Science Advances
AAAS

Supplementary Materials for

HIV-1 uncoating requires long double-stranded reverse transcription products

Ryan C. Burdick *et al.*

Corresponding author: Vinay K. Pathak, vinay.pathak@nih.gov

Sci. Adv. **10**, eadn7033 (2024)
DOI: 10.1126/sciadv.adn7033

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

Supplementary Text

Uncoating efficiency is dependent on genome length and not genome sequence

We determined the uncoating efficiency of a short HIV-1 based vector (HIV-2.7kb) that is similar to HIV-3.1kb but does not contain the P2A-Tat sequence (fig. S3A). Compared to the HIV-11.1kb control vector, the HIV-2.7kb vector produced ~3-fold more late viral DNA at 6 hours after infection, indicating that the HIV-2.7kb genome was efficiently packaged and reverse transcribed (fig. S3B). The uncoating efficiency of viral cores containing the HIV-2.7kb genome was low (~10%; fig. S3C) and similar to that of viral cores containing the HIV-2.8kb genome or no genome (Fig. 1F). The low uncoating efficiency of HIV-2.7kb and HIV-2.8kb, which are similar in length but contain different sequences, indicates that the uncoating efficiency is dependent on genome length and not the genomic sequence. The few HIV-2.7kb cores that uncoated did so with delayed uncoating kinetics (~15.2 hours after infection; fig. S3D).

HIV-11.1kb vector uncoats with an efficiency similar to HIV-10.3kb and HIV-8.1kb vectors

We determined the uncoating efficiency of an HIV-1 vector (HIV-11.1kb) that is ~1.4 kb larger than the HIV-1 genome (~9.7 kb). The uncoating efficiency of viral cores containing HIV-11.1kb was ~54% (fig. S3C) and similar to the uncoating efficiency of viral cores containing a 10.3-kb genome (~60%; Fig. 1B) or 8.4-kb genome (~52%; Fig. 1F). The uncoating kinetics of viral cores containing HIV-11.1kb (~10.0 hours after infection) was similar to viral cores containing a 10.3 kb genome (~10.6 hours after infection; Fig. 1C). These data indicate that viral cores containing an HIV-1 genome larger than wild type (HIV-11.1kb) uncoats with an efficiency similar to viral cores containing HIV-10.3kb or HIV-8.1kb genomes.

Uncoating of viral cores containing short vector genomes is slow and inefficient

Because integration of HIV-3.1kb in untreated HeLa cells was significantly delayed (Fig. 2E), we measured the number of nuclear viral cores 1-3 days after infection. Interestingly, most nuclear viral cores could still be detected in untreated HeLa cells 1-3 days after infection, with only a ~20% decrease in the number of nuclear cores 3 days after infection (fig. S5, A and B). The percentage of GFP reporter expressing cells at 4 days after infection with HIV-3.1kb remained low compared to HIV-8.4kb (fig. S5C), indicating that only small proportion of HIV-3.1kb viral cores uncoated between 1 – 4 days after infection.

