The EMBO Journal vol.13 no.11 pp.2650–2659, 1994

Evolution of a disrupted TAR RNA hairpin structure in the HIV-1 virus

Bep Klaver and Ben Berkhout[1]

University of Amsterdam, Department of Virology, Academic Medical Center, Meibergdreef 15, 1105 AZ Amsterdam, The Netherlands
[1]Corresponding author

Communicated by L.Bosch

Human immunodeficiency virus type 1 (HIV-1) RNA contains an extended hairpin structure at the 5′ end (the TAR element) that is essential for viral replication. The upper part of the stem–loop structure binds the virally encoded transcriptional activator protein Tat and cellular co-factors, but no clear function for the lower stem region has been established. Here, we report that mutant HIV-1 viruses with base substitutions in the lower stem region are dead, most likely at the level of transcription from an integrated provirus. By using large amounts of these mutant DNA constructs for transfections, revertant viruses with a great variety of genetic changes (point mutations, short deletions) could be isolated in prolonged culture experiments that lasted over 6 months. The pattern and evolution of these changes supported the notion that base-pairing of the lower stem region is essential for optimal HIV-1 replication. The functional and genetic plasticities of this RNA domain and the HIV-1 long terminal repeat promoter are discussed.

Key words: HIV-1 RNA structure/molecular evolution/TAR RNA function

Introduction

The human and simian immunodeficiency viruses (HIV and SIV) have a complex genome with a transcriptional promoter that is regulated both by viral and cellular proteins. The viral *trans*-activator protein is Tat, which is essential for viral replication (Dimitrov *et al.*, 1993, reviewed in Vaishav and Wong-Staal, 1991; Cullen, 1992). Tat is an extremely potent activator of transcription from the upstream long terminal repeat element (5′ LTR). A variety of experimental model systems has been used to study the function of Tat. However, many molecular details of the Tat-mediated *trans*-activation mechanism are still unknown and the exact mode of Tat action remains somewhat controversial. Tat has been reported to act as an anti-terminator, a transcriptional elongation factor, or as inducer of transcriptional initiation (reviewed in Vaishav and Wong-Staal, 1991; Cullen, 1992; Jeang *et al.*, 1993a). Perhaps most intriguing, Tat is tethered to the transcriptional machinery through binding to a hairpin structure formed at the 5′ end of nascent HIV transcripts (Berkhout *et al.*, 1989; Dingwall *et al.*, 1989). Like Tat, this *trans*-activation response element (TAR RNA) is essential for replication of the HIV-1 virus in T cell lines (Leonard *et al.*, 1989; Harrich *et al.*, 1990; Berkhout and Klaver, 1993), although TAR-independent replication has been reported in phorbol ester stimulated T lymphocytes (Harrich *et al.*, 1990) and astrocytic glial cells (Bagasra *et al.*, 1992; Taylor *et al.*, 1992). Since this work concentrates on the biology of this particular RNA structural element, we will summarize briefly the functional data available on TAR RNA.

Extensive mutational analyses have identified the sequence and structure requirements for an active TAR element. TAR forms an RNA hairpin structure with a well conserved 3 nt bulge and 6 nt loop. Essential nucleotide sequences are present in the single-stranded bulge and loop (Feng and Holland, 1988; Berkhout and Jeang, 1989), but also in the base-pairs flanking the bulge (Berkhout and Jeang, 1991; Weeks and Crothers, 1991). Tat interacts directly with the bulge region of the TAR element, thus explaining the phenotype of bulge mutants (Dingwall *et al.*, 1989; Frankel *et al.*, 1989; Cordingley *et al.*, 1990; Roy *et al.*, 1990; Weeks *et al.*, 1990; Mann and Frankel, 1991; Churcher *et al.*, 1993). The loop is supposed to bind a cellular factor that is involved in the *trans*-activation mechanism. Several cellular proteins that bind specifically to the loop *in vitro* have now been described (Marciniak *et al.*, 1990; Sheline *et al.*, 1991; Wu *et al.*, 1991) and several other TAR-binding proteins (Gatignol *et al.*, 1991; Sheline *et al.*, 1991; Gunnery *et al.*, 1992; Rounseville and Kumar, 1992; Masuda and Harada, 1993) and Tat-binding proteins (Nelbock *et al.*, 1990; Desai *et al.*, 1991; Jeang *et al.*, 1993b; Ohana *et al.*, 1993) have been identified. All of these observations suggest that Tat activation *in vivo* may require not only Tat binding to TAR RNA but also the interactions of multiple cellular factors with nascent TAR RNA, LTR DNA and Tat protein.

Trans-activation assays in transiently transfected tissue culture cells indicate that the upper half of the TAR hairpin (nt +19 to +42) is sufficient for function (Hauber and Cullen, 1988; Jakobovits *et al.*, 1988; Berkhout and Jeang, 1989; Berkhout *et al.*, 1989). The lower part of the stem also contributes to Tat-mediated transcriptional activation of the LTR promoter (Garcia *et al.*, 1989; Selby *et al.*, 1989). This stem region is well conserved among different members of the HIV–SIV family of retroviruses (Berkhout, 1992; Myers *et al.*, 1992), suggesting a particular role for this stem in the viral life cycle. In order to study the biological relevance of the full-length TAR hairpin, we performed infections with a mutant HIV-1 genome containing a substitution in the +3/+16 region. This mutation is expected to disrupt the lower part of the TAR stem, while leaving the critical upper part intact. This mutant was shown previously to be fully transcriptionally active in transient LTR-CAT transfection assays (Berkhout *et al.*, 1989). It was therefore surprising to find that the same mutation fully blocked viral replication. Transfected cells were cultured for up to 6 months in order to select for revertant viruses. Two

pathways of RNA structure reversion are described in detail. The combined data suggest that a closed, but not necessarily full-length, TAR stem is critical for transcription of an integrated HIV-1 provirus.

Results

The lower part of the TAR RNA stem is essential for viral replication

Figure 1 shows the LTR DNA sequence and TAR RNA secondary structure of wild-type HIV-1 and the Xho+10 mutant. This mutant is expected to retain the natural, albeit extensively abbreviated, TAR RNA hairpin. We introduced the Xho+10 substitution in both 5′ and 3′ LTRs of the full-length, infectious pLAI plasmid (Peden *et al.*, 1991). The wild-type and mutant HIV-1 plasmids were transfected by means of electroporation into the SupT1 T cell line and viral replication was monitored by the production of virus in the medium using Gag p24 ELISA (Figure 2). The Xho+10 mutant was able to produce virus transiently (Figure 2), consistent with previous results obtained with LTR-CAT constructs in transient transfection assays (Berkhout *et al.*, 1989). This Xho+10 virus, however, was unable to establish a productive infection in the CD4-positive SupT1 cell line. In contrast, rapid viral replication was measured in the culture transfected with wild-type HIV-1 and massive Env protein-induced syncytia were observed at day 4.

We decided to perform a large-scale transfection and to continue culturing the transfected cells in order to select for spontaneous revertant viruses. Transfection I was performed with 25 μg Xho+10 DNA, but again no viral replication was observed in the first month of culturing (no Gag p24, no syncytia). Over the next 5 months, increased virus replication was measured and small syncytia were first detected around day 45. We compared the replication kinetics of viruses sampled over time in a SupT1 infection using identical amounts of input virus (Figure 3). These results clearly indicate a gradual increase in the replication potential of revertants over time. Corresponding cell samples were used to isolate and analyze proviral DNA sequences present at days 32, 53, 82 and 154 (Table I). Because mutations responsible for the reversion event are expected to localize within the TAR region, we minimally sequenced HIV-1 LTR

