HIV-1 requires capsid remodelling at the nuclear pore for nuclear entry and integration

Anabel Guedán[1*], Callum D Donaldson[1*], Ophélie Cosnefroy[1], Ian A Taylor[2], Kate N Bishop[1]

[1] Retroviral Replication Laboratory, The Francis Crick Institute, London, United Kingdom

[2] Macromolecular Structure Laboratory, The Francis Crick Institute, London, United Kingdom

* These authors contributed equally

ABSTRACT

The capsid (CA) lattice of the HIV-1 core plays a key role during infection. From the moment the core is released into the cytoplasm, it interacts with a range of cellular factors that, ultimately, direct the pre-integration complex to the integration site. For integration to occur, the CA lattice must disassemble. Early uncoating or a failure to do so has detrimental effects on virus infectivity, indicating that an optimal stability of the viral core is crucial for infection. Here, we introduced cysteine residues into HIV-1 CA in order to induce disulphide bond formation and engineer hyper-stable mutants that are slower or unable to uncoat, and then followed their replication. From a panel of mutants, we identified three with increased capsid stability in cells and found that, whilst the M68C/E212C mutant had a 5-fold reduction in reverse transcription, two mutants, A14C/E45C and E180C, were able to reverse transcribe to approximately WT levels. Moreover, these mutants only had a 5-fold reduction in 2-LTR circle production, suggesting that not only could reverse transcription complete in hyper-stable cores, but that the nascent viral cDNA could enter the nuclear compartment. Furthermore, we observed significant levels of A14C/E45C mutant capsid in nuclear and chromatin-associated fractions implying that the hyper-stable cores themselves entered the nucleus. Immunofluorescence studies revealed that although the A14C/E45C mutant capsid reached the nuclear pore with the same kinetics as wild type capsid, it was then retained at the pore in association with Nup153. Crucially, infection with the hyper-stable mutants did not promote CPSF6 re-localisation to nuclear speckles, despite the mutant capsids being competent for CPSF6 binding. These observations suggest that hyper-stable cores are not able to uncoat, or remodel, enough to pass through or dissociate from the nuclear pore

and integrate successfully. This, is turn, highlights the importance of capsid lattice flexibility for nuclear entry. In conclusion, we hypothesise that during a productive infection, a capsid remodelling step takes place at the nuclear pore that releases the core complex from Nup153, and relays it to CPSF6, which then localises it to chromatin ready for integration.

AUTHOR SUMMARY

The mature viral core of human immunodeficiency virus (HIV) consists of a highly organised lattice formed by capsid molecules that encloses the viral RNA and viral enzymes. This lattice is crucial during the early stages of viral replication, as it has to break down – uncoat – at the right time and place in order for the viral DNA to integrate successfully. Lentiviruses, like HIV, can infect non-dividing cells and are able to access the host cell DNA by entering the nucleus through nuclear pores. Until recently, uncoating was thought to occur in the cytoplasm as the whole core was thought too large to pass through the nuclear pore. However, lately it has been suggested that uncoating might occur at the nuclear pore or even inside the nucleus and the site of uncoating is currently hotly debated. By investigating HIV mutants with an increased lattice stability, we have shown that lattice flexibility is crucial for nuclear entry. Provocatively, we observed hyper-stable mutant capsid in nuclear and chromatin-associated fractions suggesting that uncoating is not required for nuclear entry. Nonetheless, microscopy experiments suggested that these hyper-stable mutants were retained on the inner side of the nuclear pore, and were impaired for downstream events in the nucleus, leading to a severe infectivity defect. Therefore, we believe that an essential uncoating, or capsid lattice remodelling event normally takes place at the nuclear pore.

INTRODUCTION

Upon infection of the target cell, the human immunodeficiency virus 1 (HIV-1) core is released into the cytoplasm before trafficking to the nucleus. During this process, the viral RNA genome is reverse transcribed by the viral reverse transcriptase (RT) into double-stranded DNA forming what is known as the reverse transcription complex (RTC). Inside the nucleus, the viral DNA is the main component of the pre-integration complex (PIC) and is integrated by the viral integrase (IN) into the host cell genome to form a provirus [1].

The mature core of HIV-1 consists of a highly organised lattice of capsid (CA) molecules, encasing the viral RNA genome and associated viral proteins. The lattice is composed of approximately 1500 CA monomers assembled into about 250 hexamers and exactly 12 pentamers, forming a fullerene cone shape [2-5]. The CA monomer is a largely helical protein with two independent folded domains separated by a flexible linker: the N-terminal domain (NTD) and the C-terminal domain (CTD) [2, 6-8]. Structural studies have provided key information on how the CA lattice is organised [8-14]. Intra-hexamer and inter-hexamer interactions at different lattice interfaces contribute to an optimal CA lattice stability. The intra-hexamer interactions include NTD-NTD and NTD-CTD interactions between adjacent CA monomers within a hexamer, whereas the inter-hexamer interactions include dimeric and trimeric CTD-CTD interactions between neighbouring hexamers [11, 14].

As the outer surface of the viral core, the CA lattice protects the viral DNA from cytosolic sensors [15, 16] but it can be recognised by cellular restriction factors that inhibit viral replication [17, 18]. Moreover, from cell entry to integration, the CA lattice interacts with many cellular factors. Some of these are cytoplasmic, including Cyclophilin A (CypA) [19-21], the cellular motors dynein and kinesin-1 [1, 22, 23] and FEZ1 [24]. Others shuttle between the cytoplasm and the nucleus like Transportin-3 (TNPO3) [25-27] and Transportin-1 [28]. There are further interactions at the nuclear envelope with nuclear pore complex (NPC) proteins, particularly nucleoporins 358 (Nup358) and 153 (Nup153) [25, 26, 29-31] and with nuclear proteins such as cleavage and polyadenylation factor 6 (CPSF6); [31-35]. The roles of most of these cellular factors in viral replication are not yet fully understood.

Prior to integration, the CA lattice is assumed to disassemble in a process termed uncoating; however, the timing, location and even the exact definition of this process remains unclear. For many years, uncoating was believed to occur immediately after viral entry [36, 37] but, more recently, data has suggested that it might occur after a short time in the cytoplasm, at the nuclear pore or even, most recently, inside the nucleus [38-44]. Indeed, there is growing evidence to suggest that at least some CA is present in the nucleus, although its oligomeric state and function have yet to be fully characterised [34, 40-43, 45-48]. For this reason, using the term “remodelling” may be more appropriate than uncoating to suggest changes in the CA lattice. We, and others, have also reported a link between CA loss and reverse transcription [38, 49-51]. Uncoating is inhibited if reverse transcription is stalled. Initially, this was taken as evidence for cytoplasmic uncoating [49]. However, when combined with emerging data suggesting CA can reach the

nucleus more rapidly than previously thought [43, 44, 48], it has reopened the debate about where reverse transcription is actually completed during infection, and so does not currently preclude any specific uncoating model.

Importantly, the location and timing of uncoating together with optimal stability of the CA lattice seem to be key to successful infection. Early uncoating, or a failure to uncoat, both have detrimental effects on viral infectivity [15, 52, 53], making the CA lattice an attractive target for drug development [54, 55]. Here we set out to investigate the effects of increased core stability on the early stages of HIV-1 infection. To this end, we introduced mutations into CA predicted to stabilise the mature CA lattice at different interfaces. We show that mutations that create a hyper-stable lattice reduce virus infectivity by inhibiting integration, but only slightly impede reverse transcription. Analysis of CA protein levels within different subcellular fractions during infection revealed higher levels of hyper-stable mutant CA in all fractions over time, including in the nucleus. Finally, immunofluorescence data suggest that the hyper-stable mutant CA lattice is retained around the nuclear pore and that it is unable to promote CPSF6 re-localisation to nuclear speckles. Based upon these observations, we propose a model where hyper-stable mutants are unable to uncoat or “remodel” their capsid lattice to the required extent to successfully deliver the viral DNA for integration into the host cell genome.

RESULTS

Cysteine mutations at different CA lattice interfaces are able to stabilise the viral core in cells.

Previous *in vitro* studies introduced cysteine mutations at the CA NTD-NTD interface to create disulphide bridges in order to stabilise CA hexamers for crystallisation [11, 12]. We decided to use the same approach to examine if similar mutations would increase the stability of the CA lattice during infection of cells, to investigate the effect of core stability on the early stages of HIV-1 infection. We selected CA mutants from the literature and identified additional CA residues to substitute with cysteine residues in order to stabilise all of the different inter-hexamer interfaces with disulphide bridges (listed in Fig 1A and residues highlighted in Fig 1B). The new mutants E180C, L151C/L189C and K203C/A217C were designed based on a previous report demonstrating that disulphide bonds can have a variable Cβ-Cβ inter-residue spacing of between 3.5-4.5Å [56]. Thus, employing a cryo-EM MDFF atomic model of an *in vitro* CA tubular assembly [PDB ID: 3J34; [8]] and crystal structures [PDB ID: 3H4E and 3H47; [11]], we selected residue pairs with a Cβ-Cβ distance within this range at the various interfaces for site-directed mutagenesis. Consequently, we created a panel with at least two mutants at each CA lattice interface. In addition, the P38A mutant was included in the panel as a negative control, as it has shown to be less stable than WT [51, 52]. The original mutant used to determine the crystal structure of a hexameric CA assembly, A14C/E45C/W184A/M185A, [11, 12] was also included as a negative control.

Firstly, we synthesised viral-like particles (VLP) expressing the different CA mutants and a GFP reporter gene. We assessed VLP production by measuring the RT activity in the supernatants of

producer 293T cells using a modified RT ELISA. As expected, the A14C/E45C/W184A/M185A and W184A/M185A mutants that weaken CTD-CTD interactions were severely impaired for virus production (Supplementary Fig 1A). We confirmed that this was not due to a lack of Gag expression of these mutants, by immunoblotting 293T producer cell lysates with an anti-HIV-1 CA antibody and showing that the level of Gag expression in the mutants was similar to WT (Supplementary Fig 1B). Most of the mutants, however, showed similar titres to WT VLP, implying that these mutations did not affect Gag assembly. Of note, the inter-hexamer mutants V181C, L151C/L189C and K203C/A217C were partially impaired for VLP production (Supplementary Fig 1A), likely due to the importance of this interface on virus assembly [3, 53]. Whilst we confirmed that expression of these mutant capsid proteins in 293T producer cell lysates was similar to WT VLP (Supplementary Fig 1C), we detected an additional, smaller, CA band for both L151C/L169C and K203/A217C. This has been reported previously [12, 57] and probably represents a processing defect, such as inappropriate cleavage of CA by the viral protease due to the structural changes [57]. This may be the cause of the reduced viral titres. Thus, based on these results, the L151C/L189C and the K203C/A217C mutants as well as the A14C/E45C/W184A/M185A and W184A/M185A mutants were not included in further experiments.

Next, we performed fate-of-capsid assays, to study the effect of the cysteine mutations on viral core stability in cells (Fig 2). HeLa cells were infected with equal titres of WT and mutant VLP, and cell lysates were harvested at 2 and 20 hours post-infection (hpi). These time points were selected as they have previously been used to monitor the stability of the viral core at the early and late stages of the virus replication cycle [17]. The cell lysates were centrifuged through a

sucrose cushion to be separated into two fractions: free or soluble CA (S) and pelleted CA (P), which contains assembled cores [58]. An input (I) sample was also harvested before centrifugation through the sucrose cushion. The fractions were analysed by immunoblotting with an anti-HIV-1 CA antibody and the ratio of soluble to pelleted CA was determined. Fig 2 shows representative blots of the fate-of-capsid assay for each mutant. A WT sample was included in every assay for comparison. Although there was some variation between assays, the WT samples showed similar amounts of CA in the pellet and soluble fractions (P=S) at 2 hpi and less CA in the pellet (P<S) at 20hpi (Fig 2A), presumably reflecting a reduction in the amount of assembled CA with time due to uncoating. In agreement with previous reports, we recovered much less CA in the pellet fraction than the supernatant fraction from cells infected with the negative hypo-stable control, P38A, [51, 52]. For this mutant, P<S at both 2 and 20 hpi, indicating the small amount of assembled CA at both time points. Therefore, we set a CA ratiometric profile criteria for mutants to be considered hyper-stable as the following: Mutants should consistently show greater amounts of CA in the pellet than in the supernatant (P>S) at 2 hpi and at least equal amounts in both fractions (P=S) at 20 hpi. Three mutants showed this phenotype: M68C/E212C, A14C/E45C and E180C (highlighted with dashed boxes in Fig 2A). The profile of P207C/T216C and the V181C mutants was similar to WT (Fig 2A). At 2 hpi, A42C/T54C showed a CA ratio P=S and, Q63C/Y169C and A204C showed P<S, suggesting that none of these mutants had increased viral core stability compared to WT. Thus, taken together despite the cytoplasm being a reducing environment [59, 60], these data showed that three of the panel of cysteine mutants had increased core stability in cells and could be used to study the effects on replication of stabilizing the viral core.

Most CA mutants are less infectious, regardless of CA lattice stability.

To investigate the effect of the CA mutations on virus infectivity, we infected four cell lines (293T, HeLa, SupT1 and U937 cells) with equal RT units of VSVg-pseudo-typed GFP-reporter WT or mutant VLP and analysed the percentage of GFP+ cells at 72 hpi by flow cytometry (Fig 3). As observed previously, the P38A mutant that had reduced CA lattice stability had a marked reduction in GFP expression in all the cell lines (Fig 3, grey bars). Apart from the inter-hexamer P207C/T216C mutant that showed similar infectivity to WT, the remainder of the mutants also showed decreased infectivity, ranging from ~0.05-10% of WT infectivity. A similar pattern of infectivity was seen across all the cells lines indicating that there were no cell-type specific effects, with generally lower overall infectivity in U937 cells. Interestingly, the level of infectivity did not correlate with the CA lattice stability determined by the fate-of-CA assay (Fig 2) but there was some correlation with the CA interface modified (see colour coding Fig 1A and Fig 3), with the intra-hexamer mutants (blue and orange bars) being more defective than the inter-hexamer mutants (green and red bars). Of the three hyper-stable mutants, A14C/E45C, M68C/E212C and E180C (Fig 3, black arrowheads), the infectivity ranged from 0.7- 4%, 0.07-0.4% and 2-10% respectively between the different cell lines.

Reverse transcription can complete in hyper-stable cores but infectivity is severely inhibited.

Next, to determine which step in replication was affected by the CA mutations, we examined the ability of each mutant to reverse transcribe (Fig 4). 293T cells were synchronously infected with

equal RT units of WT or mutant VLP, and at 0, 1, 2, 4, 6 and 24 hpi, cells were harvested and the DNA extracted and analysed for early (strong stop) and late (second strand) viral cDNA products by qPCR (Fig 4). Following infection with WT VLP, the amount of viral cDNA products increased with time, peaking at 6 hpi (Fig 4A-D, black line). As seen before, infection with mutant P38A resulted in less accumulation of viral cDNA (Fig 4A, grey lines). This ~90% decrease in cDNA accumulation mirrored the 90% decrease in infectivity (Fig 3). Likewise, the P207C/T216C mutant had similar infectivity to WT VLP and had similar levels of reverse transcription (Fig 4B, red lines). However, the hyper-stable mutants A14C/E45C and M68C/E212C had a different phenotype. Surprisingly, despite a ~95% decrease in infectivity, A14C/E45C reverse transcribed to WT levels (Fig 4C, blue lines) and mutant M68C/E212C still produced about 10% the cDNA of WT (Fig 4D, orange lines), even though its infectivity was reduced to less than 1% (Fig 3). Figures 4E and 4F show a summary of the amount of strong stop cDNA accumulated at 6 and 24 hpi, respectively, for all the mutants. The E180C and V181C mutants showed WT levels of strong stop cDNA, A204C accumulated about 10% the cDNA of WT and A42C/T54C and Q63C/Y169C only accumulated ~1% the cDNA of WT. Similar results were obtained when looking at the levels of second strand cDNA (Fig S2A). In general, the level of cDNA accumulation at 24 hours compared to WT correlated with relative level of infectivity compared to WT (Fig 4G and Fig S2B). Plotting these data as a ratio of relative cDNA levels to relative infectivity shows that most mutants have a ratio of less than 3 (Fig 4H and Fig S2C). Therefore, it can be assumed that, in general, the reduction in infectivity results from the effect on reverse transcription. The only mutants where the level of reverse transcription products did not correlate with the level of infectivity was for the three hyper-stable mutants (Fig 4H and Fig S2C). These data suggest that even though increasing core stability has

236 detrimental effects on infectivity, the block is after reverse transcription. This implies that the
237 core does not need to uncoat in order to reverse transcribe. Therefore, although reverse
238 transcription promotes uncoating of WT virus [38, 49-51, 61], reverse transcription is not
239 dependant on uncoating.
240
241 We therefore decided to focus on the A14C/E45C, E180C and M68C/E212C mutants to
242 investigate the effect of core stability on the other early replication steps. First, we examined
243 whether the observed increase in core stability of these mutants was indeed due to disulphide
244 bridge formation as designed, and not due to other stabilising effects of the individual mutations.
245 HeLa cells were infected with equal RT units of WT and mutant VLP and cell lysates were analysed
246 by SDS-PAGE in reducing or non-reducing conditions followed by immunoblotting with an anti-
247 HIV-1 CA antibody (Fig 5A). Samples were either treated with iodoacetamide to prevent
248 artefactual cysteine formation (non-reducing conditions), or dithiothreitol (DTT) to reduce
249 disulphide bonds (reducing conditions) prior to SDS-PAGE. Only monomeric CA of 24KDa was
250 detected for WT VLP, in both reducing and non-reducing conditions. By contrast, disulphide
251 cross-linking of CA monomers was detected in the non-reducing conditions for all three hyper-
252 stable mutants. A strong band was evident at 48 kDa, which corresponds to a CA dimer, and
253 slower migrating species were also present, suggesting higher order oligomers existed. Upon
254 addition of DTT, the higher molecular weight bands disappeared, suggesting that intra- or inter-
255 hexamer disulphide bonds contribute to the increased core stability of these mutants.
256

257 To further investigate the infectivity block to the hyper-stable mutants, we measured the ability
258 of the mutants to complete different steps of replication between reverse transcription and
259 integration. Following synchronous infection of 293T cells with equal RT units of WT or mutant
260 VLP, samples were taken at 6 hpi, 24 hpi and two weeks post infection, and levels of early and
261 late cDNA, 2-LTR circles and integrated viral DNA were quantified by qPCR. It is widely accepted
262 that 2-LTR circles are an indicator of completion of reverse transcription and nuclear entry. Fig
263 5B shows the relative amounts of the different viral DNA products compared to WT VLP.
264 Interestingly, as well as being only slightly defective for reverse transcription, both A14C/E45C
265 and E180C mutant VLP were able to produce 2-LTR circles at ~20% the amount of 2-LTR circles
266 produced by WT VLP. This suggests that at least some of these mutant particles can enter the
267 nucleus. Importantly, these mutants integrated <2% of the amount of cDNA integrated in WT
268 infections, which correlated well with the decrease in infectivity measured. Therefore, there
269 appears to be a minor block to nuclear entry for these two mutants, but, surprisingly, the major
270 defect is after nuclear entry and before integration. In contrast, mutant M68C/E212C VLP only
271 produced 10% of the cDNA of WT, and there was a further reduction in the amount of 2-LTR
272 circles, to <1% of that produced by WT VLP. There was no further decrease in integration
273 compared to the levels of 2-LTR circles, suggesting that the M68C/E212C mutant is more severely
274 impaired earlier in replication than the other two hyper-stable mutants and is blocked for nuclear
275 entry. Altogether, these data suggests that reverse transcription can finish in a hyper-stable core
276 and that this cDNA is able to enter the nuclear compartment, but that there is a block to
277 integration that prevents these hyper-stable mutants successfully forming proviruses.
278

CA protein from the A14C/E45C mutant is detected in nuclear and chromatin fractions.

