Essential RNA Binding and Packaging Domains of the Gag-Pol Fusion Protein of the L-A Double-stranded RNA virus of Saccharomyces cerevisiae*

(Received for publication, May 5, 1994, and in revised form, July 8, 1994)

Juan Carlos Ribas‡, Tsutomu Fujimuras, and Reed B. Wickner

From the Section on Genetics of Simple Eukaryotes, Laboratory of Biochemical Pharmacology, NIDDK, National Institutes of Health, Bethesda, Maryland 20892

The crucial process in the assembly of the L-A double-stranded RNA virus is the recognition of its (+) single-stranded RNA by the Gag-Pol protein. The Pol region of this protein has RNA binding activity and is necessary for RNA packaging. Here we show that there are actually two in vitro RNA-binding domains of Pol (residues 172-190 and 770-819), and both are necessary for viral propagation, (but not for particle assembly). Furthermore, the N-terminal RNA-binding domain is necessary for in vivo packaging of viral (+) single-stranded RNA. We precisely define the extent of the Pol packaging domain (residues 67-213), which includes the N-terminal RNA-binding domain. This suggests that the N-terminal RNA-binding domain is responsible for binding the genomic RNA in the process of packaging and that additional surrounding residues are responsible for the specificity of binding.

In packaging, an RNA virus must pick out its genome from a sea of cellular mRNAs. This key process requires that a specific site(s) on the viral RNA be recognized by a viral protein(s). All dsRNA viruses package the viral (+) single-stranded RNA, converting it in the virus particle (or virus precursor particle) to the double-stranded form. The only dsRNA virus whose RNA packaging site has been determined is the L-A dsRNA virus of Saccharomyces cerevisiae. L-A has a single 4.6-kilobase segment encoding both Gag, its major coat protein, and the multifunctional Pol, expressed only as a Gag-Pol fusion protein formed by a -1 ribosomal frameshift event (see Fig. 1, Refs. 1-3). M, dsRNA, a satellite of L-A, encodes a secreted protein toxin (the "killer" toxin) and immunity to the action of that toxin. Because M, uses the L-A-encoded proteins for its propagation, the killer phenotype is useful as an indicator of the functions of L-A and of host genes affecting L-A. The 530-base pair X dsRNA is derived from L-A by a large internal deletion and has also been useful in the study of L-A (4). In vitro template-dependent replication and transcription systems have been developed for L-A, making possible characterization of the RNA sites recognized in these processes (2, 5, 6, see Fig. 1).

Because L-A replicates in S. cerevisiae, it has also been possible to learn many details of its interaction with its host (reviewed in Ref. 7).

Our complete L-A cDNA clones are capable of supporting M, dsRNA propagation, allowing testing of the in vivo requirement of any of the L-A dsRNA sequences independent of their cis requirement for replication (8). The sequence similarity of the Pol region of the L-A Gag-Pol fusion protein to the RNA-dependent RNA polymerases of all (+) ssRNA and dsRNA viruses (1) has led us to undertake the molecular dissection of this multifunctional protein. We carried out extensive localized mutagenesis of the RDRP consensus motifs defining the regions essential for M, propagation and showing that the most conserved regions were, in fact, the most essential (9).

Purified L-A dsRNA-containing virus particles, when treated with low ionic strength, pop open, losing their dsRNA. These opened empty particles, when resiolated, can specifically bind either L-A or M, viral (+) ssRNA (2, 10). This binding reaction recognizes sites with similar structures about 400 nucleotides from the 3' ends of either L-A or M, (+) ssRNAs (see Fig. 1). We have shown that these binding sites are actually the packaging signals for L-A and M, (+) ssRNAs (10, 11). We have begun to define the protein components of the packaging reaction by showing that part of the Pol region of the Gag-Pol fusion protein is necessary for packaging of the viral (+) ssRNA (12) and that the Pol region also has RNA binding activity (13).

In this work, we have precisely defined the extent of the packaging domain of the L-A Pol region, and we have found that there are actually two ssRNA-binding domains in Pol. Localization of these domains shows that one is within the packaging domain, and in vivo experiments show that it is indeed required for packaging (and for M, propagation). The other RNA-binding domain is not necessary for packaging in vivo but is necessary for viral propagation.

