ANNUAL REVIEWS

Annual Review of Virology

Multiple, Switchable Protein:RNA Interactions Regulate Human Immunodeficiency Virus Type 1 Assembly

Paul Bieniasz[1] and Alice Telesnitsky[2]

[1]Laboratory of Retrovirology and Howard Hughes Medical Institute, The Rockefeller University, New York, NY 10065, USA; email: pbieniasz@rockefeller.edu

[2]Department of Microbiology and Immunology, University of Michigan, Ann Arbor, Michigan 48109, USA; email: ateles@umich.edu

Annu. Rev. Virol. 2018. 5:165–83

First published as a Review in Advance on July 26, 2018

The *Annual Review of Virology* is online at virology.annualreviews.org

https://doi.org/10.1146/annurev-virology-092917-043448

Keywords

HIV-1, RNA, Gag, packaging, assembly

Abstract

Human immunodeficiency virus type 1 (HIV-1) particle assembly requires several protein:RNA interactions that vary widely in their character, from specific recognition of highly conserved and structured viral RNA elements to less specific interactions with variable RNA sequences. Genetic, biochemical, biophysical, and structural studies have illuminated how virion morphogenesis is accompanied by dramatic changes in the interactions among the protein and RNA virion components. The 5′ leader RNA element drives RNA recognition by Gag upon initiation of HIV-1 assembly and can assume variable conformations that influence translation, dimerization, and Gag recognition. As Gag multimerizes on the plasma membrane, forming immature particles, its RNA binding specificity transiently changes, enabling recognition of the A-rich composition of the viral genome. Initiation of assembly may also be regulated by occlusion of the membrane binding surface of Gag by tRNA. Finally, recent work has suggested that RNA interactions with viral enzymes may activate and ensure the accuracy of virion maturation.

1. INTRODUCTION

Retrovirus particles are composed of lipid, protein, and RNA, and their assembly in infected cells occurs through a series of protein:protein, protein:lipid, protein:RNA, and RNA:RNA interactions. Virions are initially constructed as immature noninfectious particles, and subsequent viral protease–catalyzed cleavage events remodel the particle into its mature infectious form. Initial particle assembly and subsequent remodeling are accompanied by dramatic changes in the nature of the interactions among protein and RNA components. Herein, we describe and discuss the protein:RNA interactions that occur during the assembly of human immunodeficiency virus type 1 (HIV-1) particles and how profound changes in these interactions ultimately enable the morphogenesis of mature virions that can productively infect new cells.

HIV-1 assembly is driven by a single, multidomain protein, Gag. Proceeding from the N terminus to the C terminus of Gag, several protein domains are encountered, each of which has a defined role (**Figure 1*a*,*b***). The matrix (MA) domain is N-terminally myristoylated and contains a basic patch (**Figure 1*c***); together these features direct and anchor the Gag protein to the plasma membrane. The capsid (CA) domain participates in a series of protein:protein interactions that first define the shape of the immature particle and then define the shape of the mature capsid. The nucleocapsid (NC) domain (**Figure 1*d***) is responsible for recruiting the viral genome and helps drive immature particle assembly through the simultaneous interaction of multiple Gag proteins with shared RNA molecules. Finally, the C-terminal p6 domain is responsible for the recruitment of the ESCRT machinery, which enables the scission of the lipid envelope of the viral particle from the infected cell plasma membrane.

Imaging studies of viral Gag and RNA molecules, as well as biochemical, pulse-chase, and cross-linking studies, have revealed much about assembly dynamics and suggest an order of events in the construction of immature HIV-1 particles (**Figure 1*b***). Upon its synthesis, Gag is detected in the cytosol and a large fraction is freely diffusible, in a monomeric or low-order multimeric form (1–3). Similarly, newly synthesized viral RNA is intrinsically mobile in the cytoplasm (4). The cytoplasmic Gag pool is relatively long lived, as pulse-chase-labeling experiments indicate that a period of hours elapses between the time a typical Gag molecule is translated and the time at which it is found in assembled, extracellular virions (5, 6). To enable assembly, Gag proteins move from the cytosolic pool, most likely via diffusion and cytoplasmic mixing, to the plasma membrane (7, 8). Some viral RNA molecules are bound to Gag monomers or low-order multimers and consequently are also recruited to the plasma membrane (4, 9, 10). This nascent, membrane-bound Gag-RNA complex, which initially contains a few molecules of Gag and two molecules of viral RNA, becomes immobile at the plasma membrane as more Gag molecules are recruited, progressively building a spherical immature virion (4, 12) (**Figure 1*b***). Once virion assembly has been initiated, individual particles are assembled over a time course of minutes (8, 11), and upon completion, the ESCRT machinery, recruited by the p6 Gag domain, mediates membrane scission and release of the immature virion (13–15).

During and after the separation of the viral particle from the plasma membrane, the viral protease is activated and begins to catalyze the cleavage of the Gag precursor such that its component MA, CA, NC, and p6 domains become separated (16). A dramatic change in particle morphology ensues, in which interaction interfaces between CA subunits are dramatically altered to form the characteristic conical core (16). Simultaneously, the viral RNA condenses into a ribonucleoprotein (RNP) complex incorporating at least some of the liberated NC protein as well as viral enzymes required to initiate a new round of viral replication (16) (**Figure 1*b***).

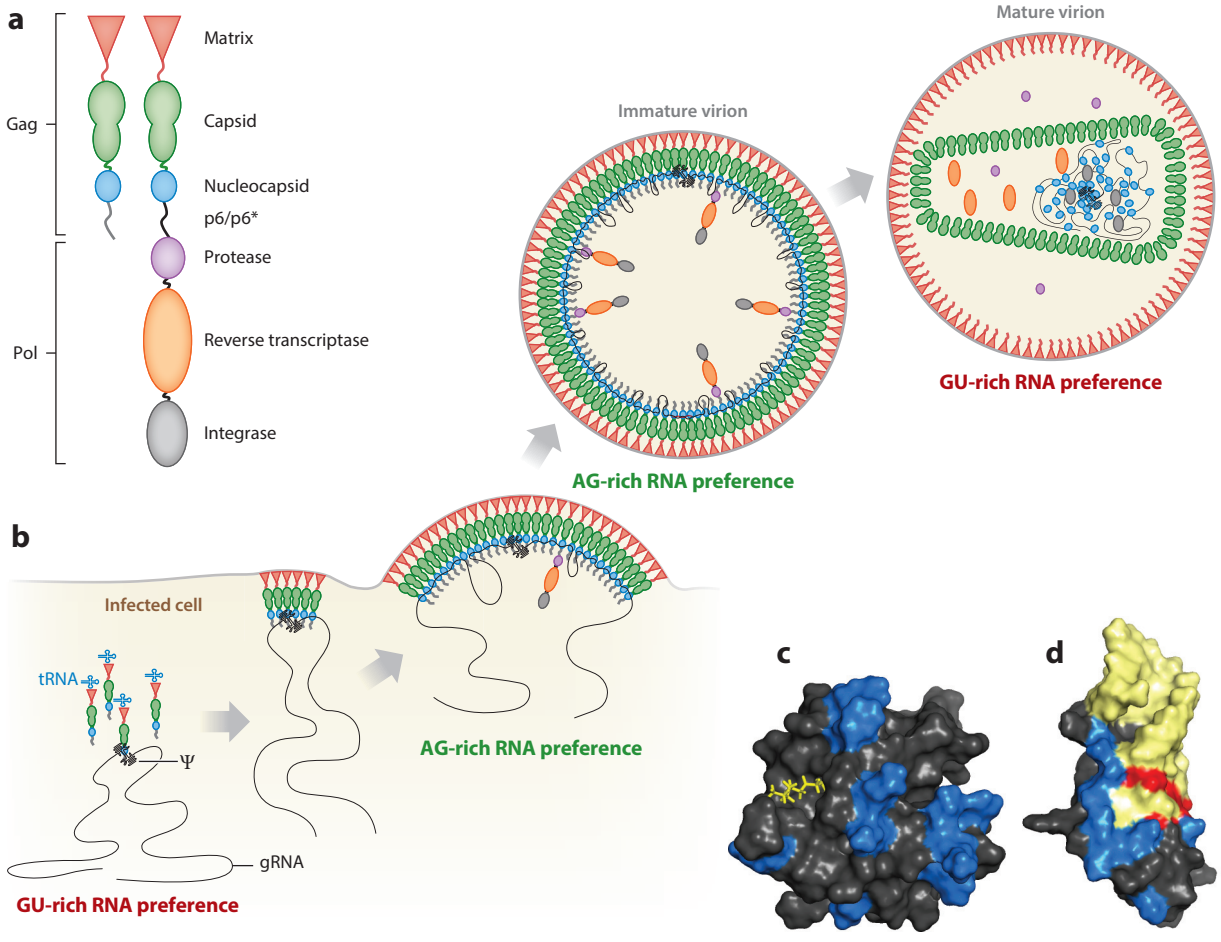

Figure 1

Switchable RNA binding specificities in the nucleocapsid (NC) and matrix (MA) domains regulate assembly and genome packaging. (*a*) Schematic representation of the HIV-1 Gag and GagPol precursor proteins. (*b*) Overview of HIV-1 particle assembly: Cytosolic Gag in infected cells preferentially binds to tRNA, via the MA domain, and to the packaging signal ψ and GU-rich mRNA elements via the NC domain. Membrane binding occludes tRNA binding by MA. Gag multimerization at the plasma membrane and in immature virions is associated with a switch in NC:RNA binding specificity to AG-rich elements whose composition matches that of the HIV-1 genome. Specificity reverts to GU-rich and ψ following proteolytic cleavage, which may be activated in virions by RNA interactions with protease. Correct positioning of the genomic RNA (gRNA):RNA ribonucleoprotein inside the conical capsid appears to require interaction between integrase and gRNA. (*c*, *d*) The major RNA binding domains in Gag: (*c*) nuclear magnetic resonance (NMR) structure of myristoylated MA (PDB ID 1A1T) viewed with the N-terminal–to–C-terminal axis perpendicular to the page surface. Blue denotes basic amino acids (K, R); yellow denotes N-terminal myristate, partly concealed in a hydrophobic pocket. (*d*) NMR structure of the HIV-1 Gag NC domain in complex with the SL3 RNA element (see **Figure 2**) (PDB ID 1A1T). Blue denotes basic amino acids (K, R), yellow denotes RNA, and red denotes unpaired G nucleotides.

