A Kissing Complex Together with a Stable Dimer Is Involved in the HIV-1_{Lai} RNA Dimerization Process *in Vitro*†

Delphine Muriaux, Philippe Fossé, and Jacques Paoletti*

Laboratoire de Pharmacologie et Physicochimie des Macromolécules Biologiques, URA 147 CNRS, Institut Gustave Roussy, rue Camille Desmoulins, 94805 Villejuif, France

Received November 29, 1995; Revised Manuscript Received February 13, 1996⊗

ABSTRACT: Retroviruses contain a dimeric RNA consisting of two identical molecules of genomic RNA. The interaction between the two monomers is thought to occur near their 5′ ends. We previously identified a region upstream from the splice donor site, comprising an autocomplementary sequence, responsible for the formation of dimeric HIV-1_{Lai} RNA [Muriaux, D., Girard, P.-M., Bonnet-Mathonière, B., & Paoletti, J. (1995) *J. Biol. Chem. 270,* 8209−8216]. This region appeared to be confined within a putative stem−loop structure. Here we report an *in vitro* model of the HIV-1 RNA dimerization process involving a two-step mechanism. We used RNA 77−402, a transcript of the HIV-1_{Lai} region, which is able to dimerize spontaneously *in vitro* under conditions of low ionic strength. Two dimers of RNA 77−402 were identified as a function of temperature, and a significant difference was found in their thermostability. Dimer D55, formed at 55 °C, is more stable than dimer D37, formed at 37 °C. RNase probing experiments confirm the involvement of a stem−loop structure in the dimerization process. In the monomer, the free G_{257}-$CGCGC_{262}$ sequence forms a loop in the 240−280 region of RNA 77−402, whereas this sequence is engaged in base pairing when D55 and D37 dimers are formed. Our results show that the loop−loop interaction of the autocomplementary $G_{257}CGCGC_{262}$ sequence, through hydrogen bonding, is responsible for the formation of dimer D37 and strongly suggest that D37 is a "kissing" complex. In contrast, in dimer D55, all the nucleotides of the two hairpin stems, 243−254/264−277, are involved in a complete interstrand interaction.

All retrovirus genomes contain a dimeric RNA consisting of two identical viral plus-strand genomic RNA molecules. The two monomers are noncovalently linked near their 5′ ends. The nature of the bonds in the dimeric structure seems weak, since the dimer is readily dissociated into monomers under mild conditions [for a review, see Coffin (1984)].

It has been proposed that retroviral RNA dimerization plays a crucial role in various phases of the viral life cycle, such as recombination (Hu & Temin, 1990), reverse transcription (Panganibam & Fiore, 1988; Temin, 1991), and translation (Bieth *et al.*, 1990). The dimeric structure is also thought to be a prerequisite for encapsidation of genomic RNA, since *cis* elements of the DLS[1] and the encapsidation site (termed E or ψ) are located within the same region in the genome of MoMuLV (Mann & Baltimore, 1985; Prats *et al.*, 1990), HaSV (Feng *et al.*, 1995), RSV (Bieth *et al.*, 1990), REV (Darlix *et al.*, 1992), BLV (Katoh *et al.*, 1993), and HIV-1 (Darlix *et al.*, 1990).

Electron microscopy studies reveal a single stable point, termed the dimer linkage structure (DLS), between the two RNA molecules where linkage occurs. This structure may involve less than 50 nucleotides and is located at approximately 300−500 nucleotides from the 5′ ends (Bender *et al.*, 1978; Murti *et al.*, 1981). Several studies indicate that short RNA molecules containing 5′ end retroviral sequences can dimerize spontaneously *in vitro* in the absence of viral proteins (Bieth *et al.*, 1990; Darlix *et al.*, 1990; Prats *et al.*, 1990; Roy *et al.*, 1990, Paoletti *et al.*, 1993). This finding by Darlix *et al.* (1990) offers a practical means of elucidating the retroviral dimerization process. This process may also be enhanced by a *trans*-acting factor, the nucleocapsid (NC) protein (Meric & Spahr, 1986; Darlix *et al.*, 1990; Prats *et al.*, 1990).

Recent studies have been conducted in an attempt to understand the HIV-1 genomic RNA dimerization process. A minimal RNA dimerization domain, downstream from the splice donor site and spanning the 5′ end of the HIV-1 *gag* gene, has been identified by deletion analysis (Darlix *et al.*, 1990). Further studies have suggested that consensus purine-rich sequences, numerous copies of which are present in most retroviral RNAs, may be involved in dimerization via the formation of purine quartets (Marquet *et al.*, 1991; Sunquist & Heaphy, 1993; Weiss *et al.*, 1993; Awang & Sun, 1993). However, Berkhout *et al.* (1993) demonstrated that HIV-2 mutant RNA which is totally lacking in purine-rich sites is fully active in dimerization *in vitro*, and more recently, we showed that RNA dimerization *in vitro* was not suppressed either with or without cDNA oligomers targeting the four purine-rich sequences in an HIV-1_{Lai} RNA fragment (Muriaux *et al.*, 1995). Data reported by Clever *et al.* (1995) also did not corroborate the premise that purine-rich tracts are required for ψ dimerization. The characterization of HIV-1

† This work was supported by the Agence Nationale de la Recherche sur le SIDA (ANRS) and a Synthelabo fellowship.

* To whom correspondence should be addressed.

⊗ Abstract published in *Advance ACS Abstracts,* March 15, 1996.

[1] Abbreviations: DLS, dimer linkage structure; PBS, primer binding site; HIV-$1_{Lai\ (or\ Mal\ or\ IIIB\ or\ HXB2)}$, human immunodeficiency virus type 1 Lai or Mal or IIIB or HXB2 isolate; MoMuLV, Moloney murine leukemia virus; HaSV, Harvey sarcoma virus; RSV, Rous sarcoma virus; ALV, avian leukemia virus; REV, reticuloendotheliosis virus; BLV, bovine leukemia virus.

