8-2007

HIV-2 RNA Dimerization is Regulated by Intramolecular Interactions in Vitro

Tayyba T. Baig

Jean-Marc Lanchy

J. Stephen Lodmell
University of Montana - Missoula, stephen.lodmell@umontana.edu

HIV-2 RNA dimerization is regulated by intramolecular interactions in vitro

TAYYBA T. BAIG, JEAN-MARC LANCHY, and J. STEPHEN LODMELL
Division of Biological Sciences, The University of Montana, Missoula, Montana 59812, USA

ABSTRACT

Genomic RNA dimerization is an essential process in the retroviral replication cycle. In vitro, HIV-2 RNA dimerization is mediated at least in part by direct intermolecular interaction at stem–loop 1 (SL1) within the 5′-untranslated leader region (UTR). RNA dimerization is thought to be regulated via alternate presentation and sequestration of dimerization signals by intramolecular base-pairings. One of the proposed regulatory elements is a palindrome sequence (pal) located upstream of SL1. To investigate the role of pal in the regulation of HIV-2 dimerization, we randomized this motif and selected in vitro for dimerization-competent and dimerization-impaired RNAs. Energy minimization folding analysis of these isolated sequences suggests the involvement of pal region in several short-distance intramolecular interactions with other upstream and downstream regions of the UTR. Moreover, the consensus predicted folding patterns indicate the altered presentation of SL1 depending on the interactions of pal with other regions of RNA. The data suggest that pal can act as a positive or negative regulator of SL1-mediated dimerization and that the modulation of base-pairing arrangements that affect RNA dimerization could coordinate multiple signals located within the 5′-UTR.

Keywords: palindrome; RNA dimerization; HIV-2; SELEX

INTRODUCTION

Retroviruses package two copies of their positive sense single-stranded RNA genome into a single viral particle. Analysis of the packaged genomes revealed the two RNA molecules are linked by noncovalent bonds and form a dimeric RNA structure inside the viral particle (for review, see Greatorex and Lever 1998; Paillart et al. 2004b; Russell et al. 2004). The dimeric nature of the genomic RNA has been characterized by sedimentation analysis (Cheung et al. 1972) and nondenaturing gel electrophoresis (Fu and Rein 1993). Moreover, electron microscopic studies indicated that the two RNA molecules are joined strongly with each other through a region close to their 5′ ends, termed the dimer linkage structure (DLS) (Bender and Davidson 1976; Hoglund et al. 1997).

The mechanism of dimerization promoted by the DLS was further studied in vitro using RNA fragments encompassing the 5′ end of the HIV-1 genomic RNA, which showed the spontaneous dimerization of these RNA fragments without any cellular or viral proteins (Darlix et al. 1990). The essential motif for dimerization of the HIV-1 genomic RNA was identified in vitro and called the dimerization initiation site (DIS) (Laughrea and Jette 1994; Paillart et al. 1994; Skripkin et al. 1994; Muriaux et al. 1995) or stem–loop 1 (SL1) (McBride and Panganiban 1996). In vitro, this motif mediates dimerization between two RNA molecules through a kissing loop interaction (Laughrea and Jette 1994; Paillart et al. 1994; Skripkin et al. 1994) that may proceed further through an extended intermolecular interaction to form a stable extended duplex at least when short RNA constructs are used (Laughrea and Jette 1996; Muriaux et al. 1996a; Greatorex and Lever 1998).

The dimerization properties of the HIV-2 leader RNA are somewhat different from those of the HIV-1 leader. Multiple dimerization elements have been described within the HIV-2 leader RNA, including stem–loop 1, which is homologous to the HIV-1 SL1 (Dirac et al. 2001; Lanchy and Lodmell 2002), and nucleotides within the primer-binding site (PBS) (Jossinet et al. 2001; Lanchy and Lodmell 2002). Furthermore, a 10-nucleotide (nt) palindrome sequence called pal within the encapsidation signal of the HIV-2 leader RNA has been proposed to be involved in dimerization or its regulation in vitro (Lanchy et al. 2003a).

Reprint requests to: Jean-Marc Lanchy, Division of Biological Sciences, The University of Montana, Missoula, MT 59812, USA; e-mail: jean-marc.lanchy@umontana.edu; fax: (406) 243 4304.

Article published online ahead of print. Article and publication date are at http://www.rnajournal.org/cgi/doi/10.1261/rna.483807.

The ability of genomic RNA to dimerize is a potentially important event for key steps of the viral replication cycle such as translation, encapsidation, and recombination (Darlix et al. 1990; Hu and Temin 1990; Stuhlmann and Berg 1992; Baudin et al. 1993; Fu and Rein 1993; Fu et al. 1994; Rein 1994; Jetzt et al. 2000; Mikkelsen et al. 2000; Abbink and Berkhout 2003; Sakuragi et al. 2003; D'Souza and Summers 2004; Flynn et al. 2004; Hibbert et al. 2004; Chin et al. 2005). Genomic RNA dimerization and encapsidation have been proposed to be regulated through different conformations of the leader RNA region that trigger or prevent these two processes. Two alternative conformations have been reported in vitro for HIV-1 leader RNA that show different dimerization properties (Berkhout and van Wamel 2000; Huthoff and Berkhout 2001). One of the proposed conformations exists as a rod like structure with a long-distance base-pairing interaction (LDI) between the poly (A) signal stem–loop and SL1 domains (Huthoff and Berkhout 2001), therefore masking the SL1 and inhibiting dimerization. In the other proposed conformation, the leader region can refold into a branched structure with multiple hairpins (BMH) that exposes both the poly (A) and SL1 hairpins and favors dimerization. Although disruption of the LDI causes phenotypic changes in HIV-1 replication (Ooms et al. 2004; Abbink et al. 2005), structural evidence for the two conformations has been elusive, as chemical probing of the HIV-1 genomic RNA structure in infected cells and in viral particles did not reveal the existence of a predominant LDI conformation (Paillart et al. 2004a).

In contrast to HIV-1, the HIV-2 leader RNA fragment does not dimerize efficiently through SL1 in vitro (Dirac et al. 2002a; Lanchy and Lodmell 2002; Lanchy et al. 2003a, b). It has been proposed that similar to HIV-1, monomeric HIV-2 RNA can adopt two alternative conformations that sequester and release SL1, respectively (Dirac et al. 2002a; Lanchy et al. 2003a). The impaired dimerization of HIV-2 RNA is thought to be caused by an energetically favored folding that sequesters SL1 (Dirac et al. 2002a; Lanchy et al. 2003b). We suggested that the intramolecular sequestration of SL1 occurs through base-pairing with the upstream 10-nt pal sequence (Lanchy et al. 2003b).

Since pal can affect the SL1-mediated dimerization of HIV-2 leader RNA fragments in vitro, we have characterized in the present study the role of pal by randomizing this motif and selecting from a population of random-pal RNAs those with enhanced or diminished dimerization properties. We cloned and sequenced individual RNAs from the selected pools and used their sequences to analyze the functionally important secondary structures that might be involved in silencing or enhancing SL1's role as a dimerization element. In particular, the predicted folding models suggest that the nonnative selected pal sequences (denoted pal* to distinguish from the wild-type pal sequence) were involved in different patterns of binding in the monomer and dimer RNA populations. Moreover, the different pal* binding patterns are associated with different structural presentation of SL1 that correlated with the observed dimerization abilities of the selected RNAs. Our observations underscore the importance of pal* as a regulator of dimerization by modulating SL1 presentation. In addition, similarities in the dimerization behavior and predicted folding of the monomer pool RNAs with wild-type RNA suggest a general model for pal's involvement in the regulation of dimerization.

RESULTS

Construction of pool 0 RNA and selections for dimerization-competent and dimerization-impaired RNAs

The HIV-2 leader RNA fragment (561-nt-long) was previously shown to dimerize inefficiently through SL1 in vitro (Lanchy and Lodmell 2002; Lanchy et al. 2003b). The existence of an alternative conformation of the monomeric leader RNA was proposed to be responsible for this inability to dimerize efficiently, in which a short-range intramolecular interaction (pal-SL1) occludes the SL1 dimerization element (Lanchy et al. 2003a). The 10-nt pal sequence (392–401) is located upstream of the SL1 in the HIV-2 RNA (Fig. 1). To further characterize the role of pal,

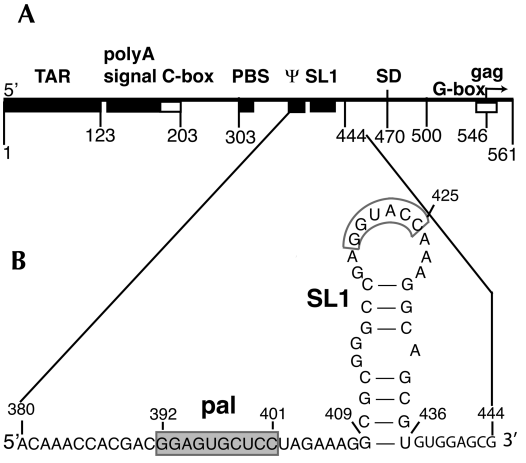

FIGURE 1. HIV-2 leader RNA and location of the pal and SL1 elements. (*A*) The 5′ leader region of HIV-2 ROD genomic RNA is represented with boxes and numbers to indicate the landmark sequences with their names indicated *above*. TAR, poly A signal, C-box, PBS, ψ, SL1, SD, G-box, and gag represent the *trans*-activation region, the poly (*A*) signal domain, C-rich sequence, the primer binding site, the encapsidation signal ψ, stem–loop 1, the major splice donor site, G-rich sequence, and the 5′ end of the Gag protein coding region, respectively. (*B*) A simplified secondary structure of nucleotides 380–444. The gray box highlights the 10-nt palindrome sequence (pal) in the encapsidation signal, and the 6-nt autocomplementary sequence essential for SL1-mediated dimerization in the apical loop of the SL1 is outlined.

we designed a 1–561 RNA construct harboring a randomized pal to determine how pal could act as a regulatory element for SL1. Nucleotides 393–401 of pal were randomized within the DNA transcription template used in this study. Hence, the preselection (pool 0) RNA species were 561-nt-long and harbored a $_{392}$GNNNNNNNNN$_{401}$ sequence (where N has an equal probability of being A, C, G, or U) instead of the wild-type $_{392}$GGAGUGCUCC$_{401}$ sequence (Figs. 1B, 2A). Randomization of this region was verified by sequencing (data not shown).

