doi:10.1016/j.jmb.2010.06.043 J. Mol. Biol. (2010) 401, 996–1014

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

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

Soumeya Ali Jaballah, Suriya J. Aktar, Jahabar Ali, Pretty Susan Phillip, Noura Salem Al Dhaheri, Aayesha Jabeen and Tahir A. Rizvi*

Department of Microbiology and Immunology, Faculty of Medicine and Health Sciences, United Arab Emirates University, P.O. Box 17666, Al Ain, United Arab Emirates

Received 24 May 2010; received in revised form 21 June 2010; accepted 21 June 2010
Available online
30 June 2010

During retroviral RNA packaging, two copies of genomic RNA are preferentially packaged into the budding virus particles whereas the spliced viral RNAs and the cellular RNAs are excluded during this process. Specificity towards retroviral RNA packaging is dependent upon sequences at the 5′ end of the viral genome, which at times extend into Gag sequences. It has earlier been suggested that the Mason–Pfizer monkey virus (MPMV) contains packaging sequences within the 5′ untranslated region (UTR) and Gag. These studies have also suggested that the packaging determinants of MPMV that lie in the UTR are bipartite and are divided into two regions both upstream and downstream of the major splice donor. However, the precise boundaries of these discontinuous regions within the UTR and the role of the intervening sequences between these dipartite sequences towards MPMV packaging have not been investigated. Employing a combination of genetic and structural prediction analyses, we have shown that region "A", immediately downstream of the primer binding site, is composed of 50 nt, whereas region "B" is composed of the last 23 nt of UTR, and the intervening 55 nt between these two discontinuous regions do not contribute towards MPMV RNA packaging. In addition, we have identified a 14-nt G–C-rich palindromic sequence (with 100% autocomplementarity) within region A that has been predicted to fold into a structural motif and is essential for optimal MPMV RNA packaging. Furthermore, we have also identified a stretch of single-stranded purines (ssPurines) within the UTR and 8 nt of these ssPurines are duplicated in region B. The native ssPurines or its repeat in region B when predicted to refold as ssPurines has been shown to be essential for RNA packaging, possibly functioning as a potential nucleocapsid binding site. Findings from this study should enhance our understanding of the steps involved in MPMV replication including RNA encapsidation process.

Keywords: Mason–Pfizer Monkey Virus (MPMV); retroviral RNA packaging; RNA secondary-structure predictions; palindromic sequence; single-stranded purines

Edited by M. F. Summers

*Corresponding author. E-mail address: tarizvi@uaeu.ac.ae.

Abbreviations used: MPMV, Mason–Pfizer monkey virus; UTR, untranslated region; ssPurine, single-stranded purine; NC, nucleocapsid; mSD, major splice donor; HIV-1, human immunodeficiency virus type 1; SIV, simian immunodeficiency virus; DIS, dimerization initiation site; LRI, long-range interaction; PBS, primer binding site; OD, optical density; RPE, relative packaging efficiency; CFU, colony-forming unit.

Introduction

Retroviral RNA packaging is a highly ordered and specific process during which unspliced genomic RNAs are preferentially encapsidated over other cellular and spliced viral mRNAs, despite the fact that retroviral genomic RNAs constitute 1% or less of the total cellular mRNAs. RNA packaging among retroviruses involves the recognition of sequences at the 5′ end of the genomic RNA, called the packaging sequences or signal (ψ), by the nucleocapsid (NC) domain of the Gag or Gag-Pol polyprotein.[1,2] The ψ in retroviruses can be a continuous region or a multipartite sequence residing in the 5′ untranslated region (UTR) and 5′ end of Gag.[1,2] Retroviral packaging determinants usually involve sequences downstream of the major spliced donor (mSD) in order to distinguish between spliced and unspliced viral transcripts (except in the case of Rous sarcoma virus and avian leukosis virus). For example, in human immunodeficiency virus type 1 (HIV-1), stem loop 2 (SL2), a major packaging determinant, binds with high affinity to NC and contains the mSD, and therefore, during splicing, the structural motif of this packaging determinant and the surrounding structures is destabilized, leading to its absence in spliced transcripts, thus minimizing their packaging ability.[1] Although RNA packaging is a ubiquitous process in all retroviruses, no sequence conservation between the packaging signals of different retroviruses has been found. However, almost invariably, it has been shown that this 5′ end RNA region (containing ψ) of all known retroviruses assumes a higher-order structure composed of different structural motifs.[1,2] Therefore, the involvement of such RNA structural motifs on the retroviral genomic transcripts (at times regardless of the primary sequence) has been strongly associated with retroviral RNA encapsidation and could explain the phenomenon of RNA cross-packaging among diverse retroviruses.[3]

In retroviral replication, the process of RNA packaging is thought to be closely linked with RNA dimerization, and in some, the former depends on the latter.[2,4] During RNA dimerization, two genomic RNAs non-covalently interact at their 5′ ends and, at least in some retroviruses, it has been shown that genomic RNAs are packaged as dimers.[4–6] Consistent with this, a number of reports have shown the overlapping packaging and dimerization initiation sites (DIS) at the 5′ end of the genomic mRNA.[6–9] Palindromic (pal) and semi-palindromic sequences have been strongly implicated in dimerization of retroviral genomic RNAs.[9,10] The autocomplementarity of pal allows base pairing between two RNAs in an antiparallel direction to initiate "kissing loop" complex that ultimately leads to a stable dimer formation.[4,10] The presence of pal in the region encompassing packaging determinants reflects the overlapping nature of packaging and dimerization sequences in a number of retroviruses.[4,11–13] Consistent with this, mutations in these overlapping regions that inhibit dimerization generally affect packaging,[14–18] and a conserved 10-nt pal sequence in HIV-2 UTR [as well as in different strains of simian immunodeficiency virus (SIV)] has been shown to influence both RNA packaging and dimerization.[17–19]

The packaging determinants of simple retroviruses are believed to be discrete and confined in a rather defined region at the 5′ end of the viral genome involving only UTR sequences or may extend into the 5′ end of Gag.[1,20,21] For example, the core packaging signals of spleen necrosis virus, murine leukemia virus, and feline leukemia virus are located downstream of the mSD, allowing for packaging of the unspliced viral RNA only,[1,22–24] whereas the packaging determinants of complex retroviruses seem to be more spread out and multipartite. In HIV-1, for example, efficient packaging requires the full 5′ UTR, which folds into several stem loop structures (SL1–SL3) that have been associated with RNA packaging.[1,2] Similarly, the major packaging determinants of feline immunodeficiency virus are composed of two discontinuous regions located within the 5′ end of the RNA genome.[25–29] The intervening sequences between these two discontinuous packaging determinants do not contribute towards RNA packaging or propagation.[29] It has been shown that the region spanning feline immunodeficiency virus packaging determinants folds into a complex RNA secondary structure composed of five conserved stem loop structures (SL1–SL5) and a conserved long-range interaction (LRI) between R/U5 and Gag. In addition, a 10-nt pal sequence within the matrix coding region of Gag has been identified and proposed to act as DIS.[13]

Mason–Pfizer monkey virus (MPMV), a simple retrovirus, offers several features, making it a potential candidate vector for human gene therapy. For example, its non-human primate nature should minimize chances of recombination with human endogenous retroviruses. In addition, the high transcriptional activity of the MPMV promoter in human cells and the presence of the constitutive transport element in MPMV vectors should facilitate efficient nucleocytoplasmic transport and expression of the therapeutic gene(s). However, before MPMV vectors can be exploited for such studies, a better and enhanced understanding of the basic steps of its life cycle, such as RNA packaging, is imperative. A generalized 624-nt region downstream of MPMV primer binding site (PBS) was initially reported to be involved in MPMV RNA packaging.[30] More recent studies have shown that the core packaging determinants of MPMV RNA are present upstream of the mSD.[31] These findings have been controversial since Schmidt *et al.*[32] have suggested that the packaging signal of MPMV is composed of two regions both upstream (region "A": 30 nt) and downstream (region "B": 23 nt) of mSD as well as the first 120 nt of Gag (Fig. 1a).[32] Therefore, the packaging determinants of MPMV seem to be multipartite and more similar to complex retroviruses in that they reside in both 5′ UTR and Gag sequences as two discontinuous regions. However, the role of the 75-nt intervening

sequences between regions A and B has not been determined in MPMV RNA packaging. This study was undertaken to ascertain the exact boundaries of these two regions and their relative role as well as the contribution of the intervening sequences towards MPMV RNA packaging by employing genetic and structural prediction approaches to establish a structure–function relationship during RNA packaging process.

Results

The relative role of regions A and B towards MPMV RNA packaging

To determine the differential role of regions A and B towards MPMV RNA packaging, we generated a mutant, SJ5, containing deletion of all 5′ UTR

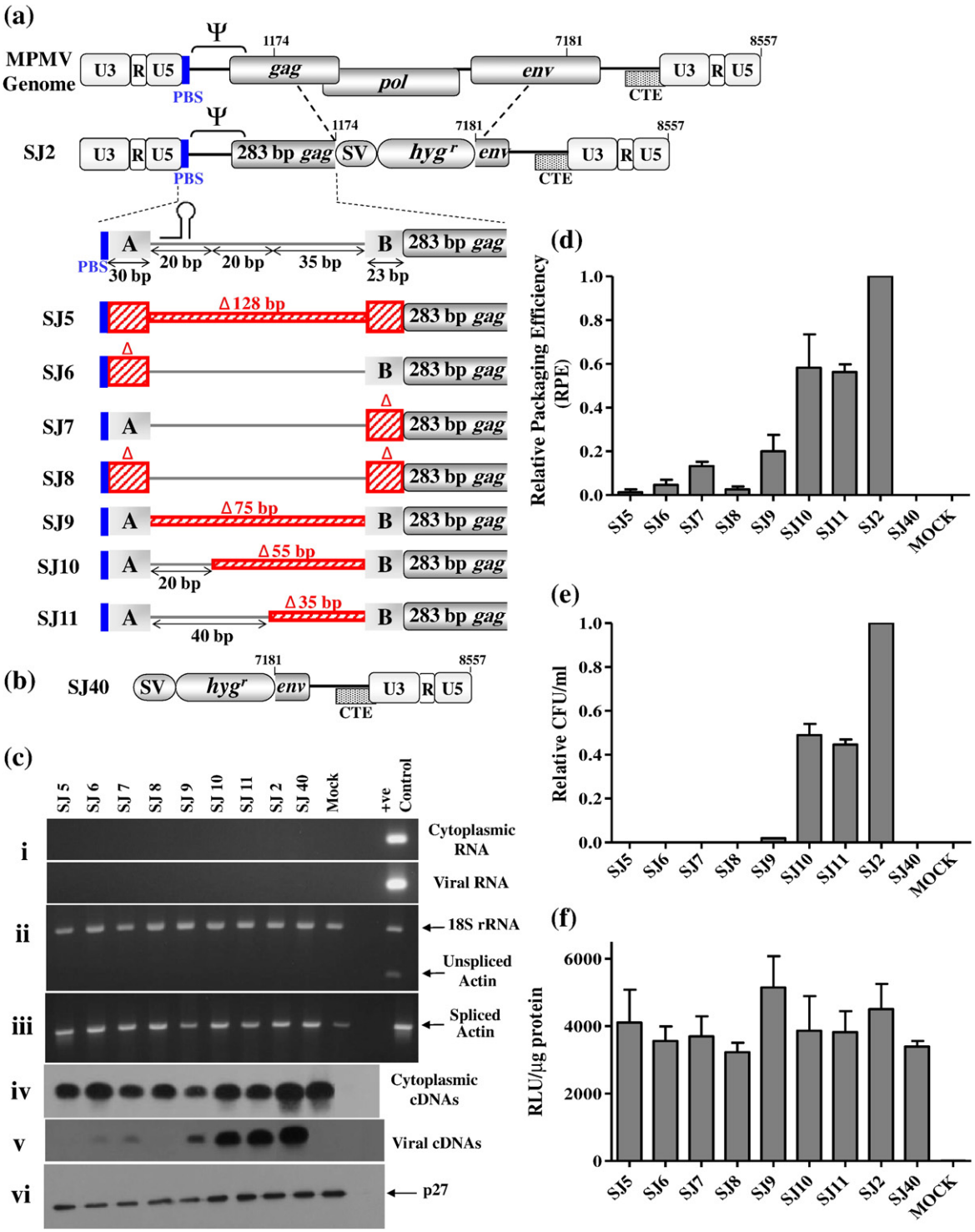

Fig. 1 *(legend on next page)*

sequences (128 nt) including regions A and B. In addition, we constructed three other mutant clones in which region A, region B, or both regions have been deleted, while maintaining the rest of the UTR sequences, resulting in SJ6, SJ7, and SJ8, respectively (Fig. 1a).

