RESEARCH ARTICLE

MMTV RNA packaging requires an extended long-range interaction for productive Gag binding to packaging signals

Suresha G. Prabhu[1], **Vineeta N. Pillai**[1], **Lizna Mohamed Ali**[1], **Valérie Vivet-Boudou**[2], **Akhil Chameettachal**[1¤], **Serena Bernacchi**[2], **Farah Mustafa**[3,4,5], **Roland Marquet**[2]*, **Tahir A. Rizvi**[1,4,5]*

1 Department of Microbiology & Immunology, College of Medicine and Health Sciences (CMHS), United Arab Emirates University (UAEU), Al Ain, United Arab Emirates, **2** Université de Strasbourg, CNRS, Architecture et Réactivité de l'ARN, UPR 9002, Strasbourg, France, **3** Department of Biochemistry & Molecular Biology, College of Medicine and Health Sciences (CMHS), United Arab Emirates University (UAEU), Al Ain, United Arab Emirates, **4** Zayed bin Sultan Center for Health Sciences, United Arab Emirates University, Al Ain, United Arab Emirates, **5** ASPIRE Research Institute in Precision Medicine, Abu Dhabi, United Arab Emirates

¤ Current address: Viral Recombination Section, HIV Dynamics and Replication Program, National Cancer Institute at Frederick, Frederick, Maryland, United States of America
* r.marquet@ibmc-cnrs.unistra.fr (RM); tarizvi@uaeu.ac.ae (TAR)

Abstract

The packaging of genomic RNA (gRNA) into retroviral particles relies on the specific recognition by the Gag precursor of packaging signals (*Psi*), which maintain a complex secondary structure through long-range interactions (LRIs). However, it remains unclear whether the binding of Gag to *Psi* alone is enough to promote RNA packaging and what role LRIs play in this process. Using mouse mammary tumor virus (MMTV), we investigated the effects of mutations in 4 proposed LRIs on gRNA structure and function. Our findings revealed the presence of an unsuspected extended LRI, and hSHAPE revealed that maintaining a wild-type–like *Psi* structure is crucial for efficient packaging. Surprisingly, filter-binding assays demonstrated that most mutants, regardless of their packaging capability, exhibited significant binding to Pr77^{Gag}, suggesting that Gag binding to *Psi* is insufficient for efficient packaging. Footprinting experiments indicated that efficient RNA packaging is promoted when Pr77^{Gag} binds to 2 specific sites within *Psi*, whereas binding elsewhere in *Psi* does not lead to efficient packaging. Taken together, our results suggest that the 3D structure of the *Psi*/Pr77^{Gag} complex regulates the assembly of viral particles around gRNA, enabling effective discrimination against other viral and cellular RNAs that may also bind Gag efficiently.

Introduction

The vast majority of virus species selectively package their DNA or RNA genome into viral particles. While DNA viruses typically assemble a procapsid first, followed by packaging of the genome in an energy-dependent process, many RNA viruses, including significant human

pathogens such as retroviruses, influenza viruses, and coronaviruses, undergo concerted processes of genomic RNA (gRNA) packaging and viral assembly [1,2].

In retroviruses, the selective packaging of gRNA, which exists as a dimer of single-stranded RNA molecules noncovalently linked near their 5′-end [3,4], from the cellular environment where it represents only a small fraction (around 1%) of total mRNA, is facilitated by the recognition of *cis*-acting sequences known as packaging signal (*Psi; ψ*) typically located at the 5′ end of the retroviral genome, sometimes extending into the Gag open reading frame (ORF) [4–27]. While the nucleocapsid (NC) domain of Gag is crucial for selective packaging [12–17,21,22,28,29], in the case of human immunodeficiency virus (HIV-1), the matrix (MA) domain [30,31], capsid domain (CA; [32,33]), p6 domain [28], and spacer peptide sp1 [31,34] also participate in the specific recognition of *Psi* by Gag.

These retroviral *Psi* sequences adopt complex higher-order structures containing various structural motifs that are essential for Gag binding and retroviral RNA packaging [4,9,10,12,14–18,21,26,35–46]. Notably, *Psi* sequences in all retroviruses feature long-range interactions (LRIs) that maintain their overall secondary structure and, in several instances, are critical for Gag binding and gRNA packaging [3,9,10,26,36,43,45–55]. It is important to note that *Psi* sequences are significantly longer than typical protein-binding sites, and without a high-resolution 3D structure of the *Psi*/Gag complex, the precise role of LRIs in these functions remains unclear. Additionally, although it is well established that *Psi*/Gag interactions are necessary for retroviral gRNA packaging [4,12,14–19,21], it is not yet clear whether these interactions alone are sufficient for this process to occur.

In this study, we tackled these questions using mouse mammary tumor virus (MMTV) as a model. MMTV, a *betaretrovirus*, induces breast cancer and, in certain instances, T-cell lymphomas in mice [56–58]. Unlike the majority of retroviruses, including lentiviruses, MMTV assembles in the cytoplasm into immature particles that later translocate to the plasma membrane for budding [59,60]. The MMTV *Psi* resides within the 5′ UTR and the first 120 nts of the *gag* gene [61], and recent studies have elucidated the structure–function relationships of several structural motifs within *Psi* [36,41,53]. Based on SHAPE (selective 2′ hydroxyl acylation analyzed by primer extension) data, the current secondary structure model of MMTV *Psi* suggests the presence of 4 evolutionary conserved LRIs [36,41,53]: LRIs I-III involve complementary base-pairing between U5 and Gag sequences that are approximately 291 nucleotides apart, while LRI IV is formed by complementary base-pairing of sequences within U5 that are approximately 190 nucleotides apart (Fig 1).

Briefly, we introduced mutations that disrupted LRIs I-IV and compensatory mutations designed to restore them, then analyzed the impact of these mutations on gRNA packaging, gRNA structure, and Pr77^Gag binding to *Psi*. Our findings indicate the presence of only 2 LRIs, which together form an extended LRI that significantly influences gRNA packaging, prompting us to propose a revised secondary structure model for MMTV *Psi*. Maintaining a *Psi* structure akin to the wild type (WT) emerges as crucial for efficient packaging. Surprisingly, most mutants displayed notable binding to Pr77^Gag regardless of their packaging ability, suggesting that Gag binding to *Psi* alone is insufficient for efficient packaging. Intriguingly, SHAPE footprinting experiments revealed that efficient RNA packaging correlated with Pr77^Gag binding to 2 specific sites within *Psi*, whereas binding elsewhere in *Psi* did not result in efficient packaging. These findings underscore the necessity for Pr77^Gag to bind to specific nucleotides within the correct structural context of *Psi* and support a model wherein the 3D structure of the *Psi*/Pr77^Gag complex governs the assembly of viral particles around gRNA, thereby facilitating effective discrimination against other viral and cellular RNA species that may also bind Gag efficiently.
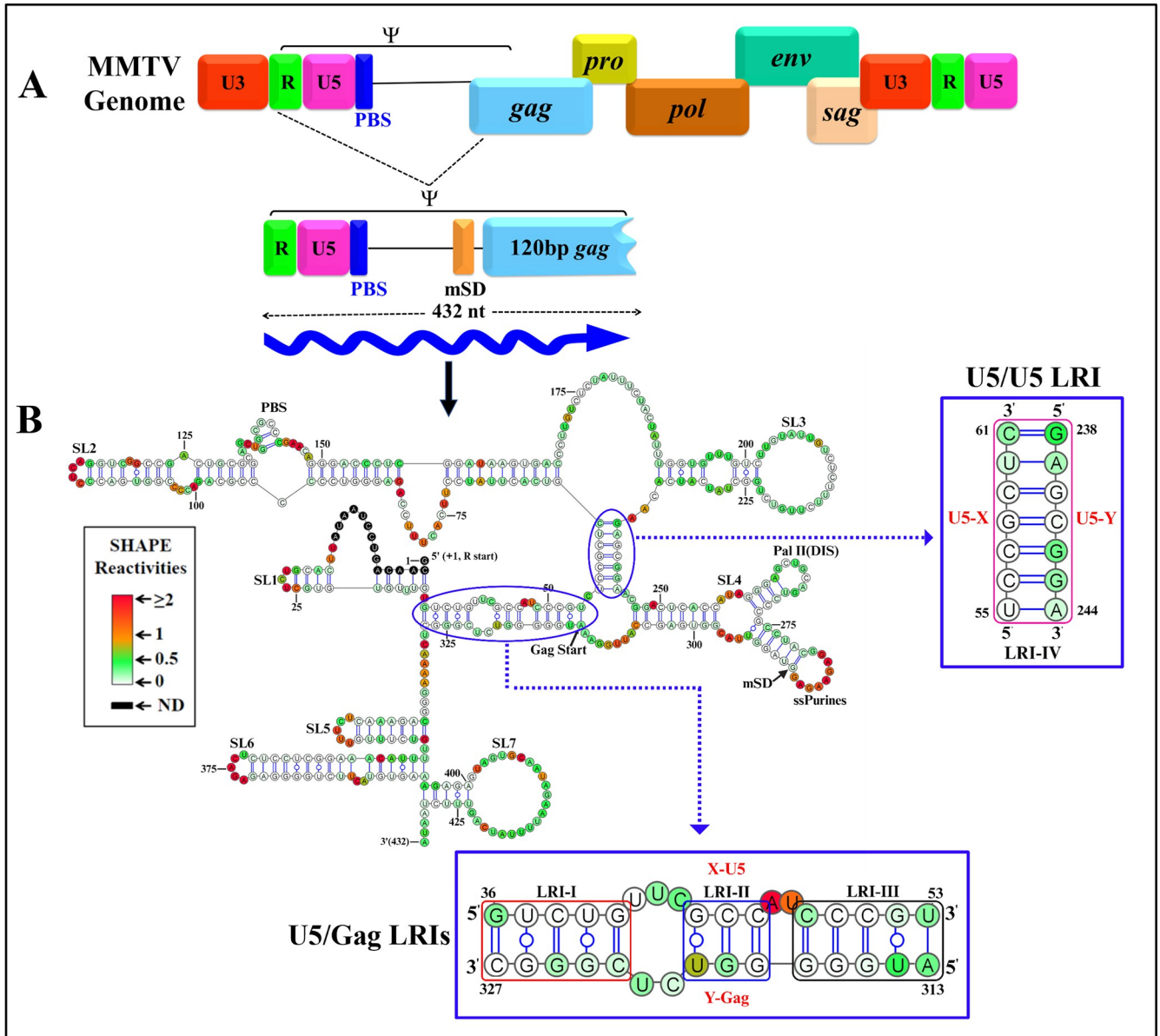
Fig 1. Schematic representation of the MMTV genome depicting the 5′ end harboring RNA packaging sequences (Ψ), along with its current RNA secondary structure model. (**A**) Genome organization of MMTV and its RNA packaging signal (Ψ) beginning from the R region to 120 nucleotides of Gag, denoted by a blue wavy arrow. (**B**) hSHAPE-validated RNA secondary structure of MMTV packaging sequences located at the 5′ end, including LRIs-I-IV depicted in blue insets. Nucleotides are color-coded according to hSHAPE reactivity, with features annotated, such as stem-loops (SL1-7), PBS, DIS, single-stranded purines (ssPurines), mSD, and Gag start site. DIS, dimerization initiation site; LRI, long-range interaction; MMTV, mouse mammary tumor virus; mSD, major splice donor; PBS, primer binding site.

Results

Experimental strategy

The three-plasmid genetic complementation assay. Conventional studies on retroviral RNA packaging using WT viruses are constrained by the involvement of the 5′ end of *gag* in *Psi*. This makes analyzing packaging using full-length WT viruses challenging, as mutations in *Psi* can affect the Gag ORF. While introducing point mutations is feasible, it often causes

unexpected *cis* effects like mRNA nuclear export alterations [62–66]. To address these limitations and study the effects of mutations in MMTV *Psi* on gRNA packaging and propagation, we used a three-plasmid genetic complementation assay (S1 Fig) [36,41,53,61,67]. This involved co-transfecting a transfer vector (DA024) containing necessary *cis*-acting sequences and a marker gene [67], a packaging construct (JA10) expressing *gag-pro-pol* genes [67], and a VSV-G expression plasmid (MD.G) [68]. This split genome strategy generates pseudotyped virus particles from JA10 and MD.G, while the transfer vector produces packageable RNA. The replication of packaged RNA is restricted to a single round, allowing monitoring of vector RNA propagation via the *hygromycin resistance* gene. RNA packaging efficiency is assessed through real-time quantitative PCR (RT-qPCR), correlating hygromycin-resistant colonies with viral RNA content. This approach allows manipulation of RNA secondary structure sequences involved in MMTV gRNA packaging without affecting *gag-pro-pol* sequences.

A TaqMan assay for estimating relative packaging efficiency (RPE). A new tailor-made MMTV TaqMan real-time qPCR assay was developed based on principles described earlier [61]. This new MMTV TaqMan assay was used along with a commercially available endogenous β-actin TaqMan assay to quantify both mutant and WT MMTV gRNAs expressed in the cytoplasm and packaged into virions. The amplification efficiency of MMTV and β-actin TaqMan assays was tested on serially diluted DA024 and a β-actin-expressing plasmid DNAs (S2A Fig; panels I and II). If the 2 assays have similar amplification efficiencies, the slope of the log input amount versus ΔCT should ideally be ≤0.1. Under our experimental conditions, the slope was 0.1076 (S2C Fig).

Structure–function analysis of the long-range interactions support a new secondary structure model of the 5′ region of MMTV gRNA

Effects of mutations designed to disrupt and restore the U5/Gag LRIs (LRI I, II, and III) on MMTV gRNA packaging and propagation. Considering the role of LRI sequences in anchoring and stabilizing RNA structures [10,35,39,52,54], we hypothesized that disrupting the base-pairing between the U5 and Gag sequences within LRI-I-IV could potentially negatively affect RNA packaging. To the contrary, stabilizing these interactions should restore structure and hence RNA packaging. To test this hypothesis, we first designed mutant SP101, in which the U•G wobble base pairs of LRI-I were disrupted (Fig 2A). Next, in mutant SP102, we restored 2 G•U wobble base pairing in LRI-I (Fig 2A).

Using our previously established three-plasmid genetic complementation assay [67], we evaluated the relative packaging efficiency (RPE) of the WT and LRI-I mutant RNAs by quantifying the amount of RNA packaged into the virions. To ensure the stable expression and successful transport of each vector RNA to the cytoplasm, nuclear and cytoplasmic RNA fractions were isolated from transfected cells and quantified. The quality of cell fractionation was assessed by the presence of unspliced β-actin mRNA in the nuclear and not in the cytoplasmic fractions [69]. Multiplex RT-PCR did not detect amplification of unspliced β-actin mRNA in the cytoplasmic fraction (Fig 2B; panel I), in contrast to the presence of 18S rRNA (Fig 2B; panel I) confirming the absence of nuclear leakage. Complementary DNAs (cDNAs) prepared from cytoplasmic RNA fractions and pelleted viral particles were amplified using specific primers (Fig 2B; panels II and III, respectively). Amplification of the desired region across all samples validated the efficient and stable expression as well as appropriate transport of transfer vector RNAs from the nucleus to the cytoplasm (Fig 2B; panel II). Finally, WT or mutant transfer vector RNAs in the cytoplasm and in the pelleted virus particles were quantified using RT-qPCR [36,41,53,61]. Part of the transfected supernatant was also used to infect HeLa CD4 + cells to evaluate the ability of the produced virions to transduce the packaged RNA into the
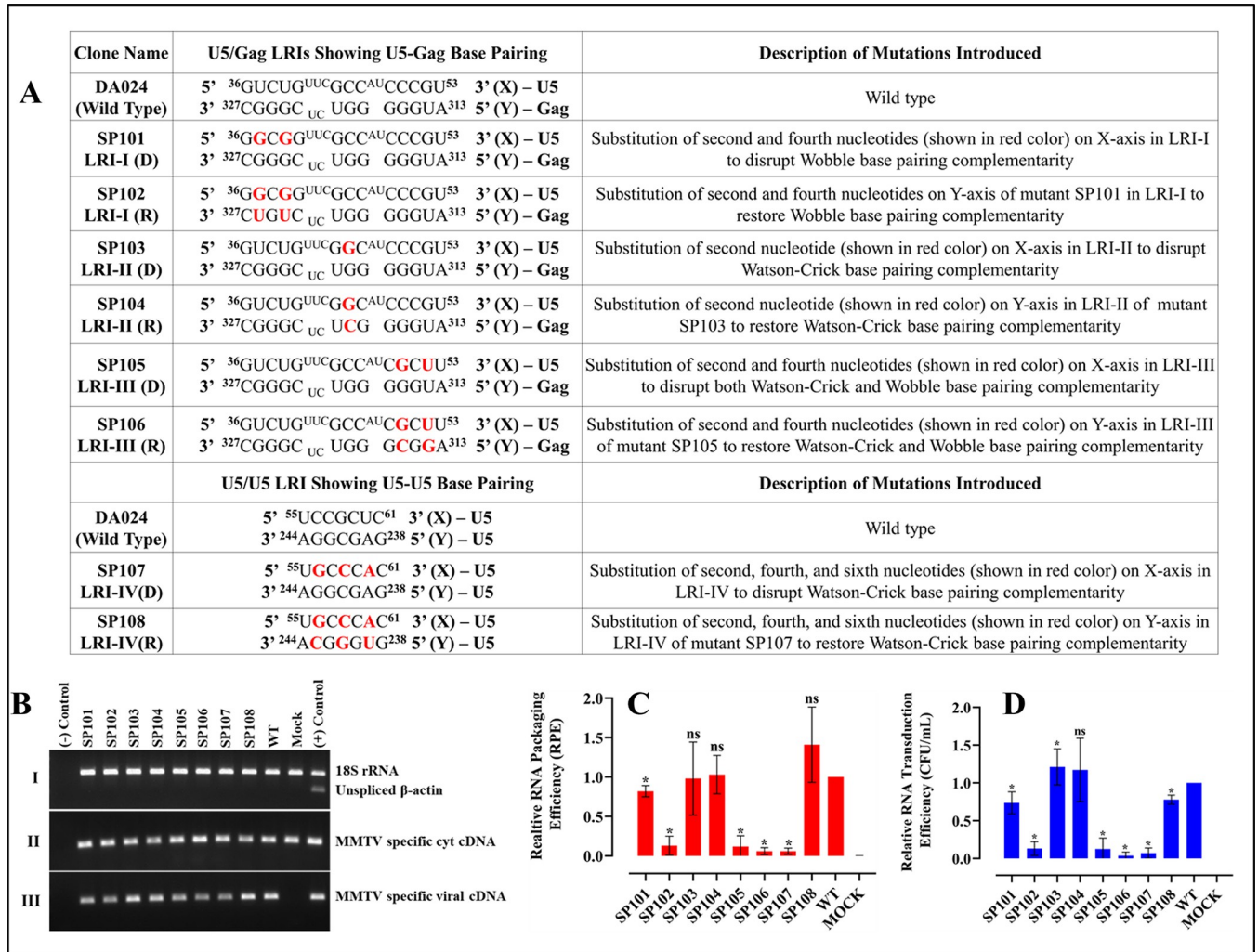
A

