Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 8873–8877, December 1986
Biochemistry

Identification of a domain within the phosphoprotein of vesicular stomatitis virus that is essential for transcription *in vitro*

(cloning in pGEM vector/transcription–translation/transcription–reconstitution/deletion mapping)

DALIP S. GILL*, DHRUBAJYOTI CHATTOPADHYAY, AND AMIYA K. BANERJEE

Roche Institute of Molecular Biology, Roche Research Center, Nutley, NJ 07110

Communicated by Herbert Weissbach, August 8, 1986

ABSTRACT **A full-length cDNA copy of the phosphoprotein (NS) mRNA of vesicular stomatitis virus (New Jersey serotype) was inserted into pGEM4 vector downstream of the promoter for bacteriophage SP6 RNA polymerase. Transcription of the cDNA *in vitro* resulted in the synthesis of NS mRNA, which was subsequently translated into NS protein in a cell-free rabbit reticulocyte system. The biological activity of the expressed NS protein was demonstrated by *in vitro* synthesis of mRNA by transcription–reconstitution with purified viral L protein and N-RNA template. Deletion mapping of the *NS* gene defined a specific domain between amino acid residues 213 and 247, which was essential for *in vitro* transcription. Removal of the COOH-terminal 21 amino acids, on the other hand, did not have a significant effect on transcription. This domain appears to be involved in efficient binding of NS protein to the N protein–RNA template.**

The phosphoprotein (NS) of vesicular stomatitis virus (VSV) is an essential polypeptide component of the RNA polymerase complex that is responsible for viral genome RNA transcription both *in vitro* and *in vivo* (1, 2). Within the purified virion, the NS protein (29 kDa) is complexed with the larger polymerase component L protein (241 kDa). This complex is tightly associated with the nucleocapsid protein N (49 kDa)-bound genome RNA template to form the transcribing ribonucleoprotein (RNP) complex (3). The RNP complex, in the presence of four ribonucleoside triphosphates, efficiently transcribes the genome RNA into a 47-base leader RNA (representing the 3′ end of the genome RNA) followed sequentially by the synthesis of five 5′ capped and 3′-polyadenylylated mRNA species coding for, in order, N-NS-M-G-L (4). A unique feature of the transcribing RNP is that the L and the NS proteins are readily dissociable from the complex, rendering the remaining N–RNA complex transcriptionally inactive (1). Effective reconstitution of *in vitro* RNA synthesis occurs only when both L and NS proteins are added to the N–RNA template (1). Thus, the complex containing the L and the NS protein constitutes the active RNA polymerase complex, whereas the N protein by its tight and specific interaction maintains the genome RNA template in a transcriptionally competent form.

Although the phenomenon of *in vitro* transcription reconstitution has been known for some time, the precise roles played by the three constituent polypeptides of the RNP complex in the transcription process *in vitro* have remained obscure. Recently, it has been shown that the purified L protein interacts with the N–RNA complex and initiates transcription but can only synthesize short uncapped oligonucleotides (5). On the other hand, addition of purified NS protein along with the L protein effectively elongates the RNA chains to form matured capped and poly(A)-containing mRNAs. The requirements of the L and the NS proteins for RNA synthesis appeared to be catalytic and stoichiometric, respectively (5). The role of NS protein is particularly interesting because it is a phosphoprotein (6, 7), and various direct and indirect studies have indicated that phosphorylation of NS protein is related to the *in vitro* transcription process (8–13). Moreover, in contrast to the other VSV genes, the *NS* gene appeared to be highly mutable. However, the NS protein of two serotypes contained acidic amino acids sequestered similarly at the NH_2-terminal half of the protein, and the COOH-terminal 21 amino acids are highly conserved (14).

To begin to define the functions of the various domains of the NS protein, we have inserted a full-length copy of the *NS* gene into the SP6 transcription vector (15). Using SP6 RNA polymerase, full-length NS mRNA was synthesized *in vitro*, which was subsequently translated *in vitro* into biologically active NS protein. Systematic deletion mapping of the gene eventually defined a region within the polypeptide that is essential for transcription.

MATERIALS AND METHODS

Construction of Plasmid pGEM-NS. The plasmids pNS50 and pNS22 contained the complete 5′ end and 3′ end of the *NS* gene of VSV(NJ), respectively, and a common (overlapping) unique *Bst*EII restriction site (Fig. 1; ref. 14). A full-length insert DNA for the *NS* gene was constructed, subcloned in M13mp9 (16), and finally cloned in pGEM4 as described in Fig. 1, using standard protocols (17).

