Dimerization of HIV-1 genomic RNA of subtypes A and B: RNA loop structure and magnesium binding.

F Jossinet, J C Paillart, E Westhof, et al.

RNA 1999 5: 1222-1234

RNA (1999), *5*:1222–1234. Cambridge University Press. Printed in the USA.

Dimerization of HIV-1 genomic RNA of subtypes A and B: RNA loop structure and magnesium binding

FABRICE JOSSINET, JEAN-CHRISTOPHE PAILLART, ERIC WESTHOF, THOMAS HERMANN,[1] EUGENE SKRIPKIN,[1] J. STEPHEN LODMELL, CHANTAL EHRESMANN, BERNARD EHRESMANN, and ROLAND MARQUET

Unité Propre de Recherche No. 9002 du Centre National de la Recherche Scientifique, Institut de Biologie Moléculaire et Cellulaire, 67084 Strasbourg cedex, France

ABSTRACT

Retroviruses encapsidate their genome as a dimer of homologous RNA molecules noncovalently linked close to their 5′ ends. The dimerization initiation site (DIS) of human immunodeficiency virus type 1 (HIV-1) RNA is a hairpin structure that contains in the loop a 6-nt self-complementary sequence flanked by two 5′ and one 3′ purines. The self-complementary sequence, as well as the flanking purines, are crucial for dimerization of HIV-1 RNA, which is mediated by formation of a "kissing-loop" complex between the DIS of each monomer. Here, we used chemical modification interference, lead-induced cleavage, and three-dimensional modeling to compare dimerization of subtype A and B HIV-1 RNAs. The DIS loop sequences of these RNAs are AGGUGCACA and AAGCGCGCA, respectively. In both RNAs, ethylation of most but not all phosphate groups in the loop and methylation of the N7 position of the G residues in the self-complementary sequence inhibited dimerization. These results demonstrate that small perturbations of the loop structure are detrimental to dimerization. Conversely, methylation of the N1 position of the first and last As in the loop were neutral or enhanced dimerization, a result consistent with these residues forming a noncanonical sheared base pair. Phosphorothioate interference, lead-induced cleavage, and Brownian-dynamics simulation revealed an unexpected difference in the dimerization mechanism of these RNAs. Unlike subtype B, subtype A requires binding of a divalent cation in the loop to promote RNA dimerization. This difference should be taken into consideration in the design of antidimerization molecules aimed at inhibiting HIV-1 replication.

Keywords: AIDS; kissing-loop complex; retrovirus; RNA dimer; RNA–RNA interactions

INTRODUCTION

Despite their different host cells and the fact that they induce very different diseases, retroviruses share several common features. In particular, they encapsidate two homologous molecules of genomic RNA physically linked in a region called the dimer linkage structure (DLS) located close to their 5′ ends (Bender & Davidson, 1976; Kung et al., 1976; Bender et al., 1978; Murti et al., 1981).

The dimerization initiation site (DIS) of the genomic RNA of the MAL isolate of the human immunodeficiency virus type 1 (HIV-1) was identified by combining chemical modification interference experiments and site-directed mutagenesis (Paillart et al., 1994; Skripkin et al., 1994). It consists of a short sequence located between the primer binding site (PBS) and the major splice donor site (SD) that folds into a stem-loop structure (Fig. 1A). The homologous sequence is responsible for in vitro dimerization of the RNA of the LAI isolate of HIV-1 (Laughrea & Jetté, 1994, 1996a; Muriaux et al., 1995; Clever et al., 1996).

When infectious HIV-1 molecular clones were mutated in the DIS, the replication rate of the viruses dramatically decreased (Berkhout & van Wamel, 1996; Haddrick et al., 1996; Paillart et al., 1996a) and their infectivity was reduced 100- to 1,000-fold (Paillart et al., 1996a; Laughrea et al., 1997). A significant fraction of the genomic RNA inside the virions was monomeric (Haddrick et al., 1996), and even though the thermal stability of the encapsidated mutated dimers was similar to that of the wild-type dimer (Berkhout & van Wamel, 1996; Laughrea et al., 1997), they appeared abnormal when assessed by nondenaturing

Reprint requests to: Roland Marquet, Unité Propre de Recherche No. 9002 du Centre National de la Recherche Scientifique, Institut de Biologie Moléculaire et Cellulaire, 15 rue René Descartes, 67084 Strasbourg cedex, France; e-mail: marquet@ibmc.u-strasbg.fr.

[1]Present address: Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10021, USA.

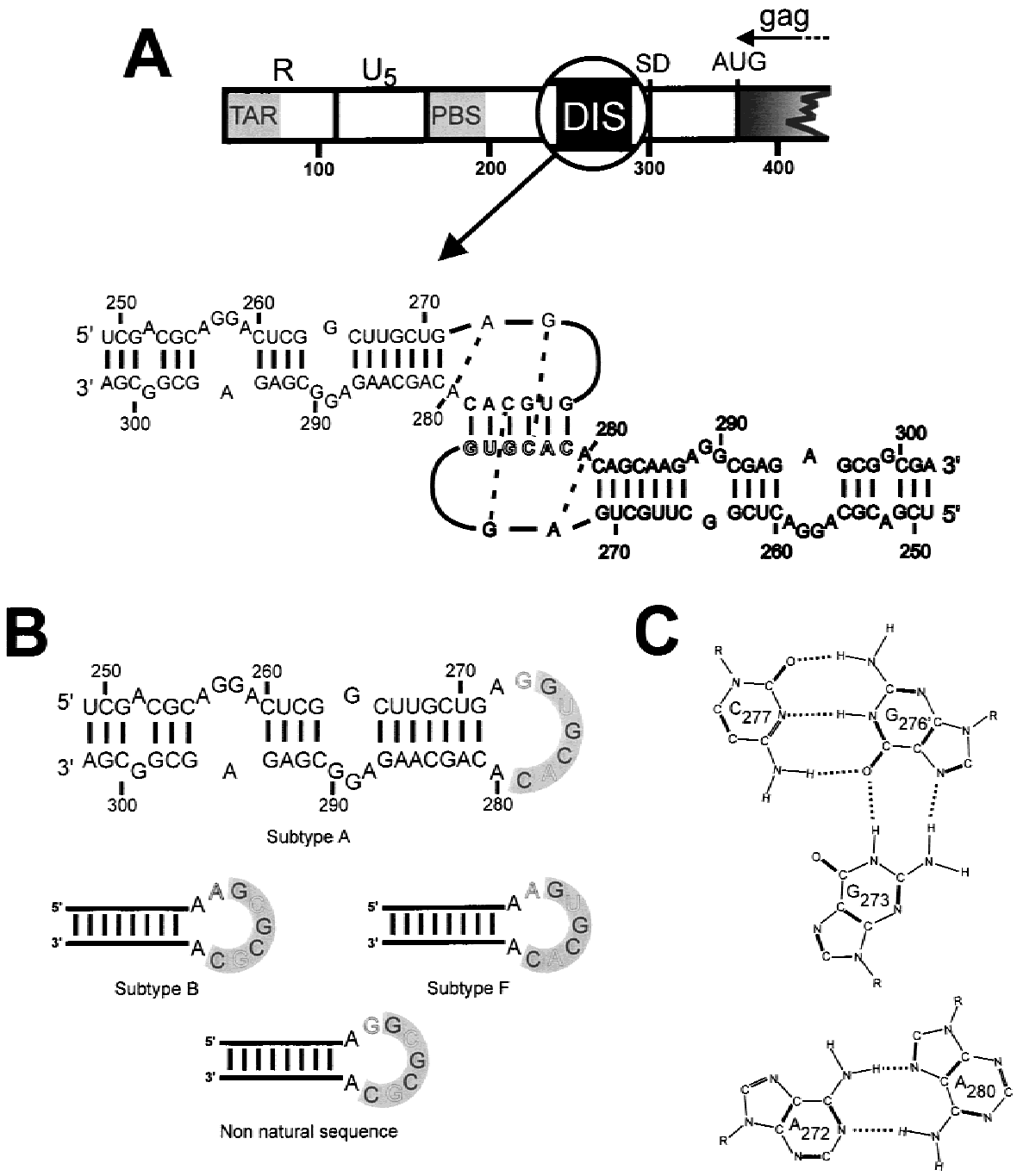

FIGURE 1. **A**: Localization of the DIS in the 5′ end of HIV-1 genomic RNA and kissing-loop structure of the RNA dimer (subtype A). TAR: tat responsive element; R: repeat sequence; U5: unique sequence at the 5′ end of the genome; PBS: primer-binding site; DIS: dimerization initiation site; SD: major splice donor site; AUG: start codon of *gag* translation. **B**: Stem-loop structure of the DIS subtypes used in this study. Noncanonical interactions proposed in Paillart et al. (1997) are indicated by dashed lines. **C**: Schematic drawing of the noncanonical interactions proposed in the dimer of subtype A RNA (Paillart et al., 1997).

gel electrophoresis (Clever & Parslow, 1997). At least two other steps of the HIV-1 replication cycle were affected (Paillart et al., 1996a; Clever & Parslow, 1997): encapsidation of the genomic RNA (Berkhout & van Wamel, 1996; McBride & Panganiban, 1996, 1997; Paillart et al., 1996a; Laughrea et al., 1997; Harrison et al., 1998) and reverse transcription (Paillart et al., 1996a).

