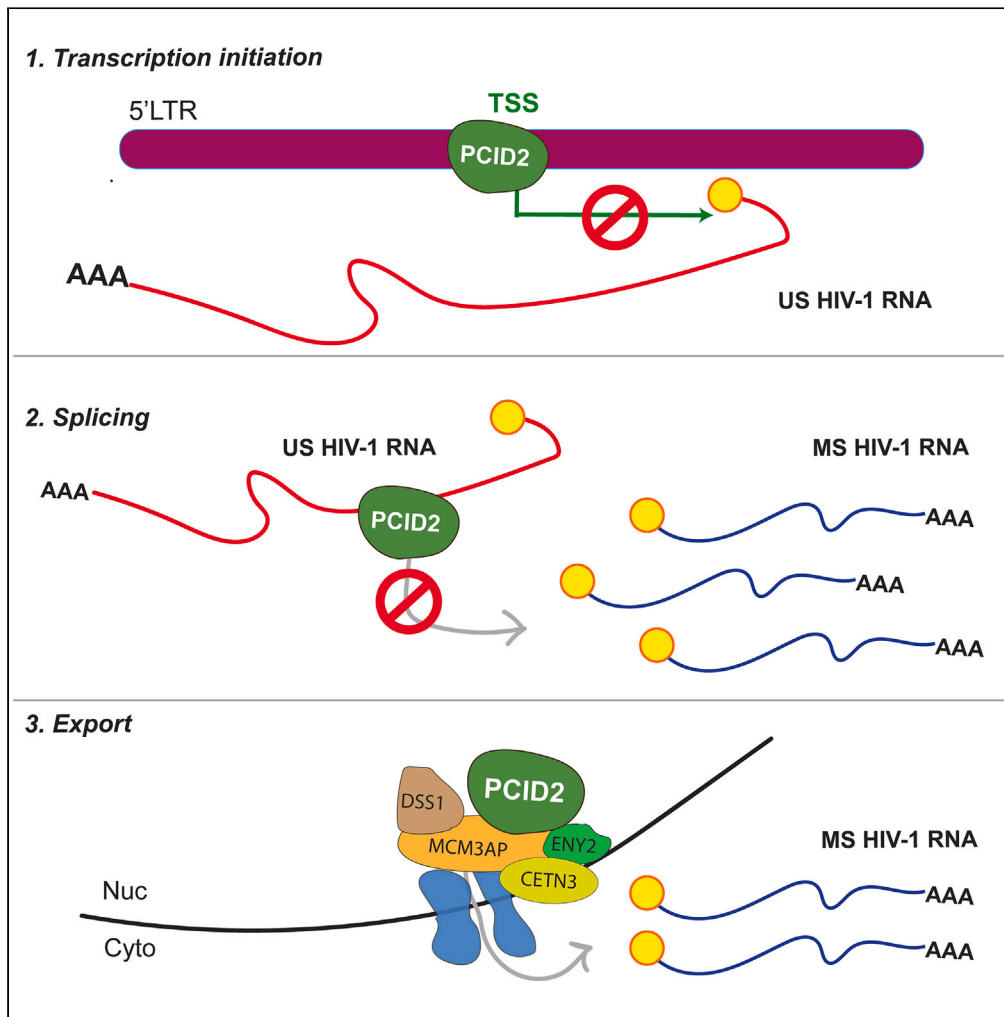**Article**

PCID2 dysregulates transcription and viral RNA processing to promote HIV-1 latency

Raquel Crespo,
Enrico Ne, Julian
Reinders, ...,
Robert-Jan Palstra,
Shringar Rao,
Tokameh
Mahmoudi

t.mahmoudi@erasmusmc.nl

Highlights

PCID2 is bound to the latent HIV-1 LTR as a transcriptional repressor

PCID2 enforces latency by acting on transcription initiation

PCID2 establishes blocks to alternative splicing during HIV-1 latency

PCID2 misregulates alternative splicing in cells obtained from people with HIV-1

Article

PCID2 dysregulates transcription and viral RNA processing to promote HIV-1 latency

Raquel Crespo,[1] Enrico Ne,[1] Julian Reinders,[1] Jenny I.J. Meier,[1] Chengcheng Li,[1] Sanne Jansen,[1] Alicja Górska,[1] Selin Koçer,[1] Tsung Wai Kan,[1,2,3] Wouter Doff,[4] Dick Dekkers,[4] Jeroen Demmers,[4] Robert-Jan Palstra,[1,2,3] Shringar Rao,[1] and Tokameh Mahmoudi[1,2,3,5,*]

SUMMARY

HIV-1 latency results from tightly regulated molecular processes that act at distinct steps of HIV-1 gene expression. Here, we characterize PCI domain-containing 2 (PCID2) protein, a subunit of the transcription and export complex 2 (TREX2) complex, to enforce transcriptional repression and post-transcriptional blocks to HIV-1 gene expression during latency. PCID2 bound the latent HIV-1 LTR (long terminal repeat) and repressed transcription initiation during latency. Depletion of PCID2 remodeled the chromatin landscape at the HIV-1 promoter and resulted in transcriptional activation and latency reversal. Immunoprecipitation coupled to mass spectrometry identified PCID2-interacting proteins to include negative viral RNA (vRNA) splicing regulators, and PCID2 depletion resulted in over-splicing of intron-containing vRNA in cell lines and primary cells obtained from PWH. MCM3AP and DSS1, two other RNA-binding TREX2 complex subunits, also inhibit transcription initiation and vRNA alternative splicing during latency. Thus, PCID2 is a novel HIV-1 latency-promoting factor, which in context of the TREX2 sub-complex PCID2-DSS1-MCM3AP blocks transcription and dysregulates vRNA processing.

INTRODUCTION

The persistence of latent HIV-1 viral reservoirs in memory CD4[+] T cells that elude elimination by the immune system and cannot be targeted by antiretroviral therapy is considered the biggest obstacle toward an HIV-1 cure.[1–3] Latency is established upon integration of the viral genome into the host cell genome and is defined by the absence of viral production.[3] The establishment and maintenance of viral latency is a layered multifactorial process that is enforced by distinct host cell pathways acting on chromatin regulation, transcription initiation and elongation, and post-transcriptional mechanisms regulating RNA metabolism and viral protein production.[4,5]

Transcription of the HIV-1 provirus is directed by its promoter in the 5′ long terminal repeat (LTR) whose activity is controlled by the chromatin landscape and relies on the availability of viral protein Tat and host cell transcription (co-)factors.[5–8] During latency, the viral protein Tat is absent, which causes RNA Pol II to pause at the TAR RNA element (+70) prohibiting transcription elongation.[9] Further blocks to HIV-1 gene expression occur at a post-transcriptional level, including inhibition of viral RNA (vRNA) splicing, nucleocytoplasmic export, translation, and trafficking (reviewed by Crespo R. et al.[4]). The contribution of these RNA processing mechanisms to the overall repression of HIV-1 gene expression has recently been characterized in both *in vitro* models of HIV-1 latency and CD4[+] T cells obtained from people living with HIV-1 (PLWH),[10–14] where it has been shown that blocks to HIV-1 gene expression occur largely at post-transcriptional levels. More recently, several unbiased and candidate studies, including our own, have successfully identified co- and post-transcriptional regulators of HIV-1 latency that control vRNA metabolism and either promote or inhibit viral gene expression.[12,15,16]

In a previous study, we performed an HIV-1 LTR locus-specific dCas9 chromatin immunoprecipitation coupled to mass spectrometry (Catchet-MS) in latent and active cells and identified putative LTR-bound transcriptional regulators of HIV-1 latency.[16] Interestingly, while a substantial number of latency-associated factors identified were, as expected, DNA associated, we also identified many putative RNA-binding latent HIV-1 promoter enriched regulators involved in RNA metabolism, including CRNKL1, recently described to post-transcriptionally regulate HIV-1 gene expression,[17] and other known regulators of HIV-1 gene expression such as DDX24.[18]

Catchet-MS also identified the host protein PCI domain containing 2 (PCID2) to be enriched at the HIV-1 promoter in the latent state. PCID2 (human homolog of yeast protein Thp1) is mainly described in literature as part of the multi-subunit transcription and export complex

[1]Department of Biochemistry, Erasmus University Medical Center, Ee622 PO Box 2040, 3000 CA Rotterdam, the Netherlands
[2]Department of Pathology, Erasmus University Medical Center, Rotterdam, the Netherlands
[3]Department of Urology, Erasmus University Medical Center, Rotterdam, the Netherlands
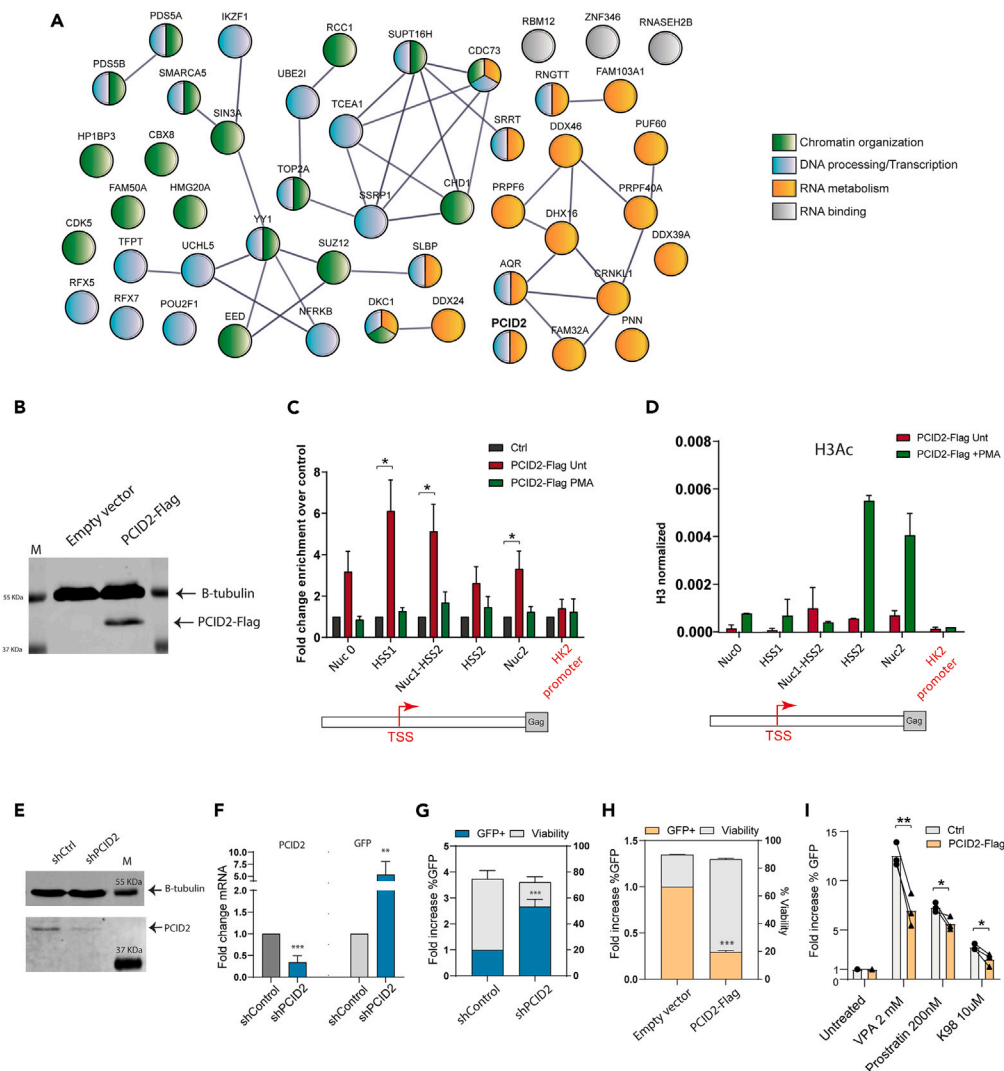[4]Proteomics Center, Erasmus University Medical Center, Ee679a PO Box 2040, 3000 CA Rotterdam, the Netherlands
[5]Lead contact
*Correspondence: t.mahmoudi@erasmusmc.nl
https://doi.org/10.1016/j.isci.2024.109152

Figure 1. PCID2 is enriched at the latent HIV-1 LTR and is a repressor of HIV-1 gene expression during latency

(A) STRING network of proteins identified to be enriched in the latent HIV-1 LTR by the Catchet-MS pipeline as published in Ne, Crespo et al. 2022 Nuc Acid Res. Proteins are colored based on GO analysis enrichment as indicated in the figure. Gray lines represent known interaction (experiments) as reported in the STRING database.

(B) Western blot of PCID2-Flag in control or stably PCID2-Flag expressing J-Lat 11.1. Beta tubulin was used as loading control.

(C) Enrichment of PCID2-Flag at the HIV-1 LTR expressed as fold enrichment in control and untreated and PMA treated PCID2-Flag expressing J-Lat 11.1 lines as assessed by chromatin immunoprecipitation (ChIP) coupled with quantitative PCR. Primers spanning across the HIV-1 promoter were used to assess relative enrichment of PCID2-Flag at sequential regions of the LTR. HK2 promoter was used as a control genomic region. Relative enrichment was normalized to Input. Mean and SD correspond to 4 ChIP replicates with independent chromatin preparations for control, untreated, and PMA treated cells.

(D) ChIP-qPCR analysis of acetylated histone 3 (F) enrichment at the HIV-1 promoter in untreated and PMA treated stably PCID2-Flag expressing J-Lat 11.1 lines. H3Ac enrichment at the HIV-1 LTR is represented as relative enrichment normalized to Total H3 as shown in Figure S1C. ChIP-qPCR corresponds to one chromatin preparation, bars and error lines represent, respectively, mean and SD of technical duplos.

(E) Western blot analysis of PCID2 in control and PCID2 knockdown J-Lat 11.1. Beta-tubulin was used as a loading control. Cells were infected with a VSV-pseudotyped lentivirus containing a scramble shRNA (shControl) or a PCID2 mRNA-targeting shRNA (shPCID2).

(F) Gene expression analysis of shRNA-mediated knockdown of PCID2 and GFP mRNA fold induction in shPCID2 cells relative to shControl and normalized to cyclophilin A. Bars represent mean of 5 independent shRNA-mediated knockdown experiments and error lines represent SEM (n = 4). Statistical significance was determined by t test; **p < 0.01, ***p < 0.001.

(G) Fold increase in the % of GFP (left y axes) and viability (right y axes) in shControl or shPCID2 J-Lat 11.1 cells as measured by flow cytometry. Bars represent mean of 4 independent shRNA-mediated knockdown experiments and error lines represent SEM. Statistical significance was determined by t-test; ***p < 0.001. Raw values for percentage of GFP are available in Table S2.

Figure 1. Continued

(H) Fold decrease in the % of GFP (left y axes) and viability (right y axes) in control or transiently PCID2-Flag overexpressing J-Lat 11.1 cells as measured by flow cytometry. Bars represent mean of 3 collections and error lines represent SEM. Statistical significance was determined by t test; ***p < 0.001. Raw values for percentage of GFP are available in Table S2.

(I) Fold change in % of GFP in control or PCID2-Flag overexpressing J-Lat 11.1 after a 48 h treatment of with latency reversing agents valproic acid (VPA), prostratin, and K98 as measured by flow cytometry. Fold increase in % GFP was assessed by normalization to untreated control. Bars represent mean and error lines represent SEM. Statistical significance was determined by t test; **p < 0.01. Raw values for percentage of GFP are available in Table S2.

