 were assigned using natural abundance ^1H,^{13}C HMQC and 2D homonuclear NOESY data collected in 100% DMSO-d_6 at 35 °C. The signals of these protons of CAP-1 in the complex were assigned in H_2O by titrating increasing amounts of CAP-1 (0–3 mM) into 700 μM CA protein and monitoring with 2D homonuclear NOESY (τ_{mix} = 120 ms) experiments. Intermolecular ^1H-^1H NOEs for the complex were obtained with 3D ^{13}C-edited HMQC-NOESY, ^{15}N-edited NOESY-HSQC and 2D homonuclear NOESY data. NMR data were processed with NMRPIPE[58] and analyzed with NMRVIEW.[59]

Joint NMR/X-ray structure determination of the CAP-1:CA^N complex

Refinement of the CAP-1:CA^N complex was performed using the AMBER-9 program package[60] by docking a CAP-1 model into the cavity of the X-ray structure and performing restrained molecular dynamics followed by

Table 2. Crystallographic and refinement data

	CA^N(A92E) real cell	CA^N(A92E) pseudo-cell	CA^N (+ CAP-1)
Data collection	NSLS X29	NSLS X29	NSLS X12-B
Space group	*P*1	*P*1	$C222_1$
Unit cell lengths: *a*, *b*, *c* (Å)	48.3, 59.0, 92.3	46.2 48.2 58.9	42.2, 62.8, 106.3
Unit cell angles: α, β, γ (°)	71.5, 88.1, 83.0	83.0, 71.3, 87.4	90, 90, 90
Wavelength (Å)	1.1000	1.1000	1.0000
Resolution (Å)	32.2–1.90 (1.97–1.90)	30.0–1.45 (1.50–1.45)	53–1.50 (1.54–1.50)
Number of observed reflections	1, 156, 031	579, 474	255, 759
Number of unique reflections	74, 959	84, 901	23, 354
Completeness (%)	96.5 (95.4)	91.0 (61.6)	95.2 (72.5)
R_{sym} (%)	5.9 (12.4)	5.6 (29.4)	4.1 (42.4)
Average $I/\sigma(I)$	10 (2)	12 (2)	10 (2)
Mosaicity (°)	0.40	0.35	0.50
Refinement			
Working R_{factor} (%)	0.204	0.170	0.163
R_{free} (%)	0.270	0.216	0.221
$R_{overall}$ (%)	0.208	0.173	0.166
No. non-hydrogen atoms	9946	5366	1,374
Number of water molecules	797	533	179
RMSD: Bond lengths (Å)	0.020	0.017	0.020
Bond angles (°)	1.724	1.664	1.873
ϕ/ψ angles, non-Gly/Pro res.			
Most favored regions (%)	94.5	93.8	93.4
Additional allowed (%)	5.3	5.8	5.8
Generously allowed (%)	0.2	0.2	0.8
Disallowed regions (%)	0.0	0.2	0.0
<*B*>: main-chain atoms (Å^2)	16.4	17.5	21.3
side chain atoms (Å^2)	19.2	20.4	24.4
Water molecules (Å^2)	26.3	29.8	36.4

Values in parentheses refer to the high-resolution shell.
$R_{sym} = \sum (I - \langle I \rangle) / \sum (I)$.
R-factor = $\sum_{hkl} | F_o - F_c | / | F_o |$. R_{free} is as for $R_{working}$ but calculated for a randomly selected 5% of reflections not included in the refinement. $R_{overall}$ is using all reflections ($R_{free} + R_{working}$).

energy minimization. Initial CAP-1 coordinates and the associated force field library were generated with X-Leap,[60] which was then used to manually dock the CAP-1 model into the vacant pocket of the X-ray structure. The resulting initial complex was subjected to restrained molecular dynamics at 350 K (50 ps). Atoms of the crystal structure with well-defined electron density were restrained to initial reference coordinates with a 1.0 kcal/mol-Å^2 potential. Other atoms of the crystal structure were unrestrained. Intermolecular NOEs with strong, medium and weak intensities were used to assign upper-limit distance restraints of 2.7, 3.3 and 5.0 Å, respectively, and were implemented with a 20 kcal/mol-Å^2 potential. Restraints involving methyl pseudo-atoms were increased by 0.5 Å.

