doi:10.1016/j.jmb.2004.09.042 J. Mol. Biol. (2004) xx, 1–18

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Anti-termination by SIV Tat Requires Flexibility of the Nascent TAR Structure

Peter Sommer[1]*, Jean-Pierre Vartanian[1], Malte Wachsmuth[2] Michel Henry[1], Denise Guetard[1] and Simon Wain-Hobson[1]*

[1]*Unité de Rétrovirologie Moléculaire, Pasteur Institute 28 rue du Dr Roux, 75724 Paris France*

[2]*Kirchhoff Institute for Physics University of Heidelberg, Im Neuenheimer Feld 227, 69120 Heidelberg, Germany*

**Corresponding authors*

Substitution of the SIVmac239 promoter/enhancer by the strong EF1α promoter results in a severe replication defect due to a failure to respond to Tat. Revertant viruses with minimal promoter sequences (two Sp1 sites and a TATA box) were obtained that had fully restored their replicative potential. Comparison of the different LTRs indicated that structural alterations in the TAR stem due to a 31 bp exon of the EF1α promoter rather than the mere presence of transcription factor binding sites within U3 were responsible for the attenuation. Structural models based on genuine RNA sequences combined with a refined algorithm to calculate the probability of the looping-mediated interaction between protein complexes bound to nucleic acid polymers indicated that the local concentration of TAR-bound Tat close to the RNA polymerase II complex was reduced more than 100-fold for the mutant as compared to SIVmac239. These results show that HIV/SIV replication requires only a minimal set of *cis*-acting elements in the promoter and suggest a hitherto unrecognised requirement of flexibility for the nascent TAR structure to allow anti-termination by Tat.

Keywords: anti-termination; EF1α promoter; LTR; primate immunodeficiency virus; Tat-TAR axis

Introduction

A single inducible promoter located in the 5′ long terminal repeat (LTR) controls HIV/SIV gene expression, which depends in the primary phase exclusively on cellular factors. The assembly of transcription factors and components of the basal transcription machinery at their target sequences in the U3 part of the LTR initiates transcription.[1] At this stage, however, viral gene expression is rather inefficient since elongation of nascent transcripts by the RNA polymerase II complex is impaired due to the formation of a stable hairpin structure at the 5′ end of all viral RNAs, termed TAR (transactivation-response element). Short, non-polyadenylated RNAs terminating just after TAR predominate viral transcripts during the early phase.[2–4] This mechanism is reminiscent of terminator or attenuator systems found in phage and bacterial operons.[5] In particular, the process of anti-termination by the protein *N* from phage λ has frequently been proposed as a model system for transcriptional regulation of lentiviral gene expression.[6–10] The *N* protein binds to a specific hairpin structure (boxB) within the *nut* site of the nascent RNA transcript. RNA looping catalyses the formation of a complex composed of the RNA polymerase, the host factor NusA and the *N* protein.[6] The effectiveness of this process required for efficient elongation of transcripts crucially depends on the flexibility and conformation of the polynucleotide linker.[11]

Likewise, the HIV/SIV transcriptional transactivator protein Tat binds to the nascent TAR stem-loop and serves as a scaffold for the assembly of a switch to efficient viral gene expression.[12] Various complexes are recruited to the LTR resulting in a tremendous increase in proviral transcription rate. This includes enzymes with histone acetyl transferase activity modifying the chromatin environment at the integration site, and the P-TEFb complex phosphorylating the Pol II carboxy-terminal domain, thus promoting efficient

Abbreviations used: LTR, long terminal repeat; STR, short terminal repeat; TAR, transactivation-response element; FACS, fluorescence activated cell sorting; hCMV, human cytomegalovirus; GFP, green fluorescent protein; rhPBMC, rhesus peripheral blood mononuclear cells; SIV, simian immunodeficiency virus; HIV, human immunodeficiency virus.

E-mail addresses of the corresponding authors: psommer@pasteur.fr; simon@pasteur.fr

0022-2836/$ - see front matter

elongation of viral transcripts.[5,13,14] Given its fundamental influence on viral gene expression, a functional Tat/TAR axis is obviously indispensable for HIV/SIV replication and hence pathogenesis.

Extensive mutational analyses have revealed a detailed picture of structural requirements in HIV TAR that allow the functional recruitment of Tat and its associated cellular partners and this knowledge is currently exploited to develop novel drugs.[14,15] In contrast, position effects of TAR for efficient Tat-transactivation are still debated. Early observations suggested that TAR needs to be immediately adjacent to the transcription initiation site in order to allow Tat to connect to the basal transcription machinery and upstream enhancer elements.[3,16] More recently, however, it was demonstrated that TAR can mediate Tat-activated transcription over long distances.[17–19] This implied that Tat is introduced into transcription complexes during their transit through TAR, thus initiating their transformation into effective, elongation competent forms. Furthermore this implied that Tat-mediated transactivation would depend on the processivity of transcription from a given upstream promoter in the absence of Tat. As long as TAR would be positioned at a location within which elongation-incompetent complexes were found, this would be sufficient to recruit Tat and its cellular partners and to switch to efficient elongation (reviewed by Karn[14]).

The major transcriptional control elements in the core promoter/enhancer (P/E) region of HIV/SIV are a TATA box and binding sites for the transcription factors Sp1 and NFκB. Upstream of these sequences are binding sites for a variety of cellular transcription factors that may modulate promoter activity in a cell type-dependent manner.[20] The relevance of this 'modulatory' region for viral transcription has been difficult to assess *in vivo*, since it overlaps the *nef* gene, which is required for multiple steps during the viral life cycle. Yet, numerous studies indicate that this region may be dispensable for efficient promoter activity and may simply serve as Nef coding sequences.[21–25] However, they can efficiently substitute for the lack of the entire core enhancer since SIVmac239 containing deletions of all NFκB and Sp1 binding sites replicated with similar kinetics like the parental strain in lymphoid cells and caused AIDS in rhesus macaques.[26] Likewise, neither NFκB nor Sp1 binding sites *per se* are absolutely required for HIV replication.[27,28] Altogether, these observations led to the notion that the promoter of HIV/SIV possesses high functional redundancy, which allows the interaction with multiple transcription factors, thereby ensuring viral replication in a variety of cellular conditions.

In contrast, engineered modifications in the LTR can significantly change its properties without affecting efficacy. The introduction of a strong, cellular enhancer upstream of the HIV core P/E allowed highly efficient transcription independent of Tat.[29] Moreover, replacement of the functional Tat/TAR axis by the tetracycline system resulted in an HIV strain that was strictly dependent on doxycycline, thus permitting replication to be turned on and off.[30] It has also been shown that HIV/SIV variants that have their native P/E region substituted by heterologous elements are replication competent but may have altered their cellular tropism.[31,32] We could recently show that the exchange of the P/E region of SIVmac239 by the major immeadiate-early (IE) promoter from human cytomegalovirus (hCMV) resulted in a highly attenuated phenotype *in vivo* and the infected animals were able to control pathogenic challenge infections.[33] These results suggested that the exchange of the entire SIV promoter by heterologous elements has a much more profound effect on the viral phenotype than mutation of the native elements, although it could not be excluded that the observed attenuation was due to particular properties of the hCMV IE promoter.

The promoter of the human eukaryotic translation elongation factor 1 alpha 1 gene (*eEF1A-1*, formerly termed *EF1α*) shares some features with both the hCMV and HIV/SIV promoter. It contains Sp1 binding sites just upstream of the TATA box and their relative position to each other does not differ significantly from that in SIVmac239. On the other hand there are no binding sites for inducible transcription factors (e.g. NFκB) and the most widely used fragment includes a short exon of 31 bp immediately downstream of the transcriptional initiation site. This promoter element has been shown to possess strong activity comparable to the hCMV promoter in various cell types, in particular in the context of integrated proviral DNA. Furthermore, it is a considerably small element, thus avoiding packaging constraints in the context of a retroviral genome.[34,35] Here, we characterised a SIVmac239 mutant which had the native promoter replaced by the rhesus macaque (rh)EF1α promoter. The virus mutant was severely attenuated for replication in established cell lines and primary lymphocyte cultures. Revertant viruses with restored replicative potential were obtained that had reconstituted the proximal region of the native SIVmac239 promoter. The replication defect of the SIVmac239-EF1α/STR mutant resulted from an unresponsiveness to transactivation by Tat. A biophysical approach based on genuine RNA sequences indicated that the 31 bp EF1α promoter exon led to structural alterations in TAR resulting in an increased stiffness of the nascent RNA chain, which diminished the looping-mediated local concentrations of TAR-bound Tat more than 100-fold. The results show that efficient replication of SIV can be achieved with a minimal set of promoter elements but there are subtle topological and structural constraints, which crucially influence viral gene expression. We hypothesise that a previously unrecognised requirement for flexibility of the nascent RNA may be involved in the switch from basal to Tat-activated gene expression.

Results

Construction of SIVmac239 mutants

To investigate effects of heterologous promoter elements on SIV replication without possible interference of *nef*/U3 overlapping sequences we first rearranged the 3′ genomic organisation of SIVmac239 to separate the *nef* open reading frame (ORF) and proximal promoter/enhancer elements. Thus, we inactivated *cis*-acting elements required for reverse transcription and integration within *nef* (U-box, PPT and *att* site) by silent point mutations and introduced the authentic functional elements just 3′ of the *nef* termination codon (Figure 1(A)). After reverse transcription these modifications resulted in a provirus with a shortened promoter region in the 5′ LTR lacking all *nef* sequences (or a short terminal repeat, STR) but keeping the coding capacity for a functional Nef protein (Figure 1(A)–(C)). Based on this reconstructed genomic organisation we replaced the entire SIV P/E region by a 240 bp fragment (position −208 to +32 as described by Wakabatashi-Ito & Nagata[34]) of the rhesus monkey EF1α promoter, termed SIVmac239-EF1α/STR. Compared to its human homologue, the rhesus macaque sequence differed in three nucleotide positions (T_{153} to C, G_{167} to T, C_{199} to T), strongly suggesting an equivalent transcriptional activity (see also below).

Replication of SIVmac239-EF1α/STR is highly attenuated

Virus stocks of SIVmac239-EF1α/STR and the parental strain SIVmac239 were obtained from transiently transfected 293T cells while stocks of the SIVmac239-STR variant originated from transiently transfected CEMx174 cells. Production of infectious particles of the SIVmac239-EF1α/STR mutant was not affected after transfection of 293T cells as evidenced by average p27 production and titration of stocks on HeLa P4P cells by limiting dilution (data not shown). This was expected, since we modified only the 3′-LTR in the molecular cloned SIVmac239-EF1α/STR construct. Virus production is driven by the wild-type promoter in the 5′-LTR in this system while progeny virus will then harbour the modification in both LTRs.[27] To assess the replicative potential of SIVmac239-EF1α/STR we infected CEMx174 cells and monitored culture fluids for SIV p27 antigen levels. In agreement with a recent report,[25] p27 production in the supernatant of cell cultures infected with SIVmac239-STR was indistinguishable from those infected with the wild-type virus. In contrast, replication of SIVmac239-EF1α/STR was severely impaired. Levels of p27 were reduced more than 1000-fold and never showed *bona fide* peak values typical for a spreading viral infection. Essentially the same picture emerged after infection of activated primary rhesus monkey PBMC (Figure 2(A)). Following several cultures over three months did not result in an adaptation to increased replication capacity and p27 levels remained very low or became undetectable. PCR analysis of infected CEMx174 cells in order to detect 2-LTR circles indicative of recent infection events and/or ongoing viral replication showed sporadically faint positive signals after about one week of culture and we never succeeded in amplifying 2-LTR circles at later time points. Sequence analysis of the PCR products confirmed the identity of the rhesus monkey promoter sequence and revealed few mutations (T_{167} to G (4/4 clones) and C_{198} to T (1/4 clones)) most likely introduced during reverse transcription (Figure 3(C) and data not shown).

