Role of the Trans-activation Response Element in Dimerization of HIV-1 RNA*

Received for publication, December 31, 2003, and in revised form, February 19, 2004
Published, JBC Papers in Press, March 9, 2004, DOI 10.1074/jbc.M314326200

Ebbe S. Andersen‡§, Sonia Antoranz Contera‡**, Bjørn Knudsen‡, Christian K. Damgaard‡§, Flemming Besenbacher‡**, and Jørgen Kjems‡§

From the ‡Interdisciplinary Nanoscience Center, Departments of §Molecular Biology and **Physics and Astronomy, †Bioinformatics Research Center, University of Aarhus, DK-8000 Aarhus C, Denmark

The HIV-1 genome consists of two identical RNA strands that are linked together through non-covalent interactions. A major determinant for efficient dimerization of the two RNA strands is the interaction between palindromic sequences in the dimerization initiation site. Here we use an interplay of bioinformatics, biochemistry, and atomic force microscopy to describe another conserved palindromic in the trans-activation response element (TAR) that functions as a strong dimerization site when transiently exposed to the viral nucleocapsid protein. In conjunction with the DIS interaction, the TAR dimerization induces the formation of a 65-nm higher-order circular structure in the dimeric HIV-1 RNA. Our results provide a molecular model for the role of TAR in packaging and reverse transcription of the viral genome. The unique structure of the TAR-TAR dimer renders it an intriguing therapeutic target for the treatment of HIV-1 infection.

The hallmark of retroviral replication is the copying of a dimeric RNA genome into double-stranded DNA by reverse transcriptase prior to integration in the host cell genome (reviewed in Ref. 1). The dimeric genome of the retrovirus allows reverse transcriptase to switch between the two templates during replication, resulting in a strong recombination potential and fast evolution of drug-resistance (reviewed in Ref. 2). The RNA sequences that control the dimerization process are located within the 5′-untranslated region of the HIV-1 genome, which contains several functionally important elements named according to their proposed function: the trans-activation response element (TAR), the 5′ polyadenylation signal (poly(A)), the primer binding site, the dimer initiation site (DIS), the major splice donor and the packaging signal that precedes the Gag open reading frame (Fig. 1A; reviewed in Ref. 3). The TAR hairpin that is the focus of this report is recognized at the three-nucleotide bulge region by the viral Tat protein (4, 5). The binding of Tat to TAR facilitates binding of cyclin T1 and cyclin-dependent protein kinase-9 to the distal loop of TAR, which concurrently induces transcriptional elongation by hyperphosphorylation of the C-terminal domain of the RNA polymerase II (reviewed in Ref. 6).

It is believed that dimerization of the RNA genomes is initiated through intermolecular base-pairing between the palindromic loop sequences of two DIS hairpins, forming a kissing-loop complex or loose dimer (7). This complex can, under appropriate conditions, be converted into a more stable tight dimer by extended base-pairing (7), a process that is potentiated by the addition of the nucleocapsid (NC) (8) or Gag protein (9). At the subsequent reverse transcription step of the viral replication cycle, the dimeric genome provides a template for the reverse transcriptase that utilizes a tRNA^Lys primer annealed to the primer binding site sequence (reviewed in Ref. 1). The cDNA synthesis proceeds until it reaches the 5′ end of the genome, which forms the so-called minus-strand strong-stop DNA. The presence of a repeated R-region in both ends of the genome enables the extending polymerase to switch templates to the 3′ terminal R-region (10). Because of the diploidy of the RNA genome, this transfer may occur either as an intramolecular or as an intermolecular event (11). It has been shown that the TAR structure at the 5′ end of the genome plays an important role in template switching, although the molecular explanation for this observation remains unclear. It has been suggested that the minus-strand strong-stop DNA, complementary to the TAR stem loop, can form a kissing-loop interaction with the TAR RNA in the 3′ end of the genome and that this interaction facilitates the first-strand transfer (12, 13). By using a combination of bioinformatics, biochemistry, and atomic force microscopy (AFM) techniques, we show that a phylogenetically conserved palindrome in the TAR RNA forms stable dimers independently of the DIS sequence when exposed to the NC protein. The interaction between two TAR hairpins, which we characterize here, provides a molecular explanation for the observed role of TAR in packaging and reverse transcription of the viral RNA.

