VIROLOGY

Suppression of HIV-1 transcription and latency reversal via ectopic expression of the viral antisense transcript AST

Rui Li[1], Kaveh Daneshvar[2]†, Xinjie Ji[3]‡, Michelle Pleet[4]§, Grace Igbinosun[1], Mohd Shameel Iqbal[1], Fatah Kashanchi[4], Alan C. Mullen[5], Fabio Romerio[1,6]*

The mechanisms that regulate HIV-1 latency are not fully elucidated. Our previous studies showed that an HIV-1 antisense transcript (AST) promotes the deposition of histone modifications at the HIV-1 5′ long terminal repeat, causing a closed chromatin state that suppresses viral transcription. Here, we report that ectopic expression of AST in CD4$^+$ T cells from people living with HIV-1 undergoing antiretroviral therapy hinders the reactivation of viral transcription in response to ex vivo stimulation with pharmacologic and T cell receptor agonists, thus preventing the reversal of latency. We defined the structural domains and sequence motifs of AST that contribute to its latency-promoting functions. Last, we carried out an unbiased proteomic screen of AST interactors that revealed an array of host factors both previously known and not known to suppress HIV-1 expression. Our studies identify AST as a first-in-class biological molecule that is capable of enforcing HIV-1 latency and with actionable curative potential.

INTRODUCTION

HIV-1 integrates preferentially in regions of the host genome that contain actively transcribed genes (*1*, *2*). At these sites, HIV-1 transcription is regulated through mechanisms that control chromatin organization at the proviral 5′ long terminal repeat (LTR), independently of the surrounding genomic context (*3*). Two nucleosomes, Nuc-0 and Nuc-1, are located at precise positions of the 5′ LTR, irrespective of the proviral integration site (*4*). Histone modifications in Nuc-0 and Nuc-1 control the chromatin structure at the HIV-1 promoter and regulate the switch between active transcription and latency (*5*, *6*). The ability to modulate the addition or removal of these histone marks can be exploited in HIV-1 cure strategies that seek either to eradicate HIV-1 through latency reversal (*7*) or to force HIV-1 into a state of irreversible latency (*8*). Various classes of latency reversal agents have been evaluated clinically with limited success (*9*–*12*), whereas strategies aiming at permanent latency via epigenetic silencing of HIV-1 transcription remain largely unexplored.

Trimethylation of histone H3 at lysine 27 (H3K27me3) in Nuc-1 induces a closed chromatin state and suppresses HIV-1 transcription (*13*–*15*). The polycomb repressor complex 2 (PRC2), which is responsible for depositing the H3K27me3 mark, lacks direct DNA binding activity (*16*). Long noncoding RNAs (lncRNAs), including natural antisense transcripts (NATs), recruit PRC2 to gene promoters, leading to nucleosome assembly, heterochromatinization, and transcriptional silencing (*17*, *18*). HIV-1 expresses an antisense transcript (AST) from a Tat-independent negative sense promoter within the 3′ LTR (*19*–*21*). AST covers 2574 nucleotides (nt) from the U3 region of the 3′ LTR through the *nef* gene and part of *env*. Studies from our group and others showed that AST is responsible for the recruitment of PRC2 to the 5′ LTR and for the establishment and maintenance of HIV-1 latency (*20*, *22*, *23*). Here, we present evidence that ectopic expression of AST restricts HIV-1 transcription and blocks reversal of latency in CD4$^+$ T cells from people living with HIV-1 (PLWH) under suppressive antiretroviral therapy (ART), and we elucidate the underlying molecular mechanisms.

RESULTS

Identification of the AST domain and motifs interacting with the HIV-1 proviral 5′ LTR

We reported that the HIV-1 AST promotes the establishment and inhibits the reversal of viral latency in in vitro cell line models (*22*). To identify the functional determinants of AST, we divided its sequence into five domains: U3 (residues 1 to 376), A (residues 377 to 926), B (residues 927 to 1476), C (residues 1477 to 2026), and D (residues 2027 to 2574) (fig. S1). We then generated a panel of deletion and substitution mutants targeting each of the five domains or motifs therein for testing in functional and molecular analyses (fig. S1 and table S1). The secondary structure (minimum free energy) of wild-type AST and its mutants was predicted with RNAfold (fig. S2) (*24*, *25*). The AST deletion and substitution mutants were then stably transduced into the latently infected cell line Jurkat E4 (JE4). The HIV-1 proviral genome in JE4 cells lacks the *gag* and *pol* genes, and it carries *gfp* inserted in place of *nef*, which allows us to rely on green fluorescent protein (GFP) expression as a surrogate marker of viral expression (fig. S3) (*13*). We have previously used this cell model to demonstrate that AST promotes HIV-1 latency via interaction with PRC2 (*22*).

[1]Department of Molecular and Comparative Pathobiology, Johns Hopkins School of Medicine, Baltimore, MD, USA. [2]Division of Gastroenterology, Massachusetts General Hospital, Boston, MA, USA. [3]Department of Biochemistry and Molecular Biology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA. [4]Laboratory of Molecular Virology, George Mason University, Manassas, VA, USA. [5]Division of Gastroenterology, University of Massachusetts Chan Medical School, Worcester, MA, USA. [6]W. Harry Feinstone Department of Molecular Microbiology and Immunology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA.
*Corresponding author. Email: fromeri2@jhmi.edu
†Present address: Arbor Biotechnologies, Cambridge, MA, USA.
‡Present address: University of Maryland, Baltimore, Baltimore, MD, USA.
§Present address: National Institutes of Health, Bethesda, MD, USA.

lncRNAs recruit epigenetic and transcription factors to the chromatin via sequence homology and base pairing with specific genomic DNA regions (*26*, *27*). The 376-nt segment at the 5′ terminus of AST that covers part of the U3 region in the 3′ LTR shares perfect sequence identity with the homologous region of the 5′ LTR, and, thus, it may be essential for the interaction between AST and 5′ LTR (fig. S4). We generated a cell line in the JE4 background, stably expressing only the 5′ terminus segment of AST (JE4-U3AST; figs. S1 and S2) (*22*). This cell line showed a markedly higher percentage of GFP-positive (GFP^+) cells compared to both parental JE4 cells and JE4 expressing full-length AST (JE4-AST) (Fig. 1A). These results were confirmed by RT-qPCR assays using primers specific for both *gfp* and *env* (fig. S5). A higher percentage of GFP^+ cells in the JE4-U3AST compared to the other two cell lines reflects the ability of U3AST to interact with the 5′ LTR, displacing the endogenous and biologically active AST that is naturally expressed by the HIV-1 molecular clone in JE4 cells (fig. S3) (*22*), and the inability of U3AST to recruit silencing factors to the 5′ LTR. Thus, U3AST acts as a dominant negative competitor of endogenous AST, ultimately causing latency reversal.

RNA-DNA interactions occur preferentially between pyrimidine-rich (Y-rich) regions in the RNA molecule and purine-rich regions in the DNA (*28*, *29*). The U3 domain of AST contains two Y-rich motifs (Fig. 1B and fig. S5): a 24-nt Y1 motif (residues 114 to 137) with 83% pyrimidine content and a 29-nt Y2 motif (residues 184 to 212) with 86% pyrimidine content. We generated an AST mutant in which we disrupted the Y1 and Y2 motifs with alternating purine-pyrimidine stretches of equal length [50 and 48% pyrimidine content, respectively; AST domain U3 mutant (ASTmutU3)]; Fig. 1B). All other domains, motifs, and sequences of AST were left unchanged. A JE4-derived cell line stably expressing ASTmutU3 showed a percentage of GFP^+ cells higher than JE4-AST and comparable to the parental JE4 cell line (Fig. 1C), suggesting that the disruption of the two Y-rich motifs abolishes the ability of AST to bind the 5′ LTR, reversing the percentage of latently infected cells to baseline levels of the parental JE4 cell line. The levels of GFP and ENV RNA detected by RT-qPCR paralleled these results (fig. S5). This conclusion was verified via chromatin isolation by RNA precipitation (ChIRP)–qPCR assays. We generated 24 biotinylated oligonucleotide probes targeting the portion of AST sequence comprised between residues 377 and 2574, which are shared in

