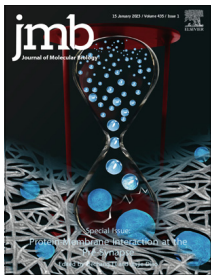

Understanding Retroviral Life Cycle and its Genomic RNA Packaging

Akhil Chameettachal [1†], Farah Mustafa [2,3*] and Tahir A. Rizvi [1,3*]

*1 - **Department of Microbiology & Immunology,** College of Medicine and Health Sciences (CMHS), United Arab Emirates University, Al Ain, United Arab Emirates*

*2 - **Department of Biochemistry,** College of Medicine and Health Sciences (CMHS), United Arab Emirates University, Al Ain, United Arab Emirates*

*3 - **Zayed bin Sultan Center for Health Sciences (ZCHS),** United Arab Emirates University, Al Ain, United Arab Emirates*

Correspondence to Farah Mustafa and Tahir A. Rizvi: Department of Biochemistry, College of Medicine and Health Sciences (CMHS), United Arab Emirates University (UAEU), P.O. Box 15551, Al Ain, United Arab Emirates (F. Mustafa). Department of Microbiology & Immunology, College of Medicine and Health Sciences (CMHS), United Arab Emirates University (UAEU), P.O. Box 15551, Al Ain, United Arab Emirates (T.A. Rizvi). *fmustafa@uaeu.ac.ae (F. Mustafa), tarizvi@uaeu.ac.ae (T.A. Rizvi)@chameettachal (A. Chameettachal)*

https://doi.org/10.1016/j.jmb.2022.167924

Edited by Eric O. Freed

Abstract

Members of the family *Retroviridae* are important animal and human pathogens. Being obligate parasites, their replication involves a series of steps during which the virus hijacks the cellular machinery. Additionally, many of the steps of retrovirus replication are unique among viruses, including reverse transcription, integration, and specific packaging of their genomic RNA (gRNA) as a dimer. Progress in retrovirology has helped identify several molecular mechanisms involved in each of these steps, but many are still unknown or remain controversial. This review summarizes our present understanding of the molecular mechanisms involved in various stages of retrovirus replication. Furthermore, it provides a comprehensive analysis of our current understanding of how different retroviruses package their gRNA into the assembling virions. RNA packaging in retroviruses holds a special interest because of the uniqueness of packaging a dimeric genome. Dimerization and packaging are highly regulated and interlinked events, critical for the virus to decide whether its unspliced RNA will be packaged as a "genome" or translated into proteins. Finally, some of the outstanding areas of exploration in the field of RNA packaging are highlighted, such as the role of epitranscriptomics, heterogeneity of transcript start sites, and the necessity of functional polyA sequences. An in-depth knowledge of mechanisms that interplay between viral and cellular factors during virus replication is critical in understanding not only the virus life cycle, but also its pathogenesis, and development of new antiretroviral compounds, vaccines, as well as retroviral-based vectors for human gene therapy.

Introduction

Retroviruses are enveloped viruses containing two copies of single-stranded positive sense RNA. These viruses belong to a unique family of viruses, the *Retroviridae*, since their life cycle includes the special step of reverse transcription of their genomic RNA (gRNA) into DNA, which is then imported into the nucleus where it integrates into the host chromosomes. According to the

International Committee on Taxonomy of Viruses (ICTV), the family *Retroviridae* is further divided into two subfamilies: *Orthoretrovirinae* and *Spumavirinae* and based on their evolutionary relatedness, they are further classified into seven genera.[1] The subfamily *Orthoretrovirinae* contains the genera *alpharetroviruses*, *betaretroviruses*, *gammaretroviruses*, *deltaretroviruses*, *epsilonretroviruses* and *lentiviruses*, while the subfamily *Spumavirinae* contains a single genus, the *spumaviruses*. Retroviruses infect a large number of host species and cause a variety of diseases, including immunodeficiency, tumors, and leukemia, both in humans and animals. Although the field of retrovirology progressed significantly after the detection of the human immunodeficiency virus type 1 (HIV-1), earlier work has contributed significantly towards the understanding of cellular oncogenes and molecular biology.[2–4]

The retroviral replication cycle includes a cascade of events which are categorized into nine main steps in this review. They are: 1) attachment, membrane fusion and entry, 2) reverse transcription, 3) nuclear import and uncoating, 4) integration, 5) transcription and processing of viral RNAs, 6) their nuclear export, 7) protein synthesis, 8) gRNA packaging and virion assembly, 9) and finally virion release and maturation (Figure 1). Ever since the discovery of the first retrovirus, numerous studies have helped identify the critical viral or/and cellular factors and processes involved in each of these steps. Retroviral infection starts with attachment of the surface subunit of Env protein to its receptor followed by fusion with the cell membrane mediated by its transmembrane domain. Once the core is released into the cytoplasm, reverse transcription begins by the virally-encoded reverse transcriptase enzyme, though in the case of HIV-1 it has been shown that both uncoating of the core and reverse transcription can be completed in the nucleus.[5–6] In the case of MLV, the pre-integration complex can access the chromosome only after disintegration of the nuclear membrane during mitosis,[7] while in HIV-1, the virion core enters the nucleus through the nuclear pore. This process involves a number of host proteins which are part of the nuclear pore complex, as well as mRNA cleavage and polyadenylation factors.[6,8] Uncoating occurs near the integration site minutes before integration and the integration process is catalyzed by the integrase enzyme. Host nuclear proteins also potentially play a role in integration site selection.[9–12]

Transcription of retroviral RNAs from the integrated virus (called a “provirus” at this stage) is facilitated by cellular transcription factors and *trans*-acting factors encoded by complex retroviruses along with their corresponding *cis*-acting elements. For example, HIV-1 encodes the *trans*-activator of transcription, Tat, protein, and its cognate response element, TAR, while human T lymphotropic virus, HTLV encodes Tax protein and its response element, TRE.[13–14] The singly spliced and unspliced viral RNAs in case of complex retroviruses are next transported out of the nucleus using *cis*-acting viral sequences which interact with virally-encoded factors. On the other hand, simple retroviruses, rely solely on cellular proteins for their RNA export by utilizing *cis*-acting sequences present within their genomes. Translation of viral transcripts is achieved by using both the general cap-dependent mechanism as well as cap-independent mechanisms using internal ribosome entry site (IRES)(reviewed in.[15]) Members of *Orthoretrovirinae* subfamily express Pol protein as a fusion protein with Gag and all of them use a ribosomal frameshift mechanism except *gamma*- and *epsilonretroviruses* as they possess a ribosomal readthrough mechanism.[15] The viral gRNA forms dimers through dimerization initiation sites (DIS) and virion assembly begins with multimerization of Gag polyprotein on the gRNA. For the *betaretrovirus* genus (e.g., mouse mammary tumor virus, MMTV), location of assembly is intracytoplasmic,[16] while for *lentiviruses*, particle assembly takes place on the plasma membrane.[17] In case of HIV-1, virion release occurs by the interaction of the p6 domain of Gag and cellular endosomal sorting complex required for transport (ESCRT I and ESCRT II) machinery.[18–19] Finally, the proteolytic maturation and structural rearrangement necessary for virion maturation are carried out by the viral protease which is essential for infectivity and continuation of the viral life cycle.[20–21] An in-depthunderstanding of each of these steps has not only led to the development of therapeutics that inhibit retroviral replication, but also has contributed to the understanding of other cellular processes, such as the mechanisms of oncogenesis and molecular biology. Moreover, a better understanding of these steps has identified mechanisms that are unique to retroviruses, mainly reverse transcription, integration, and packaging of their gRNA as a dimer.

Genomic RNA packaging in retroviruses has gained special attention because their single-stranded gRNA is packaged as a dimer (Figure 1). Dimerization and packaging are interconnected events that are highly regulated, and loss of dimerization can greatly reduce packaging (reviewed in.[22]) Specific recognition of gRNA by Gag is facilitated by *cis*-acting sequences located within the 5′ region of the gRNA termed the packaging signal, Psi (Ψ; reviewed in[23]). It is evident that not only the nucleotides that interact with Gag, but also the higher order structure of Ψ is important for efficient packaging of gRNA by the newly forming virus particles. Development of biochemical assays coupled with structural bioinformatics have identified the complex structure of HIV-1 and other retroviral 5′untranslated regions (UTRs) required for gRNA encapsidation and other steps of the viral

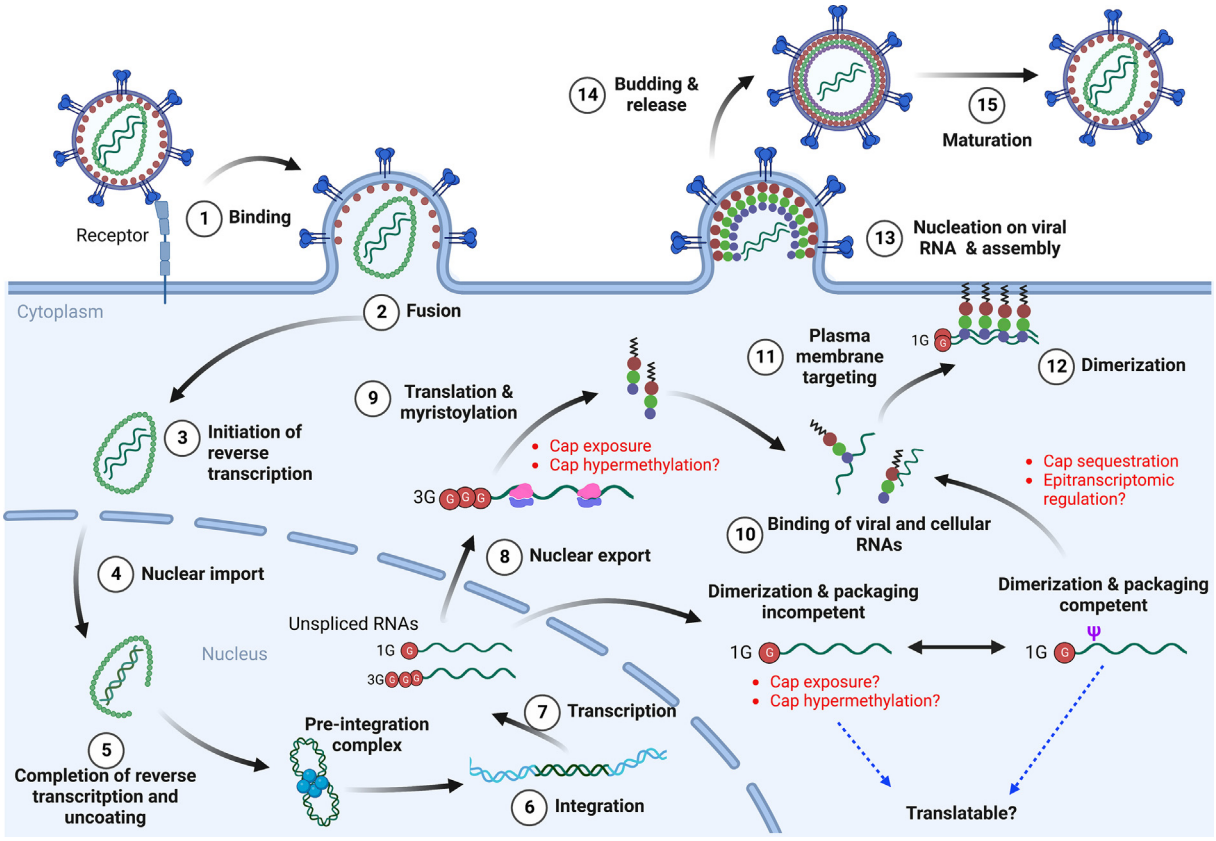

Figure 1. Schematic representation of different stages of retrovirus replication and RNA packaging using HIV-1 as a model. Infection of susceptible cells starts with attachment and fusion of the viral Env glycoprotein to the cell surface receptor, followed by the release of viral core into the cytoplasm (steps 1 & 2). However, certain retroviruses (such as avian leukosis virus and mouse mammary tumor virus) traffic through endosomes which activates their fusion with the membrane. Reverse transcription of the viral RNA begins in the cytoplasm within the core, the core then enters the nucleus, and completion of reverse transcription and uncoating occurs in nucleus near the site of integration (steps 3–5), leading to integration of the viral genome into the host chromosome (step 6). It has been shown that MLV enter the nucleus only during cell division. Transcription of viral RNA occurs from the integrated "provirus" with the help of regulatory proteins and their *cis*-acting sequences, in the case of complex retroviruses (step 7). Recent studies show that in the case of lentiviruses, different species of full-length RNAs are transcribed based on their transcription start site. The unspliced and singly spliced RNAs are then exported from nucleus by the regulatory proteins and their *cis*-acting elements with the help of host proteins (step 8). In case of HIV-1, the 3G RNA (RNA with three guanosines at their 5' end) is shown to be predominantly translated (step 9) while the 1G RNAs are encpasidated. 1G RNA exists in equilibrium between monomer and dimer conformations, in which the dimer conformation sequesters its cap. Thus, it is not captured by the translation machinery, and because of its structural advantages, it is encapsidated by Gag. However, certain studies have shown that the same pool of RNAs can act as both translating and packageable RNA. It is unknown whether the dimerization competent 1G RNA, or monomeric 1G RNA, or both, can undergo translation. Moreover, it has been shown that HIV-1 unspliced and singly spliced RNAs undergo alternate translation pathway other than eIF4E-mediated pathway. However, it is unclear whether 3G or 1G RNA preferentially undergoes this pathway. HIV-1 Gag binds to cellular and viral RNAs with equal affinity in the cytoplasm, but packaging competent RNAs (with packaging signal, psi; 1G RNA in case of HIV-1) favors assembly of Gag (step 10). The Gag-RNA complexes then localize to the plasma membrane where the RNA dimerization occurs during the initial stages of assembly (steps 11–12; in the case of RSV and MLV, it has been suggested that RNA dimerization occurs in the nucleus). Moreover, it has been shown that nucleation and assembly is efficient on packaging signal-containing RNA (step 13). Finally, the completely assembled virus particles are released in an immature form and then the proteolytic maturation and the structural rearrangement occurs (steps 14 &15).

life cycle. These assays include, selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE;[24–25]), nuclear magnetic resonance (NMR) spectroscopy,[26–27] and their modified versions; for instance, SHAPE-coupled with mutational profiling (SHAPE-Map,[28]) and long-range probing by adenosine interaction detection (lrAID;[29]). These studies, along with several other genetic and *in vitro*

approaches, have identified that purine-rich apical and/or internal loops in hairpin structures govern gRNA packaging by functioning as Gag binding sites.[30–35] Moreover, recent structural studies have shown that in HIV-1, RNA species with only one G at their 5′ end are preferentially packaged into virus particles as they fold into a packaging-competent structure.[35–37] Furthermore, the stability of the polyA stem-loop is also necessary for packaging, revealing the importance of regions other than the 'encapsidation signal' in RNA packaging, at least in context of full-length gRNA.[37–38] Thus, in this review, we summarize different steps of retroviral replication to provide a basic understanding of the viral life cycle and present the current understanding of RNA packaging among different retroviruses.

Retroviruses: A brief history and classification

The first identified retrovirus was the causative agent of leukemia in chickens (avian leukosis virus; ALV) by Ellermann and Bang in 1908.[39] In 1911, Peyton Rous reported the transmission of sarcomas in chickens through cell-free filtrates and subsequently it was named the Rous sarcoma virus (RSV;[40]). The first and only known human retrovirus that causes malignancy, Human T-lymphotropic virus type 1 (HTLV-1), was discovered and described in 1980.[41–42] Three years later, Montagnier and co-workers isolated a virus from the lymph nodes of patients with acquired immunodeficiency syndrome (AIDS) and in 1984, the link between HIV-1 and AIDS was established by Gallo and colleagues.[43–44]

Virion structure and genome organization

Retroviruses are spherical in shape measuring approximately 80–120 nm in diameter, comprising of ∼2000–5000 molecules of a structural protein called Group-specific antigen (Gag;[17,45–46]) and ∼ 14 trimers of envelope glycoprotein spikes.[47] Historically, retroviruses have been classified based on the structure and location of their nucleocapsid core within the mature particle. This classification includes: an intracellular membrane-lacking form (intracytoplasmic A particle), type-B (an extracellular eccentric, spherical core; e.g., mouse mammary tumor virus, MMTV), type-C (spherical/cone shaped core; e.g., ALV, RSV, murine leukemia virus (MLV), and type-D (a cylindrical/bar shaped core; e.g., Mason Pfizer monkey virus, MPMV). While HIV-1 is the classical example of a retrovirus with a rod/cone shaped core, the *spumaviruses* exhibit an immature core morphology.

An infectious retrovirus particle consists of two copies of single stranded, 7–12 kilobase (kb) long, linear, non-segmented, positive-sense RNA as their genome which is capped at the 5′ end and polyadenylated at the 3′ end like cellular mRNAs. The key sequences and features of the retroviral gRNA are depicted in Figure 2 and include both non-coding and coding regions described next:

Non-coding regions.

1) R (Repeat): A short sequence (15–250 nucleotides (nts) long; 15 nts in case of MMTV and ∼ 250 nts in HTLV-1) that is repeated twice in the RNA: the extreme 5′ end just after the cap and at the 3′ end immediately upstream of the polyadenylate (polyA) tail. Its flanking location in the gRNA is critical for the process of reverse transcription in all retroviruses (described below). In *lentiviruses*, R folds into a structured region called the *trans*-activation response element (TAR) that recruits the transcriptional *trans*-activator protein Tat required for efficient synthesis of the gRNA.[1,17,48]
2) U5 (Unique to 5′): A sequence of 70–250 nts located downstream of R and before the primer binding site (PBS). U5 contains one of the attachment *s*ites (U5 *att*) necessary for proviral integration into the host chromosome.[1,17,48] Furthermore, during the process of reverse transcription, the U5 region gets duplicated and transferred to the 3′ of the proviral DNA (Figure 2).
3) PBS: An ∼ 18-nt sequence complementary to the 3′ terminus of host tRNAs which acts as the primer for the initiation of reverse transcription for the synthesis of minus (-) strand of viral DNA (first strand synthesis).[1,17,48] Recent studies are unraveling its role in other steps of retrovirus replication as well, such as gRNA packaging.[32]
4) PPT (polypurine tract): During reverse transcription, PPT serves as the primer for the synthesis of the plus (+) strand of viral DNA (second strand synthesis). As the name indicates, it consists of a stretch of 'A' and 'G' residues that are usually 7–18 nts long located just upstream of the U3 region that arise out of partial nuclease digestion of the gRNA (see Section on Reverse Transcription for details).[1,17,48]
5) U3 (Unique to 3′): This region contains *cis*-acting elements important for transcription initiation and its regulation, such as the promoter and enhancer sequences, negative regulatory elements, hormone responsive elements, as well as the second *att* site (U3 *att*) required for integration. This is despite the fact it is located at the 3′ end of the viral gRNA. It is the process of reverse transcription that results in duplication of the U3 at both ends of the viral DNA, thus bringing these important transcription control elements to the 5′ end of the proviral DNA (Figure 2). U3 at times may also contain coding sequences (e.g., a portion of the *nef* gene in HIV-1, as well as the entire *sag* and a portion of MMTV *rem* genes.[1,17,48]
6) PolyA tract: Downstream to the 3′ R region is located the post-transcriptionally added polyA tail of 50–200 nts. The most common polyadenylation signal is AAUAAA and either found in the R region (e.g., HIV, Mo-MLV and *spumaviruses*) or the end

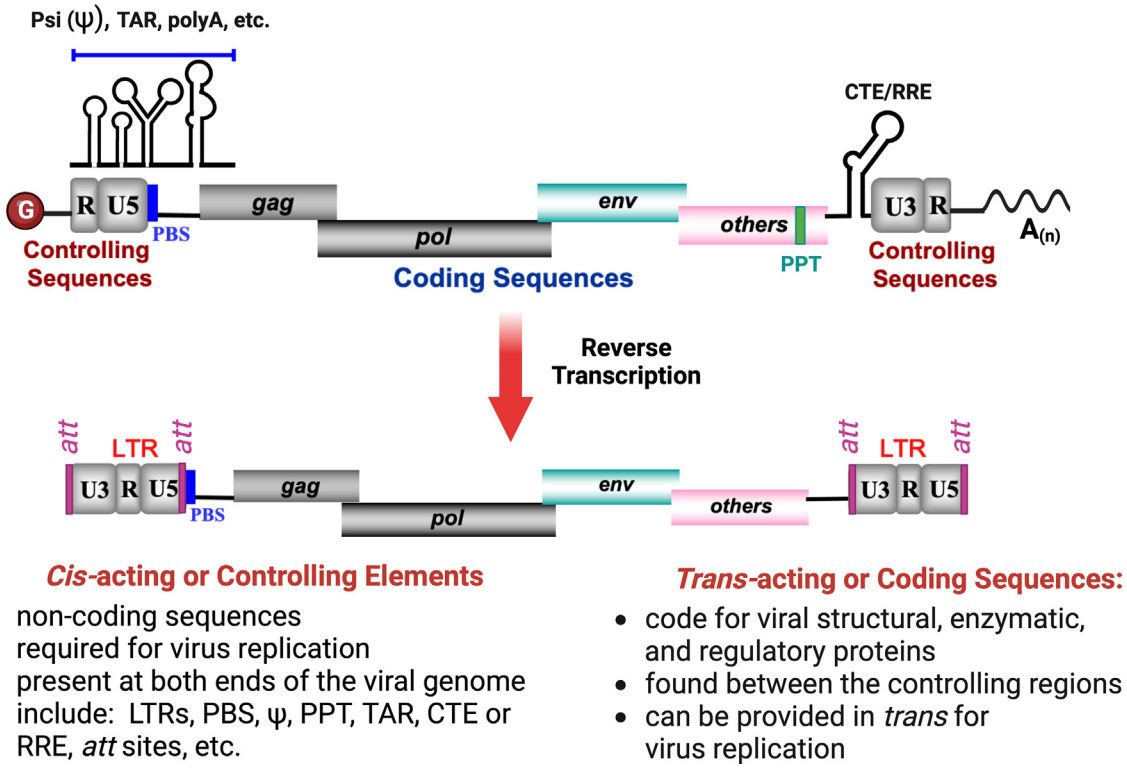

Figure 2. Organization of the coding and non-coding regions of the retroviral genomic RNA and the provirus. The single-stranded, positive sense RNA genome has a 5′ cap and its various regions can be classified into *cis*-acting (controlling) elements on both ends of the genomic RNA flanking the *trans*-acting (coding) sequences. G, the guanine cap at the 5′ end of the genomic RNA; R, repeat region at both the 5′ and 3′ ends; U5 and U3, the 5′ and 3′ unique sequences, respectively; PBS, the primer binding site where the host tRNA binds as a primer to initiate reverse transcription; Ψ, the packaging signal; PPT, the polypurine tract which is the site of initiation of plus strand DNA synthesis; $A_{(n)}$, the polyA sequence; TAR, the trans-activation response element; CTE, the constitutive transport element; RRE, the Rev responsive element; *att*, the attachment sites used for integration of the viral genome. Only the two *att* sites at the outer ends are used for this purpose.

of the U3 region (e.g., HTLV-1, ASLV and MMTV).[1,17,48]

Coding regions. All known retroviruses encode four canonical proteins which are necessary for viral particle formation and its replication: Gag, protease (PR or Pro), polymerase (Pol) and envelope (Env). Based on their genomic complexity, retroviruses are classified into either simple or complex: simple retroviruses encode only *gag, pro, pol* and *env* genes (such as ALV and MLV), while complex retroviruses (such as HIV, HTLV, and MMTV) also encode a number of small regulatory and accessory proteins which are encoded from singly or alternatively spliced mRNAs, (Figures 3 & 4;[1,17,48]).

1) Gag: Gag is expressed as a precursor polyprotein which is cleaved by viral protease during maturation into three structural proteins: matrix (MA), capsid (CA) and nucleocapsid (NC). While MA is involved in targeting the Gag or viral core to the plasma membrane during virus assembly, the CA is the critical building block that creates the viral core and NC is the protein that directly coats the gRNA (Figure 5). Gag is also critical for gRNA encapsidation into the assembling virus particle.
2) Pro and Pol: Like Gag, retroviral Pol is also a polyprotein which is cleaved by the viral protease (PR) and is found within the virus particle (Figure 5). The constituent proteins generated after PR-mediated cleavage depends on the retroviral species. In *lentiviruses*, the PR, reverse transcriptase (RT) and integrase (IN) are encoded in the *pol* gene and expressed as a Gag-Pol fusion protein, while in *alpharetroviruses*, PR is expressed as part of the Gag open reading frame (ORF; Figure 4). In *betaretroviruses*, PR is expressed from a separate ORF as a Gag-Pro or a Gag-Pro-Pol fusion protein (Figure 4;[1,17,48]).
3) Env: Env is translated from a singly spliced mRNA and is a membrane targeting protein which is necessary for retroviral attachment, fusion, and entry.

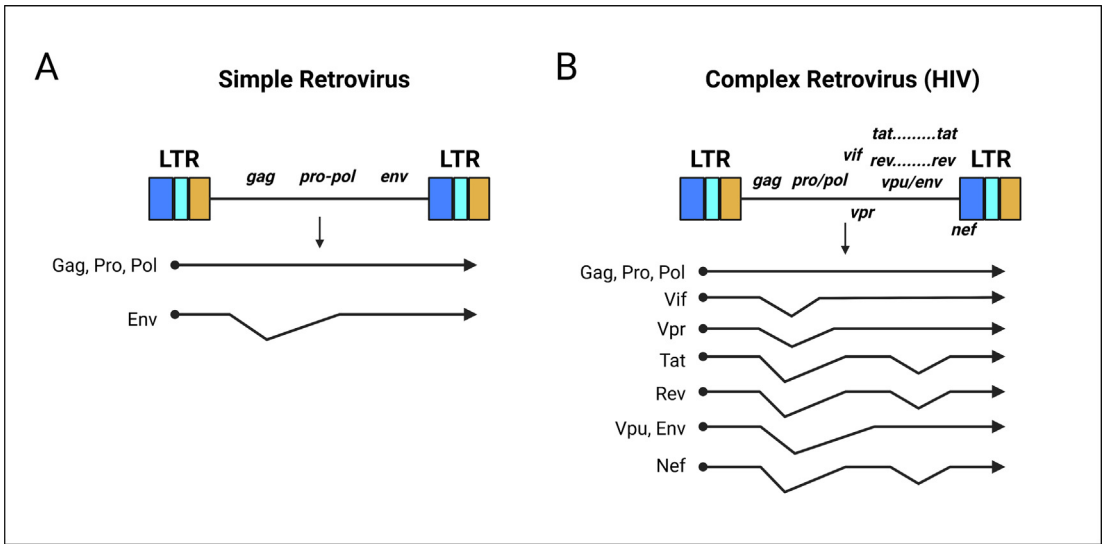

Figure 3. Genome structure and RNA splicing in simple and complex retroviruses. (A) Simple retroviruses encode *gag, pro, pol,* and *env* genes and carry out only a single splicing event to express Env. **(B)** The complex retroviruses, in addition, contain regulatory and accessory genes that are expressed from several singly- and doubly spliced mRNAs. Here, the example of human immunodeficiency virus (HIV-1) is given. LTR, long terminal repeats.

The host cellular protease, furin, cleaves Env into the surface (SU) and transmembrane (TM) domains (Figure 5;[1,17,48]).

Retroviral life cycle

Attachment, membrane fusion, and entry. This earliest step in retroviral replication involves multiple stages: initial binding of the SU subunit of the Env protein to the cellular receptor which induces a conformational change in the Env protein that results in membrane fusion (Figure 1). The initial attachment leads to a conformational change in Env which exists as a trimer, allowing the TM domain penetration into the target cell membrane, resulting in fusion of viral and cell lipid bilayers.[49–50] In many viruses, such as orthomyxoviruses, this conformational change is aided by the acidic environment within the endosomes and takes place after receptor-mediated endocytosis. But in most retroviruses, this fusion is pH-independent; thus, it can happen at the cell surface.[17] The classical experiments to study the effect of pH on fusion were carried out using lysosomotropic reagents, such as chloroquine and ammonium chloride, that raise endosomal pH. These experiments showed that HIV-1 fusion is insensitive to a wide range of pH.[51–53] On the other hand, in the case of ALV, and MMTV, a low pH-mediated activation was required for virus entry following the conformational change that occurs after Env-receptor binding.[54–55]

The type of host receptor that Env can interact with determines retroviral cell tropism. Retroviruses utilize various cell surface molecules as receptors, reflecting the wide range of host species that they can infect. HIV-1 uses the CD4 (cluster of differentiation 4) receptor which is expressed on the helper subset of T lymphocytes, macrophages, and dendritic cells, but engineered expression of CD4 alone on non-primate and rodent cells is not sufficient for productive infection.[56–57] These observations suggested that other proteins were needed to facilitate virus infection, leading to the identification of two members of the chemokine receptor family, CXCR4 and CCR5, as co-receptors for HIV-1 entry.[58] This was confirmed by a population of individuals with a mutated allele of CCR5 (a 32-bp deletion) who exhibited least susceptibility towards HIV-1 infection.[59] Moreover, some HIV-1 variants and HIV-2 strains are able to infect cells in a CD4-independent manner,[60–62] suggesting other modes of virus entry.