Clone Name	U5/Gag LRIs Showing U5-Gag Base Pairing	Description of Mutations Introduced
DA024 (Wild Type)	5' ^{36}GUCUG^{UUC}GCC^{AU}CCCGU^{53} 3' (X) – U5 3' ^{327}CGGGC $_{UC}$ UGG GGGUA^{313} 5' (Y) – Gag	Wild type
SP101 LRI-I (D)	5' ^{36}G**GCG**G^{UUC}GCC^{AU}CCCGU^{53} 3' (X) – U5 3' ^{327}CGGGC $_{UC}$ UGG GGGUA^{313} 5' (Y) – Gag	Substitution of second and fourth nucleotides (shown in red color) on X-axis in LRI-I to disrupt Wobble base pairing complementarity
SP102 LRI-I (R)	5' ^{36}G**GCG**G^{UUC}GCC^{AU}CCCGU^{53} 3' (X) – U5 3' ^{327}**CUGUC** $_{UC}$ UGG GGGUA^{313} 5' (Y) – Gag	Substitution of second and fourth nucleotides on Y-axis of mutant SP101 in LRI-I to restore Wobble base pairing complementarity
SP103 LRI-II (D)	5' ^{36}GUCUG^{UUC}G**G**C^{AU}CCCGU^{53} 3' (X) – U5 3' ^{327}CGGGC $_{UC}$ UGG GGGUA^{313} 5' (Y) – Gag	Substitution of second nucleotide (shown in red color) on X-axis in LRI-II to disrupt Watson-Crick base pairing complementarity
SP104 LRI-II (R)	5' ^{36}GUCUG^{UUC}G**G**C^{AU}CCCGU^{53} 3' (X) – U5 3' ^{327}CGGGC $_{UC}$ U**C**G GGGUA^{313} 5' (Y) – Gag	Substitution of second nucleotide (shown in red color) on Y-axis in LRI-II of mutant SP103 to restore Watson-Crick base pairing complementarity
SP105 LRI-III (D)	5' ^{36}GUCUG^{UUC}GCC^{AU}C**GC**U**U**^{53} 3' (X) – U5 3' ^{327}CGGGC $_{UC}$ UGG GGGUA^{313} 5' (Y) – Gag	Substitution of second and fourth nucleotides (shown in red color) on X-axis in LRI-III to disrupt both Watson-Crick and Wobble base pairing complementarity
SP106 LRI-III (R)	5' ^{36}GUCUG^{UUC}GCC^{AU}C**GC**U**U**^{53} 3' (X) – U5 3' ^{327}CGGGC $_{UC}$ UGG G**CGG**A^{313} 5' (Y) – Gag	Substitution of second and fourth nucleotides (shown in red color) on Y-axis in LRI-III of mutant SP105 to restore Watson-Crick and Wobble base pairing complementarity
	U5/U5 LRI Showing U5-U5 Base Pairing	**Description of Mutations Introduced**
DA024 (Wild Type)	5' ^{55}UCCGCUC^{61} 3' (X) – U5 3' ^{244}AGGCGAG^{238} 5' (Y) – U5	Wild type
SP107 LRI-IV(D)	5' ^{55}U**GCC**CA**C**^{61} 3' (X) – U5 3' ^{244}AGGCGAG^{238} 5' (Y) – U5	Substitution of second, fourth, and sixth nucleotides (shown in red color) on X-axis in LRI-IV to disrupt Watson-Crick base pairing complementarity
SP108 LRI-IV(R)	5' ^{55}U**GCC**CA**C**^{61} 3' (X) – U5 3' ^{244}A**CGG**G**UG**^{238} 5' (Y) – U5	Substitution of second, fourth, and sixth nucleotides (shown in red color) on Y-axis in LRI-IV of mutant SP107 to restore Watson-Crick base pairing complementarity

Fig 2. Role of LRIs-I-IV in MMTV gRNA packaging and propagation. (A) List of substitutions in the U5/Gag and U5/U5 LRIs, with mutations highlighted in red. (B) Representative gel images of the controls necessary for validating different aspects of the three-plasmid in vivo packaging and propagation assay: (I) multiplex amplification for nucleocytoplasmic fractionation technique, (II) PCR amplification for cDNAs prepared from cytoplasmic RNA fraction validating stability and nuclear export of transfer vector RNA, and (III) PCR amplification of packaged transfer vector RNA. (C) Packaging efficiency of mutant transfer vector RNAs relative to the wild type (DA024). (D) Relative propagation of MMTV transfer vector RNAs measured as normalized hygromycin resistant CFUs/ml for mutant transfer vectors compared to the wild type (DA024) vector. Mock samples contained only the transfer vector and no packaging construct. Data are presented as mean ± standard deviation from a minimum of 3 independent experiments performed in triplicates for RNA packaging (panel C) and in duplicates for RNA propagation (panel D). Differences compared to the wild-type were considered significant when $p < 0.05$ according to the nonparametric Mann–Whitney U test. The data underlying this Fig 2C and 2D can be found in S1 Data. CFU, colony-forming unit; gRNA, genomic RNA; LRI, long-range interaction; MMTV, mouse mammary tumor virus.

https://doi.org/10.1371/journal.pbio.3002827.g002

target cells. This was achieved by monitoring the emergence of hygromycin-resistant colonies following selection of the infected cultures with hygromycin B-containing medium.

Compared to the wild-type, SP101 RNA (Fig 2A) exhibited only moderate reductions in gRNA packaging (RPE = 0.82 ± 0.07; p-value < 0.05; Fig 2C) and propagation (CFU = 0.74± 0.14; p-value < 0.05; Fig 2D). Conversely, mutant SP102 showed 87% reduction in both RNA packaging (RPE = 0.13± 0.11; p-value < 0.05; Fig 2C) and propagation (CFU = 0.13 ± 0.08; p-value < 0.05; Fig 2D). These results indicate that mutations in the X and Y strands of LRI-I (Fig 2A) have cooperative (or additive) rather than compensatory effects suggesting that LRI-I does not exist.

Analysis of mutants SP103 and SP104, which were designed to disrupt and restore LRI-II, respectively (Fig 2A), did not allow to conclude about the existence of this LRI. Indeed, compared to the wild type, mutant SP103 did not demonstrate any defects in both RNA packaging (RPE = 0.98 ± 0.46; p-value > 0.05; Fig 2C) as well as in RNA propagation (CFU = 1.21 ± 0.23; p-value < 0.05; Fig 2D). Similarly, mutant SP104 revealed nearly identical results as the wild type for both RNA packaging and propagation (RPE = 1.03 ± 0.24; p-value > 0.05; CFU = 1.17 ± 0.41; p-value > 0.05; Fig 2C and 2D). These results indicate that the sequences mutated in SP103 and SP104 play no significant role in RNA packaging, irrespective of the existence, or not, of LRI-II.

Mutants SP105 and SP106 were designed to disrupt and restore LRI-III, respectively (Fig 2A). RNA packaging and propagation of these mutants were almost completely abolished, regardless of whether the base pairing sequence complementarity of LRI-III was disrupted (RPE = 0.12 ± 0.13; p-value < 0.05; CFU = 0.12 ± 0.14; p-value < 0.05; Fig 2C and 2D) or restored (RPE = 0.06 ± 0.03; p-value < 0.05; CFU = 0.03 ± 0.04; p-value < 0.05; Fig 2C and 2D). These results suggest that either the identity of the sequences forming LRI-III is crucial, or the proposed LRI-III may not actually exist.

Effects of mutations designed to disrupt and restore the U5/U5 LRI (LRI-IV) on MMTV gRNA packaging and propagation. To investigate the role of LRI-IV, the longest LRI in MMTV gRNA *Psi* (Fig 1B), on viral replication, we created mutants SP107 and SP108, designed to disrupt and restore 3 of the 7 base-pairs of LRI-IV, respectively (Fig 2A). SP107 revealed nearly complete abrogation of RNA packaging (RPE = 0.06 ± 0.03; p-value < 0.05; Fig 2C) and propagation (CFU = 0.07 ± 0.06; p-value < 0.05; Fig 2D). On the other hand, SP108 showed restoration of both RNA packaging (RPE = 1.41 ± 0.47; p-value > 0.05; Fig 2C) and RNA propagation (CFU = 0.78 ± 0.05; p-value < 0.05; Fig 2D) to WT levels. These results support the existence and biological significance of LRI-IV in MMTV replication. Furthermore, they highlight the importance of complementarity among nucleotides forming LRI-IV, rather than the sequence itself, for RNA packaging.

Mutations in LRIs do not affect RNA dimerization. Packaging of retroviral gRNA is closely linked to its dimerization [3,4,12,14,17,24,36,38,70,71]. Therefore, to test if the mutations were introduced in the LRIs affect RNA dimerization, 712 nucleotides from the 5′ end of the MMTV genome, encompassing wild type as well as mutant packaging sequences were cloned into a T7 promoter-containing plasmid (Fig 3A and 3B) and dimerization assays were performed using in vitro transcribed RNAs. Interestingly, both LRI packaging and non-packaging RNA mutants were found to dimerize at WT levels and showed no statistically significant differences, except for mutant SP106*i* (Fig 3C and 3D). Note that the RNA monomer species migrates slightly faster in the dimer condition than in the monomer condition, as the high ionic strength of the dimer buffer compacts the RNA. These results reveal that while the LRI destabilizing and restabilizing mutants affected RNA packaging, they did not adversely affect RNA dimerization, further validating that the effects being observed on RNA packaging were bona fide and not because of any effects on RNA dimerization.

Structure–function analyses of LRI mutants. Next, we investigated the RNA secondary structure of the 5′ end of the WT and selected LRI mutant MMTV genomes using hSHAPE [72–75]. To that goal, we treated in vitro transcribed RNAs corresponding to the 712 nucleotides at the 5′ end of the WT MMTV genome (SA35; Fig 3B) and LRI mutants (SP101*i*, 102*i*, SP105*i*-SP109*i* RNAs; Fig 3B) with BzCN. The resulting modifications of the flexible riboses were identified as stops during the extension of fluorescently labeled primers by reverse transcriptase and cDNA analysis by capillary electrophoresis. The SHAPE reactivity of each nucleotide obtained from 3 experiments using QuShape [76] (S1 Table) were utilized as
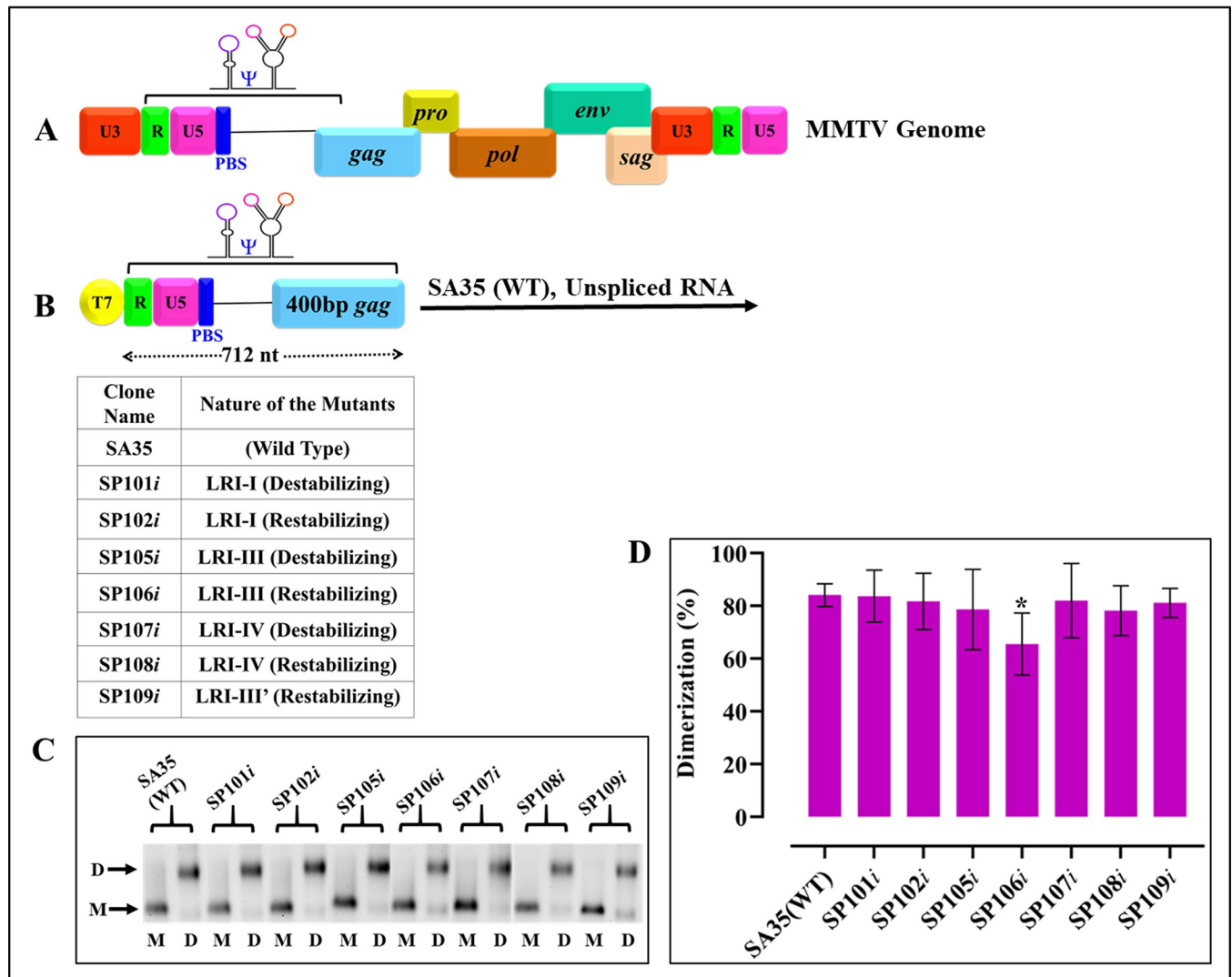
Fig 3. Dimerization of LRI mutants remains unaffected by the introduced mutations. (A) Schematic representation of the MMTV genome indicating the location of MMTV gRNA packaging determinants. (B) Illustration of the packageable vector RNA from R to 712 nucleotides expressed from a T7 expression plasmid, with the table detailing the names of the clones and the nature of the mutations. (C) Representative gel images displaying in vitro dimerization of WT (SA35) and LRI mutant RNAs in TBM buffer. M and D labels below the lanes indicate monomer and dimer buffers used for dimerization experiments. Monomeric and dimeric RNA species are denoted by letters M and D, respectively, on the gel's horizontal margin. Gels have been cropped as indicated by vertical white spaces to show the relevant areas only. (D) Histograms illustrating the dimerization efficiencies of mutant RNAs compared to WT (SA35) calculated through densitometric analysis of bands from 3 independent experiments. Dimerization efficiency was determined by dividing the intensity of the dimeric RNA band by the intensity of the band from the total RNA (i.e., sum of dimer and monomer bands). No statistically significant differences (p-values > 0.05) were observed in the ability of the mutant clones to dimerize when compared to the WT (SA35) according to the nonparametric Mann–Whitney U test, except for mutant SP106i ($p < 0.05$). The data underlying this Fig 3D can be found in S1 Data. gRNA, genomic RNA; LRI, long-range interaction; MMTV, mouse mammary tumor virus; WT, wild type.

https://doi.org/10.1371/journal.pbio.3002827.g003

pseudoenergy constraints to derive RNA secondary structure models for MMTV *Psi* of the WT and LRI mutants using the RNAStructure version 6.1 program [77].

As SP101, designed to disrupt LRI-I, exhibited minimal or no packaging defect, while SP102, designed to restore LRI-I, was severely compromised (Fig 2), we performed hSHAPE on these mutants. Stem loops (SLs) 1–4 and LRI-IV, which were present in the secondary structure of the WT MMTV gRNA proposed earlier [53], are maintained in the resulting RNA secondary structure model of SP101i, while LRIs-I-III are lost (Fig 4A). Unexpectedly, LRI-I
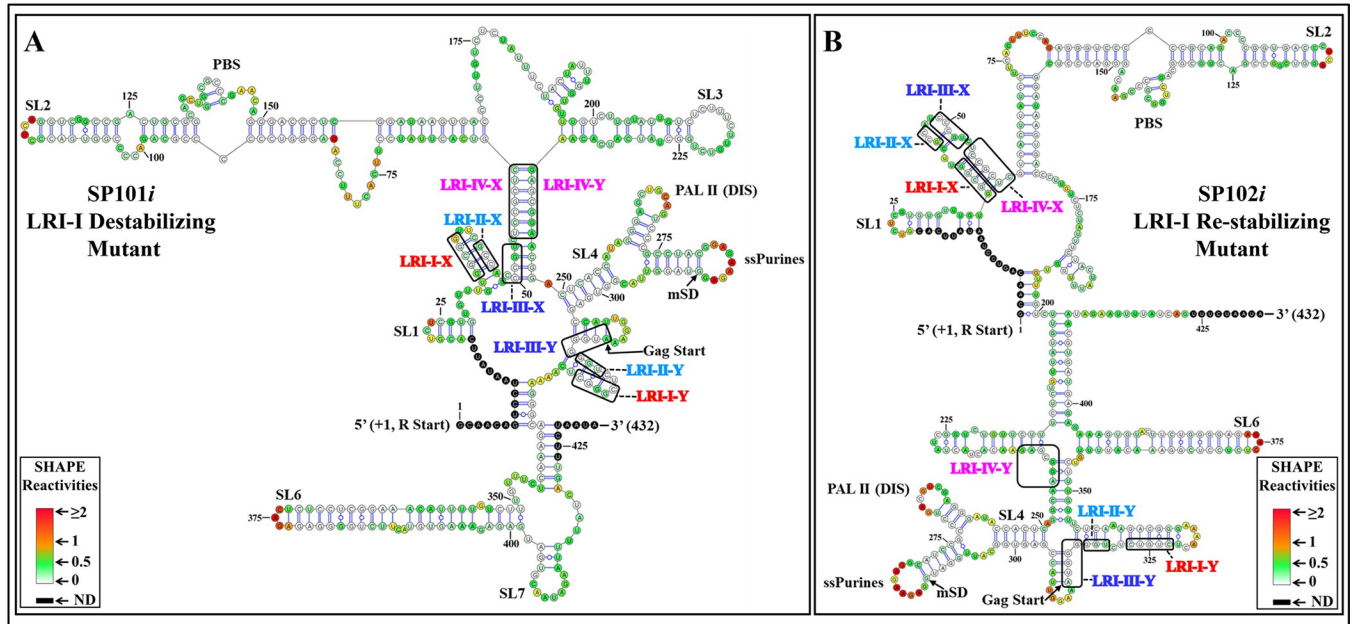
Fig 4. hSHAPE-validated RNA secondary structure model of the LRI-I mutants. The first 432 nucleotides of the 712 nt long RNA are shown. **(A)** Mutant SP101*i* was designed to destabilize LRI-I. **(B)** Mutant SP102*i* was designed to re-stabilize LRI-I. Structural elements that are also present in the WT structure (SA35) such as stem-loops (SL1-7), PBS, DIS, single-stranded purines (ssPurines), and mSD are marked as in Fig 1. The X and Y-strands of different LRIs in both destabilizing and re-stabilizing mutants are boxed and labeled in different colors for clarity. Nucleotides are color-coded according to SHAPE reactivity derived from a minimum of 3 independent experiments, with data provided in S1 Table. DIS, dimerization initiation site; LRI, long-range interaction; mSD, major splice donor; PBS, primer binding site; WT, wild type.

https://doi.org/10.1371/journal.pbio.3002827.g004

was not restored in SP102*i*; indeed, this mutant also lost the other 3 LRIs, namely LRI-II, III, and IV (Fig 4B), suggesting that the packaging defect of SP102 (Fig 2) might be due to the disruption of LRI-IV.

Next, we tested the structure of LRI-III mutant RNAs SP105*i* and SP106*i*, as these 2 mutants exhibited severe defects in RNA packaging and propagation (Fig 2). According to their hSHAPE structures (Fig 5A and 5B), all 4 LRIs were lost in these mutants, consistent with their loss of function (Fig 2).

The SP107 mutant, designed to disrupt LRI-IV, lost its ability to package and propagate RNA, while mutant SP108, designed to restore LRI-IV, restored RNA packaging and propagation to WT levels (Fig 2). Accordingly, the hSHAPE-validated structure of SP107*i* revealed not only the loss of LRI-IV but also the loss of the other 3 LRIs (Fig 6A), whereas in the case of SP108*i*, LRI-IV was restored, but not the other LRIs, while adopting a structure globally similar, though not identical, to the wild type (Fig 6B). Altogether, these results indicated that the structure of LRI-IV, but not its sequence is essential for function.

An intriguing observation from the hSHAPE structures is the consistent retention of 2 critical structural patterns, SL2 and the branched SL4, in their original locations across all mutants (Figs 4–6). These motifs encompass nucleotides recognized as primary Gag-binding sites crucial for MMTV RNA packaging [53]. Despite this, certain mutants exhibited defects in RNA packaging (Fig 2), suggesting that Pr77^Gag was unable to bind to these sites in those specific mutants.

