The Apical Loop of the HIV-1 TAR RNA Hairpin Is Stabilized by a Cross-loop Base Pair*

Received for publication, February 24, 2003, and in revised form, July 24, 2003
Published, JBC Papers in Press, July 25, 2003, DOI 10.1074/jbc.M301939200

Tadeusz Kulinski‡, Mikolaj Olejniczak‡, Hendrik Huthoff§, Lukasz Bielecki‡, Katarzyna Pachulska-Wieczorek‡, Atze T. Das¶, Ben Berkhout§, and Ryszard W. Adamiak‡

From the 1Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12-14, 61-704 Poznań, Poland and the 2Department of Human Retrovirology, Academic Medical Center, University of Amsterdam, 1105AZ Amsterdam, The Netherlands

The TAR hairpin of the HIV-1 RNA genome is indispensable for trans-activation of the viral promoter and virus replication. The TAR structure has been studied extensively, but most attention has been directed at the three-nucleotide bulge that constitutes the binding site of the viral Tat protein. In contrast, the conformational properties of the apical loop have remained elusive. We performed biochemical studies and molecular dynamics simulations, which indicate that the TAR loop is structured and stabilized by a cross-loop base pair between residues C30 and G34. Mutational disruption of the cross-loop base pair results in reduced Tat response of the LTR promoter, which can be rescued by compensatory mutations that restore the base pair. Thus, Tat-mediated transcriptional activation depends on the structure of the TAR apical loop. The C30-G34 cross-loop base pair classes TAR in a growing family of hairpins with a structured loop that was recently identified in ribosomal RNA, tRNA, and several viral and cellular mRNAs.

The HIV-1life cycle depends on specific RNA-protein interactions. The 5’-end of all HIV-1 transcripts contains a 59-nucleotide stem-loop structure termed the trans-activation response element (TAR). Understanding of the TAR RNA structure, its dynamics and interaction with proteins is of importance for the design of potential therapeutic antiviral strategies (1). TAR RNA is indispensable for HIV-1 transcriptional regulation of the long-terminal repeat (LTR) promoter. Sp1 driven basal transcription at the LTR allows the production of a low level of HIV-1 transcripts needed for the synthesis of the viral transactivator protein, Tat (2–4). The TAR hairpin serves as a binding site for Tat and the interaction between the protein and RNA strongly stimulates the activity of the LTR promoter (5–11). Several cellular proteins that stimulate the expression of TAR-containing transcripts have been identified (12–21). Currently, the best studied cellular component of the Tat-activated phase of LTR transcription is the positive transcriptional elongation factor (pTEFb). The pTEFb complex associates with TAR through its cyclin T1 subunit in a Tat-dependent manner. Binding of Tat to a three-nucleotide bulge in TAR facilitates the interaction of cyclin T1 with the TAR apical loop (13, 22, 23). The kinase component of P-TEFb, CDK9, can then phosphorylate the C-terminal domain of RNA polymerase II, which enhances the processivity of the elongating polymerase (24–26).

The binding of Tat to the TAR bulge has been studied extensively. Numerous studies have addressed the structure of both the free TAR RNA (27–29) and TAR bound to substrates such as argininamide (27), Tat-related peptides (29, 30), small inhibitor molecules (31) and cations (32, 33). In addition, molecular dynamics simulations of TAR have been reported (34, 35). From these studies it is apparent that the conformation of the bulge differs significantly in the absence and presence of the Tat protein. The bulge residues are partially stacked in the free TAR RNA. Upon binding of Tat, the bulge is extruded from the TAR helix and a base triple is formed that involves the first bulge residue U133 and the A57-U60 base pair. The binding of Tat to TAR thus coincides with a dramatic change in the tertiary structure of TAR, which may occur by an induced fit or conformational capture mechanism (36).

In contrast, little is known about the structural properties of the six-nucleotide apical loop of the TAR hairpin. NMR studies demonstrated considerable conformational flexibility in the loop with a tendency to form a compact loop structure (30, 37, 38). It appears that the most structured regions are at either end of the loop, consistent with stacking of loop residues on the stem-closing base pair (30, 37, 39). Analysis of chemical shifts observed for 13C-adenosine enriched TAR residues (38) and relaxation experiments suggested that A35 has an unusual local environment and is extruded from the apical loop (40). A few investigators have proposed distinct structural features of the apical TAR loop. The possibility of a protonated non-canonical C30-A35H base pair, yielding a tetraloop structure for TAR has been suggested (39). Alternatively, a cross-loop base pair involving residue C30 and the penultimate loop residue G34 has been proposed (41). This interaction has been observed transiently in molecular dynamics simulations (34, 35), and it was recently suggested that the C30-G34 cross-loop base pair is required for binding of cyclin T1 (42).

In this study, we specifically address the structure of the TAR loop and its involvement in the HIV-1 LTR promoter activity. We present evidence that the TAR apical loop can indeed form the C30-G34 base pair. Biochemical studies and
molecular dynamics with a model TAR hairpin, and modified analogues thereof, indicate that residue G34 is oriented inside the loop. G34 can form relatively stable hydrogen bonds with C50. The C50-G34 base pair stack on the C50-G30 closing base pair of the stem. Subsequent analysis of mutant TAR constructs in LTR-luciferase assays confirmed the importance of the cross-loop base pair in Tat-mediated transcriptional activation.