SIVmac239-EF1α/STR replication is restricted after expression of functional Tat protein

We next employed reporter cell lines, CEMx174-LTR-GFP and HeLa P4P cells, that facilitate the detection of an infection on a single cell level to extend these observations. Both cell lines harbour a reporter gene under the control of the HIV LTR that is highly expressed only in the presence of Tat. Multiple cells with bright green fluorescent protein (GFP) expression could be detected five days after infection of CEMx174-LTR-GFP cells with SIVmac239 and SIVmac239-STR by fluorescence microscopy. Moreover, cytopathic effects were already obvious as evidenced by the appearance of large cell syncytia. In contrast, only few, yet clearly positive cells are visible upon exposure to SIVmac239-EF1α/STR and cytopathic effects were difficult to assess (data not shown). Quantification of GFP expressing cells by fluorescence activated cell sorting (FACS) analysis revealed that less than 1% infected cells were present during the 15 days of observation while control cultures were characterised by a progressive increase in GFP-positive cells reaching about 75%. The percentage of GFP-positive cells at day 5 was about ten times higher for the controls and a further tenfold increase could be observed from day 5 to 10 in these cultures (Figure 2(B)).

Given that the completion of the replication cycle of SIV (from entry to release of progeny) takes about 24 hours[36] it can be assumed that five replication cycles could have been completed within five days, theoretically resulting in an up to 1000-fold increased number of infected cells. Hence, the tenfold lower percentage of infected cells in the SIVmac239-EF1α/STR cultures at day 5 is in good agreement with the assumption that the numbers of originally infected cells were similar, while a further spread of the infection was impaired. Although the frequency of infected cells was low, expression levels (as evidenced by the mean fluorescence intensity (MFI)) interestingly proved to be comparable if not superior to SIVmac239 and SIVmac239-STR infected cells during the early phase of infection (five to seven times above background; Figure 2(B)). Hence, basal transcription levels are high enough to ensure expression of sufficient functional Tat protein to boost high levels of GFP

Figure 1. Modifications introduced into the SIVmac239 genome to eliminate the *nef*/U3 overlapping region and schematic presentation of the viral constructs analysed. (A) The 239wt clone (top) contains a 517-bp U3 region that overlaps the 792-bp *nef* ORF by 407 bp (79%).[55] The *cis*-acting elements in *nef* (U-box, PPT and *att*-site) were inactivated by 14 silent point mutations (red), and the intact elements (green), including a novel MluI restriction site (in italics) as well as a second termination codon immediately downstream of *nef*, were inserted between the *nef* stop codon and the single NF-κB site (middle). The SIVmac239-STR provirus contains a U3 region of 123 bp and a *nef* gene that neither overlaps the 3′ LTR nor contains essential *cis*-regulatory elements but encodes the wild-type protein sequence (bottom). (B) Schematic presentation of the *nef*/U3 genomic organisation of the SIVmac239 constructs analysed and (C) the deduced 5′ promoter region. SIVmac239-STR and SIVmac239-EF1α/STR contain *nef* alleles with mutated *cis*-regulatory elements (abbreviated as PPT) and the intact sequences introduced between the *nef* stop codon and the single NF-κB site. After reverse transcription these changes will result in a shortened 5′ LTR (or STR, short terminal repeat) sequence. They differ in their core promoter/enhancer region being the wild-type SIVmac239 sequence (in yellow) and the minimal rhesus macaque EF1α promoter (in green), respectively. Characterised transcription factor binding sites in the EF1α promoter are named as by Wakabayashi-Ito and Nagata. Abbreviations: PPT, polypurine tract; IR, inverted repeat; RT, reverse transcription.

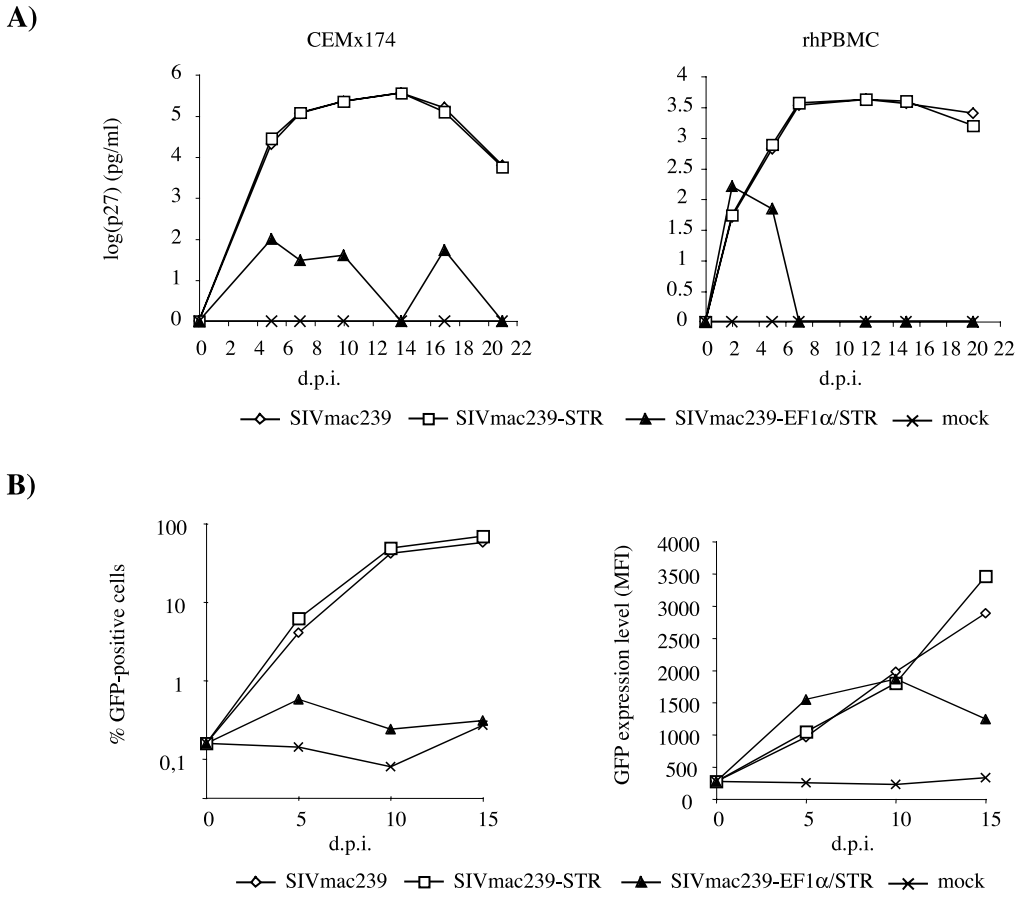
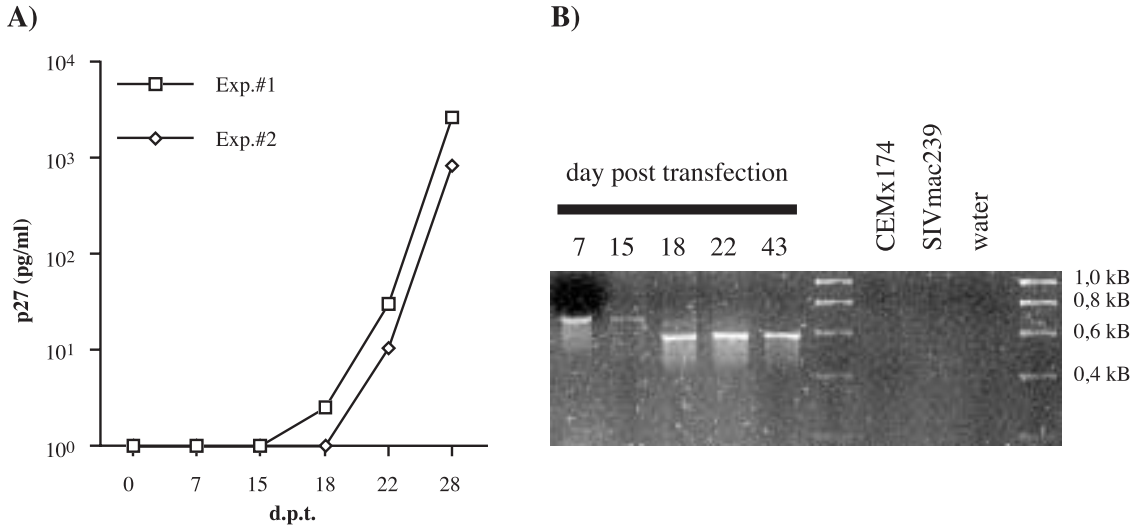


Figure 2. Replication of the SIVmac239-EF1α/STR mutant is highly attenuated. (A) p27 production of SIVmac239 variants after infection of CEMx174 cells, primary rhesus monkey peripheral blood mononuclear cells (rhPBMC) or primary human peripheral blood leukocytes (huPBL). Virus containing 10 ng of p27 was used for infection. p27 levels in supernatant were determined using a SIV Core Antigen Assay Kit, which has a detection limit of 10 pg/ml. Comparable results were obtained in three independent experiments. Abbreviations: d.p.i., days post infection. (B) FACS analysis of GFP expression after infection of CEMx174-LTR-GFP cells. Infections were performed as described above and analysed for the percentage of GFP-positive cells and their mean fluorescence intensity (MFI) at the indicated time points. The virus variants are indicated by the same symbols as in (A) and comparable results were obtained in two independent experiments.

expression mediated by the HIV LTR. Yet, there was no detectable increase in p27 levels nor in the percentage of infected cells over time (Figure 2), strongly arguing for a failure of the SIVmac239-EF1α/STR mutant to switch from basal to Tat-activated transcription. Similarly, only small numbers of single HeLa P4P cells rather weakly positive for β-galactosidase staining could be observed four days after infection with SIVmac239-EF1α/STR contrasting the numerous dark blue cells or cell syncytia present in cultures infected with the parental viruses (data not shown). Altogether, these results indicate that SIVmac239-EF1α/STR is able to infect appropriate target cells and to express a functional Tat protein, but replication occurs at too low levels to fuel a spreading infection in the cultures.

