THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 269, No. 44, Issue of November 4, pp. 27486–27493, 1994
Printed in U.S.A.

Mutational Analysis of the Bipartite Dimer Linkage Structure of Human Immunodeficiency Virus Type 1 Genomic RNA*

(Received for publication, July 19, 1994)

Jean-Christophe Paillart, Roland Marquet‡, Eugene Skripkin§, Bernard Ehresmann, and Chantal Ehresmann

From the Unité Propre de Recherches No. 9002 du Centre National de la Recherche Scientifique, Institut de Biologie Moléculaire et Cellulaire, 15 rue R. Descartes, 67084 Strasbourg-cedex, France

The genome of all retroviruses consists in two homologous RNA molecules associated near their 5′ end in a region called the dimer linkage structure. Dimerization of genomic RNA is thought to be important for several functions of the retroviral cycle such as encapsidation, reverse transcription, and translation. In human immunodeficiency virus type 1 (HIV-1), a region downstream of the splice donor site was initially postulated to mediate dimerization. However, we recently showed that the dimerization initiation site is located upstream of the splice donor site and suggested that dimerization may initiate through a loop-loop interaction. Here, we show that single base mutations in the palindromic loop of the dimerization initiation site completely abolish dimerization, while introduction of compensatory mutations restores the process. Furthermore, two single nucleotide mutants that are unable to form homodimers efficiently codimerize, while the wild type RNA and the compensatory mutant, which both form homodimers, are unable to codimerize. These results unambiguously prove the interaction between the palindromic loops of each monomer. By contrast, none of the deletions that we introduced downstream of the splice donor site abolishes dimerization. However, deletions of two purine tracts located in the vicinity of the initiation codon of the *gag* gene significantly decrease the thermal stability of the HIV-1 RNA dimer.

Retroviral genomes consist of two homologous RNA molecules noncovalently linked near their 5′ end (1–4). The functional role of the dimer linkage structure (DLS)[1] remains unclear. However, it is probably crucial since it exists in all retroviruses. It has been proposed to be involved in the regulation of several key steps in the retroviral cycle, such as translation (5, 6), encapsidation (6–11), and recombination during reverse transcription (12–14). Therefore, sequences involved in the dimerization of the retroviral genomes are most likely good targets for sense or antisense agents, and their exact localization is the first step toward potential therapeutic action.

* This work was supported by grants from the Agence Nationale de Recherches sur le SIDA and from the Fondation de la Recherche Médicale. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: UPR 9002 du CNRS, IBMC, 15 rue R. Descartes, 67084 Strasbourg-cedex, France. Tel.: 33-88-41-70-91; Fax: 33-60-22-18.

§ Fellow of the Agence Nationale de Recherches sur le SIDA.

[1] The abbreviations used are: DLS, dimer linkage structure; HIV-1, human immunodeficiency virus type 1; S.D. site, splice donor site; 3′-DLS, elements of the DLS located downstream of the S.D. site; DIS, dimerization initiation site; PCR, polymerase chain reaction.

Dimerization of retroviral RNAs can be induced *in vitro* by the nucleocapsid protein (7) and by cations (15, 16). In human immunodeficiency type I virus (HIV-1), the DLS was first localized within the 100 nucleotides downstream of the splice donor (S.D.) site (7, 15, 16). This proposed DLS is absent from the spliced mRNAs (Fig. 1*A*). From an earlier study of the *in vitro* dimerization of several RNA fragments containing this DLS, we concluded that dimerization of HIV-1 genomic RNA involves non-Watson-Crick interactions. We proposed that it may involve A and G quartets (15). Indeed, several consensus purine tracts RGGARA were found in the putative DLS of all retroviruses (15, 16). More recent studies strongly suggested that G quartets stabilize the dimers formed by short RNAs restricted to this proposed HIV-1 DLS (17, 18) and also a larger RNA (19).

However, we recently showed that dimerization of HIV-1 RNA involves sequences located upstream of the S.D. site, *i.e.* outside the previously defined DLS (20, 21). Thus, the DLS is more extended than initially thought. For clarity, we will refer in this paper to the initially proposed DLS localized downstream of the S.D. site as the 3′-DLS.

Kinetics, stability, and cation dependence experiments revealed that the dimers formed by RNAs containing the 3′-DLS but missing the sequences located upstream of the S.D. site differ from the dimers formed by RNAs containing both the upstream and downstream sequences (20). Indeed, the dimerization initiation site (DIS) of the latter RNAs was located upstream of the S.D. site (21) (Fig. 1*A*). Surprisingly, mutations in the DIS completely abolish dimerization of RNAs that contain the previously defined 3′-DLS (21). On the contrary, deletion of the RGGARA consensus tracts in the 3′-DLS does not inhibit dimerization of a HIV RNA that contains the DIS (21). These results rise the possibility that the dimers formed by RNAs restricted to the 3′-DLS are artifactual (20, 21). However, the dimers formed by RNAs that contain both the DIS and the 3′-DLS are more stable than the dimers formed by RNAs containing only the DIS (20), suggesting that some stabilizing interactions may indeed take place in the 3′-DLS.

In this paper, we studied the relative contributions of the DIS and the 3′-DLS in the HIV-1 RNA dimerization process by extensive site-directed mutagenesis. In a previous report, we proposed that the initiation step of HIV-1 RNA dimerization is the annealing of the two monomers through Watson-Crick base pairing of their palindromic DIS sequences (Fig. 1*B*) (21). Here, we unambiguously demonstrated a direct interaction between the DIS present on each monomer during dimerization. We also introduced a number of mutations in the 3′-DLS. None of the deletions in this region was found to inhibit dimerization. However, we identified two purine-rich sequences involved in the stabilization of the RNA dimer.