2. SELECTION OF VIRAL RNAs FOR PACKAGING—THE CENTRAL ROLE OF THE PACKAGING SIGNAL, Ψ

HIV-1 particle assembly is initiated when the NC domain of Gag binds with high specificity to a single dimer of viral genomic RNA (gRNA). NC interacts with nucleic acids in multiple ways: electrostatically via its basic residues with positively charged nucleic acids, and more specifically

through the NC zinc fingers' hydrophobic binding pockets that interact with exposed unpaired guanosine residues (**Figure 1*d***). The high-affinity interactions drive gRNA packaging specificity, and zinc finger differences among retroviruses are a major determinant of viral species-specific packaging (17). NC can also act as a nucleic acid chaperone that remodels RNA structures and associations during reverse transcription and other replication processes (18). By reducing electrostatic repulsion between nucleic acid strands, NC likely contributes to RNA condensation during assembly (19) (**Figure 1*b***). NC coating also may contribute to RNA integrity, as the RNA in protease minus immature virions appears more labile than that in mature virions, and RNA integrity is decreased in virus-like particles (VLPs) lacking authentic NC (20, 21).

All HIV-1 RNAs are 5′ capped and 3′ polyadenylated and are in many ways indistinguishable from the host mRNAs. HIV-1 primary transcripts have three distinct fates: They can remain unspliced and serve (*a*) as *gag* mRNAs or (*b*) as gRNAs, or they can become spliced and serve (*c*) as subgenomic mRNAs for other viral proteins. As a result, HIV-1 gRNAs are a subset of the Pol II products transcribed from a single haploid locus. Although the HIV-1 promoter can be highly active, HIV-1 RNA levels never approach the high intracellular concentrations that characterize the replication cycles of many other RNA viruses. Thus, HIV-1 packaging presents a challenge in molecular recognition, wherein gRNA must be selected from an intracellular pool that contains a vast excess of chemically similar RNAs. Nonetheless, a large majority of the virions released from HIV-1-infected cells contain precisely one dimer of gRNA (22). One potential explanation for this selectivity is *cis*-preference, a process by which newly synthesized RNA binding proteins interact with the RNAs from which they are translated. *cis*-preference may facilitate genome selection for some retroviruses and retroelements and has been suggested to confer a replication advantage to L1 retroelements by enabling proteins to associate preferentially with replication-competent gRNAs (23, 24). In concept, *cis*-packaging could explain why most HIV-1 virions contain gRNA. However, comparison of the levels of two coexpressed packageable HIV-1 RNAs—one that expresses virion proteins and another that does not—to their levels in virions has refuted a *cis*-preference model for HIV-1 RNA packaging (25). Indeed, *trans*-packaging upon coexpression of packaging-defective protein-expressing genomes with packageable ("vector") RNAs forms the basis of lentivirus vector systems, commonly used for gene transfer (26, 27).

What then is responsible for the exquisite selectivity in HIV-1 genome packaging? Early work addressed whether or not HIV-1 RNA packaging relied on strategies similar to those used by simpler retroviruses and confirmed that packaging requires interactions between *cis*-acting sequences near the 5′ end of gRNA and the NC domain of Gag (28–30). This mirrored what had been determined for gammaretroviruses such as murine leukemia virus (MLV), where the term Ψ (Psi, for packaging signal) was coined (31, 32).

Work with purified proteins and RNAs in reconstituted reactions in vitro provided evidence for viral species-specific interactions between NC and structured elements at the 5′ end of gRNA (17, 33). It is noteworthy, however, that while NC selectivity for gRNA occurs in the context of intact Gag, technical challenges limited early in vitro work to the use of truncated forms of Gag, whose structures and dynamics may differ from those of intact Gag. Progress toward an understanding of RNA interactions in the context of authentic Gag is a slow and ongoing process (34–36).

For simple retroviruses, electron micrographs had shown that encapsidated gRNAs were joined in pairs via dimer linkage structures near their 5′ ends (37, 38). Biochemical studies revealed that the linkage was noncovalent and that gRNAs were coated with a single species of protein—namely, NC. UV cross-linking studies demonstrated that while NC can bind many forms of nucleic acids in vitro, it displays a preference for U- and G-rich RNA sequences (39, 40). Subsequent work confirmed that the HIV-1 gRNA dimer organization and associations with NC were similar (41–43).

In principle, a reason that nearly every HIV-1 particle contains gRNA could be that some feature of the gRNA, e.g., Ψ, is intrinsically required for assembly. For some RNA viruses, virion assembly is nucleated by interactions between virus proteins and nucleic acids. Tobacco mosaic virus provides a classic example, in which viral particle assembly is initiated by interactions between viral coat protein discs and a specific portion of the genome called the origin of assembly sequence (44). However, several lines of evidence suggest such interactions are not required for HIV-1: Most prominently, morphologically normal particles can form in the absence of gRNA. Indeed, if packaging elements are removed from an HIV-1 genome but the open reading frames are retained, the resulting derivative yields VLPs. The majority of these VLPs are devoid of gRNA, but they are indistinguishable from native virions in appearance and rates of formation (8, 27). Although HIV-1 and gammaretrovirus VLP assembly proceeds normally without gRNA, Gag requires some form of oligonucleotide to assemble in vitro (45), and RNA:protein interactions of some sort appear to be required during replication. In the absence of gRNA, populations of virions incorporate increased levels of cellular mRNAs, and these and/or other encapsidated host noncoding RNAs provide interactions critical to assembly in the absence of gRNA (46, 47).

Even though VLPs can assemble normally in the absence of gRNA, one and only one dimer of gRNA is recruited to most particles during HIV-1 replication. Despite evidence that gRNA is dispensable for assembly, the notion that interactions between gRNA's dimer linkage and the NC domain of Gag nucleate assembly is attractive and is supported by some reports. Accelerated rates of virion release in the presence of gRNA have been reported under certain conditions, which may be masked using other experimental approaches (48, 49). Recent advances in understanding the oligomeric interactions of multidomain portions of immature Gag are furthering an as-yet-incomplete understanding of how interactions with gRNA might nucleate and actively scaffold Gag multimerization during assembly (35, 50, 51).

2.1. Complexity and Controversy Associated with HIV-1 5′ Leader Structure and Function

The HIV-1 gRNA ~400-base 5′ leader region, in which key determinants of HIV-1 packaging specificity reside, consists of the 5′ untranslated region plus initial *gag* sequences (**Figure 2*a***). Early work with MLV had demonstrated that the RNA between its splice donor and *gag* start codon was sufficient to confer packaging onto a heterologous RNA (31). Initial HIV-1 work focused on the corresponding genetic interval and concluded it was the principal packaging element for HIV-1 as well (28, 30). Thus the hairpin residing in this interval is sometimes referred to as Ψ (**Figure 2*a***). However, the HIV-1 5′ leader is organized differently and is genetically more complex than the corresponding region of MLV, and subsequent work showed neither the Ψ hairpin alone nor even the whole 5′ leader was sufficient to confer HIV-1 packaging onto a heterologous RNA.

One reason why identifying a discrete packaging signal has been challenging is that the HIV-1 5′ leader is densely packed with elements required for several replication functions (52). These include the TAR hairpin, which recruits transcriptional elongation machinery via the HIV-1 Tat RNA binding protein, and the poly(A) hairpin, which in its 3′-end repeated copy includes the site of 3′-end cleavage that precedes polyadenylation. The ~150-base RNA segment that follows the poly(A) hairpin includes the U5 region and the primer tRNA binding site (pbs) and is followed by four short predicted hairpins important in packaging. These are designated SL1, SL2, SL3, and SL4—or, alternately the DIS, SD, Ψ, and AUG hairpins, respectively (**Figure 2*a***). DIS refers to the dimer initiation sequence, which includes a 6-base GC-rich palindrome; SD refers to the major splicing donor; and AUG represents the *gag* start codon.

