


RESEARCH PAPER



The bifurcated stem loop 4 (SL4) is crucial for efficient packaging of mouse mammary tumor virus (MMTV) genomic RNA

Farah Mustafa^{a*}, Valérie Vivet-Boudou ^{b*}, Ayesha Jabeen^c, Lizna M. Ali^c, Rawan M. Kalloush^c, Roland Marquet^b, and Tahir A. Rizvi^c

^aDepartment of Biochemistry, College of Medicine and Health Sciences, United Arab Emirates University, Al Ain, UAE; ^bUniversité de Strasbourg, CNRS, Architecture et Réactivité de l'ARN, Strasbourg, France; ^cDepartment of Microbiology & Immunology, College of Medicine and Health Sciences, United Arab Emirates University, Al Ain, UAE

ABSTRACT

Packaging the mouse mammary tumor virus (MMTV) genomic RNA (gRNA) requires the entire 5' untranslated region (UTR) in conjunction with the first 120 nucleotides of the *gag* gene. This region includes several palindromic (pal) sequence(s) and stable stem loops (SLs). Among these, stem loop 4 (SL4) adopts a bifurcated structure consisting of three stems, two apical loops, and an internal loop. Pal II, located in one of the apical loops, mediates gRNA dimerization, a process intricately linked to packaging. We thus hypothesized that the bifurcated SL4 structure could constitute the major gRNA packaging determinant. To test this hypothesis, the two apical loops and the flanking sequences forming the bifurcated SL4 were individually mutated. These mutations all had deleterious effects on gRNA packaging and propagation. Next, single and compensatory mutants were designed to destabilize then recreate the bifurcated SL4 structure. A structure-function analysis using bioinformatics predictions and RNA chemical probing revealed that mutations that led to the loss of the SL4 bifurcated structure abrogated RNA packaging and propagation, while compensatory mutations that recreated the native SL4 structure restored RNA packaging and propagation to wild type levels. Altogether, our results demonstrate that SL4 constitutes the principal packaging determinant of MMTV gRNA. Our findings further suggest that SL4 acts as a structural switch that can not only differentiate between RNA for translation *versus* packaging/dimerization, but its location also allows differentiation between spliced and unspliced RNAs during gRNA encapsidation.

ARTICLE HISTORY

Received 5 April 2018
Accepted 4 June 2018

KEYWORDS

Retroviruses; Mouse mammary tumor virus; RNA secondary structure; RNA packaging and dimerization; palindrome


1. Introduction

The mouse mammary tumor virus (MMTV) is a rodent *betaretrovirus* that causes breast cancer in mice [1]. Although studied extensively, little is known about the molecular mechanism of genomic RNA (gRNA) packaging of this oncogenic retrovirus that has been proposed to be reclassified as a complex retrovirus [2,3]. In all retroviruses (irrespective of their simple or complex nature), sequences that are essential for gRNA packaging, termed 'packaging determinants and/or psi (Ψ)', reside within the 5' end of the viral genome, and more precisely within the 5' untranslated region (5'UTR) and the 5' end of the *gag* gene. Accordingly, an early study suggested that MMTV harbors sequences responsible for gRNA packaging at the 5' end of its genome [4]. Packaging signals in other retroviruses have been shown to fold into higher order structures comprising of various structural motifs [5–13]. For example, a well-defined 5' region of the murine leukemia virus (MLV) and some other retroviruses has been found to be required as well as sufficient for gRNA packaging, since inclusion of

these sequences in heterologous non-retroviral RNAs facilitates the packaging of these 'foreign' RNAs into viral particles [14–17]. Furthermore, it is the presence of these structural motifs that explains the phenomena of co- and cross-packaging among retroviruses that have no relationship to each other [11]. For example, it has been long observed that specificity of packaging can be exchanged in many retroviruses by the substitution of packaging signals that have no sequence homology [16–22]. Therefore, retroviral gRNA packaging process must involve recognition of packaging sequences at the secondary/tertiary structure level(s) rather than only at the sequence level. However, the precise nature and significance of these structural motifs during the packaging process remains largely unclear. The case of MMTV is not that different. We have recently shown that the entire 5' UTR and the first 120 nt of *gag* are required for efficient MMTV gRNA packaging and propagation [23,24]. Furthermore, these sequences were predicted to fold into higher order structures comprising of several structural motifs, which were validated by SHAPE (selective 2'-hydroxyl acylation analyzed by primer extension [25]).

CONTACT Tahir A. Rizvi  tarizvi@uae.ac.ae  Department of Microbiology & Immunology, College of Medicine and Health Sciences (CMHS), United Arab Emirates University (UAEU), P.O. Box 17666, Al Ain, UAE; Roland Marquet  r.marquet@ibmc-cnrs.unistra.fr  Université de Strasbourg, CNRS, Architecture et Réactivité de l'ARN, UPR 9002, IBMC, 15 rue René Descartes, 67084 Strasbourg cedex, France

*These authors contributed equally to this work.

 Supplemental data for this article can be accessed [here](#).

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

A distinguishing feature of the SHAPE-validated structure of the MMTV packaging signal RNA is the presence of a bifurcated stem loop 4 (SL4) containing a palindromic sequence (pal II; 5' CUGCAG 3') and a 9 nt stretch (5'GGAGAAGAG 3') of single-stranded purines (ssPurines) in adjacent apical loops [25]. Both pal II and ssPurines are phylogenetically conserved in eight different strains of MMTV at the sequence as well as the secondary structural levels [25]. Pal II (5' CUGCAG 3') has been shown to act as the dimerization initiation site (DIS) for MMTV gRNA from amongst four other palindromes within the 5' UTR of the viral genome [25]. The presence of a stretch of purines in the packaging sequences on retroviral gRNA has been proposed to facilitate RNA packaging by functioning as a potential NC binding site [26–31]. It is also interesting to note that the positioning of the purine-rich loop in MMTV adjacent to the pal II loop is reminiscent of the situation found in Mason-Pfizer monkey virus (MPMV) [32,33]. In the case of MPMV, genetic and structure-prediction analyses have suggested that either ssPurines or its partially-repeated base-paired sequence in an adjacent region plays a crucial role in gRNA packaging, possibly by functioning as a NC binding site [32,33]. Therefore, it is reasonable to propose that the bifurcated SL4 structure in MMTV functions as a major dimerization and packaging determinant during MMTV gRNA packaging process with pal II initiating the process of dimerization and ssPurines facilitating the encapsidation of the dimerized RNA into the virus particles by functioning as a potential Gag binding site.

The role of the sequences as well as the higher order structure of the MMTV SL4 region have not been tested genetically, although they offer a logical and perhaps mechanistic explanation for their central role in MMTV gRNA dimerization and packaging [24,25]. Therefore, to establish the biological significance of the sequences and structural components of SL4 and to further provide functional evidence for the role of pal II and ssPurines in MMTV RNA packaging, a series of mutations were introduced that included deletions, substitutions, and compensatory mutations in SL4. These mutations were tested employing a biologically relevant MMTV-based three plasmid *trans* complementation assay [23] to determine their direct effects on the packaging and propagation of MMTV transfer vector RNAs, and finally structure-function relationship of these mutations was established using mFold predictions and SHAPE experiments. Our results show that the bifurcated nature of SL4 structure as well as the sequence of its two apical loops are required for MMTV gRNA packaging and propagation.

2. Results

2.1. Experimental design

Mutations in SL4 were introduced into a sub-genomic MMTV transfer vector, DA024, that expresses the wild type or mutant RNAs. The potential of these RNAs to be packaged by virus particles was tested by using a 3-plasmid *trans* complementation assay developed earlier to study MMTV packaging and propagation [23–25]. In this assay, infectious, pseudotyped, virus particles were created by expression of the MMTV *gag*/

pol and vesicular stomatitis virus (VSV-G) *env* genes from two independent plasmids, while the transfer vector described above provided the wild type or mutant substrates for RNA packaging (Supplemental Fig. 1).

Nuclear export of the RNA substrates was monitored by analysis of the relative amounts of transfer vector RNA in the cytoplasmic fractions using a custom-designed qPCR assay [24,25]. In parallel, the virus particles isolated from the transfected cultures were used to: i) measure the gRNA content in the virus particles by RT-qPCR, and ii) infect target cell line HeLaT4, resulting in the transduction of these cells with the marker *hygromycin resistance* gene (propagation of the packaged RNA), as described previously [23–25]. The number of hygromycin resistant (Hyg^r) colonies obtained should be directly proportional to the amount of RNA packaged into the virus particles, providing an indirect estimate of RNA packaging, given that packaged RNA is successfully reverse transcribed and integrated. Thus, this *in vivo* packaging and propagation assay allowed us to quantify the effects of the mutations in the SL4 region on RNA packaging and propagation without any ambiguity since the defective nature of the virions produced limited the assay to a single round of replication [23–25].

2.2. The region between the mSD and gag AUG is important for MMTV RNA packaging and vector RNA propagation

To determine the biological importance of sequences encompassing SL4 or its structure in MMTV genomic RNA packaging, we initiated our investigations with a set of broad deletion and substitution mutations in the region between the MMTV major splice donor (mSD) and the Gag AUG (Supplemental Fig. 2A). Being downstream of the mSD, this region should be present exclusively in the unspliced gRNA which is destined to be packaged into nascent virions and yet be part of SL4, as highlighted in the hSHAPE-validated structure of the MMTV 5' mRNA (Supplemental Fig. 2A and B).

