Nucleic Acids Research, 1994, Vol. 22, No. 2 145–151

Dimerization of human immunodeficiency virus type 1 RNA involves sequences located upstream of the splice donor site

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

Unité Propre de Recherche 9002 du Centre National de la Recherche Scientifique, Institut de Biologie Moléculaire et Cellulaire, 15 rue R. Descartes, 67084 Strasbourg-cedex, France

Received November 9, 1993; Revised and Accepted December 13, 1993

ABSTRACT

The retroviral genome consists of two homologous RNA molecules associated close to their 5′ ends. We studied the spontaneous dimerization of four HIV-1 RNA fragments (RNAs 1-707, 1-615, 311-612, and 311-415) containing the previously defined dimerization domain, and a RNA fragment (RNA 1-311) corresponding to the upstream sequences. Significant dimerization of all RNAs is observed on agarose gels when magnesium is included in the electrophoresis buffer. In contrast to dimerization of RNAs 311-612 and 311-415, dimerization of RNAs 1-707, 1-615 and 1-311 strongly depends on the size of the monovalent cation present in the incubation buffer. Also, dimerization of RNAs 1-707, 1-615, and 1-311 is 10 times faster than that of RNAs 311-612 and 311-415. The dimers formed by the latter RNAs are substantially more stable than that of RNA 1-615, while RNA 1-311 dimer is 5–7°C less stable than RNA 1-615 dimer. These results indicate that dimerization of HIV-1 genomic RNA involves elements located upstream of the splice donor site (position 305), i.e. outside of the previously defined dimerization domain.

INTRODUCTION

The genome of the human immunodeficiency virus (HIV-1), like the genome of all known retroviruses, is formed by two homologous RNA molecules that are physically linked within the viral particles (1–7). Using genomic RNAs extracted from a variety of retroviruses, it has been shown by electron microscopy that the dimer linkage structure (DLS) is located in a region close to their 5′ end. Surprisingly, the two molecules usually appear in an apparently parallel configuration (2–5).

Dimerization of the genomic RNA is probably involved in several key steps of the retroviral cycle. First, dimerization may down regulate translation of the *gag* gene: this has been suggested for Rous sarcoma virus (RSV) (8), and HIV-1 (9). Second, it has been proposed that the DLS would constitute a positive signal for the encapsidation machinery (7, 10). Indeed, the regions involved in the dimerization of HIV-1, Moloney murine leukaemia virus (MoMuLV), and RSV genomic RNAs are also required for their encapsidation (7, 8, 11–18). The *trans*-acting nucleocapsid protein (NC), that activates retroviral RNA dimerization (7, 8, 13, 18–22) is also involved in RNA packaging (6, 10, 14, 23). Furthermore, the high level of genetic recombination observed during the synthesis of proviral DNA can be accounted by the use of a dimeric template (24–27).

Despite the fact that dimerization of the genomic RNA should constitute a suitable target against retroviruses, little is known about the mechanism of RNA dimerization. It was recently reported that synthetic RNAs corresponding to the 5′ region of RSV (8), MoMuLV (13, 19, 28, 29), HIV-1 (7, 9, 15, 17, 18, 30–32), and HIV-2 (33) genomes are able to dimerize *in vitro*. In the case of MoMuLV, the thermal stability of the RNA dimer obtained *in vitro* in the absence of NC, and of the RNA dimer extracted from viral particles are identical (28). Furthermore, chemical probing data indicated that the conformation of the dimeric MoMuLV RNA is essentially the same *in vitro* (29), and inside the viral particles (34). Moreover, a RNA stem-loop proposed as the HIV-1 packaging signal has the same structure in an *in vitro* synthesized RNA transcript (9, 35) and in the genomic RNA extracted from infected cells (36). These observations strongly support that results obtained *in vitro* reflect the biological situation. In the case of HIV-1, RNA fragments encompassing a short sequence (100 nts or less) located immediately downstream of the splice donor (SD) site (nts 305–306) are able to dimerize *in vitro* (7, 15, 17, 18, 31). However, there is no general agreement concerning the mechanism of RNA dimerization. In a first study, we suggested that the same process governs the dimerization of genomic RNA of most retroviruses, and that this process implies non-canonical interactions, probably purine quartets (15). Given the cation dependence of the HIV-1 RNA dimerization yield, we proposed that those quartets would involve guanines as well as adenines. More recently, it has been suggested that HIV-1 RNA dimerization involves purine quartets containing solely guanines

(17, 31, 32). However, no purine quartets downstream of the SD site seemed to be involved in the dimerization process of HIV-2 and BLV (33, 37).

As some of the RNAs used in these studies contained the sequences located upstream of the SD site, while others did not, we extensively compared the dimerization of both types of HIV-1 RNAs. Indeed, we found that sequences located upstream of the SD site are involved in the dimerization process. Dimerization of RNA fragments lacking the sequences upstream of the SD site differs from that of RNAs that contain these sequences in its cation dependence. In addition, the kinetics of dimerization and the thermal stability of the dimers formed by these two classes of RNAs are different. We conclude that RNAs lacking the sequences upstream of the SD site are not appropriate model systems to study dimerization of the HIV-1 genomic RNA.

MATERIALS AND METHODS

Plasmid construction and enzymatic restriction for *in vitro* transcription

Plasmid construction and restriction enzyme digestions were conducted according to standard procedures (38). Details of the construction of plasmids pHIVCG-4 and pHIVCG-61 are given elsewhere (15). Plasmid pJCB was obtained by inverse PCR on the plasmid pHIVCG-4. The PCR primers were designed to eliminate the 615−4005 HIV-1 region of pHIVCG-4 and allow a self ligation of the PCR products after restriction by *Sal* I. The resulting pJCB plasmid, confirmed by DNA sequencing, also contains a new *Pvu* II restriction site at the 3′ end of the HIV-1 sequence.