A previously described MPMV *in vivo* trans-complementation assay was employed to monitor mutant transfer vector RNA packaging and propagation.[32,33] Briefly, mutant transfer vectors, wild type (SJ2), a negative control vector, SJ40 (lacking known MPMV packaging sequences; Fig. 1b), and a control plasmid, pGL3 (to monitor for transfection efficiency), were co-transfected with the MPMV packaging construct (TR301) and the vesicular stomatitis virus glycoprotein G-based envelope expression plasmid (MD.G.) into 293T cells. The design of the transfer vector and packaging construct is primarily based on the principle that successful packaging requires the recognition of ψ present on the transfer vector RNA by the viral proteins expressed by the packaging construct. In the trans-complementation assay, the replication of the packaged RNA is limited to a single round, thus limiting reinfection of the target cells and, hence, curtailing the generation of replication-competent virus, making the assay reproducible, sensitive, and quantitative.

To ensure that the transfer vector RNA is efficiently and stably expressed and properly transported from the nucleus to the cytoplasm to be packaged, we fractionated RNAs from the transfected cells into cytoplasmic and nuclear fractions. Before cDNA preparations, cytoplasmic and viral RNAs were DNase treated and amplified using transfer vector-specific primers OTR216 and OTR217, and the lack of any amplification confirmed the absence of any detectable contaminating plasmid DNA in RNA preparations (Fig. 1c, i). To confirm that the nuclear membrane integrity was not compromised during cytoplasmic and nuclear RNA fractionation, we amplified cytoplasmic cDNA preparations as described previously[34] for the unspliced β-actin mRNA, which is exclusively found in the nuclear fraction, while the spliced form is found in both fractions.[35] As shown in Fig. 1c, ii, the lack of any amplifiable signal, following 30 cycles of PCR, for the unspliced β-actin message in cytoplasmic cDNA preparations confirmed that our fractionation technique was free of any artifact and that the nuclear membrane integrity was maintained during the fractionation process. To confirm the presence of amplifiable cDNA in the reactions, we conducted multiplex PCRs in the presence of primers/competimers for 18S ribosomal RNA. Amplification of 18S ribosomal RNA in the multiplex PCR (Fig. 1c, ii) and the presence of spliced β-actin mRNA in the cytoplasmic fractions (Fig. 1c, iii) confirmed that our cDNA preparations were amplifiable. Finally, cytoplasmic and viral cDNAs were amplified using vector-specific primers (OTR216 and OTR217), amplified products were electrophoresed on agarose gel, and Southern blotting was conducted (Fig. 1c, iv and v). The results presented in Fig. 1c, iv, confirmed that, following transfection, transfer vector RNAs were stably expressed and properly transported to the cytoplasm. Optical densities (ODs) obtained from the blots were used for calculating the relative packaging efficiencies (RPEs), which were determined by calculating the ratio of the packaged mutant RNA to the packaged wild-type RNA (SJ2) relative to the ratio of the two RNAs in the cytoplasm (detailed in Materials and Methods). The results of semiquantitative RT-PCR and the number of resulting hygromycin-resistant colony-forming units per milliliter (CFU/ml) in the transfected cultures from several independent experiments were used to determine the mean RPE and RNA propagation, respectively, for each mutant transfer vector RNA.

Our analyses revealed that the mutant SJ5, containing deletion of the entire UTR (including

Fig. 1. The role of the 3′ end of the intervening sequences between regions A and B in MPMV RNA packaging. (a) Schematic representations of the complete MPMV genome and wild-type transfer vector, SJ2, showing two major packaging determinants in the 5′ UTR labeled as regions A and B based on an earlier study.[32] Mutant transfer vectors SJ5–SJ8 contain deletions of the entire UTR, region A, region B, or both, respectively. Mutant transfer vector SJ9 contains deletion of all the intervening sequences, whereas SJ10 and SJ11 contain 55- and 35-nt deletions, respectively. Deleted regions are represented with hatched red boxes. (b) Schematic representation of the control transfer vector (SJ40) lacking viral sequences at its 5′ end to monitor nonspecific RNA packaging. (c) (i) Thirty cycle amplifications of the DNase-treated cytoplasmic RNA (upper panel) and viral RNA (lower panel) using vector-specific primers to exclude the possibility of any contaminating plasmid DNA. (ii and iii) Control for nucleocytoplasmic fractionation. Cytoplasmic cDNAs were tested for the absence of unspliced β-actin mRNA. Lack of any amplifiable signal confirmed that the cytoplasmic RNA fractions were devoid of any contaminating DNA (ii) whereas the spliced β-actin mRNA was detected in cytoplasmic cDNA (iii). Multiplex amplifications were conducted in the presence of primers/competimers for 18S ribosomal RNA as an ancillary control for the presence of amplifiable cDNA in the unspliced β-actin PCRs (ii). (iv and v) Representative Southern blots of the cytoplasmic (iv) and viral (v) cDNAs amplified using vector-specific primers in a semiquantitative RT-PCR that amplified 250-bp fragments. (vi) Representative western blot from one of the experiments that were used to calculate the RPE. (d) RPE of the mutant transfer vector RNAs. The RPE for each mutant was determined by calculating the ratio of the packaged mutant RNA to the packaged wild-type (SJ2) RNA relative to the ratio of the two RNAs in the cytoplasm after subtracting the background packaging observed in the control vector (SJ40). (e) Relative hygromycin-resistant (Hyg^r) CFU/ml for mutant transfer vectors reflecting the relative RNA propagation efficiencies. CFU/ml expressed for each mutant is relative to the wild type (SJ2) and was normalized to the luciferase expression observed in the transfected cultures. (f) Transfection efficiencies obtained for mutant and wild-type transfer vectors as described in Materials and Methods. The data represent the mean of at least three independent transfection and infection experiments. ψ, packaging signal; SV, simian virus 40 promoter; *hyg^r*, *hygromycin resistance* gene; CTE, constitutive transport element; RLU, relative light units per microgram of protein.

regions A and B), abrogated RNA packaging and consequently propagation (Fig. 1d and e) despite the fact that the transfection efficiency for this mutant as well as for the wild-type transfer vector was comparable and well within 2-fold (Fig. 1f). On the other hand, wild-type transfer vector SJ2 (containing intact UTR sequences) showed efficient RNA packaging and propagation (Fig. 1d and e), confirming the integrity of our packaging and transduction assay and further asserting its sensitive and quantitative nature. Furthermore, as shown in Fig. 1c, vi, western blot analysis revealed that similar amounts of virus particles were used to isolate viral RNA for both mutant and wild-type transfer vectors, which further authenticated our results. These observations confirmed that the UTR sequences (including regions A and B) are essential for RNA encapsidation. Deletion of either region (SJ6 and SJ7) or both regions (SJ8) severely impaired RNA packaging (RPE ranging from 0.03 to 0.13), resulting in 87–97% reduction when compared to wild type (SJ2; Fig. 1d). Consistent with the reduced level of RNA packaging of SJ6–SJ8, vector RNA propagation was also found to be ablated (Fig. 1e). The slight increase in RPE of SJ7 (deletion of region B; RPE = 0.13; $P = 0.0005$) when compared to SJ6 (deletion of region A; RPE = 0.04; $P = 0.0006$) suggested that region A is slightly more important than region B, but still not solely sufficient for optimal RNA packaging.

The packaging observed for mutant transfer vector RNAs was not due to nonspecific RNA packaging

To ensure that we did not observe any RNA packaging artifacts due to nonspecific RNA packaging through retrofection[36,37] or packaging possibly due to sequences other than the earlier suggested packaging sequences of MPMV, we tested the control transfer vector, SJ40, which lacked all the 5′ end viral sequences (Fig. 1b). RT-PCR of the RNA purified from pelleted virus particles using the same set of primers showed no detectable RNA packaging (or if packaged, the level of packaging is below the threshold level of detection; Fig. 1d). This was despite the stable expression of transfer vector RNA in the transfected cells (Fig. 1c, iv), its proper transport to the cytoplasm (Fig. 1c, ii and iii), and efficient viral particle production to package viral RNA (Fig. 1c, vi). The inability of SJ40 RNA to be packaged further confirmed the integrity of our *in vivo* packaging assay and excluded the possibility of any nonspecific RNA packaging.

Effects of the deletion of sequences at the 3′ end of the intervening sequences between regions A and B on MPMV RNA packaging

To investigate the role of the intervening sequences between regions A and B, we deleted these sequences in mutant transfer vector (SJ9) (Fig. 1a). In addition, two other mutants, SJ10 and SJ11, were created, which contained 55- and 35-nt deletions, respectively, at the 3′ end of the intervening sequences (Fig. 1a). Following genetic complementation of mutant transfer vectors (SJ9–SJ11) and after taking into account the appropriate controls (Fig. 1c and f), we calculated the effect of these mutations on RNA packaging and propagation. Data derived from several independent experiments revealed that the deletion of the entire intervening sequences between regions A and B in the case of SJ9 showed an RPE of 0.2 ($P = 0.009$), an 80% drop or 5-fold reduction in packaging efficiency when compared to the wild type (Fig. 1d). In contrast to the 5-fold reduction in packaging of SJ9, transfer vector RNA propagation was much more severely impaired (50-fold or 98% reduction) when compared to the wild type (Fig. 1e). Such a diminished RNA packaging in the case of SJ9 suggests that there are sequences within the intervening 75 nt that are essential for MPMV RNA packaging. On the other hand, deletion of 55 nt (SJ10) and 35 nt (SJ11) showed marginal reduction in RNA packaging and propagation when compared to the wild type, suggesting that the deleted sequences contribute less significantly towards MPMV RNA packaging (Fig. 1d and e).

Effects of heterologous insertions substituting deletions of sequences at the 3′ end of the intervening sequences between regions A and B on MPMV RNA packaging

To ascertain whether the intervening sequences function as spacer sequences providing stability to the higher-order structure of the RNA to be packaged, we employed an alternative approach in which the deleted regions were replaced with non-viral heterologous sequences. Mutant transfer vectors SJ5NH and SJ9NH contain a similar deletion as SJ5 (all UTR) and SJ9 (deletion of the intervening sequences only); however, the deleted sequences were substituted with heterologous sequences of equal length containing flanking NotI sites for the ease of cloning (Fig. 2a). In order to ensure that the introduction of the G–C-rich NotI site (5′ GCGGCCGC 3′) did not adversely affect RNA packaging, two other mutant transfer vectors, namely, SJ5N and SJ9N, were created as controls in which, following deletion, a NotI site was introduced (Fig. 2a).