An in vitro selection/amplification technique was used to enrich the randomized populations for RNAs that were dimerization competent or impaired. The approach used in the selections of both pools is explained in the Materials and Methods and in Figure 2B. Pool 0 RNA was fractionated into dimer and monomer species by nondenaturing gel electrophoresis. RNA migrating as a dimer species was excised from the gel, subjected to RT-PCR and transcription, and the resulting RNA was used in a subsequent round of selection for dimerization-competent RNAs. Likewise, RNA migrating as a monomer species was similarly processed, and the reamplified RNA was used in a subsequent round of selection for dimerization-impaired RNAs (Fig. 2B). After initial fractionation of pool 0 for dimer and monomer RNA species, further selections were made separately to minimize cross-contamination. The percentage of RNA migrating in dimeric form increased with the consecutive rounds of dimerization-competent RNAs selection (Fig. 3A). We stopped the dimerization-competent RNAs selection when the percentage of dimer reached a plateau and remained constant for several rounds of selection (Fig. 3C). In parallel, the percentage of dimer decreased with the successive rounds of dimerization-impaired RNAs selection (Fig. 3B). We discontinued the selection when the percentage of dimer remained the same for several rounds of selection (Fig. 3C).

Interactions of in vitro selected pal sequences (pal*) within the 561-nt-long RNA

The resulting cDNAs from the last dimerization-competent (sixth round) and dimerization-impaired (fifth round) selections were cloned into pUC18 for either sequencing or expression of the individual clonal RNAs. Sixty clones from each pool were sequenced and subjected to computer-assisted lowest-energy folding. We called these two pools monomer and dimer pools, since they originated from the last dimerization-impaired and -competent RNA selections, respectively. Mfold version 2.3 (Zuker 2003) was used to predict the most stable secondary structure at 55°C for each of the 120 individual sequences. The folding temperature was set at 55°C since the RNAs were incubated at 55°C in our selection protocol. The 120 optimal secondary structures were visually analyzed, and the organization of the pal* and SL1 elements and the base-pairing partners of the selected pal* nucleotides was recorded and shown in Figure 4 by model structures and dots adjacent to the leader RNA schematic, respectively.

One striking pattern revealed by Mfold analysis was the reproducible and relatively small number of intramolecular interactions that variant pal* made with upstream and downstream RNA sequences, as evidenced by the peaks of dots in Figure 4. Distinct structural groups are formed by pal* interacting with either the C-box, 380, SL1, 444, or 500 regions. We included another structural group called free pal*, representing the variant pal* sequences where none of pal* nucleotides are base-paired in the optimal Mfold-predicted structure (7% and 13% of dimer and monomer pool sequences, respectively) (Table 3, see below). The major partners of pal* among RNA species isolated from the dimer pool were C-box, 380-region, and SL1 (Fig. 4A). They represent 27%, 45%, and 20% of the dimer pool sequences, respectively (Table 3, see below). In the monomer pool, the major regions predicted to base-pair

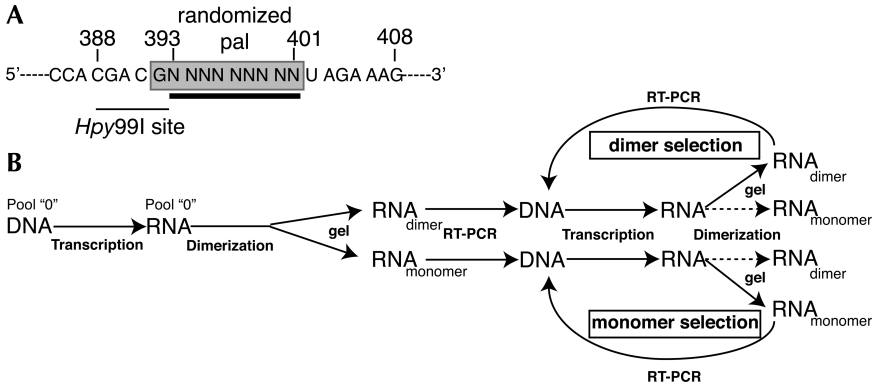

FIGURE 2. Randomized pal* sequence and in vitro selection methodology. (*A*) Detailed view of randomized pal region of the 561-nt pool 0 RNA used for in vitro selection. The pal element is a 10-nt palindrome sequence within the encapsidation signal ψ, located immediately upstream of SL1. Nine nucleotides (393–401) in the pal sequence were randomized in the template used for transcription of the pool 0 RNAs, while keeping the first nucleotide of pal (G392) constant so as not to disrupt the Hpy99I restriction site required for template construction (see Materials and Methods). (*B*) Schematic of the SELEX approach used to select and amplify the dimerization-competent and dimerization-impaired RNAs from the dimer and monomer pools, respectively. The dimeric and monomeric RNAs were partitioned on nondenaturing agarose gel. The dimeric and monomeric RNAs were excised from the pool 0 gel and used as the templates for RT-PCR to generate the DNA template to transcribe RNAs for the subsequent dimer and monomer pools. This cycle was repeated several times separately for the dimer and monomer selections. The solid arrows represent the dimeric and monomeric RNAs used for each subsequent round of selection. The dotted arrows represent those RNAs that were not processed for selection.

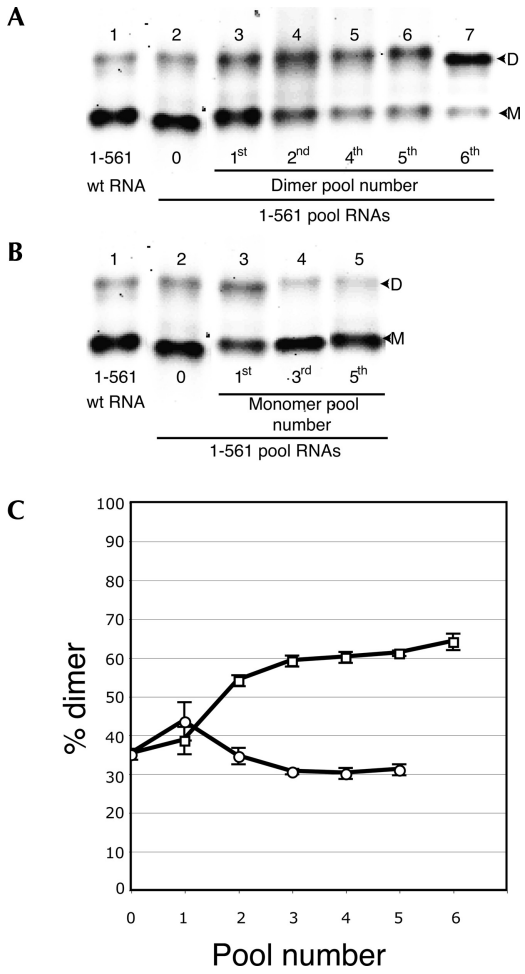

FIGURE 3. Evolution of dimerization characteristics after repeated rounds of selection for dimerization-competent and dimerization-impaired pool RNAs. (*A*) A non-denaturing dimerization gel is shown to represent the pool 0 (unselected RNA) and successive rounds of selection (first, second, fourth, fifth, and sixth) of the dimer pools. The *top* band, migrating as dimer for pool 0 RNA, was used as the starting material for each subsequent round of selection for dimerization-competent RNAs, while the *bottom* band for pool 0 RNA was used during the separate selections for dimerization-impaired RNAs. The dimerization of the wild-type RNA is shown for comparison. (*B*) A nondenaturing dimerization gel is shown to represent the pool 0 (unselected RNA) and successive rounds of selection (first, third, and fifth) of the monomer pools. The wild-type RNA is also shown for comparison. (*C*) The percentage of dimer yield from the dimer pools (open squares) and from the monomer pools (open circles) was quantified by phosphorimager scanning and analyzed with Fuji Image Gauge v3.3 software.

with pal* included the 380, SL1, 444, and 500-regions (Fig. 4B). They represent 22, 20, 30, and 12% of the monomer pool sequences respectively (Table 3, see below).

Another feature revealed by Mfold analysis was that the structural environment of SL1 differs between dimer and monomer pools and correlates with the pal*-defined structural groups. In the dimer pool, SL1 is often presented as an individualized stem–loop structure (80% of the dimer pool sequences) (Fig. 4A; Table 3, see below), while pal* interacts with sequences located upstream of itself, either the C-box element or the 380-region (27% and 45%, respectively) (Table 3, see below).

Although some pal* binding partners are located upstream of pal* (25%) (Fig. 4B; Table 3, see below), monomer pool sequences are characterized by increased interaction of pal* with sequences located downstream of SL1, either with the 444-region or the 500-region (30% and 12%, respectively) (Fig. 4B; Table 3, see below). The direct consequence is a local increase of SL1 stability, which may regulate HIV-2 RNA dimerization (see Discussion).

