JOURNAL OF VIROLOGY, Apr. 1996, p. 2184–2190 Vol. 70, No. 4
0022-538X/96/$04.00+0

Identification of Domains in Rubella Virus Genomic RNA and Capsid Protein Necessary for Specific Interaction

ZHEWEI LIU,[1]† DECHENG YANG,[1] ZHIYONG QIU,[1] KYE-TAEK LIM,[1]‡ PELE CHONG,[2] AND SHIRLEY GILLAM[1]*

Department of Pathology and Laboratory Medicine, University of British Columbia Research Centre, Vancouver, British Columbia, Canada V5Z 4H4,[1] and Connaught Research Institute, Willowdale, Ontario, Canada M2R 3T4[2]

Received 12 September 1995/Accepted 18 December 1995

In rubella virus-infected cells, genomic 40S and subgenomic 24S RNAs are present in the cytoplasm of infected cells. However, encapsidation by rubella virus capsid protein is specific for 40S genomic RNA. As a first step toward understanding the assembly of rubella virus nucleocapsid at the molecular level, the interaction between capsid protein and genomic RNA was studied by Northwestern (RNA-protein) blot analysis. RNA probes prepared by in vitro transcription were used to localize the RNA sequence that participates in binding to the capsid protein. We have identified a 29-nucleotide RNA sequence (nucleotides 347 to 375) that is essential for the binding. By using overlapping synthetic peptides of capsid protein, a peptide domain (residues 28 to 56) that displays specific RNA-binding activity of capsid protein has been located. This result suggests that the specific recognition of viral RNA during rubella virus assembly involves, at least in part, the nucleocapsid protein.

Rubella virus (RV), a positive-stranded RNA virus, is the sole member of the genus *Rubivirus* in the family *Togaviridae* (16). The virions contain three structural proteins, E1 (58 kDa), E2 (42 to 47 kDa), and capsid protein, C (33 kDa) (20). E1 and E2 are membrane glycoproteins located on the virion exterior, and capsid protein is associated with the genomic RNA, forming the nucleocapsid (20). In RV-infected cells, the genomic 40S RNA serves as a messenger for the nonstructural proteins and as a template for synthesis of a subgenomic 24S RNA that encodes the structural protein precursor, which is processed by host proteases to give rise to three structural proteins in the order C, E2, and E1 (19). The nucleotide sequence of the genomic RNA of RV (Therien strain) has been reported (5). Recently, we have obtained the complete sequence of cDNA derived from the genomic RNA of wild-type RV (M33 strain) (unpublished data). Comparison between the Therien and M33 strains reveals a homology of 97.5% at the nucleotide level (unpublished data).

Although they are not closely related, the alphaviruses (members of the *Togaviridae* family) and RV have a similar strategy for the organization and expression of their structural protein genes (21, 26). However, the alphavirus capsid protein differs from the RV capsid protein with respect to the mechanism of its cleavage from the polyprotein. The alphavirus capsid protein is an autoprotease and quickly releases itself within the cytoplasm from the polyprotein precursor (10, 26), whereas RV capsid protein requires the presence of microsomal membrane for its cleavage from the polyprotein precursor in vitro (2, 3). In addition, the capsid protein in isolated RV virions has been reported to be a disulfide-linked dimer (32). The amino acid sequence predicted from RV capsid protein cDNAs reveals a protein of 299 residues that includes the signal peptide sequence of E2 (residues 281 to 299) (2, 17, 30). RV capsid protein is very basic and rich in proline (45 residues) and arginine (33 residues) (2). The cluster of proline and arginine residues in the capsid protein region may be involved in binding to genomic RNA in the viral nucleocapsid.

In enveloped RNA viruses, two major steps are required in the assembly of virion particles. The first is interaction of the genomic RNA with the capsid protein to form the nucleocapsid; the second is interaction of nucleocapsid with viral glycoproteins embedded in the cellular membrane that results in budding and release of the virus from the plasma membrane of the cells. Recent studies with defective interfering RNA and RNA transcripts made in vitro from cDNA of the viral genome have provided evidence for a specific interaction between viral RNA and capsid protein (29, 34). In Sindbis virus, the prototype alphavirus, the RNA region that interacts specifically with the capsid protein has been shown to be located near the 5′ terminus of the genomic RNA (nucleotides 746 to 1226) in the gene for nonstructural protein nsP1 (34). Deletion analysis of capsid protein of Sindbis virus has identified a potential binding motif, a stretch of 10 amino acids from residues 97 to 106 (KPKPGKRQRM), that interacts with viral RNA (9).

Very little is known about the specificity of RNA-capsid protein interaction in RV assembly, although it is known that only the genomic RNA (40S RNA) and not the subgenomic RNA (24S RNA) of RV is encapsidated (21). As a first step to study the molecular basis of RV assembly, we have investigated the specificity of interaction between viral RNA and capsid protein. In this study, we report the localization of the RNA sequence necessary for the binding of RNA to capsid protein and the peptide domain in capsid protein that interacts with RNA.

MATERIALS AND METHODS

Construction of plasmids. RV cDNAs encoding the nonstructural (5) and structural (2) proteins are shown in Fig. 1. All cDNAs were subcloned between the *Eco*RI and *Hin*dIII sites of the vector pSPT19 (Pharmacia) under the control of phage SP6 RNA polymerase promoter. Recombinant plasmids were used for in vitro synthesis of RNA transcripts.

In vitro transcription. Transcription with SP6 bacteriophage RNA polymerase was performed with linearized plasmid templates by standard protocols (Promega Corp., Madison, Wis.). Radiolabeled RNA transcripts were synthesized by

* Corresponding author. Mailing address: Department of Pathology and Laboratory Medicine, University of British Columbia Research Centre, 950 W. 28th Ave., Vancouver, British Columbia, Canada V5Z 4H4. Phone: (604) 875-2474. Fax: (604) 875-2496.

† Present address: Capital Institute of Pediatrics, Chaoyan District, Beijing 100020, People's Republic of China.

‡ Present address: Department of Animal Science, Chonnam National University, Kwangju City, Korea.

