

# Interaction of Vesicular Stomatitis Virus P and N Proteins: Identification of Two Overlapping Domains at the N Terminus of P That Are Involved in N<sup>0</sup>-P Complex Formation and Encapsidation of Viral Genome RNA<sup>∇</sup>

Mingzhou Chen, Tomoaki Ogino, and Amiya K. Banerjee\*

*Department of Molecular Genetics, Section of Virology, Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio 44195*

Received 7 June 2007/Accepted 21 September 2007

The nucleocapsid (N) protein of nonsegmented negative-strand (NNS) RNA viruses, when expressed in eukaryotic cells, aggregates and forms nucleocapsid-like complexes with cellular RNAs. The phosphoprotein (P) has been shown to prevent such aggregation by forming a soluble complex with the N protein free from cellular RNAs (designated N<sup>0</sup>). The N<sup>0</sup>-P complex presumably mediates specific encapsidation of the viral genome RNA. The precise mechanism by which the P protein carries out this function remains unclear. Here, by using a series of deleted and truncated mutant forms of the P protein of vesicular stomatitis virus (VSV), Indiana serotype, we present evidence that the N-terminal 11 to 30 amino acids (aa) of the P protein are essential in keeping the N protein soluble. Furthermore, glutathione S-transferase fused to the N-terminal 40 aa by itself is able to form the N<sup>0</sup>-P complex. Interestingly, the N-terminal 40-aa stretch failed to interact with the viral genome N-RNA template whereas the C-terminal 72 aa of the P protein interacted specifically with the latter. With an *in vivo* VSV minigenome transcription system, we further show that a deletion mutant form of P (PΔ1-10) lacking the N-terminal 10 aa which is capable of forming the N<sup>0</sup>-P complex was unable to support VSV minigenome transcription, although it efficiently supported transcription *in vitro* in a transcription-reconstitution reaction when used as purified protein. However, the same mutant protein complemented minigenome transcription when expressed together with a transcription-defective P deletion mutant protein containing N-terminal aa 1 to 210 (PΔII+III). Since the minigenome RNA needs to be encapsidated before transcription ensues, it seems that the entire N-terminal 210 aa are required for efficient genome RNA encapsidation. Taking these results together, we conclude that the N-terminal 11 to 30 aa are required for N<sup>0</sup>-P complex formation but the N-terminal 210 aa are required for genome RNA encapsidation.

Vesicular stomatitis virus (VSV) is a negative-strand RNA virus belonging to the family *Rhabdoviridae*. Its genome RNA is enwrapped in a helical form with the nucleocapsid (N) protein and packaged within the virion together with the RNA-dependent RNA polymerase large protein (L) and the phosphoprotein (P). During cell infection, the associated RNA polymerase transcribes the genome RNA into five mRNAs that are subsequently translated into the glycoprotein (G), the matrix (M) protein, and the N, P, and L proteins. The N protein is an RNA-binding protein since it forms a high-molecular-weight complex with cellular RNAs with a quasihelical structure (1, 16) when expressed from plasmids in bacterial, mammalian, or insect cells. In contrast, it specifically encapsidates the viral genome RNA during the course of infection. It has been suggested that the P protein acts as a specific chaperone and forms a soluble complex with the N protein and prevents the latter from concentration-dependent aggregation (11, 19, 23, 25). The N-P complex, once formed, prevents the N protein from binding to cellular RNAs, keeps it in an en-

capsidation-competent form, and provides specificity to encapsidate the viral RNA during replication. This soluble complex, a vital component of replication containing RNA-free N protein, is designated N<sup>0</sup>-P (8, 11, 24, 31). Both *in vitro* and *in vivo* studies have also shown that P forms specific complexes with N<sup>0</sup> with various stoichiometries (11, 22, 24). Thus, the P protein of VSV, and presumably of other NNS RNA viruses, serves three important functions in RNA synthesis. (i) It structurally stabilizes L protein (14, 17, 28, 34) to bind to the N-RNA template to form the active L-P<sub>2</sub> holoenzyme (13); for VSV, phosphorylation of P protein is essential to form this active holoenzyme (30). (ii) P protein forms a specific complex with N<sup>0</sup> and prevents N protein from binding to cellular RNAs, which is an essential step for replication since newly formed N<sup>0</sup>-P complex initiates and encapsidates viral cRNA and newly synthesized genome RNA (14, 18, 19, 25, 31, 35). Interestingly, unlike formation of an active L-P<sub>4</sub> complex, phosphorylation of P protein is not required for N<sup>0</sup>-P complex formation (36). (iii) For VSV, the P protein forms a tripartite complex with the L and N proteins to form the replicase complex to transcribe the genome and antigenome RNAs end to end with concomitant encapsidation of RNA by the N protein elicited from the N<sup>0</sup>-P complex (32). Although it is known that the N and C termini of the P protein are involved in binding to N<sup>0</sup> and N-RNA template, respectively (36), it is not known whether N<sup>0</sup>-P complex

\* Corresponding author. Mailing address: Virology Section, Department of Molecular Genetics, Lerner Research Institute, Cleveland Clinic, Cleveland, OH 44195. Phone: (216) 444-0625. Fax: (216) 444-2998. E-mail: banerja@ccf.org.

<sup>∇</sup> Published ahead of print on 3 October 2007.

formation is necessary and sufficient for encapsidation of viral RNA. Here, we demonstrate that N<sup>0</sup>-P complex formation is necessary but a defined N-terminal region of the P protein directly regulates encapsidation of viral genome RNA.

## MATERIALS AND METHODS

**Materials.** The anti-Myc polyclonal antibody (Ab) (sc789), the monoclonal Ab (sc-40), and the anti-glyceraldehyde-3-phosphate dehydrogenase (GAPDH) monoclonal Ab (sc-32233) used were from Santa Cruz Biotechnology, Inc. The anti-glutathione S-transferase (GST) polyclonal Ab used and protein G-Sepharose 4 Fast Flow were from Amersham Biosciences. The anti-HA (clone HA-7) and anti-Flag monoclonal Abs used were from Sigma, and the anti-six-His monoclonal Ab (ab18184) used was from Abcam Inc. Anti-VSV N, P, and L polyclonal Abs were raised in rabbits and supplied by Invitrogen. ATP, GTP, UTP, and CTP were obtained from Roche Molecular Biochemical. [ $\alpha$ -<sup>32</sup>P]GTP (3,000 Ci/mmol) was from Perkin-Elmer.

**Cells and viruses.** HeLa and BHK-21 cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum. Recombinant vaccinia virus (vTF7-3) carrying the bacteriophage T7 RNA polymerase gene was propagated as described by Pattnaik et al. (30).

