The Cleavage and Polyadenylation Specificity Factor 6 (CPSF6) Subunit of the Capsid-recruited Pre-messenger RNA Cleavage Factor I (CFIm) Complex Mediates HIV-1 Integration into Genes*

Received for publication, February 12, 2016, and in revised form, March 17, 2016 Published, JBC Papers in Press, March 18, 2016, DOI 10.1074/jbc.M116.721647

Sheeba Rasheedi‡, Ming-Chieh Shun§, Erik Serrao‡, Gregory A. Sowd‡, Juan Qian‡, Caili Hao§, Twishasri Dasgupta‡, Alan N. Engelman‡, and Jacek Skowronski†‡§

From the ‡Department of Molecular Biology and Microbiology, Case Western Reserve School of Medicine, Cleveland, Ohio 44106 and §Department of Cancer Immunology and Virology, Dana-Farber Cancer Institute, Boston, Massachusetts 02215

HIV-1 favors integration into active genes and gene-enriched regions of host cell chromosomes, thus maximizing the probability of provirus expression immediately after integration. This requires cleavage and polyadenylation specificity factor 6 (CPSF6), a cellular protein involved in pre-mRNA 3′ end processing that binds HIV-1 capsid and connects HIV-1 preintegration complexes to intranuclear trafficking pathways that link integration to transcriptionally active chromatin. CPSF6 together with CPSF5 and CPSF7 are known subunits of the cleavage factor I (CFIm) 3′ end processing complex; however, CPSF6 could participate in additional protein complexes. The molecular mechanisms underpinning the role of CPSF6 in HIV-1 infection remain to be defined. Here, we show that a majority of cellular CPSF6 is incorporated into the CFIm complex. HIV-1 capsid recruits CFIm in a CPSF6-dependent manner, which suggests that the CFIm complex mediates the known effects of CPSF6 in HIV-1 infection. To dissect the roles of CPSF6 and other CFIm complex subunits in HIV-1 infection, we analyzed virologic and integration site targeting properties of a CPSF6 variant with mutations that prevent its incorporation into CFIm. We show, somewhat surprisingly, that CPSF6 incorporation into CFIm is not required for its ability to direct preferential HIV-1 integration into genes. The CPSF5 and CPSF7 subunits appear to have only a minor, if any, role in this process even though they appear to facilitate CPSF6 binding to capsid. Thus, CPSF6 alone controls the key molecular interactions that specify HIV-1 preintegration complex trafficking to active chromatin.

Integration of the DNA copy of the viral RNA into chromosomal DNA of the host cell is an integral step within the retrovirus replication cycle. Different retrovirus genera show distinct selection preferences for their integration sites, a process guided in part by interactions between components of the viral preintegration complex (PIC) and chromatin targeting proteins (1–4).

HIV-1 preferentially integrates into active genes located in gene dense chromosomal regions (5). This pathway was initially linked to the binding of the cleavage and polyadenylation specificity factor 6 (CPSF6) protein, a pre-mRNA splicing factor and member of the serine/arginine-rich (SR) protein family, to the viral capsid before the PIC translocation to the host cell nucleus (6–10). In particular, mutating the CPSF6 binding site in capsid was shown to activate an alternative mode (11–13) for PIC trafficking that leads to integration away from gene-dense regions and to decreased integration into genes (12, 14). Recent studies revealed that disruption of CPSF6 expression by RNA interference or CRISPR-mediated gene knock-out resulted in a redistribution of HIV-1 integration sites away from gene bodies and gene dense chromosomal regions, thus validating the key role of CPSF6 in HIV-1 integration targeting (15). Furthermore, direct visualization of HIV-1 replication revealed that CA’s interaction with CPSF6 enhances nuclear entry and potentiates HIV-1’s depth of nuclear invasion (54).

The biochemical events along the CPSF6-initiated pathway for HIV-1 integration site selection has only begun to be defined. It is established that this pathway requires the presence of additional cellular proteins implicated in PIC nuclear import such as Transportin 3 (TNPO3), Nucleoporin 358 (NUP358, RanBP2), and Nucleoporin 153 (NUP153) (11–13). The integrase-binding protein lens epithelium-derived growth factor (LEDGF/p75) additionally plays a key role to determine sites of HIV-1 integration along gene bodies (3, 16–19). TNPO3 is a nuclear import receptor that recognizes and binds to RS domains of SR proteins, including cleavage and polyadenylation specificity factor CPSF6, for transport to the nucleus (20–22). The importance of CPSF6 capsid binding and TNPO3-mediated CPSF6 nuclear import in HIV-1 infection is underscored by the finding that their uncoupling in variant CPSF6 proteins that retain capsid binding yet redistribute to the cytoplasm leads to perturbations in viral core uncoating, defective reverse transcription and abortive infection (8, 9, 12, 13).

* This work was supported, in whole or in part, by National Institutes of Health Grant P50 GM082251 (to A. N. E. and J. S.). Shared resources at Case Western Reserve University are supported by National Institutes of Health Center for AIDS Research Grant P30 AI036219. The authors declare that they have no conflicts of interest with the contents of this article.

† To whom correspondence should be addressed. Tel.: 216-368-8930; E-mail: jacek.skowronski@case.edu.

‡ The abbreviations used are: PIC, preintegration complex; SR, serine/arginine-rich; CPSF6, cleavage and polyadenylation specificity factor 6; CFIm, cleavage factor I; CA, capsid protein; RFP, red fluorescent protein; HtR, HIV-1tagRFP; TSS, transcription start site; RRM, RNA recognition motif.
CPSF6 Subunit of CFIm, Directs HIV-1 to Genes

23–25). NUP153 and NUP358 were implicated as important for the nuclear import step in the pathway involving CPSF6 binding to capsid (26–28). LEDGF/p75 binds directly to HIV-1 integrase in the context of the PIC and tethers it to the chromatin (2, 29, 30). Silencing expression of any of the above proteins abolishes the characteristic pattern of preferential HIV-1 integration into active genes (3, 12, 13, 18). These latter findings implicate these proteins as elements of a linear pathway that commits the HIV-1 reverse transcription complex/PIC to a particular nuclear import venue that controls PIC delivery to active chromatin. Yet, the potential role of CPSF6 binding partners in this pathway has not been investigated.

CPSF6 has emerged as a dominant player in directing HIV-1 integration into actively transcribing genes (15), yet it is not completely understood how CPSF6 exerts its effects on HIV infection at a biochemical level. CPSF6 was first identified as a subunit of cleavage factor Im (CFIm), the key regulator of mRNA 3’ end processing and polyadenylation site selection (31, 32). CFIm is a tetramer composed of two 25-kDa (CPSF5) subunits and two proteins of either 59 or 68 kDa (CPSF7 or CPSF6) (31–33). The two CPSF5 subunits form a dimer at the core of CFIm, which recognizes and binds a target site located upstream of the cleavage site at the mRNA 3’ end (33). Each CPSF5 molecule also interacts with its respective CPSF6/7 protein partner through their RM domains (34). Although CPSF6 and CPSF7 display a very similar overall domain organization, they probably play non-redundant roles in CFIm complex and mRNA 3’ end processing, as they have distinct interaction partners linking them to transcription and RNA processing machineries (34–37).

