



Published in final edited form as:

Nat Struct Mol Biol. 2023 March ; 30(3): 383–390. doi:10.1038/s41594-022-00913-5.

A molecular switch modulates assembly and host factor binding of the HIV-1 capsid

Randall T. Schirra[§],

Nayara F. B. dos Santos[§],

Kaneil K. Zadrozny,

Iga Kucharska[#],

Barbie K. Ganser-Pornillos^{*}, Owen Pornillos^{*}

Department of Molecular Physiology and Biological Physics, University of Virginia, Charlottesville, Virginia, USA

Abstract

The HIV-1 capsid is a fullerene cone made of quasi-equivalent hexamers and pentamers of the viral CA protein. Typically, quasi-equivalent assembly of viral capsid subunits is controlled by a molecular switch. Here, we identify a Thr-Val-Gly-Gly motif that modulates CA hexamer/pentamer switching by folding into a 3_{10} helix in the pentamer and random coil in the hexamer. Manipulating the coil/helix configuration of the motif allowed us to control pentamer and hexamer formation in a predictable manner, thus proving its function as a molecular switch. Importantly, the switch also remodels the common binding site for host factors that are critical for viral replication and the new ultra-potent HIV-1 inhibitor, lenacapavir. These studies reveal that a critical assembly element also modulates the post-assembly, viral replication functions of the HIV-1 capsid, and provide new insights on capsid function and inhibition.

Upon entry into a new host cell, the HIV-1 capsid performs multiple essential functions, which include shielding the genome from innate immune sensors¹, promoting reverse transcription² and transporting the core from the entry site at the plasma membrane to the integration site inside the nucleus^{3,4}. The fullerene cone architecture of the HIV-1 capsid represents an extreme case of quasi-equivalence, in which essentially each of the 1,200 or more copies of the viral CA protein occupies a different chemical environment and thus adopts a different structural configuration⁵⁻⁷. Along the body of the cone, the CA subunits are assembled on a hexagonal lattice, with the CA N-terminal domain (NTD) making hexameric rings that are linked together by the C-terminal domain (CTD). In this part of the capsid, conformational variability afforded by CA's two-domain organization accounts for quasi-equivalence, which manifests in the continuously changing curvature of the hexagonal

*Correspondence: bpornillos@virginia.edu, opornillos@virginia.edu.

[§]Equal contribution

[#]Current affiliation: The Peter Gilgan Centre for Research and Learning, Hospital for Sick Children, Toronto, Ontario, Canada
Contributions

Cloning, mutagenesis, protein purification and assembly: R.T.S., N.F.B.d.S., K.K.Z., I.K. Negative stain EM: R.T.S., N.F.B.d.S., B.K.G.-P. CryoEM sample preparation and data collection: R.T.S., N.F.B.d.S., B.K.G.-P. Image data processing: R.T.S., B.K.G.-P., O.P. Coordinate modeling and refinement: R.T.S., O.P. Project conceptualization and funding acquisition: B.K.G.-P. and O.P. Protocol design and development: R.T.S., N.F.B.d.S., I.K., B.K.G.-P., O.P. Manuscript preparation: O.P. with input from all authors.

lattice along the body of the cone^{7,8}. Twelve CA pentamers occupy sharp points of curvature – termed declinations – exactly 12 of which are required for the capsid to completely close⁵. The molecular rules that dictate CA hexamer or pentamer formation have not been previously elucidated. In other studied quasi-equivalent viruses, the capacity of genetically identical subunits to form hexamers and pentamers is conferred by molecular switches^{9,10}. Such a switch has not been previously found in retroviral CA proteins. Furthermore, the relative contributions of CA hexamers and pentamers to HIV-1 capsid interactions with host factors in the post-entry pathway have also been unknown. To address these key gaps, we performed cryoEM and biochemical analyses of pentamer-containing in vitro capsids and their complexes with host factors that modulate capsid stability, promote nuclear import, and direct integration site targeting.

Results

Preparation and single particle averaging analysis of conical HIV-1 capsids in vitro

The classic in vitro model system for the HIV-1 capsid are tubular capsid-like particles (CLPs), which assemble when purified HIV-1 CA protein is incubated in high salt concentrations (1 M NaCl) and basic pH^{6,11}. These tubes are made only of CA hexamers. Recently, the cellular metabolite inositol hexakisphosphate (IP6) was shown to stabilize the HIV-1 capsid, by binding to the central channel of the CA hexamer^{12,13}. IP6 also induces assembly of CA in vitro, even at physiological salt conditions¹². We found that IP6-induced CLPs have a range of morphologies – capped tubes and cones – when assembled at pH 8, but are primarily conical at pH 6 (Extended Data Fig. 1). These observations indicate that IP6 promotes incorporation of CA pentamers into the assembling lattice, allowing formation of declinations that close the capsid shell. By adapting deep 2D classification particle selection¹⁴ and lattice-mediated alignment strategies¹⁵, we determined the structure of the declination from projection images of the IP6-induced conical CLPs (Fig. 1a,b, Extended Data Fig. 2, Extended Data Table 1). Local refinement with imposed five-fold symmetry resulted in a high-quality map at a nominal resolution of 3.4 Å, in which the central pentamer and surrounding hexamers are well-defined (Fig. 1c, Extended Data Fig. 2d, Movie 1). Both the pentamer (Fig. 1d) and hexamer (Fig. 1e) are similar to their corresponding lower-resolution “in situ” structures that were determined by sub-tomogram averaging of capsids in intact virions^{7,16} (Extended Data Fig. 3). This correspondence indicates that the conical IP6-induced HIV-1 CA assemblies more closely mimic the architecture of the HIV-1 capsid compared to previously described in vitro model systems.

IP6 stabilizes the pentameric HIV-1 CA ring

Both the hexamer and pentamer contain two IP6 molecules inside their central channels, one above and one below the ring of positively-charged Arg18 sidechains; the lower IP6 molecules are also coordinated by a second ring of Lys25 sidechains (Fig. 1f). These support the model that charge neutralization promotes formation of the hexamer and pentamer. However, removal of the charges have different effects on CA assembly in vitro. Whereas the R18A substitution abolishes IP6-dependent assembly of CA altogether¹², the K25A mutant remains competent in assembling tubes¹⁷. These observations suggest that Arg18 interactions stabilize both the hexamer and pentamer, whereas Lys25 interactions are more

important for the pentamer. Consistent with this interpretation, the pentamer channel is narrower at the position of the Lys25 ring (dashed circles in Fig. 1f), allowing close, direct contacts between the primary amines of the lysine sidechains and the phosphates of the lower IP6 molecule. Thus, our structure indicates that IP6 contacts with Arg18 also directly stabilize the CA hexamer and contacts with both Arg18 and Lys25 stabilize the pentamer.

Quasi-equivalence of the HIV-1 CA hexamer and pentamer

In quasi-equivalent assembly systems, formation of different oligomers by the same protein is controlled by a conformational switch, which adopts different configurations that are evident from comparison of high-resolution structures of the quasi-equivalent states¹⁸. Superposition of the HIV-1 CA subunits in the hexamer and pentamer (Fig. 2a–b), allowed us to identify a putative switch in the NTD, comprising a Thr58-Val59-Gly60-Gly61 (TVGG) motif that adopts two alternative configurations – random coil in the hexamer and 3₁₀ helix in the pentamer (Fig. 2c). These configurations are well-defined in our cryoEM map (Extended Data Fig. 4a–d, Movie 2). The interpretation that the TVGG motif functions as a switch is supported by its structural context: the motif mediates alternative packing of the NTD-NTD and NTD-CTD interfaces that hold together the capsomers (Fig. 2d,g,j).

In the hexamer, the NTD-NTD interface consists of a 3-helix repeating unit (helix 2 of one subunit and helices 1 and 3 of the adjacent subunit) that makes an 18-helix barrel (Fig. 2e). In the pentamer, helix 3 is excluded from the repeating unit, which then makes a 10-helix barrel (Fig. 2f). Refolding of the TVGG motif facilitates switching from the 3-helix to the 2-helix unit through a ratcheting mechanism. In the hexamer, Thr58 together with Asn57 are the C-terminal residues of helix 3 and engage in reciprocal knobs-in-holes packing against Pro38 and Met39 in helix 2 of the neighboring subunit (Fig. 2e, boxed). In the pentamer, the protein backbone overwinds at Thr58 to facilitate the α -to-3₁₀ helix transition (Fig. 2f, boxed). The conformational change disfavors the knobs-in-holes packing of helices 2 and 3, and instead Pro38-Met39 now packs against Val24-Lys25 in helix 1. Thus, analogous to the “pawl” of a ratchet, Pro38-Met39 in helix 2 engages two alternative “gears”: Asn57-Thr58 in helix 3 to form the hexameric NTD-NTD interface or Val24-Lys25 in helix 1 to form the pentameric interface (Fig. 2g). Refolding of the TVGG motif toggles alternative packing by remodeling the hexamer gear, allowing the pawl to “slip” into the pentamer state (Fig. 2h).

The TVGG motif also mediates alternative packing of the NTD-CTD interface (Fig. 2i–k). In the hexamer, helix 8 of the CTD is adjacent to the helix 3/4 loop (Fig. 2i). The highly conserved Arg173 residue in this helix makes two hydrogen bonds (green in Fig. 2i): one that caps the C-terminal end of helix 3 (Asn57) and another to the carbonyl of Val59. These are part of an extensive hydrogen bonding network (Extended Data Fig. 5a) that we previously showed to facilitate pivoting of the CTD to accommodate the variable curvature of the hexagonal capsid lattice¹⁹. Our results suggest an additional function in stabilizing both the random coil configuration of the switch and the α -helical character of helix 3. In the pentamer, the helix 3/4 loop with the 3₁₀ helix configuration no longer interacts with helix 8 due to repositioning of the CTD (Fig. 2j). Instead, Arg173 makes hydrogen bonds (green in Fig. 2j) with a different partner (Glu28) in helix 1. Also, the C-terminal end of

helix 7 (NTD) and portions of the NTD-CTD linker pack against the groove between helices 8 and 11 (CTD) (Extended Data Fig. 5a).

In summary, the TVGG motif within the helix 3/4 loop of HIV-1 CA adopts two distinct secondary structural folds – random coil and 3_{10} helix – each of which engages a distinct set of quaternary packing interactions in the hexamer and pentamer. The location of the motif allows it to simultaneously modulate both the NTD-NTD and NTD-CTD interfaces (Fig. 2d,g,k), explaining how these two sets of interactions cooperate during capsid assembly.

The TVGG motif remodels the hydrophobic core of the NTD

Refolding of the TVGG loop moves these residues towards the protein hydrophobic core, and thereby changes or remodels the core. A large hidden pocket in the hexameric NTD core becomes smaller in the pentamer because the empty space is partly occupied by the repositioned Val59 sidechain (Extended Data Fig. 5b). Proximal to the TVGG motif, Met66 changes its rotamer configuration, which we describe here to function as a gate (Fig. 2c). In the hexamer state, the Met66 gate is partly buried in the hydrophobic core in a “closed” configuration and becomes exposed in an “open” configuration in order to avoid clash with the folded 3_{10} helix in the pentamer state. More distal to the TVGG motif, the Ala31-Phe32 peptide bond in the helix 1/2 loop rotates by 180° compared to the hexamer (Extended Data Fig. 5c). The Phe32 sidechain points towards the hydrophobic core of the protein and lines the hidden pocket in both states.

We propose that these conformational changes are functionally linked, and identify an allosteric network that communicates IP6-mediated interactions in helices 1 and 2 (NTD-NTD interface) to the TVGG switch.

Importance of the TVGG motif and associated elements in virus infectivity and inhibition

All of the residues identified above are important for viral replication or inhibition. Substitutions for Val59 and Gly60 within the TVGG motif result in non-viable or severely impaired viruses^{20–23}. Mutations in the Pro38-Met39 pawl in helix 2 abolishes CA assembly²⁴ or drastically lowers the intrinsic stability of the capsid^{25,26}. Asn57 and the Met66 gate modulate binding of the assembled capsid to host factors important for nuclear import and integration, as well as binding to one class of capsid-targeting inhibitors^{27–36}. The hidden pocket in the NTD core that is walled by Phe32 and Val59 is also a binding site for a different class of inhibitors^{20,37}. Thr58 and Met66 mutations and polymorphisms confer drug resistance, although at severe costs to viral fitness^{20,33–35}. Thus, the TVGG motif is positioned at a critical site for capsid function and inhibition, indicating that it may have post-assembly functions as well.

