Decoding the biogenesis of HIV-induced CPSF6 puncta and their fusion with the nuclear speckle.

Chiara Tomasini[1*], Celine Cuche[1*], Selen Ay[1], Maxence Collard[1], Bin Cui[2], Mohammad Rashid[2], Shaoni Bhattacharjee[1], Bruno Tello-Rubio[1], Julian Buchrieser[3], Charlotte Luchsinger[2], Cinzia Bertelli[2], Vladimir N. Uversky[4], Felipe Diaz-Griffero[2#], Francesca Di Nunzio[1#]

[1] *Institut Pasteur, Advanced Molecular Virology Unit, Department of Virology, Université Paris Cité, 75015 Paris, France*

[2]*Albert Einstein College of Medicine, Department of Immunology and Microbiology, New York 10461, USA*

[3] *Institut Pasteur,* Virus and Immunity Unit*, Department of Virology, Université Paris Cité, 75015 Paris, France.*

[4] *Department of Molecular Medicine and USF Health Byrd Alzheimer's Research Institute, Morsani College of Medicine, University of South Florida, Tampa, FL 33612, USA*

 # Address correspondence to Francesca Di Nunzio, dinunzio@pasteur.fr, or Felipe Diaz-Griffero, felipe.diaz-griffero@einsteinmed.edu

** These authors have equally contributed*

Highlights

- The formation of HIV-induced CPSF6 puncta is critical for restoring HIV-1 nuclear reverse transcription.
- CPSF6 protein lacking the FG peptide cannot bind to the viral core, thereby failing to form HIV-induced CPSF6 puncta.
- The FG peptide, rather than low-complexity regions (LCRs) or the mixed charge domains (MCDs) of the CPSF6 protein, drives the formation of HIV-induced CPSF6 puncta.
- The presence of the FG peptide plays a pivotal role in facilitating viral replication within macrophage-like cells.
- HIV-induced CPSF6 puncta form individually and later fuse with nuclear speckles (NS) via the intrinsically disordered region (IDR) of SRRM2.

Summary

Viruses rely on host cellular machinery for replication. After entering the nucleus, the HIV genome accumulates in nuclear niches where it undergoes reverse transcription and integrates into neighboring chromatin, promoting high transcription rates and new virus

progeny. Despite antiretroviral treatment, viral genomes can persist in these nuclear niches and reactivate upon treatment interruption, raising the possibility that they could play a role in the establishment of viral reservoirs. The post-nuclear entry dynamics of HIV remain unclear, and understanding these steps is critical for revealing how viral reservoirs are established.

In this study, we elucidate the formation of HIV-induced CPSF6 puncta and the domains of CPSF6 essential for this process. We also explore the roles of nuclear speckle scaffold factors, SON and SRRM2, in the biogenesis of these puncta. Through genetic manipulation and depletion experiments, we demonstrate the key role of the intrinsically disordered region of SRRM2 in enlarging nuclear speckles in the presence of the HIV capsid.

We identify the FG domain of CPSF6 as essential for both puncta formation and binding to the viral core, which serves as the scaffold for CPSF6 puncta. While the low-complexity regions (LCRs) modulate CPSF6 binding to the viral capsid, they do not contribute to puncta formation, nor do the disordered mixed charge domains (MCDs) of CPSF6. Interestingly, the FG peptide facilitates viral replication. These results demonstrate how HIV evolved to hijack host nuclear factors, enabling its persistence in the host.

Of note, this study provides new insights into the underlying interactions between host factors and viral components, advancing our understanding of HIV nuclear dynamics and offering potential therapeutic targets for preventing viral persistence.

Introduction

Since the discovery of HIV (Barre-Sinoussi et al., 1983), the initial stages of the viral life cycle, such as the reverse transcription (the conversion of the viral RNA in DNA) and the uncoating (loss of the capsid), have been understood to mainly occur within the host cytoplasm. Only the pre-integration complex, carrying the fully reverse-transcribed viral DNA, was believed to enter the nucleus for integration into the host chromatin (Suzuki and Craigie, 2007). Recent studies highlighted that the viral genome is transported in the nucleus via a shuttle that shields it from the hostile cellular environment (Rasaiyaah et al., 2013). This shuttle is constituted by the viral capsid (Blanco-Rodriguez and Di Nunzio, 2021; Blanco-Rodriguez et al., 2020; Chen et al., 2016; Selyutina et al., 2020; Yamashita and Emerman, 2004; Zila et al., 2021), which comprises 250 hexamers (depending on the size and shape of the capsid) and 12 pentamers (Pornillos et al., 2009). Within the structure of the capsid, hydrophobic pockets exist between hexamers, which serve as targets for various nucleoporins, especially the ones carrying FG repeats, facilitating the translocation of the capsid through the nuclear pore complex (NPC) (Buffone et al., 2018; Di Nunzio, 2013; Di Nunzio et al., 2012; Lelek et al., 2015; Matreyek et al., 2013; Price et al., 2014). Recent studies suggest that HIV uses multiple FG regions of several nucleoporins to translocate through the NPC, acting as a chaperone by itself (Dickson et al., 2024; Fu et al., 2024). Using a reductionist system of Nup98 condensates, it has been demonstrated that FG-mediated phase partitioning identifies specific sites on the capsid that allow it to interact autonomously with these phases. These findings complement the evolving understanding of the early stages of HIV infection, which has been revisited in recent years. They unveil that crucial stages of early viral infection occur within the host nucleus (Burdick et al., 2020; Dharan et al., 2020; Francis et al., 2020; Scoca et al., 2023; Selyutina et al., 2020). Notably, it has been shown that the pre-integration complex forms within the host nucleus (Muller et al., 2021; Scoca et al., 2023), and incoming viral RNA genomes accumulate in nuclear niches containing, cleavage and polyadenylation specificity factor subunit 6 (CPSF6), RNA-binding protein SON and Splicing component, 35 kDa (SC35, also known as serine/arginine-rich splicing factor 2 (SRSF2) (Francis et al., 2020; Rensen et al., 2021; Scoca et al., 2023). However, the mechanistic requirements that govern post-nuclear entry phases, which are crucial for a successful viral infection and the establishment of viral reservoirs, remain enigmatic.

Specifically, the mechanism underlying the formation of nuclear niches containing viral components and nuclear speckle (NS) factors, such as SC35—traditionally used as a marker for NS remains unclear. However, recent studies have revealed that SON and serine/arginine repetitive matrix protein 2 (SRRM2) are essential for NS biogenesis (Fu and Maniatis, 1990). Particularly, intrinsically disordered regions (IDRs) play a critical role in the NS formation (Ilik

et al., 2020). These membraneless organelles (MLOs) fulfil various cellular functions besides splicing. Recent evidence indicates a direct role of NSs in cellular transcription regulation, as their spatial proximity correlates with gene expression amplification, as demonstrated by live-cell imaging of heat-shock responsive genes (Chen et al., 2018; Zhang et al., 2021). HIV, being a virus capable of generating new particles through splicing and integrating into active host genes, finds NSs highly conducive for viral replication. HIV particles, along with their RNA genome, accumulate within nuclear MLOs enriched in NS factors (Ay et al., 2025; Rensen et al., 2021; Scoca et al., 2023). Notably, CPSF6, a paraspeckle factor first identified as a viral partner by KewalRamani's laboratory (Lee et al., 2010), has been clearly detected in HIV-induced CPSF6 puncta (Ay et al., 2025; Francis et al., 2020; Lee et al., 2010; Luchsinger et al., 2023; Rensen et al., 2021; Scoca et al., 2023). These puncta serve as hubs for nuclear reverse transcription and the formation of pre-integration complexes, which generate active proviruses detected outside but in close proximity to NSs (Ay et al., 2025; Li et al., 2021; Scoca et al., 2023).

In our study, we aim to elucidate how HIV-induced CPSF6 puncta form and identify the NS factors involved in their formation. CPSF6, along with NS factors, contains intrinsically disordered regions that can guide HIV-1 to the correct nuclear location for successful infection or allow the virus to remain sequestered during drug treatment, forming reservoirs. Importantly, viral reservoirs are the major bottleneck for curing the infection.

In this study, we investigate which disordered domain of CPSF6 is responsible for the recruitment of the viral core and generation HIV-induced CPSF6 puncta in the host nucleus. Simultaneously, we elucidate the key component of NSs that, through its IDRs, enable fusion with HIV-induced CPSF6 puncta, likely stabilizing them. Notably, we observed that the virus rebounds when anti-reverse transcription drugs are removed, but only if nuclear niches containing HIV, NS factors, and CPSF6 are present. If these niches are pharmacologically dismantled, viral rebound does not occur.

Overall, studying the biogenesis of HIV-induced nuclear niches is crucial for understanding how the virus navigates and persists in the nucleus of infected cells, and for designing new antiretroviral strategies.

Results

Critical Role of HIV-Induced CPSF6 Puncta in Restoring Nuclear reverse transcription After Anti-Reverse transcriptase (RT) Therapy Discontinuation.

127 Upon nuclear entry, HIV enhances the formation of CPSF6 puncta, where reverse transcription
128 ends. The treatment with the reversible RT inhibitor nevirapine (NVP) can trap the viral RNA
129 genome in these nuclear niches (Ay et al., 2025; Rensen et al., 2021; Scoca et al., 2023).
130 Once NVP is removed, the trapped vRNA can resume RT entirely within the nucleus, a process
131 we term nuclear reverse transcription. Here we demonstrate that this phenomenon depends
132 on the presence of CPSF6 puncta, which are mainly composed of CPSF6 proteins and viral
133 cores. Disruption of these puncta by high doses of PF74 (25 μM) significantly impairs the
134 rebound of the RT, as evidenced by the absence of luciferase expression, similar to what is
135 observed with complete NVP treatment (Figure 1A-B).

136

137 **CPSF6 FG Domain is Required for HIV-Induced Puncta Formation.**

138 Formation of CPSF6 puncta upon HIV-1 infection hinges on two key events: the entry of the
139 HIV-1 core into the nucleus and the binding of CPSF6 to the HIV-1 core (Blanco-Rodriguez
140 and Di Nunzio, 2021; Blanco-Rodriguez et al., 2020; Buffone et al., 2018; Zila et al., 2021). To
141 determine the contribution of CPSF6's disordered domains for the formation of CPSF6 puncta
142 upon HIV-1 infection, we correlated the binding of CPSF6 to the HIV-1 core with the formation
143 of CPSF6 puncta. To this end, we first generated CPSF6 knockout (KO) THP-1 cells (Figure
144 1C-D) to eliminate the interference from the endogenous protein, which could affect the
145 interpretation of results regarding the role of the analyzed CPSF6 domains. CPSF6 depletion
146 in THP-1 cells was performed by CRISPR Cas9 technology. To completely eliminate the
147 expression of the CPSF6 gene we selected single clones by limiting dilution. We identified a
148 clone that was completely KO for CPSF6, confirmed through western blot and
149 immunofluorescence (Fig. 1C-D) and we infected this clone and the control clone with HIV.
150 CPSF6 puncta were detected only in the control-infected cells and not in the KO clone (Fig.
151 1D; Suppl. Fig.1A). The viral integrase (IN) was observed within CPSF6 puncta, consistent
152 with previous studies (Francis et al., 2020; Rensen et al., 2021; Scoca et al., 2023), but absent
153 in CPSF6 KO cells where viral IN was observed in the cytoplasm (Fig. 1D; Suppl. Fig.1A).
154 Thus, we used KO cells for CPSF6 to assess the role of selected CPSF6 domains in HIV-
155 induced condensates (Suppl. Fig.1B). We designed various CPSF6 deletion mutants (Figure
156 2A) to specifically assess the significance of the main disordered regions of CPSF6 protein
157 (Suppl. Figure 2A) such as the FG (phenylalanine and glycine) motif, the low complexity
158 regions (LCRs), and the mixed charge domain (MCD), in the ability of CPSF6 to bind to the
159 core and facilitate the formation of CPSF6 puncta. We investigated the role of the FG peptide
160 by generating a mutant that exclusively lacks the FG peptide (ΔFG). Previous *in vitro* studies
161 have shown that the FG peptide binds to the hydrophobic pocket formed between capsid

162 hexamers (Buffone et al., 2018; Price et al., 2014). Here, we want to investigate the role of FG
163 peptide in the context of the protein during the viral infection.

164 To further explore this, we developed an alternative plasmid by expanding the FG peptide
165 deletion to include surrounding prion-like LCRs (ΔFG ΔLCR). These regions, outside the
166 CPSF6 context, have been identified as crucial for facilitating strong CPSF6 binding to capsid
167 lattices (Wei et al., 2022). In our study, we aim to evaluate their role within a more physiological
168 setting. Additionally, we assessed a CPSF6 variant that carries the 15-mer FG peptide flanked
169 by non-LCR sequences, such as those derived from Beta-adducin (ADD2), kindly provided by
170 Mamuka Kvaratskhelia (ΔLCR+ADD2). These protein segments are known for their high
171 flexibility, akin to the LCR of CPSF6. Furthermore, to elucidate the contribution of the LCRs of
172 CPSF6 in the formation of CPSF6 puncta, we generated a mutant lacking both LCRs (ΔLCR).
173 Analysis of the MCD contribution to both the ability of CPSF6 to bind to the core and formation
174 of CPSF6 puncta was achieved by deleting the MCD and adding 3 nuclear localization signals
175 (3xNLS ΔMCD) since the deletion of the MCD results in a protein that localizes mainly into the
176 cytoplasm (Figure 2A-B; Suppl. Figure 2B).

177 To correlate the ability of CPSF6 to bind to the HIV-1 core with formation of CPSF6 puncta,
178 we expressed wild type and mutant CPSF6 constructs in THP-1 cells knockout for CPSF6.
179 Subsequently, we infected these cells with HIV-1 and analyzed the presence or absence of
180 CPSF6 puncta at 24 hours post-infection. Importantly, for the imaging experiment we
181 expressed CPSF6 WT and mutants without tags to avoid the formation of aggregates that
182 could interfere with our conclusions. Our data show that HIV-induced CPSF6 puncta can form
183 extremely rarely with the deletion mutant CPSF6 ADD2ΔLCR and with the mutant lacking the
184 FG or both the FG peptide and the LCRs (Fig. 2B-D; Suppl. Fig.1B). However, when we
185 analyzed the role of the MCD domain in CPSF6 puncta formation, which was indicated to be
186 important for condensing CPSF6 in NS (Greig et al., 2020), comparing the number of CPSF6
187 WT puncta induced by HIV infection with CPSF6 mutants revealed that the MCD domain does
188 not play a critical role in HIV-induced CPSF6 puncta formation (Fig. 2B-D). In addition, we
189 observed that the majority of analysed CPSF6 3xNLSΔMCD puncta contain vRNA inside,
190 similar to CPSF6 WT puncta (Figure 2E), thus corroborating the lack of a role for this
191 intrinsically disordered domain in HIV-induced CPSF6 puncta. Since the NLS domain from
192 SV40, which replaces the MCD, is highly basic and could potentially induce condensates, we
193 fused CPSF6 with a non-basic NLS (PY-NLS) or removed the NLS entirely. Even though these
194 two proteins do not efficiently enter the nucleus, the few that do manage to reach the nucleus
195 can host viral particles, as evidenced by the presence of IN. Many viruses are typically blocked
196 in the cytoplasm due to the presence of these mutants that are mainly cytoplasmic. However,
197 because we used a high viral dose, the blockage in the cytoplasm was not complete. As a

198 result, the viruses that successfully entered the nucleus induced the formation of puncta
199 associated with CPSF6-deleted mutants, indicating that the MCD is not critical for the formation
200 of HIV-induced CPSF6 puncta (Fig.2F). Similar to the MCD, when we compared CPSF6
201 truncated for the LCRs with CPSF6 WT, we observed that the LCRs do not contribute to CPSF6
202 puncta formation. Therefore, the FG peptide alone, without the LCRs, is the only CPSF6
203 domain required for their formation (Fig. 2B-F; Suppl. Fig.1B).

204 Next, we tested the ability of the different CPSF6 deletion mutants for their ability to bind the
205 viral core using a previously described capsid binding assay (Selyutina et al., 2018). Wild type
206 and mutant CPSF6 proteins were expressed in human 293T cells at similar levels (INPUT)
207 (Figure 3A). Extracts containing wild type and mutant CPSF6 proteins were incubated with
208 stabilized HIV-1 capsid tubes for 1 h at 25˚ C in the presence of 10 μM of PF74, which is a
209 small molecule that competes with CPSF6 for binding to the hydrophobic pocket formed
210 between hexamers that constitute the viral core (Buffone et al., 2018; Price et al., 2014). HIV-
211 1 capsid stabilized tubes were washed, and the bound proteins were eluted using Laemli buffer
212 (BOUND). For every construct, the percentage of bound protein relative to input in the
213 presence or absence of PF74 is shown (Figure 3B). Our results revealed that the absence of
214 the FG peptide (ΔFG) entirely abolished CPSF6's ability to bind to the viral core. In agreement,
215 simultaneous deletion of the FG motif and LCRs (ΔFG ΔLCR) resulted in a construct unable to
216 bind to the viral core. Similar outcomes were observed when the LCRs were replaced with
217 sequences derived from ADD2, even if the FG was present.

218 LCR-FG is notably more disordered than ADD2-FG, containing a high proportion of prolines
219 (48 out of 98 residues), which makes it mostly non-foldable (Figure 4A-H). Since proline is a
220 structure-disrupting residue, LRC-FG is not expected to adopt any secondary structure. In
221 contrast, ADD2-FG contains fewer prolines (15 out of 98 residues) but has many charged
222 residues. It is predicted to form two short α helices and a ß strand, arranged as: α helix - FG –
223 ß strand - α helix (Figure 4E). ADD2-FG may form a flexible collapsed state, as its oppositely
224 charged residues are evenly distributed, potentially allowing polyelectrostatic compaction. This
225 suggests that FG within ADD2-FG may be less accessible for the interaction with the viral
226 core's hydrophobic pocket due to its involvement in this collapsed conformational ensemble
227 (Figure 4A-E, suppl. Fig.3). This aligns with the inability of CPSF6 carrying ADD2 in place of
228 the LCRs to induce CPSF6 puncta (Figure 2B). On the other hand, the deletion of only the
229 two LCRs, while keeping the FG peptide intact, resulted in unexpected findings. The ΔLCR
230 mutant exhibited a stronger binding affinity for the viral core when compared to the wild-type
231 protein (Figure 3B). These results suggest that the LCRs surrounding the FG motif are
232 modulating the affinity of CPSF6 to the viral core, which might be important for function. By
233 contrast, deletion of the MCD (ΔMCD) but retention of other regions, such as the FG peptide

234 and the LCRs, demonstrated a binding affinity to the viral core similar to that of the wild-type
235 protein. These results suggest that the MCD domain is not involved in the binding of CPSF6
236 to the viral core, which is not surprising since the CPSF6 (1-358), which does not have an
237 MCD, binds to the viral core (Lee et al., 2010).

