

Structural basis for HIV-1 capsid adaption to rescue IP6-packaging deficiency

Yanan Zhu^{1,2†}, Alex B Kleinpeter^{3†}, Juan S. Rey^{4†}, Juan Shen¹, Yao Shen¹, Jialu Xu¹, Nathan Hardenbrook¹, Long Chen¹, Anka Lucic¹, Juan R. Perilla^{4*}, Eric O. Freed^{3*}, Peijun Zhang^{1,5,6*}

¹Division of Structural Biology, Wellcome Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK

²Institute for Advanced Study in Physics, Zhejiang University, Hangzhou, Zhejiang 310027, China

³Virus-Cell Interaction Section, HIV Dynamics and Replication Program, Center for Cancer Research, National Cancer Institute, Frederick, MD 21702-1201, USA

⁴Department of Chemistry and Biochemistry, University of Delaware, Newark, DE 19716, USA

⁵Diamond Light Source, Harwell Science and Innovation Campus, Didcot OX11 0DE, UK

⁶Chinese Academy of Medical Sciences Oxford Institute, University of Oxford, Oxford OX3 7BN, UK

[†]These authors contributed equally

*Corresponding authors: peijun.zhang@strubi.ox.ac.uk; efreed@nih.gov; jperilla@udel.edu

Abstract

Inositol hexakisphosphate (IP6) promotes HIV-1 assembly via its interaction with the immature Gag lattice, effectively enriching IP6 within virions. During particle maturation, the HIV-1 protease cleaves the Gag polyproteins comprising the immature Gag lattice, releasing IP6 from its original binding site and liberating the capsid (CA) domain of Gag. IP6 then promotes the assembly of mature CA protein into the capsid shell of the viral core, which is required for infection of new target cells. Recently, we reported HIV-1 Gag mutants that assemble virions independently of IP6. However, these mutants are non-infectious and unable to assemble stable capsids. Here, we identified a mutation in the C-terminus of CA – G225R – that restores capsid formation and infectivity to these IP6-packaging-deficient mutants. Furthermore, we show that G225R facilitates the *in vitro* assembly of purified CA into capsid-like particles (CLPs) at IP6 concentrations well below those required for WT CLP assembly. Using single-particle cryoEM, we solved structures of CA hexamer and hexameric lattice of mature CLPs harbouring the G225R mutation assembled in low-IP6 conditions. The high-resolution (2.7 Å) cryoEM structure combined with molecular dynamics simulations of the G225R capsid revealed that the otherwise flexible and disordered C-terminus of CA becomes structured, extending to the pseudo two-fold hexamer-hexamer interface, thereby stabilizing the mature capsid. This work uncovers a structural mechanism by which HIV-1 adapts to a deficiency in IP6 packaging. Furthermore, the ability of G225R to promote mature capsid assembly in low-IP6 conditions provides a valuable tool for capsid-related studies and may indicate a heretofore unknown role for the unstructured C-terminus in HIV-1 capsid assembly.

Introduction

Nascent HIV-1 virions assemble at the plasma membrane of infected cells, packaging both viral and host-cell components required for the next round of infection. HIV-1 virion assembly is driven by the multimerization of the Gag precursor, a 55-kDa polyprotein (Pr55Gag) comprising matrix (MA), capsid (CA), nucleocapsid (NC), and p6 domains and two spacer peptides, SP1 and SP2, which connect CA-NC and NC-p6, respectively¹. Gag oligomerizes into a spherical lattice of Gag hexamers – the immature Gag lattice – ultimately assembling into nascent particles that undergo membrane scission and release. Viral enzymes are recruited into growing particles via co-assembly of Gag with the GagPol polyprotein precursor, which is synthesized at a GagPol:Gag ratio of ~1:20. GagPol contains the protease (PR), reverse transcriptase (RT), and integrase (IN) domains that are subsequently required for particle infectivity. Concomitant with their release from the infected cell, virions undergo maturation during which immature particles are morphologically transformed into mature virions containing a cone-shaped viral core. Maturation is initiated when PR cleaves specific sites in Gag and GagPol, liberating their constituent domains. CA then assembles into the mature capsid, a conical structure that encloses the viral RNA genome and enzymes RT and IN and serves as the protective shell of the viral core. The capsid is composed of 200-250 CA hexamers and exactly 12 CA pentamers, which are concentrated at each end to ensure capsid closure². In the next round of infection, the core is deposited into the cytosol during virus entry and remains intact, ultimately passing through a nuclear pore and uncoating near sites of integration in the nucleus. Importantly, the assembly of a stable capsid during maturation, and its stability post-entry, are crucial for reverse transcription of the HIV-1 genome, protection of the viral genome from innate immune sensing, trafficking of the viral core to the nuclear envelope, nuclear entry of the core through nuclear pores, and efficient integration of the newly synthesized viral DNA into gene-rich regions of the host genome³.

The polyanionic host-cell metabolite inositol hexakisphosphate (IP6) is a key co-factor for HIV-1 capsid assembly and stability^{4,5}. IP6 is coordinated by two positively charged amino acids in the CA N-terminal domain (CA-NTD) – Arg-18 (R18) and Lys-25 (K25) – that form electropositive rings within a pore at the centre of CA hexamers and pentamers^{4,6,7}. Mutation of R18 or K25, each of which can bind a molecule of IP6, drastically reduces capsid formation in virions and particle infectivity and prevents the efficient *in vitro* assembly of mature CA into capsid-like particles (CLPs) in the presence of IP6^{4,6,8-12}. Several studies suggest that IP6 binding to R18 and K25 rings is particularly crucial for the formation of CA

pentamers, in which these rings are more tightly packed than in hexamers, and therefore more prone to destabilization via charge repulsion in the absence of IP6^{7,10,11}. Because capsid formation occurs after particle release from a virus-producing cell and prior to fusion with a target cell, CA must rely on IP6 acquired from the virus-producing cell during particle assembly.

HIV-1 recruits IP6 during immature Gag lattice assembly via two Lys residues in the C-terminal domain of CA – Lys-158 (K158) and Lys-227 (K227) – which form positively charged rings at the centre of immature Gag hexamers^{4,13,14}. The final eight amino acids of CA (including K227) and SP1 participate in key intra-hexamer interactions stabilizing Gag hexamers and the immature Gag lattice by folding into a six-helix bundle (6HB)¹⁴⁻¹⁶. Mutations that destabilize the CA-SP1 6HB result in drastic defects in virus particle assembly, morphogenesis, and infectivity^{17,18}. Conversely, because the CA-SP1 cleavage site is located inside the 6HB, mutations that hyper-stabilize the 6HB, such as SP1-T8I, obstruct PR access to this cleavage site, preventing CA-SP1 processing and reducing virion infectivity¹⁷⁻¹⁹. The stability of the CA-SP1 6HB is thus finely tuned to allow both efficient assembly and subsequent maturation.

IP6 plays a crucial role in 6HB stability by neutralizing the highly charged pore at the top of the 6HB formed by the K158 and K227 rings. Depletion of IP6 in virus-producing cells, or mutation of either K158 or K227, results in assembly defects caused by destabilization of the 6HB and reduced IP6 incorporation into viral particles^{13,18,20}. This reliance on IP6 for efficient particle assembly ensures that IP6 is enriched in virions to promote the formation of the capsid during maturation. Our recent study showed that viruses harbouring mutations at both Lys rings (K158A/K227A or ‘KAKA’) in the immature Gag lattice could assemble independently of IP6²¹. Structural analysis revealed that KAKA virions contained an immature lattice that was largely indistinguishable from that in immature wild-type (WT) virions, apart from the absence of IP6 density^{20,21}. However, KAKA virions were not infectious due to an inability to build stable capsids during particle maturation²¹. The above findings support a model in which the requirement for IP6 during particle assembly is a mechanism to enrich IP6 in virions to promote capsid assembly and subsequent infection.

Here we utilize KAKA and KAKA/T8I mutant viruses, which fail to bind and recruit IP6 during assembly, to further delineate the separate but interdependent roles of IP6 in particle

assembly and maturation. After characterizing the effect of mutations in specific IP6-binding residues in authentic immature particles, we propagated these mutants to assess how HIV-1 could adapt to an inability to enrich particles with IP6. Propagation of KAKA and KAKA/T8I mutants led to the acquisition of a compensatory mutation, G225R, in the CA C-terminus that rescues KAKA/T8I conical capsid formation and particle infectivity without restoring IP6 enrichment during assembly. We determined the cryoEM structure of CA-KAKA/G225R CLPs at a 2.7 Å resolution, which revealed the structural mechanism by which G225R promotes mature HIV-1 capsid assembly in virions not enriched with IP6. Remarkably, we show that the G225R mutation results in an additional intermolecular interaction via the CA C-terminus, which is otherwise flexible and has thus far been refractory to structural analysis. We further show that the G225R mutation promotes the *in vitro* assembly of CLPs in low-IP6 conditions. Our results suggest that the C-terminus of CA may play a heretofore unappreciated role in the assembly of the HIV-1 capsid and the modulation of capsid stability.

Results

CA-G225R rescues the replication defect imposed by IP6-packaging deficiency

In previous studies, we investigated the effect of mutations, namely KAKA, in the IP6 binding pocket of immature Gag hexamers in the context of the WT and 6HB-stabilizing SP1 mutations T8I and M4L/T8I^{20,21}. We found that immature KAKA VLPs were composed of Gag hexamers that were largely indistinguishable from those formed by WT Gag, apart from the absence of IP6 density at the top of the 6HB. To dissect further the contributions of individual single or dual mutations on the immature Gag lattice, we analysed the structures of immature K227A and KAKA/T8I VLPs produced from 293T cells and compared them to the previously reported structures of WT and KAKA immature VLPs (Figure 1). Both K227A and KAKA/T8I assemble into immature spherical VLPs similar to those formed by WT and KAKA (Figure 1A-D). Our cryoET subtomogram averaging (STA) structure of K227A at 3.82 Å resolution showed an intermediate level of IP6 compared with the WT (full) and the KAKA mutant (empty) (Figure 1E-G, Supplementary Figure 1A, Supplementary Table 1), consistent with the previous biochemical measurements²¹. The structure of KAKA/T8I at 3.67 Å resolution displayed no IP6 density (Figure 1H, Supplementary Figure 1B, Supplementary Table 1), similar to the KAKA mutant, suggesting that the 6HB-stabilizing mutation SP1-T8I does not restore IP6 enrichment to KAKA. Consistent with this inability to restore IP6 enrichment, the addition of SP1-T8I does not substantially reverse the infectivity defect imposed by the KAKA mutations²¹ (Figure 1I).

Next, we sought to determine how HIV-1 might adapt to an inability to enrich IP6 during particle assembly. We propagated KAKA and KAKA/T8I in the highly permissive MT-4 T cell line. Briefly, MT-4 cells were initially transfected with the WT NL4-3 infectious molecular clone or derivatives harbouring the KAKA or KAKA/T8I mutations. Consistent with the drastic decrease in KAKA and KAKA/T8I infectivity, each mutant demonstrated a significant delay in replication compared to WT virus (Figure 1J). Upon propagation of virus collected from the peak of replication, both KAKA and KAKA/T8I replicated with a substantially reduced delay relative to WT (Figure 1K). This observation is suggestive of the acquisition of mutations that compensate for the defects induced by the KAKA and KAKA/T8I mutations. To confirm the acquisition of compensatory mutations, we isolated genomic DNA from infected cells at the peak of replication and PCR amplified the Gag coding region. Gag amplicons were sequenced, and potential compensatory mutations were cloned into the KAKA and KAKA/T8I NL4-3 molecular clones. This process was performed iteratively using mutant viruses selected in initial experiments to initiate downstream propagation experiments. We also performed these *in vitro* selection experiments in the C8166 T cell line and with additional mutant viruses containing a K158T mutation in place of K158A (KTKA) and/or an SP1-M4L mutation in place of T8I. Compensatory mutations selected in these experiments were found throughout CA and SP1 (Supplementary Table 2). Interestingly, several selected mutations were identified at amino acid positions associated with resistance to capsid inhibitors (e.g. CA-Thr-107, His-87, Ala-105, Thr-58, Gly-208, Thr-216)²²⁻²⁵, altered capsid assembly and stability (e.g. CA-H12Y, N21S, T216I)^{11,24,26}, and host factor dependence (e.g. CA-Ala-77, His-87, Gly-94D, Gly-208, Thr-210, Glu-187, Pro-207)²⁷⁻³⁰. We also observed mutations at amino acid positions located at critical interfaces within the mature CA lattice including the β -hairpin (Val-11 and His-12), the central pore (Asn-21 and Ala-22), the Thr-Val-Gly-Gly (TVGG) motif that controls CA hexamer/pentamer assembly (Thr-58, Gly-61), the NTD/CTD interface between adjacent CA protomers (Met-68, Ala-105, Thr-107), and the trimer interface between neighbouring CA hexamers/pentamers (Pro-207, Gly-208, Thr-210, Thr-216).

To determine whether the mutations selected in these experiments could restore fitness to KAKA and KAKA/T8I, we quantified the effect of these mutations on the single-cycle infectivity of KAKA and KAKA/T8I. Most of the mutations that we observed were unable to

significantly restore infectivity to KAKA or KAKA/T8I (Supplementary Figure 2), including two mutations – N21S and T216I – that were recently reported to restore infectivity to K25A, a mutant unable to bind IP6 in the context of the mature CA lattice¹¹. However, we did observe a mutation at position 225 in CA that significantly restored infectivity to KAKA/T8I. Initially, we observed a Gly-to-Ser substitution at residue 225 upon propagation of KAKA in MT-4 cells. After introducing the G225S mutation into multiple KAKA- and KTKA-containing clones, we observed a Ser-to-Arg substitution at CA position 225 (G225R) upon propagation of KTKA/T8I/G225S. We then introduced the G225R mutation into KAKA and KAKA/T8I to determine its effect on particle infectivity. We found that G225R conferred an ~8-fold increase in single-cycle infectivity to KAKA/T8I (from ~2% of WT to 16%) (Figure 1I). Further propagation of KAKA/T8I/G225R resulted in the acquisition of an additional mutation, A77V, that doubled the infectivity of KAKA/T8I/G225R (Figure 1I). Curiously, we also found that G225R conferred an opposing phenotype to KAKA, decreasing its infectivity to nearly undetectable levels (Figure 1I). Thus, T8I, which stabilizes the immature Gag lattice, restores KAKA/G225R single-cycle infectivity by ~100 fold. Furthermore, the G225R mutation alone imposes an infectivity defect comparable to that of KAKA/T8I/G225R, demonstrating that G225R is insensitive to the defect imposed by KAKA/T8I (Figure 1I).

G225R does not restore IP6 enrichment to the KAKA or KAKA/T8I mutants

We next sought to determine the mechanism by which G225R restores infectivity to KAKA/T8I. Because G225 is close to the mutated IP6-binding residue K227 at the top of the 6HB, the change of the Gly to a positively charged Arg at this position could introduce a new positively charged ring capable of restoring IP6 recruitment during particle assembly, thereby rescuing particle infectivity. To investigate this possibility, we produced immature KAKA/G225R, KAKA/T8I/G225R, A77V/KAKA/T8I/G225R, and G225R VLPs and solved their Gag hexamer structures by cryoET with STA (Figure 2A-D, Supplementary Figure 1, Supplementary Table 1). The structure of KAKA/T8I/G225R at 4.05 Å resolution revealed no IP6 density atop the 6HB (Figure 2A, Supplementary Figure 1C, Supplementary Table 1). The bulky Arg sidechain, which is clearly resolved in the density map, points away from the central channel of the 6HB, unlike K227 in WT Gag, and is therefore not available for IP6 coordination (Figure 2A). Indeed, no IP6 density was observed in cryoET STA maps in any of the KAKA-containing Gag mutants analysed, including KAKA/G225R and A77V/KAKA/T8I/G225R (Figure 2B-C, Supplementary Figure 1D-E, Supplementary Table 1). In contrast, the G225R Gag hexamer structure showed a clear IP6 density atop the 6HB

(Figure 2D, Supplementary Figure 1F, Supplementary Table 1). Each of the Gag mutants evaluated is capable of assembling an immature Gag lattice similar to that present in WT immature VLPs (Supplementary Figure 3). These structural data suggest that the recovery of KAKA/T8I infectivity upon introduction of the compensatory mutation G225R cannot be attributed to increased IP6 packaging.