***In Vitro* Transcription of pGEM-NS and Translation of NS mRNA.** The pGEM-NS plasmids were linearized prior to transcription reaction by digestion with the appropriate restriction enzymes. *In vitro* transcription was performed using the Riboprobe system (Promega Biotec, Madison, WI) following exactly the manufacturer's protocol except that 2 μg of plasmid DNA, 0.5 mM m^7GpppG^m, 50 μM GTP, and 30 μCi of [^{3}H]UTP (specific activity, 40 Ci/mmol; 1 Ci = 37 GBq) in a total vol of 30 μl were incubated at 40°C for 1 hr. The reaction products were extracted with phenol/chloroform, precipitated with ethanol, and dissolved in 20 μl of H_2O. Approximately 10 μg of RNA was obtained per reaction mixture.

In vitro translation was carried out using rabbit reticulocyte lysate (Amersham), 25–50 μCi of [^{35}S]methionine (specific activity, >800 Ci/mmol), and 0.8 μg of NS mRNA or its truncated versions in a 50-μl reaction mixture. Incubation was at 30°C for 90 min. The translation mixture was layered on 30% sucrose containing 10 mM Tris·HCl (pH 8.0), 0.4 M NaCl, 0.6 mM dithiothreitol, and was centrifuged at 45,000 rpm for 3 hr in an SW60 rotor to remove the ribosomes. The

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: VSV, vesicular stomatitis virus.

*Present address: Department of Virology and Molecular Biology, St. Jude Children's Research Hospital, 332 North Lauderdale, Memphis, TN 38101.

top fraction containing 90% of the radioactivity was pooled and kept at −140°C for subsequent use.

Transcription–Reconstitution Reaction and Analyses of RNA Products. The L protein and the N–RNA complex were purified as described (5, 18). The reaction mixture (0.2 ml) contained 50 mM Tris·HCl (pH 8.0), 5 mM $MgCl_2$, 4 mM CTP, 20 μCi of [α-^{32}P]CTP (specific activity, 400 Ci/mmol), 4 μg of purified N–RNA complex, 1.6 μg of purified L protein, and various amounts of NS proteins. In control experiments, virion-purified NS protein was mixed with rabbit reticulocyte lysate, recovered from the gradient as described above, and used for transcription. The reactions were at 30°C for 2 hr. The RNA products were analyzed by polyacrylamide gel electrophoresis.

Binding Studies of NS Protein with L Protein and N–RNA Complex. *In vitro*-synthesized ^{35}S-labeled NS protein or its truncated versions were incubated in a total vol of 200 μl with saturating amounts of N RNA (4 μg) and with or without L protein (1.6 μg) in the presence of 50 mM Tris·HCl, pH 8.0/100 mM NaCl/5 mM $MgCl_2$/4 mM dithiothreitol, and kept at 31°C for 45 min. The mixture was then diluted 1:5 and centrifuged through 30% (vol/vol) glycerol in the same reaction buffer at 45,000 rpm for 2 hr in an SW60 rotor onto a 100% glycerol cushion. The cores were collected from the cushion, diluted, and precipitated with acetone. The samples were then analyzed on 10% $NaDodSO_4$/polyacrylamide gels according to Laemmli (19).

RESULTS

Cloning of the Full-Length *NS* Gene and Its Expression into Biologically Active Protein. In previous studies (14), we reported the complete 856-nucleotide sequence of the *NS* gene of VSV(NJ) derived from two cDNA clones, pNS 22 and pNS 50, corresponding to the 3′- and 5′-terminal portions, respectively. Using these two clones, a full-length clone of the *NS* gene (pNS2) was constructed and inserted into pGEM4 as shown in Fig. 1. The recombinant plasmid, pGEM-NS, was digested with *Bam*HI and using SP6 polymerase in the presence of m^7GpppG^m, run-off transcription yielded capped full-length NS mRNA. As shown in Fig. 2*A*, the runoff transcription products from two different clones migrated at a similar rate and with authentic NS mRNA synthesized by purified virions. The *in vitro*-synthesized NS mRNA was purified and translated in a rabbit reticulocyte cell-free translation mixture. As shown in Fig. 2*B*, the translation product also comigrated with the *in vitro*-translated NS protein coded by authentic VSV mRNAs. In addition, polyclonal antibody against the whole virion precipitated the SP6 *Bam*HI-derived NS mRNA translation product, indicating that the protein is virus specific (Fig. 2*B*).