A new secondary structure model supported by structural and functional analysis.
Overall, the combined biological and structural data suggest that the previously proposed LRI-I and LRI-II [36,53] may not necessarily exist. LRI-III mutants SP105 and SP106 exhibited
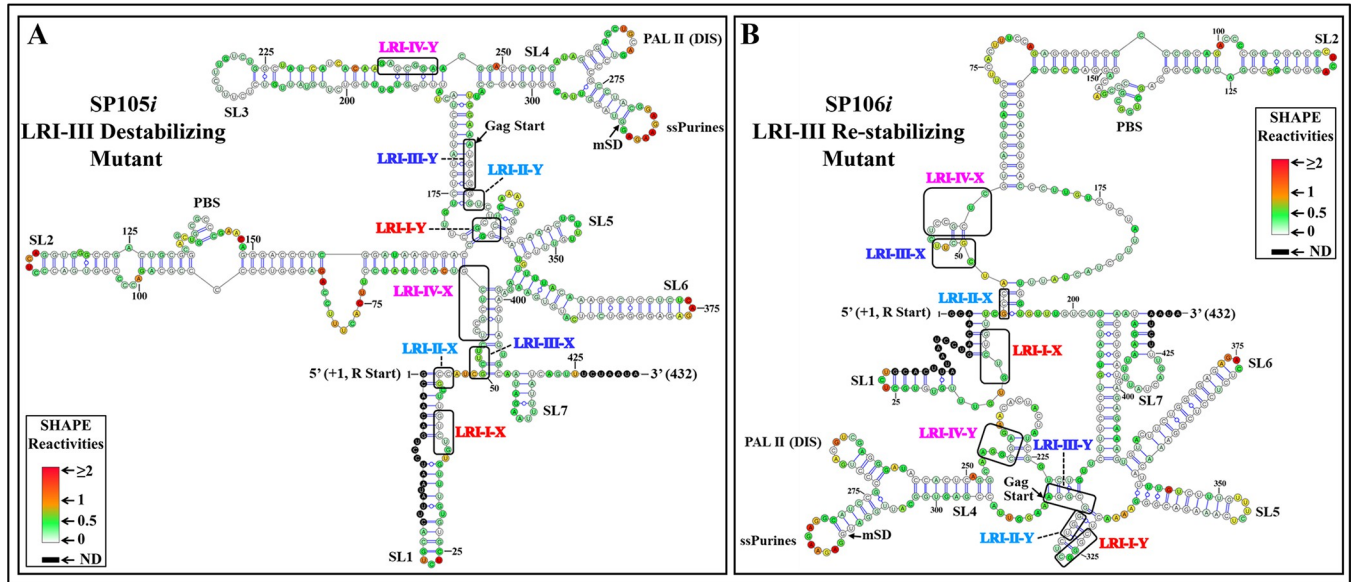
Fig 5. hSHAPE-validated RNA secondary structure model of the LRI-III mutants. The first 432 nucleotides of the 712 nt long RNA are shown. **(A)** Mutant SP105*i* was designed to destabilize LRI-III. **(B)** Mutant SP106*i* was designed to re-stabilize LRI-III. Structural elements that are also present in the WT structure (SA35) such as stem-loops (SL1-7), PBS, DIS, single-stranded purines (ssPurines), and mSD are marked as in Fig 1. The X and Y-strands of different LRIs in both destabilizing and re-stabilizing mutants are boxed and labeled in different colors for clarity. Nucleotides are color-coded according to SHAPE reactivity derived from a minimum of 3 independent experiments, with data provided in S1 Table. DIS, dimerization initiation site; LRI, long-range interaction; mSD, major splice donor; PBS, primer binding site; WT, wild type.

https://doi.org/10.1371/journal.pbio.3002827.g005
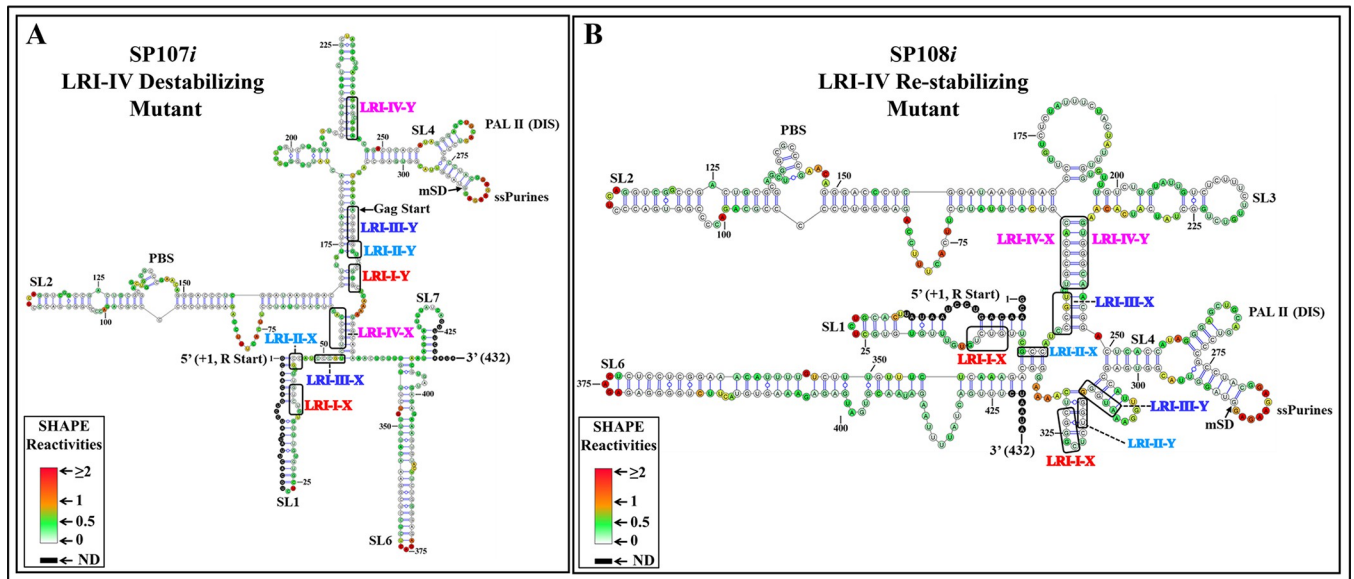
Fig 6. hSHAPE-validated RNA secondary structure model of the LRI-IV mutants. The first 432 nucleotides of the 712 nt long RNA are shown. **(A)** Mutant SP107*i* was designed to destabilize LRI-IV. **(B)** Mutant SP108*i* was designed to re-stabilize LRI-IV. Structural elements that are also present in the WT structure (SA35) such as stem-loops (SL1-7), PBS, DIS, single-stranded purines (ssPurines), and mSD are marked as in Fig 1. The X and Y-strands of different LRIs in both destabilizing and re-stabilizing mutants are boxed and labeled in different colors for clarity. Nucleotides are color-coded according to SHAPE reactivity derived from a minimum of 3 independent experiments, with data provided in S1 Table. DIS, dimerization initiation site; LRI, long-range interaction; mSD, major splice donor; PBS, primer binding site; WT, wild type.

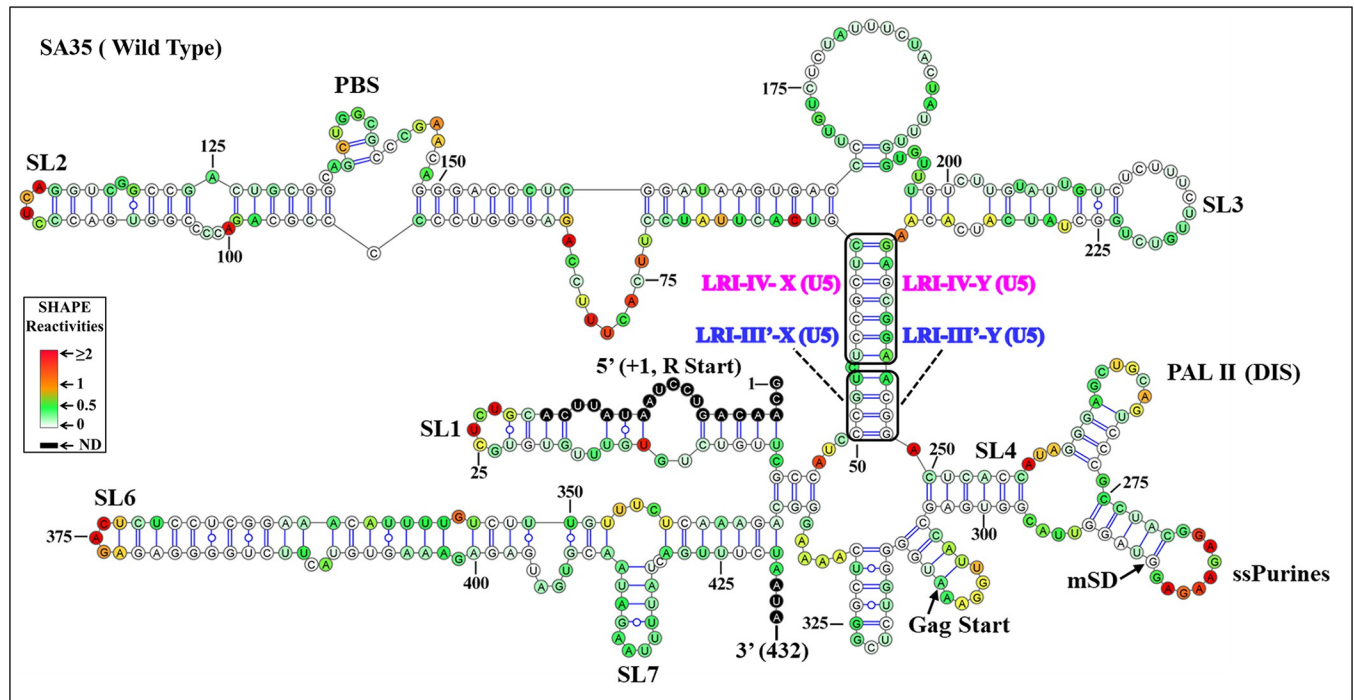https://doi.org/10.1371/journal.pbio.3002827.g006

Fig 7. The new hSHAPE-validated RNA secondary structure model of the first 432 nucleotides, obtained through biochemical re-probing of the WT (SA35) 5′ end of MMTV. The updated structure model closely resembles the previously proposed structure but exhibits notable differences in the LRIs and minor differences in SLs 5, 6, and 7. In the new structure model, LRIs I and II are absent, while LRI-IV remains unchanged. Additionally, the X-strand of the initially proposed LRI-III now base pairs with complementary sequences 195 nucleotides downstream within U5, instead of Gag. Structural elements consistent with the earlier proposed model, such as SLs1-7, PBS, DIS, single-stranded purines (ssPurines), and mSD, are present in their native positions and labeled as in Fig 1. The X and Y-strands of LRIs III′ and IV are shown in boxes and labeled with different colors for clarity. Nucleotides are color-coded according to SHAPE reactivity derived from a minimum of 3 independent experiments, with data provided in S1 Table. DIS, dimerization initiation site; LRI, long-range interaction; MMTV, mouse mammary tumor virus; mSD, major splice donor; PBS, primer binding site; SL, stem loop; WT, wild type.

https://doi.org/10.1371/journal.pbio.3002827.g007

pronounced defects in RNA packaging, which suggests that while these sequences are important for function, they may not be involved in complementary base-pairing, as initially thought [36,53]. In contrast, biological and structural data obtained with mutants SP107 and SP108 strongly support the existence of LRI-IV, as initially proposed [53].

To confirm these results, we re-probed the same region of WT (SA35) MMTV gRNA (712 nucleotides from R to 400 nucleotides of Gag). The new hSHAPE-validated structure (Fig 7) closely resembles the earlier proposed structure, but differs notably in the LRIs, with minor differences in SLs 5–7. Specifically, LRI-I is not observed, and LRI-II's U5 sequences (5′ GCC 3′; nts 44–46) are base paired with a different Gag sequence (3′ CGG 5′; nts 337–335) (Fig 7). However, the existence and biological significance of this alternative LRI remains uncertain because when we mutated the U5 sequences (which forms a LRI in both the old and the new structure models), it did not affect RNA packaging. In the new model, LRI-III is present in a modified form that we named LRI-III′: its 5′ sequence (5′ CCGU 3′; nts 50–53) is base paired with sequences within U5 (5′ ACGG 3′; 244–248) instead of Gag. Consistently, no changes are observed in LRI-IV in the new hSHAPE-validated structure, which corroborates the structure–function data of LRI-IV mutants presented above (Fig 7).

The same software and input sequences (432 nucleotides) were used to predict the secondary structure of the MMTV 5′ region in this work and our previous publication [53]. Both studies also utilized the same RNA and experimental probing conditions. However, new primers for cDNA synthesis were introduced in this study (refer to the Material and methods

section). The 3′ primer was redesigned because the one used in our previous study was unsuitable for analyzing the mutants used here. More importantly, a new 5′ primer was designed to improve the signal-to-noise ratio at the 5′ end of the RNA. The enhanced quality of the SHAPE data at the 5′ end of RNA (SA35) resulted in our new modeling, providing a slightly different secondary structure. Notably, our new modeling identified the previously published secondary structure model as the third most stable structure. For large RNAs, it is common for small differences in experimental data to result in different "most stable" structures, as multiple structures often have very similar minimal free energies [78].

To examine the role of the new LRI-III' in MMTV gRNA packaging, we utilized mutant SP105, which disrupts it, and created mutant SP109, containing the same mutations as SP105 and additional ones designed to restore LRI-III' (Fig 8A). As observed earlier, SP105 mutant exhibited severely compromised RNA packaging and propagation; however, when the new LRI-III' structure was restored, RNA packaging was restored to almost WT levels (Fig 8A). Accordingly, hSHAPE indicated that the secondary structure of mutant SP109*i* is similar to wild type, as expected (compare Fig 7 with Fig 8B). These results suggest that the secondary structure of the MMTV *Psi* RNA is held together by a long stretch of 11 Watson–Crick–Franklin base pairs (interrupted by a 1 nucleotide bulge), the sequence of which is not important for function.

Binding of Pr77^{Gag} to the mutant RNAs does not correlate with RNA packaging

Next, we performed filter-binding assays to determine whether the gRNA packaging results obtained with the LRI mutants could be correlated to their ability to bind to Pr77^{Gag}. The Pr77^{Gag} protein we used in these assays is able to form virus-like particles (VLPs) in vitro as well as in vivo [53,79]. Furthermore, the Gag VLPs formed by Pr77^{Gag}-His$_6$ fusion protein in human embryonic kidney (HEK293T) cells efficiently package RNA containing the MMTV *Psi* [53,79]. Based on the ratio of UV absorbance at 260 and 280 nm, the purified protein was observed to be devoid of nucleic acids. Finally, dynamic light scattering (DLS) revealed that the average hydrodynamic radius (Rh) determined based on volume (percent) and number (percent) distribution was around 6.00 nm, corresponding to Pr77^{Gag} trimers ([53]; S3 Fig). Thus, this purified protein was used in filter-binding assays along with radiolabeled in vitro transcribed RNA from WT (SA35) and LRI mutants (SP101*i*, 102*i*, SP105*i*-SP109*i* RNAs; Fig 3B). As is evident from Fig 9A, almost all mutant RNAs (whether packaging or non-packaging) were able to bind to Pr77^{Gag} efficiently; however, some of the non-packaging mutants exhibited a lower binding plateau, which could reflect a different stoichiometry. Interestingly, as shown in Fig 9B, all mutants except SP106*i* bind Pr77^{Gag} with a Kd and a cooperativity (h) similar to the WT. These results suggest that Pr77^{Gag} efficiently binds to the MMTV *Psi* region regardless of its ability to promote RNA packaging.

Binding of Pr77^{Gag} to specific sites is critical for efficient MMTV gRNA packaging

Since the filter-binding assays revealed that Pr77^{Gag} binding affinity does not correlate with RNA packaging efficiency, we asked whether Pr77^{Gag} binds at the same sites within *Psi* of the packaging and non-packaging mutants. To this goal, we identified specific nucleotides that bind to Pr77^{Gag} by conducting RNA modification (via BzCN), both in the presence and absence of Pr77^{Gag}. Reduced hSHAPE reactivity in the presence of Pr77^{Gag} revealed the footprints of Pr77^{Gag} on the WT and mutant *Psi*. To prevent nonspecific binding, hSHAPE was conducted with excess spliced *env* RNA (AK29; Fig 3C) at a high (4-fold) molar concentration.
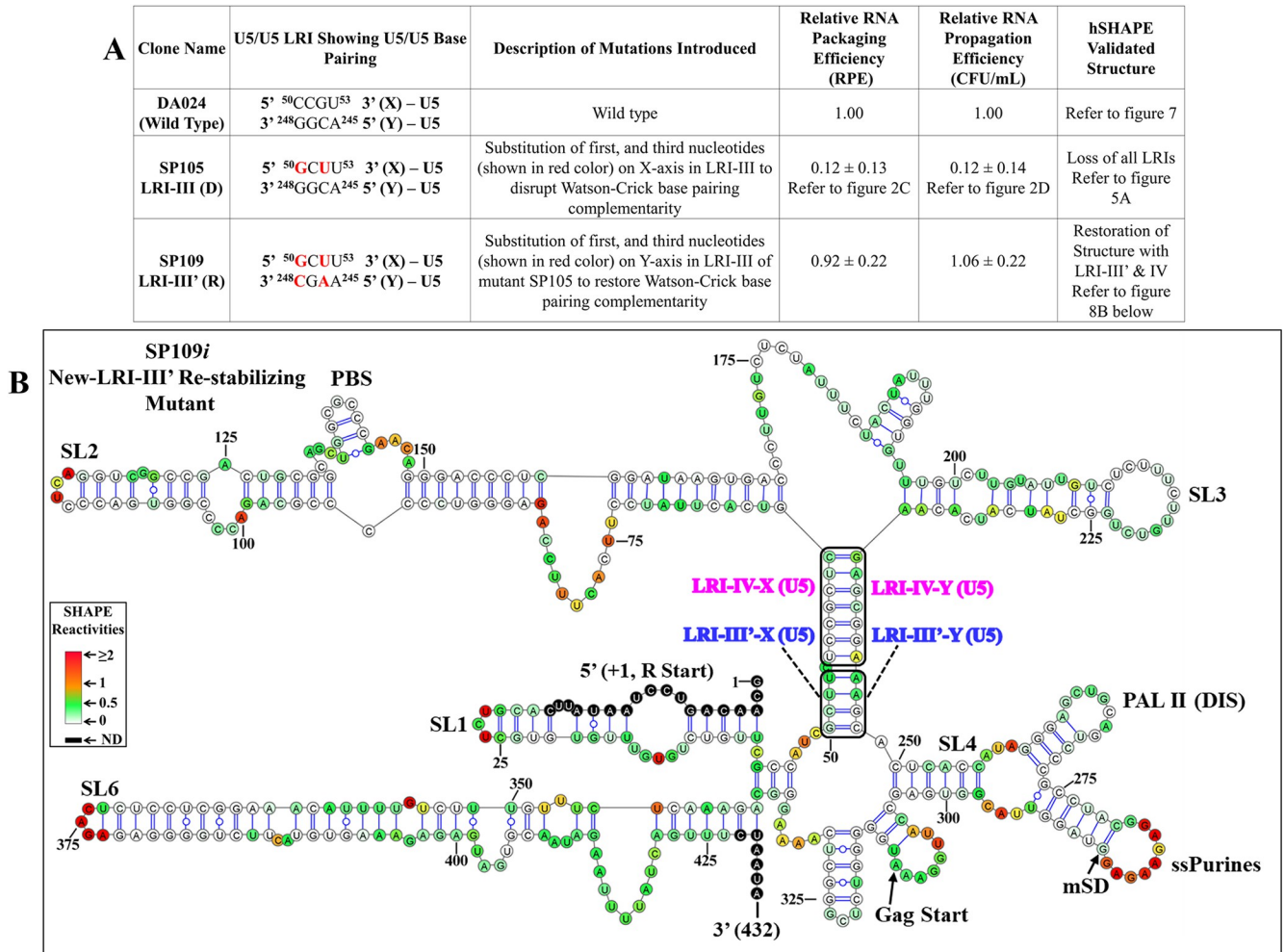
A

Clone Name	U5/U5 LRI Showing U5/U5 Base Pairing	Description of Mutations Introduced	Relative RNA Packaging Efficiency (RPE)	Relative RNA Propagation Efficiency (CFU/mL)	hSHAPE Validated Structure
DA024 (Wild Type)	5' ^{50}CCGU^{53} 3' (X) – U5 3' ^{248}GGCA^{245} 5' (Y) – U5	Wild type	1.00	1.00	Refer to figure 7
SP105 LRI-III (D)	5' ^{50}GCUU^{53} 3' (X) – U5 3' ^{248}GGCA^{245} 5' (Y) – U5	Substitution of first, and third nucleotides (shown in red color) on X-axis in LRI-III to disrupt Watson-Crick base pairing complementarity	0.12 ± 0.13 Refer to figure 2C	0.12 ± 0.14 Refer to figure 2D	Loss of all LRIs Refer to figure 5A
SP109 LRI-III' (R)	5' ^{50}GCUU^{53} 3' (X) – U5 3' ^{248}CGAA^{245} 5' (Y) – U5	Substitution of first, and third nucleotides (shown in red color) on Y-axis in LRI-III of mutant SP105 to restore Watson-Crick base pairing complementarity	0.92 ± 0.22	1.06 ± 0.22	Restoration of Structure with LRI-III' & IV Refer to figure 8B below

B

Fig 8. LRI-III' with complementary heterologous sequences restores RNA packaging and structure. (A) Description of the substitution mutants in the newly proposed LRI-III with red nucleotides indicating introduced mutations aimed at destabilizing or re-stabilizing complementarity with heterologous sequences. Columns 4, 5, and 6 show results of the effect of mutations on RNA packaging, propagation, and structure, respectively. The RNA packaging data shown here is from a minimum of 3 independent experiments performed in triplicates (±SD). The RNA propagation data shown here is from a minimum of 3 independent experiments performed in duplicates (±SD). **(B)** A hSHAPE-validated RNA secondary structure model depicts the restored LRI-III' mutant SP109i, designed to re-stabilize LRI-III' with heterologous complementary sequences. Structural elements that are also present in the new WT structure (SA35) such as stem-loops (SL1-7), PBS, DIS, single-stranded purines (ssPurines), and mSD are marked as in Fig 7. The X and Y-axes of LRIs III' and IV are boxed and labeled in different colors for clarity. Nucleotides are color-coded according to SHAPE reactivity derived from a minimum of 3 independent experiments, with data provided in S1 Table. The data underlying this Fig 8A can be found in S1 Data. DIS, dimerization initiation site; LRI, long-range interaction; mSD, major splice donor; PBS, primer binding site; SL, stem loop; WT, wild type.

https://doi.org/10.1371/journal.pbio.3002827.g008

Pr77^{Gag} was used at a concentration 10-fold higher than the Kd value, ensuring complete saturation of the high-affinity binding sites. These conditions identified high-affinity binding sites in other retroviral *Psi* footprinting experiments [46,53,80].