The co-receptors for other retroviruses have not been studied extensively. In HTLV-1, the primary entry receptor is unclear. Although initially it was shown that the glucose transporter molecule GLUT-1 is the primary receptor, heparan sulfate proteoglycans (HSPGs) and neuropilin-1 (NRP-1) also play a critical role in entry.[63–65] Additionally, the vascular cell adhesion molecule-1 (VCAM-1) has been identified as the potential co-receptor for HTLV-1 entry.[66] In case of MMTV, it uses the mouse transferrin 1 as its cellular receptor.[55] Interestingly, it has been shown that an immunoreceptor tyrosine activation motif (ITAM) within the Env is involved in transforming the mammary epithelial cells. Thus, Env not only allows the virus to infect cells, but may also have a role in mammary tumorigenesis in case of MMTV.[67]

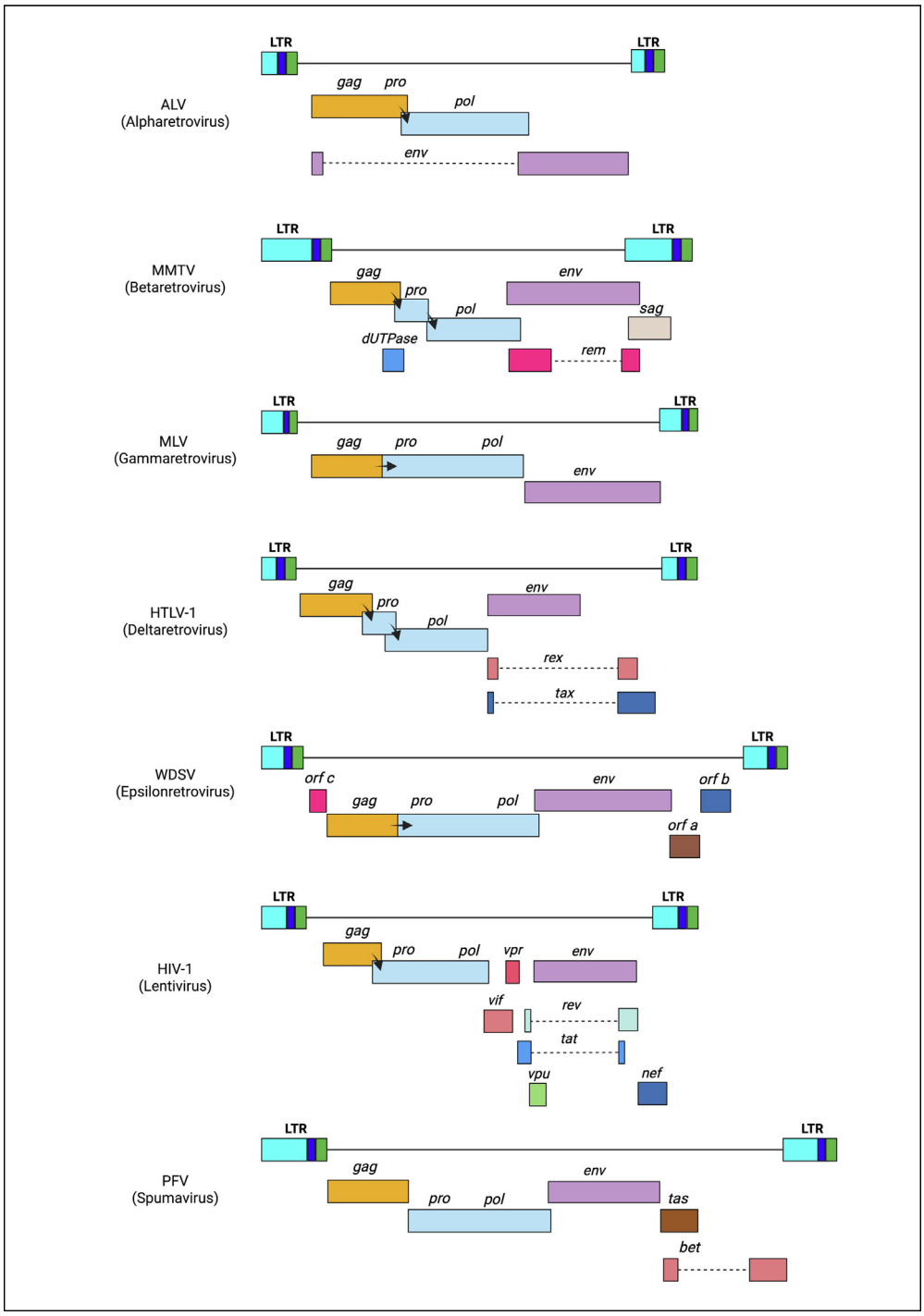

Figure 4. Genomic organization of the provirus for each retroviral genus. Schematic representation of the proviral genome of the prototypic member from each genus, showing the organization of their long terminal repeats (LTRs), the four canonical genes (*gag, pro, pol,* & *env*), and regulatory as well as accessory genes in the case of complex retroviruses. Sites of ribosomal frameshift or ribosomal readthrough are indicated as curved or right pointing arrows, respectively.

Reverse transcription. Evidence suggests that the process of reverse transcription begins in the cytoplasm, but a major fraction of it occurs and completes within the nucleus (Figure 1).[6,68–69] Treatment with inhibitors of reverse transcription results in delayed uncoating, suggesting a check-

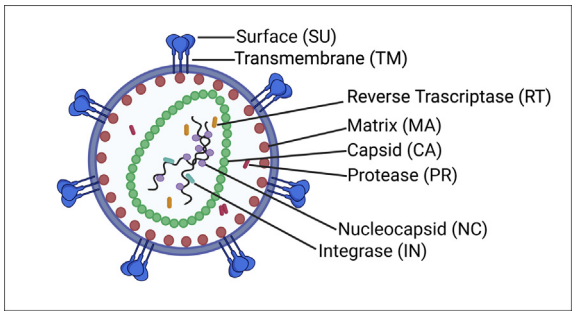

Figure 5. Schematic of a retrovirus (HIV-1) particle. The viral RNA is coated by the nucleocapsid (NC) protein. This genomic RNA-NC complex along with the reverse transcriptase (RT) and integrase (IN) are found within the viral core which is formed by a mature product of the Gag polyprotein, specifically the capsid (CA) protein. The matrix (MA) and NC are the other mature products of the Gag polyprotein. The capsid is connected to the host-derived lipid bilayer that the virion acquires as it buds out of the cell via the matrix (MA) protein. The lipid bilayer is part of the viral envelope composed additionally of the viral envelope (Env) glycoproteins that have a surface (SU) and a transmembrane (TM) subunit. The viral protease (PR/Pro) enzyme is found outside the core and responsible for maturation of the Gag polyprotein once the virus has bud out of the cell in C-type retroviruses, or within the cell for A- and B-type retroviruses. Pro is often expressed as a fusion protein with Pol within the same reading frame of reverse transcriptase and integrase (e.g., HIV-1) or from a separate reading frame (e.g., MMTV).

point that ensures the release of properly formed pre-integration complexes (PICs).[6] The functional reverse transcription complex (RTC) contains the dimeric RNA, RT, CA, NC, and IN and some accessory proteins, depending upon the virus. For example, Vpr is found in RTC in the case of HIV-1.[70] The RT enzyme possess two activities: a DNA polymerase activity that uses either RNA or DNA template to synthesize a DNA strand, and an RNase H activity that degrades RNA in the RNA:DNA hybrid.[17] The lower fidelity of the polymerase activity of RT contributes significantly to the higher mutation rate observed in HIV-1[71]; however, the cellular RNA polymerase II and host factors which are incorporated into the virus particle (such as APOBEC, a single-stranded DNA cytidine deaminase, especially APOBEC3G), induce a C-to-U change in retroviral DNA, contributing to the hypermutations observed in the HIV-1 genome.[72–74] The viral RNA has only the U5 and U3 sequences at its outer flanks, corresponding to its 5′ and 3′ ends. On the other hand, the proviral DNA contains both U3 and U5 on either end (Figure 2). Together, the U3, R and U5 at either end of the proviral DNA are known as long terminal repeats (LTRs) which are created as a by-product of the reverse transcription process (Figures 2 & 6). Since the viral promoter is located in the U3 region, duplication of U3 to the 5′ end of the DNA is necessary for further steps of viral gene expression. Hence duplication of LTR at both ends during reverse transcription is mandatory.[17]

Retroviral reverse transcription of viral RNA into double stranded DNA (dsDNA) involves several steps (Figure 6; summarized in[75]) In the first step, a short stretch of minus strand DNA synthesis occurs copying the 5′ U5 and 5′ R, using specific host tRNAs as the primer (Figure 6(A)). These tRNAs provide the free 3′ hydroxyl group to initiate reverse transcription, a requirement of DNA synthesis as observed in all DNA polymerases. The 17–18-nt long PBS binds to the 3′ acceptor and TΨC arms of the tRNA and determines which tRNA will bind there.[76–77] While HIV-1 and MMTV use $tRNA^{Lys3}$ as their primer, RSV uses $tRNA^{Trp}$, and HTLV-1uses $tRNA^{Pro}$[17] The specific packaging of tRNA is achieved by enrichment of the corresponding charged aminoacyl tRNA synthetase in the assembling virions.[78]

Next, the RNase H activity of RT degrades the RNA from the DNA-RNA hybrid, exposing the 5′ R region which is in turn complementary to the 3′R region (Figure 6(C)). Complementary base pairing of the R region on either end leads to the first template switching known as minus strand transfer (Figure 6(D)). This template switching has a significant role in retroviral recombination since it can occur either intra- or inter-molecularly, especially given the fact that retroviruses contain two copies of the gRNA.[79–80] After successful minus strand transfer, RT further extends the minus strand DNA, including the U3 located at the 3′ end (Figure 6 (E)). In the next step, the RNase H activity of RT degrades the RNA located between the PPT and 3′ end of the viral RNA, leaving a small stretch of PPT behind which is resistant to RNase H degradation due to its purine enrichment,[75] Figure 6(E)). Now PPT acts as primer for positive DNA strand synthesis (Figure 6(F)). In *lenti-* and *spumaviruses*, a central PPT/cPPT is also present in the integrase reading frame, which acts as an additional site of plus strand initiation, and mutating cPPT leads to reduced infectivity in HIV-1.[81–85] Degradation of tRNA by RNase H creates an overhanging PBS region at the 3′ end of positive strand DNA (Figure 6 (F)), which is complementary to the PBS at the 5′ end of the minus strand DNA, which leads to the second template switching or plus DNA strand transfer (Figure 6(G)). Finally, the polymerase activity of RT completes the second strand synthesis, resulting in the generation of a double-stranded copy of the gRNA except that the DNA now has two complete LTRs at both ends,[75] (Figure 6(H)).

Nuclear import and uncoating. After fusion and entry, the mature retroviral core enters the cytoplasm, leaving SU, TM and MA behind that remain associated with the membrane. Within the

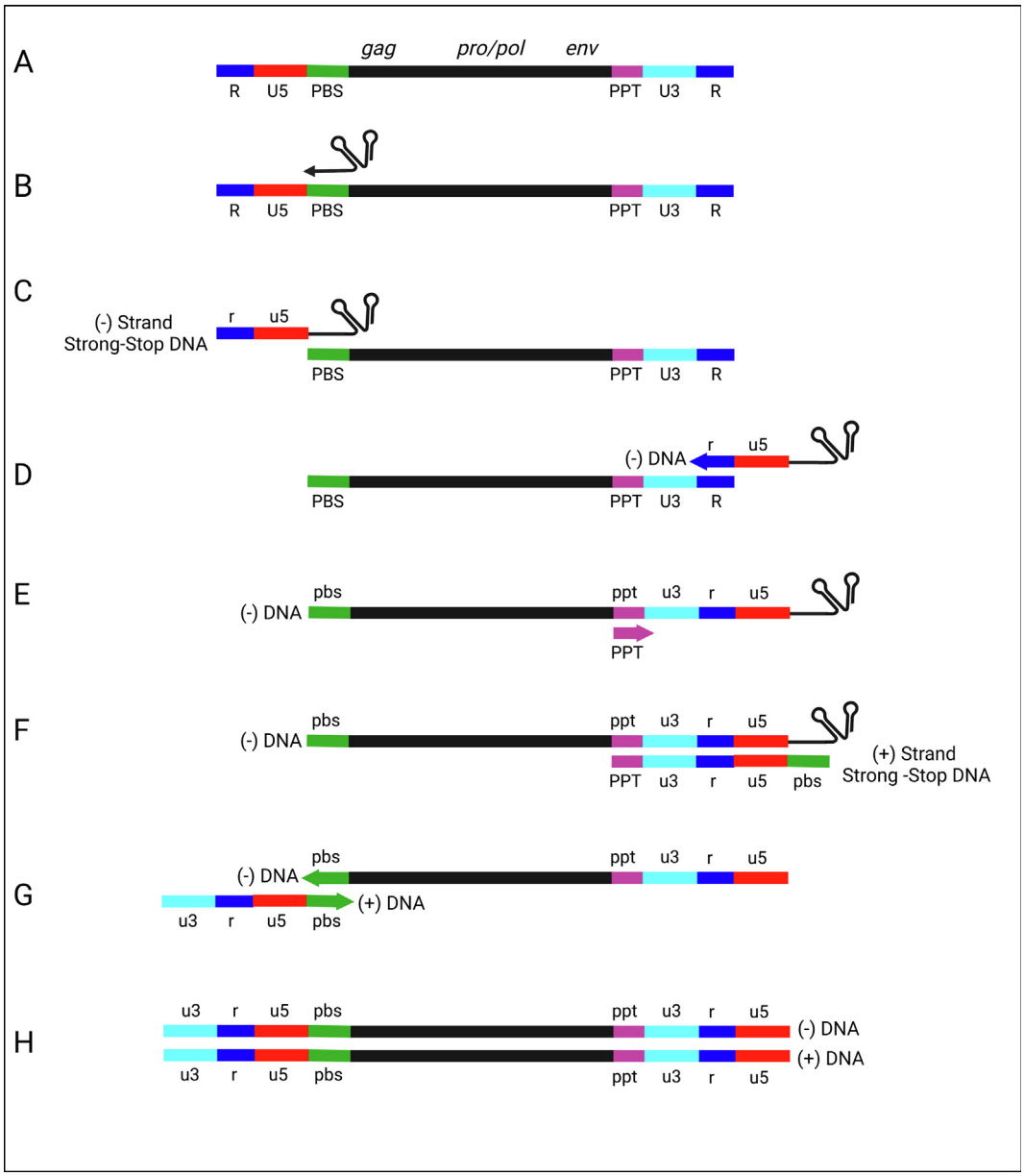

Figure 6. Steps of retroviral reverse transcription. (A) Organization of the retroviral RNA genome. **(B)** A host tRNA binds to the primer binding site (PBS) and acts as a primer for initiation of reverse transcription. **(C)** Synthesis of minus-strand strong stop DNA and removal of R and U5 region of RNA strand by the RNase H domain of the reverse transcriptase enzyme. **(D)** First strand transfer to the 3′ R. **(E)** Extension of minus-strand DNA and digestion of the RNA strand in the RNA/DNA hybrid. **(F)** The polypurine tract (PPT) is resistant to digestion by RNase H and is used for the synthesis of plus-strand DNA. **(G)** Second strand transfer and extension of minus- and plus-strand DNAs. **(H)** Completion of reverse transcription with two fully completed strands of DNA containing u3, r, u5 at either ends.

capsid core, the gRNA remains attached to NC and other enzymatic proteins. The intracytoplasmic trafficking of the virus core to the nucleus occurs through cytoskeletal fibers[86–87]. In case of MLV, the entire viral core enters the nucleus during mitosis[7]. However, HIV-1 uses an active import mechanism to enter the nucleus through nuclear pores, irrespective of whether the host cells are dividing or not,[8,88] although its precise mechanism is not clear. Due to the bigger size of the HIV-1 core (~60 nm) compared to the diameter of the nuclear pore (~30 nm), initially it was assumed that uncoating occurs within the cytoplasm and only the PIC crosses the nuclear membrane. Consistent with this, CA is not present in the PIC.[89–90] Nevertheless, the intracellular location where HIV-1 uncoating occurs is still debatable (Figure 1). Several earlier studies have suggested that the initial hypothesis is correct and uncoating happens in the cytoplasm after the entry, but before nuclear

import[91–94]. Now, it is becoming increasingly clear that HIV-1 uncoating occurs in the nucleus. A recent study demonstrated the nuclear import of intact HIV-1 core and uncoating near the integration sites that occur only after the completion of reverse transcription and formation of PIC.[6,11] The two nuclear pore complex proteins, Nup153 and Nup358, interact with CA in order to facilitate nuclear import of the viral core[95–97]. Additionally, several studies have shown that CA interacts with the cellular cleavage and polyadenylation specificity factor 6 (CPSF6) to facilitate nuclear import of intact viral cores[5–6,69,98]. Finally, CPSF6 binding to CA results in release of the core to the nucleus from the nuclear pore complexes, possibly by competing with Nup153 for CA binding.[99–100]

Interestingly, recent studies have shown that the initiation of HIV-1 core uncoating is triggered by reverse transcription.[101–103] It has been shown that the growing dsDNA builds pressure inside the viral core and leads to mechanical disruption of the viral core and different stages of reverse transcription may induce distinct mechanical changes to the core.[102–104] However, it is not possible to exclude the involvement of cellular proteins in the uncoating process since most of these studies were performed *in vitro*.[101–104]

Integration. Integration of retroviral dsDNA generated through reverse transcription into the host chromosome is necessary for the continuity of retroviral life cycle. Within the PIC (or intasome), the dsDNA is complexed primarily with the IN enzyme (reviewed in[105]). The integrated viral DNA at this point is called a "provirus" and serves as the template for transcription of viral RNAs (Figure 1). The initial seminal studies describing the fundamental steps and mechanisms involved during integration of the reverse transcribed DNA were carried out in MLV.[106] IN is the key viral protein in the integration process, expressed as a fusion protein along with Gag, RT and PR and cleaved by PR during maturation. In addition to IN and retroviral DNA, PIC also contains host-derived factors, such as lens epithelium-derived growth factor (LEDGF;[12]) barrier to auto-integration factor (BAF,[9]) and high mobility group chromosomal protein A1(HMGA1,.[107]) However, LEDGF probably is the only protein among these to directly interact with the IN enzyme in HIV-1.[10] LEDGF protects IN from proteasomal degradation, helps to tether intasome to chromosomal DNA, and enhance strand transfer during integration.[12,108–109] In case of MLV, LEDGF does not bind to IN; instead, IN binds to the bromodomain and extra terminal domain protein (BET;[110]). Interestingly, it has been shown that IN plays a role beyond integration. Class II IN mutants defective for binding to viral RNA result in altered HIV-1 particle morphology in which the viral RNA is located outside the core in mature virus particles that were not infectious.[111–112] This study also ruled out the possibility of a potential role of IN in RNA packaging as the amount of RNA in virus particles were not reduced in these mutants.

In terms of the *cis*-acting sequence required for integration of proviral DNA into the host chromosome, retroviruses use attachment sites (*att* sites) located in the U5 and U3 regions of reverse transcribed retroviral DNA, often found as imperfect inverted repeats (Figure 2;[113]). Surprisingly, the sequences of *att* sites between different retroviruses are not highly conserved, except for conserved CA and TG dinucleotides (reviewed in[105,114]). The consensus is that the conserved CA and TG dinucleotide base pairs in the U5 and U3 *att* sites play an important role in integration,[17] though it was shown that the CA dinucleotide at the 5′ end of U3 region is not necessary for integration in case of RSV.[115] In the initial step of integration, the endonuclease activity of IN cleaves at the canonical CA sequences at both the U3 and U5 *att* sites, resulting in the release of dinucleotides and generating staggered ends. In the second step, a transesterification reaction joins the 3′ OH groups of the viral DNA to the host DNA by forming a phosphodiester linkage (Figure 7(A-C)). The resulting gaps in the unjoined 5′ ends of viral DNA are repaired by the host cell machinery, though involvement of RT and IN has also been proposed.[116–117] This repair results in the duplication of a short sequence in the target sites (Figure 7(D)). The number of duplicated nucleotides vary from 4-6, depending upon the virus.[118]

Integration site selection within the host chromosome by the retroviruses has been shown to be non-random, yet virus-specific though (as mentioned earlier), no strict sequence specificity exists in the attachment sites. In general, PICs preferentially target DNA associated with histones (nucleosomes) over naked DNA.[119–120] Furthermore, based on analysis of several retroviral integration site sequences, numerous studies have shown the existence of a weak, virus-specific preference for consensus palindromic sequence during retroviral integration.[121–123] Thus, different retroviruses have evolved to prefer different integration sites based on the chromatin state (reviewed in.[105]) For example, *lentiviruses* (e.g., HIV-1, HIV-2, SIV), *gammaretroviruses* (e.g., MLV), and *spumaviruses* (foamy virus (FV)) select transcribing regions. Within the transcribing regions, lentiviruses equally prefer the entire transcribing region, but MLV prefers the 5′ end of the transcribing region.[124–125] The *alpharetroviruses* (e.g., ASLV) and *deltaretroviurses* (e.g., HTLV-1), on the other hand, show only a weak preference for the transcribing regions, while *betaretroviruses* (MMTV) reveal the most random integration pattern observed in all retroviruses.[105,126] The IN and CA proteins have been implicated in integration site

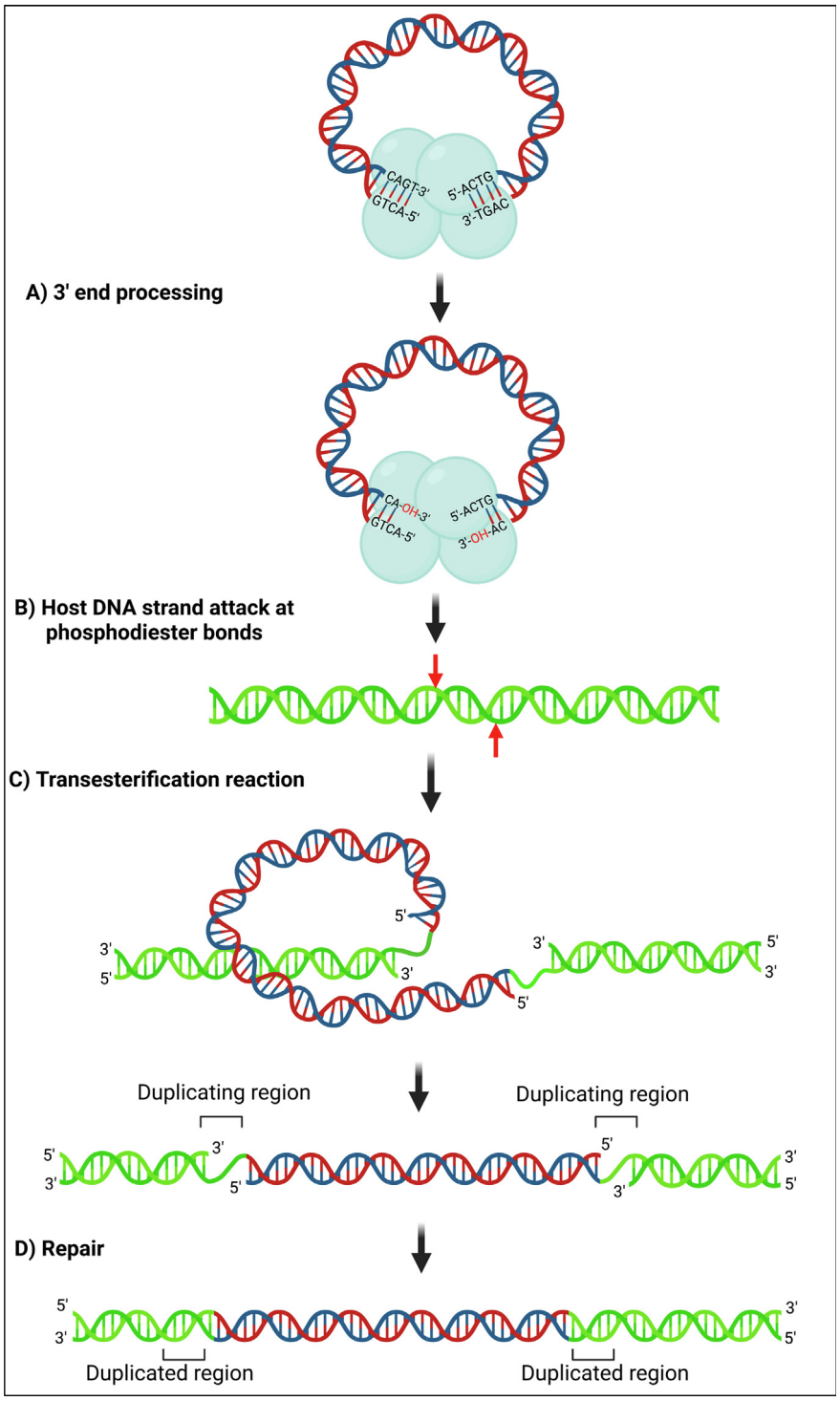

Figure 7. Different stages of the retroviral integration process. A) The integrase enzyme cleaves at the 3′ ends of the viral DNA by its endonuclease activity, releasing 'GT' dinucleotides from the 3′ end, resulting in the generation of 5′ overhangs. **B & C)** The 3′-OH groups of the ends of the viral DNA attack the phosphodiester bonds on target DNA and the viral DNA 3′ ends are joined to the target DNA by a transesterification reaction. The distance between the two phosphodiester bonds at which the transesterification occurs is dependent upon the virus (for example, for HIV-1, it is 5 nucleotides and for MMTV, 6 nucleotides). **D)** The nucleotide gaps are then repaired by the host cell machinery, resulting in duplication of the site of integration which now flanks the integrated provirus.

selection.[127–128] Moreover, the role of host factors such as LEDGF (in *lentiviruses*) and BET (in MLV) has also been established in this process.[12,129].

Transcription and processing of viral RNAs. Expression of viral RNAs from the integrated provirus is facilitated by host enzymes. The U3 region contains a promoter for RNA polymerase II and other regulatory elements. The newly transcribed retroviral mRNAs and the genomic RNA resemble eukaryotic mRNAs and are 7-methyl guanosine capped at the 5′ end and polyadenylated at the 3′ end. However, unlike eukaryotic mRNAs that normally encode one protein, retroviral full-lengh RNA gives rise to multiple proteins from a single mRNA using two different mechanisms (translational read-through and ribosomal frameshift mechanisms). Additionally, they increase their coding capacity by generating subgenomic RNAs from canonical and alternate splicing (Figure 3; reviewed in[17]). Apart from the promoter, the U3 region also contains several regions that binds to cellular transcription factors. For example, in the case of HIV-1, the cellular transcription factor NFκB induces expression from the 5′ LTR in T lymphocytes.[130] The MMTV U3 region contains binding sites for glucocorticoid hormone receptors which activate basal transcription and regulate MMTV infection in a tissue specific manner.[131]

Apart from cellular transcription factors, complex retroviruses encode for *trans*-acting regulators of transcription (transcriptional *trans*-activators). *Trans*-activators expressed at very low levels are sufficient to drive high levels of viral RNA expression. The best studied retroviral *trans*-activators are Tax and Tat proteins of HTLV and HIV, respectively. The Tax protein does not directly bind to the Tax responsive element (TRE) on the HTLV LTR; instead, it enhances binding of host transcription factor cAMP response element binding/activating transcription factor (CREB/ATF) to TRE in the LTR of the provirus.[14] The spumaviruses encode Tas (trans-activator of spumavirus), which binds to the Tas responsive elements in the U3 region and an internal promoter in the *env* gene from which Tas protein is expressed. However, the exact mechanism and the cellular factors involved, if any, in Tas-mediataed transcriptional activation is unclear.[132–133] On the other hand, in the case of HIV, Tat protein binds to a stem-loop structure located in the R region (the transactivation response element, TAR) of newly transcribing RNAs and increase the efficiency of transcriptional elongation by recruiting host super elongation complex (SEC) containing the positive transcript elongation factor, P-TEFb and the PAF-1 (polymerase-associated factor 1) complex. P-TEFb phosphorylates RNA polymerase II and host factors DRB sensitivity-inducing factor (DSIF) and negative elongation factor (NELF) to ensure transcript elongation with the help of the PAF complex recruited via a flexible AFF4 scaffold.[13,134]

Like eukaryotic mRNAs, termination of transcription of retroviral mRNAs involves polyadenylation at the 3′ end. The signal for polyadenylation is a highly conserved *cis*-acting sequence AAUAAA located 10–30 nts upstream of polyadenylation site which binds to the cellular cleavage and polyadenylation specificity factor (CPSF). A GU-rich or U-rich weakly conserved element 10–30 nts downstream of polyadenylation site is also important as a signal for mRNA cleavage and polyadenylation.[17,135] Considering that the retroviral LTRs flank the viral genome and both LTRs contain the transcript termination sequences, retroviruses have evolved various mechanisms to regulate transcription termination. This involves either spatial segregation of their *cis*-acting transcriptional control elements or interaction with splicing elements since transcript termination and splicing happen co-transcriptionally and therefore are coupled processes (reviewed in.[136]) In RSV, MMTV, and HTLV, the polyadenylation sequences are found within the U3 region of the LTR, upstream of the promoter. Thus, the polyadenylation site is not found within the transcripts and no repression of the upstream polyA signal is required. However, in case of other viruses like HIV-1, MLV and FV, the AAUAAA signal sequence is present in the R region from where transcription is initiated; thus, in these retroviruses, the termination sequences within the 5′ LTR must be repressed to prevent premature transcription termination. This is accomplished by various mechanisms in which the distance between the 5′ cap site and the polyadenylation sequences[135] and interaction with the RNA splicing machinery have been implicated.[137–139] For instance, base pairing of the U1SnRNA (part of spliceosome complex) with the major splice donor site has been observed to be the main determinant of splice-site mediated suppression of the 5' polyA site activity as the 3′ LTR does not contain the downstream major splice donor.[140–141] The same mechanism has been shown to exist in *spumaviruses* as well.[136]

After the completion of transcription, a portion of the primary transcripts undergo splicing to increase their coding capacity. In simple retroviruses, only single splicing occurs, while in complex retroviruses, both single and multiple splicing events occur in order to generate subgenomic RNAs (Figure 3;[17]).

