Supplementary Information for

HIV-1 cores retain their integrity until minutes before uncoating in the nucleus

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

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

This PDF file includes:
  Supplementary Materials and Methods
  Figures S1 to S4
  Legends for Movies S1 to S5

Other supplementary materials for this manuscript include the following:

Movies S1 to S5
Materials and Methods

Cell lines and reagents. HeLa [American Type Culture Collection (ATCC CCL-2)] and human embryonic kidney 293T cells (ATCC CRL-3216) were maintained in Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% fetal calf serum (HyClone, Logan, UT) and penicillin-streptomycin (50 units/ml and 50 µg/ml, respectively; Lonza, Walkersville, MD). Peripheral blood mononuclear cells (PBMCs) from healthy donors (ATCC) were maintained in Roswell Park Memorial Institute (RPMI) 1640 medium (CellGro, Manassas, VA) supplemented with 10% fetal calf serum and penicillin-streptomycin (50 units/ml and 50 µg/ml, respectively). PBMCs were activated with phytohaemagglutinin P (PHA-P; 5 µg/ml) for 72 hours and maintained in IL-2 (50 units/ml). All cells were maintained in a humidified incubator at 37°C with 5% CO₂.

Construction of HeLa:Bgl-mCherry, HeLa:mRuby-CPSF6, and CEM-SS:mRuby-LaminB cell lines was previously described (1). Briefly, the HeLa:Bgl-mCherry cell line expresses a truncated bacterial protein BglG that is fused to mCherry at the C-terminus and contains a nuclear localization signal (Bgl-mCherry). HeLa cells were transduced with a lentiviral vector that expresses Bgl-mCherry from a doxycycline-inducible promoter, and a single cell clone expressing low levels of Bgl-mCherry was obtained. The resulting HeLa:Bgl-mCherry cell line was maintained in complete media containing puromycin (1 µg/ml; Thermo Fisher Scientific) and G418 sulfate (200 µg/ml; Sigma-Aldrich). Doxycycline (1 µg/ml; Sigma-Aldrich) was added 24 hrs prior to imaging to induce Bgl-mCherry expression.

To generate the HeLa:mRuby-CPSF6 cell line, 3 x 10⁴ HeLa cells were transduced with a lentiviral vector that constitutively expresses an shRNA that depletes endogenous CPSF6 mRNA. Next, the 3 x 10⁴ HeLa cells were transduced with another lentiviral vector that constitutively expresses an shRNA-resistant mRuby-CPSF6. The resulting pool of HeLa:mRuby-CPSF6 cells, which expresses slightly less CPSF6 than plain HeLa cells (within twofold), was maintained in complete media containing hygromycin (200 µg/ml; Thermo Fisher Scientific) and puromycin (1 µg/ml; Thermo Fisher Scientific).
To generate the CEM-SS:mRuby-LaminB cell line, CEM-SS cells (NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH; gift from Dr. Peter L. Nara; Cat#776) were transduced with a lentiviral vector that constitutively expresses mRuby-LaminB fusion protein and selected for resistance to puromycin (1 µg/ml). The resulting pool of CEM-SS:mRuby-LaminB cells was maintained in Roswell Park Memorial Institute (RPMI) 1640 medium (CellGro, Manassas, VA) with 10% fetal calf serum and penicillin-streptomycin (50 units/ml and 50 µg/ml, respectively).

**Lentiviral vectors, protein expression vectors, and virus production.** The HIV-1 based vectors pHGFP-BglSL and pHGFP-GFPCA-BglSL were constructed previously (1). Briefly, pHGFP-BglSL vector expresses a gfp reporter and an RNA that contains 18 copies of BglG RNA stem loop (BglSL; (2)), but does not express Vif, Vpr and Env. BglSL are specifically recognized by the bacterial Bgl protein (3). pHGFP-GFPCA-BglSL is similar to pHGFP-BglSL, except that a GFP-CA fusion protein is generated after proteolytic processing. To generate pHGFP-iGFP-BglSL vector, a fragment cut with restriction enzymes MluI and BssHII from HIV Gag-iGFP ΔEnv vector [NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH, from Dr. Benjamin Chen; Cat#12455; (4)] was inserted into pHGFP-GFPCA-BglSL to replace the fragment cut with MluI and BssHII enzymes. The resulting pHGFP-iGFP-BglSL vector expresses free GFP and CA after proteolytic processing. Infectious virions that were labeled with the GFP content marker were prepared by co-transfection of 293T cells with pHGFP-iGFP-BglSL and pHGFP-BglSL at a 1:2 plasmid ratio and pHCMV-G (5), which expresses the G glycoprotein of vesicular stomatitis virus (VSV-G). Infectious virions that were labeled with GFP-CA were prepared by co-transfection of 293T cells with pHGFP-GFPCA-BglSL and pHGFP-BglSL at a 1:10 plasmid ratio, and pHCMV-G. A plasmid expressing APOBEC3F (A3F) fused to red-red vine tomato (RRvT) fluorescent protein was generated previously (1). To generate infectious virions that were co-labeled with the GFP content marker and A3F-RRvT, the A3F-RRvT expressing plasmid was co-transfected with pHGFP-iGFP-BglSL, pHGFP-BglSL, and pHCMV-G. Supernatants from the 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 (4 × 10⁴ cells/well; Ibidi, Gräfelfing, Germany) one day prior to infection. CEM-SS:mRuby-LaminB cells and PBMCs were seeded into μ-slides (1 × 10⁵ cells/well) that were pretreated with poly-L-lysine (Sigma-Aldrich; Cat#P8920). Cells were infected with viruses via spinoculation at 16ºC (6), which permitted virion binding to cell membranes but prevented virion endocytosis. CD4⁺ T cells were identified by immunofluorescence staining of CD4 using AlexaFluor647-labeled anti-human CD4 antibody (BD Pharmingen). For live-cell imaging experiments, infections with the GFP content marker- and 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). Virus infectivity was determined by flow cytometry (LSRFortessa; BD Biosciences, San Jose, CA) 24-48 hpi.

