The EMBO Journal vol.9 no.3 pp.965-970, 1990

A tripartite HIV-1 *tat-env-rev* fusion protein

Jochen Salfeld, Heinrich G.Göttlinger, Rey A.Sia, Ronald E.Park, Joseph G.Sodroski and William A.Haseltine

Division of Human Retrovirology, Dana-Farber Cancer Institute, Department of Pathology, Harvard Medical School, 44 Binney Street, Boston, MA 02115, USA

Communicated by H.Schaller

A 26 kd protein reactive with antiserum to the transactivator *tat* of the Human Immunodeficiency Virus Type 1 (HIV-1) has been detected in virus producing cells. The 26 kd protein is shown to be a tripartite fusion protein including coding sequences of the *tat*, envelope (*env*) and regulator of virion expression (*rev*) genes. Fusion of these coding sequences occurs by use of a previously undescribed exon within *env*. This 26 kd protein, designated *tnv*, has *tat* but no *rev* activity detectable with the assay used. The existence of other less abundant *tat* and *rev* related proteins in HIV-1 producing cells is also noted.

Key words: differential splicing/HIV-1/regulatory proteins/transactivation

Introduction

The genetic structure and replication cycle of the Human Immunodeficiency Virus Type 1 (HIV-1) is more complex than that of simple animal retroviruses. In addition to the capsid (*gag*) and replication proteins (*pol*) produced by translation of the full length genomic RNA, and the outer envelope product, produced by a singly spliced messenger RNA, HIV-1 specifies at least six additional proteins that are produced by multiply spliced messenger RNAs (see Figure 4A). These additional proteins regulate virus replication. Two of these supernumerary proteins, the transactivator (*tat*) protein (Arya *et al.*, 1985; Sodroski *et al.*, 1985) and the regulator of virion protein expression, the *rev* protein (Feinberg *et al.*, 1986; Sodroski *et al.*, 1986), are required for viral growth (Dayton *et al.*, 1986; Fisher *et al.*, 1986; Terwilliger *et al.*, 1988). The transactivator protein *tat* accelerates the production of all viral proteins (Rosen *et al.*, 1985) and the *rev* protein regulates the accumulation of full-length or partially spliced viral messenger RNAs (Dayton *et al.*, 1988; Malim *et al.*, 1988; Rosen *et al.*, 1988; Emerman *et al.*, 1989). An additional protein of 26 kd has been observed using monospecific anti-*tat* antisera in HIV-1 producing cells, raising the possibility that alternative forms of *tat* exist (Feinberg *et al.*, 1986). Reported here are studies that describe the structure and function of this additional 26 kd HIV-1 *tat* related protein. These studies also reveal a novel HIV-1 *env* derived exon.

Results

Detection of the 26 kd tat product

To analyze the *tat* related proteins specified by a provirus of known sequence, COS-7 cells were transfected with an infectious provirus HXBc2, in the context of a plasmid which contains an SV40 origin of replication. Under the conditions used, the plasmid replicates to high copy number yielding abundant template proviral DNA. A monospecific anti-*tat* antiserum, raised to a synthetic peptide corresponding to amino acids 1-20 of the *tat* protein, was used to precipitate the *tat* related products. The cells were labeled with [^{35}S]cysteine as the *tat* protein contains numerous cysteine residues. The data of Figure 1 (lane 3) show that under these conditions three major proteins (26, 16 and 15 kd) and two minor proteins (24 and 21 kd) are recognized by the anti-*tat* antiserum, which reacts with the amino terminus of *tat*. These proteins are not present in mock-transfected COS-7 cells (Figure 1, lane 4) nor in the transfected cells reacted with preimmune serum (Figure 1, lane 5). Under the gel conditions used the HIV-1 *tat* protein usually described as a 14 kd protein migrates with an apparent mol. wt of 16 kd. We have used this value for the description of the 87 amino acid form of *tat* throughout the paper. The 26 and 16 kd products are the most abundant proteins and are made in similar amounts.

The 26 kd protein reactive with the *tat* serum is also observed if a portion of the envelope region (nucleotide position 6619-7199) is deleted in the proviral DNA (pHXB-SV△Bg1II) (Figure 2, lane 7 and 8).

The *tat* and *rev* proteins are specified by two exons which must be joined by splicing. The reading frame of *vpu* and more than two-thirds of the reading frame for *env* is interposed between the first and second coding exons of both *tat* and *rev* (see Figure 4A). To determine whether the 26 kd protein contained *rev* specified sequences, a synthetic peptide antiserum was used that had been raised against amino acids 38-51 of *rev*, which are specified by the second *rev* coding exon. Figure 1 (lane 1) shows that this antiserum precipitates the 26 kd protein that co-migrates with that recognized by the anti-*tat* antiserum. The anti-*rev* antiserum also recognizes the 20 kd *rev* protein produced in the transfected cells. In addition, the anti-*rev* antiserum also recognizes two minor products, the 24 and 21 kd proteins that are also precipitated by the *tat* antiserum. The specificity of the anti-*rev* serum was confirmed by recognition of the 20 kd *rev* product but not 16 or 15 kd *tat* products.