RNAs were isolated from both the cytoplasmic fraction of co-transfected 293T cells and concentrated virus particles to determine cytoplasmic RNA export and relative RNA packaging efficiency (RPE) into the virus particles. After ensuring that the RNA samples were devoid of contaminating DNA (Supplemental Fig. 2C), cDNAs were prepared and tested by PCR to ensure that the cytoplasmic fractions were devoid unspliced β -actin mRNA, an mRNA that should not be transported to the cytoplasm (Supplemental Fig. 2C). This was followed by estimation of the RPE of the transfer vector RNA using the previously-validated custom-made RT-qPCR assay [24,25]. The RPE of the vector RNA was calculated by normalizing the expression of the packaged vector RNA to their cytoplasmic RNA expression (the cytoplasmic RNA expression was further normalized to the transfection efficiency), and the results were plotted relative to the packaging of the wild type (WT) vector, DA024 (brown bars in Supplemental Fig. 2D). As can be seen, none of the mutant vector RNAs could be effectively packaged into MMTV virions despite the fact that most of the mutant vector RNAs were expressed within 2–2.5-folds of the wild type levels in the cytoplasm (blue bars in Supplemental Fig. 2D). The severe

defect in vector RNA packaging was further confirmed by a corresponding defect in vector RNA propagation, as assessed by a complete lack of hygromycin-resistant colonies (CFU/ml; green bars in Supplemental Fig. 2D), further corroborating the inability of the mutant vector RNAs to be packaged into MMTV virus particles. The RNA packaging and propagation results reported here (Supplemental Fig. 2D) as well as in subsequent experiments were obtained by pooling the results from multiple independent experiments in which each sample was quantitated in triplicates. These data indicate that the region between the mSD and Gag AUG is critical for vector RNA packaging and propagation.

2.3. The sequence of the two apical loops of the bifurcated SL4 structure is critical for MMTV RNA packaging and propagation

To further interrogate the structure of SL4, another set of mutations targeting sub-regions of SL4 was introduced into DA024 (WT) to determine whether it is a particular sequence within SL4

or the whole SL4 structure that is critical for MMTV gRNA packaging. These mutations were designed to either destabilize the SL4 structure or change the sequence within different motifs of SL4 without affecting structure (Figure 1(a,b)). SL4-09 and SL4-010 were designed to destabilize the structure of SL4 by introducing a 3-nucleotide deletion in stem I or the internal loop joining stems I and II together (Figure 1(b)). SL4-11 was designed to substitute three nucleotides in the 5' part of the internal loop (AUA to UCU) and thus maintain the size of the loop. This substitution should further test whether this region could possibly serve as a binding site for Gag (purines are potential NC binding sites) or another protein important for RNA packaging. SL4-12 and SL4-13 were designed to delete either the pal II or ssPurines apical loops, respectively, while in SL4-14 ssPurines were substituted by pyrimidines (Figure 1(a,b)). These mutants were tested along with SA042P in which the whole pal II stem loop was deleted [25]. This mutant has previously been shown to be defective for MMTV vector RNA packaging and propagation [25] and was used as a reference to determine the effect of the other mutations introduced in this study.

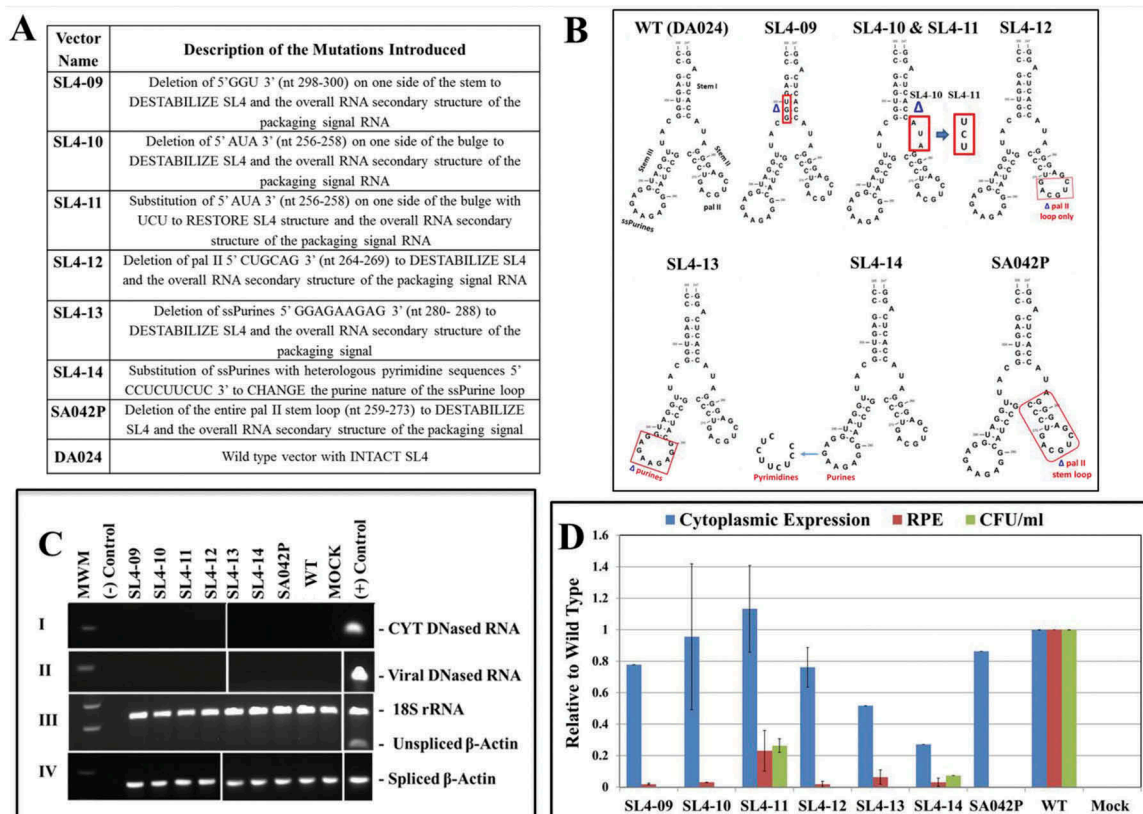


Figure 1. Sequence and structure-specific deletion and substitution mutations within SL4 and their effect on MMTV RNA packaging and propagation. A) Names and description of the mutations tested in the sub-genomic MMTV transfer vector, DA024, context. The expected effect of the mutation on SL4 structure is described in the table. B) Location of the deletion (Δ) or substitution mutation made displayed on the structure of SL4. Boxed regions show the precise location of the nucleotides changed. SA042P is a deletion mutant (deleted region is highlighted by the box) that has been tested previously [25]. C) PCR amplifications of the DNase-treated cytoplasmic (panel I) and viral (panel II) RNAs using virus-specific primers. Panel III shows amplification conducted on the cDNAs obtained from cytoplasmic RNAs using primers that amplify unspliced β -actin mRNA. Multiplex amplifications were conducted in the presence of primers/competitor for 18S ribosomal RNA. Panel IV shows RT-PCR of cytoplasmic cDNA using primers that amplify spliced β -actin mRNA. Owing to the large number of samples, gels have been spliced together as indicated by vertical white spaces in different panels to create part C for clarity. D) Cytoplasmic expression and relative packaging efficiency (RPE) of transfer vector RNAs, and hygromycin-resistance (Hyg^r) colony forming unit per ml (CFU/ml) for mutant transfer vector RNAs relative to the wild type DA024 construct. CFU/ml expressed for each mutant was normalized to the luciferase expression observed in the transfected cultures. The histograms represent data from multiple independent experiments (\pm SD). The RPE for each mutant was determined by dividing the packaged mutant viral RNA values by the luciferase-normalized cytoplasmic expression of the respective mutant (to take into account the varying transfection efficiencies). Finally, the results of RNA packaging of all the mutants were presented relative to the wild type (WT, DA024) construct.

Mutants SL4-09 to SL4-14 and SA042P along with wild type DA024 were tested in the *in vivo* packaging and propagation assay. Figure 1(c) illustrates the appropriate controls for these mutants (DNase treatment of RNA samples, nucleocytoplasmic RNA fractionation control, and presence of amplifiable cDNA), imperative for establishing the packaging and propagation potential of these mutant RNAs. Test of their packaging potential indicated that mutants SL4-09 and SL4-10 with the 3-nucleotide deletions in either stem I or the internal loop were severely defective for RNA packaging (a decrease of 30–50 folds compared to the wild type (DA024; P value <0.001)). Restoring the internal loop with a heterologous 3-nucleotide substitution (AUA to UCU) in SL4-11 resulted in partial restoration of packaging efficiency to about 20% (a 5-fold reduction compared to wild type, a difference which was statistically significant (Figure 1(d); P value <0.019). However, deletion of the pal II apical loop in SL4-12 or ssPurines apical loop in SL4-13, severely abrogated RNA packaging (by 50-folds in SL4-12, P value <0.001 and ~17-folds in SL4-13; P value <0.001). Changing the nature of ssPurines to pyrimidines in SL4-14 also abrogated RNA packaging, and RPE was decreased by >30-folds; P value <0.001 compared to the wild type (DA024). SA042P

continued to show the severe packaging defect observed earlier [25], which in fact was more severe than any of the mutants tested in this series. The propagation efficiencies of the mutant RNAs varied, but in concordance with their respective RPEs (Figure 1 (d)), confirming the RT-qPCR results. Thus, SL4-09, SL4-10, SL4-12, SL4-13, and SL4-14 were defective for vector RNA propagation, while SL4-11 with a mutation designed to maintain the size of the internal loop while changing its sequence, was partially packaged (~20%) and showed a 4-fold reduction in vector RNA propagation compared to wild type (P value <0.026; Figure 2(d)). Together, these results reveal that not only the structure of SL4 is important for RNA packaging and propagation, but specific sequences such as ssPurines, pal II, and to a lesser extent the internal loop within this structure, are also important for the packaging and propagation processes during MMTV life cycle.

