Check for updates

Retroviral RNA Dimerization: From Structure to Functions

Noé Dubois, Roland Marquet, Jean-Christophe Paillart* and Serena Bernacchi*

Architecture et Réactivité de l'ARN, UPR 9002, IBMC, CNRS, Université de Strasbourg, Strasbourg, France

The genome of the retroviruses is a dimer composed by two homologous copies of genomic RNA (gRNA) molecules of positive polarity. The dimerization process allows two gRNA molecules to be non-covalently linked together through intermolecular base-pairing. This step is critical for the viral life cycle and is highly conserved among retroviruses with the exception of spumaretroviruses. Furthermore, packaging of two gRNA copies into viral particles presents an important evolutionary advantage for immune system evasion and drug resistance. Recent studies reported RNA switches models regulating not only gRNA dimerization, but also translation and packaging, and a spatio-temporal characterization of viral gRNA dimerization within cells are now at hand. This review summarizes our current understanding on the structural features of the dimerization signals for a variety of retroviruses (HIVs, MLV, RSV, BLV, MMTV, MPMV...), the mechanisms of RNA dimer formation and functional implications in the retroviral cycle.

Keywords: retrovirus, RNA, dimerization, HIV, MuLV, structure, function
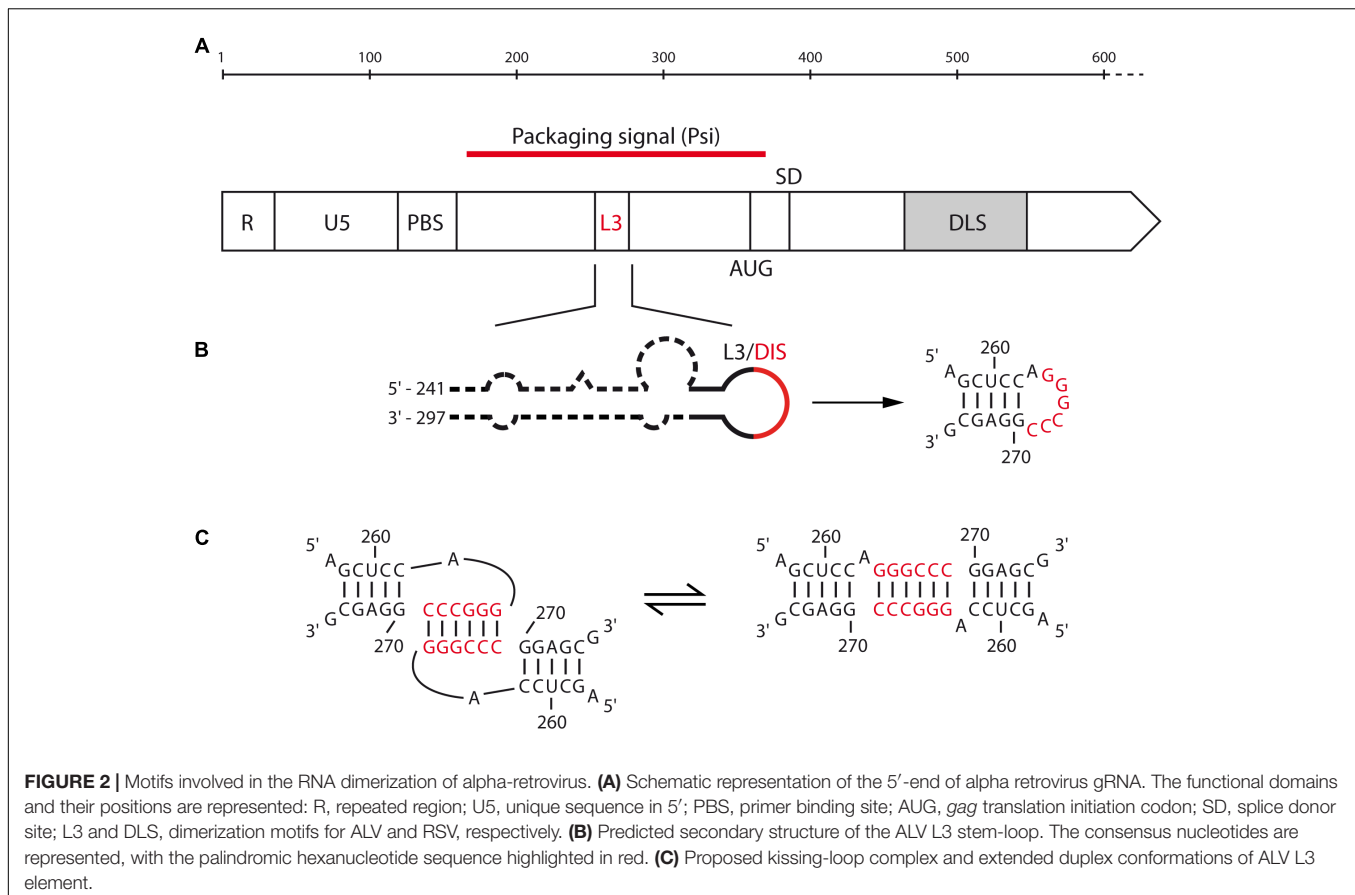
INTRODUCTION

During the late phase of their replication cycle, retroviruses package two homologous copies of their genomic RNA (gRNA) in order to produce infectious viral particles. This co-packaging of gRNA molecules is highly facilitated by the dimerization signal within the 5′-region of the viral genome and by the viral Gag precursor in order to form new viral particles (Lee et al., 1999; Cimarelli et al., 2000, for reviews see D'Souza and Summers, 2005; Mailler et al., 2016). Genome dimerization is a highly conserved process amongst retroviruses, and this feature is crucial for several important steps in the retroviral life cycle (**Figure 1**). First, for several retroviruses, e.g., human immunodeficiency virus type 1 (HIV-1) and murine leukemia virus (MuLV), gRNA dimerization is critical for selective packaging of the genome (Berkhout and van Wamel, 1996; Mougel et al., 1996; Paillart et al., 1996a; McBride and Panganiban, 1997; Mougel and Barklis, 1997; Aagaard et al., 2004; Houzet et al., 2007). Second, the conformational changes induced by RNA dimerization may also regulate translation of the unspliced gRNA (Kharytonchyk et al., 2016; Boeras et al., 2017). Third, dimerization of the viral genome also plays an important role during the reverse transcription step, allowing genome repair by strand transfer when one of the two RNA strands is damaged (Mikkelsen and Pedersen, 2000). Finally, genome dimerization presents the great advantage of increasing genetic diversity by allowing genetic recombination during reverse transcription (Mikkelsen et al., 2000; Moore et al., 2009).

FIGURE 1 | Schematic overview of the role of RNA dimerization in the retroviral life cycle. The cycle begins with the entry of the retrovirus within the target cell, followed by reverse transcription of the RNA genome into cDNA. During this step, gRNA dimerization plays an important role since RT may switch between strands, thus allowing genome repair and/or shuffling. The pre-integration complex (PIC) is then translocated into the nucleus where it is integrated in the genome of the target cell. The unspliced mRNAs are transcribed by the host machinery from the integrated provirus and transported to the cytoplasm. There, the single 5′ capped mRNAs serve as genomic RNAs that dimerize and are subsequently selected and packaged into the nascent virions, while mRNAs beginning with two or three guanosine are translated by the host machinery (Kharytonchyk et al., 2016). After budding, immature particles follow a maturation step initiated by the viral protease to produce infectious virions.

The first evidence for the existence of a dimeric genome came from ultracentrifugation sedimentation analysis of gRNA extracted Rous sarcoma virus (RSV) (Duesberg, 1968; Mangel et al., 1974), even though a tetrameric organization of the genome had also been proposed (Montagnier et al., 1969). Subsequent sedimentation and electron microscopy analyses supported the dimeric organization of the genome and extended this observation to other retroviruses families such as alpharetroviruses, gammaretroviruses, and lentiviruses (Kung et al., 1976; Bender et al., 1978; Maisel et al., 1978; Murti et al., 1981; Höglund et al., 1997), thus revealing the conservation and the importance of gRNA dimerization in the life cycle of retroviruses. Further studies showed that regions involved in gRNA dimerization, historically referred to as the dimer linkage structure (DLS), are typically close to the 5′-end of gRNA, and highly structured with multiple stem-loop motifs (Prats et al., 1990; Roy et al., 1990; Tounekti et al., 1992; Baudin et al., 1993; Mougel et al., 1993; Garzino-Demo et al., 1995; Jossinet et al., 2001; Abbink and Berkhout, 2003; D'Souza et al., 2004; Aktar et al., 2013, 2014). Even though retroviral genomes are rather large (from ∼7 to ∼12 kb), gRNA dimerization was observed to be mediated by relatively short sequences ranging from 50- to few hundreds of nucleotides (Kung et al., 1975; Bender et al., 1978; Maisel et al., 1978; Murti et al., 1981). *In vitro* analysis showed that DLS-containing RNA fragments could dimerize at temperatures ranging from 37 to 60°C depending on the virus, in the presence of monovalent (Na$^+$ or K$^+$ ranging from 0.1 to 0.3 M), and/or divalent (Mg^{2+} ranging from 1 to 10 mM) cations (Bieth et al., 1990; Darlix et al., 1990; Prats et al.,

1990; Roy et al., 1990; Marquet et al., 1991; Paoletti et al., 1993). However, a better understanding of the precise mechanisms governing the dimerization process was only obtained with the identification of the HIV-1 dimerization initiation site (DIS).

A common feature of retroviral DIS is the presence of at least one short palindromic sequence enabling intermolecular base-pairing, thus forming kissing-loop structures (Laughrea and Jetté, 1994; Paillart et al., 1994; Skripkin et al., 1994; Girard et al., 1995; Muriaux et al., 1995). In the case of HIV-1, chemical modification interference assays allowed the identification of the six nucleotides constituting the DIS (Skripkin et al., 1994). Kissing-loop complexes, often referred to as "loose dimers," are characterized by low thermal stability (Fu and Rein, 1993; Fu et al., 1994; Jalalirad and Laughrea, 2010), and can only be visualized by gel electrophoresis under native conditions, since even mild denaturing conditions were found to dissociate RNA dimers during migration (Marquet et al., 1994; Laughrea and Jetté, 1996; Polge et al., 2000). However, incubation of DLS-containing RNAs at non-physiological high temperatures (50–60°C) was found to induce formation of RNA dimers resistant to mild-denaturing electrophoresis conditions and were thus called "tight dimers." Importantly, tight dimers are also obtained at physiological temperature in the presence of the cognate nucleocapsid (NC) protein (Girard et al., 1996; Muriaux et al., 1996), well-known for its RNA chaperone properties (Feng et al., 1996; Girard et al., 1996; Muriaux et al., 1996; Rist and Marino, 2002; Mujeeb et al., 2007; Aduri et al., 2013; for a review Levin et al., 2005). These results lead to the notion that NC lowers the energy barriers and promote refolding of the 5′-end region

FIGURE 2 | Motifs involved in the RNA dimerization of alpha-retrovirus. **(A)** Schematic representation of the 5′-end of alpha retrovirus gRNA. The functional domains and their positions are represented: R, repeated region; U5, unique sequence in 5′; PBS, primer binding site; AUG, gag translation initiation codon; SD, splice donor site; L3 and DLS, dimerization motifs for ALV and RSV, respectively. **(B)** Predicted secondary structure of the ALV L3 stem-loop. The consensus nucleotides are represented, with the palindromic hexanucleotide sequence highlighted in red. **(C)** Proposed kissing-loop complex and extended duplex conformations of ALV L3 element.

of gRNA into a more stable conformation (Feng et al., 1996; Gorelick et al., 1996, 1999; Cruceanu et al., 2006). Studies on short sequences harboring the DIS suggested that these conformational changes could involve the refolding of these structures by forming cruciform intermediates that evolve into extended intermolecular base-pairing (Polge et al., 2000; Rist and Marino, 2002; Bernacchi et al., 2005).

In vivo, subsequent to viral budding, immature virions undergo a maturation step that is mandatory for viral infectivity and is mediated by the viral protease that sequentially cleaves the viral Pr55^{Gag} and Pr160^{GagPol} precursors into the mature structural and enzymatic proteins. Concomitantly to this proteolytic maturation, the viral genome undergoes a maturation process (Ohishi et al., 2011; Mailler et al., 2016). Indeed, HIV-1 and MuLV gRNA dimers extracted from immature viral particles are less stable than those extracted from mature virions (Fu and Rein, 1993; Fu et al., 1994; Jalalirad and Laughrea, 2010; Ohishi et al., 2011; Grohman et al., 2014). Interestingly, the different stabilities observed in immature and mature virions are very similar to those observed *in vitro* for loose and tight RNA dimers, respectively, suggesting these conformations may reflect the maturation process of gRNA into viral particles.

In this review, we will focus on our current understanding of the mechanisms and molecular factors involved in gRNA dimerization for different retrovirus families both *in vitro* and *in cellula*, its role during the retroviral life cycle, and finally its potential targeting by molecules aimed at inhibiting viral replication.

ALPHA-RETROVIRUSES GENOME DIMERIZATION

Even though the first evidence for retroviral RNA dimerization came from sedimentation and electron microscopy studies of alpha-retroviruses such as RSV (Duesberg, 1968; Canaani et al., 1973; Mangel et al., 1974), the precise mechanisms underlying this process remains surprisingly poorly defined in comparison with other model retroviruses such as HIV-1 and MuLV. Strikingly, while most DLS/DIS are found within the 5′-untranslated region (UTR) of gRNA, electron micrographs of gRNA dimers extracted from RSV virions located the DLS within the *gag* gene, around position 480–540 from the 5′ end (**Figure 2A**) (Murti et al., 1981). One other peculiarity lies in the fact that RSV DLS contains an imperfect palindrome that was first proposed to contribute to RNA dimerization (Schwartz et al., 1983). In rather good agreement with these findings, *in vitro* analyses of the dimerization of the first 634 nucleotides (nts) of RSV gRNA suggested that the DLS would be located between positions 544–564 (Bieth et al., 1990; Lear

et al., 1995) and would likely involve Watson–Crick base-pairing of the palindrome mentioned above (Lear et al., 1995). However, analysis of gRNA dimerization of the avian sarcoma-leucosis virus (ASLV), another alpha-retrovirus, relies on the L3 element, a conserved 19-nts hairpin harboring a perfect palindrome sequence in its apical loop and located upstream of the major Splice Donor (SD) site (**Figure 2B**) (Fossé et al., 1996). Interestingly, in the absence of NC, viral RNA fragments of 626 nts in length form loose dimers at 37°C, involving loop-loop interactions through the L3 element (**Figure 2C**, left) (Polge et al., 2000). However, when incubated at 60°C, these fragments formed tight dimers, and heterodimerization of L3 stem mutants supported the notion that *in vitro* ASLV gRNA tight dimers are characterized by the formation of an extended duplex (**Figure 2C**, right) (Polge et al., 2000).