Designed manipulations of the TVGG switch predict HIV-1 CA assembly

To test our hypothesis that the TVGG motif is a molecular switch, we designed structure-based mutations to control the coil/helix configuration, which should predict the in vitro assembly behavior of the CA protein and resulting CLP morphology (Fig. 3a). Solution conditions can also affect CLP morphology (Extended Data Fig. 1a–c), and so for uniformity we initially assayed for assembly using the IP6-induced conditions at pH 6,

which in our hands produces conical CLPs with wild type (WT) CA (Fig. 3b, Extended Data Fig. 1c).

WT HIV-1 CA is known to favor hexamer formation *in vitro*, implying that the 3_{10} helix configuration of the TVGG motif is normally disfavored. To stabilize the 3_{10} helix and thus favor the pentamer, we replaced the two glycines in the TVGG motif with alanine and proline. Crude modeling indicated that these two substitutions are compatible with the backbone configuration of the dipeptide in the pentamer state but not the hexamer state. Indeed, incubation of CA G60A/G61P with IP6 induced assembly of small spheres of around 20 nm in diameter (Fig. 3c, Extended Data Fig. 1d). CryoEM imaging indicated that the spheres are T=1 particles, or icosahedra made of 12 pentamers and no hexamers (Fig. 3f). Next, we surmised that the Met66 gate must constitute part of the activation barrier for coil-to-helix switching. Thus, a constitutively open gate (M66A) should also favor pentamers, and indeed the CA M66A mutant also assembled into T=1 particles (Extended Data Fig. 1e), as did the G60A/G61P/M66A triple mutant (Extended Data Fig. 1f). We solved the cryoEM structures of these T=1 capsids to 2.2 Å (G60A/G61P), 3.1 Å (M66A), and 2.4 Å (G60A/G61P/M66A) (Fig. 3f, Extended Data Fig. 6, Extended Data Table 1). These confirmed that the mutant pentamers have the same structure as the WT pentamer, with the switch in the 3_{10} helix configuration and the gate in the open configuration (Extended Data Fig. 4e–g).

We then asked if the CA mutants can still form hexamers by assembling them in 1 M NaCl and pH 8, conditions in which WT CA assembles only open tubes (Extended Data Fig. 1a). CA with the G60A/G61P mutations only formed aggregates, supporting the interpretation that the mutations locked the protein in its pentamer state (Extended Data Fig. 1d,f). Interestingly, the M66A mutant also failed to assemble in three independent experiments (Extended Data Fig. 1e) even though the mutation does not directly modify the TVGG switch. This result argues against a simple model wherein the TVGG motif fluctuates between the two states pre-assembly, and then gets locked in upon assembly. We speculate that the random coil configuration may be analogous to a restrained spring that is poised to refold, but is prevented from doing so by the closed Met66 gate.

We were unable to identify amino acid substitutions to rigidify the random coil configuration of the TVGG switch, which we predicted would have locked CA in its hexamer state. Nevertheless, structure-based mutations can be introduced to disfavor switching into the pentamer. As described above, Val24 stabilizes the pentameric NTD ring by mediating the knobs-in-holes interaction of helices 1 and 2 (Fig. 3d). Accordingly, the V24D mutation favored assembly of long, tubular CLPs when incubated with IP6 at pH 6 (Fig. 3d, Extended Data Fig. 1g,l).

Finally, we tested modifications to the protein hydrophobic core and its potential allosteric role in hexamer/pentamer switching. We predicted that increasing the packing density of the core would inhibit core remodeling and pentamer formation. To achieve this without modifying the TVGG motif, we replaced Leu138 with larger sidechains; this residue is in helix 7, across the internal pocket from Val59 and does not directly interact with the switch (Extended Data Fig. 5b). L138F, L138W and L138Y mutations significantly destabilized

CA, but L138F and L138W still formed cones. However, the CLPs were generally longer than WT, indicating impaired pentamer formation (Fig. 3d–e, Extended Data Fig. 1h–i,l). L138F also made very long (>1 μm) tubes. While these tubes were fewer in number, each contains tens of thousands of CA subunits in the hexamer state, again indicating a lower propensity for pentamers. We also tested the importance of the conformational change in the helix 1/2 loop involving Ala31 and Phe32 (Extended Data Fig. 5c). As with the Leu138 substitutions, A31G and F32A also destabilized the protein and reduced assembly efficiency, yet still supported cone formation. F32A had a higher proportion of tubular capsids, again indicating that pentamer formation is disfavored (Extended Data Fig. 1j–l). These results support the idea that the protein hydrophobic core allosterically connects helix 1, where IP6 binds, and the TVGG switch.

The HIV-1 CA hexamer/pentamer switch modulates ligand binding

One of the critical functions of the HIV-1 capsid is to recruit host cell factors that promote viral replication in terminally-differentiated T cells and macrophages, the natural hosts of HIV-1. Because these interactions are typically studied biochemically and structurally in context of the HIV-1 CA hexamer, and because genetic approaches cannot distinguish between the hexamer and pentamer, any potential contributions from the pentamer have been unknown. A major class of such capsid-binding factors include CPSF6, NUP153 and SEC24C, each of which contains one or more phenylalanine-glycine (FG) motifs that bind within the NTD-CTD interface of the hexamer^{28–30,32,36,38,39}. Binding of these natural FG ligands is competitively inhibited by pharmacological inhibitors, PF74 (refs. ^{29–31,39}) and GS-CA1/GS-6207/lenacapavir^{33–35}, each of which contains an analogous phenyl ring. The FG-motif phenylalanine ring binds to a hydrophobic pocket on the surface of the NTD that is adjacent the helix 3/4 loop containing the TVGG switch (Fig. 4a,c). This NTD pocket alone is sufficient for weak binding^{27,28,39}. Tighter binding is afforded by additional contacts with the CTD, which adopts a “closed” configuration and contacts ligand moieties outside of the phenyl ring^{29,30}. In essence, the NTD pocket makes the basal interaction with the FG motif or its equivalent (μM affinity), whereas surrounding elements contribute to tighter binding (pM for lenacapavir³⁴).

Our structures indicate that the alternative configurations of the TVGG motif remodels the NTD pocket (compare Fig. 4a and Fig. 4b). Furthermore, superpositions with published structures of CA in complex with FG-motif peptides indicate that the closed gate configuration of Met66 in the hexamer is compatible with binding (Fig. 4c), but the open gate configuration moves the Met66 sidechain into the phenylalanine pocket, causing a steric clash with the ligand’s phenyl ring (Fig. 4d). Thus, the CA pentamer should be impaired in binding FG motifs. To assess whether pentamers and hexamers have different FG-binding properties, we measured the thermostability of the CLPs in the presence or absence of a representative FG-motif ligand, a peptide derived from the host factor CPSF6 (refs. ^{28–30}) (Fig. 4e, Extended Data Fig. 7).

To ensure that all signals come from assembled CA, CLPs were first purified away from unassembled protein and aggregates by using size exclusion chromatography. Consistent with prediction, the M66A, G60A/G61P and G60A/G61P/M66A T=1 particles were not

stabilized by the FG peptide. In contrast, WT CLPs were stabilized by FG peptide but not by a negative control GG peptide in which the FG phenylalanine is replaced by glycine. These results indicate that CLP binding to FG motifs requires the CA hexamer and that the CA mutant pentamers lost binding, but do not rule out the formal possibility that WT CA pentamers can still bind FG.

Since bulk biochemical assays cannot distinguish between pentamers and hexamers in WT CLPs, we assayed for differential binding by using cryoEM. We incubated the WT CLPs at high concentration (0.27 mM CA) with 11-fold excess FG peptide to ensure saturation of available binding sites. Images of these complexes were collected for two independent samples, and then structures (6.0 and 6.3 Å resolution) of the declinations were solved, imposing no symmetry and using a large box size that encompassed one pentamer and 17 hexamers (Extended Data Fig. 8a,b). The hexamers serve as an internal positive control, and indeed bound peptides were clearly seen as U-shaped densities within all of the resolved hexameric NTD-CTD interfaces (Extended Data Fig. 9). To visualize the detailed interactions, we refined one data set to 3.9 Å nominal resolution (Fig. 4f–h, Extended Data Fig. 8b, Extended Data Table 1). The FG peptide is bound in the expected configuration to the hexamers (Fig. 4f), with the FG phenylalanine and CA Met66 sidechains clearly resolved by the map (Fig. 4g). We did not find peptide densities within the pentameric NTD-CTD interfaces (Fig. 4h, Extended Data Fig. 9). We therefore conclude that the CA pentamer does not bind to FG motifs.

Discussion

Here, we identify the conformational switch that modulates the ability of the HIV-1 CA protein to generate the hexameric and pentameric capsomers in the viral capsid. In other virus capsid proteins, amino and carboxyl termini are well-characterized switches that typically undergo order/disorder transitions^{9,10,40,41}. The CA termini also function as order/disorder switches that are triggered by programmed proteolysis^{42,43} but these modulate conversion of the immature capsid into the mature capsid and not hexamer/pentamer switching. The TVGG motif in CA is distinct from these known switches because it is an internal loop that refolds from one ordered configuration to another.

Although the exact trigger of the TVGG switch within virions remains to be established, our *in vitro* assembly data indicate that IP6 can perform this function. This is in line with other viral capsid switches, which are also sensitive to the assembly environment⁹. IP6 binds CA helix 1 at the central channels of the hexamer and pentamer, and so must trigger the TVGG switch through allostery, again a common theme in virus capsid assembly^{18,44}. In principle, allostery between IP6 and the TVGG switch can occur as the NTD-NTD and NTD-CTD interactions cooperatively form during assembly. Interestingly, however, our data indicate that allosteric communication may run through the protein hydrophobic core. Specifically, we find that: (1) refolding of the TVGG motif also remodels the core, (2) the helix 1/2 loop changes conformation, (3) this loop includes Phe32, whose sidechain is buried in the core, and (4) CA mutations that alter the core retain assembly but impair switching. These observations support the idea that the TVGG motif is allosterically connected to helix 1 by the hydrophobic core. A similar phenomenon has been documented in protein kinases, in

which core residues act as a non-covalent bridge between two elements (called “splines”) located on opposite sides of the protein fold⁴⁵. We speculate that CA’s helix 1/2 (NTD-NTD interface) and helix 3/4 (NTD-NTD and NTD-CTD interfaces) are analogous “splines” bridged by the residues that line its loosely packed hydrophobic core. Allosterity in kinases has been successfully exploited for pharmacological inhibition^{46,47}, and similar strategies may also be applicable to HIV-1 CA.

Another important insight from these studies is our discovery that the HIV-1 CA hexamer/pentamer switch modulates not just the assembly process, but also the capacity of the assembled capsid to bind FG- motif host cell factors that are important for the post-entry pathway of the virus. Our results suggest that the FG-motif proteins CPSF6 and NUP153, which are implicated in HIV-1 nuclear import and integration site selection⁴⁸, bind only to the hexamer. Other studies indicate that pentamer-specific factors might facilitate core trafficking in the cytoplasm or docking to the nucleus^{3,49}, pointing to division of labor between the two capsomers. Taken together, these observations support the requirement for an intact or nearly intact capsid at these viral replication steps.

The new ultra-potent HIV-1 inhibitor, lenacapavir, binds to the same site on the capsid as FG motifs. Indeed, a single point mutation in the methionine gate (M66I) confers resistance to this drug, although at severe cost to viral fitness^{34,35}. The molecular basis of this fitness cost can now be explained in light of our findings.

Methods

Protein purification and assembly

HIV-1 CA proteins were purified as described⁵⁰. Conical CLPs were assembled by incubating the protein (12-20 mg/mL) in 4-5 mM IP6, 50 mM MES, pH 6, 150 mM NaCl, 5 mM β -mercaptoethanol at 37 °C for 1-2 h. Tubular CLPs were assembled by incubating the protein in 50 mM Tris, pH 8, 1 M NaCl, 5 mM β -mercaptoethanol at 37 °C for 1-2 h. CA mutants were purified and assembled in the same way as the WT protein. Peptide-bound WT CLPs were prepared by diluting assembled CA (0.54 mM in 4 mM IP6) twofold into buffer containing 3 mM CPSF6-FG peptide (GTPVLFPGQPFGQPPLG, N-terminally acetylated and C-terminally amidated; Celltein), and incubating for 30 min in ice.

CryoEM structure determination

Conical CLPs (with or without bound FG peptide) were diluted 8 to 12-fold with 25 mM MES, pH 6, 100 μ M IP6 and then immediately applied (4 μ L) to glow-discharged lacey carbon 300-mesh Cu grids. T=1 particles were purified away from unassembled protein by using size exclusion chromatography (Superdex 200 10/300 GL; Cytiva) in 25 mM MES, pH 6, 50 mM NaCl, 10 mM β -mercaptoethanol, 100 μ M IP6, and 4 μ L of relevant fractions were applied to C-flat 1.2/1.3 300-mesh holey carbon grids. Grids were briefly blotted manually, and plunge-frozen in liquid ethane using a home-built device.