Interestingly, 20% of both monomer and dimer pools sequences optimal structures sequester SL1 by direct base-pairing with part of SL1 (pal*-SL1) (Fig. 4B; Table 3, see below). The pal*–SL1 interactions could be further divided in two subgroups, depending on whether pal* interacts with the 5′ stem of SL1 or its loop and 3′ stem (Fig. 4, cf. pal*-SL1 structures in panels A and B). Moreover, the pal*–(5′ stem SL1) interactions are more represented among dimer pool sequences than among monomer pool sequences (13% and 7%, respectively) (Table 3, see below).

Characterization of the RNAs isolated from the last dimer and last monomer pools

Sixteen sequenced DNA plasmid clones, chosen from the last dimer and monomer pool selections were transcribed into pure clonal RNAs. These clonal RNAs were characterized for their dimerization properties by nondenaturing gel electrophoresis. The HIV-2 leader RNA was shown previously to be dimerization-impaired (Lanchy et al. 2003a). As expected, the dimerization yield of all eight clonal RNAs (M01–8) from the monomer pool selection was quite similar to the wild type (Fig. 5A, cf. lanes 1 and 2–9), whereas the level of dimerization for six of the eight clonal RNAs (D01–6) from the dimer pool selection was higher than wild type (Fig. 5B, cf. lanes 1 and 2–7). The dimerization efficiency of the other two clonal RNAs (D07 and D08) was similar to wild type (Fig. 5B, cf. lanes 1 and 8,9). Although D07 and D08 RNAs were isolated from the last dimer pool selection, their impaired dimerization yields suggest that either they were selected as heterodimers with more dimerization-competent RNAs, such as D01–6 (Fig. 5B), or that the stringency of the selection allows carryover of some monomer RNA species at each round. In fact, the repartitioning of RNAs into dimer and monomer bands at each round, even though the RNA at each round is derived from template originating from solely dimer (or monomer) RNA in the previous round, suggests that the RNAs can adopt dimerization-competent or -impaired conformations.

Finally, analysis of the structural groups to which the 16 clonal RNAs belong showed the same correlation between

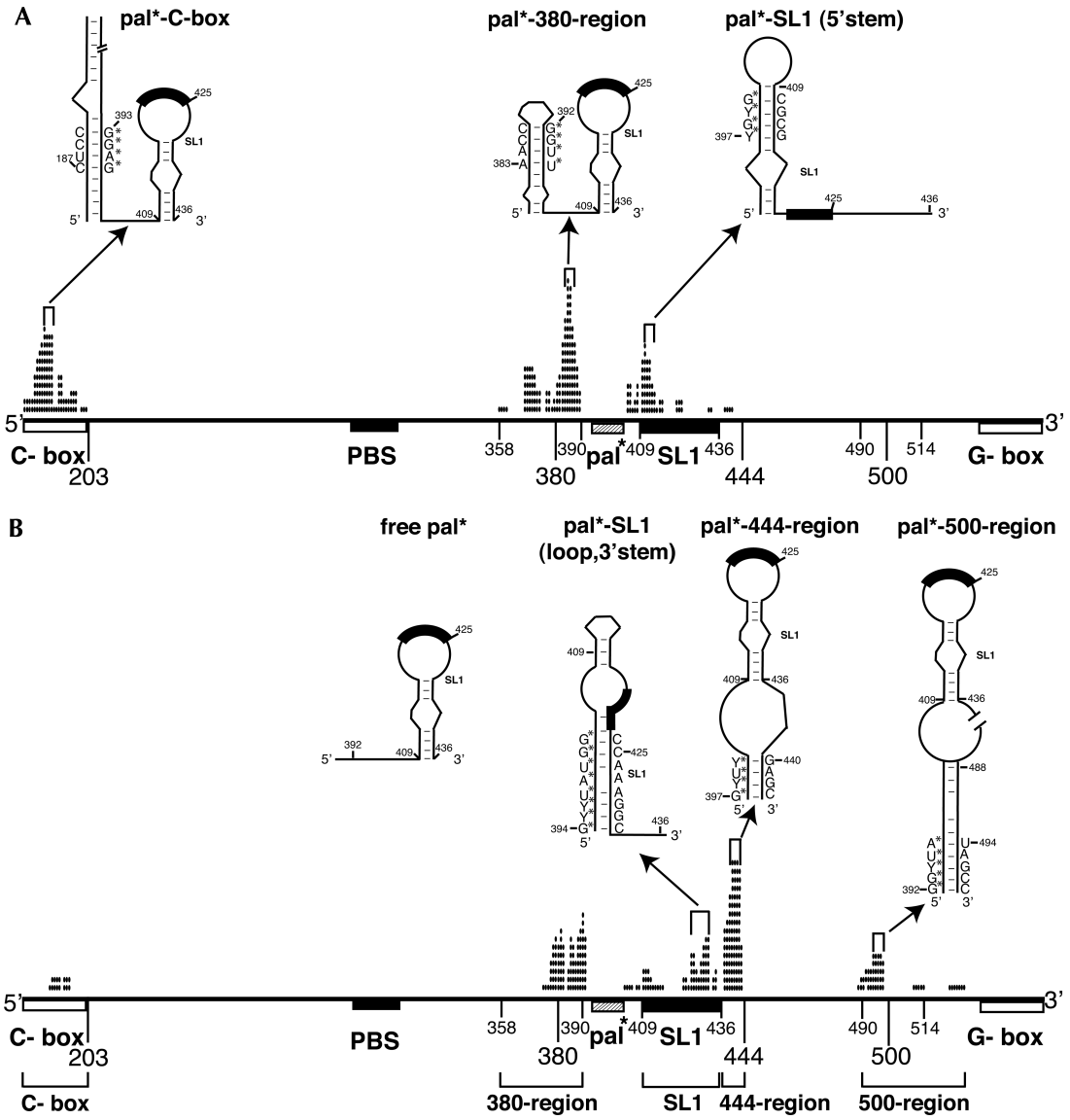

FIGURE 4. Summary of the base-pairing interactions of in vitro selected pal* in the predicted optimal 1–561 RNA structures from the last dimer pool (panel *A*) and monomer pool (panel *B*) individual sequences. The secondary structures of the selected RNAs from the dimer and monomer pools were predicted using the Mfold program, v2.3 at 55°C (Zuker 2003). The base-pairing partners of pal* (nt 392–401) were taken from the most stable predicted structure for each of the 120 cloned sequences. Nucleotides predicted to interact with pal* in all of the most stable secondary structures are indicated by dots *above* the leader region schematic (i.e., one dot equals one base-pair with one pal* nucleotide). The high peaks of dots represent the major partners of pal*, and include the C-box (nt 186–199), the 380-region (nt 358–390), SL1 (nt 409–436), the 444 region (nt 437–444), and the 500 region (nt 490–514). The major pal* binding regions for both selections are indicated with brackets *below* panel *B*. Secondary structure of pal*, SL1, and pal*-interacting partners for the structural groups are indicated *above* the dots schematic in both panels. Only the pal* nucleotides that are consistently base-paired in each structural group are indicated. Y indicates a pyrimidine residue (C or U).

structural groups and dimerization phenotypes as the one we observed for the 120 sequenced individual clones (Table 2, see below). Moreover, D07 and D08 belong to the pal*-SL1 group, a group represented equally in both monomer and dimer pools (20%) (Table 3, see below), reinforcing the hypothesis that these were selected in the dimer pool as heterodimers or that they emerged through the selection despite their marginal dimerization capacity (see above).

Wild-type 1–561 HIV-2 RNA solution structure probing

To examine the conformation(s) of the wild-type HIV-2 leader RNA in solution, we performed chemical probing of the accessibility (reactivity) of the adenine residues in dimer buffer at 55°C over time (0, 4, or 20 min) (Fig. 6A). Since the level of tight dimers is low over

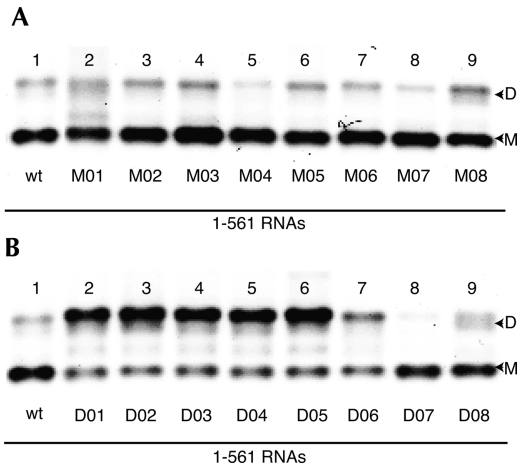

FIGURE 5. Characterization of selected RNAs from each of the last dimer and monomer pools. Eight RNAs from each pool were selected to characterize their behavior in standard dimerization assays. (*A*) Clonal RNAs from the final monomer pool are represented by M01–M08. Lanes *2–9* show the dimerization-impaired RNAs, and lane *1* shows the wild-type RNA. (*B*) Clonal RNAs from the final dimer pool are represented by D01-D08. Lanes *2–7* show the dimer-competent RNAs (D01–06). Lanes *8* and *9* show the dimerization-impaired D07 and D08 RNAs that persisted during the selection for dimerization-competent RNAs, and lane *1* represents the wt RNA.

the first 20 min of folding at 55°C (Lanchy et al. 2003a), this experiment was designed to investigate the conformation of the monomeric RNA species together with the SL1-mediated kissing loop dimer species, which can also form under these conditions. Because solution structure probing yields data on the average accessibility of each nucleotide, the readout potentially represents the superimposed reactivities of several distinct conformations if more than one conformation exists in the reaction.