EXPERIMENTAL PROCEDURES

Phylogenetic Analysis—The search for evolutionary conserved RNA palindromes in the HIV-1 leader sequence was done using a modified version of the RNA secondary structure prediction algorithm, Pfold (14). This program uses a statistical method to find RNA structures given a multiple alignment of related RNA sequences and their evolutionary history. The initial analysis was done using an alignment of 20 divergent leader sequences (15). The bioinformatic analysis of the structural model shown in Fig. 1C was based on 300 unique and unambiguous TAR sequences from the Los Alamos HIV Sequence Database (available on the World Wide Web at hiv-web.lanl.gov/).

Plasmids—The parental construct used in this study was generated by PCR amplification of the first 744 nucleotides of the full-length clone of HIV-1 NLX2 (GenBank™ accession no. K03455) and cloned into pUC18 with a T7 promoter. Site-directed mutagenesis was used to introduce mutations in the TAR and DIS palindromes. The various
HIV-1 subtype and HIV-2 constructs have been described previously (16, 17). The pHIV-1-C2-TAR that was used for footprinting was generated by site-directed mutagenesis of the subtype C2 construct. An extra G was inserted at position -1 and GGG after position 58 to generate a Smal restriction site. pYES2-His was generated by inserting a HindIII-Hind-digested fragment (AGCTTACATGCTCATCATACATCC- ACCATGGATCACCGGGAATTCCTGAT, coding strand) into a HindIII-Hind-digested pYES2 vector (Invitrogen), generating a sequence encoding a 6× His-tag upstream of a BamHI-Smal-ECori multiple cloning site. pYES2-His-Gag was generated by inserting a BamHI-EcoRI restriction fragment of pET-GTH-Gag (18) into a BamHI-EcoRI-digested pYES2-His plasmid.

**RNA Preparation**—The constructs were linearized with HindIII to obtain a single-stranded transcription of the TAR sequences (nucleotides 1–81). The HIV-1 C2 TAR used for footprinting was transcribed from pHIV-1-C2-TAR linearized with Smal, which generated a blunt-ended TAR stem with an extra G-C base pair at the proximal end. The HIV-2 TAR transcript (nucleotides 1–140) was transcribed from pHIV-2 linearized with MseI. To obtain templates for transcription of the entire leader and 409 nucleotides of the Gag open reading frame (nucleotides 1–744), the constructs were linearized with BamHI. Large scale RNA transcription was performed by using 2 μg of linearized plasmid in 20 μl reactions with 80 mM HEPES-KOH, pH 7.5, 25 mM MgCl2, 2 mM spermidine, 20 mM dithiothreitol, 7.5 mM rNTP, 10 units/ml yeast inorganic PPase, 20 units ribonuclease inhibitor (Promega), and 2 units yeast RNA polymerase. The RNA was either internally or 5′-terminally labeled according to Kjems et al. (19).

**Protein Preparation**—Synthetic NCp7 (71 amino acids) from the HXB2 strain was a generous gift from Jean-Luc Darlix. His-tagged Gag protein was expressed in Saccharomyces cerevisiae (JELI) transformed with pYES2-His-Gag growing in YM9 medium (20 g/liter peptone, 10 g/liter yeast extract, 20 g/liter raffinose (Sigma)). Cells were grown for an additional 16 h. Pelleted cells were frozen as pills in liquid nitrogen and ruptured in a bead beater (20 × 15 s) and extracted in 50 mM Tris-HCl, pH 8, 100 mM NaCl, 1 mM MgCl2, 10 mM spermidine, 20 mM dithiothreitol, 7.5 mM rNTP, 10 units/ml yeast RNA polymerase. The RNA was extracted by phenol extraction, with three steps of phenol with 0.1% SDS, one step of phenol/chloroform, and a last step of chloroform. The RNA was either internally or 5′-terminally labeled according to Kjems et al. (19).

**Enzymatic Probing**—The NC treatment and removal were basically done as described by Muriaux et al. (8). A volume of 10 μl of 200 mM RNA was denatured at 95 °C for 5 min and snap-cooled on ice for 5 min. Identical results were obtained for slowly renatured RNA (data not shown). RNA transcripts were incubated in NC buffer containing 20 μM Tris-HCl, pH 7.5, 5 mM dithiothreitol, 50 mM NaCl, and 0.2 mM MgCl2 at 37 °C for 15 min before adding 5 μM NCp7 or Gag, corresponding to ~100% of RNase A (20). After incubation, the reactions were placed on ice for 5 min. Then, at 25 °C for 5 min, S1 was added at a final concentration of 10% of ribonuclease A (Roche). After 10 min, the sample was run on native 7% PAGE and stained with SYTO-9 (Molecular probes). The TAR RNA was either 5′-terminally labeled or 5′-internally labeled using T4 polymerase. The RNA was electrophoresed at 37 °C for 2 h. The samples were treated with DNase I for 15 min prior to purification on 4% denaturing polyacrylamide gels. The RNA was then either internally or 5′-end-labeled according to Kjems et al. (19).

