Governed by surface amino acid composition: HIV capsid passage through the NPC barrier

Liran Fu[1], Shiya Cheng[2,4], Dietmar Riedel[3], Leonie Kopecny[3], Melina Schuh[2] & Dirk Görlich[1*]

[1]Department of Cellular Logistics, [2]Department of Meiosis, and [3]Laboratory for Electron Microscopy, Max Planck Institute for Multidisciplinary Sciences, Göttingen, Germany; [4]present address: School of Basic Medical Sciences, Wuhan University, China.

* Corresponding author:

goerlich@mpinat.mpg.de

Abstract

Nuclear transport receptors (NTRs) carry cargo across the permeability barrier of nuclear pore complexes (NPCs) - an FG phase condensed from disordered but cohesive FG repeats. This phase is repelling to 'normal' macromolecules but a good 'solvent' for NTRs. HIV delivers its capsid through NPCs into nuclei when infecting non-dividing cells. Those capsids are not transported like cargo but cross the barrier like NTRs. We now uncovered the molecular determinants of this NTR-behavior and found that the FG-binding pocket around CA^{N57} is insufficient. Instead, myriads of transient interactions of exposed capsid sidechains allow the capsid's outer surface to be 'wetted' by FG repeats, making capsids soluble in the phase. We have identified a critical set of such residues whose mutation to 'FG-repelling' ones prevents FG phase-partitioning, NPC-targeting, and NPC-passage. We provide direct evidence that CPSF6 releases capsids from NPCs into nuclei, resembling RanGTP in conferring directionality to canonical nuclear import.

Main

Nuclear pore complexes (NPCs) are embedded in the nuclear envelope (NE) to conduct nucleocytoplasmic transport[1-3]. This includes receptor-mediated import of nuclear proteins as well as export of mRNAs or newly assembled ribosomes[4-6]. NPCs are equipped with a barrier that can be rationalized as an FG phase[7-9] – assembled from intrinsically disordered but cohesive FG repeat domains with numerous phenylalanine-glycine (FG) motifs. This barrier is selectively permeable to nuclear transport receptors (NTRs) along with captured transport substrates, while rejecting macromolecules that are not recognized as valid cargo. Peroxisomal protein import appears to follow a similar principle[10].

Among the different FG repeat domains, the FG repeats of Nup98[11,12] appear to be most critical ones for the NPC barrier[13]. It is highly depleted of charged residues, poorly soluble in water and therefore readily phase-separating from aqueous solutions[8]. The resulting condensed FG phase shows essentially the same permeability properties as NPCs themselves[8,14]. It excludes inert macromolecules (such as mCherry), but dissolves NTRs to high partition coefficients. NTRs recognize and bind cargoes[4]. They promote barrier passage of a selected cargo by enhancing its solubility within the FG phase[7].

The importin β superfamily represents the largest NTR class[4,15,16], including both nuclear import receptors (importins), exportins, as well as biportins. Asymmetric transport cycles,

driven by the RanGTPase system, allow them to actively pump cargoes against a concentration gradient[17]. Importins, for example, capture cargo in the cytoplasm, translocate through NPCs, and release cargo into the nucleus upon RanGTP-binding. While the permeability barrier retains imported cargo inside nuclei, importin·RanGTP-complexes return to the cytoplasm, where GTP-hydrolysis releases Ran and allows the importins to import the next cargo molecule.

Importin β-type NTRs are 90-140kDa α-solenoid proteins[16]. Other NTRs include NTF2 (the importer of Ran; ref.[18]), the Mex67-Mtr2 dimer (the main exporter of RNA; ref.[19,20]), and Hikeshi (the importer of Hsp70; ref.[21]). All NTRs bind FG repeats in a multivalent fashion[20,22-24]. This binding counteracts the hydrophobic cohesive interactions between repeats[25] and thus allows an NTR to "melt" through the FG phase[26]. It was recently discovered that the HIV-1 capsid also functions as an NTR[27,28].

HIV is the retrovirus that causes AIDS by infecting and eventually eliminating immune cells[29-31]. Its capsid initially encloses two copies of the HIV genomic RNA[32], along with reverse transcriptase and integrase[33,34]. It is enveloped by a viral membrane that fuses with the plasma membrane during entry into a target cell[35]. This membrane fusion releases the capsid into the cytoplasm, where an influx of dNTPs into the capsid initiates reverse transcription[36]. The resulting vDNA is eventually integrated into a chromosome[37]. Simple retroviruses require mitosis to access chromatin. HIV and other lentiviruses, however, efficiently infect also non-dividing cells with intact NEs[38], suggesting that an incoming HIV genome can somehow cross NPCs.

It has long been believed that the 60 nm x 120 nm sized capsid was too large for NPC passage, that capsid uncoating and release of the HIV genome occurred in the cytoplasm, and that the capsid-free pre-integration complex was the species that is imported into the cell nucleus (discussed in ref.[39]). This view has recently been revised when still intact capsids were detected inside nuclei during the course of infection[40-42], when cryo-EM tomograms showed incoming capsids trapped at NPCs[43] and when the diameter of the NPC scaffold was shown to be larger in intact cells (~60 nm; ref.[44]) than in the earlier analyzed specimens from isolated nuclear envelopes[45]. Finally, mature capsids were discovered to behave like NTRs, efficiently partitioning into an FG phase and targeting NPCs without any *trans*-acting factors[27,28]. This NTR-like behavior required capsid assembly. These findings also resolved the problem that the 60 nm width of the NPC scaffold is just sufficient to accommodate a "naked" but not an importin-coated capsid. Furthermore, they provided a plausible explanation of where on the

transport trajectory the capsid can engage in its critical interaction with the nuclear host factor CPFS6[46-48].

Capsid assembly[49] begins with the translation of the gag polyprotein (reviewed in ref.[50]), which comprises the MA (matrix protein), CA (p24 capsid protein), NC (nucleocapsid) and p6 modules. MA gets myristylated, binds lipids, and targets gag to the plasma membrane[51,52]. CA assembles into the immature capsid[53] in an inositol phosphate-assisted manner[54], while NC recruits HIV genomic RNA to the nascent virions[55]. P6 initiates virus-budding from the plasma membrane[56]. The switch to mature capsids is triggered by the HIV protease[57], which cleaves gag into its individual proteins[58]. This switch involves major rearrangements in the capsomer structures[59]. This elaborate assembly sequence prevents that the capsid gets (mis-) targeted to NPCs already in virus-producing cells.

Results

Not only CA hexamers but also pentamers confer an efficient FG barrier entry

This study focuses on molecular features that allow mature HIV-1 capsids to pass the FG barrier of NPCs and thus deliver the viral genome into the nucleus. This passage should rely on interactions with FG repeats, and indeed a deep binding pocket between the subunits of the CA hexamers has been described which can accommodate the unique FG repeat of CPSF6 or a Nup153 FG repeat[48,60,61]. There, the sidechain amide of asparagine N57 of the CA protein engages in a double H-bond with the backbone of the FG peptide. A mutation of this asparagine to alanine (N57A mutation) abrogates this stable binding[48,62,63]; it impedes NPC passage as well as the partitioning into a Nup98 FG phase (see ref.[27] and figures below). The capsid comprises not only hexamers but also pentamers with a rearranged 'N57 pocket'[64,65]. It has been argued that this rearrangement causes a clash with a bound CPSF6 or Nup153 FG peptide and thus a loss of FG interaction in pentamers. This view was supported by the observations (1) that the CPSF6 FG peptide increases the thermostability of CA hexamers but not of pentamers, and (2) that cryo-EM maps showed an FG peptide density in hexamers but not in pentamers.

Still, we wondered whether pentamers could contribute to FG phase entry. To test this, we used the G60A, G61P double mutant to reconstitute 20 nm T1 capsid spheres, built of 12 pentamers[64]. For comparison, we also assembled capsids from CA wild-type[64,66], yielding large (~80 x160 nm) cone-shaped CLPs (capsid-like particles) dominated by hexamers, as well as

40 nm capsid spheres that comprise 30 hexamers and 12 pentamers[67,68]. These capsid species were all labelled by 15% GFP fusion to the CA C-terminus that points to the capsid's interior. Correct assembly was validated by negative-stain EM (Fig.1a).

As a proxy for the NPC barrier, we assembled an FG phase from Nup98-type, perfect GLFG 12mer repeats[69]. CLPs and the 40 nm capsid spheres accumulated in this phase to very high partition coefficients of 1000-2000 (Fig.1c-d) consistent with our previous observations[27]. Surprisingly, the pentamer-only spheres showed the same behavior and reached an intra-phase partition coefficient of ~750 (Fig.1c-d). This partitioning was specific, as the partition coefficient of unassembled protomers was 300 times lower, while mCherry (our internal, inert control) or non-fused GFP remained completely excluded (partition coefficients of ≤ 0.1). The pentamer-only spheres also accumulated to similarly high partition coefficients in FG phases with other common FG motifs (SLFG or FSFG; Extended Data Fig.1).

In a complementary assay, we incubated labelled capsids with digitonin-semipermeabilized HeLa cells and observed a highly efficient NPC-targeting not only of CLPs and 40 nm spheres, but also of the pentamer-only spheres (Fig.1e) - as indicated by their striking colocalization with an anti-Nup133 nanobody[70].

NPC-targeting of the capsid can occur independently of the Nup358 CypH domain

NPC-targeting could be mediated by direct FG interactions and/or by the CypH (cyclophilin homology) domain of Nup358/ RanBP2, which is known to bind the 'cyclophilin-loop" (loop 2) of the capsid[71]. To disentangle these potential contributions, we repeated the NPC-targeting experiment using the XTC-2 cell line[72] from *Xenopus laevis* – a species, which (apparently like all modern amphibians) has lost the CypH domain from its Nup358 (Extended Data Fig.2). Nevertheless, we observed efficient NPC-targeting of all three capsids (Fig.1f). This indicates a highly efficient capsid entry into the FG phase of NPCs. Remarkably, this was also true for the pentamer-only capsids, confirming that CA pentamers mediate productive FG interactions. The experiment also showed that the capsid-FG phase interactions are independent of species-specific FG domain features.

Crucial FG interactions by the N57-pocket of CA pentamers

The re-arrangement of the N57 FG-binding pocket in CA pentamers[64,65] might imply that asparagine N57 may be irrelevant for the FG phase interactions of pentamer-only spheres.

However, we observed the contrary (Fig.1g-h). The N57A mutation impeded FG phase entry not only of the hexamer-dominated CLPs and 40 nm spheres, but also of the pentamer-only spheres - to the extent that only faint staining on the FG phase surface remained. This strongly suggests that FG peptides can also dock into the N57 pockets of pentamers, with the N57 sidechain amide engaging in an energetically relevant hydrogen bond. We interpret the previous negative results[64,65] to mean that the pentameric N57 pocket binds FG repeats just in a more transient manner.

FG phase surface-arrest phenotype

To immerse into the FG phase, the capsid must locally melt cohesive interactions between FG repeats. This melting is associated with a ΔG penalty[25]. Full entry into the FG phase therefore requires FG-capsid interactions to release more free energy than this penalty. If less free energy is released, then the capsid arrests at the phase surface, where 'free' FG motifs are directly exposed to the aqueous phase.

N57A mutant capsids show this 'surface arrest' phenotype with all FG phases tested (Fig.1g-h and Extended Data Fig.3d-f), indicating a thermodynamically relevant FG interaction defect and illustrating the relevance of the two N57-mediated H-bonds for locking the peptide backbone of FG peptides. These H-bonds are evident in solved structures of CPSF6 and Nup153 FG peptides[60] as well as in AlphaFold models of CA capsomers with captured GLFG, SLFG, or FSFG peptides (Extended Data Fig.3a-c).

The amino acid composition of the highly exposed capsid's outer surface is very unusual

The residual surface-binding of the N57 mutant capsid indicates, however, that the mutant has not lost all FG interactions and that other capsid features are also relevant. In search of additional contributions, we considered a very different mode, namely the collective interactions between the surface of a client and the FG phase, which we had conceptualized and observed previously[14,73]. This concept considers (1) that the FG phase is a solvent for its clients, (2) that the entire surface of the clients, i.e. any sufficiently exposed residue, comes into contact with FG repeats when dissolved in the phase (with ~400 mg/ml FG mass[25]), and (3) that these contacts can be energetically favorable, neutral, or unfavorable (or repelling) as compared to the contacts with water.

Based on this concept, we engineered GFP to pass NPCs either very slowly or very rapidly, with extreme variants spanning a 15 000-fold difference in rate, and with a nearly perfect correlation between rates and FG phase partitioning[14]. This engineering exercise also uncovered an amino acid scale for FG interactions (Fig.2a): Negative charges (Asp and Glu) as well as lysine are strongly FG-repelling, while exposed hydrophobic residues, cysteine, methionine, histidine and also arginine attract GFP (or other clients) into the FG phase.

The FG-repelling residues Asp, Glu and Lys are typically exposed at the surface of soluble globular proteins (Fig.2). They provide topological information during the folding process and keep these globular proteins water soluble. They frequently occur in solvent-exposed parts of α-helices, β-sheets, and loops. On the protruding part of the outer surface of the HIV-1 capsid, however, they are extremely depleted (Fig.2b-e). In fact, loop 1 (the β-hairpin), loop 2 (the CypA-binding loop) and loop 3 do not contain a single Asp, Glu, or Lys (Fig.2c,e). This compositional bias is so striking that we suspected a connection to the NTR-like behavior of the capsid.

The lack of FG-repelling residues on the capsid's outer surface is key to FG phase entry

To explore this possible connection experimentally, we analyzed 13 additional capsid mutants that had previously been reported[46,74-77]. In brief, we assembled CLPs from the respective mutant CA proteins, validated the faithful assembly by size exclusion chromatography and negative-stain electron microscopy (Extended Data Fig.4), and probed the Nup98 GLFG phase-entry of these CLPs (which had been filled with sinGFP4a as a tracer for intact capsids).

Most mutants entered the FG phase as efficiently as wild-type CLPs (Fig.3a and Extended Data Fig.5a). No decrease in partitioning was observed when poorly exposed residues (other than N57) were mutated or when an exposed residue was changed to an equally or more hydrophobic one, examples being the G89V and P90A loop2 mutations, which were previously designed to abrogate a cyclophilin interaction with cyclophilin A/ the CypH domain[77]. The same was true for the E45A, G116A, L136M, or R143A mutations.

The H87Q and R132K mutations caused a ~50% reduction in partitioning. The A92E and G94D mutations in the highly exposed loop 2, however, showed a partition defect similarly strong as the N57A exchange. They arrested the capsids at the FG surface and reduced the partitioning to ~5% of the wild-type level (Fig.3a-c). This is consistent with the fact that the negatively charged aspartic and glutamic acid residues belong to the most FG-repelling surface

features, which in turn can be explained by the energetically rather unfavorable contacts between negative charges and phenylalanines.

To explore the outer surface of the capsid more systematically, we first identified additional positions that could be mutated without compromising capsid assembly. For this, we again used size exclusion chromatography and negative stain electron microscopy to validate faithful capsid assembly (Extended Data Fig.4). "Stable capsid mutants" were subsequently tested in transport assays (Fig.3d, and see below). A drastic drop in capsid-partitioning into the FG phase was evident for six individual mutations to an FG-repelling glutamate (Fig.3d and Extended Data Fig.5b); these included the Q9E mutant (an isosteric change at the tip of the β-hairpin of loop 1), the V86E, A88E, I91E, and M96E mutations (all at loop 2), as well as the P123E exchange (at loop 3).

Moreover, the R97K mutation impeded the FG phase partitioning. It preserves a positive charge; however, the two amino acids behave differently towards the FG phase[14]. Lysine is FG-repelling. Arginine is FG-attracting, the explanations being (1) that its planar guanidinium group engages more readily in a cation-π interaction with phenylalanines of the FG repeat domain, (2) that it is an excellent hydrogen bond donor for carbonyl oxygen acceptors in sidechains and the polypeptide backbone (discussed in ref.[25]), and (3) that arginine is more readily transferred from an aqueous to a hydrophobic environment[25,78]. The partitioning defect of the R97K capsid mutant reflects these differences remarkably well.

The already mentioned R132K mutant showed a similar trend but its defect was much weaker, probably because R132 is less exposed (Extended Data Fig.6). This appears to be a general pattern: R143 is located even deeper in the inter-capsomer cavity, and the R143K mutation had no effect on capsid-partitioning. Likewise, mutating leucine L136 at the bottom of the inter-capsomer cavity to glutamate (L136E) had no effect on phase entry, while the similar mutation of the highly exposed alanine 92 (A92E) was very detrimental (compared in Extended Data Fig.6). This non-equivalence emphasizes that the solubility of the HIV capsid in the hydrophobic FG phase is ruled by the surface properties of its protruding and highly exposed parts.

For the FG partitioning mutants studied here, one could argue that the effects are not due to the newly introduced FG-repelling residues, but to the loss of the original ones. To address this, we next analyzed different substitutions for the same residue (Fig.3e and Extended Data Fig.5c). The mutations of histidine 87 in loop 2 to glycine (H87G) reduced the FG partitioning of the capsid only marginally. The H87Q exchange had a stronger but still mild effect. The H87E

mutation, however, was detrimental. It reduced the partitioning ~15-fold and left transport intermediates arrested at the FG phase surface. Thus, the defect cannot be explained by the loss of the imidazole moiety alone. What mattered was indeed the exchange to the FG-repelling glutamate with its negatively charged sidechain. The same pattern was observed for mutations of V86, A88, I91, A92, and M96: Exchanges to glycine had only mild effects, while mutations to glutamate reduced the capsid-partitioning >10-fold and led to surface-arrested intermediates.

Additive effects of capsid-partitioning mutations

Up to this point, we have identified 12 individual capsid mutants with a clear defect in FG phase partitioning. Yet, all of them retained some FG interactions. A combination of partitioning mutations, however, further reduced the FG phase signal of the capsids to the point of complete loss of partitioning and surface binding. This was the case when three FG-repelling surface mutations (H87Q+A92E+G94D) were combined with the N57A pocket mutation (Fig.4a and Extended Data Fig.7a).

Likewise, there is also a clear synergy between surface mutations alone, as seen by the greatly reduced signal of the Q9E+A92E double mutant CLPs (Fig.4b and Extended Data Fig.7b). The V86E+A88E+I91E+A92E quadruple mutant capsid combines four FG-repelling loop2 mutations. It has lost all FG phase-partitioning and surface-binding. This is remarkable, because not only does the quadruple mutation keep the N57 pocket intact, but loop 2 does not even contact the FG repeat portion that docks into the N57 pocket - neither in the reported CA hexamer structures with bound CPSF6 or Nup153 FG peptides (PDB 4WYM and 4U0C) nor in structural models with other FG peptides (Extended Data Fig.3). Thus, the four negative charges dominate in their repulsion from the condensed FG phase – not only with respect to the still intact N57 pocket, but also with respect to all the other FG-attracting residues that still cover the outer surface of the capsid (e.g. H87, G94, M96, R97, or P123).

