

Diverse interactions of retroviral Gag proteins with RNAs

Alan Rein¹, Siddhartha A.K. Datta¹, Christopher P. Jones² and Karin Musier-Forsyth²

¹ HIV Drug Resistance Program, National Cancer Institute-Frederick, Frederick, MD 21702, USA

² Departments of Chemistry and Biochemistry, Center for Retroviral Research, and Center for RNA Biology, Ohio State University, Columbus, OH 43210, USA

Retrovirus particles are constructed from a single virus-encoded protein, termed Gag. Given that assembly is an essential step in the viral replication cycle, it is a potential target for antiviral therapy. However, such an approach has not yet been exploited because of the lack of fundamental knowledge concerning the structures and interactions responsible for assembly. Assembling an infectious particle entails a remarkably diverse array of interactions, both specific and nonspecific, between Gag proteins and RNAs. These interactions are essential for the construction of the particle, for packaging of the viral RNA into the particle, and for placement of the primer for viral DNA synthesis. Recent results have provided some new insights into each of these interactions. In the case of HIV-1 Gag, it is clear that more than one domain of the protein contributes to Gag–RNA interaction.

Gag proteins: the building blocks of retroviruses

Retroviruses are divided into six genera (alpha-, beta-, gamma-, delta-, epsilon-, and lenti-retroviruses) plus one outlying subfamily, the Spumaviridae, which will not be discussed here. Although there is almost no sequence conservation between genera, homologous proteins generally show striking structural resemblance. Likewise, the replication cycles of members of the different genera are similar in overall outline, but differ in many important details [1].

The Gag protein is the fundamental building block of retrovirus particles, and expression of this single virus-encoded protein is sufficient for efficient assembly of virus-like particles (VLPs) in mammalian cells (Box 1, Figure 1). During the course of virus assembly, the multidomain Gag proteins participate in interactions with lipids in the plasma membrane, with RNAs, and with other Gag molecules. This review focuses on the multiple ways in which Gag proteins interact with RNA, with an emphasis on recent findings. These interactions involve the formation of the particle itself, the selection of the correct RNA species for inclusion in the particle, and the refolding of the packaged RNAs. The interactions are largely, but (importantly) not entirely, mediated by the nucleocapsid (NC) domain of Gag.

A nonspecific interaction with RNAs drives virus particle assembly

As noted above, expression of Gag in the cell leads to particle assembly. By contrast, purified, recombinant

Gag proteins are soluble in aqueous media, exhibiting little or no tendency for the protein–protein interactions needed for virus particle formation. However, addition of nearly any single-stranded nucleic acid triggers the *in vitro* assembly of the Gag proteins into VLPs [2]. This basic observation suggests the possibility that nonspecific binding to RNA is essential for particle assembly *in vivo*.

In vivo, the dependence of particle assembly upon nucleic acids (NAs) appears to vary markedly between different retroviruses. Thus, deletion of the NC domain completely abolishes assembly in murine leukemia virus (MLV) [3], but has a much more modest effect in HIV-1 [4]. Particles formed *in vivo* from an HIV-1 Gag protein lacking its NC domain are heterogeneous, but similar in appearance to authentic particles. They contain no viral RNA and little, if any, cellular RNA. Similarly, exposure of detergent-stripped, immature virus particles to RNase disrupts the particles in the case of MLV [5], but not HIV-1 [6]. These observations are consistent with the idea that MLV particles, but not HIV-1 particles, are constructed by means of Gag–RNA as well as Gag–Gag interactions.

Despite these differences, particle assembly, at least *in vitro*, appears to depend on Gag–NA interactions. How might we explain this dependence? It is significant that NAs as short as 20–40 nucleotides can support assembly *in vitro* [7–9]. Although the ‘site sizes’ of Gags on NAs are not precisely known (see below), these NAs are so short that they can only bind a very small number of Gag molecules. Similarly, mRNAs far shorter than viral RNA (vRNA) are packaged *in vivo* by MLV or HIV-1 Gag if vRNA is not present [10]. Thus, the NA cannot be a ‘string’ upon which all the Gag molecules are aligned; rather, it appears that when several Gag molecules are bound to NA, they are primed to interact with other NA-bound Gag molecules and form a virus particle (Figure 1).

A key observation is that HIV-1 Gag molecules assemble efficiently and correctly *in vivo* if their NC domain is replaced by a leucine-zipper domain [11]. The particles formed by these chimeric proteins are nearly indistinguishable in appearance from normal immature particles, but contain very little cellular or viral RNA [12,13]. This finding suggests that oligomerization via the zipper domain has supplanted oligomerization on RNA in facilitating assembly.

The principal site of interaction with NAs is, in general, the NC domain of Gag. However, detailed structural analysis of immature particles [14–16] and genetic studies with

Corresponding author: Rein, A. (reina@mail.nih.gov)

Box 1. Retrovirus particles are made of Gag proteins

The Gag protein is the fundamental building block of the retrovirus particle (Figure 1). Gag proteins are multidomain proteins, always containing (from the N to the C terminus) an MA domain, a CA domain, and an NC domain [1]. One primary function of the MA domain is to target Gag to the plasma membrane, where assembly occurs (in most retroviruses). The CA domain plays a major role in the protein–protein interactions between Gag molecules in the virus particle, and the NC domain is crucial in the interactions of Gag proteins with NAs. The initial product of assembly is the ‘immature particle’, a roughly spherical, pleomorphic structure with a diameter of approximately 120 nm. It is

composed of several thousand Gag molecules circumscribed by a lipid bilayer. Smaller numbers of other virus-encoded proteins and a few per cent by mass of RNA are also present in the particle, but expression of Gag alone is sufficient for assembly of immature particles in mammalian cells. The Gag molecules are extended rods, arranged as radii of a sphere, with their N termini in contact with the lipid bilayer and their C termini in the interior of the particle. After the virus is released from the cell, Gag is cleaved by the virus-encoded protease, releasing free MA, CA, and NC proteins. These maturation-cleavage events are absolutely required for infectivity of the particle.