2 (TREX2), composed of PCID2, GANP, DSS1, ENY2 and CETN2/3 subunits, which links transcription with mRNA export by binding to nascent RNA containing messenger ribonucleoprotein (mRNP) complexes and mediating their transport from transcription sites to the nuclear pore for subsequent export.[19–23] Outside of the TREX2 complex, other functions have been ascribed to PCID2, including roles in genome instability and preventing R-loop formation,[24,25] in protein stability[26,27] and transcription regulation of lymphoid commitment genes.[28]

Here, we demonstrate that PCID2 plays a dual role during latency, acting both at the level of HIV-1 transcription initiation and viral RNA processing steps, resulting in repression of HIV-1 gene expression. Chromatin immunoprecipitation (ChIP) experiments indicate that, while enriched at the latent HIV-1 LTR, PCID2 is removed upon transcriptional activation and acts as a repressor of HIV-1 gene expression by inhibiting LTR-driven transcription initiation. Downregulation of PCID2 leads to overall de-repression of the HIV-1 LTR locus as indicated by removal or deposition of distinct chromatin marks concomitant with recruitment of activating transcription factors to the promoter. In a transcription-independent mechanism, PCID2 also promotes latency by blocking post-transcriptional steps of HIV-1 gene expression. Immunoprecipitation of PCID2 coupled to mass spectrometry identified PCID2 interaction partners involved in RNA splicing, export, stability, and translation. Interestingly, downregulation of PCID2 led to the over-splicing of HIV-1 vRNAs and misregulated expression of HIV-1 RNA splice variants in cell lines and primary cells obtained from people living with HIV-1, indicating that PCID2 promotes HIV-1 latency by inhibiting alternative HIV-1 RNA splicing. We find that TREX2 subunits MCM3AP and DSS1 also enforce viral latency by inhibiting transcription initiation and viral RNA splicing, indicating that PCID2-mediated repression of HIV-1 gene expression occurs in context of the RNA-binding PCID2-DSS1-MCM3AP sub-complex and providing a putative role for the TREX2 complex. Consistent with its function within the TREX2 complex, PCID2 regulates nucleocytoplasmic export of completely spliced viral RNA species. Thus, PCID2-containing complexes play a dual role to promote HIV-1 latency by blocking both transcription initiation and vRNA splicing and facilitate export of completely spliced vRNA species.

RESULTS

PCID2 is an HIV-1 LTR-bound latency promoting factor

In a previous study,[16] we identified putative latent promoter-bound regulators of HIV-1 gene expression by dCas9-targeted chromatin immunoprecipitation coupled to mass spectrometry (Catchet-MS) (Figure 1A). Because of differential enrichment at the latent HIV-1 and inactive LTR, we hypothesized that these factors may contribute to the maintenance of HIV-1 latency as putative repressors of HIV-1 gene expression. Interestingly, more than half of the latent HIV-1 LTR-bound proteins identified were categorized to be RNA binding and/or involved in host RNA metabolism (Figure 1A). Among these, endogenous PCID2 was found by the Catchet-MS pipeline to be preferentially associated with the repressed HIV-1 promoter in latent J-Lat 11.1 cells.

We first set out to confirm whether PCID2, is physically bound to the HIV-1 promoter in the latent state. For this, we generated stably transfected J-Lat 11.1 cells which express PCID2-Flag (F-PCID2 J-Lat 11.1), given lack of availability of specific antibodies to purify endogenous PCID2. We validated the exogenous expression of PCID2 by western blot analysis (Figure 1B), and RT-PCR confirmed expression levels comparable to endogenous PCID2 (Figure S1A). Consistent with Catchet-MS, which identified endogenous PCID2 enriched at the latent HIV-1 LTR, chromatin immunoprecipitation using anti-FLAG M2 beads coupled to qPCR with primers spanning the HIV-1 LTR showed that PCID2 is recruited to the HIV-1 LTR during latency (Figures 1C, S1B, and S1C; Table 1) and enriched at the transcription start site (TSS). We also performed ChIP-qPCR in PCID2-Flag expressing cells treated with the mitogen/PKC agonist phorbol 12-myristate 13-acetate (PMA) that causes strong general transcriptional activation in lymphoid cells (Figure 1C). Upon PMA treatment and transcriptional activation, PCID2 is removed from the HIV-1 promoter (Figure 1C), consistent with its putative role as a repressor and its loss from the LTR in PMA-treated J-Lat 11.1, also observed in Catchet-MS. As expected, PMA treatment of PCID2-Flag expressing cells resulted in an increase in acetylated histone 3 (Figures 1D and S1D–S1F), indicative of an open chromatin state and active transcription.

To investigate the role of PCID2 during HIV-1 latency, we depleted endogenous PCID2 in J-Lat 11.1 cells with short hairpin RNAs targeting PCID2 mRNA (shPCID2) or scramble control (shCtrl) (Figures 1E, 1F, S1G, and S1H) and observed a significant increase both in the levels of GFP mRNA and the percentage of cells expressing GFP protein as measured by flow cytometry (Figures 1F, 1G, S1G, and S1H). To rule out potential contribution of clonal effects, we also depleted PCID2 in J-Lat full-length clone 10.6 (Figures S2A and S2B) and LTR-Tat-GFP J-Lat clone A2 (Figures S2C and S2D) and confirmed the presence of significant latency reversal upon PCID2 knockdown. Consistent with these results, transient overexpression of PCID2 decreased the percentage of GFP expressing cells (Figure 1H) and prevented viral reactivation in J-Lat 11.1 cells treated with the latency reversing agents valproic acid (HDAC inhibitor), prostratin (PKC agonist), and K98 (BAF inhibitor)[29] (Figure 1I). To further validate that PCID2 negatively regulates HIV-1 gene expression, we transiently co-transfected uninfected Jurkat cells with a plasmid containing an *env* deficient full-length HIV-1 genome and a luciferase reporter replacing the *nef* gene (pNL4.3.Luc.R-E-) together with either an empty pBud-control or PCID2 expressing pBud-PCID2-Flag plasmid (Figure S3A). We observed a significant dose-dependent decrease in LTR-driven luciferase activity in PCID2 transiently overexpressing cells as compared to control cells, demonstrating

Table 1. List of primers used for quantitative PCR

Amplicon	Forward primer 5′-3′	Reverse primer 5′-3′
Nuc 0	CCACACACAAGGCTACTTCC	AACTGGTACTAGCTTGTAGCAC
HSS1	TGTGAGCCTGCATGGGATGG	GAAAGTCCCCAGCGGAAAGT
Nuc1-HSS2	CGTCTGTTGTGTGACTCTGGT	TCGAGAGAGCTCCTCTGGTT
HSS2	GCCCGAACAGGGACTTGAAA	TTGGCGTACTCACCAGTCG
Nuc2	GGTGCGAGAGCGTCAGTAT	AGCTCCCTGCTTGCCCATA
HK2 promoter	GCCGACTCTTGTATTGCCTG	TATTGTAGCACGGCCGGAAA
PCID2	AACACACCAGCTGTCTCTGG	GCCTTTGACGTGTCCCATGT
GFP	GAAGCAGCACGACTTCTTCAA	GCTTGTCGGCCATGATATAGA
Readthrough	GCCCTCAGATGCTGCATATAA	AGAGTCACACAACAGACGG
TAR	GTCTCTCTGGTTAGACCAG	TGGGTTCCCTAGTTAGCC
UTR	GCTTCAAGTAGTGTGTGCCC	GAGAGCTCCTCTGGTTTCCC
Tat-Rev/CS vRNA	GACTCATCAAGTTTCTCTATCAAA	AGTCTCTCAAGCGGTGGT
US vRNA	GACGCTCTCGCACCCATCTC	CTGAAGCGCGCACGGCAA
SS vRNA	GCGGCGACTGGAAGAAGC	CTATGATTACTATGGACCACAC
18S RNA	AAACGGCTACCACATCCAAG	CCTCCAATGGATCCTCGTTA
B2-microglobulin	AGCGTACTCCAAAGATTCAGGTT	ATGATGCTGCTTACATGTCTCGAT
GAPDH	GCAGGGGGGAGCCAAAAGGG	TGCCAGCCCCAGCGTCAAAG
PPIA	TCATCTGCACTGCCAAGACTG	CATGCCTTCTTTCACTTTGCC

that PCID2 prohibits HIV-1 gene expression in a dose-dependent manner (Figure S3B). This effect seems to be specific to HIV-1 gene expression, as overexpression of PCID2 does not affect HSP promoter-driven luciferase activity (Figure S3D).

Collectively, our results show that PCID2 is present at the HIV-1 LTR during viral latency, is removed upon transcriptional activation, and acts as a repressor of HIV-1 gene expression during HIV-1 latency.

PCID2 promotes HIV-1 latency by enforcing blocks at transcription initiation

Of the functions currently described for PCID2, two stand out as particularly relevant for its role in HIV latency: first is its function in eukaryotic RNA metabolism in context of the TREX2 complex,[19,20,22] and second its distinct role in the transcriptional regulation of lymphoid lineage commitment genes.[28] Because PCID2 differentially binds the latent LTR, we first aimed to determine whether the presence of PCID2 at the HIV-1 promoter mediates its role in repressing HIV-1 gene expression at a transcriptional level. To test this, we performed a modified version of the transcriptional profiling qPCR assay described by Telwatte et al.[10] in control and PCID2-knockdown cells using primers specific for amplification of three distinct regions in the LTR: Readthrough region before the TSS (result of readthrough transcription from the gene where the proviral genome is integrated), TAR RNA region (region after TSS where RNA Pol II pauses), and U5 UTR region (region indicative of elongated transcripts) (Figures 2A and S3E). We calculated the ratio of the relative abundance of different HIV-1 RNA amplicons to assess the relative release of blocks at transcription initiation (TAR/Readthrough) and transcription elongation (UTR/TAR). Upon PCID2 knockdown, we observed a significant increase in the TAR/Readthrough ratio but not the UTR/TAR ratio (transcriptional elongation) compared with shControl lines (Figures 2B and S3E), indicating a release of a block at transcription initiation, but not transcription elongation, and hence suggesting a role for PCID2 in blocking HIV-1 LTR transcription initiation.

We next examined the role of PCID2 as a repressor of HIV-1 transcription during latency by probing the enrichment or loss of, respectively, activating and repressing histone marks at the HIV-1 promoter in presence or absence of PCID2. We performed ChIP-qPCR for total histone 3, histone 3 acetyl (H3Ac), and H3K36 trimethyl (H3K36me3) as a marker for active transcription and H3K27 trimethyl (H3K27me3) as a marker of transcriptional repression and analyzed their presence at the LTR in J-Lat cells expressing or depleted of PCID2. Our data show that, when PCID2 is depleted, there is an enrichment of H3Ac and H3K36me3 at the HIV-1 promoter (Figures 2C–2E) indicating a release in transcriptional blocks and consistent with an active transcriptional state.[30,31] Concomitant with this, we also observed a decrease in the repressive H3K27 trimethyl histone mark at the HIV-1 promoter in PCID2 knockdown cells (Figure 2F).[31]

To assess the extent to which PCID2 regulates general transcription initiation, we depleted PCID2 via shRNA in J-Lat 11.1 cells and performed 3′ end mRNA sequencing (Figure S4). We find that PCID2 depletion in J-Lat 11.1 cells leads to a modest change in general gene expression with <40 genes up- or downregulated, thus arguing against a general role for PCID2 in blocking transcription initiation.

Altogether, our data show that reduced PCID2 levels lead to a re-organization of the HIV-1 LTR chromatin landscape consistent with increased transcriptional activation, confirming that PCID2 maintains HIV-1 latency by acting on blocks at transcription initiation.
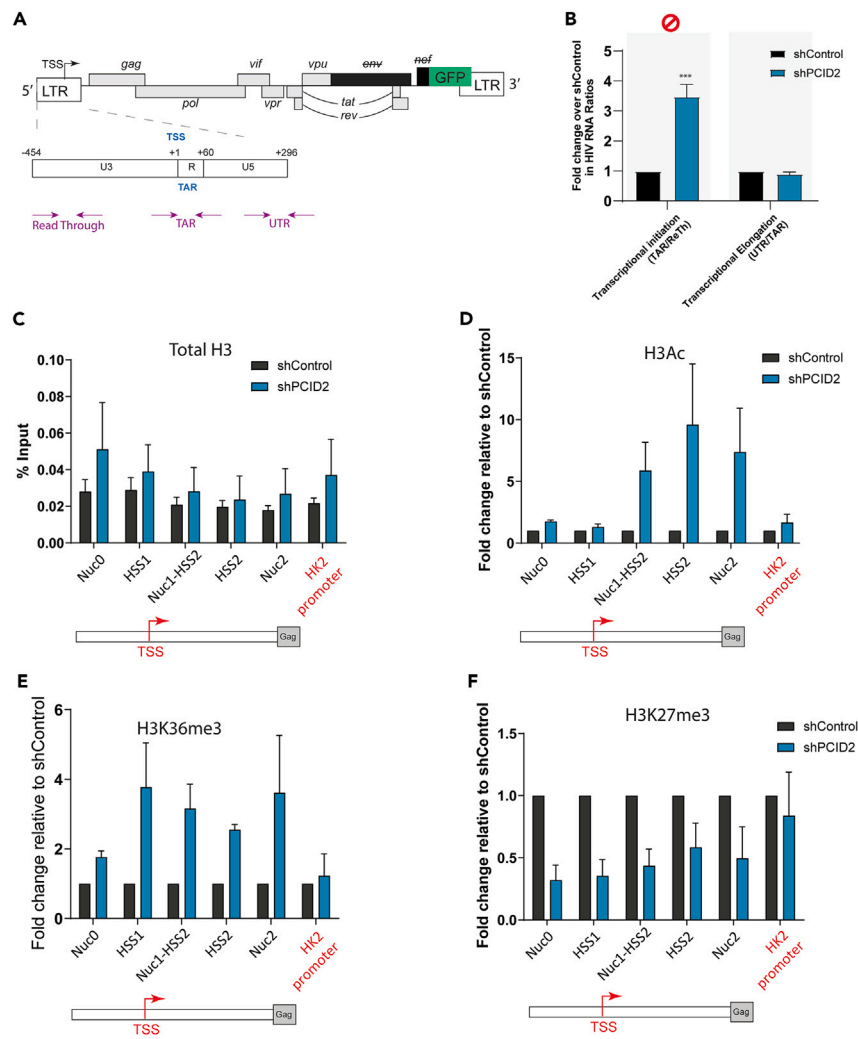
Figure 2. PCID2 is a repressor of HIV-1 transcription initiation during viral latency