**RNA Treatment**—The NC treatment and removal were basically done as described by Muriaux et al. (8). A volume of 10 μl of 200 mM RNA was denatured at 95 °C for 5 min and snap-cooled on ice for 5 min. Identical results were obtained for slowly renatured RNA (data not shown). RNA transcripts were incubated in NC buffer containing 20 μM Tris-HCl, pH 7.5, 5 mM dithiothreitol, 50 mM NaCl, and 0.2 mM MgCl2 at 37 °C for 15 min with 5 μM NCp7 or Gag, corresponding to ~100% of RNase A (20). After incubation, the reactions were placed on ice for 5 min. Then, at 25 °C for 5 min, S1 was added at a final concentration of 10% of ribonuclease A (Roche). After 10 min, the sample was run on native 7% PAGE and stained with SYTO-9 (Molecular probes). The TAR RNA was either 5′-terminally labeled or 5′-internally labeled using T4 polymerase. The RNA was electrophoresed at 37 °C for 2 h. The samples were treated with DNase I for 15 min prior to purification on 4% denaturing polyacrylamide gels. The RNA was then either internally or 5′-end-labeled according to Kjems et al. (19).

**Enzymatic Probing**—The NC treatment and removal were basically done as described by Muriaux et al. (8). A volume of 10 μl of 200 mM RNA was denatured at 95 °C for 5 min and snap-cooled on ice for 5 min. Identical results were obtained for slowly renatured RNA (data not shown). RNA transcripts were incubated in NC buffer containing 20 μM Tris-HCl, pH 7.5, 5 mM dithiothreitol, 50 mM NaCl, and 0.2 mM MgCl2 at 37 °C for 15 min with 5 μM NCp7 or Gag, corresponding to ~100% of RNase A (20). After incubation, the reactions were placed on ice for 5 min. Then, at 25 °C for 5 min, S1 was added at a final concentration of 10% of ribonuclease A (Roche). After 10 min, the sample was run on native 7% PAGE and stained with SYTO-9 (Molecular probes). The TAR RNA was either 5′-terminally labeled or 5′-internally labeled using T4 polymerase. The RNA was electrophoresed at 37 °C for 2 h. The samples were treated with DNase I for 15 min prior to purification on 4% denaturing polyacrylamide gels. The RNA was then either internally or 5′-end-labeled according to Kjems et al. (19).

**RESULTS**

**TAR Contains a Phylogenetically Conserved Palindrome**—To identify potential dimerization sites, we searched for evolutionary conserved palindromes (self-complementary sequences) in an alignment of 20 divergent HIV-1 leader sequences using a bioinformatics approach (14, 15). This algorithm predicted the presence of three significant palindromes, two of which are strongly supported by sequence variations among divergent HIV-1 strains (Fig. 1A). In addition to the previously characterized DIS palindrome (7), we found that a 10-nucleotide sequence in the distal part of the TAR hairpin can potentially base-pair with another TAR hairpin through Watson-Crick base pairs and two GU wobble base pairs (Fig. 1B, red stem). This sequence is conserved in 78% of 300 different TAR sequences from the HIV-1 database, reflecting in part the requirement of proper Tat function. Interestingly, 18% of the HIV-1 sequences contain substitutions conserving the intermolecular base-pairing potential if C-A base pairs are included (Fig. 1C). The C-A base pair has been demonstrated in other RNA structures and has unique symmetrical properties (21; reviewed in Ref. 22). The TAR intermolecular stem can be extended eight base pairs further by allowing one mismatch on each side (Fig. 1B, blue sequences). This extension is also supported by phylogenetic evidence with a high frequency of C-A base pairs (Fig. 1C, blue stem). These observations suggest that there is a strong phylogenetic pressure to conserve the intermolecular-base-pairing potential between two TAR structures.