2.4. Stem I of the bifurcated SL4 structure is important for MMTV RNA packaging and propagation

In the next series of mutants, three sets of paired mutants were tested in which the first clone contained mutations designed to

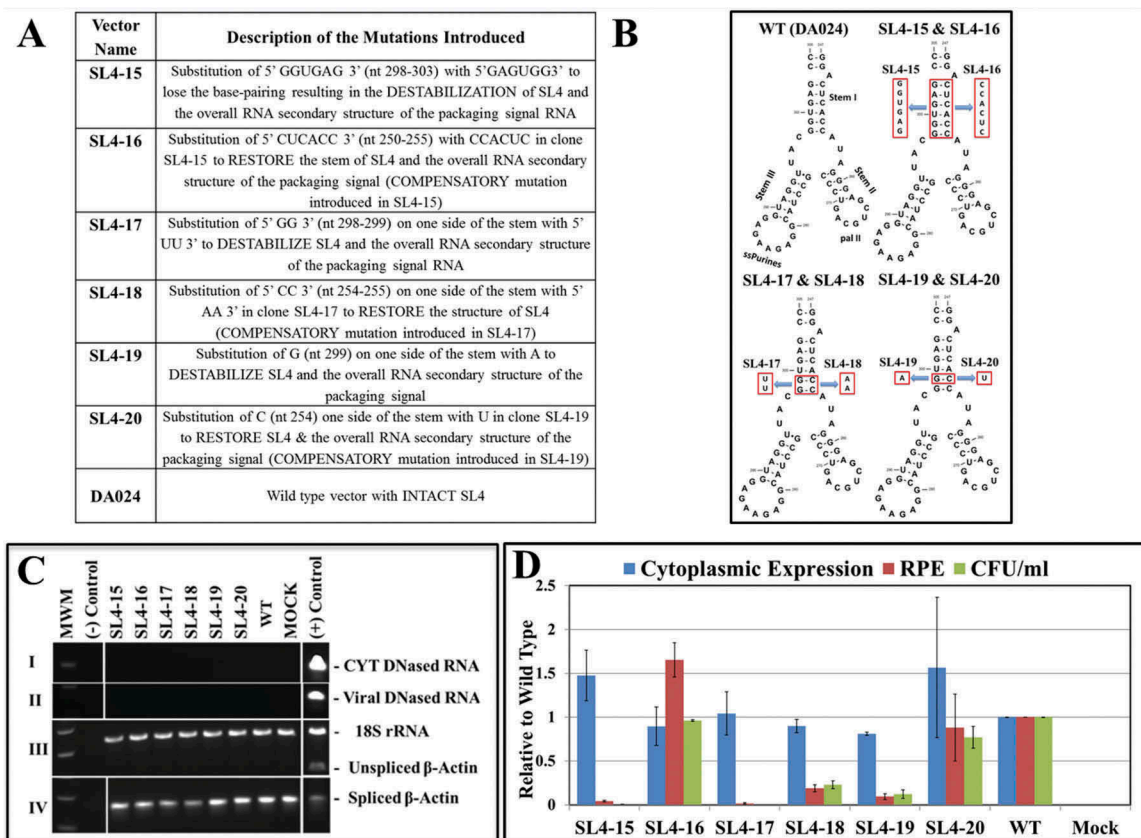


Figure 2. Stem I destabilizing deletion and corresponding compensatory substitution mutations and their effects on MMTV RNA packaging and propagation. **A)** Names and description of the mutations tested in the sub-genomic MMTV transfer vector, DA024, context. The expected effect of the mutation on SL4 structure is described in the table. **B)** Location of the deletion or substitution mutation made displayed on the structure of SL4. Boxed regions show the precise location of the nucleotides changed. **C)** PCR amplifications of the DNase-treated cytoplasmic (panel I) and viral (panel II) RNAs using virus-specific primers. Panel III shows amplification conducted on the cDNAs obtained from cytoplasmic RNAs using primers that amplify unspliced β -actin mRNA. Multiplex amplifications were conducted in the presence of primers/competimer for 18S ribosomal RNA. Panel IV shows RT-PCR of cytoplasmic cDNA using primers that amplify spliced β -actin mRNA. Owing to the large number of samples, gels have been spliced together as indicated by vertical white spaces in different panels to create part C for clarity. **D)** Cytoplasmic expression and relative packaging efficiency (RPE) of transfer vector RNAs, and hygromycin-resistance (Hyg^r) colony forming unit per ml (CFU/ml) for mutant transfer vector RNAs relative to the wild type DA024 construct. CFU/ml expressed for each mutant was normalized to the luciferase expression observed in the transfected cultures. The histograms represent data from multiple independent experiments (\pm SD). The RPE for each mutant was determined by dividing the packaged mutant viral RNA values by the luciferase-normalized cytoplasmic expression of the respective mutant (to take into account the varying transfection efficiencies). Finally, the results of RNA packaging of all the mutants were presented relative to the wild type (WT, DA024) construct.

destabilize stem I of SL4, while the corresponding paired mutant harbored an additional compensatory mutation expected to recreate the stem, but using heterologous sequences. Thus, in SL4-15, six of the 9-nucleotides of stem I (nucleotides 298–303) were substituted by flipping the sequence. The compensatory mutation was performed on the opposite side of stem I (nucleotides 250–255), generating SL4-16 (Figure 2(a,b)). In SL4-17, the lower part of stem I was targeted and the two Gs (298–299) were substituted with Us, while in SL4-18, the complementary Cs (254–255) were converted to As. These two GC pairs seemed to hold the whole SL4 structure together and our hypothesis was that lowering their free energy may result in destabilization of the whole structure despite the complementary nature of the compensatory mutations. In the third mutant pair, this relationship was investigated further by mutating only G299 to A in SL4-19, followed by a compensatory U254 substitution on the other side of the stem in SL4-20.

The relevant controls for this set of mutations are depicted in Figure 2(c). As expected, test of the packaging potential of SL4-14 and SL4-15 indicated that packaging was essentially abrogated by the substitution of 6 nucleotides in one arm of stem I in SL4-15 (Figure 2(d)). Moreover, introduction of the compensatory mutations on the opposite arm of the stem in SL4-16 restored the packaging to in fact better than the wild type levels (Figure 2(d)). When the two GGs (298–299) in stem I were mutated in SL4-17, they also resulted in a complete abrogation of RNA packaging. Interestingly, as predicted, conversion of the two G-C pairs to A-U pairs in mutant SL4-18 was unable to restore packaging to the wild type levels, though partial restoration (~20%) was observed (a 5-fold reduction compared to wild type; P value < 0.004; Figure 2(d)). Finally, when only G299 was substituted with A to destabilize the stem, packaging was disrupted in SL4-19, and restoration of the base-pairing with a compensatory mutation (U254) in SL4-20 restored packaging to the wild type levels (Figure 2(d)). In addition, propagation data closely paralleled packaging data for all mutants. These results show that despite the presence of important *cis*-acting sequences in SL4, such as pal II and ssPurines, the overall structure of SL4 was as important for efficient RNA packaging and viral propagation.

2.5. Structure-function analysis

To determine the effect of the introduced mutations on the secondary structure of the MMTV packaging signal RNA, the relevant sequences of the mutant vector RNAs starting from R to 120 nts within *gag* were folded using the MFold program that uses energy minimization to predict the energetically most stable structures [34,35]. Comparison of the mutant RNA structures SL4-01 to SL4-08 predicted by MFold with that of the wild type DA024 RNA revealed that in most cases, the structure of SL4 with its bifurcated stem loop was disrupted except in SL4-06 (Supplemental Fig. 3). In these cases, the ssPurines were base-paired with other sequences, while the pal II stem loop was largely intact except in the case of SL4-07 where both ssPurines and pal II sequences were base-paired with other RNA regions (Supplemental Figure 3). The disruption of SL4 structure in these cases correlated well with

the abrogation of packaging observed in these mutants (Supplemental Figs. 2D and 3). The only exception was SL4-06, which was as defective in packaging as the other mutants even though the native structure of SL4 was predicted to be maintained (Supplemental Figs. 2D and 3). In order to address this inconsistency, the 2D structure of mutant SL4-06 RNA was studied by hSHAPE. Our experimental data showed that the substitutions introduced in this mutant prevented the formation of the wild-type bifurcated structure of SL4 (Figure 3). The native stem-loops II and III were observed, but an additional one was also present, and moreover, the internal loop was larger than in wild type DA024 RNA structure (Figure 3). Our experimental structural probing thus confirmed that disruption of the wild type bifurcated SL4 structure abolished RNA packaging, even if the pal II and ssPurines sequences were present and exposed in apical loops.