When hSHAPE was conducted on WT (SA35) RNA in the absence of Pr77^{Gag}, the reactivity pattern was consistent with the new structure model presented in Fig 7. To identify nucleotides interacting with Gag, hSHAPE was conducted on wild type RNA (SA35; Fig 10A) in the presence of Pr77^{Gag} and spliced *env* RNA (AK29; Fig 10B) as a competitor. The differences in hSHAPE reactivity obtained in the absence and in the presence of Pr77^{Gag} were quantified and mapped onto this secondary structure model. Differences were deemed significant when hSHAPE reactivities showed a variance of ≥0.20 and a relative difference exceeding 40% to
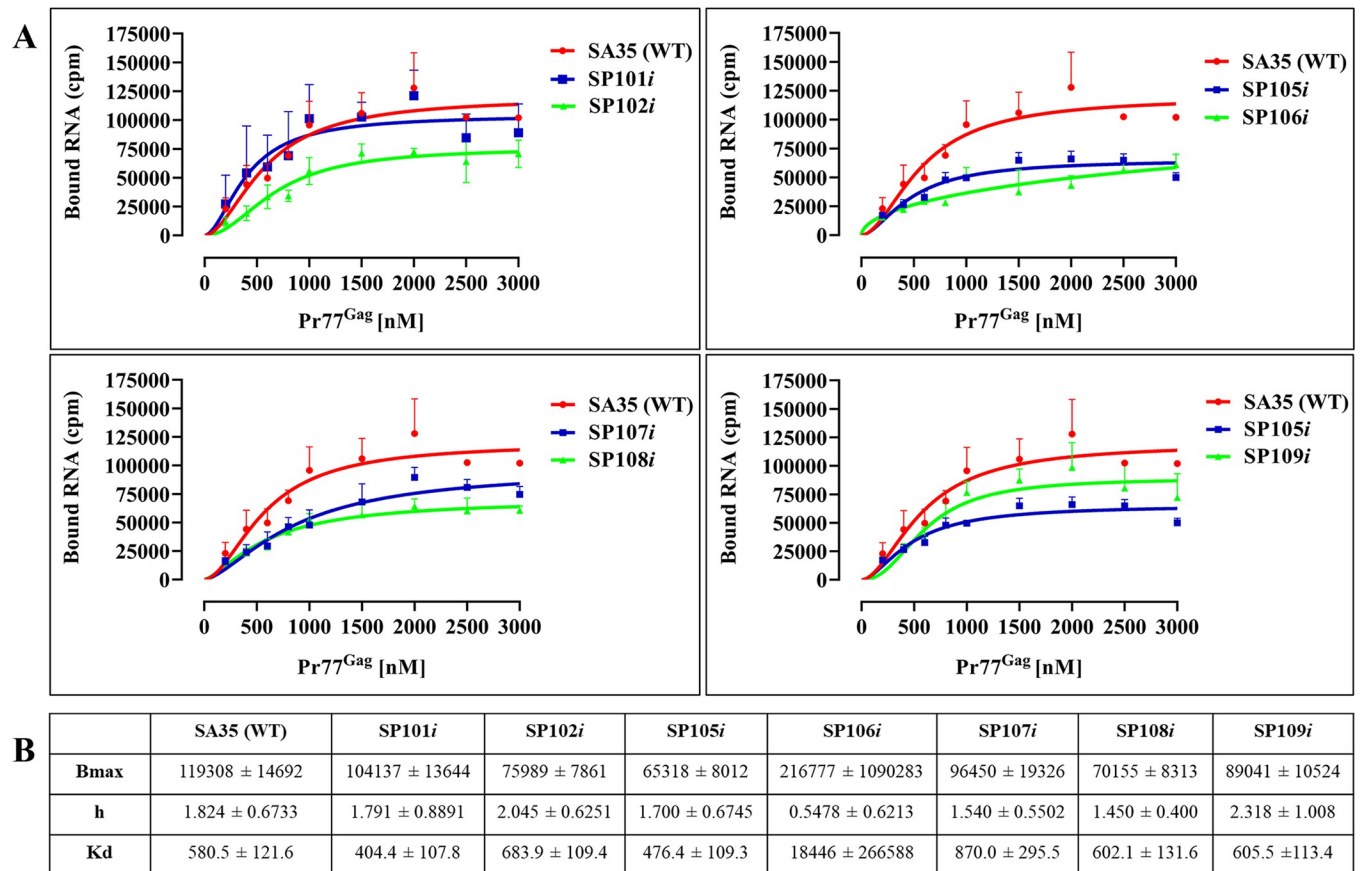
Fig 9. MMTV Pr77^{Gag} binding to the WT and LRI mutant RNAs using filter-binding assays. (A) The membrane-bound radioactivity of the wild type WT (SA35) unspliced and mutant RNAs was quantified at increasing concentrations of MMTV Pr77^{Gag}. Data points were fitted with the Hill's equation, with error bars denoting standard deviation from the mean of 3 independent experiments. **(B)** Pr77^{Gag} binding parameters to the WT and LRI mutant MMTV *Psi* region as derived using Hill's equation. Cumulative data is derived from 3 independent experiments. Bmax: represents the maximum-specific binding; h: represents the Hill slope; Kd: represents the Pr77^{Gag} concentration needed to achieve a half-maximum binding followed by their standard deviation. The data underlying this Fig 9A and 9B can be found in S1 Data. LRI, long-range interaction; MMTV, mouse mammary tumor virus; WT, wild type.

50% [81]. hSHAPE reactivities in the presence of Pr77^{Gag} consistently showed reduced reactivity in nucleotides within single-stranded purines (ssPurines) and the primer binding site (PBS), confirming Pr77^{Gag} binding to these sites, as reported earlier (Figs 10C and 11A; [53]).

Footprinting experiments conducted on SP101*i* revealed protections of most nucleotides within ssPurines and the PBS that have been shown to be the primary Gag-binding sites in the WT RNA (Figs 11B and S4). This is consistent with the mutant's ability to package RNA (Fig 2). In contrast, minimal or no protections were observed in ssPurines and PBS regions in the RNA SP102*i* (only 1 nucleotide (135C) in the PBS was protected). Surprisingly, in this mutant, other nucleotides (26G, 30G, 32U, 35U, 44G, 52G) were protected by Pr77^{Gag} (Figs 11C and S5). The lack of protection in ssPurines and the PBS, which are primary Gag-binding sites [53], aligns with the non-packaging nature of SP102 (Fig 2). However, given the packaging defect of this mutant, the binding of Pr77^{Gag} to other nucleotides was rather puzzling.

In the case of LRI-III, neither SP105*i* (destabilizing mutant) nor SP106*i* (restoring mutant) showed any Pr77^{Gag} footprint within ssPurines (S6 and S7 Figs and Fig 12B and 12C). In the PBS, only 1 (137G) or 2 nucleotides (137G and 147C) were protected by Pr77^{Gag} in SP105*i* and SP106*i*, respectively (S6 and S7 Figs and Fig 12B and 12C). The absence of footprints in the
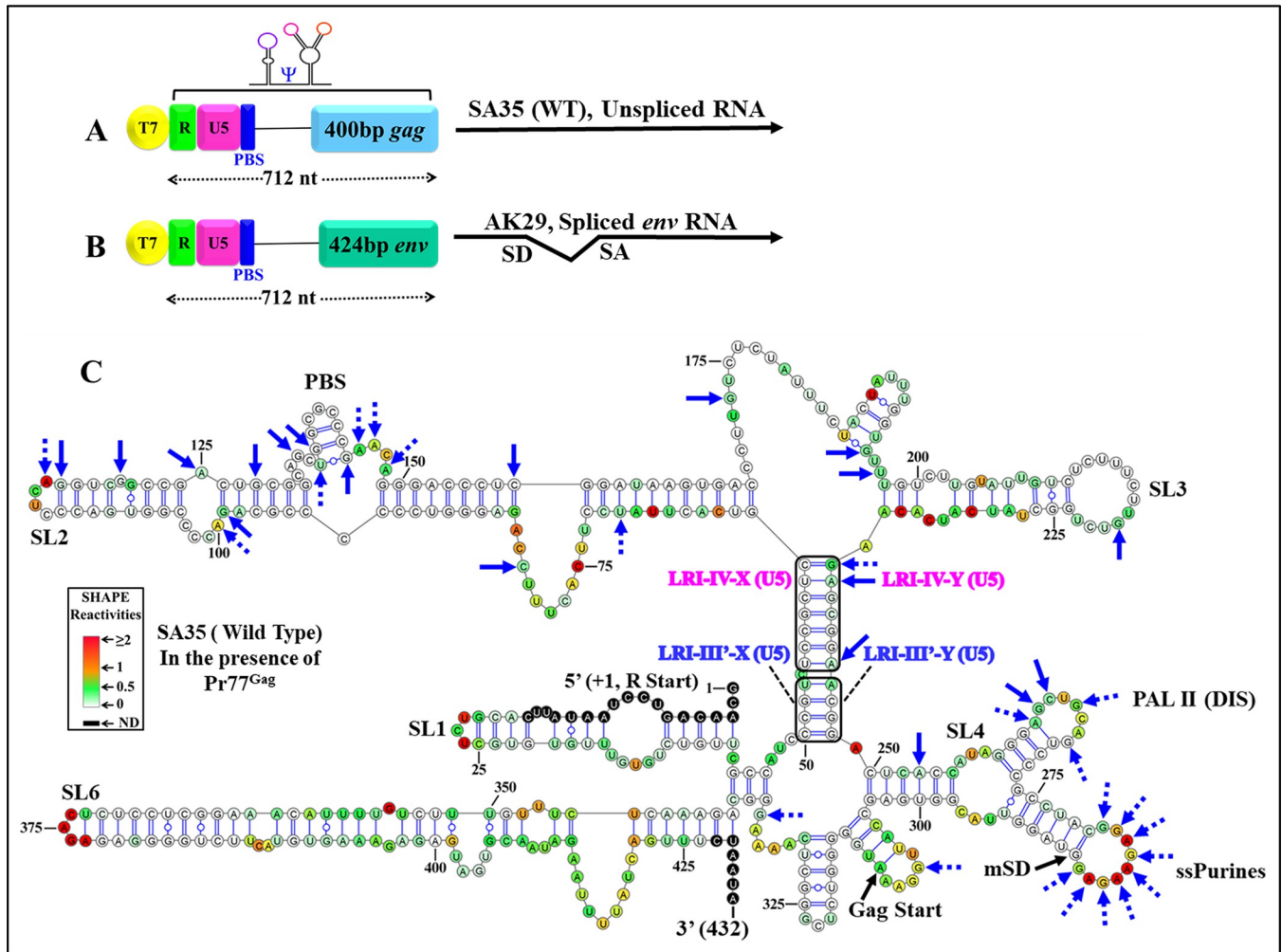
Fig 10. Footprints of Pr77^Gag on the WT structure model. (A) Schematic illustration of the MMTV WT packageable vector (SA35) RNA from R to 712 nucleotides expressed from a T7 expression plasmid. **(B)** Illustration of the *envelope* (*env*) spliced RNA (AK29) from R to 712 nucleotides expressed from a T7 expression plasmid. **(C)** hSHAPE analysis was carried out both with and without Pr77^Gag. The mean triplicate SHAPE reactivity obtained without Pr77^Gag was used to predict the RNA secondary structure model of WT (SA35) RNA. Subsequently, the mean hSHAPE reactivities obtained with Pr77^Gag were overlaid onto the RNA secondary structure model predicted in the absence of Pr77^Gag. Notably, nucleotides within the previously identified primary Gag-binding sites, such as single-stranded purines (ssPurines) and the PBS, exhibited significant reductions in hSHAPE reactivities. Nucleotides in ssPurines, PBS, and all other nucleotides marked by arrows show significant reduction in SHAPE reactivities. The hSHAPE reactivity key was developed based on the mean of hSHAPE reactivities for each nucleotide, as shown in S3 Table. The data shown is from a minimum of 3 independent experiments conducted both in the absence and presence of Pr77^Gag. All nucleotides that show a reactivity decrease >40% upon Gag addition, also show statistically significant difference according to the Mann–Whitney non parametrical U test ($p < 0.05$). MMTV, mouse mammary tumor virus; PBS, primer binding site; WT, wild type.

https://doi.org/10.1371/journal.pbio.3002827.g010

ssPurines and PBS of mutants SP105*i* and SP106*i* corroborates with their inability to package gRNA and the loss of LRIs I-IV in their secondary structure (Figs 2 and 5).

By contrast, mutant SP109, designed to restore LRI-III' based on the new RNA secondary structure model and which successfully restored both RNA packaging and structure (Fig 8), showed Pr77^Gag footprints primarily within ssPurines and the PBS, similar to the wild type (Figs 12D and S8). These results further confirm the new hSHAPE-validated RNA secondary structure model for MMTV *Psi* and stress the importance of Pr77^Gag binding to the ssPurines and the PBS for efficient RNA packaging.

This result was confirmed by footprinting experiments on mutants SP107*i* and SP108*i*. In brief, the destabilizing mutant (SP107*i*), in which LRI-IV is disrupted, showed no footprint
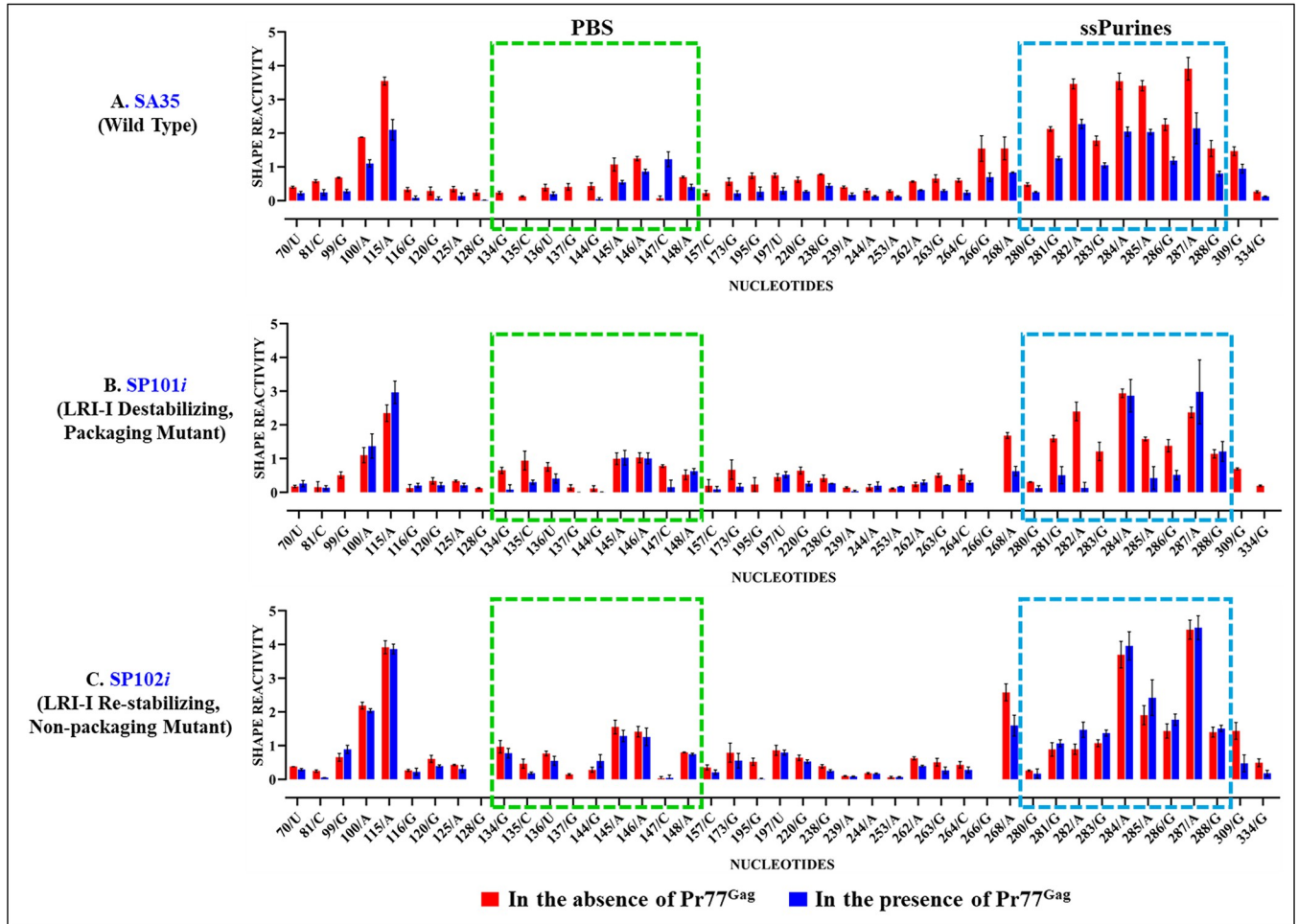
Fig 11. Footprints of Pr77^Gag on the WT and LRI-I mutant packaging signal RNAs. Histograms showing the SHAPE reactivities of nucleotides in the absence (red bars) and presence (blue bars) of Gag for the packaging signal RNAs of: **(A)** wild type (SA35), **(B)** mutant SP101*i* designed to destabilize LRI-I, and **(C)** mutant SP102*i* designed to re-stabilize LRI-I. The dashed boxes mark the already identified primary Gag-binding sites, such as single stranded purines (ssPurines) and the PBS. For full details of nucleotides attenuation of SHAPE reactivities in the absence and presence of Pr77^Gag from a minimum of 3 independent experiments, see S3 Table and S4 and S5 Figs. All nucleotides that show a reactivity decrease >40% upon Gag addition, also show statistically significant difference according to the Mann–Whitney non parametrical U test ($p < 0.05$). The data underlying this figure can be found in S1 Data. LRI, long-range interaction; PBS, primer binding site; WT, wild type.

https://doi.org/10.1371/journal.pbio.3002827.g011

within ssPurines, and only 1 nucleotide was protected by Pr77^Gag in the PBS (134G; Figs 13B and S9). Conversely, mutant SP108*i*, which restores LRI-IV, exhibited protections for a majority of the nucleotides in the ssPurines (6/9) and in the PBS (4/7) (Figs 13C and S10). These results indicate that when LRI-IV is restored, Pr77^Gag can bind to the ssPurines and the PBS and promote efficient RNA packaging (Fig 2).