MATERIALS AND METHODS

Strains and Media—YPAD, YFG, 4.7MB, S.D., and synthetic complete medium lacking appropriate amino acids (14) and LB and TB media (15) have been described. S. cerevisiae strains JR1 (MATa trp1 ura3 leu2 his3 pep4::HIS3 ura1-L-A-o BC), JR6 (JR1 pI2), JR5 pl2L2+ K' (MATa kar1 ura2 leu2 trpl L-A-o pl2L2+ M-o), and 5X47 (MATa/MATa his1+/+ trp1/+ ura3/+ M-o) were used. Escherichia coli strains MV1190 and CI206 (Bio-Rad), DH5α and JM110 (Stratagene) and WM6 (F supC) lacZ lacY mal trpB/purC himR165 photomC/+) were used. ATCC47020) (16) were used.

Plasmids and DNA Techniques—The frameshift sequence GGTT-TAGGA of the L-A cDNA expression plasmid pL2L2 (8) was modified by introducing an A to produce the sequence GGTGTTAAAGG making pol expressed in-frame with gag (pM2, (5)). We inserted the polylinker sequence XhoI-HindIII-Ncol-KpnI-ClaI 8 amino acids before the start of pol in pM2 by site-directed mutagenesis to make pJR6. The pol region from pJR6 was excised by cutting upstream with HindIII and down-stream of pol with BamHI and inserted in-frame with the 29-residue OmpA signal peptide-FLAG peptide of the pFLAG E. coli expression
plasmid (IBI) cut with HindIII and BglII to form pJR7. The Erase a Base Kit (Promega) was used to generate unidirectional deletions from CaCl2 gradient-purified plasmid DNA. For the 3' end deletions, pJR7 DNA was cut with XhoI, protected with α-phosphothioate nucleotides, and cut with EcoRI. For the 5' end deletions, DNA was cut with HindIII, protected, and cut with ClaI. pJR7 DNA was a series of pJR7 deletions starting from the 3' end of pol, leaving intact its 5' end. pJR56 refers to the series of pJR7 deletions starting from the 5' end of pol, leaving intact its 3' end and having the sequence in-frame with the FLAG peptide. All deletion mutants were sequenced.

pJR12 was made by inserting in ClaI-XhoI (ClaI Klenow filled) gag fragment from pJR8 into pFLAG cut with HindIII and XhoI (HindIII Klenow filled), leaving the protein in-frame and expressing amino acids 164–626 from Gag.

pJR26 is p5’46 with the deletion A/III-XhoI expressing the Pol amino acids 500–567. pJR27 is p5’20 with a SacI-EcoRI deletion of pol, filling its 3’ end and leaving the C terminus of Pot in-frame. pJR24 is p5’46 with a SalI-EcoRI deletion of pot, filling its 3’ end and leaving residues 242–414 and 796–860. pJR26, pJR27, pJR28 are pJR7 with deletions together, respectively. pJR58 was made by introducing the X fragment from gag into pJR7, leaving intact its 3’ end and having its C terminus. pJR41 and pJR42 are p5’46 with the termination codon TAA inserted, leaving the amino acids 100 to 715. pJR55, pJR56, and pJR57 are pJR7 with a deletion by site-directed mutagenesis from amino acids 172 to 190, 200 to 220, and 241 to 261, respectively.