Testing these mutants while taking into consideration all the appropriate controls (Fig. 2b and c) revealed that consistent with our earlier observations as in the case of SJ5 (deletion of all UTR), packaging of SJ5N and SJ5NH was also ablated (Fig. 2d), confirming the direct role of 5′ UTR in MPMV RNA packaging. On the other hand, SJ9N (deletion of the intervening sequences with an insertion of NotI site) was packaged at a much reduced level (RPE = 0.21; $P = 0.007$; Fig. 2d), which is comparable to SJ9 containing similar deletion as SJ9N (Fig. 1a and d), suggesting that the insertion of NotI site in SJ9N did not adversely affect RNA packaging. However, the substitution of the deleted 75 nt with a heterologous sequence in the case of SJ9NH drastically reduced packaging, showing an RPE of

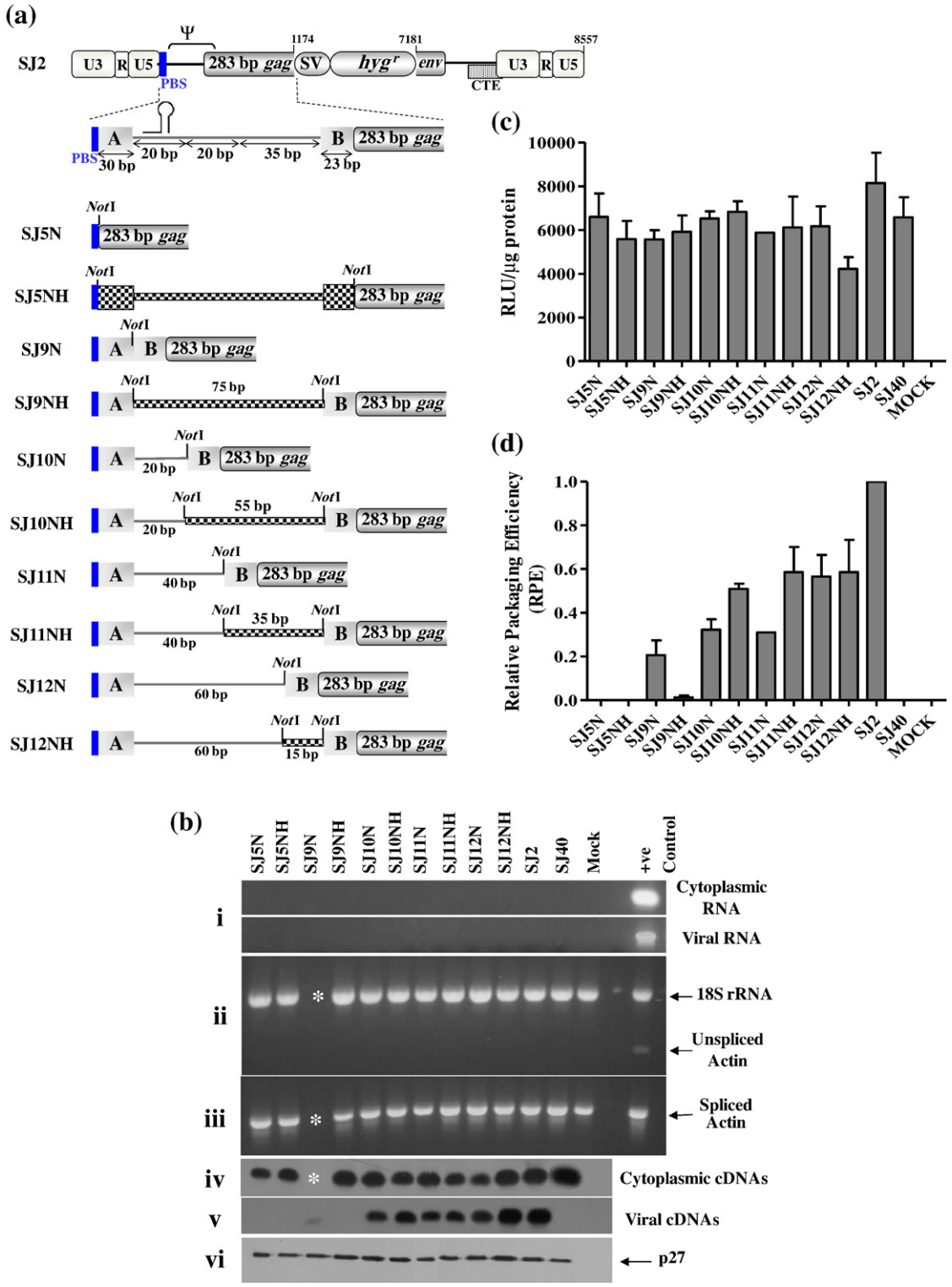

Fig. 2. Role of the deletion/substitution mutations in the 5′ end of MPMV genome in RNA packaging. (a) Schematic representations of SJ2, the wild type, and mutant transfer vectors. SJ5N and SJ9N–SJ12N contain deletions with an insertion of NotI site, whereas in SJ5NH and SJ9NH–SJ12NH, the deleted sequences have been substituted with heterologous sequences of equal length containing flanking NotI sites. Checkered boxes represent insertions. (b) Representative gels of the necessary controls (i–iii) as described in Fig. 1c. Representative Southern blots of the cytoplasmic (iv) and viral (v) cDNAs following RT-PCR and western blot (vi) of the pelleted viral particles. An asterisk (*) denotes the loss of cytoplasmic cDNA for mutant SJ9N in this particular experiment. (c and d) Mean of transfection efficiencies and relative RNA packaging efficiencies, respectively, from three independent transfection experiments except for RPE of mutant SJ11N. ψ, packaging signal; SV, simian virus 40 promoter; *hygr*, *hygromycin resistance* gene; CTE, constitutive transport element; RLU, relative light units per microgram of protein.

0.01 (Fig. 2d), which could be due to the insertion of heterologous sequences that negatively affected the overall structure of this region.

A similar approach was also employed to create mutant transfer vectors SJ10NH, SJ11NH, and SJ12NH in which deleted 55, 35, and 15 nt of the intervening sequences, respectively, were substituted by heterologous sequences of equal length containing flanking NotI sites along with their necessary NotI controls (SJ10N, SJ11N, and SJ12N; Fig. 2a). Deletion/substitution of 55, 35, and 15 nt in the case of SJ10NH, SJ11NH, and SJ12NH, respectively, only marginally affected RNA packaging. The RPEs for these mutants were calculated to be 0.51, 0.59, and 0.59, respectively, showing only less than 2-fold reduction in packaging when compared to the wild type (Fig. 2d). On the other hand, the deletion of 55, 35, and 15 nt in the case of SJ10N, SJ11N, and SJ12N with the insertion of NotI site showed RPEs of 0.32, 0.31, and 0.57, respectively (Fig. 2d). A slight decrease in RNA packaging of the mutants SJ10N (RPE = 0.32; $P = 0.0047$) and SJ11N (RPE = 0.31) containing NotI insertion when compared to SJ10 (RPE = 0.58) and SJ11 (RPE = 0.56) could be due to a possible interaction of the G–C-rich NotI site with neighboring sequences (discussed later). These results are in overall agreement with earlier observations of the 3′ end deletion mutants of the intervening sequences and further suggest that these sequences (up to 55 nt upstream of region B) do not contribute significantly in augmenting MPMV RNA packaging.

RNA secondary-structure predictions of the sequences involved in MPMV RNA packaging

Harrison *et al*. have earlier shown that the 5′ end of MPMV genomic RNA assumes a higher-order structure;[38] however, since the region used for predicting the RNA secondary structure did not include sequences starting from R, we could not rely on this structure to establish any possible structure–function correlation of the sequences involved in MPMV RNA packaging. Therefore, the region between R and the first 120 nt of Gag was folded using the RNA folding software "Mfold"[39,40] to correlate the effects of the introduced mutations on MPMV RNA packaging with the RNA secondary structure that this region may assume.

Analysis of RNA secondary-structure predictions showed that the 5′ end of MPMV genomic RNA folded into several stable stem loop (SL) structures, commonly found in the optimal and suboptimal predicted structures. As shown in Fig. 3, two stem loops (SL1 and SL2) were observed in the R/U5/PBS/UTR regions. SL1 comprises exclusively of R and U5 sequences, whereas SL2 is relatively large (96 nt) in which U5 sequences base pair with UTR sequences (including region A; 30 nt) and PBS folds into a small stem loop (PBS SL) that protrudes from SL2. Following SL2 (36 nt downstream of PBS) is a small 14-nt structural motif composed of G–C-rich palindromic (pal) sequence (5′ UCGCCGGCCGGCGA 3′) with 100% autocomplementarity and is referred to as pal SL (Fig. 3). Following pal SL is SL3 comprising exclusively of UTR sequences. Between pal SL and SL3 is a stretch of noticeable purine-rich, 16 single-stranded nucleotides (5′ UUAAAAGUGAAAGUAA 3′). Region B (23 nt), on the predicted structure, is located immediately downstream of SL3 of which 6 nt base pair with the lower part of SL3 at its 3′ end (Fig. 3). The Gag sequences folded into two distinct stem loops referred to as Gag SL1, which has a noticeable purine rich bulge, and a stable Gag SL2 with a four-uridine loop. In addition, Gag sequences both at the 5′ and at the 3′ sides of Gag SL1 are involved in LRI with U5 sequences (U5/Gag LRIs), as has been shown in complex retroviruses.[13,41]

In an attempt to compare the predicted structure with that of Harrison *et al.*,[38] we observed that, in their structure, a sequence within region A folds into a stem loop, whereas this region in the structure predicted in this study is part of a rather long SL2 in which region A base pairs with sequences in U5 extending the stem of SL2 (Fig. 3). In addition, the stretch of single-stranded purines (ssPurines) between pal SL and SL3 (Fig. 3) seems to fold into a less-defined stem loop in the earlier predicted structure with three A–U pairs at its stem. Another noticeable difference is the presence of U5/Gag LRI that could not be reported since R and U5 sequences were not included in their folding predictions.[38] Furthermore, pal SL also formed in their structure but its G–C-rich palindromic nature was not reported and correlated with its possible functional attributes towards viral replication. This could be due to the fact that only recent reports have drawn attention towards the role of pal sequences in RNA packaging and dimerization.[4,16–18,42] In order to further substantiate our structural predictions, we folded the sequences between R and 120 nt of Gag while employing the biochemical constraints published by Harrison *et al.*[38] The predicted structures with and without biochemical constraints folded almost similarly (data not shown), suggesting that our predicted structure is stable and is likely to assume during RNA packaging.

Effects of deletions and/or substitutions of the sequences at the 3′ end of the intervening sequences between regions A and B on the predicted RNA secondary structures

In order to correlate the results of packaging efficiency to the predicted structural motif(s) to establish structure–function relationships, we folded the 5′ end sequences of the mutant transfer vectors and compared them to the predicted RNA secondary structure of the wild type (SJ2). Structure prediction analyses of the mutants (SJ5–SJ9) containing deletions of either region A or B, both, or the intervening sequences showed major overall destabilization of the structure with the loss of the major structural motifs such as SL2, PBS SL, pal SL, and Gag SL1, which corresponded to the diminished or complete abrogation of RNA packaging (Fig. 4a–c).

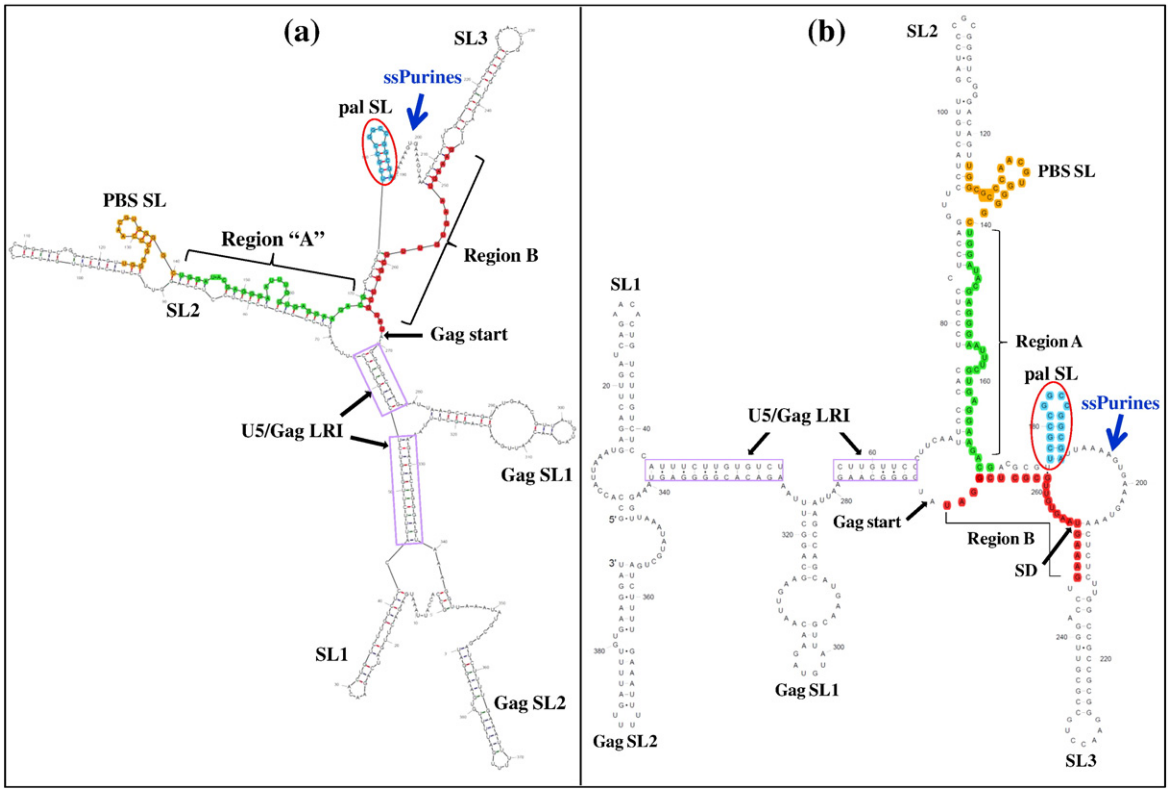

Fig. 3. RNA secondary structural model of the 5′ end of MPMV RNA genome. The region used for folding predictions included sequences from R up to 120 nt of Gag. (a) Structure predicted by Mfold, which uses free-energy minimization algorithms ($\Delta G = -124.70$ kcal/mol) showing five conserved stem loops (SLs) and two possible U5/Gag LRI regions, a 14-nt G–C-rich palindromic (pal) sequence, which acquires a stem loop motif labeled as pal SL, and a noticeable ssPurine-rich sequence. (b) The Mfold predicted structure redrawn using XRNA program. Sequences in orange, green, red, and blue represent the primer binding site (PBS), region A, region B, and pal sequences, respectively. Boxed area in purple show the predicted LRIs between U5 and Gag. SD, splice donor.

