Recognition of RNA Encapsulation Signal by the Yeast L-A Double-stranded RNA Virus*

Tsutomu Fujimura‡ and Rosa Esteban

From the Instituto de Microbiología Bioquímica / Departamento de Microbiología y Genética, Consejo Superior de Investigaciones Científicas / Universidad de Salamanca, Salamanca 37007, Spain

The encapsidation signal of the yeast L-A virus contains a 24-nucleotide stem-loop structure with a 5'-nucleotide loop and an A bulged at the 5' side of the stem. The Pol part of the Gag-Pol fusion protein is responsible for encapsidation of viral RNA. Opened empty viral particles containing Gag-Pol specifically bind to this encapsidation signal in vitro. We found that binding to empty particles protected the bulged A and the flanking-two nucleotides from cleavage by Fe(II)-EDTA-generated hydroxyl radicals. The five nucleotides of the loop sequence (4190GAUCC4194) were not protected. However, T1 RNase protection and in vitro mutagenesis experiments indicated that G4190 is essential for binding. Although the sequence of the other four nucleotides of the loop is not essential, data from RNase protection and chemical modification experiments suggested that C4194 was also directly involved in binding to empty particles rather than indirectly through its potential base pairing with G4190. These results suggest that the Pol domain of Gag-Pol contacts the encapsidation signal at two sites: one, the bulged A, and the other, G and C bases at the opening of the loop. These two sites are conserved in the encapsidation signal of M1, a satellite RNA of the L-A virus.

Double-stranded RNA (dsRNA)1 viruses encapsidate their plus single-stranded genomic RNAs into viral particles or inner cores and subsequently convert them into double-stranded form by the particle-associated RNA-dependent RNA polymerase. Since all of RNA polymerization reactions, transcription and replication, take place inside the particles, the dsRNA viruses must have efficient and selective mechanisms to encapsidate not only their genomic RNAs but also their RNA polymerase machinery into particles.

The yeast L-A dsRNA virus contains a single molecule of linear uncapped 4.6-kilobase dsRNA with two open reading frames (Fig. 1A). The 5' open reading frame is gag and encodes a 76-kDa major coat protein. The 3' open reading frame is pol and encodes RNA-dependent RNA polymerase, which is expressed only as Gag-Pol fusion protein by a -1 ribosomal frame-shifting mechanism (1–3). The Pol region of the fusion protein has a single-stranded RNA (ssRNA) binding activity (1). L-A virions have an icosahedral symmetry (4). The capsids consist of 60 asymmetric Gag dimers, and one molecule or two of Gag is replaced by Gag-Pol fusion protein.

When isolated L-A virions are exposed to low ionic strength conditions, they release L-A dsRNA (5). The resulting opened empty particles retain the Gag-Pol fusion protein and thus can perform transcription and replication reactions in vitro with specificity, if appropriate template RNAs are added (5, 6). Empty particles also bind specifically to the plus strand of viral RNA (7). Using a gel shift assay, the cis signal for binding has been identified as a 24-nucleotide stem-loop with a bulged A at the 5' side of the stem (Fig. 1C) (8). It is located near the 3' end of the L-A plus strand. Subsequently, the binding signal was identified by an in vivo assay as the encapsidation signal for the L-A virus (9). The encapsidation activity resides in the N-terminal quarter of Pol (10). According to the proposed L-A virus encapsidation model (Fig. 1D) (1, 10), the virus has developed a simple, ingenious mechanism to secure the packaging of both the viral plus strand and the RNA polymerase; the encapsidation domain (the N-terminal quarter of Pol) of the Gag-Pol specifically recognizes and binds to the cis encapsidation signal present on L-A plus strands, forming a Gag-Pol-L-A plus ssRNA complex. Then the Gag domain of the fusion protein triggers the capsid assembly by homologous interactions with free Gag proteins (or Gag dimers). The selective packaging of the viral RNA is thus governed by the interaction between Gag-Pol and the viral plus ssRNA.

Since empty particles specifically bind to the L-A encapsidation signal, this reaction provides valuable information about the nature of the protein-RNA interactions involved in the selective packaging of the viral RNA. In this paper we have analyzed the in vitro binding reaction in detail using chemical and enzymatic probes and also by in vitro mutagenesis. The results indicate that empty particles interact with the encapsidation signal at two sites: one, the bulged A and the surrounding two nucleotides at the 5' side of the stem, and the other, the G and C bases at the opening of the loop. Since the natures of the interactions at these two sites are quite different, it suggests that the fusion protein recognizes these RNA sites with different contact elements.

**Experimental Procedures**

Preparation of Empty Particles—Intact L-A virions were prepared from strain TPF29 (MATa his (3,4), leu2, sk2-2, LA-HN) as described (11). Empty particles were obtained from isolated L-A virions by low ionic treatment and then purified through a CsCl gradient as described (11).

Gel-shift Assay—Gel-shift assay was carried out as described (12). In some experiments (Fig. 2) tRNA was omitted from binding reaction mixtures.

In Vitro Mutagenesis—Oligonucleotide site-directed mutagenesis was carried out as described (8). All of the desired changes were confirmed by DNA sequencing.

