Biochemistry **1997,** *36,* 9501−9508

HIV-1 Genome Dimerization: Kissing-Loop Hairpin Dictates Whether Nucleotides Downstream of the 5′ Splice Junction Contribute to Loose and Tight Dimerization of Human Immunodeficiency Virus RNA[†]

Michael Laughrea*[‡,§] and Louis Jetté[‡]

McGill AIDS Centre, Lady Davis Institute for Medical Research, Sir Mortimer B. Davis−*Jewish General Hospital, and Department of Medicine, McGill University, Montreal, Quebec, Canada*

Received April 14, 1997; Revised Manuscript Received June 2, 1997[⊗]

ABSTRACT: The genome of all retroviruses consists of two identical RNAs noncovalently linked near their 5′ end. Adjacent genomic RNAs from human immunodeficiency virus type 1 (HIV-1) can form loose or tight dimers depending on whether their respective kissing-loop hairpins (nts 248−270 in HIV-1_{Lai}) bond via their autocomplementary sequences (ACS) or via the ACS and stem sequences [Laughrea, M., & Jetté, L. (1996a) *Biochemistry 35*, 1589−1598]. Loose dimers from HIV-1_{Mal}, but not HIV-1_{Lai}, are stabilized by a sequence (3′DLS) located downstream of the 5′ splice junction [Laughrea, M., & Jetté, L. (1996b) *Biochemistry 35*; 9366−9374]. To understand the ACS−3′DLS interplay in the formation and stability of loose and tight HIV-1 RNA dimers, we replaced the ACS of HIV-1_{Lai} (GCGCGC262) by GUGCAC, GUGCGC (two alternative HIV-1 ACS), or GAGCUC (a non-HIV ACS). For each mutant, RNAs truncated immediately upstream or downstream of the 3′DLS were prepared; their ability to dimerize and their thermal stabilities were compared. The results suggest that the ACS determines whether the 3′DLS participates in RNA dimerization: (1) GAGCUC262 led to poorly stable loose dimers due to the inability of the 3′DLS to stabilize them (the 3′DLS stabilized the GUGCAC and GUGCGC RNAs); (2) GAGCUC262 led to poor formation of tight dimers, due to an inhibitory effect of the 3′DLS (the 3′DLS had little effect on the tight dimerization of the GUGCAC, GUGCGC and GCGCGC RNAs). The results indicate that communication exists between HIV-1 RNA sequences respectively located upstream and downstream of the 5′ splice junction; they are consistent with the idea that the 3′DLS plays two ACS-dependent roles in the dimerization process: loose dimer stabilization in HIV-1 RNAs bearing an HIV ACS (unless the ACS already conferred a thermostability equal or superior to that offered by the 3′DLS), and inhibition of tight dimer formation in an HIV-1 RNA bearing a non-HIV ACS.

The primer binding site and the *gag* gene of the genomic RNA of Human Immunodeficiency Virus Type 1 (HIV-1)[1] are separated by 136 nucleotides extending from U200 to G335 in HIV-1_{Lai}. This 3′ half of the leader is called the L sequence (Coffin, 1984). It can fold into three hairpins, termed the kissing-loop domain, the 5′ splice junction hairpin, and hairpin 3 (KLD, SD, and 3 in Figure 1). Within the HIV-1 and chimpanzee immunodeficiency virus ($\mathrm{SIV}_{\mathrm{cpz}}$) lineage, the kissing-loop domain is the most conserved region of the L sequence (Laughrea et al., 1997). Located in the middle of the L sequence, it encompasses a stem-loop, named the kissing-loop hairpin (nts 248−270 in Figure 1), seated on top of a short but completely conserved stem-bulge (Figure 1; Laughrea et al., 1997). The kissing-loop hairpin is characterized by a 7 base pair stem (stem C; Figure 1) and an 8−9 nt loop (loop C; Figure 1) containing an almost invariant hexameric autocomplementary sequence (ACS) (Laughrea et al., 1997; and references therein).

Mutations in the kissing-loop hairpin have recently been shown to reduce viral infectivity (Laughrea et al., 1997; Paillart et al., 1996a), genomic RNA packaging (Berkhout & Van Wamel, 1996; Laughrea et al., 1997; McBride & Panganiban, 1997; Paillart et al., 1996a), genomic RNA dimerization (Haddrick et al., 1996; Laughrea et al., 1997), viral RNA transcription (Laughrea et al., 1997), and proviral DNA synthesis (Paillart et al., 1996a).

According to the kissing-loop model of HIV-1 genome dimerization (Laughrea & Jetté, 1994, 1996b; Skripkin et al., 1994), dimerization of two adjacent HIV-1 RNAs is initiated by Watson−Crick hydrogen-bonding between the ACS of one HIV-1 RNA monomer and the ACS of a nearby monomer (Figure 1). Adjacent HIV-1 RNAs can form two types of dimers, termed loose and tight dimers, apparently depending on whether their respective kissing-loop hairpins bond via their ACS or, more completely, via their ACS and stem C sequences, which requires the stems C to switch from intra- to interstrand bonding (Laughrea & Jetté,1996a,b; Muriaux et al., 1996a,b). Loose HIV-1 RNA dimers are fully or partially dimeric after electrophoresis in buffer TBE_2 at

---

[†] This work was supported by a grant from the Medical Research Council of Canada to Michael Laughrea.

* To whom correspondence should be addressed at the Lady Davis Institute for Medical Research, 3755 Cote Ste. Catherine Road, Montreal, Quebec, Canada H3T 1E2. Tel: (514) 340-8260. FAX: (514) 340-7502.

[‡] Sir Mortimer B. Davis−Jewish General Hospital.

[§] Department of Medicine.

[⊗] Abstract published in *Advance ACS Abstracts*, July 15, 1997.

[1] Abbreviations: HIV-1, human immunodeficiency virus type 1; nt, nucleotide; ACS, 6-nt long autocomplementary sequence located in the loop of the kissing-loop hairpin; 3′DLS, an RNA sequence located downstream of the 5′ splice junction, starting at nt 296 and ending at nt G508; T_{d}, apparent dissociation temperature; $\mathrm{SIV}_{\mathrm{cpz}}$, chimpanzee immunodeficiency virus; GNGCNC RNA, HIV-1_{Lai} RNA whose GCGCGC262 ACS has been replaced by GNGCNC; R, purine; N, any of the four canonical nts; SEM, standard error of the mean.