Capsid mutations that impede FG phase-partitioning also interfere with NPC-targeting

In a next step, we compared the targeting of wild-type and mutant CLPs to NPCs of digitonin-semipermeabilized HeLa cells (Fig.5). To label intact capsids, we again loaded them with sinGFP4a. We not only reproduced the finding that the N57A pocket mutation reduced the NPC-targeting, but also observed that all of the FG-repelling surface mutations had an at least equally deleterious effect. In fact, most of these mutations (e.g. A88E, A92E, or M96E)

impeded capsid-binding to NPCs even more. Furthermore, the effects of the mutations were, again, additive. The V86E+A88E+I91E+A92E and the N57A+H87Q+A92E+G94D quadruple mutations reduced capsid-binding to NPCs to background levels.

Large CLPs completely pass mouse oocyte NPCs when CPSF6 is overexpressed

We previously microinjected 40 nm capsid spheres into the cytoplasm of mouse oocytes[27]. These spheres passed NPCs and accumulated at intranuclear structures that probably represent nuclear speckles (or paraspeckles). When we repeated the experiment with large CLPs (covalently labelled with GFP), we observed only prominent binding to the nuclear envelope, respectively NPCs, but no intranuclear signal (Fig.6a, 'WT'). This can be explained by a very slow NPC passage. Indeed, the same capsids showed prominent speckle-binding when injected directly into the nucleus.

We reasoned that capsid release from NPCs into the nucleoplasm was rate-limiting and that this slow step could be accelerated by increasing the concentration of the host factor CPSF6[46-48,79]. To test this, we microinjected mRNA encoding either mScarlet or a CPSF6-mScarlet fusion. Four hours later, GFP-labelled CLPs were microinjected into the cytoplasm (along with a microinjection marker). Another 30 minutes later, the oocytes were imaged by confocal laser scanning microscopy. This revealed efficient translation of the mRNAs: non-fused mScarlet was evenly distributed throughout nucleus and cytoplasm, sparing only the nucleolus, while the mScarlet-CPSF6 fusion was efficiently imported into the nucleus, giving a clean nucleoplasmic signal with a notable speckle enrichment.

Strikingly, when the oocyte nuclei had been pre-loaded with CPSF6, then cytoplasmically injected CLPs not only bound to the nuclear envelope, but also passed NPCs and accumulated in nuclear speckles (Fig.6b). A complementary experiment revealed that the CPSF5-CPSF6 complex is a potent antagonist of capsid-partitioning into the FG phase (Fig.6c). Thus, CPSF6 appears to promote the completion of NPC passage by releasing the capsid from the FG phase.

FG-repelling surface mutations block NPC passage of capsids

The supplementation of mouse oocytes with CPSF6 now allowed testing capsid mutations for NPC passage defects. For this, we pre-loaded oocytes with unlabelled CPSF6, before microinjecting GFP-tagged wild-type or mutant CLPs into the cytoplasm. Wild-type CLPs reached the nuclear speckles, whereas N57A mutant ones did not (Fig.6d). The mutant effect

was confirmed by two controls. First, co-injected wild-type CLPs, labelled with mScarlet-I3, reached the speckles of the same oocyte. Second, when microinjected into the nucleus, also the N57A mutant CLPs accumulated at the speckles (Fig.6a). Thus, the N57A mutation causes a genuine NPC passage effect, even though capsid-targeting to the nuclear envelope was still prominent.

The above characterized V86E+A88E+I91E+A92E surface mutant combines four exchanges of hydrophobic residues to an FG-repelling glutamic acid. After nuclear microinjection, it effectively bound speckles (Fig.6a). Following cytoplasmic microinjection, however, it failed to reach intranuclear structures. It did not even show any detectable binding to the nuclear envelope, indicating a very tight block in NPC-targeting and NPC-passage. This was obviously a consequence of the failure of these mutant capsids to partition into the FG phase (Figs. 4-5). Thus, the presence of FG-attracting and the striking absence of FG-repelling residues in the capsid's β-hairpin and its two exposed loops are key determinants for both, entering the permeability barrier of NPCs and passing through the pores completely.

Discussion

The HIV-1 capsid has evolved into a nuclear transport receptor (NTR) that readily partitions into the FG phase-based permeability barrier of NPCs. It serves as a cargo container for the encapsulated genetic material, shielding it from anti-viral sensors in the cytoplasm, delivering it through an intact nuclear envelope into the nucleus and thus allowing the infection of non-dividing cells.

Both, hexameric and pentameric capsomers contribute to FG phase-partitioning

The capsid is very large – at the very limit of what can possibly pass through an NPC scaffold. It is cone-shaped and typically made up of 200-250 hexameric and 12 pentameric capsomers[80-85]. It was previously thought that only the hexamers engage in FG interactions[64,65]. We have now found that pentamers (represented by pentamer-only T1 spheres[64]) mediate a very efficient FG phase entry as well (Fig.1). Indeed, CA hexamers and pentamers appear very similar in their FG interactions: Efficient NPC-targeting and FG phase-partitioning of both species require assembly into a capsid structure (Fig.1c-f). Furthermore, point mutations that impede the FG partitioning of hexamer-dominated capsids also compromise the partitioning

of the pentamer-only T1 capsid; this applies to the N57A mutation as well as to the surface loop mutations (Fig.1g, Fig.4c, and Extended Data Fig.7c, Extended Data Fig.8a-b).

Although low in count, the contribution of pentamers is probably quite relevant, because they are concentrated at the capsid's narrow end, which appears to insert first[43] into NPCs and which should therefore have a particularly high FG-partitioning propensity.

The condensation of FG repeat domains into an FG phase is driven by cohesive interactions that are of a hydrophobic nature[25,26,73,86,87]. Partitioning of a client into this phase requires a local melting of these cohesive interactions. This easily happens when the client (e.g. an NTR) binds (the hydrophobic) FG motifs. Indeed, NTRs belong to the most hydrophobic soluble proteins[73]. This poses a fascinating design challenge, namely how to render a protein surface hydrophobic in an 'NTR-like' manner without risking aggregation and non-cognate interactions?

FG-binding pockets and FG-wettable surface

Binding pockets for FG motifs are one solution. They have been identified by X-ray crystallography and cryo-EM in several cellular NTRs, such as importin β, Xpo1/ CRM1 or the Mex67-Mtr2 dimer[20,22,24,88,89]. FG-binding pockets can confer excellent binding specificity without fully exposing their hydrophobicity. They contain mostly hydrophobic, but also polar residues, whose aliphatic moieties then engage in hydrophobic interactions with FG motifs, while their polar head groups hydrogen-bond, e.g., to the FG repeat backbone[20,22,89]. FG-binding pockets appear to confer rather strong FG interactions that can overcome the FG-repelling effects of the typically strong negative charge of cellular NTRs, which in turn is necessary to transport also highly positively charged cargoes, such as histones. It seems reasonable to assume that these details of FG interactions have evolved to optimize cargo transport.

The HIV-1 capsid also has a pocket. It occurs in very high multiplicity; it is located between adjacent CA monomers and is much deeper than FG-binding pockets in cellular NTRs. It makes hydrophobic contacts to the FG motif phenylalanine by the sidechains of L56, M66, L69, and I73 as well as by the aliphatic portion of K70. The sidechain amide of asparagine N57 engages in a double or triple hydrogen-bond with the backbone of the inserted FG repeat (ref.[60]; Extended Data Fig.3a-c). N57 and L56 are absolutely conserved amongst HIV-1, HIV-2 and SIV (Extended Data Fig.9). K70, M66, L69, and I73 are also conserved but show some

conservative (K→R; M→L; L→I/V; I→V) exchanges (Extended Data Fig.9). This suggests an ancestral function of this pocket and possibly a common mechanism of NPC passage of lentiviral capsids (see also below).

This pocket is, however, only one part of the story. The entire accessible outer surface of the mature capsid appears to be optimized for "wetting" by the FG phase (Figures 1-4). It has a highly unusual amino acid composition that follows a previously identified scale for FG phase entry[14]. This includes a complete absence of the FG-repelling residues Asp, Glu and Lys in the exposed loops. This absence is relevant because the introduction of repelling glutamates impeded FG phase-partitioning, NPC-targeting and also NPC passage (Figures 3-6) - up to a complete block with combined mutations. We have documented these phenotypes at nine positions: Q9 (loop1), V86, A88, H87, I91, A92, G94, M96 (loop2) and P123 (loop 3). A similar phenotype was observed for the R97K mutations, where an FG-attracting arginine was replaced by an FG-repelling lysine. None of these mutations have been observed in HIV-1 isolates from patients (Extended Data Table 1), although we assume that they still allow infection of proliferating cells, where a capsid passage through NPCs is not a strict requirement.

One should expect an enhanced FG phase-partitioning if capsid surface residues were changed to more FG-attracting ones. Such an effect is not obvious with our standard GLFG phase that already fully absorbs the wild-type CLPs. However, such enhancement is evident for a stricter FSFG phase, which does not allow full entry of the wild-type CLPs, while G89V mutant ones enter this phase 4-fold better (Extended Data Fig.8c-d).

N57 engages in a polar (H-bond) interaction with the FG repeat backbone, while the capsid surface residues analyzed here probably contribute hydrophobic contacts. The different qualities of these interactions become experimentally apparent when the salt concentration is varied in the FG phase assay (Extended Data Fig.10). Wild type CLPs enter the GLFG phase at low (50 mM), medium (150-250 mM) or high salt (600 mM NaCl). The N57A mutant fails to partition at low and medium salt, but enters the phase at 600 mM NaCl, probably because this increases the strength of the remaining hydrophobic capsid surface-FG interactions. The X→E surface mutations, however, could not be rescued by elevated salt. In fact, the Q9E mutant even shows the inverse salt effect.

A hydrophobic and yet aggregation-resistant capsid surface

The three outer loops are dominated by FG attractive residues. However, the most attractive ones, tryptophan, tyrosine, and phenylalanine are suspiciously missing (Fig.2, Extended Data Table 1). Why is this? Perhaps, because highly exposed tryptophans or tyrosines also pose the danger of highly promiscuous interactions. We observed this in our previous engineering study on transforming GFP into an NTR[14]: Placing additional tryptophans or tyrosines onto the GFP surface greatly increased the partitioning into an FG phase and the rate of NPC passage, but also conferred massive aggregation with nuclear and cytoplasmic structures.

By rational design, followed by mutagenesis and screening many variants, we eventually obtained highly NPC-/ FG phase-specific and rapidly translocating NTR-like GFP variants[14], but strikingly, they all lacked fully exposed tyrosines and tryptophans. We find it remarkable that the evolution of the capsid into a specific NTR has arrived at a similar solution.

This must have happened under strong selective pressure, since aggregation propensity scales with size and because thousands of cytoplasmic/ nuclear proteins were potential aggregation partners. Here, we see fascinating parallels to an antibody response by the immune system, which not only aims at strong target-binding but also against broad cross-reactivity with myriads of undesired (self-) targets. Antibodies observed at an early stage of the immune response often contain exposed tyrosines in their CDRs, perhaps because tyrosine-containing CDRs have such high propensity to engage in interaction[90]. In well affinity-matured antibodies, however, tyrosines are typically replaced by more sophisticated and target-specific hydrophobic arrays[91]. In this sense, we consider the HIV-1 capsid as a well selectivity-matured entity – a masterpiece of evolution, created by a very deep sampling of sequence space and by "testing" extremely large numbers of variants during infections.

Evolutionary pressure against surface lysines

To pass NPCs and thereby provide the selective advantage of infecting non-dividing cells, the capsid had to be highly optimized. Could this have evolved in a single step? Or did some other kind of evolutionary pressure pave the way? Possibly, at least as far as the suspicious absence of accessible lysines is concerned. Lysines are not only FG-repelling, they are also acceptors for ubiquitin conjugation. We assume that their elimination from the accessible outer surface (Fig.2d) came first – to protect incoming capsids from proteosomal degradation. Indeed, even

the CA N-terminus is refractory to ubiquitin-modification, since it is a secondary amine (proline) and is buried in the mature capsid[81,82,84].

Conservation of capsid features

The FG-attractive capsid signature is well conserved in lentiviruses. This not only applies to the N57 FG-binding pocket (see above), but also their outer capsid loops have a similar amino acid composition as HIV-1 (Extended Data Fig.9). We suspect that they all share the ability to partition into the FG phase and pass NPCs; at least our preliminary data with SIV capsids point in this direction. This would explain why lentiviruses are generally able to infect non-dividing cells. In contrast, the FG-attracting loop 2 is missing in simple retroviruses, such as the mouse leukemia virus MLV[92,93]. This suggests that the same way of NPC passage is not available to them and would explain why they infect quiescent cells only with marginal efficiency[38,94].

Escape from the energy well of the FG phase

The HIV-1 capsid partitions very strongly into the FG phase, so strongly that the outside signal remains non-detectable (refs.[27,28]; Figures 1,3,4**,** and Extended Data Figures 1,3,5-8,10). Thermodynamically, this state can be described as a deep energy well and it is an intriguing question of how the capsid can possibly escape from there. Cellular NTRs of the importin β superfamily use an energy input from RanGTPase, which helps at least to release cargoes from the NTRs and thus from the NPC barrier[4,69]. This release is triggered by binding of nuclear RanGTP to importins or by a cytoplasmic GTP hydrolysis in cargo·exportin·RanGTP complexes. The capsid, however, has no means of interacting with RanGTP. HIV must therefore use a different strategy to release its genetic material from NPCs. The following possibilities come to mind:

(1) Destruction of NPCs. It has been observed that the capsid can cause cracks in the NPC scaffold in some cell types[95]. However, since the NPC scaffold elements and thus (at least most of) the FG mass remain in place, it is difficult to see how the cracks could allow the capsid to escape from the FG phase. We therefore regard these cracks to be collateral damage. In fact, they may be a cell type-specific problem and not a requirement for capsid passage through NPCs: Indeed, the nucleocytoplasmic barrier of mouse oocytes remained intact and kept excluding our injection marker even after flooding these cells with large capsid species and after capsids had passed NPCs and reached the nucleoplasm (Fig.6a, b, d).

(2) Capsid disassembly leading to vDNA release. Since FG phase-partitioning is favored by capsid assembly (refs.[27,28]; Fig.1 and Extended Data Fig.1), the inverse should also hold true, namely capsid stabilization by immersion into the FG phase. This does not rule out the possibility that an FG phase-trapped capsid might burst, e.g. when reverse transcription has increased the volume of the enclosed nucleic acids[96,97]. However, this could also release the contents on the cytoplasmic side.

(3) The most plausible mechanism, however, is an energy input from a nuclear binding reaction that disengages the FG nucleoporin repeats from the capsid. Such a mechanism would allow for a directional capsid release into the nucleoplasm (Fig.7). Indeed, CPFS6 is a perfect candidate for terminating the capsid's NPC passage. It is a host factor for infection; it binds the capsid directly, it is a nuclear protein and participates in the nuclear events during the establishment of infection[46-48,79]. We have now demonstrated that the CPSF5-CPSF6 complex potently antagonizes the partitioning of capsids into the FG phase (Fig.6c) and promotes capsid release from NPCs into the nucleus: when large CLPs were microinjected into the cytoplasm of mouse oocytes, they accumulated as tightly NPC-arrested transport intermediates. However, upon CPSF6 over-expression, capsids were released and appeared at nuclear speckles (Fig.6).

It may be surprising that capsid release from NPCs was so slow when only endogenous levels of CPSF6 are available. However, late-stage oocytes might have lower CPSF6 levels than typical HIV-1 target cells. Reaching an effective concentration thus requires the exogenous supply. Possibly, CPSF6 can release large CLPs from NPCs only if it exceeds the threshold concentration to form the recently reported local co-condensates with capsids[79,98].

At non-limiting CPSF6 levels, the capsid passage through NPCs appears to be quite efficient, in particular when considering that we loaded the oocytes with far more capsid than during a genuine infection. We also acquired images already 30 minutes after capsid injection, which is fast for a transport experiment in oocytes, where diffusion distances and thus diffusion times far exceed those in somatic cells. It is also fast compared to the estimated ~1.5 hours dwell time at NPCs during a genuine infection[99]. The kinetics of CPSF6-mediated capsid release from NPCs and accumulation at speckles (Fig.6) thus appear to be within a plausible and physiologically meaningful range.

To extract the capsid from its highly multivalent FG phase interactions, CPSF6 must bind more efficiently than the FG motifs of genuine nucleoporins and compete against the extremely high local FG concentration in a condensed FG phase. The buried N57 double hydrogen bond to the CPSF6 peptide backbone certainly contributes to a strong binding[48], but it does not discriminate:

The Nup153 FG repeat-CA hexamer structure[100], Alphafold-modelling and our FG phase partitioning data (Fig.1 and Extended Data Fig.3) all suggest that 'normal' FG repeats hydrogen-bond with N57 in the same way. Compared to a normal FG motif, however, the CPSF6 FG repeat (PVLFPGQPFGQPPLG) is more hydrophobic, makes more hydrophobic contacts to the CA chains, and binds as a more compact hydrophobic cluster, which stabilizes its bound conformation. We assume that these details and the avidity-enhancing effects of additional capsid-interacting linear motifs in CPSF6, the multimeric state of the CPSF5-CPSF6 complex[101], and the high local concentration in a CPSF6 condensate make the difference (Fig.7).

The cellular function of the FG peptide in CPSF6 is unclear, but it was obviously key to the viral repurposing of CPSF6 as an NPC-release factor for the capsid (Figures 6-7). The nuclear localization of CPSF6 ensures that this release occurs into the nucleus. It makes thus the capsid transport directional. In this respect, CPSF6 appears to be analogous to RanGTP in canonical nuclear import pathways[4], where Ran acts as a nucleus-specific release factor for cellular cargoes from importins and thus from NPCs.

Figure legends

Fig.1 | Hexameric and pentameric capsomers mediate FG phase-partitioning and NPC-targeting

a, Negative stain electron micrographs of the HIV-1 capsid species used.

b, Concept of the FG phase as a permeability barrier of NPCs, reconstitution of such a barrier by phase separation of barrier-forming cohesive FG domains, and testing of its transport properties.

c, A Nup98-type FG domain comprising 52 perfect 12mer GLFG repeat units ("prf.GLFG$_{52x12}$")[69] was allowed to phase separate, forming FG phases of near-spherical shape. Indicated fluorescent species were subsequently added and detected one hour later by confocal laser scanning microscopy (CLSM). Free mCherry and EGFP remained well excluded from the FG phase, with partition coefficients (in: out) of ≤0.1. Large CLPs, 40 nm spheres, and 20 nm T1 'pentamer-only' spheres were each labeled by ~15% CA with a C-terminally fused EGFP, pointing to capsids' interior. They accumulated essentially completely inside the FG phase, reaching partition coefficients of ~1000. This was 30-300 times higher than the corresponding monomeric CA-EGFP fusions and 10 000 times higher than free EGFP.

d, Quantification of **c** but with larger datasets. Each datapoint is the integrated inside (in) signal of one FG particle and that of a corresponding outer (out) region. Numbers are means; bars indicate mean ± s.d.; n = number of quantified FG particles and outer areas, respectively.

e, f, Specific targeting of capsid species to NPCs. Human HeLa cells or frog XTC-2 cells (whose Nup358 lacks a CypH domain) were grown on coverslips, treated with digitonin to permeabilize their plasma membranes, and incubated with the indicated fluorescent species. Confocal scans were taken directly through the live samples one hour later. Laser settings were individually adjusted. NPCs were detected with an Alexa647-labelled anti-Nup133 nanobody[70]. Note that conical CLPs, 40 nm capsid spheres, and pentamer-only spheres showed highly efficient NPC-targeting in both cell types – comparable to nanobody-staining. By contrast, unassembled CA monomers showed no NPC enrichment and were evenly distributed throughout the cells.

g, h, FG phase-partitioning of all three capsid species is strongly reduced by the N57A mutation. Experiments and quantifications were performed as in **c, d** but signals at the phase surface and the center were integrated separately (see also below, Fig.3b,c). Scan settings were identical for each wildtype-N57A mutant pair. Experiments were repeated independently with identical outcomes ($N = 4$). Scale bars, 200 nm in **a**, 10 µm in **c, e, f,** and **g**.