0006-2960/96/0435-5075$12.00/0

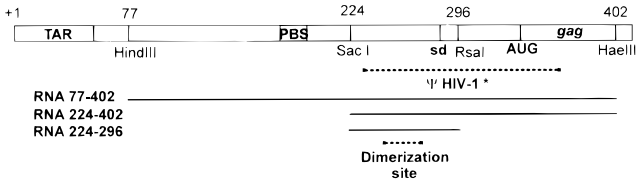

FIGURE 1: Representation of the 5′ end of HIV-1_{Lai} genomic RNA and HIV-1_{Lai} RNA transcripts used in this study. TAR, PBS, sd, and AUG indicate respectively the *trans*-activating responsive element, the primer $tRNA^{Lys3}$ binding site, the splice donor site, and the initiation codon for Pr55*gag* synthesis. ψ HIV-1* represents the 5′ packaging signal of HIV-1 recently described by Clever *et al.* (1995). Numbering is in relation to the genomic RNA cap site (+1). The restriction sites of interest are indicated: *Hind*III (+77), *Sac*I (+224), *Rsa*I (+296), and *Hae*III (+402). RNAs 77−402, 224−402, and 224−296 were generated *in vitro* by transcription of pDM2, pDM3 linearized with the appropriate enzyme. These RNA transcripts dimerize spontaneously *in vitro* in 50 mM Tris-HCl, pH 7.5, and 100 mM NaCl, at about 70%−90% (mean values from at least three experiments).

RNA dimer in viral particles shows that the HIV-1 dimeric virion RNA is not stabilized by potassium, suggesting that a guanine quartet structure is not responsible for RNA dimerization *in vivo* (Fu *et al.*, 1994). Round about the same time, it was shown that HIV-1 RNA transcripts containing a particular genomic sequence, located downstream from the primer binding site and upstream from the splice donor site, are able to dimerize *in vitro*. On the basis of an RNA−RNA recognition model [for a review, see Egushi *et al.* (1991)], it was postulated that this sequence was able to adopt a stem−loop structure and therefore could participate in the formation of a "kissing" complex leading to the formation of an HIV-1 RNA dimer (Skripkin *et al.*, 1994; Laughrea & Jetté, 1994; Paillart *et al.*, 1994; Muriaux *et al.*, 1995). However, this intermediary complex has never been isolated experimentally.

In the present report, we characterize in detail the spontaneous *in vitro* dimerization process of HIV-1_{Lai} RNA. By using temperature-dependent dimer formation, thermal stability experiments, DNA oligonucleotide interference, and enzymatic probing, we confirm that a 240−280 stem−loop structure participates in HIV-1_{Lai} RNA dimerization *in vitro*, and we propose a mechanism for dimerization with at least two essential steps, one involving the 240−280 stem−loop structure and the other, the formation of a "kissing" complex.

MATERIALS AND METHODS

Molecular Clones. Details of the plasmid pDM2 and pDM3 constructions are given elsewhere (Muriaux *et al.*, 1995).

RNA Transcripts 77−402, 224−402, and 224−296. The constructed plasmids were transcribed, after appropriate linearization, with T7 RNA polymerase under conditions stipulated by the RiboMAX large-scale RNA production system (Promega).

The pDM3 plasmid (Muriaux *et al.*, 1995) was digested by *Hae*III and transcribed by T7 RNA polymerase, giving rise to transcripts starting from position 77 of the genomic HIV-1_{Lai} RNA sequence and ending at position 402. This transcript will be referred to as RNA 77−402 (Figure 1).

The pDM2 plasmid was digested by either *Hae*III or *Rsa*I and transcribed by T7 RNA polymerase, giving rise to transcripts starting from position 224 of the genomic HIV-1_{Lai} RNA sequence and ending at positions 402 and 296, respectively. These transcripts will be referred to as RNA 224−402 and RNA 224−296 (Figure 1).

In Vitro RNA Synthesis and Purification. Exprimental procedures for HIV-1_{Lai} RNA transcript production and purification are given elsewhere (Muriaux *et al.*, 1995).

In Vitro RNA Dimerization. RNA dimerization assays were performed on 0.8 μM RNA generated *in vitro* in a buffer containing 50 mM Tris-HCl, pH 7.5, and 100 mM NaCl ("dimer" buffer), up to a final volume 10 μL, at either 37 or 55 °C. The RNA transcript can be seen in its monomeric form, at 20 °C, in a buffer containing 10 mM Tris-HCl, pH 7.5 ("monomer" buffer). RNA in water was heated for 2 min at 90 °C and chilled on ice for 2 min before addition of the buffer. Monomeric and dimeric RNA transcripts were analyzed by 1.5% agarose gel electrophoresis at 5 V/cm in a buffer containing 50 mM Tris−borate, pH 8.3, and 1 mM EDTA, at 4 °C. Ethidium bromide (0.2 μg/mL) was added to the buffer.

T_m Determination of Dimer Dissociation. After denaturation at 90 °C for 2 min in nuclease-free water, the RNA was incubated in a buffer containing 50 mM Tris-HCl, pH 7.5, and 100 mM NaCl, at a strand concentration of 0.8 μM, at either 37 or 55 °C for 90 min for optimal dimerization. The resulting dimer was then dialyzed (Millipore filters type V6, 0.025 μm) for 2 h at 4 °C against a buffer composed of 40 mM Tris-HCl, pH 7.5, 10 mM NaCl, and 1 mM EDTA. Aliquots (10 μL) of each sample were then incubated for 5 min at temperatures ranging from 4 to 90 °C and electrophoresed as described above. After fluorescent scanning of the gels by the Bioprofil apparatus (Vilber-Lourmat France), the percentage of the dimer and the monomer was estimated. The percentage of dimer was defined as the area of the dimer peak divided by the sum of the areas of monomer and dimer peaks. The melting temperature (T_m) was estimated from the plot of the amount of the dimer as a function of temperature or directly from the analysis of the electrophoresis gel.

DNA Oligomer Interference. Details of DNA oligomer 257B are described elsewhere (Muriaux *et al.*, 1995). This synthetic DNA oligomer is complementary to positions 251−267 of the HIV-1_{Lai} sequence. Numbering is relative to the genomic RNA cap site (+1). HIV-1_{Lai} RNA 77−402 was heat-denatured in the presence of the DNA oligonucleotide before the dimerization process was begun at either 37 or 55 °C *in vitro*. DNA oligonucleotide−RNA complexes were analyzed by agarose gel electrophoresis.

Enzymatic Probing. Enzymatic probing was carried out with RNase T1 (Pharmacia), specific for the cleavage of unpaired guanine residues in single-stranded RNA, and cobra venom RNase V1 (Pharmacia), specific for the cleavage of double-stranded structures or for stacked bases in single strands (Lowman & Draper, 1986). Before the addition of nucleases, RNA 77−402 (0.8 μg) in water was heated to 90 °C for 2 min, chilled on ice for 2 min, and cooled over a period of 90 min in reannealing monomer buffer at 20 °C or in reannealing dimer buffer at either 37 or 55 °C. RNA was digested with RNase T1 (0.02−0.1 unit) or RNase V1 (0.01−0.05 unit) for 10 min at 20 °C in 10 μL of monomer or dimer buffer, with 5 mM $MgCl_2$ added for RNase V1. A control without RNases was treated in parallel for each reaction. Reactions were stopped by extraction with phenol−chlorophorm twice and RNA precipitation with 0.1 volume of 3 M sodium acetate, pH 5.2, and 3 volumes of ethanol.