EXPERIMENTAL PROCEDURES
Oligoribonucleotide Synthesis—The synthesis of 29-mer HIV-1 TAR RNA and its 2-aminopurine riboside modified analogues AP34 and AP35 was performed using solid-support aided phosphoramidite chemistry and applying the 2'-O-BDMSI protection (43). After 1 μmol of scale synthesis and deprotection with fluoride ion, the RNA was purified by preparative 19% polyacrylamide gel electrophoresis under denaturing conditions, followed by electrophoresis with a Biotrapt BT 1000 and desalting on NAP columns. The sequence of the oligomers was confirmed by RNA sequencing with ribonuclease T1.

UV Melting and Thermodynamics—UV melting of the chemically synthesized oligomers was measured on a temperature controlled Beckman DU70 spectrophotometer in the temperature range from 10 to 98 °C. The RNA was dissolved in 1.0 M NaCl, 10 mM Tris/HCl (pH 7.0). The RNA concentrations, varied over 100-fold range, were calculated from the absorbance at 260 nm measured at each 0.3 °C. The buffer used for thermodynamic studies was 1 M NaCl, 10 mM Tris/HCl (pH 7.0). The RNA was renatured by incubation at 85 °C for 3 min, followed by slow cooling to room temperature. Absorbance versus temperature profiles were fit to two-state transition model with sloping baselines by using the nonlinear least-squares program MelWin. The experiments were repeated 3–5 times.

Thermal denaturation of in vitro synthesized 1/466 HIV-1 RNA was measured on a PerkinElmer Life Sciences Lambda 2 spectrophotometer at a heating rate of 0.5 °C min−1 with sampling at each 0.1 °C. Samples contained ~3.0 μg of RNA dissolved in 140 μl of 50 mM sodium Cacodylate buffer (pH 7.2). Prior to the measurement, the RNA was renatured by incubation at 37 °C for 30 min. Electrophotophoresis of the samples prior to the analysis confirmed the monomeric state of the RNA. No significant degradation of the transcripts was apparent after the measurement as judged by analysis of the RNA on a sequencing gel.

Pb2+–induced RNA Phosphodiester Bond Cleavage—The oligoribonucleotides were 5'-end-labeled using T4 polynucleotide kinase, gel purified and quantitated on a Beckmann LS5000TA scintillation counter. The reaction mixture was incubated for 15 min at room temperature and applying the 2'-O-BDMSI protection (43). After 1 μmol of scale synthesis and deprotection with fluoride ion, the RNA was purified by preparative 19% polyacrylamide gel electrophoresis under denaturing conditions, followed by electrophoresis with a Biotrapt BT 1000 and desalting on NAP columns. The sequence of the oligomers was confirmed by RNA sequencing with ribonuclease T1.

UV Melting and Thermodynamics—UV melting of the chemically synthesized oligomers was measured on a temperature controlled Beckman DU70 spectrophotometer in the temperature range from 10 to 98 °C. The RNA was dissolved in 1.0 M NaCl, 10 mM Tris/HCl (pH 7.0). The RNA concentrations, varied over 100-fold range, were calculated from the absorbance at 260 nm measured at each 0.3 °C. The buffer used for thermodynamic studies was 1 M NaCl, 10 mM Tris/HCl (pH 7.0). The RNA was renatured by incubation at 85 °C for 3 min, followed by slow cooling to room temperature. Absorbance versus temperature profiles were fit to two-state transition model with sloping baselines by using the nonlinear least-squares program MelWin. The experiments were repeated 3–5 times.

Thermal denaturation of in vitro synthesized 1/466 HIV-1 RNA was measured on a PerkinElmer Life Sciences Lambda 2 spectrophotometer at a heating rate of 0.5 °C min−1 with sampling at each 0.1 °C. Samples contained ~3.0 μg of RNA dissolved in 140 μl of 50 mM sodium Cacodylate buffer (pH 7.2). Prior to the measurement, the RNA was renatured by incubation at 37 °C for 30 min. Electrophotophoresis of the samples prior to the analysis confirmed the monomeric state of the RNA. No significant degradation of the transcripts was apparent after the measurement as judged by analysis of the RNA on a sequencing gel.
**Structure of the HIV-1 TAR Apical Loop**

**RESULTS**

**Thermal Denaturation of HIV-1 RNA with a Mutated TAR Loop**—To test if the TAR loop is structured, we performed a UV-melting experiment with transcripts corresponding to the 5’-terminal 466 nucleotides of the wild-type HIV-1 RNA and a mutant in which the apical loop residues 32–34 were changed to GGG by an AP leads to a significant destabilization of TAR hairpin. The high-temperature transition $T_m$ has previously been determined by transfection of SupT1 T cells with $0.5 \mu g$ of wild-type or TAR mutated pBlue3LTR-luc, 0.5 $\mu g$ of pRL-CMV, and 10% fetal calf serum. Therefore, the modified hairpin AP34 exhibits significantly increased mobility compared with the wild-type hairpin (Fig. 2B). In contrast, the AP35 hairpin shows practically the same mobility as the wild-type hairpin. This effect was independent of different incubation procedures used for refolding of the RNA prior to non-denaturing PAGE (results not shown). Thus, the different electrophoretic mobility of hairpin AP34 indicates a substantial change in the structural properties of its apical loop in comparison with the wild-type and AP35 hairpins.