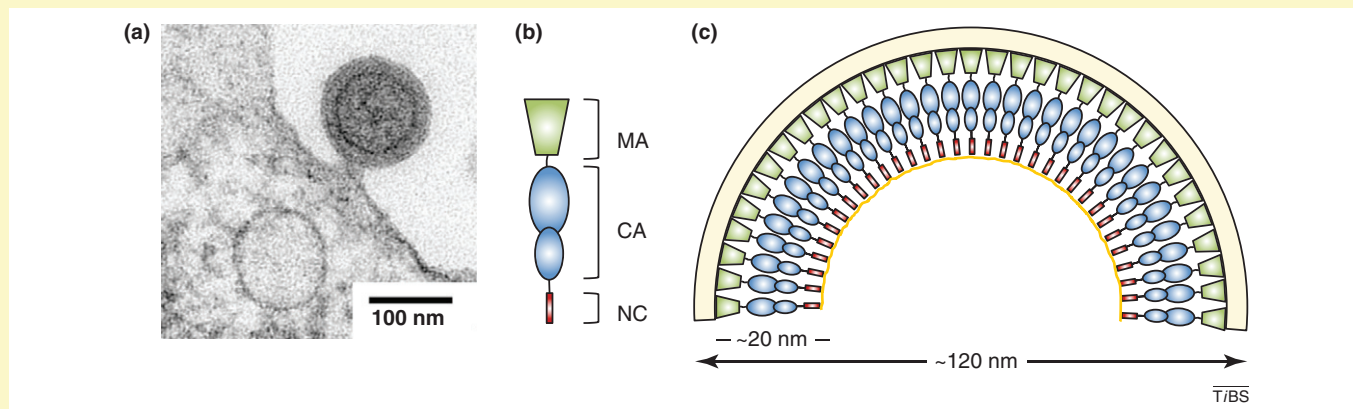


Figure 1. (a) A transmission electron micrograph of a section of a human cell expressing HIV-1 Gag, showing a single immature virus particle being released from the cell. (b) Gag is a multidomain protein always containing, from N to C terminus, the MA domain (green), the CA domain (blue), and the NC domain (red). Depending on the genus of retrovirus, one or a few additional domains or spacers may also be present. (c) The radial arrangement of Gag molecules that are in contact with membrane on the outside and RNA (yellow line) on the inside of the immature particle. There are approximately seven nucleotides per Gag molecule in the particle.

chimeric Gag proteins [17] suggests that the interfaces at which Gag proteins contact each other in the particle are largely or entirely in the capsid (CA) domain (and, in HIV-1, perhaps in SP1, a short ‘spacer’ between the CA and NC domains). These findings suggest that binding of NA to the NC domain somehow alters the conformation of the CA domain, thereby activating or exposing new interfaces for protein–protein interaction.

A recent study of a short peptide spanning the CA–SP1 junction produced data entirely consistent with this hy-

pothesis [18]. The circular dichroism spectrum of this peptide indicates that it is a coil at low concentration in an aqueous medium, but adopts an α -helical conformation when dissolved at a high concentration (approximately 5 mM). It seems likely that a similar transition occurs when Gag oligomerizes on RNA, and perhaps this change in SP1 is ‘propagated’ into the CA domain, priming it to participate in virus particle assembly.

In most retroviruses, the NC domain of Gag has the highest affinity for RNA. However, the matrix (MA) domain, at the N terminus of HIV-1 Gag, is also positively charged, and can interact with RNAs [19–21]. Interestingly, although the flexibility of the interdomain linkers has precluded detailed structural analysis of HIV-1 Gag, biochemical experiments and hydrodynamic and neutron-scattering measurements all indicate that it tends to adopt ‘folded-over’ conformations in free solution, in which its two ends are relatively close together in 3D space [21–23]. This shape is incompatible with formation of authentic immature particles, in which Gag is a rod approximately 20 nm long and approximately 2–3 nm wide.

The ‘folded’ conformation of free HIV-1 Gag protein has somewhat unexpected consequences for virus assembly, both *in vitro* and *in vivo*. When recombinant Gag is mixed with RNA, the VLPs that it forms are far smaller than authentic virus particles, with diameters of only 25–30 nm rather than approximately 120 nm [7]. These VLPs cannot be constructed from the rod-shaped Gag proteins present in authentic VLPs, given that the latter are 20–25 nm long (Box 1). Rather, the ‘shell’ of Gag visible in electron micrographs of the small VLPs appears to be only approximately 8 nm thick; presumably, these shells are composed of bent

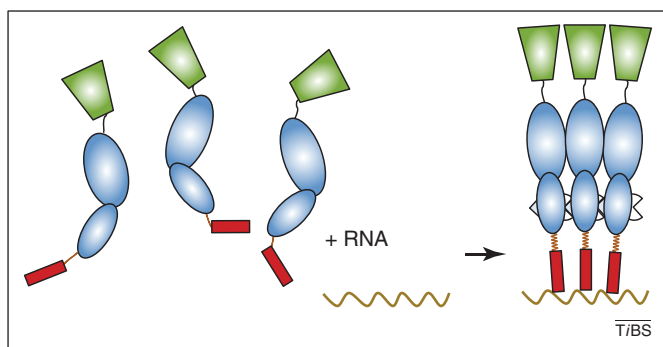


Figure 1. Proposed mechanism of dependence of virus particle assembly on nucleic acid.

Purified Gag proteins are soluble, but assemble into VLPs when NA is added. The fact that short oligonucleotides (approximately 30 bases) can support assembly suggests that oligomerization renders the Gag assembly-competent. The available data suggest that when several Gag molecules are in close proximity, the very high local Gag concentration induces a conformational change, leading to the exposure of new interfaces (shown here in the C-terminal portion of the CA domain in blue) for protein–protein interactions, and in turn to virus-particle assembly. In HIV-1, a high concentration of Gag appears to induce a helical conformation in SP1, a short ‘spacer’ between NC (red) and CA [18]. This model was originally suggested by Ma and Vogt [8]. MA domain is shown in green.

Gag molecules, in which both ends are bound to RNA. Further investigation has shown that VLPs of the correct size are assembled *in vitro* if inositol pentakisphosphate (a heavily phosphorylated compound resembling the head groups of membrane-localized phospholipids), is added to the HIV-1 Gag protein in addition to RNA [6].

Additional observations have somewhat clarified this phenomenon. Investigation of the conformation of HIV-1 Gag (lacking the p6 domain) [24] showed that it maintains its compact state when bound to a biomimetic membrane, evidently with both the N-terminal MA domain and the C-terminal NC domain bound to negatively charged lipid headgroups. Similarly, both ends can bind to a NA molecule, as in the small VLPs. Although the domains at both ends can bind either type of polyanion, the MA domain has a strong preference for lipids, and the NC domain has an especially high affinity for NAs. Thus, when lipid headgroups and NAs are both present, the ends of Gag separate, enabling the protein to adopt the rod-shaped conformation present in authentic VLPs (Figure 2). *In vivo*, the MA and NC domains of HIV-1 Gag are apparently both bound to RNA in the cytoplasm until the protein reaches the plasma membrane [20]. This membrane contains phosphatidyl inositol (4,5) bisphosphate, to which the MA domain binds with high specificity [19,25]. Perhaps the fact that Gag requires interaction with both membranes and NAs before it undergoes a conformational change explains why HIV-1 assembly occurs only upon targeting to the plasma membrane. Recent observations [26] on the ability of HIV-1 Gag to anneal tRNA to the vRNA are fully consistent with this hypothesis, as discussed below.