Nuclear export. Both unspliced and spliced RNAs in retroviruses must be exported out of the nucleus into the cytoplasm, unlike the cellular RNAs where only the spliced RNAs are transported out of the nucleus. This is necessary in retroviruses since they use their unspliced RNA to translate viral proteins as well as provide gRNA for packaging into the assembling virions. Complex

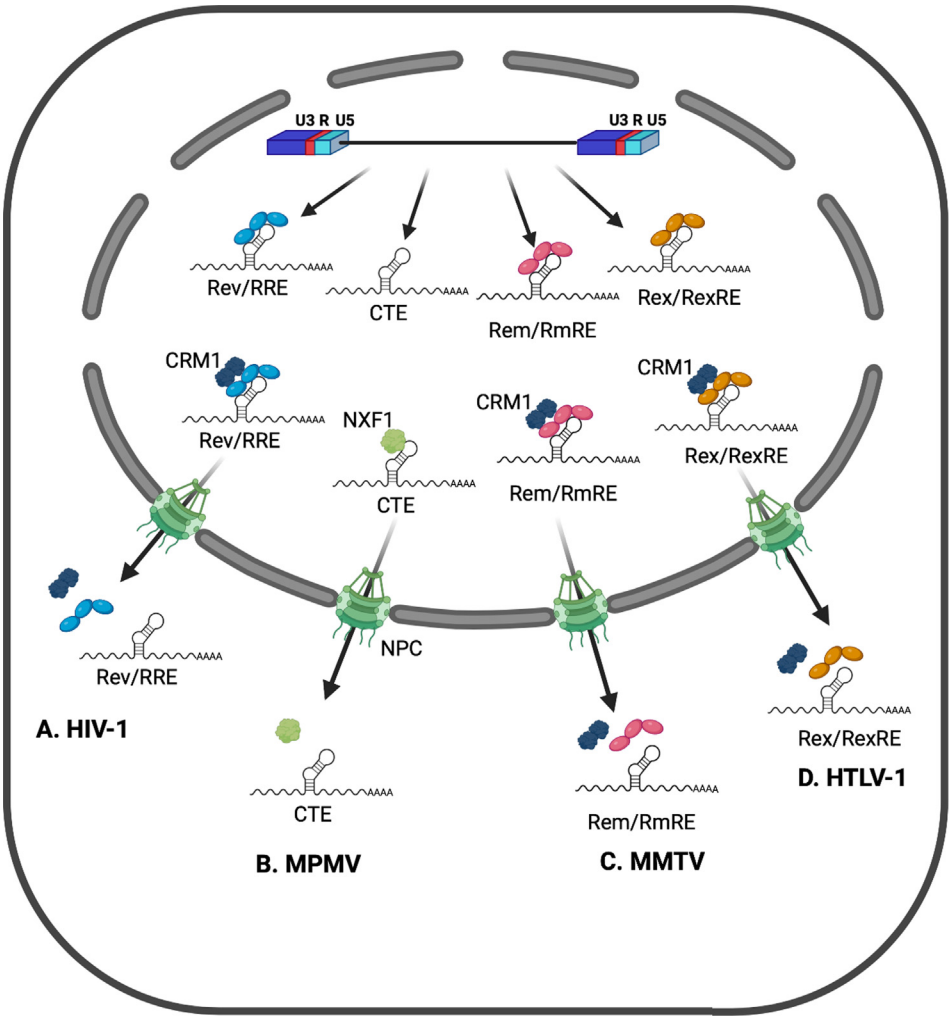

Figure 8. Nuclear export mechanisms observed in different retroviruses. Complex retroviruses such as HIV-1, HTLV-1, and MMTV have evolved to encode regulatory genes (*rev*, *rex*, *rem*, respectively) that interact with *cis*-acting regions within the viral genome (RRE, RmRE, and RexRE) to facilitate nuclear export of viral transcripts by recruiting the cellular nuclear export protein CRM1, the chromosome region maintenance 1 (CRM1). The simple retroviruses, on the other hand, have *cis*-acting regions only, that directly bind to the cellular nuclear export proteins to facilitate nuclear export of its unspliced genomic RNA. For example, MPMV encodes the constitutive transport elements (CTE) that binds NXF1 (Nuclear RNA Export Factor 1). NPC, nuclear pore complex.

retroviruses encode a regulatory protein which binds to its cognate *cis*-acting element to facilitate this nuclear export. In the case of HIV-1, the Rev protein binds to the Rev responsive element (RRE), a highly-structured, conserved RNA element located within the envelope region and consequently present in the unspliced and singly-spliced RNAs (Figure 8)[142–143]. The Rev-RRE complex binds to the cellular nuclear export proteins CRM1 (XpoI) and Ran-GTP to accomplish this task (reviewed in.[144–145] In the absence of Rev, the unspliced and singly-spliced RNAs accumulate in the nucleus.[143] In HTLV-1, the Rex protein binds to the RexRE located in the 3′ UTR and hence present in all HTLV-1 RNAs.[146–147] Similar to HIV Rev, Rex facilitates the export of unspliced and spliced mRNAs to the cytoplasm using he CRM1 cellular pathway (Figure 8).[148] Interestingly, MMTV is a *betaretrovirus* that uses a mechanism similar to these viruses (Rem/RmRE) to export its unspliced and singly spliced RNAs using the CRM1 pathway, an observation which led to MMTV being classified as a complex retrovirus.[149–151] In the case of MPMV, a simple retrovirus, the constitutive transport element (CTE) which overlaps with the 3′ UTR and 3′ LTR, acts as the *cis*-acting element for the nuclear export of viral RNAs.[152] However, CTE accomplishes this not via interaction with a viral factor, but by directly interacting with the cellular TAP/NXF1 host nuclear export pathway through its RNA binding domain to accomplish nuclear export of MPMV mRNAs via the nuclear pore (Figure 8).[153–155]

Translation. Retroviruses encode Gag-Pol or Gag-Pro-Pol proteins in the form of precursor polypeptides, which are processed by virally encoded protease to release the enzymatic proteins that are fused to Gag. Hence expression of Gag and Gag-Pol are highly regulated in order to maintain proper Gag and Gag-Pol ratios. Gag uses a stop codon to terminate translation within the Gag ORF, but skipping this stop codon leads to expression of Pol from the same ORF (e.g., MLV or feline leukemia virus, FeLV) or from a different ORF (e.g., HIV-1, HIV-2, MMTV, MPMV), resulting in Gag-Pol fusion proteins. The first mechanism for such a skipping of Gag termination codon is known as "translational read-through" (reviewed in.[15,17]) In MLV and FeLV, the amber termination codon (UAG) is occasionally mis-decoded as glutamine (CAG) by the tRNA.[156–157] In most other retroviruses, the "frameshift" mechanism exists where the ribosome slips one nucleotide backwards before the Gag termination codon, resulting in a change in ORF, consequently skipping the termination codon to continue translating in the Pol ORF.[15,17] The mechanism of ribosomal frameshifting is based on two signal sequences: a hepta-nucleotide sequence, termed "slippery sequence" of shift site followed by a higher order RNA structure called a "pseudoknot" which is required for efficient frameshifting.[158] In general, the slippery sequences have a XXXY YYZ pattern; for example, in the case of HIV-1, the slippery sequence is U UUU UUA.[159] MMTV has two such sequences; the slippery sequence for the first MMTV frameshift between Gag and Pro follows the pattern and is A AAA AAC, but the proposed second slippery sequence for the frameshift between Pro and Pol does not follow this pattern. Instead, it is a tetranucleotide sequence, UUUA.[160] Furthermore, the structured pseudoknot present in the form of a stem-loop located downstream of the frameshift site results in ribosome stalling which aids tRNA slippage.[15] In case of MMTV, FIV and simian retrovirus-1 (SRV-1), the pseudoknot assumes a more complex structure, suggesting the mechanism is more complicated than just ribosome stalling.[17,158]

Ribosomal frameshifting in HIV-1 results in expression of Gag and Gag-Pol in a ratio of 20:1 (i.e., 5% of translational events;.[159]) It has been shown that altering this ratio is detrimental to viral infectivity.[161] A recent study has shown that incorporation of Gag-Pol into virions is more efficient in *cis* (i.e., when Gag and Gag-Pol are expressed from the same mRNA) than Gag-Pol in *trans*. This suggests a possible compartmentalization of Gag and Gag-Pol during particle assembly.[162] The MMTV pseudoknot acquires a more complex structure when compared to the simple stem-loop in the case of HIV-1[17]. In MMTV, initially it was suggested that ribosome stalling at the pseudoknot aids tRNA slippage, but due to the complex structure of the pseudoknot, it was later proposed that the unpaired A in the apical loop of pseudoknot contributes to frameshifting as it reduces the coaxial stacking of helices.[17,160,163]

In addition to cap-dependent translation, it is becoming increasingly clear that the ribosome utilizes an internal ribosome entry site (IRES) located upstream of Gag initiation codon in many retroviruses for the expression of their structural proteins.[15] This cap-independent translation initiation has been well characterized in MLV.[164] Although the use of IRES for the translation initiation in HIV-1 has been somewhat controversial, HIV-1, HIV-2 and SIV possess an IRES within the Gag codon, resulting in translation of its N-terminally-truncated Gag isoforms.[15]

The Gag protein of retroviruses undergoes a number of post-translational modifications (PTMs) which are critical for different stages of the virus life cycle. A well-studied PTM in the context of retroviral assembly is myristoylation.[165] Most retroviral Gag proteins undergo myristoylation except some, including RSV, equine infectious anemia virus (EIAV) and FV. This PTM involves addition of a saturated fatty acid (myristic acid) to the second amino acid glycine in the context of a consensus sequence (reviewed in.[166]) MA trimerization and membrane binding exposes the myristic acid (presented in a sequestered conformation in the MA), which is required for anchoring Gag to the plasma membrane.[17,167] The other common PTM is phosphorylation. For example, phosphorylation of a serine residue in HIV-1 MA regulates binding of the assembling virion to the plasma membrane along with myristoylation.[168] Ubiquitination and SUMOylation are other known PTMs of retroviral Gag proteins, though their role in retroviral assembly is less well understood.[166]

The envelope protein in retroviruses is synthesized from a singly spliced RNA and is processed in the same manner as the cellular membrane and secretory proteins. It is glycosylated in the endoplasmic reticulum (ER), followed by proteolytic removal of the Env leader peptide. Within the ER, it is folded, oligomerized, and then transported to the Golgi, where the cellular protease furin mediates cleavage of the polyprotein, resulting in the generation of SU and TM domains.[17] Translation of the multiply spliced retroviral RNAs uses the classical cap-dependent pathway.[169]

Virus assembly. Retroviral particle assembly initiates with the Gag precursor and requires the gRNA for infectious virus production. The Gag polyprotein provides the basic building blocks of retroviral particle assembly, and each virus particle is made up of approximately 2000–5000 Gag molecules.[45–46] Gag alone is sufficient to make non-infectious virus-like particle (VLPs) that can be successfully released out of the cell.[16] These parti-

cles are morphologically indistinguishable from the ~100 nm diameter immature virions produced by infected cells.[170] Moreover, several retroviral Gags have been shown to assemble to VLP in presence of yeast tRNA *in vitro*[171–174]; however the size of these VLPs are smaller when compared to those formed in cells. In case of HIV-1, it has been shown that adding inositol phosphates or phosphatidylinositol phosphates into the *in vitro* assembly system rescues this defect. Furthermore, MA-mediated binding of Gag to inositol derivatives is required for proper assembly of virus particles.[175]

The Gag polyprotein is sufficient not only for VLP formation, but also for targeting the retroviral assembly to the plasma membrane (reviewed in.[176]) Towards this end, the MA domain of Gag plays a crucial role in determining the subcellular location of viral assembly. The type-C retroviruses (*alpha-, gamma-*, and *lentiviruses*) assemble on the inner leaflet of the plasma membrane, facilitated by a highly basic region in the MA domain and a hydrophobic myristic acid residue (reviewed in.[177]) Together, these regions of MA are often referred to as the "membrane binding domain" (M-domain). Gag is targeted to the membrane areas which are rich in phosphoinositide phosphatidylinositol-4,5-bi sphosphate ($PIP2_)$, sphingolipids and cholesterols, termed 'lipid rafts". Binding of MA to PIP2 results in insertion of the myristic acid to the membrane that plays a role in anchoring Gag to the membrane (reviewed in.[18,178–179]) Interestingly, substitutions of two basic amino acids in the MA domain of HIV-1 Gag resulted in partial relocation of virus assembly to the cytoplasm.[180] In case of HIV-1, further MA trimerization is important in exposing the myristic acid which anchors Gag to the plasma membrane.[167] However, RSV lacks Gag myristoylation and in this case, while multimerization of MA is important for the electrostatic interaction with plasma membrane, it may not be a strict requirement for binding to PIP2.[181]

In case of B- and D-type retroviruses, virion assembly occurs intra-cytoplasmically in the pericentriolar region and pre-assembled immature particles target to the plasma membrane for budding and release.[17,182] An early study in MPMV revealed that a single amino acid mutation in MA changed the assembly location from the cytoplasm to the plasma membrane.[183–184] Later on, similar observations were made for other members of the *betaretroviruses* genus, such as MMTV and Jaagsiekte sheep retrovirus (JSRV) as well as *spumaviruses*[185–191]. The MA domain of these viruses contains a "cytoplasmic targeting and retention signal (CTRS)" which is necessary for the intracytoplasmic assembly of virus particles.[186,190] A single amino acid substitution in CTRS results in relocalization of capsid assembly from the cytoplasm to the plasma membrane, revealing the crucial role Gag plays not only in VLP formation, but intracellular localization of retrovirus assembly.

Gag multimerization is critical for virus assembly which involves not only Gag-Gag, but also Gag-gRNA interactions. The major domain that participates in Gag multimerization is the C-terminal of the CA domain.[18] Additionally, it has been suggested that the basic residues in NC also plays a critical role in Gag multimerization.[192] Within the CA domain, the region that is involved in the Gag-Gag interaction has been termed the interaction domain (I-domain).[18] Several reports have shown that the interaction of RNA with NC is required for efficient CA-CA interaction, probably by exposing the I-domains in CA.[193–196] Furthermore, HIV-1 Gag forms oligomers on the gRNA, is transported to the plasma membrane where higher order multimerization occurs,[196] suggesting that RNA packaging and assembly occurs together.

Release and maturation. The third assembly domain of Gag which facilitates release of the fully assembled immature virions from the plasma membrane is referred to as the late domain (L-domain). In HIV-1, the p6 domain of Gag is considered as its L-domain as its deletion results in accumulation of virus particles attached to the membrane by a stalk.[197] Within the p6 domain, the region required for virion release is located to a PTAP amino acid motif at the C-terminus. This motif binds to the host tumor susceptibility gene 101 (TSG101), a part of endosomal sorting complex required for transport I (ESCRT-I).[198] Another motif, YPXL, located in the p6 domain interacts with Alix/AIP-1 host proteins, a part of ESCRT-II which also participates in virion release, but ESCRT-I remains the dominant complex in this process.[18] Additionally, the virally-encoded Vpu protein of HIV-1 enhances virus particle release from the membrane.[199] In contrast to HIV-1 L-domain, the L-domain of RSV is located near the N-terminus of Gag and contains a PPPY motif that interacts not with ESCRT but with the neuronal precursor cell-expressed developmentally downregulated 4 (Nedd4), a class of ubiquitin ligases, leading to the recruitment of TSG101[19,200]. Interestingly, the late domains of RSV and HIV-1 have been shown to be interchangebale between RSV and HIV-1 Gag irrespective of their positions[200].

Following release, the immature virions undergo proteolytic processing and extensive structural rearrangements, referred to as maturation (Figures 1 & 9). Gag alone or protease-deficient Gag-Pol are able to produce immature non-infectious virions.[20–21] Retroviral proteases belong to the aspartyl protease category and possess the conserved active site Asp-Thr-Gly (DTG; in HIV-1, HIV-2, SIV, FIV, MMTV) or Asp-Ser-Gly (DSG in RSV;.[201–202]) In the immature virus particle, Gag molecules are arranged radially, but following maturation, the CA protein reassembles to differently shaped cores depending upon the virus type. For example, the HIV-1 CA core acquires a cone shape

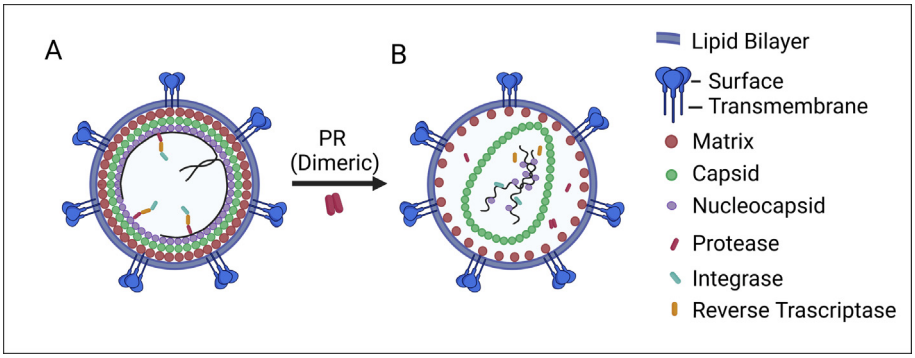

Figure 9. Structural rearrangement of a retroviral particle during maturation. (A) Organization of an immature retroviral particle. The Gag and Gag-Pol polyproteins are in full length form and genomic RNA is in a loose dimer formation. **(B)** Structure of a retrovirus particle after proteolytic maturation. Individual domains of Gag and Pol proteins are released by the activity of the dimeric protease enzyme that the virus particle carries, which result in maturation of the newly formed virion, rendering it infectious. PR, protease.

after maturation with ~250 CA hexamers and 12 pentamers.[203] The full-length CA is composed of an N-terminal domain (NTD) in the exterior of the core and a C-terminal domain (CTD) facing the interior of the core. Initiation of protease activation is a highly regulated process in retroviruses and occurs shortly after the virus buds from the cell. When the protease is part of the Gag-Pol polyprotein, its enzymatic activity is minimal. Full function of the protease requires self-cleavage at the N-terminus and folding of the protein into a stable dimeric form.[18] However, the exact mechanism that triggers protease activation after budding is unclear.

RNA packaging in retroviruses

Specific selection of full-length unspliced gRNA and assembly into the virus particle are highly intricate and coupled processes during retroviral replication termed RNA packaging. This entails selective incorporation of the genomic but not spliced or host cellular RNAs into the virus particles. This process is mediated by the interaction of the Gag polyprotein with particular sequence(s) on the gRNA located at its 5′ region, termed ψ, primarily located between R and 5′ end of the *gag* gene.[23,114,204–206]

Despite the fact that gRNA packaging is a universal step in all retroviruses and the location of the packaging signal is conserved within the retroviral gRNA, no sequence conservation has been found between the packaging signals of different retroviruses.[23,114,204–207] Conventional and novel biochemical techniques have shown that packaging sequences of retroviruses assume higher order structures comprising of different structural motifs.[23,25,205,208–211] In the following sections we describe the current understanding of the molecular mechanisms that take place during retroviral RNA packaging.

RNA binding in the context of Gag. The Gag precursor polyprotein is the basic building block of virus assembly and accounts for 50% mass of the virus particle.[212] Gag, in addition to being the major structural protein, plays a crucial role in the selective packaging of the gRNA.[23,205–206] In particular, the MA and NC domains of retroviral Gag have the ability to bind nucleic acids. Gag binding to ψ in the gRNA is facilitated by the zinc finger domain of the NC domain of the Gag polyprotein.[23,27,205,213] NC contains either one to two evolutionarily conserved, but distinct Cys-His boxes that can sequester Zn^{2+} ions, allowing high affinity NC-gRNA interactions. The Cys-His boxes contain conserved CCHC arrays (C-X_n-C-X_n-H-X_n-C, where C = Cys, H = His, Xn = n number of amino acids) that are variable among different retroviruses. Mutations of the zinc fingers drastically reduce gRNA packaging and lead to the production of non-infectious virions.[214–216] In HIV-1, apart from the zinc fingers in NC, the flanking regions of Cys-His boxes have also been shown to be important for gRNA packaging.[217–218] Interestingly, 15 basic amino acids distributed throughout the NC region can nonspecifically bind to RNA, resulting in the incorporation of cellular RNAs into the virion.[219–220] NMR studies have shown that the structure of MMTV distal zinc finger knuckle is significantly different from that of MPMV, another member of *betaretrovirus* genus[221]. Interestingly, the proximal zinc finger domains of MMTV fold similarly to those of HIV-1 and MLV.[221]

It is well known that the MA domain of several retroviruses also binds to RNA, although the precise role of this interaction in the viral life cycle remains largely unclear.[222–223] For example, it has been shown in HIV-1 that RNA-binding to MA was necessary for Gag oligomerization when two-thirds of NC was deleted from Gag.[193] Similarly, the basic amino acids in either MA or NC are suffi-

cient to package RNA without significantly reducing virion production. Moreover, in bovine leukemia virus (BLV), *in vitro* studies have shown that specific selection of gRNA is conferred by the MA protein, while NC binds non-specifically to any RNA.[224–225]

Finally, some studies have shown that the p6 domain of HIV-1 is important in selective encapsidation of gRNA by discriminating between gRNA and spliced viral RNAs since its deletion results in reduced affinity to binding gRNA.[195,226–227] However, certain other studies suggest that the p6 domain is not required for efficient gRNA binding.[228–229] In the prototype *betaretrovirus*, MPMV, a KKPKR sequence located within the pp24 domain of $Pr78^{Gag}$ has been implicated in playing a role in viral RNA packaging, though direct binding of pp24 to RNA has not yet been established.[230] It has been proposed that the presence of these basic residues enhances RNA packaging mediated by NC. Cryo-electron tomography studies have shown that the RKK residues at the C-terminus of MPMV CA has a potential role in nucleic acid binding, probably not by directly interacting with RNA, but rather by stabilizing the Gag-viral RNA complex, or possibly due to structural rearrangement of the NC region which is required for packaging.[231] Replacing the RSV NC domain with that of MLV or HIV-1 NC with that of MMTV NC, did not completely abrogate RNA packaging efficiency; rather, up to 50% of RNA packaging was retained, suggesting the involvement of other Gag domains in gRNA packaging.[220,232–233] These studies clearly demonstrate that the packaging of gRNA is an intricate and multifaceted phenomenon that occurs in the context of the whole Gag polyprotein, especially in the case of HIV-1.

Interestingly, in addition to Gag, a recent study has suggested that Gag-Pol also may play a role in RNA packaging in HIV-1. This was shown by mutating ribosomal frameshift signal to abolish frameshifting (i.e., express only Gag) which resulted in ~50% reduction in RNA packaging. This phenotype was reversed to wild type levels when Gag-Pol was provided in *trans*.[162]

The Ψ RNA: Unpaired purines are the key players. As mentioned earlier, the 5′ region of retroviral gRNA assumes higher order (secondary and tertiary) structure with several base-paired and unpaired nucleotides, creating stem loops and internal loops. Many of these structural motifs have been shown to be critical for various steps in the virus life cycle, including reverse transcription, transcriptional activation, splicing, translational regulation, dimerization, and packaging (Figure 2). Initially, the core encapsidation signals/determinants (CES) in HIV-1 were mapped to a stretch of 159 nts within the 5′ end of the gRNA.[234] This region harbored SL1-SL3 and a few nucleotides of *gag*, but excluded the role of TAR, polyA and PBS in RNA packaging.[234] However, a later study showed that the polyA loop is critically required for efficient packaging; but direct interaction with full-length HIV-1 Gag could not be identified.[235–236] Hence, it is possible that other regions within the gRNA are involved in recruiting Gag or stabilizing the Gag-gRNA interaction, rather than acting as direct binding sites. In addition, multiple regions within the packaging determinants have been shown to be required for optimal gRNA packaging by Gag. For example, in the case of HIV-1, studies have shown that TAR, the polyA hairpin, basal region of PBS, internal loop within the SL1 and the SL3 apical loop are required for its efficient packaging by $Pr55^{Gag}$.[237–240]

Furthermore, several studies in HIV-1 have shown that purine-rich regions, particularly guanosines (G) are involved in interaction with Gag.[30,33,241–242] It is also interesting that 'GA' rich motifs in 5′ UTRs are involved in mainting the ratio of unspliced RNAs to singly and multiply spliced RNAs in the case of HIV-1, by binding to cellular SFPQ/PSF protein.[243] However, this purine motifs are distinct from the Gag interacting nucleotides and functions independently of packaging.[243] Back to RNA packaging, a GGAG tetraloop in HIV-1 SL3 present downstream of the major SD has been shown to bind to the NC domain (Figure 10).[26,244–245] A recent study using NMR spectroscopy and isothermal titration calorimetry has identified a weakly base-paired [UUUU]:[GGAG] helix in the SL3 stem as the initial high affinity NC binding site.[246] In addition to NC, the p6 domain of HIV-1 has also been shown to be important in selective encapsidation of gRNA by discriminating between gRNA and spliced viral RNA.[247] Therefore, it is reasonable to suggest that the specificity and number of nucleotide interactions could vary in the full-length conformational context of Gag. For instance, it was noted that in MLV the full-length Gag binds more selectively to the CES than NC.[248] Using HIV-1 full-length $Pr55^{Gag}$, results of *in vitro* binding and footprinting experiments have shown that SL1 is the primary $Pr55^{Gag}$-binding site and the purine rich internal loop (G//AGG) in SL1 is a key determinant of $Pr55^{Gag}$ interaction during RNA packaging.[30,226] Apart from these *in vitro* studies, an earlier study using transfection-based packaging experiments have shown that mutating SL1 or SL3 alone results in mild reduction of RNA packaging efficiency, but combined mutations in SL1 and SL3 cause severe packaging defects.[249] Similarly, another study based on infected cells under competitive packaging conditions showed that mutliple 'G's act synergistically to facilitate Gag binding, rather than a single specific Gag binding site.[33]

Similar to HIV-1, in HIV-2, a 5′ GGRG 3′ motif located upstream of the DIS was shown to be important for RNA packaging and has been suggested as its Gag binding site.[250] A recent study

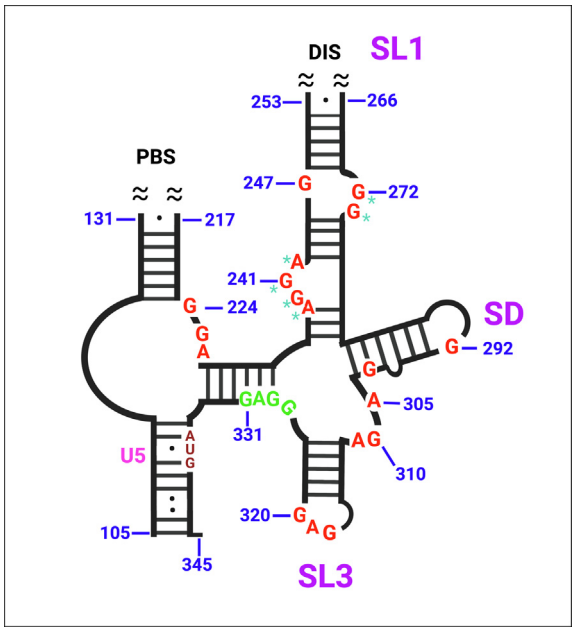

Figure 10. Purines that interact with full-length Gag or NC during HIV-1 genomic RNA packaging. Schematic of the branched multiple hairpin (BMH) structure of HIV-1 packaging signal RNA showing purines that interact with Gag or NC identified by footprinting[30] and reverse footprinting[25] approaches, respectively. Nucleotides with cyan asterisks are the suggested primary Gag/NC binding sites by both of these studies. The 5′ GGAG 3′ nucleotides (highlighted in green) are proposed as the high affinity nucleocapsid interacting sites identified by NMR and isothermal titration calorimetry.[246] Note: the structure shown here is reported by Wilkinson et al., 2008.[25] Abd El-Wahab et al., 2014[30] reported a structure without U5:AUG interactions. The nucleotides identified by Ding et al., 2020 [246] were based on a three-way junction structure in which the SD containing SL2 is sequestered by long-range interactions.[26] Only purines that showed strong chemical protection in the presence of Gag by footprinting or were deemed highly reactive by reverse footprinting are shown in this figure.

also recognized the same residues (the GGAG motif) as the primary nucleotides for packaging in HIV-2.[35] In case of SIV, a stretch of unpaired purines has been identified as crucial for RNA packaging, while stretches of other unpaired purines may also act synergistically to facilitate efficient RNA packaging.[251] The preference of NC to bind to G residues has also been demonstrated in HTLV-1 by UV crosslinking experiments.[252] But analysis of protection by bound NC from SHAPE reagents showed that only 4 'G' residues among the total 9 were being protected.[252]

MMTV, a *betaretrovirus*, exhibits a slightly different mechanism in which the PBS (that resides within stem-loop 2, SL2) and a stretch of unpaired, single-stranded purines (ssPurines) that reside within the bifurcated stem-loop 4 (SL4), are not mutually exclusive for RNA packaging.[32] Within ssPurines, a GGAG motif at the 5′ end and AG at its 3′ end are necessary for efficient packaging (Figure 11(A)). While the involvement of PBS in retroviral RNA packaging has not been observed earlier, in the case of MMTV, it functions as a high affinity Gag binding site.[32] Further studies are required to narrow down the exact role of PBS in MMTV RNA packaging, including its inter- and intra-molecular interactions that regulate RNA structures (Figure 11 (A).[32]

It is noteworthy that maintaining the native structure of MMTV ssPurine loop within SL4 and PBS within SL2, the two Gag binding sites, were not sufficient to drive MMTV gRNA packaging in many mutants which lost other structural motifs (for e.g., long range interaction (LRI) and SL3;.[32] Although Gag could bind to these RNAs with high affinity *in vitro* (because of intact ssPurines and PBS), these mutant RNAs were not packaged in virus particles, indicating the importance of the overall structure (secondary/tertiary and pseudoknots) and/or other regions of the 5′ UTR.[32] This suggests that the initial high affinity Gag binding is not enough for RNA packaging and maintaining a global 5′ UTR structure is important, which perhaps plays a role in downstream assembly steps. It is also possible that the distance between and organization of different Gag binding sites in its tertiary conformation are crucial for Gag multimerization on gRNA. It will be interesting to assess this possibility in other retroviruses such as HIV-1.