EXPERIMENTAL PROCEDURES

Plasmid Construction and RNA Synthesis—Plasmid construction and restriction enzyme digestions were conducted according to

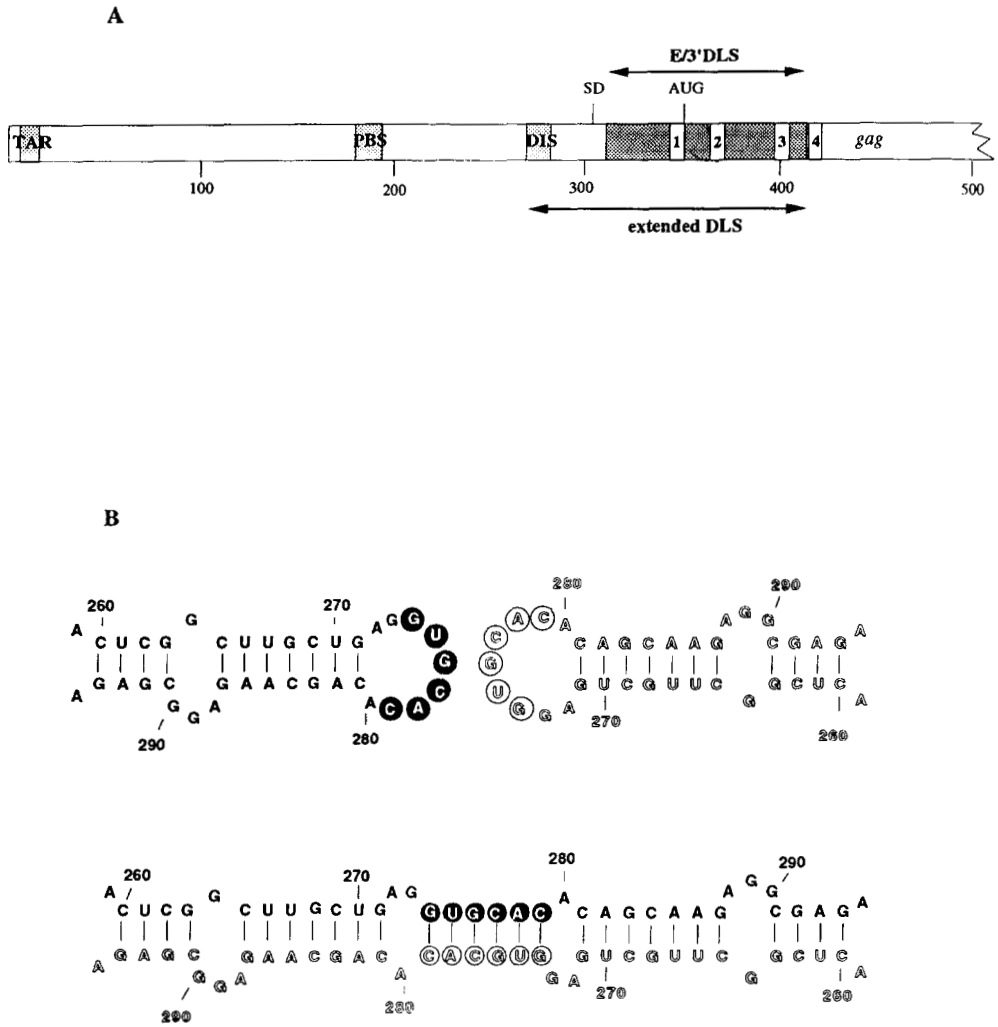


FIG. 1. *A*, scheme of the 500 5′-terminal nucleotides of HIV-1 RNA. The main RNA elements are represented: *TAR*, the trans-activation responsive element; *PBS*, the replication primer $tRNA^{Lys,3}$-binding site; *DIS*, the dimerization initiation site; *E/3′-DLS*, the elements necessary for the *in vivo* encapsidation of the unspliced viral RNA and the previously proposed dimer linkage structure, and *AUG*, the initiator codon for synthesis of gag and gag-pol polyproteins. *B*, secondary structure model of the DIS (nucleotides 256–294 of HIV-1 (Mal isolate) in monomeric and dimeric forms as proposed by Skripkin *et al.* (21). The nucleotides proposed to interact during the initiation of dimerization are *encircled*. Postulated structures of the monomer (*top*) and of the dimer (*bottom*).

standard procedures (22). Plasmid pJCB was obtained by inverse PCR on the plasmid pHIVCG-4 (15). The PCR primers were designed to eliminate the 615–4005 HIV-1 region of pHIVCG-4 and allow a self-ligation of the PCR products after restriction by *Sal*I. The resulting pJCB plasmid also contains a new *Pvu*II restriction site at the 3′ end of the HIV-1 sequence. Plasmids containing the mutations described in Fig. 2 were obtained by inverse PCR on the plasmid pJCB. The PCR products were gel purified, phosphorylated, and used to transform JM109 *Escherichia coli* cells after ligation.

All plasmids were digested with *Pvu*II before transcription to obtain RNAs corresponding to nucleotides 1–615 of the Mal isolate of HIV-1 (+1 being the first nucleotide of the genomic RNA) (23). Alternatively, pJCB was linearized with *Rsa*I in order to obtain RNA 1–311. RNA synthesis and purification was as described previously (15). Labeled RNAs were synthesized using [α^{32}P]ATP (Amersham) (50 μCi/μg of DNA template); alternatively, RNAs were labeled at their 3′ end with [α^{32}P]pCp and RNA ligase from phage T4.

In Vitro Dimerization of HIV-1 RNAs—In a typical experiment, wild type or mutant unlabeled RNA was diluted in Milli-Q (Millipore) water at 400 nM final concentration together with the corresponding radioactive RNA (3–5 nCi, 0.01–0.04 μg), heated for 2 min at 90 °C, and chilled for 2 min on ice. Dimerization was initiated by addition of 2 μl of 5-fold concentrated buffer (final concentration: 50 mM sodium cacodylate, pH 7.5, 300 mM KCl, and 5 mM $MgCl_2$), the samples were incubated at the appropriate temperature, then cooled for 2 min on ice. The dimerization yield of wild type and mutant RNAs was compared after a 30-min incubation at 37 °C. In the kinetic experiments, incubation was at 37 °C for 30 s to 90 min. For the RNA dimer stability experiments, the samples were incubated for 30 min at 30 °C, then the temperature was gradually increased by 7 °C steps. After a 5-min incubation at the appropriate temperature, an aliquot was loaded on gel after addition of glycerol (20% final concentration).

Samples were analyzed on 1.1% agarose gels. Electrophoresis were performed in 45 mM Tris borate, pH 8.3, 0.1 mM $MgCl_2$. Agarose gels were fixed for 10 min in 10% trichloroacetic acid and dried for 40 min under vacuum at room temperature. Bands corresponding to the monomeric and dimeric species were quantified using a BAS 2000 BIO-Imager (Fuji).