HIV-1 RNAs 1-311 and 1-707. In order to obtain HIV-1 RNAs 1-311 and 1-707, pHIVCG-4 was digested with *Rsa* I and *Pvu* II, respectively, and *in vitro* transcribed with T7 RNA polymerase. The resulting RNAs correspond to positions 1-311 and 1-707 of the Mal isolate of HIV-1 (+1 is the first nucleotide of the genomic RNA) (39), respectively, and contain no additional nucleotides derived from the plasmid.

HIV-1 RNA 1-615. This RNA was obtained from *in vitro* transcription of pJCB, after restriction of the plasmid with *Pvu* II, and contains 8 additional nts (CCGGGCAG) at its 3′ end.

HIV-1 RNAs 311-415 and 311-612. These RNAs were obtained from *in vitro* transcription of pHIVCG-61, after digestion with *Hae* III (RNA 311-415) or *Sma* I (RNA 311-612). Since pHIVCG-61 has been obtained by cloning the *Rsa* I fragment of HIV-1 Mal in the polylinker of pBS, the resulting RNAs contain small additional sequences that have been reduced at a minimum by the choice of the cloning sites. Both RNAs contain the sequence GGGCG upstream of the HIV-1 sequence, and RNA 311-612 has the sequence GAGCUCGGUACCC in 3′ of the HIV-1 sequence.

RNA synthesis and purification

T7 RNA polymerase was purified from the overproducing strain BL21/pAR1219, kindly supplied to us by F.W. Studier (Upton, NY, USA). *In vitro* transcription using this polymerase was for four hours, under previously described conditions (15). Transcription was followed by treatment for 30 min with RNAse free DNase I (Appligene) (2 U/μg plasmid), phenol-extraction and ethanol precipitation. RNAs were dissolved in water and purified using a FPLC (Pharmacia) system with either a Bio-Sil TSK 250 or Bio-Sil TSK 400 (Biorad) column, in a buffer containing 200 mM sodium acetate pH 6.5, 1% methanol. Transcripts were concentrated in a Centricon 10 (Amicon) unit, precipitated with ethanol and dissolved in water prior use. RNAs were checked for integrity and purity on 5 or 10% denaturing polyacrylamide gels. Labelled RNAs were synthesized using [α^{32}P] ATP (Amersham) (50 μCi/μg of template). Depending on their length, radioactive transcripts were purified on 6 to 10% denaturing polyacrylamide gels.

In vitro dimerization of HIV-1 RNAs

For studying the cation size dependence of dimerization, unlabelled RNA (0.8 μg), together with the corresponding radioactive RNA (3 to 5 nCi, 0.01 to 0.04 μg), were diluted in 8 μl of Millli-Q (Millipore) water, heated for 2 min at 90°C, and chilled for 2 min on ice. Two μl of 5-fold concentrated appropriate buffer were added (final concentration: 50 mM sodium cacodylate pH 7.5, and 300 mM of either Li^+, Na^+, K^+, or Cs^+, or the same buffer supplemented with 5 mM $MgCl_2$), the samples were incubated for 30 min at the appropriate temperature, then cooled for 2 min on ice.

For the kinetic experiments, incubation was at 37°C in 50 mM pH 7.5, 300 mM KCl and 5 mM $MgCl_2$, and the RNA concentration was 390 nM, unless otherwise stated. Dimerization was initiated by the addition of the buffer. For the RNA dimer stability experiments, the samples were incubated for 30 min at 30°C in the same buffer (unless otherwise stated), then the temperature was gradually increased by 7°C steps. After a 5 min incubation at the appropriate temperature, an aliquot was loaded on gel after addition of glycerol (20% final concentration).

Samples were analyzed either on 1.1% agarose gels, or on 5% polyacrylamide (bisacrylamide/acrylamide = 0.8/30) gels. Gels were run in 45 mM tris−borate buffer for the cation dependence experiments, and in the same buffer supplemented with 0.1 mM $MgCl_2$ for the kinetic and thermal stability experiments. Agarose gels were fixed for 10 min in 10% trichloroacetic acid and dried for 40 min under vacuum at room temperature, while polyacrylamide gels were fixed in 10% acetic acid, 6% ethanol and dried at 80°C. Fuji X-ray films were exposed without intensifying screen to obtained a low level of blackening, and the autoradiographs were scanned. The area of the peaks corresponding to the monomeric and dimeric forms of HIV-1 RNAs was quantified, and the percentage of dimer was defined as the area of the dimer peak divided by the sum of the areas of the monomer and dimer peaks. Alternatively, quantitative data were obtained using a BAS 2000 BIO-Imager (Fuji).

Kinetic analysis

The experimental data of the kinetic experiments were analyzed according to two simple models that could account for incomplete dimerization.

(i) The 'conformation model'. In this model we supposed that due to conformational heterogeneity of the monomer, a part of it (M_d) is able to dimerize, while the other part (M_i) is not:

$$M_d + M_d \xrightarrow{k_{dim}} D \quad \text{(A1)}$$

$$M_i + M_i \xrightarrow{k=0} D \quad \text{(A2)}$$

with

$$[M_0] = [M_d^0] + [M_i^0] \quad \text{(A3)}$$

$$[M_i] = [M_0] - 2\cdot[D] \quad \text{(A4)}$$

$$[M_d^0] = 2\cdot[D]_{t\to\infty} \quad \text{(A5)}$$

At any time, the concentration of M_d is given by:

$$\frac{1}{2}\cdot\frac{d[M_d]}{dt} = -k_{dim}\cdot[M_d]^2 \quad \text{(A6)}$$

$$\frac{1}{[M_d]} = \frac{1}{[M_d^0]} + 2\cdot k_{dim}\cdot t \quad \text{(A7)}$$