**TAR Dimerization Is Mediated by the Palindrome and the NC Protein**—We investigated the dimerization capacity of wild-type and mutant TAR hairpins in a gel retardation assay (Fig. 2, A and B). The TAR RNA with the central wild-type (WT) palindrome (GAGGUCUC) does not spontaneously form dimers, but upon transient incubation with NC protein from HIV-1 HXB2, it forms two dimer species with an efficiency of ~50% (Fig. 2B, lanes 1 and 2). Because the NC protein is a proteolytic product of the Gag polypeptide, we tested the ability of purified Gag from HIV-1 HXB2 to induce dimerization, but found no activity in the TAR dimerization assay (result not shown).

Next we tested the importance of the palindromic sequence in the TAR RNA. Mutating A34→U in the palindrome (Fig. 2A, TAR1, GUGGUCUC(38)) eliminates dimerization, whereas restoring the palindromic nature of the sequence by an additional mutation, U37→A (Fig. 2A, TAR2, GUGGUCUC(38)), allows dimers to form (Fig. 2B, lane 3 and 4). The increased dimerization capacity of the TAR2 restoration mutant compared with WT TAR is most likely due to the concomitant destabilization of the TAR stem and exposure of the palindrome. This interpretation is consistent with the observation that a triple mutant (Fig. 2A, TAR3), where an additional mutation of A26→U in TAR2 restores the stem, dimerizes with an efficiency comparable...
FIG. 1. HIV-1 RNA secondary structure models. A, secondary structure model of the 5'-untranslated region of HIV-1 HXB2, with indication of palindromes predicted by the Pfold program from an alignment of 20 divergent HIV-1 leader sequences. The palindromes are scored according to the strength of base-pairing and conservation of palindromic nature (highly conserved palindromic elements are indicated in red, whereas weaker or non-conserved palindromes are indicated in gray). The RNA structural elements are marked: TAR, trans-activation response element; poly(A), polyadenylation site; PBS, primer-binding site; DIS, dimerization initiation site; SD, major splice donor; PSI, packaging signal; Gag ORF, the start of the Gag open reading frame. B, model of the structural rearrangements in the TAR hairpin. TAR dimerization induced by the nucleocapsid (NC) protein gives rise to a kissing-loop complex and an extended duplex. The following sequences are marked: the central palindrome (red), extended base-pairing (blue), and secondary parallel helix (green; gray dotted lines in extended duplex). C, phylogenetic support for the dimer stems based on 300 non-redundant TAR sequences. The percentages refer to frequencies of single-nucleotide variations in the individual stems (only variations that occur in at least 2% of HIV-1 sequences are included to avoid sequencing errors). Color codes are as in panel B. Numbering starts at the transcription start site (+1) that corresponds to position 455 in the provirus of HIV-1 HXB2.
with WT TAR (Fig. 2B, lane 5). The importance of the whole TAR composition is further evidenced by the observation that the deletion of the bulged A16 (TAR4) completely eliminates dimer formation (Fig. 2B, lane 6). This deletion enables the formation of a 16-base-pair uninterrupted stem that probably renders the TAR helix too stable for NC to unwind and expose the palindrome. We suggest that the chaperone activity of the NC protein is required to expose the partially hidden palindrome within the TAR stem.

Evidence for a Two-step Progression of TAR Dimerization—

The TAR dimer forms two structurally distinct complexes in the gel retardation assay, which most likely represent an intermediate kissing-loop (KL) and extended duplex (ED) structure (Fig. 2B, lower and upper dimer bands, respectively; see model in Fig. 1B). This assignment is based on the observation that the TAR3 mutant that destabilizes the extended base-pairing (Fig. 1B, blue stem) also inhibits the formation of the ED (Fig. 2B, lane 5). The RNA sequences, which are liberated upon ED formation, may potentially form a secondary parallel duplex within the ED dimer, thus further in-

![Fig. 2. Dimerization of the TAR hairpin.](image-url)
increasing the stability of the dimeric TAR complex (Fig. 1B, green stem). This interaction is also consistent with nucleotide substitutions in divergent HIV-1 strains (Fig. 1C, green stem). Hence, our data are compatible with the model that the progression from an initial KL to an ED structure involves the formation of two parallel duplexes with two three-way junctions. The observation of trimer and tetramer, but no higher order TAR complexes, raises the possibility that the exposed nucleotides in the distal loop of TAR (U30–U32) alters the ratio between the ED and KL forms, suggesting that this nucleotide plays an important role in the progression of KL into the ED form. Moreover, comparing the two subtype B strains (HXB2 and LAI), we reproducibly saw a lower yield of ED formation for LAI, which may be explained by the unique C23→U substitution in the green stem characteristic for the ED form (compare Fig. 2B, lane 2 to Fig. 2D, lane 2). HIV-2 also forms dimers in the presence of HIV-1 NC (Fig. 2D, lanes 7 and 8), consistent with the fact that this virus also contains a conserved RNA palindrome in the TAR hairpin.

