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†

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 H3T 1E2

Received February 19, 1996⊗

ABSTRACT: The genome of all retroviruses consists of two identical RNAs noncovalently linked near their 5′ end. Dimerization of genomic RNA is thought to modulate several steps in the retroviral life cycle, such as recombination, translation, and encapsidation. The kissing-loop model of HIV-1 genome dimerization posits that the 233−285 region of the HIV-1 genome, by forming a hairpin and initiating dimerization through a loop−loop interaction, is at least the core dimerization domain of HIV-1 RNA. This region is called the *kissing-loop domain*. In addition, it can be argued that sequences within the 296−401 region [Paillart et al. (1994) *J. Biol. Chem. 269*, 27486−27493] or 5′ of the primer binding site [Laughrea & Jetté (1996) *Biochemistry 35*, 1589−1598] might play some role in the dimerization process. Accordingly, we have studied the effect of regions 1−232 and 296−790 on the dimerization kinetics and thermal stability of HIV-1_{Lai} RNAs containing the kissing-loop domain (HIV-1_{Lai} is a typical representative of North American and European HIV-1 viruses). Experiments conducted at high and low ionic strength indicate that these regions have no strongly positive effect on the dimerization process. Our experiments also indicate that the kissing-loop domain of HIV-1_{Lai} has an apparent dissociation temperature 13 °C higher than that of the HIV-1_{Mal} kissing-loop domain (HIV-1_{Mal} is a Central African virus whose kissing-loop domain has a "weak" GUGCAC autocomplementary sequence). Because the 296−401 region of HIV-1_{Mal} RNA stabilizes dimeric RNAs by ≤12 °C (Paillart et al., 1994), we infer that the contributions of sequences downstream of U295 are (at best) concealed in HIV-1_{Lai} and in most American and European HIV-1 viruses, i.e., in viruses whose kissing-loop domain is characterized by a "strong" GCGCGC autocomplementary sequence.

The genome of all retroviruses, including human immunodeficiency virus type 1 (HIV-1),[1] is formed of two identical RNAs that are noncovalently linked near their 5′ ends [see Marquet et al. (1994) and references cited therein]. Dimerization of this diploid genome might modulate several steps of the retroviral life cycle, such as (i) translation of the genome into viral proteins; (ii) preferential packaging (or encapsidation) of two genomic RNAs within the budding viral shell (or capsid) against a vast excess of cellular RNAs; (iii) recombination between the two homologous genomic RNAs in order to increase the rate of viral evolution and to bypass occasional breaks in one of the genomic RNAs; (iv) stabilization of the genome against degradation; and (v) reverse transcription [see Laughrea and Jetté (1994) and references cited therein].

A kissing-loop model has been proposed to elucidate the mechanism of HIV-1 genome dimerization (Laughrea & Jetté, 1994; Skripkin et al., 1994). According to this model, region 248−270 of the HIV-1 genome forms a hairpin consisting of a 7 bp stem (called stem C) and a nine nucleotide loop (called loop C) containing the autocomplementary sequence GCGCGC262 or GUGCAC262 depending on the HIV-1 subtype (Figure 1a). (For simplicity, all nucleotide positions are converted to HIV-1_{Lai} positions whether or not the work was done with HIV-1_{Lai}.) Dimerization would occur through Watson−Crick hydrogen-bonding between loop C of one monomer and loop C of a nearby monomer. This could lead to the formation of two different types of dimers (Figure 1b,d). Hairpin 248−270 is located in the leader sequence of the HIV-1 genome, halfway between the primer binding site (which ends at C199) and the 5′ splice junction (located at GG290). We call the 233−285 region, which can form an imperfect hairpin, the *kissing-loop domain*. Several predictions of the kissing-loop model have so far been verified, such as the need for an autocomplementary loop C (Paillart et al., 1994; Skripkin et al., 1994) and an intact stem C (Laughrea & Jetté, 1994, 1996), the adoption of two dimeric forms by HIV-1_{Lai} RNAs containing the kissing-loop domain (Laughrea & Jetté, 1996), and the finding that the isolated 248−271 region readily dimerizes (Laughrea & Jetté, 1996).

† This work was supported by a grant from the Medical Research Council of Canada to M.L.

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

‡ Lady Davis Institute for Medical Research.

§ McGill University.

⊗ Abstract published in *Advance ACS Abstracts,* July 1, 1996.

[1] Abbreviations: HIV-1, human immunodeficiency virus type 1; nt, nucleotide; bp, base pair; 3′ DLS, name of an RNA sequence starting very near the 5′ splice junction (typically at nt 296) and ending at nt 401; 5′ DLS, name of an RNA sequence ending in or near the U5 region (i.e., upstream of the primer binding site) and which can dimerize in high-salt concentrations; T_d, apparent dissociation temperature; TBE_2/27 °C, buffer TBE_2 and in a room maintained at 27 °C.