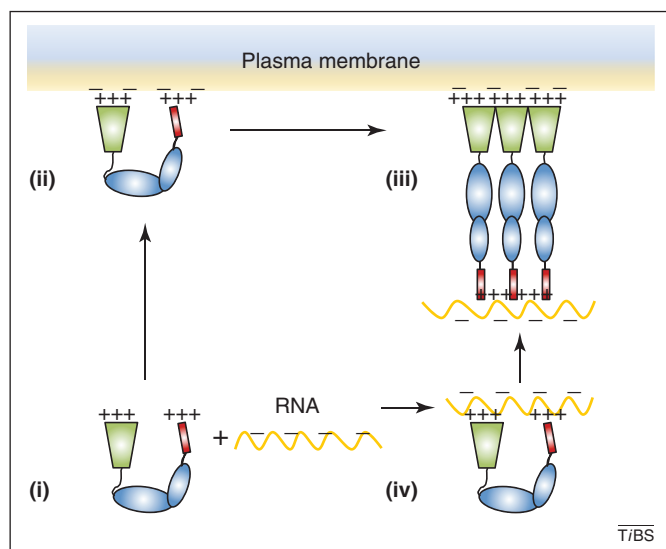


Figure 2. HIV-1 Gag undergoes major conformational changes during virus particle assembly.

(i) The HIV-1 Gag protein assumes compact conformations in solution, with both ends near each other in 3D space [22]. Both the N-terminal MA domain (green) and the C-terminal NC domain (red) are positively charged, but the MA domain (blue) has a much higher affinity for negatively charged phospholipids than for RNA, whereas the NC domain exhibits the converse pattern. Nonspecific electrostatic interactions maintain the compact conformations of the protein when it is bound (ii) to membrane with negatively charged phospholipids, or (iv) to RNA. (iii) However, when both types of ligands are present, the protein extends to the rod-shaped conformation present in authentic virus particles (see Box 1) [24].

A specific interaction leads to the selective packaging of viral RNA

Retroviral genomic RNAs (vRNAs) are also the mRNAs for the synthesis of Gag and Gag–Pol viral proteins. Like other mRNAs, they have a 5' cap and a 3' poly(A) tail. However, unlike cellular mRNAs, they are selectively packaged into assembling virus particles. What explains this selective packaging? One proposal is the proximity of translated vRNA to its gene product, that is, Gag [27]. In fact, although there is a modest preference for packaging of the Gag mRNA in HIV-1, or '*cis*-packaging' [28], the selection of vRNA occurs principally because this RNA contains a *cis*-acting 'packaging signal', generically called ' ψ '. ψ can be defined experimentally by deletion analysis, or more rigorously, by placing a stretch of viral sequence into a nonviral RNA and determining whether the resulting chimeric RNA is selectively packaged [29–31]. The ψ signals identified to date are at least 100 nucleotides long, and often significantly longer. The ψ signal in HIV-1 RNA is very poorly defined [32]; however, many studies have attributed roles in packaging to several stem-loops in the 5' untranslated region of the vRNA [33]. Interestingly, when wild-type MLV or HIV-1 Gag is expressed in a cell lacking any ψ -containing RNA, it still assembles efficiently into normal-looking particles and packages cellular mRNAs, essentially nonselectively [10]. Thus, vRNA is in competition with a large excess of cellular mRNAs for inclusion into assembling virions, and ψ confers a strong advantage to vRNA in this competition.

Encapsidated vRNAs are always dimeric, consisting of two vRNA molecules of the same (sense) polarity, held together by intermolecular base pairing near the 5' ends of the RNAs. When the virus infects a new host cell, the vRNAs are copied into a single molecule of double-stranded DNA by reverse transcriptase (RT). RT can switch between the two monomers during DNA synthesis, and this presumably provides 'insurance' against breakage of one of the monomers. Switching is also the source of genetic recombination, an important source of genetic variation in retroviruses [34].

A number of observations suggest that there is an intimate relationship between dimerization and packaging. First, ψ is generally located near the 5' end of the vRNA, and this is also the locus of the tightest connection between the monomers [1]. Moreover, in situations in which vRNA packaging is reduced, the encapsidated vRNA is still dimeric, suggesting that vRNA molecules enter virions in pairs, rather than independently. Packaging of heterodimers between distinct HIV-1 vRNAs is a function of their sequence at the 'dimer initiation site' near the 5' end of the RNAs [34]; a vRNA can be packaged as a monomer if it contains two tandem ψ regions [35]; and nonviral mRNAs containing ψ are packaged as dimers [31]. Finally, mutations that interfere with dimerization of MLV RNAs also prevent selective packaging [36]. These results seem to indicate that ψ is not simply a specific nucleotide sequence, but that it also requires a dimeric structure [33,37–39].

It has generally been assumed that ψ is a high-affinity binding site for Gag, so that binding of Gag to and encapsidation of a ψ -containing RNA is energetically more favorable

than binding to and encapsidating an RNA lacking ψ . In other words, according to this hypothesis, ψ gives its RNA a thermodynamic advantage in its competition with cellular RNAs for encapsidation. However, it is also conceivable that ψ confers an advantage to RNAs by directing them to a site in the cell where they will encounter Gag. In fact, in many RNA viruses, virus replication occurs in specialized compartments derived from cytoplasmic membranes [40]. Such physical segregation of the viral genome would certainly explain its selective packaging. It is possible that such a mechanism is at work in the alpharetrovirus Rous sarcoma virus (RSV), whose Gag protein is trafficked through the cell nucleus, and might acquire vRNA in the nucleus [41,42]. It has also been suggested that MLV ψ directs the transport of vRNA to the site of assembly [43,44].

The structural details of the molecular interactions responsible for selective vRNA packaging have been extremely difficult to analyze. As noted above, there are no molecular structures for intact Gag proteins, and RNAs as large as 100 nucleotides pose enormous challenges for analysis by nuclear magnetic resonance (NMR) or crystallography. The problem has been simplified by focusing solely on free NC protein, in essence assuming that this is the only region of Gag that is important in vRNA packaging, and that the conformation of NC is the same as that of the NC domain of Gag. NC proteins are generally very small (55 residues in HIV-1) and relatively basic, and contain one or two zinc fingers with an absolutely conserved spacing of zinc-coordinating residues, i.e., C-X₂-C-X₄-H-X₄-C. Mutational studies show that these zinc fingers are of great importance in selective packaging, as replacement of a single zinc-coordinating residue with a non-zinc-coordinating residue can lead to assembly of virus particles lacking vRNA [45–47].

The atomic-resolution structures of two complexes between HIV-1 NC and either of a pair of RNA stem-loops (believed to be elements of ψ) have been determined by NMR [48,49]. It has been found that NC adopts different structures in the two complexes: thus, it appears to be a very flexible molecule, except for the zinc fingers.

Some insight into one mechanism of specific vRNA packaging comes from studies on RNA structures in MLV ψ . Several years ago, D'Souza and Summers, using NMR to study the ψ region of MLV vRNA, observed that when MLV RNA dimerizes, there is a shift in register in some of the stem-loops; in turn, this means that there are paired bases in the monomeric vRNA that become unpaired upon dimerization [50]. These authors also noted that MLV NC protein binds to these bases with high affinity, and made the important suggestion that this 'riboswitch' could explain the selective packaging of dimeric vRNA [33,50].