Indeed, most adenine residues of the pal and SL1 elements are rather accessible to DMS methylation (on the N1 atom of their Watson–Crick side) (Fig. 6; for review, see Ehresmann et al. 1987). The most reactive adenine is A427 in SL1's loop, and the most unreactive is A387 located upstream of pal (Fig. 6). Furthermore, the A423 residue in SL1's palindromic sequence is weakly reactive, suggesting that a majority of RNA molecules may be involved in a "kissing" SL1 loop–loop interaction (loose dimers), since we know that only a small minority form tight dimers after 20 min of incubation (see Discussion and Lanchy et al. 2003a).

Finally, comparison of the change of reactivities over time (Fig. 6A, lanes 2–4) indicated that the structure involving the pal and SL1 elements formed fast and did not change drastically during the first 20 min of folding, except that the A387 residue became unreactive after 4 min (Fig. 6A). The chemical structure probing results are consistent with the extended SL1 model (Fig. 6B).

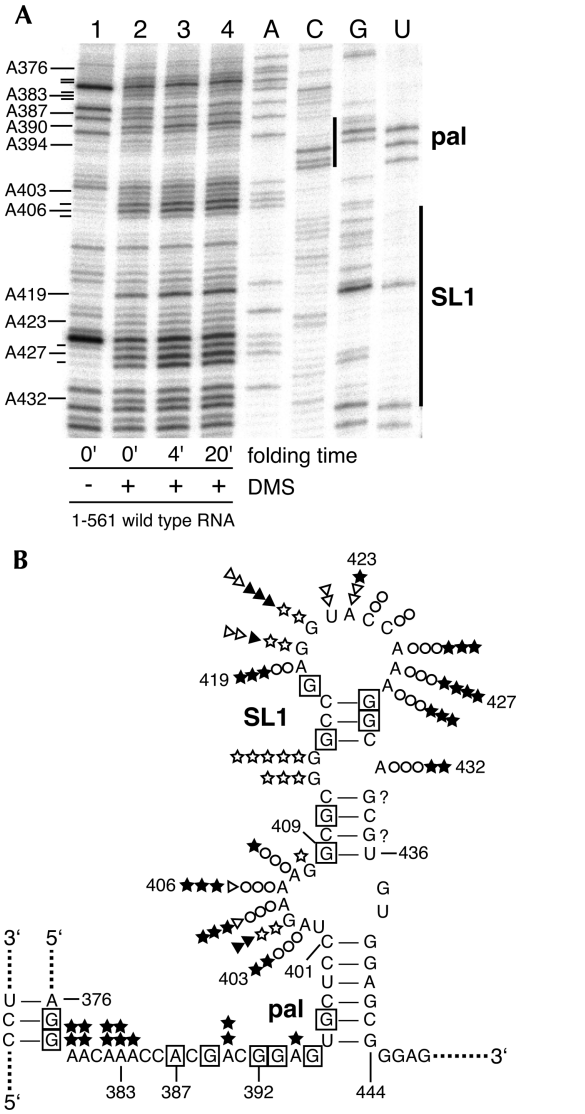

FIGURE 6. DMS probing of the pal and SL1 elements in the wild-type HIV-2 leader RNA. (*A*) Wild-type 1–561 RNA was denatured, then incubated in the presence of dimer buffer at 55°C for the times indicated. DMS was then added (0.5% final concentration), and incubation was continued at 27°C for 2 min. Reactions were stopped, and primer extensions were performed to visualize the sites and the intensities of adenine residues modifications as described in Materials and Methods (lanes *1–4*). Lanes *A*, *C*, *G*, *U* represent the DNA sequencing of the plasmid DNA used to synthesize the RNA. The pal (nt 392–401) and SL1 (nt 409–436) regions are indicated by vertical lines along the DNA sequencing lanes. (*B*) Summary of the principal nucleotide reactivities superimposed on a secondary structure model of the pal and SL1 elements for experiments done in this study (closed stars) or in previous publications: open stars for kethoxal probing of the guanine residues (Lanchy et al. 2004), open triangles for the single-strand-specific T1 nuclease (Damgaard et al. 1998), and open circles and closed triangles for DMS and T1 nuclease, respectively (Dirac et al. 2001). The boxed nucleotides indicate a lack of reactivity of the indicated adenines (this work) or guanines (Lanchy et al. 2004). The number of stars, triangles, or circles juxtaposed to a residue represent the relative levels of reactivity of this residue vis-à-vis the probing reagent (the stronger the reactivity, the greater the number of symbols).

DISCUSSION

Among the compact genomes of the RNA viruses, the viral RNA encodes not only viral proteins but also noncoding structural motifs that help to direct and regulate replicative functions. Because these regulatory entities need to be active only transiently, we and others have hypothesized that structural rearrangements of noncoding RNA lead to the alternate presentation and sequestration of signals in the HIV 5′ untranslated leader region (Berkhout and van Wamel 2000; Huthoff and Berkhout 2001; Dirac et al. 2002a; Lanchy et al. 2003a,b). We previously presented evidence that pal can interact with SL1 to inhibit HIV-2 RNA dimerization in vitro, but details of the interaction were unknown (Lanchy et al. 2003a). Because of the high degree of conservation of nucleotides in the region, comparative analysis of HIV-2 and SIV sequences was of little use. In vitro evolution of RNA molecules provides a convenient way to generate an artificial phylogeny of functionally similar molecules (Joyce 1989; Ellington and Szostak 1990; Tuerk and Gold 1990).

By randomizing pal and selecting for molecules that were either dimerization-competent or -impaired, we could examine the ability of the pal region to regulate the SL1 dimerization signal without making a priori assumptions about structures involved. During the course of the study, we evolved pools of RNAs that had two distinct dimerization phenotypes. Mfold-mediated optimal secondary structure folding of 120 individual RNAs from these pools displayed patterns of intramolecular interactions with predictive value with respect to dimerization properties, at the level of both selection pools and individual clonal RNAs. In addition, similarities in the dimerization phenotype and predicted folding of the monomer pool RNAs with wild-type RNA suggest a general model for pal's involvement in the regulation of SL1-mediated HIV-2 RNA dimerization in vitro.

Analysis of the lowest energy predicted folding structures for members of each evolved population provided general as well as specific insight into potential regulatory base-pairing interactions. The observation that most of the pal* interactions were local upstream and downstream base-pairing was striking and in line with very recent observations of essential interactions of the pal in viruses in cell culture (see below). Although such predominance of local base-pairing structures could not have been predicted, on theoretical grounds it is satisfying. To avoid large thermodynamic barriers against switching between conformations, it is likely that regulatory conformational rearrangements are nearly isoenergetic and that switching between conformations should involve rearrangement of a limited number of base-pairs; such criteria are most easily satisfied when base-pairing changes are localized to a given domain. Furthermore, in the original population of randomized pal variants, the potential existed for pal* to hybridize anywhere along the leader region. The fact that discrete and predominantly local base-pairing arrangements were found after selection suggests that either other arrangements were irrelevant to the selection or that the native folding of the 1–561 RNA limited the access of other regions of the leader RNA for hybridization with pal*, for instance, the intrinsically stable TAR and PBS domains (encompassing nt 1–123 and 197–379, respectively) (Berkhout and Schoneveld 1993).

Although most pal* base-pairing involved short-range interactions with partners located less than 45 nt upstream and downstream (Fig. 4), two exceptions were observed in both dimer and monomer pools. The first example is represented by 27% of the dimer pool sequences (pal*–C-box) (Table 3, see below), where pal* was predicted to hybridize with the so-called C-box (nt 186–199) (Fig. 4A; Table 2, see below), an element located between the poly A signal and PBS domains that we previously showed is capable of modulating dimerization through interactions with the G-box located at the 3′ end of the leader RNA (Lanchy et al. 2003b). This result corroborated previous observations because pal* hybridizing to the C-box had the same effect on dimerization as mutagenesis of the C-box or hybridization of a small complementary oligonucleotide. We hypothesize that binding of pal* to the C-box increased the accessibility of SL1 for intermolecular interactions, thus explaining the enrichment of C-box-complementary pal* sequences (Table 2, see below) in the dimer pool selection (27% vs. 3% in monomer pool) (Table 3, see below).

The other example of a nonlocal base-pairing interaction came from a structural group comprising 12% of the final monomer pool sequences (pal*-500-region) (Table 3, see below), where pal* was predicted to interact with the 500 region (nt 490–514) (Fig. 4B; Table 2, see below). Although they are slightly different, we hypothesize that the structure centered around SL1 in pal*–500 region clones is functionally similar to the one in pal*–444 region clones with regard to their effect on SL1 presentation (Fig. 4B). Indeed, the 445–485 region folds into two stable and conserved stem–loop structures (ψ1 and the major splice donor site stem–loop) (Dirac et al. 2002b) that, together with the pal*–500 region, form a constrained domain structurally extending SL1 in a way similar to the pal*–444 region fold. Such a constrained SL1 may impair the transition between monomers and tight dimers (see below), explaining the presence of the pal*–500 region selected RNAs in the monomer pool (12% versus none in dimer pool) (Table 3, see below).

The precise secondary structure(s) of the SL1 domain have not been definitively established for HIV-2 leader RNA, compared with the homologous region in HIV-1 (see below and Shen et al. 2000, 2001). Several secondary structure models have been proposed, some of which correspond to the structures predicted in this work. First, wild-type pal may interact with upstream nt 376–386 and form a stem–loop located between the PBS domain and

SL1, similar to the pal*-380-region structure shown in Figure 4A. This structure was proposed to form or to be part of a strong encapsidation signal for the HIV-2 genomic RNA (Griffin et al. 2001). Second, wild-type pal may interact with part of SL1, similar to the pal*–SL1 structures described in Figure 4A. The altered presentation of SL1's palindromic sequence may lead to a decrease of its dimerization capacity (Lanchy et al. 2003b). However, our present study using pal*–randomized RNAs found that the pal interaction with the 5′ side of SL1 was not frequently detected and, surprisingly, was slightly more represented in the final dimer pool (13% and 7% in dimer and monomer pools, respectively) (Table 3, see below). Third, the 3′ end of wild-type pal may interact with a 5′ GGAGC motif located downstream of SL1 (Fig. 6B; McCann and Lever 1997; Dirac et al. 2001) and form a structure similar to the one found in pal*–444 region species, which we called the extended SL1 (Lanchy and Lodmell 2007).