Emergence of replication competent revertants after transfection of SIVmac239-EF1α/STR DNA

The experiments described above were performed with virus stocks derived from transiently transfected 293T cells, a human embryonic kidney cell line, which cannot be infected by SIV. We also sought to produce virus in a background that allows multiple replication rounds and transfected CEMx174 cells with the molecular cloned SIVmac239-EF1α/STR construct. In contrast to our previous observations after infection, large cell syncytia became apparent after 15–18 days of culture in three of five independent samples and this correlated with a progressive increase in p27 levels (Figure 3(A) and data not shown). We analysed the promoter organisation of the emerging virus in the cultures by a PCR strategy that specifically amplified the 5′-LTR region of the SIVmac239-EF1α/STR virus, which is absent in the plasmid DNA used for transfection. A fragment of the expected size of 790 bp was detected at early time points after transfection indicating a low level of replication albeit not sufficient to result in detectable cytopathic effects or p27 levels (Figure 3(B)). Coincident with the onset of viral replication however, a deleted fragment of roughly



C)

```
huEF1alpha      1 GCATGCGTGA GGCTCCGGTG CCCGTCAGTG GGCAGAGCGC ACATCGCCCA
rhEF1alpha      1 .......... .......... .......... .......... ..........
#1(d8)          1 .......... .......... .......... .......... ..........
#1(d22)         1 .......... .......... .......... .......... ..........
#2(d22)         1 .......... .......... .......... .......... ..........

huEF1alpha     51 CAGTCCCCGA GAAGTTGGGG GGAGGGGTCG GCAATTGAAC CGGTGCCTAG
rhEF1alpha     51 .......... .......... .......... .......... ..........
#1(d8)         51 .......... .......... .......... .......... ..........
#1(d22)        51 .......... ........-- ---------- ---------- ----------
#2(d22)        51 .......... ......---- ---------- ---------- ----------

huEF1alpha    101 AGAAGGTGGC GCGGGGTAAA CTGGGAAAGT GATGTCGTGT ACTGGCTCCG
rhEF1alpha    101 .......... .......... .......... .......... ..........
#1(d8)        101 .......... .......... .......... .......... ..........
#1(d22)       101 ---------- ---------- ---------- ---------- ----------
#2(d22)       101 ---------- ---------- ---------- ---------- ----------

huEF1alpha    151 CCTTTTTCCC GAGGGTGGGG GAGAACCGTA TATAAGTGCA GTAGTCGCCG
rhEF1alpha    151 ..C....... ......T... .......... .......... ........T.
#1(d8)        151 ..C....... .......... .......... .......... ........T.
#1(d22)       151 ---------- ---------- -GGGGAGGAG CCGGTCGGGA ACGCCCACTT
#2(d22)       151 ---------- ---------- -GGGGAGGAG CCGGTCGGGA ACGCCCACTT
SIVmac239                              GGGGAGGAG CCGGTCGGGA ACGCCCACTT

huEF1alpha    201 TGAACGTTCT TTTTCGCAAC GGGTTTGCCG CCAGAACACA
rhEF1alpha    201 .......... .......... .......... ..........
#1(d8)        201 .......... .......... .......... .......... GTCGCTCTG
#1(d22)       201 TCTTGATGTA TAAATATCAC CGCATTTCGC TCTGTATTCA GTCGCTCTG
#2(d22)       201 TCTTGATGTA TAAATATCAC TGCATTTCGC TCTGTATTCA GTCGCTCTG
SIVmac239         TCTTGATGTA TAAATATCAC TGCATTTCGC TCTGTATTCA GTCGCTCTG
```

Figure 3. Emergence of replication-competent revertants after transfection of SIVmac239-EF1α/STR plasmid DNA. (A) Increasing levels of SIV p27 antigen in the supernatant of two independent CEMx174 cultures transfected with plasmid pBR239-EF1α/STR. (B) Representative PCR analysis of 5′-LTR sequences of proviruses present at the indicated time points in the culture corresponding to experiment Exp.1 in (A). The expected size of the unmodified amplified region is 790 bp. (C) Sequence analysis of cloned PCR products obtained at day 22 post-transfection. The sequences were aligned to the human (hu) EF1α promoter sequence (accession number: E02627) and the corresponding rhesus macaque (rh) sequence, the very proximal promoter sequence of SIVmac239[55] as well as a sequence obtained at day 8 after infection of CEMx174 cells with SIVmac239-EF1α/STR stock virus, termed #1(d8), are given for comparison. The sequence GGGGAG representing the recombination breakpoint in the chimeric promoter of the revertant viruses is highlighted in red while individual point mutations in the clones are shown in green.

690 bp became dominated and remained stable for up to 43 days of culture. Sequence analysis of PCR fragments obtained at day 22 from two independent cultures showed that a complex rearrangement within the U3 part of the promoter had occurred.

Most of the EF1α promoter sequence containing all described transcription factor binding motifs had been deleted and the proximal region of roughly 60 bp turned out to be identical to the SIVmac239 proximal promoter sequence including the TATA box and the first two Sp1 sites (Figure 3(C)). Two clones obtained from independent cultures differed in three nucleotides: T_{221} to C in clone 1 and a deletion of 103 bp and 105 bp, respectively. The modifications, which required a recombination event and a deletion, probably occurred simultaneously during reverse transcription. This hypothesis is supported by the fact that the sequence GGGGAG joining the EF1α and SIV part of the chimeric promoters is present in two copies in the EF1α promoter, which flank exactly the deleted region, and once in the SIV wild-type sequence. As the HIV/SIV reverse transcriptase is prone to introduce deletions between repetitive sequences and switches between the two template RNAs present in a virion it appears that an accidentally packaged RNA containing the wild-type 5′-LTR sequence derived from the transfected plasmid DNA served as the source of the SIV promoter sequence. Such RNAs may arise by rare read-through events or by transcription from the 3′-LTR present in the plasmid DNA. In fact, such a scenario is not without precedent and recombination of modified 3′-LTR sequences with wild-type 5′-LTR sequences after inoculation of molecular cloned SIV DNA have been described *in vitro* and *in vivo*.[37] Whatever the explanation for the complex reorganisations within the promoter may be, they importantly resulted in replication competent revertants that harbour extremely minimal promoter sequences essentially reduced to a TATA box and two Sp1 sites.

Revertant viruses with minimal promoter sequences efficiently replicate in CEMx174 cells and rhPBMC

Next, we used DNase-treated, cell-free supernatants from the cultures to infect fresh CEMx174 cells and prepared titrated stocks, which were subsequently used to infect CEMx174 cells and primary, activated rhPBMC. These experiments revealed that the revertants, although harbouring extremely minimal promoter sequences, replicated indistinguishable from the controls in both CEMx174 cells and rhPBMC (Figure 4(A)). Since no transcription factor binding sites have been described in the remaining EF1α sequence of the promoter chimeras present in these viruses, these observations suggest that, at least *in vitro*, efficient viral replication can be achieved with limited proximal promoter sequences. We also amplified 2-LTR circles from DNA obtained from infected CEMx174 cultures (Figure 4(B)). This PCR strategy amplifies the entire U3 region of the replicating virus and the results thus confirmed that the minimal promoter sequences as well as the rearranged 3′ genomic organisation remained stable during prolonged culture. HIV/SIV transcription depends on T cell activation and this has been correlated to the availability of inducible transcription factors, in particular NFκB. The U3 region of the revertant viruses contains only two proximal binding sites for the constitutive transcriptional activator Sp1, which is sufficient for efficient replication. Hence, we considered the possibility that these viruses could be able to replicate in unstimulated primary lymphocytes. As exemplified for SIVmac239-EF1α/STR-R1, no growth advantage of the revertant virus compared to the controls could be observed after infection of freshly isolated, unstimulated rhPBMC or human peripheral blood lymphocytes (Figure 4(C)). Hence, the failure of SIV to replicate in resting T cells does not appear solely restricted by the availability of cellular transcription factors.

The LTR of SIVmac239-EF1α/STR is unresponsive to Tat transactivation

Next, we focused on reporter assays to examine the replication defect of the SIVmac239-EF1α/STR mutant. To this end, we introduced the various promoter elements into the backbone of a self-inactivating (SIN) HIV-based lentiviral vector in front of the *GFP* gene (Figure 5(A)). VSV-G pseudotyped particles produced in 293T cells were employed to infect different cell lines and GFP expression was monitored four days later by immunofluorescence microscopy and FACS analysis. As shown in Figure 5, all promoters used demonstrated comparable basal transcriptional activity (Figure 5). Moreover, transduction of HeLa-Tat or Jurkat-Tat cells confirmed that both, the human and the rhesus macaque EF1α promoter, were not influenced by the presence of HIV-1 Tat as expected, while the activity of the SIVmac239/STR clearly increased due to the presence of TAR. The promoter of the SIVmac239-EF1α/STR revertant was equally induced by Tat, indicating that the presence of two Sp1 sites and a TATA box are sufficient to mediate the switch from basal to Tat-dependent transcription. In sharp contrast, no increase in transcriptional activity of the SIVmac239-EF1α/STR in the presence of Tat could be detected (Figure 5). These results clearly show that the replication defect of the SIVmac239-EF1α/STR mutant stems from a failure to respond to Tat despite the presence of essentially identical *cis*-acting promoter elements located at comparable positions.

Influence of the EF1α exon on the TAR structure in the SIVmac239-EF1α/STR mutant

Apart from the mere presence of promoter elements in U3 and their relative position to each

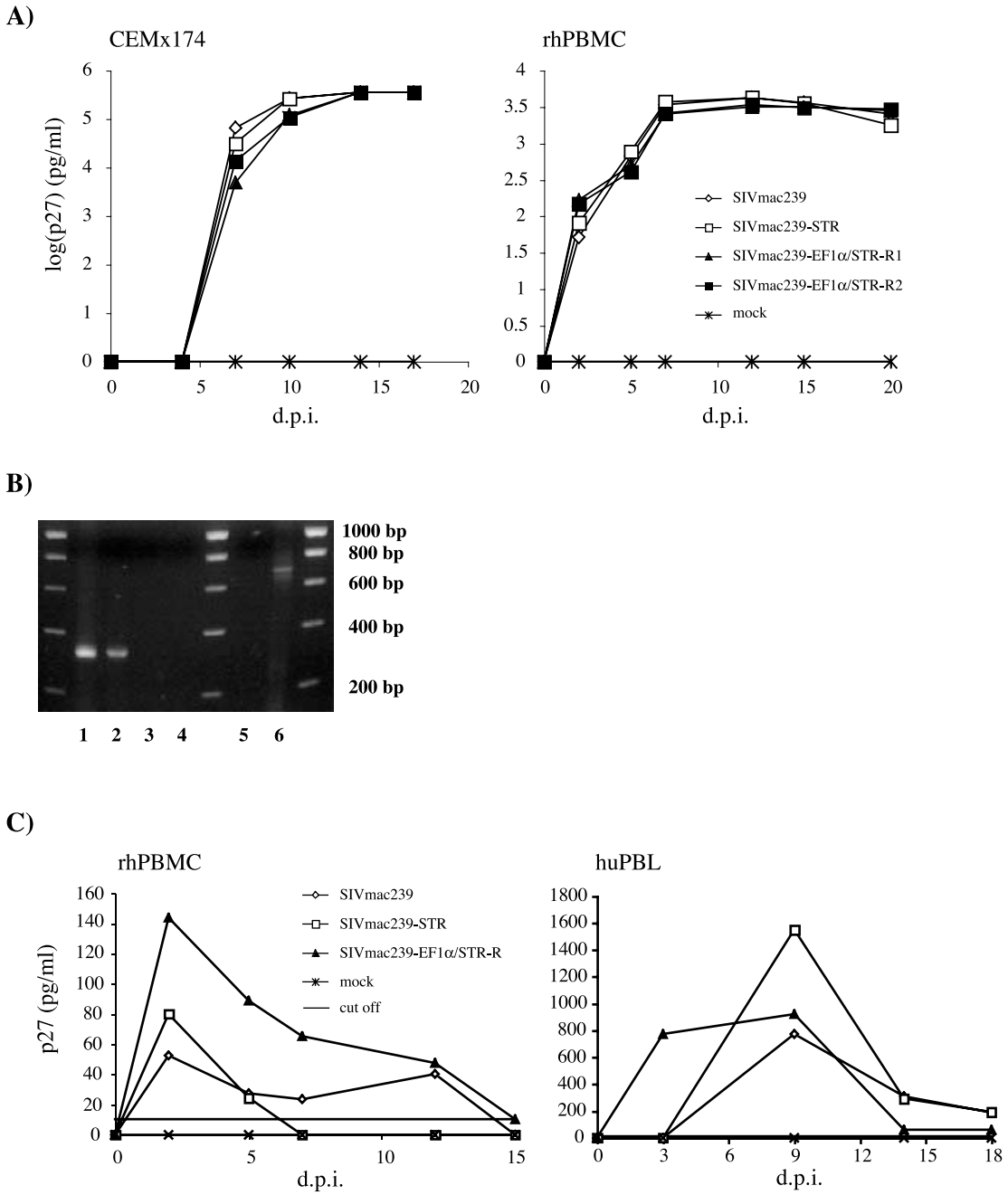


Figure 4. SIVmac239-EF1α/STR revertants are replication-competent and genetically stable. (A) Replication of SIVmac239-EF1α/STR revertants in CEMx174 cells and activated primary rhPBMC. Cell-free, DNase I-treated virus containing 10 ng of p27 was used for infection. p27 levels in supernatant were determined using a SIV Core Antigen Assay. Comparable results were obtained in three independent experiments. (B) SIV two-LTR circles were amplified from CEMx174 cultures infected 16 days before. PCR products were separated by electrophoresis through 2% agarose gels. (1) SIVmac239-EF1α/STR-R1; (2) SIVmac239-EF1α/STR-R2; (3) blank; (4) uninfected CEMx174 cells; (5) water control; (6) SIVmac239. (C) Replication of SIVmac239-EF1α/STR revertants in unstimulated rhPBMC and huPBL. Cell-free, DNase I-treated virus containing 10 ng of p27 was used for infection. p27 levels in supernatant were determined using a SIV Core Antigen Assay. Since both revertant viruses showed almost identical replication kinetics only results obtained with SIVmac239-EF1α/STR-R1 are illustrated. Comparable results were obtained in three independent experiments.