**Plasmid constructs.** HA-tagged wild-type P, P $\Delta$ 1-40, phosphorylation site mutant protein P3A (S60A, T62A, S64A), P $\Delta$ Hinge, P $\Delta$ 161-210, P $\Delta$ II+III, and P $\Delta$ 76-137; Myc-tagged wild-type N and P; and Flag-tagged L, pBS-N, pBS-P, pBS-L, and pVSV-CAT2 have been described previously (7). cDNAs encoding other mutant P proteins were amplified with pET3a-P (2) as the template and cloned in frame with the open reading frame (ORF) encoding the GAL4 activating domain and a hemagglutinin (HA) tag into pGADT7 (Clontech). cDNA encoding P $\Delta$ 1-10 was also cloned in frame with the ORF encoding the GAL4 DNA binding domain and a Myc tag into pGBKT7 (Clontech). GST-fused P peptides were amplified with pGEX-6P3 and pET3a-P as templates, respectively. The resultant PCR products were mixed as templates, and reamplified cDNAs encoding GST-fused peptides were cloned in frame with the ORF encoding the GAL4 activating domain and an HA tag into pGADT7. All constructs were verified by DNA sequencing.

**Detection of soluble N protein and in vivo coimmunoprecipitation.** A plasmid encoding the N protein was transfected alone or cotransfected with plasmids encoding wild-type P, mutant P proteins, or GST-fused peptides with Lipofectamine 2000 (Invitrogen) into HeLa cells in six-well plates, followed by vTF7-3 infection. After 24 h of transfection, cell lysates were washed once with phosphate-buffered saline and lysed in lysis buffer (6). Lysates were centrifuged at 13,000 rpm for 30 min at 4°C to separate the supernatant and pellet. The soluble N proteins in supernatants in all subsequent studies and only aggregated N proteins in pellets were analyzed by Western blotting (WB) with anti-Myc monoclonal Ab. To detect interaction between the L and P proteins, Flag-tagged L and HA-tagged P protein or mutant P proteins were coexpressed. Supernatants from lysates were subjected to immunoprecipitation with anti-L polyclonal Ab as described above. Bound proteins were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and detected by WB with monoclonal anti-Myc, anti-HA, or anti-Flag Abs.

**Separation of N<sup>0</sup>-P and N-RNA complexes by ultracentrifugation.** A 0.5- $\mu$ g sample of a plasmid encoding Myc-tagged N was transfected alone or cotransfected with increasing amounts of plasmids (0.25  $\mu$ g, 0.5  $\mu$ g, 1.0  $\mu$ g, and 2.0  $\mu$ g) encoding HA-tagged GST-1-40 (PN40) or GST-192-265 (PC72), described above. Cell pellet was incubated in 2.4 ml of lysis buffer (20 mM Tris-HCl [pH 8.0], 100 mM NaCl, 1.9% [wt/vol] Triton X-100, 1 mM 1,4-dithiothreitol, 5% [vol/vol] glycerol) for 30 min on ice. Lysates were centrifuged as described above, and supernatants were layered onto 1.6 ml 20% (vol/vol) glycerol in TE (20 mM Tris-HCl [pH 8.0], 1 mM EDTA) and ultracentrifuged further in a Sorvall TH-660 rotor at 35,000 rpm for 4 h at 4°C to separate the supernatant (U-sup) and pellet (U-ppt). The U-ppt was resuspended in 50  $\mu$ l of TE containing 10% (vol/vol) glycerol. Both the U-sup and U-ppt were boiled in Laemmli buffer and detected by WB with monoclonal anti-Myc Ab to detect N protein.

**N-RNA template purification and GST pull-down assays.** Preparation of purified N-RNA template was done as described by Ogino and Banerjee (29). HA-tagged GST, PN40, and PC72 were expressed in HeLa cells as described above. Cells were treated with lysis buffer provided with the ProFound pull-down GST protein-protein interaction kit (Pierce) and incubated for 30 min in reverse transcriptase. After centrifugation at 13,000 rpm for 30 min, equal amounts of supernatants were mixed with 200 ng purified N-RNA template. GST pull-down assays were performed with the ProFound pull-down GST protein-protein interaction kit according to the manufacturer's protocol. Proteins in supernatants

and eluted proteins from GST pull-down beads were detected by WB with anti-HA, anti-GST, or anti-N Abs, followed by gel electrophoresis.

**In vitro transcription of the VSV genome RNA.** The recombinant wild-type P and mutant P proteins (P $\Delta$ 1-10, P $\Delta$ 11-20, P $\Delta$ 21-30, and P $\Delta$ 31-40) were expressed as hexahistidine-tagged forms in Sf21 insect cells with the Bac-to-Bac baculovirus expression system (Invitrogen), purified with Ni-nitrilotriacetic acid agarose resin beads (QIAGEN), and used for in vitro transcription-reconstitution reactions with purified N-RNA (75 ng) and recombinant L (20 ng) and P (40 ng) as described by Ogino and Banerjee (29). <sup>32</sup>P-labeled transcripts were analyzed by electrophoresis in a 5% polyacrylamide gel containing 7 M urea, followed by autoradiography.

**In vivo transcription of VSV minigenome RNA.** BHK-21 cells in six-well plates were infected with vTF7-3 and transfected with plasmid pBS-N, pBS-L, or pVSV-CAT2 and a plasmid encoding HA-tagged P or mutant proteins. In the complementation assay, a plasmid encoding Myc-tagged P or P $\Delta$ 1-10 or a plasmid encoding HA-tagged P mutant protein was cotransfected with plasmid pBS-N, pBS-L, or pVSV-CAT2 (7). After 24 h, cell lysates were subjected to a chloramphenicol acetyltransferase (CAT) enzyme-linked immunosorbent assay (ELISA) for detection of the CAT expression level as described by Chen et al. (7). The expression levels of the wild-type and mutant P proteins were detected by WB with anti-HA and anti-Myc Abs.

## RESULTS

**The N-terminal 11 to 30 aa of the P protein are essential in keeping the N protein soluble.** First, we constructed a series of mutant P cDNAs that, when expressed in HeLa cells, produced a battery of mutated P proteins (Fig. 1) that were used in all subsequent experiments. To directly map the region within the N-terminal domain of the P protein that interacts with the N protein to form the N<sup>0</sup>-P complex, we coexpressed the Myc-tagged N protein and the HA-tagged, N terminus truncated P protein (P $\Delta$ 1-40) in HeLa cells as described in Materials and Methods. The supernatants and pellets that were obtained by centrifugation at 13,000 rpm represented the soluble fraction and aggregated or RNA-bound N protein, respectively (12). As shown in Fig. 2A, when expressed alone, approximately 70% of the N protein was found in the pellet fraction (Fig. 2A, lane 2) as determined from the WB analyses, whereas about 70% of the N protein was found in the supernatant fraction along with P protein when it was coexpressed with the P protein (Fig. 2A, lane 3). In contrast, about 75% of the N protein remained in the pellet fraction when it was coexpressed with P $\Delta$ 1-40, suggesting that the N-terminal 40 aa are required for N<sup>0</sup>-P formation (Fig. 2A, lanes 2 and 3). Additional mutant P proteins were then coexpressed with the Myc-tagged N protein and compared with the observed effect of P $\Delta$ 1-40 on N protein solubility. As shown in Fig. 2B, HA-tagged P3A, P $\Delta$ Hinge, P $\Delta$ 161-210, and P $\Delta$ II+III (Fig. 1) all were able to keep more than 53% of the N protein soluble (lanes 4 to 7), confirming that the N-terminal 40 aa are indeed required for its solubility and that phosphorylation of P protein is not required for N<sup>0</sup>-P complex formation. To further narrow down the region in P which can carry out this function, HA-tagged P $\Delta$ 1-20 and P $\Delta$ 1-10 (Fig. 1) were coexpressed with the N protein and the results show that P $\Delta$ 1-10 was also able to keep about 70% of the N protein soluble, which is similar to the wild-type P protein, but P $\Delta$ 1-20 was only able to keep 30% of the N protein soluble (Fig. 2C, lanes 4 and 5), indicating that residues 11 to 20 are essential for N protein solubility. To confirm this and test whether residues 21 to 40 are also required to keep the N protein soluble, we constructed three additional deletion mutant proteins, P $\Delta$ 11-20, P $\Delta$ 21-30, and P $\Delta$ 31-40, tagged with HA (Fig. 1) and coexpressed them with the N protein. As shown in