All experiments were conducted at least twice (usually three or four times), and the error ranges given in the tables correspond to two standard deviations. In the case of thermal denaturation experiments, the T_m value of the wild type RNA is slightly higher that the earlier published value (20). The origin of this difference is unknown. In order to ensure that all T_m data published in the present paper are directly comparable, the denaturation experiments were performed at the same time with three mutants and the wild type 1–615 RNA, as internal control, and all samples were loaded on the same agarose gel.

RESULTS

We used extensive mutagenesis to study the relative contributions of the DIS and the 3′-DLS in the dimerization process of HIV-1 genomic RNA. For each mutant RNA, we analyzed the dimerization yield and, when significant dimerization was ob-

TABLE I
Dimerization yields, thermal stabilities, and kinetic rate constants of HIV-1 RNA 1-615 WT and mutants of the 3'-DLS (region downstream of the S.D. site)

RNA	Dimerization yield %	Thermal stability ° C	k_{mutant}/k_{wt}
	(±5%)	(±2 °C)	(±0.1)
1-615 wild type	69	52[a]	1.00
1-311 wild type	41	42	1.35
DLS Δ306-317	44	53	2.27
DLS Δ319-323	52	52	1.03
DLS Δ326-339	38	55	1.25
DLS Δ338-349	65	46	0.77
DLS Δ351-366	50	46	0.64
DLS Δ377-385	61	52	1.18
DLS Δ396-404	43	50	0.67
DLS Δ419-433	57	49	0.61
DLS Δ4CONSEN	59	40	0.50
DLS S331-334	41	52	0.83
DLS S358-361	37	41	1.39

[a] This T_m is slightly higher than previously determined (20). See "Experimental Procedures" for details.

served, the dimerization kinetics and the thermal stability of the RNA dimer. All mutations were studied in the same context, *i.e.* HIV-1 RNA 1–615, in order to avoid the artifacts that can occur when RNAs with different 3' or 5' ends are used (20, 21). Also, since HIV-1 genomic RNA contains both the DIS and the 3'-DLS, RNA 1–615 is a better model of the genomic RNA than RNA fragments restricted to the 3'-DLS.

Mutations in the 3'-DLS—We showed that the sequences located downstream of the S.D. site increases the thermal stability of the HIV-1 RNA dimer (Table I) (20). Interactions downstream of the S.D. site may involve the consensus RGGARA tracts that are found in the putative dimerization region of most retroviruses (15, 16). First, we constructed mutants to test the role of these consensus purine tracts located downstream of the S.D. site (Fig. 2, mutants DLSΔ338–349, DLSΔ377–385, DLSΔ396–404, and DLSΔ419–433).

Deletion of the first, second, or fourth purine tract has only minor effects on the dimerization yield, while deletion of the third purine tract (DLSΔ396–404) more severely decreases the dimerization of the corresponding RNA (Table I). However, dimerization is far from being abolished as observed with the DIS mutants (see below). Analysis of the dimerization kinetics and the thermal stability of the RNA dimers indicate little influence of the deletion of the second, the third, or the fourth consensus purine tract. The only effect is a 30–40% decrease of the dimerization kinetics when the third or fourth purine tracts are deleted, while the thermal stability of the corresponding dimers are unaffected (Fig. 3 and Table I). By contrast, deletion of the first purine tract (DLSΔ338–349) reduces the melting temperature of this mutant RNA dimer by 6 °C, and the kinetic constant is decreased by 23% compared to the wild type RNA (Fig. 4 and Table I). These results indicate that the first consensus purine tract may be involved in dimerization of HIV-1 genomic RNA. However, the rather small effects produced by the deletion of the purine tracts are not sufficient to definitively prove the direct involvement of the consensus tracts in the dimerization process.

Another possibility is that some of the purine tracts play indeed an important stabilizing role in the dimer, but since the tracts are redundant, the missing tract could be replaced by the remaining ones. To test this redundancy hypothesis, we simultaneously deleted the four consensus purine tracts (DLSΔ4-CONSENS). As already observed (21), this mutant significantly dimerizes (Table I). Indeed, the dimerization yield of this mutant is similar to that of mutant DLSΔ419–433. However, the thermal stability of the dimer and the dimerization kinetics of mutant DLSΔ4CONSENS are more affected than those of any mutant containing a single purine tract deletion (Figs. 3 and 4, and Table I). The kinetic constant is decreased by 50%, and the thermal stability of the dimer by 12 °C (Figs. 3 and 4). Indeed, the thermal stability of the dimer formed by the mutant DLSΔ4CONSENS RNA is the same as the thermal stability of the dimer formed by an RNA lacking the entire 3'-DLS (RNA 1–311WT, Table I). Thus, it seems that the consensus purine tracts are functionnally redundant and that their simultaneous deletion completely abolishes the stabilization of the dimer provided by the sequences located downstream of the S.D. site.

If, as proposed in early studies (15, 17, 19, 23), dimerization of HIV-1 RNA involves purine quartets, some non-consensus G-rich sequences may also be implicated in the dimerization process. Potential interaction sites are the sequences GGAGG335, GGG355, and GAGAG362. The GGAGG335 sequence is located at the apex of a stem-loop structure (5, 24) and may therefore be easily accessible for tertiary interactions. This stem-loop structure was proposed to be an important element of the packaging signal of HIV-1 RNA (25). Given the possible link between dimerization and encapsidation (6–11), we constructed mutants containing either a deletion of the complete stem-loop (Fig. 2, *mutant DLSΔ326–339*) or a substitution of the GGAGG335 loop sequence (Fig. 2, *mutant DLSS331–334*). These mutations affect the dimerization yield but have no significant effect on the thermal stability of the dimer nor on the dimerization kinetics (Table I).

Sequences GGG355 and GAGAG362 are located at the bottom and at the apex of a small stem-loop immediately following the initiation codon of the *gag* gene, respectively (5). The negative effect of RNA dimerization on the translation of the *gag* gene (5) would be easily explained if these sequences are involved in the dimerization process. We either deleted the entire stem-loop (Fig. 2, *mutant DLSΔ351–366*) or substituted the GAGA sequence (*mutant DLSS358–361*). The thermal stability of both mutant RNAs is significantly decreased (Fig. 4*B* and Table I), suggesting that the GAGAG362 sequence may play a direct role in the stability of the RNA dimer. The dimerization yield of both mutants is strongly decreased, but the effect is more pronounced with the substitution than with the deletion. Opposite effects are observed on the kinetic constant of dimerization: the kinetics is decreased by the deletion but increased in this substitution mutant (Table I).

