Trends in Microbiology

CellPress
REVIEWS

Review

# RNA Packaging in HIV

Alan Rein [1,*]

**Successful replication of the AIDS retrovirus, HIV, requires that its genomic RNA be packaged in assembling virus particles with high fidelity. However, cellular mRNAs can also be packaged under some conditions. Viral RNA (vRNA) contains a 'packaging signal' (ψ) and is packaged as a dimer, with two vRNA monomers joined by a limited number of base pairs. It has two conformers, only one of which is capable of dimerization and packaging. Recent years have seen important progress on the 3D structure of dimeric ψ. Gag, the protein that assembles into the virus particle, interacts specifically with ψ, but this is obscured under physiological conditions by its high nonspecific affinity for any RNA. New results suggest that vRNA is selected for packaging because ψ nucleates assembly more efficiently than other RNAs.**

## Highlights

Highly selective packaging of HIV-1 RNA into virus particles is required for successful virus replication. However, cellular RNAs can also be packaged under some conditions. Thus, in infected cells, viral RNA is competing with an excess of cellular RNA for incorporation into assembling particles.

Viral RNA can fold into two alternative conformers. Only one of these can dimerize, as required for packaging into virus particles.

Virus particles are assembled from the virus-coded Gag protein. At physiological salt concentrations, Gag binds tightly to any RNA. Therefore, selective packaging of viral RNA does not result from a unique high-affinity interaction between Gag and this RNA.

Viral RNA supports virus particle assembly more efficiently than other RNAs. It probably nucleates assembly more rapidly than other RNAs; this would explain its selective packaging.

## Assembly of the HIV Virus Particle

HIV, the causative agent of the AIDS pandemic, is a retrovirus. As with all retroviruses, the virus particle is roughly spherical and around 100 nm in diameter. It is initially assembled as an immature virion, containing its RNA genome and constructed principally from ~2000 copies of a single protein, called Gag. The immature particles are released by budding from the surface of the virus-producing cell (taking a bit of plasma membrane from the cell), whereupon the virus-coded protease present in the particle cleaves Gag into three or more fragments. These maturation cleavages lead to a wholesale reorganization of the structure of the particle and are absolutely required for infectivity [1].

## The Problem: Selective Incorporation of Viral RNA into the Assembling Virus Particle

As the RNA is the genetic material to be passed on to the next generation of viruses, it is obvious that efficient viral replication requires the incorporation of this RNA into the assembling particle. In turn, this depends upon great selectivity, as the viral RNA (vRNA) is surrounded by a vast excess of cellular RNAs. Recent findings have helped to shed new light on this problem; how the selective packaging is accomplished by HIV is the subject of this review.

Viruses use a wide variety of strategies to achieve this goal. In some, such as tobacco mosaic virus, virus assembly is initiated by the specific interaction between the viral coat protein and a site on the vRNA [2]. In others, virus assembly is confined to a specific region in the cell, so that the genomes to be packaged are not admixed with nucleic acid molecules of cellular origin [3].

Remarkably, the mechanisms used by retroviruses are quite different from those mentioned above. In fact, under some experimental conditions retroviral Gag proteins can assemble into virus-like particles without packaging vRNA. It was reported over 40 years ago (before the isolation of HIV) that when cells infected with murine leukemia virus (MLV) are treated with actinomycin D, which blocks Pol II transcription, they continue to release particles, but these particles lack vRNA [4]. Subsequent studies clarified and generalized this unexpected observation: when retroviral Gag proteins are expressed from an mRNA, rather than from the intact viral genome, they successfully assemble into normal-looking particles with good efficiency. In other words, retroviral particle assembly does not depend upon the vRNA; therefore, retroviruses must employ some specific mechanism to insure the selective packaging of this RNA.

[1]HIV Dynamics and Replication Program, Center for Cancer Research, National Cancer Institute, Frederick, MD 21702-1201, USA

*Correspondence: reina@mail.nih.gov (A. Rein).