*This work was supported by Grant PB97–1121 from Dirección General de Enseñanza Superior (DGES) of the Spanish Ministry of Education. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡To whom correspondence should be addressed: Instituto de Microbiología Bioquímica CSIC, Universidad de Salamanca, Avda. del Campo Charro s/n, Salamanca 37007, Spain. Tel.: 34-923-120673; Fax: 34-923-224876; E-mail address: tfujimura@www-micro.usal.es.

1 The abbreviations used are: dsRNA, double-stranded RNA; ssRNA, single-stranded RNA; nt, nucleotide; DMS, dimethyl sulfate.

Received for publication, June 16, 2000, and in revised form, July 27, 2000
Published, JBC Papers in Press, August 22, 2000, DOI 10.1074/jbc.M005245200

This paper is available on line at http://www.jbc.org
L-A Encapsidation Signal

Plasmids—pLM1 contains most of the cDNA sequence of X, a deletion mutant of L-A, and has been described (11). It contains the L-A plus strand sequence from nt 4277 to nt 4296 and the other from nt 5925 to nt 5944. The RNA was transcribed from a plasmid with a large internal deletion (from nt 12 to nt 4090) (Fig. 1B). Between the T7 promoter and the L-A sequence there are 16 nucleotides derived from the vector.

In Vitro Transcription—Uniformly labeled RNA for gel-shift assay was made from EcoRI-digested pLM1 or its derivatives using T7 RNA polymerase and [α-32P]UTP as described (7). RNA for enzymatic cleavage or chemical modifications was made in quantity using a MEGA script T7 kit from Ambion Inc., Austin, Texas, from pLM1 or its derivatives predigested with XmnI unless otherwise stated (Fig. 1B).

pCp Labeling—RNA (2 μg) was 3′ end-labeled with [32P]pCp and T4 RNA ligase (Life Technologies, Inc.). The conditions were as suggested by the manufacturer. The labeled RNA was purified through a 5% polyacrylamide gel as described (13), except that the electrophoresis was carried out for 90 min at 8 V/cm. The RNA was cut out from the gel and extracted with 0.6 M ammonium acetate, 1 mM EDTA, and 0.1% SDS overnight at room temperature. The extracted RNA was phenol-treated, ethanol-purified, and used for ribonuclease protection experiments. For hydroxyl radical protection experiments, the RNA was further purified using an RNA purification kit from Biotools or CA. As revealed in Figs. 4 and 5, the labeled RNAs contained a small amount (about 20%) of the same RNA species but one nucleotide shorter at the 5′ end. This, however, did not affect our conclusions.

Hydroxyl Radical Cleavage—We basically followed the cleavage procedure as described (14). Since glycerol quenches hydroxyl radicals (15) a stock solution of empty particles was passed through a G-50 min spin column (Worthington Biochemical Corp.) pre-equilibrated with 10 mM Tris/0.1 M HCl, pH 7.5, 50 mM NaCl, and then used for protection experiments. RNA labeled at the 3′ end (100,000–200,000 cpm) was incubated for 10 min on ice in a buffer (25 μl) containing 10 mM Tris/0.1 M HCl, pH 7.5, 50 mM NaCl, and 0–10 μl of empty particles (−5 μg of protein/μl). One μl of each of the following four solutions was added to the top of a 15 ml Eppendorf tube containing the RNA sample. The solutions were made freshly for each experiment and were 1) 100 μM Na-EDTA, pH 7.5, 2) 50 mM Fe(NH4)2(SO4)2.6H2O (Aldrich), 3) 250 mM sodium ascorbate, and 4) 2.5% HO2. All solutions were mixed simultaneously by a brief centrifugation at 4 °C, and the tube was kept another 10 min on ice. The cleavage reaction was stopped by the addition of 10 μl of 0.1 mM thiosulfate, and 70 μl of 0.5 M sodium acetate, 0.4% SDS, and 1 μg of tRNA. The RNA was extracted with phenol, phenol/chloroform, and chloroform and precipitated with ethanol. After being washed with 70% ethanol, the RNA was separated on an 8% polyacrylamide sequencing gel. To generate a sequencing ladder, alkali hydrolysis was carried out at 95 °C for 90 s in a buffer containing 50 mM sodium bicarbonate/sodium carbonate, pH 9.2, 5 μg of tRNA, and the labeled RNA (200,000–400,000 cpm). A T1 ladder was created by incubating the labeled RNA (150,000–300,000 cpm) in a buffer containing 25 mM sodium citrate, pH 5.0, 3 μg of tRNA, 7 M urea, and 3 units of T1 nuclease at 55 °C for 15 min.

Ribonuclease Protection Experiments—3′ end-labeled RNA (100,000–200,000 cpm) was incubated on ice for 10 min in a mixture (19 μl) containing 20 μM Tris/HCl, pH 7.5, 50 mM NaCl, 5 μg of tRNA, and various amounts of empty particles (−5 μg of protein/μl). One μl of ribonuclease T1 or T2 (0.02–0.2 units/μl) was added to the mixture, and the cleavage was carried out at 30 °C for 10 min. After the reaction, the RNA was extracted with phenol and phenol/chloroform and analyzed on an 8% acrylamide sequencing gel. To ensure that each RNA molecule would receive no more than a single cleavage, the amounts of T1 and T2 used in Fig. 5 were 0.1 units. In these conditions more than 70% of the labeled RNAs remained undigested.