Identification of a mRNA which specifies the 26 kd tat product

To search for mRNAs which might specify the 26 kd *tat-rev* related protein, a cDNA library was prepared from RNA extracted from COS-7 cells transfected with HXBc2 provirus. cDNA prepared from poly(A)$^+$ selected RNA

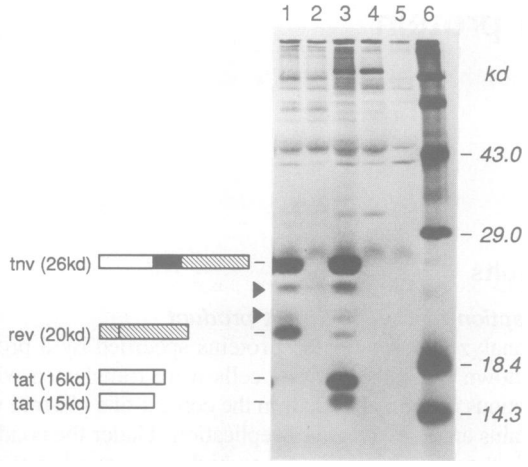

Fig. 1. *Tat* and *rev* related proteins in COS-7 cells transfected with the HXBc2 provirus. COS-7 cells were transfected with the plasmid pHXB-SV (**lanes 1, 3** and **5**) or the same amount of salmon sperm DNA (**lanes 2** and **4**) as described in Materials and methods, labeled with [^{35}S]cysteine and lysed as described (Göttlinger *et al.*, 1989). The lysates obtained were precipitated with the anti-*rev* serum (lanes 1 and 2), the anti-*tat* serum (lanes 3 and 4) or the preimmune serum (lane 5). The structure and the mol. wt of the major products is indicated on the left, the arrowheads denote minor proteins of 24 and 21 kd.

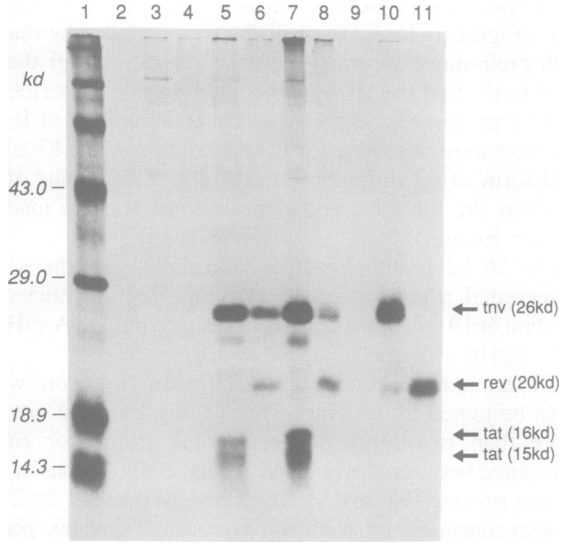

Fig. 2. Comparison of the *tat* and *rev* related proteins detected in COS-7 cells transfected with the plasmid pHXB-SV (**lanes 2, 5** and **6**), HXB-SVΔBg1II (**lanes 7** and **8**) or salmon sperm DNA (**lanes 3** and **4**) with the *rev* related protein derived from *in vitro* translation of RNA synthesized using pSP26 (**lanes 9** and **10**) and pSP*rev* (**lane 11**). The labeled proteins were subjected to immunoprecipitation using the preimmune serum (lanes 2 and 9), anti-*tat* serum (lanes 3, 5 and 7) or the anti-*rev* serum (lanes 4, 6, 8, 10 and 11). The mol. wt of the major protein bands is indicated.

was inserted into the λgt10 vector. Phage plaques were first screened for reactivity with a radioactive RNA probe which included both coding exons of *tat* and *rev* as well as the *env* gene. Plaques were also screened with a probe corresponding to *gag*. Phage in plaques that were recognized by the *tat*−*env*−*rev* probe but not by the *gag* probe were selected and purified for further study. Purification of phage which did not react with the *gag* probe helped to select against full-length cDNA clones. One cDNA clone that reacted with

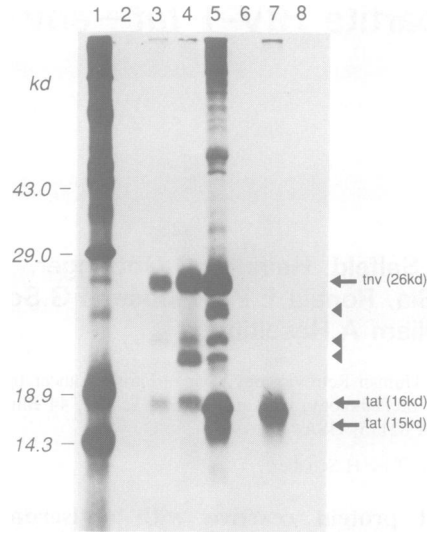

Fig. 3. Comparison of the 26 kd *tat* and 26 kd *tnv* protein found in transfected COS-7 cells with the *tat* related proteins synthesized *in vitro*. The [^{35}S]cysteine labeled proteins derived from *in vitro* translation of the *tnv in vitro* RNA (pSP26) (**lanes 2−4**) and the *tat in vitro* RNA (pSPtat) (**lanes 6−8**) were subjected to immunoprecipitation with the preimmume serum (lanes 2 and 6), the anti-*tat* serum (lanes 3 and 7) and the anti-*rev* serum (lanes 4 and 8). Those proteins precipitated were compared to the ones detected by the anti-*tat* serum in COS-7 cells transfected with pHXB-SV (lane 5). The mol. wt of the major proteins is indicated, arrowheads denote minor proteins detected.

the long terminal repeat (LTR)/*nef* probe only (AC1) was also characterized. The complete nucleotide sequence of three cDNAs designated AC1, 2 and 3 was determined. The regions of HIV-1 contained in these three cDNAs are illustrated schematically in Figure 4A.