other, Tat-mediated transactivation obviously critically depends on the integrity of its target sequence TAR. Based on the demonstration that TAR can mediate Tat-activated transcription when positioned more then 500 bp downstream from the site of transcriptional initiation,[17–19] a short exon immediately downstream of the transcriptional initiation site of the EF1α promoter was maintained in the SIVmac239-EF1α/STR mutant (Figure 6(A)). Since this constitutes the most pronounced distinction between the different promoters, we reasoned that the 31 bp insertion might affect TAR, the

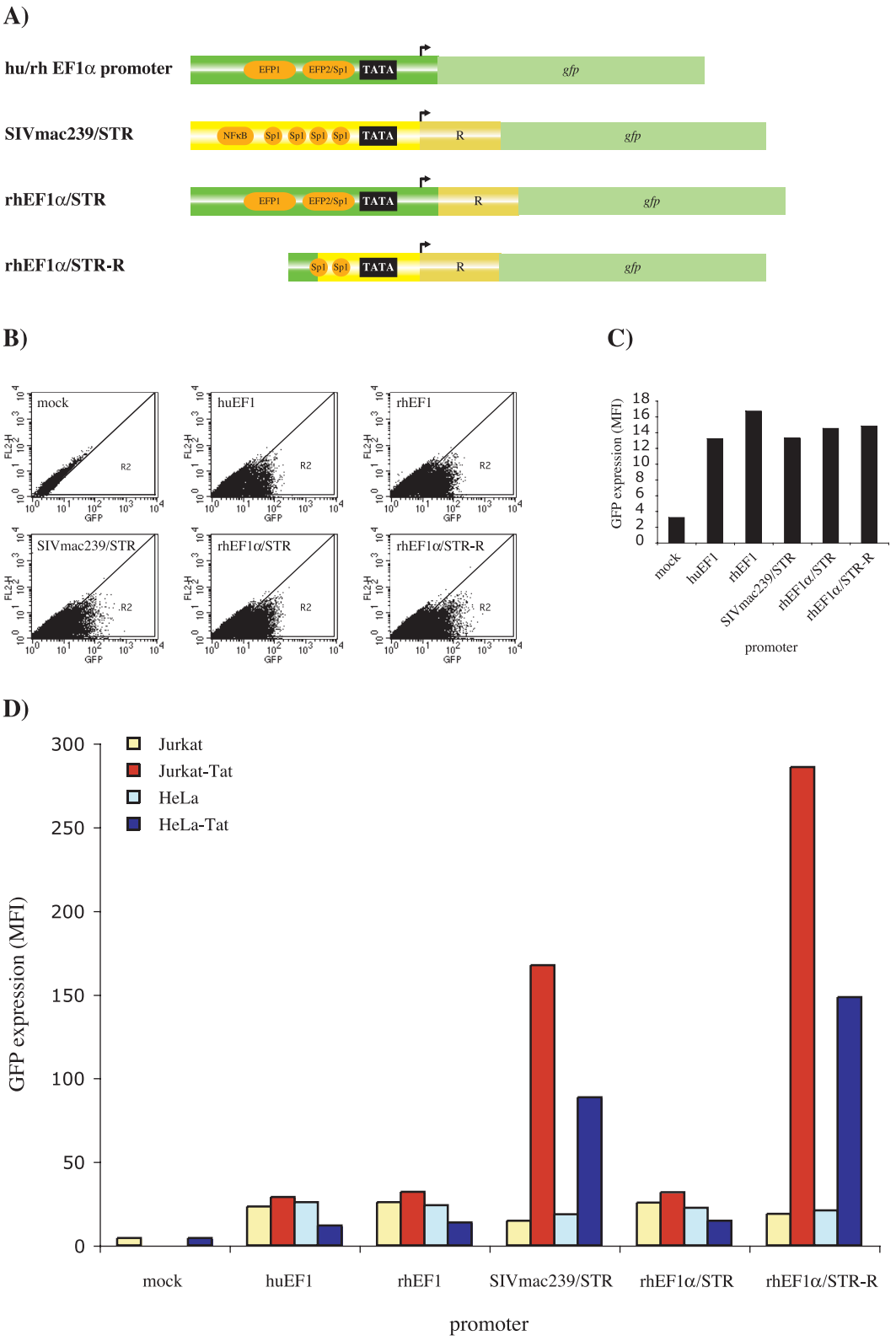


Figure 5. Substitution of the entire U3 region by the rhEF1α promoter results in an unresponsiveness to Tat-mediated transactivation. (A) Schematic presentation of the various promoters present in the HIV-based SIN vector and driving GFP expression. (B–D) Flow cytometry analysis of basal and Tat-dependent promoter activity. (B) CEMx174 cells were transduced with HIV-based SIN vector particles (corresponding to 100 ng p24) harbouring the *egfp* gene under the control of the indicated promoters and prepared for FACS analysis four days later. FL2H is the blank channel used to measure autofluorescence. (C) Bar graph illustration of the mean fluorescence intensity (MFI) of R2-gated cells shown in (A). (D) Summary of the MFI of GFP-positive cells determined by FACS analysis after transduction of Jurkat/Jurkat-Tat and HeLa/HeLa-Tat cells as described above. Representative results of at least two independent experiments are shown.

element most important for Tat responsiveness. We determined the transcription start site of the SIVmac239-EF1α/STR mutant by 5′-RACE (rapid amplification of cDNA ends) and sequence analysis. The results showed that only one transcript of the expected size could be detected, ruling out that insertion of the EF1α promoter in the SIV LTR resulted in the utilisation of cryptic or

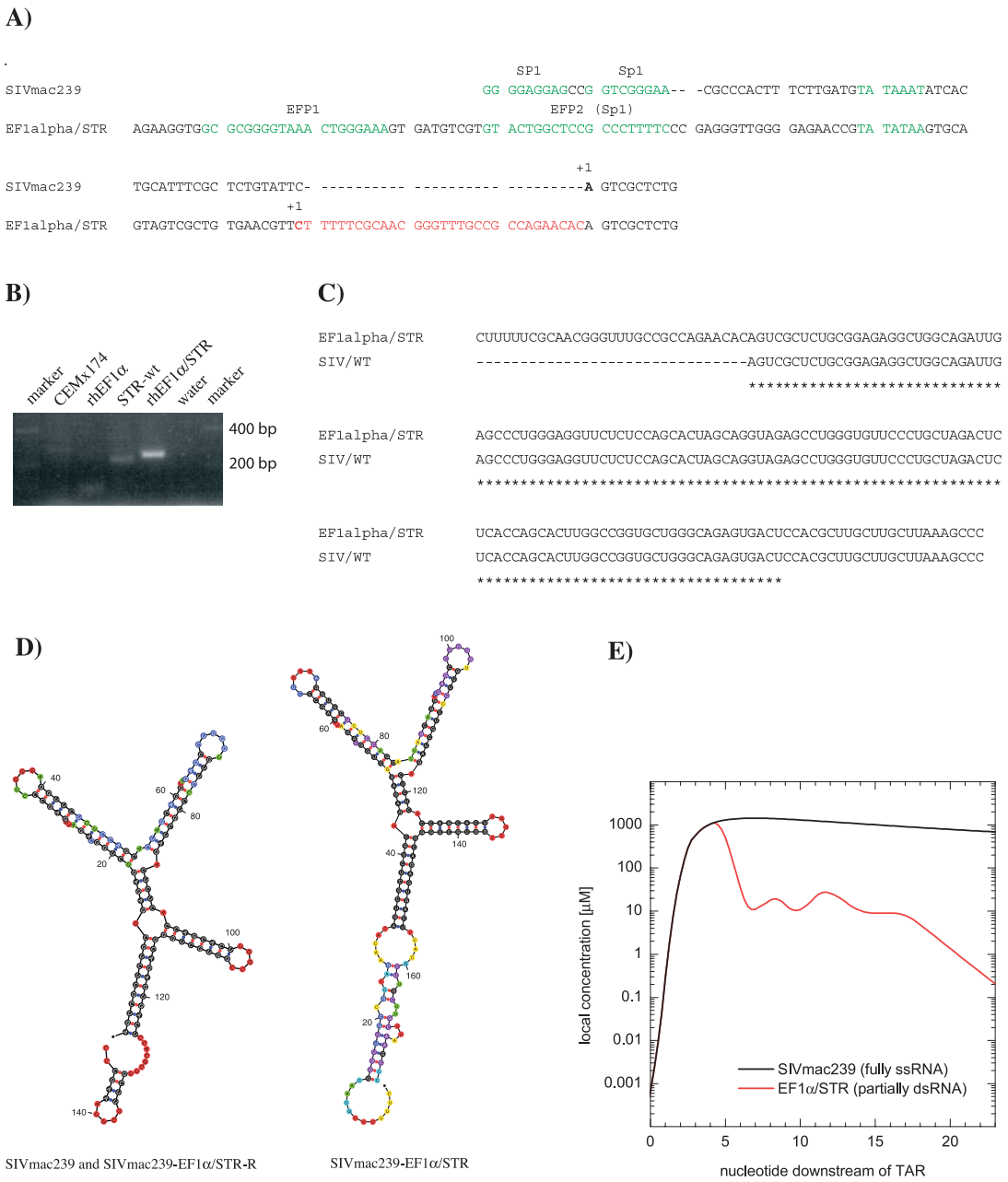


Figure 6. A short exon of 31 bp immediately downstream of the transcriptional initiation site alters structure and flexibility of TAR in the SIVmac239-EF1α/STR mutant. (A) Alignment of the proximal promoter/TAR region in SIVmac239 and SIVmac239-EF1α/STR. The relevant elements in the promoter region (e.g. transcription factor binding sites, which have been termed EFP1 and EFP2 in the case of the EF1α promoter by their discoverers[34] and the TATA box) are shown in green, while the exon of 31 bp present in the EF1α promoter is highlighted in red. The first nucleotide of the RNA transcripts is indicated by +1. (B) Agorose gel electrophoresis of 5′-RACE products. The expected size for the rhEF1α, STR-wt and rhEF1α/STR transcripts amplified are 108, 229 and 260 bp, respectively. (C) RNA sequences of the 5′-ends of transcripts originating from SIVmac239 and SIVmac239-EF1α/STR. The SIV TAR sequence is indicated by asterisks. (D) RNA structural prediction of the folding of wild-type and mutant TAR region using MFOLD. The structures are annotated using coloured base characters based on single strand count (ss-count) information. Thus base-pairings most likely to be single-stranded are drawn in red while those most likely to be double-stranded are shown in black. The detailed ss-count colour scheme can be found at http://www.bioinfo.rpi.edu/~zukerm/articles/s-annotate/colors.html. (E) Local concentration of TAR-bound Tat as a function of the neo-synthesised nucleotide downstream of TAR.

unexpected transcriptional start sites (Figure 6). Cloning of the 5′-RACE products followed by sequence analysis confirmed that only the described start site of the EF1α promoter[34] was used in all 29 clones examined. Most importantly, the TAR element was present in all clones as expected (data not shown).