Dimethyl Sulfate (DMS) Treatment—Non-labeled RNA (0.5 pmol) transcribed from EcoRI- or FspI-digested pLM1 by T7 RNA polymerase was incubated with 0, 2, or 5 μl empty particles (−10 μg of protein/μl) in a buffer (total volume 20 μl) containing 50 mM sodium cacodylate, pH 7.5 and 50 mM NaCl. Empty particles had been dialyzed twice against the same buffer for total 2 h at 4 °C. After a 10-min incubation on ice, the binding reaction mixture received 0.5 μl of 33% DMS (Aldrich) in ethanol and was incubated at 20 °C for another 10 min. The methylation reaction was terminated by the addition of 50 μl of 25 mM Tris/Cl, pH 7.5, 250 μM β-mercaptoethanol, 400 mM sodium acetate, and 2 μg of tRNA. The RNA (and protein) was precipitated with 300 μl of ethanol, and the pellet was dissolved in 200 μl of 0.3 M sodium acetate and 0.3% SDS. Then the RNA was extracted with phenol and phenol/chloroform and precipitated with ethanol. After an 80% ethanol wash, the RNA was dried and then served for reverse transcription reactions. The methylation by RNA as DMS was detected by reverse transcription using 5’ end-labeled primers. We used two primers, one complementary to the L-A plus strand sequence from nt 4277 to nt 4286 and the other from nt 4246 to nt 4259. The primers were labeled at the 5′ ends with [γ-32P]ATP and T4 polynucleotide kinase (MBI Fermentas) according to the method recommended by the enzyme supplier. To the methylated RNA (0.2 pmol), dissolved in 3 μl of water, were added 7 μl of 4.2 M NaCl, 50 mM Tris/Cl, pH 7.5, and 7 μM Na-EDTA, 4 μl of the 5′ end-labeled primer (−400,000 cpm), and 30 μl of formamide. Annealing was carried out at 45 °C overnight. The sample was diluted 6-fold with 0.3 M sodium acetate, and 20 μl of 32P-labeled RNA was precipitated with ethanol, washed once with 80% ethanol, and dried. The pellet was dissolved in water, and DNA synthesis was carried out at 45 °C for 1 h in a volume of 30 μl using 100 units of Superscript RNase H reverse transcriptase (Life Technologies, Inc.) in the conditions recommended by the supplier. The reaction was stopped by the addition of 15 μM EDTA, and the template RNA was digested at 37 °C.
for 20 min with 5 ng of RNase A. DNA was extracted with phenol/chloroform and chloroform, ethanol-precipitated, and analyzed on an 8% polyacrylamide sequencing gel. The sequence ladders were created by T7 DNA polymerase with a T7 sequencing kit (Amersham Pharmacia Biotech) using plM1 as template, the same but non-labeled primer as mentioned above, and [α-32P]dATP.

RESULTS

High Binding Affinity—Empty particles have a high affinity to bind to ssRNA with the L-A encapsidation signal. We used clones of X dsRNA, a deletion mutant of L-A, to make ssRNAs in vitro with the wild type or modified encapsidation signal (Fig. 1B). When empty particles were incubated with a ssRNA fragment having the wild type sequence, the gel-shift assay revealed a single species of binding complexes over a wide range of RNA concentrations (Fig. 2). The Lineweaver-Burk plot of binding became a straight line (not shown). These results indicate that empty particles can bind a single molecule of RNA in the reaction. From the plot, we obtained the maximum number of binding sites as 11 pmol/mg of protein. Using \(9.3 \times 10^{3}\) kDa as the molecular mass of an empty particle (118–119 subunits of Gag with a molecular mass of 76 kDa and one or two Gag-Pol fusion proteins with 170-kDa molecular mass), it is estimated that 10% of the particle population in this preparation were active in binding to the encapsidation signal. The dissociation constant \(K_d\) was calculated as 0.45 nM. A similar value has been reported for binding to the encapsidation signal of M1, a satellite RNA of L-A virus (16). This \(K_d\) gives a \(\Delta G^\circ\), Gibbs free energy, of \(-12.8\) kcal/mol.

L-A Encapsidation Signal—The cis encapsidation signal of L-A virus consists of a 24-nucleotide stem-loop with an A bulge at the 5′ side of the stem (Fig. 1C). Empty particles specifically bind to this signal. The importance of the bulged A in the reaction has been demonstrated by the inability in binding of modified RNAs whose bulged A were deleted or substituted with any of the other three nucleotides (9). When the loop sequence 4190GAUCC4194 was modified conservatively, the modified RNA with 4190AGCUU4194 failed to bind, indicating the importance of the loop sequence. On the contrary, the stem sequences themselves do not appear to be critical in binding, since the substitution of the nucleotide sequence at each side of the lower stem with the other destroyed the binding activity, whereas the restoration of the lower stem structure by simultaneously exchanging those sequences recovered the binding activity (9) (Fig. 3A). When similar exchange experiments were done to the upper stem sequences, the same results were obtained (Fig. 3B). Destruction of the upper stem structure by substituting each of its stem sequences with the other resulted in a substantial decrease in binding efficiency.