CryoEM data were collected at the University of Virginia Molecular Electron Microscopy Core. For the WT CLPs, G60A/G61P T=1 particles, and G60A/G61P/M66A particles, videos were collected using a Krios (ThermoFisher) operating at 300 kV and equipped

with an energy filter and K3 direct detector (Gatan). Data were collected using EPU (ThermoFisher) at a pixel size of 1.08 Å in counting mode, with a total dose of 50 electrons/Å² over 40 frames and target defocus of -0.5 to -2.5 μm. For M66A T=1 particles, videos were collected using a Glacios (ThermoFisher) operating at 200 kV and equipped with a Falcon 4 detector (ThermoFisher), at a pixel size of 1.5 Å in superresolution mode (effective pixel size of 0.75 Å), with a total dose of 48 electrons/Å². For CPSF6-bound CLPs, we collected two data sets with the Krios on two independent samples: Sample 1 at a pixel size of 1.62 Å and total dose of 54 electrons/Å², and Sample 2 at 1.08 Å and 50 electrons/Å².

All image processing and map calculations were performed in cryoSPARC v.3.3.1 (ref. ⁵¹) (Extended Data Table 1). Raw movies were corrected for beam-induced motion using MotionCor2 (ref. ⁵²) and CTF estimation was performed with CTFFIND4 (ref. ⁵³), as implemented in cryoSPARC. Initial particles were manually picked to generate references for subsequent template-based picking.

For the WT CA declination structure, 13,439,772 particles were initially extracted in box size 924 px (998 Å) (Extended Data Fig. 2a). Two rounds of reference-free 2D classification and selection were performed to identify and remove junk particles (mostly carbon edges) (Extended Data Fig. 2b). The selected 3,985,778 particles were then reextracted in box size 616 px (665 Å) and processed through several rounds of iterative reference-free 2D classification and selection (Extended Data Fig. 2c). Parameters were adjusted with the goal of having well-defined classes, each with around 500–5,000 particles, in the final round. This resulted in 1,345,054 particles that were used for ab initio structure calculation (5 models, C1 symmetry). Four models (1,075,923 particles) converged on similar maps and were combined into a single particle set through one round of refinement with C1 symmetry and low-pass filter of 20 Å. The resulting map was then rotated and centered on the pentamer. Heterogeneous refinement (3D classification, 2 models) was then performed in C1, with a low-pass filter of 20 Å. After discarding overlaps, the final set of 525,219 particles were reextracted in box size 308 px (333 Å) and refined with C5 symmetry and local CTF estimation, resulting in a map at a nominal resolution of 3.4 Å (Extended Data Fig. 2).

For the T=1 icosahedra, 503,025 (M66A), 1,513,851 (G60A/G61P), and 694,288 (G60A/G61P/M66A) particles were extracted (Extended Data Fig. 6). After two or three rounds of reference-free 2D classification, 116,627 (M66A), 494,057 (G60A/G61P), and 509,666 (G60A/G61P/M66A) particles were used for ab initio calculation with I symmetry imposed. Homogeneous refinement resulted in final maps at nominal resolutions of 3.1 Å (M66Å), 2.2 Å (G60A/G61P), and 2.4 Å (G61A/G61P/M66A) (Extended Data Fig. 4e–g).

For the low-resolution FG peptide-bound structures, the particle sets after junk removal and 2D classification were processed through one round of heterogeneous refinement (using as reference two copies of a WT map filtered to 20 Å) (Extended Data Fig. 8). For Sample 1, further alignment of 314,487 particles with C1 symmetry and box size 400 px (648 Å) afforded a map at ~6.3 Å resolution (Extended Data Fig. 8c). For Sample 2, alignment of 320,258 particles with C1 symmetry and box size 616 px (665 Å) afforded a map at ~6.0

Å resolution (Extended Data Fig. 8d). Sample 2 was then reprocessed through an additional heterogeneous refinement round to obtain a final set of 166,463 particles (box size 308 px, 333 Å), which were then refined with C5 symmetry to a nominal resolution of 3.9 Å (Extended Data Fig. 3e).

Model building and refinement

Coordinate models were built for the WT declination by docking PDB 4xfx [ref. ³¹] into the maps, followed by iterative manual rebuilding using Coot⁵⁴ and real space refinement using Phenix⁵⁵ (phenix.real_space_refine). Refinements were performed separately for two groups: the pentamer subunit (chain A) and the three closest three hexamer subunits to the pentamer (chains L, M and N) (Extended Data Table 1). Non-crystallographic symmetry restraints were used when appropriate. The remaining three hexamer subunits (chains O, P and Q) were modeled by rigid-body docking of the final refined model of chain M as polyalanine.

Modeling of the WT declination in complex with FG peptide was performed by rigid body docking of chains A, L, M and N of the final refined WT model into the 3.9 Å map. The peptide was initially obtained by docking PDB 4wym chain M [ref. ³⁰] into the appropriate densities. Real space refinement was performed as described above.

For the T=1 particles, one copy of PDB 4xfx was docked into the G60A/G61P/M66A map, followed by iterative manual rebuilding and real space refinement. One round of morphing was performed in the first refinement iteration. To obtain M66A and G60A/G61P models, the final refined model of the triple mutant was fitted through one round of rigid body refinement. The mutated regions were rebuilt and refit manually, and then the models were subjected to two or three rounds of real space refinement.

In all maps, modeling of IP6 densities was restricted to rigid-body fitting of the *myo* isoform as described previously¹².

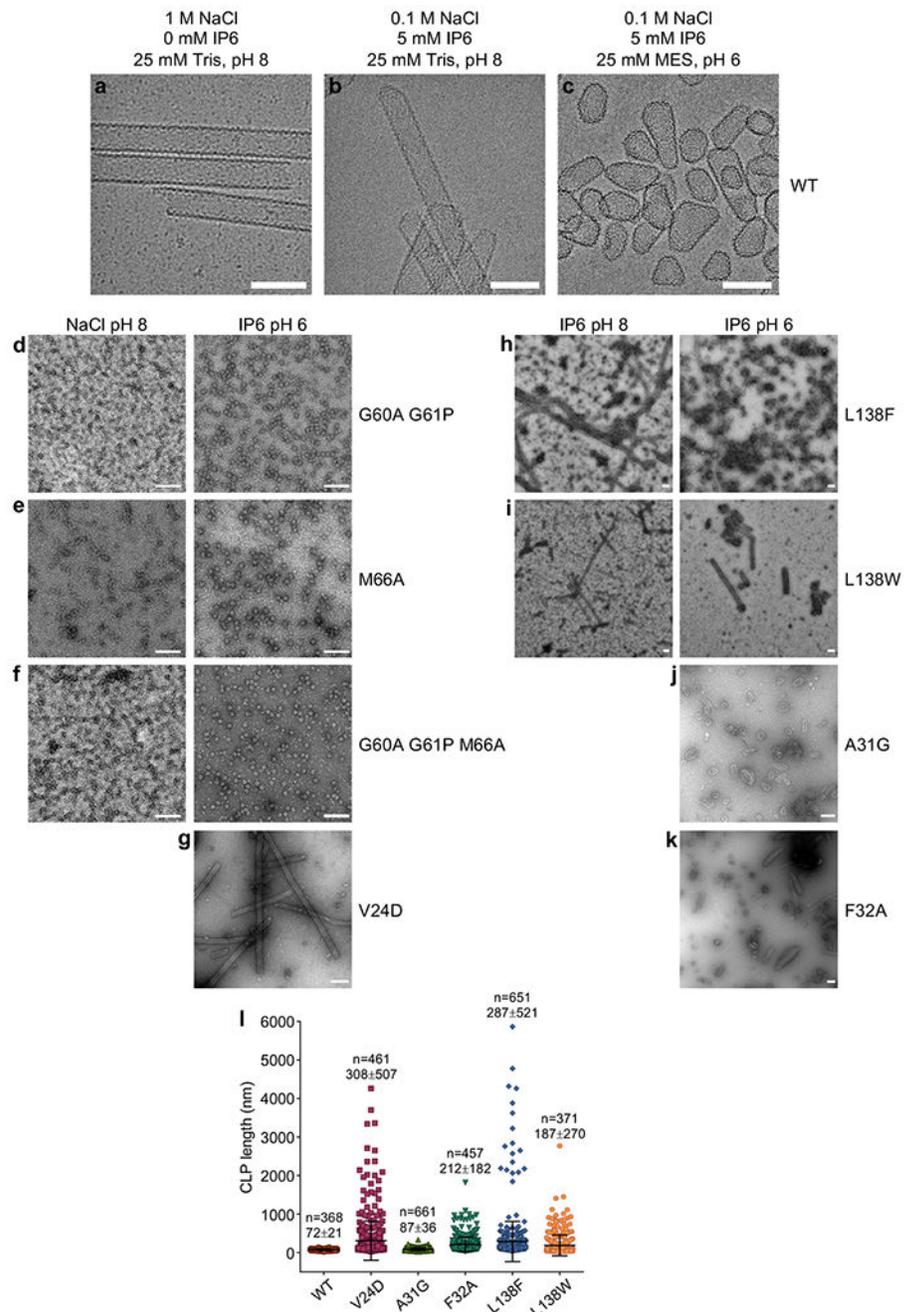
Negative stain EM

Samples (4 µL) were applied to Formvar continuous carbon 300-mesh Cu grids and incubated for 2 min. Grids were moved to a 20µL drop of 0.1 M KCl, incubated for 2 min, and blotted. Grids were then placed on a 20µL drop of 2% uranyl acetate, incubated for 2 mins, and blotted to dryness. Images were collected on a Tecnai Spirit (ThermoFisher) or an F20 microscope (ThermoFisher), both operating at 120 kV.

Deposition

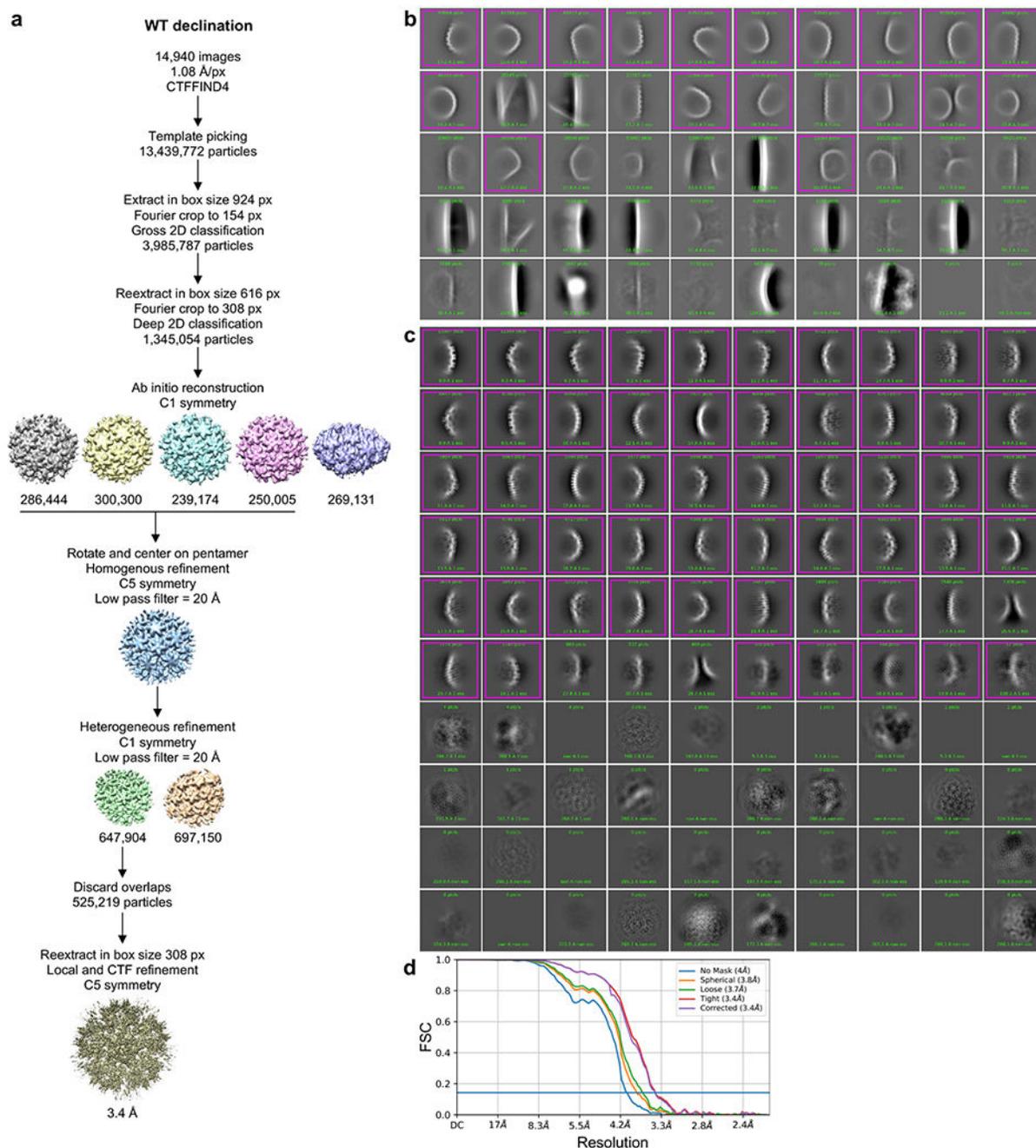
CryoEM maps are deposited at the Electron Microscopy Data Bank under accession numbers EMD- 26715 (WT declination), EMD-28054 (T=1 G60A/G61P), EMD-28057 (T=1 M66A), EMD-26718 (T=1 G60A/G61P/M66A) and EMD-28186 (WT declination bound to CPSF6-FG peptide). Coordinates are deposited at the Protein Data Bank under accession numbers 7urn (WT declination), 8ejl (WT-FG complex), 8eep (G61A/G60P), 8eet (M66A) and 7urt (G60A/G61P/M66A).

