Detailed mutational analysis of TAR RNA: critical spacing between the bulge and loop recognition domains

Ben Berkhout* and Kuan-Teh Jeang

Laboratory of Molecular Microbiology, National Institute of Allergy and Infectious Diseases,
National Institutes of Health, Bethesda, MD 20892, USA

Received September 3, 1991; Revised and Accepted October 21, 1991

ABSTRACT

Trans-activation of HIV-1 by the Tat protein is mediated through a cis-acting element (TAR) in the viral RNA. In order to obtain further insight into the molecular interactions for trans-activation, a detailed mutational analysis of TAR RNA was carried out. TAR RNA forms a hairpin structure with important sequence elements in the single-stranded bulge- and loop-domains. We found that the sequence of the base-pairs flanking the bulge is critical for Tat-mediated trans-activation. In addition, Tat-response is reduced when the bulge is forced into a base-paired configuration through the introduction of complementary nucleotides on the opposite side of the stem. Thus, the 3-nucleotide bulge and adjacent base-pairs comprise a recognition domain with both sequence- and structure-elements. Accessibility of the loop sequences is also important for Tat function, since base-pairing through the formation of a pseudoknot-like structure does inhibit Tat action. A third critical parameter that influences the magnitude of Tat response is the number of loop nucleotides. Finally, the relative spacing between the loop and the bulge is also important. We introduced additional base-pairs in the stem connecting the two domains. Such mutations progressively decreased the efficiency of Tat induction. Interestingly, activity of the HIV-2 Tat protein did markedly increase on targets with one or two additional basepairs. These results suggest that Tat interacts with a cellular loop-binding protein(s) to increase HIV gene expression.

INTRODUCTION

The human and simian immunodeficiency viruses (HIV-SIV) encode a regulatory protein Tat. The role of Tat is to activate gene expression from the long terminal repeat (LTR), thereby resulting in high levels of RNA synthesis (reviewed in 1,2). Both Tat and TAR (the cis-acting Tat responsive element located in the LTR, ref. 3) are essential for HIV replication (4–7). Tat has been proposed to enhance the rate of transcription through

a direct interaction with TAR RNA (8–13). How Tat, bound to the nascent transcript, modifies the transcription machinery remains unclear. Further elucidation of the molecular mechanism of Tat action should be aided by the identification of cellular proteins that functionally interact with Tat and/or TAR. Recently, one candidate Tat-binding protein (14) and multiple TAR RNA-binding proteins (15–18) have been described. The development of a cell-free transcription system (19), dependent on exogenously added Tat, is expected to significantly help in the identification and study of such co-factors.

Extensive mutational analyses of the 86 amino acid Tat protein has identified at least two functional domains. A highly conserved motif containing six essential cysteine residues has been proposed to bind metal ions and may facilitate protein-protein interactions (20–24). Another essential region is rich in lysine and arginine residues (20,22,23,25–29). This basic domain is required for the nucleolar localization of Tat. Furthermore, this same stretch of nine amino acids was reported to be necessary and sufficient for binding to TAR RNA in vitro (8–13).

TAR is an approximately 23 nucleotide long sequence present at the 5' end of all viral transcripts. TAR forms the upper part of an RNA hairpin structure with a characteristic 3-nucleotide bulge structure and a 6-nucleotide loop (Figure 2A, ref. 3,30). Mutational analyses suggest that the base-paired stem in TAR is absolutely required, although the actual sequence of the lower part of the stem is not critical (30–34). Essential nucleotide sequences are present in the single-stranded loop- and bulge-domains of the hairpin (34–37). Trans-activation assays in transiently transfected tissue culture cells indicate that single point mutations in either one of these domains can severely inhibit Tat function. In addition, both structure elements are conserved between HIV-1 and HIV-2 TAR (38,39). The phenotype of bulge mutants can easily be explained, since Tat interacts directly with this region of the TAR hairpin (9). At present, however, we do not understand the phenotype of loop mutants. Loop mutations do not interfere with Tat binding in vitro. It is therefore likely that a cellular loop-binding protein is intimately involved in Tat trans-activation. One candidate loop-binding protein was recently identified in HeLa nuclear extracts (17). Binding of this 68 kDa protein to TAR is sensitive to loop mutations. Furthermore, this

* Present address: Viral Regulation Laboratory, Department of Virology, University of Amsterdam, Meibergdreef 15, 1105 AZ Amsterdam, The Netherlands

partially purified protein was reported to stimulate Tat-mediated transcription *in vitro* (19).