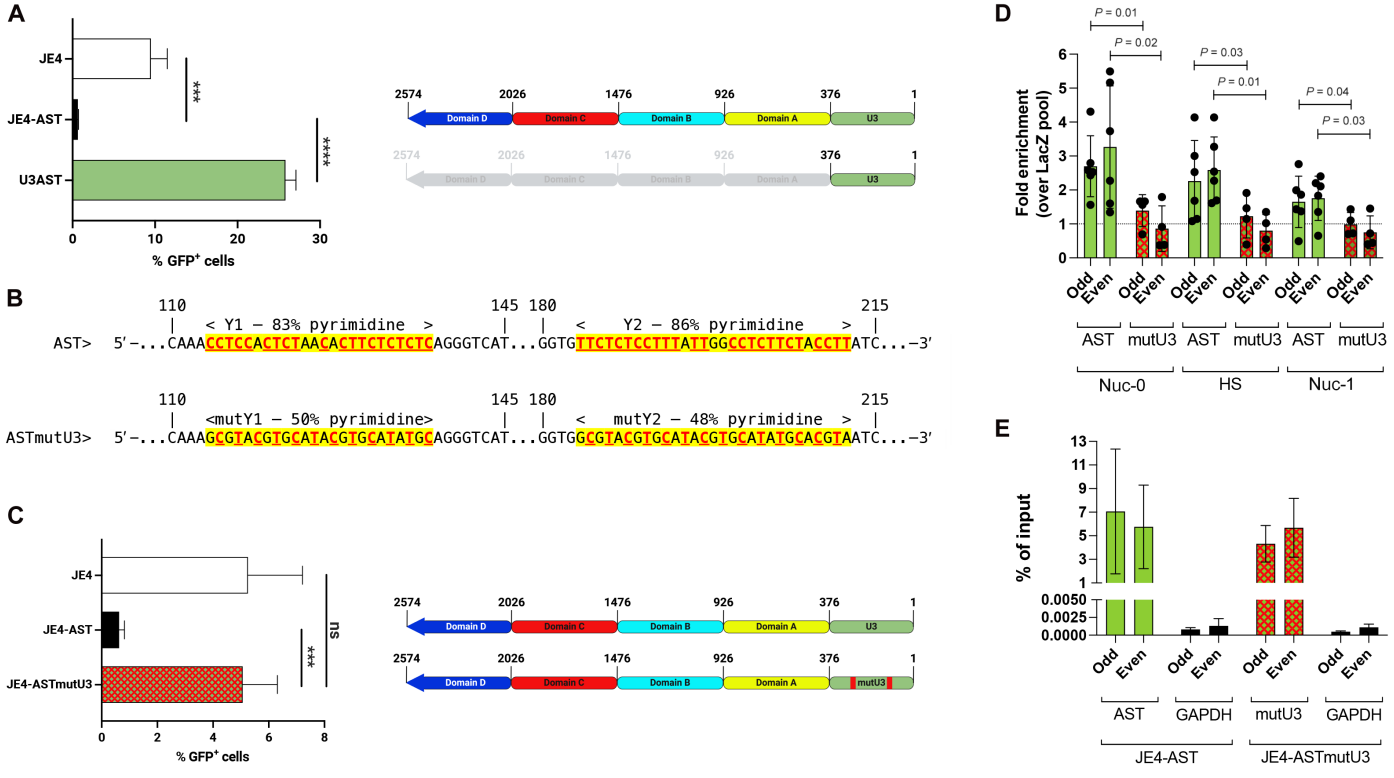

Fig. 1. The 5′ terminus domain of AST is required for interaction with the proviral 5′ LTR. (**A**) JE4 cells stably transduced with a lentiviral vector driving ectopic overexpression of AST (JE4-AST) show lower frequency of productively infected cells compared to parental JE4 cells, whereas JE4 cells stably transduced with only the 376-nt U3 domain at the 5′ terminus of AST (JE4-U3AST) contain a significantly higher percentage of productively infected cells. (**B**) Top: Nucleotide sequence and position of two Y-rich motifs (Y1 and Y2) in the U3 domain of wild-type AST, which have 83 and 86% pyrimidine content, respectively. Bottom: Nucleotide substitutions disrupting the Y-rich motifs of Y1 and Y2, reducing pyrimidine content in the Y1 and Y2 mutants (mutY1 and mutY2, respectively) sequences of the ASTmutU3 construct (50 and 48% pyrimidine content, respectively). All other domains and motifs of ASTmutU3 are the same as in the wild-type AST sequence. Pyrimidines are indicated in red, bold, and underlined letters. (**C**) Cultures of JE4 cells stably transduced with a lentiviral vector expressing the ASTmutU3 mutant contain the same percentage of productively infected cells as parental JE4 cells (~4 to 6% GFP^+ cells). (**D**) ChIRP-qPCR assays show the physical interaction of wild-type AST with the Nuc-0 and HS domains in the U3 region of the HIV-1 5′ LTR. A weak interaction between AST and the Nuc-1 domain of the 5′ LTR may reflect incomplete chromatin fragmentation. No interaction is observed between these two domains and ASTmutU3 carrying mutY1 and mutY2 mutations. Results are shown as fold enrichment over *lacZ*-specific biotinylated antisense oligonucleotide pools. (**E**) RT-qPCR shows equal enrichment of wild-type AST and ASTmutU3 using odd and even biotinylated antisense oligonucleotide pools specific for the AST sequence covering domains A to D. No enrichment of GAPDH RNA is observed with either pool of oligonucleotides. ***$P < 0.0005$; ****$P < 0.0001$; ns, not significant.

wild-type AST and ASTmutU3 (fig. S1). The probes were assigned in alternate fashion, progressing along the AST to either of two pools (odd or even; table S2) and were then used to pull down AST and ASTmutU3 from the respective whole cell lysates. Genomic DNA coprecipitating with the two transcripts was probed by qPCR assays targeting the Nuc-0 (positions −405 to −245 relative to the +1 transcription start site), the deoxyribonuclease I (DNase I) hypersensitivity (HS) regions (positions −245 to +1), and the Nuc-1 region (positions +20 to +165) of the 5′ LTR. We found that AST interacts with Nuc-0 and HS sequences in the U3 region of the 5′ LTR (~3-fold and ~2.5-fold enrichment over *lacZ* control, respectively; Fig. 1D). Weak interaction with the Nuc-1 region is likely due to incomplete chromatin fragmentation (~1.5-fold enrichment; Fig. 1D). No binding was observed between ASTmutU3 and the three regions of the 5′ LTR (Fig. 1D). RT-qPCR assays confirmed equal recovery of AST and ASTmutU3 with the two pools of probes, which did not enrich glyceraldehyde-3-phosphate dehydrogenase (GAPDH) RNA negative control (Fig. 1E). Therefore, the U3 domain of AST (residues 1 to 376) interacts with the proviral 5′ LTR. Two polypyrimidine motifs within this domain are critical for the AST-5′ LTR interaction.

Identification of the AST domain and motifs interacting with PRC2

Next, we sought to identify the domain of AST that interacts with PRC2. The JE4 line stably transduced with the AST domain B mutant (ASTmutB) showed significantly higher frequency of GFP$^+$ cells and higher levels of GFP and ENV RNA compared to both the parental JE4 and JE4-AST lines (Fig. 2A and fig. S5). Thus, we hypothesized that segment B may be involved in recruiting PRC2 and promoting latency. Replacement of segments A, C, and D (ASTmutA, ASTmutC, and ASTmutD, respectively) with the same unrelated exogenous sequence had a much more limited effect (Fig. 2A and fig. S5).

Studies have shown that PRC2 binds G-rich sequences containing $[G_3N_{1-5}]_{\geq 4}$ motifs able to form G-quadruplex structures (*30*–*32*). Using the Quadruplex-forming G-rich Sequences (QGRS) Mapper server (*33*), we identified two $[G_3N_{3-7}]_4$ motifs in AST that match quadruplex-forming G-rich sequences (a G-score of 38 for both) (Fig. 2B and fig. S6). These two motifs are overlapping and map within segment B of AST: one at residues 1048 to 1075 and the other at residues 1057 to 1082 (fig. S6). We replaced a 70-nt sequence of AST

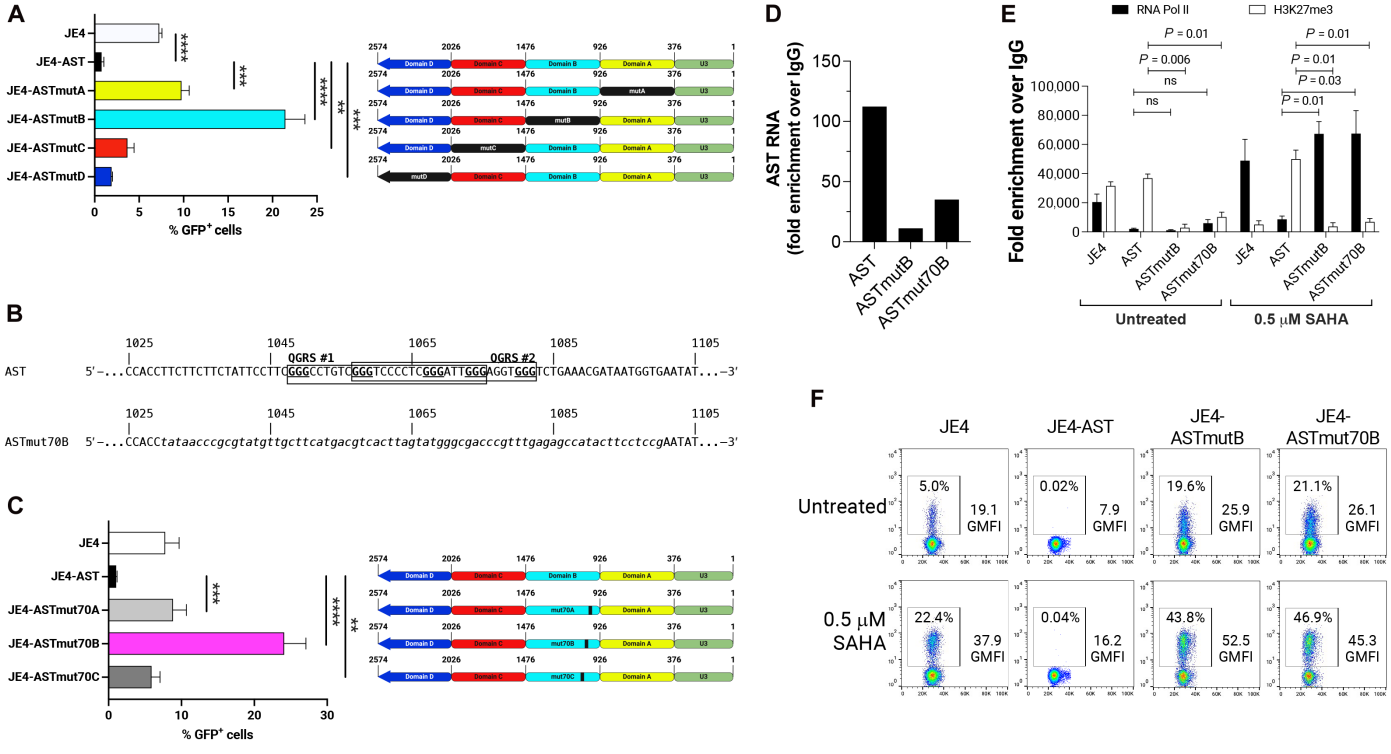

Fig. 2. Domain B of AST is primarily responsible for interaction with PRC2. (**A**) Frequency of GFP$^+$ cells in cultures of JE4 stably transduced with AST in which domains A, B, C, and D have been replaced with an exogenous sequence of equal length (ASTmutA, ASTmutB, ASTmutC, and ASTmutD). ASTmutB shows the greatest impact on AST activity. (**B**) Top: Sequence and position of two QGRS motifs within domain B of AST (QGRS #1 and QGRS #2 are enclosed in black-lined boxes). G-triplets are indicated in bold, underlined font. Bottom: Substitution of 70 nt encompassing both QGRS motifs in domain B of AST with an exogenous sequence of equal length (ASTmut70B). Sequence substitution is in lowercase italics. All other domains and motifs of ASTmut70B are maintained. (**C**) JE4 cells stably transduced with ASTmut70B contain a higher percentage of GFP$^+$ cells compared to JE4-AST cells. ASTmut70A or ASTmut70C show a much more limited impact. (**D**) RIP–RT-qPCR assays show physical interaction of wild-type AST with EZH2. This interaction is reduced in cells expressing the ASTmutB or ASTmut70B. (**E**) ChIP-qPCR assays probing H3K27me3 and RNA Pol II levels at Nuc-1 in JE4, JE4-AST, JE4-ASTmutB, and JE4-ASTmut70B before and after treatment with 0.5 μM SAHA. In JE4 cells, SAHA treatment lowered the levels of H3K27me3 at Nuc-1, whereas in JE4-AST cells, they remained high. In the ASTmutB and ASTmut70B mutants, H3K27me3 levels were low both before and after SAHA treatment. Moderate levels of RNA Pol II poised at Nuc-1 in untreated JE4 cells increased after SAHA treatment. RNA Pol II levels at Nuc-1 were low both before and after SAHA treatment in JE4-AST cells. (**F**) Flow cytometry analyses of GFP expression in JE4, JE4-AST, JE4-ASTmutB, and JE4-ASTmut70B reflect those from ChIP-qPCR (E). $^{**}P < 0.005$; $^{***}P < 0.0005$; $^{****}P < 0.0001$. IgG, immunoglobulin G.

encompassing the two overlapping motifs with an exogenous sequence of equal length (ASTmut70B; residues 1031 to 1100; Fig. 2B and figs. S1 and S2). As controls, we used the same exogenous sequence to replace 70 nt upstream (ASTmut70A; residues 946 to 1015) or downstream (ASTmut70C; residues 1141 to 1210) of the mut70B site (figs. S1 and S2). A JE4-derived cell line stably transduced with the ASTmut70B mutant expressed GFP protein, RNA, and ENV RNA at significantly higher levels than JE4-AST (Fig. 2C and fig. S5). The ASTmut70A and ASTmut70C mutations had a smaller impact (Fig. 2C and fig. S5).

RNA immunoprecipitation (RIP)–RT-qPCR assays with antibodies against EZH2, a core component and catalytic subunit of PRC2 complex, showed substantially reduced enrichment of ASTmutB and ASTmut70B compared to wild-type AST in the EZH2 immunocomplexes (Fig. 2D). Chromatin immunoprecipitation (ChIP)–qPCR assays were performed with cells untreated and treated with the latency reversal agent, suberoylanilide hydroxamic acid (SAHA; 0.5 μM), using antibodies to RNA polymerase II (Pol II) and H3K27me3 and qPCR primers for Nuc-1. JE4-AST cells showed high levels of H3K27me3 both before and after SAHA treatment, whereas JE4-ASTmutB and ASTmut70B cells showed the inverse scenario (Fig. 2E). The levels of Pol II at Nuc-1 in JE4-AST cells were very low both before and after SAHA treatment. In untreated JE4-ASTmutB and ASTmut70B cells, RNA Pol II levels at Nuc-1 were low but increased significantly following SAHA treatment (Fig. 2E). Overall, ChIP-qPCR results were consistent with the expression levels of GFP protein, RNA, and those of ENV RNA (Fig. 2F and fig. S5). In particular, the increased percentage of GFP^+ cells in untreated ASTmutB and ASTmut70B cells suggests that the low levels of RNA Pol II observed at Nuc-1 in these cultures (Fig. 2E) are consistent with a fast rate of transcription initiation, as confirmed by RT-qPCR assays (fig. S5). The increase in the percentage of GFP^+ cells and geometric mean fluorescence intensity (GMFI) after SAHA treatment reflects augmented RNA Pol II recruitment (Fig. 2E). Overall, the two overlapping $[G_3N_{3-7}]_4$ motifs in segment B strongly contribute to PRC2 binding and recruitment to the HIV-1 5′ LTR.

