Retroviral RNA dimer linkage

Jane Greatorex and Andrew Lever

Department of Medicine, University of Cambridge, Addenbrooke’s Hospital, Hills Road, Cambridge CB2 2QQ, UK

Introduction

All retroviral genomes apparently consist of two identical positive-strand molecules of RNA, making them the only known diploid viruses. The evidence for this is discussed below, but the advantages conferred by carrying two genomes are as yet unclear, although the diploid genome will permit virus recombination and template switching during reverse transcription. The conservation of the diploid genome throughout the Retroviridae argues that it has a key role in the virion life-cycle. The diploid genome appears to be physically linked (see below) and the site at which the linkage occurs is referred to as the dimer linkage site (DLS). Because of its proximity, the DLS has been proposed to contribute to the encapsidation signal of retroviruses.

The Retroviridae are a diverse family, and their individual life-cycles vary widely: thus, to assign a single generic role to the dimer linkage may be too simplistic. This review will cover the evidence for the dimeric genome, mapping of the various DLS in different retroviruses, and recent attempts to assign a role in vivo for the dimer linkage. As with many aspects of retrovirology, most data have come from studies of human immunodeficiency virus (HIV-1).

RNA dimers and dimer linkage – the evidence

The dimeric nature of the retroviral genomic RNA is supported by sedimentation analysis (Stoltzfus & Snyder, 1975) and evidence from a number of electron microscopic (EM) studies looking at viruses as diverse as HIV-1, endogenous murine ecotropic virus, endogenous murine xenotropic virus, avian reticuloendotheliosis virus, Moloney murine leukaemia virus (MoMLV) and Rous sarcoma virus (RSV) (Bender et al., 1978; Clever & Parslow, 1997; Murti et al., 1981). It has been demonstrated in several viruses, including MoMLV, HIV-1 and Harvey sarcoma virus, that when virion RNA is isolated, run out on native polyacrylamide gels, and probed by Northern blotting, two species of RNA can be detected, consistent in size with a monomer and a dimer. The dimer can be denatured by heat to become monomeric. The EM studies show that there appears to be a particularly stable contact point between the RNA molecules, close to the 5’ end of the genome. Under less stringent conditions (Mangel et al., 1974) multiple points of contact have been noted, suggesting that the single linkage observed under more restrictive conditions might be a primary association or dimerization initiation site (DIS). Recent work from Höglund et al. (1997), looking at the dimer linkage of HIV-1 showed, interestingly, that there did not appear to be any free 5’ ends, indicating that the two RNAs were interacting in addition at site(s) upstream of the recognized DLS. A combination of length determination and computer modelling suggested that more than one region might be involved in dimer linkage in this virus.

The structure/nature of the linkage

Studies with in vitro transcribed RNA showed that transcripts generated from the 5’ end of a variety of retroviral genomes exhibited electrophoretic profiles consistent with their being dimeric (Darlix et al., 1992). This observation provided a system within which the sequences and structures involved in the linkage of the two RNA molecules could be investigated. Since those initial studies, biochemical analysis, mutagenesis and physical probing of the RNAs in vitro have helped to elucidate critical sequences and predict secondary structures. Recently, the tertiary structure of at least the initial dimerization contact in HIV-1 was elucidated by NMR (Mujeeb et al., 1998) (see Fig. 3).

Early theories as to the nature of the interaction ignored the EM evidence that the molecules appeared to be aligned in a parallel fashion (Coffin, 1982). Subsequently purine tracts involved in non-canonical bonds, and apparently essential for dimer formation in vitro, were identified in the 5’ leader sequence of HIV-1 (Marquet et al., 1994). However, when identical tracts were deleted from the 5’ leader sequence of HIV-2, dimeric transcripts were still obtained (Berkhout et al., 1993). In a refinement of this work, two groups of investigators (Awang & Sen, 1993; Sundquist & Heaphy, 1993) separately suggested that since in vitro dimer formation in HIV-1 is cation dependent, this might implicate groups of guanine tetrams coordinated by cations linking the RNA molecules together, analogous to the structure of telomeres. In a number of different studies, however, it has been shown that dimer formation occurs in the absence of sequences containing these guanine motifs (Marquet et al., 1994; Muriaux et al., 1995).
This does not completely rule out a role for guanine tetrads since they might still be involved in maturation or stabilization of a dimer, but indicates that they are not pivotal in actual dimer formation.