The dimerization mechanism of RNAs containing parts of or the complete 5′ untranslated region of the HIV-1 genome has been studied extensively in vitro. It is now widely accepted that dimerization takes place by sym-

metric intermolecular interactions between the self-complementary sequences of the DIS loop, thus forming a "kissing-loop" complex (Fig. 1A) (Laughrea & Jetté, 1994, 1996a, 1996b, 1997; Paillart et al., 1994, 1996a, 1996b, 1996c, 1997; Skripkin et al., 1994; Muriaux et al., 1995, 1996a, 1996b; Clever et al., 1996; Haddrick et al., 1996; Marquet et al., 1996; Clever & Parslow, 1997). Indeed, any mutation that destroyed the self-complementarity in the DIS loop prevented dimerization (Paillart et al., 1994, 1996c, 1997; Skripkin et al., 1994; Clever et al., 1996), whereas introduction of compensatory mutations restored the process (Paillart et al., 1994, 1997). Furthermore, two RNAs with point mutations in the DIS that prevented homodimerization could complement each other (Paillart et al., 1994, 1997).

Even though the DIS stem-loop is conserved among all HIV-1 isolates, sequence variations are tolerated in the loop (Korber et al., 1997; St Louis et al., 1998) (Fig. 1B). The AAGCGCGCA sequence is found in HIV-1 subtypes B (including HIV-1 LAI) and D of the M (major) HIV-1 group, whereas the self-complementary sequence GUGCAC is found in subtypes A, C, F, and H, in some viral recombinants thereof (including HIV-1 MAL) and in HIV-1 group O (outliers) (Korber et al., 1997; St Louis et al., 1998). Variations are also observed in the nucleotides flanking the GUGCAC sequence: the most common sequence, AGGUGCACA, is found in subtype A and in recombinants thereof, whereas AAGUGCACA is found in subtypes F and H (Fig. 1A). Furthermore, in subtype C and group O (AAGUGCACU and GAGUGCACC, respectively) the first and last nucleotides of the DIS loop are complementary, thus actually reducing it to 7 nt, (Korber et al., 1997; Laughrea & Jetté 1997).

In a previous study, we showed that the second purine in the DIS loop determines the association kinetics and the stability of the RNA dimers (Paillart et al., 1997). A three-dimensional model of the DIS dimer based on the biophysical and structural analysis of a number of RNA mutants suggested that this purine interacts with the central base pairs of the self-complementary sequence, whereas the first and last adenines in the loop form a noncanonical base pair (Paillart et al., 1997) (Fig. 1C). This model is in keeping with preliminary NMR data obtained on a dimer of subtype A (Dardel et al., 1998), but differs from the NMR structure of a dimer of subtype B (Mujeeb et al., 1998). In the latter structure, the dimer is stabilized by intermolecular stacking of the first and second purines of the two DIS loops (Mujeeb et al., 1998). Furthermore, RNAs of subtypes A and B displayed different behaviors toward antisense oligonucleotides (Lodmell et al., 1998). These data show that significant differences might exist in the dimerization mechanisms of subtypes A and B.

In the present study, we compared dimerization of RNAs of subtypes A and B by monitoring interference of chemical modifications of the phosphate backbone and bases on the dimerization process, and lead-induced cleavage of the dimers. Our data highlight the importance of the DIS loop structure in the dimerization process and point toward a significant difference in the dimerization mechanism of RNAs of subtypes A and B.

RESULTS

Experimental strategy

In this study, we used chemical modification interference experiments and lead-induced cleavage to compare dimerization of RNAs encompassing the DIS of different HIV-1 subtypes. We focused on subtypes A and B, which are representative of the most widespread DIS sequences (Korber et al., 1997; St Louis et al., 1998), but also performed some experiments on subtype F and on a nonnatural sequence containing an AGGCGCGCA sequence in the DIS loop (Fig. 1B).

Modification interference previously showed that no modification of the Watson–Crick positions was tolerated in the 6-nt self-complementary sequence of the DIS loop (Skripkin et al., 1994). Here, we used the same strategy to test the importance in the dimerization process of the phosphodiester backbone, and of base positions not involved in Watson–Crick pairing of the self-complementary sequence. The phosphate groups were modified by mild treatment with *N*-ethyl-*N*-nitrosourea (ENU), which ethylates either of the nonbridging oxygens, or by statistical thiophosphate incorporation during in vitro transcription, which substitutes sulfur for the *pro*-Rp oxygen (Eckstein & Gish, 1989). DIS RNAs were also treated in mild conditions with dimethyl sulfate (DMS) to study the effects of modifying the N7-G and N1-A positions. In the latter case, we focused on the modifications of A_{272} and A_{280}, which have been proposed to form a noncanonical base pair (Fig. 1C) (Paillart et al., 1997; Dardel et al., 1998).

In all cases, the experimental conditions were chosen to statistically introduce less than one modification per RNA molecule. Modified RNAs were allowed to dimerize and the monomeric and dimeric forms of the RNAs were separated by gel electrophoresis under native conditions and eluted from the gel. To minimize degradation of the RNAs containing phosphorothioates or ethylated phosphates, the dimerization reaction was conducted at pH 6.5, rather than at pH 7.5 as in our previous studies. In addition, we noticed that dimerization of the short DIS RNAs required lower magnesium concentrations than dimerization of the larger RNAs previously used (Marquet et al., 1994; Paillart et al., 1994, 1996c, 1997; Skripkin et al., 1994). The modifications present in each form were then revealed by chemical-induced RNA scission at the modified nucleotides followed by polyacrylamide gel electrophoresis under denaturing conditions, with the exception of ad-

enines methylated at their N-1 position, which were detected by primer extension (Ehresmann et al., 1987). Modifications present in the monomeric species but absent from the dimer are those preventing dimerization (negative interference). Conversely, modifications favoring dimerization are more abundant in the dimeric form than in the monomeric form (positive interference).

Phosphate ethylation interference

In the first set of experiments, we analyzed the effect of ethylating the nonbridging oxygens of the phosphate groups of subtype B RNA. A single pronounced window of negative interference was reproducibly observed (Figs. 2A and 7). It corresponds to complete inhibition of dimerization upon ethylation of any phosphate group 3′ to nt 274 to 281. Ethylation of the phosphate 3′ to A_{282} resulted in partial interference. The single band observed in the interference window under the dimer conditions migrated between the bands corresponding to A_{278} and C_{279} in the monomer lane (Fig. 2A). Thus, this band was not taken into account, as it did not correspond to an ethylated RNA, but most likely to a hydrolysis product. Indeed, it migrated exactly at the same level as the hydrolysis product corresponding to C_{279} in the alkaline ladder (Fig. 2A). The window of negative ethylation interference did not strictly correspond to the DIS loop: modification of the first three phosphate groups in the loop ($_{271}$GpApApG$_{274}$) was neutral toward dimerization, and interference extended to the first two phosphates of the 3′ strand of the DIS stem ($_{281}$CpApG$_{283}$) (Figs. 2A and 7).

The pattern of ethylation interference observed with DIS RNA of subtype A was similar to that of subtype B, with two noticeable exceptions (Figs. 2B and 7). First, the negative effect of phosphate ethylation on dimerization was limited to the first phosphate of the 3′ strand of the DIS stem ($_{281}$CpA$_{282}$). More surprisingly, ethylation of the 5′ part of the DIS loop ($_{271}$GpApGpG$_{274}$) clearly enhanced dimerization (Figs. 2B and 7). This positive interference was reproducibly observed.

Because DIS RNAs of subtypes A and B differ by the purine at the second position in the loop, and by the self-complementary sequence (Fig. 1A), we tested the effect of phosphate ethylation on the dimerization of RNAs containing either a DIS of subtype F or the nonnatural AGGCGCGCA sequence in the loop. These two RNAs displayed very similar interference patterns