Extended Data

**Extended Data Figure 1.**

(a-c) Cryoimages WT HIV-1 CA CLPs assembled under the indicated conditions, representative of at least two independent experiments with different protein preparations. Reactions were incubated at 37 °C for 1 h. Scale bars, 100 nm. (d-k) Negative stain images of CLPs formed under the indicated conditions, representative of 2 to 3 independent experiments with different protein preparations. Scale bars, 200 nm. (l) The longest

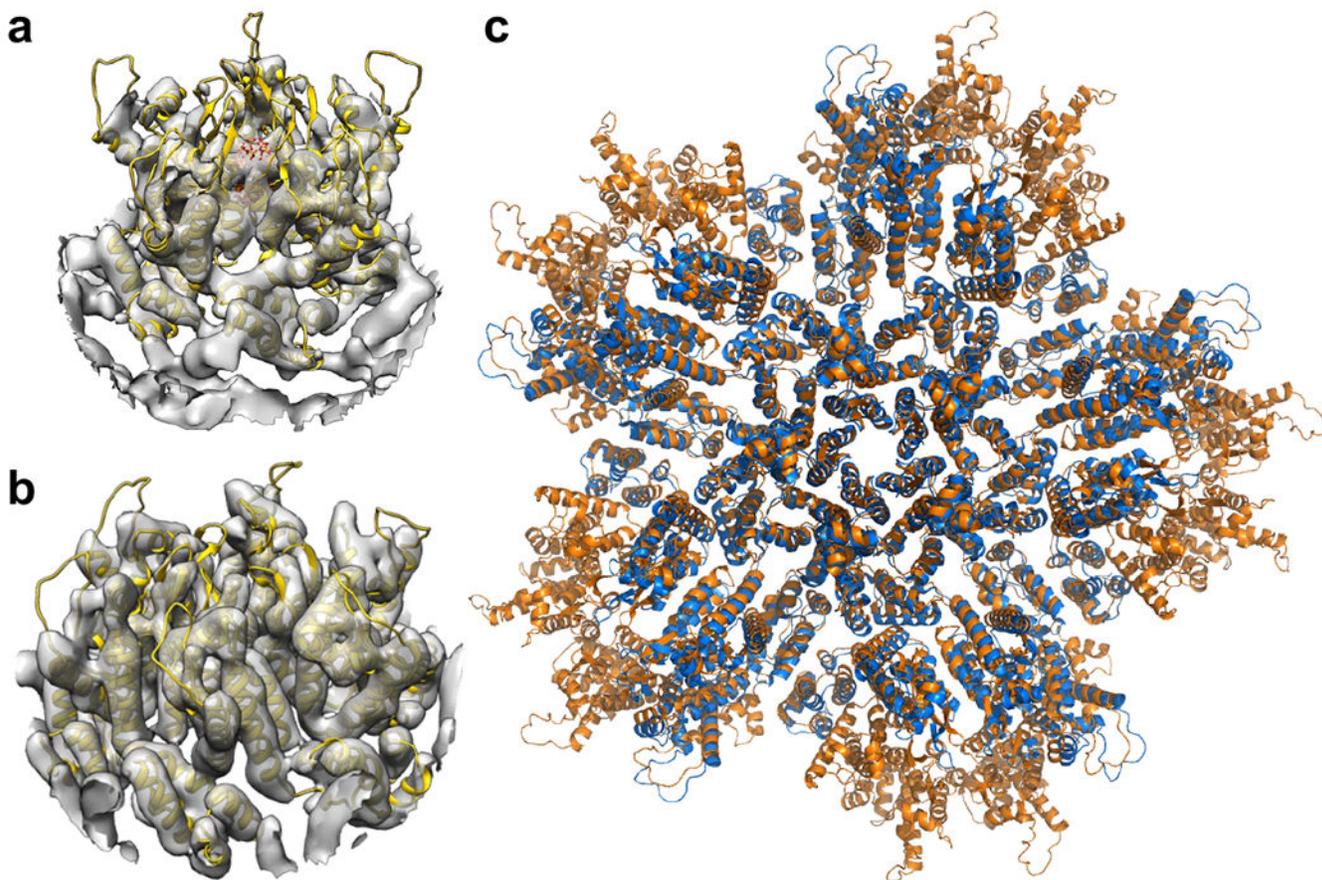
dimension of individual CLPs were quantified for each of the indicated CA assemblies in 5 mM IP6, pH 6. According to one study, the average length of native fullerene cones is around 109 nm [ref. ⁵⁶].



Extended Data Figure 2.

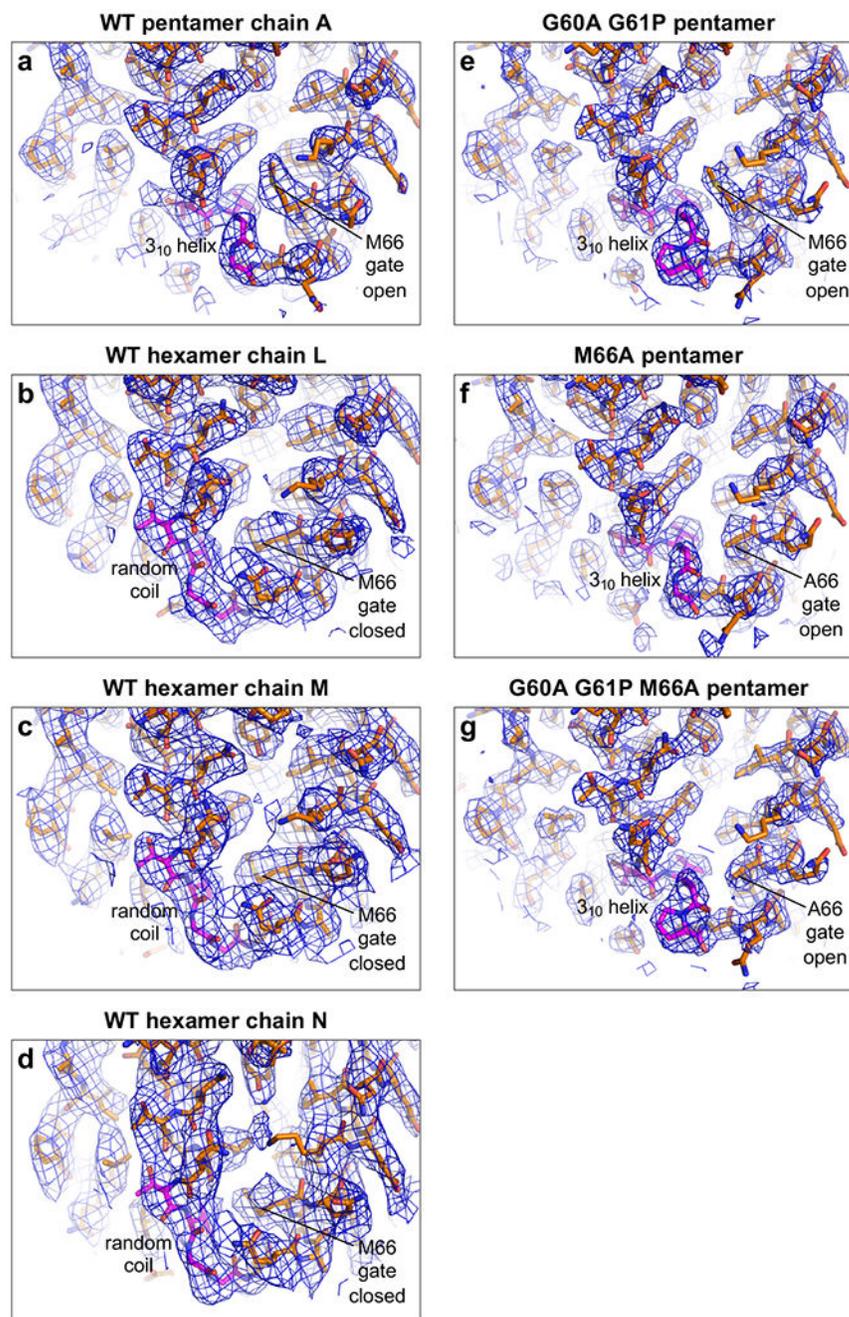
(a) Single particle averaging workflow for the WT declination. After the template picking step, particles were randomly split into 4 subsets to facilitate computations up to the ab initio reconstruction step. (b) Representative result of initial 2D classification of particles extracted

in box size 924 px (998 Å). Magenta indicates classes selected for further processing. (c) Representative result of 2D classification of particles reextracted in box size 616 px (665 Å). (d) Fourier shell correlation (FSC) plot from the final refinement round.

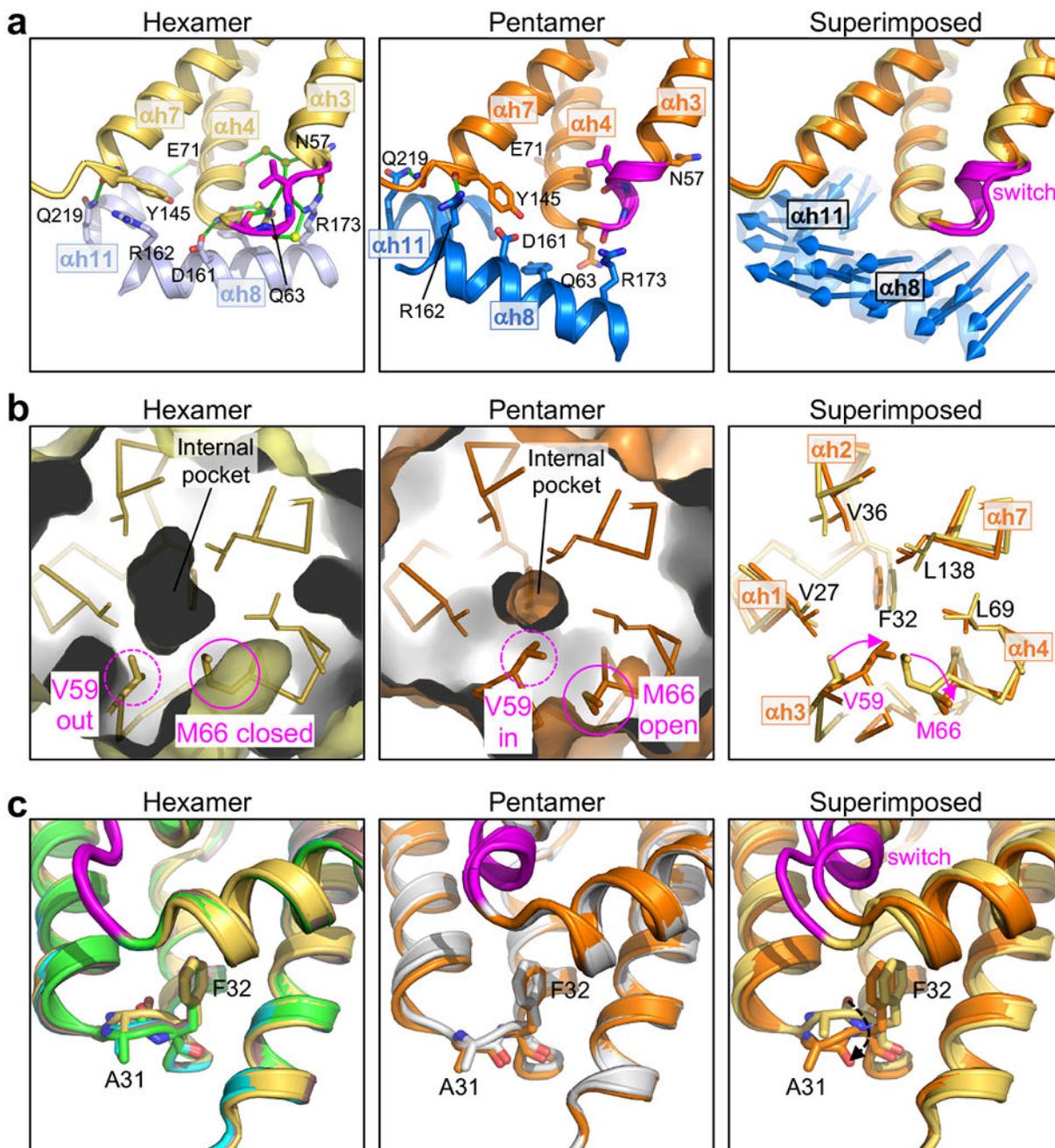


Extended Data Figure 3.