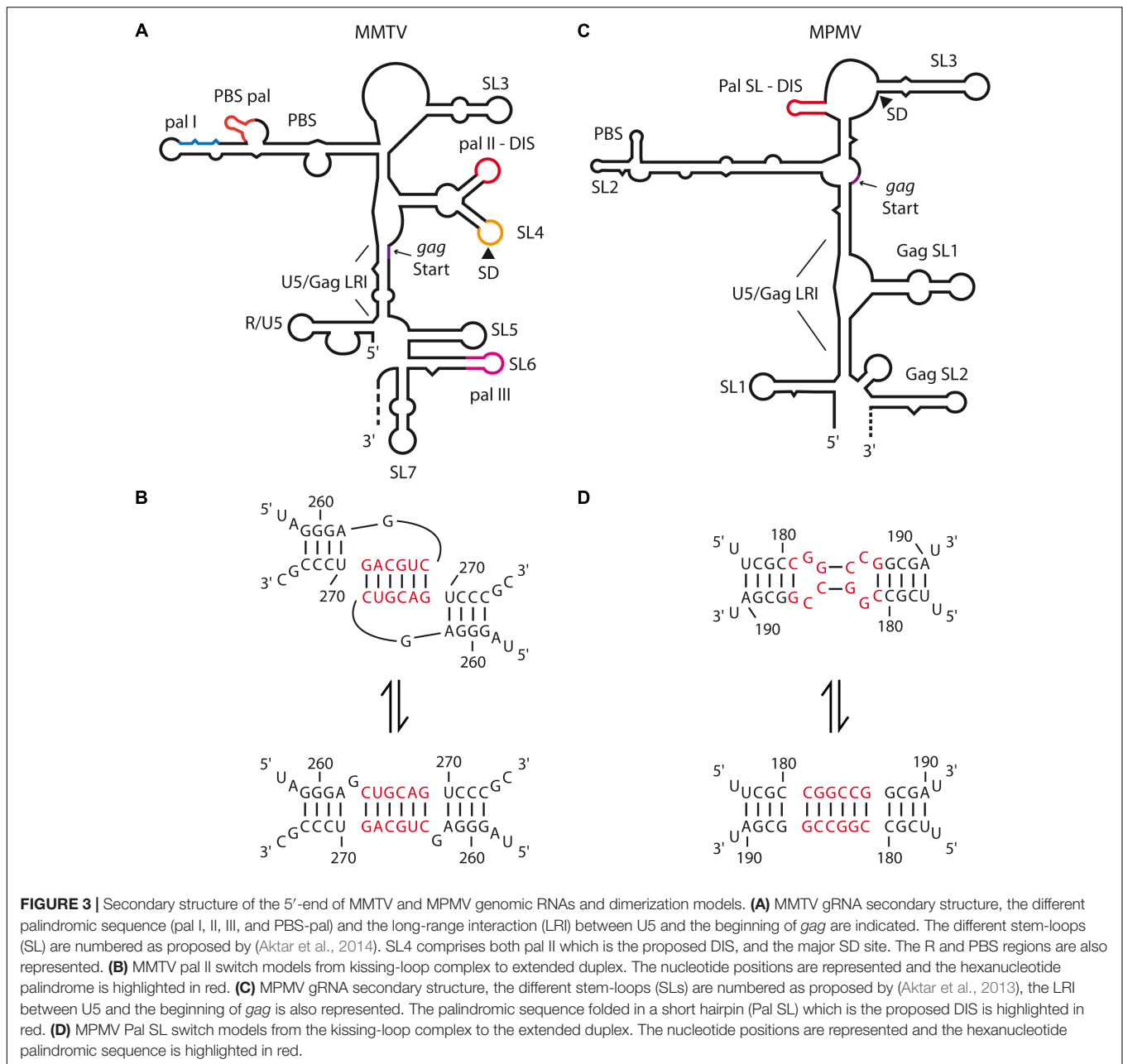## BETA-RETROVIRUSES GENOME DIMERIZATION

Recently, Rizvi and collaborators provided the first *in vitro* analyses of structure and dimerization mechanisms of the 5′-end region of the mouse mammary tumor virus (MMTV) and Mason-Pfizer monkey virus (MPMV) gRNAs (Aktar et al., 2013, 2014; Kalloush et al., 2016). Selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) and *in vitro* dimerization assays of MMTV gRNA fragments revealed that loose dimer formation is potentially mediated by two palindromic sequences, respectively, within the primer binding site (PBS-Pal) and in a bifurcated stem-loop structure (SL4) located between the PBS and the translation initiation codon of *gag* (Pal II) (**Figure 3A**). However, Pal II is the main DIS since its mutation had a greater impact on RNA dimerization than mutation of the PBS-Pal (**Figure 3B**) (Aktar et al., 2014). In MPMV, the palindromic sequence that functions as the gRNA DIS folds into a short hairpin (Pal SL) (**Figures 3C,D**) (Aktar et al., 2013). These motifs are highly conserved in MMTV and MPMV strains and their mutation leads to severe gRNA packaging and viral replication defects (Jaballah et al., 2010; Aktar et al., 2014). Interestingly, secondary structures of the 5′-end of both MPMV and MMTV gRNAs present long-range interactions (LRI) involving the 5′ unique (U5) region and a region spanning the *gag* translation initiation codon (**Figures 3A,C**) (Aktar et al., 2014; Kalloush et al., 2016), similarly to what has been observed for lentiviruses, for which such interaction was proposed to promote gRNA dimerization (see below). In MPMV, these LRIs are required for gRNA packaging and viral propagation, even though mutations destabilizing LRIs have only modest effects on RNA dimerization (Kalloush et al., 2016).

GAMMA-RETROVIRUSES GENOME DIMERIZATION

A general feature of gamma-retroviruses RNA dimerization is the presence of several palindromes within stem-loops that

contribute to dimer formation at different degrees. Within this family, murine leukemia virus (MuLV) is the most studied model, and its minimal DLS corresponds to a 170-nts long region containing four stem-loops (SL-A to SL-D) each contributing to the dimerization process (Prats et al., 1990; Roy et al., 1990; Tounekti et al., 1992; Mougel et al., 1993; Torrent et al., 1994; Girard et al., 1995; De Tapia et al., 1998; Badorrek and Weeks, 2005) (**Figure 4A**). However, comparison of the chemical reactivity between monomer and dimer also suggested that *gag* region could contribute to dimerization in a full-length genome context (Mougel et al., 1993). The first two stem-loops, SL-A and SL-B, contain palindromes of 10- and 16-nts, respectively, both critical for the *in vitro* dimerization process (**Figures 4B,C**) (Tounekti et al., 1992; Girard et al., 1995, 1996; Oroudjev et al., 1999; Ly and Parslow, 2002). Additionally, NMR and SHAPE analyses of MuLV DLS indicated that SL-C and SL-D, both presenting a highly conserved GACG apical tetraloop (Konings et al., 1992) (**Figure 4A**), form canonical heterologous loop–loop interactions that significantly contribute to dimer stability (De Tapia et al., 1998; Kim and Tinoco, 2000; D'Souza et al., 2004; Badorrek et al., 2006; Gherghe and Weeks, 2006; Miyazaki et al., 2010b) (**Figure 4B**). While MuLV initial loose dimer involves only few intermolecular base-pairs in the apical loops of SL-A and SL-B, *in vitro* analysis indicated that, when RNA fragments containing MuLV DLS are incubated at 60°C or at 37°C in the presence of NC, both contacts undergo structural rearrangements, and expand to the whole SL-A and SL-B palindromes (**Figure 4C**, right) (Roy et al., 1990; Girard et al., 1996; Ly and Parslow, 2002; Badorrek and Weeks, 2006). Interestingly, formation of the extended duplex exposes UCUG quartets that are base-paired in the monomeric structure (compare **Figures 4A** with **4C**), and that correspond to high affinity binding sites for NC protein. Furthermore, electron tomography and hydroxyl footprinting data both indicated that MuLV DLS adopts a compact structure upon dimerization (Badorrek et al., 2006; Badorrek and Weeks, 2006; Miyazaki et al., 2010b) suggesting that the stability of tight RNA dimers is also mediated by stacked helices, in addition to formation of the extended duplex.

In viro SHAPE analyses on extracted gRNA from mature MuLV particles also revealed that the gRNA dimer adopts the SL-A/SL-A′ and SL-B/SL-B′ extended duplexes conformation (Gherghe et al., 2010) (**Figure 4C**, right). Interestingly, recent chemical probing data on extracted gRNA dimers from immature MuLV particles indicated that the SL-B element would not be involved in the inter-molecular duplex (**Figure 4C**, left) (Grohman et al., 2014), thus supporting the notion of a conformational switch between immature and mature RNA dimer forms. This structural switch is likely to occur through the chaperone activity of MuLV NC, and is further facilitated by the stabilization of the duplex by SL-C and SL-D, as previously suggested (D'Souza et al., 2004; Badorrek et al., 2006; Miyazaki et al., 2010b). These data are in good agreement with findings showing that deletion of SL-A significantly decreased MuLV gRNA packaging and affected replication kinetics in cell culture, while mutation/deletion of

FIGURE 3 | Secondary structure of the 5′-end of MMTV and MPMV genomic RNAs and dimerization models. **(A)** MMTV gRNA secondary structure, the different palindromic sequence (pal I, II, III, and PBS-pal) and the long-range interaction (LRI) between U5 and the beginning of *gag* are indicated. The different stem-loops (SL) are numbered as proposed by (Aktar et al., 2014). SL4 comprises both pal II which is the proposed DIS, and the major SD site. The R and PBS regions are also represented. **(B)** MMTV pal II switch models from kissing-loop complex to extended duplex. The nucleotide positions are represented and the hexanucleotide palindrome is highlighted in red. **(C)** MPMV gRNA secondary structure, the different stem-loops (SLs) are numbered as proposed by (Aktar et al., 2013), the LRI between U5 and the beginning of *gag* is also represented. The palindromic sequence folded in a short hairpin (Pal SL) which is the proposed DIS is highlighted in red. **(D)** MPMV Pal SL switch models from the kissing-loop complex to the extended duplex. The nucleotide positions are represented and the hexanucleotide palindromic sequence is highlighted in red.

SL-B only presented moderate effect regarding these processes (Aagaard et al., 2004). Taken together, these results suggest that only the 10-nts palindromic sequence within SL-A is crucial to form the immature gRNA dimer (**Figure 4C**, left).

Finally, *in vitro* analysis of the gRNA dimer of feline endogenous RD-114, another gamma-retrovirus, revealed the presence of several palindromic stem-loops (psl-1 to psl-5) leading to RNA dimer formation (Kharytonchyk and Pedersen, 2010). Interestingly, and similarly to MuLV, the RD-114 DLS also contains two GACG tetraloops that are close to the crucial DIS elements and contribute to *in vitro* RNA dimerization (Kharytonchyk and Pedersen, 2010).

DELTA-RETROVIRUSES GENOME DIMERIZATION

Site-directed mutagenesis, antisense oligonucleotides mapping experiments, as well as structural predictions identified the DLS of bovine leukemia virus (BLV) in the 5′-region of the genome, in a region overlapping U5, the PBS and 30 bases downstream of this latter (Katoh et al., 1993). Interestingly, the DLS of BLV also overlaps the Psi (packaging signal), similarly to other retroviruses, since deletion of this region completely abrogated genome encapsidation and viral replication (Mansky et al., 1995). However, a particular feature of BLV is the fact that *in vitro* dimerization of gRNA is promoted by the viral matrix protein,
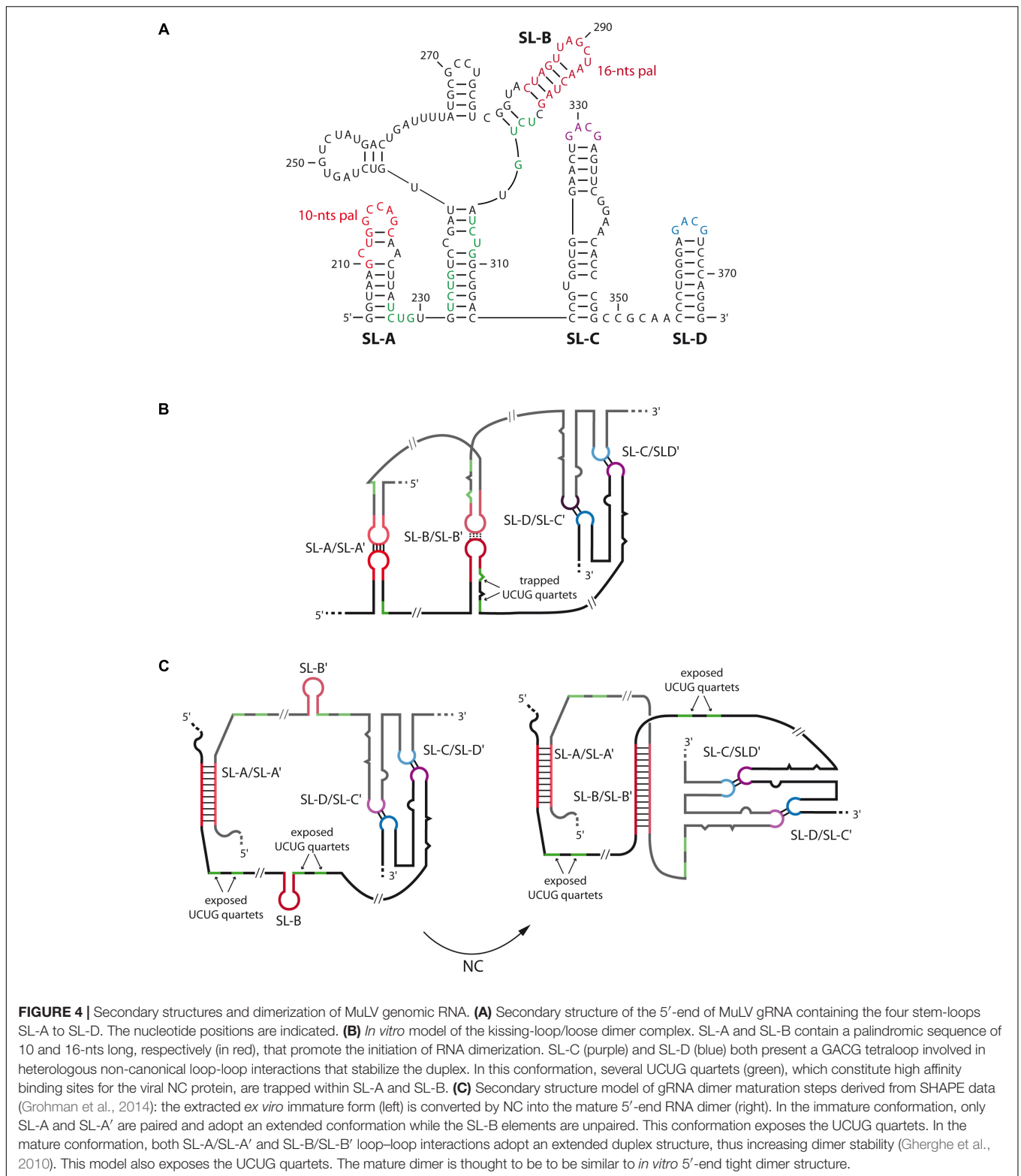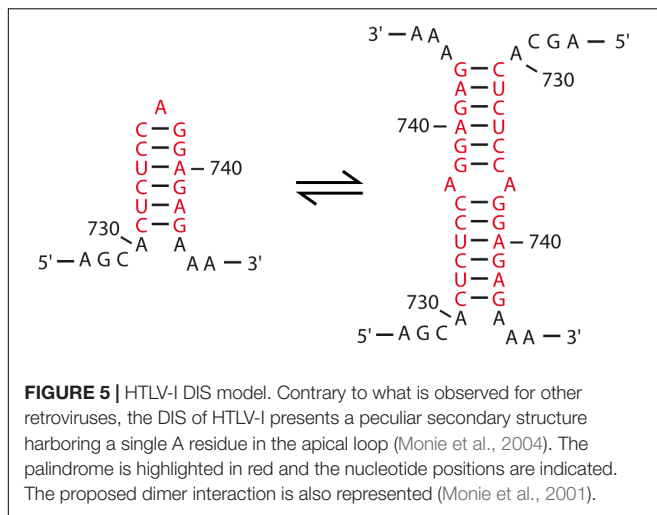
FIGURE 4 | Secondary structures and dimerization of MuLV genomic RNA. **(A)** Secondary structure of the 5′-end of MuLV gRNA containing the four stem-loops SL-A to SL-D. The nucleotide positions are indicated. **(B)** *In vitro* model of the kissing-loop/loose dimer complex. SL-A and SL-B contain a palindromic sequence of 10 and 16-nts long, respectively (in red), that promote the initiation of RNA dimerization. SL-C (purple) and SL-D (blue) both present a GACG tetraloop involved in heterologous non-canonical loop-loop interactions that stabilize the duplex. In this conformation, several UCUG quartets (green), which constitute high affinity binding sites for the viral NC protein, are trapped within SL-A and SL-B. **(C)** Secondary structure model of gRNA dimer maturation steps derived from SHAPE data (Grohman et al., 2014): the extracted *ex viro* immature form (left) is converted by NC into the mature 5′-end RNA dimer (right). In the immature conformation, only SL-A and SL-A′ are paired and adopt an extended conformation while the SL-B elements are unpaired. This conformation exposes the UCUG quartets. In the mature conformation, both SL-A/SL-A′ and SL-B/SL-B′ loop–loop interactions adopt an extended duplex structure, thus increasing dimer stability (Gherghe et al., 2010). This model also exposes the UCUG quartets. The mature dimer is thought to be to be similar to *in vitro* 5′-end tight dimer structure.

and not by NC, contrary to what is observed for other retroviruses (Katoh et al., 1991, 1993).