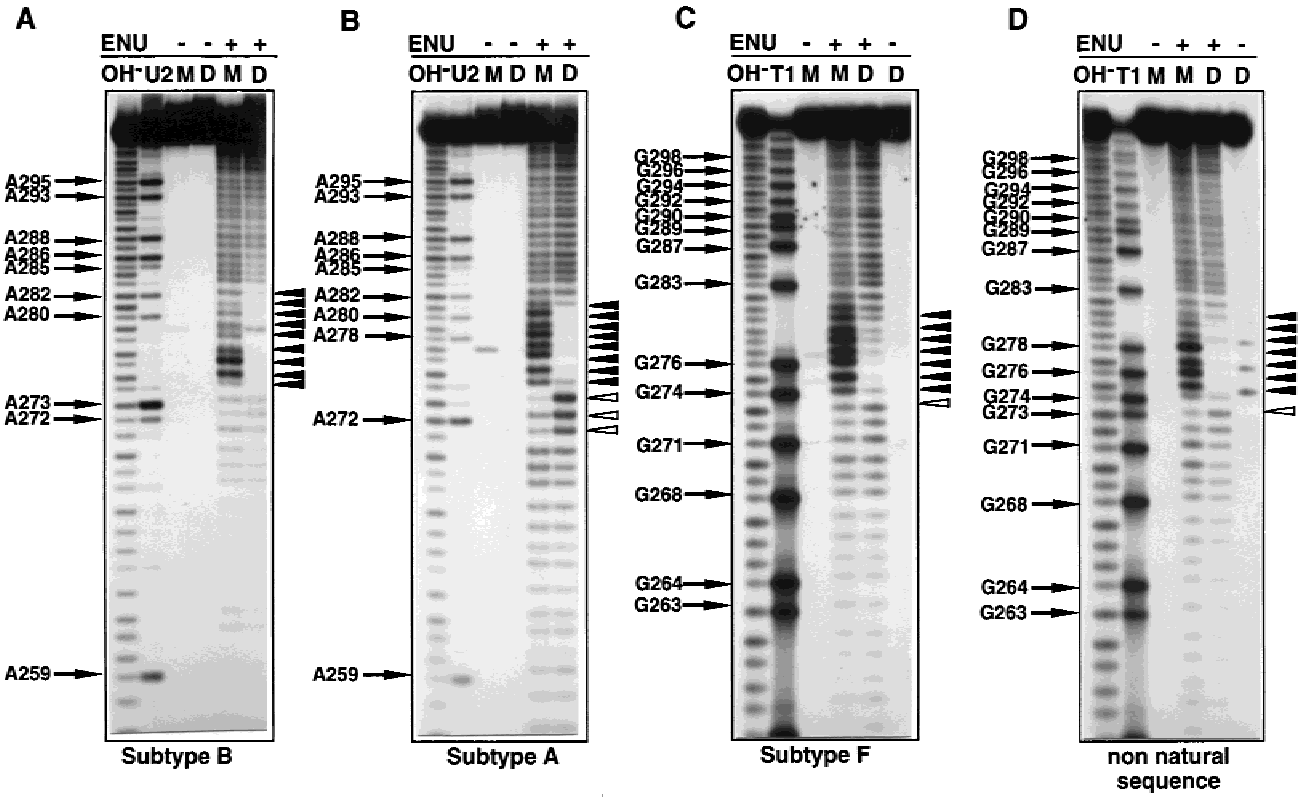

FIGURE 2. Effect of phosphate ethylation on DIS RNA dimerization. 5′-end labeled RNA of subtypes B (**A**), A (**B**), F (**C**), and RNA with the nonnatural loop sequence (**D**) were modified with ENU and submitted to dimerization. M and D refer to RNA extracted from the monomer and dimer bands, respectively, and + and − to RNA treated or untreated with ENU, respectively. Lanes labeled OH, U2, and T1 correspond to alkaline treatment, digestion with RNase U2, and digestion with RNase T1, respectively. Positive and negative interferences are indicated by open and closed arrowheads, respectively.

(Figs. 2C and 2D). In subtype F RNA, ethylation of the phosphates 3′ to nt 275–279 strongly prevented dimerization, whereas moderate negative interference was observed upon modification of phosphates 274, 280, and 281. On the other hand ethylation of phosphates 271, 272, and 273 slightly enhanced dimerization (Fig. 2C). In the nonnatural RNA, weak positive interference was restricted to modification of phosphates 272 and 273. Weak negative interference was observed upon ethylation of phosphate 274, whereas modifications of phosphates 275–280 almost totally abolished dimerization (Fig. 2D).

Lead(II)-induced cleavage

The effects of phosphate ethylation on RNA dimerization might have several nonexclusive origins, including alteration of the loop structure, disruption of metal-ion binding, or of hydrogen bonding involving the nonbridging phosphate oxygens. Given the importance of Mg^{2+} in the dimerization of HIV-1 MAL RNA (Marquet et al., 1994), we used lead(II)-induced cleavage to compare binding of metal ions to the four DIS RNAs. A similar study was previously performed in the context of RNAs encompassing the complete 5′ untranslated region of HIV-1 (Paillart et al., 1997). Using shorter RNAs in this study allows better comparison between the RNAs and, by comparison with previous data, evaluation of the effects of the sequences out of the DIS stem-loop.

In the presence of 2 mM Pb^{2+}, subtype A RNA was cleaved at a single position, at the $_{272}$ApG$_{273}$ step (Figs. 3 and 7). Increasing the Pb^{2+} concentration to 5 mM strongly enhanced this cut and revealed weaker cleavage sites at $_{254}$GpCpA$_{256}$, $_{289}$GpGpCpG$_{292}$, and $_{300}$CpG$_{301}$. These sites are all located in or immediately adjacent to internal loops or noncanonical base pairs (Fig. 1A). A very similar cleavage pattern was obtained with the nonnatural DIS RNA, except that the lead-induced cleavage 3′ to A_{272} was slightly reduced. By contrast, lead(II) induced no strong cut in the apical loop of DIS RNAs of subtypes B and F (Fig. 3). Indeed, cuts were observed at $_{271}$GpApA$_{273}$ in the presence of 5 mM Pb^{2+}, but they were weaker than the cut at $_{300}$CpG$_{301}$ (Fig. 3). The purine at the second position in the DIS loop (R_{273}) thus appeared as a crucial determinant of metal binding.

Phosphorothioate interference

Random incorporation of thiophosphates during RNA synthesis is structurally less perturbing than phosphate ethylation, because it does not neutralize a negative charge or introduce a bulky group on the phosphate. Substitution of sulfur for oxygen weakens hydrogen bonding and binding of Mg^{2+} (Pecoraro et al., 1984; Christian & Yarus, 1993; Chanfreau & Jacquier, 1994), and further, phosphate and thiophosphate groups bind "soft" metal cations such as Mn^{2+} with almost the same affinity (Pecoraro et al., 1984; Christian & Yarus, 1993). Thus, we tested the effect of phosphorothioate incorporation on the dimerization of DIS RNAs of subtypes A and B in the presence of Mg^{2+} or a mixture of Mg^{2+} and Mn^{2+}. We limited our investigations to these two RNAs because they gave the most pronounced differences in the ethylation interference and lead-induced cleavage experiments. The effects of phosphorothioate incorporation 5′ to As, Gs, Us, and Cs were tested separately by synthesizing four pools of modified RNAs (see Materials and Methods).

When looking at dimerization of subtype B RNA, no interference of thiophosphates was detected (data not shown). Similarly, introduction of thiophosphates 5′ to As and Cs had no effect on dimerization of subtype A RNA (data not shown). However, substitution of thiophosphates for phosphates 5′ to Gs negatively interfered with dimerization of this RNA (Figs. 4 and 7). Specifically, thiophosphates 5′ to G_{273}, G_{274}, and G_{276} affected dimerization, although at different levels. The negative interference of thiophosphates 5′ to G_{273} or G_{274} was strong, but not complete. In addition, partial substitution of Mn^{2+} for Mg^{2+} restored dimerization of A subtype DIS RNA containing thiophosphates 5′ to G_{273} and G_{274}. This Mn^{2+} rescue was either partial (G_{273}) or total (G_{274}) (Fig. 4). On the other hand, a thiophosphate 5′ to G_{276} totally abolished dimerization of subtype A RNA, even in the presence of Mn^{2+}. The band corresponding to this position is particularly intense in the monomer lanes. Because no band was observed at this level in the control lanes, this band actually corresponds to an iodine-induced cleavage (Fig. 4). In addition, we checked that iodine did not induce cleavage of RNA that did not contain thiophosphates (data not shown). Hence, this strong band seems to reflect a preferential thiophosphate incorporation at position 276.

Incorporation of UTPαS into subtype A RNA had no effect, except for substitution of the phosphate 5′ to U_{275}, which strongly enhanced dimerization (Figs. 4 and 7). This positive phosphorothioate interference was also observed in the presence of Mn^{2+}, suggesting that this effect was independent of cation binding.

N7-G methylation interference

Beside phosphates, the N7 position of purines might also be involved in noncanonical interactions and metal-ion binding. Thus, we tested the influence of N7-G methylation by DMS on the dimerization of DIS RNAs of subtypes A and B. Interference of the modification of N7-A positions could not be tested because modification of these positions by diethylpyrocarbonate induces base opening (Ehresmann et al., 1987).