In order to gain further insight into the molecular mechanism of Tat action, we performed a detailed mutational analysis of TAR. We systematically changed the position of TAR within the transcript, the size of the loop, the sequence of all basepairs, and the spacing between bulge and loop domains. In addition, we tested whether accessibility of the single-stranded bulge and loop is critical for function.

MATERIALS AND METHODS

Plasmid Constructions

pLTR-CAT is as previously described (30). It contains the complete HIV-1 LTR U3 region upstream, and the cat reporter gene downstream of TAR. XhoI restriction sites were introduced flanking the transcriptional initiation site in plasmids S-10 and S+10 [construct pLTR(S-10)-CAT in ref. 39 and 5' in ref. 30, respectively]. Recombination between these two XhoI sites was used to generate plasmid 5'd19. Mutants 5'd23 and 5'd24 were made by Klenow treatment of the XhoI sticky ends of 5'd19. Deletion mutant 5'd-12 was previously described (designated as 5'd in ref. 30). The complete nucleotide sequence of the inserts of the 5'i-series is provided in Figure 1B. Plasmid 5'i-63 contains three copies of the 21-nucleotide fragment shown in Figure 1B (upper sequence). TAR mutants described in Figure 2 and Table 1 were generated by cloning synthetic double-stranded DNA oligomers into the mini-TAR vector (Figure 2A) between the unique restriction sites XhoI and HindIII flanking the TAR structure. All mutations were verified by sequencing. Tat-1 and Tat-2 expression vectors were described previously (30,39).

A

wild-type	100	Bulge-G	17
BP-1	83	Bulge-G/BP-1	18
BP-2	70	Bulge-G/BP-2	15
BP-3	16	Bulge-G/BP-3	10
BP-4	12	Bulge-G/BP-4	6

B

	Tat-1	Tat-2
wild-type	100	100
Loop-CA	24	53
Loop-CG	88	141
Bulge-G	17	26
Loop-CA/Bulge-G	6	15
Loop-CG/Bulge-G	15	36

Table 1. Relative trans-activation of double TAR mutants. **A)** Activities of the BP-1,2,3,4 mutants in the context of either wild-type or Bulge-G mutant TAR. Tat-1 induction of wild-type TAR was standardized at 100%. **B)** Activities of the Loop-CA and Loop-CG insertion mutants in the context of wild-type and Bulge-G TAR. Experiments were performed with both Tat-1 and Tat-2 as trans-activator protein. For each Tat protein, the relative activity on wild-type TAR was set at 100%.

Cell Culture, Transfection, and RNA and Protein Analysis

COS and HeLa cells were maintained in Dulbecco's minimal medium containing 10% fetal bovine serum. Transfections were carried out with DEAE-dextran (400 μ g/ml for COS, 100 μ g/ml for HeLa) in the presence of chloroquine and NU-serum (40). For CAT assays, HeLa cells (40mm sub-confluent dish) were transfected with 0.5 μ g of the test plasmids in the presence of 0.2 μ g Tat expression plasmid (pSVTat). As an internal control, a β -galactosidase expression vector was included in some of the experiments. For RNA analysis, COS cells (100mm sub-confluent dish) were transfected with 8 μ g LTR-CAT construct and 1 μ g pSVTat. RNA purification and primer extension analysis were performed as described (40). CAT assays were done with butyryl-CoA in combination with thin layer chromatography (41) or with the phase-extraction method (42).

RESULTS

The position of TAR within the transcript is important for trans-activation

Figure 1A shows the partial nucleotide sequence of the HIV-1 LTR downstream of the TATA box. Sequences from position +19 to +42 relative to the transcriptional start site (+1) constitute the minimal TAR element (underlined in Figure). LTR sequences up to position +77 were cloned upstream of the cat reporter gene to produce plasmid pLTR-CAT. This wild-type construct was transfected into HeLa cells either in the presence or in the absence of a second plasmid encoding Tat. As measured by the induction of CAT enzyme (Figure 1, right panel), pLTR-CAT is dramatically activated by Tat. We questioned whether the position of TAR relative to the TATA promoter was important for this activity.