Site-directed mutagenesis (18) was done using the Bio-Rad Muta- structs, was grown in 1 liter of H-trp or H-trp-ura medium at 30 °C for 4 days. The cells were harvested, washed once with 100 ml of 0.5 M NaCl, pH 7.0, 10 mM EDTA, 15 mM phenylmethylsulfonyl fluoride, 10% glycerol. After thorough homogenization, the supernatant was centrifuged at 100,000 × g for 1 h. The pellet was resuspended in 11 ml of buffer A (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10 mM EDTA, 15 mM phenyl methylsulfonyl fluoride, 10% glycerol) and added to the cell lysis mixture. The cell lysis mixture was incubated at 37 °C for 1 h to analyze RNA binding. After protein transfer, the nitrocellulose sheet was washed three times for 15 min each at room temperature with 100 ml of 0.1% SDS, 150 mM NaCl and suspended in 50 ml of binding buffer (10 mM MES-Tris-HCl, pH 7.0, 50 mM NaCl, 1 mM EDTA, 1 × Denhardt’s solution, 0.1 μg heparin/ml). The sheet was then incubated at room temperature for 15 min in 45 μl of binding buffer with 1.5 mg (93 μg/ml) of denatured calf thymus DNA, and then the binding reaction was carried out by adding to the same buffer a 32P-labeled RNA probe. If, in place of DNA, equal amounts of mRNA or Torula total RNA were used, binding to the Pol fragments was unaffected, but the background of E. coli protein was higher. The probe was then heat-denatured, and the cells were suspended in 30 μl of 150 mM NaCl, pH 7.5, 50 μg/ml BSA, 10 mM EDTA, 1 mM dithiothreitol) by vortexing and lysed by osmotic shock. The cell suspension was centrifuged for 20 h at 130,000 × g for 1 h. The pellet was resuspended in 11 ml of buffer A by vortexing, and suspended in 1.2 ml/g wet weight of 100 mM Tris-HCl, pH 7.5, 20 mM β-mercaptoethanol, 1.5 μM sorbitol, and 4 mg/ml zymolyase 20T and incubated for 30–35 min at 37 °C with gentle shaking. The cells were collected by centrifugation at 100,000 × g for 10 min. The supernatant was then kept on ice, and the cells were suspended in 30 μl (3 μg wet weight) of buffer A (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10 mM EDTA, 1 mM dithiothreitol) by vortexing and lysed by osmotic damage. Cell debris and unbroken cells were removed by centrifugation for 10 min at 13,000 × g. The supernatant was collected and the pellet resuspended and treated with the previous lysis buffer (containing zymolyase) as before. The cell-bound beta-galactosidase was determined by incubating at 13,000 × g for 10 min, the supernatant was combined with the first lysate, and viral particles were collected by centrifugation at 100,000 × g for 1 h. The pellet was resuspended in 11 ml of buffer A and centrifuged at 10,000 × g for 20 min and the supernatant saved. Its density was adjusted to 1.32 by addition of CaCl2 (total volume 13 ml), and the suspension was centrifuged for 20 h at 130,000 × g at 4 °C. Fractions of 0.5 ml were collected and stored at −70 °C for further analysis.

Electrophoresis and Immunological Analysis. Viral Protein Gag fractions containing cDNA viral particles were diluted four times and solubilized in SDS loading buffer (final concentration 2 ×), boiled for 4 min, and analyzed by SDS-7.5% PAGE, 5 μl of each gradient fraction was used for Coomassie Blue staining and 1 μl for immunoblot analysis. After blotting to nitrocellulose (1.5 amp, 1.5 h in Tris-acetate buffer, 100 mM Tris-acetate), the sheet was exposed to a 500-816 from Gag.

RNA Binding and Packaging by Gag-Pol of L-A Virus 28421
The antibodies were detected by an alkaline phosphatase-conjugated secondary antibody (Promega).

The antibodies were detected by an alkaline phosphatase-conjugated secondary antibody (Promega).

**RESULTS**

**Pol Has at Least Two in Vitro ssRNA-binding Domains**—The Gag-Pol fusion protein has *in vitro* ssRNA binding activity, but the Gag protein does not (13), indicating that this is an activity of the Pol region (Fig. 1). To localize the region or regions responsible, we expressed the Pol open reading frame in *E. coli* using pJR7 in which Pol expression was driven by the tac promoter. We developed an *in vitro* RNA binding method, using 32P-labeled *X* (+) ssRNA as probe.

When the Pol protein was divided into two non-overlapping fragments (from amino acids 1–395 (p5'32) and 341–690 (p5'39)), both showed RNA binding activity. However, a central Pol fragment (aa 242–414 (p2D01)) and a Gag fragment (aa 164–626 of the Gag sequence (pJR12)) lacked this binding activity (Fig. 2, right). Thus, Pol possesses at least two different in vitro ssRNA-binding domains, and this activity is not due simply to the abundance of the overexpressed proteins.