Similar experimental approaches revealed that a 32-nts sequence just upstream the PBS mediates *in vitro* dimerization of

human T-cell leukemia virus type-1 (HTLV-1) gRNA (Greatorex et al., 1996). Within this region, a conserved 14-nts palindromic stem-loop sequence was identified as the DIS (Greatorex et al., 1996; Monie et al., 2001), since deletion of this motif abrogated

FIGURE 5 | HTLV-I DIS model. Contrary to what is observed for other retroviruses, the DIS of HTLV-I presents a peculiar secondary structure harboring a single A residue in the apical loop (Monie et al., 2004). The palindrome is highlighted in red and the nucleotide positions are indicated. The proposed dimer interaction is also represented (Monie et al., 2001).

in vitro RNA dimerization. Since HTLV-1 RNA dimers present a very high thermal stability (Tm ~ 70–80°C), it was proposed that Watson–Crick base-pairings could not be solely responsible for RNA dimerization, and that other non-canonical or tertiary interactions would further contribute to the *in vitro* RNA dimer stability (Greatorex et al., 1996). Interestingly, biochemical structural analysis of this motif also indicated that, unlike other retroviruses displaying DIS larger loops, the HTLV-1 DIS forms a hairpin with a single A residue in the apical loop closed by non-canonical C-synG base-pairs (**Figure 5**) (Monie et al., 2004). However, in contrast to other retroviruses, deletion of the 37-nts DLS only caused a modest decrease in viral infectivity (20–25%) (Le Blanc et al., 2000). Therefore, further studies will be necessary to identify all the determinants involved in these processes.

LENTIVIRUSES GENOME DIMERIZATION

The dimerization of lentiviral RNA genomes is the best characterized retroviral dimerization process, and several important features for gRNA dimerization are conserved in the genus. Like in all retroviruses, the DLS is located at the 5′ region of the viral genome. The DIS of primate lentiviruses such as HIV-1/2 or simian immunodeficiency virus (SIV) is located within the 5′-UTR, downstream of the PBS (Skripkin et al., 1994; Dirac et al., 2001; Whitney and Wainberg, 2006) (**Figure 6A**). Contrary to gamma-retroviruses, the lentiviral DIS consists in a single palindromic sequence, typically located within the apical loop of a hairpin structure and forming homologous intermolecular canonical Watson–Crick base-pairs (**Figure 6B**). This initial intermolecular contact generally expands to a larger sequence during the stabilization of the tight dimer (**Figure 6**). Interestingly, the DIS of feline immunodeficiency virus (FIV) seems to be located within the *gag* coding region (Kenyon et al., 2011), although a structural study proposed an alternative hairpin harboring a less conserved palindromic sequence upstream of *gag* as a potential DIS (James and Sargueil, 2008).
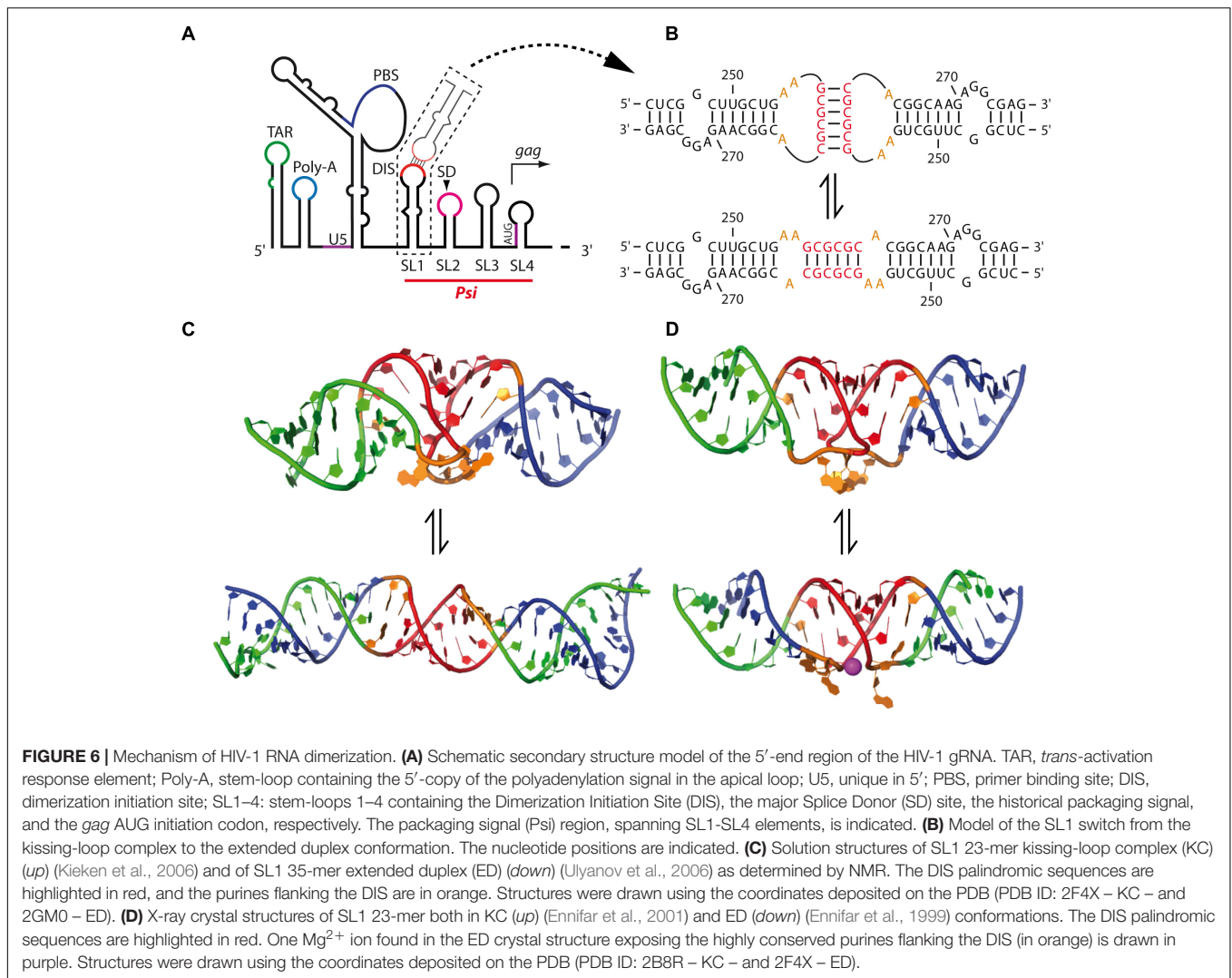
Another conserved feature in lentiviruses is the existence of a conformational switch regulating gRNA dimerization. Indeed,

gRNA fragments encompassing the 5′ region of the genome adopt alternative structures involving long-range interactions (LRI) between the R/U5 elements and regions overlapping the *gag* translation initiation codon (see below for HIV-1) to expose the DIS and promote RNA dimerization (**Figure 7**), or alternatively to prevent it (**Figure 8**) (Huthoff and Berkhout, 2001; Abbink and Berkhout, 2003; Lanchy et al., 2003a,b; Whitney and Wainberg, 2006; Kenyon et al., 2008, 2011; Lu et al., 2011; Tran et al., 2015).

Human Immunodeficiency Virus Type 1 gRNA Dimerization

The HIV-1 DIS is located in the 5′ region of the gRNA, within the packaging signal, Psi, which is composed of four stem-loops (SL1–SL4) (**Figure 6A**) (Lever et al., 1989; Aldovini and Young, 1990; Clavel and Orenstein, 1990; Clever et al., 1995; Mailler et al., 2016), although SL4, encompassing *gag* initiation codon, seems to be poorly stable and would be most likely in equilibrium with the U5 region as discussed below (Abbink and Berkhout, 2003; Lu et al., 2011). Importantly, SL1 contains in its apical loop a 6-nts palindromic sequence that have been shown by interference of chemical modifications and site-directed mutagenesis analyses to be responsible for the *in vitro* gRNA dimerization through formation of a loop-loop kissing complex (**Figure 6B**, up) SL4 (Marquet et al., 1994; Paillart et al., 1994; Skripkin et al., 1994) and is important for viral replication (Berkhout and van Wamel, 1996; Paillart et al., 1996a; Clever and Parslow, 1997; Laughrea et al., 1997, 1999). The DIS of HIV-1 is submitted to strong selection pressure, and its diversity is rather limited *in vivo*, with a strong prevalence for three GC-rich palindromes (GCGCGC, GUGCAC, and GUGCGC) (Skripkin et al., 1994; Berkhout, 1996; Paillart et al., 1996b; Hussein et al., 2010). Accordingly, *in vitro* selection experiments from randomized DIS mutants indicated that selected sequences contain the central two G-C base-pairs that are critical for kissing complex stability, while mutants presenting more than one A-U base-pair were poorly replicating (Clever et al., 1996; Laughrea et al., 1999; Lodmell et al., 2000, 2001; Hussein et al., 2010).
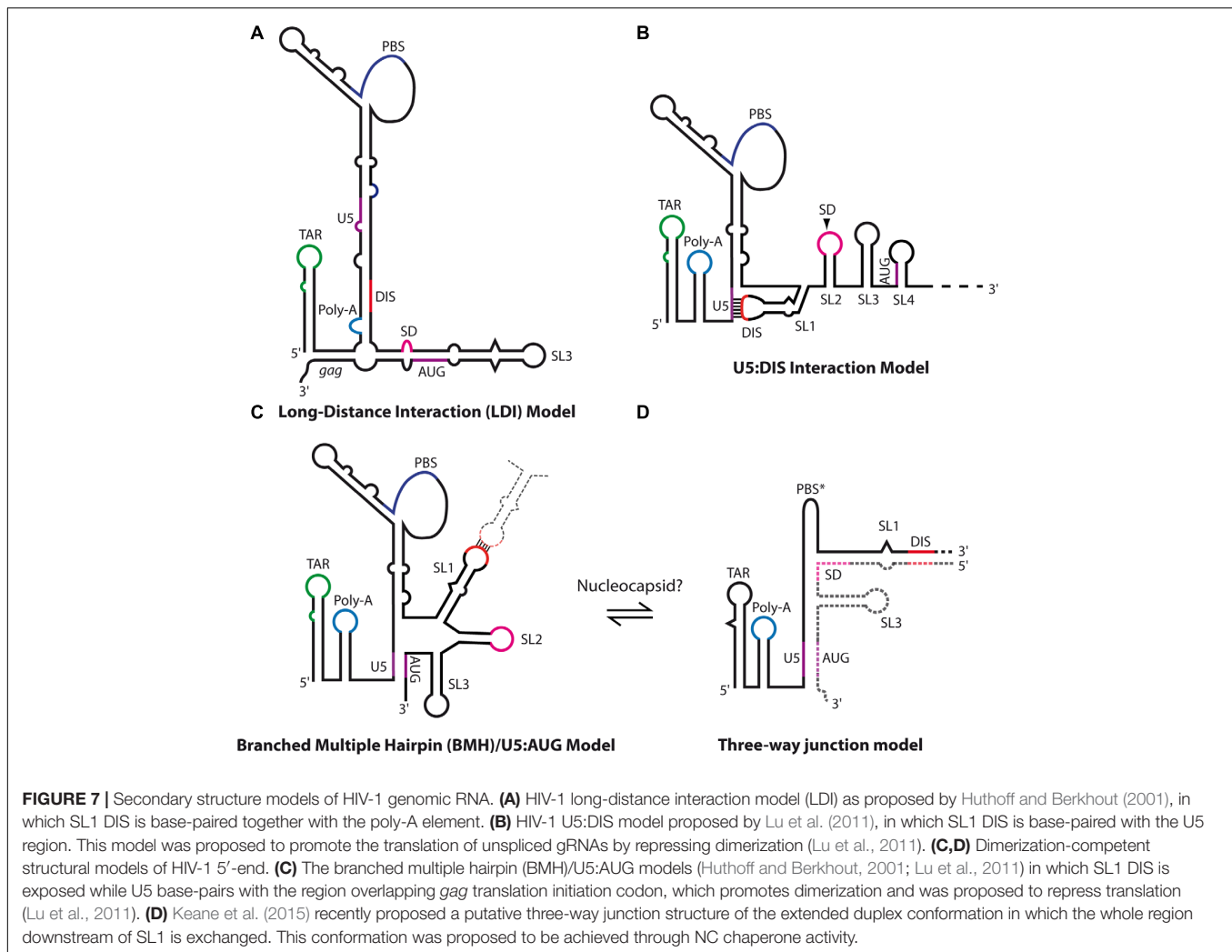
In HIV-1, the conformational switch from the kissing-loop complex to the extended duplex was extensively studied *in vitro* using 23-mer (corresponding to SL1 upper part) and 35-mer (corresponding to the whole SL1 hairpin) SL1 RNA fragments (**Figure 6B**). Both 3D structures were solved by NMR (Kieken et al., 2006; Ulyanov et al., 2006) (**Figure 6C**) and X-ray crystallography (Ennifar et al., 1999, 2001) (**Figure 6D**). Crystal structures of SL1 dimers revealed that Mg^{2+} ions promote exposure of two conserved unpaired purine residues flanking the DIS (Ennifar et al., 1999) (**Figure 6D**) that play an important role in the rate of NC-catalyzed duplex formation and were proposed to be involved in non-canonical intermolecular interactions (Paillart et al., 1997; Mihailescu and Marino, 2004; Mundigala et al., 2014). Interestingly, only the HIV-1 DIS containing the flanking purines can replace the cognate beta-retroviruses palindromic sequence (Aktar et al., 2013, 2014), further supporting the importance of these purine residues in HIV-1 palindrome dimerization, even in a heterologous context.

FIGURE 6 | Mechanism of HIV-1 RNA dimerization. **(A)** Schematic secondary structure model of the 5′-end region of the HIV-1 gRNA. TAR, *trans*-activation response element; Poly-A, stem-loop containing the 5′-copy of the polyadenylation signal in the apical loop; U5, unique in 5′; PBS, primer binding site; DIS, dimerization initiation site; SL1–4: stem-loops 1–4 containing the Dimerization Initiation Site (DIS), the major Splice Donor (SD) site, the historical packaging signal, and the *gag* AUG initiation codon, respectively. The packaging signal (Psi) region, spanning SL1-SL4 elements, is indicated. **(B)** Model of the SL1 switch from the kissing-loop complex to the extended duplex conformation. The nucleotide positions are indicated. **(C)** Solution structures of SL1 23-mer kissing-loop complex (KC) (*up*) (Kieken et al., 2006) and of SL1 35-mer extended duplex (ED) (*down*) (Ulyanov et al., 2006) as determined by NMR. The DIS palindromic sequences are highlighted in red, and the purines flanking the DIS are in orange. Structures were drawn using the coordinates deposited on the PDB (PDB ID: 2F4X – KC – and 2GM0 – ED). **(D)** X-ray crystal structures of SL1 23-mer both in KC (*up*) (Ennifar et al., 2001) and ED (*down*) (Ennifar et al., 1999) conformations. The DIS palindromic sequences are highlighted in red. One Mg²⁺ ion found in the ED crystal structure exposing the highly conserved purines flanking the DIS (in orange) is drawn in purple. Structures were drawn using the coordinates deposited on the PDB (PDB ID: 2B8R – KC – and 2F4X – ED).

Although it has been proposed, based on the crystal structure of the SL1 23-mer dimer (**Figure 6D**), that the extended duplex formation may be the result of an intermolecular *trans*-esterification reaction (Ennifar et al., 2001), NMR and UV-melting analyses on SL1 35-mers indicated that the transition from loose to tight dimers involves melting of the SL1 upper stem, without melting the kissing-loop duplex interface, followed by interstrand exchange as refolding of the stem occurs (Mundigala et al., 2014). This model is supported by several studies showing that the SL1 internal loop destabilizes the SL1 upper stem (Takahashi et al., 2000; Greatorex et al., 2002; Baig et al., 2007; Mujeeb et al., 2007) and constitutes a NC binding site that was shown to promote NC-mediated transition from the kissing-loop complex to the extended duplex (Rist and Marino, 2002; Hagan and Fabris, 2007). Interestingly, the SL1 internal loop was recently found to constitute a major binding site for Pr55^Gag (Abd El-Wahab et al., 2014; Smyth et al., 2015; Bernacchi et al., 2017), suggesting that formation of the extended duplex could be promoted by the NC domain within Pr55^Gag as well. Moreover, Mg²⁺ ions strongly impact the rate at which the duplex is formed

(Bernacchi et al., 2005) and were observed to bind the SL1 internal loop, possibly preventing premature transition from the kissing-loop to the extended duplex (Sun et al., 2007). Therefore, one could speculate on a model where Mg²⁺ and Pr55^Gag (or NC) compete for binding to the SL1 internal loop, thus stabilizing or destabilizing the upper stem. Consistent with this notion, the SL1 internal loop was shown to be important for RNA dimerization, packaging and viral infectivity in cell cultures (Clever et al., 1996; Paillart et al., 1996a; Clever and Parslow, 1997; Harrison et al., 1998; Laughrea et al., 1999; Shen et al., 2000) and stabilization of the SL1 hairpin resulted in replication defects (van Bel et al., 2014).