To move TAR closer to the TATA box, we deleted 12, 19, 23 or 24 nucleotides in the -10 region between the two elements (see 5'd12, 5'd19, 5'd23, 5'd24; Figure 1A). In 5'd12, TAR is positioned 6 nucleotides downstream of the mRNA cap site. This construct remained fully Tat-responsive. In contrast, a further removal of 7 nucleotides as in 5'd19 resulted in a dramatic loss in Tat trans-activation. The same result was obtained with the 5'd23 and 5'd24 constructions. Thus, although TAR can be moved very close to the mRNA cap site, it abruptly loses activity once the TAR sequence overlaps the +1 position. The latter result is predicted for a sequence that is active in its RNA, as opposed to DNA, form. Alternatively, loss of Tat responsiveness with 5'd19 can be the result of the progressive removal of sequences around position -10 in this construct. However, a control construct with random nucleotide substitutions in this region (plasmid S-10, Figure 1A) responded to Tat in a wild-type manner.

Previous results from Selby and Peterlin showed that TAR rapidly loses activity when its position is moved further downstream in the transcript (33). Because some essential features of Tat-trans-activation are different in their experimental system compared to ours (discussed in 43), we tested further the importance of spacing in our pLTR-CAT plasmid. To facilitate the introduction of sequences between TAR and the promoter, an unique XhoI restriction site was introduced at position +11 (plasmid S+10, Figure 1B). Synthetic DNA of randomly chosen sequence was inserted into the XhoI site. In the absence of Tat, the promoter activity of all mutant constructs was comparable

to that of the parental plasmid S+10 (data not shown). However, all insertion mutants were reduced in Tat responsiveness. Insertion of 63 nucleotides affected trans-activation most severely. Construct 5'i-63 retained 14% of the wild-type activity. Plasmids with smaller insertions showed intermediate effects on trans-activation. Our results indicate that small insertions-deletions upstream of TAR do not completely eliminate TAR activity. However, primarily because of its natural context, the positional requirements of TAR is more sensitive to upstream deletions as opposed to insertions.

Size and conformation of the loop is critical for TAR function

The minimal TAR has been mapped to be between nucleotides +19 and +42 in the HIV-1 LTR (Figure 1). This sequence forms a stem-bulge-loop structure (Figure 2A: wild-type TAR). To test only for the contribution of the nucleotides between +19 and +42, we made a construction that preserved all of the important sequence- and structure-elements (construct S+10; Figure 2A: model TAR). This mini-TAR structure was used in further mutational analyses. All of the mutations generated in S+10 were tested for both CAT enzyme activity (Figure 2) and RNA levels (Figure 3) in the presence of HIV-1 or HIV-2 Tat. The latter protein cross-trans-activates the HIV-1 LTR, albeit it at a lower efficiency.

Several studies have indicated the sequence-specific importance of the loop nucleotides CUGGGA. We asked whether the size of the loop was also important for function. Two mutants were made, each creating one additional loop nucleotide, either on the 5' side in Loop-C (Figure 2B: CCUGGGA) or on the 3' side in Loop-A (CUGGGAA). Both loop mutants showed reduced responsiveness to both Tat-1 and Tat-2 (Figure 2B). Combination of the two mutations into Loop-CA (CCUGGGAA) created a loop of 8 nucleotides which was not further reduced in the magnitude of Tat induction. In fact, Loop-CA was a better target for Tat-2 than the two single nucleotide mutants. To explain this result, we assume that the spatial symmetry of the loop relative to the stem is important for an efficient Tat response. This symmetry is restored in Loop-CA as compared to Loop-C or Loop-A. We analyzed one additional mutant with a two nucleotide insertion (Loop-CG: CCUGGGAAG). This mutant is comparable to Loop-CA, but basepairing between the inserted nucleotides is predicted to recreate the wild-type TAR loop. Indeed, induction of Loop-CG by Tat-1 reached wild-type levels. Interestingly, induction of Loop-CG by the HIV-2 Tat protein was consistently found to be more efficient than induction of wild-type TAR (see later). These results suggest that the spatial conformation of the TAR loop is important for function.

We next tested whether the TAR loop needs to be sterically accessible for hypothetical interactions with cellular proteins. We

Figure 1. Tat-mediated trans-activation of TAR spacing mutants. **A)** The nucleotide sequence of HIV-1 between TATA (indicated above sequence) and TAR (underlined) are shown. Sequences upstream of the transcriptional start site (+1) are represented by T, downstream sequences by U. Deletions in the 5'd series move TAR closer to the TATA box. Control plasmid S-10 contains a 16-nucleotide substitution (underlined). Responsiveness of these plasmids to Tat is shown in the right-hand columns. All HeLa cell transfections were performed in the presence of Tat, except for the one shown on top, which represents pLTR-CAT in the absence of Tat. Relative CAT enzyme activities (cpm) and efficiency of Tat induction (pLTR-CAT = 100%) are shown. CAT activity was quantitated using a modified CAT assay (55) and was analyzed by phase-extraction and by TLC. The positions of chloramphenicol (Cm) and butyrylated chloramphenicol (Bu-Cm) are indicated. **B)** Sequences between position +1 and TAR (underlined) are shown. Insertions of 10, 16, 21 nucleotides [two sequences for each insertion, indicated by A (in superscript) and B] and 63 nucleotides are indicated. Representative CAT TLC assays with S+10, 5'i-10A, 5'i-16B, 5'i-21A and 5'i-63 are shown.