Watson–Crick interactions are now the focus of attention. In the past 2–3 years, most work on the sequences involved in retroviral RNA dimer linkage has centred around palindromic regions found in the 5′ leader sequences of MoMLV, avian leukemia virus (ALV) and certain lentiviruses. The leader region, including the encapsidation signal of HIV-1, is predicted to consist of a series of stem–loop structures (Baudin et al., 1993; Berkhout, 1996; Clever et al., 1996; Harrison & Lever, 1992) (Fig. 1). One of these, containing a terminal palindrome (designated SL1 by some workers), was identified as being essential for dimer formation in vitro. This, and other palindromic sequences, are sometimes termed dimer initiation sites (DIS), indicating their importance in the very first stage of the interaction between the two RNA strands. As indicated below, however, the DIS structures so far postulated cannot alone be responsible for the stable intermolecular linkage seen in vitro, and the heat lability of RNAs containing only the DLS differ considerably from more extended structures containing the additional DLS elements. Those investigating the role of these sequences in vitro have proposed various theories for the nucleotide interaction. Watson–Crick bonds could form between the palindromic sequences, and the stem–loops subsequently interact in a so-called kissing-hairpin manner (Fig. 2a). In these in vitro experiments, the structure depends not only on complementarity in bases between the strands, but it is also suggested that flanking purine residues are important (Paillart et al., 1997) (Fig. 2b). This model has been supported by chemical probing of the RNA secondary structure and computer predictions of putative tertiary structure (Paillart et al., 1997; Skripkin et al., 1994) (Fig. 2b). Comparable RNA–RNA interactions based on the interaction between auto-complementary single-stranded loops have been described elsewhere for the regulation of plasmid regulation (Persson et al., 1990). Autocomplementarity between palindromes is insufficient for dimerization as some constructs containing inverted palindromes barely dimerize at all. However, since similar palindromic sequences occur in the leader regions of other lentiviruses and in ALV and MoMLV (Girard et al., 1995; Khan & Giedroc, 1992), this mechanism was postulated as being common to all retroviruses. However, the stability of in vitro produced dimeric RNA to heat and denaturing agents strongly argues against Watson–Crick pairing as being the only interaction involved in the linkage (if the in vitro system is representative of the in vivo linkage). Removal of no single element or motif consistently abolishes dimer formation. The recent tertiary structure already alluded to supports the model of two interacting stem–loop structures (Mujeeb et al., 1998). The loop structures proposed by this study, however, are atypical and are dissimilar to those seen in kissing-hairpin models. The loops are distorted by interstrand stacking of the adenosines, and the authors of this paper suggest that this might contribute to the instability of the structure, perhaps bringing about the melting of the initial contact and leading to the formation of a more stable structure. The actual NMR was obtained by proton 2-D NMR, using D₂O. It remains to be seen whether other studies back up these findings, and whether or not a structure can be solved for the more stable dimer complex.

In vivo, viral and cellular proteins may have a role to play in the dimerization process. The nucleocapsid protein (NC) has

---

**Fig. 1.** Putative secondary structure of the DIS and packaging signal (Ψ) in HIV-1 (Subject to modification after Harrison et al., 1998.).
potent annealing activity, as well as the ability to destabilize nucleic acids (Khan & Giedroc, 1992). NC protein was shown to strongly activate dimerization in vitro, presumably because of these two properties (Feng et al., 1996). It has been suggested that NC interacts with the kissing-hairpin itself (Muriaux et al., 1996).