Structures of the capsomers in IP6-induced *in vitro* CLPs are similar to their corresponding structures *in situ*^{7,16}. (a) Docking of our real-space-refined CA pentamer model (yellow ribbons, with IP6 in sticks) into EMD 3466 (gray), the first reported *in situ* structure of the HIV-1 CA pentamer at 8.8 Å resolution. (b) Docking of our real-space-refined CA hexamer model (yellow ribbons) into EMD 3465 (gray), the *in situ* structure of the HIV-1 CA hexamer at 6.8 Å resolution. (c) Superposition of the final PDB model of the declination after real space refinement into our 3.4 Å map (orange) with PDB 5mcy (blue).

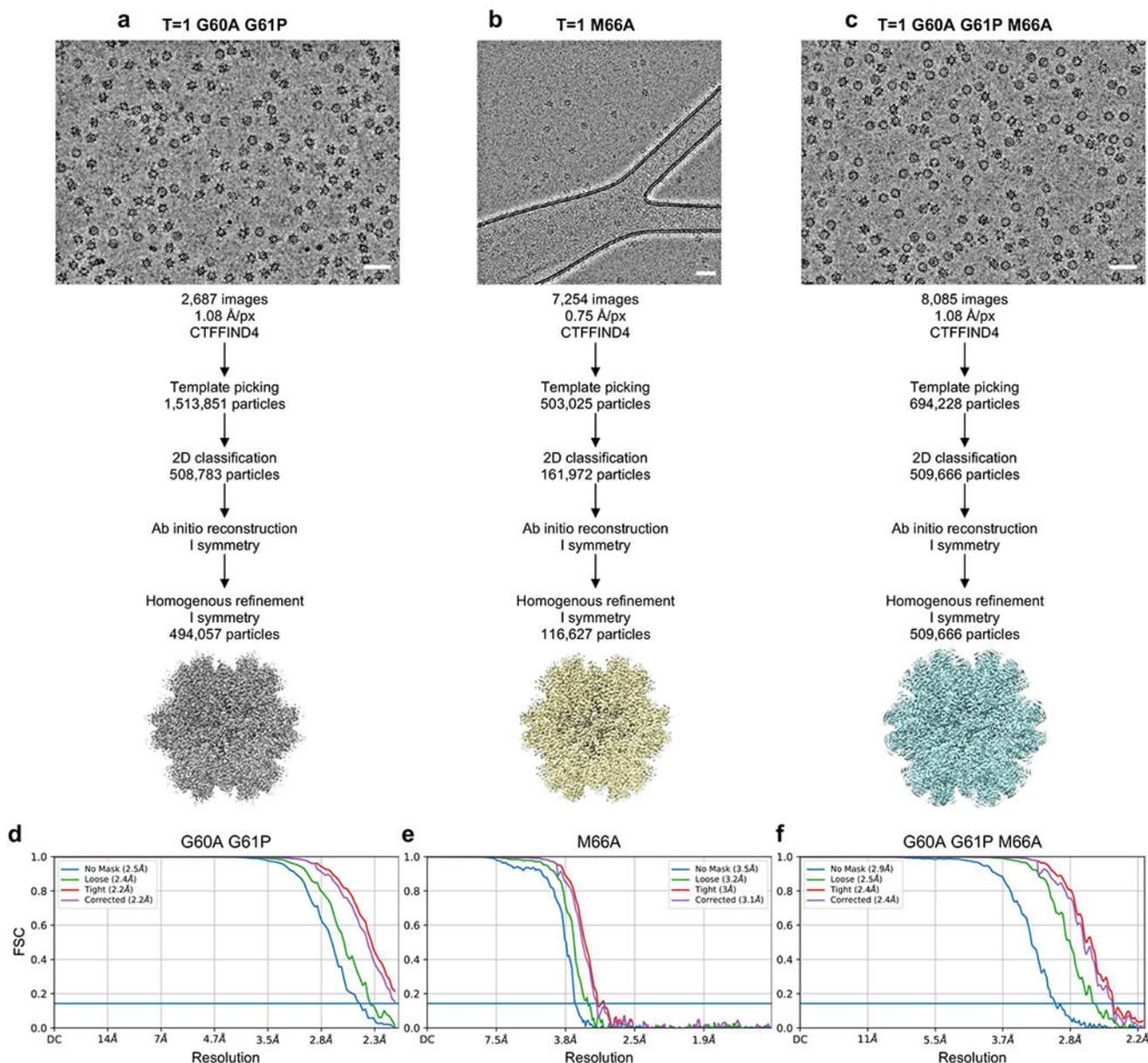
**Extended Data Figure 4.**

CryoEM maps (blue mesh) with real-space-refined models of asymmetric units (sticks). The TVGG motifs (and mutated versions) are in magenta. **(a)** WT pentamer. **(b-d)** WT hexamers. **(e)** G60A/G61P pentamer. **(f)** M66A pentamer. **(g)** G60A/G61P/M66A pentamer.

**Extended Data Figure 5.**

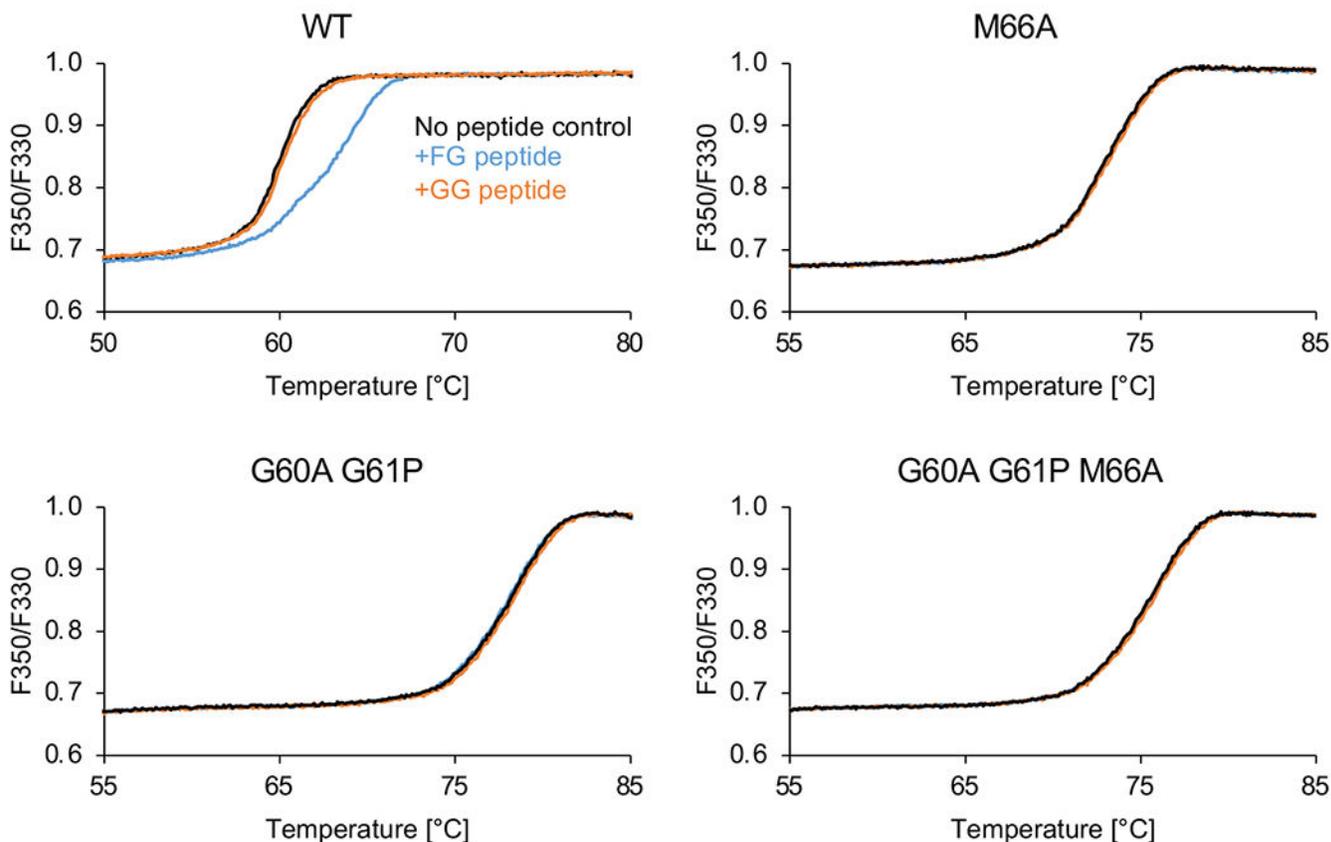
Additional distinguishing structural features of the HIV-1 CA hexamer and pentamer. (a) The NTD-CTD interfaces are distinguished by different sets of hydrogen bonds. Left: The hexameric interface is stabilized by an extensive network of hydrogen bonds involving the TVGG motif, surrounding residues and ordered water molecules. Shown is PDB 3h47, a high-resolution (1.9 Å) X-ray crystal structure¹⁹. Middle: In the pentamer, inter-subunit hydrogen bonds involving the TVGG motif are absent. Arg162 and Gln219 make hydrogen bonds with the C-terminal end of helix 7 and the linker, respectively. Right: Blue arrows

indicate the displacement vectors for each residue in helices 8 and 11. **(b)** The NTD hydrophobic core is underpacked. The hexamer pocket (Left) is larger than the pentamer pocket (Middle). Right: Residues lining the pocket are shown as sticks and labeled. **(c)** Rotation of the Ala31-Phe32 peptide bond. Left: Hexamer structures, with PDB 3h47 in yellow. The three best-defined hexamer subunits in our WT CA declination structure are shown (chain L, green; chain M, cyan; chain N, brown). Middle: Pentamer structures of WT CA (orange, 3.4 Å) and G60A/G61P/M66A (gray, 2.4 Å). Right: PDB 3h47 is superimposed with the WT pentamer. Dashed arrow indicates rotation of the peptide bond.



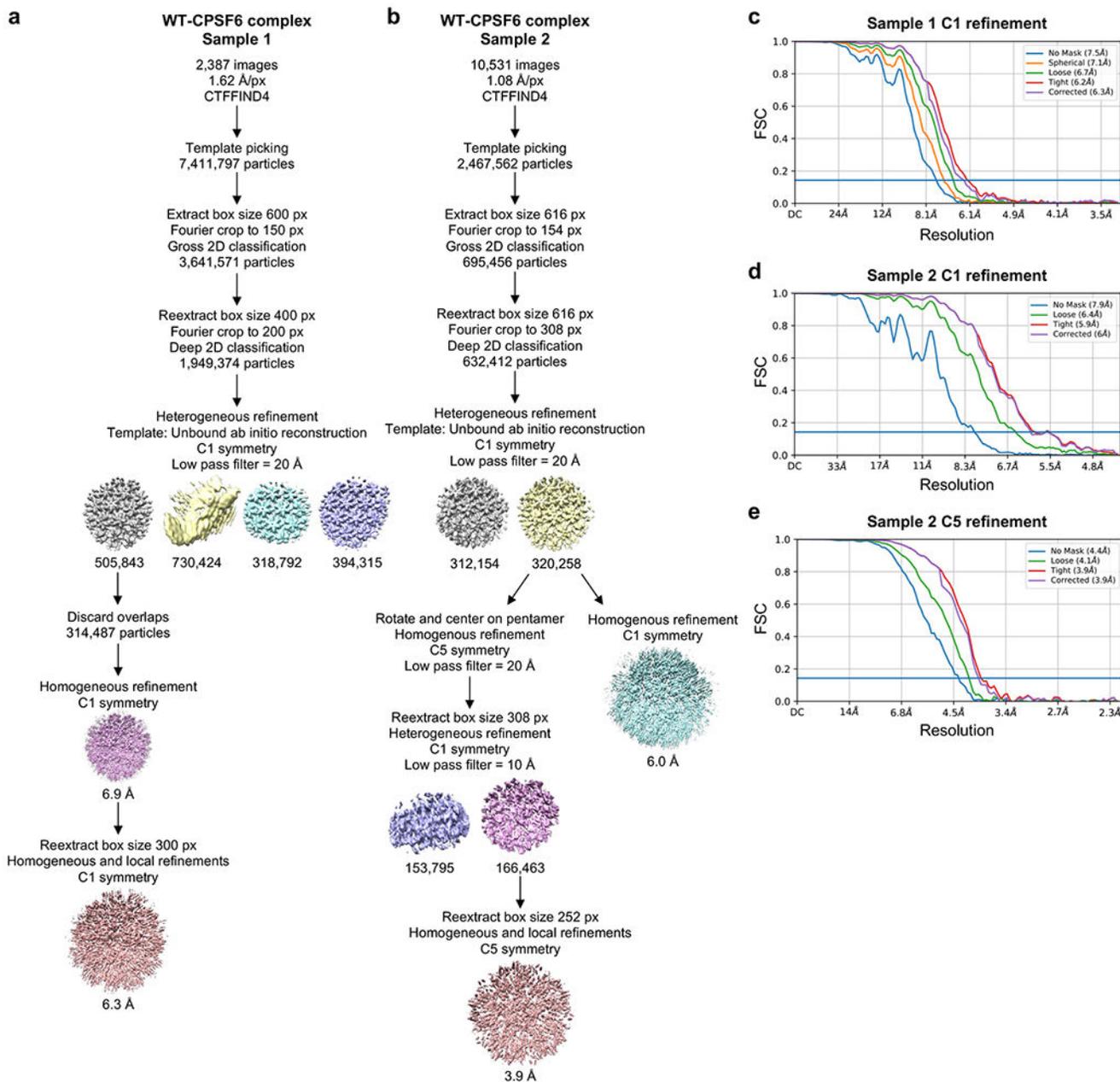
Extended Data Figure 6.

