Published online 25 December 2019 *Nucleic Acids Research, 2020, Vol. 48, No. 5* **2643–2660**
doi: 10.1093/nar/gkz1181

Tat IRES modulator of *tat* mRNA (TIM-TAM): a conserved RNA structure that controls Tat expression and acts as a switch for HIV productive and latent infection

Georges Khoury[1,2], Charlene Mackenzie[1], Lilia Ayadi[2], Sharon R. Lewin[3,4], Christiane Branlant[2] and Damian F.J. Purcell[1,*]

[1]Department of Microbiology and Immunology, The Peter Doherty Institute for Infection and Immunity - The University of Melbourne, Melbourne, Victoria 3000, Australia, [2]Ingénierie Moléculaire et Physiopathologie Articulaire (IMoPA), UMR7365 CNRS Université Lorraine, Vandoeuvre-lès-Nancy 54505, France, [3]The Peter Doherty Institute for Infection and Immunity, The University of Melbourne and Royal Melbourne Hospital, Melbourne, Victoria 3000, Australia and [4]Department of Infectious Diseases, Alfred Hospital and Monash University, Melbourne, Victoria 3010, Australia

Received May 27, 2019; Revised December 04, 2019; Editorial Decision December 06, 2019; Accepted December 18, 2019

ABSTRACT

Tat protein is essential to fully activate HIV transcription and processing of viral mRNA, and therefore determines virus expression in productive replication and the establishment and maintenance of latent infection. Here, we used thermodynamic and structure analyses to define a highly conserved sequence-structure in *tat* mRNA that functions as Tat IRES modulator of *tat* mRNA (TIM-TAM). By impeding cap-dependent ribosome progression during authentic spliced *tat* mRNA translation, TIM-TAM stable structure impacts on timing and level of Tat protein hence controlling HIV production and infectivity along with promoting latency. TIM-TAM also adopts a conformation that mediates Tat internal ribosome entry site (IRES)-dependent translation during the early phases of infection before provirus integration. Our results document the critical role of TIM-TAM in Tat expression to facilitate virus reactivation from latency, with implications for HIV treatment and drug development.

INTRODUCTION

Despite effective long-term antiretroviral therapy (ART), HIV-1 persists in a latent state as an integrated provirus in long lived and proliferating CD4+ T cells (1–4). Proviral latency is established early following infection of both resting and activated CD4 T cells (5) and constitutes the main barrier against HIV cure. Only a small fraction of the integrated HIV DNA that persists on ART is intact (6,7) and capable of yielding replication-competent virus leading to viral rebound after ART interruption (8,9).

One approach being pursued to eliminate latent infection is to activate or reverse latency to allow for production of HIV proteins or virions leading to viral immune mediated cytopathic clearance of infected cells (10).

The transition from latent to productive HIV infection is largely governed by the activities of the HIV Tat regulatory protein (11,12). Both the 72 amino acids (aa) single coding exon and the 101 aa two coding exon forms of Tat release the positive transcription elongation factor-b (P-TEFb) from an inactive complex and recruit it to the viral promoter, where it assists in phosphorylating the carboxy-terminus of RNA polymerase II (Pol II) creating the super-elongation complex (13,14). Tat also promotes the displacement of the negative elongation factors NELF (Negative Elongation Factor) and DSIF (DRB Sensitivity-Inducing Factor), and the recruitment of nucleosome remodelling complexes (SWI/SNF), which leads to a highly functional promoter (15,16). Moreover, Tat assists in post-transcriptional events required for productive replication including RNA capping (17–19) and splicing (20).

Sequencing of HIV DNA from infected individuals on suppressive ART revealed that the majority of viruses that persist on ART have large deletions and stop codons and are defective (7,21). This pool of defective proviruses (HIV_{def}) is unable to produce infectious virions but can produce 'unspliced' HIV RNA (usHIV-RNA) species in the size range of 0.5–2.6 kb, whilst also retaining appropriate translational open reading frames (ORFs) that potentially code *gag*, *pol*,

*To whom correspondence should be addressed. Tel: +61 3 8344 6753; Fax: +61 3 9347 1540; Email: dfjp@unimelb.edu.au
Present address: Georges Khoury, Division of Microbiology and Immunology, Yerkes National Primate Research Center, Emory University, Atlanta, GA 30329, USA.

env, *rev* or *nef* (22,23). Indeed, some usHIV-RNAs with internal deletions were predicted to combine parts of Gag and Env, or Gag and Nef in novel HIV chimeric proteins. Over time in individuals on ART, there is an accumulation of defective HIV genomes that retain two flanking long terminal repeats (LTRs), and several key splicing sites that allow ongoing expression of viral protein from aberrant RNAs (24) as well as chimeric host-viral mRNAs (24–27).

The production of aberrant mRNAs that translate virus polypeptides can occur through non-canonical translational pathways such as defective ribosome products (DRiPs) (28,29) or through leaky ribosome scanning (30,31). Control of HIV protein expression is also possible through a tight regulation of mRNA processing and translation by RNA structure and RNA–protein interactions. Several non-canonical translational pathways have been described for HIV, including IRESs within the 5′ leader sequence (32,33) and within the Gag ORF (34,35), which ensure successful Gag translation during G2/M phase arrest, oxidative and osmotic stress (36,37). We hypothesised that Tat could be produced at low levels from both intact and HIV_{def} proviruses through an internal ribosome entry site (IRES), and therefore can influence both the establishment and reversal of HIV latency.

Here, we investigate HIV mRNAs for their capacity to promote a pioneer round of Tat expression required for the reversal of viral latency and promotion of the ensuing phases of productive virus replication. We identify a highly conserved structural element underlying the Tat ORF, named TIM-TAM for Tat IRES Modulator of *tat* mRNA and characterize a mechanism whereby TIM-TAM controls the rate of Tat translation through cap- and IRES-dependent mechanisms, and thus shaping the output of virus reactivation from latency.

MATERIALS AND METHODS

Constructs

pNL4-3 was obtained from M. Martin (NIH, Bethesda, MD, USA). pDR-ΔEnvNef-EGFP viral clone harbours a deletion that terminates Env (nt 7037–7619, derived from pHIS-HIV-B (38)) and an EGFP introduced into KpnI site in Nef ORF, which renders the virus as a single-round EGFP virus once pseudotyped with Env (AD8).

Silaa mut and ΔTIM-TAM viral clones were generated by site directed mutagenesis (Supplementary Table S1) of a sub-cloning construct pET28a+::HIV (nt5786-7250), derived from digestion of pNL4-3 with SalI and NheI. For RNA probing, DNA fragments coding for *tat1* and *tat2* mRNAs were PCR amplified from pNL4-3 plasmid using oligonucleotides complementary to the exon junction sequences (Supplementary Figure S4A, Table S1). The amplified tat1 and tat2 DNA fragments were cloned into pcDNA3.1+/– (Invitrogen) and pSP65 (Promega) plasmids cleaved by EcoRI–XbaI/NheI–EcoRI and EcoRI–HindIII, respectively. L3U1ΔD1A3 was generated by site-directed mutagenesis allowing deletion of intronic sequences in the previously described L3U1 construct (39). Tat-lucF reporter was produced by cloning tat1 DNA fragment at EcoRI site in phase with luciferase firefly (lucF), derived from pGL4.13[luc2/SV40] (Promega). pLD-L3U1 plasmid (39) derivatives (ΔTIM-TAM and inverted) were produced by mutagenesis and Sp1–ex2inv–Δ11 construct by insertion of HIV nt5370–5388 (BRU) DNA fragment between KpnI-BamHI allowing testing for splicing activation properties *in vitro* (40). pSP65::gRNA (+IN D116N) was generated by cloning BssHII–EcoRI and EcoRI–PmlI fragments of pDR-ΔEnvNef-EGFP into pSP65::tat2 vector, followed by insertion of the integrase mutation (IN D116N) through SpeI-SalI digestion from pNL4-3 (+IN D116N) vector.

RNA production

For enzymatic and chemical probing, *tat1* and *2* WT or variant RNAs were generated by run-off transcription with T7 (pcDNA3.1::tat1) and Sp6 (pSP65::tat2 or pSP65::gRNA linearised with ClaI) MEGAscript kit (Promega). DNA templates were digested with RQ1 RNase-Free DNase (Promega), the transcripts were then phenol extracted and isopropanol precipitated. G- and A-capped RNAs were produced by *in vitro* transcription using MEGAscript T7 kit, primed with 6 mM cap structure analogues m7G(5′)ppp(5′)G and m7G(5′)ppp(5′)A (New England Biolabs), followed by polyadenylation using the poly(A) tailing kit (Ambion). RNAs were recovered by lithium chloride precipitation then dissolved in MilliQ water. For splicing assays, uniformly labelled (α32P-UTP) capped RNAs were transcribed by T7 (L3U1-PstI) and Sp6 (Sp1-HindIII) RNA polymerase, as described previously (39).

SHAPE analysis

RNA 2D structure analysis of *tat2* mRNA by SHAPE were performed as previously described (41). Briefly, 1 pmol of *in vitro* transcribed *tat2* RNA was probed in 3× folding buffer (333 mM HEPES–KOH pH 8, 333 mM NaCl, 33.3 mM MgCl2) with 1-methyl-7-nitroisatoic anhydride (1M7, 60 mM in DMSO) or DMSO for 4 min at 37°C, then recovered by ethanol precipitation. Primer extension was conducted as described previously with 0.4 μM fluorescently labelled primer (6-FAM or HEX, Sigma-Aldrich, Supplementary Table S1). The dideoxy sequencing reactions were generated using unmodified RNA, labeled primers (PET or NED, Applied Biosystems) and 0.5 mM ddGTP. cDNAs were recovered by ethanol precipitation and separated by capillary electrophoresis (ABI 3130) with LIZ500 size standard. Data was processed using the *QuShape software* (42), and the 2D structure was generated by using the normalized SHAPE reactivities as pseudo-free energy in the *RNAstructure v5.6* program (*b* and *m* parameters were –0.6 and 2.8 kcal/mol, respectively). The maximum allowed distance between paired bases was restrained to less than 300nt. The RNA model was represented using *varna*.

Enzymatic and chemical probing

RNA probing and 2D structure analysis were performed as previously described (39,43). Briefly, 1.7 pmol of *in vitro* transcribed *tat1* (500 ng) and L3U1ΔD1A3 (300 ng) RNAs were probed in the presence of 4.4 mM $MgCl_2$ and 2–5 μg of yeast tRNAs in buffer D (20 mM HEPES–KOH pH 7.9, 100 mM KCl, 0.2 mM EDTA). Prior to chemical

probing, the RNA mixture was pre-incubated for 20 min at 20°C with sodium cacodylate (50 mM) for efficient DMS and Kethoxal modifications, and with sodium borate (50 mM) for CMCT modification. Probing was performed by treatment with 0.5–0.75 U of T1 RNase, 0.5–0.75 U of T2 RNase, 10^{-4}–5×10^{-4} U of RNase V1 for 10 min at 30°C and 10–20 μl of CMCT, 1–2 μl of a 1/4 (V/V) DMS/EtOH solution, 2–10 μl of Kethoxal (37 mg/ml). The reactions were stopped as described previously (44) and immediately followed by phenol extraction. Positions of enzymatic and chemical cleavages were identified by primer extension analyses using AMV RT (Q Biogene) and the 5′-end labelled primers (Supplementary Table S1). Experiments were repeated several times using different batches of RNA. Stable secondary structures showing the best fit with experimental data were identified using *Mfold v3.2* software (45). Probing data were introduced as a constraint in the search.

Virus production and infectivity assays

Viral stocks were generated by transfecting the proviral constructs into HEK 293T cells with Lipofectamine 2000 (Invitrogen) in serum free media (Opti-MEM, Gibco). pCMV-EGFP vector was added to each transfection mixture as a normalizer for transfection efficiency. Supernatants were collected after 72 h, filtered through a 0.45 μm to clear cell debris, then concentrated using microcon centrifugal filter device (30K, Merck Millipore) and stored at –80°C.

Virus titres were quantified by measuring $p24^{CA}$ levels by capture ELISA. Briefly, Nunc MaxiSorp flat-bottom 96-well plates (Thermo Fisher Scientific, Waltham, MA) were pre-coated with anti-p24 sheep antibody D7320 (300 ng/well; Aalto Bio Reagents) overnight at 4°C. Following 1 h blockage with 5% skim milk/PBS, 100 μl of pre-diluted virus stock or standards (AG6054, Aalto Bio Reagents) with 1% HI-FBS/0.1% Triton X-100/PBS were added in duplicates and incubated for 3 h at room temperature. The samples were then incubated for 2 h with anti-p24 mouse antibody BC1071 (EH12E1 clone, 1:2000; Aalto Bio Reagents) in PBS supplemented with 5% skim milk, washed, and HRP-rabbit conjugated anti-mouse IgG antibody (1:3000 in PBS supplemented with 1% sheep serum, 5% skim milk, Tween-20 0.05%; Invitrogen) was added for 1 h. Following washes with PBS–Tween 0.1% then PBS, 100 μl of TMB substrate (Sigma-Aldrich) was added to the wells and incubated for 30 min at room temperature. The colorimetric reaction was stopped by the addition of 50 μl of 1 M H_2SO_4, and the optical density at 450 nm, against a reference of 690 nm, was measured using an ELISA plate reader (Thermo Fisher Scientific). The assay lower limit of detection was determined as 25 pg/ml.

Virus stocks were titrated by performing a 5-fold serial dilutions in quadruplicate in TZM-bl cells (obtained from the NIH AIDS Research and Reagent Program #8129) as previously described (46). Infectivity of the different viral particles was assessed using equal amounts of p24 (∼6 pg per well). Cells were lysed 48 h post-infection and luciferase activity was measured with Britelite Plus (PerkinElmer) using a FLUOstar plate reader (1sec, Gain 4095, BMG Labtech). The presence of infectious particles is indicated by ≥2.5-fold luciferase activity in the infected cells compared to the non-infected cells (6000 LU). For Tat expression, 100 ng of Tat recombinant protein (Diatheva, clade B) or pcDNA3.1::Tat86-Flag (generous gift from D. Harrich) were co-transfected with proviral DNA into the producer cells. The level of protein expression was estimated by western-blot (WB) analysis of total cell extracts, harvested in lysis buffer (50 mM Tris–HCl pH 8, 1 mM EDTA, 150 mM NaCl, 0.1% IgePal) supplemented with protease inhibitors (Roche) for 30 min at 4°C. Equal amounts of total cellular proteins measured with Bradford reagent (Bio-Rad) were resolved by 12.5% SDS–PAGE, transferred onto nitrocellulose membranes (Bio-Rad) and probed with primary polyclonal HIV-1 antibody (generous gift from J. Mak, 1:5000), antibodies directed against Tat (ANT0003 Diatheva 1:2000 – ab43014 1:1000 for 293T cells and 1:20 for PBMCs) and GAPDH (14C10 Cell Signalling, 1:1000). Proteins were detected using goat anti-human and anti-rabbit IgG secondary antibodies conjugated to HRP (1:5,000, Invitrogen) and developed with SuperSignal West Pico Chemiluminescent Substrate (Thermo Fisher Scientifc). Images were visualised using an *MF-ChemiBis 3.2 imaging system (DNR)*.

