

A new NMR solution structure of the SL1 HIV-1_{Lai} loop–loop dimer

Fabien Kieken, Françoise Paquet, Fabienne Brulé, Jacques Paoletti
and Gérard Lancelot*

Centre de Biophysique Moléculaire (Affiliated to the Orléans University), Rue Charles Sadron,
45071 Orléans Cedex2, France

Received as resubmission November 25, 2005; Revised and Accepted December 15, 2005

ABSTRACT

Dimerization of genomic RNA is directly related with the event of encapsidation and maturation of the virion. The initiating sequence of the dimerization is a short autocomplementary region in the hairpin loop SL1. We describe here a new solution structure of the RNA dimerization initiation site (DIS) of HIV-1_{Lai}. NMR pulsed field-gradient spin-echo techniques and multidimensional heteronuclear NMR spectroscopy indicate that this structure is formed by two hairpins linked by six Watson–Crick GC base pairs. Hinges between the stems and the loops are stabilized by intra and intermolecular interactions involving the A8, A9 and A16 adenines. The coaxial alignment of the three A-type helices present in the structure is supported by previous crystallography analysis but the A8 and A9 adenines are found in a bulged in position. These data suggest the existence of an equilibrium between bulged in and bulged out conformations in solution.

INTRODUCTION

Retroviruses are unique in having a diploid genome (1). This diploid feature is directly involved in different stages of the virion life cycle such as recombination and encapsidation (2–7). Electron microscopy of HIV-1 has shown the RNA genome to consist of two identical genomic RNA strands joined within their 5' leader, at a 100 bases long region defined as the dimer linkage structure (DLS) (8–11). In the HIV-1_{Lai} strain, the dimerization initiation site (DIS) is a 6 nt autocomplementary sequence located in the highly conserved 35 bases long sequence stem–loop (SL1) (Figure 1) (12,13). Mutations in the DIS region or deletion of the entire autocomplementary sequence have been shown to affect the yield of dimer

formation consequently reducing virion infectivity (14–22). Further more this deficit in virulence is closely implicated with defects in encapsidation, indicating a direct involvement between dimerization and formation of the capsid.

In vitro studies with 200 nt RNA sequences containing the DIS region show that HIV-1 dimer formation is initiated via a loop–loop interaction between two RNA molecules (23,24). This metastable complex is converted at 55°C, or at lower temperatures in the presence of the NCp7 protein, into a more stable conformation, an extended dimer (23,25–27). Shorter sequences comprising the entire 35 nt SL1 sequence (28) or its extremity of 23 nt (29,30) remain able to dimerize into two different conformations of different stabilities. The structure of the loop–loop dimer or 'kissing complex' has been resolved by NMR (31) and crystallography (32). Both studies utilize different RNA sequences. The NMR sequence presents a modification of the base pair order in the stem contrary to the crystallography sequence. Although both studies report identical nucleotide sequences of the loop, discrepancies appear when both proposed structures are compared. The NMR solution structure shows a symmetrical homodimer where the average plane of the loops is roughly perpendicular to the axis of the stems whereas in the crystallographic structure it shows a perfect coaxial alignment with the stems. Unlike the NMR structure where all the bases are inside the RNA molecule, the crystallographic structure presents the A8 and A9 adenines as stacked in a bulged out conformation. The NMR structure shows that the last base pair of the stem (G7-C17) is not paired contrary to the model proposed by crystallography where bases appear paired. Previous NMR findings from our laboratory have confirmed the presence of base pair (G7-C17) in the loop–loop complex in the 23 nt sequence in the wild-type. To clarify these contradictory reports, we solved the 23 nt structure of the SL1 region in the HIV-1_{Lai} wild-type without base pair rearrangement in the stem by multidimensional NMR spectroscopy. We found a new solution structure of the 'kissing complex'. Its structure is composed by two hairpins, coaxially aligned, linked by

*To whom correspondence should be addressed. Tel: +33 2 38 25 55 71; Email: lancelot@cns-orleans.fr

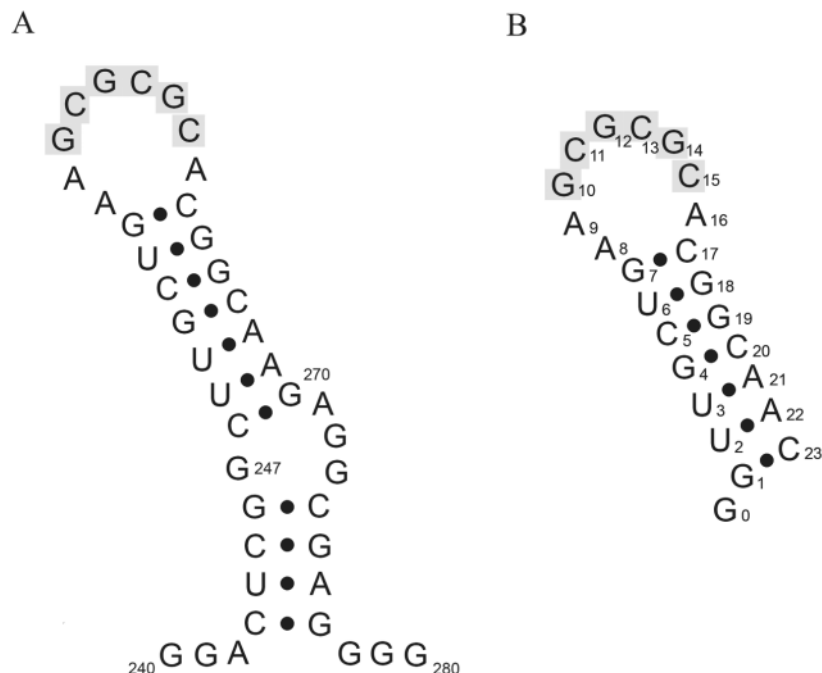


Figure 1. (A) The SL1 sequence of the genomic RNA HIV-1_{Lai}. (B) The SL1 sequence synthesized for NMR studies. Bases in SL1 are numbered with respect to the cap site in the HIV-1 genomic transcript. The 6 nt autocomplementary sequence is highlighted in grey. In the numbering of the Mal variant genome, the adenines 256, 257 and 264 are numbered 272, 273 and 280, respectively.

their autocomplementary sequence with the loop adenines appearing in a bulged in position.