(A) Schematic figure of the HIV-1 genome in J-Lat 11.1. Arrows represent primer location and amplicons across the HIV-1 genome used in the transcriptional profiling assay.
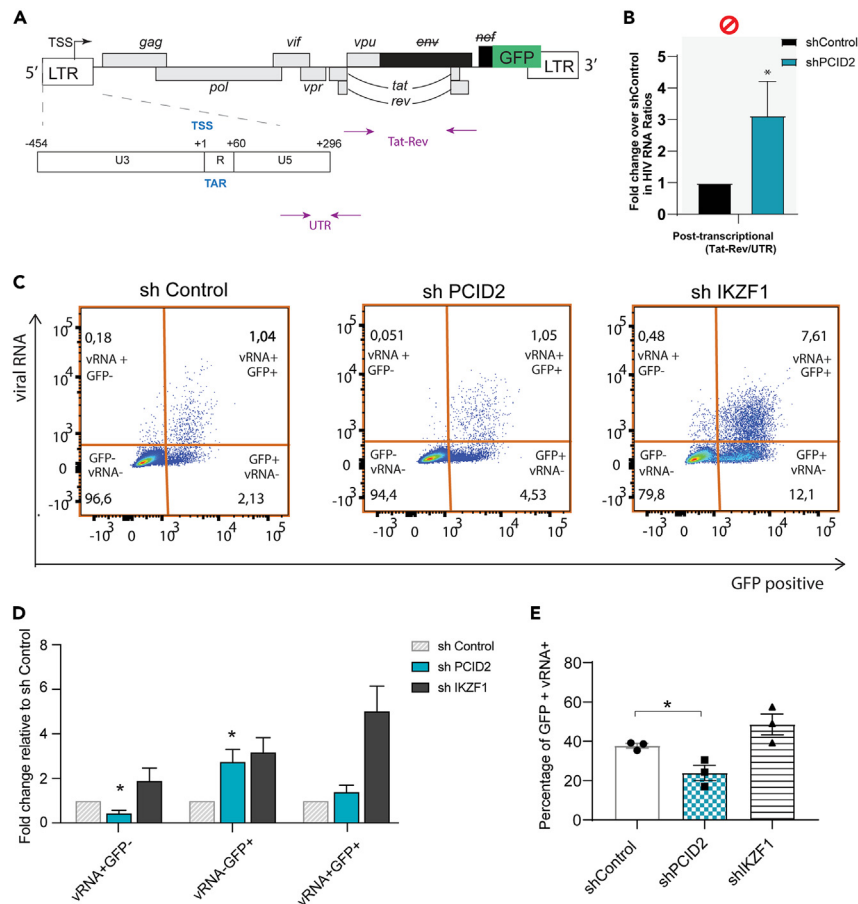
(B) Transcriptional profiling assay of shControl and shPCID2 J-Lat 11.1. Gene expression blocks at transcriptional initiation, elongation, and post-transcriptional steps are assessed by calculating the ratio of the relative abundance of HIV-1 RNA species as shown in the figure. Data are presented as fold change in HIV-1 RNA ratios in shPCID2 relative to shControl. Bars represent mean of 5 independent shRNA-mediated knockdown experiments and error lines represent SEM. Statistical significance was determined by t test; $*p < 0.05$, $***p < 0.001$.

(C–F) ChIP-qPCR analysis of total histone 3 (C), acetylated histone 3 (D), Histone 3 K36-trimethyl (E), and Histone 3 K27-trimethyl (F) enrichment at the HIV-1 promoter in control (shControl) and PCID2-knockdown (shPCID2) J-Lat 11.1. Total H3 at the HIV-1 LTR is represented as % input. H3Ac, H3K36me3, and H3K27me3 enrichment at the HIV-1 LTR is represented as relative enrichment normalized to Total H3 as shown in A and relative to control. Bars and error lines represent, respectively, mean and SEM of four ChIP-qPCR experiments in four independent chromatin preparations and shRNA-mediated knockdown experiments except for H3K36me3 mark where ChIP-qPCR was performed three times.

PCID2 represses HIV-1 gene expression at post-transcriptional steps of gene regulation and interacts with proteins involved in multiple steps of RNA processing

Once transcription is initiated, nascent RNA is co-transcriptionally packed into messenger ribonucleoparticles (mRNPs) for its efficient trafficking and export. The TREX-2 complex then facilitates transport of export-competent mRNPs to the nuclear pore complex.[20–22] Surveillance, trafficking and export of HIV-1 viral RNA containing mRNPs are tightly regulated during viral latency.[4] It has recently been shown that the main blocks to HIV-1 gene expression in cells obtained from PLWH occur at post-transcriptional steps.[11,12] Hence, in order to assess the possible contribution of PCID2 in promoting HIV-1 latency at a post-transcriptional level, we analyzed the release of blocks at this specific step in PCID2 depleted J-Lat 11.1 cells with a transcriptional profiling assay as described in Figure 3A. For this, we calculated the relative change in the ratios of Tat-Rev RNA (completely spliced viral RNA) over U5 UTR region. Upon PCID2 knockdown, we observe a significant increase in Tat-Rev/UTR RNA ratios (Figures 3B and S3E), indicative of a release in blocks at post-transcriptional steps of gene expression.

Figure 3. PCID2 blocks post-transcriptional steps of HIV-1 latency

(A) Schematic figure of the HIV-1 genome in J-Lat 11.1. Arrows represent primer location and amplicons across the HIV-1 genome used in the transcriptional profiling assay.

(B) Transcriptional profiling assay of shControl and shPCID2 J-Lat 11.1. Gene expression blocks at post-transcriptional steps are assessed by calculating the ratio of the relative abundance of HIV-1 RNA species as shown figure and A. Data are presented as fold change in HIV-1 RNA ratios in shPCID2 relative to shControl. Bars represent mean of 5 independent shRNA-mediated knockdown experiments and error lines represent SEM. Statistical significance was determined by t test; *p < 0.01.

(C) Representative FISH-Flow plots of shControl, shPCID2, or shIKZF1 knockdown J-Lat 11.1 cells.

(D) Bar graph showing fold change in vRNA+GFP-, vRNA-GFP+, and vRNA+GFP+ percentages in shControl, shPCID2, and shIKZF1 cells. Bars and error lines represent mean and SEM (n = 3). Statistical significance was determined by t test, *p < 0.05.

(E) Percentage of GFP+ cells that express viral RNA (vRNA+) in shControl, shPCID2, and shIKZF1 as analyzed by FISH-Flow. Bars and error lines represent mean and SEM respectively. Statistical significance was calculated by t test; *p < 0.05.

To better distinguish between the role of PCID2 in transcription and post-transcriptional steps of HIV-1 gene expression regulation, we performed FISH-Flow (fluorescence *in situ* hybridization-flow cytometry), a technology that discriminates between and quantifies transcriptional and post-transcriptional proviral activity at a single cell level, in shControl and shPCID2 J-Lat 11.1 cells (Figures 3C–3E and S5). FISH-Flow uses probes specific for the GagPol region to measure unspliced HIV-1 viral RNA (US vRNA) to characterize LTR transcriptional activity by flow cytometry. Simultaneously, GFP, which replaces the *nef* gene in J-Lat cells, is a translation product of completely spliced HIV-1 viral RNA (CS vRNA) and can be measured as a marker of early viral reactivation and viral RNA splicing. In latent cells, and consistent with literature, there is a basal level of residual transcription and basal expression of GFP[32] (Figures 3C and 3D), with about 37.7% (mean) of GFP+ cells also expressing vRNA (Figure 3E). To demonstrate the effect of a release in a transcriptional block specifically, we performed FISH-Flow in J-Lat 11.1 cells shRNA depleted of Ikaros zinc finger 1 (IKZF1), a recently identified HIV-1 transcriptional repressor.[16] Indeed, we observed strong viral reactivation in IKZF1-knockdown cells as shown by a marked increase in vRNA and GFP producing cells (Figure 3D), of which 48.6% (mean) also express vRNA (Figure 3E). Interestingly, when PCID2 is depleted, we observed a significant decrease in vRNA+ GFP- cells and an increase in cells only expressing GFP (Figure 3D). In addition, the percentage of GFP+ cells that also express vRNA was significantly reduced (23.8%, mean) (Figure 3E). This indicates that, while depletion of PCID2 in J-Lat cells leads to early viral reactivation, there is also an accompanying misregulation in splicing, resulting in a decrease in intron-containing viral RNA in both GFP- and GFP+ cell populations, but an overall increase in GFP+ cells, indicative of increased abundance of completely spliced HIV-1 vRNA species.
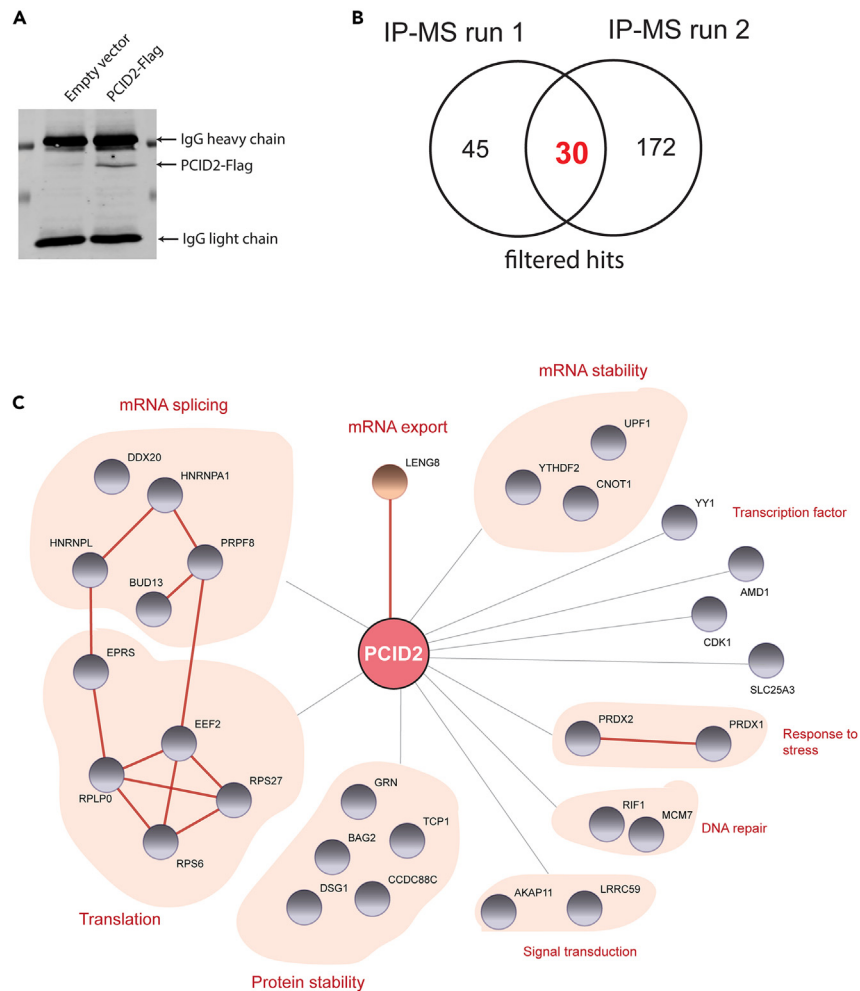
Figure 4. PCID2 interacts with proteins involved in multiple stages of RNA processing

(A) IP-western blot of PCID2-Flag in control and stably PCID2-Flag expressing J-Lat 11.1 samples for subsequent mass spectrometry analysis.

(B) Venn diagram of the total individual and common number of identified hits by mass spectrometry in PCID2-Flag over control samples from two independent IP-MS runs.

(C) Network of PCID2 interacting hits identified by mass spectrometry. Hits are classified by cellular function. Gray lines represent PCID2 interacting proteins as analyzed by IP-MS in this study. Red lines represent protein-protein interactions predicted by STRING analysis on experimental evidence only.

These results point to a potential role for PCID2 in regulation of HIV-1 latency at distinct post-transcriptional steps of gene expression, such as splicing and export. Although PCID2 and the TREX2 complex have mostly been studied for their role in linking transcription with transport of export-competent mRNPs, it is currently unknown whether other steps of RNA processing are influenced by PCID2. Actually, members of the THO/TREX complex, that interact with TREX-2 and also mediate transcription with mRNP export,[33] are known to interact with and recruit splicing factors to mRNPs.[34,35] Hence, in order to identify possible molecular pathways by which PCID2 regulates HIV-1 gene expression at a post-transcriptional level in an unbiased manner, we performed immunoprecipitation of exogenously expressed PCID2-Flag in J-Lats 11.1 cells coupled with semi-quantitative mass spectrometry analysis (IP-MS). As a control, we used J-Lat cells transfected with an empty vector. We performed the IP-MS pipeline two times with independent PCID2-Flag immunoprecipitations (Figures 4A and S6A; Table S1) followed by application of stringent unbiased filtering to decrease the possibility of enrichment of false interactors (Figure S6B). The raw hits found in both conditions (Control and PCID2-Flag) were filtered for common contaminants and presence of unique peptides >1. Only the hits present in the PCID2-Flag IP with a minimum 1.2-fold threshold were taken into consideration for further analysis. After filtering, 75 and 202 proteins were found in the two independent PCID2-Flag IP-MS runs and 30 proteins are common to both experiments (Figure 4B). As expected, the top hit found in both PCID2-Flag IP-MS runs was PCID2 (Table S1). Importantly, protein LENG8, recently suggested to be a mammalian ortholog for yeast factor Sac3 (also named GANP) and associated with PCID2 and other members of the TREX2 complex *in vitro*,[36] was found in both IP-MS runs to be one of the top hits uniquely present in PCID2-FLAG IP samples (Figures 4C; Table S1). In the common PCID2-Flag IP-MS interactor list (Figure 4C), PCID2 interactors included those involved in mRNA

metabolism pathways such as translation (ribosomal factors such as RPLP0 and RPS26, RPS6), mRNA splicing (spliceosome complex proteins PRPF8 and BUD13), and mRNA surveillance and stability (UPF1). We also found PCID2 interacting proteins involved in DNA repair, protein stability, cell response to stress, signal transduction and transcription.

Thus, in addition to its role in transcriptional repression of HIV-1 gene expression, our data suggested a role for PCID2 in maintenance of HIV-1 latency at post-transcriptional steps of HIV-1 gene expression via interaction with proteins involved in RNA processing pathways including RNA export, splicing, surveillance and translation.

PCID2 inhibits alternative splicing during HIV-1 latency

In line with our previous data, the list of PCID2 interactors discovered in our IP-MS approach suggested that, besides RNA export, PCID2 takes part in other steps of RNA trafficking and processing such as splicing. In support of this, we found, among the identified PCID2 interactors, several proteins involved in RNA splicing, including members of the spliceosome complex (PRPF8 and DDX20) and HNRNP family members, known to act as negative regulators of HIV-1 vRNA splicing (Figure 4C).[37–39] To determine if PCID2 indeed plays a role in HIV-1 alternative splicing, we performed a PCR-based splicing assay as previously described[40] to measure the relative abundance of the three main HIV-1 viral RNA splicing variants: full-length US, single spliced (SS), and CS HIV-1 vRNA species. HIV-1 splicing is tightly regulated during latency and US HIV-1 RNA species are more abundant than spliced HIV-1 vRNA species (reviewed by Pasternak A.O. and Berkhout B.[41]). Upon PCID2 knockdown, we observed a relative decrease in unspliced vRNA accompanied by a significant increase in completely spliced HIV-1 vRNA (Figure 5A) and a shift in the abundances of unspliced/spliced viral RNA (Figure 5B), in agreement with the data we observed by FISH-Flow (Figures 3C–3E). Overexpression of PCID2 in Jurkat lines transfected with an *env* deficient full-length HIV-1 plasmid shows the opposite effect, with a significant increase in US vRNA and decrease in CS vRNA species (Figure S7). Importantly, the effect observed in abundance of viral RNA species upon PCID2 knockdown is reflected at the protein level, as shown by an increase in percentage of GFP expressing cells and decrease in percentage of intracellular p24 expressing cells (Figure 5C), encoded by, respectively, US and CS vRNA.