S0006-2960(97)00862-3 CCC: $14.00

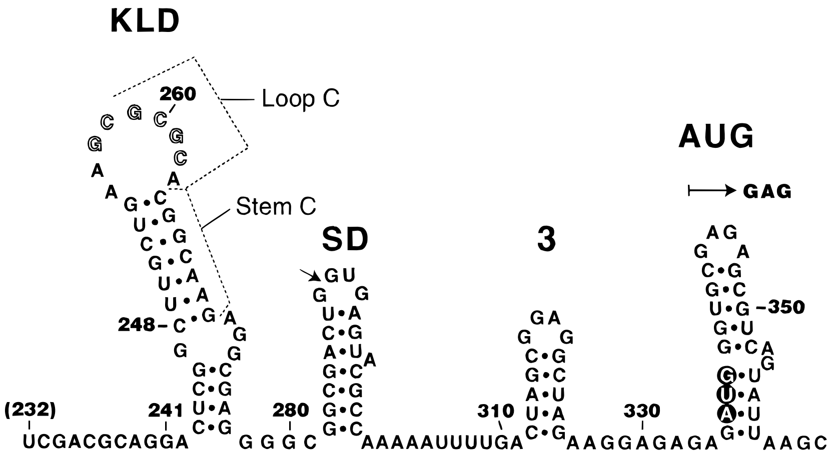

FIGURE 1: Secondary RNA structure model for the HIV-1$_{\text{Lai}}$ L sequence. The cleavage site within the 5′ major splice donor is marked by an arrow within the SD hairpin. The AUG initiation codon of the gag gene is highlighted and somewhat arbitrarily presented as hydrogen-bonded in the interest of a compact figure. This model is based on published structure probing and/or phylogenetic data (Laughrea et al., 1997, and references therein). The nucleotides of the GCGCGC262 autocomplementary sequence (ACS) are highlighted.

4 °C but completely monomeric after electrophoresis in buffer TBE$_2$ at 27 °C; tight dimers are as dimeric after electrophoresis in buffer TBE$_2$ at 27 °C as after electrophoresis in buffer TBE$_2$ at 4 °C (Laughrea & Jetté, 1996a; see below). Loose dimers of HIV-1$_{\text{Mal}}$ RNAs (which bear a GUGCAC ACS), but not of HIV-1$_{\text{Lai}}$ RNAs (which bear a GCGCGC ACS), are stabilized by the 3′ DLS, a sequence located within the 296−508 region, i.e. downstream of the 5′ splice junction (Laughrea & Jetté, 1996b; Paillart et al., 1994; Figure 1). However, stabilized loose dimers are less stable than tight dimers (Laughrea & Jetté, 1996b).

In 49 sequenced loops C from HIV-1 and SIV$_{\text{cpz}}$, the ACS GCGCGC is found in 27 viruses (SIV$_{\text{cpzant}}$ and all HIV-1 subtypes B, D, and F); GUGCAC in 21 viruses (SIV$_{\text{cpzgab}}$, all HIV-1 subtypes C, E, G, O, U, and one HIV-1 subtype A); GUGCGC is found in the loop C of HIV-1$_{\text{Ibng}}$, a subtype A virus. In contrast, GAGCUC is absent from the kissing-loop hairpins of all human and simian immunodeficiency viruses so far sequenced (Berkhout & Van Wamel, 1996; Laughrea et al., 1997; and references therein). According to simple statistical modeling, 49 HIV-1 and SIV$_{\text{cpz}}$ loop C sequences have a ≥99.9% chance of uncovering at least 22 different strong ACS (216 different ACS can be imagined,[2] 41 of which are defined as strong because their predicted free-energy changes of duplex formation (ΔG°_{37}) are equal to or more favorable than −6.2 kcal/mol, the ΔG°_{37} associated with GUGCGC, the weakest wild-type ACS (see Discussion)]. Since only 3 different ACS have thus far been uncovered in HIV-1 and SIV$_{\text{cpz}}$, it follows that some important property of HIV-1 genomic RNA might be exquisitely sensitive to the exact shape of the kissing-loop hairpin.

This paper has two main goals concurrently approached by studying the same RNA mutants. First, we wish to investigate whether HIV-1$_{\text{Lai}}$ RNAs bearing any of the three wild-type HIV-1/SIV$_{\text{cpz}}$ ACS have dimerization properties in common, and if so, whether these shared properties differ from those of an HIV-1$_{\text{Lai}}$ RNA bearing GAGCUC, a strong ACS (ΔG°_{37} of −7.6 kcal/mol) not found in any human or simian immunodeficiency virus. Second, we wish to better understand the interplay of the ACS and the 3′DLS (sequences respectively located upstream and downstream of the 5′ splice junction) in the formation and stability of loose and tight dimers of HIV-1 RNAs. Accordingly, we replaced the ACS of HIV-1$_{\text{Lai}}$ by GUGCAC, GUGCGC, or GAGCUC. For each mutant, HIV-1 RNAs truncated immediately upstream or downstream of the 3′DLS were prepared; their ability to dimerize and their thermal stabilities were compared in order to test if the 3′DLS makes an ACS-dependent contribution to the thermal stability of loose dimers (hypothesized by Laughrea & Jetté,1996b) or to the formation of tight dimers. An ancillary purpose was to investigate if the reported differences between the dimerization properties of HIV-1$_{\text{Lai}}$ and HIV-1$_{\text{Mal}}$ RNAs (Laughrea & Jetté, 1996a, 1996b; Marquet et al., 1994; Paillart et al., 1994) are essentially due to their different ACS.

Our results show that a seemingly minor change in the ACS of HIV-1$_{\text{Lai}}$ considerably influences the loose and tight dimerization processes via altering the function of the "distant" 3′DLS (located ≥35 nts downstream of the ACS mutations). (1) GAGCUC led to poorly stable loose dimers, due to the inability of the 3′DLS to stabilize them (the 3′DLS stabilized the GUGCAC and GUGCGC RNAs[1]), as if a communication between the kissing-loop hairpin and the 3′DLS was altered. (2) GAGCUC led to poor formation of tight dimers, due to an inhibitory effect of the 3′DLS (the 3′DLS had little effect on tight dimerization of GUGCAC, GUGCGC and GCGCGC RNAs); this points to another form of communication betweeen the kissing-loop hairpin and the 3′DLS, at least when tight dimerization is studied in the abscence of NCp7.

MATERIALS AND METHODS

Buffers. Buffer L (Laughrea & Jetté, 1994; Marquet et al., 1991): 50 mM sodium cacodylate, pH 7.5, 40 mM KCl, and 0.1 mM MgCl$_2$. Buffer H (Laughrea & Jetté, 1994; Marquet et al., 1991): 50 mM sodium cacodylate, pH 7.5, 300 mM KCl, and 5 mM MgCl$_2$. TBE$_2$, TBE$_{0.5}$, and TB (Peacock & Dingman, 1967): 89 mM Tris, 89 mM borate, and respectively 2, 0.5 mM, and no EDTA. TBM (Laughrea & Jetté, 1994; Laughrea & Moore, 1977): same as TBE$_2$ but with the EDTA replaced by 1 mM MgCl$_2$.