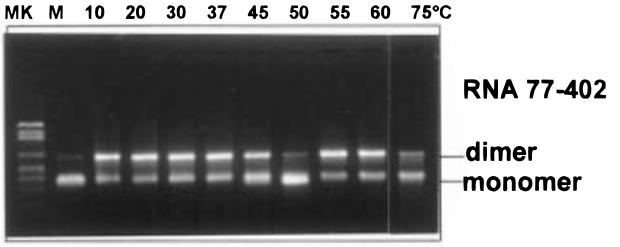

FIGURE 2: RNA 77−402 dimer formation as a function of temperature. Experimental conditions are 50 mM Tris-HCl, pH 7.5, and 100 mM NaCl. The total RNA strand concentration is 0.8 μM.

Pellets were washed with 80% ethanol and vacuum dried. The precipitated RNA was redissolved in doubly distilled water. Cuts in the RNA 77−402 were detected by extension analyses of an oligonucleotide complementary to nucleotides 312−326 of RNA 77−402. Primer extension reactions were carried out with reverse transcription as described (Ehresmann *et al.*, 1987) with minor changes. The primer was labeled at its 5′ end with [γ-^{32}P]ATP and purified by polyacrylamide gel electrophoresis. 5′-end labeled primer (200 000 cpm) was hybridized to RNA in 7.5 μL of doubly distilled water for 2 min at 90 °C. The mixture was then frozen in liquid CO_2 for 2 min before the addition of 7.5 μL of a solution containing 100 mM Tris−HCl, pH 8.5, 16 mM $MgCl_2$, 60 mM KCl, 2 mM DTT, and 5 mM each of dXTP and 2 units of avian myeloblastosis virus reverse transcriptase (Boehringer Mannheim). Reverse transcriptions were performed at 37 °C for 30 min. The cDNA fragments were then precipitated with ethanol, washed twice with 80% ethanol, dried, and dissolved in 4 μL of water and 6 μL of loading buffer (98% formamide, 10 mM EDTA, 0.025% xylene cyanol, 0.025% bromophenol blue). The samples were heated at 100 °C before being loaded onto 10% polyacrylamide−8 M urea gels. Sequencing of the unmodified RNA was done according to the method described by Sanger *et al.* (1977).

RNA Secondary Structure Analysis. The secondary structure of the HIV-1_{Lai} 240−280 region within the RNA 77−402 transcript was predicted using the MFOLD program which uses Zuker's energy minimization method (1989) and energy rules developed by Turner *et al.* (1988).

RESULTS

We have previously shown that the HIV-1_{Lai} RNA transcript spanning nucleotides 77−402 is able to dimerize spontaneously *in vitro* at 37 °C under conditions of low ionic strength (Muriaux *et al.*, 1995). Our results suggested that mutual recognition of the two identical RNA molecules occurs via an interaction through nucleotides $G_{257}CGCGC_{262}$. This sequence (dimerization site) is localized upstream from the splice donor site of the 5′ end of the HIV-1 genomic RNA (Figure 1). We conducted biochemical and structural experiments on dimer formation to characterize accurately the mechanism whereby two RNA strands are converted into a stable dimer via direct RNA−RNA interactions.

Spontaneous Dimerization of HIV-1_{Lai} RNA 77−402 as a Function of Temperature. We analyzed the dimerization process as a function of temperature. After heat denaturation, RNA 77−402 is in a monomeric form (Figure 2, lane M) when incubated in monomer buffer. Upon incubation in the dimer buffer, spontaneous dimerization is triggered and the rate of RNA 77−402 dimerization can be analyzed. For

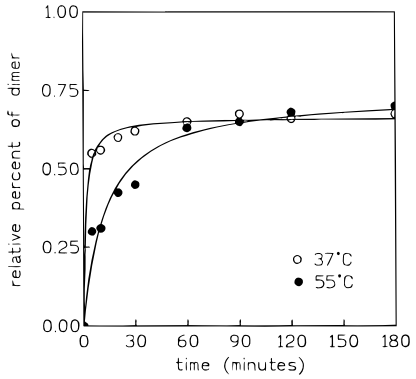

FIGURE 3: RNA 77−402 dimerization kinetics at either 37 °C (○) or 55 °C (●). Heat-denatured RNAs were incubated at either 37 or 55 °C in 50 mM Tris-HCl, pH 7.5, and 100 mM NaCl for 0−180 min. Samples were analyzed by 1.5% agarose gel electrophoresis. The percent of dimeric RNA was plotted as a function of the incubation time.

temperatures lower than or equal to 37 °C, 80% of the RNA was dimeric (Figure 2, lanes 10, 20, 30, and 37). At 45 °C, the amount of dimer decreased (Figure 2, lane 45) up to 5% at 50 °C (Figure 2, lane 50). Interestingly, for temperatures equal to or higher than 55 °C, dimer reappeared (Figure 2, lanes 55 and 60) and was totally denatured at temperatures above 75 °C (Figure 2, lane 75). These results suggest that there is a difference between two dimers in the optimum temperature required for formation and probably also in their thermal stability. One of these dimers, termed D37, is formed at a temperature of 37 °C whereas the other, D55, is formed at 55 °C. Furthermore, the melting temperature of D37 should be lower than that required for denaturation of D55.

Kinetics of the formation of both dimers are shown (Figure 3). The rate constant (k) of D37 formation was faster than that of D55 formation. At 90 min, dimerization was optimal at both temperatures for D37 and D55. It should be pointed out, however, that when D37 was formed, it could not evolve toward D55 unless the temperature was increased to 55 °C.