Other retroviral DLS which have been mapped include those of bovine leukaemia virus (BLV) and human T-cell lymphotropic virus type I (HTLV-I). Although these two viruses are related, the evidence to date does not point to a similar mechanism for dimer linkage (Greatorex et al., 1996; Katoh et al., 1991, 1993). In HTLV-I, sequences required for the dimerization of in vitro transcribed RNAs have been mapped to a region downstream of the splice donor (SD) and upstream of the primer binding site (PBS). Potential secondary structures identified in this region of the HTLV-1 genome include palindromic sequences, and a conserved stem–loop. However, disruptive mutagenesis of these putative structures reduced but did not abolish dimer formation, suggesting that they were not essential (Greatorex et al., 1996). In contrast to HTLV-I,
BLV, like RSV, appears to have a discontinuous DLS (Katoh et al., 1991, 1993; Kurg et al., 1995) although work is still required to pinpoint the site of dimer association.

In RSV, the dimer linkage has been visualized by EM (Murti et al., 1981). In vitro mapping studies indicate that it is located between nucleotides 544 and 564 from the 5' end of the genome (Bieth et al., 1990). There appear to be discontinuous sequences within positions 208–270 and 400–460 which also promote dimer formation.

To date, there is no known single type of nucleotide interaction which can explain all the findings in vivo and in vitro concerning the nature of the dimer linkage interaction. The frequent occurrence of palindromic sequences in regions shown to be associated with dimer linkage strongly supports their involvement, but it must be remembered that there is also much experimental evidence (described above) which indicates that Watson–Crick bonding alone cannot be the sole linkage.

### The role of the dimer linkage in the virus life-cycle

In all retroviruses, dimer linkage and the diploid genome provide the opportunity for switching from a damaged to a physically linked intact template, and this is probably advantageous. In highly variable viruses like HIV where variation is linked to ‘fitness’, the ability to recombine is also an asset. In MoMLV the DLS overlap with sequences encoding a hyper-variable region, and it has been proposed (Mikkelsen et al., 1996) that the DLS might, in this instance, be involved in recombination, although this is difficult to reconcile with the highly conserved genome of MoMLV.

Recent reports have further investigated whether or not the kissing-hairpin sequences have a role in virion dimer formation, and also if they have a biological significance (see below). Summarized in Table 1 are some of the data from the six studies which have attempted to examine the role of the dimer linkage in the HIV-1 life-cycle. The conclusions vary, possibly reflecting differences in assay methods. Two of the five studies quantifying relative amounts of dimer demonstrated that this was decreased in viruses mutated in the kissing-hairpin domain (Haddrick et al., 1996; Laughrea et al., 1997). However, this decrease was nowhere near as profound as that seen in vitro, and three of the five studies showed no difference at all (Berkhout & van Wamel, 1996; Clever & Parslow, 1997; Sakuragi & Panganiban, 1997). In vitro, the melting temperatures of the transcripts in which the kissing-hairpin sequences have been deleted differ from that of undeleted transcripts (Laughrea & Jette, 1994; Muriaux et al., 1995). There is conflicting evidence as to whether or not this is the case in vivo, however, some authors seeing no difference between wild-type and mutated viruses (Berkhout, 1996; Clever & Parslow,

### Table 1. Summary of data from studies investigating the role of the kissing-hairpin sequences in some aspects of the HIV-1 life-cycle