(ii) The equilibrium model. In this model, we supposed that dimerization and dissociation occur in the same time scale and that an equilibrium is reached between these two processes. If f_D is the dimer fraction at any time, and f_D^{eq} the dimer fraction at equilibrium, we have:

$$M + M \xrightarrow{k_{dim}} D \quad \text{(B1)}$$

and

$$D \xrightarrow{k_{diss}} M + M \quad \text{(B2)}$$

since

$$2\cdot[D] = [M_0]\cdot f_D \quad \text{(B3)}$$

$$\frac{d[D]}{dt} = \frac{1}{2}\cdot[M_0]\cdot\frac{df_D}{dt} \quad \text{(B4)}$$

$$\frac{d[D]}{dt} = k_{dim}\cdot[M_0 - f_D\cdot M_0]^2 - \frac{1}{2}k_{diss}\cdot f_D\cdot[M_0] \quad \text{(B5)}$$

and at equilibrium :

$$\frac{d[f_D]}{dt} = 0 \quad \text{and} \quad k_{diss} = 2\cdot k_{dim}\cdot[M_0]\cdot\frac{(1-f_D^{eq})^2}{f_D^{eq}} \equiv 2\cdot k_{dim}\cdot[M_0]\cdot w \quad \text{(B6)}$$

At any time the dimer fraction is given by:

$$\frac{d[f_D]}{dt} = 2\cdot k_{dim}\cdot[M_0]\cdot\left\{f_D^2 - f_D\cdot\left(2+\frac{w}{2}\right)+1\right\} \quad \text{(B7)}$$

let be

$$q = w\cdot\left(2+\frac{w}{4}\right) \quad \text{(B8)}$$

integration yields:

$$\frac{1}{\sqrt{q}}\cdot\ln\frac{\left\{2\cdot f_D - \left(2+\frac{w}{2}\right) - \sqrt{q}\right\}\cdot\left\{-\left(2+\frac{w}{2}\right)+\sqrt{q}\right\}}{\left\{2\cdot f_D - \left(2+\frac{w}{2}\right)+\sqrt{q}\right\}\cdot\left\{-\left(2+\frac{w}{2}\right)-\sqrt{q}\right\}} = 2\cdot k_{dim}\cdot[M_0]\cdot t \quad \text{(B9)}$$

Note that this model gives a pseudo first order kinetic equation that is very different from the typical second order kinetic of the 'conformation model'. Fitting of the experimental kinetic data with equations A7 and B9 should allow us to understand the origin of the incomplete HIV-1 RNA dimerization.

RESULTS

By using *in vitro* transcription with T7 RNA polymerase, we synthesized HIV-1 RNAs differing by their length. HIV-1 RNAs 1-707 and 1-615 span the complete R and U5 sequences, the primer binding site (nts 179−196), the leader region containing the SD site (nts 305−306), the initiation site of the translation of the *gag* gene (nts 350−352) and the beginning of its coding sequences. HIV-1 RNAs 311-612 and 311-415 start just downstream of the SD site. These RNAs all contain the previously proposed dimerization region (7, 15, 30), and the postulated packaging signal, which has been proposed to correspond to nts 313−356 of HIV-1 Mal (11, 12, 14, 16). On the contrary, RNA 1-311 encompasses the sequences located upstream of the previously defined dimerization region. Dimerization of all RNAs can be analyzed on 1.1% agarose gels. Alternatively, the monomer and dimer species of RNA 311-415 can be resolved on 5% polyacrylamide gels, with a lower background. All experiments were reproduced from 2 to 5 times with essentially the same results.

Magnesium and cation size dependence of HIV-1 RNA dimerization

We previously reported that RNAs 1-707, 311-612, and 311-415 undergo a salt induced dimerization, while RNA 1-311 does not (15). To ensure that the RNA dimers did not dissociate during electrophoresis, we checked the effect of including 0.1 mM $MgCl_2$ in the agarose gels and electrophoresis buffer. The stabilizing effect of $MgCl_2$ was clearly evidenced by measuring the dimerization yield of RNA 1-615 after a 30 min incubation at 37°C in 50 mM sodium cacodylate pH 7.5, 300 mM KCl and 5 mM $MgCl_2$. The measured dimerization yield of RNA 1-615 was strongly enhanced when the electrophoresis buffer contained 0.1 mM $MgCl_2$ (Fig. 1, lanes 4 and 10). Unexpectedly, we could evidence a significant dimerization of RNA 1-311 under these electrophoresis conditions (Fig. 1, lane 8), while no dimer of this RNA was observed in the absence of Mg^{2+} in the gel

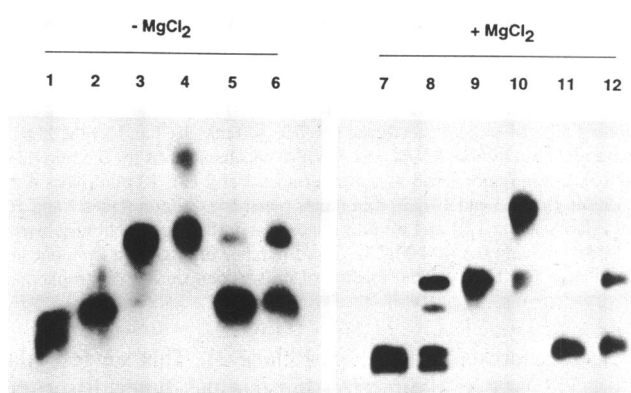

Figure 1. Influence of the magnesium in the electrophoresis buffer on the dimerization yield of RNAs 1-311, 1-615, and 311-612. RNAs 1-311 (lanes 1, 2, 7, 8), 1-615 (lanes 3, 4, 9, 10), and 311-612 (lanes 5, 6, 11, 12) were incubated under monomer conditions (50 mM sodium cacodylate pH 7.5, 40 mM KCl and 0.1 mM $MgCl_2$) (lanes 1, 3, 5, 7, 9, 11) or under dimer conditions (50 mM sodium cacodylate pH 7.5, 300 mM KCl and 5 mM $MgCl_2$) (lanes 2, 4, 6, 8, 10, 12). Electrophoresis was run in 45 mM tris−borate (lanes 1−6) or in 45 mM tris−borate, 0.1 mM $MgCl_2$ (lanes 7−12). Note that RNA 1-311 monomer displayed two distinct bands. The shorter band results from the cleavage of the 125 5′ nts of RNA 1-311 during incubation (Skripkin, unpublished results).