G225R does not affect the production of KAKA/T8I virions from IP6-depleted cells

To confirm that the KAKA-containing mutants rescued by G225R remain IP6-independent during particle assembly, we conducted experiments to investigate the effect of IP6 depletion in virus-producing cells on virus production efficiency and particle infectivity. Briefly, we produced virus in HEK 293T knockout (KO) cell lines lacking either inositol polyphosphate multikinase (IPMK), which phosphorylates IP3 and IP4 to generate IP4 and IP5, or inositol pentakisphosphate 2-kinase (IPPK), which phosphorylates IP5 to generate IP6¹³. To further decrease IP6 levels in IPMK or IPPK KO cells, we simultaneously overexpressed multiple inositol-polyphosphate 1 (MINPP1), which removes phosphates from IP6, IP5, and IP4. After transfecting NL4-3 infectious molecular clones, we measured the efficiency of virus production in these IP6-depleted cells by quantifying the levels of p24 (CA) present in the supernatant relative to total Gag present in the cells and supernatant. Consistent with previous reports^{20,21,31}, we observed a significant reduction in the efficiency of WT virus production from IPMK KO and IPPK KO HEK 293T cells overexpressing MINPP1 relative to parental 293T cells co-transfected with an empty vector (Figure 2E, Supplementary Figure 4). In agreement with our structural data, G225R exhibited an IP6-dependent assembly phenotype similar to WT. In contrast, virus production efficiency of all KAKA-containing mutants was unaffected by producer-cell IP6 depletion, confirming that G225R does not rescue KAKA/T8I infectivity by restoring IP6 enrichment. Again, these data are in agreement with our cryo-ET/STA results and demonstrate that irrespective of the presence of G225R, KAKA mutants assemble in an IP6-independent manner.

We then investigated the effects of producer-cell IP6 depletion on particle infectivity. Consistent with previous results^{20,21,31} producer-cell IP6 depletion drastically reduces WT particle infectivity, while KAKA and KAKA/T8I infectivity is more mildly reduced (Figure 2F). Our assay was unable to detect KAKA/G225R infectivity under any condition. The rescued mutants KAKA/T8I/G225R and A77V/KAKA/T8I/G225R also showed only mild reductions in infectivity when produced from IP6-depleted cells. However,

KAKA/T8I/G225R virions produced from IP6-depleted cells remained ~10-fold more infectious than KAKA/T8I virions produced under the same conditions (Figure 2F). Since neither mutant is able to enrich IP6 into virions, this suggests that G225R may rescue KAKA/T8I infectivity by mitigating the consequences of reduced IP6 during particle maturation, perhaps by facilitating capsid formation at low IP6 concentrations.

G225R increases the efficiency of KAKA/T8I capsid formation in virions and *in vitro*

To test the hypothesis that G225R facilitates capsid formation in the absence of enriched IP6, we next examined the effects of G225R on the morphology of mature, Env(-) viral particles by cryoET (Figure 2G-H). CryoET reconstructions revealed a variety of distinct morphologies of mature WT HIV-1 particles produced from HEK 293T cells, which we categorized into Mature, Mature (eccentric), Mature (abnormal), and Immature (Figure 2G). All the mutants tested displayed reduced numbers of viral particles containing normal mature conical cores relative to the WT. The revertant mutants KAKA/T8I/G225R and A77V/KAKA/T8I/G225R exhibited higher numbers of viral particles containing normal mature conical cores than KAKA/T8I (Figure 2H). The occurrence of normal mature cores (Figure 2H, light blue) correlates well with the observed infectivity measurements for these mutants (Figure 1I). These analyses demonstrate that G225R increases the formation of normal mature cores in the absence of IP6 enrichment into viral particles.

The findings that the G225R mutation improves mature particle formation without IP6 enrichment led us to hypothesize that this mutation may increase the efficiency of mature capsid assembly under low-IP6 conditions. To test this, we expressed and purified recombinant CA protein carrying KAKA and G225R mutations and conducted *in vitro* assembly assays at IP6 concentrations ranging from 0 μ M to 5mM (Figure 3A-D). Assembly properties of each mutant were evaluated by determining the amount of pelletable CA in each reaction. While WT and KAKA CA require high IP6 concentrations to assemble (Figure 3A-B), the CA proteins carrying the G225R mutation assemble at low IP6 concentrations (Figure 3C-D). Negative stain transmission EM images show that the KAKA/G225R CA starts forming tubes and cones at an IP6 concentration as low as 5 μ M and the G225R CA at 50 μ M, whereas no such structures were observed with WT and KAKA CA at 150 μ M IP6 (Figure 3E). Further statistical analyses on the number of assembled tubes and cones for each variant indicate that KAKA/G225R and G225R CA form significantly more tubes and cones than

WT and KAKA CA at low IP6 concentrations (Figure 3F). The data suggest that the G225R mutation reduces IP6 dependency for mature capsid assembly.

CryoEM structure of KAKA/G225R CLP reveals previously unresolved CA C-terminus stabilizing the dimer interface

To understand the molecular mechanism by which the G225R mutation promotes the assembly of stable mature capsids under low-IP6 conditions, we prepared KAKA/G225R CLPs in the presence of 100 μ M IP6 to optimize the sample for cryoEM structural analysis, noting that WT CA shows no assembly at this IP6 concentration (Figure 3E-F). The *in vitro* assembled KAKA/G225R CLPs are well dispersed and highly ordered (Figure 4A), from which we determined the structure of KAKA/G225R CA hexamer at 2.7 Å resolution and built an atomic model (PDB 9I8I) using single-particle cryoEM (Figure 4B-C, Supplementary Figure 5, Supplementary Table 3).

The structure reveals two IP6 densities in the hexamer centre, coordinated by R18 at the top density and K25 at the bottom density (Figure 4D), consistent with previous studies^{4,6,7,9,32,33}. It is worth noting that the IP6 densities are rather weak compared to those in WT CA CLPs assembled with 2.5-5 mM IP6^{9,32,33} and in native mature cores within perforated virions with 1mM IP6⁷, suggesting a low IP6 occupancy in KAKA/G225R CLPs. This is expected, as our KAKA/G225R CLPs were assembled in the presence of 25- to 50-fold lower IP6 concentrations than WT CLPs. In the hexamer structure, the β -hairpin exhibits an open conformation similar to those observed in previous cryoEM structures derived from CLPs as well as the crystal structure from the non-pandemic strain HIV-1 (O) at pH 6, distinct from the closed conformation in cryoEM structures determined from tubular assemblies and the crystal structure of the WT CA hexamer at pH 7^{8,9,32-35} (Figure 4E-F). The reduced pH in the CLP assemblies was suggested to affect the β -hairpin conformation^{4,6,8-12}. The NTD-CTD interface around the binding pocket for Lenacapavir and FG-motif-containing host factors appears preserved (Figure 4G-H). The linker that connects CA-NTD to CA-CTD resembles that of the WT CA (Figure 4I). The CypA binding loop is very well ordered (Figure 4I), in contrast to the majority of CA hexamer structures determined previously. Intriguingly, the density of KAKA/G225R CA C-terminus extends to G223 which can be unambiguously modelled (Figure 4I, inset), along with additional densities further extending to the inter-hexamer interfaces but not well-resolved. Interestingly, this C-terminal extension adopts the same configuration as our previous CA-CTD NMR solution structure (PDB 2KOD, the 2nd

conformer) (Figure 4I inset, grey)³⁶. These results suggest that the otherwise flexible C-terminus could adopt a structured conformation to mediated intermolecular interactions, thus strengthening the capsid stability when required.

To further clarify the inter-hexamer densities, cryoEM density maps of the tri-hexamer interface were obtained (Figure 5A). The dimer and trimer interfaces of KAKA/G225R are very similar to those of WT (PDB 8G6M), with an RMSD of 0.27 Å and 0.44 Å, respectively (Figure 5B-C). A further 3D classification and refinement resulted in two major classes, one without extra density (Figure 5D) and another with an extra density extending from the end of H11, reaching below the dimer interface (Figure 5E, red density). This density most likely corresponds to the previously unresolved CA C-terminal fragment (220-231), which was too flexible to be resolved in all previous crystal or cryoEM structures.

Molecular dynamics simulation of interactions involving the CA C-terminal segment

To explore the interactions between the C-terminal fragment with the neighbouring CA monomers, we used Rosetta³⁷⁻³⁹ to derive density-guided models of a KAKA/G225R CA trimer of dimers with an extended C-terminal. Furthermore, we perform molecular dynamics simulations of the KAKA/G225R CA trimer of dimers, as well as a KAKA CA and WT CA trimer of dimers, to assess the stability of the C-terminal fragment in the CA dimer interface. Through 400ns of unrestrained equilibration MD simulations, we observed that the C-terminal region (residues 220 to 232) is highly flexible and transiently occupies the CA dimer interface. In the KAKA/G225R CA trimer of dimers simulation, we measure occupancies of 62.5%, 30.6% and 1.7% at each dimer interface (avg. occupancy of 32.6%), consistent with the multiple 3D classes observed in the cryoEM analysis. In contrast, for KAKA CA and WT CA, the dimer interface occupancies were notably lower, with maximum occupancies of 12.9% and 22.1% and average occupancies of 6.8% and 9.8%, respectively (Supplementary Figure 6).

Additionally, we analysed the residence times of C-terminal segment contacts in all simulations. While C-terminal interactions at the dimer interface generally have short residence times (<5ns), the residence time distribution for KAKA/G225R CA is broader, with several interactions lasting over 20ns and one event lasting up to 100ns (Supplementary Figure 6).

We then performed a residue-level contact analysis of the C-terminal segment and CA dimer interactions, to identify the specific interactions that stabilize the C-terminal tail at the dimer interface. Notably, we observed that residues L231 (the C-terminal residue), R229 and G225R form salt bridges with high occupancies with residues K199, R154 and D152 at the dimer interface (Figure 5F, Supplementary Table 4). These interactions increase the stability of conformations where the C-terminal segment resides at the dimer interface. In contrast, interactions with those residues have significantly reduced occupancies in the KAKA CA and WT CA simulations (Supplementary Table 4).

Thus, the extra density observed in the cryoEM map of KAKA/G225R CA tri-hexamer likely stems from the transient interactions of the C-terminal segment at the dimer interface, which are facilitated by the added charge of the G225R mutation on the C-terminal segment. The presence of the C-terminal segment at the dimer interface, along with the formation of salt bridge interactions therein might explain why the KAKA/G225R CA mutation stabilizes the mature capsid even with minimal IP6 present.

KAKA/G225R CA forms compact assembly interfaces

To further evaluate the effect of the KAKA/G225R CA mutation in capsid interfaces, we extracted all possible capsomer-capsomer interfaces (hexamer-hexamer-hexamer, hexamer-hexamer-pentamer and hexamer-pentamer-pentamer) from a full capsid cone⁷ and compared them with the KAKA/G225R CA trimer of dimers assembly (Figure 6). For each capsomer-capsomer dimer interface, we measured the centre of mass distances and tilt angles between alpha helices 9 and 10 and the 3-10 helix. Compared to all WT dimer interfaces, the helices 9 and 10 in KAKA/G225R CA are packed closer, resembling the distance for pentamer-pentamer-hexamer interfaces. However, these helices are oriented at a more acute angle (for helix 9), or closer to a right angle (for helix 10). In conclusion, the KAKA/G225R CA dimer interfaces are more tightly packed than those measured in a WT CA cone, likely due to the interactions with the C-terminal flexible tail described above.

Infectivity of the KAKA and G225R mutants is insensitive to target-cell IP6 depletion

We further investigated whether an inability to enrich IP6 into viral particles results in sensitivity to depletion of IP6 in target cells, and whether G225R could reverse this sensitivity. Consistent with previous reports⁴⁰, we found that the infectivity of the hypostable CA mutant P38A was severely reduced in HEK 293T IPPK KO target cells relative to its infectivity in

parental HEK 293T cells (Supplementary Figure 7). In contrast, WT and KAKA/T8I infectivity was unaffected in HEK 293T IPPK KO cells relative to parental cells, and the addition of G225R to KAKA/T8I did not alter this phenotype. We did observe a statistically significant decrease in G225R infectivity in HEK 293T IPPK KO cells relative to parental cells. However, the effect size was much smaller than that of P38A. These data demonstrate that the addition of G225R to KAKA/T8I did not eliminate insensitivity to target cell IP6 depletion.

Discussion

Here we investigated how HIV-1 adapts to an inability to specifically package and enrich IP6 into particles during virus assembly. Propagation of the IP6-binding-deficient mutants KAKA and KAKA/T8I resulted in the acquisition of a mutation in the C-terminus of CA – G225R – that rescued KAKA/T8I particle infectivity by increasing the efficiency of capsid formation in virions. *In vitro* assembly assays revealed that CA protein harbouring the G225R mutation assembled into CA tubes and cones at substantially lower IP6 concentrations than required for WT assembly. These findings suggest that G225R facilitates capsid assembly in virions containing limited IP6, explaining the ability of G225R to restore KAKA/T8I capsid formation and infectivity. Importantly, there was no indication of CA assembly in the absence of IP6 (Figure S3A-D), suggesting that G225R does not confer fully IP6-independent assembly to CA. Some IP6 is still required to assemble mature tubes and cones. In a biological context, virions unable to actively incorporate IP6 during assembly, such as KAKA/T8I, likely have access to low levels of passively incorporated IP6. Based on our structural, morphological, and *in vitro* assembly data, we propose that G225R restores KAKA/T8I capsid assembly and infectivity by drawing upon the small pool of passively incorporated IP6 available during maturation, considering that the cellular IP6 concentration is around 50-100 μM , which is sufficient to assemble KAKA/G225R capsids. To address the mechanism by which G225R increases the efficiency of capsid assembly, we solved a tri-hexamer structure of *in vitro* assembled KAKA/G225R CLPs. We observed a strip of electron density corresponding to the CA C-terminus extending from the end of helix 11 at the trimer interface between hexamers into the dimer interface between hexamers. Molecular dynamics analysis revealed that the KAKA/G225R mutations promotes the formation of transient salt bridges between residues in the C-terminal segment (residues 220 to 231) and residues in helix 10 and the 3-10 helix at the dimer interface, in particular G225R and R229 with D152.

These interactions increase the occupancy of the C-terminal at the dimer interface and contribute to forming a more compact interface. These results demonstrate that the unstructured and flexible CA C-terminus is capable of altering the properties of mature CA assembly and may indicate a role for the CA C-terminus during HIV-1 particle maturation.

The stability of the HIV-1 capsid is critical for several post-entry functions including reverse transcription, nuclear entry, and interactions with host factors. Thus, the assembly of a stable capsid during maturation is required to set the stage for these key steps in the HIV-1 replication cycle. IP6, which is packaged into virions via its interaction with the Gag hexamers that compose the immature Gag lattice, plays a significant role in mature capsid assembly and the maintenance of capsid stability. Our observation of a mutation in the CA C-terminus that increases the efficiency of capsid formation in IP6-packaging-deficient virions and CLP assembly *in vitro* at low IP6 concentrations suggests that the flexible C-terminus may allosterically regulate capsid assembly and stability.

IP6 has also been proposed to play a critical role in CA pentamer formation, a requirement for the assembly of a closed conical capsid^{7,9-11,18,32,33,41}. It is possible that mutations in the C-terminus of CA may promote pentamer formation under limiting IP6 conditions. The mutation T216I, which lies in helix 11 near the unstructured C-terminus, was recently proposed to reverse a defect in CA pentamer formation induced by mutation of the IP6-interacting residue K25¹¹ and was also described as a compensatory mutation for the hypostable capsid mutant P38A⁴². Interestingly, T216I also arose during propagation of the IP6-packaging-deficient mutants KAKA and KAKA/T8I in this study (Supplementary Table 2), though its addition to KAKA/T8I only mildly restored infectivity (Supplementary Figure 2). Propagation of IP6-packaging-deficient mutants also resulted in the acquisition of mutations at several other amino acid positions previously implicated in pentamer formation. These include the N21S mutation within the central pore, which exhibits the same phenotype described above for T216I¹¹; the T58I and G61E mutations, which are part of the so-called “Thr-Val-Gly-Gly (TVGG) switch” that controls pentamer formation³³; and the M68I mutation, which lies near the Met-66 residue that plays a critical role in the conformation of the TVGG switch³³.

While the data clearly show that G225R enhances mature capsid assembly *in vitro* and in virions, the size of the effect in virions does not fully explain the observed increase in

infectivity. Upon the addition of G225R, we observed a roughly 30% increase in the number of KAKA/T8I virions containing mature conical capsids. However, KAKA/T8I/G225R is ~8-fold more infectious than KAKA/T8I. This suggests that KAKA/T8I/G225R capsids are inherently more stable than KAKA/T8I capsids. We also note that propagation of KAKA/T8I/G225R resulted in the acquisition of the A77V mutation, which doubled the infectivity of KAKA/T8I/G225R to ~30% of WT infectivity. A77 lies at the NTD-CTD interface near a binding pocket critical for HIV-1 capsid interactions with host factors including CPSF6⁴³, Nup153⁴⁴, and Sec24C⁴⁵, and capsid inhibitors such as PF74⁴⁶ and lenacapavir (LEN)⁴⁷. Interestingly, A77V has been previously described as a CPSF6-binding-deficient mutant that maintains WT-level fitness in primary cells and humanized mice⁴⁸. The interaction between the HIV-1 capsid and CPSF6 has been implicated in several viral functions including nuclear entry, uncoating, and integration site selection⁴⁹⁻⁵¹. CPSF6-binding-deficient mutant cores have been shown to dock at nuclear pores but not enter the nucleus, completing reverse transcription and uncoating at the nuclear envelope⁵⁰. It is unclear how a loss of CPSF6 binding would increase the infectivity of KAKA/T8I/G225R. One possibility is that KAKA/T8I/G225R cores are not stable enough to efficiently pass through nuclear pores intact, and the addition of A77V spares these cores the stress of passing through nuclear pores, allowing reverse transcription and uncoating to occur more efficiently. Alternatively, A77V may simply enhance capsid assembly or the stability of KAKA/T8I/G225R cores or perform some other infectivity-promoting function in this context. The mutants found in these selections will provide fertile ground for further virological and structural investigation.