S0006-2960(96)00395-9 CCC: $12.00

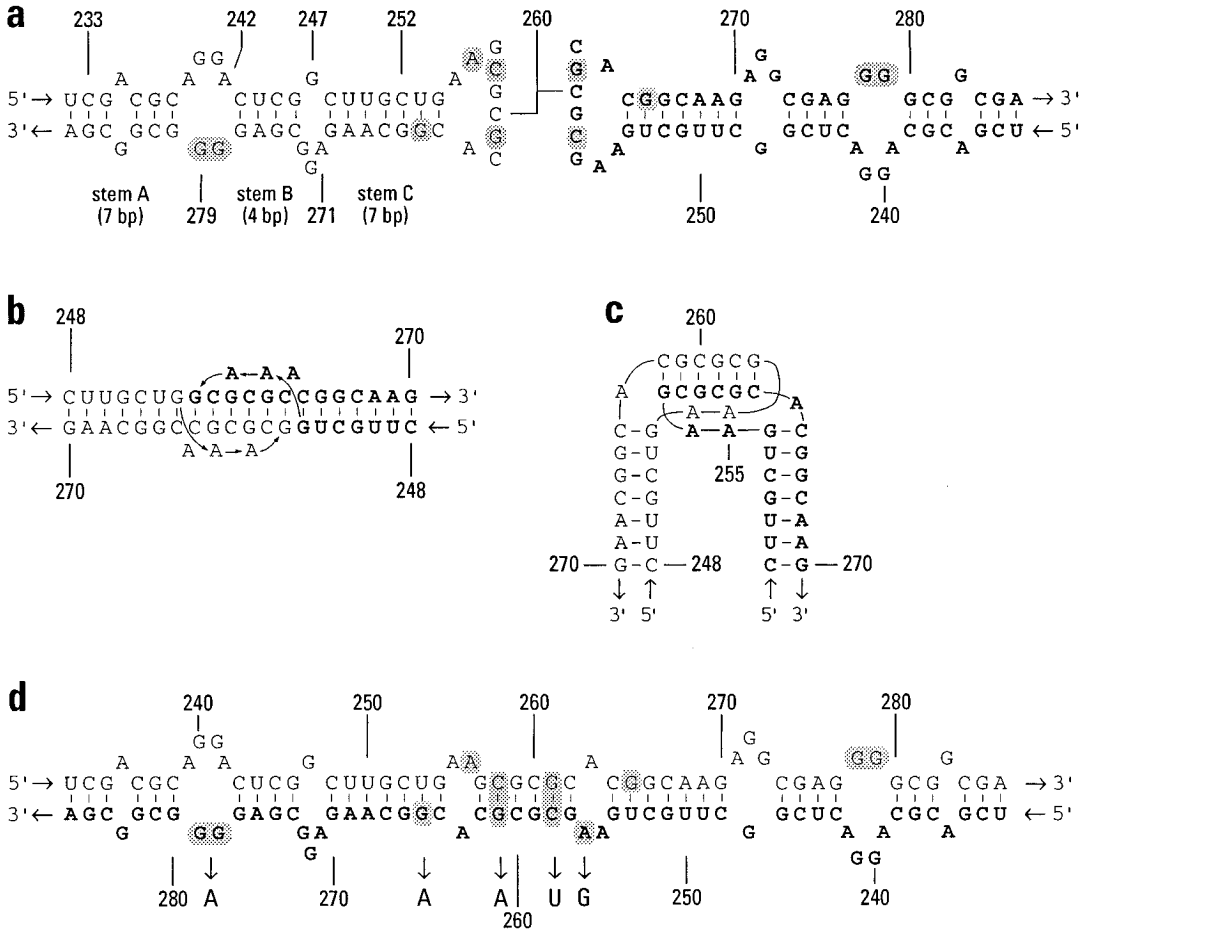

FIGURE 1: Secondary and tertiary structure model for the 232−286 region of HIV-1_{Lai} RNA in (a) monomeric, (b and c) loose dimeric, and (d) tight dimeric form (Laughrea & Jetté, 1996). In (a) and (d), nucleotides that differ in HIV-1_{Mal} and HIV-1_{Lai} are shadowed, and the corresponding HIV-1_{Mal} nucleotides are indicated by arrows in the bottom part of (d). (a) Postulated stem−loop structures of two monomers with their respective GCGCGC262 almost ready to hydrogen-bond to each other. (b) and (c) Possible secondary structures of the loose dimer. (d) Possible secondary structure of the tight dimer with one strand in boldface characters and the other in light face characters. Loose and tight dimers are defined a little further into the introduction.

The kissing-loop model is appealing because it has as least *two physiological connections*. First, it suggests a mechanism for the observation that immature virions have a dimeric HIV-1 genome that is less stable than in mature virions (Fu et al., 1994; Laughrea & Jetté, 1996). Second, the postulated hairpin overlaps with an encapsidation signal (Kim et al., 1994). Yet it is not clear if the kissing-loop domain is a *core* or the *whole* in vitro dimerization domain of HIV-1 genomic RNA. The notion that the kissing-loop domain might only be a core domain stems from the fact that HIV-1 RNAs ending upstream of the primer binding site (Laughrea & Jetté, 1996) or starting downstream of the 5′ splice junction (Marquet et al., 1991) can be shown to dimerize under some ionic conditions. This led to the notion of a 5′ DLS (dimer linkage sequence) and a 3′ DLS within the HIV-1 genome (Laughrea & Jetté, 1996). The 3′ DLS is located within the 296−401 region (Paillart et al., 1994; Laughrea & Jetté, 1996). We have shown that the 5′ and 3′ DLS sequences lose the ability to dimerize when respectively 3′ and 5′ extended to a small extent (Laughrea & Jetté, 1994, 1996) and that the 5′ and 3′ DLS do not improve the dimerization yields of RNAs containing the kissing-loop domain (Laughrea & Jetté, 1996). However, this does not rule out a possible influence on the dimerization kinetics and/or the thermal stability of the formed dimers. It was recently shown at high ionic strength that sequences within the 3′ DLS increased the thermal stability of HIV-1_{Mal} RNA transcripts (Paillart et al., 1994). But HIV-1_{Mal} is an unusual strain, at least in an American/Asian/European context. It is a hybrid, hard to classify, Central African virus with a subtype A *gag*, a subtype D *env*, and a group O leader sequence[2] characterized by a large 20 nt insert between A217 and C218, i.e., halfway between the primer binding site and the kissing-loop domain (Myers et al., 1993, 1994). The dimerization of HIV-1_{Mal} RNAs might thus display a mixture of universal and idiosyncratic properties. In addition, the influence of the 5′ DLS on the dimerization kinetics and the thermostability of HIV-1 RNAs had never been studied, nor had the putative

[2] Phylogenetically, HIV-1 isolates usually divide themselves into two divergent groups: group M, for major, and group O, for outlier (Charneau et al., 1994). Group M isolates can be divided into eight equidistant genotypes, from subtype A to subtype H (Myers et al., 1994). Group O isolates are considerably divergent and could themselves be divided into at least two subtypes (Charneau et al., 1994). HIV-1_{Mal} and subtype A leaders have the same GUGCAC autocomplementary sequence. Subtypes B and D (which represent 90% of all HIV-1 viruses sequenced) bear a GCGCGC autocomplementary sequence (Myers et al., 1993, 1994; Laughrea & Jetté, 1996). The autocomplementary sequence of the other subtypes is unknown. Though there are no international borders for HIV-1, group O and subtype D viruses are mostly found in Africa. Subtype A viruses are found in Africa, Asia, and Europe. Subtype B viruses are found on all continents and are strongly prevalent in North America.

stabilizing influence of the 3′ DLS been verified at low ionic strength.

The purpose of this paper is to test whether the 5′ and/or 3′ DLS sequences might strongly influence the dimerization kinetics of HIV-1_{Lai} RNA transcripts or the apparent dissociation temperature (T_d) of the formed dimers. HIV-1_{Lai} is a typical subtype B virus, and subtype B viruses account for most HIV-1 infections in North America, South America, and Europe (Myers et al., 1994). Specifically, we have studied the effect of sequences upstream of C233 and downstream of U295 on the dimerization kinetics and the T_d of HIV-1_{Lai} RNAs containing the kissing-loop domain. Dimerization kinetics of retroviral RNAs have always been assayed using nondenaturing gel electrophoresis to distinguish dimers from monomers (Roy et al., 1990; Marquet et al., 1994; Paillart et al., 1994; Girard et al., 1995). Identically, thermal analysis of retroviral dimeric RNAs has always involved nondenaturing gel electrophoresis, no matter whether the RNA under study was from HIV-1 (Darlix et al., 1990; Marquet et al., 1991, 1994; Awang & Sen, 1993; Sundquist & Heaphy, 1993; Fu et al., 1994; Paillart et al., 1994; Muriaux et al., 1995) or from other retroviruses (Bender et al., 1978; Maisel et al., 1978; Prats et al., 1990; Roy et al., 1990; Fu & Rein, 1993; Feng et al., 1995; Girard et al., 1995; Lear et al., 1995), except for one early study (Stoltzfus & Snyder, 1975), which involved separation of monomers from dimers by glycerol gradient centrifugation. Though alternative methods might be available, especially to estimate T_d (e.g., Romby et al., 1985; Riesner et al., 1991; Jaeger et al., 1993), we felt that it would be advantageous to obtain data directly comparable to previous studies of retroviral RNA dimerization, thus, our use of the economic technique of nondenaturing gel electrophoresis to assess the rates of dimerization and the thermal resistance of dimeric HIV-1_{Lai} RNA transcripts.