The secondary structure of nucleotides 205–374 of MLV vRNA, a crucial part of ψ , has also been analyzed in detail, in both monomers and dimers (Figure 3), using a chemical structure-probing technique termed 'selective 2'-hydroxyl acylation analyzed by primer extension' (SHAPE) [51,52]. The intermolecular base pairs in the dimer include both extended palindromic stretches (PAL1 and PAL2) and loop-loop 'kissing' interactions between the CG sequences in GACG loops (SL1 and SL2). The RNA contains two copies of the sequence UCG-UPu-UCUG (Figure 3, blue

boxes), and each of the four UCG motifs is partially or fully base-paired in the monomer, but unpaired in the dimer. Each copy of this 10-base sequence is flanked on both sides by base-paired stretches in the dimer. Recent studies indicate that these motifs are binding sites for recombinant MLV Gag and NC protein *in vitro*, and that they are occupied by NC protein within mature virus particles [52]. Moreover, a filter-binding assay showed that the apparent affinity of Gag protein for dimeric RNA is reduced by approximately sixfold if the four UCG motifs are changed to UCUA. Finally, introducing these four mutations into vRNA effectively disrupted ψ , preventing selective packaging of the vRNA [52]. Taken together, these results show that the UCG motifs, and their surrounding context, are crucial for specific binding by Gag and for selective encapsidation. This was the first report in which a packaging signal was destroyed by such a small number of single-base changes. These results make it clear, at least for gammaretroviruses such as MLV, that ψ does indeed contain a specific, high-affinity binding site for the Gag protein, and that the binding properties of Gag are similar to those of NC. The importance of the exposed UCG sequence also explains the selective packaging of dimeric, rather than monomeric, vRNA.

As noted above, retroviral genera are often remarkably different from each other in the details of their biology and replication. For example, the unmasking of the UCG motif upon dimerization was the clue that suggested its special significance in MLV ψ ; however, no differences in secondary structure between monomeric and dimeric HIV-1 RNA have yet been detected [53]. Another striking exception from the general rule is that although the NC domain of Gag is positively charged and is essential for NA interactions in most retroviruses, it is neutral in the deltaretroviruses; the MA domain seems to perform the principal NA-binding functions in these viruses [54].

Gag proteins refold RNAs during virus assembly

In addition to the nonspecific Gag–RNA interaction required for particle assembly and the specific Gag–RNA interaction required for selecting vRNA, there are nonspecific interactions in which Gag remodels the RNA, by virtue of its NA chaperone activity [55]. In general, an NA chaperone binds NAs transiently and, with no need for ATP, catalyzes their refolding into thermodynamically favored states [56–58]. The NA chaperone activity of free HIV-1 NC protein has been analyzed in considerable detail [59,60], and recent studies have begun to dissect the differences between the chaperone activities of free NC and Gag [26,55,61].

Two RNA annealing events accompany retrovirus assembly: vRNA dimerization and tRNA annealing. All retroviruses use host cell tRNAs as primers for reverse transcription. To act as a primer, the tRNA must be partially unwound and annealed to a complementary sequence, the primer-binding site (PBS), on the viral genome (Figure 4). In the case of HIV-1, the 18 nucleotides at the 3' end of human tRNA^{Lys3} are perfectly complementary to the PBS, and both NC [59,60] and Gag [55] can facilitate the annealing of this tRNA to vRNA *in vitro*. However, free NC protein does not catalyze these essential reactions *in vivo*:

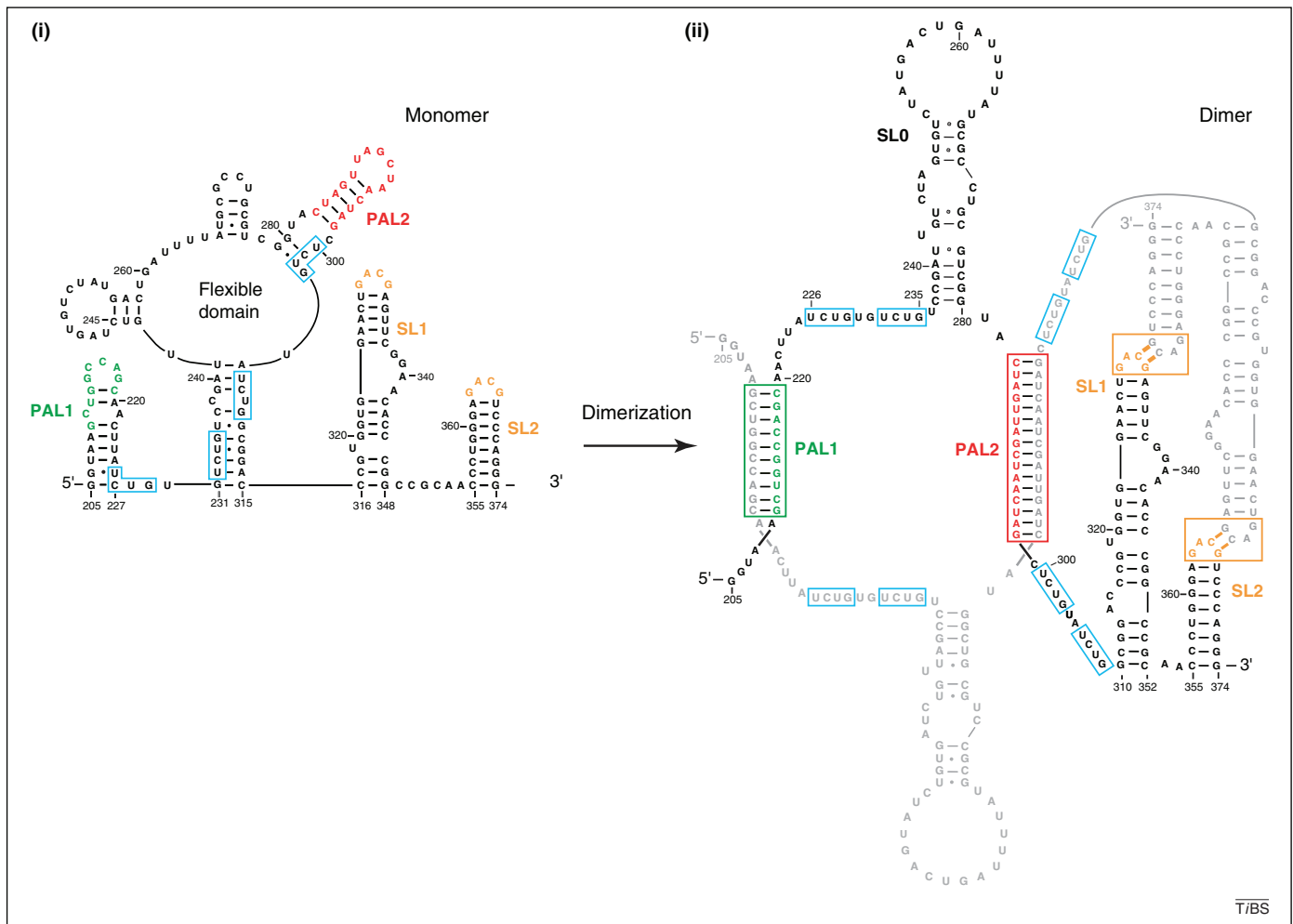


Figure 3. The MLV packaging signal contains a riboswitch.