(a-c) Single particle averaging workflow for the indicated T=1 particles. (d-f) Fourier shell correlation (FSC) plots from the final refinement rounds.

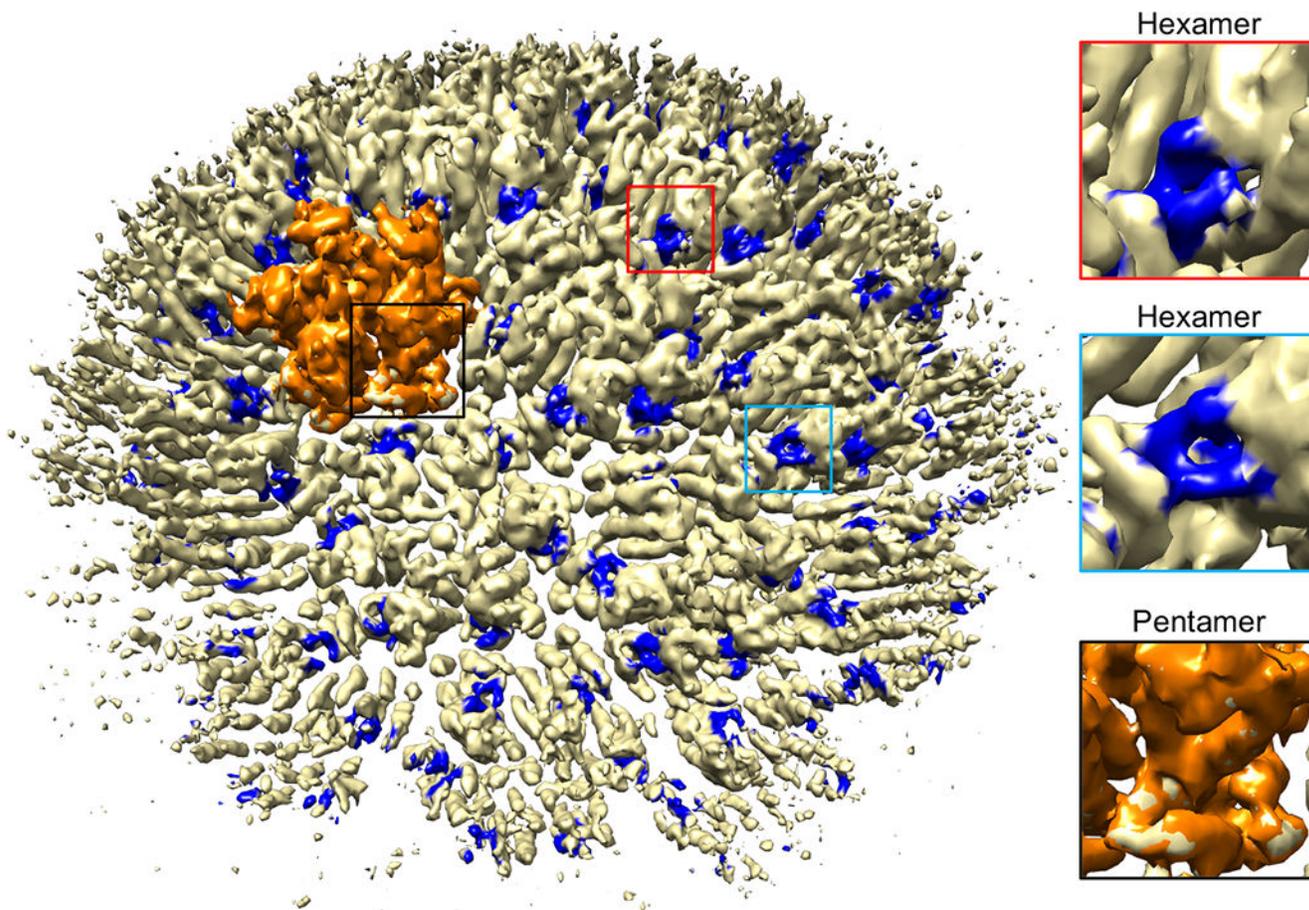


Extended Data Figure 7.

Representative raw thermal melting profiles. Indicated CLPs (12–22 μM) were heated in the absence of peptide (black curves), in the presence of 500 μM FG peptide (blue curves) or in the presence of 500 μM GG peptide (orange curves). The FG peptide induced a clear shift in the melting profile of WT CLPs (upper left panel), thus indicating binding. In all other cases, the melting curves are virtually superimposable with their respective no-peptide controls, indicating lack of binding.

**Extended Data Figure 8.**

(a-b) Single particle averaging workflow for the WT declinations in complex with CPSF6-FG peptide. (c-e) Fourier shell correlation (FSC) plots.

**Extended Data Figure 9.**

CryoEM structure (Sample 2, 6 Å resolution) of the WT declination in complex with CPSF6-FG peptide (0.4 contour). The pentamer is in orange. Peptide densities are in blue.

Extended Data Table 1.

Data collection, processing and refinement statistics

	WT declination	WT declination + CPSF6-FG	T=1 G60A/G61P	T=1 M66A	T=1 G60A/G61P/M66A
EMDB	EMD-26715	EMD-28186	EMD-28054	EMD-28057	EMD-26718
PDB	7urn	8ejl	8eep	8eet	7urt
Data collection, processing and map calculation					
Voltage (kV)	300	300	300	200	300
Electron exposure (e-/Å ²)	50	54	50	48	50
Defocus range (μm)	0.5 to 2.5	0.75 to 2.5	0.5 to 2.5	0.5 to 2.5	0.5 to 2.5
Pixel size (Å)	1.08	1.08	1.08	0.75	1.08
Symmetry imposed	C5	C5	I	I	I

	WT declination	WT declination + CPSF6-FG	T=1 G60A/G61P	T=1 M66A	T=1 G60A/G61P/M66A	
Particles	525,219	166,463	494,057	116,627	509,666	
Map Resolution (Å)	3.4	3.9	2.2	3.1	2.4	
FSC threshold	0.143	0.143	0.143	0.143	0.143	
Coordinate modeling and refinement						
	Chain A (pentamer) PDB 4xfx	Chains L,M,N (half-hexamer) PDB 4xfx	PDB 7urn	PDB 7urt	PDB 7urt	PDB 4xfx
Initial model						
Model resolution (Å)	3.4	3.6	4.1	2.5	3.4	2.5
FSC threshold	0.5	0.5	0.5	0.5	0.5	0.5
Map sharpening B-factor (Å ²)	146.9	146.9	159.3	97.0	144.4	109.7
Model composition						
Non-hydrogen atoms	1,725	5,175	4,431	1,726	1,726	1,726
Protein residues	221	663	619	221	221	221
IP6	2	0	0	2	2	2
B-factors (Å ²)	73.2	78.8	117.0	42.9	42.6	
Root mean squared deviations						
Bond lengths (Å)	0.006	0.004	0.005	0.002	0.004	0.004
Bond angles (°)	0.623	0.508	0.545	0.647	1.019	0.554
Validation						
Molprobrity score	1.57	1.42	1.81	1.33	1.37	1.34
Clash score	8.40	5.79	7.23	3.95	6.23	4.05
Poor rotamers (%)	0.43	1.24	3.04	1.59	1.07	1.60
Ramachandran plot						
Favored (%)	97.26	97.87	97.82	100	99.54	99.09
Outliers (%)	0	0	0	0	0	0
Z-score	1.00	1.79	1.74	1.98	1.87	1.22

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Transmission electron micrographs were recorded at the University of Virginia Molecular Electron Microscopy Core facility (RRID: SCR_019031), which was built in part with NIH grant G20-RR31199. Purchase of the Krios was funded in part by S10-RR025067. This study was funded by R21-AI167756, P50-AI150464 and U54-AI170856 (O.P. and B.K.G.-P.). R.T.S. was supported by T32-GM008136. We especially thank K. Dryden and M. Purdy for invaluable assistance in EM data collection and processing, and M. Purdy for help with optimizing server and software set up.

References

1. Hilditch L & Towers GJ A model for cofactor use during HIV-1 reverse transcription and nuclear entry. *Curr Opin Virol* 4, 32–36 (2014). [PubMed: 24525292]
2. Christensen DE, Ganser-Pornillos BK, Johnson JS, Pornillos O & Sundquist WI Reconstitution and visualization of HIV-1 capsid-dependent replication and integration in vitro. *Science* 370, abc8420 (2020).
3. Zila V, Müller TG, Müller B & Kräusslich HG HIV-1 capsid is the key orchestrator of early viral replication. *PLoS Pathog* 17, e1010109 (2021). [PubMed: 34968390]
4. Naghavi MH HIV-1 capsid exploitation of the host microtubule cytoskeleton during early infection. *Retrovirology* 18, 19 (2021). [PubMed: 34229718]
5. Ganser BK, Li S, Klishko VY, Finch JT & Sundquist WI Assembly and analysis of conical models for the HIV-1 core. *Science* 283, 80–83 (1999). [PubMed: 9872746]
6. Li S, Hill CP, Sundquist WI & Finch JT Image reconstructions of helical assemblies of the HIV-1 CA protein. *Nature* 407, 409–413 (2000). [PubMed: 11014200]
7. Mattei S, Glass B, Hagen WJ, Kräusslich HG & Briggs JA The structure and flexibility of conical HIV-1 capsids determined within intact virions. *Science* 354, 1434–1437 (2016). [PubMed: 27980210]
8. Pornillos O, Ganser-Pornillos BK & Yeager M Atomic-level modelling of the HIV capsid. *Nature* 469, 424–427 (2011). [PubMed: 21248851]
9. Johnson JE & Speir JA Quasi-equivalent viruses: a paradigm for protein assemblies. *J Mol Biol* 269, 665–675 (1997). [PubMed: 9223631]
10. Dokland T Freedom and restraint: themes in virus capsid assembly. *Structure* 8, R157–162 (2000). [PubMed: 10997898]
11. Ehrlich LS, Agresta BE & Carter CA Assembly of recombinant human immunodeficiency virus type 1 capsid protein in vitro. *J Virol* 66, 4874–4883 (1992). [PubMed: 1629958]
12. Dick RA et al. Inositol phosphates are assembly co-factors for HIV-1. *Nature* 560, 509–512 (2018). [PubMed: 30069050]
13. Mallery DL et al. IP6 is an HIV pocket factor that prevents capsid collapse and promotes DNA synthesis. *eLife* 7, e35335 (2018). [PubMed: 29848441]
14. Yao X, Fan X & Yan N Cryo-EM analysis of a membrane protein embedded in the liposome. *Proc Natl Acad Sci U S A* 117, 18497–18503 (2020). [PubMed: 32680969]
15. Briggs JA et al. Structure and assembly of immature HIV. *Proc Natl Acad Sci U S A* 106, 11090–11095 (2009). [PubMed: 19549863]
16. Ni T et al. Structure of native HIV-1 cores and their interactions with IP6 and CypA. *Sci Adv* 7, eabj5715 (2021). [PubMed: 34797722]
17. Renner N et al. A lysine ring in HIV capsid pores coordinates IP6 to drive mature capsid assembly. *PLoS Pathog* 17, e1009164 (2021). [PubMed: 33524070]
18. Domitrovic T et al. Virus assembly and maturation: auto-regulation through allosteric molecular switches. *J Mol Biol* 425, 1488–1496 (2013). [PubMed: 23485419]
19. Pornillos O et al. X-ray structures of the hexameric building block of the HIV capsid. *Cell* 137, 1282–1292 (2009). [PubMed: 19523676]
20. Lemke CT et al. Distinct effects of two HIV-1 capsid assembly inhibitor families that bind the same site within the N-terminal domain of the viral CA protein. *J Virol* 86, 6643–6655 (2012). [PubMed: 22496222]
21. Manochewa S, Swain JV, Lanxon-Cookson E, Rolland M & Mullins JI Fitness costs of mutations at the HIV-1 capsid hexamerization interface. *PLoS One* 8, e66065 (2013). [PubMed: 23785468]
22. Rihn SJ et al. Extreme genetic fragility of the HIV-1 capsid. *PLoS Pathog* 9, e1003461 (2013). [PubMed: 23818857]
23. Al-Mawsawi LQ et al. High-throughput profiling of point mutations across the HIV-1 genome. *Retrovirology* 11, 124 (2014). [PubMed: 25522661]

