HIV-1 uncoats in the nucleus near sites of integration

Ryan C. Burdick[a], Chenglei Li[a], MohamedHusen Munshi[a], Jonathan M. O. Rawson[b], Kunio Nagashima[c], Wei-Shau Hu[b], and Vinay K. Pathak[a,1]

[a]Viral Mutation Section, HIV Dynamics and Replication Program, Center for Cancer Research, National Cancer Institute at Frederick, Frederick, MD 21702; [b]Viral Recombination Section, HIV Dynamics and Replication Program, Center for Cancer Research, National Cancer Institute at Frederick, Frederick, MD 21702; and [c]Electron Microscopy Laboratory, Cancer Research Technology Program, Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research (FNLCR), Frederick, MD 21702

Edited by Stephen P. Goff, Columbia University Medical Center, New York, NY, and approved January 29, 2020 (received for review November 22, 2019)

HIV-1 capsid core disassembly (uncoating) must occur before integration of viral genomic DNA into the host chromosomes, yet remarkably, the timing and cellular location of uncoating is unknown. Previous studies have proposed that intact viral cores are too large to fit through nuclear pores and uncoating occurs in the cytoplasm in coordination with reverse transcription or at the nuclear envelope during nuclear import. The capsid protein (CA) content of the infectious viral cores is not well defined because methods for directly labeling and quantifying the CA in viral cores have been unavailable. In addition, it has been difficult to identify the infectious virions because only one of ∼50 virions in infected cells leads to productive infection. Here, we developed methods to analyze HIV-1 uncoating by direct labeling of CA with GFP and to identify infectious virions by tracking viral cores in living infected cells through viral DNA integration and proviral DNA transcription. Astonishingly, our results show that intact (or nearly intact) viral cores enter the nucleus through a mechanism involving interactions with host protein cleavage and polyadenylation specificity factor 6 (CPSF6), complete reverse transcription in the nucleus before uncoating, and uncoat <1.5 h before integration near (<1.5 μm) their genomic integration sites. These results fundamentally change our current understanding of HIV-1 postentry replication events including mechanisms of nuclear import, uncoating, reverse transcription, integration, and evasion of innate immunity.

HIV-1 | capsid | uncoating | integration | transcription

The HIV-1 mature conical capsid core, composed of 250 CA hexamers and 12 pentamers (1–3), enters the cytoplasm upon fusion of the viral and host membranes and contains viral RNA and enzymes needed to complete viral replication. Determining where and when uncoating occurs is fundamental to understanding essential postentry replication events, including reverse transcription, evasion of host innate immunity, nuclear import, and integration. For the past four decades, retroviral uncoating has been thought to occur in the cytoplasm (4). Previous studies have proposed that intact viral cores with diameters of ∼61 nm (5) are too large to fit through the ∼39-nm inner diameter of nuclear pores (reviewed in ref. 6) and uncoating occurs in the cytoplasm in coordination with reverse transcription (7–12). A few recent studies have concluded that uncoating occurs at the nuclear envelope (NE) during nuclear import (12–17). Importantly, no published studies have concluded that HIV-1 viral cores remain intact or nearly intact during nuclear import and uncoat in the nucleus.

Previous studies of HIV-1 uncoating have been hampered for two major reasons. First, only one of ∼50 reverse transcription complexes/preintegration complexes (RTCs/PICs) in infected cells leads to provirus formation and productive infection (18); consequently, biochemical analyses of the population of RTCs/PICs may not reflect the properties of infectious viral complexes. Second, most previous studies of HIV-1 uncoating using imaging assays have relied on indirect CA detection methods to infer loss of CA from intracellular viral nucleoprotein complexes composed of a partial or intact viral core, viral genomic RNA or DNA, and enzymes (for simplicity, referred to as viral complexes hereafter). These indirect methods include immunofluorescence assays with anti-CA antibodies (14, 18–21) or fluorescent labeling of viral core-associated host protein cyclophilin A (CypA) (12, 15). However, immunofluorescence assays may be compromised by accessibility of the CA epitope in viral complexes, and loss of fluorescent CypA may report dissociation of the CypA from the viral core rather than viral core disassembly. Some previous studies have suggested that low levels of CA are associated with nuclear complexes (14, 15, 17, 19, 20, 22) and are thought to influence integration site selection (23, 24); in these studies, the nuclear CA amounts may have been underestimated because of reduced CA epitope accessibility (25). Although a few studies have directly labeled CA with a tetracysteine tag (26), this approach has not been widely used because of potential issues of nonspecific labeling and photobleaching.

Here, we developed methods to directly label CA with GFP and track viral cores in infected cells through viral DNA integration and proviral DNA transcription by live-cell microscopy. Astonishingly, our results show that infectious viral cores in the nucleus are intact (or nearly intact) and complete reverse transcription in the nucleus before uncoating. These observations fundamentally change our current understanding of HIV-1 postentry replication events including mechanisms of nuclear import and uncoating as well as reverse transcription, integration, and evasion of innate immunity. We also probed the mechanism of viral core nuclear import and show that intact or nearly intact

Significance

For several decades, retroviral core uncoating has been thought to occur in the cytoplasm in coordination with reverse transcription, and while some recent studies have concluded that HIV-1 uncoating occurs at the nuclear envelope during nuclear import, none have concluded that uncoating occurs in the nucleus. Here, we developed methods to study HIV-1 uncoating by direct labeling and quantification of the viral capsid protein associated with infectious viral cores that produced transcriptionally active proviruses. We find that infectious viral cores in the nuclei of infected cells are largely intact and uncoat near their integration sites just before integration. These unexpected findings fundamentally change our understanding of HIV-1 postentry replication events.

Author contributions: R.C.B. and V.K.P. designed research; R.C.B., C.L., M.M., and K.N. performed research; J.M.O.R. and W.-S.H. contributed new reagents/analytic tools; R.C.B., C.L., M.M., K.N., W.-S.H., and V.K.P. analyzed data; and R.C.B. and V.K.P. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: pathakv@mail.nih.gov.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1920631117/-/DCSupplemental.

MICROBIOLOGY

viral cores gain nuclear entry through a mechanism involving interactions at the NE with host protein CPSF6.

Results

Development of a Method to Directly Label HIV-1 CA in Viral Complexes. To directly label HIV-1 CA in viral complexes, we inserted GFP between matrix and CA as previously described (27, 28). In addition, mutation of the protease cleavage site between GFP and CA resulted in expression of a GFP-CA fusion protein enabling direct visualization of the viral core (Fig. 1 *A*, *Left*). This method produced virions that were labeled with a high efficiency and retained ~50% of their infectivity in HeLa cells, CEM-SS T cells, and THP-1–derived macrophages compared to virions made in the absence of GFP-CA (Fig. 1 *A*, *Right*). Most HIV-1 virions produced by cotransfecting the GFP-CA–expressing vector and a vector expressing wild-type (WT) *gag-pol* at a 1:15 ratio were labeled with GFP-CA (~96%; *SI Appendix*, Fig. S1 *A–C*). To determine whether GFP-CA remained associated with viral complexes through nuclear import, cells were infected with virions colabeled with GFP-CA and virion core-associated host restriction factor APOBEC3F (A3F) (18) fused to fluorescent protein red–red vine tomato (RRvT). The percentage of A3F-RRvT–labeled nuclear

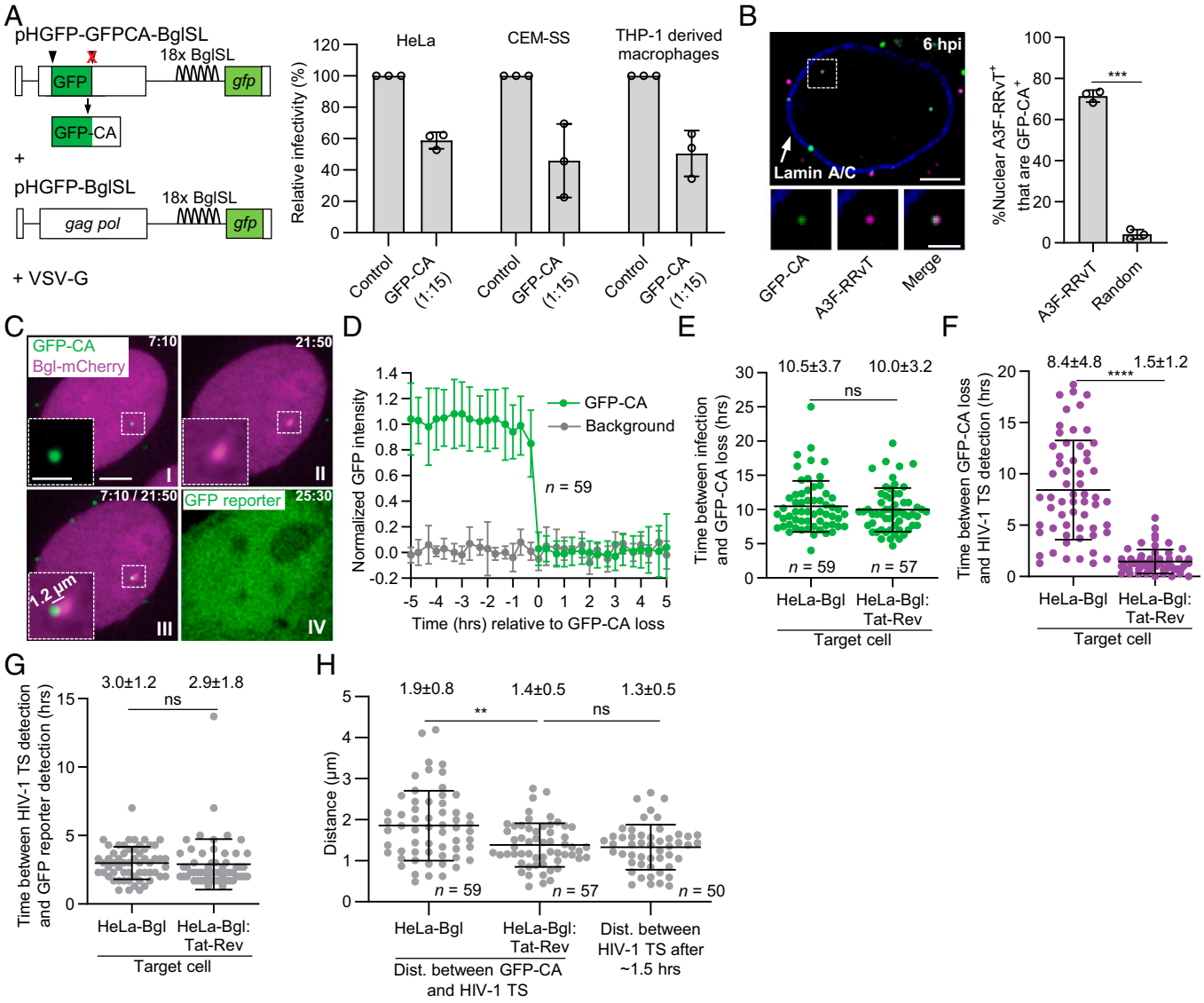

Fig. 1. GFP-CA–labeled viral complexes uncoat in the nucleus within ~1.5 μm of HIV-1 transcription sites. (*A*) HIV-1 vectors (*Left*) used to produce GFP-CA–labeled virions with high infectivity in HeLa, CEM-SS, and THP-1–derived macrophages (*Right*) compared to unlabeled control virions (set to 100%). (*B*) Nucleus of a HeLa cell infected with virions colabeled with GFP-CA + A3F-RRvT and immunostained with anti-Lamin A/C antibody (*Left*). Most nuclear A3F-RRvT viral complexes (~70%) have detectable GFP-CA signals 6-h postinfection (hpi) compared to random locations in the nucleus (*Right*). (*C*) Representative live-cell microscopy images of a HeLa-Bgl cell infected with GFP-CA–labeled virions. A GFP-CA–labeled nuclear viral complex uncoated and lost the GFP-CA signal 7:10 hpi (I) and HIV-1 TS appeared near the site of GFP-CA disappearance 21:50 hpi (II). The HIV-1 TS appeared 1.2 μm from the GFP-CA signal (III). GFP reporter expression detected 25:30 hpi (IV). (*D*) Average normalized GFP-CA intensities are stable before abrupt GFP-CA loss within a single frame (<20 min). (*E*) Time between infection and nuclear GFP-CA loss, (*F*) nuclear GFP-CA loss and HIV-1 TS appearance, and (*G*) HIV-1 TS appearance and *gfp* reporter detection for 59 and 57 infectious GFP-CA–labeled viral complexes in HeLa-Bgl cells and HeLa-Bgl:Tat-Rev cells, respectively. (*H*) Distance between GFP-CA signal (time point prior to GFP-CA loss) and HIV-1 TS (first time point of detection) in HeLa-Bgl cells (~8.4 h) and HeLa-Bgl:Tat-Rev cells (~1.5 h) compared to HIV-1 TS movements in ~1.5 h. (Scale bars, 5 μm; *Inset*, 2 μm.) For *A* and *B*, data are mean ± SD from three independent experiments; *P* values are from paired *t* tests. For (*E–H*), lines are mean ± SD; *P* values are from Welch's *t* tests. ****$P < 0.0001$; ***$P < 0.001$; **$P < 0.01$; *$P < 0.05$; ns, not significant ($P > 0.05$).

viral complexes that colocalized with detectable levels of GFP-CA was high (71%; Fig. 1*B*), indicating that GFP-CA remained associated with viral complexes through nuclear import at a high efficiency. The GFP-CA labeling at the 1:15 ratio did not have a significant impact on virion morphology since labeled and unlabeled virions displayed similar ratios of virions with mature and immature morphologies (*SI Appendix*, Fig. S1*D*). Sucrose-gradient fractionation of GFP-CA–labeled and unlabeled virions revealed similar proportions of intact viral cores, suggesting that GFP-CA labeling did not alter the in vitro stability of the viral cores (*SI Appendix*, Fig. S1 *E* and *F*).

To determine the effect of GFP-CA labeling on the efficiency of NE docking and nuclear import, viral cores composed of WT CA were labeled by incorporation of integrase-YFP (14) or A3F-YFP, and their efficiency of docking at the NE and nuclear import were compared to GFP-CA–labeled viral complexes as previously described (*SI Appendix*, Fig. S1*G*) (14). Briefly, infected cells were fixed at 6 hpi and viral complexes at the NE and in the nucleus were quantified using a custom MATLAB program. Similar NE docking and nuclear import efficiency of integrase-YFP-, A3F-YFP-, and GFP-CA–labeled viral complexes indicated that the GFP-CA labeling did not significantly influence viral complex association with the nuclear pores or nuclear import.

HIV-1 Uncoating Occurs ∼1.5 μm of Integration Sites <1.5 h before Integration. To image HIV-1 integrated proviruses, HIV-1 transcription sites (TSs) were visualized by specific recognition of RNA stem loops (29) in the HIV-1 vector RNA with mCherry-tagged bacterial protein (Bgl-mCherry; Fig. 1*A*). HeLa cells expressing Bgl-mCherry (HeLa-Bgl; Fig. 1*C*) infected with GFP-CA–labeled virions at a low multiplicity of infection (<0.1 GFP-expressing proviruses/cell; *SI Appendix*, Fig. S1*H*) were analyzed by live-cell imaging (Fig. 1*C* and Movie S1) to quantify nuclear GFP-CA–labeled viral complexes, HIV-1 TS, distances between GFP-CA–labeled viral complexes and HIV-1 TS, and *gfp* reporter expression. Live-cell imaging from ∼4–24 hpi identified intranuclear GFP-CA–labeled viral complexes that maintained a steady level of GFP-CA for several hours and abruptly lost the GFP-CA signal ∼10.5 hpi, indicating nuclear uncoating (Fig. 1 *D* and *E*; $n = 59$). HIV-1 TSs were detected near the sites of GFP-CA disappearance ∼8.4 h later (Fig. 1*F*) followed by detection of *gfp* reporter expression ∼3.0 h later (Fig. 1*G* and Movie S1).