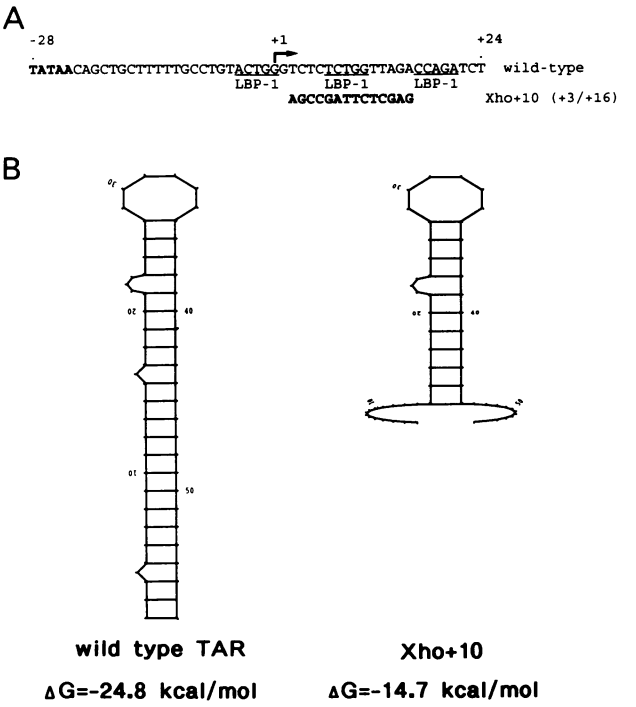

Fig. 1. LTR DNA sequence and TAR RNA structure models for wild-type HIV-1 and the Xho+10 mutant. (**A**) The wild-type LTR DNA sequence is shown from the TATAA box (indicated in bold) at position −28 up to position +24. Nucleotide numbers are relative to the transcriptional start site at +1. This position marks the border between the upstream U3 and downstream R regions of the LTR. Three binding sites for the cellular LBP-1 protein are indicated. The +3/+16 substitution present in the Xho+10 mutant is shown below in bold. This sequence contains a diagnostic *Xho*I restriction site at positions +11 to +16. (**B**) Secondary RNA structure models as predicted by the Zuker algorithm (Zuker, 1989). The calculated free energy of the stem−loop structures is indicated below. There is experimental support for the wild-type TAR structure as shown (Muesing *et al.*, 1987; Berkhout *et al.*, 1989; Baudin *et al.*, 1993). Both wild-type and Xho+10 fold the upper part of the TAR stem−bulge−loop structure, including the minimal TAR element (positions +19 to +42) as defined in transient transfection assays.

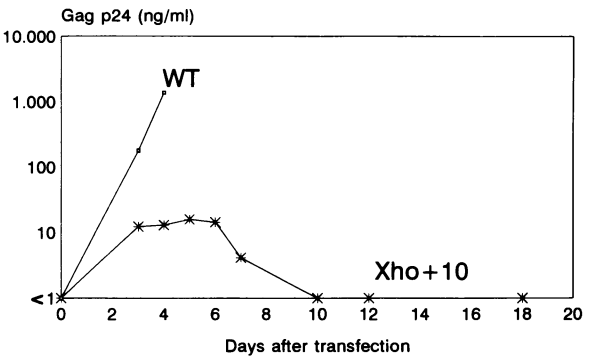

Fig. 2. Replication potential of the Xho+10 mutant virus in a T cell line. The SupT1 cell line was electroporated with 5 μg of the wild-type (WT) and Xho+10 mutant plasmids. Supernatant samples were collected at the times indicated and assayed for the HIV-1 capsid protein Gag p24. Note the logarithmic scale used to plot Gag p24 values. Identical results were obtained in reverse transcriptase enzymatic assays (not shown). Similar phenotypes were scored in transfections of other T cell lines (H9 and A3.01) and peripheral blood lymphocyte cultures (data not shown).

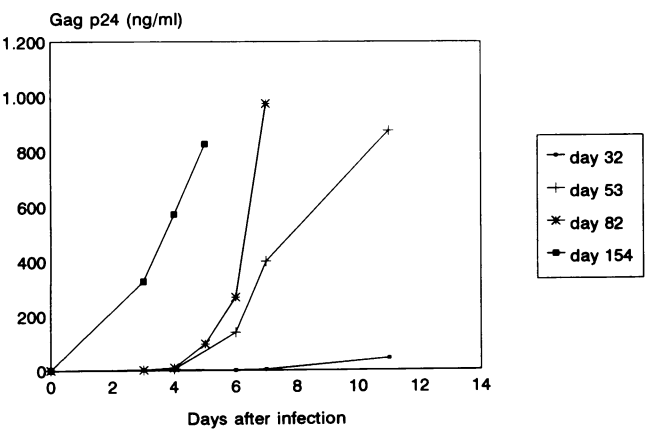

Fig. 3. Growth kinetics of revertant viruses isolated at different time points after transfection. Frozen supernatant samples taken at days 32, 53, 82 and 154 post-transfection were analyzed for Gag p24 levels and subsequently used to infect SupT1 cells. For all samples, equal amounts of input virus were used (2.5 ng Gag p24). Virus replication was assayed by Gag p24 ELISA.

DNA from the TATAA box at position −28 up to the *Hin*dIII site at position +77. Multiple clones were sequenced for every time point and the analysis was performed on three parallel cultures that were split at day 10 post-transfection (A, B and C in Table I). It is therefore expected that the initial steps of the reversion pathway will be identical for these three samples, which is indeed what we observed.

An increasing number of mutations were found in the Xho+10 genome over time (Table I). Most mutations were positioned in the TAR domain that carried the original Xho+10 substitution. One cannot exclude, however, the possibility that additional critical mutations are present in other regions of the HIV-1 genome, notably in the LTR promoter or the gene encoding the Tat *trans*-activator protein. To verify that these TAR sequence changes were sufficient for the restoration of viral replication, some revertant TAR genotypes were introduced into the replication-incompetent Xho+10 plasmid. These DNA constructs were tested for the ability to generate replication-competent virus after introduction of the DNA into SupT1 cells (Figure 4). The recombinant viruses yielded the expected phenotypes. The Ia2 variant isolated at a very early time point (day 32) produced no detectable virus, but rapid replication kinetics were measured for the Id15 variant isolated at day 154. In addition, we measured slow virus replication for the IIa6 revertant that was obtained at day

Table I. Nucleotide sequences of revertant viruses in evolution experiment I

		-28	-8	+1	+33	+53	+73	+82	
				Xho+10					
Xho+10		**TATAA**GCAGCTGCTTTTTGC	CTGTACTGGG**AGCCGATTCT**	**CGAG**CCAGATCTGAGCCTGG	GAGCTCTCTGGCTAACTAGG	GAACCCACTGCTTAAGCCTC	AATA**AAGCTT**		day
Ia1			------G----------						
Ia2									
Ia3									
Ia4									32
Ia5									
Ia6									
Ia7									
Ia8			------G----------						
Ic5	A		----------TT--------						
Ic6			----------TT--------						82
Ic7			----------TT--------						
Ic8			----------TT--------						
Id5			--------C-TT--------						
Id6			------G---TT--------						154
Id7			------G---TT--------						
Id8			--------C-TT--------						
Ib7		-------------------G							
Ib8									
Ib9									
Ib10									53
Ib11									
Ib12			----------T---------						
Ib13									
Ib14			----------T---------				-----T----------		
Ic14	B		------G-T-TT--------						
Ic15			------G---TT--------						
Ic16			--------T-TT--------						82
Ic17			--------C-TT--------						
Ic18			------G-C-TT--------						
Id9			------G-T-TT---G----						
Id10			------G-T-TT---G----						154
Id11			--------C-TT---G----						
Id12			--------C-TT---G----						
Ia9									
Ia10						A-			
Ia11									
Ia12									32
Ia13									
Ia14									
Ia15									
Ia16			--------A----------						
Ib1			----------TT--------						
Ib2			----------TT--------						
Ib3			----------TT--------						53
Ib4			----------TT--------						
Ib5	C		----------TT--------						
Ib6			----------TT--------						
Ic9			----------TT--------						
Ic10			----------TTA-------						
Ic11			----------TT--------						82
Ic12			----------TT--------						
Ic13			----------TT--------						
Id13			------G---TT--------						
Id14			------G---TT--------						154
Id15			------G---TT---G----						
Ie1			--------C-TT---G----			C-A-			
Ie2			--------C-TT---G----	----▲----------		C-A-			day 40
Ie3	mix		--------C-TT---G----			C-A-			after
Ie4			--------C-TT---G----			A-			mixing
Ie5			--------C-TT---G----	----------C--------		C-A-			

The complete nucleotide sequences of the Xho+10 mutant from position −28 to position +82 are shown. The transcriptional start site at +1 marks the border between the upstream U3 and downstream R domains of the HIV-1 LTR. A line above the sequence indicates the Xho+10 substitution. Indicated in bold are the TATAA box (position −28), Xho+10 substitution (+3/+16) and *Hin*dIII site (+77). The name of the culture (A, B or C) and the variant is indicated on the left, the day of isolation is shown in the right-hand column. Deletions are indicated by ▲. The Ie isolates were obtained after mixing of three virus populations (day 154 supernatants of cultures A, B and C), followed by 40 days of co-culture.

44 of a second transfection experiment (see Table II and Discussion). These results indicate that genetic changes in the TAR region can fully account for the phenotypic reversion.