Other elements potentially involved in the dimerization of HIV-1 RNA are a small stem-loop (nucleotides 305–316) encompassing the splice donor site (3' of nucleotide 305) (5, 26), and an U tract located immediately downstream of this stem loop (nucleotides 319–323) that was cross-linked to the nucleocapsid protein (7). Deletions in these regions drastically reduced packaging of the genomic HIV-1 RNA into the viral particles (27–29). Although both deletions decrease the dimerization yield (Table I, mutants DLSΔ306–317 and DLSΔ319–323), they do not significantly affect the thermal stability of the RNA dimer. Surprisingly, the dimerization rate is 2-fold enhanced by deletion of the splice donor site stem-loop.

Mutations in the DIS—The initial step of dimerization of HIV-1 RNA was proposed to be the annealing of the central palindromic sequence of the DIS, which is present in a loop on each RNA monomer, through Watson-Crick base pairing, followed by the subsequent melting and annealing of the monomer stems (21). In the dimeric structure, the two DIS sequences form three helices separated by internal loops (Fig. 1*B*). Mutants DISΔ265–287, DISΔ279–287, DISS274–276, DISA276, DISC275, DISG278, and DISCOMPENS were designed to further test the involvement of the DIS sequence in the dimerization process of HIV-1 RNA.

A. 3'DLS MUTANTS

```
                  310         320         330         340         350         360         370
                   |           |           |           |           |           |           |
WT          GGUGAG UACGCCAAUU UUUGACUAGC GGAGGCUAGA AGGAGAGAGA UGGGUGCGAG AGCGUCAGUA
DLSΔ306-317 G----- -------AUU UUUGACUAGC GGAGGCUAGA AGGAGAGAGA UGGGUGCGAG AGCGUCAGUA
DLSΔ319-323 GGUGAG UACGCCAA-- ---GACUAGC GGAGGCUAGA AGGAGAGAGA UGGGUGCGAG AGCGUCAGUA
DLSΔ326-339 GGUGAG UACGCCAAUU UUUGA----- ---------A AGGAGAGAGA UGGGUGCGAG AGCGUCAGUA
DLSΔ338-349 GGUGAG UACGCCAAUU UUUGACUAGC GGAGGCU--- ---------A UGGGUGCGAG AGCGUCAGUA
DLSΔ351-366 GGUGAG UACGCCAAUU UUUGACUAGC GGAGGCUAGA AGGAGAGAGA ---------- ------AGUA
DLSΔ377-385 GGUGAG UACGCCAAUU UUUGACUAGC GGAGGCUAGA AGGAGAGAGA UGGGUGCGAG AGCGUCAGUA
DLSΔ396-404 GGUGAG UACGCCAAUU UUUGACUAGC GGAGGCUAGA AGGAGAGAGA UGGGUGCGAG AGCGUCAGUA
DLSΔ419-433 GGUGAG UACGCCAAUU UUUGACUAGC GGAGGCUAGA AGGAGAGAGA UGGGUGCGAG AGCGUCAGUA
DLSΔ4CONSEN GGUGAG UACGCCAAUU UUUGACUAGC GGAGGCU--- ---------A UGGGUGCGAG AGCGUCAGUA
DLSS331-334 GGUGAG UACGCCAAUU UUUGACUAGC UUCGGCUAGA AGGAGAGAGA UGGGUGCGAG AGCGUCAGUA
DLSS358-361 GGUGAG UACGCCAAUU UUUGACUAGC GGAGGCUAGA AGGAGAGAGA UGGGUGCUUC GGCGUCAGUA

                  380         390         400         410         420         430
                   |           |           |           |           |           |
WT          UUAAGCGGGG GAAAAUUAGA UGCAUGGGAG AAAAUUCGGU UAAGGCCAGG GGGAAAGAAA AAAUAU
DLSΔ306-317 UUAAGCGGGG GAAAAUUAGA UGCAUGGGAG AAAAUUCGGU UAAGGCCAGG GGGAAAGAAA AAAUAU
DLSΔ319-323 UUAAGCGGGG GAAAAUUAGA UGCAUGGGAG AAAAUUCGGU UAAGGCCAGG GGGAAAGAAA AAAUAU
DLSΔ326-339 UUAAGCGGGG GAAAAUUAGA UGCAUGGGAG AAAAUUCGGU UAAGGCCAGG GGGAAAGAAA AAAUAU
DLSΔ338-349 UUAAGCGGGG GAAAAUUAGA UGCAUGGGAG AAAAUUCGGU UAAGGCCAGG GGGAAAGAAA AAAUAU
DLSΔ351-366 UUAAGCGGGG GAAAAUUAGA UGCAUGGGAG AAAAUUCGGU UAAGGCCAGG GGGAAAGAAA AAAUAU
DLSΔ377-385 UUAAGC---- -----UUAGA UGCAUGGGAG AAAAUUCGGU UAAGGCCAGG GGGAAAGAAA AAAUAU
DLSΔ396-404 UUAAGCGGGG GAAAAUUAGA UGCAU----- ----UUCGGU UAAGGCCAGG GGGAAAGAAA AAAUAU
DLSΔ419-433 UUAAGCGGGG GAAAAUUAGA UGCAUGGGAG AAAAUUCGGU UAAGGCCA-- ---------- ---UAU
DLSΔ4CONSEN UUAAGC---- -----UUAGA UGCAU----- ----UUCGGU UAAGGCCA-- ---------- ---UAU
DLSS331-334 UUAAGCGGGG GAAAAUUAGA UGCAUGGGAG AAAAUUCGGU UAAGGCCAGG GGGAAAGAAA AAAUAU
DLSS358-361 UUAAGCGGGG GAAAAUUAGA UGCAUGGGAG AAAAUUCGGU UAAGGCCAGG GGGAAAGAAA AAAUAU
```