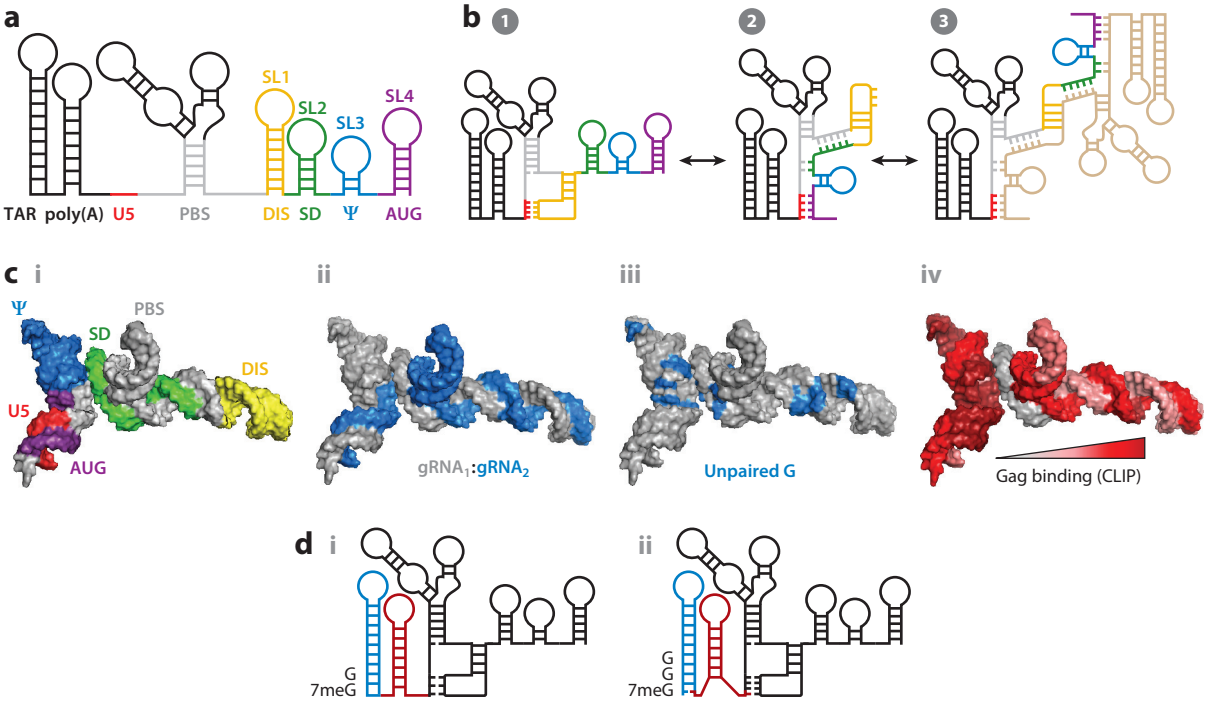

Figure 2

Switchable secondary and tertiary structures in the HIV-1 5′ leader regulate genomic RNA (gRNA) packaging. (*a*) Schematic representation of RNA secondary structure elements in the HIV-1 5′ leader. RNA sequences omitted from the core encapsidation signal (ψ^{CES}) nuclear magnetic resonance (NMR) structures are black; other elements are color coded consistently. (*b*) Switchable secondary structure of ψ^{CES}. In conformer ❶, the palindromic sequence in the DIS/SL1 loop is base-paired with U5 and thus is not available to initiate dimerization; in conformer ❷, the DIS palindrome is exposed and accessible to form "kissing" interactions with a second RNA that initiate dimer ❸ formation. (*c*) NMR structure of ψ^{CES} with various features highlighted: (*i*) secondary structural elements (colored as in panels *a* and *b*); (*ii*) contributions from different, dimerized gRNA molecules; (*iii*) unpaired G nucleotides important for packaging; and (*iv*) frequency of Gag binding in the cytosol of infected cells, as determined by CLIP assays. (*d*) Effect of alternate transcription start sites on 5′ leader RNA structures. (*i*) Two G:C and 7meG:C base pairs are present at the base of the TAR hairpin (*blue*). Few base pairs stabilize the interaction between the DIS hairpin and U5, enabling RNA isomerization to the dimer-competent conformer (panel *b*). (*ii*) Alternatively, an additional 5′ G base-pairs with a C at the base of the poly(A) hairpin, liberating a G, which extends and strengthens base-pairing between U5 and DIS, favoring an RNA conformation resistant to dimerization.

Genetic analyses of packaging determinants suggest that together, the 5′ leader and sequences that mediate RNA nuclear export are sufficient to enable packaging (53). These studies also support the notion that HIV-1 RNAs are constitutively packaged as dimers, as dimerization and packaging were difficult, if not impossible, to genetically separate (54). The initial interaction between two RNA dimerization partners consists of loop-loop "kissing" base-pairing interactions between DIS palindromes, followed by more extended duplex formation enabled by complementarity in the DIS hairpin stems (**Figure 2*b***). Evidence that it is the recognition of a dimer linkage and not RNA number or mass that dictates packaging includes findings that monomeric RNAs are packaged when packaging signals are placed in tandem (55), and that for both gammaretroviruses and HIV-1, RNA length does not alter the stoichiometry of encapsidation (32, 56). The presumption, supported by structural evidence, is that high-affinity binding sites for NC are exposed and available when RNAs adopt a dimer-competent fold but are masked when RNAs adopt alternate folds such as those that promote translation (57, 58). Formation of a dimer of gRNAs thus increases the

diversity of potential RNA structures beyond what would be possible with a monomeric genome and provides additional advantages during replication, such as a backup copy of genetic information to serve in genetic recombination and genome patch repair (59, 60).

Throughout 30 years of work defining the RNA elements that specify HIV-1 packaging, dozens of structural models for portions of the 5′ leader have been proposed (61). In the absence of three-dimensional data, RNA structures were assigned by using chemical probing or nuclease sensitivity, which could differentiate paired from unpaired residues but did not directly address pairing partners. These results could be mapped onto predicted secondary structures to determine which was most consistent with the data, but interpretation can require a subjective component and may lead to incorrect inferences. The structural clarity provided by recent nuclear magnetic resonance (NMR) studies (see below) required better definition of the minimal genetic elements required. However, defining minimal genetic elements for packaging has been difficult, due in part to the multiple structured elements in the 5′ leader (**Figure 2*a***). Roles in packaging for secondary structures and sequence motifs throughout the first ~500 residues of HIV-1 RNA have been suggested. Nearly every portion of the 5′ leader has been implicated in packaging in one study or another, including the TAR region (62, 63), the poly(A) hairpin (64), portions of the pbs region (65), and *gag* coding sequences (66, 67), in addition to the SL1–4 region (68). Some of these conflicting reports likely reflect how disrupting one strand of a base-paired duplex can disrupt adjacent RNA folds. For example, the entire TAR hairpin can be removed without affecting structures that regulate RNA fates, but smaller deletions or other alterations that destabilize the TAR hairpin can interfere with the proper folding of downstream elements (69, 70).

Adding to challenges in defining Ψ sequences is the apparent promiscuity in HIV-1 RNA packaging. When normalized for abundance within cells, HIV-1 gRNA is packaged with several orders of magnitude greater selectivity than any cellular mRNA, yet under some conditions, deletion of the entire region encompassing SL1–4 elements only marginally reduces HIV-1 RNA packaging (71). Nevertheless, when provided with pairwise combinations of candidate gRNAs, HIV-1 particles show pronounced selectivity for authentic gRNA (25). As a result, packaging competition approaches have proven invaluable in defining minimal 5′ leader elements sufficient to compete with authentic gRNAs.

2.2. Structure of the HIV-1 Core Encapsidation Signal, Ψ^{CES}

Recent landmark advances in NMR-based RNA structure determination have transformed our understanding of the HIV-1 packaging signal and its recognition (72, 73). These studies involved novel NMR approaches and integrated information gleaned from large sets of partially deuterated, segmentally labeled, or otherwise modified versions of a 155-base NL4–3 strain–derived core encapsidation signal (Ψ^{CES}) that lacks TAR, the poly(A) hairpin, and much of the pbs-containing interval and was designed based on competitive packaging studies (25, 58). The results illuminate the structural bases of HIV-1 RNA fate determination and RNA recognition for packaging (**Figure 2*b*,*c***).

The Ψ^{CES} sequence can adopt two distinct conformations, each with distinct secondary structure elements (57). Competitive packaging properties of mutants designed to stabilize one fold or the other have validated the structures, which are different from, but highly reminiscent of, previously predicted alternate 5′ leader structures (74). In one of these two NMR structures, the DIS loop is exposed and available for base-pairing with the DIS on a secondary RNA, while in the second structure, the DIS loop is sequestered by intramolecular base-pairing to upstream sequences on the same RNA, thus preventing intermolecular dimerization (**Figure 2*b***) (57). These two structures can be resolved from one another by agarose gel electrophoresis. When Ψ^{CES}

sequences are incubated at high concentrations in physiologic-like salts, the monomer RNA undergoes a structural switch to adopt the dimer-competent fold over time, until an equilibrium is reached. Although this structural switch occurs independently of any proteins in vitro, differences in mature and immature virions' dimer linkages suggest a role of proteins in dimer maturation during virus replication (20).

A subsequent report described a three-dimensional structure of the entire packaging-competent Ψ^{CES}, with surprising features that helped explain RNA packaging specificity and other previous enigmas, such as why the SD hairpin was not detected in prior NMR studies (58) (**Figure 2*c***). This three-dimensional structure revealed that packaging signal RNA adopts a novel tandem three-way junction structure that includes many unpredicted noncanonical base interactions. Surprisingly, sequences flanking the splice donor, which chemical probing had consistently predicted to be paired and which had been assumed to constitute a stem-loop structure (the SD/SL2 hairpin), instead participate in long-range base-pairing interactions at the heart of the structure (**Figure 2*c***). The splice donor itself is sequestered in base-pairing interactions that would preclude recruitment of the splicing machinery, thus providing a structural explanation for why HIV-1 RNA packaging and splicing are, for the most part, mutually exclusive. Subsequent structural work described the extended dimer structure and revealed its extensive intermolecular interface (**Figure 2*c*,*b***) (73). A high-resolution structure for the monomeric RNA is yet to be defined.