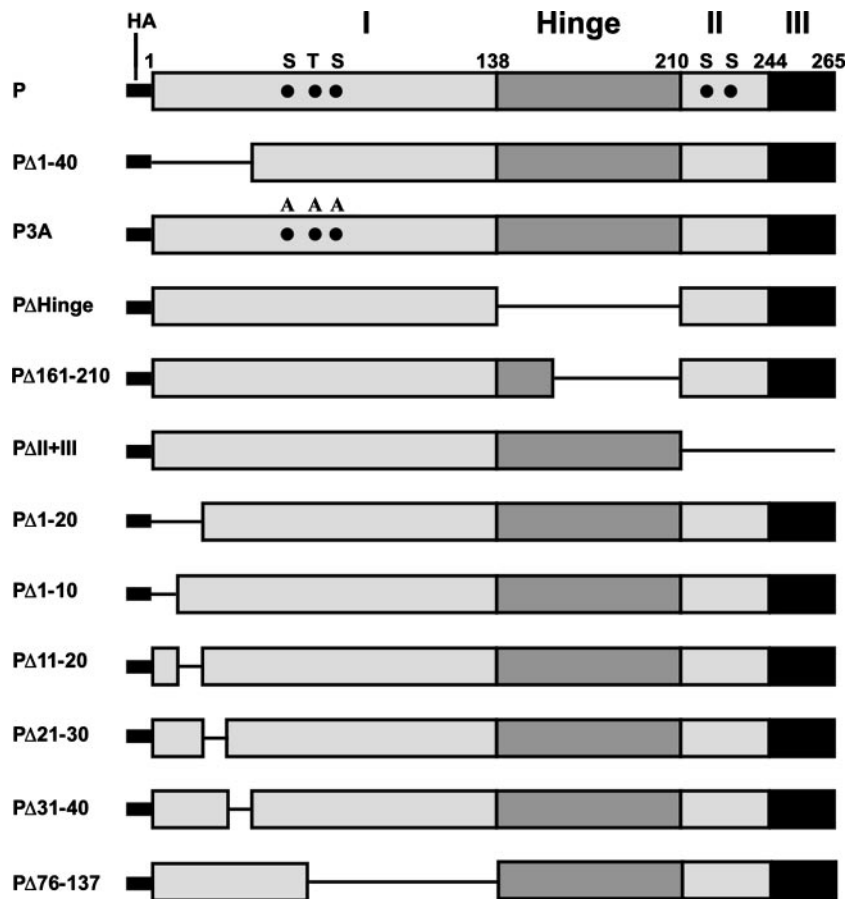


FIG. 1. Schematic representation of the wild-type P protein and the mutant P proteins used in this study. The wild-type P protein is divided into four regions designated I (aa 1 to 137), Hinge (aa 138 to 210), II (aa 211 to 244), and III (aa 245 to 265). The letters S at position 60, T at position 62, S at position 64, S at position 226, and S at position 227 represent the phosphorylation sites of the P protein. Wild-type P and the mutant P proteins were fused with an HA tag. The numbers represent amino acid positions.

Fig. 2D, P $\Delta$ 11-20 and P $\Delta$ 21-30 were able to keep only 36% and 40% of the N protein soluble, respectively, whereas P $\Delta$ 31-40 was similar to the wild-type P protein (lanes 4 to 6) and was able to keep 70% of the N protein soluble. If these results are taken together, it seems that N-terminal residues 11 to 30 of the P protein are essential for N<sup>0</sup>-P complex formation.

**The N-terminal 40 aa of the P protein by themselves are capable of forming the N<sup>0</sup>-P complex.** Next, we wanted to find out whether expressing only the N-terminal 40 aa will render the N protein soluble. Initially, we attempted to express the N-terminal 60 aa to avoid possible degradation of a smaller peptide of 40 aa in cells. However, we failed to detect any expression of the peptide in HeLa cells. We then constructed three chimeric proteins with two N-terminal peptides and one C-terminal peptide fused with GST and an HA tag as shown in Fig. 3A. HA-tagged GST, GST-1-40 (PN40), GST-1-60 (PN60), and GST-192-265 (PC72) were efficiently expressed when transfected into HeLa cells (data not shown). When these fusion peptides were coexpressed with the Myc-tagged N protein and the supernatants were analyzed for N protein by WB with anti-Myc Ab, both PN40 and PN60 were able to keep at least 70% of the N protein soluble whereas GST and PC72 were only able to keep 35% of the N soluble (Fig. 3B). From

these results, it seems that both PN40 and PN60 could form N<sup>0</sup>-P complexes whereas PC72 possibly interacts with the N-RNA complex only as shown by Green et al. (15). To further confirm that PN40 binds to N<sup>0</sup> and prevents the N protein to bind to cellular RNAs, increasing amounts of plasmids expressing PN40 or PC72 were coexpressed with the N protein and supernatant containing the soluble N protein after centrifugation at 13,000 rpm was further centrifuged at 35,000 rpm (see Materials and Methods) and analyzed for N protein by WB. When expressed alone, at least 80% of the N protein was pelleted during ultracentrifugation (U-ppt), which clearly indicated that it was bound to cellular RNAs (Fig. 3C, lane 1). We cannot, however, rule out the possibility that some small aggregated N complex also remained in the pellet. On the other hand, with increasing concentrations of PN40, nearly 90% of the N protein shifted to the U-sup fractions following ultracentrifugation, indicating that PN40 prevents N protein from binding to cellular RNAs (Fig. 3C, lanes 2 to 5). In contrast, almost 80% of the N protein remained in the pellet fraction with increasing concentrations of PC72 (Fig. 3C, lanes 6 to 9), indicating that PN40 fails to interact with the N-RNA complex whereas PC72 probably does (14). To directly demonstrate that PC72 indeed binds to the N-RNA complex, we