All retroviral vRNAs are packaged into virus particles in dimeric form. (ii) In MLV, the contacts between the two monomers in the dimer include the extended base-paired stretches PAL1 and PAL2 (green and red boxes, respectively), and 'kissing-loop' interactions between SL1 in one monomer and SL2 in the other (orange boxes). Each monomer contains two UCUg-UPu-UCUG motifs (blue boxes). Dimerization involves a 'switch', in that the UCUg-UPu-UCUG motifs are partially or fully base-paired in (i) the monomer, but unpaired in the dimer. These motifs are high-affinity binding sites for MLV Gag protein and are bound by NC protein in mature MLV particles; replacement of the four G residues with A residues prevents selective packaging of vRNA [50–52].

both tRNA primer annealing and genome dimerization occur in protease-deficient virions [62–67]. Further, sequence changes within the NC domain of Gag greatly reduce tRNA annealing and the fraction of vRNA extracted from virions in dimeric form [65,68,69]. These observations suggest that Gag, via its NC domain, catalyzes the two annealing steps *in vivo*. Other cofactors, such as RNA helicase A, might also contribute to tRNA annealing by acting upon vRNA [70].

The molecular mechanisms underlying the NA chaperone activity of HIV-1 NC protein have been investigated in great detail, using a number of experimental systems and techniques. The activity stems from at least three properties of the protein: (i) its polycationic character, which, through electrostatic interactions, helps bring NA strands together; (ii) its ability to weakly destabilize base pairs within NA and reduce the cooperativity of strand separation; and (iii) its very rapid kinetics of binding to and dissociation from NA molecules [59,60]. NC can readily adapt to bind to different RNA structures [48,49] by virtue of its intrinsically disordered regions, a feature it shares with many NA chaperones [71].

Recombinant HIV-1 Gag can catalyze both vRNA dimerization and tRNA annealing *in vitro* [26,72,73]. Less is known about the chaperone activities of other retroviral Gag proteins, although their NC counterparts show a wide range of activities [74–77]. For example, the rates at which the NC proteins of HIV-1 and RSV catalyze annealing of complementary NAs are roughly equivalent, whereas the annealing rate in the presence of MLV NC protein is lower, and that of HTLV-1 (a deltaretrovirus) is greatly reduced [77]. Whether a similar relationship exists for the corresponding Gag proteins remains to be determined.

In MLV, the conversion of vRNA monomers to dimers involves extensive remodeling (Figure 3). Interestingly, dimerization also unmasks the UCUg-UPu-UCUG high-affinity binding sites for MLV Gag; thus, in addition to acting as a catalyst for dimerization, MLV Gag might shift the equilibrium in the cell in favor of dimer formation. As noted above, however, structural differences between monomer and dimer were not apparent from SHAPE studies of the HIV-1 genome [53]. It is known that the dimeric linkage in HIV-1 involves a kissing-loop structure with six base pairs (including at least four GC pairs), but

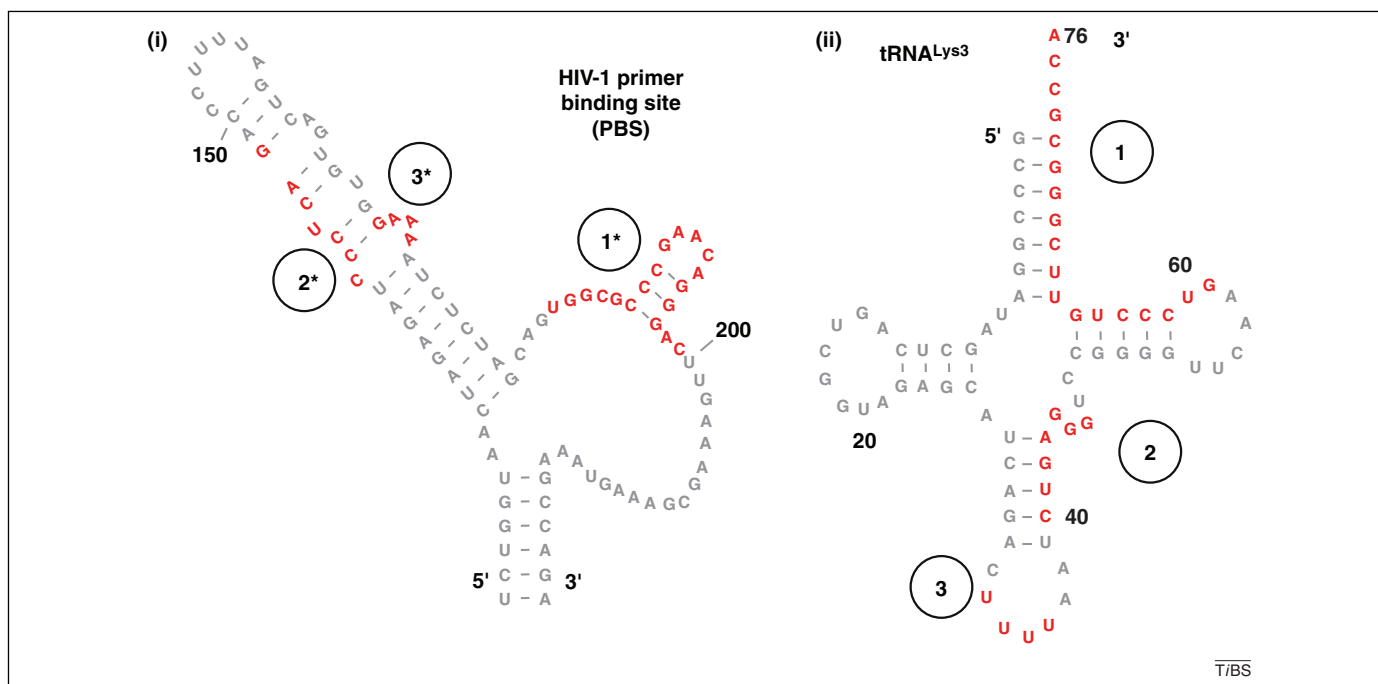


Figure 4. Substrates for Gag-facilitated RNA refolding events that occur during HIV-1 assembly.