Identification of additional AST interactors at the 5′ LTR

Mutation of domains C and D (ASTmutC and ASTmutD, respectively) individually and the deletion of domain D (ASTdelD; figs. S1 and S2) had limited impact on HIV-1 expression (Figs. 2A and 3A and fig. S5). However, mutation or deletion of both domains concurrently (ASTmutCD and ASTdelCD; figs. S1 and S2) had a profound effect (Fig. 3A and fig. S5), suggesting that domains C and D may contribute to the interaction between AST and additional host factors.

To address this hypothesis, we fractionated JE4-AST whole cell lysates by size exclusion chromatography. RT-qPCR analyses showed that ~80% of AST eluted in a very narrow range of fractions containing macromolecular complexes of ≥2.2 MDa in size (Fig. 3B). Given that the combined molecular weight of AST and the core PRC2 complex is estimated at ~1 MDa (*34*), >1.2 MDa remained unaccounted for in the AST-containing complexes that elute in fractions 11 to 15. This suggests that AST associates with several other host factors in large multimolecular complexes.

To identify additional AST interactors, we appended four copies of the streptavidin-binding S1m RNA aptamer at the 3′ terminus of AST and scrambled sequence (SCR) RNA (AST-4×S1m and SCR-4×S1m; Fig. 3C) (*35*, *36*). 4×S1m-containing RNA molecules were pulled down with streptavidin beads from lysates of human embryonic kidney 293T cells stably transduced with vectors expressing the two chimeric constructs. Copurifying proteins in the 30- to 150-kDa range significantly enriched in the AST-4×S1m pulldown were identified by mass spectrometry (MS) (data S1). These include transcription factors and subunits of epigenetic and chromatin remodeling complexes previously shown to promote HIV-1 silencing and latency, such as components of BAF/PBAF complex (SMARCD1-3), Yin Yang 1 (YY1), histone deacetylase 2 (HDAC2), and components of the polycomb complex (RBBP4 and RNF2) (Fig. 3D and Table 1) (*5*, *37*–*40*). The MS screen also identified host factors involved in heterochromatinization of the X chromosome via interaction with the *Xist* lncRNA (*41*): TAR DNA binding protein 43 (TDP-43), human nuclear ribonuclear protein K (hnRNPK), matrin 3 (MATR3), and polypyrimidine tract binding protein 1 (PTBP1). RIP assays confirmed the interaction between AST and some of these factors (Fig. 3E). TDP-43 was first identified as a DNA binding factor binding the TAR sequence of the HIV-1 proviral genome (*42*), and, therefore, it represents a special case. To ensure that the interaction between AST and TDP-43 is not simply due to residual low amounts of contaminating TAR DNA in the cell lysates, we performed RIP assays for TDP-43 using cell lysates treated with DNase I. The results show that AST binds TDP-43 with equal efficiency in DNase-treated and untreated cell lysates (fig. S7). Therefore, this interaction is not mediated by residual TAR DNA. Together, the HIV-1–silencing properties of AST may occur via interaction with and recruitment of a macromolecular complex of host factors acting cooperatively.

Blockade of ex vivo latency reversal by AST in cells from ART-suppressed PLWH

We reported that AST promotes the establishment and inhibits the reversal of viral latency in in vitro cell line models (*22*). Here, we assessed whether ectopic expression of AST restricts HIV-1 transcription and prevents latency reversal ex vivo. A vector expressing either AST or a control SCR under the eukaryotic translation initiation factor 1A (EIF1α) promoter was introduced via nucleofection into resting $CD4^+$ T cells freshly isolated from peripheral blood of PLWH undergoing suppressive ART for at least 1 year (median, 6 years; range, 1 to 18 years; table S3). Both vectors also expressed GFP under the human cytomegalovirus promoter (Fig. 4A). Latency reversal was then triggered by treating the cells with the latency reversal agents SAHA and panobinostat or through T cell receptor stimulation with anti-CD3/CD28–coated magnetic beads (Fig. 4A). All three treatments reactivated HIV-1 transcription to various degrees in donor cells expressing SCR but not AST ($P < 0.002$, $P = 0.016$, $P = 0.031$ for SAHA, panobinostat, and anti-CD3/CD28, respectively; Fig. 4B and fig. S10). Efficiency of nucleofection was assessed on the basis of the percentage of GFP^+ cells, and it ranged between 20 and 40% (Fig. 4C and fig. S8). Given the expected low levels of protein translation in resting $CD4^+$ T cells, the expression of GFP is likely to underestimate the efficiency of nucleofection. For a more accurate estimate that can account for the near-complete blockade of latency reversal in cells ectopically expressing AST, we labeled the two plasmid vectors with the fluorescent dye, cyanine-5 (Cy5), and we assessed the percentage of Cy5-positive cells after nucleofection (Fig. 4C and fig. S7). The estimate of nucleofection efficiency through this method was consistently between 75 and 90%, and it was 1.5- to 3-fold greater than the one based on GFP expression ($P = 0.001$ and $P = 0.0005$ for cells transfected with SCR and AST vectors,

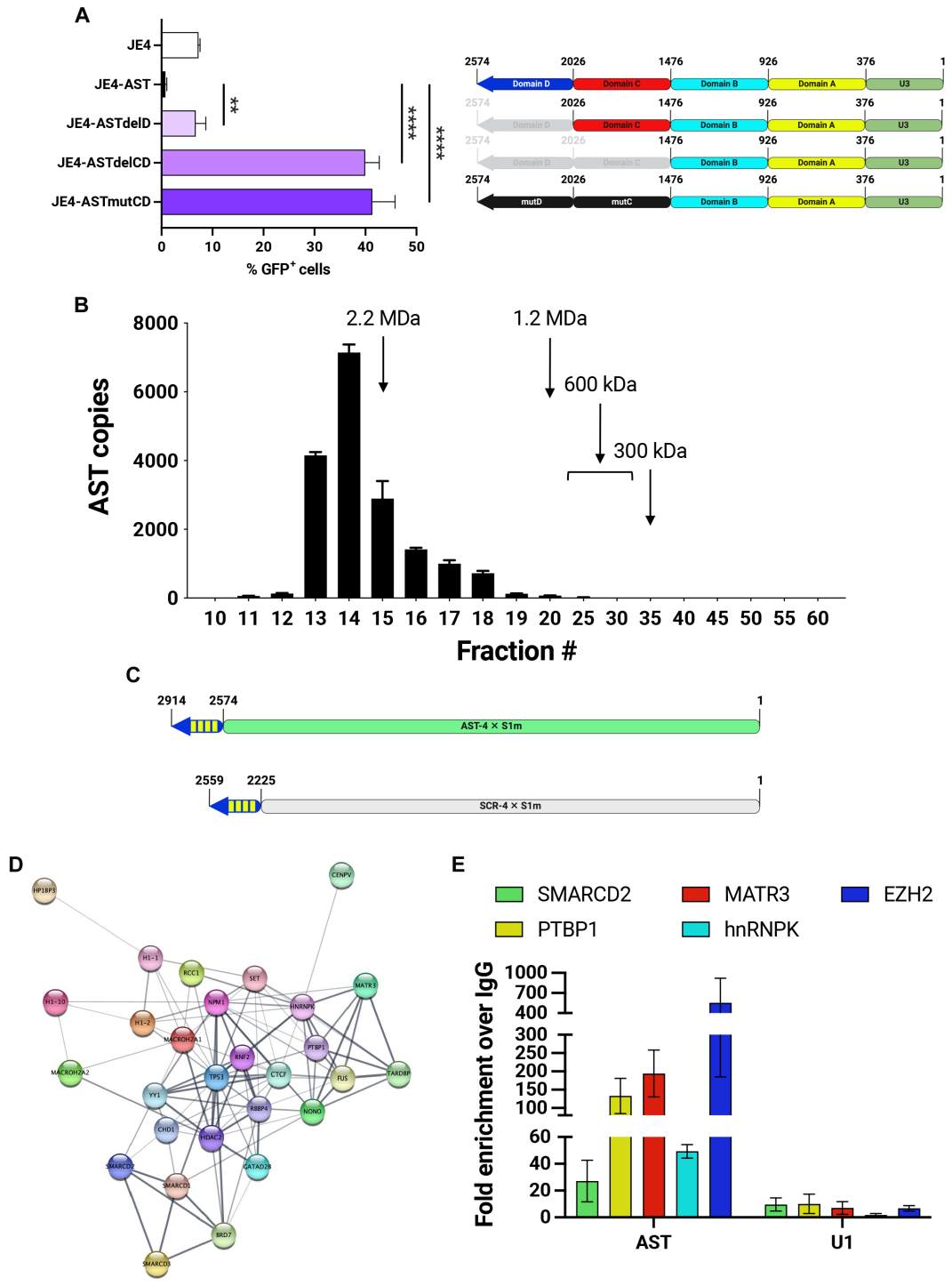

Fig. 3. Domains C and D of AST contribute to the recruitment of additional chromatin factors. (**A**) Limited impact of delD of AST (ASTdelD) on the frequency of productively infected cells (~5 to 7% GFP$^+$ cells). This is consistent with results from ASTmutD cultures. Deletion or substitution of both domains C and D (ASTdelCD and ASTmutCD, respectively) profoundly affects AST activity (36 to 38 and 46 to 48% GFP$^+$ cells, respectively). (**B**) RT-qPCR assays measuring AST levels in size exclusion chromatography fractions from lysates of JE4-AST cells. Greater than 80% of AST molecules elute in fractions 11 to 15, which contain complexes of estimated molecular weight ≥2.2 MDa. (**C**) Schema of AST and SCR molecules with four copies of the streptavidin-binding S1m RNA aptamer (4×S1m) fused at the 3′ terminus. (**D**) Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) network of selected AST interactors included in Table 1. (**E**) RIP–RT-qPCR assays validating the interaction between AST and four additional interactors, SMARCD2, PTBP1, MATR3, and hnRNPK, selected from the list in Table 1. EZH2 serves as a positive control. U1 small nuclear RNA (snRNA) is used as a negative control. $^{**}P < 0.005$; $^{****}P < 0.0001$.

Table 1. Subset of AST interactors. AST-binding factors identified by MS in AST pulldown streptavidin complexes. The list of proteins reported here is known to play a role in transcription regulation in various capacities, and it is a subset of AST interactors reported in data S1. Fold enrichment is relative to SCR-4×S1m control.