2.6. The two apical loops within SL4 are important *cis*-acting single-stranded RNA elements in the RNA packaging process

Functional analysis of the second set of mutants revealed that deletion of the three nucleotides (298–300) on the 5' side of stem I in SL4-09 or 3-nucleotides (256–258) in the internal loop in SL4-10 both resulted in a complete loss of packaging (Figure 1). For mutant SL4-09, this phenotype was in agreement with the complete loss of the SL4 structure predicted by mFold in which both the ssPurines and pal II stem loops were found to be base-paired with other sequences (Figures 1(d) and 4). However, for mutant SL4-10, the hSHAPE experiments were in disagreement with the mFold prediction. While pal II and ssPurine sequences were predicted to be base-paired in this RNA, our hSHAPE data indicated that the two apical loops and the remaining half internal loop were highly reactive, indicative of a bifurcated structure with a shortened internal loop (Figure 3). Moreover, this structure with pal II and ssPurine in apical loops is in keeping with our gRNA dimerization data, which showed SL4-10 RNA dimerized as efficiently as the wild type (DA024) (Supplemental Fig. 4). We would expect that this RNA would be unable to dimerize if the pal II sequence was base-paired as in the structure predicted by mFold. Thus, our SHAPE data indicate that the size of the SL4 internal loop or/and the geometry of this junction are important for biological activity, perhaps acting as a binding site for viral or cellular proteins necessary for RNA encapsidation.

In mutant SL4-11, substitution of three nucleotides (AUA) of the internal loop with heterologous nucleotides (UCU) was tested. The mFold structure prediction was confirmed by hSHAPE experiments and indicated that the substitution did not affect the structure, which was undistinguishable from the wild type structure (Figures 3 and 4); however, packaging efficiency of this mutant decreased by 75% compared to the wild type (DA024) but not dimerization (Supplementary Fig. 4), suggesting that the sequence of the internal loop is important but not absolutely necessary, for either RNA-RNA or RNA-protein interactions and consequently for packaging.

Deletion of the apical loop of the pal II stem loop in SL4-12 resulted in a complete loss of packaging which

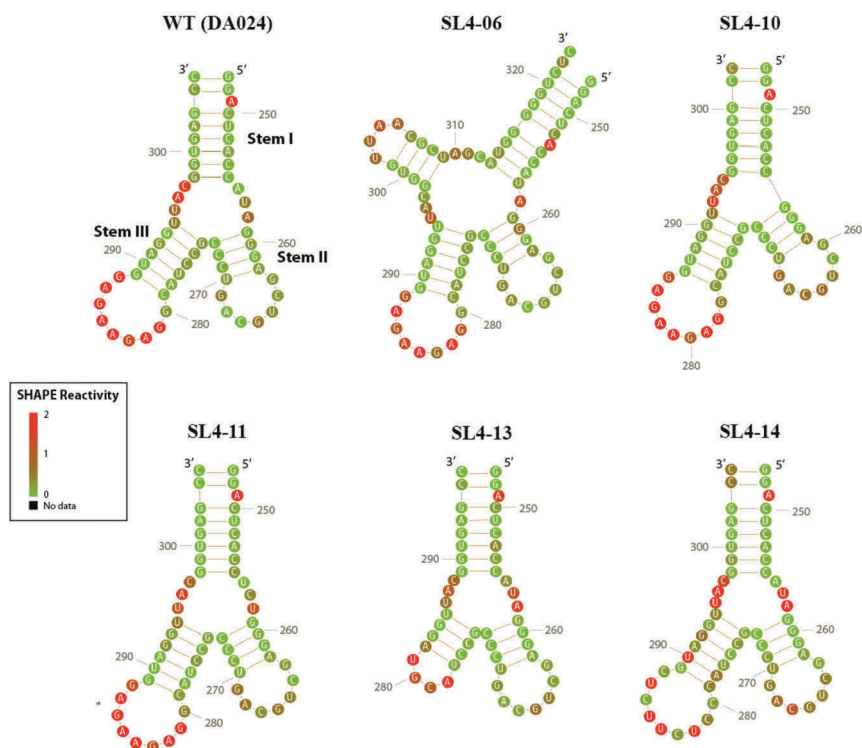


Figure 3. hSHAPE-validated structure of the SL4 domain in wild type and selected mutant transfer vector RNAs. hSHAPE experiments were performed in triplicate on *in vitro* transcribed RNA corresponding to nucleotides 1–713 of the wild type (DA024) and mutant (SL4-06, SL4-10, SL4-11, SL4-13 and SL4-14) transfer vectors. hSHAPE reactivity data (Supplemental Table 2) were introduced as constraints in the RNAstructure folding software (version 5.3; 40). For the sake of clarity, only the SL4 domain of the resulting structures was drawn using the Assemble graphical tool [69] and presented here. The hSHAPE reactivity values from 0 to 2 or greater are indicated by a linear color scale from green to red. WT: wild type.

corresponded to a complete loss of SL4 structure (Figures 1(d) and 4). Consistent with this, our previously published observations demonstrated that the deletion of pal II stem loop in totality (SA042P) also led to a complete loss of SL4 structure [25] and a corresponding loss of packaging (Figures 1(d) and 4). Interestingly, substitution of pal II sequences with HIV-1 pal containing additional purines (SA047P) did not restore RNA packaging and propagation, but was found to be competent for RNA dimerization [25]. These results reveal that the primary sequence of this loop is important for MMTV RNA packaging and suggests that pal II apical loop is involved in RNA-protein interactions during genome encapsidation in addition to dimerization. Similarly, either deletion of ssPurines or their substitution with pyrimidines in SL4-13 and SL4-14 mutants respectively, resulted in a loss of packaging despite folding of a SL4-like structure predicted by mFold and validated by hSHAPE experiments (Figure 1(d), 3, and 4). In SL4-13, the 9 nt deletion of ssPurines resulted in a loop consisting of four nucleotides from stem III comprising of two purines and two pyrimidines (Figures 3 and 4). In SL4-14 on the other hand, the ssPurines apical loop was essentially replaced by a ssPyrimidine loop (see yellow highlighted loop in Figure 4). Despite the conservation of the SL4 bifurcated structure in these mutants, and their ability to dimerize as efficiently as the wild type RNA (Supplemental Fig. 4), packaging was essentially lost, revealing the critical role of the ssPurine sequences in the RNA packaging process.

2.7. The central stem (I) is important for holding the SL4 bifurcation together

Structure-function analysis of the last set of mutants confirmed the observations made earlier that the unique bifurcated structure of SL4 is critical for function; in particular, the central stem I that holds the two bifurcated stem loops together is critical for the stability of the overall RNA secondary structure and consequently its function. This includes its ability to allow proper dimerization (employing pal II as DIS) and packaging of the dimeric genomic RNA into nascent virus particles [25]. In these set of mutants (SL4-15 to SL4-20), SL4 structure destabilizing mutations were introduced in stem I followed by compensatory mutations on the other arm of the stem to restore the structure without affecting the sequence of the two apical SL4 loops (pal II and ssPurines). The mFold prediction and the hSHAPE analysis of the SL4-15 mutant revealed that indeed the bifurcated structure of SL4 had been compromised by the destabilizing mutation, correlating well with the loss of RNA packaging (Figures 2(d), 5, and 6). The subsequent compensatory mutation in SL4-16 restored both the structure as well as RNA packaging using heterologous sequences (Figures 2(d), 5, and 6). This was not the case with the second set of mutants (SL4-17 and SL4-18) in which the destabilizing mutant (SL4-17) did result in a loss of SL4 structure and corresponding RNA packaging, but the compensatory mutant (SL4-18) was unable