FIG. 2. Quantitative analysis of the binding reaction. Empty particles (20 µg of protein) were incubated with 100,000 (1) or 200,000 (2–8) cpm of 32P-labeled truncated X plus strand RNA made by T7 RNA polymerase from EcoRI-digested plM1 (1.75 \(\times 10^6\) cpm/pmol). The same but non-labeled transcripts were also added to the reaction (3–8). The amounts of cold transcripts added were 33 (3), 65 (4), 130 (5), 260 (6), 550 (7), and 1100 (8) nmol. Sample 9 contained only labeled RNA (200,000 cpm) without empty particles as control. After incubation, RNA was separated in an agarose gel and visualized by autoradiography. Bound and unbound RNAs were scanned with an AGFA DUOSCAN and quantified using the computer program MacBas 2.0, Koshin Graphic systems, Inc.

FIG. 3. Nucleotide sequences of the lower (A) and upper (B) stems of the encapsidation signal are not important for binding. 32P-Labeled RNA containing the encapsidation signal or modifications thereof were incubated in the presence or absence of empty particles, and the binding complexes were separated on agarose gels. Autoradiograms of the gels are shown. The wild type sequence (a in each case) and modified sequences (underlined) are shown.
in the failure in binding. Again, swapping the upper sequences simultaneously, thus restoring the stem structure, recovered the binding activity, although slightly less compared with the original structure. Therefore, these results suggest that the stem sequences themselves are not important, but that the stem structure is important to present the bulged A and the loop sequence at a certain distance or spatial configuration.

Hydroxyl Radical Footprinting—We used hydroxyl radical footprinting to analyze the physical interactions between the encapsidation signal and empty particles. Hydroxyl radicals generated by a Fenton reaction from Fe(II)-EDTA and H₂O₂ are quite ideal as a probe to examine protein-RNA interactions (14, 17). First, the radicals cleave the sugar backbone, and the cleavage is not affected by the bases of the sequence. Second, the size of the probe is small; thus it is suitable for a fine mapping. More importantly, the cleavage is not affected by the secondary structure of RNA as far as the sugar backbone is accessible to the aqueous phase (18, 19).

A small L-A fragment with the encapsidation signal was made in vitro and labeled at the 3′ end with ³²P-pCp by T4 RNA ligase. Then the labeled RNA was treated with hydroxyl radicals in the presence or absence of empty particles. As shown in Fig. 4, the bulged A and its surrounding two nucleotides were protected from cleavage upon binding to particles. Interestingly, the 3′ side stem sequence and the loop sequence were not protected. As a control, the same RNA fragment but with a modified loop sequence (4190AGCUU4194) was used. Since this RNA has a much reduced binding affinity, almost no protection by empty particles was observed (Fig. 4). Therefore these results indicate that empty particles physically contact with the bulged A and the surrounding few nucleotides but that the loop sequence is not protected from hydroxyl radical cleavage by binding.

The Loop Sequence Is Also Important—Although the loop sequence of the encapsidation signal is important for binding (9), hydroxyl radical footprinting failed to detect interactions between the loop sequence and empty particles. Since the radicals cleave the sugar backbone, this may suggest that the particles contact with the nucleotide bases of the loop. In fact, when similar protection experiments were done using T1 ribonuclease, we could detect such interactions. As shown in Fig. 5A, the first G (base 4190) of the loop is accessible to T1 nuclease. The extent of the cleavage, however, was less compared with other unpaired G residues, thus suggesting its partially double-stranded nature. In the presence of empty particles, this G was completely protected from T1 cleavage. This protection is related to the binding events, since the same G in the bulged A-deleted RNA was not protected (Fig. 5C).
Ribonuclease T2 cleaved the RNA well at the A and U of the loop sequence and, at a lesser extent, at the C of base 4193. These T2 cleavages, however, were not inhibited by binding to the particles (Fig. 5B). Therefore, these results suggest that the first G (base 4190) of the loop is important for binding. There are two plausible explanations for this protection. Empty particles directly interact with this G and protect it from T1 cleavage. Alternatively, the binding of empty particles may stabilize the upper stem structure and induce base pairing of this G with C4194, thus forming a loop with three bases instead of five. Since T1 does not cleave double-stranded regions, this alternative structure would result in apparent protection of the G from T1 cleavage.