Primary cell isolation, culture and infection

Peripheral blood mononuclear cells (PBMCs) were purified from healthy donor blood packs (Red Cross Blood Bank, Melbourne, Australia) by Ficoll-Plaque PLUS (GE Healthcare) gradient centrifugation. For PBMC infections, cells were stimulated with 2.5 μg/ml PHA (Remel) for 3 days in RPMI 1640 medium (Gibco) supplemented with 10% FBS and IL-2 (10 U/ml, Roche). Infected cells (20 pg p24 per 10^6 cells) were fed every 3–4 days with fresh medium supplemented with 10U/ml IL-2 (47), and virus production was monitored by measuring $p24^{CA}$ titres in the culture fluid and TCID50 in TZM-bl cells. Cells were also harvested at each media change for WB, US (*pol*) and MS (*tat1*) RNA quantification by ddPCR. The ddPCR mix consisted of 12 μl 2× ddPCR super mix for probes (no dUTP, Bio-Rad); 900 nM of each primer; 250 nM probe (FAM-MGBNFQ, Applied Biosystems, Supplementary Table S1) and up to 7.2 μl of cDNA (50–140 ng) into 24 μl final volume. Ribonuclease P/MRP 30 kDa (RPP30) was used as a reference gene (HEX, Bio-Rad). Following droplets generation (15–18 000 on average), thermal cycling were conducted as follows: 95°C for 10 min, 40 cycles of 94°C for 30 s and 60°C for 60 s, followed by 98°C for 10 min (ramp rate 2°C/s for each step) on a C1000 Touch Thermal cycler (Bio-Rad). The droplets were subsequently read by a QX200 droplet-reader (Bio-Rad) and the data were analysed with *QuantaSoft 1.7.4* software. The limit of detection of our assay was of 0.5 copies/μl. Integrated HIV DNA was also quantified from extracted genomic DNA (AIQamp DNA Mini kit, Qiagen) using a nested real-time Alu-LTR assay, as described previously (48). HIV fragments (nt 5590–6042) were PCR amplified, gel extracted (Marcherey-Nagel) and sequenced by BDV3.1 cycle sequencing reactions (AGD, Department of Pathology, UoM) using primers listed in Supplementary Table S1. Competition assays started with equal amount of virus particles. The virus was cultured for 2 weeks then passaged onto fresh activated PBMCs and maintained in cul-

ture until virus-induced syncytia (day 28). The integrated proviral DNA isolated from infected cells (day 28; output) and input (day 1) virus were sequenced, as described above.

A total of 10^6 Jurkat T cells were infected with 40 and 80 ng of viral particles via spinoculation at 1500 x g for 2 h at 23°C, followed by 72 h of incubation at 37°C.

CD4+ T cells were purified from PBMCs by negative selection using CD4+ T cell isolation kit (Miltenyi Biotec), and cultured in RPMI 1640 medium supplemented with 10% FBS and 30 U/ml IL-2. Cells were activated with anti-CD3/CD28 (coated 1 μg/ml OKT3 – soluble 0.5 μg/ml L293, BD) for 48 h and infected by spinoculation at 1200 g using 90 and 180 ng DuoF (+ Env 92HT593.1; #3077, NIH reagents) per 10^6 cells. At 72 h after infection, cells were washed and stained with Near-IR Live/Dead fixable dead cell staining (Invitrogen), anti-CD4 OKT4 BV650, anti-CD3 UCHT1 A700 according to the manufacturer's instructions. Cells were fixed with 1% formaldehyde and analysed on a Fortessa (BD) flow cytometer.

Resting CD4+ T cells were isolated by magnetic bead depletion (goat anti-mouse IgG microbeads, MACS) and cell sorting using a cocktail of antibodies (49), with purity >95% as assessed by LSRII flow cytometer using anti-CD4 OKT4 FITC or RPA-T4 PE, CD3 UCHT1 A700, HLA-DR TU36 PE-Cy5, CD69 L78 APC (BD Biosciences) and Live/Dead Near-IR. Purified resting CD4+ T cells were treated one day before infection with CCL19 (30 nM, R&D Systems) (49,50), PHA (10 μg/ml) combined with IL-2 (10 U/ml) (50) or left unactivated. Cells were infected with NL4-3 virus (MOI 0.1 infectious particle per cell assessed by TCID50 in TZM-bl cells). The method used for infection of resting CD4+ T cells and reactivation with PMA/PHA (10 nM-10 μg/ml) is summarized in Figure 7D. Virus production was monitored by measuring the RT activity (51) and cell associated US RNAs.

In vitro splicing assays

In vitro splicing assays were performed with HeLa cell nuclear extracts using 10 fmol of uniformly [^{32}P]-labeled L3U1 pre-mRNA per assay with or without the addition of recombinant SRSF2 protein (300 ng), in the conditions previously described (40). Splicing products were fractionated by electrophoresis on a 5% polyacrylamide/8 M urea gels. The experiments were repeated three times using different preparations of pre-mRNA transcript. Splicing efficiencies (MA3/P, amount of Mature/Pre-mRNA) were established by quantifying the radioactivity using a Typhoon 9410 PhosphorImager, the *ImageQuant* Software Version 5.2 (1999) Molecular Dynamics. The fold activation of splicing by SRSF2 was estimated by dividing the MA3/P ratio in the presence of the recombinant protein by that obtained in its absence.

Translation assays

TZM-bl and HeLa cells were transfected respectively with equimolar amounts of G- or A-capped and polyadenylated *tat* and *tat-lucF* RNAs (2.5 μg) using DMRIE-C reagent (1:2.5, Invitrogen). LucRenilla RNAs (1 μg) were co-transfected for normalisation of luciferase activity. Cells were collected 48 h after transfection and divided into two parts for RNA and protein extractions. Translation levels were monitored by a dual-luciferase reporter assay (Promega) using equal amounts of protein lysate (20 μl per condition) on a FLUOstar plate reader. Each assay was carried out in triplicates with a minimum of three independent transfections. For RNA stability assay, cells were treated with actinomycin D (5 μg/ml) for 4 h before collection. Total cellular RNAs were isolated using the TRIzol Reagent (Invitrogen) and treated with 2 U of RQ1 RNase-Free DNase (Promega). One μg of total RNA was reverse-transcribed using Omniscript-reverse transcriptase (Qiagen) and d(T)15 primer following the manufacturers specifications. Real-time PCR was performed on CFX Connect real-time PCR detection system (Bio-Rad) using FAST SYBR Green Master Mix (Life Technologies) and analysed with *CFX Manager 3.0* software. Quantitative PCR (qPCR) analysis of *tat* transcripts was carried out in duplicates using specific primer pairs (Supplementary Table S1). Each sample was normalised on the basis of *GAPDH* mRNA content. When mentioned, the luciferase activities (lucF/lucR) were normalized on *tat* mRNA levels. gRNA electroporation was completed using the Neon transfection system (Invitrogen). For that, 5 x 10^5 TZM-bl cells were resuspended in 100 μl of buffer R and immediately electroporated (2 pulses, 1005 V, 35 ms) with 150 pmol of capped gRNAs. Cells were incubated in RPMI medium for 24 h at 37°C, then lysed with 100 μl passive lysis buffer (Promega) before measuring the luciferase activity using Britelite (PerkinElmer, lucF) and Stop&Glo (Promega, lucR).

Analysis of HIV_{def} proviruses

Sequences of defective proviruses were obtained by single-genome amplification and sequencing from four previously published studies (6,7,21,22). GenBank accession numbers for HIV_{def} proviruses used in our analysis were reported in (21) (355 sequences, with accession numbers KX505390 to KX505744), (7) (220 sequences, KF526120-KF526339), (22) (208 sequences, KU677989–KU678196) and (6) (448 sequences, KY766193–KY778681 and KY766150–KY766212). Acute, chronic, non-induced, induced, PVL >40 and PVL <40 were defined as previously described in the corresponding studies.

Sequencing reads were first aligned and compared to the reference genome HXB2 using *HIVAlign* (MAFFT) (52) on the Los Alamos database. Clones that did not align to *tat-exon4* (Tat-) were excluded from further analyses. Hypermutations (indels, syn or non-syn, codon stops) were determined using the Tat+ clones and the Los Alamos *Highlighter* (53) and *SNAP v2.1.1* (54) algorithms. Heatmaps were generated using *GraphPad Prism 7*, while sequence conservation and probability (frequency of each nucleic acid at that position) were generated using *WebLogo* (55). Secondary structure propensity profiles of RNA sequences at single nucleotide resolution were determined with *CROSS* (56). A score in the range (0,1) means that the nucleotide is predicted to be paired, while a score in the range (–1,0) indicates that the nucleotide is single stranded.

Statistical analysis

Two-tailed Student's *t*-test, one- and two-way ANOVA with a Sidak–Bonferroni correction method were used for comparisons between populations using *GraphPad Prism 7*. $P < 0.05$ was considered significant.

RESULTS

HIV_{def} proviruses harbour intact Tat coding region

Patient derived cells containing HIV intact and defective proviruses transcribe viral pre-mRNAs *in vitro* and *ex vivo* containing active and defective splice donors (SD) and acceptors (SA) whose use during mRNA splicing leads to the production of RNA splice variants (23). We questioned whether aberrant spliced HIV mRNAs could encode the essential HIV Tat protein needed to restore productive viral replication. To examine the potential for HIV_{def} to produce Tat protein, we first sought to quantify the level of intact Tat ORF in the HIV proviral DNA from four previous studies of HIV infected individuals using limiting dilution then sequencing of full-length proviruses (6,7,21,22). The proportion of Tat+ clones that harbor HIV exon 4 and Tat ORF was low (32.33%, 398 clones out of 1231, Figure 1A). While the majority of these proviruses presented a mutated Tat start codon or internal stop codons, only a small subset of clones presented intact Tat ORFs and a start codon (20–43.5%) (Figure 1A). As previously reported, the proportion of hypermutated proviruses ($HIV_{hypermut}$) accumulates rapidly during acute infection (6,7,21,22), simultaneous with the higher frequency detection of insertions-deletions (indels) and pre-mature stop codons (Supplementary Figure S1A, B). In addition, the non-induced clones and patients with low proviral loads (PVL<40 copies per milliliter) presented a higher rate of indels when compared to those from the induced and high PVL (PVL > 40 copies per milliliter) clones. Most indels were detected in the C-terminal region of Tat that has a low impact on Tat trans-activation (57), coinciding with a higher proportion of non-synonymous mutations (non-syn) at the C-terminus (Figure 1B, Supplementary Figure S1C). On the other hand, the N-terminus segment of Tat needed for trans-activation activity of patient provirus had a relative absence of indels and higher frequency of synonymous (syn) mutations (Figure 1B, Supplementary Figure S1D), implying a positive selection of Tat residues encoded by *tat-exon4* (Figure 2A, highlighted in gray). These data suggest a relative conservation of the known Tat trans-activation domain within defective proviruses and mRNAs, and potentially any embedded mRNA elements.

Conserved stem–loop structure in Tat coding region

HIV proviruses use alternative splice sites to produce mRNA transcripts that code for essential viral proteins, such as Tat (58). However, HIV 3′ splice sites (3′ss) are suboptimal and are frequently included into stable RNA structures (59). However, while HIV_{def} proviruses can be transcribed, it remains unclear whether these could be spliced efficiently. To analyse the potential splicing of *tat* mRNAs from HIV_{def} transcripts, we used the approach of computational recognition of secondary structures (CROSS) to profile *tat* RNA structure compared to HXB2 reference strain at a single nucleotide resolution (56). Despite presenting conserved A3 splice site (SA3) and polypyrimidine sequences compared to HXB2 SA3 consensus (Supplementary Figure S1E), major differences were observed in ribonucleotides pairing into *tat* mRNA structure overlapping the *tat-exon4* (highlighted in grey, Supplementary Figure S1F). These changes were predominantly in the N-terminal region of *tat-exon4* that are unique to Tat coding mRNA, but were not apparent in distal HIV coding regions that overlap with *rev* and *nef* mRNAs (Figure 2A). This warranted a deeper investigation of the role of potential conserved mRNA structure in the different multiply spliced RNAs.

We compared the relative level of predicted conserved RNA structure (sfc) and nucleotide sequence (nfc) within multiply spliced mRNAs of 34 replication competent HIV strains from different viral subtypes (Figure 2B). This analysis indicates the RNA regions common between 2 kb *tat*, *rev* and *nef* mRNAs share both conserved predicted folded structure and sequence. As expected, the 5′ and 3′ untranslated regions (UTRs) contained a greater proportion of conserved RNA structure than primary sequence. These terminal regions contain conserved RNA structures with known functions, such as trans-activation response (TAR) RNA domain known for Tat-binding. The TAR displays an unpaired RNA motif embedded into stable paired RNA structure in this analysis (Figure 2B). In contrast, the splice donor (D1) and acceptors (A3, A4a and A5) presented a greater proportion of conserved sequences (Figure 2B). This analysis revealed another interesting and unique stem–loop RNA element underlying the N-terminal region of the Tat ORF (nt 5440–5455 HXB2, Figure 2) fulfilling the features for a functional element with highly conserved RNA structure. We called this element TIM-TAM (for Tat IRES Modulator of *tat* mRNA), as it was included in the 2kb mRNA for *tat*, but not *rev* and *nef* mRNA. However, this element was also present in the previously reported HIV genomic RNA structure (60) (Figure 2C). RNA structure prediction and thermodynamic analysis revealed folding of TIM-TAM into a stem–loop structure displaying an apical tetraloop. Furthermore, comparing 2504 HIV isolates in the Los Alamos HIV sequence database revealed this element had low-level sequence changes within the stem (Figure 2D and Supplementary Figure S2A). Most mutations were detected within the tetraloop and therefore did not have a drastic effect on the predicted thermodynamic stability of this stem–loop (Supplementary Figure S2B, C). Altogether, our results indicate highly conserved stem–loop TIM-TAM structure underlying the Tat ORF, and this is also likely to exist in aberrant mRNAs produced from HIV_{def} proviruses that accumulate during ART.