---

[2] 64 if one does not allow G−U base pairing; 160 if one allows ≤2 G−U base pairs per duplex; 216 if one allows up to six G−U base pairs per duplex. Among the duplexes containing ≤2 G−U base pairs, 32 have a 5′ A, 32 a 5′ C, 48 a 5′ G, and 48 a 5′ U.

Plasmid Construction, RNA Synthesis, and Recovery. Standard techniques were used for molecular cloning (Sambrook et al., 1989), and nucleotide positions are those of the HIV-1$_{\text{Lai}}$ genomic RNA. Plasmid pBH contains the 18−1556 region of the HIV-1$_{\text{Lai}}$ genome cloned in front of a strong T7 RNA polymerase promotor (Laughrea et al., 1997). Plasmid pBL has been described in Laughrea & Jetté (1994). It yields a T7 transcript starting at G1 of the HIV-1$_{\text{Lai}}$ genome. Its dimerization properties are undistinguishable from those of T7 transcripts from plasmid pBH (not shown). The kissing-loop hairpin of pBH was altered by cassette mutagenesis as described in Laughrea et al. (1997). Three mutant plasmids were thus prepared, respectively termed pBHGUGCAC, pBHGUGCGC, and pBHGAGCUC (GCGCGC262 respectively replaced by GUGCAC, GUGCGC, or GAGCUC). All constructs were sequenced to verify that the correct mutations had been achieved. Prior to T7 RNA polymerase transcription, plasmids pBL, pBH, pBHGUGCAC, pBHGUGCGC, and pBHGAGCUC were cleaved with restriction endonucleases *Rsa*I or *Acc*I (Pharmacia, Montreal, Québec) to obtain RNAs ending at U295 and G508, respectively, and starting at G1 or C18. Transcription and RNA recovery was done using the MEGAscript kit of Ambion (Austin, Texas). In short, after transcription had proceeded for 6 h, DNase I was added, followed 15 min later by the addition of ammonium acetate stop solution and extraction with phenol/chloroform.

In Vitro Dimerization of HIV-1 RNAs. A 500-ng to 1-µg amount of RNA dissolved in 8 µL of water was heated for 2 min at 92 °C and chilled for 2 min on ice. For preparation of loose dimers, 2 µL of 5× concentrated buffer L or H was added, and (unless otherwise stated) the samples were incubated for 30 min at 0 °C in buffer L (low salt) or at 30 °C in buffer H (high salt). For preparation of tight dimers, 2 µL of 5× concentrated buffer L or H was added, and the samples were preincubated 30 min at 0 °C and incubated for 10 min at the indicated temperatures (which ranged from 43 to 70 °C). The samples were loaded on 2.5% agarose gels after addition of 2 µL of loading buffer containing glycerol. Electrophoresis was run at 4−5 W at 4 °C, room temperature or at 27 °C, as indicated, using the GNA-100 submarine gel electrophoresis unit of Pharmacia. The electrophoresis (and gel) buffers were TBM, TB, TBE$_{0.5}$, or TBE$_2$, as indicated. Electrophoresis lasted 70−100 min, i.e. long enough to accumulate 6000−7000 V-min. After electrophoresis, the gels were stained in 2 µg of ethidium bromide/mL for 10 min, and the percentage of dimerization was estimated by scanning the photographic negatives with an LKB Ultroscan XL laser densitometer (Laughrea & Jetté, 1994, 1996b). To measure the apparent dissociation temperature of the loose dimers, four or five different 10 µL samples of the same dimeric RNAs were simultaneously postincubated for 10 min at four or five different temperatures (e.g., 41, 45, 49, 53, and 57 °C). These RNA samples were next loaded on a 2.5% agarose gel without delay and with the voltage on (Laughrea & Jetté, 1996b).

RESULTS

Mutating GCGCGC262 of HIV-1$_{\text{Lai}}$ RNA into GUGCAC, GUGCGC, or GAGCUC Results in Loose Dimers Which Are Unstable at Low Mg^{2+}: Similarity between Loose Dimers of GAGCUC, GUGCAC, and GUGCGC RNAs under These Conditions. Plasmid pBH contains the 18−1556 region of the HIV-1$_{\text{Lai}}$ genome cloned in front of a strong T7 RNA polymerase promotor (Laughrea et al., 1997). We replaced its ACS (Figure 1) by either GUGCAC, the HIV-1$_{\text{Mal}}$ ACS, GUGCGC, a subtype A ACS, or GAGCUC, an ACS not found in any Human or Simian Immunodeficiency Virus. Prior to T7 RNA polymerase transcription, the wild-type and mutant plasmids were cleaved with RsaI or AccI to yield HIV-1 RNAs ending at U295 or G508. To produce loose dimers, the RNAs were incubated at 30 °C in buffer H (a high ionic strength buffer containing 350 mM monovalent cations and 5 mM Mg^{2+}) prior to electrophoresis on 2.5% agarose gels containing various concentrations of EDTA or Mg^{2+} (Figure 2).

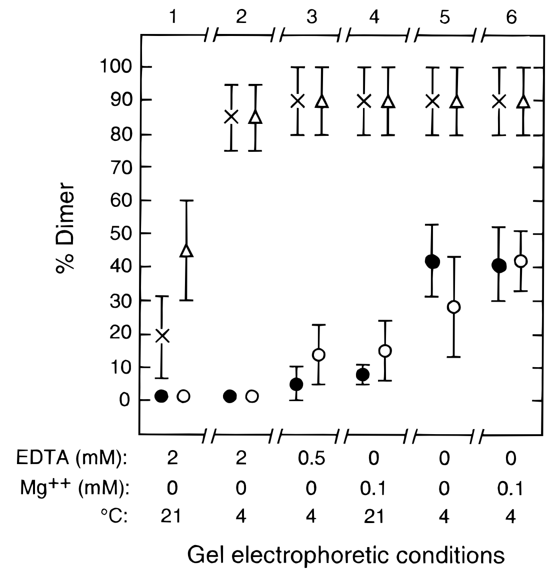

FIGURE 2: Electrophoretic dissociation of loose dimers of HIV-1$_{\text{Lai}}$ RNAs and HIV-1$_{\text{Lai}}$ RNAs whose GCGCGC262 ACS had been replaced by GUGCAC, GUGCGC, or GAGCUC. All RNAs were incubated for 30 min at 30 °C in buffer H (to form dimers) and submitted to 2.5% agarose gel electrophoresis (4−5 W for 70−100 min) in the indicated buffers at the indicated temperatures: TBE$_2$ at 21 °C (lane 1), TBE$_2$ at 4 °C (lane 2), TBE$_{0.5}$ at 4 °C (lane 3), TBM at 21 °C (lane 4), TB at 4 °C (lane 5), and TBM at 4 °C (lane 6). ×: HIV-1$_{\text{Lai}}$ RNAs 1−508 and 18−508. △: HIV-1$_{\text{Lai}}$ RNAs 1−295 and 18−295. ●: RNAs GUGCAC 18−508, GUGCGC 18−508 and GAGCUC 18−508. ○: RNAs GUGCAC 18−295, GUGCGC 18−295 and GAGCUC 18−295. (RNAs bearing mutant ACS have been regrouped because no significant difference was detected between them.) Consistent with lane 6, Clever et al. (1996) have recently found that HIV-1$_{\text{Lai}}$ RNA 178−384 mutated to bear a GUGCGC ACS are substantially dimeric when incubated in buffer H and electrophoresed in buffer TBM at 4 °C. Error bars represent one standard deviation (except for RNA that were ≥85% dimeric, where they reflect the full range of the results).