Previous investigations had shown that insertions of 10–88 bp between the transcription start site and HIV-1 TAR diminish Tat-dependent transcription to levels ranging from 90% to 20% of wild-type levels.[3,38] Others described constructs with insertions between 25 bp up to some 500 bp that showed no defects in Tat-dependent transcription at all.[17–19] Hence we reasoned that spacing constraints alone are likely not sufficient to explain the essentially complete unresponsiveness of the SIVmac239-EF1α/STR mutant to Tat and thus explored the possibility that structural alterations also may have occurred. To this end, we performed a RNA secondary structure prediction analysis using the MFOLD 3.0 software†.[39] This indicated that the proximal 12 bases of the EF1α exon displace TAR from the transcription initiation site and the distal 19 bases fold back to the SIV RNA, resulting in an extended stem structure of TAR (Figure 6(D)). The free energy (ΔG) of the entire TAR structure changes from −65.3 kcal/mol in the wild-type SIV to −76 kcal/mol in the SIVmac239-EF1α/STR mutant. The additional stem structure itself has a ΔG of −7.1 kcal/mol and increases ΔG for the lower TAR stem from −12.8 kcal/mol to −22.6 kcal/mol. The native structural features of SIV TAR, however, are preserved in the predicted structure, suggesting that Tat as well as all necessary cellular cofactors could still be recruited.

Discussion

Heterologous exchange of the SIV promoter may constitute a novel means of attenuation as evidenced by our recent demonstration that the substitution of the P/E region by the strong, major IE promoter from human cytomegalovirus (hCMV) resulted in a highly attenuated phenotype *in vivo*.[33] While this virus mutant, termed SIVmegalo, was able to establish a persistent non-pathogenic infection in rhesus monkeys importantly resulting in the induction of protective immune responses against pathogenic challenge infections, it remained possible that the observed phenotype was due to particular properties of the hCMV IE promoter. Our results with the SIVmac239-EF1α/STR mutant now extend the concept of SIV promoter exchange as a means of attenuation to another strong, constitutive promoter of cellular origin.

On the other hand, viruses with minimal promoter sequences (two Sp1 sites and a TATA box) were recovered that had fully restored their replicative potential. Of note, the promoters present in the SIVmac239-EF1α/STR revertants are among the most minimal U3 sequences so far described that allow efficient replication of SIV variants in cell culture.[40] They have a length of 135 bp and 137 bp for the two variants, respectively, that are composed of 66 bp and 68 bp derived from the rhEF1α promoter with no described *cis*-acting elements. Inspection of these sequences with the matrix search program MatInspector V2.2‡ resulted in the prediction of potential binding sites for AP2 and TCF11 with rather low matrix similarity values of 0.866 and 0.852, respectively. This suggests that the remaining proximal 69 bp harbouring two Sp1 sites and the TATA box derived from SIVmac239 are sufficient to support efficient replication, at least *in vitro*. Although the U3 region of these viruses contains only binding sites for the constitutive transcriptional activator Sp1, no significantly increased potential to replicate in unstimulated primary lymphocytes could be noted. This suggests that the failure of SIV to replicate in resting T cells might not only be restricted by the availability of cellular transcription factors. Alternatively, the P-TEFb complex might be limiting in resting cells[13,41] or inducible transcription factor binding sites located downstream of the transcription initiation site in the U5 region of the LTR may also be required for effective replication.[1,42,43]

In contrast to SIVmegalo, which is highly attenuated *in vivo* but replicates well compared to the wild-type virus *in vitro*,[33] replication of the SIVmac239-EF1α/STR mutant turned out to be already strongly curtailed in various cell culture systems due to a failure to respond to Tat. Yet, it is noteworthy that the SIVmegalo promoter is a weak Tat responder, too, even though the TAR sequence was not disturbed. Spontaneous deletion of a large part of the SIVmegalo promoter improved the resultant viral replication and the deleted promoter became highly responsive to Tat. These deletions were not in the core promoter but in a part of the enhancer region, which has been shown to negatively affect transcription from the CMV promoter (as discussed by Blancou *et al.*). Presence and relative position of the core promoter elements (e.g. Sp1 binding sites, TATA box) in SIVmegalo, SIVmegaloΔclone61, SIVmac239-EF1α/STR and SIVmac239 are all very similar. Yet, only the SIVmac239-EF1α/STR shows the severe defect in Tat responsiveness. We have confirmed that only the described start site of the EF1α promoter is used in the context of the virus variant and that the TAR element is present as expected (Figure 6 and data not shown). This is also important since the 5′ terminal oligopyrimidine tract of the EF1α promoter, which is present in the SIVmac239-EF1α/STR transcripts, has been shown to function as a transcriptional initiator.[44] It has been shown that the sequence G^{+1}GGTCT represents the core of a HIV-1 initiator element (INR) required

† http://bioweb.pasteur.fr

‡ http://transfac.gbf.de/TRANSFAC/index.html

for Tat-responsive transcription.[18] This sequence differs significantly from mammalian INR elements and is surprisingly not conserved in SIVmac239 (the corresponding sequence is C^{+1}AGTCG). To our knowledge, functional data on the presence of a SIV INR element comparable to HIV-1 are unfortunately not published. Work in our laboratory, however, has shown that replacement of the SIVmac239 core promoter/enhancer region by the corresponding elements of HIV-1 clade B, C and E allowed replication indistinguishable from the parental strain (M. Centlivre *et a*l., unpublished results). These viruses harbour the HIV-1 promoter sequences from nt −147 to −1 directly fused to the SIVmac239 TAR sequences and hence lack the core HIV-1 INR element. This may indicate that the SIV sequence C^{+1}AGTCG constitutes a functional INR element that can substitute for the HIV-1 INR element. On the other hand, SIV transcription may be less dependent on a functional INR element.

Independent of the functional presence/importance of a potential INR element in SIVmac239 it is important to emphasise that mutations in the HIV-1 core INR sequence reduced transcription initiation rates dramatically and substantially affected basal transcription levels.[18] In contrast, our results clearly demonstrate very similar expression levels of SIVmac239/STR and the SIVmac239-EF1α/STR mutant in the absence of Tat (Figure 5). Therefore, we think that it is rather unlikely that the unresponsiveness of the SIVmac239-EF1α/STR mutant to Tat is related to deleterious changes in a potential INR element within the SIVmac239 promoter. In conclusion all promoter elements required for efficient initiation and basal expression are present and properly located in the SIVmac239-EF1α/STR mutant (see Figure 6), which is supported by indistinguishable basal expression levels compared to the parental strain (Figure 5). The replication defect of the SIVmac239-EF1α/STR mutant is most likely linked to the essentially complete Tat unresponsiveness.

Thus the 31 bp exon in the SIVmac239-EF1α/STR mutant appears to be the most logical candidate to explain the observed deleterious effects. A simple explanation could be the differential TAR positioning that is detrimental to HIV transcription.[3,38] On the other hand, there is evidence that the TAR region of HIV-1 can be functional when displaced downstream of the start of transcription.[17–19] In these studies, TAR was separated from the transcription initiation site from 25 bp up to >500 bp without substantial effects on Tat-dependent transcription. Moreover, Selby *et al.* showed a sequential decrease of Tat-dependent transcription with increasing length of the inserted sequences (ranging from 11 bp to 88 bp), but even the longest insertion retained about 20% of Tat-dependent activity. Similarly, all constructs (insertions of 10–63 bp) described by Berkhout and Jeang showed Tat-dependent activities in the range of 15–60% of the wild-type. Interestingly, significant differences were observed for insertions of identical length but different sequences (e.g. a 21 bp insert caused an 80% defect in one but only a 40% defect in a second case). Hence, there is convincing evidence that spacing constraints alone may not be sufficient to explain the essentially complete Tat unresponsiveness of the SIVmac239-EF1α/STR mutant.

The Tat/TAR axis has many features similar to anti-termination mediated by the phage λ*N* protein, which involves the formation of a complex composed of the RNA polymerase, the host factor NusA and the *N* protein. The effectiveness of this process is catalysed by RNA looping and crucially depends on the flexibility and conformation of the polynucleotide linker.[9,10] It can be quantified by the concept of the local concentration of one molecule in the vicinity of the other.[11]

An interesting hypothesis emerging from our data is that, in analogy to the lamba system, RNA looping could also be involved in Tat-mediated transcription. As a first step to support this hypothesis we calculated the local Tat concentrations based on the predicted RNA structures of SIVmac239 and the SIVmac239-EF1α/STR mutant (Figure 6). To accommodate for the variations in the polymer properties for both the EF1α/STR mutant and the SIVmac239 RNA we developed an iterative algorithm to determine the effective physical properties for a heterogeneous polymer and the corresponding looping behaviour. As shown in Figure 6(E), the local concentration of TAR-bound Tat near the polymerase increases rapidly during transcription for the wild-type sequence running to a plateau of above 1 mM. For the EF1α/STR mutant sequence it follows the same path but drops significantly after the internal loop is closed at nt 130, then running into values around 20 μM and after formation of the rather rigid structure downstream of nt 142 into values below 1 μM, two to three orders of magnitude smaller. Assuming a transcription rate of $\sim 20\,\mathrm{nt\,s^{-1}}$,[46] this occurs within less than a second. With the further assumption that the Tat concentration is at the level of the dissociation constant $K_{\mathrm{diss}} \sim 40$ nM or below and that for the reported maximum transcription rate of $1000\,\mathrm{h^{-1}}$ (Ref. 36) the off rate of Tat from TAR is the same ($k_{\mathrm{off}} \sim 0.28\,\mathrm{s^{-1}}$) the association probability of Tat to TAR is ∼1/4 per second. Hence, the sequence taken into account here is fully synthesised when a Tat molecule binds to TAR, and the final concentrations can be used to estimate the dissociation constant of Tat to the polymerase to be between 1 μM and 1000 μM. These calculations suggest that the inherent flexibility of the nascent TAR RNA may catalyse the looping-mediated connection of Tat and its associated factors to the basal transcription machinery allowing anti-termination, efficient enhancement of promoter activity and gene expression. Thus, an intrinsic flexibility of the nascent TAR RNA may be required to provide high local concentrations of Tat and its associated cellular factors at the promoter. This is particularly intriguing in the case of P-TEFb, which is a limiting transcription factor.[13,47–49] It is

also interesting to note that the TAR position immediately adjacent to the start of transcription as well as a stretch of unpaired, single-stranded nucleotides immediately downstream of the nascent TAR hairpin is a conserved feature of HIV/SIV and related lentiviruses.[50–52]

In conclusion our results demonstrate that efficient replication of SIV can be achieved with an astonishingly minimal set of *cis*-acting elements in the promoter. At least *in vitro*, only two Sp1 sites, a TATA box and the TAR element are sufficient to allow basal and Tat-enhanced expression levels, resulting in replication kinetics indistinguishable from the wild-type virus. It is important to note that not only the mere presence of these elements determines promoter activity but that there are subtle topological and structural constraints, which crucially influence viral gene expression.[3,53] Our data are consistent with the notion that Tat acts as an anti-terminator and, as the phage λ *N* protein, uses RNA looping to make contact with the transcription complex.[6,8] A detailed analysis of the limitations for the assembly of highly processive transcription complexes at the LTR should help to identify functional mechanisms regulating gene expression in the context of a higher order structure and may help to understand the regulation of HIV/SIV latency.