The AC1 cDNA contains the HIV-1 5′ leader sequences, from nucleotide 1 to 16, fused by splicing from the donor site at nucleotide 287 to the acceptor site at nucleotide 7956. The AC1 cDNA is most likely to correspond to the mRNA for *nef* as the *nef* AUG codon is the first initiation codon of the RNA. A similar *nef* mRNA was recently described in Simian Immunodeficiency Virus (SIV) infected cells (Columbini *et al.*, 1989). The sequence of the AC2 and AC3 cDNAs reveals the existence of a novel exon within *env*. In addition to the HIV-1 leader sequences from 91 to 287, AC2 contains sequences from 4970 to 5043 derived from the *vif* gene, a non-coding exon that has been previously described (Muesing *et al.*, 1985); from 5557 to 5625, derived from the 3′ portion of the first coding exons of *tat* and *rev*; from 6185 to 6300, a novel *env* derived exon, and from 7956 to 9210, corresponding to the 3′ exon common to all HIV-1 mRNAs (Muesing *et al.*, 1985). AC3 is incomplete as it lacks the 5′ leader sequence. The 5′ end of AC3 begins within the first coding exon of *tat* at nucleotide 5450 and includes the remainder of the *tat* to nucleotide 5625. AC3 contains the same novel *env* exon and 3′ derived sequences as does AC2.

The sequence of AC3 predicts that the corresponding mRNA will yield a protein which contains the amino terminal *tat* exon fused in-frame to *env* sequences, which in turn are fused in-frame to the carboxy terminus of *rev* (see Figure 4A). To determine whether such a hybrid protein can be made, a plasmid was constructed in which the 5′ coding sequences of the first coding exon of *tat* were restored by

linking nucleotides 5367 (*Sal*I) to 5608 (*Hin*dIII) derived from the parental HXBc2 provirus to the AC2 cDNA. This entire DNA fragment was placed 3′ to an SP6 promoter, creating the pSP26 plasmid (see Figure 4B), so that mRNA derived from the sequence could be used to program an *in vitro* protein synthesis reaction. RNA derived from two additional plasmids, pSP*tat* and pSP*rev* which contain the first and second coding exons of *tat* and *rev*, respectively, joined in-frame were also used. The pSP*tat* plasmid contains the entire coding sequence of *tat*, whereas in pSP*rev* the initiation codon of *tat* was deleted (see Figure 4B). RNA derived from these plasmids was used to direct protein synthesis in a reticulocyte lysate which contained [^{35}S]cysteine.

In vitro translation of the pSP26 mRNA yields a 26 kd protein as the major product (Figure 3). The 26 kd protein made in the reticulocyte extract is detected by anti-*tat* (lane 3) as well as anti-*rev* serum (lane 4) and co-migrates with the 26 kd protein identified in COS-7 cells transfected with the entire provirus (lane 5). As expected, the 16 and 15 kd *tat* and two minor products recognized by the anti-*tat* sera in the transfected cells are not present in these extracts. Two minor proteins of ~21 and 16.5 kd, respectively, visible in Figure 2B (lanes 3 and 4) have not been characterized further. A 20 kd protein is recognized by the *rev* serum only (Figure 3, lane 4; Figure 2, lane 10). *In vitro* translation of the *tat* and *rev* mRNA yields only 16 (Figure 3, lane 7) and 20 kd proteins respectively (Figure 2, lane 11). These two proteins migrate with the authentic *tat* and *rev* products produced in COS-7 cells transfected with the HXBc2 provirus (Figures 2 and 3).

Activity of the 26 kd protein

To determine whether the 26 kd *tat–env–rev* fusion protein retains *tat* or *rev* activity, a cDNA fragment capable of producing the 26 kd protein was placed 3′ to the HIV LTR to create the plasmid pIIIex26. The *tat* activity of the *tnv* protein was examined by co-transfection of pIIIex26 with the indicator plasmids pHIV1CAT or pHIV2CAT. The pHIV1CAT contains the HIV-1 LTR sequences (nucleotides −650 to +80) placed 5′ to the gene for chloramphenicol acetyl transferase (CAT) while pHIV2CAT contains the full-length 5′ LTR of HIV-2 until nucleotide +860 fused to the CAT gene. For comparison, transfections with pIIIextat expressing the 16 kd *tat* protein were done in parallel. The data in Table I show that in the presence of pIIIex26 there is a dramatic increase in CAT activity directed by the HIV-1 and HIV-2 promoters but not by the metallothionein promoter. The stimulatory effect of *tnv* was found to be 45–60% of the activity of *tat*; Table I shows a representative experiment. Transactivation of the HIV-1 LTR as compared to the HIV-2 LTR by the 26 kd *tnv* product is similar to what has been reported for the 16 kd *tat* protein (Emerman *et al.*, 1987). The *rev* activity of the 26 kd protein was compared to that of the 20 kd *rev* protein (expressed from plasmid pIIIexrev) using as an indicator the plasmid pIIIAR which contains the CAT gene located 5′ to the 3′ portion of the HIV-1 genome (Rosen *et al.*, 1988). Expression of CAT activity directed by this plasmid has been shown to depend on *rev*. The plasmid pIIIARΔCAR was used as a control. This plasmid is identical to pIIIAR except for a 64 bp deletion in the *rev* responsive sequence (CAR) that renders this construct unresponsive to *rev*. The data of Table I show that the 26 kd *tnv* protein has no detectable *rev* activity as judged by this assay. It remains to be seen whether more sensitive *rev* assays will detect some residual *rev* activity.