Since HIV-1_{Lai} RNAs can form two types of dimers, respectively called loose and tight dimers (Laughrea & Jetté, 1996), we have studied the dimerization kinetics and the T_d of both types of dimers. Loose dimers remain dimeric after electrophoresis in a Mg^{2+}-free/EDTA-containing buffer at 4 °C, but not after electrophoresis in the same buffer at 27 °C (Laughrea & Jetté, 1996). Tight dimers resist either condition (Laughrea & Jetté, 1996). At physiological temperatures, and with the RNAs to be studied in this paper, only loose dimers are formed in the absence of nucleocapsid protein (Laughrea & Jetté, 1996). Rapid formation of tight dimers in the absence of nucleocapsid protein usually requires incubation at ≥55 °C (Laughrea & Jetté, 1996).

Kinetic and thermal analysis was done at low ionic strength (0.1 mM Mg^{2+} and 90 mM monovalent cations) and at high ionic strength (5 mM Mg^{2+} and 350 mM monovalent cations). The high ionic strength buffer, though it can sometimes promote apparently artifactual interactions between some RNA molecules (Bloomfield et al., 1974; Laing & Draper, 1994; Laughrea & Jetté, 1996), allows direct comparison between the kinetic and thermal properties of loose HIV-1_{Lai} dimers and those, published by Marquet et al. (1994) and Paillart et al. (1994), of HIV-1_{Mal} RNA dimers. We will note significant differences that will be rationalized under Discussion.

Our results can be summarized as follows. (i) Sequences upstream of C233 (which include the 5′ DLS) do not strongly improve the dimerization process. (ii) The previously proposed dimerization domain (Darlix et al., 1990; Marquet et al., 1991; Awang & Sen, 1993; Sundquist & Heaphy, 1993), i.e., region 296−401 (formerly called DLS and now called 3′ DLS), does not increase the thermal stability of HIV-1_{Lai} RNA dimers in both low- and high-salt buffers, and does not improve the dimerization kinetics of loose dimers. It might improve by at most 2- or 3-fold the dimerization kinetics of tight dimers. We will propose that the differences between our results and those of Paillart et al. (1994) are related to the fact that subtype B and D viruses have a kissing-loop domain with a stronger autocomplementary sequence than subtype A and HIV-1_{Mal} viruses.

MATERIALS AND METHODS

Buffers. Buffer L (Marquet et al., 1991; Laughrea & Jetté, 1994): 50 mM sodium cacodylate, pH 7.5, 40 mM KCl, and 0.1 mM $MgCl_2$. Buffer H (Marquet et al., 1991; Laughrea & Jetté, 1994): 50 mM sodium cacodylate, pH 7.5, 300 mM KCl, and 5 mM $MgCl_2$. TBE_2 and $TBE_{0.5}$ (Peacock & Dingman, 1967): 89 mM Tris, 89 mM borate, and respectively 2 and 0.5 mM EDTA. TBM (Laughrea & Moore, 1977; Laughrea & Jetté, 1994; Marquet et al., 1994): same as TBE_2 but with the EDTA replaced by 1 mM $MgCl_2$.

Materials. Restriction enzymes were from Pharmacia (Montreal, Quebec), except for *Hga*I (New England Biolabs, Mississauga, Ontario) and *Tru*9I and *Mae*I (Boehringer Mannheim, Montreal, Quebec). Agarose, electrophoresis grade, was from ICN.

Plasmid Construction, RNA Synthesis, and Recovery. All plasmids described in this paper have a fraction of the HIV-1_{Lai} genome cloned immediately (or a few nucleotides) downstream from a strong T7 RNA polymerase promoter. Plasmids pBL, pB233, pLJ230, and pLJ242 have been described in Laughrea and Jetté (1994). They yield T7 transcripts respectively starting at G1, C233, U230, and A242 of the HIV-1_{Lai} genome. Prior to T7 RNA polymerase transcription, plasmids were cleaved with *Rsa*I, *Mae*I, *Hga*I, *Tru*9I, *Sau*3AI, *Hae*III, *Acc*I, *Hin*dIII, or *Dra*I to obtain a collection of RNAs ending at U295, A314, G342, A359, C378, G401, G508, U635, and U790, respectively. Transcription and RNA recovery were done using the MEGAscript kit of Ambion (Austin, TX). 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 150 ng−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 the samples were incubated at 37 °C for 30 min (or other times if a kinetic experiment was performed). For preparation of tight dimers (Laughrea & Jetté, 1996), 2 μL of 5× concentrated buffer L was added, and the samples were preincubated 30 min at 0 °C and incubated at 58−60 °C for 10 min (unless a kinetic experiment was performed). For kinetic measurements, incubations in a 10 μL volume proceeded for periods ranging from 10 s to 40 min. The rate constant for the dimerization was then derived assuming a simple bimolecular reaction: monomer + monomer ↔ dimer. For T_d measurements, different 10 μL samples of the same dimeric RNAs were simultaneously postincubated

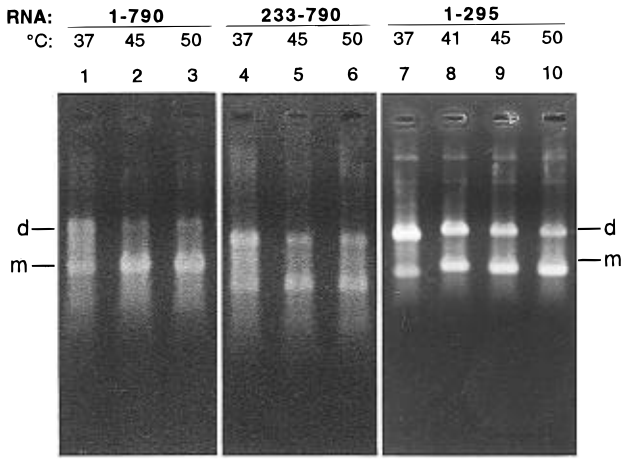

FIGURE 2: Thermal stability of loose dimers of RNAs 1−790, 233−790, and 1−295 at low ionic strength. 500 ng of RNA was preincubated at 37 °C for 30 min in buffer L and incubated at the indicated temperatures for 10 min immediately before loading. Lanes 1−3: RNA 1−790. Lanes 4−6: RNA 233−790. Lanes 7−10: RNA 1−295. Electrophoresis was in TBE_2/4 °C and on 2.2% agarose (lanes 1−6) or 3% agarose (lanes 7−10).

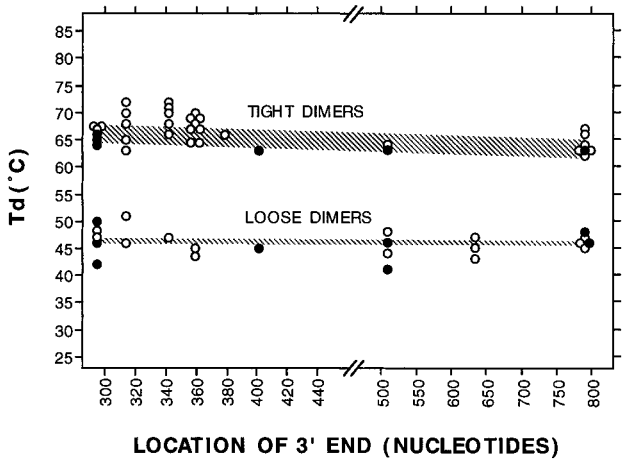

FIGURE 3: Low ionic strength thermostability of dimeric HIV-1 RNAs starting at G1 (●), or in the 230−242 region (○), as a function of the location of their 3′ end (U295, A314, G342, A359, C378, G401, G508, U635, or U790). The bottom part of the graph displays the temperature at which half of the loose dimers formed in buffer L were dissociated (T_d). The upper part displays the T_d of the tight dimeric form of the same RNAs. For estimating the T_d of loose dimers, gel electrophoresis was in 3% agarose in TBE_2/4 °C or in 6% acrylamide/bis(acrylamide) gel in TBE_2/4 °C. The results were independent of the electrophoretic procedure (not shown). To estimate the T_d of tight dimers, gel electrophoresis was as above but in a 27 °C room (Laughrea & Jetté, 1996).