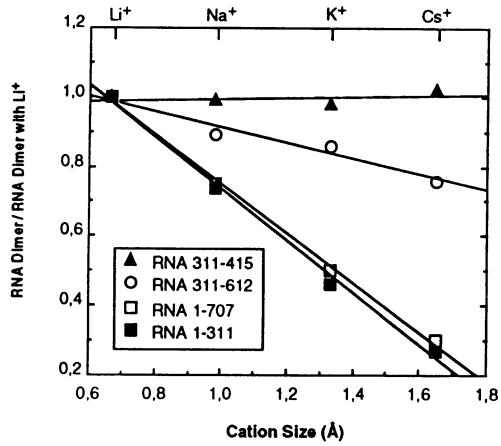

Figure 2. Influence of the cation size on the dimerization process. RNA 1-707, 1-311, 311-612, or 311-415 (see insert) were incubated 30 min at 37°C, in buffers containing 50 mM sodium cacodylate pH 7.5, and either 300 mM of either Li^+, Na^+, K^+, or Cs^+. Samples were loaded on 1.1% agarose or 5% polyacrylamide gels and electrophoresis was conducted in a 45 mM tris−borate (RNAs 1-707, 311-612, and 311-415), or in 45 mM tris−borate, 0.1 mM $MgCl_2$ (RNA 1-311, see text). Each dimerization yield was divided by the dimerization yield obtained when the same RNA was incubated with Li^+, and the resulting ratios were plotted as a function of the ionic radius of the cation contained in the buffer.

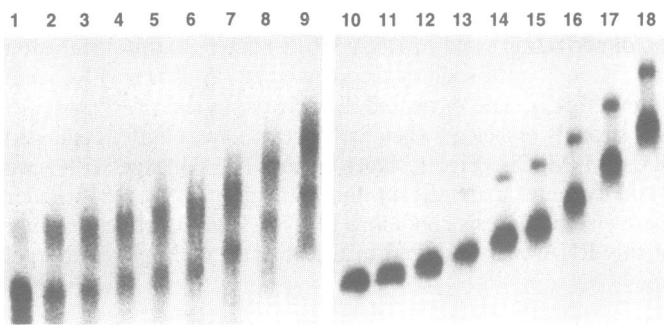

Figure 3. Kinetics of dimerization of RNAs 1-615 and 311-612. RNA 1-615 (lanes 1−9) and RNA 311-612 (lanes 10−18) were incubated at 37°C in a buffer containing 50 mM sodium cacodylate pH 7.5, 300 mM KCl and 5 mM $MgCl_2$. The concentration of both RNAs was 390 nM. Incubation was for 0.5 min (lanes 1 and 10), 2 min (lanes 2 and 11), 5 min (lanes 3 and 12), 10 min (lanes 4 and 13), 15 min (lanes 5 and 14), 20 min (lanes 6 and 15), 30 min (lanes 7 and 16), 45 min (lanes 8 and 17), and 60 min (lanes 9 and 18). Electrophoresis was in 45 mM tris−borate, 0.1 mM $MgCl_2$. The difference of migration from one lane to the other is due to sequential loading of the samples during electrophoresis.

and electrophoresis buffer (Fig. 1, lane 2). This suggests that RNA 1-311 is able to dimerize, but that the dimer dissociates during electrophoresis in the absence of Mg^{2+}. On the contrary, the dimerization yield of RNA 311-612 was not markedly affected by the addition of Mg^{2+} in the electrophoresis buffer (Fig. 1, lanes 6 and 12).

We previously noticed that dimerization of HIV-1 RNA 1-707 strongly depends on the size of the monovalent cation present in the dimerization buffer (15). Here, we compared the dimerization efficiency of RNAs 1-707, 311-612, 311-415, and 1-311 in buffers containing either Li^+, Na^+, K^+, or Cs^+ (Fig. 2). In order to facilitate comparison of the results, the

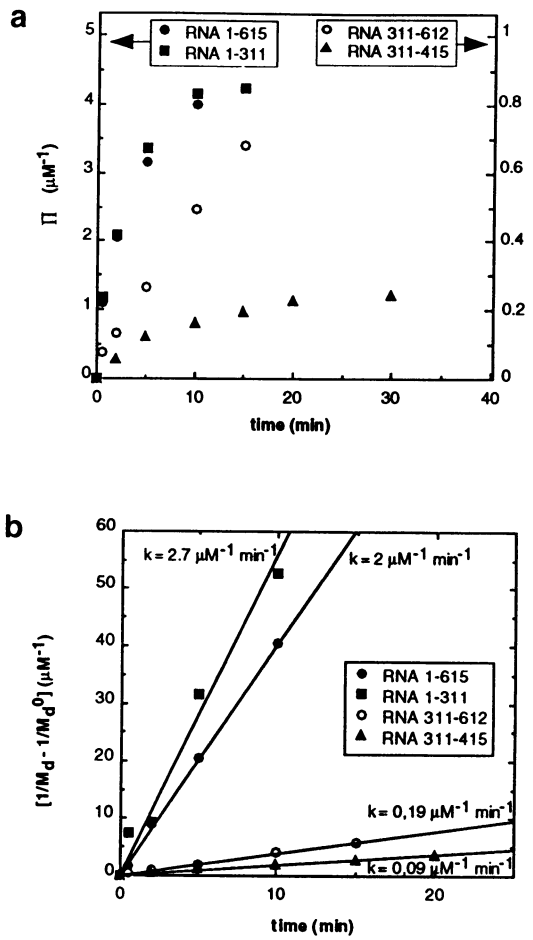

Figure 4. Analysis of the dimerization kinetics. The experimental dimerization data of RNAs 1-615, 1-311, 311-612, and 311-415 were introduced into equation B9 **(a)**, corresponding to the 'equilibrium model', or into equation A7 **(b)**, corresponding to the 'conformation model'. In figure 4a,