Whereas it has been established that in mRNA 3’ end processing CPSF6 functions in concert with other CFIm complex subunits, it remains unknown whether the function of CPSF6 in post-entry steps of HIV-1 infection is mediated by CPSF6 alone or through the CFIm complex, and, if so, whether the CFIm CPSF6 and/or CPSF7 subunits play any role in directing the HIV-1 PIC to the nucleus and targeting integration to genes is unclear. Here, as a first step in this direction, we attempted to determine the quaternary form of CPSF6 that binds HIV-1 capsid to guide the PIC within the nucleus for integration into active chromatin. We show that the vast majority of cellular CPSF6 is sequestered in the CFIm complex. CPSF6 binding to its docking site in HIV-1 capsid leads to the recruitment of CFIm tetramer, suggesting that CFIm mediates CPSF6 function(s) in integration site targeting. Nevertheless, we find that CPSF6 association with CPSF5 and CPSF7 is not required for HIV-1 integration targeting into active chromatin. Hence, CPSF6 is the only subunit of the capsid-recruited CFIm complex that is required for preferential HIV-1 integration into genes.

Materials and Methods

Expression Vectors and Viruses—Genes encoding HA-epitope (or HFA-epitope: HA, FLAG, AU1 triple epitope); Ref. 38–tagged human CPSF6 protein were cloned into MSCV(GFP) retroviral vector. The CPSF6(Δ116–122) variant (39) was constructed by site-directed mutagenesis with Quick-Change XLII kit (Agilent) and cloned into MSCV(GFP). VSV-G pseudotyped MSCV(GFP) viruses were produced from HEK293T cells as described previously (40). Cell culture supernatants containing infectious virus were collected 24 h post transfection, passed through a 0.45-mm filter, and concentrated by centrifugation with Centricon Plus-70 filter unit (Millipore) at 3000 × g until concentrated 3–4-fold. HIV-1tagRFP (vif-, vpr-, vpu-, nef-) reporter virus expressing tagRFP marker protein from the nef locus was kindly provided by Nicolas Manel. HIV-1tagRFP with N74D capsid mutation was constructed by site-directed mutagenesis using a 3-kb AatII-Spel fragment comprising the capsid protein (CA) coding sequence as template. Mutations were confirmed by DNA sequencing and reintroduced into HIV-1tagRFP proviral backbone. VSV-G pseudotyped virus stocks were produced from HEK293T cells co-transfected with the proviral clone and a VSV-G glycoprotein expression plasmid by a calcium co-precipitation method (38). LKO.1tagRFP lentiviral vector was also produced from HEK293T cells co-transfected with BH10 packaging vector and VSV-G, as we previously described (38). Culture supernatants were harvested 24 h post-transfection, and viral particles were partially purified and concentrated by pelleting through 20% sucrose in 10 mM Tris-HCl, pH 7.4, 100 mM NaCl, 1 mM EDTA cushion at 27,000 rpm for 3 h (41). Viruses were normalized by Western blotting for p24 CA and stored at −80 °C.

Cell Lines and Infections—Human embryonic kidney HEK293T cells and parental and CPSF6 knock-out cells (HEK293T.CPSF6.KO; referred to as B8 cells in Ref. 15) were maintained in Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% fetal bovine serum (FBS), penicillin at 100 units/ml, and streptomycin at 100 µg/ml. THP1, U937, and Jurkat T cells were maintained in RPMI1640 medium supplemented as described above, at cell densities between 3 × 10⁴ and 5 × 10⁵ cells/ml.

Immunoprecipitations, Immunoblotting, and Antibodies—Whole cell extracts from U937 and Jurkat T cells were prepared as previously described (40). Protein complexes were immunoprecipitated via FLAG or HA epitope tag (38, 41). Cell extracts and immune complexes were separated by SDS-PAGE along with SeeBluePlus2 Prestained Proteins Standard (ThermoFisher) and transferred to PVDF membrane for immunoblotting. Unless indicated otherwise, volume-normalized amounts of extracts prepared from equivalent numbers of cells were resolved by SDS-PAGE. Proteins were detected with antibodies specific for epitope tags: anti-HA (12CA5) and anti-FLAG M2 (Sigma) or reacting directly with CPSF6 (ab99347 Abcam, F3, sc376228 Santa Cruz Biotechnology), CPSF7 (A301–359A, Bethesda; sc79426, Santa Cruz Biotechnology), and CPSF5 (2203C3, sc81109; Santa Cruz Biotechnology). Immune complexes were revealed with HRP-conjugated antibodies specific for the Fc fragment of mouse or rabbit immunoglobulin G (Jackson ImmunoResearch Laboratories) and enhanced chemiluminescence (Amersham Biosciences).

Glycerol Gradient Centrifugation—Eluates from FLAG-M2 beads (Sigma) were overlaid onto a 5–ml 5–40% glycerol gradient in LB buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.7% CHAPS, 5% glycerol) supplemented with Protease Inhibitor Mixture (Sigma) centrifuged for 8 h at 4 °C at 45,000 rpm in an SW55 Ti rotor, and fractions (0.5 ml) were collected from the
top of the gradient. Aliquots of fractions were resolved by SDS-PAGE and analyzed by immunoblotting. Gradients were calibrated by sedimenting a set of standards including albumin (66 kDa), aldolase (158 kDa homotetramer of 40-kDa subunits), catalase (250-kDa homotetramer of 62.5-kDa subunits), and thyroglobulin (669-kDa homotetramer of 165-kDa subunits).

**Capsid Binding Assays—**Recombinant HIV-1 NL4–3 CA and its N74D variant with cysteine substitutions at positions 15 and 45 expressed in *Escherichia coli* and purified as previously described (42, 43) were kindly provided by Drs. Jinwoo Ahn and Chris Aiken. High molecular weight CA assemblies were produced by mixing 26 μl of PBS containing 100 μg of CA, 1 mM DTT, 0.02% NaN₃, 10% glycerol, supplemented with Complete Protease Inhibitor (Sigma), with equal volume of CA assembly buffer composed of 50 mM Tris-HCl, pH 8.0, 1 mM NaCl, 15 mM β-mercaptoethanol followed by incubation at 37 °C for 2 h or at 4 °C overnight. Whole cell extracts were prepared by extracting TPH-1, HEK293T cells, or HEK293T.CPSF6.KO cells with LB buffer. Aliquots of extracts (4 mg, 300 μl) were mixed with 20 μg of capsid assemblies. The total volume was adjusted to 1 ml with PBS in 10% glycerol, and the binding reactions were incubated at 4 °C overnight. Next, binding reactions were overlaid onto a 100-μl, 35% sucrose in a PBS cushion in a 1.5-ml microcentrifuge tube. The tubes were centrifuged at 4750 rpm for 30 min at 4 °C in a Beckman Allegra centrifuge. The pellets were washed twice with 500 μl of PBS, 0.1% Triton X-100, 10% glycerol followed by centrifugation at 3,000 rpm for 5 min to recover the capsid assemblies and their associated proteins.