The nuclear pore has generally been considered too small to allow passage of a complete HIV core, and yet we have shown here that hyper-stable CA mutants can synthesise cDNA and that at least some of this cDNA can reach the nucleus to form 2-LTR circles. This suggests that either the cDNA can escape the core in order to enter the nucleus, despite the lattice being more resistant to disassembly, or that the whole stable lattice can enter the nucleus. Therefore, in order to determine the cellular localisation of components of a hyper-stable core, we measured CA and IN protein levels in different cellular compartments during infection (Fig 6). HeLa cells were synchronously infected with equal RT units of WT or A14C/E45C VLP and cells were harvested at 4, 8, 24 and 30 hpi. The total levels of both CA and IN in whole cell lysates were similar for WT and A14C/E45C (CC) infection at all time points, both decreasing between 8 and 24 hpi (Fig 6A). The infected cell lysates were next processed into the following subcellular fractions: cytoplasm, membranes, nucleus and chromatin-bound. The separation resolution of the subcellular fractionation was determined by looking for distinct markers in the different fractions: HSP90 for the cytoplasm, Calnexin for the membranes, HDAC2 for the nucleus and Histone H3 for chromatin-bound. The representative immunoblot in Fig 6B shows that the fractionation procedure was highly effective, and the markers were predominantly found in the expected fractions. We also probed for Tubulin-α and Lamin B1 to monitor the distribution of the cytoskeleton and the nuclear envelope, respectively. Tubulin-α was highly enriched in the cytoplasmic fraction while Lamin B1 was mainly localised in the nuclear fraction and a small portion seemed to be chromatin-bound. Furthermore, we confirmed that CPSF6 is mainly found

in the soluble nuclear fraction. Having determined that the cellular compartments had been separated successfully, we analysed CA and IN protein levels in the different subcellular fractions by immunoblotting with anti-HIV-1 CA and anti-HIV-1 IN antibodies. The amount of CA in the cytoplasm remained relatively constant throughout the course of the infection (Fig 6C), with noticeably more CA detected following infection with A14C/E45C VLP, especially at 8 and 24 hpi. Conversely, IN levels decreased markedly between 8 and 24 hpi, but there was little difference between WT and A14C/E45C infections (Fig 6C). A similar pattern was observed in the membrane fraction (Fig 6D). This confirms the increased stability of the A14C/E45C CA lattice compared to WT CA, and suggests that IN protein has a shorter half-life than CA. Surprisingly, we detected both CA and IN in nuclear fractions as early as 4 hpi (Fig 6E). Particularly surprising was the amount of A14C/E45C CA detected. Also unexpected was that more CA and IN were detected in the nucleus at the earlier time points than at 24 hours. This suggested that viral cores travel to the nucleus faster than previously thought, and before the peak of reverse transcription at 6 hpi (see Fig 4). Other groups have also recently reported this phenomenon [39, 40, 43, 44, 48]. Finally, we observed a similar, but more extreme, pattern of detection of CA in the chromatin fractions (Fig 6F). Here we observed significantly more A14C/E45C CA than WT CA present at all time points, and, again, A14C/E45C CA was present event at 4 hpi. Interestingly, there was more IN present in this fraction from the WT infections than the A14C/E45C infections at all time points. The association of A14C/E45C CA with nuclear fractions, particularly with the chromatin-bound fraction, suggests that, despite being hyper-stable, this CA protein could enter the nucleus. However, as the nuclear envelope marker, Lamin B1, was also distributed in these two fractions, it could also be hypothesised that A14C/E45C was accumulating at the nuclear

envelope rather than inside the nucleus. To help resolve these options, we performed cellular localisation experiments.

The A14C/E45C mutant is impeded at nuclear entry.

In order to investigate nuclear entry of the A14C/E45C mutant further, we monitored the interactions between CA and the nuclear pore proteins Nup358 and Nup153 during infection. These proteins are part of the nuclear pore complex (NPC) where Nup358 faces the cytoplasm and Nup153 forms part of the nuclear basket of the NPC, facing the nucleus. Furthermore, both proteins have previously been described to bind HIV-1 CA directly and to be involved in HIV-1 replication [25, 26, 30]. HeLa cells were synchronously infected with equal RT units of WT or A14C/E45C VLP for 2, 4, 6, 8 and 10 hpi. At each time point, cells were fixed, and proximity ligation assays (PLA) were performed with specific antibody combinations: anti-HIV-1 CA and anti-Nup358 (Fig 7A-C), and anti-HIV-1 CA and anti-Nup153 (Fig 7D-F). In this assay, direct protein-protein interactions or proteins spaced < 40 nm apart can be visualised as foci by immunofluorescence. In addition, by doing *post hoc* analysis, the localisation of these interactions can be analysed in relation to the nuclear envelope, in this case approximated by DAPI staining. Fig 7A shows representative cells assayed for CA-Nup358 co-localisation at 8 hpi. Counting the number of foci per cell (Fig 7B) revealed that there were similar levels of interaction between CA and Nup358 throughout the time course of infection, and there was no significant difference between WT and A14C/E45C infections. The localisation of these foci (Fig 7C) seemed to be around the DAPI edge for both VLP, but with a tendency for WT foci to be further into the

nucleus than A14C/E45C foci. This suggests that there is little difference in the early stages of infection, up to reaching the nuclear pore, between WT and A14C/E45C VLP, and, once again, that cores reach the nucleus within 2 hpi, sooner than previously expected. In contrast, there were clear differences in the staining for CA-Nup153 co-localisation between WT and A14C/E45C infections (Fig 7D-F). Analysis of the number of foci per cell showed that there was increased A14C/E45C foci compared to WT infections at all time points (Fig 7E). As equal numbers of cores were arriving at the nucleus, as measured by CA-Nup358 staining (Fig 7B), this suggests that the A14C/E45C cores were spending more time at the nuclear pore. Moreover, the A14C/E45C foci appeared to localise at the cytoplasmic side of the DAPI edge while the WT foci were mainly on the nuclear side (Fig 7F). Given that a layer of lamin and the nuclear envelope surround the chromatin in the nucleus, the DAPI edge does not indicate the exact nuclear/cytoplasmic boundary, but rather the inner side of the nuclear pore. Thus, being on the cytoplasmic side of the DAPI staining, together with increased CA-Nup153 staining, suggests that A14C/E45C hyper-stable cores are being trapped at the NPC.

CPSF6 is not re-localised to nuclear speckles during infection with hyper-stable mutants.

In addition to binding nuclear pore proteins, CA interacts with CPSF6, which has been reported to direct HIV-1 integration site specificity. Specifically, CPSF6 is proposed to locate HIV-1 to highly transcribing regions of the genome identified as nuclear speckles, or speckle-associated domains (SPADs) [9, 34, 35]. The absence of CPSF6, or blocking of CPSF6-CA binding, promotes integration into lamina-associated domains (LADs), regions of heterochromatin near the nuclear envelope

[9, 34]. Unfortunately, despite repeated attempts, we were unable to optimise our PLA assay to reliably visualise CA-CPSF6 co-localisation. However, recently, it has been observed that CPSF6 only redistributes to SPADs during WT HIV-1 infection and not during infection with CPSF6-binding deficient mutants A77V and N74D [35, 48]. This suggests that the CA-CPSF6 interaction is driving CPSF6 reorganisation in the nucleus. Therefore, to investigate the interaction between our hyper-stable mutants and CPSF6, we monitored the reorganisation and redistribution of CPSF6 to SPADs during infection. HeLa cells were synchronously infected with equal RT units of WT, A14C/E45C, E180C or M68C/E212C VLP. At 16 hpi, cells were fixed and immuno-stained with antibodies against HIV-1 CA, CPSF6 and SC35, also called serine and arginine rich splicing factor 2 (SRSF2), which is a marker for SPADs [62]. Due to the species of the antibodies, we could only co-stain for pairs of markers at a time. Fig 8A and Fig S3 show that cells infected with WT-HIV-1 exhibit a compelling redistribution of CPSF6 into puncta that co-localise with SC35-positive nuclear speckles, confirming previous reports [35, 48]. However, there was no such redistribution of CPSF6 during infection with any of the hyper-stable mutants which showed similar CPSF6 staining to uninfected cells. Importantly, when we co-stained for CA and CPSF6, ~80% of WT CA positive cells contained CPSF6 puncta (Fig 8B), compared to only 5% of hyper-stable CA positive cells. Although we were able to detect some CA signal in the nucleus of WT HIV-1 infected cells that colocalised with the CPSF6 signal by immunofluorescence (Fig 8B), the CA signal was relatively weak and we cannot say whether this CA was part of a core or not. The CA staining in the nucleus was not as evident during infection with the hyper-stable mutants. Together, these data show that WT CA is able to interact with CPSF6 and induce its redistribution to SPADs, whilst the hyper-stable CA fail to alter CPSF6 localisation.

DISCUSSION

It has been known for many years that HIV-1 can infect non-dividing cells and so must cross the nuclear envelope during infection. Until recently, the dogma was that HIV-1 cores uncoated in the cytoplasm during reverse transcription, because they were too large to cross the nuclear pore. However, new evidence has led people to re-examine this order of events. When and where the CA lattice breaks down is currently highly controversial. Theories range from cytoplasmic uncoating to uncoating at the nuclear pore (reviewed in [63]) [35, 38, 39, 42, 64], and more recently, uncoating inside the nucleus [43, 44, 48, 65]. Along these lines, emerging electron microscopy images appear to show intact cores going through nuclear pores [65]. Some open questions are (1) Can reverse transcription complete inside a core? (2) Can a complete core cross the nuclear pore? (3) What replication events require CA? and (4) When is uncoating required? Unfortunately, uncoating is difficult to measure directly. Therefore, in this study, we have taken a genetic approach to investigating uncoating by comparing the replication of WT HIV-1 with hyper-stable CA mutants that are potentially slower or unable to uncoat. Thus, we have not asked "where/when does uncoating occur?", but rather "where/when does viral replication get blocked if uncoating is prevented?".

We compared a panel of mutants designed to induce disulphide bonds at different CA lattice interfaces. Many of these have previously been used for structural studies and clearly induced stability of *in vitro* assembled CA [7, 8, 11, 12, 57, 66-69], but anecdotally were thought unlikely

411 to make stabilising disulphide bonds in the reducing environment of the cytoplasm. Nevertheless,
412 we found that two previously described intra-hexamer mutants, A14C/E45C and M68C/E212C
413 [11, 12, 57], and a novel inter-hexamer mutant, E180C, all showed increased lattice stability in
414 cells (Fig 2A and Fig 6) and maintained disulphide bonds following infection (Fig 5A). Interestingly,
415 these three mutants represent stabilising different interfaces within the CA lattice. Thus, it
416 appears that no particular interface is dominant in lattice stability. It was not surprising that most
417 of the mutants did not increase core stability, as there is a lot of variation in the exact Cβ-Cβ
418 distances between individual residues, depending on the region of the lattice that they reside
419 (i.e. depending on the curvature of that part of the fullerene cone [5, 66]). This is especially true
420 around the trimeric interface. It is intriguing that the NTD-NTD mutant A14C/E45C resulted in
421 hyper-stable cores whilst the NTD-NTD mutant A42C/T54C did not, as they have near identical
422 hexamer crystal structures [12, 67]. However, the geometry and exact chemistry of the
423 surrounding environment likely affect whether disulphide bonds form in cells.
424
425 **Uncoating is required between reverse transcription and integration**
426
427 As expected, although they produced normal viral titres as measured by particle release (Fig S1),
428 most of the mutants in the panel had reduced infectivity (Fig 3). The intra-hexamer mutants were
429 the least infectious, perhaps highlighting the importance of residues within the CA N-terminal
430 domain for early replication steps. The severity of the infectivity defect did not correlate with
431 core stability, as measured by the fate-of-CA assay, but did generally reflect the level of reverse
432 transcription (Fig 4G and H). The exceptions to this were the three mutants identified as hyper-

433 stable. Surprisingly, despite uncoating previously being linked to reverse transcription [38, 49,
434 51, 61, 70], A14C/E45C and E180C were able to reverse transcribe to approximately WT levels
435 and, furthermore, showed only a partial reduction in 2-LTR circle production (Fig 4 and Fig 5B).
436 This shows that reverse transcription can complete in a hyper-stable core and implies that it is
437 not necessary for the core to uncoat in order to accommodate the double stranded DNA. These
438 results are in line with previous studies showing that hyper-stable mutants E45A, Q63A/Q67A,
439 5Mut (Q67H/K70R/H87P/T107N/L111) and A14C/E45C can reverse transcribe [48, 52, 71, 72] and
440 with a recent report saying that reverse transcription can complete in whole WT cores [43].
441 However, it is possible that the hyper-stable cores are partially opened. Indeed, the production
442 of 2-LTR circles would imply that the core was open to some extent in order for the cellular ligases
443 to access the viral DNA. Recent *in vitro* studies have suggested that reverse transcription
444 increases the pressure inside the core [61], inducing mechanical changes in the capsid that
445 progressively remodel the lattice [73, 74] and may result in viral DNA loops bursting out of
446 partially uncoated cores [75]. This agrees with our previous assessment that uncoating is
447 triggered after first strand transfer during reverse transcription [51]. By inducing disulphide bond
448 formation, we may have prevented some or all of these remodelling changes occurring in the
449 hyper-stable mutant cores, and therefore prevented more extensive uncoating from occurring.
450 This implies that whilst reverse transcription promotes uncoating, it is not dependent on the
451 disassembly of the lattice. Instead, the virus appears to be using reverse transcription to time
452 uncoating.
453 Nonetheless, hyper-stability can be detrimental for reverse transcription as the M68C/E212C
454 mutant showed a 10-fold decrease in reverse transcription products compared to WT (Fig 4D).

455 The cysteine residues introduced in M68C/E212C are located on the intra-hexamer NTD-CTD
456 interface, which is important for the curvature of the lattice. Christensen *et al*. recently reported
457 that the compound GS-CA1, which binds at the NTD-CTD interface, affected capsid integrity and
458 strongly inhibited reverse transcription *in vitro*, and proposed that affecting the NTD-CTD
459 interface introduces lattice strain and promotes capsid fracturing [75]. Thus, it could be possible
460 that the formation of disulphide bonds at this specific interface affects the structure of the CA
461 lattice in such a way that is not compatible with reverse transcription.
462 Importantly, even the mutants that reverse transcribed well were still markedly impaired for
463 integration (Fig 5B), presumably meaning that the cores are unable to open fully, or in the correct
464 manner to allow integration. This agrees with various recent reports suggesting that a final capsid
465 uncoating reaction needs to occur before integration can take place [35, 39, 42-44, 48, 64, 65].
466 Stabilising the CA lattice by either mutation, or drugs, probably affects flexibility as well as lattice
467 break down, so it is hard to separate whether the core needs to restructure in some way or to
468 break apart completely. We therefore use the term “CA remodelling” to cover both of these
469 events.
470
471 The link between reverse transcription and uncoating has recently taken a new twist, as the
472 location of reverse transcription in the cell has been questioned. There is increasing evidence
473 that CA-containing viral complexes reach the nucleus early, before reverse transcription has
474 completed [35, 43, 44, 48, 76, 77], and that reverse transcription actually completes in the
475 nucleus [44]. In support of nuclear-associated reverse transcription, here, we also detected CA

476 and, importantly, IN in nuclear fractions at 4 hpi (Fig 6E) and detected CA interacting with nuclear
477 pore proteins at 2 hpi (Fig 7), well before the peak of reverse transcription at 6hpi (Fig 4).
478
479 **Hyper-stable CA mutants are compromised for nuclear entry**
480
481 As two of our hyper-stable mutants seemed to have a major defect after reverse transcription,
482 we asked whether these hyper-stable cores could enter the nucleus. 2-LTR circles are considered
483 a surrogate for nuclear entry [78, 79], and as A14C/E45C and E180C produced 2-LTR circles, albeit
484 with a 5-fold decrease compared to WT (Fig 5B), it suggested that they could indeed enter the
485 nucleus. Our subcellular fractionation experiments confirmed that both WT and hyper-stable CA
486 were detected in all fractions following infection, including both soluble and chromatin-
487 associated nuclear fractions (Fig 6). To confirm where the nuclear envelope fractionated, we
488 blotted for lamin B1 which is located inside the inner layer of the nuclear envelope. Lamin B1 was
489 detected in the nucleus and, to a lesser extent, in the chromatin fraction (Fig 6B), suggesting that
490 the CA in these fractions could be at the nuclear envelope as well as inside the nucleus.
491 Importantly, chromatin is known to be tightly linked with nuclear pores and lamin [80]. There
492 was noticeably less WT HIV-1 CA present at each time point than the hyper-stable mutant CA,
493 particularly in the nuclear fractions, suggesting that it was turned over faster (Fig 6E and F),
494 although it is also possible that WT CA was more susceptible to degradation during the
495 fractionation process. Conversely, the levels of IN were similar at all time points between WT and
496 hyper-stable infections in all fractions except the chromatin associated fraction, where there was
497 more IN detected following WT infections. As WT CA was barely detectable here (Fig 6F), this

498 could reflect WT cores being able to breakdown and deliver IN more efficiently to the chromatin
499 than A14C/E45C cores, and that once disassembled, the CA is degraded. Together, these data
500 show that although the hyper-stable cores retain more CA than WT cores, the IN levels are
501 comparable until the cores meet chromatin. This implies that uncoating has little impact on
502 replication until a chromatin associated event.
503
504 To explore whether hyper-stable cores could truly enter the nucleus or were associating with
505 nuclear membranes, we studied the dynamics of WT and A14C/E45C CA interactions with nuclear
506 pore proteins (Fig 7). We chose Nup358 and Nup153 because they are located on the outer and
507 inner face of the nuclear pore, respectively, and because their interaction with CA is well
508 characterised [25, 26, 29, 30]. We found that WT and A14C/E45C CA showed similar levels of
509 interaction with Nup358 over time (Fig 7B) suggesting that both cores could reach the nuclear
510 pore with equivalent kinetics, agreeing with our fractionation experiments. As seen in a previous
511 report [81], we observed some CA-Nup358 foci in the cytoplasm, which was more noticeable
512 during A14C/E45C infection, but there were no significant differences between the WT and
513 mutant CA distribution (Fig 7C). In contrast, there was clear divergence in the co-localisation of
514 CA and Nup153 between WT and hyper-stable cores. There were significantly increased numbers
515 of foci for the A14C/E45C CA at all time points. As equal numbers of cores were arriving at the
516 nucleus (as measured by Nup358 co-localisation), this suggests that the A14C/E45C cores were
517 spending more time at the nuclear pore. Moreover, the A14C/E45C foci appeared to localise
518 more towards the cytoplasmic side of the DAPI edge while the WT foci were mainly on the nuclear
519 side, suggesting that WT cores travel further into the nucleus and that A14C/E45C hyper-stable

cores are trapped at the NPC. It is worth pointing out that quantifying the distance of the PLA foci to the DAPI edge has caveats. The PLA signal is the result of a complex of rolling circle amplification products together with antibodies bound to two proteins of interest. As this is a big complex, the foci distance to the DAPI edge cannot be taken as an absolute distance. However, the same caveats apply to both WT and mutant cores, so a relative comparison can be made. Thus, it seems that uncoating or lattice flexibility is required for nuclear entry and points to a CA remodelling event at the nuclear pore, as suggested by some other labs [42, 65]. It is interesting to recall that the A14C/E45C mutant still produces 2-LTR circles, despite apparent limited nuclear entry, questioning what this product really represents.