Overall, footprinting experiments conducted on all mutants revealed that efficient RNA packaging correlates with strong Pr77^Gag binding to nucleotides in ssPurines and the PBS, known as primary Gag-binding sites during WT MMTV gRNA packaging [53]. In stark contrast, mutants that failed to package RNA efficiently revealed Pr77^Gag binding to nucleotides scattered in regions other than ssPurines and the PBS. Hence, while Gag binding during retroviral RNA packaging is essential, successful RNA packaging relies on binding to specific nucleotides that can only take place within the appropriate structural context.
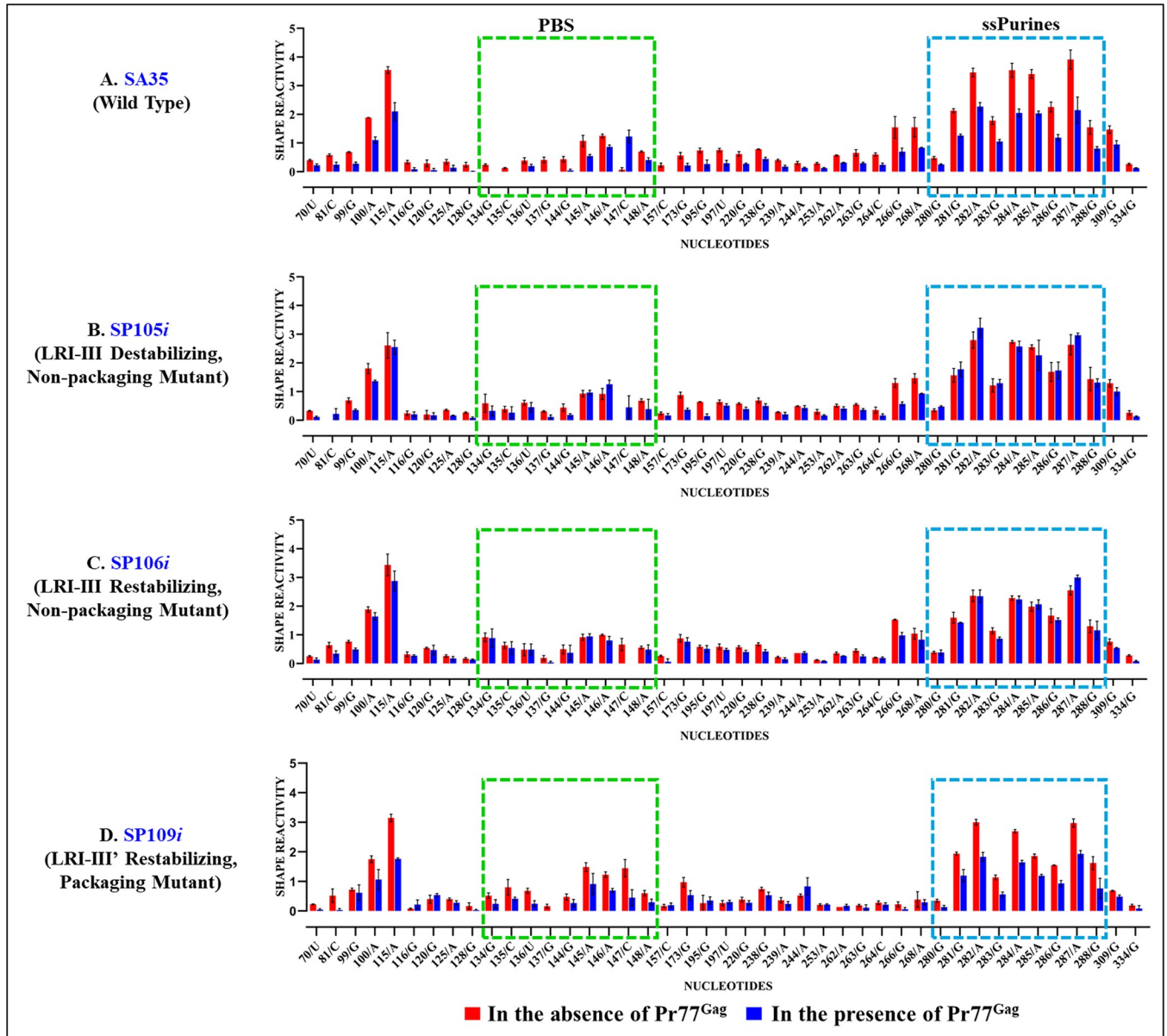
Fig 12. Footprints of Pr77^Gag on the WT and LRI-III mutant packaging signal RNAs. Histograms showing the SHAPE reactivities of nucleotides in the absence (red bars) and presence (blue bars) of Gag for the packaging signal RNAs of: **(A)** wild type (SA35), **(B)** mutant SP105*i* designed to destabilize LRI-III, **(C)** mutant SP106*i* designed to re-stabilize LRI-III, and **(D)** mutant SP109*i* designed to restore the LRI-III'. The dashed boxes depict the already identified primary Gag-binding sites, such as single stranded purines (ssPurines) and the PBS. For full details of nucleotides attenuation of SHAPE reactivities in the absence and presence of Pr77^Gag from a minimum of 3 independent experiments, see S3 Table and S6, S7 and S8 Figs. All nucleotides that show a reactivity decrease >40% upon Gag addition, also show statistically significant difference according to the Mann–Whitney non parametrical U test ($p < 0.05$). The data underlying this figure can be found in S1 Data. LRI, long-range interaction; PBS, primer binding site; WT, wild type.

Discussion

RNA viruses rely on conserved structural information found in noncoding and occasionally coding sequences to execute crucial events during their life cycle. The 5′ end of the genome in numerous RNA viruses, including crucial pathogens for humans, animals, and plants, harbors a wealth of *cis*-acting information conveyed through diverse higher-order structures, often maintained by LRIs [82,83].
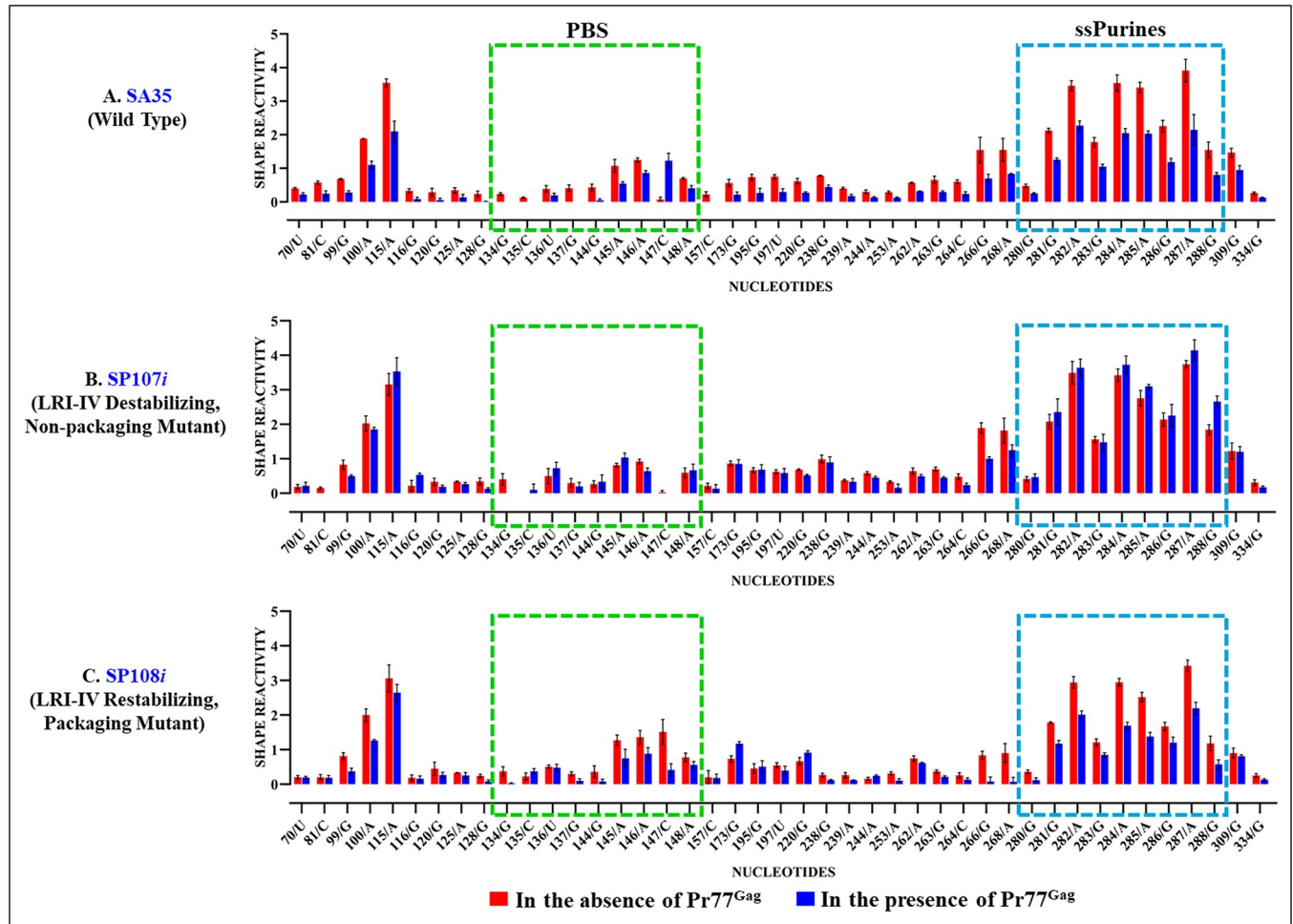
Fig 13. Footprints of Pr77^Gag on WT and LRI-IV mutant packaging signal RNAs. Histograms showing SHAPE reactivities of nucleotides in the absence (red bars) and presence (blue bars) of Gag for the packaging signal RNAs of: **(A)** wild type (SA35), **(B)** mutant SP107*i* designed to destabilize LRI-IV, and **(C)** mutant SP108*i* designed to restore LRI-IV. The dashed boxes depict the already identified primary Gag-binding sites such as single stranded purines (ssPurines) and the PBS. For full details of nucleotides attenuation of SHAPE reactivities in the absence and presence of Pr77^Gag from a minimum of 3 independent experiments, see S3 Table and S9 and S10 Figs. All nucleotides that show a reactivity decrease >40% upon Gag addition, also show statistically significant difference according to the Mann–Whitney non parametrical U test ($p < 0.05$). The data underlying this figure can be found in S1 Data. LRI, long-range interaction; PBS, primer binding site; WT, wild type.

https://doi.org/10.1371/journal.pbio.3002827.g013

Notably, LRIs have been identified as conserved structural motifs in various retroviruses [9,10,36,43,45,47–54,84]. Despite significant sequence and structural heterogeneities, the persistence of LRIs in isolates of HIV-1, HIV-2, MPMV, MMTV, and FIV provides evidence for their functional importance in the retroviral life cycle [9,10,36,43,47,48,52–54]. Indeed, mutations destabilizing the complementarity of these LRIs adversely affect crucial steps in the retroviral life cycle, including RNA packaging and dimerization [10,47,48,52,54].

The propensity of retroviral *Psi* to fold into complex secondary structures underscores their crucial role during retroviral gRNA packaging. In most retroviruses, including lentiviruses, the efficiency and selectivity of gRNA packaging in part is governed by the multimerization of the Gag precursor at the plasma membrane [32,85–88], complicating the analysis of this process. MMTV constitutes an attractive system to study membrane-independent retroviral RNA packaging, as it assembles in the cytoplasm of infected cells [59,60]. Within the 5′ end of the MMTV genome, various sequence and structural motifs have been identified as pivotal for

gRNA packaging and dimerization ([36,41,53,61]; Fig 1). A distinctive feature of the previously published RNA secondary structure of the MMTV gRNA *Psi* is the presence of several LRIs predicted to anchor the overall secondary RNA structure ([36,41,53]; Fig 1).

The initial goal of this study was to test the role of the 4 proposed LRIs (LRI-I-IV) in MMTV gRNA packaging by combining structural and functional approaches. Among the LRIs that involve only complementary U5 sequences, we were unable to confirm the presence of LRI-II. However, our findings indicate that the proposed LRI-I and LRI-III do not manifest as originally proposed (see Fig 2). On the other hand, our functional data support the existence of LRI-IV, formed by complementary base-pairing of U5/U5 sequences (Fig 2). Our structural data (Figs 4–7) supported these conclusions and demonstrated the existence of an alternative LRI involving U5/U5 complementary sequences (approximately 198 nucleotides apart) instead of U5/Gag as initially proposed ([36,53], compare Fig 1 with Fig 7) that we named as LRI-III'. Interestingly, LRI-III' and LRI-IV are both important for gRNA packaging (Figs 2 and 8), and our results suggest these LRIs function mechanistically in a similar manner, as their sequences can be substituted with heterologous sequences without having any adverse effects on RNA packaging as long as base pairing is maintained (Figs 2, 6B and 8). Interestingly, the 7 nucleotide LRI-IV in MMTV exhibit functional similarities to the R/U5-Gag heptanucleotide LRI observed in FIV, a lentivirus, since base-pairing, but not sequence, of the FIV LRI is crucial for RNA packaging [10,54].

Of note, the MMTV LRI-III' and LRI-IV are contiguous in the new secondary structure model of the MMTV *Psi*, being only separated by a one nucleotide bulge (Fig 7). Accordingly, when 2 nucleotides (out of 4) in LRI-III' or 3 nucleotides (out of 7) in LRI-IV were substituted independently to destabilize the respective LRIs, both LRIs were lost, as well as function (Figs 2, 5A and 6A). Furthermore, in a compensatory approach, when we designed mutants aimed at restoring LRI-III' and LRI-IV individually (mutants SP109 and SP108, respectively), both function and structure were restored (Figs 2, 6B and 8). The data presented here thus suggest that any perturbation designed to destroy complementarity in this 11 nucleotide long extended LRI severely compromise MMTV gRNA packaging. In light of these observations, we reviewed the structure function analysis of all LRI mutants and observed that the LRI-I destabilizing mutant (SP 101), which did not affect RNA packaging maintained both LRI-III' and IV as a long continuous stretch as discussed above (Figs 2 and 4A). Taken together, these observations strongly argue that LRIs III' and IV must be regarded as 1 extended LRI rather than 2 separate LRIs. In this respect, MMTV resembles HIV and FIV, which possess a single identified long-range RNA-RNA interaction [9,10,43,47,48,52,54]. However, unlike other retroviruses such as HIV-1, HIV-2, MPMV, and FIV, where U5/Gag sequences are involved in forming LRIs, [9,10,43,47,48,52,54], the extended MMTV LRI III'-IV only involves U5/U5 sequences (198 nucleotides apart). MPMV, another *betaretrovirus* harbors 2 LRIs that are important for RNA packaging, and in contrast with MMTV, both the structure as well as the sequence of one of these LRIs are important for MPMV gRNA packaging [35,52].

Results from filter-binding experiments demonstrated that both WT and LRI mutants, irrespective of their RNA packaging phenotype, could efficiently bind to Pr77[Gag] (Fig 9). This observation is in strong contrast with a number of previous studies on several retroviruses, including HIV [44,46], FIV [54], and MMTV [53], in which a clear correlation was observed between binding of the Gag precursor to *Psi* and gRNA packaging efficiency. Our present results thus indicate that while Gag binding to the packaging signal is necessary, it is not sufficient to ensure efficient RNA packaging.

Several previous studies have proposed that high affinity binding of HIV-1 Gag to *Psi*-containing RNAs cannot explain selective packaging of HIV-1 genomic RNA [89–92]. However, a Gag precursor lacking the p6 domain (GagΔp6) was used in all these studies. When a full-

length Gag precursor was used, specific binding of Gag to *Psi*-containing RNA was observed [28,44,46] that is consistent with enrichment of HIV-1 genomic RNA in viral particles [93]. In HIV-1 GagΔp6, nonspecific electrostatic interactions overweigh specific interactions [89,92]. It is likely that the p6 domain, which is negatively charged, contributes to the specific binding of full-length Gag to *Psi* by neutralizing the positive charges of the NC domain [28]. Besides, Mutational Interference Mapping Experiment (MIME), an unbiased exhaustive approach, revealed a good correlation between mutations that decreased the binding affinity of full-length HIV-1 Gag for *Psi*-containing RNA [45] with those that decreased packaging of the genomic RNA [26] indicating that, in HIV-1, specific binding of Gag to *Psi* does contribute to packaging. However, some mutations affected RNA packaging without affecting Gag binding [26]. While some of these mutations were found to affect RNA metabolism, thus reducing packaging indirectly, the effect of mutations in the PBS domain remained unexplained. Therefore, it is possible that specific Gag binding to *Psi* is not the only mechanism ensuring selective packaging of the genomic HIV-1 RNA, and kinetics could also affect selectivity of the packaging process [91]. Hence, in the case of MMTV, selection of the gRNA may indeed primarily rely on the kinetic advantage provided by *Psi*.

Indeed, our footprinting experiments (Figs 10–13 and S4–S10) showed that efficient gRNA packaging correlates with Pr77^{Gag} binding in the PBS and ssPurines regions of *Psi*. These results are in line with our previous study that showed that mutations in the PBS and ssPurine regions of MMTV *Psi* dramatically reduced Pr77^{Gag} binding and gRNA packaging [53]. It is also noteworthy that in the case of HIV-1 and many other retroviruses, such as HIV-2, SIV, and MPMV, multiple unpaired purines (single-stranded purines) play a key role in the selective encapsidation of viral RNA [80,94–98]. This points towards a highly conserved role of unpaired purines in selective packaging of retroviral RNAs irrespective of their assembly mode in different parts of the cell. With our non-packaging mutants, we observed scattered footprints even though the local structure of the PBS and ssPurines regions was not altered (Figs 4–6). This suggests that alteration of the overall 3D structure of the *Psi* of the non-packaging mutants, caused by disruption of the extended LRI III'-IV, prevents Pr77^{Gag} binding at these specific sites and promotes binding to nonspecific nucleotides. This indicates that for MMTV gRNA to be selectively packaged, Pr77^{Gag} must bind to specific nucleotides in the correct structural context rather than binding to any nucleotide(s) when the proper structural context is lost, as has been proposed for HIV-1 [99].

Why does efficient Pr77^{Gag} binding to nonspecific nucleotides not promote RNA packaging? RNA is known to be a crucial structural element of retroviral particles [100] and it induces Gag multimerization and Gag assembly [91,101]. Previous experiments suggested that HIV-1 *Psi* more efficiently promotes in vitro Gag/RNA assemblies than heterologous RNAs [91]. Our present study supports a model in which Gag must bind to *Psi* in the correct 3D context to promote efficient assembly of viral particles. Indeed, computer simulations showed that the RNA folding geometry of the packaging signal affects the assembly activation energy barrier, allowing kinetic selectivity of the genomic RNA [102].

Conclusions

Here, we investigated the relationship between the structure of MMTV gRNA and its ability to interact with Pr77^{Gag}. Our study demonstrates that combining genetic, biological, biochemical, and structure–function approaches can provide detailed insights into RNA–RNA and RNA–protein interactions that govern fundamental biological processes during retroviral replication. Specifically, we show that Pr77^{Gag} must bind to specific nucleotides in the correct structural context, which leads to the downstream events of Gag assembly, facilitating selective

MMTV gRNA packaging. Furthermore, our work also suggests that Pr77^{Gag} possesses the ability to engage with various RNA conformations with different consequences, suggesting that these mechanistic distinctions could have significant implications for MMTV gRNA packaging. This highlights the importance of a native fold towards identifying well-defined protein-binding sites that promotes function. Thus far, to the best of our knowledge, this is the first demonstration of the intimate connection between retroviral RNA structure, proper Gag binding to specific sites in *Psi*, and efficient encapsidation of the viral RNA in the nascently forming virions.

Material and methods

MMTV strain and nucleotide numbers

The MMTV nucleotide numbering system is based on the sequence of HYBMTV, a molecular clone created earlier [103]. This clone was employed for constructing wild-type and mutant transfer vectors used in this study.