Next, we performed melting experiments with the 29-mer hairpins and derived the thermodynamic parameters by numerical analysis (Table I). Each hairpin melts in a single transition and melting temperatures ($T_m$) were independent on the RNA concentration in the range from 0.3 to 15 $\mu M$, which indicates the absence of dimeric or higher order RNA complexes. Both AP34 and AP35 are less stable than the wild-type RNA. The difference of melting temperature $T_m$ between the wild-type TAR and AP35 is small (0.6 °C), but the $\Delta G^{\circ}$ of AP35 is $-0.81$ kcal/mol higher than that of the wild-type RNA. Thermodynamic parameters of AP34 and AP35 are more evidently different. The $T_m$ of AP34 is 3.2 °C lower and $\Delta G^{\circ}$ is 1.0 kcal/mol higher than that of wild-type TAR. Thus, the replacement of purine-substituted TAR Hairpins—To study the structure of the TAR loop in more detail, we synthesized 29-mer TAR hairpins with the wild-type sequence and 2-aminopurine substitutions at position 34 or 35 (Fig. 2A). When these oligoribonucleotides were checked for their purity in denaturing PAGE, all three fragments demonstrated the same electrophoretic mobility (results not shown). However, under non-denaturing PAGE conditions, the modified hairpin AP34 exhibits significantly increased mobility compared with the wild-type hairpin (Fig. 2B). In contrast, the AP35 hairpin shows practically the same mobility as the wild-type hairpin. This effect was independent of different incubation procedures used for refolding of the RNA prior to non-denaturing PAGE (results not shown). Thus, the different electrophoretic mobility of hairpin AP35 indicates a substantial change in the structural properties of its apical loop in comparison with the wild-type and AP35 hairpins.

**Electrophoretic and Thermodynamic Properties of 2-Aminopurine-substituted TAR Hairpins**—To study the structure of the TAR loop in more detail, we synthesized 29-mer TAR hairpins with the wild-type sequence and 2-aminopurine substitutions at position 34 or 35 (Fig. 2A). When these oligoribonucleotides were checked for their purity in denaturing PAGE, all three fragments demonstrated the same electrophoretic mobility (results not shown). However, under non-denaturing PAGE conditions, the modified hairpin AP34 exhibits significantly increased mobility compared with the wild-type hairpin (Fig. 2B). In contrast, the AP35 hairpin shows practically the same mobility as the wild-type hairpin. This effect was independent of different incubation procedures used for refolding of the RNA prior to non-denaturing PAGE (results not shown). Thus, the different electrophoretic mobility of hairpin AP35 indicates a substantial change in the structural properties of its apical loop in comparison with the wild-type and AP35 hairpins.

**Cells and Transfection**—Transfection of C33A was performed the calcium phosphate method as previously described (53). SupT1 T cells were grown in RPMI 1640 medium containing 10% fetal calf serum at 37 °C and 5% CO$_2$. Cells were transfected by electroporation. 5$\times$10$^6$ cells were washed in RPMI 1640 with 20% fetal calf serum, mixed with DNA in 0.4-cm cuvettes and electroporated at 250 volts and 960 $\mu$F, and then resuspended in RPMI 1640 with 10% fetal calf serum. The transcriptional activity of the wild-type and TAR-mutated HIV-1LTRs was determined by transfection of SupT1 T cells with 5 $\mu g$ of wild-type or TAR mutated pBlue3LTR-luc, 0.5 $\mu g$ of pRL-CMV, with or without 0.5 $\mu g$ of pcDNA3-Tat. At 2 days after transfection, cells were isolated by centrifugation at 260 $g$ for 10 min. The Dual-Luciferase Reporter assay system was used to determine the firefly and Renilla luciferase levels.

**Comparative Sequence Analysis**—Sequences of different HIV-1 isolates were obtained from the HIV data base at Los Alamos (lanl.gov/). Only isolates with a complete TAR sequence were considered in the analysis, yielding a total of 654 sequences from most of the prevailing HIV-1 subtypes.
than A35 in maintaining a structured apical loop.

Lead-induced Cleavage of 2-Aminopurine-substituted TAR Hairpins—Metal ion promoted RNA cleavage can be used to map single stranded regions and metal coordination sites in an RNA structure (54). We previously used this approach to demonstrate that the TAR bulge binds metal ions in solution (33). Here, we used Pb2+-induced cleavage to probe for changes in the TAR loop imposed by the 2-aminopurine substitutions. Overall, the cleavage pattern in the loop is quite similar for the wild-type and AP35 hairpins (Fig. 3). At the 5’-side of the loop, reactivity of the phosphodiester bond between C30-U31 and U31-G32 dominates the cleavage pattern. Cleavage is weaker at position C39, reflecting less conformational freedom of this base pair residue. Toward the 3’-side of the loop, there are weak cleavages at G34, G33, and G34, but backbone scission is more pronounced at residue G34. The 2-aminopurine substitution in the AP34 hairpin strongly enhances the cleavage at position 34. Indeed, residue 34 is most prone to cleavage in the loop of the AP34 hairpin. This result indicates that the AP34 substitution results in an increased conformational flexibility of the TAR loop. In contrast, the intensity of cleavage at position 35 in AP35 is moderately decreased, indicating that the integrity of the apical loop is less sensitive to modification of residue 35. These results confirm that the loop structure is highly dependent on the nature of residue 34.