24. Ganser-Pornillos BK, von Schwedler UK, Stray KM, Aiken C & Sundquist WI Assembly properties of the human immunodeficiency virus type 1 CA protein. *J Virol* 78, 2545–2552 (2004). [PubMed: 14963157]
25. Forshey BM, von Schwedler U, Sundquist WI & Aiken C Formation of a human immunodeficiency virus type 1 core of optimal stability is crucial for viral replication. *J Virol* 76, 5667–5677 (2002). [PubMed: 11991995]
26. Eschbach JE et al. Capsid lattice destabilization leads to premature loss of the viral genome and integrase enzyme during HIV-1 infection. *J Virol* 95, 00984–20 (2020).
27. Blair WS et al. HIV capsid is a tractable target for small molecule therapeutic intervention. *PLoS Pathog* 6, e1001220 (2010). [PubMed: 21170360]
28. Price AJ et al. CPSF6 defines a conserved capsid interface that modulates HIV-1 replication. *PLoS Pathog* 8, e1002896 (2012). [PubMed: 22956906]
29. Price AJ et al. Host cofactors and pharmacologic ligands share an essential interface in HIV-1 capsid that is lost upon disassembly. *PLoS Pathog* 10, e1004459 (2014). [PubMed: 25356722]
30. Bhattacharya A et al. Structural basis of HIV-1 capsid recognition by PF74 and CPSF6. *Proc Natl AcadSci USA* 111, 18625–18630 (2014).
31. Gres AT et al. STRUCTURAL VIROLOGY. X-ray crystal structures of native HIV-1 capsid protein reveal conformational variability. *Science* 349, 99–103 (2015). [PubMed: 26044298]
32. Buffone C et al. Nup153 unlocks the nuclear pore complex for HIV-1 nuclear translocation in nondividing cells. *J Virol* 92, 00648–18 (2018).
33. Yant SR et al. A highly potent long-acting small-molecule HIV-1 capsid inhibitor with efficacy in a humanized mouse model. *Nat Med* 25, 1377–1384 (2019). [PubMed: 31501601]
34. Link JO et al. Clinical targeting of HIV capsid protein with a long-acting small molecule. *Nature* 584, 614–618 (2020). [PubMed: 32612233]
35. Bester SM et al. Structural and mechanistic bases for a potent HIV-1 capsid inhibitor. *Science* 370, 360–364 (2020). [PubMed: 33060363]
36. Rebenburg SV et al. Sec24C is an HIV-1 host dependency factor crucial for virus replication. *Nat Microbiol* 6, 435–444 (2021). [PubMed: 33649557]
37. Kelly BN et al. Structure of the antiviral assembly inhibitor CAP-1 complex with the HIV-1 CA protein. *J Mol Biol* 373, 355–366 (2007). [PubMed: 17826792]
38. Lee K et al. HIV-1 capsid-targeting domain of cleavage and polyadenylation specificity factor 6. *J Virol* 86, 3851–3860 (2012). [PubMed: 22301135]
39. Matreyek KA, Yucel SS, Li X & Engelman A Nucleoporin NUP153 phenylalanine-glycine motifs engage a common binding pocket within the HIV-1 capsid protein to mediate lentiviral infectivity. *PLoS Pathog* 9, e1003693 (2013). [PubMed: 24130490]
40. Barcena J et al. The coat protein of Rabbit hemorrhagic disease virus contains a molecular switch at the N-terminal region facing the inner surface of the capsid. *Virology* 322, 118–134 (2004). [PubMed: 15063122]
41. Luque D et al. Infectious bursal disease virus capsid assembly and maturation by structural rearrangements of a transient molecular switch. *J Virol* 81, 6869–6878 (2007). [PubMed: 17442720]
42. von Schwedler UK et al. Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly. *EMBO J* 17, 1555–1568 (1998). [PubMed: 9501077]
43. Gross I et al. A conformational switch controlling HIV-1 morphogenesis. *EMBO J* 19, 103–113 (2000). [PubMed: 10619849]
44. Zlotnick A & Mukhopadhyay S Virus assembly, allostery and antivirals. *Trends Microbiol* 19, 14–23 (2011). [PubMed: 21163649]
45. Kim J et al. A dynamic hydrophobic core orchestrates allostery in protein kinases. *Sci Adv* 3, e1600663 (2017). [PubMed: 28435869]
46. Zhang J, Yang PL & Gray NS Targeting cancer with small molecule kinase inhibitors. *Nat Rev Cancer* 9, 28–39 (2009). [PubMed: 19104514]
47. Dar AC & Shokat KM The evolution of protein kinase inhibitors from antagonists to agonists of cellular signaling. *Annu Rev Biochem* 80, 769–795 (2011). [PubMed: 21548788]

48. Engelman AN HIV Capsid and Integration Targeting. *Viruses* 13, 125 (2021). [PubMed: 33477441]
49. Santos da Silva E et al. HIV-1 capsids mimic a microtubule regulator to coordinate early stages of infection. *EMBO J* 39, e104870 (2020). [PubMed: 32896909]
50. Yoo S et al. Molecular recognition in the HIV-1 capsid/cyclophilin A complex. *J Mol Biol* 269, 780–795 (1997). [PubMed: 9223641]
51. Punjani A, Rubinstein JL, Fleet DJ & Brubaker MA cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* 14, 290–296 (2017). [PubMed: 28165473]
52. Zheng SQ et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* 14, 331–332 (2017). [PubMed: 28250466]
53. Rohou A & Grigorieff N CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* 192, 216–221 (2015). [PubMed: 26278980]
54. Emsley P, Lohkamp B, Scott WG & Cowtan K Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66, 486–501 (2010). [PubMed: 20383002]
55. Liebschner D et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* 75, 861–877 (2019). [PubMed: 31588918]
56. Briggs JA et al. The mechanism of HIV-1 core assembly: insights from three-dimensional reconstructions of authentic virions. *Structure* 14, 15–20 (2006). [PubMed: 16407061]

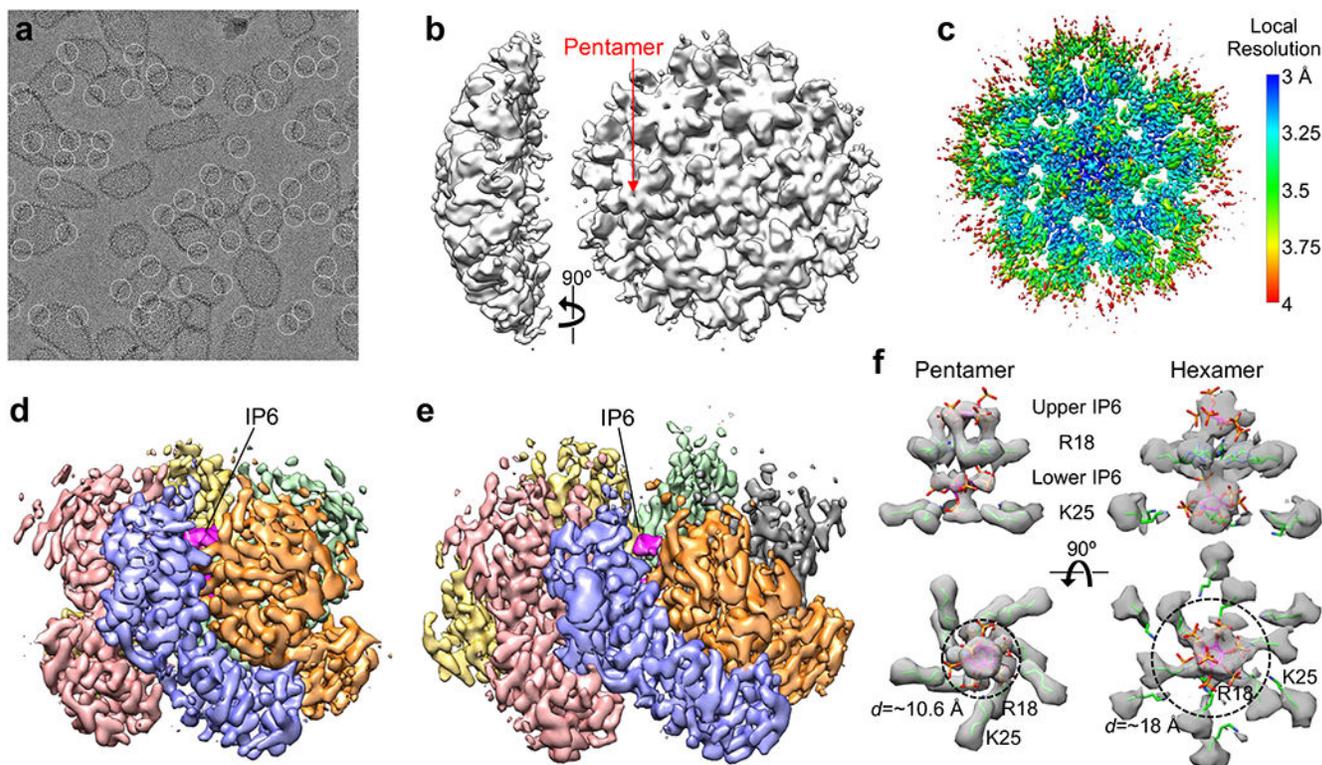


Figure 1.

Assembly and structural analysis of IP6-induced HIV-1 CA capsid-like particles (CLPs). **(a)** In vitro assembled CLPs, with declinations encircled. **(b)** Initial ab initio map calculated from 670,480 particles. **(c)** Map after further local refinement with imposed five-fold symmetry, colored according to local resolution. At this contour level (0.4), the five central pentamer subunits and their fifteen closest hexameric subunits are very well defined. **(d)** Structure of the pentamer (0.28 contour). Each subunit is in a different color, and IP6 is colored in magenta. **(e)** Structure of the hexamer (0.22 contour). **(f)** Side views (upper) and bottom views (lower) of the central channels of the capsomers. Maps (gray) are contoured at 0.4 (pentamer) and 0.21 (hexamer). Green sticks indicate the Arg18 and Lys25 sidechains. IP6 molecules were rigid-body docked as the *myo* isoform. Dashed circles with indicated diameters (d) circumscribe the positions of the Lys25 epsilon amines.