Mutant RNAs are not categorized according to their ACS in Figure 2 because they had the same proportion of loose dimers, no matter their ACS (not shown) or the location of their 3′ end (compare ○ to ● in Figure 2). After electrophoresis at 4 °C in buffer TBE$_2$ (a Tris-borate buffer containing 2 mM EDTA and no Mg^{2+}), mutant RNAs were monomeric while HIV-1$_{\text{Lai}}$ RNAs 18−295 and 18−508 were almost fully dimeric (lane 2) and still moderately dimeric at 21 °C (lane 1). This suggests that under severe Mg^{2+} depletion, the dissociation temperature of HIV-1$_{\text{Lai}}$ RNA is > 17 °C higher than that of the three mutants (compare × and △ of lane 1 to ○ and ● of lanes 2 and 3). After electrophoresis in buffer TBM (no EDTA and 0.1 mM

Table 1: Apparent Dissociation Temperature (T_d) of Dimeric HIV-1_{Lai} RNAs Ending at G508 or U295 and Containing Either No Mutation (GCGCGC) or Having the GCGCGC262 ACS Replaced by the First Three Indicated Sequences[a]

ACS (nts 257−262)	T_d (in °C) of RNAs ending at G508	T_d (in °C) of RNAs ending at U295
GUGCAC	51 ± 1	42.5 ± 2
GUGCGC	51.5 ± 1.5	42.5 ± 1
GAGCUC	43.5 ± 0.5	43± 1
GCGCGC	55 ± 2.5[b]	55 ± 2.5[b]

[a] The RNAs were dimerized by incubation at 30 °C in buffer H and were assayed by electrophoresis in buffer TB at 4 °C (Materials and Methods). [b] From Laughrea and Jetté (1996b).

Mg^{2+}), mutant RNAs were less dimeric than HIV-1_{Lai} RNAs electrophoresed in buffer TBE_2 (compare ○ and ● in lanes 4 and 6 to × and △ in lanes 1 and 2). Mutant RNAs were ∼40% dimeric prior to electrophoresis for two reasons: (i) they appeared 40% dimeric after electrophoresis at 4 °C in buffer TB or TBM (lanes 5−6); (ii) if incubated at room temperature in buffer TBE_2 or $TBE_{0.5}$ prior to electrophoresis at 4 °C in TB or TBM, they appeared less than 5% dimeric, indicating that dimerization did not occur during electrophoresis (not shown). Incubating the samples of lanes 5−6 at room temperature or at 37 °C in buffer H prior to electrophoresis did not significantly decrease (37 °C) or did not increase (room temperature) the dimerization levels, as if the <100% dimer yield was not due to an incubation temperature too close to the apparent dissociation temperature (T_d) of the dimeric RNAs (not shown).

Loose Dimers of GAGCUC RNAs Are Poorly Stable Because the 3′DLS Does Not Stabilize Them. To compare more systematically our eight RNA dimers, we measured their T_d in buffer H. The RNAs were preincubated for 30 min at 30 °C in buffer H and incubated for 10 min at either 37, 41, 45, 49, 53, or 57 °C prior to electrophoresis at 4 °C on a 2.5% agarose gel containing buffer TB (Laughrea & Jetté, 1996b). T_d[3] was scored as the temperature at which % dimerization was half the level obtained at 30 °C. The first column of Table 1 shows that RNAs ending at G508 formed two classes: GCGCGC, GUGCAC, and GUGCGC RNA dimers had a high T_d of 51−55 °C while GAGCUC RNA dimers had a low T_d of ∼43 °C. The second column of Table 1 displays the T_ds of RNAs ending at U295 and identifies the source of this differential behaviour: the 3 mutant RNAs had an identical and low T_d of ∼42.5−43 °C, indicating that the 296−508 region had stabilized the long GUGCAC and GUGCGC RNA dimers. HIV-1_{Lai} RNA dimers ending at U295 are less revealing because they are as thermostable as long HIV-1_{Lai} RNA dimers (Table 1; Laughrea & Jetté, 1996b). Their high T_d, presumably a consequence of the highly favorable $\Delta G°_{37}$ of the GCGCGC duplex (−10.4 kcal/mol; Laughrea & Jetté, 1996b), masks any putative contribution from the 3′DLS (see Discussion).

In sum, the 3′DLS stabilized RNA dimers containing a low T_d kissing-loop hairpin, except the one bearing a non-HIV ACS. Table 1 therefore suggests that stabilized loose dimeric HIV-1 RNAs are physiologically preferred, and that this stabilization is mediated by only a *subset* of the 41 "strong" ACS defined in the introduction. This subset includes the three ACS so far found in HIV-1 and SIV_{cpz} kissing-loop hairpins but apparently not GAGCUC, which is absent from the kissing-loop hairpin of any human or simian immunodeficiency virus sequenced so far. As suggested by thermodynamics of duplex formation, the dimeric GAGCUC kissing-loop hairpin is not intrinsically less stable than the dimeric GUGCAC and GUGCGC kissing-loop hairpins (second column of Table 1); however, it prevents a communication between the kissing-loop hairpin and the 3′DLS which exists in the other mutant RNAs and, perhaps also (see Discussion), in HIV-1_{Lai} RNA.

Incubation in buffer L (a low ionic strength buffer containing 90 mM monovalent cations and 0.1 mM Mg^{2+}) did not yield enough dimeric mutant RNAs to allow reliable T_d estimates (not shown). This is consistent with lane 4 of Figure 2, which showed that mutant RNAs appeared poorly dimeric after electrophoresis at room temperature in TBM, a buffer ionically closer to buffer L than to buffer H.