To study whether the pGEM-NS-expressed NS protein was biologically active, the *in vitro*-synthesized NS protein was used in a transcription–reconstitution reaction containing purified L protein and N–RNA complex. As shown in Fig. 3, no detectable RNA was synthesized when purified L protein alone was added to the N–RNA complex. However, significant amounts of both the leader RNA and at least four mRNA species (coding for N, NS, G, and M) were detectable after addition of SP-6-derived NS. In fact, the yield of RNA synthesized *in vitro* was ≈70% that of the control reconstitution reaction containing saturating amounts of NS, L, and N–RNA complex purified from the virion and in the presence of ribosome-free reticulocyte lysate (Fig. 3). These results

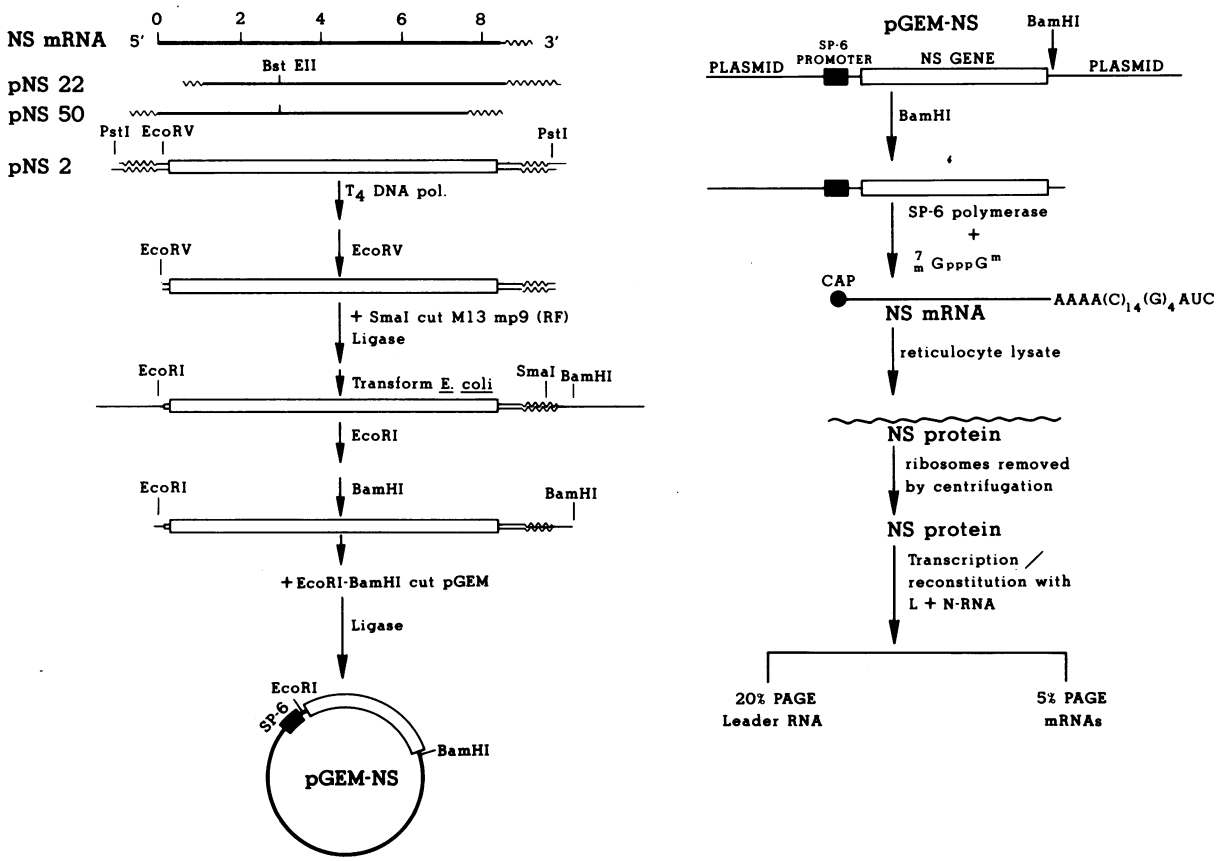

FIG. 1. Cloning and expression of the VSV *NS* gene. Two NS clones, pNS22 and pNS50, were cut with *Bst*EII, and appropriate fractions were ligated and inserted into the *Pst* I site of pBR322. The full-length *NS* gene was then subcloned and inserted into a pGEM4 vector. For expression, the vector was cut with *Bam*HI and, using SP6 polymerase, the gene was transcribed and subsequently translated in rabbit reticulocyte lysate and assayed for biological activity by transcription reconstitution.