MATERIALS AND METHODS

RNA synthesis and folding

RNA strands were synthesized by *in vitro* transcription with T7 RNA polymerase using synthetic templates as described previously (33). Template oligonucleotides (5'-GTTGCCGT-GCGCGCTTCAGCAACCTATAGTGAGTCGTATTA and 5'-TAATACACT CACTATAG) were purchased from Eurogentec (Belgium). T7 RNA polymerase was isolated from *Escherichia coli* strain BL21/PAR1219 as described in Davanloo *et al.* (34). RNAs were synthesized with unlabelled NTPs (Sigma) or labelled ¹⁵N/¹³C NTPs (Silantes). Following phenol extraction and ethanol precipitation, all RNA samples were purified by PAGE under denaturing conditions with 8 M Urea. The SL1 24 nt RNA (Figure 1) was recovered by diffusion in water. RNA was desalted by reverse phase chromatography on Sep-Pak C18 (Waters) and by gel filtration on DG-10 column (Bio-Rad). A total of 10 μM desalted sample was refolded by heating at 90°C (3 min) and snap-frozen at -20°C. The sample was concentrated by lyophilization. The final concentrations of the unlabelled and labelled SL1 24 nt were 2 mM and 0.6 mM, respectively. Each sample was placed in a Shigemi tube in 170 μl of 90% water, 10% deuterium oxide at pH 5.7, without addition of any salt for the experiments involving exchangeable protons. For the non-exchangeable protons experiments, samples were lyophilized and solved again in 100% deuterium oxide. The NMR experiment was generally run below 12°C to avoid the transition toward the extended structure.

NMR spectroscopy

NOESY, DQF-COSY, TOCSY, ¹⁵N and ¹³C HSQC, HNN-COSY, HMBC, NMR pulsed field-gradient spin-echo experiments (35) were run on a Varian 600 MHz UNITY INOVA spectrometer. ¹⁵N and ¹³C double-half-filter NOESY experiments were run on a Varian 600 MHz and on an 800 MHz UNITY INOVA spectrometer.

Pf1 phage was purchased from ASLA Ltd (Riga, Latvia). Phage concentrations were adjusted by weighing appropriate amounts of a 50 mg/ml stock solution in H₂O. Phage alignment was verified by measuring the splitting of the HOD signal in 90% H₂O/10% D₂O. Resonance assignments are reported relative to DSS.

Structure refinement

The structure of the loop-loop complex was calculated using torsion angle molecular dynamics (TAMD) as described by Stallings and Moore (36). First, one hundred random structures were generated using CNS 1.0 (37). The basic started structures were subjected to a simulated annealing-restrained molecular dynamics protocol in torsion angle space. TAMD consisted of a high-temperature torsion angle dynamics for a total of 15 ps at 20 000 K with a Van der Waals scale factor of 0.1. The molecules were then cooled to 300 K in 4000 steps for 60 ps of torsion angle dynamics with a Van der Waals scale factor increased from 0.1 to 1. This was followed by another simulated annealing step in Cartesian space from 2000 K to 300 K in 3000 steps for 15 ps where the Van der Waals scale factor increased from 1 to 4, finalized by 2000 steps of Powell minimization with full electrostatic and Van der Waals terms. The structures with zero violation on NOE distance (0.2 Å), dihedral (5°) and with the lowest energies were selected.

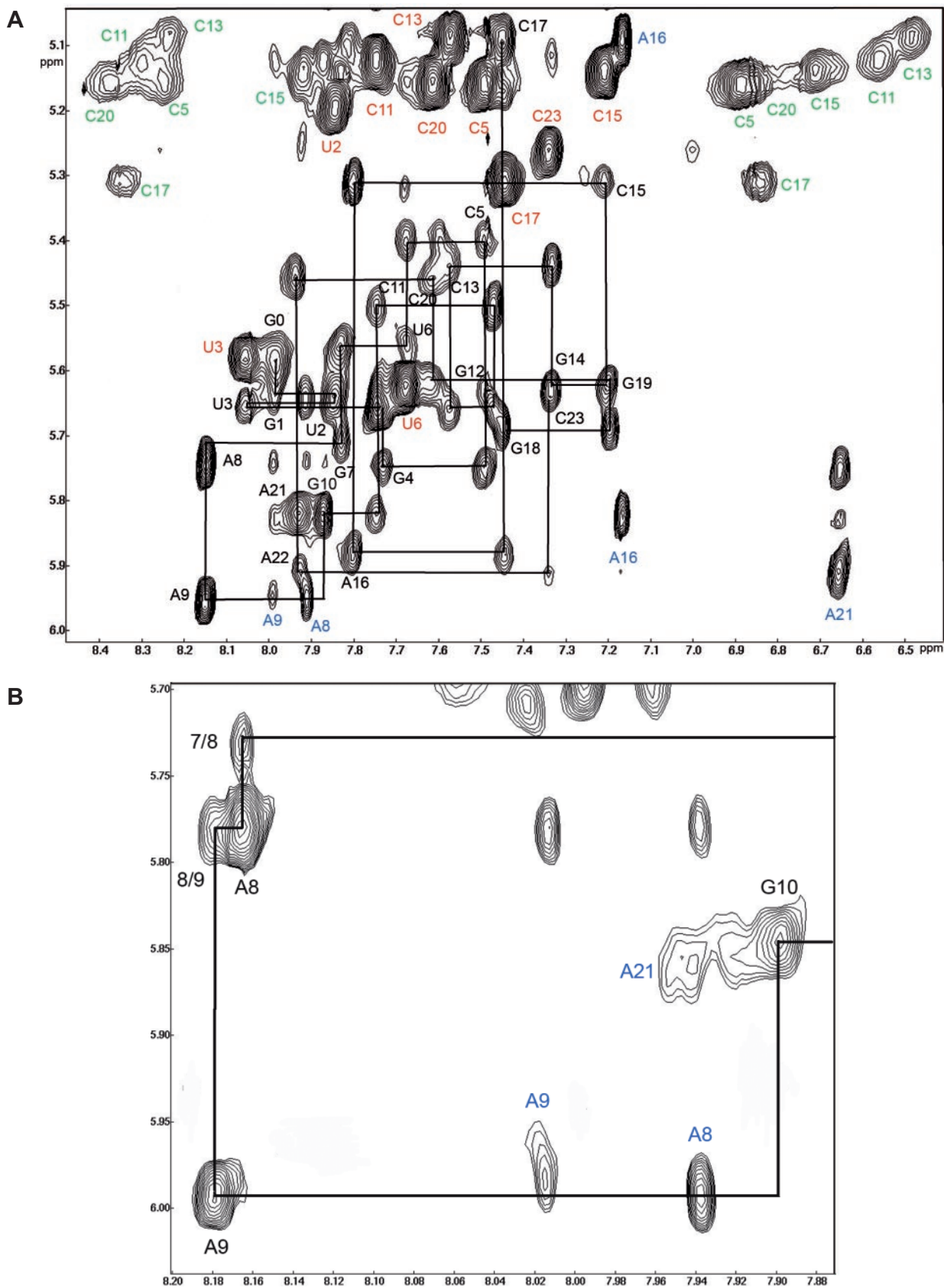


Figure 2. (A) Sequential assignment of the NOESY spectra run in H₂O at 10°C with 80 ms mixing time on the loose complex formed by SL1. The sequential assignment of the H8, H6 and H1' protons as well as H5, H2 and the NH2 of cytosines is shown. (B) Expansion of a region shows clearly the connectivity H8(A9)-H1'(A8) at 20°C with 300 ms mixing time in D₂O.