Arguably, the effect on abundance of the different vRNA splicing variants upon PCID2 depletion could also result from PCID2-mediated stabilization of US HIV-1 vRNA. We found, in fact, several proteins that interact with PCID2 involved in RNA surveillance and stability, such as UPF1, previously identified to regulate HIV-1 latency at a post-transcriptional level.[32] Hence, we next aimed to assess whether the PCID2-mediated change in the relative abundance of splicing variants is explained by a role of PCID2 in HIV-1 RNA stability and decay during latency. To test this, we inhibited general transcription in PCID2-Flag overexpressing and control J-Lat 11.1 lines (Figures 5D–5F) by treatment with the DNA intercalator actinomycin D and measured the relative decay of US HIV-1 vRNA over time. We observed a comparable steady decay both in PCID2-Flag overexpressing and control J-Lat 11.1 lines (Figure 5E). As control for mRNA decay, we also measured decay of B2-microglobulin mRNA and observed, similar to that for US HIV-1 vRNA, a steady decay both in PCID2-Flag overexpressing and control J-Lat 11.1 lines (Figure 5F). These data show that PCID2 has no apparent role in US HIV-1 vRNA stability and that the effect we observe on the abundance and ratios of HIV-1 vRNA splicing variants upon PCID2 knockdown or overexpression results from its role as an inhibitor of viral RNA splicing.

To further confirm the role of PCID2 on HIV-1 latency and specifically misregulation of HIV-1 vRNA splicing, we set out to investigate the effect of PCID2 depletion in a more physiologically relevant system for HIV-1 infection. For this, we performed an shRNA-mediated knockdown of PCID2 in primary *ex vivo* CD4[+] T cells obtained from 3 donors living with HIV-1 under suppressive antiretroviral therapy (Figures 5G–5I) and assessed the effect of PCID2 depletion on the ratios of CS and US vRNA species. Consistent with our results in cell lines (Figures 3B, 5A and 5B; Table 2), we show that PCID2 knockdown results in a shift in abundances of CS and US vRNA, as observed by an increase in the ratio of CS to US HIV-1 vRNA species, indicative of a release in post-transcriptional blocks (Figures 5G–5I).

Taken together, our results indicate that, besides transcription, PCID2 promotes HIV-1 latency at post-transcriptional steps of HIV-1 gene expression regulation by inhibiting alternative splicing.

TREX2 complex subunits MCM3AP and DSS1 block transcription and misregulate splicing resulting in repression of HIV-1 gene expression

While transcriptional control by PCID2 has been studied as an independent subunit, its involvement in regulation of RNA processing has been almost exclusively linked to its activity as part of the TREX2 complex. Mammalian TREX2 complex is formed by the PCI-containing subunits MCM3AP, DSS1, and PCID2 that form the scaffold and RNA binding platform of the TREX2 complex and subunits ENY2 and CETN2/3 that form the nuclear pore binding dock. TREX2 is hence recruited to nascent RNA and facilitates export of mRNPs via the NXF1 pathway.[19–22,42] We first addressed to what extent viral RNA export is regulated by PCID2 and examined the distribution of intron-containing and completely spliced HIV-1 vRNA species in the nuclear and cytoplasmic compartments upon PCID2 depletion. Indeed, knockdown of PCID2 causes partial nuclear retention of completely spliced HIV-1 vRNA species, as observed by a shift in the distribution between the nuclear and cytoplasmic fractions (Figures 6A–6C). This is in line with the role of the PCID2/TREX2 complex in mRNA export, as only completely spliced HIV-1 vRNA variants, because of their small size, are exported via the NXF1 pathway, while intron-containing vRNAs need the viral protein Rev (encoded by CS HIV-1 vRNA) for efficient export.[4] Accordingly, upon PCID2 knockdown the relative distribution of US HIV-1 vRNA between the cytosolic and nuclear fraction remains stable (Figure 6C), as US HIV-1 vRNA is retained in the nucleus during viral latency (Figure 6C).

Still, the question remains whether the observed suppression of HIV-1 proviral transcription and/or viral RNA alternative splicing by PCID2 occur independently or within the platform of the TREX2 complex. In fact, we previously performed an insertional mutagenesis genetic screen in latent cells and identified the TREX2 complex member MCM3AP as a host factor involved in maintenance of HIV-1 latency.[15] Hence, to
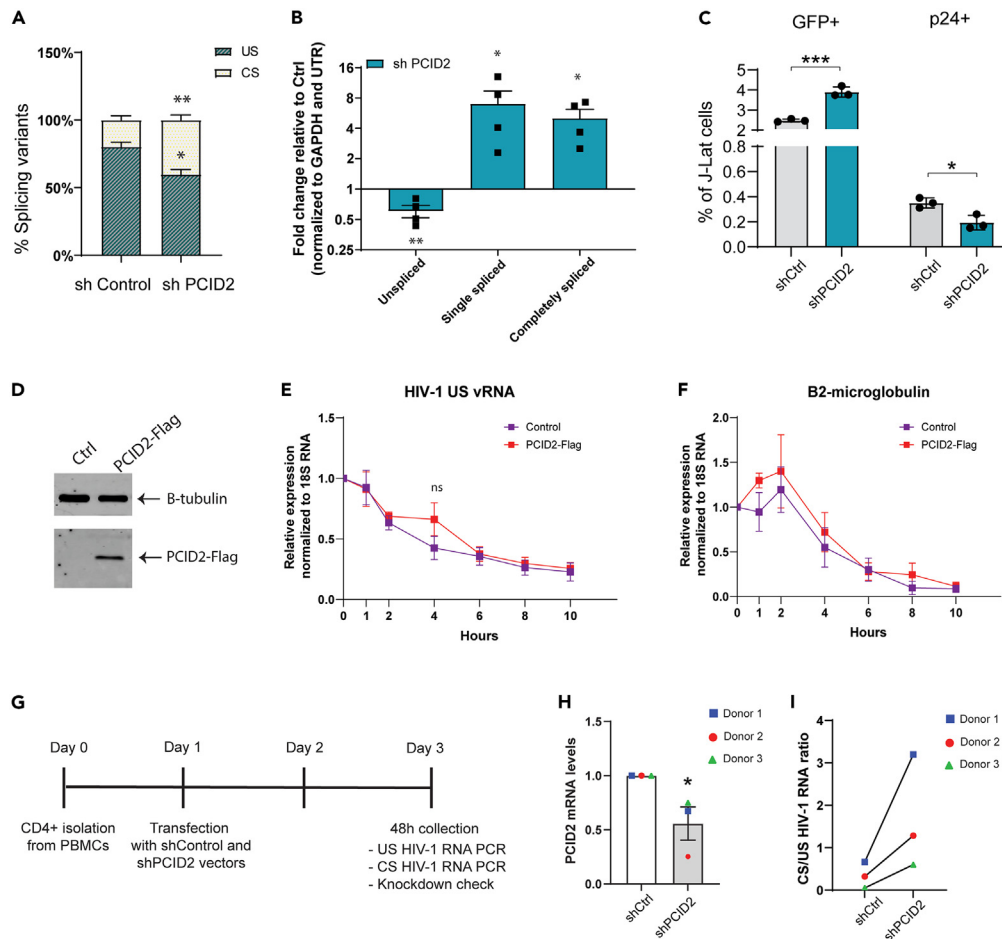
Figure 5. PCID2 blocks HIV-1 RNA alternative splicing during latency in cell line models and cells obtained from people living with HIV

(A) Relative percentage of intron-containing unspliced (US) and completely spliced (CS) HIV-1 RNA in shControl and shPCID2 knockdown J-Lat 11.1. Bars and error lines represent mean and SEM respectively (n = 4). Statistical significance was determined by t test, *p < 0.05, **p < 0.01.

(B) Fold change in HIV-1 RNA splicing variants upon PCID2 knockdown relative to shControl. Values were normalized to GAPDH and UTR HIV-1 transcript. Bars represent mean and error lines represent SEM. Statistical significance was determined by t test; *p < 0.05, **p < 0.01.

(C) Percentage of sh Control and shPCID2 J-Lat cells expressing GFP and intracellular p24 as analyzed by flow cytometry. Bars represent mean of three independent infections and error lines represent SEM. Statistical significance was determined by t test; *p < 0.05, ***p < 0.001.

(D) Western blot of PCID2-Flag in control or PCID2-Flag overexpressing J-Lat 11.1. Beta tubulin was used as loading control.

(E) HIV-1 unspliced (US) RNA decay dynamics in control and PCID2-Flag overexpressing J-Lat 11.1. Cells were treated with Actinomycin D (10 μg/mL) and RNA was collected at 0, 1, 2, 4, 6, 8 and 10 h after treatment. Gene expression of unspliced RNA was normalized to 18S RNA and values are calculated relative to time point 0 h. Symbols represent the mean of three independent experiments and error bars represent SEM. Statistical significance was determined by t test; ns = not significant.

(F) B2-microglobulin RNA decay dynamics in control and PCID2-Flag overexpressing J-Lat 11.1 as described in B. Gene expression of B2-microglobulin RNA was normalized to 18S RNA and values are calculated relative to time point 0 h. Symbols represent the mean of three independent experiments and error bars represent SEM.

(G) Timeline schematic of shRNA mediated PCID2 knockdown in CD4+ T cells obtained from people living with HIV-1.

(H) Gene expression analysis of PCID2 mRNA in shControl and shPCID2 knockdown CD4+ T cells from three PLWH donors 48 h after transfection. Bars and error lines represent mean and SEM respectively.

(I) Release in post-transcriptional blocks to latency calculated as a ratio of relative abundance in HIV-1 RNA splicing variants CS HIV-1 RNA and US HIV-1 RNA normalized to housekeeping in shControl and shPCID2 knockdown CD4+ T cells from three PLWH donors.

determine the potential role of the TREX2 complex as a whole in suppressing HIV-1 gene expression, we assessed the overall contribution of the other TREX2 complex members to HIV-1 latency. We therefore depleted the individual subunits via shRNA knockdown in J-Lat 11.1 cells and interrogated latency reversal at multiple levels of HIV-1 gene expression (Figures 6D–6F and S8). We found that depletion of TREX2 nuclear pore binding subunits ENY2 and CETN3 did not lead to HIV-1 reactivation from latency in J-Lat 11.1 (Figure 6D). On the contrary, significant reactivation of HIV-1 latency, as assessed by production of GFP both at the protein and RNA level, was observed consequent to depletion of endogenous DSS1 and MCM3AP that, together with PCID2, form the TREX2 RNA binding dock (Figure 6D).

Table 2. List of primers used for US and CS vRNA nested qPCR

Nested US forward	TCAGCCCAGAAGTAATACCCATGT
Nested US reverse 1	TGCTATGTCAGTTCCCCTTGGTTCTCT
Nested US reverse 2	CACTGTGTTTAGCATGGTGTTT
US probe	[6FAM]ATTATCAGAAGGAGCCACCCCACAAGA[BHQ1]
Nested CS forward	CTTAGGCATCTCCTATGGCAGGAA
Nested CS forward 2	ACAGTCAGACTCATCAAGTTTCTCTATCAAAGCA
Nested CS reverse	GGATCTGTCTCTGTCTCTCTCTCCACC
CS probe	[6FAM]TTCCTTCGGGCCTGTCGGGTCCC[TAMRA]

When looking at the relative release of specific blocks to HIV-1 gene expression by a PCR-based transcriptional profiling assay, we observe that depletion of endogenous DSS1 and MCM3AP, but not ENY2 or CETN3, led to release of blocks in transcription initiation and alternative splicing, similar to what we observed upon PCID2 depletion (Figure 6E). The effect of DSS1 and MCM3AP on suppressing splicing is also reflected by distinct changes in the abundances of viral RNA splicing variants with significant increase in spliced viral RNA upon depletion, while, consistent with our previous data, depletion of ENY2 and CETN3 does not lead to changes in the abundance of viral RNA splicing variants in J-Lat 11.1 (Figure 6F).

Our results are consistent with the notion that PCID2 function in NXF1-mediated mRNA export and in regulation of HIV-1 latency at transcriptional and post-transcriptional levels of HIV-1 gene expression occur in context of and within functional PCID2-DSS1-MCM3AP RNA-binding docking subunit of the TREX2 complex.

DISCUSSION

In a previous study, we discovered novel LTR-bound regulators of HIV-1 transcription by means of locus-specific chromatin immunoprecipitation coupled to mass spectrometry (Catchet-MS) in latent and active infected cells.[16] Here, we delineate the functions of PCID2, a novel factor we identified to be differentially bound to the latent HIV-1 promoter, as a multifactorial protein that blocks distinct transcriptional and post-transcriptional steps of HIV-1 gene expression to promote viral latency.

In latent cells, endogenous PCID2 is bound to the HIV-1 LTR and removed upon transcriptional activation triggered by PMA treatment, consistent with our findings by the Catchet MS pipeline.[16] We demonstrate that PCID2 acts as a repressor of HIV-1 transcription initiation during latency and prevents latency reactivation (Figure 1). In literature, PCID2 has been shown to regulate gene expression at multiple levels. The main role of PCID2 described in literature concerns co-transcriptional mRNA processing, and is linked to its function as part of the transcription and export complex 2 that couples transcription with transport of mRNPs containing nascent RNA to the nuclear pore complex for its subsequent export to the cytoplasm.[19–23] Independent of its role as part of the TREX2 complex, PCID2 has been shown to have a role in genome and protein stability,[24,25] regulation of embryonic stem cell development, and transcriptional regulation of specific genes in hematopoietic stem cells.[27,28] At a transcriptional level, Ye et al. showed that PCID2 interacts with ZNHIT1, a member of the SRCAP complex, and silences transcription of genes involved in lymphoid fate commitment.[28] This study also demonstrated that absence of PCID2 increases chromatin accessibility at promoters of lymphoid commitment genes. In agreement, we find that PCID2 promotes latency by enforcing blocks at transcription initiation specifically (Figure 2). Downregulation of endogenous PCID2 re-starts LTR-driven transcription in latent cells and results in changes in the epigenetic state of the promoter as shown by enrichment and decrease in enrichment of, respectively, activating H3Ac and H3K36me3 and repressive H3K27me3 histone marks at the LTR.

The yeast homolog TREX2 complex has been characterized in literature to affect RNA-Pol II mediated transcriptional elongation and absence of several members, including the PCID2 ortholog Thp1, results in significant decrease in elongation of nascent RNA.[43,44] Transcription elongation is tightly regulated during HIV-1 latency where, in absence of viral protein Tat, RNA-Pol II pauses a few nucleotides after the transcription start site as it encounters the secondary RNA element TAR, and elongation of nascent RNA transcripts is prohibited.[5,45] Interestingly, in our system, PCID2 depletion does not lead to an increase in elongated viral RNA transcripts. The functions of the mammalian PCID2, however, may very well differ from its yeast ortholog, at least in context of transcriptional elongation at the HIV-1 promoter. Future studies are needed to address the overall role of PCID2 independently or as part of the TREX2 complex on RNA-Pol II-mediated transcription elongation in metazoans.