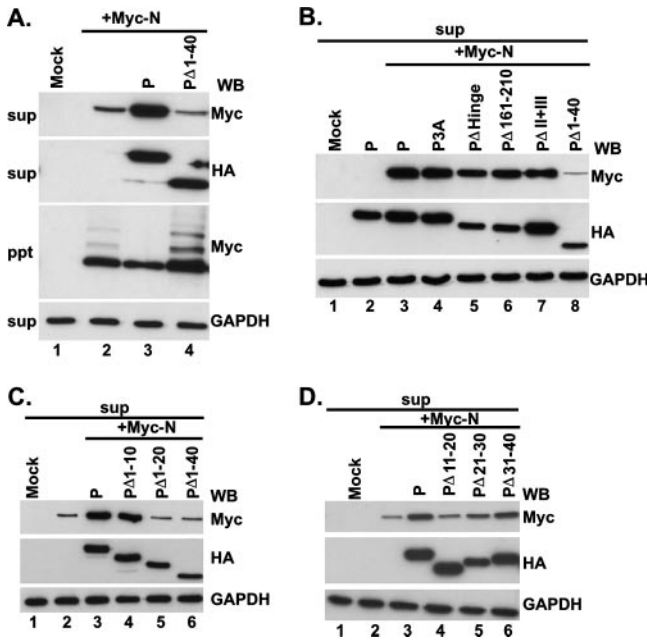


FIG. 2. Amino acids 11 to 30 of the P protein are essential to keep N protein soluble. Myc-tagged N protein and HA-tagged P protein or mutant P proteins were coexpressed in HeLa cells as indicated. Supernatants (sup) from centrifugation at 13,000 rpm (A, B, C, and D) were detected by WB with anti-Myc, anti-HA, and anti-GAPDH monoclonal Abs. GAPDH was used as a loading control. (A) The pellet (ppt) was also detected by WB with an anti-Myc monoclonal Ab.

used an N-RNA template purified from VSV virions (Fig. 4A and B) and studied its interaction with PN40 and PC72. Equal amounts of HeLa cell lysates expressing GST, PN40, and PC72 were mixed with 200 ng of purified N-RNA template for GST pull-down assays. As shown in Fig. 4C, PC72 was able to pull down the N-RNA template (lane 3, right side) but GST and PN40 were not (lanes 1 and 2, right side) although they had similar levels of expression (left side). Thus, it appears that PC72 bound to N protein prebound to RNA and PN40 interacted specifically with N<sup>0</sup> when both were coexpressed with N in HeLa cells.

**The N-terminal 40 aa of the P protein are not sufficient for VSV minigenome transcription in vivo, and the N-terminal 210 aa are required to make the minigenome competent for transcription.** Next, we wanted to determine whether all of the N-terminal 40 aa are also required for VSV genome RNA transcription. First, we studied whether the mutant protein PΔ1-10, PΔ11-20, PΔ21-30, or PΔ31-40 would interact with the L protein to form an active RNA polymerase complex when coexpressed in HeLa cells. As shown in Fig. 5A, all of the mutant P proteins efficiently interacted with the L protein, since Ab-recognizing L protein brought down the wild-type P protein and all of the mutant P proteins (lanes 3 to 7). The expression levels of the P and L proteins were similar in the two cases (Fig. 5B), indicating that the RNA polymerase holoenzyme complex was formed in each case. We then tested all four mutant proteins for the ability to transcribe genome RNA in *in vitro* transcription-reconstitution assays. His-tagged mutant P proteins His-PΔ1-10, His-PΔ11-20, His-PΔ21-30, and His-PΔ31-40 were efficiently expressed in insect cells with re-

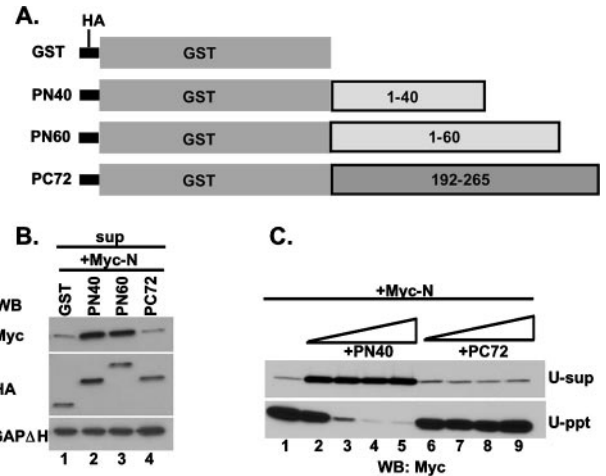


FIG. 3. The N-terminal 40 aa of the P protein are sufficient to keep N protein soluble. (A) Schematic representation of the N-terminal 40 and 60 aa and the C-terminal 72 aa fused with GST and an HA tag. Numbers represent amino acid positions. (B) Myc-tagged N and HA-tagged GST, GST-1-40 (PN40), GST-1-60 (PN60), or GST-192-265 (PC72) were coexpressed in HeLa cells as indicated. Supernatants (sup) from centrifugation at 13,000 rpm were detected by WB with anti-Myc, anti-HA, and anti-GAPDH monoclonal Abs. GAPDH was used as a loading control. (C) PN40 binds to N<sup>0</sup>, and PC72 binds to the N-RNA complex. Myc-tagged N protein was expressed alone or coexpressed with increasing concentrations of PN40 or PC72. Supernatants from centrifugation at 13,000 rpm were further centrifuged at 35,000 rpm (see Materials and Methods). N<sup>0</sup> and the N-RNA complex were detected in U-sup or U-ppt by WB with anti-Myc Ab.

combinant baculovirus expressing the mutant proteins as described in Materials and Methods and were purified (Fig. 6A). The purified P proteins were then used in transcription-reconstitution reactions *in vitro* with purified L protein and the N-RNA template. As shown in Fig. 6B, with the N-RNA template alone or in the presence of L protein only, no transcription was observed (lanes 1 and 2) whereas the transcribed mRNAs (G, N, and P/M) were clearly detected in the presence of L and P, as well as with the four mutant proteins (lanes 3 to 7). The relative transcription activities of His-PΔ1-10, His-

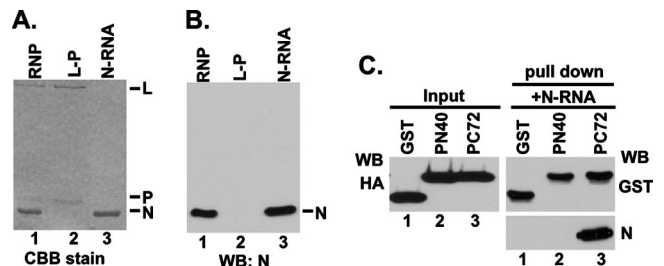


FIG. 4. The N-terminal 40-aa region is unable to bind to the N-RNA template *in vitro*. Purified RNP, N-RNA, and L-P complexes were analyzed by SDS-PAGE, followed by Coomassie brilliant blue (CBB) staining (A), and detected by WB with an anti-N polyclonal Ab (B). (C) PC72 interacted with the N-RNA template *in vitro*, but PN40 did not. Equal amounts of HeLa cell lysates expressing GST, PN40, or PC72 were detected by WB with anti-HA Ab (left side), and the same lysates were mixed with 200 ng of the purified N-RNA template in GST pull-down assays. Eluted GST beads were assayed by WB with anti-GST and anti-N Abs.