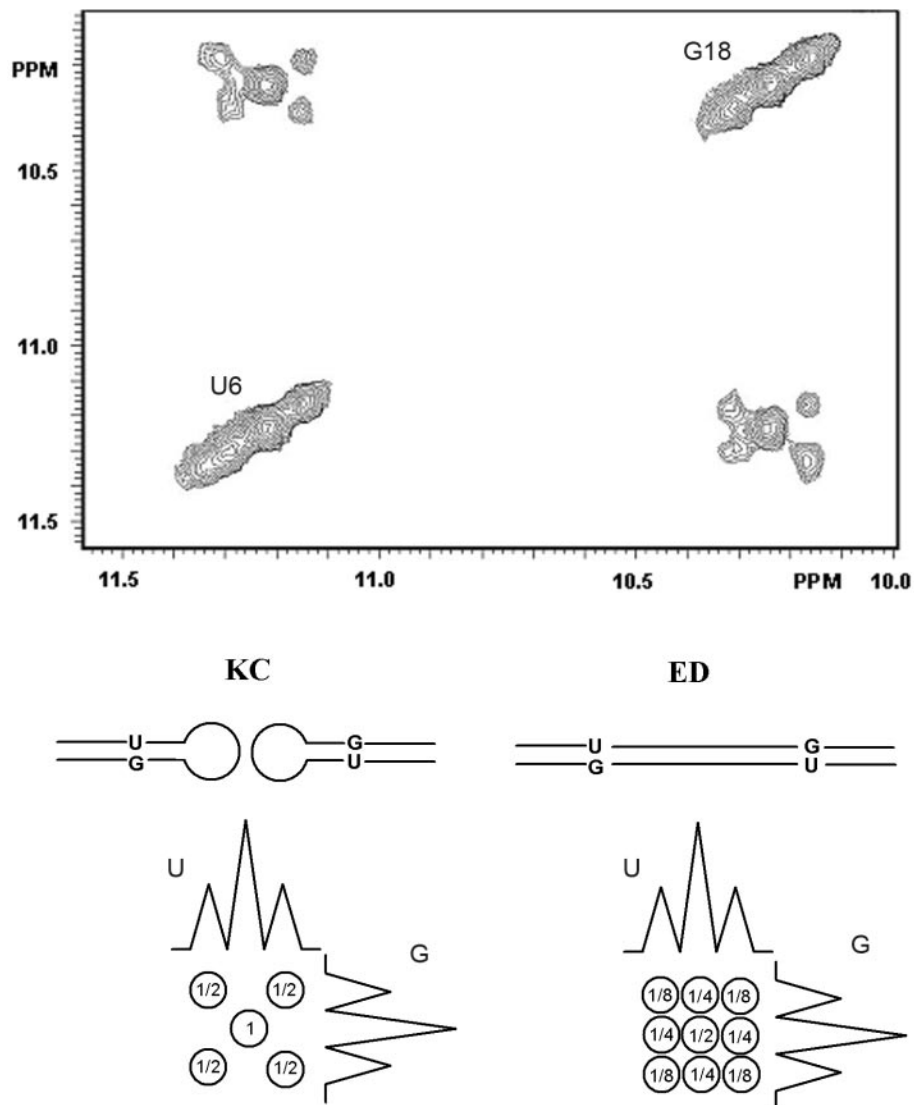


Figure 3. Pattern of the NOE connectivity between N3H(U6) and N1H(G18) obtained on a NOESY spectra. This was run at 10°C without ^{15}N decoupling during t1 and t2 on a mixture 1:1 formed by an unlabelled SL1 oligonucleotide and a $^{15}\text{N}/^{13}\text{C}$ labelled SL1 oligonucleotide. For an extended duplex, 25% of ^{14}N - ^{14}N , 25% of ^{15}N - ^{15}N and 50% of ^{14}N - ^{15}N connectivities are expected due to the interaction between strands of different labelling. For the 'kissing complex', 50% of ^{14}N - ^{14}N and 50% of ^{15}N - ^{15}N connectivities are expected, due to the interaction between strands of same labelling. Following these data and taking into account the ^{15}N - ^1H coupling constant, a scheme shows the two different patterns predicted for an extended duplex (ED) and for a 'kissing complex' (KC), demonstrating that in our experimental conditions SL1 formed a loop-loop complex.

Molecular graphics image were visualized with the MOLMOL program (38). All helical parameters were determined using CURVES (39,40).

Coordinates

The coordinates of the eighteenth structures have been deposited in the Protein Data Bank (PDB ID code is **2F4X**).

RESULTS

NMR spectroscopy and assignments

The sequential assignment of the imino and amino protons of the dimer was obtained by using a 80 ms mixing time NOESY spectrum recorded at 10°C. The three guanine imino protons

of the loop were linked by hydrogen bonds to cytosines C11, C13, C15 and a sequential assignment G10-G14-G12 was observed. These data were confirmed by using the ^{15}N HSQC and HNN-COSY spectra. ^{15}N HSQC discriminated the ^{15}NH belonging to T, G or U by their ^{15}N chemical shifts (41). HNN-COSY (42) allowed the direct observation of hydrogen coupling by a cross hydrogen scalar coupling between the imino $^{15}\text{N}1$ of the donor guanines with the hydrogen bond acceptor $^{15}\text{N}4$ atom of the complementary cytosines, in the Watson-Crick base pairs. The pairing of G7 and of C17 bases, which is essential for the modelling, has been shown by several spectra. First, the connectivity $^{15}\text{N}1$ - ^1H of G7 clearly appears on the ^{15}N HSQC spectrum, demonstrating the pairing of G17. Second, the observation of two NH2-H5 connectivities on the 80 ms NOESY spectrum showed a rotation hindrance of the amino group of C17 and strongly suggested its pairing.

Lastly, weak connectivity appeared between $^{15}\text{N1(G7)}$ and the $^{15}\text{N4 (C17)}$ on the HNN-COSY.

All the H6/8—H1' connectivities, except the sequential way between H8(A9) and H1'(G10), were directly assigned on the NOESY spectra at 80 ms mixing time in H₂O (Figure 2). H2 assignment of A8, A9 and A16 were discriminated from the corresponding H8 resonances by using their $^{13}\text{C2}$ and $^{13}\text{C8}$ chemical shifts on the HSQC spectra at 10°C. The assignments were confirmed using NOESY-HSQC, HCCH-TOCSY to connect H1' to H2', H1' to H6 or H8 in HCN (43) as well as H8 to H2 of adenines in HMBC (44).

Size of the complex and special features

The diffusion coefficient was obtained by NMR pulsed field-gradient spin-echo techniques (35). The size of the loose complex was determined by comparing this coefficient with the extended complex and the mutated SL1-A12. These two structures were already characterized as an extended duplex of 2×24 nucleotides (28,45) and as a stem-loop monomer of 23 nt (46), respectively. At 10°C, the diffusion coefficients of the extended duplex and those of the loose complex were the same ($0.52 \pm 0.05 \cdot 10^{-6}$ and $0.51 \pm 0.05 \cdot 10^{-6} \text{ cm}^2/\text{s}$), showing a similar size (2×24 mer). This result confirms the data of Muriaux *et al.* (27) showing that the loose duplex migration is similar to the extended duplex on non-denaturing PAGE on longer DIS sequences.