$$\Pi = \frac{1}{[M_0]} \cdot \frac{1}{\sqrt{q}} \cdot \ln \frac{\left\{2 \cdot f_D - \left(2 + \frac{w}{2}\right) - \sqrt{q}\right\} \cdot \left\{-\left(2 + \frac{w}{2}\right) + \sqrt{q}\right\}}{\left\{2 \cdot f_D - \left(2 + \frac{w}{2}\right) + \sqrt{q}\right\} \cdot \left\{-\left(2 + \frac{w}{2}\right) - \sqrt{q}\right\}}$$

When Π is plotted as a function of t, a linear plot with a slope of 2 k_{dim} is expected if the equilibrium model applies. Note that two y scales are used for RNAs 1-615 and 1-311, and RNAs 311-612 and 311-415, respectively, to better evidence the negative curvature obtained with RNA 311-415. The RNA concentration was 390 nM (RNAs 1-615, 1-311, and 311-612) or 2.2 μM (RNA 311-415). We determined from the experimental data that f_D^{eq} (in equation B9) and $[M_d^o]/[M_o]$ (in equation A7) are equal to 0.62, 0.62, 0.45, and 0.40 for RNAs 1-615, 1-311, 311-612, and 311-415, respectively.

dimerization yields are normalized to those obtained with Li^+, and the results were plotted as a function of the ionic radius of the cation. To allow a direct comparison with our previous work (15), the gels and electrophoresis buffer used in this experiment did not contain magnesium, except when we studied RNA 1-311. The dimerization yields of RNAs 1-707, 311-612, and 311-415 depends almost linearly on the cation size, the dimerization process being favoured by the smaller cations. However, the important point is that the slope of the lines strongly differ from

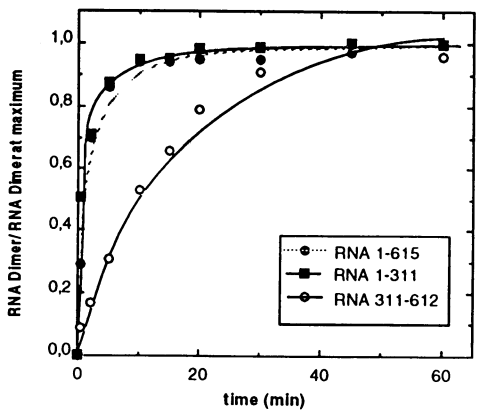

Figure 5. Fitting of dimerization data of RNAs 1-615, 1-311 and 311-612. Experimental data (circles and squares) obtained from quantification of gels such as those of figure 2 were fitted with the conformation model, using the rate constant determined from figure 4b. Data are normalized to the dimer yield at the maximum of the curve, in order to better visualize the differences in the kinetics of dimerization of RNAs 1-615 and 311-612.

one RNA to the other. While Cs^+ is 3.3 times less efficient than Li^+ in the dimerization of RNA 1-707, it is only 1.3 less efficient than Li^+ in the dimerization of RNA 311-612, and it is as efficient as Li^+ in the dimerization of RNA 311-415. RNA 1-615 gave results similar to RNA 1-707 (data not shown). Dimerization of RNA 1-311 displayed a cation size dependence similar to that of RNA 1-707, even though 0.1 mM $MgCl_2$ was included in the electrophoresis buffer for the analysis of the shorter RNA (Fig. 2).

Using RNA 1-707 and 1-615, we reproduced the same cation dependence experiments, except that 0.1 mM $MgCl_2$ was included in the gel and in the electrophoresis buffer (data not shown). In these conditions, Li^+ and Na^+ were also much more efficient than K^+ and Cs^+ to promote HIV-1 RNA dimerization, but the difference of efficiency between Li^+ and Na^+ disappeared. However, when 0.1 mM $MgCl_2$ was present in both the incubation and electrophoresis buffers, the dimerization yields of RNAs 1-707 and 1-615 became almost independent of the nature of the monovalent cation present in the incubation buffer.

Kinetics of HIV-1 RNA dimerization

We compared the dimerization kinetics of RNAs 1-707, 1-615, 1-311, 311-612 and 311-415 at 37°C. Large variations of the dimerization rates of these RNAs can be observed (Fig. 3). We analyzed the kinetic data using two simple models that could account for incomplete dimerization. In the first one (conformation model), we assume that due to conformational heterogeneity of the monomer, only part of the monomer is able to dimerize (see materials and methods). In the second model (equilibrium model), we assume that the dimerization is reversible and that an equilibrium exists between dimerization and dissociation. Introducing the experimental dimerization data into equation B9 does not yield a linear relationship (Fig. 4a), except for RNA 311-612. The strong deviation from linearity observed for RNAs 1-311, 1-615, and 311-415 indicates that dimerization of these RNAs does not follow the equilibrium model. On the contrary, introducing the experimental data into equation A7 yields a linear relationship for all RNAs (Fig. 4b), indicating that under our experimental conditions only a sub-set of the RNA molecules adopt a conformation that allow dimerization. The dimerization constants of RNAs 1-615 and 1-311 are ten times higher than that of RNA 311-612, while the dimerization constants of RNAs 311-612 and 311-415 only differ by a factor of two (Fig. 4b). The dimerization constants of RNAs 1-615 and 1-311 are almost indistinguishable, in the limit of the experimental errors. The dimerization constant of RNA 311-415 was found to be only slightly dependent of the cation present in the dimerization buffer: the dimerization constant was found ~ 35% greater in presence of K^+ than in presence of Li^+ (data not shown).