**Fractionation of viral cores using sucrose gradients.** The fractionation of viral cores using sucrose gradients was performed as previously described, with slight modifications (7). Concentrated virions were prepared as described above and subjected to ultracentrifugation (100,000 × g for 16 hrs at 4 ºC) through a layer of 1% Triton X-100 into a linear 30% to 70% sucrose density gradient containing 10 ml of 10 mM Tris-HCl [pH 7.4], 100 mM NaCl, and 1 mM EDTA. Following ultracentrifugation, 0.5 ml fractions were collected from the top of the gradient. CA and GFP content marker 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) followed by goat anti-rabbit antibody (IRDye-800CW; LI-COR, Lincoln, NE) and goat anti-mouse antibody (IRDye-680RD; LI-COR, Lincoln, NE). Western blots were imaged and quantitated using the Odyssey infrared imaging system (LI-COR).
**Microscopy and image processing.** Confocal images were acquired using Nikon Eclipse Ti-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 (Alexa Fluor 405 dye), 488-nm (GFP), 561-nm (mRuby/RRvT), 594-nm (mCherry), and 647-nm (Cy5 dye) 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. Time-lapse images of the infected cells (described below) were examined using Nikon Elements or ImageJ (8). 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 analysis of GFP content marker loss, HIV-1 transcription site detection, and gfp reporter detection.** To visualize the nuclear GFP content marker-labeled 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 h starting at ~4 hpi. The GFP signals and HIV-1 TS were manually identified upon extensive analysis of the entire movie. GFP intensities for individual particles were determined using a custom-written MATLAB program. Briefly, the 3-dimensional (3D) position of the highest intensity GFP-CA signal in the z-stack was manually determined and then the pixel intensities at the location of the GFP signal were determined after local-background subtraction. Overall background intensities were determined by selecting random positions in the nuclei of infected cells and extracting pixel intensities. The GFP content marker-labeled nuclear complexes in CEM-SS::mRuby-LaminB cells were identified manually from z-stacks acquired between 6 and 10 hpi; iGFP intensities for these complexes were determined as described above.

Distances between the GFP content marker-labeled viral complexes and associated HIV-1 transcription sites were determined using a custom written MATLAB program. Briefly, the image of the nucleus in the last frame the GFP content marker signal was detected prior to disappearance
and the image of the nucleus in the first frame in which the associated HIV-1 transcription site was detected were aligned to account for cell movement. Next, the 3D positions of the GFP-labeled particle and the HIV-1 transcription site were determined using 3D gaussian fitting, which were then used to calculate distances.

**Live-cell imaging of fusion of GFP content marker-labeled virus.** To determine the extent of the GFP content marker signal trapped inside the capsid upon fusion, HeLa:Bgl-mCherry cells seeded in ibiTreated µ-slides (4 × 10^4 cells/well) were infected with the GFP content marker-labeled virus at very low MOI (≤ 0.01; <2 GFP^+ particles/cell) via spinoculation at 16°C. The low MOI infection allows unambiguous tracking of most of the particles during the observation period. To capture the rapid fusion kinetics, z-stacks of the infected cells (13 slices at 0.4 µm interval) were acquired every 1 min for 30 min starting 2 min after the temperature was equilibrated from 16°C to 37°C.