Tight Dimers of GAGCUC RNAs Are Inefficiently Formed Due to 3′DLS Interference. In the absence of NCp7, formation of tight dimers requires incubation at >45 °C (Laughrea & Jetté, 1996a; Muriaux et al., 1996b). Laughrea and Jetté (1996a) have suggested, and Muriaux et al. (1996a) have recently shown, that NCp7 can activate at 30−37 °C the formation of tight dimers from preformed loose dimers of HIV-1_{Lai} RNAs.[4] This suggests that tight dimerization of HIV-1 genomic RNA might play a physiological role. If this role is important, then all wild-type HIV-1 RNAs should be able to form tight dimers at 37 °C (in the presence of NCp7) or perhaps even at >45 °C in the absence of NCp7, whereas judiciously chosen kissing-loop hairpin mutants might not.

Accordingly, we tested whether our RNAs truncated at G508 could form tight dimers at temperatures ranging from 43 to 61 °C in buffer L (Figure 3A) and from 46 to 70 °C in buffer H (Figure 3B). As assessed by gel electrophoresis in buffer TBE_2 at 27 °C, GUGCGC, GUGCAC, and GCGCGC RNAs could efficiently form tight dimers in both buffers; GAGCUC RNA could not (compare continuous to dotted lines in Figure 3A and 3B). This differential tight dimerization efficiency was at least as detectable at relatively low temperature (49−52 °C in buffer L; 55−58 °C in buffer H) as at higher temperature, suggesting a possible physiological significance. In contrast, all RNAs truncated at U295 efficiently formed tight dimers in buffers L (Figure 4A) and H (Figure 4B). Thus, the 296−508 region strongly reduced tight dimerization of GAGCUC RNA: in buffer L, the inhibition was 80% at 49 °C and 55−70% at 52 to 58 °C (dotted line of Figure 3A divided by dotted line of Figure

---

[3] Out of 27 published thermal analyses of retroviral dimeric RNAs (16 on HIV-1 RNAs), 26 have involved nondenaturing gel electrophoresis and one has involved glycerol gradient centrifugation, but none has involved alternative methods to distinguish monomeric from dimeric RNAs (Laughrea & Jetté, 1996b, and references therein; Laughrea et al., 1997; Berkhout & Van Wamel, 1996; Feng et al., 1996; Fossé et al., 1996; Haddrick et al., 1996; Muriaux et al., 1996a,b; Paillart et al., 1996a,b). We are aware that T_ds obtained by such economical techniques cannot be regarded as absolute because the data are not taken at equilibrium and the incubation buffer usually differs from the electrophoresis buffer. However, such T_ds are useful when the goal is to compare RNAs of related sequences or to compare our results to previous thermal analyses of retroviral RNA dimerization. Thus, the meaningful results of the present T_d study are not so much the exact numerical T_d values obtained but *the sequence dependence of these values.*

[4] NCp7 is also able to stabilize dimeric RNA obtained from a synthetic fragment of Harvey sarcoma virus RNA (Feng et al., 1996).

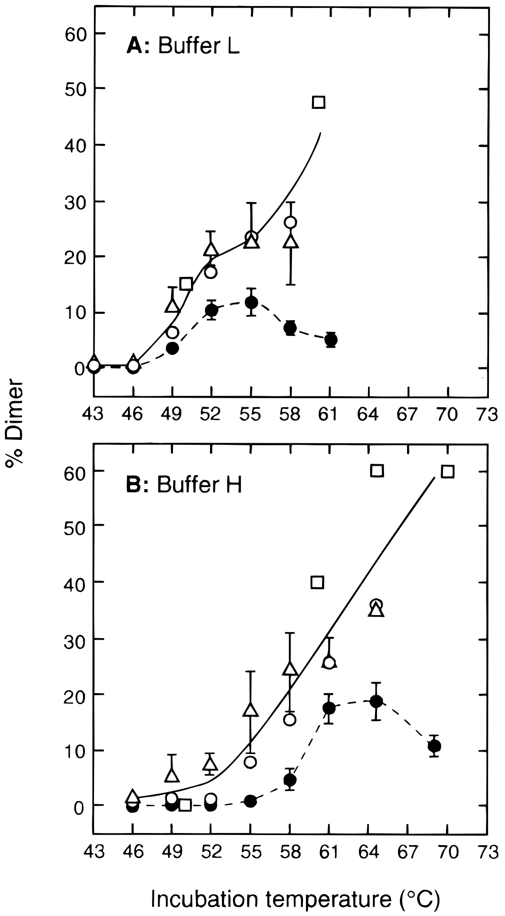

FIGURE 3: Tight dimerization of HIV-1 RNAs ending at G508 as a function of incubation buffer and incubation temperature. Incubation in buffer L (panel A) or buffer H (panel B) was for 10 min at the indicated temperatures prior to 2.5% agarose gel electrophoresis in TBE$_2$ at 27 °C (Materials & Methods). ○: RNA GUGCAC; △: RNA GUGCGC. ●: RNA GAGCUC. □: RNA HIV-1$_{Lai}$. Error bars have been included for the GAGCUC and the GUGCAC RNAs (the error bars for the other two RNAs were slightly smaller than for the GUGCAC RNAs). Error bars represent SEM of three to five data points.

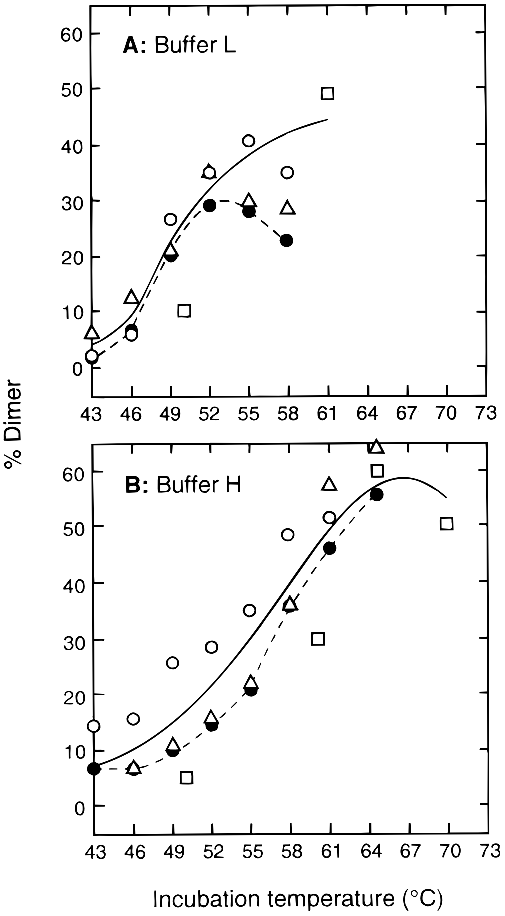

FIGURE 4: Tight dimerization of HIV-1 RNAs ending at U295 as a function of incubation buffer and incubation temperature. Symbols and experimental conditions as in Figure 3.