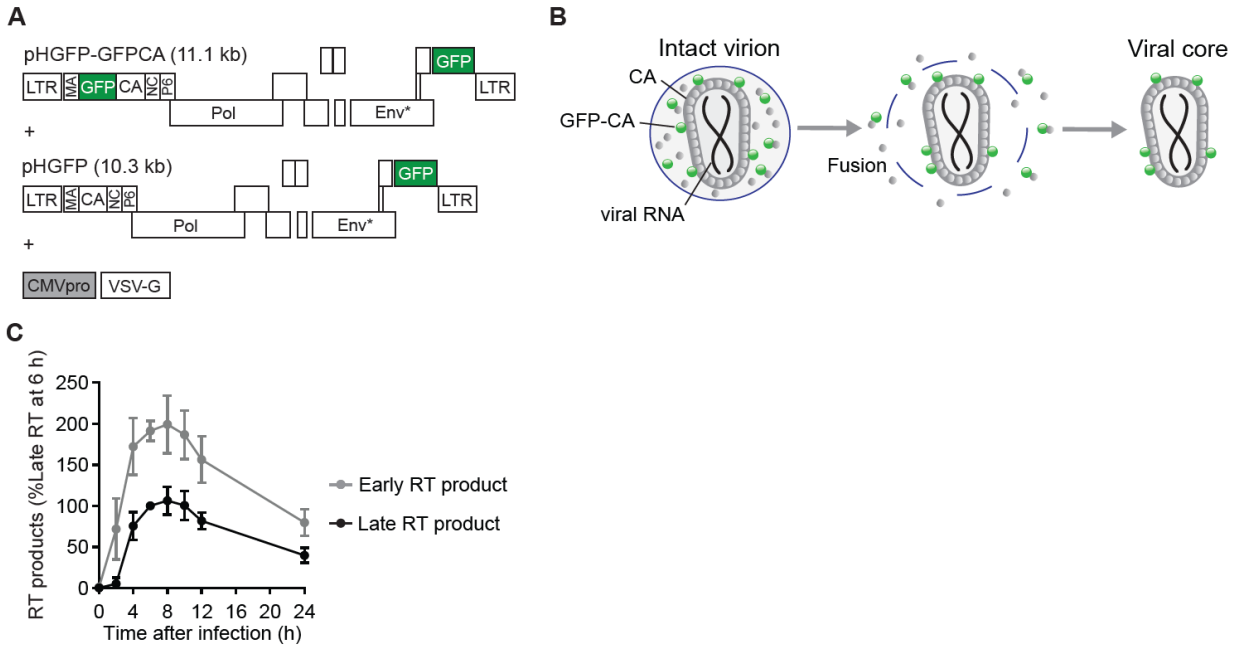

Fig. S1. Determination of reverse transcription kinetics and HIV-1 uncoating kinetics after NVP washout by PF74 time-of-addition and live-cell imaging assays. (A) Schematic of the vectors used to generate GFP-CA-labeled virus. Infectious virions labeled with GFP-CA were prepared by co-transfection of 293T cells with pHGFP-GFPCA and pHGFP at a 1:10 plasmid ratio, and a plasmid expressing the G glycoprotein of vesicular stomatitis virus (VSV-G). Asterisk indicates mutation in the *env* gene that introduces a premature stop codon. **(B)** GFP-CA is a marker that can be used to detect HIV-1 cores after fusion of the viral and cell membranes. **(C)** Measurement of early RT and late RT products in aphidicolin-treated HeLa cells at indicated time points after infection (mean ± SD, $N = 3$ biological replicates).

