VIROLOGY

Kinetic implications of IP_6 anion binding on the molecular switch of HIV-1 capsid assembly

Manish Gupta[1]†, Arpa Hudait[1]†, Mark Yeager[2,3], Gregory A. Voth[1]*

HIV-1 capsid (CA) proteins self-assemble into a fullerene-shaped CA, enabling cellular transport and nuclear entry of the viral genome. A structural switch comprising the Thr-Val-Gly- Gly (TVGG) motif either assumes a disordered coil or a 3_{10} helix conformation to regulate hexamer or pentamer assembly, respectively. The cellular polyanion inositol hexakisphosphate (IP_6) binds to a positively charged pore of CA capsomers rich in arginine and lysine residues mediated by electrostatic interactions. Both IP_6 binding and TVGG coil-to-helix transition are essential for pentamer formation. However, the connection between IP_6 binding and TVGG conformational switch remains unclear. Using extensive atomistic simulations, we show that IP_6 imparts structural order at the central ring, which results in multiple kinetically controlled events leading to the coil-to-helix conformational change of the TVGG motif. IP_6 facilitates the helix-to-coil transition by allowing the formation of intermediate conformations. Our results suggest a key kinetic role of IP_6 in HIV-1 pentamer formation.

INTRODUCTION

HIV-1 particle maturation involves sequential cleavage of the Gag polyprotein, leading to a conical capsid (CA) assembly with the liberated CA protein (*1*–*4*). The CA protects the viral RNA from the cytosolic immune defense during transport within the cell and mediates nuclear import through the nuclear pore complex (NPC) (*5*–*9*). There is an increasing degree of evidence that predominantly intact CA enters the nucleus, and reverse transcription of the viral genome occurs in the CA interior, leading to the release of DNA at the sites of integration (*8*–*11*). The conical shape of the CA is particularly relevant for nuclear entry, as it can dock at the NPC central channel when approaching from the narrow end (*8*). This allows the NPC pore to progressively dilate, relieving stress as the scaffold nucleoporins encounter the wider sections of the CA (*6*, *8*). The CA uses varying surface curvature to selectively interact with cellular host factors (*12*, *13*). When the CA is docked at the NPC central channel, the host factor nucleoporin 153 prevalent at the nuclear basket interacts with the narrow end to facilitate translocation into the nucleus. These features suggest that there is a distinct evolutionary advantage of the conical shape of the CA for productive viral infection.

HIV-1 CA comprises hexamers with exactly 12 pentamers at the wide and narrow ends to stabilize the curvature required for CA closure (*4*, *14*). The CA monomers consist of the N-terminal domain (CA_{NTD}) and C-terminal domain (CA_{CTD}). The CA_{NTD} consists of seven α helices and a β hairpin, and the CA_{CTD} consists of four α helices. The CA_{NTD} and CA_{CTD} are connected by a linker, which imparts conformational flexibility (*15*). The CA monomer leverages conformational flexibility to control quasi-equivalent capsomer association, imparting varying curvature in the CA lattice (*15*, *16*). The interactions between CA_{NTD} allow the formation of the capsomers, while the interactions between CA_{CTD} regulate the assembly of hexamers and pentamers to form the CA lattice. The small-molecule polyanionic inositol hexakisphosphate (IP_6) is an assembly cofactor for both immature and mature HIV-1 CA lattices (*17*–*21*). Structural characterization of the mature CA identified that IP_6 coordinates R18 residues in the central pore of CA_{NTD}, enabling nucleotide diffusion into the CA for reverse transcription (*22*, *23*). In addition, IP_6 can also bind to the ring of K25 residues below the R18 ring (*24*). During in vitro reconstitutions of CA assembly, the addition of IP_6 promotes cone-shaped morphologies by coordinating the positively charged ring of R18 and K25 in the central pore of both the hexamer and pentamers (*17*, *18*, *25*). Several studies have established the mechanistic models of the role of IP_6 in CA assembly. In the pentamers, IP_6 can bind to R18 and K25 rings with comparable probability. In contrast, in the hexamers, IP_6 preferentially binds to the R18 ring because the pore size for the K25 ring is substantially larger (*12*). Recent cryo–electron microscopy (cryo-EM) and molecular dynamics (MD) simulations demonstrated that pentamer formation is dependent on the reduction of electropositive repulsion in the central pore by IP_6 binding (*25*–*27*). Free energy quantification reveals a more stable binding of IP_6 to R18 in CA pentamers relative to hexamers (*25*, *27*). The stabilization of IP_6 is facilitated by smaller pore volume in the pentamer, which allows tighter coordination (*27*).

The relative arrangement of CA_{NTD} and CA_{CTD} is mostly identical in hexamers and pentamers. The residues 58 to 61 at the base of helix 3 constitute the Thr-Val-Gly- Gly (TVGG) motif. The TVGG motif is in a coil conformation in the hexamer, and it forms a 3_{10} helix in the pentamer (*12*, *28*). The TVGG motif is juxtaposed at the NTD-NTD and NTD-CTD interfaces that maintain capsomer stability (*29*, *30*). The interprotomer NTD-NTD interface involves contact between P38/M39 in helix 2 and N57/T58 in helix 3 of neighboring subunits in the hexamer (*29*), whereas P38/M39 contacts V24/K25 in the adjacent helix 1 in the pentamer (*28*). Notably, cryo-EM and biochemical analyses revealed that a structural switch in the TVGG motif of CA modulates the relative distribution of hexamers and pentamers in the polyhedral CA (*28*). Furthermore, the 3_{10} helix formation results in NTD-NTD and NTD-CTD interactions that are unique to CA pentamers.

It has been proposed that IP_6-induced conformational switching may occur through allosteric communication through the hydrophobic

[1]Department of Chemistry, Chicago Center for Theoretical Chemistry, Institute for Biophysical Dynamics, and James Franck Institute, The University of Chicago, Chicago, IL 60637, USA. [2]Frost Institute for Chemistry and Molecular Science, University of Miami, Coral Gables, FL 33124, USA. [3]Department of Chemistry, University of Miami, Coral Gables, FL 33146, USA.
*Corresponding author. Email: gavoth@uchicago.edu
†These authors contributed equally to this work.

core mediated by a network of interactions. Ligand-induced allostery can occur through changes in the protein structure or dynamics (*31*, *32*). Key questions are as follows: Does IP_6 binding trigger the conformational switching of the TVGG motif? If IP_6 influences the conformational switching, what structural and dynamic alternation of the CA subunit does it induce that culminates in the coil-to-helix transition?

To answer these mechanistic questions, we performed extensive unbiased all-atom (AA) MD simulations, conformational sampling, and well-tempered metadynamics (WT-MetaD) (*33*–*35*) biased free energy sampling simulations to investigate the network events triggered by the binding of IP_6 and conformational switching of the TVGG motif. Our results identify a three-step mechanism initiated by IP_6 binding to CA pentamers that results in the folding of the TVGG random coil to the 3_{10} helix. In the pore of the pentamer, IP_6 occupies midway between the planes of the basic R18 and K25 rings. In the presence of IP_6, the K25 residues are conformationally rigidified, coordinating to IP_6 and allowing stable K25-M39 contacts. Hence, M39 points away from the TVGG motif, creating a pocket for the TVGG motif in a coil conformation to fold into a partial 3_{10} helix. During both unbiased and WT-MetaD simulations in the presence of IP_6, the TVGG motif undergoes the coil-to-helix transition mediated via a partially folded intermediate state. Our results uncover several key mechanistic details of how IP_6 regulates TVGG motif conformational switching in CA pentamers through allostery. These results also may point to previously underexplored druggable targets that could impede HIV-1 CA formation.

RESULTS

IP_6 induces folding of the TVGG motif in the CA pentamer

Our atomic model of the mature HIV-1 CA system contains roughly 100 million atoms (*36*), making it computationally intractable for exhaustive, unbiased AA MD simulations. Instead, to limit computational expense, we simulated a global pentamer structure to investigate the relationship between IP_6 binding and the TVGG coil-to-helix transition. Here, the "global pentamer" patch comprised a central pentamer surrounded by five peripheral hexamers (Fig. 1A) in the solvent. To simulate the IP_6 effect, another global pentamer system was prepared by adding IP_6 anions roughly 3 Å above the central R18 ring in each capsomere. In total, four replicas of the global pentamer were simulated for a total of ~12 μs with and without IP_6.