Protein	Fold enrichment	Function
SMARCD3	15.250	Chromatin remodeling; SWI/SNF complex
CENPV	14.000	Heterochromatin assembly
YY1	11.724	Transcription factor
CHD1	10.417	Chromatin remodeling
SMARCD1	10.339	Chromatin remodeling; SWI/SNF complex
SMARCD2	5.370	Chromatin remodeling; SWI/SNF complex
TP53	4.974	Transcription factor
CTCF	4.524	3D chromatin regulator
HIST1H1A	4.286	Histone H1 isoform; higher-order chromatin
RCC1	3.837	Chromatin condensation
H2AFY2	3.811	Histone H2A isoform; transcription repressor
GATAD2B	3.484	Transcription repressor; NuRD complex
RNF2	3.465	Transcription repressor; PcG complex
PTBP1	3.367	Gene silencing; biomolecular condensates
NONO	3.535	Scaffold protein; biomolecular condensates
HDAC2	3.262	NuRD complex
SET	3.011	Gene silencing; biomolecular condensates
BRD7	2.996	Transcription factor
HP1BP3	2.757	Heterochromatin organization
HNRNPK	2.459	Gene silencing; biomolecular condensates
HIST1H1C	2.308	Histone H1 isoform; higher-order chromatin
MATR3	2.308	Gene silencing; biomolecular condensates
NPM1	2.207	Scaffold protein; biomolecular condensates
H2AFY	2.176	Histone H2A isoform; transcription repressor
TARDBP	2.174	Transcription repressor; TAR DNA binding
FUS	2.115	RBP; biomolecular condensates
RBBP4	2.065	Chromatin remodeling; NuRD, PRC2 complex
H1FX	2.036	Histone H1 isoform; higher-order chromatin

respectively). Expression levels of both AST and SCR were analyzed by strand-specific RT-qPCR in both cell cultures (Fig. 4D). As previously reported (*22*), we detected endogenous expression of AST in $CD4^+$ T cells from PLWH after nucleofection with SCR-expressing vector. Much higher expression levels were detected after nucleofection of the AST expression vector ($P = 0.0003$). SCR RNA was readily detected following nucleofection of the SCR-expressing vector but only at background levels in AST-nucleofected cells ($P = 0.008$). Together, these results show that ectopic overexpression of AST in cells from ART-suppressed PLWH effectively blocks the reversal of latency and viral transcription in response to various stimuli.

DISCUSSION

Antisense RNA–based mechanisms ensure a tight control of gene expression by establishing a threshold of activation that buffers weak stimuli and represses stochastic expression of the target gene (*43*, *44*). Stimuli that exceed such thresholds switch on gene expression, which rapidly reaches maximal levels (*45*, *46*). As stimuli subside, antisense transcripts accelerate the return to baseline levels (*47*). *Xist*, *ANRIL*, and *Kcnq1ot1* are examples of antisense transcripts that function in such manner via the recruitment of suppressive epigenetic factors that trigger a closed chromatin state and silence transcription (*27*, *48*, *49*).

The impact of AST on HIV-1 transcription involves mechanisms similar to those described for eukaryotic antisense transcripts. AST is expressed constitutively from a Tat-independent negative sense promoter in the 3′ LTR (*19*–*22*). However, we showed that AST levels increase markedly in response to several stimuli (*50*), similar to other antisense transcripts with gene regulatory function (*27*). AST has a predominantly nuclear localization due to inefficient polyadenylation (*51*) and, possibly, interaction with nuclear RNA binding proteins (RBPs) (Table 1 and data S1) (*52*). Our previous studies using cell line models demonstrated that AST promotes HIV-1 latency by recruiting the chromatin remodeling complex PRC2 to the 5′ LTR (*22*). Here, we provided proof that AST hinders latency reversal in an ex vivo model based on $CD4^+$ T cells from ART-suppressed PLWH. This ability is especially remarkable in consideration of the HIV-1 sequence variability within and among infected individuals.

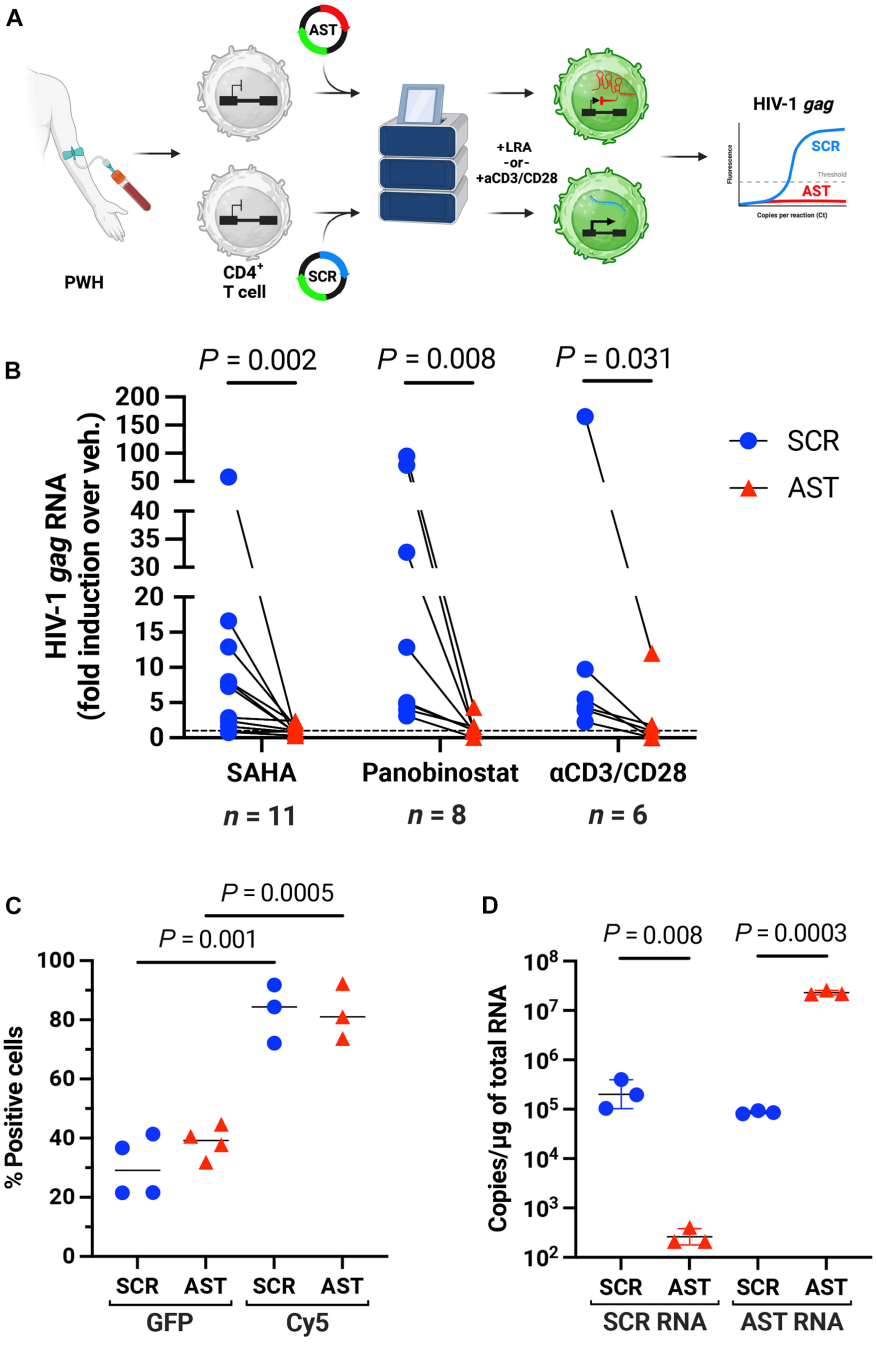

Fig. 4. Blockade of latency reversal in CD4$^+$ T cells from PLWH after ectopic expression of AST. (**A**) Diagram of experimental design. CD4$^+$ T cells were isolated by negative selection from the peripheral blood of 15 PLWH under ART (table S3). After nucleofection with vectors expressing either AST or SCR, cells were treated with SAHA, panobinostat, or anti-CD3/CD28 magnetic beads. HIV-1 *gag* RNA levels were analyzed by seminested RT-qPCR. (**B**) HIV-1 *gag* RNA expression in CD4$^+$ T cells from PLWH nucleofected with vectors expressing AST (red triangles) or SCR (blue circles) and then treated with SAHA (left; $n = 11$), panobinostat (middle; $n = 8$), or anti-CD3/CD28 antibodies (right; $n = 6$). RNA levels are expressed as fold increase over vehicle (veh.) control [dimethyl sulfoxide (DMSO) for SAHA and panobinostat and phosphate-buffered saline (PBS) for anti-CD3/CD28]. *P* values were determined with the Wilcoxon matched-pairs signed-rank test. (**C**) Efficiency of nucleofection in freshly isolated, resting CD4$^+$ T cells from PLWH. This was determined by both the expression of GFP from the same vector also expressing AST or SCR (fig. S8) and by intracellular Cy5 content when using Cy5-labeled vectors in nucleofection (fig. S9). Frequency of transfected cells measured with the two methods were significantly different both for cells transfected with SCR vector and for cells transfected with AST vector ($P = 0.001$ and $P = 0.0005$, respectively). (**D**) Quantitative RT-qPCR assays measuring AST and SCR RNA expression levels in freshly isolated, resting CD4$^+$ T cells from PLWH nucleofected with SCR- or AST-expression vectors. The levels of the two transcripts were significantly different between cells nucleofected with each of the transcript-expressing vectors and cells nucleofected with the other transcript-expressing vector ($P = 0.008$ for SCR levels in SCR-nucleofected cells versus AST-nucleofected cells; $P = 0.0003$ for AST levels in AST-nucleofected cells versus SCR-nucleofected cells). Higher baseline levels of AST in SCR-nucleofected cells reflect endogenous expression of AST in PLWH-derived samples.

Investigation of the functional properties of lncRNAs has traditionally relied on genome-level ablation with programmable nucleases and transcript-level knockdown with small interfering RNA molecules and antisense oligonucleotides (*53*). In the case of HIV-1, genome-level manipulation is complicated by the fact that ablation or mutation of the AST gene would also affect other viral genetic elements. In addition, classical RNA knockdown approaches are inefficient for lncRNAs that have primarily nuclear localization and that are highly structured (*54*). On the other hand, ectopic expression has been used less commonly to investigate the function of an lncRNA but is more suitable to study the mechanisms of action through expression of the lncRNA with targeted mutations (*53*, *55*).

Because of their complex secondary and tertiary structure and their length, lncRNAs serve as scaffolds for the recruitment of proteins that regulate multiple biological functions, including DNA repair and replication, transcription, and chromatin remodeling (*56*, *57*). Above a threshold concentration, lncRNAs can selectively partition specific proteins into biomolecular condensates that are formed by phase separation (*58*, *59*). By contrast, knocking down expression of lncRNAs or deleting essential protein-interacting domains prevents formation of the condensates and their function (*60*, *61*). The lncRNAs *eRNA* (*61*), *HSATII* (*62*), *HOTAIR* (*63*), *NEAT1* (*64*), *DIGIT* (*36*), and *Xist* (*65*) regulate transcription through their interaction with RBPs that contribute to the recruitment of chromatin remodeling factors (*27*, *56*).

Similar to many other lncRNAs, AST is a highly structured molecule. Deletion or substitution of even short sequence motifs may have short- and long-range effects on AST folding that affect its function. All AST mutants we have generated and tested so far have shown some loss of activity compared to wild-type AST. This supports the hypothesis that AST serves as a scaffold in the formation of biomolecular condensates at the 5′ LTR through the recruitment of multiple host factors. Our studies showed that AST is associated exclusively with large macromolecular complexes of >2 MDa that include RBPs such as hnRNPK, MATR3, PTBP1, TDP-43, FUS, NONO, and NPM1 (Table 1 and data S1), which are known to cooperate with lncRNA during phase separation and the formation of biomolecular condensates (*27*, *56*, *57*). Deletion or mutation of structural domains and sequence motifs of AST eliminates its ability to promote latency. In a previous study, we also identified and mapped AST epigenetic and posttranscriptional modifications, including 2′-*O*-ribose methylation, adenosine-to-inosine, and pseudouridine (*66*), which are catalyzed by enzymes identified in the screen of AST interactors [fibrillarin (FBL), FtsJ RNA 2′-O-Methyltransferase 3 (FTSJ3), and dyskerin pseudouridine synthase 1 (DKC1); data S1] and have been shown to facilitate lncRNA-RBP interactions (*67*–*72*).