To discriminate between the presence of the loop-loop complex and that of the extended complex we performed NOESY WATERGATE. This was done without ^{15}N decoupling during t1 and t2 at 80 ms mixing time on a 1:1 mixture of a labelled $^{13}\text{C}^{15}\text{N}$ SL1-24mer and an unlabelled SL1-24mer. As shown by Takashashi *et al.* (28), the N3H(U6)-N1H(G18) NOESY connectivity is predicted to show quite different patterns for the extended duplex and the loop-loop complex due to non existent ^{14}NH - ^{15}NH connectivities in the loop complex. Figure 3 shows a clear pattern with a central connectivity $^{14}\text{N1H}$ - $^{14}\text{N3H}$ surrounded by four smallest cross-peaks $^{15}\text{N1H}$ - $^{15}\text{N3H}$ with $J(^{15}\text{N}, ^1\text{H}) = 95 \text{ Hz}$, in agreement with a loop-loop association. Confirmation of our NH assignments in the 'kissing complex' was obtained by ^{15}N double-half-filter NOESY experiment on a mixing (1:1) of unlabelled SL1 and $^{13}\text{C}^{15}\text{N}$ labelled SL1 (47).

As expected for a loop-loop complex, the $^{15}\text{N} \rightarrow ^{15}\text{N}$ or the $^{14}\text{N} \rightarrow ^{14}\text{N}$ sub spectra showed only the connectivities between residues belonging to the stem while the $^{15}\text{N} \rightarrow ^{14}\text{N}$ or the $^{14}\text{N} \rightarrow ^{15}\text{N}$ sub spectra showed the connectivities NH(G10)-NH(G12) and NH(G12)-NH(G14) as well as the corresponding NH-NH2 connectivities.

Several unique features were noticed on the NOESY, COSY or TOCSY spectra. First, A8 and A9 were not found in *anti* conformation and their sugar in north conformation. Second, the H2 proton of A16 showed medium or strong connectivities with the H2 proton of A9 and with the H8 and H1' protons of G10. Third, the imino proton of G10 showed connectivities with the H1', H2 and H8 protons of A16.

NMR structure determination and general description of the loose complex

The part of the dimer corresponding to the stems (G0 to G7 and C17 to C23) and to the loops (G10 to C15) showed a typical

Table 1. NMR restraints and statistics for the 18 converged structures

| | |
|---|--------------------------------------|
| Distance and dihedral restraints | |
| Intra-residue distance restraints | 120 |
| Inter-residue distance restraints | 236 |
| Hydrogen bonding distance restraints | 108 |
| Total distance restraints | 464 |
| Torsion angle restraints for sugar pucker | 92 |
| Backbone torsion angle restraints | 200 |
| Total angle restraints | 292 |
| Total restraints | 756 |
| Precision of lowest-energy structures | |
| Number of structures | 18 |
| NMR distance restraint violations (up to 0.2 Å) | 0 |
| R.m.s.d ^a | |
| Stem helix (bases G1-G7, C17-C23) | $0.59 \pm 0.15 \text{ \AA}$ |
| Central stem helix (bases G10-C15) | $0.53 \pm 0.17 \text{ \AA}$ |
| Adenines of junction | $0.48 \pm 0.20 \text{ \AA}$ |
| Whole kissing complex | $0.59 \pm 0.16 \text{ \AA}$ |
| Energy | |
| Total Energy | $-1255 \pm 29 \text{ kcal.mol}^{-1}$ |

^aR.m.s.d: root mean square deviation \pm SD in position of heavy atoms compared to the global average structure.

double helix A-type conformation and no ambiguity with the intra and intermolecular connectivities assignment. The A8 and A9 nucleotides adopt a different conformation accommodating the transition from an A-RNA conformation to a loop-loop association with a minimum of constraints. They present original connectivities whose discrimination between inter-strand and intra-strand is not trivial. In order to choose between intra and inter connectivities, four [^1H , ^1H]NOESY spectra were obtained using a $^{13}\text{C}(\omega_1, \omega_2)$ double-half-filter where each half-filter was applied alternatively. A linear combination of these four spectra gave separately the $^{13}\text{C} \rightarrow ^{13}\text{C}$, $^{13}\text{C} \rightarrow ^{12}\text{C}$ and $^{12}\text{C} \rightarrow ^{12}\text{C}$ connectivities (48). The strong NOE connectivities H8(A8)-H1'(A8) and H8(A9)-H1'(A9) were found intra-strand and H1'(G10)-H2(A16), H8(G10)-H2(A16), H2(A8)-H1'(A9) inter-strand. We attempted to obtain supplementary information concerning the relative orientation of the loop and the stem by measuring residual dipolar coupling. This experiment, which requires partial orientation of the molecule, produced a transition of the 'kissing complex' toward the extended dimer conformation in presence of phages. Classical A-helix connectivities were observed for G1 to G7, G10 to C15 and C17 to C23. For these nucleotides, the following dihedral restraints were used: $\alpha = -85 \pm 50^\circ$, $\beta = -165 \pm 50^\circ$, $\gamma = 53 \pm 50^\circ$, $\epsilon = 180 \pm 50^\circ$ and $\zeta = -60 \pm 50^\circ$ (Table 1). Cross-validation of our structure was obtained by removing 25% of the restraints involving G7, A8, A9, G10, A16 and C17. The eighteen converged structures showed no violation of distance restraints $>0.2 \text{ \AA}$. A very weak connectivity H2(A9)-H1'(G10) was observed at 300 ms mixing time, corresponding to a distance of 6.2 \AA in our structures. It can be explained by the spin diffusion way H2(A9)→H2(A16)→H1'(G10).

The solution structure of the loose complex of HIV-1_{Lai} showed a symmetric homodimer fully paired on the stems and the loops (Figure 4). These two loops, which are hydrogen bonded by six central GC base pairs, were roughly located in the axe of the stems. The overall phosphodiester backbone was similar to the canonical A-type geometry, except for the