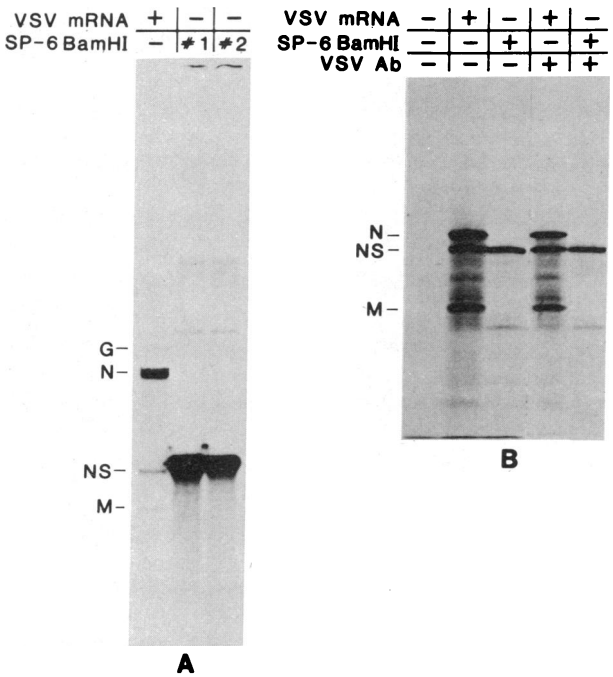

FIG. 2. Transcription and translation of the *NS* gene. (*A*) The *Bam*HI-cut pGEM-NS from two clones (lanes 1 and 2) was transcribed *in vitro* and the products were analyzed by 5% PAGE. The migration of transcription products coding for G, N, NS, and M obtained from *in vitro* transcription of purified virion is shown. (*B*) The SP6-*Bam*HI NS mRNA was translated *in vitro* and the product was analyzed by 10% PAGE. The migration of the translation products of *in vitro* VSV mRNAs is shown. Immunoprecipitation of the NS protein was carried out as described (13). Migration of the immunoprecipitated proteins is shown in the last two lanes.

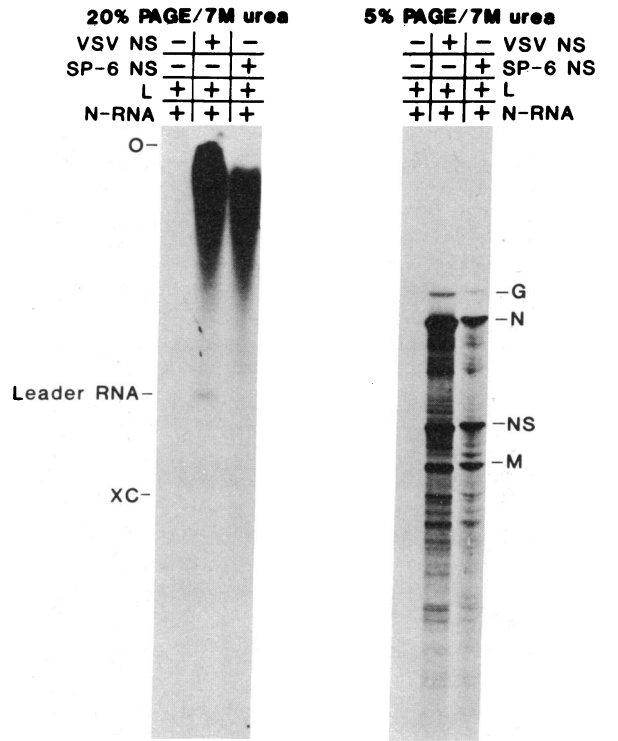

FIG. 3. Transcription reconstitution *in vitro* with SP6 NS. The SP6-expressed NS was added to purified L and N-RNA complex in the presence of [α-^{32}P]CTP and three other ribonucleoside triphosphates (see *Materials and Methods*). (*A*) An aliquot of the RNA product was analyzed by 20% PAGE for detecting leader RNA. (*B*) The other aliquot was subjected to RNase H treatment to remove poly(A) tail (5) and was analyzed by 5% PAGE. Additions (+) and omissions (−) of components are shown. The migration positions of leader RNA; mRNAs coding for G, N, NS, and M; and xylene cyanol (XC) are shown. O, origin.

clearly indicate that pGEM-NS-derived NS protein is biologically active and capable of binding with the L protein and N–RNA template to synthesize virus-specific mRNAs *in vitro*.