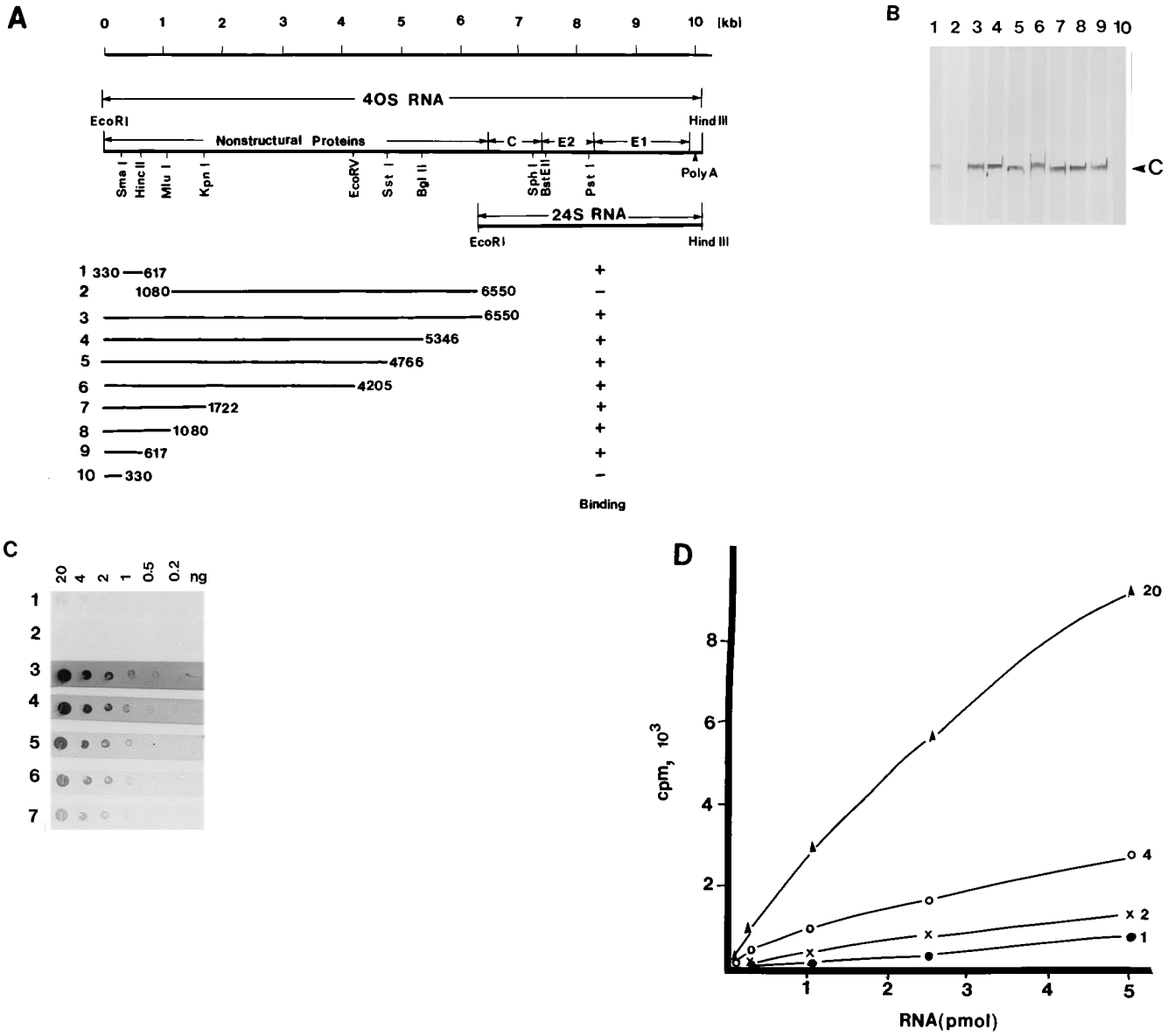

FIG. 1. Localization of the capsid protein binding region in the viral RNA. (A) Schematic diagrams of the RNA transcripts and RV genomic and subgenomic RNAs. Numbers on the genome scale refer to distances from the 5′ end in kilobases. The locations of the deleted regions are shown as nucleotides, numbered from the 5′ end of the RV genome. Restriction sites used in the synthesis of the deletion RNA transcripts are shown at the bottom of the diagram. (B) Northwestern blot analysis of RNA transcripts. Purified capsid protein (3 pmol per strip) immobilized on nitrocellulose was incubated with ^{35}S-labeled RNA transcripts (from 2 to 0.1 pmol in 1 ml of probe buffer) as described in Materials and Methods. Each lane corresponds to the transcripts generated as shown in panel A. The position of monomeric capsid protein (C) is shown. (C) Effect of RNA and capsid protein concentrations on the interaction between RNA and capsid protein. Capsid protein (20 to 0.2 ng, equivalent to 0.7 to 0.007 pmol) was spotted onto nitrocellulose and probed with labeled RNA transcripts. Lane 1, 487/1779 RNA, 0.5 pmol; lane 2, 1584/2714 RNA, 0.6 pmol; lanes 3 to 7, 330/617 RNA at 5, 2.5, 1, 0.25, and 0.05 pmol, respectively. (D) Dependence of RNA-capsid protein interaction on concentrations of RNA and capsid protein. The bound ^{35}S-labeled RNAs in the dot blots (panel C, lanes 3 to 7) were quantitated by cutting out the spots and scintillation counting. Capsid protein: 20 ng (0.7 pmol, ▲); 4 ng (0.14 pmol, ○); 2 ng (0.07 pmol, x); 1 ng (0.035 pmol, ●).

addition of 50 μCi of [^{35}S]CTP (New England Nuclear; 1290 Ci/mmol) in a 20-μl volume. The synthesized RNA was not capped. All synthesized RNA transcripts were analyzed for size and intactness by gel electrophoresis.

Northwestern (RNA-protein) blot analysis. RV (M33 strain) was grown in Vero cells and isolated from the culture supernatant (2). Viral proteins from purified virus particles were separated by preparative sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) under reducing conditions. Capsid protein was electroeluted and used as the antigen in the binding assay. Purified capsid protein suspended in RIPA buffer (50 mM Tris-HCl [pH 7.4], 1% Triton X-100, 150 mM NaCl, 0.1% SDS, 1% sodium deoxycholate, 10 mM EDTA) was separated by SDS-PAGE in a 10% polyacrylamide gel (13) and electrophoretically transferred onto a nitrocellulose membrane (31). The membrane strips were washed with probe buffer (10 mM Tris-HCl [pH 7.5], 50 mM NaCl, 1 mM EDTA, 1× Denhardt's solution) for 10 min at room temperature and further incubated with probe buffer containing *Escherichia coli* tRNA (250 μg/ml) for 1 h to block nonspecific binding. ^{35}S-labeled RNAs (10^6 cpm/μg) were added to the probe buffer and further incubated for 1 h with gentle shaking. Unbound radiolabeled probe was removed by three consecutive 10-min washes with probe buffer. The strips were dried, and the bound RNA was visualized by autoradiography.