for 10 min at three or four different temperatures. The percent dimerization was assayed by gel electrophoresis in TBE_2 at 4 °C (TBE_2/4 °C) or in TBE_2 at 27 °C (TBE_2/27 °C). In short, after each kinetic and T_d experiment, 150−900 ng of RNA was loaded on agarose or acrylamide gels without delay and with the voltage on. Agarose gel electrophoresis has been described previously (Laughrea & Jetté, 1996). Electrophoretic separation of RNAs <600 nt in monomeric length was often performed using a XCELL II mini-cell electrophoresis apparatus (Novex/Helixx, Scarborough, Ontario) and 6% acrylamide precast gels in TBE_2 (Novex/Helixx). Half the RNA material appropriate for agarose gel electrophoresis could be loaded on acrylamide gels and yet give an equivalent ethidium bromide staining response. All gels were stained as in Laughrea and Jetté (1996), and the percent dimerization was estimated by scanning the photographic negatives or the autoradiograms with an LKB Ultroscan XL laser densitometer.

RESULTS

Effect of the 5′ and 3′ DLS Sequences on the Apparent Dissociation Temperature (T_d) of Loose Dimers. Plasmids pBL, pLJ230, pB233, and pLJ242 were linearized at U790 or at various positions upstream of A636 prior to transcription. Eighteen RNAs were produced. They started at G1, U230, C233, or A242 of the HIV-1_{Lai} genome and ended at U295, A314, G342, A359, C378, G401, G508, U635, or U790. They were preincubated at 37 °C in buffer L, a low ionic strength buffer containing 90 mM monovalent cations and 0.1 mM Mg^{2+}, and electrophoresed at 4 °C in an agarose gel containing buffer TBE_2. Some examples are shown in Figure 2. The numbers in the name of an RNA refer to its first and last HIV-1 RNA nucleotide. The low mobility material marked "d" represents loose dimers (Laughrea & Jetté, 1996). No tight dimers are formed under these conditions (Laughrea & Jetté, 1996). [For yields of loose dimers that approach 100%, preincubation in buffer H, a high ionic strength buffer containing 350 mM monovalent cations and 5 mM Mg^{2+}, is recommended (Laughrea & Jetté, 1996).] Electrophoresis in TBE_2 at 4 °C (TBE_2/4 °C) was adopted because an important requirement for estimating relative T_ds by gel analysis is the use of electrophoretic conditions that neither break down dimers nor favor dimer formation. Preliminary experiments showed that TBE_2/4 °C was superior to $TBE_{0.5}$/4 °C and TBM/4 °C in this respect (not shown). TBE_2/4 °C did not stimulate the formation of dimers from monomers and minimally dissociated some formed dimers (not shown; Laughrea & Jetté, 1996).

Figure 2 shows that HIV-1_{Lai} RNA 1−295 does not have a lower T_d than RNA 1−790 and that RNA 233−790 has the same T_d as RNA 1−790. These results were confirmed by the bottom part of Figure 3, which displays a larger range of experiments. Figure 3 shows that at low ionic strength the T_d of loose dimers is 46 ± 2.3 °C no matter the location of the 5′ end within the range 1−242 (compare white dots to black dots at bottom of Figure 3), and no matter the location of the 3′ end within the range 295−790. Thus, the 5′ and 3′ DLS do not stabilize loose dimers formed at low ionic strength or stabilize them by <2.5 °C. Under identical ionic conditions, HIV-1_{Mal} RNAs were completely monomeric at 37 °C (Paillart et al., 1994), suggesting that the dimeric kissing-loop domain of HIV-1_{Mal} has a T_d <37 °C.

In buffer H, the T_d of all RNAs studied in Figure 3 was 55 ± 2.5 °C, no matter the location of their 5′ or 3′ end (data not shown). Thus, the 5′ and 3′ DLS do not stabilize loose dimers formed at high ionic strength by HIV-1_{Lai} RNAs or stabilize them by <2.5 °C. Under identical ionic conditions, HIV-1_{Mal} RNAs 1−295 and 1−601 had respective T_ds of 41 and 52 °C (Paillart et al., 1994). Together, these results suggest that short HIV-1_{Lai} dimeric RNAs are at least as stable as long HIV-1_{Mal} dimeric RNAs and much more stable than short HIV-1_{Mal} dimeric RNAs.

Effect of the 5′ and 3′ DLS on the T_d of Tight Dimers. To form tight dimers, HIV-1 RNA transcripts were preincubated in buffer L at 0 °C and incubated at 60 °C for 10 min (Laughrea & Jetté, 1996). To estimate their T_d, the samples

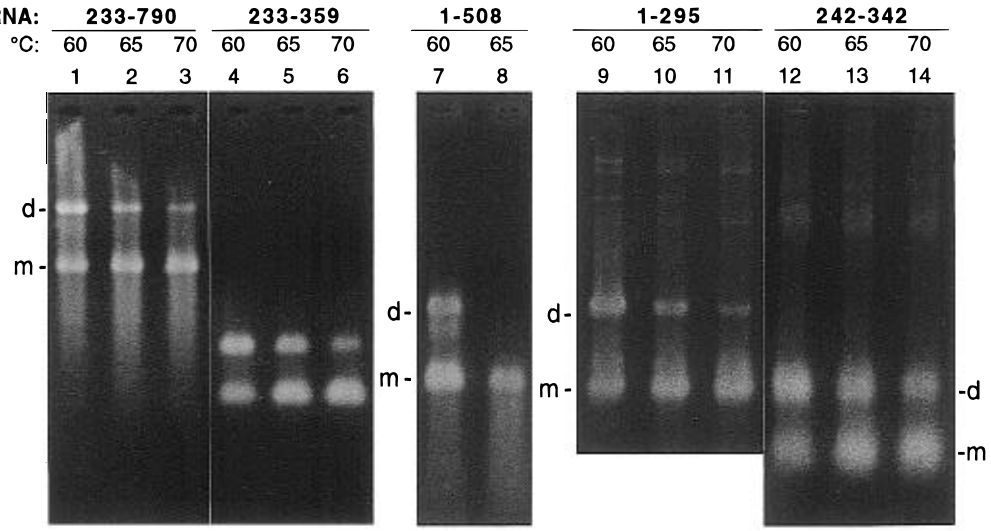

FIGURE 4: Thermal stability of tight dimers. To form tight dimers, all RNAs were preincubated 30 min in buffer L at 0 °C and incubated at 60 °C for 10 min. To estimate the T_d, this was followed by a 10 min postincubation at the indicated temperatures and electrophoresis (70 V) in TBE_2/27 °C on 3% agarose (lanes 1−6 and 9−14) or on 2.2% agarose (lanes 7−8). Lanes 1−8 were electrophoresed for 100 min and lanes 9−14 for 86 min. The quantities of RNA loaded were 900 ng of 233−790 and 1−508, 800 ng of 1−295, 650 ng of 242−342, and 500 ng of 233−359.

were postincubated for 10 min at 60, 65, 70, and 75 °C, and electrophoresed at 27 °C in a gel containing TBE_2. TBE_2/27 °C was chosen because it is a fool-proof way to study tight dimers: no matter how the RNAs of this paper are handled between the postincubation step and the time of electrophoresis, no tight dimer can form from monomers (unless the handling is done at ≥55 °C), and if loose dimers are inadvertently formed, they disappear upon electrophoresis in TBE_2/27 °C (Laughrea & Jetté, 1996).