Thermal Stability of RNA Dimers. To check the difference in thermostability between D37 and D55, dimer was formed *in vitro* at either 37 or 55 °C in the dimer buffer for 90 min. Dimerization was complete at both temperatures. We then studied the thermal stability of both dimers under the same experimental conditions in a buffer containing 40 mM Tris-HCl, pH 7.5, 10 mM NaCl, and 1 mM EDTA. These conditions, which correspond to a 10-fold lower NaCl concentration than that of the dimer buffer, were selected to reduce the range of the dimer T_m value so that the melting of D55 could be observed. Agarose gels corresponding to the experiments are shown (Figure 4a,b) together with the denaturation curves (Figure 4c). Half of D55 was denatured at 53 °C whereas D37 was dissociated at 33 °C. However, in the case of D37, at temperatures above the T_m value, another dimer was formed (Figure 4a, lane 50), whose denaturation curve corresponded to that of D55 (Figure 4c, m). This confirmed the existence of two dimers with different T_m values.

We also examined the thermal stability of shorter RNA dimers (Figure 5). RNA 224−402 and RNA 224−296 have been described to dimerize spontaneously and completely *in vitro* in the same experimental conditions as RNA 77−402 (Muriaux *et al.*, 1995). We observed the same behavior

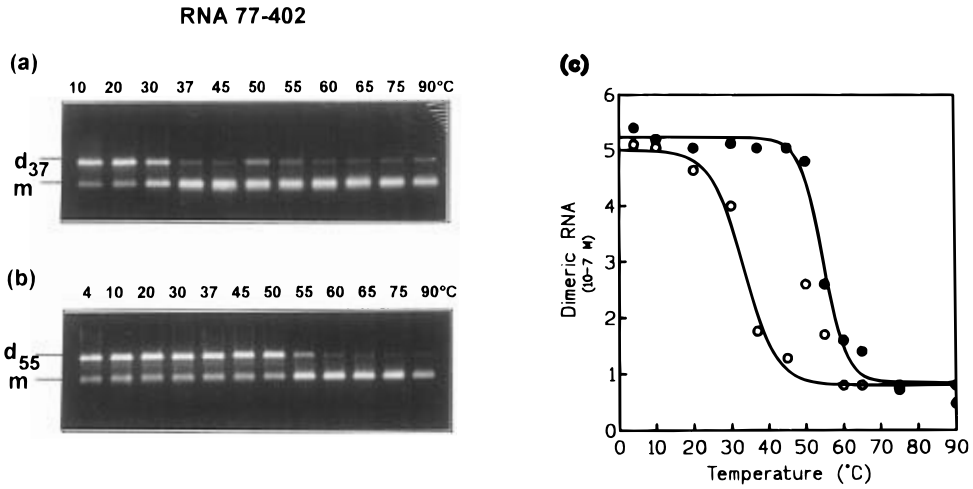

FIGURE 4: Thermal stability of the dimer RNA 77−402 formed at either 37 °C (a) or 55 °C (b) as a function of temperature. After dimerization, RNA was microdialyzed into 40 mM Tris-HCl, pH 7.5, and 10 mM NaCl buffer, and aliquots were incubated for 5 min at the indicated temperatures. Samples were analyzed by 1.5% agarose gel electrophoresis, and the percent of each species was determined from the gel. The temperatures on the plot (c) correspond to the temperatures shown above the gels (a, ○) and (b, ●). d, dimer; m, monomer.

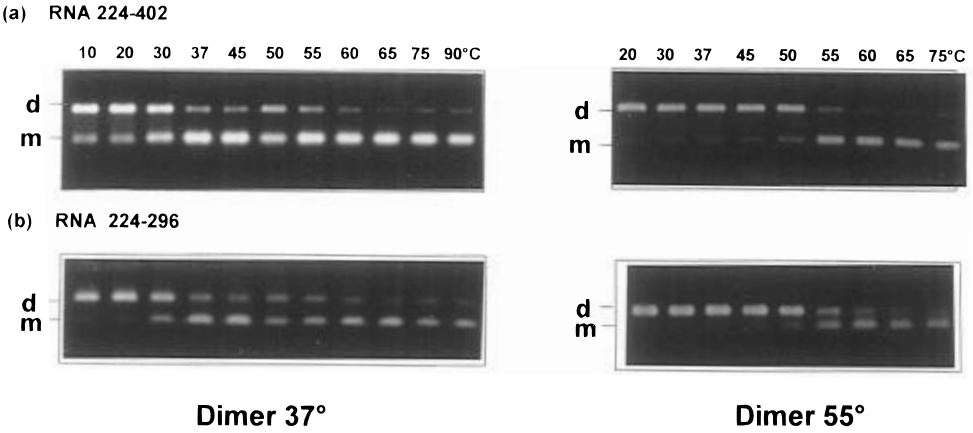

FIGURE 5: Melting profiles of RNA 224−402 (a) and RNA 224−296 (b). The effects of the dimerization temperature upon melting profiles are shown. RNA transcripts were incubated in the dimer buffer at either 37 or 55 °C and treated as described in Figure 3. d, dimer; m, monomer.

for the three RNAs. As shown in Figure 4, the melting temperature of the RNA 224−402 dimer (Figure 5a) and the RNA 224−296 dimer (Figure 5b) formed at 37 °C appeared between 30 and 37 °C, whereas the T_m values of dimers formed at 55 °C appeared between 50 and 55 °C. Consequently, the minimal region in HIV-1_{Lai} RNA fragments where possible differences may be observed between T_m values is located within nucleotides 224−296.

A Complementary DNA Oligomer Reveals That the Same Region Is Involved in the Formation of Dimer 37 and Dimer 55. An explanation for the difference observed between the T_m values of the two dimers would be to assume that the same sequence is not involved in the formation of the two dimers.

We previously defined a complementary DNA oligonucleotide, oligomer 257B, which totally inhibits the dimerization process of RNA 224−402 (Muriaux *et al.*, 1995). This oligonucleotide targets the autocomplementary sequence 257−262 containing nucleotides $G_{257}CGCGC_{262}$. We used oligomer 257B to interfere with the dimerization process of RNA 77−402 (Figure 6). Total inhibition of dimerization was observed for the RNA 77−402 dimer at an oligonucleotide 257B/RNA ratio of 1/1, whether the dimer was formed at 37 °C or at 55 °C. We checked that oligomer 257B was

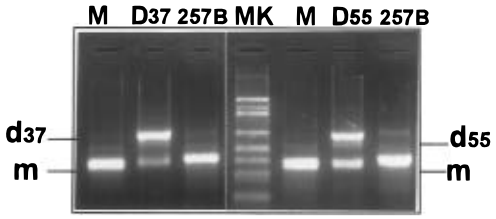

FIGURE 6: Inhibition of RNA 77−402 dimer formation at either 37 or 55 °C in the presence of the DNA oligomer 257B (ratio 1:1).

only hybridizing to RNA 77−402 in its monomeric form (data not shown) as described by Muriaux *et al.* (1995).