Deletion Mapping and Functional Analyses of Various Domains of the NS Protein. The successful reconstitution of transcription activity by the pGEM-NS protein described above provided an opportunity to study the function(s) of various domains within the polypeptide by deletion mapping of the *NS* gene. Using several restriction enzymes, the *NS* gene (thus the mRNA) was truncated (Fig. 4*A*), and the *in vitro* transcription reactions and cell-free translations of truncated RNAs were carried out as described above. As shown in Fig. 4 *B* and *C*, the truncated mRNAs and their corresponding polypeptides migrated according to their predicted sizes. The translation product of *Bst*EII–NS, upon PAGE, released three polypeptides. However, only the smaller polypeptide was precipitable by antibody (data not shown). The precise origin of the larger polypeptide is presently unclear.

Transcription–reconstitution was carried out with each of the truncated polypeptides and the results are shown in Fig. 5. It can be seen that both *Alu* I (253 amino acids) and *Nar* I (247 amino acids) NS protein, each of which lacks the entire homologous domain III, effectively transcribed (≈75%) leader RNA and mRNAs *in vitro* (lanes D, E, d, and e). However, a total loss of RNA synthesis occurred when an additional 34 amino acids (*RSa* I–NS) was removed from the *Nar* I–NS fragment (lanes F and f). As expected, the *Bst*NI–NS (193 amino acids) (lanes G and g) and *Bst*EII–NS (97 amino acids) (lanes H and h) did not support RNA synthesis *in vitro*. These results clearly indicate that the COOH-terminal highly homologous domain in the NS protein is not required in the *in vitro* transcription process. By contrast, a polypeptide stretch (34 amino acids) spanning amino acid residues 213–247 (Fig. 4) is essential to obtain transcriptionally active NS protein.

Binding of the Truncated Proteins with the N–RNA Complex. Next, we examined whether the apparent lack of transcription capability of the *Rsa* I–NS and the smaller polypeptides was due to their inability to bind to the N–RNA complex. The *in vitro*-translated NS protein and the variously truncated polypeptides were incubated individually with the N–RNA template. Complexes formed were recovered from the cushion by centrifugation and the constituent polypeptides were analyzed by 10% PAGE. As shown in Fig. 6*A*, the control *Bam*HI–NS bound efficiently to the N–RNA complex. Based on the recovery of labeled NS protein, a 50% binding with the N–RNA complex was observed. In sharp contrast, the *Alu* I and *Nar* I–NS failed to bind stably with the N–RNA template, in spite of the fact that these polypeptides supported transcription efficiently. As expected, *Rsa* I–NS, *Bst*NI–NS, and *Bst*EII–NS did not bind to the N–RNA template. These results indicate that the binding of *Alu* I–NS and *Nar* I–NS with the N–RNA template is not strong enough to be recovered by the procedure described. However, they are bound adequately to carry out transcription in the presence of the L protein (Fig. 5).

To investigate whether the putative loose binding of *Alu* I–NS and *Nar* I–NS with the N–RNA template was due to lack of L protein, similar binding experiments were carried out in the presence of excess L protein and an increasing concentration of NS protein. It can be seen in Fig. 6*B* that the addition of L protein substantially increased the binding of NS protein to the complex. In the presence of saturating