The top part of Figure 3 shows that sequences 3′ of U295 or 5′ of U230 did not stabilize tight dimeric RNAs. Specifically, RNAs starting at G1 had an average T_d of 64 ± 1.3 °C (black dots at top of Figure 3), and RNAs lacking the 5′ DLS (i.e., starting within the 230−242 region) had an average T_d of 67 ± 2.8 °C (white dots at top of Figure 3). The T_ds were independent of whether the RNAs started at U230, C233, or G242 (Figure 3 and data not shown), indicating that the inactivity of the 5′ DLS is unlikely to be an idiosyncratic truncation artifact. Similarly, RNAs ending upstream of G315 have an average T_d of 66.8 ± 2.6 °C while RNAs possessing a complete 3′ DLS (i.e., ending downstream of G400) had an average T_d of 63.8 ± 1.5 °C. These conclusions are underlined by Figure 4, which shows some gel lanes from which the line at the top of Figure 3 was constructed. For example, RNA 1−508 (lanes 7−8) had a slightly lower T_d than RNA 1−295 (lanes 9−11), despite possessing a complete 3′ DLS, and RNA 233−790 (lanes 1−3) had a T_d slightly lower than that of RNA 233−359 (lanes 4−6) and clearly lower than that of RNA 242−342 (lanes 12−14), in conformity with some of the data points of Figure 3. One also notes that RNA 1−508, despite possessing a complete 5′ DLS, had a T_d at best equal to that of RNA 233−790 (compare lanes 7−8 to lanes 1−3).

Effect of the 5′ and 3′ DLS Sequences on the Dimerization Kinetics of Loose Dimers. Figure 5a compares in high salt conditions the dimerization kinetics at 37 °C of RNAs starting at G1 (black dots) or at C233 (white dots) and ending within the 295−790 region. The kinetics of formation of loose dimers were somewhat variable, a reminder of previous results obtained with HIV-1_{Mal} RNA transcripts (Paillart et al., 1994), so that we feel unable to distinguish two k_1s that differ less than 2-fold. With this caveat in mind, Figure 5a shows that the kinetics of formation of loose dimers were not improved by sequences downstream of U295. The rate constants were undistinguishable from those obtained by Marquet et al. (1994) and Paillart et al. (1994) with HIV-1_{Mal} RNAs 1−295 and 1−601, as if dimerization kinetics in buffer H at 37 °C were relatively independent of whether the autocomplementary sequence of the kissing-loop domain

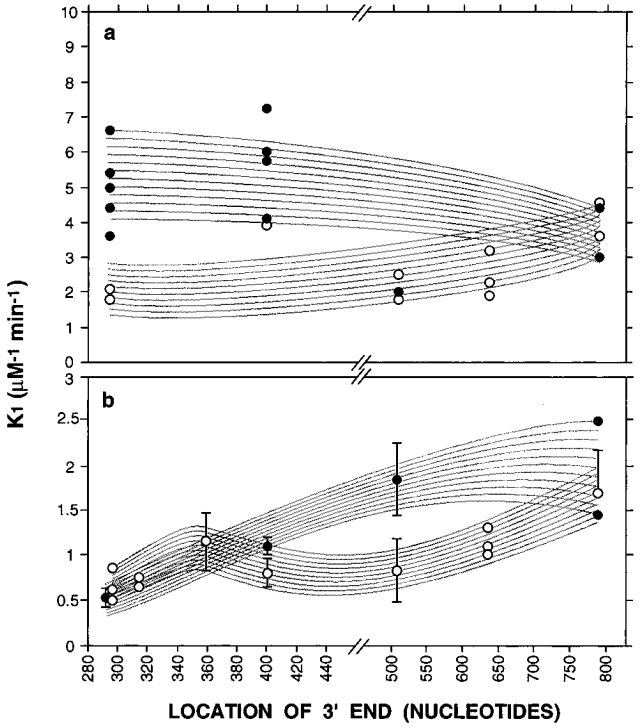

FIGURE 5: Kinetic rate constants of formation of (a) loose dimers and (b) tight dimers as a function of the location of the 3′ end of RNAs starting at G1 (●) or at C233 (○). Dimerization of loose dimers was studied in buffer H at 37 °C. Dimerization of tight dimers was studied in buffer L at 60 °C, after a 30 min preincubation at 0 °C. The RNAs studied ended at U295, A314, A359, G401, G508, U635, or U790. For some of them, the data are expressed as average k_1 ± SD.

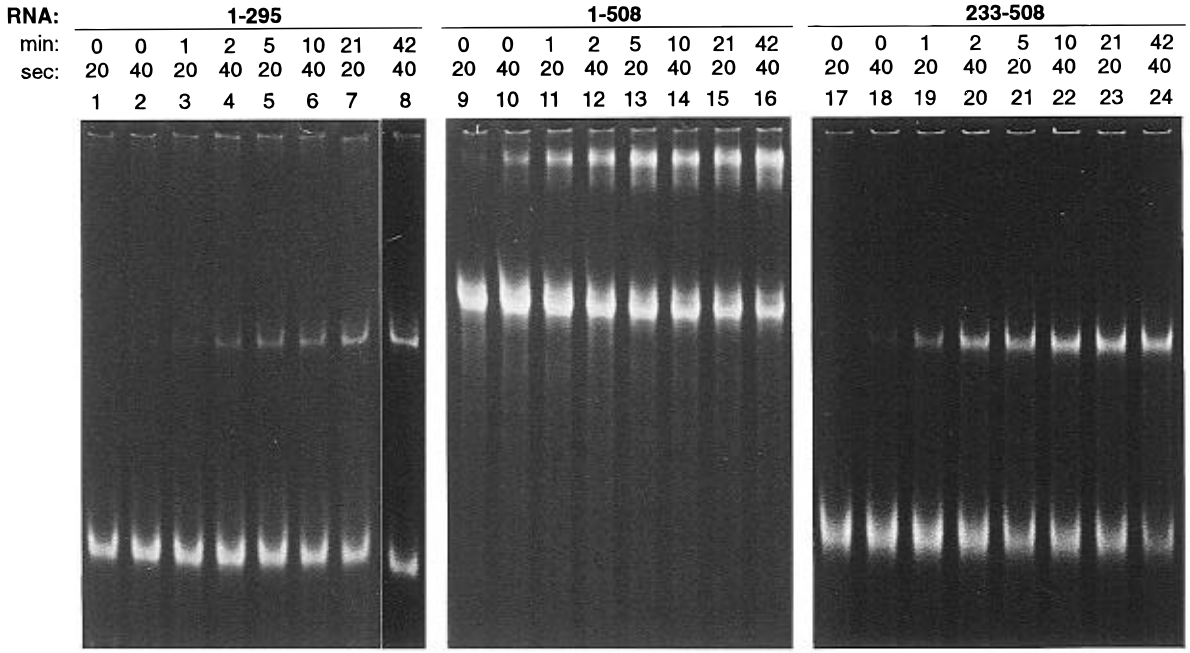

FIGURE 6: Rate of formation of tight dimers from RNAs 1−295, 1−508, and 233−508. All RNAs were preincubated 30 min in buffer L at 0 °C before incubation at 60 °C for times varying from 20 s to 21 min, 20 s. 150 ng of RNA 1−295, 200 ng of RNA 233−508, and 300 ng of RNA 1−508 were loaded on the 6% acrylamide gel in TBE_2/27 °C which was run at 112 V for 140 min.

is GCGCGC (as in HIV-1_{Lai}, subtype B and subtype D viruses) or GUGCAC (as in HIV-1_{Mal} and subtype A viruses). It is interesting to note that the rate constants were 20−50 times higher than those for the dimerization of Moloney murine leukemia virus RNA transcripts at 37 °C in 50 mM Tris and 100 mM NaCl (Girard et al., 1995). This suggests that loose dimerization of HIV-1 RNAs involves smaller conformational changes than the dimerization of Moloney murine leukemia virus RNA transcripts.