Expression plasmids and transfer vector

JA10 is an MMTV packaging construct that expresses the MMTV *gag-pro-pol* genes ([67]; S1 Fig); the MD.G plasmid serves as an expression vector for the vesicular stomatitis virus glycoprotein G (VSV-G) ([68]; S1 Fig), and the DA024 plasmid ([67]; S1 Fig) is the MMTV subgenomic transfer vector which functions as a source of packageable RNA. DA024 also contains the *hygromycin B phosphotransferase* gene, which is used to monitor the propagation of the packaged RNA in the transduced cells [67].

Cloning and mutagenesis

Mutations designed to disrupt or re-establish LRIs were introduced through splice overlap extension (SOE) PCR, utilizing DA024 as the template, using outer sense (S; OTR 249) and antisense (AS; OTR 552) primers (S2 Table). Custom mutation-specific inner primers (S2 Table) were employed along with outer primers in initial 2 separate amplification reactions. Subsequently, the 2 amplified products were combined in a final PCR using outer primers. The final amplified products, flanked by *Spe*I sites, were cleaved and cloned into DA024, previously cleaved with the same restriction enzyme. Mutant clones were confirmed by sequencing.

Mutations in MMTV packaging sequences that were employed for in vivo biological studies were cloned into DA024 [67], which is referred to as the WT and the resultant mutant clones are named as SP101-SP109. For biochemical studies, nucleotides 1–712 of the MMTV gRNA (with nucleotide +1 denoting the start of the R (repeat) region) were PCR-amplified from DA024 or its mutant counterparts, employing already established procedures [53,54]. This amplification utilized primers OTR 984 (S) harboring T7 promoter sequences and a *Hind*III restriction site, and OTR 985 (AS) with a *Xma*I restriction site for T7 promoter insertion at its 5′ end and for the ease of cloning (S2 Table). The resulting amplified products were cloned into pIC19R [104], creating SP101*i*, 102*i*, SP105*i*-SP109*i*. Mutants containing the T7 promoter were also confirmed through sequencing.

Genetic complementation assay to study single round of MMTV replication

For monitoring the effects of the mutations in the LRIs on gRNA packaging and RNA transduction efficiencies in a biologically relevant assay, we employed a previously established in vivo three-plasmid genetic complementation system ([61,67]; S1 Fig). The MMTV sub-

genomic transfer vector (DA024; [67]), harboring the essential *cis*-acting sequences needed for RNA packaging, reverse transcription, and integration, served as the source of RNA that is competent for packaging in virus particles produced by JA10 and pseudotyped by vesicular stomatitis virus envelope glycoprotein (VSV-G) expressed by MD.G [61,67,68]. An additional plasmid DNA (pSEAP2-control vector, Clontech, USA) expressing the *secreted alkaline phosphatase* (*SEAP*) gene was utilized to assess transfection efficiency of each culture. The 4 plasmids were co-transfected into human embryonic kidney (HEK293T) using calcium phosphate transfection kit (Thermo Fisher Scientific, USA). *SEAP* expression was measured employing the Great EscAPe SEAP Chemiluminescence Kit 2.0, (Clontech, USA), as previously described [79,105,106].

Since the replication of the packaged RNA is confined to a single round (S1 Fig), the inclusion of the *hygromycin B phosphotransferase* gene in DA024 allows the quantitative monitoring of propagation of the packaged RNA by counting hygromycin-resistant colonies. Therefore, a portion of the supernatant from the transfected cultures was used to infect human cervical cancer cells (Hela CD4+) in the presence of 1 μg/ml diethylaminoethyl (DEAE)-dextran. This involved selecting the infected cells with cell culture medium containing 200 μg/ml hygromycin B (Hyclone). Cells transduced by the packageable transfer vector RNA expressing the *hygromycin B phosphotransferase* gene were then counted and expressed as colony-forming units (CFUs/ml). The CFU/ml values were normalized to the transfection efficiency of each transfected culture, determined by the relative SEAP values, and divided by the values obtained from the WT DA024 vector to obtain the transduction of mutant RNAs relative to the WT (relative CFU/ml). The number of hygromycin-resistant (Hyg^R) colonies obtained is expected to be directly proportional to the amount of RNA packaged into the virus particles, providing a good estimate of RNA packaging efficiency.

Nucleocytoplasmic fractionation of the transfected cells, isolation of cellular and viral RNAs and cDNA synthesis

Transfected cells were fractionated into nuclear and cytoplasmic fractions, following previously established protocols [107]. In summary, virus particles produced from transfected cultures were purified and pelleted through ultracentrifugation, and viral RNA was extracted, treated with TURBO DNase (Invitrogen, Thermo Fisher Scientific) and amplified for 30 cycles employing conventional PCR with established conditions using vector-specific primers OTR 671(S) and OTR 672 (AS) (S2 Table) to test for any residual contaminating plasmid DNA in RNA preparations [36]. In parallel, DNase-treated RNAs were reverse transcribed into cDNAs [53,96,97] and the amplifiability of cDNA preparations was tested through multiplex PCR as described earlier [8,61]. Integrity of the nuclear membrane during fractionation was assessed by multiplex PCR, amplifying cDNAs from the cytoplasmic fraction with oligos OTR 582 (S) and OTR 581 (AS) (S2 Table), specifically targeting unspliced β-actin mRNA. Additionally, as a control for the presence of amplifiable cDNA in multiplex PCRs, cDNAs were amplified using primers/competimer for 18S ribosomal RNA (18S Quantum competimer control, Ambion). Finally, the stability and proper nuclear export of mutant transfer vector RNAs were confirmed by amplifying cytoplasmic cDNA preparations using OTR 671/OTR 672 MMTV-specific primer pair (S2 Table).

Quantitation of relative RNA packaging efficiency (RPE) using RT-qPCR

To assess the relative levels of transfer vector RNAs in the cytoplasm and in virions, we developed a real-time PCR custom expression assay in a region that was common to all the RNAs and away from the mutation sites, employing minor groove binding (MGB)/FAM chemistry

(Applied Biosystems, USA). The Applied Biosystems RT-qPCR assays utilize carboxy-X-rhodamine (ROX) as an internal fluorescence reference dye which is used for normalizing the target reporter dye signal during data analysis. Primers for the PCR were designed to allow annealing to the 101-nucleotide (nt) long region at the beginning of MMTV U5 region shared among all mutant and WT transfer vectors tested. This design enabled their simultaneous relative quantitative assessment using the same probe/primer combination. Detailed information about the probe and primers is provided in S2 Table. In silico validation using the "pcrEfficiency" tool [108] and ABI design pipeline bioinformatic tools confirmed the assay's specificity, reproducibility, and amplification efficiency. An optimized β-actin MGB FAM labeled assay from ABI (TaqMan 4351368; Applied Biosystems, USA) was employed as an endogenous control.

To standardize the comparative (ΔΔCt) method for relative quantification (RQ) analysis, DNA from plasmids (DA024 and one expressing β-actin) were serially diluted and subjected to non-multiplex RT-qPCRs in triplicates (S2A Fig; panels I and II). The resulting Ct values were then plotted on a scatter plot against the μg amount of plasmid DNA in each dilution (in triplicates) to generate standard curves for both assays (S2B Fig; panels I and II). Subsequently, ΔCt values were calculated (Ct values from the target MMTV assay minus Ct value of the β-actin endogenous control) and plotted against the input plasmid DNA to determine the slope of the curve (S2C Fig).

To quantify gRNA packaged in virus particles, cDNAs derived from viral and cytoplasmic RNAs were tested in triplicates using the Taqman Universal Master Mix (Applied Biosystems) and the 7500 Real Time PCR System (Applied Biosystems, USA). Prior to cDNA preparation, extracted RNA samples were DNase treated and subjected to PCR using vector-specific primers. No detectable amplification could be observed in these DNase-treated RNA samples after 30 cycles of PCR, ensuring lack of any measurable plasmid DNA contamination in these samples, under these conditions. Next, cDNAs were prepared and assessed for vector expression in cytoplasmic fractions and their ability to package into viral particles using the real-time PCR assay under the following cycling conditions: an initial denaturation step of 10 min at 94°C, followed by 40 cycles of denaturation and annealing/extension steps at 95°C for 15 s and 60°C for 1 min. The cytoplasmic RNA expression was normalized to control for transfection efficiency. The determination of the RPE for each mutant transfer vector RNA was achieved by dividing the viral RQ with that of the normalized cytoplasmic RQ and reported relative to that of the WT.

In vitro transcription and purification of in vitro transcribed RNA

Plasmids carrying the WT (SA35) and mutant MMTV packaging sequences under the control of the T7 promoter were linearized with *Sma*I and utilized for in vitro transcription using MEGAscript T7 Transcription Kit (Thermo Fisher Scientific). The quality of produced RNA was assessed by testing 2 μl of the in vitro transcription reaction on 8% denaturing (8 M urea) polyacrylamide gels. The remaining in vitro transcribed product was precipitated overnight at −20°C using ethanol and the pellets were resuspended in 500 μl of Milli-Q water. Next, these RNAs were purified through gel filtration chromatography using a TSK Gel G4000SW column (Tosoh Bioscience) in 0.2 M sodium acetate (pH 6.5) containing 1% (v/v) methanol, following previously established protocols [109–111]. Subsequently, the appropriate fractions were pooled, ethanol precipitated, and examined for purity and integrity as described above.

Internally labeled RNAs were prepared by conducting in vitro transcription in the presence of (α-^{32}P)-ATP, following previously described methods [46,110,111]. The DNase-treated labeled RNAs were purified using Micro Bio-Spin chromatography columns (BioRad), as per manufacturer's instructions.

Expression and purification of full-length recombinant MMTV Pr77^{Gag} and western blot analysis

Recombinant full-length MMTV Gag containing hexa-histidine tag at the C-terminus (Pr77^{Gag}) was expressed and purified as described previously [79,105,106,112]. The purified Pr77^{Gag} was characterized by western blotting using anti-His$_6$ as well as MMTV anti-p27 monoclonal antibodies to confirm purity [79].

Characterization of Pr77^{Gag} by dynamic light scattering (DLS)

MMTV Pr77^{Gag} was further characterized in the storage buffer (50 mM Tris-HCl (pH 8.0) and 1 M NaCl) by DLS as described earlier [46]. By assimilating the proteins in solution to spheres, their diffusion coefficient was correlated to their hydrodynamic radius (R_h) using the Stokes–Einstein equation:

$$D = \frac{kT}{6\pi\mu R h}.$$

In this equation, k represents the Boltzmann constant, T is the absolute temperature, and μ is the viscosity of the solvent. Before acquiring reading, the buffer underwent filtration through 0.2-micron filters (Millex) and buffer solvent offset was measured for subsequent data analysis.

Filter-binding assay

For filter-binding experiments, 25,000 cpm of internally ^{32}P-labeled RNAs were denatured in the presence of 5 nM of the cognate unlabeled RNA and 0.4 μg of yeast tRNA at 90°C for 2 min, following which the samples were chilled on ice for 2 min. The denatured RNAs were incubated at 37°C for 30 min in 1× RNA folding buffer (30 mM Tris-HCl (pH 8), 300 mM NaCl, 5 mM MgCl$_2$), along with 5U of Rnasin and 0.01% Triton X-100. Reaction mixtures underwent an additional 30-min incubation with increasing concentrations of protein in 1× protein buffer (30 mM Tris-HCl (pH 8), 300 mM NaCl, 5 mM MgCl$_2$, 10 mM DTT, 0.02 mg/ml BSA) to form protein-RNA complexes, which were stabilized by incubating on ice for 30 min.

Filter-binding assays were performed utilizing 0.45 μm nitrocellulose membranes (BioRad) that were pre-wetted with 1× TBS (Tris buffered saline: 20 mM Tris-HCl (pH 7.5) and 500 mM NaCl) for 10 min at room temperature. After drying the membrane on a filter paper, it was placed onto a Bio-Dot Microfiltration unit (BioRad) attached to a vacuum suction drainage. Each well was washed with 100 μl of 1× buffer before applying the protein-RNA complex. Following the addition of 40 μl of cold 1× Gag-binding buffer (30 mM Tris-HCl (pH 8), 300 mM NaCl, 5 mM MgCl$_2$), 20 μl of the reaction mixture was applied to each well and incubated at room temperature for 10 min. Following sample filtration, wells were washed 3 times with 100 μl cold 1× Gag-binding buffer to remove any unbound RNA, the nitrocellulose membrane was removed from filtration unit and air-dried. Finally, the membranes were exposed with an imaging plate (Fujifilm), scanned using a FLA 5000 (Fuji) scanner, and quantification was performed using ImageQuant (Cytiva) software, as described earlier [46]. The GraphPad Prism version 8 (v8) software was employed to fit the experimental data with the Hill equation:

$$Y = \frac{B_{max} \times X^h}{\left(K_d{}^h + X^h\right)}.$$

In this equation, B_{max} represents the maximum specific binding, K_d represents the equilibrium dissociation constant, and h is the Hill slope.

High-throughput selective 2′-hydroxyl acylation analyzed by primer extension (hSHAPE) and Pr77 [Gag] RNA footprints on MMTV gRNA

The structure of RNAs was interrogated using the hSHAPE methodology, enabling a detailed examination of each nucleotide through structure-dependent RNA modification [35,36,46,52–54,72–75,80,96,113]. hSHAPE utilizes a single reagent such as benzoyl cyanide (BzCN), to modify all 4 nucleotides simultaneously, which are more reactive when unpaired. In contrast, the base-paired nucleotides show minimal reactivity to the hSHAPE reagent. To get the Pr77[Gag] footprint on WT and LRI mutant RNAs, BzCN was used to modify WT RNA and LRI mutant RNAs both in the absence and presence of Pr77[Gag]. Nucleotides displaying reduced reactivity in the presence of Pr77[Gag] indicated protein's binding to these nucleotides. This approach has proven successful in identifying Gag-binding sites on the respective gRNAs of HIV-1, MMTV, and MPMV [46,53,80]. Footprinting assays were conducted in the presence of *env* mRNA to mimic the physiological conditions, resulting in obtaining only highly specific Gag-binding sites.

Briefly, for hSHAPE the WT or mutant RNAs were modified in the absence or presence of Pr77[Gag] and reverse transcription was performed as described previously [53] using VIC-labeled primers (MMTV_WT_VIC and MMTV_329_VIC) (S2 Table). Sequencing reactions using ddGTP were prepared in parallel using NED-labeled primers (MMTV_WT_NED and MMTV_329_NED) (S2 Table) as previously described [53,81]. The resulting cDNAs from both reactions were combined, precipitated, and denatured before analysis using an Applied Biosystems 3130xl genetic analyzer. The obtained electropherograms were then analyzed with the QuShape software [76]. Normalized hSHAPE reactivity data from a minimum of 3 to 4 independent experiments was applied as pseudo-energy constraints for folding the secondary structure of the MMTV packaging signal RNA. The RNA folding was achieved using the RNAstructure version 6.1 program [77], while the VARNA v3-93 software [114] was used to redraw the RNA structures. The hSHAPE reactivity data for each nucleotide was incorporated into the structure as described previously [114]. For RNA footprinting experiments, the hSHAPE reactivities obtained in the presence of 4 μm Pr77[Gag] was applied onto the RNA structure obtained using reactivity in the absence of Pr77[Gag]. The Mann–Whitney non parametrical U test was conducted to assess statistically significant differences in hSHAPE reactivities in the absence and presence of Pr77[Gag]. Variations in hSHAPE reactivity showing a variance of ≥ 0.20 and a relative difference exceeding 40% to 50% were deemed significant, as reported earlier for HIV-1, MMTV, and MPMV [46,53,80,81].

In vitro dimerization assays

Dimerization of 300 nM of the purified WT and LRI mutant RNAs in either a dimer buffer (30 mM Tris (pH 7.5), 300 mM NaCl, 5 mM $MgCl_2$) or monomer buffer (30 mM Tris (pH 7.5), 300 mM NaCl, 0.1 mM $MgCl_2$) was performed using previously established protocols [35,53]. Following electrophoresis of the samples through a native 1% agarose gel in TBM (50 mM Tris base, 45 mM boric acid, 0.1 mM $MgCl_2$) at 4°C, gels were scanned using a Gel Doc EZ Imager (BioRad) and monomeric and dimeric bands were quantified using Image Lab (BioRad) software, as described previously [53,54]. The weight/weight percentage (%) of dimerization efficiency was determined by dividing the intensity of the dimeric RNA band by the sum of the intensities of the dimer and monomer bands. These values were plotted as the percent of dimerization relative to the wild type for each LRI mutation introduced.

Statistical analysis

We used GraphPad Prism version 8 (v8) software to plot individual data set points. Statistical analysis was performed using the Mann–Whitney non parametrical U test. Results were considered significant when the *p*-value was <0.05.

Supporting information

S1 Fig. Schematic representation of the MMTV three-plasmid genetic complementation assay. The MMTV three-plasmid genetic complementation assay was designed with the following rationale: virus particles generated from the MMTV Gag/Pro/Pol expression plasmid (JA10), pseudotyped with vesicular stomatitis virus envelope glycoprotein (VSV-G) expressed by MD.G, facilitate packaging of MMTV subgenomic transfer vector (DA024) RNA, containing a functional RNA packaging sequences (Ψ). HEK293T cells were co-transfected with these 3 plasmids to produce pseudotyped infectious virus particles capable of only 1 round of replication. Transfected cells were fractionated into nuclear and cytoplasmic fractions to isolate cytoplasmic RNA and analyzed for RNA stability and efficient nuclear export. Viral RNA isolated from the pelleted virus particles was used to quantify the packaged RNA using RT-qPCR. Viral supernatants were used to infect target HeLa CD4+ cells in order to assess the ability of propagation of the encapsidated RNA via selection with media supplemented with hygromycin B antibiotic. This allowed selection of cells transduced by the packaged RNA containing the *hygromycin resistance* gene cassette and appearing as hygromycin resistant colonies. (PDF)

S2 Fig. Validation of the custom made MMTV TaqMan assay for relative quantification of WT and mutant transfer vector RNAs. Estimation of the amplification efficiency of **(A: panel I)** the custom-made MMTV TaqMan assay, **(A: panel II)** the commercially available β-actin TaqMan assay. ΔRn = Normalized Reporter (Rn)—baseline). Standard curves were generated for both **(B: panel I)** MMTV, and **(B: panel II)** β-actin TaqMan assays. **(C)** Relative amplification efficiency plot of MMTV and β-actin TaqMan assays. To ensure similar amplification efficiencies, the slope of log input amount vs. ΔCt should be close to zero (ideally ≤0.1). In our experimental conditions, this slope was calculated to be 0.1076, validating the assay for relative quantification analysis. The data underlying this figure can be found in S1 Data. (PDF)

S3 Fig. Characterization of the recombinant MMTV full-length Pr77^{Gag}-His$_6$-tag fusion precursor protein. (A) Coomassie brilliant blue stained SDS-PAGE of the purest form of the recombinant full-length MMTV Pr77^{Gag}-His6-tag fusion protein post-size exclusion chromatography. **(B)** Western blot analysis using α-His$_6$ and MMTV α-p27 monoclonal antibodies. **(C, D)** Characterization of the full-length MMTV Pr77^{Gag}-His$_6$-tag fusion protein conducted via dynamic light scattering (DLS) in binding buffer, showing: protein mass vs. size distribution represented as hydrodynamic radius (Rh) distribution and protein number vs. hydrodynamic radius (Rh) distribution, respectively. The data underlying this S3C and S3D Fig can be found in S1 Data. (PDF)

S4 Fig. Footprints of Pr77^{Gag} on LRI-I destabilizing mutant SP101*i* RNA secondary structure model. hSHAPE analysis was carried out both with and without Pr77^{Gag}. The mean triplicate SHAPE reactivity obtained without Pr77^{Gag} was used to predict the RNA secondary structure model. Subsequently, the mean hSHAPE reactivities obtained with Pr77^{Gag} were overlaid onto the RNA secondary structure model predicted in the absence of Pr77^{Gag}.