Figure 2. Primary sequences and RNA secondary structures of wild-type and mutant TAR constructs. Nucleotide insertions are indicated in bold. Relative trans-activation efficiencies with both the HIV-1 and HIV-2 Tat proteins were measured in transfected HeLa cells. All DNA transfections were performed on at least five separate occasions. The results shown represent average values.

introduced sequences complementary to the loop (UCCCA) downstream of the mini-TAR stem in TAR-PK (Figure 4A). Basepairing between these two stretches can result in the formation of a so-called pseudoknot RNA structure (44) which would inhibit loop recognition and TAR function. Either wild-type TAR or pseudoknot TAR (TAR-PK) was co-transfected with increasing amounts of a Tat-1 producing plasmid into HeLa cells. CAT enzyme activity was assayed after two days (Figure 4A,

right-hand panel). We found that Tat trans-activation was only marginally affected by the insertion of the loop-antisense sequences in TAR-PK. One explanation is that the expected tertiary RNA structure did not efficiently form. For instance, one critical parameter in pseudoknot formation is the size of the loop, and the six-nucleotide TAR loop might be too small for this type of basepairing. To circumvent this problem, we generated another set of complementary mutants: TAR-AC and

Figure 3. Primer extension analysis of cat mRNA synthesized in COS cells transfected with the indicated plasmids. The wild-type TAR construct was tested both in the absence (–) and presence (+) of Tat (lanes 3 and 4, respectively), all other transfections were performed in the presence of Tat. Loop-135 (lane 5) and Bulge-123 (lane 6) are 3-base substitution mutants (loop CUGGGA to AUUGUA and bulge UCU to AAG, respectively). BP4 is in lane 7, BP+1 in lane 8, BP+2 in lane 9, BP+4 in lane 10, BP+6 in lane 11, Bulge-GGA in lane 12, Bulge-G in lane 13 and Loop-CA in lane 14. A 31-nucleotide cat-specific primer was used (primer alone in lane 2). The position of the different cDNA products are indicated. Base insertions in some of the TAR mutants are appropriately reflected in changes in cDNA lengths. End-labeled pBR322-HpaII fragments were used as molecular size markers in lanes 1 and 15.

TAR-AC-PK (Figure 4B). TAR-AC contains an eight-nucleotide loop and should adopt a pseudoknot structure more easily than wild-type TAR. A sequence complementary to GUCCAG was inserted downstream of the TAR-AC stem to create TAR-AC-PK. These paired constructs did show a 2- to 3-fold difference in Tat responsiveness over a wide range of Tat protein (Figure 4B). We should point out that, due to the additional loop-nucleotides in TAR-AC compared to TAR, the overall magnitude of Tat induction is approximately 6-fold lower (Figure 4, compare panel B to A). These results, together with those presented in Figure 2B, indicate that the TAR loop participates in trans-activation by Tat not only as a primary sequence motif, but also as an appropriately configured single-stranded loop at the apex of the TAR RNA hairpin.

Figure 4. Inhibition of TAR function through pseudoknot formation. Schematic secondary/tertiary RNA structures are shown. The loop-complementary sequences introduced in TAR-PK and TAR-AC-PK are underlined. In order to measure trans-activation efficiencies, a constant amount of the test plasmids was co-transfected into HeLa cells with increasing amounts of a Tat-1 expression vector. TAR is ■ and TAR-PK is □ in panel A, TAR-AC is ● and TAR-AC-PK is ○ in panel B.