**The N-terminal ssRNA-binding Domain Is Located between Residues 172 and 190**—To localize more precisely the N-terminal domain involved in the ssRNA binding activity, we made 3' end deletions of the Pol-encoding sequence in pJR7 (Fig. 3). All mutant proteins were efficiently expressed (Fig. 3, left). A protein with the first 190 amino acids of Pol efficiently bound ssRNA, but a deletion of 24 more amino acids resulted in loss of most binding activity and longer deletions had no binding activity at all (Fig. 3, right). This locates the N-terminal boundary of the N-terminal-binding domain at residue 180.

Delimitation of the N-terminal end of the N-terminal-binding domain required removal of the C-terminal half of Pol in order to detect lack of binding when the N-terminal-binding domain was disrupted. We made 5' end deletions starting with pJR7D which lacks the C-terminal half of Pol (SalI to 5' end, i.e. residues 415–860, deleted). Sequencing identified the one-third of clones in which the remaining part of Pol was in-frame with the translation initiation AUG. These clones efficiently expressed protein fragments of the expected size (Fig. 4, left), and their RNA binding activity showed that residues 1–172 were not necessary, but that deletion of an additional 10 residues further eliminated binding activity (Fig. 4, right). Thus, residue 172 is the N-terminal border of the N-terminal-binding domain.

**Residues 770–819 Contain the C-terminal ssRNA-binding Domain**—5' end deletion mutants of pJR7 were sequenced to determine which were in-frame with the translation initiation AUG. All of those deleted beyond residue 697 were out of frame, perhaps because of toxicity to *E. coli*. We also noted that fragments in the 18–29 kDa range, independent of their amino acid sequence, were more difficult to express than larger fragments. All the deletion mutants in this series had ssRNA binding activity (Fig. 5) showing that the C-terminal-binding domain is downstream of residue 697.

To locate the N-terminal end of the C-terminal-binding domain, pJR44, pJR45, and pJR46 were made (Fig. 6, left). They included at their N-terminus residues 1–99, which do not bind RNA (see Fig. 3, pJR50), in order to increase the size of the protein and avoid expression problems. A further deletion mutant (pJR27) was obtained by in-frame deletion of a SalI-EcoRI fragment (residues 414–796). The expressed proteins had the expected sizes, and only the shortest deletion lacked RNA binding activity. Thus, the N-terminal boundary of the C-terminal-binding domain is residue 770.

The C-terminal end of the C-terminal-binding domain was determined using p5'46 and its derivatives (Fig. 6, right). The expressed proteins had the expected sizes, and their RNA binding activities show the C-terminal-binding domain lies upstream of residue 819. Interestingly, the most basic cluster in all of Pol, 785KRK789, was not sufficient to give binding activity (pJR28, Fig. 6, right). The two other most basic regions, 789RRK793 and 800KVKRR799, also lie outside the RNA-binding domains.

**Deletion of Two Domains Is Necessary and Sufficient to Make Pol Unable to Bind RNA**—To confirm that the N- and C-terminal RNA-binding domains are the only domains responsible for the RNA binding of the intact Pol protein, we removed them individually or together and assayed RNA binding activity (Fig. 7). Deletion of just the N-terminal domain (residues 172–190) left binding unchanged, while deletion of just the C-terminal domain (770–819) left reduced, but definite, binding activity. Only removal of both domains eliminated binding activity of Pol. Apparently these are the only RNA-binding domains in Pol.

The weaker binding of Pol lacking the C-terminal domain may be due to weaker RNA affinity of the N-terminal domain or to altered accessibility of the RNA to this site in the mutated protein.
Fig. 2. Pol has at least two in vitro RNA-binding domains. Left, the entire Pol fragment was expressed from pJR7 in E. coli. Soluble (Sup) and insoluble (Ppt) fractions from E. coli without or with expression of Pol were analyzed by SDS-12% PAGE and Coomassie Blue-stained. The Pol fragment was highly expressed but only appeared in the insoluble fraction as the major protein component. Right, the proteins examined are diagrammed below. Proteins made in E. coli were isolated and analyzed by SDS-PAGE. The gels were either stained with Coomassie Blue (numbers to the left give the size of marker proteins in kDa), or transferred to a nitrocellulose membrane, treated with urea, and probed with $^{32}$P X (+) strand RNA. Except for protein D, the amino acid residue number refers to Pol and begins with the 1st residue of Pol after the frameshift (R1 S2 G3 R4 S5 Y6 D7...). + and – denote ssRNA binding activity.