TIM-TAM activates Tat IRES translation, but restricts cap-dependent translation

Due to the close proximity to the A3 3′ss, we first examined the role of this stem–loop in mRNA splicing *in vitro* using the L3U1 construct (40), which harbours *tat-exon1* and *4*. Deletion or inversion of TIM-TAM in this reporter construct did not significantly affect the splicing at site

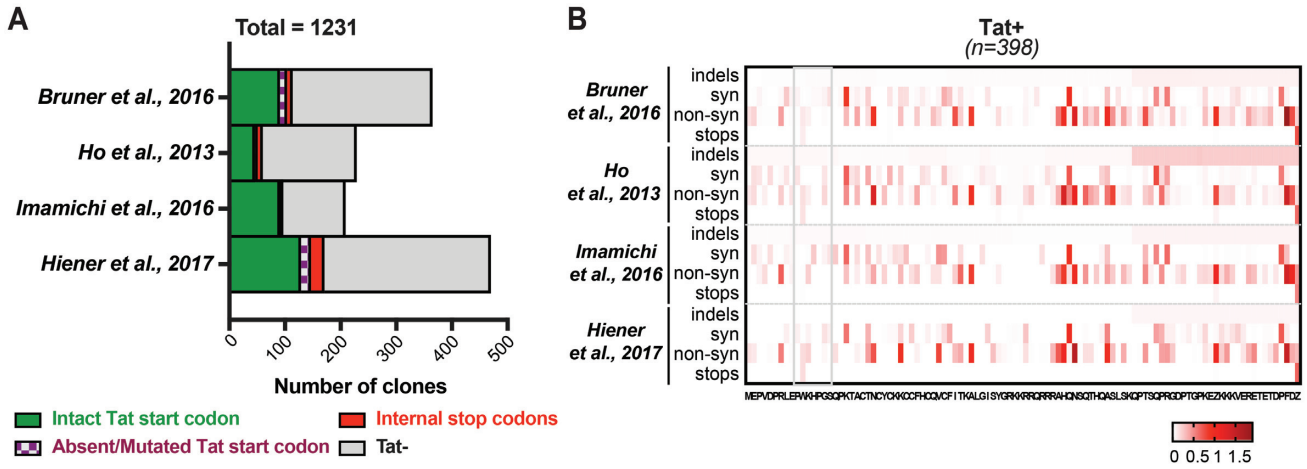

Figure 1. HIV_{def} proviruses can harbour intact Tat coding region. (**A**) Number of Tat+ clones from ART treated individuals. Fraction of hypermutated proviruses that contain an intact Tat start codon are shown in green. Sequences with absent or mutated Tat start codon as well as an internal stop are shown in grey, purple and red, respectively. (**B**) Hypermutations (indels, synonymous or non-synonymous mutations and stop codons) of Tat+ clones were determined using the Los Alamos *SNAP* algorithm (Korber B. *et al.*, 2000).

A3 (MA3/P 0.2–0.31) or activation by the splicing factor SRSF2, commonly known as SC35 (fold activation 15.6 or 14.5 instead of 16.7, Supplementary Figure S3A). Because RNA structures also play an essential role in modulating protein translation and especially IRES activity, we next examined the ability of this conserved sequence-structure to modulate Tat production through the IRES mechanism. We used *in vitro* transcribed and ribosome-scanning active G-capped or inactive A-capped polyadenylated *tat* mRNAs to test for IRES translational activity. RNA transfection of G- and A-capped RNA into TZM-bl cells allowed us to distinguish between cap- and IRES-dependent translation of Tat through expression of Luciferase Firefly (LucF) from an integrated reporter provirus (Figure 3A). Our results confirmed the production of Tat protein through an IRES-mediated mechanism, although at a 30% efficiency of the G-capped mRNA substrate for ribosome scanning. IRES-expressed Tat was functional and transactivated transcription at the HIV 5′LTR (Figure 3A).

To further investigate the role of the TIM-TAM stem–loop in Tat IRES translation, we inserted silent point mutations (Silaa mut) that significantly disrupted the RNA structure as predicted by mfold ($\Delta G = -2.6$ kcal/mol for Silaa mut and +3.1 kcal/mol for ΔTIM-TAM versus −7.7 kcal/mol for the WT sequence, Figure 3B), however did not affect the amino acid sequence of Tat protein. RNA structure mapping of Silaa and ΔTIM-TAM mutants using SHAPE reactivity revealed the introduced changes had a significant impact on *tat-exon4* SHAPE reactivity and RNA structure compared to the WT RNA (Figure 3C). These RNA structure studies show a dominant role for the TIM-TAM stem–loop in overall *tat* mRNA folding.

When *in vitro* transcribed A-capped and polyadenylated Tat-luciferase Firefly RNA was transfected into reporter cells, a significant reduction in Tat expression was observed for the Silaa mut and ΔTIM-TAM mutants due to a loss in IRES activity, confirming a role of the TIM-TAM stem–loop in Tat IRES-dependent translation (Figure 3D). As both Silaa and ΔTIM-TAM mutations induce drastic changes to *tat* mRNA structure, it is plausible that these effects on Tat translation might be due to changes in the accessibility of ribosomes to Tat initiation codon. While TIM-TAM conferred a cap independent translation initiation activity, silent point mutations that disrupted the RNA structure induced a paradoxical increase in Tat-Firefly luciferase cap-mediated translation (G-cap + pA, Figure 3D). This effect was not observed when we deleted a distal conserved sequence motif, Δ11 (nt 5373–5383 BRU), located in the vicinity of Tat start codon (Supplementary Figure S3C). We have previously shown Δ11 as a binding motif for various SR and hnRNP proteins, such as SRSF2, SRSF5 and hnRNP A1 proteins (61) that play a key role in splicing (40,61) and translational control of *tat1* mRNA (62). These data demonstrate a second role of the stable stem–loop structure in blocking efficient ribosome progression during the cap-dependent translation initiation of *tat* mRNAs. Therefore, TIM-TAM confers moderately efficient cap-independent translation activity for Tat and additionally impedes fully efficient cap-dependent Tat translation.

TIM-TAM required for an authentic *tat* mRNA folding

Due to the strong sequence-structure conservation and role in Tat production, we hypothesised that TIM-TAM may be formed both in unspliced genomic RNA and spliced *tat* mRNAs. To test this, we studied in parallel *tat1* pre-mRNA and mRNA 2D structure, as the surrounding sequences might influence the folding of TIM-TAM. Very similar profiles of enzymatic digestions by RNase T1, T2 (for single-stranded) and V1 (double-stranded), as well as chemical modifications by DMS, CMCT and Kethoxal (for single-stranded) were observed for both RNA. Representative examples of the primer extension analyses of *tat*1 mRNA are provided in Supplementary Figure S5. Furthermore, to test if TIM-TAM interacts with neighboring HIV RNA sequences, a smaller *tat*1 RNA that included only *tat*-exons1/4 and a larger RNA, *tat*2, harboring an additional exon upstream of A3 3′ss (exons1/2/4, nt 4494–

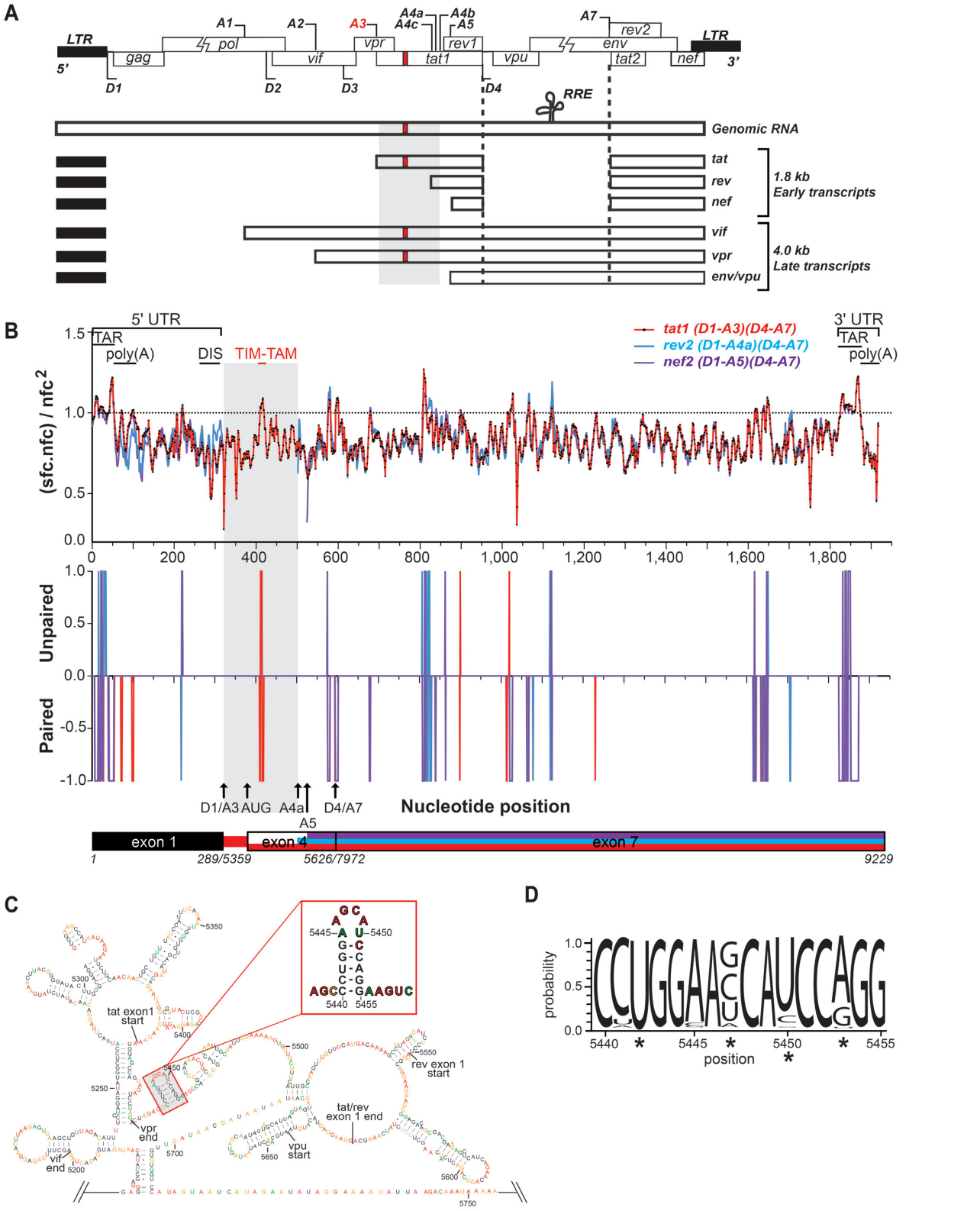

Figure 2. Conserved stem–loop structure in Tat coding region. (**A**) Schematic representation of HIV-1 genome and a sample of the mRNAs produced during early (1.8 kb) and late phases (4 kb) of infection. ORFs are represented by white boxes, while LTRs are in black. HIV-1 genome contains four major donor (D) and 8 acceptor splice sites (A). The *cis* regulatory element TIM-TAM is highlighted in red. (**B**) Secondary structure propensity profiles of multiply spliced mRNAs. Structure (sfc) – sequence (nfc) conservation and pairing probabilities (unpaired, paired) of *tat1*, *rev2* and *nef2* mRNAs are shown with 8 nt window employed for smoothing. Structure conservation over sequence is highlighted by (sfc.nfc)/nfc^2 values above 1, while values below 1 indicate a preferential sequence conservation over structure. Known structured regions in the 5′ and 3′ UTR regions are indicated above. (**C**) Secondary structure of TIM-TAM within the HIV RNA genome determined by (J. Watts *et al.*, 2009). (**D**) Sequence conservation of TIM-TAM (nt 5440–5455, BRU) within 2504 HIV strains on the Los Alamos HIV databases. HXB2 strain was used as a reference.