Table I. Comparison of *tat* and *rev* activities of *tat*, *rev* and 26 kd *tnv*

Expressor plasmid	Assay for *tat* activity			Assay for *rev* activity	
	pHIV1CAT	pHIV2CAT	pMT3CAT	pIIIAR	pIIIARΔCAR
control	1	1	1	1	1
pIIIex tat	119	101	1	–	–
pIIIex rev	–	–	–	39	1
pIIIex 26	68	78	1	1	1

Values represent fold transactivation as compared to control. The assay for *tat* function done in Jurkat cells was done as described (Rosen *et al.*, 1985). The *rev* assay was done in HeLa TatIII cells as described (Rosen *et al.*, 1988). The plasmid pGem4 (Promega) was used as a control.

Discussion

The experiments reported here reveal a hitherto unknown exon in the HIV-1 *env* coding sequence. mRNAs can be made which contain an exon included between nucleotides 6185 and 6300. mRNAs have been isolated in which the splice donor at the 3′ end of the first *tat–rev* coding exon is joined to the 5′ end of the novel *env* exon. The 3′ splice donor of the novel *env* exon is spliced to the 5′ region of the second *tat–rev* coding exon. The sequences of the 3′ and 5′ end of the novel *env* exon fit well with the consensus sequences (Mount *et al.*, 1982). Figure 5 shows a comparison of the HIV-1 and HIV-2 splice donor sites including the potential base-pairing with the U1 snRNA involved in recognition of splice donor sites (Sharp, 1987).

The 26 kd *tat* related protein that has been previously reported to be present in HIV-1 producing cells (Feinberg *et al.*, 1986) is shown to be a product of a mRNA which contains the novel *env* exon. A protein is produced by the mRNA containing the novel exon and is recognized by antisera which recognize the amino terminus of *tat* and the carboxy terminus of *rev*. The size of the protein produced by HIV-1 producing cells is the same as that made upon *in vitro* translation of the mRNA which contains the *tat–env–rev* exons. Truncation of the 3′ sequences of such a message reduces the apparent size of all the proteins produced in an *in vitro* translation reaction (data not shown). It is proposed that the 26 kd *tat* related protein be designated *tnv* for *t*at–en*v*–re*v* fusion product.

The 26 kd *tnv* product has *tat* but not *rev* activity. This observation is consistent with other studies which show that the second coding exon of *tat* is not required for activity (Sodroski *et al.*, 1985; Seigel *et al.*, 1986). Mutations in the amino terminus of *rev* have been shown to inactivate the activity of this protein (Malim *et al.*, 1989). The co-migration of *in vivo* and *in vitro tnv* proteins indicates that the *tnv* protein is not modified post-transcriptionally. Absence of a hydrophobic signal sequence probably accounts for a failure to modify the protein by glycosylation despite the presence of potential glycosylation sites within the *env* derived sequences.

Noteworthy are other *tat* and *env* related proteins in the HIV-1 producing cells. The *tat* related protein which