The same involvement of ssPurines as being important for RNA packaging and Gag binding has been observed in other simple retroviruses. For example, in the case of MPMV, a simple retrovirus in the *betaretrovirus* family, the $Pr78^{Gag}$ binds to a stretch of ssPurines and a 'GU' rich region downstream of the major splice donor (Figure 11 (B); 11 out of 13 (84.6%) of the $Pr78^{Gag}$-interacting nucleotides were purines; the remaining 2 were U residues).[34] It was further shown that these two elements (ssPurines and GU-rich region) act redundantly in their Gag binding to facilitate specific encapsidation of gRNA since removal of the GU-rich element upon splicing disrupted the structure of the ssPurine region (Figure 11(B))[34].

In case of the prototypic simple retrovirus MLV, the core packaging determinants have been mapped to ~100 nt within the 5′ UTR comprising three stem loops (Figure 12;[253–256]). Structural studies using the NC domain of MLV Gag, and the core encapsidation signal (CES) RNA have identified a UCUG motif which is base paired in the monomeric conformation of RNA and unpaired following dimerization (Figure 12). Moreover, the interaction of 'G' residue within this UCUG motif with the zinc knuckle has been well-characterized.[253,257] Footprinting experiments using MLV NC and full-length

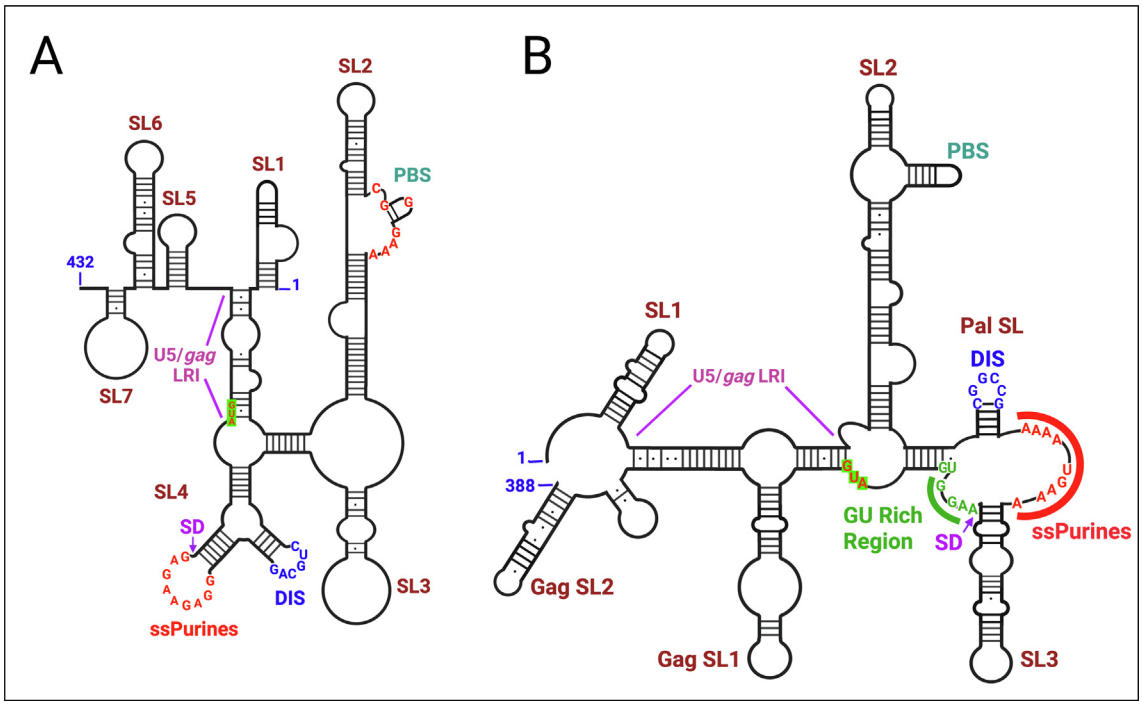

Figure 11. Secondary structures of packaging signal RNAs and nucleotides important for gRNA packaging of two different members of *betaretroviruses* (A) Nucleotides that interact with $Pr77^{Gag}$ of the mouse mammary tumor virus (MMTV) identified by SHAPE-footprinting.[32] Both single stranded purines (ssPurines; highlighted in red) and the primer binding site (PBS; also highlighted in red) are critical for MMTV gRNA packaging as mutating either one completely abrogates RNA packaging. **(B)** Nucleotides that interact with $Pr78^{Gag}$ of Mason-Pfizer monkey virus (MPMV) identified by SHAPE-footprinting.[34] Both the MPMV ssPurines (highlighted in red) and a 'GU' rich region (highlighted in green) act as Gag binding sites in a redundant manner, as deleting either one does not affect MPMV gRNA packaging significantly. SL; stem-loop, PBS; primer binding site, DIS; dimer initiation site, LRI; long-range interaction, Pal; palindromic.

Gag which required using a longer RNA compared to the NMR study identified two UCUG-UR-UCUG motifs, both of which were required for high affinity NC binding (Figure 12[248]). Interestingly, within these UCUG-UR-UCUG motifs, the 'G's were shown to be crucial for *in vitro* binding by both NC and Gag.[248]

In RSV, another simple retrovirus, only a 82-nt long region in the 5′ UTR and no sequences in *gag* has been implicated in gRNA packaging.[258–259] NMR analysis of this region has revealed a small four-way junction structure present that has the ability to bind the zinc knuckles in RSV NC. More specifically, the N-terminal zinc knuckle preferentially interacted with a UGCG tetraloop on SL-C present in this structure with the second 'G' making an extensive contact with NC, while an interaction of the C-terminal zinc knuckle with an 'A' was also required for high affinity binding.[258–259].

Discrimination of unspliced RNA from spliced RNAs. Although numerous studies have shown how unspliced gRNAs are selectively packaged into the newly forming virus particles, only limited studies have shown the mechanisms that prevent spliced RNAs from their encapsidation into the assembling virions. These studies have shown that several potential mechanisms could be involved in this process (Figure 13), such as the packaging signal in spliced RNAs could either be removed (Figure 13(A)), or architecturally disrupted (Figure 13(B)), as has been observed in MPMV and MMTV[32,34]. The former possibility is applicable to HIV-1 based on a three-way junction structure identified by NMR in CES RNA.[246] This study identified a [UUUU]:[GGAG] helical stem that acts as the initial high affinity binding site for NC. In this case, the 'GGAG' stem is located downstream of the SD and thus could be spliced out. However, the case in HIV-1 is a lot more complex. For example, other studies have identified an internal loop in SL1 and a region upstream of SL1 that are the main packaging determinants since they serve as the primary Gag binding sites; interestingly, these sequences are present both in the unspliced and spliced RNAs.[25,30] It is also possible that HIV-1 precludes the spliced mRNAs from the virus particle due to a 45-nt sequence downstream of the Gag AUG which is present only in unspliced RNAs. This sequence acts as positive regulator for packaging to compensate the negative regulatory effect exerted by sequences upstream of Ψ (Figure 13(C))[30].

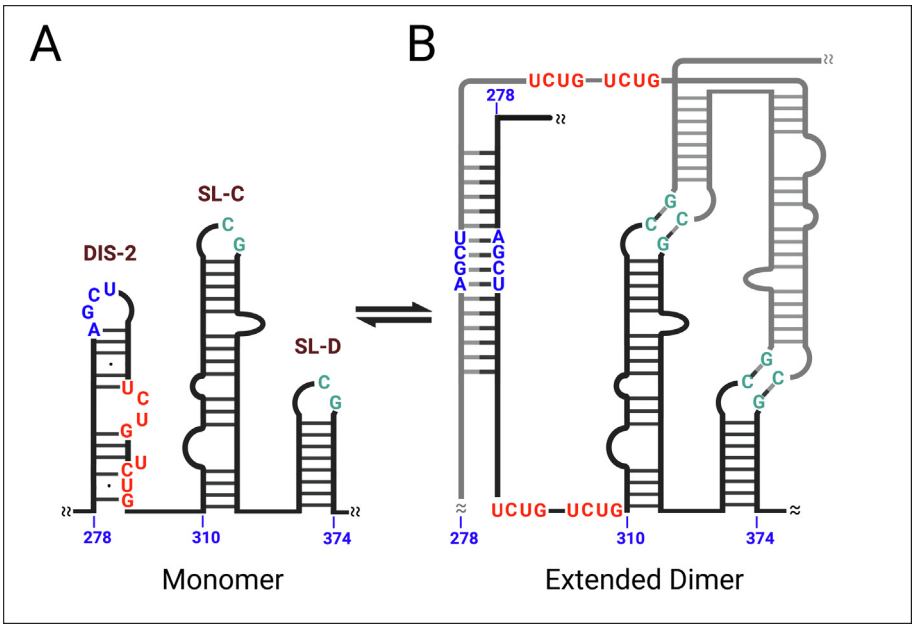

Figure 12. Dimerization induces structural rearrangement of the core encapsidation signal (Ψ^{CES}) of MLV gRNA (A) Secondary structure of the monomeric MLV Ψ^{CES}. The four nucleotides involved in the initial kissing loop interaction within the 16-nt long palindromic sequence of DIS-2 are shown in blue. Additionally, two other dinucleotides on the non-palindromic apical loops of SL-C and SL-D that participate in kissing-interaction are shown in green. The two runs of Gag-interacting nucleotides "UCUG" are shown in red which are base-paired and not available for interacting with Gag in the monomer. **(B)** The extended dimer structure of the MLV Ψ^{CES}.

Since all the unpaired guanosines involved in Gag binding and sequences that form U5:AUG interactions were maintained and the structure of the key mutants was not disrupted in their experiment, it was suggested that the negative effect is a steric hindrance that prevents binding of HIV-1 Gag to the lower part of SL1.[30] However, disruption of the Ψ RNA structure in the presence of downstream sequences from spliced RNA is another possibility that cannot be excluded from these experiments.

Another possible explanation for selective packaging of unspliced RNA that has been investigated is that several unpaired purines, both upstream and downstream of splice donor site, act synergistically in Gag binding and loss of multiple Gs downstream of the splice donor site in spliced RNAs greatly reduces Gag binding affinity to spliced RNAs.[33] In this scenario, it is not necessary that a significant structural change in the region upstream of splice donor site occurs because the loss of downstream Gag binding nucleotide should reduce the binding affinity *per se* (Possibility 1, Figure 13(D)). However, the structure of HIV-1 spliced RNAs must be elucidated to completely understand how the spliced RNAs are prevented from being packaged. As mentioned elsewhere in this review, the stability of polyA stem is important for RNA packaging; thus, it is possible that the structure of polyA stem and other critical elements, such as the DIS in spliced RNA, are altered preventing their packaging. In the case of HIV-2, both the primary Gag binding nucleotides (i.e., mutation of these nucleotides reduced packaging efficiency to 50%) and secondary Gag binding site (mutation of this site when combined with mutated primary site resulted in 70% reduction of packaging efficiency) are located upstream of the major SD.[35] Mutating multiple unpaired Gs downstream of the SD (categorized as "tertiary sites") did not greatly reduce the packaging efficiency.[35] Thus, it is reasonable to suggest that spliced RNAs are prevented from encapsidation in HIV-2 by the combined effect of structural changes which result in base-pairing of Gs that interact with Gag upstream of the splice donor site, instability of other structural elements such as DIS, and the loss of Gag interacting Gs downstream of the splice donor site (Possibility 2, Figure 13(D)).

Ψ RNA acts as the nucleation point for Gag multimerization. The retroviral Gag proteins have the ability to assemble and form VLPs even in the absence of ψ-containing viral gRNA. Several studies have suggested that at physiological salt concentrations, HIV-1 Gag can bind to both ψ-containing and non-ψ RNAs with almost equal affinity, including both specific and nonspecific interactions (reviewed in.[207]) However, nonspecific binding was significantly reduced when the ionic strength of the experimental conditions was increased, thereby identifying only the specific high affinity binding sites.[30,32,34,260] Hence, it has been

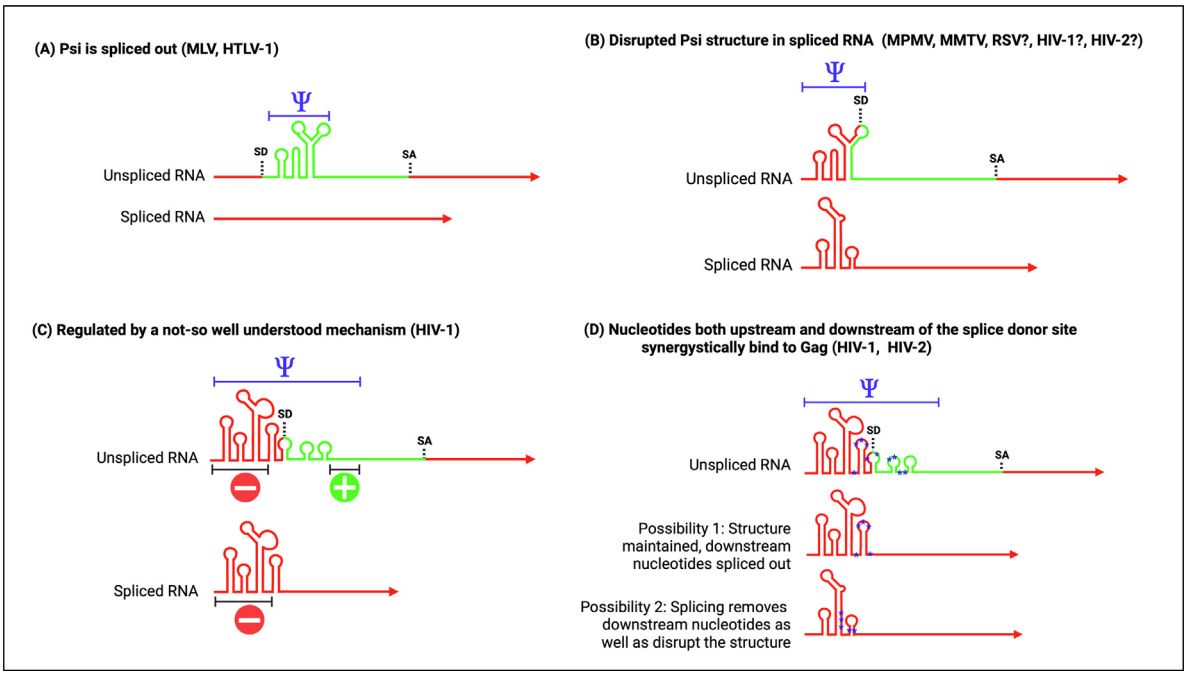

Figure 13. Different models of selective packaging of unspliced RNA over spliced RNAs. (A) The packaging signal (Ψ) in this model is located downstream of the splice donor site (SD) which is removed due to splicing. **(B)** The Gag binding nucleotides are located upstream of the SD in an unpaired conformation. After splicing, the conformation of Ψ is disrupted and the nucleotides that bind Gag are base paired. **(C)** This model suggests that the 5′ end of Ψ shown in red and composed of TAR, poly A and/or PBS domain(s) of HIV-1 RNA, exerts an inhibitory effect on Gag binding. On the other hand, a short region of 45-nt at the 3′ end of Ψ suppresses this effect (shown in green). **(D)** In this model, the nucleotides both upstream and downstream of Ψ participate synergistically in Gag binding. Removal of nucleotides downstream of the splice donor site significantly reduce Gag binding. Blue asterisks represent nucleotides that interact with Gag during RNA packaging. SA, splice acceptor.

suggested that Gag has the potential to bind to all RNAs, but nucleation and assembly following Gag-RNA binding occurs more rapidly on ψ-containing RNAs than nonspecific RNAs, probably in a two-stage process.[207]

A study based on labelling Gag and gRNA in cells and quantifying the empty and gRNA-containing virus particles released (single virion analysis) showed that when Gag was expressed in cells at levels similar to those in cells containing one provirus, the presence of HIV-1 RNA greatly enhanced viral particle production.[261] However, when Gag was overexpressed, this RNA-specific enhancement disappeared, suggesting assembly of Gag at higher concentrations did not require viral RNA.[261] As previously mentioned, this study also underlines the fact that RNA with specific packaging signal acts as the "nucleation point" for Gag assembly. A recent study from the same group has shown that plasma membrane anchoring and Gag-Gag multimerization on viral RNAs are critical steps during this nucleation process and the nucleation event occurs at the plasma membrane.[229] This study used a complementation system that functionally separated assembling and RNA binding properties of Gag, which avoids challenges due to mutations in Gag to assess RNA packaging since they can lead to assembly defects.[229] Another study based on single molecule localization microscopy proposed a simplistic model of gRNA interaction with the Gag polyprotein, preceding to Gag multimerization and assembly at the plasma membrane.[196] It showed that the assembly process begins in the cytoplasm where a small number of Gag molecules interact with the gRNA, which is then trafficked to the plasma membrane, the place at which higher order multimerization of Gag takes place. The study clearly demonstrated that the initial interaction with gRNA is the key step for the entire assembly process since the ΔNC-Gag freely migrated rather than clustering.[196] This observation is previously supported by other studies where the deletion of NC resulted in delayed particle production without any morphological differences. Interestingly, the delayed particle production can also be explained by the weaker affinity of nucleic acid interacting capacity of MA.[194] Thus, Gag-RNA interactions play an essential role in Gag assembly, using ψ as the point of nucleation.

Nucleus as the initial site of gRNA selection. The intracellular location where the initial interaction of

gRNA with Gag occurs remains a fundamental question in the retroviral life cycle. It has been assumed for a long time that this interaction occurred either in the cytoplasm or at the plasma membrane. Many retroviral Gag proteins, including RSV, HIV-1, MMTV, MPMV, FIV and MLV have been identified in the nucleus, raising the possibility that the initial selection of gRNA occurs in the nucleus. RSV has been considered the prototypic model for studying the mechanism of Gag nuclear trafficking. These studies have shown that nuclear trafficking of RSV Gag protein plays an important role in the efficient packaging of RSV RNA. Recently, it has been established that within the nucleus, the site of active viral RNA transcription is the place where initial interaction between gRNA and RSV Gag takes place.[262–263]

Despite the fact that several studies have explored the nuclear trafficking of HIV-1 $Pr55^{Gag}$, the precise location where the initial interaction between HIV-1 Gag and gRNA occurs still remains largely controversial. Some studies have shown that the nuclear localization signal (NLS) of HIV-1 is part of the MA domain of $Pr55^{Gag}$, the same domain involved in recruiting Gag-RNA complexes to the plasma membrane.[264–265] Deletion of NLS results in accumulation of gRNA in the nucleus and reduced RNA packaging efficiency, suggesting that the initial Gag-RNA interaction starts in the nucleus,[266] while another study observed the co-localization of RNA and Gag at the perinuclear region.[267] However, live cell imaging studies have shown that HIV-1 Gag-RNA complexes are not associated with the nuclear envelope.[268] A recent study proposes that HIV-1 Gag interacts with newly transcribed RNA, leading to the formation of ribonucleoprotein (RNP) complexes in the nucleus, though this study does not confirm that nuclear Gag-RNA RNP complex formation is necessary for RNA packaging.[269]

Role of nuclear export pathways in RNA packaging. In case of the simple retrovirus MLV, using specific nuclear export pathway blockers, it has been shown that unspliced RNAs use both the CRM1 and NXF1 export machinery for their nuclear export[270]. Interestingly, it was further shown that unspliced RNAs exported through the CRM1 pathway were selected for packaging and those exported through the NXF1 were coupled to translation[270]. However, the exact mechanism that drives the unspliced RNA to two different export pathways and whether the nuclear export pathway itself or pre/post-exporting events mark the RNA for their cytoplasmic fate remains unknown.[270]

Binding of HIV-1 Rev protein to the RRE is a critical requirement for the nuclear export of the unspliced and singly spliced viral RNAs through the CRM1 export pathway. In addition to its nuclear export function, RRE has also been shown to play an important role in RNA packaging.[271–273] Additionally, binding of Gag to RRE located in the 3′ UTR has also been demonstrated, although a role of this interaction in packaging remains elusive.[274] It has further been suggested that the Rev-RRE interaction is not directly involved in RNA packaging; rather it is involved indirectly by influencing the nuclear export pathway[275]. Interestingly, replacement of RRE with CTE of MPMV that mediates the nuclear export via NXF1 pathway, can efficiently complement the nuclear export function of Rev-RRE[152,240,276–277]. However, a recent study showed that the RRE-containing RNAs packaged more efficiently when compared to CTE-containing RNAs under competitive experimental conditions.[273] Hence, further studies are required to elucidate the precise role of the Rev-RRE system in the process of retroviral RNA encapsidation.

Role of assembly intermediates in gRNA packaging. Studies pertaining to the interaction of unspliced gRNA and HIV-1 Gag as assembly intermediates have also been carried out using the highly sensitive and specific proximity ligation assays (PLAs) that can detect protein-RNA interactions with single-molecule resolution *in situ* in unmodified cells. These assays have demonstrated that the Gag-RNA complex is first present in the second assembly intermediate (80S, not in the ~10S soluble Gag), suggesting that association of HIV-1 Gag and unspliced HIV-1 RNA occurs within host RNA granules that may act as assembly precursors[278]. These granules also contain non-translating mRNAs and cellular proteins ABCE1 and DDX6 which have been shown to facilitate capsid assembly[279]. However, further studies are required to confirm whether host-derived RNA granules are the initial packaging complexes or not. Also, it is not clear whether Gag makes direct contact with gRNA in these complexes. Similar to HIV-1, assembly intermediates derived from ABCE1- and DDX6-containing granules has also been detected in FIV[279]. Another host protein, Staufen1, has been shown to interact with HIV-1 Gag and found to colocalize in the viral RNP complex, suggesting its role during HIV-1 gRNA encapsidation[280–282]. Taken together, these data indicate that host-derived factors in the RNP complex may play a role in RNA encapsidation and virion assembly during the retroviral life cycle. However, a recent study suggests that a major fraction of these complexes may not represent the on-pathway (leading to completely assembled virus particles) assembly intermediates; instead, they represent lower oligomers of Gag (monomeric and dimeric Gag complexed with ribosomes), making it difficult to conclude that the previously-reported Gag containing host granules represent true assembly intermediates.[283] Hence further studies are required to understand whether assembly intermediates and host-granule derived proteins play a

role in gRNA encapsidation and assembly processes.

Retroviral RNA dimerization

Retroviral gRNA dimerization is an essential step in the virus life cycle during which gRNAs are packaged as dimers, non-covalently joined through their 5′ ends. The first observation of the existence of a retroviral RNA in a dimer form in the virus particle was made in RSV.[284–285] Better understood in HIV-1, genomic RNA dimerization is initiated by the palindromic sequence DIS which allows interaction between the 5′ ends of the two RNA genomes.[286–287] Following initial kissing-loop interactions mediated by DIS, a dimer linkage structure (DLS, or extended-duplex; Figure14) conformation forms involving base pairing of 50 to a few hundred nucleotides.[288] Further studies using similar *in virio* assays coupled to mutational and native northern blot analyses have led to proposals for more refined 2D and 3D models of the DLS element with a unique thermodynamically stable pseudoknot at the core of the structure along with the possibility of long-range interactions between a 5-nt sequence upstream of U5 and a 5-nt-sequence downstream of PBS that occur either intra-or intermolecularly.[289] The RNA dimers that are formed soon after the formation of the kissing-loop complex are often considered as weak dimers. Dimeric RNAs isolated from protease-deficient HIV-1 particles have lower thermal stability and different conformations, suggesting that further stabilization of the kissing-loop complex is achieved by the nucleic acid chaperone activity of NC, which is generated as a result of proteolytic maturation of Gag.[290–291] Furthermore, it has been shown that HIV-1 RNA stable dimerization occurs at the plasma membrane in the presence of Gag and is initiated by two Gag-RNA complexes rather than two RNA monomers.[292] In contrast, in RSV and MLV, dimerization occurs in the nucleus, and in MLV, dimerization occurs co-transcriptionally.[293–295] Whether Gag has any role to play in RSV or MLV genome dimerization like HIV-1 remains to be determined.

In vitro dimerization studies and analysis of virion RNAs have been widely used to understand the mechanism involved during gRNA dimerization. Several studies have shown that the *in vitro* derived RNA dimers have similar characteristics with those derived from virions. Furthermore, *in virio* RNA structural studies reveal that RNA structures exist in virus particles which are not significantly rearranged when compared to their *in vitro* conformations.[25,210] This makes *in vitro* dimerization assays a reliable and simple technique to study retroviral gRNA dimerization.[296–297] The presence of magnesium ions (Mg^{2+}), Gag, or NC

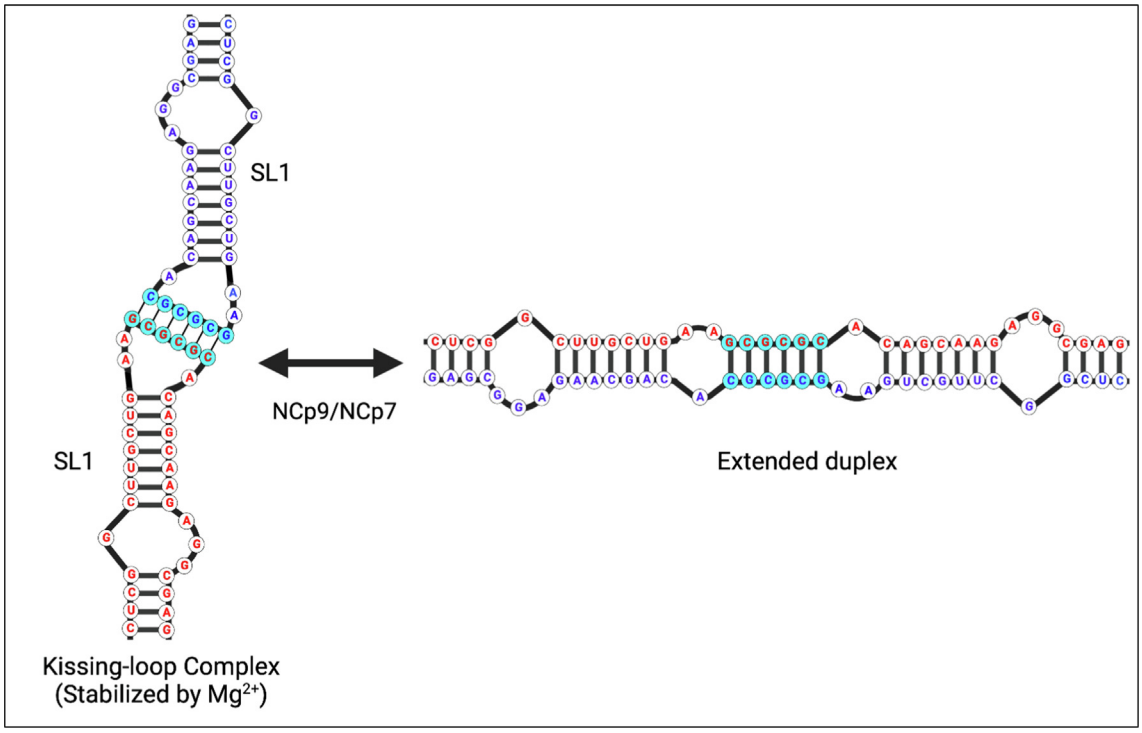

Figure 14. Transition of initial kissing-loop complex to the extended duplex during HIV-1 RNA dimerization. The kissing loop complex is stabilized by Mg^{2+} ions which bind to the internal loop of SL1 and the purines that flank the dimerization initiation site (DIS; highlighted in cyan; GCGCGC). The chaperon activity of nucleocapsid (NC) converts the kissing-loop complex to the extended-duplex formed by two 35-mer SL1s.

strongly influence dimer formation *in vitro*.[227,296–299] While Mg^{2+} ions have been implicated in binding to an internal loop where the HIV-1 $Pr55^{Gag}$ also binds, the purines flanking the DIS stabilize the kissing-loop complex.[300–301] Furthermore, the chaperone activity of Gag/NC converts Mg^{2+}-stabilized dimer to a more stable extended-duplex (Figure 14;[302,290,303]). However, it has been suggested that the chaperon activity of NC is higher than that of full-length Gag[304] and among different NC maturation intermediate products, NCp9 and final NCp7 (fully matured) are more efficient in converting the kissing-complex dimer into a more stable extended-dimer (Figure 14;[305–306]).