These experiments allowed us to conclude that the same autocomplementary sequence is involved in the dimerization mechanism of both dimers.

HIV-1_{Lai} RNA Dimer Formation Involves an Autocomplementary Sequence in a Stem−Loop Structure. The susceptibility of RNA nucleotides 240−280 to RNase T1 and RNase V1 was monitored on the RNA 77−402 transcript. RNA 77−402 reactivity to these enzymes was tested in the monomer buffer and in the dimer buffer under conditions in which most of the molecules were cleaved less than once. We checked that RNA 77−402 was in a monomeric form in 10 mM Tris-HCl, pH 7.5, at 20 °C (Figure 2, lane M)

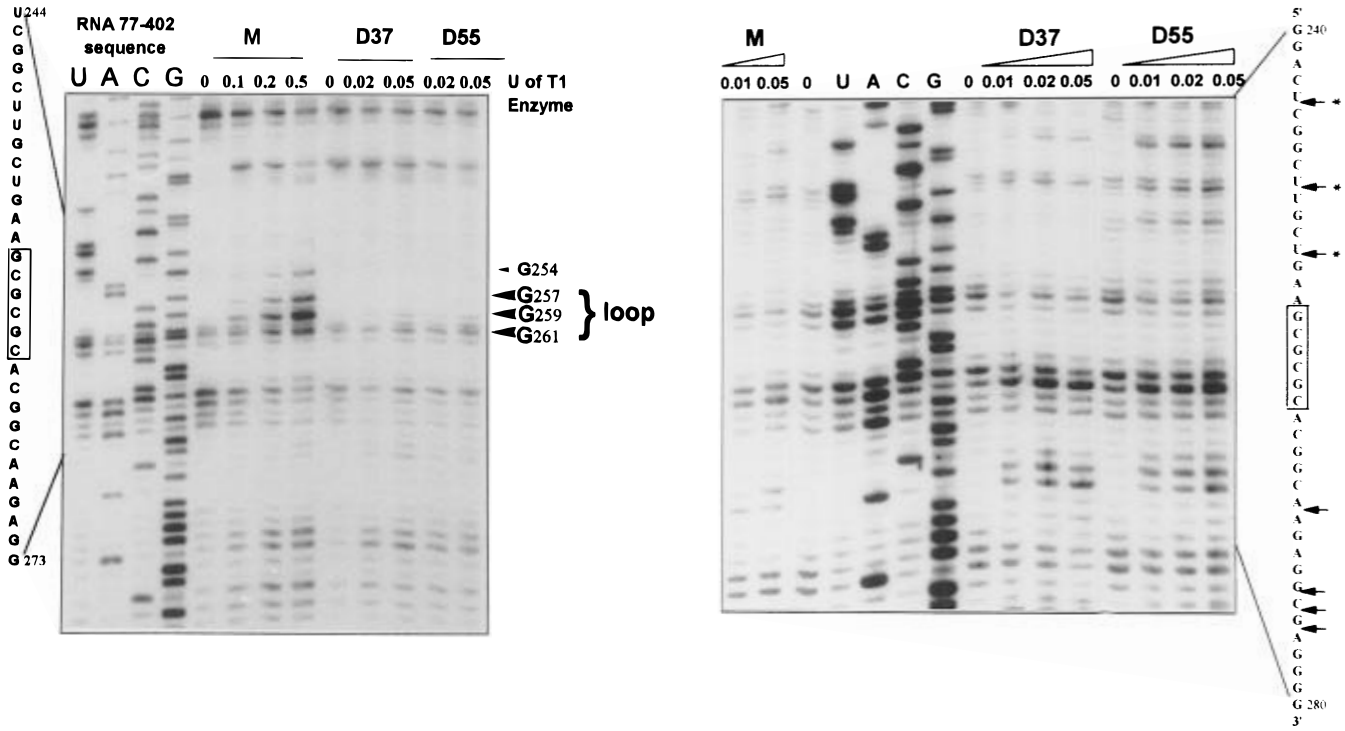

FIGURE 7: T1 (a, left) and V1 (b, right) RNase accessibility mapping of secondary structures in the HIV-1_{Lai} RNA 77−402 transcript *in vitro*. Autoradiograms of a gel showing comparison of reverse-transcribed RNA templates. In (a), untreated RNA 77−402 (lanes 0), dimeric RNA digested with 0.02 and 0.05 unit of RNase T1, and monomeric RNA digested with 0.1, 0.2, and 0.5 unit of RNase T1 are shown. Stops specific to cleaved RNA are labeled (G*xxx*). The sequence between nucleotides U244 and G273 is numbered on the left. In (b), untreated RNA 77−402 (lanes 0), dimeric RNA D37 and D55 digested with 0.01, 0.02, and 0.05 unit of RNase V1, and monomeric RNA with 0.01 and 0.05 unit of RNase V1 are shown. V1 cleavages are indicated by an arrow on the sequence. Specific cleavages only observed in D55 are noted with a star. The sequence between nucleotides G240 and G280 is indicated.

and in a dimeric form in a buffer containing 50 mM Tris-HCl, pH 7.5, and 100 mM NaCl, at either 37 or 55 °C, so that up to 80% of the dimer could be obtained reproducibly (Figure 2, lanes 37 and 55, respectively). Probing experiments are shown in Figure 7. To obtain an equivalence of reactivity to RNase T1 on the monomer and on both dimers, the enzyme was concentrated 10-fold in the monomer buffer.

RNase T1 was used to probe the accessibility of the three guanines, G257, G259, and G261, which are in the loop of the putative stem−loop structure of the monomeric RNA (Figure 8). The results in Figure 7a clearly show that the monomer cleavage pattern, generated by RNase T1, is different from cleavage patterns obtained with both dimers D37 and D55. The main differences between the monomer and both dimers were restricted to guanine residues G254, G257, G259, and G261. RNase T1 reactivities of these residues were strongly increased in the monomer (Figure 7a, lanes 0.1, 0.2, and 0.5 of M) compared with that of both dimers (Figure 7a, lanes 0.02 and 0.05 of D37 and D55). Among the different optimal and suboptimal conformations of the RNA 77−402 transcript, generated by the MFOLD program of Zucker (1989), only one RNA secondary structure was compatible with RNase T1 accessibility mapping of the 240−280 region (Figure 8, monomer hairpin). The cleavage patterns of dimers D37 and D55, generated by RNase T1, were fairly similar. The same guanine residues became inaccessible to RNase T1 cleavage (Figure 7a, lanes 0.02 and 0.05 of D37 and D55) upon dimer formation, indicating the involvement of the $G_{257}CGCGC_{262}$ sequence in the linkage of the two RNA subunits in both cases.