## The Players

### Gag Protein

The two principal participants in RNA packaging are, of course, the Gag protein and the vRNA. A retroviral Gag protein is ~50–60 kDa and is composed of several independently folded domains connected by flexible linkers. During viral maturation, the virus-coded protease cleaves Gag between these domains; the resulting fragments are the principal constituents of the mature particle. The domains always include, from N to C-terminal, Matrix (MA), Capsid (CA), and Nucleocapsid (NC). In HIV, there is one additional domain, called 'p6', at the extreme C terminus, and two very short (<20 residue) cleavage products, called 'spacers' (Gag is shown schematically in Figure 1). The Gag proteins in the immature particle are rod-shaped and are arranged radially in the particle, with their N termini in contact with the lipid bilayer covering the surface of the particle and their C termini facing inwards.

To a first approximation, the MA domain (which in most retroviruses has a fatty acid modification at its N terminus) functions in the interactions of Gag with the plasma membrane of the virus-producing cell. The CA domain, together with the SP1 'spacer' between CA and NC, plays a major role in the interactions between Gag protein molecules leading to particle assembly. The NC domain functions predominantly in the interactions of Gag with RNA, while p6 functions in the release of the assembled immature particle from the surface of the virus-producing cell.

Studies attempting to understand RNA packaging in retroviruses have focused largely on the NC domain. It is very basic and contains two zinc fingers (except in some genera, where it has only one). Point mutations disrupting the zinc fingers interfere with the specific packaging of vRNA [5,6]. As it is only 55 of the 500 amino acids in HIV Gag, reducing the problem from Gag-RNA to NC-RNA interactions is a tremendous simplification, for example, in analysis of molecular structures.

### vRNA

Retroviral genomes are roughly 8–10-kb RNA molecules. Just like most cellular mRNA molecules, they are capped at their 5′ ends and have a poly A tail at their 3′ ends; in fact, the full-length viral genomic RNA is not only packaged in assembling progeny virus particles, but is also the mRNA in the cell for Gag and for a larger fusion protein containing both Gag and the three virus-coded enzymes (protease, reverse transcriptase, and integrase). This RNA is selected for packaging because it contains, near its 5′ end, a region called the 'packaging signal', or ψ. Packaging signals may be several hundred nucleotides in length; although they are largely in the 5′ untranslated region of the vRNA, they frequently extend into the region coding for Gag (the gene closest to the 5′ end of the RNA).

## RNA Packaging

All studies of the genomic RNA within retrovirus particles have shown that the RNA is packaged as a dimer. That is, virions contain two genomic RNA molecules of the same (positive- or sense-

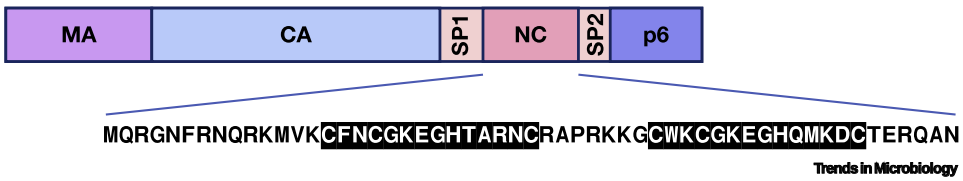


Figure 1. HIV Gag Protein and its Nucleocapsid Domain. Gag is the building block of HIV virus particles. After the virus is released from the cell, Gag is cleaved by the viral protease into six fragments, which are all present in the mature particle. These are (from N to C terminus) Matrix (MA), Capsid (CA), 'Spacer 1', Nucleocapsid (NC), 'Spacer 2', and p6. The interactions of Gag with RNA are mediated principally by its NC domain. The lower portion of the figure shows the sequence of the NC protein of the BH10 isolate of HIV-1. The two zinc fingers in NC are highlighted in the sequence.

strand) polarity, linked together by a limited number of base pairs. When the RNA is copied into double-stranded DNA during infection, the reverse transcriptase jumps repeatedly between the two RNA copies, so that the DNA form of the viral genome, produced during infection, is a recombinant between the two RNA copies in the virion. This high-frequency recombination is a major source of variation in the viral sequence; the presence of two copies in the virion is presumably also insurance against random RNA strand breakage [7].

The universality with which dimers are packaged raises the possibility that dimerization is a requirement for packaging; perhaps the real packaging signal consists not merely of the linear RNA sequence of ψ, but also includes the 3D structure formed upon dimerization. This idea drew strong support from the elegant experiments of Sakuragi *et al.*, in which two packaging signals were placed in tandem on a single RNA molecule; these RNAs were then, in contrast with normal genomes, packaged as monomers. Presumably, the two tandem packaging signals were able to interact with each other, forming an intramolecular structure mimicking the intermolecular dimeric structure present in normal packaged genomes [8].