B. DIS MUTANTS

```
                260         270         280         290
                 |           |           |           |
WT               C UCGGCUUGCU GAGGUGCACA  CAGCAAGAGG CGAG
DISΔ265-287      C UCGG------ ----------  -------AGG CGGG
DISΔ279-287      C UCGGCUUGCU GAGGUGCA--  -------AGG CGAG
DISS274-276      C UCGGCUUGCU GAGAAACACA  CAGCAAGAGG CGAG
DISA276          C UCGGCUUGCU GAGGUACACA  CAGCAAGAGG CGAG
DISC275          C UCGGCUUGCU GAGGCGCACA  CAGCAAGAGG CGAG
DISG278          C UCGGCUUGCU GAGGUGCGCA  CAGCAAGAGG CGAG
DISCOMPENS       C UCGGCUUGCU GAGGCGCGCA  CAGCAAGAGG CCAG
```

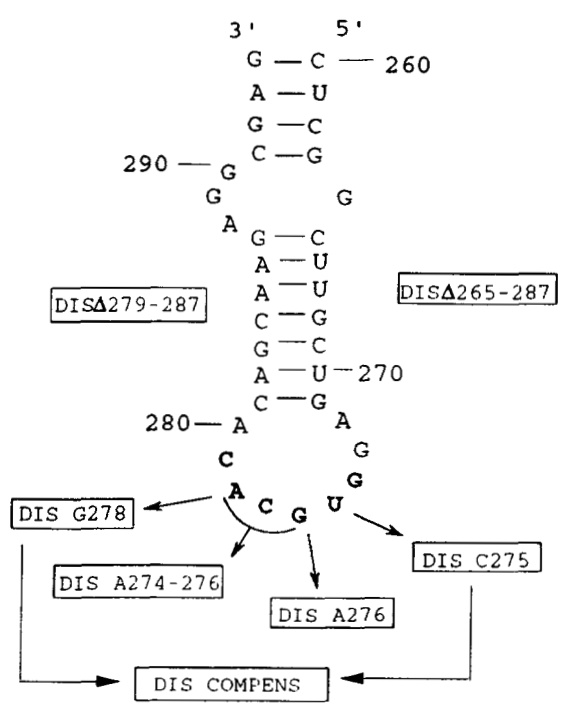


FIG. 2. **HIV-1 RNA mutants synthesized *in vitro*.** *A*, localization of the mutations in the 3'-DLS region (*311–440*). *B*, localization of the deleted (–) or substituted (*underlined*) sequences in the DIS region (*260–290*). These mutations are also indicated on the monomer model of the DIS.

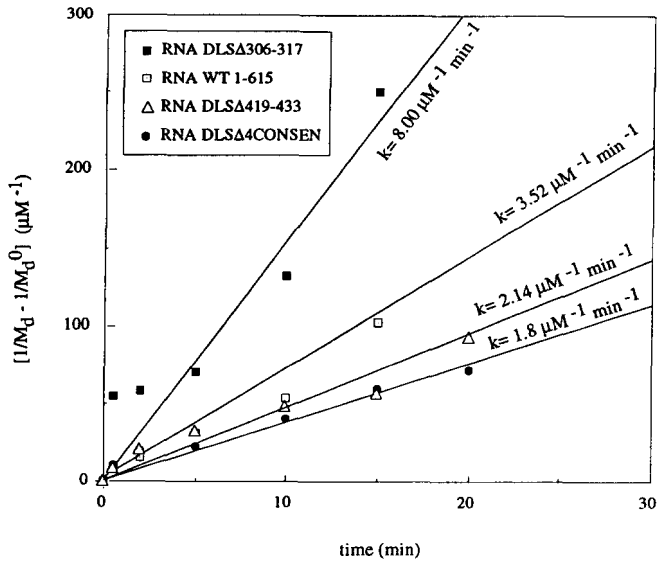


FIG. 3. **Analysis of the dimerization kinetics.** The experimental dimerization data of RNAs (WT) 1–615, DLS Δ306–317, DLS Δ419–433, and DLS Δ4CONSEN were introduced into equation $1/M_d = 1/M_d^0 + 2 \cdot k_{dim} \cdot t$ corresponding to the second order "conformation model" (20).

We first confirmed our previous results (21), indicating that, contrary to what was observed with the deletion mutants in the 3′-DLS, total and partial deletions of the DIS stem-loop (DISΔ265–287, DISΔ279–287) almost totally inhibit dimerization of HIV-1 RNA (Table II). Substitution of nucleotides 274–276 that destroys the proposed six central base pairings of the dimeric form of the DIS (Fig. 2, *mutant DISS274–276*), also abolishes dimerization of 1–615 RNA (Table II). Here, we further show that dimerization is also inhibited by single base substitutions (Fig. 2, *mutants DISA276* and *DISC275*) that introduce two A/C mismatches instead of classical Watson-Crick base pairs in the DIS dimeric form (Table II). When two A-U base pairs are replaced by G-U pairs (Fig. 2, *mutant DISG278*), dimerization is strongly reduced, but not totally abolished (Table II). Accordingly, a strong decrease of the thermal stability of this DISG278 RNA dimer is observed (Table II).

In order to further validate the implication of the DIS sequence in dimerization of HIV-1 RNA, we introduced a compensatory mutation at position 278 in mutant DISC275 (Fig. 2, *mutant DISCOMPENS*) and tested its ability to dimerize. As shown in Table II, substitution of A278 for G restores dimerization of HIV-1 RNA at wild type levels. Compared to the wild type RNA, the compensatory mutant RNA displays no significant differences neither in the thermal stability of the dimer, nor in the dimerization kinetics (Table II). These results strongly support the fact that the palindromic DIS present in each monomer directly interacts within the RNA dimer.

However, one cannot totally exclude an indirect role of the DIS. The crucial point could be a particular structure of the DIS loop in the monomer maintained by the base pairing of nucleotides 275–278, that would be required for dimerization, without any interaction of these sequences in the dimer. To test this possibility, we investigated the capability of the mutant DISCOMPENS RNA to form heterodimer with wild type RNA. Since both RNAs are able to dimerize with themselves, they should also be able to form heterodimers, if the interaction between nucleotides 275 and 278 is intramolecular. On the contrary, no heterodimer should form between these RNAs if this interaction is intermolecular. We used RNA of either 311 or 615 nucleotides in order to distinguish between homo- and heterodimers (Fig. 5*A*). We previously showed that 1–311WT RNA and 1–615WT RNA were able to form hetero length dimers (21). Similarly, DISCOMPENS and 1–311DISCOMPENS RNAs also form hetero length dimers (Fig. 5*A*, *lane 5*). The band migrating between the positions of the dimeric forms of 1–311 RNA and 1–615 RNA in *lane 5* of Fig. 5*A* corresponds to a dimer containing 1 molecule of DISCOMPENS RNA and 1 molecule of 1–311DISCOMPENS RNA. On the contrary, no heterodimer can be form between wild type RNAs and the compensatory mutant RNAs (Fig. 5*A*, *lanes 6* and *7*), irrespective of whether the wild type RNA is the large (*lane 6*), or the short (*lane 7*) molecule.