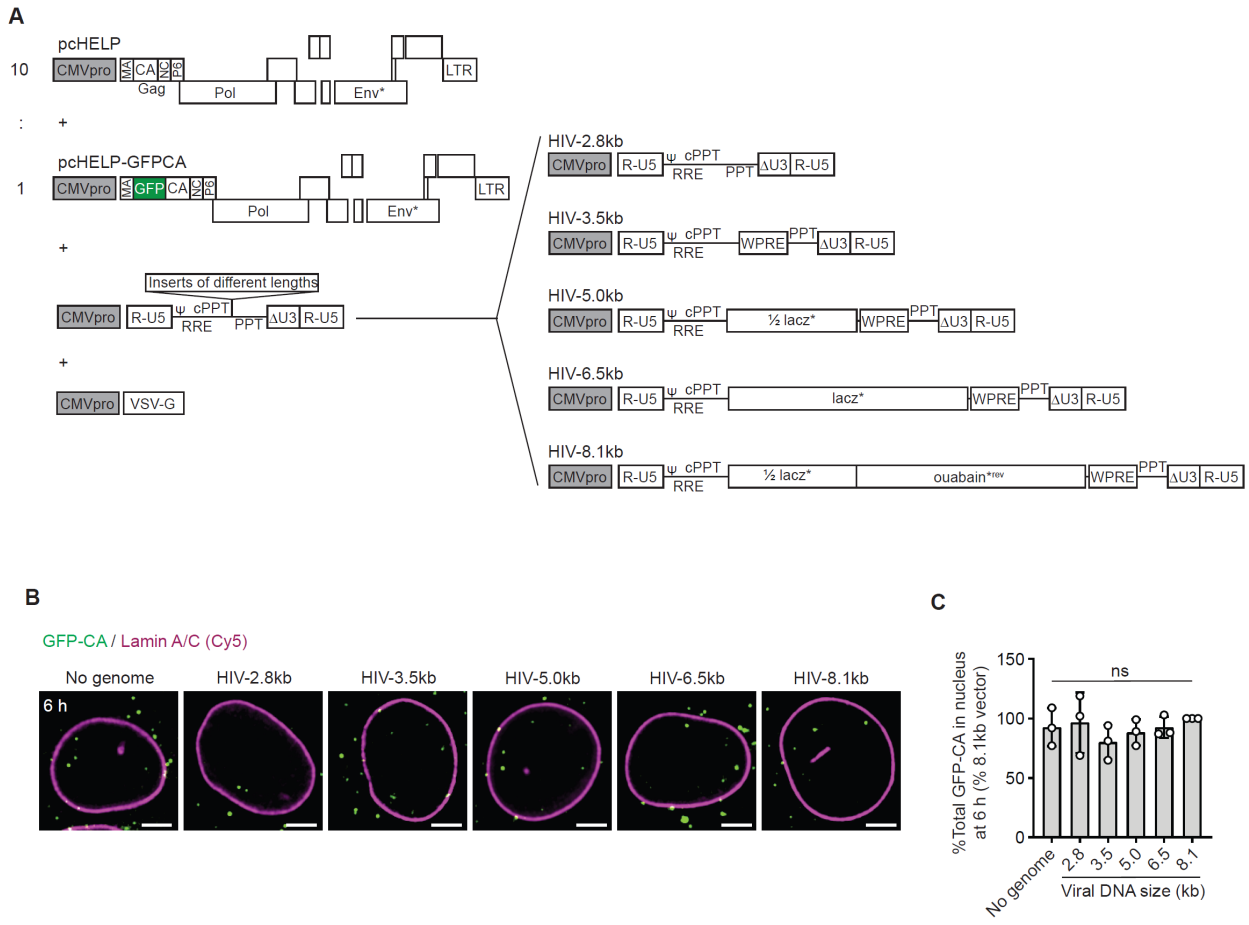

Fig. S2. Schematic of the vectors used to generate GFP-CA-labeled virus containing no viral genome or different-sized viral genomes and nuclear import efficiency. (A) Different-sized lentiviral vectors were constructed by addition of non-HIV-1 sequences, including *LacZ*, *ouabain*, and Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element (WPRE) sequences. Asterisk indicates absence of start codon (i.e. protein is not expressed). pcHELP and pcHELP-GFPCA are HIV-1 helper constructs that lack a packaging signal and primer-binding site; the helper construct expresses all the viral proteins except Nef and Env. pcHELP expresses WT CA whereas pcHELP-GFPCA expresses a GFP-CA fusion protein (after proteolytic processing of Gag during virion maturation). A 1:10 ratio of pcHELP-GFPCA to pCHELP and a VSV-G expression plasmid was used to generate GFP-CA-labeled virions that packaged no genome or the different-sized genomes (HIV-2.8kb-HIV-8.1kb). The HIV-1 vectors contain all of the cis-acting elements needed to complete reverse transcription. The expected genome size (kb) after RT is indicated in each vector name. **(B)** HeLa cells were challenged with GFP-CA-labeled HIV-1 particles containing no genome or different-sized genomes, fixed 6 hours after infection, the nuclear envelope was stained using an anti-Lamin A/C antibody followed by a secondary antibody labeled with Cy5, confocal images were acquired, and the GFP-CA spots were automatically detected. Scale bar, 10 µm. **(C)** Percentage of total (cytoplasm + nucleus) GFP-CA spots inside the nucleus 6 hours after infection. Data are shown relative to HIV-8.1kb (set to 100%; mean ± SD, $N = 3$ biological replicates, Welch's t-test; ns, not significant).