Structure requirements for the trinucleotide bulge element

It has been proposed that Tat interacts with TAR RNA at the three-nucleotide bulge. We tested the importance of accessibility to this single-stranded UCU sequence. Mutant TAR plasmids were made with bulge-complementary sequences inserted on the opposite strand of the TAR stem (Figure 2C: Bulge-G and Bulge-GGA). We found that both structural changes interfered with Tat-mediated gene expression (Figure 2C and Figure 3: compare Bulge-GGA, lane 13 and Bulge-G, lane 14 to wild-type TAR, lane 4). Bulge-G compared to Bulge-GGA was slightly less efficient in blocking Tat action. It is likely that the two-nucleotide bulge that is left in mutant TAR-G still retained activity. We constructed a third mutant with an altered bulge structure: Bulge-UCU (Figure 2C). In this plasmid, the bulge conformation is changed into an internal loop. We found this mutant to be severely defective in the Tat response. These results suggest that very precise structural constraints are imposed on the bulge recognition domain, and that single strandedness per se of UCU (see Bulge-UCU) is not wholly sufficient for trans-activation.

Spacing between bulge and loop and the importance of sequence-specificity of the interposed basepairs

The preceding experiments reinforce the concept of two important recognition domains in TAR, the bulge for binding of Tat and the loop for binding of critical cellular components. Although either protein may interact well with TAR RNA independently, it is likely that cooperativity in binding exists. It is also probable that a functional interaction between both proteins occurs during Tat trans-activation. If these assumptions are correct, then the spacing between loop and bulge should be critical for function. To test this, we initially moved the bulge up or down the basepaired stem (Figure 2D). In our nomenclature, basepairs are numbered starting at the loop. Thus, wild-type TAR is termed

4B7 (4 basepairs-Bulge-7 basepairs). The bulge was either moved one basepair-position closer to the loop (3B8), or one or two 'basepairs' down the stem structure (5B6 and 6B5). Interestingly, moving the bulge only one basepair up or down the TAR stem did markedly inhibit Tat function. More surprisingly, shifting the bulge to two basepairs below its normal position (6B5) restored activity. A close examination of the actual sequences of 6B5 provides a possible explanation for these perplexing results. In 6B5, unlike the other mutants, the basepairs flanking the bulge are identical to those in wild-type TAR. Thus, these results strongly suggest that the actual sequence of the basepairs flanking the bulge are important for full activity.

To verify this, we constructed a series of seven TAR mutants with individual basepair changes (BP1 to 7, Figure 2E). The actual sequence of the nucleotides was reversed in a systematic manner without disrupting basepair interactions (e.g. C-G to G-C or A-U to U-A). Each test plasmid was introduced with either a Tat-1 or Tat-2 plasmid into HeLa cells. We found that all basepair-changes negatively affect the Tat response. The most dramatic relative reduction was seen for mutants BP3, 4 and 5. Primary sequence changes here disproportionately interfere with Tat activity. Thus, besides having a structural role, important sequence-specific information is contained in the basepairs flanking the bulge. We note that the trans-activation defects seen with these BP mutants are at least as severe as the defects seen with loop and bulge mutants (Figure 3, compare BP4; lane 7, to mutant Loop-135; lane 5, and Bulge-123; lane 6).

One implication of these results is that the UCU-bulge cannot independently function between just any set of basepairs. In a different approach, one or multiple basepairs were inserted into the stem region between bulge and loop segments (Figure 2F: BP+1, +2, +4, +6) in order to preserve the immediate sequences flanking the bulge. Thus, we moved the loop and bulge apart, without affecting the basepairs flanking these domains. When these plasmids were tested in cells, we found a progressive decrease in Tat-mediated transcriptional induction (Figure 2F). It is possible that the extension of the mini-TAR stem with six basepairs does create a structure in the leader transcript that interferes with mRNA translation. In order to exclude this possibility, RNA levels were quantitatively measured for all these constructs (Figure 3). These results indicate that optimal spacing between bulge and loop is important for an efficient transcriptional induction.

In contrast to Tat-1, Tat-2 trans-activation actually increased upon the insertion of one or two additional basepairs (Figure 2F, BP+1 and BP+2 and Figure 2B, Loop-CG). However, like Tat-1, a sharp decline in Tat-2 response was seen with mutants BP+4 and BP+6. This small difference in specificity between the two trans-activator proteins is striking when one considers the fact that all other TAR mutants, tested in this and other studies (37,38), qualitatively responded in a consistent fashion to either Tat-1 or Tat-2.