4A); in buffer H, the inhibition was >85% at 49−58 °C and 60−65% at 61−65 °C (dotted line of Figure 3B divided by dotted line of Figure 4B). After subtracting from Figures 3 and 4 a small but significant temperature-dependent background formation of tight dimers (see Discussion), these four percentages respectively increase to >85%, 70−80%, >85%, and 75−80%, indicating that kissing-loop hairpin-directed tight dimerization was reduced ~5-fold by the 3′DLS of the GAGCUC RNA. In comparison, the 3′DLS reduced tight dimerization of the GUGCGC and GUGCAC RNAs by typically a little less than 50%, often a non-significant difference (compare continuous lines of Figure 4 to continuous lines of Figure 3; data not shown). Taken together, Table 1, Figure 3, and Figure 4 suggest that the 3′DLS of GAGCUC RNA *misinteracted* with its kissing-loop hairpin, the result being no positive effect on loose dimerization and a (5-fold) negative effect on tight dimerization.

DISCUSSION

Replacing the ACS by a non-autocomplementary sequence inhibits the in vivo dimerization of HIV-1 genomic RNA (Haddrick et al., 1996; Laughrea et al., 1997). Here, we replaced the GCGCGC ACS of HIV-1$_{Lai}$ RNA (Figure 1) by the two other HIV-1/SIV$_{cpz}$ ACS or by a thermodynamically equivalent ACS not found in any human or simian immunodeficiency virus; we studied the effect of these mutations on the dimerization of the resulting HIV-1 RNA transcripts. We reasoned that if HIV-1 and SIV$_{cpz}$ sequences have so far uncovered only three different ACS, the dimerization of HIV-1 RNA might require the ACS to be more than any autocomplementary hexamer able to form a stable duplex.[5] We found non-trivial in vitro differences supporting this reasoning. Specifically, our results point to the existence of two types of ACS-dependent communications between the kissing-loop hairpin and the 3′DLS (respectively located upstream and downstream of the 5′ splice junction): positive influence of the 3′DLS on the stability of loose dimers bearing an HIV-1/SIV$_{cpz}$ ACS, and negative influence of the 3′DLS on the tight dimerization of an RNA bearing a non-HIV ACS. It is not known if loose and tight dimerization are influenced by the same sequence elements of the 3′DLS.

Conditional Role of the 3′DLS in Loose Dimerization (via the ACS). The results of Table 1 support the dimerization model of Laughrea and Jetté (1996b), according to which

---

[5] The 41 autocomplementary hexamers predicted to form stable duplexes ($\Delta G°_{37} \leq -6.2$ kcal/mol) are >66% G/C-rich: 20 start with G, 11 with A or U, and 10 with C; they all form ≤2 nonadjacent G−U base pairs. Not all >66% G/C-rich autocomplementary hexamers forming ≤2 G−U base pairs are able to form stable duplexes: 14 can not (Freier et al., 1986; Turner et al., 1988; He et al., 1991).

Table 2: Four Groups of Kissing-Loop Hairpins among 47 HIV-1 Leaders So Far Sequenced[a]

	no. of base pairs in stem C	no. of G−U base pairs in stem C[b]	loop C sequence[c]
group 1: subtypes A, E, G, and U	7	0	$_{255}$RG**GUGCRC**A$_{263}$
group 2: subtypes B, D, and F	7	0 or 1	$_{255}$AA**GCGCGC**R$_{263}$
group 3: subtype C	8[d]	1	$_{255}$AA**GUGCAC**U$_{263}$
group 4: subtype O	8[d]	0	$_{255}$GA**GUGCAC**C$_{263}$

[a] Laughrea et al., 1997, and references therein. [b] The first seven base pairs of stem C (i.e., not counting the extra base pair present in groups 3 and 4 kissing-loop hairpins) include four G−C base pairs and either three A−U or two A−U plus one G−U base pairs. [c] The ACS is indicated in bold letters. [d] The longer stem C comes from base pairing between the 5′ and 3′ bases of loop C.

mere dimerization of the kissing-loop hairpin is not sufficient to trigger either dimerization of the 3′ DLS or any other mechanism whereby the 3′ DLS might stabilize dimeric HIV-1 RNAs. If one assumes that the 3′DLS makes a functional (but concealed) contribution to the dimerization of HIV-1$_{\text{Lai}}$ RNA (Laughrea & Jetté, 1996b), then Table 1 shows that the action of the 3′DLS on loose dimers is inhibited by two transversions but not by one or two transitions within the ACS; if one assumes that the 3′DLS makes no contribution to the dimerization of HIV-1$_{\text{Lai}}$ RNA, then the action of the 3′DLS on loose dimers would oddly be unaffected by two transversions in the ACS but be stimulated by one or two transitions. The first alternative is simpler: it means that the kissing-loop hairpin positively influences the 3′DLS unless the ACS is substantially modified.

Are HIV-1$_{\text{Lai}}$ RNAs bearing another HIV-1 ACS wild-type or hybrid RNAs? If they are hybrid, do they significantly differ from other wild-type RNAs? The 47 HIV-1 kissing-loop hairpins so far sequenced can be divided into four groups according to the sequence of their loop C (Table 2). (i) Group 1 kissing-loop hairpins have a U/A containing ACS coupled to a strong stem C (free of G−U base pairs). (ii) Group 2 have a GCGCGC ACS flanked by AA256 and R263. (iii) Group 3 have a U/A containing ACS and a weak but long stem C (stem C contains one G265−U253 base pair; A255 and U263 can form a base pair). (iv) Group 4 have a U/A containing ACS coupled to a strong and long stem C (G255 and C263 can form a base pair). Since HIV-1$_{\text{Lai}}$ has a weak stem C (Figure 1), it follows that HIV-1$_{\text{Lai}}$ RNAs bearing another HIV-1 ACS are, strictly speaking, hybrid rather than wild-type RNAs: GUGCAC and GUGCGC RNAs have the ACS of group 1 kissing-loop hairpin but not its strong stem C and its G256. We assume that GUGCAC and GUGCGC RNAs, though of hybrid sequence, dimerize in a way that cannot be distinguished from group 1 RNAs. To support this assertion, we note that GUGCAC RNA has four functional properties reminiscent of HIV-1$_{\text{Mal}}$ RNA (Marquet et al., 1994; Paillart et al., 1994; Skripkin et al., 1994), a group 1 sequence bearing a GUGCAC ACS: (i) inability of GUGCAC RNA loose dimers to appear dimeric after electrophoresis in a Mg^{2+}-depleted buffer (Figure 2); (ii) stabilization of its loose dimers by the 3′DLS (Table 1); (iii) T_d identical to the T_d of HIV-1$_{\text{Mal}}$ RNA dimers; (iv) inefficient tight dimerization relative to that of HIV-1$_{\text{Lai}}$ RNA (Figure 3; Laughrea & Jetté, 1996a; Muriaux et al., 1996b) [the results of Paillart et al. (1994, 1996b) suggest that HIV-1$_{\text{Mal}}$ RNAs do not form tight dimers or, at best, form them inefficiently]. The kissing-loop hairpins of HIV-1$_{\text{Lai}}$ and HIV-1$_{\text{Mal}}$ differ by four transitions, one at G265 in the stem, and three in loop C: one at A256 and the other two in the ACS. Our results are consistent with the idea that the two transitions located in the ACS are largely responsible for the reported differential dimerization behaviour of HIV-1$_{\text{Lai}}$ and HIV-1$_{\text{Mal}}$ RNAs (Laughrea & Jetté, 1996a,b; Marquet et al., 1994; Paillart et al., 1994), despite the fact that the HIV-1$_{\text{Mal}}$ sequence differs considerably from that of HIV-1$_{\text{Lai}}$, most notably by having a large 20 nt insert between A217 and C218, i.e. half-way between the primer binding site and the kissing-loop domain.