RNA structural analysis of revertant viruses: repair of the TAR stem

In order to study evolution of the disrupted Xho+10 TAR structure, it is important to note that, although the mutation was designed to present the lower part of the TAR stem as two single-stranded domains, alternative base-pairing schemes may exist. Computer analysis of the secondary structure of mutant Xho+10 RNA using the Zuker algorithm (Zuker, 1989) showed that this RNA can form three alternative structures with comparable stability (Figure 5, left). All three structures conserve the upper part of the TAR hairpin, but differ in the lower domain. This domain is either fully single-stranded ($\Delta G = -14.7$ kcal/mol), or partially structured using two alternative base-pairing interactions. First, complementary segments in the +10 and +50 regions can recombine to form a new stem segment with five additional base-pairs ($\Delta G = -16.1$ kcal/mol). Second, the +10 sequences may fold into a small hairpin, leaving the +50 sequences single-stranded ($\Delta G = -15.4$ kcal/mol). We will first discuss the structural rearrangements seen over time in the transfection experiment I (Table I and upper part of Figure 5). A second reversion pathway is described in detail below (Table II and lower part of Figure 5).

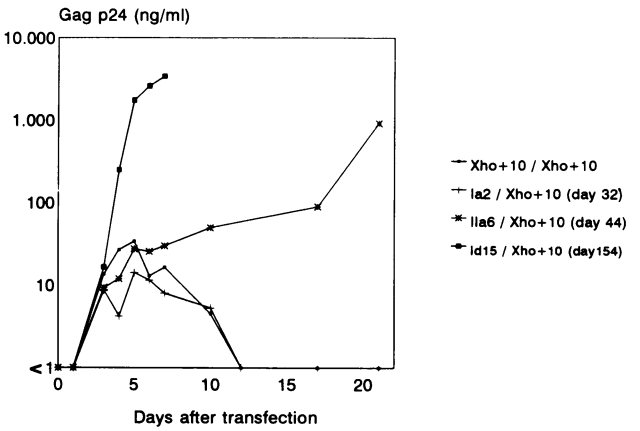

Fig. 4. Mutations in the TAR region account for the phenotypic reversion of Xho+10. The TAR sequences of variants Ia2, IIa6 and Id15 were introduced into the 5′ LTR of the Xho+10 construct. These plasmids were transfected into SupT1 cells by means of electroporation. Supernatant samples were collected at the times indicated and assayed for Gag p24. Note the logarithmic scale used to plot Gag p24 values. The progeny is expected to inherit the R sequences of the 5′ LTR and U3 sequences of the 3′ LTR (Klaver and Berkhout, 1994). To verify this, we analyzed proviral DNA samples at day 5 (Id15 transfection) and day 12 (IIa6 transfection). Indeed, of the multiple mutations present in these two variants (see Tables I and II), only the mutations downstream of +1 were inherited. The sudden increase in replication rate of variant IIa6 around day 20 suggests that an additional mutation was acquired by this progeny during culturing.

The first mutation observed (G51→A) results in destabilization of the most stable Xho+10 conformation ($\Delta G = -16.1$ kcal/mol). Thus, TAR folding is shifted to either of the two alternative metastable conformations ($\Delta G = -14.7$ kcal/mol or $\Delta G = -15.4$ kcal/mol; the latter is shown in Figure 5 for the Ia2 variant). As shown in Figure 4, such a variant is unable to replicate efficiently. Two subsequent mutations (A3→U and G4→U) make it possible for a TAR-like stem to fold, with a substantial increase in stability (Ib12, $\Delta G = -16.0$ kcal/mol; Ib1, $\Delta G = -20.9$ kcal/mol). The resulting hairpin is closed by six base-pairs, but maintains a large internal loop element. Upon prolonged culturing, a single A8→G mutation allows four additional base-pairs to be formed in this region, with an increase in stability of −1.0 kcal/mol (Id15, $\Delta G = -21.9$ kcal/mol). This TAR variant fully supports HIV-1 replication (Figure 4).

The reversion pathway I as depicted in Figure 5 holds for all three culture samples (A, B and C), but not all mutations can be incorporated into this model. For instance, we observed an additional G+1→C/U mutation in samples A and B, which will destabilize the structure as shown for Ib1 by 1.1 kcal/mol. Interestingly, in a competition experiment, the C+1 variant was able to outgrow the other viruses within 40 days of co-culture (Table I), suggesting a strong selective advantage of this C+1 change. Because this mutation maps

Table II. Nucleotide sequences of revertant viruses in evolution experiment II

	-28	-8	+1 (Xho+10)	+33	+53	+73	+82	
Xho+10	TATAAGCAGCTGCTTTTTGC	CTGTACTGGGAGCCGATTCT	CGAGCCAGATCTGAGCCTGG	GAGCTCTCTGGCTAACTAGG	GAACCCACTGCTTAAGCCTC	AATAAAGCTT		
IIa1	----------	---------A----------	----------	----------	---T----------	----------		day
IIa2	----------	---------AA----------	----------	----------A	----------	----------		44
IIa5	----------	---------A----------	----------	----------A	----------	----------		
IIa6	----------	---------A----------	----------	----------A	----------	----------		
IIa'1	----------	▲▲▲▲▲----A----------	----------	----------A	----------	----------		
IIa'2	----------	▲▲▲▲▲----A----------	----------	----------A	----------	----------		
IIa'3	----------	▲▲▲▲▲----A----------	----------	----------A	----------	----------		54
IIa'4	----------	▲▲▲▲▲----A---G------	----------	----------A	----------	----------		
IIa'5	----------	▲▲▲▲▲----A---G------	----------	----------A	----------	----------		
IIb10	----------	▲▲▲▲▲----A---G------	----------	----------A	----------	----------		
IIb11	----------	▲▲▲▲▲---AA---G------	----------	----------A	----------C---------	----------		
IIb12	----------	▲▲▲▲▲---AA---G------	----------	----------A	----------	----------		79
IIb13	----------	▲▲▲▲▲---AA---G------	----------	----------A	----------	----------		
IIb14	----------	▲▲▲▲▲----A---G------	----------	----------A	----------G---------	----------		
IIc12	----------	▲▲▲▲▲---AA---G------	----------T---------	----------A	----------	----------		
IIc13	----------C--	▲▲▲▲▲---AA---G------	----------T---------	----------A	----------	----------		
IIc14	----------	▲▲▲▲▲---AA---G------	----------T---------	----------A	----------	----------		117
IIc15	----------	▲▲▲▲▲---AA---G------	----------T---------	----------A	----------	----------		
IIc16	----------	▲▲▲▲▲---AA---G------	----------T---------	----------A	----------	----------		

See Table I for details. The 5 nt deletion is shown at positions −8/−4, but due to duplication of the CTG sequence in this region, the deletion can arbitrarily be defined as −8 to −4, −7 to −3, −6 to −2, or −5 to −1. Inspection of the sequences surrounding the mutation may suggest a reverse transcriptase-mediated mechanism for deletion. Upon copying of the $^{-3}$CTG sequence, the nascent cDNA may have dissociated, followed by reannealing to the short direct repeat $^{-8}$CTG (Pilsinelli and Temin, 1991).

in the transcriptional initiation region, we analyzed the +1 usage of these variants by primer extension analysis (Figure 6). However, we could not detect a difference in start site usage for these mutants relative to wild-type HIV-1. Although we currently do not understand the selective advantage of the C+1 change, we note that this mutation does not disprove the overall RNA evolution scheme depicted in Figure 5. In fact, the final samples obtained in the C+1 culture had acquired an additional mutation (Table I, U49→C) that further stabilizes the proposed structure by changing a G-U into a G-C pair with an increase in stability of −1.7 kcal/mol. Thus, despite the G+1→C change, this structure will regain a stability (ΔG = −21.5 kcal/mol) comparable to that of the Id15 variant (ΔG = −21.9 kcal/mol). These results indicate a strong evolutionary pressure to close an opened TAR stem.

An alternative pathway for reversion

In order to achieve a more fundamental mechanistic understanding of the requirements for a closed TAR stem, we performed a second Xho+10 reversion experiment (Transfection II). Again, it took more than a month of culturing to detect some level of virus replication, and fully replication-competent virus was selected over a 4 month period. Samples were analyzed at days 44, 54, 79 and 117 (Table II and the lower panel of Figure 5). This pathway II seemed to have been initiated on the basis of the most stable Xho+10 conformation (Figure 5, ΔG = −16.1 kcal/mol). The G52→A mutation present at day 44 changes a U-G base-pair to U-A, with a concomitant increase in thermodynamic stability (IIa6, ΔG = −17.3 kcal/mol). We measured low levels of virus replication for this mutant (Figure 4). A putative problem with the IIa6 structure is the single-stranded character of the 5′ end. Furthermore, because a large 5 nt bulge is present on the 3′ side of the stem, only two 3′ nucleotides (positions 56 and 57) remain for further repair of the stem. Of course, HIV-1 RNA is extended beyond position 57, but these sequences are not expected to be available for TAR-repair because they are involved in a hairpin structure containing the poly(A) signal (Muesing *et al.*, 1987). Consistent with this idea, we never observed fixation of mutations in the region downstream of position 57.