To further substantiate this finding we tested the same set of TAR mutants using different forms of Tat-1 or Tat-2 proteins. A full-length Tat-1 was compared to a truncated exon-1 form of Tat-1 (Tat-1 Δ C) and to a Tat protein in which the basic domain was replaced with that from the HIV-1 Rev protein (Tat-Rev, 28). The wild-type Tat-2 protein was compared to two shortened polypeptides (39): Tat-2 Δ C (a deletion of the complete exon-2), and Tat-2 Δ N (an internal deletion over amino acids 8–33). Trans-activation assays showed that all modified forms of Tat-1 and Tat-2 proteins retained activity. We next measured

Figure 5. Trans-activation of TAR spacing mutants by different forms of Tat-1 and Tat-2. Equal amounts of the wild-type TAR plasmid and the BP+1, +2, +4, +6 mutants were individually co-transfected into HeLa cells with plasmids encoding the different Tat proteins. For each Tat protein, the relative activity on wild-type TAR was set at 100%.

their relative activity on the BP+ mutant series (Figure 5). The data show that all Tat-1 forms progressively lose activity with an increase in the number of basepairs between the bulge and loop. Tat-2 polypeptides, on the other hand, seem to favor a target with one or two additional basepairs (i.e. BP+1, BP+2). These results also indicate that the observed difference in specificity of Tat-2 cannot be attributed to either the N-terminal or C-terminal portion of this protein.

Double mutations in TAR produce a more severe phenotype

We previously reported that a triple mutation in the bulge (B123, Figure 3, lane 6) affected trans-activation more dramatically than the individual point mutations (37). Analogous results were seen for the loop domain, where the triple mutant L135 (Figure 3, lane 5) combines the inhibitory effects of three individual point mutants (data not shown). We next tested whether a similar result can be seen upon combination of mutations in the bulge and the stem basepairs (Table 1A). The activities measured with the BP1,2,3,4 mutants were assayed both in the context of wild-type TAR (left panel) and Bulge-G TAR mutant (right panel). Indeed, the inhibitory effects of bulge and basepair changes become more pronounced when combined. Another set of double TAR mutants was constructed with Bulge-G and either the Loop-CA or Loop-CG mutant (Table 1B). Activity of the bulge-loop mutants was assayed with either Tat-1 or Tat-2. The data in Table 1 clearly demonstrate that TAR function becomes increasingly hampered in loop-bulge double mutants.

DISCUSSION

This study, designed to analyze the sequence and structure requirements for TAR function, suggests that responsiveness to Tat requires multiple motifs within this 23-nucleotide RNA hairpin element. Base changes within the single-stranded bulge and loop domains were previously shown to reduce Tat-mediated gene induction. In this mutational analysis, additional recognition nucleotides have been identified in the basepairs adjacent to the bulge. For instance, reversion of the basepair on top of the bulge (BP4; G-C to C-G) reduced TAR activity by 10-fold.

In vitro binding experiments have indicated that Tat recognition of TAR requires the presence of a U-rich bulge (8–13) in the context of a stem with the appropriate basepair-sequences (12,45). The multiplicity of important nucleotides in the bulge-duplex domain of TAR argues that Tat looks at the entire surface of this part of TAR RNA. Given the behavior of TAR double mutants, we assume that each of the essential nucleotides contributes to Tat binding independently. This sequence-specific contribution may exist in two forms. First, base-specific contacts may exist between these nucleotides and the Tat protein. Second, important structural properties may be imposed on TAR RNA by specific nucleotide composition. The Tat protein may specifically recognize the three-dimensional RNA shape, for instance through sensing of a special arrangement of the phosphate backbone at the bulge-stem junction (13). Despite the widespread occurrence of bulge elements in many natural RNA molecules, relatively little is known about their thermodynamic parameters (46). The observation that helix sequence is important only at basepairs adjacent to the bulge, may well indicate that these positions are more accessible to Tat due to a distortion of the regular helix shape (45).

Another TAR sequence required for maximal trans-activation is the six-nucleotide loop (34,36,37). Our mutational data support this notion. Base insertions that increase the loop size by one inhibit TAR activity. In addition, distortion of the loop through basepairing into a pseudoknot conformation interferes with trans-activation. These combined results suggest that accessibility and spatial presentation of the loop sequences are important for function. Since the loop sequences are not directly involved in Tat binding (9), a simple explanation for the phenotype of loop mutants is that a cellular co-factor (LBP, loop binding protein) specifically recognizes this domain.