We systematically analyzed the structural changes in the TVGG motif of the central pentamer of the CA patch. In the initial pentamer structure, the TVGG motif resembled a random coil configuration, equivalent to the conformation in hexamers. Although the disordered conformation of the TVGG motif is termed "random

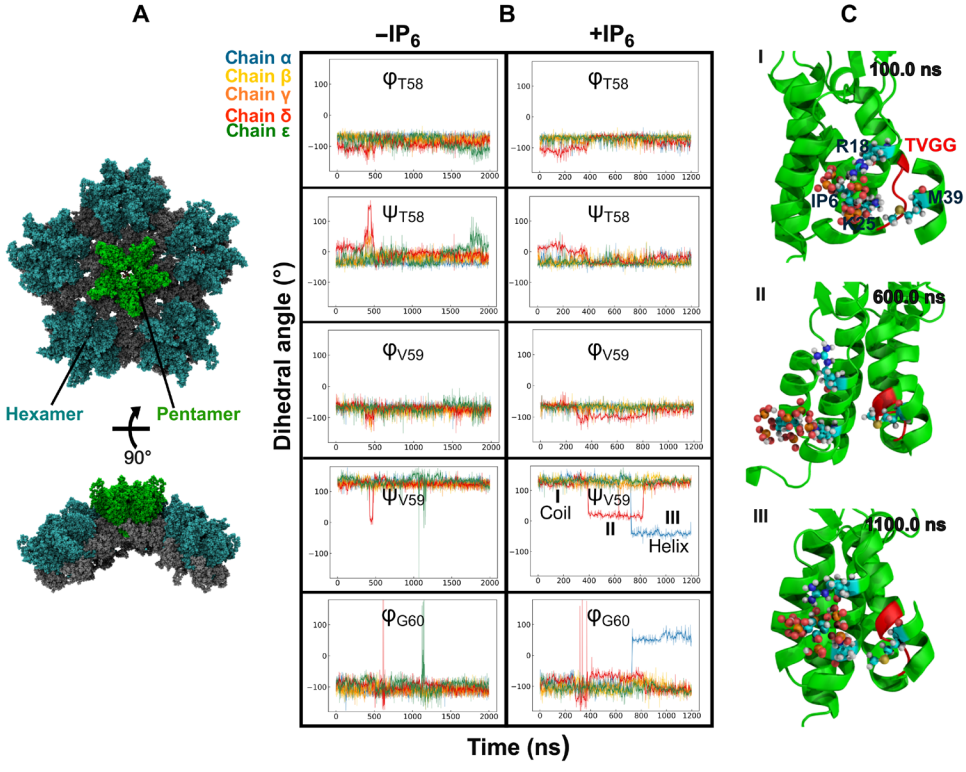

Fig. 1. IP_6 facilitates coil-to-helix transition in the TVGG motif of the CA pentamer. (**A**) The "global pentamer," made of a central CA pentamer (NTD in green and CTD in gray) and five adjacent hexamers (NTD in blue and CTD in gray), forms a curved lattice. (**B**) Time series of TVGG backbone (φ, ψ) dihedral angles of each subunit in the central pentamer without (left panel) and with IP_6 (right panel) for times longer than 100 ns. In the absence of IP_6, the TVGG (φ, ψ) dihedral angle space briefly explored a metastable state but primarily stayed in the coil conformation. In all the IP_6-bound pentamer simulations, the TVGG motif transitioned either to a metastable state with ψ_{V59} and φ_{G60} values of ~10° and −60°, respectively, or to a partial helix conformation with ψ_{V59} and φ_{G60} values of −45° and 60°, respectively. (**C**) Snapshots from the IP_6-bound global pentamer MD trajectory show the conformational transition in the TVGG motif of the CA pentamer. Zoomed-in view of the TVGG motif (red) of α subunit in the coil conformation is shown in (I) at the top. During the course of the simulation, TVGG motifs in the δ subunit and α subunit in the pentamer transition to metastable (II) and partial 3_{10} helix (III) conformations, respectively. Time steps at which the snapshots are recorded are labeled at the upper-right corner of each panel.

coil," there is an ensemble of backbone geometries in the (φ, ψ) dihedral angle space (Fig. 1B). The pentamer TVGG backbone (φ, ψ) dihedral angles fluctuated persistently in the absence of IP_6. T58 (φ, ψ) dihedral angles gradually converged to approximate values of −60° and −45°, respectively, which are consistent with a 3_{10} helix conformation. Notably, φ_{V59} resembled a helix-like structure with a value of −60°, but the ψ_{V59} value of ~125° deviated from the 3_{10} helix. φ_{G60} converged to a large negative value of −100°, but ψ_{G60} and G61 (φ, ψ) adopted a random coil conformation.

In the presence of IP_6, the fluctuation of the TVGG backbone (φ, ψ) dihedral angles is significantly altered. T58 (φ, ψ) and φ_{V59} dihedral angles resembled a 3_{10} helix. However, major transitions in ψ_{V59} and φ_{G60} values were observed in the α subunit. ψ_{V59} and φ_{G60} transitioned to values of −45° and 60° (Fig. 1B), respectively, to allow a coil-to-helix switch within the TVGG motif (Fig. 1C). In the δ subunit, ψ_{V59} transitioned to a value of 10°, and φ_{G60} concurrently transitioned to −60°, indicating the existence of a transiently populated and partially folded intermediate state between the coil and the 3_{10} helix. The relaxation of ψ_{V59} to a 3_{10} helix configuration was more pronounced for pentamers with IP_6, suggesting that the TVGG coil-to-helix switch might be directly related to IP_6 binding at the central channel. The anionic phosphate groups of IP_6 remain bound to the R18 and K25 rings throughout the duration of the simulation. In all except one IP_6-bound pentamer simulation, ψ_{V59} switched to a value of −45°, and φ_{G60} switched to 60° concomitantly, assuming a 3_{10} helix structure.

IP_6 imparts order in the positively charged central pore

The CA hexamer and pentamers contain two electropositive rings composed of R18 and K25 residues, mutually stacked in the central pore. The abundance of basic residues renders the pore strongly electropositive, inducing asymmetry at the center (*13*). Visual inspection of the simulation trajectories revealed that R18 residues are highly orientationally dynamic (Fig. 2A). To characterize the orientational fluctuations, we estimated the variation in root mean square fluctuation of the helix 1 center of mass relative to the other helix 1 in the pore (Fig. 3C). Root mean square fluctuation calculations revealed a persistent disorder in the central ring of the pentamer in the absence of IP_6. Evidently, the R18 residue fluctuated between three conformations, directing either toward the exterior, center, or interior of the CA pore. Snapshots of the pentamer central pore without and with IP_6 are shown in Fig. 2 (A and B, respectively), and the dynamics are shown in movies S1 and S2. The orientational disorder of the R18 residues is instigated by the electrostatic repulsion between the charged guanidino groups at the pore. Consequently, the guanidino groups fluctuate rapidly to orient in the opposite direction of the adjacent positively charged moieties of arginine.

In the presence of IP_6, the guanidinium side chains of R18 were ordered and flipped toward the channel's center (Fig. 2B). Despite the smaller pentamer pore size, IP_6 could not stabilize all five R18 side chains simultaneously, leaving at least one R18 residue uncoordinated. K25 side chains were also oriented toward the center of the pore to coordinate IP_6. Notably, IP_6 binding had a significant influence on the central pore, as its dynamic fluctuations were reduced (Fig. 3, B-I and D-I). Previous ligand density maps have identified the metastable interactions that guide free IP_6 toward the CA central pore (*27*). However, longer simulations of IP_6-bound CA were required to access accurate spatial distribution and dynamics of IP_6 in the binding pocket. We calculated a three-dimensional ligand density map and compared the IP_6 binding modes at the hexamer (Fig. 4A) and pentamer central pore (Fig. 4B). We observed that IP_6 was preferably located above the R18 ring, which is known as the primary binding site in the hexamer pore. Only a weak IP_6 ligand density was observed below the R18 ring, likely due to the relatively larger diameter of the K25 ring for the hexamer. Intriguingly, significant IP_6 density was observed below the R18 ring, highlighting the differential mode of binding in the pentamer central pore.