Methylation of N7-G positions interferes with dimerization of DIS RNAs of both subtypes (Figs. 5 and 7). Modification of G_{274} and G_{276} strongly reduced dimer-

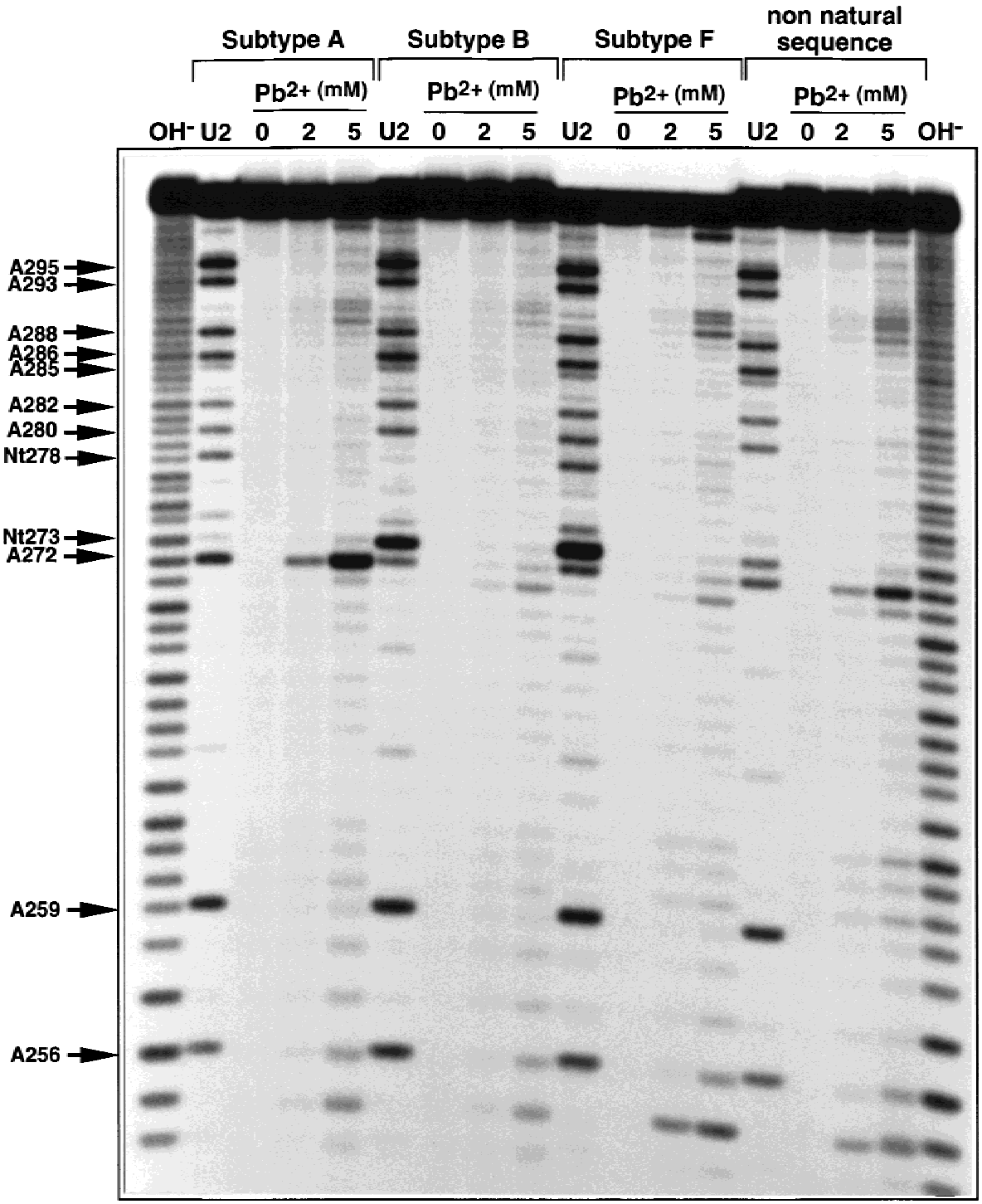

FIGURE 3. Lead-induced cleavage of the RNA dimers. Cleavage of 5′-end labeled RNAs was for 5 min in the presence of 2 or 5 mM lead acetate. A control without lead(II) was run in parallel.

ization of subtype A RNA, whereas modification of G_{271}, G_{273}, and G_{283} was neutral. Thus, the N7 position of G_{273} is not involved in the binding of Mg^{2+} and Pb^{2+}. Similarly, methylation of G_{271} and G_{283} did not affect dimerization of subtype B RNA, whereas modification of G_{274}, G_{276} and G_{278} significantly decreased the efficiency of the process.

N1-A methylation interference

Finally, we analyzed the effect of N1-A methylation by DMS on RNA dimerization. As previously observed with a large RNA encompassing the complete 5′ untranslated region of HIV-1 MAL (Skripkin et al., 1994), methylation of A_{278}, located in the 6-nt self-complementary

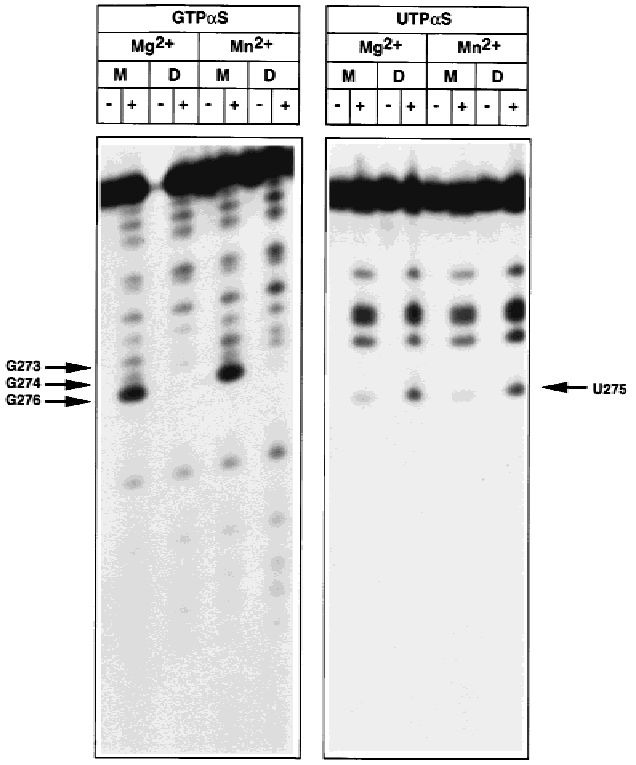

FIGURE 4. Effect of GTPαS and UTPαS incorporation on dimerization of subtype A RNA. Dimerization was conducted in the presence of Mg^{2+} or a mixture of Mn^{2+} and Mg^{2+}. Scission of the 3′-end labeled RNA at the thiophosphate positions was induced by iodine treatment (+). An untreated control (−) was run in parallel. M and D refer to RNA extracted from the monomer and dimer bands, respectively.

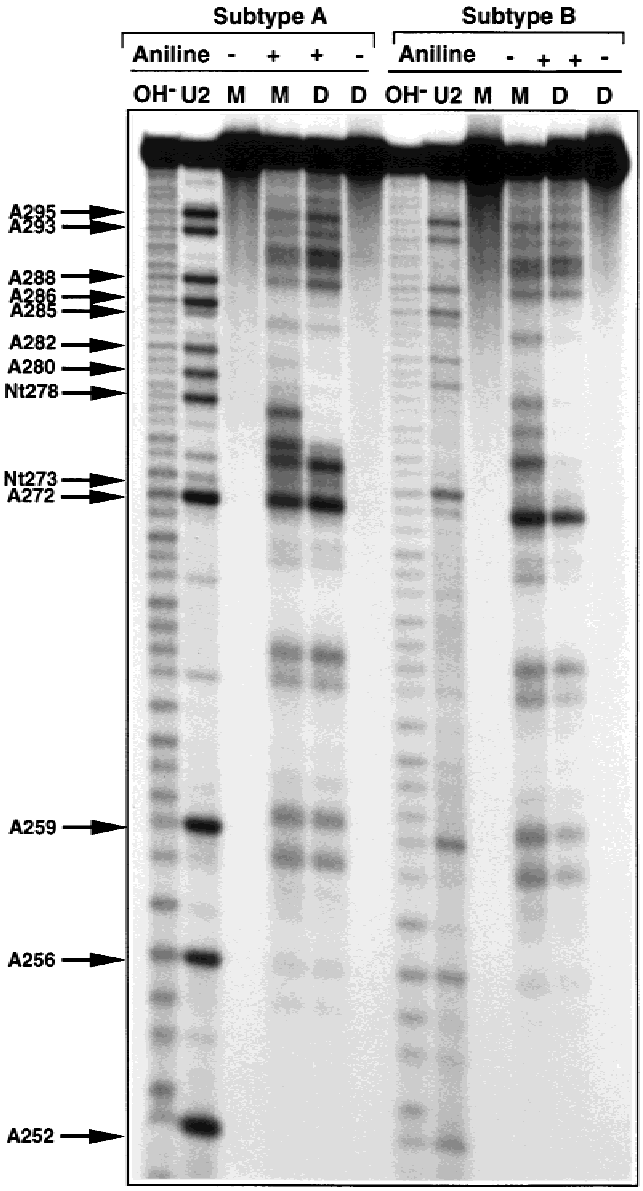

FIGURE 5. Interference of N7-G methylation on RNA dimerization. Scission of 5′-end labeled RNAs of subtypes A and B at methylated Gs was performed by treatments with $NaBH_4$ and aniline (+). Modified RNAs treated with $NaBH_4$, but not with aniline (−), were run in parallel as controls. M and D refer to RNA extracted from the monomer and dimer bands, respectively.

sequence, inhibited dimerization of DIS RNA of subtype A (Figs. 6 and 7). On the other hand, methylation of A_{280} was neutral, whereas that of A_{272} significantly enhanced the dimerization process (Figs. 6 and 7). Methylation of A_{272} and A_{280} had limited or no effects on dimerization of subtype B RNA. In RNA of subtype B, A_{273} was only marginally modified, thus making difficult the interpretation of the results concerning this position (Fig. 6). RNA dimerization was either unaffected or slightly enhanced upon methylation of A_{273}.