We observed that G225R-mediated rescue of capsid assembly in low IP6 conditions was dependent on the presence of the SP1-T8I mutation. This is curious because T8I did not restore IP6 enrichment in the context of KAKA or KAKA/G225R (Figure 1 G-H and Figure 2 A-B) and does not participate in capsid assembly after its removal from CA during the final step of Gag processing. G225R drastically reduces KAKA virus production efficiency and infectivity, while having no effect on KAKA/T8I virus production efficiency and increasing its infectivity (Figure 1K and Figure 2 E&F). Our interpretation of these results is that G225R imposes a destabilizing effect on IP6-independent immature Gag lattice assembly that is restored by the addition of the T8I mutation. T8I promotes efficient IP6-independent particle assembly and G225R subsequently promotes efficient capsid formation at reduced IP6 concentrations during particle maturation. These results suggest that the G225R mutation

imposes contrasting IP6-dependent effects on immature and mature assembly. Interestingly, the T216I mutation demonstrated the same T8I-dependent phenotype as G225R (Supplementary Figure 2). Together, these findings highlight the evolutionary pressure faced by HIV-1 CA to balance immature and mature assembly to maximize virion production and infectivity.

Methods and materials

Cells

MT-4, SupT1, and C8166 human T cell lines were acquired from American Type Culture Collection (ATCC) and cultured at 37°C with 5% CO₂ in Roswell Park Memorial Institute (RPMI) 1640 medium (Corning) supplemented with 10% fetal bovine serum (GenClone), 2mM L-glutamine, penicillin (100U/mL; Gibco), and streptomycin (100µg/mL). HEK 293T cells were acquired from ATCC and cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented as above. HEK 293T Parental, IPMK KO, and IPPK KO cells, described in previous studies^{13,20,21}, were cultured in DMEM supplemented as above.

Plasmids

pNL4-3, a lab adapted, subtype B infectious molecular clone was used for transfection of MT-4, C8166, and SupT1 T-cell lines to initiate replication and forced evolution experiments. pNL4-3 was also used to transfect 293T cells to produce virus for virus production efficiency and infectivity experiments. pNL4-3 plasmids harbouring deletions in the *pol* and *env* genes (pNL4-3 Δ*pol*Δ*env*) were used to transfect 293T cells for the production of immature virions for cryo-ET analysis. pNL4-3 KFS (*env*-) was used to transfect HEK 293T cells to produce mature virions for cryo-ET analysis. An *env*-/*vpr*-, luciferase-encoding pNL4-3 plasmid (pNL4-3.Luc.R-.E-) was co-transfected into 293T cells along with a plasmid encoding VSV-G to generate VSV-G pseudo-typed HIV-1 particles for infectivity experiments in 293T Parental and IPPK KO target cells. HIV-1 CA mutants were generated by subcloning of DNA fragments purchased from Twist Bioscience.

Forced-evolution experiments

MT-4 or C8166 cells were transfected with WT or mutant replication-competent pNL4-3 infectious molecular clones. Replication kinetics were quantified using a 32P-based RT assay as described previously⁵². Virus collected at the peak of replication was used to infect fresh cells for further propagation. Genomic DNA isolated from cells collected at the peak of replication was used as the template for PCR amplification of the HIV-1 Gag-coding region. PCR-amplified Gag fragments were treated with ExoSAP-IT PCR product cleanup reagent (Thermo Fisher Scientific Catalog #78201.1ML) and subjected to Sanger sequencing to identify potential compensatory mutations. Selected mutations were incorporated into pNL4-3 in the presence and absence of the original mutations by subcloning Gag fragments

purchased from Twist Bioscience. The resultant molecular clones were used to transfect MT-4 cells to measure their inherent replication kinetics.

Virus Production Efficiency

Parental, IPMK KO, and IPPK KO HEK 293T cells were co-transfected with pNL4-3 WT or mutants plasmids with an expression plasmid encoding MINPP1 (IPMK KO and IPPK KO) or an empty vector (Parental) as described previously²⁰. Viruses were produced for 24 hours at 37°C with 5% CO₂. Viral supernatants were filtered through a 0.45µm syringe filter, ultracentrifuged and virus pellets and cell monolayers were lysed and subjected to western blot analysis. Virion-associated and cell-associated Gag was detected using the HIV-Immunoglobulin (HIV-Ig), imaged using a Sapphire Biomolecular Imager (Azure Biosystems), and quantified using Azure Spot analysis software. Virus production efficiency was calculated using the formula below and normalized to WT virus production efficiency in parental cells.

$$\frac{\text{virus } p24}{\text{virus } p24 + \text{cell } p24 + \text{cell } p41 + \text{cell } Pr55}$$

Single-Cycle Infectivity

Parental, IPMK KO, or IPPK KO HEK 293T cells were transfected as above. Virus present in filtered supernatants was quantified by RT assay, and single-cycle infectivity assays were performed as described previously²⁰. Briefly, virus stocks were serially diluted and used to infect TZM-bl cells in the presence of 10µg/mL DEAE-dextran. Specific infectivity was determined by measuring luminescence after lysis of the infected cells 36-48 hours post-infection. All infectivity measurements were expressed as a percentage of the infectivity of WT virions produced from parental HEK 293T cells. To measure single-cycle infectivity in IP6 depleted target cells, HEK 293T cells were co-transfected with pNL4-3.Luc.R-.E- and a plasmid expressing VSV-G. VSV-G-pseudo-typed, replication-incompetent virus was quantified by RT assay and used to infect HEK 293T Parental or IPPK KO target cells as described above.

Purification of mature and immature virions

HEK 293T cells were transfected with the pNL4-3-derived vectors using GenJet (Ver. II) transfection reagent as per the manufacturer's protocol. The media was changed 8-16 hours post-transfection and virions were produced for 36-48 hours at 37°C with 5% CO₂.

Supernatants were centrifuged briefly and filtered through a 0.45µm syringe filter to remove cellular debris. Virions were pelleted by ultracentrifugation through an 8% iodixanol cushion and resuspended in phosphate buffered saline (PBS). Virions were then purified by ultracentrifugation through a 10-30% iodixanol gradient, diluted with PBS, and subjected to additional centrifugation step. Pelleted virions were resuspended in a small volume of PBS and loaded onto grids for cryo-ET analysis.

CA protein purification and assembly

The cDNAs encoding HIV-1 capsid (CA) with G225R, K158A/K227A and K158A/G225R/K227A (KAKA/G225R) were mutated from the CA-pNL4-3-pET21 vector and verified by sequencing. Mutated CA proteins were purified as described in Liu, C. et al⁵³ and concentrated to 30 mg/ml. For CA cones/tubes assembly, 5 µl mutated CA proteins, 5 µl 500 mM MES pH=6.0 and 5 µl IP6 at a desired concentration were mixed and incubated at 37 °C for 1 hour. For supernatant and pellet assay, 10 µl of each assembled mixture was centrifuged at 12,000g for 10 minutes, supernatants and pellets were separated and diluted 1:10 with SDS loading dye before loading to SDS PAGE.

Negative stain grids preparation and image

For negative stain grids preparation, 400 mesh carbon coated copper grids were glow discarded. 3 µl assembled CA mutation mixture was loaded onto the grid, incubated for 1 minute. The grid was then blotted to dry and incubated with uranyl formate solution for 1 minute. The grid was blotted to dry again and loaded to a JEOL 120 keV TEM (Sir William Dunn School of Pathology, University of Oxford) for imaging. Number of assembled tubes and cones were counted from 5 micrographs recorded from each sample and averaged.

CryoET grids preparation

To prepare a frozen hydrated grids for tomography analysis using cryo-electron microscopy analysis, vitrification was performed using a Vitrobot (FEI, Thermo Scientific) operating at 4 °C and 98% relative humidity (RH). 10ul of concentrated gold fiducial was prepared by removing 90 µl of supernatant after the centrifugation of 100 µl of gold fiducial (Electron Microscopy Sciences, Inc.) at 15000 rpm for 30 min at 4 °C. 3 µl of freshly prepared HIV virions and 6 nm size of gold nanoparticles (Electron Microscopy Sciences, Inc.) was mixed at a ratio of 9:1 (v/v) was applied to lacey carbon film copper 300 mesh, 100-micron grids

(Electron Microscopy Sciences, Inc.) which was previously glow discharged at 15mA for 5 seconds with a Pelco easiGlow Glow discharge cleaning system (Ted Pella, Inc.). After 5 seconds, grids were blotted for 2 seconds and plunged into liquid ethane. Frozen sample grids were preserved under liquid nitrogen temperature for imaging.

CryoET data collection and subtomogram averaging.

For K227A and KAKA/T8I samples, cryo-ET data collection was acquired using a Thermo Fisher Titan Krios microscope operated at 300 keV equipped with a Gatan Quantum post-column energy filter (Gatan Inc) operated in zero-loss mode with 20 eV slit width, and Gatan K3 direct electron detector in eBIC (Electron BioImaging Centre, Diamond). Tilt series images were collected with SerialEM⁵⁴ with a nominal magnification of 64k and a physical pixel size of 1.34 Å per pixel. For G225R, KAKA/G225R, KAKA/T8I/G225R and A77V/KAKA/T8I/G225R samples, cryo-ET data collection was acquired using a Thermo Fisher Titan Krios microscope operated at 300 keV equipped with Facon4i detector and Selectrix with 10eV slit width in eBIC (Electron BioImaging Centre, Diamond). All the datasets were acquired using a dose-symmetric tilt-scheme⁵⁵ starting from 0° with a 3° tilt increment by a group of 3 and an angular range of ± 60°. The accumulated dose of each tilt series was around 123 e⁻/Å² with a defocus range between -1.5 and -6 μm. In total, 71 tilt series from KAKA sample, 81 tilt series from KAKA/T8I sample, 78 tilt series from G225R sample, 86 tilt series from KAKA/G225R sample, 61 tilt series from KAKA/T8I/G225R and 89 tilt series from A77V/KAKA/T8I/G225R sample were collected. Each projection image was dose-fractioned into ten frames. Details of data collection parameters are listed in Supplementary Table 1.

The automated cryo-ET pipeline developed in-house was used for pre-processing (https://github.com/ffyr2w/cet_toolbox) through performing motion correction⁵⁶ of the raw frames, tilt-series alignment with IMOD⁵⁷. The gold beads were manually inspected to ensure the centering of fiducial markers for each tilt series in eTOMO.

Subtomogram averaging for all the datasets was performed following the workflow of emClarity^{58,59}. The in vitro assembled HIV-1 Gag structure (EMD-8403)⁶⁰ was low-pass filtered to 30 Å and used as the initial template for template search in 6x binned tomograms with a pixel size of 8.04Å for K227A and KAKA/T8I samples, 9.006Å for G225R, KAKA/G225R, KAKA/T8I/G225R and A77V/KAKA/T8I/G225R samples. 141,486

subtomograms based on the initial position and orientations were selected from 71 tilt series for K227A dataset, 159,659 subtomograms were selected from 81 tilt series for the KAKA/T8I dataset, 284,966 subtomograms were selected from 78 tilt series for the G225R dataset, 165,561 subtomograms were selected from 86 tilt series for the KAKA/G225R dataset, 116,132 subtomograms were selected from 61 tilt series for the KAKA/T8I/G225R dataset, and 228,256 subtomograms were selected from 89 tilt series for the A77V/KAKA/T8I/G225R dataset. The subtomogram averaging and alignment were performed iteratively using 6x, 5x, 4x, 3x, 2x, and 1x binned tomograms. A cylindrical alignment mask including 7 hexamers and a 6-fold symmetry was applied throughout the whole procedure. The final density maps were reconstructed at bin 1 and sharpened with a b-factor of -50.

CryoET data collection and analysis of mature VLPs

Cryo-ET data collection was acquired using a Thermo Fisher Titan Krios microscope operated at 300 keV equipped with a Gatan Quantum post-column energy filter (Gatan Inc) operated in zero-loss mode with 20 eV slit width, and Gatan K3 direct electron detector in eBIC (Electron BioImaging Centre, Diamond). Tilt series images were collected with SerialEM⁵⁴ with a nominal magnification of 42k and a physical pixel size of 2.18 Å per pixel. All the datasets were acquired using a dose-symmetric tilt-scheme⁵⁵ starting from 0° with a 3° tilt increment by a group of 3 and an angular range of ± 60°. The accumulated dose of each tilt series was around 123 e⁻/Å² with a defocus range between -1.5 and -6 μm. In total, ~30 tilt series from each variant were collected. Each projection image was dose-fractioned into ten frames.

The automated cryo-ET pipeline developed in-house was used for pre-processing (https://github.com/ffyr2w/cet_toolbox) through performing motion correction⁵⁶ of the raw frames, tilt-series alignment with IMOD⁵⁷. The gold beads were manually inspected to ensure the centering of fiducial markers for each tilt series in eTOMO. The morphological analysis was conducted manually from reconstructed 3D tomograms.

CryoEM data collection and processing

CryoEM movie data were collected using EPU software on a Gatan K3 direct detector camera in super-resolution mode. Each movie contains 40 frames with an accumulated dose of 40 electrons/Å². The calibrated physical pixel size and the super-resolution pixel size were

1.34 and 0.67 Å per pixel, respectively. The defocus was prescribed in the range from -0.8 to -2.5 µm. A total of 8,443 movies in super-resolution mode were collected for data analysis.

All frames of the raw movies were corrected for their gain using a gain reference recorded within 3 days of the acquired movie to generate a single micrograph using CryoSPARC Patch Motion Correction. The micrographs were used for the determination of the actual defocus using CryoSPARC Patch CTF Estimation. 14,047,760 particles were picked and extracted using CryoSPARC Template Picker, in which the templates were averaged from manually picked ~1000 particles. 4,195,547 particles were selected after several rounds of CryoSPARC 2D classification. CryoSPARC heterogeneous refinement was conducted with C6 symmetry applied. One class with 1,646,193 particles was selected for non-uniform refinement. The final hexamer structure was obtained by 3D reconstruction at super-resolution mode to a global resolution of 2.75 Å.

Then the particles were re-centred on the trimer interface by moving the coordinates for further heterogeneous refinement. One class of 1,320,752 particles without CA C-terminus density was selected for non-uniform refinement to the final map with a global resolution of 3.59 Å. Another class of 309,508 particles with CA C-terminus density was selected for non-uniform refinement to the final map with a global resolution of 4.24 Å.

Atomic model of KAKA/G225R CA hexamer and structural analysis

The model was initially predicted from AlphaFold3⁶¹ and improved in Coot by manually main-chain and sidechain fitting. The refinement of the atomic model was carried out in real space with program Phenix.real_space_refine⁶², with secondary structure and geometry restraints to prevent overfitting. Structural comparison and visualization were conducted in UCSF ChimeraX⁶³ and Chimera⁶⁴. All figures of the structures were plotted in UCSF ChimeraX⁶³ and Chimera.

KAKA/G225R CA extended C-terminal model building

The initial model for a KAKA/G225R CA trimer of dimer with an extended C-terminal was derived from a WT CA hexamer (PDB ID 4XFX)⁶⁵ using Rosetta³⁷, by the following procedure: First, three CA hexamers were rigid-body fitted in the density derived for the KAKA/G225R CA tri-hexamer using ChimeraX⁶³. Next, the density corresponding to a

single monomer at the trimer interface, along with the corresponding rigid-body fitted CA monomer coordinates, was extracted. Residues 1-12 and 220 -231 were then removed from the CA monomer, and missing regions were remodelled into the density using RosettaES³⁹ with the KAKA/G225R CA sequence. In total, 200 models were sequentially generated by building the missing regions one residue at a time and choosing the most favourable conformation through a Monte Carlo procedure³⁹.

Furthermore, a trimer of dimers assembly was built from the KAKA/G225R CA monomer by applying a C3 symmetric Monte Carlo refinement of the interface sidechain orientation and the distance between monomers. The Monte Carlo refinement was implemented using the Generic Monte Carlo mover in RosettaScripts³⁸. The cryoEM density is used as a guiding potential at all steps of the modelling process in Rosetta by using the elec_dens_wt scoring function with a weight of 35. A total of 40 symmetrized and refined models were generated, and the model with the lowest Rosetta score was selected for further study via molecular dynamics.

Molecular dynamics simulations setup

In preparation for MD simulations of CA trimer of dimers, we prepared models for the KAKA/G225R, KAKA and WT CA as follows: First, we assigned the protonation state of the C-terminal extended KAKA/G225R CA trimer of dimers model at pH 7.0 and added hydrogen atoms as appropriate using the PDB2PQR⁶⁶ software version 3.6.1. We then generated models for a KAKA and WT CA trimer of dimers by reverting G225R or all mutations in the KAKA/G225R CA model using the Mutator plugin in the Visual Molecular Dynamics (VMD) software⁶⁷. Afterwards, we solvated each system in a water box using TIP3P⁶⁸ water molecules and added Na⁺/Cl⁻ ions around the protein and in bulk using the cionize and autoionize VMD plugins, to a 150mM NaCl concentration. We then employed the hydrogen mass repartitioning (HMR) procedure⁶⁹ to redistribute the mass of heavy atoms bound to hydrogen, allowing us to use a timestep of 4.0 fs during MD simulations for increased sampling efficiency. This procedure resulted one MD-ready CA trimer of dimers model for each CA sequence. Each MD-ready CA trimer of dimer models consisted of ~268,000 atoms, including solvent, spanning a unit cell of dimensions 173 Å x 173 Å x 88 Å.