Bal 31 deletion. 5′ and 3′ deletion mutants for mapping the essential RNA region that binds to capsid protein were constructed by using Bal 31 nuclease (14). The plasmid containing the *Sma*I-*Hin*cII fragment (residues 330 to 617) was linearized at the *Sma*I site (5′-end deletion) or the *Hin*cII site (3′-end deletion) and digested with Bal 31 nuclease (see Fig. 3A). In a typical reaction (100 μl), 2 μg of linearized plasmid DNA was incubated with Bal 31 (1.5 U) at 30°C, and

samples were withdrawn at 1, 2, 3, and 5 min into EGTA (ethylene glycol tetraacetic acid) solution (14). The DNA samples were treated with T4 DNA polymerase and recircularized by using T4 DNA ligase. Plasmids containing different lengths of deletion were digested with *Eco*RI and *Hin*dIII and analyzed by agarose gel electrophoresis to select the plasmids containing cDNAs of desired sizes.

PCR mutagenesis. Deletion of 31 nucleotides, from nucleotides 376 to 406, was carried out by PCR (11) with the mutagenic oligonucleotide pAGGGCGG TGGCGGCGACCCAGCGGCGC (complementary to nucleotides 417 to 406 and 375 to 361) and oligonucleotide primer 5′-ATTCGAATTCCAATGGAAG CTATCGGAC (the underlined nucleotides correspond to RV genomic RNA nucleotides 1 to 18). RV cDNA containing nucleotides 1 to 406 in vector M13mp19 was used as a template in mutagenesis. The conditions for PCR mutagenesis were four cycles at 95°C for 30 s, 25°C for 2 min, and 72°C for 1 min and 35 cycles at 95°C for 30 s and 65°C for 2 min. The amplified 375-nucleotide fragment was isolated from the agarose gel, treated with T4 DNA polymerase, digested with *Eco*RI, and inserted into the pSPT19 vector which had been treated with *Eco*RI and *Sma*I. The *Nco*I-*Sma*I fragment (nucleotides 38 to 330) was removed from the resulting plasmid, and the ends were filled in by repair and religated. The deletion was confirmed by DNA sequencing. This plasmid was named pSPT19Δ376/406.

Synthetic peptides and peptide-RNA binding assay. The synthesis of 11 overlapping peptides (C1 to C11, 26 to 34 residues in length) covering the capsid protein (residues 1 to 280, excluding the E2 signal peptide at the C terminus of the capsid protein) has been reported (22). RNA binding to peptides was performed with nitrocellulose (Hybond C; Amersham). Equal molar portions of each peptide were applied to the filter. The blots were incubated with probe buffer containing *E. coli* tRNA (250 μg/ml) for 1 h and then incubated with labeled RNA probes. Binding of RNA was carried out as described for Northwestern blot analysis.

Mobility shift analysis. Gel mobility shift analysis was performed as described by del Angel et al. (4). Briefly, the binding of RNA to synthetic peptide was carried out in a volume of 15 μl containing 10 mM HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid, pH 7.9), 20 mM NaCl, 150 mM KCl, 2 mM $MgCl_2$, 0.5 mM EDTA, 10% glycerol, 1 mM dithiothreitol, 120 ng of bovine serum albumin (BSA), 15 U of RNasin, peptide (as indicated in the legends), and 1 μg of *E. coli* tRNA. The reaction mixture was incubated on ice for 10 min before ^{35}S-labeled RNAs (1 pmol) which had been heated at 90°C for 2 min and cooled to room temperature were added. The mixtures were incubated for an additional 10 min on ice. For competition experiments, unlabeled RNAs were added to the reaction mixtures prior to the addition of labeled RNAs. The samples were loaded directly onto a nondenaturing 10% polyacrylamide gel, and electrophoresis was allowed to proceed for 2 h at a constant current of 15 mA (4). The gel was dried, and bound RNA was visualized by autoradiography.

RESULTS

Localization of the RNA region responsible for interaction with capsid protein. The capsid protein of RV interacts with the genomic RNA to form the icosahedral nucleocapsid. This interaction is specific; only the genomic RNA, not the subgenomic RNA, is encapsidated. In order to investigate the possibility that there is a specific binding site on the viral RNA, we studied the specificity of interaction between RNA and purified capsid protein immobilized on nitrocellulose. The cDNA fragment (6550 nucleotides) encoding the nonstructural proteins (Fig. 1A) was subcloned between the *Eco*RI and *Hin*dIII sites of the vector pSPT19. A series of ^{35}S-labeled RNA transcripts (3 to 10) were generated by cutting the cDNA within nonstructural genes at a *Bgl*II, *Sst*I, *Eco*RV, *Kpn*I, *Mlu*I, *Hin*cII, or *Sma*I site (Fig. 1A). RNA transcripts 1 and 2 were made from plasmid containing *Sma*I-*Hin*cII and *Mlu*I-*Hin*dIII fragments, respectively (Fig. 1A). These transcripts consisted of a nested set of RNAs differing in the length of the 3′ or 5′ sequences that they retained. Specific binding of RNA transcripts to capsid protein was detected by incubating the ^{35}S-labeled RNAs (2 to 0.1 pmol) with capsid protein (3 pmol) blotted onto nitrocellulose filters in Northwestern blot analysis (24). Specific binding was observed when the RNA transcripts contained nucleotides 330 to 617 (Fig. 1B). No binding was observed when this binding region was not present in the RNA transcripts (Fig. 1B, lanes 2 and 10). These data suggest that the sequence on the RNA that confers the binding specificity is located within the 288-nucleotide fragment between nucleotides 330 and 617 (Fig. 1A). A plasmid, pSPT19-330/617, containing this region (*Sma*I-*Hin*cII fragment) was constructed for synthesis of the RNA transcripts that were used in the experiments described below.

Since the RNA transcripts used in the binding assay differed in length, from 330 to 6550 nucleotides, the linearity of the RNA-capsid protein binding was determined by dot-blot analysis. Capsid protein (0.7 to 0.07 pmol per dot) was spotted onto nitrocellulose and probed with different concentrations (5 to 0.05 pmol/ml) of ^{35}S-labeled 330/617 RNA transcript. RNA transcripts (nucleotides 487 to 1779 or 1584 to 2714) were used as negative controls (Fig. 1C, lanes 1 and 2). The bound labeled RNAs (Fig. 1C, lanes 3 to 7) were quantitated by scintillation counting of the dot-blots. The interaction between capsid protein and RNA was found to be dependent on the concentrations of both RNA and capsid protein (Fig. 1D). No binding was observed with control RNA transcripts that did not contain the RNA sequence between nucleotides 330 and 617 (Fig. 1C, lanes 1 and 2).

