RNA (2001), *7*:143–157. Cambridge University Press. Printed in the USA.

Two alternating structures of the HIV-1 leader RNA

HENDRIK HUTHOFF and BEN BERKHOUT
Department of Human Retrovirology, Academic Medical Center, University of Amsterdam, The Netherlands

ABSTRACT

In this study we demonstrate that the HIV-1 leader RNA exists in two alternative conformations, a branched structure consisting of several well-known hairpin motifs and a more stable structure that is formed by extensive long-distance base pairing. The latter conformation was first identified as a compactly folded RNA that migrates unusually fast in nondenaturing gels. The minimally required domains for formation of this conformer were determined by mutational analysis. The poly(A) and DIS regions of the leader are the major determinants of this RNA conformation. Further biochemical characterization of this conformer revealed that both hairpins are disrupted to allow extensive long-distance base pairing. As the DIS hairpin is known to be instrumental for formation of the HIV-1 RNA dimer, the interplay between formation of the conformer and dimerization was addressed. Formation of the conformer and the RNA dimer are mutually exclusive. Consequently, the conformer must rearrange into a branched structure that exposes the dimer initiation signal (DIS) hairpin, thus triggering formation of the RNA dimer. This structural rearrangement is facilitated by the viral nucleocapsid protein NC. We propose that this structural polymorphism of the HIV-1 leader RNA acts as a molecular switch in the viral replication cycle.

Keywords: dimer; HIV-1; NC protein; retrovirus replication; RNA structure rearrangement

INTRODUCTION

The full-length 9-kb HIV-1 RNA fulfills a dual role as mRNA and as the viral genome. The untranslated leader region of this RNA contains a multitude of regulatory structure and sequence elements that exert their respective effects at various stages during the viral replication cycle (Fig. 1). These regulatory functions can be roughly divided into two groups: regulation of gene expression (transcriptional activation, polyadenylation, translation, etc.) and virion-associated functions (dimerization, encapsidation, and reverse transcription). The dual nature of this RNA may correlate with distinct structural conformations that constitute the intracellular and virion-encapsidated phases of the RNA. It is well established that retroviral genomic RNA is found in virion particles as a dimer (Bender & Davidson, 1976).

Several secondary structure elements in the HIV-1 leader RNA have been well described. The first 97 nt of the HIV-1 leader constitute the R region, which is reiterated at the 3′ end of viral transcripts and is instrumental during the first strand transfer of reverse transcription. This domain contains the TAR and poly(A) hairpin structures (Berkhout, 1996). The TAR hairpin structure is supported by biochemical and biophysical data and fulfills an important function in transactivation of the LTR promoter (Muesing et al., 1985; Fisher et al., 1986; Berkhout et al., 1989; Aboul-ela et al., 1996; Wei et al., 1998). The poly(A) hairpin structure is supported by phylogenetic analysis and viral replication studies and has been shown to regulate polyadenylation of viral transcripts (Berkhout et al., 1995; Das et al., 1997, 1999). Immediately downstream of the R region lies a large central domain of the HIV-1 leader RNA that contains the primer binding site (PBS) that is complementary to the cellular $tRNA^{lys3}$, the natural primer of reverse transcription. Based on biochemical probing data and phylogeny, at least two different structures have been proposed for this domain (Isel et al., 1995; Berkhout, 1997; Damgaard et al., 1998). Four smaller hairpin structures have been described downstream of the PBS domain: the dimer initiation signal (DIS), the major splice donor (SD), the core packaging signal (Ψ) and a hairpin containing the AUG initiation codon of the Gag open reading frame (AUG) (reviewed in Berkhout, 1996). Of these, the DIS hairpin represents the most extensively studied element, and its role in initiating dimerization of HIV-1 transcripts is supported by a wide range of stud-

Reprint requests to: Ben Berkhout, Department of Human Retrovirology, University of Amsterdam, Academic Medical Center, Meibergdreef 15, 1105 AZ Amsterdam, The Netherlands; e-mail: b.berkhout@amc.uva.nl.

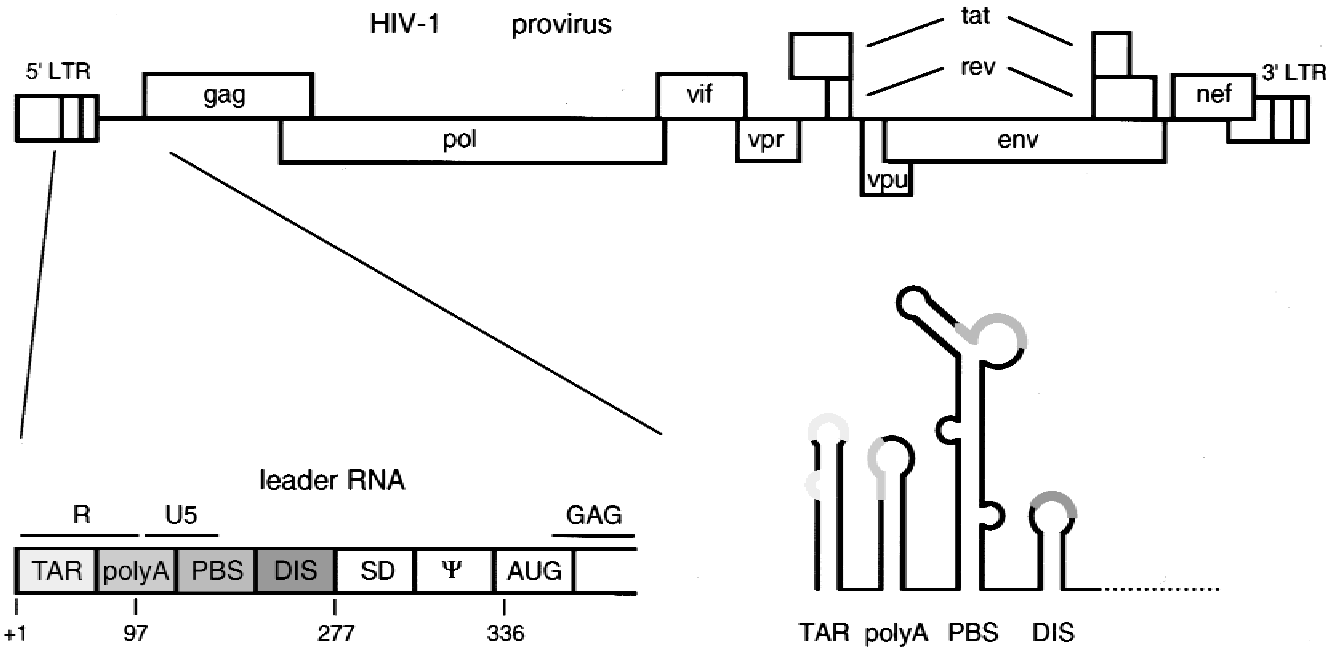

FIGURE 1. Organization of the HIV-1 provirus and untranslated leader RNA. For the provirus, the 5′ and 3′ LTRs and all open reading frames are indicated. For the leader RNA, previously proposed structured domains are boxed and identified by the corresponding abbreviations (see the text). On the right, a schematic representation highlights the hairpin structures for TAR, poly(A), PBS, and DIS.

ies, including NMR analyses (Laughrea & Jette, 1994; Skripkin et al., 1994; Mujeeb et al., 1998; Girard et al., 1999). Recently, the structures of the Ψ and SD hairpins have been studied by NMR (Amarasinghe et al., 2000; Zeffmandagger et al., 2000). Despite the apparent wealth of evidence for the individual secondary structure elements, relatively few studies have addressed the overall structure of the HIV-1 leader RNA (Baudin et al., 1993; Rizvi & Panganiban, 1993; Berkhout, 1996). As many of the biochemical studies, in particular NMR analyses, use short RNA fragments that do not represent the full leader RNA in its native state, it remains possible that important features such as tertiary structure and alternative conformations have been overlooked. We recently reported that the HIV-1 leader RNA adopts a compactly folded structure through a long-range RNA–RNA interaction (Berkhout & van Wamel, 2000). In this study, we have sought to define the determinants of this RNA conformation.

RESULTS

The loop and upper stem sequences of the DIS hairpin are required for conformer formation

It has previously been reported that the unusual fast migration of HIV-1 leader transcripts on nondenaturing gels is dependent on a long-range contact involving nucleotides 1/105 and 245/290 (Berkhout & van Wamel, 2000). In this study, we aimed to further fine-map the domains required for this long-range contact. We have used the characteristic migration on nondenaturing gels for identifying the interaction domains of the conformer. A set of HIV-1 leader transcripts with a common 5′ end (+1) and different 3′ ends between positions 245 and 290 were synthesized with T7 polymerase. The transcripts produce a ladder of increasing transcript size on denaturing gels containing urea (Fig. 2A). On nondenaturing gels, this pattern is lost, as transcripts that form the conformer migrate unusually fast through the gel (Fig. 2B). Formation of the fast-migrating conformer (marked C*) is not observed for transcripts that end at positions 245 and 256 (Fig. 2B, lanes 1 and 2), but efficient conformer formation is observed upon inclusion of 14 or more downstream nucleotides (Fig. 2B, lanes 3–5). This region of the HIV-1 leader RNA contains the DIS hairpin structure, and the critical 257–270 region corresponds to the loop and upper stem of this hairpin. Throughout this study we have used primer CN1, an antisense DNA oligo that disrupts the RNA conformer (Berkhout & van Wamel, 2000), as a denaturing control in nondenaturing gel electrophoresis (Fig. 2B, lanes 6–10).