To confirm the importance of the first G of the loop in binding, we modified the loop sequence by *in vitro* mutagenesis and tested their binding activity. A mutation in the loop (4191GCU4193) did not significantly affect binding (Fig. 6). This result is consistent with T2 cleavage experiments (Fig. 5B); that is, empty particles did not protect these three nucleotides. Since a mutant RNA with 4190AGCU4194 failed to bind to empty particles, these results suggest that the remaining two bases of the loop (G at 4190 and/or C at 4194) are important for binding. In fact, when these two nucleotides were simultaneously substituted with C (4190) and G (4194) or A (4190) and U (4194), none of these mutant RNAs showed binding activity (Fig. 6). To evaluate their importance in binding, we mutated these two nucleotides individually. When G at 4190 was substituted by any of the other three nucleotides, all of them failed to bind to empty particles (Fig. 7A), confirming that G at 4190 is essential for binding. On the other hand, when C at 4194 was modified, all of them retained the binding activity substantially. Especially, the RNA with C substitution (GG mutant) showed a binding activity only slightly less compared with the wild type RNA. These *in vitro* mutagenesis experiments therefore indicate (i) that the G at 4190 is essential for binding, whereas (ii) that C at 4194 is exchangeable with other nucleotides in binding.

**Binding to Empty Particles Does Not Induce GC Base Pairing in the Loop**—As shown above C at 4194 can be replaced by other nucleotides in binding. This does not necessarily mean that this nucleotide does not contact with the empty particles. We rather think that it directly interacts with them, based on the following evidence. When the GG mutant RNA with G at 4194 was treated with T1 nuclease, the two G residues (at 4190 and 4194) in the loop were protected from T1 by binding to empty particles (Fig. 5D). The extent of the protection of each G was the same at a given concentration of empty particles (Fig. 5E). This result thus indicates that the particles directly interact with the first G (at 4190) of the loop and protect it from T1 cleavage rather than inducing the GC base pair at the opening of the loop. The result also indicates that empty particles directly interact with C4194 in the GG mutant RNA (and perhaps with C4194 in the wild type RNA).

To confirm this we did DMS modification experiments. DMS methylates A (N1) and C (N3) at the Watson-Crick base-pairing positions in the single-stranded regions, and these modifications can be detected by reverse transcriptase using 5′ end-labeled primers (17, 20). As shown in Fig. 8, the methylation of two C residues at positions 4194 and 4206 was greatly stimulated upon binding to empty particles. We used two primers.
One of them was complementary to the L-A plus strand sequence from nt 4277 to nt 4296 (Fig. 8) and the other, from 4246 to 4265 (not shown). Both primers gave the same results. We also used two RNA templates for DMS methylation. One contained the intact L-A 3'9 end region (from nt 4091 to nt 4579), including the encapsidation signal and the correct L-A 3'9 end without extra sequences (Fig. 8). The other was the same RNA but truncated at nt 4353 (not shown). Again, in both cases we observed the same results. From these we can mention several points. First, the N3 atom of C at base 4194 was more methylated by binding to empty particles, thus indicating that the binding does not induce the Watson-Crick GC base pairing at the opening of the loop. Therefore it confirms our previous conclusion that empty particles directly interact with G at 4190 and protect it from T1 cleavage. Second, since G4190 and C4194 are partially double-stranded in the absence of empty particles, according to the T1 protection experiment (Fig. 5A), this stimulation in methylation rather indicates that the binding of empty particles brakes the base pairing between these two nucleotides and increases the accessibility of the N3 of C4194 toward DMS. Finally, the modification of C at 4206 (and A4207) by DMS was also stimulated by empty particles. Since the particles did not protect this C and the surrounding nucleotides from hydroxyl radical or T2 cleavage (Figs. 4 and 5B), this stimulation is likely to be a secondary effect caused by the binding. Perhaps the tension induced by binding is released in this single-stranded region directly attached to the 3'9 lower stem. This may be correlated with our previous observation that the addition of the 10-nucleotide upstream and downstream sequences to the 24-nucleotide encapsidation signal increased the binding affinity to empty particles (9).

**DISCUSSION**

**Interactions at Two RNA Sites**—In this paper we have analyzed physical interactions between the cis encapsidation signal of L-A virus and empty L-A viral particles. It is reasonable to assume that most of the interactions (if not all) exerted by empty particles can be ascribed to the Pol part of the fusion protein, since the packaging activity of RNA with the encapsidation signal resides in the N-terminal quarter of Pol (10) and neither Gag nor the Gag part of the fusion protein appears to be involved in the recognition of the encapsidation signal (21). Empty particles specifically bind to the encapsidation signal with high affinity (Kd; 5 x 10^-10 M) and they physically contact it at two sites. One is the bulged A4186 and the surrounding few nucleotides (U4185 and G4187), including the encapsidation signal and the correct L-A 3' end without extra sequences (Fig. 8). The other was the same RNA but truncated at nt 4353 (not shown). Again, in both cases we observed the same results. From these we can mention several points. First, the N3 atom of C at base 4194 was more methylated by binding to empty particles, thus indicating that the binding does not induce the Watson-Crick GC base pairing at the opening of the loop. Therefore it confirms our previous conclusion that empty particles directly interact with G at 4190 and protect it from T1 cleavage. Second, since G4190 and C4194 are partially double-stranded in the absence of empty particles, according to the T1 protection experiment (Fig. 5A), this stimulation in methylation rather indicates that the binding of empty particles brakes the base pairing between these two nucleotides and increases the accessibility of the N3 of C4194 toward DMS. Finally, the modification of C at 4206 (and A4207) by DMS was also stimulated by empty particles. Since the particles did not protect this C and the surrounding nucleotides from hydroxyl radical or T2 cleavage (Figs. 4 and 5B), this stimulation is likely to be a secondary effect caused by the binding. Perhaps the tension induced by binding is released in this single-stranded region directly attached to the 3' lower stem. This may be correlated with our previous observation that the addition of the 10-nucleotide upstream and downstream sequences to the 24-nucleotide encapsidation signal increased the binding affinity to empty particles (9).
stem structure is intact (Fig. 3). We have not observed the protection of the 3’ side stem by empty particles from hydroxyl radical cleavage. Therefore, these results suggest that empty particles have very localized contact with the 5’ side of the stem, namely, at the bulged A and the sugar moieties of its flanking nucleotides.