Similarly, we tested the possibility of forming heterodimers between DISC275 and DISG278 mutant RNAs: in this case, the heterodimer should be favored over the homodimers if the interaction between nucleotides 275 and 278 is intermolecular (Fig. 5*B*). While DISC275 and DISG278 RNAs do not dimerize and poorly dimerize, respectively (Table II and Fig. 5*B*, *lanes 1–4*), it clearly appears that they form heterodimers very efficiently (Fig. 5*B*, *lane 5*). Since only low levels of homodimers are formed by the mutant RNAs used in the experiments, the heterodimers can be evidenced by using RNAs of the same length. These results unambiguously demonstrate the direct interaction of the two palindromic sequences of the DIS in the dimer, as was proposed in Fig. 2*B*.

DISCUSSION

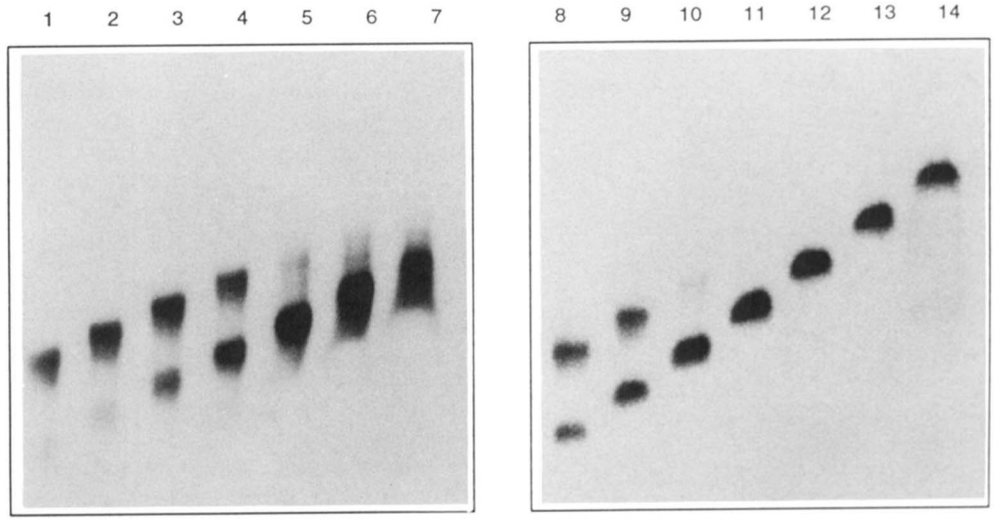
Our results clearly show the importance of the DIS in the dimerization process of HIV-1 RNA, *in vitro*, and confirm the mechanism of dimerization initiation that we recently proposed (Fig. 1*B*) (21). Not only deletions, but also single base substitutions totally abolish or strongly reduce dimerization of 1–615 RNA. This is consistent with our previous results, since we initially identify the DIS by chemical interference experiments (21). In this technique, at most one nucleotide is statistically modified per RNA molecule, so that the DIS could be found only if one modification in the DIS significantly decreases the dimerization yield.

The fact that a compensatory mutation in the palindromic sequence of the DIS restores dimerization is in total agreement with the dimerization scheme that we previously proposed (21). Furthermore, we now unambiguously demonstrated the interaction of the central palindromic sequence of the DIS within the dimer by testing the formation of heterodimers with mutant RNAs. These results demonstrate that we can favor the formation of either homo- or heterodimers. We are presently investigating whether the DIS is also the major determinant of HIV-1 genomic RNA dimerization *ex vivo*. If it is the case, mutations of the DIS will be valuable tools to either favor, or hinder, copackaging of mutated RNAs in cells simultaneously transfected or infected with several retroviral vectors. Thus, mutations in the DIS may be useful in controlling recombination between these vectors.

Effects of the mutations in the 3′-DLS are not so dramatic than those observed with the DIS mutants: no deletion in the 3′-DLS abolishes dimerization of 1–615 RNA. Some of the mutations in the 3′-DLS have no detectable effects (DLSΔ377–385). For most of the mutants (DLSΔ396–404, DLSΔ326–339, DLSS331–334, DLSΔ306–317, and DLSΔ319–323), the dimerization yield, and in some cases the dimerization rate, are affected, but the thermal stability of the RNA dimer is not significantly affected. These results probably reflect an indirect conformational effect of the mutation, rather than a direct involvement of the mutated sequences, in the dimerization process. Chemical and enzymatic probing of the mutant RNA dimers indicates only limited structural perturbations in the vicinity of the mutated sites.[2] The monomeric form of these

[2] E. Skripkin, unpublished results.