HIV-1 Gag catalyzes annealing of (ii) the primer tRNA^{Lys3} to (i) complementary regions on the viral genome. Three regions of tRNA^{Lys3} (1, 2 and 3; red) interact with three complementary regions on the genome (1*, 2* and 3*; red). (1) The highly conserved, 18-nucleotide PBS (1*) is complementary to the 3–18 nucleotides of tRNA^{Lys3}. Reverse transcription initiates from the 3-hydroxyl group of A76. Regions 2* and 3* have been proposed to interact with the anticodon stem and variable arm (2) and the anticodon loop (3), respectively, but their roles in tRNA^{Lys3} priming are not well established. The 2/2* interaction is important for efficient initiation of cDNA synthesis [83], and a similar interaction occurs in RSV [84]. The 3/3* interaction between the U-rich tRNA^{Lys3} anticodon and the A-rich bulge upstream of the PBS has been proposed to help stabilize the tRNA^{Lys3} primer [85]. These interactions are consistent with SHAPE analysis of the genome [53].

other intermolecular contacts or intramolecular rearrangements in the dimers have not yet been identified [34].

The ability of HIV-1 Gag to facilitate the annealing *in vitro* of tRNA^{Lys3} to vRNA was recently investigated in greater detail. This activity depends on the NC domain, as Gag variants lacking NC failed to catalyze the annealing reaction [26]. Although Gag is a more efficient chaperone than NC on a molar basis, the rate of Gag-facilitated annealing of the tRNA primer was approximately an order of magnitude lower (at saturating protein concentrations) than NC-facilitated annealing [26]. This reduced rate was due to the presence of the MA domain, as constructs that harbored partial deletions of, or replacements of basic residues within, MA showed enhanced rates of annealing [26]. Moreover, Gag-facilitated tRNA^{Lys3} annealing was enhanced in the presence of inositol phosphate (IP) compounds, which resemble the head groups of membrane lipids with which MA interacts. [25,26]. As discussed above, HIV-1 MA can bind to NAs, but has a higher affinity for IPs [19,21,26], consistent with the general model for Gag–NA interactions (Figure 2). Although further studies are needed, modulation of the chaperone activity of Gag by binding of IP or phospholipids might regulate important events in the HIV lifecycle, such as tRNA primer annealing.

Although the NA chaperone activity of HIV-1 Gag and its cleavage product, NC, are comparable in most assays *in vitro*, the dimeric vRNA in immature virus particles is less compact and less thermostable than that in mature particles [64,65,67]. Similarly, the tRNA–vRNA complex is apparently different in immature and mature particles [62,78,79]. Thus, the annealing attributable to Gag in immature VLPs is somehow impaired relative to that

ascribed to NC in mature particles. It is not known whether these differences arise from intrinsic differences in the NA chaperone activities of the two proteins, or from differences in the structures of immature and mature particles. For example, it seems plausible that, because Gag molecules are immobilized in a lattice in immature particles, they have less access to RNA than does NC in mature particles [67].

Concluding remarks

It is clear from the work of many laboratories that the interactions of Gag proteins with RNA are of great significance for successful viral replication. Briefly, binding to RNA facilitates the assembly of virus particles from Gag protein molecules. Recent findings indicate that the Gag protein undergoes conformational changes leading to assembly when it attains a high local concentration, presumably as a result of cooperative binding to RNA. These changes in the protein are an important subject for future investigations. Gag must also select the vRNA out of an excess of other mRNAs, and it is now clear that in MLV, this selection stems from specific recognition of a sequence motif and its 3D context in a dimer of vRNA. This RNA structure remains to be elucidated, and it is not known whether other retroviruses follow a similar paradigm. Finally, Gag refolds RNAs, unwinding a tRNA molecule and catalyzing its hybridization to vRNA. Further refolding occurs when Gag is cleaved during virus maturation, but the mechanism by which refolding is coordinated with maturation is not understood.

Although the interactions of Gag with RNAs are crucial for retroviral replication, they are not yet well defined.

There have been many studies of the interactions of free HIV-1 NC protein with NAs in solution [75,80,81], but these interactions are remarkably complex [82]. The behavior of Gag is presumably complicated further by the presence of the MA domain, which also interacts with NAs, and of the CA domain, which mediates interactions between the Gag molecules themselves. To date, attempts to characterize Gag–NA interactions *in vitro* have yielded conflicting results. For example, the ‘site size’ of Gag proteins on RNA is not precisely known, nor is there clarity regarding the relative affinities of Gag and free NC proteins for specific NAs. The binding of Gag to RNA is expected to be cooperative, but this has never been demonstrated. One reason that the analysis of Gag–RNA interactions has proved so difficult is undoubtedly the fact that binding of Gag to NAs longer than approximately 20–40 bases leads to VLP assembly, so that the reactants are removed from the solution phase. The development of methods in which one of the binding partners is immobilized should ultimately help to overcome this problem.

Acknowledgments

This work was supported in part by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research, and by NIH Grant R01 GM065056 (to K.M.-F.). C.P.J. was supported by NIH Training Grant T32 GM008512. We thank Kevin Weeks for Figure 3.