The second site is two nucleotides, G4190 and C4194, at the opening of the loop. The other three nucleotides of the loop (AUC4193) do not have direct contact with empty particles. G4190 is essential for binding since it cannot be substituted with any of the other three nucleotides. Furthermore, empty particles protected G4190 from T1 cleavage, thus indicating their physical interaction. We have ruled out the possibility that this protection was caused by base pairing of G4190 with C4194, since empty particles protected G4190 well from T1 cleavage even in a modified GG mutant RNA whose C4194 was substituted with G. In addition, methylation at N3 of C4194 by DMS was greatly increased by binding to empty particles. If Watson-Crick base pairing between C4190 and C4194 were involved in the protection, we should observe the opposite, a decrease of methylation at C4194 upon binding to empty particles.

We have several lines of evidence that suggests direct interactions between C4194 and empty particles. In the GG mutant RNA, T1 cleavage at G4190 was inhibited by empty particles, and the pattern of its inhibition was identical to the one observed at G4190, thus implying that the same binding event elicited the protection at both G residues simultaneously. This result therefore indicates that empty particles directly contact with G at position 4194 in the GG mutant. And it is perhaps the same with the wild type sequence, since the binding affinities decrease in the order of C, G, U, and A at position 4194. Bruenn and co-workers (22) have also observed that not only the bulged A and G4190 but also C4194 are conserved among RNA isolates with binding activity in in vitro selection. M1, a satellite RNA of the L-A virus, has these three nucleotides exactly at the same positions as L-A in its encapsidation site (9).

A Possible Mechanism for the Interaction between C4194 and Empty Particles—The contribution of a single hydrogen bonding in EcoRI endonuclease/DNA substrate binding is estimated to be -1.5 kcal/mol on the average (23). The substitution of C4194 with A decreased its binding affinity about 10-fold. It corresponds to a decrease in the binding energy of less than 1.4 kcal/mol. Therefore, it suggests that the interactions between empty particles and C4194 are not extensive. If we assume that empty particles recognize O2 of C4194, some of the data presented here could be explained. First, binding of empty particles to G4190 and O2 of C4194 would break a Watson-Crick base pairing between them, and it may increase the accessibility of N3 of C4194 to DMS methylation. Second, the binding affinities of RNAs substituted at 4194 decreased in the order of C, G, U, and A. Since adenosine has no keto oxygens, the A substitution has the lowest affinity. Uracil has the keto oxygen (O2) similar to cytosine, but the hydrogen atom at N3 causes unfavorable conditions for binding. Then, how is the G substitution? If two G residues at 4190 and 4194 in the GG mutant form a non-canonical GG N7-imino base pair, the relative distance and spatial configuration of O6 of G4194 (or O4) toward G4190 may be comparable with those of O2 of C toward G in a normal Watson-Crick GC base pairing (Fig. 9). Although G4190 and C4194 of the wild type RNA do not form a Watson-Crick base pairing in the complex, O6 of G4194 of the GG mutant may occupy a position similar to that of O2 of C4194 in the complex and interact with empty particles as a hydrogen bond acceptor. This also explains why binding to empty particles protected G4194 in the GG mutant from T1 cleavage (Fig. 5D). Since T1 nuclease interacts with O6 in addition to N1, N2, and N7 of guanine base (24), such an interaction would surely inhibit T1 cleavage at G4194 because of steric hindrance.

The minor groove of an A-type RNA helix is wide and shallow but provide poor information on hydrogen bond donors and acceptors to interact with proteins. The major groove, on the contrary, has rich information but is very deep and narrow (25). At the end of the helix the major groove becomes wider and accessible to reagents such as diethyl pyrocarbonate, which is comparable in size with protein side chains such as arginine (26). The second binding site G4190 and C4194 is placed at the end of the upper stem or at the opening of the five-nucleotide loop. Although the binding does not require a Watson-Crick base pairing between these nucleotides, T1 cleavage of the wild type sequence suggests their partially double-stranded nature in the absence of empty particles. Perhaps it is important for these bases to be held at a certain spatial configuration or in a dynamic state at the boundary between the upper helix and the loop in order to interact with empty particles. The upper stem between these two sites consists of three base pairs. Since the rotation of an A-type helix is 33° per residue, the sugar backbones of the bulged A and the G4190 may form an angle of about 115°. If a single Pol peptide in the empty particles is responsible for the interactions at these two sites, the contact of empty particles with the GC (4190 and 4194) site may take place through its major groove side.