On the other hand, structure prediction analyses of SJ10 and SJ11 in which 55 and 35 nt were deleted from the 3′ end of the intervening sequences showed that the overall wild-type structure was maintained with the loss of only SL3 due to the deletion of its sequence (Figs. 4d and 5b). Such a modest alteration in the RNA secondary-structure predictions of these mutants is in agreement with marginal effects (less than 2-fold) on packaging data (Figs. 4d and 5b).

Careful analyses of mutants' structural predictions suggested that, in addition to regions A and B,[32] a 14-nt G–C-rich pal sequence is also needed for MPMV RNA packaging, since the maintenance of pal SL (while maintaining regions A and B) increased the packaging efficiency considerably in the case of SJ10 and SJ11 compared to SJ9, in which pal sequences were deleted (Figs. 1d, 4d, and 5b). Furthermore, data presented here also suggest that SL2 and Gag SL1 may also have a role in MPMV RNA packaging because mutants (SJ8 and SJ9) that exhibited changes in either of these structural motif(s) resulted in a decreased RNA packaging efficiency (Fig. 4b and c).

Structure prediction analyses of the deletion/substitution mutant clones revealed a similar observation that we obtained for the earlier series of clones SJ5–SJ11 with a noticeable exception in the case of SJ10N and SJ11N. The structure predictions of SJ10N and SJ11N (which contain insertion of NotI site following deletion of 55 and 35 nt, respectively, at the 3′ end of the intervening sequences) showed the loss of the native pal SL, due to the interaction of the G–C-rich NotI sequence with the G–C-rich pal sequence resulting in a chimeric stem loop structure (Fig. 5c; data not shown for SJ11N). Besides this chimeric SL, the overall wild-type structure was maintained except for the absence of structures that are likely to be formed by the deleted sequences. The loss of the native pal SL in the case of SJ10N and SJ11N could explain the reduced RPE (0.32 and 0.31, respectively; Figs. 2d and 5c) when compared to SJ10 and SJ11 (RPEs = 0.58 and 0.56), which did not contain the NotI insertion, thus maintaining the native pal SL motif (Figs. 4d and 5b). In SJ10NH, SJ11NH, and SJ12NH, the substitution of 55, 35, and 15 nt with heterologous sequences containing flanking NotI sites (NotI sites paired forming an alternate structure involving heterologous sequences) maintained native pal SL (Fig. 5d) and did not affect the overall structure as well as packaging (RPE = 0.51, 0.59, and 0.59, respectively; Fig. 2d). Consistent with structure predictions, a comparable RPE was observed in the case of SJ10, SJ10NH, SJ11 and SJ11NH, and SJ12N

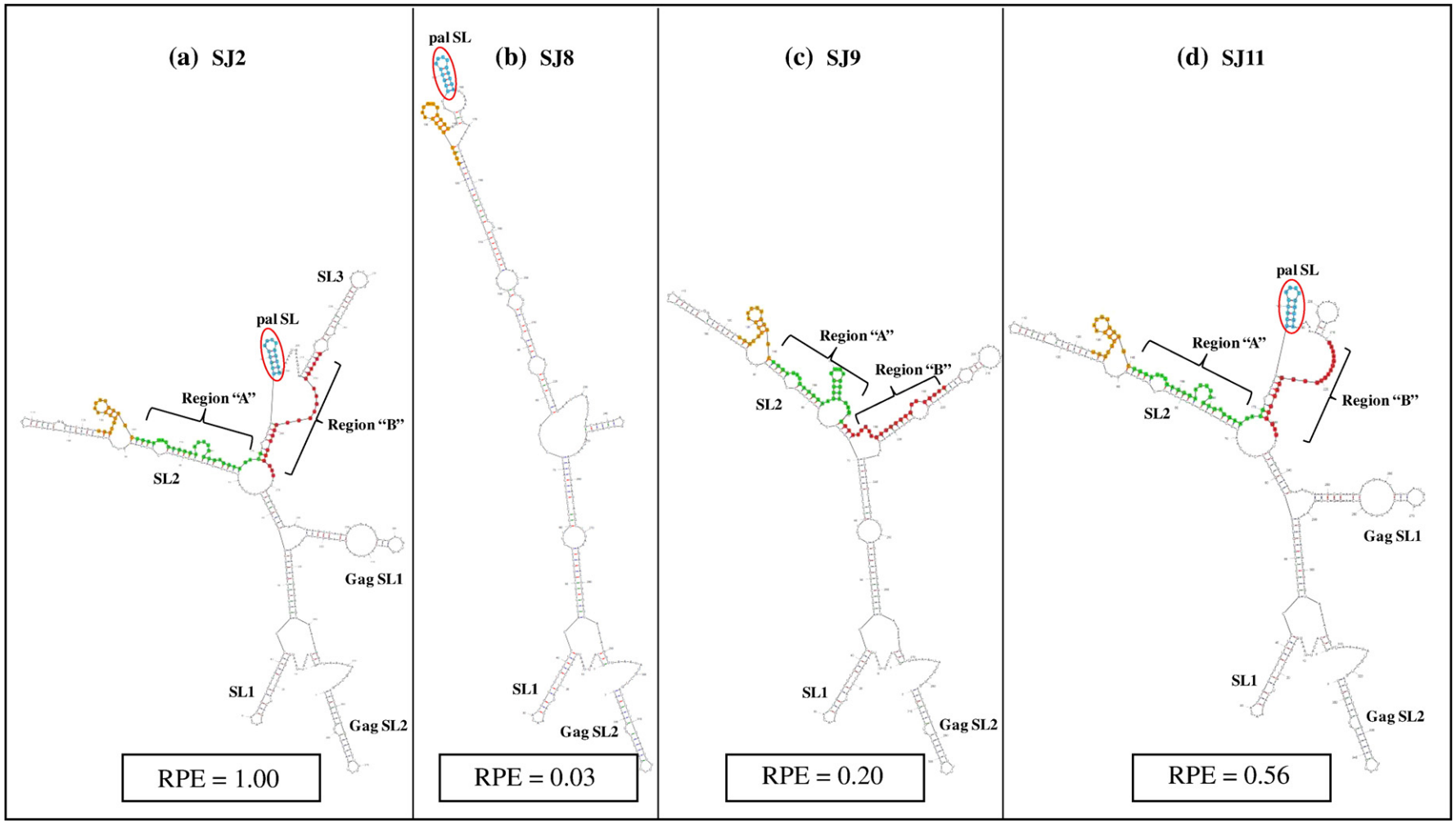

Fig. 4. Mfold structural predictions of the 5′ end MPMV RNA genome containing deletions at the 3′ end of the intervening sequences. (a) Wild type, SJ2, $\Delta G = -124.70$ kcal/mol. (b) SJ8 containing deletions of both regions A and B, $\Delta G = -104.30$ kcal/mol. (c) SJ9 with deletion of all the intervening sequences (75 nt) between regions A and B, $\Delta G = -94.30$ kcal/mol. (d) SJ11 containing 35-nt deletion at the 3′ end of the intervening sequence between regions A and B, $\Delta G = -106.90$ kcal/mol. Note the correlation of the restoration of the overall RNA secondary structure in SJ11, which is comparable to wild type (a), except for the deleted region, with the significant increase in RNA packaging efficiency. Orange, green, red, and blue colors highlight the primer binding site (PBS), region A, region B, and palindrome (pal) sequences, respectively.

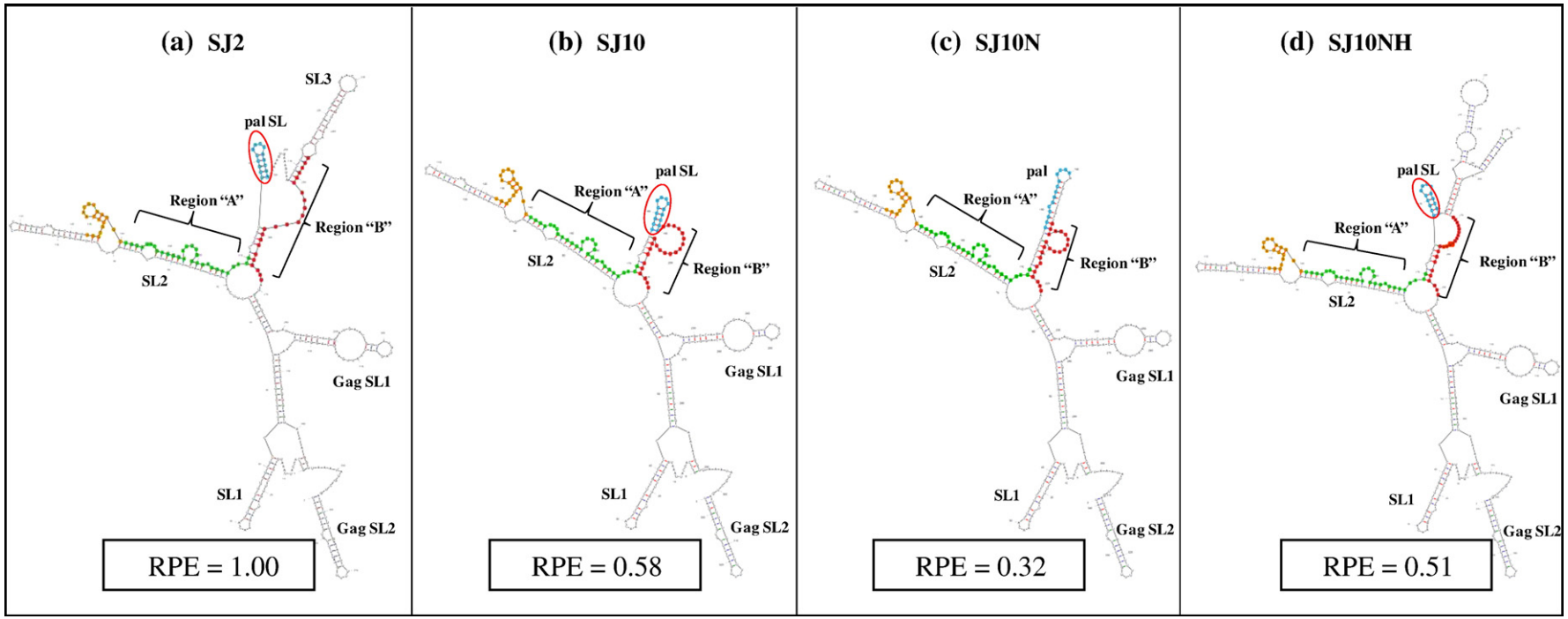

Fig. 5. Mfold structural predictions of the 5′ end MPMV RNA genome with deletion/substitution mutants showing the maintenance, loss, and/or reemergence of palindromic stem loop (pal SL). (a) Wild type, SJ2, $\Delta G = -124.70$ kcal/mol. (b) SJ10 containing 55-nt deletion at the 3′ end of the intervening sequences between region A and region B, $\Delta G = -105.50$ kcal/mol. (c) SJ10N containing the similar 55-nt deletion as in the case of SJ10, but replacing deletion with the insertion of a NotI site, $\Delta G = -112.90$ kcal/mol. (d) SJ10NH containing substitution of the deleted 55 nt with heterologous sequences containing flanking NotI sites, $\Delta G = -128.80$ kcal/mol. Note the loss of native pal stem loop (pal SL) in SJ10N (c) and appearance of an alternate chimeric SL due to the interaction of G–C-rich pal and NotI sequence, lowering the RNA packaging efficiency in SJ10N, whereas the reemergence of native pal SL in SJ10NH (d) restored the packaging efficiency to the level of the mutant clone SJ10 (b). Orange, green, red, and blue colors represent the primer binding site (PBS), region A, region B, and palindrome (pal) sequences, respectively.

and SJ12NH (Figs. 1d and 2d; data not shown for all structure predictions).