As mentioned above, when Gag is expressed in cells in the absence of vRNA, it assembles efficiently into virus-like particles. Strikingly, these particles are not empty, but contain the same amount of RNA as normal wild type particles. Detailed analysis showed that the RNAs packaged in place of vRNA in these particles are cellular mRNA molecules. The packaged mRNAs are nearly a random sampling of the mRNAs present in the cells, except that: (i) mitochondrial mRNAs are excluded, and (ii) there is a slight preference for packaging mRNAs with long 3′ untranslated regions [9,10].

Retrovirus particle assembly has also been studied in defined *in vitro* systems and the results of these studies generally dovetail well with the findings described above. As first observed in an avian retroviral system, a retroviral Gag protein purified from bacteria assembles into virus-like particles upon addition of nucleic acid [11]. This was subsequently demonstrated with recombinant HIV Gag protein; for technical reasons, this protein differed from authentic Gag in lacking the fatty acid moiety at its N terminus and the p6 domain at its C terminus [12]. Also, in the HIV case, the particles are morphologically abnormal, but their structural defect was corrected if there was a sizable deletion within the MA domain, or if an additional cofactor, inositol pentakisphosphate, was added to the assembly reaction [13]. Remarkably, nearly any single-stranded nucleic acid longer than ~30 bases can support particle assembly under these conditions [12]. Thus, both within the cell and *in vitro*, particle assembly seems to require nucleic acid, but there is no obvious specificity in this requirement.

## Structural Advances

Ultimately, one would hope to understand the mechanisms underlying selective packaging of vRNA at a detailed molecular level. Thus, many investigators have devoted years of effort to structural characterization of Gag, vRNA, and their interactions. These studies have been very challenging, but important progress has been made in recent years.

### Gag Protein

As noted above, Gag is a multidomain protein with flexible linkers between the independently folded domains. Thus, while there is extensive, detailed structural information about the individual domains, the flexibility of the connections between them means that there is no unique structure for the full-length protein.

The domain with the greatest relevance for RNA packaging is the NC domain. The NC protein has been exhaustively analyzed from many points of view and performs a remarkable array of

functions during retroviral replication. While the NC domain of Gag participates in selecting the genomic RNA during assembly, the free NC protein is believed to coat all of the RNA within a mature virus particle. It is also an essential cofactor during the synthesis of DNA in the initial stages of infection. This stems largely from its nucleic acid chaperone activity (i.e., its ability to catalyze the rearrangement of nucleic acids into the most thermodynamically stable structures) [14]. In fact, during virus assembly the Gag protein, using its NC domain, catalyzes the unwinding of a specific cellular transfer RNA (tRNA) and its annealing to a complementary 18-base sequence on the viral genomic RNA: this is essential for replication as the annealed tRNA will serve as the primer for DNA synthesis when the virus infects a new host cell [15].

The sequence of HIV NC protein is shown in Figure 1. It is highly basic (pI 9.93); its basic character is, of course, a major factor in its interactions with nucleic acids. It also contains two zinc fingers, with the zinc-chelating cysteine and histidine residues arranged as C-$X_2$-C-$X_4$-H-$X_4$-C. In fact, all retroviral NC proteins (except those of the spumaretroviruses, an outlier group of viruses not considered here) contain one or two zinc fingers, and this spacing of the cysteines and histidines is absolutely conserved among them. The basic residues are found particularly in the N-terminal portion of the protein and between the two fingers. The fingers, as well as the basic residues, are of great importance in interactions of NC (and therefore of Gag) with nucleic acids. As electrostatic interaction is nearly independent of the nucleotide sequence of the target nucleic acid, it is the fingers that are largely responsible for specificity in nucleic acid-binding. Thus, viruses whose zinc fingers are disrupted by single point mutations in the conserved cysteines or histidines can successfully assemble into virus-like particles, but they fail to selectively package genomic RNA [5,6].