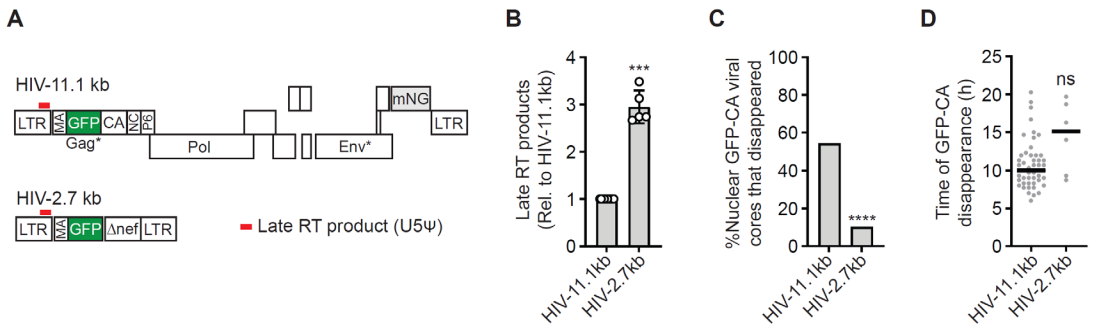

Fig. S3. Measurement of viral DNA levels and uncoating of viral cores containing a short genome. (A) Schematic of HIV-2.7kb, a short HIV-1-based vector that expresses an MA-GFP fusion protein under the control of the LTR, and HIV-11.1kb, a control vector. An asterisk indicates mutation in gene that introduces a premature stop codon. Size (kb) of the genome is indicated in the vector name. The location of the primer/probe target of late RT product (U5Ψ; red line) is indicated. HIV-1 helper plasmids were used to generate virus particles; the viral genomes were packaged *in trans*. **(B)** Measurement of late RT products 6 hours after infection of aphidicolin-treated HeLa cells with p24-normalized virus input. Data are shown relative to HIV-11.1kb (set to 1; mean ± SD, Welch's t-test, $N = 5$ biological replicates). **(C)** Percentage of nuclear HIV-1 cores that disappeared during the live-cell movies for the indicated virus. An average of 75 HIV-1 cores was analyzed for each sample (Fisher's exact test). **(D)** Time (h) after infection of GFP-CA disappearance (line = median, Mann-Whitney U test). ***, $P < 0.001$, ****, $P < 0.0001$; ns, not significant.

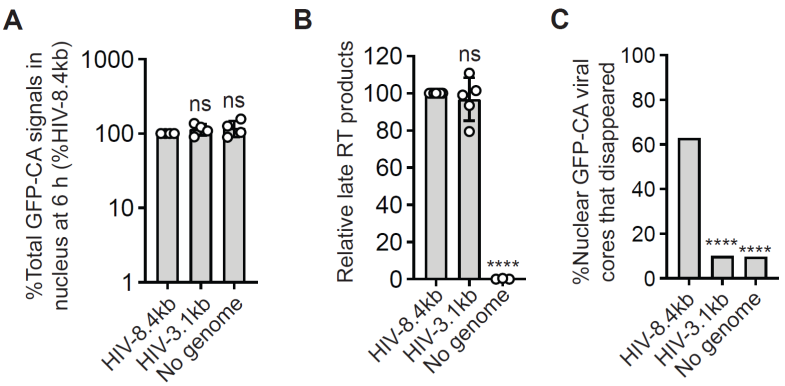

Fig. S4. Effect of genome size on nuclear import efficiency of HIV-1 cores, RT product levels, and uncoating efficiency of nuclear HIV-1 cores. (A) Percentage of total (cytoplasm + nucleus) GFP-CA spots inside the nucleus 6 hours after infection. HeLa cells were challenged with GFP-CA-labeled HIV-1 particles containing no genome or different-sized genomes (HIV-8.4kb or HIV-3.1kb), fixed 6 hours after infection, the nuclear envelope was stained using an anti-Lamin A/C antibody followed by a secondary antibody labeled with Cy5, confocal images were acquired, and the GFP-CA spots were automatically detected. Data are shown relative to HIV-8.4kb (mean ± SD, Welch's t-test, $N = 4$ biological replicates). **(B)** Quantitation of late RT products for indicated viruses 6 hours after infection with p24-normalized virus input. Data are shown relative to HIV-8.4kb (set to 100%; mean ± SD, Welch's t-test, $N = 3$-7 biological replicates). **(C)** Percentage of nuclear HIV-1 cores that disappeared during live-cell movies for the indicated virus. Aphidicolin-treated HeLa cells were challenged with GFP-CA-labeled virus containing HIV-8.4kb, HIV-3.1kb, or no genome, and live-cell microscopy was performed. An average of 60 HIV-1 cores was analyzed for each sample. Fisher's exact test was used to compare statistical significance to WT in untreated cells. ****, $P < 0.0001$; ns, not significant.