238 Thus, the viral capsid, through the FG peptide of CPSF6, constitutes the scaffold of HIV-
239 induced CPSF6 puncta.

240 In summary, our results suggest that the FG peptide is the main determinant involved in the
241 binding of CPSF6 to the viral capsid. Interestingly, our work implies that the LCRs may be
242 modulating the affinity of the FG motif for the viral core (Figure 3B). Recognition motifs that
243 mediate protein-protein interactions, such as the FG motif of CPSF6, are usually embedded
244 within longer intrinsically disordered regions that can modulate affinity of the interaction
245 (Karlsson et al., 2022). Taken together, our data show that the FG peptide coordinates both
246 the binding to the viral core and the induction of CPSF6 puncta. This coordination suggests
247 that the FG peptide plays a critical dual role in recognizing the viral capsid and facilitating the
248 cellular clustering of CPSF6, which may be part of the cellular response to viral entry.

249

250 **Role of FG domain in viral replication.**

251 We aimed to investigate the role of FG domain of CPSF6 in viral infection. To correctly perform
252 the infectivity assay, we generated stable cell clones expressing protein levels comparable to
253 WT cells (Fig.5 A-B). Importantly, we assessed the reproducibility of our experiments by
254 freezing and thawing these clones. Thus, we performed infectivity assays in THP-1
255 macrophage-like cells WT, KO for CPSF6 or complemented with CPSF6ΔFG, using both a
256 single-round virus and a replication-competent virus. The results, shown in Figure 5 C-D,
257 indicate that complete depletion of CPSF6 reduces infectivity, as measured by luciferase
258 expression in a single-round infection (KO: ~65%; ΔFG: ~74%; compared to WT: 100% on
259 average). Notably, a more pronounced defect in viral particle production was observed when
260 WT virus was used for infection (KO: ~21%; ΔFG: ~16%; compared to WT: 100% on average).
261 According to our results, depletion of full-length CPSF6, or expression of a CPSF6 variant
262 lacking only the FG domain, both reduce viral infection compared to cells expressing full-length
263 CPSF6, indicating that the FG domain plays an important role in HIV-1 infection.

264

265 **HIV-induced CPSF6 mutants puncta and Nuclear Speckles.**

266 The NS factor SC35, commonly used as a marker of NS, has been detected in HIV-induced
267 CPSF6 puncta (Figure 6A)**.** In this study, we investigated whether HIV-induced CPSF6 mutant

268 puncta are associated with SC35. We infected cells expressing CPSF6 WT, CPSF6 3xNLS
269 ΔMCD, CPSF6 ΔLCR, CPSF6 ΔMCD, and CPSF6 PY NLS ΔMCD, and examined whether the
270 nuclear puncta formed by the various CPSF6 proteins associate with SC35. Imaging analysis
271 revealed no significant difference in the association of SC35 with CPSF6 WT or the CPSF6
272 mutants (Figure 6B), confirming that the disordered domains, LCRs and MCDs, are
273 dispensable for the formation of HIV-induced CPSF6 puncta that localize in the canonical
274 nuclear niches marked by SC35.
275

276 **Biogenesis of HIV-Induced CPSF6 Puncta carrying Nuclear Speckle factors.**

277 Nuclear speckle factors, particularly those involved in their biogenesis, such as SON and
278 SRRM2 (Ilik et al., 2020), have been identified as constituents of HIV-induced CPSF6 puncta.
279 However, the specific role of NSs in these puncta remains unclear. Most of the existing results
280 have been obtained through immunofluorescence experiments conducted several days post-
281 infection. In this study, we conducted a time course experiment to investigate the biogenesis
282 of HIV-induced CPSF6 puncta, which contain NS factors. Our goal was to capture the fusion
283 event between the HIV-induced CPSF6 puncta and NS, providing insights into the dynamics
284 of how HIV manipulates host nuclear structures during infection. We hypothesized that NS
285 factors might either be recruited during the initial formation of HIV-induced CPSF6 puncta,
286 shortly after the virus is released from the nuclear basket of the NPC, or later via the fusion
287 between NSs and HIV-induced CPSF6 puncta. To investigate this, we performed live imaging
288 to track CPSF6-mNeonGreen and NSs in cells expressing endogenous SRRM2 fused with a
289 Halo tag, using CRISPR Paint (courtesy of Roy Parker) (Lester et al., 2021).We fixed and
290 labeled samples at different time points post-infection, ranging from 6 h.p.i. to 30 h.p.i. (Figure
291 7A, Suppl. Fig.4A-B). At 6 h.p.i., 27% of HIV-induced CPSF6 puncta were still individual,
292 compared to only 9% at 30 h.p.i. (Figure 7B). Concurrently, 61% of HIV-induced CPSF6 puncta
293 were fused with NS at 6 h.p.i., rising to 75% at 30 h.p.i. (Figure 7B). This indicates a
294 progressive increase in the number of HIV-induced CPSF6 puncta fusing with NS over time
295 (Figure 7C). Overall, we detected individual CPSF6 puncta (green) and NSs (red) that quickly
296 fused, confirming that this fusion occurs within the two independent puncta, CPSF6 and
297 SRRM2, rather than during the formation of HIV-induced CPSF6 puncta (Figure 7A-C, movies
298 1A, 1B).
299 Taken together, these results suggest that HIV-induced CPSF6 puncta first form independently
300 of NS and later fuse with NS, causing an enlargement of NS as part of the hijacking process
301 by HIV-1.
302

303 **Role of SON and SRRM2 in the Fusion and Stabilization of HIV-Induced CPSF6 Puncta**
304 **within Nuclear Speckles.**

305 Macrophage-like cells, THP-1, were depleted for SON or SRRM2 using AUM*silence*™ ASO
306 technology (Gao et al., 2024; Marasca et al., 2022; Mazzeo et al., 2024; Zhang et al.). The
307 level of depletion of SON and SRRM2 was evaluated by immunofluorescence and western
308 blot using antibodies against SON and SRRM2 (Figure 8A). Both depleted cells were analyzed
309 also for the presence of NS, labelled by SC35 (Suppl. Fig.6A). However, recent findings
310 suggest that the primary target of the SC35 mAb is SRRM2 (Ilik et al., 2020). To confirm this,
311 cells depleted of SRRM2 were labelled with antibodies against both SRRM2 and SC35 (Suppl.
312 Figure 6 A-B). We have observed that the reduction of SRRM2 resulted in a slight decrease in
313 the mean intensity of SC35, whereas the depletion of SON (Suppl. Fig.6A) did not have the
314 same effect.

315 Subsequently, we infected THP-1 depleted cells for SRRM2 and SON or control cells with HIV-
316 1 and we fixed them at 48 h post-infection for immunofluorescence. We calculated the
317 percentage of CPSF6 puncta in HIV infected THP-1 control cells (approximately 78%) and in
318 HIV infected THP-1 cells depleted for SRRM2 (about 43%) and SON (around 66%). Results
319 from our experimental conditions indicate that the partial depletion of SRRM2 affects the
320 formation of HIV-induced CPSF6 puncta, while the depletion of SON slightly reduces their
321 establishment (Figure 8B).

322

323 **The IDR of SRRM2 is a crucial element for the fusion of HIV-Induced CPSF6 puncta to**
324 **the nuclear speckles.**

325 CPSF6 contains several disordered regions (Di Nunzio et al., 2023) as well as NS factors (Ilik
326 et al., 2020). Previous studies have established the role of IDR of SRRM2 in the biogenesis of
327 NS (Ilik et al., 2020).

328 Here, we investigate the role of IDRs within NS factors, with a specific emphasis on SRRM2,
329 in the fusion and stabilization of HIV-induced CPSF6 puncta. To address this inquiry, we
330 utilized previously published HEK293 cell lines generated using the CRISPaint system (Lester
331 et al., 2021), comprising two distinct lines: HaloTag SRRM2, SRRM2 full-length (FL) (1-2748
332 aa) fused with Halo tag(insertion at amino acid [aa] 2,708), and a cell line lacking the C-
333 terminal IDR of SRRM2, known as ∆IDR HaloTag SRRM2(1-429 aa, with halo insertion at aa
334 430) .

335 As anticipated, the truncated form of SRRM2 displayed a more diffuse distribution within the
336 nucleus, without recruitment to NS (suppl. Fig.5 A-E), and consequently lacked nuclear puncta

337 (Figure 8C top panels). On the other hand, the number of SON puncta was highly similar
338 between the two cell lines (Figure 8C, panels on the bottom, Suppl. Fig.7).

339 Subsequently, we quantified the formation of CPSF6 puncta in both cell lines infected with HIV-
340 1. No significant difference in HIV-induced CPSF6 puncta formation was observed between
341 HEK293 and HEK293 carrying the SRRM2 halo tag (~27%). However, a substantial reduction
342 in HIV-induced CPSF6 puncta was evident in the cell line carrying the SRRM2 form that lacks
343 the IDR (~11%) (Figure 8D). Collectively, these results underscore the pivotal role of the
344 SRRM2 IDR in the stabilization of HIV-induced CPSF6 puncta through their fusion with NSs
345 (Figure 7A-C).

346

347

348

Discussion

HIV-1 capsid has transformed long-standing assumptions in the field. Previously considered an undruggable viral target due to the belief that it disassembled shortly after HIV entered target cells, whereas recent findings have revealed its essential role in nuclear import (Ay and Di Nunzio, 2023; Blanco-Rodriguez and Di Nunzio, 2021; Blanco-Rodriguez et al., 2020; Chen et al., 2016; Taylor and Fassati, 2024; Yamashita and Emerman, 2004; Zila et al., 2021). Notably, the first-in-class antiretroviral capsid inhibitor, Lenacapavir, has shown remarkable results in patients, demonstrating that the capsid, contrary to previous belief, can indeed be a viable therapeutic target. While Lenacapavir improves patients' quality of life with only two injections per year (Link et al., 2020; Segal-Maurer et al., 2022), no current antiretroviral drugs provide a cure. This may be due to our incomplete understanding of certain aspects of HIV biology. In this study, we shed light on the post-nuclear entry steps, a critical phase for the establishment of viral reservoirs, which represent the main barrier to a cure. Recent findings have shown that not only the HIV viral capsid translocates through the NPC, but that the viral nuclear entry also enhances the formation of CPSF6 puncta. Additionally, it has been revealed that RT is completed within the nucleus (Burdick et al., 2020; Dharan et al., 2020; Rensen et al., 2021; Scoca et al., 2023; Selyutina et al., 2020). Furthermore, incoming viral RNA has been observed to be sequestered in nuclear niches in cells treated with the reversible reverse transcriptase inhibitor, NVP. When macrophage-like cells are infected in the presence of NVP, the incoming viral RNA is held within the nucleus (Ay et al., 2025; Rensen et al., 2021; Scoca et al., 2023). This scenario could also occur in patients receiving antiretroviral therapy. Notably, we found that pharmacological dismantling of CPSF6 puncta prevents the restoration of nuclear reverse transcription, indicating that HIV-induced CPSF6 puncta are critical for the viral life cycle. Conversely, THP-1 cells either depleted of CPSF6 or expressing a CPSF6 variant lacking the FG domain exhibit impaired viral replication. Understanding the biogenesis of these puncta could be a significant step towards deepening our knowledge of HIV biology and providing additional tools to combat this pandemic virus.

Here we identify the disordered FG peptide essential for the binding with the viral capsid as the inducer of HIV-induced CPSF6 puncta. Notably, CPSF6 protein lacking the FG peptide is incapable of forming nuclear puncta. Moreover, we discovered that the two major intrinsically disordered regions of CPSF6, the MCDs and the LCRs, are dispensable for the formation of viral nuclear puncta. The MCDs of CPSF6 have been shown to provide cohesion for NS condensation (Greig et al., 2020). A study performed in HEK293T-derived CPSF6 knockout (CKO) cells (clone B8) complemented with a tagged CPSF6 reported that the MCD is critical for CPSF6 cluster formation (Jang et al., 2024). In contrast, we chose to work with a more physiologically relevant cell line for HIV infection, macrophage-like cells, and we employed

385 untagged CPSF6 proteins (WT and deletion mutants) to avoid potential artifacts introduced by
386 protein tagging. Under our experimental conditions, the MCD did not appear to play a critical
387 role in the formation of HIV-induced CPSF6 puncta. CPSF6 puncta induced by CPSF6
388 depleted for the MCD upon infection are highly similar to those formed by wild-type CPSF6,
389 suggesting that MCDs do not play a significant role in this process. Additionally, the LCRs of
390 CPSF6 also do not appear to influence puncta formation, as their depletion does not reduce
391 the number of CPSF6 puncta. This indicates that neither the MCDs nor the LCRs are involved
392 in CPSF6 puncta formation during HIV infection. Surprisingly, when we assessed the ability of
393 CPSF6 domains to bind to the viral capsid, we observed that the deletion of LCRs increases
394 CPSF6's ability to bind to the viral capsid. We hypothesize that a change in charges may alter
395 the binding mechanism of CPSF6 when LCRs are absent. In scenarios where the FG motif is
396 depleted, we observed a dramatic inhibition of HIV-induced CPSF6 puncta formation and a
397 lack of binding to the viral core *in vitro*. On the flip side, the linkage between CPSF6 entities is
398 facilitated by the FG peptides' interaction with certain hydrophobic CA pockets along adjoining
399 hexamers (Wei et al., 2022). Therefore, it's conceivable that FG peptides, not involved in the
400 capsid's binding, could coalesce similarly to FG-Nups. These undergo phase separation,
401 forming condensates with nuclear pore complex-like permeability barrier features (Hülsmann
402 et al., 2012).

403 We have also identified HIV-induced CPSF6 puncta formation independent of NS at early
404 stages post-infection, with a progressive increase in CPSF6 puncta co-localizing with NS over
405 time. These results were obtained by focusing our studies on two scaffold proteins involved in
406 NS biogenesis: SON and SRRM2. Over the past approximately 0.6–1.2 billion years of
407 metazoan evolution, these two factors have undergone significant lengthening, unlike many
408 other proteins involved in splicing. This extension primarily occurred within their IDRs, which
409 are commonly associated with liquid-liquid phase separation (LLPS) and the formation of
410 biomolecular condensates (Rai et al., 2018). Co-depleting SRRM2 with SON, or depleting SON
411 in a cell line where it is deleted the intrinsically disordered C-terminus of SRRM2, abolished
412 the formation of NSs (Ilik et al., 2020). However, the depletion of only one factor does not
413 abolish NSs. Consistent with this finding, we observed that the depletion of SRRM2 does not
414 affect the presence of SON nuclear puncta. Similar results were obtained with cells genetically
415 modified to express the truncated form of SRRM2 lacking the IDRs. When the ΔIDR SRRM2-
416 halo tag was detected by the halo ligand, there was no recruitment of the truncated SRRM2
417 form in NSs. However, if SC35 is used as a target for antibodies, it can still be detected, albeit
418 with a much lower intensity signal than in cells expressing the full-length SRRM2. Additionally,
419 we observed a significant reduction in the detection of HIV-induced CPSF6 puncta. The few
420 CPSF6 puncta detected in these cells colocalized with the weak SC35 signal. These results

421 suggest that the IDR of SRRM2 plays an important role in HIV-induced CPSF6 puncta
422 stabilization, but rare puncta can appear, likely induced by redundant NS factors. This indicates
423 that further investigation is needed to better understand the hijacking of NS by HIV.

424 Taken together, our results reveal the intricate interplay between individual CPSF6 domains
425 and the viral capsid in dictating the nuclear fate of the virus. Concurrently, the IDR domain of
426 SRRM2 contributes to the enlargement of particular nuclear speckles with the fusion of HIV-
427 induced CPSF6 puncta. Lastly, these HIV-induced CPSF6 puncta necessitate of FG peptide
428 to engage the viral core.

429 Overall, this study could provide insights into the understanding of viral invasion and
430 persistence within the host.

431

432

433 **Materials and methods**

434

435 **Cell lines**

436 THP-1 cells are immortalized monocytic cells, which, once seeded, differentiate into
437 macrophage-like cells under phorbol 12-myristate 13-acetate (PMA) treatment (160 nM). THP-
438 1 cells were also engineered knocking out CPSF6. These cells were cultivated in RPMI 1640
439 medium supplemented with 10% fetal bovine serum (FBS) and 1% penicillin–streptomycin
440 solution (100 U/ml). HEK293T cells are human embryonic kidney cells used to produce LVs.
441 For figure 2, we used two engineered HEK293 strains, Halo tagged SRRM2 HEK293 cells and
442 Halo tagged SRRM2 ΔIDR HEK293 cells, kind gifts from Roy Parker's lab (Lester et al., 2021),
443 characterized by the Halo tagged SRRM2 protein. In Halo tagged SRRM2 ΔIDR HEK293 cells
444 Parker's lab deleted also the SRRM2 sequence encoding for aa 430-2748. All the HEK293
445 strains were cultivated in Dulbecco's modified Eagle medium supplemented with 10% FBS and
446 1% penicillin–streptomycin (100 U/ml).

447

448 **Bacteria strains**

449 All E. coli bacteria strains were grown in Luria-Bertani (LB) medium at 37°C. DHα competent
450 cells and Stellar Competent Cells were used for molecular cloning, while E.coli One-Shot
451 BL21star (DE3) cells were exploited for protein production.