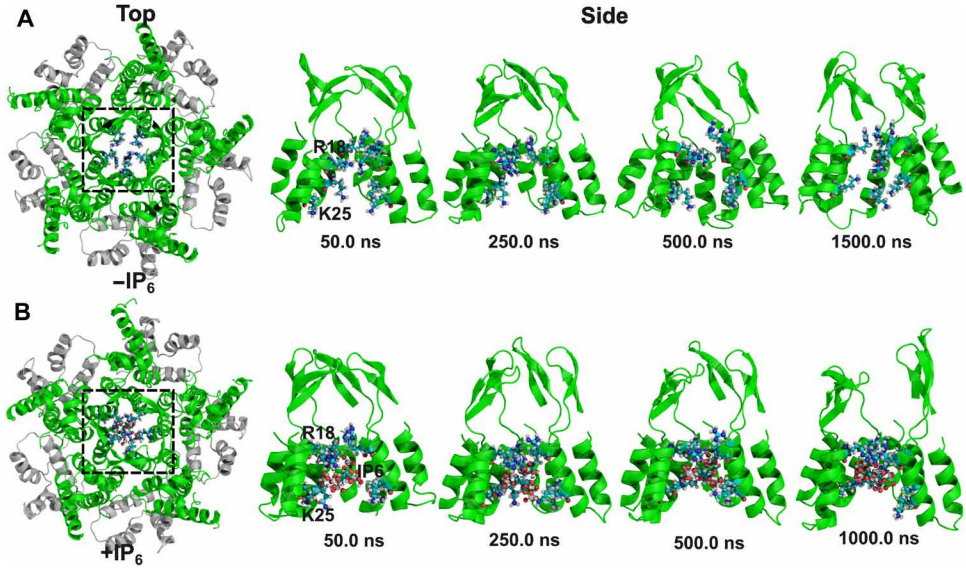

Fig. 2. Comparison of central pores from the CA pentamer and IP_6-bound pentamer. (**A**) Snapshots from (−IP_6) simulations show that the central pore is rendered asymmetric by the repulsion of positively charged residues (NTD is shown in green, and CTD is shown in gray). Close-up views of the central channel show that in the absence of any stabilizing force, R18 residues fluctuate between three different configurations, whereas K25 is oriented toward the interior of the CA. The ε subunit is not shown for clarity. (**B**) IP_6 imparts fivefold symmetry in the central channel. IP_6 is bound below the R18 ring, coordinating multiple R18 and K25 residues. R18 and K25 side chains point toward the center of the CA to contact IP_6.

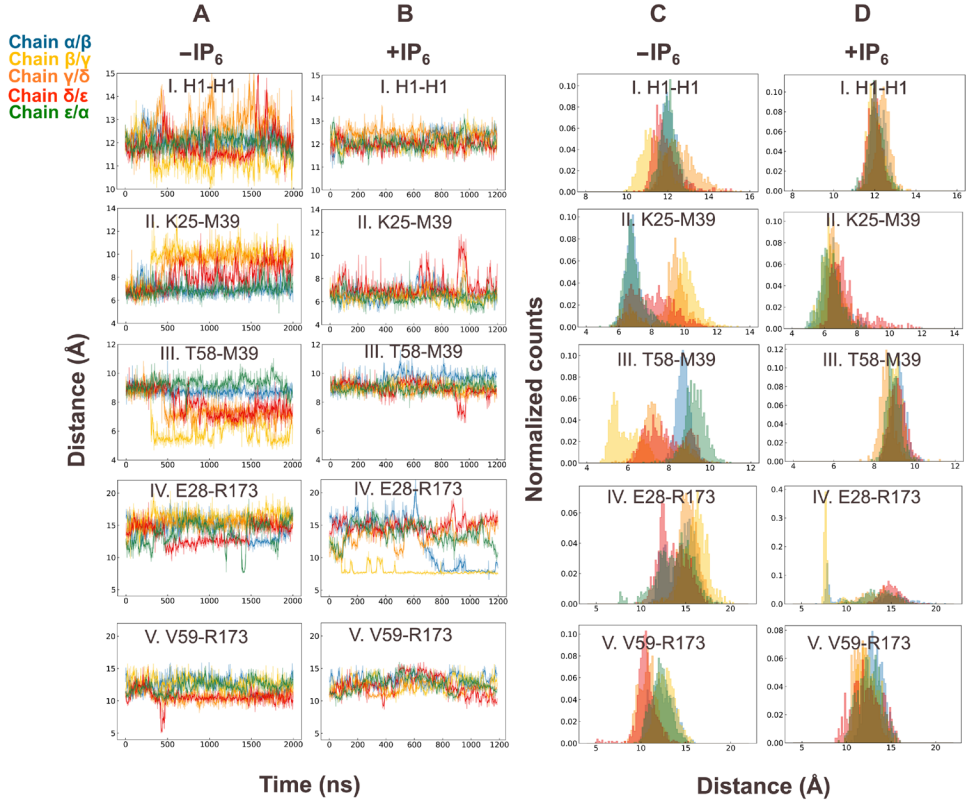

Fig. 3. Structural arrangement in NTD-NTD and NTD-CTD interfaces in the CA pentamer. (**A**) Time courses of (I) H1(CA_i)-H1(CA_{i+1}), (II) K25(CA_i)-M39(CA_{i+1}), (III) T58(CA_i)-M39(CA_{i+1}), (IV) E28(CA_i)-R173(CA_{i+1}), and (V) V59(CA_i)-R173(CA_{i+1}) interactions in the pentamer. (**B**) Time courses of (I) H1(CA_i)-H1(CA_{i+1}), (II) K25(CA_i)-M39(CA_{i+1}), (III) T58(CA_i)-M39(CA_{i+1}), (IV) E28(CA_i)-R173(CA_{i+1}), and (V) V59(CA_i)-R173(CA_{i+1}) interactions in the IP_6-bound pentamer. (**C**) Histograms of (I) H1(CA_i)-H1(CA_{i+1}), (II) K25(CA_i)-M39(CA_{i+1}), (III) T58(CA_i)-M39(CA_{i+1}), (IV) E28(CA_i)-R173(CA_{i+1}), and (V) V59(CA_i)-R173(CA_{i+1}) interaction distances in the pentamer. (**D**) Histograms of (I) H1(CA_i)-H1(CA_{i+1}), (II) K25(CA_i)-M39(CA_{i+1}), (III) T58(CA_i)-M39(CA_{i+1}), (IV) E28(CA_i)-R173(CA_{i+1}), and (V) V59(CA_i)-R173(CA_{i+1}) interaction distances in the IP_6-bound pentamer.

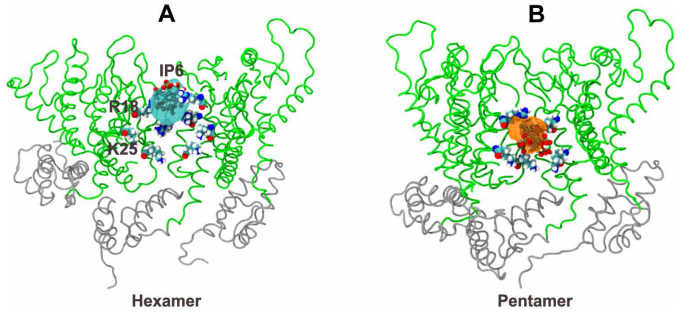

Fig. 4. Characterization of IP_6 binding modes and structural changes in the pentamer. (**A**) The IP_6 density map is overlaid with the CA hexamer structure. Densities contribute to IP_6 binding above the R18 ring in the hexamer with only a small ligand density below the R18 ring. (**B**) The ligand density map in the pentamer contains IP_6 primarily below the R18 ring.

IP_6 binding promotes pentamer-like NTD-NTD and NTD-CTD packing

Recent high-resolution cryo-EM structures have shown distinctly different packing at the NTD-NTD interface for the HIV-1 hexamer and pentamer (*28*). The NTD-NTD interaction between adjacent CA subunits in the hexamer is mediated by contact between M39 in helix 2 and T58 in helix 3. In CA pentamers, M39 in helix 2 preferentially contacts K25 in helix 1 (*37*). In our unbiased AA MD simulation trajectories, we analyzed the molecular contacts of K25, M39, and T58 in the presence or absence of IP_6.

The time series plots of the K25(CA_i)-M39(CA_{i+1}) and M39(CA_i)-T58(CA_{i+1}) distances between adjacent subunits are shown in Fig. 3. In the absence of IP_6, the root mean square distance between the centers of mass of K25-M39 between adjacent CA subunits increases. Simultaneously, the M39-T58 distances also decrease, which indicates the formation of contact. The distribution of the K25-M39 fluctuations revealed that stable K25-M39 packing is not favored in the absence of IP_6 (Fig. 3, A-II and C-II). In the absence of IP_6, the disordered central pore prevented the stable packing of the M39 residue in the hole formed by V24 and K25, even though this configuration is preferred in pentamers. Histograms of H1(CA_i)-H1(CA_{i+1}) and K25(CA_i)-M39(CA_{i+1}) distances in Fig. 3 distinctly demonstrate that the fluctuations within the H1-H1 contacts in the β/γ, γ/δ, and δ/ε CA subunits are relayed to K25-M39 contacts in the same subunits. The loss of K25-M39 contact moved the M39 residue toward N57 and T58, forming a hexamer-like NTD-NTD interface (Fig. 3, A-II, A-III, C-II, and C-III).