Molecular dynamics simulations

MD-ready systems were minimized and equilibrated by the following procedure: First, each system was minimized using a conjugate gradient algorithm for 50,000 steps. Convergence of the minimization process was verified by ensuring that the energy gradient achieved values below $0.1 \text{ kcal/mol-}\text{\AA}^{-2}$. Next, each system was thermalized by gradually raising the temperature from 50 K to 310 K at a rate of 20 K/ns in NPT ensemble (constant number of particles, pressure $P=1 \text{ atm}$, and temperature). During thermalization, the position of protein backbone atoms was harmonically restrained with a spring constant of $10 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$. Following thermalization, the positional harmonic restraints are gradually released at a rate of $-1 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ every 0.2 ns in an NPT simulation at $T = 310 \text{ K}$ and $P = 1 \text{ atm}$.

Equilibration MD simulations were then performed for each CA trimer of dimers system for 400ns in an NPT ensemble at $T= 310 \text{ K}$ and $P= 1 \text{ atm}$. During equilibration, harmonic positional restraints with a spring constant of $0.5 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ were applied to the carbon alpha atoms of all alpha helices from residues 1 to 218; to maintain the stability of the trimer of dimers in the absence of a continuous CA lattice and to conserve the fitting to the cryoEM density, while linkers and the extended C-terminal segment (residue 220 to 231) were unrestrained and unbiased. In all simulations, pressure control was applied using a Nose-Hoover Langevin piston with a 200 ps period and a 100 ps decay time, maintaining a constant area in the XY plane of the unit cell while volume was allowed to fluctuate in the Z axis; temperature control was applied using a Langevin thermostat with 0.1 ps^{-1} damping coefficient.

All simulations were performed with the NAMD 3.0.1⁷⁰ molecular dynamics engine with the CHARMM36m⁷¹ force field parameters for proteins, water molecules and ions. MD simulations used a 4.0 fs timestep, in accordance with the HMR scheme. Bonds to hydrogen atoms were constrained using the SHAKE⁷² and SETTLE⁷³ algorithms. Nonbonded interactions were calculated with a 12 \AA cutoff, 10 \AA switching distance, and 14 \AA pairlist distance for short-range electrostatics, with long-range electrostatics calculated every 8.0 fs using the particle mesh Ewald summation method⁷⁴ with a 1 \AA grid spacing.

After completing the MD simulations, the trajectories were analysed using in-house scripts in VMD⁶⁷ to measure contact occupancies, with a threshold of 3.5 \AA for salt bridge interactions. Occupancy of the C-terminal segment at the dimer interface was considered when at atoms of the C-terminal region established 3.5 \AA contacts with atoms from the

neighbouring CA monomers' 3-10 helix or helix 9 simultaneously. Volume occupancies were calculated using the volmap tool in VMD⁶⁷.

Data Availability

The immature and mature HIV-1 capsid lattice structures are deposited in the Electron Microscopy Data Bank (EMDB) and Protein Data Bank (PDB) under accession codes: EMD-51826 for the K227A immature CA hexamer, EMD-51825 for the KAKA/T8I immature CA hexamer, EMD-51821 for G225R immature CA hexamer, EMD-51822 for the KAKA/G225R immature CA hexamer, EMD-51823 for the KAKA/T8I/G225R immature CA hexamer, EMD-51824 for the A77V/KAKA/T8I/G225R immature CA hexamer, EMD-52724 and PDB 9I8I for the KAKA/G225R mature CA hexamer, EMD-52725 for the KAKA/G225R mature CA tri-hexamer Class 1, and EMD-52726 for the KAKA/G225R mature CA tri-hexamer Class 2.

Acknowledgements

We are grateful to Yuewen Sheng and Yun Song for technical support for cryoEM data collection, and to Tapan Kanai for assistance with cryoEM grid preparation. We thank Diamond Light source for access and support of the cryoEM facilities at the UK National Electron Bio-Imaging Centre (eBIC) (proposal NT29812). Computation was performed at the Oxford Biomedical Research Computing (BMRC) facility, a joint development between the Wellcome Centre for Human Genetics (Wellcome Trust Core Award Grant Number 203141/Z/16/Z) and the Big Data Institute (BDI) supported by Health Data Research UK and the NIHR Oxford Biomedical Research Centre. This work was supported by the National Institutes of Health (P50AI150481), the UK Wellcome Trust Investigator Award 206422/Z/17/Z, the UK Biotechnology and Biological Sciences Research Council grant BB/S003339/1 and ERC AdG grant 101021133. Research in the Freed laboratory is supported by the Intramural Research Program of the Center for Cancer Research, National Cancer Institute, National Institutes of Health. AK was supported in part by an Intramural AIDS Research Fellowship and the National Institute of Allergy and Infectious Disease (K99AI174891-01). Additionally, we would like to acknowledge our collaborative interactions with the Pittsburgh Center for HIV Protein Interactions (U54AI170791). Work in the J.R.P. laboratory was funded in part by the National Institutes of Health through grant R01AI178846. This work used the Delta supercomputing resources at the University of Illinois Urbana-Champaign as well as the Stampede3 supercomputer at the Texas Advanced

Computing Center (TACC) through allocation MCB-170096 from the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program, which is supported by National Science Foundation awards #2138259, #2138286, #2138307, #2137603, and #2138296. We acknowledge computational support through the Delaware Advanced Research Workforce and Innovation Network (DARWIN) at the University of Delaware.

Author contributions

P.Z., E.O.F., and A.B.K. conceived the research. A.B.K prepared the immature and mature VLP samples and conducted the forced evolution, infectivity, virus production experiments. J.S. purified WT CA and mutant variants, conducted the *in vitro* CLP assembly assay. J.S., Y.S. and Y.Z. prepared cryoEM grids and collected cryoEM data. Y.Z. collected cryoET data. Y.Z., L.C., A.L. and N.H. determined the immature Gag structures. Y.Z. carried out the cryoEM single particle analysis, determined the structures and built the model. J.X. conducted the statistical analysis of virion morphology. J.S.R performed molecular dynamics (MD) simulations. Y.Z, P.Z., A.B.K., E.O.F, J.S.R. and J.R.P wrote the paper with the input from all the authors.

Reference

- 1 Bell, N. M. & Lever, A. M. HIV Gag polyprotein: processing and early viral particle assembly. *Trends Microbiol* **21**, 136-144 (2013). <https://doi.org/10.1016/j.tim.2012.11.006>
- 2 Zhao, G. *et al.* Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature* **497**, 643-646 (2013). <https://doi.org/10.1038/nature12162>
- 3 Jang, S. & Engelman, A. N. Capsid-host interactions for HIV-1 ingress. *Microbiol Mol Biol Rev* **87**, e0004822 (2023). <https://doi.org/10.1128/membr.00048-22>
- 4 Dick, R. A. *et al.* Inositol phosphates are assembly co-factors for HIV-1. *Nature* **560**, 509-512 (2018). <https://doi.org/10.1038/s41586-018-0396-4>
- 5 Mallery, D. L. *et al.* IP6 is an HIV pocket factor that prevents capsid collapse and promotes DNA synthesis. *Elife* **7** (2018). <https://doi.org/10.7554/eLife.35335>

- 6 Renner, N. *et al.* A lysine ring in HIV capsid pores coordinates IP6 to drive mature capsid assembly. *Plos Pathog* **17**, e1009164 (2021). <https://doi.org/10.1371/journal.ppat.1009164>
- 7 Ni, T. *et al.* Structure of native HIV-1 cores and their interactions with IP6 and CypA. *Science Advances* **7** (2021). <https://doi.org/10.1126/sciadv.abj5715>
- 8 Jacques, D. A. *et al.* HIV-1 uses dynamic capsid pores to import nucleotides and fuel encapsidated DNA synthesis. *Nature* **536**, 349-353 (2016). <https://doi.org/10.1038/nature19098>
- 9 Highland, C. M., Tan, A., Ricana, C. L., Briggs, J. A. G. & Dick, R. A. Structural insights into HIV-1 polyanion-dependent capsid lattice formation revealed by single particle cryo-EM (vol 120, e2220545120, 2023). *P Natl Acad Sci USA* **120** (2023). <https://doi.org/10.1073/pnas.2312510120>
- 10 Xu, C. *et al.* Permeability of the HIV-1 capsid to metabolites modulates viral DNA synthesis. *PLoS Biol* **18**, e3001015 (2020). <https://doi.org/10.1371/journal.pbio.3001015>
- 11 Kleinpeter, A. *et al.* HIV-1 adapts to lost IP6 coordination through second-site mutations that restore conical capsid assembly. *Nature Communications* **15** (2024). <https://doi.org/10.1038/s41467-024-51971-w>
- 12 Schirra, R. T., dos Santos, N. F. B., Ganser-Pornillos, B. K. & Pornillos, O. Arg18 Substitutions Reveal the Capacity of the HIV-1 Capsid Protein for Non-Fullerene Assembly. *Viruses-Basel* **16** (2024). <https://doi.org/10.3390/v16071038>
- 13 Mallery, D. L. *et al.* Cellular IP Levels Limit HIV Production while Viruses that Cannot Efficiently Package IP Are Attenuated for Infection and Replication. *Cell Rep* **29**, 3983 (2019). <https://doi.org/10.1016/j.celrep.2019.11.050>
- 14 Mendonça, L. *et al.* CryoET structures of immature HIV Gag reveal six-helix bundle. *Commun Biol* **4** (2021). <https://doi.org/10.1038/s42003-021-01999-1>
- 15 Wagner, J. M. *et al.* Crystal structure of an HIV assembly and maturation switch. *Elife* **5** (2016). <https://doi.org/10.7554/eLife.17063>
- 16 Schur, F. K. M. *et al.* An atomic model of HIV-1 capsid-SP1 reveals structures regulating assembly and maturation. *Science* **353**, 506-508 (2016). <https://doi.org/10.1126/science.aaf9620>
- 17 Waki, K. *et al.* Structural and Functional Insights into the HIV-1 Maturation Inhibitor Binding Pocket. *Plos Pathog* **8** (2012). <https://doi.org/10.1371/journal.ppat.1002997>
- 18 Mallery, D. L. *et al.* A stable immature lattice packages IP(6) for HIV capsid maturation. *Sci Adv* **7** (2021). <https://doi.org/10.1126/sciadv.abe4716>
- 19 Fontana, J. *et al.* Identification of an HIV-1 Mutation in Spacer Peptide 1 That Stabilizes the Immature CA-SP1 Lattice. *J Virol* **90**, 972-978 (2016). <https://doi.org/10.1128/Jvi.02204-15>
- 20 Kleinpeter, A. B. *et al.* The Effect of Inositol Hexakisphosphate on HIV-1 Particle Production and Infectivity can be Modulated by Mutations that Affect the Stability of the Immature Gag Lattice. *J Mol Biol* **435** (2023). <https://doi.org/10.1016/j.jmb.2023.168037>
- 21 Renner, N. *et al.* HIV-1 is dependent on its immature lattice to recruit IP6 for mature capsid assembly. *Nat Struct Mol Biol* (2023). <https://doi.org/10.1038/s41594-022-00887-4>
- 22 Blair, W. S. *et al.* HIV Capsid is a Tractable Target for Small Molecule Therapeutic Intervention. *Plos Pathog* **6** (2010). <https://doi.org/10.1371/journal.ppat.1001220>

- 23 Lemke, C. T. *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). <https://doi.org/10.1128/Jvi.00493-12>
- 24 Siddiqui, M. A. *et al.* A Novel Phenotype Links HIV-1 Capsid Stability to cGAS-Mediated DNA Sensing. *J Virol* **93** (2019). <https://doi.org/10.1128/JVI.00706-19>
- 25 Lamorte, L. *et al.* Discovery of Novel Small-Molecule HIV-1 Replication Inhibitors That Stabilize Capsid Complexes. *Antimicrob Agents Ch* **57**, 4622-4631 (2013). <https://doi.org/10.1128/Aac.00985-13>
- 26 Lopez, C. S., Tsagli, S. M., Sloan, R., Eccles, J. & Barklis, E. Second site reversion of a mutation near the amino terminus of the HIV-1 capsid protein. *Virology* **447**, 95-103 (2013). <https://doi.org/10.1016/j.virol.2013.08.023>
- 27 Busnadiego, I. *et al.* Host and Viral Determinants of Mx2 Antiretroviral Activity. *J Virol* **88**, 7738-7752 (2014). <https://doi.org/10.1128/Jvi.00214-14>
- 28 Saito, A. *et al.* Roles of Capsid-Interacting Host Factors in Multimodal Inhibition of HIV-1 by PF74. *J Virol* **90**, 5808-5823 (2016). <https://doi.org/10.1128/Jvi.03116-15>
- 29 Aberham, C., Weber, S. & Phares, W. Spontaneous mutations in the human immunodeficiency virus type 1 gag gene that affect viral replication in the presence of cyclosporins. *J Virol* **70**, 3536-3544 (1996). <https://doi.org/10.1128/Jvi.70.6.3536-3544.1996>
- 30 Yoo, S. H. *et al.* Molecular recognition in the HIV-1 capsid/cyclophilin a complex. *J Mol Biol* **269**, 780-795 (1997). <https://doi.org/10.1006/jmbi.1997.1051>
- 31 Ricana, C. L., Lyddon, T. D., Dick, R. A. & Johnson, M. C. Primate lentiviruses require Inositol hexakisphosphate (IP6) or inositol pentakisphosphate (IP5) for the production of viral particles. *Plos Pathog* **16** (2020). <https://doi.org/10.1371/journal.ppat.1008646>
- 32 Stacey, J. C. V. *et al.* Two structural switches in HIV-1 capsid regulate capsid curvature and host factor binding. *P Natl Acad Sci USA* **120** (2023). <https://doi.org/10.1073/pnas.2220557120>
- 33 Schirra, R. T. *et al.* A molecular switch modulates assembly and host factor binding of the HIV-1 capsid. *Nat Struct Mol Biol* **30**, 383-390 (2023). <https://doi.org/10.1038/s41594-022-00913-5>
- 34 Zuliani-Alvarez, L. *et al.* Evasion of cGAS and TRIM5 defines pandemic HIV. *Nat Microbiol* **7**, 1762-1776 (2022). <https://doi.org/10.1038/s41564-022-01247-0>
- 35 Ni, T. *et al.* Intrinsic curvature of the HIV-1 CA hexamer underlies capsid topology and interaction with cyclophilin A. *Nat Struct Mol Biol* **27**, 855-862 (2020). <https://doi.org/10.1038/s41594-020-0467-8>
- 36 Byeon, I. J. L. *et al.* Structural Convergence between Cryo-EM and NMR Reveals Intersubunit Interactions Critical for HIV-1 Capsid Function. *Cell* **139**, 780-790 (2009). <https://doi.org/10.1016/j.cell.2009.10.010>
- 37 Leman, J. K. *et al.* Macromolecular modeling and design in Rosetta: recent methods and frameworks. *Nature methods* **17**, 665-680 (2020).
- 38 Fleishman, S. J. *et al.* RosettaScripts: a scripting language interface to the Rosetta macromolecular modeling suite. *PloS one* **6**, e20161 (2011).
- 39 Frenz, B., Walls, A. C., Egelman, E. H., Veesler, D. & DiMaio, F. RosettaES: a sampling strategy enabling automated interpretation of difficult cryo-EM maps. *Nature methods* **14**, 797-800 (2017).