RNase V1 mapping was used to determine whether there was any difference in the structures of D37 and D55 and allowed us to confirm the predictive secondary structure of stem−loop 240−280 in the monomer (Figure 7b). If D37 was a kissing complex, i.e., had a secondary structure different from that of D55 (Figure 8), then we would have expected to observe a difference in the V1 cleavage pattern of the region spanning nucleotides 248−270. The comparison of the two dimer cleavage patterns clearly revealed differences. RNase V1 cuts between residues U253 and G254 in D55 but not in D37. The cutting site between U244 and C245 was increased in D55 compared to that of D37. A weak increase in reactivity to RNase V1 was also observed between U249 and U250 in D55 (Figure 7b). These differences in accessibility to RNase V1 could signify that RNA helixes 243−246/274−277 and 248−254/264−270 have a more stable structure in D55. The formation of the 257−262/262−257 helixes in this dimer must therefore be contributing to its increased stability. On the other hand, the loop−loop interaction in D37 probably destabilizes the 243−246/274−277 and 248−254/264−270 helixes. In particular, the absence of the U253/G254 cutting site in D37, which exists in D55, could be interpreted as a disruption of the G254/C264 pair of the 248−254/264−270 helical stem in D37. This disruption could mean that the GCGCGC loop−loop interaction is present. Overall, these results suggest that D55 has a more compact and stable structure than D37.

Concerning the monomer, enzymatic cuts were observed between U249 and U250, A268 and A269, C274 and G275, and G275 and A276 residues. Since RNase V1 cuts the regions which are supposed to be in a helix form, our results are in good agreement with the proposed stem−loop structure model (Figure 8, monomer hairpin).

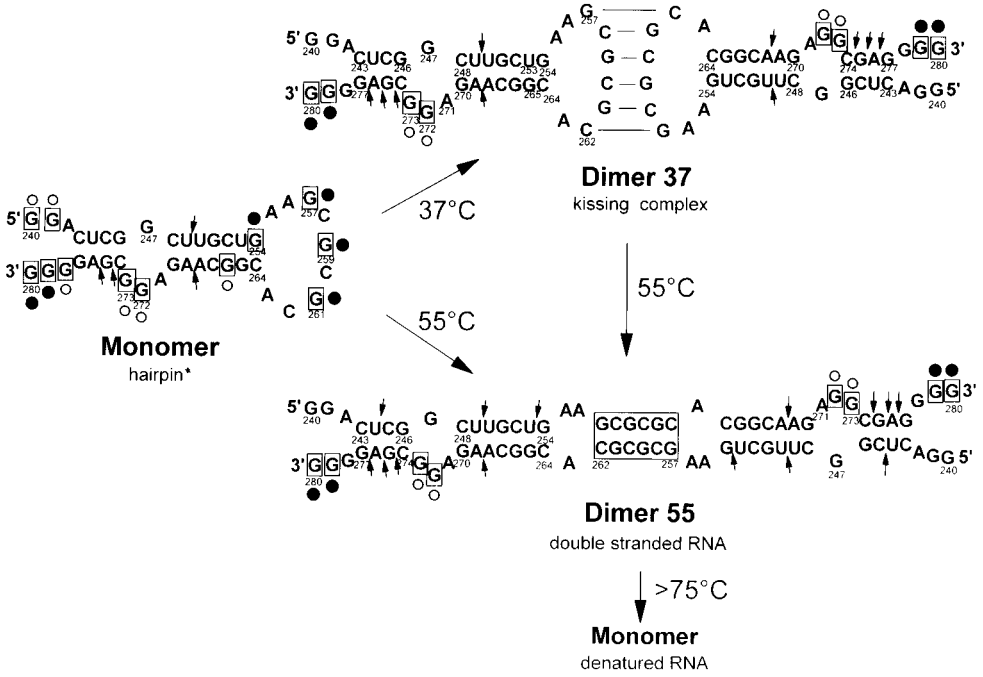

FIGURE 8: Model of the HIV-1_{Lai} RNA loop−loop interaction giving rise to a stable dimer. Sites of enzymatic cleavage are indicated. The boxed bases represent guanine residues cleaved by RNase T1 in the RNA 77−402 transcript. Full circles indicate strong or moderate reactivity to RNase T1 whereas empty ones indicate weak reactivity. The arrows represent RNase V1 cutting sites in the RNA 77−402 transcript. Monomer hairpin* represents the only conformation of the 240−280 region which agrees both with the experimental data obtained through RNase probing and with a suboptimal structure of RNA 77−402 generated with the MFOLD program.

DISCUSSION

Our results can be briefly summarized as follows. Dimerization of transcripts from the 77−402 region of the HIV-1_{Lai} genome generated two different dimers according to the temperature of the process (Figure 2). One of these dimers (D37) corresponds to a "kissing" complex.

The RNA 77−402 transcript can dimerize spontaneously *in vitro*, in conditions of low ionic strength, at 37 and 55 °C (Figure 2, lanes 37 and 55). Dimer dissociates between these temperatures, suggesting that at around 50 °C, the temperature might be energetically sufficient to disrupt dimer D37 but not efficient enough to provide conformational changes needed to allow the formation of the stable dimer D55. The two dimers therefore exhibited different T_m values (Figure 4). Furthermore, the same T_m values were also found when the shorter RNA fragments 224−402 and 224−296 were studied, suggesting that this latter sequence (or part of it) is involved in the process (Figure 5). Finally, oligomer 257B totally inhibited dimer formation of HIV-1 RNA 77−402 both at 37 °C and at 55 °C (Figure 6). On the other hand, a 40% decrease in the amount of dimer was observed when oligomer 257B was added to already formed D37 or D55 (oligo/RNA: 5/1) (data not shown). This is consistent with the implication of nucleotides 251−267 in the formation of both dimers. These nucleotides contain all the sequences and the structure required for the "kissing" reaction to occur, and this is confirmed by the differences we found between the structure of each dimer using RNase probing experiments (Figure 7).