Interestingly, studies with the CPSF6 binding-defective HIV-1 CA mutants, N74D and A77V, have reported that a longer residence time at the nuclear envelope promotes integration into heterochromatin regions close to the nuclear envelope [9, 34, 35, 43, 48]. Furthermore, unlike WT HIV-1, neither the N74D nor the A77V mutant can re-localise CPSF6 to SPADs, showing that this is a CA dependent event [9, 34, 35, 48]. Since A14C/E45C cores were stalled at the nuclear pore, we were intrigued to see whether this mutant would be able to promote CPSF6 redistribution to nuclear speckles. In agreement with very recent reports [35, 48], we found that only WT-infected cells showed a redistribution of CPSF6 to SC35 positive puncta (Fig 8). Importantly, A14C/E45C is still able to bind CPSF6 *in vitro* [82, 83], suggesting that either it does not have access to CPSF6 in cells, or that it is unable to relocate to nuclear speckles because it is retained elsewhere. CPSF6 binds to the same pocket on the CA lattice as Nup153 [29, 31, 32] and Bejarano *et al* have recently reported the consecutive binding of the hexameric CA lattice to

Nup153 and then CPSF6 in macrophages [40]. Thus, we speculate that if A14C/E45C is still binding Nup153 at the nuclear pore, it might not be able to uncouple and move on to binding CPSF6. We hypothesize that the CA lattice needs to remodel at the pore in order to be released from Nup153 and move into the nucleus, where it can then interact with CPSF6 and move to an optimal site for integration. As CA is required for the CPSF6 interaction, it follows that a fraction of CA must be retained by the pre-integration complex until chromatin binding, but sufficient CA must be removed to allow IN and the viral cDNA to access chromatin for integration itself to occur.

Models of HIV-1 uncoating

We have illustrated the current models of HIV-1 uncoating in Fig 9. From left to right, uncoating in the cytoplasm, uncoating at the nuclear pore, uncoating inside the nucleus or uncoating at the integration site. Early uncoating or a failure to uncoat both have detrimental effects on virus infectivity. The data presented here point to a remodelling event at the nuclear envelope that, although it is not required for reverse transcription to complete, is needed to completely pass through the nuclear pore, allow binding to CPSF6 and, ultimately, for integration. We therefore think it likely that some sort of uncoating event begins at the nuclear pore and finishes at the site of integration. We have also previously shown that the murine leukaemia virus (MLV) p12 protein binds directly to both CA and nucleosomes, tethering the MLV core to chromatin during mitosis, and have proposed that p12 could be acting in a similar capacity to CPSF6 in HIV-1 infection [84]. Figure 9 includes our model of MLV uncoating. As MLV cannot traverse nuclear pores [85] it is unlikely that it undergoes a similar pore-induced CA remodelling event, suggesting that MLV

uncoating might be triggered by chromatin binding. As MLV must wait for mitosis, it is possible that an intact CA lattice is needed until integration to provide a protective environment for the viral cDNA.

In conclusion, we have demonstrated that HIV-1 with a hyper-stable CA lattice is able to reverse transcribe successfully but is stalled at nuclear entry, which has a negative effect on CPSF6 binding and integration. We suggest that an uncoating or CA remodelling event normally occurs at the nuclear pore and that this is essential for the later replication events that take place in the nucleus. Furthermore, our observations suggest that viral cores are present at the nucleus before reverse transcription is completed. Therefore, it is plausible that both reverse transcription and uncoating finish in the nucleus. Further work is needed to fully understand the state of the CA lattice in the nucleus and the exact relationship between uncoating and reverse transcription.

MATERIAL & METHODS

Cell lines

Adherent cell lines, 293T, HeLa and Vero cells, were maintained in Dulbecco's modified Eagle medium (Thermo Fisher), and suspension cell lines, SupT1 and U937 cells, were maintained in RPMI-1640 (Thermo Fisher). All cell lines were authenticated and tested mycoplasma-free from Bishop laboratory cell stocks. Media was supplemented with 10% heat-inactivated foetal bovine

serum (FBS; Biosera) and 1% Penicillin/Streptomycin (Sigma). Cells were grown in a humidified incubator at 37°C and 5% CO_2.

Plasmids and site-directed mutagenesis

The plasmids used to produce HIV-1 VLP, pVSV-G, pCMVΔR8.91 and pCSGW, have been described previously [86, 87]. To create Gag-Pol plasmids carrying cysteine-substitution mutations in CA, site-directed mutagenesis was performed on pCMVΔR8.91 using the QuickChange II-XL site-directed mutagenesis kit (Agilent) according to manufacturer's instructions and using the primers listed in Supplementary Table 1. Repeated site-directed mutagenesis was performed to create the double, triple and quadruple mutants. The introduction of the desired mutations was confirmed by Sanger sequencing (Source Bioscience).

Virus-like particle (VLP) production

HIV-1 virus-like particles (VLP) were produced by co-transfecting 293T cells with a 1:1:1 ratio of three plasmids: pVSV-G, pCSGW (GFP-reporter) and pCMVΔ8.91 (or pCMVΔ8.91 mutants). Approximately 16h post-transfection, cells were treated with 10mM sodium butyrate for 8h and VLP-containing supernatants were harvested 24 h later. VLP titres were analysed using a Lenti RT ELISA kit (Cavidi) following manufacturer's instructions. For the fate-of-capsid, cell fractionation and PLA assays, VLPs were concentrated by ultracentrifugation through a 20% (w/w) sucrose cushion using a Beckman SW32Ti rotor (Beckman-Coulter) at 20,200 rpm for 2h at 4°C.

607

608 **Single round infectivity assay**

609

610 293T, HeLa, SupT1 and U937 cells were challenged with normalised amounts of WT and mutant
611 VLP based on their RT activity and incubated for 72 h at 37°C. The percentage of GFP-expressing
612 cells was analysed by flow cytometry using a FACS VERSE, LSR or Fortessa A analysers (BD
613 Biosciences). Data was analysed using FlowJo software.

614

615 **Immunoblotting**

616

617 Cells were lysed in ice-cold radioimmunoprecipitation assay (RIPA) buffer (Thermo Fisher)
618 supplemented with a protease inhibitor cocktail (Roche) and proteins were separated on 4-12%
619 or 10% Bis-Tris SDS-PAGE gels (Thermo Fisher). To assess disulphide cross-linking in cells, 4x SDS
620 sample buffer (Thermo Scientific) was added to cell lysates prior to treatment with 50mM
621 Iodoacetamide (Sigma) or 0.1M DTT (Sigma). Samples treated with Iodoacetamide were
622 incubated for 30min at room temperature followed by a 15 min incubation at 65°C. Samples
623 treated with DTT were boiled at 95°C for 5 min. Samples were then applied to 8% Bolt Bis-Tris
624 Plus gels (Thermo Fisher) and electrophoresed in Bolt MOPS buffer (Thermo Fisher). Primary
625 antibodies used were: anti-HIV-1 CA (in house), anti-HIV-1 IN (in house), anti-HSP90 (CST; #4874),
626 anti-Calnexin (CST; #2679), anti-HDAC2 (CST; #5113), anti-Histone 3 (CST; #14269), anti-Tubulin-
627 α (Bio-Rad; VMA00051) and anti-Lamin B1 (Proteintech; 66095-1-Ig). Secondary antibodies used
628 were: anti-mouse and anti-rabbit HRP-conjugated secondary antibodies (Thermo Fisher; 61-6520

and 31460); anti-mouse IRDye 800CW secondary antibodies (LICOR). Blots were analysed on a Chemidoc imaging system (Bio-Rad) or by X-ray film exposure or on an Odyssey CLx imaging system (LICOR).

Fate-of-capsid assay

The fate-of-capsid assay was performed as previously described (Yang *et al*, 2014) but with some modifications. HeLa cells were seeded at 10^6 cells/well in 6-well plates one day prior infection and spinoculated (1600rpm at 16°C for 30min, followed by 37°C for 1h) with equal amounts of VLP based on their RT activity. Cells were harvested at 2 and 20 hpi and cell pellets were lysed by adding hypotonic buffer (10mM Tris-HCl pH 8.0, 10mM KCl and 1mM EDTA pH 8.0 supplemented with a protease inhibitor cocktail (Roche)) and passing through a Qiashredder column (Qiagen). At this point, 50 µL was harvested as the input (I) and resuspended in 2x Laemmli's SDS-PAGE sample loading buffer (Sigma). The remaining lysate was layered on top of a 30% (w/w) sucrose cushion and centrifuged using a Beckman SW41 rotor (Beckman Coulter) at 32,000rpm and 4°C for 1 h, to separate soluble and assembled CA. After centrifugation, 500 µL of the uppermost portion of the supernatant was harvested as the soluble fraction (S). Following aspiration of the sucrose cushion, the pellet (P) was resuspended in 100 µL of 2x sample loading buffer. The soluble fraction was precipitated using methanol-chloroform extraction and resuspended in 100 µL of 2x sample loading buffer. Then, the I, S and P samples were analysed by western blotting with an anti-HIV-1 CA antibody.

651 **Quantitative PCR analysis to measure RT products**

652

653 Quantitative PCR analysis was conducted as previously described (Bishop *et al*, 2008; Cosnefroy
654 *et al*, 2016). Briefly, VLP were treated with 20 units/mL RQ1-DNase (Promega) in 10mM $MgCl_2$
655 for 1h at 37°C before infection. 293T cells were spinoculated (1600rpm at 16°C for 30min)
656 followed by a 30min incubation at 37°C. The reverse transcriptase inhibitor Nevirapine (NVP;
657 Sigma) was used as a negative control, prepared in DMSO and used at a final concentration of 10
658 µM. Cells were harvested at the indicated times post-infection and total DNA was extracted using
659 the DNeasy Blood & Tissue kit (Qiagen). The extracted DNA was digested with 1 unit/µL DpnI
660 (Thermo Fisher) for 2.5 h at 37°C. qPCR was performed in TaqMan real-time PCR master mix
661 (Thermo Fisher) with 900 nM primers and 250 nM probes. The reactions were performed on a
662 7500 fast real-time PCR system (Applied Biosystems). To calculate DNA copy numbers, standard
663 curves were generated from serial dilutions of pCSGW or p2-LTR junction in 293T cellular DNA.
664 The following primers and probes were used; strong stop cDNA products: *for* 5'-
665 *TAACTAGGGAACCCACTGC*, *rev* 5'-*GCTAGAGATTTTCCACACTG* and *probe* 5'-*FAM-*
666 *ACACAACAGACGGGCACACACTA-TAMRA*; second strand cDNA products: *for* 5'-
667 *TAACTAGGGAACCCACTGC*, *rev* 5'-*CTGCGTCGAGAGAGCTCCTCTGGTT* and *probe* 5'-*FAM-*
668 *ACACAACAGACGGGCACACACTA-TAMRA*; 2-LTR junction: *for* 5'-*GTGTGTGCCCGTCTGTTG*, *rev* 5'-
669 *CAGTACAAGCAAAAAGCAGATC* and *probe* 5'-*FAMGGTAACTAGAGATCCCTCAGACCTAMRA*.

670

671 **Cell fractionation assay**

672

HeLa cells were infected with equal amounts of WT and mutant VLP (based on their RT activity) by spinoculation (1600rpm, 16°C for 2h) followed by a 30 min incubation at 37°C. At the indicated times post-infection, cells were harvested in parallel either as a whole cell lysate (WCL) or for processing with the subcellular protein fractionation kit for cultured cells (Thermo Fisher) following manufacturer's instructions. Protein content in the WCL and in the different fractions was measured by BCA assay (Thermo Fisher) using a FLUOstar Omega plate reader (BMG Labtech). Relative amounts of each fraction compared to the WCL were analysed by immunoblotting.

Immunofluorescence

HeLa cells were seeded at $0.8x10^5$ cells/well on 13 mm glass coverslips in 24-well plates the day prior to infection. Cells were infected with equal RT units of WT or mutant VLP by spinoculation (1600rpm, 16°C for 2h) followed by a 30min incubation at 37°C. At the indicated times post-infection, cells were washed twice with ice-cold PBS, fixed with PBS supplemented with 4% paraformaldehyde for 5min at room temperature followed by an ice-cold methanol incubation for 5min at -20°C, and washed twice again with ice-cold PBS. After permeabilization with 0.5% saponin (Sigma) in PBS for 30min at room temperature, cells were blocked in 5% donkey serum (DS; Sigma) and 0.5% saponin in PBS for, at least, 1h at room temperature. Then, cells were incubated with the following primary antibodies: anti-HIV-1 CA (in house), anti-CPSF6 (Atlas; HPA039973) and anti-SC35 (Abcam; ab11826), diluted in 1% DS with 0.5% saponin in PBS (antibody buffer) for 1h at RT. After three washes with PBS, cells were incubated with the

following secondary antibodies: goat anti-mouse-AF488 (abcam; ab150117), donkey anti-rabbit-AF568 (abcam; ab175692), donkey anti-mouse-AF647 (Thermo Fisher; A-31571) and donkey anti-rabbit-AF647 (Thermo Fisher; A-31573) in antibody buffer for 1h at RT. After three washes with PBS, the coverslips were mounted on glass slides with ProLong gold antifade mountant with DAPI (Thermo Fisher).

Proximity ligation assay (PLA)

HeLa cells were seeded at 10^5 cells/well on 13mm glass coverslips in 24-well plates the day prior to infection. Cells were infected with equal RT units of WT or mutant VLP by spinoculation (1600rpm, 16°C for 2h) followed by a 30min incubation at 37°C. At the indicated times post-infection, cells were washed twice with ice-cold PBS, fixed with PBS supplemented with 4% paraformaldehyde for 5min at room temperature followed by an ice-cold methanol incubation for 5min at -20°C, and washed twice again with ice-cold PBS. Then the immunofluorescence protocol was followed using the following pairs of primary antibodies: anti-HIV-1 CA (in house) and anti-Nup358 (Abcam; ab64276); anti-HIV-1 CA (in house) and anti-Nup153 (Abcam; ab84872), diluted in 1% DS with 0.5% saponin in PBS for 1h at RT. From this point, the Duolink PLA fluorescence detection kit's protocol (Sigma) was followed. After three washes with PBS, coverslips were incubated with secondary antibodies conjugated to PLA probes (anti-mouse PLUS or anti-rabbit MINUS; Sigma) in antibody buffer (Sigma) for 1h at 37°C. Coverslips were incubated with ligase for 30 min at 37°C followed by amplification by polymerase for 100 min at 37°C. From the primary antibody incubation, coverslips were washed with buffers A and B (Sigma) as

717 indicated in the manufacturer's protocol. Finally, coverslips were mounted with the Duolink *in*
718 *situ* mounting media with DAPI (Sigma) on glass slides (Menzel-Gläser) and sealed. Samples were
719 visualised on a SP5 inverted confocal microscope (Leica) using a 63X objective. Longitudinal Z-
720 series were acquired with 0.5 µm step sizes. Image analysis was performed using the GIANI plug-
721 in in Fiji (described in [88]). The number of foci and their distance to the edge of the DAPI staining,
722 were quantified taking into account the 3D cell volume reflected by the longitudinal Z-series.
723
724 **Statistics**
725
726 Statistical analyses were caried out using GraphPad Prism 9 software. Differences between
727 conditions was estimated by one-way ANOVA complemented with Turkey's *post hoc* test (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$).
728
729
730 ## ACKNOWLEDGEMENTS
731
732 We thank David J Barry (Advanced Light Microscopy, The Francis Crick Institute) for help with
733 microscopy data analysis. We are grateful to Joe Brock (Communications, The Francis Crick
734 Institute) for helping design Fig 9. We acknowledge Jonathan Stoye for helpful discussions. For
735 the purpose of Open Access, the authors have applied a CC BY public copyright licence to any
736 Author Accepted Manuscript version arising from this submission.
737

FUNDING

This work was supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK (FC001042, FC001178), the UK Medical Research Council (FC001042, FC001178), and the Wellcome Trust (FC001042, FC001178). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

1. McDonald D, Vodicka MA, Lucero G, Svitkina TM, Borisy GG, Emerman M, et al. Visualization of the intracellular behavior of HIV in living cells. J Cell Biol. 2002;159(3):441-52.

2. Ganser BK, Li S, Klishko VY, Finch JT, Sundquist WI. Assembly and analysis of conical models for the HIV-1 core. Science. 1999;283(5398):80-3.

3. Ganser-Pornillos BK, von Schwedler UK, Stray KM, Aiken C, Sundquist WI. Assembly properties of the human immunodeficiency virus type 1 CA protein. J Virol. 2004;78(5):2545-52.

4. Pornillos O, Ganser-Pornillos BK, Yeager M. Atomic-level modelling of the HIV capsid. Nature. 2011;469(7330):424-7.

5. Mattei S, Glass B, Hagen WJ, Krausslich HG, Briggs JA. The structure and flexibility of conical HIV-1 capsids determined within intact virions. Science. 2016;354(6318):1434-7.