- 40 Sowd, G. A., Shi, J., Fulmer, A. & Aiken, C. HIV-1 capsid stability enables inositol phosphate-independent infection of target cells and promotes integration into genes. *Plos Pathog* **19** (2023). <https://doi.org/10.1371/journal.ppat.1011423>
- 41 Gupta, M., Pak, A. J. & Voth, G. A. Critical mechanistic features of HIV-1 viral capsid assembly. *Science Advances* **9** (2023). <https://doi.org/10.1126/sciadv.add7434>
- 42 Yang, R. F. *et al.* Second-site suppressors of HIV-1 capsid mutations: restoration of intracellular activities without correction of intrinsic capsid stability defects. *Retrovirology* **9** (2012). <https://doi.org/10.1186/1742-4690-9-30>
- 43 Price, A. J. *et al.* CPSF6 Defines a Conserved Capsid Interface that Modulates HIV-1 Replication. *Plos Pathog* **8** (2012). <https://doi.org/10.1371/journal.ppat.1002896>
- 44 Price, A. J. *et al.* Host Cofactors and Pharmacologic Ligands Share an Essential Interface in HIV-1 Capsid That Is Lost upon Disassembly. *Plos Pathog* **10** (2014). <https://doi.org/10.1371/journal.ppat.1004459>
- 45 Rebensburg, S. V. *et al.* Sec24C is an HIV-1 host dependency factor crucial for virus replication. *Nat Microbiol* **6**, 435-444 (2021). <https://doi.org/10.1038/s41564-021-00868-1>
- 46 Bhattacharya, A. *et al.* Structural basis of HIV-1 capsid recognition by PF74 and CPSF6. *P Natl Acad Sci USA* **111**, 18625-18630 (2014). <https://doi.org/10.1073/pnas.1419945112>
- 47 Link, J. O. *et al.* Clinical targeting of HIV capsid protein with a long-acting small molecule. *Nature* **584**, 614-618 (2020). <https://doi.org/10.1038/s41586-020-2443-1>
- 48 Saito, A. *et al.* Capsid-CPSF6 Interaction Is Dispensable for HIV-1 Replication in Primary Cells but Is Selected during Virus Passage. *J Virol* **90**, 6918-6935 (2016). <https://doi.org/10.1128/Jvi.00019-16>
- 49 Sowd, G. A. *et al.* A critical role for alternative polyadenylation factor CPSF6 in targeting HIV-1 integration to transcriptionally active chromatin. *P Natl Acad Sci USA* **113**, E1054-E1063 (2016). <https://doi.org/10.1073/pnas.1524213113>
- 50 Burdick, R. C. *et al.* HIV-1 uncoats in the nucleus near sites of integration. *P Natl Acad Sci USA* **117**, 5486-5493 (2020). <https://doi.org/10.1073/pnas.1920631117>
- 51 Lee, K. *et al.* Flexible Use of Nuclear Import Pathways by HIV-1. *Cell Host Microbe* **7**, 221-233 (2010). <https://doi.org/10.1016/j.chom.2010.02.007>
- 52 Willey, L. R. S., H.D.; Lasky, A.L.; Theodore, S.T.; Earl, L.P.; Moss, B.; Capon, J.D.; Martin, A.M. In Vitro Mutagenesis Identifies a Region within the Envelope Gene of the Human Immunodeficiency Virus That Is Critical for Infectivity. *J Virol* **62**, 139-147 (1988).
- 53 Liu, C. *et al.* Cyclophilin A stabilizes the HIV-1 capsid through a novel non-canonical binding site. *Nature Communications* **7** (2016). <https://doi.org/10.1038/ncomms10714>
- 54 Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J Struct Biol* **152**, 36-51 (2005). <https://doi.org/10.1016/j.jsb.2005.07.007>
- 55 Hagen, W. J. H., Wan, W. & Briggs, J. A. G. Implementation of a cryo-electron tomography tilt-scheme optimized for high resolution subtomogram averaging. *J Struct Biol* **197**, 191-198 (2017). <https://doi.org/10.1016/j.jsb.2016.06.007>
- 56 Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332 (2017). <https://doi.org/10.1038/nmeth.4193>

- 57 Mastrorarde, D. N. & Held, S. R. Automated tilt series alignment and tomographic reconstruction in IMOD. *J Struct Biol* **197**, 102-113 (2017).
<https://doi.org/10.1016/j.jsb.2016.07.011>
- 58 Himes, B. A. & Zhang, P. emClarity: software for high-resolution cryo-electron tomography and subtomogram averaging. *Nat Methods* **15**, 955-961 (2018).
<https://doi.org/10.1038/s41592-018-0167-z>
- 59 Ni, T. *et al.* High-resolution in situ structure determination by cryo-electron tomography and subtomogram averaging using emClarity. *Nat Protoc* **17**, 421-444 (2022). <https://doi.org/10.1038/s41596-021-00648-5>
- 60 Ning, J. *et al.* In vitro protease cleavage and computer simulations reveal the HIV-1 capsid maturation pathway. *Nat Commun* **7**, 13689 (2016).
<https://doi.org/10.1038/ncomms13689>
- 61 Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493-500 (2024). <https://doi.org/10.1038/s41586-024-07487-w>
- 62 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). <https://doi.org/10.1107/S2059798319011471>
- 63 Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci* **27**, 14-25 (2018). <https://doi.org/10.1002/pro.3235>
- 64 Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612 (2004).
<https://doi.org/10.1002/jcc.20084>
- 65 Gres, A. T. *et al.* X-ray crystal structures of native HIV-1 capsid protein reveal conformational variability. *Science* **349**, 99-103 (2015).
- 66 Dolinsky, T. J., Nielsen, J. E., McCammon, J. A. & Baker, N. A. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic acids research* **32**, W665-W667 (2004).
- 67 Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J Mol Graph* **14**, 33-38, 27-38 (1996). [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5)
- 68 Jorgensen, W. L. & Jenson, C. Temperature dependence of TIP3P, SPC, and TIP4P water from NPT Monte Carlo simulations: Seeking temperatures of maximum density. *Journal of computational chemistry* **19**, 1179-1186 (1998).
- 69 Hopkins, C. W., Le Grand, S., Walker, R. C. & Roitberg, A. E. Long-time-step molecular dynamics through hydrogen mass repartitioning. *Journal of chemical theory and computation* **11**, 1864-1874 (2015).
- 70 Phillips, J. C. *et al.* Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J Chem Phys* **153**, 044130 (2020). <https://doi.org/10.1063/5.0014475>
- 71 Huang, J. *et al.* CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nature methods* **14**, 71-73 (2017).
- 72 Elber, R., Ruymgaart, A. P. & Hess, B. SHAKE parallelization. *The European Physical Journal Special Topics* **200**, 211-223 (2011).
- 73 Miyamoto, S. & Kollman, P. A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of computational chemistry* **13**, 952-962 (1992).
- 74 Essmann, U. *et al.* A smooth particle mesh Ewald method. *The Journal of chemical physics* **103**, 8577-8593 (1995).

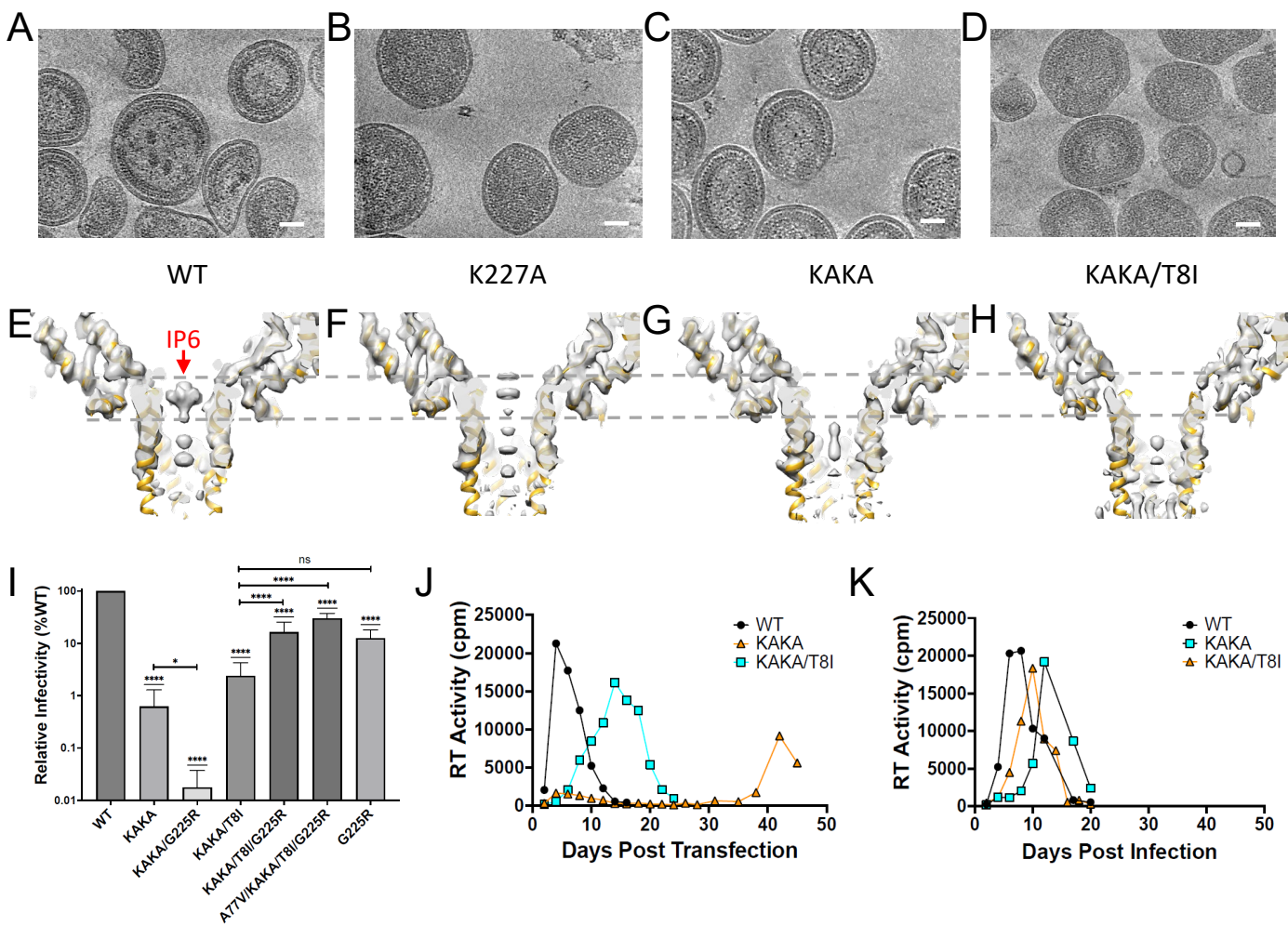


Figure 1 | Characterization of Gag mutants defective in IP6 binding and a compensatory mutation G225R. (A-D) Tomographic slices of WT, K227A, K158A/K227A (KAKA) and KAKA/T8I VLPs respectively. Scale bar, 50nm. (E-H) CryoET STA structures of immature Gag hexamers from WT, K227A, KAKA and KAKA/T8I VLPs superimposed with the model (PDB: 7ASH). Dashed grey lines mark the height position for IP6. I) Specific infectivity of KAKA and KAKA/T8I in the presence and absence of G225R measured in TZM-bl cells at 36-48 hours post-infection. Data are the mean of at least three independent replicates for each mutant and error bars depict +/- SEM. Statistical analysis was performed using GraphPad Prism. Statistical significance was determined by a one-sample Student's t-test with a hypothetical value set to 100 when comparing groups to WT. All other comparisons were made using unpaired Student's t-tests. (p-value summary: >0.05 = not significant; <0.05 = *; <0.01 = **; <0.001 = ***; <0.0001 = ****). J) Representative replication kinetics of KAKA and KAKA/T8I mutants in MT-4 cells showing a delay in viral replication relative to WT. K) Representative re-passage of KAKA and KAKA/T8I in fresh MT-4 cells showing an increase in replication kinetics relative to the initial passage.

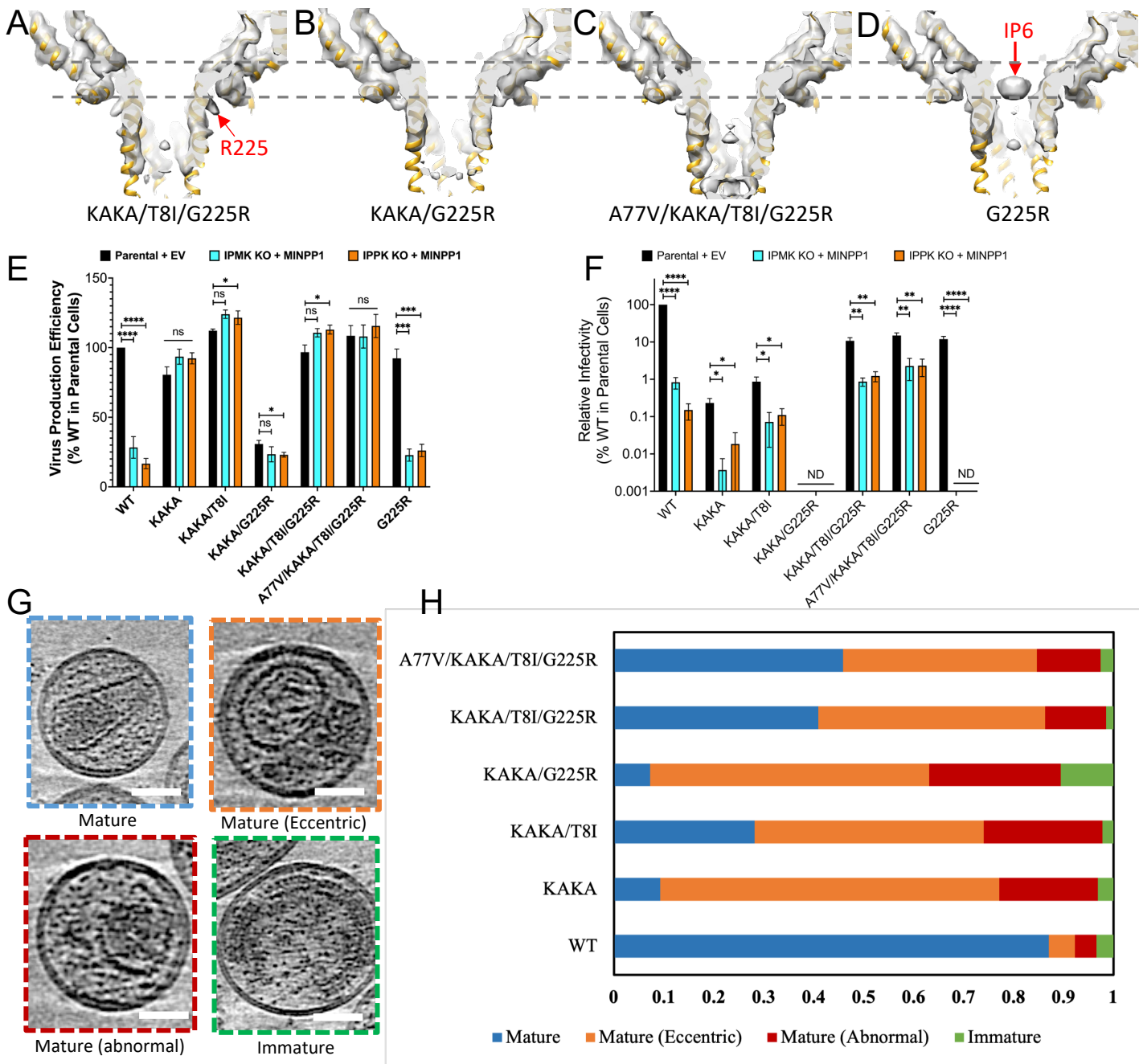


Figure 2 | G225R does not restore IP6 recruitment in IP6 deficient mutants. (A-D) CryoET STA structures of immature Gag from KAKA/T8I/G225R, KAKA/G225R, A77V/KAKA/T8I/G225R and G225R VLPs. E) Cell lysates and concentrated virus lysates were collected after a 24-hour incubation following media change after co-transfection of pNL4-3 with 500 ng of empty vector (EV) or MINPP1 expression vector (IPMK KO and IPPK KO). Cell and virus Gag levels were quantified by western blot. F) Virus from cells transfected as above was collected 24 hours post-transfection and assessed for specific infectivity on TZM-bl cells. Infectivity was normalized to the infectivity of WT produced from parental cells. Data are the mean of at least three independent replicates for each mutant and error bars depict +/- SEM. Statistical analysis was performed using GraphPad Prism. Statistical significance was determined by a one-sample Student's t-test with a hypothetical value set to 100 when comparing groups to WT. All other comparisons were made using unpaired Student's t-tests. (p-value summary: >0.05 = not significant; <0.05 = *; <0.01 = **; <0.001 = ***; <0.0001 = ****). (G) A Gallery of distinct morphologies of HIV-1 particles produced from HEK-293 cells, shown in tomographic slices. Particle morphologies are classified as indicated. The slice thickness is 4.36 nm. Scale bar, 50 nm. (H) Distribution of particle morphologies of WT (n=442) and KAKA (n=162), KAKA/T8I (n=181), KAKA/G225R (n=152), KAKA/T8I/G225R (n=66), A77V/KAKA/T8I/G225R (n=687) mutant Gag virions.