To determine the effect of exogenous Tat and Rev expression on the time of HIV-1 TS appearance, we constructed HeLa-Bgl cells that constitutively express HIV-1 Tat and Rev proteins (HeLa-Bgl:Tat-Rev). We found that expression of Tat and Rev did not affect the kinetics of GFP-CA loss (∼10.0 hpi; Fig. 1*E*) or the time between HIV-1 TS detection and *gfp* reporter expression (∼2.9 h; Fig. 1*G*), but HIV-1 TSs were detected much faster (∼1.5 h) after GFP-CA loss (Fig. 1*F* and Movie S2). These results indicate that ∼6.9 h were needed after GFP-CA loss for Tat to reach sufficient levels of expression to produce detectable HIV-1 TSs. Treatment of cells with integrase inhibitor raltegravir (RAL) showed that most of the HIV-1 TSs detected were from integrated proviruses (*SI Appendix*, Fig. S1*I*). Interestingly, exogenous Tat-Rev expression promoted detectable transcription from unintegrated DNAs (*SI Appendix*, Fig. S1*I*), suggesting that silencing of unintegrated HIV-1 DNAs by the human silencing hub complex (30) can be suppressed or reversed by Tat expression. Comparisons with control vectors indicated that the BglSL stem loops, Vif and Vpr expression, and GFP-CA fusion protein did not influence the kinetics of nuclear uncoating or *gfp* reporter expression (*SI Appendix*, Fig. S1 *J–L*).

We compared the locations of HIV-1 uncoating and integration sites by adjusting for cell movements and then measuring the distances between the last frame in which GFP-CA puncta were detected and the first frame in which HIV-1 TSs were detected (Fig. 1*H*). The average distance in the Tat-Rev-expressing cells (∼1.4 μm) was similar to the average distance HIV-1 TS moved in ∼1.5 h (∼1.3 μm), and the previously reported constrained diffusion of genes within a 1.5-μm radius (31). The average distance in HeLa-Bgl cells was slightly higher (∼1.9 μm), perhaps due to the longer observation time (∼8.4 h vs. ∼1.5 h). These results demonstrate that viral complexes uncoat within ∼1.5 μm of the sites of integration.

Nuclear Uncoating Confers Resistance to CA-Binding Inhibitor PF-3450074 and Is Delayed by Inhibiting Reverse Transcription. The CA-binding inhibitor PF-3450074 (PF74) binds to the N-terminal domain of one CA subunit and the C-terminal domain of an adjacent CA subunit within a hexamer and destabilizes that viral cores; interestingly, CPSF6 binds to the same site at which PF74 binds (reviewed in ref. 32). Treating infected cells after nuclear import of GFP-CA–labeled viral complexes with PF74 resulted in rapid disappearance (∼12.9 min) of 86% of the nuclear viral complexes (Fig. 2 *A* and *B* and Movie S3), indicating that nuclear viral complexes contained CA hexamers.

Next, we performed time-of-addition experiments with PF74, reverse transcriptase inhibitor nevirapine (NVP), or integrase inhibitor RAL. Addition of PF74 to cells infected with unlabeled HIV-1 virions at various times after infection showed that 50% of the viral complexes became PF74 resistant ∼11.5 hpi (Fig. 2 *C* and *D*). This average time of PF74 sensitivity loss was similar to the average time of GFP-CA loss (∼10.5 hpi; Fig. 1*E*), indicating that nuclear uncoating was correlated with PF74 resistance. The loss of PF74 sensitivity occurred ∼3.0 h after the loss of sensitivity to NVP (Fig. 2*D*), suggesting that nuclear uncoating occurred ∼3 h after completion of reverse transcription. Nevertheless, inhibition of reverse transcription with NVP was correlated with a delay or inhibition of nuclear uncoating (Fig. 2 *E* and *F*).

RAL time-of-addition experiments showed that integration was completed ∼1.0 h after uncoating (Fig. 2*D*) and that inhibiting integration did not affect uncoating (Fig. 2 *E* and *F*). We also determined that unlabeled- and GFP-CA–labeled viral complexes became PF74 resistant with similar kinetics (∼11.4 hpi; Fig. 2*G*). Since loss of sensitivity to PF74 was correlated with uncoating, the result indicated that GFP-CA labeling of viral complexes did not significantly influence their uncoating kinetics. GFP-CA–labeled and unlabeled virions exhibited the same sensitivity to NVP, PF74, and RAL with nearly identical 50% inhibitory concentrations (*SI Appendix*, Fig. S1*M*).

Finally, we determined that the average time at which GFP-CA–labeled viral complexes are imported into the nucleus was ∼4.4 hpi (Fig. 2*H*), which was not significantly different from our previously determined average time of import (∼4.3 hpi) for viral complexes labeled with integrase-YFP or A3F-YFP (14). Since the average time of nuclear import (∼4.4 hpi) was ∼4 h earlier than the average time of reverse transcription completion (∼8.5 hpi), we conclude that most viral complexes complete reverse transcription after nuclear import.

Nuclear Viral Complexes Retain Most of the CA Present in Intact Viral Cores. Previous studies have suggested that one uncoating step occurs at the NE during import (12–17). To determine whether some uncoating occurs during nuclear import, we compared the average GFP-CA intensities of viral complexes while they were docked at the NE for six frames before and six frames after nuclear entry and found no detectable loss of GFP-CA (Fig. 3 *A* and *B*, $n = 18$). Modeling 5–20% loss of GFP-CA intensities indicated that a ≥10% GFP-CA loss would have been detectable under the imaging conditions (*SI Appendix*, Fig. S2*A*); thus, most of the CA was retained by the viral complexes during nuclear import.

To compare the GFP-CA intensities of nuclear viral complexes and intact viral cores, intact virions were lysed in vitro as previously reported (33), resulting in ∼55% loss of free GFP-CA that was not incorporated into viral cores (Fig. 3*C* and *SI Appendix*,

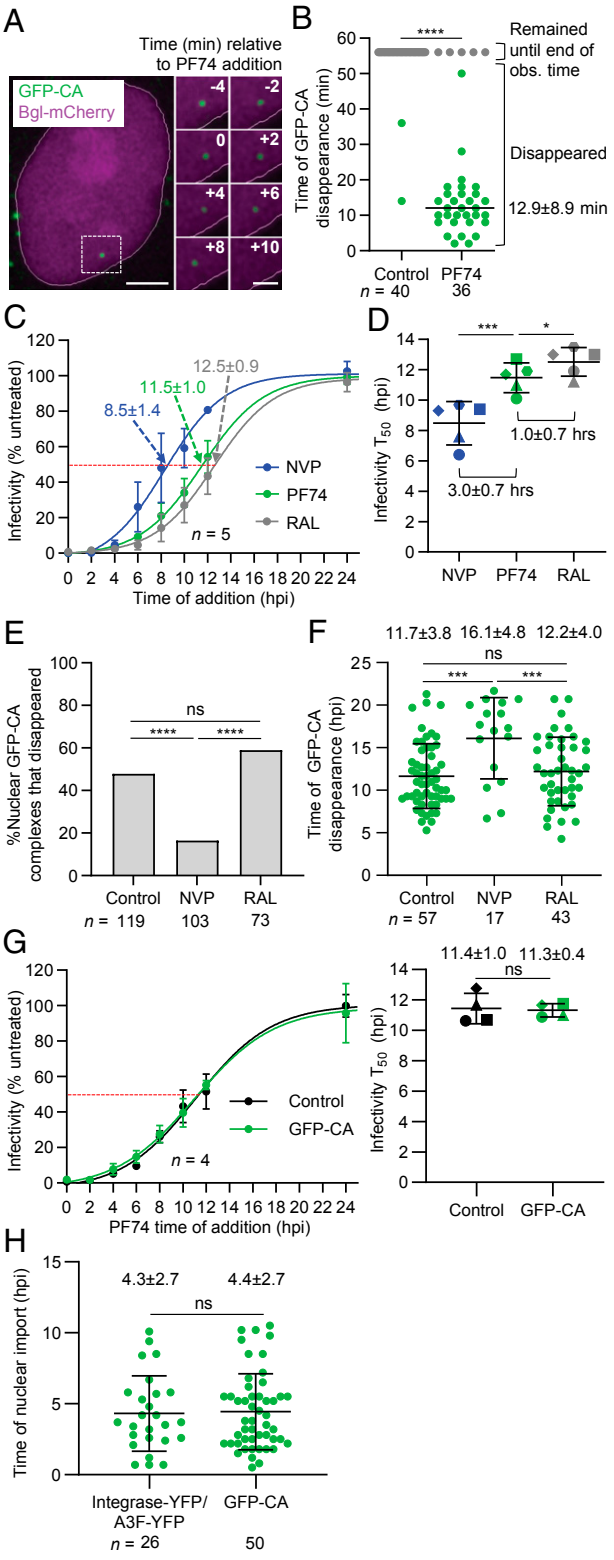

Fig. 2. Determination of the sensitivity of nuclear GFP-CA–labeled viral complexes to capsid, reverse transcriptase and integrase inhibitors. (*A*) Representative live-cell microscopy images of a nuclear GFP-CA–labeled viral complex before and after addition of PF74 (10 μM). Numbers in white indicate time (min) relative to the time of PF74 addition. (Scale bar, 5 μm; *Inset*, 2 μm.) (*B*) Time of disappearance of GFP-CA–labeled viral complexes in untreated control cells and PF74-treated cells during ∼1-h observation time. (*C*) Time-of-addition assays with NVP, PF74, or RAL. The numbers indicate the time at which 50% of the viral complexes became resistant to the inhibitors (infectivity T_{50}). (*D*) Comparison of average infectivity T_{50} for NVP, PF74, and RAL from five independent experiments. (*E*) Proportion of nuclear GFP-CA–labeled complexes that disappeared during the observation time (21.6 hpi). (*F*) Average time of GFP-CA disappearance. Lines are mean ± SD; *P* values are from Welch's *t* tests. (*G*) PF74 time-of-addition assays with GFP-CA–labeled and unlabeled virions. Comparison of average time at which 50% of the viral complexes became resistant to PF74 (infectivity T_{50}) from four independent experiments (*Right*). (*H*) Comparison of the time of nuclear import previously determined for integrase-YFP- or A3F-YFP-labeled viral complexes (14) and GFP-CA–labeled viral complexes. For *B* and *E*, *P* values are from Fisher's exact tests comparing the proportion nuclear GFP-CA complexes that disappeared. For *D* and *G*, lines are mean ± SD; *P* values are from paired *t* tests. ****$P < 0.0001$; ***$P < 0.001$; *$P < 0.05$; ns, not significant ($P > 0.05$).

Fig. S2 *B–D*). The mean GFP-CA intensity of 116 nuclear viral complexes from HeLa cells or 223 nuclear viral complexes from CEM-SS T cells was not significantly different from the intact viral cores (Fig. 3 *C* and *D*). Modeling 2–10% GFP-CA loss from intact viral cores indicated that ≥6% GFP-CA loss would have been detectable (*SI Appendix*, Fig. S2*E*), indicating that infectious nuclear RTCs/PICs retained ≥90% of their viral core-associated CA. In addition, the results also indicated that most reverse transcription is completed within an intact (or nearly intact) viral core since most viral complexes became NVP resistant (∼8.5 hpi) before GFP-CA loss (∼10.5 hpi).

Disruption of the CA-CPSF6 Interaction Results in Uncoating at the NE. CPSF6 is a nuclear host factor that interacts with the viral core (34) and disruption of the CA-CPSF6 interaction alters the target sites of HIV-1 integration (23, 24, 35). To elucidate the role of CPSF6 in nuclear import and intranuclear trafficking, we generated GFP-CA–labeled virions containing CA mutations N74D (34) or A77V (36), which substantially reduce CPSF6 binding. As expected, the N74D and A77V mutations had minimal effects on infectivity in HeLa cells, CEM-SS cells, or THP-1–derived macrophages (*SI Appendix*, Fig. S3*A*). However, the A77V mutation has been reported to revert in humanized mice (36), indicating a lower fitness in vivo. We infected HeLa cells expressing Bgl-mCherry with the N74D and A77V mutants and observed that the kinetics of GFP-CA loss, HIV-1 TS appearance, and *gfp* reporter detection (Fig. 4 *A* and *B*, *SI Appendix*, Fig. S3 *B–D*, and Movie S4) were not different from the kinetics of WT virions that were described in Fig. 1 *E–G*.

Next, we compared the locations of the WT, N74D, and A77V GFP-CA puncta. Bgl-mCherry protein is predominantly localized to the nucleus (*SI Appendix*, Fig. S3*E*), and the NE is localized to the periphery of the Bgl-mCherry signal. We found that, in contrast to the WT, the N74D and A77V GFP-CA puncta did not enter the nucleus but disappeared at the periphery of the Bgl-mCherry signal (*SI Appendix*, Fig. S3 *E* and *F*), indicating that they uncoated at the NE (Fig. 4*A* and *SI Appendix*, Fig. S3*B*). Interestingly, the N74D and A77V mutants exhibited GFP-CA loss at the NE at approximately the same time after infection (∼9.7 and ∼9.4 hpi, respectively) as the GFP-CA loss exhibited by WT viral complexes in the nucleus (∼10.5 hpi) (Fig. 4*B*). This observation suggested that the molecular events that trigger uncoating occur in the nucleus and at the NE with the same kinetics. Consistent with this finding, the NE residence time of GFP-CA–labeled viral complexes prior to nuclear import was much longer for the N74D (∼5.5 h) and A77V (∼4.9 h) mutants than for the WT complexes (∼1.9 h) (Fig. 4*C*). Subsequently, HIV-1 TS appeared near the site of N74D and A77V GFP-CA puncta disappearance at the NE followed by *gfp* reporter expression. The HIV-1 TSs were much closer to the NE (∼0.8 μm) in cells infected with N74D and A77V mutants compared to WT HIV-1 TSs (∼1.5 μm; Fig. 4*D*). These observations suggested that, after uncoating at the NE, the N74D/A77V PICs (without the viral cores) integrated into nearby chromatin either by entering the nucleus or by accessing the adjacent chromatin while docked at the NPC.

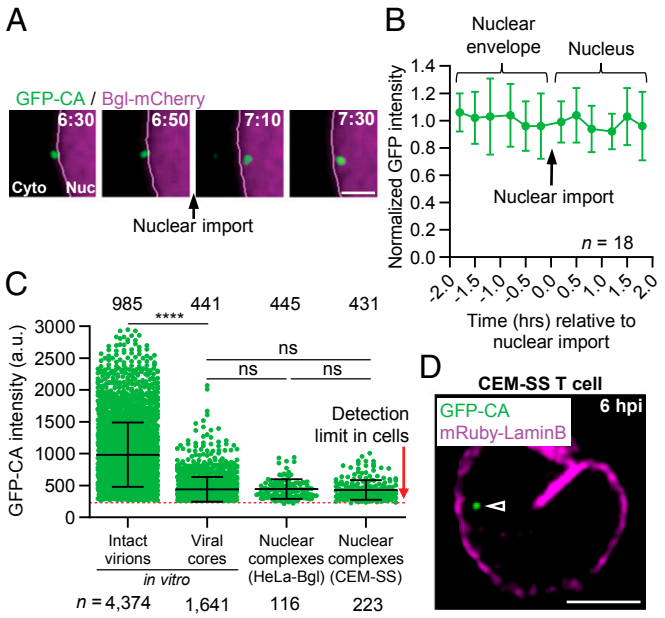

Fig. 3. Nuclear viral complexes retain most of the GFP-CA associated with in vitro viral cores. (*A*) Live-cell microscopy images of a GFP-CA–labeled viral complex docked at the NE and in the nucleus after import. Numbers in white indicate time postinfection. (Scale bar, 2 μm.) (*B*) Normalized mean GFP-CA intensities of six frames before and after nuclear import indicate no significant loss of GFP-CA. (*C*) Comparison of the mean GFP-CA intensities (arbitrary units; a.u.) of intact virions, in vitro viral cores, and nuclear viral complexes in infected HeLa and CEM-SS T cells. Intact virions and in vitro viral cores with GFP-CA intensities below the detection limit (<265 a.u.) in HeLa and CEM-SS cell nuclei were removed. Lines are mean ± SD; *P* values are from Welch's *t* tests. ****$P < 0.0001$; ns, not significant ($P > 0.05$). (*D*) Representative image of a GFP-CA–labeled nuclear complex in an infected CEM-SS T cell expressing mRuby-LaminB 6 hpi. (Scale bar, 5 μm.)