DISCUSSION

Comparison of the dimerization of HIV-1 RNAs corresponding to subtypes of group M, by interference techniques and lead-induced cleavage, revealed both common and specific features (Fig. 7). First, only modifications in the DIS loop and in the adjacent nucleotides of the DIS stem interfere with RNA dimerization. This observation is consistent with our early interference study of the modifications of the Watson–Crick positions of large RNAs encompassing the complete 5′ untranslated region of HIV-1 RNA, and with the concept that HIV-1 RNA dimerization proceeds through formation of a kissing-loop complex (Skripkin et al., 1994). Indeed, ethylation of any phosphate of the 6-nt self-complementary sequence of the DIS almost totally abolished dimerization. In principle, ethylation interference may result from weakening of hydrogen bonding or coordination of metal ion or from perturbation of the tertiary structure (Schnitzer & von Ahsen, 1997). However, given the number of modifications (7–9) that prevent dimerization of the RNAs, it is likely that the negative interference at least partially originates from perturbation of the DIS loop structure. From previous mutagenesis and oligonucleotide-mediated inhibition studies

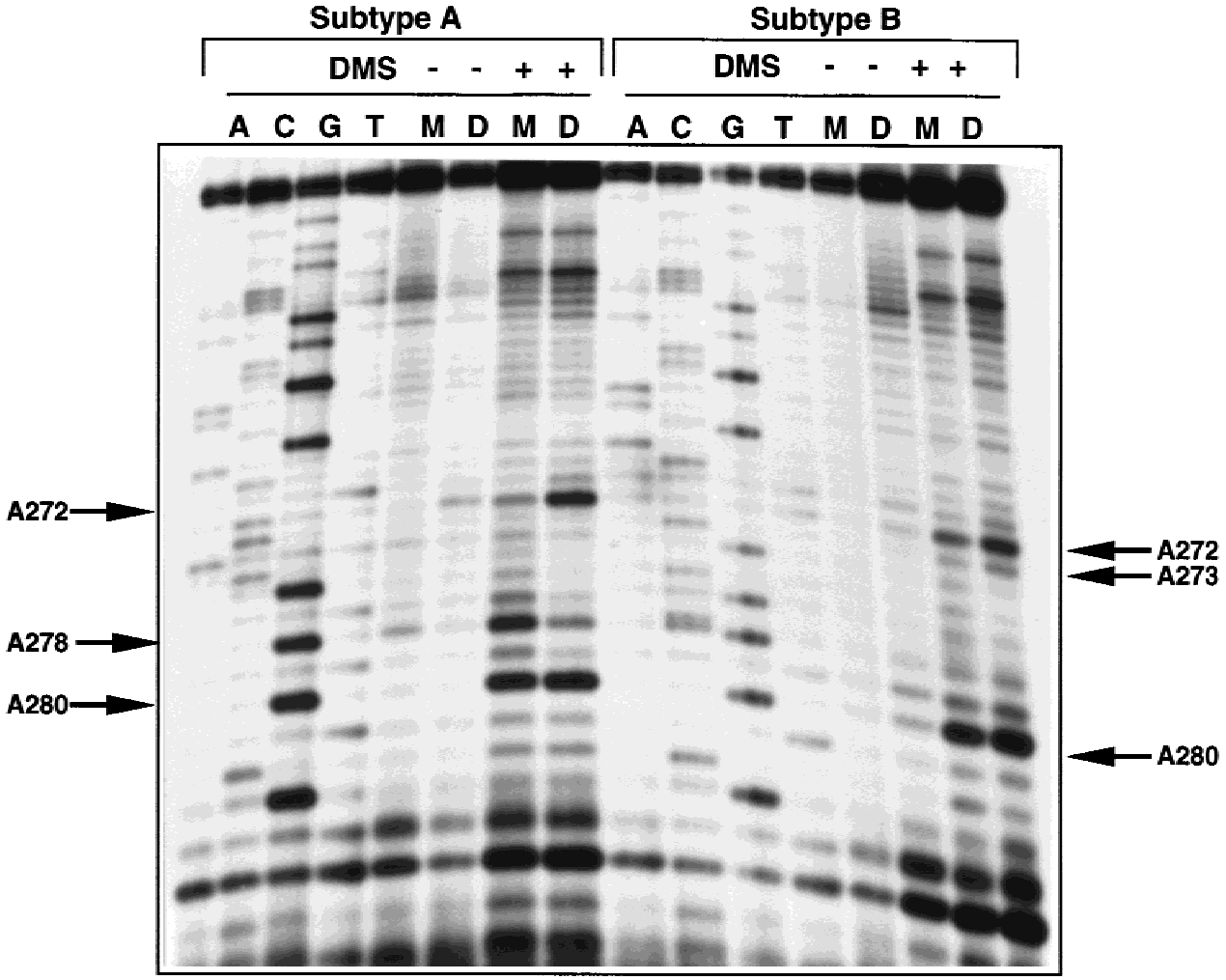

FIGURE 6. Interference of N1-A methylation on RNA dimerization. Methylated As present in RNA extracted from the monomer (M) and dimer (D) were detected by primer extension. As a control (−), primer extension was performed on unmodified RNA. Dideoxy sequencing lanes were run in parallel.

(Skripkin et al., 1996; Paillart et al., 1997), we concluded that the structure of the DIS loop of subtype A was important for the dimerization process. The present study stresses this conclusion and extends it to other HIV-1 subtypes.

A feature common to DIS RNAs of subtypes A and B is the inhibition of dimerization upon N7-methylation of the Gs in the self-complementary sequence of the DIS loop. This observation suggests that the structure of the kissing loop is very compact, or that these positions are involved in noncanonical interactions. This result is in keeping with the three-dimensional model of the DIS that we previously constructed to account for our solution of structural probing and physicochemical data (Paillart et al., 1997). We proposed that in the subtype A RNA, the guanosines at the second position of the loop in each monomer (G_{273} and G_{273}') lie into the major groove of the helix formed by the pairing of the two self-complementary sequences and form intermolecular base triples with the C_{277}–G_{276}' and G_{276}–C_{277}' base pairs (Fig. 1C). (The prime refers to the second RNA molecule in the dimer). The existence of such base triples is supported by preliminary NMR data of the subtype A dimer (Dardel et al., 1998). In the case of DIS RNA of subtype B, we proposed that A_{273} and A_{273}' might form intramolecular base triples with G_{276}–C_{277}' and C_{277}–G_{276}', respectively. These base triples are not present in the NMR structure model of the subtype B dimer (Mujeeb et al., 1998). In this structure, the kissing-loop complex is intrinsically bent and the DIS stems converge towards the major groove of the helix formed by the self-complementary sequences. The resulting narrow major groove might explain the effect of N7-G methylation on dimerization. Alternatively, the NMR structure, which was solved in the absence of divalent cations, might correspond to an partially unfolded intermediate between a more stable kissing-loop complex and an extended duplex (Mujeeb et al., 1998).

With both subtype A and B RNAs, we observed that N1-methylation of A_{272} and A_{280} (and for type B

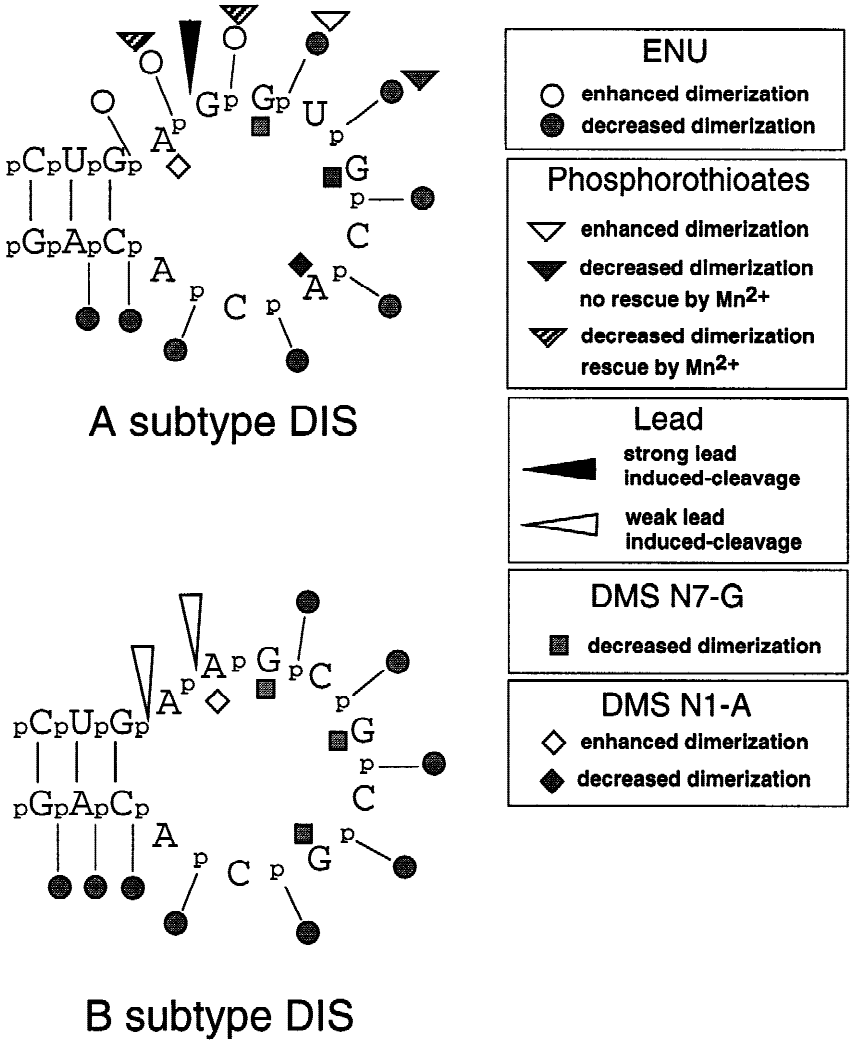

FIGURE 7. Summary of the interference and lead-cleavage experiments conducted on the RNAs of subtypes A and B. Modifications that were neutral towards dimerization are not indicated.