Comparison of L-A Encapsulation Signal with MS2 Coat Protein Binding Site—The L-A encapsidation signal is quite similar to the RNA hairpin structure found in the MS2 bacteriophage replicase operator to which the MS2 coat protein dimer binds. Extensive structural analyses have been done on the latter system. The MS2 hairpin structure contains a four-nucleotide loop and a bulged A at the 5’ side of the stem separated from the loop by a two-base pair upper stem. Interestingly, this bulged A, when unligated, stacks in the double-stranded helix (27). On the contrary, when it forms a complex with the coat protein dimer, the bulged A becomes extrahelical and interacts with the dimer through extensive hydrogen bonding (28, 29). The encapsidation signal of the L-A virus, therefore, may undergo a similar structural change at the bulged A upon binding to empty particles. In the absence of empty par-
ticles, the bulged A would be intercalated between the neighboring base pairs. Its resistance to T2 digestion and non-susceptibility to DMS methylation may be indicative of its stacking nature. In the complex the bulged A may become extrahelical so that empty particles can contact extensively with its adenine base to discriminate it from other bases. The extrahelical so that empty particles can contact extensively with the hairpin structure second MS2 binding site is the last adenine of the loop. The with its adenine base to discriminate it from other bases. The extrahelical so that empty particles can contact extensively with the hairpin and why the two A residues are quasi-2-fold-related in structure in the complex. In the case of the L-A encapsidation signal, the contact sites with empty particles are quite different: one, the bulged A, but the other, G4190 (and C4194), at the opening of the loop. As shown in the hydroxyl radical protection experiments, empty particles interact extensively with the bulged A including its sugar backbone; on the other hand, their interactions with the second site appear to be confined to the nucleotide bases. The Gag part of the fusion protein directs its assembly into capsids, whereas the region necessary and sufficient for in vivo RNA packaging has been narrowed down to residues 67–213 of the Pol part (30) (Fig. 1A), still larger than the MS2 coat protein. Therefore, these data suggest that the encapsidation domain of Pol has two different sites to contact with the encapsidation signal.

**Does a Gag-Pol Homodimer Exist in the Virion?—**As discussed above, available data suggest that the encapsidation domain of the fusion protein has two different regions or sites to contact with the L-A RNA encapsidation signal. It has been proposed that the fusion protein forms a homodimer that is responsible for encapsidation (31, 32). If it is the case, one of these different regions in each subunit may contact with one of the two RNA sites. Alternatively, there is still the possibility that, like the MS2 coat protein, a single region in each subunit of the fusion protein dimer can contact with either the protruding A or GC site. However, we think that neither of them is the case, as discussed below. Two types of data can be mentioned as the supporting evidence for dimer formation. First, the efficiency of the –1 ribosomal frame-shifting, measured in an RNA fragment encompassing the frame-shifting site using a reporter gene, was 1.9%. Since L-A particles contain 120 copies of Gag, it roughly corresponds to two Gag-Pol molecules per particle. Furthermore, this frame-shifting efficiency was claimed to be critical for viral reproduction since either increasing or decreasing the efficiency more than 2-fold disrupted M1 satellite RNA propagation (31). When the frame-shifting was measured in a different RNA fragment or even in the same fragment but in a different cell background, however, much higher efficiencies (10–20%) were observed (33, 34). One may also wonder if the frame-shifting efficiency should be the same as the ratio of the coat proteins assembled. Several host mutants that increased the frame-shifting efficiency have been isolated. Most of them could actually maintain M1 at 20 °C. More importantly, none of them had any effects on L-A virus, even on its copy number (35).

The other line of evidence might be the incorporation of a truncated Gag-Pol protein into active (M1 dsRNA-synthesizing) particles. Since the truncated protein lacks the C-terminal half of Pol including the RNA polymerase domain, thus defective in RNA polymerization, those active particles containing the truncated protein should also contain the intact fusion protein (36). From these results, however, one cannot generalize that L-A virions contain two intact fusion protein molecules per particle. Because the truncated protein can be incorporated into capsids without viral RNA or the intact fusion protein by virtue of its intact Gag domain, the results can also be interpreted as a random distribution of the truncated protein into viral capsids. Furthermore, if the formation of a fusion protein dimer is responsible for the incorporation of the truncated protein into active M1 virions, then the results strongly suggest that RNA polymerase does not require dimerization for its activity. Pol contains a leucine zipper-like sequence (residues 216 to 244) that could be involved in dimerization (2). RNA encapsidation, however, does not require the C-terminal three-quarters of Pol, including this leucine zipper-like sequence (10). Then, why is the fusion protein dimerization necessary? Or is the N-terminal one-quarter of Pol still sufficient for formation of a Gag-Pol homodimer? As far as we know, there are no convincing data in the literature that show the presence of two fusion protein molecules per intact L-A virion. In our hands, the quantification of coat proteins strongly suggests that intact L-A virions contain a single Gag-Pol molecule per particle.2 Thus we believe that a single fusion polypeptide is responsible for the interactions with the two RNA sites in the encapsidation signal.