These results suggest that the intervening sequences at the 5′ end, which are involved in forming pal SL, play a significant role in MPMV RNA packaging at the structural level. This assertion is based on the fact that the loss of pal SL in the case of SJ10N (Fig. 5c) and SJ11N (data not shown) while maintaining its sequence reduced packaging further compared to the deletion/substitution clones SJ10, SJ11, and SJ10NH (Figs. 4d and 5b and d) and SJ11NH, SJ12N, and SJ12NH (data not shown), in which pal SL was maintained. In addition to the role of pal SL, our results suggest that the 3′ end of the intervening sequences does not contribute critically in RNA packaging since the deletion of these sequences (55, 35, and 15 nt) did not abrogate RNA packaging.

The role of the sequences at the 5′ end of the intervening sequences between regions A and B in MPMV RNA packaging

To further investigate the role of the sequences at the 5′ end of the intervening sequences (just downstream of region A) towards MPMV RNA packaging, we created a series of mutant transfer vectors containing deletions at the 5′ end with and without deletions at the 3′ end (Fig. 6a). These mutants were tested in the *in vivo* packaging assay, and the RPEs of these mutant transfer vectors were calculated after taking into consideration all the necessary controls (Fig. 6b and e). Mutant transfer vectors SJ42 and SJ43 containing deletion of 35 and 55 nt (including deletion of pal) from the 5′ end of the intervening sequences, respectively, showed a marked reduction in RNA packaging (RPE ranging from 0.14 to 0.11; $P<0.001$), representing 7- to 9-fold (86–89%) reduction when compared to the wild type (Fig. 6c). Consistent with the RNA packaging results, vector RNA propagation data of these mutants showed correspondingly 9- to 25-fold (89–96%) reduction in vector RNA propagation (Fig. 6d). These results are in support of our earlier assumption that sequences and/or structural motif(s) at the 5′ end of the intervening sequences have more significant role in augmenting MPMV RNA packaging compared to the 3′ end intervening sequences.

To further analyze the role of sequences between pal and the sequences involved in base pairing in SL3 at its 3′ end, we created a series of mutant transfer vectors. SJ44 contains a deletion of 20 nt (including the 16 ssPurines followed by pal SL), while SJ47 contains a double mutation, a deletion of 15 nt just upstream of region B in addition to SJ44 deletion; thus, SJ47 maintained both pal and 20 nt that form the 5′ side of SL3 (Fig. 6a). Finally, SJ46 contains deletions in the intervening sequences between regions A and B both at the 5′ (deletion of pal and the following 20 nt) and deletion of 15 nt just upstream of region B (Fig. 6a). Test of these mutants revealed that SJ44 showed an RPE of 0.55 (Fig. 6c; less than 2-fold reduction compared to the wild type) that is similar to the one observed for 3′ end deletion/insertion mutants (Fig. 1d). Corroborating with the RNA packaging data, the propagation was also reduced by less than 2-fold (49% reduction) when compared to the wild type (Fig. 6d). SJ46 showed an RPE of 0.04 ($P=0.0001$) that is 96% reduction in packaging, whereas 98% reduction in propagation was observed when compared to the wild type (Fig. 6c and d). The inclusion of pal sequence in the case of SJ47, while deleting 15 nt upstream of region B, increased packaging significantly (compared to SJ46) and showed an RPE of 0.21 (5-fold reduction; $P=0.0009$) when compared to wild type (SJ2) (Fig. 6c). Although SJ47 maintains pal sequence, the reduced packaging efficiency of this mutant was still lower compared to SJ44 and such a drop in packaging could possibly be explained either by structural analysis (discussed later) or due to the mere fact that it had double mutations. The propagation of SJ47 was found to be reduced further (23-fold compared to 5-fold reduction; Fig. 6d), suggesting that the double mutation may have also affected post-packaging steps in virus replication such as reverse transcription and integration since our RNA propagation read-out assay is dependent on successful completion of these steps.

Taken together, these results suggest that the 5′ end of the sequences intervening between regions A and B (especially pal) plays a more important role in augmenting RNA packaging than sequences at the 3′ end, since deletion of pal (SJ42, SJ43, and SJ46) significantly compromised RNA packaging than mutants containing deletions at the 3′ end, which showed only a modest decrease in RNA packaging, further substantiating the importance of the G–C-rich pal in MPMV RNA packaging.

Effects of deletions of the sequences at the 5′ end of the intervening sequences between regions A and B on the predicted RNA secondary structures

Structure prediction analysis of the 5′ end deletion mutants (SJ42 and SJ43) containing deletion of pal destabilized the overall RNA secondary structure (Fig. 7b; data for SJ42 are not shown). A careful examination of the destabilized RNA secondary-structure predictions revealed that the reduced packaging efficiency in the case of SJ42 (RPE = 0.11) and SJ43 (RPE = 0.14) was due to the loss of pal SL.

The structure prediction analysis of SJ44, which contained 20 nt deletion (including the 16 ssPurines after pal SL) in the middle of the intervening sequences while maintaining 5′ and 3′ end sequences (pal SL and SL3), showed the preservation of almost all the wild-type structural motifs (SL2, PBS SL, pal SL, SL3, and Gag SL1; Fig. 7c). Corroborating well with the structural predictions, the packaging efficiency of SJ44 was only reduced by less than 2-fold (RPE = 0.55), suggesting the importance of these structural elements in addition to pal SL in MPMV RNA packaging. Structure

prediction of SJ44 revealed an interesting phenomenon: the deletion of ssPurines was compensated by its repeat in region B, which became single stranded and presumably contributed towards its efficient RNA packaging (Figs. 7c and 8f), a pattern similarly observed in SJ10 and SJ10NH (Fig. 8d and e). Finally, RNA structure prediction analysis of SJ46 containing double deletions (deletion of 5′ end sequences while maintaining the deletion at the 3′ end) revealed destabilization of the overall structure and consequently reduced packaging by 25-fold when compared to the wild type (Fig. 6c and structure data not shown). In summary, the well-corroborated structural predictions with the packaging data further suggest that sequences at the 3′ end of the intervening region that are involved in

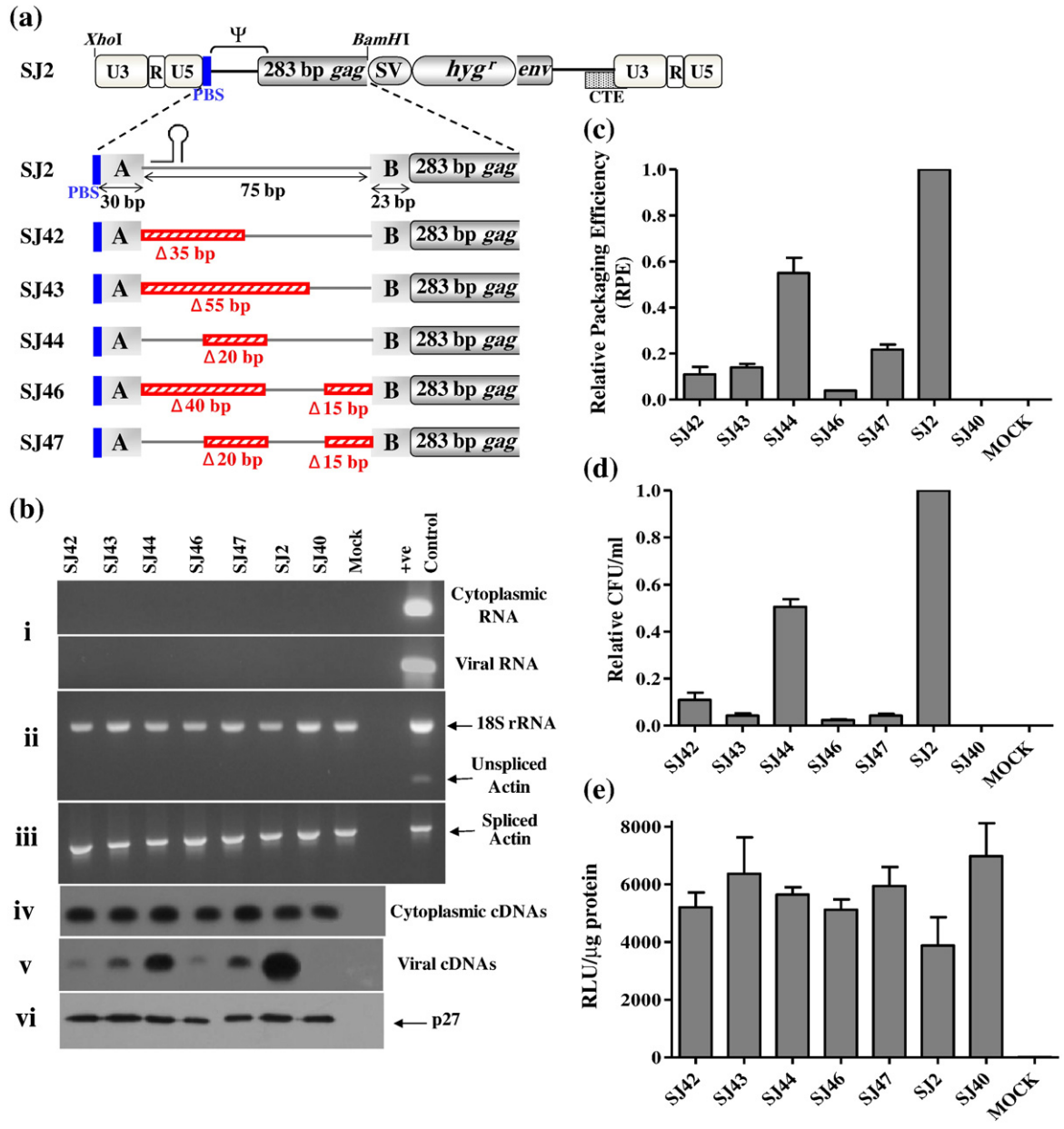

Fig. 6. The role of the 5′ end of the intervening sequences between regions A and B in MPMV RNA packaging. (a) Schematic representations of SJ2, wild type, and mutant transfer vectors. SJ42 and SJ43 contain 35- and 55-nt deletions, respectively, starting from the 5′ end of the intervening sequences. Mutant transfer vectors SJ44 contains deletion of 20 nt in the middle of the intervening sequences, and SJ46 contains 40-nt deletion at the 5′ end and 15-nt deletion at the 3′ end of the intervening sequences. SJ47 contains the same 20-nt deletion as in SJ44 and in addition contains the same 15-nt deletion as in SJ46. Deleted regions are represented with hatched red boxes. (b) Representative gels of the necessary controls (i–iii) as described in Fig. 1c. Representative Southern blots of the cytoplasmic (iv) and viral (v) cDNAs following RT-PCR and western blot (vi) of the pelleted viral particles. (c–e) The relative RNA packaging, RNA propagation, and mean transfection efficiencies, respectively, from three independent experiments. ψ, packaging signal; SV, simian virus 40 promoter; *hygr*, *hygromycin resistance* gene; CTE, constitutive transport element; RLU, relative light units per microgram of protein.