It is unclear whether nucleotides upstream of C233 improve the dimerization kinetics of loose dimers: they might increase the rate constant of short RNAs, but not that of long ones (compare white dots to black dots in Figure 5a). The general kinetic picture was unchanged at low ionic strength except that the rate constants were about 50% smaller. Specifically, the k_1 values of RNAs 1−295 and 1−401 were respectively 2.5 ± 0.9 and 3.3 ± 1.6 μM^{-1} min^{-1} in buffer L at 37 °C (not shown). Those of RNAs 233−295 and 233−401 were 0.8 ± 0.2 and 1.6 ± 0.5 μM^{-1} min^{-1} (not shown).

Effect of the 5′ and 3′ DLS on the Dimerization Kinetics of Tight Dimers. Figure 5b shows that sequences upstream of C233 usually did not improve the kinetics of tight dimerization (compare white dots to black dots). On the other hand, Figure 5b is consistent with the idea that the 315−359 or the 315−790 region increases k_1 by a factor of perhaps 2 or 3. This tendency is more convincing with RNAs starting at G1 (black dots in Figure 5b): RNAs 1−508 and 1−790 had a k_1 of 1.5−2.5 μM^{-1} min^{-1} vs a k_1 of 0.4−0.6 μM^{-1} min^{-1} for RNA 1−295. It is worth noting that the rate of tight dimerization at 60 °C was slower or at best equal to the rate of loose dimerization at 37 °C (compare Figure 5b to Figure 5a). Three examples of tight dimerization kinetics are shown in Figure 6. RNA 1−295 had the lowest rate constant among the three displayed RNAs (k_1 = 0.44 μM^{-1} min^{-1} at 60 °C and in low salt conditions). RNA 1−508 had the highest rate constant (2.0 μM^{-1} min^{-1}), and RNA 233−508 had an intermediate one of 1.3 μM^{-1} min^{-1}.

DISCUSSION

Relevance of in Vitro Studies. In this and the preceding papers (Laughrea & Jetté, 1994, 1996), we have studied 69 HIV-1 RNA transcripts containing different sequences from the 1−790 region of the HIV-1_{Lai} genome. By and large, *consistent* results were obtained with these RNAs. This self-consistency, coupled to the fact that in all retroviruses studied by electron microscopy the dimer linkage structure has been localized within the first 600 nts of the genome (Bender & Davidson, 1976; Kung et al., 1976; Bender et al., 1978; Murti et al., 1981), gives confidence that the common folding of these RNAs is not an artifact and that it bears on the in vivo conformation of the dimerization domain of the HIV-1 genome.

Adequacy of Gel Analysis To Estimate Relative Thermal Stability. It was not the goal of this study to derive absolute thermodynamic parameters for the HIV-1 RNA dimerization process [thermodynamic parameters, such as ΔS and ΔG, are anyhow solvent-dependent (Puglisi & Tinoco, 1989)] but to identify, within a given large (790 nucleotides) sequence domain, which HIV-1 RNA sequences play a dominant role in the *in vitro* dimerization process and which do not. This was done by studying the *differential* behavior of RNAs truncated at many different positions. Thus, the meaningful results of this paper are not so much the exact numerical values obtained for apparent dissociation temperatures and apparent rate constants but *the sequence dependence of these values*.

We are aware that T_ds obtained by our gel electrophoretic procedure cannot be regarded as absolute because the data are not taken at equilibrium and because the incubation buffer is usually different from the electrophoresis buffer. However, such T_ds are useful when the goal is (i) to compare RNAs of different sequences and (ii) to compare our results to previous *thermal analyses* of retroviral RNA dimerization. Had we not used the present analysis to obtain T_ds, we would have been unable to compare our results to the apparently

variant results of Paillart et al. (1994). As will be seen below, the possibility of a direct comparison will lead to a global interpretation that would have otherwise been impossible.

For estimating T_ds, differential UV absorbance melting is an economic technique of potential interest. It has recently been profitably used to study RNA molecules as large as 399 nt long group 1 ribozymes (Jaeger et al., 1990; Banergee et al., 1993). But it remains undemonstrated that optical melting could reliably detect the melting of a short duplex within RNA dimers as large as those studied here (Draper & Gluick, 1995). Even tRNAs have multiphasic melting profiles that differ dramatically from one isoaccepting tRNA to the other within the same species (Privalov & Filiminov, 1978). And a 98 nt domain of the 23S RNA of *E. coli* ribosomes has a quite diffuse absorbance melting profile (Laing & Draper, 1994), suggesting gradual unfolding over a temperature range from 20 to 80 °C. It appeared likely that in RNAs as large as those studied in this paper, the monomerization transition would easily get lost within overlapping tertiary folding transitions. And had we against all odds succeeded, we would have been unable to directly compare the T_ds obtained by optical melting to published T_ds of HIV-1 RNA dimers.

Tetra- and Dimorphism of the 3′ DLS: Which Is the Physiological Conformation? From an overview of available data on the role of the 3′ DLS in the dimerization process, it appears that at high ionic strength this region can adopt four different conformations depending on the sequence context: three dimeric conformations (Skripkin et al., 1994) and one unable to dimerize (Laughrea & Jetté, 1994, 1996). It also appears that the two most stable dimeric conformations are truncation artifacts because their high T_ds have only been seen with specifically shortened RNAs (Skripkin et al., 1994). The 3′ DLS of HIV-1 RNAs with an intact 1−500 region would thus be *at most dimorphic*. When located within a physiological HIV-1_{Lai} sequence context, the 3′ DLS does not make any strongly positive contribution to the dimerization process at high and low ionic strength (shown in the present paper), as if it were unable to dimerize or as if its putative contribution was in all respects smaller or at best comparable to that of the kissing-loop domain. When located within a physiological HIV-1_{Mal} sequence context, the 3′ DLS contributes to the stabilization of the dimeric RNA at high ionic strength, increasing its T_d by 10 °C (Paillart et al., 1994). HIV-1_{Mal} RNAs do not easily dimerize at low ionic strength (Marquet et al., 1994), contrary to HIV-1_{Lai} RNAs (Laughrea & Jetté, 1994; Clever et al., 1995; Muriaux et al., 1995).

A Dimerization Model for All HIV-1 Subtypes: The Contribution of the 3′ DLS Is Triggered by More Than Mere Dimerization of the Kissing-Loop Domain. How to accommodate the variances between HIV-1_{Mal} and HIV-1_{Lai} results? Perhaps the non-HIV-1-coded purine-rich tail located at the 3′ end of the internally deleted RNAs of Paillart et al. (1994) caused artifactual dimerization in some mutants. More plausibly, we believe that the T_d of the dimeric kissing-loop domain depends on the HIV-1 subtype under study, i.e., at the molecular level, on the sequence of the autocomplementary sequence of the kissing-loop domain.