Cure and treatment strategies that seek to permanently suppress viral transcription remain underdeveloped compared to those proposing to flush out or disable proviruses hiding in cellular and anatomical reservoirs (*7*, *8*). This is especially true for approaches that aim to achieve epigenetic silencing of the HIV-1 promoter (*73*, *74*). Our published studies and the ones described here show that AST establishes a tight control of HIV-1 transcription, inhibiting latency reversal in response to external stimuli and accelerating reestablishment of latency when such stimuli wane (*22*). For future in vivo studies, we envision an approach similar to the one we have used in our in vitro studies, namely, ectopic overexpression of AST from a lentiviral vector because stable, long-term nuclear expression of AST is likely needed to stabilize HIV-1 latency. Ectopic expression of AST at high levels is expected to augment the local concentration of epigenetic and transcriptional repressors at the 5′ LTR and to suppress latency reversal and viral expression. Together, this provides a rationale for testing the use of AST as a curative agent that can restrict HIV-1 transcription and stabilize viral latency.

MATERIALS AND METHODS

Participant cohorts, sample collection, and study approval

Peripheral blood was collected from PLWH on ART, who were continually suppressed (<50 HIV-1 RNA copies/ml) for at least 1 year at the time of blood collection (table S3). Donors were enrolled at the Johns Hopkins Hospital after approval from the Johns Hopkins Institutional Review Board (protocol no. IRB00280013). All donors provided written informed consent.

Isolation of resting CD4$^+$ T cells, nucleofection, latency reversal, and HIV-1 RT-qPCR

Peripheral blood mononuclear cells were isolated from peripheral blood by Ficoll centrifugation and used to enrich resting CD4$^+$ T cells by negative selection via magnetic-activated cell sorting (Miltenyi Biotec). Untouched CD4$^+$ T cells were then nucleofected with a vector expressing AST (*22*) or SCR under the EF1α promoter using the Amaxa P3 Primary Cell nucleofection kit and an Amaxa Nucleofector four-dimensional (4D) (Lonza) programmed for an EH-100 pulse code. After 48 hours, cells were treated with 0.5 μM SAHA [BEI Resources, National Institutes of Health (NIH) HIV Reagent Program, Division of AIDS, National Institute of Allergy and Infectious Diseases (NIAID), NIH, ARP-12130], 30 nM panobinostat (Selleckchem), and anti-CD3/CD28 Dynabeads (1 bead per cell; Thermo Fisher Scientific). As a vehicle control for cells treated with SAHA and panobinostat, cells were treated with dimethyl sulfoxide (DMSO; 1:2000; Sigma-Aldrich). As a vehicle control for cells treated with anti-CD3/CD28 and panobinostat, cells were treated with an equal volume of phosphate-buffered saline (PBS; Thermo Fisher Scientific). All treatments were performed in the presence of 2 μM raltegravir (BEI Resources, NIH HIV Reagent Program, NIAID, NIH, HRP-11680) and interleukin-2 (1 U/ml; Thermo Fisher Scientific). Nucleofected cells treated with vehicle were used to determine baseline HIV-1 transcription. After 72 hours, cells were collected in TRIzol (Thermo Fisher Scientific), and total RNA was extracted using the Direct-zol RNA Microprep kit (Zymo Research) with on-column DNase treatment. Cell-associated, unspliced HIV-1 gag RNA levels were then determined by RT-qPCR relative to vehicle-treated samples (DMSO for SAHA and panobinostat; PBS for anti-CD3/CD28 Dynabeads). Total RNA (50 ng) was reverse transcribed into cDNA using Superscript III Reverse Transcriptase (Thermo Fisher Scientific). Briefly, RNA was incubated at 65°C for 5 min with 1.5 μg of random hexamers, 0.25 μg of oligo(dT) (12 to 18) before adding the first-strand buffer, 5 mM dithiothreitol (DTT), 20 U of RNAseOUT RNAse inhibitor, and 5% RT enzyme in a volume of 20 μl. The reaction was incubated at 42°C for 45 min and 80°C for 5 min. The cDNA was then used to perform a seminested real-time qPCR assay with primers in table S4, as previously described (*11*, *75*). The PCR conditions for Gag RNA were 10% AmpliTaq Buffer II, 2.5 mM $MgCl_2$, 0.2 mM of each deoxynucleotide triphosphate, 0.4 μM of each primer, and 0.5% AmpliTaq Gold polymerase in a reaction volume of 50 μl and run at 95°C for 10 min, followed by 15 cycles of 94°C for 20 s, 55°C for 40 s, and 72°C for 40 s. For the

second round of PCR, 2 μl of first-round PCR product was used as a DNA template with 47% PowerUp SYBR Green master mix (Thermo Fisher Scientific) and 0.75 μM of each primer. qPCR conditions were as follows: 50°C for 2 min and then 95°C for 10 min, followed by 40 cycles of 95°C for 20 s and 60°C for 40 s. All samples were run together with no RT control wells to assess for DNA contamination. Expression data for the fold change over vehicle control were calculated using the $2^{-\Delta\Delta Ct}$ method.

Cell lines, reagents, and lentiviral transduction

Lenti-X 293T cells (Takara) were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum (FBS), penicillin, streptomycin, and L-glutamine. The latently infected cell line JE4 (a gift from J. Karn, Case Western Reserve University) (*6*, *13*) was cultured in RPMI 1640 supplemented with 10% FBS and antibiotics. Lentiviral transduction was carried out as we described previously (*22*). Briefly, AST-derived sequences were cloned into the pLVX-Puro vector (Takara) and then transfected into Lenti-X 293T cells using the Lenti-X Packaging Single Shots system (Takara). Culture supernatants were collected, and viral particles were concentrated using the PEG-it reagent (System Biosciences). JE4 cells were then transduced with purified lentiviral particles overnight in complete RPMI 1640 medium containing polybrene (2 μg/ml). Stably transduced cells were selected and maintained in medium containing puromycin (1 μg/ml).

Flow cytometry

Plasmid DNA was labeled with the DNA Label IT Tracker Intracellular Nucleic Acid Localization kit (Mirus, Madison, WI) using 0.5 μl of Cy5 per microgram of plasmid DNA according to the manufacturer's protocol. Unreacted Label IT Tracker reagent was removed by ethanol precipitation. For flow cytometry analyses, samples were acquired on an LSRII Fortessa flow cytometer (BD Biosciences, San Jose, CA). Doublet cells were excluded using forward scatter and side scatter. DNA uptake into cells was determined by analysis of anaphase-promoting complex–positive cells. Raw data were analyzed with FlowJo software.

RNA isolation, reverse transcription, and quantitative PCR

RNA isolation and RT-qPCR were carried out as previously described (*22*). Briefly, RNA was isolated using the RNeasy Mini Kit (QIAGEN), digested with Turbo DNase (Ambion), and further purified with the ribonuclease (RNase) Mini Kit. RNA (100 to 500 ng) was used for reverse transcription using the iScript Select cDNA Synthesis Kit (Bio-Rad) and the AST-tagged RT primer. For the TaqMan PCR reaction, we used the TaqMan Gene Expression Master Mix (Applied Biosystems) or the PowerUp SYBR Green master mix (Thermo Fisher Scientific) according to the manufacturer's instructions. For quantitative PCR, the AST amplicon carrying the tag sequence was cloned into pUC57. The plasmid was then linearized and used at 10^0 to 10^5 copies to generate a PCR standard curve. The copy number per 3 μl of cDNA was then normalized to the equivalent number of cells, which were measured manually with a hemacytometer. For RT-qPCR of the U1 small nuclear RNA (snRNA), RT reaction was carried out with the same kit as above and with the random primers contained in the kit. The sequence of primers and probes is reported in table S4.

RNA immunoprecipitation

A total of 5 million cells were washed twice with ice-cold PBS and lysed with a buffer containing 100 mM KCl, 5 mM $MgCl_2$, 10 mM Hepes (pH 7), 0.5% IGEPAL CA-630, 1 mM DTT, protease inhibitors, and RNaseOUT in molecular biology grade water. After 5 min incubation on ice, the lysate was frozen at −80°C overnight. For TDP-43 RIP assays, cell lysates were treated with 400 U/ml of DNase I (Roche, catalog no. 04716728001) for 30 min at room temperature before freezing. Protein A + G Magnetic Beads (50 μl) were washed with 500 μl of NT-2 buffer [50 mM tris-HCl (pH 7.4), 150 mM NaCl, 1 mM $MgCl_2$, and 0.05% IGEPAL CA-630 in molecular biology grade water] and incubated with 5 μg of antibody of preimmune rabbit immunoglobulin G (IgG) or target-specific antibodies in table S5 with rotation at room temperature for 1 hour. Excess antibody was removed with six washes with NT-2 buffer on ice. After the last wash, 900 μl of NET-2 buffer [20 mM EDTA (pH 8), 1 mM DTT, and RNaseOUT in NT-2 buffer] was added to the antibody-Dynabeads complex. The lysate was thawed at room temperature and centrifuged at 14,000 rpm for 10 min at 4°C. Supernatant (100 μl) was added to the antibody-Dynabeads complex, followed by overnight incubation with rotation at 4°C, and 10 μl of supernatant was saved at −80°C as input control. After washing six times on ice, the protein-antibody-Dynabeads complex and input were digested with 150 μl of Proteinase K buffer (1% SDS and 180 μg of Proteinase K solution in NT-2 buffer) at 55°C for 30 min. The supernatants were transferred to new Eppendorf tubes and vortexed with 250 μl of NT-2 Buffer and 400 μl of Phenol/Chloroform 5:1 (Fisher Bioreagents) for 15 s, followed by a 10-min centrifugation at 14,000 rpm at room temperature. Three hundred fifty microliters of the aqueous phase was collected and vortexed with 400 μl of chloroform for 15 s, followed by another 10-min centrifugation at 14,000 rpm at room temperature. Three hundred microliters of the aqueous phase was collected for an overnight RNA precipitation at −80°C with 50 μl of 5 M ammonium acetate, 15 μl of 7.5 M LiCl, 5 μl of glycogen (1 mg/ml), and 850 μl of ethanol. RNA was pelleted down by centrifuging at 14,000 rpm for 30 min at 4°C and washed with 500 μl of 80% ethanol. The RNA pellet was centrifuged again at 14,000 rpm for 15 min at 4°C to remove the supernatant and left on the bench to air dry. Last, RNA was resuspended in 15 μl of molecular biology grade water for RT-qPCR analysis.