The packaging signal structures' topology exposes small clusters of unpaired guanosine residues near the bases of each of the three-way junctions. Consistent with a role for these guanosines in NC domain interactions, mutating these junctional residues rendered the RNA unable to compete with wild-type RNAs for packaging. RNA crosslinking-immunoprecipitation (CLIP) studies have confirmed that these same residues in Ψ^{CES} likely constitute the primary site of Gag interaction with gRNA in the cytoplasm of infected cells (75) (**Figure 2*c***).

2.3. Switchable 5′ Leader RNA Structures Regulated by the Transcription Start Site Influence Fates of HIV-1 RNA

The packaging signal structures described above provide information consistent with several known biological properties of the HIV-1 5′ leader. However, because they were derived using truncated leader RNAs, it remained important to address whether or not the defined structural elements were retained when rebuilt into their authentic context. Unexpectedly, a structural switch between dimerization-competent and monomeric leaders was found to be triggered by the presence or absence of a single extra guanosine or 7-methyl guanosine (7meG) cap residue on the RNA's 5′ end, even though these very minor changes were more than 100 bases upstream of Ψ^{CES} (76) (**Figure 2*d***).

HIV-1 transcription initiates within a run of three guanosine residues in the proviral promoter, but the precise residue used as the transcription start site had alternately been reported as the first, second, or third of these Gs or as a combination of more than one (77–80). By comparison to a series of synthetic RNA standards, it was determined that infected cells contained a mixed population of RNAs resulting from the use of the first and third Gs, generating RNAs with either one or three 5′ guanosines. Remarkably, and consistent with dimerization properties of the purified RNAs, virions contained exclusively one-guanosine RNAs, while the two-base-longer RNA was found to be enriched on polysomes.

What caused this functional separation of RNAs? The dimerization properties of RNAs with mutations designed to stabilize alternate folds supported a model consistent with earlier reports that showed destabilizing TAR reduces packaging (63, 69). According to this model, the nature of HIV-1 RNA's 5′ end initiates a chain of folding events that ultimately dictate whether or not

the DIS is exposed and available for dimer initiation (**Figure 2*d***). Together, these observations contribute to a growing understanding of how HIV-1 controls several highly regulated RNA functions using dynamic structural properties of its RNA's 5′ end.

3. SWITCHABLE GAG:RNA BINDING SPECIFICITY—A POSSIBLE ROLE FOR NUCLEOTIDE COMPOSITION IN GENOME PACKAGING

Even though Ψ plays a key initiating role in the recognition and recruitment of the viral genome by Gag, several findings suggest that Ψ is not the sole feature of the viral genome that drives packaging. In considering the potential role played by RNA sequences other than Ψ in viral genome packaging, two facts are especially pertinent. First, electron microscopic images of immature particles present an electron-dense layer coincident with the interior of the immature Gag shell (16). This electron-dense layer likely represents RNA bound to the pseudo-two-dimensional arrayed NC domains that line the interior of the Gag shell (**Figure 1*b***). Consistent with this idea, "tomo-bubblegrams" indicate the presence of radiosensitive material at this location in immature virions (81). The total RNA binding capacity of the NC domains contained within this layer of the immature virion vastly exceeds that which could be occupied by the dimeric Ψ element in each virion. It follows therefore that the vast majority of the Gag molecules in an immature virion particle must be bound to viral RNA sequences other than Ψ (**Figure 1*b***). A key question, then, is whether these additional interactions between Gag and viral RNA are purely incidental, and secondary to Gag:Ψ binding, or whether they play a facilitating role in viral genome packaging.

CLIP studies show that the specificity with which Gag binds to viral RNA as opposed to cellular RNA, measured in the cytoplasm of an infected cell, is clearly insufficient to explain the remarkable specificity with which viral genomes are packaged into virions (75). Specifically, viral RNA represents <1% of the total mRNA in an infected cell, yet viral RNA represents ∼60% to 75% of the RNA in virion particles (47, 75). However, CLIP experiments indicate that in cells Gag is bound to viral RNA at a frequency that is only a few-fold greater than the frequency with which it is bound to cellular mRNAs (75).

Evidence that RNA sequences other than Ψ facilitate packaging derives from studies showing that perturbation of Ψ reduces, but does not abolish, specific incorporation of viral RNA (71, 75, 82, 83). Spliced viral RNAs, in which much of the Ψ element has been deleted, are also preferentially packaged into virions over cellular RNAs, particularly in the absence of a full-length viral RNA (84). Studies that have attempted to define minimal viral sequences that enable viral genome transduction in the context of HIV-1 vectors suggest that viral sequences outside the Ψ region facilitate genome packaging (66, 67, 69, 85–88); however, attempts to map non-Ψ sequences that are important for packaging have not reached concordant conclusions (66, 86–91). Overall these studies suggest that some feature(s) of the viral genome outside the Ψ element play an accessory role in packaging. Notably, in immature virions, Gag is bound to many places on the HIV-1 genome (**Figure 1*b***), and fluctuations in binding frequency across the genome are strikingly similar in the presence or absence of the Ψ element (75). Thus, the interactions of Gag with the non-Ψ portion of the viral RNA are at least partly independent of Ψ.

CLIP studies that analyzed a side reaction during viral genome packaging, namely the interaction between Gag and cellular mRNA, strongly suggest the RNA binding specificity of Gag changes during the assembly of immature virions (75). In particular, the RNA sequences bound by monomeric Gag in the cytoplasm reveal a preference for G-rich sequences and motifs. This finding is consistent with in vitro biochemical analyses carried out using an isolated NC domain, and is consistent with the notion that the NC zinc knuckles are predisposed to bind to unpaired G nucleotides (41, 92, 93) (**Figure 1*d***). In contrast, even though the cellular mRNAs that are

packaged into HIV-1 particles are largely representative of those present in the cytoplasm of infected cells, the sites on the RNA that are bound by Gag differ markedly in the cell cytoplasm versus immature virions (75). Most strikingly, the nucleotide composition of the preferred Gag binding sites on cellular mRNAs found in immature virions is very similar to the A-rich, C-poor nucleotide composition of the HIV-1 genome (~35% A, ~22.5% U, ~24.5% G, ~18% C). Moreover, sequence motifs that are enriched in Gag binding sites in immature virions are both G- and A-rich.

Notably, this apparent change in Gag:RNA binding specificity is transient and largely reverts upon virion maturation. In mature virions the pattern of NC binding on the viral RNA is similar, albeit not identical, to the G-rich specificity exhibited by the intact Gag protein in the cytoplasm of infected cells (75). The NC domain of Gag undergoes large changes in its configuration during assembly that could plausibly alter its RNA binding specificity. Upon multimerization at the plasma membrane, NC domains become highly constrained in a pseudo-two-dimensional array as Gag molecules are assembled into a hexameric lattice that characterizes the immature virion. Following budding and proteolytic cleavage of Gag, NC domains are liberated and are then, in principle, free to diffuse within the interior of the viral particle, constraints imposed by CA and the viral RNA notwithstanding. Thus, these CLIP studies suggest that the apparently reversible change to an A-rich/G-rich RNA binding specificity in Gag:NC RNA binding is governed by Gag multimerization, and perhaps by attendant changes in the exposure of NC amino acids that participate in RNA binding (**Figure 1*b*,*d***).

Despite the fact that HIV-1 sequences are highly variable, the biased nucleotide composition of its genome is remarkably conserved among primate lentiviruses (94), suggesting a functional constraint. The fact that Gag in immature virions selectively binds to cellular mRNA elements whose composition corresponds to that of the HIV-1 genome suggests the possibility that selective packaging of viral RNA may have provided the impetus for the HIV-1 genome to evolve to its current characteristic composition. However, it is unlikely that individual Gag molecules can "sense" the overall nucleotide composition of a large RNA. More likely is that many individual Gag molecules within an assembling array preferentially bind discrete motifs that are overrepresented in sequences with HIV-1-like composition. Mononucleotide, dinucleotide, trinucleotide, or longer oligonucleotide sequences could plausibly serve as Gag recognition sites, perhaps with a hierarchy of affinities (e.g., GA > AA > UA > CA). This premise is, in effect, an extension of theoretical proposals for other RNA viruses in which numerous low-affinity, degenerate, scattered motifs are invoked to explain genome packaging (95–97). Thus, nucleotide composition could confer overrepresentation of preferred Gag binding motifs, or vice versa. Moreover, nucleotide composition in HIV-1 likely represents a compromise of the dual demands of binding to Gag during packaging while simultaneously encoding functional proteins. An alternative scenario is that some other selective pressure drove acquisition of the A-rich, C-poor genome character, and Gag then acquired a multimerization-driven specificity for A-rich RNA to accentuate packaging. Sources of evolutionary pressure that may have favored an A-rich, C-poor genome include the APOBEC3 proteins (98). These cytidine deaminase proteins catalyze minus-strand dC-to-dU mutations that are transcribed to generate plus-strand G-to-A mutations. Another antiviral protein, zinc finger antiviral protein, recognizes CG dinucleotides, causes cytoplasmic depletion of CG-rich RNA, and may have driven the removal of C nucleotides from the HIV-1 genome (99). An A-rich, C-poor composition may also have been selected because of potential advantages associated with maintenance of a single-stranded character (100) in the context of a virus whose NC protein preferentially recognizes unpaired G nucleotides (41, 92, 93). Thus, it is conceivable that the A-rich, G-rich motifs preferred by Gag in immature virions reflect preferential binding to G nucleotides in single-stranded regions of RNA.