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laughrea et al. (1997)</td>
<td>Stem–loop mutations Deletion of stem and palindrome</td>
<td>Saw difference in amount of dimer Mutants packaged to a lesser extent than wild-type Transcription efficiency reduced Splicing reduced Infectivity reduced ~ 99%</td>
</tr>
<tr>
<td>Paillart et al. (1996)</td>
<td>Deletion of KLD* ± bulge Mutations in palindrome</td>
<td>2–5 × decrease in packaging 10–1000 × decrease in infectivity</td>
</tr>
<tr>
<td>Berkhout &amp; van Wamel (1996)</td>
<td>Disruption of palindrome Insertion of larger palindrome</td>
<td>Dimers had similar thermal stabilities Mutants showed 2 × decrease in packaging Mutants showed 10 × decrease in infectivity</td>
</tr>
<tr>
<td>Clever &amp; Parslow (1997)</td>
<td>Mutations in stem and loop</td>
<td>Dimers had similar thermal stabilities Mutants showed slight packaging defects Mutants package more spliced RNA Mutants show infectivity defect</td>
</tr>
<tr>
<td>Haddrick et al. (1996)</td>
<td>Mutation in palindrome</td>
<td>Reduction in the amount of dimeric RNA Delayed replication</td>
</tr>
<tr>
<td>Sakuragi &amp; Panganiban (1997)</td>
<td>Stem–loop mutations KLD mutations</td>
<td>Dimers had similar thermal stabilities</td>
</tr>
</tbody>
</table>

* KLD, kissing loop domain.
1997), others seeing less dimeric RNA (Haddrick et al., 1996; Laughrea et al., 1997) (see Table 1). Likewise, replication was also impaired when the viral genome was mutated in the kissing-loop domain. However, if assays are continued for a period of time longer than that reported in these studies (<10 days), it is apparent that replication of the mutants is only delayed and can reach the same levels as wild-type (Harrison et al., 1998).

As other authors have observed, the proximity of the DLS to the packaging signal in HIV-1 makes it tempting to speculate that dimerization and encapsidation are linked; it is thus difficult to separate the effect of a mutation on one function from its effect on the other. Similarly, independently of this, dimerization may affect other processes linked to infectivity. The authors of the various studies in Table 1 used different assays to assess viral infectivity, but 4/6 studies showed a marked decrease in infectivity (ranging from 10–100%) (Berkhout & van Wamel, 1996; Clever & Parslow, 1997; Laughrea et al., 1997; Paillart et al., 1996). Pertinent observations on the link between encapsidation and dimerization have been made in other viruses: RSV is one of the few retroviruses in which virion RNA has been directly analysed by electrophoresis on native polyacrylamide gels (Lear et al., 1995), and this study appeared to show that initially monomers are packaged, maturing over time to dimers. Thus, in this virus, encapsidation does not depend on dimerization or a DLS.

Fu & Rein (1993) showed in MoMuLV that, in newly released virions, the RNA was dimeric but dissociated to monomers at a lower temperature than in mature virions. Using protease negative and NC mutants, they suggested that maturation of dimeric RNA requires the cleavage of the Gag precursor or the presence of an intact cysteine array in the released NC protein.

However, both RSV and MoMuLV are simple retroviruses and it is hazardous to extrapolate findings in either of these to complex retroviruses such as HIV where the encapsidation process may be quite distinct.

The way forward

The dimer linkage is a consistent feature of retroviruses as evidenced by biochemical analysis, electron microscopy, and examination of virion RNA. Its conservation, proximity and potential role in RNA encapsidation also single it out as one of a number of possible targets for gene and antiviral therapy.

Much effort has gone into in vitro studies teasing out the possible nature of the actual interaction but a complete understanding of the DLS awaits solution of all the tertiary and quaternary structures. Until that time, molecular modelling by mutagenesis provides only a tantalizing but incomplete glimpse with which to speculate about the structure. This information hopefully will supplement biological studies of virus replication and infectivity and between them further our understanding of the role of RNA dimer linkage in the virus life-cycle.

We thank Ray Hicks for his help with the figures in this Review, and Teresa Barnes for secretarial assistance. We would also like to thank Paul Digard for critical reading of the manuscript. We are grateful to Anwar Muyeeb and Tris Parslow for Fig. 3.

References