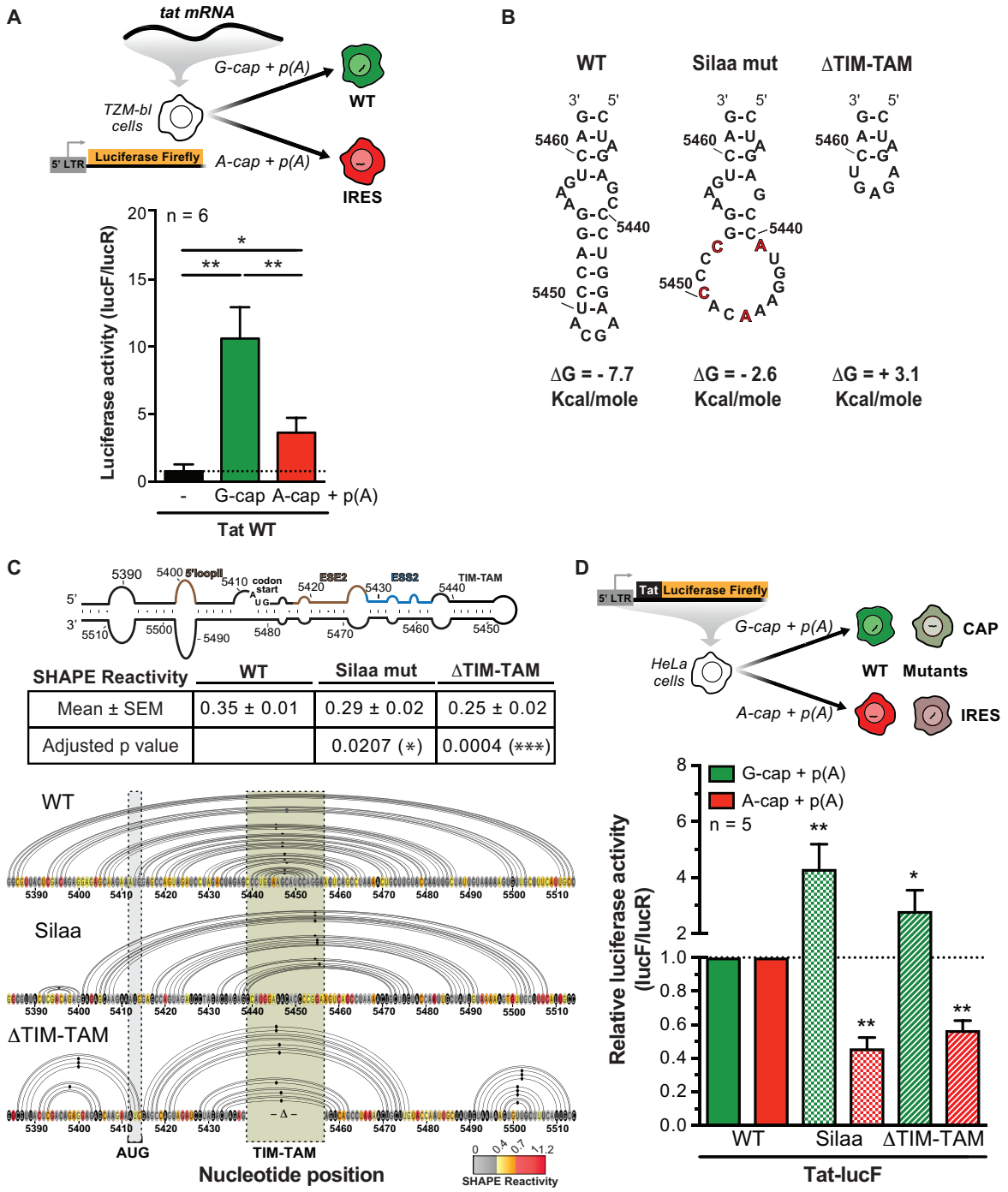

Figure 3. TIM-TAM activates Tat IRES translation. (**A**) Ability of *tat* mRNA to induce Tat production in a cap-independent manner. RNA transfection of TZM-bl cells with G- or A-capped and polyadenylated *tat* mRNAs followed by luminescence assay 48 h post-transfection. Cells were co-transfected with luciferase Renilla RNAs where results are reported as relative luciferase Firefly activity over Renilla (lucF/lucR). Data represent mean ± SEM, $n = 6$. (**B**) WT, silent mutant (Silaa mut) and ΔTIM-TAM RNAs used in this study. Silent point mutations introduced in TIM-TAM are represented in red. The free energies (ΔG in kcal/mol) were calculated using Mfold webserver (37°C in 1 M NaCl). (**C**) Effect of TIM-TAM substitution or deletion on *tat2* mRNA structure. 2D structure of the stem–loop harbouring Tat start codon and TIM-TAM is represented on top of this panel. The *cis* regulatory elements; ESS2 (in blue), ESE2 and 5′loop II (in brown), located within this stem–loop are shown. SHAPE reactivities of WT, Silaa mut and ΔTIM-TAM mRNAs are represented as a function of nucleotide position surrounding TIM-TAM over a window of nt 5384 and 5514 (BRU). Nucleotides are colored by their respective shape reactivities, with highly reactive nucleotides report flexible positions (SHAPE reactivity ≥ 0.7 in red and 0.4 ≤ SHAPE reactivity > 0.7 in orange), while low reactive nucleotides are represented in grey (SHAPE reactivity < 0.4). Arcs indicate highly probable helices. Data represents mean ± SEM. Statistical significance was determined using Friedman test. $^*P < 0.05$; $^{***}P < 0.001$. (**D**) Influence of silent point mutations or deletion of TIM-TAM on Tat cap and IRES dependent translation. HeLa cells were transfected with G- or A-capped and polyadenylated Tat-luciferase Firefly RNAs followed by luminescence assay 48h post-transfection. Cells were co-transfected with Renilla luciferase RNA (G-cap+pA), and results are reported as relative luciferase activity (lucF/lucR) compared to WT RNA, defined as 1. Data represents mean ± SEM, $n = 5$. Statistical significance was determined using two-tailed Student's *t*-test. $^*P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; ns, not significant.

4544 BRU) were studied by limited enzymatic digestion and SHAPE analysis, respectively. Normalised SHAPE reactivities and probing data were then used to determine the secondary structure model for the entire *tat*2 (Figure 4) and *tat*1 (Supplementary Figure S6) mRNA by converting the SHAPE reactivities and enzymatic-chemical modifications to pseudo-energy constraints. In the proposed models (Figure 4A; Supplementary Figures S4 and S6), most of the nucleotides assigned as single-stranded are reactive to 1M7, RNase T1 and T2; while paired nucleotides presented low reactivity to 1M7 and high reactivity to RNase V1. Agreement across various RNA structure probing methods further strengthens the accuracy of TIM-TAM structure obtained from these models. When presented as 2D structures, the 5′ and 3′ UTR regions harbour several positions with low SHAPE reactivity and high pairing probabilities such as TAR, polyA, U5 and DIS stem–loop structures (Figure 4A). Interestingly, TIM-TAM presented the highest pairing probability (> 0.98 vs. 0.9 for 5′ TAR, data not shown). As illustrated in Figure 4B and C, our RNA structure probing demonstrates that TIM-TAM forms identically in *tat*2 and *tat*1 alternative spliced mRNAs the apical part of a long irregular stem–loop structure that harbors Tat start codon. In addition, these data support the thermodynamic dominance of TIM-TAM ensuring formation of the same structure in the full-length viral RNA (Figure 2C) and *tat* mRNAs (Figure 4). This suggests TIM-TAM acts to support Tat translation from full-length genomic RNA (gRNA) through an IRES mechanism. We therefore tested IRES-mediated Tat expression from G- or A-capped *in vitro* transcribed full-length gRNA after electroporation into TZM-bl cells. The gRNA included an integrase mutation, IN(D116N), that blocks provirus integration to model the effects from uncoated virion RNA. The results showed that Tat could be produced efficiently, and to a similar level from G- and A-capped unintegrated gRNAs (Figure 5). In the presence of Silaa mutations, a significant reduction in Tat translation and transactivation was detected for both G- and A-capped Silaa mut gRNAs. Thus, these data suggest the role of the long irregular stem–loop $SLS3_{A3}$ and its apical motif, TIM-TAM, in modulating both gRNA and *tat* mRNA folding to both impede ribosome progression on m7G-capped mRNA, and to facilitate translation of RNA lacking an m7G-cap by a privileged cap independent translation initiation mechanism. We proposed that this mechanism might provide an immediate early expression of Tat to pioneer viral trans-activation functions.

TIM-TAM controls early and late stages of productive infection

Due to the central role of Tat protein in the virus life cycle, we examined the ability of TIM-TAM to modulate virus production and infectivity during both the early and late stages using a pseudotyped envelope (Env) defective HIV provirus, competent for only a single round of infection, pDR-ΔEnv Nef-EGFP. During the productive phase of a single round infection, the presence of Silaa mut within pDR-ΔEnv Nef-EGFP resulted in a significant increase in virion production (Figure 6A, Supplementary Figure S7A). In addition, we observed a significant increase in virus infectivity when TZM-bl cells (Supplementary Figure S7B) and activated PBMCs (Figure 6B) were infected with equal amounts of virus particles. The increase in virus production was coupled with an increase in Tat protein production (Figure 6C). However, as shown in our kinetic studies during a spreading infection in PHA-activated PBMC several days post infection (7–10 dpi), a significant decrease in Silaa mut virus transmission and production was observed compared to WT NL4.3 (Figure 6D). This suggests a role of TIM-TAM in regulating the time to maximal Tat expression both at the early and late stages of the infection (Figure 6E), and thus the level of virus accumulation in spreading infection. To confirm these changes in virus production, we measured the amount of cell-associated unspliced (*pol*) and multiply spliced (*tat1*) RNAs by droplet digital PCR (ddPCR). The quantification of unspliced RNAs indicative of the late stages of infection revealed a decrease in *pol* RNA levels in the presence of Silaa mut and ΔTIM-TAM, while no significant change was observed for *tat*1 mRNA (data not shown). Similar kinetic profiles and unspliced RNA levels were found in all four independent infections conducted using activated PBMCs.

While these changes in virus production were not due to variation in the sequence of the provirus (Supplementary Figure S7C, D), it is plausible that Tat would exert a negative feedback on its own production during the late stage of infection. In fact, overexpression or addition of recombinant Tat protein induced a significant decrease in virus production of NL4.3 WT that to a similar level of Silaa mut virus (Supplementary Figure S7E). These results suggest the role of TIM-TAM in modulating Tat translation in primary lymphocytes by inducing its production in the pioneer rounds of infection through an IRES activity, but followed by a reduction during active viral replication through structure-mediated impedance of cap-dependent ribosome progression.

TIM-TAM controls latent infection and virus reactivation

Tat protein exerts an important role during latent infection, as low levels of the protein or nuclear retention of the multiply spliced RNAs that encode Tat contribute in maintaining the cells in a latent state (63). One the other hand, Tat delivery into latently infected cells can inhibit the establishment of HIV latency (11), and thus induce virus reactivation (64). To investigate the role of TIM-TAM during latency, we used a dual color HIV reporter virus, R7GEmC, that contains an LTR-dependent EGFP and an mCherry constitutively expressed from an EF1α promoter, independently of the HIV 5′LTR activity. This single round HIV reporter distinguishes between productively (eGFP+ mCherry+) and latently (mCherry+) infected cells (65). When infections of Jurkat T cells were carried with the dual colour reporter virus harboring the silent mutation, a drastic reduction in the single mCherry+ population was observed (1.27% for WT versus 0.32% for Silaa mut, Figure 7A), independent of the virus titre used (Supplementary Figure S7F), indicative of a restriction on the establishment of a latent infection in the absence of full TIM-TAM function.

We then tested these viral constructs in primary CD4+ T cells and their ability to generate latent infection. Once

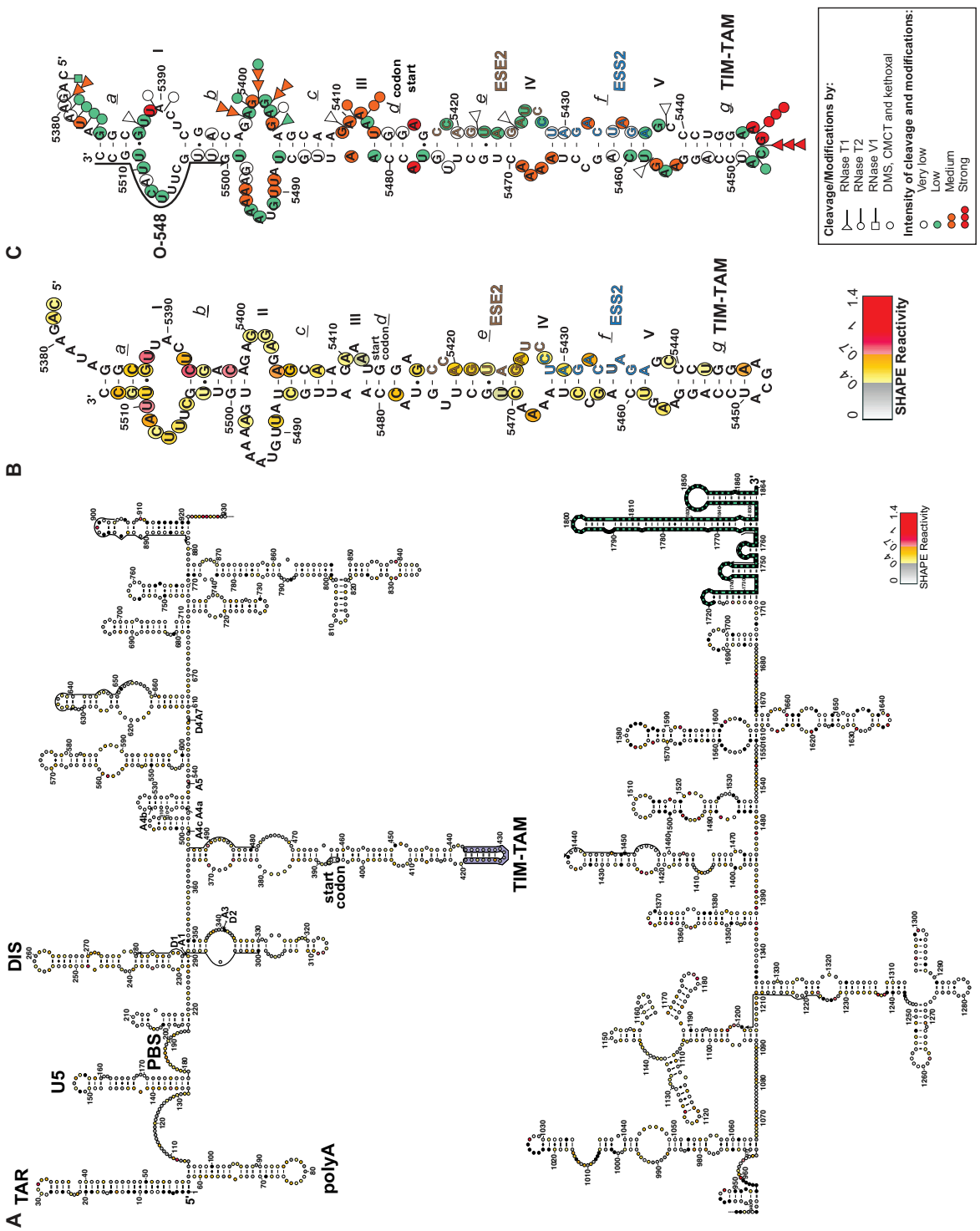

Figure 4. TIM-TAM required for an authentic *tat* mRNA folding. (**A**) 2D mRNA structure of full-length *tat2* mRNA determined by SHAPE. The SHAPE reactivities were superimposed on the 2D model with highly reactive nucleotides report flexible positions (SHAPE reactivity ≥ 0.7 in red), while low reactive nucleotides represent paired nucleotides (SHAPE reactivity < 0.4 in grey). (**B**) Secondary structure model of *tat2* mRNA in the vicinity of Tat start codon (nt 5377–5514) that we established by SHAPE. Nucleotides are colored as in (a), according to their SHAPE reactivity. (**C**) Model of *tat1* mRNA structure around Tat start codon (nt 5377–5514) that we established by chemical (DMS, CMCT, Kethoxal) and enzymatic (RNase T1, T2, V1) modifications. RNase T1, T2 and V1 cleavages are represented by arrows surmounted by triangles, circles and squares, respectively. Nucleotides modified by DMS, CMCT or Kethoxal are circled. Colours of the arrows and circles indicate the yield of cleavage and modifications: white, green, orange and red for very low, low, medium and strong cleavage or modification, respectively. Positions are numbered in reference to HIV-1 BRU (K02013).