A



B

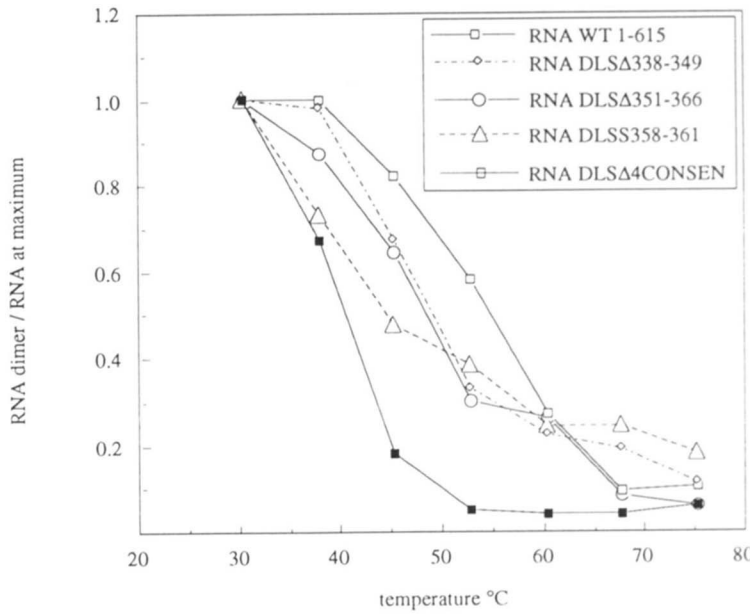


FIG. 4. **Thermal stabilities of dimers of RNAs 1–615 (wild type) and mutants.** *A*, WT RNA 1–615 (*lanes 1–7*) and RNA DLS Δ4CONSEN (*lanes 8–14*) as example of experimental data. Melting experiments were conducted as described under "Experimental Procedures." *Lanes 1* and *8*, 30 °C; *lanes 2* and *9*, 37 °C; *lanes 3* and *10*, 44 °C; *lanes 4* and *11*, 52 °C; *lanes 5* and *12*, 58 °C; *lanes 6* and *13*, 65 °C; and *lanes 7* and *14*, 72 °C. *B*, thermal stability curves of RNAs 1–615 (WT), DLS Δ338–349, DLS Δ351–366, DLS S358, and DLS Δ4CONSEN. For the sake of comparison, the dimerization yields were normalized according to the dimerization yield at 30 °C.

RNAs is structurally heterogeneous,[2] as already observed for the wild type RNA monomer (5). We previously showed that the incomplete dimerization of HIV-1 RNA fragments is due to the presence of monomers that have a conformation incompatible with dimerization, rather than to the limited stability of the dimers (20). Thus, variations of the dimerization yield and kinetics among the mutants RNAs probably also reflect differences in the monomer structures.

On the contrary, important changes in the thermal stability of the dimers probably indicate that the mutated sequences are directly involved in the dimerization process. Only two tested regions downstream of the S.D. site are directly involved in the thermal stability of HIV-1 RNA dimer: one is located immediately upstream of the initiation codon of the *gag* gene, and the second one is located at the apex of the short stem-loop immediately 3′ of the *gag* gene start. Interestingly, both sequences

TABLE II
Dimerization yields, thermal stabilities, and kinetic rate constants of HIV-1 RNA 1-615 WT and mutants of the DIS (region 260–290)

RNA	Dimerization yield %	Thermal stability °C	k_{mutant}/k wt
	±5%	*±2 °C*	*±0.1*
Wild type	69	52	1.00
DIS G278	21	38	nd
DIS COMPENS	61	51	0.79
DIS Δ265-287	<5		
DIS Δ279-287	<5		
DIS S274-276	<5		
DIS A276	<5		
DIS C275	<5		

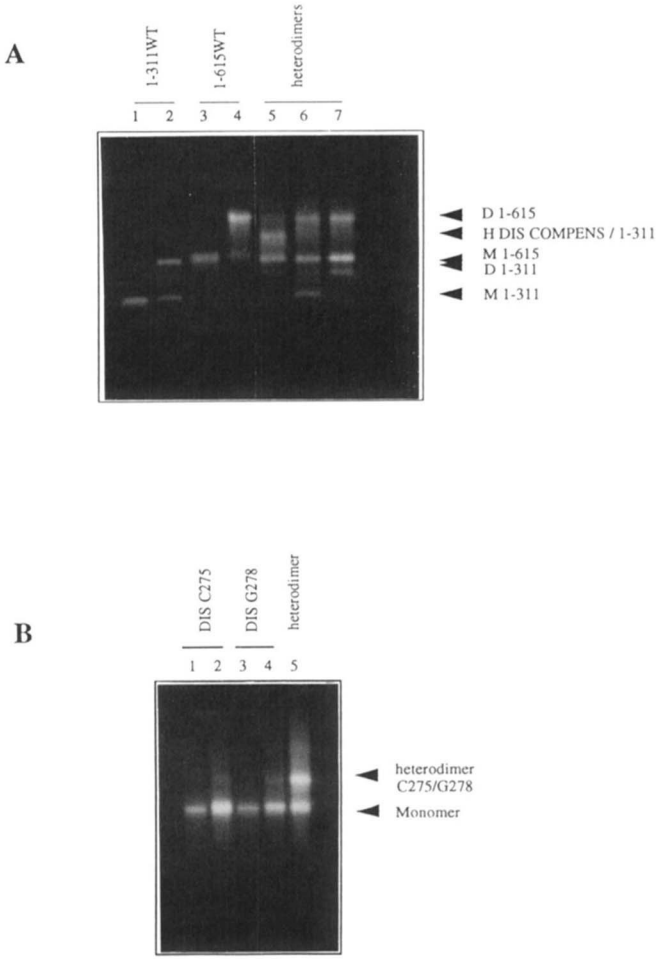


FIG. 5. **Separation on ethidium bromide-stained agarose gels of heterodimers formed between various HIV-1 RNA fragments.** *A*, RNA 1–311WT and RNA 1–615WT were incubated in monomer (*lanes 1* and *3*) or dimer (*lanes 2* and *4*) buffer. RNA DISCOMPENS was coincubated in dimer buffer with either RNA 1–311DISCOMPENS (*lane 5*) or RNA 1–311WT (*lane 7*). RNA1–615WT and RNA DIS 1–311COMPENS were coincubated in dimer buffer (*lane 6*). *M*, monomer; *D*, dimer; and *H*, heterodimer. *B*, RNA DISC275 and RNA DISG278 were either incubated separately, in monomer (*lanes 1* and *3*) or dimer (*lanes 2* and *4*) buffer or coincubated in dimer buffer (*lane 5*).

are purine tracts containing alternating As and Gs. Moreover, the results obtained with mutant DLSΔ4CONSEN indicated that the consensus RGGARA tracts are functionally redundant. The involvement of these sequences in G quartets is incompatible with the cation dependence of RNA 1–615 dimerization (20) and with the stability of dimers of genomic HIV-1 RNA extracted from viral particles (30). Whether the alternating G and A tracts may form G and A quartets, as we first proposed (15), remains an open question. These alternating G and A sequences may also form other unusual structures such as parallel duplexes (31). Of course, they could also stabilize the dimer by classical Watson-Crick interactions.