The existence of an extended duplex in large HIV-1 RNA fragments encompassing the whole DLS constitutes a question of debate and remains to be clearly demonstrated in the viral context. First, probing techniques cannot discriminate intramolecular from intermolecular base-pairings, thus making difficult the discrimination of both conformations. Second, SL1 *trans*-complementary mutants that are not able to form the extended duplex efficiently dimerized *in vitro* and were found

FIGURE 7 | Secondary structure models of HIV-1 genomic RNA. **(A)** HIV-1 long-distance interaction model (LDI) as proposed by Huthoff and Berkhout (2001), in which SL1 DIS is base-paired together with the poly-A element. **(B)** HIV-1 U5:DIS model proposed by Lu et al. (2011), in which SL1 DIS is base-paired with the U5 region. This model was proposed to promote the translation of unspliced gRNAs by repressing dimerization (Lu et al., 2011). **(C,D)** Dimerization-competent structural models of HIV-1 5′-end. **(C)** The branched multiple hairpin (BMH)/U5:AUG models (Huthoff and Berkhout, 2001; Lu et al., 2011) in which SL1 DIS is exposed while U5 base-pairs with the region overlapping *gag* translation initiation codon, which promotes dimerization and was proposed to repress translation (Lu et al., 2011). **(D)** Keane et al. (2015) recently proposed a putative three-way junction structure of the extended duplex conformation in which the whole region downstream of SL1 is exchanged. This conformation was proposed to be achieved through NC chaperone activity.

to be as stable as wild-type RNA fragments (Paillart et al., 1996c). Besides, the conformational switch from the kissing-loop complex to the extended duplex in full-length gRNA or even in complete 5′-UTR context presents huge topological and steric constraints. Therefore, it is reasonable to consider that HIV-1 RNA tight dimer formation would rather be the result of conformational rearrangements at a larger scale, with the formation of additional contacts (see below). The notion of additional inter-genomic contacts outside SL1/DIS is also supported by electron microscopy studies on full-length gRNA dimers extracted from virions and atomic force microscopy (AFM) analysis on gRNA fragments encompassing the first 744 nts of the HIV-1 genome showing that HIV-1 gRNA dimers are not Y-shaped, like in other retroviruses, but rather formed a loop toward the genome 5′-end (Höglund et al., 1997; Andersen et al., 2004; Pallesen, 2011). Besides the DIS, the most extensively described gRNA motif potentially having a role in the dimerization process is the TAR element at the 5′ extremity of the genome (**Figure 6A**). Although HIV-1 TAR contains a 10-nts palindromic sequence in its apical loop that mediates dimerization of a 57-nts long TAR fragment (Andersen

et al., 2004), it seems this link with gRNA dimerization is rather indirect. Indeed, even though electrophoretic mobility analyses of mutant TAR RNA dimers extracted from virions presented dimerization defects (Song et al., 2008; Jalalirad et al., 2012), compensatory mutations failed to restore the putative intermolecular TAR-TAR interaction (Jalalirad et al., 2012). Additionally, other *in cellula* and *in vitro* studies indicated that TAR has a moderate effect on RNA dimerization (Das et al., 2007; Lu et al., 2011; Heng et al., 2012), suggesting that TAR destabilization impacts the overall gRNA structure, thus inducing aberrant gRNA dimerization and packaging (Ooms et al., 2004; Vrolijk et al., 2008; Das et al., 2012). Therefore, the intermolecular TAR-TAR interactions would promote gRNA structure stability rather than RNA dimerization.

HIV-1 gRNA can alternatively serve as a template for the synthesis of Gag precursors, and/or be selected for encapsidation within virions as dimers. Consistent with this notion, it was suggested that the regulation of HIV-1 gRNA translation or dimerization and packaging would be due to RNA structural switches. Using biochemical and chemical probing, Huthoff and Berkhout (2001), Berkhout et al. (2002), Abbink and
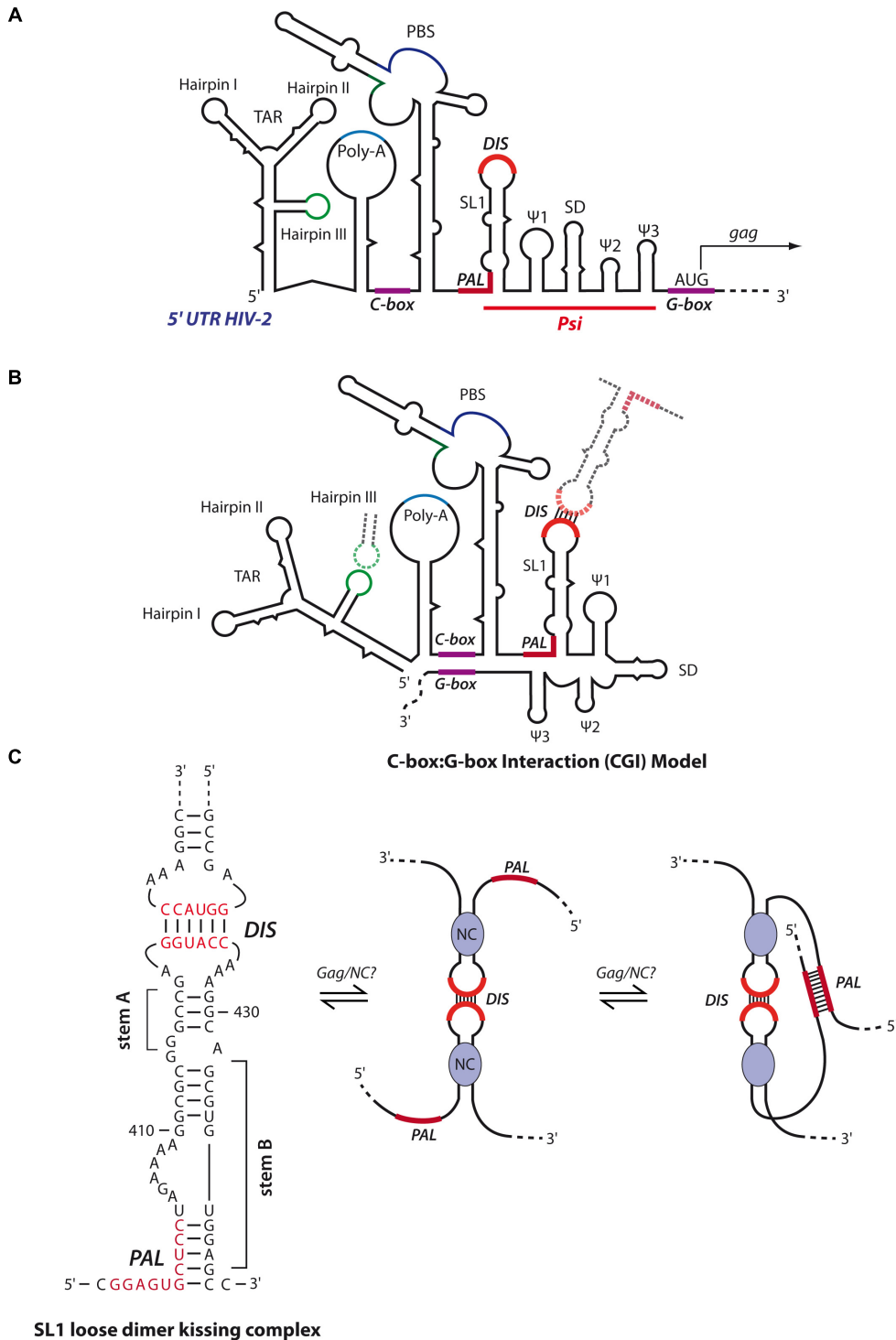
FIGURE 8 | Secondary structure models of HIV-2 genomic RNA. **(A)** Schematic representation of the secondary structure model of the 5′- region of the HIV-2 gRNA. The red line delimits the packaging signal (Psi) region containing the five stem-loops SL1, Ψ1 to Ψ3 and SD. Similar to HIV-1, SL1 contains in its apical loop a hexanucleotide palindromic sequence constituting the DIS. SL1 also contains in its basal part a 10-nts palindromic sequence partially entrapped (PAL). The *gag* translation initiation codon is located within a G-rich region (G-box). Upstream of Psi, are found the TAR, the poly-A, a C-rich region important for RNA dimerization (C-box) and the PBS. **(B)** In the CGI dimer structural conformation, the C-box and G-Box are base-paired. This conformation restricts HIV-2 RNA dimerization and is adopted by loose dimers when the DIS and TAR hairpin III are involved in kissing-loop interactions (Lanchy et al., 2003a,b). **(C)** The RNA switch model from loose to tight dimer, as proposed by Purzycka et al. (2011). HIV-2 SL1 stem B is melted through Gag/NC chaperone activity, freeing the PAL region that can thus forms additional intermolecular base-pairings to stabilize the RNA dimer.

Berkhout (2003), Abbink et al. (2005) proposed an RNA-switch mechanism from a long-distance interaction (LDI) to a branched multiple hairpin (BMH) conformation that would regulate gRNA dimerization and translation (**Figures 7A,B**). In the LDI conformation, SL1 is disrupted and base-paired with the poly-A element, thus sequestering the DIS sequence and potentially promoting translation (**Figure 7A**). Alternatively, in the BMH conformation, the *gag* AUG initiation codon together with the U5 region forms the so-called U5:AUG interaction, displaying SL1-SL3 to fold hairpin structures (**Figure 7B**) and promoting gRNA dimerization. Importantly, if the U5:AUG interaction is supported by phylogenic studies showing its conservation and co-variation throughout HIV-1 isolates (Abbink and Berkhout, 2003; Russell et al., 2003; Damgaard et al., 2004; Song et al., 2008; Wilkinson et al., 2008; Lu et al., 2011; Sakuragi et al., 2012; Keane et al., 2015, 2016; Tran et al., 2015; Mueller et al., 2016) and other related lentiviruses such as SIV (Whitney and Wainberg, 2006; Tran et al., 2015) or HIV-2 (see below), the LDI conformation has never been characterized *in vivo*, and RNA mutants aiming to disrupt LDI failed to affect HIV-1 unspliced RNA translation (Abbink et al., 2005), thus questioning the biological relevance of the LDI conformation. Also, mutational analysis of two highly conserved sequences, upstream of SL1 and downstream of SL4, suggested that these are part of a pseudoknot-like conformation contributing to HIV-1 DLS structure and promoting dimerization (Sakuragi et al., 2012).

An alternative *in vitro* RNA conformational switch model derived from the previous one was proposed by the group of Michael Summers by using a new NMR approach (long-range probing by adenosine interaction detection or lr-AID) which is based on selective isotopic labeling of each RNA molecule in the dimer thus allowing the discrimination between intra- and inter-molecular base-pairs (Lu et al., 2011). In this model, contrary to the LDI-BMH conformational switch, the DIS and *gag* initiation codon are alternatively base-paired with the U5 region (**Figures 7B,C**), promoting translation or RNA dimerization, respectively. Even though the dimerization competent U5:AUG RNA structure appears rather similar to the BMH conformation (Lu et al., 2011) (**Figure 7B**), a striking difference corresponds to the three-way junction RNA structure at the base of SL1, in which SL2 base-pairs with the bottom region of the PBS from the other gRNA molecule (**Figure 7D**) (Keane et al., 2015, 2016). However, this latter conformation was very recently challenged by a phylogenic study supporting the existence of SL2 hairpin structure (Mueller et al., 2016).

Human Immunodeficiency Virus Type 2 gRNA Dimerization

The secondary structure and *in vitro* dimerization properties of HIV-2 gRNA present similarities to HIV-1 gRNA (**Figure 8A**). First, loose dimer formation involves several contacts including the SL1 region of Psi. HIV-2 SL1 also contains a 6-nts palindromic sequence within the apical loop driving the dimerization of 5′ RNA fragments (Dirac et al., 2001). Second, the HIV-2 TAR hairpin III (**Figures 8A,B**) palindromic sequence was also shown to contribute to the stabilization of the kissing-complex

by forming homologous intermolecular base-pairs (**Figure 8B**) (Purzycka et al., 2011). Interestingly, in the context of RNA fragments encompassing the whole 5′-UTR (up to position ∼560), another palindromic sequence in the PBS was also described to be required for RNA dimerization (Jossinet et al., 2001; Lanchy and Lodmell, 2002; Lanchy et al., 2003b).

However, *in vitro* tight dimers of HIV-2 RNA fragments seem to be quite different from HIV-1. Indeed, RNA fragments encompassing the whole 5′-UTR, in the absence of HIV-2 NC protein and at physiological temperature, were shown to form only loose dimers through the PBS palindrome (Jossinet et al., 2001; Lanchy and Lodmell, 2002; Lanchy et al., 2003b), while shorter transcripts (up to position ∼440, i.e., ending immediately 3′ of SL1) formed tight dimers through the SL1 apical loop (Dirac et al., 2001; Lanchy et al., 2003a). A possible explanation relies on the fact that HIV-2 RNA secondary structure forms a LRI between a C-rich region upstream of the PBS and a G-rich region overlapping the *gag* translation initiation codon termed C box-G box interaction (CGI), homologous to the HIV-1 U5:AUG interaction (compare **Figures 7B**, **8B**). Interestingly, while this latter was proposed to promote HIV-1 RNA dimerization through SL1 (Dirac et al., 2001; Lanchy et al., 2003a), the CGI seems to prevent formation of tight HIV-2 RNA dimers since targeting the CGI with oligonucleotides restores SL1-dependent dimerization *in vitro* (Lanchy et al., 2003a,b). Moreover, incubation of RNA fragments either at high temperature (>50°C) or in presence of HIV-2 NC has been shown to destabilize the CGI and restore SL1-dependent tight dimerization *in vitro* (Lanchy and Lodmell, 2002; Lanchy et al., 2003b). Interestingly, in addition to the apical loop of SL1, an upstream 10 nt-long conserved palindromic sequence (termed PAL) also contributes to *in vitro* RNA tight dimer formation (**Figure 8C**). Indeed, by using antisense oligonucleotides targeting the SL1 apical DIS or PAL, this latter was shown to serve as an alternative stabilization element when the apical DIS is repressed (Lanchy et al., 2003a). Importantly, *in vitro* SHAPE and biochemical analyses revealed that stem B of SL1 (**Figure 8C**), formed by base-pairing between the 3′-end of PAL and a region downstream of SL1 stem A (**Figure 8C**), regulates HIV-2 RNA dimerization (Baig et al., 2007; Purzycka et al., 2011). Indeed, in the proposed model, upon opening of SL1 stem B through the chaperone activity of Gag/NC, the PAL region becomes available to build interstrand contacts. Interestingly, a comparative study revealed that the dimerization properties of HIV-1, HIV-2 and SIV RNAs are regulated through the stability of SL1 stem B (Baig et al., 2008).

Consistently with these *in vitro* results, mutations in the PAL region strongly impaired HIV-2 RNA dimerization and infectivity *in cellula*, while mutations in the SL1 apical loop had limited impact (L'Hernault et al., 2007, 2012). Similarly, the integrity of SL1 stem B was shown to be crucial for HIV-2 gRNA packaging (Lanchy and Lodmell, 2007). Finally, it was also proposed that, *in cellula*, translation of HIV-2 gRNA by the ribosome may disrupt the CGI to allow RNA dimerization, a notion consistent with a study proposing that HIV-2 genome is packaged co-translationally (Griffin et al., 2001). However, this *cis*-packaging model was recently challenged and it was elegantly shown by visualization HIV-2 RNA in individual

particles that *trans*-packaging is the major mechanism of HIV-2 RNA packaging, as it was observed for HIV-1 (Ni et al., 2011).