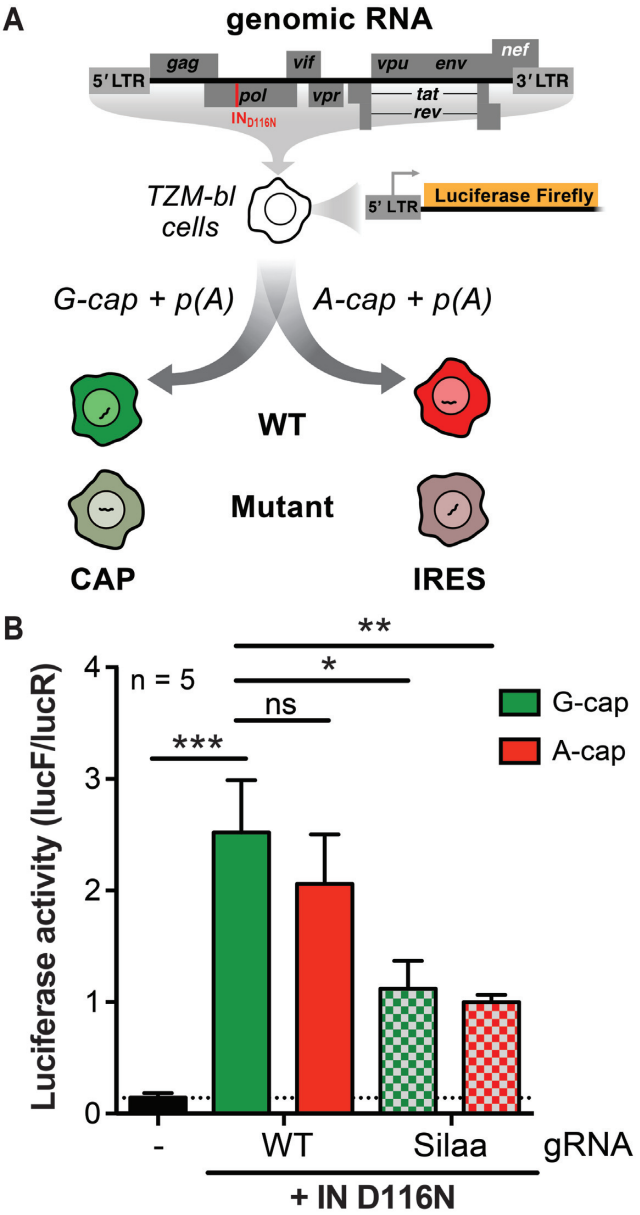

Figure 5. Ability of TIM-TAM to induce Tat translation from full-length genomic RNA. TZM-bl cells were electroporated with capped gRNA (+IN D116N) and luminescence assay performed 24 h post-transfection. The results are reported as relative luciferase activity (lucF/lucR). Bars show mean values and SEM ($n = 5$). *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ns, not significant (two-tailed Student's *t*-test).

more, maximal productive infection was impeded by the full TIM-TAM structure (WT 13.2% versus Silaa 15.5% eGFP + mCherry+), while mutations in TIM-TAM led to a significant reduction in latent infection (WT 0.43% versus Silaa 0.30% mCherry+, Figure 7B). A significant reduction of latent/productive infection ratio was also observed when TIM-TAM was mutated (Figure 7C), confirming a role of TIM-TAM in modulating Tat translation during latent infection, and therefore important in the establishment and maintenance of latency.

We next evaluated the extent of virus reactivation from latently infected cells by phorbol myristate acetate and phytohemagglutinin (PMA/PHA) stimulation in the presence WT or mutated TIM-TAM. This was tested in the CCL19 primary model of latency described previously to display the broadest response to reactivation stimuli compared to other latency models (66). The CCL19 model most closely resembles resting CD4+ (rCD4+) T cells from aviraemic patients (66). In this assay, rCD4+ T cells were pre-treated with the CCL19 chemokine and infected with NL4.3 WT or Silaa (MOI 0.1) (49,50). Four days post-infection, rCD4+ T cells were stimulated with PMA/PHA and virus reactivation was quantified by measuring the RT activity in the culture supernatants, as well as the cell associated US RNA levels (Figure 7D). Analysis of RT activity at day 5 post-infection of activated, resting and CCL19 treated cells revealed minimal virus production in CCL19 treated infected CD4+ T cells, consistent with latent infection (data not shown). Interestingly, the lack of cap-independent translation in the case of Silaa mut induced a significant reduction in virus reactivation in the majority of donors tested. Altogether, these results suggest an essential role of TIM-TAM and Tat IRES translation in controlling latent infection and virus reactivation.

DISCUSSION

We have identified a highly conserved RNA structure, TIM-TAM, underlying the ORF in the first coding exon of Tat that reduces the efficiency of Tat protein translation from conventional *tat* mRNA which pushes the virus towards latency. Our data also support a novel mechanism for translation of Tat from unconventional viral RNAs. TIM-TAM assists in the pioneer round of Tat expression by a cap-independent mechanism from unintegrated 9 kb unspliced genomic RNA promoting establishment of productive viral replication (Figure 8). In addition, this alternative translation mechanism could permit Tat expression from HIV_{def} proviruses retaining the TIM-TAM RNA element and the Tat ORF. Furthermore, other aberrant chimeric host-HIV RNAs containing TIM-TAM may facilitate production of viral peptides that impact upon ongoing immune activation and temporal changes in the structure of the latent HIV reservoir during treated HIV disease (24).

The role of TIM-TAM in modulating Tat expression provides a pathway whereby unconventional sources for HIV RNAs may contribute to viral expression, replication and pathogenesis. Our results support a mechanism to prioritize Tat expression from any Tat ORF-containing RNA arising, even from the unintegrated HIV DNA at the earliest stages of infection. The TIM-TAM would permit Tat expression through canonical and/or cap-independent translation, thereby promoting high viral gene expression and productive replication upon completion of provirus integration. These results complement earlier studies that have demonstrated the ability of unintegrated HIV DNA or HIV_{def} integrase mutants to serve as a template for antigen production in activated and resting CD4+ T cells (67,68), as well as macrophages for as long as 20 to 30 days (69,70).

TIM-TAM functions may also impact on viral rebound after treatment interruption. For example, in patients un-

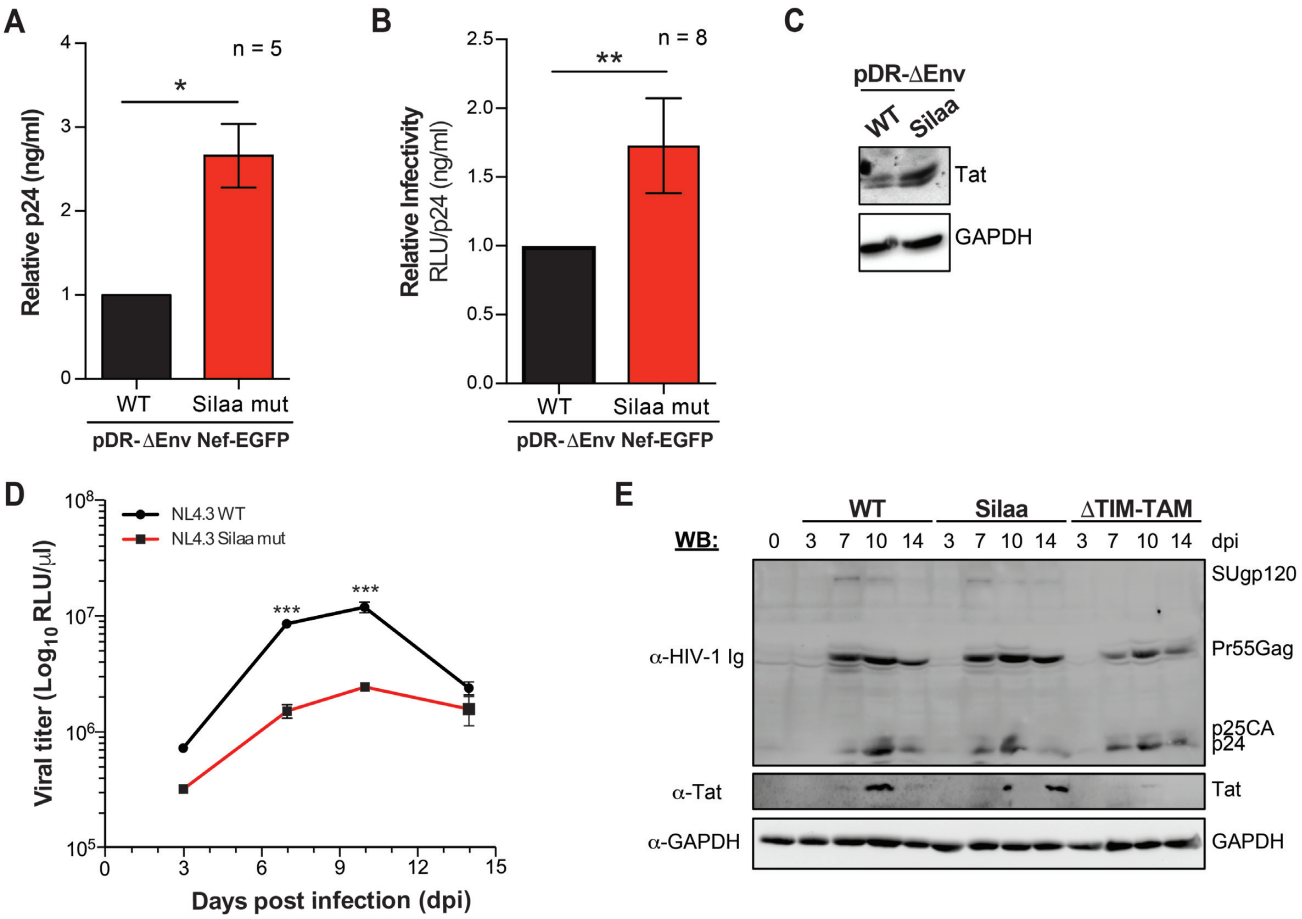

Figure 6. TIM-TAM controls productive infection. (**A**) Virus titres of single-round WT and Silaa mutant proviral clones pDR-ΔEnvNef-EGFP pseudotyped with AD8 Env. EGFP was used for normalisation of transfection efficiency. Virion production was quantified by $p24^{CA}$-ELISA of the supernatant 72 h post-transfection (mean ± SEM, $n = 5$). (**B**) Viral infectivity of single round EGFP virus was assessed by infection of PHA stimulated PBMCs with equal amounts of viral particles (20 ng per 10^6 cells). Viral infectivity was then determined by $p24^{CA}$-ELISA of the supernatant harvested 72 h post-infection (mean ± SEM, $n = 8$). (**C**) Western blot quantification of the HIV Tat protein and the endogenous protein GAPDH in the infected PBMCs with single-round infectious clones. (**D**) Virus kinetics of HIV pNL4–3 WT and silent mutant. Activated PBMCs from four healthy donors were infected with equal amount of pNL4–3 WT and Silaa mut viral particles. Supernatants were collected several days post-infection and assayed for virion production by infection of TZM-bl cells over a six-point series dilution with equal volumes of the supernatant harvested at the different time points. Luciferase activity (RLU) was measured 48 h post-infection of TZM-bl cells. The graph is representative of one kinetic infection of PBMCs. All four independent infections gave similar kinetics. (**E**) Western blot quantification of the HIV p24, Pr55Gag, Env SUgp120 (αHIV-sera) and Tat proteins, and the endogenous protein GAPDH in the infected PBMCs with pNL4–3 virus several days post-infection (dpi). For panels A, B, statistical significance was determined using two-tailed Student's t-test, while panel D, One Way ANOVA. $*P < 0.05$; $**P < 0.01$; $***P < 0.001$; ns, not significant.

dertaking integrase inhibitor therapy, RNAs arising from episomal unintegrated 1- and 2-LTR circular HIV DNAs (71,72) that abundantly accumulate in macrophages (73), or in infected T cells (74), lymphoid tissues and the brain during therapy (75,76). Tat sequences in rebounding viruses are highly similar to sequences found in unintegrated episomal DNA (77). Unintegrated HIV DNA is gradually silenced into repressive episomal minichromatin, however it remains inducible by HDAC inhibitors and Tat to promote persistent viral DNA transcription in primary macrophages (78). Thus, unintegrated DNA may serve as residual viral reservoir in resting T cells and macrophages. Indeed, when stimulated in the presence of cytokines, low-level viral replication can occur from unintegrated DNA (79–81).

Our data demonstrate a dual mode of action for TIM-TAM with a moderate IRES-dependent translation activity and a strong block of cap-dependent Tat translation due to a stable stem–loop structure (Figure 8). TIM-TAM impedes Tat cap-dependent translation by serving as a thermostable barrier to progressive ribosomal catalysis of amino acids needed for expression of Tat from large levels of 2 and 4 kb HIV mRNAs expressed in the early phase of productive HIV replication.