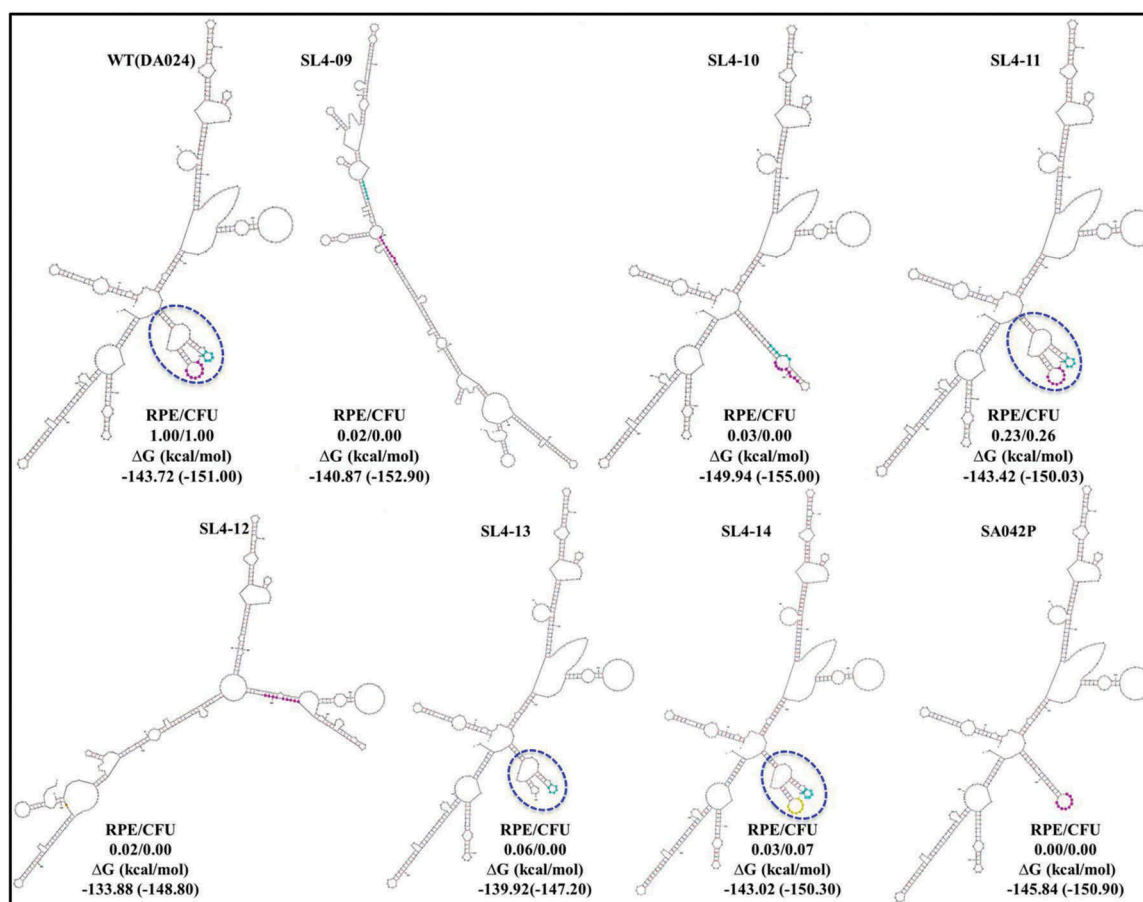


Figure 4. Predicted effect of specific SL4 mutations on MMTV RNA packaging signal structure. MFold structural predictions of the 5' end of the MMTV RNA containing deletions and substitutions in between the major splice donor (mSD) and Gag initiation codon (ATG) residing within stem loop SL4. Fuchsia color highlights the single-stranded purines (ssPurines) within SL4 structure, while peacock blue highlights the palindrome II (pal II) sequences that form the dimerization initiation site (DIS) in MMTV, respectively. The yellow color highlights the substituted ssPurine sequences with single-stranded pyrimidine sequences. The dashed circle highlights SL4 in the various wild type and mutant structures. WT: wild type; RPE: Relative Packaging Efficiency of transfer vector RNAs; CFU/ml: Colony Forming Unit per ml.

to restore either structure or RNA packaging (Figures 2(d) and 5). Actually, the hSHAPE experiments performed on these two mutants indicated that, as expected, the bifurcated structure was not conserved in SL4-17, but contrary to the predictions, the compensatory mutation introduced in SL4-18 only allowed the formation of stem-loop III with ssPurines in the loop, but not stem-loop II with the pal II loop (Figure 6). Consistent with base-pairing of pal II within a stem, dimerization of these mutant RNAs was strongly affected compared to the wild type (Supplemental Fig. 4). Energetically, it can be argued that replacing two strong G-C base pairs by two U-A base pairs at such a critical point where the central stem is bifurcating into two stem loops should reduce the strength of the bonds holding the structure together. This indeed had been predicted and observed. The last set of mutants confirmed this assertion since substitution of G299 for A in mutant SL4-19 led to the loss of the bifurcated SL4 structure (Figures 5 and 6) and the concomitant loss of function (Figure 2(d)), whereas introducing the complementary substitution of C254 by U in mutant SL4-20 (only the more distal G-C pair is changed with A-U, while maintaining the G-C pair proximal to the

internal loop) resulted in a complete restoration of structure as well as function (Figures 2(d), 5, and 6).

3. Discussion

Work presented in this study shows that the bifurcated SL4 structure within the 5' end of the MMTV genome is a key element for the packaging of MMTV gRNA, bringing two important single-stranded *cis*-acting elements of the viral genome together, DIS and ssPurines. Deletion of sequences immediately downstream of the DIS and ssPurines resulted in a complete abrogation of RNA packaging and propagation in mutants SL4-01 to SL4-08 despite the actual presence of these *cis*-acting sequences important for MMTV RNA packaging (Supplementary Fig. 2). These mutations resulted in a loss of the SL4 structure in all the mutants tested (Supplementary Fig. 3 and Figure 3), suggesting that the secondary structure of the SL4 region is important for MMTV RNA packaging. However, this could be due to the artifact of the introduced mutations since destabilization of secondary RNA structures at times can result in misfolding of hairpin structures, leading to aberrant phenotypes not associated with the RNA structure *per se*. Such erroneous effects have been previously reported

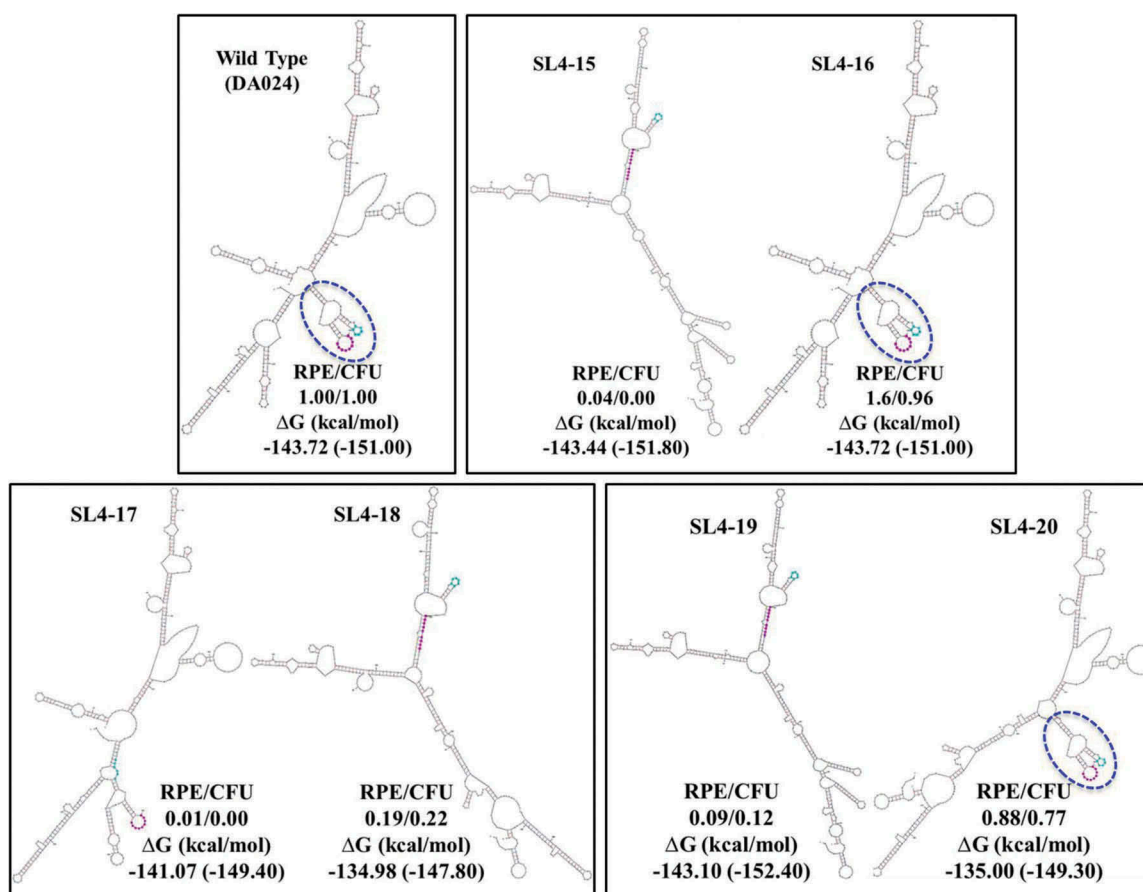


Figure 5. Predicted effect of destabilizing and compensatory mutations on MMTV RNA packaging signal structure. Mfold structural predictions of the 5' end of the MMTV RNA containing destabilizing deletion and compensatory substitution mutations in SL4. Fuchsia color highlights the single-stranded purines (ssPurines) within SL4 structure, while peacock blue highlights the palindrome II (pal II) sequences that form the dimerization initiation site (DIS) in MMTV, respectively. The dashed circle highlights SL4 in the various wild type and mutant structures. WT: wild type; RPE: Relative Packaging Efficiency of transfer vector RNAs; CFU/ml: Colony Forming Unit per ml.

with partial deletions of HIV-1 TAR that have included effect of TAR on RNA dimerization, packaging, and polyadenylation [36–38]. Therefore, we went on to introduce two sets of more directed single and multiple nucleotide changes targeting the stems and loops of the SHAPE-validated structure of SL4 to determine whether their effect on MMTV RNA packaging was consistent (Figures 1 and 2).

The importance of the bifurcated structure of SL4 for MMTV RNA packaging and propagation is supported by the analysis of SL4-09 mutant as well as the analysis of paired mutations that either destabilized the stem I of SL4 or recreated it by compensatory mutations without affecting the sequence of DIS or ssPurines (SL4-15 to 20) (Figures 1 and 2). Further interrogation of the sub-regions of SL4 revealed that the size, and to a lesser extent, the sequence of the internal loop or/and the exact geometry of the three-stem junction are important factors for RNA packaging and propagation (Figures 1 and 2). Indeed, mutants SL4-06 and SL4-10 in which stem-loops II and III are maintained together by a larger (SL4-06) or a smaller (SL4-10) internal loop lost their activity (Supplementary Fig. 3; Figures 3–4). The sequence as well as the single-stranded loop nature of ssPurines is critical for the packaging potential of MMTV since its direct deletion or substitution still resulted in a complete loss of RNA packaging potential (mutants SL4-13 and

SL4-14) even when the bifurcated structure was maintained (Figures 1, 3, and 4). The data presented in this study cannot conclude on the importance of the pal II sequence by its own since its deletion always led to the loss of the SL4 bifurcated structure (Figures 1 and 4); however, our previous study clearly demonstrates the importance of both the sequence of pal II and its single-stranded character [25]. Also, although we observed good correlation between our packaging and propagation data, one cannot exclude that the propagation defects may partially result from other defective step(s), such as reverse transcription or integration, in addition to defective packaging.