452

453 **Plasmids**

454 To express the WT CPSF6/WT CPSF6-mNeonGreen and the mutant CPSF6/mutant CPSF6-
455 mNeonGreen clones, the correspondent coding sequences were engineered in pSICO
456 plasmids. The two original plasmids used were pSICO CPSF6-mNeonGreen and pLPCX
457 CPSF6 ADD2, gift from Mamuka's lab. HIV-1ΔEnvIN$_{HA}$ LAI (BRU) plasmid encodes the
458 ΔEnvHIV-1 LAI (BRU) viral genome where the IN protein is fused to the HA tag while pNL4.3
459 ∆env ∆Nef IRES GFP plasmid encodes the ΔEnvHIV-1 NL4.3 viral genome and contains also
460 a GFP sequence headed by an IRES. The pNL4.3 ∆env ∆NefLuc has the Luciferase cDNA as
461 reporter gene.

462

463 **CRISPR-Cas9 knockout in THP-1 cells**

464 To target CPSF6, three different crRNAs were used simultaneously (specific sequence: 5'-
465 TCGGGCAAATGGCCAGTCAAAGG-3', 5'-AGGACGGGGCCGTTTTCCAGGGG-3', and 5'-
466 CATGTAATCTCGGTCTTCTGGGG-3', all ordered from Integrated DNA Technologies, IDT).
467 Pre-designed unspecific crRNA was used as control (IDT). crRNA and tracrRNA were
468 resuspended in IDT Duplex Buffer according to the manufacturer's instructions. On the day of
469 the nucleofection, duplexes were formed by mixing equimolar concentration of crRNA and
470 tracrRNA, followed by 5-min annealing at 95°C. RNA duplexes were then mixed (1:2) with
471 TrueCut™ Cas9 Protein v2 for 10 min at room temperature to generate ribonucleoprotein
472 (RNP) complexes. 2×10^5 THP1 cells were resuspended in P3 Primary Cell Nucleofector™
473 Solution, mixed with RNP and Alt-R® Cas9 Electroporation Enhancer (90 pmol, IDT), and
474 nucleofected in a 4D-Nucleofector™ System using the P3 Primary Cell 4D-Nucleofector™X Kit
475 S (program FI-110). After nucleofection, cells were seeded in complete RPMI medium with
476 20% FBS. Three days after nucleofection, cells were plated for clonal selection.

477

478 **Clonal selection of KO cell lines**

479 Seventy-two hours post nucleofection, cells were diluted in RPMI medium containing 20% FBS
480 and plated in five 96-well plates at 1 and 5 cells/well condition. After one-month, selected
481 microcolonies (50-100) are placed into 24- wells plate. Once that wells were near confluence,
482 cells were transferred into the well of a 6- well plate. After growing for another one-month, cells
483 were proceeded for western blot.

484

485 **Western blot**

486 Proteins were extracted on ice from THP-1 cells using RIPA buffer (20 mM HEPES, pH 7.6,
487 150 mM NaCl, 1% sodium deoxycholate, 1% Nonidet P-40, 0.1% SDS, 2mM EDTA, complete
488 protease inhibitor), and protein concentration was quantified using a detergent-compatible
489 (DC) protein assay (Pierce™ BCA Protein Assay Kit) with bovine serum albumin (BSA) as a
490 standard. 90 µg of total protein lysate were loaded onto SDS-PAGE 4- 12% Bis-Tris gels
491 (Invitrogen); an Ab rabbit anti-CPSF6 (1:500) and an anti-rabbit HRP-conjugated (1:5000)
492 were used for the detection of CPSF6, whereas the normalization was done by an Ab anti-
493 actin HRP-conjugated (1:3000). Visualization was carried out using an ECL solution .

494 **AUMsilence™ ASO**

495 AUMsilenceTM 352 ASOs were synthesized by AUM BioTech, LLC (Philadelphia, USA). THP-
496 1 negative control, THP-1 KD SRRM2 and THP-1 KD SON cells were differentiated with PMA
497 (160 nM) for 48 h then incubated 72 h with 10 µM of a AUMsilenceTM ASOs complementary
498 to the mRNA of SRRM2 and SON, respectively (scramble control AUMscrambleTM). All cells
499 were kept in incubator at 37°C and 5% CO2. These cells were cultivated in RPMI 1640 medium
500 supplemented with 10% fetal bovine serum (FBS) and 1% penicillin–streptomycin solution (100
501 U/ml).

502 **Cloning**

503 pSICO CPSF6-mNeonGreen was used to generate deletion mutants by exceeding different
504 regions from the original sequence according to Table2. The mutants were produced with and
505 without the mNeonGreen tag, except for pSICO CPSF6-ΔMCD and pSICO CPSF6-ΔMCD PY
506 NLS. All primers are specified in Table1.

507 For pSICO CPSF6-mNeonGreen ΔLCR ADD2, CPSF6 ΔLCR ADD2 sequence was amplified
508 by PCR from pLPCX CPSF6 ADD2 and specific primers were designed to add also BamHI
509 restriction sites at the extremities. Phusion Flash High-Fidelity PCR Master Mix was used and
510 the reaction was performed according to the manufacture datasheet (100 ng template DNA
511 and 56° annealing temperature). PCR products were treated with DpnI for 1h at 37°C and then
512 digested with BamHI restriction enzyme at 37°C for 1h. 2 ul of the backbone pSICO CPSF6-
513 mNeonGreen were digested with BamHI at 37°C for 3h and further treated with CIP. After gel
514 extraction of the DNA and purification, insert and backbone were ligated with T4 Ligase for 2
515 h at 22°C. 5 ul of the product were used to transform 50 ul of DH5α bacteria (30 min at 4°C,
516 45 sec at 42°C, 2 min at 4°C, incubation in SOC medium for 1h at 37°C and plating on LB agar
517 dishes).

518 All the other mutants were obtained using In-Fusion®Snap Assembly protocol, which allowed
519 to amplify the original plasmid deleting the small region of interest. The different primers used
520 are reported in Table1. As reported in the In-Fusion®Snap Assembly protocol

521 (https://www.takarabio.com/documents/User%20Manual/In/InFusion%20Snap%20Assembly
522 %20User%20Manual.pdf), 5 ng of the original plasmid were amplified with PrimeSTARMax
523 DNA polymerase in 35 PCR cycles (10 sec at 98°C, 15 sec at 55°C, 5sec/kb at 72°C). The
524 PCR products were then digested for 1h at 37°C using DpnI to get rid of the original plasmid.
525 After plasmid cleaning up, it was circularized through a ligation of 15 min at 50°C. 2.5 ul of
526 products were used in Stellar™ Competent Cells' transformation, following the same
527 procedure used for DH5α transformation, previously explained.

528 **Lentiviral vectors and viral productions**

529 LVs and HIV-1 viruses were produced by transient transfection of HEK293T cells through
530 calcium chloride co-precipitation. Co-transfection was performed as follows: for LVs, 10 µg of
531 transfer vector, 10 µg of packaging plasmid (gag-pol-tat-rev), and 2.5 µg of pHCMV-VSV-G
532 envelope plasmid; for VSV-G / HIV-1ΔEnvIN$_{HA}$ LAI (BRU) -vpx viruses and VSV-G/pNL4.3
533 ∆env ∆Nef IRES GFP-vpx, 10 µg HIV-1ΔEnvIN$_{HA}$ LAI (BRU) plasmid or pNL4.3 ∆env ∆Nef
534 IRES GFP or pNL4.3 ∆env∆R∆Nef Luc, 2.5 µg of pHCMV-VSV-G plasmid and 3 ug of SIVMAC
535 vpx (Durand et al., 2013) or full length virus 10 µg HIV-1 NL4.3AD8 and 3 ug of SIVMAC vpx.
536 After the collection of the supernatant 48 h post-transfection, lentiviral particles were
537 concentrated by ultracentrifugation for 1 h at 22000 rpm at 4°C and stored at −80°C. LVs and
538 viruses were tittered by qPCR in HEK293T cells 3 days post-transduction.

539

540 **Cell transduction and infection**

541 THP-1 ctrl CRISPR clone 2 cells and THP-1 (duplex1-2-3 CRISPR) KO clone 4 cells were
542 differentiated with PMA (160 nM) for 72 h then transduced with different mutants of CPSF6 for
543 72h (MOI=1) and then infected for 30h with VSV-G / HIV-1ΔEnvIN$_{HA}$ LAI (BRU) -vpx (MOI=10)
544 in presence of Nevirapine (10uM). The medium was always supplemented with PMA (160 nM)
545 .

546 For Halo tagged SRRM2 HEK 293 cells and Halo tagged SRRM2 ∆IDR HEK 293 cells, 2×10^5
547 cells were seeded on coverslips coated with polylysine in complete growth medium (DMEM,
548 GlutaMAX^{TM}-I, 10% FBS, and 1% P/S) and incubated at 37°C (5% CO_2) for 24 h. Cells were
549 then infected with the VSV-G / HIV-1ΔEnvIN$_{HA}$ LAI (BRU) (MOI 10) in complete growth medium
550 supplemented with Nevirapine (10µM) for 24h.

551 THP-1 control (scramble), THP-1 KD SRRM2 Cells and THP-1 KD SON cells were seeded on
552 coverslips and differentiated with PMA (160 nM) for 72 h. Then incubated 48 h with 10 µM of
553 a FANA ASOs (scramble control FANA (SCR-FANA), SRRM2-FANA and *SON*-FANA). ASOs
554 used in this study were designed and synthesized by AUM LifeTech (Philadelphia, PA, USA).

555 Next, cells were infected with the pNL4.3 ∆env ∆Nef IRES GFP-VPX (MOI 25) in complete
556 growth medium and incubated at 37°C in 5% CO_2 for 4 days.

557 In the four timepoints experiment, THP-1 cells were differentiated with PMA (160 nM) for 72 h.
558 Cells were then infected with VSV-G / HIV-1ΔEnvIN$_{HA}$ LAI (BRU) -vpx (MOI=10) in presence
559 of Nevirapine (10uM), in complete growth medium supplemented with PMA (160 nM), and
560 incubated at 37°C in 5% CO_2 for 6 h, 9 h, 12 h, 30 h post infection.

561 **THP-1 infection with HIV-1 replication competent virus and viral RNA real-time**
562 **quantification.**

563 THP-1 cells differentiate into macrophage-like cells within 48 h under phorbol 12-myristate 13-
564 acetate (PMA) treatment (160 nM). THP1 cells were cultured at 37 °C in 5% CO2 in RPMI
565 Medium 1640 supplemented with GlutaMAXTM-I, 10% FBS, 1% P/S (100 U/mL), MEM Non-
566 essential Amino Acid Solution (0.1 mM), Sodium Pyruvate (1 mM), HEPES buffer solution (10
567 mM). THP1 cells differentiated into macrophages are infected with HIV-1 NL4.3AD8 with Vpx
568 (MOI 6). 24 hours after infection, the cells were washed and RPMI is added to the cells. Three
569 days post- infection, the supernatant is collected and centrifuged. Total viral RNA was extracted
570 from the supernatant via the QIAamp Viral RNA Mini Kit (Qiagen, Hilden Germany) following
571 the manufacturer's protocol. Total viral RNA was eluted into 60 µL of DEPC. vRNA is reverse
572 transcribed into DNA with the Maxima H minus reverse transcriptase (Thermoscientific,
573 EP0752) following the manufacturer's protocol and the following primers were used (forward:
574 GCCTCAATAAAGCTTGCCTTGA, reverse: TGACTAAAAGGGTCTGAGGGATCT). DNA
575 copies for each sample were determined by qPCR and the SYBR green dye. The absolute
576 DNA copies were calculated from a standard curve run in parallel using HIV-1 NL4.3AD8
577 plasmid. The qPCR cycle was as follows: 15 min at 95 °C, 40 cycles (15 s at 95 °C, 30 s at 58
578 °C, 30 s at 72 °C), 15 s at 95 °C, 15 s at 55 °C, 20 min melting curve until reaching 95 °C, 15
579 s at 90 °C.

580

581 **Immunofluorescense Microscopy**

582 *Immunostaining*

583 On the day of fixation, the cells were washed with PBS and fixed with 4% PFA for 15 min. Cells
584 were treated with glycine 0.15% for 10 min, permeabilized with 0.5% Triton X-100 for 30 min,
585 and blocked with 1% BSA for 30 min. All antibody incubations were carried out at room
586 temperature in a dark humid chamber, for 1 h with primary antibodies and for 45 min with
587 secondary antibodies. Washes between antibody incubations and antibody dilutions were
588 done in 1% BSA.

589 Primary antibodies were diluted as follows: anti-HA 1:500, anti-CPSF6 1:400, anti-SC35 1:200,
590 anti-SON 1:200, anti-SRRM2 1:200, Secondary antibodies used were goat anti-rabbit Alexa-
591 488 1:300 and donkey anti-Rabbit Cy3 1:1000 for CPSF6, SON and SRRM2, goat anti-rat
592 Alexa-647 1:100 for IN-HA, goat anti-mouse Alexa-647 1:300 or goat anti-mouse Alexa-488
593 1:300 for SC35, donkey anti-mouse Alexa-647 1:300 for SRRM2.

594 Finally, cells were stained with Hoechst 33342 1:10000 for 5 min. Coverslips were mounted
595 on glass slides (Star Frost) with Prolong Diamond Antifade Mountant.

596 *HaloTag Labelling*

597 To detect the HaloTag in Halo tagged SRRM2 HEK 293 cells and Halo tagged SRRM2 ΔIDR
598 HEK 293 cells, we used HaloTag®TMR Ligand following the Technical Manual available at
599 www.promega.com/protocols. Specifically, cells were incubated in DMEM supplemented with
600 HaloTag®TMR Ligand (5 μM) for 15 min at 37°C. The ligand-containing medium was then
601 removed and replaced with an equal volume of 1X PBS, repeating the step twice and ending
602 with warm complete medium. Cells were incubated in complete culture medium for 30 min at
603 37°C. The medium was then removed and replaced with an equal volume of fresh warm culture
604 medium.

605 *Immuno-RNA FISH*

606 On the day of fixation, fixed cells were incubated in Permeabilization/Blocking buffer (1% BSA,
607 0.3% Triton X-100, 2mM Vanadyl Ribonucleoside complexes (VRCs) in RNase-free PBS) for
608 1 h before the antibodies' incubations. Antibodies were diluted in Permeabilization/Blocking
609 buffer. After the primary and secondary antibody staining (respectively of 1h and 45 min) with
610 the respective washes, cells were fixed for a second time in PFA 4% (in RNase-free PBS) for
611 10 min at room temperature with subsequent washes with RNase-free PBS. In the meantime,
612 40 pmol of primary smiFISH probes (Tsanov et al., 2016) (24 smiFISH probes designed against
613 HIV-1 pol sequence (Rensen et al., 2021) were hybridized with 50 pmol of secondary FLAP
614 probe conjugated to a Cy5 fluorophore
615 (Cy5/AATGCATGTCGACGAGGTCCGAGTGTAA/Cy5Sp/) in 1X NEBuffer 3 (diluted in RNase-
616 free H_2O) using a thermocycler. The program setting follows: 3 min at 85°C, 3 min at 65°C and
617 5 min at 25°C. FISH-probe solution was then diluted 1:50 in Hybridization buffer (90%
618 Stellaris® RNA-FISH Hybridization Buffer, 10% Deionized Formamide). After the samples were
619 washed in Wash A buffer (20% Stellaris® RNA-FISH Wash Buffer A, 10% Deionized

620 Formamide, in RNase-free H_2O) at room temperature for 5 min, they were placed on parafilm,
621 covered with 50 ul of FISH-probe solution in Hybridization buffer and incubated overnight at
622 37°C. The next day, cells were washed with Wash A buffer in the dark at 37°C for 30 min.
623 Afterwards, the samples were incubated for 10 min in Hoechst 333342 diluted 1:10000 in
624 RNase-free H_2O and then washed with Wash B buffer for 5 min at room temperature in the
625 dark. Finally, the cells were washed in RNase-free H_2O before the coverslips were mounted
626 on microscopy slides using ProLongTM Diamond Antifade mounting medium. The mounting
627 medium was cured overnight at room temperature under the chemical hood and away from
628 light.

629 *Images acquisition*

630 Images were acquired using a Zeiss LSM700 confocal inverted microscope, with a 63×
631 objective (Plan Apochromat, oil immersion, NA = 1.4), using diode lasers at 405 nm, 488 nm,
632 555 nm and 639 nm for the respective fluorophores. A pixel size of 0.07 µm was used.

633

634 **Live imaging on HIV-1 infected Halo tagged SRRM2 HEK 293 cells**

635 Halo tagged SRRM2 HEK293 cells were seeded and transduced with CPSF6-mNeonGreen
636 lentiviral vector (MOI 0.5) for 24 h. 0.3 x 10^6 Cells were then transferred on poly-L-Lysin coated
637 Ibidi-dishes and infected with VSV-G / HIV-1ΔEnvIN$_{HA}$ LAI (BRU) (MOI = 40) in presence of
638 Nevirapine (10 uM) for 2 h. After having changed the medium and labeled with HaloTag®TMR
639 Ligand (5 µM) (see previous "HaloTag labelling" section) and Hoechst 333342 diluted 1:80000
640 in complete DMEM medium, 4D movies were acquired.

641 3D movies were acquired using a Nikon Ti2-E Confocal Inverted Spinning Disk microscope,
642 using a 63x objective (Plan Apochromat, oil immersion, NA = 1.4) and a sCMOS Hamamatsu
643 camera, Orca Flash 4. Pixel size 6.5 µm, 2048 x 2044 pixel, QE 82%. For the Z stuck a Z piezo
644 stage was used, with 0.33 µm interval. For live imaging acquisitions, cells were placed in an
645 environmental chamber with 37°C temperature, 5% CO_2 and 21% O_2.