After having established that the TAR hairpin alone can form dimers, we investigated the role of TAR in dimerization of a more extended region spanning the 5′ part of the HIV-1 genome (nucleotides 1–744). Incubation of this RNA with the NC protein induces the formation of higher order RNA-protein complex unable to enter the gel (Fig. 2E, lane 1). When incubating the HIV-1 RNA in buffer without NC, dimerization occurs inefficiently in a DIS-dependent manner (Fig. 2E, lane 2 and 3). Incubation of the WT RNA with NC followed by phenol extraction leads to an increased amount of dimer (Fig. 2E, lane 4). Disrupting the TAR palindrome in the TAR1 mutant does not alter the dimerization efficiency probably because the dimerization can occur via the DIS palindrome. The TAR2 mutant RNA forms mostly multimers, suggesting that the increased TAR dimerization capacity of this mutant may favor TAR and DIS dimerization between different HIV-1 RNA molecules (Fig. 2E, lane 6). The TAR3 mutant dimerizes to a similar extent as wild-type (Fig. 2E, lane 7). Analyzing all of the TAR mutants in the absence of a functional DIS palindrome yielded a dimerization pattern that is very similar to what was observed for the TAR hairpin alone, except that the overall dimerization level was clearly reduced (Fig. 2E, lanes 8–11). This shows that, although the DIS constitutes the major dimerization site, the extended HIV-1 RNA can dimerize in the absence of DIS and that this dimerization is clearly dependent upon the

---

**Fig. 3. Enzymatic structure probing of the TAR dimer.** A, dimerization of TAR RNA from HIV-1 subtype C2 in the absence and presence of NC treatment under the footprinting conditions used in panel B. B, autoradiogram showing enzymatic probing of wild-type TAR from HIV-1 subtype C2. The RNases that were applied include: single-strand specific RNase T1 (T1), T2 (T2) or double-strand specific RNase V1 (V1); RNase V2 fragments migrate ~½ nucleotide position slower than RNase T1 and T2 fragments; C, control without RNase; A, alkaline hydrolysis ladder. The enhancements of RNase cleavages in the TAR dimer are indicated by arrowheads; protections are indicated by circles using the following color code: T1 (blue), T2 (green), and V1 (red). C, summary of the result shown in panel B. The NC-induced enhancements and protections are plotted on the ED structure of TAR from C2 using the same symbols as in panel B. The terminal base pair shown in brackets is included to create an SmaI site used for linearization of the template. This allows generation of TAR RNA without protruding 5′ or 3′ ends.
Role of TAR in HIV-1 Dimerization

TAR is known to be a multifunctional region involved in transcriptional activation, packaging, and reverse transcription of the HIV-1 genome (3). In this study, we show that, in addition, the TAR hairpin harbors a dimerization signal that is induced by the NC protein. Based on the appearance of two dimeric TAR complexes with different mobility in a native gel, we propose a model implying that dimerization occurs in two steps. Our phylogenetic study, mutational analysis, and enzymatic probing to investigate the structures of monomeric and dimeric TAR RNAs from HIV-1 subtypes B-HXB2 (data not shown) and A-F11006 with a diameter of 65 ± 15 nm. Another significant structural change occurs in the bulged annealing (Fig. 3) clearly supports the role of the palindrome in intermolecular association of proteins with double-stranded RNA (23–27). To image the dimer linkage region of HIV-1 RNA at the single-molecule level, we applied AFM to the RNA spanning nucleotides 1–744 of the HIV-1 genome. TAR and DIS palindromes are marked in red.