In case of ALV, MPMV, MMTV and HIV-1, a 6-nucleotide GC-rich region has been identified which functions as the DIS, while in FIV, a 10-nt long palindromic sequence acts as the DIS (Figure 15;[208–209,307–309]). In the case of MLV, multiple stem loops are involved in dimer initiation,

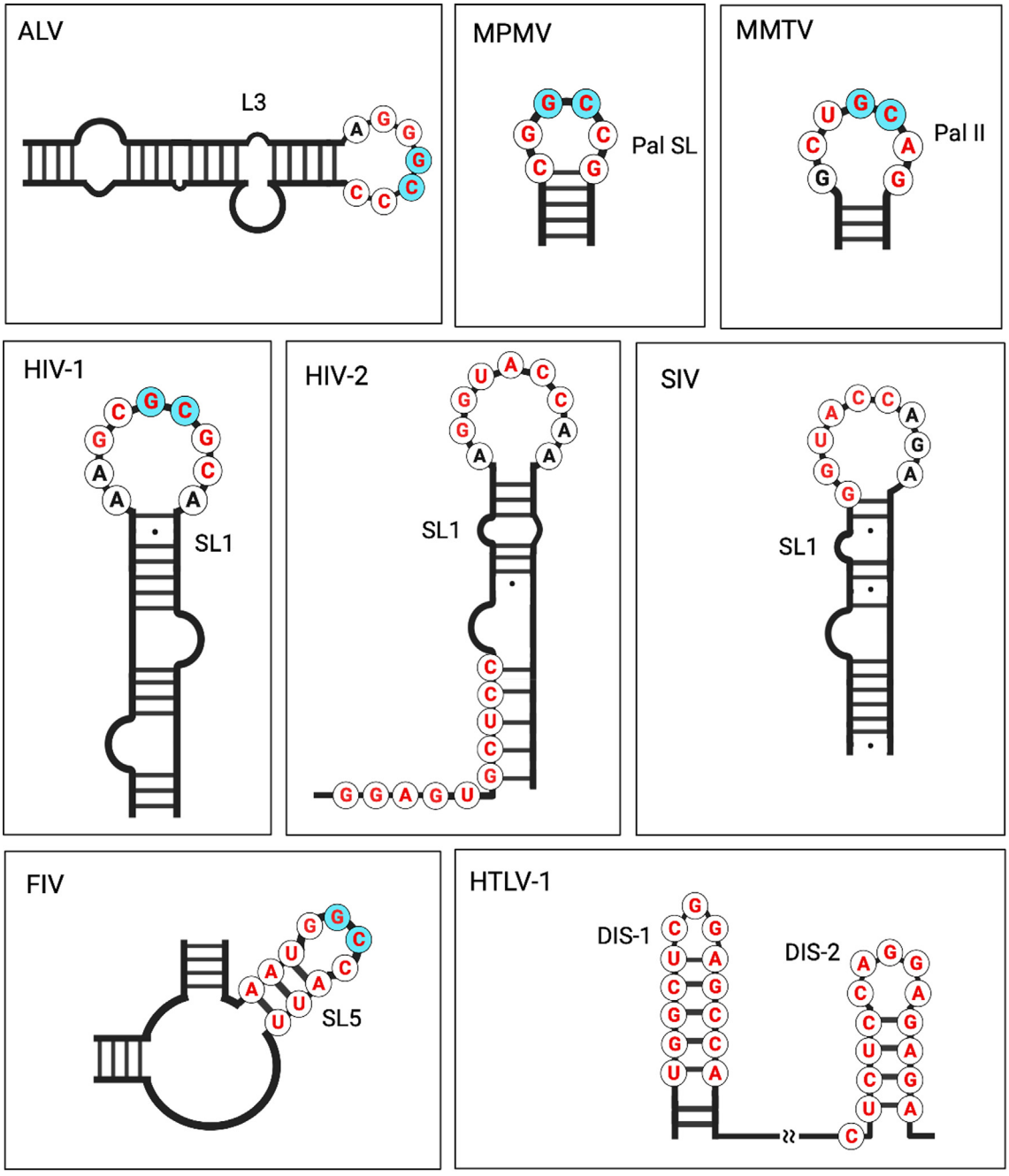

Figure 15. The dimer initiation site (DIS) of different retroviruses. DIS is the region that participates in the initial kissing-loop interaction between retroviral gRNAs. The "GC" dinucleotide highlighted in cyan is highly conserved among different strains of the specified retroviruses. However, this "GC" dyad is not observed in other retroviruses, such as HIV-2, SIV, or HTLV-1.

namely DIS-1 and DIS-2, both containing palindromic sequences.[310] It has also been shown that 2 sets of non-palindromic dinucleotides located in the apical loop of two different stem loops interact with each other and contribute to dimerization initiation in MLV.[311–312] Interestingly, a GC dyad in the DIS is highly conserved among these viruses (blue highlight in Figure 15). Two 13-nt and 14-nt long imperfect palindromes have also been identified as the DIS in the case of HTLV-1 (Figure 15,[252,313–314]); however, the exact mechanism of dimerization initiation is still unknown for HTLV-1. A recent study has revealed the presence of a bipartite DIS in RSV 5′ UTR: one is a 6-nt GC rich palindrome, while the other one is not a palindrome. Both parts of the DIS are functionally redundant and neither of these regions are necessary for virus replication in cell culture.[315] This observation is similar to that made in HIV-1 in which DIS has been shown to be dispensable for virus replication in PBMCs,[316] suggesting role of other elements during dimerization.

The process of DIS-mediated dimerization and gRNA encapsidation are highly interlinked, as evident by defective packaging of DIS mutants.[317] In HIV-1, the gRNA containing an additional region involved in DLS formation has been shown to be successfully packaged as a monomer, indicating that DLS formation is important for packaging, irrespective of whether it is via an intra-or intermolecular interaction.[318] The proximity or overlap between the DIS and the packaging sequences also suggests the existence of an interplay between dimerization and packaging of gRNA.[319–320] The exact mechanism in the selective encapsidation of RNA dimers by Gag remains unknown; however, it is assumed that following dimerization, conformational changes result in the proper exposure of the Gag binding sites within the gRNA which otherwise remain obscured in the monomeric RNAs.[253] This scenario is best understood in the case of MLV, in which dimerization induces a conformational change that exposes high affinity NC binding UCUG motifs which are base-paired in the monomeric RNA state.[253] Ensuring the encapsidation of RNA in a dimeric state has several implications in retroviral life cycle. For example, dimeric RNA: i) greatly enhances the rate of reverse transcription,[287,321] ii) maintains the functional genome during replication by repairing single-strand nicks or other RNA damage by template switching,[322] and iii) allows recombination due to co-packaging of heterodimer (of two different genotypes; in case of HIV-1), leading to enhanced genetic variability that contributes to increased replicative fitness and drug resistance of retroviruses.[323–326]

Heterogeneity in retroviral RNA transcripts: Functional implications

Over the years, numerous *in vitro* studies have suggested that an RNA-conformational switch exists in HIV-1 that regulates the fate of the gRNA, whether to undergo dimerization and packaging or translation. These experiments have revealed the existence of two structures of HIV-1 gRNA with different functional implications: the long-distance interaction (LDI) and the branched multiple hairpin (BMH).[327–329] The LDI conformation involves base-pairing of sequences several hundred nucleotides apart. The DIS in this conformation is base-paired with the polyA region, thus not well-suited for dimerization. On the other hand, in the BMH conformation, several stem loops exist, including SL1 (containing the DIS) that exposes Gag binding sites, thus facilitating both dimerization and packaging. Structural studies based on NMR on short RNAs have demonstrated that either the Gag AUG or the DIS in HIV-1 base-pairs with U5; the U5:Gag AUG interaction favors the dimerization competent structure, and it was proposed that this structure does not serve as translating mRNAs.[27] On the other hand, the structure with U5:DIS interaction was suggested to be involved in translation.[27] However, later it was proposed that in the case of HIV-1, two different pools of unspliced RNAs exist with two functional implications: one acts as mRNA and the other as gRNA for packaging.[327] Despite the different structures that have been proposed for the dimerization and packaging-competent HIV-1 RNAs, the existence of TAR, polyA, SL1 and SL3 (initially referred as ψ) are common features among all these structures.[25,27,30,210,327,329–330] Moreover, all these studies are in agreement in terms of the U5:AUG long-range interactions except a few.[30,210] It is also worth mentioning that similar U5-Gag long-range interactions have also been reported for *betaretroviruses*[208–209] as well as other lentiviruses, such as FIV.[331]

Apart from the *in vitro* studies, several groups over the years have used *in vivo* studies using HIV-1-infected cells to determine when the full-length viral RNA is used as mRNA or as a substrate for packaging into virions. An earlier study showed that prior translation was not required to generate packagable RNA and inhibiting translation rather increased the utilization of RNA for packaging in HIV-1.[332] This study also suggested that only a single pool of RNA exists that functions interchangeably as mRNA and packagable RNA, with possibly the translational machinery preventing the RNA from being packaged.[332] This conclusion was supported by another study that showed that inihibition of transcription resulted in reduction of RNA packaging and Gag-Pol expression comparable to cyoplasmic level of RNA, suggesting RNAs are not functionally separated in cells.[333] Poon et al., further supported these findings by suggesting that HIV-1 RNAs are packaged in *cis* where the newly synthesized Gag packages its cognate RNA; thus, this model supports the theory of a single pool of RNA with dual function.[334] However, it is possible that the reduced packaging

of RNA mutants deficient in Gag translation observed in this study was due to the specific mutations introduced in and around the Gag initiation codon and not due to its inability to translate Gag, as shown by another study which documented that HIV-1 packaging occurs mainly in *trans*.[335] A more recent study has shown that altering U5:AUG or U5:DIS interactions did not change translation levels observed,[336] suggesting the two pool hypothesis[327]. Alternatively, it is also possible that both monomer and dimers can be translated, but some portion of dimers can be captured by Gag and serve as packagable RNA. On the other hand, using live cell imaging assays, the theory of functional heterogeneity in the RNA transcripts is backed up by the latest results that HIV-1 Gag preferentially encapsidates non-translating RNAs into the assembling virion, though this study did not confirm the existence of two distinct pools of RNA.[337]

Consistent with the 'two-pool' hypothesis, using NMR, recent studies have shown that different pools of unspliced RNAs exist based on their transcription initiation site.[36,338] Using nuclease protection assays or S1 nuclease mapping, earlier studies have shown that HIV-1 transcription initiation occurs either at the 1st G or at 2nd G of U3-R junction; however, these assays were designed to detect a single RNA population.[339–340] More recently, two other studies have confirmed the alternative usage of transcription initiation sites in HIV-1, based on 5′ rapid amplification of cDNA ends (5′ RACE).[341–342] Further verification of the 5′-RACE is necessary to obtain more confirmatory results as both of these studies reported issues with copying cap, misincorporating or deleting nucleotides at the 5′-end, etc. A recent study confirms the observations of Masuda et al, showing heterogenous transcription sites and preferential packaging of 1G RNA conserved among different HIV-1 founder viruses.[343] They also noted that different strains of HIV-2 and SIV exhibit a similar phenomenon.[343] It is proposed that if the transcript has only 1G at the 5′ end, the RNA undergoes packaging and translation is inhibited by adopting a structure in which the 5′ cap is sequestered from the eukaryotic translation initiation factor 4E (eIF4E) interaction by coaxial stacking of TAR and polyA helices.[344] Moreover, NMR studies on CES RNAs (starting with 1G) have demonstrated a three-way junction structure in which the U5:Gag AUG interaction favors the dimerization-competent structure, as previously observed.[26] On the other hand, if the transcript starts with either 2 or 3Gs, the RNAs are used as mRNA and are enriched in polysomes.[36,338] Additionally, these RNAs exhibit U5:DIS interaction that prevent dimerization and packaging of these RNAs.[36]

Another recent study supports the observation of structural heterogeneity based on the transcription start site in HIV-1, using in gel probing-SHAPE and *in vivo* RNA-packaging experiments.[37] This study showed that transcripts with only 1G can form dimers, and this RNA can adopt more than one structure at the same time, but almost all these structures had exposed polyA, DIS and multiple Gs that could interact with Gag, although a small fraction of these conformations had only DIS exposed (with the polyA refolded), revealing the dynamic nature of the HIV-1 5′UTR and the ensemble of structures that can co-exist simultaneously. On the other hand, it showed that the 3G RNAs could not form dimers, and the dimerization-incompetent multiple monomeric RNAs showed structures with destabilized polyA and DIS sites. This study also concluded that the stability of the polyA stem is the main determinant that drives RNA packaging[37], unlike the previous model that suggested U5:DIS base pairing prevents dimerization and packaging of 3G RNAs.[338] Underlining this finding, a study based on a high-throughput approach called functional analysis of RNA structure-sequencing (FARS-seq) has recently shown that the 3G RNA exhibits reconfigured polyA and SL1.[38] Interestingly, the overall structure of the HIV-1 3G RNA in the 5′ UTR published by Ye and colleagues is different[38] from the one published by Nikolatchik and coworkers,[37] especially in the SL2 and polyA regions as well as their complementary binding sites.

Role of other mechanisms like cap sequestration and epitranscriptomics in gRNA packaging

Recent studies are revealing other parts of the host cellular machinery involved in regulating RNA processing and translation that could modulate gRNA packaging. For example, it has been shown that cap sequestration is important for packaging by a mechanism which is more complicated than structural heterogeneity of 1G, 2G and 3G RNAs (Figure 1).[344–345] The authors conclude that the stability of polyA is important for packaging in a way that the coaxial stacking of polyA stem and TAR sequester the cap efficiently. They also speculate that the cellular RNA processing and translation machinery captures the cap-exposed RNA in an irreversible manner so that they are not available for packaging.[344] However, more detailed studies are required to confirm this hypothesis.

The role of epitranscriptomics in gRNA packaging has recently been proposed by two groups.[169,346] The first group has shown that hypermethylation of mRNA cap leads to an alternate translation mechanism using the cap binding protein 80/nuclear cap binding protein 3 (CBP80/NCBP3) heterodimers, rather than eIF4.[169] However, the mechanism of cap hypermethylation is not clear, an aspect that needs further investigation such as does it occur on the 3G RNAs where cap is available for modification or other RNA substrates. The other study, on the other hand, has observed that the most common RNA methylation mark at adenosine residues (m6A) on the unspliced RNA is associated with increased Gag expression, while

demethylated RNAs are packaged into HIV-1 particles.[346] This suggests an elegant mechanism that the virus can use to identify pools of unspliced retroviral RNA that is used for translation *versus* packaging.

The existence of two pools of RNA (one for translation and other for packaging) was established much earlier in the case of MLV by treating cells with the transcription inhibitor, actinomycin D. This treatment did not reduce Gag translation, but reduced the amount of packaged RNA.[347–348] Although it has been shown that the nuclear export pathway plays a role in determining cytoplasmic fate of unspliced RNAs,[270] whether the heterogeneous transcription start site segregates unspliced RNA into two pools in MLV has not yet been elucidated. In HIV-2, the situation may be unique since the same RNA has been shown to be packaged that is translated simultaneously. Thus, HIV-2 may discriminate between packageable genomic and spliced mRNAs by using a co-translation packaging mechanism where the genomic RNA is captured by Gag in "*cis*" from the same genomic RNA that is translated at the same time.[349–350] However, a later study disputes these results and shows that HIV-2 packages its gRNA in *trans* like HIV-1.[351] Therefore, whether the co-translational packaging mechanism exists at all for HIV-2 or other retroviruses remains to be explored.

Perspectives/Concluding Remarks

Replication of retroviruses involves a multitude of intricate processes that are highly integrated and occur both inside the host cell and outside it to generate infectious virus particles. This necessitates exploitation of the host machinery involved in transcription, nuclear import/export and processing of RNAs, as well as translation. It also includes the unique and obligatory steps of gRNA dimerization and its encapsidation into the virus particles, areas attractive for therapeutics development in the context of increasing resistance to currently available anti-retroviral therapies. As our understanding of retroviral replication and the packaging process evolves, older questions give rise to new ones, such as the nature of packageable competent and incompetent RNAs. This includes the emerging roles of epitranscriptomics (such as m6A RNA modification and cap methylation), heterogeneity of transcript start sites, and the necessity of functional polyA sequences, as well as their relationship to gRNA dimerization and packaging *versus* translation. Another emerging question relates to multiple high affinity Gag binding sites being not sufficient to drive gRNA packaging,[32] suggesting that not just the mere presence, but the positioning of these sites in the tertiary structural context of RNA may be critical for Gag multimerization on the gRNA to ensure its appropriate capture into the assembling virion. Hence, in addition to tackling the "how", it is equally important to explore the "where" (subcellular location where the initial Gag-gRNA interactions occurs) and "with whom" (role of cellular factors that facilitate this process) in understanding the mechanism of gRNA packaging within cells, aspects that are all interconnected in time and space. With continued advancement in technologies, many such challenging questions will be answered, and newer ones will arise. However, this knowledge should lead to not only a better understanding of retrovirus replication and its pathogenesis, but also the development of retroviral-based vectors for gene therapy, as well as novel anti-retrovirals and vaccines targeted towards combating virus replication and as potential therapeutics.

Funding Details

Research in TAR's laboratory is funded by Abu Dhabi Precision Medicine Virtual Research Institute (VRI) grant (21R098), Abu Dhabi Department of Education and Knowledge (ADEK) ASPIRE Grant (21M150), College of Medicine and Health Sciences (CMHS), United Arab Emirates University (UAEU) grants (NP-22-33/12M131, NP-20-29/31M476/12M040, and NP-18-30/31M388), UAEU Program for Advanced Research-UPAR grant (12M103) and UAEU Zayed bin Sultan Center for Health Sciences (ZCHS) grants (UCBR-31R123 and UCBR-31R143). AC was supported by UAEU ZCHS grant (UCBR-31R123). Research in FM's laboratory is funded by the Al Jalila Foundation (AJF2020006), Sheikh Hamdan Award for Medical Sciences (MRG-39/2011-2012), CMHS, UAEU grants (31M421 and 12M092), and UAEU ZCHS grants (UCBR-31R140 and UCBR-31R122).

Conflict of Interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Akhil Chameettachal: Methodology, Visualization, Writing – original draft, Writing – review & editing. **Farah Mustafa:** Conceptualization, Methodology, Supervision, Writing – review & editing, Funding acquisition. **Tahir A. Rizvi:** Conceptualization, Methodology, Supervision, Writing – review & editing, Project administration, Funding acquisition.

Received 17 October 2022;
Accepted 12 December 2022;
Available online 16 December 2022

Keywords:
retroviruses;
retroviral replication;
RNA Packaging;
RNA nuclear export;
RNA dimerization

† Current Address: Viral Recombination Section, HIV Dynamics and Replication Program, National Cancer Institute at Frederick, Frederick, MD 21702, USA.

References

1. Knipe, D.M., Howley, P., (2013). Fields Virology. Lippincott Williams & Wilkins.
2. Baltimore, D., (1970). Viral RNA-dependent DNA Polymerase: RNA-dependent DNA Polymerase in Virions of RNA Tumour Viruses. *Nature* **226**, 1209–1211. https://doi.org/10.1038/2261209a0.
3. Stehelin, D., Varmus, H.E., Bishop, J.M., Vogt, P.K., (1976). DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. *Nature* **260**, 170–173. https://doi.org/10.1038/260170a0.
4. Temin, H.M., Mizutani, S., (1970). Viral RNA-dependent DNA Polymerase: RNA-dependent DNA Polymerase in Virions of Rous Sarcoma Virus. *Nature* **226**, 1211–1213. https://doi.org/10.1038/2261211a0.
5. Dharan, A., Bachmann, N., Talley, S., Zwikelmaier, V., Campbell, E.M., (2020). Nuclear pore blockade reveals that HIV-1 completes reverse transcription and uncoating in the nucleus, Nature. *Microbiology* **5**, 1088–1095. https://doi.org/10.1038/s41564-020-0735-8.
6. Burdick, R.C., Li, C., Munshi, M., Rawson, J.M.O., Nagashima, K., Hu, W.-S., Pathak, V.K., (2020). HIV-1 uncoats in the nucleus near sites of integration. *PNAS* **117**, 5486–5493. https://doi.org/10.1073/pnas.1920631117.
7. Roe, T., Reynolds, T.C., Yu, G., Brown, P.O., (1993). Integration of murine leukemia virus DNA depends on mitosis. *EMBO J.* **12**, 2099–2108.
8. Bukrinsky, M.I., Sharova, N., Dempsey, M.P., Stanwick, T. L., Bukrinskaya, A.G., Haggerty, S., Stevenson, M., (1992). Active nuclear import of human immunodeficiency virus type 1 preintegration complexes. *PNAS* **89**, 6580–6584.
9. Chen, H., Engelman, A., (1998). The barrier-to-autointegration protein is a host factor for HIV type 1 integration. *Proc. Nat. Acad. Sci.* **95**, 15270–15274.
10. Cherepanov, P., Maertens, G., Proost, P., Devreese, B., Van Beeumen, J., Engelborghs, Y., De Clercq, E., Debyser, Z., (2003). HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. *J. Biol. Chem.* **278**, 372–381.
11. Li, C., Burdick, R.C., Nagashima, K., Hu, W.-S., Pathak, V.K., (2021). HIV-1 cores retain their integrity until minutes before uncoating in the nucleus. *PNAS* **118** https://doi.org/10.1073/pnas.2019467118.
12. Shun, M.-C., Raghavendra, N.K., Vandegraaff, N., Daigle, J.E., Hughes, S., Kellam, P., Cherepanov, P., Engelman, A., (2007). LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration. *Genes Dev.* **21**, 1767–1778. https://doi.org/10.1101/gad.1565107.
13. Feinberg, M.B., Baltimore, D., Frankel, A.D., (1991). The role of Tat in the human immunodeficiency virus life cycle indicates a primary effect on transcriptional elongation. *PNAS* **88**, 4045–4049. https://doi.org/10.1073/pnas.88.9.4045.
14. Suzuki, T., Fujisawa, J.I., Toita, M., Yoshida, M., (1993). The trans-activator tax of human T-cell leukemia virus type 1 (HTLV-1) interacts with cAMP-responsive element (CRE) binding and CRE modulator proteins that bind to the 21-base-pair enhancer of HTLV-1. *PNAS* **90**, 610–614. https://doi.org/10.1073/pnas.90.2.610.
15. Balvay, L., Lopez Lastra, M., Sargueil, B., Darlix, J.-L., Ohlmann, T., (2007). Translational control of retroviruses. *Nat. Rev. Microbiol.* **5**, 128–140. https://doi.org/10.1038/nrmicro1599.
16. Swanstrom, R., Wills, J.W., (1997). Synthesis, Assembly, and Processing of Viral Proteins. In: Coffin, J.M., Hughes, S.H., Varmus, H.E. (Eds.), *Retroviruses*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY) (accessed September 29, 2022).
17. Coffin, J.M., Hughes, S.H., Varmus, H.E. (Eds.). *Retroviruses*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY) (accessed November 20, 2020).
18. Freed, E.O., (2015). HIV-1 assembly, release and maturation. *Nat. Rev. Microbiol.* **13**, 484–496. https://doi.org/10.1038/nrmicro3490.
19. Xiang, Y., Cameron, C.E., Wills, J.W., Leis, J., (1996). Fine mapping and characterization of the Rous sarcoma virus Pr76gag late assembly domain. *J. Virol.* **70**, 5695–5700. https://doi.org/10.1128/JVI.70.8.5695-5700.1996.
20. Katoh, I., Yoshinaka, Y., Rein, A., Shibuya, M., Odaka, T., Oroszlan, S., (1985). Murine leukemia virus maturation: protease region required for conversion from “immature” to “mature” core form and for virus infectivity. *Virology* **145**, 280–292. https://doi.org/10.1016/0042-6822(85)90161-8.
21. Kohl, N.E., Emini, E.A., Schleif, W.A., Davis, L.J., Heimbach, J.C., Dixon, R.A., Scolnick, E.M., Sigal, I.S., (1988). Active human immunodeficiency virus protease is required for viral infectivity. *PNAS* **85**, 4686–4690. https://doi.org/10.1073/pnas.85.13.4686.
22. Moore, M.D., Hu, W.S., (2009). HIV-1 RNA dimerization: It takes two to tango. *AIDS Rev.* **11**, 91–102.
23. D’Souza, V., Summers, M.F., (2005). How retroviruses select their genomes. *Nat. Rev. Microbiol.* **3**, 643. https://doi.org/10.1038/nrmicro1210.
24. Merino, E.J., Wilkinson, K.A., Coughlan, J.L., Weeks, K. M., (2005). RNA structure analysis at single nucleotide resolution by selective 2’-hydroxyl acylation and primer extension (SHAPE). *J. Am. Chem. Soc.* **127**, 4223–4231. https://doi.org/10.1021/ja043822v.
25. Wilkinson, K.A., Gorelick, R.J., Vasa, S.M., Guex, N., Rein, A., Mathews, D.H., Giddings, M.C., Weeks, K.M., (2008). High-throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states. *PLoS Biol.* **6**, e96.
26. Keane, S.C., Heng, X., Lu, K., Kharytonchyk, S., Ramakrishnan, V., Carter, G., Barton, S., Hosic, A., et al., (2015). RNA structure. Structure of the HIV-1 RNA packaging signal. *Science* **348**, 917–921. https://doi.org/10.1126/science.aaa9266.
27. Lu, K., Heng, X., Garyu, L., Monti, S., Garcia, E.L., Kharytonchyk, S., Dorjsuren, B., Kulandaivel, G., et al., (2011). NMR detection of structures in the HIV-1 5’-leader

RNA that regulate genome packaging. *Science* **334**, 242–245. https://doi.org/10.1126/science.1210460.
28. Siegfried, N.A., Busan, S., Rice, G.M., Nelson, J.A.E., Weeks, K.M., (2014). RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). *Nat. Methods* **11**, 959–965. https://doi.org/10.1038/nmeth.3029.
29. Lu, K., Heng, X., Summers, M.F., (2011). Structural determinants and mechanism of HIV-1 genome packaging. *J. Mol. Biol.* **410**, 609–633. https://doi.org/10.1016/j.jmb.2011.04.029.
30. Abd El-Wahab, E.W., Smyth, R.P., Mailler, E., Bernacchi, S., Vivet-Boudou, V., Hijnen, M., Jossinet, F., Mak, J., et al., (2014). Specific recognition of the HIV-1 genomic RNA by the Gag precursor. *Nat. Commun.* **5**, 4304. https://doi.org/10.1038/ncomms5304.
31. Ali, L.M., Pitchai, F.N.N., Vivet-Boudou, V., Chameettachal, A., Jabeen, A., Pillai, V.N., Mustafa, F., Marquet, R., et al., (2020). Role of Purine-Rich Regions in Mason-Pfizer Monkey Virus (MPMV) Genomic RNA Packaging and Propagation. *Front. Microbiol.* **11** https://doi.org/10.3389/fmicb.2020.595410.
32. Chameettachal, A., Vivet-Boudou, V., Pitchai, F.N.N., Pillai, V.N., Ali, L.M., Krishnan, A., Bernacchi, S., Mustafa, F., et al., (2021). A purine loop and the primer binding site are critical for the selective encapsidation of mouse mammary tumor virus genomic RNA by Pr77Gag. *Nucleic Acids Res.*. https://doi.org/10.1093/nar/gkab223.
33. Nikolaitchik, O.A., Somoulay, X., Rawson, J.M.O., Yoo, J. A., Pathak, V.K., Hu, W.-S., (2020). Unpaired Guanosines in the 5′ Untranslated Region of HIV-1 RNA Act Synergistically to Mediate Genome Packaging. *J. Virol.*. https://doi.org/10.1128/JVI.00439-20.
34. Nuzra Nagoor Pitchai, F., Chameettachal, A., Vivet-Boudou, V., Mohamed Ali, L., Pillai, V.N., Krishnan, A., Bernacchi, S., Mustafa, F., et al., (2021). Identification of Pr78Gag binding sites on the Mason-Pfizer monkey virus genomic RNA packaging determinants. *J. Mol. Biol.*. https://doi.org/10.1016/j.jmb.2021.166923.
35. Umunnakwe, C.N., Duchon, A., Nikolaitchik, O.A., Rahman, S.A., Liu, Y., Chen, J., Tai, S., Pathak, V.K., et al., (2021). Specific Guanosines in the HIV-2 Leader RNA are Essential for Efficient Viral Genome Packaging. *J. Mol. Biol.* **433**, https://doi.org/10.1016/j.jmb.2020.11.017 166718.
36. Brown, J.D., Kharytonchyk, S., Chaudry, I., Iyer, A.S., Carter, H., Becker, G., Desai, Y., Glang, L., et al., (2020). Structural basis for transcriptional start site control of HIV-1 RNA fate. *Science* **368**, 413–417. https://doi.org/10.1126/science.aaz7959.
37. Nikolaitchik, O.A., Liu, S., Kitzrow, J.P., Liu, Y., Rawson, J.M.O., Shakya, S., Cheng, Z., Pathak, V.K., et al., (2021). Selective packaging of HIV-1 RNA genome is guided by the stability of 5' untranslated region polyA stem. *Proc. Natl. Acad. Sci. U S A* **118** https://doi.org/10.1073/pnas.2114494118. e2114494118.
38. Ye, L., Gribling-Burrer, A.-S., Bohn, P., Kibe, A., Börtlein, C., Ambi, U.B., Ahmad, S., Olguin-Nava, M., et al., (2022). Short- and long-range interactions in the HIV-1 5' UTR regulate genome dimerization and packaging. *Nat. Struct. Mol. Biol.* **29**, 306–319. https://doi.org/10.1038/s41594-022-00746-2.
39. Ellerman, C., Bang, O., (1908). Centralbl. *Bakteriol* **46**, 595–609.
40. Rous, P., (1911). A sarcoma of the fowl transmissible by an agent separable from the tumor cells. *J. Exp. Med.* **13**, 397–411. https://doi.org/10.1084/jem.13.4.397.
41. Poiesz, B.J., Ruscetti, F.W., Gazdar, A.F., Bunn, P.A., Minna, J.D., Gallo, R.C., (1980). Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *PNAS* **77**, 7415–7419. https://doi.org/10.1073/pnas.77.12.7415.
42. Yoshida, M., Miyoshi, I., Hinuma, Y., (1982). Isolation and characterization of retrovirus from cell lines of human adult T-cell leukemia and its implication in the disease. *PNAS* **79**, 2031–2035.
43. Barré-Sinoussi, F., Chermann, J.C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., et al., (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* **220**, 868–871. https://doi.org/10.1126/science.6189183.
44. Gallo, R.C., Salahuddin, S.Z., Popovic, M., Shearer, G.M., Kaplan, M., Haynes, B.F., Palker, T.J., Redfield, R., et al., (1984). Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science* **224**, 500–503. https://doi.org/10.1126/science.6200936.
45. Briggs, J.A.G., Simon, M.N., Gross, I., Kräusslich, H.-G., Fuller, S.D., Vogt, V.M., Johnson, M.C., (2004). The stoichiometry of Gag protein in HIV-1. *Nat. Struct. Mol. Biol.* **11**, 672–675. https://doi.org/10.1038/nsmb785.
46. Carlson, L.-A., Briggs, J.A.G., Glass, B., Riches, J.D., Simon, M.N., Johnson, M.C., Müller, B., Grünewald, K., et al., (2008). Three-Dimensional Analysis of Budding Sites and Released Virus Suggests a Revised Model for HIV-1 Morphogenesis. *Cell Host Microbe* **4**, 592–599. https://doi.org/10.1016/j.chom.2008.10.013.
47. Zhu, P., Liu, J., Bess, J., Chertova, E., Lifson, J.D., Grisé, H., Ofek, G.A., Taylor, K.A., et al., (2006). Distribution and three-dimensional structure of AIDS virus envelope spikes. *Nature* **441**, 847–852. https://doi.org/10.1038/nature04817.
48. Hunter, E., (2008). Retroviruses: General Features. In: Mahy, B.W.J., Van Regenmortel, M.H.V. (Eds.), *Encyclopedia of Virology*. third ed. Academic Press, Oxford, pp. 459–467. https://doi.org/10.1016/B978-012374410-4.00677-4.
49. Doms, R.W., Moore, J.P., (2000). HIV-1 Membrane Fusion. *J. Cell Biol.* **151**, f9–f14.
50. White, J.M., Delos, S.E., Brecher, M., Schornberg, K., (2008). Structures and Mechanisms of Viral Membrane Fusion Proteins. *Crit. Rev. Biochem. Mol. Biol.* **43**, 189–219. https://doi.org/10.1080/10409230802058320.
51. McClure, M.O., Marsh, M., Weiss, R.A., (1988). Human immunodeficiency virus infection of CD4-bearing cells occurs by a pH-independent mechanism. *EMBO J.* **7**, 513–518.
52. Sinangil, F., Loyter, A., Volsky, D.J., (1988). Quantitative measurement of fusion between human immunodeficiency virus and cultured cells using membrane fluorescence dequenching. *FEBS Lett.* **239**, 88–92. https://doi.org/10.1016/0014-5793(88)80551-9.
53. Stein, B.S., Gowda, S.D., Lifson, J.D., Penhallow, R.C., Bensch, K.G., Engleman, E.G., (1987). pH-independent HIV entry into CD4-positive T cells via virus envelope