Consistent with this hypothesis, only WT GFP-CA labeled viral complexes, but not N74D/A77V GFP-CA–labeled viral complexes, were detected in the nuclei of infected CEM-SS and HeLa cells (Fig. 4 *E* and *F* and *SI Appendix*, Fig. S3*G*). Overall, these results indicated that N74D/A77V GFP-CA–labeled viral cores did not enter the nucleus and uncoated while they were docked at the NE; furthermore, the results suggested that the CA-CPSF6 interaction is necessary for nuclear import of intact or nearly intact viral cores.

CA-CPSF6 Interaction at the NE Facilitates Nuclear Entry of Viral Cores. To visualize CA-CPSF6 interactions, we stably knocked down endogenous CPSF6 and expressed short hairpin RNA-resistant mRuby-CPSF6, which did not significantly influence virus infectivity, efficiency of NE docking, or nuclear import efficiency (Fig. 4*G* and *SI Appendix*, Fig. S4 *A–C*). Live-cell imaging of 18 GFP-CA–labeled WT viral complexes that entered the nucleus showed that they all accumulated mRuby-CPSF6 ∼0.6 h after NE docking and the dual-labeled complexes translocated to the nucleus ∼2.2 h later (Fig. 4*H*, *SI Appendix*, Fig. S5 *A–C*, and Movie S5); similar kinetics were observed for A3F-mNeonGreen-labeled viral complexes, indicating that GFP-CA labeling of viral complexes did not influence the kinetics of accumulation of mRuby-CPSF6 or the translocation of the dual-labeled complexes into the nucleus (*SI Appendix*, Fig. S4*D*). Interestingly, a high proportion of the WT viral complexes that did not enter the nucleus (47%) were also associated with mRuby-CPSF6 (Fig. 4*I*), indicating that the CA-CPSF6 association at the NE is necessary but is not sufficient for nuclear import of the viral core. None of the N74D or A77V GFP-CA–labeled viral complexes at the NE colocalized with mRuby-CPSF6 (0/50), confirming that a specific CA-CPSF6 interaction is required to accumulate CPSF6 at the NE.

We sought to determine how long after nuclear import CPSF6 dissociated from the viral complexes. Analysis of 24 GFP-CA–labeled viral complexes that entered the nucleus showed that the GFP-CA and mRuby-CPSF6 signals disappeared simultaneously, indicating that CPSF6 dissociated from the viral complexes at the time of uncoating (Fig. 4*J*, *SI Appendix*, Figs. S4*E* and S5 *D–F*, and Movie S6).

Since the HeLa:mRuby-CPSF6 cells contained 62% the level of CPSF6 as the parental HeLa cells (*SI Appendix*, Fig. S4*A*), we asked whether the reduced CPSF6 levels had any effect on HIV-1 uncoating. GFP-CA–labeled viral complexes uncoated in the nuclei of HeLa:mRuby-CPSF6 cells with the same efficiency and kinetics as in HeLa-Bgl cells (*SI Appendix*, Fig. S4 *F* and *G*), indicating that the reduced CPSF6 levels did not significantly influence the timing or efficiency of uncoating. Finally, the mRuby-CPSF6 levels associated with GFP-CA- or integrase-superfolder GFP-labeled nuclear viral complexes were similar (*SI Appendix*, Fig. S4 *H* and *I*), suggesting that the nuclear viral complexes contained similar amounts of CA, which resulted in similar amounts of mRuby-CPSF6 association.

Discussion

Here, we show that intact (or nearly intact) viral cores enter the nucleus and uncoat <1.5 h before integration within ∼1.5 μm of their chromosomal integration sites (model shown in Fig. 5). These results shift the current paradigm of HIV-1 postentry replication events and have important implications for the mechanisms of nuclear import and uncoating as well as reverse transcription, integration, and evasion of host innate immunity.

Our studies provide a robust method for fluorescently labeling viral cores in infected cells. GFP-CA–labeled virions were not significantly different from unlabeled virions with respect to 1) the ratio of mature and immature virions, 2) in vitro stability of viral cores during sucrose gradient fractionation, 3) proportion of viral cores that stably associated with the NE and imported into the nucleus, 4) the timing of GFP reporter expression, and 5) sensitivity to reverse transcriptase, capsid, and integrase inhibitors. Furthermore, the GFP-CA–labeling efficiency was high (96%), and virion infectivity was within twofold (∼50%) of the unlabeled virions. Most importantly, for these studies, time-of-addition assays showed that the GFP-CA–labeled and unlabeled virions became resistant to PF74 with similar kinetics, which was correlated to uncoating, indicating that GFP-CA labeling did not significantly affect the viral uncoating kinetics. Although we cannot exclude the possibility that GFP-CA labeling has some effects on HIV-1 replication that were not revealed in our experiments, we conclude that GFP-CA labeling does not substantially influence most aspects of HIV-1 replication.

Our observation that infectious viral cores uncoat in the nucleus was unexpected since most previous studies have concluded that uncoating occurs in the cytoplasm (7–12, 37), while a few recent studies have concluded that uncoating occurs at the NE during nuclear import (12–17). Previous imaging studies of HIV-1 uncoating have been hampered by the inability to fluorescently label CA directly without adversely affecting uncoating and viral replication. Although a few studies have labeled CA with a tetracysteine tag (7, 26), the method has not been widely used because of technical challenges, such as nonspecific labeling and rapid photobleaching. Immunofluorescence assays using anti-CA antibodies are widely used, and loss of the fluorescent signal is interpreted as loss of CA from viral complexes. However, reduced CA epitope accessibility by conformational changes in the viral core or association with host proteins can also result in loss of the fluorescent signal and potentially be misinterpreted as loss of CA from viral complexes. Recently, loss of fluorescently labeled CypA (CypA-dsRed) from viral complexes at the NE was interpreted as uncoating

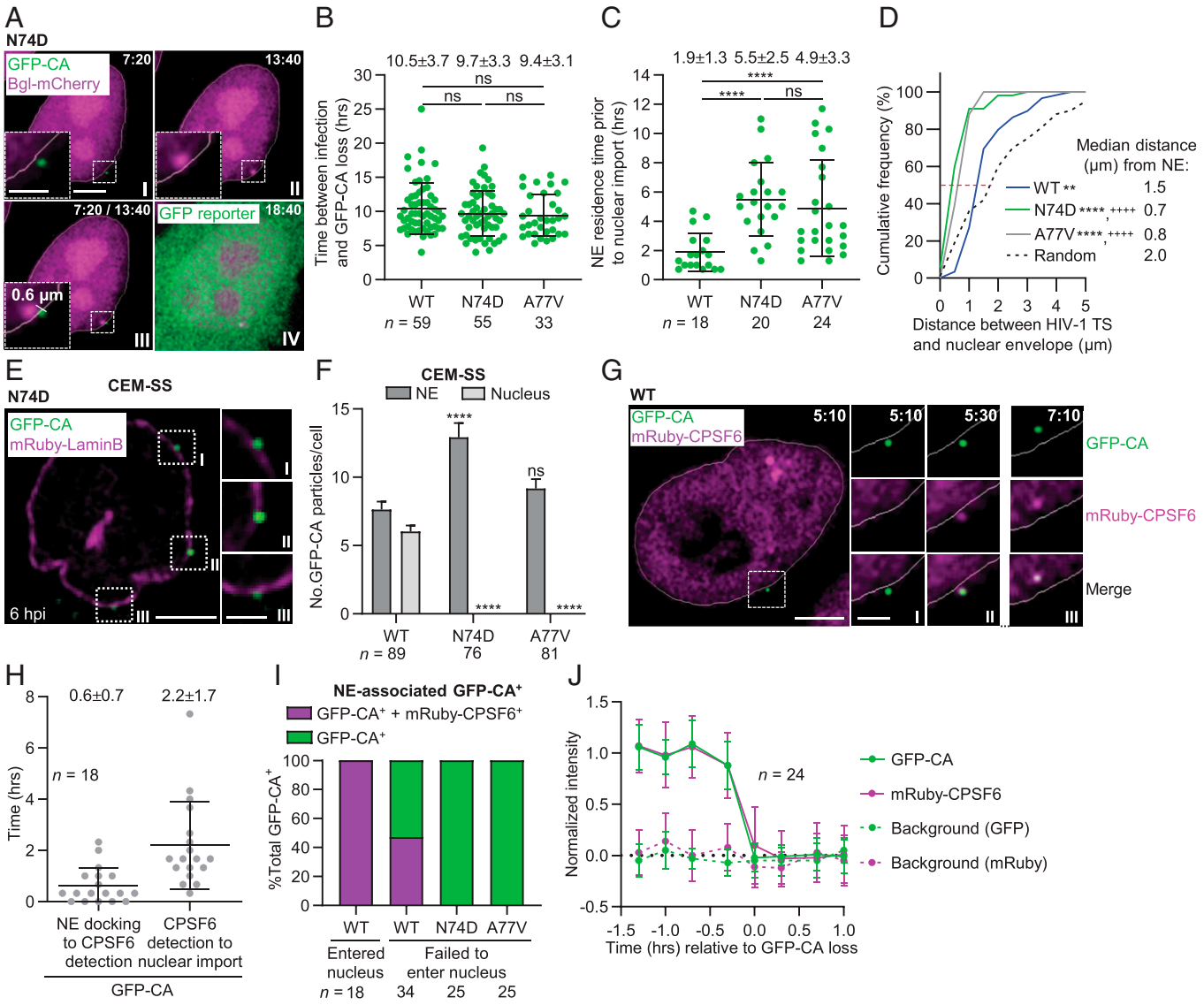

Fig. 4. CA-CPSF6 interaction at the NE facilitates nuclear import of viral complexes, the location of their uncoating, and the location of HIV-1 TS. (*A*) Representative live-cell microscopy images of a HeLa-Bgl cell infected with GFP-CA–labeled virions of CA mutant N74D. A GFP-CA–labeled viral complex uncoated at the edge of the nuclear Bgl-mCherry signal, 7:20 hpi (I) and HIV-1 TS appeared near the site of GFP-CA disappearance 13:40 hpi (II). The HIV-1 TS appeared 0.6 μm from the GFP-CA signal (III). GFP reporter expression detected 18:40 hpi (IV). (*B*) Time between infection and GFP-CA loss; data for WT the same as in Fig. 1*E*. Lines are mean ± SD; *P* values are from Welch's *t* tests. (*C*) NE residence time of GFP-CA–labeled viral complexes prior to nuclear import. For N74D/A77V mutants, the time of nuclear import was assumed to occur at the time of GFP-CA loss (i.e., uncoating). (*D*) Cumulative frequency distribution of distances (μm) between HIV-1 TS and NE and random sites in the nuclei to NE; median distances are indicated by the red dotted line. *P* values are from Kolmogorov–Smirnov tests. $^{**}P < 0.01$ compared to random; $^{****}P < 0.0001$ compared to random; $^{++++}P < 0.0001$ compared to WT. (*E*) N74D GFP-CA–labeled viral complexes localize to the NE but not in the nuclei of CEM-SS cells. (*F*) Quantitation of GFP-CA–labeled viral complexes at the NE and in the nucleus. Data are pooled from three independent infections (*n* = number of cells analyzed); *P* values are from paired *t* tests. $^{****}P < 0.0001$; ns, not significant ($P > 0.05$). (*G*) Representative live-cell microscopy images of infected HeLa:mRuby-CPSF6 cells show mRuby-CPSF6 recruitment to a GFP-CA–labeled viral complex located at or near the NE (I and II); dual-labeled GFP-CA and mRuby-CPSF6 complexes are imported into the nucleus (III). (*H*) Time between NE docking to CPSF6 detection and CPSF6 detection to nuclear import for 18 GFP-CA–labeled viral complexes. Lines are mean ± SD (*I*) Proportion of NE-associated GFP-CA$^+$ viral complexes that are CPSF6$^+$. The viral complexes that entered the nucleus during the observation time and those that were docked at the NE but failed to enter the nucleus were analyzed. (*J*) Simultaneous disappearance of intranuclear GFP-CA and mRuby-CPSF6 signals. (Scale bars, 5 μm; *Inset*, 2 μm.)

of the viral complexes at the NE during nuclear import; however, it is conceivable that CypA-dsRed dissociated from the viral complexes at the NE prior to nuclear import. Biochemical assays have also been used to study viral core uncoating in infected cells (38); however, only one in ∼50 virions leads to productive infection (18), and disassembly of a bulk population of viral cores may not reflect the behavior of the rare infectious viral cores.

Our studies provide direct evidence that CPSF6 recruitment is a critical requirement for nuclear import of intact or nearly intact viral cores. These results provide essential mechanistic insights into previous observations indicating that the CA-CPSF6 interaction is critical for integration into gene-rich euchromatin regions that are located in the interior regions of the nucleus (23, 24, 35, 39, 40). We propose the nuclear import of intact (or nearly intact) viral cores, and uncoating in the nucleus is essential for integration in gene-rich euchromatin regions since the N74D/A77V mutants, which uncoat at the NE, integrate into gene-sparse heterochromatin regions within ∼0.8 μm of the NE. The mechanism by which the

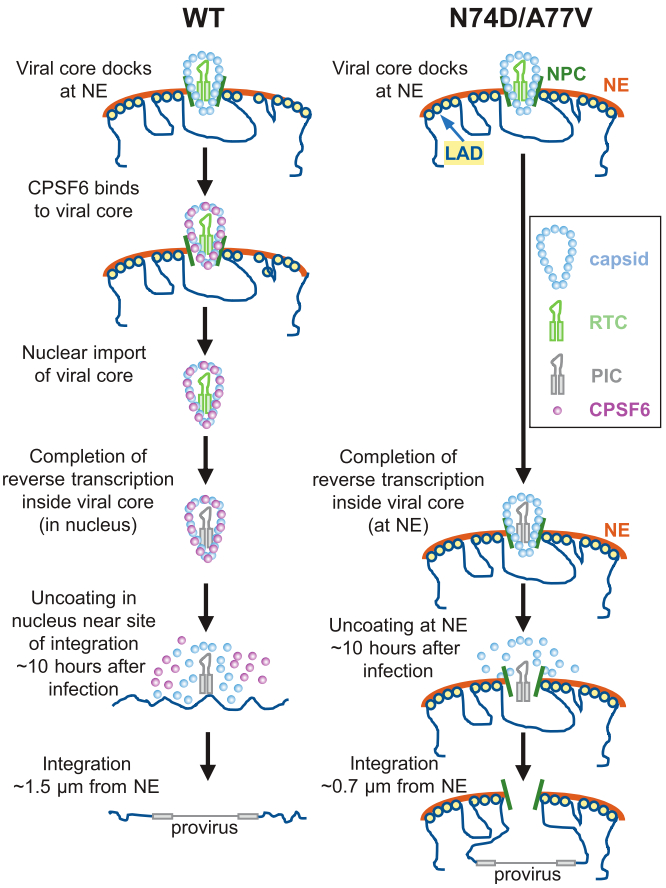

Fig. 5. Model for nuclear import and uncoating of HIV-1. WT and N74D/A77V viral cores dock at the NE. CPSF6 is recruited to the WT viral cores at the NE but not to the N74D/A77V viral cores. WT viral cores are imported into the nucleus ~1.9 h after docking at the NE; the N74D/A77V GFP-CA–labeled viral cores remain associated with the NE and are not imported into the nucleus. Reverse transcription is completed inside the intact (or nearly intact) viral core for WT (in the nucleus) and N74D/A77V mutants (at NE). The nuclear WT viral complexes and NE-associated N74D/A77V viral complexes uncoat ~10 h after infection. WT PIC integrates into chromatin near the sites of uncoating ~1.5 μm from the NE; the N74D/A77V PIC integrates into chromatin associated with lamina-associated domains (LADs) ~0.7 to 0.8 μm from the NE. In addition to the localization of transcriptionally active WT and N74D/A77V proviruses to the nuclear periphery in these studies, localization of WT viral DNA (50, 51) and N74D/A77V DNA (23, 25) to the nuclear periphery was previously reported.