The N-terminal in Vitro RNA-binding Domain Is Needed for in Vivo Viral RNA Packaging—To correlate the in vitro RNA binding activities with in vivo functions, we first measured RNA packaging by a modification of our previous method (Fig. 8A, ref. 12). Two plasmids are simultaneously expressed in yeast strain JR1 lacking the normal L-A virus. The L-A protein expression plasmid (pI2L2 or a derivative) is the source of either wild type or mutant L-A viral particle proteins, and the packaging substrate expression plasmid (pJR58) produces a transcript, detectable on a Northern blot, containing the L-A packaging site. Expression of both plasmids should produce viral particles that encapsidate the small transcript carrying the packaging signal provided they have an intact Pol packaging domain.

Viral particles were purified in a CsCl gradient with an initial density of 1.32 g/ml. The gradient was divided into 24-25 fractions, and the particles appeared concentrated in three to five fractions (Fig. 8B, upper panel). The density of the peak was 1.32, slightly denser than the free proteins ($\rho = 1.31$). Those particles were empty as they did not contain any pI2L2 transcript (data not shown). Had they packaged their mRNA, they would have been denser, as the L-A full particles have a density of 1.41 g/ml. Although empty, they sediment as particles and not as free proteins (12, and data not shown).
When the L-A proteins were expressed from the wild type plasmid pI2L2 or mutant plasmids lacking either the N-terminal (172–190) or C-terminal (770–819) RNA-binding domains and the packaging substrate was expressed from another plasmid (pJR58), the particle distribution was essentially the same as that for pI2L2 alone, peaking in all cases at 1.32 g/ml. The presence of particles in all cases shows that the deletion mutations are not affecting the assembly or stability of the particles, a result expected from the fact that expression of Gag alone is sufficient to produce morphologically normal (albeit empty) particles (12). As expected, when the transcript from the packaging substrate plasmid was packaged, it was found in particles slightly denser than the bulk of particles in the gradient (Fig. 8B, compare third and fourth panels).

In each case the Gag-Pol fusion protein was expressed and packaged in particles in similar amount (Fig. 8B, third panel), but the packaging substrate RNA transcript was only found in particles made from the wild-type plasmid or from the C-
RNA Binding and Packaging by Gag-Pol of L-A Virus

The C-terminal RNA-binding domain is between Pol residues 770 and 819. The deletion mutants on the left demarcate the left border of the C-terminal RNA-binding domain, while those on the right set the right border. See legend of Fig. 2.

Fig. 6. The C-terminal RNA-binding domain is between Pol residues 770 and 819. The deletion mutants on the left demarcate the left border of the C-terminal RNA-binding domain, while those on the right set the right border. See legend of Fig. 2.

The in Vivo Packaging Domain Comprises Residues 67–213 of Pol—The N-terminal one-fourth of Pol (along with Gag) is sufficient for packaging RNA bearing the specific L-A packaging signal (12). A mutant expressing only the first 213 residues (pJR10) had this in vivo packaging activity, but one truncated 8 residues further (pTF139) was inactive (12). Of course, none of these constructs were able to support the propagation of M1 (Fig. 9).

Here we show that deletion of the 19-amino-acid N-terminal RNA-binding domain, inside the N-terminal one-fourth of Pol results in inability to package the signal-containing RNA (Fig. 8) and also to propagate M1 (Fig. 9). This mutant has the rest of the Pol sequence, and the lack of activity shows that the N-terminal RNA-binding domain is necessary and is the only one that is sufficient for packaging.

To more accurately determine the extent of the in vivo packaging domain, we made a series of deletions shown in Fig. 9. In all cases, particle assembly and incorporation of the Gag-Pol fusion protein were normal (data not shown). Deletion of either residues 209–213 (pJR75) or 204–208 (pJR74) destroyed packaging activity and ability to propagate M1. Thus, the C terminus of the packaging domain is between residues 209 and 213. The localization of the N-terminal end was accomplished by making a series of five Pol deletion mutants, from residue 39 (just at the end of the overlapping gag open reading frame) to residue 182 (11 amino acids inside the N-terminal in vivo RNA-binding domain). None of these mutants were able to propagate M1 in vivo and all except pJR69, the most N-terminal deletion, were unable to package in vivo. That pJR69 does not propagate M1 indicates that the 39–66 deletion is affecting some function of the viral cycle, but it is not packaging ability itself. The deletion 67–95 (pJR70) is the most N-terminal deletion that affects packaging and therefore must overlap the N-terminal end of the packaging domain (Fig. 9). Thus, the Pol region necessary and sufficient for viral packaging comprises residues 67–213.