Materials and Methods

Plasmids and recombinant DNA work

The plasmids p239SpSp5′, p239SpE3′ harbouring the 5′ and 3′-half of SIVmac239, respectively, as well as plasmid pBR239 containing the full length molecular cloned SIVmac239 DNA have been described.[22,54,55] To generate a SIVmac239 variant without *nef*/U3 overlapping sequences we amplified the *env*-3′ LTR region in three steps using mutated oligonucleotides to introduce the desired modifications and restriction sites. A detailed description of the introduced sequence modifications is illustrated in Figure 1(A). Briefly, the *env-nef* region of SIVmac239 was amplified using primers 239(8998)+RI and 239/9680-PPT (sequences of oligonucleotids are given in Table 1). The *nef*/U3 overlapping region was amplified employing primers 239/935-Xho and 239/650-Mlu. The minimal LTR region including the core promoter/enhancer and extending to the end of U5 was amplified with primer STR-Mlu and 239-NSE. The PCR products were gel purified and directly cloned into the vector pCR2.1-TOPO included in the TOPO TA Cloning Kit (Invitrogen Corp., San Diego, CA). In order to assemble the three fragments, the pTOPO constructs were digested with the appropriate restriction enzymes (NheI/XhoI, XhoI/MluI, MluI/EcoRI, respectively). If not indicated, all enzymes used were purchased from New England Biolabs, Beverly, USA. The fragments were gel purified, mixed in equimolar amounts and subjected to a brief ligation reaction. This ligation mixture was then employed in a PCR reaction to amplify the entire, modified 3′ region using primers 239(8998)+RI and 239-NSE. The purified PCR product was subcloned into pCR2.1-TOPO and sequenced to confirm that only the intended changes were present. Next, the fragment was subcloned into the plasmid p239SpE3′ using unique NheI and EcoRI sites resulting in the plasmid pSTRwt-3′.

To replace the entire SIV P/E region by the rhesus monkey EF1α (rhEF1α) promoter, we first amplified a 424 bp fragment with primer EF1-5′ and EF1-3′ from rhesus monkey PBMC DNA. The PCR product was cloned into pCR2.1-TOPO and sequenced to allow the design of specific primer for amplification of the desired promoter region. The rhEF1α sequence was then directly fused to the R-U5 region of SIVmac239 by spliced overlap extension PCR. In short, 240 bp of the rhEF1α promoter was amplified with primers EF1-Xma and EF1-P2R. The SIVmac239 R-U5 sequence was amplified with primers EF1-P3F and 239-NSE. The left- and right-half PCR products were gel purified, mixed in equimolar amounts, and subjected to a further PCR with primers EF1-Xma and 239-NSE. The purified PCR product was subcloned into pCR2.1-TOPO and sequenced to confirm that only the intended changes were present. Next, this fragment was introduced into plasmid pUC-Bbox, a "shuttle" construct that harbours the U-box, PPT and *att* site of SIVmac239 followed by the minimal core promoter/enhancer region of HIV-1_{LAI} fused to the R-U5 region of SIVmac239. Moreover, appropriate restriction sites (e.g. MluI, XmaI and EcoRI) are present that facilitate the exchange of the minimal promoter sequences in the context of SIVmac239/STR. A detailed description of this plasmid can be provided by the authors upon request. The MluI/EcoRI fragment from the pUC-EF1α/STR construct was then cloned into pSTRwt-3′ replacing the native core P/E sequences of SIVmac239 by the rhEF1α promoter and generating plasmid pEF1α/STR-3′. Finally, the NheI/EcoRI fragment derived from pEF1α/STR-3′ was employed to replace the corresponding sequence in pBR239, resulting in the full-length molecular clone pBR239-EF1α/STR.

The HIV-based SIN vector construct TRIPΔU3-EF1α encoding GFP under the control of the human EF1α promoter has been described.[35] The MluI/BamHI promoter fragment of this construct was replaced by the following sequences: (i) rhEF1α promoter, (ii) SIVmac239-STR, (iii) EF1α/STR, (iv) EF1α/STR-R1 and (v) EF1α/STR-R2. These sequences were obtained by PCR using pBR239-EF1α/STR ((i) and (iii)), pBR239 (ii) and pTOPO constructs harbouring EF1α/gag PCR fragments (see below) from revertant viruses present at day 22 after transfection of CEMx174 cells ((iv) and (v)), respectively. The primers used were Mlu-EF1, Bam-EF1, Mlu-STR and Bam-TAR, respectively. Before cloning into TRIPΔU3-EF1α the purified PCR products were subcloned into pCR2.1-TOPO and sequenced.

Cells, virus stocks and infections

CEMx174, Jurkat and 293T cells were purchased from the American Type Culture Collection (ATCC). Jurkat-Tat and HeLa-Tat cells were obtained from the NIH AIDS Research and Reference Program, while HeLa-P4P cells were a kind gift from P. Charneau (Pasteur Institute, Paris, France). CEMx174-LTR-GFP cells are a clonal derivative of CEMx174 cells transduced with HIV-based retroviral vector particles encoding eGFP under the control of the HIV-1_{LAI} LTR (kindly provided by A. Boese, Pasteur Institute, Paris, France). These cells are characterised by low basal GFP expression, which drastically increases upon challenge with SIVmac239 due to transactivation of the HIV-1_{LAI} LTR by SIV Tat (P.S., unpublished results). CEMx174, CEMx174-LTR-GFP, Jurkat and Jurkat-Tat cells were cultured in RPMI 1640 (Gibco-BRL, Grand Island,

Table 1. Oligonucleotides used in this study

Name	Sequence (5′→3′)
239(8998)+RI	GGACGGAATTCAATGCTAGCTAAGTTAAGG
239/9680-PPT	ATACCCTCGAGACCACCTTTCTCCTTAATGAAGTGAGACATGTCTATTG
239/935-Xho	GGTGGTCTCGAGGGTATATACTATAGTGCAAGAAGACATAG
239/650-Mlu	ATAAAATACGCGTTATTAGCGAGTTTCCTTCTTGTCAG
STR-Mlu	GCTAATAACGCGTATTTTATAAAAGAAAAGGGGGGACTGGAAGGGATTTAAACAGCAGGGACTTTCCA CAAG
239-NSE	ACTCGGAATTCGTCGACGGCGCCAATCTGCTAGGGATTTTCCTGCTTCGG
EF1-5′	GTAGAACCCAGAGATCGCTG
EF1-3′	CTCGGGATCAAGAATCACGT
EF1-Xma	CCCGGGCATGCGTGAGGCTCCGGTG
EF1-P2R	CAGCCTCTCCGCAGAGCGACTGTGTTCTGGCGGCAAAC
EF1-P3F	GTTTGCCGCCAGAACACAGTCGCTCTGCGGAGAGGCTG
Mlu-EF1	ACGCGTCATGCGTGAGGCTCCGGTGCCCGT
Bam-EF1	GGATCCTGTGTTCTGGCGGCAAACCCGTTGC
Mlu-STR	ACGCGTGCAGGGACTTTCCACAAGGGGATG
Bam-TAR	GGATCCTGAAGAGGGCTTTAAGCAAGCAAGC
U5out	AGCCGCCGCCTGGTCAACTC
R2Cout	CACCCAGGCTCTACCTGCTAG
U5/2C	TGTTAGGACCCTTTCTGCTTTGGG
R/2C-rev	ATCTGCCAGCCTCTCCGCAGAG
gag1201	TCTTCCTCTCCCGCTCCTCACGCCGT

NY) supplemented with 10% (v/v) foetal calf serum (FCS, Gibco-BRL), penicillin (50 units/ml) and streptomycin (50 μg/ml). HeLa-P4P, HeLa-Tat and 293T cells were cultured in DMEM (Gibco-BRL, Grand Island, NY) supplemented with 10% FCS, penicillin (50 units/ml) and streptomycin (50 μg/ml). Primary PBMC were obtained from fresh blood samples by Ficoll-Hypaque. Cells were maintained in RPMI 1640 supplemented with 10% FCS, penicillin (50 units/ml) and streptomycin (50 μg/ml). Cells were activated by addition of 2 μg/ml phytohemagglutinin (PHA; SIGMA-ALDRICH, Saint-Quentin Fallavier, France) for the first two days of culture after which 50 units/ml of human recombinant IL-2 (GIBCO-BRL Invitrogen corporation) were supplied for the remainder. To obtain PBLs, monocytes were removed from freshly isolated PBMC by plastic adherence for two hours at 37 °C.

SIVmac239 and SIVmac239/STR stocks were generated by electroporation of CEMx174 cells. For these experiments, 10 μg of each 5′ and 3′ SIV sequences were digested with the restriction enzyme SphI (New England Biolabs, Beverly, MA). The resulting fragments were joined with T4 DNA ligase (New England Biolabs) prior to electroporation of 2×10^7 CEMx174 cells at 250 V and 950 μF using a Bio-Rad Gene Pulser II (BioRad, Marnes-la-Coquette, France). Transfected CEMx174 cells were further cultured in complete medium and virus-containing supernatants were harvested at or near the peak of virus production as evidenced by the appearance of cytopathic effects. Alternatively, 10^6 293T cells were transfected with 5 μg of plasmids pBR239 and pBR239-EF1α/STR by Fugene6 (Roche Molecular Biochemicals, Indianapolis, USA) as recommended by the manufacturer and virus-containing supernatants were collected three days after transfection. Supernatants were treated with 10 units/ml of DNase I (Amersham Biosciences Europe GmbH, Orsay, France) in the presence of 10 mM $MgCl_2$ at 37 °C for one hour to remove residual plasmid DNA. The culture fluids were filtered (0.45 μm) and aliquots were stored at −80 °C. SIV p27 antigen levels were quantified using a SIV Core Antigen Assay Kit (Coulter IMMUNOTECH S.A., Miami, FL USA).

Infections were generally performed by spinoculation essentially as described.[56] Briefly, 1×10^6–2×10^6 cells were resuspended in a small volume of virus-containing supernatant corresponding to 10 ng of p27 core antigen. After centrifugation for two hours at 1200*g* and 25 °C, cells were extensively washed with phosphate buffered saline (PBS) to remove residual viral particles and further cultured in the appropriate medium. To assess the replication potential of the viruses, culture fluids were monitored for SIV p27 antigen levels as described above.

PCR analysis, 5′-RACE and sequencing

2-LTR was amplified using primers U5out and R2Cout in the first round and U5/2C and R/2C-rev in the second round. PCR was performed in 50 mM Tris–HCl (pH 8.0), 50 mM KCl, 2.5 mM $MgCl_2$, 2.5 units of Taq polymerase, 200 μM of each deoxynucleotide triphosphate, 100 pmol of each primer. Cycling parameters were 95 °C for ten minutes, 35 cycles at 95 °C for 30 seconds, 60 °C for 30 seconds and 72 °C for 45 seconds each; 72 °C for ten minutes. Aliquots of 2 μl of the first PCR were employed for nested amplification using the same cycling conditions except that the annealing temperature was 55 °C. Specific amplification of the 5′-EF1α/STR sequence was achieved by a single PCR using primers EF1-Xma and gag120. Cycling parameters were 95 °C for ten minutes, 35 cycles at 95 °C for 45 seconds, 60 °C for 30 seconds and 72 °C for 45 seconds each; 72 °C for ten minutes. Amplification of 5′ cDNA ends was achieved with the GeneRacer Kit (Invitrogen) as recommended by the manufacturer. The RNA derived from CEMx174 cells transduced with HIV-based SIN vector constructs harbouring the *egfp* gene under the control of the SIVmac239-STR, the EF1α promoter and the EF1α/STR, respectively. The sequence of the gene-specific primer mapping in the *egfp* gene were EGR1: 5′-GTCCAGCTCGACCAGGATGGG-3′ and EGR2: 5′-ACAGCTCCTCGCCCTTGCTCAC-3′. PCR fragments were purified from agarose gels (QiaQuick Gel Purification Kit, Qiagen, Chatsworth, CA) and subcloned into the pCR2.1-TOPO vector (Invitrogen Corp., San Diego, CA). After transformation of *Escherichia coli* TOP10F′, colonies were picked and prepared for DNA sequencing performed by Genome Express (Montreuil, France).