Duplex formation by GAGCUC, GUGCGC, and GUGCAC oligonucleotides in 1 M NaCl leads to respective $\Delta G°_{37}$ values of −7.6, −6.2, and −7.4 kcal/mol vs a much more favorable −10.4 kcal/mol for GCGCGC (Freier et al., 1986; He et al., 1991; Turner et al., 1988). Should not this 3−4 kcal/mol in favor of the GCGCGC duplex allow HIV-1$_{\text{Lai}}$ RNAs to band apart at least in some assays? It does. First, in buffer H, HIV-1$_{\text{Lai}}$ RNAs ending at U295 had a T_d that was 12 °C higher than other RNAs ending at U295 (Table 1). Second, under stringent electrophoretic conditions (such as electrophoresis at room temperature and low Mg^{2+} or even at 4 °C in the presence of EDTA and no Mg^{2+}), loose HIV-1$_{\text{Lai}}$ RNA dimers were more stable than the other loose dimers, no matter whether the RNAs ended at U295 or G508 (lanes 1−4 of Figure 2). The contribution of Mg^{2+} to the formation and preservation of loose dimers has been discussed before (Laughrea & Jetté, 1996a). Table 1 and Figure 2 thus confirm thermodynamic expectations that, for the purpose of duplex formation, GAGCUC is as good as GUGCAC and GUGCGC, but weaker than GCGCGC. The apparent paradox between Figure 2 and the left side of Table 1 probably means that constructive interactions between the kissing-loop hairpin and the 3′DLS are impossible when Mg^{2+} concentration is unphysiologically low.

Conditional Role of the 3′DLS in Tight Dimerization (via the ACS): Physiological Significance of the Tight Dimerization Process. Our data on loose and tight dimerization, together with those of Muriaux et al. (1996a), favor the idea that tight dimerization of HIV-1 genomic RNA has physiological significance and is exquisitely sensitive to the ACS sequence. Namely (i) tight dimerization of HIV-1$_{\text{Lai}}$ RNA (and possibly Harvey sarcoma virus RNA) occurs at 30−37 °C in the presence of NCp7 (Muriaux et al., 1996a; Feng et al., 1996); (ii) mature HIV-1 genomic RNA has a higher T_d than immature dimeric HIV-1 genomic RNA (Fu et al., 1994); (iii) even in the abscence of NCp7, tight dimers are efficiently formed by HIV-1$_{\text{Lai}}$ RNAs and HIV-1$_{\text{Lai}}$ RNAs mutated to bear any of the two other HIV-1/SIV$_{\text{cpz}}$ ACS, but inefficiently formed by an HIV-1$_{\text{Lai}}$ RNA carrying a similar but non-HIV ACS (this paper). The effective difference between "efficient" and "inefficient" tight dimerization is 50% larger than the ∼2-fold difference obtained by directly comparing continuous to dotted lines in Figure 3. The reason is that heating HIV-1 RNAs deprived of a functional kissing-loop hairpin at 49−58 °C in buffer L or at 58−65 °C in buffer H led to background formation of tight dimers, probably due to some stem-loops having a melting temperature close to the incubation temperature and occasionally switching from intrastrand to interstrand hydrogen bonding (Laughrea et al., 1997; data not shown). To properly interpret the results of Figure 3, this 3 to 9% tight dimeric RNA background[6] has to be subtracted: as a result, tight dimerization specifically due to the kissing-loop hairpin of the GAGCUC RNA was half the level shown by the dotted lines in Figure 3, while tight dimerization specifically due to the kissing-loop hairpins of the GUGCAC and GUGCGC RNAs was ∼75% of the level shown by the continuous lines in Figure 3. Thus, kissing-loop hairpin-related tight dimerization of GUGCAC and GUGCGC RNAs was ∼3 times more efficient than that of GAGCUC RNA. This and the above observations lead us to the following assertions. (1) Tight dimerization might occur with all wild-type HIV-1 RNAs (especially in the presence of NCp7). (2) Since dimerization of HIV-1_{Lai} RNAs in the presence of NCp7 significantly improves tight dimer yields—from ∼60% to >95% (Muriaux et al., 1996a,b; Laughrea & Jetté, 1996a,b)—and since no artefactual tight dimerization occurs at 37 °C (Laughrea & Jetté, 1996b), tight dimerization studied at 37 °C in the presence of NCp7 might reveal sharper differences between the respective abilities of HIV-1 RNAs bearing an HIV or a non-HIV ACS. These assertions can be tested by studying the differential influence of kissing-loop hairpin mutations on the tight dimerization of HIV-1 RNAs in the presence of NCp7.