References

- Coffin, J.M. *et al.* (1997) *Retroviruses*, Cold Spring Harbor Laboratory Press
- Campbell, S. and Vogt, V.M. (1995) Self-assembly *in vitro* of purified CA-NC proteins from Rous sarcoma virus and human immunodeficiency virus type 1. *J. Virol.* 69, 6487–6497
- Muriaux, D. *et al.* (2004) Role of murine leukemia virus nucleocapsid protein in virus assembly. *J. Virol.* 78, 12378–12385
- Ott, D.E. *et al.* (2003) Elimination of protease activity restores efficient virion production to a human immunodeficiency virus type 1 nucleocapsid deletion mutant. *J. Virol.* 77, 5547–5556
- Muriaux, D. *et al.* (2001) RNA is a structural element in retrovirus particles. *Proc. Natl. Acad. Sci. U.S.A.* 98, 5246–5251
- Campbell, S. *et al.* (2001) Modulation of HIV-like particle assembly *in vitro* by inositol phosphates. *Proc. Natl. Acad. Sci. U.S.A.* 98, 10875–10879
- Campbell, S. and Rein, A. (1999) *In vitro* assembly properties of human immunodeficiency virus type 1 Gag protein lacking the p6 domain. *J. Virol.* 73, 2270–2279
- Ma, Y.M. and Vogt, V.M. (2002) Rous sarcoma virus Gag protein-oligonucleotide interaction suggests a critical role for protein dimer formation in assembly. *J. Virol.* 76, 5452–5462
- Ma, Y.M. and Vogt, V.M. (2004) Nucleic acid binding-induced Gag dimerization in the assembly of Rous sarcoma virus particles *in vitro*. *J. Virol.* 78, 52–60
- Rulli, S.J., Jr *et al.* (2007) Selective and Nonselective Packaging of Cellular RNAs in Retrovirus Particles. *J. Virol.* 81, 6623–6631
- Zhang, Y. *et al.* (1998) Analysis of the assembly function of the human immunodeficiency virus type 1 gag protein nucleocapsid domain. *J. Virol.* 72, 1782–1789
- Crist, R.M. *et al.* (2009) Assembly properties of human immunodeficiency virus type 1 Gag-leucine zipper chimeras: implications for retrovirus assembly. *J. Virol.* 83, 2216–2225
- Keene, S.E. *et al.* (2010) 7SL RNA is retained in HIV-1 minimal virus-like particles as an S-domain fragment. *J. Virol.* 84, 9070–9077
- Briggs, J.A. *et al.* (2006) Cryo-electron microscopy reveals conserved and divergent features of gag packing in immature particles of Rous sarcoma virus and human immunodeficiency virus. *J. Mol. Biol.* 355, 157–168
- Briggs, J.A. *et al.* (2009) Structure and assembly of immature HIV. *Proc. Natl. Acad. Sci. U.S.A.* 106, 11090–11095
- Wright, E.R. *et al.* (2007) Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells. *EMBO J.* 26, 2218–2226
- Ako-Adjei, D. *et al.* (2005) The retroviral capsid domain dictates virion size, morphology, and coassembly of gag into virus-like particles. *J. Virol.* 79, 13463–13472
- Datta, S.A.K. *et al.* (2011) On the role of the SP1 domain in HIV-1 particle assembly: A molecular switch? *J. Virol.* 85, 4111–4121
- Alfadhli, A. *et al.* (2009) Analysis of human immunodeficiency virus type 1 matrix binding to membranes and nucleic acids. *J. Virol.* 83, 12196–12203
- Chukkappalli, V. *et al.* (2010) Opposing mechanisms involving RNA and lipids regulate HIV-1 Gag membrane binding through the highly basic region of the matrix domain. *Proc. Natl. Acad. Sci. U.S.A.* 107, 1600–1605
- Shkriabai, N. *et al.* (2006) Interactions of HIV-1 gag with assembly cofactors. *Biochemistry* 45, 4077–4083
- Datta, S.A.K. *et al.* (2007) Conformation of the HIV-1 gag protein in solution. *J. Mol. Biol.* 365, 812–824
- Datta, S.A.K. *et al.* (2007) Interactions between HIV-1 gag molecules in solution: an inositol phosphate-mediated switch. *J. Mol. Biol.* 365, 799–811
- Datta, S.A. *et al.* (2011) HIV-1 Gag extension: conformational changes require simultaneous interaction with membrane and nucleic acid. *J. Mol. Biol.* 406, 205–214
- Saad, J.S. *et al.* (2006) Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. *Proc. Natl. Acad. Sci. U.S.A.* 103, 11364–11369
- Jones, C.P. *et al.* (2011) Matrix domain modulates Hiv-1 Gag’s nucleic acid chaperone activity via inositol phosphate binding. *J. Virol.* 85, 1594–1603
- Kaye, J.F. and Lever, A.M. (1999) Human immunodeficiency virus types 1 and 2 differ in the predominant mechanism used for selection of genomic RNA for encapsidation. *J. Virol.* 73, 3023–3031
- Poon, D.T. *et al.* (2002) Human immunodeficiency virus type 1 preferentially encapsidates genomic RNAs that encode Pr55(Gag): functional linkage between translation and RNA packaging. *Virology* 293, 368–378
- Adam, M.A. and Miller, A.D. (1988) Identification of a signal in a murine retrovirus that is sufficient for packaging of nonretroviral RNA into virions. *J. Virol.* 62, 3802–3806
- Banks, J.D. *et al.* (1999) An Mpsi-containing heterologous RNA, but not env mRNA, is efficiently packaged into avian retroviral particles. *J. Virol.* 73, 8926–8933
- Hibbert, C.S. *et al.* (2004) mRNA molecules containing murine leukemia virus packaging signals are encapsidated as dimers. *J. Virol.* 78, 10927–10938
- Berkowitz, R.D. *et al.* (1995) 5′ regions of HIV-1 RNAs are not sufficient for encapsidation: implications for the HIV-1 packaging signal. *Virology* 212, 718–723
- D’Souza, V. and Summers, M.F. (2005) How retroviruses select their genomes. *Nat. Rev. Microbiol.* 3, 643–655
- Moore, M.D. and Hu, W.S. (2009) HIV-1 RNA dimerization: It takes two to tango. *AIDS Rev.* 11, 91–102
- Sakuragi, J.-i. *et al.* (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
- Miyazaki, Y. *et al.* (2009) An RNA structural switch regulates diploid genome packaging by Moloney murine leukemia virus. *J. Mol. Biol.* 396, 141–152
- Johnson, S.F. and Telesnitsky, A. (2010) Retroviral RNA dimerization and packaging: the what, how, when, where, and why. *PLoS Pathog.* 6
- Paillart, J.C. *et al.* (2004) Dimerization of retroviral RNA genomes: an inseparable pair. *Nat. Rev. Microbiol.* 2, 461–472
- Russell, R.S. *et al.* (2004) Is HIV-1 RNA dimerization a prerequisite for packaging? Yes, no, probably? *Retrovirology* 1, 23
- den Boon, J.A. *et al.* (2010) Cytoplasmic viral replication complexes. *Cell Host Microbe* 8, 77–85
- Garbitt-Hirst, R. *et al.* (2009) Genetic evidence for a connection between Rous sarcoma virus gag nuclear trafficking and genomic RNA packaging. *J. Virol.* 83, 6790–6797