Two interesting outcomes of this study are that the dimerization-impaired population is enriched with RNA species (30% of the monomer pool sequences) (Table 3, see below) showing a predicted, extended SL1 structure (i.e., the pal*-444-region) (Fig. 4B) and that the wild-type HIV-2 pal and SL1 elements may fold into a similar structure (Fig. 6B). Wild-type RNA, in line with this observation, forms tight dimers rather inefficiently, to an extent similar to the yields of monomer pool RNA species (Figs. 3B, 5A), yet SL1 in the extended SL1 structure seems clearly presented as an individualized stem–loop with its palindromic sequence in the apical loop (Fig. 6B). This apparent paradox between a presentation of SL1 apparently optimal for dimerization and an impaired dimerization phenotype may be explained if one considers both the mechanism of SL1-mediated tight dimerization and the relative stabilities of the SL1 structures (Fig. 7). Although the mechanism of the transition between monomers and tight dimers is still a matter of discussion, studies using a 39-mer HIV-1 SL1 or larger leader RNA fragments revealed that this transition could be regulated by the stability of the stem of SL1 during the transition from kissing loops dimers to tight dimers in vitro (Muriaux et al. 1996b; Laughrea et al. 1999; Shubsda et al. 1999; Takahashi et al. 2000; Baba et al. 2001; Huthoff and Berkhout 2002; Bernacchi et al. 2005; Mujeeb et al. 2007). Accordingly, the increased stability of extended (−11.3 kcal/mol) versus short (−7.5 kcal/mol) HIV-2 SL1 structures could have a similar negative effect on tight dimerization efficiency. Furthermore, the extended SL1 structure is even more stable in HIV-2 than the comparable 39-mer HIV-1 SL1 (−11.3 and −6.4 kcal/mol, respectively; Mfold v2.3, 55°C; data not shown), which could increase the energy barrier of the transition from kissing to tight dimers.

Although the extended SL1 structure characterized only one of the different RNA structural models revealed during the in vitro selections, recent results suggest that this conformation of the SL1 domain is functionally important in vivo. In particular, we showed that the formation of the extended SL1 structure is necessary for HIV-2 replication and genomic RNA encapsidation (Lanchy and Lodmell 2007). Our experiments showed that one important role of the 3′ end of pal is to base-pair with the 5′-GGAGC motif located downstream of SL1 and that this interaction is necessary for an effective genomic RNA encapsidation to occur (Lanchy and Lodmell 2007). Although a strong selective pressure during viral replication for an adenine

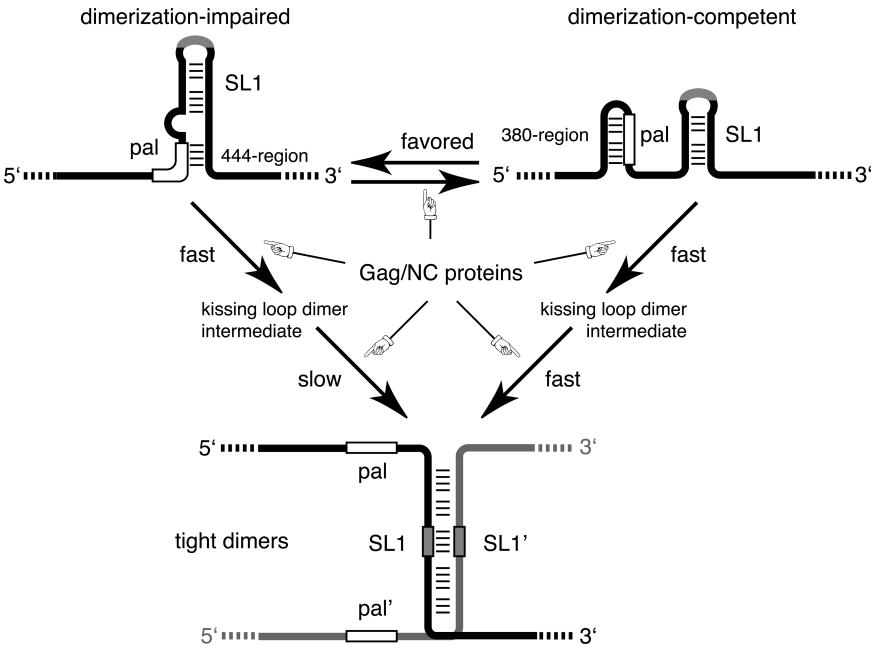

FIGURE 7. Model of pal's regulatory function in HIV-2 leader RNA SL1-mediated dimerization. Different conformations of the pal and SL1 elements and different stages of the RNA tight dimerization process are represented. The *top left* structure corresponds to the extended SL1 structure where the 3′ end of pal base-pairs to a sequence immediately downstream of SL1 (Fig. 6B; Dirac et al. 2001). The *top right* structure shows individualized intra-ψ and SL1 small stem–loop structures (Fig. 4A; Griffin et al. 2001). The dimerization-impaired phenotype for RNAs bearing extended SL1 may be due to a thermodynamically disfavored transition from kissing loop dimers to tight dimers, compared with dimerization-competent RNA species. However, one cannot rule out that a favored folding of extended SL1 may further amplify the dimerization-impaired phenotype of wild type HIV-2 leader RNA during in vitro dimerization assay. Finally, the virally encoded nucleocapsid protein NC and its Gag polyprotein precursor have RNA chaperone activity and may play a role in either the initial intramolecular folding or the different stages of SL1-mediated RNA dimerization.

at position 394 was observed upon mutation to a guanine residue (Lanchy and Lodmell 2007), the precise role of the 5′ end of pal is less understood. Nevertheless, it is notable that the extended SL1 structure is surrounded by two 5′-GGAG motifs, including one in pal (Fig. 6B). It is possible that these purine motifs might be involved in the specific RNA–protein interactions taking place during genomic RNA selection and encapsidation. Likewise, deletion mutagenesis studies suggest that an extended SL1 element is involved in encapsidation and dimerization of the SIVmac239 genomic RNA (Strappe et al. 2003; Whitney and Wainberg 2006).

In conclusion, our study shows that a few nucleotide changes in the pal element of HIV-2 leader RNA lead to altered dimerization phenotypes in vitro. The efficiency of SL1-mediated tight dimerization seems to correlate with the folding of the pal and SL1 elements, especially with the degree of structural constraint around the SL1 element. Extrapolating to wild type, our results suggest that the monomeric form of HIV-2 leader RNA adopts several local conformations that regulate the usage of SL1 as a dimerization element (Fig. 7). Factors that contribute to the transition from dimerization-impaired to dimerization-competent RNA species possibly include viral proteins binding to this dynamic region, either Gag or NC or both (Fig. 7; Clever et al. 2000; D'Souza and Summers 2005). Testing the protein–RNA interactions in this region represents our next endeavor.

MATERIALS AND METHODS

Production of pool 0 DNA template

Several PCR reactions and ligation steps were used to build the pool 0 DNA library, which was degenerate at 9 nt in the palindrome sequence (pal). First, a sense primer containing a *Bam*HI site with a T7 RNA polymerase promoter (Table 1, sBamT7R) and an anti-sense primer containing the unique *Hpy*99I site (nt 388–392) (Table 1, asHpy) were used to amplify the first 392-nt-long fragment of the HIV-2 leader region (ROD isolate, GenBank M15390). Second, a sense primer containing the *Hpy*99I site with the degenerate sequence in pal (sHpy-random) (Table 1) and an anti-sense primer containing an *Eco*R1 site (asEco561) (Table 1) were used to amplify the last 169-nt-long fragment of the HIV-2 1–561 region (PfuUltra High Fidelity DNA polymerase, Stratagene). The 10-nt pal was degenerate at only 9 nt (393–401), because the first nucleotide (G392) of pal belongs to the *Hpy*99I recognition site used in our cloning strategy and thus needed to be kept intact to get cohesive *Hpy*99I ends of the fragments (see Fig. 2A). Third, the agarose gel-purified PCR fragments were digested with Hpy99I and ligated together using the Quick ligation kit (New England Biolabs). Another gel purification was used to purify the correct ligation product, i.e., the 392-nt fragment ligated to the 169-nt fragment. Fourth, the 561-nt-long ligation product was reamplified using sBamT7R and asEco561 primers, gel-purified, and digested with *Bam*HI and *Eco*R1 to get the pool 0 DNA template.

Pool 0 RNA synthesis and purification

Pool 0 RNA was synthesized from the pool 0 DNA template with the AmpliScribe T7 transcription kit (Epicentre). After transcription, the DNA template was digested with RNase-free DNase. The RNA transcripts were purified by phenol/chloroform extraction, ammonium acetate precipitation, and ethanol precipitation followed by denaturing gel electrophoresis, excision, and extraction of the 1–561 RNA. A sample of the pool 0 DNA template was sequenced to verify degeneracy of the pal region (data not shown).