Taking the dimerization rate constants determined from figure 4b, one can re-calculate the theoretical curve from equation A7 and compare it to the complete set of experimental data (including the last points of the kinetics that were not included for the determination of the rate constants). Indeed, the fit is quite good (Fig. 5). This further supports the use of the 'conformation model' to analyze our kinetic experiments. However, we are aware that the 'conformation model' is an oversimplification that cannot explain all our results. For example, it does not easily explain why the dimerization yield depends on RNA and salt concentrations in the incubation buffer (15). Indeed, a more complex scheme taking into account the kinetics of renaturation of the initially denaturated monomer could explain these data. However, the rate constants obtained with the more simple 'conformation model' allow a semi-quantitative comparison of the dimerization kinetics of the different RNAs.

Thermal stability of the HIV-1 RNA dimer

In order to further characterize the RNA dimers formed by the various HIV-1 RNAs, we compared their thermal stabilities. Melting experiments of RNAs 1-707, 1-615, 1-311 and 311-612 were conducted at the same RNA molar concentration (390 nM) and are thus directly comparable (Fig. 6). The thermal stabilities of RNA dimers 1-707 and 1-615 are almost indistinguishable: denaturation starts above 30°C and is almost complete at 60°C, half of the initial dimer being melted at ~47°C. RNA 1-311 appears slightly less stable: half of the RNA 1-311 dimer is denaturated at ~ 40°C. On the contrary, the dimerization yield of RNA 311-612 increases between 30 and 44°C. This is consistent with the fact that maximal dimerization is obtained at higher temperature (by about 10°C) for RNAs 311-415 and 311-612 than for RNAs 1-707 and 1-615 (data not shown). The dimer of RNA 311-612 starts to dissociate above 44°C and half of the dimer is melted at ~ 61°C. Since the dimerization yield of RNA 311-415 is quite low at a RNA concentration of 390 nM, the stability of this dimer has been measured at higher concentration (2.2 μM). In these conditions, no denaturation of the RNA 311-415 dimer was observed up to 80°C (data not shown).

In order to test whether the cation dependence of HIV-1 RNA 1-707 dimerization depicted in figure 1 is linked to the stability of the dimer or not, we investigated the thermal stability of the RNA 1-707 dimer in buffers containing 300 mM of either Li^+, Na^+, K^+, or Cs^+, or 300 mM K^+ supplemented with 5 mM $MgCl_2$ (Fig. 7). The dimer has almost the same stability in 300 mM Li^+ and in 300 mM K^+ + 5 mM $MgCl_2$, but was 5–7°C less stable in buffers containing 300 mM Na^+ or 300 mM K^+. The dimer was even less stable in Cs^+ buffer, but the melting temperature was difficult to measure precisely due to low dimerization yields (data not shown).

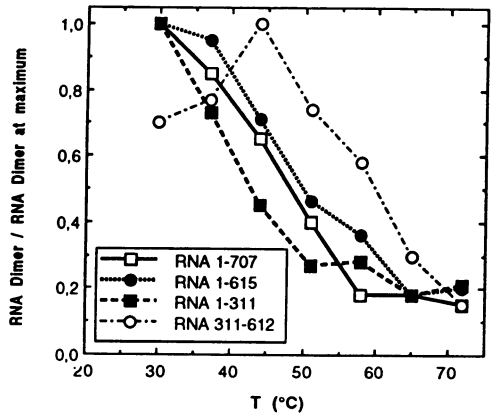

Figure 6. Thermal stabilities of dimers of RNAs 1-707, 1-615, 1-311 and 311-612. To facilitate comparison, the dimerization yields were divided by the dimerization at the maximum of the melting curves. All RNAs were at a concentration of 390 nM.

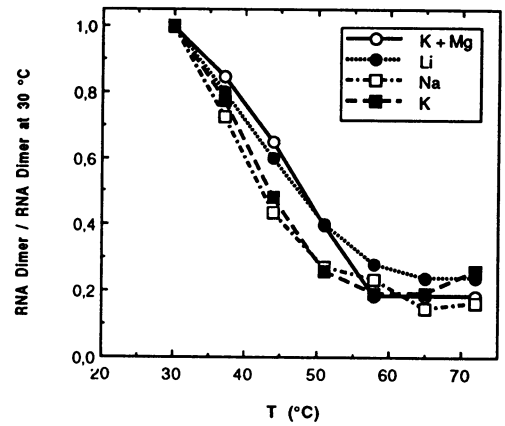

Figure 7. Influence of the buffer composition on the thermal stability of RNA 1-707 dimer. Melting curves were obtained as described in materials and methods in buffer containing 50 mM of cacodylate pH 7.5 and 300 mM KCl + 5 mM $MgCl_2$, or 300 mM LiCl, or 300 mM NaCl, or 300 mM KCl.

DISCUSSION

The dimerization region of HIV-1 RNA was originally located in the 100 nts immediately downstream of the SD site (position 305 in the Mal isolate) (7, 15, 30). This localization was further supported by the demonstration that RNAs corresponding to nts 294–418 (17), 305–415 (31) or 298–339 (18) are able to dimerize (all numberings are according to the HIV-1 Mal RNA). In initial studies, no dimerization of RNA 1-311 was reported (7, 15). However, here we showed that high levels of RNA 1-311 dimerization can be detected, provided that 0.1 mM $MgCl_2$ were included in the electrophoresis buffer. Therefore, the localization and size of the dimerization region of HIV-1 RNA must be re-examined.

In all retroviruses studied by electron microscopy, the DLS has been localized within the first 600 nts of the genome (2–4), so that one can almost be sure that RNAs 1-707 and 1-615 contain the authentic dimerization site of the genomic HIV-1 RNA. Comparison of the dimerization process of RNAs 1-311, 311-612, 311-415, 1-615, and 1-707 indicates that dimerization of RNA 1-311 closely resembles that of RNAs 1-707 and 1-615, while dimerization of RNAs 311-612 and 311-415 is clearly different.