In addition to silencing HIV-1 transcription initiation, our findings demonstrate that PCID2 enforces blocks at post-transcriptional steps of latency (Figures 3, 4, 5, and 6). Co- and post-transcriptional regulation of HIV-1 gene expression are crucial steps during the establishment and maintenance of viral latency and include blocks at 5'capping, alternative splicing, nucleocytoplasmic export, vRNA surveillance, translation, and assembly.[4] We distinctively show that knockdown of PCID2 misregulates the abundance of intron-containing and CS vRNA species, resulting in a marked decrease in levels of US vRNA and accumulation of spliced vRNAs in latently infected cell lines and, crucially, in primary cells obtained from people living with HIV-1. Our data also demonstrates that this effect is not a consequence of a decrease in stability of US vRNA upon PCID2 depletion but rather is the result of a release in blocks at alternative splicing. This is further supported by our finding that PCID2 interacts with several members of the spliceosome complex and other splicing modulators such as PRPF8, HNRNPA1, and BUD13. A similar over-splicing effect has been reported in literature as a result of altering the abundance or functionality of SR protein kinases,[46,47] and absence or mutation of *cis*-acting elements that prevent binding of the spliceosome complex (reviewed by Emery A. and Swanstrom R.[48]).
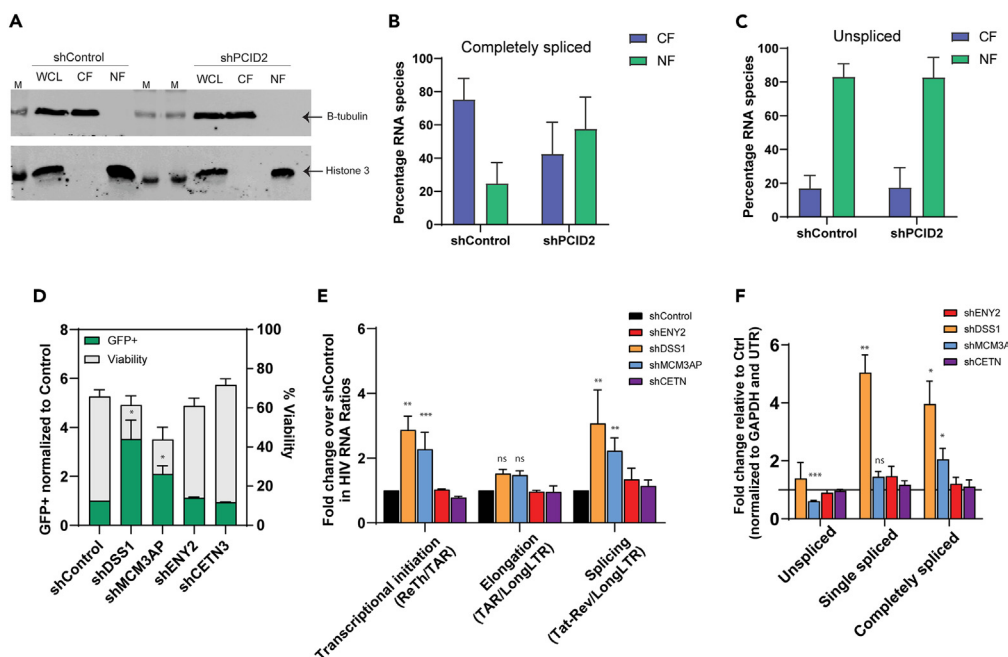
Figure 6. Effect of PCID2 and other TREX2 complex members on transcriptional and post-transcriptional steps of HIV-1 gene expression during viral latency

(A) Western blot of whole cell lysate (WCL) and cell fractions corresponding to cytoplasm fraction (CF) and nuclear fraction (NF) in shControl or shPCID2 knockdown J-Lat 11.1 to probe for B-tubulin and histone 3. M refers to protein marker.

(B and C) Percentage of RNA species unspliced HIV-1 vRNA (B) and completely spliced HIV-1 vRNA (C) in the cytosolic fraction (CF) and nuclear fraction (NF). Bars and error lines represent mean and SEM (n = 3).

(D) Fold increase in the % of GFP (left y axes) and viability (right y axes) in shControl or shDSS1, shMCM3AP, shENY2, and shCETN3 J-Lat 11.1 cells as measured by flow cytometry. Bars represent mean of three independent shRNA-mediated knockdown experiments and error lines represent SEM. Statistical significance was determined by t test; *p < 0.05. Raw values for percentage of GFP are available in Table S2.

(E) Transcriptional profiling assay of shControl and shDSS1, shMCM3AP, shENY2, and shCETN3 J-Lat 11.1 cells. Gene expression blocks at transcriptional initiation, elongation, and post-transcriptional steps are assessed by calculating the ratio of the relative abundance of HIV-1 RNA species as shown in the figure. Data are presented as fold change in HIV-1 RNA ratios in knocked down cells relative to shControl. Bars represent mean of three to five independent shRNA-mediated knockdown experiments and error lines represent SEM. Statistical significance was determined by ANOVA test; *p < 0.05, **p < 0.01, ***p < 0.001.

(F) Fold change in HIV-1 RNA splicing variants upon DSS1, MCM3AP, ENY2, and CETN3 knockdown relative to shControl. Values were normalized to GAPDH and UTR HIV-1 transcript. Bars represent mean and error lines represent SEM. Statistical significance was determined by t test; *p < 0.05, **p < 0.01, ***p < 0.001.

Importantly, because PCID2 depletion enables a partial release of HIV-1 from latency, it is unlikely that PCID2 absence would lead to production of full HIV-1 viral particles for two reasons: (1) Levels of unspliced vRNA and Gag protein are significantly downregulated, hence negatively impacting the production of viral particles; (2) PCID2 depletion results in misregulation of vRNA species at splicing and export, likely impacting the formation of infectious viral particles.

The role of PCID2 in RNA metabolism has been primarily linked to its function in nucleocytoplasmic RNA export and other putative functions of PCID2 alone or in context of the TREX2 complex in prior or further steps of RNA processing have been unexplored. Notably, another TREX2 complex member, protein Sem1 (yeast homolog for DSS1) has been shown previously to influence mRNA alternative splicing in yeast.[49] This raised the question of whether PCID2 exerts its function in transcriptional and/or post-transcriptional steps of HIV-1 gene expression regulation independently or as part of the TREX2 complex. Interestingly, when we examined the relative contribution of the other TREX2 complex members to HIV-1 latency individually, we observed a prominent role for TREX2 subunits MCM3AP and DSS1 in regulating viral latency. This agrees with our previous study in which, by performing an insertional mutagenesis genetic screen in latent cells, we identified MCM3AP as a putative host factor that enforces HIV-1 latency.[15] We further show that, mirroring the effects observed with PCID2, suppression of HIV-1 gene expression by MCM3AP and DSS1 results from hindered transcription initiation and inhibition of viral RNA alternative splicing. In yeast, Sac3 (yeast homolog for MCM3AP) and Sem1 (DSS1) are also PCI-containing proteins that bind Thp1 (PCID2), and respectively, form the scaffold and stabilize Thp1 (PCID2) in the Sac3-Thp1-Sem1 complex.[50,51] The sub-complex DSS1-MCM3AP-PCID2 simultaneously forms part of the RNA binding platform of the larger TREX2 complex. Our findings imply a specific role for the PCID2-DSS1-MCM3AP sub-complex in interacting with viral RNA and modulating viral latency at both transcriptional and post-transcriptional steps of gene expression.[20,22,50] In our experimental system, we did not observe a clear effect on HIV-1 latency reactivation upon depletion of TREX2 subunits ENY2 and CETN3 that bind the nuclear pore complex for efficient nucleocytoplasmic export of mRNP complexes.[20–22] Thus, these observations are consistent
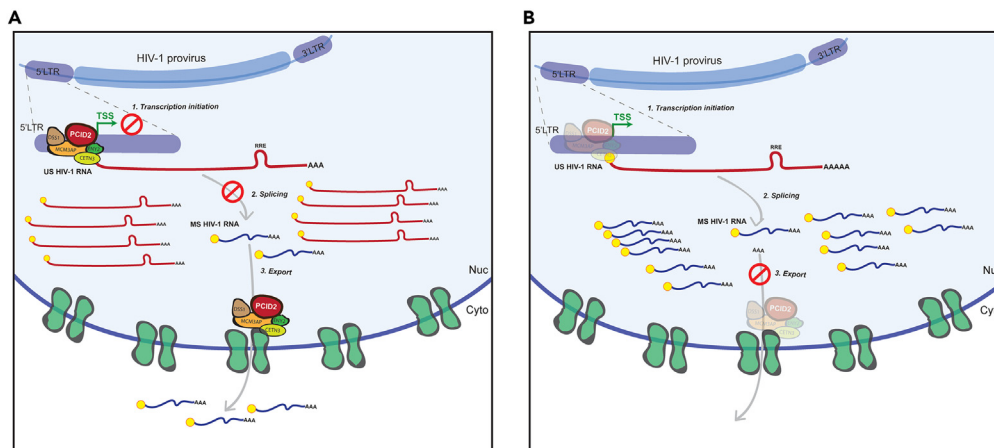
Figure 7. Proposed model for the role of PCID2/TREX2 in HIV-1 gene expression during viral latency

(A) In latent cells, PCID2/TREX2 is present at the HIV-1 LTR transcription start site (TSS) and blocks HIV-1 gene expression and promotes viral latency at two distinct levels: transcription initiation and RNA alternative splicing.

(B) Upon PCID2-DSS1-MCM3AP downregulation, LTR-driven transcription is re-started. In addition, absence of PCID2-DSS1-MCM3AP leads to increased splicing of intron-containing unspliced (US) viral RNA and accumulation of completely spliced (CS) HIV-1 RNA. Because PCID2 facilitates export of mRNPs as part of the TREX2 complex via the NXF1 pathway, its downregulation leads to nuclear retention of CS HIV-1 vRNA.

with the hypothesis that repression of HIV-1 gene expression by PCID2 is an RNA-mediated process that likely occurs in context of the PCID2-DSS1-MCM3AP sub-complex of the TREX2 complex and is dependent on nucleic acid binding.

Lastly, a well-described role for PCID2 is its role as part of the TREX2 complex in nucleocytoplasmic export of spliced mRNAs via the NXF1-NXT1 pathway.[20,22,23] During HIV-1 infection, only CS vRNAs, because of their 2kb size, are exported via NXF1-NXT1. Larger, intron-containing vRNA species need viral protein Rev and other co-factors for their efficient export to the cytoplasm.[4,52] In yeast, absence or loss of function of Thp1 (PCID2 ortholog) leads to overall inhibition of spliced mRNA export.[23] Consistent with this, we show that downregulation of PCID2 causes nuclear retention of CS vRNA but not US vRNA species. Of note, we still observe an increase in GFP, encoded by the CS vRNA species, when PCID2 is depleted from J-Lat cells. This could be explained by the fact that our shRNA-mediated knockdown is not 100% efficient and some residual PCID2 still remains, and/or that other compensatory pathways are in place that sustain export of CS vRNA species to the cytoplasm.

In conclusion, our data support a model in which PCID2-containing complexes maintain HIV-1 latency at two essential steps of HIV-1 gene expression regulation (Figure 7). PCID2 acts as a repressor of HIV-1 LTR-driven transcription initiation and blocks viral RNA alternative splicing. Hence, upon PCID2 depletion, transcription initiation from the HIV-1 LTR is re-started and blocks at splicing are released, resulting in higher levels of spliced vRNAs and overall viral reactivation from latency. Finally, our study furthers current knowledge on the molecular pathways involved in HIV-1 latency regulation, emphasizing the role of novel multifactorial host cell factors in governing viral reactivation across various steps, including post-transcriptionally.

Limitations of the study

Our data demonstrates a prominent role for PCID2 in blocking transcription initiation and alternative splicing during HIV-1 latency. Importantly, we show that PCID2 misregulates splicing in primary CD4$^+$ T cells obtained from PWH. However, assessing the role of PCID2 in blocking HIV-1 transcription initiation in cells from PWH is currently met with technical limitations: assessment of epigenetic changes at the LTR locus in a PCID2 knockdown or overexpression system would require very large scale and potentially single-cell technologies since the percentage of HIV-1 infected cells in blood from PWH is remarkably low.

In addition, the exact molecular mechanisms by which PCID2-mediates repression of HIV-1 transcription and alternative splicing remain to be elucidated, whether this occurs by PCID2 alone, in context of the TREX2 complex or via association with other factors. Future studies are therefore needed to explore the molecular mechanism and overall role of PCID2 and the TREX2 complex in HIV-1 gene expression regulation in relevant model systems.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.109152.

AUTHOR CONTRIBUTIONS

Conceptualization: R.C., E.N., R.J.P., S.R., and T.M.; experimental design, formal analysis, methodology, and validation: R.C., E.N., J.R., J.M., C.L., S.J., A.G., S.K., T.W.K., W.D., D.D., and J.D.; writing – original draft and review and editing: R.C., R.J.P., S.R., and T.M.

DECLARATION OF INTERESTS

T.M. is an inventor on US patent application serial no. 17/259,107. International application no. PCT/US2019/041466. Entitled: Methods for reversing HIV latency using BAF complex modulating compounds. Inventors: Emily Dykhuizen (University of Purdue), Tokameh Mahmoudi (Erasmus MC), Gerald Crabtree (Stanford University).

REFERENCES

1. Chun, T.W., Davey, R.T., Jr., Ostrowski, M., Shawn Justement, J., Engel, D., Mullins, J.I., and Fauci, A.S. (2000). Relationship between pre-existing viral reservoirs and the re-emergence of plasma viremia after discontinuation of highly active anti-retroviral therapy. Nat. Med. 6, 757–761. https://doi.org/10.1038/77481.

2. Chun, T.W., Engel, D., Berrey, M.M., Shea, T., Corey, L., and Fauci, A.S. (1998). Early establishment of a pool of latently infected, resting CD4(+) T cells during primary HIV-1 infection. Proc. Natl. Acad. Sci. USA 95, 8869–8873. https://doi.org/10.1073/pnas.95.15.8869.

3. Siliciano, J.D., and Siliciano, R.F. (2022). In Vivo Dynamics of the Latent Reservoir for HIV-1: New Insights and Implications for Cure. Annu. Rev. Pathol. 17, 271–294. https://doi.org/10. 1146/annurev-pathol-050520-112001.

4. Crespo, R., Rao, S., and Mahmoudi, T. (2022). HibeRNAtion: HIV-1 RNA Metabolism and Viral Latency. Front. Cell. Infect. Microbiol. 12, 855092.

5. Ne, E., Palstra, R.J., and Mahmoudi, T. (2018). Transcription: Insights From the HIV-1 Promoter. Int. Rev. Cell Mol. Biol. 335, 191–243. https://doi.org/10.1016/bs.ircmb. 2017.07.011.

6. Verdin, E. (1991). DNase I-hypersensitive sites are associated with both long terminal repeats and with the intragenic enhancer of integrated human immunodeficiency virus type 1. J. Virol. 65, 6790–6799. https://doi.org/10.1128/JVI.65.12.6790-6799.1991.

7. Verdin, E., and Van Lint, C. (1995). Internal transcriptional regulatory elements in HIV-1 and other retroviruses. Cell. Mol. Biol. 41, 365–369.

8. Van Lint, C., Bouchat, S., and Marcello, A. (2013). HIV-1 transcription and latency: an update. Retrovirology 10, 1–38. https://doi.org/10.1186/1742-4690-10-67.

9. Ott, M., Geyer, M., and Zhou, Q. (2011). The control of HIV transcription: keeping RNA polymerase II on track. Cell Host Microbe 10, 426–435. https://doi.org/10.1016/j.chom.2011.11.002.

10. Telwatte, S., Morón-López, S., Aran, D., Kim, P., Hsieh, A., Joshi, S., Montano, M., Greene, W.C., Butte, A.J., Wong, J.K., and Yukl, S.A. (2019). Heterogeneity in HIV and cellular transcription profiles in cell line models of latent and productive infection: implications for HIV latency. Retrovirology 16, 32. https://doi.org/10.1186/s12977-019-0494-x.