646

647 **Protein expression and purification.**

648 pET-11a vectors were used to express the HIV-1 capsid protein. Point mutations, A14C and
649 E45C were introduced using the QuikChange II site-directed mutagenesis kit (Stratagene)
650 according to the manufacturer's instructions. All proteins were expressed in *Escherichia coli*

651 one-shoot BL21star™ (DE3) cells (Invitrogen). Briefly, LB medium was inoculated with
652 overnight cultures, which were grown at 30°C until mid log-phase (A_{600}, 0.6-0.8). Protein
653 expression was induced with 1 mM isopropyl-β-d-thiogalactopyranoside (IPTG) overnight at
654 18°C. Cells were harvested by centrifugation at 5,000 x g for 10 min at 4°C, and pellets were
655 stored at −80°C until purification. Purification of capsid was carried out as follows. Pellets from
656 2 L of bacteria were lysed by sonication (Qsonica microtip: 4420; A=45; 2 minutes; 2 seconds
657 on; 2 seconds off for 12 cycles), in 40 ml of lysis buffer (50 mM Tris pH=8, 50mM NaCl, 100mM
658 β-mercaptoethanol and Complete EDTA-free protease inhibitor tablets). Cell debris were
659 removed by centrifugation at 40,000g for 20 min at 4°C. Proteins from the supernatant were
660 precipitated by incubation with 1/3 of volume of saturated ammonium sulfate containing 100
661 mM β-mercaptoethanol for 20 min at 4°C and centrifugation at 8,000g for 20 min at 4°C.
662 Precipitated proteins were resuspended in 30 ml of buffer A (25mM MES pH6.5, 100 mM β-
663 mercaptoethanol) and sonicated 2-3 times (Qsonica microtip: 4420; A=45; 2 minutes; 1 second
664 on; 2 seconds off). Sample was dialyzed 3 times in buffer A (2 h, overnight, 2 h). The sample
665 was sonicated and diluted in 500 ml of buffer A and was chromatographed sequentially on a 5
666 ml HiTrap™ Q HP column and on a 5 ml HiTrap™ SP FF column (GE Healthcare), both pre-
667 equilibrated with buffer A. The capsid protein was eluted from HiTrap™ SP FF column using a
668 linear gradient from 0-2 M of NaCl. Absorbance at 280 nm was checked to take the eluted
669 fraction that had higher protein levels. Pooled fractions were dialyzed 3 times (2 h, overnight,
670 2 h) in storage buffer (25mM MES, 2M NaCl, 20 mM β-mercaptoethanol). Sample was
671 concentrated using centricons to a concentration of 20 mg/ml and stored at -80°C.

672

673 **Assembly of stabilized HIV-1 capsid tubes**

674 1 mL of monomeric capsid (3 mg/mL or 1 mg/mL) was dialyzed in SnakeSkin dialysis tubing
675 10,000 MWCO (Thermo Scientific) against a buffer that is high in salt and contains a reducing
676 agent (buffer 1: 50 mM Tris, pH 8, 1 M NaCl, 100 mM β-mercaptoethanol) at 4°C for 8 hours.
677 Subsequently the protein was dialyzed against the same buffer without the reducing agent β-
678 mercaptoethanol (buffer 2: 50 mM Tris, pH 8, 1 M NaCl) at 4°C for 8 hours. The absence of β-
679 mercaptoethanol in the second dialysis allows formation of disulfide bonds between Cysteine
680 14 and 43 inter-capsid monomers in the hexamer. Finally the protein is dialyzed against buffer
681 3 (20 mM Tris, pH 8,0, 40 mM NaCl) at 4°C for 8 hours. Assembled complexes were kept at
682 4°C up to 1 month.

683

684 **Capsid binding assay protocol**

685 Human HEK293T cells were transfected for 24 h with a plasmid expressing the specified
686 CPSF6 variant tagged with mNeonGreen. Cell media was completely removed and cells were

687 lysed in 300 µL of capsid binding buffer (CBB: 10 mM Tris, pH 8,0, 1,5 mM MgCl2, 10 mM
688 KCl) by scrapping off the plate. Cells were rotated at 4°C for 15 min and then centrifuged to
689 remove cellular debris (21,000 *x g* ,15 minutes, 4 °C). Cell lysates were incubated with
690 stabilized HIV-1 capsid tubes for 1 h at 25 °C. Subsequently, stabilized HIV-1 capsid tubes
691 were washed by pelleting the complexes by centrifugation at 21,000 x g for 2 min. Pellets were
692 washed using by resuspension in CBB or PBS. Pellets were resuspended in Laemmli buffer
693 1X and analyzed by Western blotting using anti-p24 or anti-mNeonGreen antibodies.

694

695 **Bioinformatics analysis of CPSF6**

696 Intrinsic disorder propensity of CPSF6 was evaluated using the Rapid Intrinsic Disorder
697 Analysis Online platform (RIDAO) (https://ridao.app/) designed to predict disordered residues
698 and regions in a query protein based on its amino acid sequence (Dayhoff and Uversky, 2022).
699 RIDAO yields results by combining the outputs of several commonly used per-residue disorder
700 predictors, such as PONDR® VLXT (Romero et al., 2001), PONDR® VL3(Peng et al., 2006),
701 PONDR® VLS2B(Peng et al., 2005), PONDR® FIT (Xue et al., 2010), as well as IUPred2
702 (Short) and IUPred2 (Long)(Dosztanyi et al., 2005a, b). RIDAO also computes a mean disorder
703 score for each residue based on these. In the resulting intrinsic disorder profile, the disorder
704 score of 0.5 is the threshold between order and disorder, where residues/regions above 0.5
705 are disordered, and residues/regions below 0.5 are ordered. The disorder score of 0.15 is the
706 threshold between order and flexibility, where residues/regions with the disorder scores above
707 0.15 but below 0.5 are flexible, and residues/regions below 0.15 are highly ordered.

708 Amino acid compositions of the intrinsically disordered C-terminal domain (residues 261-358)
709 of human CPSF6 and its different variants (CPSF6 ΔFG ΔLCR, CPSF6 ΔLCR, CPSF6 ΔFG,
710 and CPSF6 ADD2 ΔLCR) were analyzed to evaluate the relative abundance of prion-like low
711 complexity region (LCR) defining uncharged, charged, and Pro residues in these protein
712 regions. The corresponding values of the relative abundance of these residue groups were
713 calculated by dividing numbers of prion-like LCR defining uncharged (Ala, Gly, Val, Phe, Tyr,
714 Leu, Ile, Ser, Thr, Pro, Asn, Gln), charged (Asp, Glu, Lys, Arg), and Pro residues by total
715 number of amino acids in the corresponding protein fragments. As references, we used the
716 corresponding data for protein sequences deposited to the UniProtKB/Swiss-Prot database
717 that provides information on the overall distribution of amino acids in nature (Bairoch et al.,
718 2005); PDB Select 25 (Berman et al., 2000), which is a subset of structures from the Protein
719 Data Bank with less than 25% sequence identity, biased towards the composition of proteins
720 amenable to crystallization studies; and DisProt 3.4 that is comprised of a set of consensus
721 sequences of experimentally determined disordered regions (Sickmeier et al., 2007). Per-
722 residue intrinsic disorder propensities of the LCR-FG and ADD2-FG sequences were

723 evaluated by PONDR® VLXT (Romero et al., 2001), which is sensitive to local peculiarities of
724 the amino acid sequences freely available at http://www.pondr.com/ (accessed on August 03,
725 2024). Linear distribution of the net charge per residue (NCPR) within the LCR-FG and ADD2-
726 FG sequences were evaluated by CIDER (Holehouse et al., 2017), which is a webserver for
727 the analysis of a wide range of the physicochemical properties encoded by IDP sequences
728 freely available at http://pappulab.wustl.edu/CIDER (accessed on August 03, 2024).
729 Secondary structure propensities of the LCR-FG and ADD2-FG sequences were evaluated by
730 PSIPRED (McGuffin et al., 2000), which is a highly accurate secondary structure prediction
731 method freely available to non-commercial users at http://globin.bio.warwick.ac.uk/psipred/
732 (accessed on August 03, 2024).

733

734

735 **IMAGING AND STATISTICAL ANALYSIS**

736 All images were analysed using Fiji Software. More in detail, for the count of the nuclear
737 speckles a macro was computed to segment cellular nuclei and to select and count the nuclear
738 speckles. For the nuclei segmentation a Gaussian Blur with sigma=2 was used and for the
739 channel related to the nuclear speckles the set threshold was at 7000-46012 (min and max).
740 For the counting, a "size =10 – Infinity summarize" was used for the nuclei, whereas for the
741 nuclear speckles the size was reduced to "size=0-Infinity summarize".

742 For live imaging analysis, Arivis software was used to reconstruct the 3D movies.

743 All data were statistically analyzed with GraphPad Prism 9 (GraphPad Software, La Jolla
744 California USA, www.graphpad.com). Calculations were performed and figures were drawn
745 using Excel 365 or GraphPad Prism 8.0. Statistical analysis was performed, with Wilcoxon
746 matched paired t-tests or Mann–Whitney unpaired t-tests. Spearman correlation coefficients
747 (r) were calculated using GraphPad Prism.

748

749

750 **Figure legends**

751 **Figure 1. Role of HIV-induced CPSF6 puncta in the nuclear reverse transcription upon**
752 **removal of NVP. A) A)** THP-1 cells, infected with VSV-G ∆env HIV-1 (NL4.3) ∆R LUC (MOI
753 10) in presence or not of Nevirapine (10 μM) for 5 days, or in presence of Nevirapine (10 μM)
754 for 2 days and then the remaining 3 days without drug or in presence of Nevirapine (10 μM)
755 for 2 days then in presence of PF74 (25 μM). Confocal microscopy images, to verify the
756 presence of CPSF6 puncta, the cells are stained with anti-CPSF6 antibody (green). Nuclei are

757 stained with Hoechst (blue). Scale bar 10 µm. **B)** Luciferase Assay, to verify luciferase
758 expression in the aforementioned samples. Luciferase values were normalized by total
759 proteins revealed with the Bradford kit. One-way ANOVA statistical test with multiple
760 comparison was performed (****=p<0.0001; *=p<0.05; ns=p>0.05). Data are representative of
761 two independent experiments. **C)** Western blots demonstrate CPSF6 depletion using a specific
762 antibody against CPSF6 in THP-1 cells subjected to CRISPR Cas9 methods: CRISPR Cas9
763 bulk (left), and CRISPR Cas9 clones selected by limiting dilution (right). Each condition is
764 normalized for actin labeling. The ratio between the intensity signal of CPSF6 and actin was
765 analyzed via ImageJ and is plotted below each western blot. **D)** Confocal microscopy images
766 of THP-1 ctrl CRISPR clone 2 cells (Ctrl 2) and THP-1 duplex1-2-3 CRISPR clone 4 cells
767 (CPSF6 KO 4) infected with VSV-G / HIV-1ΔEnvIN$_{HA}$ LAI (BRU) -vpx (MOI 10) in the presence
768 of Nevirapine (10 µM). The cells are stained 30 h p.i. with anti-CPSF6 antibody and anti-HA
769 antibodies to detect HA tagged integrase (IN).
770

771 **Figure 2. Role of CPSF6 Domains in HIV-Induced CPSF6 Puncta. A)** Schema of CPSF6
772 isoform 588aa deletion mutants. **B)** Confocal microscopy images of THP-1 CPSF6 KO cells,
773 transduced with different mutants of CPSF6, infected with VSV-G / HIV-1ΔEnvIN$_{HA}$ LAI (BRU)
774 -vpx (MOI 10) in presence of Nevirapine (10 µM). The cells are stained with CPSF6 and HA
775 antibody 30 h p.i.. Scale bar 5µm. **C)** Analysis of the number of CPSF6 puncta in THP-1 CPSF6
776 KO cells transduced with different mutants of CPSF6, not infected or infected in presence of
777 Nevirapine (10 µM) (the number of analyzed cells is shown under the x-axis). Statistical test:
778 ordinary one-way ANOVA (****=p<0.0001; ***=p<0.001; *=p<0.05; ns=p>0.05). **D)** The plot
779 compares the number of CPSF6 puncta per cell in THP-1 CPSF6 KO cells transduced with
780 different mutants of CPSF6, infected with HIV-1 in the presence of Nevirapine (10 µM).
781 Statistical test: ordinary one-way ANOVA (****=p<0.0001; ns=p>0.05). **E)** Confocal microscopy
782 images of THP-1 CPSF6 KO clone 4, non-transduced and non-infected or transduced with WT
783 CPSF6 and CPSF6 3xNLSΔMCD and infected with VSV-G / HIV-1ΔEnvIN$_{HA}$ LAI (BRU) -vpx
784 (MOI 10) in presence of Nevirapine (10 µM). Immuno-RNA FISH: the cells are stained with
785 CPSF6 (green) antibody and with 24 probes against HIV-1 Pol sequence (gray)(RNA-FISH)
786 25 h p.i. Nuclei are stained with Hoechst (blue). Scale bar 10 µm. Violin plot presenting the
787 percentage of CPSF6 puncta colocalizing with the viral RNA in THP-1 CPSF6 KO clone 4 cells
788 transduced with LVs expressing CPSF6 WT or CPSF6 3xNLSΔMCD (respectively n=73 and
789 n=103) and infected with VSV-G / HIV-1ΔEnvIN$_{HA}$ LAI (BRU) -vpx (MOI 10) in presence of
790 Nevirapine (10 µM). A total of 198 CPSF6 WT puncta and 264 CPSF6 3xNLSΔMCD puncta
791 were counted. Statistical test: unpaired t-test, ns=p>0.05. **F)** Confocal microscopy images of
792 THP-1 KO CPSF6 cells transduced with WT CPSF6 and CPSF6 ΔMCD without NLS, with

793 3xNLS or with PY NLS, respectively. Cells were differentiated for 3 days, transduced with
794 CPSF6 lentiviral vectors (MOI 1) for 3 days and infected for 24 h with VSV-G / HIV-1ΔEnvIN$_{HA}$
795 LAI (BRU) -vpx (MOI 10) in presence of Nevirapine (10 uM) (left panels). The panels on the
796 right show transduced and uninfected cells. CPSF6 and the IN tagged with the HA are labeled
797 with anti-CPSF6 (green) and anti-HA (white) antibodies, respectively. Nuclei are stained with
798 Hoechst (blue). The arrows show CPSF6 puncta in colocalization with IN-HA. Scale bar 10µm.

799

800 **Figure 3. Evaluation of CPSF6 Deletion Mutants' Binding Capacity to the Viral Core. A)**
801 Ability of wild type and mutant CPSF6 proteins to bind to the HIV-1 core. Cellular extracts
802 derived from human 293T cells expressing similar levels of the indicated CPSF6 proteins
803 (INPUT) were incubated with HIV-1 capsid stabilized tubes for 1 hour at room temperatures in
804 the presence and absence of 10 µM PF74, as described in materials and methods. As a carrier
805 control, we utilized DMSO. Subsequently, HIV-1 capsid stabilized tubes were washed, and the
806 bound proteins eluted 1X Laemmli buffer 1X. The BOUND fractions were analysed by Western
807 blotting using antibodies against neon-GFP and the HIV-1 capsid. **B)** Experiments were
808 repeated at least three times and the average BOUND fraction relative to the INPUT fraction
809 normalized to wild type binding are shown for the different CPSF6 mutants. *** indicates a p-
810 value < 0.001; **** indicates a p-value < 0.0001; and ns indicates no significant difference as
811 determined by unpaired t-tests.

812 **Figure 4. Comparison of second structures of ADD2 and LCR. A)** Physicochemical
813 characteristics of the LCR-FG and ADD2-FG sequences. Intrinsic disorder predispositions
814 evaluated by PONDR® VLXT. Position of the FR segment within the LCR-FG and ADD2-FG
815 sequences is shown as gray shaded area. **B)** Linear distribution of the net charge per residue
816 (NCPR) within the LCR-FG sequence evaluated by CIDER. **C)** Linear distribution of the net
817 charge per residue (NCPR) within the ADD2-FG sequence evaluated by CIDER. **D)** Secondary
818 structure propensity of the LCR-FG sequence evaluated by PSIPRED. **E)** Secondary structure
819 propensity of the ADD2-FG sequence evaluated by PSIPRED. **F)** Analysis of the peculiarities
820 of the amino acid compositions of the intrinsically disordered C-terminal domain (residues 261-
821 358) of human CPSF6 and its different mutants. Relative abundance of prion-like LCR defining
822 uncharged residues in analyzed protein segments. **G)** Relative abundance of proline residues
823 in analyzed protein segments. **H)** Relative abundance of charged residues in analyzed protein
824 segments. The values were calculated by dividing numbers of prion-like LCR defining
825 uncharged (Ala, Gly, Val, Phe, Tyr, Leu, Ile, Ser, Thr, Pro, Asn, Gln), Pro, and charged (Asp,
826 Glu, Lys, Arg) residues by the total number of amino acids in the respective protein fragments.

827 Corresponding values for all protein sequences deposited in the UniProtKB/Swiss-Prot
828 database, PDB Select25, and DisProt are shown for comparison.

829

830 **Figure 5. Role of FG motif in viral replication. A)** WB showing CPSF6 protein from several
831 single clones derived from CPSF6 KO clone obtained upon complementation with CPSF6 ΔFG
832 and normalized with beta-actin. **B)** Quantification of the expression of CPSF6 ΔFG protein in
833 different single clones compared to CPSF6 WT (value 1). **C)** Infectivity essay using a single
834 round infectious virus carrying the cDNA of Luciferase as reporter gene. Values are expressed
835 as% of RLU compared to WT cells. **D)** Infectivity essay of a replication competent virus: vRNA
836 from new viruses produced after infection of WT, CPSF6 KO and CPSF6ΔFG cells was
837 analyzed and showed in the histograms as % of vRNA copies compared to vRNA in WT THP-
838 1 cells considered 100%.

839

840 **Figure 6. Depletion of MCD or LCR does not affect the formation of HIV-induced CPSF6**
841 **puncta. A)** Epifluorescence microscopy images of both infected and non-infected
842 differentiated THP-1 cells showing the presence of CPSF6 puncta only in the infected
843 condition. CPSF6 and SC35 are labelled with anti-CPSF6 (green) and anti-SC35 (red)
844 antibodies, respectively. Nuclei are stained with Hoechst (blue). Scale bar 10µm. **B)** Confocal
845 microscopy images of THP-1 KO CPSF6 cells, differentiated for 3 days, transduced with
846 CPSF6 lentiviral vector (MOI 1) (specifically WT CPSF6, CPSF6 ΔLCRs, CPSF6 ΔMCD with
847 3xNLS, without NLS, or with PY NLS) for 3 days and infected for 24 h with VSV-G / HIV-
848 1ΔEnvIN$_{HA}$ LAI (BRU) -vpx MOI 10) in presence of Nevirapine (10 uM). CPSF6 and nuclear
849 speckles were labelled with anti-CPSF6 (green) and anti-SC35 (red) antibodies, respectively.
850 Nuclei are stained with Hoechst (blue). Scale bar 10µm. The percentage of CPSF6 puncta
851 associated with SC35 per field of view is shown in the graph. N cells were counted in each
852 condition and a one-way ANOVA statistical test with multiple comparison was performed; ns=
853 p value > 0.05.