A second set of 3′ deletion mutants was synthesized to determine whether the loop or stem sequences of the DIS contribute to the formation of the fast-migrating conformer (Fig. 3). We also included two mutants based on the 1/270 transcript, in which the DIS hairpin was stabilized by the introduction of 3′-terminal nucleotides

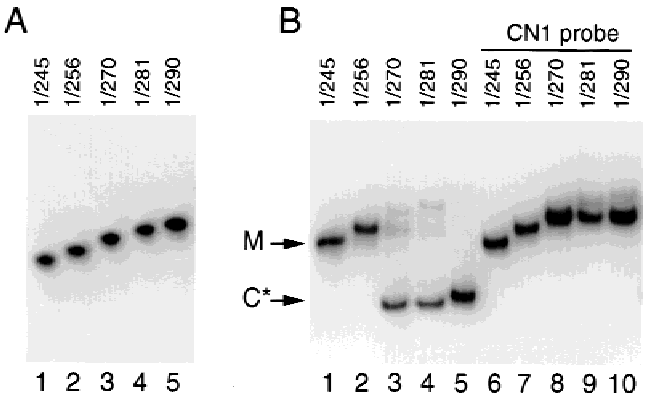

FIGURE 2. Deletion mapping of 3′ terminal sequences required for formation of the fast-migrating RNA species. Transcripts are indicated at the top of each lane. **A**: Transcripts as visualized on a denaturing sequencing gel. **B**: Nondenaturing gel electrophoresis of transcripts renatured in TEN buffer (lanes 1–5) and annealed to the CN1 primer (lanes 6–10). Inclusion of the DIS stem-loop (transcript 1/270, lane 3) leads to the formation of the fast-migrating conformer (marked C*). Transcripts unable to form the conformer are marked M, including transcripts denatured by primer CN1.

that increase the number of consecutive base pairs in the DIS stem (Fig. 3B, 270+4 and 270+8). Efficient formation of the conformer is apparent for the 1/267 transcript and all longer transcripts (Fig. 3A). As expected, no conformer is observed for the 1/256 transcript, but transcript 1/263 displays an intermediate electrophoretic mobility. Transcript 1/263 produces a smear on nondenaturing gels that migrates between the position of the 1/267 conformer and the 1/256 transcript (Fig. 3A, compare lane 3 to lanes 1 and 5). We conclude that inclusion of nt 257–263 at the 3′ end of HIV-1 leader transcripts introduces the minimally required sequence for formation of the fast-migrating conformer, although downstream nucleotides do significantly stabilize this RNA conformation. This minimally required sequence corresponds to the loop of the DIS hairpin (257-GCGCGCA-263). Both mutant transcripts 1/270+4 and 1/270+8 contain the minimal sequence required for efficient formation of the fast-migrating conformer (1/270), but the high-mobility band appears very faint for these transcripts (Fig. 3A, lanes 11 and 13). Thus, further stabilization of the DIS hairpin interferes with the folding of the conformer. We expect that the GCGCGC palindrome is properly exposed in the loop of these stabilized hairpins, which argues against conformer formation by means of an intramolecular loop-loop or triple helix interaction involving the DIS loop.

The loop and upper stem sequences of the poly(A) hairpin are required for conformer formation

With the minimally required 3′ sequences now established, we proceeded to test an extensive set of HIV-1 leader transcripts with 5′-terminal and internal mutations for formation of the fast-migrating conformer. Mutations in the central PBS domain had little effect on formation of the conformer (not shown), which is consistent with the results of antisense probing experiments (Berkhout & van Wamel, 2000). The 5′ TAR domain does contribute to conformer formation, but a low level of conformer can be formed in the absence of TAR, and the partial defect of TAR mutants can be rescued by altering salt conditions (H. Huthoff & B. Berkhout, in prep.). Based on these results, we focused our analysis on the domain between TAR and the PBS, which contains the poly(A) hairpin. Figure 4C shows the secondary structure models of the wild-type and mutant poly(A) hairpins used in this study. Mutant A contains a stabilized poly(A) hairpin due to two bulge deletions and one nucleotide substitution. The poly(A) hairpin has been destabilized in mutants B, C, and D. Destabilization of the poly(A) hairpin is modest for mutant B, which has four nucleotide substitutions. Destabilization is more severe in mutants C and D because of the disruption of the lower stem of the hairpin. The base pairing capacity is restored in the double mutant CD as the mutations in C and D are complementary. In mutant E, the upper stem and loop domain of the poly(A) hairpin has been deleted. We also included a transcript corresponding to the sequence of a revertant virus, which was obtained after prolonged culturing of mutant A (Berkhout et al., 1997). Figure 4A shows a nondenaturing gel analysis of 1/290 transcripts containing these mutations. No conformer is formed by mutants A and E, and slight smearing is observed for mutants B and CD. Mutants C and D form the fast-migrating conformer at wild-type efficiency, indicating that the lower stem region of the poly(A) hairpin is not required for its formation. The inability of mutant E to form the conformer demonstrates that important sequences for the formation of the conformer must reside in the loop or upper stem of the poly(A) hairpin.

To investigate whether the inability of mutant A to form the conformer is caused by mutation of the primary sequence or by stabilization of the poly(A) hairpin, we analyzed the A4 revertant (Fig. 4B). This revertant maintains the original mutations, but has reduced the stability of the poly(A) hairpin by acquisition of additional mutations that partially disrupt the stem (Fig. 4C). A significant amount of the A4 revertant transcripts form the fast-migrating conformer, indicating that the stability of the poly(A) hairpin affects formation of the conformer (Fig. 4B, lane 5). The altered stability of the poly(A) hairpin of mutants B and CD could also explain the smearing observed for these mutants. Whereas the overall stability of the poly(A) hairpin in mutant B is decreased, the upper stem is in fact stabilized by an additional G-U base pair. Mutant CD is also more stable than the wild-type hairpin and produces a smear on the gel. Interestingly, both mutants C and D show no

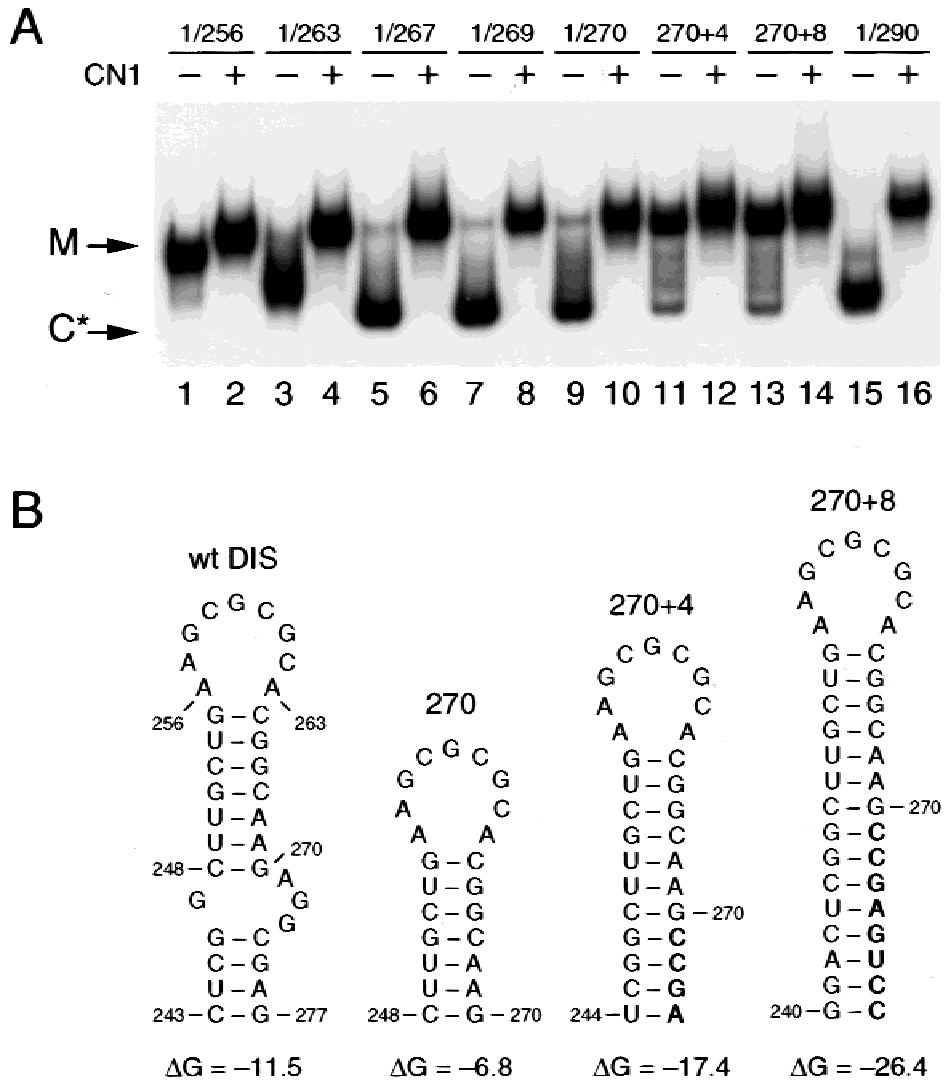

FIGURE 3. Fine mapping of DIS sequence elements required for formation of the fast-migrating conformer. **A**: Nondenaturing gel analysis of transcripts with 5′ and 3′ extremities as indicated at the top of the lanes. Presence and absence of the denaturing CN1 primer is indicated by + or −, respectively. C* marks the position of the conformer. Transcripts unable to form the conformer are marked M, including transcripts denatured by primer CN1. Two mutant transcripts are included in which the helical stem of the DIS stem-loop structure has been extended (270+4, lanes 11 and 12, and 270+8, lanes 13 and 14). Formation of the fast-migrating conformer is severely impaired for these mutants, although they do contain the minimally required sequence for the conformer. **B**: Secondary structure models of the wild-type and mutant DIS elements with the predicted thermodynamic stability (kcal/mol at $T = 37\,^{\circ}C$) indicated below. Mutants 270+4 and 270+8 contain additional sequences 3′ to position 270 that extend the DIS helical stem.

such defects, again suggesting that the defect of mutant CD is related to the stability of the poly(A) hairpin. This indicates that increased stability of the poly(A) hairpin results in a reduced ability to form the conformer, and this is reminiscent of the results obtained with the stabilized DIS hairpin.

Characterization of the long-distance poly(A)–DIS interaction

With the two major domains involved in the formation of the conformer determined by mutational analysis, we aimed to characterize the structural nature of this RNA conformation. Several modes of long-distance interactions between the poly(A) and DIS domains are possible. The HIV-leader may adopt an intricate tertiary structure involving non-Watson–Crick interactions. Alternatively, the conformer may form through canonical base pairing of complementary sequences in the poly(A) and DIS domains. Such a long-distance base-pairing interaction may be incompatible with formation of the individual hairpin motifs or it may consist of a loop-loop interaction between the poly(A) and DIS hairpins. However, the observation that stabilization of the poly(A) and DIS stems inhibits formation of the conformer argues against the possibility of a loop-loop