fusion to the plasma membrane. *Cell* **49**, 659–668. https://doi.org/10.1016/0092-8674(87)90542-3.
54. Mothes, W., Boerger, A.L., Narayan, S., Cunningham, J. M., Young, J.A., (2000). Retroviral entry mediated by receptor priming and low pH triggering of an envelope glycoprotein. *Cell* **103**, 679–689. https://doi.org/10.1016/s0092-8674(00)00170-7.
55. Ross, S.R., Schofield, J.J., Farr, C.J., Bucan, M., (2002). Mouse transferrin receptor 1 is the cell entry receptor for mouse mammary tumor virus. *PNAS* **99**, 12386–12390. https://doi.org/10.1073/pnas.192360099.
56. Dalgleish, A.G., Beverley, P.C., Clapham, P.R., Crawford, D.H., Greaves, M.F., Weiss, R.A., (1984). The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature* **312**, 763–767. https://doi.org/10.1038/312763a0.
57. McDougal, J.S., Kennedy, M.S., Sligh, J.M., Cort, S.P., Mawle, A., Nicholson, J.K., (1986). Binding of HTLV-III/LAV to T4+ T cells by a complex of the 110K viral protein and the T4 molecule. *Science* **231**, 382–385. https://doi.org/10.1126/science.3001934.
58. Bleul, C.C., Wu, L., Hoxie, J.A., Springer, T.A., Mackay, C.R., (1997). The HIV coreceptors CXCR4 and CCR5 are differentially expressed and regulated on human T lymphocytes. *PNAS* **94**, 1925–1930. https://doi.org/10.1073/pnas.94.5.1925.
59. McNicholl, J.M., Smith, D.K., Qari, S.H., Hodge, T., (1997). Host genes and HIV: the role of the chemokine receptor gene CCR5 and its allele. *Emerg. Infect. Dis.* **3**, 261–271. https://doi.org/10.3201/eid0303.970302.
60. Borsetti, A., Parolin, C., Ridolfi, B., Sernicola, L., Geraci, A., Ensoli, B., Titti, F., (2000). CD4-Independent Infection of Two CD4−/CCR5−/CXCR4+ Pre-T-Cell Lines by Human and Simian Immunodeficiency Viruses. *J. Virol.* **74**, 6689–6694.
61. Dragic, T., Alizon, M., (1993). Different requirements for membrane fusion mediated by the envelopes of human immunodeficiency virus types 1 and 2. *J. Virol.* **67**, 2355–2359. https://doi.org/10.1128/JVI.67.4.2355-2359.1993.
62. Weiss, R.A., (1993). Cellular Receptors and Viral Glycoproteins Involved in Retrovirus Entry. In: Levy, J.A. (Ed.), *The Retroviridae*. Springer US, Boston, MA, pp. 1–108. https://doi.org/10.1007/978-1-4899-1627-3_1.
63. Ghez, D., Lepelletier, Y., Lambert, S., Fourneau, J.-M., Blot, V., Janvier, S., Arnulf, B., van Endert, P.M., et al., (2006). Neuropilin-1 is involved in human T-cell lymphotropic virus type 1 entry. *J. Virol.* **80**, 6844–6854. https://doi.org/10.1128/JVI.02719-05.
64. Jones, K.S., Petrow-Sadowski, C., Bertolette, D.C., Huang, Y., Ruscetti, F.W., (2005). Heparan sulfate proteoglycans mediate attachment and entry of human T-cell leukemia virus type 1 virions into CD4+ T cells. *J. Virol.* **79**, 12692–12702. https://doi.org/10.1128/JVI.79.20.12692-12702.2005.
65. Manel, N., Kim, F.J., Kinet, S., Taylor, N., Sitbon, M., Battini, J.-L., (2003). The ubiquitous glucose transporter GLUT-1 is a receptor for HTLV. *Cell* **115**, 449–459. https://doi.org/10.1016/s0092-8674(03)00881-x.
66. Hildreth, J.E., Subramanium, A., Hampton, R.A., (1997). Human T-cell lymphotropic virus type 1 (HTLV-1)-induced syncytium formation mediated by vascular cell adhesion molecule-1: evidence for involvement of cell adhesion molecules in HTLV-1 biology. *J. Virol.* **71**, 1173–1180. https://doi.org/10.1128/JVI.71.2.1173-1180.1997.
67. Ross, S.R., Schmidt, J.W., Katz, E., Cappelli, L., Hultine, S., Gimotty, P., Monroe, J.G., (2006). An Immunoreceptor Tyrosine Activation Motif in the Mouse Mammary Tumor Virus Envelope Protein Plays a Role in Virus-Induced Mammary Tumors. *J. Virol.* **80**, 9000–9008. https://doi.org/10.1128/JVI.00788-06.
68. Lee, Y.M., Coffin, J.M., (1991). Relationship of avian retrovirus DNA synthesis to integration in vitro. *Mol. Cell Biol.* **11**, 1419–1430. https://doi.org/10.1128/mcb.11.3.1419-1430.1991.
69. Selyutina, A., Persaud, M., Lee, K., KewalRamani, V., Diaz-Griffero, F., (2020). Nuclear Import of the HIV-1 Core Precedes Reverse Transcription and Uncoating. *Cell Rep.* **32**, https://doi.org/10.1016/j.celrep.2020.108201 108201.
70. Fassati, A., Goff, S.P., (2001). Characterization of intracellular reverse transcription complexes of human immunodeficiency virus type 1. *J. Virol.* **75**, 3626–3635. https://doi.org/10.1128/JVI.75.8.3626-3635.2001.
71. Preston, B.D., Poiesz, B.J., Loeb, L.A., (1988). Fidelity of HIV-1 reverse transcriptase. *Science* **242**, 1168–1171. https://doi.org/10.1126/science.2460924.
72. Mangeat, B., Turelli, P., Caron, G., Friedli, M., Perrin, L., Trono, D., (2003). Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature* **424**, 99–103. https://doi.org/10.1038/nature01709.
73. O'Neil, P.K., Sun, G., Yu, H., Ron, Y., Dougherty, J.P., Preston, B.D., (2002). Mutational analysis of HIV-1 long terminal repeats to explore the relative contribution of reverse transcriptase and RNA polymerase II to viral mutagenesis. *J. Biol. Chem.* **277**, 38053–38061. https://doi.org/10.1074/jbc.M204774200.
74. Venkatesan, S., Rosenthal, R., Kanu, N., McGranahan, N., Bartek, J., Quezada, S.A., Hare, J., Harris, R.S., et al., (2018). Perspective: APOBEC mutagenesis in drug resistance and immune escape in HIV and cancer evolution. *Ann. Oncol.* **29**, 563–572. https://doi.org/10.1093/annonc/mdy003.
75. Hu, W.-S., Hughes, S.H., (2012). HIV-1 reverse transcription. *Cold Spring Harb. Perspect. Med.* **2**, https://doi.org/10.1101/cshperspect.a006882 a006882.
76. Das, A.T., Klaver, B., Berkhout, B., (1995). Reduced replication of human immunodeficiency virus type 1 mutants that use reverse transcription primers other than the natural tRNA(3Lys). *J. Virol.* **69**, 3090–3097.
77. Wakefield, J.K., Kang, S.M., Morrow, C.D., (1996). Construction of a type 1 human immunodeficiency virus that maintains a primer binding site complementary to tRNA(His). *J. Virol.* **70**, 966–975. https://doi.org/10.1128/JVI.70.2.966-975.1996.
78. Jin, D., Musier-Forsyth, K., (2019). Role of host tRNAs and aminoacyl-tRNA synthetases in retroviral replication. *J. Biol. Chem.* **294**, 5352–5364. https://doi.org/10.1074/jbc.REV118.002957.
79. Hu, W.S., Bowman, E.H., Delviks, K.A., Pathak, V.K., (1997). Homologous recombination occurs in a distinct retroviral subpopulation and exhibits high negative interference. *J. Virol.* **71**, 6028–6036.
80. Jones, J.S., Allan, R.W., Temin, H.M., (1994). One retroviral RNA is sufficient for synthesis of viral DNA. *J. Virol.* **68**, 207–216. https://doi.org/10.1128/JVI.68.1.207-216.1994.

81. Charneau, P., Alizon, M., Clavel, F., (1992). A second origin of DNA plus-strand synthesis is required for optimal human immunodeficiency virus replication. *J. Virol.* **66**, 2814–2820. https://doi.org/10.1128/JVI.66.5.2814-2820.1992.
82. Hu, C., Saenz, D.T., Fadel, H.J., Walker, W., Peretz, M., Poeschla, E.M., (2010). The HIV-1 central polypurine tract functions as a second line of defense against APOBEC3G/F. *J. Virol.* **84**, 11981–11993. https://doi.org/10.1128/JVI.00723-10.
83. Kupiec, J.J., Tobaly-Tapiero, J., Canivet, M., Santillana-Hayat, M., Flügel, R.M., Périès, J., Emanoil-Ravier, R., (1988). Evidence for a gapped linear duplex DNA intermediate in the replicative cycle of human and simian spumaviruses. *Nucleic Acids Res.* **16**, 9557–9565. https://doi.org/10.1093/nar/16.20.9557.
84. Stetor, S.R., Rausch, J.W., Guo, M.J., Burnham, J.P., Boone, L.R., Waring, M.J., Le Grice, S.F., (1999). Characterization of (+) strand initiation and termination sequences located at the center of the equine infectious anemia virus genome. *Biochemistry* **38**, 3656–3667. https://doi.org/10.1021/bi982764l.
85. Whitwam, T., Peretz, M., Poeschla, E., (2001). Identification of a central DNA flap in feline immunodeficiency virus. *J. Virol.* **75**, 9407–9414. https://doi.org/10.1128/JVI.75.19.9407-9414.2001.
86. Kizhatil, K., Albritton, L.M., (1997). Requirements for different components of the host cell cytoskeleton distinguish ecotropic murine leukemia virus entry via endocytosis from entry via surface fusion. *J. Virol.* **71**, 7145–7156.
87. McDonald, D., Vodicka, M.A., Lucero, G., Svitkina, T.M., Borisy, G.G., Emerman, M., Hope, T.J., (2002). Visualization of the intracellular behavior of HIV in living cells. *J. Cell Biol.* **159**, 441–452. https://doi.org/10.1083/jcb.200203150.
88. Katz, R.A., Greger, J.G., Boimel, P., Skalka, A.M., (2003). Human immunodeficiency virus type 1 DNA nuclear import and integration are mitosis independent in cycling cells. *J. Virol.* **77**, 13412–13417. https://doi.org/10.1128/jvi.77.24.13412-13417.2003.
89. Farnet, C.M., Haseltine, W.A., (1991). Determination of viral proteins present in the human immunodeficiency virus type 1 preintegration complex. *J. Virol.* **65**, 1910–1915. https://doi.org/10.1128/JVI.65.4.1910-1915.1991.
90. Levy, N., Eiler, S., Pradeau-Aubreton, K., Crucifix, C., Schaetzel, A., Drillien, R., Parissi, V., Emiliani, S., et al., (2013). Structural and functional studies of the HIV-1 pre-integration complex. *Retrovirology* **10**, P76. https://doi.org/10.1186/1742-4690-10-S1-P76.
91. Bukrinsky, M., (2004). A hard way to the nucleus. *Mol. Med.* **10**, 1–5.
92. Dvorin, J.D., Malim, M.H., (2003). Intracellular trafficking of HIV-1 cores: journey to the center of the cell. *Curr. Top. Microbiol. Immunol.* **281**, 179–208. https://doi.org/10.1007/978-3-642-19012-4_5.
93. Lehmann-Che, J., Saïb, A., (2004). Early stages of HIV replication: how to hijack cellular functions for a successful infection. *AIDS Rev.* **6**, 199–207.
94. Suzuki, Y., Craigie, R., (2007). The road to chromatin - nuclear entry of retroviruses. *Nat. Rev. Microbiol.* **5**, 187–196. https://doi.org/10.1038/nrmicro1579.
95. Buffone, C., Martinez-Lopez, A., Fricke, T., Opp, S., Severgnini, M., Cifola, I., Petiti, L., Frabetti, S., et al., (2018). Nup153 Unlocks the Nuclear Pore Complex for HIV-1 Nuclear Translocation in Nondividing Cells. *J. Virol.* **92**, e00648–e718. https://doi.org/10.1128/JVI.00648-18.
96. Dharan, A., Talley, S., Tripathi, A., Mamede, J.I., Majetschak, M., Hope, T.J., Campbell, E.M., (2016). KIF5B and Nup358 Cooperatively Mediate the Nuclear Import of HIV-1 during Infection. *PLoS Pathog.* **12**, e1005700.
97. Zila, V., Margiotta, E., Turonova, B., Müller, T.G., Zimmerli, C.E., Mattei, S., Allegretti, M., Börner, K., et al., (2020). Cone-shaped HIV-1 capsids are transported through intact nuclear pores 2020.07.30.193524 *BioRxiv*.. https://doi.org/10.1101/2020.07.30.193524.
98. Francis, A.C., Marin, M., Prellberg, M.J., Palermino-Rowland, K., Melikyan, G.B., (2020). HIV-1 Uncoating and Nuclear Import Precede the Completion of Reverse Transcription in Cell Lines and in Primary Macrophages. *Viruses* **12**, 1234. https://doi.org/10.3390/v12111234.
99. Bejarano, D.A., Peng, K., Laketa, V., Börner, K., Jost, K. L., Lucic, B., Glass, B., Lusic, M., et al., (2019). HIV-1 nuclear import in macrophages is regulated by CPSF6-capsid interactions at the nuclear pore complex. *Elife* **8**, e41800.
100. Matreyek, K.A., Yücel, S.S., Li, X., Engelman, A., (2013). Nucleoporin NUP153 phenylalanine-glycine motifs engage a common binding pocket within the HIV-1 capsid protein to mediate lentiviral infectivity. *PLoS Pathog.* **9**, e1003693.
101. Müller, T.G., Zila, V., Peters, K., Schifferdecker, S., Stanic, M., Lucic, B., Laketa, V., Lusic, M., et al., (2021). HIV-1 uncoating by release of viral cDNA from capsid-like structures in the nucleus of infected cells. *Elife* **10**, e64776.
102. Rankovic, S., Varadarajan, J., Ramalho, R., Aiken, C., Rousso, I., (2017). Reverse Transcription Mechanically Initiates HIV-1 Capsid Disassembly. *J. Virol.* **91**, e00289–e317. https://doi.org/10.1128/JVI.00289-17.
103. Yu, A., Lee, E.M.Y., Briggs, J.A.G., Ganser-Pornillos, B. K., Pornillos, O., Voth, G.A., (2022). Strain and rupture of HIV-1 capsids during uncoating. *Proc. Natl. Acad. Sci. U S A.* **119** https://doi.org/10.1073/pnas.2117781119. e2117781119.
104. Rankovic, S., Deshpande, A., Harel, S., Aiken, C., Rousso, I., (2021). HIV-1 uncoating occurs via a series of rapid biomechanical changes in the core related to individual stages of reverse transcription JVI.00166-21 *J. Virol.*. https://doi.org/10.1128/JVI.00166-21.
105. Desfarges, S., Ciuffi, A., (2010). Retroviral Integration Site Selection. *Viruses* **2**, 111–130. https://doi.org/10.3390/v2010111.
106. Brown, P.O., Bowerman, B., Varmus, H.E., Bishop, J.M., (1989). Retroviral integration: structure of the initial covalent product and its precursor, and a role for the viral IN protein. *PNAS* **86**, 2525–2529. https://doi.org/10.1073/pnas.86.8.2525.
107. Farnet, C.M., Bushman, F.D., (1997). HIV-1 cDNA integration: requirement of HMG I(Y) protein for function of preintegration complexes in vitro. *Cell* **88**, 483–492. https://doi.org/10.1016/s0092-8674(00)81888-7.
108. Emiliani, S., Mousnier, A., Busschots, K., Maroun, M., Van Maele, B., Tempé, D., Vandekerckhove, L., Moisant, F., et al., (2005). Integrase mutants defective for interaction with LEDGF/p75 are impaired in

chromosome tethering and HIV-1 replication. *J. Biol. Chem.* **280**, 25517–25523. https://doi.org/10.1074/jbc.M501378200.
109. Llano, M., Vanegas, M., Hutchins, N., Thompson, D., Delgado, S., Poeschla, E.M., (2006). Identification and characterization of the chromatin-binding domains of the HIV-1 integrase interactor LEDGF/p75. *J. Mol. Biol.* **360**, 760–773. https://doi.org/10.1016/j.jmb.2006.04.073.
110. Gupta, S.S., Maetzig, T., Maertens, G.N., Sharif, A., Rothe, M., Weidner-Glunde, M., Galla, M., Schambach, A., et al., (2013). Bromo- and extraterminal domain chromatin regulators serve as cofactors for murine leukemia virus integration. *J. Virol.* **87**, 12721–12736. https://doi.org/10.1128/JVI.01942-13.
111. Elliott, J.L., Eschbach, J.E., Koneru, P.C., Li, W., Puray-Chavez, M., Townsend, D., Lawson, D.Q., Engelman, A. N., et al., (2020). Integrase-RNA interactions underscore the critical role of integrase in HIV-1 virion morphogenesis. *Elife* **9**, e54311.
112. Kessl, J.J., Kutluay, S.B., Townsend, D., Rebensburg, S., Slaughter, A., Larue, R.C., Shkriabai, N., Bakouche, N., et al., (2016). HIV-1 Integrase Binds the Viral RNA Genome and Is Essential during Virion Morphogenesis. *Cell* **166**, 1257–1268.e12. https://doi.org/10.1016/j.cell.2016.07.044.
113. Brown, H.E., Chen, H., Engelman, A., (1999). Structure-based mutagenesis of the human immunodeficiency virus type 1 DNA attachment site: effects on integration and cDNA synthesis. *J. Virol.* **73**, 9011–9020. https://doi.org/10.1128/JVI.73.11.9011-9020.1999.
114. Ali, L.M., Rizvi, T.A., Mustafa, F., (2016). Cross- and Co-Packaging of Retroviral RNAs and Their Consequences. *Viruses* **8**, 276. https://doi.org/10.3390/v8100276.
115. Oh, J., Chang, K.W., Wierzchoslawski, R., Alvord, W.G., Hughes, S.H., (2008). Rous sarcoma virus (RSV) integration in vivo: a CA dinucleotide is not required in U3, and RSV linear DNA does not autointegrate. *J. Virol.* **82**, 503–512. https://doi.org/10.1128/JVI.01441-07.
116. Acel, A., Udashkin, B.E., Wainberg, M.A., Faust, E.A., (1998). Efficient Gap Repair Catalyzed In Vitro by an Intrinsic DNA Polymerase Activity of Human Immunodeficiency Virus Type 1 Integrase. *J. Virol.* **72**, 2062–2071. https://doi.org/10.1128/JVI.72.3.2062-2071.1998.
117. Yoder, K.E., Bushman, F.D., (2000). Repair of Gaps in Retroviral DNA Integration Intermediates. *J. Virol.* **74**, 11191–11200. https://doi.org/10.1128/JVI.74.23.11191-11200.2000.
118. Lesbats, P., Engelman, A.N., Cherepanov, P., (2016). Retroviral DNA Integration. *Chem. Rev.* **116**, 12730–12757. https://doi.org/10.1021/acs.chemrev.6b00125.
119. Maskell, D.P., Renault, L., Serrao, E., Lesbats, P., Matadeen, R., Hare, S., Lindemann, D., Engelman, A. N., et al., (2015). Structural basis for retroviral integration into nucleosomes. *Nature* **523**, 366–369. https://doi.org/10.1038/nature14495.
120. Naughtin, M., Haftek-Terreau, Z., Xavier, J., Meyer, S., Silvain, M., Jaszczyszyn, Y., Levy, N., Miele, V., et al., (2015). DNA Physical Properties and Nucleosome Positions Are Major Determinants of HIV-1 Integrase Selectivity. *PLoS One* **10**, e0129427.
121. Holman, A.G., Coffin, J.M., (2005). Symmetrical base preferences surrounding HIV-1, avian sarcoma/leukosis virus, and murine leukemia virus integration sites. *PNAS* **102**, 6103–6107. https://doi.org/10.1073/pnas.0501646102.
122. Stevens, S.W., Griffith, J.D., (1994). Human immunodeficiency virus type 1 may preferentially integrate into chromatin occupied by L1Hs repetitive elements. *PNAS* **91**, 5557–5561.
123. Wu, X., Li, Y., Crise, B., Burgess, S.M., Munroe, D.J., (2005). Weak Palindromic Consensus Sequences Are a Common Feature Found at the Integration Target Sites of Many Retroviruses. *J. Virol.* **79**, 5211–5214. https://doi.org/10.1128/JVI.79.8.5211-5214.2005.
124. Mitchell, R.S., Beitzel, B.F., Schroder, A.R.W., Shinn, P., Chen, H., Berry, C.C., Ecker, J.R., Bushman, F.D., (2004). Retroviral DNA Integration: ASLV, HIV, and MLV Show Distinct Target Site Preferences. *PLoS Biol.* **2** https://doi.org/10.1371/journal.pbio.0020234.
125. Shih, C.-C., Stoye, J.P., Coffin, J.M., (1988). Highly preferred targets for retrovirus integration. *Cell* **53**, 531–537. https://doi.org/10.1016/0092-8674(88)90569-7.
126. Konstantoulas, I., (2014). Mouse mammary tumor virus-based vector transduces non-dividing cells, enters the nucleus via a TNPO3-independent pathway and integrates in a less biased fashion than other retroviruses. *Retrovirology* **11**, 34. https://doi.org/10.1186/1742-4690-11-34.
127. Lewinski, M.K., Yamashita, M., Emerman, M., Ciuffi, A., Marshall, H., Crawford, G., Collins, F., Shinn, P., et al., (2006). Retroviral DNA Integration: Viral and Cellular Determinants of Target-Site Selection. *PLoS Pathog.* **2** https://doi.org/10.1371/journal.ppat.0020060.
128. Schaller, T., Ocwieja, K.E., Rasaiyaah, J., Price, A.J., Brady, T.L., Roth, S.L., Hué, S., Fletcher, A.J., et al., (2011). HIV-1 capsid-cyclophilin interactions determine nuclear import pathway, integration targeting and replication efficiency. *PLoS Pathog.* **7**, e1002439.
129. De Rijck, J., de Kogel, C., Demeulemeester, J., Vets, S., El Ashkar, S., Malani, N., Bushman, F.D., Landuyt, B., et al., (2013). The BET family of proteins targets moloney murine leukemia virus integration near transcription start sites. *Cell Rep.* **5**, 886–894. https://doi.org/10.1016/j.celrep.2013.09.040.
130. Nabel, G., Baltimore, D., (1987). An inducible transcription factor activates expression of human immunodeficiency virus in T cells. *Nature* **326**, 711–713. https://doi.org/10.1038/326711a0.
131. Günzburg, W.H., Salmons, B., (1992). Factors controlling the expression of mouse mammary tumour virus. *Biochem. J.* **283**, 625–632.
132. Winkler, I., Bodem, J., Haas, L., Zemba, M., Delius, H., Flower, R., Flügel, R.M., Löchelt, M., (1997). Characterization of the genome of feline foamy virus and its proteins shows distinct features different from those of primate spumaviruses. *J. Virol.* **71**, 6727–6741.
133. Omoto, S., Brisibe, E.A., Okuyama, H., Fujii, Y.R., (2004). Feline foamy virus Tas protein is a DNA-binding transactivator. *J. Gen. Virol.* **85**, 2931–2935. https://doi.org/10.1099/vir.0.80088-0.
134. Chou, S., Upton, H., Bao, K., Schulze-Gahmen, U., Samelson, A.J., He, N., Nowak, A., Lu, H., et al., (2013). HIV-1 Tat recruits transcription elongation factors dispersed along a flexible AFF4 scaffold. *PNAS* **110**, E123–E131. https://doi.org/10.1073/pnas.1216971110.

135. Guntaka, R.V., (1993). Transcription termination and polyadenylation in retroviruses. *Microbiol. Rev.* **57**, 511–521.
136. Schrom, E.-M., Moschall, R., Hartl, M.J., Weitner, H., Fecher, D., Langemeier, J., Bohne, J., Wöhrl, B.M., et al., (2013). U1snRNP-mediated suppression of polyadenylation in conjunction with the RNA structure controls poly (A) site selection in foamy viruses. *Retrovirology* **10**, 55. https://doi.org/10.1186/1742-4690-10-55.
137. Ashe, M.P., Griffin, P., James, W., Proudfoot, N.J., (1995). Poly(A) site selection in the HIV-1 provirus: inhibition of promoter-proximal polyadenylation by the downstream major splice donor site. *Genes Dev.* **9**, 3008–3025. https://doi.org/10.1101/gad.9.23.3008.
138. Iwasaki, K., Temin, H.M., (1990). The efficiency of RNA 3'-end formation is determined by the distance between the cap site and the poly(A) site in spleen necrosis virus. *Genes Dev.* **4**, 2299–2307. https://doi.org/10.1101/gad.4.12b.2299.
139. Weichs an der Glon, C., Monks, J., Proudfoot, N.J., (1991). Occlusion of the HIV poly(A) site. *Genes Dev.* **5**, 244–253. https://doi.org/10.1101/gad.5.2.244.
140. Ashe, M.P., Pearson, L.H., Proudfoot, N.J., (1997). The HIV-1 5' LTR poly(A) site is inactivated by U1 snRNP interaction with the downstream major splice donor site. *EMBO J.* **16**, 5752–5763. https://doi.org/10.1093/emboj/16.18.5752.
141. Ashe, M.P., Furger, A., Proudfoot, N.J., (2000). Stem-loop 1 of the U1 snRNP plays a critical role in the suppression of HIV-1 polyadenylation. *RNA* **6**, 170–177. https://doi.org/10.1017/s1355838200991957.
142. Hadzopoulou-Cladaras, M., Felber, B.K., Cladaras, C., Athanassopoulos, A., Tse, A., Pavlakis, G.N., (1989). The rev (trs/art) protein of human immunodeficiency virus type 1 affects viral mRNA and protein expression via a cis-acting sequence in the env region. *J. Virol.* **63**, 1265–1274. https://doi.org/10.1128/JVI.63.3.1265-1274.1989.
143. Malim, M.H., Hauber, J., Le, S.Y., Maizel, J.V., Cullen, B. R., (1989). The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature* **338**, 254–257. https://doi.org/10.1038/338254a0.
144. Dayton, A.I., (2004). Within you, without you: HIV-1 Rev and RNA export. *Retrovirology* **1**, 35. https://doi.org/10.1186/1742-4690-1-35.
145. Neville, M., Stutz, F., Lee, L., Davis, L.I., Rosbash, M., (1997). The importin-beta family member Crm1p bridges the interaction between Rev and the nuclear pore complex during nuclear export. *Curr. Biol.* **7**, 767–775. https://doi.org/10.1016/s0960-9822(06)00335-6.
146. Grassmann, R., Berchtold, S., Aepinus, C., Ballaun, C., Boehnlein, E., Fleckenstein, B., (1991). In vitro binding of human T-cell leukemia virus rex proteins to the rex-response element of viral transcripts. *J. Virol.* **65**, 3721–3727. https://doi.org/10.1128/JVI.65.7.3721-3727.1991.
147. Hanly, S.M., Rimsky, L.T., Malim, M.H., Kim, J.H., Hauber, J., Duc Dodon, M., Le, S.Y., Maizel, J.V., et al., (1989). Comparative analysis of the HTLV-I Rex and HIV-1 Rev trans-regulatory proteins and their RNA response elements. *Genes Dev.* **3**, 1534–1544. https://doi.org/10.1101/gad.3.10.1534.
148. Nakano, K., Watanabe, T., (2012). HTLV-1 Rex: the courier of viral messages making use of the host vehicle. *Front. Microbiol.* **3** https://doi.org/10.3389/fmicb.2012.00330.
149. Indik, S., Günzburg, W.H., Salmons, B., Rouault, F., (2005). A novel, mouse mammary tumor virus encoded protein with Rev-like properties. *Virology* **337**, 1–6. https://doi.org/10.1016/j.virol.2005.03.040.
150. Mertz, J.A., Simper, M.S., Lozano, M.M., Payne, S.M., Dudley, J.P., (2005). Mouse Mammary Tumor Virus Encodes a Self-Regulatory RNA Export Protein and Is a Complex Retrovirus. *J. Virol.* **79**, 14737–14747. https://doi.org/10.1128/JVI.79.23.14737-14747.2005.
151. Müllner, M., Salmons, B., Günzburg, W.H., Indik, S., (2008). Identification of the Rem-responsive element of mouse mammary tumor virus. *Nucleic Acids Res.* **36**, 6284–6294.
152. Bray, M., Prasad, S., Dubay, J.W., Hunter, E., Jeang, K. T., Rekosh, D., Hammarskjöld, M.L., (1994). A small element from the Mason-Pfizer monkey virus genome makes human immunodeficiency virus type 1 expression and replication Rev-independent. *PNAS* **91**, 1256–1260.
153. Braun, I.C., Herold, A., Rode, M., Conti, E., Izaurralde, E., (2001). Overexpression of TAP/p15 heterodimers bypasses nuclear retention and stimulates nuclear mRNA export. *J. Biol. Chem.* **276**, 20536–20543. https://doi.org/10.1074/jbc.M100400200.
154. Cullen, B.R., (2003). Nuclear RNA export. *J. Cell Sci.* **116**, 587–597. https://doi.org/10.1242/jcs.00268.
155. Grüter, P., Tabernero, C., von Kobbe, C., Schmitt, C., Saavedra, C., Bachi, A., Wilm, M., Felber, B.K., et al., (1998). TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export from the nucleus. *Mol. Cell* **1**, 649–659. https://doi.org/10.1016/s1097-2765(00)80065-9.
156. Yoshinaka, Y., Katoh, I., Copeland, T.D., Oroszlan, S., (1985). Murine leukemia virus protease is encoded by the gag-pol gene and is synthesized through suppression of an amber termination codon. *PNAS* **82**, 1618–1622.
157. Yoshinaka, Y., Katoh, I., Copeland, T.D., Oroszlan, S., (1985). Translational readthrough of an amber termination codon during synthesis of feline leukemia virus protease. *J. Virol.* **55**, 870–873.
158. Chamorro, M., Parkin, N., Varmus, H.E., (1992). An RNA pseudoknot and an optimal heptameric shift site are required for highly efficient ribosomal frameshifting on a retroviral messenger RNA. *PNAS* **89**, 713–717.
159. Jacks, T., Power, M.D., Masiarz, F.R., Luciw, P.A., Barr, P.J., Varmus, H.E., (1988). Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature* **331**, 280–283. https://doi.org/10.1038/331280a0.
160. Jacks, T., Townsley, K., Varmus, H.E., Majors, J., (1987). Two efficient ribosomal frameshifting events are required for synthesis of mouse mammary tumor virus gag-related polyproteins. *PNAS* **84**, 4298–4302.
161. Shehu-Xhilaga, M., Crowe, S.M., Mak, J., (2001). Maintenance of the Gag/Gag-Pol ratio is important for human immunodeficiency virus type 1 RNA dimerization and viral infectivity. *J. Virol.* **75**, 1834–1841. https://doi.org/10.1128/JVI.75.4.1834-1841.2001.
162. Benner, B.E., Bruce, J.W., Kentala, J.R., Murray, M., Becker, J.T., Garcia-Miranda, P., Ahlquist, P., Butcher, S. E., et al., (2022). Perturbing HIV-1 Ribosomal Frameshifting Frequency Reveals a cis Preference for Gag-Pol Incorporation into Assembling Virions. *J. Virol.* **96**, e0134921.