SPUMAVIRUSES GENOME DIMERIZATION

Spumaretroviruses, or foamy viruses (FVs), have the largest genome in the *Retroviridae* family (∼12 kb) and are singularly different from other retroviruses. Amongst these differences, FV Gag proteins do not present the NC Cys-His zinc fingers, a hallmark of other retroviral Gag proteins, but have instead Gly-Arg rich domains binding nucleic acids with equal affinity and required for viral replication (Yu et al., 1996; Linial, 1999). Furthermore, reverse transcription is a late event in FV replication cycle and infectious viral particles contain DNA genomes (Linial, 1999; Yu et al., 1999). Despite these differences, three different sites (SI, SII, and SIII) were shown to contribute to *in vitro* dimerization of RNA fragments corresponding to the 5′ region of FV gRNA (Erlwein et al., 1997). Amongst these domains, SII maps to positions 391–410 and contains a highly conserved 10-nts UCCUAGGA palindrome that is crucial for *in vitro* RNA dimerization since antisense oligonucleotides or mutations targeting this motif abolish RNA dimer formation (Erlwein et al., 1997; Cain et al., 2001). SI and SIII also contribute to RNA dimerization, but to a lesser extent since antisense oligonucleotide targeting of these elements alone failed to completely abrogate *in vitro* RNA dimerization of primate FV 5′-RNA fragments (Erlwein et al., 1997; Cain et al., 2001). However, targeting bovine FV (BFV) SI element by antisense oligonucleotides inhibited *in vitro* RNA dimerization similarly to what was observed when targeting the SII palindrome (Yu et al., 2007).

Importantly, mutation of the SII palindrome in cell culture resulted in a strong viral replication defect, even though genome packaging was not affected (Park and Mergia, 2000; Cain et al., 2001), differently from other retroviruses such as HIV-1 which showed defects in gRNA packaging when mutation of the DIS occurred (Berkhout and van Wamel, 1996; Paillart et al., 1996a; Laughrea et al., 1997, 1999; Houzet et al., 2007). Since it is well known that RNA dimerization plays an important role in reverse transcription (Paillart et al., 1996a; Berkhout et al., 1998; Parent et al., 2000), it seems likely that this replication defect may come from reverse transcription defects (Park and Mergia, 2000).

IN CELLULA GENOME DIMERIZATION: WHERE AND WHEN?

Retroviruses select specifically two copies of their genome from the pool of cellular and viral spliced RNAs for packaging. Since aberrant gRNA dimerization may interfere with different steps in the retroviral life cycle, including viral assembly, the timing and localization of this process must be finely regulated within the host cell. An important question has been whether two monomers are packaged and dimerize during/after viral assembly, or if a pre-formed RNA dimer is selected and packaged into viral particles. However, since gRNA dimerization and packaging are interconnected events (Aagaard et al., 2004; Houzet et al., 2007; Moore et al., 2007; Miyazaki et al., 2010a; Mailler et al., 2016), gRNA dimerization was supposed to initiate upon gRNA recognition by the Gag precursor. In the following section, we will review the current understanding of the spatio-temporal regulation of retroviral RNA dimerization.

Various studies using protease-deficient (PR⁻) virions or different Gag mutants suggested that HIV-1 gRNA is packaged as two monomers (Shehu-Xhilaga et al., 2001; Song et al., 2007). It is, however, possible that gRNA extraction protocols may disrupt immature loose dimers, especially when the stability of RNA mutants is reduced (Höglund et al., 1997). Although it has been reported that, in some cell types, HIV-1 gRNA DIS mutants presenting packaging defects can still be dimeric (Laughrea et al., 1997; Hill et al., 2003), a majority of studies showed that disruption of SL1 kissing-loop interface impaired or affected gRNA packaging (Berkhout and van Wamel, 1996; Paillart et al., 1996a; Houzet et al., 2007; Moore et al., 2007). Moreover, duplication of HIV-1 DLS leads to production of partially monomeric genomes, due to gRNA circularization by self-dimerization (Sakuragi et al., 2001, 2007), contradicting the notion that HIV-1 packages its genome as RNA monomers. The comparison of HIV-1 gRNA structure *in cellula* and *in viro* revealed little structural rearrangements between these two conditions (Paillart et al., 2004), and recombination analyses supported the notion that gRNA dimerization initiates in the cytoplasm (Moore et al., 2009).

An extensive set of analyses support the idea that gRNA dimerization occurs prior budding of viral particle (Moore et al., 2007; Jouvenet et al., 2009; Moore and Hu, 2009; Sardo et al., 2015; Ferrer et al., 2016). Because retroviral assembly involves only two copies of gRNA, detecting RNA dimerization in cells is challenging. Several bio-imaging approaches based on fluorescence have been used to address this issue. To allow specific gRNA observation *in cellula*, the viral genome is labeled by incorporating numerous stem-loops that specifically bind the coat protein of MS2 bacteriophage (Bertrand et al., 1998; Fusco et al., 2003), or the phage lambda protein λN22 (Daigle and Ellenberg, 2007) fused to fluorescent protein(s). Although it was suggested that HIV-1 gRNA migrates at the plasma membrane as a pre-formed dimer (Jouvenet et al., 2009; Chen et al., 2016), further *ex vivo* analyses showed that HIV-1 gRNA dimer reaches the plasma membrane in association with low-order multimers of Gag precursor (Jouvenet et al., 2008, 2009; Kutluay and Bieniasz, 2010; Kutluay et al., 2014). Importantly, Ferrer et al. (2016) recently combined Fluorescence *In Situ* Hybridization (FISH), TIRF-M, 3D-super-resolution microscopy and Fluorescence Cross-Correlation Spectroscopy (FCCS) to show that HIV-1 gRNA dimerization already occurs in the cytosol but that RNA dimers are more easily detected at the plasma membrane. One possible explanation could be the concentration of Gag at the plasma membrane, which might promote gRNA dimer stabilization through the chaperone activity of its NC domain (Ferrer et al., 2016). In this context, Hu and co-workers used a similar approach and proposed that HIV-1 gRNA dimerization would occur preferentially at

the plasma membrane (Chen et al., 2016). Their comparison between TIRF-M and *in silico* simulations showed that RNA heterodimerization would occur at frequencies similar to those observed in Ferrer et al. (2016) study (∼10–15%). Since the co-localization frequencies increased at the plasma membrane in a Gag-dependent manner, they thus proposed that HIV-1 genome diffuses into the cytoplasm and reaches the plasma membrane as a monomer. However, a common point for both studies resides in the differences between heterodimers frequencies at the plasma membrane (∼10–15% co-localization) and in virions (∼40–50%) (Chen et al., 2009; Ferrer et al., 2016), suggestive of an RNA dimer stabilization step at the plasma membrane *prior* to encapsidation. It is, however, possible that technical difficulties for the detection of RNA molecules in a highly dynamic environment, such as the cytosol and the plasma membrane, compared to the stable environment within viral particles, could affect these estimations. Further analyses are definitively needed to clarify spatiotemporal HIV-1 gRNA dimerization.

Finally, studies suggest that MuLV gRNA dimerization, unlike HIV-1, occurs in the nucleus. Indeed, heterodimerization is enhanced when two proviruses are spatially close, suggesting that MuLV gRNA dimerization is coupled to transcription and splicing processes (Maurel and Mougel, 2010). Consistent with these observations, MuLV RNAs transcribed from the same locus form dimers at high frequencies (Flynn et al., 2004; Kharytonchyk et al., 2005; Rasmussen and Pedersen, 2006; Maurel et al., 2007). Interestingly, mutations of RNA elements involved in dimerization or packaging processes impact the intracellular transport of viral genome and result in aberrant accumulation in the nucleus or in the cytoplasm (Basyuk et al., 2005; Smagulova et al., 2005).

RNA DIMERIZATION AND REVERSE TRANSCRIPTION/RECOMBINATION

Reverse transcription is initiated at the PBS immediately 3′ to the 5′ copy of the R (repeat) region of the genome by the viral RT enzyme and generates the complete viral cDNA with duplicated LTR. In order to achieve cDNA synthesis, the negative strong-stop DNA must be translocated from the 5′ to the 3′-end of the gRNA during the first strand transfer occurring in an intra- or inter-molecular manner (Panganiban and Fiore, 1988; Hu and Temin, 1990; Berkhout and van Wamel, 1996). During reverse transcription, recombination occurs as the RT enzyme switches between the two RNA templates. This process is due to RNA sequence homology and promoted by particular RNA structures, nicks in the viral genome or RT pausing (Galetto and Negroni, 2005). Importantly, recombination provides a mechanism for genome repair or, in the case of a heterodiploid genome, a way to increase genetic diversity by genomes shuffling. This leads to immune escape and drug resistance, especially in the case of HIV-1 (Morris et al., 1999; Rambaut et al., 2004), for which it has previously been shown that gRNA heterodimerization rate can reach ∼40–50% *in viro* (Chen et al., 2009; Ferrer et al., 2016), and that recombination occurs at least three times during a single viral replication cycle

(Zhuang et al., 2002; Schlub et al., 2010, 2014; Smyth et al., 2014).

Genomic RNA dimerization may play an important role during reverse transcription since RSV mutant particles containing monomeric genomes display a 100-fold decrease in cDNA synthesis (Parent et al., 2000) and heat-induced dissociation of HIV-1 gRNA dimers inhibits the first strand transfer (Berkhout et al., 1998). In MuLV, template switching occurs preferentially at direct RNA–RNA interactions mediated by palindromic sequences (Mikkelsen et al., 1998a,b, 2000). In the case of a co-infection by two genetically different viruses, the subsequent packaging of a heterozygous dimer increases the viral diversity by enabling the production of new recombinant strains (Mikkelsen et al., 2000). In HIV-1, *trans*-complementary DIS mutants showed that the recombination rate is enhanced when two palindromic mutants can form Watson–Crick base-pairs (from 50 to 90% depending on the mutants) (Moore et al., 2009). Similarly, both the *in vitro* and *in cellula* HIV-1 recombination rates are directly linked to the dimeric state of the viral genome (Balakrishnan et al., 2001, 2003; Sakuragi et al., 2015, 2016). Taken together, these findings highlight the central role of gRNA dimerization in the reverse transcription and recombination processes.

RNA DIMERIZATION AS A POTENTIAL ANTIRETROVIRAL TARGET

While the current antiretroviral treatments have proved successful in extending the life expectancy of infected patients, the increasing number of multi-drug resistant HIV-1 mutants highlights the need for the discovery of new antiretroviral drugs. In this context, gRNA dimerization is an attractive target since it is crucial for several key steps in the retrovirus life cycle. Interestingly, the crystal structure of the HIV-1 kissing-loop complex revealed strong structural similarities with the 16S ribosomal A site, a natural aminoglycoside binding site (Ennifar et al., 2003). Indeed, some aminoglycosides, such as neomycin and lividomycin, bind the HIV-1 kissing-loop complex *in vitro* with high affinity and strongly stabilize it, preventing its conversion to an extended (Ennifar et al., 2003, 2006, 2007; Bernacchi et al., 2007). Interestingly, footprint analysis also revealed that the DIS of HIV-1 gRNA is also protected by aminoglycosides in cells and in virions (Ennifar et al., 2006, 2007), thus providing an interesting basis for the development of new antiretroviral strategies aiming at targeting retroviral gRNA dimerization (Ennifar et al., 2007, 2013; Blond et al., 2014).

CONCLUDING REMARKS

Retroviral gRNA dimerization is a highly conserved mechanism in retroviruses, and plays a critical role in several key steps of their replicative cycle. RNA dimerization, which is a pre-requisite for packaging, plays a critical role in the generation of multi-drug resistant recombinant viruses during reverse transcription. In this context, a general feature of retroviruses is the existence

of complex structural switches regulating the transition between a loose kissing-loop complex to a stable extended duplex conformation, and the equilibrium between gRNA translation and dimerization for packaging. Nevertheless, *in vitro* structural models most likely do not exactly reflect the *in vivo* situation and, despite recent progresses, much remains to be unraveled to reach complete understanding of the RNA dimerization process. This understanding could also allow the development of new antiretroviral strategies that are essential to counterbalance the increasing prevalence of multi-drug resistant viruses.

AUTHOR CONTRIBUTIONS

ND, RM, J-CP, and SB conceived the review topic. ND drafted the manuscript and generated the figures. SB, J-CP, and RM

corrected and edited the manuscript. All authors read and approved the final version of the manuscript.

REFERENCES

Aagaard, L., Rasmussen, S. V., Mikkelsen, J. G., and Pedersen, F. S. (2004). Efficient replication of full-length murine leukemia viruses modified at the dimer initiation site regions. *Virology* 318, 360–370. doi: 10.1016/j.virol.2003. 09.008

Abbink, T. E. M., and Berkhout, B. (2003). A novel long distance -pairing interaction in human immunodeficiency virus type 1 RNA occludes the Gag start codon. *J. Biol. Chem.* 278, 11601–11611. doi: 10.1074/jbc.M210291200

Abbink, T. E. M., Ooms, M., Haasnoot, P. C. J., and Berkhout, B. (2005). The HIV-1 leader RNA conformational switch regulates RNA dimerization but does not regulate mRNA translation. *Biochemistry* 44, 9058–9066. doi: 10.1021/bi0502588

Abd El-Wahab, E. W., Smyth, R. P., Mailler, E., Bernacchi, S., Vivet-Boudou, V., Hijnen, M., et al. (2014). Specific recognition of the HIV-1 genomic RNA by the Gag precursor. *Nat. Commun.* 5:4304. doi: 10.1038/ncomms5304

Aduri, R., Briggs, K. T., Gorelick, R. J., and Marino, J. P. (2013). Molecular determinants of HIV-1 NCp7 chaperone activity in maturation of the HIV-1 dimerization initiation site. *Nucleic Acids Res.* 41, 2565–2580. doi: 10.1093/nar/gks1350

Aktar, S. J., Jabeen, A., Ali, L. M., Vivet-Boudou, V., Marquet, R., and Rizvi, T. A. (2013). SHAPE analysis of the 5' end of the Mason-Pfizer monkey virus (MPMV) genomic RNA reveals structural elements required for genome dimerization. *RNA* 19, 1648–1658. doi: 10.1261/rna.040931.113

Aktar, S. J., Vivet-Boudou, V., Ali, L. M., Jabeen, A., Kalloush, R. M., Richer, D., et al. (2014). Structural basis of genomic RNA (gRNA) dimerization and packaging determinants of mouse mammary tumor virus (MMTV). *Retrovirology* 11:96. doi: 10.1186/s12977-014-0096-6

Aldovini, A., and Young, R. A. (1990). Mutations of RNA and protein sequences involved in human immunodeficiency virus type 1 packaging result in production of noninfectious virus. *J. Virol.* 64, 1920–1926.

Andersen, E. S., Contera, S. A., Knudsen, B., Damgaard, C. K., Besenbacher, F., and Kjems, J. (2004). Role of the trans-activation response element in dimerization of HIV-1 RNA. *J. Biol. Chem.* 279, 22243–22249. doi: 10.1074/jbc.M31432 6200

Badorrek, C. S., Gherghe, C. M., and Weeks, K. M. (2006). Structure of an RNA switch that enforces stringent retroviral genomic RNA dimerization. *Proc. Natl. Acad. Sci. U.S.A.* 103, 13640–13645. doi: 10.1073/pnas.0606156103

Badorrek, C. S., and Weeks, K. M. (2005). RNA flexibility in the dimerization domain of a gamma retrovirus. *Nat. Chem. Biol.* 1, 104–111. doi: 10.1038/nchembio712

Badorrek, C. S., and Weeks, K. M. (2006). Architecture of a gamma retroviral genomic RNA dimer. *Biochemistry* 45, 12664–12672. doi: 10.1021/bi060521k

Baig, T. T., Lanchy, J.-M., and Lodmell, J. S. (2007). HIV-2 RNA dimerization is regulated by intramolecular interactions in vitro. *RNA* 13, 1341–1354. doi: 10.1261/rna.483807

Baig, T. T., Strong, C. L., Lodmell, J. S., and Lanchy, J.-M. (2008). Regulation of primate lentiviral RNA dimerization by structural entrapment. *Retrovirology* 5:65. doi: 10.1186/1742-4690-5-65

Balakrishnan, M., Fay, P. J., and Bambara, R. A. (2001). The kissing hairpin sequence promotes recombination within the HIV-I 5' leader region. *J. Biol. Chem.* 276, 36482–36492. doi: 10.1074/jbc.M102860200

Balakrishnan, M., Roques, B. P., Fay, P. J., and Bambara, R. A. (2003). Template dimerization promotes an acceptor invasion-induced transfer mechanism during human immunodeficiency virus type 1 minus-strand synthesis. *J. Virol.* 77, 4710–4721. doi: 10.1128/JVI.77.8.4710-4721.2003

Basyuk, E., Boulon, S., Skou Pedersen, F., Bertrand, E., and Vestergaard Rasmussen, S. (2005). The packaging signal of MLV is an integrated module that mediates intracellular transport of genomic RNAs. *J. Mol. Biol.* 354, 330–339. doi: 10.1016/j.jmb.2005.09.071

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. doi: 10.1006/jmbi.1993.1041

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.