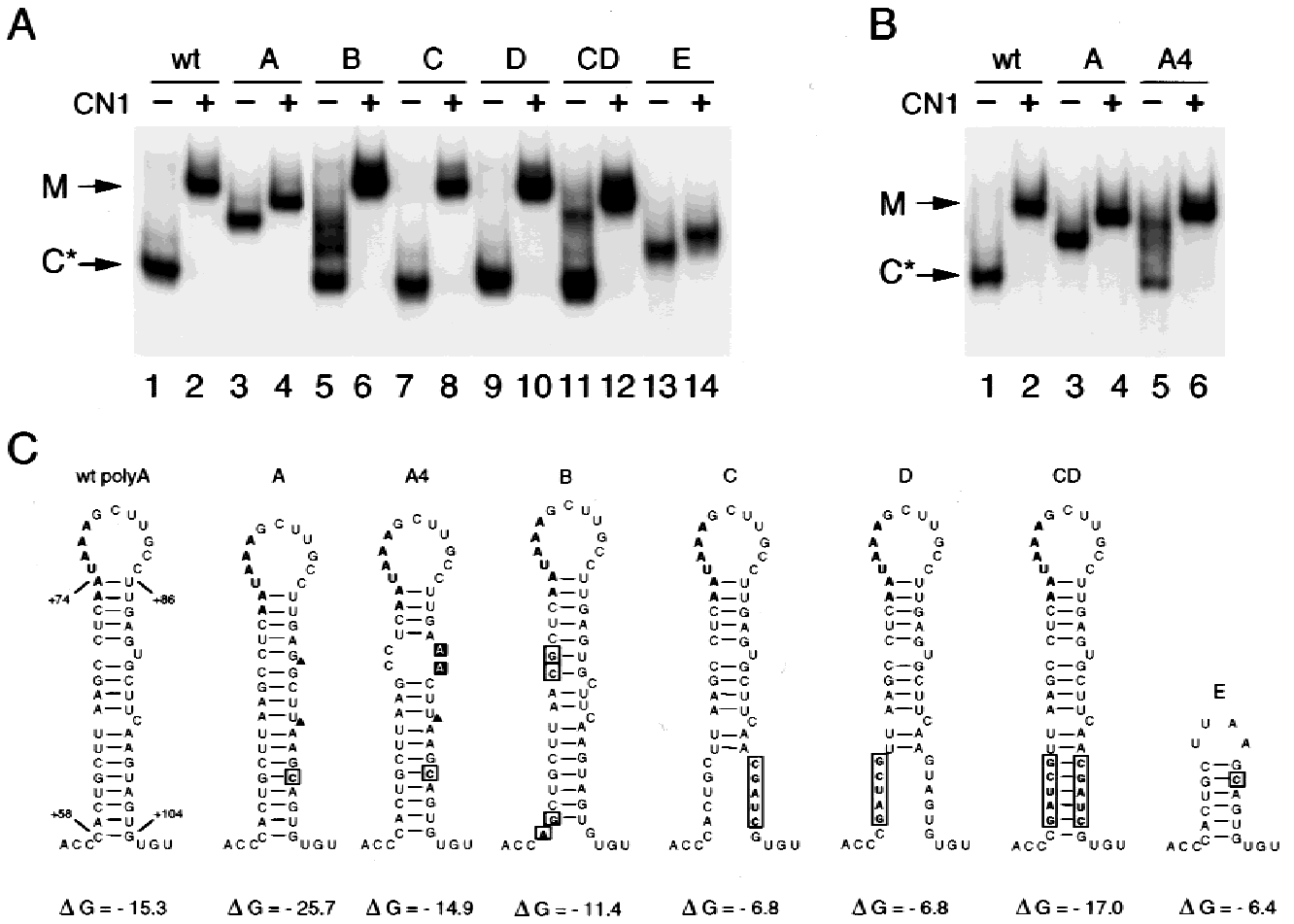

FIGURE 4. Analysis of transcripts with wild-type and mutant poly(A) hairpin structures. **A**: Wild-type and mutant transcripts (all 1/290) as analyzed on a nondenaturing gel with (+) and without (−) the CN1 primer. C* marks the position of the conformer. Transcripts unable to form the conformer are marked M, including transcripts denatured by primer CN1. **B**: Analysis of wild-type (wt), A-mutant, and A4 revertant transcripts on a nondenaturing gel. The A4 viral sequence was obtained by reversion of the mutant A virus. Formation of the fast-migrating RNA species is partially restored in the A4 revertant RNA. **C**: Secondary structure models of the wild-type and mutant poly(A) hairpins with the predicted thermodynamic stability (kcal/mol at $T = 37\,^{\circ}C$) indicated below. The AAUAAA hexamer that forms the polyadenylation signal is indicated in bold. Mutated nucleotides are boxed and deletions are indicated by ▲. Reversion mutations in A4 are indicated by black boxes.

interaction. We have employed a combination of biochemical and biophysical methods and computer-assisted RNA folding to distinguish between these possibilities.

To investigate the secondary structure of the wild-type leader RNA, we performed an RNA structure prediction analysis with the Mfold 3.0 algorithm (Mathews et al., 1999; Zuker et al., 1999). A single Mfold analysis produces a set of solutions for each RNA sequence, which are ranked by their corresponding free-energy values. Interestingly, the lowest energy solution for the HIV-1 leader RNA predicts base pairing of the GCGCGC palindrome to nt 88–93 of the poly(A) region, and this results in the loss of the poly(A) and DIS hairpins. This correlates strikingly with the results of the mutational analysis, wherein the poly(A) and DIS domains are critical for formation of the conformer. The predicted overall structure is an elongated rodlike fold that contains the intact TAR element and an extended stem interrupted by internal loops (Fig. 5). The poly(A) and DIS hairpins are predicted only by suboptimal solutions of the Mfold algorithm, producing a branched RNA structure similar to a previously proposed multiple-hairpin model of the HIV-1 leader (Berkhout, 1996). Participation of the poly(A) and DIS domains in either long-distance base pairing or formation of their respective hairpins is the sole difference between the rodlike structure and the branched structure. Structure predictions for mutant transcripts that are unable to form the conformer, such as mutant A and 1/270+4, consistently favored the branched fold with the individual hairpins over the rodlike fold.

To assess the validity of the structure models generated by Mfold, we performed chemical RNA structure

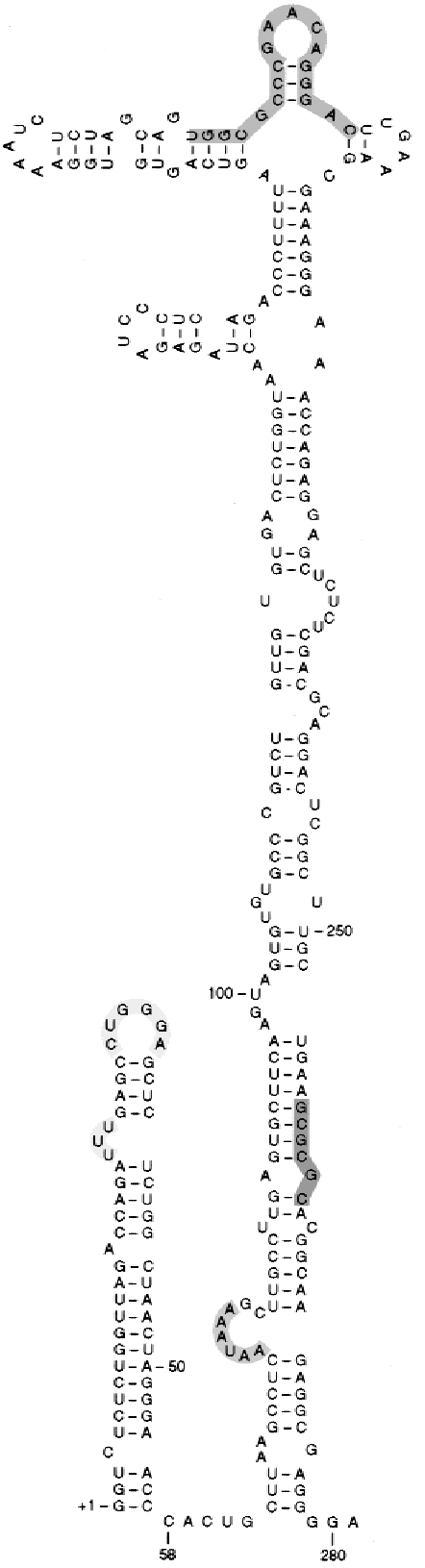

FIGURE 5. Secondary structure model of the HIV-1 leader RNA. This model represents the lowest energy solution provided by the Mfold 3.0 algorithm and is supported by several experimental findings (see the text). This structure predicts the base pairing of the poly(A) to the DIS domain. Several structure and sequence motifs are highlighted by shading: the bulge and loop of TAR, the AAUAAA polyadenylation signal, the PBS, and the DIS palindrome.

probing analyses. Both the wild-type and mutant A transcripts were used to compare the chemical accessibility of the conformer versus a conformer-impaired transcript (Fig. 6). The DEPC and DMS modification patterns of the wild-type transcript are in good agreement with the rodlike structure predicted by Mfold (Fig. 7A). The AAUAAA polyadenylation signal is predicted to be entirely exposed in an internal loop in the wild-type rodlike structure. Indeed, all A residues of this polyadenylation signal are highly reactive towards DEPC and DMS (Fig. 6A). The modification pattern of mutant A showed several differences compared with the wild-type transcript. These differences are located exclusively in the poly(A) and DIS domains (not shown), reinforcing the model of a long-distance interaction between these domains. Within the poly(A) domain, A73 and A74 of the polyadenylation signal show no reactivity in mutant A. Indeed, the branched structure predicts base pairing of these residues in the poly(A) hairpin. Increased reactivity in the wild-type compared with the mutant A transcript was also observed for nt A66 and A67. These residues are in an internal loop in the wild-type rodlike structure, whereas they are base paired within the poly(A) hairpin for mutant A (Fig. 7A,B).

A more subtle difference in the modification pattern is observed within the DIS domain (Fig. 6B). The susceptibility of A263 to DMS modification is significantly higher for mutant A than the wild-type transcript. In addition, we measured an elevated level of reactivity at C260 and C262 of the GCGCGC palindrome for mutant A. These nucleotides are predicted to be loop exposed in the DIS hairpin for mutant A. Thus, the mutations introduced in the poly(A) domain cause structural alterations some 150 nt further downstream in the DIS domain. We note that the distribution of RT pausing products in the control reactions with untreated transcripts also differs significantly for the wild-type and mutant transcript in the vicinity of the DIS. For example, pausing is dramatically increased at A239 and decreased at A268 for mutant A (Fig. 6B, compare lanes 1 and 3). We have previously reported that structure-induced pausing of RT in the poly(A) domain correlates with hairpin stability (Klasens et al., 1999), but this long-distance effect has gone hitherto unnoticed. These combined results indicate that the mutant A folds the branched multihairpin structure and that the wild-type HIV-1 RNA adopts the rodlike structure.