**Live-cell imaging of nuclear capsids labeled with GFP content marker or GFP-CA in PF74-treated cells.** To determine the sensitivity of nuclear capsids to PF74, HeLa-mRuby-LaminB cells were infected with GFP- or GFP-CA-labeled virions. Starting at 4 hpi, time-lapse images of the cells were acquired every 2-3 min for 60 min using spinning disk confocal microscopy. After the sixth frame, the movie was paused, cells were treated with 2, 5, or 10 µM PF74, and the imaging was resumed for the remaining 50 min. For longer movies, HeLa-mRuby-LaminB cells were infected with GFP-labeled virions and 2 µM PF74 was added 4 hpi. Time-lapse images of the cells were acquired every 20 min for 20 hours immediately after PF74 addition. The loss of the GFP content marker or GFP-CA signals was manually determined.

**Live-cell imaging of nuclear capsids co-labeled with core-associated mRuby-CPSF6 and GFP content marker or GFP-CA.** To distinguish the timing of capsid integrity loss and core disassembly, HeLa:mRuby-CPSF6 cells infected with GFP content marker- or GFP-CA-labeled viruses were imaged at 5 min/frame using 100 millisecond exposure for 4 h starting at ~8 hpi. The loss of the GFP content marker or GFP-CA and core-associated mRuby-CPSF6 signals was
manually scored. The probability of capturing GFP and mRuby-CPSF6 signals disappearing in the same 5-min period between adjacent frames or the GFP signal disappearing in the 5-min period just before the 5-min period during which mRuby-CPSF6 disappeared was modeled for comparison.

**Single Virion Analysis.** 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 (9). The positions of the spots were used to determine colocalization; spots were considered colocalized if the centers of the spots were within 3 pixels.

**Fixed-cell imaging and image analysis.** A custom-written MATLAB program was used to determine the colocalization of A3F-RRvT and GFP signals inside the nuclei of infected cells. First, the A3F-RRvT signals were detected using Localize (9). 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 Alexa Fluor 405-labeled secondary antibody [Thermo Fisher Scientific, Cat#A-31553 or Cat#A-31556]) was determined using a custom-written MATLAB program as previously described (10). 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 background 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 the GFP content marker. The percentage of GFP-labeled particles that colocalized with the NE mask and the percentage of particles that colocalized with the nucleus mask were determined.
In vitro analysis of intact virions and viral cores. GFP content marker-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+ particles and determine their signal intensities. Virus particles that lost ≥33% of the initial GFP content marker signal were considered to undergo membrane fusion and have mature capsids (typically 90% of particles) and were used for further analysis.

Transmission Electron Microscope (TEM) analysis of virus pellets. Virus pellets labeled with or without GFP content marker were obtained from two 10-cm plates of transfected 293T cells and subjected for TEM analysis following previously described procedures (11, 12). 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 thin-sectioned and stained as previously described (1). The sections were examined and images were captured by digital camera in Hitachi 7650 TEM (Hitachi, Tokyo, Japan) operated at 75kv (12). Virions exhibiting the mature and immature phenotype were scored manually.