**CPSF6 Add-back and Flow Cytometry—**HEK293T.CPSF6. KO cells (5 × 10⁶ cells/well of 12-well plate) were spinoculated with MSCV(GFP) expressing HA-CPSF6 or HA-CPSF6(M116→122) in the presence of Polybrene (4 μg/ml) at 2437 × g for 90 min. Two days later cells were harvested for immunoblot analysis of CPSF6 expression or challenged with HIV-1tagRFP and harvested 2 days later for flow cytometry analysis or 5 days later for integration site mapping. For flow cytometry analyses cells were harvested by trypsinization, and GFP and RFP fluorescence was recorded with LSRFortessa (BD Biosciences) and analyzed with FlowJo software (Tree Star, Inc.).

**Integration Site Analysis—**Sequencing of HIV-1 integration sites was conducted essentially as previously described (44). Briefly, 10 μg of genomic DNA from each infected cell culture was digested overnight at 37 °C with 100 units each of Msel and BgII restriction endonucleases and purified the next day with the QIAquick PCR Purification kit (Qiagen). Double-stranded, asymmetric linkers containing a compatible 5′-TA overhang were prepared by heating 10 μM concentrations of each oligonucleotide strand in 35 μl of 10 mM Tris-HCl, pH 8.0, 0.1 mM EDTA and slowly cooling to room temperature in steps of 1 °C per min. For parallel ligation reactions per each DNA sample contained 1.5 μM linker, 1 μg of digested DNA, and 800 units of T4 DNA ligase. Ligation was conducted overnight at 12 °C and purified the next day with the QIAquick PCR purification kit. Semi-nested PCR was then carried out in quadruplicate using primers specific to both the linker sequence and the HIV-1 U5 sequence. The first and second rounds of PCR utilized nested U5 primers, whereas the same linker-specific primer was used for both rounds. The linker primer and second round U5 primer each encoded adapter sequences necessary for DNA clustering as well as sequencing primer binding sites. To afford the identification of unique library samples from multiplexed sequencing runs, unique linker DNAs and linker-specific primers were employed for each sample, and the nested U5 primer additionally encoded a unique 6-bp index sequence. The sequences of oligonucleotides will be provided upon request. PCRs, which contained 100 ng of template DNA, 1.9 μM U5 primer, and 0.375 μM linker primer, were carried out using the Advantage 2 Polymerase Mix (Clontech) according to the manufacturer’s instructions. PCRs were purified with the QIAquick PCR purification Kit, and second round reactions were submitted to the Dana-Farber Cancer Institute Molecular Biology Core Facilities for quality control and Illumina MiSeq single-end 150-bp next-generation sequencing. Resulting sequences were mapped to the hg19 version of the human genome using BLAT, allowing for a minimum of 97% sequence identity match. Correlations of integration site distributions relative to various genomic features was conducted primarily using BEDTools (45). Statistical analysis of the resulting integration frequencies was determined using R (44), with p values being calculated by Fisher’s exact test and Wilcoxon rank sum test. The matched random control was constructed from 50,000 computer-generated sites as described (44).

**Results**

**CPSF6 Is Sequestered into the CFIₘ Complex—**CPSF6 is a subunit of a heterotetrameric CFIₘ complex (Fig. 1A). A crystal structure of a CPSF6 peptide bound to its docking site on the CFIm complex is not known. As a first step to address this question, we assessed CPSF6 participation in CFIₘ and other high molecular weight complexes.

U937 monocytes and Jurkat T cells were engineered to stably express HFA-tandem epitope-tagged CPSF6. HFA-CPSF6 was immunoprecipitated together with its associated proteins from whole cell extracts via its FLAG tag and eluted under native conditions. Separation of the eluates by SDS-PAGE followed by immunoblotting revealed that CPSF6 co-precipitated the CPSF5 and CPSF7 CFIₘ subunits from extracts of both U937 and Jurkat T cells as expected (Fig. 1B). The eluted CPSF6 and its associated proteins were further separated by centrifugal sedimentation through a 5–40% glycerol gradient. Immunoblot analysis of the gradient fractions informed that the vast majority of CPSF6 (~68 kDa) co-sedimented with CPSF5 (25 kDa) and CPSF7 (52 kDa) with the levels of all three proteins peaking in fractions 6 and 7 (Fig. 1C). The distribution of the three CFIₘ complex subunits in the gradient was consistent with that expected for CPSF6-containing CFIₘ tetramers comprising either CPSF6-CPSF5(2)-CPSF7 (172 kDa) or CPSF6(2)-CPSF5(2) (178 kDa) subunits. A slight broadening of the CFIₘ peak toward the bottom of the gradient could be caused by the association of a subset of complexes with one or more of the previously identified CFIₘ binding partners (34). Significantly, no CPSF6 was detected at the gradient position expected for free CPSF6 monomer (fraction 3). Together, these data are con-
CPSF6 Subunit of CFI\textsubscript{m} Directs HIV-1 to Genes

**FIGURE 1. HIV-1 capsid assemblies recruit CFI\textsubscript{m} complex.** A, schematic representation of the CFI\textsubscript{m} complex. CFI\textsubscript{m} heterotrimer comprises a CPSF5 homodimer and a combination of two large CPSF6 and/or CPSF7 subunits. B, CPSF6 co-precipitates CPSF5 and CPSF7. HFA-tagged CPSF6 and its associated proteins were immunoprecipitated (IP) from monocytic U937 cells (M) or T lymphocytic Jurkat T cells (T), and immune complexes were analyzed by immunoblotting for CPSF6, CPSF5, and CPSF7 subunits of the CFI\textsubscript{m} complex. C, the vast majority of CPSF6 co-sediments with other subunits of the CFI\textsubscript{m} complex. HFA-tagged CPSF6 was immunopurified from U937 cells, and immune complexes were sedimented through 10–40% glycerol gradients. Aliquots of fractions were immunoblotted for the indicated subunits of CFI\textsubscript{m} complex. The positions of molecular weight standards sedimented in parallel are indicated, and fraction numbers are also indicated below the gel image. D, CPSF6 recruits CFI\textsubscript{m} to HIV-1 CA assemblies. HeLa whole cell extract (Extract) was incubated with A14C/E45C cross-linked wild type (wt) or CPSF6-binding site-mutated (N74D) HIV-1 NL4-3 CA assemblies or in their absence (mock) as a negative control. Proteins associated with CA assemblies were isolated by centrifugation and immunoblotted for the CFI\textsubscript{m} complex subunits. B and D, representative results of three independent experiments are shown.

Consistent with the possibility that the vast majority of CPSF6 is sequestered into the CFI\textsubscript{m} complex.

**HIV-1 Capsid Assemblies Recruit CFI\textsubscript{m} via the CPSF6 Docking Site**—The lack of detectable monomeric CPSF6 suggests that capsid binds CPSF6 that is incorporated into a CFI\textsubscript{m} tetramer. To assess this possibility, we carried out an *in vitro* binding assay using recombinant A14C/E45C HIV-1 CA, which readily forms large cross-linked tubular assemblies that can be used as capsid surrogates in co-sedimentation experiments from whole cell extracts (43). As expected, CPSF6 was specifically precipitated by wild type CA assemblies but not by those with the N74D substitution, which disrupts the CPSF6 docking site in the CA hexamer (Fig. 1D; Ref. 8–10). Significantly, CA assemblies also efficiently precipitated CPSF5 and CPSF7 in a CPSF6-dependent manner. We conclude that HIV-1 capsid recruits the CFI\textsubscript{m} tetramers by binding to their CPSF6 subunit.