Specificity of RNA-capsid protein interaction. To explore the specificity of the RNA-capsid protein interaction, we examined the binding of transcripts prepared in vitro from linearized plasmid pSPT19-330/617 in both the presence and absence of the same RNA in unlabeled form as a competitor. The binding of ^{35}S-labeled RNA to capsid protein was blocked by the presence of unlabeled RNA from the same source (Fig. 2A, lane 5), confirming that this RNA sequence (nucleotides 330 to 617) confers the specificity for RNA binding to capsid protein.

To determine whether this region (nucleotides 330 to 617) is actually a recognition site for capsid protein binding, we inserted the 288-bp *Sma*I-*Hin*cII cDNA fragment (Fig. 1A) into plasmid pSPT19-24S (2) immediately downstream from the start site for in vitro transcription. Binding activity to capsid protein was observed only in the modified 24S RNA containing this region and not in the unmodified 24S RNA (Fig. 2B), indicating that the binding specificity on the RNA is located within this region. This RNA binding to capsid protein was also found to occur when this region was inserted at the 3′ end of the 24S RNA (data not shown).

In RV, the nucleocapsid capsomere apparently consists of two disulfide-linked dimers of the capsid protein (6). It has been suggested that the central core of the RV virion may be composed of multiple dimeric units of capsid protein that associate with RV genomic RNA to form an icosahedral structure (32). In cells infected with RV, synthesized capsid protein spontaneously forms homodimers in the absence of other RV proteins. These dimers appears to become disulfide linked after virus budding or if exposed to an oxidizing environment (1). To determine if there is a difference in binding affinity between the dimeric and monomeric forms of capsid protein to viral RNA, we performed Northwestern blot analyses with reduced and nonreduced RV structural proteins immobilized to nitrocellulose. The blots were probed with ^{35}S-labeled RNA transcripts from nucleotides 330 to 617 or 1080 to 6550 (Fig. 1A). As expected, capsid protein but not E2 and E1 glycoproteins specifically bound to the 330/617 RNA transcript that contained the capsid protein binding site (Fig. 2C). No difference in binding affinity was observed between the dimeric and monomeric forms of capsid protein (Fig. 2C, lanes 1 and 2), indicating that the oligomeric state of the protein is not involved in its capacity to bind to RNA and that specificity in RNA binding is not dependent on the conformation of the capsid protein.

Mapping the binding site on viral RNA. The fine localization of the binding site was performed by Bal 31 nuclease deletion mapping. The plasmid containing the cDNA insert (nucleo-

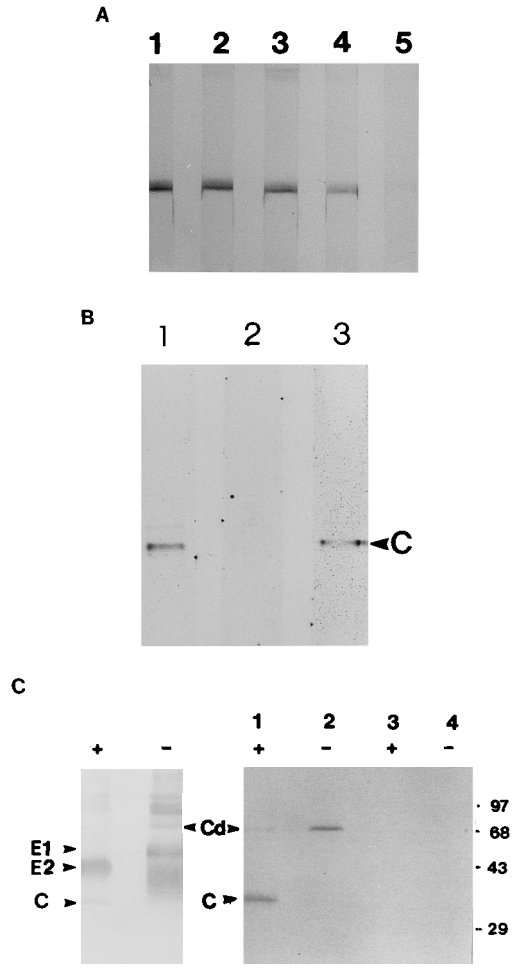

FIG. 2. Northwestern blot analysis of interaction between RNA and capsid protein. Immobilized capsid protein (3 pmol per strip) was used in the binding assay (A and B) as described in Materials and Methods. (A) Competition with unlabeled RNA transcript. Strips (0.5 cm) were incubated with ^{35}S-labeled 330/617 RNA (0.5 pmol) in the absence (lane 1) or presence (lanes 2 to 5) of various concentrations of unlabeled 330/617 RNA. The concentrations of competitor RNA were 4 (lane 2), 8 (lane 3), 16 (lane 4), and 32 (lane 5) pmol. The amount of input RNA recovered in lanes 1 to 5 was 95, 90, 78, 50, and 5%, respectively. (B) Effect of insertion of the 330/617 RNA fragment on the binding activity of 24S RNA to the capsid protein. The 288-bp *Sma*I-*Hin*cII fragment (corresponding to nucleotides 330 to 617) was inserted into the *Eco*RI site at the 5′ end of 24S cDNA. Modified and unmodified 24S RNAs were synthesized in the presence of [^{35}S]CTP and used as probes. Lane 1, labeled 330/617 RNA; lane 2, unmodified 24S RNA; lane 3, modified 24S RNA containing the 330/617 insert. The position of monomeric capsid protein (C) is shown. (C) Interaction of reduced and nonreduced capsid protein with viral RNA. Purified RV particles were treated (+) or not (−) with 2-mercaptoethanol prior to electrophoresis. Transferred blots were probed with labeled RNA transcripts of nucleotides 330 to 617 (right-hand panel, lanes 1 and 2) or 1080 to 6550 (right-hand panel, lanes 3 and 4). The RV structural proteins were detected with human anti-RV serum (shown on the left). The positions of protein molecular size standards are shown on the right (in kilodaltons). C, monomeric capsid protein; Cd, dimeric capsid protein. E1 and E2 refer to the structural proteins of RV. The electrophoretic mobilities of E1 and E2 are lower under nonreducing conditions.

tides 1 to 805) was linearized at the *Sma*I or *Hin*cII site and digested with Bal 31 nuclease to generate a series of mutant cDNAs differing in length at their respective 5′ and 3′ ends (Fig. 3A). Radiolabeled RNA transcripts were synthesized from selected clones and used as probes to test for their capacity to bind to capsid protein. The cDNA sequences of deletion mutants were analyzed to determine the endpoint of each deletion. The location of the endpoint of each deletion is shown as nucleotides numbered from the 5′ end of the RV genome (Fig. 3A). At the 5′ end (starting from nucleotide 330), downstream deletion of more than 46 nucleotides (mutant 376/617) resulted in a loss of binding activity (Fig. 3A, 5′-end deletion). Comparison between the endpoints of cDNA sequences and Northwestern blot analysis of truncated RNA transcripts suggest that the sequence exhibiting specific capsid protein binding is located between nucleotides 347 and 375 (Fig. 3A, 5′-end deletion). At the 3′ end (starting from nucleotide 617), upstream deletion of 211 nucleotides (mutant 330/406) did not affect the binding between RNA and capsid protein (Fig. 3A, 3′-end deletion), indicating that the RNA sequence that interacts with the capsid protein is located between nucleotides 330 and 406 (Fig. 3A, 3′-end deletion).