RNase accessibility mapping, coupled with computerized sequence analysis of the HIV-1_{Lai} RNA 77−402 transcript, suggests that there is a structure in the region 240−280 comprising a stem−loop structure in the monomer (Figure 8, monomer hairpin). The folding patterns of the monomer and of both dimers are in accordance with the results of our T1 and V1 RNase mapping experiments on the 240−280 region within RNA 77−402 (Figures 7 and 8) and are in good agreement with the model of HIV-1_{Mal} dimerization proposed by Skripkin *et al.* (1994). In general, this stem−loop structure can be formed at the same location in sequences from nearly all published HIV-1 isolates (Myers *et al.*, 1992). The sequences required for the stem formation are particularly well conserved among isolates, and when variations are present in the sequence of the loop, sequence autocomplementarity is maintained (except in the HXB2 strain).

The recognition mechanism we describe is operating via a loop−loop interaction through nucleotides $G_{257}CGCGC_{262}$ situated in the loop. This transient complex, called dimer 37 (Figure 8), is formed with these complementary nucleotides and is no longer stable at 50 °C (Figure 2). Both stem−loops then open and generate a stable double-stranded dimer, via Watson−Crick base pairing: dimer 55 (Figure 8).

Such a model implies kinetic differences between dimer formation at 37 °C and at 55 °C. The formation of D37, at 37 °C, is based on a loop−loop interaction without any further conformational changes in the stem−loop. In this case, the rate constant is expected to be fast, as long as the loop is fully accessible. On the contrary, when the dimer is formed at 55 °C, complete opening of the stem−loop structure should occur. This implies a slower rate of dimer formation if this conformational change is taken in account. Our experiments confirmed this hypothesis (Figure 3). D37 formation at 37 °C was fast ($k_{37} = 6 \times 10^3$ M^{-1} s^{-1}), and the formation of a more stable dimer was no longer possible even after an extended incubation time, indicating that, at 37 °C, D55 is kinetically inaccessible from monomer or from D37. At 55 °C, the formation of D55 exhibited a kinetic profile which corresponds to a slower rate constant ($k_{55} = 9 \times 10^2$ M^{-1} s^{-1}).

This kissing mechanism is comparable to those described for the formation of a duplex between RNA I and RNA II from plasmid ColE1 [for a review, see Eguchi *et al.* (1991)] and for the antisense RNA CopA and its target RNA CopT in the replication of plasmid R1 (Persson *et al.*, 1990).

The dimerization site is located in the region of the 5′ RNA packaging signal. This suggests possible interference between packaging and dimerization. In the HIV-1_{IIIB} packaging signal region, Harrison and Lever (1992) proposed a stable secondary structure of stem−loop IId which is homologous to our 240−280 stem−loop. This structure was also found in the ψ locus of HIV-1_{Lai} (Figure 1) recently defined by Clever *et al.* (1995) as stem−loop SL1. Their results of chemical and RNase accessibility mapping of the 5′ HIV-1_{Lai} ψ region suggest that it is composed of four independent stem−loops, SL1 to SL4. These stem−loops occupy a 110-nucleotide region spanning residues 243−352 of HIV-1_{Lai} RNA. Like Clever *et al.* (1995), the fold pattern of the RNA 77−402 transcript in our study shows the same four independent stem−loops in this region (data not shown). Clever *et al.* (1995) also examined some ψ-related RNA oligonucleotides (termed B, C, K, and H) that either contained or lacked purine-rich tracts (Awang & Sen, 1993; Sundquist & Heaphy, 1993) or AT-rich sequences (Sakagushi *et al.*, 1993), in their ability to dimerize spontaneously *in vitro*. The four ψ subfragments tested formed dimers. We note that all of these ψ subfragments comprise the 240−280 stem−loop structure (or SL1), which has been defined in our model as being involved in HIV-1 dimer formation. All these results are in accordance with our findings and confirm that the dimerization domain is located in the 5′ packaging signal of HIV-1 RNA and further support the hypothetical role of dimerization in retroviral RNA packaging.

HIV-1 dimer formation appears to be an intrinsic property of viral RNA. The stable dimer dissociates into two RNA monomer molecules at about 53 °C. This T_m value is in good agreement with that found by Darlix *et al.* (1990) with full-length HIV-1_{Lai} genomic RNA isolated from wild-type virus particles. However, 55 °C is not a physiological temperature, and it has been suggested that intracellular dimerization occurs via the stimulatory effects of the viral NCp7 protein. The stem structure of the 240−280 hairpin needs to be opened up to form the stable dimer (Figure 8, dimer 55). This RNA conformational change represents a major energy barrier which must be overcome to allow the formation of this stable dimer. In our *in vitro* experiments, the temperatures served as a substitute for an effector such as a protein. The nucleocapsid protein NCp7 has been shown to exert nucleic acid annealing activities (Lapadat-Tapolsky *et al.*, 1995) and nucleic acid strand renaturation activity (Dib-Hajj *et al.*, 1993) and to be involved in the HIV-1 RNA dimerization process (Darlix *et al.*, 1990). This protein is a good candidate for this role. We therefore propose that the "kissing" complex shown on Figure 8 (dimer 37) might either be formed or be recognized by the nucleocapsid protein.

A mechanism of dimerization, based on RNA loop−loop recognition, has already been described in other retroviral systems. Girard *et al.* (1995) proposed the $U_{288}AGCUA_{293}$ sequence in a loop within the 280−330 region of the ψ locus of MoMuLV as being implicated in the dimerization process of MoMuLV RNA. This loop is well conserved among murin isolates (Tounekti *et al.*, 1992). The dimer formation of Harvey sarcoma virus RNA, another murine virus, has been studied *in vitro* (Torrent *et al.*, 1994; Feng *et al.*, 1995). A single region spanning nucleotides 205−272 was found to dimerize only if both the extreme 5′ and 3′ ends of this RNA were present. Torrent *et al.* (1994) suggested that the purine-rich sequence at the 5′ end of this sequence plays a crucial role in dimerization of this transcript. However, the possible secondary structures of the 205−272 region in HaSV RNA reveal two adjacent stem−loops, as proposed by Feng *et al.* (1995). One of these stem-loops, spanning nucleotides 205−226, has an autocomplementary sequence, 5′ GGCC 3′, in the loop. As in the dimerization process of HIV-1 RNA transcripts, a "kissing" mechanism, based on the recognition of the GGCC sequence, can be proposed for the dimerization of HaSV. Interestingly, the study by Feng *et al.* (1995) showed that the dimer of nucleotides 37−378 formed at 37 °C is less thermostable than that of the same RNA formed at 55 °C, as we observed for HIV-1 RNA transcripts (Figures 4 and 5). These authors discussed a dimer stabilization mechanism at 55 °C based on the break of intramolecular bonds present in the monomer at 37 °C (Feng *et al.*, 1995).