Probing the TAR-TAR Interaction—We next used enzymatic probing to investigate the structures of monomeric and dimeric TAR RNAs from HIV-1 subtypes B-HXB2 (data not shown) and C2 (Fig. 3). The ladder subtype forms dimers very efficiently under the probing conditions (Fig. 3A) and, therefore, allows a better distinction between the two forms. A comparison of the probing results from the WT TAR hairpin with and without NC-treatment reveals a dramatic structural reorganization of the distal loop upon dimerization. A significant increase in double-strand specific RNase V1 cleavage at positions G31  G32  G33  and G30  G31  G32  G33  G34, respectively, clearly supports the role of the palindrome in intermolecular annealing (Fig. 3B; lanes 3-5 and 7-9, summarized in Fig. 3C). Another significant structural change occurs in the bulged region. The T2 RNase cleavages at C23  and U24  within the bulge were strongly reduced upon dimer formation as one would expect from the model (Fig. 3B, lanes 4 and 8). Accordingly, G25 and G27, which are base-paired in the monomer stem, become moderately accessible to single-strand specific RNase T1 in the dimer (Fig. 3B, compare lane 3 and 7). However, the observed induction of T1 RNase cleavage at 27G  was completely absent when probing subtype B. Therefore, we favor a model in which G42-U45 is base-paired to A26-C29 (Fig. 3C, blue stem in the dimer), although it is possible that C44-U45 alternatively can bind A14-G15. We conclude that the probing data on the monomer and dimer TAR RNA is compatible with the structural model shown in Fig. 1B.

Direct Visualization of Intermolecular Linkages in HIV-1 Dimerization—AFM has previously been used to study the association of proteins with double-stranded RNA (23–27). To image the dimer linkage region of HIV-1 RNA at the single-molecule level, we applied AFM to the RNA spanning nucleotides 1–744 of the HIV-1 genome. RNA constructs, containing wild-type or mutated TAR palindromes, were transiently incubated with or without NC protein and deposited on a spermine-treated mica surface. The AFM images of the NC-treated WT RNA exhibited distinct circular structures with a diameter of 65 ± 15 nm that were assigned as RNA dimers based on length measurements (Fig. 4, A and B). Notably, AFM imaging of either RNA that contains the disruptive TAR1 mutation in the palindrome or RNA that is incubated without NC protein revealed no circular structures (Fig. 4, C and D). RNA with mutations in the DIS palindrome always yielded AFM images in which the RNA appeared partially aggregated without sign of circular structures (data not shown). This shows that the TAR palindrome and the NC-incubation is important for formation of the circular structure, and the most likely interpretation of this result is that the circular structure is formed through intermolecular linkages involving the TAR and the DIS palindromes (Fig. 4E).

**DISCUSSION**

TAR is known to be a multifunctional region involved in transcriptional activation, packaging, and reverse transcription of the HIV-1 genome (3). In this study, we show that, in addition, the TAR hairpin harbors a dimerization signal that is induced by the NC protein. Based on the appearance of two dimeric TAR complexes with different mobility in a native gel, we propose a model implying that dimerization occurs in two steps. Our phylogenetic study, mutational analysis, and enzymatic probing to investigate the structures of monomeric and dimeric TAR RNAs from HIV-1 subtypes B-HXB2 (data not shown) and C2 (Fig. 3). The ladder subtype forms dimers very efficiently under the probing conditions (Fig. 3A) and, therefore, allows a better distinction between the two forms. A comparison of the probing results from the WT TAR hairpin with and without NC-treatment reveals a dramatic structural reorganization of the distal loop upon dimerization. A significant increase in double-strand specific RNase V1 cleavage at positions G31  G32  G33  and G30  G31  G32  G33  G34, respectively, clearly supports the role of the palindrome in intermolecular annealing (Fig. 3B; lanes 3-5 and 7-9, summarized in Fig. 3C). Another significant structural change occurs in the bulged region. The T2 RNase cleavages at C23  and U24  within the bulge were strongly reduced upon dimer formation as one would expect from the model (Fig. 3B, lanes 4 and 8). Accordingly, G25 and G27, which are base-paired in the monomer stem, become moderately accessible to single-strand specific RNase T1 in the dimer (Fig. 3B, compare lane 3 and 7). However, the observed induction of T1 RNase cleavage at 27G  was completely absent when probing subtype B. Therefore, we favor a model in which G42-U45 is base-paired to A26-C29 (Fig. 3C, blue stem in the dimer), although it is possible that C44-U45 alternatively can bind A14-G15. We conclude that the probing data on the monomer and dimer TAR RNA is compatible with the structural model shown in Fig. 1B.
matic probing data suggest that the initial KL interaction involves 10 base pairs. In a second step, this interaction was extended by four base pairs on either side, followed by the formation of a second parallel helix, which resulted in two three-way helical junctions. In this two-step process, 20 base pairs were broken and 28 new ones were formed, making the reaction thermodynamically favorable. The required chaperone activity of the NC protein may reflect that the TAR palindrome is situated partly in a helical region and not readily available for initial intermolecular base-pairing. The two-step dimerization reaction resembles the dimerization of the DIS stem-loop that also proceeds from an initial loose dimer that involves kissing of the distal loops into an extended dimer where the DIS stem dissociates and becomes engaged in intermolecular base-pairing (7). The TAR dimerization model also bears some resemblance to the suggested structure of the CopA-CopT complex, which is functional in bacterial plasmid-replication control where a loop-loop interaction progresses to form a four-way helical junction (21). Interestingly, the non-canonical A-C base pair, which is observed with high frequency in the TAR dimer complex, is also found in the CopA-CopT and other related complexes (28) and might provide special symmetrical properties in the topology of these characteristic knot-structures.