11. Yukl, S.A., Kaiser, P., Kim, P., Telwatte, S., Joshi, S.K., Vu, M., Lampiris, H., and Wong, J.K. (2018). HIV latency in isolated patient CD4(+) T cells may be due to blocks in HIV transcriptional elongation, completion, and splicing. Sci. Transl. Med. 10. https://doi.org/10.1126/scitranslmed.aap9927.

12. Moron-Lopez, S., Telwatte, S., Sarabia, I., Battivelli, E., Montano, M., Macedo, A.B., Aran, D., Butte, A.J., Jones, R.B., Bosque, A., et al. (2020). Human splice factors contribute to latent HIV infection in primary cell models and blood CD4+ T cells from ART-treated individuals. PLoS Pathog. 16, e1009060. https://doi.org/10.1371/journal.ppat.1009060.

13. Rao, S., Lungu, C., Crespo, R., Steijaert, T.H., Gorska, A., Palstra, R.J., Prins, H.A.B., van Ijcken, W., Mueller, Y.M., van Kampen, J.J.A., et al. (2021). Selective cell death in HIV-1-infected cells by DDX3 inhibitors leads to depletion of the inducible reservoir. Nat. Commun. 12, 2475. https://doi.org/10.1038/s41467-021-22608-z.

14. Sarracino, A., Gharu, L., Kula, A., Pasternak, A.O., Avettand-Fenoel, V., Rouzioux, C., Bardina, M., De Wit, S., Benkirane, M., Berkhout, B., et al. (2018). Posttranscriptional Regulation of HIV-1 Gene Expression during Replication and Reactivation from Latency by Nuclear Matrix Protein MATR3. mBio 9, e02158-18. https://doi.org/10.1128/mBio.02158-18.

15. Röling, M., Mollapour Sisakht, M., Ne, E., Moulos, P., Crespo, R., Stoszko, M., De Crignis, E., Bodmer, H., Kan, T.W., Akbarzadeh, M., et al. (2021). A Two-Color Haploid Genetic Screen Identifies Novel Host Factors Involved in HIV-1 Latency. mBio 12, e0298021. https://doi.org/10.1128/mBio.02980-21.

16. Ne, E., Crespo, R., Izquierdo-Lara, R., Rao, S., Koçer, S., Górska, A., van Staveren, T., Kan, T.W., Dekkers, D., Rokx, C., et al. (2022). Catchet-MS identifies IKZF1-targeting Thalidomide analogues as novel HIV-1 latency reversal agents. Nucleic Acids Res. 50, 5577–5598. https://doi.org/10.1093/nar/gkac407.

17. Xiao, H., Wyler, E., Milek, M., Grewe, B., Kirchner, P., Ekici, A., Silva, A., Jungnickl, D., Full, F., Thomas, M., et al. (2021). CRNKL1 Is a Highly Selective Regulator of Intron-Retaining HIV-1 and Cellular mRNAs. mBio 12. https://doi.org/10.1128/mBio.02525-20.

18. Ma, J., Rong, L., Zhou, Y., Roy, B.B., Lu, J., Abrahamyan, L., Mouland, A.J., Pan, Q., and Liang, C. (2008). The requirement of the DEAD-box protein DDX24 for the packaging of human immunodeficiency virus type 1 RNA. Virology 375, 253–264. https://doi.org/10.1016/j.virol.2008.01.025.

19. Glukhova, A.A., Kurshakova, M.M., Nabirochkina, E.N., Georgieva, S.G., and Kopytova, D.V. (2021). PCID2, a subunit of the Drosophila TREX-2 nuclear export complex, is essential for both mRNA nuclear export and its subsequent cytoplasmic trafficking. RNA Biol. 18, 1969–1980. https://doi.org/10.1080/15476286.2021.1885198.

20. Stewart, M. (2019). Structure and Function of the TREX-2 Complex. Subcell. Biochem. 93, 461–470. https://doi.org/10.1007/978-3-030-28151-9_15.

21. Umlauf, D., Bonnet, J., Waharte, F., Fournier, M., Stierle, M., Fischer, B., Brino, L., Devys, D., and Tora, L. (2013). The human TREX-2 complex is stably associated with the nuclear pore basket. J. Cell Sci. 126, 2656–2667. https://doi.org/10.1242/jcs.118000.

22. Jani, D., Lutz, S., Hurt, E., Laskey, R.A., Stewart, M., and Wickramasinghe, V.O. (2012). Functional and structural characterization of the mammalian TREX-2 complex that links transcription with nuclear messenger RNA export. Nucleic Acids Res. 40, 4562–4573. https://doi.org/10.1093/nar/gks059.

23. Fischer, T., Strässer, K., Rácz, A., Rodriguez-Navarro, S., Oppizzi, M., Ihrig, P., Lechner, J., and Hurt, E. (2002). The mRNA export machinery requires the novel Sac3p-Thp1p complex to dock at the nucleoplasmic entrance of the nuclear pores. EMBO J. 21, 5843–5852. https://doi.org/10.1093/emboj/cdf590.

24. Bhatia, V., Barroso, S.I., García-Rubio, M.L., Tumini, E., Herrera-Moyano, E., and Aguilera, A. (2014). BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. Nature 511, 362–365. https://doi.org/10.1038/nature13374.

25. Nakaya, T., Kuwahara, K., Ohta, K., Kitabatake, M., Toda, T., Takeda, N., Tani, T., Kondo, E., and Sakaguchi, N. (2010). Critical role of Pcid2 in B cell survival through the regulation of MAD2 expression. J. Immunol. 185, 5180–5187. https://doi.org/10.4049/jimmunol.1002026.

26. Zhang, J., Zhai, J., Wong, C.C., Chen, H., Wang, X., Ji, J., and Yu, J. (2021). A novel amplification gene PCI domain containing 2 (PCID2) promotes colorectal cancer through directly degrading a tumor suppressor promyelocytic leukemia (PML). Oncogene 40, 6641–6652. https://doi.org/10.1038/s41388-021-01941-z.

27. Ye, B., Dai, Z., Liu, B., Wang, R., Li, C., Huang, G., Wang, S., Xia, P., Yang, X., Kuwahara, K., et al. (2014). Pcid2 inactivates developmental genes in human and mouse embryonic stem cells to sustain their pluripotency by modulation of EID1 stability. Stem Cell. 32, 623–635. https://doi.org/10.1002/stem.1580.

28. Ye, B., Liu, B., Yang, L., Huang, G., Hao, L., Xia, P., Wang, S., Du, Y., Qin, X., Zhu, P., et al. (2017). Suppression of SRCAP chromatin remodelling complex and restriction of lymphoid lineage commitment by Pcid2. Nat. Commun. 8, 1518. https://doi.org/10.1038/s41467-017-01788-7.

29. Marian, C.A., Stoszko, M., Wang, L., Leighty, M.W., de Crignis, E., Maschinot, C.A., Gatchalian, J., Carter, B.C., Chowdhury, B., Hargreaves, D.C., et al. (2018). Small Molecule Targeting of Specific BAF (mSWI/SNF) Complexes for HIV Latency Reversal. Cell Chem. Biol. 25, 1443–1455.e14. https://doi.org/10.1016/j.chembiol.2018.08.004.

30. Black, J.C., Van Rechem, C., and Whetstine, J.R. (2012). Histone lysine methylation dynamics: establishment, regulation, and biological impact. Mol. Cell 48, 491–507. https://doi.org/10.1016/j.molcel.2012.11.006.

31. Turner, A.M.W., and Margolis, D.M. (2017). Chromatin Regulation and the Histone Code in HIV Latency Yale. J Biol Med 90, 229–243.

32. Rao, S., Amorim, R., Niu, M., Temzi, A., and Mouland, A.J. (2018). The RNA surveillance proteins UPF1, UPF2 and SMG6 affect HIV-1 reactivation at a post-transcriptional level. Retrovirology 15, 42. https://doi.org/10.1186/s12977-018-0425-2.

33. Rondón, A.G., Jimeno, S., and Aguilera, A. (2010). The interface between transcription and mRNP export: from THO to THSC/TREX-2. Biochim. Biophys. Acta 1799, 533–538. https://doi.org/10.1016/j.bbagrm.2010.06.002.

34. Huang, Y., Yario, T.A., and Steitz, J.A. (2004). A molecular link between SR protein dephosphorylation and mRNA export. Proc. Natl. Acad. Sci. USA 101, 9666–9670. https://doi.org/10.1073/pnas.0403533101.

35. Hurt, E., Luo, M.J., Röther, S., Reed, R., and Strässer, K. (2004). Cotranscriptional recruitment of the serine-arginine-rich (SR)-like proteins Gbp2 and Hrb1 to nascent mRNA via the TREX complex. Proc. Natl. Acad. Sci. USA 101, 1858–1862. https://doi.org/10.1073/pnas.0308663100.

36. Yongxu Zhao, X.W., Liu, Y., Li, N., Wang, S., Sun, Z., Gao, Z., Zhang, X., Mao, L., Tang, R., Xue, W., et al. (2021). LENG8 regulation of mRNA processing, is responsible for the control of mitochondrial activity. Preprint at bioRxiv. https://doi.org/10.1101/2021.07.17.452750.

37. Stoltzfus, C.M., and Madsen, J.M. (2006). Role of viral splicing elements and cellular RNA binding proteins in regulation of HIV-1 alternative RNA splicing. Curr. HIV Res. 4, 43–55. https://doi.org/10.2174/157016206775197655.

38. Caputi, M., Mayeda, A., Krainer, A.R., and Zahler, A.M. (1999). hnRNP A/B proteins are required for inhibition of HIV-1 pre-mRNA splicing. EMBO J. 18, 4060–4067. https://doi.org/10.1093/emboj/18.14.4060.

39. Purcell, D.F., and Martin, M.A. (1993). Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. J. Virol. 67, 6365–6378. https://doi.org/10.1128/JVI.67.11.6365-6378.1993.

40. Cochrane, A. (2014). HIV-1 Rev function and RNA nuclear-cytoplasmic export. Methods Mol. Biol. 1087, 103–114. https://doi.org/10.1007/978-1-62703-670-2_9.

41. Pasternak, A.O., and Berkhout, B. (2021). The Splice of Life: Does RNA Processing Have a Role in HIV-1 Persistence? Viruses 13, v13091751. https://doi.org/10.3390/v13091751.

42. Cunningham, C.N., Schmidt, C.A., Schramm, N.J., Gaylord, M.R., and Resendes, K.K.

(2014). Human TREX2 components PCID2 and centrin 2, but not ENY2, have distinct functions in protein export and co-localize to the centrosome. Exp. Cell Res. *320*, 209–218. https://doi.org/10.1016/j.yexcr.2013.11.015.

43. Schneider, M., Hellerschmied, D., Schubert, T., Amlacher, S., Vinayachandran, V., Reja, R., Pugh, B.F., Clausen, T., and Köhler, A. (2015). The Nuclear Pore-Associated TREX-2 Complex Employs Mediator to Regulate Gene Expression. Cell *162*, 1016–1028. https://doi.org/10.1016/j.cell.2015.07.059.

44. González-Aguilera, C., Tous, C., Gómez-González, B., Huertas, P., Luna, R., and Aguilera, A. (2008). The THP1-SAC3-SUS1-CDC31 complex works in transcription elongation-mRNA export preventing RNA-mediated genome instability. Mol. Biol. Cell *19*, 4310–4318. https://doi.org/10.1091/mbc.e08-04-0355.

45. Mbonye, U., Kizito, F., and Karn, J. (2023). New insights into transcription elongation control of HIV-1 latency and rebound. Trends Immunol. *44*, 60–71. https://doi.org/10.1016/j.it.2022.11.003.

46. Jacquenet, S., Decimo, D., Muriaux, D., and Darlix, J.L. (2005). Dual effect of the SR proteins ASF/SF2, SC35 and 9G8 on HIV-1 RNA splicing and virion production. Retrovirology *2*, 1–13. https://doi.org/10.1186/1742-4690-2-33.

47. Wong, R.W., Balachandran, A., Ostrowski, M.A., and Cochrane, A. (2013). Digoxin suppresses HIV-1 replication by altering viral RNA processing. PLoS Pathog. *9*, e1003241. https://doi.org/10.1371/journal.ppat.1003241.

48. Emery, A., and Swanstrom, R. (2021). HIV-1: To Splice or Not to Splice, That Is the Question. Viruses *13*. https://doi.org/10.3390/v13020181.

49. Wilmes, G.M., Bergkessel, M., Bandyopadhyay, S., Shales, M., Braberg, H., Cagney, G., Collins, S.R., Whitworth, G.B., Kress, T.L., Weissman, J.S., et al. (2008). A genetic interaction map of RNA-processing factors reveals links between Sem1/Dss1-containing complexes and mRNA export and splicing. Mol. Cell *32*, 735–746. https://doi.org/10.1016/j.molcel.2008.11.012.

50. Ellisdon, A.M., Dimitrova, L., Hurt, E., and Stewart, M. (2012). Structural basis for the assembly and nucleic acid binding of the TREX-2 transcription-export complex. Nat. Struct. Mol. Biol. *19*, 328–336. https://doi.org/10.1038/nsmb.2235.

51. Ellisdon, A.M., and Stewart, M. (2012). Structural biology of the PCI-protein fold. BioArchitecture *2*, 118–123. https://doi.org/10.4161/bioa.21131.

52. Truman, C.T.S., Järvelin, A., Davis, I., and Castello, A. (2020). HIV Rev-isited. Open Biol. *10*, 200320. https://doi.org/10.1098/rsob.200320.

53. Jordan, A., Bisgrove, D., and Verdin, E. (2003). HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. EMBO J. *22*, 1868–1877. https://doi.org/10.1093/emboj/cdg188.

54. Naldini, L., Blömer, U., Gallay, P., Ory, D., Mulligan, R., Gage, F.H., Verma, I.M., and Trono, D. (1996). In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. Science *272*, 263–267. http://www.jstor.org/stable/2889637.

55. Schmittgen, T.D., and Livak, K.J. (2008). Analyzing real-time PCR data by the comparative C(T) method. Nat. Protoc. *3*, 1101–1108. https://doi.org/10.1038/nprot.2008.73.

56. Lagundžin, D., Krieger, K.L., Law, H.C.H., and Woods, N.T. (2022). An optimized co-immunoprecipitation protocol for the analysis of endogenous protein-protein interactions in cell lines using mass spectrometry. STAR Protoc. *3*, 101234. https://doi.org/10.1016/j.xpro.2022.101234.

57. Tyanova, S., Temu, T., and Cox, J. (2016). The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nat. Protoc. *11*, 2301–2319. https://doi.org/10.1038/nprot.2016.136.

58. Prins, H.A.B., Crespo, R., Lungu, C., Rao, S., Li, L., Overmars, R.J., Papageorgiou, G., Mueller, Y.M., Stoszko, M., Hossain, T., et al. (2023). The BAF complex inhibitor pyrimethamine reverses HIV-1 latency in people with HIV-1 on antiretroviral therapy. Sci. Adv. *9*, eade6675. https://doi.org/10.1126/sciadv.ade6675.

59. Pasternak, A.O., Adema, K.W., Bakker, M., Jurriaans, S., Berkhout, B., Cornelissen, M., and Lukashov, V.V. (2008). Highly sensitive methods based on seminested real-time reverse transcription-PCR for quantitation of human immunodeficiency virus type 1 unspliced and multiply spliced RNA and proviral DNA. J. Clin. Microbiol. *46*, 2206–2211. https://doi.org/10.1128/Jcm.00055-08.