The next step in evolution scheme II came as a surprise. A deletion of five nucleotides in the region between TATAA box (−28) and transcriptional initiation site (+1) was observed at day 54, followed by selection of a C6→G change at days 54−79. Because RNA polymerase always chooses the start site at ~28 nt downstream from the TATAA box (Gluzman *et al.*, 1980; Benoist and Chambon, 1981), the deletion may result in a downstream shift of transcriptional initiation. This was verified by primer extension analysis of viral RNA isolated from infected cells at different time points along pathway II (Figure 6). The day 52 sample exhibited a somewhat diffuse initiation pattern around positions +6 and +11 (lane 11). The day 77 sample, however, showed

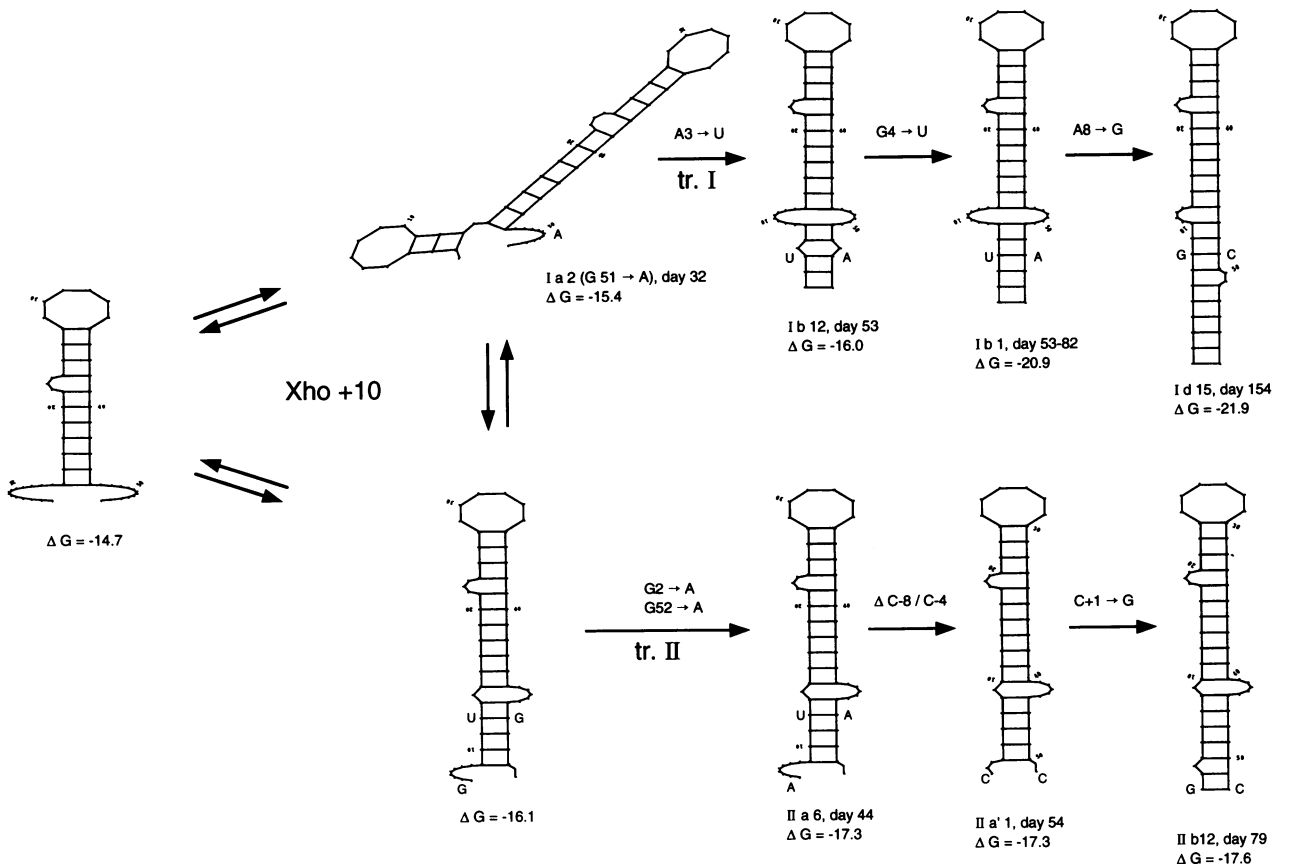

Fig. 5. Repair of a disrupted TAR stem. Shown are schematic drawings of the RNA secondary structures of mutant Xho+10 (left) and revertants observed in evolution experiments I (right, upper part) and II (right, lower part). The free energies of the structures and the day at which revertants were first detected are indicated. Three alternative base-pairing schemes are shown for the Xho+10 mutant. For the revertants only the structure of lowest free energy is shown. The mutated nucleotides and newly formed base-pairs are indicated.

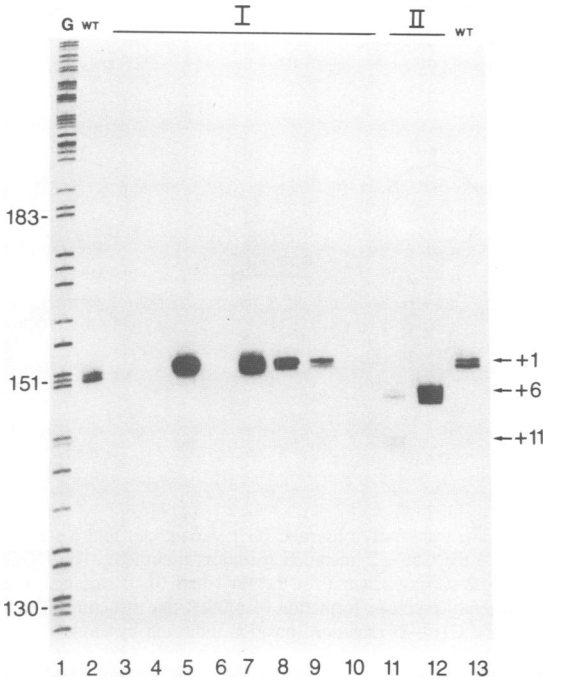

Fig. 6. Primer extension analysis of transcripts of wild-type and revertant viruses. Total cellular RNA was isolated and analyzed with a primer complementary to the U5 region of HIV-1 transcripts. This primer generates a 151 nt cDNA product on wild-type HIV-1 transcripts (lanes 2 and 13). Samples were from transfection I, culture C at days 34, 60 and 150 (lanes 3, 4 and 5, respectively), culture A at days 92 and 150 (lanes 7 and 8, respectively), culture B at day 150 (lane 9); and transfection II, days 52 and 77 (lanes 11 and 12, respectively). The RNA of mock-infected SupT1 cells was analyzed in lanes 6 and 10. For reference, a dideoxy sequencing lane (G-track only) of the wild-type HIV-1 plasmid using the same primer is shown (lane 1). The +1 start and alternative initiation sites (+6 and +11) are indicated on the right.

a more precise +1 usage at the $^{+6}$GGA sequence (lane 12). Note that the numbering is changed accordingly for the truncated IIa′1 variant in Figure 5. These results suggest that the deletion forces the transcriptional apparatus to search for an appropriate downstream initiation site. Initially, the +1 usage is rather diffuse, but it is markedly improved upon acquisition of the C+1→G mutation. Since the selected G actually forms the new start position, and because wild-type HIV-1 also uses a G+1, it can be argued that HIV-mediated transcriptional initiation favors a G at the +1 position.

The rationale for selection of the deletion variant comes from analysis of the RNA structure. 5′ Truncation is an ideal way of removing the dangling 5′ end, although this modification does not increase the overall stability of TAR RNA (IIa′1, ΔG = −17.3 kcal/mol). The subsequent C+1→G mutation results in base-pairing of the bottom part of the stem and modestly increases its stability (IIb12, ΔG = −17.6 kcal/mol). Thus, the results obtained in the two reversion experiments are in accordance with the hypothesis that a base-paired 5′ end is important for efficient TAR function in HIV infection assays.