6. Du S, Betts L, Yang R, Shi H, Concel J, Ahn J, et al. Structure of the HIV-1 full-length capsid protein in a conformationally trapped unassembled state induced by small-molecule binding. J Mol Biol. 2011;406(3):371-86.

760 7. Deshmukh L, Schwieters CD, Grishaev A, Ghirlando R, Baber JL, Clore GM. Structure and
761 dynamics of full-length HIV-1 capsid protein in solution. J Am Chem Soc. 2013;135(43):16133-47.
762 8. Zhao G, Perilla JR, Yufenyuy EL, Meng X, Chen B, Ning J, et al. Mature HIV-1 capsid
763 structure by cryo-electron microscopy and all-atom molecular dynamics. Nature.
764 2013;497(7451):643-6.
765 9. Li W, Singh PK, Sowd GA, Bedwell GJ, Jang S, Achuthan V, et al. CPSF6-Dependent
766 Targeting of Speckle-Associated Domains Distinguishes Primate from Nonprimate Lentiviral
767 Integration. mBio. 2020;11(5).
768 10. Briggs JA, Wilk T, Welker R, Krausslich HG, Fuller SD. Structural organization of authentic,
769 mature HIV-1 virions and cores. EMBO J. 2003;22(7):1707-15.
770 11. Pornillos O, Ganser-Pornillos BK, Kelly BN, Hua Y, Whitby FG, Stout CD, et al. X-ray
771 structures of the hexameric building block of the HIV capsid. Cell. 2009;137(7):1282-92.
772 12. Pornillos O, Ganser-Pornillos BK, Banumathi S, Hua Y, Yeager M. Disulfide bond
773 stabilization of the hexameric capsomer of human immunodeficiency virus. J Mol Biol.
774 2010;401(5):985-95.
775 13. Ganser-Pornillos BK, Yeager M, Pornillos O. Assembly and architecture of HIV. Adv Exp
776 Med Biol. 2012;726:441-65.
777 14. Gres AT, Kirby KA, KewalRamani VN, Tanner JJ, Pornillos O, Sarafianos SG. STRUCTURAL
778 VIROLOGY. X-ray crystal structures of native HIV-1 capsid protein reveal conformational
779 variability. Science. 2015;349(6243):99-103.
780 15. Le Sage V, Mouland AJ, Valiente-Echeverria F. Roles of HIV-1 capsid in viral replication and
781 immune evasion. Virus Res. 2014;193:116-29.

782 16. Rasaiyaah J, Tan CP, Fletcher AJ, Price AJ, Blondeau C, Hilditch L, et al. HIV-1 evades innate
783 immune recognition through specific cofactor recruitment. Nature. 2013;503(7476):402-5.
784 17. Stremlau M, Owens CM, Perron MJ, Kiessling M, Autissier P, Sodroski J. The cytoplasmic
785 body component TRIM5alpha restricts HIV-1 infection in Old World monkeys. Nature.
786 2004;427(6977):848-53.
787 18. Malim MH, Bieniasz PD. HIV Restriction Factors and Mechanisms of Evasion. Cold Spring
788 Harb Perspect Med. 2012;2(5):a006940.
789 19. Franke EK, Yuan HE, Luban J. Specific incorporation of cyclophilin A into HIV-1 virions.
790 Nature. 1994;372(6504):359-62.
791 20. Bosco DA, Eisenmesser EZ, Pochapsky S, Sundquist WI, Kern D. Catalysis of cis/trans
792 isomerization in native HIV-1 capsid by human cyclophilin A. Proc Natl Acad Sci U S A.
793 2002;99(8):5247-52.
794 21. Gamble TR, Vajdos FF, Yoo S, Worthylake DK, Houseweart M, Sundquist WI, et al. Crystal
795 structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. Cell.
796 1996;87(7):1285-94.
797 22. Sabo Y, Walsh D, Barry DS, Tinaztepe S, de Los Santos K, Goff SP, et al. HIV-1 induces the
798 formation of stable microtubules to enhance early infection. Cell Host Microbe. 2013;14(5):535-
799 46.
800 23. Lukic Z, Dharan A, Fricke T, Diaz-Griffero F, Campbell EM. HIV-1 uncoating is facilitated by
801 dynein and kinesin 1. J Virol. 2014;88(23):13613-25.

802 24. Malikov V, da Silva ES, Jovasevic V, Bennett G, de Souza Aranha Vieira DA, Schulte B, et al.
803 HIV-1 capsids bind and exploit the kinesin-1 adaptor FEZ1 for inward movement to the nucleus.
804 Nat Commun. 2015;6:6660.
805 25. Brass AL, Dykxhoorn DM, Benita Y, Yan N, Engelman A, Xavier RJ, et al. Identification of
806 host proteins required for HIV infection through a functional genomic screen. Science.
807 2008;319(5865):921-6.
808 26. Konig R, Zhou Y, Elleder D, Diamond TL, Bonamy GM, Irelan JT, et al. Global analysis of
809 host-pathogen interactions that regulate early-stage HIV-1 replication. Cell. 2008;135(1):49-60.
810 27. Christ F, Thys W, De Rijck J, Gijsbers R, Albanese A, Arosio D, et al. Transportin-SR2 imports
811 HIV into the nucleus. Curr Biol. 2008;18(16):1192-202.
812 28. Fernandez J, Machado AK, Lyonnais S, Chamontin C, Gartner K, Leger T, et al. Transportin-
813 1 binds to the HIV-1 capsid via a nuclear localization signal and triggers uncoating. Nat Microbiol.
814 2019;4(11):1840-50.
815 29. Matreyek KA, Engelman A. The requirement for nucleoporin NUP153 during human
816 immunodeficiency virus type 1 infection is determined by the viral capsid. J Virol.
817 2011;85(15):7818-27.
818 30. Di Nunzio F, Danckaert A, Fricke T, Perez P, Fernandez J, Perret E, et al. Human
819 nucleoporins promote HIV-1 docking at the nuclear pore, nuclear import and integration. PLoS
820 One. 2012;7(9):e46037.
821 31. Price AJ, Jacques DA, McEwan WA, Fletcher AJ, Essig S, Chin JW, et al. Host cofactors and
822 pharmacologic ligands share an essential interface in HIV-1 capsid that is lost upon disassembly.
823 PLoS Pathog. 2014;10(10):e1004459.

824 32. Bhattacharya A, Alam SL, Fricke T, Zadrozny K, Sedzicki J, Taylor AB, et al. Structural basis
825 of HIV-1 capsid recognition by PF74 and CPSF6. Proc Natl Acad Sci U S A. 2014;111(52):18625-30.
826 33. Chin CR, Perreira JM, Savidis G, Portmann JM, Aker AM, Feeley EM, et al. Direct
827 Visualization of HIV-1 Replication Intermediates Shows that Capsid and CPSF6 Modulate HIV-1
828 Intra-nuclear Invasion and Integration. Cell Rep. 2015;13(8):1717-31.
829 34. Achuthan V, Perreira JM, Sowd GA, Puray-Chavez M, McDougall WM, Paulucci-
830 Holthauzen A, et al. Capsid-CPSF6 Interaction Licenses Nuclear HIV-1 Trafficking to Sites of Viral
831 DNA Integration. Cell Host Microbe. 2018;24(3):392-404 e8.
832 35. Francis AC, Marin M, Singh PK, Achuthan V, Prellberg MJ, Palermino-Rowland K, et al. HIV-
833 1 replication complexes accumulate in nuclear speckles and integrate into speckle-associated
834 genomic domains. Nat Commun. 2020;11(1):3505.
835 36. Miller MD, Farnet CM, Bushman FD. Human immunodeficiency virus type 1 preintegration
836 complexes: studies of organization and composition. J Virol. 1997;71(7):5382-90.
837 37. Fassati A, Goff SP. Characterization of intracellular reverse transcription complexes of
838 human immunodeficiency virus type 1. J Virol. 2001;75(8):3626-35.
839 38. Mamede JI, Cianci GC, Anderson MR, Hope TJ. Early cytoplasmic uncoating is associated
840 with infectivity of HIV-1. Proc Natl Acad Sci U S A. 2017;114(34):E7169-E78.
841 39. Francis AC, Melikyan GB. Single HIV-1 Imaging Reveals Progression of Infection through
842 CA-Dependent Steps of Docking at the Nuclear Pore, Uncoating, and Nuclear Transport. Cell Host
843 Microbe. 2018;23(4):536-48 e6.

844 40. Bejarano DA, Peng K, Laketa V, Borner K, Jost KL, Lucic B, et al. HIV-1 nuclear import in
845 macrophages is regulated by CPSF6-capsid interactions at the nuclear pore complex. Elife.
846 2019;8.
847 41. Zurnic Bonisch I, Dirix L, Lemmens V, Borrenberghs D, De Wit F, Vernaillen F, et al. Capsid-
848 Labelled HIV To Investigate the Role of Capsid during Nuclear Import and Integration. J Virol.
849 2020;94(7).
850 42. Blanco-Rodriguez G, Gazi A, Monel B, Frabetti S, Scoca V, Mueller F, et al. Remodeling of
851 the Core Leads HIV-1 Preintegration Complex into the Nucleus of Human Lymphocytes. J Virol.
852 2020;94(11).
853 43. Burdick RC, Li C, Munshi M, Rawson JMO, Nagashima K, Hu WS, et al. HIV-1 uncoats in the
854 nucleus near sites of integration. Proc Natl Acad Sci U S A. 2020;117(10):5486-93.
855 44. Dharan A, Bachmann N, Talley S, Zwikelmaier V, Campbell EM. Nuclear pore blockade
856 reveals that HIV-1 completes reverse transcription and uncoating in the nucleus. Nat Microbiol.
857 2020;5(9):1088-95.
858 45. Schaller T, Ocwieja KE, Rasaiyaah J, Price AJ, Brady TL, Roth SL, et al. HIV-1 capsid-
859 cyclophilin interactions determine nuclear import pathway, integration targeting and replication
860 efficiency. PLoS Pathog. 2011;7(12):e1002439.
861 46. Hulme AE, Kelley Z, Foley D, Hope TJ. Complementary Assays Reveal a Low Level of CA
862 Associated with Viral Complexes in the Nuclei of HIV-1-Infected Cells. J Virol. 2015;89(10):5350-
863 61.

47. Burdick RC, Delviks-Frankenberry KA, Chen J, Janaka SK, Sastri J, Hu WS, et al. Dynamics and regulation of nuclear import and nuclear movements of HIV-1 complexes. PLoS Pathog. 2017;13(8):e1006570.

48. Selyutina A, Persaud M, Lee K, KewalRamani V, Diaz-Griffero F. Nuclear Import of the HIV-1 Core Precedes Reverse Transcription and Uncoating. Cell Rep. 2020;32(13):108201.

49. Hulme AE, Perez O, Hope TJ. Complementary assays reveal a relationship between HIV-1 uncoating and reverse transcription. Proc Natl Acad Sci U S A. 2011;108(24):9975-80.

50. Yang Y, Fricke T, Diaz-Griffero F. Inhibition of reverse transcriptase activity increases stability of the HIV-1 core. J Virol. 2013;87(1):683-7.

51. Cosnefroy O, Murray PJ, Bishop KN. HIV-1 capsid uncoating initiates after the first strand transfer of reverse transcription. Retrovirology. 2016;13(1):58.

52. Forshey BM, von Schwedler U, Sundquist WI, Aiken C. Formation of a human immunodeficiency virus type 1 core of optimal stability is crucial for viral replication. J Virol. 2002;76(11):5667-77.

53. von Schwedler UK, Stray KM, Garrus JE, Sundquist WI. Functional surfaces of the human immunodeficiency virus type 1 capsid protein. J Virol. 2003;77(9):5439-50.

54. Link JO, Rhee MS, Tse WC, Zheng J, Somoza JR, Rowe W, et al. Clinical targeting of HIV capsid protein with a long-acting small molecule. Nature. 2020;584(7822):614-8.

55. Bester SM, Wei G, Zhao H, Adu-Ampratwum D, Iqbal N, Courouble VV, et al. Structural and mechanistic bases for a potent HIV-1 capsid inhibitor. Science. 2020;370(6514):360-4.

56. Srinivasan N, Sowdhamini R, Ramakrishnan C, Balaram P. Conformations of disulfide bridges in proteins. Int J Pept Protein Res. 1990;36(2):147-55.

886 57. Yufenyuy EL, Aiken C. The NTD-CTD intersubunit interface plays a critical role in assembly
887 and stabilization of the HIV-1 capsid. Retrovirology. 2013;10:29.
888 58. Yang Y, Luban J, Diaz-Griffero F. The fate of HIV-1 capsid: a biochemical assay for HIV-1
889 uncoating. Methods Mol Biol. 2014;1087:29-36.
890 59. Rietsch A, Beckwith J. The genetics of disulfide bond metabolism. Annu Rev Genet.
891 1998;32:163-84.
892 60. Sevier CS, Kaiser CA. Formation and transfer of disulphide bonds in living cells. Nat Rev
893 Mol Cell Biol. 2002;3(11):836-47.
894 61. Rankovic S, Varadarajan J, Ramalho R, Aiken C, Rousso I. Reverse Transcription
895 Mechanically Initiates HIV-1 Capsid Disassembly. J Virol. 2017;91(12).
896 62. Fei J, Jadaliha M, Harmon TS, Li ITS, Hua B, Hao Q, et al. Quantitative analysis of multilayer
897 organization of proteins and RNA in nuclear speckles at super resolution. J Cell Sci.
898 2017;130(24):4180-92.
899 63. Campbell EM, Hope TJ. HIV-1 capsid: the multifaceted key player in HIV-1 infection. Nat
900 Rev Microbiol. 2015;13(8):471-83.
901 64. Zila V, Muller TG, Laketa V, Muller B, Krausslich HG. Analysis of CA Content and CPSF6
902 Dependence of Early HIV-1 Replication Complexes in SupT1-R5 Cells. mBio. 2019;10(6).
903 65. Zila V, Margiotta E, Turonova B, Muller TG, Zimmerli CE, Mattei S, et al. Cone-shaped HIV-1
904 capsids are transported through intact nuclear pores. Cell. 2021.
905 66. Byeon IJ, Meng X, Jung J, Zhao G, Yang R, Ahn J, et al. Structural convergence between
906 Cryo-EM and NMR reveals intersubunit interactions critical for HIV-1 capsid function. Cell.
907 2009;139(4):780-90.

908 67. Yeager M. Design of in vitro symmetric complexes and analysis by hybrid methods reveal
909 mechanisms of HIV capsid assembly. J Mol Biol. 2011;410(4):534-52.
910 68. Zhao G, Ke D, Vu T, Ahn J, Shah VB, Yang R, et al. Rhesus TRIM5alpha disrupts the HIV-1
911 capsid at the inter-hexamer interfaces. PLoS Pathog. 2011;7(3):e1002009.
912 69. Ramalho R, Rankovic S, Zhou J, Aiken C, Rousso I. Analysis of the mechanical properties of
913 wild type and hyperstable mutants of the HIV-1 capsid. Retrovirology. 2016;13:17.
914 70. Francis AC, Marin M, Shi J, Aiken C, Melikyan GB. Time-Resolved Imaging of Single HIV-1
915 Uncoating In Vitro and in Living Cells. PLoS Pathog. 2016;12(6):e1005709.
916 71. Shi J, Zhou J, Shah VB, Aiken C, Whitby K. Small-molecule inhibition of human
917 immunodeficiency virus type 1 infection by virus capsid destabilization. J Virol. 2011;85(1):542-
918 9.
919 72. Yang R, Shi J, Byeon IJ, Ahn J, Sheehan JH, Meiler J, et al. Second-site suppressors of HIV-
920 1 capsid mutations: restoration of intracellular activities without correction of intrinsic capsid
921 stability defects. Retrovirology. 2012;9:30.
922 73. Rankovic S, Deshpande A, Harel S, Aiken C, Rousso I. HIV-1 uncoating occurs via a series
923 of rapid biomechanical changes in the core related to individual stages of reverse transcription.
924 bioRxiv. 2021.
925 74. Rankovic S, Ramalho R, Aiken C, Rousso I. PF74 Reinforces the HIV-1 Capsid To Impair
926 Reverse Transcription-Induced Uncoating. J Virol. 2018;92(20).
927 75. Christensen DE, Ganser-Pornillos BK, Johnson JS, Pornillos O, Sundquist WI.
928 Reconstitution and visualization of HIV-1 capsid-dependent replication and integration in vitro.
929 Science. 2020;370(6513).

930 76. Francis AC, Marin M, Prellberg MJ, Palermino-Rowland K, Melikyan GB. HIV-1 Uncoating
931 and Nuclear Import Precede the Completion of Reverse Transcription in Cell Lines and in Primary
932 Macrophages. Viruses. 2020;12(11).
933 77. Rensen E, Mueller F, Scoca V, Parmar JJ, Souque P, Zimmer C, et al. Clustering and reverse
934 transcription of HIV-1 genomes in nuclear niches of macrophages. EMBO J. 2021;40(1):e105247.
935 78. Sharkey ME, Teo I, Greenough T, Sharova N, Luzuriaga K, Sullivan JL, et al. Persistence of
936 episomal HIV-1 infection intermediates in patients on highly active anti-retroviral therapy. Nat
937 Med. 2000;6(1):76-81.
938 79. Butler SL, Hansen MS, Bushman FD. A quantitative assay for HIV DNA integration in vivo.
939 Nat Med. 2001;7(5):631-4.
940 80. Kuhn TM, Capelson M. Nuclear Pore Proteins in Regulation of Chromatin State. Cells.
941 2019;8(11).
942 81. Dharan A, Talley S, Tripathi A, Mamede JI, Majetschak M, Hope TJ, et al. KIF5B and Nup358
943 Cooperatively Mediate the Nuclear Import of HIV-1 during Infection. PLoS Pathog.
944 2016;12(6):e1005700.
945 82. Price AJ, Fletcher AJ, Schaller T, Elliott T, Lee K, KewalRamani VN, et al. CPSF6 defines a
946 conserved capsid interface that modulates HIV-1 replication. PLoS Pathog. 2012;8(8):e1002896.
947 83. Rasheedi S, Shun MC, Serrao E, Sowd GA, Qian J, Hao C, et al. The Cleavage and
948 Polyadenylation Specificity Factor 6 (CPSF6) Subunit of the Capsid-recruited Pre-messenger RNA
949 Cleavage Factor I (CFIm) Complex Mediates HIV-1 Integration into Genes. J Biol Chem.
950 2016;291(22):11809-19.