Regarding the dimerization of avian retroviruses, RSV (Lear *et al.*, 1995) and ALV (Fossé *et al.*, unpublished experiments) RNA transcripts reveal an autocomplementary sequence which can adopt a stem−loop structure and which is involved in the dimerization process.

Katoh *et al.* (1993) have mapped the BLV RNA for dimer formation: RNA dimerization is mediated by a region surrounding the primer binding site. The RNA secondary structure analysis of this sequence 228−366 reveals a multibranched stem−loop structure in which an autocomplementary sequence, 5′ UGAUCA 3′, is also found in loop V, located downstream from the PBS of BLV RNA.

All these works support the concept that an autocomplementary stem−loop, located downstream from the PBS, is involved in a dimerization mechanism likely shared by genomic RNA of all retroviruses.

ACKNOWLEDGMENT

We thank E. Lescot (IGR, Villejuif) for the synthesis of the 257B DNA oligonucleotide and Lorna St. Ange (IGR, Villejuif) for editing the manuscript.

REFERENCES

Awang, G., & Sen, D. (1993) *Biochemistry 32*, 11453−11457.
Bender, W., Chien, Y. H., Chattopadkyay, S., Vogt, P. K., Gardner, M. R., & Davidson, N. (1978) *J. Virol. 25*, 888−896.
Berkout, B., Oude Essink, B. B., & Schoneveld, I. (1993) *FASEB J. 7*, 181−187.
Bieth, E., Garbus, C., & Darlix, J. L. (1990) *Nucleic Acids Res. 18*, 119−127.
Clever, J., Sassetti, C., & Parslow, T. G. (1995) *J. Virol. 69*, 2101−2109.
Coffin, J. M. (1984) in *RNA Tumor Viruses* (Weiss, R., Teich, N., Varmus, H., & Coffin, G., Eds.) Vol. 1, pp 261−368, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Darlix, J. L., Gabus, C., Nugeyre, M. T., Clavel, F., & Barré-Sinoussi, F. (1990) *J. Mol. Biol. 216*, 689−699.
Darlix, J.-L., Gabus, C., & Allain, B. (1992) *J. Virol. 66*, 7245−7252.
Dib-Hajj, F., Khan, R., & Giedroc, D. P. (1993) *Protein Sci. 2*, 231−243.
Eguchi, Y., Itoh, T., & Tomizawa, J. (1991) *Annu. Rev. Biochem. 60*, 631−652.

Ehresmann, C., Baudin, F., Mougel, M., Romby, P., Ebel, J. P., & Ehresmann, B. (1987) *Nucleic Acids Res. 15*, 9109−9128.
Feng, Y., Fu, W., Winter A. J., Levin, J. G., & Rein, A. (1995) *J. Virol. 69*, 2486−2490.
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.
Harrison, G. P., & Lever, A. M. L. (1992) *J. Virol. 66*, 4144−4153.
Hu, W. S., & Temin, H. M. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 1556−1560.
Katoh, I., Yasunaga, T., & Yshinaka, Y. (1993) *J. Virol. 67*, 1830−1839.
Lapadat-Tapolsky, M., Pernelle, C., Borie, C., & Darlix, J.-L. (1995) *Nucleic Acids Res. 23*, 2434−2441.
Laughrea, M., & Jetté, L. (1994) *Biochemistry 33*, 13464−13474.
Lear, A. L., Haddrick, M., & Heaphy, S. (1995) *Virology 212*, 47−57.
Lowman, H. B., & Draper, D. E. (1986) *J. Biol. Chem. 261*, 5396−5403.
Mann, R., & Baltimore, D. (1985) *J. Virol. 54*, 401−407.
Marquet, R., Baudin, F., Gabus, C., Darlix, J. L., Mougel, M., Ehresmann, C., & Ehresmann, B. (1991) *Nucleic Acids Res. 18*, 2349−2357.
Méric, C., & Spahr, P.-F. (1986) *J. Virol. 60*, 450−459.
Muriaux, D., Girard, P.-M., Bonnet-Mathonière, B., & Paoletti, J. (1995) *J. Biol. Chem. 270*, 8209−8216.
Murti, K. G., Bondurant, M., & Tereba, A. (1981) *J. Virol. 37*, 411−419.
Myers, G., Korber, B., Berzofsky, J. A., Smith, R. F., & Pavlakis, G. N., Eds. (1992) *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.
Panganibam, A., & Fiore, D. (1988) *Science 241*, 1064−1069.
Paoletti, J., Mougel, M., Tounekti, N., Girard, P.-M., Ehresmann, C., & Ehresmann, B. (1993) *Biochimie 75*, 681−686.
Persson, C., Wagner, E. G. H., & Nordström, K. (1990) *EMBO J. 9*, 3777−3785.
Prats, A. C., Roy, C., Wang, P., Erard, M., Housset, V., Gabus, C., Paoletti, C., & Darlix, J. L. (1990) *J. Virol. 64*, 774−783.
Roy, C., Tounekti, N., Mougel, M., Darlix, J. L., Paoletti, C., Ehresmann, C., Ehresmann, B., & Paoletti, J. (1990) *Nucleic Acids Res. 18*, 7287−7292.
Sakagushi, K., Zambrano, N., Baldwin, E. T., Shapiro, B. A., Erickson, J. E., Omichinski, J. G., Clore, G. M., Gronenborn, A. M., & Appela, E. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 5219−5223.
Sanger, F. S., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463−5467.
Skripkin, E., Paillart, J. C., Marquet, R., Ehresmann, B., & Ehresmann, C. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 4945−4949.
Sundquist, W., & Heaphy, S. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 3393−3397.
Temin, H. M. (1991) *Trends Genet. 7*, 71−74.
Torrent, C., Bordet, T., & Darlix, J.-L. (1994) *J. Mol. Biol. 240*, 434−444.
Tounekti, N., Mougel, M., Roy, C., Marquet, R., Darlix, J. L., Paoletti, J., Ehresmann, B., Ehresmann, C. (1992) *J. Mol. Biol. 223*, 205−220.
Turner, D. H., Sugimoto, N., & Freier, S. M. (1988) *Annu. Rev. Biophys. Chem. 17*, 167−192.
Weiss, S., Häusl, G., Famulok, M., & König, B. (1993) *Nucleic Acids Res. 21*, 4879−4885.
Zuker, M. (1989) *Science 244*, 48−52.

BI952822S