The conformation of structured RNAs can also be monitored by absorbance melting measurements (Puglisi & Tinoco, 1989; Brion & Westhof, 1997). This method makes use of the hypochromicity of native RNA relative to its denatured state and allows for the determination of the thermal denaturation profile. The melting curves in a 20–85 °C trajectory are shown for the wild-type HIV-1 RNA and several conformer-impaired mutants as a $\delta A/\delta T$ plot, which allows the accurate determination of the melting transition and the T_m

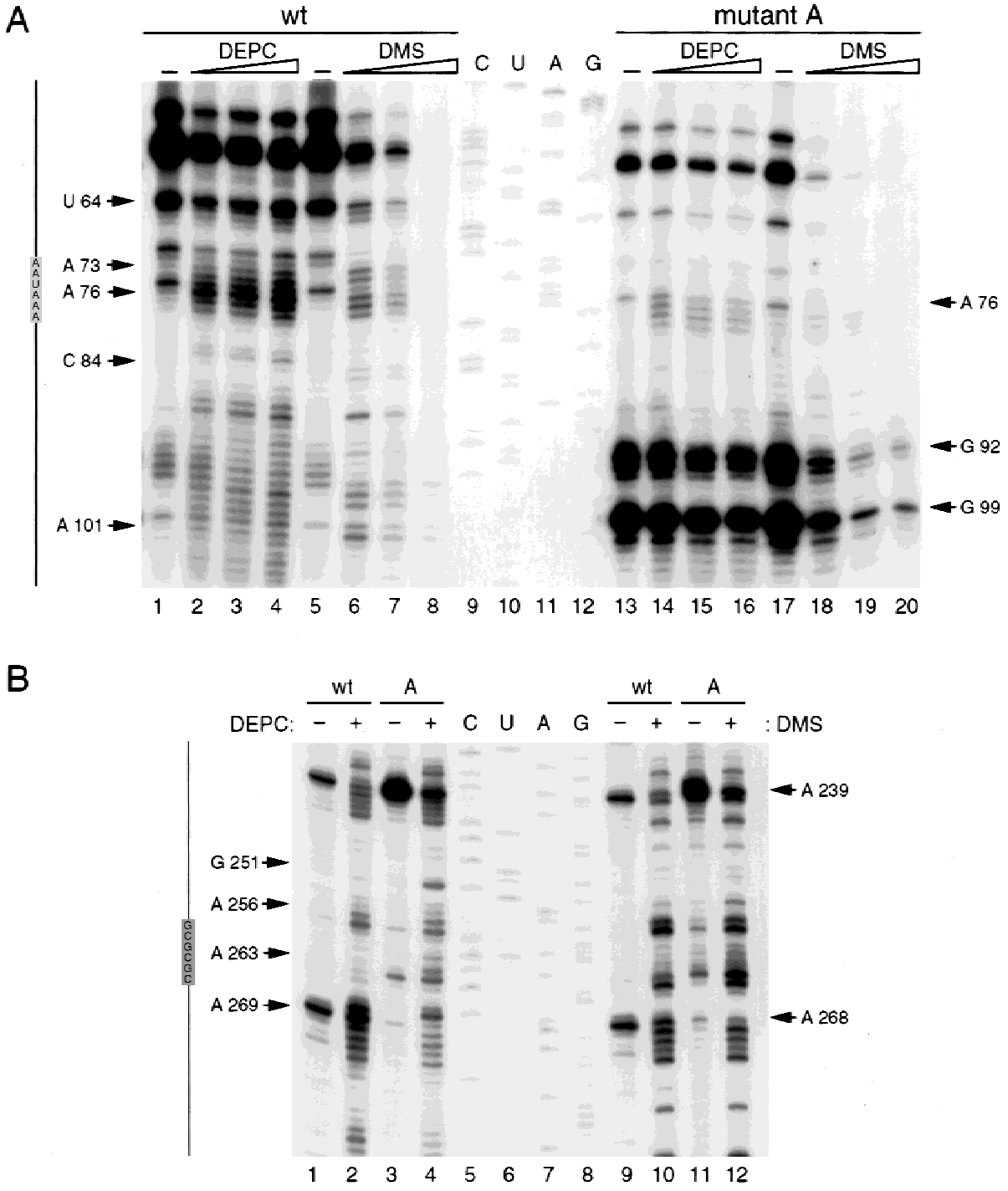

FIGURE 6. Chemical structure probing of wild-type and mutant A HIV-1 leader transcripts. Transcripts (1/368) were treated with limiting amounts of DEPC and DMS. **A**: Analysis of the poly(A) region. Modification sites were detected by primer extension with the Lys 21 primer. Lanes 1, 5, 13, and 17 represent the control samples with untreated transcripts. A sequence reaction is shown in lanes 9–12. To the left a linear representation of the transcript indicates the position of the AAUAAA polyadenylation signal. **B**: Analysis of the DIS region using the R368 primer in primer extension analysis. Only the modification reactions with the highest concentration of DEPC and DMS are shown (4 μL/100 μL reaction volume). Lanes 1, 3, 9, and 11 represent the control samples with untreated transcripts. A sequence reaction is shown in lanes 5–8. To the left, a linear representation of the transcript indicates the position of the DIS palindrome. Nucleotide positions that are discussed in the text are indicated by arrows.

(Fig. 8). No changes in absorbance were observed below 20 °C and above 85 °C. The wild-type RNA shows two discrete transitions in the melting profile at T_m = 49.1 °C and T_m = 66.6 °C (Fig. 8A). The separated peaks suggest the existence of two independently structured domains for the wild-type 1/290 transcript. The high-

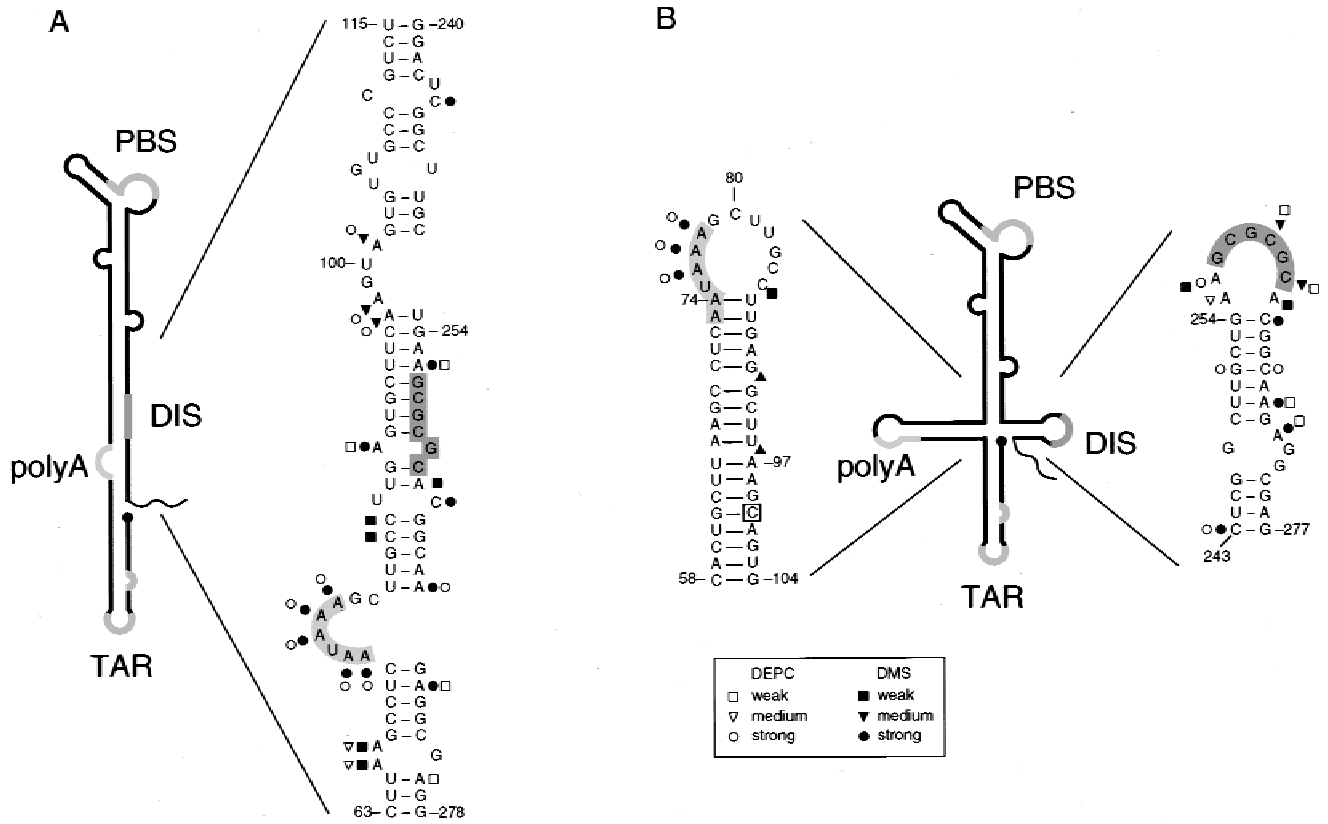

FIGURE 7. Secondary structure model for the wild-type and mutant A transcripts with a summary of the chemical modification data. Several structure and sequence motifs are highlighted by shading: the bulge and loop of TAR, the AAUAAA polyadenylation signal, the PBS, and the DIS palindrome. **A**: Schematic representation of the rodlike structure of the wild-type HIV-1 leader RNA. The duplex that is formed by base pairing of the poly(A) and DIS domains is shown in detail with chemical modification data. **B**: Schematic representation of the branched RNA conformation of the mutant A. Hairpin structures for poly(A) and DIS are highlighted and chemical modification data of mutant A are indicated. The nucleotide substitution in mutant A is indicated by a box. Note that ▲ indicates deleted nucleotides and does not represent chemical reactivity in the stem of the hairpin. Two nucleotides have been deleted in mutant A, but numbering of the nucleotides corresponds to the wild-type positions.

intensity transition at $T_m = 49.1\,^{\circ}C$ of the wild-type transcript is in agreement with the thermal stability of the conformer as assayed by gel electrophoresis (Berkhout & van Wamel, 2000), and this peak is lost in mutant transcripts unable to form the conformer. The loss of this transition for mutant A results in the appearance of several low intensity transitions in the 27–62 °C range (Fig. 8B). A somewhat similar pattern is observed for the transcript lacking TAR (58/290) and the transcript containing a stabilized DIS hairpin (1/270+4), although fewer additional transitions are observed (Fig. 8C,D).

The second transition in the wild-type melting profile at $T_m = 66.6\,^{\circ}C$ is likely to correspond to the stable TAR hairpin. The transcript lacking TAR (58/290) has no melting transitions above 57 °C, whereas both the poly(A) mutant A and DIS mutant 1/270+4 do contain this characteristic high-temperature transition. Interestingly, the TAR element appears to be of higher stability ($T_m = 72.6\,^{\circ}C$) in the context of the mutant A. It is possible that TAR gains stability by coaxial stacking interactions with the immediately adjacent poly(A) hairpin of mutant A (Berkhout, 1996). Additional stacking interactions have been reported to cause an increase in the T_m of melting transitions (LeCuyer & Crothers, 1993). This observation underlines the difficulties in interpreting and assigning transitions from absorbance melting measurements, and a more detailed analysis is currently in progress. Nevertheless, the biophysical data presented do suggest that the wild-type HIV-1 RNA conformer consists of two independently structured domains, one of which is likely to correspond to TAR. The extended long-distance base pairing between the poly(A) and DIS domains may well correspond to the high-intensity transition at $T_m = 49.1\,^{\circ}C$. The helical domains within this extended stem do not exceed eight consecutive base pairs, indicating that the extended structure is not excessively stable, that is, less stable than TAR. In general, the absorbance melting measurements are compatible with the proposed conformer structure.