No 3D structure has been determined for free NC protein. However, NMR studies indicate that both N-terminal and C-terminal regions of the protein are flexible, while the zinc fingers are rigid. Michael Summers and his colleagues have obtained two NMR-derived structures of NC in complex with two different RNA stem-loops, which are parts of ψ [16,17]. It is striking that the conformation of NC is different in the two complexes: in other words, the overall structure of NC is not fixed, but can adapt to optimize binding to RNA. Specifically, the highly basic N-terminal tail of the protein forms a $3_{10}$ helix in both structures, but the position of this helix is different in the two structures. The relative orientations of the two zinc fingers are also different in the structures.

While this discussion focuses on the NC domain, it has recently become clear that the MA domain can also bind RNAs, and the biological significance of these interactions is under active investigation [18–20].

## RNA

As noted above, retroviral genomic RNAs are always packaged as dimers. The full-length vRNA is not only the genome to be packaged into progeny virus particles, but also serves as the mRNA for Gag and for the Gag-Pol fusion protein. Computer predictions suggest that nucleotides 233–370 of the RNA (i.e., a substantial portion of the ψ region) could fold into four stem-loops, termed SL1–SL4. The loop of SL1 would contain the six-base palindromic sequence ('dimerization initiation site' or 'DIS'), by which the two monomers in the vRNA dimer are initially base-paired with each other. In this structure, the initiation codon for Gag (nt 337) is at the base of SL4, and thus might be in a poor context for translational initiation. However, as first pointed out by Berkhout and colleagues, the same sequence can be folded into a different secondary structure. In this alternate conformation, the DIS palindrome is occluded by intramolecular base-pairing and thus unavailable for intermolecular interactions [21,22], while the Gag initiation codon is partially exposed. They proposed that the two structures correspond to the two functions of this RNA species: (i) packaging into virions, which evidently requires dimerization; and (ii) translation into

Gag and the Gag-Pol fusion protein. They called the two structures 'branched multiple hairpin (BMH)' and 'long distance interaction (LDI)', respectively. It has recently been found that HIV RNA molecules in the infected cell differ with respect to the exact transcriptional start site, and that this in turn influences the proportions of the two conformations and the probability that a given RNA molecule will be packaged [23,24].

While the details of the two proposed structures have evolved with time, the basic idea is now widely accepted: the RNA can either assume a conformation in which the DIS is exposed, permitting dimerization as required for packaging, while the Gag AUG is base-paired with upstream sequences; or alternatively, a conformation in which the DIS is buried but the Gag AUG is exposed (see Figure 2A, in which the two structures are designated 'DIS-sequestered' and 'DIS-exposed'). There has followed a tremendous effort, using many approaches and extending over many years, to determine the 3D structure of the 'BMH' packageable form. In a major step forward in 2015, Keane *et al*. presented a model for a similar structure [25]. The model is based on NMR analysis of smaller RNAs designed to mimic the relevant portions of the real structure; in addition, these investigators constructed the RNA from fragments, some of which were deuterated and thus invisible to the NMR. In support of this strategy, they show that binding of NC protein to this RNA resembles its binding to the full-length RNA. Even with the judicious removal of unwanted sequences, the structure encompasses 155 bases, and is thus considerably larger than any RNA structure previously determined by NMR.

The structure is shown in Figure 2B. Its most notable feature is a tandem pair of three-way junctions: one between the SL3 stem-loop (sometimes called 'ψ'), the U5:AUG stem, and the H1 connecting helix; and the other between H1, the primer-binding site stem-loop, and the DIS stem-loop (in which the major splice donor is buried). This RNA is designed to represent one of the monomers in the dimeric RNA that will be packaged: in this model structure, the DIS stem-