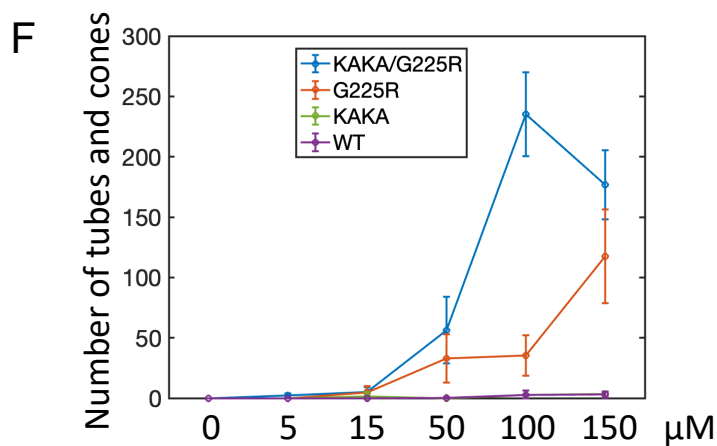
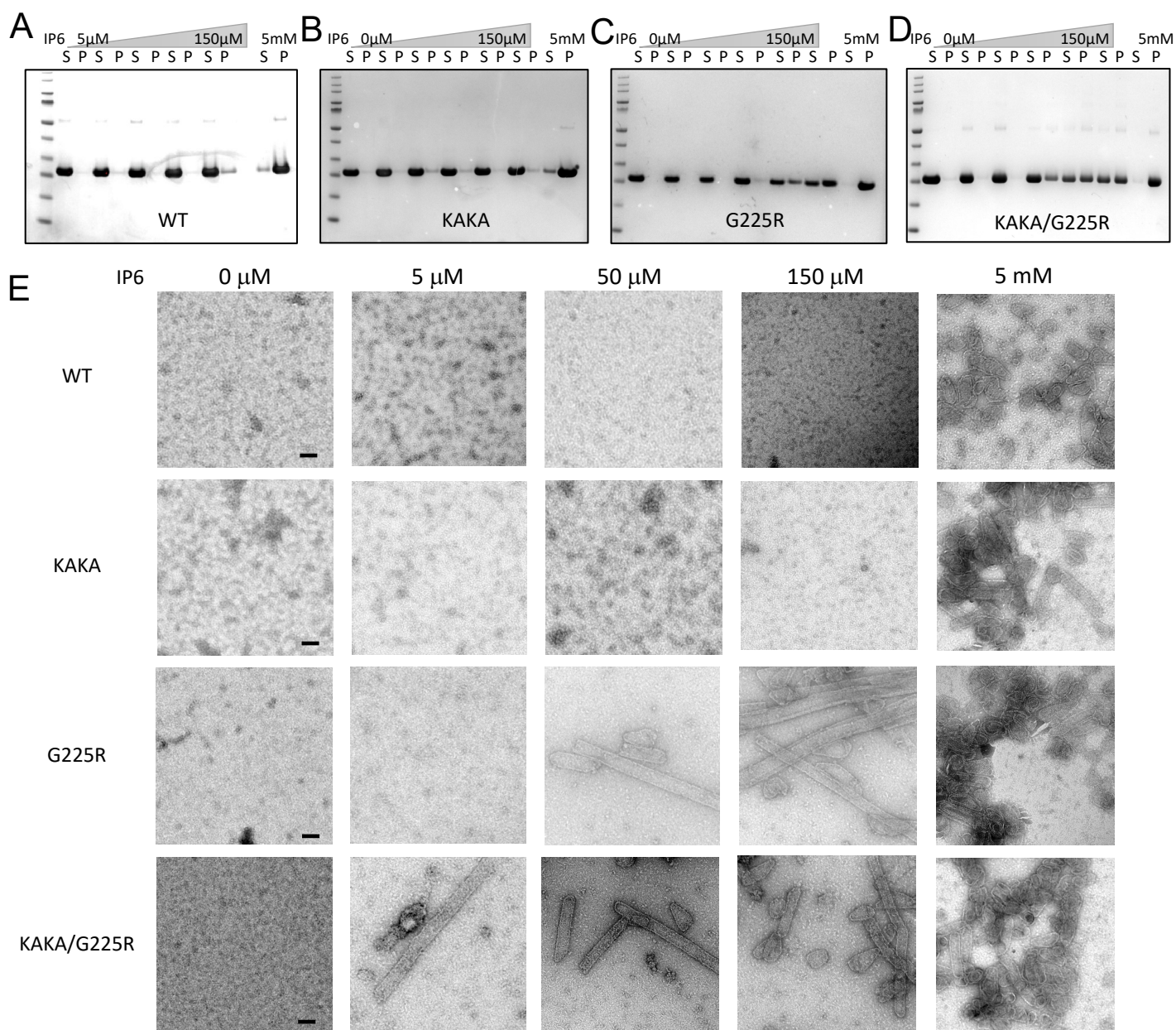


Figure 3 | G225R increases the efficiency of capsid formation *in vitro*. (A-D) Sup-pellet arrays for the *in vitro* assembly of CA WT, KAKA, G225R and KAKA/G225R at 0, 5, 15, 50, 150 μ M and 5 mM IP6 concentrations. (E) Negative stain images of CA WT, KAKA, G225R and KAKA/G225R with IP6 concentrations from left to right are 0, 5, 50 and 150 μ M respectively. (F) The numbers of assembled cones and tubes in each micrographs of WT and mutant CA at different IP6 concentrations (n=5). Scale bars, 100nm

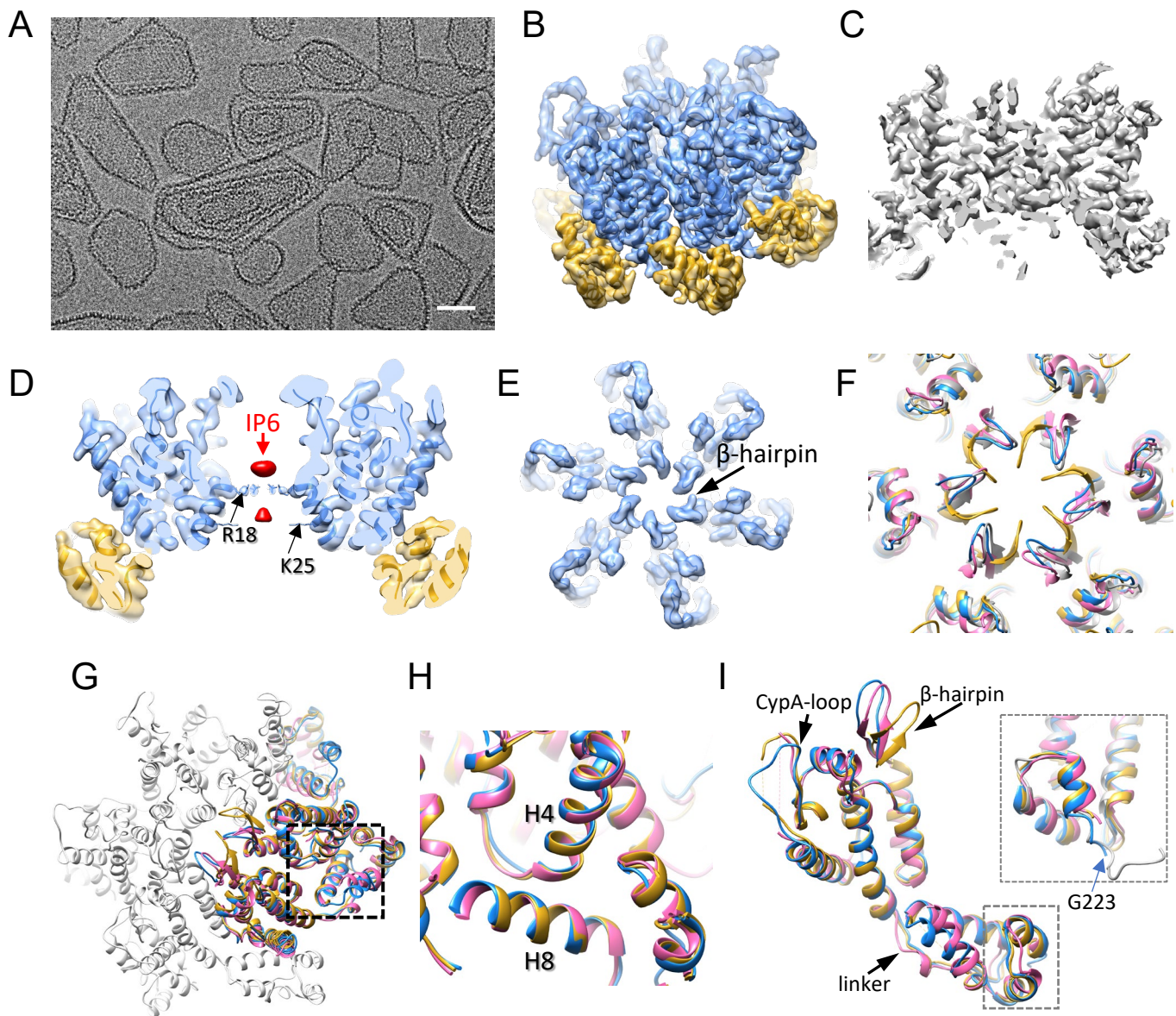


Figure 4 | CryoEM structure of CA KAKA/G225R hexamer. (A) A representative cryoEM micrograph of *in vitro* assembled KAKA/G225R capsid with 100 μ M IP6. (B) The overview KAKA/CA G225R density map at 2.7 Å resolution superimposed with the refined molecular model (PDB 9I8I), with CA NTD colored blue and CA CTD colored gold. (C) The central slice of KAKA/G225R CA density map. (D) The central slice of KAKA/G225R CA hexamer density map superimposed with the refined model with IP6 density colored pink. R18 sidechain is indicated. (E) The top view of KAKA/G225R CA hexamer density map superimposed with the refined model, showing β -hairpins in an open conformation. (F) Comparison of β -hairpin among HIV-1 CA hexamers from crystal of WT (M) (PDB 5HGN, gold) and HIV-1 (O) (PDB 7T12, pink), cryoEM of WT CLP (PDB 7URN, grey) and KAKA/G225R (PDB 9I8I, blue). (G) Comparison of CA NTD-CTD interfaces among HIV-1 KAKA/G225R, HIV-1 WT (M) and HIV-1 (O). (H) An enlarged view of NTD-CTD interfaces. Helices H4 from CA1 NTD and H8 from CA2 CTD are labelled. (I) Comparison of CA monomers among HIV-1 KAKA/G225R (PDB 9I8I, blue), HIV-1 (M) WT (PDB 5HGN, gold) and HIV-1 (O) (PDB 7T12, pink). Inset: an enlarged view of the C-terminus, additionally overlapped with an NMR HIV-1 WT CTD₁₄₄₋₂₃₁ model (PDB 2KOD, grey). The last resolved residue G223 in the cryoEM structure (blue) is marked. Scale bar, 50 nm

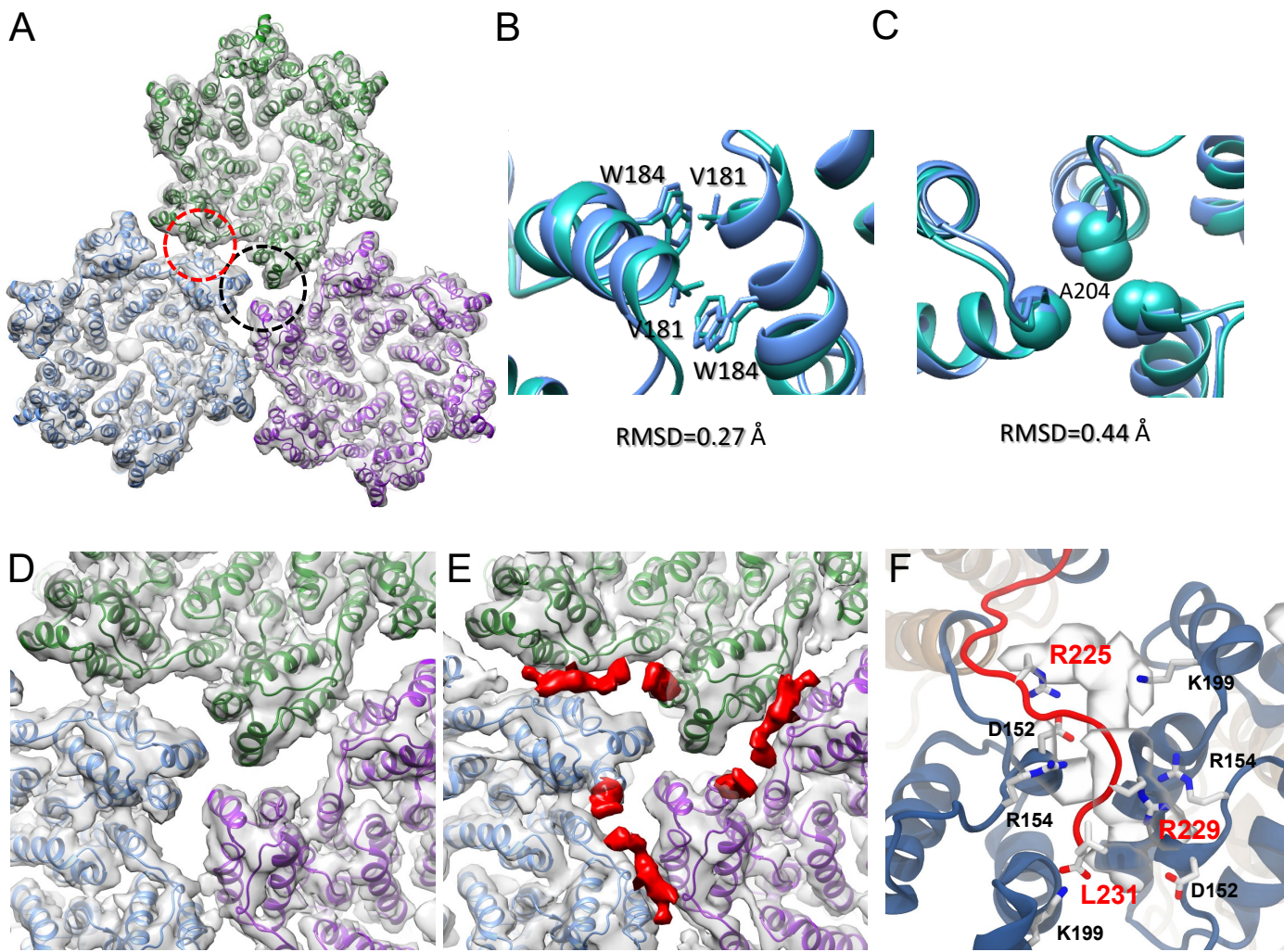


Figure 5 | CryoEM structures of CA KAKA/G225R tri-hexamer. (A) The overview KAKA/CA G225R tri-hexamer density map at 3.63 Å resolution superimposed with the refined molecular model (PDB 9I8I). Three hexamer models are colored blue, green and purple. The dimer and trimer interfaces are marked with red and black dashed circles. (B) Comparison of dimer interfaces between WT (PDB 6SKN, green) and KAKA/G225R (PDB 9I8I, blue). (C) Comparison of trimer interfaces between WT (PDB 6SKN, green) and KAKA/G225R (PDB 9I8I, blue). (D-E) Two 3D classes of CA KAKA/G225R tri-hexamer maps, superimposed with model (PDB 9I8I), viewed from CTD side (inside of capsid). Class 2 (E) distinguishes from Class 1 (D) by additional densities (segmented in red) occupying the space from H11 to the dimer interface. (F) Interactions observed between C-terminal region (220-231, red) and residues at the dimer interface in unbiased molecular dynamics simulation of the KAKA/G225R CA trimer of dimers superimposed with the density map. Interacting residues are labelled.

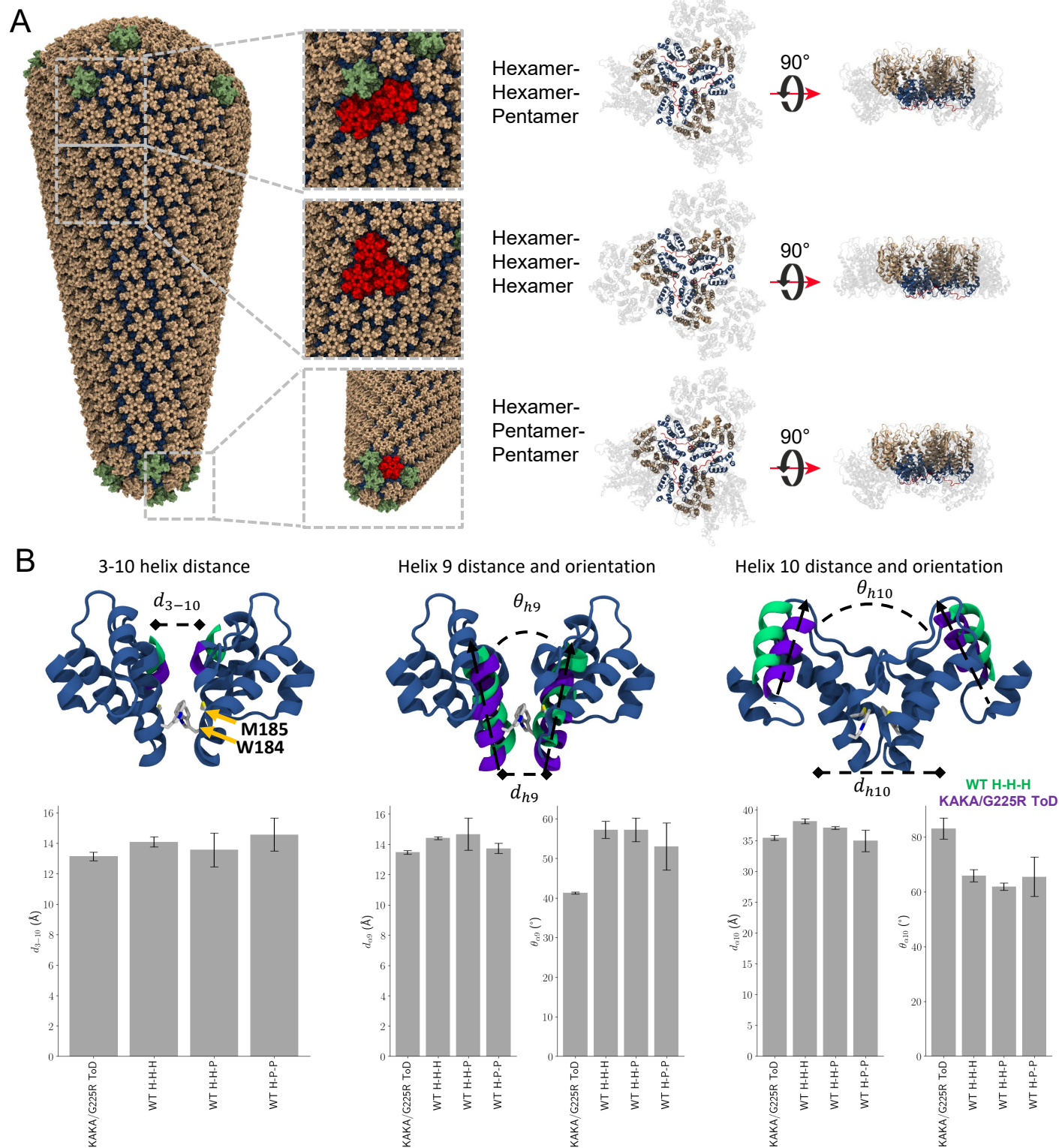


Figure 6 | KAKA/G225R CA dimer interfaces compared to assembly interfaces from full size capsid cone. (A) All cases of capsomer interfaces in the full WT capsid (left) overlaid on the KAKA/G225R CA trimer of dimers (right). NTD is colored tan on hexamers and green on pentamers; CTD is colored dark blue. (B) Distance and orientation measurements of the CTD structural elements at the dimer interface for the KAKA/G225R CA trimer of dimers (ToD) or the WT capsomer interfaces involving hexamers (H) and pentamers (P). Helices for a WT hexamer-hexamer-hexamer are colored in teal and helices for the KAKA/G225R CA ToD are colored in purple. Dimer interface residues M185 and W184 are visualized in licorice representation.