**CPSF6(Δ116–122) RRM Domain Mutant Does Not Assemble into CFI\textsubscript{m}**—The findings described above suggest that the known effects of CPSF6 in HIV-1 infection are mediated by the CFI\textsubscript{m} complex. This in turn raised the question of whether the integrity of the CFI\textsubscript{m} complex and/or other subunits has an important role(s) in HIV-1 integration targeting. To address these issues, we first constructed a variant CPSF6 protein that retained binding to CA assemblies but was unable to bind CPSF5 and, therefore, unable to assemble into the CFI\textsubscript{m} complex.

Two mutations in the CPSF6 RRM domain were previously reported to diminish binding to CPSF5 (34, 39). For further *in vivo* studies, we selected one of these mutants, CPSF6(Δ116–122), which harbors a deletion of residues 116–122 residing in the β2-β3 loop in the RRM domain, which was shown to play a key role in CFI\textsubscript{m} complex assembly *in vitro* (Fig. 2A) (39). Importantly, this deletion does not impact the previously identified CA or TNPO3 binding residues (9, 22). First, we assessed the effect of the Δ116–122 deletion on the ability of CPSF6 to associate with other CFI\textsubscript{m} subunits. HA epitope-tagged CPSF6 and the CPSF6(Δ116–122) variant were transiently expressed in HEK293T cells, immunoprecipitated via their HA epitope tags, and analyzed by immunoblotting for their association with CFI\textsubscript{m} complex subunits (Fig. 2B). Whereas full-length CPSF6 co-precipitated CPSF7 and CPSF5, the CPSF6(Δ116–122) variant showed, at most, only minor residual binding to these proteins. Hence, the Δ116–122 RRM deletion effectively abolished CPSF6 incorporation into CFI\textsubscript{m} *in vivo*.

To address the possibility that the CPSF6(Δ116–122) variant participates in other high molecular complexes, the HFA-tagged CPSF6 and CPSF6(Δ116–122) variants were expressed in CPSF6 knock-out (HEK293T.CPSF6.KO) cells by retroviral transduction. The proteins were purified by immunoprecipitation via their FLAG tags and separated by centrifugal sedimentation through a 10–40% glycerol gradient. Immunoblot analysis of the gradient fractions showed that CPSF6 co-sedimented with other subunits of the CFI\textsubscript{m} complex as expected. In contrast, the CPSF6(Δ116–122) variant sedimented at a much slower rate and was detected at the gradient position expected for free CPSF6 monomer (Fig. 2C). Together, these data confirm that the majority of CPSF6(Δ116–122) is not sequestered into CFI\textsubscript{m} or other high molecular weight protein complexes.

**CPSF6(Δ116–122) Binds CA Assemblies in the Absence of Interaction with Other CFI\textsubscript{m} Subunits**—Next, we tested whether the Δ116–122 RRM mutation affects CPSF6 binding to A14C/E45C cross-linked CA assemblies in pulldown assays. The full-length CPSF6 or CPSF6(Δ116–122) proteins were expressed by retroviral transduction in HEK293T.CPSF6.KO cells. Whole cell extracts were incubated with CA assemblies for up to 6 h in a time course experiment, and the bound CPSF6 and CPSF5 were determined by immunoblotting. As shown in
CPSF6 Inhibits HIV-1 Infection in a Manner Dependent on \( C_{Fi_m} \) Binding—We asked whether CPSF6 interaction with CPSF5/7 modulates HIV-1 infection. HA epitope-tagged CPSF6 and CPSF6(\( \Delta 116–122 \)) variant were expressed in HEK293T.CPSF6.KO cells by retroviral transduction in a dose-response experiment. As shown in Fig. 4A, this titration permitted expression of each CPSF6 protein at levels comparable with those seen with the endogenous CPSF6 protein. Next, all transduced cell populations as well as the parental HEK293T.CPSF6.KO cells were challenged with a VSV-G pseudotyped HIV-1tagRFP (HtR) single cycle reporter virus. In parallel, cells were challenged with a control reporter virus, which was isogenic except for the N74D point mutation in capsid (HtR\( (N74D) \)), and transduction was quantified by flow cytometry 2 days later.

As shown in Fig. 4B, expression of full-length CPSF6 had a modest (~2-fold) inhibitory effect on HtR transduction, as measured by the fraction of cells expressing the RFP marker protein (compare flow cytometry panels for CPSF6 shown in the upper row). The inhibition was already seen at ectopic CPSF6 expression levels lower than those of the endogenous CPSF6 (compare lane 7 with lane 1 in Fig. 4A, upper panel) and, therefore, was not likely to be an artifact of CPSF6 overexpression. Significantly, the N74D capsid mutation relieved, to a large extent, the CPSF6-mediated inhibition of HtR infection, indicating that the inhibition involves CPSF6 binding to viral capsid (Fig. 4, B and C). In contrast, the CPSF6(\( \Delta 116–122 \)) variant did not have such an inhibitory effect on HtR infection. We conclude that CPSF6 can moderately inhibit HIV infection in our experimental system and that this inhibition correlates with capsid binding by CPSF6 and could involve the presence of other \( C_{Fi_m} \) complex subunits, albeit to a lesser extent.

CPSF6 Subunit of \( C_{Fi_m} \) Directs HIV-1 to Genes in Gene-rich Chromosomal Regions—Our findings indicated that CPSF6 docks to capsid while in complex with CPSF5 and/or CPSF7 \( C_{Fi_m} \) subunits. Moreover, the \( C_{Fi_m} \)-incorporated CPSF6 bound capsid better than the CPSF6(\( \Delta 116–122 \)) variant, which does not participate in \( C_{Fi_m} \). Thus, it was conceivable that CPSF5 and/or CPSF7 might facilitate CPSF6-capsid binding and/or play another essential role(s) in directing HIV-1 integration into gene-rich chromosomal regions. To address this latter possibility, we characterized the distribution of HIV-1 integration sites relative to selected chromosomal marks in cells expressing the CPSF5 binding-deficient CPSF6(\( \Delta 116–122 \)) variant and compared it to that in cells expressing full-length, \( C_{Fi_m} \)-incorporated CPSF6.

HtR-transduced HEK293T.CPSF6.KO cell populations ectopically expressing CPSF6(\( \Delta 116–122 \)) or full-length CPSF6 at near endogenous levels (shown in Fig. 4) were selected for the integration site analysis. Parenteral HEK293T and HEK293T.CPSF6.KO cells transduced with an empty MSCV-GFP vector at a comparable multiplicity of infection provided additional controls. The analysis of integration sites was carried out as previously described (44). In brief, chromosomal sequences linked to the U5 region of the downstream HIV-1 LTR were identified by high throughput sequencing and aligned to human genome build 19 (hg19), and the loci were annotated for their locations in gene bodies, the vicinity of transcription start sites (TSSs), CpG islands, and gene density around the integration sites (see Table 1 and Fig. 5). As previously shown, the vast majority of the integration sites in parental HEK293T cells were within reference genes (81.9%; Ref. 15). The frequency of such integration events was significantly lower in cells lacking CPSF6 (57.0%) but was restored almost completely when full-length CPSF6 expression was restored.