Nucleotides marked by arrows show significant reduction in hSHAPE reactivities according to the Mann–Whitney non parametrical U test ($p < 0.05$). The hSHAPE reactivity key was developed based on the mean of hSHAPE reactivities for each nucleotide, as shown in S3 Table. The data shown is from a minimum of 3 independent experiments conducted both in the absence and presence of Pr77Gag.
(PDF)

S5 Fig. Footprints of Pr77Gag on LRI-I re-stabilizing mutant SP102*i* RNA secondary structure model. hSHAPE analysis was carried out both with and without Pr77Gag. The mean triplicate SHAPE reactivity obtained without Pr77Gag was used to predict the RNA secondary structure model. Subsequently, the mean hSHAPE reactivities obtained with Pr77Gag were overlaid onto the RNA secondary structure model predicted in the absence of Pr77Gag. Nucleotides marked by arrows show significant reduction in hSHAPE reactivities according to the Mann–Whitney non parametrical U test ($p < 0.05$). The hSHAPE reactivity key was developed based on the mean of hSHAPE reactivities for each nucleotide, as shown in S3 Table. The data shown is from a minimum of 3 independent experiments conducted both in the absence and presence of Pr77Gag.
(PDF)

S6 Fig. Footprints of Pr77Gag on LRI-III destabilizing mutant SP105*i* RNA secondary structure model. hSHAPE analysis was carried out both with and without Pr77Gag. The mean triplicate SHAPE reactivity obtained without Pr77Gag was used to predict the RNA secondary structure model. Subsequently, the mean hSHAPE reactivities obtained with Pr77Gag were overlaid onto the RNA secondary structure model predicted in the absence of Pr77Gag. Nucleotides marked by arrows show significant reduction in hSHAPE reactivities according to the Mann–Whitney non parametrical U test ($p < 0.05$). The hSHAPE reactivity key was developed based on the mean of hSHAPE reactivities for each nucleotide, as shown in S3 Table. The data shown is from a minimum of 3 independent experiments conducted both in the absence and presence of Pr77Gag.
(PDF)

S7 Fig. Footprints of Pr77Gag on LRI-III restabilizing mutant SP106*i* RNA secondary structure model. hSHAPE analysis was carried out both with and without Pr77Gag. The mean triplicate SHAPE reactivity obtained without Pr77Gag was used to predict the RNA secondary structure model. Subsequently, the mean hSHAPE reactivities obtained with Pr77Gag were overlaid onto the RNA secondary structure model predicted in the absence of Pr77Gag. Nucleotides marked by arrows show significant reduction in hSHAPE reactivities according to the Mann–Whitney non parametrical U test ($p < 0.05$). The hSHAPE reactivity key was developed based on the mean of hSHAPE reactivities for each nucleotide, as shown in S3 Table. The data shown is from a minimum of 3 independent experiments conducted both in the absence and presence of Pr77Gag.
(PDF)

S8 Fig. Footprints of Pr77Gag on LRI-III' re-stabilizing mutant SP109*i* RNA secondary structure model. hSHAPE analysis was carried out both with and without Pr77Gag. The mean triplicate SHAPE reactivity obtained without Pr77Gag was used to predict the RNA secondary structure model. Subsequently, the mean hSHAPE reactivities obtained with Pr77Gag were overlaid onto the RNA secondary structure model predicted in the absence of Pr77Gag. Nucleotides marked by arrows show significant reduction in hSHAPE reactivities according to the Mann–Whitney non parametrical U test ($p < 0.05$). The hSHAPE reactivity key was developed based on the mean of hSHAPE reactivities for each nucleotide, as shown in S3 Table. The data

shown is from a minimum of 3 independent experiments conducted both in the absence and presence of Pr77^{Gag}.
(PDF)

S9 Fig. Footprints of Pr77^{Gag} on LRI-IV destabilizing mutant SP107i RNA secondary structure model. hSHAPE analysis was carried out both with and without Pr77^{Gag}. The mean triplicate SHAPE reactivity obtained without Pr77^{Gag} was used to predict the RNA secondary structure model. Subsequently, the mean hSHAPE reactivities obtained with Pr77^{Gag} were overlaid onto the RNA secondary structure model predicted in the absence of Pr77^{Gag}. Nucleotides marked by arrows show significant reduction in hSHAPE reactivities according to the Mann–Whitney non parametrical U test ($p < 0.05$). The hSHAPE reactivity key was developed based on the mean of hSHAPE reactivities for each nucleotide, as shown in S3 Table. The data shown is from a minimum of 3 independent experiments conducted both in the absence and presence of Pr77^{Gag}.
(PDF)

S10 Fig. Footprints of Pr77^{Gag} on LRI-IV re-stabilizing mutant SP108i RNA secondary structure model. hSHAPE analysis was carried out both with and without Pr77^{Gag}. The mean triplicate SHAPE reactivity obtained without Pr77^{Gag} was used to predict the RNA secondary structure model. Subsequently, the mean hSHAPE reactivities obtained with Pr77^{Gag} were overlaid onto the RNA secondary structure model predicted in the absence of Pr77^{Gag}. Nucleotides marked by arrows show significant reduction in hSHAPE reactivities according to the Mann–Whitney non parametrical U test ($p < 0.05$). The hSHAPE reactivity key was developed based on the mean of hSHAPE reactivities for each nucleotide, as shown in S3 Table. The data shown is from a minimum of 3 independent experiments conducted both in the absence and presence of Pr77^{Gag}.
(PDF)

S1 Table. Mean hSHAPE reactivities with standard deviations (SD) of first 432 nucleotides of WT unspliced RNA (SA35) and LRI mutant clones (SP101i, SP 102i, SP 105i-SP 109i) from 3 independent experiments. The yellow highlighted region shows mutations introduced in LRI mutants.
(PDF)

S2 Table. Primer details used for introduction of mutations, construction of clones, hSHAPE, RT-qPCR, and other amplifications.
(PDF)

S3 Table. Mean hSHAPE reactivities from 3 experiments in the absence and presence of Pr77^{Gag}. The yellow highlighted nucleotides showed ≥1.5-fold reduction of hSHAPE reactivities in the presence of Pr77^{Gag}.
(PDF)

S1 Data. Excel spreadsheet containing, in separate sheets for each figure, the underlying and individual numerical data used for Figs 2C, 2D, 3D, 8A, 9A, 9B, 11A, 11B, 11C, 12A, 12B, 12C, 12D, 13A, 13B, 13C, S2A–S2C, S3C and S3D.
(XLSX)

S1 Raw Images. Uncropped gel images used to make panel I of Fig 2B, panels II and III of Figs 2B, 3C, S3A and S3B.
(PDF)

Author Contributions

Conceptualization: Farah Mustafa, Roland Marquet, Tahir A. Rizvi.

Data curation: Suresha G. Prabhu.

Formal analysis: Suresha G. Prabhu, Vineeta N. Pillai, Lizna Mohamed Ali, Valérie Vivet-Boudou, Akhil Chameettachal, Serena Bernacchi.

Funding acquisition: Roland Marquet, Tahir A. Rizvi.

Investigation: Vineeta N. Pillai, Lizna Mohamed Ali, Valérie Vivet-Boudou, Akhil Chameettachal, Serena Bernacchi, Tahir A. Rizvi.

Methodology: Roland Marquet, Tahir A. Rizvi.

Project administration: Roland Marquet, Tahir A. Rizvi.

Resources: Roland Marquet, Tahir A. Rizvi.

Supervision: Farah Mustafa, Roland Marquet, Tahir A. Rizvi.

Writing – original draft: Suresha G. Prabhu.

Writing – review & editing: Suresha G. Prabhu, Vineeta N. Pillai, Lizna Mohamed Ali, Valérie Vivet-Boudou, Akhil Chameettachal, Serena Bernacchi, Farah Mustafa, Roland Marquet, Tahir A. Rizvi.

References

1. Hunter E. Chapter 6, Viral Assembly. 6th Edition. Fields Virology. 6th Edition. 2013. p. 144–146.

2. Flint SJ, Enquist LW, Racaniello VR, Skalka AM. Chapter 13. 2nd Edition. Principles of Virology. 2nd Edition. 2000. p. 466–474.

3. Paillart J-C, Shehu-Xhilaga M, Marquet R, Mak J. Dimerization of retroviral RNA genomes: an inseparable pair. Nat Rev Microbiol. 2004; 2:461–472. https://doi.org/10.1038/nrmicro903 PMID: 15152202

4. Dubois N, Marquet R, Paillart J-C, Bernacchi S. Retroviral RNA Dimerization: From Structure to Functions. Front Microbiol. 2018; 9:527. https://doi.org/10.3389/fmicb.2018.00527 PMID: 29623074

5. Berkhout B, van Wamel JL. Role of the DIS hairpin in replication of human immunodeficiency virus type 1. J Virol. 1996; 70:6723–6732. https://doi.org/10.1128/JVI.70.10.6723-6732.1996 PMID: 8794309

6. Das AT, Klaver B, Klasens BI, van Wamel JL, Berkhout B. A conserved hairpin motif in the R-U5 region of the human immunodeficiency virus type 1 RNA genome is essential for replication. J Virol. 1997; 71:2346–2356. https://doi.org/10.1128/JVI.71.3.2346-2356.1997 PMID: 9032371

7. Clever JL, Miranda D, Parslow TG. RNA Structure and Packaging Signals in the 5′ Leader Region of the Human Immunodeficiency Virus Type 1 Genome. J Virol. 2002; 76:12381–12387. https://doi.org/10.1128/jvi.76.23.12381-12387.2002 PMID: 12414982

8. Mustafa F, Ghazawi A, Jayanth P, Phillip PS, Ali J, Rizvi TA. Sequences Intervening between the Core Packaging Determinants Are Dispensable for Maintaining the Packaging Potential and Propagation of Feline Immunodeficiency Virus Transfer Vector RNAs. J Virol. 2005; 79:13817–13821. https://doi.org/10.1128/JVI.79.21.13817-13821.2005 PMID: 16227303

9. Kenyon JC, Ghazawi A, Cheung WKS, Phillip PS, Rizvi TA, Lever AML. 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. 2008; 14:2597–2608. https://doi.org/10.1261/rna.1284908 PMID: 18974279

10. Rizvi TA, Kenyon JC, Ali J, Aktar SJ, Phillip PS, Ghazawi A, et al. Optimal packaging of FIV genomic RNA depends upon a conserved long-range interaction and a palindromic sequence within gag. J Mol Biol. 2010; 403:103–119. https://doi.org/10.1016/j.jmb.2010.08.019 PMID: 20732330

11. Jouvenet N, Lainé S, Pessel-Vivares L, Mougel M. Cell biology of retroviral RNA packaging. RNA Biol. 2011; 8:572–580. https://doi.org/10.4161/rna.8.4.16030 PMID: 21691151

12. Ali LM, Rizvi TA, Mustafa F. Cross- and Co-Packaging of Retroviral RNAs and Their Consequences. Viruses. 2016; 8:276. https://doi.org/10.3390/v8100276 PMID: 27727192

13. Bieniasz P, Telesnitsky A. Multiple, Switchable Protein:RNA Interactions Regulate Human Immunodeficiency Virus Type 1 Assembly. Annu Rev Virol. 2018; 5:165–183. https://doi.org/10.1146/annurev-virology-092917-043448 PMID: 30048218

14. Chameettachal A, Mustafa F, Rizvi TA. Understanding Retroviral Life Cycle and its Genomic RNA Packaging. J Mol Biol. 2023; 435:167924. https://doi.org/10.1016/j.jmb.2022.167924 PMID: 36535429

15. Comas-Garcia M, Davis SR, Rein A. On the Selective Packaging of Genomic RNA by HIV-1. Viruses. 2016; 8:246. https://doi.org/10.3390/v8090246 PMID: 27626441

16. D'Souza V, Summers MF. How retroviruses select their genomes. Nat Rev Microbiol. 2005; 3:643. https://doi.org/10.1038/nrmicro1210 PMID: 16064056

17. Johnson SF, Telesnitsky A. Retroviral RNA Dimerization and Packaging: The What, How, When, Where, and Why. PLoS Pathog. 2010; 6:e1001007. https://doi.org/10.1371/journal.ppat.1001007 PMID: 20949075

18. Kaddis Maldonado RJ, Parent LJ. Orchestrating the Selection and Packaging of Genomic RNA by Retroviruses: An Ensemble of Viral and Host Factors. Viruses. 2016; 8. https://doi.org/10.3390/v8090257 PMID: 27657110

19. Mailler E, Bernacchi S, Marquet R, Paillart J-C, Vivet-Boudou V, Smyth RP. The Life-Cycle of the HIV-1 Gag-RNA Complex. Viruses. 2016; 8. https://doi.org/10.3390/v8090248 PMID: 27626439

20. Kutluay SB, Zang T, Blanco-Melo D, Powell C, Jannain D, Errando M, et al. Global changes in the RNA binding specificity of HIV-1 gag regulate virion genesis. Cell. 2014; 159:1096–1109. https://doi.org/10.1016/j.cell.2014.09.057 PMID: 25416948

21. Lever AML. HIV-1 RNA packaging. Adv Pharmacol San Diego Calif. 2007; 55:1–32. https://doi.org/10.1016/S1054-3589(07)55001-5 PMID: 17586311

22. Lu K, Heng X, Summers MF. Structural determinants and mechanism of HIV-1 genome packaging. J Mol Biol. 2011; 410:609–633. https://doi.org/10.1016/j.jmb.2011.04.029 PMID: 21762803

23. Lu K, Heng X, Garyu L, Monti S, Garcia EL, Kharytonchyk S, et al. NMR detection of structures in the HIV-1 5'-leader RNA that regulate genome packaging. Science. 2011; 334:242–245. https://doi.org/10.1126/science.1210460 PMID: 21998393

24. Miyazaki Y, Miyake A, Nomaguchi M, Adachi A. Structural dynamics of retroviral genome and the packaging. Front Microbiol. 2011; 2:264. https://doi.org/10.3389/fmicb.2011.00264 PMID: 22232618

25. Rulli SJ, Hibbert CS, Mirro J, Pederson T, Biswal S, Rein A. Selective and nonselective packaging of cellular RNAs in retrovirus particles. J Virol. 2007; 81:6623–6631. https://doi.org/10.1128/JVI.02833-06 PMID: 17392359

26. Smyth RP, Smith MR, Jousset A-C, Despons L, Laumond G, Decoville T, et al. 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. 2018. https://doi.org/10.1093/nar/gky152 PMID: 29514260

27. Browning MT, Mustafa F, Schmidt RD, Lew KA, Rizvi TA. Delineation of sequences important for efficient packaging of feline immunodeficiency virus RNA. J Gen Virol. 2003; 84:621–627. https://doi.org/10.1099/vir.0.18886-0 PMID: 12604814

28. Dubois N, Khoo KK, Ghossein S, Seissler T, Wolff P, McKinstry WJ, et al. The C-terminal p6 domain of the HIV-1 Pr55Gag precursor is required for specific binding to the genomic RNA. RNA Biol. 2018; 15:923–936. https://doi.org/10.1080/15476286.2018.1481696 PMID: 29954247

29. Rawson JMO, Nikolaitchik OA, Yoo JA, Somoulay X, Brown MA, Mbuntcha Bogni FS, et al. Adaptation of HIV-1/HIV-2 Chimeras with Defects in Genome Packaging and Viral Replication. MBio. 2022; 13: e0222022. https://doi.org/10.1128/mbio.02220-22 PMID: 36036631

30. Ott DE, Coren LV, Gagliardi TD. Redundant roles for nucleocapsid and matrix RNA-binding sequences in human immunodeficiency virus type 1 assembly. J Virol. 2005; 79:13839–13847. https://doi.org/10.1128/JVI.79.22.13839-13847.2005 PMID: 16254319

31. Ristic N, Chin MPS. Mutations in matrix and SP1 repair the packaging specificity of a Human Immunodeficiency Virus Type 1 mutant by reducing the association of Gag with spliced viral RNA. Retrovirology. 2010; 7:73. https://doi.org/10.1186/1742-4690-7-73 PMID: 20825656

32. Carlson L-A, Bai Y, Keane SC, Doudna JA, Hurley JH. Reconstitution of selective HIV-1 RNA packaging in vitro by membrane-bound Gag assemblies. Elife. 2016; 5:e14663. https://doi.org/10.7554/eLife.14663 PMID: 27343348

33. Guo X, Roy BB, Hu J, Roldan A, Wainberg MA, Liang C. The R362A mutation at the C-terminus of CA inhibits packaging of human immunodeficiency virus type 1 RNA. Virology. 2005; 343:190–200. https://doi.org/10.1016/j.virol.2005.08.031 PMID: 16183096

34. Russell RS, Roldan A, Detorio M, Hu J, Wainberg MA, Liang C. Effects of a Single Amino Acid Substitution within the p2 Region of Human Immunodeficiency Virus Type 1 on Packaging of Spliced Viral RNA. J Virol. 2003; 77:12986–12995. https://doi.org/10.1128/jvi.77.24.12986-12995.2003 PMID: 14645555

35. Aktar SJ, Jabeen A, Ali LM, Vivet-Boudou V, Marquet R, Rizvi TA. SHAPE analysis of the 5′ end of the Mason-Pfizer monkey virus (MPMV) genomic RNA reveals structural elements required for genome dimerization. RNA. 2013; 19:1648–1658. https://doi.org/10.1261/rna.040931.113 PMID: 24152551

36. Aktar SJ, Vivet-Boudou V, Ali LM, Jabeen A, Kalloush RM, Richer D, et al. Structural basis of genomic RNA (gRNA) dimerization and packaging determinants of mouse mammary tumor virus (MMTV). Retrovirology. 2014; 11:96. https://doi.org/10.1186/s12977-014-0096-6 PMID: 25394412

37. Clever J, Sassetti C, Parslow TG. RNA secondary structure and binding sites for gag gene products in the 5' packaging signal of human immunodeficiency virus type 1. J Virol. 1995; 69:2101–2109. https://doi.org/10.1128/JVI.69.4.2101-2109.1995 PMID: 7884856

38. D'Souza V, Summers MF. Structural basis for packaging the dimeric genome of Moloney murine leukaemia virus. Nature. 2004; 431:586. https://doi.org/10.1038/nature02944 PMID: 15457265

39. Jaballah SA, Aktar SJ, Ali J, Phillip PS, Al Dhaheri NS, Jabeen A, et al. A G–C-Rich Palindromic Structural Motif and a Stretch of Single-Stranded Purines Are Required for Optimal Packaging of Mason–Pfizer Monkey Virus (MPMV) Genomic RNA. J Mol Biol. 2010; 401:996–1014. https://doi.org/10.1016/j.jmb.2010.06.043 PMID: 20600114

40. Mustafa F, Lew KA, Schmidt RD, Browning MT, Rizvi TA. Mutational analysis of the predicted secondary RNA structure of the Mason-Pfizer monkey virus packaging signal. Virus Res. 2004; 99:35–46. https://doi.org/10.1016/j.virusres.2003.09.012 PMID: 14687944

41. Mustafa F, Vivet-Boudou V, Jabeen A, Ali LM, Kalloush RM, Marquet R, et al. The bifurcated stem loop 4 (SL4) is crucial for efficient packaging of mouse mammary tumor virus (MMTV) genomic RNA. RNA Biol. 2018; 15:1047–1059. https://doi.org/10.1080/15476286.2018.1486661 PMID: 29929424

42. Zeffman A, Hassard S, Varani G, Lever A. The major HIV-1 packaging signal is an extended bulged stem loop whose structure is altered on interaction with the Gag polyprotein. J Mol Biol. 2000; 297:877–893. https://doi.org/10.1006/jmbi.2000.3611 PMID: 10736224

43. Kenyon JC, Tanner SJ, Legiewicz M, Phillip PS, Rizvi TA, Le Grice SFJ, et al. SHAPE analysis of the FIV Leader RNA reveals a structural switch potentially controlling viral packaging and genome dimerization. Nucleic Acids Res. 2011; 39:6692–6704. https://doi.org/10.1093/nar/gkr252 PMID: 21546549

44. Bernacchi S, Abd El-Wahab EW, Dubois N, Hijnen M, Smyth RP, Mak J, et al. HIV-1 Pr55Gagbinds genomic and spliced RNAs with different affinity and stoichiometry. RNA Biol. 2017; 14:90–103. https://doi.org/10.1080/15476286.2016.1256533 PMID: 27841704

45. Smyth RP, Despons L, Huili G, Bernacchi S, Hijnen M, Mak J, et al. Mutational interference mapping experiment (MIME) for studying RNA structure and function. Nat Methods. 2015; 12:866–872. https://doi.org/10.1038/nmeth.3490 PMID: 26237229

46. Abd El-Wahab EW, Smyth RP, Mailler E, Bernacchi S, Vivet-Boudou V, Hijnen M, et al. Specific recognition of the HIV-1 genomic RNA by the Gag precursor. Nat Commun. 2014; 5:4304. https://doi.org/10.1038/ncomms5304 PMID: 24986025