Selections for dimerization-competent and dimerization-impaired RNAs

The selection for dimerization-competent RNAs was carried out essentially as previously described (Lodmell et al. 2000). In this study, additional selections were made separately for dimerization-impaired RNAs. Briefly, an aliquot of the pool 0 degenerate RNA was denatured in 8 μL H_2O for 2 min at 90°C and then snap cooled on ice for 2 min. Dimerization was allowed to proceed for 30 min at 55°C after adding the dimer buffer (final concentration: 50 mM Tris-HCl at pH 7.5 at 37°C, 300 mM KCl, and 5 mM $MgCl_2$). Following 30 min of incubation, samples were loaded with 2 μL of 6× glycerol loading dye (40% glycerol, 44 mM Tris-borate at pH 8.3, 0.25% Bromophenol blue) onto a 0.8% agarose gel. Electrophoresis was carried out for 90 min at room temperature (24°C) in 44 mM Tris-borate (pH 8.3) and 1 mM EDTA (0.5× TBE). The ethidium bromide–stained gel was visualized on a Fuji FLA3000G Image Analyzer. The bands corresponding to the dimer and monomer RNA species were excised from the gel. To obtain the templates for further rounds of transcription and selection, a slice of each excised band was subjected to reverse transcription and polymerase chain reaction (RT-PCR) using AccuScript High Fidelity RT-PCR kit (Stratagene). The reverse transcription step was performed in a final volume of 10 μL to synthesize the cDNA from a 5 μL equivalent slice of each excised band with the asEco561 antisense primer (0.6 μM) (Table 1), dNTP mix (1 mM of each dNTP), 1× AccuScript RT buffer, and 0.5 μL AccuScript High Fidelity RT enzyme according to the

TABLE 1. Oligonucleotides used in this study

sBamT7R	5′-TAG GAT CCT AAT ACG ACT CAC TAT AGG TCG CTC TGC GGA GAG-3′
asEco561	5′-AAG AAT TCA GTT TCT CGC GCC CAT CTC CC-3′
sHpy-random	5′-CAA CCA CGA CGN NNN NNN NNT AGA AAG GCG CGG GCC GAG G-3′
asHpy	5′-AGG CAC TCC GTC GTG GTT TGT TCC TGC CGC CC-3′

The nucleotides forming the Hpy99I site are underlined (5′-CGACG-3′).

TABLE 2. Structural groups of the individual pal* sequences from monomer and dimer pool selections

Dimer pool selection																		
pal*–C-box																		
	193	T	A	C	C	C	T	C	G	T	G	T	G	T	G			
D01/45/60		G	T	A	G	G	A	G	A	G	A							
D17					G	G	A	G	A	G	A	C	A	T				
D54				G	G	G	A	G	A	G	A	T	A					
D23			G	T	G	G	A	G	A	G	A	T						
D41				G	G	G	A	G	C	A	T	A	A					
D51				G	G	G	A	G	C	A	T	T	A					
D29				G	G	G	A	G	C	A	A	T	G					
D30				G	G	G	A	G	T	A	C	A	A					
D31/38					G	G	A	G	C	A	C	A	C	A				
D56				G	G	G	A	G	C	T	A	A	C					
D04				G	G	G	A	G	A	G	A	T	A					
D39						G	T	G	C	A	C	T	T	A	G			
D03					G	G	G	G	C	G	T	T	T	A				
pal*–380 region																		
	392	G	C	A	G	C	A	C	C	A	A	A	C	A	A	G	G	A
D14						G	T	G	G	T	T	T	A	C	A			
D02/34						G	T	G	G	T	G	C	T	A	T			
D22						G	T	G	G	T	C	A	T	A	C			
D10/21								G	G	T	T	T	G	C	T	T	A	
D19								G	G	T	T	G	G	A	T	G	C	
D44								G	G	T	C	A	C	G	T	A	C	
D42			G	C	G	T	T	G	G	T	T	A						
D06				G	G	A	A	G	G	T	T	A	T					
D40					G	A	T	G	G	T	T	C	G	T				
D58						G	A	G	G	T	T	C	A	G	T			
D50				G	C	C	T	G	G	T	T	T	T					
D18		G	A	T	C	A	G	G	G	T	T							
D57	G	C	G	A	A	T	G	T	G	T								
D12						G	T	C	G	T	G	G	A	A	T			
D16				G	T	A	T	C	G	T	A	T	G					
D05		G	G	A	T	G	A	C	G	T	A							
D48				G	A	T	G	C	G	T	G	T	A					
D59									G	G	A	C	T	A	T	C	A	A
D35									G	G	G	A	T	G	T	T	A	A
D46/47						G	C	C	T	A	C	A	A	C	T			
D25						G	C	C	C	A	C	A	T	A	G			
D09/11						G	C	C	G	T	C	T	T	C	G			
D53				G	T	G	C	T	C	A	A	A	G					

Monomer pool selection																						
pal*–C-box																						
	193	T	A	C	C	C	T	C	G	T	G	T	G	T	G							
M18						G	A	G	T	T	G	G	G	C	T							
M33						G	A	G	T	T	T	G	G	G	A							
pal*–380 region																						
	392	G	C	A	G	C	A	C	C	A	A	A	A	C	A	A						
M25							G	G	G	T	A	C	T	T	C	T						
M28				G	T	C	C	G	G	T	A	T	G									
M53						G	T	G	G	A	T	G	T	A	C							
M20		G	A	T	C	C	T	G	G	A	T											
M56					G	A	T	C	G	T	G	A	A	T								
M32		G	C	A	A	A	T	C	G	T	C											
M01			G	G	C	A	T	C	G	A	G	T										
M51			G	A	G	T	A	T	G	C	A	G										
M58			G	G	A	A	C	G	G	T	T	G										
M23/43			G	T	T	T	T	T	A	A	T	G										
M16			G	T	G	T	G	A	T	C	T	A										
M12			G	G	G	T	G	T	T	A	A	T										
pal*–SL1 (5′ stem)																						
	419	A	G	C	C	G	G	G	C	G	C	G	G	A	A	A	G					
M07		G	G	T	C	C	A	T	G	T	C											
M26							G	C	G	C	A	C	A	T	T	T						
M29		G	A	C	T	A	T	G	C	A	T											
M50								G	G	C	T	T	C	T	C	T	T					
pal*–SL1 (loop, 3′ stem)																						
	434	C	G	A	C	G	G	A	A	A	A	C	C	A								
M04/15		G	A	G	T	T	T	A	T	G	G											
M54				G	G	T	C	A	A	T	G	G	T									
M11				G	G	C	C	T	T	A	A	G	G									
M13			G	C	G	C	T	C	G	A	T	A										
M55					G	C	A	G	A	C	G	G	A	T								
M08					G	C	G	T	G	C	A	C	G	T								
M14					G	G	C	C	T	T	G	G	T	A								
pal*–444 region																						
	451	G	A	G	G	A	G	G	C	G	A	G	G	T	G	T	G	C		G	A	C
M10										G	C	T	T	T	G	T	T	A	G			
M40										G	C	T	C	T	T	T	G	A	G			
M21										G	T	T	T	G	A	C	A	A	T			

(*continued*)

TABLE 2. *Continued*

		pal*–SL1 (5′ stem)																			
	419	A	G	C	C	G	G	G	C	G	C	G	G	A	A	A	G	A	T	C
D49									G	C	G	A	G	G	A	T	T	T		
D13									G	C	G	C	A	A	C	T	T	T		
D37									G	C	G	T	C	C	T	T	C	G		
D24		G	A	T	T	T	T	T	G	C	G									
D28				G	G	A	C	T	G	T	G	T	A							
D07				G	G	T	C	C	T	T	G	C	A							
D36/43											G	C	C	T	G	G	T	T	T	T
		pal*–SL1 (loop, 3′ stem)																			
	434	C	G	A	C	G	G	A	A	A	C	C									
D32/33			G	G	A	A	A	T	G	T	G	G									
D55		G	T	G	A	G	A	G	A	A	T										
D08		G	C	G	A	T	C	T	T	G	A										
		pal*–444 region																			
	451	G	A	G	G	A	G	G	G	C	G	A	G	G	T	G	T	G	C	G
D26								G	G	A	C	T	C	C	T	T	T			
		pal*–500 region																			
	498	C	C	G	A	T	G	T	C	A	A										
		No sequence																			
		pal*-none (free pal*)																			
		no target																			
D20		G	A	G	A	T	T	T	A	G	C										
D27		G	A	T	T	T	T	A	A	A	C										
D15		G	A	A	C	A	A	G	A	G	A										
D52		G	A	T	G	A	G	T	A	A	A										

M24														G	T	T	T	C	A	T	T	T	A		
M27						G	C	A	T	T	G	C	T	C	T										
M05						G	A	C	G	A	G	C	T	T	A										
M45		G	T	T	C	A	C	A	T	G	C														
M52		G	A	C	G	T	G	T	A	G	C														
M47											G	T	A	C	T	A	C	C	T	T				
M41											G	C	A	G	A	A	G	C	A	T				
M02											G	A	G	C	T	A	C	A	A	A				
M38/42											G	A	C	C	T	C	C	A	C	A				
M48						G	A	G	T	C	C	A	T	T	C										
M37											G	A	A	T	C	C	A	T	T	G				
M57													G	G	C	C	A	T	T	T	A	T		
M34/39											G	C	A	A	G	T	A	C	G	T				
		pal*–500 region																			
	498	C	C	G	A	T	G	T	C	A	A										
M06		G	G	T	T	A	C	A	A	T	T										
M44		G	G	T	T	A	C	C	C	A	G										
M22/31		G	G	C	T	A	A	G	T	T	C										
M36		G	A	G	T	C	A	T	A	G	T										
M03		G	T	T	T	C	G	G	T	A	G										
M59		G	G	A	T	A	T	C	T	G	T										
		pal*–none (free pal*)																			
		no target																			
M30		G	A	C	T	G	A	A	A	C	G										
M35		G	A	A	T	C	G	G	C	A	A										
M19/46		G	A	A	G	C	G	A	T	A	G										
M49		G	A	G	C	A	T	T	A	C	A										
M17		G	T	A	A	C	T	T	C	C	C										
M60		G	G	A	A	G	T	A	A	G	T										
M09		G	A	T	T	A	T	A	A	T	G										