To summarize, in the absence of IP_6, the molecular environment in the pentamer drifts toward the hexamer. Fluctuations of residues adjacent to the central pore of pentamers are significantly reduced in the presence of IP_6 (Fig. 3, B-I and D-I). This allowed stable packing of M39 against the V24/K25 hole (Fig. 3, B-II and D-II),

trapping it in the pentamer state. The formation of molecular contacts between helices 1 and 2 excludes helix 3 from the NTD-NTD packing, creating an empty pocket and allowing the TVGG motif to fold into a 3_{10} helix. A similar trend is observed in all four replica simulations (figs. S1 to S6).

The TVGG motif is located at the NTD-CTD interface that stabilizes the intra-capsomer contacts between adjacent CA subunits. The pocket is the binding site for host cell factors responsible for nuclear import and small molecule inhibitors that target the HIV-1 CA (*38*, *39*). In the hexamer, V59 of the TVGG motif interacts with R173 of the neighboring CA domain, contributing to the packing of helices 3 and 8 at the NTD-CTD interface (*12*, *28*). In contrast, NTD-CTD packing in the pentamer is facilitated by interaction between E28(CA_i) in helix 1 and R173(CA_{i+1}) in helix 8. We hypothesized that the ordered central pore should favor contact between E28(CA_i) and R173(CA_{i+1}) in the IP_6-bound pentamer. We observed close contact between E28(CA_i) and R173(CA_{i+1}) in multiple chains (Fig. 3, B-IV and D-IV, and figs S2, S4, and S6) in the IP_6-bound pentamer. In the CA patch simulations, the distance between E28 and R173 residues in adjacent CA chains increased to a value of ~12 to 20 Å (Fig. 3, A-IV and C-IV), and the V59-R173 distance decreased in the absence of IP_6 (figs. S2, S4, and S6), signifying a tendency to adopt a hexameric arrangement. This suggests that IP_6 promotes pentamer-like NTD-NTD and NTD-CTD packing, while in the absence of IP_6, the pentamer structure drifts toward the hexameric state.

To further investigate the structural rearrangements effected by IP_6, we compared contacts made between CA_i helix 1 or 3 and helix 2 of CA_{i+1} (fig. S7). In the IP_6-bound pentamer, residues located at the base of helix 1 engage residues 34 to 36 in helix 2 of the neighboring protomer, favoring pentamer-specific contacts. In contrast, a hexamer-like NTD-NTD interface was adopted in the absence of IP_6 as residues 57 to 60 located at the base of helix 3 slipped toward helix 2 of the adjacent CA subunit. These findings suggest that IP_6 binding can influence the TVGG switch in the pentamer by modulating interactions within the NTD-NTD and NTD-CTD interfaces.

Partially folded intermediates facilitate coil-to-helix transition in the IP_6-bound pentamer

In long unbiased MD simulations in the presence of IP_6, the TVGG motif fluctuates between these states: coil → partially folded → 3_{10} helix. In contrast, in the absence of IP_6, the TVGG motif undergoes a coil ⇌ 3_{10} helix transition, albeit rarely, and does not form a stable 3_{10} helix. Because of the limited timescales accessible in our unbiased atomistic simulations, we observe coil-to-helix transitions in only a few CA protomers, as the complete folding of TVGG in all CA subunits of the pentamer occurs on a substantially longer timescale. To assist in the interpretation of the above unbiased atomistic MD simulations, we performed WT-MetaD simulations of a CA pentamer in the presence and absence of IP_6. Briefly, WT-MetaD simulation is an enhanced sampling method that can accelerate the timescales of transitions between states, which are often not sampled in unbiased simulations (*35*, *40*). WT-MetaD simulations are performed as a function of appropriate reaction coordinates (often called "collective variables" or CVs), which provide an energetic landscape of the conformational transition. In the WT-MetaD simulations, the central pentamer is surrounded by partial hexamers (see Materials and Methods for details of system construction). The reaction coordinates in the WT-MetaD simulations are the distances between K25 of CA_1 and M39 of CA_2 denoted as $d_{K25-M39}$, and the helicity parameter (h_{V59}) is calculated from ψ_{V59} of CA_1 (time series shown in figs. S8 and S10). Because of computational constraints, we only performed the biased simulation for a single chain CA_1. The helicity 0 and 1 correspond to the coil and helix conformations, respectively. The choice of the reaction coordinates is informed by the analysis of the unbiased simulations. The potential of mean force (PMF; i.e., conditional free energy) of TVGG motif folding (Fig. 5A and fig. S9; energy scale, 0.2 kcal/mol) is presented as a function of the reaction coordinates h_{V59} and $d_{K25-M39}$.

In the PMFs (in both the presence and absence of IP_6), there is a deep energy minimum for the coil conformation of the TVGG motif. The minimum corresponds to a $d_{K25-M39}$ value ranging from 6 to 6.8 Å. While remaining in the coil conformation, fluctuations of the K25(CA_1)-M39(CA_2) contact can lead to higher $d_{K25-M39}$ values. When in the helix conformation, the energetically stable basin corresponds to a $d_{K25-M39}$ value ranging from 6 to 6.8 Å. Similar to that in the coil state, fluctuations allow for the K25(CA_1)-M39(CA_2) contacts to weaken, exploring higher $d_{K25-M39}$ values. Therefore, both coil and helix conformations allow for heterogeneous distribution of the K25-M39 contact, consistent with the results of the unbiased simulations. In the presence of IP_6 (for both replicas), coil-to-helix transition occurs through a funnel-like landscape. As the TVGG motif in the coil conformation begins to fold, along the reaction coordinates h_{V59}, the K25(CA_1)-M39(CA_2) distance fluctuations narrow to 6 to 6.8 Å. In the presence of IP_6, the barrier corresponds to an h_{V59} value of ~0.3. Beyond this barrier, there is a distinct intermediate state corresponding to the h_{V59} value range of ~0.52 to 0.65. The barrier to forming the partially folded intermediate is ~2 kcal/mol, while forming the 3_{10} helix from the intermediate state is essentially barrierless. The presence of this partially folded state is

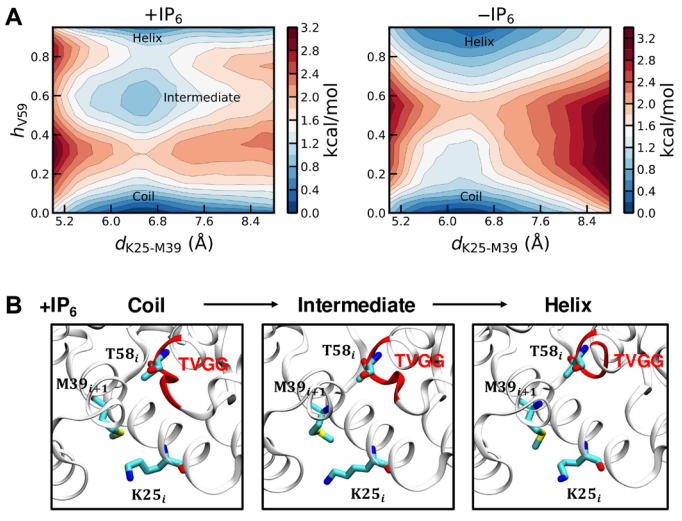

Fig. 5. Molecular pathway of coil-to-helix transition in the CA pentamer modulated by IP_6. (**A**) Two-dimensional plot of the PMF of the conformation landscape of the TVGG motif as a function of the helicity (h_{V59}) of V59 and distance between K25 of CA_1 and M39 of CA_2 denoted as $d_{K25-M39}$. The bar on the right has the color code for energy level contours in kilocalories per mole. The PMFs were calculated from the WT-MetaD simulations performed in the presence (+IP_6) and absence of IP_6 (−IP_6). (**B**) Snapshots of the folding of the TVGG motif (shown in red ribbon) in the presence of IP_6. The CA monomers (except for the TVGG motif) are shown in silver ribbon. The residues $K25_i$, $M58_i$, and $M39_{i+1}$ are highlighted in the snapshots.

observed in the presence of IP_6, consistent with the observations of the unbiased MD simulations. In the absence of IP6, in one replica, the coil-to-helix transition occurs through a funnel-like landscape, analogous to that of the PMF in the presence of IP_6, albeit without the intermediate state. In the absence of IP_6, the barrier corresponds to an h_{V59} value of ~0.5. The barrier of conformational change is ~2 kcal/mol. As the barrier is crossed along the h_{V59} coordinate, the K25(CA_1)-M39(CA_2) distances gradually become more heterogeneous. In the second replica (in the absence of IP_6), the states along the coordinates h_{V59} are sampled in a barrierless fashion in a smooth free energy landscape. This can be attributed to orientational disorder (akin to structural plasticity) of the relevant residues between helices 1 and 3 in the absence of IP_6 and is consistent with the unbiased simulations. Our enhanced sampling simulations demonstrate that IP_6 binding to the pore of the pentamer directs the pathway of coil-to-helix transition via an intermediate partially folded state.