The TAR dimerization efficiency of the different HIV-1 subtypes varies between 20–80%. Differences in the dimerization efficiency have also been observed for the DIS hairpin when analyzing different subtypes (16). The HIV-1 subtypes may thus adjust the dimerization capability of individual sites to obtain a certain over-all stability of the dimeric genome in a cooperative manner.

The AFM analysis indicated that the TAR-TAR interaction, together with the dimer linkage site in the DIS region, is responsible for the formation of ~65 nm circular dimeric structures. Because the length of the RNA strand between the TAR and DIS palindromes is 216 nucleotides, we estimate the length per nucleotide to be 0.47 ± 0.11 nm within the RNA circle. This value is in agreement with previous x-ray and AFM studies of single-stranded homopolymer RNA (29, 30) and is smaller than the length of a fully extended backbone (0.6 nm/nucleotide). Therefore, a significant proportion of the RNA spanning the loop region must be single-stranded. The circles observed by AFM resemble the structures visualized by electron microscopy in RNA isolated from virus particles by Högland et al. (31). They speculated that the 5′ end interaction involves the palindromic crowning the poly(A) hairpin (Fig. 1A). However, this sequence was subsequently shown not to be important for dimerization in vitro (32). In the light of our results, the circles observed by Högland et al. may also involve the TAR-TAR interaction.

NC is produced by proteolytic cleavage from the Gag precursor protein during maturation of the viral particle. The finding that the TAR-TAR interaction is dependent upon the NC protein may therefore ensure that TAR dimerization does not occur during synthesis of the HIV-1 RNA in the infected cell, where TAR has to fulfill its role in enhancing transcription through binding of the viral Tat and cellular proteins (reviewed in Ref. 6). During virion assembly or immediately after budding from the host cell, the NC protein is released and becomes closely associated with the RNA genome. At this stage, NC can facilitate structural rearrangements that stabilize the TAR dimer into an extended duplex structure. This may serve as an additional element to stabilize the RNA dimer at the packaging step. TAR dimerization may also function to bring together the TAR elements, which are positioned at each end of the HIV-1 RNA genome. Our observation of tetrameric TAR complexes may even suggest that all four TAR copies present in the virus particle can become associated. Association of the TAR sequences located at the termini of the HIV-1 genome could facilitate the essential transfer of the minus-strand strong-stop DNA transcription unit from the 5′ to the 3′ R region in the genome (10). This model is in agreement with recent reports (16, 33) showing that RNA template dimerization strongly increases template switching by reverse transcriptase and that the native structure of HIV-1 RNA extracted from virus particles is required for the first-strand transfer (34). Thus, the NC protein might play an important role in setting the stage for reverse transcription of the HIV-1 genome. The importance of the TAR dimerization for viral replication remains to be studied. If crucial, the unique structural properties of the TAR dimer may provide a novel therapeutic target for inhibiting several fundamental steps in the replication cycle of HIV-1.

Acknowledgments—We thank Jean-Luc Darlix for the generous donation of synthetic NCp7, Rita Rosendahl Hansen for excellent technical assistance, and Ray Brown and Finn Skou Pedersen for critical reading of the manuscript.