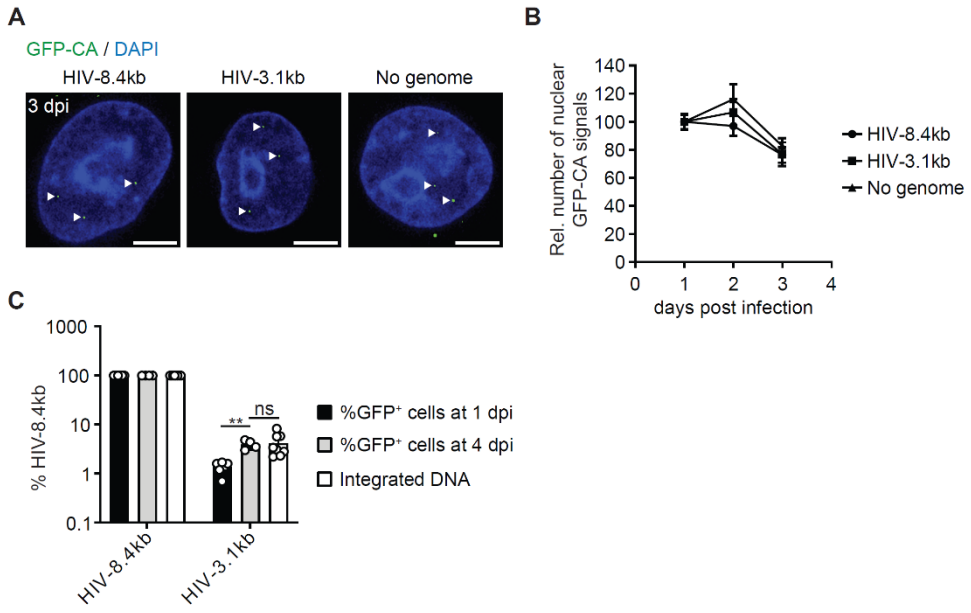

Fig. S5. Detection of HIV-1 cores 1-3 days after infection and measurement of GFP reporter expression. (A-B) Detection of nuclear HIV-1 cores 1, 2, and 3 days after infection of aphidicolin-treated HeLa cells with GFP-CA-labeled HIV-1 particles containing HIV-8.4kb, HIV-3.1kb, or no genome. Infected cells were fixed 1, 2, or 3 days after infection, nuclei were stained using DAPI, confocal images were acquired, and the nuclear GFP-CA spots were automatically detected. **(A)** Examples of nuclear HIV-1 cores 3 days after infection are shown for indicated viruses. Scale bars, 10 µm. **(B)** Number of nuclear HIV-1 cores/cell. Data is shown relative to the number of nuclear HIV-1 cores/cell 1 day after infection for each virus (set to 100%). A total of ~75 cells from 3 independent infections were analyzed (mean ± SEM). **(C)** Percentage of GFP reporter-expressing cells measured 1 and 4 days after infection and integrated viral DNA levels measured using late RT primers and probe 7 days after infection of HeLa cells with the indicated virus (HIV-8.4kb set to 100%; mean ± SD, Welch's t-test, $N = 4$-8 biological replicates). **, $P < 0.01$, ns, not significant.