- 42 Gudleski, N. *et al.* (2010) Directionality of nucleocytoplasmic transport of the retroviral gag protein depends on sequential binding of karyopherins and viral RNA. *Proc. Natl. Acad. Sci. U.S.A.* 107, 9358–9363
- 43 Basyuk, E. *et al.* (2005) The packaging signal of MLV is an integrated module that mediates intracellular transport of genomic RNAs. *J. Mol. Biol.* 354, 330–339
- 44 Smagulova, F. *et al.* (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
- 45 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
- 46 Gorelick, R.J. *et al.* (1988) Point mutants of Moloney murine leukemia virus that fail to package viral RNA: evidence for specific RNA recognition by a “zinc finger-like” protein sequence. *Proc. Natl. Acad. Sci. U.S.A.* 85, 8420–8424
- 47 Gorelick, R.J. *et al.* (1990) Noninfectious human immunodeficiency virus type 1 mutants deficient in genomic RNA. *J. Virol.* 64, 3207–3211
- 48 Amarasinghe, G.K. *et al.* (2000) NMR structure of the HIV-1 nucleocapsid protein bound to stem-loop SL2 of the psi-RNA packaging signal. Implications for genome recognition. *J. Mol. Biol.* 301, 491–511
- 49 De Guzman, R.N. *et al.* (1998) Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. *Science* 279, 384–388
- 50 D'Souza, V. and Summers, M.F. (2004) Structural basis for packaging the dimeric genome of Moloney murine leukaemia virus. *Nature* 431, 586–590
- 51 Gherghe, C. *et al.* (2010) Secondary structure of the mature ex virio Moloney murine leukemia virus genomic RNA dimerization domain. *J. Virol.* 84, 898–906
- 52 Gherghe, C. *et al.* (2010) Definition of a high-affinity Gag recognition structure mediating packaging of a retroviral RNA genome. *Proc. Natl. Acad. Sci. U.S.A.* 107, 19248–19253
- 53 Wilkinson, K.A. *et al.* (2008) High-Throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states. *PLoS Biol.* 6, e96
- 54 Wang, H. *et al.* (2003) Involvement of the matrix and nucleocapsid domains of the bovine leukemia virus Gag polyprotein precursor in viral RNA packaging. *J. Virol.* 77, 9431–9438
- 55 Rein, A. (2010) Nucleic acid chaperone activity of retroviral Gag proteins. *RNA Biol.* 7, 685–690
- 56 Cristofari, G. and Darlix, J.L. (2002) The ubiquitous nature of RNA chaperone proteins. *Prog. Nucleic Acid Res. Mol. Biol.* 72, 223–268
- 57 Rajkowitsch, L. *et al.* (2007) RNA chaperones, RNA annealers and RNA helicases. *RNA Biol.* 4, 118–130
- 58 Woodson, S.A. (2010) Taming free energy landscapes with RNA chaperones. *RNA Biol.* 7, 662–671
- 59 Levin, J.G. *et al.* (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
- 60 Levin, J.G. *et al.* (2010) Role of HIV-1 nucleocapsid protein in HIV-1 reverse transcription. *RNA Biol.* 7, 754–774
- 61 Wu, T. *et al.* (2010) Fundamental differences between the nucleic acid chaperone activities of HIV-1 nucleocapsid protein and Gag or Gag-derived proteins: Biological implications. *Virology* 405, 556–567
- 62 Cen, S. *et al.* (2000) Roles of Pr55(gag) and NCp7 in tRNA(3)(Lys) genomic placement and the initiation step of reverse transcription in human immunodeficiency virus type 1. *J. Virol.* 74, 10796–10800
- 63 Crawford, S. and Goff, S.P. (1985) A deletion mutation in the 5' part of the pol gene of Moloney murine leukemia virus blocks proteolytic processing of the gag and pol polyproteins. *J. Virol.* 53, 899–907
- 64 Fu, W. *et al.* (1994) Characterization of human immunodeficiency virus type 1 dimeric RNA from wild-type and protease-defective virions. *J. Virol.* 68, 5013–5018
- 65 Fu, W. and Rein, A. (1993) Maturation of dimeric viral RNA of Moloney murine leukemia virus. *J. Virol.* 67, 5443–5449
- 66 Huang, Y. *et al.* (1997) Effect of mutations in the nucleocapsid protein (NCp7) upon Pr160(gag-pol) and tRNA(Lys) incorporation into human immunodeficiency virus type 1. *J. Virol.* 71, 4378–4384
- 67 Stewart, L. *et al.* (1990) Properties of avian retrovirus particles defective in viral protease. *J. Virol.* 64, 5076–5092
- 68 Kafaie, J. *et al.* (2008) Mapping of nucleocapsid residues important for HIV-1 genomic RNA dimerization and packaging. *Virology* 375, 592–610
- 69 Meric, C. and Spahr, P.F. (1986) Rous sarcoma virus nucleic acid-binding protein p12 is necessary for viral 70S RNA dimer formation and packaging. *J. Virol.* 60, 450–459
- 70 Xing, L. *et al.* (2011) Coordinate roles of Gag and RNA helicase A in promoting the annealing of tRNA_{Lys3} to HIV-1 RNA. *J. Virol.* 85, 1847–1860
- 71 Tompa, P. and Csermely, P. (2004) The role of structural disorder in the function of RNA and protein chaperones. *FASEB J.* 18, 1169–1175
- 72 Feng, Y.X. *et al.* (1999) The human immunodeficiency virus type 1 Gag polyprotein has nucleic acid chaperone activity: possible role in dimerization of genomic RNA and placement of tRNA on the primer binding site. *J. Virol.* 73, 4251–4256
- 73 Roldan, A. *et al.* (2005) A HIV-1 minimal gag protein is superior to nucleocapsid at in vitro annealing and exhibits multimerization-induced inhibition of reverse transcription. *J. Biol. Chem.* 280, 17488–17496
- 74 Ali, M.B. *et al.* (2007) Structural requirements for nucleocapsid protein-mediated dimerization of avian leukosis virus RNA. *J. Mol. Biol.* 372, 1082–1096
- 75 Egele, C. *et al.* (2007) The single-finger nucleocapsid protein of Moloney murine leukemia virus binds and destabilizes the TAR sequences of HIV-1 but does not promote efficiently their annealing. *Biochemistry* 46, 14650–14662
- 76 Qualley, D.F. *et al.* (2010) C-terminal domain modulates the nucleic acid chaperone activity of human T-cell leukemia virus type 1 nucleocapsid protein via an electrostatic mechanism. *J. Biol. Chem.* 285, 295–307
- 77 Stewart-Maynard, K.M. *et al.* (2008) Retroviral nucleocapsid proteins display nonequivalent levels of nucleic acid chaperone activity. *J. Virol.* 82, 10129–10142
- 78 Guo, F. *et al.* (2009) Roles of Gag and NCp7 in facilitating tRNA(Lys)(3) Annealing to viral RNA in human immunodeficiency virus type 1. *J. Virol.* 83, 8099–8107
- 79 Kleiman, L. *et al.* (2004) The selective packaging and annealing of primer tRNA_{Lys3} in HIV-1. *Curr. HIV Res.* 2, 163–175
- 80 Hagan, N. and Fabris, D. (2003) Direct mass spectrometric determination of the stoichiometry and binding affinity of the complexes between nucleocapsid protein and RNA stem-loop hairpins of the HIV-1 Psi-recognition element. *Biochemistry* 42, 10736–10745
- 81 Shubsda, M.F. *et al.* (2002) Affinities of packaging domain loops in HIV-1 RNA for the nucleocapsid protein. *Biochemistry* 41, 5276–5282
- 82 Fisher, R.J. *et al.* (2006) Complex interactions of HIV-1 nucleocapsid protein with oligonucleotides. *Nucleic Acids Res.* 34, 472–484
- 83 Iwatani, Y. *et al.* (2003) Efficient initiation of HIV-1 reverse transcription in vitro Requirement for RNA sequences downstream of the primer binding site abrogated by nucleocapsid protein-dependent primer-template interactions. *J. Biol. Chem.* 278, 14185–14195
- 84 Morris, S. and Leis, J. (1999) Changes in Rous sarcoma virus RNA secondary structure near the primer binding site upon tRNA_{Trp} primer annealing. *J. Virol.* 73, 6307–6318
- 85 Bilbille, Y. *et al.* (2009) The structure of the human tRNA_{Lys3} anticodon bound to the HIV genome is stabilized by modified nucleosides and adjacent mismatch base pairs. *Nucleic Acids Res.* 37, 3342–3353