Fig.2 | Surface amino acid composition of the HIV-1 capsid

a, An amino acid scale for partitioning into an FG phase[14]. FG-repelling residues are coloured in red, FG-neutral residues in yellow, and FG-attracting residues in green or dark green.

b, Comparison of the mCherry, GAPDH and HIV-1 capsid surfaces with residues coloured according to the FG phase partitioning scale. Note that FG-repelling residues (shown in red) are very abundant on the surface of mCherry, GAPDH (an example of a cytoplasmic mass proteins), and on the inner surface of the capsid; but they are absent from protruding parts of the outer capsid surface. There are only a few in remaining outer surface, and these are not well accessible but mostly salt-bridged to FG-attracting R and H (E113-R97; E98-H84/ H87).

c, Amino acids on the exposed loops of HIV-1 capsid are either FG-attracting or FG-neutral.

d, Comparison of lysines exposed on the surfaces of mCherry, GAPDH and HIV-1 capsid.

e, Diagrams compare total and surface amino acid compositions, classified by the scale of **a**.

Fig.3 | Identification of capsid mutations that impede FG phase-partitioning

Partitioning experiments into the GLFG 12mer FG phase were performed as in Fig.1c, but using indicated (wildtype or mutant) CLPs non-covalently filled with sinGFP4a. Experiments were repeated independently with identical outcomes (N = 3). Scale bars, 10 µm. Negative stain electron micrographs, validating proper capsid assembly, are shown in Extended Data Fig.4.

a, Comparison of wildtype CLPs and CLPs carrying previously reported mutations. Note that the N57A, A92E and G94D mutant CLPs show a strong FG phase-partitioning defect.

b, Illustration of the observed FG phase-partitioning phenotypes.

c, Illustration of the quantification strategy. Fluorescence signals for the capsid species were integrated separately at the rim and in the center of FG particles, normalized to the wildtype values, and listed for each mutant in **a, d** and **e**. For more detailed quantifications, see Extended Data Fig.5.

d, Drastic FG-partitioning defects in rationally designed capsid mutants, where FG-attracting residues had been exchanged for FG-repelling ones (E or K).

e, Selected CA positions were mutated to G, Q, or E, as indicated. Drastic FG phase-partitioning defects were evident only when an FG-repelling glutamate was introduced. Mutations to an FG-neutral glycine or glutamine had mild effects at best.

f, A CA-hexamer (PDB 4WYM) viewed from the side and the capsid's outside. Surface mutant positions with FG-partitioning defects are colored in green, the N57 FG-binding pocket in cyan.

Fig.4 | Synergy between 'loss-of-FG-partitioning' mutations in hexamer-dominated CLPs and pentamer-only spheres

a, b, Assay was performed as in Fig.3, analyzing WT CLPs and CLPs carrying the indicated single or multiple mutations. Individual mutations caused already strong FG phase-partitioning defects. Combining mutations aggravated the defect up to a complete loss of partitioning.

c, The assay was performed as in **a** and **b**, the difference being that pentamer-only spheres with a covalent EGFP-label were analyzed. These showed a very similar response to the mutations as the CLPs.

Experiments were repeated independently with identical outcomes (N = 3). More detailed quantifications are shown in Extended Data Fig.7. Scale bars, 10 µm.

Fig.5 | NPC-binding requires a lack of FG-repelling residues on the capsid surface

a, Targeting of conical CLPs (non-covalently filled with sinGFP4) carrying indicated mutations to NPCs of HeLa cells was as in Fig.1e. Scan settings were identical for wild-type (WT) and all mutations. **b,** Quantification of GFP signals on NPCs with larger datasets. Numbers are means; bars indicate mean ± s.d.; n = number of quantified nuclei. *P*-values were calculated by one-way ANOVA with Tukey's post hoc test. *P-values*: ****<0.0001. Experiments were repeated independently with identical outcomes (N = 3). Scale bar, 10 µm.

Fig.6 | Requirements for NPC-passage of CLPs

a, Conical CLPs, carrying a C-terminal EGFP label and indicated mutations, were microinjected into fully grown mouse oocytes, and imaged 30 min later. The injected compartment, nucleus (N) or cytoplasm(C), was marked by co-injected Alexa647-labelled MBP dimers that do not cross the nuclear envelope (NE). Wildtype, or N57A, or V86E+A88E+I91E+A92 quadruple surface mutant CLPs accumulated in nuclear speckles after nuclear microinjection. This is a crucial control for the experiment of panel **c**. None of the large CLPs reached the speckles when microinjected into the cytoplasm. WT and N57A CLPs accumulated at the NE, the quadruple surface mutant did not – consistent with its failure to partition into an FG phase (see Fig.4b).

b, mScarlet or a CPSF6-mScarlet fusion were over-expressed from mRNA that was microinjected into mouse oocytes. After four hours, EGFP-labelled wildtype CLPs were microinjected into the cytoplasm and imaged 30 minutes later. The CPSF6 fusion allowed the CLPs to pass NPCs and to accumulate in speckles, where they co-localized with CPSF6.

c, The CPSF5·CPSF6 complex is a potent antagonist of capsid-partitioning into an FG phase. The assay was performed as in Fig.3, but CLPs were pre-mixed with a recombinant CPSF5-CPSF6 complex before adding to the GLFG phase. Molar ratio was calculated for a tetrameric (A_2B_2) CPSF5·CPSF6 complex.

d, Unlabelled CPSF6 was overexpressed from microinjected mRNA for four hours. Then, mScarlet-I3-labelled wildtype CLPs were co-injected with indicated EGFP-labelled mutant CLPs into the cytoplasm, and imaged 30 minutes later. The mScarlet-I3-labelled wildtype CLPs bound NPCs, crossed the NE, and accumulated in speckles. EGFP-labelled wildtype CLPs behaved the same. N57A mutant CLPs bound to the NE, but failed to pass. Quadruple surface mutant CLPs failed to bind NPCs/ the NE and failed to accumulate in speckles – again consistent with their failure to partition into an FG phase. Experiments were repeated independently with identical outcomes ($N = 3$). Scale bars, 10 µm.

Fig.7 | Scheme of HIV-1 capsid passage through NPCs

To reach the nucleus, the HIV capsid has to dock to NPCs first. The initial interaction might be a capsid-CypH interaction or a direct targeting to the FG barrier. The capsid then fully partitions into the FG phase, which probably requires disengagement from the CypH domain. CPSF6 promotes capsid egress from NPCs into the nucleus through a competing FG peptide and condensation into a CPSF6 phase that has a sufficient local concentration to extract the capsid from the NPC barrier.

Methods

DNA sequences for recombinant protein expression

All recombinant proteins used in this study were produced in *E.coli*, using codon-optimized genes, His_{14}-NEDD8 or His_{14}-SUMO tags and a purification strategy that includes binding to a Ni^{2+} chelate matrix and proteolytic release by NEDP1 or SenP1/ Ulp1[102,103]. Expression vectors are listed in Supplementary Tables S2-S5.

Assembly and purification of conical CLPs

His_{14}-bdSUMO tagged CA^{P1A} was expressed in NEB Express *E. coli* cells (New England Biolabs, C2523). Here, the P1A mutation is required for tag-removal, i.e. to to allow SUMO cleavage. Induction was with 0.1 mM IPTG (isopropyl-β-thiogalactoside) at 18°C for 16 hours. Cells were harvested by centrifugation, resuspended in lysis buffer (50 mM Tris/HCl pH 8.0, 300 mM NaCl, 20 mM imidazole, 1 mM TCEP), and lysed by a freeze-thaw cycle followed by sonication. The lysates were cleared by ultracentrifugation, the soluble fractions were bound to Ni^{2+} chelate beads for 1 hr at 4 °C, beads were washed with wash buffer (50 mM Tris/HCl pH 8.0, 40 mM imidazole, 300 mM NaCl, 1 mM TCEP), and proteins eluted by tag cleavage with 100 nM of bdSENP1 protease in cleavage buffer (50 mM Tris/HCl pH 8.0, 20 mM imidazole, 300 mM NaCl, 0.5 mM TCEP) for 3 hours at 4 °C.

The tag-free proteins were concentrated to approximately 20-30 mg/ml. IP6 (D-myo-inositol hexakisphosphate)-assisted assembly[54] into CLPs essentially followed a published protocol[64] with minor modifications. The buffer of the CA protein was exchanged to 25 mM MES pH 6.0, 50 mM NaCl and 1 mM TECP. Assembly was initiated by adding 0.5 volumes of 75 mM MES pH 6.0, 150 mM NaCl, 15 mM IP6, 3 mM TCEP and shifting the temperature to 37°C for two hours. The final volume was 500µl and the CA concentration was 12 mg/ml.

Assembled CLPs were pelleted by centrifugation in an 5424R Eppendorf microcentrifuge (FA-45-24-11 rotor) at 21 000 g for 10 min, resuspended in gel filtration buffer (25 mM Tris/HCl pH 8.0, 150 mM NaCl, 0.5 mM IP6 and 0.5 mM TCEP), aggregates were removed by centrifugation @ 21 000 g for 5 minutes. Note that the pH-shift prevents the pelleting of properly assembled CLPs in the second centrifugation step. The CLPs in the supernatant were further purified by size exclusion chromatography on a Superose6 Increase 10/300 GL column (equilibrated in gel filtration buffer), where they eluted in the void volume. Assembly was performed either in the presence of 2 mM soluble sinGFP4a (ref.[14]), as a non-covalently

encapsulated tracer, or with CA-EGFP or CA-mScarlet-I3 fusion (ref.[104]) used as a tracer in a 1:6 molar ratio to the unlabeled CA.

CLPs with the Q9E-A92E, H87Q-A92E-G94D, N57A-A92E-G94D, N57A-H87Q-A92E-G94D and V86E-A88E-I91E-A92E mutations do not pellet in the post-assembly centrifugation step, probably because of charge-repulsion. They were therefore directly applied to the Superose 6 column, after removing aggregates by a 5 minutes 21 000 g centrifugation step.

Assembly and purification of 40 nm capsid spheres

His_{14}-bdSUMO tagged CA^{P1A} with the additional N21C, A22C mutations[67] was expressed and purified as described above and concentrated to approximately 6 mg/ml. Assembly was performed by dialysis against 50 mM Tris/HCl pH 8.0, 1 M NaCl and 0.1 mM IP6 for 24 to 48 hours. Assembled CLPs were further purified by size exclusion chromatography on a Superose6 Increase 10/300 GL column equilibrated in 25 mM Tris/HCl pH 8.0, 500 mM NaCl, 0.5 mM IP6, where they eluted near the void volume.

Assembly and purification of 20 nm pentamer-only spheres

For assembly of 20 nm pentamer-only spheres, His_{14}-bdSUMO tagged $CA^{P1A,G60A,G61P}$ was purified as described above and concentrated to approximately 20-30 mg/ml. Assembly into 20 nm capsids followed a published protocol[64] with minor modifications. Briefly, the buffer of the CA protein was exchanged to 25 mM MES pH 6.0, 50 mM NaCl and 1 mM TECP, assembly was initiated by adding 0.5 volumes of 75 mM MES pH 6.0, 150 mM NaCl, 15 mM IP6, 3 mM TCEP, and allowed to proceed in a volume of 500 µl for two hours at 37°C and a CA concentration of 12 mg/ml. Assembled CLPs were further purified by size exclusion chromatography on a Superose6 Increase 10/300 GL column equilibrated in 25 mM Tris/HCl pH 8.0, 150 mM NaCl, 0.5 mM IP6 and 0.5 mM TCEP.

Fluorescence labeling

The anti-Nup133 nanobody xhNup133-Nb2t was labelled with Alexa Fluor™ 647 C_2 maleimide (ThermoFisher) through two ectopic cysteines at the N- and C-termini as desribed[70], reaching a density of labeling (DOL) of 2. The MBP tandem dimer (2xMBP) was also labeled with Alexa Fluor™ 647 C_2 maleimide but through a single cysteine at the C terminus.

Negative stain electron microscopy

Samples were bound to a glow-discharged carbon foil covered 400 mesh copper grid. After successive washes with water, samples were stained with 1% uranyl acetate aq. and examined at room temperature on a Talos L120C transmission electron microscope (Thermo Fisher Scientific).

Digitonin-permeabilized cell assays

HeLa K cells (RRID: CVCL_1922) and XTC-2 cells (RRID: CVCL_5610; ref.[72]) were obtained from the European Cell Culture Collection, authenticated by the manufacturer, and tested negative for mycoplasma. HeLa cells were grown at 37°C in Dulbecco's modified Eagle's Medium (DMEM, high glucose), supplemented with 10% heat-inactivated fetal calf serum (FCS), antibiotics ("AAS", Sigma-Aldrich) and 5% CO_2. XTC-2 cells were cultivated at 22°C in 70% Leibovitz medium (diluted with water), 10% FCS, AAS antibiotics.

Cells were seeded on 8-well µ-Slides (IBIDI, Germany) to 70% confluence. Plasma membranes were permeabilized[26,105] by treating the cells with 30 µg/ml digitonin (water-soluble fraction) in transport buffer (20 mM HEPES/KOH pH 7.5, 110 mM (HeLa) or 80 mM potassium acetate (XTC-2), 5 mM magnesium acetate, 0.5 mM EGTA, 250 mM sucrose) for 3 or 6 mins at 25°C (with gentle shaking), followed by three washing steps in transport buffer. Permeabilized cells were then incubated for 30 minutes with 40 nM Alexa647-labeled xhNup133-Nb2t nanobody and either EGFP (3 µM), CA-EGFP (1 µM), conical CLPs, 40 nm capsid spheres of HIV-1 or 20 nm pentamer-only spheres (with CA concentrations of 5, 1, and 2 µM, respectively). The samples were then directly scanned with a Leica SP8 confocal laser-scanning microscope (equipped with a 63x oil objective and HyD GaAsP detectors), with sequential excitation at 488 and 638 nm.

mRNA for microinjection

Mouse CPSF6 (CCDS Database: CCDS78898.1) and mScarlet[106] cDNAs were cloned into pGEMHE[107], which contains a T7 promoter, *Xenopus* globin 5' and 3'UTRs, and poly(A) tail. Plasmids were linearized by AscI (NEB) before *in vitro* transcription using the HiScribe T7 ARCA mRNA Kit (NEB #E2060S). mRNAs were purified with the RNeasy Mini Kit (Qiagen #74104). 4 pl of 1 µM mRNA was cytoplasmically injected.

Microinjections

Mouse oocytes were obtained from ovaries of 9-week-old CD1 mice that were maintained in a specific pathogen-free environment according to the Federation of European Laboratory Animal Science Association guidelines and recommendations, as previously described[108]. Fully-grown oocytes were kept arrested in prophase in homemade M2 medium supplemented with 250 µM dibutyryl cyclic AMP under paraffin oil (NidaCon) at 37°C. Labelled CLPs, along with the Alexa647 labelled 2xMBP injection marker, were microinjected into cytosol or nucleus of oocytes, as previously described[27]. Oocytes were imaged about 30 minutes after microinjection.

FG phase assays

The assays were performed as previously described[69] with minor modifications. In brief, 1mM FG domain stocks were prepared in 2 M (GLFG repeats) or 4 M guanidinium hydrochloride (all other repeat domains). Phase separation was initiated by rapid dilution of the FG domain stock with 25 volumes (GLFG repeats) or 50 volumes (other repeats) of assay buffer (50 mM Tris/HCl pH 7.5, 150 or 250 mM NaCl, 0.5 mM IP6), followed by a further fourfold dilution in buffer with indicated fluorescent probes. The resulting mixture was pipetted on collagen-coated µ-slides 18-well (IBIDI, Germany), and FG particles were allowed to settle on the bottom for one hour, before confocal scans were taken. Salt effects of the assay are detailed in Extended Data Fig.10.

Partition coefficients were calculated as the integrated raw signal within independent FG particles (In) divided by the signals reference areas in outside regions (Out). Background was not subtracted, which means that the numbers of high partition coefficients are still underestimated. Plots are shown for representative FG particles (with 5-10 µm diameters). Images were analyzed in FIJI 2.9.0 and the exported data were further processed in GraphPad Prism 9.5.1.

The sequences of FG domains used in FG phase assays are shown in Supplementary Table S1.

Structure modeling

The Alphafold 3 Server[109] was used for all modellings. FG-binding was modelled with 6 copies of the wildtype HIV-1:M CA and 6 copies of the respective FG peptide. The following sequences were used: CPSF6 (PVLFPGQPFGQPPLG), GLFG (QPATGGLFGGNTQ), SLFG (QPATGSLFGGNTQ), FSFG (NTQPATGFSFGGNTQPATG)

CA (PIVQNLQGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRLHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTHNPPIPVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQNANPDCKTILKALGPGATLEEMMTACQGVGGPGHKARVL)

For the structural alignment and assessment of accessibilities (Extended Data Fig.9), the sequences shown in the figure were used as inputs to model mature hexamers.

Competing interests: The authors have no competing interests to declare.

Acknowledgements

We wish to thank Thomas U. Schwartz and Onno Ackermans for fruitful collaborations, discussions, and support, Sheung Chun Ng and Jürgen Schünemann for supply with FG domains, Kathrin Gregor for the labelled xhNup133-Nb2t nanobody, Waltraud Taxer for producing the CPSF5-CPSF6 complex, Zhenglin Fu for advice in surface calculations, Stefanie Reiter and Svetlana Agafonova for cell culture support, Susanne Brandfass and Gaby Hawlitschek for technical help, and Thomas U. Schwartz, Sheung Chun Ng, and Loren Andreas for critical reading of the manuscript.