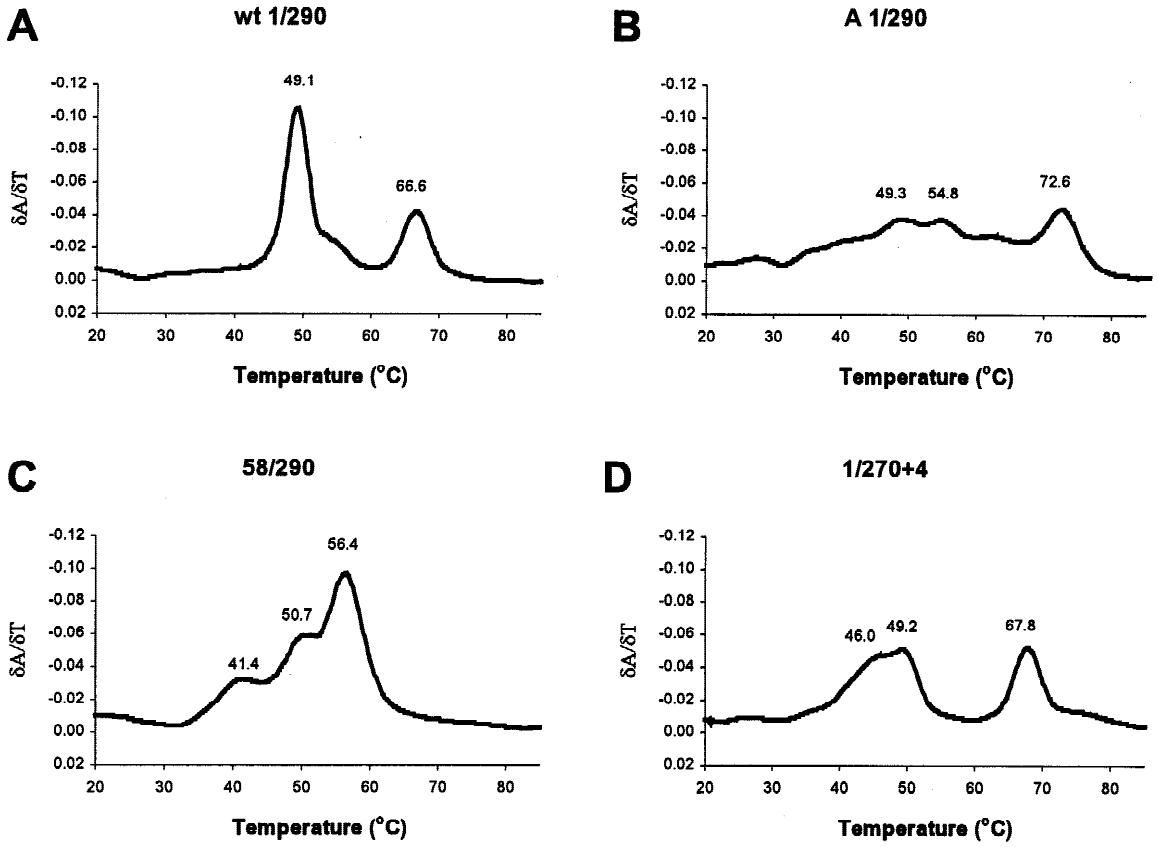

FIGURE 8. Thermal melting profiles of the wild-type HIV-1 leader RNA and mutant transcripts, plotted as $\delta A/\delta T$ curves. **A**: Wild-type 1/290. **B**: Poly(A) mutant A 1/290. **C**: Deletion of TAR, 58/290. **D**: Stabilized DIS hairpin 1/270+4. Numbers at the top of the peaks correspond to the temperature at the peak maximum.

The structures of the fast-migrating conformer and the HIV-1 RNA dimer are mutually exclusive

The results described above indicate that the DIS and poly(A) motifs are base paired to form the rodlike conformer. We next wondered whether this conformation is maintained during dimerization of HIV-1 leader transcripts. We performed in vitro dimerization assays with the wild-type 1/290 transcript and three mutant transcripts that are impaired in their ability to form the conformer: a TAR deletion mutant (10/290), the stabilized poly(A) mutant (A 1/290) and a DIS mutant (1/256). The latter transcript also serves as a negative control for RNA dimerization, as it terminates just upstream of the DIS palindrome. In addition to the dimerization conditions (Fig. 9, lanes 9–12), transcripts were dissolved in TE buffer (Fig. 9, lanes 1–4) or heat denatured in formamide before electrophoresis (Fig. 9, lanes 5–8).

One striking result from this analysis is that the conformer-impaired transcripts 10/290 and A 1/290 form dimers (marked D) much more efficiently than wild-type RNA. Even in the TE buffer, a substantial amount of dimers is observed for these mutants (Fig. 9, lanes 2 and 4), but not for the wild-type transcript (Fig. 9, lane 1). When these mutant transcripts are subjected to the in vitro dimerization conditions, approximately 90% of the RNA is dimeric (Fig. 9, lanes 10 and 12), whereas the wild-type transcript produces approximately 40% dimer (Fig. 9, lane 9). Furthermore, both mutant transcripts form substantial amounts of multimeric complexes. As indicated above, these transcripts are either fully (mutant A) or partially defective in conformer formation (10/290; note the presence of small amounts of C*, Fig. 9, lane 2). The previous analyses indicate that these mutant transcripts favor the branched fold, and their more efficient dimerization thus correlates with exposure of the DIS hairpin. The results indicate that the ability of the wild-type transcript to dimerize is hindered by the conformer structure, which masks the DIS region.

Another interesting observation comes from the comparison of the relative electrophoretic mobility of monomeric and dimeric transcripts. Whereas the wild-type conformer migrates faster through native gels than mutant transcripts, the wild-type dimer is retarded as compared to the mutant dimers (Fig. 9, compare lane 9 to lanes 10 and 12). In fact, the electrophoretic mobility of the wild-type and mutant dimers accurately mimics the electrophoretic mobility of the formamide-denatured transcripts (Fig. 9, lanes 6–10) and correlates precisely

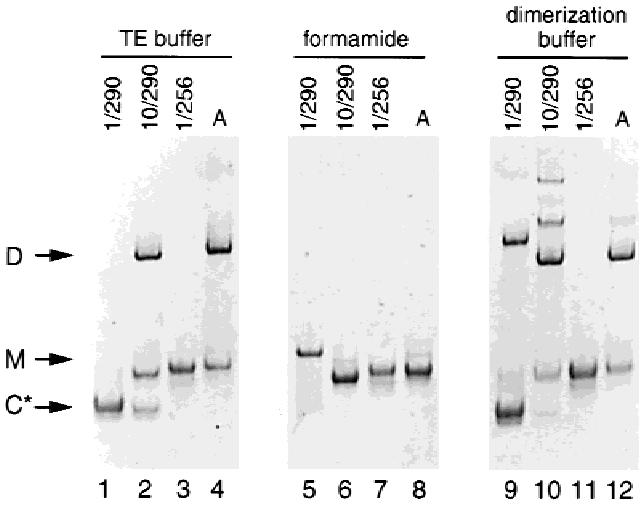

FIGURE 9. Dimerization of the wild-type and mutant HIV-1 transcripts. Nondenaturing gel analysis of the transcripts after incubation in TE buffer (lanes 1–4), denaturation in formamide (lanes 5–8), and incubation in dimerization buffer (lanes 9–12). The RNA was visualized by ethidium bromide staining of the gel. 5′ and 3′ extremities of the transcripts are indicated at the top of the lanes. The 1/290 transcript with the mutant A poly(A) hairpin is indicated by A (top of lanes 4, 8, and 12). D marks the position of HIV-1 leader RNA dimer, M marks monomeric transcripts not able to form the conformer and C* marks the position of conformer. Transcript 1/256 lacks the GCGCGC palindrome and can neither form dimers nor the conformer. Transcripts 10/290 and A are both impaired for formation of the fast-migrating conformer but form dimers very efficiently, even in the absence of cations (lanes 2 and 4). In addition, several multimeric forms are observed for these mutant transcripts (lanes 10 and 12).

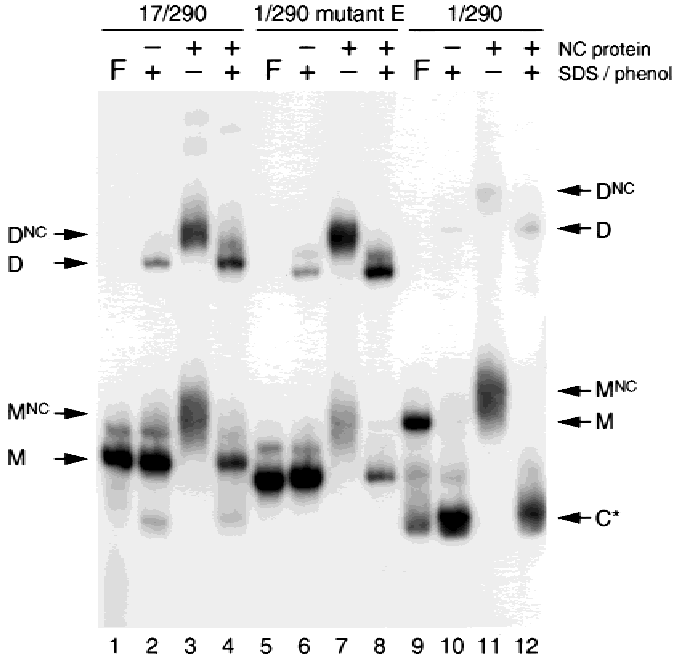

FIGURE 10. NC protein disrupts the conformer and facilitates dimerization. Analysis of the effect of NC protein (+/− indicated at the top of lanes) on dimerization of wild-type and mutant transcripts, the TAR deletion mutant (17/290), and the poly(A) deletion mutant (mutant E). Controls are included in which the transcript is denatured in formamide (lanes 1, 5, and 9). Transcripts were heat-dimerized without NC (lanes 2, 6, and 10), and subsequently treated with SDS/phenol to demonstrate the stability of the dimer (D) during protein removal. Transcripts were incubated with NC protein at 37 °C and aliquots were either loaded directly on the gel (lanes 3, 7, and 11) or treated with SDS/phenol prior to loading (lanes 4, 8, and 12). Coating by NC causes a shift in the position of the RNA, which is indicated by M^{NC} for the monomer and D^{NC} for the dimer. Coating by NC of the wild-type 1/290 transcript (lane 11), which forms the fast-migrating conformer (C*), leads to a dramatic shift above the position of the denatured control (lane 9).

with the differences in transcript length. Thus, the fast electrophoretic mobility typical of the conformer is lost upon RNA dimerization, indicating that the RNA structures that constitute the conformer and the dimer are mutually exclusive.