RNA, A_{273}) did not prevent dimerization. This result was rather unexpected, because we proposed a noncanonical interaction between these two adenosines involving the N1 position of A_{272} (Paillart et al., 1997) (Fig. 1C). A noncanonical interaction between these 2 nt is supported by chemical probing and phylogenetic analysis (Paillart et al., 1996b, 1997), and an NMR study conducted on subtype A RNA (Dardel et al., 1998). Our interference results suggest that, either this A–A interaction is not important for dimerization, or, most likely, it is not a *trans* base pair between the Watson–Crick positions of A_{272} and the Hoogsteen sites of A_{280} (Fig. 1C). In the NMR structure of subtype B RNA, this A–A pairing was not observed (Mujeeb et al., 1998). The NMR structure is thus in agreement with our results of N1-A methylation interference, but hard to reconcile with the strong conservation of A_{280} (Paillart et al., 1996b) and with our chemical probing data (Paillart et al., 1997).

Beside these common features, our results also highlight distinct characteristics of the RNA dimers of subtypes A and B. Ethylation of the phosphates in the $_{271}$GpApRpG$_{274}$ sequence enhanced dimerization of subtype A RNA, whereas it was neutral towards dimerization of subtype B RNA. Indeed, the positive effect on dimerization of subtype A RNA is likely because of neutralization of the phosphate negative charge by the ethyl group. Consistent with the fact that dimerization of subtype A RNA requires divalent cations, unlike that of subtype B RNA (Lodmell et al., 1998), we propose that Mg^{2+} prevents electrostatic repulsion between these phosphates in the type A RNAs.

The strong lead-induced cleavage at $_{272}$ApR$_{273}$ in subtype A RNA and in the nonnatural RNA, but not in DIS RNAs of subtypes B and F, suggests the existence of a strong ion-binding site in the former RNAs. The weak cleavages observed in the DIS loop of the latter RNAs, as well as cleavages in the internal loops, likely reflect hydrolysis of flexible RNA regions by unbound Pb^{2+} ions (Gornicki et al., 1989; Paillart et al., 1997). This interpretation is in agreement with the results of the phosphorothioate interference experiments. In sub-

type A RNA, substitution of the pro(R) nonbridging oxygen by sulfur 5′ to G_{273}, G_{274}, and G_{276} inhibited dimerization. This negative interference might be due to weakening of hydrogen bonding or/and metal-ion binding. However, dimerization of RNAs containing a phosphorothioate 5′ to G_{273} and G_{274} in the presence of Mn^{2+} is a strong indication that these phosphate groups are directly involved in metal-ion binding (Christian & Yarus, 1993; Chanfreau & Jacquier, 1994). The interaction between the phosphate group 5′ to G_{276} and the metal ion is likely indirect, maybe via the hydration sphere of the latter.

On the other hand, we observed no interference of phosphorothioates on dimerization of subtype B RNA. This result indicates either that the phosphate groups of this RNA are not involved in Mg^{2+} binding, or that Mg^{2+} binding is not crucial for dimerization of subtype B RNA. The absence of strong lead-induced cleavage in this RNA favors the first interpretation, although one cannot totally exclude that Pb^{2+} actually binds to the RNA with a geometry that does not allow efficient hydrolysis. Additional lead-induced cleavage data obtained with subtype F RNA and a nonnatural sequence show that the identity of R_{273} (A or G) is a major determinant in the binding of metal ions.

We used three-dimensional modeling in an attempt to rationalize our data in structural terms. First, we looked for a conformation of the A_{272}–A_{280} noncanonical base pair that would explain the complete set of probing and chemical interference data. Molecular modeling showed that these nucleotides might exist in a sheared conformation, the N-3 and H-2 positions of A_{272} interacting with the NH_2-6 and N-7 positions of 280 (Fig. 8). This conformation accounts for the fact that the N-7 position of A_{272} is more accessible to chemical modification than that of A_{280} (Paillart et al., 1997). The N1 positions of those As are not involved in hydrogen bonding; however, the three-dimensional model shows that while N1-A_{280} lies at the surface of the shallow groove, N1-A_{272} is buried inside the deep groove (Fig. 8B). Thus, this model also accounts for the modification pattern of A_{272} and A_{280} by DMS (Baudin et al., 1993; Paillart et al., 1997). In addition, unlike the previously proposed A–A base pair (Fig. 1C), the sheared conformation is in keeping with the chemical interference data. Indeed, because the N-1 positions of A_{272} and A_{280} are not involved in hydrogen bonding, our model is consistent with the observation that these positions can be methylated without interfering with dimerization. The present model is also consistent with the NMR data obtained on a DIS RNA of subtype A (Dardel et al., 1998) that reveal continuous stacking at the junction between the DIS loop and the 3′ part of the stem.

In the next step, we used Brownian-dynamics simulation of cation diffusion to identify potential Mg^{2+}-binding sites in the DIS kissing-loop complex (Hermann & Westhof, 1998). The three-dimensional model of the DIS of subtype A was used as a target for cations diffusing under the influence of Brownian motion within the electrostatic field of the RNA. Using cationic spheres of 2- or 2.5-Å radius, we identified a potential Mg^{2+}-binding site in each DIS loop, in the sharp turn delineated by the $_{273}$GGUG$_{276}$ sequence (Fig. 8b). Thus, the predicted Mg^{2+} location is at the center of the pocket identified by the phosphorothioate interference experiments. Binding of divalent cations in the sharp turn in another kissing-loop complex has been proposed (Lee & Crothers, 1998). However, our experimental data, together with the Brownian-dynamics simulation, represent the first strong evidence of such binding. Furthermore, our results with the DIS of subtypes B and F indicate that subtle differences in the loop sequence affect Mg^{2+} binding. Indeed, R_{273} probably promotes different noncanonical interactions in these DIS subtypes that affect the precise conformation of the sharp turn in the DIS loop (Paillart et al., 1997). In keeping with this interpretation, Brownian-dynamics simulation conducted on the DIS of subtype B reveals a reduced probability of cation binding, as compared to the DIS of subtype A (data not shown).

The functional role of the DIS (Berkhout & van Wamel, 1996; Paillart et al., 1996a; Clever & Parslow, 1997; Laughrea et al., 1997; McBride & Panganiban, 1997) suggests that it might be a valuable target for antiretroviral therapy. As phosphate ethylation interference indicates that the structure of the DIS loop is crucial for dimerization, we anticipate that ligands of the DIS loop in the monomeric form might inhibit dimerization by altering its structure. The difference in the fine dimerization mechanism of subtype A and B RNAs should also be taken into account for the design and/or screening of such DIS loop ligands.

MATERIALS AND METHODS

RNA synthesis and labeling

Primers TGGAGTAATACGACTCACTATAGGGTCGACGC AGGACTCGGCTTGC (sense) and TCGCCGCTCTCGCC TCTTGCTG (antisense) were used to amplify the DIS region of plasmids pJCB (Paillart et al., 1994), pDISCompA, pDISA, and pDISComp (Paillart et al., 1997) downstream of the T7 RNA polymerase promoter. in vitro transcription of gel-purified PCR products was used to generate 55-mer RNAs encompassing subtype A, B, and F DIS, as well as the RNA containing the nonnatural AGGCGCGC sequence in the DIS loop, respectively (Fig. 1B). Unmodified RNAs were obtained by in vitro transcription of 1 μg of PCR product in 40 mM Tris-HCl, pH 8.0, 50 mM NaCl, 15 mM $MgCl_2$, 1 mM spermidine, 0.05 mg/mL BSA, 50 mM DTE, 1.2 U/mL RNasine, 0.25 mg/mL of home-purified T7 RNA polymerase, 4 mM of ATP, UTP, and CTP, 1 mM GTP and 5 mM ApG. Incorporation of ApG at the 5′ terminus of the transcripts allowed efficient labeling with [γ-^{32}P]ATP and T4 polynucleotide kinase. Phosphorothioate-containing RNAs were obtained by in vitro transcription in the