Supplementary figures and tables

Supplementary Table 1. CryoET data collection and structure determination of immature Gag

Data acquisition						
samples	K227A	KAKA/T8I	G225R	KAKA /G225R	KAKA/T8I /G225R	A77V/KAKA /T8I/G225R
Microscope	Titan Krios					
Voltage (keV)	300					
Detector	Gatan Quantum K3		Facon4i			
Energy-filter	Yes					
Slit width (eV)	20		10			
Super-resolution mode	No					
Å/pixel	1.34	1.34	1.501	1.501	1.501	1.501
Defocus range (µm)	-1.5 to -6					
Defocus increment (µm)	0.3					
Acquisition scheme	Dose-Symmetric, -60/60, 3 ^o step, group 3					
Total Dose (electrons/Å ²)	123					
Number of Frames	10					
Number of Tomograms	71	81	78	86	61	89
Structure determination						
No. of subtomograms	141,486	159,659	284,966	165,561	116,132	228,256
Resolution (Å)	3.82	3.67	3.79	4.10	4.05	3.78
b-factor applied	-10	-10	-10	-10	-10	-10
Data deposition	EMD-51826	EMD-51825	EMD-51821	EMD-51822	EMD-51823	EMD-51824

Supplementary Table 2. Mutants deficient for IP6 packaging acquire mutations throughout CA upon propagation in T cell lines. The listed IP6-packaging-deficient mutants were transfected into the indicated T cells and propagated. Virus from the peak of replication was collected and re-passaged in fresh cells. Sanger sequencing of genomic DNA isolated from infected cells was used to identify mutations acquired during propagation. .

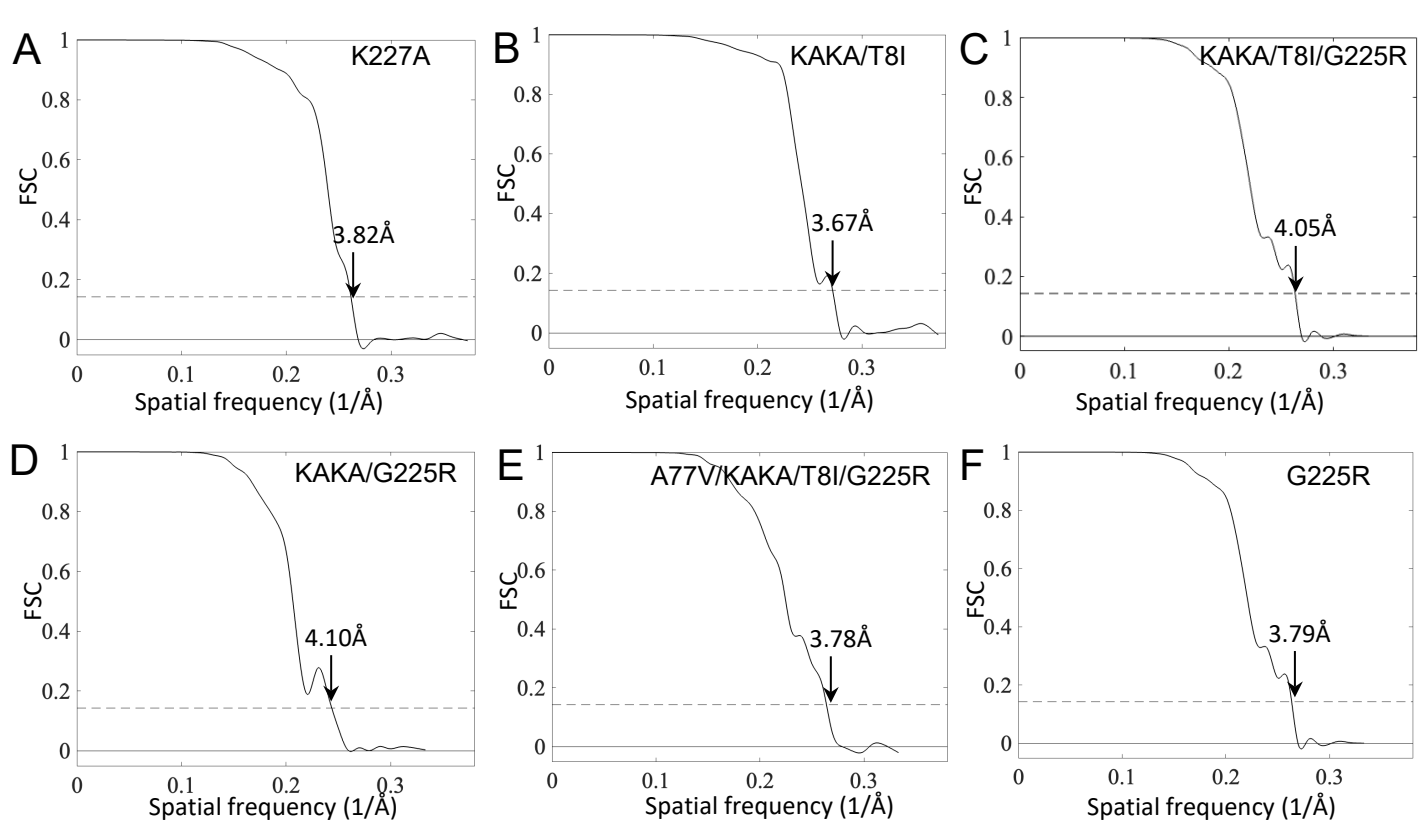
Transfected Clone	Passages	Cells	Mutations
KTKA/M4L	2	C8166	N21S, A105T, T210S
KTKA/T8I	2	C8166	M68I
KAKA/T8I	1	C8166	M68I, T107N, G208R
KAKA/M4L	1	C8166	G208R, S146N
KTKA/T8I	1	C8166	M68I, H87Q, G208R
KTKA/M4L	1	C8166	G208R, Q192K, P207S
KAKA	2	MT4	G225S
KTKA	2	MT4	E187Q, P207T
KAKA/T8I	2	MT4	T216I
KTKA/T8I	2	MT4	G61E
KAKA/M4L/T8I	5	MT4	M68I, G94D
KAKA/M4L/T8I	2	MT4	A31D,G61E
KTKA/T8I/G225S	3	MT4	G61E
KAKA/T8I/G225S	3	MT4	T58I
KAKA/T216I/G225S	3	MT4	T58I
KAKA/T8I	3	MT4	T58I,M68I
KTKA/T8I	3	MT4	
KAKA	4	MT4	V11I, SP1-M4I
KTKA/T216I	4	MT4	SP1-M4I
KTKA/T8I/G225S	4	MT4	S225R
KTKA/T216I	2	C8166	H12Y, D156N, G225S
KTKA/T8I/G225S	2	C8166	M68I, G61E, S146N
KAKA/G225S	2	C8166	MA E40K
KAKA/T8I/G225S	2	C8166	A22T, P207S
KAKA/T8I/G225R	3	MT4	A77V,A88E

Supplementary Table 3. CryoEM data collection and structure determination of assembled mature capsid.

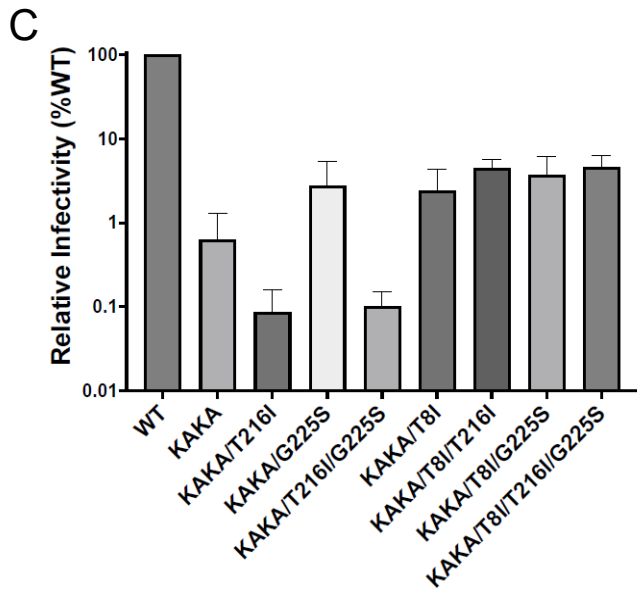
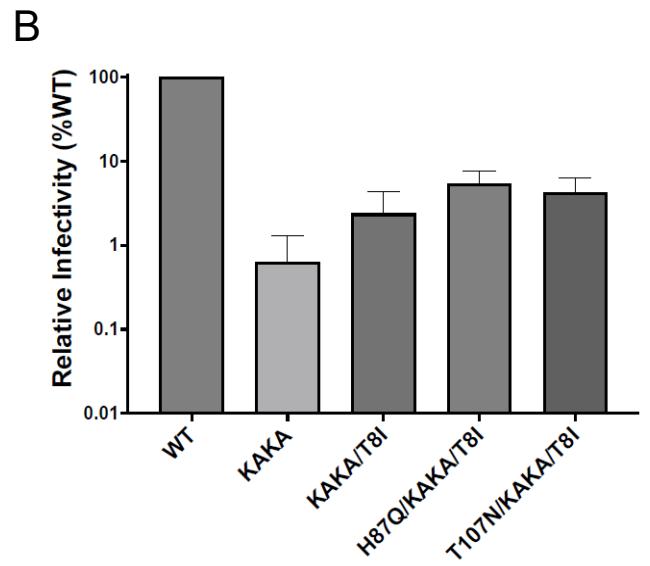
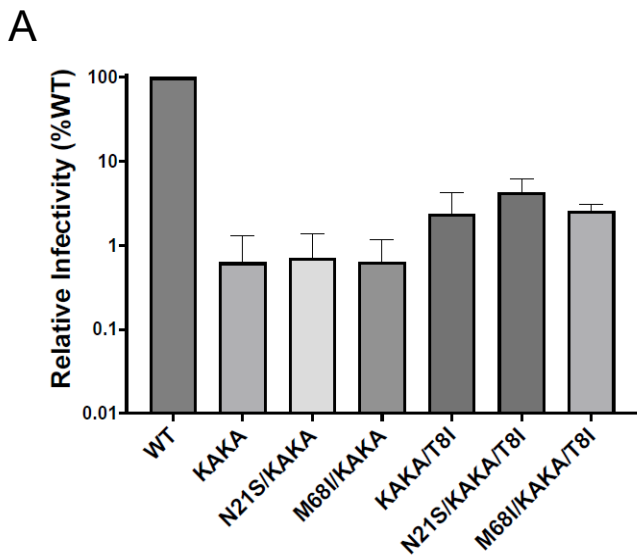
Data collection and processing			
KAKA/G225R CA	Hexamer	Tri-hexamer Classs 1	Tri-hexamer Classs 2
Microscope	Tian Krios		
Magnification	64,000		
Voltage	300 kV		
Electron dose	40 e-/Å ²		
Detector	Gatan Quantum K3		
Defocus Range	-0.8 to -2.5 μm		
Pixel Size	1.34 Å		
Symmetry	C6	C3	C3
Particle numbers	1,646,193	1,320,752	309,508
Map resolution(Å)	2.75	3.59	4.24
FSC threshold	0.143	0.143	0.143
Local resolution range(Å)	2.3-4.0	3.2-6.0	3.7-6.0
Refinement			
Initial Model used (PDB code)	AlphaFold3 Prediction		
Model resolution(Å)	3.5		
FSC threshold	0.5		
Map sharpening <i>B</i> factor(Å ²)	-130.3		
Model composition			
Non-hydrogen atoms	10374		
Residues	1338		
R.m.s.d deviations			
Bond length(Å)	0.003		
Bond angles(°)	0.941		
Validation			
MolProbity score	2.24		
Clashscore	12.16		
Rotamers Outliers (%)	2.23		
Ramachandran Plot			
Favored(%)	94.49		
Allowed(%)	5.51		
Outliers(%)	0.00		
Data deposition	EMD-52724 PDB:9I8I	EMD-52725	EMD-52726

Supplementary Table 4. Highest C-terminal segment (residues 220-231) contact occupancies at the CA dimer interface for KAKA/G225R CA, KAKA CA and WT CA from 400ns MD simulations.

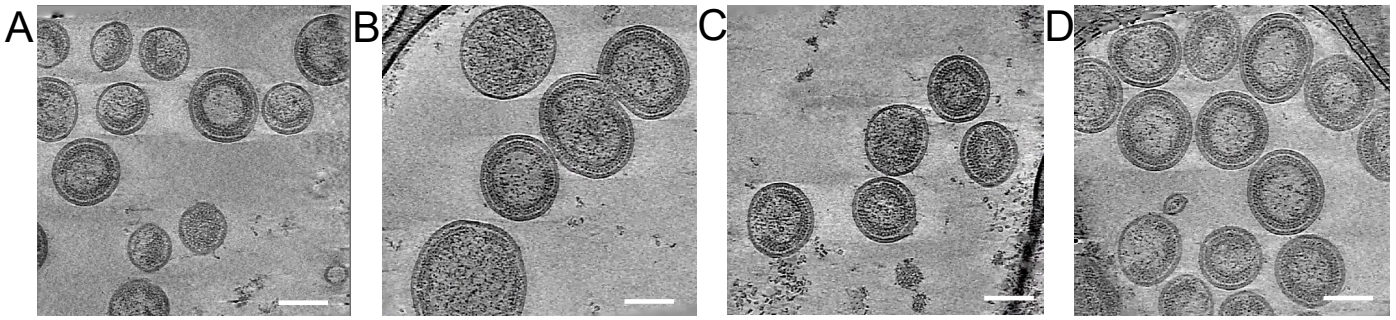
KAKA/G225R CA		KAKA CA		WT CA	
Residue interaction	Occupancy range (%)	Residue interaction	Occupancy range (%)	Residue interaction	Occupancy range (%)
R225-D152	2 – 16	G225-D152	0 - 7	K227-D152	0 – 16
R229-N193	3 – 57	R229-D152	2 – 4	R229-N193	0 – 2
R229-Q192	3 – 57	L231-R154	1 – 4	R229-Q192	0 – 3
R229-D152	9 – 46	L231-K199	1 – 9	R229-D152	0 – 12
L231-R154	17 – 50			R229-D197	0 – 9
L231-K199	3 - 27			L231-Q192	0 – 16
				L231-R154	0 – 11
				L231-K199	0 - 7



Supplementary Figure 1 | Fourier Shell Correlation (FSC) plot of immature Gag density maps. (A-F) Fourier Shell Correlation (FSC) plot of cryoET STA density maps from K227A, KAKA/T8I, KAKA/T8I/G225R, KAKA/G225R, A77V/KAKA/T8I/G225R and G225R VLPs. The resolution is indicated at the FSC value of 0.143.



Supplementary Figure 2 | Infectivity of compensatory mutants in combination with KAKA and KAKA/T8I. Mutations acquired during propagation of IP6-packaging-deficient mutants were introduced into KAKA and KAKA/T8I pNL4-3 vectors and infectivity in TZM-bl cells was expressed as a percentage of WT.