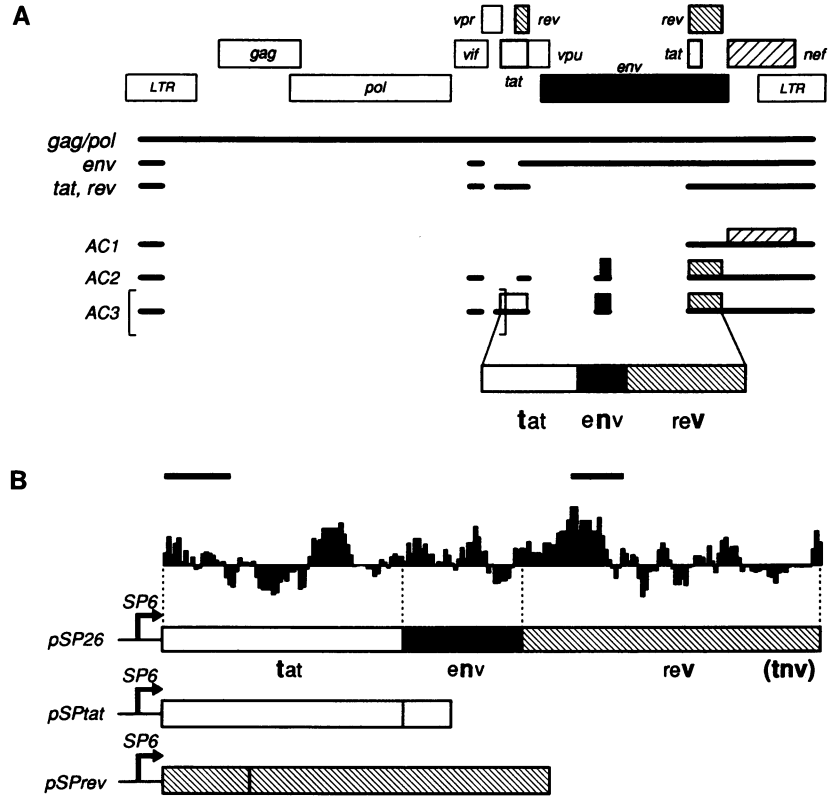

Fig. 4. Genome organization of HIV-1. (**A**) The known open reading frames flanked by the two LTRs are shown, the reading frames for *tat*, *rev*, *env* and *nef* are highlighted. The transcripts described by Muesing *et al.* (1985) are shown, the transcripts AC1, AC2 and AC3 are shown as characterized and described in this paper. The reading frames expressed from each individual mRNA are highlighted. The assumed 5′ end of the AC3 transcript is shown in brackets. The structure of the 26 kd *tnv* protein is indicated. (**B**) The hydrophilicity (up)/hydrophobicity (down) prediction for the *tnv* protein was computed according to Hopp and Woods (1981). The horizontal black bars indicate the position of the peptide sequences used for raising the antisera. The structure of the plasmids used for *in vitro* transcription is shown; the exon boundaries are indicated by vertical lines.

migrates slightly faster than the 16 kd *tat* product is likely to be the product of the first coding exon of *tat* previously described (Malim *et al.*, 1988). This protein is not recognized by the *rev* antiserum. Two minor proteins, 24 and 21 kd, are recognized by both the *tat* and *rev* antiserum (Figure 1, lanes 1 and 3). Neither of these two proteins are evident in the products of the *in vitro* translation of *tnv* RNA. It is possible that these two proteins represent alternative *tat*−*rev* fusion products. In this regard we note an alternative splice donor at the 3′ end of the first coding exon of *tat* in position 5594 which, if fused to the splice acceptor in position 7956, may result in the in-frame fusion of the first coding exon of *tat* with the second coding exon of *rev* without additional *env* sequences. Either the 24 or 21 kd protein may be the result of the usage of this alternative splice donor.

A protein migrating more rapidly than 26 kd at ~20 kd is recognized by the anti-*rev* serum in reticulocyte extracts programmed with the *tnv* RNA (Figure 3, lane 4; Figure 2, lane 10). The electrophoretic mobility of this product is decreased by removal of 3′ coding sequences from the *tnv* RNA (data not shown). This observation suggests that these proteins are the result of internal initiation at one of the two frame AUG codons in the envelope exon. Use of any of these internal initiation sites could result in alternative forms of *tat* and *rev* related proteins in HIV-1 producing cells. It is notable that the first methionine codons in the mRNA corresponding to its AC2 cDNA clone are located within these *env* derived sequences (Figure 4A). The other smaller

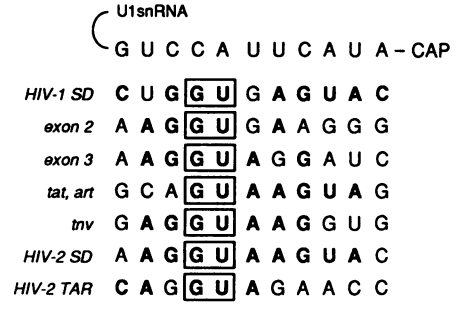

Fig. 5. Comparison of the splice donor sites of HIV-1 and HIV-2. The major splice donor is labeled SD, the other ones corresponding to the exons involved. The HIV-1 sequences shown are derived from Myers *et al.* (1989). The sequences for HIV-2 are derived from cDNA cloning and sequencing (J.Salfeld, unpublished results) and are in positions identical to the ones described for SIV (Colombini *et al.*, 1989). The nucleotides shown in bold type are potentially involved in base-pairing with the U1 snRNA (Sharp *et al.*, 1987).

proteins of 21 and 16.5 kd evident in this experiment (Figure 3, lanes 3 and 4) may be either products of premature termination or degradation of *tnv*.

These experiments reveal a surprising flexibility in utilization of the coding information contained within *tat* and *rev* exons. Multiple *tat*−*rev* products have also been reported in cells producing SIV. The multiple SIV *tat* and *rev* products result from use of alternative splice acceptor sequences within the second *tat*−*rev* exons (Viglianti and Mullins, 1988). However, no *tat*−*rev* or *tat*−*env*−*rev* fusion products have

yet been reported for SIV. The consensus splice acceptor sequence in the envelope region is conserved in many, but not all, HIV-1 isolates (Myers *et al.*, 1989). Absence of a consensus splice acceptor in some HIV-1 strains may reflect *in vitro* selection against the expression of the exon. In this regard a high frequency of defective or modulated *tat* sequences *in vivo* as well as in cell culture has been reported recently (Meyerhans *et al.*, 1989). Other HIV-1 genes including *vrp*, *vpu* and *nef* are counter-selected by *in vitro* growth (Cohen *et al.*, 1988). The existence of multiple forms of these highly important regulatory genes raises the possibility that the activity of these proteins is subtly modified during the course of infection and/or by the host cell splicing apparatus. There are similar examples in other viral systems, most prominently in Bovine Papilloma Virus (BPV), where regulatory proteins are made by fusing domains from one or two different reading frames via splicing of the corresponding mRNA (Lambert *et al.*, 1987, 1989; Stephens and Harlow, 1987). The implication of the existence of multiple forms of the *tat* and *rev* proteins for development of antiviral drugs that are designed to inhibit the *tat* and *rev* activity must be explored.