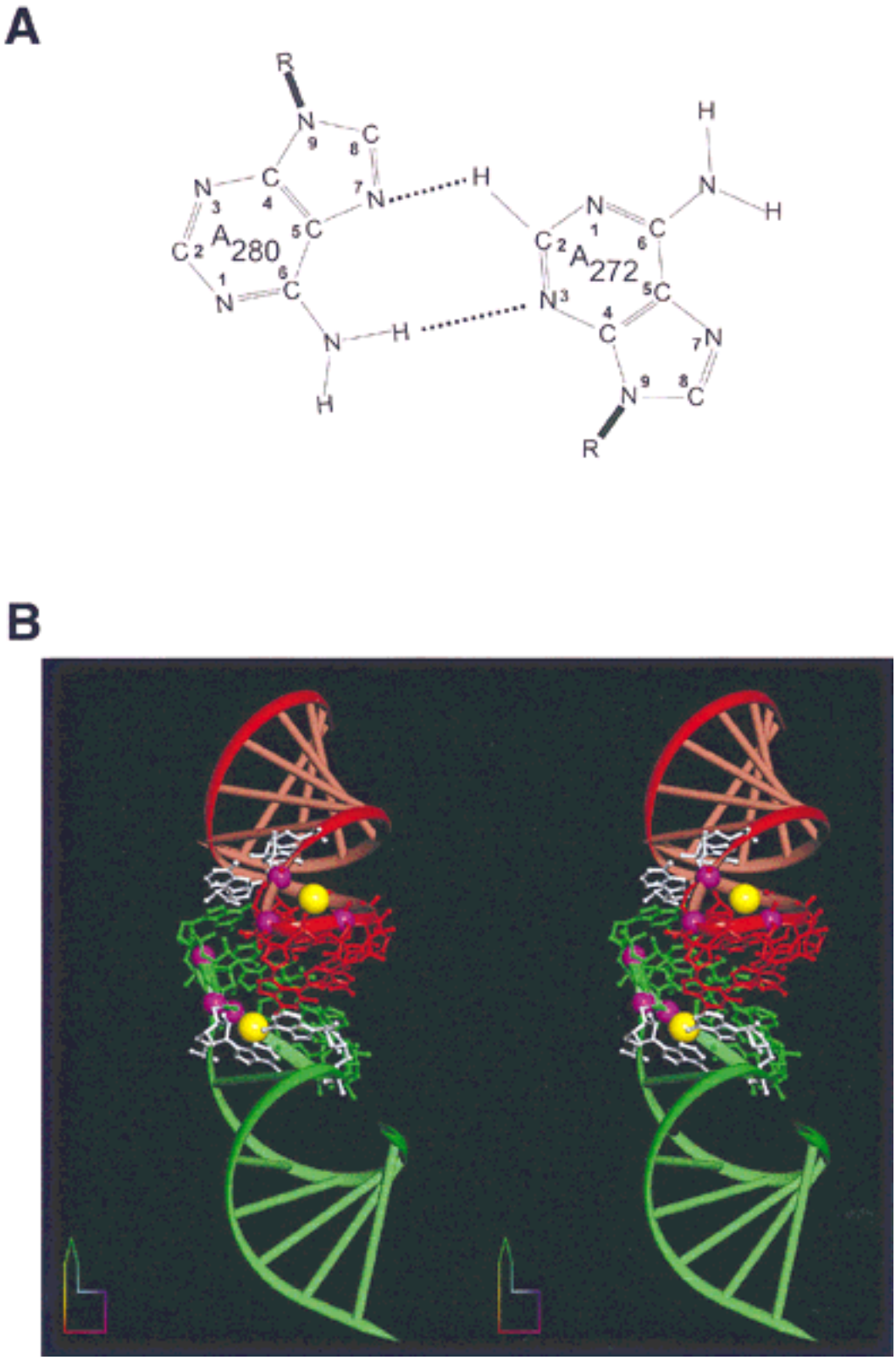

FIGURE 8. **A**: Schematic drawing of the proposed sheared A–A base pair. **B**: Stereo view of the three-dimensional model of the kissing-loop complex formed by the DIS of subtype A. The two monomers are in green and red, respectively, except the A–A base pairs, which are shown in white. The DIS stems are shown as ribbons and nucleotides are shown in the loops. The small purple spheres correspond to the phosphates displaying negative interference upon phosphorothioate substitution. The yellow spheres correspond to the magnesium-binding sites identified by Brownian-dynamics simulation using cationic spheres with a radius of 2 (top site) or 2.5 Å (bottom site).

presence of 3 mM of the four unmodified NTP and 0.5 mM of one NTPαS. The modified RNAs were 3′-end labeled overnight at 4 °C with [^{32}P]pCp and T4 RNA ligase.

RNA modifications

Modification of 55-mer RNAs (6.5 μM) with 0.5 vol of *N*-ethyl-*N*-nitrosourea (ENU)-saturated ethanol was for 1 min at 90 °C in 50 mM sodium cacodylate, pH 6.5, 50 mM KCl, 1 mM EDTA. The reaction was stopped by precipitation with 0.3 M sodium acetate and 3 vol of ethanol. The pellet was redissolved in water and used in the dimerization reaction. An aliquot of RNA was submitted to the same treatment except that ENU was omitted.

In methylation experiments, RNA was first denatured for 1 min at 90 °C in 50 mM sodium cacodylate, pH 6.5, 50 mM KCl, 1 mM EDTA and modified either 10 min at room temperature or 1 min at 90 °C with 0.01 vol dimethyl sulfate (DMS). Modification was stopped by precipitation with sodium acetate and ethanol, the pellet was redissolved in water and used in the dimerization reaction. In parallel, an aliquot of RNA was submitted to the same treatment except for omission of DMS.

Dimerization of the modified RNAs and identification of the modifications interfering with this process

Modified labeled RNAs (5.5 pmol, $>5 \times 10^5$ cpm) were denatured in 8 μL water for 2 min at 90 °C. After addition of 2 μL buffer (final concentrations: 50 mM sodium cacodylate, pH 6.5, 50 mM KCl, 0.005–0.1 mM $MgCl_2$), dimerization was allowed to proceed for 20 min at 37 °C. For each RNA, the $MgCl_2$ concentration was adjusted to obtain ~50% dimer (e.g., 5 μM for B subtype DIS RNA and 0.1 mM for A subtype DIS RNA). In the manganese rescue experiments, one-third of the Mg^{2+} cations were replaced by Mn^{2+}. The monomeric and dimeric forms were separated by electrophoresis at 4 °C on a 10% nondenaturing polyacrylamide gel in 45 mM Tris-borate, pH 8.3, supplemented with the same $MgCl_2$ concentration as the dimerization buffer. The corresponding bands were identified by autoradiography, cut, and the RNA was eluted overnight at 4 °C in a buffer containing 500 mM ammonium acetate, 0.1 mM EDTA, and $\frac{1}{6}$ vol phenol and precipitated.

Identification of the modifications interfering with dimerization

Phosphorothioates were cleaved for 2 min at room temperature by adding $\frac{1}{10}$ vol of 10 mM I_2 in ethanol to the monomeric and dimeric RNAs dissolved in 5 μL water. A control reaction was performed in parallel by adding $\frac{1}{10}$ vol ethanol instead of the iodine solution.

Ethylated phosphates were identified by incubating the modified RNA in 10 μL Tris-HCl 0.1 M, pH 9.0 at 50 °C for 10 min. As a control, unmodified RNA was submitted to the same treatment.

To identify methylated N7-G positions, modified RNA dissolved in 10 μL water was first incubated in the dark at 4 °C for 10 min with 8 μg/μL $NaBH_4$. The reaction mixture was ethanol precipitated and split into two parts. The first was used as a control, and the second was treated with aniline for 10 min at 60 °C in the dark.

Lead-induced cleavage

RNA was allowed to dimerize for 20 min at 37 °C in 25 mM HEPES, pH 7.0, 300 mM potassium acetate. Cleavage was induced by addition of 2 or 5 mM lead(II) acetate and stopped after 5 min by addition of 29 mM EDTA and precipitation.

In all experiments, RNA fragments were separated by denaturing gel electrophoresis on a 15% acrylamide gel and quantified with a BioImager BAS2000 (Fuji).

Three-dimensional molecular modeling

Programs used for construction of the type A DIS model were previously described (Westhof, 1993). The assembly and connection of the secondary elements into a three-dimensional fold was realized on a Silicon Graphics computer station, using the program MANIP (Massire & Westhof, 1999). Finally, the generated models were subjected to restrained least-squares refinement (Konnert & Hendrickson, 1980) with the programs NUCLIN and NUCLSQ (Westhof et al., 1985) to ensure geometry and stereochemistry with allowed distances between interacting atoms and to avoid steric conflicts. The potential metal-ion binding sites in the DIS three-dimensional model were determined by Brownian-dynamics simulations of cation diffusion, as recently described (Hermann & Westhof, 1998). The color views were generated with the program DRAWNA (Massire et al., 1994).

ACKNOWLEDGMENTS

This work was supported by grants from the “Agence Nationale de Recherches sur le SIDA” (ANRS). J.S.L. is a fellow of the ANRS.