4. POTENTIAL ROLES OF A SWITCHABLE MA:RNA INTERACTION IN THE SPATIOTEMPORAL REGULATION OF HIV-1 ASSEMBLY

A key function of the MA domain of Gag is to specify the timing and location at which HIV-1 assembly occurs in infected cells. MA contains a bipartite membrane binding signal, consisting of an N-terminal myristoyl modification and clustered basic amino acids on one face of the MA globular head (101) (**Figure 1*c***). MA:membrane binding is somewhat promiscuous in vitro with respect to the preferred lipid composition of target membranes (102–104), but the targeting of Gag and virion assembly in living cells is highly specific. Indeed, while the plasma membrane constitutes only a few percent of the total membrane content of the cell, HIV-1 particle assembly occurs nearly exclusively at this site (7). A particular plasma membrane–specific phospholipid, namely phosphatidylinositol 4,5-bisphosphate [$PI(4,5)P_2$], is crucial for this specific targeting, as overexpression of a $PI(4,5)P_2$-depleting enzyme reduces extracellular virion yield and redirects assembly to intracellular organelles (105). Presumably, $PI(4,5)P_2$ must bind to MA with higher affinity than do other acidic phospholipids that are present in other cellular membranes.

It has long been known that, in addition to binding negatively charged membrane surfaces, MA can bind to RNA in vitro (102, 106–109). The N-terminal basic amino acids in MA that are responsible for directing Gag to the plasma membrane are also required for in vitro RNA binding (101, 102, 110). For many years, the RNA binding activity of MA was not thought to be sequence specific. However, recent CLIP studies have shown that the MA domain binds nearly exclusively to tRNAs in infected cells (75). In particular, a subset of ~10 tRNAs are preferred binding targets for HIV-1 MA, including $tRNA^{Glu}$, $tRNA^{Gly}$, $tRNA^{Lys}$, and $tRNA^{Val}$. The molecular basis for this apparent specificity has not been determined, but the identity of the cross-linked nucleotides in the targeted tRNAs suggests that the dihydrouridine loop contacts MA (75). Notably, MA:tRNA interactions constitute the most frequent interaction between Gag molecules and RNA in the cytosol of infected cells (**Figure 1*b***).

What are the potential functional consequences of MA:RNA binding? One model for HIV-1 assembly, based on analyses of Gag monomer conformations in solution (111), invokes an intermediate conformational step in which Gag molecules are folded into a compact structure with NC and MA domains positioned proximally to each other. In such a conformation, the MA and NC RNA binding domains within an individual monomer could be bridged by viral RNA, a tRNA, or even the viral RNA with an annealed $tRNA^{LysUUU}$ primer. However, CLIP assays do not lend support to these models—there is little to no binding detected between MA and viral RNA, while tRNA constitutes a small fraction of the RNA bound by NC. Moreover, although one of the ~10 preferred tRNAs bound by MA is $tRNA^{LysUUU}$, the observed MA:tRNA binding occurs independently of NC and the pbs. Therefore, MA:tRNA binding cannot be confined to interactions with the tRNA primer that is annealed to the viral genome (75).

If some level of nonspecific interaction with RNA is an inescapable consequence of possessing a basic patch (whose primary function is to target the plasma membrane), then the presence of two RNA binding domains, in the context of a single Gag protein that also has an intrinsic ability to multimerize, might predispose Gag to aggregate in the context of a cell cytoplasm. Such an event would likely inhibit virion assembly. Thus, the acquisition of a specific, monovalent interaction with a small RNA molecule (e.g., tRNA) might serve as a mechanism to buffer the basic surface on the MA globular head, preventing inappropriate interaction with larger RNA molecules that might trigger aggregation. Moreover, tRNA could inhibit inappropriate interaction between MA and intracellular membranes. Indeed, in vitro studies indicate that RNA can indeed inhibit the interaction between Gag and authentic cellular membranes, as well as between Gag and model membranes reconstituted as liposomes in vitro (75, 102, 103, 110). In the latter context, RNA is better able to block MA binding to membranes that are devoid of $PI(4,5)P_2$. Therefore,

tRNA:MA interactions could increase the specificity with which Gag molecules are targeted to the plasma membrane. Such a scenario could potentially be realized through a hierarchy of affinities, whereby intermediate-affinity MA:tRNA binding could successfully compete with lower-affinity intracellular membrane:MA binding, while higher-affinity $PI(4,5)P_2$ at the plasma membranes might successfully compete with tRNA.

An alternative role for MA:tRNA binding might be to temporally, rather than spatially, regulate Gag membrane binding. In addition to inhibition by RNA, Gag:membrane binding is regulated by a myristoyl switch mechanism, whereby myristate is concealed in a hydrophobic pocket when Gag is monomeric (**Figure 1*b***), and multimerization promotes myristate exposure and membrane binding (112, 113). Both tRNA-mediated occlusion of the basic amino acids (75, 102) and the sequestration of myristate in the hydrophobic pocket (112–114) contribute to the apparent autoinhibition of membrane binding that is mediated by the MA globular head. While N-terminal myristoylation is clearly not required for tRNA binding (75), it is possible that tRNA binding and myristate exposure might be reciprocally regulated. In any case, removal of negative regulatory mechanisms through mutations in MA or, most clearly, deletion of the MA globular head promotes Gag membrane binding and accelerates virion release (115, 116). This feature is most prominent at the low Gag concentrations encountered in the cytoplasm of infected cells shortly after the onset of its expression, when Gag is largely monomeric (117, 118). As Gag concentrations progressively increase during the late stages of the replication cycle, a greater fraction of the protein is driven into multimers, leading to myristate exposure, membrane binding, and virion assembly. Thus, occlusion of membrane binding features has the effect of delaying the accumulation of Gag in infected cell membranes, lengthening the HIV-1 replication cycle by several hours (119). It is counterintuitive that this feature would be naturally selected, unless there was some benefit to HIV-1 replication that is conferred by delayed virion assembly. The fact that HIV-1 target cells express antiviral proteins that can reduce the yield and infectiousness of virions from infected cells may provide an impetus for this apparently selected property. For example, proteins that inhibit the release of HIV-1 (Tetherin) or infiltrate and poison HIV-1 particles (APOBEC3 and SERINC3/5) would clearly curtail HIV-1 replication were they not removed from their site of action in infected cells (98, 120–122). HIV-1 employs antagonists of antiviral proteins (Nef, Vif, and Vpu) for this purpose. These viral accessory proteins deplete or relocalize antiviral proteins, reducing their effectiveness. A consequence of the MA-imposed delay in HIV-1 assembly is to postpone virion assembly to a time at which APOBEC3G and likely other antiviral proteins have been removed from their sites of action (119). Without the removal of these antiviral proteins, the assembly of new HIV-1 particles would be at best futile, and might even reduce the overall yield of virions from a given infected cell.

5. RNA INTERACTIONS WITH ENZYMES DURING VIRION MATURATION

In addition to the interactions with the Gag protein, recent findings have suggested RNA interaction with viral enzymes may regulate virion maturation. The HIV-1 protease (PR) that catalyzes the proteolytic cleavages to enable maturation is a homodimer, and both subunits are essential for the formation of the active site. Various sites in the Gag and Pol proteins (**Figure 1*a*,*b***) are cleaved at different rates, determined at least in part by amino acid sequences immediately surrounding the cleaved bond, as well as by the broader protein context, which presumably influences cleavage site accessibility. Initially, it was found that the cleavage of certain intermediate substrates in vitro that contained the NC domain was accelerated by including RNA or DNA in the reaction (123–125). This could be explained by a model in which NC:RNA interactions facilitate access to the cleaved

peptide bond. Conversely, more recent findings have shown that RNA stimulates PR activity irrespective of the substrate, and even short peptides are cleaved more efficiently in the presence of RNA (126). While RNA does not appear to promote PR dimerization, longer RNAs appear more effective than oligonucleotides at stimulating PR activity. Overall these findings are best explained by a model in which RNA binds to an allosteric site (or sites) in the dimeric PR enzyme. As such, the packaged RNA may be a key regulator of PR activity and promote maturation.

Another RNA:protein interaction that appears critical for accurate virion morphogenesis involves the viral integrase (IN) protein (**Figure 1*b***). A number of studies of so-called class II, or pleiotropic, IN mutants, and other studies with allosteric integrase inhibitors (ALLINIs), have shown that perturbations of IN can disrupt virion morphogenesis (81, 127–129). Virions assembled under these conditions are, for the most part, morphologically accurate, in that the NC protein is bound to the viral RNA and the CA lattice is correctly assembled (130). Strikingly, however, the spatial organization of the virion is often aberrant, with the NC:RNA RNP positioned outside the capsid (81, 128, 129). The ability of IN to mediate correct positioning of the RNP within mature virions appears to be linked to its ability to bind to RNA (131). CLIP analyses in mature virions show clearly that IN can be cross-linked at many positions to the viral RNA. The IN footprint on the viral genome differs significantly from that observed with the NC protein, with IN tending to favor regions that are double stranded (e.g., the Tat response element TAR and RRE). A variety of in vitro analyses support the conclusions that IN:RNA interactions might facilitate proper maturation. For example, IN tetramers can bridge RNA molecules, forming large but discrete condensed complexes that are visible by atomic force microscopy (131). Additionally, mass spectrometry–based footprinting shows that lysine residues in the IN C-terminal domain are specifically occluded when IN is bound by RNA. Mutation of these residues abolishes RNA binding in virions and recapitulates the characteristic morphogenesis defect that is induced by ALLINIs. Unlike other class II mutants, an RNA binding IN mutant is fully capable of carrying out concerted integration in vitro, suggesting that the aberrant morphogenesis is not due to pleiotropic effects but is caused by a loss of IN:RNA binding.