REFERENCES

1. Coffin, J. M., Hughes, S. H., and Varmus, H. E., eds (1997) Retroviruses, pp. 121–160, Cold Spring Harbor Laboratory Press, New York
2. Temin, H. M. (1991) Trends Genet. 7, 71–74
3. Berkhour, R. (2000) Adv. Pharmacol. 48, 29–73
4. Dingwall, C., Ernberg, I., Gust, M. J., Green, S. M., Heaphy, S., Karn, J., Lowe, A. D., Singh, M., and Skinner, M. A. (1990) EMBO J. 9, 4145–4153
5. Weicks, E. M., Ampe, C., Schultz, S. C., Steinz, T. A., and Crothers, D. M. (1990) Science 249, 1281–1285
6. Karn, J. (1999) J. Mol. Biol. 293, 235–254
7. Skripkin, E., Paillart, J. C., Marquet, R., Ehresmann, B., and Ehresmann, C. (1999) Proc. Natl. Acad. Sci. U.S.A. 96, 4945–4949
8. Muriaux, D., de Rocquigny, H., Roques, B. P., and Paoletti, J. (1996) J. Biol. Chem. 271, 33686–33692
9. Feng, Y. X., Campbell, S., Harvin, D., Ehresmann, B., Ehresmann, C., and Rein, A. (1999) J. Virol. 73, 4251–4256
10. Gilboa, E., Mitr, S. W., Goff, S., and Baltimore, D. (1979) Cell 18, 93–100
11. van Wamel, J. L., and Berkhout, B. (1998) Virology 244, 245–251
12. Jeeninga, R. E., Huthoff, H. T., Gultyaev, A. P., and Berkhout, B. (1993) Nucleic Acids Res. 21, 5472–5479
13. Berkhour, B., Vestenhoudt, N. L., Klaasen, I. B., and Houthof, H. (2001) RNA 7, 1097–1114
14. Knudsen, B., and Hein, J. (2003) Nucleic Acids Res. 31, 3423–3428
15. Damgaard, C. K., Andersen, E. S., Knudsen, B., Gorodkin, J., and Kjems, J. (2004) J. Mol. Biol. 333, 207–218
16. Andersen, E. S., Jeeninga, R. E., Damgaard, C. K., Berkhour, B., and Kjems, J. (2003) J. Virol. 77, 3020–3030
17. Damgaard, C. K., Dyhr-Mikkelsen, H., and Kjems, J. (1998) Nucleic Acids Res. 26, 3676–3677
18. Jensen, T. H., Jensen, A., and Kjems, J. (1998) Gene 162, 235–237
19. Kjems, J., Egelberg, J., and Christiansen, J. (1996) Analysis of RNA protein Complexes in Vitro (van der Vliet, P. C., ed.) p. 26, Elsevier Science B. V., Amsterdam, The Netherlands
20. You, J. C., and McHenry, C. S. (1994) J. Biol. Chem. 269, 31491–31495
21. Kolb, F. A., Engdahl, H. M., Slatter-Jüger, J. G., Ehresmann, B., Ehresmann, C., Westhof, E., Wagner, E. G., and Romby, P. (2000) EMBO J. 19, 5905–5915
22. Massguida, B., and Westhof, E. (2000) RNA 6, 9–15
23. Hess, A., Medalia, O., Shi, S. P., Steinberg, M., Franceschi, F., and Sagi, I. (2001) Proc. Natl. Acad. Sci. U.S.A. 98, 5007–5012
24. Bonin, M., Oberstrass, J., Lukacs, N., Ewert, K., Oesterschulze, E., Kassing, R., and Nellen, W. (2000) RNA 6, 563–570
25. Bonin, M., Oberstrass, J., Vogt, U., Wassenedegger, M., and Nellen, W. (2001) Bioc. Chem. 382, 1157–1162
26. Shen, C. L., Galie, D. R., Le, H., and Hansma, P. K. (1997) J. Struct. Biol. 119, 109–117
27. Kisseljova, O. I., Yaminzky, I. V., Karger, E. M., Frulova, O. Y., Dorokhov, Y. L., and Atabekov, J. G. (2001) J. Gen. Virol. 82, 1503–1508
28. Kolb, F. A., Westhof, E., Ehresmann, B., Ehresmann, C., Wagner, E. G., and Romby, P. (2000) J. Mol. Biol. 290, 605–614
29. Saenger, W. (1984) in Principles of Nucleic Acid Structures (Cantor, C. R., ed) pp. 298–330, Springer, New York
30. Hansma, G. H., Revenko, I., Kim, K., and Laney, D. E. (1996) Nucleic Acids Res. 24, 713–729
31. Högland, S., Ohagen, A., Gonçalves, J., Panganiban, A. T., and Gabuzda, D. (1997) Virology 233, 271–279
32. Laughrue, M., Shen, N., Jette, L., Darlix, J. L., Kleiman, L., and Wainberg, M. A. (2001) Virology 281, 109–116
33. Balakrishnan, M., Roques, B. P., Pay, P. J., and Bambara, R. A. (2003) J. Virol. 77, 4710–4721
34. Berkhour, B., Das, A. T., and van Wamel, J. L. (1998) Virology 249, 211–218