To further confirm that this RNA domain (nucleotides 347 to 375) is responsible for specific interaction with capsid protein, PCR mutagenesis was used to delete the RNA region from nucleotides 376 to 406. The construct (pSPT19Δ376/406) from PCR mutagenesis contained nucleotides 330 to 375 and 407 to 417 (total, 57 nucleotides). It is interesting that deletion of this 31-nucleotide fragment not only did not lead to any loss in capsid protein binding activity but instead enhanced its binding activity (Fig. 3B), indicating that the RNA domain that participates in the formation of the RNA-capsid protein complex is within the RNA sequence between nucleotides 347 and 375 (Fig. 3B). The RNA sequence in the region between nucleotides 376 and 406 can be folded into a base-paired structure by using the energy minimization program of Jaeger et al. (12) and Zuker (36) (Fig. 3C). The observed enhanced binding activity of the RNA to the capsid protein after deletion of the 31-nucleotide fragment (nucleotides 376 to 406) suggests that this stem-loop structure may obstruct the accessibility of the RNA binding site to the capsid protein or perturb an overall structure of the RNA binding site (Fig. 3C).

Identification of a domain in capsid protein that interacts with viral RNA. To examine the amino acid residues involved in capsid protein-RNA binding, we carried out peptide-RNA dot-blot analysis with 11 overlapping synthetic peptides (C1 to C11) covering the capsid protein from residues 1 to 280 (22). Peptides were spotted onto nitrocellulose and probed with ^{35}S-labeled RNA (77- or 57-nucleotide RNA transcript) (see Fig. 3B). An RNA transcript (nucleotides 417 to 617) was used as a negative control. At the level tested (65 pmol), C2 peptide interacted strongly with both RNA transcripts, whereas peptides C3, C10, and C11 interacted weakly with the 57-nucleotide RNA (Fig. 4A). Quantitation of the radioactivity by scintillation counting from the dot blot indicated that peptides C3, C10, and C11 bound 10, 5, and 15% of the labeled RNA bound by the C2 peptide, respectively. No binding was observed when the blot was probed with RNA transcript from nucleotides 417 to 617 (data not shown). The binding of peptide to RNA was compared with that of the capsid protein (2 pmol) in dot blots. It was found that the binding affinity of C2 peptide to RNA was lower than that of the capsid protein from the virion when the binding activity was compared on a molar basis (Fig. 4A), suggesting that although the other region(s) of the capsid protein does not confer specificity, it probably enhances interaction between the capsid protein and RNA.

In order to examine the specificity of RNA-peptide interaction, we carried out mobility shift experiments. Binding of the ^{35}S-labeled 57-nucleotide RNA (1 pmol) to synthetic peptides C1 to C11 (140 pmol) was carried out in solution in the presence of *E. coli* tRNA (250 μg/ml). Formation of a complex between the RNA and the peptide was analyzed by electrophoresis in a nondenaturing polyacrylamide gel (4). It was