It is informative to examine whether the intermediate partially folded conformation constitutes a metastable state in the presence of IP_6. To this end, we picked five independent configurations at the intermediate state basin from the WT-MetaD simulation trajectory. We simulated these configurations for 1000 ns to examine the evolution of the intermediate-to-helix or intermediate-to-coil state populations. The system starts at $t = 0$ at the intermediate state with the h_{V59} value ranging from ~0.54 to 0.69. In one replica (Fig. 6), the TVGG motif continuously jumped between the helix, coil, and intermediate states. Typically, the residence time at each state is ~10 to 30 ns, which corresponds to the fast interstate fluctuations and demonstrates the metastable nature of the partially folded intermediate. These fast interstate fluctuations can be interpreted as attempts to convert to either the disordered coil or folded helix conformations. These fluctuations persist up to 300 ns, after which the TVGG motif converts to the coil conformation and remains as such up to 1000 ns. In contrast, in another replica, the TVGG motif starting from the intermediate state converts to the helix within 10 ns and does not undergo any further helix-to-coil transition within 1000 ns (fig. S10A).

In the other three replicas, the TVGG motif at the intermediate state converts to the random coil conformation within 50 ns (fig. S10, B to D). These simulations establish that the partially folded intermediate conformation is a metastable state, as demonstrated by the short lifetime. In addition, when starting from the intermediate conformation, the TVGG motif is more likely to relax to a random coil conformation.

Our results provide the following mechanistic picture. Folding of a peptide motif leads to a decrease in conformational entropy, which is compensated by favorable enthalpic interactions. For the folding kinetics of a small chain, such as the TVGG motif, conformational entropy is the likely factor that determines the energetic landscape (*41*). The barrier to partially folded → helix transformation is barrierless, yet in our simulations, the partially folded intermediate preferentially converts to the random coil conformation. This can be rationalized by considering the gain of conformational entropy for the partially folded → coil transformation. In addition, in a folding pathway, the formation of intermediates can accelerate the kinetics (*42*). We therefore propose that IP_6 facilitates pentamer formation by accelerating the kinetics of the coil-to-helix transition mediated by an intermediate state during CA self-assembly.

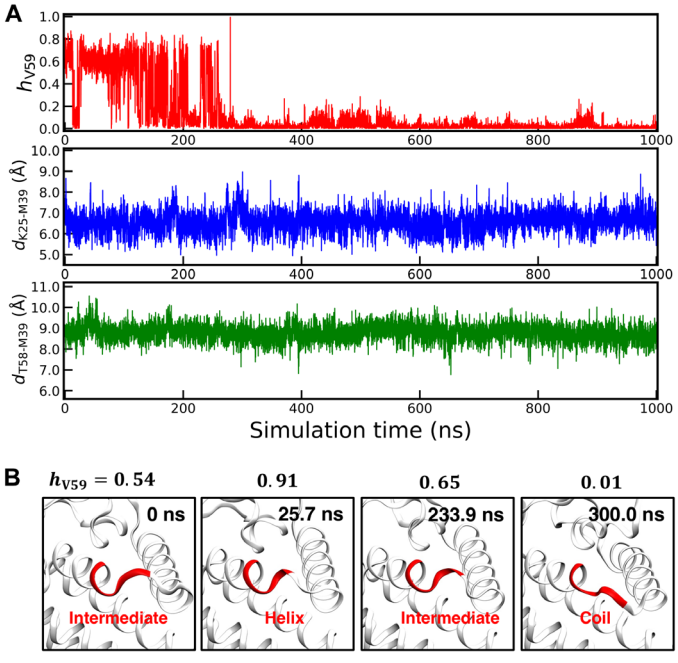

Fig. 6. Dynamics of coil-to-helix transition from the intermediate partially folded state. (**A**) Time evolution of the helicity parameter (h_{V59}), the distance between K25 and M39 ($d_{K25-M39}$) of CA_1 and CA_2, and the distance between T58 and M39 ($d_{T58-M39}$) of CA_1 and CA_2. The initial configuration corresponds to the intermediate partially folded state obtained from the WT-MetaD simulations in the presence of IP_6. In the initial configuration, the values of h_{V59}, $d_{K25-M39}$, and $d_{T58-M39}$ are 0.54, 7.2 Å, and 8.7 Å, respectively. (**B**) The snapshots show the continuous folding and unfolding of the TVGG motif (highlighted in red ribbon) within the first 300-ns simulations. The rest of the CA protein is shown in silver ribbon. For each snapshot, the title indicates the helicity parameter (h_{V59}) value.

DISCUSSION

The typical morphology of the canonical mature CA is a fullerene cone. In accordance with Euler's theorem, the HIV-1 CA cone requires exactly 12 pentamers to ensure the closure of the polyhedron (*43*). The way the CA protein adjusts its structure to pentamers and hexamers for different regions of the lattice curvature remains an important area of investigation in HIV-1 structural biology. Analysis of the CA monomers from low-resolution structures of cores proposed that the interdomain hinge at helix 9 likely provides the requisite flexibility to CA monomers to adapt to lattice curvature (*15*, *44*). High-resolution structures revealed that CA_{NTD} in the hexamer and pentamer are nearly identical, with similar positioning of the hinge (*12*). Recent structural studies have identified the TVGG motif as a molecular switch that regulates hexamer or pentamer formation (*12*, *28*). CA tubes composed exclusively of hexamers are formed in abundance, while pentamers are rarely formed during spontaneous CA self-assembly (*18*, *45*). Therefore, CA monomers predominantly self-assemble to hexamers in the absence of IP_6. In contrast, IP_6 promotes the formation of cone-like structures (*25*), which require pentamers for closure. Hence, a minimal combination of CA protein and IP_6 is necessary for pentamer formation. Because, structurally, the CA hexamers and pentamers in the mature CA cone differ at the TVGG motif, IP_6 likely triggers the TVGG conformational change.

After the formation of the pentamer during CA self-assembly (*25*), the following molecular pathways are conceivable for the pentamer/hexamer switch: (i) coil-to-helix transition followed by IP_6 binding at the pore locking the TVGG motif in the 3_{10} helix state or (2) IP_6 binding at the pore triggering the coil-to-helix transition. In this study, we tested both these pathways by simulating CA patch with and without IP_6. In the absence of IP_6, the TVGG motif primarily adopted the coil conformation with no coil-to-helix transition. This likely explains the rarity of pentamers during CA self-assembly and eliminates the first pathway. In the presence of IP_6, the likelihood of coil-to-helix transition increases, thus implicating IP_6 as the primary trigger for the TVGG switch.

Our group's previous atomistic simulation suggested that IP_6 binding below the R18 ring is more favored in the hexamer than the pentamer (*27*). The ligand density calculations of these simulation trajectories identified higher IP_6 density below the R18 ring in the hexamer than in the pentamer. In our more extensive simulations of this paper, IP_6 coordinated positively charged moieties from both R18 and K25 rings in the pentamer, leaving the coordination site above the R18 ring available for additional IP_6 binding. This result is consistent with recent cryo-EM analysis that revealed a distinctly strong IP_6 density below the R18 ring in the pentamer but a relatively weaker IP_6 density in the hexamer (*26*). K25A mutants can form tubes, but their ability to assemble conical CA is significantly diminished (*45*, *46*). If electrostatic repulsion within the central pore were the predominant factor in mature CA assembly, then K25A should exhibit enhanced conical assembly. In contrast, both the K25N mutant and wild-type CA can produce mature CA in the presence of IP_6 (*23*), suggesting a further role of IP_6 binding at the K25 ring. This is further supported by recent studies showing that K25A compensatory mutations can restore pentamer assembly (*46*), while R18L mutants adopt an alternative pentamer arrangement with hexamer-like NTD-NTD packing (*47*). These findings indicate the presence of an allosteric network that regulates the pentamer/hexamer equilibrium. Notably, the following steps have been identified to govern pentamer formation: (i) IP_6 binding, (ii) 10-helix bundle formation involving helix 1/helix 2 of the adjacent NTDs, and (iii) coil-to-helix transition at the TVGG motif in helix 3, although not necessarily in that order. In a previous study, coarse-grained simulations of the entire CA self-assembly process demonstrated that transient pentamers can form initially and are trapped upon IP_6 binding (*25*). In the absence of IP_6, the M39 residues fluctuated between interactions with K25 in helix 2 and T58 in helix 3 of adjacent CA monomers, indicating that the CA molecules had not yet locked in the pentamer state. The placement of IP_6 at the central pore of the pentamer stabilized the packing of helices 1 and 2, enabling a coil-to-helix transition of the TVGG motif in multiple CA subunits. Collectively, these findings suggest that IP_6 stabilizes the central pore of the pentamer, facilitating the packing of helices 1 and 2 and triggering the structural rearrangement of the TVGG motif in helix 3 into a partial helix. This transition ultimately locks the CA into the pentamer state. It is important to note that the time required to observe the full coil-to-helix transition of the TVGG motif across the entire pentamer could not be accessed in our unbiased simulations because of computational limitations.