![Image](image_url)
Remarkably, the CPSF6(Δ116–122) variant also supported preferential integration of HIV-1 into genes, almost to the levels seen with full-length CPSF6 (73.2%). These values were statistically significant compared with cells lacking CPSF6, indicating that CPSF6(Δ116–122) retains the ability to target HIV-1 integration into genes in the absence of binding to other subunits of the CFIm complex.

CPSF6(Δ116–122) also efficiently rescued other characteristics of CPSF6-mediated HIV-1 integration site distribution, including proximal targeting of TSSs and CpG islands and preferential integration into gene-dense chromosomal regions (Fig. 5). In particular, the frequencies of HIV-1 integration in the proximity of TSSs in cells with full-length CPSF6 (73.2%) and in cells with the CPSF6(Δ116–122) variant (5.1% and 3.9%) were close to those in control cells expressing endogenous CPSF6 (4.6%), all being much higher than those in cells lacking CPSF6 expression (1.6%). The frequencies of HIV-1 integration within 2.5 kb of CpG-rich islands in cells with ectopic full-length or variant CPSF6 (6.6% or 5.3%) were similar to the frequency observed in parental cells (6.1%), and each of these values was well above that seen with CPSF6 knock-out cells (0.9%). Finally, the average gene density within 1 Mb of the integration site in cells with CPSF6(Δ116–122) (18.7) was similar to those in ectopic or endogenous CPSF6-expressing cells (20.6 or 21.4), all the values far higher than that seen in the absence of CPSF6 expression (6.0). Thus, the CPSF6(Δ116–122) variant supported almost in full the previously reported preference of HIV-1 for integration into chromosomal regions that are relatively enriched in genes despite being excluded from the CFIm complex.

Similar results were obtained in an independent experiment in which experimental and control cell populations were transduced with a minimal HIV-1-based LKO-1 lentiviral vector (46). As shown in Table 1, CPSF6(Δ116–122) efficiently mediated LKO-1 integration into genes in gene-enriched chromosomal regions (70.5% versus 73.5% and 56.1% and 17.7 versus 17.9 and 6.1 genes/Mb, in variant versus full-length expressing KO cells and in KO cells). CPSF6(Δ116–122) also supported the characteristic distribution of LKO-1 integration sites in the vicinity of TSS and CpG islands (TSS: 3.8% versus 4.3% and 1.8%; CpG: 5.1% versus 5.5% and 1.0% in the CPSF6 variant versus full-length expressing KO cells and control KO cells). Together, these findings indicate that CPSF6 alone possesses the determinants that direct preferential HIV-1 integration into genes, with other CFIm subunits being dispensable for this process.

Discussion

HIV-1 favors integration into active genes and gene-enriched regions of host cell chromosomes (5). Whereas this characteristic HIV-1 integration pattern is known to be conferred in part
by CPSF6 binding to HIV-1 capsid (15), the underlying biochemical events remain poorly defined. Here, we studied the roles of CFI<sub>m</sub> complex subunits, the abundant binding partners of CPSF6, in this pathway. We show that the CFI<sub>m</sub> complex accounts for most of the cellular CPSF6 pool. This finding strongly suggests that the CFI<sub>m</sub> complex mediates the known effects of CPSF6 in HIV-1 infection. Our studies also reveal, somewhat surprisingly, that CPSF6 is the only subunit of the CFI<sub>m</sub> complex that is required for preferential HIV-1 integration into genes. CPSF6 incorporation into CFI<sub>m</sub> is not essential, and other subunits of this complex appear to have only a minor, if any, role in this process even

**FIGURE 4.** Reconstitution of CPSF6(Δ116–122) expression to near wild type levels in HEK293T.CPSF6.KO cells. A, immunoblot analysis of ectopic CPSF6 expression in HEK293T.CPSF6.KO cells. HEK293T.CPSF6.KO cells were transduced with MSCV.GFP retroviral vectors expressing full-length CPSF6 (upper panel) or CPSF6(Δ116–122) (lower panel) in a dose-response experiment, and whole cell extracts were analyzed by immunoblotting for the CFI<sub>m</sub> complex subunits. Serial 2-fold dilutions of extracts from HEK293T cells expressing endogenous CPSF6 (lanes 1–4) provided a reference for the quantification of ectopic full-length CPSF6(ec) and CPSF6(Δ116–122) (the latter indicated as CPSF6.Δ(ec)) levels in extracts from HEK293T.KO cells that were reconstituted to express these proteins at different levels (lanes 5–8). HEK293T.CPSF6.KO cell extract was also analyzed as a control (lane 8). Aliquots of extracts containing 12 μg of proteins were resolved in lanes 1 and 5–8 and 6 μg, 3 μg, and 1.5 μg in lanes 2–4, respectively. B, effects of wild type and Δ116–122-deleted CPSF6 on HIV-1 infection. HEK293T.CPSF6.KO cell populations reconstituted to express full-length CPSF6 or the CPSF6(Δ116–122) variant, characterized in panel A, were challenged with VSV-G pseudotyped single cycle HIV-1tagRFP reporter virus containing wild type (HTR) or N74D-substituted (HTR<sub>N74D</sub>) capsid, and transduction was characterized 2 days later by flow cytometry. The dot plots show fluorescence of GFP marker protein, correlated with CPSF6 expression levels, on the y axis and of RFP marker protein, expressed from HIV-1tagRFP reporter viruses, on the x axis. A representative of three independent experiments summarized in panel C is shown. C, CPSF6 inhibited HIV-1 infection in a capsid and CFI<sub>m</sub> binding-dependent manner. HIV-1 transduction efficiencies observed in three independent experiments such as that shown in panels A and B are plotted as a function of CPSF6 or CPSF6(Δ116–122) expression levels. The values were normalized to that in the absence of CPSF6 expression.