Molecular Dynamics Simulations of the TAR Apical Loop Structure—To evaluate the structural trends within the apical loop we have undertaken molecular dynamics (MD) simulations. We used two different approaches in preparing the initial model for MD simulations, thereby extending the conformational space searched for possible loop structures. In the first approach we have selected 1 out of 20 available NMR structures of the TAR RNA (34). We modeled the AP34 and AP35 substitutions onto the TAR structure to evaluate their effect on the loop conformation. In the second approach, we built models of a 14-mer TAR mini-hairpin (position 26–39) in which we imposed base pairing between C30 and G34, followed by running the MD simulation without any constraints to investigate the stability of a C30-G34 base pair. The 29-mer wild-type TAR, AP34, and AP35 hairpins were built based on the coordinates of model No. 17 in PDB file 1ANR (30). The in aquea MD simulation for the wild-type hairpin was conducted for 1.6, and 1.0 ns simulations were performed for the AP34 and AP35 hairpins at 300 K. These simulations revealed considerable dynamics of the bulge and apical loop in all three hairpins. The C3’-endo sugar puckering was maintained largely in the residues of the double stranded stem. In the apical loop, sugar rings were considerably unstable. Frequent switches to the C2’-endo pucker and subsequent stabilization of this conformation was observed. Interestingly, loop residue 34 uniquely shows a stable C3’-endo puck for all three hairpins. In all three cases, residue 35 retains its looped-out orientation. In the wild-type and AP35 simulations, residue 34 has a tendency to interact with the upper part of the double stranded region by stacking on the stem closing C39-G36 base pair, but this stacking is much less pronounced for the AP34 hairpin (Fig. 4A). This may explain the contribution of residue 34 to the hairpin stability. Only few of the potential hydrogen-bonding contacts within the loop are stably observed, and these mainly involve residue G34 in the wild-type and AP35 hairpins. G34 can form a relatively stable hydrogen bond involving its O6 and the N(4)H2 of C30. In parallel, the exo-NH2 group of G34 interacts with the sugar moiety of C30. In the AP35 hairpin, transient stacking interactions of C30 with U31, and G32 with G33, were observed despite considerable flexibility. In the final stages of the simulation with the AP34 hairpin, the 2-aminopurine at position 34 starts to interact more strongly with the looped out residue A35. In general terms, the AP34 hairpin
shows the least tendency toward a structured loop conformation during the simulation, which is consistent with the results from our experimental analyses.

To further investigate the possible interactions between C30 and G34, we modeled the C30–G34 base pair in a 14-mer TAR hairpin and then performed extensive MD simulations at different temperatures (Fig. 4, B and C). The initial models were built based on fitting the TAR sequence on the NMR structure of the HIV-1 A-rich hexaloop (PDB code 1bvj) (48). We selected this structure as a template for our modeling because of its well-defined structural motif characterized by 3′-stacking of four adenosines in this hexaloop. The resulting structure was remodeled by imposing constraints for the C30–G34 interaction, followed by running a short MD. This provided a set of initial structures for further MD investigations. All constraints were then removed and full MD trajectories were collected running simulations for 2 up to 8 ns, depending on the convergence of the simulated structures, at three different temperatures: 300, 320, and 340 K. This yielded a range of conformational motifs characterized by different internal interactions and exhibiting diverse stability and convergence within the MD trajectories in a total of 36 ns of simulation time.

The most stable MD trajectories were obtained with models that maintain the interaction between C30 and G34 (Fig. 4C). We consistently observed that C30 is repositioned slightly by stacking on the upper stem toward the major groove, exposing its Watson-Crick face across the loop toward G34. Due to the local reorientation of the backbone around G34, this base is shifted with respect to A35 and points across the loop toward C30. In this orientation, all three Watson-Crick C–G hydrogen bonds across the loop were observed with high occupancy. As a consequence of the interaction between C30 and G34, two sharp turns in the phosphodiester backbone are observed: the first between U31 and G32 and the second between G33 and G34. We observed a strong tendency toward stacking interactions between G32 and G33. This two-base stack is frequently oriented toward the outside of the loop, forming a structural module of relatively high mobility, although there are numerous events that disrupt this stacking in a reversible manner. These results indicate that hydrogen bonds between C30 and G34 are highly favorable and play a pivotal role in forming a structured conformation of the apical TAR loop. Additional interactions may stabilize this base pair. In particular, we observed additional hydrogen bonds between C30 and A35 in conjunction with the C30–G34 base pair.