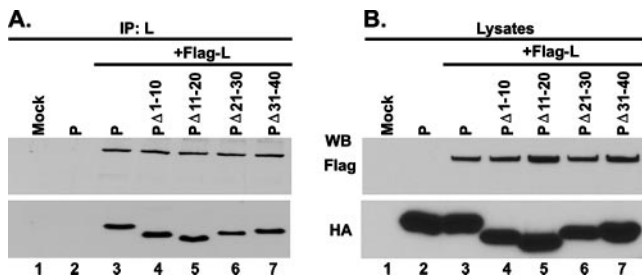


FIG. 5. P $\Delta$ 1-10, P $\Delta$ 11-20, P $\Delta$ 21-30, and P $\Delta$ 31-40 all interact with L protein in vivo. HA-tagged wild-type P protein or mutant P proteins were coexpressed with Flag-tagged L protein (Flag-L) in HeLa cells as indicated. Protein expression was detected in lysates by WB with anti-HA and anti-Flag Abs (B). Immunoprecipitation (IP) was performed with anti-L polyclonal Ab, and immune complexes were detected by WB with anti-HA and anti-Flag Abs (A).

P $\Delta$ 11-20, His-P $\Delta$ 21-30, and His-P $\Delta$ 31-40 were 103%, 88%, 72%, and 48% of that of wild-type P, which was taken as 100% (Fig. 6C). These results clearly demonstrated that all of the mutant proteins were active in supporting VSV genome transcription in vitro and that the N-terminal 30 aa are dispensable; His-P $\Delta$ 31-40, however, showed only 48% activity. To test whether these mutant proteins are active in transcription in vivo, we used the VSV minigenome transcription assay system previously used (7), as detailed in Materials and Methods.  $\nu$ TF7-3-infected BHK-21 cells were transfected with plasmids pVSV-CAT2, pBS-N, and pBS-L and a plasmid encoding HA-tagged P or the four mutant proteins, and cell lysates were subjected to a CAT ELISA for the detection of CAT expression levels. Surprisingly, all of the mutant proteins, including P $\Delta$ 1-10 and P $\Delta$ 31-40, which form the N<sup>0</sup>-P complex (Fig. 2), were unable to support VSV minigenome transcription in vivo (Fig. 6D, top). To detect transcription of the minigenome, it needs to be encapsidated by the N protein before transcription ensues; these results demonstrate that N<sup>0</sup>-P complex formation is necessary but not sufficient to encapsidate the genome RNA in vivo. All of the mutant proteins are transcriptionally active in vitro because they interact with the preformed N-RNA template and L protein. To address the possibility that regions beyond the N-terminal 40 aa are required to encapsidate VSV genome RNA, we used Myc-tagged P $\Delta$ 1-10, which is inactive in supporting minigenome transcription in vivo, and several mutant P proteins tagged with HA and performed a functional complementation assay. Myc-tagged wild-type P protein was used as a positive control.  $\nu$ TF7-3-infected BHK-21 cells were transfected with plasmids pVSV-CAT2, pBS-N, and pBS-L and a plasmid encoding Myc-P, Myc-P $\Delta$ 1-10, or HA-tagged mutant proteins P $\Delta$ II+III, P $\Delta$ 161-210, and P $\Delta$ 76-137 (Fig. 1). As shown in Fig. 7A, all of the mutant proteins were unable to support minigenome transcription compared to the Myc-P protein (lanes 1 to 6) at similar levels of expression (Fig. 7B). However, when plasmids encoding Myc-P $\Delta$ 1-10 and HA-tagged P $\Delta$ II+III or P $\Delta$ 161-210 were coexpressed in the presence of other supporting plasmids, the relative CAT expression levels were 82.3% and 40.7%, respectively (lanes 7 and 8), strongly suggesting that an additional region of the P protein is necessary for encapsidation of minigenome RNA. It is interesting that in the complementation with Myc-P $\Delta$ 1-10 and HA-

P $\Delta$ 76-137, no CAT expression was detected (Fig. 7A, lane 9), which indicated that aa 76 to 137 are also required for the encapsidation of VSV genome RNA. We have previously demonstrated that P $\Delta$ 138-160 is active in supporting VSV minigenome transcription in vivo (7). Taken together, the findings indicate that the entire N-terminal 137 aa and the self-association domain (residues 161 to 210) are required for efficient encapsidation of VSV genome RNA.

## DISCUSSION

The N protein of VSV, as well as all viruses belonging to the NNS RNA virus family, plays a pivotal role in the life cycle of the virus. Not only does it encapsidate the genome RNA (and render it inaccessible to RNase action) which serves as the template for both transcription and replication to occur, but it also associates with the P protein to form a soluble N<sup>0</sup>-P complex, which is the required intermediate that provides the N protein to encapsidate newly replicated genome RNA. Additionally, the complex also interacts with the L protein to form the replicase complex that transcribes the genome or antigenome RNA end to end (32). It has been proposed that the putative replicase holoenzyme complex, i.e., L-N-P, initiates replication, which is followed by concomitant encapsidation of nascent RNA by N<sup>0</sup>-P (32). Thus, formation of the N<sup>0</sup>-P complex is crucial in the production of functional genomic and antigenomic templates. Several studies have been done previously to gain insight into the interaction of P protein with the N protein to form the N<sup>0</sup>-P complex, in particular, the roles of the N and C termini of the P protein in this process (36). However, whether formation of the N<sup>0</sup>-P complex is sufficient to encapsidate the genome RNA or there exists a separate region in P that is additionally required to regulate the N protein for encapsidation has not been investigated.

To address this possibility, we used a battery of mutant P proteins and studied their interactions with N protein expressed in HeLa cells. Our primary criteria were based on the fact the N protein, expressed alone, binds to cellular RNAs (5, 20, 21, 33, 27) and remains in the pellet fraction when the lysates are centrifuged (12). In contrast, when coexpressed with the wild-type P protein, a major fraction, if not all, of N is rendered soluble and RNA free (8, 11, 24, 25, 31) and is found in the supernatant fraction. By these criteria, we have shown that the N-terminal 11 to 30 aa of the P protein, Indiana serotype, are required to keep N protein soluble (Fig. 2C and D). Importantly, the N-terminal peptide PN40 by itself is sufficient to form the N<sup>0</sup>-P complex (Fig. 3 and 4), whereas the C-terminal 72 aa of the P protein interact only with N protein that is already bound to RNA, i.e., the N-RNA complex (Fig. 4) (15). Since encapsidation of genome RNA by N protein is vital for subsequent transcription or replication of the N-RNA template, we used this criterion to further probe the presence of a putative encapsidation-facilitating domain within the P protein that would help the N<sup>0</sup>-P complex to carry out this function. With an in vivo minigenome transcription system, we demonstrated for the first time that N<sup>0</sup>-P complex formation is necessary but not sufficient to encapsidate the genome RNA. We showed that the mutant P protein P $\Delta$ 1-10, which is perfectly capable of forming the N<sup>0</sup>-P complex and highly active in an in vitro transcription assay, is incapable of carrying out