The viral NC protein disrupts the long-distance RNA interaction and promotes dimerization

A consequence of the mutual exclusiveness of the conformer and RNA dimer structures is that the wild-type HIV-1 leader RNA must undergo a structural rearrangement to form dimers. The viral NC protein is a nucleic acid chaperone (Herschlag et al., 1994) and a likely candidate for facilitating this conformational RNA switch in vivo. We investigated this possibility by assaying the effect of recombinant NC protein on in vitro dimerization (Fig. 10). NC stimulates dimerization of HIV-1 RNA at physiological conditions, whereas either high salt conditions or incubation at high temperatures is required in the absence of NC (Darlix et al., 1990; Muriaux et al., 1995). We compared wild-type and mutant transcripts for dimerization by heat treatment and incubation with NC (Fig. 10, lanes 2, 6, and 10 and lanes 3, 7, and 11, respectively). Two conformer-impaired transcripts were arbitrarily chosen as negative controls, the TAR mutant 17/290 and the poly(A) deletion mutant E. As expected, these transcripts display higher levels of heat-induced dimers compared with the wild-type RNA (Fig. 10, compare lanes 2 and 6 with lane 10). Dimerization is markedly more efficient in the presence of NC, and this stimulatory effect is observed both for the wild-type and mutant transcripts (Fig. 10, lanes 3, 7, and 11). Binding of the NC protein to RNA causes a shift of both monomeric (M) and dimeric (D) RNA bands upon electrophoresis. The HIV-1 RNA dimer has been reported to withstand removal of protein by SDS/phenol treatment (Muriaux et al., 1996). We have included such deproteinized samples in our analysis (Fig. 10, lanes 4, 8, and 12). As expected, the NC-induced bandshift is lost upon phenol extraction of the samples. The two mutant transcripts that do not form the conformer show a relatively minor NC bandshift. In contrast, the wild-type transcript that does form the conformer exhibits an extreme NC shift (Fig. 10, compare lanes 10 and 11), suggesting that there is an additive effect of NC binding and disruption of the fast-migrating conformer. Consis-

tent with this idea, the wild-type RNA–NC complex (Fig. 10, lane 11) runs just above the denatured control sample (Fig. 10, lane 9). After removal of NC, the fast-migrating band reappears (Fig. 10, lane 12). This is due to rapid refolding of the conformer, which is commonly observed for transcripts that are heat denatured in formamide (not shown). These combined results suggest that the NC protein disrupts the long-distance interaction of the wild-type conformer, thereby facilitating dimerization through exposure of the DIS hairpin.

DISCUSSION

Structure of the fast-migrating HIV-1 RNA conformer

We show that a previously identified conformer of the HIV-1 untranslated leader (Berkhout & van Wamel, 2000) is, in fact, the most stable conformation of this RNA. This conformer is formed through long-range base pairing that involves both 5′ and 3′ domains of the HIV-1 leader RNA. A mutational analysis identified the poly(A) and DIS domains as the main determinants for formation of the conformer. Mutations that stabilize the individual poly(A) and DIS hairpin motifs severely impaired formation of the conformer despite retaining the minimally required sequences. This indicates that the poly(A) and DIS hairpin structures are incompatible with formation of the conformer.

An interaction between the poly(A) and DIS domains was readily predicted by computer-assisted RNA folding of the HIV-LAI sequence. The lowest energy solutions from Mfold predict base pairing of the GCGCGC palindrome in the DIS domain to the poly(A) domain, resulting in the loss of the hairpin motifs and formation of a rodlike structure with an extended stem (Fig. 5). Similar structures were predicted for other HIV-1 subtype B isolates and viral isolates that belong to different HIV-1 subtypes (not shown). This structure is also in good agreement with chemical RNA probing experiments. A most notable result from this analysis is the high chemical accessibility of all A residues in the AAUAAA polyadenylation signal of wild-type transcripts. The rodlike structure predicts the presentation of the complete AAUAAA sequence in an internal loop. A mutant unable to form the conformer (poly(A) mutant A) due to stabilization of the poly(A) hairpin displayed no reactivity at the two most 5′ A residues of this sequence motif, which is consistent with the poly(A) hairpin structure. In addition, this mutant displayed an alternative chemical modification pattern within the DIS domain that is in good agreement with the branched multiple hairpin structure. Chemical modification in other domains of the HIV-1 leader RNA was unaffected for this mutant, underlining that the poly(A) and DIS domains interact. Formation of the long-distance poly(A)-DIS interaction is also compatible with absorbance melting measurements. We found that the wild-type HIV-1 leader undergoes two discrete melting transitions, which suggests the presence of two independently structured domains. Indeed, these transitions are selectively lost upon disruption of the conformer and deletion of TAR.

The HIV-1 leader RNA as a dynamic molecule and potential molecular switch

The finding that the poly(A) and DIS domains are involved in a long-range base pairing interaction seems to conflict with previous reports that support hairpin structures for these domains. The poly(A) hairpin structure is supported by phylogenetic analysis (Berkhout et al., 1995) and viral replication studies (Berkhout et al., 1997; Das et al., 1997). The DIS hairpin has been studied by NMR analysis (Mujeeb et al., 1998) and is also supported by phylogenetic (Berkhout, 1996) and biochemical analysis (Laughrea & Jette, 1994; Skripkin et al., 1994; Jossinet et al., 1999). One way to resolve this contradiction is by conceptualizing the HIV-1 leader as a conformationally polymorph and dynamic molecule. That is, the poly(A) and DIS domains may be able to switch between the long-distance base pairing and hairpin conformations (Fig. 11). In an extensive probing comparison of the HIV-1 RNA monomer and dimer, it has been reported that differences in modification reside in the poly(A) and DIS regions (Baudin et al., 1993). We find that the characteristic fast electrophoretic mobility of the HIV-1 leader RNA is lost upon formation of the RNA dimer, indicating a rearrangement of the conformer structure during dimerization. Mutants unable to form the conformer expose the poly(A) and DIS hairpins and dimerize very efficiently. This result is consistent with an important role of the DIS hairpin in dimerization (Skripkin et al., 1994; Mujeeb et al., 1998). Thus, disruption of the long-distance interaction between poly(A) and DIS may be a prerequisite for dimerization of wild-type HIV-1 RNA. Dimerization is strongly stimulated by the viral nucleocapsid protein NCp7 (Darlix et al., 1990), which possesses nucleic acid chaperoning activity (Herschlag et al., 1994). Binding of NC causes an extreme band shift for wild-type transcripts that fold the conformer structure. We have interpreted this result as being caused by an additive effect of NC coating of the RNA and a NC-induced structural rearrangement of the conformer. This NC-mediated conformational change coincides with increased dimerization of the HIV-1 leader RNA. These combined results suggest that NC enhances dimerization through stabilization of the branched multiple hairpin structure (Fig. 11B). Although RNA chaperones have been proposed to facilitate the search for the most stable folding (Herschlag, 1995), NC apparently favors the branched folding over the

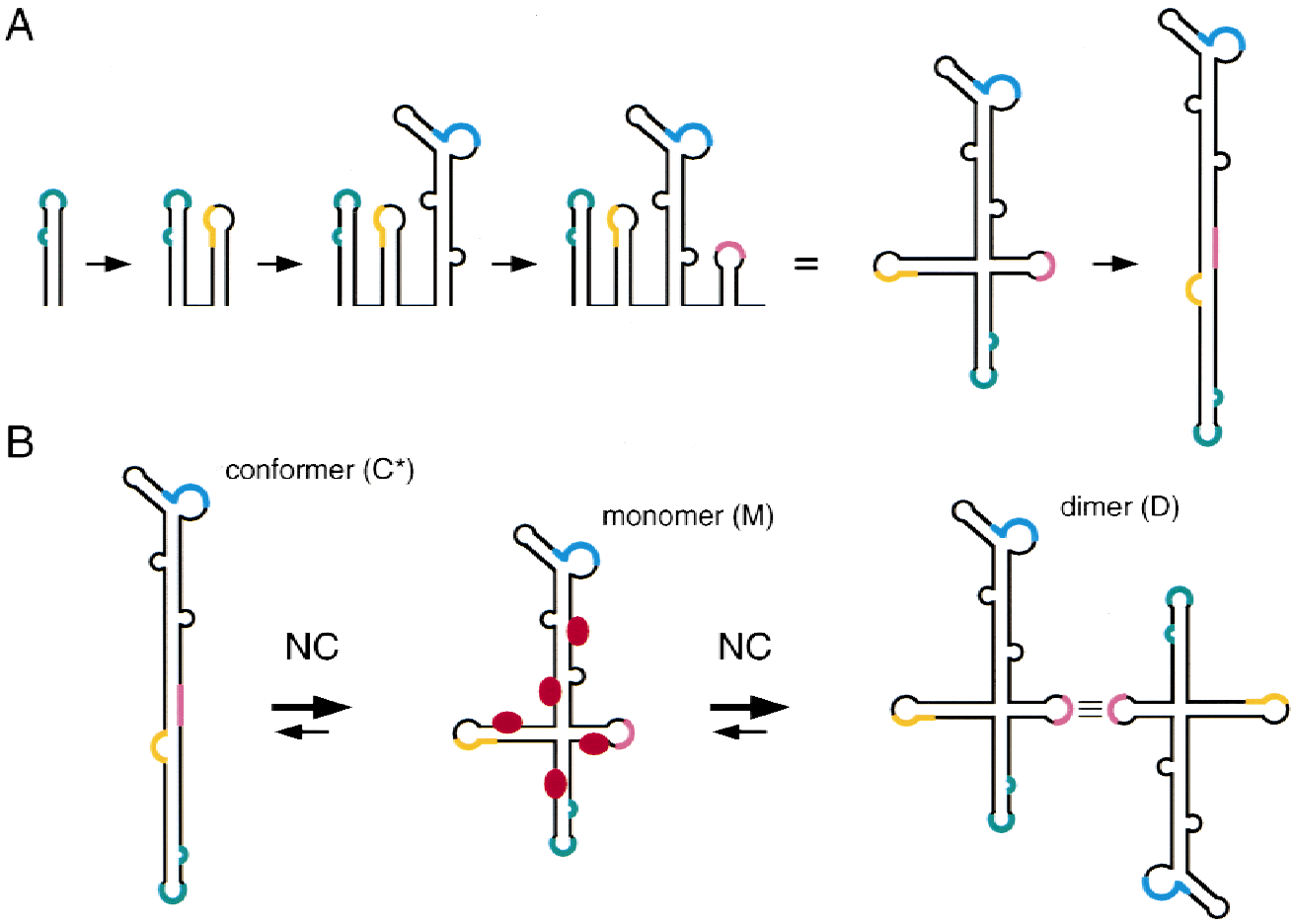

FIGURE 11. Alternating structures of the HIV-1 leader RNA. **A**: RNA folding during transcription of the HIV-1 leader sequence. Individual hairpin motifs are formed as kinetic intermediates before the poly(A) and DIS hairpins are disrupted to form the thermodynamically favored conformer. As part of the nascent transcript, the TAR and poly(A) hairpins play important roles in transcriptional activation (Berkhout et al., 1989) and down regulation of polyadenylation (Das et al., 1999). **B**: NC induced rearrangement of the HIV-1 leader RNA structure. Several structure and sequence motifs are highlighted by colored shading: the bulge and loop of TAR (green), the AAUAAA polyadenylation signal (orange), the PBS (blue), and the DIS palindrome (pink). The HIV-1 leader is in the most stable rodlike conformation in the absence of NC. Upon binding of NC (red ovals) the branched conformation with the poly(A) and DIS hairpins is favored, leading to dimerization through loop-loop kissing by the autocomplementary DIS loops.

thermodynamically most stable rodlike structure of the HIV-1 leader RNA.