163. Chen, X., Chamorro, M., Lee, S.I., Shen, L.X., Hines, J.V., Tinoco, I., Varmus, H.E., (1995). Structural and functional studies of retroviral RNA pseudoknots involved in ribosomal frameshifting: nucleotides at the junction of the two stems are important for efficient ribosomal frameshifting. *EMBO J.* **14**, 842–852.
164. Berlioz, C., Darlix, J.L., (1995). An internal ribosomal entry mechanism promotes translation of murine leukemia virus gag polyprotein precursors. *J. Virol.* **69**, 2214–2222. https://doi.org/10.1128/JVI.69.4.2214-2222.1995.
165. Göttlinger, H.G., Sodroski, J.G., Haseltine, W.A., (1989). Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1. *PNAS* **86**, 5781–5785. https://doi.org/10.1073/pnas.86.15.5781.
166. Bussienne, C., Marquet, R., Paillart, J.-C., Bernacchi, S., (2021). Post-Translational Modifications of Retroviral HIV-1 Gag Precursors: An Overview of Their Biological Role. *Int. J. Mol. Sci.* **22**, 2871. https://doi.org/10.3390/ijms22062871.
167. Hill, C.P., Worthylake, D., Bancroft, D.P., Christensen, A. M., Sundquist, W.I., (1996). Crystal structures of the trimeric human immunodeficiency virus type 1 matrix protein: implications for membrane association and assembly. *PNAS* **93**, 3099–3104. https://doi.org/10.1073/pnas.93.7.3099.
168. Yu, G., Shen, F.S., Sturch, S., Aquino, A., Glazer, R.I., Felsted, R.L., (1995). Regulation of HIV-1 gag protein subcellular targeting by protein kinase C. *J. Biol. Chem.* **270**, 4792–4796. https://doi.org/10.1074/jbc.270.9.4792.
169. Singh, G., Seufzer, B., Song, Z., Zucko, D., Heng, X., Boris-Lawrie, K., (2022). HIV-1 hypermethylated guanosine cap licenses specialized translation unaffected by mTOR. *Proc. Nat. Acad. Sci.* **119** https://doi.org/10.1073/pnas.2105153118. e2105153118.
170. Larson, D.R., Johnson, M.C., Webb, W.W., Vogt, V.M., (2005). Visualization of retrovirus budding with correlated light and electron microscopy. *PNAS* **102**, 15453–15458. https://doi.org/10.1073/pnas.0504812102.
171. Chameettachal, A., Pillai, V., Ali, L., Pitchai, F., Ardah, M., Mustafa, F., Marquet, R., Rizvi, T., (2018). Biochemical and Functional Characterization of Mouse Mammary Tumor Virus Full-Length Pr77Gag Expressed in Prokaryotic and Eukaryotic Cells. *Viruses* **10**, 334.
172. Krishnan, A., Pillai, V.N., Chameettachal, A., Mohamed Ali, L., Nuzra Nagoor Pitchai, F., Tariq, S., Mustafa, F., Marquet, R., et al., (2019). Purification and Functional Characterization of a Biologically Active Full-Length Feline Immunodeficiency Virus (FIV) Pr50Gag. *Viruses.* **11** https://doi.org/10.3390/v11080689.
173. McKinstry, W.J., Hijnen, M., Tanwar, H.S., Sparrow, L.G., Nagarajan, S., Pham, S.T., Mak, J., (2014). Expression and purification of soluble recombinant full length HIV-1 Pr55Gag protein in Escherichia coli. *Protein Expr. Purif.* **100**, 10–18. https://doi.org/10.1016/j.pep.2014.04.013.
174. Pitchai, F.N.N., Ali, L., Pillai, V.N., Chameettachal, A., Ashraf, S.S., Mustafa, F., Marquet, R., Rizvi, T.A., (2018). Expression, purification, and characterization of biologically active full-length Mason-Pfizer monkey virus (MPMV) Pr78 Gag. *Sci. Rep.* **8**, 11793.
175. Campbell, S., Fisher, R.J., Towler, E.M., Fox, S., Issaq, H.J., Wolfe, T., Phillips, L.R., Rein, A., (2001). Modulation of HIV-like particle assembly in vitro by inositol phosphates. *PNAS* **98**, 10875–10879. https://doi.org/10.1073/pnas.191224698.
176. Freed, E.O., (1998). HIV-1 gag proteins: diverse functions in the virus life cycle. *Virology* **251**, 1–15.
177. Ono, A., Orenstein, J.M., Freed, E.O., (2000). Role of the Gag matrix domain in targeting human immunodeficiency virus type 1 assembly. *J. Virol.* **74**, 2855–2866. https://doi.org/10.1128/jvi.74.6.2855-2866.2000.
178. Ono, A., Ablan, S.D., Lockett, S.J., Nagashima, K., Freed, E.O., (2004). Phosphatidylinositol (4,5) bisphosphate regulates HIV-1 Gag targeting to the plasma membrane. *PNAS* **101**, 14889–14894. https://doi.org/10.1073/pnas.0405596101.
179. Saad, J.S., Miller, J., Tai, J., Kim, A., Ghanam, R.H., Summers, M.F., (2006). Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. *PNAS* **103**, 11364–11369. https://doi.org/10.1073/pnas.0602818103.
180. Cannon, P.M., Matthews, S., Clark, N., Byles, E.D., Iourin, O., Hockley, D.J., Kingsman, S.M., Kingsman, A. J., (1997). Structure-function studies of the human immunodeficiency virus type 1 matrix protein, p17. *J. Virol.* **71**, 3474–3483. https://doi.org/10.1128/JVI.71.5.3474-3483.1997.
181. Dalton, A.K., Murray, P.S., Murray, D., Vogt, V.M., (2005). Biochemical characterization of rous sarcoma virus MA protein interaction with membranes. *J. Virol.* **79**, 6227–6238. https://doi.org/10.1128/JVI.79.10.6227-6238.2005.
182. Zábranský, A., Hadravová, R., Stokrová, J., Sakalian, M., Pichová, I., (2009). Premature processing of mouse mammary tumor virus Gag polyprotein impairs intracellular capsid assembly. *Virology* **384**, 33–37. https://doi.org/10.1016/j.virol.2008.10.038.
183. Rhee, H., (1990). A single amino acid substitution within the matrix protein of a type D retrovirus converts its morphogenesis to that of a type C retrovirus. *Cell* **63**, 77–86. https://doi.org/10.1016/0092-8674(90)90289-Q.
184. Rhee, H., (1991). Amino acid substitutions within the matrix protein of type D retroviruses affect assembly, transport and membrane association of a capsid. *EMBO J.* **10**, 535–546. https://doi.org/10.1002/j.1460-2075.1991.tb07980.x.
185. Arnaud, F., Murcia, P.R., Palmarini, M., (2007). Mechanisms of late restriction induced by an endogenous retrovirus. *J. Virol.* **81**, 11441–11451. https://doi.org/10.1128/JVI.01214-07.
186. Choi, G., Park, S., Choi, B., Hong, S., Lee, J., Hunter, E., Rhee, S.S., (1999). Identification of a Cytoplasmic Targeting/Retention Signal in a Retroviral Gag Polyprotein. *J. Virol.* **73**, 5431–5437. https://doi.org/10.1128/JVI.73.7.5431-5437.1999.
187. Murcia, P.R., Arnaud, F., Palmarini, M., (2007). The transdominant endogenous retrovirus enJS56A1 associates with and blocks intracellular trafficking of Jaagsiekte sheep retrovirus Gag. *J. Virol.* **81**, 1762–1772. https://doi.org/10.1128/JVI.01859-06.
188. Sfakianos, J.N., LaCasse, R.A., Hunter, E., (2003). The M-PMV cytoplasmic targeting-retention signal directs nascent Gag polypeptides to a pericentriolar region of the cell. *Traffic* **4**, 660–670. https://doi.org/10.1034/j.1600-0854.2003.00125.x.
189. Sfakianos, J.N., Hunter, E., (2003). M-PMV capsid transport is mediated by Env/Gag interactions at the

pericentriolar recycling endosome. *Traffic* **4**, 671–680. https://doi.org/10.1034/j.1600-0854.2003.00126.x.
190. Yu, S.F., Eastman, S.W., Linial, M.L., (2006). Foamy virus capsid assembly occurs at a pericentriolar region through a cytoplasmic targeting/retention signal in Gag. *Traffic* **7**, 966–977. https://doi.org/10.1111/j.1600-0854.2006.00448.x.
191. Zhang, G., Sharon, D., Jovel, J., Liu, L., Wine, E., Tahbaz, N., Indik, S., Mason, A., (2015). Pericentriolar Targeting of the Mouse Mammary Tumor Virus GAG Protein. *PLoS One* **10**, e0131515.
192. Dawson, L., Yu, X.F., (1998). The role of nucleocapsid of HIV-1 in virus assembly. *Virology* **251**, 141–157. https://doi.org/10.1006/viro.1998.9374.
193. Burniston, M.T., Cimarelli, A., Colgan, J., Curtis, S.P., Luban, J., (1999). Human immunodeficiency virus type 1 Gag polyprotein multimerization requires the nucleocapsid domain and RNA and is promoted by the capsid-dimer interface and the basic region of matrix protein. *J. Virol.* **73**, 8527–8540.
194. Ott, D.E., Coren, L.V., Shatzer, T., (2009). The Nucleocapsid Region of Human Immunodeficiency Virus Type 1 Gag Assists in the Coordination of Assembly and Gag Processing: Role for RNA-Gag Binding in the Early Stages of Assembly. *J. Virol.* **83**, 7718–7727. https://doi.org/10.1128/JVI.00099-09.
195. Tanwar, H.S., Khoo, K.K., Garvey, M., Waddington, L., Leis, A., Hijnen, M., Velkov, T., Dumsday, G.J., et al., (2017). The thermodynamics of Pr55Gag-RNA interaction regulate the assembly of HIV. *PLoS Pathog.* **13**, e1006221.
196. Yang, Y., Qu, N., Tan, J., Rushdi, M.N., Krueger, C.J., Chen, A.K., (2018). Roles of Gag-RNA interactions in HIV-1 virus assembly deciphered by single-molecule localization microscopy. *PNAS* **115**, 6721–6726. https://doi.org/10.1073/pnas.1805728115.
197. Göttlinger, H.G., Dorfman, T., Sodroski, J.G., Haseltine, W.A., (1991). Effect of mutations affecting the p6 gag protein on human immunodeficiency virus particle release. *PNAS* **88**, 3195–3199.
198. Garrus, J.E., von Schwedler, U.K., Pornillos, O.W., Morham, S.G., Zavitz, K.H., Wang, H.E., Wettstein, D. A., Stray, K.M., et al., (2001). Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell* **107**, 55–65. https://doi.org/10.1016/s0092-8674(01)00506-2.
199. Göttlinger, H.G., Dorfman, T., Cohen, E.A., Haseltine, W. A., (1993). Vpu protein of human immunodeficiency virus type 1 enhances the release of capsids produced by gag gene constructs of widely divergent retroviruses. *PNAS* **90**, 7381–7385. https://doi.org/10.1073/pnas.90.15.7381.
200. Parent, L.J., Bennett, R.P., Craven, R.C., Nelle, T.D., Krishna, N.K., Bowzard, J.B., Wilson, C.B., Puffer, B.A., et al., (1995). Positionally independent and exchangeable late budding functions of the Rous sarcoma virus and human immunodeficiency virus Gag proteins. *J. Virol.* **69**, 5455–5460. https://doi.org/10.1128/jvi.69.9.5455-5460.1995.
201. Konvalinka, J., Kräusslich, H.-G., Müller, B., (2015). Retroviral proteases and their roles in virion maturation. *Virology* **479–480**, 403–417. https://doi.org/10.1016/j.virol.2015.03.021.
202. Wlodawer, A., Gustchina, A., (2000). Structural and biochemical studies of retroviral proteases. *Biochim. Biophys. Acta* **1477**, 16–34. https://doi.org/10.1016/s0167-4838(99)00267-8.
203. Zhao, G., Perilla, J.R., Yufenyuy, E.L., Meng, X., Chen, B., Ning, J., Ahn, J., Gronenborn, A.M., et al., (2013). Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature* **497**, 643–646. https://doi.org/10.1038/nature12162.
204. Comas-Garcia, M., Davis, S.R., Rein, A., (2016). On the Selective Packaging of Genomic RNA by HIV-1. *Viruses* **8**, 246. https://doi.org/10.3390/v8090246.
205. Johnson, S.F., Telesnitsky, A., (2010). Retroviral RNA Dimerization and Packaging: The What. *How, When, Where, and Why, PLOS Pathogens.* **6**, e1001007.
206. Lever, A.M.L., (2007). HIV-1 RNA packaging. *Adv. Pharmacol.* **55**, 1–32. https://doi.org/10.1016/S1054-3589(07)55001-5.
207. Rein, A., (2019). RNA Packaging in HIV. *Trends Microbiol.* **27**, 715–723. https://doi.org/10.1016/j.tim.2019.04.003.
208. Aktar, S.J., Jabeen, A., Ali, L.M., Vivet-Boudou, V., Marquet, R., Rizvi, T.A., (2013). SHAPE analysis of the 5′ end of the Mason-Pfizer monkey virus (MPMV) genomic RNA reveals structural elements required for genome dimerization. *RNA* **19**, 1648–1658. https://doi.org/10.1261/rna.040931.113.
209. Aktar, S.J., Vivet-Boudou, V., Ali, L.M., Jabeen, A., Kalloush, R.M., Richer, D., Mustafa, F., Marquet, R., et al., (2014). Structural basis of genomic RNA (gRNA) dimerization and packaging determinants of mouse mammary tumor virus (MMTV). *Retrovirology* **11**, 96. https://doi.org/10.1186/s12977-014-0096-6.
210. Paillart, J.-C., Dettenhofer, M., Yu, X., Ehresmann, C., Ehresmann, B., Marquet, R., (2004). First Snapshots of the HIV-1 RNA Structure in Infected Cells and in Virions. *J. Biol. Chem.* **279**, 48397–48403. https://doi.org/10.1074/jbc.M408294200.
211. Rizvi, T.A., Kenyon, J.C., Ali, J., Aktar, S.J., Phillip, P.S., Ghazawi, A., Mustafa, F., Lever, A.M.L., (2010). Optimal packaging of FIV genomic RNA depends upon a conserved long-range interaction and a palindromic sequence within gag. *J. Mol. Biol.* **403**, 103–119. https://doi.org/10.1016/j.jmb.2010.08.019.
212. Bell, N.M., Lever, A.M.L., (2013). HIV Gag polyprotein: processing and early viral particle assembly. *Trends Microbiol.* **21**, 136–144. https://doi.org/10.1016/j.tim.2012.11.006.
213. Miyazaki, Y., Miyake, A., Nomaguchi, M., Adachi, A., (2011). Structural dynamics of retroviral genome and the packaging. *Front. Microbiol.* **2**, 264. https://doi.org/10.3389/fmicb.2011.00264.
214. Aldovini, A., 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.
215. Gorelick, R.J., Henderson, L.E., Hanser, J.P., Rein, A., (1988). Point mutants of Moloney murine leukemia virus that fail to package viral RNA: evidence for specific RNA recognition by a “zinc finger-like” protein sequence. *PNAS* **85**, 8420–8424.
216. Wang, C.T., Barklis, E., (1993). Assembly, processing, and infectivity of human immunodeficiency virus type 1 gag mutants. *J. Virol.* **67**, 4264–4273. https://doi.org/10.1128/jvi.67.7.4264-4273.1993.

217. Cimarelli, A., Sandin, S., Höglund, S., Luban, J., (2000). Basic residues in human immunodeficiency virus type 1 nucleocapsid promote virion assembly via interaction with RNA. *J. Virol.* **74**, 3046–3057.
218. Housset, V., De Rocquigny, H., Roques, B.P., Darlix, J.L., (1993). Basic amino acids flanking the zinc finger of Moloney murine leukemia virus nucleocapsid protein NCp10 are critical for virus infectivity. *J. Virol.* **67**, 2537–2545. https://doi.org/10.1128/jvi.67.5.2537-2545.1993.
219. Aronoff, R., Linial, M., (1991). Specificity of retroviral RNA packaging. *J. Virol.* **65**, 71–80.
220. Poon, D.T., Wu, J., Aldovini, A., (1996). Charged amino acid residues of human immunodeficiency virus type 1 nucleocapsid p7 protein involved in RNA packaging and infectivity. *J. Virol.* **70**, 6607–6616.
221. Klein, R., Ruttkowski, B., Schwab, S., Peterbauer, T., Salmons, B., Günzburg, W.H., Hohenadl, C., (2008). Mouse Mammary Tumor Virus Promoter-Containing Retroviral Promoter Conversion Vectors for Gene-Directed Enzyme Prodrug Therapy are Functional In Vitro and In Vivo. *J. Biomed. Biotechnol.* **2008** https://doi.org/10.1155/2008/683505.
222. Alfadhli, A., Barklis, E., (2014). The roles of lipids and nucleic acids in HIV-1 assembly. *Front. Microbiol.* **5**, 253. https://doi.org/10.3389/fmicb.2014.00253.
223. Chukkapalli, V., Oh, S.J., 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**, 1600–1605. https://doi.org/10.1073/pnas.0908661107.
224. Katoh, I., Kyushiki, H., Sakamoto, Y., Ikawa, Y., Yoshinaka, Y., (1991). Bovine leukemia virus matrix-associated protein MA(p15): further processing and formation of a specific complex with the dimer of the 5'-terminal genomic RNA fragment. *J. Virol.* **65**, 6845–6855.
225. Wang, H., Norris, K.M., Mansky, L.M., (2003). Involvement of the matrix and nucleocapsid domains of the bovine leukemia virus Gag polyprotein precursor in viral RNA packaging. *J. Virol.* **77**, 9431–9438. https://doi.org/10.1128/jvi.77.17.9431-9438.2003.
226. Bernacchi, S., Abd El-Wahab, E.W., Dubois, N., Hijnen, M., Smyth, R.P., Mak, J., Marquet, R., Paillart, J.-C., (2017). HIV-1 Pr55Gagbinds genomic and spliced RNAs with different affinity and stoichiometry. *RNA Biol.* **14**, 90–103. https://doi.org/10.1080/15476286.2016.1256533.
227. Dubois, N., Marquet, R., Paillart, J.-C., Bernacchi, S., (2018). Retroviral RNA Dimerization: From Structure to Functions. *Front. Microbiol.* **9**, 527. https://doi.org/10.3389/fmicb.2018.00527.
228. Sarni, S., Biswas, B., Liu, S., Olson, E.D., Kitzrow, J.P., Rein, A., Wysocki, V.H., Musier-Forsyth, K., (2020). HIV-1 Gag protein with or without p6 specifically dimerizes on the viral RNA packaging signal. *J. Biol. Chem.* **295**, 14391–14401. https://doi.org/10.1074/jbc.RA120.014835.
229. Duchon, A., Santos, S., Chen, J., Brown, M., Nikolaitchik, O.A., Tai, S., Chao, J.A., Freed, E.O., et al., (2021). Plasma Membrane Anchoring and Gag: Gag Multimerization on Viral RNA Are Critical Properties of HIV-1 Gag Required To Mediate Efficient Genome Packaging. *MBio* **12**, e0325421.
230. Bohl, C.R., Brown, S.M., Weldon, R.A., (2005). The pp24 phosphoprotein of Mason-Pfizer monkey virus contributes to viral genome packaging. *Retrovirology* **2**, 68. https://doi.org/10.1186/1742-4690-2-68.
231. Füzik, T., Píchalová, R., Schur, F.K.M., Strohalmová, K., Křížová, I., Hadravová, R., Rumlová, M., Briggs, J.A.G., et al., (2016). Nucleic Acid Binding by Mason-Pfizer Monkey Virus CA Promotes Virus Assembly and Genome Packaging. *J. Virol.* **90**, 4593–4603. https://doi.org/10.1128/JVI.03197-15.
232. Dupraz, P., Spahr, P.F., (1992). Specificity of Rous sarcoma virus nucleocapsid protein in genomic RNA packaging. *J. Virol.* **66**, 4662–4670.
233. Poon, D.T.K., Li, G., Aldovini, A., (1998). Nucleocapsid and Matrix Protein Contributions to Selective Human Immunodeficiency Virus Type 1 Genomic RNA Packaging. *J. Virol.* **72**, 1983–1993.
234. Heng, X., Kharytonchyk, S., Garcia, E.L., Lu, K., Divakaruni, S.S., LaCotti, C., Edme, K., Telesnitsky, A., et al., (2012). Identification of a minimal region of the HIV-1 5'-leader required for RNA dimerization, NC binding, and packaging. *J. Mol. Biol.* **417**, 224–239. https://doi.org/10.1016/j.jmb.2012.01.033.
235. Smyth, R.P., Despons, L., Huili, G., Bernacchi, S., Hijnen, M., Mak, J., Jossinet, F., Weixi, L., et al., (2015). Mutational interference mapping experiment (MIME) for studying RNA structure and function. *Nat. Methods* **12**, 866–872. https://doi.org/10.1038/nmeth.3490.
236. Smyth, R.P., Smith, M.R., Jousset, A.-C., Despons, L., Laumond, G., Decoville, T., Cattenoz, P., Moog, C., et al., (2018). In cell mutational interference mapping experiment (in cell MIME) identifies the 5' polyadenylation signal as a dual regulator of HIV-1 genomic RNA production and packaging. *Nucleic Acids Res.*. https://doi.org/10.1093/nar/gky152.
237. Clever, J.L., Daniel Miranda, Parslow, T.G., (2002). RNA Structure and Packaging Signals in the 5′ Leader Region of the Human Immunodeficiency Virus Type 1 Genome. *J. Virol.* **76**, 12381–12387. https://doi.org/10.1128/JVI.76.23.12381-12387.2002.
238. Didierlaurent, L., Racine, P.J., 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**, 8915–8927. https://doi.org/10.1093/nar/gkr577.
239. Helga-Maria, C., Hammarskjöld, M.-L., 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**, 4127–4135.
240. McBride, M.S., Schwartz, M.D., Panganiban, A.T., (1997). Efficient encapsidation of human immunodeficiency virus type 1 vectors and further characterization of cis elements required for encapsidation. *J. Virol.* **71**, 4544–4554.
241. Comas-Garcia, M., Kroupa, T., Datta, S.A., Harvin, D.P., Hu, W.-S., Rein, A., (2018). Efficient support of virus-like particle assembly by the HIV-1 packaging signal. *Elife* **7** https://doi.org/10.7554/eLife.38438.
242. Keane, S.C., Summers, M.F., (2016). NMR Studies of the Structure and Function of the HIV-1 5'-Leader. *Viruses* **8** https://doi.org/10.3390/v8120338.
243. Singh, G., Rife, B.D., Seufzer, B., Salemi, M., Rendahl, A., Boris-Lawrie, K., (2018). Identification of conserved, primary sequence motifs that direct retrovirus RNA fate. *Nucleic Acids Res.* **46**, 7366–7378. https://doi.org/10.1093/nar/gky369.