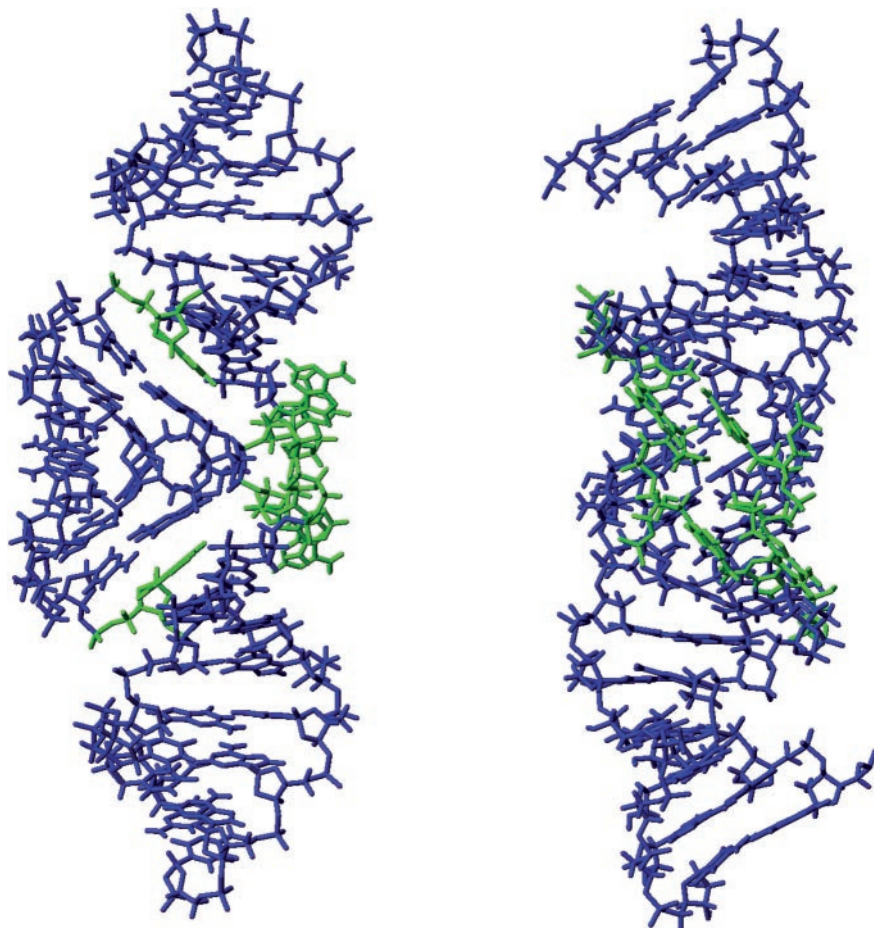


Figure 4. Perpendicularly views of the average structure obtained with the 18 converged structures. G0 that was not restrained was not presented in the figure. The A8, A9 and A16 nucleotides are shown in green.

residues A8 and A9 whose sugars were in south conformation and positioned *syn*.

The molecular edifice was stabilized by stacking of A16 with G7-C17 of the other strand, in agreement with the observed upfield shift resonance of H1'(C17) (5.1 p.p.m.).

DISCUSSION

The solution structure presented in our NMR study is very similar to the crystallographic structure (32) (Figure 5), however significant differences emerged when comparisons were made with the NMR structure obtained in the study by Mujeeb *et al.* (31). The orientation of the loops versus the stems represented a major difference between these two NMR structures (Figure 5). A valid explanation accounting for this difference lies with the discrepancies in obtaining the restraints for the structures in both studies (Table 2). The overall conformation of our 'kissing complex' presented similarities with the crystallographic structure (23), particularly in the co-linearity of the three helices as well as the stacking of A16 with G7-C17. However, several elements involving the A8, A9 and A16 nucleotides were different. A9 was found in proximity of A16 in agreement with a medium NOE between H2(A9) and H2(A16) (Table 2). In addition, the sequential way

between A8-A9 indicated the proximity of A8 to A9 of the same strand. These data did not support the conformational model where A8 and A9 are found in the bulged out position as observed in the crystallographic structure. It is worth noting that in our structure, the relative positions of the four adenines A8 and A9 of the two strands were such that their H2 and H8 resonances were exposed to small ring current effect supported by their chemical shift (Figure 6). On the contrary, in the crystallographic structure A8 and A9 were found stacked. The bulged out conformation of A8 and A9 generated a 'hole' between G7-C17 and A16 ($C1'(A16)-C1'(C17) = 15 \text{ \AA}$), which did not exist in our solution structure ($C1'(A16)-C1'(C17) = 5 \text{ \AA}$) (Figure 5). The bulging out phenomenon was a destabilizing factor, which can be compensated by the interdimer stacking of the adenines in the crystalline form. In solution, the absence of such a self-association could explain the bulged in conformation where the corresponding hole is filled by A8 and A9.

Even though a weak part of the bulged out conformation may coexist in solution, our data show clearly that the majority conformation did not own extra adenines at pH 5.7. However, it can be pointed out that $^1\text{H}-^{13}\text{C}$ HSQC spectra showed that adenine A8 underwent a protonation-deprotonation equilibrium with an apparent pK of 6.4 in 3.3 mM MgCl₂ for DIS23 (Mal) (49). Besides, fluorescence studies and mutation on A8

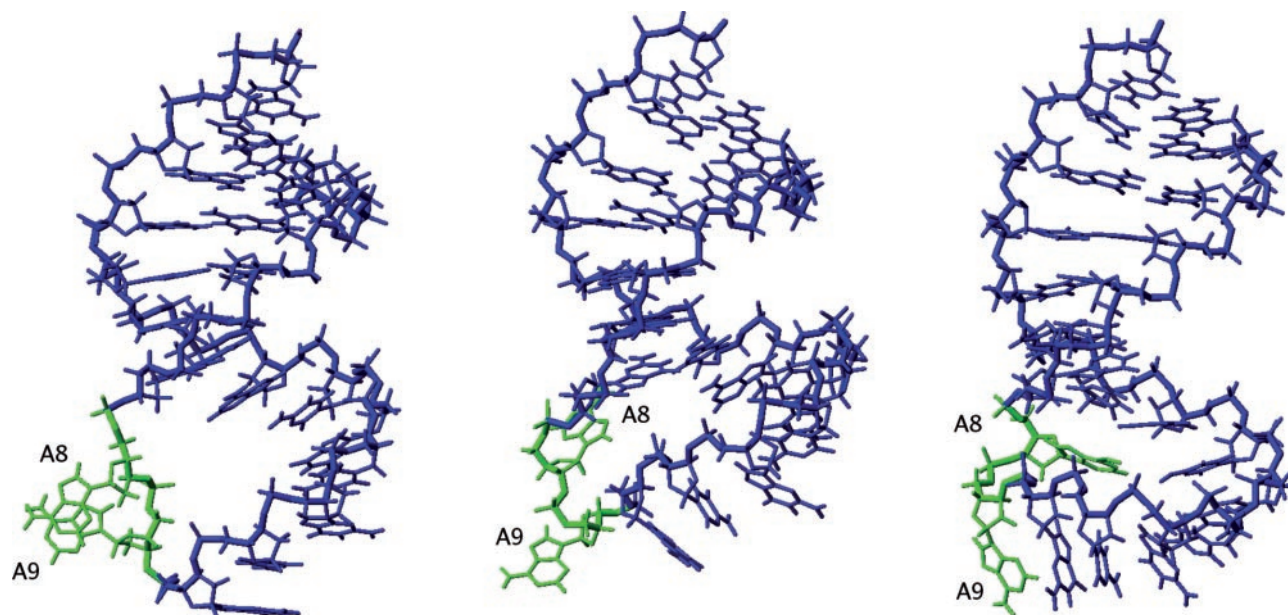


Figure 5. Comparison between the crystallographic structure by Ennifar *et al.* (32) (left), our solution structure (middle) and the NMR structure by Mujeeb *et al.* (31) (right). The hole generated by the bulged out conformation can be observed on the left-structure. In order to improve the sight of each nucleotide, only one strand is shown.