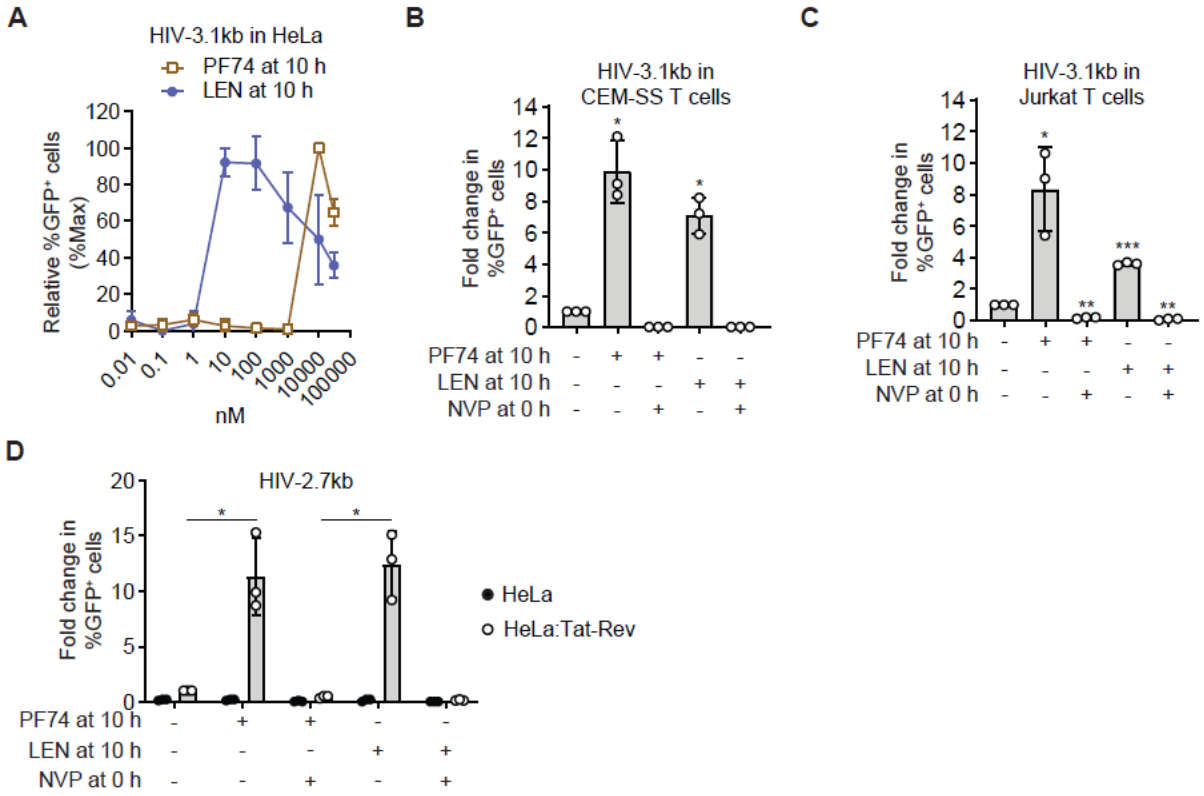

Fig. S6. Disruption of HIV-1 cores containing HIV-3.1kb with PF74 and LEN in HeLa cells and T cells leads to GFP reporter expression. (A) Percentage of GFP reporter-expressing cells measured ~1 day after infection of HeLa cells with HIV-3.1kb treated with different concentrations of PF74 or LEN added 10 hours after infection (mean ± SD, N = 3 biological replicates). The percentage of GFP reporter-expressing cells in untreated control cells (~1-2%) was subtracted from all samples and data is shown relative to the maximal value (10 μM for PF74 and 10 nM or 100 nM for LEN; set to 100%). (**B-C**) Fold change in the percentage of GFP reporter-expressing cells measured 1 day after infection of CEM-SS T cells **(B)** or Jurkat T cells **(C)** with HIV-3.1kb and treatment of cells with 10 μM PF74 or 10 nM LEN 10 hours after infection. Data are shown relative to the percentage of GFP reporter-expressing cells in untreated cells (set to 1; mean ± SD, Welch's t-test, N = 3 biological replicates). **(D)** Fold change in the percentage of GFP reporter-expressing cells measured 1 day after infection of HeLa cells or HeLa cells constitutively expressing Tat and Rev (HeLa:Tat-Rev) with HIV-2.7kb (short viral genome that is similar to HIV-3.1kb except does not encode P2A-Tat) and treatment with PF74 10 hours after infection. Data are shown relative to the percentage of GFP reporter-expressing cells in untreated HeLa:Tat-Rev cells (set to 1; mean ± SD, Welch's t-test, N = 3 biological replicates). *, $P < 0.05$, **, $P < 0.01$, ***, $P < 0.001$.