Atomic coordinates

The atomic coordinates have been deposited in the Protein Data Bank‖. PDB ID codes: 2pxr, CA^N crystallized in the presence of CAP-1; 2pwm, CA^N A92E true cell; 2pwo, CA^N A92E pseudo cell; 2jpr, CAP-1:CAN complex determined by the hybrid NMR/Xray crystallography approach.

Acknowledgements

This work was supported by the NIH (grants AI 30917 to M.F.S., AI 45405 to W.I.S. and C.P.H.), and the Offices of Biological and Environmental Research and of Basic Energy Sciences of the US Department of Energy/National Center for Research Resources. S.K. was supported by an NIH NRSA F31 fellowship (GM 076979). We thank Barbie Ganser-Pornillos and Mark Yeager (Scripps) for sharing their structure of HIV-1 CA prior to publication and David King (HHMI, University of California, Berkeley) and the beamline staff and organizers of RapiData 2005 for technical support. Use of the National Synchrotron Light Source, Brookhaven National Laboratory, was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under contract no. DE-AC02-98CH10886.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2007.07.070

References

1. UNAIDS & WHO (2007). IDS Epidemic Update - Dec 06. Joint United Nations Programme on HIV/AIDS (UNAIDS) and World Health Organization (WHO) 2006.
2. Richman, D. D. (2001). HIV chemotherapy. *Nature*, **410**, 995–1001.
3. Pillay, D., Taylor, S. & Richman, D. D. (2000). Incidence and impact of resistance against approved antiretroviral drugs. *Rev. Med. Virol.* **10**, 231–253.
4. Lengauer, T. & Sing, T. (2006). Bioinformatics-assisted anti-HIV therapy. *Nature Rev. Microbiol.* **4**, 790–797.
5. Mansky, L. M., Pearl, D. K. & Gajary, L. C. (2002). Combination of drugs and drug-resistant reverse transcriptase results in a multiplicative increase of human immunodeficiency virus type 1 mutant frequencies. *J. Virol.* **76**, 9253–9259.
6. Coffin, J. (1995). HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. *Science*, **267**, 483–489.
7. Kuritzkes, D. R. (1996). Clinical significance of drug resistance in HIV-1 infection. *AIDS*, **10**, S27–S33.
8. Lucas, G. M. (2005). Antiretroviral adherence, drug resistance, viral fitness and HIV disease progression: a tangled web is woven. *J. Antimicrob. Chemother.* **55**, 413–416.
9. Rossmann, M. G. (1988). Antiviral agents targeted to interact with viral capsid proteins and a possible application to human immunodeficiency virus. *Proc. Natl Acad. Sci. USA*, **85**, 4625–4627.
10. Briggs, J. A. G., Simon, M. N., Gross, I., Krausslich, H.-G., Fuller, S. D., Vogt, V. M. & Johnson, M. C. (2004). The stoichiometry of Gag protein in HIV-1. *Nature Struct. Mol. Biol.* **11**, 672–675.
11. Forshey, B. M., von Schwedler, U., Sundquist, W. I. & Aiken, C. (2002). Formation of a human immunodeficiency virus type 1 core of optimal stability is crucial for viral replication. *J. Virol.* **76**, 5667–5677.
12. Gitti, R. K., Lee, B. M., Walker, J., Summers, M. F., Yoo, S. & Sundquist, W. I. (1996). Structure of the amino-terminal core domain of the HIV-1 capsid protein. *Science*, **273**, 231–235.
13. Ganser, B. K., Li, S., Klishko, V. Y., Finch, J. T. & Sundquist, W. I. (1999). Assembly and analysis of conical models for the HIV-1 core. *Science*, **283**, 80–83.
14. Li, S., Hill, C. P., Sundquist, W. I. & Finch, J. T. (2000). Image reconstructions of helical assemblies of the HIV-1 CA protein. *Nature*, **407**, 409–413.
15. Briggs, J. A., Wilk, T., Welker, R., Krausslich, H.-G. & Fuller, S. D. (2003). Structural organization of authentic, mature HIV-1 virions and cores. *EMBO J.* **22**, 1707–1715.
16. Gamble, T. R., Yoo, S., Vajdos, F. F., von Schwedler, U. K., Korthylake, D. K., Wang, H. *et al.* (1997). Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. *Science*, **278**, 849–853.
17. Mortuza, G., Haire, L. F., Stevens, A., Smerdon, S. J., Stoye, J. P. & Taylor, I. A. (2004). High-resolution structure of a retroviral capsid hexameric amino-terminal domain. *Nature*, **431**, 481–485.
18. Worthylake, D. K., Wang, H., Yoo, S., Sundquist, W. I. & Hill, C. P. (1999). Structures of the HIV-1 capsid protein dimerization domain at 2.6 A resolution. *Acta Crystallog. sect. D*, **55**, 85–92.
19. Ivanov, D., Stone, J. R., Maki, J. L., Collins, T. & Wagner, G. (2005). Mammalian SCAN domain dimer is a domain-swapped homolog of the HIV capsid C-terminal domain. *Mol. Cell*, **17**, 137–143.
20. Ivanov, D., Tsodikov, O. V., Kasanov, J., Ellenberger, T., Wagner, G. & Collins, T. (2007). Domain-swapped dimerization of the HIV-1 capsid C-terminal domain. *Proc. Natl Acad. Sci. USA*, **104**, 4353–4358.
21. Lanman, J., Lam, T. T., Barnes, S., Sakalian, M., Emmett, M. R., Marshall, A. G. & Prevelige, P. E., Jr