951 84. Wanaguru M, Barry DJ, Benton DJ, O'Reilly NJ, Bishop KN. Murine leukemia virus p12
952 tethers the capsid-containing pre-integration complex to chromatin by binding directly to host
953 nucleosomes in mitosis. PLoS Pathog. 2018;14(6):e1007117.
954 85. Miller DG, Adam MA, Miller AD. Gene transfer by retrovirus vectors occurs only in cells
955 that are actively replicating at the time of infection. Mol Cell Biol. 1990;10(8):4239-42.
956 86. Zufferey R, Nagy D, Mandel RJ, Naldini L, Trono D. Multiply attenuated lentiviral vector
957 achieves efficient gene delivery in vivo. Nat Biotechnol. 1997;15(9):871-5.
958 87. Bainbridge JW, Stephens C, Parsley K, Demaison C, Halfyard A, Thrasher AJ, et al. In vivo
959 gene transfer to the mouse eye using an HIV-based lentiviral vector; efficient long-term
960 transduction of corneal endothelium and retinal pigment epithelium. Gene Ther.
961 2001;8(21):1665-8.
962 88. Barry DJ, Gerri C, Bell DM, D'Antuono R, Niakan KK. GIANI: open-source software for
963 automated analysis of 3D microscopy images. bioRxiv. 2020.
964

965 ## FIGURE CAPTIONS

966

967 **Figure 1. Panel of CA mutants.** (A) The panel of CA mutants used in this study and the lattice
968 interface at which the mutations reside. New CA mutants with cysteine mutations (E180C,
969 L151C/L189C and K203C/A217C) were designed based on a previously published cryo-EM MDFF
970 atomic model (PDB ID: 3J34; Zhao et al, 2013) and crystal structures (PDB ID: 3H4E and 3H47;
971 Pornillos et al, 2009). The remaining CA mutants were selected from previous publications as
972 indicated in the references column. (B) Structure of the CA lattice from PDB ID: 3J34, showing 18

973 CA monomers arranged into three hexameric rings. Residues where mutations were made are
974 highlighted and colour-coded according to the lattice interface type, as in (A).
975
976 **Figure 2. Effect of the CA mutations on viral core stability in cells.** (A, B) Representative
977 immunoblots of Fate-of-capsid assays comparing WT VLP to mutant VLP. HeLa cells were infected
978 with equal RT units of WT or mutant VLP and cell lysates were harvested at 2 hpi (A, B) and 20
979 hpi (A). Cell lysates were centrifuged through a sucrose cushion to separate viral CA into free
980 (soluble, S) and assembled (pellet, P) fractions. An input (I) sample was also harvested before the
981 centrifugation through the sucrose cushion. CA was detected by immunoblotting using an anti-
982 HIV-1 CA antibody. Mutant VLP with a CA ratio of P>S at 2h and P=S at 20h were considered
983 "hyper-stable" (surrounded with a dashed-line box). Each assay was performed at least three
984 times independently.
985
986 **Figure 3. Effect of CA mutations on VLP infectivity.** 293T (A), HeLa (B), SupT1 (C) and U937 (D)
987 cells were infected with equal RT units of GFP-reporter WT or mutant VLP. The percentage of
988 GFP+ cells was measured by flow cytometry at 72hpi and plotted relative to WT VLP. Points
989 indicate individual biological repeats and lines show the mean ± SEM. Bars are colour coded
990 according to the lattice interface at which the cysteines have been introduced, as in Fig 1. Hyper-
991 stable mutants based on the fate of capsid assay are indicated with black arrow heads.
992
993 **Figure 4. Effect of CA mutations on reverse transcription.** 293T cells were synchronously infected
994 with equivalent RT units of WT or mutant VLP. At the indicated times post-infection, cells were

harvested for DNA extraction, and viral cDNA products were measured by qPCR. (A-D) Graphs show the levels of early (strong stop) cDNA (upper panels) and late (second strand) cDNA (lower panels) in 293T cells following infection with WT VLP (Black line) and mutants P38A (A), P207C/T216C (B), A14C/E45C (C) or M68C/E212C (D). The data is shown as mean ± SEM from at least two independent experiments. (E, F) Bar charts show the levels of strong stop cDNA at 6h (E) and 24h (F) post infection for the panel of mutants relative to WT infections. (G) Bar chart showing the levels of strong stop cDNA at 24 h (left y-axis) and infectivity (from Fig 3) at 72 h (right y-axis) compared to WT VLP for each mutant. Individual points represent biological repeats and lines indicate the mean +/- SEM. (H) Bar chart showing the ratio of relative levels of strong stop cDNA to infectivity, from (G). Dashed line indicates a ratio of 3. Bars are colour coded according to the lattice interface at which the cysteines have been introduced, as in Fig 1. Hyper-stable mutants are indicated with a black arrow heads.

Figure 5. Effect of hyper-stable CA mutations on the early stages of infection. (A) Disulphide cross-linking of CA monomers in cells. HeLa cells were infected with WT VLPs or the hyper-stable mutants M68C/E212C, A14C/E45C or E180C, and cell lysates were analysed by non-reducing SDS-PAGE and immunoblotting with an HIV-1 CA antibody. Samples were either treated with 10µM-50µM (indicated by gradient arrows in first panel) or 50µM (indicated by "+") Iodoacetamide, to prevent further disulphide bond formation, or with 0.1M DTT (indicated by "+"), to reduce existing disulphide bonds, prior to SDS-PAGE. The prominent bands at 24 and 48 kDa correspond to monomeric and dimeric CA, respectively. (B) 293T cells were synchronously infected with equivalent RT units of WT or mutant VLP. At the indicated times post-infection, cells were

harvested for DNA extraction, and early and late cDNA (6hpi), 2-LTR circles (24hpi) and integrated proviral DNA (2 weeks post infection) were measured by qPCR. Data were plotted relative to WT infections (shown as a dashed line at 100%). Points indicate individual biological repeats and lines show the mean ± SEM. Viral infectivity from parallel infections (see Fig 3) was plotted for comparison.

Figure 6. Cellular localisation of CA and IN proteins during WT and A14C/E45C infections. HeLa cells were synchronously infected with equal RT units of WT or A14C/E45C mutant (CC in the figure) VLP. At 4, 8, 24 and 30 h post-infection, cells were harvested in parallel to be processed as a whole cell lysate or to undergo subcellular fractionation. Protein levels were quantified by BCA assay, proportional amounts of the fractions related to the WCL were loaded on SDS-PAGE gels and analysed by immunoblotting using the following antibodies: Anti-CA and anti-IN for HIV-1 proteins, HSP90 for cytoplasm, Calnexin for membranes, HDAC2 for nucleus, Histone 3 for chromatin and Tubulin α for cytoskeleton. Lamin B1 was used as a nuclear envelope marker. "-" indicates uninfected cells. Panels show representative immunoblots of: (A) WCL probed for HIV-1 CA and IN. HSP90 was blotted as a loading control. (B) Subcellular fractions from uninfected cells, probed for control proteins to confirm successful sample fractionation. (C-E) Subcellular fractions probed for HIV-1 CA and IN and the appropriate cellular marker: (C) cytoplasm, (D) membranes, (E) nucleus and (F) chromatin fraction.

Figure 7. CA co-localisation of CA with Nup358 and Nup153 during infection. HeLa cells were synchronously infected with equal RT units of WT or A14C/E45C VLP and fixed at 2, 4, 6, 8 and 10

hpi. Cells were incubated with primary antibodies to HIV-1 CA and either Nup358 (A-C) or Nup153 (D-F), followed by specific secondary antibodies conjugated to the PLUS and MINUS PLA oligonucleotides (probes). Each foci represents a positive PLA signal generated by the amplication of the interaction between the PLUS and MINUS probes. (A) and (D) show representative images of CA-Nup358 and CA-Nup153 co-localisation at 8 hpi, respectively (the scale bar is 10µm). Longitudinal Z-series were acquired with a 63X objective using a confocal microscope followed by 3D image analysis performed with the GIANI plug-in in FIJI. The number of foci per cell (B and E) and the distance of those foci to the DAPI edge (C and F) were quantified. In the plots, each point represents the mean foci data from all the cells within an image (75 to 100 cells/image). At least 12 images were acquired per condition over at least three independent biological repeats (plotted in different colours). Overall means ± SEM are shown in black. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001, ns = not significant.

Figure 8. HIV-1 CA, CPSF6 and SC35 staining during infection with WT and hyper-stable mutant VLP. HeLa cells were synchronously infected with equal RT units of WT, A14C/E45C, E180C or M68C/E212C VLP and fixed at 16hpi. Cells were incubated with primary antibodies against HIV-1 CA, CPSF6 and SC35 followed by specific secondary antibodies conjugated to Alexa Fluor fluorophores. (A) Representative images of SC35 and CPSF6 staining. White arrows point to SC35 and CPSF6 co-localisation in the merged WT image. (B) Representative images of HIV-1 CA and CPSF6 staining (left panels). White boxes in WT columns indicate zoomed-in regions labelled 1 and 2 (right panels). The scale bar in (A) and (B) represents 10µm. Bar chart shows the number of CA positive cells that showed CPSF6 redistribution to puncta. Points indicate individual

biological repeats (~150 cells per experiment were used for quantification) and lines show the mean ± SEM; ****p<0.0001.

Figure 9. Current models for HIV-1 and MLV uncoating. Incoming viral cores (HIV-1, left columns and MLV, right column) surrounded by a CA lattice (orange hexamers) contain the viral RNA (red) coated with nucleocapsid (light blue), as well as the viral enzymes, protease (yellow), reverse transcriptase (light green) and integrase (purple) (step 1). Reverse transcription of the RNA to double stranded DNA (dark blue) starts following infection and the core travels towards the nucleus down microtubules (step 2). HIV-1 DNA can cross the nuclear pore (step 3), and in association with CPSF6 (green dots) (step 4), move to nuclear speckles (darker pink shaded regions) and integrate into the host cell chromatin (grey with orange ovals) (step 5) to form a provirus (step 6). When and where the CA lattice disassembles and where reverse transcription finishes is still debated and the possible scenarios for HIV-1 uncoating (loss of orange hexagons) are illustrated here: from left to right; uncoating in the cytoplasm, uncoating at the nuclear pore, uncoating inside the nucleus or uncoating at the site of integration. In contrast, MLV cannot pass through nuclear pores and must wait for mitosis to occur before accessing the chromatin. Our current model for MLV uncoating is shown on the right: At mitosis, the N-terminal domain of the MLV p12 protein (red) binds directly to CA and the C-terminal domain (green) binds to nucleosomes, tethering the likely intact MLV core to the host chromatin. Following mitosis, MLV uncoats to allow integration, which, as it is already associated with chromatin, probably occurs at the integration site.

SUPPORTING INFORMATION

Supplementary table 1. Primers used to make CA mutations by site-directed mutagenesis. Table listing forward and reverse primers used to introduce cysteine mutations on CA by site-directed mutagenesis.

Supplementary figure 1. VLP production and Gag expression. (A) GFP-reporter gene-expressing HIV-1 WT and mutant VLP were produced by transient transfection of 293T cells and the VLP titres in the cell supernatants were calculated by measuring RT activity using a modified RT ELISA. The bar chart shows the RT activity of the mutants relative to WT. Points indicate individual biological repeats and lines show the mean ± SEM. Colour coding is as in Fig 1. (B) Immunoblot of transfected 293T producer cell lysates probed with an anti-HIV-CA antibody showing expression of WT and mutant Gag proteins from CA mutants A14C/E45C, W184A/M185A and A14C/E45C/W184A/M185A. The blot was imaged using a LiCor Odyssey CLx imager. (C) Immunoblot of transfected 293T producer cell lysates probed with anti-HIV-CA antibody showing expression of WT and mutant Gag proteins from CA mutants A42C/T54C, Q63C/Y169C, E180C, V181C, L151C/L169C and K203C/A217C. The blot was imaged by exposure to X-ray film.

Supplementary figure 2. Effect of CA mutations on late reverse transcription. 293T cells were synchronously infected with equivalent RT units of WT or mutant VLP. Cells were harvested and DNA extracted and analysed for viral late cDNA products (second strand) by qPCR. (A) Bar chart shows the levels of second strand cDNA at 6 h post infection relative to WT infection. (B) Bar

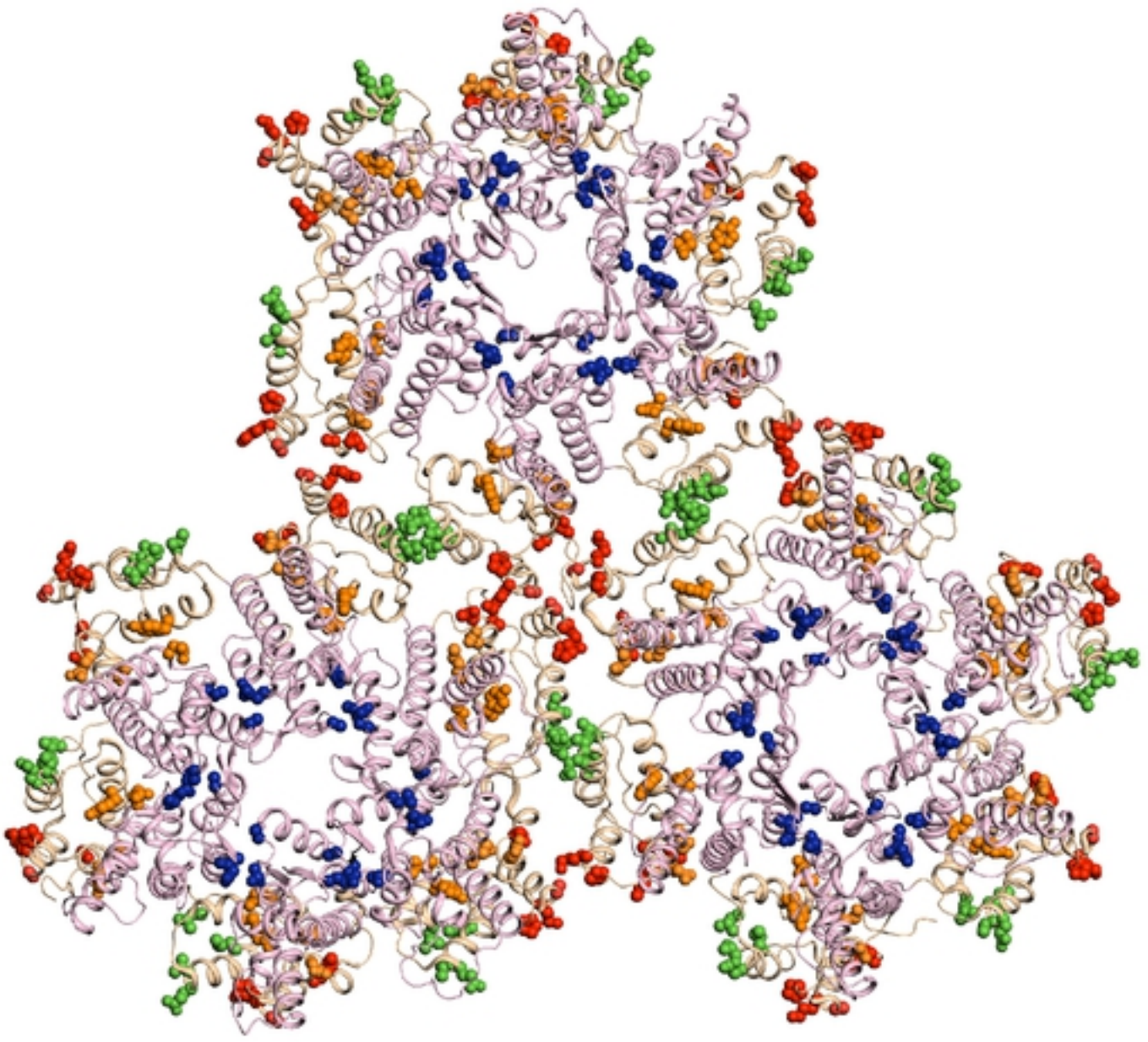
chart shows the levels of second strand cDNA at 24 h (left y-axis) and infectivity at 72 h (right y-axis) compared to WT VLP for each mutant. Individual points represent biological repeats and lines indicate the mean +/- SEM. (H) Bar chart shows the ratio of relative levels of second strand cDNA to infectivity, from (B). Dashed line indicates a ratio of 3. Bars are colour coded according to the lattice interface at which the cysteines have been introduced, as in fig 1. Hyper-stable mutants are indicated with black arrow heads.

Supplementary figure 3. CPSF6 and SC35 staining during infection. HeLa cells were synchronously infected with equal RT units of WT, A14C/E45C, E180C or M68C/E212C VLP and fixed at 16hpi. Cells were incubated with primary antibodies against CPSF6 and SC35 followed by specific secondary antibodies conjugated to Alexa Fluor fluorophores. The figure shows representative images of CPSF6 and SC35 staining at a lower magnification (63X) than in Fig 8A. The scale bar represents 20 µm.