Chromatin immunoprecipitation

A total of 20 million cells were placed in two 150-mm culture dishes in 20 ml of medium, and 540 μl of 37% formaldehyde (Sigma-Aldrich, catalog no. 252549) was added to each culture dish and incubated for 10 min at room temperature with gentle shaking. Two milliliters of 10× glycine (Cell Signaling Technology, 7005) was added to each culture dish and incubated for 5 min at room temperature with gentle shaking. Cross-linked cells were combined and pelleted at 1000*g* for 5 min at 4°C and washed twice with ice-cold PBS. Cell pellet was resuspended in 1 ml of 1× ChIP Sonication Cell Lysis Buffer (Cell Signaling Technology, 96529) with Protease Inhibitor Cocktail (PIC; Cell Signaling Technology, 7012) and stored at −80°C if not processed immediately. Cell suspension was incubated on ice for 10 min and then pelleted down at 5000*g* for 5 min at 4°C. The pellet was resuspended in 1 ml of 1× ChIP Sonication Cell Lysis Buffer with PIC and incubated on ice for another 10 min. The cell suspension was centrifuged at 5000*g* for 5 min at 4°C. The pellet

was resuspended in 1 ml of ice-cold 1× ChIP Sonication Nuclear Lysis Buffer (Cell Signaling Technology, 28778) with PIC and aliquoted into four 500-µl PCR tubes for sonication. Lysates were sonicated at 70% amplitude with 15-s on and 30-s off cycles for 25 min. After sonication, lysates were combined into a 1.5-ml Eppendorf tube and clarified by centrifugation at 21,100*g* for 10 min at 4°C. Supernatant was collected to a new tube and stored at −80°C if not processed immediately. The clarified lysate (50 µl) was mixed with 100 µl of nuclease-free water, 6 µl of 5 M NaCl (Cell Signaling Technology, 7010), and 2 µl of RNase A (Cell Signaling Technology, 7013) and incubated at 37°C for 30 min. Then, 2 µl of Proteinase K (Cell Signaling Technology, 10012) was added, and the sample was incubated at 65°C for 2 hours to overnight. After incubation, chromatin was purified with QIAquick PCR Purification Kit (QIAGEN, 28106). Purified chromatin (15 µl) was taken to determine DNA fragment size by electrophoresis on a 1% agarose gel. ChIP was performed if 60 to 90% of total DNA fragments are <1 kb. For ChIP, sonicated chromatin was diluted with 1× ChIP Buffer (Cell Signaling Technology, 7008) at a dilution factor of 1:4. Diluted chromatin (10 µl) was transferred to a microcentrifuge tube as 2% input and stored at −80°C until future use. Diluted chromatin (500 µl) was transferred to microcentrifuge tubes with immunoprecipitating antibodies (table S5) and incubated at 4°C with rotation for 4 hours to overnight. Then, 30 µl of Protein G Magnetic Beads (Cell Signaling Technology, 9006) was added to each IP reaction and incubated 4°C with rotation for 2 hours. Beads were recovered with a magnetic separation rack and washed three times with low salt wash (1 ml of 1× ChIP Buffer) and once with high salt wash (1 ml of 1× ChIP Buffer plus 70 µl of 5 M NaCl). One hundred fifty microliters of 1× ChIP Elution Buffer (Cell Signaling Technology, 7009) was added to each IP sample and 2% input sample. To elute chromatin from the antibody/Protein G Magnetic Beads, samples were incubated for 30 min at 65°C with gentle vortexing (1200 rpm). Supernatant was collected, and reverse cross-linking was performed to all samples including the 2% input by adding 6 µl of 5 M NaCl and 2 µl of Proteinase K, followed by a 2-hour to overnight incubation at 65°C. DNA samples were purified with QIAquick PCR Purification Kit and used for qPCR.

Chromatin isolation by RNA precipitation

ChIRP assays were carried out using the Magna ChIRP RNA Interactome Kit (MilliporeSigma) following the manufacturer's instructions. Briefly, 2×10^7 cells were washed with PBS, resuspended in 1% glutaraldehyde in PBS, and incubated at room temperature for 10 min. The cross-linking agent was blocked by the addition of glycine, and the cells were washed with PBS. After resuspension in PBS, the cells were lysed, and genomic DNA was sheared by sonication with a Q800R Sonicator (Qsonica). Cell lysates were then incubated with odd and even pools of 24 AST-specific biotinylated probes (ChIRP Probe Designer, www.biosearchtech.com; table S2) or *lacZ*-specific biotinylated probes as a negative control (provided in the kit). Oligonucleotides were then assigned to the "odd" or "even" pool in an alternate fashion, progressing along the AST map. This strategy allowed us to obtain two pools of oligonucleotides that do not span the overlapping sequences of AST and control for hybridization to off-target sequences with sufficient complementarity, thus generating a "merged" dataset with low noise (*76*, *77*). After 4 hours of hybridization at 37°C with rotation, streptavidin magnetic beads were added, and the hybridization mixture was incubated for an additional 30 min at 37°C. The beads and captured material were collected with a magnet and then washed four times. After resuspension in wash buffer, 90% of beads were used for DNA extraction using reagents provided in the ChIRP kit and the remaining 10% were used for RNA extraction using a miRNeasy Mini Kit (QIAGEN). Last, DNA was analyzed by qPCR using primers specific for Nuc-0, HS, and Nuc-1 (table S4). Results were expressed as enrichment with AST-specific probes over *lacZ*-control pools. RNA was analyzed by RT-qPCR with primers specific for AST and GAPDH (table S4) to evaluate efficiency and specificity of RNA recovery.

Fractionation of whole cell lysates by size exclusion chromatography

JE4-AST cells were washed three times with PBS without Mg^{2+} and Ca^{2+} and lysed in a buffer containing 50 mM tris-HCl (pH 7.5), 120 mM NaCl, 5 mM EDTA, 0.5% NP-40, 50 mM NaF, 0.2 mM Na_3VO_4, 1 mM DTT, and one complete protease cocktail tablet per 50 ml. After incubation on ice for 20 min with gentle vortexing, cell lysates were cleared by centrifugation at 10,000 rpm for 10 min at 4°C. Supernatants were transferred to a clean tube, and protein concentration was determined by Bradford assay (Bio-Rad, Hercules, CA, USA). For size exclusion fractionation, 2.5 to 5.0 mg of total protein was diluted to 1 ml of total volume with chromatography running buffer [0.2 M tris-HCl (pH 7.5), 0.5 M NaCl, and 5% glycerol] and run on a Superose 6 10/300 size exclusion chromatography column (GE Healthcare Bio-Sciences, Uppsala, Sweden) using the ÄKTA Purifier system (GE Healthcare Bio-Sciences, Piscataway, NJ, USA). A quarter-inch gap was introduced to the top of the Superose 6 column to better separate small–molecular weight complexes from fractions, eluting off the far-right side of the chromatogram. After sample injection (using a 1-ml loop), the running buffer was set at a flow rate of 0.3 ml/min, and 0.5 ml of fractions of the flow-through was collected at 4°C for a total of approximately 60 fractions. Total RNA was extracted with an RNeasy Mini Kit (QIAGEN) from 100 µl of every fifth chromatography fraction, and AST copies were measured in 5 µl of purified RNA by RT-qPCR. Given that AST copies peaked at fraction #15, RNA extraction and RT-qPCR were repeated with chromatography fractions 11 to 19.

Identification of AST-binding partners by 4×S1m-mediated enrichment and MS

The AST-4×S1m and SCR-4×S1m constructs were cloned into the pLVX-Puro vector and stably transduced into Lenti-X 293T cells as described above. For RNA pulldown and MS analyses, cells were plated on ten 10-cm^2 plates and allowed to grow to near confluency. Cells were washed with DPBS and irradiated with 400 mJ/cm^2 of energy on ice in an ultraviolet cross-linker (Stratalinker). Nuclei of ultraviolet cross-linked cells were isolated by suspending the cells in a hypotonic buffer [20 mM tris-HCl (pH 7.4), 10 mM NaCl, and 3 mM $MgCl_2$] for 15 min, adding NP-40 to a final concentration of 0.5% and then centrifuging for 10 min at 850*g*. Nuclei were lysed in a lysis buffer [150 mM KCl, 25 mM tris-HCl (pH 7.4), 5 mM EDTA, 5 mM $MgCl_2$, 1% NP-40, 0.5 mM DTT, Roche minitablet protease inhibitor, and RNaseOUT (100 U/ml)], and debris was cleared by centrifugation at 16,000*g*. The supernatant containing the nuclear lysate was then cleared with 50 µl of Avidin Agarose beads (Thermo Fisher Scientific, 20219) to deplete biotin from the lysates (*78*). The cleared lysate was incubated with 150 µl of streptavidin C Dynabeads (Thermo Fisher Scientific) for 4 hours on a rocker at 4°C. The

beads were collected using a magnet and washed three times with a wash buffer of the same composition as the lysis buffer, except KCl was increased to 350 mM. To release the proteins, beads were incubated with RNase A and resuspended in 2× LDS sample buffer. The entire lysates were resolved on a polyacrylamide gel. Segments of the gel from 30 to 150 kDa were excised and analyzed using MS, which was performed at Harvard Medical School Taplin Mass Spectrometry Core. Analysis of the proteomic data was performed using the CRAPome proteomic analysis (*79*), using a threshold of >90% confidence and greater than fourfold enrichment over controls.

Generation of AST mutants

All AST mutants were generated by DNA synthesis (GeneScript), cloned into the lentiviral vector pLVX-Puro (Takara) under the EIF1α promoter, and stably transduced into JE4 cells as previously described (*22*). All experiments were carried out with the bulk population of stably transduced cells obtained under puromycin selection without isolation of single cell–derived clones. The complete sequence of the wild-type AST and of each AST mutant is reported in table S1. To generate the ASTmutA, ASTmutB, ASTmutC, and ASTmutD mutants, we replaced each domain with the same exogenous sequence of 550 nt. To generate the ASTmutCD mutant, domains C and D were replaced with two copies of the 550-nt exogenous sequence placed in tandem. To generate the ASTmut70A, ASTmut70B, and ASTmut70C mutants, we used the first 70 nt of the exogenous sequence above. The ASTmutU3 mutant was generated by replacing the Y1 (residues 114 to 137 of the full-length AST; 83% pyrimidine content) and Y2 motifs (residues 184 to 212 of the full-length AST; 86% pyrimidine content) with sequences of alternating purine-pyrimidine residues (50 and 48% pyrimidine content, respectively) of equal length.

In silico identification of G-quadruplex motif in AST and structural modeling of wild-type and mutant AST sequences

To identify potential G-quadruplex motifs in AST, we used the open-source web-based tool QGRS Mapper (https://bioinformatics.ramapo.edu/QGRS/index.php) (*33*). The search was performed using the following parameters: maximum length, 30 nt; minimum tetrads, 3 ($G_3N_{y1}G_3N_{y2}G_3N_{y3}G_3$); loop size, 0 to 30. The server identified a QGRS of 28 nt at positions 1048 to 1075 (domain B) of AST with a G-score of 38 (fig. S6). A second 26-nt QGRS with a G-score of 38 was found that overlaps the first and covers positions 1057 to 1082 of AST (fig. S6). The maximum G-score possible is 105 with a $G_6TG_6TG_6TG_6$. The 70-nt substitution in ASTmut70B encompasses both sequences. QGRS Mapper did not identify any G-quadruplex motif in ASTmut70B.

Statistical analyses

Statistical analyses were performed with the Wilcoxon matched-pairs signed-rank test and parametric paired and unpaired *t* tests using Prism GraphPad 10.