Fluorescence microscopy and FACS analysis

At the indicated time points after infection of CEMx174-LTR-GFP cells or transduction of various other cell lines with viral vector particles, culture medium was removed, cells were washed with 1× PBS and then fixed with 2% (w/v) paraformaldehyde/PBS for ten minutes. Adherent cells were grown directly on glass slides, while non-adherent cells were dropped on slides after fixation. GFP expression was examined using a Leica DMRXA fluorescence microscope and images acquisition was done with an Hamamatsu C4742-95 cooled charge-coupled device (CCD) camera controlled by the Openlab® software (version 2.2.4, Improvision, Coventry, UK). Images were captured with fixed exposure times and processed identically with Adobe Photoshop as TIFF files.

For fluorescence-activated cell sorting (FACS) cells were harvested, washed, and fixed in 2% (w/v) paraformaldehyde/PBS. GFP expression was analysed on a FACSCALIBUR cytometer using CellQuest software (both from Becton Dickinson, Pont de Claix, France).

HIV-derived vectors

Virus particles containing the HIV-based SIN vectors were produced by transient transfection of 293T cells employing calcium phosphate precipitation with the vector plasmid, an encapsidation plasmid (pCMV-ΔR8.2) and a vesicular stomatitis virus-G protein (VSV-G) envelope expression plasmid (pHCMV-G) as described.[57] Culture supernatants were collected at 48 h after transfection and cleared of cellular debris by low-speed centrifugation. Supernatants were treated with DNase I to remove residual plasmid DNA and filtered (0.45 μm). The amount of VSV-G-pseudotyped particles was evaluated by measuring HIV-1 Gag p24 with an antigen capture enzyme-linked immunosorbent assay (NEN, Life Science). Aliquots of vector stocks were stored at −80 °C until needed.

Calculation of looping-induced local concentration

The looping-induced local concentration can be computed directly only for homogeneous polymers with well-defined Kuhn and contour lengths.[11] For a polymer consisting of two parts with different properties the first two moments of the end-to-end distance distribution can be calculated[58] and compared to those of a homogeneous chain, resulting in an effective Kuhn and contour length, i.e. those of a homogeneous polymer with the same looping behaviour. Iteratively adding monomer units like nucleotides or base-pairs with known properties (length per unit, Kuhn length) and applying this algorithm allows us to determine the looping behaviour and thus the local concentration of a heterogeneous polymer. In the case of the SIVmac239 and EF1α/STR mutant transcripts the sequence is known and the secondary structure state (ss, ds, or loop integration) of each unit is taken as predicted by MFOLD, leading to the local concentration of TAR-bound Tat near the transcription complex as a function of transcription progress.

The local concentration for molecules bound at distant sites of a polymer separated by a number of monomers *nt* is calculated as follows:[11]

$$j_{\mathrm{M}}(\mathrm{nt}) = 0.53\, n(\mathrm{nt})^{-3/2} \exp\left(-\frac{2 - d(L_{\mathrm{K}}(\mathrm{nt}))}{n(\mathrm{nt})^2 + d(L_{\mathrm{K}}(\mathrm{nt}))}\right) \times L_{\mathrm{K}}(\mathrm{nt})^{-3}\ \mathrm{nm}^3\ \mathrm{mol/l} \quad (1)$$

with n representing the number of polymer segments, d the parameter of facilitated interaction accounting for intrinsic looping and/or interaction radii of involved molecules, and L_{K} the Kuhn length. The number of segments is given by the contour length x(nt) and the Kuhn length L_{K}(nt) according to[59]:

$$n(\mathrm{nt}) = \frac{x(\mathrm{nt})}{L_{\mathrm{K}}(\mathrm{nt})} \quad (2)$$

Equation (1) is based on the end-to-end distance distribution, that of a rigid worm-like chain for short and that of a flexible Gaussian chain for long homogeneous polymers. For a heterogeneous polymer, this end-to-end distance distribution is generally not known. However, at least the first two moments of this distribution are accessible. For looping of a chain, one end of which is kept fixed, it is useful to define as the first moment the average projection of the end-to-end vector $\vec{R}$ on the first segment $\vec{u}(0)$ (in this case the first segment means the lastly added during transcription because this is still attached to the elongation complex while the rest of the molecule can move within its physical constraints):

$$\langle \vec{u}(0)\vec{R}\rangle = \int_0^x \mathrm{d}s \langle \vec{u}(0)\vec{u}(s)\rangle \quad (3)$$

where $\vec{u}(s)$ is the unit vector tangent to the chain at position s along the contour. The second moment can be defined as the mean squared end-to-end distance:

$$\langle \vec{R}^2\rangle = \int_0^x \mathrm{d}s \int_0^x \mathrm{d}s' \langle \vec{u}(s)\vec{u}(s')\rangle \quad (4)$$

For a homogeneous chain this gives:[58]

$$\langle \vec{u}(0)\vec{R}\rangle = \frac{L_{\mathrm{K}}}{2}(1 - \mathrm{e}^{-2x/L_{\mathrm{K}}}) \quad (5)$$

and

$$\langle \vec{R}^2\rangle = L_{\mathrm{K}} x\left(1 - \frac{L_{\mathrm{K}}}{2x}(1 - \mathrm{e}^{-2x/L_{\mathrm{K}}})\right) \quad (6)$$

i.e. they are determined by the Kuhn and the contour length. These moments can also be calculated analytically for a polymer consisting of a first part with Kuhn length $L_{\mathrm{K},1}$ and contour length x_1 and a second part with Kuhn length $L_{\mathrm{K},2}$ and contour length x_2:

$$\langle \vec{u}(0)\vec{R}\rangle = \frac{L_{\mathrm{K},1}}{2}(1 - \mathrm{e}^{-2x_1/L_{\mathrm{K},1}}) + \mathrm{e}^{-2x_1/L_{\mathrm{K},1}} \frac{L_{\mathrm{K},2}}{2} \times (1 - \mathrm{e}^{-2x_2/L_{\mathrm{K},1}}) \quad (7)$$

and

$$\langle \vec{R}^2\rangle = L_{\mathrm{K},1} x_1 \left(1 - \frac{L_{\mathrm{K},1}}{2x_1}(1 - \mathrm{e}^{-2x_1/L_{\mathrm{K},1}})\right) + L_{\mathrm{K},2} x_2 \left(1 - \frac{L_{\mathrm{K},2}}{2x_2}(1 - \mathrm{e}^{-2x_2/L_{\mathrm{K},2}})\right) + \frac{L_{\mathrm{K},1} L_{\mathrm{K},2}}{2} \times (1 - \mathrm{e}^{-2x_1/L_{\mathrm{K},1}})(1 - \mathrm{e}^{-2x_2/L_{\mathrm{K},2}}) \quad (8)$$

By comparing the moments of a two-component chain with equations (5) and (6) one can compute the Kuhn length L_{K}(nt) and the contour length x(nt) of a homogeneous polymer that would have the same moments of the end-to-end distance distribution and therefore the same or a very similar looping behaviour.

This allows us to define an algorithm for the calculation of the looping behaviour of a heterogeneous chain: starting with a homogeneous segment, one adds the next monomer and determines the resulting effective

Kuhn and contour length, i.e. the two-segment chain is treated as homogeneous. This is repeated for all following monomers, resulting in the effective Kuhn and contour length and the looping behaviour of the whole chain.

Note that the result is not necessarily symmetric, i.e. depends on the direction since it is assumed that the first segment is kept fixed. Adding flexible monomers to a rigid starting segment is different from adding rigid monomers to a flexible starting segment.

The values for length per nucleotide (nt) Δx and Kuhn length L_K are 0.65 nm per nt and 4 nm for ssRNA, 0.65 nm per nt and 8 nm for two parallel single strands, and 0.27 nm per nt and 75 nm for dsRNA (values from Rippe[11]). The Kuhn length for two parallel single strands equals two Kuhn lengths of ssRNA for short chains, assuming that the bending energy is simply additive. The looping-relevant part of the EF1α/STR mutant TAR structure is assumed to consist of all nucleotides before nt 11 and after 114, respectively, i.e. including the lower stem of the wild-type TAR (nt 1–11 and 114–124 in SIVmac239, respectively; see Figure 6). Then, starting with the base-pair (bp) 124 : 11, nucleotides are added iteratively, *nt* indexing the currently added nucleotide.

The facilitated reaction parameter was determined numerically[11] and found to be 0.130 for dsDNA (Kuhn length 100 nm, reaction radius 10 nm). Thus, we assume 0.130 for dsRNA and approximately zero for ssRNA and interpolate it linearly depending on the Kuhn length:

$$d(L_K(\mathrm{nt})) = \frac{0.13}{75\ \mathrm{nm}} L_K(\mathrm{nt}) \qquad (9)$$

Using this, the polymer properties and local concentrations are calculated by adding nucleotide by nucleotide and allowing the predicted secondary structure to form, starting with nt 125 and taking nt 114–124 into account as dsRNA. The resulting local concentrations of Tat close to the polymerase are plotted for the SIVmac239 sequence (which is single-stranded for all nt after position 124), for the EF1α/STR mutant sequence (which is expected to be single-stranded during synthesis of nt 125–129 and then to switch to the above-mentioned secondary structure) and for the EF1α/STR mutant sequence assuming that it would take immediately the secondary structure. For a fully double-stranded structure the model is not applicable, however, the local concentration is expected to be well below 10 nM.

Acknowledgements

We thank Annette Boese and Pierre Charneau for the kind gift of reagents and all members of the Wain-Hobson laboratory as well as Karsten Rippe for fruitful discussions. Several reagents were obtained through the NIH AIDS Research and Reference Reagent Program. This work was supported by the Agence Nationale de la Recherche sur le SIDA (A.N.R.S.) and the Pasteur Institute. P.S. acknowledges the receipt of a postdoctoral fellowship from the A.N.R.S. as well as a Marie Curie Fellowship of the European Community programme IHP (contract number HPMF-CT-1999-00267). M.W. was supported by the Volkswagenstiftung and a Boehringer Ingelheim Fonds Travel Allowence.

References

1. Tang, H., Kuhen, K. L. & Wong-Staal, F. (1999). Lentivirus replication and regulation. *Annu. Rev. Genet.* **33**, 133–170.
2. Bengal, E. & Aloni, Y. (1991). Transcriptional elongation by purified RNA polymerase II is blocked at the *trans*-activation-responsive region of human immunodeficiency virus type 1 *in vitro*. *J. Virol.* **65**, 4910–4918.
3. Selby, M. J., Bain, E. S., Luciw, P. A. & Peterlin, B. M. (1989). Structure, sequence, and position of the stem-loop in tar determine transcriptional elongation by tat through the HIV-1 long terminal repeat. *Genes Dev.* **3**, 547–558.
4. Toohey, M. G. & Jones, K. A. (1989). *In vitro* formation of short RNA polymerase II transcripts that terminate within the HIV-1 and HIV-2 promoter-proximal downstream regions. *Genes Dev.* **3**, 265–282.
5. Brigati, C., Giacca, M., Noonan, D. M. & Albini, A. (2003). HIV Tat, its TARgets and the control of viral gene expression. *FEMS Microbiol. Letters*, **220**, 57–65.
6. Greenblatt, J., Nodwell, J. R. & Mason, S. W. (1993). Transcriptional antitermination. *Nature*, **364**, 401–406.
7. Jones, K. A. (1997). Taking a new TAK on tat transactivation. *Genes Dev.* **11**, 2593–2599.
8. Kao, S. Y., Calman, A. F., Luciw, P. A. & Peterlin, B. M. (1987). Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. *Nature*, **330**, 489–493.
9. Van Gilst, M. R. & von Hippel, P. H. (2000). Quantitative dissection of transcriptional control system: N-dependent antitermination complex of phage lambda as regulatory paradigm. *Methods Enzymol.* **323**, 1–31.
10. Van Gilst, M. R., Rees, W. A., Das, A. & von Hippel, P. H. (1997). Complexes of N antitermination protein of phage lambda with specific and nonspecific RNA target sites on the nascent transcript. *Biochemistry*, **36**, 1514–1524.
11. Rippe, K. (2001). Making contacts on a nucleic acid polymer. *Trends Biochem. Sci.* **26**, 733–740.
12. Richter, S., Ping, Y. H. & Rana, T. M. (2002). TAR RNA loop: a scaffold for the assembly of a regulatory switch in HIV replication. *Proc. Natl Acad. Sci. USA*, **99**, 7928–7933.
13. Garriga, J. & Grana, X. (2004). Cellular control of gene expression by T-type cyclin/CDK9 complexes. *Gene*, **337**, 15–23.
14. Karn, J. (1999). Tackling Tat. *J. Mol. Biol.* **293**, 235–254.
15. Lind, K. E., Du, Z., Fujinaga, K., Peterlin, B. M. & James, T. L. (2002). Structure-based computational database screening, *in vitro* assay, and NMR assessment of compounds that target TAR RNA. *Chem. Biol.* **9**, 185–193.
16. Berkhout, B., Gatignol, A., Rabson, A. B. & Jeang, K. T. (1990). TAR-independent activation of the HIV-1 LTR: evidence that tat requires specific regions of the promoter. *Cell*, **62**, 757–767.
17. Churcher, M. J., Lowe, A. D., Gait, M. J. & Karn, J. (1995). The RNA element encoded by the *trans*-activation-responsive region of human immunodeficiency virus type 1 is functional when displaced downstream of the start of transcription. *Proc. Natl Acad. Sci. USA*, **92**, 2408–2412.
18. Rittner, K., Churcher, M. J., Gait, M. J. & Karn, J. (1995). The human immunodeficiency virus long