DISCUSSION

L-A shares with retroviruses the formation of a Gag-Pol fusion protein using ribosomal frameshifting, but it resembles the pararetroviruses like hepatitis B virus in requiring Pol for

Fig. 7. Only deletion of both N-terminal and C-terminal RNA-binding domains eliminates RNA binding by Pol. Legend is as for Fig. 2.

terminal-binding domain deletion plasmid (Fig. 8B, lower panel). Even in much longer autoradiographic exposures, packaging was not detected with the mutant deleted for the N-terminal-binding domain. Thus, the C-terminal RNA-binding domain is unrelated to packaging, but the N-terminal RNA-binding domain is necessary for packaging.

Confirmation that these domains are necessary for the in vivo viral replication cycle is that either C- or N-terminal-binding domain deletion mutant failed to propagate M1 dsRNA when tested by cytoduction experiments (see "Materials and Methods"). Two substitution mutations within the C-terminal RNA binding domain, 795KKNEFANK800 → 795AAAAF AAA800 and 803WERTMYKAYR812 → 803AAAAAMAAAA812 (pJR80 and pJR81), also failed to maintain M1. However, the substitution, 785KRKAR789 → 785AAAA786, does not affect propagation of M1. Within the N-terminal RNA-binding domain the substitution
Fig. 8. *in vivo* method used to test the role of specific regions of Pol in RNA packaging. Viral proteins made from one plasmid package an RNA transcript with the L-A packaging site made from another plasmid. Isolated virus particles are tested by Northern blot hybridization for packaging of the second plasmid's transcript. 

A, packaging of RNA requires the N-terminal *in vitro* RNA-binding domain but not the C-terminal domain. The L-A protein expression plasmid (see A) was either the intact L-A cDNA expression clone, pL2L2 (w.t.) (8) or N-termΔ=pJR59, lacking Pol residues 172–190, or C-termΔ=pJR60, lacking residues 770–819. The packaging substrate expression plasmid (pJR58, below in A) has the X sequence (including the L-A packaging site). Virus particles were isolated by CsCl equilibrium gradient centrifugation. Particles made from the pL2L2 wild-type L-A cDNA clone in the absence of the packaging substrate expression plasmid are shown at the top of B. Fractions of the CsCl gradients were analyzed by SDS-7.5% PAGE and either Coomassie Blue staining or by immunoblot analysis with antibodies to either Pol or Gag. Aliquots were also analyzed by Northern blot analysis probed with ³²P-labeled X (−) strand RNA.

**B** CsCl gradient of L-A particles made from the cDNA clone

packaging of genomic RNA (24, 25). Nonetheless, L-A is a dsRNA virus, typical of a rapidly growing group of viruses having similar architecture and found in fungi and parasitic microorganisms (reviewed in Ref. 26).

We show here that the Pol region of the Gag-Pol fusion protein has two *in vitro* ssRNA-binding domains as detected by Northwestern blot RNA binding assay of Pol fragments expressed in E. coli. The C-terminal *in vitro* RNA-binding domain, comprising amino acids 770–819, is not necessary for *in vivo* packaging, but is needed for *in vivo* M₁ propagation (Fig. 10). Its proximity to the consensus RNA-dependent RNA polymerase domains suggests that it may have to do with the transcriptase or replicase activities. Two substitution mutants within this domain are unable to propagate M₁, supporting its physiological importance. Interestingly, the most basic cluster of amino acids in all of Pol, ³⁸⁵KRKAR³⁹⁰, was not sufficient to
in vivo propagation and could be part of the RNA-dependent RNA polymerase domains in Pol. The N-terminal in vitro RNA-binding domain is part of the in vivo packaging domain. The C-terminal in vitro RNA-binding domain, though not related to packaging, is essential for viral propagation and could be part of the RNA-dependent RNA polymerase domains, the most conserved parts of which are residues 541–561 and 565–583 (9).