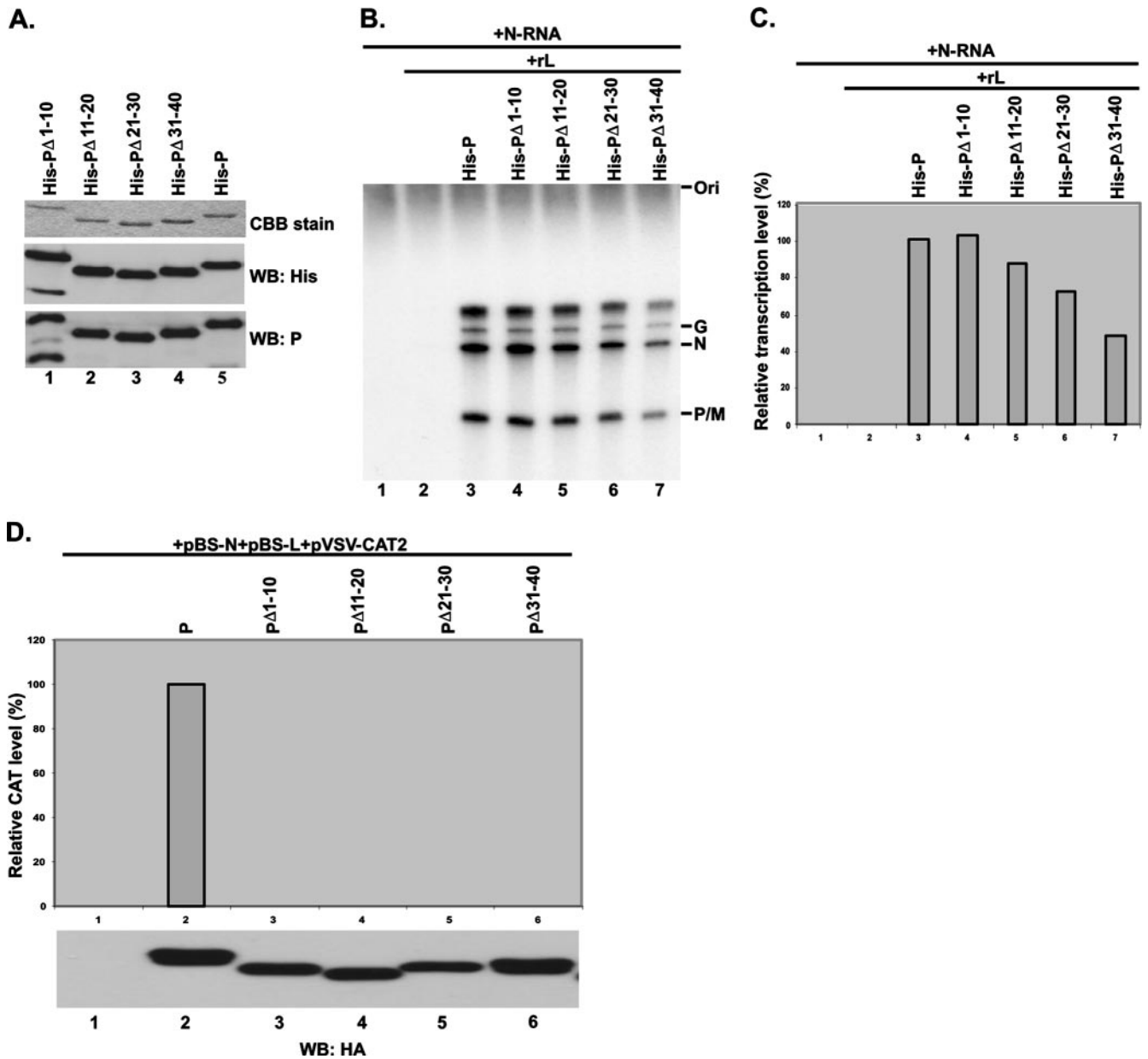


FIG. 6. The N-terminal 40-aa region is required for encapsidation of viral genome RNA. (A) Purified, recombinant, His-tagged wild-type P protein and mutant P proteins were analyzed by SDS-PAGE, followed by Coomassie brilliant blue (CBB) staining, and by WB with anti-His and anti-P Abs. (B) In vitro transcription reconstituted with an N-RNA template and recombinant P protein or mutant P proteins and L protein. The recombinant P protein or mutant P proteins and L protein were subjected to in vitro transcription with the N-RNA template. <sup>32</sup>P-labeled transcripts were analyzed by urea-5% PAGE, followed by autoradiography. The positions of the viral mRNAs are indicated on the right. (C) The relative transcription activities (percent) of wild-type P protein and mutant P proteins in transcription reactions are indicated. The relative transcription activity of wild-type P protein was defined as 100%. (D) Roles of wild-type P protein and mutant P proteins in VSV minigenome transcription in vivo. BHK cells expressing T7 polymerase were transfected with pBS-N, pBS-L, pVSV-CAT2, and a plasmid encoding HA-tagged wild-type P protein or mutant P proteins as indicated. A CAT ELISA was performed with lysates of the transfected cells to measure relative CAT expression levels as viral minigenome transcription activities. The relative CAT expression level of the wild-type P protein was defined as 100%. The expression levels of the wild-type P protein and mutant P proteins were detected by WB with an anti-HA monoclonal Ab.

minigenome transcription in vivo (Fig. 6B and D). In striking contrast, when a mutant protein lacking domains II and III (P $\Delta$ II+III, which is transcription negative but N<sup>0</sup>-P positive) can efficiently complement the transcription function of P $\Delta$ 1-10 when added in *trans* (Fig. 7). These results strongly suggest that the encapsidation-facilitating domain resides be-

yond the N-terminal 40 aa, which is the minimum region required for N<sup>0</sup>-P complex formation. All of the N-terminal 210 aa (P $\Delta$ II+III) containing the self-association domain (aa 161 to 210) seem to be essential for transcription complementation to occur since deletion of aa 161 to 210 reduced transcription complementation to 40%. Furthermore, deletion of aa 76 to

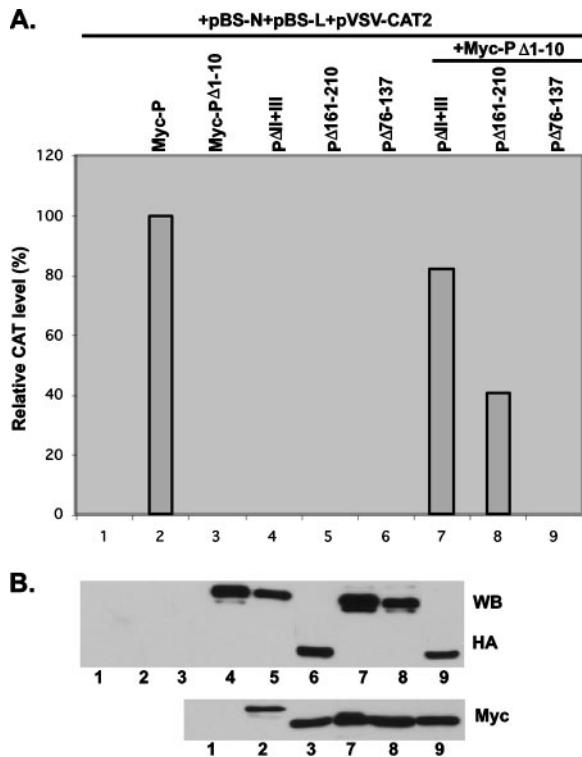


FIG. 7. Complementation assay of the ability of mutant proteins to support the transcription of the VSV minigenome *in vivo*. BHK cells expressing T7 polymerase were transfected with pBS-N, pBS-L, pVSV-CAT2, and a plasmid encoding the wild-type P protein or mutant P proteins as indicated. (A) A CAT ELISA was performed with lysates of transfected cells to measure relative CAT expression levels as viral minigenome transcription activities. The relative CAT expression level of the wild-type P protein was defined as 100%. (B) Expression levels of HA-tagged mutant proteins and the Myc-tagged P protein and PΔ1-10 were detected by WB with anti-HA and anti-Myc monoclonal Abs, respectively.