Berkhout, B. (1996). Structure and function of the human immunodeficiency virus leader RNA. *Prog. Nucleic Acid Res. Mol. Biol.* 54, 1–34. doi: 10.1016/S0079-6603(08)60359-1

Berkhout, B., Das, A. T., and van Wamel, J. L. (1998). The native structure of the human immunodeficiency virus type 1 RNA genome is required for the first strand transfer of reverse transcription. *Virology* 249, 211–218. doi: 10.1006/viro.1998.9321

Berkhout, B., Ooms, M., Beerens, N., Huthoff, H., Southern, E., and Verhoef, K. (2002). In vitro evidence that the untranslated leader of the HIV-1 genome is an RNA checkpoint that regulates multiple functions through conformational changes. *J. Biol. Chem.* 277, 19967–19975. doi: 10.1074/jbc.M200950200

Berkhout, B., and van Wamel, J. L. (1996). Role of the DIS hairpin in replication of human immunodeficiency virus type 1. *J. Virol.* 70, 6723–6732.

Bernacchi, S., Abd El-Wahab, E. W., Dubois, N., Hijnen, M., Smyth, R. P., Mak, J., et al. (2017). HIV-1 Pr55(Gag) binds genomic and spliced RNAs with different affinity and stoichiometry. *RNA Biol.* 14, 90–103. doi: 10.1080/15476286.2016. 1256533

Bernacchi, S., Ennifar, E., Tóth, 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. doi: 10.1074/jbc.M503230200

Bernacchi, S., Freisz, S., Maechling, C., Spiess, B., Marquet, R., Dumas, P., et al. (2007). Aminoglycoside binding to the HIV-1 RNA dimerization initiation site: thermodynamics and effect on the kissing-loop to duplex conversion. *Nucleic Acids Res.* 35, 7128–7139. doi: 10.1093/nar/gkm856

Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S. M., Singer, R. H., and Long, R. M. (1998). Localization of ASH1 mRNA particles in living yeast. *Mol. Cell* 2, 437–445. doi: 10.1016/S1097-2765(00)80143-4

Bieth, E., Gabus, C., and Darlix, J. L. (1990). A study of the dimer formation of Rous sarcoma virus RNA and of its effect on viral protein synthesis in vitro. *Nucleic Acids Res.* 18, 119–127. doi: 10.1093/nar/18.1.119

Blond, A., Ennifar, E., Tisné, C., and Micouin, L. (2014). The design of RNA binders: targeting the HIV replication cycle as a case study. *ChemMedChem* 9, 1982–1996. doi: 10.1002/cmdc.201402259

Boeras, I., Seufzer, B., Brady, S., Rendahl, A., Heng, X., and Boris-Lawrie, K. (2017). The basal translation rate of authentic HIV-1 RNA is regulated by 5'UTR nt-pairings at junction of R and U5. *Sci. Rep.* 7:6902. doi: 10.1038/s41598-017-06883-9

Cain, D., Erlwein, O., Grigg, A., Russell, R. A., and McClure, M. O. (2001). Palindromic sequence plays a critical role in human foamy virus dimerization. *J. Virol.* 75, 3731–3739. doi: 10.1128/JVI.75.8.3731-3739.2001

Canaani, E., Helm, K. V., and Duesberg, P. (1973). Evidence for 30-40S RNA as precursor of the 60-70S RNA of Rous sarcoma virus. *Proc. Natl. Acad. Sci. U.S.A.* 70, 401–405. doi: 10.1073/pnas.70.2.401

Chen, J., Nikolaitchik, O., Singh, J., Wright, A., Bencsics, C. E., Coffin, J. M., et al. (2009). High efficiency of HIV-1 genomic RNA packaging and heterozygote formation revealed by single virion analysis. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13535–13540. doi: 10.1073/pnas.0906822106

Chen, J., Rahman, S. A., Nikolaitchik, O. A., Grunwald, D., Sardo, L., Burdick, R. C., et al. (2016). HIV-1 RNA genome dimerizes on the plasma membrane in the presence of Gag protein. *Proc. Natl. Acad. Sci. U.S.A.* 113, E201–E208. doi: 10.1073/pnas.1518572113

Cimarelli, A., Sandin, S., Höglund, S., and Luban, J. (2000). Basic residues in human immunodeficiency virus type 1 nucleocapsid promote virion assembly via interaction with RNA. *J. Virol.* 74, 3046–3057. doi: 10.1128/JVI.74.7.3046-3057.2000

Clavel, F., and Orenstein, J. M. (1990). A mutant of human immunodeficiency virus with reduced RNA packaging and abnormal particle morphology. *J. Virol.* 64, 5230–5234.

Clever, J., Sassetti, C., and Parslow, T. G. (1995). RNA secondary structure and binding sites for gag gene products in the 5' packaging signal of human immunodeficiency virus type 1. *J. Virol.* 69, 2101–2109.

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.

Clever, J. L., Wong, M. L., and Parslow, T. G. (1996). Requirements for kissing-loop-mediated dimerization of human immunodeficiency virus RNA. *J. Virol.* 70, 5902–5908.

Cruceanu, M., Urbaneja, M. A., Hixson, C. V., Johnson, D. G., Datta, S. A., Fivash, M. J., et al. (2006). Nucleic acid binding and chaperone properties of HIV-1 Gag and nucleocapsid proteins. *Nucleic Acids Res.* 34, 593–605. doi: 10.1093/nar/gkj458

Daigle, N., and Ellenberg, J. (2007). LambdaN-GFP: an RNA reporter system for live-cell imaging. *Nat. Methods* 4, 633–636. doi: 10.1038/nmeth1065

Damgaard, C. K., Andersen, E. S., Knudsen, B., Gorodkin, J., and Kjems, J. (2004). RNA interactions in the 5' region of the HIV-1 genome. *J. Mol. Biol.* 336, 369–379. doi: 10.1016/j.jmb.2003.12.010

Darlix, J. L., Gabus, C., Nugeyre, M. T., Clavel, F., and Barré-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. doi: 10.1016/0022-2836(90)90392-Y

Das, A. T., Harwig, A., Vrolijk, M. M., and Berkhout, B. (2007). The TAR hairpin of human immunodeficiency virus type 1 can be deleted when not required for Tat-mediated activation of transcription. *J. Virol.* 81, 7742–7748. doi: 10.1128/JVI.00392-07

Das, A. T., Vrolijk, M. M., Harwig, A., and Berkhout, B. (2012). Opening of the TAR hairpin in the HIV-1 genome causes aberrant RNA dimerization and packaging. *Retrovirology* 9:59. doi: 10.1186/1742-4690-9-59

De Tapia, M., Metzler, V., Mougel, M., Ehresmann, B., and Ehresmann, C. (1998). Dimerization of MoMuLV genomic RNA: redefinition of the role of the palindromic stem-loop H1 (278-303) and new roles for stem-loops H2 (310-352) and H3 (355-374). *Biochemistry* 37, 6077–6085. doi: 10.1021/bi9800303

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. doi: 10.1074/jbc.M103462200

D'Souza, V., Dey, A., Habib, D., and Summers, M. F. (2004). NMR structure of the 101-nucleotide core encapsidation signal of the Moloney murine leukemia virus. *J. Mol. Biol.* 337, 427–442. doi: 10.1016/j.jmb.2004.01.037

D'Souza, V., and Summers, M. F. (2005). How retroviruses select their genomes. *Nat. Rev. Microbiol.* 3, 643–655. doi: 10.1038/nrmicro1210

Duesberg, P. H. (1968). Physical properties of Rous sarcoma virus RNA. *Proc. Natl. Acad. Sci. U.S.A.* 60, 1511–1518. doi: 10.1073/pnas.60.4.1511

Ennifar, E., Aslam, M. W., Strasser, P., Hoffmann, G., Dumas, P., and van Delft, F. L. (2013). Structure-guided discovery of a novel aminoglycoside conjugate targeting HIV-1 RNA viral genome. *ACS Chem. Biol.* 8, 2509–2517. doi: 10.1021/cb400498n

Ennifar, E., Paillart, J.-C., Bernacchi, S., Walter, P., Pale, P., Decout, J.-L., et al. (2007). A structure-based approach for targeting the HIV-1 genomic RNA dimerization initiation site. *Biochimie* 89, 1195–1203. doi: 10.1016/j.biochi.2007.03.003

Ennifar, E., Paillart, J.-C., Bodlenner, A., Walter, P., Weibel, J.-M., Aubertin, A.-M., et al. (2006). Targeting the dimerization initiation site of HIV-1 RNA with aminoglycosides: from crystal to cell. *Nucleic Acids Res.* 34, 2328–2339. doi: 10.1093/nar/gkl317

Ennifar, E., Paillart, J.-C., Marquet, R., Ehresmann, B., Ehresmann, C., Dumas, P., et al. (2003). HIV-1 RNA dimerization initiation site is structurally similar to the ribosomal A site and binds aminoglycoside antibiotics. *J. Biol. Chem.* 278, 2723–2730. doi: 10.1074/jbc.M205726200

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. doi: 10.1038/nsb727

Ennifar, E., Yusupov, M., Walter, P., Marquet, R., Ehresmann, B., Ehresmann, C., et al. (1999). The crystal structure of the dimerization initiation site of genomic HIV-1 RNA reveals an extended duplex with two adenine bulges. *Structure* 1993, 1439–1449. doi: 10.1016/S0969-2126(00)80033-7

Erlwein, O., Cain, D., Fischer, N., Rethwilm, A., and McClure, M. O. (1997). Identification of sites that act together to direct dimerization of human foamy virus RNA in vitro. *Virology* 229, 251–258. doi: 10.1006/viro.1997.8438

Feng, Y. X., Copeland, T. D., Henderson, L. E., Gorelick, R. J., Bosche, W. J., Levin, J. G., et al. (1996). HIV-1 nucleocapsid protein induces "maturation" of dimeric retroviral RNA in vitro. *Proc. Natl. Acad. Sci. U.S.A.* 93, 7577–7581. doi: 10.1073/pnas.93.15.7577

Ferrer, M., Clerté, C., Chamontin, C., Basyuk, E., Lainé, S., Hottin, J., et al. (2016). Imaging HIV-1 RNA dimerization in cells by multicolor super-resolution and fluctuation microscopies. *Nucleic Acids Res.* 44, 7922–7934. doi: 10.1093/nar/gkw511

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. doi: 10.1128/JVI.78.22.12129-12139.2004

Fossé, P., Motté, N., Roumier, A., Gabus, C., Muriaux, D., Darlix, J. L., et al. (1996). A short autocomplementary sequence plays an essential role in avian sarcoma-leukosis virus RNA dimerization. *Biochemistry* 35, 16601–16609. doi: 10.1021/bi9613786

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.

Fu, W., and Rein, A. (1993). Maturation of dimeric viral RNA of Moloney murine leukemia virus. *J. Virol.* 67, 5443–5449.

Fusco, D., Accornero, N., Lavoie, B., Shenoy, S. M., Blanchard, J.-M., Singer, R. H., et al. (2003). Single mRNA molecules demonstrate probabilistic movement in living mammalian cells. *Curr. Biol.* 13, 161–167. doi: 10.1016/S0960-9822(02)01436-7

Galetto, R., and Negroni, M. (2005). Mechanistic features of recombination in HIV. *AIDS Rev.* 7, 92–102.

Garzino-Demo, A., Gallo, R. C., and Arya, S. K. (1995). Human immunodeficiency virus type 2 (HIV-2): packaging signal and associated negative regulatory element. *Hum. Gene Ther.* 6, 177–184. doi: 10.1089/hum.1995.6.2-177

Gherghe, C., Leonard, C. W., Gorelick, R. J., and Weeks, K. M. (2010). Secondary structure of the mature ex virio Moloney murine leukemia virus genomic RNA dimerization domain. *J. Virol.* 84, 898–906. doi: 10.1128/JVI.01602-09

Gherghe, C., and Weeks, K. M. (2006). The SL1-SL2 (stem-loop) domain is the primary determinant for stability of the gamma retroviral genomic RNA dimer. *J. Biol. Chem.* 281, 37952–37961. doi: 10.1074/jbc.M607380200

Girard, P. M., Bonnet-Mathonière, B., Muriaux, D., and Paoletti, J. (1995). A short autocomplementary sequence in the 5' leader region is responsible for dimerization of MoMuLV genomic RNA. *Biochemistry* 34, 9785–9794. doi: 10.1021/bi00030a016

Girard, P. M., de Rocquigny, H., Roques, B. P., and Paoletti, J. (1996). A model of PSI dimerization: destabilization of the C278-G303 stem-loop by the nucleocapsid protein (NCp10) of MoMuLV. *Biochemistry* 35, 8705–8714. doi: 10.1021/bi952454s

Gorelick, R. J., Benveniste, R. E., Gagliardi, T. D., Wiltrout, T. A., Busch, L. K., Bosche, W. J., et al. (1999). Nucleocapsid protein zinc-finger mutants of simian immunodeficiency virus strain mne produce virions that are replication defective in vitro and in vivo. *Virology* 253, 259–270. doi: 10.1006/viro.1998.9513

Gorelick, R. J., Chabot, D. J., Ott, D. E., Gagliardi, T. D., Rein, A., Henderson, L. E., et al. (1996). Genetic analysis of the zinc finger in the Moloney murine leukemia virus nucleocapsid domain: replacement of zinc-coordinating residues with other zinc-coordinating residues yields noninfectious particles containing genomic RNA. *J. Virol.* 70, 2593–2597.

Greatorex, J., Gallego, J., Varani, G., and Lever, A. (2002). Structure and stability of wild-type and mutant RNA internal loops from the SL-1 domain of the HIV-1 packaging signal. *J. Mol. Biol.* 322, 543–557. doi: 10.1016/S0022-2836(02)00776-3

Greatorex, J. S., Laisse, V., Dockhelar, M. C., and Lever, A. M. (1996). Sequences involved in the dimerisation of human T cell leukaemia virus type-1 RNA. *Nucleic Acids Res.* 24, 2919–2923. doi: 10.1093/nar/24.15.2919

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. doi: 10.1128/JVI.75.24.12058-12069.2001

Grohman, J. K., Gorelick, R. J., Kottegoda, S., Allbritton, N. L., Rein, A., and Weeks, K. M. (2014). An immature retroviral RNA genome resembles a kinetically trapped intermediate state. *J. Virol.* 88, 6061–6068. doi: 10.1128/JVI.03277-13

Hagan, N. A., and Fabris, D. (2007). Dissecting the protein-RNA and RNA-RNA interactions in the nucleocapsid-mediated dimerization and isomerization of HIV-1 stemloop 1. *J. Mol. Biol.* 365, 396–410. doi: 10.1016/j.jmb.2006.09.081

Harrison, G. P., Miele, G., Hunter, E., and Lever, A. M. (1998). Functional analysis of the core human immunodeficiency virus type 1 packaging signal in a permissive cell line. *J. Virol.* 72, 5886–5896.