N74D/A77V PICs (without the viral cores) integrate into chromatin near the NE is unclear; however, it is likely that either the mutant PICs enter the nucleus or access the adjacent chromatin while docked at the NPC. Interestingly, the N74D/A77V mutants retain their infectivity in single-cycle assays, suggesting that integration into euchromatin may be essential for maintaining viral fitness in vivo. These results indicating that HIV-1 uncoating and its regulation are linked to the selection of HIV-1 integration sites have important implications for the regulation of HIV-1 transcription and the establishment of a latent reservoir of infected cells, which are major impediments to HIV-1 eradication and cure (41).

In contrast to the current prevailing view that reverse transcription is completed in the cytoplasm after uncoating and is followed by nuclear import of the viral preintegration complex, our results indicate that viral DNA synthesis and the formation of a preintegration complex occurs within an intact (or nearly intact) core and that these steps are completed in the nucleus. Although it is generally thought that reverse transcription is completed in the cytoplasm before nuclear import, some previous studies have provided evidence indicating that viral DNA synthesis is initiated in the cytoplasm but completed in the nucleus (42–45). In support of reverse transcription taking place in an intact or nearly intact viral core, a recent study showed that CA hexamers form a positively charged channel for transport of deoxynucleotide triphosphates into viral cores, providing the substrates necessary for reverse transcription within intact viral cores (46). Our observation that uncoating occurs <1.5 h before integration implies that the viral preintegration complex is exposed to the nuclear environment for <1.5 h before viral DNA integration. We propose that the viral core may remain intact until <1.5 h before integration to ensure completion of reverse transcription, formation of a functional preintegration complex, and potential evasion of innate sensing by cytoplasmic (47) and/or nuclear DNA sensors to suppress cellular immune responses (48).

It is unclear how an intact viral core with a width of ~61 nm (5) can be imported through a NPC with an inner diameter of ~39 nm (49). Lack of CPSF6 recruitment to viral complexes at the NE was correlated with a failure to import viral cores, and uncoating at the NE, suggesting that the intact (or nearly intact) viral cores are, indeed, too large to be imported through nuclear pores in the absence of the CA-CPSF6 interaction at the NE. Our results suggest that CPSF6 recruitment to the viral complexes at the NE results in alteration of the NPC or the viral core structure to facilitate the nuclear import of intact (or nearly intact) viral cores in HeLa cells as well as T cells, which are the major target cells for HIV-1 infection in patients. Bejarano and colleagues recently proposed that CPSF6 binds to CA multimers and facilitates nuclear import of viral complexes in primary monocyte-derived macrophages (21) but did not determine if the nuclear viral complexes were composed of intact viral cores, partially uncoated viral cores, or CA-derived subviral structures. Although the structure of the viral complexes at the NE or after nuclear import is not known, they contain most, if not all, of the viral core-associated CA and must retain CA hexamers since they associate with CPSF6 and rapidly disassemble upon PF74 treatment.

The N74D/A77V viral complexes uncoat at the NE with the same kinetics as WT nuclear viral complexes (~10 hpi), suggesting that the molecular trigger for uncoating is intrinsic to the viral complex and independent of its intracellular location. Inhibiting reverse transcription prevented uncoating even though it is completed ~2 to 3 h earlier, suggesting that uncoating is not initiated upon completion of reverse transcription but requires the completion of viral DNA synthesis.

Overall, our studies help resolve a long-standing question in HIV-1 replication regarding the location and timing of uncoating and fundamentally change our understanding of HIV-1 postentry replication events. We propose that the viral core may remain intact or nearly intact until just before integration to maintain high concentrations of reverse transcriptase and integrase near the viral nucleic acid to complete DNA synthesis and ensure assembly of an integration-competent viral complex. Additionally, maintaining an intact or nearly intact viral core may facilitate evasion of innate sensing by DNA sensors to suppress cellular immune responses and ensure integration into the preferred sites of integration located in gene-rich euchromatin regions.

Materials and Methods

Experimental details and methods can be found in the *SI Appendix*, including sources of cell lines and procedures for their maintenance, generation of HeLa-Bgl, HeLa-Bgl:Tat-Rev, HeLa:mRuby-CPSF6, and CEM-SS-mRuby-Lamin B cell lines, construction of lentiviral vectors pHGFP-GFPCA-BglSL, pHGFP-BglSL, pHGFP(N74D)-GFPCA-BglSL, pHGFP(N74D)-BglSL, pHGFP(A77V)-GFPCA-BglSL, pHGFP(A77V)-BglSL, and procedures for virus production and infection. Details of microscopy and image processing, live-cell imaging and image analysis using custom written MATLAB programs, and fixed-cell imaging and image analysis are described in the *SI Appendix*. Methods for single virion analysis, in vitro analysis of intact virions and viral cores, transmission electron

microscope analysis of virus pellets, fractionation of viral cores using sucrose gradients, and data analysis and statistics are also described in the *SI Appendix*.

Data Availability Statement. All data generated in this study are included in the paper and *SI Appendix*.

ACKNOWLEDGMENTS. We thank John Coffin, Eric Freed, and Tom Misteli for valuable discussions and suggestions during paper preparation. This work was supported, in part, by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research, by Intramural AIDS Targeted Antiviral Program grant funding (to V.K.P. and to W.-S.H.) under Contract HHSN26120080001E.

1. A. T. Gres *et al*., STRUCTURAL VIROLOGY. X-ray crystal structures of native HIV-1 capsid protein reveal conformational variability. *Science* **349**, 99–103 (2015).
2. S. Mattei, B. Glass, W. J. Hagen, H. G. Kräusslich, J. A. Briggs, The structure and flexibility of conical HIV-1 capsids determined within intact virions. *Science* **354**, 1434–1437 (2016).
3. G. Zhao *et al*., Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature* **497**, 643–646 (2013).
4. M. Aboud, R. Shoor, S. Salzberg, Adsorption, penetration, and uncoating of murine leukemia virus studied by using its reverse transcriptase. *J. Virol.* **30**, 32–37 (1979).
5. J. A. Briggs, T. Wilk, R. Welker, H. G. Kräusslich, S. D. Fuller, Structural organization of authentic, mature HIV-1 virions and cores. *EMBO J.* **22**, 1707–1715 (2003).
6. E. M. Campbell, T. J. Hope, HIV-1 capsid: The multifaceted key player in HIV-1 infection. *Nat. Rev. Microbiol.* **13**, 471–483 (2015).
7. J. I. Mamede, G. C. Cianci, M. R. Anderson, T. J. Hope, Early cytoplasmic uncoating is associated with infectivity of HIV-1. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E7169–E7178 (2017).
8. Z. Lukic, A. Dharan, T. Fricke, F. Diaz-Griffero, E. M. Campbell, HIV-1 uncoating is facilitated by dynein and kinesin 1. *J. Virol.* **88**, 13613–13625 (2014).
9. A. E. Hulme, O. Perez, T. J. Hope, Complementary assays reveal a relationship between HIV-1 uncoating and reverse transcription. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 9975–9980 (2011).
10. Y. Yang, T. Fricke, F. Diaz-Griffero, Inhibition of reverse transcriptase activity increases stability of the HIV-1 core. *J. Virol.* **87**, 683–687 (2013).
11. O. Cosnefroy, P. J. Murray, K. N. Bishop, HIV-1 capsid uncoating initiates after the first strand transfer of reverse transcription. *Retrovirology* **13**, 58 (2016).
12. A. C. Francis, M. Marin, J. Shi, C. Aiken, G. B. Melikyan, Time-resolved imaging of single HIV-1 uncoating in vitro and in living cells. *PLoS Pathog.* **12**, e1005709 (2016).
13. J. Rasaiyaah *et al*., HIV-1 evades innate immune recognition through specific cofactor recruitment. *Nature* **503**, 402–405 (2013).
14. R. C. Burdick *et al*., Dynamics and regulation of nuclear import and nuclear movements of HIV-1 complexes. *PLoS Pathog.* **13**, e1006570 (2017).
15. A. C. Francis, G. B. Melikyan, Single HIV-1 imaging reveals progression of infection through CA-dependent steps of docking at the nuclear pore, uncoating, and nuclear transport. *Cell Host Microbe* **23**, 536–548.e6 (2018).
16. J. Fernandez *et al*., Transportin-1 binds to the HIV-1 capsid via a nuclear localization signal and triggers uncoating. *Nat. Microbiol.* **4**, 1840–1850 (2019).
17. N. J. Arhel *et al*., HIV-1 DNA Flap formation promotes uncoating of the pre-integration complex at the nuclear pore. *EMBO J.* **26**, 3025–3037 (2007).
18. R. C. Burdick, W. S. Hu, V. K. Pathak, Nuclear import of APOBEC3F-labeled HIV-1 preintegration complexes. *Proc. Natl. Acad. Sci. U.S.A.* **110**, E4780–E4789 (2013).
19. A. E. Hulme, Z. Kelley, D. Foley, T. J. Hope, Complementary assays reveal a low level of CA associated with viral complexes in the nuclei of HIV-1-infected cells. *J. Virol.* **89**, 5350–5361 (2015).
20. K. Peng *et al*., Quantitative microscopy of functional HIV post-entry complexes reveals association of replication with the viral capsid. *eLife* **3**, e04114 (2014).
21. D. A. Bejarano *et al*., HIV-1 nuclear import in macrophages is regulated by CPSF6-capsid interactions at the nuclear pore complex. *eLife* **8**, e41800 (2019).
22. L. Zhou *et al*., Transportin 3 promotes a nuclear maturation step required for efficient HIV-1 integration. *PLoS Pathog.* **7**, e1002194 (2011).
23. V. Achuthan *et al*., Capsid-CPSF6 interaction licenses nuclear HIV-1 trafficking to sites of viral DNA integration. *Cell Host Microbe* **24**, 392–404.e8 (2018).
24. G. A. Sowd *et al*., A critical role for alternative polyadenylation factor CPSF6 in targeting HIV-1 integration to transcriptionally active chromatin. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E1054–E1063 (2016).
25. C. R. Chin *et al*., Direct visualization of HIV-1 replication intermediates shows that capsid and CPSF6 modulate HIV-1 intra-nuclear invasion and integration. *Cell Rep.* **13**, 1717–1731 (2015).
26. E. M. Campbell, O. Perez, J. L. Anderson, T. J. Hope, Visualization of a proteasome-independent intermediate during restriction of HIV-1 by rhesus TRIM5alpha. *J. Cell Biol.* **180**, 549–561 (2008).
27. W. Hübner *et al*., Sequence of human immunodeficiency virus type 1 (HIV-1) Gag localization and oligomerization monitored with live confocal imaging of a replication-competent, fluorescently tagged HIV-1. *J. Virol.* **81**, 12596–12607 (2007).
28. B. Müller *et al*., Construction and characterization of a fluorescently labeled infectious human immunodeficiency virus type 1 derivative. *J. Virol.* **78**, 10803–10813 (2004).
29. J. Chen *et al*., High efficiency of HIV-1 genomic RNA packaging and heterozygote formation revealed by single virion analysis. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 13535–13540 (2009).
30. Y. Zhu, G. Z. Wang, O. Cingöz, S. P. Goff, NP220 mediates silencing of unintegrated retroviral DNA. *Nature* **564**, 278–282 (2018).
31. E. H. Finn *et al*., Extensive heterogeneity and intrinsic variation in spatial genome organization. *Cell* **176**, 1502–1515.e10 (2019).
32. Z. Ambrose, C. Aiken, HIV-1 uncoating: Connection to nuclear entry and regulation by host proteins. *Virology* **454-455**, 371–379 (2014).
33. C. L. Márquez *et al*., Kinetics of HIV-1 capsid uncoating revealed by single-molecule analysis. *eLife* **7**, e34772 (2018).
34. K. Lee *et al*., Flexible use of nuclear import pathways by HIV-1. *Cell Host Microbe* **7**, 221–233 (2010).
35. T. Schaller *et al*., HIV-1 capsid-cyclophilin interactions determine nuclear import pathway, integration targeting and replication efficiency. *PLoS Pathog.* **7**, e1002439 (2011).
36. A. Saito *et al*., Capsid-CPSF6 interaction is dispensable for HIV-1 replication in primary cells but is selected during virus passage in vivo. *J. Virol.* **90**, 6918–6935 (2016).
37. H. Xu *et al*., Evidence for biphasic uncoating during HIV-1 infection from a novel imaging assay. *Retrovirology* **10**, 70 (2013).
38. Y. Yang, J. Luban, F. Diaz-Griffero, The fate of HIV-1 capsid: A biochemical assay for HIV-1 uncoating. *Methods Mol. Biol.* **1087**, 29–36 (2014).
39. K. E. Ocwieja *et al*., HIV integration targeting: A pathway involving transportin-3 and the nuclear pore protein RanBP2. *PLoS Pathog.* **7**, e1001313 (2011).
40. A. Zhyvoloup *et al*., Digoxin reveals a functional connection between HIV-1 integration preference and T-cell activation. *PLoS Pathog.* **13**, e1006460 (2017).
41. M. Lusic, R. F. Siliciano, Nuclear landscape of HIV-1 infection and integration. *Nat. Rev. Microbiol.* **15**, 69–82 (2017).
42. P. Barbosa, P. Charneau, N. Dumey, F. Clavel, Kinetic analysis of HIV-1 early replicative steps in a coculture system. *AIDS Res. Hum. Retroviruses* **10**, 53–59 (1994).
43. Y. M. Lee, J. M. Coffin, Relationship of avian retrovirus DNA synthesis to integration in vitro. *Mol. Cell. Biol.* **11**, 1419–1430 (1991).
44. M. I. Bukrinsky *et al*., Association of integrase, matrix, and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6125–6129 (1993).
45. A. E. Galvis, H. E. Fisher, T. Nitta, H. Fan, D. Camerini, Impairment of HIV-1 cDNA synthesis by DBR1 knockdown. *J. Virol.* **88**, 7054–7069 (2014).
46. D. A. Jacques *et al*., HIV-1 uses dynamic capsid pores to import nucleotides and fuel encapsidated DNA synthesis. *Nature* **536**, 349–353 (2016).
47. K. M. Monroe *et al*., IFI16 DNA sensor is required for death of lymphoid CD4 T cells abortively infected with HIV. *Science* **343**, 428–432 (2014).
48. B. A. Diner, K. K. Lum, I. M. Cristea, The emerging role of nuclear viral DNA sensors. *J. Biol. Chem.* **290**, 26412–26421 (2015).
49. K. E. Knockenhauer, T. U. Schwartz, The nuclear pore complex as a flexible and dynamic gate. *Cell* **164**, 1162–1171 (2016).
50. C. Di Primio *et al*., Single-cell imaging of HIV-1 provirus (SCIP). *Proc. Natl. Acad. Sci. U.S.A.* **110**, 5636–5641 (2013).
51. B. Marini *et al*., Nuclear architecture dictates HIV-1 integration site selection. *Nature* **521**, 227–231 (2015).

PNAS
www.pnas.org

Supplementary Information for

HIV-1 uncoats in the nucleus near sites of integration

Ryan C. Burdick, Chenglei Li, MohamedHusen Munshi, Jonathan Rawson, Kunio Nagashima, Wei-Shau Hu, and Vinay K. Pathak

Vinay K. Pathak
Email: vinay.pathak@nih.gov

This PDF file includes:

Other supplementary materials for this manuscript include the following:

www.pnas.org/cgi/doi/10.1073/pnas.1920631117

Materials and Methods

Cell lines and reagents. HeLa [American Type Culture Collection (ATCC CCL-2), TZM-bl cells [NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH, from Dr. John C. Kappes and Dr. Xiaoyun Wu; Cat#8129; (1)], and Human embryonic kidney 293T cells (ATCC CRL-3216) were maintained as previously described (2, 3). CEM-SS cells (NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH; gift from Dr. Peter L. Nara; Cat#776) and THP-1 cells (ATCC TIB-202) were maintained in Roswell Park Memorial Institute (RPMI) 1640 medium (CellGro, Manassas, VA) supplemented with 10% fetal calf serum (HyClone, Logan, UT) and 1% penicillin-streptomycin (penicillin 50 U/ml and streptomycin 50 µg/ml; Lonza, Walkersville, MD). THP-1 cells were differentiated with phorbol myristyl acetate treatment (Sigma-Aldrich, St. Louis, MO; 100 ng/ml) for 48 hours to generate THP-1-derived macrophages.