In our simulations, the NTD-NTD interactions between helices 1 and 2 are mediated by K25-M39 packing of adjacent subunits at the hydrophobic core. These molecular contacts play an important role in modulating the coil-to-helix transition, consistent with the recently proposed ratchet mechanism of the hexamer/pentamer switch (*28*). In one of the replica simulations, IP_6 binding failed to impose order at the pentamer's central pore during the simulation period, leading to unstable packing between K25 and M39 (fig. S6). We observed a transient coil-to-helix transition of the TVGG motif, underscoring the importance of stable packing between adjacent K25 in helix 1 and M39 in helix 2 in regulating the TVGG switch. Our results so far have established the steps by which IP_6 binding to the central pore triggers coil-to-helix conformational switching of the TVGG motif. IP_6 binding allows K25-M39 packing at the hydrophobic core, regulating the coil-to-helix transition of the TVGG motif. The allosteric activation of protein because of ligand binding from the inactive to active state is coupled with the change in the energetic landscape of the conformational change. Allosteric activation via ligand binding can occur through enthalpic regulation by modulation of the relative stability of the initial and final states (*48*). In contrast, allosteric activation can occur through entropic regulation, where the relative stability of the end states remains unaltered (*49*). Instead, additional intermediate conformations appear during the conformational change. Our results demonstrate that in the presence of IP_6, an additional intermediate conformation regulates the coil-to-helix transition. Simultaneously, the barrier remains unchanged in the presence of IP_6. Our results show that allostery between IP_6 and TVGG motif conformational switching is entropically regulated.

To conclude, IP_6 is necessary for the fullerene mode of assembly, while its absence results in the remodeling of pentameric defects into the hexamer. The precise molecular mechanism underlying the insertion of a CA protomer into a transient metastable pentamer and the formation of a stable hexamer is of therapeutic interest. Our analysis reveals a high degree of fluctuation within the pentamer pore because of electropositive repulsion in the absence of IP_6. Specifically, interprotomer contact is weakened in helix 1 (movie S1), leading to the distortion of the compact pentameric molecular iris made of β hairpins. A plausible mode of pentamer-to-hexamer remodeling during CA self-assembly can then be via the placement of a CA monomer at this defect site. Future investigations will require new, more highly resolved coarse-grained MD simulations, which can access timescales that are substantially longer than atomistic MD simulations in this work and can better elucidate the mechanistic pathways of pentamer-to-hexamer remodeling.

MATERIALS AND METHODS

System and simulation setup

The initial fully atomistic structures of the CA hexamers and pentamers (residues 1 to 220) were constructed from cryo–electron tomography structures with Protein Data Bank (PDB) nos. 5MCX and 5MCY, respectively (*50*). These atomic structures are derived from intact virus particles. Residues 220 to 231 were obtained from the nuclear magnetic resonance structure of CA_{CTD} (PDB: 2KOD) (*51*). The atomistic structures of the CA hexamer and pentamers were aligned to an equilibrated atomistic structure of a CA cone (*36*) to construct a CA patch with a central pentamer and five surrounding hexamers. In the IP_6 simulations, six IP_6 were positioned roughly 3 Å above the R18 pore in each capsomer based on the x-ray crystal structure (*18*). The protein-cofactor complex was placed in the center of the simulation cell such that the distance between the edge of the protein and the box boundary is 1.2 nm. The protein-cofactor complex was solvated by 795,958 water molecules, and the salt

concentrations were adjusted to maintain 150 mM NaCl and produce a charge-neutral system, resulting in a total of 2,600,199 atoms.

Energy minimization was performed using the steepest descent algorithm, and the minimization was terminated when the maximum force was smaller than 1000.0 kJ/mol per nanometer. The composite system was then equilibrated in a constant NVT ensemble by applying harmonic positional restraint (spring constant value of 1000 kJ/mol per square nanometer) on protein heavy atoms for 10 ns. The temperature of the system was maintained at 310 K using a stochastic velocity rescaling thermostat with a time constant of 1 ps (*52*). The harmonic positional restraints were removed, and the system was equilibrated in a constant NPT ensemble at 310 K and 1 bar for an additional 100 ns. The configurations at the end point of the NPT equilibration were used as initial configurations for the production runs. The production runs were performed at 310 K and 1 bar. In the NPT simulations, the temperature was maintained using a Nose-Hoover chain thermostat with a 2-ps time constant and the pressure with a Parrinello-Rahman barostat with a 10-ps time constant (*53*, *54*).

We prepared a partial CA patch containing a central CA pentamer and two closest CA monomers of each hexamer surrounding the pentamer. The partial CA patch was used to perform conformational sampling simulations and WT-MetaD simulations. In these simulations, the CA monomers are part of the hexameric lattice by applying harmonic positional restraint (spring constant value of 1000 kJ/mol per square nanometer). The solvated partial CA patch system was prepared with the identical protocol as the full CA patch system. The system was solvated by 274,274 water molecules. Salt concentrations were adjusted to maintain 150 mM NaCl, resulting in a system with a total of 878,428 atoms.

All MD simulations were performed with periodic boundary conditions in all directions. The protein was modeled with the CHARMM36m force field (*55*), and water was modeled with TIP3P parameters (*56*). The parameters of IP_6 were generated with cgenff (*57*). The bonds between the protein heavy atoms and hydrogen atoms were constrained with the LINCS algorithm. Electrostatic interactions were computed using the particle mesh Ewald method with a cutoff of 1.2 nm (*58*). The van der Waals force was truncated smoothly to zero between 1.0 and 1.2 nm. The equations of motion were integrated using the velocity Verlet algorithm with a time step of 2 fs (*59*). The unbiased simulations were performed using GROMACS version 2019.6 (*60*). All restrained simulations and WT-MetaD simulations were performed using GROMACS version 2020.4 patched with PLUMED 2.6 (*61*).

WT-MetaD simulation setup

We estimated the free energy landscape of the coil-to-helix transition of the TVGG motif with and without IP_6 bound at the central pore using WT-MetaD simulations (*40*). We used two CVs to perform our WT-MetaD simulation: (i) the distances between the center of mass of the side-chain heavy atoms of K25 of CA_1 and M39 of CA_2 denoted as $d_{K25-M39}$ and (ii) helicity parameter (h_{V59}) of the V59 residue of CA_1. Here, CA_1 and CA_2 are two adjacent monomers in a pentamer. The helicity parameter (h_{V59}) is defined as $0.5 \times \left[1+\cos\left(\psi_{V59}-\psi^{red}_{V59}\right)\right]$. Here, ψ_{V59} is the dihedral angle centered at the V59 residue of CA_1. ψ^{ref}_{V59} is the reference dihedral angle that corresponds to the value of the ideal α helix. The selection of the CVs was informed by unbiased atomistic simulation results.

We estimated the free energy profile of coil-to-helix transition using the approach in (*34*). Briefly, the sampled space in WT-MetaD with microscopic coordinate R is associated with a reweighting factor $w(R) = e^{V[s(R),t]-c(t)}$. Here, $V(x)$ is the bias and $c(t)$ is referred to as a time-dependent bias offset. The term $c(t)$ is estimated "on the fly" in the MetaD simulations. In our simulations, the biasing factor was set to 10, the bias height was set to 0.2 kcal/mol, and the biasing frequency was set to 5000 time steps. These parameters were chosen to ensure that the biasing potential was not added too aggressively. The variances of Gaussian bias were set to be 0.05 Å and 0.05 for $d_{K25-M39}$ and h_{V59}, respectively. The upper bound of $d_{K25-M39}$ was set to 14 Å. In the WT-MetaD simulations, the equations of motion of CA_1 and CA_2 are integrated, and rest of the CA monomers of the pentamers are harmonically restrained.