Deletion of the second consensus purine tract, which was proposed to be involved in the dimerization of short RNAs restricted to the 3′-DLS (17, 18), has no effect neither on the dimerization of RNA 1–615, which also contains the DIS, nor on the thermal stability of the corresponding dimer. This result further stresses the importance of working with RNAs containing the whole 5′ region of HIV-1 genomic RNA.

Although the DIS appears more important than the 3′-DLS in promoting RNA dimerization, the interactions involving sequences downstream of the S.D. site may be crucial for the functions linked to dimerization. Since these interactions involve sequences located immediately upstream and downstream of the initiation codon of the *gag* gene, they may easily explain the negative effect of the RNA dimerization on the translation of this gene that was previously reported (5). In addition, deletions downstream of the S.D. site drastically affect RNA packaging (27–29). Recombination and complete reverse transcription require strand transfer of reverse transcriptase (12, 14). This process should be greatly enhanced by an overall parallel orientation of the two genomic RNA molecules. Indeed, this orientation could be obtained by the intermediate of purine-mediated non-canonical interactions.

Finally, the stabilization of the dimer by the sequences downstream of the S.D. site that we observed *in vitro* may be involved in the maturation step observed *in vivo*. It has been recently shown that murine leukemia virus and HIV-1 RNA dimers undergo a protease-dependent maturation, following the release of the viral particles (30, 32). Although our data are not directly comparable with those of Fu *et al.* (30), one can notice that the T_m of the immature (less than 50 °C) and mature (55 °C) HIV-1 RNA dimers are quite close from the values that we determined for wild type RNAs 1–311 and 1–615, respectively.

Since we showed that the interactions in the 3′-DLS cannot take place within the context of a large RNA in the absence of a functional DIS, the DIS is probably a best target than the 3′-DLS for antiretroviral drugs such as sense or antisense agents.

Acknowledgments—C. Isel is acknowledged for fruitful discussions. D. Mignot is acknowledged for skillful technical assistance.

REFERENCES

1. Bender, W., Chien, Y. H., Chattopadkyay, S., Vogt, P. K., Gardner, M. B., and Davidson, N. (1978) *J. Virol.* **25,** 888–896
2. Bender, W., and Davidson, N. (1976) *Cell* **7,** 595–607
3. Canaani, E., Helm, K. V. D., and Duesberg, P. (1973) *Proc. Natl. Acad. Sci. U. S. A.* **72,** 401–405
4. Murti, K. G., Bondurant, M., and Tereba, A. (1981) *J. Virol.* **37,** 411–419
5. Baudin, F., Marquet, R., Isel, C., Darlix, J. L., Ehresmann, B., and Ehresmann, C. (1993) *J. Mol. Biol.* **229,** 382–397
6. Bieth, E., Gabus, C., and Darlix, J. L. (1990) *Nucleic Acids Res.* **18,** 119–127
7. Darlix, J. L., Gabus, C., Nugeyre, M. T., Clavel, F., and Barré-Sinoussi, F. (1990) *J. Mol. Biol.* **216,** 689–699
8. Prats, A. C., Sarih, L., Gabus, C., Litvak, S., Keith, G., and Darlix, J. L. (1988) *EMBO J.* **7,** 1777–1783
9. Prats, A. C., Roy, C., Wang, P., Erard, M., Housset, V., Gabus, C., Paoletti, C., and Darlix, J. L. (1990) *J. Virol.* **64,** 774–783
10. Prats, A. C., Housset, V., de Billy, G., Cornille, F., Prats, H., Roques, B., and Darlix, J. L. (1991) *Nucleic Acids Res.* **19,** 3533–3541
11. Roy, C., Tounekti, N., Mougel, M., Darlix, J. L., Paoletti, C., Ehresmann, C., Ehresmann, B., and Paoletti, J. (1990) *Nucleic Acids Res.* **18,** 7287–7292
12. Hu, W. S., and Temin, H. M. (1990) *Science* **250,** 1227–1233
13. Luo, G., and Taylor, J. (1990) *J. Virol.* **64,** 4321–4328
14. Temin, H. M. (1991) *Trends Genet.* **7,** 71–74
15. Marquet, R., Baudin, F., Gabus, C., Darlix, J. L., Mougel, M., Ehresmann, C., and Ehresmann, B. (1991) *Nucleic Acids Res.* **19,** 2349–2357
16. Ehresmann, B., Marquet, R., Mougel, M., and Ehresmann, C. (1992) *Bull. Inst. Pasteur* **90,** 109–124
17. Sundquist, W. I., and Heaphy, S. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90,** 3393–3397

18. Awang, G., and Sen, D. (1993) *Biochemistry* **32,** 11453–11457
19. Weiss, S., Häusl, G., Famulok, M., and König, B. (1993) *Nucleic Acids Res.* **21,** 4879–4885
20. Marquet, R., Paillart, J. C., Skripkin, E., Ehresmann, C., and Ehresmann, B. (1994) *Nucleic Acids Res.* **22,** 145–151
21. Skripkin, E., Paillart, J. C., Marquet, R., Ehresmann, B., and Ehresmann, C. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 4945–4949
22. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning: a Laboratory Manual,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
23. Alizon, M., Wain-Hobson, S., Montagnier, L., and Sonigo, P. (1986) *Cell* **46,** 63–74
24. Hayashi, T. Y. U., and Okamoto, T. (1993) *FEBS Lett.* **327,** 213–218
25. Hayashi, T., Shioda, T., Iwakura, Y., and Shibuta, H. (1992) *Virology* **188,** 590–599
26. Harrison, G. P., and Lever, A. M. L. (1992) *J. Virol.* **66,** 4144–4153
27. Lever, A., Gottlinger, H., Haseltine, W., and Sodroski, J. (1989) *J. Virol.* **63,** 4085–4087
28. Aldovini, A., and Young, R. A. (1990) *J. Virol.* **64,** 1920–1926
29. Clavel, F., and Orenstein, J. M. (1990) *J. Virol.* **64,** 5230–5234
30. Fu, W., Gorelick, R. J., and Rein, A. (1994) *J. Virol.* **68,** 5013–5018
31. Rippe, K., Fritsch, V., Westhof, E., and Jovin, T. M. (1992) *EMBO J.* **11,** 3777–3786
32. Fu, W., and Rein, A. (1993) *J. Virol.* **67,** 5443–5449