References

1. Feldherr, C. M. The nuclear annuli as pathways for nucleocytoplasmic exchanges. *The Journal of Cell Biology* **14**, 65-72 (1962).
2. Knockenhauer, K. E. & Schwartz, T. U. The nuclear pore complex as a flexible and dynamic gate. *Cell* **164**, 1162-1171 (2016).
3. Lin, D. H. & Hoelz, A. The Structure of the Nuclear Pore Complex (An Update). *Annu Rev Biochem* **88**, 725-783 (2019).
4. Görlich, D. & Kutay, U. Transport between the cell nucleus and the cytoplasm. *Annu Rev Cell Dev Biol* **15**, 607-660 (1999).
5. Kimura, M. & Imamoto, N. Biological significance of the importin-β family-dependent nucleocytoplasmic transport pathways. *Traffic* **15**, 727-748 (2014).
6. Baade, I. & Kehlenbach, R. H. The cargo spectrum of nuclear transport receptors. *Curr Opin Cell Biol* **58**, 1-7 (2019).

7. Frey, S. & Görlich, D. A saturated FG-repeat hydrogel can reproduce the permeability properties of nuclear pore complexes. *Cell* **130**, 512-523 (2007).

8. Schmidt, H. B. & Görlich, D. Nup98 FG domains from diverse species spontaneously phase-separate into particles with nuclear pore-like permselectivity. *eLife* **4**, e04251 (2015).

9. Musacchio, A. On the role of phase separation in the biogenesis of membraneless compartments. *EMBO J* **41**, e109952 (2022).

10. Gao, Y., Skowyra, M. L., Feng, P. & Rapoport, T. A. Protein import into peroxisomes occurs through a nuclear pore-like phase. *Science* **378**, eadf3971 (2022).

11. Powers, M. A., Forbes, D. J., Dahlberg, J. E. & Lund, E. The vertebrate GLFG nucleoporin, Nup98, is an essential component of multiple RNA export pathways. *J Cell Biol* **136**, 241-250 (1997).

12. Wente, S. R., Rout, M. P. & Blobel, G. A new family of yeast nuclear pore complex proteins. *J Cell Biol* **119**, 705-723 (1992).

13. Hülsmann, B. B., Labokha, A. A. & Görlich, D. The permeability of reconstituted nuclear pores provides direct evidence for the selective phase model. *Cell* **150**, 738-751 (2012).

14. Frey, S. et al. Surface properties determining passage rates of proteins through nuclear pores. *Cell* **174**, 202-217.e9 (2018).

15. Mattaj, I. W. & Englmeier, L. Nucleocytoplasmic transport: the soluble phase. *Annu Rev Biochem* **67**, 265-306 (1998).

16. Cook, A., Bono, F., Jinek, M. & Conti, E. Structural biology of nucleocytoplasmic transport. *Annu Rev Biochem* **76**, 647-671 (2007).

17. Izaurralde, E., Kutay, U., von Kobbe, C., Mattaj, I. W. & Görlich, D. The asymmetric distribution of the constituents of the Ran system is essential for transport into and out of the nucleus. *EMBO J* **16**, 6535-6547 (1997).

18. Ribbeck, K., Lipowsky, G., Kent, H. M., Stewart, M. & Görlich, D. NTF2 mediates nuclear import of Ran. *EMBO J* **17**, 6587-6598 (1998).

19. Segref, A. et al. Mex67p, a novel factor for nuclear mRNA export, binds to both poly(A)+ RNA and nuclear pores. *EMBO J* **16**, 3256-3271 (1997).

20. Fribourg, S. & Conti, E. Structural similarity in the absence of sequence homology of the messenger RNA export factors Mtr2 and p15. *EMBO Rep* **4**, 699-703 (2003).

21. Kose, S., Furuta, M. & Imamoto, N. Hikeshi, a nuclear import carrier for Hsp70s, protects cells from heat shock-induced nuclear damage. *Cell* **149**, 578-589 (2012).

22. Bayliss, R., Littlewood, T. & Stewart, M. Structural Basis for the Interaction between FxFG Nucleoporin Repeats and Importin-β in Nuclear Trafficking. *Cell* **102**, 99-108 (2000).

23. Isgro, T. A. & Schulten, K. Binding dynamics of isolated nucleoporin repeat regions to importin-β. *Structure* **13**, 1869-1879 (2005).

24. Port, S. A. et al. Structural and functional characterization of CRM1-Nup214 interactions reveals multiple FG-binding sites involved in nuclear export. *Cell Rep* **13**, 690-702 (2015).

25. Ng, S. C. et al. Barrier properties of Nup98 FG phases ruled by FG motif identity and inter-FG spacer length. *Nat Commun* **14**, 747 (2023).

26. Ribbeck, K. & Görlich, D. Kinetic analysis of translocation through nuclear pore complexes. *EMBO J* **20**, 1320-1330 (2001).

27. Fu, L. et al. HIV-1 capsids enter the FG phase of nuclear pores like a transport receptor. *Nature* **626**, 843-851 (2024).

28. Dickson, C. F. et al. The HIV capsid mimics karyopherin engagement of FG-nucleoporins. *Nature* **626**, 836-842 (2024).

29. Barré-Sinoussi, F. et al. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* **220**, 868-871 (1983).

30. Gallo, R. C. et al. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science* **224**, 500-503 (1984).

31. Coffin, J. M., Hughes, S. H. & Varmus, H. *Retroviruses* (CSHL Press, Cold Spring Harbor, NY, 1997).

32. Ratner, L. et al. Complete nucleotide sequence of the AIDS virus, HTLV-III. *Nature* **313**, 277-284 (1985).

33. Hu, W. S. & Hughes, S. H. HIV-1 reverse transcription. *Cold Spring Harb Perspect Med* **2**, a006882 (2012).

34. Maertens, G. N., Engelman, A. N. & Cherepanov, P. Structure and function of retroviral integrase. *Nat Rev Microbiol* **20**, 20-34 (2022).

35. Wilen, C. B., Tilton, J. C. & Doms, R. W. HIV: cell binding and entry. *Cold Spring Harb Perspect Med* **2**, a006866 (2012).

36. Aiken, C. & Rousso, I. The HIV-1 capsid and reverse transcription. *Retrovirology* **18**, 29 (2021).

37. Dwivedi, R., Prakash, P., Kumbhar, B. V., Balasubramaniam, M. & Dash, C. HIV-1 capsid and viral DNA integration. *mBio* **15**, e0021222 (2024).

38. Yamashita, M. & Emerman, M. Retroviral infection of non-dividing cells: old and new perspectives. *Virology* **344**, 88-93 (2006).

39. Fassati, A. HIV infection of non-dividing cells: a divisive problem. *Retrovirology* **3**, 74 (2006).

40. Burdick, R. C. et al. HIV-1 uncoats in the nucleus near sites of integration. *Proc Natl Acad Sci U S A* **117**, 5486-5493 (2020).

41. Li, C., Burdick, R. C., Nagashima, K., Hu, W. S. & Pathak, V. K. HIV-1 cores retain their integrity until minutes before uncoating in the nucleus. *Proc Natl Acad Sci U S A* **118**, e2019467118 (2021).

42. Müller, T. G. et al. HIV-1 uncoating by release of viral cDNA from capsid-like structures in the nucleus of infected cells. *Elife* **10**, e64776 (2021).

43. Zila, V. et al. Cone-shaped HIV-1 capsids are transported through intact nuclear pores. *Cell* **184**, 1032-1046.e18 (2021).

44. Schuller, A. P. et al. The cellular environment shapes the nuclear pore complex architecture. *Nature* **598**, 667-671 (2021).

45. Kosinski, J. et al. Molecular architecture of the inner ring scaffold of the human nuclear pore complex. *Science* **352**, 363-365 (2016).

46. Lee, K. et al. Flexible use of nuclear import pathways by HIV-1. *Cell Host Microbe* **7**, 221-233 (2010).

47. Lee, K. et al. HIV-1 capsid-targeting domain of cleavage and polyadenylation specificity factor 6. *J Virol* **86**, 3851-3860 (2012).

48. Price, A. J. et al. CPSF6 defines a conserved capsid interface that modulates HIV-1 replication. *PLoS Pathog* **8**, e1002896 (2012).

49. Sundquist, W. I. & Kräusslich, H. G. HIV-1 assembly, budding, and maturation. *Cold Spring Harb Perspect Med* **2**, a006924 (2012).

50. Freed, E. O. HIV-1 gag proteins: diverse functions in the virus life cycle. *Virology* **251**, 1-15 (1998).

51. Klingler, J. et al. How HIV-1 Gag Manipulates Its Host Cell Proteins: A Focus on Interactors of the Nucleocapsid Domain. *Viruses* **12**, 888 (2020).

52. Olety, B. & Ono, A. Roles played by acidic lipids in HIV-1 Gag membrane binding. *Virus Res* **193**, 108-115 (2014).

53. 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).

54. Campbell, S. et al. Modulation of HIV-like particle assembly in vitro by inositol phosphates. *Proc Natl Acad Sci U S A* **98**, 10875-10879 (2001).

55. Rein, A. RNA Packaging in HIV. *Trends Microbiol* **27**, 715-723 (2019).

56. Chen, X. & Wang, X. The HIV-1 gag p6: a promising target for therapeutic intervention. *Retrovirology* **21**, 1 (2024).

57. Kohl, N. E. et al. Active human immunodeficiency virus protease is required for viral infectivity. *Proc Natl Acad Sci U S A* **85**, 4686-4690 (1988).

58. Copeland, T. D. & Oroszlan, S. Genetic locus, primary structure, and chemical synthesis of human immunodeficiency virus protease. *Gene Anal Tech* **5**, 109-115 (1988).

59. Bharat, T. A. M. et al. Cryo-electron microscopy of tubular arrays of HIV-1 Gag resolves structures essential for immature virus assembly. *Proc Natl Acad Sci U S A* **111**, 8233-8238 (2014).

60. Matreyek, K. A., Yücel, S. S., Li, X. & Engelman, A. Nucleoporin NUP153 phenylalanine-glycine motifs engage a common binding pocket within the HIV-1 capsid protein to mediate lentiviral infectivity. *PLoS Pathog* **9**, e1003693 (2013).

61. Bhattacharya, A. et al. Structural basis of HIV-1 capsid recognition by PF74 and CPSF6. *Proc Natl Acad Sci U S A* **111**, 18625-18630 (2014).

62. Buffone, C. et al. Nup153 unlocks the nuclear pore complex for HIV-1 nuclear translocation in nondividing cells. *J Virol* **92**, e00648-18 (2018).

63. Fischer, D. K. et al. CA mutation N57A has distinct strain-specific HIV-1 capsid uncoating and infectivity phenotypes. *J Virol* **93**, e00214-19 (2019).

64. 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).

65. Stacey, J. C. V. et al. Two structural switches in HIV-1 capsid regulate capsid curvature and host factor binding. *Proc Natl Acad Sci U S A* **120**, e2220557120 (2023).

66. Dick, R. A. et al. Inositol phosphates are assembly co-factors for HIV-1. *Nature* **560**, 509-512 (2018).

67. Pornillos, O., Ganser-Pornillos, B. K., Banumathi, S., Hua, Y. & Yeager, M. Disulfide bond stabilization of the hexameric capsomer of human immunodeficiency virus. *J Mol Biol* **401**, 985-995 (2010).

68. Zhang, Z. et al. T = 4 icosahedral HIV-1 capsid as an immunogenic vector for HIV-1 V3 loop epitope display. *Viruses* **10**, 667 (2018).

69. Ng, S. C., Güttler, T. & Görlich, D. Recapitulation of selective nuclear import and export with a perfectly repeated 12mer GLFG peptide. *Nat Commun* **12**, 4047 (2021).

70. Solà Colom, M. et al. A checkpoint function for Nup98 in nuclear pore formation suggested by novel inhibitory nanobodies. *EMBO J* **43**, 2198-2232 (2024).

71. Schaller, T. et al. HIV-1 Capsid-Cyclophilin Interactions Determine Nuclear Import Pathway, Integration Targeting and Replication Efficiency. *PLoS Pathogens* **7**, e1002439 (2011).

72. Pudney, M., Varma, M. G. & Leake, C. J. Establishment of a cell line (XTC-2) from the South African clawed toad, Xenopus laevis. *Experientia* **29**, 466-467 (1973).

73. Ribbeck, K. & Görlich, D. The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion. *EMBO J* **21**, 2664-2671 (2002).

74. Forshey, B. M., von Schwedler, U., Sundquist, W. I. & Aiken, C. Formation of a human immunodeficiency virus type 1 core of optimal stability is crucial for viral replication. *J Virol* **76**, 5667-5677 (2002).

75. von Schwedler, U. K., Stray, K. M., Garrus, J. E. & Sundquist, W. I. Functional surfaces of the human immunodeficiency virus type 1 capsid protein. *J Virol* **77**, 5439-5450 (2003).

76. Saito, A. et al. Capsid-CPSF6 Interaction Is Dispensable for HIV-1 Replication in Primary Cells but Is Selected during Virus Passage In Vivo. *J Virol* **90**, 6918-6935 (2016).

77. Franke, E. K., Yuan, H. E. & Luban, J. Specific incorporation of cyclophilin A into HIV-1 virions. *Nature* **372**, 359-362 (1994).

78. Villegas, J. A. & Levy, E. D. A unified statistical potential reveals that amino acid stickiness governs nonspecific recruitment of client proteins into condensates. *Protein Sci* **31**, e4361 (2022).

79. Jang, S. et al. HIV-1 usurps mixed-charge domain-dependent CPSF6 phase separation for higher-order capsid binding, nuclear entry and viral DNA integration. *Nucleic Acids Res* **52**, 11060-11082 (2024).

80. Ganser, B. K., Li, S., Klishko, V. Y., Finch, J. T. & Sundquist, W. I. Assembly and analysis of conical models for the HIV-1 core. *Science* **283**, 80-83 (1999).

81. Li, S., Hill, C. P., Sundquist, W. I. & Finch, J. T. Image reconstructions of helical assemblies of the HIV-1 CA protein. *Nature* **407**, 409-413 (2000).

82. Briggs, J. A. G., Wilk, T., Welker, R., Kräusslich, H.-G. & Fuller, S. D. Structural organization of authentic, mature HIV-1 virions and cores. *EMBO J* **22**, 1707-1715 (2003).

83. Benjamin, J., Ganser-Pornillos, B. K., Tivol, W. F., Sundquist, W. I. & Jensen, G. J. Three-dimensional structure of HIV-1 virus-like particles by electron cryotomography. *J Mol Biol* **346**, 577-588 (2005).

84. Pornillos, O. et al. X-ray structures of the hexameric building block of the HIV capsid. *Cell* **137**, 1282-1292 (2009).

85. Zhao, G. et al. Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature* **497**, 643-646 (2013).

86. Frey, S., Richter, R. P. & Görlich, D. FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties. *Science* **314**, 815-817 (2006).

87. Najbauer, E. E., Ng, S. C., Griesinger, C., Görlich, D. & Andreas, L. B. Atomic resolution dynamics of cohesive interactions in phase-separated Nup98 FG domains. *Nat Commun* **13**, 1494 (2022).

88. Bayliss, R., Littlewood, T., Strawn, L. A., Wente, S. R. & Stewart, M. GLFG and FxFG nucleoporins bind to overlapping sites on importin-beta. *J Biol Chem* **277**, 50597-50606 (2002).

89. Kehlenbach, R. H., Neumann, P., Ficner, R. & Dickmanns, A. Interaction of nucleoporins with nuclear transport receptors: a structural perspective. *Biol Chem* **404**, 791-805 (2023).

90. Fellouse, F. A., Wiesmann, C. & Sidhu, S. S. Synthetic antibodies from a four-amino-acid code: A dominant role for tyrosine in antigen recognition. *Proceedings of the National Academy of Sciences* **101**, 12467-12472 (2004).

91. Aksu, M. et al. Nanobodies to multiple spike variants and inhalation of nanobody-containing aerosols neutralize SARS-CoV-2 in cell culture and hamsters. *Antiviral Res* **221**, 105778 (2024).

92. Mortuza, G. B. et al. High-resolution structure of a retroviral capsid hexameric amino-terminal domain. *Nature* **431**, 481-485 (2004).

93. Qu, K. et al. Structure and architecture of immature and mature murine leukemia virus capsids. *Proc Natl Acad Sci U S A* **115**, E11751-E11760 (2018).

94. Salas-Briceno, K., Zhao, W. & Ross, S. R. Murine leukemia virus infection of non-dividing dendritic cells is dependent on nucleoporins. *PLoS Pathog* **20**, e1011640 (2024).

95. Kreysing, J. P. et al. Passage of the HIV capsid cracks the nuclear pore. *Cell* S0092-8674(24)01421 (2025).

96. Rankovic, S., Varadarajan, J., Ramalho, R., Aiken, C. & Rousso, I. Reverse Transcription Mechanically Initiates HIV-1 Capsid Disassembly. *J Virol* **91**, e00289-17 (2017).

97. Müller, T. G., Zila, V., Müller, B. & Kräusslich, H.-G. Nuclear Capsid Uncoating and Reverse Transcription of HIV-1. *Annu Rev Virol* **9**, 261-284 (2022).

98. Ay, S. & Di Nunzio, F. HIV-Induced CPSF6 Condensates. *J Mol Biol* **435**, 168094 (2023).

99. Burdick, R. C. et al. Dynamics and regulation of nuclear import and nuclear movements of HIV-1 complexes. *PLoS Pathog* **13**, e1006570 (2017).

100. 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**, e1004459 (2014).

101. Yang, Q., Coseno, M., Gilmartin, G. M. & Doublié, S. Crystal structure of a human cleavage factor CFI(m)25/CFI(m)68/RNA complex provides an insight into poly(A) site recognition and RNA looping. *Structure* **19**, 368-377 (2011).

102. Frey, S. & Görlich, D. A new set of highly efficient, tag-cleaving proteases for purifying recombinant proteins. *J Chromatogr A* **1337**, 95-105 (2014).

103. Frey, S. & Görlich, D. Purification of protein complexes of defined subunit stoichiometry using a set of orthogonal, tag-cleaving proteases. *J Chromatogr A* **1337**, 106-115 (2014).

104. Gadella, T. W. J. et al. mScarlet3: a brilliant and fast-maturing red fluorescent protein. *Nat Methods* **20**, 541-545 (2023).

105. Adam, S. A., Marr, R. S. & Gerace, L. Nuclear protein import in permeabilized mammalian cells requires soluble cytoplasmic factors. *J Cell Biol* **111**, 807-816 (1990).

106. Bindels, D. S. et al. mScarlet: a bright monomeric red fluorescent protein for cellular imaging. *Nat Methods* **14**, 53-56 (2017).

107. Liman, E. R., Tytgat, J. & Hess, P. Subunit stoichiometry of a mammalian K+ channel determined by construction of multimeric cDNAs. *Neuron* **9**, 861-871 (1992).

108. Schuh, M. & Ellenberg, J. Self-organization of MTOCs replaces centrosome function during acentrosomal spindle assembly in live mouse oocytes. *Cell* **130**, 484-498 (2007).