HIV-1_{Lai} RNAs can form loose and tight dimers (Laughrea & Jetté, 1996a,b; Muriaux et al., 1996b) while HIV-1_{Mal} RNAs apparently form no or background levels of tight dimers (Paillart et al., 1994, 1996b; Laughrea & Jetté, 1996a). This variance raises an important question: is tight dimerization a property of all or only some HIV-1 RNAs? Consistent with the fact that HIV-1_{Lai} RNA 248−271 (Figure 1) can form loose and tight dimers (Laughrea & Jetté, 1996a), formation of loose HIV-1_{Mal} RNA dimers does not involve melting of the two adjacent stems C (Paillart et al., 1996b), from which it follows that formation of tight HIV-1_{Lai} RNA dimers most likely involve melting of the stems C, as suggested earlier (Laughrea & Jetté, 1996a). It is wrong to conclude that HIV-1_{Mal} RNAs are unable to form dimers that involve melting of the stems C (i.e. tight dimers), for at least two related reasons. (1) HIV-1_{Mal} RNAs might form tight dimers only in the presence of NCp7 or similar RNA binding proteins. This would not be astonishing because: (i) the T_{d} of HIV-1_{Lai} RNA tight dimers is only ∼5 °C higher than their optimal protein-free formation temperature (Laughrea & Jetté, 1996b) (thus formation temperature and T_{d} of HIV-1_{Mal} RNA tight dimers might be too close to allow significant protein-free tight dimerization); and (ii) GUGCAC RNA, which has the same ACS as HIV-1_{Mal}, did not tight dimerize as efficiently as HIV-1_{Lai} RNA (Figure 3 shows that its tight dimerization plateaued at ∼35% tight dimer vs ∼60% with HIV-1_{Lai} RNA). (2) Tight dimerization of protein-free HIV-1_{Mal} RNA might be detectable only within a narrow temperature window, a window which might have been missed by the incubation protocol of Paillart et al. (1994, 1996b), where dimerization was studied at intervals never smaller than 7 °C. It is correct to conclude that protein-free HIV-1_{Lai} RNA transcripts form tight dimers more easily than protein-free HIV-1_{Mal} RNA transcripts. Further elaboration amounts to speculation, including our hypothesis that, because GUGCAC RNA could form tight dimers (Figure 3), HIV-1_{Mal} RNAs will eventually be shown to tight dimerize.

Conclusion. Our results account for the natural selection of few autocomplementary sequences (GCGCGC, GUGCAC, and GUGCGC) in the kissing-loop hairpin of HIV-1 and $\mathrm{SIV}_{\mathrm{cpz}}$ viruses and for the apparent non-selection of GAGCUC: the three select sequences allow better loose dimerization (via a positive action of the 3′DLS) and better tight dimerization (via prevention of a negative impact from the 3′DLS). This specific influence of all three wild-type ACS on sequences downstream of the 5′ splice junction has physiological merits: if the kissing-loop hairpin was solely responsible for genomic RNA dimerization (and the dimerization-dependent physiological processes), both spliced and unspliced RNAs would dimerize as efficiently. Our results suggest that genomic RNA dimerizes more efficiently than spliced HIV-1 RNA because the dimeric kissing-loop hairpin, though common to both RNAs, influences, if it has the right ACS, downstream sequences found only in unspliced RNAs.

ACKNOWLEDGMENT

We thank Dr. Thierry Ducruet for statistical modeling and Kay Berckmans for the preparation of the typed text.

REFERENCES

Berkhout, B., & Van Wamel, J. L. P. (1996) *J. Virol. 70*, 6723−6732.

Clever, J. L., Wong, M. L., & Parslow, T. G. (1996) *J. Virol. 70*, 5902−5908.

Coffin, J. (1984) in *RNA Tumor Viruses* (Weiss, R., Teich, N., Varmus, H., & Coffin, J., Eds.) pp 261−368, Cold Spring Harbor Laboratory Press, Plainview, NY.

Feng, Y.-X., Copeland, T. D., Henderson, L. E., Gorelick, R. J., Bosche, W. J., Levin, J. G., & Rein, A. (1996) *Proc. Nat. Acad. Sci. U.S.A. 93*, 7577−7581.

Fossé, P., Motté, N., Roumier, A., Gabus, C., Muriaux, D., Darlix, J.-L., & Paoletti, J. (1996) *Biochemistry 35*, 16601−16609.

Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., & Turner, D. H. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 9373−9377.

Fu, W., Gorelick, R. J., & Rein, A. (1994) *J. Virol. 68*, 5013−5018.

Haddrick, M. A., Lear, A. L., Cann, A. J., & Heaphy, S. (1996) *J. Mol. Biol. 259*, 58−68.

He, L., Kierzek, R., Santa Lucia, J., Walter, A. E., & Turner, D. H. (1991) *Biochemistry 30*, 11124−11132.

Laughrea, M., & Moore, P. B. (1977) *J. Mol. Biol. 112*, 399−421.

Laughrea, M., & Jetté, L. (1994) *Biochemistry 33*, 13464−13474.

Laughrea, M., & Jetté, L. (1996a) *Biochemistry 35*, 1589−1598.

Laughrea, M., & Jetté, L. (1996b) *Biochemistry 35*, 9366−9374.

Laughrea, M., Jetté, L., Mak, J., Kleiman, L., Liang, C., & Wainberg, M. A. (1997) *J. Virol. 71*, 3397−3406.

---

[6] Background level of tight dimerization is defined for each temperature as the average level among three HIV-1_{Lai} RNAs ending at G508: one whose ACS has been replaced by GGGCGC (unpublished results); another whose GAAGCGCGC262 has been replaced by AGAUCACUC (Laughrea et al., 1997); and a third lacking the 5′ strand of stem C and most of loop C (Laughrea et al., 1997). This average background was 3% tight dimer at 49 °C in buffer L, 6% at 52 and 55 °C in buffer L, 3% at 58 °C in buffers L and H, 9% at 61 and 65 °C in buffer H, and <3% at other incubation conditions (Laughrea et al., 1997; and data not shown).

Marquet, R., Baudin, F., Gabus, C., Darlix, J.-C., Mougel, M., Ehresmann, C., & Ehresmann, B. (1991) *Nucleic Acids Res. 19*, 2349−2357.

Marquet, R., Paillart, J.-C., Skripkin, E., Ehresmann, C., & Ehresmann, B. (1994) *Nucleic Acids Res. 22*, 145−151.

McBride, M. S., & Panganiban, A. T. (1997) *J. Virol. 71*, 2050−2058.

Muriaux, D., De Rocquigny, H., Roques, B.-P., & Paoletti, J. (1996a) *J. Biol. Chem. 271*, 33686−33692.

Muriaux, D., Fossé, P., & Paoletti, J. (1996b) *Biochemistry 35*, 5075−5082.

Paillart, J.-C., Marquet, R., Skripkin, E., Ehresmann, B., & Ehresmann, C. (1994) *J. Biol. Chem. 269*, 27486−27493.

Paillart, J.-C., Berthoux, L., Ottmann, M., Darlix, J.-L., Marquet, R., Ehresmann, B., & Ehresmann, C. (1996a) *J. Virol. 70*, 8348−8354.

Paillart, J.-C., Skripkin, E., Ehresmann, B., Ehresmann, C., & Marquet, R. (1996b). *Proc. Nat. Acad. Sci. USA 93*, 5572−5577.

Peacock, A. C., & Dingman, C. W. (1967) *Biochemistry 6*, 1818−1827.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Mammal*, 2nd ed., Cold Spring Harbor Laboratory Press, Plainview, NY.

Skripkin, E., Paillart, J.-C., Marquet, R., Ehresmann, B., & Ehresmann, C. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 4945−4949.

Turner, D. H., Sugimoto, N., & Freier, S. M. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 167−192.

BI970862L