Site-directed Mutagenesis and Structure Probing of the C30–G34 Base Pair—Because the previous analyses suggest the formation of a base pair between residues C30 and G34 within the TAR loop, we decided to directly test this possibility by generating mutant TAR constructs in which this base pair is either disrupted or replaced by another base pair. Six different mutants were made, comprising two sets in which the base pair is either disrupted or restored (Fig. 5A). Mutants G30 and C34 disrupt the proposed cross-loop base pair, which is potentially restored in the double mutant G30C34. Likewise, mutants U30 and A35 disrupt cross-loop pairing, but the double mutant U30A35 could form an alternative base pair.

We performed RNA structure probing on the wild-type and mutant TAR hairpins as 1–80 transcripts to evaluate the structural effects of these mutations in the loop. For the wild-type,
our data are in good agreement with the secondary structure model of TAR. The single-nucleotide bulges C5 and A17 are both identified by DMS modification, as are U24 and U25 in the three-nucleotide bulge (Fig. 5A, lane 3). In addition, we observed DMS reactivity at A22, which forms a distorted base pair with U40 just below the bulge and its sensitivity to DMS and DEPC has been documented repeatedly (10, 55). These features are maintained in all the mutant TAR elements, demonstrating that the loop mutations do not result in gross rearrangement of the TAR secondary structure. Structure probing of the wild-type TAR loop shows some characteristics of a structured loop. The three guanine residues are differentially modified by kethoxal, with reactivity being high at G32, intermediate at G33 and low at G34 (Figs. 5A, lane 2 and 6B, wt). The same pattern is observed when the RNA is probed with RNase T1 (results not shown) (11, 13, 37), and this observation was previously used to propose that G34 pairs with C30 (41). Indeed, C30 is protected from modification by DMS (Fig. 5A, lane 3 and C, wt). U31 and A35 are sensitive to DMS modification, which confirms the exposure of the functional groups of these residues to the solvent.

Among the set of TAR loop mutants, significant changes in the probing profile are observed. The most severe effect is observed for mutants that have a base change at position 30 (G30, G30C34, U30, and U30A34), which results in the loss of the characteristic reactivity of the guanine residues (Fig. 5A, lanes 8, 11, 17, and 20). G32 and G33 are no longer differentially modified and G34 is more reactive than in the wild-type context (Fig. 5B). Furthermore, the pattern of reverse transcriptase stops is redistributed by a loss of the stop at position 31 and the emergence of a novel stop at position 38 for all mutations at position 30 (Fig. 5A). These results indicate that C30 is instrumental in maintaining a structured TAR loop.

Mutation of position 34 has a less severe effect on the overall reactivities in the TAR loop, but it is apparent that both mutants C34 and A34 are reactive at position 34, whereas the wild-type G34 is not reactive (Fig. 5A, lanes 6 and 15, and Fig. 6C). In addition, both mutants exhibit increased reactivity of A35, which indicates that mutation of residue 34 disturbs the local loop structure. For mutant A34, there is an additional increase in the reactivity of C30, which is protected in the wild-type and in mutant C34 (Fig. 5C, mutant A34). Thus, substituting G34 with an A results in exposure of residue 34 and the absence of increased reactivity at C30 in mutant C34 could be due to the presence of a pyrimidine at position 34 in this mutant, which may be insufficiently large to displace C30.

Structure probing with the single mutants G30, U30, C33, and A34 demonstrates that the loop is significantly altered by these sequence changes and these results are compatible with the disruption of a base pair involving residues C30 and G34. The double mutants G30C34 and U30A34 can form an alternative cross-loop base pair that may restore the loop structure. Close inspection of the base reactivities at positions 34 and 35 provides evidence that this may indeed be the case. We already showed that substituting G34 with C or A results in a high reactivity of the mutant residue 34 and an increase in reactivity of A35 (Fig. 5D and E, respectively). Furthermore, substituting residue C30 with G or U reduces the reactivity of A35 (Fig. 5D and E). In the double mutants G30C34 and U30A34, the reactivity of residue A35 is restored to levels similar to that...
of the wild-type TAR. In addition, residue 34 is slightly protected in the double mutants compared with the single mutants. Thus, the double mutants restore the wild-type behavior of residue A35 and possibly the base pairing between residues 30 and 34.

Transcriptional Activity of LTR Promoters with a Wild-type or Mutant TAR Loop—We next assessed the effect of mutations in the TAR loop on the activity of the LTR promoter. C33A cells were transiently transfected with the wild-type and mutant LTR-luciferase reporter plasmids with and without a Tat expression plasmid. In the absence of Tat, the wild-type and mutant TAR elements have similar LTR activity, indicating that basal transcription is not influenced by the structure of the TAR loop (Fig. 6A). The wild-type LTR activity with Tat was set at 100%, corresponding to a 24-fold activation over the basal activity without Tat (Fig. 6, A and B). Mutants C34 and A34 cause the most severe reduction in Tat-mediated transcriptional activation, with a modest 2-fold LTR activation. Mutants G30 and U30 have a less severe effect, but the Tat response is significantly reduced to 6-fold. The double mutant
G30C34 reaches a Tat response of about 5-fold, a phenotype that is intermediate between that of the two individual mutants G30 and C34. Thus, the double mutant partially compensates for the defect caused by the C34 mutation. This compensatory effect is more prominent for the double mutant U30A34, which demonstrates a Tat response that exceeds the level observed for the individual mutants U30 and A34. These results indicate that the base pair between residues 30 and 34 plays a role in Tat-mediated activation of the LTR promoter, but has no effect on basal transcription. Similar results were obtained in the SupT1 T-cell line (Fig. 6, C and D), indicating that these effects hold true in the cell type that is physiologically relevant to HIV-1 infection.