109. Abramson, J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493-500 (2024).

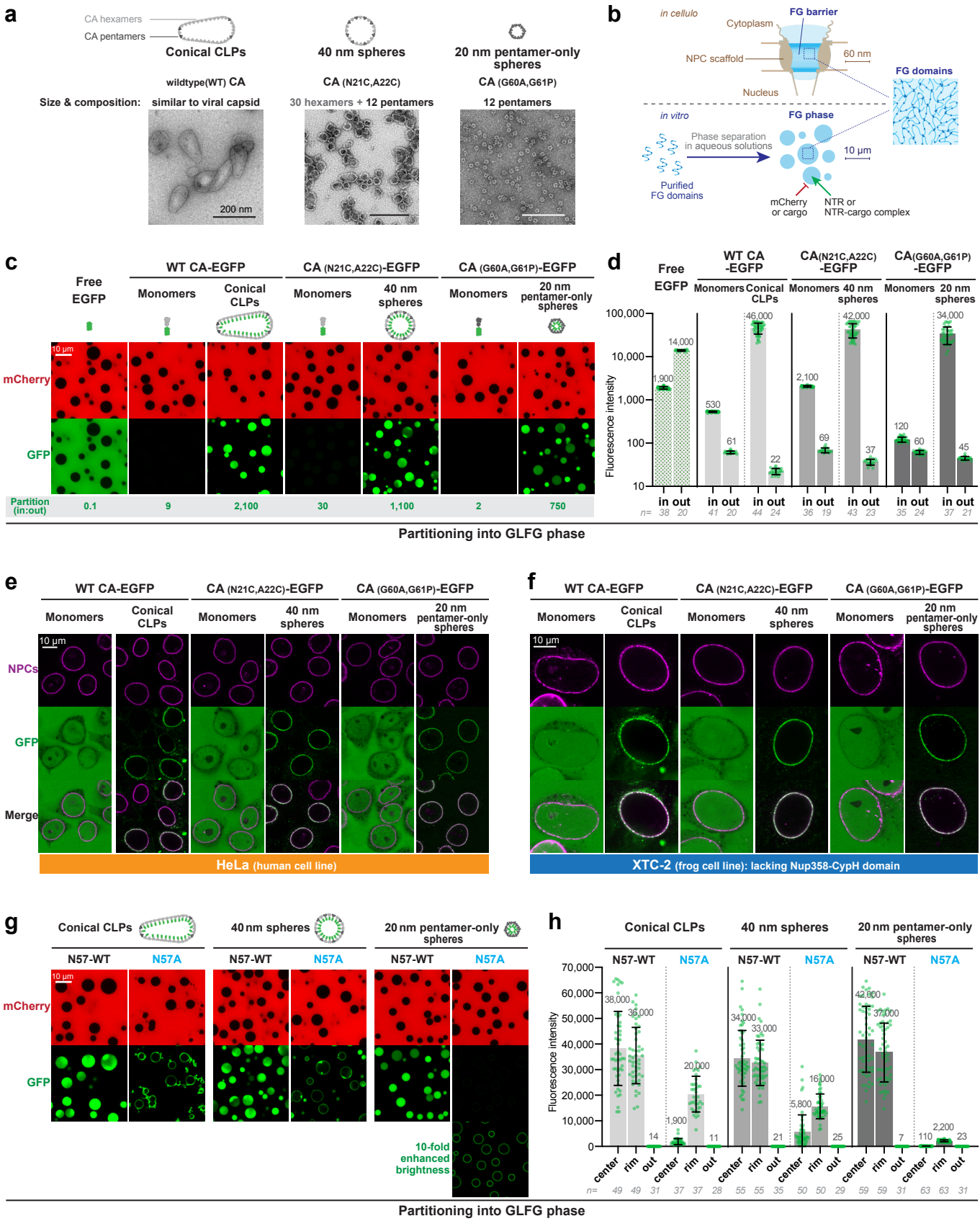

a
CA hexamers
CA pentamers
Conical CLPs
40 nm spheres
20 nm pentamer-only spheres
wildtype(WT) CA
CA (N21C,A22C)
CA (G60A,G61P)
Size & composition:
similar to viral capsid
30 hexamers + 12 pentamers
12 pentamers
200 nm
b
in cellulo
Cytoplasm
FG barrier
NPC scaffold
60 nm
Nucleus
FG domains
in vitro
FG phase
Phase separation in aqueous solutions
10 µm
Purified FG domains
mCherry or cargo
NTR or NTR-cargo complex
c
Free EGFP
WT CA-EGFP
CA (N21C,A22C)-EGFP
CA (G60A,G61P)-EGFP
Monomers
Conical CLPs
40 nm spheres
20 nm pentamer-only spheres
mCherry
GFP
10 µm
Partition (in:out)
0.1
9
2,100
30
1,100
2
750
d
Fluorescence intensity
100,000
10,000
1,000
100
10
1,900
14,000
530
61
46,000
22
2,100
69
42,000
37
120
60
34,000
45
in out
n= 38 20 41 20 44 24 36 19 43 23 35 24 37 21
Partitioning into GLFG phase
e
WT CA-EGFP
CA (N21C,A22C)-EGFP
CA (G60A,G61P)-EGFP
Monomers
Conical CLPs
40 nm spheres
20 nm pentamer-only spheres
NPCs
GFP
Merge
10 µm
HeLa (human cell line)
f
WT CA-EGFP
CA (N21C,A22C)-EGFP
CA (G60A,G61P)-EGFP
Monomers
Conical CLPs
40 nm spheres
20 nm pentamer-only spheres
10 µm
XTC-2 (frog cell line): lacking Nup358-CypH domain
g
Conical CLPs
40 nm spheres
20 nm pentamer-only spheres
N57-WT
N57A
mCherry
GFP
10 µm
10-fold enhanced brightness
h
Conical CLPs
40 nm spheres
20 nm pentamer-only spheres
N57-WT
N57A
Fluorescence intensity
70,000
60,000
50,000
40,000
30,000
20,000
10,000
0
38,000
36,000
14
1,900
20,000
11
34,000
33,000
21
5,800
16,000
25
42,000
37,000
7
110
2,200
23
center rim out
n= 49 49 31 37 37 28 55 55 35 50 50 29 59 59 31 63 63 31
Partitioning into GLFG phase

Figure 1

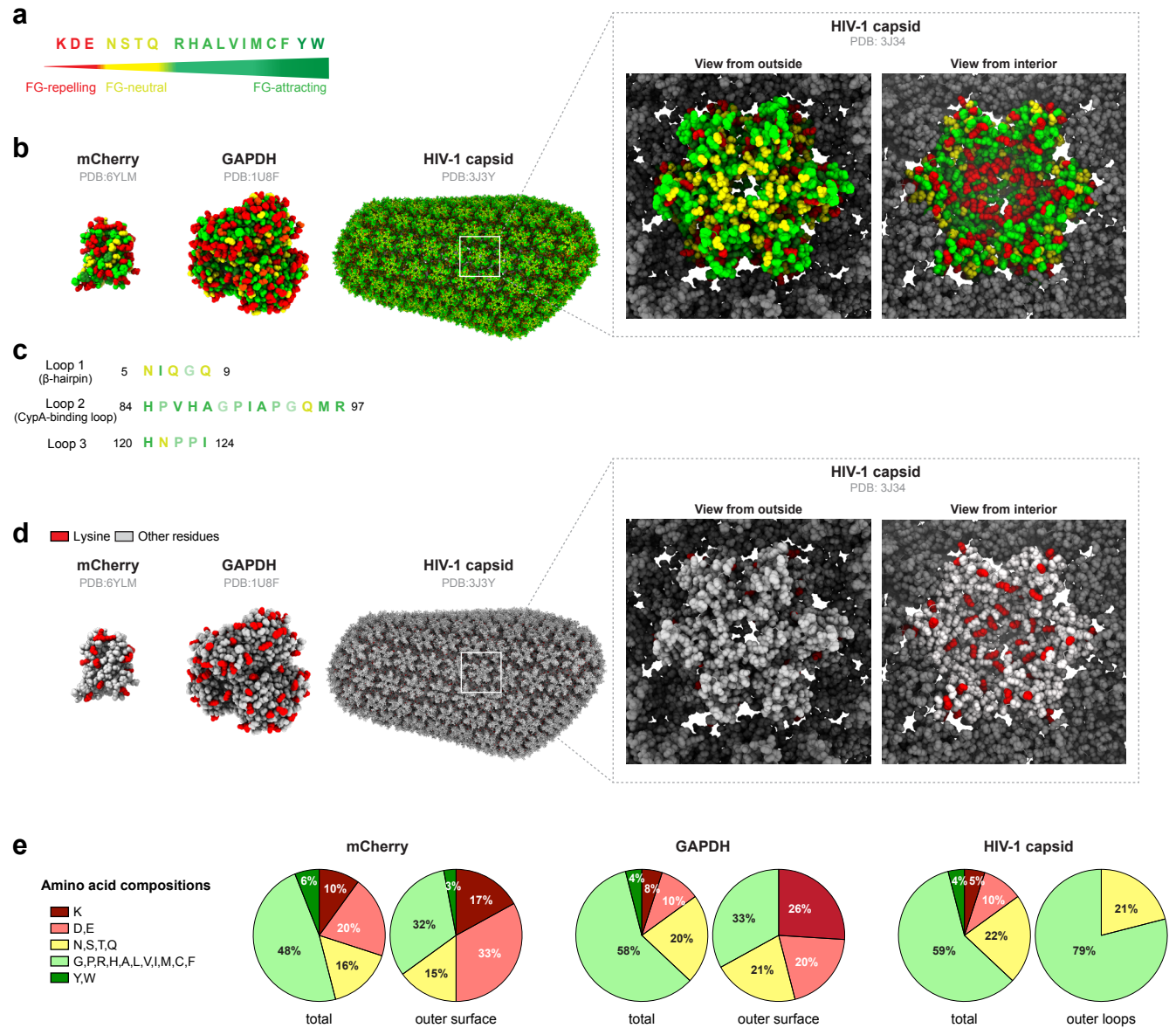

a
K D E N S T Q R H A L V I M C F Y W
FG-repelling
FG-neutral
FG-attracting
b
mCherry
PDB:6YLM
GAPDH
PDB:1U8F
HIV-1 capsid
PDB:3J3Y
HIV-1 capsid
PDB: 3J34
View from outside
View from interior
c
Loop 1 (β-hairpin) 5 N I Q G Q 9
Loop 2 (CypA-binding loop) 84 H P V H A G P I A P G Q M R 97
Loop 3 120 H N P P I 124
d
Lysine
Other residues
mCherry
PDB:6YLM
GAPDH
PDB:1U8F
HIV-1 capsid
PDB:3J3Y
HIV-1 capsid
PDB: 3J34
View from outside
View from interior
e
Amino acid compositions
K
D,E
N,S,T,Q
G,P,R,H,A,L,V,I,M,C,F
Y,W
mCherry
6%
10%
20%
16%
48%
3%
17%
33%
15%
32%
total
outer surface
GAPDH
4%
8%
10%
20%
58%
26%
20%
21%
33%
total
outer surface
HIV-1 capsid
4%
5%
10%
22%
59%
21%
79%
total
outer loops

Figure 2

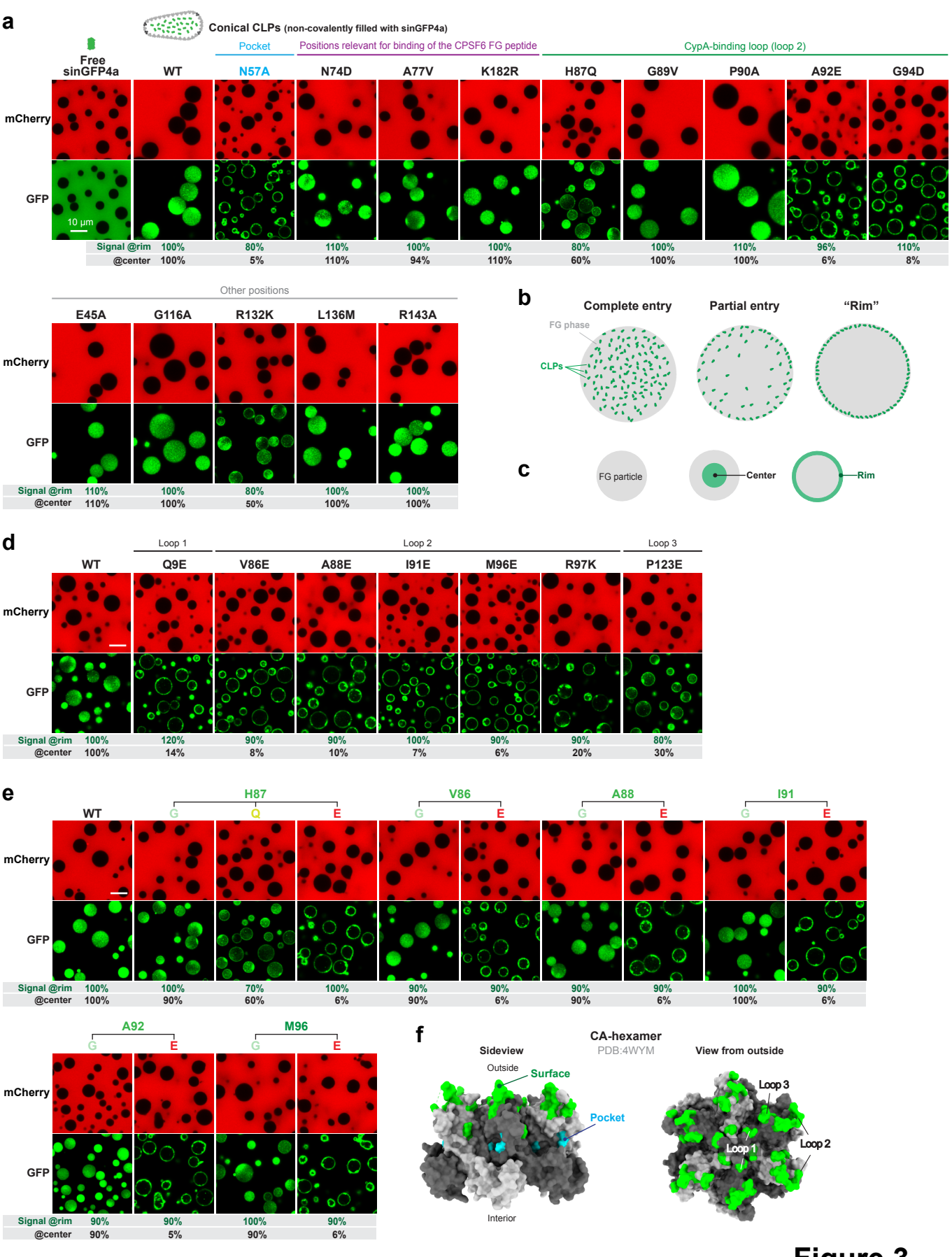
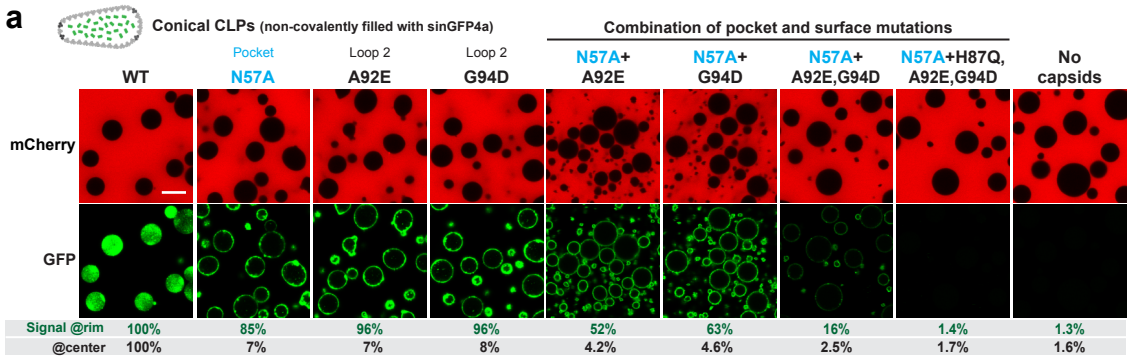
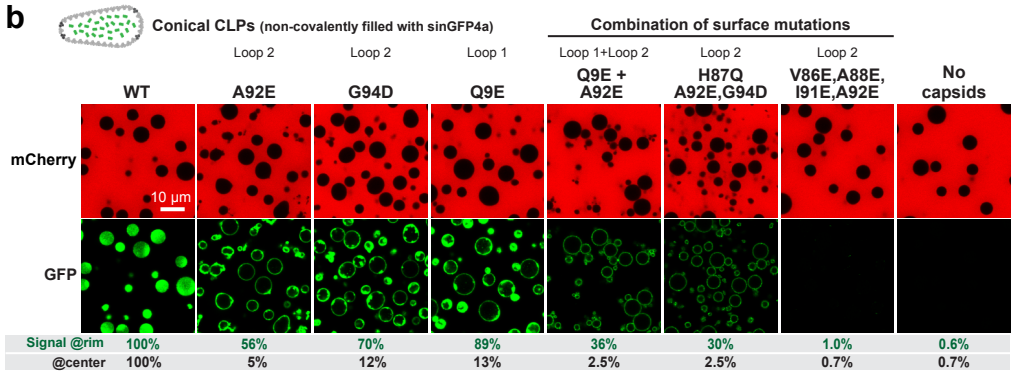

a
Conical CLPs (non-covalently filled with sinGFP4a)
Pocket
Positions relevant for binding of the CPSF6 FG peptide
CypA-binding loop (loop 2)
Free sinGFP4a
WT
N57A
N74D
A77V
K182R
H87Q
G89V
P90A
A92E
G94D
mCherry
GFP
10 µm
Signal @rim 100% 80% 110% 100% 100% 80% 100% 110% 96% 110%
@center 100% 5% 110% 94% 110% 60% 100% 100% 6% 8%
Other positions
E45A
G116A
R132K
L136M
R143A
Signal @rim 110% 100% 80% 100% 100%
@center 110% 100% 50% 100% 100%
b
Complete entry
Partial entry
"Rim"
FG phase
CLPs
c
FG particle
Center
Rim
d
Loop 1
Loop 2
Loop 3
WT
Q9E
V86E
A88E
I91E
M96E
R97K
P123E
Signal @rim 100% 120% 90% 90% 100% 90% 90% 80%
@center 100% 14% 8% 10% 7% 6% 20% 30%
e
H87
V86
A88
I91
G
Q
E
Signal @rim 100% 100% 70% 100% 90% 90% 90% 90% 100% 90%
@center 100% 90% 60% 6% 90% 6% 90% 6% 100% 6%
A92
M96
Signal @rim 90% 90% 100% 90%
@center 90% 5% 90% 6%
f
CA-hexamer
PDB:4WYM
Sideview
View from outside
Outside
Surface
Pocket
Interior
Loop 3
Loop 1
Loop 2