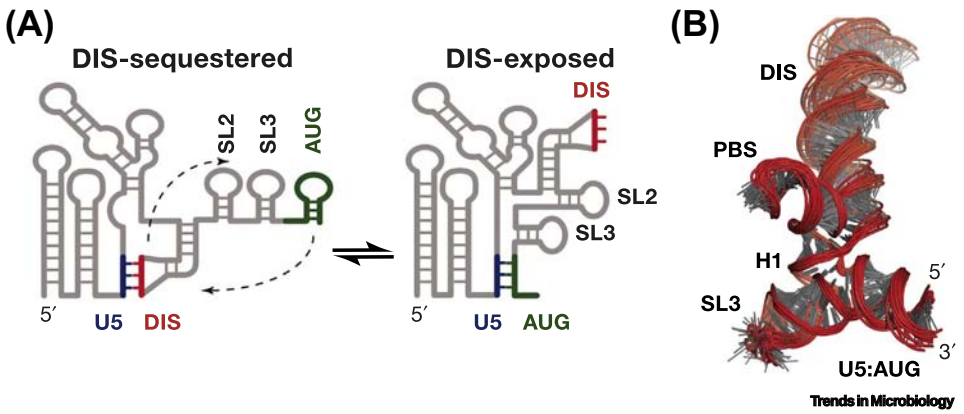


Figure 2. The Packaging Signal within HIV RNA. (A) The 5′ end of HIV RNA (nucleotides ~1–345) can be folded into two alternative structures. In one of these, indicated here as 'DIS-sequestered' but resembling the earlier proposal of 'long distance interaction' (LDI) [21,22], the AUG codon for initiation of Gag translation is exposed, while the palindromic dimer initiation site (DIS) is base-paired to the 'U5' region of the molecule. In contrast, in the 'DIS-exposed' structure (resembling the branched multiple hairpin or BMH model [21,22]), the Gag AUG is occluded by base-pairing and the DIS is exposed. Exposure enables the DIS to engage in intermolecular base-pairing with the DIS of another viral RNA molecule; the resulting RNA dimer is a substrate for selective packaging into assembling virus particles. (B) NMR-based model of 3D structure of the DIS-exposed conformer, representing nucleotides 105–344 of HIV RNA. It contains two consecutive three-way junctions: one between SL3, the U5:AUG helix, and the connecting helix H1, and the other between H1, the primer-binding site (PBS) stem-loop, and the DIS stem-loop. Although nucleotides 282–300 can theoretically be folded into a small stem-loop, indicated as SL2 in (A), in the DIS-exposed structure they actually pair with bases immediately 3′ of 'U5' to form H1. Adapted from [25,35].

loop and the U5:AUG stem-loop are both intramolecular, but in the authentic dimer, the stems are apparently formed by intermolecular base-pairing [26].

As mentioned above, structures were previously determined for complexes of NC protein with individual RNA stem-loops [16,17]. A striking feature of these structures was the intimate interaction between the NC zinc fingers and unpaired or mispaired guanosine residues in the RNA, and it has often been suggested that these bases might be important in the recognition of genomic RNA by the NC domain of Gag. In fact, yet another attractive feature of the new structure [25] is that it contains a number of unpaired G residues. These bases are tightly bound by NC protein within mature, infectious HIV virions [27]. Mutant vRNAs in which these Gs are replaced by As are only very inefficiently packaged into virions [25]; thus, some or all of these Gs are necessary for selective packaging of genomic RNA. Interestingly, independent studies of RNA packaging in MLV (from a distinct retroviral genus) have arrived at analogous conclusions [28].

## Mechanism of Selective Packaging

It was noted above that in the absence of genomic RNA, cellular mRNAs are packaged into virus-like particles. This observation shows that these RNAs, like vRNA, are capable of supporting particle assembly. Thus, in the virus-producing cell, the genomic RNA is in competition with a very large excess of cellular RNAs for incorporation into the assembling particle; it is selectively packaged because ψ gives it an advantage in this competition. The nature of this advantage is evidently the key to selective packaging.

One possibility could lie in affinities: perhaps ψ is a high-affinity binding site for Gag. It is technically challenging to measure the affinity of Gag for an RNA, since under many experimental conditions Gag-RNA complexes assemble into virus-like particles, precluding the determination of an equilibrium constant for the binding. This difficulty can be overcome, however, by making measurements at extremely low concentrations.

A pioneering study was carried out by Webb *et al*., who used fluorescence anisotropy measurements on very dilute, fluorophore-tagged RNAs to detect binding of recombinant Gag protein [29]. Comas-Garcia *et al*. used different techniques to measure binding [30], but the results of these two studies are broadly consistent with each other and will be summarized together. Both studies used recombinant HIV Gag protein lacking the myristate modification at its N terminus and the p6 domain at its C terminus. (As it has been suggested that p6 influences interactions with RNA [31], its absence in this work adds one caveat to the conclusions.)

These analyses showed that at physiological salt concentrations, Gag binds to any RNA, either with or without ψ, with very high affinity (with a $K_D$ on the order of 25–50 nM [30]). This basic result appears to exclude the hypothesis that selective packaging is due to a special affinity of Gag for ψ.