Some RNA binding proteins bind RNA directly, while others require the presence of another protein. For instance, such a tripartite interaction has been proposed for the signal recognition particle (47) and the U2 small nuclear ribonucleoprotein complex (48,49). The presence of a second protein binding site on TAR is therefore of considerable interest. One experimental approach, insertion of additional basepairs in between the two protein binding sites, indicates that the spacing between TAR-bound Tat and the putative LBP protein is critical for trans-activation. One explanation is that LBP induces a structural change in TAR such that the Tat binding site becomes fully accessible. In this model, the structural rearrangement becomes less efficient with increasing distance. The main advantage of such a tri-partite mechanism would be to ensure a high degree in binding specificity, and this may in fact explain the modest 10-fold specificity of Tat binding measured *in vitro* in the absence of LBP protein. An alternative explanation is that Tat directly interacts with LBP. Such a protein-protein interaction is also expected to add stability to the TAR-Tat-LBP complex (50). We favor this latter model because Tat-2, unlike Tat-1, acts with increasing efficiency on TAR targets with one or two additional basepairs. This difference in specificity between the two trans-activator proteins suggests that Tat-2 is able to contact LBP over a longer distance. This property of Tat-2 is reflected in HIV-2 TAR, which contains five instead of four basepairs between loop and bulge (38,39). More detailed mutational analyses of Tat are needed to identify the amino acids that confer this spacing-specificity and the protein motif that is required for the interaction with LBP. A candidate motif for the proposed protein-protein interaction is the cysteine-rich domain present in all Tat proteins.

ACKNOWLEDGEMENTS

We thank Cees Pleij for advise on RNA pseudoknots, Dr. G. Chinnadurai for the generous gift of Tat-Rev plasmid and Alicia Buckler-White for oligonucleotide synthesis. This project was supported by the intramural AIDS targeted anti-viral program from the office of the director of the National Institutes of Health. This work is dedicated to the memory of Peter H. van Knippenberg.

REFERENCES

1. Sharp, P.A. and Marciniak, R.A. (1989) *Cell* 59, 229–230.
2. Cullen, B.R. (1990) *Cell* 63, 655–657.
3. Muesing, M.A., Smith, D.H. and Capon, D.J. (1987) *Cell* 48, 691–701.
4. Dayton, A.I., Sodroski, J.G., Rosen, C.A., Goh, W.C. and Haseltine, W.A. (1986) *Cell* 44, 941–947.
5. Fisher, A.G., Feinberg, S.F., Josephs, S.F., Harper, M.E., Marselle, L.M., Reyes, G., Gonda, M.A., Aldovini, A., Debouk, C., Gallo, R.C. and Wong-Staal, F. (1986) *Nature* 320, 367–371.
6. 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.
7. Harrich, D., Garcia, J., Mitsuyasu, R. and Gaynor, R. (1990) *EMBO J.* 9, 4417–4423.
8. 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.
9. Dingwall, C., Ernberg, I., Gait, M.J., Green, S.M., Heaphy, S., Karn, J., Lowe, A.D., Singh, M. and Skinner, M.A. (1990) *EMBO J.* 9, 4155–4153.
10. Roy, S., Delling, U., Chen, C.H., Rosen, C.A. and Sonenberg, N. (1990) *Genes Dev.* 4, 1365–1373.
11. Cordingley, M.G., LaFemina, R.L., Callahan, P.L., Condra, J.H., Sardana, V.V., Graham, D.J., Nguyen, T.M., LeGrow, K., Gotlib, L., Schlabach, A.J. and Colonna, R.J. (1990) *Proc. Natl. Acad. Sci. USA* 87, 8985–8989.
12. Weeks, K.M., Ampe, C., Schultz, S.C., Steitz, T.A. and Crothers, D.M. (1990) *Science* 249, 1281–1285.
13. Calnan, B.J., Biancalana, S., Hudson, D. and Frankel, A.D. (1991) *Genes Dev.* 5, 201–210.
14. Calnan, B.J., Tidor, B., Biancalana, S., Hudson, D. and Frankel, A.D. (1991) *Science* 252, 1167–1171.
15. Nelbock, P., Dillon, P.J., Perkins, A. and Rosen, C.A. (1990) *Science* 248, 1650–1653.
16. Gaynor, R., Soultanakis, E., Kuwabara, M., Garcia, J. and Sigman, D.S. (1989) *Proc. Natl. Acad. Sci. USA* 86, 4858–4862.
17. Gatignol, A., Kumar, A., Rabson, A.R. and Jeang, K.-T. (1989) *Proc. Natl. Acad. Sci. USA* 86, 7828–7832.
18. Marciniak, R.A., Garcia-Blanco, M.A. and Sharp, P.A. (1990) *Proc. Natl. Acad. Sci. USA* 87, 3642–3628.
19. Gatignol, A., Buckler-White, A., Berkhout, B. and Jeang, K.-T. (1991) *Science* 251, 1597–1600.
20. Marciniak, R.A., Calnan, B.J., Frankel, A.D. and Sharp, P.A. (1990) *Cell* 63, 791–802.
21. Garcia, J.A., Harrich, D., Pearson, L., Mitsuyasu, R. and Gaynor, R.B. (1988) *EMBO J.* 7, 3143–3147.
22. Frankel, A.D., Bredt, D.S. and Pabo, C.O. (1988) *Science* 240, 70–73.
23. Kuppuswamy, M., Subramanian, T., Srinivasan, A. and Chinnadurai, G. (1989) *Nucl. Acids Res.* 17, 3551–3561.
24. Ruben, S., Perkins, A., Purcell, R., Joung, K., Sia, R., Burghoff, R., Haseltine, W.A. and Rosen, C.A. (1989) *J. Virol.* 63, 1–8.
25. Rice, A.P. and Charlotti, F. (1990) *J. Virol.* 64, 1864–1868.
26. Dang, C.V. and Lee, W.M.F. (1989) *EMBO J.* 7, 3143–3147.
27. Hauber, J., Malim, M.H. and Cullen, B.R. (1989) *J. Virol.* 63, 1181–1187.
28. Siomi, H., Shida, H., Maki, M. and Hatanaka, M. (1990) *J. Virol.* 64, 1803–1807.
29. Subramanian, T., Kuppuswamy, M., Venkatesh, L., Srinivasan, A. and Chinnadurai, G. (1990) *Virology* 176, 178–183.
30. Subramanian, T., Govindarajan, R. and Chinnadurai, G. (1991) *EMBO J.* 10, 2311–2318.
31. Berkhout, B., Silverman, R.H. and Jeang, K.-T. (1989) *Cell* 59, 273–282.
32. Hauber, J. and Cullen, B.R. (1988) *J. Virol.* 62, 673–679.
33. Garcia, J.A., Harrich, D., Soultanakis, E., Wu, F., Mitsuyasu, R. and Gaynor, R.B. (1989) *EMBO J.* 8, 765–778.