Figure 3


a
Conical CLPs (non-covalently filled with sinGFP4a)
Combination of pocket and surface mutations
Pocket
Loop 2
Loop 2
WT
N57A
A92E
G94D
N57A+ A92E
N57A+ G94D
N57A+ A92E,G94D
N57A+H87Q, A92E,G94D
No capsids
mCherry
GFP
Signal @rim 100% 85% 96% 96% 52% 63% 16% 1.4% 1.3%
@center 100% 7% 7% 8% 4.2% 4.6% 2.5% 1.7% 1.6%


b
Conical CLPs (non-covalently filled with sinGFP4a)
Combination of surface mutations
Loop 2
Loop 2
Loop 1
Loop 1+Loop 2
Loop 2
Loop 2
WT
A92E
G94D
Q9E
Q9E + A92E
H87Q A92E,G94D
V86E,A88E, I91E,A92E
No capsids
mCherry
GFP
10 µm
Signal @rim 100% 56% 70% 89% 36% 30% 1.0% 0.6%
@center 100% 5% 12% 13% 2.5% 2.5% 0.7% 0.7%

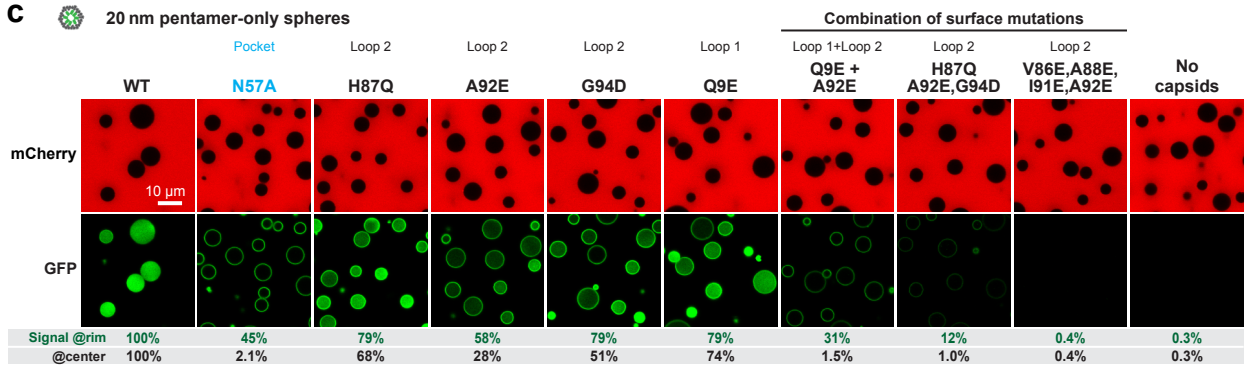

c
20 nm pentamer-only spheres
Combination of surface mutations
Pocket
Loop 2
Loop 2
Loop 2
Loop 1
Loop 1+Loop 2
Loop 2
Loop 2
WT
N57A
H87Q
A92E
G94D
Q9E
Q9E + A92E
H87Q A92E,G94D
V86E,A88E, I91E,A92E
No capsids
mCherry
GFP
10 µm
Signal @rim 100% 45% 79% 58% 79% 79% 31% 12% 0.4% 0.3%
@center 100% 2.1% 68% 28% 51% 74% 1.5% 1.0% 0.4% 0.3%

Figure 4

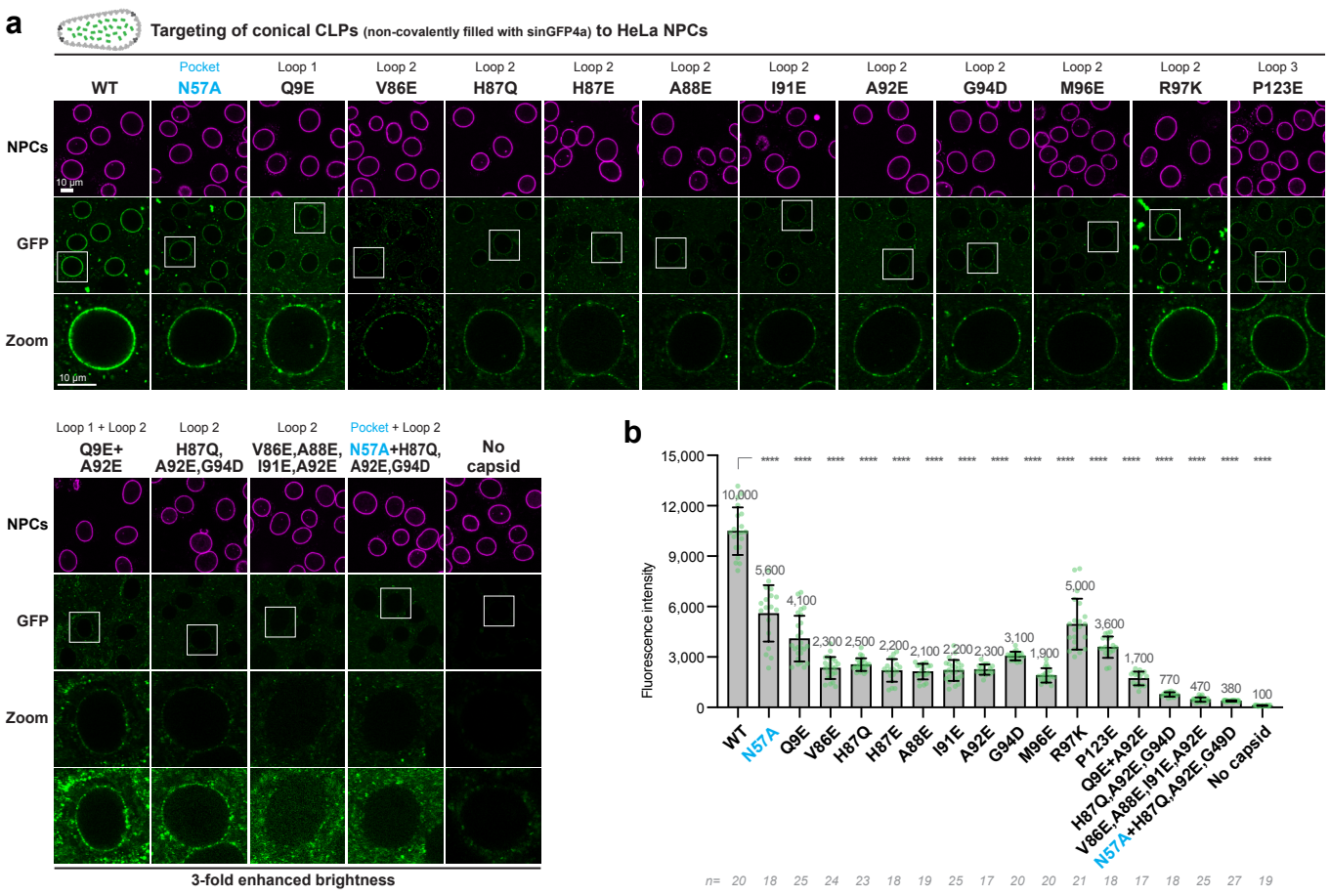

a
Targeting of conical CLPs (non-covalently filled with sinGFP4a) to HeLa NPCs
WT
Pocket N57A
Loop 1 Q9E
Loop 2 V86E
Loop 2 H87Q
Loop 2 H87E
Loop 2 A88E
Loop 2 I91E
Loop 2 A92E
Loop 2 G94D
Loop 2 M96E
Loop 2 R97K
Loop 3 P123E
NPCs
GFP
Zoom
10 µm
Loop 1 + Loop 2 Q9E+ A92E
Loop 2 H87Q, A92E,G94D
Loop 2 V86E,A88E, I91E,A92E
Pocket + Loop 2 N57A+H87Q, A92E,G94D
No capsid
3-fold enhanced brightness
b
Fluorescence intensity
15,000
12,000
9,000
6,000
3,000
0
10,000
5,600
4,100
2,300
2,500
2,200
2,100
2,200
2,300
3,100
1,900
5,000
3,600
1,700
770
470
380
100
****
WT
N57A
Q9E
V86E
H87Q
H87E
A88E
I91E
A92E
G94D
M96E
R97K
P123E
Q9E+A92E
H87Q,A92E,G94D
V86E,A88E, I91E,A92E
N57A+H87Q,A92E,G49D
No capsid
n= 20 18 25 24 23 18 19 25 17 20 20 21 18 17 18 25 27 19

Figure 5

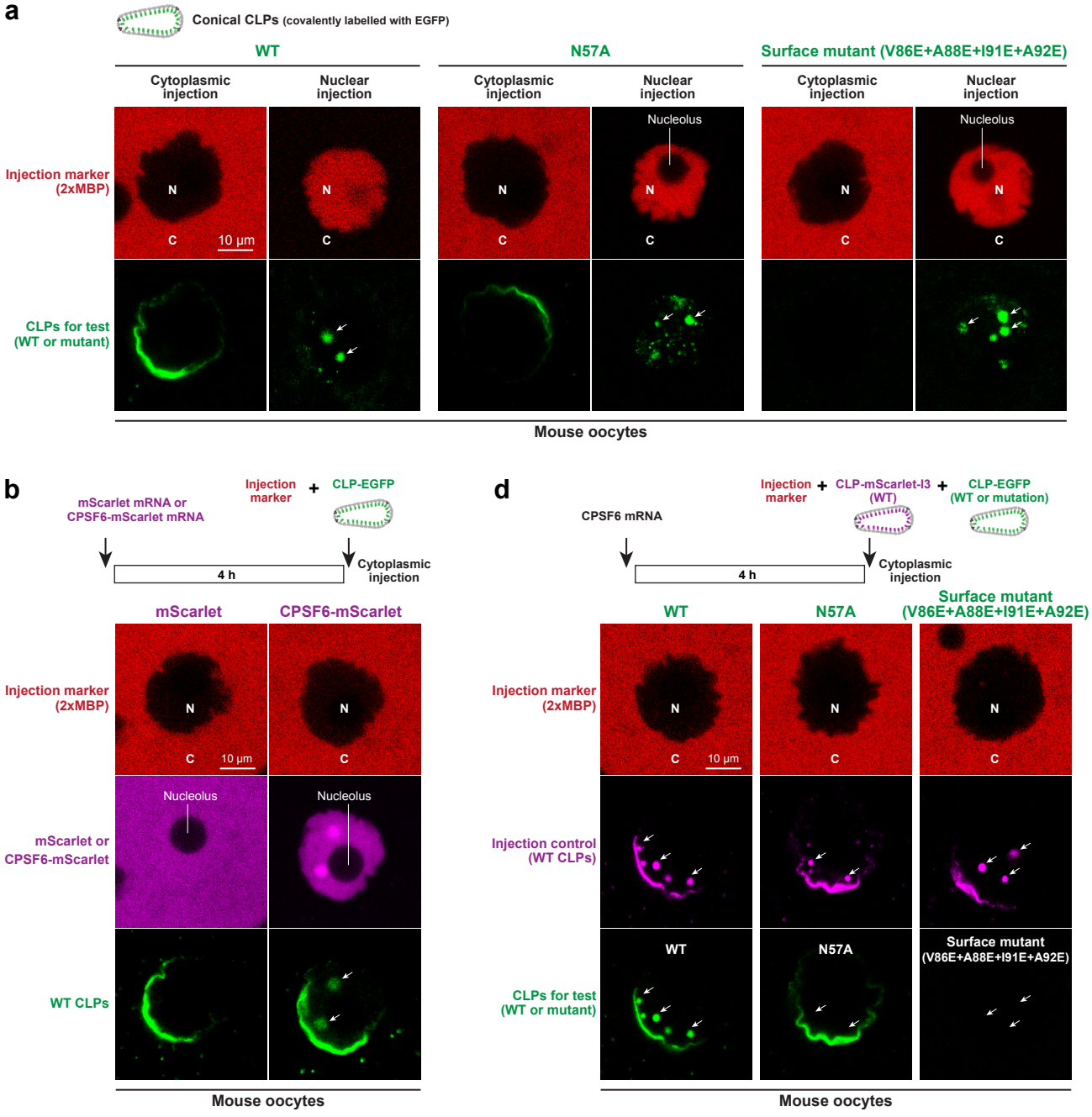

a
Conical CLPs (covalently labelled with EGFP)
WT
N57A
Surface mutant (V86E+A88E+I91E+A92E)
Cytoplasmic injection
Nuclear injection
Injection marker (2xMBP)
CLPs for test (WT or mutant)
N
C
10 μm
Nucleolus
Mouse oocytes
b
mScarlet mRNA or CPSF6-mScarlet mRNA
Injection marker
CLP-EGFP
4 h
Cytoplasmic injection
mScarlet
CPSF6-mScarlet
mScarlet or CPSF6-mScarlet
WT CLPs
Mouse oocytes
d
CPSF6 mRNA
CLP-mScarlet-I3 (WT)
CLP-EGFP (WT or mutation)
Injection control (WT CLPs)
Surface mutant (V86E+A88E+I91E+A92E)
Mouse oocytes
c
+mCherry
+CPSF5/CPSF6-sin2AmCherry
Molar ratio of CPSF5/6 complex : CA
0.25 : 1
0.5 : 1
1.25 : 1
2.5 : 1
mCherry or CPSF5/CPSF6 -sin2AmCherry
Conical CLPs
GLFG phase

Figure 6

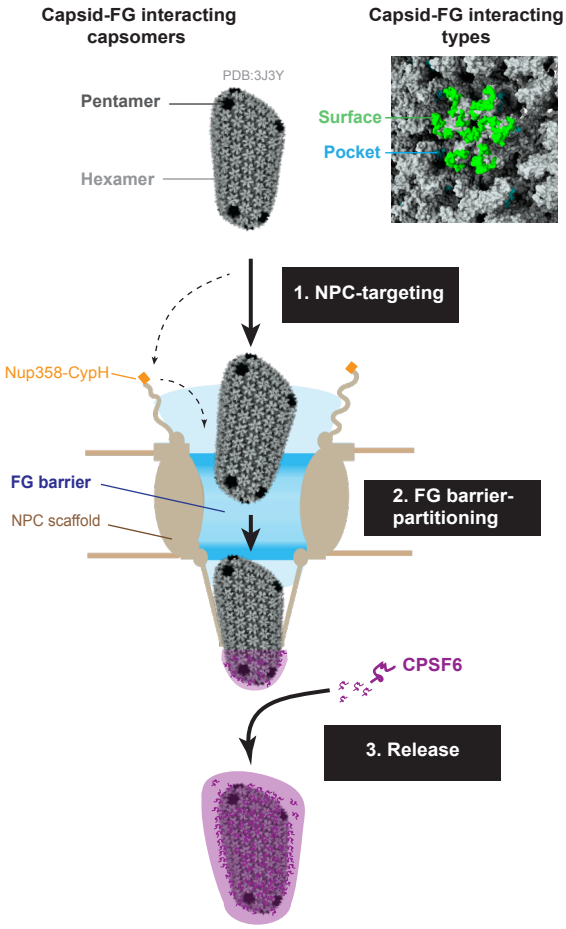
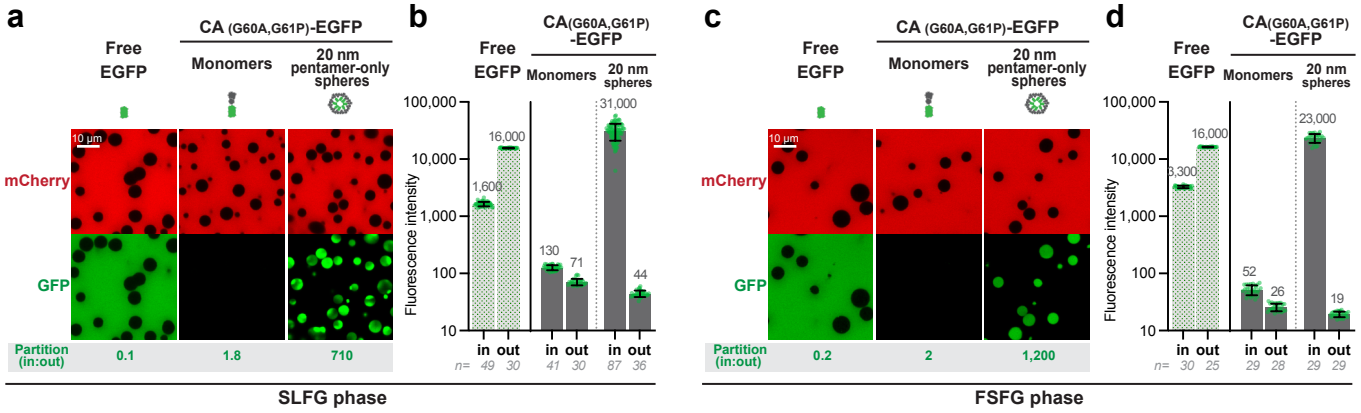

Capsid-FG interacting capsomers
Capsid-FG interacting types
PDB:3J3Y
Pentamer
Hexamer
Surface
Pocket
1. NPC-targeting
Nup358-CypH
FG barrier
NPC scaffold
2. FG barrier-partitioning
CPSF6
3. Release

Figure 7

Extended Data Table 1. Capsid mutations studied and corresponding sequence variations in HIV-1 M subgroup strains

Mutation	Position	FG phase-partitioning [a]	NPC-targeting [b]	Substitutions in naturally occurring strains [c]
N57A	helix 3 (FG pocket)	Severe defect	Reduced	-
V86E	loop 2	Severe defect	Reduced	A, P, I, Q, T, M, L
H87Q	loop 2	Reduced	Reduced	Q, P
H87E	loop 2	Severe defect	Reduced	Q, P
A88E	loop 2	Severe defect	Reduced	V
G89V	loop 2	Increased	Reduced	-
P90A	loop 2	N/A	Reduced	-
I91E	loop 2	Severe defect	Reduced	V, N, A, F, H, L, P, Y, T, Q, M
A92E	loop 2	Severe defect	Reduced	P, Q, V
G94D	loop 2	Severe defect	Reduced	N
M96E	loop 2	Severe defect	Reduced	I, L, V
R97K	loop 2	Reduced	Reduced	-
Q9E	loop 1 (β-hairpin)	Severe defect	Reduced	H, A
P123E	loop 3	Reduced	Reduced	A, S, V, G

a: See Fig.3 and Extended Data Fig.8; **b**: see Fig.5; **c**: Substitutions found in HIV-1 M subgroup strains from patients' isolates, shown as frequency from high to low (as summarized in ref.[102, 103]; see Extended Data Fig.9). Please note the absence of FG-repelling substitutions in naturally occurring strains.