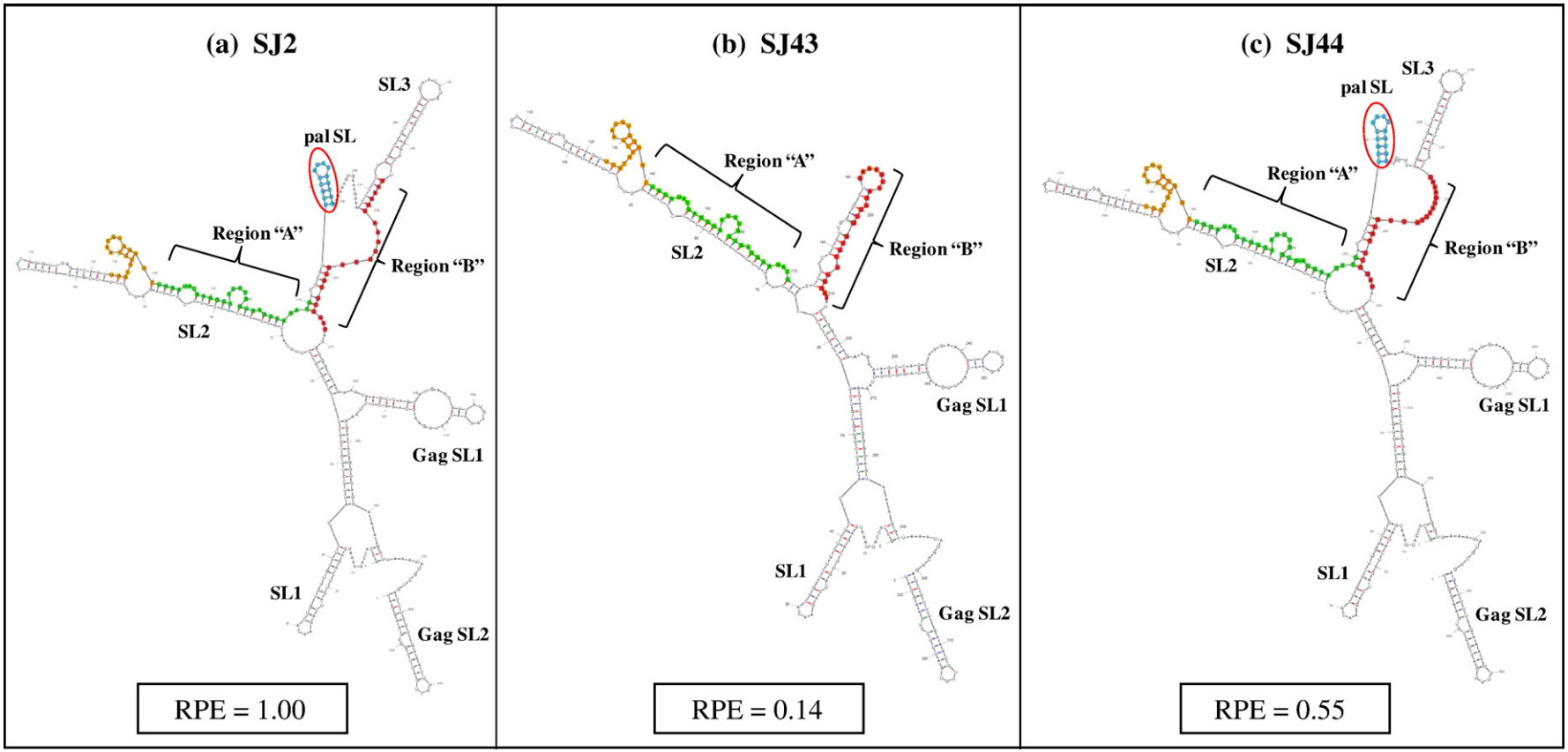

Fig. 7. Mfold structural predictions of the 5′ end of MPMV RNA genome illustrating that part of the 5′ end of the intervening sequence between regions A and B is needed to maintain the overall RNA secondary structure for optimum RNA packaging. (a) Wild type, SJ2, $\Delta G = -124.70$ kcal/mol. (b) Mutant clone SJ43 containing a 55-nt deletion 3′ to region A, $\Delta G = -98.90$ kcal/mol. (c) SJ44 containing 20-nt deletion in the middle of the intervening sequence between regions A and B, $\Delta G = -121.60$ kcal/mol. Note the reemergence of native pal stem loop (pal SL) in SJ44 (c) when only 20-nt downstream region A were maintained, which increased packaging significantly compared to SJ43 (b) containing deletion of pal sequence. The primer binding site (PBS), region A, region B, and palindrome (pal) sequences are represented by the orange, green, red, and blue colors, respectively.

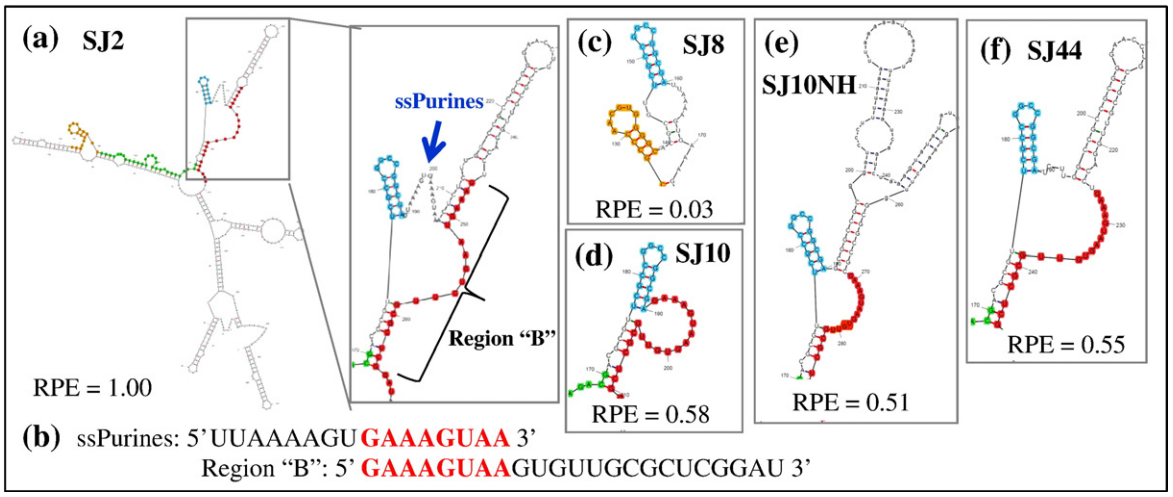

Fig. 8. Role of ssPurines in MPMV RNA packaging. (a) Mfold structural predictions of the 5′ end of MPMV RNA genome showing that a stretch of ssPurine-rich sequence (enlarged area) is needed to maintain the overall RNA secondary structure and optimal RNA packaging. (b) A mirror image of 8 nt of the ssPurine-rich sequence is present in region B. (c) SJ8 containing deletions of both regions A and B, which forced the ssPurines to base pair with other regions in the predicted structure, abrogated RNA packaging (for full structure prediction of SJ8, please see Fig. 4b). (d) SJ10 containing 55-nt deletion at the 3′ end of the intervening region (including the ssPurines) forced the 8-nt mirror image in region B to become single stranded in order to increase RNA packaging significantly. Similar phenomenon of region B purines becoming single stranded was observed to increase RNA packaging significantly in several other mutants, for example, SJ10NH (e) and SJ44 (f) both containing the deletion of ssPurines. The primer binding site (PBS), region A, region B, and palindrome (pal) sequences are represented by the orange, green, red, and blue colors, respectively.

base pairing in SL3 contribute less significantly; on the other hand, these analyses suggest an eminent role of a G–C-rich pal SL at the 5′ end (downstream of region A) in MPMV RNA packaging at the structural level.

Discussion

In an attempt to enhance our understanding of MPMV RNA packaging process, a combination of genetic and structural prediction approaches, which offered additional refinement of the cis-acting encapsidation sequences of MPMV and validated that the MPMV packaging determinants consist of two discontinuous regions both upstream and downstream of mSD, was employed. Based on mutational and structural prediction analyses, we propose that boundaries of region A should be further extended at its 3′ end and should comprise the first 50 nt of UTR (to include a G–C-rich pal sequence) as opposed to the 30 nt suggested earlier.[32] Such discrepancy between the two studies could be due to the less sensitive technique of slot blot that was employed in the earlier study.[32] The intervening sequences between the newly defined boundaries of regions A and B within 5′ UTR do not contribute significantly towards MPMV RNA packaging. RNA structural predictions of the 5′ end (from R up to 120 nt of Gag) of the MPMV wild-type genomic RNA revealed several structural motifs including stem loops (SLs), ssPurines, and U5/Gag LRIs (Fig. 3). Among the predicted structural motifs, SL2 and pal SL were found to be most important for MPMV RNA packaging since mutants containing deletions and/or alterations of the sequences involving some of these structures influenced RNA packaging significantly. In addition, the presence of ssPurines was found to be directly correlated with optimal MPMV RNA packaging.

The deletion of region A in SJ6 and region B in SJ7 reduced packaging by 25- and 7-fold, respectively (Fig. 1d). Such difference in RNA packaging reflects the importance of region A over region B. This could be due to the fact that region A is longer than region B and that region A sequences are part of a distinct SL2 where it base pairs with U5 sequences extending its stem, possibly providing stability to the overall predicted RNA secondary structure (Fig. 3).

The diminished RNA packaging in SJ7 (Fig. 1d) when region B was deleted confirmed the importance of this region in MPMV RNA encapsidation. The presence of this region in unspliced MPMV mRNA, where it overlaps with the mSD, should favor the preferential packaging of unspliced viral genomic mRNAs over spliced viral mRNAs. The contribution of region B towards RNA packaging could also be attributed to the presence of an 8-nt purine-rich sequence (5′ GAAAGUAA 3′). The presence of a stretch of purines on retroviral genomic RNA has been proposed to be a potential NC binding site, thereby facilitating its efficient encapsidation.[10,42–44] Interestingly, our structural predictions have also revealed a prominent stretch of ssPurines (5′ UUAAAAGUGAAAGUAA 3′) between pal SL and SL3 (Figs. 3 and 8a) containing a repeat sequence (underlined) of the purines observed in region B (Fig. 8b). The deletion of these ssPurines in the case of SJ10 showed an interesting folding pattern in which the stretch of

purines in region B became single stranded (Figs. 5b and 8d) similar to the ssPurines in the wild type while maintaining the overall RNA secondary structure except for the loss of the deleted SL3, which correlated well with the marginal effect on RNA packaging (Figs. 5b and 8d). Similar observation pertaining to the emergence of alternate ssPurines was also observed in the case of SJ44 (Figs. 7c and 8f) in which only the ssPurines were deleted, in addition to several other mutants, for example, SJ10 and SJ10NH (Figs. 5b and d and 8d and e). Since the stretch of purines in region B compensated for the conformational structure that would have been lost with the deletion of the ssPurines, it is reasonable to hypothesize that a stretch of native or alternate ssPurines in MPMV function at the structural level in mediating RNA packaging possibly as a potential NC binding site as has been proposed for some other retroviruses. Deletion of sequences just upstream of region B (SJ10 and SJ11; Figs. 5b and 4d) or deletion of 12 nt of region B downstream of the purine-rich sequence in an earlier study did not affect RNA packaging significantly,[31] which further attributes the importance of the primary sequence of region B (especially purines) towards MPMV RNA encapsidation.

Our structure–function analysis suggests that in addition to the essential roles of regions A and B and the presence of a stretch of ssPurines, a 14-nt G–C-rich pal sequence (5′ UCGCCGGCCGGCGA 3′) is important for optimum RNA packaging, since its deletion in SJ9, SJ9N, SJ9NH, SJ42, and SJ43 ablated RNA packaging (Figs. 1d, 2d, and 6c). Consistent with this, in an earlier study, without identifying its palindromic nature, Mustafa *et al.*[45] introduced mutations in pal SL resulting in the maintenance of the palindromic nature, which showed minimal effect on RNA packaging.[45] Their finding further argues for the importance of its palindromic nature rather than its primary sequence towards MPMV RNA packaging. In addition, a deletion at the 5′ end of MPMV genome that included the last 3 nt of pal (comparable to SJ10) resulting in an alternate 14-nt pal (with only one G–C to G–U change)[31] showed packaging defect comparable to SJ10 (less than two fold; Fig. 1a and d) further validating our observations and the accuracy of our assay.

Since the identified 14-nt G–C-rich pal sequence assumes a stand-alone structural motif (pal SL), it is reasonable to propose that the pal sequence also functions at the structural level. Such an assumption is substantiated by the efficient RNA packaging of mutants in which pal SL was maintained (Figs. 4d, 5b and d, and 7c), whereas the deletion of pal in SJ42 and SJ43 (Figs. 6c and 7b) diminished RNA packaging. The reduced packaging efficiency associated with the loss of pal SL due to its interaction with the neighboring G–C-rich NotI site (5′ GCGGCCGC 3′) in SJ10N and SJ11N (Fig. 5; data not shown for SJ11N) or with a nearby G–C-rich sequence in SJ47 (data not shown) further advocates the importance of pal SL for efficient RNA packaging. However, the functional contribution of pal SL towards MPMV RNA packaging seems to be only in the presence of the other two core packaging determinants in the UTR (regions A and B) since the presence of pal sequence with or without the formation of pal SL in mutants that had deletions in regions A or B or both (SJ6, SJ7, and SJ8) showed either abrogated or diminished RNA packaging (Figs. 1d and 4b). The inability of pal to promote packaging by itself argues for its indirect role in RNA packaging and a possible role in other steps of retroviral life cycle such as RNA dimerization. Furthermore, sequences that are involved in RNA dimerization are often intermingled with the sequences that are responsible in augmenting RNA packaging[2,4] and pal sequences in a number of retroviruses have been identified as DISs. For example, a 6-nt G–C-rich pal sequence has been found to be phylogenetically conserved in more than 50 isolates of HIV-1, HIV-2, and SIV.[4] In addition, a 10-nt pal in the UTR is found to be phylogenetically conserved in HIV-2, macaque, and sooty mangabey SIVs,[46] which modulates steps in viral replication including RNA packaging and dimerization.[16–18] Hence, it is conceivable to propose that MPMV pal is a potential DIS. In a recent study, deletion of HIV-1 SL1, which contains 6-nt pal DIS (5′ GCGCGC 3′), while maintaining the core packaging determinants, surprisingly reduced packaging by 4- to 5-fold.[47] Such a scenario in RNA packaging is comparable to that of SJ9 (Figs. 1a and d and 4c), where the deletion of pal when regions A and B were maintained, reduced packaging by 5-fold as well. Reduction in packaging observed by Houzet *et al.*[47] by deleting HIV-1 SL1 suggests that intermolecular interactions between two HIV-1 genomic mRNAs are a prerequisite for RNA packaging, as has also been proposed for HIV-2.[48,49] Therefore, it is plausible to hypothesize that MPMV RNA packaging might be dependent on the presence of a dimeric genome, consistent at least with some retroviruses.[1,50] Although we did not mutate pal sequence, the structure–function analysis presented in this study has suggested the possible role of pal in RNA packaging and possibly dimerization where pal SL potentially functions as DIS. Therefore, it would be interesting to investigate the effects of pal mutations on MPMV RNA dimerization and packaging, which may differentiate the relative role of pal on either of these processes.