Heng, X., Kharytonchyk, S., Garcia, E. L., Lu, K., Divakaruni, S. S., LaCotti, C., et al. (2012). Identification of a Minimal Region of the HIV-1 5'-Leader required for RNA Dimerization, NC Binding and Packaging. *J. Mol. Biol.* 417, 224–239. doi: 10.1016/j.jmb.2012.01.033

Hill, M. K., Shehu-Xhilaga, M., Campbell, S. M., Poumbourios, P., Crowe, S. M., and Mak, J. (2003). The dimer initiation sequence stem-loop of human immunodeficiency virus type 1 is dispensable for viral replication in peripheral blood mononuclear cells. *J. Virol.* 77, 8329–8335. doi: 10.1128/JVI.77.15.8329-8335.2003

Höglund, S., Ohagen, A., Goncalves, J., Panganiban, A. T., and Gabuzda, D. (1997). Ultrastructure of HIV-1 genomic RNA. *Virology* 233, 271–279. doi: 10.1006/viro.1997.8585

Houzet, L., Paillart, J. C., Smagulova, F., Maurel, S., Morichaud, Z., Marquet, R., et al. (2007). HIV controls the selective packaging of genomic, spliced viral and cellular RNAs into virions through different mechanisms. *Nucleic Acids Res.* 35, 2695–2704. doi: 10.1093/nar/gkm153

Hu, W. S., and Temin, H. M. (1990). Retroviral recombination and reverse transcription. *Science* 250, 1227–1233. doi: 10.1126/science.1700865

Hussein, I. T. M., Ni, N., Galli, A., Chen, J., Moore, M. D., and Hu, W.-S. (2010). Delineation of the preferences and requirements of the human immunodeficiency virus type 1 dimerization initiation signal by using an in vivo cell-based selection approach. *J. Virol.* 84, 6866–6875. doi: 10.1128/JVI.01930-09

Huthoff, H., and Berkhout, B. (2001). Two alternating structures of the HIV-1 leader RNA. *RNA* 7, 143–157. doi: 10.1017/S1355838201001881

Jaballah, S. A., Aktar, S. J., Ali, J., Phillip, P. S., Al Dhaheri, N. S., Jabeen, A., et al. (2010). A G-C-rich palindromic structural motif and a stretch of single-stranded purines are required for optimal packaging of Mason-Pfizer monkey virus (MPMV) genomic RNA. *J. Mol. Biol.* 401, 996–1014. doi: 10.1016/j.jmb.2010.06.043

Jalalirad, M., and Laughrea, M. (2010). Formation of immature and mature genomic RNA dimers in wild-type and protease-inactive HIV-1: Differential roles of the Gag polyprotein, nucleocapsid proteins NCp15, NCp9, NCp7, and the dimerization initiation site. *Virology* 407, 225–236. doi: 10.1016/j.virol.2010.08.013

Jalalirad, M., Saadatmand, J., and Laughrea, M. (2012). Dominant role of the 5' TAR bulge in dimerization of HIV-1 genomic RNA, but no evidence of TAR-TAR kissing during in vivo virus assembly. *Biochemistry* 51, 3744–3758. doi: 10.1021/bi300111p

James, L., and Sargueil, B. (2008). RNA secondary structure of the feline immunodeficiency virus 5'UTR and Gag coding region. *Nucleic Acids Res.* 36, 4653–4666. doi: 10.1093/nar/gkn447

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. doi: 10.1074/jbc.M008642200

Jouvenet, N., Bieniasz, P. D., and Simon, S. M. (2008). Imaging the biogenesis of individual HIV-1 virions in live cells. *Nature* 454, 236–240. doi: 10.1038/nature06998

Jouvenet, N., Simon, S. M., and Bieniasz, P. D. (2009). Imaging the interaction of HIV-1 genomes and Gag during assembly of individual viral particles. *Proc. Natl. Acad. Sci. U.S.A.* 106, 19114–19119. doi: 10.1073/pnas.0907364106

Kalloush, R. M., Vivet-Boudou, V., Ali, L. M., Mustafa, F., Marquet, R., and Rizvi, T. A. (2016). Packaging of Mason-Pfizer monkey virus (MPMV) genomic RNA depends upon conserved long-range interactions (LRIs) between U5 and gag sequences. *RNA* 22, 905–919. doi: 10.1261/rna.055731.115

Katoh, I., Kyushiki, H., Sakamoto, Y., Ikawa, Y., and Yoshinaka, Y. (1991). Bovine leukemia virus matrix-associated protein MA(p15): further processing and formation of a specific complex with the dimer of the 5'-terminal genomic RNA fragment. *J. Virol.* 65, 6845–6855.

Katoh, I., Yasunaga, T., and Yoshinaka, Y. (1993). Bovine leukemia virus RNA sequences involved in dimerization and specific gag protein binding: close relation to the packaging sites of avian, murine, and human retroviruses. *J. Virol.* 67, 1830–1839.

Keane, S. C., Heng, X., Lu, K., Kharytonchyk, S., Ramakrishnan, V., Carter, G., et al. (2015). RNA structure. Structure of the HIV-1 RNA packaging signal. *Science* 348, 917–921. doi: 10.1126/science.aaa9266

Keane, S. C., Van, V., Frank, H. M., Sciandra, C. A., McCowin, S., Santos, J., et al. (2016). NMR detection of intermolecular interaction sites in the dimeric 5'-leader of the HIV-1 genome. *Proc. Natl. Acad. Sci. U.S.A.* 113, 13033–13038. doi: 10.1073/pnas.1614785113

Kenyon, J. C., Ghazawi, A., Cheung, W. K. S., Phillip, P. S., Rizvi, T. A., and Lever, A. M. L. (2008). The secondary structure of the 5' end of the FIV genome reveals a long-range interaction between R/U5 and gag sequences, and a large, stable stem-loop. *RNA* 14, 2597–2608. doi: 10.1261/rna.1284908

Kenyon, J. C., Tanner, S. J., Legiewicz, M., Phillip, P. S., Rizvi, T. A., Le Grice, S. F. J., et al. (2011). SHAPE analysis of the FIV Leader RNA reveals a structural switch potentially controlling viral packaging and genome dimerization. *Nucleic Acids Res.* 39, 6692–6704. doi: 10.1093/nar/gkr252

Kharytonchyk, S., Monti, S., Smaldino, P. J., Van, V., Bolden, N. C., Brown, J. D., et al. (2016). Transcriptional start site heterogeneity modulates the structure and function of the HIV-1 genome. *Proc. Natl. Acad. Sci. U.S.A.* 113, 13378–13383. doi: 10.1073/pnas.1616627113

Kharytonchyk, S., and Pedersen, F. S. (2010). A unique, thermostable dimer linkage structure of RD114 retroviral RNA. *RNA* 16, 572–584. doi: 10.1261/rna.1495110

Kharytonchyk, S. A., Kireyeva, A. I., Osipovich, A. B., and Fomin, I. K. (2005). Evidence for preferential copackaging of Moloney murine leukemia virus genomic RNAs transcribed in the same chromosomal site. *Retrovirology* 2:3. doi: 10.1186/1742-4690-2-3

Kieken, F., Paquet, F., Brulé, F., Paoletti, J., and Lancelot, G. (2006). A new NMR solution structure of the SL1 HIV-1Lai loop-loop dimer. *Nucleic Acids Res.* 34, 343–352. doi: 10.1093/nar/gkj427

Kim, C. H., and Tinoco, I. (2000). A retroviral RNA kissing complex containing only two G.C base pairs. *Proc. Natl. Acad. Sci. U.S.A.* 97, 9396–9401. doi: 10.1073/pnas.170283697

Konings, D. A., Nash, M. A., Maizel, J. V., and Arlinghaus, R. B. (1992). Novel GACG-hairpin pair motif in the 5' untranslated region of type C retroviruses related to murine leukemia virus. *J. Virol.* 66, 632–640.

Kung, H. J., Bailey, J. M., Davidson, N., Nicolson, M. O., and McAllister, R. M. (1975). Structure, subunit composition, and molecular weight of RD-114 RNA. *J. Virol.* 16, 397–411.

Kung, H. J., Hu, S., Bender, W., Bailey, J. M., Davidson, N., Nicolson, M. O., et al. (1976). RD-114, baboon, and woolly monkey viral RNA's compared in size and structure. *Cell* 7, 609–620. doi: 10.1016/0092-8674(76)90211-7

Kutluay, S. B., and Bieniasz, P. D. (2010). Analysis of the initiating events in HIV-1 particle assembly and genome packaging. *PLoS Pathog.* 6:e1001200. doi: 10.1371/journal.ppat.1001200

Kutluay, S. B., Zang, T., Blanco-Melo, D., Powell, C., Jannain, D., Errando, M., et al. (2014). Global changes in the RNA binding specificity of HIV-1 gag regulate virion genesis. *Cell* 159, 1096–1109. doi: 10.1016/j.cell.2014.09.057

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., 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. doi: 10.1016/S0022-2836(02)00369-8

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. doi: 10.1128/JVI.02025-06

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. doi: 10.1021/bi0271190

Laughrea, M., and Jetté, 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. doi: 10.1021/bi00249a035

Laughrea, M., and Jetté, 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. doi: 10.1021/bi960395s

Laughrea, M., Jetté, L., Mak, J., Kleiman, L., Liang, C., and Wainberg, M. A. (1997). Mutations in the kissing-loop hairpin of human immunodeficiency virus type 1 reduce viral infectivity as well as genomic RNA packaging and dimerization. *J. Virol.* 71, 3397–3406.

Laughrea, M., Shen, N., Jetté, L., and Wainberg, M. A. (1999). Variant effects of non-native 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. doi: 10.1021/bi981728j

Le Blanc, I., Greatorex, J., Dokhélar, M. C., and Lever, A. M. (2000). A 37 base sequence in the leader region of human T-cell leukaemia virus type I is a high affinity dimerization site but is not essential for virus replication. *J. Gen. Virol.* 81, 105–108. doi: 10.1099/0022-1317-81-1-105

Lear, A. L., Haddrick, M., and Heaphy, S. (1995). A study of the dimerization of Rous sarcoma virus RNA in vitro and in vivo. *Virology* 212, 47–57. doi: 10.1006/viro.1995.1452

Lee, E. G., Yeo, A., Kraemer, B., Wickens, M., and Linial, M. L. (1999). The gag domains required for avian retroviral RNA encapsidation determined by using two independent assays. *J. Virol* 73, 6282–6292.

Lever, A., Gottlinger, H., Haseltine, W., and Sodroski, J. (1989). Identification of a sequence required for efficient packaging of human immunodeficiency virus type 1 RNA into virions. *J. Virol.* 63, 4085–4087.

Levin, J. G., Guo, J., Rouzina, I., and Musier-Forsyth, K. (2005). Nucleic acid chaperone activity of HIV-1 nucleocapsid protein: critical role in reverse transcription and molecular mechanism. *Prog. Nucleic Acid Res. Mol. Biol.* 80, 217–286. doi: 10.1016/S0079-6603(05)80006-6

L'Hernault, A., Greatorex, J. S., Crowther, R. A., and Lever, A. M. L. (2007). Dimerisation of HIV-2 genomic RNA is linked to efficient RNA packaging, normal particle maturation and viral infectivity. *Retrovirology* 4:90. doi: 10.1186/1742-4690-4-90

L'Hernault, A., Weiss, E. U., Greatorex, J. S., and Lever, A. M. (2012). HIV-2 genome dimerization is required for the correct processing of Gag: a second-site reversion in matrix can restore both processes in dimerization-impaired mutant viruses. *J. Virol.* 86, 5867–5876. doi: 10.1128/JVI.00124-12

Linial, M. L. (1999). Foamy viruses are unconventional retroviruses. *J. Virol.* 73, 1747–1755.

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. doi: 10.1017/S1355838200000698

Lodmell, J. S., Ehresmann, C., Ehresmann, B., and Marquet, R. (2001). Structure and dimerization of HIV-1 kissing loop aptamers. *J. Mol. Biol.* 311, 475–490. doi: 10.1006/jmbi.2001.4879

Lu, K., Heng, X., Garyu, L., Monti, S., Garcia, E. L., Kharytonchyk, S., et al. (2011). NMR detection of structures in the HIV-1 5'-leader RNA that regulate genome packaging. *Science* 334, 242–245. doi: 10.1126/science.1210460

Ly, H., and Parslow, T. G. (2002). Bipartite signal for genomic RNA dimerization in Moloney murine leukemia virus. *J. Virol.* 76, 3135–3144. doi: 10.1128/JVI.76.7.3135-3144.2002

Mailler, E., Bernacchi, S., Marquet, R., Paillart, J.-C., Vivet-Boudou, V., and Smyth, R. P. (2016). The life-cycle of the HIV-1 Gag–RNA complex. *Viruses* 8:E248. doi: 10.3390/v8090248

Maisel, J., Bender, W., Hu, S., Duesberg, P. H., and Davidson, N. (1978). Structure of 50 to 70S RNA from Moloney sarcoma viruses. *J. Virol.* 25, 384–394.

Mangel, W. F., Delius, H., and Duesberg, P. H. (1974). Structure and molecular weight of the 60-70S RNA and the 30-40S RNA of the Rous sarcoma virus. *Proc. Natl. Acad. Sci. U.S.A.* 71, 4541–4545. doi: 10.1073/pnas.71.11.4541

Mansky, L. M., Krueger, A. E., and Temin, H. M. (1995). The bovine leukemia virus encapsidation signal is discontinuous and extends into the 5' end of the gag gene. *J. Virol.* 69, 3282–3289.

Marquet, R., Baudin, F., Gabus, C., Darlix, J. L., Mougel, M., Ehresmann, C., et al. (1991). Dimerization of human immunodeficiency virus (type 1) RNA: stimulation by cations and possible mechanism. *Nucleic Acids Res.* 19, 2349–2357. doi: 10.1093/nar/19.9.2349

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. doi: 10.1093/nar/22.2.145

Maurel, S., Houzet, L., Garcia, E. L., Telesnitsky, A., and Mougel, M. (2007). Characterization of a natural heterodimer between MLV genomic RNA and the SD' retroelement generated by alternative splicing. *RNA* 13, 2266–2276. doi: 10.1261/rna.713807

Maurel, S., and Mougel, M. (2010). Murine leukemia virus RNA dimerization is coupled to transcription and splicing processes. *Retrovirology* 7:64. doi: 10.1186/1742-4690-7-64

McBride, M. S., and Panganiban, A. T. (1997). Position dependence of functional hairpins important for human immunodeficiency virus type 1 RNA encapsidation in vivo. *J. Virol.* 71, 2050–2058.

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. U.S.A.* 101, 1189–1194. doi: 10.1073/pnas.0307966100

Mikkelsen, J. G., Lund, A. H., Duch, M., and Pedersen, F. S. (1998a). Recombination in the 5' leader of murine leukemia virus is accurate and influenced by sequence identity with a strong bias toward the kissing-loop dimerization region. *J. Virol.* 72, 6967–6978.

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. doi: 10.1128/JVI.74.2.600-610.2000

Mikkelsen, J. G., Lund, A. H., Dybkaer, K., Duch, M., and Pedersen, F. S. (1998b). Extended minus-strand DNA as template for R-U5-mediated second-strand transfer in recombinational rescue of primer binding site-modified retroviral vectors. *J. Virol.* 72, 2519–2525.

Mikkelsen, J. G., and Pedersen, F. S. (2000). Genetic reassortment and patch repair by recombination in retroviruses. *J. Biomed. Sci.* 7, 77–99. doi: 10.1007/BF02256615

Miyazaki, Y., Garcia, E. L., King, S. R., Iyalla, K., Loeliger, K., Starck, P., et al. (2010a). An RNA structural switch regulates diploid genome packaging by Moloney Murine Leukemia Virus. *J. Mol. Biol.* 396, 141–152. doi: 10.1016/j.jmb.2009.11.033

Miyazaki, Y., Irobalieva, R. N., Tolbert, B. S., Smalls-Mantey, A., Iyalla, K., Loeliger, K., et al. (2010b). Structure of a conserved retroviral RNA packaging element by NMR spectroscopy and cryo-electron tomography. *J. Mol. Biol.* 404, 751–772. doi: 10.1016/j.jmb.2010.09.009

Monie, T., Greatorex, J., and Lever, A. M. (2001). Oligonucleotide mapping of the core genomic RNA dimer linkage in human T-cell leukaemia virus type-1. *Virus Res.* 78, 45–56. doi: 10.1016/S0168-1702(01)00283-0

Monie, T. P., Greatorex, J. S., Zacharias, M., and Lever, A. M. L. (2004). The human T-cell lymphotropic virus type-I dimerization initiation site forms a hairpin loop, unlike previously characterized retroviral dimerization motifs. *Biochemistry* 43, 6085–6090. doi: 10.1021/bi030237i

Montagnier, L., Goldé, A., and Vigier, P. (1969). A possible subunit structure of Rous sarcoma virus RNA. *J. Gen. Virol.* 4, 449–452. doi: 10.1099/0022-1317-4-3-449

Moore, M. D., Fu, W., Nikolaitchik, O., Chen, J., Ptak, R. G., and Hu, W.-S. (2007). Dimer initiation signal of human immunodeficiency virus type 1: its role in partner selection during RNA copackaging and its effects on recombination. *J. Virol.* 81, 4002–4011. doi: 10.1128/JVI.02589-06

Moore, M. D., and Hu, W. S. (2009). HIV-1 RNA dimerization: it takes two to tango. *AIDS Rev.* 11, 91–102.