854

855

856 **Figure 7. Dynamics of the HIV-induced CPSF6 puncta formation and their fusion with**
857 **NSs. A)** Time course of infection of THP-1, 6h.p.i., 9h.p.i., 12h.p.i., 30h.p.i. or non infected.
858 Cells were stained with antibodies against CPSF6 (green) and SC35 (red). **B)** The graph
859 shows the percentage of CPSF6 puncta associated with NS or adjacent to NS or isolated from
860 NS at different time post-infection. **C)** The graph shows the progression of CPSF6 puncta

associated to NS during the time post-infection. N indicates the number of cells analyzed. One-way ANOVA statistical test with multiple comparison was performed; ns= p value > 0.05; **** indicates a p-value < 0.0001.

Figure 8. Role of SRRM2 and SON in the Formation of HIV-Induced CPSF6 puncta. A) Depletion of SON and SRRM2 in THP-1 cells using AUM*silence*™ ASO technology. The degree of depletion is quantified by WB and the mean intensity through immunofluorescence using antibodies against SON and SRRM2, respectively. Scale bar 5µm. **B)** The percentage of CPSF6 puncta formation is quantified by IF in THP-1 cells knocked down for SON, SRRM2, and control (Ctrl) infected with HIV-1 (MOI 25) for 48 h. CPSF6 is stained with an antibody against CPSF6 (green), and nuclei are stained with Hoechst (blue). The graph on the right reports the percentage of CPSF6 puncta calculated from more than 100 cells. Scale bar 10µm. Experiments were performed at least twice. **C)** (Top panels) confocal microscopy images of ∆IDR HaloTag SRRM2 HEK293 and HaloTag SRRM2 HEK293 cells stained with the halo tag ligand (red), and nuclei (blue). Scale bar 10µm. (Bottom panels) confocal microscopy images of HaloTag SRRM2 HEK293 and ∆IDR HaloTag SRRM2 HEK293 cells, both labeled with anti-SRRM2 (red) and anti-SON (gray) antibodies. Nuclei are stained with Hoechst (blue). Scale bar 10µm. Statistical studies are summarized in the violin plot which displays the distribution of the number of SON puncta per cell in the two conditions. N cells were counted and Kolmogorov Smirnov test was performed, ns= p>0.05. **D)** Confocal microscopy images of HaloTag SRRM2 HEK293 and ∆IDR HaloTag SRRM2 HEK293 cells, either non-infected or infected for 24 h with VSV-G / HIV-1ΔEnvIN$_{HA}$ LAI (BRU) (MOI 10) in the presence of Nevirapine (10 µM). CPSF6 and SC35 are labeled with anti-CPSF6 (red) and anti-SC35 (gray) antibodies, respectively. Nuclei are stained with Hoechst (blue). The plot shows the mean ± SD of the percentage of cells with CPSF6 puncta calculated in n fields of view (n=24, 29, 32); N is the number of cells analyzed for each of the three different cell lines; an unpaired t-test was performed, ****=p<0.0001, ns=p>0.05. Scale bar 10µm. Experiments were performed at list twice.

Movies 1 A and B. CPSF6 MLOs can form independently from SRRM2 MLOs in HEK293 SRRM2 HaloTag cells. Spinning Disk confocal images of HEK293 SRRM2 HaloTag cells acquired 9 hours after VSV-G / HIV-1ΔEnvIN$_{HA}$ LAI (BRU) (MOI 10) infection, occurred in presence of Nevirapine (10 µM). Cells were previously transduced for 24h with CPSF6-mNeonGreen lentiviral vector (MOI 0.5) to detect CPSF6 (visualized in green). SRRM2 MLOs were detected using the TMR-HaloTag Ligand (in red) and the nuclei were stained with

896 Hoechst (blue). Acquisitions were performed continuously for 27 minutes, which allowed to see
897 the independent formation of CPSF6 puncta, their migration toward the SRRM2 MLOs and the
898 final fusion of CPSF6 puncta and SRRM MLOs.

899 **Supplementary Figures**

900 **Suppl. Figure 1. A)** Multiple examples of confocal microscopy images of THP-1 ctrl CRISPR
901 clone 2 cells and THP-1 CPSF6 KO 4 cells infected with VSV-G / HIV-1ΔEnvIN$_{HA}$ LAI (BRU) -
902 vpx (MOI 10) in the presence of Nevirapine (10 μM). The cells are stained 30 h p.i. with CPSF6
903 (green) and HA (red) antibodies to detect integrase (IN). **B)** Multiple examples of THP-1 CPSF6
904 KO clone 4 cells transduced with different LVs carrying CPSF6 WT or mutants and stained
905 with CPSF6 and HA antibody 30 h p.i.. Scale bar 5μm.

906 **Suppl. Figure 2. A)** Per-residue intrinsic disorder propensity of the CPSF6 isoform 588 aa
907 evaluated by the Rapid Intrinsic Disorder Analysis Online platform (RIDAO) (Dayhoff and
908 Uversky, 2022) that yields results for IU-Pred_short (yellow line), IUPred_long (blue line),
909 PONDR® VL3 (green line), PONDR® VLXT (black line), PONDR® VSL2 (red line), and
910 PONDR® FIT (pink line) and computes a mean disorder score for each residue based on these
911 predictors (MDP, thick, dark pink, dashed line). Light pink shadow represents MDP error
912 distribution. The thin black line at the disorder score of 0.5 is the threshold between order and
913 disorder, where residues/regions with disorder scores above 0.5 are disordered, and
914 residues/regions below 0.5 are ordered. The dashed line at the disorder score of 0.15 is the
915 threshold between order and flexibility, where residues/regions above 0.15 are flexible, and
916 residues/regions below 0.15 are highly ordered (upper). Schema of the deletion mutants of
917 CPSF6 (bottom). **B)** Lentiviral Vector Transduction of PMA-Differentiated THP-1 Cells
918 expressing CPSF6 ΔMCD fused to mNeonGreen (left), CPSF6 NLSΔMCD fused to
919 mNeonGreen (center), CPSF6 3xNLSΔMCD fused to mNeonGreen (right). CPSF6 is
920 represented in green, and nuclei are stained in blue. Scale bar 5μm.

921 **Suppl. Figure 3.** Sequences of FG and LCRs or substituted amino acids sequences analyzed
922 in figure 4.

923 **Suppl. Figure 4. A)** THP-1 infected with VSV-G / HIV-1ΔEnvIN$_{HA}$ LAI (BRU) -vpx were labelled
924 after 6h, 9h, 12h, 30h p.i. with specific antibodies against CPSF6 (green) and SC35 (red),
925 nuclei are stained with Hoechst (blue). Scale bar 10μm. **B)** Timelapse of Movie B which
926 recapitulates the key steps of the dynamics of HIV induced-CPSF6 puncta biogenesis. CPSF6
927 was visualized transducing HEK293 SRRM2 HaloTag cells for 24h with CPSF6-mNeonGreen
928 lentiviral vector (MOI 0.5). SRRM2-HaloTag allowed the use of TMR-Halo Tag Ligand to
929 visualize the SRRM2 MLOs. After 9 hours after VSV-G / ΔEnvHIV-1 LAI (BRU) (MOI 10)
930 infection in presence of Nevirapine, CPSF6 puncta formed independently (9h 01m), migrating

afterwards to the SRRM2 MLOs and leading to a final fusion between CPSF6 puncta and SRRM2 MLOs (9h14m).

Suppl. Figure 5. Multiple examples (**A-E**) of confocal microscopy images of HaloTag SRRM2 HEK293 and ∆IDR HaloTag SRRM2 HEK293 cells, either non-infected or infected for 24 h with VSV-G / HIV-1ΔEnvIN$_{HA}$ LAI (BRU) (MOI 10) in the presence of Nevirapine (10 µM). CPSF6 and SC35 are labeled with anti-CPSF6 (red) and anti-SC35 (gray) antibodies, respectively. Nuclei are stained with Hoechst (blue). Scale bar 10µm.

Suppl. Figure 6. SRRM2 depletion is confirmed using antibodies against both SRRM2 and SC35 by IF. Statistical analysis: One-Way ANOVA (****=p<0.0001; *=p<0.05; ns=p>0.05).

Suppl. Figure 7. HEK 293 HaloTag SRRM2 or ∆IDR were labelled with Halo ligand (red) and an anti-SON antibody (gray), nuclei are stained by Hoechst (blue). Scale bar 10µm.

ACKNOWLEDGMENTS

F.D.N. is supported by the Institut Pasteur and ANRS grants (ECTZ192036, ECTZ137593), ANR-PRCI grant, Sidaction grant. C.T. is supported by fellowships Sidaction, ANR-PRCI. S.A. is supported by ANRS fellowship ECTZ204694. F.D.-G., B.C., M.R., C.L., C.B. are supported by NIH Grants R01AI087390 and R01AI150455. We gratefully acknowledge the UtechS Photonic BioImaging platform (Imagopole), C2RT at Institut Pasteur. We thank the NIH AIDS Reagents program to support us with precious reagents and Addgene.

Bibliography

Ay, S., Burlaud-Gaillard, J., Gazi, A., Tatirovsky, Y., Cuche, C., Diana, J.S., Scoca, V., Di Santo, J.P., Roingeard, P., Mammano, F.*, et al.* (2025). In vivo HIV-1 nuclear condensates safeguard against cGAS and license reverse transcription. EMBO J *44*, 166-199.

Ay, S., and Di Nunzio, F. (2023). HIV-Induced CPSF6 Condensates. J Mol Biol *435*, 168094.

Bairoch, A., Apweiler, R., Wu, C.H., Barker, W.C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M.*, et al.* (2005). The Universal Protein Resource (UniProt). Nucleic Acids Res *33*, D154-159.

Barre-Sinoussi, F., Chermann, J.C., Rey, F., Nugeyre, M.T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C.*, et al.* (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science *220*, 868-871.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. Nucleic Acids Res *28*, 235-242.

968 Blanco-Rodriguez, G., and Di Nunzio, F. (2021). The Viral Capsid: A Master Key to Access the Host
969 Nucleus. Viruses *13*.
970 Blanco-Rodriguez, G., Gazi, A., Monel, B., Frabetti, S., Scoca, V., Mueller, F., Schwartz, O., Krijnse-
971 Locker, J., Charneau, P., and Di Nunzio, F. (2020). Remodeling of the Core Leads HIV-1 Preintegration
972 Complex into the Nucleus of Human Lymphocytes. J Virol *94*.
973 Buffone, C., Martinez-Lopez, A., Fricke, T., Opp, S., Severgnini, M., Cifola, I., Petiti, L., Frabetti, S.,
974 Skorupka, K., Zadrozny, K.K.*, et al.* (2018). Nup153 Unlocks the Nuclear Pore Complex for HIV-1
975 Nuclear Translocation in Nondividing Cells. J Virol *92*.
976 Burdick, R.C., Li, C., Munshi, M., Rawson, J.M.O., Nagashima, K., Hu, W.S., and Pathak, V.K. (2020).
977 HIV-1 uncoats in the nucleus near sites of integration. Proc Natl Acad Sci U S A *117*, 5486-5493.
978 Chen, N.Y., Zhou, L., Gane, P.J., Opp, S., Ball, N.J., Nicastro, G., Zufferey, M., Buffone, C., Luban, J.,
979 Selwood, D.*, et al.* (2016). HIV-1 capsid is involved in post-nuclear entry steps. Retrovirology *13*, 28.
980 Chen, Y., Zhang, Y., Wang, Y., Zhang, L., Brinkman, E.K., Adam, S.A., Goldman, R., van Steensel, B., Ma,
981 J., and Belmont, A.S. (2018). Mapping 3D genome organization relative to nuclear compartments
982 using TSA-Seq as a cytological ruler. J Cell Biol *217*, 4025-4048.
983 Dayhoff, G.W., 2nd, and Uversky, V.N. (2022). Rapid prediction and analysis of protein intrinsic
984 disorder. Protein Sci *31*, e4496.
985 Dharan, A., Bachmann, N., Talley, S., Zwikelmaier, V., and Campbell, E.M. (2020). Nuclear pore
986 blockade reveals that HIV-1 completes reverse transcription and uncoating in the nucleus. Nature
987 Microbiology *5*, 1088-1095.
988 Di Nunzio, F. (2013). New insights in the role of nucleoporins: a bridge leading to concerted steps
989 from HIV-1 nuclear entry until integration. Virus Res *178*, 187-196.
990 Di Nunzio, F., Danckaert, A., Fricke, T., Perez, P., Fernandez, J., Perret, E., Roux, P., Shorte, S., Charneau,
991 P., Diaz-Griffero, F.*, et al.* (2012). Human nucleoporins promote HIV-1 docking at the nuclear pore,
992 nuclear import and integration. PLoS One *7*, e46037.
993 Di Nunzio, F., Uversky, V.N., and Mouland, A.J. (2023). Biomolecular condensates: insights into early
994 and late steps of the HIV-1 replication cycle. Retrovirology *20*, 4.
995 Dickson, C.F., Hertel, S., Tuckwell, A.J., Li, N., Ruan, J., Al-Izzi, S.C., Ariotti, N., Sierecki, E., Gambin, Y.,
996 Morris, R.G.*, et al.* (2024). The HIV capsid mimics karyopherin engagement of FG-nucleoporins.
997 Nature.
998 Dosztanyi, Z., Csizmok, V., Tompa, P., and Simon, I. (2005a). IUPred: web server for the prediction of
999 intrinsically unstructured regions of proteins based on estimated energy content. Bioinformatics *21*,
1000 3433-3434.
1001 Dosztanyi, Z., Csizmok, V., Tompa, P., and Simon, I. (2005b). The pairwise energy content estimated
1002 from amino acid composition discriminates between folded and intrinsically unstructured proteins. J
1003 Mol Biol *347*, 827-839.
1004 Durand, S., Nguyen, X.N., Turpin, J., Cordeil, S., Nazaret, N., Croze, S., Mahieux, R., Lachuer, J., Legras-
1005 Lachuer, C., and Cimarelli, A. (2013). Tailored HIV-1 vectors for genetic modification of primary human
1006 dendritic cells and monocytes. J Virol *87*, 234-242.
1007 Francis, A.C., Marin, M., Singh, P.K., Achuthan, V., Prellberg, M.J., Palermino-Rowland, K., Lan, S.,
1008 Tedbury, P.R., Sarafianos, S.G., Engelman, A.N.*, et al.* (2020). Publisher Correction: HIV-1 replication
1009 complexes accumulate in nuclear speckles and integrate into speckle-associated genomic domains.
1010 Nat Commun *11*, 6165.
1011 Fu, L., Weiskopf, E.N., Akkermans, O., Swanson, N.A., Cheng, S., Schwartz, T.U., and Görlich, D. (2024).
1012 HIV-1 capsids enter the FG phase of nuclear pores like a transport receptor. Nature *626*, 843-851.
1013 Fu, X.D., and Maniatis, T. (1990). Factor required for mammalian spliceosome assembly is localized to
1014 discrete regions in the nucleus. Nature *343*, 437-441.
1015 Gao, L., Zhang, W., Zhang, L., Gromova, B., Chen, G., Csizmadia, E., Cagle, C., Nastasio, S., Ma, Y.,
1016 Bonder, A.*, et al.* (2024). Silencing of aryl hydrocarbon receptor repressor restrains Th17 cell
1017 immunity in autoimmune hepatitis. Journal of Autoimmunity *143*, 103162.