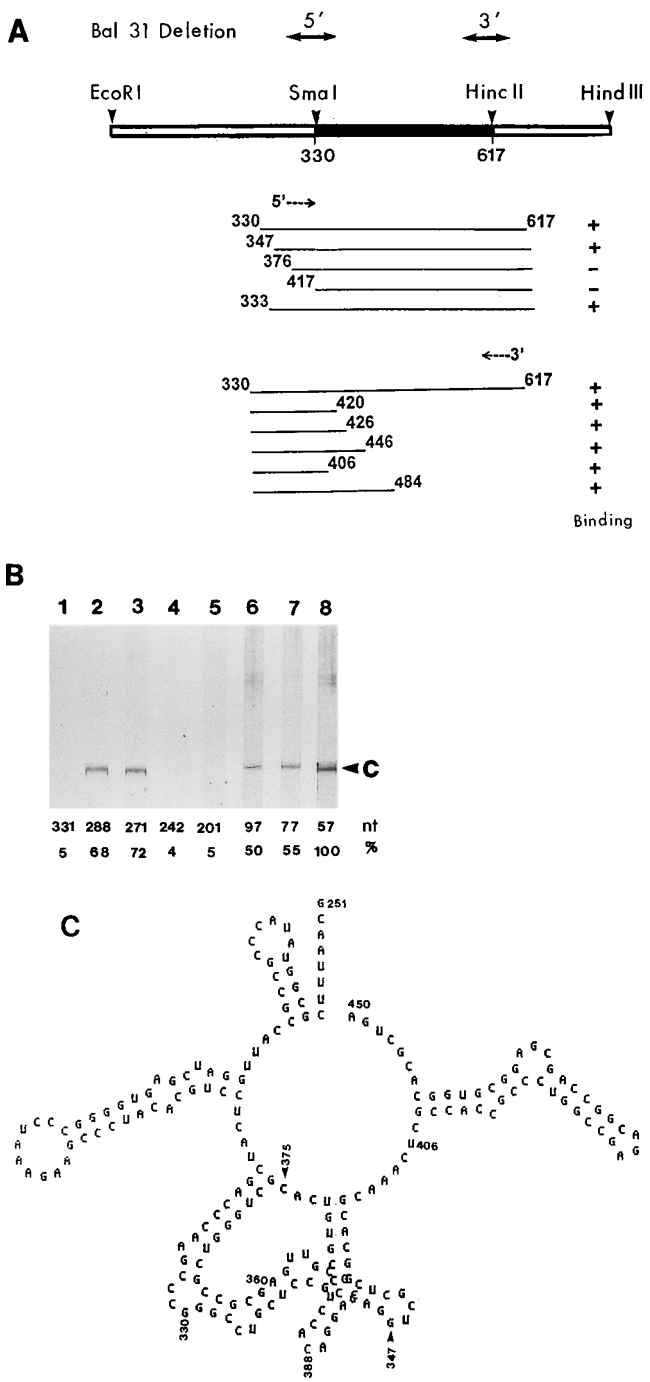

FIG. 3. Bal 31 deletion and analysis of the truncated RNA transcripts. (A) Schematic diagrams of Bal 31 deletion. The plasmid was linearized at the *Sma*I site (5′ deletion) or at the *Hin*cII site (3′ deletion) and digested with Bal 31 nuclease. Selected deletion plasmids were sequenced and used to prepare labeled RNA transcripts for Northwestern blot analysis. The location of the endpoint of each deletion is shown as nucleotides, numbered from the 5′ end of the RV genome. The binding activity of the truncated RNA is indicated (+, positive; −, negative). (B) Northwestern blot analysis of representative deletion RNA transcripts. Purified capsid protein (3 pmol per strip) immobilized on nitrocellulose was incubated with ^{35}S-labeled RNA transcripts (12 pmol in 1 ml of probe buffer) as described in Materials and Methods. RNA transcripts were from nucleotides 1 to 330 (lane 1), 330 to 617 (lane 2), 347 to 617 (lane 3), 376 to 617 (lane 4), 417 to 617 (lane 5), 330 to 426 (lane 6), 330 to 406 (lane 7), and 330 to 376 plus 407 to 417 (lane 8). The position of monomeric capsid protein (C) is shown. The length of each RNA transcript is shown in nucleotides (nt). The amount of bound ^{35}S-labeled RNA corresponding to each RNA transcript in the blots was quantitated by scintillation counting and normalized to that of the 57-nucleotide RNA, which was set at 100%. (C) Secondary structure of the region between nucleotides 251 and 450, predicted with the program of Jaeger et al. (12) and Zuker (36). The numbers present in this structure refer to Fig. 1A. The RNA sequence that interacts with the capsid protein is shown by the arrowhead.

found that the C2 and C3 peptides were retarded, whereas the C9, C10, and C11 peptides were minimally retarded (Fig. 4B, lanes 3, 4, 10, and 11). More than 90% of the input RNA was recovered in peptides C1, C4, C5, C6, C7, C8, and C9, while in peptides C2, C10, and C11, the recovered RNA was about 65%. In contrast, with peptide C3, only 40% of the RNA was recovered.

The results from dot-blot analysis (Fig. 4A) and the mobility shift assay (Fig. 4B) suggested that the binding of RNA to peptides C3, C10, and C11 may be due to nonspecific protein-RNA interaction through the neutralization of the negative charge on the phosphate backbone. Therefore, we performed the gel shift assay at a lower peptide concentration. With 14 pmol of peptide, no mobility shift was observed for any peptide except C2, which was slightly shifted (Fig. 5A). The difference in mobility of the RNA-C2 peptide complex was found to be dependent on the concentration of the peptide (Fig. 5B). The dependence of the mobility shift on the C2 peptide concentration seems to suggest that interaction between RNA and peptide C2 may produce a complex of a different shape or different electrostatic properties, depending on the ratio of RNA to C2 peptide.

To further explore the specificity of RNA-C2 peptide interaction, we examined the binding of the ^{35}S-labeled 57-nucleotide RNA transcript in the presence and absence of competitor RNAs. In vitro transcripts of unlabeled RNA from the same plasmid and nucleotides 617 to 1050 were used as competitor RNAs in gel mobility shift assays. The binding of ^{35}S-labeled RNA to peptide C2 was blocked by the presence of unlabeled RNA from the same source (Fig. 5C, lanes 3 to 6), whereas the RNA from nucleotides 617 to 1050 was ineffective in inhibiting the binding (Fig. 5C, lanes 7 to 9), confirming that this RNA sequence (nucleotide 330 to 375) confers the specificity for RNA binding to peptide C2. We have also analyzed the interaction of the 57-nucleotide RNA transcript with capsid protein in mobility shift assays. The interaction between capsid protein and RNA resulted in the formation of a ribonucleoprotein complex with a retarded mobility. A 50-fold molar excess of the unlabeled RNA inhibited its formation (data not shown).

DISCUSSION

Viral capsid protein primarily provides a protective shell for the virus genome in its journey from one host cell to another. The results presented here show that there is a specific interaction between RV genomic RNA and its capsid protein immobilized on nitrocellulose and that the specificity of binding is associated with a short sequence at the 5′ terminus of the viral genome between nucleotides 330 and 617 (Fig. 1). Modified 24S RNA, into which the segment required for binding (nucleotides 330 to 617) was inserted, showed binding activity (Fig. 2B), suggesting that at least some of the sequence is involved in binding recognition. Deletion at either end of the 330-617 fragment by Bal 31 nuclease digestion has narrowed down the capsid protein binding site to 29 nucleotides between nucleotides 347 and 375 in the viral RNA (Fig. 3). Encapsidation signals have now been identified for a wide spectrum of RNA viruses. They range in size from about 30 for the bacteriophage R17 and QβRNAs (35) to 60 to 100 nucleotides for

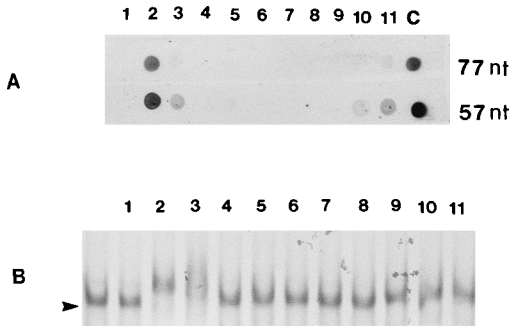

FIG. 4. Dot-blot analysis of peptide-RNA binding. (A) Synthetic peptides C1 to C11 (65 pmol per dot, lanes 1 to 11) and purified capsid protein (2 pmol per dot, lane C) were applied to nitrocellulose. Nitrocellulose strips were blocked with probe buffer containing *E. coli* tRNA (250 μg/ml) for 1 h. ^{35}S-labeled RNA (77 or 57 nucleotides; 24 pmol) was added to the probe buffer (1 ml) and further incubated for 1 h with gentle shaking. The blots were washed as described in Materials and Methods. (B) Polyacrylamide gel analysis of the complexes formed between RNA and peptide. ^{35}S-labeled 57-nucleotide RNA (1 pmol) was incubated with 140 pmol of peptide in the presence of *E. coli* tRNA (250 μg/ml) as described in Materials and Methods. Free RNA is indicated by the arrowhead on the left. Numbers 1 to 11 refer to synthetic peptides C1 to C11. The amount of input RNA and RNA-peptide complex in each lane was quantitated by scintillation counting. The amount of input RNA recovered for peptides C1 to C11 was 95, 65, 40, 93, 92, 94, 93, 92, 90, 62, and 64%, respectively.