‖ http://www.pdb.org

(2003). Identification of novel interactions in HIV-1 capsid protein assembly by high-resolution mass spectrometry. *J. Mol. Biol.* **325**, 759–772.
22. Lanman, J., Lam, T. T., Emmett, M. R., Marshall, A. G., Sakalian, M. & Prevelige, P. E. J. (2004). Key interactions in HIV-1 maturation identified by hydrogen-deuterium exchange. *Nature Struct. Mol. Biol.* **11**, 676–677.
23. Ganser, B. K., Cheng, A., Sundquist, W. I. & Yeager, M. (2003). Three-dimensional structure of the M-MuLV CA protein on a lipid monolayer: a general model for retroviral capsid assembly. *EMBO J.* **22**, 2886–2892.
24. Bowzard, J. B., Wills, J. W. & Craven, R. C. (2001). Second-site suppressors of Rous sarcoma virus CA mutations: evidence for interdomain interactions. *J. Virol.* **75**, 6850–6856.
25. Deres, K., Schroder, C. H., Paessens, A., Goldmann, S., Hacker, H. J., Weber, O. *et al.* (2003). Inhibition of hepatitis B virus replication by drug-induced depletion of nucleocapsids. *Science*, **299**, 893–896.
26. Smith, T. J., Kremer, M. J., Luo, M., Vriend, G., Arnold, E., Kamer, G. *et al.* (1986). The site of attachment in human rhinovirus 14 for antiviral agents that inhibit uncoating. *Science*, **233**, 1286–1293.
27. Sticht, J., Humbert, M., Findlow, S., Bodem, J., Muller, B., Dietrich, U. *et al.* (2005). A peptide inhibitor of HIV-1 assembly in vitro. *Nature Struct. Mol. Biol.* **12**, 671–677.
28. Ternois, F., Sticht, J., Duquerroy, S., Krausslich, H.-G. & Rey, F. A. (2005). The HIV-1 capsid protein C-terminal domain in complex with a virus assembly inhibitor. *Nature Struct. Mol. Biol.* **12**, 678–682.
29. Aiken, C. & Chen, C. H. (2005). Betulinic acid derivatives as HIV-1 antivirals. *Trends Mol. Med.* **11**, 31–36.
30. Tang, C., Loeliger, E., Kinde, I., Kyere, S., Mayo, K., Barklis, E. *et al.* (2003). Antiviral inhibition of the HIV-1 capsid protein. *J. Mol. Biol.* **327**, 1013–1020.
31. Howard, B. R., Vajdos, F. F., Li, S., Sundquist, W. I. & Hill, C. P. (2003). Structural insights into the catalytic mechanism of cyclophilin A. *Nature Struct. Mol. Biol.* **10**, 475–481.
32. Kelly, B. N., Howard, B. R., Wang, H. E., Robinson, H., Sundquist, W. I. & 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.
33. Aberham, C., Weber, S. & Phares, W. (1996). Spontanous mutation in the human immunodeficiency type 1 gag gene that affect viral replication in the presence of cyclosporins. *J. Virol.* **70**, 3536–3544.
34. Braaten, D., Aberham, C., Franke, E. K., Yin, L., Phares, W. & Luban, J. (1996). Cyclosporin A-resistant human immunodeficiency virus type 1 mutants demonstrate that Gag encodes the functional target of cyclophilin A. *J. Virol.* **70**, 5170–5176.
35. Campos-Olivas, R. & Summers, M. F. (1999). Backbone dynamics of the N-terminal domain of the HIV-1 Capsid protein and comparison with the G94D mutant conferring cyclosporin resistance/dependance. *Biochemistry*, **38**, 10262–10271.
36. Herzberg, O. & Moult, J. (1991). Analysis of the steric strain in the polypeptide backbone of protein molecules. *Proteins: Struct. Funct. Genet.* **11**, 223–229.