The first clue is that HIV-1_{Lai} and HIV-1_{Mal} RNAs deprived of the 3′ DLS and studied at the same concentration (∼500 nM) have widely different T_ds at high ionic strength: 55 °C in HIV-1_{Lai} (this paper) and 41 °C in HIV-1_{Mal} (Paillart et al., 1994). The second clue is that the presence of the 3′ DLS in HIV-1_{Mal} RNAs increases their T_d at high ionic strength to a temperature *equal or slightly inferior (52 ± 2 °C)* to that of the dimeric HIV-1_{Lai} kissing-loop domain (Paillart et al., 1994; this paper). The third clue is that the kissing-loop domains of HIV-1_{Lai} and HIV-1_{Mal} have different autocomplementary sequences. HIV-1_{Lai}, like subtype B and D viruses (and 90% of all HIV-1 leaders sequenced), bears a strong GCGCGC262 autocomplementary sequence (Myers et al., 1993, 1994; Laughrea & Jetté, 1996). HIV-1_{Mal} and subtype A viruses bear a presumably weaker GUGCAC262 autocomplementary sequence. This leads to a generic kissing-loop model governed by four parameters. First, the 3′ DLS, if and when it contributes to the dimerization process, increases the T_d of ∼500 nM dimeric HIV-1 RNAs to at most 54 °C at high ionic strength (Paillart et al., 1994). Second, in the absence of any 3′ DLS, the T_d of the dimeric kissing-loop domain is 13−14 °C higher with GCGCGC262 than with GUGCAC262 as autocomplementary sequence [this paper and Paillart et al. (1994)]. [Experiments with 0.1 mM oligonucleotides show that a GCGCGC duplex should have a T_d 15 °C higher than a GUGCAC duplex and 20 °C higher than a GUGCGC duplex (Freier et al., 1986; Turner et al., 1988; He et al., 1991).] Third, whether or not the 3′ DLS stabilizes formed dimers is determined by the precise association constant or the precise shape/charge distribution of the dimeric kissing-loop domain: mere dimerization 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 (see below). Finally, the T_d of dimeric HIV-1 RNAs is 9 °C lower in buffer L than in buffer H (this paper).

Five Accounts from the Generic Kissing-Loop Model Viewpoint. (i) The generic kissing-loop model accounts for the failure of ∼500 nM HIV-1_{Mal} RNAs to dimerize when incubated at 37 °C at low ionic strength (Marquet et al., 1994): because the T_d of the HIV-1_{Mal} dimeric kissing-loop domain is about 41 °C in buffer H (Marquet et al., 1994; Paillart et al., 1994), it follows that its T_d shall be ∼32 °C in buffer L and that HIV-1_{Mal} RNAs will stay monomeric in buffer L at 37 °C. (ii) The generic model accounts for the failure of the 3′ DLS to make a strong contribution to the HIV-1_{Lai} dimerization process (this paper): the most the 3′ DLS could do in an RNA possessing the HIV-1_{Lai} kissing-loop domain is a concealed contribution, i.e., one overshadowed by the high T_d associated with the GCGCGC262 autocomplementary sequence of HIV-1_{Lai}. It is an open question whether the 3′ DLS makes a functional contribution to the dimerization of HIV-1_{Lai} RNAs, because the strong dimeric kissing-loop domain of HIV-1_{Lai} may not allow as good a functional connection with the 3′ DLS as the weaker dimeric kissing-loop domain of HIV-1_{Mal}. (iii) The generic model accounts for the absence of T_d change associated with the replacement of GUGCAC262 of HIV-1_{Mal} RNA 1−601 by GCGCGC (Paillart et al., 1994): the high T_d of wild-type HIV-1_{Mal} RNA 1−601 is due to a stabilizing contribution of the 3′ DLS while the identical T_d of GCGCGC HIV-1_{Mal} RNA 1−601 is due to its stronger kissing-loop domain (whether or not the 3′ DLS contributes anything). It will be interesting to know if the T_d of GCGCGC HIV-1_{Mal} RNA 1−295 is about 10−15 °C higher than that of wild-type HIV-1_{Mal} RNA 1−295. (iv) The generic model accounts for the large change in T_d (14 °C) associated with the replacement

of A261 of HIV-1_{Mal} RNA 1−601 by G261, to make a theoretically barely weaker GUGCGC262 autocomplementary sequence (Paillart et al., 1994): the functional link to the 3′ DLS is lost because of the different association constant or the different shape/charge distribution of the mutant dimeric kissing loop. As a result, the 3′ DLS has lost its stabilizing influence, and the T_d of mutant RNA 1−601 is postulated to be the same as that of the dimeric GUGCGC kissing-loop domain. It will be interesting to know if the T_ds of GUGCGC HIV-1_{Mal} RNA 1−295 and GUGCGC HIV-1_{Mal} RNA 1−601 are identical. (v) Finally, the generic model accounts for our observation that an internal deletion within stem C of the HIV-1_{Lai} kissing-loop domain reduces to an equal extent the T_d of all tested dimeric RNAs no matter whether they contain the 3′ DLS (in RNAs 1−401Δ242−251 and 1−790Δ242−251) or lack it (in RNA 1−295Δ242−251). These three RNAs have the 242−251 region deleted. Their T_ds at ∼500 nM were 48 °C in buffer H and 39 °C in buffer L (data not shown). The absence of any impact of the 3′ DLS in these RNAs has an interesting implication. The mutant HIV-1_{Lai} kissing-loop domain was not a weak one: its T_d was 7 °C higher than that of the HIV-1_{Mal} kissing-loop domain. Thus, if the HIV-1_{Mal} connection between the kissing-loop domain and the 3′ DLS was due to the self-affinity of the kissing-loop domain, the Δ241−252 HIV-1_{Lai} kissing-loop domain had a fair chance of triggering the dimer-stabilizing conformation of the 3′ DLS and to make this contribution visible (contrary to wild-type HIV-1_{Lai} RNAs). Yet this hypothetical contribution was again invisible, as if the dimer-stabilizing conformation of the HIV-1_{Lai} 3′ DLS was weaker (i.e., only able to lift T_d to ≤48 °C) than that of the HIV-1_{Mal} 3′ DLS (able to lift T_d to 52 °C). It will be interesting to see if the T_ds of GUGCAC HIV-1_{Lai} RNAs 1−295 and 1−790 are identical or differ by 10 °C.

Conclusion. Even though the isolated 5′ and 3′ DLS can dimerize, no sequence within the 1−790 region makes a striking kinetic or stabilizing contribution to the protein-free dimerization of subtype B and D HIV-1 genomes (i.e., genomes whose kissing-loop domain has a "strong" GCGCGC262 autocomplementary sequence), other than the kissing-loop domain. It is thus an *open question* whether the 3′ DLS participates in the dimerization of all HIV-1 RNAs. The temperature of formation of tight dimers of HIV-1_{Lai} RNAs is not lowered by the presence of the 3′ DLS (Laughrea & Jetté, 1996), indicating that the 3′ DLS of HIV-1_{Lai} RNA does not stimulate the transition from loose to tight dimer under the conditions of our studies. Whether or not the 3′ DLS stabilizes dimeric HIV-1 RNAs might depend on a precise (minimal?) association constant or a precise shape/charge distribution of the dimeric kissing-loop domain. Perhaps this is why among 24 possible autocomplementary hexamers which begin and end by a G/C and which have at least a 66.6% G/C content, only 2 different ones have so far been encountered among the sequenced HIV-1 kissing-loop domains. Finally, our studies do not rule out the possibility that sequences downstream of U790 might make important contributions to the dimerization process, though preliminary studies with RNAs ending at C1403 have so far uncovered no new dimeric properties (data not shown).