To generate the HeLa-Bgl cell line, pLVX-TRE3G-Bgl-mCherry, a lentiviral vector that expresses truncated bacterial protein BglG that is fused to mCherry at the C-terminus and contains a nuclear localization signal (Bgl-mCherry) from a doxycytidine-inducible promoter was first created. Bgl-mCherry, which specifically binds to Bgl RNA stem loops (4), was cloned into lentiviral vector pLVX-TRE3G (Clontech, Mountainview, CA) to generate pLVX-TRE3G-Bgl-mCherry. HeLa cells were transduced with pLVX-TRE3G-Bgl-mCherry and Tet-On 3G, which expresses the transactivator protein, and the transduced cells were maintained in complete media containing puromycin (1 µg/ml; Thermo Fisher Scientific) and G418 Sulfate (200 µg/ml; Sigma-Aldrich); a single cell clone expressing a low level of Bgl-mCherry was obtained by limiting dilution to generate the HeLa-Bgl cell line.

To generate the HeLa-Bgl:Tat-Rev cell line, a murine leukemia virus (MLV)-based vector that expressed HIV-1 Tat and Rev was first created by transferring a cassette that expresses a hygromycin b phosphotransferase gene (5) from an internal ribosomal entry site (IRES) from WH450, a spleen necrosis virus vector, into MLV-based vector pAR2 (6). Next, a cassette containing codon-optimized HIV-1 Tat and Rev separated by an in-frame self-cleaving peptide

from porcine teschovirus 2A [P2A; (7)] was inserted upstream of IRES-Hygro to generate MLV-Tat-P2A-Rev-IRES-Hygro. Infectious virions were produced by co-transfection of MLV-Tat-P2A-Rev-IRES-Hygro along with MLV *gag-pol* expression plasmid pLGPS (8) and amphotropic MLV envelope expression plasmid pSV-A-MLV-*env* [NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: from Dr. Nathaniel Landau and Dr. Dan Littman; Cat#1065 (9)]. HeLa-Bgl cells were transduced with MLV-Tat-P2A-Rev-IRES-Hygro to generate HeLa-Bgl:Tat-Rev cell line, which expresses Bgl-mCherry under the control of a doxycycline-inducible promoter and constitutively expresses HIV-1 Tat and Rev. The HeLa-Bgl:Tat-Rev cells were maintained in complete media containing hygromycin (200 µg/µl; Thermo Fisher Scientific), G418 Sulfate (200 µg/ml), and puromycin (1µg/ml). Doxycycline (Sigma-Aldrich; 1 µg/ml) was added for 24 hrs prior to imaging to induce Bgl-mCherry expression.

To generate the HeLa:mRuby-CPSF6 cell line, HeLa cells constitutively expressing an shRNA targeting CPSF6 were first generated by transduction with a lentiviral vector (pLKO.1 hygro; gift from Bob Weinberg; Addgene plasmid #24150), which was modified to express the shRNA targeting CPSF6 (5'-AGACCGTCATGACGATTATTA-3'). A control cell line that expresses a non-targeting control shRNA from the human U6 promoter was also constructed. The transduced cells were selected for resistance to hygromycin (200 µg/ml). Next, pLVX-EF1-mRuby-CPSF6-P2A-Puro, a lentiviral bicistronic vector that expresses mRuby-CPSF6 that is resistant to the CPSF6 shRNA under the control of the human EF1 promoter was created. Briefly, the IRES-Puro cassette in the lentiviral vector pLVX-IRES-Puro (Cat# 632183, Clontech) was replaced with mRuby, which was amplified by PCR from mRuby-LaminB1-10 (Addgene plasmid #55869), and the CMV promoter was replaced with the human EF1 promoter. Next, an shRNA-resistant CPSF6 (GenBank NM_007007.2; synonymous mutations in the nucleotide sequence targeted by CPSF6 shRNA were introduced) and an in-frame P2A fused to Puro were synthesized as multiple gBlocks (Integrated DNA Technologies, Inc., Coralville, Iowa). These fragments were assembled downstream of mRuby by Gibson Assembly (New England Biolabs, Inc., Ipswich, MA), generating the lentiviral vector pLVX-EF1-mRuby-CPSF6-P2A-Puro. The

HeLa cells stably expressing the CPSF6 shRNA were transduced with the lentiviral vector pLVX-EF1-mRuby-CPSF6-P2A-Puro and selected for puromycin resistance (1 µg/ml). The mRuby-CPSF6 fusion protein was expressed in HeLa cells stably expressing CPSF6 shRNA to reduce the amount of endogenous unlabeled CPSF6 and increase the mRuby-CPSF6 labeling efficiency. To validate the cell lines, endogenous CPSF6 and mRuby-CPSF6 were detected by SDS-PAGE and western blot analysis using an anti-CPSF6 antibody (Sigma-Aldrich; Cat#HPA039973) followed by goat anti-rabbit antibody (IRDye-800CW; LI-COR, Lincoln, NE); HSP90 was used as a loading control and detected using an anti-HSP90 antibody (Santa Cruz Biotechnology, Inc., Dallas, TX; Cat#sc-69703) followed by goat anti-mouse antibody (IRDye-680RD; LI-COR). Western blots were imaged and quantitated using the Odyssey infrared imaging system (LI-COR).

To generate the CEM-SS-mRuby-LaminB cell line, CEM-SS cells were transduced with the lentiviral vector pLVX-CMV-mRuby-LaminB-P2A-Puro, a bicistronic lentiviral vector that expresses a mRuby-LaminB fusion protein and Puro under the control of the CMV promoter. Briefly, the mRuby-CPSF6 cassette in pLVX-EF1-mRuby-CPSF6-P2A-Puro was replaced with mRuby-LaminB, which was amplified by PCR from mRuby-LaminB1-10 (Addgene plasmid #55869) and the EF1 promoter was replaced with the CMV promoter. Transduced cells were selected for resistance to puromycin (1 µg/ml). All cells were maintained in humidified 37º C incubators with 5% CO_2. Nevirapine (NVP) and raltegravir (RAL) were obtained through the NIH AIDS Reagent Program and were used at final concentrations of 5 µM and 10 µM, respectively. PF-3450074 (PF74; Sigma) was used at a final concentration of 10 µM.

Lentiviral vectors, protein expression vectors, and virus production. The HIV-1 based vector pHGFP-BglSL was generated by inserting a cassette containing 18 RNA stem loops [BglSL; (4)], which are specifically recognized by the bacterial Bgl protein[21], into pHGFP. The RNA stem loops were separated by short (~10 nucleotide) randomized linkers and were inserted into the Vpr-Vif region of the HIV-1 based vector pHGFP. pHGFP is an HIV-1 vector derived from

pHL (10) that contains a *gfp* reporter gene in place of *nef* and does not express *env*. The resulting HIV-1 vector expresses a *gfp* reporter gene and an RNA that contains 18 BglSL, but does not express Vif, Vpr, and Env. To generate the HIV-1 based vector, pHGFP-GFPCA-BglSL, a Y132I mutation in the protease cleavage site between GFP and CA, which prevents cleavage (11) was introduced into HIV Gag-iGFP [NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: gift from Dr. Benjamin Chen; Cat#12457; (12)] by site-directed mutagenesis. A cassette containing MA-GFP-CA (containing the Y132I mutation) was inserted in place of MA-CA in pHGFP-BglSL, generating a HIV-1 vector that is similar to pHGFP-BglSL, except that a GFP-CA fusion protein is generated after proteolytic processing. CA mutants N74D and A77V were generated by site-directed mutagenesis of pHGFP and cassettes containing CA with the mutations were inserted into pHGFP-GFPCA-BglSL and pHGFP-BglSL, generating pHGFP(N74D)-GFPCA-BglSL, pHGFP(N74D)-BglSL, pHGFP(A77V)-GFPCA-BglSL, and pHGFP(A77V)-BglSL, respectively. Infectious virions that were labeled with GFP-CA were prepared by co-transfection of 293T cells with pHGFP-GFPCA-BglSL and pHGFP-BglSL (or similar vectors containing CA mutation), which expresses wild-type gag-pol, at a 1:15 plasmid ratio, and pHCMV-G (13), which expresses the G glycoprotein of vesicular stomatitis virus (VSV-G). To generate infectious virions that are labeled with GFP-CA, but do not contain BglSL and express Vif and Vpr, a cassette containing MA-GFP-CA (containing the Y132I mutation) was inserted in place of MA-CA in pHGFP, generating a vector (pHGFP-GFPCA) that expresses a GFP-CA fusion protein after proteolytic processing that also expresses Vif and Vpr. Infectious virions that were labeled with GFP-CA and produced in the presence of Vif and Vpr were prepared by co-transfection of 293T cells with pHGFP-GFPCA and pHGFP at a 1:10 plasmid ratio, and pHCMV-G. Infectious virions containing only wild-type *gag-pol* were produced by co-transfection of 293T cells with pHGFP and pHCMV-G.

A plasmid expressing APOBEC3F (A3F) fused to red-red vine tomato (RRvT) fluorescent protein was generated by replacing YFP in A3F-YFP (3) with RRvT, which was amplified by PCR from pBad-HisB-RRvT [addgene plasmid #87364; (14)]. To generate infectious virions that were

labeled with GFP-CA and A3F-RRvT, a plasmid expressing an A3F-RRvT (1.25 µg) was co-transfected with pHGFP-GFPCA-BglSL, pHGFP-BglSL, and pHCMV-G. To generate single-labeled infectious virions, plasmids expressing either A3F-YFP, A3F-mNeonGreen (mNG; mNeonGreen was amplified by PCR from pmNeonGreen-N1 (Allele Biotechnology, San Diego, CA) and used to replace YFP in A3F-YFP), Vpr-Integrase-YFP (2) , or Vpr-Integrase-superfolder(sf)GFP [kindly provided by Dr. Greg Melikyan; (15)] were co-transfected with pHGFP-BglSL or pHGFPΔVifΔVpr, a pHGFP-derived vector in which the *vif* and *vpr* genes were deleted, and pHCMV-G.

To generate cell lines, the lentiviral vectors used to create the cell lines (*e.g.* pLVX-TRE3G-Bgl-mCherry) were co-transfected with pC-Help (16), which is an HIV-1 helper construct that expresses *gag-pol* but lacks several cis-acting elements needed for viral replication, and pHCMV-G. Supernatants from transfected 293T cells were filtered and the HIV-1 particles were concentrated by ultracentrifugation (100,000 × g) for 1.5 hrs at 4 ºC through a 20% sucrose cushion (wt/vol) in 1X phosphate buffered saline (PBS).

Virus infection. HeLa-based cell lines were seeded in ibiTreated µ-slides (30,000 cells/well; Ibidi, Gräfelfing, Germany) one day prior to infection. The CEM-SS:mRuby-LaminB cell line was seeded into µ-slides (100,000 cells/well) that were pretreated with poly-L-lysine (Sigma-Aldrich; Cat#P8920). THP-1 cells were seeded into µ-slides (50,000 cells/well) and treated with PMA (100 ng/ml) for 48-h prior to spinoculation. Cells were infected with viruses via spinoculation at 16ºC (17), which permitted virion binding to cell membranes but prevented virion endocytosis. For live-cell imaging experiments, infections with the GFP-CA-labeled virions were performed at a multiplicity of infection (MOI) of ≤0.1 GFP-expressing proviruses/cell and in the presence of aphidicolin (2 µg/ml) to prevent cell division during the long movies. After spinoculation, the media was replaced with prewarmed media to allow internalization of the virus (defined as the 0-h time point) and thereafter incubated at 37ºC. Time-lapse images of the infected cells were acquired by spinning disk confocal microscopy (described below) or the cells were fixed at

various time points post-infection with 4.0% (wt/vol) paraformaldehyde (PFA). The half maximal inhibitory concentrations (IC50) for NVP, PF74, and RAL were determined for GFP-CA-labeled and unlabeled virions in TZM-bl cells. Virus infectivity was determined by flow cytometry (LSRFortessa; BD Biosciences, San Jose, CA) 24-48 hpi or after infection of TZM-bl cells and measurement of luciferase activity using the britelite plus Reporter Gene Assay System (Perkin Elmer) 48 hpi.

Microscopy and image processing. Confocal images were acquired using Nikon Eclipse T*i*-E microscope equipped with a Yokogawa CSU-X1 spinning disk unit with a Plan-Apochromat 100x N.A. 1.49 oil objective, using 405-nm (DAPI/AF405), 488-nm (GFP), 514-nm (YFP), 561-nm (mRuby/RRvT), 594-nm (mCherry), and 647-nm (Cy5) lasers for illumination. Images were captured using a TwinCam system (Cairn, Faversham, UK) equipped with a 565-nm splitter and two iXon Ultra (Andor, Belfast, UK) cameras. A Tokai Hit microscope stage top incubator (Tokai, Japan) was used for all live-cell imaging experiments. To visualize the nuclear GFP-CA puncta, HIV-1 transcription sites, and *gfp* reporter expression, z-stacks of the infected cells (13 slices at 0.4 μm interval) were acquired every 20 min for 24 hrs starting at ~4 hrs after infection. Time-lapse images of the infected cells were examined using Nikon Elements or ImageJ (18). For display, a pixel-averaging filter was applied to the images and the contrast was adjusted; unmodified images were used for intensity analyses.

Live-cell imaging and image analysis. GFP-CA signals and HIV-1 TS were manually identified upon extensive analysis of the entire movie. GFP-CA intensities for individual particles were determined using a custom-written MATLAB program. Briefly, the 3D position of the GFP-CA signal in the z-stack was manually determined and then the local background-subtracted pixel intensities at location the GFP-CA signal were determined. Background intensities were determined by selecting random positions in the nuclei of infected cells and extracting pixel intensities. The GFP-CA-labeled nuclear complexes in CEM-SS:mRuby-LaminB cells were identified manually from z-stacks acquired between 6 and 10 hpi; GFP-CA intensities for these

complexes were determined as described above. Of the 116 HIV-1 TS observed (59 in HeLa-Bgl cells and 57 in HeLa-Bgl:Tat-Rev cells), 81 TS were the only TS detected in the cells. When cells had >1 TS, only the timing of the 1st detectable TS was considered. To visualize the nuclear import of GFP-CA or A3F-mNG particles, similar time-lapse images of the infected cells were acquired as described above, except the movies were started ~10 min after infection.

Distances between the GFP-CA-labeled viral complexes and associated HIV-1 TS were determined using a custom written MATLAB program. Briefly, the image of the nucleus in the last frame GFP-CA signal was detected prior to disappearance and the image of the nucleus in the first frame in which the associated HIV-1 TS was detected were aligned to account for cell movement in between GFP-CA loss and detection of HIV-1 TS. Next, the 3D positions of the GFP-CA particle and HIV-1 TS were determined using 3D gaussian fitting, which were then used to calculate distances. A similar method was used to determine the distance HIV-1 TS moved in ~1.5 hrs.