The data reported here suggest that the RNA-protein interactions between the L-A fusion protein and the encapsidation signal are more complex than those of the RNA bacteriophage. Further fine structural analyses such as x-ray diffraction or NMR studies would be indispensable to understand these interactions in detail.

**Acknowledgments—**We thank Drs. F. Leal and M. Medarde for valuable comments on the manuscript.

REFERENCES
1. Fujimura, T., and Wickner, R. B. (1988) Cell 55, 663–671
2. Icho, T., and Wickner, R. B. (1989) J. Biol. Chem. 264, 6716–6723
3. Dinman, J. D., Icho, T., and Wickner, R. B. (1991) Proc. Natl. Acad. Sci. U. S. A. 82, 174–178
4. Cheng, R. H., Caston, J. R., Wang, G., Gu, F., Smith, T. J., Baker, T. S., Rothard, R. F., Trus, R. L., Cheng, N., Wickner, R. E., and Steven, A. G. (1994) J. Mol. Biol. 244, 255–258
5. Fujimura, T., and Wickner, R. B. (1988) J. Biol. Chem. 263, 454–460
6. Fujimura, T., and Wickner, R. B. (1989) J. Biol. Chem. 264, 10872–10877
7. Esteban, R., Fujimura, T., and Wickner, R. B. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 4411–4415
8. Esteban, R., Fujimura, T., and Wickner, R. B. (1989) EMBO J. 8, 947–954
9. Fujimura, T., Esteban, R., Esteban, I. M., and Wickner, R. B. (1990) Cell 62, 819–828
10. Fujimura, T., Ribas, J. C., Makkov, A. M., and Wickner, R. B. (1992) Nature 359, 746–749
11. Esteban, R., Wickner, T., and Wickner, R. B. (1992) J. Biol. Chem. 267, 2708–2713
12. Esteban, R., and Wickner, R. B. (1988) J. Virol. 62, 1278–1286
13. Rodrigues-Coutinho, N., and Esteban, R. (1992) Nucleic Acids Res. 20, 2761–2766
14. Powers, T., and Noller, H. F. (1995) RNA (N. Y.) 1, 194–209
15. Bashkin, J. S., and Tullius, T. D. (1993) Footprinting of Nucleic Acid-Protein Complexes (Bevazin, A., ed) pp. 75–106, Academic Press, Inc., New York
16. Shen, Y., and Brunn, J. A. (1993) Virology 195, 481–491
17. Moine, H., Ehresmann, B., Ehresmann, C., and Rompy, B. (1997) RNA Structure and Function (Simons, R. W., and Grunberg-Manago, M., eds) pp. 77–115, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
18. Latham, J. A., and Cech, T. R. (1999) Science 285, 274–282
19. Celerand, D. W., and Cech, T. R. (1990) Biochemistry 29, 1355–1361
20. Ehresmann, C., Baudin, F., Mougel, M., Rompy, P., Ebel, J.-P., and Ehresmann, B. (1987) Nucleic Acids Res. 15, 9109–9128
21. Yao, W., Muqattad, K., and Brunn, J. A. (1995) J. Virol. 69, 1917–1919
22. Yao, W., Adelman, K., Gao, K., and Brunn, J. A. (1997) J. Virol. 71, 2151–2162
23. Lesser, D. R., Kurpiewski, M. R., Waters, T., Connolly, B. A., and Jen-Jacobson, L. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 5748–5752
24. Arni, R., Heinemann, U., Tsuchiya, R., and Saenger, W. (1988) J. Biol. Chem. 263, 13588–13598
25. Steitz, J. A. (1999) in The RNA World (gestel, R. F., Cech, F. C., and Atkins, J. F., eds) 2nd Ed., pp. 427–450, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
26. Weeks, K. M., and Crothers, D. M. (1993) J. Biol. Chem. 268, 1574–1577
27. Borger, P. N., Lin, Y., Wang, S., Roggenbuck, M. W., Gott, J. M., Uhlenbeck, O. C., and Pelzer, I. (1995) Biochemistry 34, 6488–6503
28. Valegard, K., Murray, J., Stockley, P. G., Stonehouse, N. J., and Liljas, L.2

2 T. Fujimura and R. Esteban, unpublished observations.
L-A Encapsidation Signal

(1994) Nature 371, 623–626
29. Valegard, K., Murray, J. B., Stonehouse, N. J., van den Worm, S., Stockley, P. G., and Liljas, L. (1997) J. Mol. Biol. 270, 724–738
30. Ribas, J. C., Fujimura, T., and Wickner, R. B. (1994) J. Biol. Chem. 269, 28420–28428
31. Dinman, J. D., and Wickner, R. B. (1992) J. Virol. 66, 3669–3676
32. Wickner, R. B. (1996) Microbiol. Rev. 60, 250–265
33. Tzeng, T.-H., Tu, C.-L., and Bruenn, J. A. (1992) J. Virol. 66, 999–1006
34. Balasundaram, J., Dinman, J. D., Wickner, R. B., Tabor, C. W., and Tabor, H. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 172–176
35. Dinman, J. D., and Wickner, R. B. (1994) Genetics 136, 75–86
36. Ribas, J. C., and Wickner, R. B. (1998) J. Biol. Chem. 273, 9306–9311