Our predicted structure also revealed a stem loop structure labeled as SL3 (Fig. 3); a similar stem loop was also predicted by Harrison *et al.*[38] Based on SL3's considerable G–C-rich stem, its conservation in type D retroviruses, and the presence of a GAYC (Y stands for pyrimidine) motif, which is conserved across many distantly related retroviruses,[31,38,51–54] Harrison *et al.* suggested its seminal role in MPMV encapsidation. However, deletion mutants in which the region encompassing SL3 (SJ10, SJ10N, SJ10NH, SJ11, SJ11N, and SJ11NH) was deleted did not affect RNA packaging significantly (Figs. 1d and 2d). Consistent with this, destabilizing and compensatory

substitution mutational analysis of SL3 and deletions and substitution mutations in the MPMV ACC motif by Mustafa *et al.*[45] did not affect RNA packaging considerably. Taken together, it is reasonable to propose that contrary to the suggestion of Harrison *et al.*,[38] the role of SL3 in MPMV RNA packaging is minimal.

Deletion and/or substitution mutation and structure prediction analyses presented here confirmed that the 5′ end of MPMV RNA genome assumes a higher-order structure and identified structural motifs that are needed for its efficient RNA encapsidation. The results presented in this study should provide a better understanding of MPMV RNA packaging process and the role of structural motifs in relation to RNA–RNA and RNA–protein interactions that take place during MPMV replication. A further refinement of these steps in MPMV life cycle is essential if MPMV-based vectors were to be used for human gene therapy.

Materials and Methods

MPMV Gag-Pol expression plasmid (packaging construct), envelope expression plasmid, and MPMV transfer vectors

The MPMV packaging construct, TR301, which expresses Gag-Pol proteins and the vesicular stomatitis virus glycoprotein G envelope expression plasmid, MD.G, has been described earlier.[33,55] The overall design of SJ2, a subgenomic MPMV transfer vector, is similar to KAL011[33] and contains necessary cis-acting sequences needed for genome replication including transcription, polyadenylation, encapsidation, reverse transcription, and integration (Fig. 1a). MPMV nucleotide positions refer to GenBank accession number M12349.[56]

In order to study the role of the intervening sequences between regions A and B (Fig. 1a), we generated a series of deletion mutants (SJ5–SJ11) through splice overlap extension PCR[57] using SJ2 as the template as described earlier.[32,45] The final amplified product for each introduced mutation was digested with XhoI and BamHI and cloned into SJ2 that has been previously cleaved with the same restriction enzymes to replace the wild-type sequences with sequences containing mutations (Fig. 1a).

To ascertain the role of the deleted sequences between regions A and B, we substituted the deleted sequences with heterologous sequences in order to delineate whether these regions are needed for RNA packaging or function as spacer to provide the overall stability to the RNA secondary structure of this region. For the convenience of cloning, an artificial NotI site was introduced at the junction of the deleted sequences resulting in clones SJ5N and SJ9N–SJ12N (Fig. 2a). To introduce heterologous sequences, we amplified a region of pcDNA3 that was equal in size to the deleted sequence and contained flanking NotI sites. Amplified products were cleaved with NotI and cloned into the similarly digested SJ5N and SJ9N–SJ12N to generate final clones SJ5NH and SJ9NH–SJ12NH, respectively (Fig. 2a).

Mutant transfer vectors containing deletions at the 5′ end of the intervening sequences consist of two groups. One group contains incremental deletions starting from the 5′ end of the intervening sequences between regions A and B (SJ42 and SJ43; Fig. 6a). The other group contains, in addition to 5′ end deletions, deletions at the 3′ end as well as in the middle of the intervening sequences (Fig. 6a). Once again, the final amplified products containing desired mutations following splice overlap extension PCR were digested with XhoI and BamHI and the fragments containing the desired mutations were cloned into SJ2 to obtain the final clones that were referred to as SJ42, SJ43, SJ44, SJ46, and SJ47.

Control transfer vector to monitor nonspecific RNA packaging

In order to exclude the possibility of nonspecific RNA packaging of an overexpressed RNA possibly through retrofection, we created a control transfer vector named SJ40. As shown in Fig. 1b, SJ40 is a derivative of the wild-type transfer vector (SJ2) and lacks all of the viral sequences at its 5′ end. It contains the SV40 promoter/enhancer and *hygromycin resistance* gene at the 5′ end followed by MPMV viral sequences from nucleotide 7181 until nucleotide 8557 (end of MPMV 3′ long terminal repeat). Such a cloning strategy allowed the use of 3′ end MPMV long terminal repeat for proper transcript termination of SV-hygro RNA whose packaging (if any) could then be monitored.

All the resulting mutant clones were confirmed by sequencing. The details of PCRs, primers, and the intermediate cloning steps can be obtained from the authors upon request.

Transfection, virus production, and infection

Producer cells (293T) were transfected using the calcium phosphate as described previously.[58] Briefly, DNA cocktail used for transfection contained 2 μg/well each of (1) wild type or mutant transfer vector DNA, (2) packaging construct (TR301), and (3) envelope expression plasmid (MD.G). An additional plasmid, pGL3 control (Promega), expressing firefly luciferase, was added to the DNA cocktail to measure transfection efficiency as described earlier.[58] Using the three-plasmid trans-complementation assay,[32,33] transfected 293T cells produced pseudotyped virus particles containing transfer vector RNA, which allowed the monitoring of both RNA packaging and propagation.

Seventy-two hours following transfection, supernatants obtained from the transfected cultures were pelleted to isolate viral RNA in order to determine the RPE through RT-PCR. Part of the viral supernatants was also used to infect HeLa $CD4^+$ cells to monitor transfer vector RNA propagation efficiency by selecting transduced cells with medium containing Hygromycin B antibiotic after 42 h of infection. The number of resulting hygromycin-resistant (Hyg^r) CFU/ml of transfected supernatants for each transfer vector was normalized to the transfection efficiency in order to determine the relative CFU/ml (relative propagation efficiency compared to the wild type) for each transfer vector RNA.

To pellet viral particles, we cleared supernatants from transfected cultures of cellular debris via low-speed centrifugation at 4000 rpm for 10 min followed by filtration through 0.2-μm cellulose acetate syringe filters. Based on transfection efficiencies, different volumes of supernatants were used to pellet the viral particles using 20% sucrose cushion and resuspended in TNE buffer as described previously.[34,58] One-third of the viral particle suspension was saved for analysis of virus particles by

western blot, while the remaining two-thirds was used to isolate virion RNA.

Nucleocytoplasmic fractionation of transfected cells, RNA isolation, and RT-PCR

A portion of the cells collected from transfected cultures were fractionated into cytoplasmic and nuclear fractions as described previously.[34,58] Cytoplasmic fractions and viral particle suspensions were added to Trizol and Trizol LS reagents, respectively, and RNAs were isolated as described earlier.[34,58] Prior to cDNA preparation, both cytoplasmic and viral RNAs were DNase treated and amplified for 30 cycles using the vector-specific primers OTR216 (S; 5′ GGAATTCAGACAGTGGTGGCTCATA 3′) and OTR217 (AS; 5′ CCTCTAGAAGGAAATACGC-CATCTTA 3′) in order to check for the absence of any contaminating plasmid DNA. To ascertain the integrity of the RNA fractionation process, cytoplasmic RNA fractions were amplified to look for the absence of unspliced β-actin mRNA, which is exclusively nuclear[35] using primers OTR582 and OTR581 as described earlier.[34] Towards this end, a multiplex PCR was performed in the presence of primers/competimers for 18S ribosomal RNA (18S Quantum competimer control, Ambion) as a control for the presence of amplifiable cDNA.

After having confirmed the integrity of our fractionation process, PCR amplifications of cDNA samples were performed using the vector-specific primers OTR216 and OTR217 for 15, 20, and 25 cycles. Amplified products were gel-electrophoresed and transferred to a nylon membrane for Southern blot analysis as described previously[34] using a probe that was amplified using the same primers and SJ2 as the template.

RPE of mutant transfer vector RNA

Southern blots were scanned using the Biometra gel documentation system. ODs were measured using the BioDoc Analyze software version 2.0 followed by further calculations to determine the RPE of each transfer vector. The RPE of each transfer vector was determined by calculating the ratio of the packaged mutant RNA to the wild-type RNA (SJ2) relative to the ratio of the two RNAs in the cytoplasm as follows: RPE = [(OD of mutant transfer vector RNA − Bkgrd OD)/(OD of wild-type transfer vector RNA − Bkgrd OD)]/[(OD of cellular mutant RNA normalized to RTE − Bkgrd OD)/(OD of wild-type transfer vector RNA normalized to RTE − Bkgrd OD)]. OD is the optical density observed in the specific RT-PCR product, Bkgrd is the background OD observed in the control transfer vector (SJ40) lane, and RTE is the relative transfection efficiency.

Western blot analysis

To ensure that equal amounts of viral particles were pelleted and used for viral RNA isolation by using normalized volumes of transfected supernatant, we analyzed pelleted viral particles by western blot using anti-MPMV Gag/Pol Pr78 polyserum as described previously.[32,34]

Secondary RNA structure prediction and statistical analysis

Mfold sequence analysis software[39,40] was used to predict the secondary RNA structure of the 5′ end of MPMV RNA genome and also to analyze the effect of the introduced mutations on the folding potential of the proposed RNA secondary structure. The statistical analysis was performed by one-way ANOVA with Dunnett's post test, keeping SJ2 as the control, using GraphPad Prism (version 5.01) for Windows (GraphPad software, San Diego, CA, USA).

Acknowledgements

This research was funded primarily by a grant from Emirates Foundation (2009/044 to T.A.R.) and in part by new project grants (NP 08/17 and NP 09/18) from the Faculty of Medicine and Health Sciences, UAE University, to T.A.R. We express our thanks to Dr. Didier Trono (Ecole Polytechnique Fédérale de Lausanne, Switzerland) for providing MD.G. and Dr. Eric Hunter (Emory University, Atlanta, GA) for providing MPMV molecular clone and MPMV Gag/Pol polyclonal serum. We also wish to thank Ms. Jaicy George, Department of Microbiology and Immunology, Faculty of Medicine and Health Sciences, UAE University, for her assistance during the course of these experiments and Ms. Akela Ghazawi, Department of Microbiology and Immunology, Faculty of Medicine and Health Sciences, UAE University, for stimulating discussions. We also wish to thank Professor Andrew Lever and Dr. Julia Kenyon, Addenbrook's Hospital, Cambridge University, Cambridge, UK, for their input on RNA structure predictions.