Table 2.

| Nature of the restraints | Experimental data | Crystallographic structure |
|--|---------------------------|---|
| Differences between our experimental data and the structural elements of the crystallographic structure (32) | | |
| H8(A9)-H1'(A8) | Weak | AA > 5 and AB > 10 Å |
| H1'(A9)-H2(A8) | Weak | AA > 5 and AB > 10 Å |
| H2(A9)-H2(A16) | Medium | AA and AB > 10 Å |
| H8(A8)-H1'(A8) | <i>syn</i> | <i>anti</i> |
| H8(A9)-H1'(A9) | <i>syn</i> | <i>anti</i> |
| H8(G7)-H8(A8) | Medium | AA and AB > 10 Å |
| H2(A9)-H1'(G10) | Very weak | AA and AB > 10 Å no possible pathway for spin diffusion |
| H2(A8)-H8(A9) | 4.5 Å | >10 Å |
| H2(A8)-H2(A9) | No NOE | 4.3 Å |
| A8 | South | North |
| G7 | North | South |
| δ H2A8 | 8.17 p.p.m. ^{a*} | A8 and A9 stacked |
| δ H8A8 | 7.91 p.p.m. ^{a*} | |
| δ H2A9 | 8.18 p.p.m. ^{a*} | |
| δ H8A9 | 8.0 p.p.m. ^{a*} | |
| | A8 and A9 unstacked | |
| Nature of the restraints | Experimental data | Other NMR structure |
| Differences between our experimental data and the structural elements of the Mujeeb <i>et al.</i> NMR structure (31) | | |
| H2(A16)-H1'(C17) | Strong | >7 Å |
| H5(C17)-H1'(A16) | Weak | 2.6 Å |
| H1'(A9)-H2(A8) | Medium | >7 Å |
| H2(A8)-H8(A9) | Weak | >7 Å |
| H6(C17)-H8(A16) | Medium | 6.3 Å |
| A8 | <i>syn</i> | <i>anti</i> |
| A9 | <i>syn</i> | <i>anti</i> |
| C17 | <i>anti</i> | <i>syn</i> |
| H2(A8)-H2(A16) | No NOE | 3.5 Å |
| H2(A8)-H2(A9) | No NOE | 3.4 Å |
| H2(A16)-H5(C17) | No NOE | 3.7 Å |
| G7-C17 | Paired | Unpaired |
| A8,A9 | South | North |
| G7 | North | South |

AA and AB characterize distances involving intra- and inter-strand residues.

Strong, medium, weak and very weak are relative to the intensity of the NOEs.

^{a*}The value of the chemical shift of H2 and H8 of A8 and A9 indicated that their H2 and H8 proton are not efficiently shifted upfield by a ring current effect and consequently are far from ring of other bases.

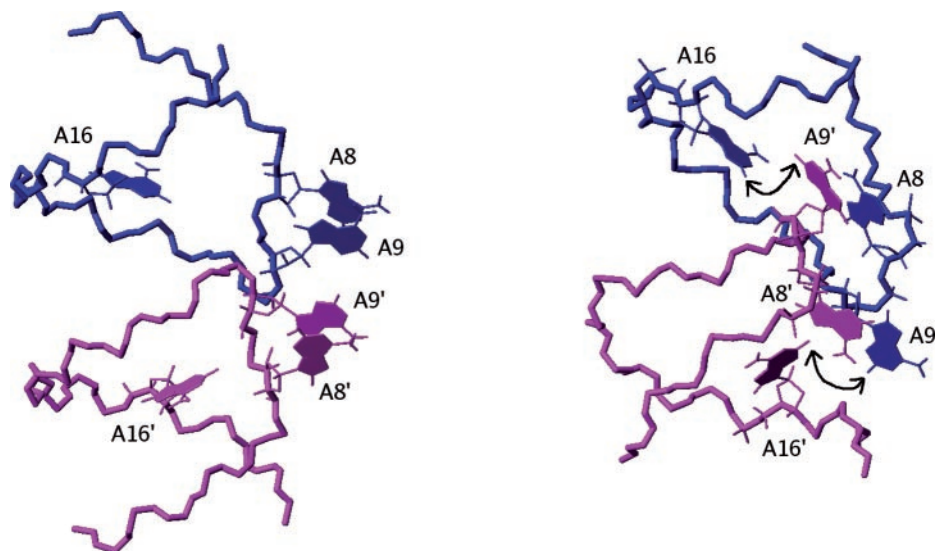


Figure 6. Comparison between the bulged in (32) and bulged out conformations showing the relative location of the A8, A9 and A16 adenines. The nearest adenines are A8 and A9 in the crystalline structure (left) and A8 and A9' in the solution structure (right). The double-arrow shows the proximity of H2(A9) and H2(A16).

indicated that the pH dependence of the NCp7-catalyzed DIS isomerization was correlated to the protonation state of A8 (49). Although it may be hazardous to deduce the behaviour of DIS23(Lai) from those of DIS23(Mal), because the transition toward the extended conformation can be easily induced by increasing the temperature for Lai (30) contrary to Mal (29), these data suggest that the ratio between bulged out and bulged in conformation for DIS23(Lai) should be different at physiological pH.

In order to explain the discrepancy in A8 and A9 conformation between the solution structure and the crystal structure we suggest that the crystallization process shifts the equilibrium between a bulged in and a bulged out conformation. The transition from a bulged in to a bulged out conformation generated a change in the position of the base from *syn* to *anti* for A8 and A9 as well as the conformation of the sugar from south to north for A8. This reversed the adenine positions and allowed the bulged out conformation (Figure 6). Further differences between the NMR and crystallographic studies emerge from addition of salt during sample preparation. Our sample was analyzed in the absence of salts whereas magnesium chloride was present in the mixture for crystallographic procedure. Addition of magnesium (1 ion for every 5 nt) on the 'kissing complex' structure did not change the NOESY spectrum, however it broadened the resonance lines. Increased magnesium concentrations caused precipitation prevented high-resolution NMR and bringing evidence of self-association.

Several results concerning the chemical probes support our structure: (i) chemical probes showed that N1(A8) is hyper reactive whereas N7(A8) is not (50). This added to the idea that N1 of A8 is not involved in a hydrogen bond and that N7 of A8 has limited accessibility to the chemical probes, (ii) methylation of N1 of the adenine bases A8, A9 and A16 did not impair dimerization in agreement with the fact that these three adenines are not paired, (iii) the three phosphates 5'-P(A8)-P(A9)-P-3' may be ethylated without

impairing dimerization. This region allows the ethylation without destroying the initial conformation.