The distance between the HIV-1 TS and the nucleus boundary was determined using a custom-written MATLAB program. Bgl-mCherry contains a nuclear localization signal; thus, most of the Bgl-mCherry localized inside the nucleus and very low amounts of Bgl-mCherry localized with the NE. We identified nuclear pore complexes (NPCs) in HeLa cells expressing Bgl-mCherry by immunostaining with anti-NPC antibody Mab414 (Abcam, Cambridge, UK; Cat#ab24609) followed by AlexaFluor488-labeled secondary antibody (Thermo Fisher Scientific, Waltham, MA; Cat#A-11001) and compared the Bgl-mCherry signals localized to the NPCs and inside the nucleus (Fig. S3E-F). As expected, the Bgl-mCherry signals inside the nucleus were significantly higher than at the NPCs (414 and 145, respectively; Fig. S3F), indicating that the reduction in the Bgl-mCherry signal can be used to visualize the nucleus boundary. The position of the HIV-1 TS at the time of detection was determined using 3D gaussian fitting. Next, the nearest edge of the Bgl-mCherry signal to the HIV-1 TS was determined and then the 3D distance between the nucleus boundary and the HIV-1 TS was determined.

Fixed-cell imaging and image analysis. A custom-written MATLAB program was used to determine the colocalization of A3F-RRvT and GFP-CA signals inside the nuclei of infected cells. First, the A3F-RRvT signals were detected using Localize (19). Colocalization of the fluorescently-labeled virus particles with a mask of the nucleus interior (based on immunofluorescence staining using an anti-Lamin A/C antibody [Sigma-Aldrich; Cat#L1293 or Thermo Fisher Scientific; Cat#MA3-1000] followed by detection using a AlexaFluor405-labeled secondary antibody [Thermo Fisher Scientific, Cat#A-31553 or Cat#A-31556) was determined using a custom-written MATLAB program as previously described (2). Next, to determine background signals in the GFP channel, the intensity values in the GFP channel at random positions inside >50 nuclei were determined; the threshold was determined as the mean + 2 SD of the random intensities. The intensity values of the GFP channel at the position of each nuclear A3F-RRvT particle or at random positions in the nuclei were determined; A3F-RRvT particles co-localizing with GFP signals that were above the threshold intensity value were considered positive for GFP-CA. For some experiments, a mask of the NE was also created using the Lamin A/C signal. The percentage of GFP-CA, Integrase-YFP, or A3F-YFP particles that colocalized with the NE mask and the percentage of particles that colocalized with the nucleus mask were determined. To determine the percentage of GFP-CA- or Integrase-sfGFP-labeled viral complexes that were mRuby-CPSF6^{+} and the intensities of those complexes, the GFP-CA- or Integrase-sfGFP-labeled nuclear complexes in HeLa:mRuby-CPSF6 cells were identified manually from z-stacks acquired 6 hpi. Pixel intensities at the locations of the GFP-CA or Integrase-sfGFP particle signals were determined after subtracting the local background. Nuclear background intensities were determined by selecting 85 random positions from the nuclei of 5 cells and extracting pixel intensities in the mRuby channel; the threshold intensity was determined as the mean + 1 SD of the random intensities in the mRuby channel. GFP-CA or Integrase-sfGFP particles co-localizing with mRuby signals that were above the threshold value were considered positive for mRuby-CPSF6.

Single Virion Analysis (SVA). Fluorescently-labeled virus particles were centrifuged onto ibiTreated µ-slides (1,200 x g for 1 hr). CA was detected using an anti-CA antibody (AG3.0; NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: gift from Dr. Jonathan Allan; Cat#4121) followed by a Cy5-labeled secondary antibody (Thermo Fisher Scientific, Cat#A-10524), and then imaged by spinning disk confocal microscopy. The diffraction-limited spots were detected, and their positions were determined in each image using Localize. The positions of the spots were used to determine colocalization; spots were considered colocalized if the centers of the spots were within 3 pixels.

***In vitro* analysis of intact virions and viral cores.** GFP-CA-labeled virus particles were centrifuged onto ibiTreated µ-slides (1,200 × g for 1 hr). Intact virions bound to the slides were lysed by saponin treatment *in vitro* to remove viral membranes and free CA that was not incorporated into viral cores; an equal volume of PBS containing 2X saponin detergent (0.1% final concentration) was added to the wells containing virus particles between the 2nd and 3rd time points. Time-lapse images of the virus particles were acquired every 10 seconds for 5 minutes. A custom-written MATLAB program was used to identify GFP-CA particles and determine their signal intensities. Virus particles that lost ≥33% of the initial GFP-CA signal were considered to lose the viral membrane (typically 90% of particles) and were used for further analysis. Next, the virus particles were separated into three categories; first, some particles (4%) lost all GFP-CA signal at the time of saponin treatment; second, some particles (44%) lost the GFP-CA signal in two discrete steps, losing free GFP-CA that was not incorporated into viral cores at the time of saponin treatment and subsequently losing viral core-associated GFP-CA signal during the 5-min observation time (i.e. uncoated); third, some particles (52%) lost free GFP-CA that was not incorporated into viral cores and retained viral core-associated GFP-CA until the end of the 5-min observation time. The background-subtracted GFP-CA intensities of 4,441 intact virions (avg. 974; before saponin treatment) and 3,681 viral cores (avg. 291; ~1 min after saponin treatment) were determined. The detection limit in cells was determined using intensity values in the GFP channel at 120 random positions (5 random positions/nucleus inside the nuclei of 24 infected

cells). The threshold was determined as the mean + 3 SD of the random intensities and the 4,374 intact virions and 1,641 viral cores with GFP-CA signals above this threshold were considered detectable in cells. Thus, 98.5% of the GFP-CA-labeled intact virions (4,374/4,441 = 98.5%) and 45% of the GFP-CA-labeled viral cores (1,641/3,681 = 45%) would be detectable inside cells. The GFP-CA intensities for the 4,374 intact virions and 1,641 viral cores that would be detectable in cells were used to compare to the GFP-CA intensity of 116 nuclear complexes in Hela cells (59 from HeLa-Bgl and 57 from Hela-Bgl:Tat-Rev) and 223 nuclear complexes in CEM-SS:mRuby-LaminB cells (Fig. 3C).

Transmission Electron Microscope (TEM) analysis of virus pellets. Virus pellets were obtained by ultracentrifugation as described above, washed once in PBS, processed for thin-sectioned EM analysis as previously described (20, 21). Briefly, virus pellets were fixed in 2% (v/v) glutaraldehyde (Tousimis, Rockville, MD) in 0.1 M sodium cacodylate (pH 7.4; Electron Microscopy Sciences [EMS], Fort Washington, PA), followed by Osmium post fixation (1% Osmium tetroxide v/v in same buffer for 1 hr). The pellets were dehydrated in a series of ethanol solutions (e.g., 35%, 50%, 70%, 95%, 100%) and 100% propylene oxide. An infiltration was made in 1:1 mixture of propylene oxide and epoxy resin (EMS) overnight, and the pellets were embedded in pure resin. The resin was cured for 48 hrs in 55ºC and thin sections (60 to 70 nm) were made with an ultra-microtome equipped with a diamond knife and mounted on naked copper grid. The thin sections were stained in uranyl acetate (EMS) and lead citrate (Leica, Bannockburn, IL), stabilized by carbon evaporation. The sections were examined and images were captured by digital camera in Hitachi 7650 TEM (Hitachi, Tokyo, Japan) operated at 75kv (21). Virions exhibiting the mature and immature phenotype were scored manually.

Fractionation of viral cores using sucrose gradients. The fractionation of viral cores using sucrose gradients was performed as previously described, with slight modifications (22). Concentrated virions were prepared as described above and subjected to ultracentrifugation (100,000 × g for 16 hr at 4 ºC) through a layer of 1% Triton X-100 into a linear sucrose density

gradient (10 mls of 10 mM Tris-HCl [pH 7.4], 100 mM NaCL, 1mM EDTA containing 30% to 70% sucrose). Following ultracentrifugation, fractions (1 ml) were collected from the top of the gradient. CA and GFP-CA protein were detected in each fraction by western blot analysis using antibodies against CA (NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH, HIV-1 p24 Gag Monoclonal (#24-3) from Dr. Michael H. Malim; Cat#6458) and GFP (Thermo Fisher Scientific; Cat#A-6455). The amount of CA in the peak fractions (fractions 8 and 9), which represent CA associated with viral cores, was divided by the total amount of CA in the fractions to determine the yield of cores.

Data analysis and statistics. The Welch's unpaired *t*-test and paired *t*-test were used to analyze parametric data. The Mann Whitney *U* test was used to analyze nonparametric data and the Kolmogorov-Smirnov test was used to analyze cumulative frequencies. A Fisher's exact test was used to analyze 2 x 2 contingency tables. All statistical tests were performed in Prism 8 (GraphPad Software, San Diego, CA). *P* values <0.05 were considered significant.

SI References

1. Wei X*, et al.* (2002) Emergence of resistant human immunodeficiency virus type 1 in patients receiving fusion inhibitor (T-20) monotherapy. *Antimicrob Agents Chemother* 46(6):1896-1905.
2. Burdick RC*, et al.* (2017) Dynamics and regulation of nuclear import and nuclear movements of HIV-1 complexes. *PLoS Pathog* 13(8):e1006570.
3. Burdick RC, Hu WS, & Pathak VK (2013) Nuclear import of APOBEC3F-labeled HIV-1 preintegration complexes. *Proc Natl Acad Sci U S A* 110(49):E4780-4789.
4. Chen J*, et al.* (2009) High efficiency of HIV-1 genomic RNA packaging and heterozygote formation revealed by single virion analysis. *Proc Natl Acad Sci U S A* 106(32):13535-13540.
5. Gritz L & Davies J (1983) Plasmid-encoded hygromycin B resistance: the sequence of hygromycin B phosphotransferase gene and its expression in Escherichia coli and Saccharomyces cerevisiae. *Gene* 25(2-3):179-188.
6. Yin PD, Pathak VK, Rowan AE, Teufel RJ, 2nd, & Hu WS (1997) Utilization of nonhomologous minus-strand DNA transfer to generate recombinant retroviruses. *J Virol* 71(3):2487-2494.
7. Kim JH*, et al.* (2011) High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLoS One* 6(4):e18556.
8. Miller AD*, et al.* (1991) Construction and properties of retrovirus packaging cells based on gibbon ape leukemia virus. *J Virol* 65(5):2220-2224.
9. Landau NR, Page KA, & Littman DR (1991) Pseudotyping with human T-cell leukemia virus type I broadens the human immunodeficiency virus host range. *J Virol* 65(1):162-169.

10. Nikolenko GN*, et al.* (2007) Mutations in the connection domain of HIV-1 reverse transcriptase increase 3'-azido-3'-deoxythymidine resistance. *Proc Natl Acad Sci U S A* 104(1):317-322.
11. Lee SK, Harris J, & Swanstrom R (2009) A strongly transdominant mutation in the human immunodeficiency virus type 1 gag gene defines an Achilles heel in the virus life cycle. *J Virol* 83(17):8536-8543.
12. Hubner W*, et al.* (2007) Sequence of human immunodeficiency virus type 1 (HIV-1) Gag localization and oligomerization monitored with live confocal imaging of a replication-competent, fluorescently tagged HIV-1. *J Virol* 81(22):12596-12607.
13. Yee JK*, et al.* (1994) A general method for the generation of high-titer, pantropic retroviral vectors: highly efficient infection of primary hepatocytes. *Proc Natl Acad Sci U S A* 91(20):9564-9568.
14. Wiens MD*, et al.* (2016) A Tandem Green-Red Heterodimeric Fluorescent Protein with High FRET Efficiency. *Chembiochem* 17(24):2361-2367.
15. Francis AC & Melikyan GB (2018) Single HIV-1 Imaging Reveals Progression of Infection through CA-Dependent Steps of Docking at the Nuclear Pore, Uncoating, and Nuclear Transport. *Cell Host Microbe* 23(4):536-548 e536.
16. Mochizuki H, Schwartz JP, Tanaka K, Brady RO, & Reiser J (1998) High-titer human immunodeficiency virus type 1-based vector systems for gene delivery into nondividing cells. *J Virol* 72(11):8873-8883.
17. O'Doherty U, Swiggard WJ, & Malim MH (2000) Human immunodeficiency virus type 1 spinoculation enhances infection through virus binding. *J Virol* 74(21):10074-10080.
18. Schneider CA, Rasband WS, & Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9(7):671-675.
19. Zenklusen D, Larson DR, & Singer RH (2008) Single-RNA counting reveals alternative modes of gene expression in yeast. *Nat Struct Mol Biol* 15(12):1263-1271.
20. Mariner JM, McMahon JB, O'Keefe BR, Nagashima K, & Boyd MR (1998) The HIV-inactivating protein, cyanovirin-N, does not block gp120-mediated virus-to-cell binding. *Biochem Biophys Res Commun* 248(3):841-845.
21. Tobin GJ, Nagashima K, & Gonda MA (1996) Immunologic and Ultrastructural Characterization of HIV Pseudovirions Containing Gag and Env Precursor Proteins Engineered in Insect Cells. *Methods* 10(2):208-218.
22. Forshey BM, von Schwedler U, Sundquist WI, & Aiken C (2002) Formation of a human immunodeficiency virus type 1 core of optimal stability is crucial for viral replication. *J Virol* 76(11):5667-5677.

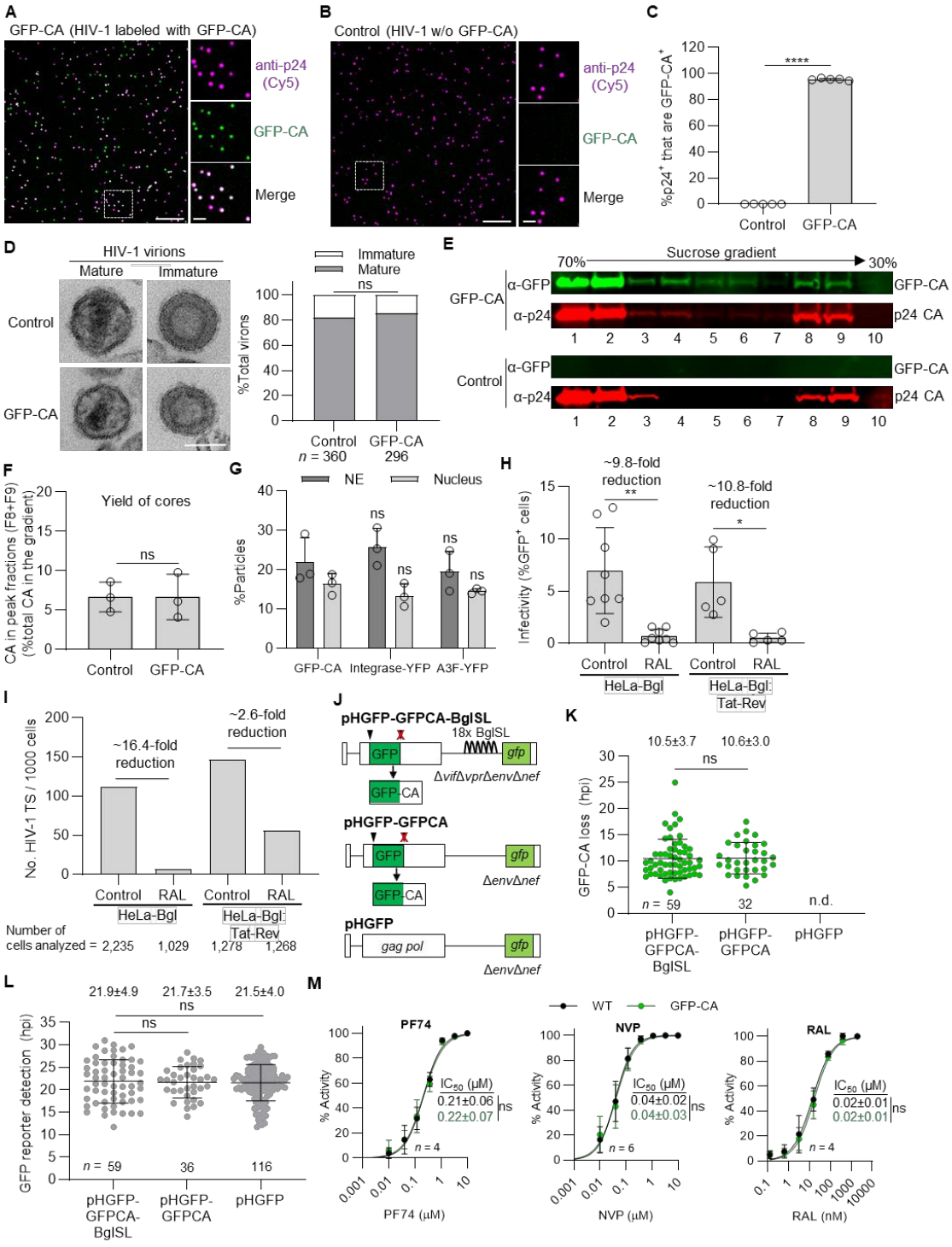

Fig. S1. Characterization of GFP-CA labeled virions and infectivity and detection of TS in HeLa-Bgl and HeLa-Bgl:Tat-Rev cells. (**A-B**) Single virion analysis of HIV-1 virions labeled with and