**TABLE 1**

Effects of CPSF6 and CPSF6(Δ116–122) variant on HIV-1 integration site targeting

Tabulation of HIV-1tagRFP and minimal HIV-1-based vector LKO-1.RFP integration sites in HEK293T.CPSF6.KO cells was reconstituted to express full-length CPSF6 or CPSF6(Δ116–122) variant at near endogenous levels. The numbers of unique integration sites and their distribution relative to RefSeq gene bodies, TSSs, CpG islands, and gene density in the vicinity of integration loci are indicated. The results of experiments with HIV-1tagRFP are shown in Fig. 5. MRS, matched random control.

| Cell line/CPSF6 | HIV-1 vector | Unique sites | Within RefSeq genes | Within 5 kb of CpG | Within 5 kb of TSS | Gene density within 1 Mb |
|-----------------|--------------|--------------|---------------------|--------------------|---------------------|------------------------|
| CPSF6 KO/CPSF6 | HIV-1tagRFP  | 35,842       | 27,371 (76.4)       | 2,352 (6.6)        | 1,813 (5.1)        | 20.6                   |
| CPSF6 KO/CPSF6(Δ116–122) | HIV-1tagRFP | 21,801       | 15,965 (73.2)       | 1,148 (5.3)        | 851 (3.9)          | 18.7                   |
| CPSF6 KO(−)    | HIV-1tagRFP  | 50,610       | 28,837 (57.0)       | 453 (9.9)          | 805 (1.6)          | 6.0                    |
| CPSF6 WT(−)    | HIV-1tagRFP  | 68,267       | 55,892 (81.9)       | 4,142 (6.1)        | 3,109 (4.6)        | 21.4                   |
| CPSF6 KO/CPSF6 | LKO-1.RFP    | 18,154       | 13,344 (73.5)       | 1,007 (5.5)        | 797 (4.3)          | 17.9                   |
| CPSF6 KO(−)/LKO-1.RFP | 30,156 | 21,247 (70.5) | 1,539 (5.1) | 1,131 (3.8) | 17.7 |
| CPSF6 WT(−)/LKO-1.RFP | 38,903 | 21,821 (56.1) | 390 (1.0) | 682 (1.8) | 6.1 |
| CPSF6 KO/−     | LKO-1.RFP    | 10,505       | 8,235 (78.4)        | 826 (7.9)          | 618 (5.9)          | 23.6                   |
| MRC             | LKO-1.RFP    | 50,000       | 22,326 (44.7)       | 2,100 (4.2)        | 2010 (4.0)         | 8.7                    |
though our data support the possibility that they facilitate CPSF6 binding to capsid.

Our data indicate that the vast majority of CPSF6 is sequestered in the CFIm complex. In particular, glycerol gradient centrifugation studies demonstrated that CPSF6 co-sediments with other CFIm subunits. The possibility that a major fraction of CFIm is sequestered in another protein complex is not supported by our proteomic analyses of CPSF6-associated proteins, which did not identify abundant binding partners that could constitute such a complex (data not shown). Moreover, we show that the vast majority of CPSF6(ΔH9004116–122) variant, which does not participate in CFIm, behaves like a monomer during centrifugal sedimentation, consistent with the possibility that the bulk of CPSF6 does not participate in a high molecular weight protein complex(es) other than CFIm. The above evidence and our finding that CPSF6 recruits CPSF5 and CPSF7 to CA-assemblies indicate that CFIm mediates the effects of CPSF6 at early steps in HIV-1 infection.

Whereas previous studies focused on the interaction of CPSF6 alone with its docking site in CA-hexamers, our findings suggest that CPSF5 and possibly also CPSF7 contribute to CPSF6 binding to capsid. Data in Fig. 3 clearly show that CPSF6 binding to A14C/E45C cross-linked CA assemblies was enhanced when it was incorporated into CFIm together with CPSF5 and CPSF7. The latter subunits may contact the CA directly, thus synergizing with CPSF6 binding, or contribute to the binding indirectly by improving the presentation of the CA binding determinant in CPSF6. Of note, whereas the possibility that the A14C/E45C cross-link modifies the interaction with full-length CPSF6 and/or CFIm complex cannot be excluded, the published CA structures do not raise a concern that this modification could interfere with CPSF6 binding via the previously defined canonical CA-binding peptide (7, 47, 48). It is also not likely that the small RRM domain deletion present in the CPSF6(ΔH9004116–122) variant has a direct negative effect on binding to CA assemblies, as this deletion was reported not to result in a detectable misfolding of the variant protein (39). Furthermore, this deletion in the RRM and the peptide that docks to the binding pocket in CA-hexamers (CPSF6 residues 314–326) are located almost 200 residues apart in distinct domains of the CPSF6 molecule. Together, our evidence suggests a model in which CPSF5 acts in concert with CPSF6 in binding to HIV-1 capsid. Whether or not CPSF5 binds CA directly, the CA interface for CPSF6 and/or CFIm binding is likely to be larger than that defined in the co-crystal structures of CA with CPSF6-derived pocket-binding peptides (7, 10).

The observation that CFIm binding to capsid is associated with a modest inhibition of HIV-1 infection was unexpected.

FIGURE 5. HIV-1 integration patterns in control- and CPSF6(Δ116–122)-expressing HEK293T cells. A, percent of HIV-1tagRFP integration sites within RefSeq genes, within 5 kb of TSS or CpG islands (CpG) in HEK293T cells expressing endogenous or ectopic CPSF6 proteins. The CPSF6 genotypes of the target cell populations are shown below the histograms: C6(en), endogenous CPSF6 gene; C6(ec), retrovirally transduced ectopic CPSF6; (−), transduced with a control empty vector; ko, knock-out; wt, wild type; Δ, CPSF6(Δ116–122). B, percent integration sites (y axis) plotted as a function of genes per Mb (x axis) for infections of HEK293T populations expressing endogenous or ectopically expressed wild type and mutant CPSF6 proteins. MRC, matched randomized control. C, statistic analysis of the results of integration site mapping experiments. a, p values for the analyses of integrations into RefSeq genes and in the vicinity of CpG or TSS were calculated by Fisher’s exact text. b, p values for integrations as a function as gene dense (G.D.) were calculated by Wilcoxon Rank-Sum test. The superscripts refer to p values calculated by (a) Fisher’s exact test or (b) Wilcoxon rank-sum test.
This inhibition was relieved by CPSF6 mutation that disrupts binding to CPSF5 or by capsid N74D mutation that disrupts CPSF6 binding, indicating that it results from the binding of CFI\textsubscript{m}, not CPSF6 alone, to the viral capsid. The capsid CPSF6 binding pocket, formed by the N/C-terminal domain interface in the CA-hexamer, was implicated in controlling capsid stability (7, 8, 10, 23, 24, 49). Because our data suggest that CFI\textsubscript{m} incorporate, and free CPSF6 bind to CA assemblies in different modes, these differences may have different outcomes for processes controlled by the N/C-terminal domain interface.

Our studies demonstrate that CPSF6 targets HIV-1 integration into genes whether or not it is incorporated into the CFI\textsubscript{m} complex, implying that CPSF5 and CPSF7 CFI\textsubscript{m} subunits have only minor, if any, roles in CPSF6-dependent HIV-1 PIC integration targeting. Whereas the possibility remains that the CPSF6(Δ116–122) retains residual CPSF5 binding, it is evident that CPSF5 recruitment to capsid by this variant is almost completely abolished (See Fig. 3). Therefore, we do not favor the possibility that CPSF6(Δ116–122) requires CFI\textsubscript{m} to target HIV integration into genes.