At the same time, further experimentation showed that this high-affinity binding to RNA represents the sum of both specific and nonspecific interactions. There are several ways that these interactions can be distinguished experimentally. For example, as the nonspecific interaction is largely electrostatic, it is ablated by raising the salt concentration in the assay buffer. Assays at increased ionic strength (or with other modifications suppressing nonspecific binding) showed that the specific component of binding to ψ is far higher-affinity than the specific component of binding to other RNAs. In other words, there is a large difference between binding to ψ and binding to other RNAs, but this difference is virtually undetectable under physiological conditions because it is obscured by nonspecific interactions. Webb *et al*. also reported that nonspecific binding to a control RNA involves more positive charges in the protein than specific binding to

ψ, and suggested the possibility that specific binding can be attributed largely or entirely to the NC domain, while nonspecific binding engages both the MA and NC domains [29].

The similarity in affinities for ψ and control RNAs under physiological conditions implies that the advantage of ψ-containing RNA in the competition for incorporation into virions is not thermodynamic in nature; but of course, it could still be kinetic. The data are all consistent with the hypothesis that ψ supports particle assembly more efficiently than cellular RNAs. This hypothesis was recently tested experimentally in an *in vitro* assembly system [32] and all of the data were strongly supportive of this proposal. Thus, Gag was titrated into solutions of a ψ-containing RNA and control RNAs *in vitro*, and the extent of particle assembly was assessed by monitoring the shift of the RNA into large, rapidly sedimenting structures. As shown in Figure 3, there was significantly more particle assembly in the presence of ψ than in the other RNAs; this difference is particularly notable at low Gag concentrations [32]. Directly analogous results were obtained in transfected human cells: when Gag is at low, limiting concentrations, the presence in the cells of a ψ-containing RNA markedly enhances particle release [33].

## Concluding Remarks

Selective packaging of ψ-containing RNA might seem difficult to explain in view of the facts that virtually any RNA can support assembly and that under physiological conditions, Gag does not bind much more tightly to ψ than to control RNA. However, particle assembly is accomplished by interactions among a large number of Gag molecules. The riddle is easily solved if one assumes that binding to ψ promotes these Gag–Gag interactions more efficiently than binding to other RNAs. Moreover, it seems plausible that there are two stages in particle assembly: (i) initiation or nucleation and (ii) polymerization. As the particle is a hexameric lattice of Gag molecules, one could also imagine that the nucleating event is formation of an initial hexamer, while polymerization is the accretion of Gag molecules onto this hexamer.

These ideas are all consistent with the hypothesis that ψ-containing RNAs are selectively packaged because nucleation, the rate-limiting step in assembly, occurs more rapidly on ψ than on other RNAs. Presumably the specific interaction of Gag with ψ leads to nucleation faster than the nonspecific interactions with other RNAs. This idea also appears to explain why ψ is so large: it must contain binding sites for multiple (perhaps six) Gag molecules in order to promote formation of the initial hexamer. A cartoon depicting this speculative mechanism is shown in Figure 4. While we have couched this proposal in qualitative terms, it is similar to the scenarios

## Outstanding Questions

What is the complete 3D structure of the dimeric ψ (the real substrate of selective packaging)?

What role do regions of Gag other than NC, such as MA and p6, play in selective packaging?

What is the role of the cellular context in RNA packaging and particle assembly? Assembly is completed in association with the plasma membrane, but is it initiated elsewhere in the cell?

We are postulating the existence of a discrete intermediate complex, the product of the 'nucleation' step in particle assembly. Does this exist? What is its composition and structure?

The selectivity for vRNA appears to arise from quantitative, rather than qualitative, differences between the interactions of this RNA with Gag and those of other RNAs with Gag. What is the magnitude of these differences and how exactly are they manifested?

Can new understanding of RNA packaging be exploited in new antiviral strategies?