terminal repeat includes a specialised initiator element which is required for Tat-responsive transcription. *J. Mol. Biol.* **248**, 562–580.
19. Wright, S. & Luccarini, C. (1996). Effect of the position of TAR on transcriptional activation by HIV-1 Tat *in vivo*. *J. Mol. Biol.* **263**, 1–7.
20. Pereira, L. A., Bentley, K., Peeters, A., Churchill, M. J. & Deacon, N. J. (2000). A compilation of cellular transcription factor interactions with the HIV-1 LTR promoter. *Nucl. Acids Res.* **28**, 663–668.
21. Ilyinskii, P., Daniel, M., Simon, M., Lackner, A. & Desrosiers, R. (1994). The role of upstream U3 sequences in the pathogenesis of simian immunodeficiency virus-induced AIDS in rhesus monkeys. *J. Virol.* **68**, 5933–5944.
22. Kestler, H. r., Ringler, D., Mori, K., Panicali, D., Sehgal, P., Daniel, M. R. & Desrosiers (1991). Importance of the nef gene for maintenance of high virus loads and for development of AIDS. *Cell*, **65**, 651–662.
23. Kirchhoff, F., Kestler, H. r. & Desrosiers, R. (1994). Upstream U3 sequences in simian immunodeficiency virus are selectively deleted *in vivo* in the absence of an intact nef gene. *J. Virol.* **68**, 2031–2037.
24. Kirchhoff, F., Greenough, T., Brettler, D., Sullivan, J. & Desrosiers, R. (1995). Brief report: absence of intact nef sequences in a long-term survivor with nonprogressive HIV-1 infection. *N. Engl. J. Med.* **332**, 228–232.
25. Munch, J., Adam, N., Finze, N., Stolte, N., Stahl-Hennig, C., Fuchs, D. *et al*. (2001). Simian immunodeficiency virus in which nef and U3 sequences do not overlap replicates efficiently *in vitro* and *in vivo* in rhesus macaques. *J. Virol.* **75**, 8137–8146.
26. Ilyinskii, P., Simon, M., Czajak, S., Lackner, A. & Desrosiers, R. (1997). Induction of AIDS by simian immunodeficiency virus lacking NF-kappaB and SP1 binding elements. *J. Virol.* **71**, 1880–1887.
27. Chen, B., Feinberg, M. & Baltimore, D. (1997). The kappaB sites in the human immunodeficiency virus type 1 long terminal repeat enhance virus replication yet are not absolutely required for viral growth. *J. Virol.* **71**, 5495–5504.
28. Zhang, L., Huang, Y., Yuan, H., Chen, B., Ip, J. & Ho, D. (1997). Identification of a replication-competent pathogenic human immunodeficiency virus type 1 with a duplication in the TCF-1alpha region but lacking NF-kappaB binding sites. *J. Virol.* **71**, 1651–1656.
29. West, M. J. & Karn, J. (1999). Stimulation of Tat-associated kinase-independent transcriptional elongation from the human immunodeficiency virus type-1 long terminal repeat by a cellular enhancer. *EMBO J.* **18**, 1378–1386.
30. Marzio, G., Verhoef, K., Vink, M. & Berkhout, B. (2001). *In vitro* evolution of a highly replicating, doxycycline-dependent HIV for applications in vaccine studies. *Proc. Natl Acad. Sci. USA*, **98**, 6342–6347.
31. Chang, L. J., McNulty, E. & Martin, M. (1993). Human immunodeficiency viruses containing heterologous enhancer/promoters are replication competent and exhibit different lymphocyte tropisms. *J. Virol.* **67**, 743–752.
32. Guan, Y., Whitney, J., Detorio, M. & Wainberg, M. (2001). Construction and *in vitro* properties of a series of attenuated simian immunodeficiency viruses with all accessory genes deleted. *J. Virol.* **75**, 4056–4067.
33. Blancou, P., Chenciner, N., Fang, R. H., Monceaux, V., Cumont, M. C., Guetard, D. *et al*. (2004). Simian immunodeficiency virus promoter exchange results in a highly attenuated strain that protects against uncloned challenge virus. *J. Virol.* **78**, 1080–1092.
34. Wakabayashi-Ito, N. & Nagata, S. (1994). Characterization of the regulatory elements in the promoter of the human elongation factor-1 alpha gene. *J. Biol. Chem.* **269**, 29831–29837.
35. Dardalhon, V., Herpers, B., Noraz, N., Pflumio, F., Guetard, D., Leveau, C. *et al*. (2001). Lentivirus-mediated gene transfer in primary T cells is enhanced by a central DNA flap. *Gene Ther.* **8**, 190–198.
36. Reddy, B. & Yin, J. (1999). Quantitative intracellular kinetics of HIV type 1. *AIDS Res. Hum. Retroviruses*, **15**, 273–283.
37. Kent, S. J., Dale, C. J., Preiss, S. & Purcell, D. F. (2002). Evidence of recombination between 3′ and 5′ LTRs in macaques inoculated with SIV DNA. *AIDS Res. Hum. Retroviruses*, **18**, 227–230.
38. Berkhout, B. & Jeang, K. T. (1991). Detailed mutational analysis of TAR RNA: critical spacing between the bulge and loop recognition domains. *Nucl. Acids Res.* **19**, 6169–6176.
39. Zuker, M., Mathews, D. H. & Turner, D. H. (1999). Algorithms and Thermodynamics for RNA Secondary Structure Prediction: A Practical Guide in RNA Biochemistry and Biotechnology. In *NATO ASI Series* (Barciszewski, J. & Clark, B. F. C., eds), Kluwer Academic, Dodretch.
40. Pohlmann, S., Floss, S., Ilyinskii, P., Stamminger, T. & Kirchhoff, F. (1998). Sequences just upstream of the simian immunodeficiency virus core enhancer allow efficient replication. *J. Virol.* **72**, 5589–5598.
41. Garriga, J., Peng, J., Parreno, M., Price, D. H., Henderson, E. E. & Grana, X. (1998). Upregulation of cyclin T1/CDK9 complexes during T cell activation. *Oncogene*, **17**, 3093–3102.
42. Rabbi, M. F., Al-Harthi, L. & Roebuck, K. A. (1997). TNFalpha cooperates with the protein kinase A pathway to synergistically increase HIV-1 LTR transcription *via* downstream TRE-like cAMP response elements. *Virology*, **237**, 422–429.
43. Rabbi, M. F., Saifuddin, M., Gu, D. S., Kagnoff, M. F. & Roebuck, K. A. (1997). U5 region of the human immunodeficiency virus type 1 long terminal repeat contains TRE-like cAMP-responsive elements that bind both AP-1 and CREB/ATF proteins. *Virology*, **233**, 235–245.
44. Shibui-Nihei, A., Ohmori, Y., Yoshida, K., Imai, J., Oosuga, I., Iidaka, M. *et al*. (2003). The 5′ terminal oligopyrimidine tract of human elongation factor 1A-1 gene functions as a transcriptional initiator and produces a variable number of Us at the transcriptional level. *Gene*, **311**, 137–145.
46. Tolic-Norrelykke, S. F., Engh, A. M., Landick, R. & Gelles, J. (2004). Diversity in the rates of transcript elongation by single RNA polymerase molecules. *J. Biol. Chem.* **279**, 3292–3299.
47. Isel, C. & Karn, J. (1999). Direct evidence that HIV-1 Tat stimulates RNA polymerase II carboxyl-terminal domain hyperphosphorylation during transcriptional elongation. *J. Mol. Biol.* **290**, 929–941.
48. Marshall, N. F., Peng, J., Xie, Z. & Price, D. H. (1996). Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. *J. Biol. Chem.* **271**, 27176–27183.
49. Marshall, N. F. & Price, D. H. (1995). Purification of P-TEFb, a transcription factor required for the transition into productive elongation. *J. Biol. Chem.* **270**, 12335–12338.

50. Smith, C. A., Calabro, V. & Frankel, A. D. (2000). An RNA-binding chameleon. *Mol. Cell*, **6**, 1067–1076.
51. Xie, B., Wainberg, M. A. & Frankel, A. D. (2003). Replication of human immunodeficiency viruses engineered with heterologous Tat-transactivation response element interactions. *J. Virol.* **77**, 1984–1991.
52. Chen, H., Wilcox, G., Kertayadnya, G. & Wood, C. (1999). Characterization of the Jembrana disease virus tat gene and the *cis*- and *trans*-regulatory elements in its long terminal repeats. *J. Virol.* **73**, 658–666.
53. Roy, S., Parkin, N. T., Rosen, C., Itovitch, J. & Sonenberg, N. (1990). Structural requirements for trans activation of human immunodeficiency virus type 1 long terminal repeat-directed gene expression by tat: importance of base pairing, loop sequence, and bulges in the tat-responsive sequence. *J. Virol.* **64**, 1402–1406.
54. Gundlach, B. R., Linhart, H., Dittmer, U., Sopper, S., Reiprich, S., Fuchs, D. *et al*. (1997). Construction, replication, and immunogenic properties of a simian immunodeficiency virus expressing interleukin-2. *J. Virol.* **71**, 2225–2232.
55. Regier, D. & Desrosiers, R. (1990). The complete nucleotide sequence of a pathogenic molecular clone of simian immunodeficiency virus. *AIDS Res. Hum. Retroviruses*, **6**, 1221–1231.
56. O'Doherty, U., Swiggard, W. J. & Malim, M. H. (2000). Human immunodeficiency virus type 1 spinoculation enhances infection through virus binding. *J. Virol.* **74**, 10074–10080.
57. Naldini, L., Blomer, U., Gallay, P., Ory, D., Mulligan, R., Gage, F. H. *et al*. (1996). *In vivo* gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science*, **272**, 263–267.
58. Rivetti, C., Walker, C. & Bustamante, C. (1998). Polymer chain statistics and conformational analysis of DNA molecules with bends or sections of different flexibility. *J. Mol. Biol.* **280**, 41–59.
59. Grosberg, A. & Khoklov, A. (1994). *Statistical Physics of Macromolecules*. AIP Press, New York.

Edited by J. Karn

(Received 23 June 2004; received in revised form 19 September 2004; accepted 20 September 2004)