1018 Greig, J.A., Nguyen, T.A., Lee, M., Holehouse, A.S., Posey, A.E., Pappu, R.V., and Jedd, G. (2020).
1019 Arginine-Enriched Mixed-Charge Domains Provide Cohesion for Nuclear Speckle Condensation. Mol
1020 Cell *77*, 1237-1250.e1234.
1021 Holehouse, A.S., Das, R.K., Ahad, J.N., Richardson, M.O., and Pappu, R.V. (2017). CIDER: Resources to
1022 Analyze Sequence-Ensemble Relationships of Intrinsically Disordered Proteins. Biophys J *112*, 16-21.
1023 Hülsmann, B.B., Labokha, A.A., and Görlich, D. (2012). The permeability of reconstituted nuclear
1024 pores provides direct evidence for the selective phase model. Cell *150*, 738-751.
1025 Ilik, I.A., Malszycki, M., Lubke, A.K., Schade, C., Meierhofer, D., and Aktas, T. (2020). SON and SRRM2
1026 are essential for nuclear speckle formation. eLife *9*.
1027 Jang, S., Bedwell, Gregory J., Singh, Satya P., Yu, Hyun J., Arnarson, B., Singh, Parmit K.,
1028 Radhakrishnan, R., Douglas, AidanDarian W., Ingram, Zachary M., Freniere, C.*, et al.* (2024). HIV-1
1029 usurps mixed-charge domain-dependent CPSF6 phase separation for higher-order capsid binding,
1030 nuclear entry and viral DNA integration. Nucleic Acids Research.
1031 Karlsson, E., Schnatwinkel, J., Paissoni, C., Andersson, E., Herrmann, C., Camilloni, C., and Jemth, P.
1032 (2022). Disordered Regions Flanking the Binding Interface Modulate Affinity between CBP and NCOA.
1033 J Mol Biol *434*, 167643.
1034 Lee, K., Ambrose, Z., Martin, T.D., Oztop, I., Mulky, A., Julias, J.G., Vandegraaff, N., Baumann, J.G.,
1035 Wang, R., Yuen, W.*, et al.* (2010). Flexible use of nuclear import pathways by HIV-1. Cell Host Microbe
1036 *7*, 221-233.
1037 Lelek, M., Casartelli, N., Pellin, D., Rizzi, E., Souque, P., Severgnini, M., Di Serio, C., Fricke, T., Diaz-
1038 Griffero, F., Zimmer, C.*, et al.* (2015). Chromatin organization at the nuclear pore favours HIV
1039 replication. Nat Commun *6*, 6483.
1040 Lester, E., Ooi, F.K., Bakkar, N., Ayers, J., Woerman, A.L., Wheeler, J., Bowser, R., Carlson, G.A.,
1041 Prusiner, S.B., and Parker, R. (2021). Tau aggregates are RNA-protein assemblies that mislocalize
1042 multiple nuclear speckle components. Neuron *109*, 1675-1691 e1679.
1043 Li, C., Burdick, R.C., Nagashima, K., Hu, W.S., and Pathak, V.K. (2021). HIV-1 cores retain their integrity
1044 until minutes before uncoating in the nucleus. Proc Natl Acad Sci U S A *118*.
1045 Link, J.O., Rhee, M.S., Tse, W.C., Zheng, J., Somoza, J.R., Rowe, W., Begley, R., Chiu, A., Mulato, A.,
1046 Hansen, D.*, et al.* (2020). Clinical targeting of HIV capsid protein with a long-acting small molecule.
1047 Nature *584*, 614-618.
1048 Luchsinger, C., Lee, K., Mardones, G.A., KewalRamani, V.N., and Diaz-Griffero, F. (2023). Formation of
1049 nuclear CPSF6/CPSF5 biomolecular condensates upon HIV-1 entry into the nucleus is important for
1050 productive infection. Sci Rep *13*, 10974.
1051 Marasca, F., Sinha, S., Vadalà, R., Polimeni, B., Ranzani, V., Paraboschi, E.M., Burattin, F.V., Ghilotti, M.,
1052 Crosti, M., Negri, M.L.*, et al.* (2022). LINE1 are spliced in non-canonical transcript variants to regulate
1053 T cell quiescence and exhaustion. Nature Genetics *54*, 180-193.
1054 Matreyek, K.A., Yucel, S.S., Li, X., and Engelman, A. (2013). Nucleoporin NUP153 phenylalanine-
1055 glycine motifs engage a common binding pocket within the HIV-1 capsid protein to mediate lentiviral
1056 infectivity. PLoS Pathog *9*, e1003693.
1057 Mazzeo, L., Ghosh, S., Di Cicco, E., Isma, J., Tavernari, D., Samarkina, A., Ostano, P., Youssef, M.K.,
1058 Simon, C., and Dotto, G.P. (2024). ANKRD1 is a mesenchymal-specific driver of cancer-associated
1059 fibroblast activation bridging androgen receptor loss to AP-1 activation. Nature Communications *15*,
1060 1038.
1061 McGuffin, L.J., Bryson, K., and Jones, D.T. (2000). The PSIPRED protein structure prediction server.
1062 Bioinformatics *16*, 404-405.
1063 Muller, T.G., Zila, V., Peters, K., Schifferdecker, S., Stanic, M., Lucic, B., Laketa, V., Lusic, M., Muller, B.,
1064 and Krausslich, H.G. (2021). HIV-1 uncoating by release of viral cDNA from capsid-like structures in the
1065 nucleus of infected cells. Elife *10*.
1066 Peng, K., Radivojac, P., Vucetic, S., Dunker, A.K., and Obradovic, Z. (2006). Length-dependent
1067 prediction of protein intrinsic disorder. BMC Bioinformatics *7*, 208.
1068 Peng, K., Vucetic, S., Radivojac, P., Brown, C.J., Dunker, A.K., and Obradovic, Z. (2005). Optimizing long
1069 intrinsic disorder predictors with protein evolutionary information. J Bioinform Comput Biol *3*, 35-60.

1070 Pornillos, O., Ganser-Pornillos, B.K., Kelly, B.N., Hua, Y., Whitby, F.G., Stout, C.D., Sundquist, W.I., Hill,
1071 C.P., and Yeager, M. (2009). X-ray structures of the hexameric building block of the HIV capsid. Cell
1072 *137*, 1282-1292.
1073 Price, A.J., Jacques, D.A., McEwan, W.A., Fletcher, A.J., Essig, S., Chin, J.W., Halambage, U.D., Aiken, C.,
1074 and James, L.C. (2014). Host cofactors and pharmacologic ligands share an essential interface in HIV-1
1075 capsid that is lost upon disassembly. PLoS Pathog *10*, e1004459.
1076 Rai, A.K., Chen, J.X., Selbach, M., and Pelkmans, L. (2018). Kinase-controlled phase transition of
1077 membraneless organelles in mitosis. Nature *559*, 211-216.
1078 Rasaiyaah, J., Tan, C.P., Fletcher, A.J., Price, A.J., Blondeau, C., Hilditch, L., Jacques, D.A., Selwood, D.L.,
1079 James, L.C., Noursadeghi, M.*, et al.* (2013). HIV-1 evades innate immune recognition through specific
1080 cofactor recruitment. Nature *503*, 402-405.
1081 Rensen, E., Mueller, F., Scoca, V., Parmar, J.J., Souque, P., Zimmer, C., and Di Nunzio, F. (2021).
1082 Clustering and reverse transcription of HIV-1 genomes in nuclear niches of macrophages. EMBO J *40*,
1083 e105247.
1084 Romero, P., Obradovic, Z., Li, X., Garner, E.C., Brown, C.J., and Dunker, A.K. (2001). Sequence
1085 complexity of disordered protein. Proteins *42*, 38-48.
1086 Scoca, V., Morin, R., Collard, M., Tinevez, J.Y., and Di Nunzio, F. (2023). HIV-induced membraneless
1087 organelles orchestrate post-nuclear entry steps. J Mol Cell Biol *14*.
1088 Segal-Maurer, S., DeJesus, E., Stellbrink, H.J., Castagna, A., Richmond, G.J., Sinclair, G.I., Siripassorn,
1089 K., Ruane, P.J., Berhe, M., Wang, H.*, et al.* (2022). Capsid Inhibition with Lenacapavir in Multidrug-
1090 Resistant HIV-1 Infection. N Engl J Med *386*, 1793-1803.
1091 Selyutina, A., Bulnes-Ramos, A., and Diaz-Griffero, F. (2018). Binding of host factors to stabilized HIV-1
1092 capsid tubes. Virology *523*, 1-5.
1093 Selyutina, A., Persaud, M., Lee, K., KewalRamani, V., and Diaz-Griffero, F. (2020). Nuclear Import of the
1094 HIV-1 Core Precedes Reverse Transcription and Uncoating. Cell Rep *32*, 108201.
1095 Sickmeier, M., Hamilton, J.A., LeGall, T., Vacic, V., Cortese, M.S., Tantos, A., Szabo, B., Tompa, P., Chen,
1096 J., Uversky, V.N.*, et al.* (2007). DisProt: the Database of Disordered Proteins. Nucleic Acids Res *35*,
1097 D786-793.
1098 Suzuki, Y., and Craigie, R. (2007). The road to chromatin — nuclear entry of retroviruses. Nature
1099 Reviews Microbiology *5*, 187-196.
1100 Taylor, I.A., and Fassati, A. (2024). The capsid revolution. J Mol Cell Biol *15*.
1101 Tsanov, N., Samacoits, A., Chouaib, R., Traboulsi, A.M., Gostan, T., Weber, C., Zimmer, C., Zibara, K.,
1102 Walter, T., Peter, M.*, et al.* (2016). smiFISH and FISH-quant - a flexible single RNA detection approach
1103 with super-resolution capability. Nucleic Acids Res *44*, e165.
1104 Wei, G., Iqbal, N., Courouble, V.V., Francis, A.C., Singh, P.K., Hudait, A., Annamalai, A.S., Bester, S.,
1105 Huang, S.W., Shkriabai, N.*, et al.* (2022). Prion-like low complexity regions enable avid virus-host
1106 interactions during HIV-1 infection. Nat Commun *13*, 5879.
1107 Xue, B., Dunbrack, R.L., Williams, R.W., Dunker, A.K., and Uversky, V.N. (2010). PONDR-FIT: a meta-
1108 predictor of intrinsically disordered amino acids. Biochim Biophys Acta *1804*, 996-1010.
1109 Yamashita, M., and Emerman, M. (2004). Capsid is a dominant determinant of retrovirus infectivity in
1110 nondividing cells. J Virol *78*, 5670-5678.
1111 Zhang, L., Zhang, Y., Chen, Y., Gholamalamdari, O., Wang, Y., Ma, J., and Belmont, A.S. (2021). TSA-seq
1112 reveals a largely conserved genome organization relative to nuclear speckles with small position
1113 changes tightly correlated with gene expression changes. Genome Res *31*, 251-264.
1114 Zhang, Z.-Y., Harischandra, D.S., Wang, R., Ghaisas, S., Zhao, J.Y., McMonagle, T.P., Zhu, G., Lacuarta,
1115 K.D., Song, J., Trojanowski, J.Q.*, et al.* TRIM11 protects against tauopathies and is down-regulated in
1116 Alzheimer's disease. Science *381*, eadd6696.
1117 Zila, V., Margiotta, E., Turonova, B., Muller, T.G., Zimmerli, C.E., Mattei, S., Allegretti, M., Borner, K.,
1118 Rada, J., Muller, B.*, et al.* (2021). Cone-shaped HIV-1 capsids are transported through intact nuclear
1119 pores. Cell *184*, 1032-1046 e1018.

1120

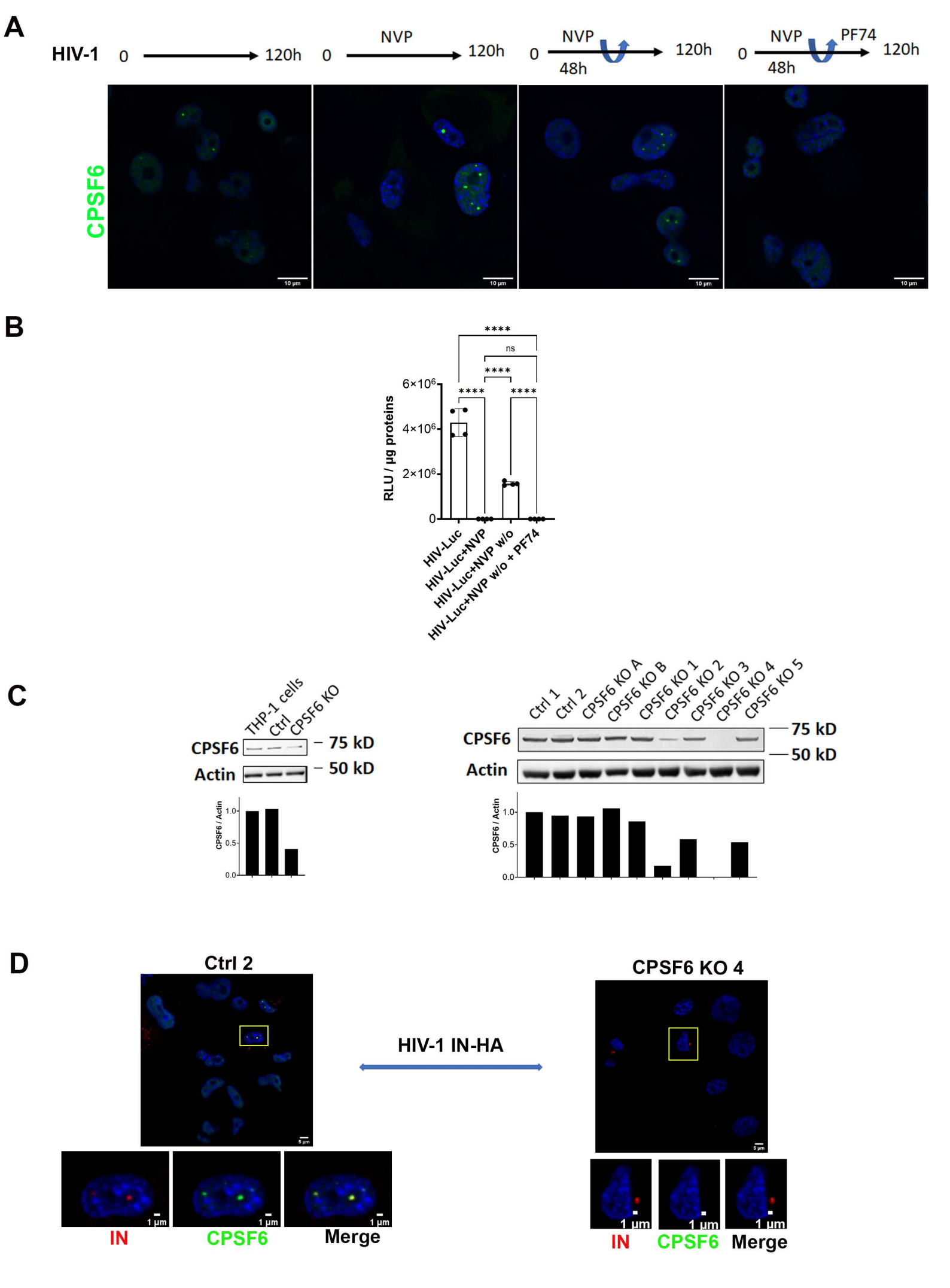

A
HIV-1
0
120h
NVP
48h
PF74
CPSF6
10 µm
B
RLU / µg proteins
6×10⁶
4×10⁶
2×10⁶
0
****
ns
HIV-Luc
HIV-Luc+NVP
HIV-Luc+NVP w/o
HIV-Luc+NVP w/o + PF74
C
THP-1 cells
Ctrl
CPSF6 KO
CPSF6
Actin
75 kD
50 kD
CPSF6 / Actin
1.0
0.5
0.0
Ctrl 1
Ctrl 2
CPSF6 KO A
CPSF6 KO B
CPSF6 KO 1
CPSF6 KO 2
CPSF6 KO 3
CPSF6 KO 4
CPSF6 KO 5
D
Ctrl 2
CPSF6 KO 4
HIV-1 IN-HA
1 µm
IN
CPSF6
Merge

Figure 1

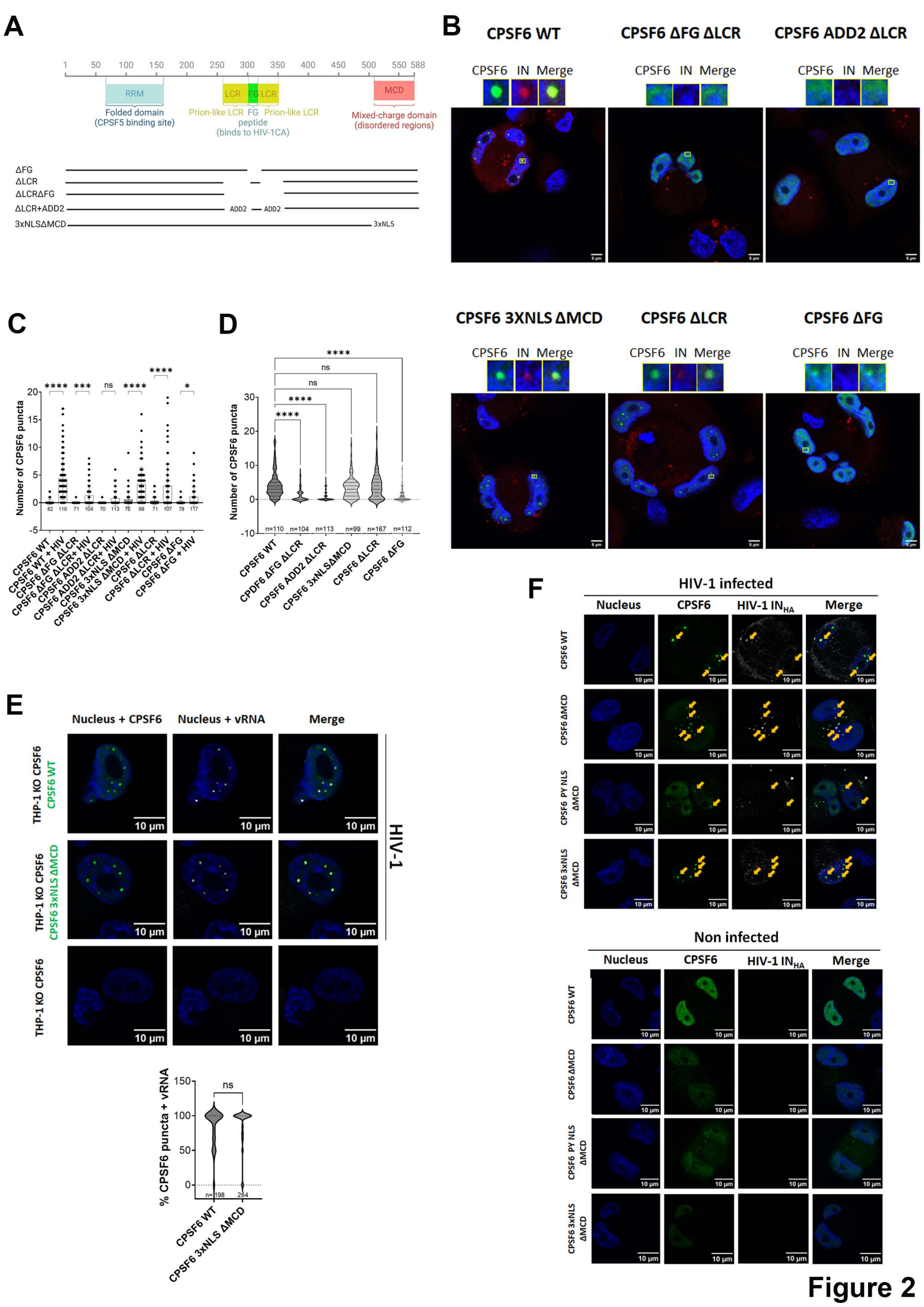
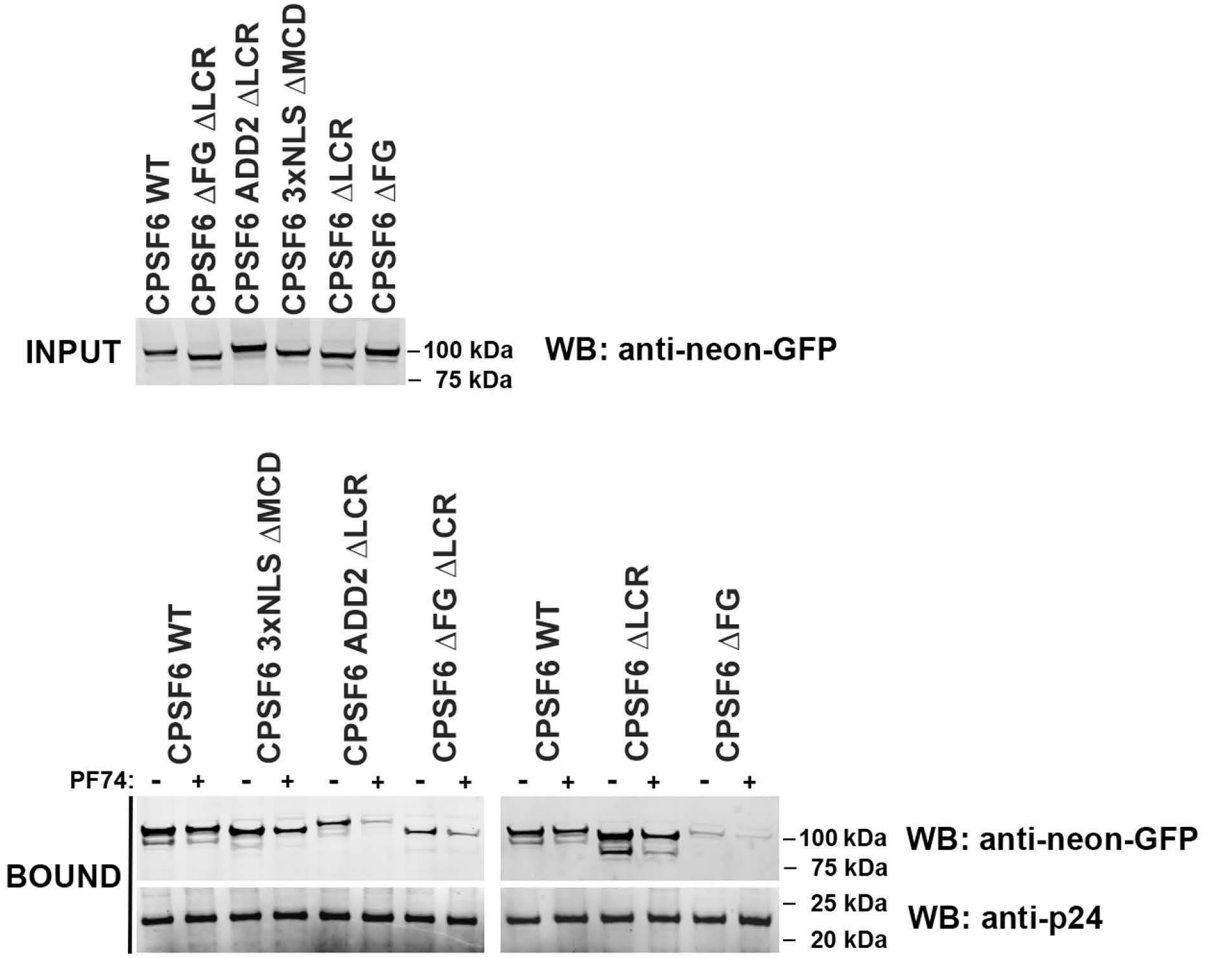