137 resulted in complete abrogation of transcription complementation, suggesting that the entire acidic domain is also critical to the encapsidation reaction. The precise boundaries of the N-terminal domain covering aa 1 to 137 required for encapsidation, however, need to be further investigated by additional deletion experiments. Taken together, it seems that the N-terminal 210 aa contain two overlapping domains, one N-terminal 40-aa domain for N<sup>0</sup>-P complex formation (NP40) and the other for encapsidation of genome RNA (NP1-210). It will be of interest to examine whether it is possible to rescue infectious virus by expressing PΔ1-10 and P1-210 as two independent genes which would confirm our minigenome-based assay. It is noteworthy that NP1-210 contains the self-association domain of P (aa 161 to 210). Thus, it seems that within the N<sup>0</sup>-P complex, the self-association domain in P must be masked by the N protein to prevent it from oligomerizing and consequently binding to L to form a transcription complex (L-P<sub>2</sub>). The L protein, on the other hand, may bind directly to the N<sup>0</sup>-P complex to form the putative replicase. The concerted action of the replicase complex and the N<sup>0</sup>-P complex may trigger the transcription-encapsidation reaction. The precise domain(s) within the N-terminal 1 to 210 aa of the P

protein directly involved in the interaction with the L protein and the formation of an encapsidation-competent N<sup>0</sup>-P complex is currently under investigation.

A challenging question that remains is how the N<sup>0</sup>-P complex confers specificity to encapsidate only its cognate genome RNA. For example, the P protein of human parainfluenza virus type 3 (HPIV3) fails to form an N<sup>0</sup>-P complex with the N protein of VSV, and conversely, PN40 of VSV fails to form an N<sup>0</sup>-P complex with the N protein of HPIV3 (unpublished observations). These findings clearly demonstrate that the N and P proteins of NNS RNA viruses interact via a specific amino acid sequence or recognition fold, as well as recognize a specific *cis*-acting RNA sequence motif at the 5' end of the respective genome RNAs. Structural-biology studies would certainly provide insight into these interactive processes. Two recent publications of the crystal structure of N-RNA complexes of VSV (16) and rabies virus (1) purified from bacteria and insect cells, respectively, following expression of the N protein demonstrated that the RNA is tightly sequestered in a cavity at the interface between N protein lobes and shield it from the environment. The structure, however, does not reveal any sequence specificity in RNA recognition. It is noteworthy that a recent study of N<sup>0</sup>-P formation in Chandipura virus, which also belongs to the *Rhabdoviridae* family, demonstrated that the N protein in a monomeric form promotes specific viral genome RNA recognition. Once N oligomerizes or aggregates, it interacts with cellular RNA even in the presence of P (4). The P protein also interacts with leader RNA in its unphosphorylated form and initiates nucleocapsid assembly onto a nascent RNA chain (3). Thus, the interaction between P and leader RNA provides the specificity that successively recruits N protein onto RNA while itself being replaced. Whether VSV P also can keep N in a monomeric form and can recognize viral leader RNA in its unphosphorylated form needs to be studied.

N-P interaction studies with other NNS viruses show some similarity to our findings. It was shown for Sendai virus that the N-terminal half of the P protein is essential for N<sup>0</sup>-P complex formation but not for mRNA synthesis, and later studies demonstrated that this region could be further narrowed down to aa 33 to 41 (9, 10). The N<sup>0</sup>-P complex-forming domain in P of HPIV3 has been mapped within the N-terminal 40 aa (12). A similar study with rabies virus P protein by Mavrikakis et al. (26) has recently shown that a peptide containing residues 4 to 40 remained associated with N within the N<sup>0</sup>-P complex following sequential protease digestion of the complex, confirming the requirement of this region for N<sup>0</sup>-P formation only. However, in each case, the precise region(s) of the P protein that is required for genome RNA encapsidation has not been determined.

In summary, we propose a model in which the P protein of VSV acts as a chaperone that keeps the N protein soluble and free from cellular RNA by interaction with a specific recognition motif in N<sup>0</sup> through the N-terminal 40 aa (aa 11 to 30 are essential). For efficient encapsidation of the viral genome RNA, the N-terminal 137 aa and the self-association domain (aa 161 to 210) are required. However, the exact N-terminal boundary between aa 41 and 137 needed for encapsidation remains to be determined. Additional mutations are required to address this question. Once the viral genome RNA is transcribed for replication mediated by the replicase, the encapsidation

ation of nascent RNA is initiated by the N protein from the N<sup>0</sup>-P complex via recognition of a specific *cis*-acting RNA motif present in the 5' end of the nascent RNA. The P protein probably dissociates from the N<sup>0</sup>-P complex during the encapsidation process or after it is complete. The P protein may subsequently bind to the L protein to form the RNA polymerase holoenzyme complex and bind to the newly formed N-RNA template via the C-terminal 72 aa or a shorter region of P to carry out its function as the RNA polymerase cofactor.

#### ACKNOWLEDGMENT

This work was supported by a grant from the National Institutes of Health (AI26585) to A.K.B.