**DISCUSSION**

The structure of the HIV-1 TAR element has been studied extensively, but the conformation of the apical loop has remained elusive. Based on structure probing and modeling, it has been suggested that the TAR loop is structured through a cross-loop base pair involving nucleotides C30 and G34 (41), but this model has received little attention. Recently, hydrogen bonds involving G34 have been implied in stabilizing the apical loop (34) and the C30–G34 base pair was observed transiently in molecular dynamics simulations (56). Furthermore, binding of cyclin T1 to TAR appears to require this base pair (42). To date, it has not been demonstrated whether a structured conformation of the TAR loop is required for trans-activation of the LTR promoter.

The results we present in this study further support the notion that the TAR loop is stabilized by the C30–G34 cross-loop base pair. We observed destabilization of the TAR hairpin upon mutation of the loop, indicating that the wild-type sequence contributes to the overall stability of the TAR hairpin. Studies with model TAR hairpins containing 2-aminopurine riboside substitutions in the loop support this idea. Non-denaturing PAGE analysis and lead-probing showed that G34 plays a key role in maintaining a structured TAR loop. Introducing the modified base at position 34 (AP34) alters two putative Watson-Crick hydrogen bonds because the acceptor-donor properties of the functional groups at the position 1 and 6 of the guanosine are changed. The corresponding change in free energy that was measured for AP34 is within the range associated with the elimination of one or two hydrogen bonds (57).

The molecular modeling and MD simulations demonstrate that the TAR loop has a tendency to adopt a stable structure through internal hydrogen bonds. We observed a base-specific interaction between C30 and G34 involving the O(6) of G34 and the N(4)H2 of C30 in MD simulations of the 29-mer TAR RNA. Further MD simulations on the 14-mer TAR model hairpin indicated that all Watson-Crick hydrogen bonds between C30 and G34 can be formed. The geometry of these interactions may be distorted, which could explain why the base pair was not readily detected in previous NMR studies. Close inspection of these NMR-derived structure models provides further evidence that hydrogen bonds between residues C30 and G34 are possible within the TAR apical loop. These TAR models show a strong tendency for an in-loop orientation of G34 by stacking on G36, and this exposes the Watson-Crick side toward C30 (30). Indeed, this is consistent with a NOE sequential contact between H8–G34 and H8–G36 (30). C30 efficiently stacks on C29, but commonly adopts an orientation away from G34. However, several structure models in the ensemble allow for hydrogen bonding between C30 and G34. Previously, it has proven difficult to assign the imino proton resonances of G32, G33, and G34 in NMR studies of the HIV-1 TAR hairpin. In one study these remained unassigned (11.05, 10.86, and 10.62 ppm) (39), while in another the assignment was non-specific (11.58 (G32), 10.82 (G33), and 10.59 (G34) ppm) (29). Two of these resonances are within the range of unpaired guanines, but one consistently appears off-range above 11.00 ppm. The participation of G34 in a pairing interaction with C30 would be consistent with assigning the off-range resonance to G34.

RNA structure probing of the wild-type TAR RNA and mutants with a disrupted or substituted base pair at positions 30 and 34 showed base reactivity changes that are consistent with the presence of a C30–G34 base pair. Reactivity changes at positions 34 and 35 in the double mutants compared with single mutants indicated that the double mutants may partially restore an interaction between residues 30 and 34. In LTR-luciferase assays, we observed a partial rescue of the Tat response with the G30C34 double mutant and true rescue with
the U30A34 double mutant as compared with constructs with the individual mutations. This demonstrates that the cross-loop base pair is required for Tat-mediated trans-activation of the HIV-1 promoter. The mutant analysis also suggests that this requires a 5′-pyrimidine and a 3′-purine in the cross-loop base pair, since mutant G30C34 only partially rescues transcriptional activation. Because the U30A34 mutant does not reach wild-type levels of promoter activity in the presence of Tat, it is likely that the sequence as well as the structure of the TAR loop contribute to the optimal Tat response of the LTR promoter.

Recently, two groups have independently recognized the frequent occurrence of RNA loops containing a cross-loop base pair capped by a tri-loop in ribosomal RNA, and each proposed a different nomenclature for this motif, the T-loop or the lone pair capped by a tri-loop in ribosomal RNA, and each proposed a different nomenclature for this motif, the T-loop or the lone pair triloop motif. Crick pairing across the loop are common (59). Our results demonstrate that the loss of protein binding due to mutation of residue 34 is partially rescued by a complementary sequence at position 30. Throughout these different analyses, C30 and G34 consistently emerge as conserved and functionally important residues, which is compatible with the structured conformation of the TAR loop containing a C30–G34 cross-loop base pair.