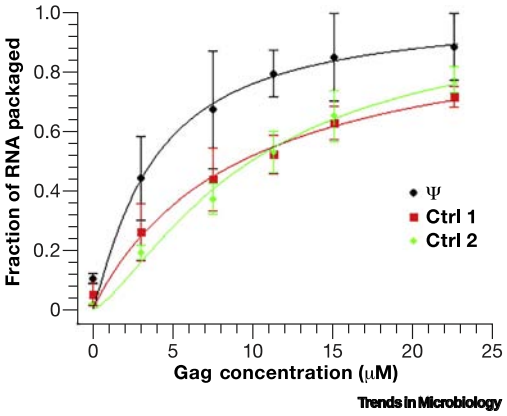


Figure 3. Ψ-Containing RNA Supports Particle Assembly More Efficiently than Control RNAs. Recombinant HIV-1 Gag protein was titrated into solutions of 401-base fluorophore-tagged RNAs. One of these, representing nucleotides 200–600 of HIV RNA, contained Ψ, while the other two did not. The incorporation of the RNAs into virus-like particles was monitored by sedimenting them through sucrose gradients. The results show that the Ψ-containing RNA supports particle formation more efficiently than the control RNAs. Adapted from [32].

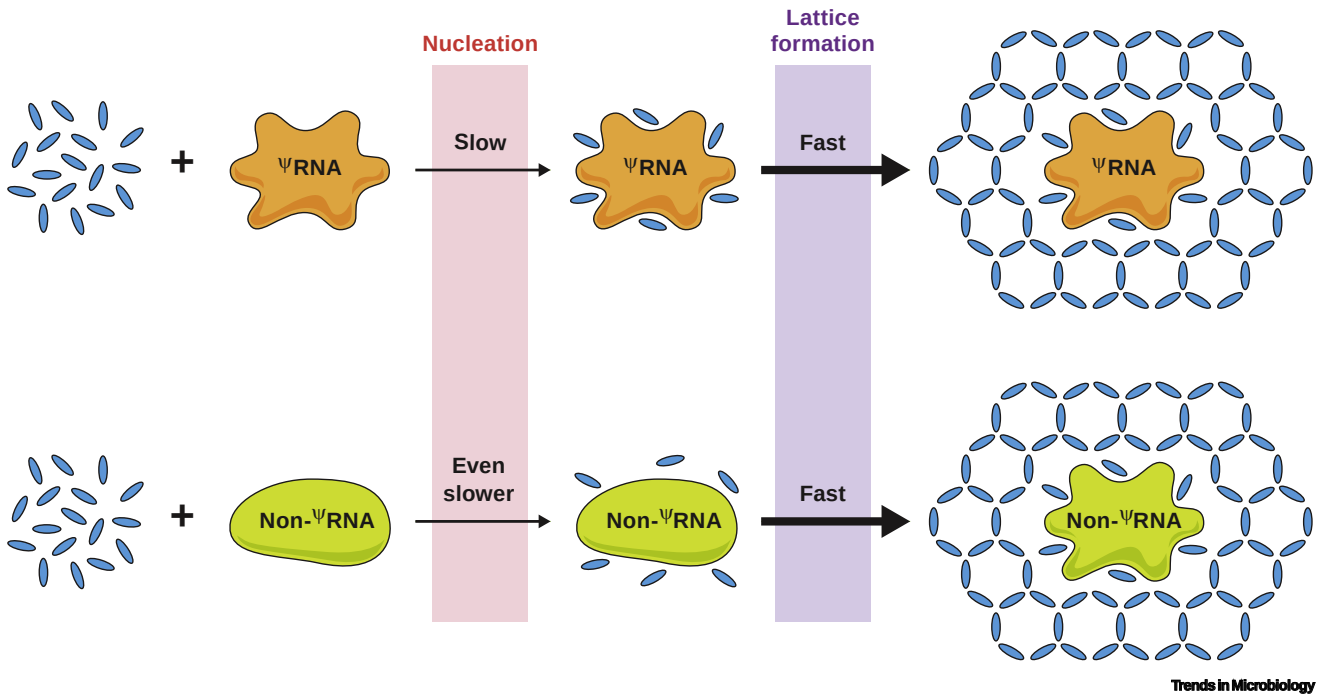


Figure 4. Model Explaining Selective Packaging of Ψ-Containing RNA. Either viral genomic RNA, which contains Ψ, or cellular mRNAs, which lack Ψ, can be packaged into assembling virus particles. We propose that assembly proceeds in two steps, nucleation and lattice formation. We suggest that nucleation is the rate-limiting step in assembly and that it occurs more rapidly on Ψ than on other RNAs. Gag molecules are represented as blue ovals, while the RNAs are shown as larger colored blobs.

considered quantitatively by Hagan [34]. As several aspects of the overall problem have become much clearer in recent years, we can hope that a comprehensive understanding will be available soon (see Outstanding Questions).