The N-terminal in vitro RNA-binding domain comprises residues 172–190, and it is necessary for M protein propagation and for in vivo packaging. This sequence is 172RPRKHFKGRLRX-TSKVTK190. Although the presence of basic residues, with tyrosine, phenylalanine, glycine, and proline residues is like many known RNA-binding domains, we could not find homology of this domain, or the C-terminal domain with such regions of other proteins. Although residues to the left or right of the two RNA-binding domains defined in this study may contribute to the binding in the absence of the other side, the deletion mutations define the only residues of each domain essential for the activity.

The packaging model for the L-A virus was based on the existence of the Gag-Pol fusion protein and the presence of single-stranded RNA binding activity attributable to the Pol region (13). The essence of this model was that the Pol region of single-stranded RNA binds to the viral (+) strand, and the Gag region would associate with free Gag protein resulting in the packaging of the viral genome (and the RNA polymerase). Several points of this model have now been verified, including the existence of a specific packaging site on the viral (+) ssRNA, and the requirement for packaging for a part of Pol. We show here that Pol has two ssRNA-binding domains, one of which is indeed required for packaging as predicted by the model.

We precisely defined the extent of the in vivo packaging domain, finding that it comprises residues 67–213, much larger than the included N-terminal in vitro RNA-binding domain (Fig. 9). We also find that the deletion of just 5 residues (780SPRRK213) eliminates in vivo packaging, but not in vitro RNA binding. A similar sequence (SPKK) has been shown to form an Asx-turn, a special type of β turn, and to interact with the minor groove of double-stranded DNA. The SPKK motif has been identified in many nucleic acid-binding proteins including Xenopus H1 histone and hepatitis B virus core protein (28, 29). Since the L-A virus cis-packaging site consists of a stem-loop with an A bulge, the packaging domain of Pol must recognize and bind to this complex structure. This could explain why the in vivo packaging domain consists of not only the N-terminal in vitro RNA-binding domain but also its surrounding regions. For example, the SPRRK sequence might interact with the stem structure at one of the two grooves such that the RNA-binding domain comes into close contact with the loop sequence of the packaging site. Alternatively, the surrounding regions might be necessary for the packaging domain to fold properly, so that the RNA-binding domain can interact with the stem-loop structure. These possibilities should be examined when soluble native Pol protein becomes available.

RNA-binding domains of many RNA virus-encoded proteins have been defined, but in most cases their functions remain unclear. The packaging of retroviruses involves the nucleocapsid portion of the Gag precursor protein. Conserved C.C...H...C sequences recognize the viral packaging site (30, 31), and packaging of the genomic RNA ensues.

Several of the five rotavirus nonstructural proteins have ssRNA binding activity. NS34 (NSP3) protein has ssRNA binding activity that is sequence nonspecific in vitro (32), but in vivo cross-linking experiments show that it is specifically associated with the conserved 3′ ends of viral (+) ssRNA (33). NS35 has both ssRNA and dsRNA binding activity, but again has no sequence specificity in vitro (34). In vivo, NS35 is associated with viral mRNA in the cytoplasmic region where viral replication is proceeding, the viroplasm (35). Ts mutants in this protein produce empty particles at the non-permissive temperature, suggesting that it has a role in packaging of viral (+) strands (36).
RNA Binding and Packaging by Gag-Pol of L-A Virus

The NS53 (NSP1) protein has single-stranded RNA binding activity with specific affinity in vitro for the 5' part of viral (+) strands (37). This activity is due to a cysteine-rich region near the N terminus that binds zinc (38, 37). The function of NS53 is not yet clear, but its association with the cytoskeleton and not the viroplasm suggests that it may be involved in initiating the organization of new replicase particles (37).

It is not yet clear which rotavirus RNA-binding nonstructural protein corresponds to the L-A ssRNA binding regions defined here, but presumably all of the RNA binding functions of L-A's multifunctional Pol region are encompassed in the more complex rotavirus replication and packaging apparatus. The availability in the L-A system of both in vitro template-dependent replication and transcription systems, in vivo expression and packaging assay systems, and the easily manipulated yeast host system, may allow dissection of the details of these processes more easily than is possible at this time with other dsRNA viruses.

REFERENCES