Moore, M. D., Nikolaitchik, O. A., Chen, J., Hammarskjöld, M.-L., Rekosh, D., and Hu, W.-S. (2009). Probing the HIV-1 genomic RNA trafficking pathway and dimerization by genetic recombination and single virion analyses. *PLoS Pathog.* 5:e1000627. doi: 10.1371/journal.ppat.1000627

Morris, A., Marsden, M., Halcrow, K., Hughes, E. S., Brettle, R. P., Bell, J. E., et al. (1999). Mosaic structure of the human immunodeficiency virus type 1 genome infecting lymphoid cells and the brain: evidence for frequent in vivo recombination events in the evolution of regional populations. *J. Virol.* 73, 8720–8731.

Mougel, M., and Barklis, E. (1997). A role for two hairpin structures as a core RNA encapsidation signal in murine leukemia virus virions. *J. Virol.* 71, 8061–8065.

Mougel, M., Tounekti, N., Darlix, J. L., Paoletti, J., Ehresmann, B., and Ehresmann, C. (1993). Conformational analysis of the 5' leader and the gag initiation site of Mo-MuLV RNA and allosteric transitions induced by dimerization. *Nucleic Acids Res.* 21, 4677–4684. doi: 10.1093/nar/21.20.4677

Mougel, M., Zhang, Y., and Barklis, E. (1996). cis-active structural motifs involved in specific encapsidation of Moloney murine leukemia virus RNA. *J. Virol.* 70, 5043–5050.

Mueller, N., Das, A. T., and Berkhout, B. (2016). A phylogenetic survey on the structure of the HIV-1 leader RNA domain that encodes the splice donor signal. *Viruses* 8:E200. doi: 10.3390/v8070200

Mujeeb, A., Ulyanov, N. B., Georgantis, S., Smirnov, I., Chung, J., Parslow, T. G., et al. (2007). Nucleocapsid protein-mediated maturation of dimer initiation complex of full-length SL1 stemloop of HIV-1: sequence effects and mechanism of RNA refolding. *Nucleic Acids Res.* 35, 2026–2034. doi: 10.1093/nar/gkm097

Mundigala, H., Michaux, J. B., Feig, A. L., Ennifar, E., and Rueda, D. (2014). HIV-1 DIS stem loop forms an obligatory bent kissing intermediate in the dimerization pathway. *Nucleic Acids Res.* 42, 7281–7289. doi: 10.1093/nar/gku332

Muriaux, D., Girard, P.-M., Bonnet-Mathonière, 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. doi: 10.1074/jbc.270.14.8209

Muriaux, D., Rocquigny, H. D., Roques, B.-P., and Paoletti, J. (1996). NCp7 activates HIV-1Lai RNA dimerization by converting a transient loop-loop complex into a stable dimer. *J. Biol. Chem.* 271, 33686–33692. doi: 10.1074/jbc.271.52.33686

Murti, K. G., Bondurant, M., and Tereba, A. (1981). Secondary structural features in the 70S RNAs of Moloney murine leukemia and Rous sarcoma viruses as observed by electron microscopy. *J. Virol.* 37, 411–419.

Ni, N., Nikolaitchik, O. A., Dilley, K. A., Chen, J., Galli, A., Fu, W., et al. (2011). Mechanisms of human immunodeficiency virus type 2 RNA packaging: efficient trans packaging and selection of RNA copackaging partners. *J. Virol.* 85, 7603–7612. doi: 10.1128/JVI.00562-11

Ohishi, M., Nakano, T., Sakuragi, S., Shioda, T., Sano, K., and Sakuragi, J. (2011). The relationship between HIV-1 genome RNA dimerization, virion maturation and infectivity. *Nucleic Acids Res.* 39, 3404–3417. doi: 10.1093/nar/gkq1314

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. doi: 10.1128/JVI.78.19.10814-10819.2004

Oroudjev, E. M., Kang, P. C., and Kohlstaedt, L. A. (1999). An additional dimer linkage structure in Moloney murine leukemia virus RNA. *J. Mol. Biol.* 291, 603–613. doi: 10.1006/jmbi.1999.2984

Paillart, J. C., Berthoux, L., Ottmann, M., Darlix, J. L., Marquet, R., Ehresmann, B., et al. (1996a). A dual role of the putative RNA dimerization initiation site of human immunodeficiency virus type 1 in genomic RNA packaging and proviral DNA synthesis. *J. Virol.* 70, 8348–8354.

Paillart, J.-C., Dettenhofer, M., Yu, X., Ehresmann, C., Ehresmann, B., and Marquet, R. (2004). First snapshots of the HIV-1 RNA structure in infected cells and in virions. *J. Biol. Chem.* 279, 48397–48403. doi: 10.1074/jbc.M408294200

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., Marquet, R., Skripkin, E., Ehresmann, C., and Ehresmann, B. (1996b). Dimerization of retroviral genomic RNAs: structural and functional implications. *Biochimie* 78, 639–653.

Paillart, J. C., Skripkin, E., Ehresmann, B., Ehresmann, C., and Marquet, R. (1996c). A loop-loop "kissing" complex is the essential part of the dimer linkage of genomic HIV-1 RNA. *Proc. Natl. Acad. Sci. U.S.A.* 93, 5572–5577.

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. doi: 10.1006/jmbi.1997.1096

Pallesen, J. (2011). Structure of the HIV-1 5' untranslated region dimer alone and in complex with gold nanocolloids: support of a TAR-TAR-containing 5' dimer linkage site (DLS) and a 3' DIS-DIS-containing DLS. *Biochemistry* 50, 6170–6177. doi: 10.1021/bi200488h

Panganiban, A. T., and Fiore, D. (1988). Ordered interstrand and intrastrand DNA transfer during reverse transcription. *Science* 241, 1064–1069. doi: 10.1126/science.2457948

Paoletti, J., Mougel, M., Tounekti, N., Girard, P. M., Ehresmann, C., and Ehresmann, B. (1993). Spontaneous dimerization of retroviral MoMuLV RNA. *Biochimie* 75, 681–686. doi: 10.1016/0300-9084(93)90099-E

Parent, L. J., Cairns, T. M., Albert, J. A., Wilson, C. B., Wills, J. W., and Craven, R. C. (2000). RNA dimerization defect in a Rous sarcoma virus matrix mutant. *J. Virol.* 74, 164–172. doi: 10.1128/JVI.74.1.164-172.2000

Park, J., and Mergia, A. (2000). Mutational analysis of the 5' leader region of simian foamy virus type 1. *Virology* 274, 203–212. doi: 10.1006/viro.2000.0423

Polge, E., Darlix, J. L., Paoletti, J., and Fossé, P. (2000). Characterization of loose and tight dimer forms of avian leukosis virus RNA. *J. Mol. Biol.* 300, 41–56. doi: 10.1006/jmbi.2000.3832

Prats, A. C., Roy, C., Wang, P. A., Erard, M., Housset, V., Gabus, C., et al. (1990). cis elements and trans-acting factors involved in dimer formation of murine leukemia virus RNA. *J. Virol.* 64, 774–783.

Purzycka, K. J., Pachulska-Wieczorek, K., and Adamiak, R. W. (2011). The in vitro loose dimer structure and rearrangements of the HIV-2 leader RNA. *Nucleic Acids Res.* 39, 7234–7248. doi: 10.1093/nar/gkr385

Rambaut, A., Posada, D., Crandall, K. A., and Holmes, E. C. (2004). The causes and consequences of HIV evolution. *Nat. Rev. Genet.* 5, 52–61. doi: 10.1038/nrg1246

Rasmussen, S. V., and Pedersen, F. S. (2006). Co-localization of gammaretroviral RNAs at their transcription site favours co-packaging. *J. Gen. Virol.* 87, 2279–2289. doi: 10.1099/vir.0.81759-0

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. doi: 10.1021/bi0267240

Roy, C., Tounekti, N., Mougel, M., Darlix, J. L., Paoletti, C., Ehresmann, C., et al. (1990). An analytical study of the dimerization of in vitro generated RNA of Moloney murine leukemia virus MoMuLV. *Nucleic Acids Res.* 18, 7287–7292. doi: 10.1093/nar/18.24.7287

Russell, R. S., Hu, J., Bériault, V., Mouland, A. J., Kleiman, L., Wainberg, M. A., et al. (2003). Sequences downstream of the 5' Splice donor site are required

for both packaging and dimerization of human immunodeficiency virus type 1 RNA. *J. Virol.* 77, 84–96. doi: 10.1128/JVI.77.1.84-96.2003

Sakuragi, J., Ode, H., Sakuragi, S., Shioda, T., and Sato, H. (2012). A proposal for a new HIV-1 DLS structural model. *Nucleic Acids Res.* 40, 5012–5022. doi: 10.1093/nar/gks156

Sakuragi, J., Shioda, T., and Panganiban, A. T. (2001). Duplication of the primary encapsidation and dimer linkage region of human immunodeficiency virus type 1 RNA results in the appearance of monomeric RNA in virions. *J. Virol.* 75, 2557–2565. doi: 10.1128/JVI.75.6.2557-2565.2001

Sakuragi, J.-I., Sakuragi, S., and Shioda, T. (2007). Minimal region sufficient for genome dimerization in the human immunodeficiency virus type 1 virion and its potential roles in the early stages of viral replication. *J. Virol.* 81, 7985–7992. doi: 10.1128/JVI.00429-07

Sakuragi, S., Shioda, T., and Sakuragi, J. (2015). Properties of human immunodeficiency virus type 1 reverse transcriptase recombination upon infection. *J. Gen. Virol.* 96, 3382–3388. doi: 10.1099/jgv.0.000265

Sakuragi, S., Yokoyama, M., Shioda, T., Sato, H., and Sakuragi, J.-I. (2016). SL1 revisited: functional analysis of the structure and conformation of HIV-1 genome RNA. *Retrovirology* 13:79. doi: 10.1186/s12977-016-0310-9

Sardo, L., Hatch, S. C., Chen, J., Nikolaitchik, O., Burdick, R. C., Chen, D., et al. (2015). Dynamics of HIV-1 RNA near the plasma membrane during virus assembly. *J. Virol.* 89, 10832–10840. doi: 10.1128/JVI.01146-15

Schlub, T. E., Grimm, A. J., Smyth, R. P., Cromer, D., Chopra, A., Mallal, S., et al. (2014). Fifteen to twenty percent of HIV substitution mutations are associated with recombination. *J. Virol.* 88, 3837–3849. doi: 10.1128/JVI.03136-13

Schlub, T. E., Smyth, R. P., Grimm, A. J., Mak, J., and Davenport, M. P. (2010). Accurately measuring recombination between closely related HIV-1 genomes. *PLoS Comput. Biol.* 6:e1000766. doi: 10.1371/journal.pcbi.1000766

Schwartz, D. E., Tizard, R., and Gilbert, W. (1983). Nucleotide sequence of Rous sarcoma virus. *Cell* 32, 853–869. doi: 10.1016/0092-8674(83)90071-5

Shehu-Xhilaga, M., Crowe, S. M., and Mak, J. (2001). Maintenance of the Gag/Gag-Pol ratio is important for human immunodeficiency virus type 1 RNA dimerization and viral infectivity. *J. Virol.* 75, 1834–1841. doi: 10.1128/JVI.75.4.1834-1841.2001

Shen, N., Jetté, 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. doi: 10.1128/JVI.74.12.5729-5735.2000

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. U.S.A.* 91, 4945–4949. doi: 10.1073/pnas.91.11.4945

Smagulova, F., Maurel, S., Morichaud, Z., Devaux, C., Mougel, M., and Houzet, L. (2005). The highly structured encapsidation signal of MuLV RNA is involved in the nuclear export of its unspliced RNA. *J. Mol. Biol.* 354, 1118–1128. doi: 10.1016/j.jmb.2005.10.021

Smyth, R. P., Despons, L., Huili, G., Bernacchi, S., Hijnen, M., Mak, J., et al. (2015). Mutational interference mapping experiment (MIME) for studying RNA structure and function. *Nat. Methods* 12, 866–872. doi: 10.1038/nmeth.3490

Smyth, R. P., Schlub, T. E., Grimm, A. J., Waugh, C., Ellenberg, P., Chopra, A., et al. (2014). Identifying recombination hot spots in the HIV-1 genome. *J. Virol.* 88, 2891–2902. doi: 10.1128/JVI.03014-13

Song, R., Kafaie, J., and Laughrea, M. (2008). Role of the 5' TAR stem–loop and the U5-AUG duplex in dimerization of HIV-1 genomic RNA. *Biochemistry* 47, 3283–3293. doi: 10.1021/bi7023173

Song, R., Kafaie, J., Yang, L., and Laughrea, M. (2007). HIV-1 viral RNA is selected in the form of monomers that dimerize in a three-step protease-dependent process; the DIS of stem–loop 1 initiates viral RNA dimerization. *J. Mol. Biol.* 371, 1084–1098. doi: 10.1016/j.jmb.2007.06.010

Sun, X., Zhang, Q., and Al-Hashimi, H. M. (2007). Resolving fast and slow motions in the internal loop containing stem-loop 1 of HIV-1 that are modulated by Mg^{2+} binding: role in the kissing–duplex structural transition. *Nucleic Acids Res.* 35, 1698–1713. doi: 10.1093/nar/gkm020

Takahashi, K. I., Baba, S., Chattopadhyay, P., Koyanagi, Y., Yamamoto, N., Takaku, H., et al. (2000). Structural requirement for the two-step dimerization of human immunodeficiency virus type 1 genome. *RNA* 6, 96–102. doi: 10.1017/S1355838200991635

Torrent, C., Gabus, C., and Darlix, J. L. (1994). A small and efficient dimerization/packaging signal of rat VL30 RNA and its use in murine leukemia virus-VL30-derived vectors for gene transfer. *J. Virol.* 68, 661–667.

Tounekti, N., Mougel, M., Roy, C., Marquet, R., Darlix, J. L., Paoletti, J., et al. (1992). Effect of dimerization on the conformation of the encapsidation Psi domain of Moloney murine leukemia virus RNA. *J. Mol. Biol.* 223, 205–220. doi: 10.1016/0022-2836(92)90726-Z

Tran, T., Liu, Y., Marchant, J., Monti, S., Seu, M., Zaki, J., et al. (2015). Conserved determinants of lentiviral genome dimerization. *Retrovirology* 12:83. doi: 10.1186/s12977-015-0209-x

Ulyanov, N. B., Mujeeb, A., Du, Z., Tonelli, M., Parslow, T. G., and James, T. L. (2006). NMR structure of the full-length linear dimer of stem-loop-1 RNA in the HIV-1 dimer initiation site. *J. Biol. Chem.* 281, 16168–16177. doi: 10.1074/jbc.M601711200

van Bel, N., Das, A. T., Cornelissen, M., Abbink, T. E. M., and Berkhout, B. (2014). A short sequence motif in the 5' leader of the HIV-1 genome modulates extended RNA dimer formation and virus replication. *J. Biol. Chem.* 289, 35061–35074. doi: 10.1074/jbc.M114.621425

Vrolijk, M. M., Ooms, M., Harwig, A., Das, A. T., and Berkhout, B. (2008). Destabilization of the TAR hairpin affects the structure and function of the HIV-1 leader RNA. *Nucleic Acids Res.* 36, 4352–4363. doi: 10.1093/nar/gkn364

Whitney, J. B., and Wainberg, M. A. (2006). Impaired RNA incorporation and dimerization in live attenuated leader-variants of SIVmac239. *Retrovirology* 3:96. doi: 10.1186/1742-4690-3-96

Wilkinson, K. A., Gorelick, R. J., Vasa, S. M., Guex, N., Rein, A., Mathews, D. H., et al. (2008). High-Throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states. *PLoS Biol.* 6:e96 doi: 10.1371/journal.pbio.0060096

Yu, H., Li, T., Qiao, W., Chen, Q., and Geng, Y. (2007). Guanine tetrad and palindromic sequence play critical roles in the RNA dimerization of bovine foamy virus. *Arch. Virol.* 152, 2159–2167. doi: 10.1007/s00705-007-1047-5

Yu, S. F., Edelmann, K., Strong, R. K., Moebes, A., Rethwilm, A., and Linial, M. L. (1996). The carboxyl terminus of the human foamy virus Gag protein contains separable nucleic acid binding and nuclear transport domains. *J. Virol.* 70, 8255–8262.

Yu, S. F., Sullivan, M. D., and Linial, M. L. (1999). Evidence that the human foamy virus genome is DNA. *J. Virol.* 73, 1565–1572.

Zhuang, J., Jetzt, A. E., Sun, G., Yu, H., Klarmann, G., Ron, Y., et al. (2002). Human immunodeficiency virus type 1 recombination: rate, fidelity, and putative hot spots. *J. Virol.* 76, 11273–11282. doi: 10.1128/JVI.76.22.11273-11282.2002