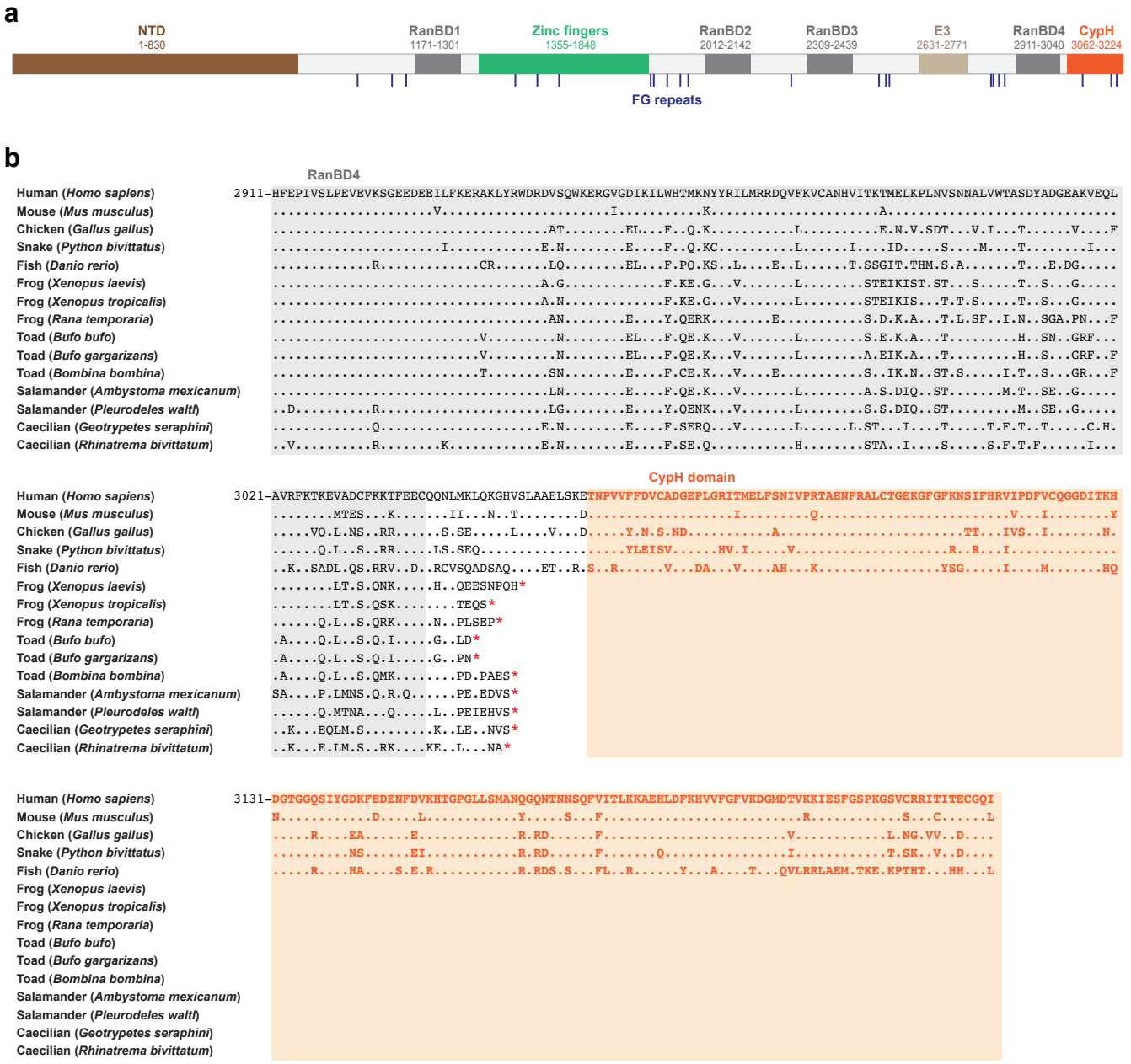


Extended Data Fig.1 | Pentamer-only capsids partition efficiently into SLFG and FSFG phases.

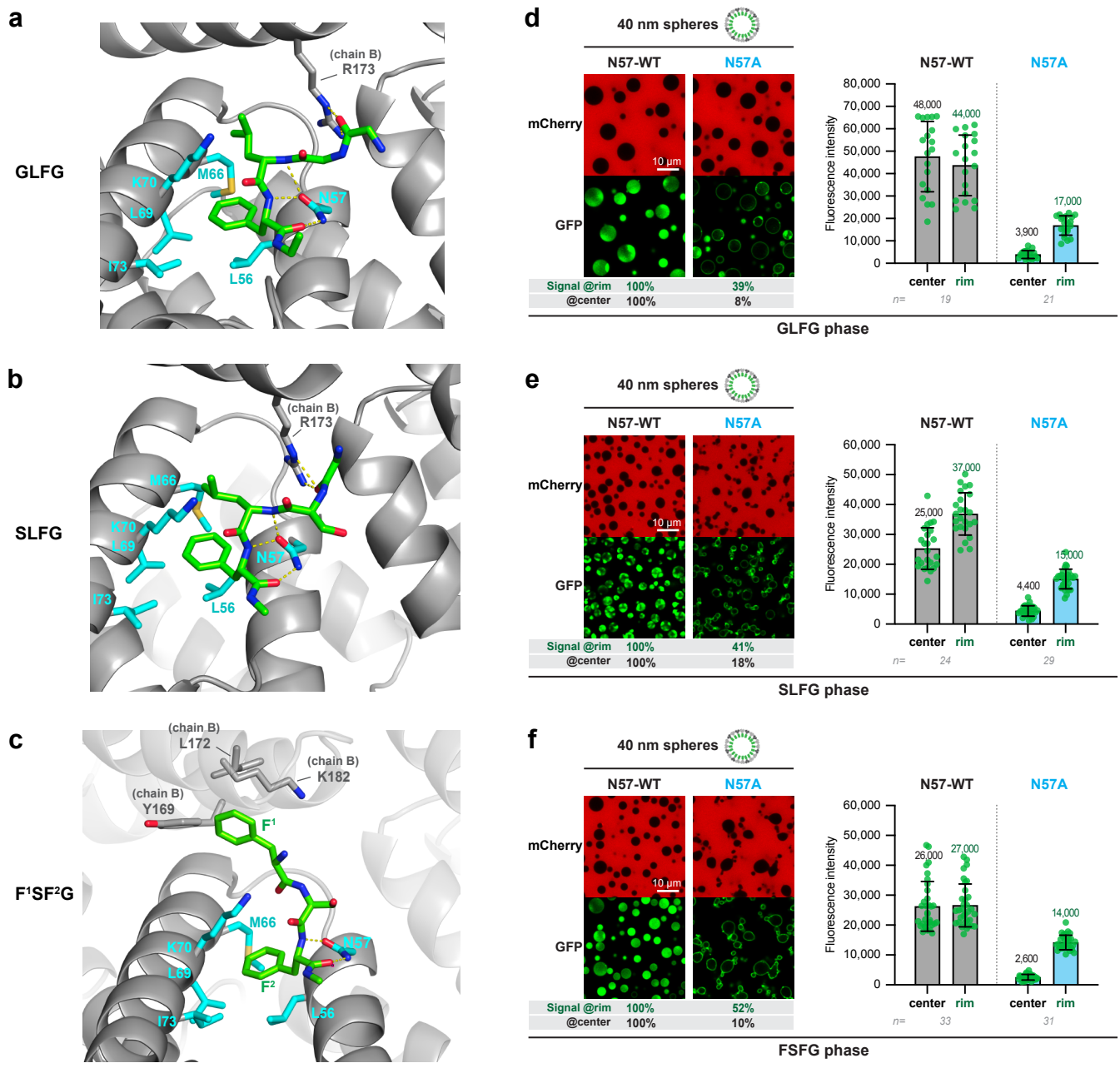
FG phase-partitioning experiments and quantifications were performed as in Fig.1c-d, but FG phases were assembled from different FG repeat domains. Note that fully assembled T1 spheres partitioned very well also in these alternative FG phases – 500 times better than the unassembled CA monomers. Scan settings were identical for each monomer-capsid pair. Experiments were repeated independently with identical outcomes (*N* = 3). Scale bars, 10 µm.

a,b, Instead of the 52x GLFG 12mer repeats (G**GLFG**GNTQPAT)$_{52}$ as in Fig.1, a domain with 52 SLFG 12mer repeat units was used (ref.[25]; see Supplementary Table S1 for sequence). The inter-FG spacer was also free of charged amino acids – a typical feature of Nup98-type FG repeats. SLFG is the second most abundant FG motif in Nup98 FG repeat domains.

c,d, An FG repeat domain of 669 residues with non-charged spacers and 36 FSFG motifs was used. The lower motif density (1 motif per 18 residues) considers that FSFG motifs are more hydrophobic and thus more cohesive than GLFG or SLFG motifs. For sequence see ref.[25] or Supplementary Table S1.



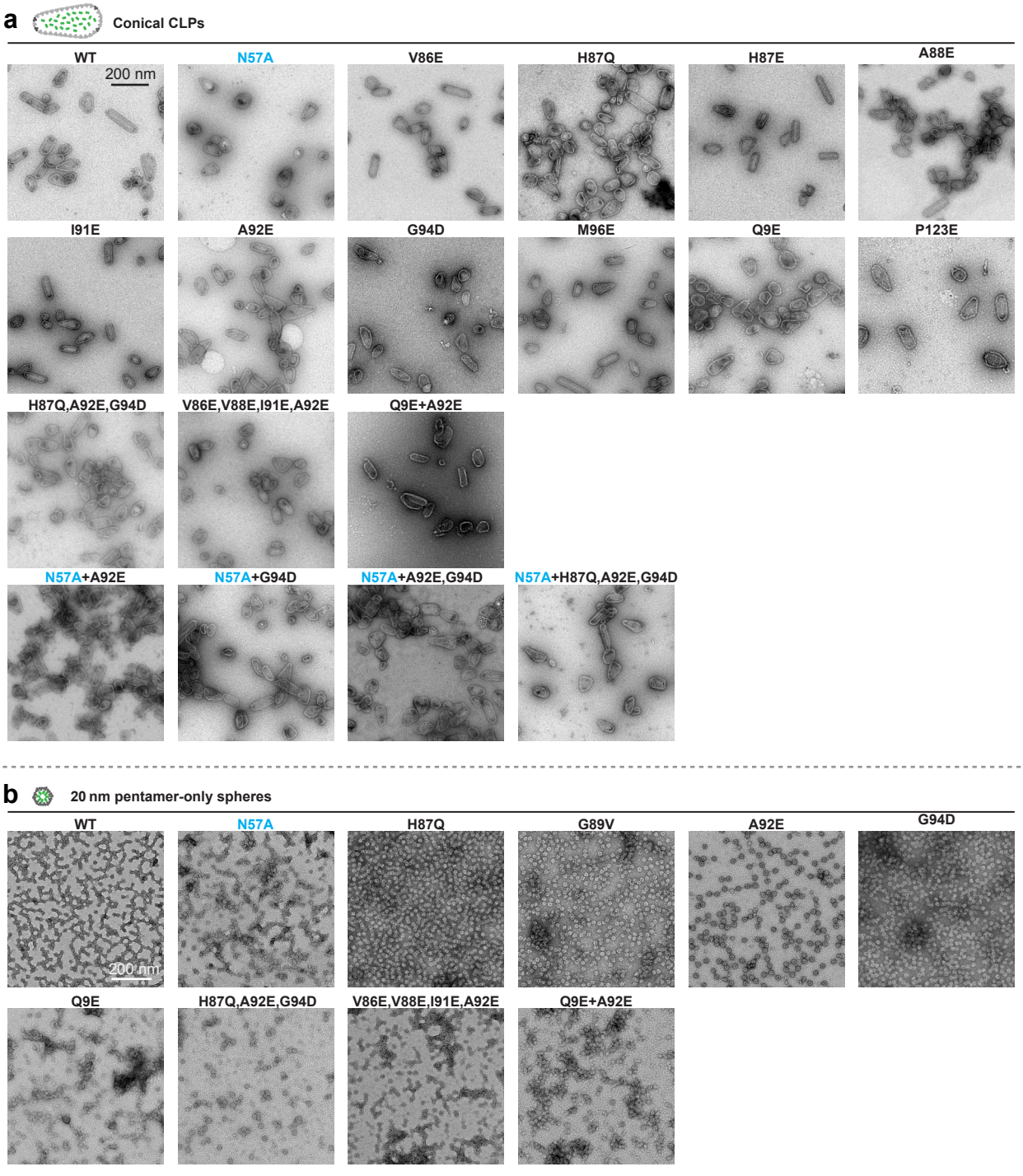
Extended Data Fig.2 | *Xenopus* Nup358/ RanBP2 lacks the C-terminal CypH (cyclophilin homology) domain.
a, Domain organization of human Nup358, which includes an N-terminal NPC-anchor domain, four RanBP1-like Ran-binding domains (RanBD1-4), 8 Ran-binding zinc finger domains, two SUMO-E3 ligase domains, several FG repeat domains, and a C-terminal CypH domain.
b, Sequence alignment of the C terminus part of Nup358 from indicated species. Note that Nup358 from mammals, birds, reptiles, or fish all have a C-terminal CypH domain. However, the CypH domain has been lost in modern amphibians such as frogs, toads, salamanders, and caecilians, where a translation stop is found immediately after the last RanBD. Cells from the frog/ toad *Xenopus laevis* can therefore be considered as a naturally occurring ΔCypH mutant. Sequence identifiers are listed in Supplementary Table S6.



Extended Data Fig.3 | Alphafold 3-modeling confirms that the N57 pocket binds diverse FG repeats in a similar hydrogen bond-stabilized manner.

a,b,c, A mature full-length CA-hexamer was modeled with either GLFG, SLFG, or FSFG peptides bound. Images show just the two CA chains (chain A and chain B) that accommodate the corresponding peptide. The N57 sidechain amide forms two hydrogen-bonds to the backbone of the FG phenylalanine – virtually identically to the FG motif of the CPSF6 peptide (PDB 4B4N, ref.[48]). In the case of the GLFG and SLFG peptide, the N57 sidechain oxygen forms a third H-bond to the backbone nitrogen of the preceding leucine. Additional H-bonds are formed between the guanidine group of R173 (chain B) and the backbone carbonyl oxygen of the glycine or serine that precedes the LFG motif. The sidechain of each FG phenylalanine engages in extensive hydrophobic contact with L56, M66, L69, (the aliphatic part of) K70, and I73 of chain A. The first phenylalanine (F^1) of the FSFG peptide makes additional hydrophobic contacts with Y169, L172, and (the aliphatic part of) K182 of chain B.

d,e,f, Wildtype and N57A mutant 40 nm spheres were allowed to partition into GLFG, SLFG and FSFG phases, essentially as described in Fig.1 and ref.[27]. The N57A mutations caused clear partitioning defects, validating the relevance of the N57 hydrogen bonds for the FG interaction.

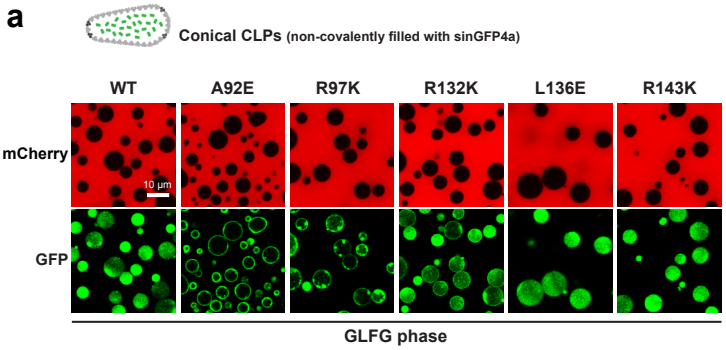


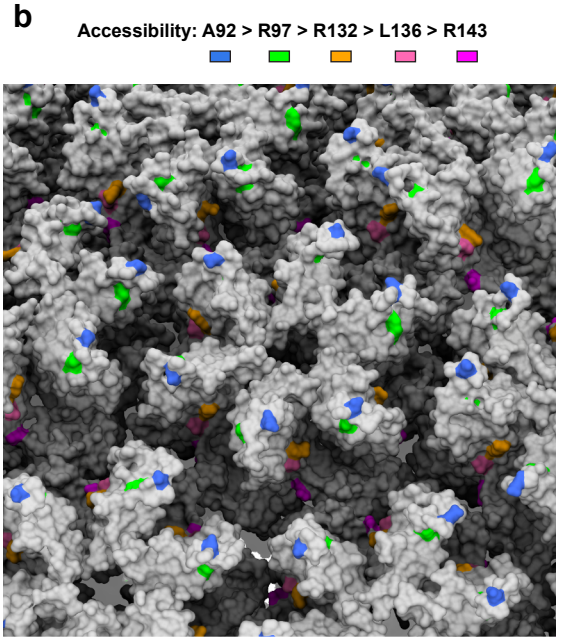
Extended Data Fig.4 | Negative stain electron micrographs of mutant HIV-1 capsids.

a, Wildtype or mutant CA was recombinantly expressed, purified, and allowed to assemble into cone-shaped CLPs (in the presence of IP6; see Methods). The assemblies were applied to a Superose 6 size exclusion column, where CLPs elute close to the void volume. This can be considered as a first criterion for a faithful CLP assembly. The corresponding fractions were pooled and further analyzed by negative stain electron microscopy. Representative images are shown in this panel and confirm the proper assembly of all mutants for which we describe a transport phenotype. Several other mutants showed defects in capsid assembly (for example, I6E, Q7E, M10E/K, P122E, I124E). These were excluded from the analysis.

b, Wildtype and mutant pentamer-only spheres were assembled as analyzed in **a** the differences being (1) that the CA protein contained in addition the G60A,G61P double mutation, which is required to assemble pentamer-only spheres[64], (2) that the assemblies were smaller than the CLPs (20 nm instead of 100-200 nm), and (3) that these smaller assemblies also eluted later from the Superose 6 column. Scale bars, 200 nm.