The presence of a stretch of purines in the packaging sequences on retroviral gRNA has been proposed to facilitate RNA packaging by functioning as a potential NC binding site [27–29,31]. The specificity for gRNA during packaging, thus, has been ascribed to high-affinity interactions between the NC domain of Gag and the packaging signal [5–13]. We have made a similar observation in the case of MPMV, where genetic and structure-prediction analyses have suggested that either ssPurines (Region A) or its partially repeated but base-paired sequence (Region B) plays a crucial role in gRNA packaging, possibly by functioning as a NC binding site [32,33].

The significance of the SL4 structure to RNA packaging is highlighted by the observation that even a one nucleotide

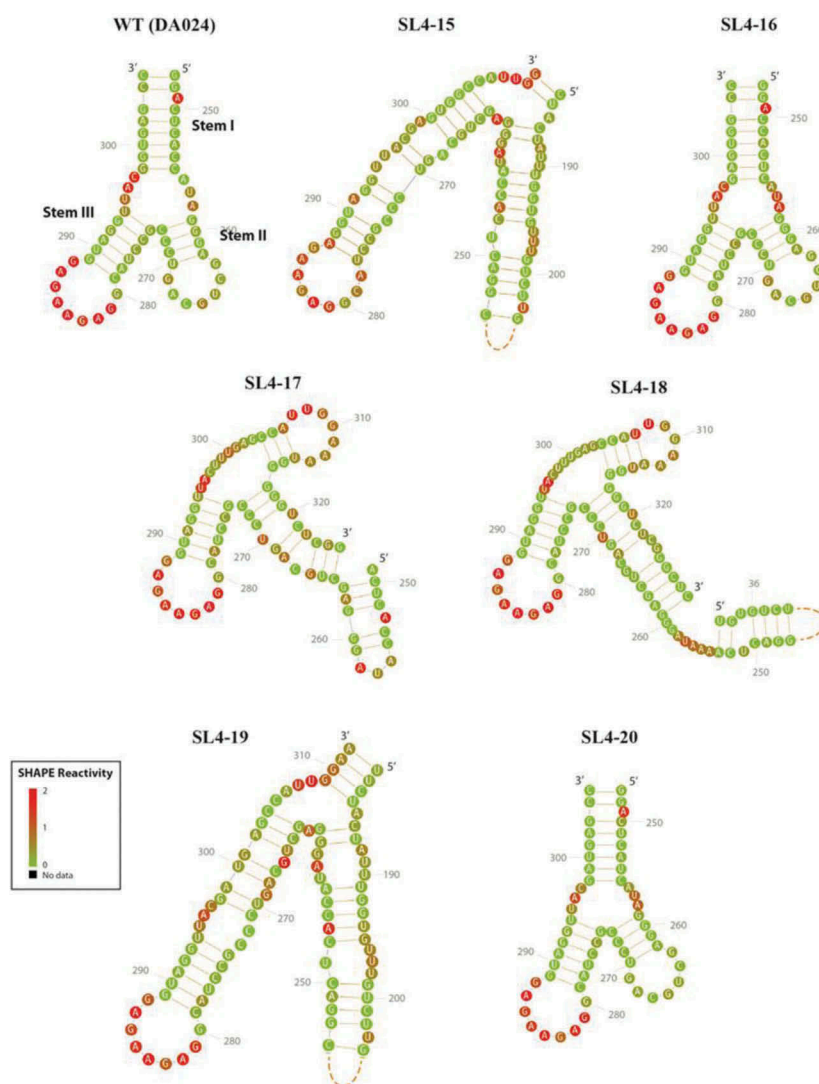


Figure 6. hSHAPE-validated structure of the SL4 domain in the wild type and selected mutant transfer vector RNAs. hSHAPE experiments were performed in triplicate on *in vitro* transcribed RNA corresponding to nucleotides 1–713 wild type (DA024) and mutant (SL4-15, SL4-16, SL4-17, SL4-18, SL4-19 and SL4-20) transfer vectors. hSHAPE reactivity data (Supplemental Table 2) were introduced as constraints in the RNAstructure folding software (version 5.3; 40). For the sake of clarity, only the SL4 domain of the resulting structures was drawn using the Assemble graphical tool [69] and presented here. The hSHAPE reactivity values from 0 to 2 or greater are indicated by a linear color scale from green to red. WT: wild type.

mutation in a destabilizing region of the structure, a G299 to A mutation at the base of stem I in SL4-19, was sufficient to disrupt the structure of SL4 and its subsequent RNA packaging and propagation despite the presence of intact ssPurines and pal II (Figures 2, 5, and 6). This was evidenced by the fact that RNA packaging could be restored by a corresponding compensatory mutation (C254 to U mutation) on the other arm of the stem in SL4-20. The effect of the one nucleotide SL4-19 mutation was as drastic as the six-nucleotide stem mutation in SL4-15, that could also be balanced by a compensatory mutation on the opposite arm of the stem in SL4-16 (Figures 2, 5, and 6). Furthermore, we should also point out that the distal G299 in SL4-19 may be less important than the G298 proximal to the internal loop that was also mutated along with the distal G in SL4-17 (Figure 2). This assertion is made since G299 was substituted along with the compensatory mutations in two clones SL4-16 (along with other sequences in this case) and SL4-20 (on its own), and in both cases these mutants restored

RNA packaging, but this was not the case when the proximal base-pair was restored in SL4-18. Only partial packaging was restored in this mutant, suggesting that this G-C base pair is critical in holding the whole SL4 structure together, as indicated by our hSHAPE experiments (Figure 6).

Work presented in this study also clarifies how MMTV may specifically select the gRNA for encapsidation; i.e., deletion of short sequences between the major splicing donor (mSD) and start codon of Gag completely abolished RNA packaging, even though the *cis*-acting packaging signals such as ssPurines and pal II were still present in the RNA (mutants SL4-01 to SL4-08; Supplementary Fig. 2). This allows us to propose a model where the full-length, unspliced, genomic RNA is preferentially packaged not only due to the presence of these *cis*-acting sequences, but also due to a structurally competent encapsidation sequence (i.e., a bifurcated SL4 comprising of pal II and ssPurines apical loops). The SL4 structure is disrupted during splicing, thereby limiting the ability of spliced RNAs to package into the nascently-

forming virus particles. Thus, in MMTV, a combination of both *cis*-acting sequences and structural elements are responsible for the recognition and differential packaging of the unspliced RNA during the encapsidation process. The case of human immunodeficiency virus type 1 (HIV-1) may somewhat be similar to MMTV: the 5' UTR in HIV-1 implicated in RNA packaging folds into several stem loops, SL1-4 [39,40] of which SL1 harbors a GC-rich 6-nt pal sequence in the form of a loop that has been shown to function as the DIS [41,42]. Until recently SL3, which is present downstream of mSD, had been proposed to be the main HIV-1 packaging signal since it harbors a conserved region containing a GGAG tetraloop that has been shown to bind NC [43]. However, it has now been shown that SL1, in addition to containing the HIV-1 DIS, also harbors a single-stranded purine rich internal loop (ssPurine: GAGG) that acts as the main site for HIV-1 Gag binding during gRNA packaging [26,30,44,45]. This internal ssPurine loop as well as the DIS, however, are located upstream of the mSD and therefore are part of both the spliced and unspliced viral mRNA. Thus, in the case of HIV-1, selection of the gRNA for packaging is likely to be dependent upon the higher order structure of the major packaging determinant of the gRNA which is disrupted upon splicing [26,30,44,45]. In the case of HIV-2, even though both the spliced and unspliced viral mRNAs harbor RNA packaging sequences, only the unspliced message is packaged into nascently formed virus particle since HIV-2 Gag preferentially package gRNA in *cis* [46–48].

Interestingly, in almost all retroviruses, sequences augmenting gRNA dimerization and packaging have been mapped nearly to the same 100 to 400 nucleotides in the 5' region of the gRNA [5–13], which also harbors the dimer linkage structure (DLS) [5,7–13]. Based on our findings, it is reasonable to propose that the bifurcated SL4 in MMTV packaging signal RNA that contains phylogenetically-conserved pal II and ssPurines [25] functions as a major dimerization and packaging determinant during MMTV gRNA packaging process, during which pal II initiates gRNA dimerization and ssPurines functions as a potential Gag binding site for specific incorporation of the dimeric RNA into the newly forming virus particles. Thus, SL4 could interlink the process of dimerization with packaging both physically and mechanistically.

This observation lends support to the developing hypothesis that retroviral packaging determinants are dynamic in nature and require structural flexibility to allow successful recognition of the packaging substrate RNA that is most likely recognized as a dimer by the Gag proteins [5–7,49]. Thus, packaging signals may act as 'structural switches' that favor either translation or dimerization/packaging of the gRNA, and also help the virus differentiate between spliced and unspliced viral RNAs. Several riboswitch models have been proposed for HIV-1 [50–53], and in the case of FIV, the existence of dual structures has been validated by SHAPE and RNA dimerization assays [54]. Flexibility in retroviral gRNA is created *via* long-range interactions that have now been observed in several retroviruses, from HIV-1, FIV, MLV, to MPMV and MMTV [6,25,26,33,55–59]. The DIS in one conformation is occluded by base pairing and not available for dimerization, thus earmarking this RNA for translation, while the DIS in the other conformation is

available as an apical loop on a stem loop, thus allowing dimerization and packaging of the dimeric RNA [54]. In a similar manner, we propose that SL4 may be part of an RNA structural switch that may facilitate the ability of MMTV to convert its genomic RNA from a form that favors translation to a dimeric form that may act as a substrate for RNA packaging. Work is in progress to support this assertion experimentally in MMTV.