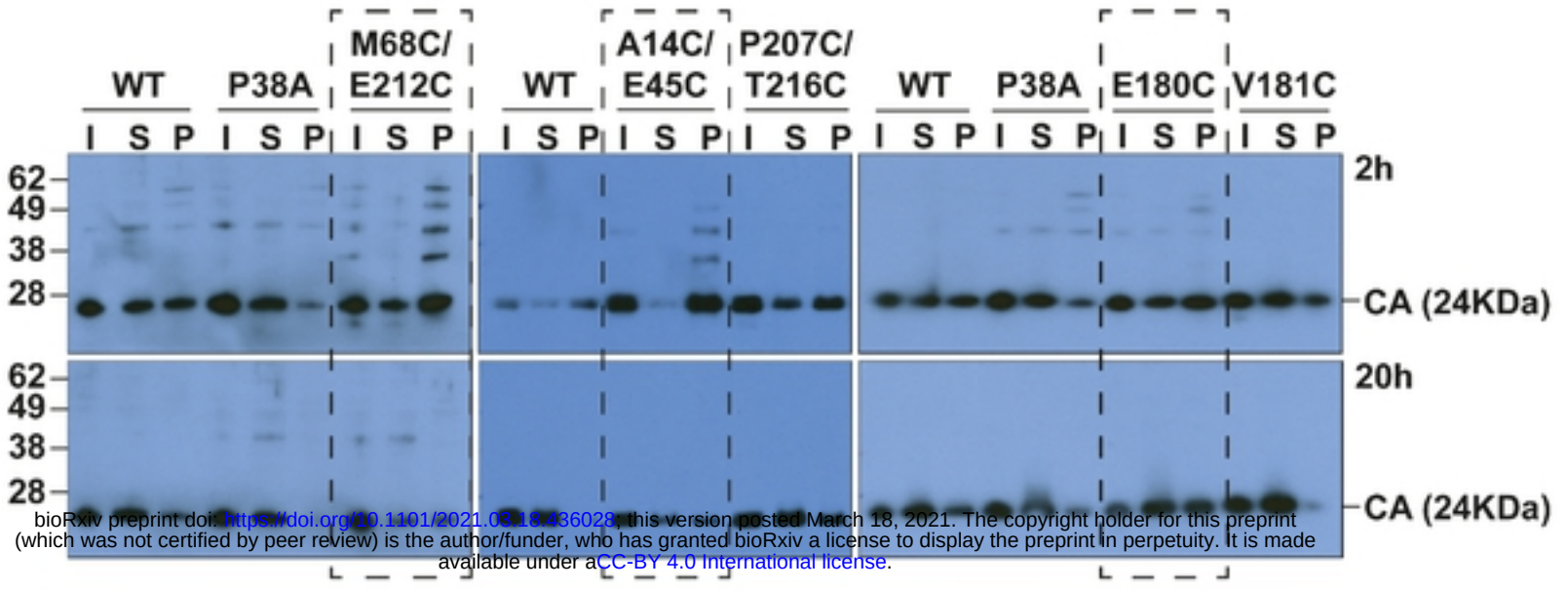
A

LATTICE INTERFACE		CA MUTATION	REFERENCES
INTRA-HEXAMER	NTD	P38A	Forshey *et al*, 2002; Cosnefroy *et al*, 2016
	NTD-NTD	A14C/E45C	Pornillos *et al*, 2009; Pornillos *et al*, 2010
		A42C/T54C	Pornillos *et al*, 2010; Yeager, 2011
	NTD-CTD	M68C/E212C	Yufenyuy & Aiken, 2013
		Q63C/Y169C	Yufenyuy & Aiken, 2013
INTER-HEXAMER	DIMERIC CTD-CTD	E180C	New CA mutant
		T181C	Pornillos *et al*, 2009
		L151C/L189C	New CA mutant
	TRIMERIC CTD-CTD	A204C	Zhao *et al*, 2013; Ramalho *et al*, 2016
		P207C/T216C	Byeon *et al*, 2009; Zhao *et al*, 2011
		K203C/A217C	New CA mutant
INTRA-HEXAMER + INTER-HEXAMER	NTD-NTD + DIMERIC CTD-CTD	A14C/E45C + W184A/M185A	Pornillos *et al*, 2009; Pornillos *et al*, 2010