60. Suzuki, K., Bose, P., Leong-Quong, R.Y., Fujita, D.J., and Riabowol, K. (2010). REAP: A two minute cell fractionation method. BMC Res. Notes *3*, 1–6. https://doi.org/10.1186/1756-0500-3-294.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-M2-FLAG	Sigma Aldrich	F3165; RRID:AB_259529
Anti-β-tubulin	Sigma Aldrich	T5168; RRID:AB_477579
Anti-PCID2	Genetex	GTX52023
Anti-Histone H3	Abcam	ab1791; RRID:AB_302613
Anti-Flag® M2 Affinity gel	Sigma Aldrich	A-2220; RRID:AB_10063035
Anti-Histone H3Ac	Millipore	17-615; RRID:AB_1163437
Anti-Histone H3K36me3	Abcam	ab9050; RRID:AB_306966
Anti-Histone H3K27me3	Active Motif	39055; RRID: AB_2561020
Anti-p24 KC57-RD1	Beckman Coulter	6604667; RRID:AB_1575989
Bacterial and virus strains		
DH5α	Thermo Fisher	18265017
Chemicals, peptides, and recombinant proteins		
Phorbol 12-myristate 13-acetate	Sigma Aldrich	P8139
Valproic acid sodium salt	Sigma Aldrich	P4543
Prostratin	Sigma Aldrich	P007
K98	N/A	N/A
Actinomycin D	Sigma Aldrich	A4262
Anhydrous DMSO	Sigma Aldrich	276855-100mL
Molecular grade methanol-free 16% formaldehyde	Polysciences Inc.	18814
EDTA-free protease inhibitor cocktail	Roche	5056489001
DAPI-containing Fluorescent Mounting Media	Dako Omnis	S3023
TRI Reagent	Sigma Aldrich	93289
Critical commercial assays		
Invitrogen™ PureLink™ HiPure Plasmid Miniprep Kit	Thermo Fisher	K210016
Dual-Glo® Luciferase Assay System	Promega	E2920
Protein A sepharose beads	GE Healthcare	17-0618-02
Protein G sepharose beads	GE Healthcare	17-0963-03
PrimeFlow RNA assay	Thermo Fisher	88-18005-204
EasySep Human CD4+ T cell enrichment kit	Stemcell Technologies	100-0696
Human T cell Nucleofector™ kit	Lonza	VPA-1002
Cell Line Nucleofector ™ kit R	Lonza	VCA-1001
High-capacity cDNA Reverse Transcription Kit	Applied Biosystems	4368814
GoTaq qPCR Master mix kit	Promega	A6001
DNAse I amplification grade Kit	Invitrogen	18068015
SuperScrip™t II reverse transcriptase	Invitrogen	18064014
Platinum™ Taq DNA Polymerase	Invitrogen	15966005
3' mRNA-seq Library Prep Kit Protocol for Ion Torrent	QuantSeq-LEXOGEN	012
DNA High Sensitivity Kit	Agilent Technologies	5067-4626
Ion PI Hi-Q OT2 200 Kit	Thermo Fisher	A26434

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
MS raw data	This study	PRIDE: PXD043334
RNA sequencing data	This study	GEO: GSE249455
Experimental models: Cell lines		
Jurkat cells	ATCC	TIB-152
J-Lat 11.1	(Jordan et al.[53]) Gifted by Eric Verdin	N/A
J-Lat 10.6	(Jordan et al.[53]) NIH Reagent Program	9854
J-Lat A2	(Jordan et al.[53]) NIH Reagent Program	9854
HEK 293T	ATCC	CRL-3216
Oligonucleotides		
PCR primers	Integrated DNA Technologies	Tables 1 and 2
Recombinant DNA		
sh Control	Sigma Aldrich	SHC002
sh PCID2	Sigma Aldrich	TRCN0000167267
sh PCID2	Sigma Aldrich	TRCN0000168656
sh IKZF1	Sigma Aldrich	TRCN0000107874
sh DSS1	Sigma Aldrich	TRCN0000083695
sh MCM3AP	Sigma Aldrich	TRCN0000147338
sh ENY2	Sigma Aldrich	TRCN0000285223
sh CETN3	Sigma Aldrich	TRCN0000055841
Lentiviral packaging construct pCMVΔR8.91	Naldini et al.[54]	N/A
VSV-G glycoprotein-expressing vector	Naldini et al.[54]	N/A
pBudCE4.1 (pBud control)	Sigma Aldrich	V53220
pBud-PCID2-Flag	This study	N/A
pNL4.3.Luc.R-.E-	NIH Reagent program	ARP-3418
pGL3-HSP	N/A	N/A
pRL Renilla	Promega	E2231
GagPol FISH-Flow probe	Thermo Fisher	VF10-10884
RPP30 FISH-Flow probe	Thermo Fisher	
Software and algorithms		
MaxQuant v2.1.3.0	Max Plank Institute of Biochemistry	https://www.maxquant.org/
Prism v8.3.0	GraphPad	https://www.graphpad.com/scientific-software/prism/
FlowJo v10	BD Biosciences	https://www.flowjo.com/
STRING database	STRING Consortium	https://string-db.org/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Tokameh Mahmoudi (t.mahmoudi@erasmusmc.nl)

Materials availability

Plasmids generated in this study are available upon request to lead contact, Tokameh Mahmoudi (t.mahmoudi@erasmusmc.nl).

Data and code availability

- Mass spectrometry data is available in PRIDE repository and RNA Sequencing data is available in GEO repository and can be accessed as of the date of publication. Accession numbers are listed in the key resources table.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this work paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Jurkat cells and latent HIV-1 infected Jurkat clones J-Lat 11.1, 10.6 and A2[53] were cultured in RPMI-1640 media supplemented with heat in-activated 7% Fetal Bovine Serum and 100 μg/mL Penicillin-Streptomycin at 37°C in a humidified 95% air 5% CO_2 incubator. Newly generated stable control or PCID2-Flag expressing J-Lat 11.1 cells were cultured in RPMI-1640 media supplemented with heat inactivated 7% Fetal Bovine Serum (FBS) and 100 μg/mL Penicillin-Streptomycin at 37°C in a humidified 95% air 5% CO_2 incubator. HEK 293T cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with heat inactivated 7% Fetal Bovine Serum and 100 μg/mL Penicillin-Streptomycin at 37°C in a humidified 95% air 5% CO_2 incubator.

METHOD DETAILS

Reagents

Cells were treated as indicated in the figures or figure legends with the following compounds: PMA (phorbol 12-myristate 13-acetate, Sigma Aldrich), Valproic acid (Sigma Aldrich), Prostratin (Sigma Aldrich), K98 (kind gift from Emily Dykhuizen), Actinomycin D (Sigma Aldrich). All chemicals were reconstituted following manufacturer's instructions.

Western blot

1-2 million cells (up to 5 million cells for PCID2 knockdown western blot) were lysed using an NP-40 IP lysis buffer (1% NP-40, 25mM Tris pH 7.4, 150 mM NaCl, 1mM EDTA, 5% glycerol, 1 U/mL EDTA-free protease inhibitor cocktail (Roche) and 1mM DTT) for 30 minutes on ice and centrifuged for 10 minutes at 14000 rpm in a cold centrifuge. Supernatants were collected, 1x Laemmli loading buffer was added and lysates were boiled for 5 min at 95°C and subjected to 12% SDS-PAGE separation. The following antibodies were used for detection of proteins by western blot: anti-M2-FLAG (Sigma, F3165), anti-β-tubulin (Sigma, T5168), anti-PCID2 (Genetex, GTX52023) and anti-Histone 3 (Abcam, ab1791) antibodies.

Lentiviral shRNA-mediated knockdown

Lentiviral constructs containing the shRNA targeting the gene of interest were obtained from the MISSIONx shRNA library (Sigma), facilitated by Erasmus Center for Biomics (key resources table). Plasmids were amplified in DH5α bacterial cells and isolated with Invitrogen™ PureLink™ HiPure Plasmid Miniprep Kit. Pseudotyped lentivirus were obtained by co-transfecting ∼7x10⁶ HEK293T cells with 6 μg of lentiviral construct with 4,5 μg of envelope plasmid pCMVΔR8.9 and 2 μg of packaging plasmid pCMV-VSVg mixed with 10 mM polyethyleneimine containing serum free-DMEM (transfection mix was incubated 15 min at room temperature) for 12 hours. Medium was then replaced with RPMI-1640 supplemented with 7% Fetal Bovine Serum (FBS) and 100 μg/mL Penicillin-Streptomycin, and supernatant was harvested 36, 48 and 60 hours post-transfection, filtered through a cellulose acetate membrane (0.45 μm pore) and stored at -80°C for later use.

Establishment of stable PCID2-Flag expressing lines

Stable PCID2-Flag expressing J-Lat 11.1 cell lines were generated by cloning the PCID2 open reading frame (ORF) in a pBud backbone (pBud Tag3 C-ter) behind an EF1α promoter and introducing a 2x FLAG tag at the C-term of the PCID2 ORF, using DH5α bacterial cells to propagate the plasmids. The plasmid contains a Bleomycin resistance cassette for bacterial selection and a Geneticin® cassette for mammalian cell selection. Stable cell lines were generated by nucleofecting 2 μg of pBud-PCID2-Flag plasmid or empty pBud as control using Amaxa Nucleofector (Lonza) and Nucleofector Kit R (Lonza) following manufacturer instructions. Briefly, 5x10⁶ were centrifuged at 1500 rpm for 5 min at room temperature and resuspended in 100 μL of solution R, and nucleofected in the presence of the plasmid mix containing pBud control or pBud-PCID2-Flag using program O28. Nucleofected cells were immediately incubated in pre-warmed serum free and antibiotic free RPMI media for 15 to 30 min at 37°C in a humidified 95% air 5% CO_2 incubator and then transferred to 5 mL of pre-warm RPMI-F7. 4 days after nucleofection cells were selected for two weeks with 0.5 mg/mL Geneticin® and expanded. A clone was selected based on exogenous expression of PCID2-Flag confirmed by western blot and comparable levels of PCID2 mRNA assessed by RT-PCR.

Effect of transient PCID2 overexpression on HIV-1 gene expression and latency reactivation

To assess the effect of PCID2 overexpression on HIV-1 reactivation from latency and viral RNA stability as described below, J-Lat 11.1 were nucleofected with 2 µg of pBud-PCID2-Flag plasmid or empty pBud as control as described above. After 4 days cells were selected for two weeks with 0.5 mg/mL Geneticin® and used for further experiments.

To determine the effect of PCID2 overexpression on HIV-1 gene expression, Jurkat cells were co-nucleofected with the following mix of plasmid constructs: 300 and 600 ng of empty pBud control or pBud-PCID2-Flag, 800 ng pNL4.3.Luc.R-.E- (NIH Reagents program) or pGL3-HSP vector and 200 ng pRL *Renilla* (Promega E2231) as control for nucleofection efficiency. Nucleofection was performed as described above. After 2 days, cells were processed for luciferase readings or RNA extraction as described below.

Dual luciferase assay

One million Jurkat cells transiently expressing pBud control or pBud-PCID2-Flag, pNL4.3.Luc.R-.E- or pGL3-HSP and pRL *Renilla* were collected 48 hours post-nucleofection and processed for luciferase readings using Dual-Glo® Luciferase Assay System (Promega) following manufacturer instructions. Renilla luciferase activity was used as a nucleofection efficiency control for normalization of Firefly luciferase readings.

Chromatin preparation

In order to obtain chromatin, 80-100 million cells (control and untreated or PMA-treated PCID2-Flag expressing cells; shControl or shPCID2 cells) were centrifuged and resuspended in PBS supplemented with 1mM $CaCl_2$ and 1mM $MgCl_2$ (PBS+/+). Cells were crosslinked by addition of Buffer A (0.1M NaCl, 1mM EDTA, 0.5 mM EGTA, 20mM HEPES) with formaldehyde (molecular grade methanol-free 16% formaldehyde, Polysciences Inc.) at a final percentage of 1% for 30 minutes at room temperature in a vertical rotator. The crosslinking reaction was stopped with 150 mM final concentration of Glycine and cells were washed once with PBS+/+. The pellets were washed consecutively once with cold Buffer B (0.25% Triton x-100, 1mM EDTA, 0.5 mM EGTA, 20mM HEPES pH 7.6) and once with cold Buffer C (150 mM NaCl, 1mM EDTA, 0.5 mM EGTA, 20mM HEPES pH 7.6) for 10 min on a vertical rotator at 4°C and centrifuged for 5 min 1500 rpm at 4°C. Pellets were resuspended in 2.5mL of ChIP Incubation buffer (1% Triton x-100, 150 mM NaCl, 1mM EDTA, 0.5 mM EGTA, 20mM HEPES pH 7.6) supplemented with 1% SDS and protease inhibitors (EDTA-free protease inhibitor cocktail; Roche) and transferred to 15 mL polystyrene Falcon tubes compatible with sonication. Chromatin was sonicated with Bioruptor (Diagenode) using 15mL probes for 12-15 cycles of 30 seconds ON 30 seconds OFF intervals at high intensity to achieve fragments in between 100 and 500bp. The sonicated chromatin was centrifuged for 15 min 14000 rpm at 4°C to get rid of cell debris.

To check for chromatin quality and fragment size, 100 µL of chromatin were added to 300 µL of ChIP Elution buffer (1% SDS, 0.1M $NaHCO_3$), brought to 200 mM NaCl final concentration, and decrosslinked for a minimum of 4 hours or overnight at 65°C on a heat block with shaker at 750 rpm. DNA extraction was performed using a standard phenol-chloroform isoamyl alcohol isolation protocol. Briefly, an equal volume of phenol:chloroform:isoamyl alcohol was added to the decrosslinked chromatin, mixed vigorously, and centrifuged at 14000 rpm for 5 min at 4°C. The aqueous phase was collected and an equal volume of 24:1 chloroform:isoamyl alcohol was added, mixed thoroughly and spun at 14000 rpm for 5 min at 4°C. The DNA contained in the aqueous phase was then precipitated with 100% ethanol in presence of glycogen as a carrier and NaAc, and snap frozen with liquid N_2. The DNA pellet was washed once with 70% ethanol, dried, and resuspended in 100 µL nuclease-free DEPC-treated water. Quality and size of the fragments were analyzed on a 1.2% agarose gel.