Experimental Prospect for Uncovering Concealed Quaternary Contributions. Temperature gradient gel electrophoresis (TGGE) can resolve various temperature-dependent conformations of one same RNA duplex (Riesner et al., 1991). Thus, if an HIV-1 sequence makes concealed contributions to the quaternary structure of HIV-1 RNA, TGGE might uncover them. For example, if the putative quaternary contribution of the 3′ DLS "melts" at a lower temperature than that of the kissing-loop domain, one electrophoretic consequence might be the specific retardation of the partially melted dimer, i.e., a retardation not seen in a control monomeric sample of the same RNA. This retardation might be detectable by TGGE, especially if one uses very low ionic strength (but not very physiological) buffers in order to magnify the effect of a given conformational change on electrophoretic mobility (Riesner et al., 1991). One lingering uncertainty (not resolvable by TGGE) is whether dimerization induces noninterfacial conformational changes that produce new secondary or tertiary structures that melt confusingly close to the T_d of the hypothetical "weak" quaternary interactions. Nevertheless, we hope to be able to pursue some investigations along these lines in the near future.

ACKNOWLEDGMENT

We thank Dr. Léa Brakier-Gingras for critical reading of the manuscript and Mary Tomaras for the preparation of the typed text.

REFERENCES

Awang, G., & Sen, D. (1993) *Biochemistry 32*, 11453−11457.
Banerjee, A. R., Jaeger, J. A., & Turner, D. H. (1993) *Biochemistry 32*, 153−163.
Bender, W., & Davidson, N. (1976) *Cell 7*, 595−607.
Bender, W., Chien, Y. H., Chattopadkyary, S., Vogt, P. K., Gardner, M. B., & Davidson, N. (1978) *J. Virol. 25*, 888−896.
Bloomfield, V. A., Crothers, D. M., & Tinoco, I., Jr. (1974) *Physical Chemistry of Nucleic Acids*, Harper & Row, New York.
Charneau, P., Borman, A. M., Quillent, C., Guetard, D., Chamaret, S., Cohen, J., Remy, G., Montagnier, L., & Clavel, F. (1994) *Virology 205*, 247−253.
Clever, J., Sassetti, C., & Parslow, T. G. (1995) *J. Virol. 69*, 2101−2109.
Darlix, J.-L., Gabus, C., Nugeyre, M.-T., Clavel, F., Barré-Sinoussi, F. (1990) *J. Mol. Biol. 216*, 689−699.
Draper, D. E., & Gluick, T. C. (1995) *Methods Enzymol. 259*, 281−304.
Feng, Y.-X., Fu, W., Winter, A. J., Levin, J. G., & Rein, A. (1995) *J. Virol. 69*, 2486−2490.
Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Crauthers, M. H., Neilson, T., & Turner, D. H. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 9373−9377.
Fu, W., & Rein, A. (1993) *J. Virol. 67*, 5443−5449.
Fu, W., Gorelick, R. J., & Rein, A. (1994) *J. Virol. 68*, 5013−5018.
Girard, P.-M., Bonnet-Mathonière, B., Muriaux, D., & Paoletti, J. (1995) *Biochemistry 34*, 9785−9794.
He, L., Kierzek, R., Santa Lucia, J., Walter, A. E., & Turner, D. H. (1991) *Biochemistry 30*, 11124−11132.
Jaeger, J. A., Santalucia, J., Jr., & Tinoco, I., Jr. (1993) *Annu. Rev. Biochem. 62*, 255−287.
Kim, H.-J., Lee, K., & O'Rear, J. J. (1994) *Virology 198*, 336−340.
Kung, H. J., Hu, S., Bender, W., Bailey, J. M., Davidson, N., Nicolson, M. O., & McAllister, R. M. (1976) *Cell 7*, 609−620.
Laing, L. G., & Draper, D. E. (1994) *J. Mol. Biol. 237*, 560−575.
Laughrea, M., & Jetté, L. (1994) *Biochemistry 33*, 13464−13474.
Laughrea, M., & Jetté, L. (1996) *Biochemistry 35*, 1589−1598.
Laughrea, M., & Moore, P. B. (1977) *J. Mol. Biol. 112*, 399−421.
Lear, A. L., Haddrick, M., & Heaphy, S. (1995) *Virology 212*, 47−57.

Maisel, J., Bender, W., Hu, S., Duesberg, P. H., & Davidson, N. (1978) *J. Virol. 25*, 384−394.
Marquet, R., Baudin, F., Gabus, C., Darlix, J.-L., 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.
Muriaux, D., Girard, P.-M., Bonnet-Mathonière, B., & Paoletti, J. (1995) *J. Biol. Chem. 270*, 8209−8216.
Murti, K. G., Bondurant, N., & Tereba, A. (1981) *J. Virol. 37*, 411−419.
Myers, G., Korber, B., Wain-Hobson, S., Smith, R. F., & Pavlakis, G. N., Eds. (1993) *Human Retroviruses and AIDS: A compilation and Analysis of Nucleic Acid and Amino Acid Sequences*, Los Alamos National Laboratory, Los Alamos, NM.
Myers, G., Korber, B., Wain-Hobson, S., Jeang, K.-T., Henderson, L. E., & Pavlakis, G. N., Eds. (1994) *Human Retroviruses and AIDS: A compilation and Analysis of Nucleic Acid and Amino Acid Sequences*, Los Alamos National Laboratory, Los Alamos, NM.
Paillart, J. C., Marquet, R., Skripkin, E., Ehresmann, B., & Ehresmann, C. (1994) *J. Biol. Chem. 269*, 27486−27493.
Peacock, A. C., & Dingman, C. W. (1967) *Biochemistry 6*, 1818−1827.
Prats, A.-C., Roy, C., Wang, P., Erard, M., Housset, V., Gabus, C., Paoletti, C., & Darlix, J.-L. (1990) *J. Virol. 64*, 774−783.
Privalov, P. L., & Filiminov, V. V. (1978) *J. Mol. Biol. 122*, 447−464.
Puglisi, J. D., & Tinoco, I., Jr. (1989) *Methods Enzymol. 180*, 304−325.
Riesner, D., Henco, K., & Steger, G. (1991) *Adv. Electrophoresis 4*, 169−250.
Romby, P., Giegé, R., Houssier, C., & Grosjean, H. (1985) *J. Mol. Biol. 184*, 107−118.
Roy, C., Tounekti, N., Mougel, M., Darlix, J.-L., Paoletti, C., Ehresmann, C., Ehresmann, B., & Paoletti, J. (1990) *Nucleic Acids Res. 18*, 7287−7292.
Skripkin, E., Paillart, J.-C., Marquet, R., Ehresmann, B., & Ehresmann, C. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 4945−4949.
Stoltzfus, C. M., & Snyder, P. N. (1975) *J. Virol. 16*, 1161−1170.
Sundquist, W. I., & Heaphy, S. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 3393−3397.
Turner, D. H., Sugimoto, N., & Freier, S. M. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 167−192.

BI960395S