## Acknowledgments

Work in the author's laboratory is supported by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research, and in part with funds from the Intramural AIDS Targeted Antiviral Therapy Program.

## References

1. Swanstrom, R. and Wills, J.W. (1997) Synthesis, assembly, and processing of viral proteins. In *Retroviruses* (Coffin, J.M., et al., eds), pp. 263–334, Cold Spring Harbor Laboratory Press
2. Butler, P.J. (1984) The current picture of the structure and assembly of tobacco mosaic virus. *J. Gen. Virol.* 65, 253–279
3. den Boon, J.A. *et al.* (2010) Cytoplasmic viral replication complexes. *Cell Host Microbe* 8, 77–85
4. Levin, J.G. *et al.* (1974) Deficiency of 60 to 70S RNA in murine leukemia virus particles assembled in cells treated with actinomycin D. *J. Virol.* 14, 152–161
5. 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
6. Gorelick, R.J. *et al.* (1990) Noninfectious human immunodeficiency virus type 1 mutants deficient in genomic RNA. *J. Virol.* 64, 3207–3211
7. Rhodes, T. *et al.* (2003) High rates of human immunodeficiency virus type 1 recombination: near-random segregation of markers one kilobase apart in one round of viral replication. *J. Virol.* 77, 11193–11200
8. Sakuragi, J. *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
9. Rulli Jr., S.J. *et al.* (2007) Selective and nonselective packaging of cellular RNAs in retrovirus particles. *J. Virol.* 81, 6623–6631
10. Comas-Garcia, M. *et al.* (2016) On the selective packaging of genomic RNA by HIV-1. *Viruses* 8, E246
11. 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
12. 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
13. 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
14. 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
15. Rein, A. (2010) Nucleic acid chaperone activity of retroviral Gag proteins. *RNA Biol.* 7, 61–66
16. 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
17. 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
18. Alfadhli, A. *et al.* (2011) HIV-1 matrix protein binding to RNA. *J. Mol. Biol.* 410, 653–666
19. Chukkapalli, 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
20. 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
21. Abbink, T.E. and Berkhout, B. (2003) A novel long distance base-pairing interaction in human immunodeficiency virus type 1 RNA occludes the Gag start codon. *J. Biol. Chem.* 278, 11601–11611
22. Huthoff, H. and Berkhout, B. (2001) Two alternating structures of the HIV-1 leader RNA. *RNA* 7, 143–157
23. Kharytonchyk, S. *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
24. Masuda, T. *et al.* (2015) Fate of HIV-1 cDNA intermediates during reverse transcription is dictated by transcription initiation site of virus genomic RNA. *Sci. Rep.* 5, 17680
25. Keane, S.C. *et al.* (2015) RNA structure. Structure of the HIV-1 RNA packaging signal. *Science* 348, 917–921
26. Keane, S.C. *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
27. 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
28. 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
29. Webb, J.A. *et al.* (2013) Distinct binding interactions of HIV-1 Gag to Psi and non-Psi RNAs: implications for viral genomic RNA packaging. *RNA* 1078–1088
30. Comas-Garcia, M. *et al.* (2017) Dissection of specific binding of HIV-1 Gag to the 'packaging signal' in viral RNA. *eLife* 6, e27055
31. Dubois, N. *et al.* (2018) The C-terminal p6 domain of the HIV-1 Pr55(Gag) precursor is required for specific binding to the genomic RNA. *RNA Biol.* 15, 923–936
32. Comas-Garcia, M. *et al.* (2018) Efficient support of virus-like particle assembly by the HIV-1 packaging signal. *Elife* 7, e38438
33. Dilley, K.A. *et al.* (2017) Interactions between HIV-1 Gag and viral RNA genome enhance virion assembly. *J. Virol.* 91, e02319-02316
34. Perlmutter, J.D. and Hagan, M.F. (2015) The role of packaging sites in efficient and specific virus assembly. *J. Mol. Biol.* 427, 2451–2467
35. Lu, K. *et al.* (2011) NMR detection of structures in the HIV-1 5'-leader RNA that regulate genome packaging. *Science* 334, 242–245