We note that in the rescue experiments CPSF6(Δ116–122) was slightly less effective than the full-length ectopic CPSF6 in supporting HIV-1 integration preferences, and the differences were statistically significant. This deficiency could be due to the poorer binding of the CPSF6(Δ116–122) variant to capsid, revealed by our CA assembly pulldown experiments, resulting in less effective recruitment of the PIC to the CPSF6-controlled integration targeting pathway and, as a consequence, slightly lower frequency of integration into genes, compared with that seen for the full-length CFI\textsubscript{m}-incorporated CPSF6. Significantly, however, re-expression of full-length CPSF6 in the CPSF6.KO cells also did not fully rescue the distribution of HIV-1 integration sites seen in the parental cell line expressing the endogenous CPSF6 protein. Although the exact reason for this is not known, a slightly different CPSF6 expression level, the presence of the N-terminal epitope tag in the ectopic CPSF6 or other experimental modalities that are difficult to fully control could be responsible for the observed differences. Overall, our evidence indicates that CPSF6 is the CFI\textsubscript{m} subunit that controls the key molecular interactions that specify the trafficking of the PIC to active chromatin and that the CPSF5 and CPSF7 subunits contribute little to these processes.

Whereas the exact biochemical events targeting the PIC to active genes for integration remain elusive, we show that all the key interactions resulting from CFI\textsubscript{m} binding are likely carried out by the CPSF6 subunit of this complex alone. Our prior bioinformatic analysis of the role of CPSF6 in HIV-1 integration targeting had already suggested that its function in integration targeting is independent from its known role in 3\textsuperscript{-}UTR length determination, which is a CFI\textsubscript{m}-dependent process (31, 32, 50). Although CPSF6 was found to play a role in directing integration into genes that contained 3\textsuperscript{-}UTRs, integration into genes with CPSF6-dependent changes in 3\textsuperscript{-}UTRs was unaffected by CPSF6 knock-out (50). Here, we utilized the biochemical properties of CFI\textsubscript{m} binding defective CPSF6(Δ116–122) mutant to reach a similar conclusion.

CPSF6 is known to be recruited to sites of RNA PolIII-mediated transcription and co-transcriptional RNA processing (51, 52), and CPSF6-mediated trafficking of the PIC to active genes is probably executed by the same molecular mechanism(s). Indeed, CPSF6-mediated import of the PIC to the nucleus requires TNPO3, which transports CPSF6 and other SR-family proteins to the nucleus (20–22). Whether CPSF6 interaction with TNPO3 is sufficient to direct the PIC to the vicinity of active genes or interaction with an additional partner(s) is required is not known. If the latter, potential candidates include members of the SR family of splicing factors, such as SRp20, 9G8, and hTRA2b, which bind the CPSF6 C-terminal SR domain (34), or Thoc5, a subunit of the TREX complex, which couples mRNA transcription and processing, for the control of polyadenylation site choice and was proposed to recruit CPSF6 co-transcriptionally to active genes (53). One or more CPSF6 binding partner or other proteins critical for CPSF6 nuclear function could connect the PIC to the vicinity of active transcription units for integration. Future studies will aim to identify the precise mechanism of PIC intranuclear trafficking initiated by CPSF6 binding to HIV-1 capsid.

Author Contributions—J. S. and A. N. E. conceived and coordinated the study and wrote the paper. S. R., M. S., J. Q., and C. H. designed, performed, and analyzed the experiments shown in Figs. 1, 2, and 4. J. Q. and T. D. performed the experiment shown in Fig. 3. E. S. performed the experiments and analyzed the data for Fig. 5 and Table 1. G. A. S. and A. N. E. contributed new reagents. All authors reviewed the results and approved the final version of the manuscript.

Acknowledgments—We thank Selene Swanson, Laurence Flores, and Michael Washburn for proteomic characterization of CPSF6 associated proteins, Jinwoo Ahn and Chris Aiken for providing recombinant capsid proteins, Nicolas Manel for the gift of HIV-1tagRFP, Chunaping Wang for excellent technical assistance, and Junpeng Yan and Weifeng Wang for help with experiments at various stages of these studies, and we are indebted to Teresa Brosenitsch for critical reading of the manuscript and editorial help.

References
1. Mitchell, R. S., Beitzel, B. F., Schroder, A. R., Shinn, P., Chen, H., Berry, C. C., Ecker, J. R., and Bushman, F. D. (2004) Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. PLoS Biol. 2, e234
2. Cherepanov, P., Maertens, G., Proost, P., Devreeze, B., Van Beeumen, J., Engelborghs, Y., De Clercq, E., and Debyser, Z. (2003) HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. J. Biol. Chem. 278, 372–381
3. Ciuffi, A., Llanos, M., Poeschla, E., Hoffmann, C., Leipzig, J., Shinn, P., Ecker, J. R., and Bushman, F. (2005) A role for LEDGF/p75 in targeting HIV DNA integration. Nat. Med. 11, 1287–1289
4. Lewinski, M. K., Yamashita, M., Emerman, M., Ciuffi, A., Marshall, H., Crawford, G., Collins, F., Shinn, P., Leipzig, J., Hannenhalli, S., Berry, C. C., Ecker, J. R., and Bushman, F. D. (2006) Retroviral DNA integration: viral and cellular determinants of target-site selection. PLoS Pathog. 2, e60
5. Schroder, A. R., Shinn, P., Chen, H., Berry, C., Ecker, J. R., and Bushman, F. (2002) HIV-1 integration in the human genome favors active genes and local hotspots. Cell 110, 521–529
6. Ambrose, Z., Lee, K., Ndjomou, J., Xu, H., Oztos, J., Matous, J., Takemura, T., Unutmaz, D., Engelman, A., Hughes, S. H., and KewalRamani, V. N. (2012) Human immunodeficiency virus type 1 capsid mutation N74D alters cyclophilin A dependence and impairs macrophage infection. J. Virol. 86, 4708–4714
7. Bhattacharya, A., Alam, S. L., Fricke, T., Zadrozny, K., Sedzicki, J., Taylor,
CPSF6 Subunit of CFIm Directs HIV-1 to Genes

A. B. Domeler, B. Pornillos, O. Ganser-Pornillos, B. K. Diaz-Griñero, F. Ivanov, D. N., and Yeager, M. (2014) Structural basis of HIV-1 capsid recognition by PF74 and CPSF6. Proc. Natl. Acad. Sci. U.S.A. 111, 18625–18630.

8. Lee, K., Ambose, Z., Martin, T. D., Oztop, L., Mulky, A., Julias, J. G., Vandegraaff, N., Baumann, J. G., Wang, R., Yuen, W., Takemura, T., Shelton, K., Taniuchi, I., Li, Y., Sodroski, J., Littman, D. R., Coffin, J. M., Hughes, S. H., Unutmaz, D., Engelman, A., and KselawiRamani, V. N. (2010) Flexible use of nuclear import pathways by HIV-1. Cell Host Microbe 7, 221–233.

9. Lee, K., Mulky, A., Yuen, W., Martin, T. D., Meyerson, N. R., Choi, L., Yu, H., Sawyer, S. L., and KselawiRamani, V. N. (2012) HIV-1 capsid-targeting domain of cleavage and polyadenylation specificity factor 6. J. Virol. 86, 3851–3860.

10. Price, A. J., Fricke, A. J., Schaller, T., Elliott, T., Lee, K., KselawiRamani, V. N., Chin, I. W., Towers, G. J., and James, L. C. (2012) CPSF6 defines a conserved capsid interface that modulates HIV-1 replication. PLoS Pathog. 8, e1002966.