A
1 50 100 150 200 250 300 350 400 450 500 550 588
RRM
LCR FG LCR
MCD
Folded domain (CPSF5 binding site)
Prion-like LCR FG Prion-like LCR peptide (binds to HIV-1CA)
Mixed-charge domain (disordered regions)
ΔFG
ΔLCR
ΔLCRΔFG
ΔLCR+ADD2
ADD2 ADD2
3xNLSΔMCD
3xNLS
B
CPSF6 WT
CPSF6 ΔFG ΔLCR
CPSF6 ADD2 ΔLCR
CPSF6 3XNLS ΔMCD
CPSF6 ΔLCR
CPSF6 ΔFG
CPSF6 IN Merge
C
Number of CPSF6 puncta
D
Number of CPSF6 puncta
n=110 n=104 n=113 n=99 n=167 n=112
CPSF6 WT
CPDF6 ΔFG ΔLCR
CPSF6 ADD2 ΔLCR
CPSF6 3xNLSΔMCD
CPSF6 ΔLCR
CPSF6 ΔFG
E
Nucleus + CPSF6
Nucleus + vRNA
Merge
THP-1 KO CPSF6 CPSF6 WT
THP-1 KO CPSF6 CPSF6 3xNLS ΔMCD
THP-1 KO CPSF6
HIV-1
10 µm
% CPSF6 puncta + vRNA
ns
n=198 264
F
HIV-1 infected
Nucleus
CPSF6
HIV-1 IN$_{HA}$
Merge
CPSF6 WT
CPSF6 ΔMCD
CPSF6 PY NLS ΔMCD
CPSF6 3xNLS ΔMCD
Non infected

Figure 2

A


CPSF6 WT
CPSF6 ΔFG ΔLCR
CPSF6 ADD2 ΔLCR
CPSF6 3xNLS ΔMCD
CPSF6 ΔLCR
CPSF6 ΔFG
INPUT
100 kDa
75 kDa
WB: anti-neon-GFP
CPSF6 WT
CPSF6 3xNLS ΔMCD
CPSF6 ADD2 ΔLCR
CPSF6 ΔFG ΔLCR
CPSF6 WT
CPSF6 ΔLCR
CPSF6 ΔFG
PF74: - + - + - + - + - + - + - +
BOUND
100 kDa
75 kDa
WB: anti-neon-GFP
25 kDa
20 kDa
WB: anti-p24

B

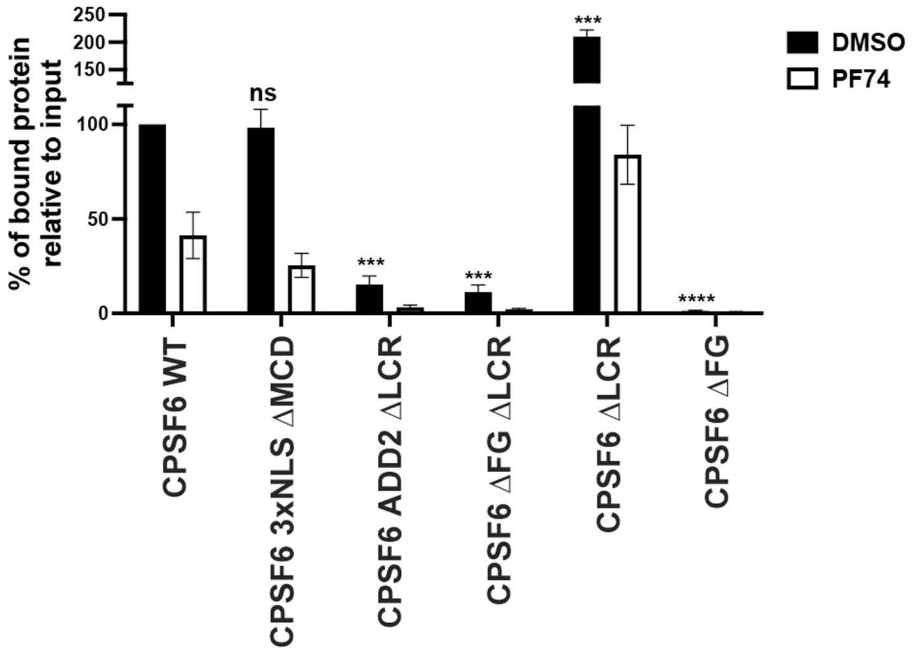
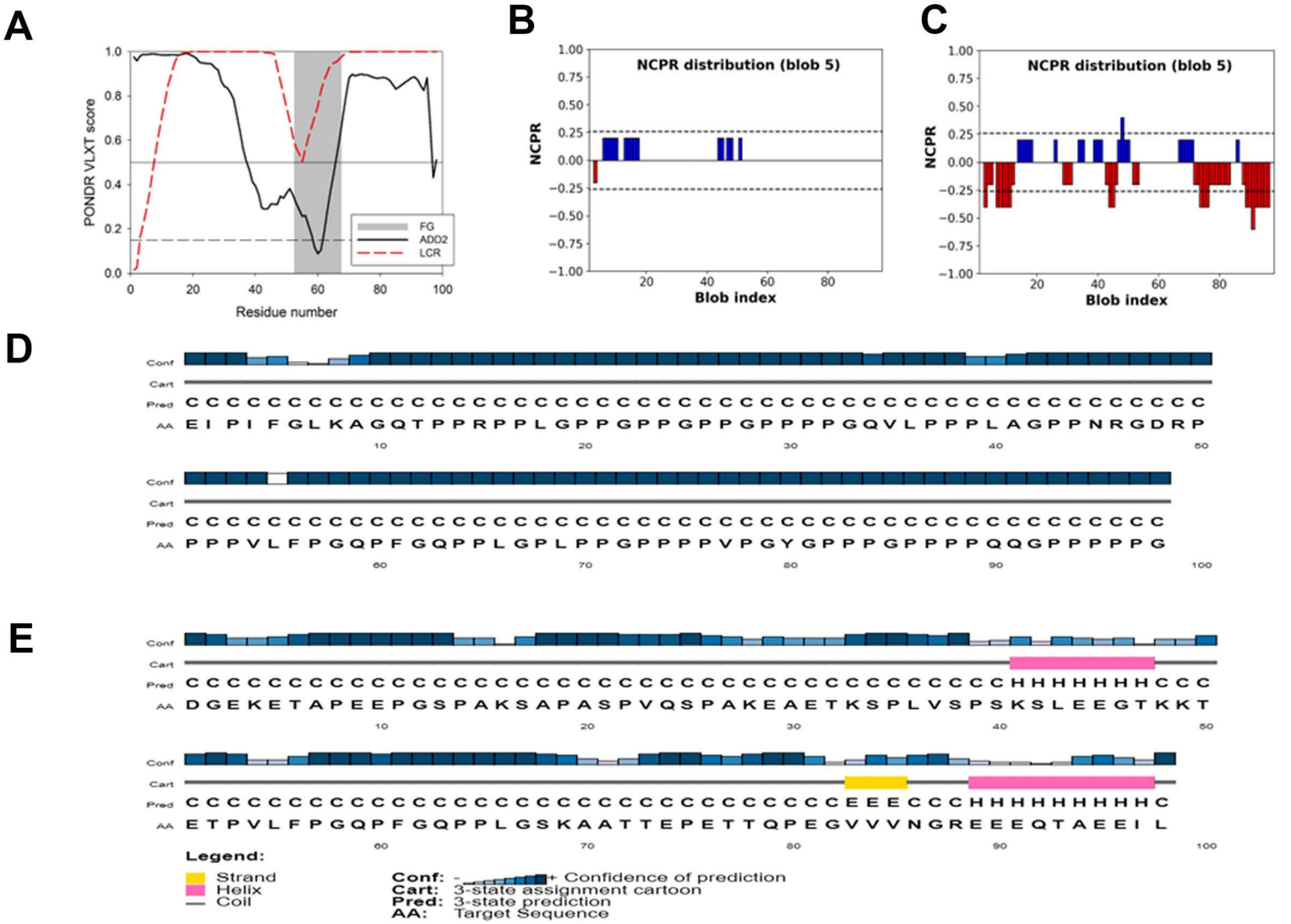

% of bound protein relative to input
250
200
150
100
50
0
ns
***
***
***
****
DMSO
PF74
CPSF6 WT
CPSF6 3xNLS ΔMCD
CPSF6 ADD2 ΔLCR
CPSF6 ΔFG ΔLCR
CPSF6 ΔLCR
CPSF6 ΔFG

Figure 3


A
PONDR VLXT score
Residue number
FG
ADD2
LCR
B
NCPR distribution (blob 5)
NCPR
Blob index
C
NCPR distribution (blob 5)
NCPR
Blob index
D
Conf
Cart
Pred
AA
E
Legend:
Strand
Helix
Coil
Conf: Confidence of prediction
Cart: 3-state assignment cartoon
Pred: 3-state prediction
AA: Target Sequence

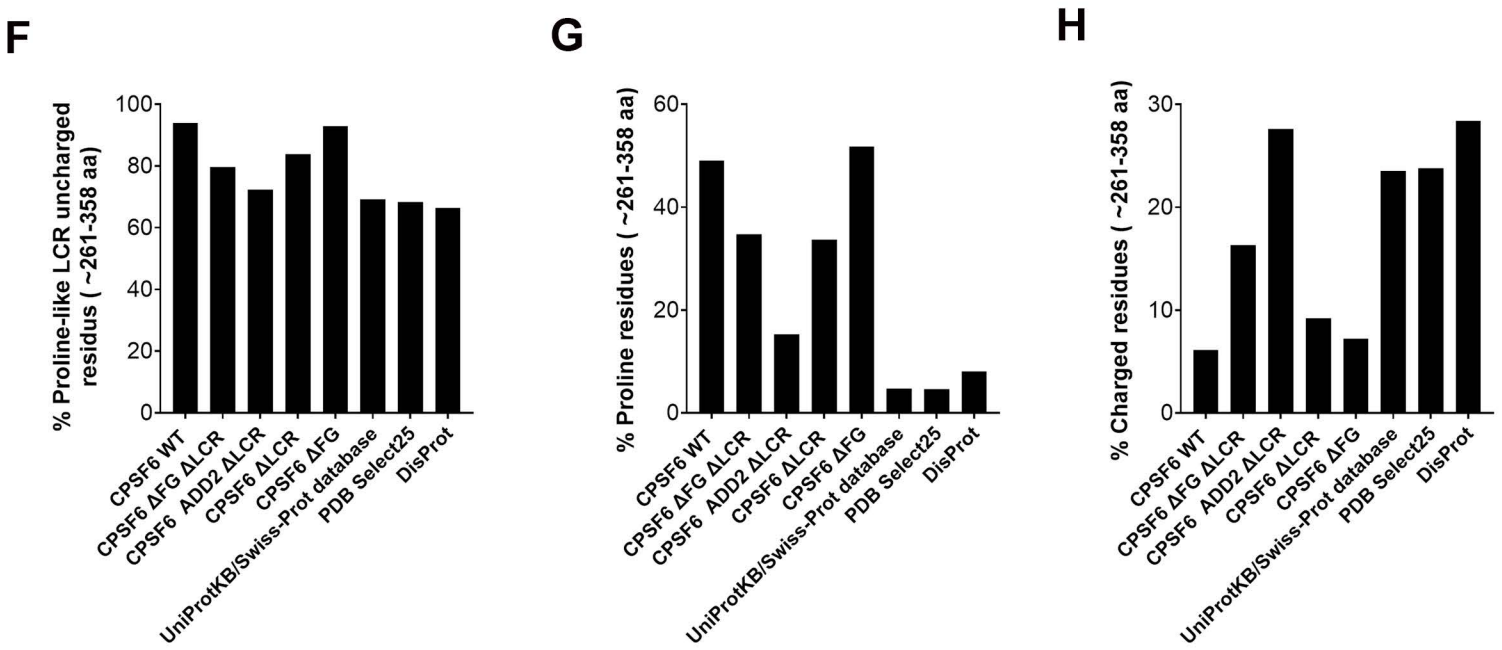
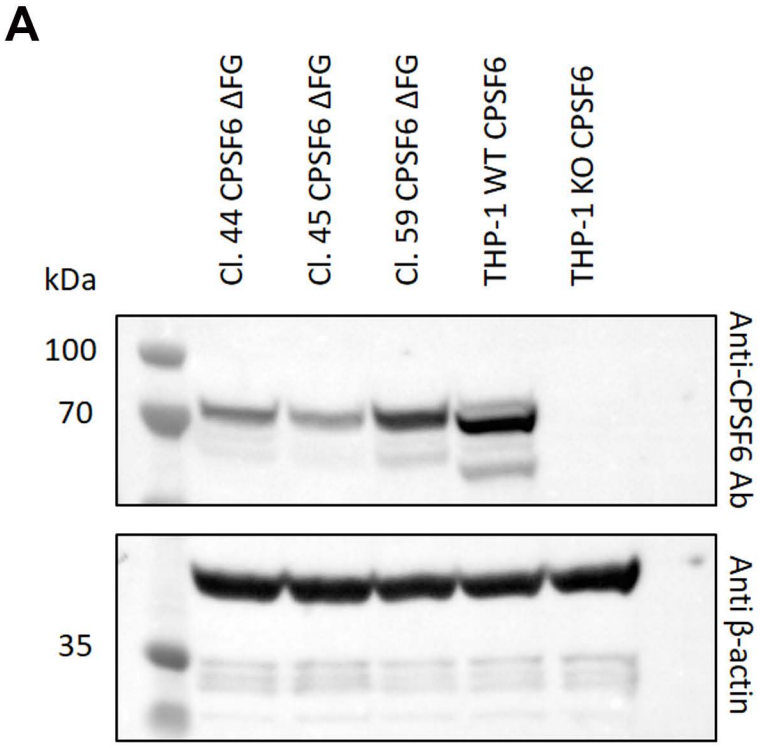
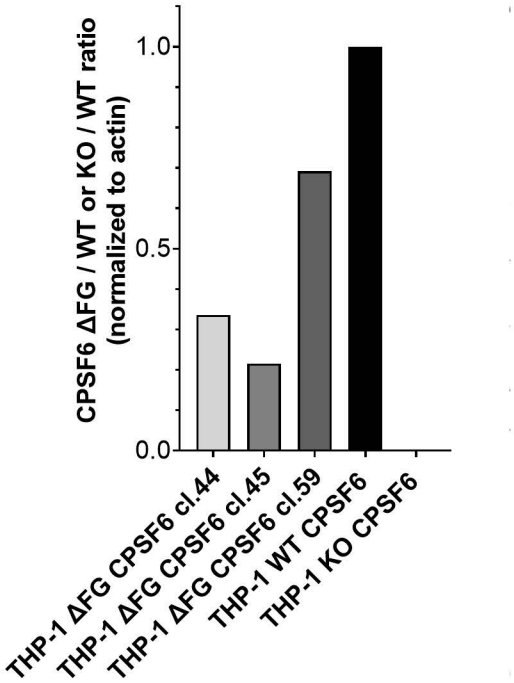
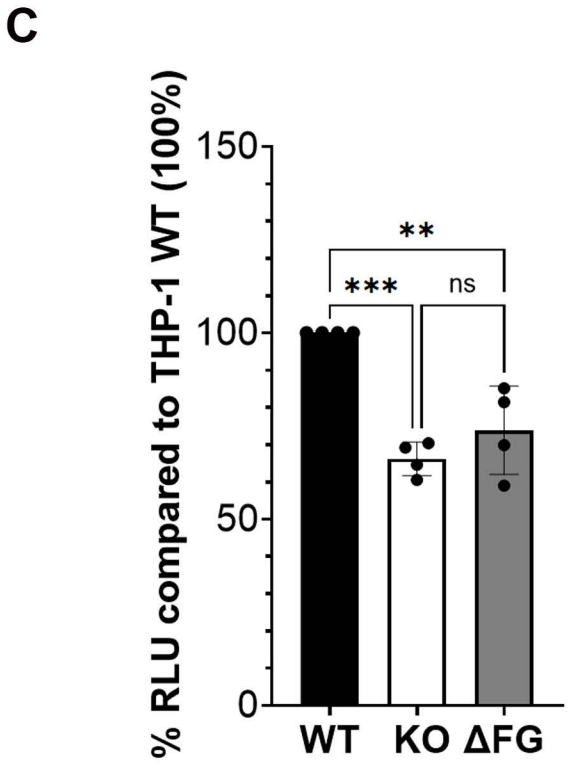