In our structure, the unpaired adenines of A8, A9 and A16 had weaker interactions than those of the extended dimer where they formed a zipper like motif. Moreover weaker stacking was observed for the six central GC base pairs. These features would explain the state of 'kissing complex' as less thermostable than the linear duplex.

We have previously shown that the conformational change from 'kissing complex' to extended dimer (induced by an increase of the temperature) kept the six central base pairs GC paired during the transition (30). We suggest that increased temperature lead to the weakness of adenines A8, A9 and A16 interaction and the fading of G7-C17 pairing. The mechanism of this transition starts by distancing of A9 from A16, allowing an increase in the stacking of the central GC base pairs. The stems are less thermostable than the central mini double helix, because they comprise two AT and one wobble GU base pairs. Their melting may allow for molecular rearrangements from paired loops to extended duplex.

Our study provides a new structure of the SL1 HIV-1_{Lai} loop-loop dimer. The next step in understanding the RNA dimerization is to solve the structure of the entire SL1 hairpin (51). This region contains a bulge of well-conserved residues. Mutations in the bulge cause defects in encapsidation (15). Antiretroviral therapy research has not yet targeted the process of dimerization of the viral RNA, thereby providing an avenue for future drug design strategies.

ACKNOWLEDGEMENTS

The authors thank the Laboratoire de RMN, Institut de Biologie Structurale in Grenoble, France, for running NMR spectra at 800 MHz. This work was supported by Agence Nationale de la Recherche sur le Sida grant (J.P.). F.K. was supported by a fellowship from the Ministère Français de l'Éducation

Nationale, de la Recherche et de la Technologie. The Open Access publication charges for this article were waived by Oxford University Press.

Conflict of interest statement. None declared.

REFERENCES

- Coffin, J.M. (1984) Structure of the retroviral genome. In Weiss, R., Teich, N., Varmus, H. and Coffin, J. (eds), *Tumor Viruses*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, Vol. I, pp. 261–368.
- Hu, W.S. and Temin, H.M. (1990) Genetic consequences of packaging two RNA genomes in one retroviral particle: pseudodiploidy and high rate of genetic recombination. *Proc. Natl Acad. Sci. USA*, **87**, 1556–1560.
- Hu, W.S. and Temin, H.M. (1990) Retroviral recombination and reverse transcription. *Science*, **250**, 1227–1233.
- Panganiban, A.T. and Fiore, D. (1988) Ordered interstrand and intrastrand DNA transfer during reverse transcription. *Science*, **241**, 1064–1069.
- Stuhlmann, H. and Berg, P. (1992) Homologous recombination of copackaged retrovirus RNAs during reverse transcription. *J. Virol.*, **66**, 2378–2388.
- Temin, H.M. (1991) Sex and recombination in retroviruses. *Trends Genet.*, **7**, 71–74.
- Temin, H.M. (1993) Retrovirus variation and reverse transcription: abnormal strand transfers result in retrovirus genetic variation. *Proc. Natl Acad. Sci. USA*, **90**, 6900–6903.
- Bender, W. and Davidson, N. (1976) Mapping of poly(A) sequences in the electron microscope reveals unusual structure of type C oncornavirus RNA molecules. *Cell*, **7**, 595–607.
- Bender, W., Chien, Y.H., Chattopadhyay, S., Vogt, P.K., Gardner, M.B. and Davidson, N. (1978) High-molecular-weight RNAs of AKR, NZB, and wild mouse viruses and avian reticuloendotheliosis virus all have similar dimer structures. *J. Virol.*, **25**, 888–896.
- Hoglund, S., Ohagen, A., Goncalves, J., Panganiban, A.T. and Gabuzda, D. (1997) Ultrastructure of HIV-1 genomic RNA. *Virology*, **233**, 271–279.
- Chien, Y.H., Junghans, R.P. and Davidson, N. (1980) Molecular biology of RNA tumor viruses. *Electron Microscopic Analysis of the Structure of RNA Tumor Virus Nucleic Acids*. Academic Press, NY, pp. 395–446.
- Marquet, R., Paillart, J.C., Skripkin, E., Ehresmann, C. and Ehresmann, B. (1994) Dimerization of human immunodeficiency virus type 1 RNA involves sequences located upstream of the splice donor site. *Nucleic Acids Res.*, **22**, 145–151.
- Skripkin, E., Paillart, J.C., Marquet, R., Ehresmann, B. and Ehresmann, C. (1994) Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization *in vitro*. *Proc. Natl Acad. Sci. USA*, **91**, 4945–4949.
- Bernacchi, S. and Mely, Y. (2001) Exciton interaction in molecular beacons: a sensitive sensor for short range modifications of the nucleic acid structure. *Nucleic Acids Res.*, **29**, E62–E62.
- Clever, J.L. and Parslow, T.G. (1997) Mutant human immunodeficiency virus type 1 genomes with defects in RNA dimerization or encapsidation. *J. Virol.*, **71**, 3407–3414.
- Haddrick, M., Lear, A.L., Cann, A.J. and Heaphy, S. (1996) Evidence that a kissing loop structure facilitates genomic RNA dimerisation in HIV-1. *J. Mol. Biol.*, **259**, 58–68.
- Laughrea, M., Shen, N., Jette, L. and Wainberg, M.A. (1999) Variant effects of nonnative kissing-loop hairpin palindromes on HIV replication and HIV RNA dimerization: role of stem-loop B in HIV replication and HIV RNA dimerization. *Biochemistry*, **38**, 226–234.
- Liang, C., Rong, L., Laughrea, M., Kleiman, L. and Wainberg, M.A. (1998) Compensatory point mutations in the human immunodeficiency virus type 1 Gag region that are distal from deletion mutations in the dimerization initiation site can restore viral replication. *J. Virol.*, **72**, 6629–6636.
- McBride, M.S., Schwartz, M.D. and Panganiban, A.T. (1997) Efficient encapsidation of human immunodeficiency virus type 1 vectors and further characterization of cis elements required for encapsidation. *J. Virol.*, **71**, 4544–4554.
- Paillart, J.C., Skripkin, E., Ehresmann, B., Ehresmann, C. and Marquet, R. (1996) A loop-loop “kissing” complex is the essential part of the dimer linkage of genomic HIV-1 RNA. *Proc. Natl Acad. Sci. USA*, **93**, 5572–5577.
- Paillart, J.C., Marquet, R., Skripkin, E., Ehresmann, C. and Ehresmann, B. (1996) Dimerization of retroviral genomic RNAs: structural and functional implications. *Biochimie*, **8**, 639–653.
- Shen, N., Jette, L., Liang, C., Wainberg, M.A. and Laughrea, M. (2000) Impact of human immunodeficiency virus type 1 RNA dimerization on viral infectivity and of stem-loop B on RNA dimerization and reverse transcription and dissociation of dimerization from packaging. *J. Virol.*, **74**, 5729–5735.
- Muriaux, D., Girard, P.M., Bonnet-Mathoniere, B. and Paoletti, J. (1995) Dimerization of HIV-1LAI RNA at low ionic strength. An autocomplementary sequence in the 5' leader region is evidenced by an antisense oligonucleotide. *J. Biol. Chem.*, **270**, 8209–8216.
- Windbichler, N., Werner, M. and Schroeder, R. (2003) Kissing complex-mediated dimerisation of HIV-1 RNA: coupling extended duplex formation to ribozyme cleavage. *Nucleic Acids Res.*, **31**, 6419–6427.
- Laughrea, M. and Jette, L. (1996) Kissing-loop model of HIV-1 genome dimerization: HIV-1 RNAs can assume alternative dimeric forms, and all sequences upstream or downstream of hairpin 248–271 are dispensable for dimer formation. *Biochemistry*, **35**, 1589–1598.
- Laughrea, M. and Jette, L. (1996) HIV-1 genome dimerization: formation kinetics and thermal stability of dimeric HIV-1LAI RNAs are not improved by the 1–232 and 296–790 regions flanking the kissing-loop domain. *Biochemistry*, **35**, 9366–9374.
- Muriaux, D., Fosse, P. and Paoletti, J. (1996) A kissing complex together with a stable dimer is involved in the HIV-1LAI RNA dimerization process *in vitro*. *Biochemistry*, **35**, 5075–5082.
- Takahashi, K.I., Baba, S., Chattopadhyay, P., Koyanagi, Y., Yamamoto, N., Takaku, H. and Kawai, G. (2000) Structural requirement for the two-step dimerization of human immunodeficiency virus type 1 genome. *RNA*, **6**, 96–102.
- Rist, M.J. and Marino, J.P. (2002) Mechanism of nucleocapsid protein catalyzed structural isomerization of the dimerization initiation site of HIV-1. *Biochemistry*, **41**, 14762–14770.
- Theilleux-Delalande, V., Girard, F., Huynh-Dinh, T., Lancelot, G. and Paoletti, J. (2000) The HIV-1(LAI) RNA dimerization. Thermodynamic parameters associated with the transition from the kissing complex to the extended dimer. *Eur. J. Biochem.*, **267**, 2711–2719.
- Mujeeb, A., Clever, J.L., Billeci, T.M., James, T.L. and Parslow, T.G. (1998) Structure of the dimer initiation complex of HIV-1 genomic RNA. *Nat. Struct. Biol.*, **5**, 432–436.
- Ennifar, E., Walter, P., Ehresmann, B., Ehresmann, C. and Dumas, P. (2001) Crystal structures of coaxially stacked kissing complexes of the HIV-1 RNA dimerization initiation site. *Nat. Struct. Biol.*, **8**, 1064–1068.
- Milligan, J.F., Groebe, D.R., Witherell, G.W. and Uhlenbeck, O.C. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.*, **15**, 8783–8798.
- Davanloo, P., Rosenberg, A.H., Dunn, J.J. and Studier, F.W. (1984) Cloning and expression of the gene for bacteriophage T7 RNA polymerase. *Proc. Natl. Acad. Sci. USA*, **81**, 2035–2039.
- Lapham, J., Rife, P.B., Moore, J.P. and Crothers, D.M. (1997) Measurement of diffusion constants for nucleic acids by NMR. *J. Biomol. NMR*, **10**, 255–262.
- Stallings, S.C. and Moore, P.B. (1997) The structure of an essential splicing element: stem loop IIa from yeast U2 snRNA. *Structure*, **5**, 1173–1185.
- Brünger, A.T., Adams, P.D., Clore, G.M., De Lano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S. et al. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D. Biol. Crystallog.*, **54**, 905–921.
- Koradi, R., Billeter, M. and Wuthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics*, **14**, 51–55, 29–32.
- Lavery, R. and Sklenar, H. (1989) Defining the structure of irregular nucleic acids: conventions and principles. *J. Biomol. Struct. Dyn.*, **6**, 655–667.
- Lavery, R. and Sklenar, H. (1988) The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.*, **6**, 63–91.