Analysis of MD simulations

In the production runs, the configuration of the system was saved every 500-ps interval. The MDAnalysis package (*62*) was used to access and align the trajectories. We excluded the first 100 ns of the trajectories from calculations to ensure that the analysis reflects the system's true equilibrium behavior. To calculate the dihedral angles, we aligned the trajectories to the first frame and conducted a Ramachandran analysis on residues 58 to 61 using the dihedral class in MDAnalysis. Positions of the C_α center of mass of the relevant residues were used for the calculation of intersubunit distances. For the histogram calculations, we extracted the distances of interest from the trajectories and divided the data into bins based on the minimum and maximum values of the distance data with a specified bin width. To normalize the histogram, we calculated weights by assigning an equal weight to each data point, normalized by the total number of data points, ensuring that the histogram represented a probability distribution. The histograms were then plotted with the number of bins set to 100. This approach ensured a consistent representation of the distance distributions. The density of IP_6 was calculated by mapping the atomic densities of the ligand on a discretized grid with a volume of 0.5 Å by 0.5 Å by 0.5 Å.

Supplementary Materials

The PDF file includes:

Figs. S1 to S10

Legends for movies S1 and S2

Other Supplementary Material for this manuscript includes the following:

Movies S1 and S2

REFERENCES AND NOTES

1. E. O. Freed, HIV-1 assembly, release and maturation. *Nat. Rev. Microbiol.* **13**, 484–496 (2015).
2. W. I. Sundquist, H. G. Krausslich, HIV-1 assembly, budding, and maturation. *Cold Spring Harb. Perspect. Med.* **2**, a006924 (2012).
3. B. K. Ganser-Pornillos, M. Yeager, O. Pornillos, Assembly and architecture of HIV. *Adv. Exp. Med. Biol.* **726**, 441–465 (2012).
4. J. A. Briggs, H. G. Kräusslich, The molecular architecture of HIV. *J. Mol. Biol.* **410**, 491–500 (2011).
5. J. Rasaiyaah, C. P. Tan, A. J. Fletcher, A. J. Price, C. Blondeau, L. Hilditch, D. A. Jacques, D. L. Selwood, L. C. James, M. Noursadeghi, G. J. Towers, HIV-1 evades innate immune recognition through specific cofactor recruitment. *Nature* **503**, 402–405 (2013).
6. V. Zila, E. Margiotta, B. Turoňová, T. G. Müller, C. E. Zimmerli, S. Mattei, M. Allegretti, K. Börner, J. Rada, B. Müller, Cone-shaped HIV-1 capsids are transported through intact nuclear pores. *Cell* **184**, 1032–1046.e18 (2021).
7. A. Dharan, N. Bachmann, S. Talley, V. Zwikelmaier, E. M. Campbell, Nuclear pore blockade reveals that HIV-1 completes reverse transcription and uncoating in the nucleus. *Nat. Microbiol.* **5**, 1088–1095 (2020).
8. A. Hudait, G. A. Voth, HIV-1 capsid shape, orientation, and entropic elasticity regulate translocation into the nuclear pore complex. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2313737121 (2024).