Materials and methods

Preparation of the λgt10 library

The poly(A)$^+$ RNA prepared from COS-7 cells transfected with HIV-1 HXBc2, carried the plasmid pHXB-SV (Dayton *et al.*, 1988) which contains an SV40 origin of replication, was used to prepare double-stranded cDNA essentially as described (Gubler and Hoffman, 1983). RNA was harvested from the transfected cells 48 h post-transfection. Subsequently the cDNA was linked to synthetic *Eco*RI adaptors and size fractionated on a 1.0% agarose gel. The cDNA between 1000 and 3000 bp was eluted from the gel, ethanol precipitated and linked to *Eco*RI digested λgt10 arms. Further steps in the preparation of the λgt10 cDNA library were done as described (Maniatis *et al.*, 1982). The library contained $>5 \times 10^5$ clones and was amplified to obtain phage titers of 1.6×10^{11} ml^{-1}.

Screening of the cDNA library

Sequence specific RNA probes were used for screening the cDNA library. To produce the RNA probes specific for different regions of the HIV genome, *Pst*I−*Bgl*II fragments (nucleotides 969−1641) were subcloned from HIV HXBc2 into pGem3 resulting in pGem3gag; the corresponding SP6 probe will hybridize downstream of the HIV-1 splice donor within the *gag* intron and can be used to select against cDNAs of unspliced RNAs. The *Sal*I−*Bam*HI fragment (nucleotides 5367−8053) was cloned into pGem3 resulting in pGem3env. Finally, the *Bam*HI−*Hin*dIII fragment (nucleotides 8053−9194) was cloned in the same vector resulting in pGem3LTR/nef.

The RNA was synthesized *in vitro* using the SP6 polymerase. The typical reaction is: 4 μl 5 × transcription buffer (200 mM Tris−HCl, pH 7.5, 30 mM $MgCl_2$, 10 mM spermidine, 50 mM NaCl), 2 μl 100 μM CTP, 1 μl linearized pGem plasmid, 5 μl 10 mCil/ml [α-^{32}P]CTP (3000 Cil/mmol), 1 μl SP6 polymerase. After 30 min incubation at 37°C an additional 0.5 μl of SP6 polymerase was added. After completion of the 60 min incubation the DNA template was digested with RNase free DNase (Promega) as described. The labeled RNA was purified using a Du Pont Nensorb^{TM} 20 nucleic acid purification cartridge as suggested by the manufacturer.

Replica filters of the phage plates were prepared as described (Maniatis *et al.*, 1982), baked and then prehybridized in 50% formamide, 5 × Denhardt's solution, 5 × SSP, 0.1% SDS, 100 μg/ml denatured salmon sperm DNA for 4 h. The dried RNA probe was then dissolved in 1 ml prehybridization solution, heated to 90°C, chilled on ice and added to the hybridization solution containing the filter. All hybridizations were done at 42°C. The washing procedures were done as described (Maniatis *et al.*, 1982). Multiple filters of each phage plate were screened in pairs of two with three different probes each. Single positive plaques or plaques hybridizing to the *gag* probe were ignored, the other phage plaques were picked, diluted and replated for a second round of screening.

Double positive plaques after the second screening were grown in liquid cultures, the DNA was prepared and *Eco*RI digested, the cDNA inserts were purified on a LMA gel and subcloned into M13mp18. White M13 plaques on a XL1 blue lawn were picked, grown up and single-strand DNA was prepared as described.

The inserts were sequenced using the Sequenase kit (US Biochemicals) as suggested by the manufacturer.