Replication characteristics of revertant viruses

The revertants obtained in evolution experiment I finally reached TAR RNA stabilities (ΔG = −21.9 kcal/mol) comparable to that of wild-type TAR (ΔG = −24.8 kcal/mol). The observation that the shortened TAR hairpin obtained in transfection II (ΔG = −17.6 kcal/mol) was compatible with virus replication suggests that the length and overall stability of the TAR stem are less important than having an HIV-1 transcript with a base-paired 5′ end. We have continued both phylogeny experiments beyond the sixth month of culture, but so far no additional mutations have been detected in the TAR sequences (data not shown). Through the first months of culture, approximately one advantageous mutation became fixed in the viral population

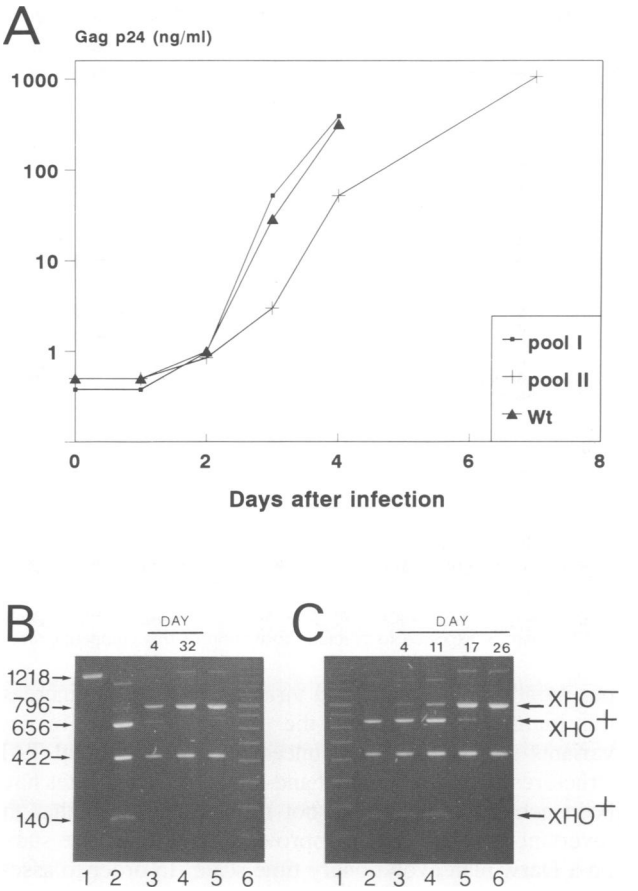

Fig. 7. Replication kinetics of revertant and wild-type viruses. (**A**) Growth curves of revertant pool I (day 194 sample), pool II (day 137 sample) and wild-type HIV-1 virus. Equal amounts of input virus (0.5 ng Gag p24 each) were used to infect SupT1 cells. (**B**) A mixture of the wild-type and Id15 molecular clones (1 μg each) was electroporated into SupT1 cells. Total cellular DNA was isolated at days 4 and 32 after transfection (indicated at the top of the panel). Viral LTR sequences were PCR-amplified, digested with *XhoI* and analyzed on agarose gels (lanes 3 and 4, respectively). Control PCR samples were generated on the Id15 LTR (Xho$^+$, lane 2) and the wild-type LTR (Xho$^-$, lane 5) as templates. The PCR fragment is 1218 bp in length (lane 1), and will produce two or three bands depending on the presence of the Xho+10 restriction site (either 422 + 796 or 422 + 656 + 140). A 100 bp ladder is shown in lane 6. (**C**) SupT1 cells were infected with a mixture of three viruses; pool I (day 194), pool II (day 137) and wild-type HIV-1 LAI (0.5 ng Gag p24 each). LTR sequences were PCR-amplified from chromosomal DNA at days 4, 11, 17 and 26 post-infection and analyzed for the presence of the Xho+10 site (lanes 3−6, days indicated at the top of the panel). Lane 2 contains a control Xho$^+$ sample (Id15 LTR DNA) and lane 1 a 100 bp ladder. The positions of the Xho$^+$ and Xho$^-$ specific bands are indicated on the right.

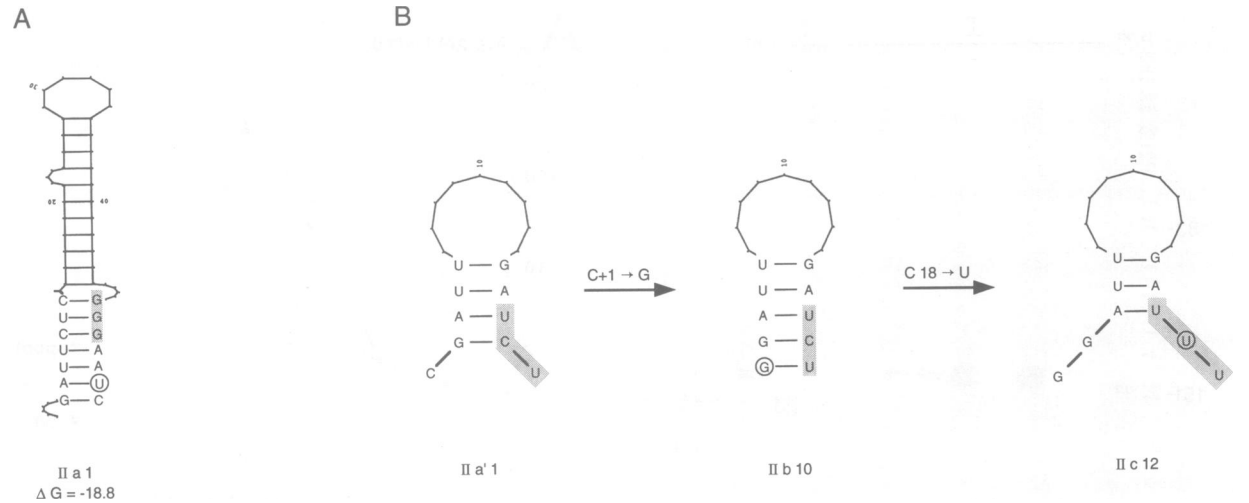

Fig. 8. Alternative base-pairing interactions determine the fate of some revertants. (**A**) The transiently observed IIa′1 variant can fold a stem similar to Xho+10 (Figure 5, $\Delta G = -16.1$ kcal/mol) but with two additional base-pairs due to the C56→U mutation (position is circled). The ^{51}GGG motif, which is now fully engaged in base-pairing, is shaded in gray. (**B**) Alternative base-pairing scheme for the IIa′1 and IIb10 variants. The TAR-bulge element (UCU, marked by gray boxes) is engaged in base-pairing that will transiently preclude formation of a TAR-like structure. The next IIc12 variant is expected to preclude formation of this competing structure because of the C18→U mutation (position indicated by circle).

per month. While enhanced viral replication will increase this mutational frequency, the growth advantage of new variants will be less pronounced once (sub)optimal TAR structures have been formed and high replication rates have been achieved. This does not necessarily mean that the revertants cannot further improve and that they are stable on a Darwinian evolutionary time scale. In order to assess the replication potential of the revertant viruses we performed several experiments.

Revertant virus stocks obtained at days 194 and 137 in evolution experiments I and II, respectively, and the wild-type HIV-1 virus were used to infect the SupT1 T cell line and virus production was assayed over time (Figure 7A). No major differences in replication kinetics were observed between wild-type and class I revertants. However, a marked delay in virus production was observed for revertant II, suggesting that a truncated TAR stem can only function suboptimally. In order to measure accurately relatively small differences in virus production we performed sensitive co-culture experiments in which the fittest virus will outgrow other variants as a function of time (Koken *et al.*, 1992).

In the first competition experiment we transfected T cells with an equimolar plasmid mixture (1 μg each) of wild-type HIV-1 and the class I revertant Id15 (Table I, Figure 5). Virus production was monitored by the appearance of syncytia, at which time point a sample of the cell-free supernatant was used to infect fresh T cells. Total cellular DNA was isolated at different time points and viral LTR sequences were PCR-amplified, digested with *Xho*I and subsequently analyzed on agarose gels (Figure 7B). Control digests indicate that this test is diagnostic for the presence of wild-type (Xho$^-$, lane 5) or Id15 genomes (Xho$^+$, lane 2). While the Id15 virus was present in the initial phase of the co-culture (day 4 after transfection, lane 3), we could only detect wild-type virus after 32 days (lane 4).

In the second competition assay, we infected T cells with an equimolar mixture of three virus stocks; pool I (day 194 sample), pool II (day 137) and the wild-type virus. LTR DNA was isolated at several time points after infection and analyzed for the Xho phenotype (Figure 7C). Both type I and II revertants stably contain the Xho$^+$ restriction site. Based on the individual growth curves (Figure 7A) and the competition experiment of Figure 7B, a loss of both revertant viruses is expected, which should result in an increased Xho$^-$:Xho$^+$ ratio. Indeed, we observed a gradual increase in this ratio upon prolonged culturing (days 4, 11, 17 and 26 in lanes 3–6). The wild-type virus was able to outcompete the two TAR variants within 3 weeks, as Xho$^-$ DNA became the predominant component of progeny LTR sequences. Based on these direct competition assays we estimate a small (~10%) difference in virus production levels between revertant I and wild-type HIV-1 (Koken *et al.*, 1994).