M01–M60 and D01–D60 represent the selected pal* sequence of the 120 clones isolated from the final dimer (left panel) and monomer (right panel) pool selections written in a 5′ to 3′ orientation. The pal* sequence also includes the constant G392. The optimal (i.e., most stable) predicted secondary structures of the 120 individual clones were visually analyzed, and the organization of the pal* and SL1 elements was recorded, allowing segregation of the 120 pal* sequences into distinct structural groups as described in Materials and Methods. Duplicated or triplicated sequences are represented by only one row (e.g., D01/45/60). The sequence of the target region is indicated below the name of each structural group together with the number of the 3′ nucleotide number, in a 3′ to 5′ orientation. The names of the individual clonal RNAs (M01–8 and D01–8) tested for dimerization are bolded. In order to highlight base-pairing consensus in each structural group, pal* nucleotides base-paired to target nucleotides in the predicted structures are aligned relative to the indicated target sequence. For example, a majority of sequences in the dimer pool's pal*-380-region group showed a 5′-GGTT motif base-paired to a target 3′-386-CCAA motif.

manufacturer's protocol. The reverse transcription reaction was incubated for 30 min at 42°C. The PCR step was performed in a final volume of 50 μL containing 4 μL of the cDNA template, 0.2 μM of sBamT7R and asEco561 primers (Table 1), dNTP mix (0.2 mM of each dNTP), 1× PCR buffer, and 2.5 units of PfuUltra High Fidelity DNA polymerase (Stratagene). Forty cycles of amplification with a 55°C annealing temperature were used. The resulting RT-PCR products were digested with BamHI and EcoR1 prior to transcription or cloning. We carried out five and six successive rounds of selections for dimerization-impaired and dimerization-competent RNAs, respectively (see Fig. 2B). For sequence analysis, the digested RT-PCR products of the final rounds for both selections were ligated to the BamHI and EcoRI sites of the pUC18 plasmid. The ligation reaction was transformed into competent DH5α *Escherichia coli* bacteria. Sixty individual DNA clones from each selection were obtained by DNA isolation from the bacterial colonies and sequenced.

Prediction of secondary structures

Mfold version 2.3 (Zuker 2003) was used to predict the most stable secondary structure for 1–561 wild-type HIV-2 ROD and the 120 individual clonal 1–561 RNA sequences isolated from the final dimer (seventh) and monomer (sixth) pools. The software used is found on the Mfold server (http://frontend.bioinfo.rpi.edu/applications/mfold/cgi-bin/rna-form1-2.3.cgi) (Zuker 2003). The folding temperature was set at 55°C since the RNAs were incubated at 55°C in our dimerization protocol. The 120 most stable (i.e., optimal) secondary structures were visually analyzed, and the organization of the pal* and SL1 elements and the base-pairing partners of the selected pal* nucleotides (nt 392–401) were recorded, allowing segregation of the 120 pal* sequences into distinct structural groups (see Fig. 4; Tables 2, 3).

Synthesis and purification of selected RNAs

Eight clones out of the 60 individual DNAs from each final dimer and monomer pools were chosen for characterization of dimerization properties. The plasmids were linearized with EcoRI prior to in vitro transcription. RNAs were synthesized with AmpliScribe T7 transcription kit (Epicentre). After synthesis, RNAs were purified by ammonium acetate precipitation and size exclusion chromatography (Bio-Gel P-4, Bio-Rad), followed by ethanol precipitation.

In vitro dimerization of selected RNAs

The dimerization efficiency of the 16 individual clonal RNAs was checked by nondenaturing gel electrophoresis as described for the in vitro selection of dimer and monomer pools.

RNA solution structure probing

In a standard experiment, 5 pmol of 1–561 wild-type HIV-2 genomic RNA was heated in water for 2 min at 90°C and then quench-cooled on ice. The folding was started by the addition of dimer buffer. After 0-, 4-, or 20-min incubation at 55°C, dimethylsulfate (DMS, Aldrich) was added to the RNA and the mixture incubated 2 min at room temperature (27°C). The final concentration of DMS was 0.5%. Chemical probing was stopped by the addition of 10 μg glycogen, EDTA (2 mM, final concentration), and sodium acetate (300 mM, final concentration), followed by two and a half volumes of absolute ethanol. After a 30-min precipitation at –25°C, the samples were pelleted by centrifugation at 15,000 rpm for 30 min, ethanol washed, vacuum dried, and resuspended in water. One fourth of the resuspended material was then reverse-transcribed with the avian myeloblastosis virus reverse transcriptase (Promega) and a 5′ end ^{32}P-labeled radioactive DNA oligonucleotide primer (complementary to nt 470–494). After a 30-min primer extension at 42°C, the samples were precipitated as described above, and the dried pellet was resuspended in formamide loading and tracking dye and analyzed by denaturing polyacrylamide gel electrophoresis. A DNA sequencing of the plasmid DNA used for the transcription was loaded together with the primer extension to identify the modified nucleotides.

TABLE 3. Proportional representation of each structural group in the final dimer and monomer pools

	Pool selection (%)	
Structural groups	Dimer	Monomer
pal*–C-box	27	3
pal*–380 region	45	22
pal*–SL1		
SL1: 5′ stem	13	7
SL1: loop, 3′ stem	7	13
pal*–444 region	1	30
pal*–500 region	0	12
pal*–none (free pal*)	7	13
Total	100	100

The percentages are obtained from the inventory of the structural organization of the pal* and SL1 elements in the Mfold-predicted optimal structures of the individual monomer and dimer pools selected sequences (60 each; see Table 2 and Materials and Methods for details).

ACKNOWLEDGMENTS

We thank Dr. Michael Zuker for consultation using the several versions of Mfold. This work is funded by the National Institutes of Health grant AI45388 to J.S.L. The plasmid pROD10 used as the original template to clone the wt 1-561 HIV-2 sequence (Lanchy and Lodmell 2002) was provided by the EU Programme EVA/MRC Centralised Facility for AIDS Reagents, NIBSC, UK (Grant Number QLK2-CT-1999-00609 and GP828102).

Received January 26, 2007; accepted May 15, 2007.

REFERENCES

Abbink, T.E. and Berkhout, B. 2003. A novel long distance base-pairing interaction in human immunodeficiency virus type 1 RNA occludes the Gag start codon. *J. Biol. Chem.* **278:** 11601–11611.

Abbink, T.E., Ooms, M., Haasnoot, P.C., and Berkhout, B. 2005. The HIV-1 leader RNA conformational switch regulates RNA dimerization but does not regulate mRNA translation. *Biochemistry* **44:** 9058–9066.

Baba, S., Takahashi, K., Nomura, Y., Noguchi, S., Koyanagi, Y., Yamamoto, N., Takaku, H., and Kawai, G. 2001. Conformational