11. De Iaco, A., Santoni, F., Vannier, A., Guipponi, M., Antonarakis, S., and De Iaco, A. (2013) Role of human immunodeficiency virus type 1 capsid and cellular factors nucleoprin 53 and LEDGF/p75 in the efficiency and specificity of viral DNA integration. J. Virol. 87, 648–658.

12. Cechova, K. E., Brady, T. L., Ronen, K., Hugel, A., Roth, S. L., Schaller, T., James, L. C., Towers, G. J., Young, A. J., Chanda, S. K., König, R., Malani, N., Berry, C. C., and Bushman, F. D. (2011) HIV integration targeting: a pathway involving Transportin-3 and the nuclear pore protein RanBP2. PLoS Pathog. 7, e1002133.

13. Schaller, T., Cechova, K. E., Rasaiaja, J., Price, A. J., Brady, T. L., Roth, S. L., Hugel, A., Schaller, T., James, L. C., Towers, G. J., Young, A. J., Chanda, S. K., König, R., Malani, N., Berry, C. C., and Bushman, F. D. (2011) HIV integration targeting: a pathway involving Transportin-3 and the nuclear pore protein RanBP2. PLoS Pathog. 7, e1002133.

14. Gruber, A. R., Martin, G., Keller, W., and Zavolan, M. (2012) Cleavage factor Im is a key regulator of 3′-UTR length. RNA Biol. 9, 1465–1412.

15. Millevoi, S., Loulergue, C., Dettwiler, S., Karas, S. Z., Antoniou, M., and Vagner, S. (2006) An interaction between U2AF65 and CFIm150 links the splicing and 3′ end processing machineries. EMBO J. 25, 4854–4864.

16. Millevoi, S., Loulergue, C., Dettwiler, S., Karas, S. Z., Antoniou, M., and Vagner, S. (2006) An interaction between U2AF65 and CFIm150 links the splicing and 3′ end processing machineries. EMBO J. 25, 4854–4864.

17. Schrijvers, R., De Rijck, J., Demeulemeester, J., Adachi, N., Vets, S., Ronen, K., Christ, F., Bushman, F. D., Debzer, Y., and Gijsbers, A. W. (2012) LEDGF/p75-independent HIV-1 replication demonstrates a role for HRP-2 and remains sensitive to inhibition by LEDGINs. PLoS Pathog. 8, e1002558.

18. Sun, M. C., Raghavendra, N. K., Vandegraaff, N., Daigle, J. E., Hughes, S. H., Kellam, P., Cherepanov, P., and Engelman, A. (2007) Distinct sequence motifs within the 68-kDa subunit of cleavage and polyadenylation specificity factor 6/5'UTR-25 and implications for a functional role in mRNA 3′ end processing. J. Biol. Chem. 282, 2287–2297.

19. Kataoka, N., Bachior, J. L., and Dreysse, G. (1999) Transportin-SR, a nuclear import receptor for SR proteins. J. Cell Biol. 145, 1145–1146.

20. Lai, M. C., Lin, R. I., Huang, S. Y., Tsai, C. W., and Tarn, W. Y. (2000) A human importin-β family protein, transportin-SR2, interacts with the phosphorylated RS domain of SR proteins. J. Biol. Chem. 275, 7950–7957.

21. Maertens, G. N., Cook, N. J., Wang, W., Hare, S., Gupta, S. S., Öztop, I., Lee, K., Pye, V. E., Cosnefroy, O., Snijders, A. P., KselawiRamani, V. N., Fassati, A., Engelman, A., and Cherepanov, P. (2014) Structural basis for nuclear import of splicing factors by human Transportin 3. Proc. Natl. Acad. Sci. U.S.A. 111, 2728–2733.
Palczewski, K., Gronenborn, A. M., Ahn, J., and Skowronski, J. (2013) Tetramerization of SAMHD1 is required for biological activity and inhibition of HIV infection. *J. Biol. Chem.* **288**, 10406–10417

41. Srivastava, S., Swanson, S. K., Manel, N., Flores, L., Washburn, M. P., and Skowronski, J. (2008) Lentiviral Vpx accessory factor targets VprBP/DCAF1 substrate adaptor for cullin 4 E3 ubiquitin ligase to enable macrophage infection. *PLoS Pathog.* **4**, e1000059

42. Byeon, I. J., Meng, X., Jung, J., Zhao, G., Yang, R., Ahn, J., Shi, J., Concel, J., Aiken, C., Zhang, P., and Gronenborn, A. M. (2009) Structural convergence between Cryo-EM and NMR reveals intersubunit interactions critical for HIV-1 capsid function. *Cell* **139**, 780–790

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

44. Matreyek, K. A., Wang, W., Serrao, E., Singh, P. K., Levin, H. L., and Engelman, A. (2014) Host and viral determinants for MxB restriction of HIV-1 infection. *Retrovirology* **11**, 90

45. Quinlan, A. R., and Hall, I. M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842

46. Moffat, J., Grueneberg, D. A., Yang, X., Kim, S. Y., Kloepfer, A. M., Hinkle, G., Pipani, B., Eisenhaure, T. M., Luo, B., Grenier, J. K., Carpenter, A. E., Foo, S. Y., Stewart, S. A., Stockwell, B. R., Hacohen, N., Hahn, W. C., Lander, E. S., Sabatini, D. M., and Root, D. E. (2006) A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell* **124**, 1283–1298

47. Gres, A. T., Kirby, K. A., KewalRamani, V. N., Tanner, J. J., Pornillos, O., and Sarafianos, S. G. (2015) Structural virology: x-ray crystal structures of native HIV-1 capsid protein reveal conformational variability. *Science* **349**, 99–103

48. Price, A. J., Jacques, D. A., McEwan, W. A., Fletcher, A. J., Essig, S., Chin, J. W., Halambage, U. D., Aiken, C., and James, L. C. (2014) Host cofactors and pharmacologic ligands share an essential interface in HIV-1 capsid that is lost upon disassembly. *PLoS Pathog.* **10**, e1004459

49. Hulme, A. E., Perez, O., and Hope, T. J. (2011) Complementary assays reveal a relationship between HIV-1 uncoating and reverse transcription. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 9975–9980

50. Sowd, G. A., Serrao, E., Wang, H., Wang, W., Fadel, H. J., Poeschla, E. M., and Engelman, A. N. (2016) 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

51. Dantonel, J. C., Murthy, K. G., Manley, J. L., and Tora, L. (1997) Transcription factor TFIIID recruits factor CPSF for formation of 3’ end of mRNA. *Nature* **390**, 399–402

52. Hsin, J. P., and Manley, J. L. (2012) The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev.* **26**, 2119–2137

53. Katahira, J., Okuzaki, D., Inoue, H., Yoneda, Y., Maehara, K., and Ohkawa, Y. (2013) Human TREX component Thoc5 affects alternative polyadenylation site choice by recruiting mammalian cleavage factor I. *Nucleic Acids Res.* **41**, 7060–7072

54. Chin, C. R., Perreira, J. M., Savidis, G., Portmann, J. M., Aker, A. M., Feeley, E. M., Smith, M. C., and Brass, A. L. (2015) 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