The dynamic nature of the HIV-1 RNA conformation is compatible with previous viral replication studies that addressed the structure and function of the poly(A) hairpin (Berkhout et al., 1997; Das et al., 1997, 1999). These studies provide compelling evidence for the existence of the hairpin and demonstrate that the thermodynamic stability of the poly(A) hairpin is maintained within narrow limits. In particular, mutants with altered poly(A) hairpins were found to revert by acquisition of mutations that restore the stability of the hairpin to wild-type levels, but not necessarily by restoring the primary sequence. We included the revertant A4 of the A mutant with a stabilized poly(A) hairpin in our analysis. Although mutant A is incapable of forming the conformer, presumably because the hairpin conformation is favored, revertant A4 has partially restored the capacity to form the conformer. A dynamic model of the HIV-1 leader thus provides a rationale for the observed requirement of the poly(A) hairpin not to become excessively stable. Apparently, the poly(A) hairpin is maintained below a certain threshold of thermodynamic stability to allow the formation of the alternative conformation, that is, the long-distance interaction with the DIS. Preliminary data from computer-assisted RNA folding using the STAR algorithm indicate that the poly(A) hairpin is a metastable kinetic intermediate during folding of the nascent HIV-1 leader RNA (A.P. Gultyaev, pers. comm.; Fig. 11A). The existence of the poly(A) hairpin as a kinetic intermediate is of interest, as it has previously been demonstrated that this hairpin prevents polyadenylation at the 5′ end of nascent viral transcripts (Klasens et al., 1998).

The intrinsically different structures for the HIV-1 RNA rodlike monomer (conformer) and branched dimer (hair-

pins) creates the potential for the leader RNA to act as a regulatory switch during viral replication. The HIV-1 leader RNA contains a multitude of regulatory RNA signals that exert diverse effects during retroviral replication. In general, these RNA motifs are required for the regulation of either intracellular gene-expression functions or virion-associated functions such as packaging and reverse transcription. In some cases regulation by RNA structure can be assigned to discrete motifs, as is the case for TAR-dependent transcriptional activation of the LTR promoter (Muesing et al., 1985; Berkhout et al., 1989; Wei et al., 1998). On the other hand, mutation of regulatory RNA domains in the HIV-1 leader frequently causes a pleiotropy of defects for multiple processes during viral replication (Berkhout, 2000). One would expect this to occur when these regulatory RNA motifs can adopt different conformations that function in distinct processes. In addition, multipartite RNA signals are commonly encountered in the HIV-1 leader. A well-documented example is the HIV-1 packaging signal, which requires several upstream domains and downstream Gag coding sequences in addition to the core Ψ hairpin structure (Berkhout & van Wamel, 1996; McBride & Panganiban, 1996; Clever & Parslow, 1997; Das et al., 1997; Clever et al., 1999; Helga-Maria et al., 1999). This suggests that the overall structure of the HIV-1 leader RNA is crucial for selective recognition and packaging into virion particles. Furthermore, it has recently been described that binding of the Gag polyprotein to the Ψ site induces a structural rearrangement of the RNA (Zeffmandagger et al., 2000). The biology of HIV may necessitate distinct conformations of the leader RNA because of the dual role of HIV-1 transcripts as mRNA and viral genome. The structural transition from the HIV-1 RNA conformer into the branched structure may coincide with the down regulation of RNA signals that control gene expression and render the RNA in a conformation that allows dimerization, packaging, and, ultimately, reverse transcription.

MATERIALS AND METHODS

In vitro transcription

RNAs were in vitro transcribed from PCR-generated templates containing a T7 promoter directly upstream of the natural HIV-1 LAI $+1$ transcriptional start site or the indicated nucleotide positions. PCR was performed on the pBluescript 5′ LTR plasmid containing the wild-type or mutant 5′ LTR (Das et al., 1997). The T7 promoter sequence was present in the sense primers, thus introducing the T7 promoter at desired positions during PCR. PCR fragments were excised from agarose gels and purified using the QIAEX II DNA isolation system according to the manufacturer's instructions. Transcription was carried out using the Ambion megashortscript T7 transcription kit. Radiolabeled transcripts were synthesized in the presence of 1 μL α-^{32}P-UTP. Transcripts were subsequently excised from a 4% denaturing polyacrylamide gel (visualized by UV shadowing or autoradiography) and eluted from the gel fragment by overnight incubation in 0.3 M sodium acetate buffer (pH 5.2) at room temperature. The RNA was then precipitated and dissolved in water. Quantification of the RNA was done by UV-absorbance measurements and scintillation counting in case of radiolabeled transcripts.

In vitro RNA studies and nondenaturing electrophoresis

Purified transcripts in water were heated at 85 °C and allowed to slowly cool to room temperature. The stock solutions of RNA were stored at -20 °C and aliquots were taken for subsequent analysis. In conformer assays, approximately 10 ng of radiolabeled RNA was incubated in a final volume of 10 μL TEN buffer (0.1 M NaCl, 10 mM Tris-HCl, pH 7.5, and 1 mM EDTA, pH 8.0) at 37 °C for 30 min. Annealing of antisense DNA was also done in TEN buffer by heating at 60 °C and subsequent cooling to room temperature. We used 100 ng of the antisense oligo CN1 (5′-GGTCTGAGGGATCTCTAGTT ACCAGAGTC-3′), which is complementary to nt 123–151 of the HIV-1 leader RNA and was selected from a set of antisense DNA oligos that effectively denature the RNA conformer (Berkhout & van Wamel, 2000).

Dimerization assays were carried out in a final volume of 10 μL dimerization buffer (40 mM NaCl, 0.1 mM $MgCl_2$, and 10 mM Tris-HCl, pH 7.5). Approximately 1 μg of RNA was incubated at 60 °C for 30 min and subsequently allowed to cool to room temperature. Dimerization in the presence of the NCp7 protein was carried out at 37 °C for 30 min using 0.5 μg radiolabeled RNA. Dimerization is less efficient in this assay using radiolabeled RNA compared to dimerization with relatively high concentrations of unlabeled RNA. In vitro dimerization is dependent on RNA concentration (Berkhout et al., 1993). NC was present at a concentration of 12.5 μM, corresponding to a molar ratio RNA:NC of 1:250. NC protein was removed in some samples by addition of 2 μL 5% SDS and subsequent PCI extraction; these manipulations were carried out on ice. Reaction samples were chilled on ice and diluted with 5 μL loading buffer (30% glycerol with BFB dye). Heat denaturation of control samples was carried out in formamide loading buffer by heating at 85 °C. These samples were loaded on the gel immediately, because of rapid refolding of the conformer (not shown). Samples were analyzed on a 4% polyacrylamide gel with 0.25$\times$ TBE in the gel and running buffer. Native gels were run at 150 V at room temperature. Some samples were also analyzed on 6% sequencing gels containing urea.

RNA secondary structure prediction

Computer-assisted RNA secondary structure prediction was done using the Mfold version 3.0 algorithm (Mathews et al., 1999; Zuker et al., 1999) offered by the MBCMR Mfold server (http://mfold.burnet.edu.au/). Settings were standard for all folding jobs, corresponding to conditions at 37 °C and 1.0 M NaCl, using a 5% suboptimality range. The sequence of the HIV-1 B LAI isolate was downloaded from the HIV database (http://hiv-web.lanl.gov/). Folding jobs were submitted on sequences starting at the $+1$ transcriptional start site up to the

gag AUG start codon and with further 3′ extensions up to 3 kb. Furthermore, all mutant and truncated RNA species used in this study were analyzed by Mfold.

RNA absorbance melting curves

RNA thermal denaturation was monitored by measuring the absorbance of UV light at 260 nm in a quartz cuvette with a standard 1-cm path length. Absorbance was measured over a continuous temperature range from 10 to 95 °C on a temperature controlled Perkin Elmer Lambda 2 spectrophotometer. Temperature was increased at a rate of 0.5 °C/min with sampling at each 0.1 °C. Samples contained approximately 3.0 μg of RNA dissolved in 140 μL of 50 mM Na cacodylate buffer (pH 7.2). Prior to the measurement, the RNA was renatured by incubation at 37 °C for 30 min. Electrophoretic analysis of the samples prior to the analysis confirmed the monomeric state of the RNA (not shown). No significant degradation of the transcripts was apparent after the measurement as judged by analysis of the RNA on a sequencing gel (not shown).

NC protein

Recombinant NCp7 (MN isolate, amino acids 1–55, MQRGN FRNQRKIIKCFNCGKEGHIAKNCRAPRKRGCWKCGKEGH QMKDCTERQAN) was a kind gift from Dr. L. Henderson (National Cancer Institute, USA). Stock solutions of 1 μg/μL in NC buffer (60 mM NaCl, 30 mM $ZnCl_2$, 5.0 mM DTT, and 20 mM Tris-HCl, pH 7.5) were stored at −80 °C. Aliquots were taken and diluted with the NC buffer for use in dimerization assays and stored at −20 °C.