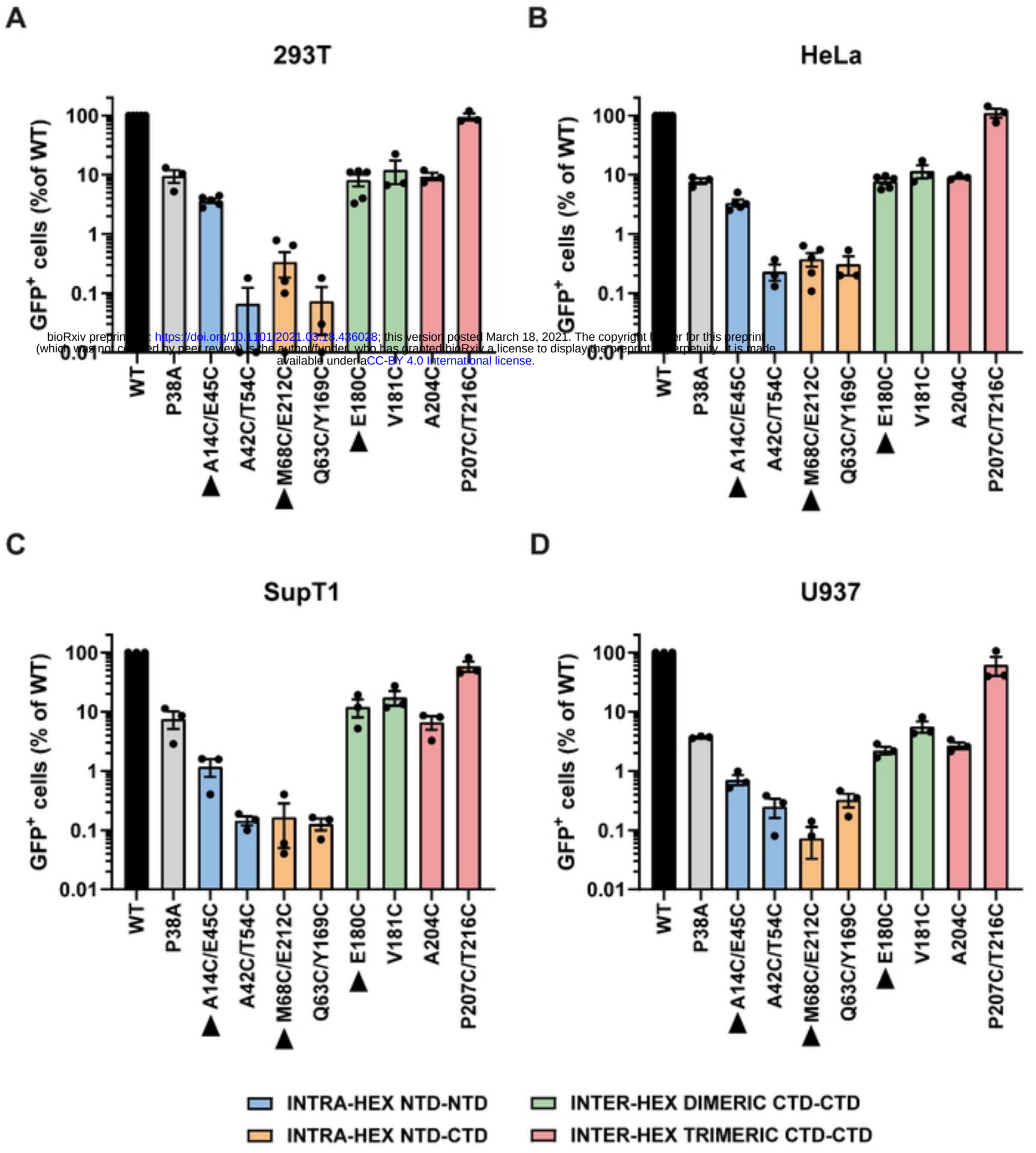
B

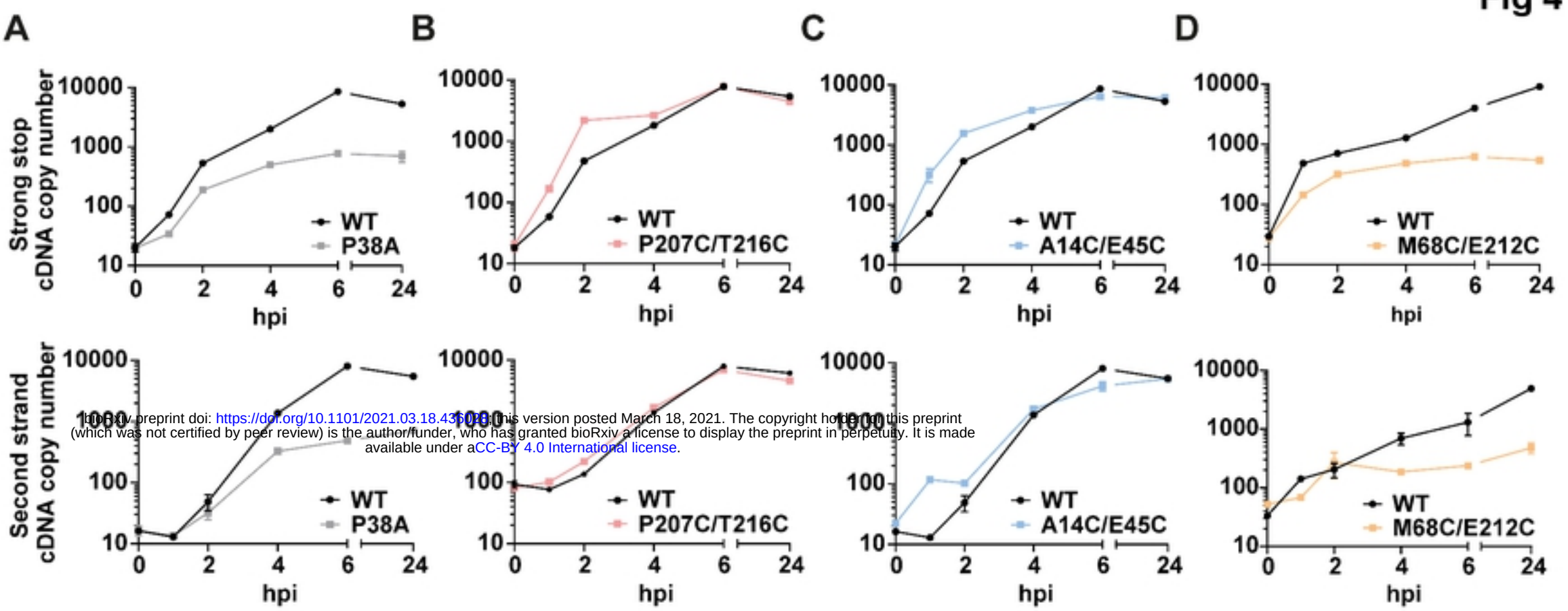


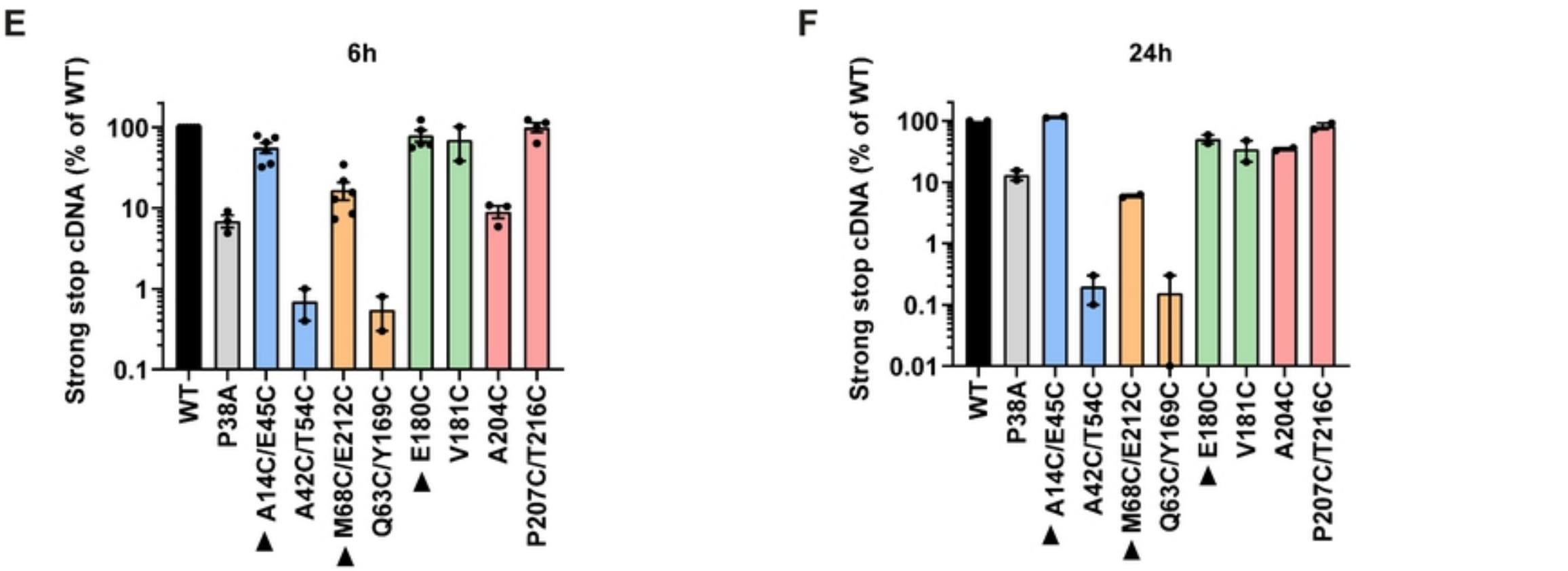
A

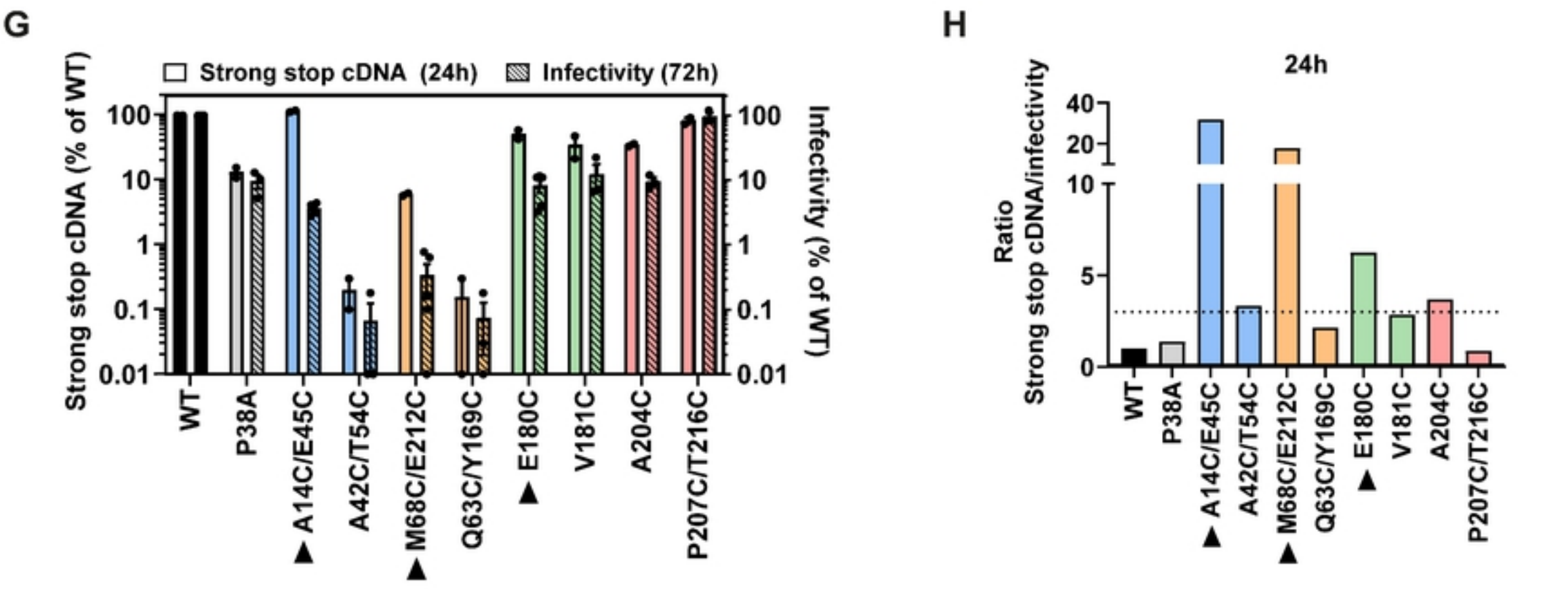


B

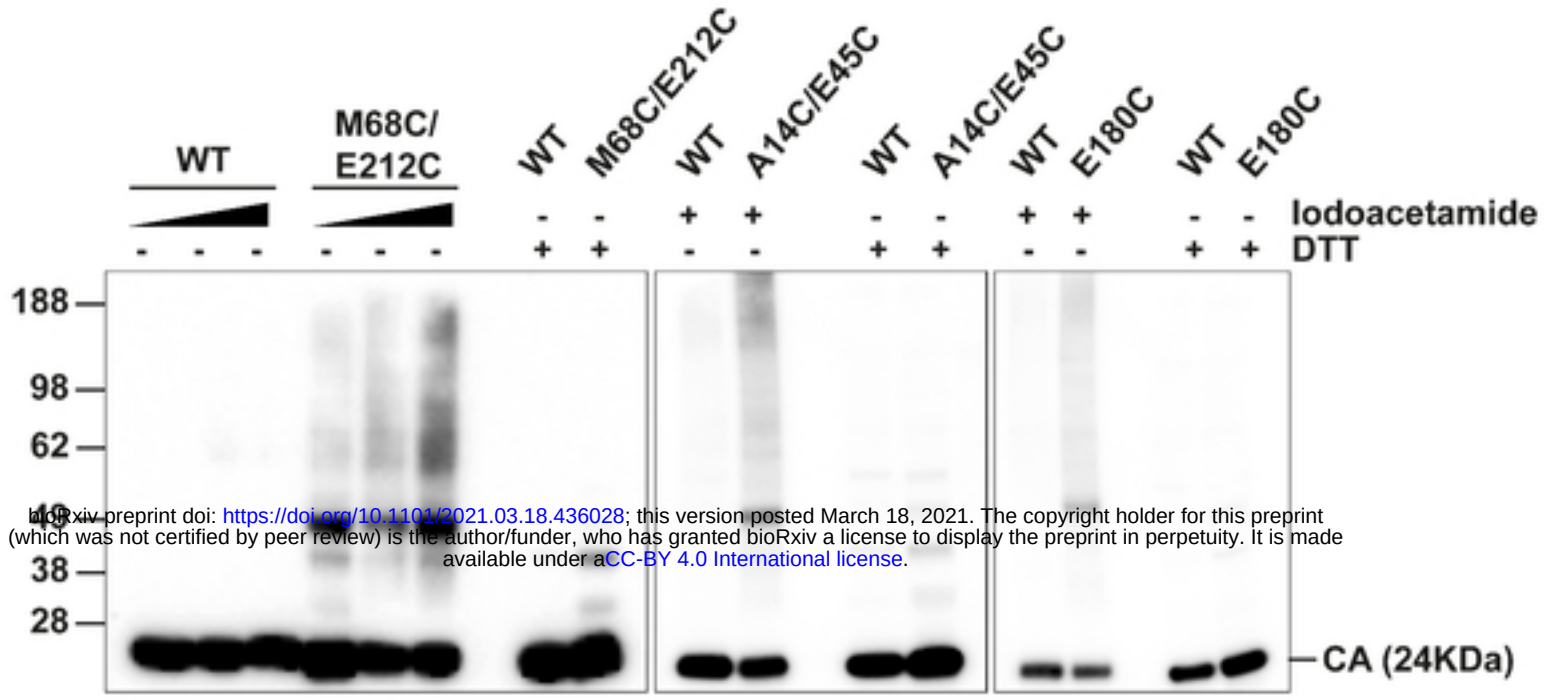








A



B

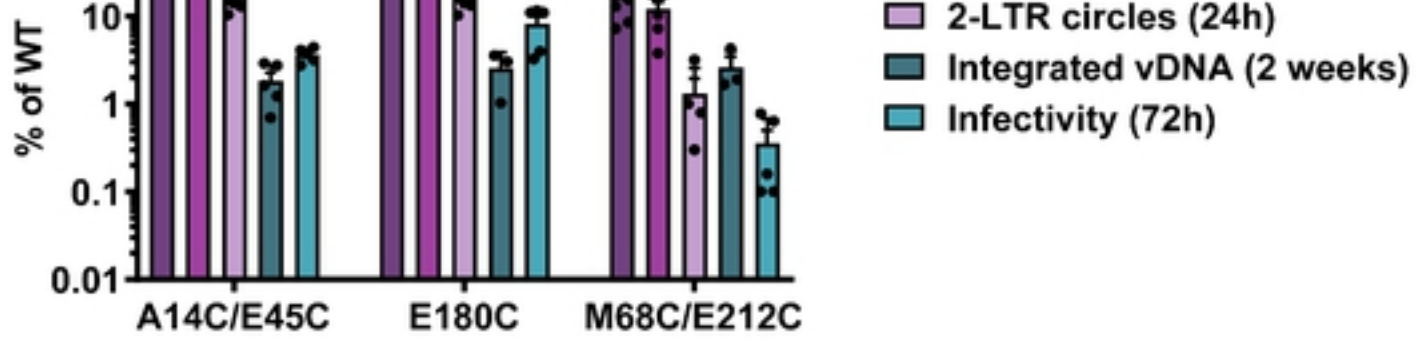
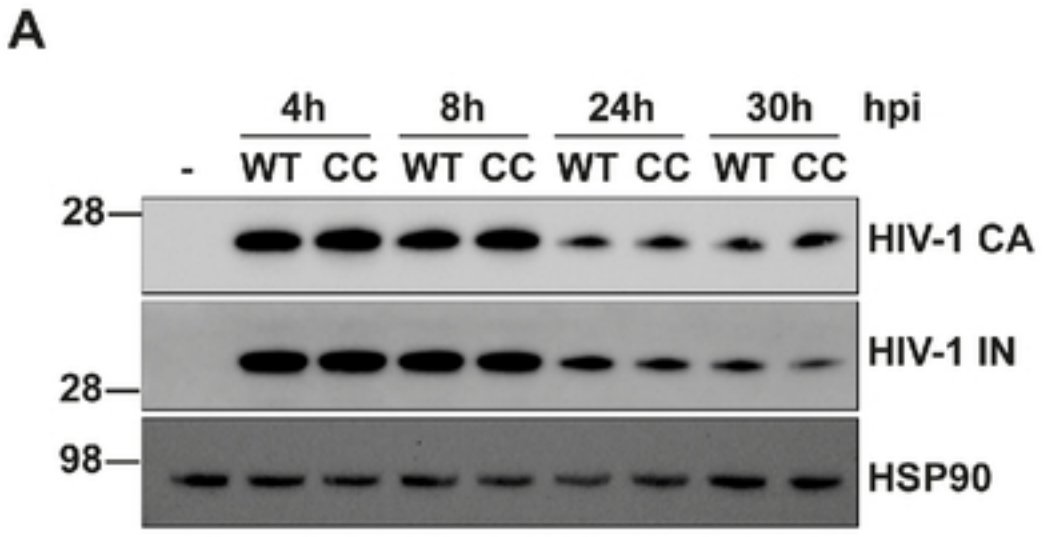
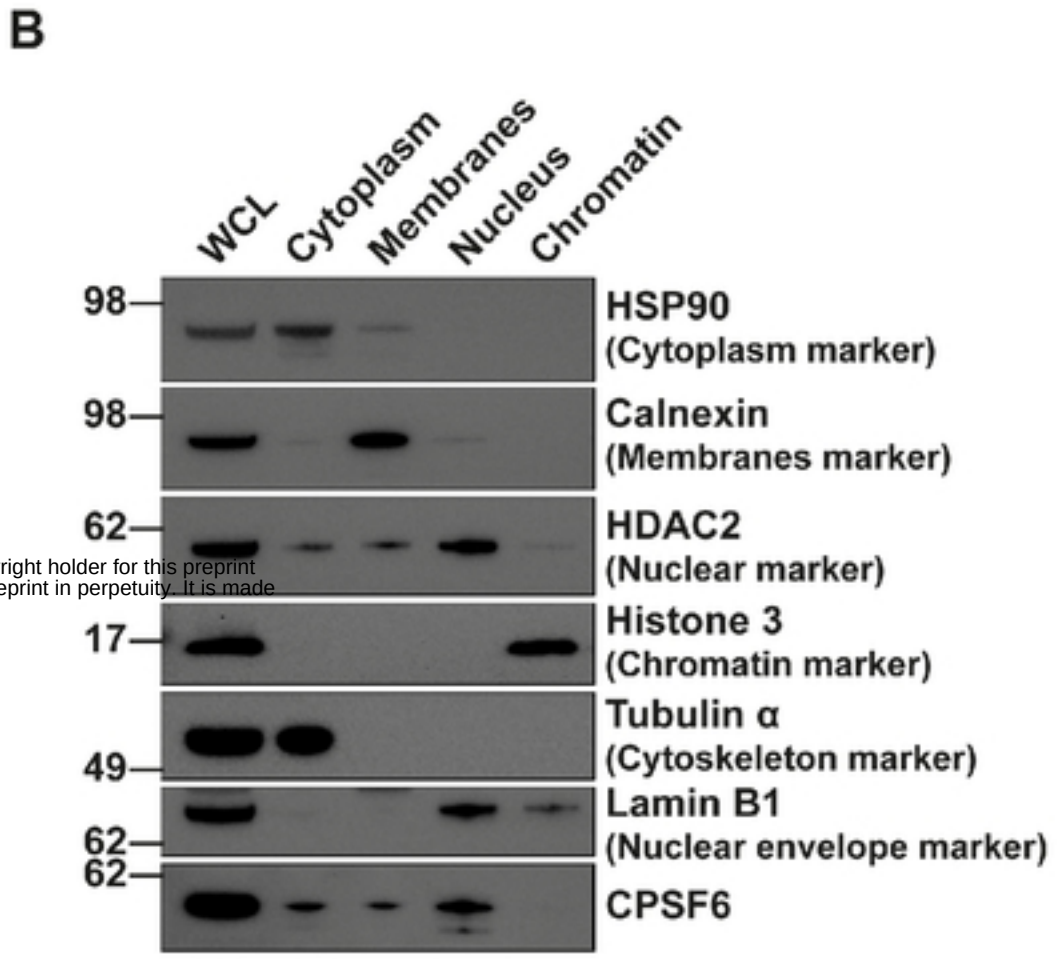
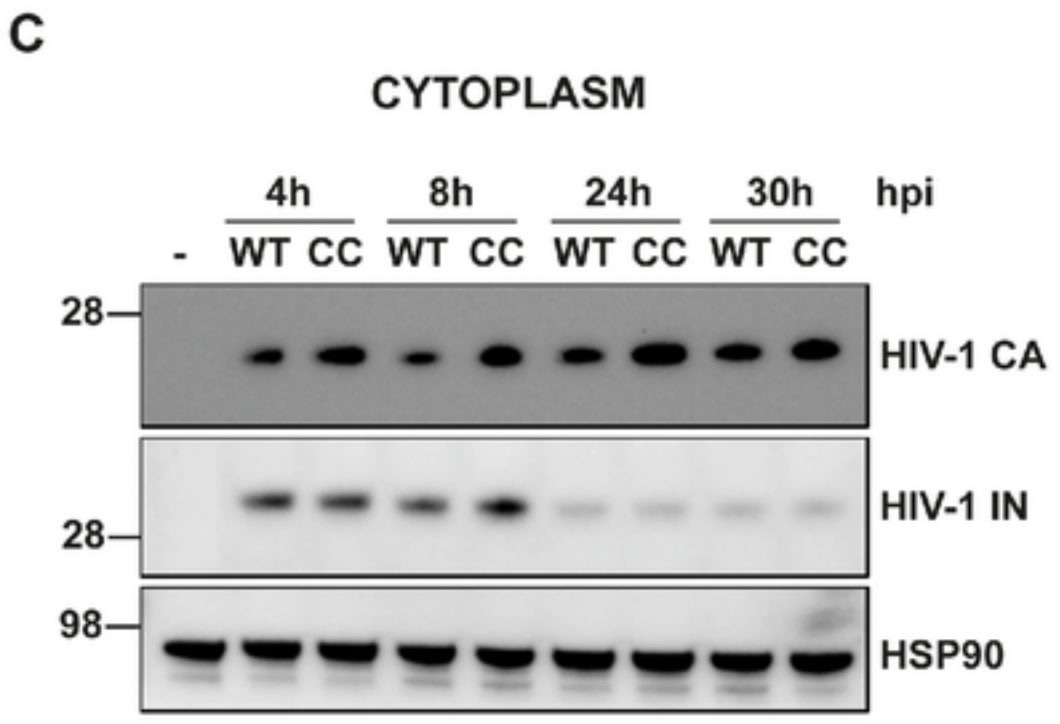
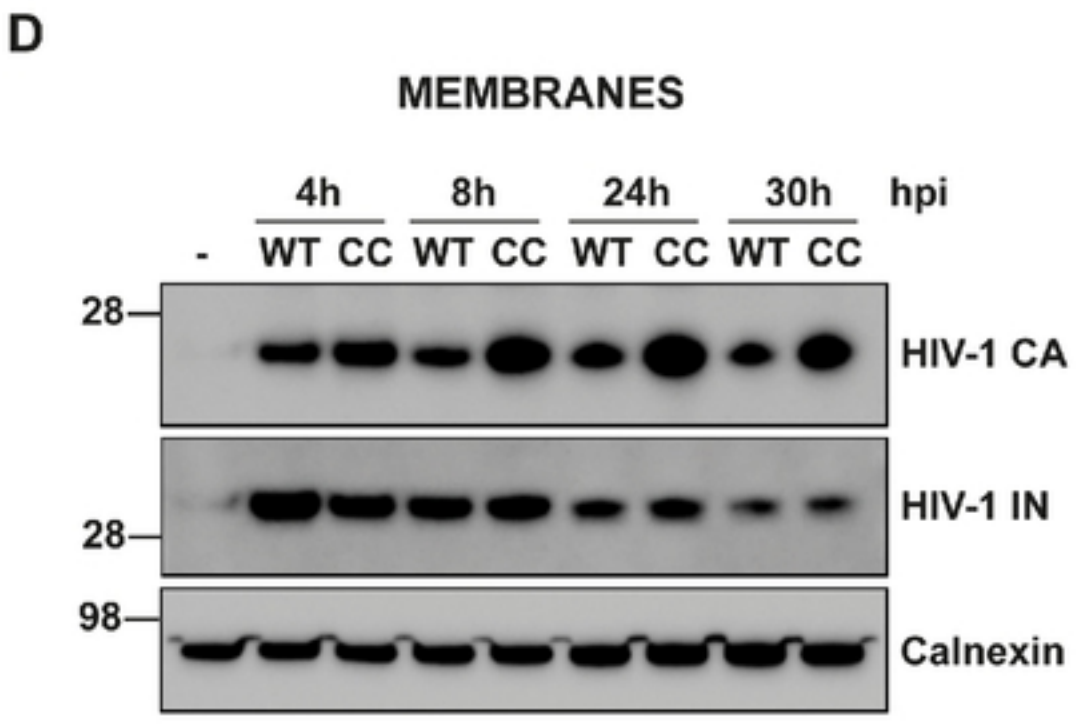