Construction of plasmids

In order to construct pSP26 the cDNA clone of λAC2 was digested with *Hin*dIII and *Eco*RI and the resulting fragment was purified and ligated to the *Sal*I−*Hin*dIII fragment derived from pHXB-SV (Dayton *et al.*, 1988) to restore the *tat* reading frame. The resulting *Sal*I−*Eco*RI fragment was subcloned into pGem3 (Promega Biotec, Madison, WI). Plasmid pSP*rev* was constructed by subcloning the *rev* cDNA as a *Sal*I−*Kpn*I fragment from plasmid pH3-art (Rosen *et al.*, 1988) into pGem3. This cDNA fragment does not contain the *tat* initiation codon because this has been removed by BAL31 digestion in creating pH3-art (Rosen *et al.*, 1988). The plasmid pSP*tat* was constructed by subcloning the *Sal*I−*Bam*HI fragment from pCV1 (Arya *et al.*, 1985) into pSP64 (Melton *et al.*, 1984). The plasmid pIIIex26 was created by subcloning the *Sal*I−*Eco*RI fragment of pSP26 into pIIIextat under deletion of the *tat* cDNA. The plasmid pIIIextat has been previously described as pΔ(83-5365/8033-9296) (Sodroski *et al.*, 1985). The plasmid pIIIexrev has been previously described as pH3-art (Rosen *et al.*, 1988). The pHIV2CAT plasmid was constructed by subcloning the *Pvu*I−*Nar*I fragment of pROD27 (Guyader *et al.*, 1987) to the *Cla*I−*Pvu*I fragment of pMT3CAT. The pHIV1CAT plasmid has been previously described as pU3R-III (Sodroski *et al.*, 1986). The structure of pMT3-CAT is essentially the same as pSV2CAT (Gorman *et al.*, 1982) with the SV40 promoter being replaced by the human metallothionein promoter (H.Schaller, personal communication). The plasmid containing the full-length provirus pHXB-SV has been described (Dayton *et al.*, 1988). The plasmid pHXB-SVΔBg1II is identical to pHXB-SV except for a deletion of a portion of the envelope gene from nucleotide position 6619 to 7199. The plasmid pIIIAR has been described recently (Rosen *et al.*, 1988). The plasmid pIIIARΔCAR used as a control is identical except for a deletion between nucleotide position 7351 and 7407 in the *rev* responsive element CAR. This deletion renders the plasmid non-responsive to *rev* (Dayton *et al.*, 1988). All DNA manipulations were done as described (Maniatis *et al.*, 1982).

In vitro translation and immunoprecipitations

The *in vitro* translation and the subsequent immunoprecipitations were done as previously described (Weimer *et al.*, 1987). The 15% SDS−PAGE was done as described (Laemmli, 1970). The gels were subsequently fixed for 30 min in 20% v/v methanol and 20% v/v acetic acid, incubated for 1 h in Enhance (Du Pont, NJ), rehydrated for 30 min in *aqua bidest.*, dried and exposed on X-Omat film (Kodak, Rochester, NY). The anti-*rev* serum has been described earlier (Goh *et al.*, 1987) and was prepared using a synthetic peptide corresponding to the sequence Arg_{38}-Arg-Asn-Arg-Arg-Arg-Arg-Trp-Arg-Glu-Arg-Gln-Arg-Gln_{51} derived from the second *rev* exon. The anti-*tat* serum was prepared against a synthetic peptide corresponding to the sequence Met_1-Glu-Pro-Val-Asp-Pro-Arg-Leu-Glu-Pro-Trp-Lys-His-Pro-Gly-Ser-Gln-Pro-Lys-Thr_{20} from the first exon of *tat*. The metabolic labeling of COS-7 cells and the preparation of the lysates for immunoprecipitation were done as previously described (Göttlinger *et al.*, 1989).

Cell lines and transfections

The Jurkat-tatIII cell line (Rosen *et al.*, 1985) was used for the functional *rev* studies, the Jurkat cell line was used for the transactivation studies as described. The transfections were done as described (Cullen *et al.*, 1987). COS-7 cells were purchased from ATCC and used for initial expression studies and the preparation of the mRNA for the cDNA library. Transfection and detection of HIV-1 in COS-7 cells were done as described (Dayton *et al.*, 1988).

Assays for tat and rev

The assay for transactivation was done as described (Rosen *et al.*, 1985) in Jurkat cells, the assay for *rev* has been described recently (Rosen *et al.*, 1988).

Nomenclature

The nomenclature in this paper follows the one agreed upon recently (Gallo *et al.*, 1988) and the numbering system is as in the complete sequence of HXBc2 (Ratner *et al.*, 1987). The +1 position is consequently the start site of transcription.

Acknowledgements

We thank Drs E.Beck and H.Schaller, ZMBH, Heidelberg, for laboratory space and support in preparing the cDNA library, Drs H.Schaller and

M.Junker for providing the plasmid pMTC3CAT and Dr E.A.Cohen for the plasmids pSPrev and pSPtat. We thank Andrew Hirsch for technical help in some of the experiments and, in particular, Amy Emmert for help with the artwork. Special thanks to Drs Y.Lu, M.Bowman and D.Blohm for constant interest and helpful discussions, and L.C.Wang-Salfeld as well as Drs Y.Lu and H.Schaller for critically reading the manuscript. This work was supported by NIH grants AI24845, AI28785, CA44460, AI24755 and AI27702. J.S. was supported by NIH grants AI28785 and CA44460. W.A.H. was supported by grants AI24845 and AI28785. J.G.S. was supported by the Leukemia Society of America and NIH grants AI24775 and AI27702. H.G.G. was supported by a Sheldon Andelson/AMFar scholar award (no. 700111-6-RF).

References

Arya,S., Guo,C., Josephs,S. and Wong-Staal,F. (1985) *Science*, **229**, 69–73.

Cohen,E., Terwilliger,E., Sodroski,J. and Haseltine,W.A. (1988) *Nature*, **334**, 532–534.

Colombini,S., Arya,S., Reitz,M., Jagodzinski,L., Beaver,B. and Wong-Staal,F. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 4813–4817.