Sequences resembling the TAR loop are not allowed in a single-stranded conformation

The two independent reversion pathways described in Figure 5 suggest that a mutation in the ^{51}GGG^{53} sequence occurs as an obligatory first step of reversion. A G51→A change was detected early in experiment I and a G52→A mutation was observed in experiment II. Although these individual changes do not fully restore the replication defect of the Xho+10 mutant (Figure 4), their rapid appearance does suggest that these mutations form a prerequisite for the subsequent phenotypic reversion. As discussed above, we can understand the G52→A mutation in terms of TAR structure evolution (Figure 5, clone IIa6), but no such effect is apparent for the G51→A mutation. In fact, the selected A51 is maintained as an unpaired nucleotide in the stable RNA structure present at the end of evolution experiment I (Id15, Figure 5).

Inspection of the sequences surrounding the ^{51}GGG element may suggest an alternative explanation for the rapid mutation of this motif when present in combination with the Xho+10 substitution. We observed a striking sequence homology between the ^{51}GGG region and the loop domain of the TAR hairpin (^{31}GGG). It is therefore possible that the ^{51}GGG sequence is toxic in the single-stranded Xho+10

conformation because it may abusively bind proteins that should recognize the TAR loop. Alternatively, the free ^{51}GGG domain may interfere with the proper folding of other RNA structures in HIV-1 transcripts. It follows that mutation of the pseudo-loop sequences in the Xho+10 context will help to restore TAR function. To verify that the ^{51}GGG sequence is not toxic in the closed conformation of wild-type TAR RNA, we introduced the G51→A mutation in the wild-type pLAI plasmid. We did not observe any difference in viral replication kinetics between wild-type HIV-1 and the +51 mutant in SupT1 transfections (data not shown).

There is some additional evidence for the hypothesis that the ^{51}GGG sequence is a negative element in the context of an opened lower TAR stem. Of all the revertant sequences obtained, one variant disobeyed the 'rule' of first inactivating the ^{51}GGG domain (variant IIa1, Table II). This sequence was transiently detected at day 44, and was apparently outcompeted by other variants present in the culture. Analysis of the RNA secondary structure of the IIa1 intermediate indicated that this mutant had followed a different strategy to neutralize the exposed ^{51}GGG region (Figure 8A, $\Delta G = -18.8$ kcal/mol). The ^{51}GGG motif and downstream nucleotides are base-paired in a relatively stable stem structure. This structure is based on the most favorable Xho+10 conformation (Figure 5, $\Delta G = -16.1$ kcal/mol), with a significant increase in stability due to the C56→U change and formation of two additional base-pairs.

Competing RNA structures are discriminated against

A surprising base change was observed in the final phase of evolution experiment II (Table II). The C23→U mutation affects the sequence of the trinucleotide TAR bulge, without having structural consequences for TAR folding. Both variants are expected to be fully replication-competent based on the following arguments. First, both sequences occur in natural HIV-1 isolates (~80% UCU and >15% UUU; Berkhout, 1992; Myers *et al.*, 1992). Second, no difference in TAR activity was measured for these two sequences in transient LTR-CAT assays (Dingwall *et al.*, 1989) and in HIV-1 infection experiments (Klaver and Berkhout, data not shown). Thus, outgrowth of the UUU variant is remarkable because there is no obvious selective advantage for this sequence change.

We propose a speculative mechanism for selection of the UUU bulge in transfection II based on the formation of alternative, transient RNA structures. We previously used a mutational approach in which TAR folding was inhibited by the transient formation of overlapping, alternative structures (Berkhout *et al.*, 1989). Metastable non-TAR structures can also be proposed for the IIa'1 variant and in particular for the next IIb10 variant with the additional C+1→G mutation (Figure 8B). Stabilization of this alternative hairpin as in variant IIb10 may be an unwanted side-effect of the C+1→G mutation. We propose that the C23→U change is favorable in this context because it strongly interferes with formation of the alternative hairpin (Figure 8B, IIc12), without affecting TAR folding and functioning.

Discussion

The HIV infectivity data presented here show a strong requirement for a base-paired lower part of the TAR RNA hairpin structure. Previous studies with LTR-CAT plasmids indicated that the lower part of the TAR stem is not essential for Tat-mediated *trans*-activation of the viral LTR promoter (Jakobovits *et al.*, 1988; Berkhout *et al.*, 1989). In fact, we routinely performed TAR mutational analyses on such a mini-TAR template (Berkhout and Jeang, 1991). The above observations suggest that TAR-mediated transcriptional induction of the LTR promoter is significantly different for unintegrated LTR-CAT plasmids compared with stably integrated HIV-1 proviral DNA. There are, however, other possible explanations of this finding. One should realize that LTR-CAT assays test only for the functions of TAR-mutated transcripts in transcription, transport, stability and/or translation. In virus replication experiments, however, it is possible that the +10 mutation is inhibitory to HIV at other stages, notably during genome packaging or reverse transcription. For instance, the TAR sequences, as part of the viral R (repeat) element, are pivotal in one of the initial steps of reverse transcription, in which the cDNA of the 5′ R region is translocated to the complementary 3′ R element. We note that our HIV-1 construct contained the Xho+10 substitution in both 5′ and 3′ LTRs, thereby avoiding such problems. Furthermore, transiently produced Xho+10 viruses were able to infect cells and to complete the reverse transcription reaction (B.Klaver and B.Berkhout, unpublished data). These results indicate that the blockade in Xho+10 replication is at the level of LTR transcription from an integrated, chromosomal position, suggesting that the requirement for the HIV TAR sequences in proviral transcription is stricter than in plasmid transcription. This is perhaps not surprising given the influence of chromatin structure on gene expression in general (McPherson *et al.*, 1993; reviewed by Felsenfeld, 1992; Adams and Workman, 1993), and that of integrated retroviruses in particular (Sorge *et al.*, 1984; Taketo *et al.*, 1985; Laimins *et al.*, 1986). Some recent studies reported other characteristics of HIV transcription that differed for integrated versus unintegrated LTR promoters (Jeang *et al.*, 1993a; Kim *et al.*, 1993). For instance, we found that the second coding exon of the Tat *trans*-activator protein, previously regarded as functionally dispensable based on LTR-CAT transfections, plays a role in optimal activation of integrated LTRs (Jeang *et al.*, 1993a).

The analysis of revertant viruses establishes the importance of HIV-1 transcripts to have a closed TAR RNA stem. As discussed above, we favor a transcriptional role for an extended TAR RNA stem, which might be instrumental in the optimal presentation of bulge-bound Tat and loop-bound cellular factors to the transcriptional machinery. We cannot exclude, however, some other mechanistic explanations. For instance, the importance of an extended TAR stem may also reflect the need for binding of a cellular protein to the lower stem region. This RNA−protein interaction, however, should not depend on sequence-specific contacts because the revertant TAR structures do not resemble the wild-type TAR sequence. One candidate stem-binding factor is the cellular protein kinase DAI, the double-stranded RNA-activated inhibitor of translation. This enzyme is induced by interferon and plays a key role in cellular defense against viruses. TAR RNA has been reported to interact with the DAI protein (SenGupta *et al.*, 1990; S.Roy *et al.*, 1991) and to inhibit its function (Gunnery *et al.*, 1990, 1992). Interestingly, a duplex structure of at least 14 bp was reported to be

necessary for DAI inactivation (Gunnery *et al.*, 1992), suggesting that our Xho+10 mutant is unable to block DAI, whereas all revertant structures, including the truncated type II TAR stems, will have at least partially restored this property. How TAR interacts with DAI in infected cells and its subsequent effects on viral replication await further study.

The Xho+10 mutation may also affect binding of a cellular transcription factor to LTR DNA. Multiple host transcription factors interact with the promoter-enhancer region of the LTR (U3 region), but some factors have been identified that bind DNA motifs in the transcribed DNA (R region). The cellular protein LBP-1 (Jones *et al.*, 1988) or UBP-1 (Garcia *et al.*, 1987; Wu *et al.*, 1988) recognizes direct repeat sequences in the region surrounding the transcription start site (Figure 1A, positions −4/+1, 7/11 and 17/22). The high affinity site at position 7/11, which is mutated in Xho+10, has been reported to act as a positive control element both *in vivo* (Jones *et al.*, 1988; Kato *et al.*, 1991; Moses *et al.*, 1994) and *in vitro* (Kato *et al.*, 1991). The other sites are occupied at high protein concentrations, causing inhibition of LTR activity. Overlapping the LBP-1 binding sites is the pyrimidine-rich initiator element (positions −5/+9; Smale *et al.*, 1990), which has been reported to bind the activator protein USF (Du *et al.*, 1993) and the related TFII-I factor (A.L.Roy *et al.*, 1991, 1993). Because the class I revertants obtained in this study replicated efficiently without restoration of the wild-type LTR sequence, it is obvious that HIV-1 replication in T cell lines is not critically dependent on the initiator element and the high-affinity LBP-1 site.