Chromatin immunoprecipitation (ChIP) – qPCR

To assess the enrichment of PCID2 at the HIV-1 LTR in control and exogenously expressing PCID2-Flag cells, we used the equivalent of 20-30 million cells as chromatin input. Chromatin was diluted with ChIP incubation buffer with no SDS to reach a final concentration of 0.15% SDS and pre-cleared overnight with 50 µL of a mix of Protein A/Protein G sepharose beads washed two times with ChIP incubation buffer 0.15% SDS. 100 µL per sample of anti-FLAG® M2 affinity beads (Sigma) were washed twice with ChIP incubation buffer 0.15% SDS and blocked overnight with 0.1% bovine serum albumin and 400 ng/mL of sheared salmon sperm DNA. An aliquot of the pre-cleared chromatin was set aside for quality control and input quantification. PCID2-Flag bound complexes were isolated by precipitating the pre-cleared chromatin with ~100µL of blocked affinity beads overnight at 4°C. The day after, beads were washed twice (10 minutes 4°C, 1500 rpm in a vertical rotator) with Buffer 1 (0.1% SDS, 0.1% DOC, 1% Triton x-100, 150mM NaCl, 1 mM EDTA pH 8.0, 0.5 mM EGTA, 20mM HEPES pH8.0) once per buffer with Buffer 2 (500mM NaCl: 0.1% SDS, 0.1% DOC, 1% Triton x-100, 500mM NaCl, 1 mM EDTA pH 8.0, 0.5mM EGTA, 20mM HEPES pH8.0) and Buffer 3 (0.25M LiCL, 0.5%DOC, 0.5% NP-40, 1mM EDTA pH 8.0, 0.5mM EGTA pH 8.0, 20mM HEPES pH 8.0), and twice with Buffer 4 (1mM EDTA pH8.0, 0.5mM EGTA pH 8.0, 20mM HEPES pH 8.0). Immunoprecipitated complexes were then eluted with ChIP Elution buffer (1% SDS, 0.1M $NaHCO_3$) for 30 minutes at room temperature in a vertical rotator.

For ChIP of core and modified histones and other factors in untreated and PMA-treated PCID2-Flag expressing cells, or in knockdown lines we used less chromatin input, corresponding to 10-20 million cells. Chromatin was pre-cleared and 100 µL per sample of a mix of Protein A/Protein G sepharose beads were blocked overnight as described above. The day after, bound complexes were isolated by overnight precipitation of the pre-cleared chromatin with 100 µL of blocked beads in combination with 3-4 µg of antibodies (anti- Histone H3 (Abcam ab17913, µg), Histone H3Ac (Millipore 17-615, 3µg), Histone H3K27me3 (Active Motif 39055, 3µg), Histone H3K36me3 (Abcam ab9050,

4μg), washed twice with Buffer 1, once per buffer with Buffer 2 and Buffer 3, and twice with Buffer 4, and eluted with ChIP Elution Buffer as described above.

DNA extraction was performed the day after using a standard phenol-chloroform isoamyl alcohol isolation protocol as described above. Enrichment of proteins, core and modified histones and transcription factors at the HIV-1 LTR was assessed by quantitative PCR using primers spanning the full promoter (Table 1) with GoTaq qPCR Master mix kit (Promega) in a CFX Connect Real-Time PCR thermocycler (BioRad). Relative enrichment over the DNA input was calculated with the $2^{-\Delta Ct}$ method.[55]

PCID2 immunoprecipitation for mass spectrometry

Twenty million control and stable exogenously expressing PCID2-Flag cells were lysed in 1 mL IP lysis buffer (1% NP-40, 25mM Tris pH 7.4, 150mM NaCl, 1mM EDTA, 5% glycerol, 1 U/mL EDTA-free protease inhibitor cocktail (Roche) and 1mM DTT) for 30 min on ice and centrifuged for 10 min 14000rpm at 4°C. Lysates were pre-cleared for 6 hours with 50 μL of a mix of Protein A and Protein G sepharose beads. 80 μL of Anti-FLAG M2 affinity gel per sample were washed twice in IP lysis buffer and incubated overnight with pre-cleared lysates. The day after, beads were washed twice with Wash buffer A (10 mM HEPES, pH 7.4, 10mM KCl, 10mM NaCl, 1 mM MgCl2, 0.05% Nonidet P-40) once with Wash buffer B (10 mM HEPES, pH 7.4, 10mM KCl, 0.07% Nonidet P-40), as described by Lagundzin, D. et al.,[56] and two times with PBS. Beads were further processed for Mass spectrometry as described below.

Mass spectrometry

Proteins were on-bead digested with sequencing grade trypsin (1:100 (w:w), Roche) overnight at room temperature. Protein digests were then desalted using a Sep-Pak tC18 Vac cartridge (Waters) and eluted with 80% acetonitrile (AcN). Peptides were then analyzed by nanoflow LC-MS/MS as described below.

Nanoflow LC-MS/MS was performed on an EASY-nLC system (Thermo) coupled to a Fusion Lumos Tribrid Orbitrap mass spectrometer (Thermo), operating in positive mode and equipped with a nanospray source. Peptide mixtures were trapped on a ReproSil C18 reversed phase column (Dr Maisch GmbH; column dimensions 1.5 cm × 100 μm, packed in-house) at a flow rate of 8 μl/min. Peptide separation was performed on ReproSil C18 reversed phase column (Dr Maisch GmbH; column dimensions 15 cm × 50 μm, packed in-house) using a linear gradient from 0 to 80% B (A = 0.1% FA; B = 80% (v/v) AcN, 0.1 % FA) in 70 or 120 min and at a constant flow rate of 250 nl/min. The column eluent was directly sprayed into the ESI source of the mass spectrometer. All mass spectra were acquired in profile mode. The resolution in MS1 mode was set to 120,000 (AGC: 4E5), the m/z range 350-1400. Fragmentation of precursors was performed in 2 s cycle time data-dependent mode by HCD with a precursor window of 1.6 m/z and a normalized collision energy of 30.0; MS2 spectra were recorded in the orbitrap at 30,000 resolution. Singly charged precursors were excluded from fragmentation and the dynamic exclusion was set to 60 seconds.

Data analysis: Mass spectrometric raw data were analyzed using the MaxQuant software suite (version 2.1.3.0[57] for identification and relative quantification of proteins. A false discovery rate (FDR) of 0.01 for proteins and peptides and a minimum peptide length of 6 amino acids were required. The Andromeda search engine was used to search the MS/MS spectra against the *Homo sapiens* Uniprot database (version May 2022) concatenated with the reversed versions of all sequences and a contaminant database listing typical background proteins. A maximum of two missed cleavages were allowed. MS/MS spectra were analyzed using MaxQuant's default settings for Orbitrap and ion trap spectra. The maximum precursor ion charge state used for searching was 7 and the enzyme specificity was set to trypsin. Further modifications were cysteine carbamidomethylation (fixed) as well as methionine oxidation. The minimum number of peptides for positive protein identification was set to 2. The minimum number of razor and unique peptides set to 1. Only unique and razor non-modified, methionine oxidized and protein N-terminal acetylated peptides were used for protein quantitation. The minimal score for modified peptides was set to 40 (default value).

MS raw data and data for protein identification and quantification were submitted as supplementary tables to the ProteomeExchange Consortium via the PRIDE partner repository with the data identifier PXD043334.

The identified hits in both IP-MS runs were first filtered for common contaminants, including filtering of keratin and immunoglobulin proteins. Next, hits were filtered by unique peptides larger than 1. Lastly, a fold enrichment was calculated of iBAQ value in PCID2-Flag samples over control samples, and only hits enriched >1.2 fold were included for further analysis. Common hits for both IP-MS runs were identified and pathway and interaction analysis was performed using STRING database.

RNA stability assay

Twenty million control or overexpressing PCID2-Flag lines were plated in 6 mL of RPMI-F7 medium and 10 μg/μL of Actinomycin D was added to the cultures. An aliquot of 1 million cells was collected and lysed in TRI reagent (Sigma) immediately, corresponding to timepoint 0 hours. Cells were then collected and lysed at 1 hour post-treatment and in several sequential timepoints of 2, 4, 6, 8 and 10 hours post-treatment initiation. Total RNA was extracted and 10 μL of RNA was used for cDNA synthesis as described above/below. Decay of HIV-1 unspliced RNA and beta-2-microglobulin was assessed by relative decrease in gene expression from timepoint 0 hours. Values calculated using the $2^{-\Delta\Delta Ct}$ and normalized to 18S RNA to account for RNA input variability.

FISH-Flow

To analyze the dynamics of viral RNA (vRNA) and/or GFP-producing cells by FISH-Flow, five million control and PCID2-knockdown cells were collected, fixed, permeabilized and processed as described before[13] using the PrimeFlow RNA assay (ThermoFisher Scientific) according to

manufacturer protocol. Viral RNA producing cells were detected by labelling HIV-1 unspliced mRNA with a set of 40 probe pairs against the GagPol region (Affymetrix eBioscience catalogue number GagPol HIV-1 VF10-10884) and RPL13A (Affymetrix eBioscience, catalogue number VA-13187) as a control diluted 1:5 in diluent provided in the kit and hybridized to the target HIV-1 vRNA for 2 hours at 40°C. As a control for hybridization, we used a probe set against the ribosomal RNA RPP. After hybridization, the excess probes were washed and cells were stored overnight in the presence of RNAsin. The next day, signal was amplified by incubating samples with sequential 1.5 hours, 40°C incubations with pre-amplification and amplification mix, followed by labelling of the amplified RNA with fluorescently-tagged probes for 1 hour at 40°C. Cells were acquired on a BD LSR Fortessa Analyzer and analysis was performed using FlowJo V10 software (Treestar). Gates to quantify viral RNA+ only cells, GFP+ only cells, and vRNA+ GFP+ cells were set using the shControl sample.

Intracellular p24 staining

Two million shControl or shPCID2 J-Lat 11.1 cells were pelleted and washed in PBS 3% FBS, then fixed and permeabilized with eBioscience™ IC Intracellular Fixation and Permeabilization Buffer Set following manufacturers instructions. Cells were stained with anti-p24 KC57-RD1 antibody (Beckman Coulter) for 45 minutes at 4°C in the dark, then washed with PBS and analysed by flow cytometry.

Confocal imaging

For confocal imaging of viral RNA and GFP in control and PCID2-knockdown cells, we used 100.000 cells processed for FISH-Flow as described above. Cells were dried on a slide, mounted with DAPI-containing Fluorescent Mounting Media (Dako Omnis), and visualized using a LSM700 (Zeiss) confocal microscope.

Total RNA isolation and RT quantitative PCR analysis

Cells were lysed with TRI reagent (Sigma) and total cell-associated RNA was isolated following manufacturer instructions. Up to one microgram of RNA was used to synthetized cDNA using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems 4368814) according to kit protocol. Real time quantitative PCR was performed using GoTaq qPCR Master mix kit (Promega) in a CFX Connect Real-Time PCR thermocycler (BioRad). Normalized relative gene expression was calculated with the $2^{-\Delta\Delta Ct}$ method using PPIA as reference gene. Primer list is available in Table 1.

Transcriptional assay in primary cells from people living with HIV-1

CD4+ T cells were isolated from frozen PBMCs by negative selection with EasySep Human CD4+ T cell enrichment kit (STEMCELL Technologies) following manufacturer's protocol and left to rest overnight in RPMI media supplemented with 10% FBS and 1% Penicillin/Streptomycin. 10-15 million CD4+ T cells were then transfected with 5 μg of shControl or shPCID2 #2 vector and 2 μg of pmaxGFP™ vector using Human T cell Nucleofector™ kit following manufacturer's instructions. 48 hours after transfection, cells were collected and lysed in TRI reagent (Sigma) and total RNA was isolated as described before.[58] We then DNAse treated the RNA and synthetized cDNA using, respectively, DNAse I amplification grade kit (Invitrogen) and SuperScript II reverse transcriptase (Invitrogen) following the provided protocols.

To analyse the change in US and CS vRNA we performed a nested qPCR assay targeting Gag (US vRNA) and Tat-Rev (CS vRNA) as described before[58,59] with minor modifications. Briefly, for pre-amplification of US and CS vRNA 10 μL of diluted cDNA were used in a final volume of 25 μL using Platinum™ Taq DNA polymerase kit (Invitrogen) and nested US forward and reverse 1 primers for US vRNA detection or nested CS forward and reverse primers for CS vRNA detection (Table 2). The PCR reaction was the following: 95°C for 5 min followed by 15 cycles at 95°C for 30 seconds, 55°C for 30 seconds and 72°C for 15 minutes. For the second amplification, 2 μL of pre-amplified cDNA were used in a final volume of 25 μL using Platinum™ Taq DNA polymerase kit and nested US forward and reverse 2 primers and nested US probe or CS forward 2 and reverse primers and CS probe at 95°C for 5 min followed by 45 cycles at 95°C for 30 seconds and 60°C for 1 minute. To calculate the release in blocks at post-transcriptional levels of gene expression we first assessed the relative abundance of US vRNA and CS vRNA species normalized to housekeeping gene and then calculated the ratios of CS vRNA over CS vRNA as described in the figure legends. Knockdown efficiency was assessed by qPCR with primers targeting PCID2 (Table 1) using GoTaq qPCR Master mix kit as described below.

3' mRNA sequencing and analysis

2 million shControl and shPCID2 J-Lat 11.1 cells were collected for RNA sequencing analysis in duplicate. Total RNA was isolated as described above and cDNA libraries were generated using the 3' mRNA-seq Library Prep Kit Protocol for Ion Torrent (QuantSeq-LEXOGEN) following manufacturer's instructions. Library quality was assessed on a Bionalayzer using the DNA High Sensitivity Kit reagent and protocol (Agilent Technologies). Libraries were pooled and template and enriched using an Ion Proton One Touch system with Ion PI Hi-Q OT2 200 Kit (Thermo Fisher Scientific). Sequencing was performed using the Ion PI Hi-Q Sequencing 200 Kit on Ion Proton PI V2 chips (Thermo Fisher Scientific) according to manufacturer's protocols. Reads were mapped on the UCSC hg19 reference genome and processed for normalization as described before (elife paper). Differentially expressed gene analysis of obtained reads was performed using edgeR package under Galaxy (https://usegalaxy.eu/). Cut-off for false discovery rate was 0.01 and a 1.5 fold change for differentially expressed. Volcano plots were generated using Galaxy (https://usegalaxy.eu/).

Subcellular fractionation

To assess the relative abundance of HIV-1 viral RNA species in the nucleus and cytoplasm of shControl and shPCID2 J-Lat 11.1 cells, we used a modified version of REAP protocol by Suzuki et al.[60] 2 million cells were lysed in 900 μL of PBS 0.1% NP-40 and mechanically triturated for 4-5 times with a micropipette. 300 μL are taken aside as whole cell lysate of which 100 μL are added to 900 μL of TRI reagent for RNA isolation. The remaining 600 μL are spun for 1 minute at 14000 rpm in a cold centrifuge. 300 μL of the supernatant are set aside as cytoplasmic fraction of which 100 μL are added to 900 μL of TRI reagent for RNA isolation. The rest of the supernatant is discarded and 900 μL of PBS 0.1% NP-40 is added to wash the pellet, which corresponds to the nuclear fraction, for 1 minute at 14000 rpm in a cold centrifuge. Supernatant is discarded and pellet is resuspended in 1 mL of cold TRI reagent. Total RNA was isolated from samples in TRI reagent and processed for cDNA synthesis as described above. RNA concentration of each fraction independently was normalized amongst samples. The relative abundance of unspliced and completely spliced HIV-1 viral RNA species in the nucleus and cytoplasm was calculated with the $2^{-\Delta Ct}$ method. Ratios were calculated for each vRNA species in the nucleus and cytoplasm and presented as percentages. To check the efficiency of the subcellular fractionation at a protein level, a parallel fractionation protocol was performed as described above except 4x SDS loading buffer with 100 mM DTT is added to the collected whole cell lysate and cytoplasmic fractions. The pellet corresponding to the nuclear fraction is dissolved in 300 μL of 1x SDS loading buffer with 100 mM DTT and sonicated with Bioruptor (Diagenode) 5 cycles of 30 seconds ON 30 seconds OFF intervals at high intensity. Samples were then processed for western blot analysis.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis was performed as indicated in the figure legends using Graphpad Prism v8.3.0.