#### REFERENCES

- Albertini, A. A., A. K. Wernimont, T. Muziol, R. B. Ravelli, C. R. Clapier, G. Schoehn, W. Weissenhorn, and R. W. Ruigrok. 2006. Crystal structure of the rabies virus nucleoprotein-RNA complex. *Science* **313**:360–363.
- Barik, S., and A. K. Banerjee. 1991. Cloning and expression of the vesicular stomatitis virus phosphoprotein gene in *Escherichia coli*: analysis of phosphorylation status versus transcriptional activity. *J. Virol.* **65**:1719–1726.
- Basak, S., T. Raha, D. Chattopadhyay, A. Majumder, M. S. Shaila, and D. J. Chattopadhyay. 2003. Leader RNA binding ability of Chandipura virus P protein is regulated by its phosphorylation status: a possible role in genome transcription-replication switch. *Virology* **307**:372–385.
- Bhattacharya, R., S. Basak, and D. J. Chattopadhyay. 2006. Initiation of encapsidation as evidenced by deoxycholate-treated nucleocapsid protein in the Chandipura virus life cycle. *Virology* **349**:197–211.
- Bhella, D., A. Ralph, L. B. Murphy, and R. P. Yeo. 2002. Significant differences in nucleocapsid morphology within the Paramyxoviridae. *J. Gen. Virol.* **83**:1831–1839.
- Chen, M., J. C. Cortay, I. R. Logan, V. Sapountzi, C. N. Robson, and D. Gerlier. 2005. Inhibition of ubiquitination and stabilization of human ubiquitin E3 ligase PIRH2 by measles virus phosphoprotein. *J. Virol.* **79**:11824–11836.
- Chen, M., T. Ogino, and A. K. Banerjee. 2006. Mapping and functional role of the self-association domain of vesicular stomatitis virus phosphoprotein. *J. Virol.* **80**:9511–9518.
- Curran, J., H. Homann, C. Buchholz, S. Rochat, W. Neubert, and D. Kolakofsky. 1993. The hypervariable C-terminal tail of the Sendai paramyxovirus nucleocapsid protein is required for template function but not for RNA encapsidation. *J. Virol.* **67**:4358–4364.
- Curran, J., R. Boeck, N. Lin-Marq, A. Lupas, and D. Kolakofsky. 1995. Paramyxovirus phosphoproteins form homotrimers as determined by an epitope dilution assay, via predicted coiled coils. *Virology* **214**:139–149.
- Curran, J., T. Pelet, and D. Kolakofsky. 1994. An acidic activation-like domain of the Sendai virus P protein is required for RNA synthesis and encapsidation. *Virology* **202**:875–884.
- Davis, N. L., H. Arnheiter, and G. W. Wertz. 1986. Vesicular stomatitis virus N and NS proteins form multiple complexes. *J. Virol.* **59**:751–754.
- De, B. P., M. A. Hoffman, S. Choudhary, C. C. Huntley, and A. K. Banerjee. 2000. Role of NH<sub>2</sub>- and COOH-terminal domains of the P protein of human parainfluenza virus type 3 in transcription and replication. *J. Virol.* **74**:5886–5895.
- Ding, H., T. J. Green, S. Lu, and M. Luo. 2006. Crystal structure of the oligomerization domain of the phosphoprotein of vesicular stomatitis virus. *J. Virol.* **80**:2808–2814.
- Errington, W., and P. T. Emmerson. 1997. Assembly of recombinant Newcastle disease virus nucleocapsid protein into nucleocapsid-like structures is inhibited by the phosphoprotein. *J. Gen. Virol.* **78**:2335–2339.
- Green, T. J., S. Macpherson, S. Qiu, J. Lebowitz, G. W. Wertz, and M. Luo. 2000. Study of the assembly of vesicular stomatitis virus N protein: role of the P protein. *J. Virol.* **74**:9515–9524.
- Green, T. J., X. Zhang, G. W. Wertz, and M. Luo. 2006. Structure of the vesicular stomatitis virus nucleoprotein-RNA complex. *Science* **578**:357–360.
- Horikami, S. M., and S. A. Moyer. 1995. Alternative amino acids at a single site in the Sendai virus L protein produce multiple defects in RNA synthesis in vitro. *Virology* **211**:577–582.
- Horikami, S. M., J. Curran, D. Kolakofsky, and S. A. Moyer. 1992. Complexes of Sendai virus NP-P and P-L proteins are required for defective interfering particle genome replication in vitro. *J. Virol.* **66**:4901–4908.
- Howard, M., and G. Wertz. 1989. Vesicular stomatitis virus RNA replication: a role for the NS protein. *J. Gen. Virol.* **70**:2683–2694.
- Iseni, F., A. Barge, F. Baudin, D. Blondel, and R. W. Ruigrok. 1998. Characterization of rabies virus nucleocapsids and recombinant nucleocapsid-like structures. *J. Gen. Virol.* **79**:2909–2919.
- Karlin, D., S. Longhi, and B. Canard. 2002. Substitution of two residues in the measles virus nucleoprotein results in an impaired self-association. *Virology* **302**:420–432.
- La Ferla, F. M., and R. W. Peluso. 1989. The 1:1 N-NS protein complex of vesicular stomatitis virus is essential for efficient genome replication. *J. Virol.* **63**:3852–3857.
- Majumder, A., S. Basak, T. Raha, S. P. Chowdhury, D. Chattopadhyay, and S. Roy. 2001. Effect of osmolytes and chaperone-like action of P-protein on folding of nucleocapsid protein of Chandipura virus. *J. Biol. Chem.* **276**:30948–30955.
- Masters, P. S., and A. K. Banerjee. 1988. Resolution of multiple complexes of phosphoprotein NS with nucleocapsid protein N of vesicular stomatitis virus. *J. Virol.* **62**:2651–2657.
- Masters, P. S., and A. K. Banerjee. 1988. Complex formation with vesicular stomatitis virus phosphoprotein NS prevents binding of nucleocapsid protein N to nonspecific RNA. *J. Virol.* **62**:2658–2664.
- Mavrakis, M., S. Mehous, E. Real, F. Iseni, D. Blondel, N. Tordo, and R. W. Ruigrok. 2006. Rabies virus chaperone: identification of the phosphoprotein peptide that keeps nucleoprotein soluble and free from non-specific RNA. *Virology* **349**:422–429.
- Mavrakis, M., L. Kolesnikova, G. Schoehn, S. Becker, and R. W. Ruigrok. 2002. Morphology of Marburg virus NP-RNA. *Virology* **296**:300–307.
- Mellon, M. G., and S. U. Emerson. 1978. Rebinding of transcriptase components (L and NS proteins) to the nucleocapsid template of vesicular stomatitis virus. *J. Virol.* **27**:560–567.
- Ogino, T., and A. K. Banerjee. 2007. Unconventional mechanism of mRNA capping by the RNA-dependent RNA polymerase of vesicular stomatitis virus. *Mol. Cell* **25**:85–97.
- Pattnaik, A. K., L. Hwang, T. Li, N. Englund, M. Mathur, T. Das, and A. K. Banerjee. 1997. Phosphorylation within the amino-terminal acidic domain I of the phosphoprotein of vesicular stomatitis virus is required for transcription but not for replication. *J. Virol.* **71**:8167–8175.
- Peluso, R. W. 1988. Kinetic, quantitative, and functional analysis of multiple forms of the vesicular stomatitis virus nucleocapsid protein in infected cells. *J. Virol.* **62**:2799–2807.
- Qanungo, K. R., D. Shaji, M. Mathur, and A. K. Banerjee. 2004. Two RNA polymerase complexes from vesicular stomatitis virus-infected cells that carry out transcription and replication of genome RNA. *Proc. Natl. Acad. Sci. USA* **101**:5952–5957.
- Schoehn, G., F. Iseni, M. Mavrakis, D. Blondel, and R. W. Ruigrok. 2001. Structure of recombinant rabies virus nucleoprotein-RNA complex and identification of the phosphoprotein binding site. *J. Virol.* **75**:490–498.
- Smallwood, S., K. W. Ryan, and S. A. Moyer. 1994. Deletion analysis defines a carboxyl-proximal region of Sendai virus P protein that binds to the polymerase L protein. *Virology* **202**:154–163.
- Spehner, D., R. Drillien, and P. M. Howley. 1997. The assembly of the measles virus nucleoprotein into nucleocapsid-like particles is modulated by the phosphoprotein. *Virology* **232**:260–268.
- Takacs, A. M., T. Das, and A. K. Banerjee. 1993. Mapping of interacting domains between the nucleocapsid protein and the phosphoprotein of vesicular stomatitis virus by using a two-hybrid system. *Proc. Natl. Acad. Sci. USA* **90**:10375–10379.