6. CONCLUDING REMARKS

A set of protein:RNA interactions, some of which change dramatically during particle morphogenesis, appear to play key roles in the assembly and maturation of HIV-1 particles. These interactions vary widely in their character, ranging from specific recognition of the highly structured Ψ element to less specific interactions with the viral RNA that are, or may be, influenced by the unusual nucleotide composition of the HIV-1 genome. These RNA:protein interactions appear to have an extraordinary diversity of roles, including RNA packaging, spatial and/or temporal regulation of the assembly of immature and mature virion structures, and regulation of viral enzymes. The structural details of these RNA:protein interactions and the details of the molecular mechanisms by which they promote or regulate HIV-1 assembly and maturation are only just beginning to be understood.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

P.B. and A.T. acknowledge support from National Institutes of Health grant U54 GM103297.

LITERATURE CITED

1. Gomez CY, Hope TJ. 2006. Mobility of human immunodeficiency virus type 1 Pr55Gag in living cells. *J. Virol.* 80:8796–806
2. Kutluay SB, Bieniasz PD. 2010. Analysis of the initiating events in HIV-1 particle assembly and genome packaging. *PLOS Pathog.* 6:e1001200
3. Hendrix J, Baumgartel V, Schrimpf W, Ivanchenko S, Digman MA, et al. 2015. Live-cell observation of cytosolic HIV-1 assembly onset reveals RNA-interacting Gag oligomers. *J. Cell Biol.* 210:629–46
4. Jouvenet N, Simon SM, Bieniasz PD. 2009. Imaging the interaction of HIV-1 genomes and Gag during assembly of individual viral particles. *PNAS* 106:19114–19
5. Huang M, Orenstein JM, Martin MA, Freed EO. 1995. p6Gag is required for particle production from full-length human immunodeficiency virus type 1 molecular clones expressing protease. *J. Virol.* 69:6810–18
6. Schubert U, Ott DE, Chertova EN, Welker R, Tessmer U, et al. 2000. Proteasome inhibition interferes with Gag polyprotein processing, release, and maturation of HIV-1 and HIV-2. *PNAS* 97:13057–62
7. Jouvenet N, Neil SJ, Bess C, Johnson MC, Virgen CA, et al. 2006. Plasma membrane is the site of productive HIV-1 particle assembly. *PLOS Biol.* 4:e435
8. Jouvenet N, Bieniasz PD, Simon SM. 2008. Imaging the biogenesis of individual HIV-1 virions in live cells. *Nature* 454:236–40
9. Chen J, Grunwald D, Sardo L, Galli A, Plisov S, et al. 2014. Cytoplasmic HIV-1 RNA is mainly transported by diffusion in the presence or absence of Gag protein. *PNAS* 111:E5205–13
10. Sardo L, Hatch SC, Chen J, Nikolaitchik O, Burdick RC, et al. 2015. Dynamics of HIV-1 RNA near the plasma membrane during virus assembly. *J. Virol.* 89:10832–40
11. Ivanchenko S, Godinez WJ, Lampe M, Krausslich HG, Eils R, et al. 2009. Dynamics of HIV-1 assembly and release. *PLOS Pathog.* 5:e1000652
12. Chen J, Rahman SA, Nikolaitchik OA, Grunwald D, Sardo L, et al. 2016. HIV-1 RNA genome dimerizes on the plasma membrane in the presence of Gag protein. *PNAS* 113:E201–8
13. Garrus JE, von Schwedler UK, Pornillos OW, Morham SG, Zavitz KH, et al. 2001. Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell* 107:55–65
14. Jouvenet N, Zhadina M, Bieniasz PD, Simon SM. 2011. Dynamics of ESCRT protein recruitment during retroviral assembly. *Nat. Cell Biol.* 13:394–401
15. Baumgartel V, Ivanchenko S, Dupont A, Sergeev M, Wiseman PW, et al. 2011. Live-cell visualization of dynamics of HIV budding site interactions with an ESCRT component. *Nat. Cell Biol.* 13:469–74
16. Sundquist WI, Krausslich HG. 2012. HIV-1 assembly, budding, and maturation. *Cold Spring Harb. Perspect. Med.* 2:a006924
17. Berkowitz R, Fisher J, Goff SP. 1996. RNA packaging. *Morphog. Matur. Retrovir.* 214:177–218
18. Levin JG, Guo J, Rouzina I, Musier-Forsyth K. 2005. Nucleic acid chaperone activity of HIV-1 nucleocapsid protein: critical role in reverse transcription and molecular mechanism. *Prog. Nucleic Acid Res. Mol. Biol.* 80:217–86
19. Johnson SF, Telesnitsky A. 2010. Retroviral RNA dimerization and packaging: the what, how, when, where, and why. *PLOS Pathog.* 6:e1001007
20. Fu W, Gorelick RJ, Rein A. 1994. Characterization of human immunodeficiency virus type 1 dimeric RNA from wild-type and protease-defective virions. *J. Virol.* 68:5013–18
21. Keene SE, King SR, Telesnitsky A. 2010. 7SL RNA is retained in HIV-1 minimal virus-like particles as an S-domain fragment. *J. Virol.* 84:9070–77
22. Chen J, Nikolaitchik O, Singh J, Wright A, Bencsics CE, et al. 2009. High efficiency of HIV-1 genomic RNA packaging and heterozygote formation revealed by single virion analysis. *PNAS* 106:13535–40
23. Kulpa DA, Moran JV. 2006. *cis*-preferential LINE-1 reverse transcriptase activity in ribonucleoprotein particles. *Nat. Struct. Mol. Biol.* 13:655–60
24. Kaye JF, Lever AM. 1999. Human immunodeficiency virus types 1 and 2 differ in the predominant mechanism used for selection of genomic RNA for encapsidation. *J. Virol.* 73:3023–31
25. Heng X, Kharytonchyk S, Garcia EL, Lu K, Divakaruni SS, et al. 2012. Identification of a minimal region of the HIV-1 5′-leader required for RNA dimerization, NC binding, and packaging. *J. Mol. Biol.* 417:224–39

26. Naldini L, Blomer U, Gallay P, Ory D, Mulligan R, et al. 1996. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* 272:263–67
27. Muhlebach MD, Schule S, Gerlach N, Schweizer M, Buchholz CJ, et al. 2010. Gammaretroviral and lentiviral vectors for gene delivery. In *Retroviruses: Molecular Biology, Genomics and Pathogenesis*, ed. R Kurth, N Bannert, pp. 347–70. Poole, UK: Caister Acad.
28. Aldovini A, Young RA. 1990. Mutations of RNA and protein sequences involved in human immunodeficiency virus type 1 packaging result in production of noninfectious virus. *J. Virol.* 64:1920–26
29. Gorelick RJ, Nigida SM Jr., Bess JW Jr., Arthur LO, Henderson LE, Rein A. 1990. Noninfectious human immunodeficiency virus type 1 mutants deficient in genomic RNA. *J. Virol.* 64:3207–11
30. Lever A, Gottlinger H, Haseltine W, Sodroski J. 1989. Identification of a sequence required for efficient packaging of human immunodeficiency virus type 1 RNA into virions. *J. Virol.* 63:4085–87
31. Mann R, Mulligan RC, Baltimore D. 1983. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. *Cell* 33:153–59
32. Adam MA, Miller AD. 1988. Identification of a signal in a murine retrovirus that is sufficient for packaging of nonretroviral RNA into virions. *J. Virol.* 62:3802–6
33. Clever J, Sassetti C, Parslow TG. 1995. RNA secondary structure and binding sites for *gag* gene products in the 5′ packaging signal of human immunodeficiency virus type 1. *J. Virol.* 69:2101–9
34. Abd El-Wahab EW, Smyth RP, Mailler E, Bernacchi S, Vivet-Boudou V, et al. 2014. Specific recognition of the HIV-1 genomic RNA by the Gag precursor. *Nat. Commun.* 5:4304
35. Pak AJ, Grime JMA, Sengupta P, Chen AK, Durumeric AEP, et al. 2017. Immature HIV-1 lattice assembly dynamics are regulated by scaffolding from nucleic acid and the plasma membrane. *PNAS* 114:E10056–65
36. Carlson LA, Bai Y, Keane SC, Doudna JA, Hurley JH. 2016. Reconstitution of selective HIV-1 RNA packaging in vitro by membrane-bound Gag assemblies. *eLife* 5:e14663
37. Kung HJ, Bailey JM, Davidson N, Vogt PK, Nicolson MO, McAllister RM. 1974. Electron-microscope studies of tumor-virus RNA. *Cold Spring Harb. Symp. Quant. Biol.* 39:827–34
38. Murti KG, Bondurant M, Tereba A. 1981. Secondary structural features in the 70S RNAs of Moloney murine leukemia and Rous sarcoma viruses as observed by electron microscopy. *J. Virol.* 37:411–19
39. Darlix JL, de Rocquigny H, Mauffret O, Mely Y. 2014. Retrospective on the all-in-one retroviral nucleocapsid protein. *Virus Res.* 193:2–15
40. Darlix JL, Spahr PF. 1982. Binding sites of viral protein P19 onto Rous sarcoma virus RNA and possible controls of viral functions. *J. Mol. Biol.* 160:147–61
41. Vuilleumier C, Bombarda E, Morellet N, Gerard D, Roques BP, Mely Y. 1999. Nucleic acid sequence discrimination by the HIV-1 nucleocapsid protein NCp7: a fluorescence study. *Biochemistry* 38:16816–25
42. You JC, McHenry CS. 1993. HIV nucleocapsid protein. Expression in *Escherichia coli*, purification, and characterization. *J. Biol. Chem.* 268:16519–27
43. Rein A, Datta SA, Jones CP, Musier-Forsyth K. 2011. Diverse interactions of retroviral Gag proteins with RNAs. *Trends Biochem. Sci.* 36:373–80
44. Butler PJ. 1984. The current picture of the structure and assembly of tobacco mosaic virus. *J. Gen. Virol.* 65(Pt. 2):253–79
45. Gross I, Hohenberg H, Wilk T, Wiegers K, Grattinger M, et al. 2000. A conformational switch controlling HIV-1 morphogenesis. *EMBO J.* 19:103–13
46. Telesnitsky A, Wolin SL. 2016. The host RNAs in retroviral particles. *Viruses* 8:E235
47. Rulli SJ Jr., Hibbert CS, Mirro J, Pederson T, Biswal S, Rein A. 2007. Selective and nonselective packaging of cellular RNAs in retrovirus particles. *J. Virol.* 81:6623–31
48. Ueno T, Tokunaga K, Sawa H, Maeda M, Chiba J, et al. 2004. Nucleolin and the packaging signal, ψ, promote the budding of human immunodeficiency virus type-1 (HIV-1). *Microbiol. Immunol.* 48:111–18
49. Dilley KA, Nikolaitchik OA, Galli A, Burdick RC, Levine L, et al. 2017. Interactions between HIV-1 Gag and viral RNA genome enhance virion assembly. *J. Virol.* 91:e02319-16
50. Tanwar HS, Khoo KK, Garvey M, Waddington L, Leis A, et al. 2017. The thermodynamics of Pr55Gag-RNA interaction regulate the assembly of HIV. *PLOS Pathog.* 13:e1006221
51. Wagner JM, Zadrozny KK, Chrustowicz J, Purdy MD, Yeager M, et al. 2016. Crystal structure of an HIV assembly and maturation switch. *eLife* 5:e17063