47. Paillart J-C, Skripkin E, Ehresmann B, Ehresmann C, Marquet R. In Vitro Evidence for a Long Range Pseudoknot in the 5′-Untranslated and Matrix Coding Regions of HIV-1 Genomic RNA*. J Biol Chem. 2002; 277:5995–6004. https://doi.org/10.1074/jbc.M108972200 PMID: 11744696

48. Abbink TEM, Berkhout B. A novel long distance base-pairing interaction in human immunodeficiency virus type 1 RNA occludes the Gag start codon. J Biol Chem. 2003; 278:11601–11611. https://doi.org/10.1074/jbc.M210291200 PMID: 12458192

49. Siegfried NA, Busan S, Rice GM, Nelson JAE, Weeks KM. RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). Nat Methods. 2014; 11:959–965. https://doi.org/10.1038/nmeth.3029 PMID: 25028896

50. Keane SC, Heng X, Lu K, Kharytonchyk S, Ramakrishnan V, Carter G, et al. RNA structure. Structure of the HIV-1 RNA packaging signal. Science. 2015; 348:917–921. https://doi.org/10.1126/science.aaa9266 PMID: 25999508

51. Tran T, Liu Y, Marchant J, Monti S, Seu M, Zaki J, et al. Conserved determinants of lentiviral genome dimerization. Retrovirology. 2015; 12:83. https://doi.org/10.1186/s12977-015-0209-x PMID: 26420212

52. Kalloush RM, Vivet-Boudou V, Ali LM, Mustafa F, Marquet R, Rizvi TA. Packaging of Mason-Pfizer monkey virus (MPMV) genomic RNA depends upon conserved long-range interactions (LRIs) between U5 and gag sequences. RNA. 2016; 22:905–919. https://doi.org/10.1261/rna.055731.115 PMID: 27095024

53. Chameettachal A, Vivet-Boudou V, Pitchai FNN, Pillai VN, Ali LM, Krishnan A, et al. 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. 2021. https://doi.org/10.1093/nar/gkab223 PMID: 33836091

54. Krishnan A, Ali LM, Prabhu SG, Pillai VN, Chameettachal A, Vivet-Boudou V, et al. Identification of a putative Gag binding site critical for feline immunodeficiency virus genomic RNA packaging. RNA N Y N. 2023; 30:68–88. https://doi.org/10.1261/rna.079840.123 PMID: 37914398

55. Ye L, Gribling-Burrer A-S, Bohn P, Kibe A, Börtlein C, Ambi UB, et al. Short- and long-range interactions in the HIV-1 5' UTR regulate genome dimerization and packaging. Nat Struct Mol Biol. 2022; 29:306–319. https://doi.org/10.1038/s41594-022-00746-2 PMID: 35347312

56. Cardiff RD, Kenney N. Mouse mammary tumor biology: a short history. Adv Cancer Res. 2007; 98:53–116. https://doi.org/10.1016/S0065-230X(06)98003-8 PMID: 17433908

57. Medina D. Of mice and women: A short history of mouse mammary cancer research with an emphasis on the paradigms inspired by the transplantation method. Cold Spring Harb Perspect Biol. 2010; 2: a004523. https://doi.org/10.1101/cshperspect.a004523 PMID: 20630995

58. Dudley JP, Golovkina TV, Ross SR. Lessons Learned from Mouse Mammary Tumor Virus in Animal Models. ILAR J. 2016; 57:12–23. https://doi.org/10.1093/ilar/ilv044 PMID: 27034391

59. Coffin JM, Hughes SH, Varmus HE. Purification, Composition, and Morphology of Virions. Cold Spring Harbor Laboratory Press; 1997. Available from: https://www.ncbi.nlm.nih.gov/books/NBK19361/

60. Smith GH. Evidence for a precursor-product relationship between intracytoplasmic A particles and mouse mammary tumour virus cores. J Gen Virol. 1978; 41:193–200. https://doi.org/10.1099/0022-1317-41-1-193 PMID: 81268

61. Mustafa F, Amri DA, Ali FA, Sari NA, Suwaidi SA, Jayanth P, et al. Sequences within Both the 5' UTR and Gag Are Required for Optimal In Vivo Packaging and Propagation of Mouse Mammary Tumor Virus (MMTV) Genomic RNA. PLoS ONE. 2012; 7:e47088. https://doi.org/10.1371/journal.pone.0047088 PMID: 23077548

62. Swanson CM, Puffer BA, Ahmad KM, Doms RW, Malim MH. Retroviral mRNA nuclear export elements regulate protein function and virion assembly. EMBO J. 2004; 23:2632–2640. https://doi.org/10.1038/sj.emboj.7600270 PMID: 15201866

63. Stake MS, Bann DV, Kaddis RJ, Parent LJ. Nuclear trafficking of retroviral RNAs and Gag proteins during late steps of replication. Viruses. 2013; 5:2767–2795. https://doi.org/10.3390/v5112767 PMID: 24253283

64. Maldonado RJK, Rice B, Chen EC, Tuffy KM, Chiari EF, Fahrbach KM, et al. Visualizing Association of the Retroviral Gag Protein with Unspliced Viral RNA in the Nucleus. MBio. 2020; 11. https://doi.org/10.1128/mBio.00524-20 PMID: 32265329

65. Garbitt-Hirst R, Kenney SP, Parent LJ. Genetic evidence for a connection between Rous sarcoma virus gag nuclear trafficking and genomic RNA packaging. J Virol. 2009; 83:6790–6797. https://doi.org/10.1128/JVI.00101-09 PMID: 19369339

66. Gudleski N, Flanagan JM, Ryan EP, Bewley MC, Parent LJ. Directionality of nucleocytoplasmic transport of the retroviral gag protein depends on sequential binding of karyopherins and viral RNA. Proc Natl Acad Sci U S A. 2010; 107:9358–9363. https://doi.org/10.1073/pnas.1000304107 PMID: 20435918

67. Rizvi TA, Ali J, Phillip PS, Ghazawi A, Jayanth P, Mustafa F. Role of a heterologous retroviral transport element in the development of genetic complementation assay for mouse mammary tumor virus (MMTV) replication. Virology. 2009; 385:464–472. https://doi.org/10.1016/j.virol.2008.12.027 PMID: 19157480

68. Naldini L, Blömer U, Gallay P, Ory D, Mulligan R, Gage FH, et al. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. Science. 1996; 272:263–267. https://doi.org/10.1126/science.272.5259.263 PMID: 8602510

69. Tan W, Felber BK, Zolotukhin AS, Pavlakis GN, Schwartz S. Efficient expression of the human papillomavirus type 16 L1 protein in epithelial cells by using Rev and the Rev-responsive element of human immunodeficiency virus or the cis-acting transactivation element of simian retrovirus type 1. J Virol. 1995; 69:5607–5620. https://doi.org/10.1128/JVI.69.9.5607-5620.1995 PMID: 7637007

70. Moore MD, Hu WS. HIV-1 RNA dimerization: It takes two to tango. AIDS Rev. 2009; 11:91–102. PMID: 19529749

71. Nikolaitchik OA, Dilley KA, Fu W, Gorelick RJ, Tai S-HS, Soheilian F, et al. Dimeric RNA recognition regulates HIV-1 genome packaging. PLoS Pathog. 2013; 9:e1003249. https://doi.org/10.1371/journal.ppat.1003249 PMID: 23555259

72. McGinnis JL, Duncan CDS, Weeks KM. High-Throughput Analysis of RNA Structure and Ribonucleoprotein Assembly. Methods Enzymol. 2009; 468:67–89. https://doi.org/10.1016/S0076-6879(09)68004-6 PMID: 20946765

73. Low JT, Weeks KM. SHAPE-Directed RNA Secondary Structure Prediction. Methods San Diego Calif. 2010; 52:150–158. https://doi.org/10.1016/j.ymeth.2010.06.007 PMID: 20554050

74. Mailler E, Paillart J-C, Marquet R, Smyth RP, Vivet-Boudou V. The evolution of RNA structural probing methods: From gels to next-generation sequencing. WIREs RNA. 2019; 10:e1518. https://doi.org/10.1002/wrna.1518 PMID: 30485688

75. Weeks KM, Mauger DM. Exploring RNA Structural Codes with SHAPE Chemistry. Acc Chem Res. 2011; 44:1280–1291. https://doi.org/10.1021/ar200051h PMID: 21615079

76. Karabiber F, McGinnis JL, Favorov OV, Weeks KM. QuShape: rapid, accurate, and best-practices quantification of nucleic acid probing information, resolved by capillary electrophoresis. RNA. 2013; 19:63–73. https://doi.org/10.1261/rna.036327.112 PMID: 23188808

77. Reuter JS, Mathews DH. RNAstructure: software for RNA secondary structure prediction and analysis. BMC Bioinformatics. 2010; 11:129. https://doi.org/10.1186/1471-2105-11-129 PMID: 20230624

78. Boyd PS, Brown JB, Brown JD, Catazaro J, Chaudry I, Ding P, et al. NMR Studies of Retroviral Genome Packaging. Viruses. 2020; 12:1115. https://doi.org/10.3390/v12101115 PMID: 33008123

79. Chameettachal A, Pillai V, Ali L, Pitchai F, Ardah M, Mustafa F, et al. Biochemical and Functional Characterization of Mouse Mammary Tumor Virus Full-Length Pr77Gag Expressed in Prokaryotic and Eukaryotic Cells. Viruses. 2018; 10:334. https://doi.org/10.3390/v10060334 PMID: 29912170

80. Pitchai FNN, Chameettachal A, Vivet-Boudou V, Mohamed Ali L, Pillai VN, Krishnan A, et al. Identification of Pr78Gag binding sites on the Mason-Pfizer monkey virus genomic RNA packaging determinants. J Mol Biol. 2021;166923. https://doi.org/10.1016/j.jmb.2021.166923 PMID: 33713677

81. Gilmer O, Mailler E, Paillart J-C, Mouhand A, Tisné C, Mak J, et al. Structural maturation of the HIV-1 RNA 5' untranslated region by Pr55Gag and its maturation products. RNA Biol. 2022; 19:191–205. https://doi.org/10.1080/15476286.2021.2021677 PMID: 35067194

82. Nicholson BL, White KA. 3' Cap-independent translation enhancers of positive-strand RNA plant viruses. Curr Opin Virol. 2011; 1:373–380. https://doi.org/10.1016/j.coviro.2011.10.002 PMID: 22440838

83. Fricke M, Marz M. Prediction of conserved long-range RNA-RNA interactions in full viral genomes. Bioinforma Oxf Engl. 2016; 32:2928–2935. https://doi.org/10.1093/bioinformatics/btw323 PMID: 27288498

84. Paillart J-C, Dettenhofer M, Yu X-F, Ehresmann C, Ehresmann B, Marquet R. First snapshots of the HIV-1 RNA structure in infected cells and in virions. J Biol Chem. 2004; 279:48397–48403. https://doi.org/10.1074/jbc.M408294200 PMID: 15355993

85. Sardo L, Hatch SC, Chen J, Nikolaitchik O, Burdick RC, Chen D, et al. Dynamics of HIV-1 RNA Near the Plasma Membrane during Virus Assembly. J Virol. 2015; 89:10832–10840. https://doi.org/10.1128/JVI.01146-15 PMID: 26292321

86. Pak AJ, Grime JMA, Sengupta P, Chen AK, Durumeric AEP, Srivastava A, et al. Immature HIV-1 lattice assembly dynamics are regulated by scaffolding from nucleic acid and the plasma membrane. Proc Natl Acad Sci U S A. 2017; 114:E10056–E10065. https://doi.org/10.1073/pnas.1706600114 PMID: 29114055

87. Duchon A, Santos S, Chen J, Brown M, Nikolaitchik OA, Tai S, et al. Plasma Membrane Anchoring and Gag:Gag Multimerization on Viral RNA Are Critical Properties of HIV-1 Gag Required To Mediate Efficient Genome Packaging. MBio. 2021; 12:e0325421. https://doi.org/10.1128/mbio.03254-21 PMID: 34872357

88. Ying Y, Yang Y, Chen AK. Roles of RNA scaffolding in nanoscale Gag multimerization and selective protein sorting at HIV membranes. Sci Adv. 2024; 10:eadk8297. https://doi.org/10.1126/sciadv.adk8297 PMID: 38394201

89. Webb JA, Jones CP, Parent LJ, Rouzina I, Musier-Forsyth K. Distinct binding interactions of HIV-1 Gag to Psi and non-Psi RNAs: Implications for viral genomic RNA packaging. RNA. 2013; 19:1078–1088. https://doi.org/10.1261/rna.038869.113 PMID: 23798665

90. Comas-Garcia M, Datta SA, Baker L, Varma R, Gudla PR, Rein A. Dissection of specific binding of HIV-1 Gag to the "packaging signal" in viral RNA. Goff SP, editor. Elife. 2017; 6: e27055. https://doi.org/10.7554/eLife.27055 PMID: 28726630

91. Comas-Garcia M, Kroupa T, Datta SA, Harvin DP, Hu W-S, Rein A. Efficient support of virus-like particle assembly by the HIV-1 packaging signal. Elife. 2018; 7:e38438. https://doi.org/10.7554/eLife.38438 PMID: 30070634

92. Kroupa T, Datta SAK, Rein A. Distinct Contributions of Different Domains within the HIV-1 Gag Polyprotein to Specific and Nonspecific Interactions with RNA. Viruses. 2020; 12:394. https://doi.org/10.3390/v12040394 PMID: 32252233

93. Houzet L, Paillart JC, Smagulova F, Maurel S, Morichaud Z, Marquet R, et al. HIV controls the selective packaging of genomic, spliced viral and cellular RNAs into virions through different mechanisms. Nucleic Acids Res. 2007; 35:2695–2704. https://doi.org/10.1093/nar/gkm153 PMID: 17426127

94. Wilkinson KA, Gorelick RJ, Vasa SM, Guex N, Rein A, Mathews DH, et al. High-throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states. PLoS Biol. 2008; 6:e96. https://doi.org/10.1371/journal.pbio.0060096 PMID: 18447581

95. Nikolaitchik OA, Somoulay X, Rawson JMO, Yoo JA, Pathak VK, Hu W-S. Unpaired Guanosines in the 5′ Untranslated Region of HIV-1 RNA Act Synergistically to Mediate Genome Packaging. J Virol. 2020 [cited 2020 Sep 14]. https://doi.org/10.1128/JVI.00439-20 PMID: 32796062

96. Ali LM, Pitchai FNN, Vivet-Boudou V, Chameettachal A, Jabeen A, Pillai VN, et al. Role of Purine-Rich Regions in Mason-Pfizer Monkey Virus (MPMV) Genomic RNA Packaging and Propagation. Front Microbiol. 2020;11. https://doi.org/10.3389/fmicb.2020.595410 PMID: 33250884

97. Pillai VN, Ali LM, Prabhu SG, Krishnan A, Chameettachal A, Pitchai FNN, et al. 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. 2021; 433:167293. https://doi.org/10.1016/j.jmb.2021.167293 PMID: 34624298

98. Umunnakwe CN, Duchon A, Nikolaitchik OA, Rahman SA, Liu Y, Chen J, et al. Specific Guanosines in the HIV-2 Leader RNA are Essential for Efficient Viral Genome Packaging. J Mol Biol. 2021; 433:166718. https://doi.org/10.1016/j.jmb.2020.11.017 PMID: 33221337

99. Kenyon JC, Prestwood LJ, Lever AML. A novel combined RNA-protein interaction analysis distinguishes HIV-1 Gag protein binding sites from structural change in the viral RNA leader. Sci Rep. 2015; 5:14369. https://doi.org/10.1038/srep14369 PMID: 26449409

100. Muriaux D, Mirro J, Harvin D, Rein A. RNA is a structural element in retrovirus particles. Proc Natl Acad Sci U S A. 2001; 98:5246–5251. https://doi.org/10.1073/pnas.091000398 PMID: 11320254

101. Mattei S, Flemming A, Anders-Össwein M, Kräusslich H-G, Briggs JAG, Müller B. RNA and Nucleocapsid Are Dispensable for Mature HIV-1 Capsid Assembly. J Virol. 2015; 89:9739–9747. https://doi.org/10.1128/JVI.00750-15 PMID: 26178992

102. Mizrahi I, Bruinsma R, Rudnick J. Packaging contests between viral RNA molecules and kinetic selectivity. PLoS Comput Biol. 2022; 18:e1009913. https://doi.org/10.1371/journal.pcbi.1009913 PMID: 35363785

103. Shackleford GM, Varmus HE. Construction of a clonable, infectious, and tumorigenic mouse mammary tumor virus provirus and a derivative genetic vector. Proc Natl Acad Sci U S A. 1988; 85:9655–9659. https://doi.org/10.1073/pnas.85.24.9655 PMID: 2849114

104. Marsh JL, Erfle M, Wykes EJ. The pIC plasmid and phage vectors with versatile cloning sites for recombinant selection by insertional inactivation. Gene. 1984; 32:481–485. https://doi.org/10.1016/0378-1119(84)90022-2 PMID: 6099327

105. Pitchai FNN, Ali L, Pillai VN, Chameettachal A, Ashraf SS, Mustafa F, et al. Expression, purification, and characterization of biologically active full-length Mason-Pfizer monkey virus (MPMV) Pr78 Gag. Sci Rep. 2018; 8:11793. https://doi.org/10.1038/s41598-018-30142-0 PMID: 30087395

106. Krishnan A, Pillai VN, Chameettachal A, Mohamed Ali L, Nuzra Nagoor Pitchai F, Tariq S, et al. Purification and Functional Characterization of a Biologically Active Full-Length Feline Immunodeficiency Virus (FIV) Pr50Gag. Viruses. 2019; 11. https://doi.org/10.3390/v11080689 PMID: 31357656

107. Akhlaq S, Panicker NG, Philip PS, Ali LM, Dudley JP, Rizvi TA, et al. A cis-Acting Element Downstream of the Mouse Mammary Tumor Virus Major Splice Donor Critical for RNA Elongation and Stability. J Mol Biol. 2018; 430:4307–4324. https://doi.org/10.1016/j.jmb.2018.08.025 PMID: 30179605

108. Mallona I, Weiss J, Egea-Cortines M. pcrEfficiency: a Web tool for PCR amplification efficiency prediction. BMC Bioinformatics. 2011; 12:404. https://doi.org/10.1186/1471-2105-12-404 PMID: 22014212

109. Marquet R, Baudin F, Gabus C, Darlix JL, Mougel M, Ehresmann C, et al. Dimerization of human immunodeficiency virus (type 1) RNA: stimulation by cations and possible mechanism. Nucleic Acids Res. 1991; 19:2349–2357. https://doi.org/10.1093/nar/19.9.2349 PMID: 1645868

110. Paillart JC, Marquet R, Skripkin E, Ehresmann B, Ehresmann C. Mutational analysis of the bipartite dimer linkage structure of human immunodeficiency virus type 1 genomic RNA. J Biol Chem. 1994; 269:27486–27493. PMID: 7961663

111. Paillart JC, Skripkin E, Ehresmann B, Ehresmann C, Marquet R. A loop-loop "kissing" complex is the essential part of the dimer linkage of genomic HIV-1 RNA. Proc Natl Acad Sci U S A. 1996; 93:5572–5577. https://doi.org/10.1073/pnas.93.11.5572 PMID: 8643617

112. Pillai VN, Ali LM, Prabhu SG, Krishnan A, Tariq S, Mustafa F, et al. Expression, purification, and functional characterization of soluble recombinant full-length simian immunodeficiency virus (SIV) Pr55Gag. Heliyon. 2023; 9:e12892. https://doi.org/10.1016/j.heliyon.2023.e12892 PMID: 36685375

113. Kalloush RM, Vivet-Boudou V, Ali LM, Pillai VN, Mustafa F, Marquet R, et al. Stabilizing role of structural elements within the 5´ Untranslated Region (UTR) and gag sequences in Mason-Pfizer monkey virus (MPMV) genomic RNA packaging. RNA Biol. 2019; 16:612–625. https://doi.org/10.1080/15476286.2019.1572424 PMID: 30773097

114. Darty K, Denise A, Ponty Y. VARNA: Interactive drawing and editing of the RNA secondary structure. Bioinformatics. 2009; 25:1974–1975. https://doi.org/10.1093/bioinformatics/btp250 PMID: 19398448