We observed that the structure of the TAR apical loop affects the Tat-dependent phase of transcription from the LTR promoter and this suggests that the structured loop conformation may not be exclusively attributable to the TAR-cyclin T1 interaction. Indeed, cyclin T1 also affects transcription at the LTR promoter independently of its interaction with TAR (69).

In addition, other cellular proteins that bind to the TAR loop have been described (20, 70–72) and may require the C30–G34 cross-loop base pair for recognition of TAR.

Acknowledgments—We thank Wim van Est for preparing the figures and René Oldsohn and Joost Haasnoot for fruitful discussions. Access to CRAY J916 at the Poznan Supercomputing and Networking Centre is acknowledged.

REFERENCES

1. Hamy, F., Felder, E. R., Heizmann, G., Lazdins, J., Aboul-ela, F., Varani, G., Karm, J., and Klitmøller, T. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 3545–3553
2. Jones, K. A., Kadonaga, J. T., Luciw, P. A., and Tjian, R. (1986) Science 232, 755–759
3. Harrich, D., Garcia, J., Wu, F., Mitsuyasu, R., Gonzalez, J., and Gaynor, R. (1989) J. Virol. 63, 2555–2591
4. Berkhourt, B., and Jeang, K. T. (1992) J. Virol. 66, 139–149
5. Berkhourt, B., and Jeang, K. T. (1989) J. Virol. 63, 5501–5504
6. Frankel, A. D., Bingel, C., and Alston, B. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 7397–7401
7. Roy, S., Parkin, N. T., Rosen, C., Hovi, T., and Sonenberg, N. (1990) J. Virol. 64, 1403–1406
8. Dingwall, C., Emberg, I., Gait, M. J., Green, S. M., Heaphy, S., Karm, J., Lowe, A. D., Singh, M., and Skinner, M. A. (1990) EMBO J. 9, 4145–4153
9. Sumner-Smith, M., Roy, S., Barnett, R., Besd, L. S., Kuperman, R., Delling, U., and Sonenberg, N. (1991) J. Virol. 65, 5186–5202
10. Weeks, K. M., and Crothers, D. M. (1991) Cell 66, 577–588
11. Harper, J. W., and Logsdon, N. J. (1991) Biochemistry 30, 8060–8066
12. Mansingh, M. A., Smith, D. H., and Capon, D. J. (1987) Cell 48, 691–701
13. Wei, P., Garber, M. E., Fang, S.-M., Fisher, W. H., and Jones, K. A. (1998) Cell 92, 451–462
14. Zhang, J., Tumiaras, N., Hwang, S., Garber, M. E., Huq, I., Jones, K. A., and Rana, T. M. (2000) J. Biol. Chem. 275, 34314–34319
15. Rana, T. M., and Sonenberg, N. (1991) J. Virol. 65, 5186–5202
16. Weeks, K. M., and Crothers, D. M. (1991) Cell 66, 577–588
17. Chen, G. Y., Tretiaikov, A. P., Krakhmarov, C. P., Johnson, E. M., and Khalili, K. (1998) Gene (Amst.) 210, 37–44
18. Han, X. M., Laras, A., Bounseville, M. P., Kumar, A., and Shank, P. R. (1992) J. Virol. 66, 4065–4072.
19. Wu-Baer, F., Sigman, D. A., and Gaynor, B. R. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 7153–7157.
20. Wu, F., Garcia, J., Sigman, D., and Gaynor, R. (1991) Genes Dev. 5, 2128–2140.
21. Gatkof, A., Buckler-White, A., Berkhot, B., and Jeang, K. T. (1991) Science 251, 1597–1600.
22. Wimmer, J., Fujinaga, K., Tsubo, R., Cuje, T. P., Zhi, Y., Peng, J., Price, D. H., and Peterlin, B. M. (1999) Virology 255, 182–189.
23. Richter, S., Ping, Y. H., and Rana, T. M. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 7928–7933.
24. Bieniawsk, P. D., Godina, T. A., Bogerd, H. P., and Cullen, B. R. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 7791–7796.
25. Parada, C. A., and Roeder, R. G. (1998) Nature 394, 375–378.
26. Marciniak, R. A., and Sharp, P. A. (1991) EMBO J. 10, 4189–4196.
27. Puglisi, J. D., Tan, R., Calnan, B. J., Frankel, A. D., and Williamson, J. R. (1992) Science 257, 76–80.
28. Aboul-ela, F., Karn, J., and Varani, G. (1996) Nucleic Acids Res. 24, 3974–3981.
29. Long, K. S., and Crothers, D. M. (1999) Biochemistry 38, 10059–10069.
30. Aboul-ela, F., Karn, J., and Varani, G. (1995) J. Mol. Biol. 253, 313–332.
31. Faber, C., Sticht, H., Schweimer, K., and Roeh, P. (2000) J. Biol. Chem. 275, 20660–20666.
32. Papolito, J. A., and Steitz, T. A. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 9819–9824.