We described in detail two reversion pathways that repair a disrupted TAR RNA structure. In this tissue culture evolution system, mutations are generated either during transcription of the integrated provirus or during reverse transcription of the viral genome. Although the diversity will be random in nature, the population is shaped by selection of replication-competent viruses (Skinner *et al.*, 1989). Thus, the mechanism of mutation fixation is the positive selection of the fittest variants. Unlike *in vitro* evolution approaches (reviewed in Burke and Berzal-Herranz, 1993), where a pool of randomized sequences is available at the start of the experiment, our evolution protocol will strongly bias the advantageous mutations that happen to occur relatively early after transfection. Novel mutations augment existing variation, so that the evolutionary search is biased, in an appropriate fashion, by selection events that have already occurred. Thus, the viruses selected for in our evolution system represent only one class of revertants. Indeed, another reversion pathway was identified in a second, independent transfection. In general, we believe these two cases represent only a minor fraction of the potential routes to restore function to a disrupted TAR hairpin structure.

Although the TAR evolution experiments may not have lasted long enough to create the optimal TAR configuration, the data indicate that structural irregularities in the stem region are not detrimental to TAR function. Inspection of naturally existing TAR stems within the HIV−SIV virus family also indicates the presence of multiple bulges (Berkhout, 1992). This may indicate that the viral TAR RNA structure is a compromise between two opposing tendencies. On the one hand, folding and functioning of TAR may be best served by a strong, perfectly base-paired stem. The presence of a stable structure at the 5′ end of HIV genomic and messenger RNA may also confer protection against cellular RNases. On the other hand, extremely stable RNA structures may inhibit other critical functions of the viral RNA. As part of the leader of all HIV mRNA species, TAR should not interfere with the scanning movement of ribosomes (Kozak, 1983; Berkhout *et al.*, 1989). As part of the genomic RNA, TAR should not inhibit progress of the RT enzyme during strong-stop minus strand DNA synthesis (Pathak and Temin, 1992; Klaver and Berkhout, 1994). Another hypothetical role for a semi-stable TAR stem is in the next step of the reverse transcription reaction, where the cDNA of 5′ TAR is translocated to the complementary 3′ TAR element ('the first jump'; Gilboa *et al.*, 1979). Reverse transcription may be aborted at this step when the target 3′ TAR RNA is occluded in stable base-pairing.

Materials and methods

Cells, transfection and viral infection

The lymphocytic T cell line SupT1 was grown in RPMI 1640 medium containing 10% fetal calf serum. T cells were electroporated with HIV-1 DNA as described by Koken *et al.* (1992). Infections with the HIV-1 virus were performed at 37°C with SupT1 cells (~1 × 10^6 cells) in a small volume (1 ml) of RPMI medium. The volume was increased to 10 ml after 60 min.

In order to select for revertant viruses of the replication-incompetent Xho+10 mutant, transfected cells were maintained for several months. Cells were grown at 37°C and diluted 1 in 10 into fresh medium twice a week. Cell and supernatant samples were collected at each passage and stored at −70°C for future analysis. In a typical Xho+10 transfection, syncytia were not observed until the second month of maintenance. When HIV-induced cytopathic effects were seen, high level virus replication was maintained by passage of the cell-free culture supernatant onto an uninfected 10 ml SupT1 culture. We initially used 1 ml per passage to allow new syncytia to be formed within ~1 week. At later time points, we gradually used less culture supernatant per passage. For the fully replication-competent revertants present at months 4−6, we used only 0.1 μl to infect a fresh 10 ml SupT1 culture. Infections were monitored by measuring Gag p24 protein in the culture supernatant by ELISA within the linear range of the assay (Koken *et al.*, 1992).

Construction of mutant HIV-1 genomes

The wild-type, infectious HIV-1 plasmid pLAI was described previously (Peden *et al.*, 1991). The following strategy was used to introduce the Xho+10-mutated LTR of plasmid S+10 (Berkhout and Jeang, 1991) into both 5′ and 3′ LTRs of pLAI. First, the mutant TAR sequence was cloned as a *Bsp*EI−*Hin*dIII fragment (position −147 to +77 relative to the transcriptional start site) into the subclones Blue-5′ LTR and Blue-3′ LTR (Klaver and Berkhout, 1994). Second, the 5′ LTR of wild-type pLAI was replaced by the mutant LTR as an *Xba*I−*Cla*I fragment, thus generating the plasmid Xho+10/wt. Finally, the 3′ LTR of Xho+10/wt was replaced with the mutant *Xho*I−*Bgl*I fragment to produce the isogenic plasmid Xho+10/Xho+10, henceforth referred to as Xho+10.

Proviral DNA sequence analysis

The proviral LTR region was PCR-amplified using total cellular DNA and an Env-U5 primer pair [oligo *Bam*HI-Env and C(N1); Berkhout and Klaver, 1993] that specifically amplifies the 3′ LTR of integrated HIV-1 proviruses. The double-stranded DNA product was digested with *Bam*HI (within the *env* gene) and *Hin*dIII (position +77), cloned into Bluescript and sequenced on a Biosystems automated sequencer using the Dye Deoxy Terminator Cycle kit and the U3-region primer 5′ CE or the Bluescript primer KS (Berkhout and Klaver, 1993). This strategy allows for the analysis of the complete TAR region up to the downstream *Hin*dIII site. Upstream sequences were usually obtained up to position −220, but only sequences up to the TATAA box (−28) are represented in Tables I and II.

RNA secondary structure and primer extension analyses

RNA structure predictions and free energy calculations were performed using the MFOLD program in the University of Wisconsin Genetics Computer Group software package version 7.2 (Devereux *et al.*, 1984). The MFOLD program (Zuker, 1989) uses the energy rules as defined by Freier *et al.* (1986).

Primer extension reactions were performed to map the transcriptional initiation sites used by the several mutant HIV viruses. Total cellular RNA was isolated by the hot phenol method and analyzed by primer extension as described previously (Berkhout *et al.*, 1990) with some minor modifications. Instead of an end-labelled primer, we used 0.3 μl of [α-^{32}P]dCTP (10 mCi/ml) in the extension reaction. The C(N1) primer, complementary to the U5 region of the HIV-1 leader RNA (positions 122–150), was described previously (Berkhout and Klaver, 1993).

Acknowledgements

We thank Keith Peden for the generous gift of the pLAI plasmid, Atze Das for computer advice and Wim van Est for excellent artwork and photography. This research was supported by the Netherlands Organization for Scientific Research (NWO) and the Dutch Cancer Society (KWF).