The TAR stem–loop present in all viral transcripts can also impede translation either by blocking access of translation initiation factors to the 5′-cap structure (82) or by activating the dsRNA-dependent protein kinase R (PKR) (83–85), which results in phosphorylation and inactivation of the α-subunit of eukaryotic initiation factor 2 (eIF2α) to inhibit translation initiation. TAR RNA folds into a stable stem–loop structure that engages with double-stranded RNA binding proteins (dsRBPs), such as PKR (86–88), RNA helicase A (RHA) (89), nuclear factor 90 (NF90) (90,91), TAR RNA-binding pro-

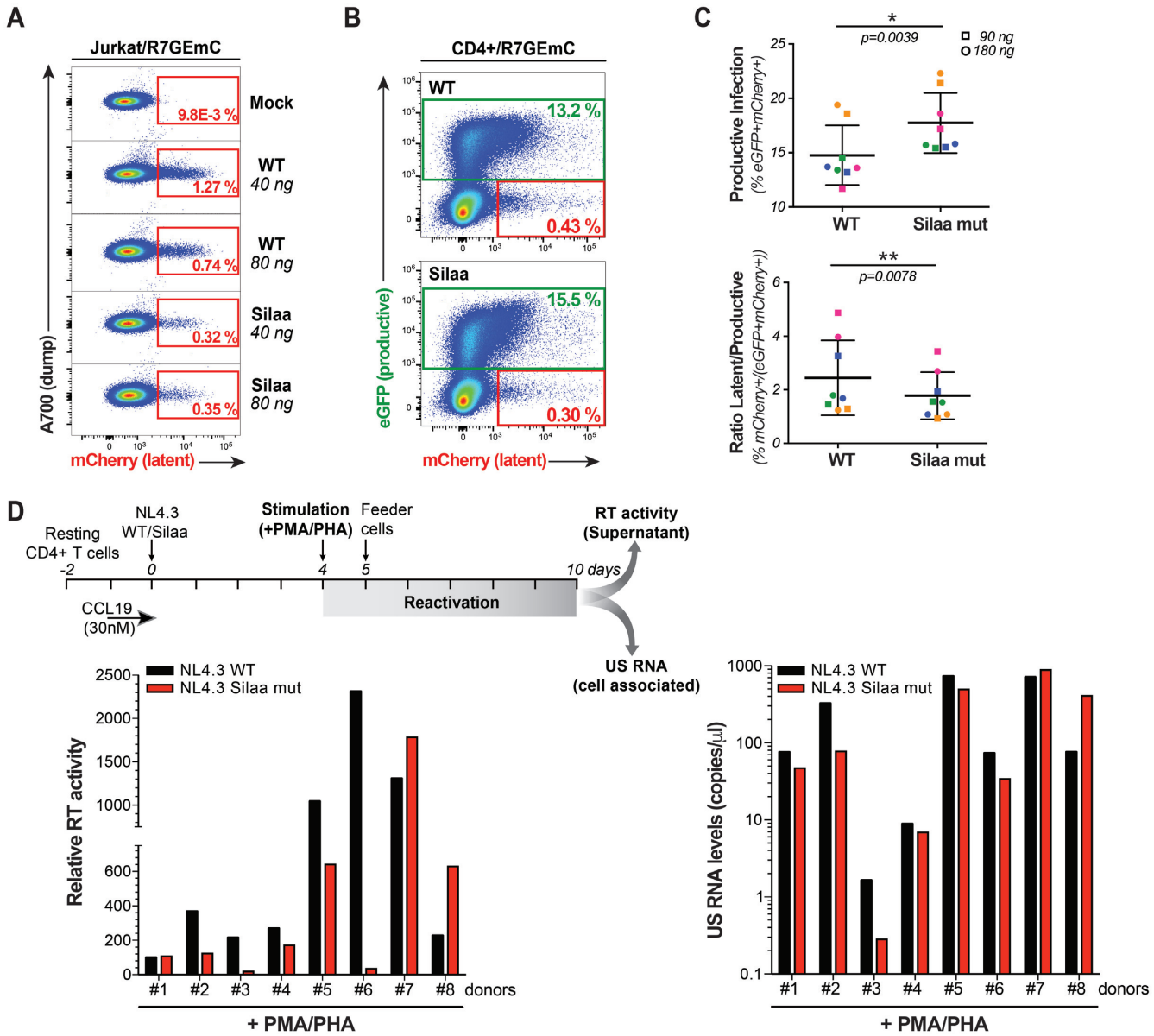

Figure 7. TIM-TAM controls provirus reactivation from latency. (**A**) Dot plot of mCherry+ population from Jurkat T cells 3 days post-infection with 40 or 80 ng of R7GEmC WT or Silaa virus. (**B**) Infection profiles of stimulated primary CD4+ T cells infected with R7GEmC WT or Silaa virus. Double-positive (eGFP+mCherry+, actively infected) and single positive (mCherry+, latently infected) cells are shown. Data shown are from a single donor, however are representative of four separate donors. (**C**) Quantification of productive infection (%eGFP+mCherry+) and ratio of latent to productive infection %mCherry+/(%eGFP+mCherry+) from panel (b). Data represents the mean ± SEM of four donors. * $P < 0.05$, ** $P < 0.01$ (two-tailed Student's *t*-test). (**D**) Diagram of the experimental procedure used for infection and re-stimulation of CCL19 treated resting CD4+ T cells. Resting CD4+ T cells were treated for 1 day with CCL19 (30 nM) before infection with WT and Silaa mutant pNL4–3 virus. The infected cells were then cultured with media containing 1 U/ml IL-2 for 4 days before re-stimulation with PMA/PHA (10 µg/ml, 10 nM respectively). PHA activated feeder PBMCs depleted of CD8 T cells were added 24 h after stimulation to enhance detection of virus reactivation. Virus production was quantified by measurement of the RT activity in the culture supernatant at day 10, as well as the cell associated unspliced RNA levels (US RNA copies/µl).

tein (TRBP) (92,93), adenosine deaminase acting on RNA (ADAR) and protein activator of PKR (PACT) (94), leading subsequently to activation or inhibition of translation initiation.

Our sequence and structural analysis revealed that TIM-TAM stem–loop presents an apical tetraloop, with conserved AG nucleotides at position 5446–5447 for subtype B (Supplementary Figure S2). More work is needed to define whether TIM-TAM folds into an AGNN tetraloop and serves as a platform for the binding of dsRBPs such as PKR and thus permitting a common mode of recognition of AGNN tetraloops by proteins with dsRNA binding domains (dsRBD). Moreover, extended RNA helix is required for PKR activation (95,96) and therefore, it would be interesting to assess the potential of TIM-TAM in PKR recruitment and activation through the potential interactions between TIM-TAM and TAR elements that may lead to an extended dsRNA helix structure that is optimal for

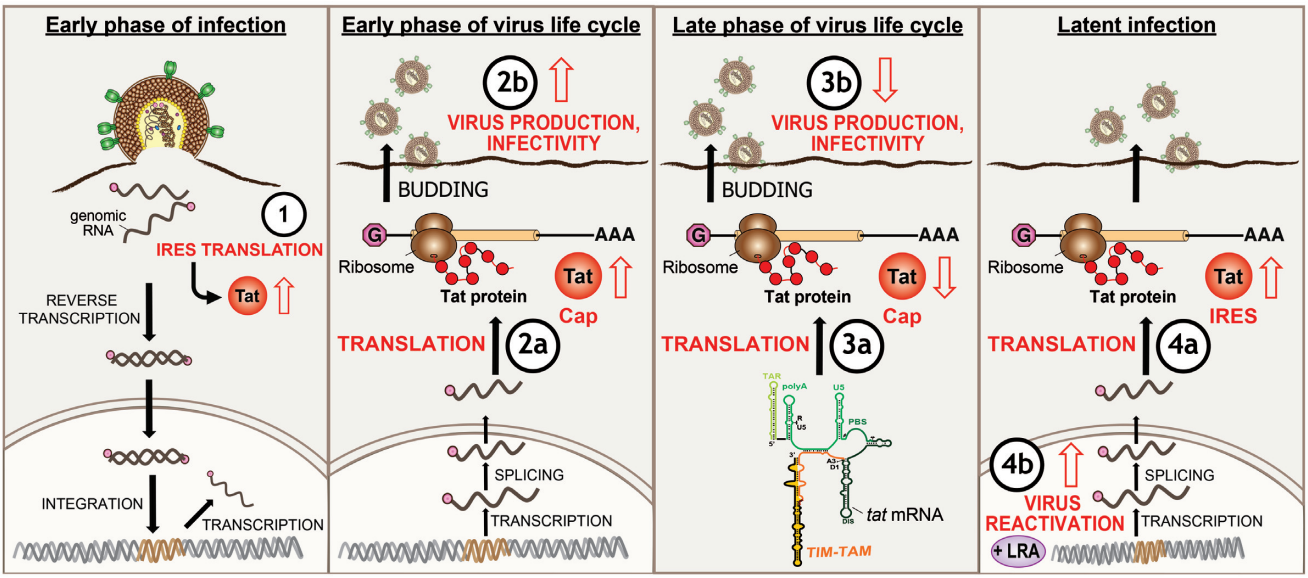

Figure 8. Model depicting TIM-TAM roles in controlling Tat expression. (1) At early rounds of infection before provirus integration, TIM-TAM enhances Tat translation from genomic RNA through an IRES-dependent mechanism thus providing an immediate early expression of Tat to pioneer its viral transactivation functions. (2, 3) By controlling Tat translation, TIM-TAM impacts on timing and level of Tat protein expression at the early and late phase of the virus life cycle to control HIV production and infectivity. (3a, 4a) Due to the strong structure stability, TIM-TAM impedes fully efficient cap-dependent Tat translation but confers moderately efficient IRES-dependent translation activity from spliced *tat* mRNA. Thus, TIM-TAM acts as a switch for HIV productive and latent infection. (4) Tat IRES enables non-conventional viral mRNA to express HIV peptides or Tat protein to assist reactivation from latency in the presence of latency reversing agents (LRA).

PKR dimerisation and activation for Tat translation inhibition. A similar model of dynamic refolding of *interferon-γ* (*IFN-γ*) mRNA enables it to function as a translation template and as a PKR activator preventing cytokine expression (97). Interestingly, the structured TIM-TAM element is located downstream of the authentic Tat AUG initiation site, therefore Tat expression would require the dynamic folding of *tat* mRNA and/or the 'backward' scanning of the 40S small ribosomal subunit. This aforementioned phenomenon has been recently reported for HIV-1 (Gag p40-p55) (98) and HIV-2 (Gag p57-p50 and p44) (99) Gag IRES elements.

In vivo, the residual provirus remaining on ART varies widely depending on the individual and the reservoir site (26). Latently infected cells are enriched with Tat variants comprising impaired transactivation activity with several of the mutated residues (P10, W11, K/W12, G15) located within TIM-TAM (100). Given that the frequency of HIV infected cells increases over time due to clonal expansion of latently infected CD4+ T cells (3,4,101), and the majority of the HIV-1 proviruses are defective (7,21), an IRES activity embedded within the TIM-TAM motif is an important determinant of whether or not any of these defective proviruses could lead to the production of viral proteins in the setting of prolonged viral suppression. As HIV integration plays an important role in the expansion and persistence of latently infected cells, further studies are needed to determine the role of Tat-IRES mediated translation on the expression of the adjacent cellular genes including *MKL2*, *BACH2*, *STAT5B*, and *HORMAD2* that have been linked to clonal expansion (3,4,101).

Recent studies revealed that HIV_{def} proviruses may bypass major splice donor defects to produce novel HIV RNA transcripts that can be translated *in vitro* and *ex vivo* (22,23), and become targets for HIV-specific cytotoxic T lymphocytes (CTLs) (23). Given that Tat is associated with T cell activation, and even if a small number of latently infected cells contained intact Tat sequence and function, Tat expression through IRES-dependent translation may drive the ongoing immune activation and pathogenesis (22,23). Alternatively aberrant mRNAs that capture the TIM-TAM element, such as may occur from deleted proviruses, or aberrantly spliced viral mRNAs, or even aberrant HIV–cellular mRNAs may result in viral peptide expression that enables noise driven events (102) and thus shapes the structure of the latent HIV reservoir (22–24).

Tuning Tat levels could be sufficient to control HIV expression and therefore used to 'flush out' the latently infected cells independently of the cell activation state (64,103). On the other hand, inhibition of Tat expression and/or activity could induce a prolonged transcriptional silencing, hence preventing viral rebound and replenishment of the reservoir by maintaining a permanent latent state. Indeed, the use of didehydro-cortistatin A (dCA), a Tat dependent transcription inhibitor, significantly diminished the capacity of virus reactivation from latently infected cells (92% of αCD3/CD28) and hence the size of the viral reservoir (104,105). Thus, TIM-TAM constitutes a potential drug target for residual Tat expression and latency reversal, which can be targeted by new generation of latency reversing agents such as the 2-acylaminothiazole class that activates Tat expression in HIV-1 latency models (106). In

conclusion, Tat IRES-translation may shape the outcome of provirus reactivation from latency.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank M. Zucker, K. Weeks and E. Bell for RNA folding and analysis software, S. Saleh and A. Gatignol for their insight into primary T-cell infections, D. Allen, and J. Howard for their technical assistance. We are grateful to S. Sonza of UoM for helpful discussions and comments. We also thank the DMI Flow Facility staff for their advice and generous assistance in maintaining the flow cytometers used in this study.

Author contributions: Conceptualization, G.K., D.FJ.P. and C.B.; Methodology, G.K., S.R.L. and D.FJ.P.; Investigation, G.K., C.M., D.FJ.P. and L.A.; Writing – Original Draft, G.K. and D.FJ.P.; Writing – Review & Editing, G.K., S.R.L., D.FJ.P.; Funding Acquisition, D.FJ.P. and C.B.; Resources, D.FJ.P. and C.B.

FUNDING

National Health and Medical Research Council of Australia (NHMRC) program and project grants [1052979, 1129320 to D.FJ.P.]; National Agency for Research on AIDS and Hepatitis (ANRS) [05195-06194, 08061-09069 to C.B.]; G.K. was a fellow of the french 'Ministère de la Recherche et de l'Enseignement' and the Sidaction and Pierre Bergé foundation. Funding for open access charge: NHMRC project grant [1129320].

Conflict of interest statement. None declared.