Chemical RNA structure probing

Probing assays were carried out with 0.5 μg of the target RNA (1–368) and 1.5 μg of carrier antisense human 18S ribosomal RNA (transcribed from the Ambion control template). The RNA was dissolved in 100 μL TEN buffer. We then added 0, 1, 2, or 4 μL of DEPC or DMS and the reaction mixture incubated at 20 °C for 15 min. Reactions were stopped by ethanol precipitation of the RNA. The RNA was dissolved in 10 μL water of which 2 μL was used in subsequent primer extension analysis. Primer extension was done by using ^{32}P-end-labeled primers and AMV RT. Primers Lys21 (5′-CAAGTCCCTGTTCGGGCGCCA-3′) and R:A368-A347 (5′-TCCCCCGCTTAATACTGACGCT-3′) were end-labeled with T4 PNK kinase. Approximately 2 ng of the labeled probe was annealed to the target RNA by heating to 60 °C and subsequent cooling to room temperature in hybridization buffer (83 mM Tris-HCl, pH 7.5, and 125 mM KCl) in a total volume of 10 μL. Primer extension reactions were initiated by addition of 5 μL 3× concentrated RT buffer (9 mM $MgCl_2$, 30 mM DTT, 30 μM of each dNTP, and actinomycin D at 150 μg/mL) and 0.5 μL of AMV reverse transcriptase (10 U). The reaction was performed at 43 °C for 15 min and stopped by precipitation. As a size marker, a sequencing reaction on the pBluescript 5′ LTR plasmid was performed in the presence of ^{35}S-dATP and the same primers as in the reverse transcription reaction. Sequencing was performed with the USB T7 sequenase kit according to the manufacturer's instructions. The samples were analyzed on 6% sequencing gels.

ACKNOWLEDGMENTS

We thank Jord Nagel and Cees Pleij for access to and assistance with the temperature-controlled spectrophotometer and Damian Purcell for providing an excellent Mfold server. We also thank Wim van Est for the artwork and Koen Verhoef and Nancy Beerens for fruitful discussions and critical reading of the manuscript. The NC protein was a kind gift from Dr. L. Henderson. This work was funded by the Netherlands Organisation for Scientific Research (NWO-CW).

Received September 15, 2000; returned for revision October 2, 2000; revised manuscript received October 26, 2000

REFERENCES

Aboul-ela F, Karn J, Varani G. 1996. Structure of HIV-1 TAR RNA in the absence of ligands reveals a novel conformation of the trinucleotide bulge. *Nucleic Acids Res 24*:3974–3981.

Amarasinghe GK, De Guzman RN, Turner RB, Summers MF. 2000. NMR structure of stem-loop SL2 of the HIV-1 RNA packaging signal reveals a novel A-U-A base-triple platform. *J Mol Biol 299*:145–156.

Baudin F, Marquet R, Isel C, Darlix JL, Ehresmann B, Ehresmann C. 1993. Functional sites in the 5′ region of human immunodeficiency virus type 1 RNA form defined structural domains. *J Mol Biol 229*:382–397.

Bender W, Davidson N. 1976. Mapping of poly(A) sequences in the electron microscope reveals unusual structure of type C oncornavirus RNA molecules. *Cell 7*:595–607.

Berkhout B. 1996. Structure and function of the Human Immunodeficiency Virus leader RNA. *Prog Nucleic Acid Res Mol Biol 54*:1–34.

Berkhout B. 1997. The primer-binding site on the RNA genome of human and simian immunodeficiency viruses is flanked by an upstream hairpin structure. *Nucleic Acids Res 25*:4013–4017.

Berkhout B. 2000. Multiple biological roles associated with the repeat (R) region of the HIV-1 RNA genome. *Adv Pharmacol 48*:29–73.

Berkhout B, Klaver B, Das AT. 1995. A conserved hairpin structure predicted for the poly(A) signal of human and simian immunodeficiency viruses. *Virology 207*:276–281.

Berkhout B, Klaver B, Das AT. 1997. Forced evolution of a regulatory RNA helix in the HIV-1 genome. *Nucleic Acids Res 25*:940–947.

Berkhout B, Oude Essink BB, Schoneveld I. 1993. In vitro dimerization of HIV-2 leader RNA in the absence of PuGGAPuA motifs. *FASEB J 7*:181–187.

Berkhout B, Silverman RH, Jeang KT. 1989. Tat trans-activates the human immunodeficiency virus through a nascent RNA target. *Cell 59*:273–282.

Berkhout B, van Wamel JLB. 1996. Role of the DIS hairpin in replication of human immunodeficiency virus type 1. *J Virol 70*:6723–6732.

Berkhout B, van Wamel JLB. 2000. The leader of the HIV-1 RNA genome forms a compactly folded tertiary structure. *RNA 6*:282–295.

Brion P, Westhof E. 1997. Hierarchy and dynamics of RNA folding. *Annu Rev Biophys Biomol Struct 26*:113–137.

Clever JL, Eckstein DA, Parslow TG. 1999. Genetic dissociation of the encapsidation and reverse transcription functions in the 5′R region of human immunodeficiency virus type 1. *J Virol 73*:101–109.

Clever JL, Parslow TG. 1997. Mutant human immunodeficiency virus type 1 genomes with defects in RNA dimerization or encapsidation. *J Virol 71*:3407–3414.

Damgaard CK, Dyhr-Mikkelsen H, Kjems J. 1998. Mapping the RNA binding sites for human immunodeficiency virus type-1 Gag and

NC proteins within the complete HIV-1 and -2 untranslated leader regions. *Nucleic Acids Res 26*:3667–3676.

Darlix JL, Gabus C, Nugeyre MT, Clavel F, Barre-Sinoussi F. 1990. *Cis* elements and *trans*-acting factors involved in the RNA dimerization of the human immunodeficiency virus HIV-1. *J Mol Biol 216*:689–699.

Das AT, Klaver B, Berkhout B. 1999. A hairpin structure in the R region of the Human Immunodeficiency Virus type 1 RNA genome is instrumental in polyadenylation site selection. *J Virol 73*:81–91.

Das AT, Klaver B, Klasens BIF, van Wamel JLB, Berkhout B. 1997. A conserved hairpin motif in the R-U5 region of the human immunodeficiency virus type 1 RNA genome is essential for replication. *J Virol 71*:2346–2356.

Fisher AG, Feinberg MB, Josephs SF, Harper ME, Marselle LM, Reyes G, Gonda MA, Aldovini A, Debouck C, Gallo RC. 1986. The *trans*-activator gene of HTLV-III is essential for virus replication. *Nature 320*:367–371.

Girard F, Barbault F, Gouyette C, Huynh-Dinh T, Paoletti J, Lancelot G. 1999. Dimer initiation sequence of HIV-1 LAI genomic RNA: NMR solution structure of the extended duplex. *J Biomol Struct Dyn 16*:1145–1157.

Helga-Maria C, Hammarskjold ML, Rekosh D. 1999. An intact TAR element and cytoplasmic localization are necessary for efficient packaging of human immunodeficiency virus type 1 genomic RNA. *J Virol 73*:4127–4135.

Herschlag D. 1995. RNA chaperones and the RNA folding problem. *J Biol Chem 270*:20871–20874.

Herschlag D, Khosla M, Tsuchihashi Z, Karpel RL. 1994. An RNA chaperone activity of non-specific RNA binding proteins in hammerhead ribozyme catalysis. *EMBO J 13*:2913–2924.

Isel C, Ehresmann C, Keith G, Ehresmann B, Marquet R. 1995. Initiation of reverse transcription of HIV-1: Secondary structure of the HIV-1 RNA/tRNA(3Lys) template/primer. *J Mol Biol 247*:236–250.

Jossinet F, Paillart JC, Westhof E, Hermann T, Skripkin E, Lodmell JS, Ehresmann C, Ehresmann B, Marquet R. 1999. Dimerization of HIV-1 genomic RNA of subtypes A and B: RNA loop structure and magnesium binding. *RNA 5*:1222–1234.

Klasens BIF, Das AT, Berkhout B. 1998. Inhibition of polyadenylation by stable RNA secondary structure. *Nucleic Acids Res 26*:1870–1876.

Klasens BIF, Huthoff HT, Das AT, Jeeninga RE, Berkhout B. 1999. The effect of template RNA structure on elongation by HIV-1 reverse transcriptase. *Biochim Biophys Acta 1444*:355–370.

Laughrea M, Jette L. 1994. A 19-nucleotide sequence upstream of the 5′ major splice donor is part of the dimerization domain of human immunodeficiency virus 1 genomic RNA. *Biochemistry 33*:13464–13474.

LeCuyer KA, Crothers DM. 1993. The *leptomonas collosoma* spliced leader RNA can switch between two alternate structural forms. *Biochemistry 32*:5301–5311.

Mathews DH, Sabina J, Zuker M, Turner DH. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J Mol Biol 288*:911–940.

McBride MS, Panganiban AT. 1996. The human immunodeficiency virus type 1 encapsidation site is a multipartite RNA element composed of functional hairpin structures. *J Virol 70*:2963–2973.

Muesing MA, Smith DH, Cabradilla CD, Benton CV, Lasky LA, Capon DJ. 1985. Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovirus. *Nature 313*:450–458.

Mujeeb A, Clever JL, Billeci TM, James TL, Parslow TG. 1998. Structure of the dimer initiation complex of HIV-1 genomic RNA. *Nat Struct Biol 5*:432–436.

Muriaux D, De Rocquigny H, Roques B, Paoletti J. 1996. NCp7 activates HIV-1 LAI RNA dimerization by converting a transient loop-loop complex into a stable dimer. *J Biol Chem 271*:33686–33692.

Muriaux D, Girard PM, Bonnet-Mathoniere B, Paoletti J. 1995. Dimerization of HIV-1 LAI RNA at low ionic strength. *J Biol Chem 270*:8209–8216.

Puglisi JD, Tinoco I. 1989. Absorbance melting curves of RNA. *Methods Enzymol 180*:304–325.

Rizvi TA, Panganiban AT. 1993. Simian immunodeficiency virus RNA is efficiently encapsidated by human immunodeficiency virus type 1 particles. *J Virol 67*:2681–2688.

Skripkin E, Paillart JC, Marquet R, Ehresmann B, Ehresmann C. 1994. Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro. *Proc Natl Acad Sci USA 91*:4945–4949.

Wei P, Garber ME, Fang SM, Fisher WH, Jones KA. 1998. A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity loop-specific binding to TAR RNA. *Cell 92*:451–462.

Zeffmandagger A, Hassardagger S, Varani G, Lever A. 2000. The major HIV-1 packaging signal is an extended bulged stem loop whose structure is altered on interaction with the Gag polyprotein. *J Mol Biol 297*:877–893.

Zuker M, Mathews DH, Turner DH. 1999. Algorithms and thermodynamics for RNA secondary structure prediction: A practical guide. In: Barciszewski J, Clark BFC, eds. *RNA biochemistry and biotechnology*. Dordrecht/Boston/London: Kluwer Academic Publishers. pp 11–43.