change of dimerization initiation site of HIV-1 genomic RNA by NCp7 or heat treatment. *Nucleic Acids Res. Suppl.* 155–156.
Baudin, F., Marquet, R., Isel, C., Darlix, J.L., Ehresmann, B., and Ehresmann, C. 1993. Functional sites in the 5′ region of human immunodeficiency virus type 1 RNA form defined structural domains. *J. Mol. Biol.* **229:** 382–397.
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.
Berkhout, B. and Schoneveld, I. 1993. Secondary structure of the HIV-2 leader RNA comprising the tRNA-primer binding site. *Nucleic Acids Res.* **21:** 1171–1178.
Berkhout, B. and van Wamel, J.L. 2000. The leader of the HIV-1 RNA genome forms a compactly folded tertiary structure. *RNA* **6:** 282–295.
Bernacchi, S., Ennifar, E., Toth, K., Walter, P., Langowski, J., and Dumas, P. 2005. Mechanism of hairpin-duplex conversion for the HIV-1 dimerization initiation site. *J. Biol. Chem.* **280:** 40112–40121.
Cheung, K.S., Smith, R.E., Stone, M.P., and Joklik, W.K. 1972. Comparison of immature (rapid harvest) and mature Rous sarcoma virus particles. *Virology* **50:** 851–864.
Chin, M.P., Rhodes, T.D., Chen, J., Fu, W., and Hu, W.S. 2005. Identification of a major restriction in HIV-1 intersubtype recombination. *Proc. Natl. Acad. Sci.* **102:** 9002–9007.
Clever, J.L., Taplitz, R.A., Lochrie, M.A., Polisky, B., and Parslow, T.G. 2000. A heterologous, high-affinity RNA ligand for human immunodeficiency virus Gag protein has RNA packaging activity. *J. Virol.* **74:** 541–546.
Damgaard, C.K., Dyhr-Mikkelsen, H., and Kjems, J. 1998. Mapping the RNA binding sites for human immunodeficiency virus type-1 gag and NC proteins within the complete HIV-1 and -2 untranslated leader regions. *Nucleic Acids Res.* **26:** 3667–3676.
Darlix, J.L., Gabus, C., Nugeyre, M.T., Clavel, F., and Barre-Sinoussi, F. 1990. *cis* elements and *trans*-acting factors involved in the RNA dimerization of the human immunodeficiency virus HIV-1. *J. Mol. Biol.* **216:** 689–699.
Dirac, A.M., Huthoff, H., Kjems, J., and Berkhout, B. 2001. The dimer initiation site hairpin mediates dimerization of the human immunodeficiency virus, type 2 RNA genome. *J. Biol. Chem.* **276:** 32345–32352.
Dirac, A.M., Huthoff, H., Kjems, J., and Berkhout, B. 2002a. Regulated HIV-2 RNA dimerization by means of alternative RNA conformations. *Nucleic Acids Res.* **30:** 2647–2655.
Dirac, A.M., Huthoff, H., Kjems, J., and Berkhout, B. 2002b. Requirements for RNA heterodimerization of the human immunodeficiency virus type 1 (HIV-1) and HIV-2 genomes. *J. Gen. Virol.* **83:** 2533–2542.
D'Souza, V. and Summers, M.F. 2004. Structural basis for packaging the dimeric genome of Moloney murine leukaemia virus. *Nature* **431:** 586–590.
D'Souza, V. and Summers, M.F. 2005. How retroviruses select their genomes. *Nat. Rev. Microbiol.* **3:** 643–655.
Ehresmann, C., Baudin, F., Mougel, M., Romby, P., Ebel, J.P., and Ehresmann, B. 1987. Probing the structure of RNAs in solution. *Nucleic Acids Res.* **15:** 9109–9128.
Ellington, A.D. and Szostak, J.W. 1990. In vitro selection of RNA molecules that bind specific ligands. *Nature* **346:** 818–822.
Flynn, J.A., An, W., King, S.R., and Telesnitsky, A. 2004. Nonrandom dimerization of murine leukemia virus genomic RNAs. *J. Virol.* **78:** 12129–12139.
Fu, W. and Rein, A. 1993. Maturation of dimeric viral RNA of Moloney murine leukemia virus. *J. Virol.* **67:** 5443–5449.
Fu, W., Gorelick, R.J., and Rein, A. 1994. Characterization of human immunodeficiency virus type 1 dimeric RNA from wild-type and protease-defective virions. *J. Virol.* **68:** 5013–5018.
Greatorex, J. and Lever, A. 1998. Retroviral RNA dimer linkage. *J. Gen. Virol.* **79:** 2877–2882.
Griffin, S.D., Allen, J.F., and Lever, A.M. 2001. The major human immunodeficiency virus type 2 (HIV-2) packaging signal is present on all HIV-2 RNA species: Cotranslational RNA encapsidation and limitation of Gag protein confer specificity. *J. Virol.* **75:** 12058–12069.
Hibbert, C.S., Mirro, J., and Rein, A. 2004. mRNA molecules containing murine leukemia virus packaging signals are encapsidated as dimers. *J. Virol.* **78:** 10927–10938.
Hoglund, S., Ohagen, A., Goncalves, J., Panganiban, A.T., and Gabuzda, D. 1997. Ultrastructure of HIV-1 genomic RNA. *Virology* **233:** 271–279.
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.* **87:** 1556–1560.
Huthoff, H. and Berkhout, B. 2001. Two alternating structures of the HIV-1 leader RNA. *RNA* **7:** 143–157.
Huthoff, H. and Berkhout, B. 2002. Multiple secondary structure rearrangements during HIV-1 RNA dimerization. *Biochemistry* **41:** 10439–10445.
Jetzt, A.E., Yu, H., Klarmann, G.J., Ron, Y., Preston, B.D., and Dougherty, J.P. 2000. High rate of recombination throughout the human immunodeficiency virus type 1 genome. *J. Virol.* **74:** 1234–1240.
Jossinet, F., Lodmell, J.S., Ehresmann, C., Ehresmann, B., and Marquet, R. 2001. Identification of the in vitro HIV-2/SIV RNA dimerization site reveals striking differences with HIV-1. *J. Biol. Chem.* **276:** 5598–5604.
Joyce, G.F. 1989. Amplification, mutation and selection of catalytic RNA. *Gene* **82:** 83–87.
Lanchy, J.M. and Lodmell, J.S. 2002. Alternate usage of two dimerization initiation sites in HIV-2 viral RNA in vitro. *J. Mol. Biol.* **319:** 637–648.
Lanchy, J.M. and Lodmell, J.S. 2007. An extended stem–loop 1 is necessary for human immunodeficiency virus type 2 replication and affects genomic RNA encapsidation. *J. Virol.* **81:** 3285–3292.
Lanchy, J.M., Ivanovitch, J.D., and Lodmell, J.S. 2003a. A structural linkage between the dimerization and encapsidation signals in HIV-2 leader RNA. *RNA* **9:** 1007–1018.
Lanchy, J.M., Rentz, C.A., Ivanovitch, J.D., and Lodmell, J.S. 2003b. Elements located upstream and downstream of the major splice donor site influence the ability of HIV-2 leader RNA to dimerize in vitro. *Biochemistry* **42:** 2634–2642.
Lanchy, J.M., Szafran, Q.N., and Lodmell, J.S. 2004. Splicing affects presentation of RNA dimerization signals in HIV-2 in vitro. *Nucleic Acids Res.* **32:** 4585–4595.
Laughrea, M. and Jette, L. 1994. A 19-nucleotide sequence upstream of the 5′ major splice donor is part of the dimerization domain of human immunodeficiency virus 1 genomic RNA. *Biochemistry* **33:** 13464–13474.
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., 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.
Lodmell, J.S., Ehresmann, C., Ehresmann, B., and Marquet, R. 2000. Convergence of natural and artificial evolution on an RNA loop-loop interaction: The HIV-1 dimerization initiation site. *RNA* **6:** 1267–1276.
McBride, M.S. and Panganiban, A.T. 1996. The human immunodeficiency virus type 1 encapsidation site is a multipartite RNA element composed of functional hairpin structures. *J. Virol.* **70:** 2963–2973.
McCann, E.M. and Lever, A.M. 1997. Location of *cis*-acting signals important for RNA encapsidation in the leader sequence of human immunodeficiency virus type 2. *J. Virol.* **71:** 4133–4137.
Mikkelsen, J.G., Lund, A.H., Duch, M., and Pedersen, F.S. 2000. Mutations of the kissing-loop dimerization sequence influence the site specificity of murine leukemia virus recombination in vivo. *J. Virol.* **74:** 600–610.

Mujeeb, A., Ulyanov, N.B., Georgantis, S., Smirnov, I., Chung, J., Parslow, T.G., and James, T.L. 2007. Nucleocapsid protein-mediated maturation of dimer initiation complex of full-length SL1 stem–loop of HIV-1: Sequence effects and mechanism of RNA refolding. *Nucleic Acids Res.* **35:** 2026–2034.

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.

Muriaux, D., De Rocquigny, H., Roques, B.P., and Paoletti, J. 1996a. NCp7 activates HIV-1Lai RNA dimerization by converting a transient loop-loop complex into a stable dimer. *J. Biol. Chem.* **271:** 33686–33692.

Muriaux, D., Fosse, P., and Paoletti, J. 1996b. A kissing complex together with a stable dimer is involved in the HIV-1Lai RNA dimerization process in vitro. *Biochemistry* **35:** 5075–5082.

Ooms, M., Huthoff, H., Russell, R., Liang, C., and Berkhout, B. 2004. A riboswitch regulates RNA dimerization and packaging in human immunodeficiency virus type 1 virions. *J. Virol.* **78:** 10814–10819.

Paillart, J.C., Marquet, R., Skripkin, E., Ehresmann, B., and Ehresmann, C. 1994. Mutational analysis of the bipartite dimer linkage structure of human immunodeficiency virus type 1 genomic RNA. *J. Biol. Chem.* **269:** 27486–27493.

Paillart, J.C., Dettenhofer, M., Yu, X.F., Ehresmann, C., Ehresmann, B., and Marquet, R. 2004a. First snapshots of the HIV-1 RNA structure in infected cells and in virions. *J. Biol. Chem.* **279:** 48397–48403.

Paillart, J.C., Shehu-Xhilaga, M., Marquet, R., and Mak, J. 2004b. Dimerization of retroviral RNA genomes: An inseparable pair. *Nat. Rev. Microbiol.* **2:** 461–472.

Rein, A. 1994. Retroviral RNA packaging: A review. *Arch. Virol. Suppl.* **9:** 513–522.

Russell, R.S., Liang, C., and Wainberg, M.A. 2004. Is HIV-1 RNA dimerization a prerequisite for packaging? Yes, no, probably? *Retrovirology* **1:** 23.

Sakuragi, J., Ueda, S., Iwamoto, A., and Shioda, T. 2003. Possible role of dimerization in human immunodeficiency virus type 1 genome RNA packaging. *J. Virol.* **77:** 4060–4069.

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.

Shen, N., Jette, L., Wainberg, M.A., and Laughrea, M. 2001. Role of stem B, loop B, and nucleotides next to the primer binding site and the kissing-loop domain in human immunodeficiency virus type 1 replication and genomic-RNA dimerization. *J. Virol.* **75:** 10543–10549.

Shubsda, M.F., McPike, M.P., Goodisman, J., and Dabrowiak, J.C. 1999. Monomer-dimer equilibrium constants of RNA in the dimer initiation site of human immunodeficiency virus type 1. *Biochemistry* **38:** 10147–10157.

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.* **91:** 4945–4949.

Strappe, P.M., Greatorex, J., Thomas, J., Biswas, P., McCann, E., and Lever, A.M. 2003. The packaging signal of simian immunodeficiency virus is upstream of the major splice donor at a distance from the RNA cap site similar to that of human immunodeficiency virus types 1 and 2. *J. Gen. Virol.* **84:** 2423–2430.

Stuhlmann, H. and Berg, P. 1992. Homologous recombination of copackaged retrovirus RNAs during reverse transcription. *J. Virol.* **66:** 2378–2388.

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.

Tuerk, C. and Gold, L. 1990. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249:** 505–510.

Whitney, J.B. and Wainberg, M.A. 2006. Impaired RNA incorporation and dimerization in live attenuated leader-variants of SIVmac239. *Retrovirology* **3:** 96.

Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31:** 3406–3415.