37. Jabs, A., Weiss, M. S. & Hilgenfeld, R. (1999). Non-proline cis peptide bonds in proteins. *J. Mol. Biol.* **286**, 291–304.
38. Leitner, T., Foley, B., Hahn, B., Marx, P., McCutchan, F., Mellors, J. *et al.* (2006). *Theoretical Biology and Biophysics Group.* Los Alamos National Laboratory, LA-UR 06-0680.
39. Wright, E. R., Schooler, J. B., Ding, H. J., Kieffer, C., Fillmore, C., Sundquist, W. I. & 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.
40. Lanman, J. & Prevelige, P. E. J. (2005). Kinetic and mass spectrometry-based investigation of human immunodeficiency virus type 1 assembly and maturation. *Adv. Virus Res.* **64**, 285–309.
41. Ganser-Pornillos, B. K., von Schwedler, U. K., Stray, K. M., Aiken, C. & Sundquist, W. I. (2004). Assembly properties of the human immunodeficiency virus type 1 CA protein. *J. Virol.* **78**, 2545–2552.
42. Binkowski, T. A., Naghibzadeh, S. & Liang, J. (2003). CASTp: Computed Atlas of Surface Topography of proteins. *Nucl. Acids Res.* **31**, 3352–3355.
43. Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A. & Steitz, T. A. (1992). Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science*, **256**, 1783–1790.
44. Ren, J., Esnouf, R., Garman, E., Somers, D., Ross, C., Kirby, I. *et al.* (1995). High resolution structures of HIV-1 RT from four RT-inhibitor complexes. *Nature Struct. Biol.* **2**, 293–302.
45. Ding, J., Das, K., Moereels, H., Koymans, L., Andries, K., Janssen, P. A. J. *et al.* (1995). Structure of HIV-1 RT/TIBO R 86183 complex reveals similarity in the binding of diverse non-nucleoside inhibitors. *Nature Struct. Biol.* **2**, 407–415.
46. Artico, M., Silvestri, R., Pagnozzi, E., Bruno, B., Novellino, E., Greco, G. *et al.* (2000). Structure-based design, synthesis, and biological evaluation of novel pyrrolyl aryl sulfones: HIV-1 non-nucleoside reverse transcriptase inhibitors active at nanomolar concentrations. *J. Med. Chem.* **43**, 1886–1891.
47. von Schwedler, U. K., Stemmler, T. L., Klishko, V. Y., Albertine, K. H., Davis, D. R. & Sundquist, W. I. (1998). Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly. *EMBO J.* **17**, 1555–1568.
48. Leslie, A. G. (2006). The integration of macromolecular diffraction data. *Acta Crystallog. sect. D*, **62**, 48–57.
49. Vagin, A. & Teplyakov, A. (2000). An approach to multi-copy search in molecular replacement. *Acta Crystallog. sect. D*, **56**, 1622–1624.
50. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. sect. D*, **53**, 240–255.
51. McRee, D. E. (1999). XtalView/Xfit–A versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **125**, 156–165.
52. Van Woerkom, R., Dixon, A., Oslund, R. & Howard, B. R. (2007). Identification, optimization and preliminary X-ray diffraction of a new crystal form of the N-terminal domain of the HIV-1 CA protein. *Amer. J. Undergrad. Res.* **6**(1), 1–6.
53. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–327.
54. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). Likelihood-enhanced fast translation functions. *Acta Crystallog. sect. D*, **61**, 458–464.
55. Collaborative Computational Project, No. 4 (1994). The CCP4 suite: Programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.

56. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallog. sect. D*, **60**, 2126–2132.
57. Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids.* John Wiley and Sons, New York.
58. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
59. Johnson, B. A. & Blevins, R. A. (1994). NMRview: a computer program for the visualization and analysis of NMR data. *J. Biomol. NMR*, **4**, 603–614.
60. Case, D. A., Cheatham, T. E. I., Darden, T., Gohlke, H., Luo, R., Merz, K. M. J. *et al.* (2005). The AMBER biomolecular simulation programs. *J. Computat. Chem.* **26**, 1668–1688.
61. Brunger, A. T., Adams, P. D. & Rice, L. M. (1997). New applications of simulated annealing in X-ray crystallography and solution NMR. *Structure*, **5**, 325–336.

SUPPLEMENTARY MATERIAL

for

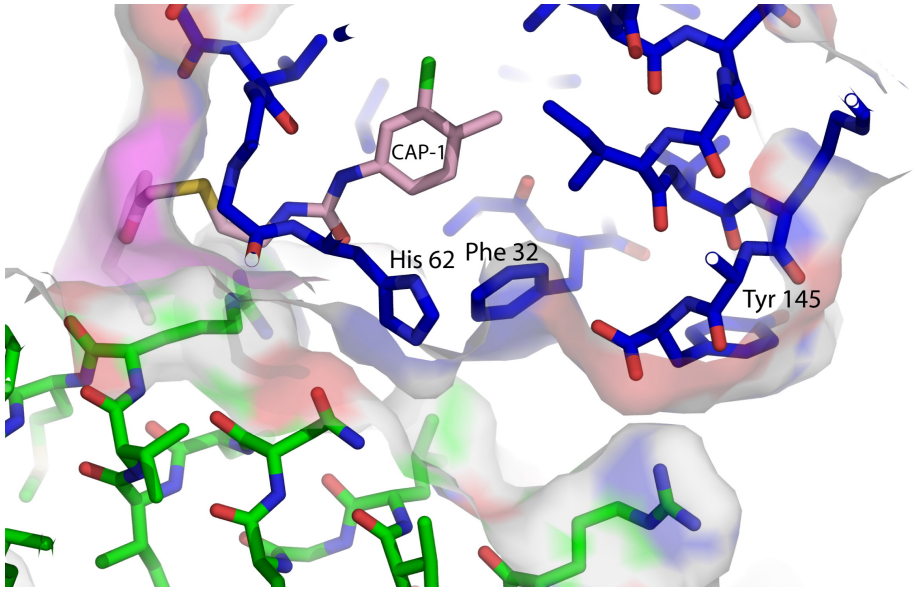
Structure of the antiviral assembly inhibitor CAP-1 bound to the HIV-1 CA protein

Brian N. Kelly[=§], Sampson Kyere[=#], Isaac Kinde[#], Chun Tang[#], Bruce R. Howard[%] Howard Robinson[∞], Wesley I. Sundquist[*§], Michael F. Summers[*#], Christopher P. Hill[*§]

[§]Department of Biochemistry, University of Utah, Salt Lake City, UT 84112-5650, USA; [#]Department of Chemistry and Biochemistry, University of Maryland, Baltimore County, Baltimore, MD 21250, USA; [%]Department of Physical Science, Southern Utah University, Cedar City, UT 84720; and [∞]Biology Department, Brookhaven National Laboratory, Upton, NY 11973-5000

[=]These authors contributed equally to this work.

*To whom correspondence should be addressed: W.I.S.: tel, 801-585-5402; e-mail, wes@biochem.utah.edu. M.F.S: tel 410-455-2880; e-mail summers@hhmi.umbc.edu. C.P.H: tel, 801-585-5536; e-mail, chris@biochem.utah.edu.



Supplemental Figure S1. Crystal lattice contacts near the CAP-1 binding site. The unique CA^N molecule in the CAP-1:CA^N crystal asymmetric unit (blue) and a neighboring symmetry-related molecule (green) are shown. The CAP-1 model obtained from NMR data and modeling is shown in stick representation (magenta). Residues that are displaced relative to CA^N crystallized in the absence of CAP-1 (Phe32, His62, Tyr145) are near this interface and labeled explicitly.