Supplementary Materials

The PDF file includes:

Figs. S1 to S10
Tables S1 to S5
Legend for data S1

Other Supplementary Material for this manuscript includes the following:

Data S1

REFERENCES AND NOTES

1. A. R. W. Schroder, P. Shinn, H. Chen, C. Berry, J. R. Ecker, F. Bushman, HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* **110**, 521–529 (2002).
2. Y. Han, K. Lassen, D. Monie, A. R. Sedaghat, S. Shimoji, X. Liu, T. C. Pierson, J. B. Margolick, R. F. Siliciano, J. D. Siliciano, Resting $CD4^+$ T cells from human immunodeficiency virus type 1 (HIV-1)-infected individuals carry integrated HIV-1 genomes within actively transcribed host genes. *J. Virol.* **78**, 6122–6133 (2004).
3. U. Mbonye, J. Karn, The molecular basis for human immunodeficiency virus latency. *Annu. Rev. Virol.* **4**, 261–285 (2017).
4. E. Verdin, P. Paras Jr., C. Van Lint, Chromatin disruption in the promoter of human immunodeficiency virus type 1 during transcriptional activation. *EMBO J.* **12**, 3249–3259 (1993).
5. J. J. Coull, F. Romerio, J. M. Sun, J. L. Volker, K. M. Galvin, J. R. Davie, Y. Shi, U. Hansen, D. M. Margolis, The human factors YY1 and LSF repress the human immunodeficiency virus type 1 long terminal repeat via recruitment of histone deacetylase 1. *J. Virol.* **74**, 6790–6799 (2000).
6. R. Pearson, Y. K. Kim, J. Hokello, K. Lassen, J. Friedman, M. Tyagi, J. Karn, Epigenetic silencing of human immunodeficiency virus (HIV) transcription by formation of restrictive chromatin structures at the viral long terminal repeat drives the progressive entry of HIV into latency. *J. Virol.* **82**, 12291–12303 (2008).
7. D. M. Margolis, Latency reversal and clearance of persistent HIV infection. *Methods Mol. Biol.* **2407**, 375–389 (2022).
8. L. Mori, S. T. Valente, Cure and long-term remission strategies. *Methods Mol. Biol.* **2407**, 391–428 (2022).
9. G. Lehrman, I. B. Hogue, S. Palmer, C. Jennings, C. A. Spina, A. Wiegand, A. L. Landay, R. W. Coombs, D. D. Richman, J. W. Mellors, J. M. Coffin, R. J. Bosch, D. M. Margolis, Depletion of latent HIV-1 infection in vivo: A proof-of-concept study. *Lancet* **366**, 549–555 (2005).
10. N. M. Archin, J. J. Eron, S. Palmer, A. Hartmann-Duff, J. A. Martinson, A. Wiegand, N. Bandarenko, J. L. Schmitz, R. J. Bosch, A. L. Landay, J. M. Coffin, D. M. Margolis, Valproic acid without intensified antiviral therapy has limited impact on persistent HIV infection of resting CD4+ T cells. *AIDS* **22**, 1131–1135 (2008).
11. J. H. Elliott, F. Wightman, A. Solomon, K. Ghneim, J. Ahlers, M. J. Cameron, M. Z. Smith, T. Spelman, J. M. Mahon, P. Velayudham, G. Brown, J. Roney, J. Watson, M. H. Prince, J. F. Hoy, N. Chomont, R. Fromentin, F. A. Procopio, J. Zeidan, S. Palmer, L. Odevall, R. W. Johnstone, B. P. Martin, E. Sinclair, S. G. Deeks, D. J. Hazuda, P. U. Cameron, R.-P. Sékaly, S. R. Lewin, Activation of HIV transcription with short-course vorinostat in HIV-infected patients on suppressive antiretroviral therapy. *PLOS Pathog.* **10**, e1004473 (2014).
12. T. A. Rasmussen, M. Tolstrup, C. R. Brinkmann, R. Olesen, C. Erikstrup, A. Solomon, A. Winckelmann, S. Palmer, C. Dinarello, M. Buzon, M. Lichterfeld, S. R. Lewin, L. Østergaard, O. S. Søgaard, Panobinostat, a histone deacetylase inhibitor, for latent-virus reactivation in HIV-infected patients on suppressive antiretroviral therapy: A phase 1/2, single group, clinical trial. *Lancet HIV* **1**, e13–e21 (2014).
13. J. Friedman, W. K. Cho, C. K. Chu, K. S. Keedy, N. M. Archin, D. M. Margolis, J. Karn, Epigenetic silencing of HIV-1 by the histone H3 lysine 27 methyltransferase enhancer of zeste 2. *J. Virol.* **85**, 9078–9089 (2011).
14. K. Nguyen, B. Das, C. Dobrowolski, J. Karn, Multiple histone lysine methyltransferases are required for the establishment and maintenance of HIV-1 latency. *MBio* **8**, e00133-17 (2017).
15. F. Lu, U. Zankharia, O. Vladimirova, Y. Yi, R. G. Collman, P. M. Lieberman, Epigenetic landscape of HIV-1 infection in primary human macrophage. *J. Virol.* **96**, e0016222 (2022).
16. R. Margueron, D. Reinberg, The polycomb complex PRC2 and its mark in life. *Nature* **469**, 343–349 (2011).
17. C. Davidovich, T. R. Cech, The recruitment of chromatin modifiers by long noncoding RNAs: Lessons from PRC2. *RNA* **21**, 2007–2022 (2015).
18. R. Li, R. Sklutuis, J. L. Groebner, F. Romerio, HIV-1 natural antisense transcription and its role in viral persistence. *Viruses* **13**, 795 (2021).
19. N. L. Michael, M. T. Vahey, L. d'Arcy, P. K. Ehrenberg, J. D. Mosca, J. Rappaport, R. R. Redfield, Negative-strand RNA transcripts are produced in human immunodeficiency virus type 1-infected cells and patients by a novel promoter downregulated by Tat. *J. Virol.* **68**, 979–987 (1994).
20. M. Kobayashi-Ishihara, M. Yamagishi, T. Hara, Y. Matsuda, R. Takahashi, A. Miyake, K. Nakano, T. Yamochi, T. Ishida, T. Watanabe, HIV-1-encoded antisense RNA suppresses viral replication for a prolonged period. *Retrovirology* **9**, 38 (2012).
21. K. Bentley, N. Deacon, S. Sonza, S. Zeichner, M. Churchill, Mutational analysis of the HIV-1 LTR as a promoter of negative sense transcription. *Arch. Virol.* **149**, 2277–2294 (2004).
22. J. C. Zapata, F. Campilongo, R. A. Barclay, C. DeMarino, M. D. Iglesias-Ussel, F. Kashanchi, F. Romerio, The human immunodeficiency virus 1 ASP RNA promotes viral latency by recruiting the polycomb repressor complex 2 and promoting nucleosome assembly. *Virology* **506**, 34–44 (2017).