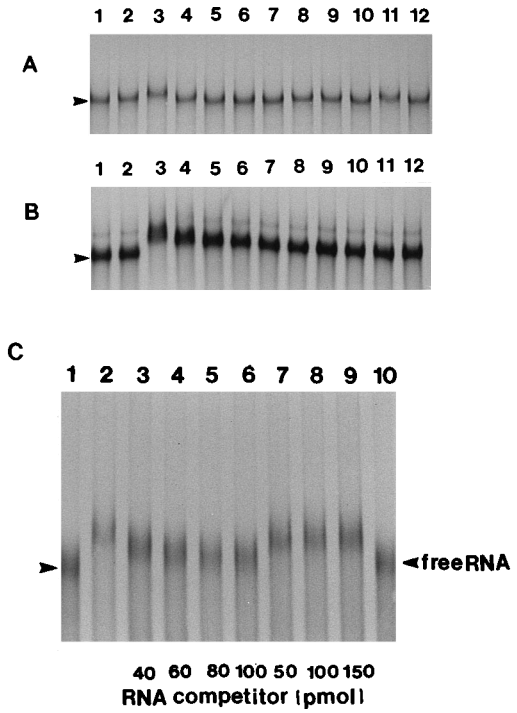

FIG. 5. RNA mobility shift analysis. (A) RNA-peptide interaction at lower peptide concentration. ^{35}S-labeled 57-nucleotide RNA (1 pmol) was incubated with 14 pmol of peptide as described in the legend to Fig. 4B. Lane 1, free RNA; lanes 2 to 12, peptides C1 to C11, respectively. (B) Effect of increasing the concentration of peptide C2 on complex formation. Assays were carried out with peptide C2 (lanes 3 to 11) at 140, 28, 7, 2, 0.7, 0.35, 0.2, 0.1, and 0.05 pmol, respectively. Lanes 1 and 12, free RNA; lane 1, peptide C1 (140 pmol). More than 90% of the input RNA was recovered in each lane. (C) RNA mobility shift in the presence of competitor RNAs. The assay was carried out at the indicated concentrations of unlabeled RNAs in the presence of 1 pmol of ^{35}S-labeled 57-nucleotide RNA, 140 pmol of peptide C2, and *E. coli* tRNA (250 μg/ml). Lanes 1 and 10, labeled RNA in the absence of C2 peptide; lane 2, in the presence of C2 peptide; lanes 3 to 6, in the presence of unlabeled 57-nucleotide transcript; lanes 7 to 9, in the presence of nonspecific RNA transcript (nucleotides 617 to 1050). The level of competitor RNA is shown. The position of unbound labeled RNA is indicated by arrowheads. More than 90% of the input labeled RNA was recovered in lanes 3 to 9. The recovered RNA in lane 2 was 75% of the input labeled RNA, as determined by scintillation counting.

Sindbis virus (9) and coronavirus (8). It is not known whether this stretch of 29 nucleotides is sufficient for RV genomic RNA packaging. It has been shown that RNAs of turnip crinkle virus (33) and Rous sarcoma virus (28) have multiple sites that interact to induce packaging. It would not be surprising if another region(s) of the RV genome is also necessary for RNA packaging.

Using 11 overlapping capsid peptides (C1 to C11), we have mapped the RNA-binding domain on capsid protein to residues 28 to 56 (C2; AAGASQSRRPRPPRHARAQHLPEMTP AVT) (Fig. 4). Capsid proteins from a number of RNA viruses have a common feature of arginine-rich regions, postulated to bind nucleic acids (9, 18, 25). Examination of the amino acid sequence of RV capsid protein suggested that the interaction of RV capsid protein with its viral RNA is not due solely to the high density of basic residues in this region (residues 28 to 56). Our results show that other highly basic regions, residues 205 to 233 (C9), 231 to 257 (C10), and 255 to 280 (C11), interact poorly with viral RNA. It is likely that the interaction between capsid protein and viral RNA is due to a nucleation event which confers specificity on the binding process and that other regions of capsid protein may then function in subsequent steps to lead to the formation of nucleocapsids. It has been suggested that the capsid protein in alphaviruses shows two types of RNA interactions: one major type of interaction which is nonspecific, and another which is specific for the encapsidation signal (7). Nonspecific protein-RNA interactions have been widely found for viral RNAs. It has been suggested that a lack of specificity may aid in virion assembly by compaction of the RNA through the neutralization of the negative charge on the phosphate backbone.

Although peptide C2 was found to interact specifically with RV viral RNA, its lower RNA binding affinity compared with that of the intact capsid protein suggests that another region(s) of the capsid protein may be necessary for efficient RNA binding. At present, we cannot rule out the possibility that regions in peptides C3, C10, and C11 are not involved in the RNA binding. Further studies are needed to use peptides including both C2 and C3 or C10 and C11 regions in the RNA-binding assay.

Cloning of the cDNAs of the Sindbis virus genome (23) and its smaller defective interfering RNA (15) and studies on the biological activity of their RNA transcripts have made it possible to define sequences in these genomes that are required for recognition in replication and encapsidation. Studies are in progress to use the RV cDNA expression system to examine the functions of the identified RNA segment and capsid protein domain in genomic encapsidation and assembly of RV.

ACKNOWLEDGMENTS

This work was supported by a grant from the Medical Research Council of Canada. S. Gillam is an investigator of the B.C.'s Children's Hospital Foundation. Z. Qiu was a predoctoral trainee receiving support from the B.C.'s Children's Hospital Foundation.

REFERENCES

1. **Baron, M., and K. Forsell.** 1991. Oligomerization of the structural proteins of rubella virus. Virology **185:**811–819.
2. **Clarke, D., T. Loo, I. Hui, P. Chong, and S. Gillam.** 1987. Nucleotide sequence and *in vitro* expression of rubella virus 24 S subgenomic mRNA encoding the structural proteins E1, E2 and C. Nucleic Acids Res. **15:**3041–3057.
3. **Clarke, D. M., T. W. Loo, H. McDonald, and S. Gillam.** 1988. Expression of