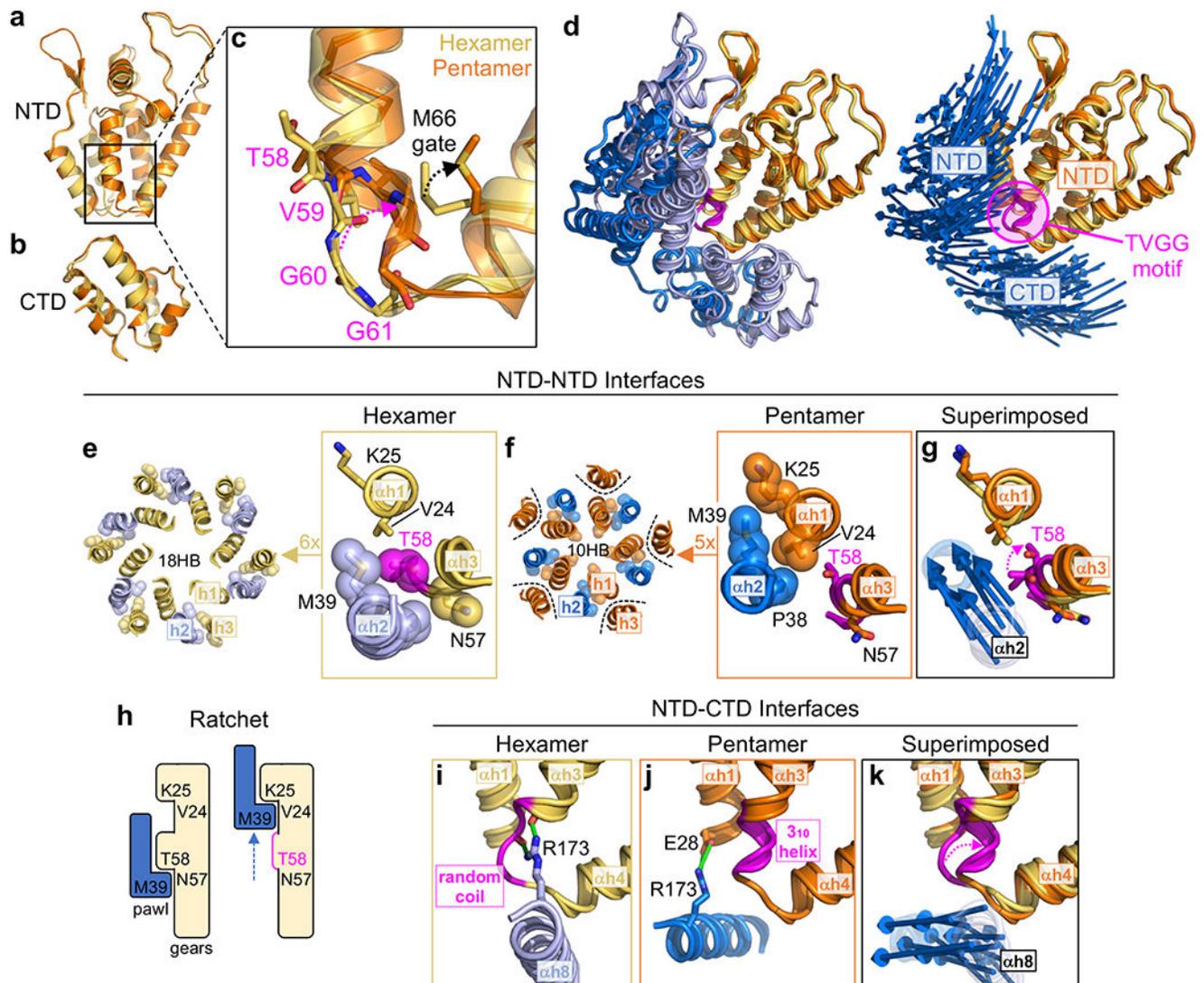


Figure 2. Comparison of HIV-1 CA pentamer and hexamer structures identify a molecular switch. **(a-b)** The NTD and CTD from hexameric (yellow) and pentameric (orange) subunits are superimposed. **(c)** Distinct secondary structures of the Thr58-Val59-Gly60-Gly61 (TVGG) motif in the hexamer and pentamer. **(d)** The TVGG motif (magenta) is at the juncture of the NTD-NTD and NTD-CTD interfaces that hold together both capsomers. Reference pentamer and hexamer NTDs are superimposed. Right: Displacement vectors (blue arrows) indicate positional changes of the NTD and CTD of the neighboring subunit. **(e-g)** Details of the NTD-NTD interfaces in the hexamer (yellow and light blue) and pentamer (orange and blue). Dashed magenta arrow in **g** signifies overwinding of Thr58. Displacement vectors (blue arrows) signify “slippage” of helix 2 from the hexamer state into the pentamer. **(h)** Linear ratchet mechanism of hexamer/pentamer switching at the NTD-NTD interface. **(i-k)** Details of the NTD-CTD interfaces in the hexamer (yellow and light blue) and pentamer (orange and blue). Green indicates hydrogen bonds. Dashed magenta arrow in **k** signifies

refolding of the TVGG motif. Displacement vectors (blue arrows) indicate movement of helix 8.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

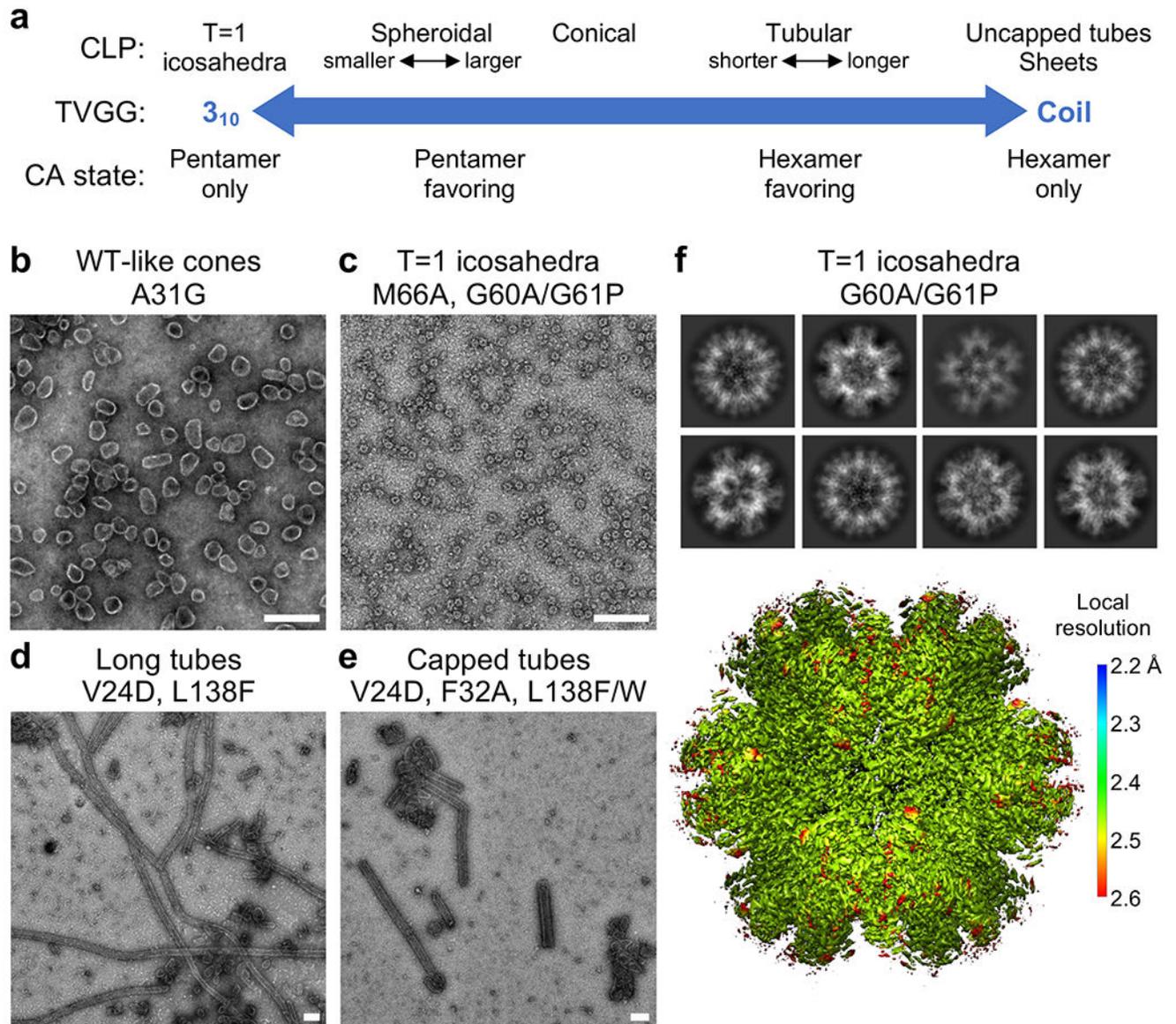


Figure 3. Designed manipulations favor or disfavor pentamer formation during HIV-1 CA assembly. **(a)** Predicted phenotypes of CA assembly behavior in vitro and CLP shape, as determined by the configuration of the TVGG motif. All but T=1 icosahedra have been observed previously. **(b-e)** Morphologies of CLPs formed by the indicated CA constructs in 25 mM MES, pH 6, 150 mM NaCl, 5 mM β -mercaptoethanol, 5 mM IP6. Scale bars, 200 nm. **(f)** 2D class averages (upper) and cryoEM structure (lower) of T=1 icosahedra made by CA G60A/G61P. The map is colored according to local resolution.

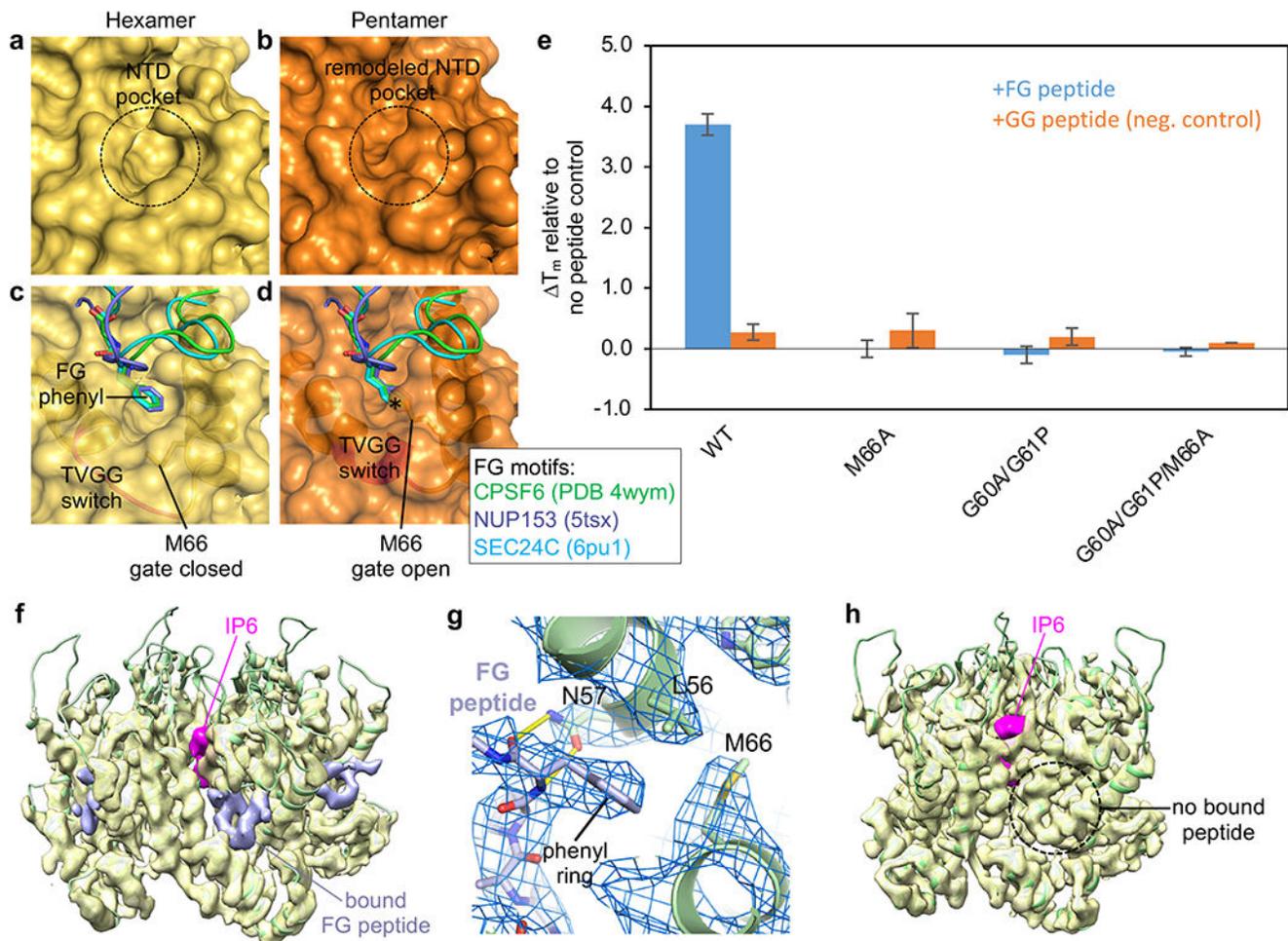


Figure 4. Conformational switching of the TVGG motif remodels the FG-motif binding site. **(a,b)** Connolly surface representations of our WT hexamer and pentamer structures. **(c,d)** The hexamer and pentamer are superimposed with the indicated crystal structures of CA with bound FG ligands. * indicates steric clash. **(e)** Thermostabilization of indicated CLPs by CPSF6-FG and CPSF6-GG peptides. **(f,h)** Views of the hexamer and pentamer from a cryoEM reconstruction (3.9 Å) of the declination from WT CLPs incubated with excess CPSF6-FG peptide (0.4 contour). Pink indicates bound IP6, found in both hexamer and pentamer. Lavender indicates that the peptide is only bound to the hexamer. **(g)** Close-up of the binding interaction in the hexamer pocket.