4. Material and methods

4.1. Genome nucleotide numbering system

The MMTV nucleotide numbering system refers to the nucleotide positions to HYBMTV, a molecular clone created by Shackleford and Varmus [60].

4.2. Plasmid construction

The MMTV packaging construct, JA10, expresses the MMTV *gag/pol* genes under the transcriptional control of the human cytomegalovirus (hCMV) intron A promoter/enhancer [23–25]; Supplemental Figure 1). MD.G is a vesicular stomatitis virus glycoprotein G (VSV-G) expression vector [61]; Supplemental Figure 1). MMTV sub-genomic transfer vector, DA024 [23–25]; Supplemental Fig. 1) harbors the *hygromycin B phosphotransferase* gene expressed from an internal simian virus 40 (SV40) early promoter (SV-Hyg^r) (Supplemental Fig. 1).

Several mutations, including deletions and substitutions were introduced in sub-genomic transfer vector (DA024) within the sequences involved in forming SL4 (Supplementary Fig. 2; Figures 1–2) through splice overlap extension (SOE) polymerase chain reaction (PCR), as previously described [24,25]. Briefly, SOE PCR required two separate rounds of amplifications. Round 1 PCR reactions in our case involved two independent PCRs, A and B in which PCR (A) was performed using the universal outer (sense; S) primer OTR 249 along with the inner antisense (AS) primer that varied for different mutant clones (Supplemental Table 1) so that a desired mutation could be introduced. In PCR (B), the AS outer primer OTR 552 was used along with the inner S primer which again varied for different mutant clones (Supplemental Table 1). Amplified products from PCRs A and B (containing overlapping complementary sequences allowing them to anneal and function as a template) were used as target for amplification in round 2 PCR using S and AS outer primers (OTR 249/OTR 552; Supplemental Table 1) to generate a final amplified product harboring the desired mutations. To introduce the MMTV sequences into the sub-genomic transfer vector, the PCR-amplified mutant products were cleaved with *SpeI* restriction endonuclease and cloned into a similarly-digested wild type transfer vector, DA024. Such an approach resulted in substituting the wild type sequences with the PCR-amplified products harboring the desired mutations. All mutant clones were confirmed by DNA sequencing.

4.3. Nucleocytoplasmic fractionation, virion and RNA isolation, and cDNA preparation

Following transfection, cultured cells were fractionated into nuclear and cytoplasmic fractions, as described previously [24,25,32,57,59,62,63]. Briefly, packaged viral RNA from the pelleted viral particles and cellular RNAs from the cytoplasmic fractions were isolated using Trizol and Trizol LS reagents (Invitrogen), respectively. RNA samples were treated with TURBO™ DNase (Invitrogen) and amplified for 30 cycles by conventional PCR using vector-specific primers OTR 671(S) or and OTR 672 (AS; Supplemental Table 1) employing PCR conditions described previously [25]. Following confirmation by PCR that no detectable amount of contaminating plasmid DNA was present in RNA samples, the DNase-treated RNAs were converted into cDNAs and tested to establish the amplifiability of cDNA preparations through PCR. Integrity of the nuclear membrane during the fractionation process was established by amplifying cDNAs prepared from the cytoplasmic fraction in multiplex PCR using oligos OTR 582(S) and OTR 581 (AS; Supplemental Table 1) which should only amplify unspliced β -actin mRNA. Additionally, as an ancillary control for the presence of amplifiable cDNA in the multiplex PCRs, cDNAs were also amplified using primers/competimer for 18S ribosomal RNA (18S Quantum competimer control, Ambion). Finally, to ascertain the stability as well as proper nuclear export of mutant transfer vector RNAs, cytoplasmic cDNA preparations were amplified using MMTV-specific vector primer pair OTR 671/OTR 672 (Supplemental Table 1).

4.4. Real time quantitative RT-PCR (RT-qPCR) of transfer vector RNA

To quantitate the relative levels of transfer vector RNAs expressed in the cytoplasm and subsequently packaged into the viral particles, a custom-made Taqman gene expression assay (Applied Biosystems Inc., ABI) was developed, as described previously [24,25]. This assay uses β -actin (human β -actin endogenous control assay, ABI #4326315E) as an endogenous control for qPCR. Briefly, this assay employed a FAM/MGB-labeled probe along with primers within the U5 region, a region common to the wild type and all mutant transfer vector RNAs and away from the site of the mutations analyzed. Although this region is upstream of the major splice donor of MMTV, there are no known splice acceptor sites in the wild type as well as mutant transfer vectors, including the SV-Hyg^r cassette, thereby preventing the generation of any spliced RNAs.

Once the RT-qPCR assay was conducted on both the cytoplasmic and viral cDNA samples, the RPE for each mutant was determined by dividing the packaged mutant viral RNA relative quantification (RQ) values by the luciferase-normalized cytoplasmic expression of the respective mutant (to take into account the varying transfection efficiencies of the different constructs). Finally, the results of RNA packaging of all the mutants were presented relative to that of the wild type (DA024) construct.

4.5. High throughput selective 2'-hydroxyl acylation analyzed by primer extension (hSHAPE) methodology

hSHAPE experiments [64–66] were performed on *in vitro* transcribed RNAs corresponding to nucleotides 1 to 713 using benzoyl cyanide (BzCN) and following the protocol described previously [25,33,57]. Briefly, RNAs were modified in a buffer favoring gRNA dimerization (50 mM sodium cacodylate (pH 7.5), 300 mM KCl and 5 mM MgCl₂) in the presence of 2 μ g total yeast tRNA (Sigma Aldrich). The modified RNAs were reverse transcribed using AS primers OTRs 9 and 10: (5' AACAGATTTGGCTTCTGCGG 3'; MMTV nt 614–633) labelled with VIC or NED, respectively. These primers allowed analysis of the RNA structure from approximately nucleotide 560 to 100. The primer extension products were loaded onto an Applied Biosystems 3130xl genetic analyzer and the electropherograms were analyzed with the QuShape software [67]. The hSHAPE reactivities obtained with QuShape from three independent experiments were highly reproducible. These reactivities were averaged (Supplemental Table 2) before being used as constraints to fold the RNA secondary structure of the MMTV packaging signal with the RNAstructure software version 5.3 [68] to determine the effects of mutations on the overall higher order structure. Based on the RNAstructure data, the SL4 sub-region of WT and mutants RNAs were drawn using the Assemble graphical tool [69].

4.6. In silico analysis of MMTV packaging signal RNA secondary structure

The secondary structure of the 5' region of the MMTV gRNA of the wild type, DA024, (containing the region between R and the first 120 nt of *gag*; a total of 432 nt of the 5' end of the genome from R) and mutant MMTV genomes were folded using MFold [34,35] to correlate the effects of the introduced mutations in SL4 on the overall packaging signal RNA secondary structure and establish structure-function relationship of this region during MMTV RNA packaging.

4.7. Statistical analysis

For quantitation and determination of statistically significant differences in relative packaging efficiencies between the wild type and mutant clones, the standard paired, two-tailed Students *t*-test was performed. A *P*-value of 0.05 was considered to be significant.

Acknowledgments

Authors would like to thank Ms. Faiza Moureen, Dr. Jaleel Kizhakkayil, and Dr. Suriya Aktar (Department of Microbiology & Immunology, CMHS, UAEU) for their help in some of the cloning and experimental procedures at the initial stages of the project as a part of their training and Dr Fabrice Jossinet for his help in drawing the secondary structure models of Figures 3 and 5 with Assemble. The authors would also like to thank the editorial assistance of Akhil Chameettachal, Fathima Nuzra Nagoor Pitchai, and Vineeta Pillai (Department of Microbiology & Immunology, CMHS, UAEU) for reading the manuscript and providing their inputs.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was funded by a joint grant from the United Arab Emirates University (UAEU; Zayed Center for Health Sciences) and Terry Fox Funds for Cancer Research grants (fund codes 31R020 and 21M095 respectively) and College of Medicine & Health Sciences grant (NP-14-34) to TAR and in part from and UAEU-National Research Foundation grant (31M101) and UAEU CMHS grant 31M331 to FM.