rubella virus cDNA coding for the structural proteins. Gene **65:**23–30.
4. **del Angel, R. M., A. G. Papavassilian, C. Fernandez-Tomas, S. J. Silverstein, and V. R. Racaniello.** 1989. Cell protein binds to multiple sites within the 5′-untranslated region of poliovirus RNA. Proc. Natl. Acad. Sci. USA **86:** 8299–8303.
5. **Dominguez, G., C. Y. Wang, and T. K. Frey.** 1990. Sequence of the genomic RNA of rubella virus: evidence for genetic rearrangement during togavirus evolution. Virology **177:**225–238.
6. **Dorsett, P. H., D. C. Miller, K. Y. Green, and F. I. Byrd.** 1985. Structure and function of the rubella virus proteins. Rev. Infect. Dis. **7**(Suppl. 1)**:**S150–S157.
7. **Forsell, K., M. Suomalainen, and H. Garoff.** 1995. Structure-function relation of the NH_2-terminal domain of the Semliki Forest virus capsid protein. J. Virol. **69:**1556–1563.
8. **Fosmire, J. A., K. Hwang, and S. Makino.** 1992. Identification and characterization of a coronavirus packaging signal. J. Virol. **66:**3522–3530.
9. **Geigenmuller-Gnirke, U., H. Nitschko, and S. Schlesinger.** 1993. Deletion analysis of the capsid protein of Sindbis virus: identification of the RNA binding region. J. Virol. **67:**1620–1626.
10. **Hahn, C. S., and J. H. Strauss.** 1990. Site-directed mutagenesis of the proposed catalytic amino acids of the Sindbis virus capsid protein autoprotease. J. Virol. **64:**3069–3073.
11. **Horton, R. M., and L. R. Pease.** 1991. Recombination and mutagenesis of DNA sequence using PCR, p. 217–247. *In* M. J. McPherson (ed.), Directed mutagenesis: a practical approach. IRL Press, Oxford.
12. **Jaeger, J. A., D. H. Turner, and M. Zuker.** 1989. Improved predictions of secondary structures for RNA. Proc. Natl. Acad. Sci. USA **86:**7706–7710.
13. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–685.
14. **Legerski, R. J., J. L. Hodnett, and H. B. Gray, Jr.** 1978. Extracellular nucleases of *Pseudomonas* Bal 31. III. Use of the double-strand deoxyribonuclease activity as the basis of a convenient method for the mapping of fragments of DNA produced by cleavage with restriction enzymes. Nucleic Acids Res. **5:**1445–1452.
15. **Levis, R., B. C. Weiss, M. Tsing, H. Huang, and S. Schlesinger.** 1986. Deletion mapping of Sindbis virus DI RNAs derived from cDNAs defines the sequence essential for replication and packaging. Cell **44:**137–145.
16. **Matthews, R. E. F.** 1982. Classification and nomenclature of viruses: third report of the international committee on taxonomy of viruses. Intervirology **17:**1–199.
17. **McDonald, H., T. C. Hobman, and S. Gillam.** 1991. The influence of capsid protein cleavage on the processing of E2 and E1 glycoproteins of rubella virus. Virology **183:**52–60.
18. **Nassal, M.** 1992. The arginine-rich domain of the hepatitis B virus core protein is required for pregenome encapsidation and productive viral positive-strand DNA synthesis but not for virus assembly. J. Virol. **66:**4107–4116.
19. **Oker-Blom, C.** 1984. The gene order for rubella virus structural proteins is NH_2-C-E2-E1-COOH. J. Virol. **51:**354–358.
20. **Oker-Blom, C., N. Kalkkinen, L. Kaariainen, and R. F. Pettersson.** 1983. Rubella virus contains one capsid protein and three envelope glycoproteins, E1 and E2a and E2b. J. Virol. **46:**964–973.
21. **Oker-Blom, C., I. Ulmanen, L. Kaariainen, and R. F. Pettersson.** 1984. Rubella virus 40 S genome RNA specifies a 24 S subgenomic mRNA that codes for a precursor to structural proteins. J. Virol. **49:**403–408.
22. **Ou, D., P. Chong, B. Tripet, and S. Gillam.** 1992. Analysis of T- and B-cell epitopes of capsid protein of rubella virus by using synthetic peptides. J. Virol. **66:**1674–1681.
23. **Rice, C. M., R. Levis, J. H. Strauss, and H. V. Huang.** 1987. Production of infectious RNA transcripts from Sindbis virus cDNA clones: mapping of lethal mutations, rescue of a temperature-sensitive marker, and in vitro mutagenesis to generate defined mutants. J. Virol. **61:**3809–3819.
24. **Robbins, S. G., M. F. Frana, J. J. McGowan, J. F. Boyle, and K. V. Holmes.** 1986. RNA-binding proteins of MHV: detection of monomeric and multimeric N protein with an RNA overlay-protein blot assay. Virology **150:**402–410.
25. **Sacher, R., and P. Ahlquist.** 1989. Effect of deletions in the N-terminal basic arm of brome mosaic virus coat protein on RNA packaging and systemic infection. J. Virol. **63:**4545–4552.
26. **Simmons, D. T., and J. H. Strauss.** 1974. Translation of Sindbis virus 26 S and 49 S RNA in lysates of rabbit reticulocytes. J. Mol. Biol. **86:**397–409.
27. **Soderlund, H., K. Takkinen, A. Jalanko, and N. Kalkkinen.** 1985. The expression and organization of the alphavirus genome, p. 323–337. *In* Y. Becker (ed.), Viral messenger RNA. Martinus Nijhoff, Boston.
28. **Sorge, J., R. Ricci, and S. H. Hughes.** 1983. *cis*-acting RNA packaging locus in the 115-nucleotide direct repeat of Rous sarcoma virus. J. Virol. **48:**667–675.
29. **Stohlman, S., R. Barric, G. Nelson, L. H. Soe, L. M. Welter, and R. J. Deans.** 1988. Specific interaction between coronavirus leader RNA and nucleocapsid protein. J. Virol. **62:**4288–4295.
30. **Suomalainen, M., H. Garoff, and M. Baron.** 1990. The E2 signal peptide sequence of rubella virus remains part of the capsid protein and confers membrane association in vitro. J. Virol. **64:**5500–5509.
31. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA **76:**4350–4354.
32. **Waxham, M. N., and J. S. Wolinsky.** 1985. A model of the structural organization of rubella virions. Rev. Infect. Dis. **7**(Suppl. 1)**:**S133–S139.
33. **Wei, N., L. A. Heaton, T. J. Morris, and S. C. Harrison.** 1990. Structure and assembly of turnip crinkle virus. VI. Identification of coat protein virus binding sites on the RNA. J. Mol. Biol. **214:**85–95.
34. **Weiss, B., H. Nitschko, I. Ghattas, R. Wright, and S. Schlesinger.** 1989. Evidence for specificity in encapsidation of Sindbis virus RNAs. J. Virol. **63:**5310–5318.
35. **Witherell, G. W., and O. C. Uhlenbeck.** 1989. Specific RNA binding by Qβ coat protein. Biochemistry **28:**71–76.
36. **Zuker, M.** 1989. On finding all suboptimal folding of an RNA molecule. Science **244:**48–52.