without GFP-CA. HIV-1 virions produced by co-transfection of 293T cells with pHGFP-GFPCA-BglSL and pHGFP-BglSL at a 1:15 ratio along with VSV-G expressing plasmid (**A**) or produced in the absence of pHGFP-GFPCA-BglSL (**B**) were centrifuged onto a slide, fixed, and immunostained with anti-CA antibody followed by Cy5-labeled secondary antibody. Scale bars, 10 μm; inset scale bars, 2 μm. (**C**) Quantitation of the proportion of p24 CA^+ virions that are GFP-CA^+ showed that 96% of the virions produced in the presence of pHGFP-GFPCA-BglSL were GFP-CA^+, whereas <0.1% of the virions produced in the absence of pHGFP-GFPCA-BglSL were GFP-CA^+. Data are mean ± s.d. from 5 independent experiments; *P* values are from paired *t*-tests. (**D**) Quantitative electron microscopy of HIV-1 virions shows that similar high proportions (>80%) of GFP-CA-labeled and control virions have a capsid core, indicative of a mature morphology. Scale bar, 100 nm. Statistical significance was determined using Fisher's exact test. (**E-F**) Sucrose gradient fractionation of HIV-1 virions with WT and GFP-CA Gag (1:15 ratio) and with only WT Gag (**E**) and the quantitation of the proportion of CA in fractions 8 and 9 of the sucrose gradient, which contain intact viral cores, relative to total CA in the gradient (**F**). The results show that similar proportions of CA associated with intact viral cores were recovered, indicating that GFP-CA labeling of viral cores did not influence their *in vitro* stability. Data are mean ± s.d. from 3 independent experiments; *P* values are from Welch's *t*-tests. (**G**) Comparison of the proportion of viral complexes in infected HeLa cells that stably associated with the NE and entered the nucleus at 6 hpi. The results show that similar proportions of virions labeled with GFP-CA, Integrase-YFP, or A3F-YFP were stably associated with the NE and entered the nucleus, indicating that GFP-CA labeling of viral cores did not significantly affect their association with the NE or their nuclear entry. Data are mean ± s.d. from 3 independent experiments; *P* values are from Welch's *t*-tests. (**H**) Infection of Hela-Bgl and HeLa-Bgl:Tat-Rev cells with GFP-CA-labeled virions at low MOIs resulted in GFP reporter expression in ~7% of the infected cells (MOIs ~0.05 – 0.1). Approximately 5- to 10-fold lower amounts of virions were used to infect HeLa:Tat-Rev cells to achieve a similar proportion of GFP^+ cells. Treatment of target cells with RAL significantly reduced GFP-expressing HeLa-Bgl cells by ~9.8-fold and HeLa-Bgl:Tat-Rev

cells by ~10.8-fold, indicating that integration was required for GFP reporter expression. Data are mean ± s.d. from ≥5 independent experiments; *P* values are from Welch's *t*-tests. (**I**) Number of HIV-1 TS/1,000 cells. Approximately 111 HIV-1 TS were detected per 1,000 infected Hela-Bgl cells and 147 TS were detected per 1,000 infected HeLa-Bgl:Tat-Rev cells. RAL treatment reduced HIV-1 TS in HeLa-Bgl cells ~16.4-fold and in HeLa-Bgl:Tat-Rev cells ~2.6-fold, indicating that the majority of HIV-1 TS represented transcription from integrated viral DNA. RAL treatment reduced HIV-1 TS to a lesser extent in HeLa-Bgl:Tat-Rev cells, indicating that constitutive expression of Tat increased detectable transcription from unintegrated DNA. (**J**) Structures of pHGFP-GFPCA-BglSL, pHGFP-GFPCA and pHGFP. pHGFP-GFPCA-BglSL does not express Vif, Vpr, Env, or Nef and contains 18 copies of stem loops that specifically bind to Bgl protein. pHGFP-GFPCA expresses Vif and Vpr but not Env or Nef and does not contain any Bgl stem loops. pHGFP expresses Vif and Vpr, but not Env or Nef, and does not contain any Bgl stem loops; in addition, it does not express a GFP-CA fusion protein. (**K**) GFP-CA loss occurs with the same kinetics in cells infected with pHGFP-GFPCA-BglSL and pHGFP-GFPCA, indicating that the presence of Bgl stem loops or absence of Vif and Vpr expression does not affect the kinetics of nuclear uncoating. (**L**) The kinetics of GFP reporter detection was not affected by the presence of Bgl stem loops, absence of Vif or Vpr expression, or GFP-CA labeling of viral complexes, indicating that labeling of virions with GFP-CA did not significantly affect the kinetics of GFP reporter expression. For (**K**) and (**L**), lines are mean ± s.d.; *P* values are from Welch's *t*-tests. n.d., not determined. (**M**) Sensitivity of GFP-CA-labeled and unlabeled virions to PF74 (left), NVP (middle), and RAL (right). The half maximal inhibitory concentrations (IC50) were determined for each experiment (n ≥ 4) and statistical significance was determined using Welch's *t*-tests. ****, $P < 0.0001$; **, $P < 0.01$; *, $P < 0.05$; ns, not significant ($P > 0.05$). Open circles in (**C**), (**F-H**) indicate independent experiments.

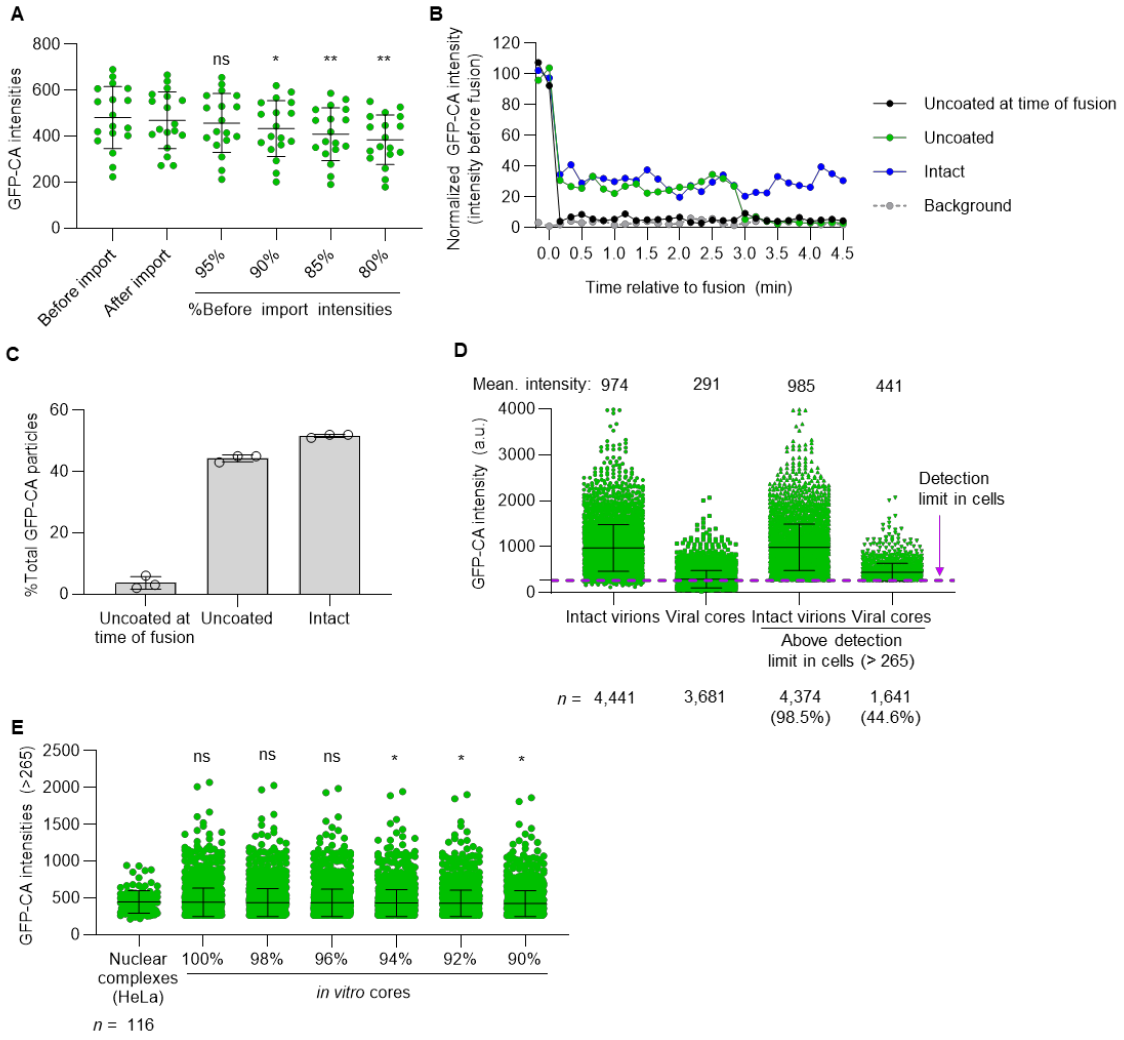

Fig. S2. Modeling GFP-CA loss to determine sensitivity of GFP-CA detection and determination of *in vitro* GFP-CA intensities of intact virions and viral cores. (**A**) Modeling GFP-CA loss during nuclear import. GFP-CA intensities of 18 viral complexes before nuclear import were adjusted to reflect 5-20% GFP-CA loss, particles with GFP-CA intensities below the limit of detection (<265) were removed, and the adjusted GFP-CA intensities were compared to the intensities of the 18 viral complexes after nuclear import. The results indicated that ≥10% loss of GFP-CA would be statistically significant ($P < 0.05$). (**B-C**) Representative graphs of GFP-CA intensities of viral complexes after 0.1% saponin treatment leading to disruption of the viral membrane, mimicking fusion of the viral and host membranes. A small proportion of viral complexes (~4%) lost all GFP-signal upon saponin treatment, indicative of unstable viral cores that disassembled upon fusion;

44% of the viral complexes exhibited a ≥33% loss of GFP-CA signal at the time of fusion, followed by loss of the remaining GFP-CA signal within 4.5 min after saponin treatment, indicative of uncoating; another 52% of the particles lost ≥33% of the GFP-CA signal at the time of fusion, and retained the remaining GFP-CA signal during the 4.5 min observation time. Open circles in (**C**) indicate independent experiments. (**D**) Comparison of the GFP-CA intensities (arbitrary units; a.u.) of intact virions and *in vitro* viral cores. Intact virions and *in vitro* viral cores with GFP-CA intensities below the detection limit in cells (<265 a.u.) were removed (right). The percentage of signals that were above the detection limit in cells is shown. Lines are mean ± s.d. (**E**) Modeling GFP-CA loss of nuclear viral complexes in comparison to *in vitro* cores. GFP-CA intensities of *in vitro* viral cores were adjusted to reflect 2-10% GFP-CA loss, viral cores with GFP-CA intensities below the limit of detection (<265) were removed, and the adjusted GFP-CA intensities were compared to the intensities of 116 GFP-CA labeled nuclear complexes in HeLa-Bgl and HeLa-Bgl:Tat-Rev cells (Fig. 3C). The results indicated that ≥6% loss of GFP-CA would be statistically significant ($P < 0.05$). Lines are mean ± s.d.; P values are from Welch's t-tests. **, $P < 0.01$; *, $P < 0.05$; ns, not significant ($P > 0.05$).

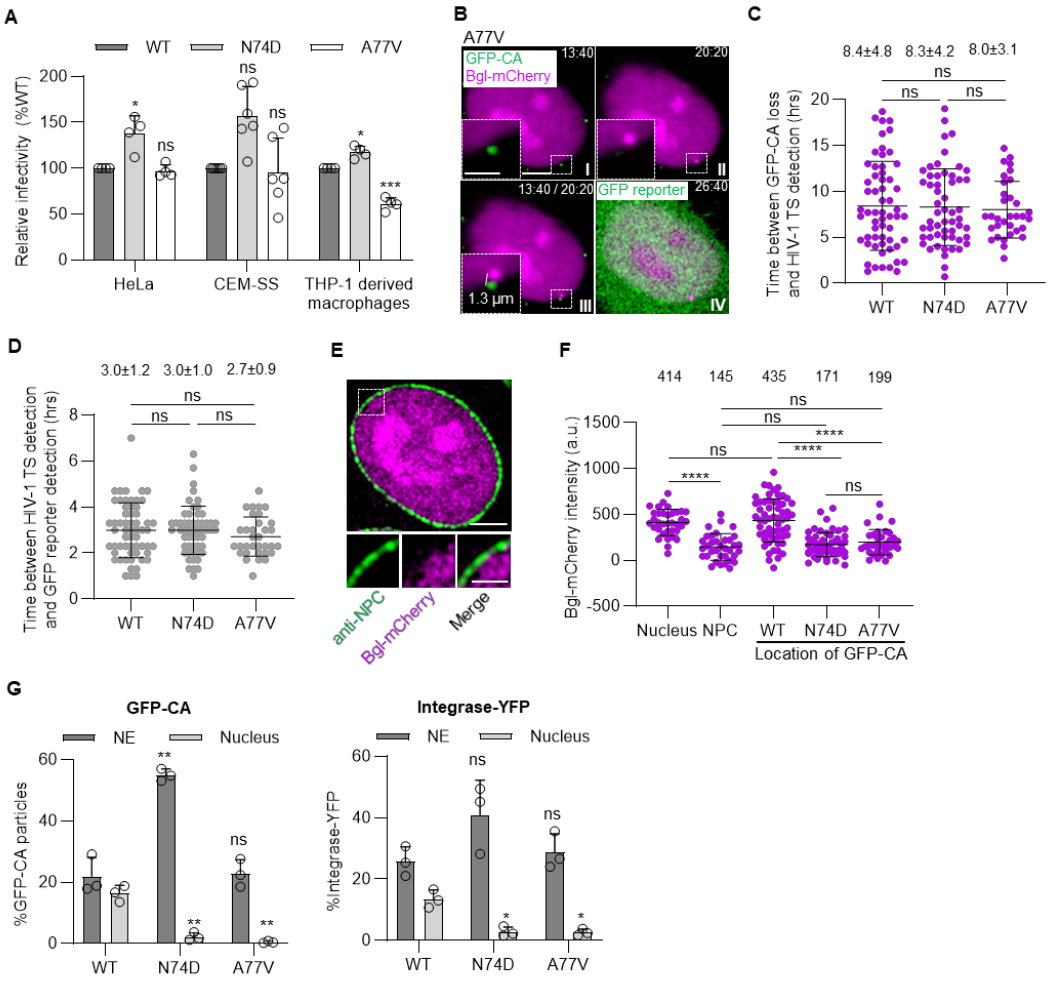

Fig. S3. Characterization of CA mutants N74D and A77V, which disrupt viral core interactions with CPSF6. (**A**) Relative infectivities of WT, N74D, and A77V mutants show that the CA mutations have <2-fold effect on virus infectivity in HeLa cells, CEM-SS cells, and THP-1-derived macrophages. Data are mean ± s.d. from ≥3 independent experiments; *P* values are from Welch's *t*-tests. (**B**) Live-cell microscopy (20 min/frame) images of Hela:Bgl-mCherry cells infected with GFP-CA-labeled virions of CA mutant A77V. A GFP-CA-labeled viral complex uncoated at the edge of the nuclear Bgl-mCherry signal, 13:40 hpi (I) and HIV-1 TS appeared near the site of GFP-CA disappearance 20:20 hpi (II). The HIV-1 TS appeared 1.3 µm from site of GFP-CA disappearance (III). GFP reporter expression was detected 26:40 hpi (IV). (**C**) Comparison of WT GFP-CA-labeled virions with N74D and A77V mutant GFP-labeled virions