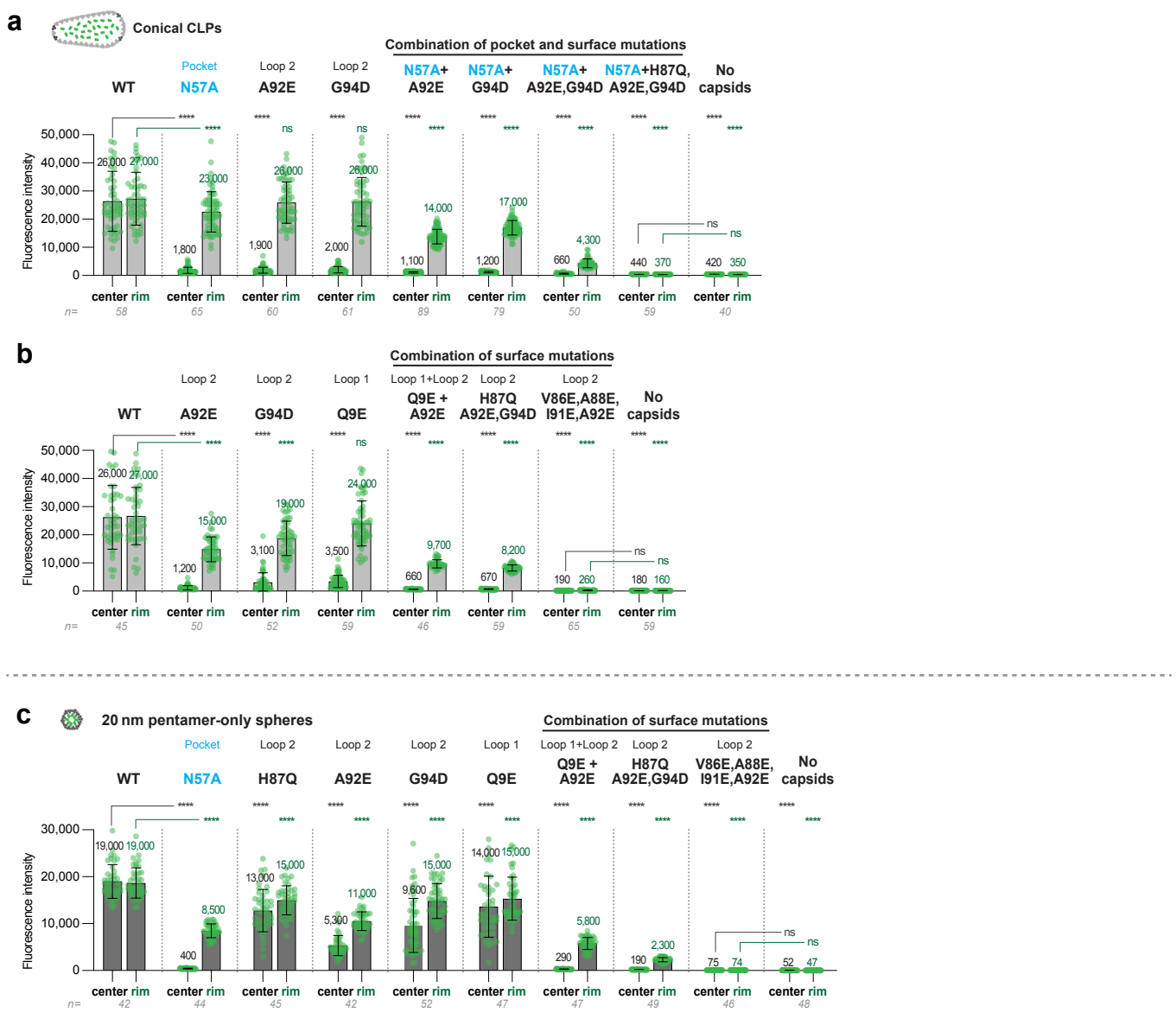
Extended Data Fig.5 | Identification of capsid mutations that caused defects in FG phase-partitioning.
The panels provide more detailed quantifications for the partitioning of capsid mutants into the GLFG phase of Fig.3. They are based on larger datasets. Capsid signals at the FG phase surface and in the FG phase center were integrated separately. *P-values* were calculated using one-way ANOVA with Tukey's post hoc test. *P-values*: * < 0.05, ** < 0.01, *** < 0.001, ****<0.0001, ns: not significant (*P-value* ≥0.05).





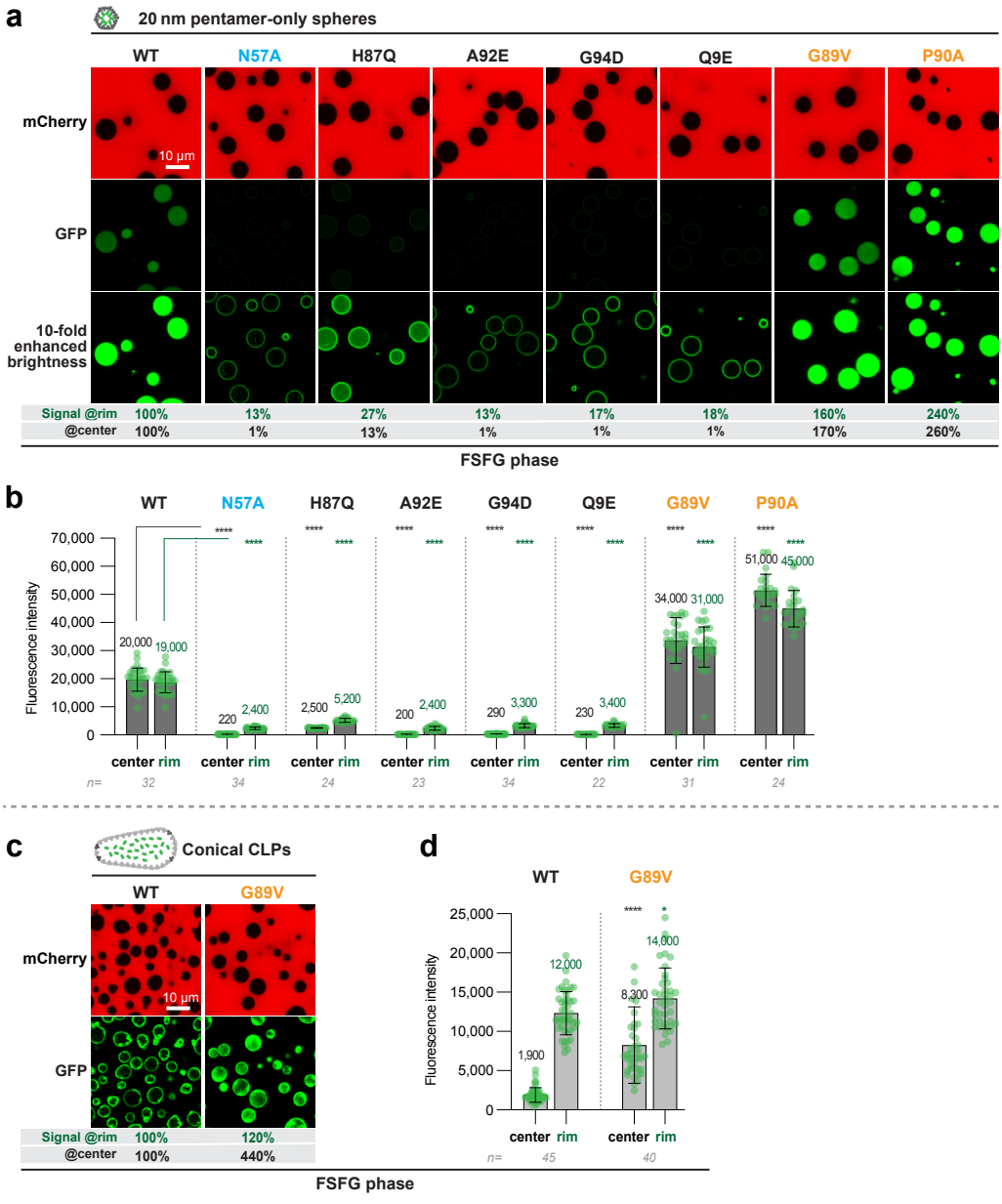
Extended Data Fig.6 | Reduction of FG phase-solubility depends on the accessibility of FG-repelling mutations on the capsid.

a, The FG phase-partitioning of CLPs was tested as in Fig.3, using wild-type and indicated mutants. Partitioning was strongly inhibited by FG-repelling mutations of highly accessible residues (A92E and R97K). When similar mutations were introduced at less accessible positions, partitioning was little affected (R132K) or unaffected (L136E or R143K). **b,** A92, R97, R132, L136 and R143 are shown with indicated coloring in a capsid structure (PDB 3J3Y).



Extended Data Fig.7 | Synergy between surface residues and the FG-binding pocket in FG phase-partitioning of the capsid.

Quantification of the capsid-partitioning into the GLFG phase as shown in Fig.4 but based on larger datasets. Quantification was performed as in Extended Data Fig.5. *P-values*: ****<0.0001, ns: not significant (*P-value* ≥0.05).



Extended Data Fig.8 | Mutations that block or enhance the capsid-partitioning into an FSFG phase.

a,b, Partitioning of pentamer-only spheres into the FSFG phase was as in Extended Data Fig.1, but the indicated CA mutants were also analyzed. The resulting partitioning defects by N57A, A92E, H87Q, G94D, and Q9E exchanges were considerably more severe as compared to the GLFG phase, showing an up to 100-fold reduction of WT. This can be explained by the stronger cohesion of the more hydrophobic FSFG motifs, which necessitates stronger capsid-FG interactions for melting into the phase. Interestingly, the experiment also revealed that the G89V and P90A mutations enhance the partitioning of this capsid species. The effects can be explained by G89V making the surface more hydrophobic and by the P90A exchange allowing for an additional hydrogen-bonding between the alanine amine group and the FG repeats. The P90A enhancement is specific for the pentamer-only spheres. The enhancing effects were not obvious in the GLFG phase experiments, because there, the partitioning was already complete without additional mutations. Quantifications were as in Extended Data Fig.5.

c,d, Partitioning of cone-shaped CLPs into the FSFG phase. The stronger cohesion of this phase results in partitioning defect of wildtype CLPs, which arrest at the surface of the phase. This defect can be compensated by the G89V mutation, which allows for stronger hydrophobic capsid-FG interactions. P-values were calculated using unpaired two-tailed Student's t test.

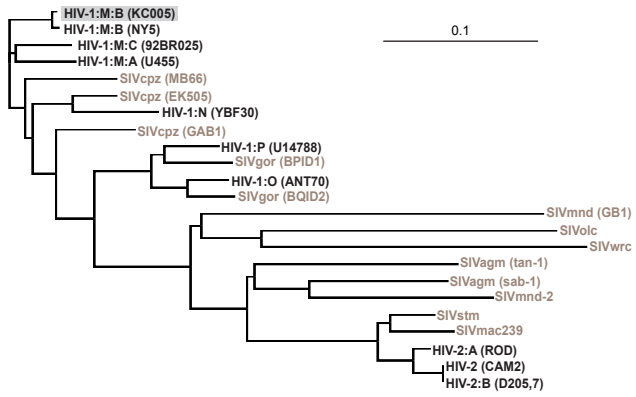
P-values: * < 0.05, ****<0.0001.

Experiments were repeated independently with identical outcomes ($N = 3$). Scale bars, 10 µm.

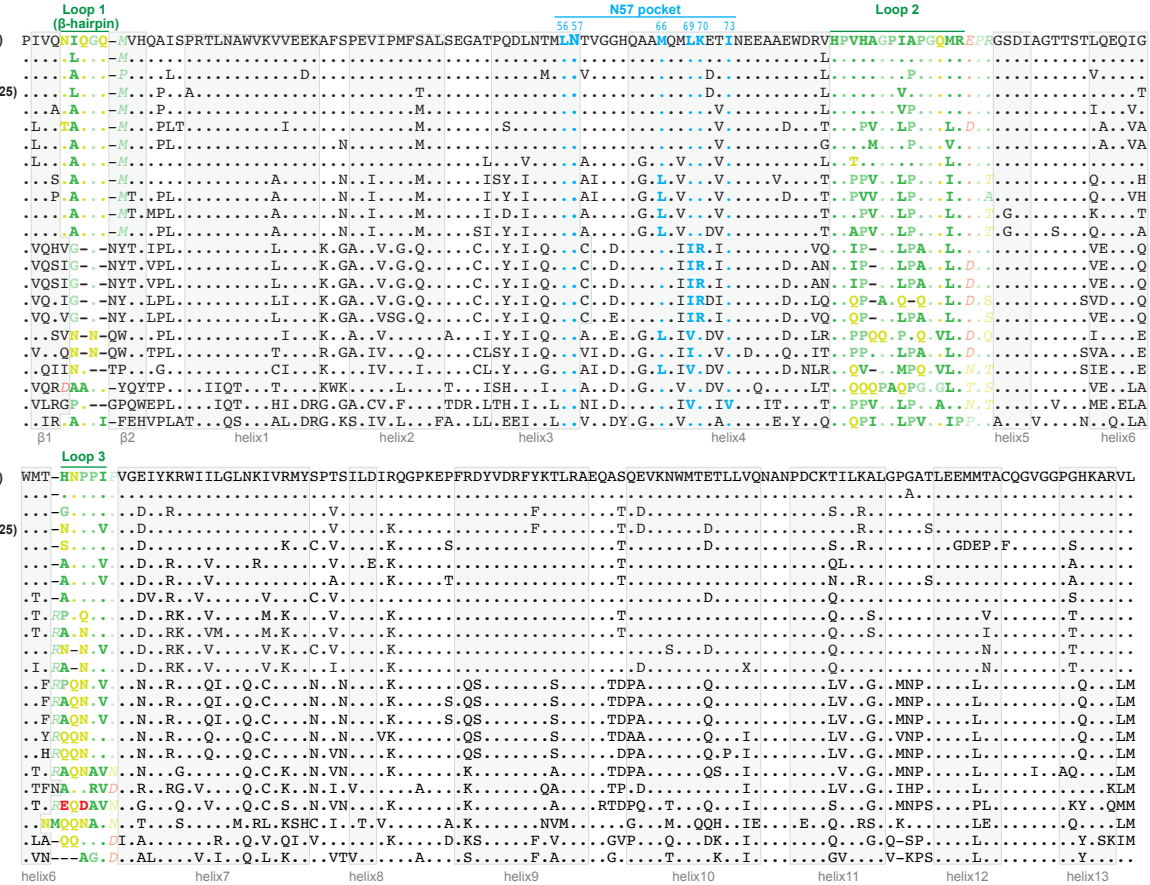
a

Lentivirus	Host
HIV-1:M:B (KC005; NY5)	Human
HIV-1:M:A (U455)	
HIV-1:M:C (92BR025)	
HIV-1:N (YBF30)	
HIV-1:O (ANT70)	
HIV-1:P (U14788)	
HIV-2:A (ROD)	
HIV-2:B (D205,7)	
HIV-2 (CAM2)	
SIVcpz (MB66; EK505; GAB1)	Chimpanzee
SIVgor (BQID2; BPID1)	Gorilla
SIVmac239	Rhesus macaque
SIVstm	Stump-tailed macaque
SIVagm (sab-1; tan-1)	Green monkey
SIVmnd (GAB1; SIVmnd-2)	Mandrill
SIVwrc	Red colobus monkey
SIVolc	Olive colobus monkey

b



c

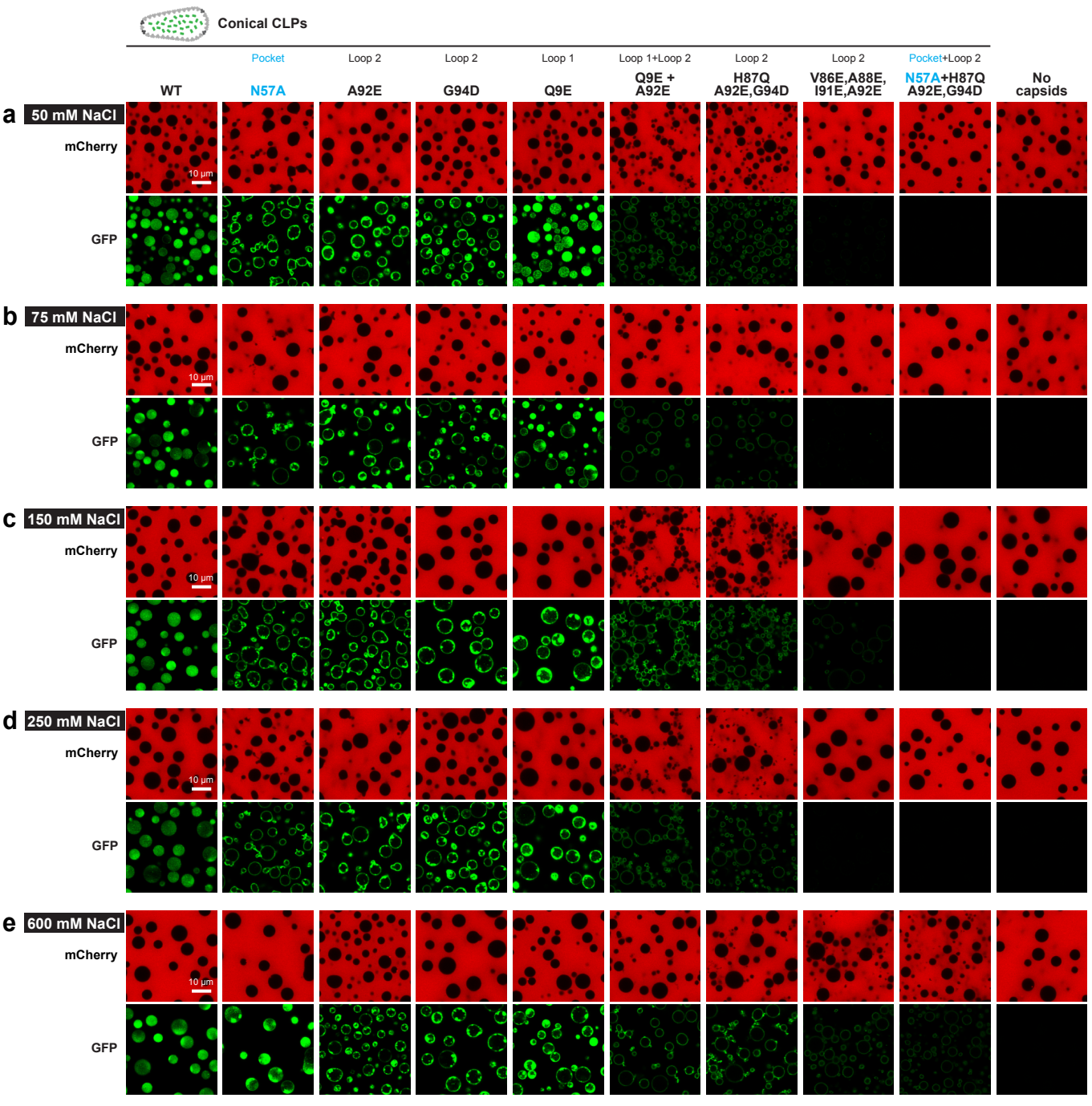


Extended Data Fig.9 | Sequence alignment and phylogenetic analysis of representative lentiviral CA proteins.

a, Representative lentiviruses and their hosts.

b, Neighbor-joining phylogenetic tree of lentiviral CA proteins.

c, A structural alignment of lentiviral CA sequences (based on Alphafold3-modelling). HIV-1 isolate KC005 (M group, subtype B) is the CA species analyzed in this study, therefore serving as a reference. Residues of the N57 pocket that contact the FG motif are colored in cyan. Loop1-3 residues are colored according to the amino acid scale of Fig.2a. Highly exposed residues on the outer loops are in bold. Residues that are not well exposed for FG interaction (either due to forming existing H-bond or salt bridge or due to orientation) are in italic. Secondary structures are indicated by light gray boxes. Sequence identifiers are listed in Supplementary Table S7. For sequence variations within the HIV-1: M group, see also ref.[102,103] and Extended Data Table 1.



Extended Data Fig.10 | FG pocket and capsid surface residues contribute different qualities of FG interactions. Partitioning experiments into the GLFG 12mer FG phase of CLP mutations were performed as in Fig.3, but the NaCl concentration was varied from 50 mM to 600 mM as indicated in each panel. Note that wildtype CLPs showed a perfect FG phase partitioning at all salt concentrations. N57A mutant CLPs showed a strong partitioning defect at 50 mM, 75 mM, 150 mM, and 250 mM NaCl; however, at 600 mM NaCl, the N57A mutant CLPs efficiently entered the FG phase. The A92E and G94D mutations showed a strong partitioning defect at any salt concentration. The Q9E mutant CLPs entered the phase at low salt, but showed a partitioning defect at higher salt. These different behaviors indicate that N57 and the critical surface residues mediate interactions of different qualities.