33. Selby, M.J., Bain, E.S., Luciw, P. and Peterlin, B.M. (1989) *Genes Dev.* 3, 547–555.
34. Roy, S., Parkin, N.T., Rosen, C.A., Itovich, J. and Sonenberg, N. (1990) *J. Virol.* 64, 1402–1406.
35. Jakobovits, A., Smith, D.H., Jakobovits, E.B. and Capon, D.J. (1988) *Mol. Cell. Biol.* 8, 2555–2561.
36. Feng, S. and Holland, E.C. (1988) *Nature* 334, 165–167.
37. Berkhout, B. and Jeang, K.-T. (1989) *J. Virol.* 63, 5501–5504.
38. Emerman, M., Guyader, M., Montagnier, L., Baltimore, D. and Muesing, M.A. (1987) *EMBO J.* 6, 3755–3760.
39. Berkhout, B., Gagnon, A., Silver, J. and Jeang, K.-T. (1990) *Nucl. Acids Res.* 18, 1839–1846.
40. Berkhout, B. and Jeang, K.-T. (1990) *Nucl. Acids Res.* 18, 6903–6907.
41. Gorman C., Moffat, L. and Howard, B. (1982) *Mol. Cell. Biol.* 2, 1044–1051.
42. Seed, B. and Sheen, J.Y. (1988) *Gene* 67, 271–277.
43. Berkhout, B. and K.-T. Jeang (1991) in *Advances in Molecular Biology and Targeted Treatment for AIDS* (Ed. A. Kumar) pp. 107–121, Plenum Press, New York.
44. Pleij, C.W.A., Rietveld, K. and Bosch, L. (1985) *Nucl. Acids Res.* 13, 1717–1731.
45. Weeks, K.M. and Crothers, D.M. (1991) *Cell* 66, 577–588.
46. Longfellow, C.E., Kierzek, R. and Turner, D.H. (1990) *Biochemistry* 29, 278–285.
47. Romisch, K., Webb, J., Lingelbach, K., Gausepohl, H. and Dobberstein, B. (1990) *J. Cell. Biol.* 111, 1793–1802.
48. Scherly, D., Dathan, N.A., Boelens, W., van Venrooij, W.J. and Mattaj, I.W. (1990) *EMBO J.* 11, 3675–3681.
49. Boelens, W., Scherly, D., Beijer, R.P., Jansen, E.J.R., Dathan, N.A., Mattaj, I.W. and van Venrooij, W.J. (1990) *Nucl. Acids Res.* 19, 455–460.
50. Witherell, G.W., Wu, H.-N. and Uhlenbeck, O.C. (1990) *Biochemistry* 29, 11051–11057.