First, in the absence of Mg^{2+}, dimerization of RNAs 1-707, 1-605, and 1-311 strongly depends on the ionic radius of the cation present in the incubation medium, while dimerization of RNAs 311-612 and 311-415 does not. The cation size dependence of RNA 1-707 dimerization was one of the arguments that led us to propose that purine quartets containing both adenines and guanines may be involved in the dimerization process (15, 30). However, our present results indicate that this cation size dependence is linked to sequences located upstream of the SD site. Since the purine tracts that could be involved in the formation of G and A quartets are located downstream of the SD site (15, 30), the origin of the strong influence of the cation size on the dimerization of RNAs 1-707 and 1-615 remains unknown. However, it is not linked to the formation of G and A quartets downstream of the SD site, as we previously postulated (15, 30).

Second, RNA 311-612 and 311-415 dimers also differentiate from the dimers formed by RNAs 1-707 and 1-615 by their thermal stability, since they are at least 10–15°C more stable. This result can hardly be considered simply as an effect of the length of the RNA on the dimer stability: in the case of MoMuLV, two RNAs of 300 and 700 nucleotides in length that contain the dimerization region form dimers of identical thermal stability (28). Therefore, one must consider that at least some of the interactions stabilizing RNA 311-612 and 311-415 dimers are different from those stabilizing the RNA 1-707 and 1-615 dimers. The dimers formed by RNAs 294–418 (17) and 305–415 (31) are also more stable than the RNA 1-707 dimer. By contrast, RNA 1-311 forms a dimer that is about 7°C less stable than those formed by RNAs 1-707 and 1-615. This result suggests that RNA 1-707 and 1-615 dimers involve the same interactions as RNA 1-311, but that additional interactions take place downstream of the SD site.

The Li^+ stimulation of RNAs 1-707 and 1-615 dimerization is directly linked to the thermal stability of the corresponding dimers. Indeed, we found that RNA 1-707 dimer is more stable in presence of Li^+ than in presence of Na^+ or K^+. Furthermore, similar yields of RNA 1-707 dimers are obtained in buffers containing either 300 mM Li^+, or 300 mM K^+ + 5 mM $MgCl_2$, and the thermal stability of this dimer is identical in the two buffers. The similar effect of Li^+ and Mg^{2+} on the RNA dimer stability may be related to the fact that the ionic radius of these cations are identical. Understanding the exact role of these cations in maintaining the DLS will probably be possible when the nucleotides involved in the DLS will be precisely defined. The enhanced stability of the RNA 1-707 dimer in presence of Li^+ contrasts with the strong stabilizing effect of K^+ evidenced on RNA 294–418 (17) and 305–415 (31) dimers. This result strongly suggests that the dimers formed by HIV-1 RNA 1-707 and HIV-1 RNA 294–418 or 305–415 are different.

The third criterion that distinguishes RNAs 311-612 and 311-415 from RNAs 1-707, 1-615, and 1-311 is the kinetic of dimerization. The dimerization kinetics are ten times faster when the HIV-1 RNAs contain the sequences upstream of the SD site. Since the structure of region 1-305 is almost the same in RNAs 1-311 and 1-707 (9), and region 312-415 adopts essentially the same conformation in RNAs 1-615 and 311-415 (E.Skripkin, unpublished results), the differences in the dimerization kinetics among these RNAs cannot be attributed to differences in the

conformation of the different RNAs. The most likely explanation is that sequences located in the 1-311 region hasten the dimerization process.

The main conclusion of this study is that sequences located upstream of the SD site are involved in the dimerization of HIV-1 RNAs 1-707 and 1-615, and most probably of HIV-1 genomic RNA. Sequences located upstream of the SD site that could be involved in the dimerization of HIV-1 RNA have not been identified by previous studies (7, 15, 17, 18, 30, 31). However, a common dimerization initiation site of RNAs 1-1333, 1-707, and 1-311 has been recently evidenced in our laboratory by damage selection experiments (Skripkin *et al.*, to be published). Also, the involvement of sequences located upstream of the SD site has been demonstrated in the dimerization of RSV (8), HIV-2 (33) and Bovine Leukaemia Virus (37) RNAs. Dimerization of short RNA fragments limited to the encapsidation signal domain, and missing these sequences, clearly differs from the dimerization of RNAs 1-707 and 1-615. Thus, these short RNAs are not adequate models to study the dimerization of genomic RNA.

The involvement of sequences located upstream of the SD site in the dimerization process somewhat questions the proposed link between dimerization and encapsidation. Indeed, the proposed HIV-1 packaging signal is located downstream of the SD site (16), and the sequences located upstream of the SD site are present in both sub-genomic and genomic RNAs. However, comparison of the thermal stability of RNA 1-615 dimer and RNA 1-311 dimer indicates that some interactions also take place downstream of the SD site. These additional interactions may involve the conserved purine tracts downstream of the SD site (15, 30), and may be crucial for the packaging and reverse transcription processes. In this context, it may be important to note that HIV-1 packaging signal has been recently shown to extend in the *env* gene (40). Also, our kinetic analysis suggests that depending on their conformation, some monomers are able to dimerize, while other are not. These conformers may also behave differently in the translation process, the later being translated more efficiently than the former.

ACKNOWLEDGEMENTS

Claude Houssier (Université de Liège, Belgium) is greatly acknowledged for fruitful discussions. This work was supported by the Centre National de la Recherche Scientifique (CNRS) and by grants from the 'Agence Nationale de Recherches sur le SIDA' (ANRS) and the 'Fondation de la Recherche Médicale' (FRM).