indicates that the CA mutations did not influence the average time between GFP-CA loss and HIV-1 TS detection. Data for WT same as Fig. 1F. (**D**) Comparison of WT GFP-CA labeled virions with N74D and A77V mutant GFP-labeled- virions indicates that the CA mutations did not influence the average time between HIV-1 TS detection and GFP reporter detection. Data for WT same as Fig. 1G. (**E**) Immunofluorescence staining of HeLa:Bgl-mCherry cells with an anti-NPC antibody shows that the NE is located at the edge of the nuclear Bgl-mCherry signal. (**F**) Average Bgl-mCherry intensities in the nucleus were significantly higher than the Bgl-mCherry intensities at the NPCs (414 vs. 145 a.u.). The Bgl-mCherry intensities at the locations of WT GFP-CA complexes prior to uncoating were similar to those in the nucleus (435 vs. 414) indicating that the WT GFP-CA viral complexes are localized to the nucleus. The Bgl-mCherry intensities at the locations of the N74D and A77V GFP-CA-labeled viral complexes prior to uncoating were similar to those at the NPCs (171 and 199 vs. 145), indicating that the CA mutant GFP-CA viral complexes are localized at the NPCs. (**G**) Comparison of the proportion of WT, N74D and A77V viral complexes in infected HeLa cells that stably associated with the NE and entered the nucleus at 6 hpi. The results show that viral the N74D and A77V viral complexes stably associated with the NE, but a very low frequency of the mutant viral complexes could be detected in the nucleus compared to WT viral complexes. Both GFP-CA-labeled and Integrase-YFP-labeled N74D/A77V viral complexes were not detected in the nucleus, indicating that the low frequency of nuclear viral complexes was not a result of labeling viral complexes with GFP-CA. Scale bars, 5 μm; inset scale bars, 2 μm. For (**A**) and (**G**), data are mean ± s.d. from ≥3 independent experiments; *P* values are from Welch's *t*-tests. For (**C-D**), and (**F**), lines are mean ± s.d.; *P* values are from Welch's *t*-tests. ****, $P < 0.0001$; ***, $P < 0.001$; **, $P < 0.01$; *, $P < 0.05$; ns, not significant ($P > 0.05$). Open circles in (**A**) and (**G**) indicate independent experiments.

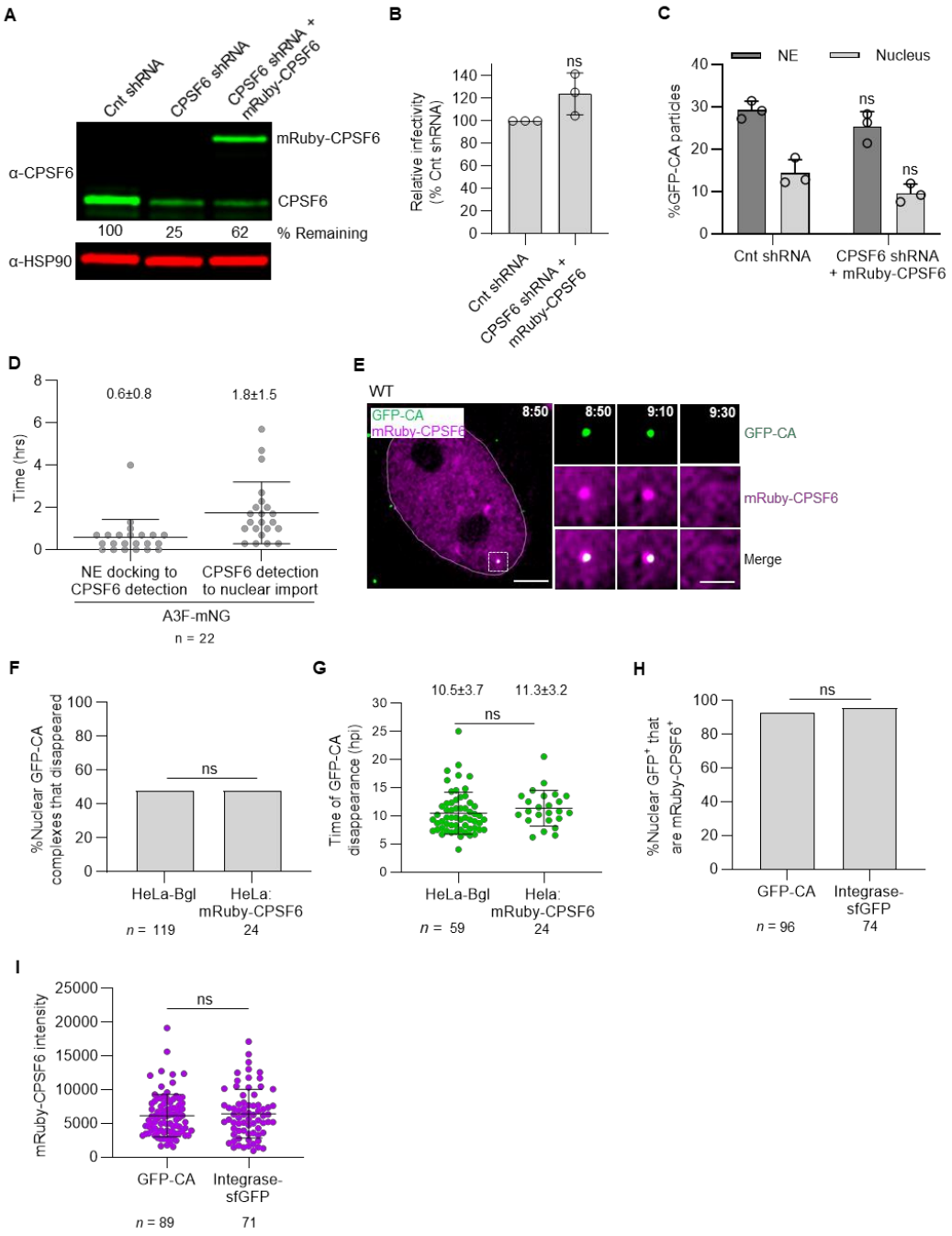

Fig. S4. Characterization of HeLa:mRuby-CPSF6 cells. (**A**) Western blot analysis of HeLa cells stably expressing control shRNA, CSPF6 shRNA, and CPSF6 shRNA + shRNA-resistant mRuby-CPSF6. The results show that stable expression of CPSF6 shRNA reduced endogenous CPSF6 expression to 25% of the control shRNA, and expression of the shRNA-resistant mRuby-CPSF6

restored the CPSF6 levels to 62% of the control shRNA. (**B-C**) CPSF6 shRNA + mRuby-CPSF6 expression did not influence virus infectivity (**B**), the NE-docking efficiency or nuclear import efficiency of GFP-CA-labeled viral complexes (**C**). Data are mean ± s.d. from 3 independent experiments; *P* values are from Welch's *t*-tests. (**D**) Analysis of 22 A3F-mNG-labeled viral complexes showed recruitment of mRuby-CPSF6 to the NE-associated viral complexes 0.6 ± 0.8 hrs after NE docking. In addition, nuclear import of the A3F-mNG + mRuby-CPSF6-labeled complexes occurred 1.8 ± 1.5 hrs after detection of CPSF6 recruitment to the NE. These values are not different than the values in Fig. 4g ($P > 0.05$). Lines are mean ± s.d. (**E**) Representative images of a nuclear GFP-CA-labeled viral complex that was labeled with mRuby-CPSF6. The images show simultaneous loss of both GFP-CA and mRuby-CPSF6 at the time of nuclear uncoating (9:30 hpi). Scale bar, 5 μm; inset scale bar, 2 μm. (**F-G**) CPSF6 shRNA + mRuby-CPSF6 expression did not influence the proportion of nuclear GFP-CA-labeled viral complexes that uncoated (**F**) or the time of GFP-CA disappearance after infection (**G**). (**H-I**) A comparison of the percentage of GFP-CA- or Integrase-sfGFP-labeled nuclear viral complexes that had detectable levels of mRuby-CPSF6 (**H**) and their associated mRuby-CPSF6 intensities (**I**) indicating that the levels of mRuby-CPSF6 associated with nuclear viral complexes are not altered with GFP-CA labeling. For (**F**) and (**H**), *P* values are from Fisher's exact test. For (**D**) and (**I**), lines are mean ± s.d.; *P* values are from Welch's *t*-tests. ns, not significant ($P > 0.05$). Open circles in (**B**) and (**C**) indicate independent experiments.

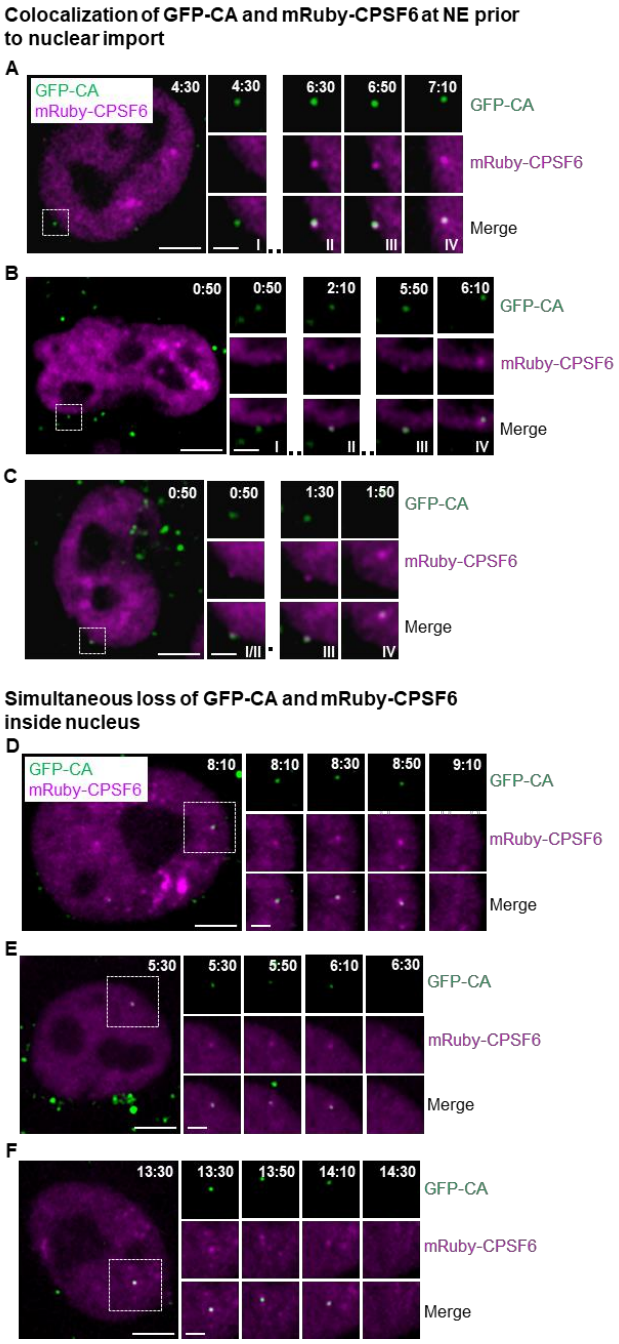

Fig. S5. Additional examples showing the recruitment of mRuby-CPSF6 to GFP-CA-labeled viral complexes at the NE prior to nuclear import and simultaneous disappearance of mRuby-CPSF6 and GFP-CA in the nucleus. (**A-C**) Representative images taken from movies showing mRuby-CPSF6 recruitment to GFP-CA-labeled viral complex. The first frame the GFP-CA-labeled viral complex appeared at the NE (I), first frame of mRuby-CPSF6 detection (II), frame immediately

prior to (III) and after (IV) nuclear import are shown. (**D-F**) Representative images taken from movies showing simultaneous loss of mRuby-CPSF6 and GFP-CA in the nucleus. The three frames prior to and one frame after GFP-CA disappearance are shown. Time scale, hours:minutes post-infection. Scale bar, 5 μm; inset scale bar, 2 μm.

Supplementary Movie Legends

Movie S1. Uncoating of an infectious HIV-1 complex. HeLa-Bgl cells were infected with GFP-CA-labeled virions. Time-lapse images of the cells were acquired every 20 min using spinning disk confocal microscopy. A nuclear GFP-CA-labeled viral complex uncoated and lost the GFP-CA signal 7:10 hpi, a HIV-1 transcription site appeared near the site of GFP-CA disappearance 21:50 hpi, and *gfp* reporter was detected 23:50 hpi. For display, the z-slice closest to the viral complex was extracted from the z-stack for each time point. Time scale, hours:minutes post-infection; scale bar, 5 μm.

Movie S2. Uncoating of an infectious HIV-1 complex in cells expressing exogenous Tat and Rev. HeLa-Bgl:Tat-Rev cells were infected with GFP-CA-labeled virions. Time-lapse images of the cells were acquired every 20 min using spinning disk confocal microscopy. A nuclear GFP-CA-labeled viral complex uncoated and lost the GFP-CA signal 6:20 hpi, a HIV-1 transcription site appeared near the site of GFP-CA disappearance 7:20 hpi, and *gfp* reporter was detected 11:40 hpi. For display, the z-slice closest to the viral complex was extracted from the z-stack for each time point. Time scale, hours:minutes post-infection; scale bar, 5 μm.

Movie S3. Disappearance of a nuclear GFP-CA-labeled viral complex <10 min after PF74 treatment. HeLa-Bgl cells were infected with GFP-CA-labeled virions. Starting at 4 hpi, time-

lapse images of the cells were acquired every 2 min for 1 hr using spinning disk confocal microscopy. After the second frame, the movie was paused, cells were treated with 10 μM PF74, and the imaging was resumed for the remaining time. For display, the z-slice closest to the viral complex was extracted from the z-stack for each time point. Time scale, hours:minutes relative to PF74 treatment; scale bar, 2 μm.

Movie S4. Uncoating of an infectious HIV-1 complex containing the N74D CA mutation. HeLa-Bgl cells were infected with GFP-CA-labeled virions containing the N74D CA mutation. Time-lapse images of the cells were acquired every 20 min using spinning disk confocal microscopy. A GFP-CA-labeled viral complex uncoated and lost the GFP-CA signal at or near the NE (7:20 hpi), a HIV-1 transcription site appeared near the site of GFP-CA disappearance 13:40 hpi, and *gfp* reporter was detected 17:00 hpi. For display, the z-slice closest to the viral complex was extracted from the z-stack for each time point. Time scale, hours:minutes post-infection; scale bar, 5 μm.

Movie S5. Visualizing the nuclear import of a GFP-CA-labeled viral complex in cells expressing mRuby-CPSF6. HeLa:mRuby-CPSF6 cells were infected with GFP-CA-labeled virions. Time-lapse images of the cells were acquired every 20 min using spinning disk confocal microscopy. A GFP-CA-labeled viral complex docked at the NE (5:10 hpi), colocalized with mRuby-CPSF6 (5:30 hpi), and entered the nucleus (7:10 hpi). For display, the z-slice closest to the viral complex was extracted from the z-stack for each time point. Time scale, hours:minutes post-infection; scale bar, 5 μm.

Movie S6. Visualizing the simultaneous loss of GFP-CA and associated mRuby-CPSF6 in the nucleus. HeLa:mRuby-CPSF6 cells were infected with GFP-CA-labeled virions. Time-lapse

images of the cells were acquired every 20 min using spinning disk confocal microscopy. A nuclear GFP-CA-labeled viral complex was visualized for ~4 hrs prior to simultaneous disappearance of GFP-CA and the associated mRuby-CPSF6 (9:30 hpi). For display, the z-slice closest to the viral complex was extracted from the z-stack for each time point. Time scale, hours:minutes post-infection; scale bar, 5 μm.