REFERENCES

1. Deeks,S.G., Lewin,S.R., Ross,A.L., Ananworanich,J., Benkirane,M., Cannon,P., Chomont,N., Douek,D., Lifson,J.D., Lo,Y.R. *et al.* (2016) International AIDS Society global scientific strategy: Towards an HIV cure 2016. *Nat. Med.*, **22**, 839–850.
2. Chomont,N., El-Far,M., Ancuta,P., Trautmann,L., Procopio,F.A., Yassine-Diab,B., Boucher,G., Boulassel,M.R., Ghattas,G., Brenchley,J.M. *et al.* (2009) HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nat. Med.*, **15**, 893–900.
3. Maldarelli,F., Wu,X., Su,L., Simonetti,F.R., Shao,W., Hill,S., Spindler,J., Ferris,A.L., Mellors,J.W., Kearney,M.F. *et al.* (2014) HIV latency. Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. *Science*, **345**, 179–183.
4. Wagner,T.A., McLaughlin,S., Garg,K., Cheung,C.Y., Larsen,B.B., Styrchak,S., Huang,H.C., Edlefsen,P.T., Mullins,J.I. and Frenkel,L.M. (2014) HIV latency. Proliferation of cells with HIV integrated into cancer genes contributes to persistent infection. *Science*, **345**, 570–573.
5. Chavez,L., Calvanese,V. and Verdin,E. (2015) HIV latency is established directly and early in both resting and activated primary CD4 T cells. *PLoS Pathog.*, **11**, e1004955.
6. Hiener,B., Horsburgh,B.A., Eden,J.S., Barton,K., Schlub,T.E., Lee,E., von Stockenstrom,S., Odevall,L., Milush,J.M., Liegler,T. *et al.* (2017) Identification of genetically intact HIV-1 proviruses in specific CD4(+) T Cells from effectively treated participants. *Cell Rep.*, **21**, 813–822.
7. Ho,Y.C., Shan,L., Hosmane,N.N., Wang,J., Laskey,S.B., Rosenbloom,D.I., Lai,J., Blankson,J.N., Siliciano,J.D. and Siliciano,R.F. (2013) Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell*, **155**, 540–551.
8. Crooks,A.M., Bateson,R., Cope,A.B., Dahl,N.P., Griggs,M.K., Kuruc,J.D., Gay,C.L., Eron,J.J., Margolis,D.M., Bosch,R.J. *et al.* (2015) Precise quantitation of the latent HIV-1 reservoir: implications for eradication strategies. *J. Infect. Dis.*, **212**, 1361–1365.
9. Siliciano,J.D., Kajdas,J., Finzi,D., Quinn,T.C., Chadwick,K., Margolick,J.B., Kovacs,C., Gange,S.J. and Siliciano,R.F. (2003) Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells. *Nat. Med.*, **9**, 727–728.
10. Zerbato,J.M., Purves,H.V., Lewin,S.R. and Rasmussen,T.A. (2019) Between a shock and a hard place: challenges and developments in HIV latency reversal. *Curr. Opin. Virol.*, **38**, 1–9.
11. Donahue,D.A., Kuhl,B.D., Sloan,R.D. and Wainberg,M.A. (2012) The viral protein Tat can inhibit the establishment of HIV-1 latency. *J. Virol.*, **86**, 3253–3263.
12. Khoury,G., Darcis,G., Lee,M.Y., Bouchat,S., Van Driessche,B., Purcell,D.F. and Van Lint,C. (2018) *HIV Vaccine and Cure - The Path Towards Finding an Effective Cure and Vaccine*. Springer Nature, Vol. **1075**.
13. Muniz,L., Egloff,S., Ughy,B., Jady,B.E. and Kiss,T. (2010) Controlling cellular P-TEFb activity by the HIV-1 transcriptional transactivator Tat. *PLoS Pathog.*, **6**, e1001152.
14. Peterlin,B.M. and Price,D.H. (2006) Controlling the elongation phase of transcription with P-TEFb. *Mol. Cell*, **23**, 297–305.
15. Lu,X., Zhu,X., Li,Y., Liu,M., Yu,B., Wang,Y., Rao,M., Yang,H., Zhou,K., Wang,Y. *et al.* (2016) Multiple P-TEFbs cooperatively regulate the release of promoter-proximally paused RNA polymerase II. *Nucleic Acids Res.*, **44**, 6853–6867.
16. Ping,Y.H. and Rana,T.M. (2001) DSIF and NELF interact with RNA polymerase II elongation complex and HIV-1 Tat stimulates P-TEFb-mediated phosphorylation of RNA polymerase II and DSIF during transcription elongation. *J. Biol. Chem.*, **276**, 12951–12958.
17. Chiu,Y.L., Coronel,E., Ho,C.K., Shuman,S. and Rana,T.M. (2001) HIV-1 Tat protein interacts with mammalian capping enzyme and stimulates capping of TAR RNA. *J. Biol. Chem.*, **276**, 12959–12966.
18. Chiu,Y.L., Ho,C.K., Saha,N., Schwer,B., Shuman,S. and Rana,T.M. (2002) Tat stimulates cotranscriptional capping of HIV mRNA. *Mol. Cell*, **10**, 585–597.
19. Zhou,M., Deng,L., Kashanchi,F., Brady,J.N., Shatkin,A.J. and Kumar,A. (2003) The Tat/TAR-dependent phosphorylation of RNA polymerase II C-terminal domain stimulates cotranscriptional capping of HIV-1 mRNA. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 12666–12671.
20. Jablonski,J.A., Amelio,A.L., Giacca,M. and Caputi,M. (2010) The transcriptional transactivator Tat selectively regulates viral splicing. *Nucleic Acids Res.*, **38**, 1249–1260.
21. Bruner,K.M., Murray,A.J., Pollack,R.A., Soliman,M.G., Laskey,S.B., Capoferri,A.A., Lai,J., Strain,M.C., Lada,S.M., Hoh,R. *et al.* (2016) Defective proviruses rapidly accumulate during acute HIV-1 infection. *Nat. Med.*, **22**, 1043–1049.
22. Imamichi,H., Dewar,R.L., Adelsberger,J.W., Rehm,C.A., O'Doherty,U., Paxinos,E.E., Fauci,A.S. and Lane,H.C. (2016) Defective HIV-1 proviruses produce novel protein-coding RNA species in HIV-infected patients on combination antiretroviral therapy. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, 8783–8788.
23. Pollack,R.A., Jones,R.B., Pertea,M., Bruner,K.M., Martin,A.R., Thomas,A.S., Capoferri,A.A., Beg,S.A., Huang,S.H., Karandish,S. *et al.* (2017) Defective HIV-1 proviruses are expressed and can be recognized by cytotoxic T lymphocytes, which shape the proviral landscape. *Cell Host Microbe*, **21**, 494–506.
24. Pinzone,M.R., VanBelzen,D.J., Weissman,S., Bertuccio,M.P., Cannon,L., Venanzi-Rullo,E., Migueles,S., Jones,R.B., Mota,T., Joseph,S.B. *et al.* (2019) Longitudinal HIV sequencing reveals reservoir expression leading to decay which is obscured by clonal expansion. *Nat. Commun.*, **10**, 728.
25. Pasternak,A.O., DeMaster,L.K., Kootstra,N.A., Reiss,P., O'Doherty,U. and Berkhout,B. (2016) Minor contribution of

chimeric host-HIV readthrough transcripts to the level of HIV cell-associated gag RNA. *J. Virol.*, **90**, 1148–1151.
26. Telwatte,S., Lee,S., Somsouk,M., Hatano,H., Baker,C., Kaiser,P., Kim,P., Chen,T.H., Milush,J., Hunt,P.W. *et al.* (2018) Gut and blood differ in constitutive blocks to HIV transcription, suggesting tissue-specific differences in the mechanisms that govern HIV latency. *PLoS Pathog.*, **14**, e1007357.
27. 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**, eaap9927.
28. Goldwich,A., Hahn,S.S., Schreiber,S., Meier,S., Kampgen,E., Wagner,R., Lutz,M.B. and Schubert,U. (2008) Targeting HIV-1 Gag into the defective ribosomal product pathway enhances MHC class I antigen presentation and CD8+ T cell activation. *J. Immunol.*, **180**, 372–382.
29. Yewdell,J.W., Anton,L.C. and Bennink,J.R. (1996) Defective ribosomal products (DRiPs): a major source of antigenic peptides for MHC class I molecules? *J. Immunol.*, **157**, 1823–1826.
30. Bullock,T.N. and Eisenlohr,L.C. (1996) Ribosomal scanning past the primary initiation codon as a mechanism for expression of CTL epitopes encoded in alternative reading frames. *J. Exp. Med.*, **184**, 1319–1329.
31. Schwartz,S., Felber,B.K., Fenyo,E.M. and Pavlakis,G.N. (1990) Env and Vpu proteins of human immunodeficiency virus type 1 are produced from multiple bicistronic mRNAs. *J. Virol.*, **64**, 5448–5456.
32. Brasey,A., Lopez-Lastra,M., Ohlmann,T., Beerens,N., Berkhout,B., Darlix,J.L. and Sonenberg,N. (2003) The leader of human immunodeficiency virus type 1 genomic RNA harbors an internal ribosome entry segment that is active during the G2/M phase of the cell cycle. *J. Virol.*, **77**, 3939–3949.
33. Charnay,N., Ivanyi-Nagy,R., Soto-Rifo,R., Ohlmann,T., Lopez-Lastra,M. and Darlix,J.L. (2009) Mechanism of HIV-1 Tat RNA translation and its activation by the Tat protein. *Retrovirology*, **6**, 74.
34. Buck,C.B., Shen,X., Egan,M.A., Pierson,T.C., Walker,C.M. and Siliciano,R.F. (2001) The human immunodeficiency virus type 1 gag gene encodes an internal ribosome entry site. *J. Virol.*, **75**, 181–191.
35. Locker,N., Chamond,N. and Sargueil,B. (2011) A conserved structure within the HIV gag open reading frame that controls translation initiation directly recruits the 40S subunit and eIF3. *Nucleic Acids Res.*, **39**, 2367–2377.
36. Gendron,K., Ferbeyre,G., Heveker,N. and Brakier-Gingras,L. (2011) The activity of the HIV-1 IRES is stimulated by oxidative stress and controlled by a negative regulatory element. *Nucleic Acids Res.*, **39**, 902–912.
37. Monette,A., Ajamian,L., Lopez-Lastra,M. and Mouland,A.J. (2009) Human immunodeficiency virus type 1 (HIV-1) induces the cytoplasmic retention of heterogeneous nuclear ribonucleoprotein A1 by disrupting nuclear import: implications for HIV-1 gene expression. *J. Biol. Chem.*, **284**, 31350–31362.
38. Dale,C.J., De Rose,R., Wilson,K.M., Croom,H.A., Thomson,S., Coupar,B.E., Ramsay,A., Purcell,D.F., Ffrench,R., Law,M. *et al.* (2004) Evaluation in macaques of HIV-1 DNA vaccines containing primate CpG motifs and fowlpoxvirus vaccines co-expressing IFNgamma or IL-12. *Vaccine*, **23**, 188–197.
39. Jacquenet,S., Ropers,D., Bilodeau,P.S., Damier,L., Mougin,A., Stoltzfus,C.M. and Branlant,C. (2001) Conserved stem-loop structures in the HIV-1 RNA region containing the A3 3′ splice site and its cis-regulatory element: possible involvement in RNA splicing. *Nucleic Acids Res.*, **29**, 464–478.
40. Ropers,D., Ayadi,L., Gattoni,R., Jacquenet,S., Damier,L., Branlant,C. and Stevenin,J. (2004) Differential effects of the SR proteins 9G8, SC35, ASF/SF2, and SRp40 on the utilization of the A1 to A5 splicing sites of HIV-1 RNA. *J. Biol. Chem.*, **279**, 29963–29973.
41. Mortimer,S.A. and Weeks,K.M. (2007) A fast-acting reagent for accurate analysis of RNA secondary and tertiary structure by SHAPE chemistry. *J. Am. Chem. Soc.*, **129**, 4144–4145.
42. Karabiber,F., McGinnis,J.L., Favorov,O.V. and Weeks,K.M. (2013) QuShape: rapid, accurate, and best-practices quantification of nucleic acid probing information, resolved by capillary electrophoresis. *RNA*, **19**, 63–73.
43. Mougin,A., Gregoire,A., Banroques,J., Segault,V., Fournier,R., Brule,F., Chevrier-Miller,M. and Branlant,C. (1996) Secondary structure of the yeast Saccharomyces cerevisiae pre-U3A snoRNA and its implication for splicing efficiency. *RNA*, **2**, 1079–1093.
44. Maenner,S., Blaud,M., Fouillen,L., Savoye,A., Marchand,V., Dubois,A., Sanglier-Cianferani,S., Van Dorsselaer,A., Clerc,P., Avner,P. *et al.* (2010) 2-D structure of the A region of Xist RNA and its implication for PRC2 association. *PLoS Biol.*, **8**, e1000276.
45. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
46. Sarzotti-Kelsoe,M., Bailer,R.T., Turk,E., Lin,C.L., Bilska,M., Greene,K.M., Gao,H., Todd,C.A., Ozaki,D.A., Seaman,M.S. *et al.* (2014) Optimization and validation of the TZM-bl assay for standardized assessments of neutralizing antibodies against HIV-1. *J. Immunol. Methods*, **409**, 131–146.
47. Clerzius,G., Shaw,E., Daher,A., Burugu,S., Gelinas,J.F., Ear,T., Sinck,L., Routy,J.P., Mouland,A.J., Patel,R.C. *et al.* (2013) The PKR activator, PACT, becomes a PKR inhibitor during HIV-1 replication. *Retrovirology*, **10**, 96.
48. Lewin,S.R., Murray,J.M., Solomon,A., Wightman,F., Cameron,P.U., Purcell,D.J., Zaunders,J.J., Grey,P., Bloch,M., Smith,D. *et al.* (2008) Virologic determinants of success after structured treatment interruptions of antiretrovirals in acute HIV-1 infection. *J. Acquir. Immune Defic. Syndr.*, **47**, 140–147.
49. Saleh,S., Wightman,F., Ramanayake,S., Alexander,M., Kumar,N., Khoury,G., Pereira,C., Purcell,D., Cameron,P.U. and Lewin,S.R. (2011) Expression and reactivation of HIV in a chemokine induced model of HIV latency in primary resting CD4+ T cells. *Retrovirology*, **8**, 80.
50. Cameron,P.U., Saleh,S., Sallmann,G., Solomon,A., Wightman,F., Evans,V.A., Boucher,G., Haddad,E.K., Sekaly,R.P., Harman,A.N. *et al.* (2010) Establishment of HIV-1 latency in resting CD4+ T cells depends on chemokine-induced changes in the actin cytoskeleton. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 16934–16939.
51. Goff,S., Traktman,P. and Baltimore,D. (1981) Isolation and properties of Moloney murine leukemia virus mutants: use of a rapid assay for release of virion reverse transcriptase. *J. Virol.*, **38**, 239–248.
52. Katoh,K. and Standley,D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.*, **30**, 772–780.
53. Keele,B.F., Giorgi,E.E., Salazar-Gonzalez,J.F., Decker,J.M., Pham,K.T., Salazar,M.G., Sun,C., Grayson,T., Wang,S., Li,H. *et al.* (2008) Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 7552–7557.
54. Korber,B. (2001) *HIV Signature and Sequence Variation Analysis. Computational Analysis of HIV Molecular Sequences*. 1st edn. Springer, US.
55. Crooks,G.E., Hon,G., Chandonia,J.M. and Brenner,S.E. (2004) WebLogo: a sequence logo generator. *Genome Res.*, **14**, 1188–1190.
56. Delli Ponti,R., Marti,S., Armaos,A. and Tartaglia,G.G. (2017) A high-throughput approach to profile RNA structure. *Nucleic Acids Res.*, **45**, e35.
57. Jeang,K.T., Xiao,H. and Rich,E.A. (1999) Multifaceted activities of the HIV-1 transactivator of transcription, Tat. *J. Biol. Chem.*, **274**, 28837–28840.
58. 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.
59. Saliou,J.M., Bourgeois,C.F., Ayadi-Ben Mena,L., Ropers,D., Jacquenet,S., Marchand,V., Stevenin,J. and Branlant,C. (2009) Role of RNA structure and protein factors in the control of HIV-1 splicing. *Front. Biosci. (Landmark Ed.)*, **14**, 2714–2729.
60. Watts,J.M., Dang,K.K., Gorelick,R.J., Leonard,C.W., Bess,J.W. Jr., Swanstrom,R., Burch,C.L. and Weeks,K.M. (2009) Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature*, **460**, 711–716.
61. Hallay,H., Locker,N., Ayadi,L., Ropers,D., Guittet,E. and Branlant,C. (2006) Biochemical and NMR study on the competition between proteins SC35, SRp40, and heterogeneous nuclear ribonucleoprotein A1 at the HIV-1 Tat exon 2 splicing site. *J. Biol. Chem.*, **281**, 37159–37174.