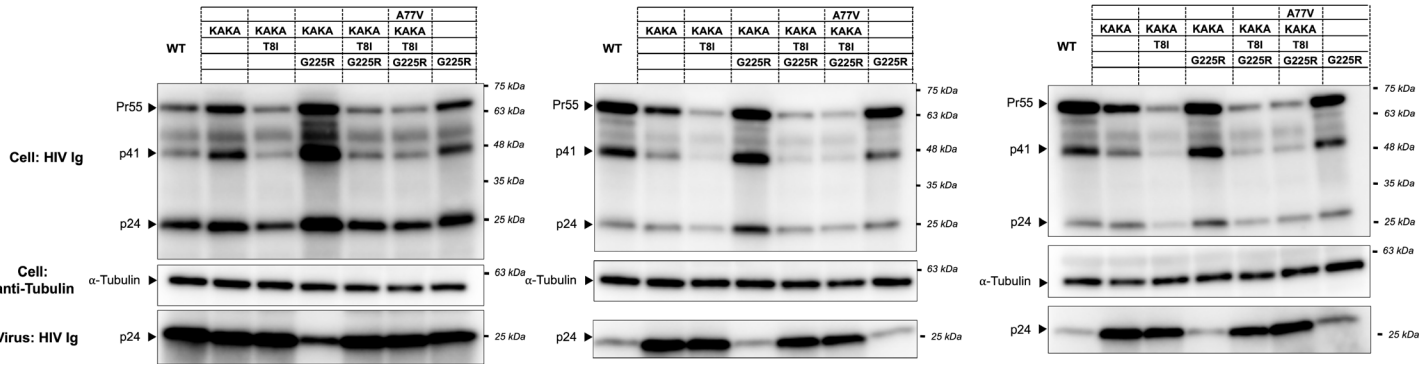


Supplementary Figure 3 | CryoET slices of these immature VLPs. (A-D) Tomographic slices of KAKA/T8I/G225R, KAKA/G225R, A77V/KAKA/T8I/G225R and G225R VLPs with slice thickness 4.02 nm. Scale bar, 100 nm.

Parental + EV

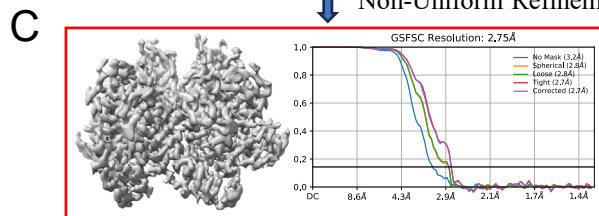
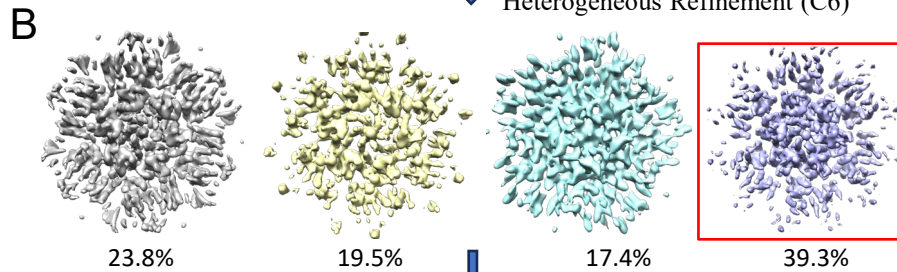
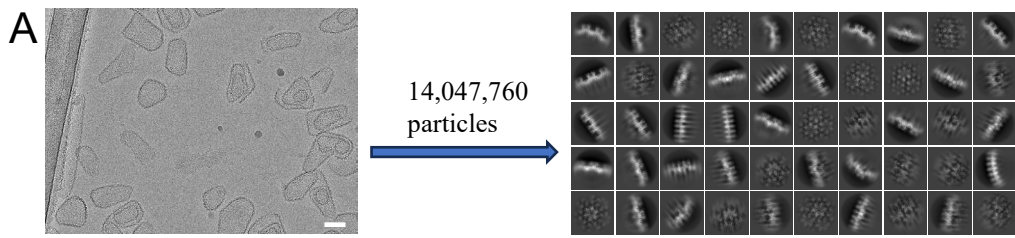
IPMK KO + MINPP1

IPPK KO + MINPP1

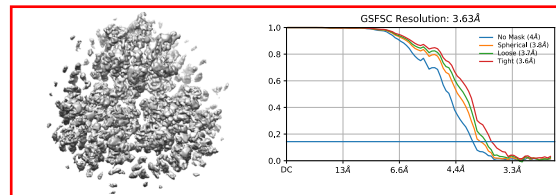


Supplementary Figure 4 | Western blot analysis for virus production efficiency measurements.

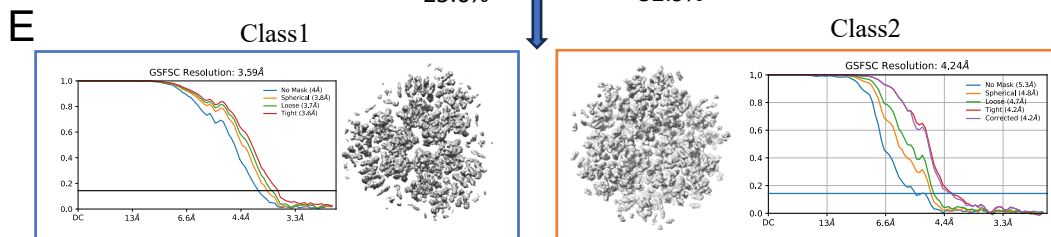
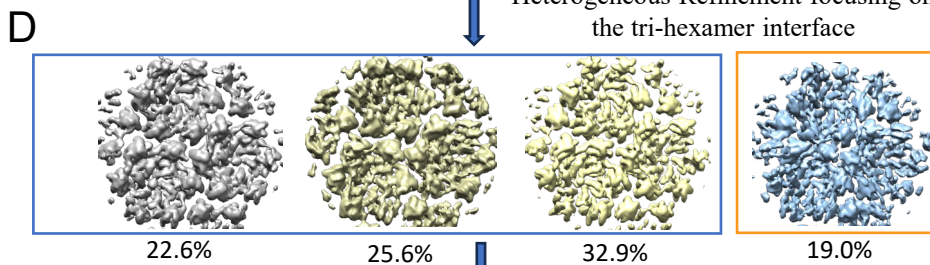
Representative western blots are shown for cell and viral lysates collected from 293T Parental cells co-transfected with pNL4-3 and empty vector (EV) (left), and 293T IPMK KO (middle) and IPPK KO cells co-transfected with pNL4-3 and a MINPP1 expression vector. Cellular Gag and virus p24 were quantified, and virus production efficiency was calculated in each cell line as described in the Materials and Methods.



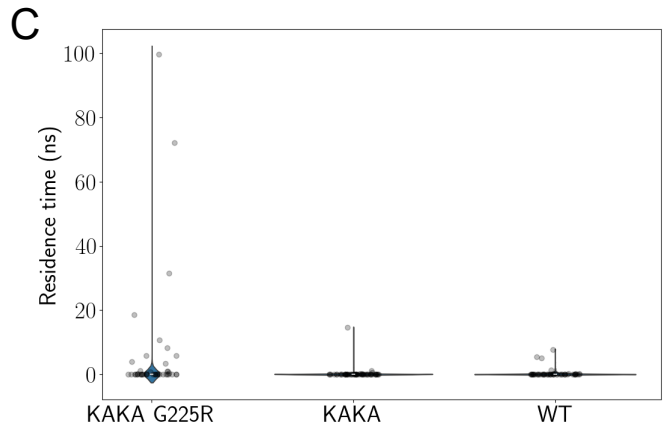
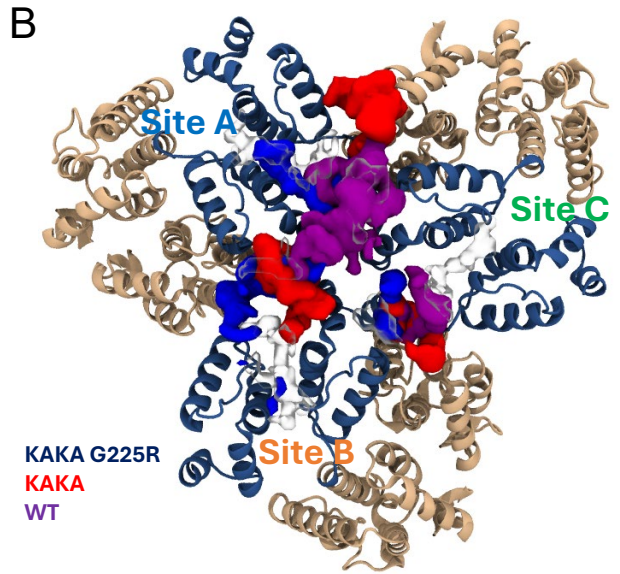
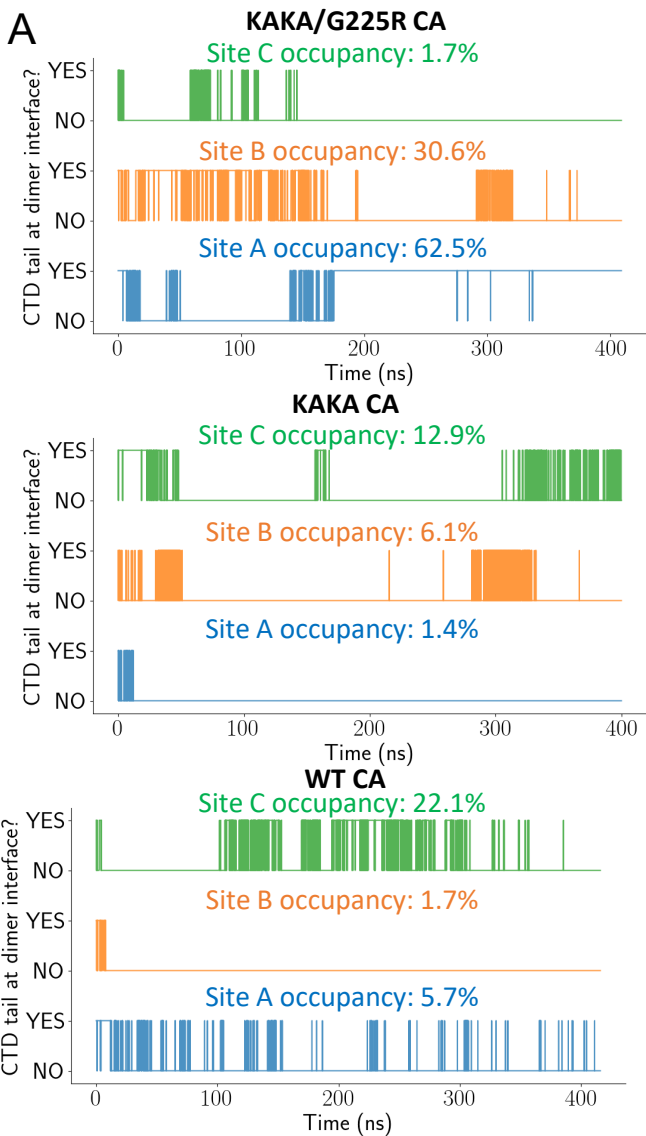
Re-center and Non-Uniform Refinement focusing on the tri-hexamer interface



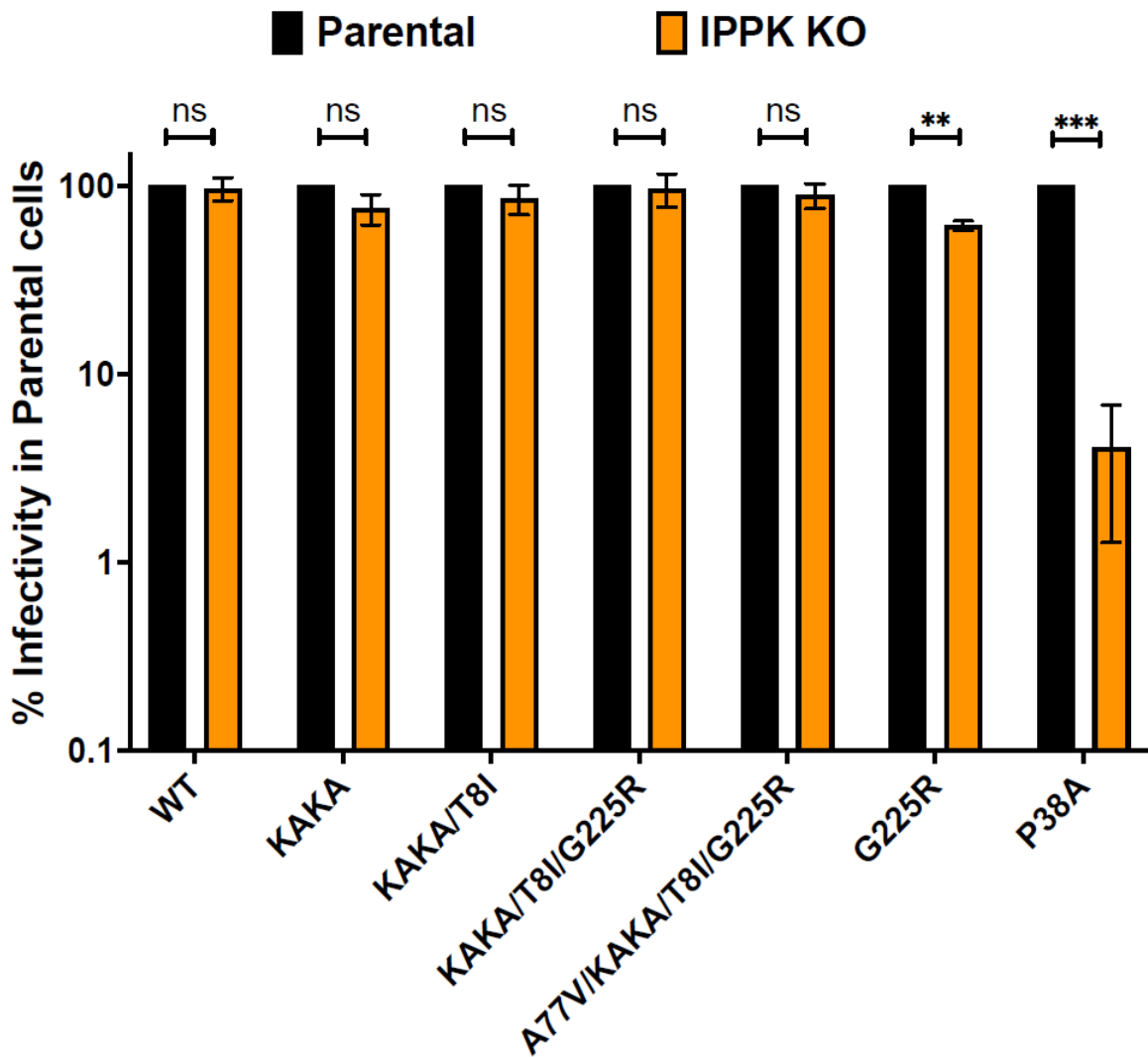
Heterogeneous Refinement focusing on the tri-hexamer interface



Supplementary Figure 5 | CryoEM structural determination of KAKA/G225R CA capsid hexamer. (A) Step1: Reference-free particle picking and 2d class averaging to remove bad particles. **(B) Step2:** Heterogeneous refinement with C6 symmetry applied. **(C) Step3:** Non-uniform refinement with the best class selected for Step2. **(D) Step4:** Re-center the particles on the tri-hexamer interface and perform heterogeneous refinement focusing on the tri-hexamer interface. **(E) Step5:** Subgroup to two classes based on the density of CA-C terminus tail and perform Non-uniform refinement respectively.



Supplementary Figure 6 | C-terminal segment occupancy from KAKA/G225R, KAKA and WT CA trimer of dimer MD simulations. (A) Traces of C-terminal contacts at the CA dimer interface through 400ns MD simulations for KAKA/G225R CA, KAKA CA and WT CA. The three dimer interfaces in the CA trimer of dimers are colored separately in green, orange and blue. (B) Volume occupancy map derived for the position of atoms in the C-terminal segment through 400ns for KAKA/G225R CA (dark blue), KAKA CA (red) and WT CA (purple). Each dimer interface is labeled according to the colors in (A). (C) Distribution of residence times for C-terminal interactions with residues in the CA dimer interface.



Supplementary Figure 7 | Infectivity of IP6-packaging-deficient mutants and rescue mutants is unaffected by target-cell IP6 depletion. VSV-G-pseudotyped pNL4-3.Luc.R-E- virions were produced in HEK 293T cells and used to infect HEK 293T Parental and IPPK KO cells. Specific infectivity was measured at 48 hours post-infection and compared between each cell type for each mutant. Infectivity for each mutant in IPPK KO target cells is expressed as a percentage of that mutant's infectivity in parental target cells. Data are the mean of three independent replicates for each mutant and error bars depict +/- SEM. Statistical analysis was performed using GraphPad Prism. Statistical significance was determined by a one-sample Student's t-test with a hypothetical value set to 100 (p-value summary: >0.05 = not significant; <0.05 = *; <0.01 = **; <0.001 = ***; <0.0001 = ****).