244. De Guzman, R.N., Wu, Z.R., Stalling, C.C., Pappalardo, L., Borer, P.N., Summers, M.F., (1998). Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. *Science* **279**, 384–388.
245. Kenyon, J.C., Prestwood, L.J., Lever, A.M.L., (2015). A novel combined RNA-protein interaction analysis distinguishes HIV-1 Gag protein binding sites from structural change in the viral RNA leader. *Sci. Rep.* **5**, 14369. https://doi.org/10.1038/srep14369.
246. Ding, P., Kharytonchyk, S., Waller, A., Mbaekwe, U., Basappa, S., Kuo, N., Frank, H.M., Quasney, C., et al., (2020). Identification of the initial nucleocapsid recognition element in the HIV-1 RNA packaging signal. *PNAS* **117**, 17737. https://doi.org/10.1073/pnas.2008519117.
247. Dubois, N., Khoo, K.K., Ghossein, S., Seissler, T., Wolff, P., McKinstry, W.J., Mak, J., Paillart, J.-C., et al., (2018). The C-terminal p6 domain of the HIV-1 Pr55Gag precursor is required for specific binding to the genomic RNA. *RNA Biol.* **15**, 923–936. https://doi.org/10.1080/15476286.2018.1481696.
248. Gherghe, C., Lombo, T., Leonard, C.W., Datta, S.A.K., Bess, J.W., Gorelick, R.J., Rein, A., Weeks, K.M., (2010). Definition of a high-affinity Gag recognition structure mediating packaging of a retroviral RNA genome. *PNAS* **107**, 19248–19253. https://doi.org/10.1073/pnas.1006897107.
249. McBride, M.S., Panganiban, A.T., (1996). The human immunodeficiency virus type 1 encapsidation site is a multipartite RNA element composed of functional hairpin structures. *J. Virol.* **70**, 2963–2973. https://doi.org/10.1128/JVI.70.5.2963-2973.1996.
250. Baig, T.T., Lanchy, J.-M., Lodmell, J.S., (2009). Randomization and in vivo selection reveal a GGRG motif essential for packaging human immunodeficiency virus type 2 RNA. *J. Virol.* **83**, 802–810. https://doi.org/10.1128/JVI.01521-08.
251. Pillai, V.N., Ali, L.M., Prabhu, S.G., Krishnan, A., Chameettachal, A., Pitchai, F.N.N., Mustafa, F., Rizvi, T. A., (2021). A Stretch of Unpaired Purines in the Leader Region of Simian Immunodeficiency Virus (SIV) Genomic RNA is Critical for its Packaging into Virions. *J. Mol. Biol.* **433**, https://doi.org/10.1016/j.jmb.2021.167293 167293.
252. Wu, W., Hatterschide, J., Syu, Y.-C., Cantara, W.A., Blower, R.J., Hanson, H.M., Mansky, L.M., Musier-Forsyth, K., (2018). Human T-cell leukemia virus type 1 Gag domains have distinct RNA-binding specificities with implications for RNA packaging and dimerization. *J. Biol. Chem.* **293**, 16261–16276. https://doi.org/10.1074/jbc.RA118.005531.
253. D'Souza, V., Summers, M.F., (2004). Structural basis for packaging the dimeric genome of Moloney murine leukaemia virus. *Nature* **431**, 586. https://doi.org/10.1038/nature02944.
254. Mann, R., Mulligan, R.C., Baltimore, D., (1983). Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. *Cell* **33**, 153–159. https://doi.org/10.1016/0092-8674(83)90344-6.
255. Miyazaki, Y., Garcia, E.L., King, S.R., Iyalla, K., Loeliger, K., Starck, P., Syed, S., Telesnitsky, A., et al., (2010). An RNA Structural Switch Regulates Diploid Genome Packaging by Moloney Murine Leukemia Virus. *J. Mol. Biol.* **396**, 141–152. https://doi.org/10.1016/j.jmb.2009.11.033.
256. Mougel, M., Barklis, E., (1997). A role for two hairpin structures as a core RNA encapsidation signal in murine leukemia virus virions. *J. Virol.* **71**, 8061–8065. https://doi.org/10.1128/JVI.71.10.8061-8065.1997.
257. Miyazaki, Y., Irobalieva, R.N., Tolbert, B., Smalls-Mantey, A., Iyalla, K., Loeliger, K., D'Souza, V., Khant, H., et al., (2010). Structure of a Conserved Retroviral RNA Packaging Element by NMR Spectroscopy and Cryo-Electron Tomography. *J. Mol. Biol.* **404**, 751–772. https://doi.org/10.1016/j.jmb.2010.09.009.
258. Zhou, J., McAllen, J.K., Tailor, Y., Summers, M.F., (2005). High affinity nucleocapsid protein binding to the muPsi RNA packaging signal of Rous sarcoma virus. *J. Mol. Biol.* **349**, 976–988. https://doi.org/10.1016/j.jmb.2005.04.046.
259. Zhou, J., Bean, R.L., Vogt, V.M., Summers, M.F., (2007). Solution Structure of the Rous Sarcoma Virus Nucleocapsid Protein:μΨ RNA Packaging Signal Complex. *J. Mol. Biol.* **365**, 453–467. https://doi.org/10.1016/j.jmb.2006.10.013.
260. Comas-Garcia, M., Datta, S.A., Baker, L., Varma, R., Gudla, P.R., Rein, A., (2017). Dissection of specific binding of HIV-1 Gag to the "packaging signal" in viral RNA. *Elife* **6**, e27055.
261. Dilley, K.A., Nikolaitchik, O.A., Galli, A., Burdick, R.C., Levine, L., Li, K., Rein, A., Pathak, V.K., et al., (2017). Interactions between HIV-1 Gag and Viral RNA Genome Enhance Virion Assembly. *J. Virol.* **91** https://doi.org/10.1128/JVI.02319-16.
262. Garbitt-Hirst, R., Kenney, S.P., Parent, L.J., (2009). Genetic evidence for a connection between Rous sarcoma virus gag nuclear trafficking and genomic RNA packaging. *J. Virol.* **83**, 6790–6797. https://doi.org/10.1128/JVI.00101-09.
263. Maldonado, R.J.K., Rice, B., Chen, E.C., Tuffy, K.M., Chiari, E.F., Fahrbach, K.M., Hope, T.J., Parent, L.J., (2020). Visualizing Association of the Retroviral Gag Protein with Unspliced Viral RNA in the Nucleus. *MBio* **11** https://doi.org/10.1128/mBio.00524-20.
264. Bukrinsky, M.I., Haggerty, S., Dempsey, M.P., Sharova, N., Adzhubel, A., Spitz, L., Lewis, P., Goldfarb, D., et al., (1993). A nuclear localization signal within HIV-1 matrix protein that governs infection of non-dividing cells. *Nature* **365**, 666–669. https://doi.org/10.1038/365666a0.
265. Kräusslich, H.G., Welker, R., (1996). Intracellular transport of retroviral capsid components. *Curr. Top. Microbiol. Immunol.* **214**, 25–63.
266. Dupont, S., Sharova, N., DéHoratius, C., Virbasius, C.M., Zhu, X., Bukrinskaya, A.G., Stevenson, M., Green, M.R., (1999). A novel nuclear export activity in HIV-1 matrix protein required for viral replication. *Nature* **402**, 681–685. https://doi.org/10.1038/45272.
267. Poole, E., Strappe, P., Mok, H.-P., Hicks, R., Lever, A.M. L., (2005). HIV-1 Gag-RNA interaction occurs at a perinuclear/centrosomal site; analysis by confocal microscopy and FRET. *Traffic* **6**, 741–755. https://doi.org/10.1111/j.1600-0854.2005.00312.x.
268. Kemler, I., Meehan, A., Poeschla, E.M., (2010). Live-cell coimaging of the genomic RNAs and Gag proteins of two lentiviruses. *J. Virol.* **84**, 6352–6366. https://doi.org/10.1128/JVI.00363-10.
269. Tuffy, K.M., Maldonado, R.J.K., Chang, J., Rosenfeld, P., Cochrane, A., Parent, L.J., (2020). HIV-1 Gag Forms Ribonucleoprotein Complexes with Unspliced Viral RNA

at Transcription Sites. *Viruses* **12** https://doi.org/10.3390/v12111281.

270. Mougel, M., Akkawi, C., Chamontin, C., Feuillard, J., Pessel-Vivares, L., Socol, M., Laine, S., (2020). NXF1 and CRM1 nuclear export pathways orchestrate nuclear export, translation and packaging of murine leukaemia retrovirus unspliced RNA. *RNA Biol.* **17**, 528–538. https://doi.org/10.1080/15476286.2020.1713539.
271. Brandt, S., Bliißenbach, M., Grewe, B., Konietzny, R., Grunwald, T., Überla, K., (2007). Rev Proteins of Human and Simian Immunodeficiency Virus Enhance RNA Encapsidation. *PLoS Pathog.* **3**, e54.
272. 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.
273. Kharytonchyk, S., Brown, J.D., Stilger, K., Yasin, S., Iyer, A.S., Collins, J., Summers, M.F., Telesnitsky, A., (2018). Influence of gag and RRE Sequences on HIV-1 RNA Packaging Signal Structure and Function. *J. Mol. Biol.* **430**, 2066–2079. https://doi.org/10.1016/j.jmb.2018.05.029.
274. Kutluay, S.B., Zang, T., Blanco-Melo, D., Powell, C., Jannain, D., Errando, M., Bieniasz, P.D., (2014). Global changes in the RNA binding specificity of HIV-1 gag regulate virion genesis. *Cell* **159**, 1096–1109. https://doi.org/10.1016/j.cell.2014.09.057.
275. Liu, Y., Nikolaitchik, O.A., Rahman, S.A., Chen, J., Pathak, V.K., Hu, W.-S., (2017). HIV-1 Sequence Necessary and Sufficient to Package Non-viral RNAs into HIV-1 Particles. *J. Mol. Biol.* **429**, 2542–2555. https://doi.org/10.1016/j.jmb.2017.06.018.
276. Blissenbach, M., Grewe, B., Hoffmann, B., Brandt, S., Überla, K., (2010). Nuclear RNA Export and Packaging Functions of HIV-1 Rev Revisited. *J. Virol.* **84**, 6598–6604. https://doi.org/10.1128/JVI.02264-09.
277. Rizvi, T.A., Schmidt, R.D., Lew, K.A., Keeling, M.E., (1996). Rev/RRE-Independent Mason-Pfizer Monkey Virus Constitutive Transport Element-Dependent Propagation of SIVmac239 Vectors Using a Single Round of Replication Assay. *Virology* **222**, 457–463. https://doi.org/10.1006/viro.1996.0444.
278. Barajas, B.C., Tanaka, M., Robinson, B.A., Phuong, D.J., Chutiraka, K., Reed, J.C., Lingappa, J.R., (2018). Identifying the assembly intermediate in which Gag first associates with unspliced HIV-1 RNA suggests a novel model for HIV-1 RNA packaging. *PLoS Pathog.* **14**, e1006977.
279. Reed, J.C., Westergreen, N., Barajas, B.C., Ressler, D.T. B., Phuong, D.J., Swain, J.V., Lingappa, V.R., Lingappa, J.R., (2018). Formation of RNA Granule-Derived Capsid Assembly Intermediates Appears To Be Conserved between Human Immunodeficiency Virus Type 1 and the Nonprimate Lentivirus Feline Immunodeficiency Virus. *J. Virol.* **92** https://doi.org/10.1128/JVI.01761-17.
280. Abrahamyan, L.G., Chatel-Chaix, L., Ajamian, L., Milev, M.P., Monette, A., Clément, J.-F., Song, R., Lehmann, M., et al., (2010). Novel Staufen1 ribonucleoproteins prevent formation of stress granules but favour encapsidation of HIV-1 genomic RNA. *J. Cell Sci.* **123**, 369–383. https://doi.org/10.1242/jcs.055897.
281. Cochrane, A.W., McNally, M.T., Mouland, A.J., (2006). The retrovirus RNA trafficking granule: from birth to maturity. *Retrovirology* **3**, 18. https://doi.org/10.1186/1742-4690-3-18.
282. Mouland, A.J., Mercier, J., Luo, M., Bernier, L., DesGroseillers, L., Cohen, É.A., (2000). The Double-Stranded RNA-Binding Protein Staufen Is Incorporated in Human Immunodeficiency Virus Type 1: Evidence for a Role in Genomic RNA Encapsidation. *J. Virol.* **74**, 5441–5451.
283. Deng, Y., Hammond, J.A., Pauszek, R., Ozog, S., Chai, I., Rabuck-Gibbons, J., Lamichhane, R., Henderson, S.C., et al., (2021). Discrimination between Functional and Non-functional Cellular Gag Complexes involved in HIV-1 Assembly. *J. Mol. Biol.* **433**, https://doi.org/10.1016/j.jmb.2021.166842 166842.
284. Canaani, E., Helm, K.V., Duesberg, P., (1973). Evidence for 30–40S RNA as precursor of the 60–70S RNA of Rous sarcoma virus. *PNAS* **70**, 401–405. https://doi.org/10.1073/pnas.70.2.401.
285. Mangel, W.F., Delius, H., Duesberg, P.H., (1974). Structure and molecular weight of the 60–70S RNA and the 30–40S RNA of the Rous sarcoma virus. *PNAS* **71**, 4541–4545. https://doi.org/10.1073/pnas.71.11.4541.
286. Skripkin, E., Paillart, J.C., Marquet, R., Ehresmann, B., Ehresmann, C., (1994). Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro. *PNAS* **91**, 4945–4949.
287. Sakuragi, J., Sakuragi, S., Shioda, T., (2007). Minimal Region Sufficient for Genome Dimerization in the Human Immunodeficiency Virus Type 1 Virion and Its Potential Roles in the Early Stages of Viral Replication. *J. Virol.* **81**, 7985–7992. https://doi.org/10.1128/JVI.00429-07.
288. Murti, K.G., 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**, 411–419.
289. Sakuragi, J., Ode, H., Sakuragi, S., Shioda, T., Sato, H., (2012). A proposal for a new HIV-1 DLS structural model. *Nucleic Acids Res.* **40**, 5012–5022. https://doi.org/10.1093/nar/gks156.
290. Feng, Y.X., Copeland, T.D., Henderson, L.E., Gorelick, R. J., Bosche, W.J., Levin, J.G., Rein, A., (1996). HIV-1 nucleocapsid protein induces “maturation” of dimeric retroviral RNA in vitro. *PNAS* **93**, 7577–7581. https://doi.org/10.1073/pnas.93.15.7577.
291. Fu, W., Gorelick, R.J., Rein, A., (1994). Characterization of human immunodeficiency virus type 1 dimeric RNA from wild-type and protease-defective virions. *J. Virol.* **68**, 5013–5018. https://doi.org/10.1128/JVI.68.8.5013-5018.1994.
292. Chen, J., Rahman, S.A., Nikolaitchik, O.A., Grunwald, D., Sardo, L., Burdick, R.C., Plisov, S., Liang, E., et al., (2016). HIV-1 RNA genome dimerizes on the plasma membrane in the presence of Gag protein. *PNAS* **113**, E201–E208. https://doi.org/10.1073/pnas.1518572113.
293. Kharytonchyk, S.A., Kireyeva, A.I., Osipovich, A.B., Fomin, I.K., (2005). Evidence for preferential copackaging of Moloney murine leukemia virus genomic RNAs transcribed in the same chromosomal site. *Retrovirology* **2**, 3. https://doi.org/10.1186/1742-4690-2-3.
294. Flynn, J.A., An, W., King, S.R., Telesnitsky, A., (2004). Nonrandom Dimerization of Murine Leukemia Virus Genomic RNAs. *J. Virol.* **78**, 12129–12139. https://doi.org/10.1128/JVI.78.22.12129-12139.2004.

295. Chen, E.C., Maldonado, R.J.K., Parent, L.J., (2021). Visualizing Rous Sarcoma Virus Genomic RNA Dimerization in the Nucleus, Cytoplasm, and at the Plasma Membrane. *Viruses* **13**, 903. https://doi.org/10.3390/v13050903.
296. Marquet, R., Baudin, F., Gabus, C., Darlix, J.L., Mougel, M., Ehresmann, C., Ehresmann, B., (1991). Dimerization of human immunodeficiency virus (type 1) RNA: stimulation by cations and possible mechanism. *Nucleic Acids Res.* **19**, 2349–2357. https://doi.org/10.1093/nar/19.9.2349.
297. Paillart, J.-C., Shehu-Xhilaga, M., Marquet, R., Mak, J., (2004). Dimerization of retroviral RNA genomes: an inseparable pair. *Nat. Rev. Microbiol.* **2**, 461–472. https://doi.org/10.1038/nrmicro903.
298. Huthoff, H., Berkhout, B., (2002). Multiple secondary structure rearrangements during HIV-1 RNA dimerization. *Biochemistry* **41**, 10439–10445. https://doi.org/10.1021/bi025993n.
299. Rein, A., (2010). Nucleic acid chaperone activity of retroviral Gag proteins. *RNA Biol.* **7**, 700–705.
300. Aduri, R., Briggs, K.T., Gorelick, R.J., Marino, J.P., (2013). Molecular determinants of HIV-1 NCp7 chaperone activity in maturation of the HIV-1 dimerization initiation site. *Nucleic Acids Res.* **41**, 2565–2580. https://doi.org/10.1093/nar/gks1350.
301. Sun, X., Zhang, Q., Al-Hashimi, H.M., (2007). Resolving fast and slow motions in the internal loop containing stem-loop 1 of HIV-1 that are modulated by Mg2+ binding: role in the kissing-duplex structural transition. *Nucleic Acids Res.* **35**, 1698–1713. https://doi.org/10.1093/nar/gkm020.
302. Muriaux, D., De Rocquigny, H., Roques, B.P., Paoletti, J., (1996). NCp7 activates HIV-1Lai RNA dimerization by converting a transient loop-loop complex into a stable dimer. *J. Biol. Chem.* **271**, 33686–33692. https://doi.org/10.1074/jbc.271.52.33686.
303. Feng, Y.X., Campbell, S., Harvin, D., Ehresmann, B., Ehresmann, C., Rein, A., (1999). The human immunodeficiency virus type 1 Gag polyprotein has nucleic acid chaperone activity: possible role in dimerization of genomic RNA and placement of tRNA on the primer binding site. *J. Virol.* **73**, 4251–4256. https://doi.org/10.1128/JVI.73.5.4251-4256.1999.
304. Cruceanu, M., Urbaneja, M.A., Hixson, C.V., Johnson, D. G., Datta, S.A., Fivash, M.J., Stephen, A.G., Fisher, R.J., et al., (2006). Nucleic acid binding and chaperone properties of HIV-1 Gag and nucleocapsid proteins. *Nucleic Acids Res.* **34**, 593–605. https://doi.org/10.1093/nar/gkj458.
305. Gilmer, O., Mailler, E., Paillart, J.-C., Mouhand, A., Tisné, C., Mak, J., Smyth, R.P., Marquet, R., et al., (2022). Structural maturation of the HIV-1 RNA 5' untranslated region by Pr55Gag and its maturation products. *RNA Biol.* **19**, 191–205. https://doi.org/10.1080/15476286.2021.2021677.
306. Jalalirad, M., Laughrea, M., (2010). Formation of immature and mature genomic RNA dimers in wild-type and protease-inactive HIV-1: differential roles of the Gag polyprotein, nucleocapsid proteins NCp15, NCp9, NCp7, and the dimerization initiation site. *Virology* **407**, 225–236. https://doi.org/10.1016/j.virol.2010.08.013.
307. Clever, J.L., Wong, M.L., Parslow, T.G., (1996). Requirements for kissing-loop-mediated dimerization of human immunodeficiency virus RNA. *J. Virol.* **70**, 5902–5908.
308. Kenyon, J.C., Tanner, S.J., Legiewicz, M., Phillip, P.S., Rizvi, T.A., Le Grice, S.F.J., Lever, A.M.L., (2011). SHAPE analysis of the FIV Leader RNA reveals a structural switch potentially controlling viral packaging and genome dimerization. *Nucleic Acids Res.* **39**, 6692–6704. https://doi.org/10.1093/nar/gkr252.
309. Muriaux, D., Girard, P.M., Bonnet-Mathonière, B., Paoletti, J., (1995). Dimerization of HIV-1Lai RNA at low ionic strength. An autocomplementary sequence in the 5' leader region is evidenced by an antisense oligonucleotide. *J. Biol. Chem.* **270**, 8209–8216. https://doi.org/10.1074/jbc.270.14.8209.
310. Ly, H., Parslow, T.G., (2002). Bipartite Signal for Genomic RNA Dimerization in Moloney Murine Leukemia Virus. *J. Virol.* **76**, 3135–3144. https://doi.org/10.1128/JVI.76.7.3135-3144.2002.
311. Kim, C.-H., Tinoco, I., (2000). A retroviral RNA kissing complex containing only two GśC base pairs. *PNAS* **97**, 9396–9401. https://doi.org/10.1073/pnas.170283697.
312. Konings, D.A., Nash, M.A., Maizel, J.V., Arlinghaus, R.B., (1992). Novel GACG-hairpin pair motif in the 5' untranslated region of type C retroviruses related to murine leukemia virus. *J. Virol.* **66**, 632–640. https://doi.org/10.1128/JVI.66.2.632-640.1992.
313. Greatorex, J.S., Laisse, V., Dockhelar, M.C., Lever, A.M., (1996). Sequences involved in the dimerisation of human T cell leukaemia virus type-1 RNA. *Nucleic Acids Res.* **24**, 2919–2923.
314. Monie, T., Greatorex, J., Lever, A.M., (2001). Oligonucleotide mapping of the core genomic RNA dimer linkage in human T-cell leukaemia virus type-1. *Virus Res.* **78**, 45–56. https://doi.org/10.1016/s0168-1702(01)00283-0.
315. Liu, S., Kaddis Maldonado, R., Rye-McCurdy, T., Binkley, C., Bah, A., Chen, E.C., Rice, B.L., Parent, L.J., et al., (2020). Rous Sarcoma Virus Genomic RNA Dimerization Capability In Vitro Is Not a Prerequisite for Viral Infectivity. *Viruses* **12**, 568. https://doi.org/10.3390/v12050568.
316. Hill, M.K., Shehu-Xhilaga, M., Campbell, S.M., Poumbourios, P., Crowe, S.M., Mak, J., (2003). The dimer initiation sequence stem-loop of human immunodeficiency virus type 1 is dispensable for viral replication in peripheral blood mononuclear cells. *J. Virol.* **77**, 8329–8335. https://doi.org/10.1128/jvi.77.15.8329-8335.2003.
317. Laughrea, M., Jetté, L., Mak, J., Kleiman, L., Liang, C., Wainberg, M.A., (1997). Mutations in the kissing-loop hairpin of human immunodeficiency virus type 1 reduce viral infectivity as well as genomic RNA packaging and dimerization. *J. Virol.* **71**, 3397–3406.
318. Sakuragi, J., Shioda, T., Panganiban, A.T., (2001). Duplication of the primary encapsidation and dimer linkage region of human immunodeficiency virus type 1 RNA results in the appearance of monomeric RNA in virions. *J. Virol.* **75**, 2557–2565. https://doi.org/10.1128/JVI.75.6.2557-2565.2001.
319. Paillart, J.C., Westhof, E., Ehresmann, C., Ehresmann, B., Marquet, R., (1997). Non-canonical interactions in a kissing loop complex: the dimerization initiation site of HIV-1 genomic RNA. *J. Mol. Biol.* **270**, 36–49. https://doi.org/10.1006/jmbi.1997.1096.

320. Sakuragi, J.-I., Ueda, S., Iwamoto, A., Shioda, T., (2003). Possible Role of Dimerization in Human Immunodeficiency Virus Type 1 Genome RNA Packaging. *J. Virol.* **77**, 4060–4069. https://doi.org/10.1128/JVI.77.7.4060-4069.2003.
321. Parent, L.J., Cairns, T.M., Albert, J.A., Wilson, C.B., Wills, J.W., Craven, R.C., (2000). RNA dimerization defect in a Rous sarcoma virus matrix mutant. *J. Virol.* **74**, 164–172. https://doi.org/10.1128/jvi.74.1.164-172.2000.
322. Onafuwa-Nuga, A., Telesnitsky, A., (2009). The Remarkable Frequency of Human Immunodeficiency Virus Type 1 Genetic Recombination. *Microbiol. Mol. Biol. Rev.* **73**, 451–480. https://doi.org/10.1128/MMBR.00012-09.
323. Konings, F.A.J., Burda, S.T., Urbanski, M.M., Zhong, P., Nadas, A., Nyambi, P.N., (2006). Human immunodeficiency virus type 1 (HIV-1) circulating recombinant form 02_AG (CRF02_AG) has a higher in vitro replicative capacity than its parental subtypes A and G. *J. Med. Virol.* **78**, 523–534. https://doi.org/10.1002/jmv.20572.
324. Kozal, M.J., Shah, N., Shen, N., Yang, R., Fucini, R., Merigan, T.C., Richman, D.D., Morris, D., et al., (1996). Extensive polymorphisms observed in HIV-1 clade B protease gene using high-density oligonucleotide arrays. *Nat. Med.* **2**, 753–759. https://doi.org/10.1038/nm0796-753.
325. Njai, H.F., Gali, Y., Vanham, G., Clybergh, C., Jennes, W., Vidal, N., Butel, C., Mpoudi-Ngolle, E., et al., (2006). The predominance of Human Immunodeficiency Virus type 1 (HIV-1) circulating recombinant form 02 (CRF02_AG) in West Central Africa may be related to its replicative fitness. *Retrovirology* **3**, 40. https://doi.org/10.1186/1742-4690-3-40.
326. Chin, M.P.S., Chen, J., Nikolaitchik, O.A., Hu, W.-S., (2007). Molecular determinants of HIV-1 intersubtype recombination potential. *Virology* **363**, 437–446. https://doi.org/10.1016/j.virol.2007.01.034.
327. Abbink, T.E.M., Ooms, M., Haasnoot, P.C.J., Berkhout, B., (2005). The HIV-1 leader RNA conformational switch regulates RNA dimerization but does not regulate mRNA translation. *Biochemistry* **44**, 9058–9066. https://doi.org/10.1021/bi0502588.
328. Abbink, T.E.M., 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. https://doi.org/10.1074/jbc.M210291200.
329. Huthoff, H., Berkhout, B., (2001). Two alternating structures of the HIV-1 leader RNA. *RNA* **7**, 143–157.
330. Damgaard, C.K., Andersen, E.S., Knudsen, B., Gorodkin, J., Kjems, J., (2004). RNA interactions in the 5' region of the HIV-1 genome. *J. Mol. Biol.* **336**, 369–379. https://doi.org/10.1016/j.jmb.2003.12.010.
331. Kenyon, J.C., Ghazawi, A., Cheung, W.K.S., Phillip, P.S., Rizvi, T.A., Lever, A.M.L., (2008). The secondary structure of the 5' end of the FIV genome reveals a long-range interaction between R/U5 and gag sequences, and a large, stable stem-loop. *RNA* **14**, 2597–2608. https://doi.org/10.1261/rna.1284908.
332. Butsch, M., Boris-Lawrie, K., (2000). Translation is not required To generate virion precursor RNA in human immunodeficiency virus type 1-infected T cells. *J. Virol.* **74**, 11531–11537. https://doi.org/10.1128/jvi.74.24.11531-11537.2000.
333. Dorman, N., Lever, A., (2000). Comparison of Viral Genomic RNA Sorting Mechanisms in Human Immunodeficiency Virus Type 1 (HIV-1), HIV-2, and Moloney Murine Leukemia Virus. *J. Virol.* **74**, 11413–11417.
334. Poon, D.T.K., Chertova, E.N., Ott, D.E., (2002). Human immunodeficiency virus type 1 preferentially encapsidates genomic RNAs that encode Pr55(Gag): functional linkage between translation and RNA packaging. *Virology* **293**, 368–378. https://doi.org/10.1006/viro.2001.1283.
335. Nikolaitchik, O., Rhodes, T.D., Ott, D., Hu, W.-S., (2006). Effects of Mutations in the Human Immunodeficiency Virus Type 1 gag Gene on RNA Packaging and Recombination. *J. Virol.* **80**, 4691–4697. https://doi.org/10.1128/JVI.80.10.4691-4697.2006.
336. Boeras, I., Seufzer, B., Brady, S., Rendahl, A., Heng, X., Boris-Lawrie, K., (2017). The basal translation rate of authentic HIV-1 RNA is regulated by 5'UTR nt-pairings at junction of R and U5. *Sci. Rep.* **7**, 6902. https://doi.org/10.1038/s41598-017-06883-9.
337. Chen, J., Liu, Y., Wu, B., Nikolaitchik, O.A., Mohan, P.R., Chen, J., Pathak, V.K., Hu, W.-S., (2020). Visualizing the translation and packaging of HIV-1 full-length RNA. *PNAS* **117**, 6145–6155. https://doi.org/10.1073/pnas.1917590117.
338. Kharytonchyk, S., Monti, S., Smaldino, P.J., Van, V., Bolden, N.C., Brown, J.D., Russo, E., Swanson, C., et al., (2016). Transcriptional start site heterogeneity modulates the structure and function of the HIV-1 genome. *PNAS* **113**, 13378–13383. https://doi.org/10.1073/pnas.1616627113.
339. Starcich, B., Ratner, L., Josephs, S.F., Okamoto, T., Gallo, R.C., Wong-Staal, F., (1985). Characterization of long terminal repeat sequences of HTLV-III. *Science* **227**, 538–540. https://doi.org/10.1126/science.2981438.
340. Muesing, M.A., Smith, D.H., Cabradilla, C.D., Benton, C. V., Lasky, L.A., Capon, D.J., (1985). Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovirus. *Nature* **313**, 450–458. https://doi.org/10.1038/313450a0.
341. Masuda, T., Sato, Y., Huang, Y.-L., Koi, S., Takahata, T., Hasegawa, A., Kawai, G., Kannagi, M., (2015). Fate of HIV-1 cDNA intermediates during reverse transcription is dictated by transcription initiation site of virus genomic RNA. *Sci. Rep.* **5**, 1–15. https://doi.org/10.1038/srep17680.
342. Menees, T.M., Müller, B., Kräusslich, H.-G., (2007). The major 5' end of HIV type 1 RNA corresponds to G456. *AIDS Res. Hum. Retroviruses* **23**, 1042–1048. https://doi.org/10.1089/aid.2006.0275.
343. Rawson, J.M.O., Nikolaitchik, O.A., Shakya, S., Keele, B. F., Pathak, V.K., Hu, W.-S., (2022). Transcription Start Site Heterogeneity and Preferential Packaging of Specific Full-Length RNA Species Are Conserved Features of Primate Lentiviruses. *Microbiol Spectr.* **10**, e0105322.
344. Ding, P., Kharytonchyk, S., Kuo, N., Cannistraci, E., Flores, H., Chaudhary, R., Sarkar, M., Dong, X., et al., (2021). 5'-Cap sequestration is an essential determinant of HIV-1 genome packaging. *Proc. Natl. Acad. Sci. U S A* **118** https://doi.org/10.1073/pnas.2112475118.
345. Ding, P., Summers, M.F., (2022). Sequestering the 5′-cap for viral RNA packaging 2200104 *Bioessays*.

346. Pereira-Montecinos, C., Toro-Ascuy, D., Ananías-Sáez, C., Gaete-Argel, A., Rojas-Fuentes, C., Riquelme-Barrios, S., Rojas-Araya, B., García-de-Gracia, F., et al., (2022). Epitranscriptomic regulation of HIV-1 full-length RNA packaging. *Nucleic Acids Res.* **50**, 2302–2318. https://doi.org/10.1093/nar/gkac062.
347. Levin, J.G., Rosenak, M.J., (1976). Synthesis of murine leukemia virus proteins associated with virions assembled in actinomycin D-treated cells: evidence for persistence of viral messenger RNA. *PNAS* **73**, 1154–1158. https://doi.org/10.1073/pnas.73.4.1154.
348. Levin, J.G., Grimley, P.M., Ramseur, J.M., Berezesky, I. K., (1974). Deficiency of 60 to 70S RNA in Murine Leukemia Virus Particles Assembled in Cells Treated with Actinomycin D. *J. Virol.* **14**, 152–161.
349. Griffin, S.D., Allen, J.F., Lever, A.M., (2001). The major human immunodeficiency virus type 2 (HIV-2) packaging signal is present on all HIV-2 RNA species: cotranslational RNA encapsidation and limitation of Gag protein confer specificity. *J. Virol.* **75**, 12058–12069. https://doi.org/10.1128/JVI.75.24.12058-12069.2001.
350. Kaye, J.F., Lever, A.M., (1999). Human immunodeficiency virus types 1 and 2 differ in the predominant mechanism used for selection of genomic RNA for encapsidation. *J. Virol.* **73**, 3023–3031.
351. Ni, N., Nikolaitchik, O.A., Dilley, K.A., Chen, J., Galli, A., Fu, W., Prasad, V.V.S.P., Ptak, R.G., et al., (2011). Mechanisms of human immunodeficiency virus type 2 RNA packaging: efficient trans packaging and selection of RNA copackaging partners. *J. Virol.* **85**, 7603–7612. https://doi.org/10.1128/JVI.00562-11.