52. Berkhout B. 1996. Structure and function of the human immunodeficiency virus leader RNA. *Prog. Nucleic Acid Res. Mol. Biol.* 54:1–34
53. Blissenbach M, Grewe B, Hoffmann B, Brandt S, Uberla K. 2010. Nuclear RNA export and packaging functions of HIV-1 Rev revisited. *J. Virol.* 84:6598–604
54. Russell RS, Liang C, Wainberg MA. 2004. Is HIV-1 RNA dimerization a prerequisite for packaging? Yes, no, probably? *Retrovirology* 1:23
55. Sakuragi J, Shioda T, Panganiban AT. 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.* 75:2557–65
56. Shin NH, Hartigan-O'Connor D, Pfeiffer JK, Telesnitsky A. 2000. Replication of lengthened Moloney murine leukemia virus genomes is impaired at multiple stages. *J. Virol.* 74:2694–702
57. Lu K, Heng X, Garyu L, Monti S, Garcia EL, et al. 2011. NMR detection of structures in the HIV-1 5′-leader RNA that regulate genome packaging. *Science* 334:242–45
58. Keane SC, Heng X, Lu K, Kharytonchyk S, Ramakrishnan V, et al. 2015. Structure of the HIV-1 RNA packaging signal. *Science* 348:917–21
59. Onafuwa-Nuga A, Telesnitsky A. 2009. The remarkable frequency of human immunodeficiency virus type 1 genetic recombination. *Microbiol. Mol. Biol. Rev.* 73:451–80
60. King SR, Duggal NK, Ndongmo CB, Pacut C, Telesnitsky A. 2008. Pseudodiploid genome organization AIDS full-length human immunodeficiency virus type 1 DNA synthesis. *J. Virol.* 82:2376–84
61. Lu K, Heng X, Summers MF. 2011. Structural determinants and mechanism of HIV-1 genome packaging. *J. Mol. Biol.* 410:609–33
62. Andersen ES, Contera SA, Knudsen B, Damgaard CK, Besenbacher F, Kjems J. 2004. Role of the trans-activation response element in dimerization of HIV-1 RNA. *J. Biol. Chem.* 279:22243–49
63. Helga-Maria C, Hammarskjold ML, Rekosh D. 1999. An intact TAR element and cytoplasmic localization are necessary for efficient packaging of human immunodeficiency virus type 1 genomic RNA. *J. Virol.* 73:4127–35
64. Russell RS, Hu J, Laughrea M, Wainberg MA, Liang C. 2002. Deficient dimerization of human immunodeficiency virus type 1 RNA caused by mutations of the u5 RNA sequences. *Virology* 303:152–63
65. Clever JL, Miranda D Jr., Parslow TG. 2002. RNA structure and packaging signals in the 5′ leader region of the human immunodeficiency virus type 1 genome. *J. Virol.* 76:12381–87
66. Liu Y, Nikolaitchik OA, Rahman SA, Chen J, Pathak VK, Hu WS. 2017. HIV-1 sequence necessary and sufficient to package non-viral RNAs into HIV-1 particles. *J. Mol. Biol.* 429:2542–55
67. McBride MS, Schwartz MD, Panganiban AT. 1997. Efficient encapsidation of human immunodeficiency virus type 1 vectors and further characterization of cis elements required for encapsidation. *J. Virol.* 71:4544–54
68. Lever AM. 2007. HIV-1 RNA packaging. *Adv. Pharmacol.* 55:1–32
69. Das AT, Klaver B, Klasens BI, van Wamel JL, Berkhout B. 1997. A conserved hairpin motif in the R-U5 region of the human immunodeficiency virus type 1 RNA genome is essential for replication. *J. Virol.* 71:2346–56
70. Vrolijk MM, Ooms M, Harwig A, Das AT, Berkhout B. 2008. Destabilization of the TAR hairpin affects the structure and function of the HIV-1 leader RNA. *Nucleic Acids Res.* 36:4352–63
71. Laham-Karam N, Bacharach E. 2007. Transduction of human immunodeficiency virus type 1 vectors lacking encapsidation and dimerization signals. *J. Virol.* 81:10687–98
72. Keane SC, Summers MF. 2016. NMR studies of the structure and function of the HIV-1 5′-leader. *Viruses* 8:338
73. Keane SC, Van V, Frank HM, Sciandra CA, McCowin S, et al. 2016. NMR detection of intermolecular interaction sites in the dimeric 5′-leader of the HIV-1 genome. *PNAS* 113:13033–38
74. Abbink TE, 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.* 278:11601–11
75. Kutluay SB, Zang T, Blanco-Melo D, Powell C, Jannain D, et al. 2014. Global changes in the RNA binding specificity of HIV-1 Gag regulate virion genesis. *Cell* 159:1096–109
76. Kharytonchyk S, Monti S, Smaldino PJ, Van V, Bolden NC, et al. 2016. Transcriptional start site heterogeneity modulates the structure and function of the HIV-1 genome. *PNAS* 113:13378–83