F
% Proline-like LCR uncharged residus (~261-358 aa)
G
% Proline residues (~261-358 aa)
H
% Charged residues (~261-358 aa)
CPSF6 WT
CPSF6 ΔFG ΔLCR
CPSF6 ADD2 ΔLCR
CPSF6 ΔLCR
CPSF6 ΔFG
UniProtKB/Swiss-Prot database
PDB Select25
DisProt

Figure 4


A
Cl. 44 CPSF6 ΔFG
Cl. 45 CPSF6 ΔFG
Cl. 59 CPSF6 ΔFG
THP-1 WT CPSF6
THP-1 KO CPSF6
kDa
100
70
35
Anti-CPSF6 Ab
Anti β-actin


B
CPSF6 ΔFG / WT or KO / WT ratio
(normalized to actin)
1.0
0.5
0.0
THP-1 ΔFG CPSF6 cl.44
THP-1 ΔFG CPSF6 cl.45
THP-1 ΔFG CPSF6 cl.59
THP-1 WT CPSF6
THP-1 KO CPSF6


C
% RLU compared to THP-1 WT (100%)
150
100
50
0
**
***
ns
WT
KO
ΔFG

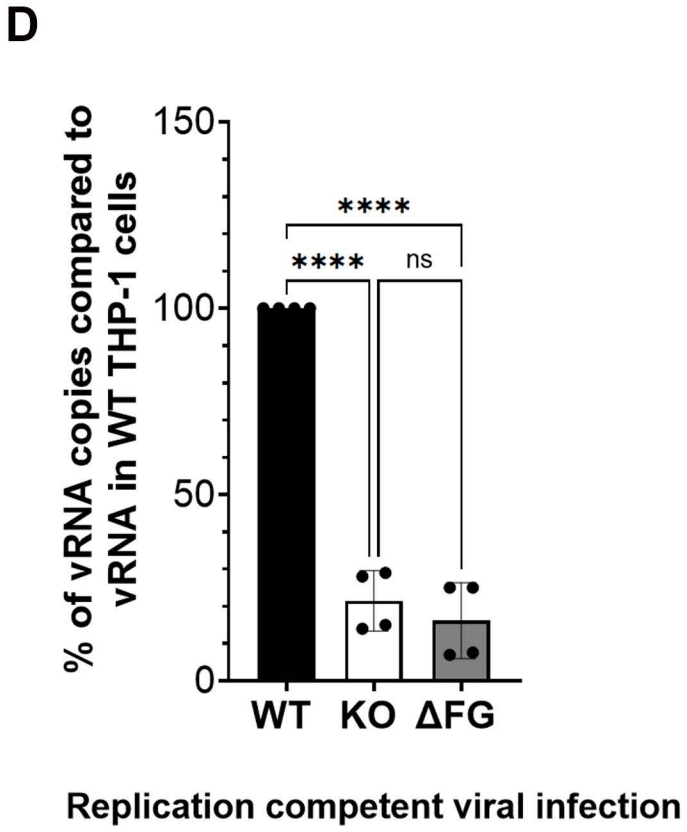
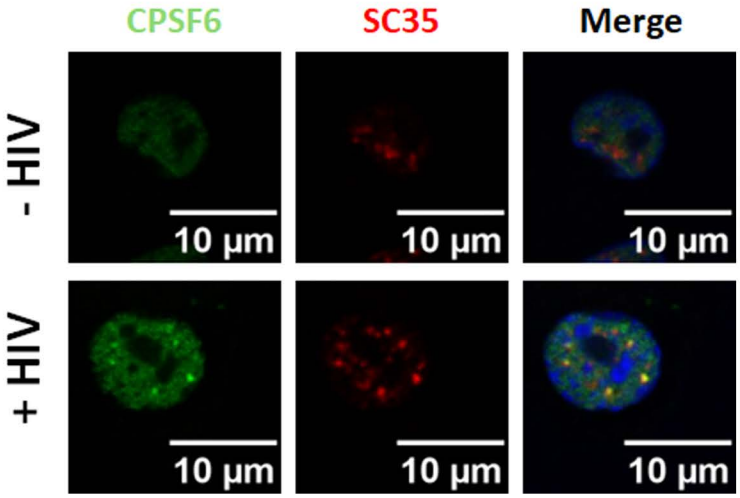
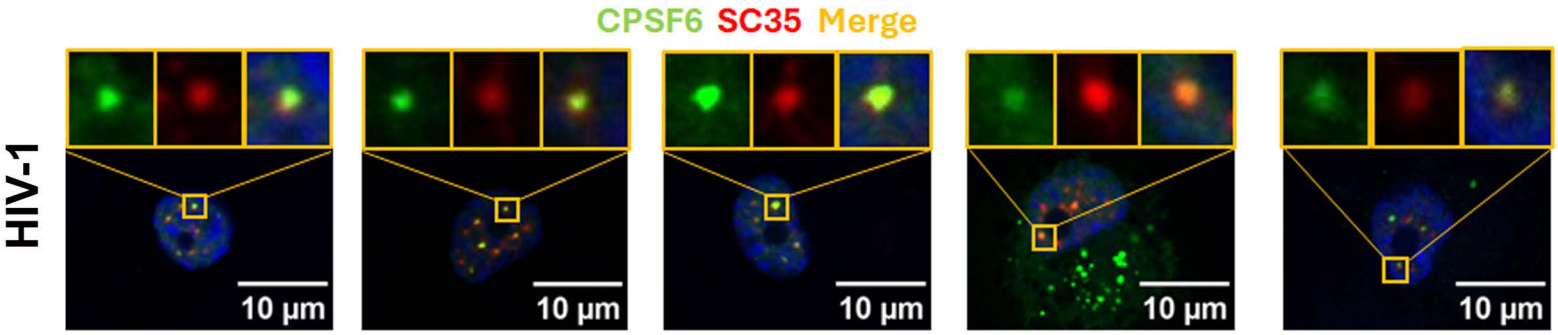

D
% of vRNA copies compared to vRNA in WT THP-1 cells
150
100
50
0
****
****
ns
WT
KO
ΔFG
Replication competent viral infection

Figure 5

A


CPSF6
SC35
Merge
- HIV
+ HIV
10 µm

B


CPSF6 SC35 Merge
HIV-1
10 µm
CPSF6 WT
CPSF6 3xNLS ΔMCD
CPSF6 ΔLCR
CPSF6 ΔMCD
CPSF6 PY NLS ΔMCD

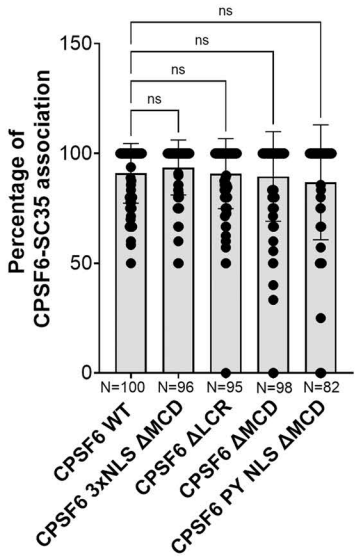

Percentage of CPSF6-SC35 association
150
100
50
0
ns
N=100
N=96
N=95
N=98
N=82
CPSF6 WT
CPSF6 3xNLS ΔMCD
CPSF6 ΔLCR
CPSF6 ΔMCD
CPSF6 PY NLS ΔMCD

Figure 6

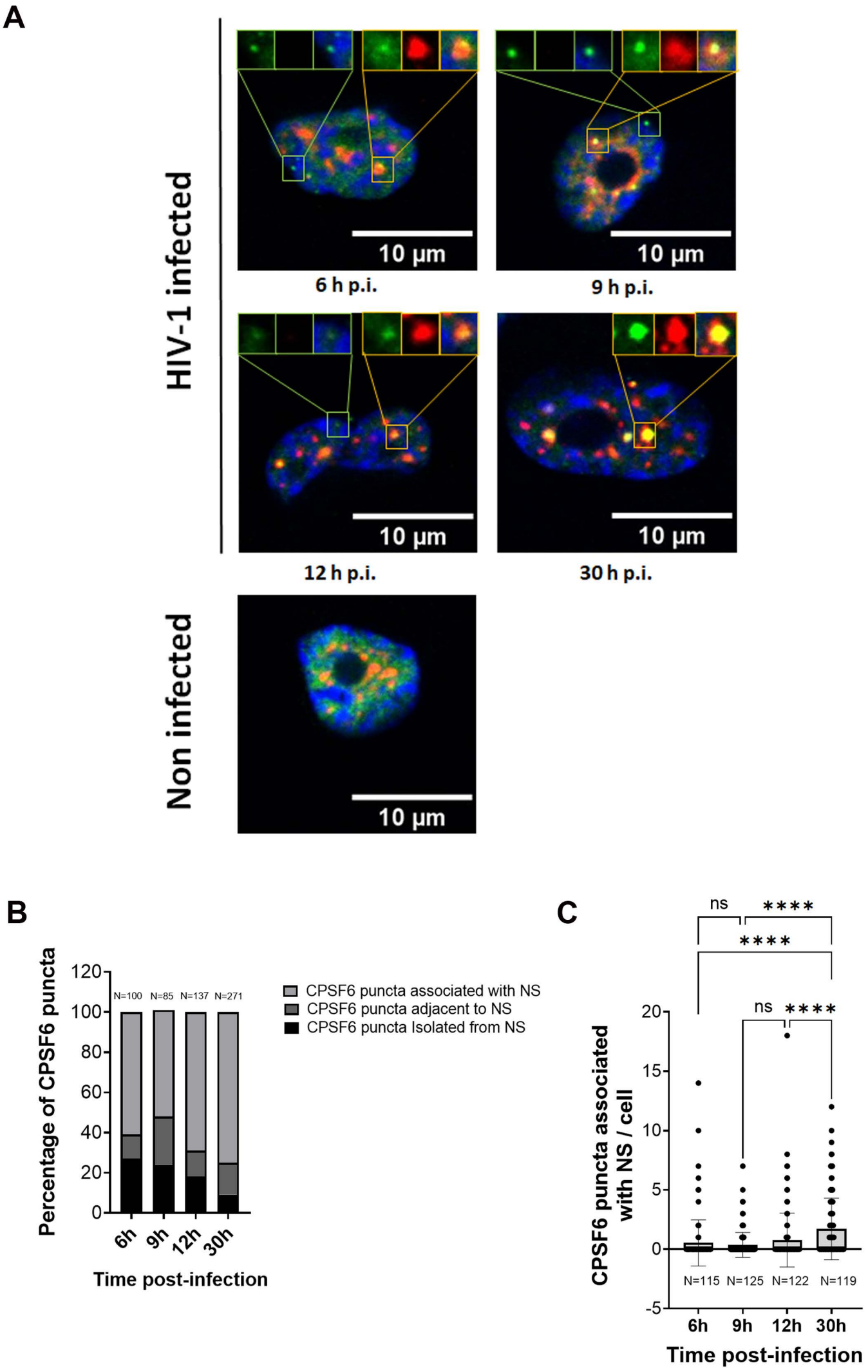

A
HIV-1 infected
10 µm
6 h p.i.
9 h p.i.
12 h p.i.
30 h p.i.
Non infected
B
Percentage of CPSF6 puncta
N=100 N=85 N=137 N=271
CPSF6 puncta associated with NS
CPSF6 puncta adjacent to NS
CPSF6 puncta Isolated from NS
6h 9h 12h 30h
Time post-infection
C
ns
****
CPSF6 puncta associated with NS / cell
N=115 N=125 N=122 N=119
6h 9h 12h 30h
Time post-infection

Figure 7

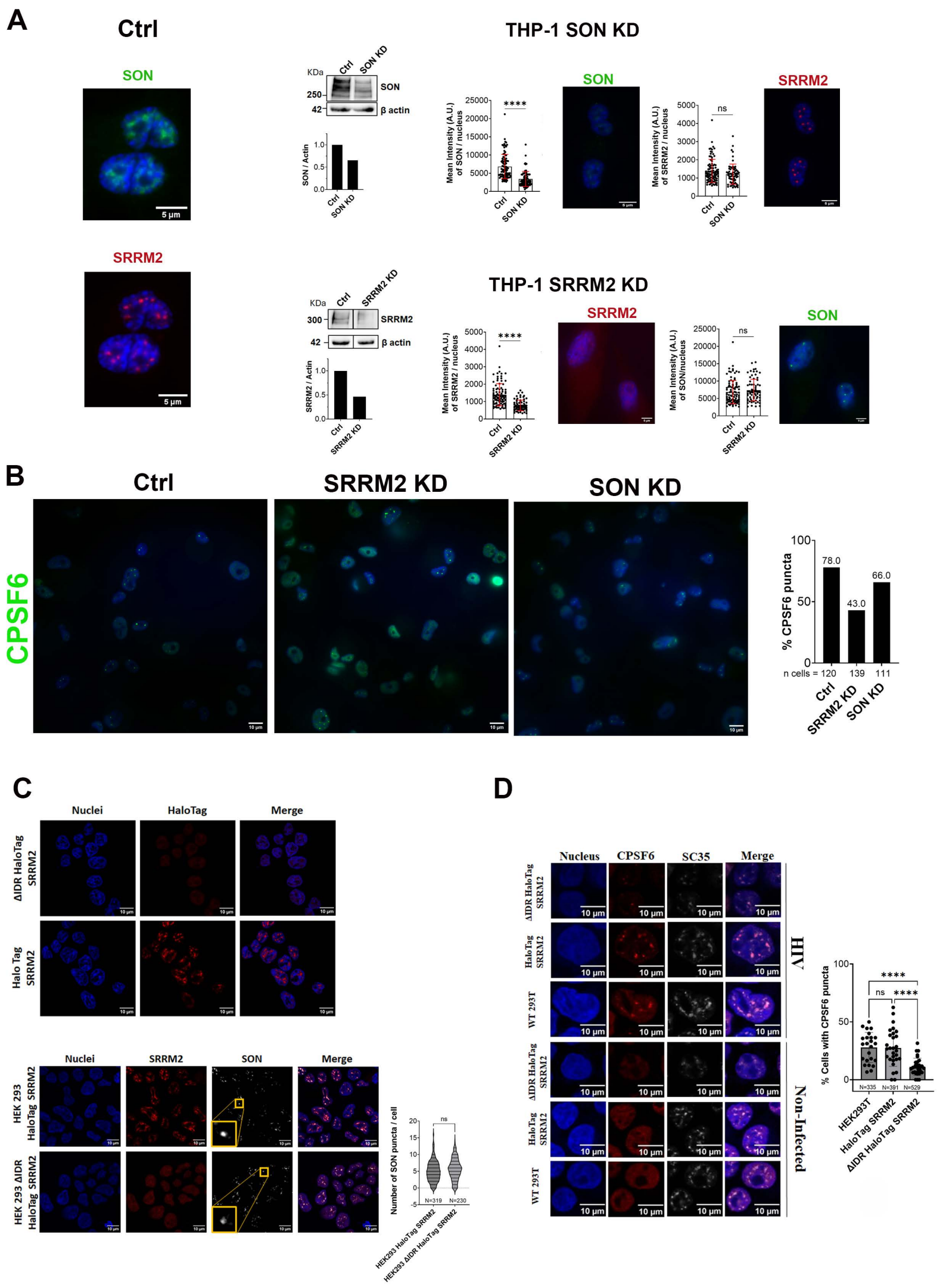

A
Ctrl
THP-1 SON KD
SON
SRRM2
5 µm
KDa
Ctrl
SON KD
250
42
β actin
SON / Actin
1.0
0.5
0.0
Mean Intensity (A.U.) of SON / nucleus
25000
20000
15000
10000
5000
0
****
Mean Intensity (A.U.) of SRRM2 / nucleus
5000
4000
3000
2000
1000
ns
THP-1 SRRM2 KD
SRRM2 KD
300
SRRM2 / Actin
Mean Intensity (A.U.) of SON/nucleus
B
Ctrl
SRRM2 KD
SON KD
CPSF6
10 µm
% CPSF6 puncta
100
50
78.0
43.0
66.0
n cells = 120
139
111
C
Nuclei
HaloTag
Merge
ΔIDR HaloTag SRRM2
Halo Tag SRRM2
SRRM2
SON
HEK 293 HaloTag SRRM2
HEK 293 ΔIDR HaloTag SRRM2
Number of SON puncta / cell
20
15
10
5
0
-5
N=319
N=230
HEK293 HaloTag SRRM2
HEK293 ΔIDR HaloTag SRRM2
D
Nucleus
CPSF6
SC35
WT 293T
HIV
Non-Infected
% Cells with CPSF6 puncta
N=335
N=391
N=529
HEK293T
HaloTag SRRM2
ΔIDR HaloTag SRRM2

Figure 8