ORCID

Valérie Vivet-Boudou  <http://orcid.org/0000-0001-8702-1047>

References

- Dudley JP, Golovkina TV, Ross SR. Lessons learned from mouse mammary tumor virus in animal models. *ILAR J.* 2016;57:12–23.
- Mertz JA, Simper MS, Lozano MM, et al. Mouse mammary tumor virus encodes a self-regulatory RNA export protein and is a complex retrovirus. *J Virol.* 2005;79:14737–14747.
- Indik S. Mouse mammary tumor virus-based vector for efficient and safe transgene delivery into mitotic and non-mitotic cells. *Cell & Gene Thera Insights.* 2016;2:589–597.
- Salmons B, Moritz-Legrand S, Garcha I, et al. Construction and characterization of a packaging cell line for MMTV-based conditional retroviral vectors. *Biochem Biophys Res Commun.* 1989;159:1191–1198.
- Kuzembayeva M, Dilley K, Sardo L, et al. Life of psi: how full-length HIV-1 RNAs become packaged genomes in the viral particles. *Virology.* 2014;454–455:362–370.
- Mailler E, Bernacchi S, Marquet R, et al. The life-cycle of the HIV-1 Gag-RNA complex. *Viruses.* 2016;8(9):248.
- 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(9):257.
- D'Souza V, Summers MF. How retroviruses select their genomes. *Nat Rev Microbiol.* 2005;3:643.
- Johnson SF, Telesnitsky A. Retroviral RNA dimerization and packaging: the what, how, when, where, and why. *PLOS Pathog.* 2010;6:e1001007.
- Lever AML. HIV-1 RNA packaging. *Adv Pharmacol San Diego Calif.* 2007;55:1–32.
- Ali LM, Rizvi TA, Mustafa F. Cross- and Co-Packaging of retroviral RNAs and their consequences. *Viruses.* 2016;8:276.
- Comas-Garcia M, Davis SR, Rein A. On the selective packaging of genomic RNA by HIV-1. *Viruses.* 2016;8:246.
- Dubois N, Marquet R, Paillart J-C, et al. Retroviral RNA dimerization: from structure to functions. *Front Microbiol.* 2018;9:527.
- Adam MA, Miller AD. Identification of a signal in a murine retrovirus that is sufficient for packaging of nonretroviral RNA into virions. *J Virol.* 1988;62:3802–3806.
- Hibbert CS, Mirro J, Rein A. mRNA molecules containing murine leukemia virus packaging signals are encapsidated as dimers. *J Virol.* 2004;78:10927–10938.
- Al Dhaheri NS, Phillip PS, Ghazawi A, et al. Cross-packaging of genetically distinct mouse and primate retroviral RNAs. *Retrovirology.* 2009;6:66.
- Al Shamsi IR, Al Dhaheri NS, Phillip PS, et al. Reciprocal cross-packaging of primate lentiviral (HIV-1 and SIV) RNAs by heterologous non-lentiviral MPMV proteins. *Virus Res.* 2011;155:352–357.
- Rizvi TA, Panganiban AT. Simian immunodeficiency virus RNA is efficiently encapsidated by human immunodeficiency virus type 1 particles. *J Virol.* 1993;67:2681–2688.
- Yin PD, Hu W-S. RNAs from genetically distinct retroviruses can copackage and exchange genetic information in vivo. *J Virol.* 1997;71:6237–6242.
- White SM, Renda M, Nam N-Y, et al. Lentivirus vectors using human and simian immunodeficiency virus elements. *J Virol.* 1999;73:2832–2840.
- Parveen Z, Mukhtar M, Goodrich A, et al. Cross-packaging of human immunodeficiency virus type 1 vector RNA by spleen necrosis virus proteins: construction of a new generation of spleen necrosis virus-derived retroviral vectors. *J Virol.* 2004;78:6480–6488.
- Moore MD, Fu W, Nikolaitchik O, et al. Dimer initiation signal of human immunodeficiency virus type 1: its role in partner selection during RNA copackaging and its effects on recombination. *J Virol.* 2007;81:4002–4011.
- Rizvi TA, Ali J, Phillip PS, et al. 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.
- Mustafa F, Amri DA, Ali FA, 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.
- Aktar SJ, Vivet-Boudou V, Ali LM, et al. Structural basis of genomic RNA (gRNA) dimerization and packaging determinants of mouse mammary tumor virus (MMTV). *Retrovirology.* 2014;11:96.
- Abd El-Wahab EW, Smyth RP, Mailler E, et al. Specific recognition of the HIV-1 genomic RNA by the Gag precursor. *Nat Commun.* 2014;5:4304.
- Lever AM. RNA packaging in lentiviruses. *Retrovirology.* 2009;6:113.
- Moore MD, Hu WSHIV-1. RNA dimerization: it takes two to tango. *AIDS Rev.* 2009;11:91–102.
- Paillart JC, Westhof E, Ehresmann C, et al. Non-canonical interactions in a kissing loop complex: the dimerization initiation site of HIV-1 genomic RNA. *J Mol Biol.* 1997;270:36–49.
- Smyth RP, Despons L, Huili G, et al. Mutational interference mapping experiment (MIME) for studying RNA structure and function. *Nat Methods.* 2015;12:866–872.
- Zeffman A, Hassard S, Varani G, et al. 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.
- Jaballah SA, Aktar SJ, Ali J, et al. 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.
- Aktar SJ, Jabeen A, Ali LM, et al. 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.
- Mathews DH, Sabina J, Zuker M, et al. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J Mol Biol.* 1999;288:911–940.
- Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 2003;31:3406–3415.
- Das AT, Vrolijk MM, Harwig A, et al. Opening of the TAR hairpin in the HIV-1 genome causes aberrant RNA dimerization and packaging. *Retrovirology.* 2012;9:59.
- Vrolijk MM, Harwig A, Berkhout B, et al. Destabilization of the TAR hairpin leads to extension of the polyA hairpin and inhibition of HIV-1 polyadenylation. *Retrovirology.* 2009;6:13.
- Vrolijk MM, Ooms M, Harwig A, et al. Destabilization of the TAR hairpin affects the structure and function of the HIV-1 leader RNA. *Nucleic Acids Res.* 2008;36:4352–4363.
- Clever J, Sasseti 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.
- McBride MS, Schwartz MD, Panganiban AT. Efficient encapsidation of human immunodeficiency virus type 1 vectors and further characterization of cis elements required for encapsidation. *J Virol.* 1997;71:4544–4554.

41. Paillart JC, Skripkin E, Ehresmann B, et al. 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.
42. Skripkin E, Paillart JC, Marquet R, et al. Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro. *Proc Natl Acad Sci U S A*. 1994;91:4945–4949.
43. De Guzman RN, Wu ZR, Stalling CC, et al. Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. *Science*. 1998;279:384–388.
44. Bernacchi S, Abd El-Wahab EW, Dubois N, et al. HIV-1 Pr55Gag binds genomic and spliced RNAs with different affinity and stoichiometry. *RNA Biol*. 2017;14:90–103.
45. Smyth RP, Smith MR, Jousset A-C, 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.
46. Balvay L, Lopez Lastra M, Sargueil B, et al. Translational control of retroviruses. *Nat Rev Microbiol*. 2007;5:128–140.
47. Griffin SD, Allen JF, Lever AM. The major human immunodeficiency virus type 2 (HIV-2) packaging signal is present on all HIV-2 RNA species: cotranslational RNA encapsidation and limitation of Gag protein confer specificity. *J Virol*. 2001;75:12058–12069.
48. Kaye JF, Lever AM. Human immunodeficiency virus types 1 and 2 differ in the predominant mechanism used for selection of genomic RNA for encapsidation. *J Virol*. 1999;73:3023–3031.
49. Rein A, Datta SAK, Jones CP, et al. Diverse interactions of retroviral Gag proteins with RNAs. *Trends Biochem Sci*. 2011;36:373–380.
50. Berkhout B, Ooms M, Beerens N, et al. In vitro evidence that the untranslated leader of the HIV-1 genome is an RNA checkpoint that regulates multiple functions through conformational changes. *J Biol Chem*. 2002;277:19967–19975.
51. 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.
52. Ooms M, Huthoff H, Russell R, et al. A riboswitch regulates RNA dimerization and packaging in human immunodeficiency virus type 1 virions. *J Virol*. 2004;78:10814–10819.
53. Lu K, Heng X, Garyu L, et al. NMR detection of structures in the HIV-1 5′-leader RNA that regulate genome packaging. *Science*. 2011;334:242–245.
54. Kenyon JC, Tanner SJ, Legiewicz M, 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.
55. D’Souza V, Summers MF. Structural basis for packaging the dimeric genome of Moloney murine leukaemia virus. *Nature*. 2004;431:586.
56. Gherghe C, Lombo T, Leonard CW, et al. Definition of a high-affinity Gag recognition structure mediating packaging of a retroviral RNA genome. *Proc Natl Acad Sci U S A*. 2010;107:19248–19253.
57. Kalloush RM, Vivet-Boudou V, Ali LM, et al. 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.
58. Kenyon JC, Ghazawi A, Cheung WKS, et al. 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.
59. Rizvi TA, Kenyon JC, Ali J, 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.
60. 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*. 1988;85:9655–9659.
61. Naldini L, Blömer U, Gally P, et al. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science*. 1996;272:263–267.
62. Ghazawi A, Mustafa F, Phillip PS, et al. Both the 5′ and 3′ LTRs of FIV contain minor RNA encapsidation determinants compared to the two core packaging determinants within the 5′ untranslated region and gag. *Microbes Infect*. 2006;8:767–778.
63. Mustafa F, Ghazawi A, Jayanth P, et al. 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.
64. Merino EJ, Wilkinson KA, Coughlan JL, et al. RNA structure analysis at single nucleotide resolution by selective 2′-hydroxyl acylation and primer extension (SHAPE). *J Am Chem Soc*. 2005;127:4223–4231.
65. Mortimer SA, Weeks KM. Time-resolved RNA SHAPE chemistry: quantitative RNA structure analysis in one-second snapshots and at single-nucleotide resolution. *Nat Protoc*. 2009;4:1413–1421.
66. Mortimer SA, Weeks KM. A fast-acting reagent for accurate analysis of RNA secondary and tertiary structure by SHAPE chemistry. *J Am Chem Soc*. 2007;129:4144–4145.
67. Karabiber F, McGinnis JL, Favorov OV, et al. QuShape: rapid, accurate, and best-practices quantification of nucleic acid probing information, resolved by capillary electrophoresis. *RNA*. 2013;19:63–73.
68. Reuter JS, Mathews DH. RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics*. 2010;11:129.
69. Jossinet F, Ludwig TE, Westhof E. Assemble: an interactive graphical tool to analyze and build RNA architectures at the 2D and 3D levels. *Bioinforma Oxf Engl*. 2010;26:2057–2059.