References

Adams,C. and Workman,J.L. (1993) *Cell*, **72**, 305–308.
Bagasra,O., Khalili,K., Seshamma,T., Taylor,J.P. and Pomerantz,R.J. (1992) *J. Virol.*, **66**, 7522–7528.
Baudin,F., Marquet,R., Isel,C., Darlix,J.L., Ehresmann,B. and Ehresmann,C. (1993) *J. Mol. Biol.*, **229**, 382–397.
Benoist,C. and Chambon,P. (1981) *Nature*, **290**, 304–310.
Berkhout,B. (1992) *Nucleic Acids Res.*, **20**, 27–31.
Berkhout,B. and Jeang,K.-T. (1989) *J. Virol.*, **63**, 5501–5504.
Berkhout,B. and Jeang,K.-T. (1991) *Nucleic Acids Res.*, **19**, 6169–6176.
Berkhout,B. and Klaver,B. (1993) *Nucleic Acids Res.*, **21**, 5020–5024.
Berkhout,B., Silverman,R. and Jeang,K.-T. (1989) *Cell*, **59**, 273–282.
Berkhout,B., Gatignol,A., Silver,J. and Jeang,K.-T. (1990) *Nucleic Acids Res.*, **18**, 1839–1846.
Burke,J.M. and Berzal-Herranz,A. (1993) *FASEB J.*, **7**, 106–112.
Churcher,M.J., Lamont,C., Hamy,F., Dingwall,C., Green,S.M., Lowe,A.D., Butler,J.G., Gait,M.J. and Karn,J. (1993) *J. Mol. Biol.*, **230**, 90–110.
Cordingley,M.G., LaFemina,R.L., Callahan,P.L., Condra,J.H., Schlabach,A.J. and Colonno,R.J. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 8985–8989.
Cullen,B.R. (1992) *Microbiol. Rev.*, **56**, 375–394.
Desai,K., Loewenstein,P.M. and Green,M. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 8875–8879.
Devereux,J., Haeberli,P. and Smithies,O. (1984) *Nucleic Acids Res.*, **12**, 387–395.
Dimitrov,D.S., Willey,R.I., Sato,H., Chang,L.-J., Blumenthal,R. and Martin,M.A. (1993) *J. Virol.*, **67**, 2182–2190.
Dingwall,C., Ernberg,I., Gait,M.J., Green,S.M., Heaphy,S., Karn,J., Lowe,A.D., Singh,M., Skinner,M.A. and Valerio,R. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 6925–6929.
Du,H., Roy,A.L. and Roeder,R.G. (1993) *EMBO J.*, **12**, 501–511.
Felsenfeld,G. (1992) *Nature*, **355**, 219–224.
Feng,S. and Holland,E.C. (1988) *Nature*, **334**, 165–167.
Frankel,A.D., Biancalana,S. and Hudson,D. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 7397–7401.
Freier,S.M., Kierzek,R., Jaeger,J.A., Sugimoto,N., Caruthers,M.H., Nielson,T. and Turner,D.H. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 9373–9377.
Garcia,J.A., Wu,F.K., Mitsuyasu,R. and Gaynor,R.B. (1987) *EMBO J.*, **6**, 3761–3770.
Garcia,J.A., Harrich,D., Soultanakis,E., Wu,F., Mitsuyasu,R. and Gaynor,R.B. (1989) *EMBO J.*, **8**, 765–778.
Gatignol,A., Buckler-White,A., Berkhout,B. and Jeang,K.-T. (1991) *Science*, **251**, 1597–1600.
Gilboa,E., Mitra,S.W., Goff,S.P. and Baltimore,D. (1979) *Cell*, **18**, 93–100.
Gluzman,Y., Sambrook,J.F. and Frisque,R.J. (1980) *Proc. Natl Acad. Sci. USA*, **77**, 3898–3902.
Gunnery,S., Rice,A.P., Robertson,H.D. and Mathews,M.B. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 8687–8691.
Gunnery,S., Green,S.R. and Mathews,M.B. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 11556–11561.
Harrich,D., Garcia,J., Mitsuyasu,R. and Gaynor,R. (1990) *EMBO J.*, **9**, 4417–4423.
Hauber,J. and Cullen,B.R. (1988) *J. Virol.*, **62**, 673–679.
Jakobovits,A., Smith,D.H., Jakobovits,E.B. and Capon,D.J. (1988) *Mol. Cell. Biol.*, **8**, 2555–2561.
Jeang,K.-T., Berkhout,B. and Dropulic,B. (1993a) *J. Biol. Chem.*, **268**, 24940–24949.
Jeang,K.-T., Chun,R., Lin,N.H., Gatignol,A., Glabe,C.G. and Fan,H. (1993b) *J. Virol.*, **67**, 6224–6233.
Jones,K.A., Luciw,P.A. and Duchange,N. (1988) *Genes Dev.*, **2**, 1101–1114.
Kato,H., Horikoshi,M. and Roeder,R.G. (1991) *Science*, **251**, 1476–1479.
Kim,J.Y., Gonzalez-Scarano,F., Zeichner,S.L. and Alwine,J.C. (1993) *J. Virol.*, **67**, 1658–1662.
Klaver,B. and Berkhout,B. (1994) *Nucleic Acids Res.*, **22**, 137–144.
Koken,S.E.C., van Wamel,J.L.B., Goudsmit,J., Berkhout,B. and Geelen J.L.M.C. (1992) *Virology*, **191**, 968–972.
Koken,S.E.C., van Wamel,J.L.B., Geelen,J.L.M.C. and Berkhout,B. (1994) *J. Biomed. Sci.*, **1**, 83–92.
Kozak,M. (1983) *Cell*, **34**, 971–978.
Laimins,L., Holmgren-Konig,M. and Khoury,G. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 3151–3155.
Leonard,J., Parrott,C., Buckler-White,A.J., Turner,W., Ross,E.K., Martin,M.A. and Rabson,A.R. (1989) *J. Virol.*, **63**, 4919–4924.
Mann,D.A. and Frankel,A.D. (1991) *EMBO J.*, **10**, 1733–1739.
Marciniak,R.A., Garcia-Blanco,M.A. and Sharp,P.A. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 3624–3628.
Masuda,T. and Harada,S. (1993) *Virology*, **192**, 696–700.
McPherson,C.E., Shim,E.-Y., Friedman,D.S. and Zaret,K.S. (1993) *Cell*, **75**, 387–398.
Moses,A.V., Ibanez,C., Gaynor,R.B., Ghazal,P. and Nelson,J.A. (1994) *J. Virol.*, **68**, 298–307.
Muesing,M.A., Smith,D.H. and Capon,D.J. (1987) *Cell*, **48**, 691–701.
Myers,G., Berzofsky,J.A., Pavlakis,G.N., Korber,B. and Smith,R.F. (eds) (1992) *Human Retroviruses and AIDS*. Los Alamos National Laboratory, Los Alamos, New Mexico, USA.
Nelbock,P., Dillon,P.J., Perkin,A. and Rosen,C.A. (1990) *Science*, **248**, 1650–1653.
Ohana,B., Moore,P.A., Ruben,S.M., Southgate,C.D., Green,M.R. and Rosen,C.A. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 138–142.
Pathak,V.K. and Temin,H.M. (1992) *J. Virol.*, **66**, 3093–3100.
Peden,K., Emerman,M. and Montagnier,L. (1991) *Virology*, **185**, 661–672.
Pilsinelli,G.A. and Temin,H.M. (1991) *J. Virol.*, **65**, 4786–4797.
Rounseville,M.P. and Kumar,A. (1992) *J. Virol.*, **66**, 1688–1694.
Roy,A.L., Meisterernst,M., Pognonec,P. and Roeder,R.G. (1991) *Nature*, **354**, 245–248.
Roy,A.L., Malik,S., Meisterernst,M. and Roeder,R.G. (1993) *Nature*, **365**, 355–359.
Roy,S., Delling,U., Chen,C.H., Rosen,C.A. and Sonenberg,N. (1990) *Genes Dev.*, **4**, 1365–1373.
Roy,S., Agy,M., Hovanessian,A.G., Sonenberg,N. and Katze,M.G. (1991) *J. Virol.*, **65**, 632–640.
Selby,M.J., Bain,E.S., Luciw,P.A. and Peterlin,B.M. (1989) *Genes Dev.*, **3**, 547–558.
SenGupta,D.N., Berkhout,B., Gatignol,A., Zhou,A. and Silverman,R.H. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 7492–7496.
Sheline,C.T., Miloco,L.H. and Jones,K.A. (1991) *Genes Dev.*, **5**, 2508–2520.
Skinner,M.A., Racaniello,V.R., Dunn,G., Cooper,J., Minor,P.D. and Alond,J.W. (1989) *J. Mol. Biol.*, **207**, 379–392.
Smale,S.T., Schmidt,M.C., Berk,A.J. and Baltimore,D. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 4509–4513.
Sorge,J., Cutting,A.E., Erdman,V.D. and Gautsch,J.W. (1984) *Proc. Natl Acad. Sci. USA*, **81**, 6627–6631.
Taketo,M., Gilboa,E. and Sherman,M.I. (1985) *Proc. Natl Acad. Sci. USA*, **82**, 2422–2426.
Taylor,J.P., Pomerantz,R., Bagasra,O., Chowdhury,M., Rappaport,J., Khalili,K. and Amini,S. (1992) *EMBO J.*, **11**, 3395–3403.
Vaishav,Y.N. and Wong-Staal,F. (1991) *Annu. Rev. Biochem.*, **60**, 577–630.
Weeks,K.M. and Crothers,D.M. (1991) *Cell*, **66**, 577–588.
Weeks,K.M., Ampe,C., Schultz,S.C., Steitz,T.A. and Crothers,D.M. (1990) *Science*, **249**, 1281–1285.
Wu,F.J., Garcia,D., Harrich,D. and Gaynor,R. (1988) *EMBO J.*, **7**, 2117–2129.
Wu,F.J., Garcia,D., Sigman,D. and Gaynor,R. (1991) *Genes Dev.*, **5**, 2128–2140.
Zuker,M. (1989) *Science*, **244**, 48–52.

Received on February 21, 1994; revised on March 24, 1994