33. Olejniczak, M., Gdaniec, Z., Fischer, A., Grabarkiewicz, T., Bielecki, L., and Adamiak, R. W. (1999) Collect. Czech. Chem. Commun. Symp. Ser. 2, 191–196.
34. Nifosi, R., Reyes, C. M., and Kollman, P. A. (2000) Nucleic Acids Res. 28, 4944–4955.
35. Leulliot, N., and Varani, G. (2001) Biochemistry 40, 7947–7956.
36. Colvin, R. A., and Garcia-Blanco, M. A. (1992) J. Virol. 66, 930–935.
37. Michnicka, M. J., Harper, J. W., and Cullen, G. C. (1993) Biochemistry 32, 385–400.
38. Jaeger, J. A., and Tineci, I., Jr. (1993) Biochemistry 32, 12522–12530.
39. King, G. C., Harper, J. W., and Xi, Z. (1995) Methods Enzymol. 261, 436–450.
40. Critchley, A. D., Hanford, I., Cowan, D. J., and Stockley, P. G. (1993) J. Mol. Graph. 11, 92–97.
41. Richter, S., Cao, H., and Rana, T. M. (2002) Biochemistry 41, 6391–6397.
42. Usman, N., Ogilvie, K. K., Jiang, M. Y., and Cedergren, R. J. (1987) J. Am. Chem. Soc. 109, 7845–7854.
43. Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., DeBoh, S., Ferguson, D., Seibel, G., and Kollman, P. (1995) Comput. Phys. Commun. 74, 1–41.
44. Frish M. J., Trucks G. W., Schlegel H. B., Gill P. M. W., Johnson B. G., Robb M. A., Cheeseman J. R., Raghavachari K., Al-Laham M. A., zakrzewski V. G., Ortiz J. V., Foresman J. B., Ciosiowski J., Stefanov B. B., Nanayakkara A., Challacombe M., Peng C. Y., Ayala P. Y., Chen W., Wong M. W., Andrews J. L., Replogle E. S., Gompper R., Martin R. L., Fox D. J., Binkley J. S., Defrees D. J., Baker J., Stewart J. J. P., Head-Gordon M., Gonzalez C., and Pope J. A. (1996) Gaussian 94, Gaussian, Inc., Pittsburgh, PA.
45. Ciesiolka, J., Marciniec, T., and Krzyzosiak, W. (1989) Eur. J. Biochem. 182, 445–450.
46. Baudin, F., Marquet, R., Jiel, C., Darlin, J. L., Ehresmann, B., and Ehresmann, C. (1993) J. Mol. Biol. 229, 382–397.
47. Reyes, C. M., Nifosi, R., Frankel, A. D., and Kollman, P. A. (2001) Biophys. J. 80, 2833–2842.
48. Ciesiolka, J., Marciniec, T., and Krzyzosiak, W. (1999) J. Biol. Chem. 274, 1597–1600.
49. Puglisi, J. D., Tan, R., Calnan, B. J., Frankel, A. D., and Williamson, J. R. (1992) Science 257, 76–80.
50. Richter, S., Ping, Y. H., and Rana, T. M. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 7928–7933.
51. Parada, C. A., and Roeder, R. G. (1996) Nature 382, 375–378.
52. Marciniak, R. A., and Sharp, P. A. (1991) EMBO J. 10, 4189–4196.
53. Puglisi, J. D., Tan, R., Calnan, B. J., Frankel, A. D., and Williamson, J. R. (1992) Science 257, 76–80.
54. Ciesiolka, J., Marciniec, T., and Krzyzosiak, W. (1999) Eur. J. Biochem. 182, 445–450.
55. Colvin, R. A., and Garcia-Blanco, M. A. (1992) J. Virol. 66, 930–935.
56. Michnicka, M. J., Harper, J. W., and Cullen, G. C. (1993) Biochemistry 32, 385–400.
57. Jaeger, J. A., and Tineci, I., Jr. (1993) Biochemistry 32, 12522–12530.
58. King, G. C., Harper, J. W., and Xi, Z. (1995) Methods Enzymol. 261, 436–450.
59. Critchley, A. D., Hanford, I., Cowan, D. J., and Stockley, P. G. (1993) J. Mol. Graph. 11, 92–97.
60. Richter, S., Cao, H., and Rana, T. M. (2002) Biochemistry 41, 6391–6397.
61. Usman, N., Ogilvie, K. K., Jiang, M. Y., and Cedergren, R. J. (1987) J. Am. Chem. Soc. 109, 7845–7854.
62. Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., DeBoh, S., Ferguson, D., Seibel, G., and Kollman, P. (1995) Comput. Phys. Commun. 74, 1–41.
63. Frisch M. J., Trucks G. W., Schlegel H. B., Gill P. M. W., Johnson B. G., Robb M. A., Cheeseman J. R., Raghavachari K., Al-Laham M. A., zakrzewski V. G., Ortiz J. V., Foresman J. B., Ciosiowski J., Stefanov B. B., Nanayakkara A., Challacombe M., Peng C. Y., Ayala P. Y., Chen W., Wong M. W., Andrews J. L., Replogle E. S., Gompper R., Martin R. L., Fox D. J., Binkley J. S., Defrees D. J., Baker J., Stewart J. J. P., Head-Gordon M., Gonzalez C., and Pope J. A. (1996) Gaussian 94, Gaussian, Inc., Pittsburgh, PA.
64. Ciesiolka, J., Marciniec, T., and Krzyzosiak, W. (1999) Eur. J. Biochem. 182, 445–450.