41. Nikonowicz, E.P. and Pardi, A. (1993) An efficient procedure for assignment of the proton, carbon and nitrogen resonances in $^{13}\text{C}/^{15}\text{N}$ labeled nucleic acids. *J. Mol. Biol.*, **232**, 1141–1156.
42. Dingley, A.J. and Grzesiek, S. (1998) Direct observation of Hydrogen bonds in nucleic acid base pairs by internucleotide $^2\text{J}_{\text{NN}}$ Couplings. *J. Am. Chem. Soc.*, **120**, 8293–8297.
43. Sklenar, V., Peterson, R.D., Rejante, M.R. and Feigon, J. (1993) Two- and three-dimensional HCN experiments for correlating base and sugar resonances in $^{15}\text{N}, ^{13}\text{C}$ -labeled RNA oligonucleotides. *J. Biomol. NMR*, **3**, 721–727.
44. van Dongen, M.J., Wijmenga, S.S., Eritja, R., Azorin, F. and Hilbers, C.W. (1996) Through-bond correlation of adenine H2 and H8 protons in unlabeled DNA fragments by HMBC spectroscopy. *J. Biomol. NMR*, **8**, 207–212.
45. Girard, F., Barbault, F., Gouyette, C., Huynh-Dinh, T., Paoletti, J. and Lancelot, G. (1999) Dimer initiation sequence of HIV-1Lai genomic RNA: NMR solution structure of the extended duplex. *J. Biomol. Struct. Dyn.*, **16**, 1145–1157.
46. Kieken, F., Arnoult, E., Barbault, F., Paquet, F., Huynh-Dinh, T., Paoletti, J., Genest, D. and Lancelot, G. (2002) HIV-1(Lai) genomic RNA: combined used of NMR and molecular dynamics simulation for studying the structure and internal dynamics of a mutated SL1 hairpin. *Eur. Biophys. J.*, **31**, 521–531.
47. Slijper, M., Kaptein, R. and Boelens, R. (1996) Simultaneous ^{13}C and ^{15}N isotope editing of biomolecular complexes. Application to a mutant lac repressor headpiece DNA complex. *J. Magn. Reson. B.*, **111**, 199–203.
48. Otting, G., Senn, H., Wagner, G. and Wüthrich, K. (1986) Editing of 2D ^1H NMR spectra using X half-filters. combined use with residue-selective ^{15}N labeling of proteins. *J. Magn. Reson.*, **70**, 500–505.
49. Mihailescu, M.R. and Marino, J.P. (2004) A proton-coupled dynamic conformational switch in the HIV-1 dimerization initiation site kissing complex. *Proc. Natl Acad. Sci. USA*, **101**, 1189–1194.
50. Paillart, J.C., Westhof, E., Ehresmann, C., Ehresmann, B. and Marquet, R. (1997) Non-canonical interactions in a kissing loop complex: the dimerization initiation site of HIV-1 genomic RNA. *J. Mol. Biol.*, **270**, 36–49.
51. Kuiken, C., Thakallapalli, R., Esklid, A. and de Ronde, A. (2000) Genetic analysis reveals epidemiologic patterns in the spread of human immunodeficiency virus. *Am. J. Epidemiol.*, **152**, 814–822.