23. S. Saayman, A. Ackley, A.-M. Turner, M. Famiglietti, A. Bosque, M. Clemson, V. Planelles, K. V. Morris, An HIV-encoded antisense long noncoding RNA epigenetically regulates viral transcription. *Mol. Ther.* **22**, 1164–1175 (2014).
24. W. L. Yu, J. J. Yao, Z. Z. Xie, Y. J. Huang, S. Xiao, lncRNA PRNCR1 rs1456315 and CCAT2 rs6983267 polymorphisms on 8q24 associated with lung cancer. *Int. J. Gen. Med.* **14**, 255–266 (2021).
25. Y. Zheng, M. Wang, S. Wang, P. Xu, Y. Deng, S. Lin, N. Li, K. Liu, Y. Zhu, Z. Zhai, Y. Wu, Z. Dai, G. Zhu, lncRNA MEG3 rs3087918 was associated with a decreased breast cancer risk in a Chinese population: A case-control study. *BMC Cancer* **20**, 659 (2020).
26. A. Blank-Giwojna, A. Postepska-Igielska, I. Grummt, lncRNA *KHPS1* activates a poised enhancer by triplex-dependent recruitment of epigenomic regulators. *Cell Rep.* **26**, 2904–2915.e4 (2019).
27. J. Ferrer, N. Dimitrova, Transcription regulation by long non-coding RNAs: Mechanisms and disease relevance. *Nat. Rev. Mol. Cell Biol.* **25**, 396–415 (2024).
28. Z. Zhou, K. E. Giles, G. Felsenfeld, DNA·RNA triple helix formation can function as a *cis*-acting regulatory mechanism at the human β-*globin* locus. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 6130–6139 (2019).
29. A. Hao, Y. Wang, D. B. Stovall, Y. Wang, G. Sui, Emerging roles of lncRNAs in the EZH2-regulated oncogenic network. *Int. J. Biol. Sci.* **17**, 3268–3280 (2021).
30. X. Wang, K. J. Goodrich, A. R. Gooding, H. Naeem, S. Archer, R. D. Paucek, D. T. Youmans, T. R. Cech, C. Davidovich, Targeting of polycomb repressive complex 2 to RNA by short repeats of consecutive guanines. *Mol. Cell* **65**, 1056–1067.e5 (2017).
31. X. Wang, K. J. Goodrich, E. G. Conlon, J. Gao, A. H. Erbse, J. L. Manley, T. R. Cech, C9orf72 and triplet repeat disorder RNAs: G-quadruplex formation, binding to PRC2 and implications for disease mechanisms. *RNA* **25**, 935–947 (2019).
32. K. Lyu, E. Y.-C. Chow, X. Mou, T.-F. Chan, C. K. Kwok, RNA G-quadruplexes (rG4s): Genomics and biological functions. *Nucleic Acids Res.* **49**, 5426–5450 (2021).
33. O. Kikin, L. D'Antonio, P. S. Bagga, QGRS Mapper: A web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Res.* **34**, W676–W682 (2006).
34. C. Ciferri, G. C. Lander, A. Maiolica, F. Herzog, R. Aebersold, E. Nogales, Molecular architecture of human polycomb repressive complex 2. *eLife* **1**, e00005 (2012).
35. M. Jazurek, A. Ciesiolka, J. Starega-Roslan, K. Bilinska, W. J. Krzyzosiak, Identifying proteins that bind to specific RNAs - Focus on simple repeat expansion diseases. *Nucleic Acids Res.* **44**, 9050–9070 (2016).
36. K. Daneshvar, M. B. Ardehali, I. A. Klein, F. K. Hsieh, A. J. Kratkiewicz, A. Mahpour, S. O. L. Cancelliere, C. Zhou, B. M. Cook, W. Li, J. V. Pondick, S. K. Gupta, S. P. Moran, R. A. Young, R. E. Kingston, A. C. Mullen, lncRNA DIGIT and BRD3 protein form phase-separated condensates to regulate endoderm differentiation. *Nat. Cell Biol.* **22**, 1211–1222 (2020).
37. H. Rafati, M. Parra, S. Hakre, Y. Moshkin, E. Verdin, T. Mahmoudi, Repressive LTR nucleosome positioning by the BAF complex is required for HIV latency. *PLOS Biol.* **9**, e1001206 (2011).
38. F. Romerio, M. N. Gabriel, D. M. Margolis, Repression of human immunodeficiency virus type 1 through the novel cooperation of human factors YY1 and LSF. *J. Virol.* **71**, 9375–9382 (1997).
39. S. R. Jefferys, S. D. Burgos, J. J. Peterson, S. R. Selitsky, A.-M. W. Turner, L. I. James, Y.-H. Tsai, A. R. Coffey, D. M. Margolis, J. Parker, E. P. Browne, Epigenomic characterization of latent HIV infection identifies latency regulating transcription factors. *PLOS Pathog.* **17**, e1009346 (2021).
40. S. E. Kauder, A. Bosque, A. Lindqvist, V. Planelles, E. Verdin, Epigenetic regulation of HIV-1 latency by cytosine methylation. *PLOS Pathog.* **5**, e1000495 (2009).
41. Y. Markaki, J. G. Chong, Y. Wang, E. C. Jacobson, C. Luong, S. Y. X. Tan, D. Maestrini, A. Banerjee, B. A. Mistry, I. Dror, F. Dossin, J. Schöneberg, E. Heard, M. Guttman, T. Chou, K. Plath, *Xist* nucleates local protein gradients to propagate silencing across the X chromosome. *Cell* **184**, 6174–6192.e32 (2021).
42. S. H. Ou, F. Wu, D. Harrich, L. F. Garcia-Martinez, R. B. Gaynor, Cloning and characterization of a novel cellular protein, TDP-43, that binds to human immunodeficiency virus type 1 TAR DNA sequence motifs. *J. Virol.* **69**, 3584–3596 (1995).
43. F. Romerio, Origin and functional role of antisense transcription in endogenous and exogenous retroviruses. *Retrovirology* **20**, 6 (2023).
44. V. Pelechano, L. M. Steinmetz, Gene regulation by antisense transcription. *Nat. Rev. Genet.* **14**, 880–893 (2013).
45. S. Legewie, D. Dienst, A. Wilde, H. Herzel, I. M. Axmann, Small RNAs establish delays and temporal thresholds in gene expression. *Biophys. J.* **95**, 3232–3238 (2008).
46. U. Dühring, I. M. Axmann, W. R. Hess, A. Wilde, An internal antisense RNA regulates expression of the photosynthesis gene *isiA*. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 7054–7058 (2006).
47. P. Mehta, S. Goyal, N. S. Wingreen, A quantitative comparison of sRNA-based and protein-based gene regulation. *Mol. Syst. Biol.* **4**, 221 (2008).
48. C. Patrat, J. F. Ouimette, C. Rougeulle, X chromosome inactivation in human development. *Development* **147**, dev183095 (2020).
49. X. Zhou, X. Han, A. Wittfeldt, J. Sun, C. Liu, X. Wang, L. M. Gan, H. Cao, Z. Liang, Long non-coding RNA ANRIL regulates inflammatory responses as a novel component of NF-κB pathway. *RNA Biol.* **13**, 98–108 (2016).
50. R. Li, I. Caico, Z. Xu, M. S. Iqbal, F. Romerio, Epigenetic regulation of HIV-1 sense and antisense transcription in response to latency-reversing agents. *Noncoding RNA* **9**, 5 (2023).
51. G. Ma, J. I. Yasunaga, K. Shimura, K. Takemoto, M. Watanabe, M. Amano, H. Nakata, B. Liu, X. Zuo, M. Matsuoka, Human retroviral antisense mRNAs are retained in the nuclei of infected cells for viral persistence. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2014783118 (2021).
52. J. L. Rinn, H. Y. Chang, Genome regulation by long noncoding RNAs. *Annu. Rev. Biochem.* **81**, 145–166 (2012).
53. T. Srinivas, E. Siqueira, S. Guil, Techniques for investigating lncRNA transcript functions in neurodevelopment. *Mol. Psychiatry* **29**, 874–890 (2024).
54. L. Li, H. Y. Chang, Physiological roles of long noncoding RNAs: Insight from knockout mice. *Trends Cell Biol.* **24**, 594–602 (2014).
55. J. S. Mattick, P. P. Amaral, P. Carninci, S. Carpenter, H. Y. Chang, L. L. Chen, R. Chen, C. Dean, M. E. Dinger, K. A. Fitzgerald, T. R. Gingeras, M. Guttman, T. Hirose, M. Huarte, R. Johnson, C. Kanduri, P. Kapranov, J. B. Lawrence, J. T. Lee, J. T. Mendell, T. R. Mercer, K. J. Moore, S. Nakagawa, J. L. Rinn, D. L. Spector, I. Ulitsky, Y. Wan, J. E. Wilusz, M. Wu, Long non-coding RNAs: Definitions, functions, challenges and recommendations. *Nat. Rev. Mol. Cell Biol.* **24**, 430–447 (2023).
56. K. Somasundaram, B. Gupta, N. Jain, S. Jana, lncRNAs divide and rule: The master regulators of phase separation. *Front. Genet.* **13**, 930792 (2022).
57. N. Fernandes, J. R. Buchan, RNAs as regulators of cellular matchmaking. *Front. Mol. Biosci.* **8**, 634146 (2021).
58. S. F. Banani, H. O. Lee, A. A. Hyman, M. K. Rosen, Biomolecular condensates: Organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017).
59. A. S. Lyon, W. B. Peeples, M. K. Rosen, A framework for understanding the functions of biomolecular condensates across scales. *Nat. Rev. Mol. Cell Biol.* **22**, 215–235 (2021).
60. X. Huo, L. Ji, Y. Zhang, P. Lv, X. Cao, Q. Wang, Z. Yan, S. Dong, D. Du, F. Zhang, G. Wei, Y. Liu, B. Wen, The nuclear matrix protein SAFB cooperates with major satellite RNAs to stabilize heterochromatin architecture partially through phase separation. *Mol. Cell* **77**, 368–383.e7 (2020).
61. J. H. Lee, R. Wang, F. Xiong, J. Krakowiak, Z. Liao, P. T. Nguyen, E. V. Moroz-Omori, J. Shao, X. Zhu, M. J. Bolt, H. Wu, P. K. Singh, M. Bi, C. J. Shi, N. Jamal, G. Li, R. Mistry, S. Y. Jung, K. L. Tsai, J. C. Ferreon, F. Stossi, A. Caflisch, Z. Liu, M. A. Mancini, W. Li, Enhancer RNA m6A methylation facilitates transcriptional condensate formation and gene activation. *Mol. Cell* **81**, 3368–3385.e9 (2021).
62. L. L. Hall, M. Byron, D. M. Carone, T. W. Whitfield, G. P. Pouliot, A. Fischer, P. Jones, J. B. Lawrence, Demethylated HSATII DNA and HSATII RNA foci sequester PRC1 and MeCP2 into cancer-specific nuclear bodies. *Cell Rep.* **18**, 2943–2956 (2017).
63. A. Bhan, S. S. Mandal, lncRNA HOTAIR: A master regulator of chromatin dynamics and cancer. *Biochim. Biophys. Acta* **1856**, 151–164 (2015).
64. C. M. Clemson, J. N. Hutchinson, S. A. Sara, A. W. Ensminger, A. H. Fox, A. Chess, J. B. Lawrence, An architectural role for a nuclear noncoding RNA: *NEAT1* RNA is essential for the structure of paraspeckles. *Mol. Cell* **33**, 717–726 (2009).
65. A. Pandya-Jones, Y. Markaki, J. Serizay, T. Chitiashvili, W. R. Mancia Leon, A. Damianov, C. Chronis, B. Papp, C. K. Chen, R. McKee, X. J. Wang, A. Chau, S. Sabri, H. Leonhardt, S. Zheng, M. Guttman, D. L. Black, K. Plath, A protein assembly mediates *Xist* localization and gene silencing. *Nature* **587**, 145–151 (2020).
66. M. Estevez, R. Li, B. Paul, K. Daneshvar, A. C. Mullen, F. Romerio, B. Addepalli, Identification and mapping of post-transcriptional modifications on the HIV-1 antisense transcript *Ast* in human cells. *RNA* **28**, 697–710 (2022).
67. Y. Zhao, W. Dunker, Y. T. Yu, J. Karijolich, The role of noncoding RNA pseudouridylation in nuclear gene expression events. *Front. Bioeng. Biotechnol.* **6**, 8 (2018).
68. K. S. Rajan, K. Adler, H. Madmoni, D. Peleg-Chen, S. Cohen-Chalamish, T. Doniger, B. Galili, D. Gerber, R. Unger, C. Tschudi, S. Michaeli, Pseudouridines on *Trypanosoma brucei* mRNAs are developmentally regulated: Implications to mRNA stability and protein binding. *Mol. Microbiol.* **116**, 808–826 (2021).
69. P. P. Vaidyanathan, I. AlSadhan, D. K. Merriman, H. M. Al-Hashimi, D. Herschlag, Pseudouridine and N^6-methyladenosine modifications weaken PUF protein/RNA interactions. *RNA* **23**, 611–618 (2017).
70. N. I. Vlachogiannis, M. Sachse, G. Georgiopoulos, E. Zormpas, D. Bampatsias, D. Delialis, F. Bonini, G. Galyfos, F. Sigala, K. Stamatelopoulos, A. Gatsiou, K. Stellos, Adenosine-to-inosine *Alu* RNA editing controls the stability of the pro-inflammatory long noncoding RNA *NEAT1* in atherosclerotic cardiovascular disease. *J. Mol. Cell. Cardiol.* **160**, 111–120 (2021).
71. E. Wanowska, K. Samorowska, M. W. Szcześniak, Emerging roles of long noncoding RNAs in breast cancer epigenetics and epitranscriptomics. *Front. Cell Dev. Biol.* **10**, 922351 (2022).
72. D. G. Dimitrova, L. Teysset, C. Carre, RNA 2′-O-methylation (Nm) modification in human diseases. *Genes (Basel)* **10**, 117 (2019).

73. C. Li, G. Mousseau, S. T. Valente, Tat inhibition by didehydro-Cortistatin A promotes heterochromatin formation at the HIV-1 long terminal repeat. *Epigenetics Chromatin* **12**, 23 (2019).
74. C. F. Kessing, C. C. Nixon, C. Li, P. Tsai, H. Takata, G. Mousseau, P. T. Ho, J. B. Honeycutt, M. Fallahi, L. Trautmann, J. V. Garcia, S. T. Valente, In vivo suppression of HIV rebound by didehydro-Cortistatin A, a "block-and-lock" strategy for HIV-1 treatment. *Cell Rep.* **21**, 600–611 (2017).
75. J. M. Zerbato, G. Khoury, W. Zhao, M. J. Gartner, R. D. Pascoe, A. Rhodes, A. Dantanarayana, M. Gooey, J. Anderson, P. Bacchetti, S. G. Deeks, J. McMahon, M. Roche, T. A. Rasmussen, D. F. Purcell, S. R. Lewin, Multiply spliced HIV RNA is a predictive measure of virus production ex vivo and in vivo following reversal of HIV latency. *EBioMedicine* **65**, 103241 (2021).
76. C. Chu, K. Qu, F. L. Zhong, S. E. Artandi, H. Y. Chang, Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol. Cell* **44**, 667–678 (2011).
77. J. J. Quinn, H. Y. Chang, In situ dissection of RNA functional subunits by domain-specific chromatin isolation by RNA purification (dChIRP). *Methods Mol. Biol.* **1262**, 199–213 (2015).
78. K. Leppek, G. Stoecklin, An optimized streptavidin-binding RNA aptamer for purification of ribonucleoprotein complexes identifies novel ARE-binding proteins. *Nucleic Acids Res.* **42**, e13 (2014).
79. D. Mellacheruvu, Z. Wright, A. L. Couzens, J.-P. Lambert, N. St-Denis, T. Li, Y. V. Miteva, S. Hauri, M. E. Sardiu, T. Y. Low, V. A. Halim, R. D. Bagshaw, N. C. Hubner, A. Al-Hakim, A. Bouchard, D. Faubert, D. Fermin, W. H. Dunham, M. Goudreault, Z.-Y. Lin, B. G. Badillo, T. Pawson, D. Durocher, B. Coulombe, R. Aebersold, G. Superti-Furga, J. Colinge, A. J. R. Heck, H. Choi, M. Gstaiger, S. Mohammed, I. M. Cristea, K. L. Bennett, M. P. Washburn, B. Raught, R. M. Ewing, A.-C. Gingras, A. I. Nesvizhskii, The CRAPome: A contaminant repository for affinity purification–mass spectrometry data. *Nat. Methods* **10**, 730–736 (2013).

Acknowledgments

Funding: This work was supported by National Institutes of Health grant no. R01AI144893 (F.R.), National Institutes of Health grant R01AI120008 (F.R.), American Foundation for AIDS Research 109612-62-RGRL (F.R.), National Institutes of Health grant no. R01DK116999 (A.C.M.), National Institutes of Health grant no. R01MH134389 (F.K.), National Institutes of Health grant R01AI043894 (F.K.), National Institutes of Health grant no. R21AI074410 (F.K.), National Institutes of Health grant no. R21AI078859 (F.K.), National Institutes of Health grant no. R21AI127351 (F.K.), and National Institutes of Health grant no. R01NS099029 (F.K.). **Author contributions:** Conceptualization: R.L., K.D., F.K., and F.R. Methodology: R.L., K.D., M.P., F.K., A.C.M., and F.R. Validation: R.L., K.D., F.K., and F.R. Formal analysis: R.L., K.D., G.I., F.K., and F.R. Investigation: R.L., K.D., M.P., X.J., G.I., M.S.I., and F.K. Resources: R.L., K.D., F.K., A.C.M., and F.R. Data curation: R.L., K.D., F.K., and F.R. Writing—original draft: F.R. Writing—review and editing: R.L., K.D., M.S.I., F.K., A.C.M., and F.R. Visualization: R.L., M.P., F.K., and F.R. Supervision: K.D., F.K., F.K., A.C.M., and F.R. Project administration: F.K. and F.R. Funding acquisition: F.K., A.C.M., and F.R. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 21 November 2024
Accepted 4 April 2025
Published 9 May 2025
10.1126/sciadv.adu8014