Data analysis and statistics. The Welch’s unpaired t-test and paired t-test were used to analyze parametric data. A Fisher’s exact test was used to analyze 2 × 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. R. C. Burdick et al., HIV-1 uncoats in the nucleus near sites of integration. Proc Natl Acad Sci U S A 117, 5486-5493 (2020).
2. 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).
3. F. Houman, M. R. Diaz-Torres, A. Wright, Transcriptional antitermination in the bgl operon of E. coli is modulated by a specific RNA binding protein. Cell 62, 1153-1163 (1990).
4. W. Hubner 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).
5. J. K. Yee et al., 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, 9564-9568 (1994).
6. U. O'Doherty, W. J. Swiggard, M. H. Malim, Human immunodeficiency virus type 1 spinoculation enhances infection through virus binding. J Virol 74, 10074-10080 (2000).
7. V. B. Shah, C. Aiken, In vitro uncoating of HIV-1 cores. J Vis Exp 10.3791/3384 (2011).
8. C. A. Schneider, W. S. Rasband, K. W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis. Nat Methods 9, 671-675 (2012).
9. D. Zenklusen, D. R. Larson, R. H. Singer, Single-RNA counting reveals alternative modes of gene expression in yeast. Nat Struct Mol Biol 15, 1263-1271 (2008).
10. R. C. Burdick et al., Dynamics and regulation of nuclear import and nuclear movements of HIV-1 complexes. PLoS Pathog 13, e1006570 (2017).
11. J. M. Mariner, J. B. McMahon, B. R. O'Keefe, K. Nagashima, M. R. Boyd, The HIV-inactivating protein, cyanovirin-N, does not block gp120-mediated virus-to-cell binding. Biochem Biophys Res Commun 248, 841-845 (1998).
12. G. J. Tobin, K. Nagashima, M. A. Gonda, Immunologic and Ultrastructural Characterization of HIV Pseudovirions Containing Gag and Env Precursor Proteins Engineered in Insect Cells. Methods 10, 208-218 (1996).
**Fig. S1** Analysis of GFP-labeled virions before and 60 seconds after treatment with saponin detergent. (A) GFP-labeled virions were spun onto slides, treated with saponin detergent and imaged every 10 seconds for 5 minutes. The GFP intensities (arbitrary units [a.u.]) for intact virions (n=1608) and viral cores (n = 762) 60 seconds after initial loss of GFP are shown. The GFP signal completely disappeared for 846 virions. (B) The intact virions were divided up into four quartiles based on GFP intensity and ranked from lowest 25% (Q1) to the highest 25% (Q4). The numbers of detectable GFP-labeled intact virions and viral cores and the mean GFP intensities of the intact virions and viral cores were determined (shown below x-axis). The GFP intensities of the detectable viral cores directly correlate with the GFP intensities of the intact virions. (C) The percentage of GFP-labeled viral cores that were detected 60 seconds after initial loss of GFP increases with increasing intact virion intensity. These results indicate the loss of GFP-labeled viral cores is largely a result of GFP levels falling below the detection limit.
Fig. S2 PF74 treatment leads to an abrupt (< 2 min) loss of the GFP content marker. The GFP intensities for 10 GFP-labeled capsids that disappeared in cells treated with 2 µM PF74 treatment and 10 GFP-labeled capsids that disappeared in cells treated with 10 µM PF74 treatment were determined and normalized to the GFP intensity in the frame immediately before loss of detection.
Fig. S3 Effect of PF74 treatment on nuclear viral cores labeled with GFP content marker or GFP-CA. (A) The percentage of nuclear viral cores labeled with GFP content marker or GFP-CA that disappeared <50 min of treatment with different concentrations of PF74. (B) Time of GFP disappearance for nuclear capsids labeled with GFP content marker or GFP-CA relative to the time of 10 µM PF74 addition. Lines are mean ± SD; *P* values are from Welch’s *t*-tests; ns, not significant (*P* > 0.05).
**Fig. S4** High time-resolution analysis of the loss of GFP, GFP-CA, and capsid-associated mRuby-CPSF6. (A-B) Representative live-cell microscopy images (1 frame/5 minutes) of infected HeLa:mRuby-CPSF6 cells showing disappearance of GFP signal one frame before mRuby-CPSF6 disappearance (A) and disappearance of both GFP-CA and mRuby-CPSF6 signals in the same frame (B). Scale bars for A and B, 5 µm; inset, 1 µm. (C) Illustrations summarizing the two-step process of uncoating starting with loss of capsid integrity (loss of GFP) ~1-3 min before core disassembly (uncoating; loss of GFP-CA and mRuby-CPSF6), which occurs within ~1 min.
Supplementary Movie Legends

**Movie S1.** Fusion of an GFP-labeled virion results in dramatic loss of GFP signal and detection of an GFP-labeled capsid. HeLa:Bgl-mCherry cells were infected with GFP-labeled virions. Time-lapse images of the cells were acquired every 2 min using spinning disk confocal microscopy. Fusion occurred between 10- and 11-min post-infection, as indicated by the significant drop in GFP signal, and the GFP-labeled capsid was tracked for the remainder of the observation time (24 min post-infection). 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.** Loss of capsid integrity inside the nucleus. HeLa:Bgl-mCherry cells were infected with GFP-labeled virions. Time-lapse images of the cells were acquired every 20 min using spinning disk confocal microscopy. A GFP-labeled capsid docked at the nuclear envelope and entered the nucleus 3:50 hours post-infection (hpi), lost integrity inside the nucleus 6:30 hpi, an HIV-1 transcription site appeared near the site of GFP disappearance 16:10 hpi, and gfp reporter was detected 19:30 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.

**Movie S3.** Visualizing the simultaneous loss of GFP and associated mRuby-CPSF6 in the nucleus. HeLa:mRuby-CPSF6 cells were infected with GFP-labeled virions. Time-lapse images of the cells were acquired every 5 min using spinning disk confocal microscopy. A nuclear GFP-labeled capsid was tracked and the GFP and associated mRuby-CPSF6 signals disappeared in the period between 08:05-08:10 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 S4.** Visualizing the loss of GFP one frame before associated mRuby-CPSF6 in the nucleus. HeLa:mRuby-CPSF6 cells were infected with GFP-labeled virions. Time-lapse images of the cells were acquired every 5 min using spinning disk confocal microscopy. A nuclear GFP-labeled capsid was tracked, and the GFP signal disappeared in the period between 09:15-09:20 hpi and the associated mRuby-CPSF6 signal disappeared in the period between 09:20-09:25 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 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 5 min using spinning disk confocal microscopy. A nuclear GFP-CA-labeled capsid was tracked, and the GFP-CA and associated mRuby-CPSF6 signals disappeared in the period between 08:35-08:40 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.