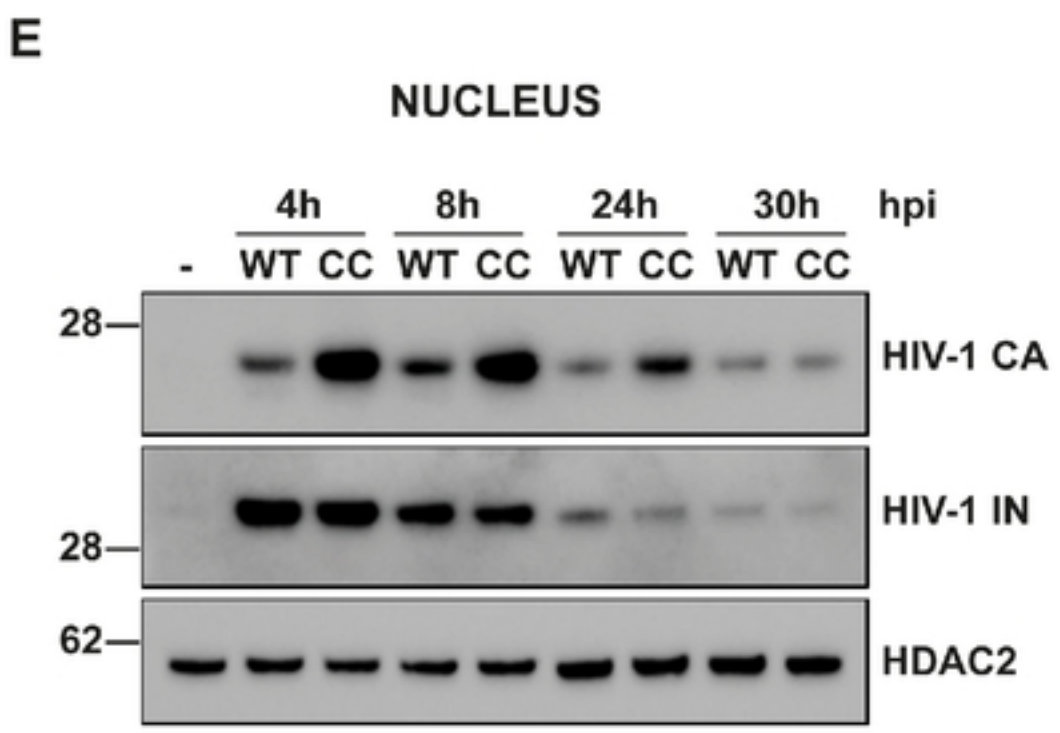
Fig 6

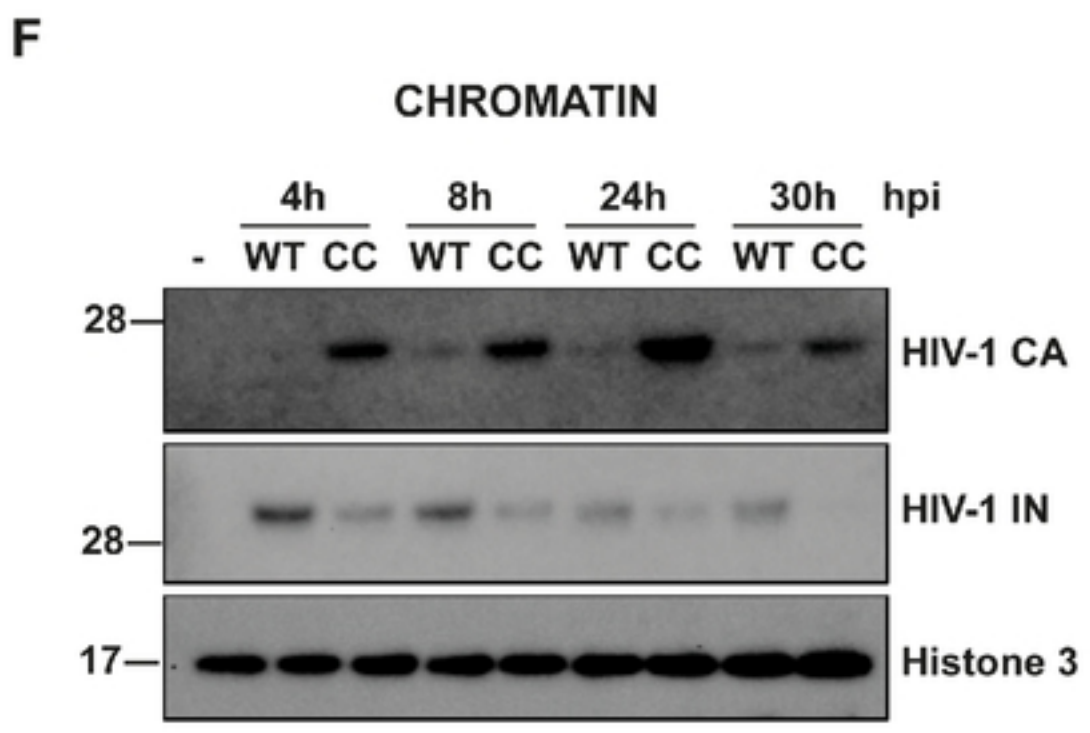




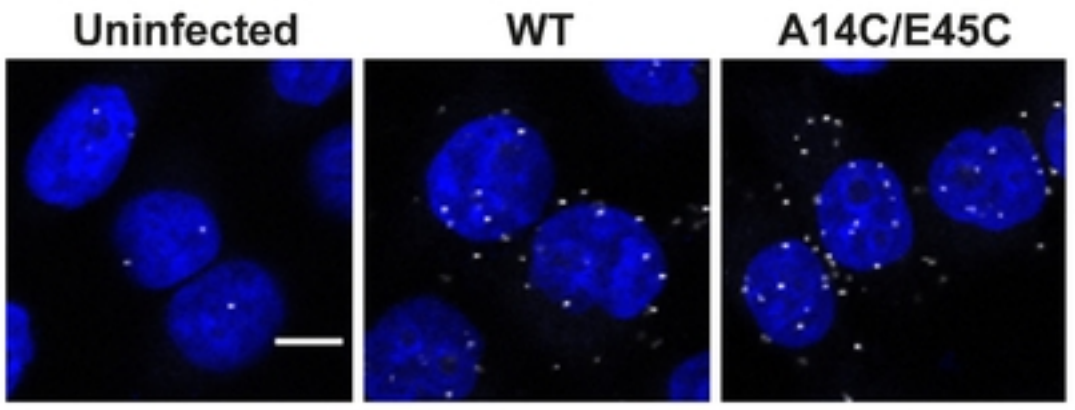




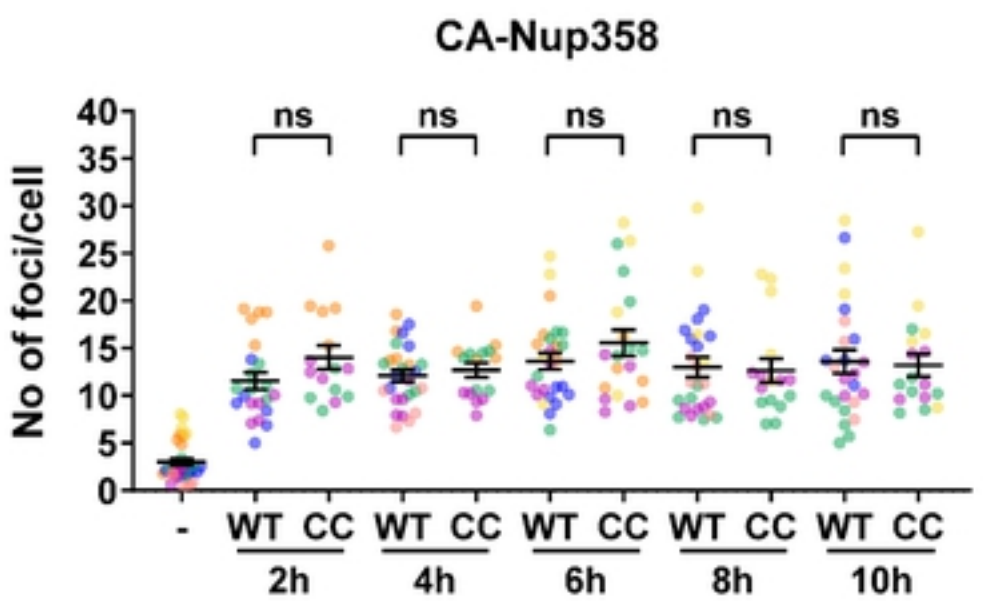




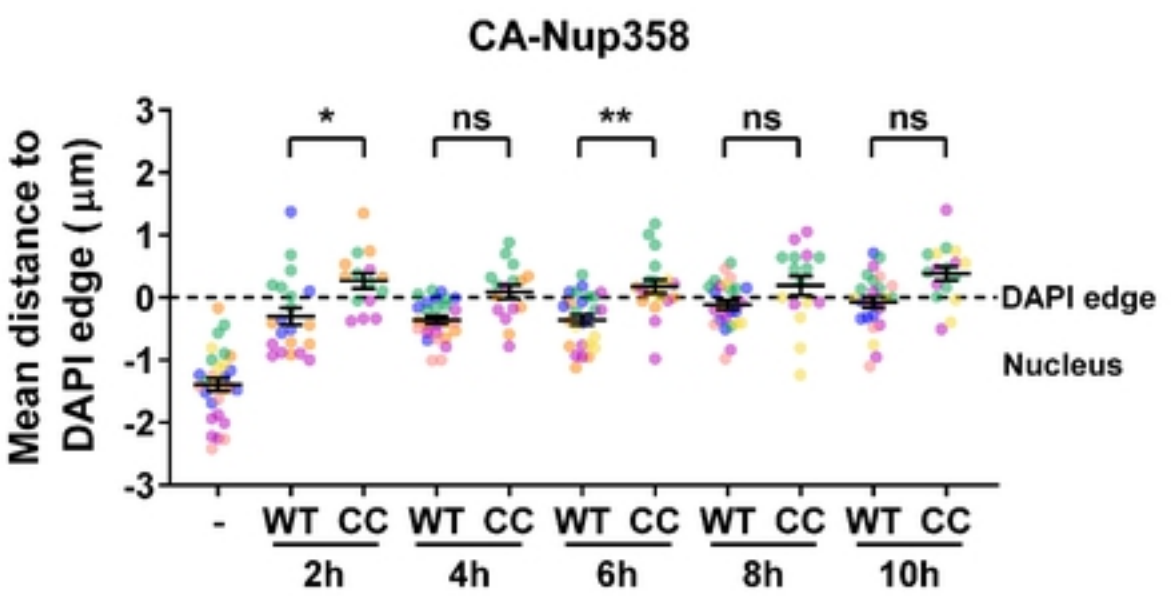
A



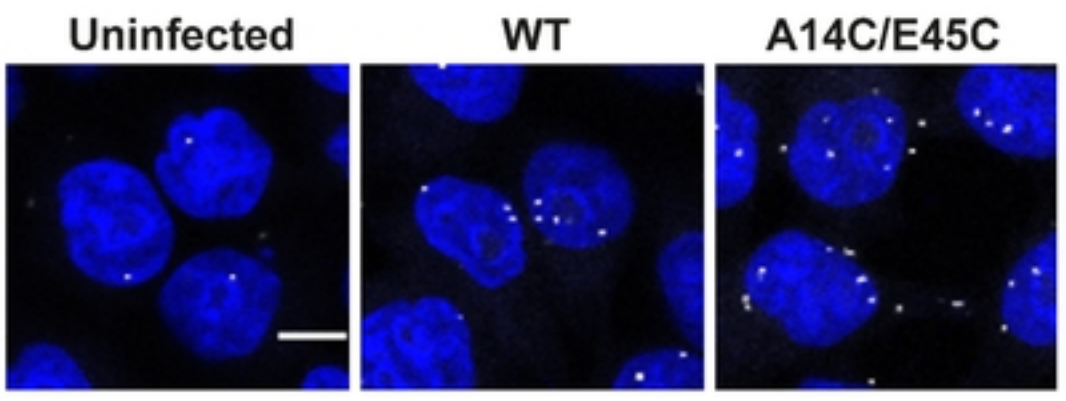
B



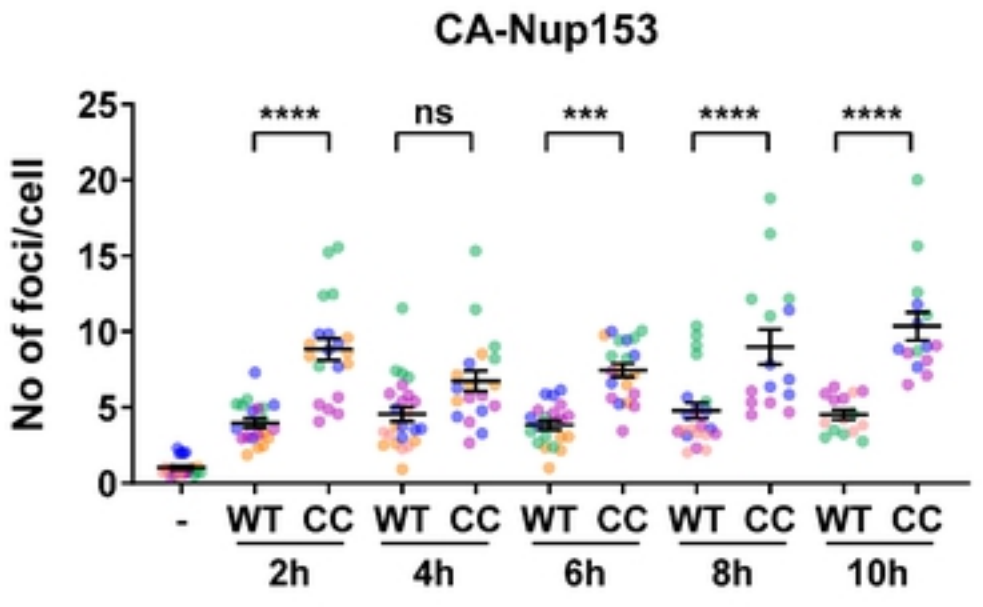
C



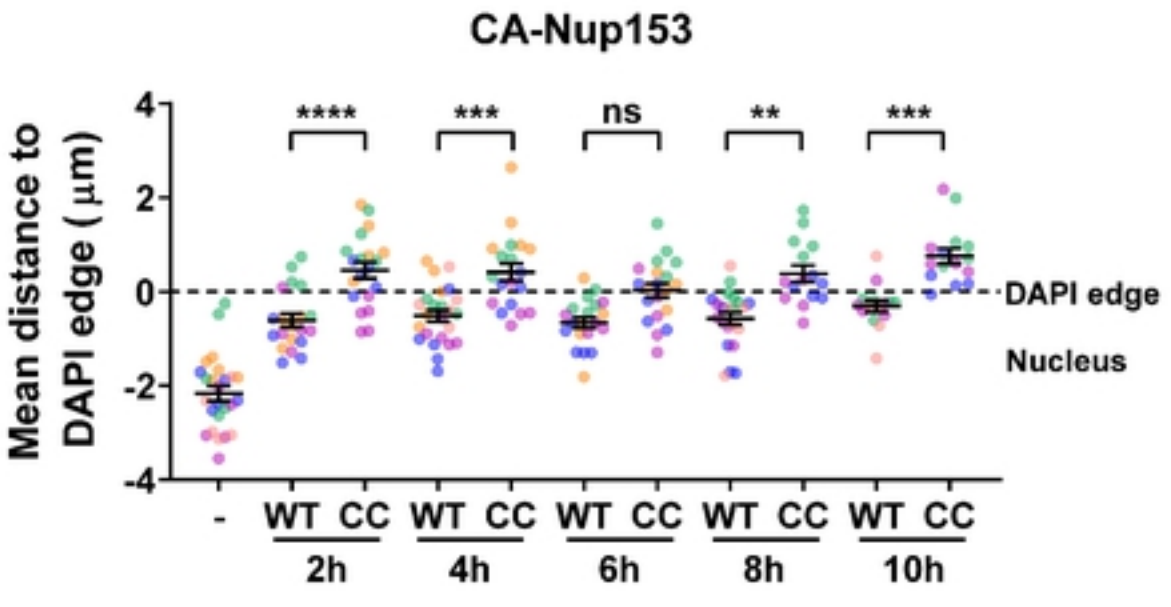
D



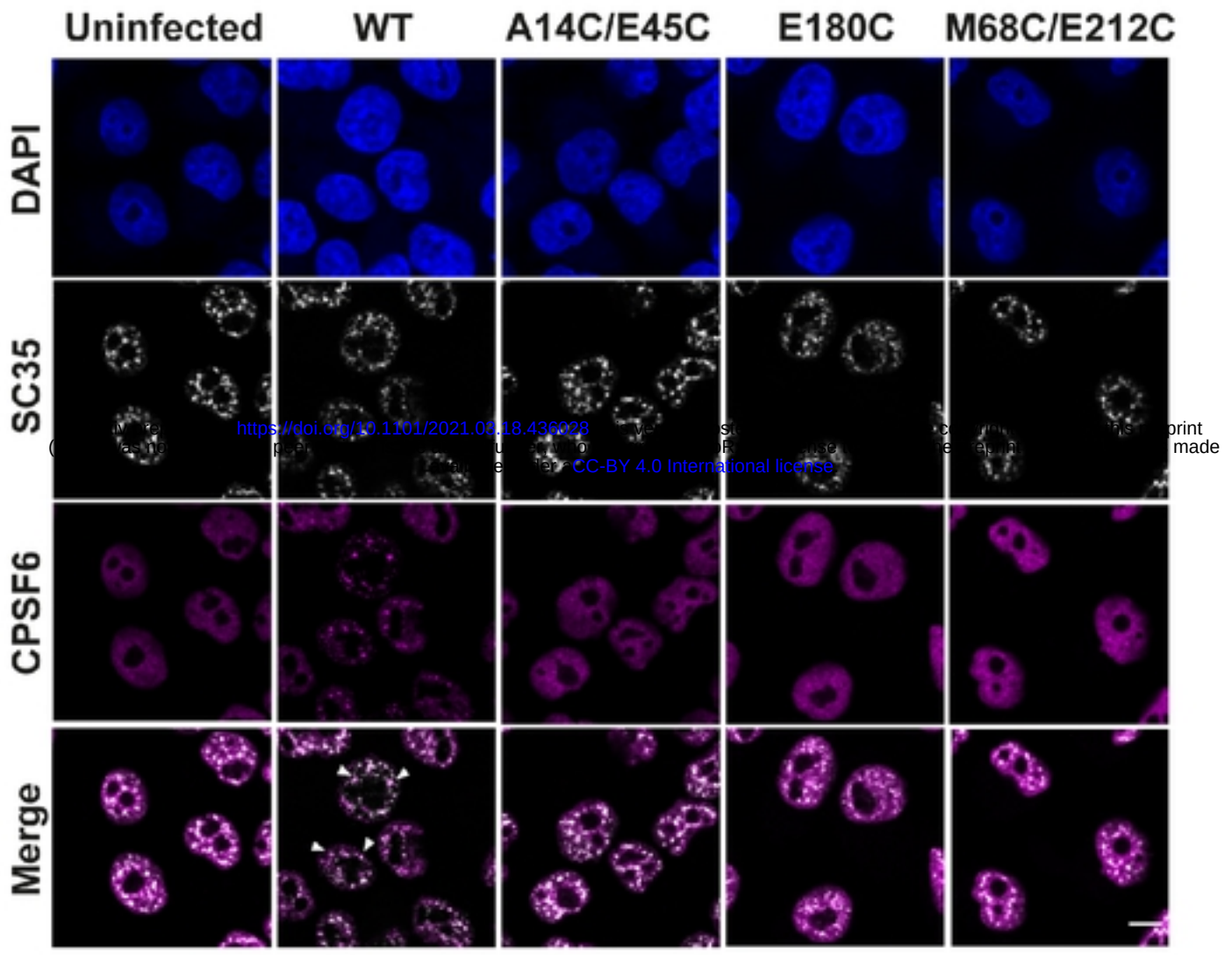
E



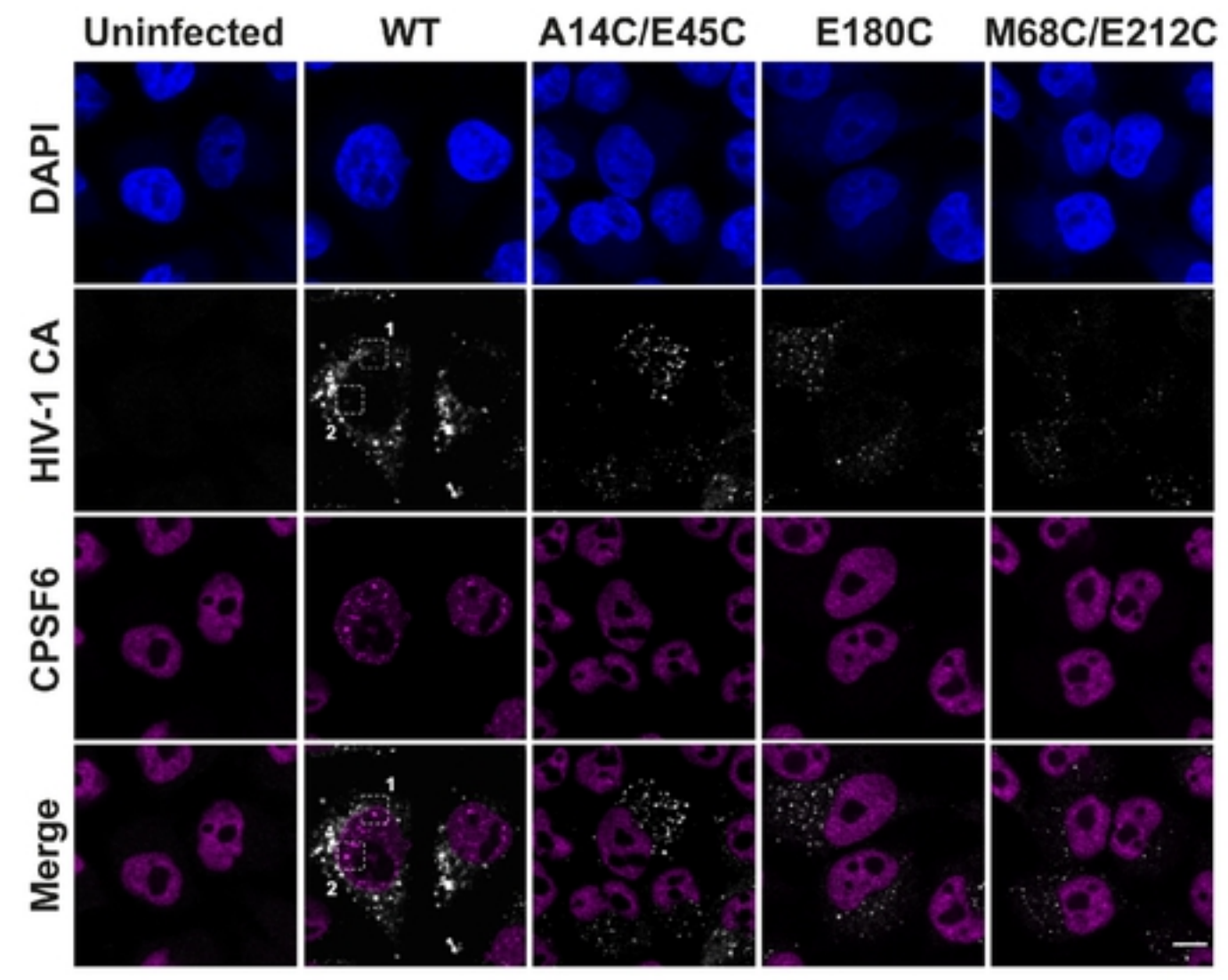
F

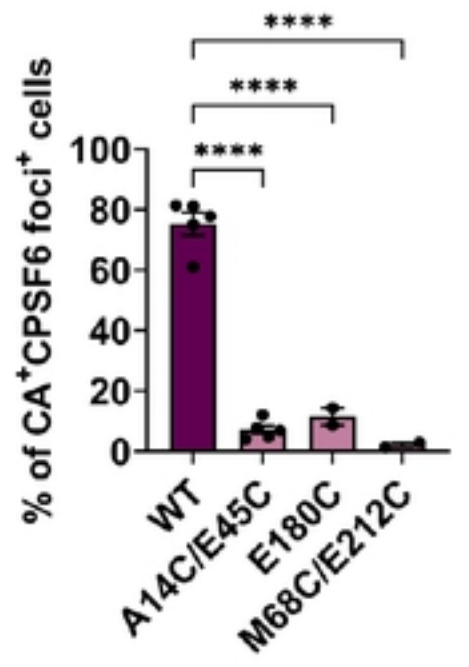


A



B





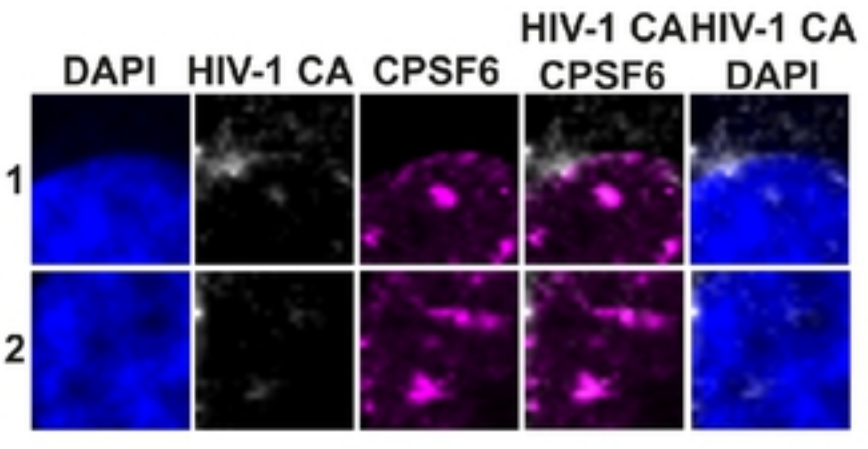


Fig 9