62. Khoury,G., Ayadi,L. and Branlant,C. (2011) Translational control mechanism of HIV-1 tat1 mRNA. *Retrovirology*, **8**(Suppl. 2), P39.
63. Lassen,K.G., Ramyar,K.X., Bailey,J.R., Zhou,Y. and Siliciano,R.F. (2006) Nuclear retention of multiply spliced HIV-1 RNA in resting CD4+ T cells. *PLoS Pathog.*, **2**, e68.
64. Razooky,B.S., Pai,A., Aull,K., Rouzine,I.M. and Weinberger,L.S. (2015) A hardwired HIV latency program. *Cell*, **160**, 990–1001.
65. Calvanese,V., Chavez,L., Laurent,T., Ding,S. and Verdin,E. (2013) Dual-color HIV reporters trace a population of latently infected cells and enable their purification. *Virology*, **446**, 283–292.
66. Spina,C.A., Anderson,J., Archin,N.M., Bosque,A., Chan,J., Famiglietti,M., Greene,W.C., Kashuba,A., Lewin,S.R., Margolis,D.M. *et al.* (2013) An in-depth comparison of latent HIV-1 reactivation in multiple cell model systems and resting CD4+ T cells from aviremic patients. *PLoS Pathog.*, **9**, e1003834.
67. Sloan,R.D., Donahue,D.A., Kuhl,B.D., Bar-Magen,T. and Wainberg,M.A. (2010) Expression of Nef from unintegrated HIV-1 DNA downregulates cell surface CXCR4 and CCR5 on T-lymphocytes. *Retrovirology*, **7**, 44.
68. Wu,Y. and Marsh,J.W. (2003) Early transcription from nonintegrated DNA in human immunodeficiency virus infection. *J. Virol.*, **77**, 10376–10382.
69. Gillim-Ross,L., Cara,A. and Klotman,M.E. (2005) HIV-1 extrachromosomal 2-LTR circular DNA is long-lived in human macrophages. *Viral Immunol.*, **18**, 190–196.
70. Kelly,J., Beddall,M.H., Yu,D., Iyer,S.R., Marsh,J.W. and Wu,Y. (2008) Human macrophages support persistent transcription from unintegrated HIV-1 DNA. *Virology*, **372**, 300–312.
71. Buzon,M.J., Massanella,M., Llibre,J.M., Esteve,A., Dahl,V., Puertas,M.C., Gatell,J.M., Domingo,P., Paredes,R., Sharkey,M. *et al.* (2010) HIV-1 replication and immune dynamics are affected by raltegravir intensification of HAART-suppressed subjects. *Nat. Med.*, **16**, 460–465.
72. Hatano,H., Strain,M.C., Scherzer,R., Bacchetti,P., Wentworth,D., Hoh,R., Martin,J.N., McCune,J.M., Neaton,J.D., Tracy,R.P. *et al.* (2013) Increase in 2-long terminal repeat circles and decrease in D-dimer after raltegravir intensification in patients with treated HIV infection: a randomized, placebo-controlled trial. *J. Infect. Dis.*, **208**, 1436–1442.
73. Chun,T.W., Carruth,L., Finzi,D., Shen,X., DiGiuseppe,J.A., Taylor,H., Hermankova,M., Chadwick,K., Margolick,J., Quinn,T.C. *et al.* (1997) Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. *Nature*, **387**, 183–188.
74. Pauza,C.D., Galindo,J.E. and Richman,D.D. (1990) Reinfection results in accumulation of unintegrated viral DNA in cytopathic and persistent human immunodeficiency virus type 1 infection of CEM cells. *J. Exp. Med.*, **172**, 1035–1042.
75. Pang,S., Koyanagi,Y., Miles,S., Wiley,C., Vinters,H.V. and Chen,I.S. (1990) High levels of unintegrated HIV-1 DNA in brain tissue of AIDS dementia patients. *Nature*, **343**, 85–89.
76. Teo,I., Veryard,C., Barnes,H., An,S.F., Jones,M., Lantos,P.L., Luthert,P. and Shaunak,S. (1997) Circular forms of unintegrated human immunodeficiency virus type 1 DNA and high levels of viral protein expression: association with dementia and multinucleated giant cells in the brains of patients with AIDS. *J. Virol.*, **71**, 2928–2933.
77. Sharkey,M., Babic,D.Z., Greenough,T., Gulick,R., Kuritzkes,D.R. and Stevenson,M. (2011) Episomal viral cDNAs identify a reservoir that fuels viral rebound after treatment interruption and that contributes to treatment failure. *PLoS Pathog.*, **7**, e1001303.
78. Meltzer,B., Dabbagh,D., Guo,J., Kashanchi,F., Tyagi,M. and Wu,Y. (2018) Tat controls transcriptional persistence of unintegrated HIV genome in primary human macrophages. *Virology*, **518**, 241–252.
79. Chan,C.N., Trinite,B., Lee,C.S., Mahajan,S., Anand,A., Wodarz,D., Sabbaj,S., Bansal,A., Goepfert,P.A. and Levy,D.N. (2016) HIV-1 latency and virus production from unintegrated genomes following direct infection of resting CD4 T cells. *Retrovirology*, **13**, 1.
80. Kantor,B., Ma,H., Webster-Cyriaque,J., Monahan,P.E. and Kafri,T. (2009) Epigenetic activation of unintegrated HIV-1 genomes by gut-associated short chain fatty acids and its implications for HIV infection. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 18786–18791.
81. Trinite,B., Ohlson,E.C., Voznesensky,I., Rana,S.P., Chan,C.N., Mahajan,S., Alster,J., Burke,S.A., Wodarz,D. and Levy,D.N. (2013) An HIV-1 replication pathway utilizing reverse transcription products that fail to integrate. *J. Virol.*, **87**, 12701–12720.
82. Dorin,D., Bonnet,M.C., Bannwarth,S., Gatignol,A., Meurs,E.F. and Vaquero,C. (2003) The TAR RNA-binding protein, TRBP, stimulates the expression of TAR-containing RNAs in vitro and in vivo independently of its ability to inhibit the dsRNA-dependent kinase PKR. *J. Biol. Chem.*, **278**, 4440–4448.
83. Geballe,A.P. and Gray,M.K. (1992) Variable inhibition of cell-free translation by HIV-1 transcript leader sequences. *Nucleic Acids Res.*, **20**, 4291–4297.
84. Park,H., Davies,M.V., Langland,J.O., Chang,H.W., Nam,Y.S., Tartaglia,J., Paoletti,E., Jacobs,B.L., Kaufman,R.J. and Venkatesan,S. (1994) TAR RNA-binding protein is an inhibitor of the interferon-induced protein kinase PKR. *Proc. Natl. Acad. Sci. U.S.A.*, **91**, 4713–4717.
85. Silverman,R.H. and Sengupta,D.N. (1990) Translational regulation by HIV leader RNA, TAT, and interferon-inducible enzymes. *J. Exp. Pathol.*, **5**, 69–77.
86. Carpick,B.W., Graziano,V., Schneider,D., Maitra,R.K., Lee,X. and Williams,B.R. (1997) Characterization of the solution complex between the interferon-induced, double-stranded RNA-activated protein kinase and HIV-I trans-activating region RNA. *J. Biol. Chem.*, **272**, 9510–9516.
87. Kim,I., Liu,C.W. and Puglisi,J.D. (2006) Specific recognition of HIV TAR RNA by the dsRNA binding domains (dsRBD1-dsRBD2) of PKR. *J. Mol. Biol.*, **358**, 430–442.
88. Maitra,R.K., McMillan,N.A., Desai,S., McSwiggen,J., Hovanessian,A.G., Sen,G., Williams,B.R. and Silverman,R.H. (1994) HIV-1 TAR RNA has an intrinsic ability to activate interferon-inducible enzymes. *Virology*, **204**, 823–827.
89. Bolinger,C., Sharma,A., Singh,D., Yu,L. and Boris-Lawrie,K. (2010) RNA helicase A modulates translation of HIV-1 and infectivity of progeny virions. *Nucleic Acids Res.*, **38**, 1686–1696.
90. Agbottah,E.T., Traviss,C., McArdle,J., Karki,S., St Laurent,G.C. 3rd and Kumar,A. (2007) Nuclear Factor 90(NF90) targeted to TAR RNA inhibits transcriptional activation of HIV-1. *Retrovirology*, **4**, 41.
91. Hoque,M., Shamanna,R.A., Guan,D., Pe'ery,T. and Mathews,M.B. (2011) HIV-1 replication and latency are regulated by translational control of cyclin T1. *J. Mol. Biol.*, **410**, 917–932.
92. Benkirane,M., Neuveut,C., Chun,R.F., Smith,S.M., Samuel,C.E., Gatignol,A. and Jeang,K.T. (1997) Oncogenic potential of TAR RNA binding protein TRBP and its regulatory interaction with RNA-dependent protein kinase PKR. *EMBO J.*, **16**, 611–624.
93. Gatignol,A., Buckler-White,A., Berkhout,B. and Jeang,K.T. (1991) Characterization of a human TAR RNA-binding protein that activates the HIV-1 LTR. *Science*, **251**, 1597–1600.
94. Chukwurah,E., Handy,I. and Patel,R.C. (2017) ADAR1 and PACT contribute to efficient translation of transcripts containing HIV-1 trans-activating response (TAR) element. *Biochem. J.*, **474**, 1241–1257.
95. Heinicke,L.A., Wong,C.J., Lary,J., Nallagatla,S.R., Diegelman-Parente,A., Zheng,X., Cole,J.L. and Bevilacqua,P.C. (2009) RNA dimerization promotes PKR dimerization and activation. *J. Mol. Biol.*, **390**, 319–338.
96. Pfaller,C.K., Li,Z., George,C.X. and Samuel,C.E. (2011) Protein kinase PKR and RNA adenosine deaminase ADAR1: new roles for old players as modulators of the interferon response. *Curr. Opin. Immunol.*, **23**, 573–582.
97. Cohen-Chalamish,S., Hasson,A., Weinberg,D., Namer,L.S., Banai,Y., Osman,F. and Kaempfer,R. (2009) Dynamic refolding of IFN-gamma mRNA enables it to function as PKR activator and translation template. *Nat. Chem. Biol.*, **5**, 896–903.
98. Deforges,J., de Breyne,S., Ameur,M., Ulryck,N., Chamond,N., Saaidi,A., Ponty,Y., Ohlmann,T. and Sargueil,B. (2017) Two ribosome recruitment sites direct multiple translation events within HIV1 Gag open reading frame. *Nucleic Acids Res.*, **45**, 7538.
99. Herbreteau,C.H., Weill,L., Decimo,D., Prevot,D., Darlix,J.L., Sargueil,B. and Ohlmann,T. (2005) HIV-2 genomic RNA contains a novel type of IRES located downstream of its initiation codon. *Nat. Struct. Mol. Biol.*, **12**, 1001–1007.
100. Yukl,S., Pillai,S., Li,P., Chang,K., Pasutti,W., Ahlgren,C., Havlir,D., Strain,M., Gunthard,H., Richman,D. *et al.* (2009) Latently-infected

CD4+ T cells are enriched for HIV-1 Tat variants with impaired transactivation activity. *Virology*, **387**, 98–108.
101. Cohn,L.B., Silva,I.T., Oliveira,T.Y., Rosales,R.A., Parrish,E.H., Learn,G.H., Hahn,B.H., Czartoski,J.L., McElrath,M.J., Lehmann,C. *et al.* (2015) HIV-1 integration landscape during latent and active infection. *Cell*, **160**, 420–432.
102. Hansen,M.M.K., Wen,W.Y., Ingerman,E., Razooky,B.S., Thompson,C.E., Dar,R.D., Chin,C.W., Simpson,M.L. and Weinberger,L.S. (2018) A post-transcriptional feedback mechanism for noise suppression and fate stabilization. *Cell*, **173**, 1609–1621.
103. Khoury,G., Mota,T.M., Li,S., Tumpach,C., Lee,M.Y., Jacobson,J., Harty,L., Anderson,J.L., Lewin,S.R. and Purcell,D.F.J. (2018) HIV latency reversing agents act through Tat post translational modifications. *Retrovirology*, **15**, 36.
104. Mousseau,G., Kessing,C.F., Fromentin,R., Trautmann,L., Chomont,N. and Valente,S.T. (2015) The Tat inhibitor didehydro-cortistatin A prevents HIV-1 reactivation from latency. *MBio*, **6**, e00465.
105. Kessing,C.F., Nixon,C.C., Li,C., Tsai,P., Takata,H., Mousseau,G., Ho,P.T., Honeycutt,J.B., Fallahi,M., Trautmann,L. *et al.* (2017) In vivo suppression of HIV rebound by didehydro-cortistatin A, a "Block-and-Lock" strategy for HIV-1 treatment. *Cell Rep.*, **21**, 600–611.
106. Nguyen,W., Jacobson,J., Jarman,K.E., Jousset Sabroux,H., Harty,L., McMahon,J., Lewin,S., Purcell,D.F.J. and Sleebs,B.E. (2019) Identification of 5-substituted 2-acylaminothiazoles that activate Tat mediated transcription in HIV-1 latency models. *J. Med. Chem*, **10**, 5148–5175.