Cullen,B.R. (1987) *Methods Enzymol.*, **152**, 648–704.

Dayton,A., Sodroski,J., Goh,W. and Haseltine,W.A. (1986) *Cell*, **44**, 941–947.

Dayton,A., Terwilliger,E., Potz,J., Kowalski,M., Sodroski,J. and Haseltine,W. (1988) *J. Acquired Immune Deficiency Syndromes*, **1**, 441–452.

Emerman,M., Guyader,M., Montagnier,L., Baltimore,D. and Muesing,M.A. (1987) *EMBO J.*, **6** 3755–3760.

Emerman,M., Vazeux,R. and Peden,K. (1989) *Cell*, **57**, 1155–1165.

Feinberg,M., Jarett,R., Aldovini,A. and Gallo,R. (1986) *Cell*, **46**, 807–817.

Fisher,A.G., Feinberg,M.B., Josephs,S.F., Harper,M.E., Marselle,L.M., Reyes,M., Gonda,M.A., Aldovini,A., Debouk,C., Gallo,R.C. and Wong-Staal,F. (1986) *Nature*, **320**, 367–371.

Gallo,R.C., Wong-Staal,F., Montagnier,L., Haseltine,W.A. and Yoshida,M. (1988) *Nature*, **333**, 504.

Goh,W., Sodroski,J., Rosen,C. and Haseltine,W. (1987) *J. Virol.*, **61**, 633–637.

Gorman,C.M., Moffat,L.F. and Howard,B.H. (1982) *Mol. Cell. Biol.*, **2**, 1044–1051.

Göttlinger,H.G., Sodroski,J. and Haseltine,W.A. (1987) *J. Virol.*, **61**, 633–637.

Gubler,U. and Hoffman,B. (1983) *Gene*, **25**, 263–269.

Guyader,M., Emerman,M., Sonigo,P., Clavel,F., Montagnier,L. and Alizon,M. (1987) *Nature*, **326**, 662–669.

Hopp,T. and Woods,K. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 3824–3828.

Laemmli,U.K. (1970) *Nature*, **227**, 680–685.

Lambert,P.F., Spalholz,B.A. and Howley,P.M. (1987) *Cell*, **50**, 69–78.

Lambert,P.F., Hubbert,N.L., Howley,P.M. and Schiller,J.T. (1989) *J. Virol.*, **63**, 3151–3154.

Malim,M.H., Hauber,J., Fenrick,R. and Cullen,B.R. (1988) *Nature*, **335**, 181–183.

Malim,M.H., Boehnlein,S., Hauber,J. and Cullen,B.R. (1989) *Cell*, **58**, 205–214.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Melton,D.A., Krieg,P.A., Rebagliati,M.R., Maniatis,T., Zinn,K. and Green,M.R. (1984) *Nucleic Acids Res.*, **12**, 7035–7056.

Meyerhans,A., Cheynier,R., Albert,J., Seth,M., Kwok,S., Sninsky,J., Mortfeldt-Manson,L., Asjo,B. and Wain-Hobson,S. (1989) *Cell*, **58**, 901–910.

Mount,S.M. (1982) *Nucleic Acids Res.*, **10**, 459–472.

Muesing,M., Smith,D., Cabradilla,C., Benton,C., Lasky,L. and Capon,L. (1985) *Nature*, **313**, 450–457.

Myers,G., Rabson,A.B., Josephs,S.F., Smith,T. and Wong-Staal,F. (eds) (1989) *Human Retroviruses and AIDS 1989. A Compilation and Analysis of Nucleic Acid and Amino Acid Sequences*. Los Alamos National Laboratory, Los Alamos, NM.

Ratner,L., Fisher,A., Jagodinski,L.L., Mitsuya,H., Liou,R., Gallo,R.C. and Wong-Staal,F. (1987) *AIDS Res. Human Retroviruses*, **3**, 57–69.

Rosen,C., Sodroski,J. and Haseltine,W. (1985) *Cell*, **41**, 813–823.

Rosen,C., Terwilliger,E., Dayton,A., Sodroski,J. and Haseltine,W. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 2071–2075.

Seigel,L.J., Ratner,L., Josephs,S.F., Derse,D., Feinberg,M.B., Reyes,G.R., O'Brien,S.J. and Wong-Staal,F. (1986) *Virology*, **148**, 226–231.

Sharp,P.A. (1987) *Science*, **235**, 766–771.

Sodroski,J., Patarca,R., Rosen,C., Wong-Staal,F. and Haseltine,W. (1985) *Science*, **229**, 74–77.

Sodroski,J., Goh,W., Rosen,C., Dayton,A., Terwilliger,E. and Haseltine,W. (1986) *Nature*, **321**, 412–417.

Stephens,C. and Harlow,E. (1987) *EMBO J.*, **6**, 2027–2035.

Terwilliger,E., Burghoff,R., Sia,R., Sodroski,J., Haseltine,W. and Rosen,C. (1988) *J. Virol.*, **62**, 655–658.

Viglianti,G.A. and Mullins,J.I. (1988) *J. Virol.*, **62**, 4523–4532.

Weimer,Th., Salfeld,J. and Will,H. (1987) *J. Virol.*, **61**, 3109–3113.

Received on October 20, 1989; revised on December 13, 1989