Received May 10, 1999; returned for revision May 26, 1999; revised manuscript received June 4, 1999

REFERENCES

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, Chien YH, Chattopadkyay S, Vogt PK, Gardner MB, Davidson N. 1978. High-molecular weight RNAs of AKR, NZB and wild mouse viruses and avian reticuloendotheliosis virus all have similar dimer structures. *J Virol 25*:888–896.

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

Chanfreau G, Jacquier A. 1994. Catalytic site components common to both splicing steps of a group II intron. *Science 266*:1383–1387.

Christian EL, Yarus M. 1993. Metal coordination sites that contribute to structure and catalysis in the group I intron from *Tetrahymena*. *Biochemistry 32*:4475–4480.

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

Clever JL, Wong ML, Parslow TG. 1996. Requirements for kissing-loop-mediated dimerization of human immunodeficiency virus RNA. *J Virol 70*:5902–5908.

Dardel F, Marquet R, Ehresmann C, Ehresmann B, Blanquet S. 1998. Solution studies of the dimerization initiation site of HIV-1 genomic RNA. *Nucleic Acids Res 26*:3567–3571.

Eckstein F, Gish G. 1989. Phosphorothioates in molecular biology. *Trends Biochem Sci 14*:97–100.

Ehresmann C, Baudin F, Mougel M, Romby P, Ebel JP, Ehresmann B. 1987. Probing the structure of RNA in solution. *Nucleic Acids Res 15*:9109–9128.

Gornicki P, Baudin F, Romby P, Wiewiorowski M, Krzyzosiak JW, Ebel J-P, Ehresmann C, Ehresmann B. 1989. Use of lead(II) to probe the structure of large RNAs. Conformation of the 3′ terminal domain of *E. coli* 16S rRNA and its involvement in building the tRNA binding sites. *J Biomol Struct Dyn 5*:971–984.

Haddrick M, Lear AL, Cann AJ, Heaphy S. 1996. Evidence that a kissing-loop structure facilitates genomic RNA dimerization in HIV-1. *J Mol Biol 259*:58–68.

Harrison GP, Miele G, Hunter E, Lever AML. 1998. Functional analysis of the core human immunodeficiency virus type 1 packaging signal in a permissive cell line. *J Virol 72*:5886–5896.

Hermann T, Westhof E. 1998. Exploration of metal ion binding sites in RNA folds by Brownian-dynamics simulations. *Structure 6*:1303–1314.

Konnert JH, Hendrickson WA. 1980. A restrained-parameter thermal-factor refinement procedure. *Acta Crystallogr 36*:344–350.

Korber B, Hahn B, Foley B, Mellors JW, Leitner T, Myers G, McCutchan F, Kuiken C (eds.). 1997. *Human retroviruses and AIDS. A compilation and analysis of nucleic acid and amino acid sequences.* Los Alamos, New Mexico: Theoretical Biology and Biophysics Group T-10.

Kung HJ, Hu S, Bender W, Bailey JM, Davidson N, Nicolson MO, McAllister RM. 1976. RD-114, baboon, and woolly monkey viral RNAs compared in size and structure. *Cell 7*:609–620.

Laughrea M, Jetté 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.

Laughrea M, Jetté L. 1996a. HIV-1 genome dimerization: Formation kinetics and thermal stability of dimeric HIV-1 LAI RNAs are not improved by the 1–232 and 296–790 regions flanking the kissing-loop domain. *Biochemistry 35*:9366–9374.

Laughrea M, Jetté L. 1996b. Kissing-loop model of HIV-1 genome dimerization: HIV-1 RNA can assume alternative dimeric forms, and all sequences upstream or downstream of hairpin 248–271 are dispensable for dimer formation. *Biochemistry 35*:1589–1598.

Laughrea M, Jetté L. 1997. HIV-1 genome dimerization: Kissing-loop hairpin dictates wether nucleotides downstream of the 5′ splice junction contribute to loose and tight dimerization of human immunodeficiency virus RNA. *Biochemistry 36*:9501–9508.

Laughrea M, Jetté L, Mak J, Kleiman L, Liang C, Wainberg MA. 1997. Mutations in the kissing-loop hairpin of human immunodeficiency virus type 1 reduce viral infectivity as well as genomic RNA packaging and dimerization. *J Virol 71*:3397–3406.

Lee AJ, Crothers DM. 1998. The solution structure of an RNA loop-loop complex: The *ColE1* inverted loop sequence. *Structure 6*:993–1005.

Lodmell JS, Paillart JC, Mignot D, Ehresmann B, Ehresmann C, Marquet R. 1998. Oligonucleotide-mediated inhibition of genomic RNA dimerization of HIV-1 strains MAL and LAI: A comparative analysis. *Antisense Nucleic Acid Drug Dev 8*:517–529.

Marquet R, Paillart J-C, Skripkin E, Ehresmann C, Ehresmann B. 1994. Dimerization of human immunodeficiency virus type 1 RNA involves sequences located upstream of the splice donor site. *Nucleic Acids Res 22*:145–151.

Marquet R, Paillart J-C, Skripkin E, Ehresmann C, Ehresmann B 1996. Localization of the dimerization initiation site of HIV-1 genomic RNA and mechanism of dimerization. In: Sarma RH, Sarma MH, eds. *Biological structure and dynamics.* Albany, New York: Adenine Press. pp 61–72.

Massire C, Gaspin C, Westhof E. 1994. DRAWNA: A program for drawing schematic views of nucleic acids. *J Mol Graph 12*:201–207.

Massire C, Westhof E. 1999. MANIP: An interactive tool for modelling RNA. *J Mol Graph & Modelling*. In press.

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.

McBride MS, Panganiban AT. 1997. Position dependence of functional hairpins important for human immunodeficiency virus type 1 RNA encapsidation in vivo. *J Virol 71*:2050–2058.

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 BP, Paoletti J. 1996b. 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, Fossé P, Paoletti J. 1996a. A kissing complex together with a stable dimer is involved in the HIV-1_{Lai} RNA dimerization process in vitro. *Biochemistry 35*:5075–5082.

Muriaux D, Girard P-M, Bonnet-Mathonière B, Paoletti J. 1995. Dimerization of HIV-1Lai RNA at low ionic strength. An autocomplementary sequence in the 5′ leader region is evidenced by an antisense oligonucleotide. *J Biol Chem 270*:8209–8216.

Murti KG, Bondurant M, Tereba A. 1981. Secondary structural features in the 70 S RNAs of Moloney murine leukemia and Rous sarcoma viruses as observed by electron microscopy. *J Virol 37*:411–419.

Paillart J-C, Berthoux L, Ottmann M, Darlix J-L, Marquet R, Ehresmann B, Ehresmann C. 1996a. A dual role of the putative RNA dimerization initiation site of human immunodeficiency virus type 1 in genomic RNA packaging and proviral DNA synthesis. *J Virol 70*:8348–8354.

Paillart JC, Marquet R, Skripkin E, Ehresmann B, Ehresmann C. 1994. Mutational analysis of the bipartite dimer linkage structure of HIV-1 genomic RNA. *J Biol Chem 269*:27486–27493.

Paillart JC, Marquet R, Skripkin E, Ehresmann C, Ehresmann B. 1996b. Dimerization of retroviral genomic RNAs: Structural and functional implications. *Biochimie 78*:639–653.

Paillart J-C, Skripkin E, Ehresmann B, Ehresmann C, Marquet R. 1996c. A loop-loop "kissing" complex is the essential part of the dimer linkage of genomic HIV-1 RNA. *Proc Natl Acad Sci USA 93*:5572–5577.

Paillart JC, Westhof E, Ehresmann C, Ehresmann B, Marquet R. 1997. Non-canonical interactions in a kissing-loop complex: The dimerization initiation site of HIV-1 genomic RNA. *J Mol Biol 270*:36–49.

Pecoraro VL, Hermes JD, Cleland WW. 1984. Stability constants of Mg^{2+} and Cd^{2+} complexes of adenine nucleotides and thionucleotides and rate constants for formation and dissociation of MgATP and MgADP. *Biochemistry 23*:5262–5271.

Schnitzer W, von Ahsen U. 1997. Identification of specific Rp-phosphate oxygens in the tRNA anticodon loop required for ribosomal P-site binding. *Proc Natl Acad Sci USA 94*:12823–12828.

Skripkin E, Paillart JC, Marquet R, Blumenfeld M, Ehresmann B, Ehresmann C. 1996. Mechanisms of inhibition of in vitro dimerization of HIV type I RNA by sense and antisense oligonucleotides. *J Biol Chem 271*:28812–28817.

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

St Louis DC, Gotte D, Sanders-Buell E, Ritchey DW, Salminen MO, Carr JK, McCutchan FE. 1998. Infectious molecular clones with the nonhomologous dimer initiation sequences found in different subtypes of human immunodeficiency virus type 1 can recombine and initiate a spreading infection in vitro. *J Virol 72*:3991–3998.

Westhof E. 1993. Modelling the three-dimensional structure of ribonucleic acids. *J Mol Struct 286*:203–210.

Westhof E, Dumas P, Moras D. 1985. Crystallographic refinement of yeast aspartic acid transfer RNA. *J Mol Biol 184*:119–145.