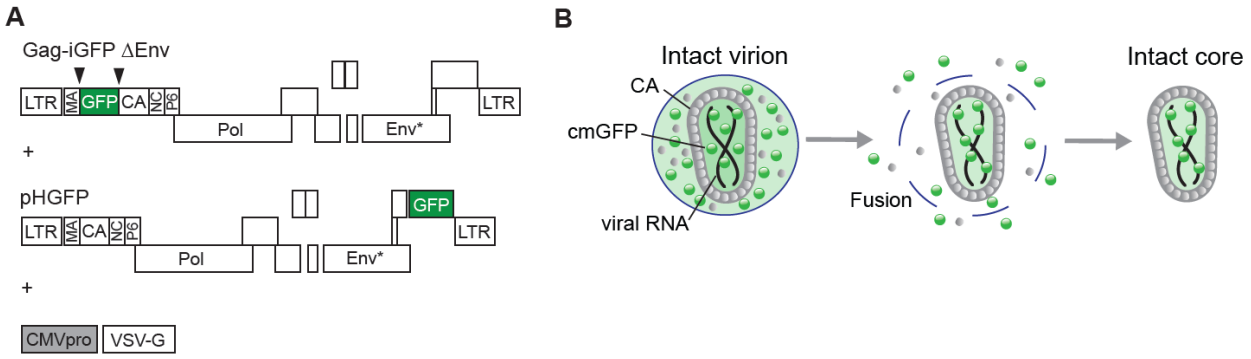

Fig. S7. Detection of intact HIV-1 cores with a GFP content marker. (A) Schematic of the vectors used to generate content marker GFP (cmGFP)-labeled virus. Infectious virions labeled with cmGFP were prepared by co-transfection of 293T cells with an HIV-1-based vector that contains an internal GFP (iGFP) between MA and CA (Gag-iGFP ΔEnv) and pHGFP at a 1:2 plasmid ratio, and a VSV-G expression plasmid. Asterisk indicates mutation in the *env* gene that introduces a premature stop codon. **(B)** Proteolytic processing of the internal GFP during virus maturation produces free GFP, which serves as a content marker that can be used to detect intact HIV-1 cores after fusion of the viral and cell membranes.