77. Muesing MA, Smith DH, Cabradilla CD, Benton CV, Lasky LA, Capon DJ. 1985. Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovirus. *Nature* 313:450–58
78. Starcich B, Ratner L, Josephs SF, Okamoto T, Gallo RC, Wong-Staal F. 1985. Characterization of long terminal repeat sequences of HTLV-III. *Science* 227:538–40
79. Menees TM, Muller B, Krausslich HG. 2007. The major 5′ end of HIV type 1 RNA corresponds to G456. *AIDS Res. Hum. Retrovir.* 23:1042–48
80. Masuda T, Sato Y, Huang YL, Koi S, Takahata 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
81. Fontana J, Jurado KA, Cheng N, Ly NL, Fuchs JR, et al. 2015. Distribution and redistribution of HIV-1 Nucleocapsid protein in immature, mature, and integrase-inhibited virions: a role for integrase in maturation. *J. Virol.* 89:9765–80
82. Clever JL, Parslow TG. 1997. Mutant human immunodeficiency virus type 1 genomes with defects in RNA dimerization or encapsidation. *J. Virol.* 71:3407–14
83. McBride MS, Panganiban AT. 1997. Position dependence of functional hairpins important for human immunodeficiency virus type 1 RNA encapsidation in vivo. *J. Virol.* 71:2050–58
84. Houzet L, Paillart JC, Smagulova F, Maurel S, Morichaud Z, et al. 2007. HIV controls the selective packaging of genomic, spliced viral and cellular RNAs into virions through different mechanisms. *Nucleic Acids Res.* 35:2695–704
85. Berkowitz RD, Hammarskjold ML, Helga-Maria C, Rekosh D, Goff SP. 1995. 5′ regions of HIV-1 RNAs are not sufficient for encapsidation: implications for the HIV-1 packaging signal. *Virology* 212:718–23
86. Chamanian M, Purzycka KJ, Wille PT, Ha JS, McDonald D, et al. 2013. A *cis*-acting element in retroviral genomic RNA links Gag-Pol ribosomal frameshifting to selective viral RNA encapsidation. *Cell Host Microbe* 13:181–92
87. Richardson JH, Child LA, Lever AM. 1993. Packaging of human immunodeficiency virus type 1 RNA requires *cis*-acting sequences outside the 5′ leader region. *J. Virol.* 67:3997–4005
88. Kaye JF, Richardson JH, Lever AM. 1995. *cis*-acting sequences involved in human immunodeficiency virus type 1 RNA packaging. *J. Virol.* 69:6588–92
89. Didierlaurent L, Racine PJ, Houzet L, Chamontin C, Berkhout B, Mougel M. 2011. Role of HIV-1 RNA and protein determinants for the selective packaging of spliced and unspliced viral RNA and host U6 and 7SL RNA in virus particles. *Nucleic Acids Res.* 39:8915–27
90. Grewe B, Ehrhardt K, Hoffmann B, Blissenbach M, Brandt S, Uberla K. 2012. The HIV-1 Rev protein enhances encapsidation of unspliced and spliced, RRE-containing lentiviral vector RNA. *PLOS ONE* 7:e48688
91. Nikolaitchik OA, Hu WS. 2014. Deciphering the role of the Gag-Pol ribosomal frameshift signal in HIV-1 RNA genome packaging. *J. Virol.* 88:4040–46
92. South TL, Summers MF. 1993. Zinc- and sequence-dependent binding to nucleic acids by the N-terminal zinc finger of the HIV-1 nucleocapsid protein: NMR structure of the complex with the Psi-site analog, dACGCC. *Protein Sci.* 2:3–19
93. De Guzman RN, Wu ZR, Stalling CC, Pappalardo L, Borer PN, Summers MF. 1998. Structure of the HIV-1 nucleocapsid protein bound to the SL3 Ψ-RNA recognition element. *Science* 279:384–88
94. van der Kuyl AC, Berkhout B. 2012. The biased nucleotide composition of the HIV genome: a constant factor in a highly variable virus. *Retrovirology* 9:92
95. Patel N, Dykeman EC, Coutts RH, Lomonossoff GP, Rowlands DJ, et al. 2015. Revealing the density of encoded functions in a viral RNA. *PNAS* 112:2227–32
96. Rolfsson O, Middleton S, Manfield IW, White SJ, Fan B, et al. 2016. Direct evidence for packaging signal-mediated assembly of bacteriophage MS2. *J. Mol. Biol.* 428:431–48
97. Stockley PG, White SJ, Dykeman E, Manfield I, Rolfsson O, et al. 2016. Bacteriophage MS2 genomic RNA encodes an assembly instruction manual for its capsid. *Bacteriophage* 6:e1157666
98. Sheehy AM, Gaddis NC, Choi JD, Malim MH. 2002. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* 418:646–50
99. Takata MA, Goncalves-Carneiro D, Zang TM, Soll SJ, York A, et al. 2017. CG dinucleotide suppression enables antiviral defence targeting non-self RNA. *Nature* 550:124–27

100. van Hemert F, van der Kuyl AC, Berkhout B. 2016. Impact of the biased nucleotide composition of viral RNA genomes on RNA structure and codon usage. *J. Gen. Virol.* 97:2608–19
101. Zhou W, Parent LJ, Wills JW, Resh MD. 1994. Identification of a membrane-binding domain within the amino-terminal region of human immunodeficiency virus type 1 Gag protein which interacts with acidic phospholipids. *J. Virol.* 68:2556–69
102. Chukkapalli V, Oh SJ, Ono A. 2010. Opposing mechanisms involving RNA and lipids regulate HIV-1 Gag membrane binding through the highly basic region of the matrix domain. *PNAS* 107:1600–5
103. Dick RA, Kamynina E, Vogt VM. 2013. Effect of multimerization on membrane association of Rous sarcoma virus and HIV-1 matrix domain proteins. *J. Virol.* 87:13598–608
104. Barros M, Heinrich F, Datta SA, Rein A, Karageorgos I, et al. 2016. Membrane binding of HIV-1 matrix protein: dependence on bilayer composition and protein lipidation. *J. Virol.* 90:4544–55
105. Ono A, Ablan SD, Lockett SJ, Nagashima K, Freed EO. 2004. Phosphatidylinositol (4,5) bisphosphate regulates HIV-1 Gag targeting to the plasma membrane. *PNAS* 101:14889–94
106. Cimarelli A, Luban J. 1999. Translation elongation factor 1-alpha interacts specifically with the human immunodeficiency virus type 1 Gag polyprotein. *J. Virol.* 73:5388–401
107. Ott DE, Coren LV, Gagliardi TD. 2005. Redundant roles for nucleocapsid and matrix RNA-binding sequences in human immunodeficiency virus type 1 assembly. *J. Virol.* 79:13839–47
108. Alfadhli A, Still A, Barklis E. 2009. Analysis of human immunodeficiency virus type 1 matrix binding to membranes and nucleic acids. *J. Virol.* 83:12196–203
109. Ramalingam D, Duclair S, Datta SA, Ellington A, Rein A, Prasad VR. 2011. RNA aptamers directed to human immunodeficiency virus type 1 Gag polyprotein bind to the matrix and nucleocapsid domains and inhibit virus production. *J. Virol.* 85:305–14
110. Chukkapalli V, Inlora J, Todd GC, Ono A. 2013. Evidence in support of RNA-mediated inhibition of phosphatidylserine-dependent HIV-1 Gag membrane binding in cells. *J. Virol.* 87:7155–59
111. Datta SA, Curtis JE, Ratcliff W, Clark PK, Crist RM, et al. 2007. Conformation of the HIV-1 Gag protein in solution. *J. Mol. Biol.* 365:812–24
112. Zhou W, Resh MD. 1996. Differential membrane binding of the human immunodeficiency virus type 1 matrix protein. *J. Virol.* 70:8540–48
113. Tang C, Loeliger E, Luncsford P, Kinde I, Beckett D, Summers MF. 2004. Entropic switch regulates myristate exposure in the HIV-1 matrix protein. *PNAS* 101:517–22
114. Saad JS, Miller J, Tai J, Kim A, Ghanam RH, Summers MF. 2006. Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. *PNAS* 103:11364–69
115. Paillart JC, Gottlinger HG. 1999. Opposing effects of human immunodeficiency virus type 1 matrix mutations support a myristyl switch model of Gag membrane targeting. *J. Virol.* 73:2604–12
116. Reil H, Bukovsky AA, Gelderblom HR, Gottlinger HG. 1998. Efficient HIV-1 replication can occur in the absence of the viral matrix protein. *EMBO J.* 17:2699–708
117. Perez-Caballero D, Hatziioannou T, Martin-Serrano J, Bieniasz PD. 2004. Human immunodeficiency virus type 1 matrix inhibits and confers cooperativity on Gag precursor-membrane interactions. *J. Virol.* 78:9560–63
118. Hatziioannou T, Martin-Serrano J, Zang T, Bieniasz PD. 2005. Matrix-induced inhibition of membrane binding contributes to human immunodeficiency virus type 1 particle assembly defects in murine cells. *J. Virol.* 79:15586–89
119. Holmes M, Zhang F, Bieniasz PD. 2015. Single-cell and single-cycle analysis of HIV-1 replication. *PLOS Pathog.* 11:e1004961
120. Neil SJ, Zang T, Bieniasz PD. 2008. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* 451:425–30
121. Rosa A, Chande A, Ziglio S, De Sanctis V, Bertorelli R, et al. 2015. HIV-1 Nef promotes infection by excluding SERINC5 from virion incorporation. *Nature* 526:212–17
122. Usami Y, Wu Y, Gottlinger HG. 2015. SERINC3 and SERINC5 restrict HIV-1 infectivity and are counteracted by Nef. *Nature* 526:218–23
123. Mirambeau G, Lyonnais S, Coulaud D, Hameau L, Lafosse S, et al. 2007. HIV-1 protease and reverse transcriptase control the architecture of their nucleocapsid partner. *PLOS ONE* 2:e669

124. Sheng N, Erickson-Viitanen S. 1994. Cleavage of p15 protein in vitro by human immunodeficiency virus type 1 protease is RNA dependent. *J. Virol.* 68:6207–14
125. Sheng N, Pettit SC, Tritch RJ, Ozturk DH, Rayner MM, et al. 1997. Determinants of the human immunodeficiency virus type 1 p15NC-RNA interaction that affect enhanced cleavage by the viral protease. *J. Virol.* 71:5723–32
126. Potempa M, Nalivaika E, Ragland D, Lee SK, Schiffer CA, Swanstrom R. 2015. A direct interaction with RNA dramatically enhances the catalytic activity of the HIV-1 protease in vitro. *J. Mol. Biol.* 427:2360–78
127. Engelman A. 1999. In vivo analysis of retroviral integrase structure and function. *Adv. Virus Res.* 52:411–26
128. Desimmie BA, Schrijvers R, Demeulemeester J, Borrenberghs D, Weydert C, et al. 2013. LEDGINs inhibit late stage HIV-1 replication by modulating integrase multimerization in the virions. *Retrovirology* 10:57
129. Jurado KA, Wang H, Slaughter A, Feng L, Kessl JJ, et al. 2013. Allosteric integrase inhibitor potency is determined through the inhibition of HIV-1 particle maturation. *PNAS* 110:8690–95
130. Madison MK, Lawson DQ, Elliott J, Ozanturk AN, Koneru PC, et al. 2017. Allosteric HIV-1 integrase inhibitors lead to premature degradation of the viral RNA genome and integrase in target cells. *J. Virol.* 91:e00821-17
131. Kessl JJ, Kutluay SB, Townsend D, Rebensburg S, Slaughter A, et al. 2016. HIV-1 integrase binds the viral RNA genome and is essential during virion morphogenesis. *Cell* 166:1257–68