References

1. D'Souza, V. & Summers, M. F. (2005). How retroviruses select their genomes. *Nat. Rev. Microbiol.* **3**, 643–655.
2. Lever, A. M. L. (2007). HIV RNA packaging. *Adv. Pharmacol.* **55**, 1–32.
3. Al Dhaheri, N. S., Phillip, P. S., Ghazawi, A., Ali, J., Beebi, E., Jaballah, S. A. & Rizvi, T. A. (2009). Cross-packaging of genetically distinct mouse and primate retroviral RNAs. *Retrovirology*, **6**, 66.
4. Russell, R. S., Liang, C. & Wainberg, M. A. (2004). Is HIV-1 RNA dimerization a prerequisite for packaging? Yes, no, probably? *Retrovirology*, **1**, 23.
5. Hibbert, C. S., Mirro, J. & Rein, A. (2004). mRNA molecules containing murine leukemia virus packaging signals are encapsidated as dimers. *J. Virol.* **78**, 10927–10938.
6. Paillart, J. C., Shehu-Xhilaga, M., Marquet, R. & Mak, J. (2004). Dimerization of retroviral RNA genomes: an inseparable pair. *Nat. Rev., Microbiol.* **2**, 461–472.
7. Paillart, J. C., Berthoux, L., Ottmann, M., Darlix, J. L., Marquet, R., Ehresmann, B. & Ehresmann, C. (1996). A dual role of the putative RNA dimerization initiation site of human immunodeficiency virus type 1 in genomic RNA packaging and proviral DNA synthesis. *J. Virol.* **70**, 8348–8354.
8. Greatorex, J. & Lever, A. (1998). Retroviral RNA dimer linkage. *J. Gen. Virol.* **79**, 2877–2882.

9. Greatorex, J. (2004). The retroviral RNA dimer linkage: different structures may reflect different roles. *Retrovirology*, **1**, 22.
10. Moore, M. D. & Hu, W. S. (2009). HIV-1 RNA dimerization: it takes two to tango. *AIDS Rev.* **11**, 91–102.
11. Skripkin, E., Paillart, J. C., Marquet, R., Ehresmann, B. & Ehresmann, C. (1994). Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro. *Proc. Natl Acad. Sci. USA*, **91**, 4945–4949.
12. Laughrea, M. & Jetté, L. (1994). A 19-nucleotide sequence upstream of the 5′ major splice donor is part of the dimerization domain of human immunodeficiency virus 1 genomic RNA. *Biochemistry*, **33**, 13464–13474.
13. Kenyon, J. C., Ghazawi, A., Cheung, W. K. S., Phillip, P. S., Rizvi, T. A. & Lever, A. M. L. (2008). The secondary structure of the 5′ end of the FIV genome reveals a long-range interaction between R/U5 and gag sequence, and a large, stable stem loop. *RNA*, **14**, 2597.
14. Berkhout, B. (1996). Structure and function of the human immunodeficiency virus leader RNA. *Prog. Nucleic Acid Res. Mol. Biol.* **54**, 1–34.
15. Laughrea, M., Jetté, L., Mak, J., Kleiman, L., Liang, C. & Wainberg, M. A. (1997). Mutations in the kissing-loop hairpin of human immunodeficiency virus type 1 reduce viral infectivity as well as genomic RNA packaging and dimerization. *J. Virol.* **71**, 3397–3406.
16. Lanchy, J. M., Ivanovitch, J. D. & Lodmell, J. S. (2003). A structural linkage between the dimerization and encapsidation signals in HIV-2 leader RNA. *RNA*, **9**, 1007–1018.
17. Baig, T. T., Lanchy, J. M. & Lodmell, J. S. (2007). HIV-2 RNA dimerization is regulated by intramolecular interactions in vitro. *RNA*, **8**, 1341–1354.
18. Lanchy, J. M. & Lodmell, J. S. (2007). An extended stem-loop 1 is necessary for human immunodeficiency virus type 2 replication and affects genomic RNA encapsidation. *J. Virol.* **7**, 3285–3292.
19. Baig, T. T., Lanchy, J. & Lodmell, J. S. (2009). Randomization and in vivo selection reveal a GGRG motif essential for packaging human immunodeficiency virus type 2 RNA. *J. Virol.* **83**, 802–810.
20. Berkowitz, R., Fisher, J. & Goff, S. P. (1996). RNA packaging. *Curr. Top. Microbiol. Immunol.* **214**, 177–218.
21. Banks, J. D., Beemon, K. L. & Linial, M. L. (1997). RNA regulatory elements in the genomes of simple retroviruses. *Semin. Virol.* **8**, 194–204.
22. Watanabe, S. & Temin, H. M. (1982). Encapsidation sequences for spleen necrosis virus, an avian retrovirus, are between the 5′ long terminal repeat and the start of the gag gene. *Proc. Natl Acad. Sci. USA*, **79**, 5986–5990.
23. Adam, M. A. & Miller, A. D. (1988). Identification of a signal in murine retrovirus that is sufficient for packaging of nonretroviral RNA into virions. *J. Virol.* **62**, 3802–3806.
24. Burns, C. C., Moser, M., Banks, J., Alderete, J. P. & Overbaugh, J. (1996). Identification and deletion of sequences required for feline leukemia virus RNA packaging and construction of a high-titer feline leukemia virus packaging cell line. *Virology*, **222**, 14–20.
25. Kemler, I., Barraza, R. & Poeschla, E. M. (2002). Mapping the encapsidation determinants of feline immunodeficiency virus. *J. Virol.* **76**, 11889–11903.
26. Browning, M. T., Mustafa, F., Schmidt, R. D., Lew, K. A. & Rizvi, T. A. (2003). Sequences within the gag gene of feline immunodeficiency virus (FIV) are important for efficient RNA encapsidation. *Virus Res.* **93**, 199–209.
27. Browning, M. T., Mustafa, F., Schmidt, R. D., Lew, K. A. & Rizvi, T. A. (2003). Delineation of sequences important for efficient FIV RNA packaging. *J. Gen. Virol.* **84**, 621–627.
28. Kemler, I., Azmi, I. & Poeschla, E. M (2004). The critical role of proximal gag sequences in feline immunodeficiency virus genome encapsidation. *Virology*, **327**, 111–120.
29. Mustafa, F., Ghazawi, A., Jayanth, P., Phillip, P. S., Ali, J. & Rizvi, T. A (2005). Sequences intervening between the core packaging determinants are dispensable for maintaining the packaging potential and propagation of feline immunodeficiency virus transfer vector RNA's. *J. Virol.* **79**, 13817–13821.
30. Vile, R. G., Ali, M., Hunter, E. & McClure, M. O. (1992). Identification of a generalized packaging sequence for D-type retroviruses and generation of a D-type retroviral vector. *Virology*, **189**, 786–791.
31. Guesdon, F. M. J., Greatorex, J., Rhee, S. R., Fisher, R., Hunter, E. & Lever, A. M. L. (2001). Sequences in the 5′ leader of Mason–Pfizer monkey virus which affect viral particle production and genomic RNA packaging: development of MPMV packaging cell lines. *Virology*, **288**, 81–88.
32. Schmidt, R. D., Mustafa, F., Lew, K. A., Browning, M. T. & Rizvi, T. A. (2003). Sequences within both the 5′ untranslated region and the gag gene are important for efficient encapsidation of Mason–Pfizer monkey virus RNA. *Virology*, **309**, 166–178.
33. Browning, M. T., Schmidt, R. D., Lew, K. A. & Rizvi, T. A. (2001). Primate and feline lentiviral vector RNA packaging and propagation by heterologous lentiviral virions. *J. Virol.* **75**, 5129–5140.
34. Ghazawi, A., Mustafa, F., Phillip, P. S., Jayanth, P., Ali, J. & Rizvi, T. A. (2006). 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.* **3**, 767.
35. Tan, W., Felber, B. K., Zolotukhin, A. S., Pavlakis, G. N. & Schwartz, S. (1995). 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.* **69**, 5607–5620.
36. Linial, M. (1987). Creation of a processed pseudogene by retroviral infection. *Cell*, **49**, 93–102.
37. Dornburg, R. & Temin, H. M. (1988). Retroviral vector system for the study of cDNA gene formation. *Mol. Cell. Biol.* **8**, 2328–2334.
38. Harrison, G. P., Hunter, E. & Lever, A. M. L. (1995). Secondary structure model of the Mason–Pfizer monkey virus 5′ leader sequence: identification of a structural motif common to a variety of retroviruses. *J. Virol.* **69**, 2175–2186.
39. Mathews, D. H., Sabina, J., Zuker, M. & Turner, D. H. (1999). Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* **288**, 911–940.
40. Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406.
41. Paillart, J. C., Skripkin, E., Ehresmann, B., Ehresmann, C. & Marquet, R. (2002). *In vitro* evidence for a long

range pseudoknot in the 5′-untranslated and matrix coding regions of HIV-1 genomic RNA. *J. Biol. Chem.* **277**, 5995–6004.
42. Paillart, J. C., Westhof, E., Ehresmann, C., Ehresmann, B. & Marquet, R. (1997). Non-canonical interactions in a kissing loop complex: the dimerization initiation site of HIV-1 genomic RNA. *J. Mol. Biol.* **270**, 36–49.
43. Russell, R. S., Hu, J., Laughrea, M., Wainberg, M. A. & Liang, C. (2002). Deficient dimerization of human immunodeficiency virus type 1 RNA caused by mutations of the U5 RNA sequences. *Virology*, **303**, 152–163.
44. Lever, A. M. L. (2009). RNA packaging in lentiviruses. *Retrovirology*, **6**, 113.
45. Mustafa, F., Lew, K. A., Schmidt, R. D., Browning, M. T. & Rizvi, T. A. (2004). Mutational analysis of the predicted secondary RNA structure of the Mason–Pfizer monkey virus packaging signal. *Virus Res.* **99**, 35–46.
46. Leitner, T., Foley, B., Hahn, B., Marx, P., McCutchan, F., Mellors, J. *et al.* (2005). *HIV Sequence Compendium 2005*. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM.
47. Houzet, L., Paillart, J. C., Smagulova, F., Maurel, S., Morichaud, Z., Marquet, R. & Mougel, M. (2007). HIV controls the selective packaging of genomic, spliced viral and cellular RNAs into virions through different mechanisms. *Nucleic Acids Res.* **35**, 2695–2704.
48. Huthoff, H. & Berkhout, B. (2002). Multiple secondary structure rearrangements during HIV-1 RNA dimerization. *Biochemistry*, **41**, 10439–10445.
49. L'Hernault, A., Greatorex, J. S., Crowther, R. A. & Lever, A. M. L. (2007). Dimerisation of HIV-2 genomic RNA is linked to efficient RNA packaging, normal particle maturation and viral infectivity. *Retrovirology*, **4**, 90.
50. Miyazaki, Y., Garcia, E. L., King, S. R., Iyalla, K., Loeliger, K., Starck, P. *et al.* (2010). An RNA structural switch regulates diploid genome packaging by Moloney murine leukemia virus. *J. Mol. Biol.* **396**, 141–152.
51. Daniel, M. D., King, N. W., Letvin, N. L., Hunt, R. D., Sehgal, P. K. & Desrosiers, R. C. (1984). A new type D retrovirus isolated from macaques with an immunodeficiency syndrome. *Science*, **223**, 602–605.
52. Desrosiers, R. C., Daniel, M. D., Butler, C. V., Schmidt, D. K., Letvin, N. L., Hunt, R. D. *et al.* (1985). Retrovirus D/New England and its relation to Mason–Pfizer monkey virus. *J. Virol.* **54**, 552–560.
53. Konings, D. A., Nash, M. A., Maizel, J. V. & Arlinghaus, R. B. (1992). Novel GACG-hairpin pair motif in the 5′ untranslated region of type C retroviruses related to murine leukemia virus. *J. Virol.* **66**, 632–640.
54. Yang, S. & Temin, H. M. (1994). A double hairpin structure is necessary for the efficient encapsidation of spleen necrosis virus retroviral RNA. *EMBO J.* **13**, 713–726.
55. Naldini, L., Blomer, U., Gallay, P., Ory, D., Mulligan, R., Gage, F. H. *et al.* (1996). In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science*, **272**, 263–267.
56. Sonigo, P., Barker, C., Hunter, E. & Wain-Hobson, S. (1986). Nucleotide sequence of Mason–Pfizer monkey virus: an immunosuppressive D-type retrovirus. *Cell*, **45**, 375–385.
57. Gibbs, J. S., Regier, D. A. & Desrosiers, R. C. (1994). Construction and in vitro properties of SIVmac mutants with deletions in "nonessential" genes. *AIDS Res. Hum. Retroviruses*, **10**, 607.
58. Mustafa, F., Jayanth, P., Phillip, P. S., Ghazawi, A., Schmidt, R. D., Lew, K. A. & Rizvi, T. A. (2005). Relative activity of the feline immunodeficiency virus promoter in feline and primate cell lines. *Microbes Infect.* **7**, 233–239.