9. T. G. Müller, V. Zila, K. Peters, S. Schifferdecker, M. Stanic, B. Lucic, V. Laketa, M. Lusic, B. Müller, H.-G. Kräusslich, HIV-1 uncoating by release of viral cDNA from capsid-like structures in the nucleus of infected cells. *eLife* **10**, e64776 (2021).
10. R. C. Burdick, C. Li, M. Munshi, J. M. O. Rawson, K. Nagashima, W.-S. Hu, V. K. Pathak, HIV-1 uncoats in the nucleus near sites of integration. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 5486–5493 (2020).
11. C. Li, R. C. Burdick, K. Nagashima, W.-S. Hu, V. K. Pathak, HIV-1 cores retain their integrity until minutes before uncoating in the nucleus. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2019467118 (2021).
12. J. C. V. Stacey, A. Tan, J. M. Lu, L. C. James, R. A. Dick, J. A. G. Briggs, Two structural switches in HIV-1 capsid regulate capsid curvature and host factor binding. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2220557120 (2023).
13. T. Ni, S. Gerard, G. Zhao, K. Dent, J. Ning, J. Zhou, J. Shi, J. Anderson-Daniels, W. Li, S. Jang, A. N. Engelman, C. Aiken, P. Zhang, Intrinsic curvature of the HIV-1 CA hexamer underlies capsid topology and interaction with cyclophilin A. *Nat. Struct. Mol. Biol.* **27**, 855–862 (2020).
14. B. K. Ganser, S. Li, V. Y. Klishko, J. T. Finch, W. I. Sundquist, Assembly and analysis of conical models for the HIV-1 core. *Science* **283**, 80–83 (1999).
15. L. Deshmukh, C. D. Schwieters, A. Grishaev, R. Ghirlando, J. L. Baber, G. M. Clore, Structure and dynamics of full-length HIV-1 capsid protein in solution. *J. Am. Chem. Soc.* **135**, 16133–16147 (2013).
16. A. T. Gres, K. A. Kirby, V. N. KewalRamani, J. J. Tanner, O. Pornillos, S. G. Sarafianos, X-ray crystal structures of native HIV-1 capsid protein reveal conformational variability. *Science* **349**, 99–103 (2015).
17. R. A. Dick, D. L. Mallery, V. M. Vogt, L. C. James, IP6 regulation of HIV capsid assembly, stability, and uncoating. *Viruses* **10**, 640 (2018).
18. R. A. Dick, K. K. Zadrozny, C. Xu, F. K. M. Schur, T. D. Lyddon, C. L. Ricana, J. M. Wagner, J. R. Perilla, B. K. Ganser-Pornillos, M. C. Johnson, O. Pornillos, V. M. Vogt, Inositol phosphates are assembly co-factors for HIV-1. *Nature* **560**, 509–512 (2018).
19. O. Pornillos, B. K. Ganser-Pornillos, M. Yeager, Atomic-level modelling of the HIV capsid. *Nature* **469**, 424–427 (2011).
20. D. L. Mallery, K. R. Faysal, A. Kleinpeter, M. S. Wilson, M. Vaysburd, A. J. Fletcher, M. Novikova, T. Böcking, E. O. Freed, A. Saiardi, Cellular IP6 levels limit HIV production while viruses that cannot efficiently package IP6 are attenuated for infection and replication. *Cell Rep.* **29**, 3983–3996.e84 (2019).
21. S. Campbell, R. J. Fisher, E. M. Towler, S. Fox, H. J. Issaq, T. Wolfe, L. R. Phillips, A. Rein, Modulation of HIV-like particle assembly in vitro by inositol phosphates. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10875–10879 (2001).
22. D. A. Jacques, W. A. McEwan, L. Hilditch, A. J. Price, G. J. Towers, L. C. James, HIV-1 uses dynamic capsid pores to import nucleotides and fuel encapsidated DNA synthesis. *Nature* **536**, 349–353 (2016).
23. C. Xu, D. K. Fischer, S. Rankovic, W. Li, R. A. Dick, B. Runge, R. Zadorozhnyi, J. Ahn, C. Aiken, T. Polenova, A. N. Engelman, Z. Ambrose, I. Rousso, J. R. Perilla, Permeability of the HIV-1 capsid to metabolites modulates viral DNA synthesis. *PLOS Biol.* **18**, e3001015 (2020).
24. T. Ni, Y. Zhu, Z. Yang, C. Xu, Y. Chaban, T. Nesterova, J. Ning, T. Böcking, M. W. Parker, C. Monnie, J. Ahn, J. R. Perilla, P. Zhang, Structure of native HIV-1 cores and their interactions with IP6 and CypA. *Sci. Adv.* **7**, eabj5715 (2021).
25. M. Gupta, A. J. Pak, G. A. Voth, Critical mechanistic features of HIV-1 viral capsid assembly. *Sci. Adv.* **9**, eadd7434 (2023).
26. C. M. Highland, A. Tan, C. L. Ricaña, J. A. G. Briggs, R. A. Dick, Structural insights into HIV-1 polyanion-dependent capsid lattice formation revealed by single particle cryo-EM. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2220545120 (2023).
27. A. Yu, E. M. Y. Lee, J. Jin, G. A. Voth, Atomic-scale characterization of mature HIV-1 capsid stabilization by inositol hexakisphosphate (IP6). *Sci. Adv.* **6**, eabc6465 (2020).
28. R. T. Schirra, N. F. B. Dos Santos, K. K. Zadrozny, I. Kucharska, B. K. Ganser-Pornillos, O. Pornillos, A molecular switch modulates assembly and host factor binding of the HIV-1 capsid. *Nat. Struct. Mol. Biol.* **30**, 383–390 (2023).
29. O. Pornillos, B. K. Ganser-Pornillos, B. N. Kelly, Y. Hua, F. G. Whitby, C. D. Stout, W. I. Sundquist, C. P. Hill, M. Yeager, X-ray structures of the hexameric building block of the HIV capsid. *Cell* **137**, 1282–1292 (2009).
30. M. G. Mateu, The capsid protein of human immunodeficiency virus: Intersubunit interactions during virus assembly. *FEBS J.* **276**, 6098–6109 (2009).
31. A. Cooper, D. T. F. Dryden, Allostery without conformational change. *Eur. Biophys. J.* **11**, 103–109 (1984).
32. J. P. Changeux, Allostery and the Monod-Wyman-Changeux model after 50 years. *Annu. Rev. Biophys.* **41**, 103–133 (2012).
33. A. Barducci, G. Bussi, M. Parrinello, Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Phys. Rev. Lett.* **100**, 020603 (2008).
34. P. Tiwary, M. Parrinello, A time-independent free energy estimator for metadynamics. *J. Phys. Chem. B* **119**, 736–742 (2015).
35. J. F. Dama, M. Parrinello, G. A. Voth, Well-tempered metadynamics converges asymptotically. *Phys. Rev. Lett.* **112**, 240602 (2014).
36. A. Yu, E. M. Y. Lee, J. A. G. Briggs, B. K. Ganser-Pornillos, O. Pornillos, G. A. Voth, Strain and rupture of HIV-1 capsids during uncoating. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2117781119 (2022).
37. A. T. Gres, K. A. Kirby, W. M. McFadden, H. Du, D. Liu, C. Xu, A. J. Bryer, J. R. Perilla, J. Shi, C. Aiken, X. Fu, P. Zhang, A. C. Francis, G. B. Melikyan, S. G. Sarafianos, Multidisciplinary studies with mutated HIV-1 capsid proteins reveal structural mechanisms of lattice stabilization. *Nat. Commun.* **14**, 5614 (2023).
38. A. Bhattacharya, S. L. Alam, T. Fricke, K. Zadrozny, J. Sedzicki, A. B. Taylor, B. Demeler, O. Pornillos, B. K. Ganser-Pornillos, F. Diaz-Griffero, D. N. Ivanov, M. Yeager, Structural basis of HIV-1 capsid recognition by PF74 and CPSF6. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 18625–18630 (2014).
39. L. Lamorte, S. Titolo, C. T. Lemke, N. Goudreau, J.-F. Mercier, E. Wardrop, V. B. Shah, U. K. von Schwedler, C. Langelier, S. S. R. Banik, C. Aiken, W. I. Sundquist, S. W. Mason, Discovery of novel small-molecule HIV-1 replication inhibitors that stabilize capsid complexes. *Antimicrob. Agents Chemother.* **57**, 4622–4631 (2013).
40. O. Valsson, P. Tiwary, M. Parrinello, Enhancing important fluctuations: Rare events and metadynamics from a conceptual viewpoint. *Annu. Rev. Phys. Chem.* **67**, 159–184 (2016).
41. A. R. Dinner, A. Sali, L. J. Smith, C. M. Dobson, M. Karplus, Understanding protein folding via free-energy surfaces from theory and experiment. *Trends Biochem. Sci.* **25**, 331–339 (2000).
42. C. Wagner, T. Kiefhaber, Intermediates can accelerate protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 6716–6721 (1999).
43. S. Li, C. P. Hill, W. I. Sundquist, J. T. Finch, Image reconstructions of helical assemblies of the HIV-1 CA protein. *Nature* **407**, 409–413 (2000).
44. G. Zhao, J. R. Perilla, E. L. Yufenyuy, X. Meng, B. Chen, J. Ning, J. Ahn, A. M. Gronenborn, K. Schulten, C. Aiken, P. Zhang, Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature* **497**, 643–646 (2013).
45. N. Renner, D. L. Mallery, K. M. R. Faysal, W. Peng, D. A. Jacques, T. Böcking, L. C. James, A lysine ring in HIV capsid pores coordinates IP6 to drive mature capsid assembly. *PLOS Pathog.* **17**, e1009164 (2021).
46. A. Kleinpeter, D. L. Mallery, N. Renner, A. Albecka, J. O. Klarhof, E. O. Freed, L. C. James, HIV-1 adapts to lost IP6 coordination through second-site mutations that restore conical capsid assembly. *Nat. Commun.* **15**, 8017 (2024).
47. R. T. Schirra, N. F. Dos Santos, B. K. Ganser-Pornillos, O. Pornillos, Arg18 substitutions reveal the capacity of the HIV-1 capsid protein for non-fullerene assembly. *Viruses* **16**, 1038 (2024).
48. M. F. Perutz, Stereochemistry of cooperative effects in haemoglobin: Haem–Haem interaction and the problem of allostery. *Nature* **228**, 726–734 (1970).
49. S. E. Reichheld, Z. Yu, A. R. Davidson, The induction of folding cooperativity by ligand binding drives the allosteric response of tetracycline repressor. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 22263–22268 (2009).
50. S. Mattei, B. Glass, W. J. Hagen, H.-G. Kräusslich, J. A. Briggs, The structure and flexibility of conical HIV-1 capsids determined within intact virions. *Science* **354**, 1434–1437 (2016).
51. I. J. Byeon, X. Meng, J. Jung, G. Zhao, R. Yang, J. Ahn, J. Shi, J. Concel, C. Aiken, P. Zhang, A. M. Gronenborn, Structural convergence between Cryo-EM and NMR reveals intersubunit interactions critical for HIV-1 capsid function. *Cell* **139**, 780–790 (2009).
52. G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
53. D. J. Evans, B. L. Holian, The Nose–Hoover thermostat. *J. Chem. Phys.* **83**, 4069–4074 (1985).
54. M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
55. J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. de Groot, H. Grubmüller, A. D. MacKerell, CHARMM36m: An improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71–73 (2017).
56. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
57. K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, A. D. Mackerell Jr., CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **31**, 671–690 (2010).
58. T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
59. W. C. Swope, H. C. Andersen, P. H. Berens, K. R. Wilson, A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *J. Chem. Phys.* **76**, 637–649 (1982).
60. M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19–25 (2015).
61. G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, G. Bussi, PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **185**, 604–613 (2014).

62. N. Michaud-Agrawal, E. J. Denning, T. B. Woolf, O. Beckstein, MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **32**, 2319–2327 (2011).
63. D. Stanzione, J. West, R. T. Evans, T. Minyard, O. Ghattas, D. K. Panda, "Frontera: The evolution of leadership computing at the National Science Foundation," in *Practice and Experience in Advanced Research Computing 2020: Catch the Wave* (Association for Computing Machinery, 2020), pp. 106–111.

Acknowledgments: We thank Y. Wang for helpful discussions during simulation setup. **Funding:** This research was supported in part by the National Institute of Allergy and Infectious Diseases (NIAID) of the National Institutes of Health (NIH) grant R01 AI178850 (to G.A.V.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Partial support was also received from the Frost Institute for Chemistry and Molecular Science at the University of Miami (to M.Y.). Simulations were performed using resources by the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support program, which is supported by NSF grant numbers 2138259, 2138286, 2138307, 2137603, and 2138296, and Frontera (*63*) (at TACC) funded by NSF OAC-1818253 (G.A.V. allocations). **Author contributions:** Conceptualization: all authors. Investigation: M.G. and A.H. Supervision: G.A.V. Writing—original draft: G.A.V., M.G., and A.H. Writing—review and editing: all authors. Methodology: M.G. and A.H. Resources: G.A.V., A.H., and M.G. Data curation: G.A.V., A.H., and M.G. Validation: G.A.V., A.H., and M.G. Formal analysis: G.A.V., A.H., and M.G. Software: A.H. and M.G. Project administration: G.A.V. and M.G. Visualization: G.A.V., A.H., and M.G. Funding acquisition: G.A.V. and M.Y. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Simulation input files and models are available in Zenodo (https://doi.org/10.5281/zenodo.15010373).

Submitted 10 October 2024
Accepted 12 March 2025
Published 16 April 2025
10.1126/sciadv.adt7818