REFERENCES

1. Canaani, E., Helm, K.V.D. & Duesberg, P. (1973) *Proc. Natl. Acad. Sci. USA*, **72**, 401–405.
2. Bender, W. & Davidson, N. (1976) *Cell*, **7**, 595–607.
3. Kung, H.J., Hu, S., Bender, W., Bailey, J.M., Davidson, N., Nicolson, M.O. & McAllister, R.M. (1976) *Cell*, **7**, 609–620.
4. Bender, W., Chien, Y.H., Chattopadkyay, S., Vogt, P.K., Gardner, M.B. & Davidson, N. (1978) *J. Virol.*, **25**, 888–896.
5. Murti, K.G., Bondurant, M. & Tereba, A. (1981) *J. Virol.*, **37**, 411–419.
6. Méric, C. & Spahr, P.F. (1986) *J. Virol.*, **60**, 450–459.
7. Darlix, J.L., Gabus, C., Nugeyre, M.T., Clavel, F. & Barré-Sinoussi, F. (1990) *J. Mol. Biol.*, **216**, 689–699.
8. Bieth, E., Gabus, C. & Darlix, J.L. (1990) *Nucleic Acids Res.*, **18**, 119–127.
9. Baudin, F., Marquet, R., Isel, C., Darlix, J.L., Ehresmann, B. & Ehresmann, C. (1993) *J. Mol. Biol.*, **229**, 382–397.
10. Gorelick, R.J., Nigada, S.M., Arthur, L.O., Henderson, L.E. & Rein, A. (1991) *Advances in molecular biology and targeted treatment of AIDS*, Kumar A., ed., Plenum Press, N.Y., pp 257–272.
11. Lever, A., Gottlinger, H., Haseltine, W. & Sodroski, J. (1989) *J. Virol.*, **63**, 4085–4087.
12. Clavel, F. & Orenstein, J.M. (1990) *J. Virol.*, **64**, 5230–5234.
13. Prats, A.C., Roy, C., Wang, P., Erard, M., Housset, V., Gabus, C., Paoletti, C. & Darlix, J.L. (1990) *J. Virol.*, **64**, 774–783.
14. Aldovini, A. & Young, R.A. (1990) *J. Virol.*, **64**, 1920–1926.
15. Marquet, R., Baudin, F., Gabus, C., Darlix, J.L., Mougel, M., Ehresmann, C. & Ehresmann, B. (1991) *Nucleic Acids Res.*, **19**, 2349–2357.
16. Hayashi, T., Shioda, T., Iwakura, Y. & Shibuta, H. (1992) *Virology*, **188**, 590–599.
17. Sundquist, W.I. & Heaphy, S. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 3393–3397.
18. Sakaguchi, K., Zambrano, N., Baldwin, E.T., Shapiro, B.A., Erickson, J.W., Omichinski, J.G., Clore, G.M., Gronenborn, A.M. & Appella, E. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 5219- 5223.
19. Prats, A.C., Sarih, L., Gabus, C., Litvak, S., Keith, G. & Darlix, J.L. (1988) *EMBO J.*, **7**, 1777–1783.
20. Prats, A.C., Housset, V., de Billy, G., Cornille, F., Prats, H., Roques, B. & Darlix, J.L. (1991) *Nucleic Acids Res.*, **19**, 3533–3541.
21. De Rocquigny, H., Gabus, C., Vincent, A., Fournié-Zaluski, M.C., Roques, B. & Darlix, J.L. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 6472–6476.
22. Weiss, S., Konig, B., Morikawa, Y. & Jones, I. (1992) *Gene*, **121**, 203–212.
23. Méric, C. & Goff, S. (1989) *J. Virol.*, **63**, 1558–1568.
24. Hu, W.S. & Temin, H.M. (1990) *Science*, **250**, 1227–1233.
25. Luo, G. & Taylor, J. (1990) *J. Virol.*, **64**, 4321–4328.
26. Panganiban, A. & Fiore, D. (1988) *Science*, **241**, 1064–1069.
27. Temin, H.M. (1991) *Trends in Genetics*, **7**, 71–74.
28. Roy, C., Tounekti, N., Mougel, M., Darlix, J.L., Paoletti, C., Ehresmann, C., Ehresmann, B. & Paoletti, J. (1990) *Nucleic Acids Res.*, **18**, 7287–7292.
29. Tounekti, N., Mougel, M., Roy, C., Marquet, R., Darlix, J.L., Paoletti, J., Ehresmann, B. & Ehresmann, C. (1992) *J. Mol. Biol.*, **223**, 205–220.
30. Ehresmann, B., Marquet, R., Mougel, M. & Ehresmann, C. (1992) *Bull. Inst. Pasteur*, **90**, 109–124.
31. Awang, G. & Sen, D. (1993) *Biochemistry*, **32**, 11453–11457.
32. Weiss, S., Häusl, G., Famulok, M. & König, B. (1993) *Nucleic Acids Res.*, **21**, 4879–4885.
33. Berkhout, B., Oude Essink, B.B. & Schoneveld, I. (1993) *FASEB J.*, **7**, 181–187.
34. Alford, R.L., Honda, S., Lawrence, C.B. & Belmont, J.W. (1991) *Virology*, **183**, 611–619.
35. Harrison, G.P. & Lever, A.M.L. (1992) *J Virol*, **66**, 4144–4153.
36. Hayashi, T., Y., U. & Okamoto, T. (1993) *FEBS Lett.*, **327**, 213–218.
37. Katoh, I., Yasunaga, T. & Yoshinaka, Y. (1993) *J. Virol.*, **67**, 1830–1839.
38. Maniatis, T., Fritsch, E.F. & Sambrook, J. (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
39. Alizon, M., Wain-Hobson, S., Montagnier, L. & Sonigo, P. (1986) *Cell*, **46**, 63–74.
40. Richardson, J.H., Child, L.A. & Lever, A.M.L. (1993) *J. Virol.*, **67**, 3997–4005.