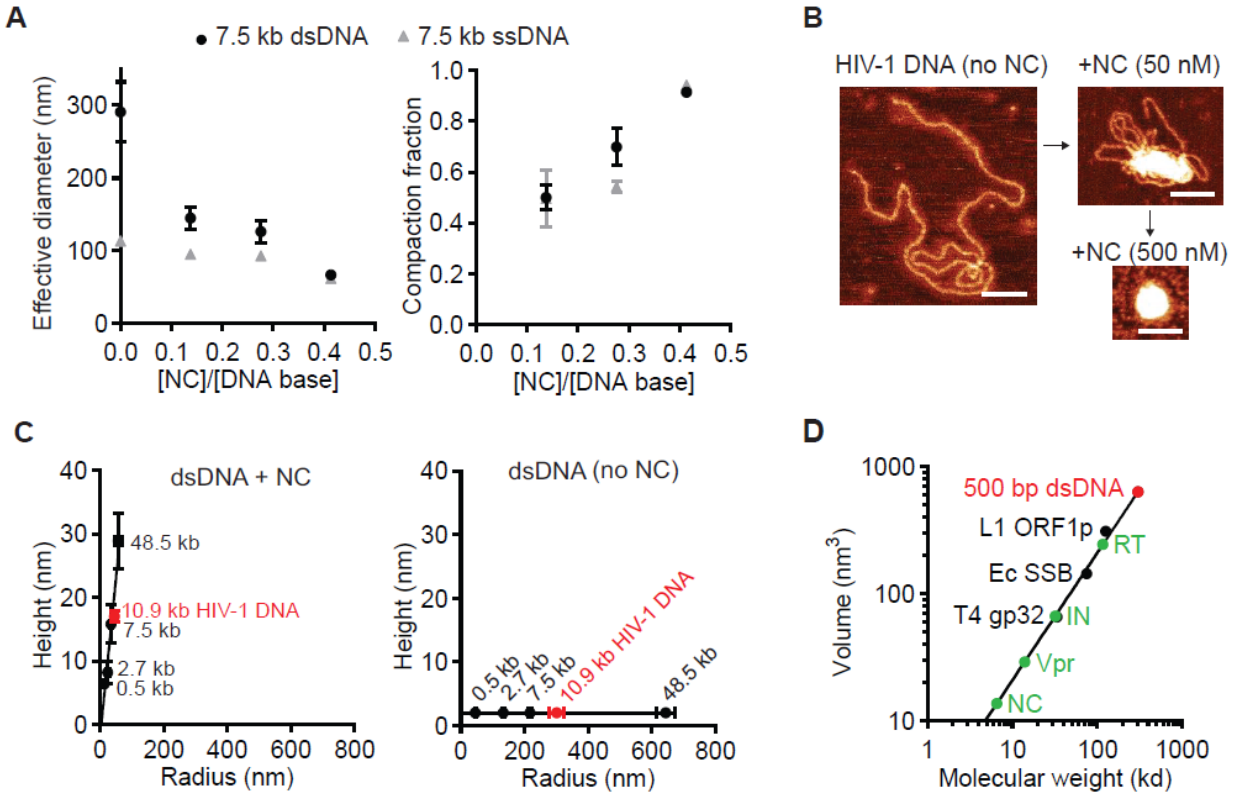

Fig. S8. AFM measurements of NC-DNA condensation and protein size. (A) Diameter (left) and fraction of DNA in central globule (right) after incubation of 7.5 kb dsDNA or 7.5 kb ssDNA with different NC amounts (mean ± SD, N = 3 biological replicates). The [NC]/[DNA base] of 0.28 corresponds to 200 nM NC, though limiting NC-DNA incubation to one minute to prevent aggregation lowers the total quantity of NC bound. **(B)** AFM images of HIV-1 DNA (10.7 kb dsDNA) incubated with or without NC. Scale bars, 100 nm. **(C)** Relationship between height and radius of phage dsDNA of different lengths (black) or HIV-1 DNA (red) incubated with 500 nM NC (left) or no NC (right) (mean ± SEM, N = 12-26 molecules). **(D)** Quantification of the minimal bounding volume of viral proteins using AFM. The volume of 500 bp dsDNA and previously measured proteins (mean ± SEM, N = 180-364 molecules) was measured using AFM and used to confirm the linear relationship between protein volume and molecular weight. The proportionality (volume/molecular weight = 2.09) was used to determine the volumes of indicated viral proteins (green) using the known molecular weight of the viral proteins.

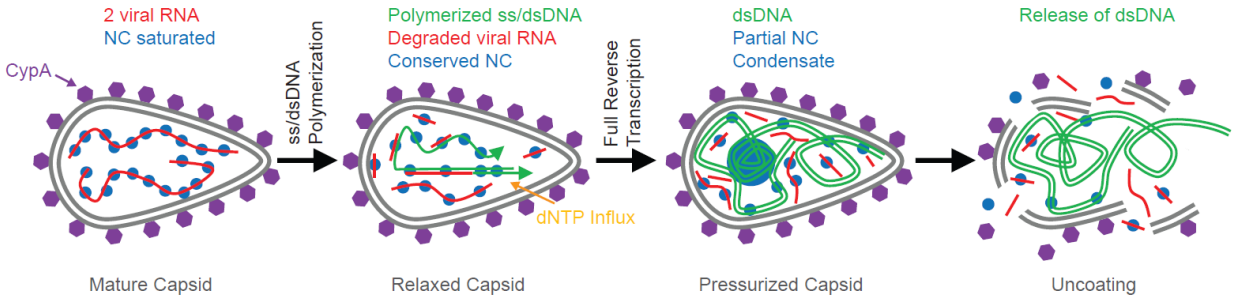

Fig. S9. Model of reverse transcription-induced uncoating of HIV-1 cores. Two copies of single-stranded (ss) viral RNA (red) with sufficient NC (blue) are originally contained in the mature capsid. The viral DNA (green) is polymerized, and viral RNA is cleaved during reverse transcription. During reverse transcription, the capsid contains a combination of single-stranded RNA, RNA/DNA hybrid, single-stranded DNA, and eventually dsDNA. Single-stranded RNA and single-stranded DNA are flexible, and lengths of stiffer duplexed DNA are initially short and discontinuous so that all nucleic acids are easily contained within the capsid. After reverse transcription completion, the quantity of nucleic acid bases contained in the capsid doubles while the number of NC proteins is conserved and is insufficient to fully saturate the completed viral DNA substrate. Uncondensed dsDNA has a persistence length roughly equal to the diameter of the capsid and must be forcibly bent inwards by contact with the capsid, straining the capsid and leading to the loss of capsid integrity. Cyclophilin A (CypA), or another CsA-sensitive cyclophilin, binding stabilizes capsids in the nucleus, ensuring reverse transcription of long dsDNA is completed before uncoating. The location of viral DNA ends appears random because it is unknown whether the intasome has formed at the time of uncoating.

Data S1. Raw data for main and supplementary figures.

The Excel file includes all rad data for generating the main and supplementary figures in this manuscript.