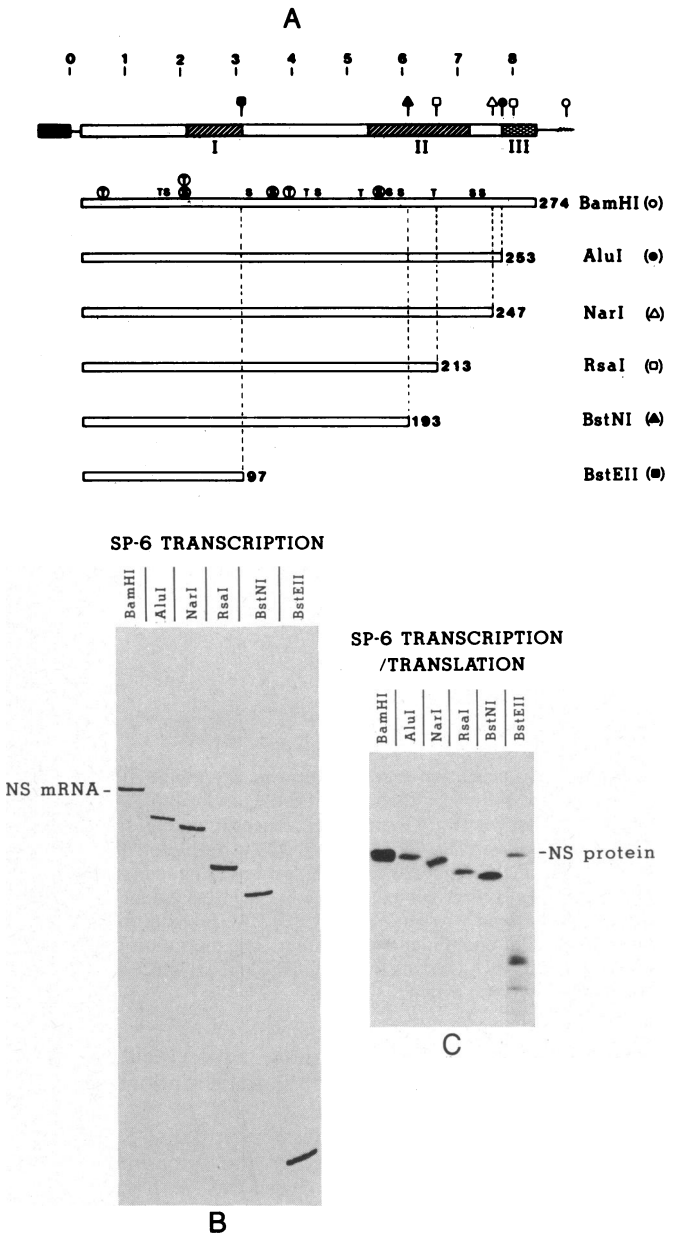

FIG. 4. Deletion mapping of the *NS* gene and analyses of truncated NS RNAs and their corresponding proteins. (*A*) The full-length *NS* gene, in hundreds of nucleotides, and the site of various restriction enzymes are shown. Hatched regions I and II represent the nonhomologous regions, and III represents the highly homologous region in the *NS* gene (14). Solid rectangle represents the SP6 promoter region. T and S, positions of invariant threonine and serine, respectively, within the *NS* genes of both VSV(NJ) and VSV(IND). Circled T or S symbolizes that the residue at that position in VSV(IND) is serine or threonine, respectively. The deletion of the *NS* gene was carried out by various restriction enzymes and the sizes of the respective truncated genes (in number of amino acid residues) are shown. The various truncated SP6 *NS* genes were transcribed and subsequently translated *in vitro*. The RNA (*B*) and the protein products (*C*) were analyzed by 5% and 10% PAGE, respectively, and then autoradiographed.

amounts of L protein, the binding of control *Bam*HI–NS increased to almost 90% with increasing concentrations of NS product (based on the recovery of radioactivity). Similarly, the binding of *Alu* I–NS or *Nar* I–NS (not shown) increased to ≈50% of the control in the presence of L protein. In contrast, the presence of L protein had no effect on the binding of *Rsa* I–NS to the N–RNA complex. Thus, the

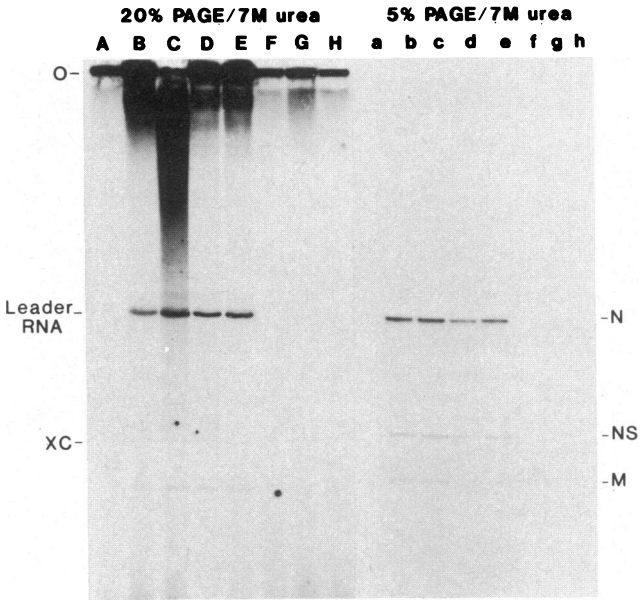

FIG. 5. Transcription reconstitution *in vitro* with truncated NS proteins. The variously truncated proteins obtained from Fig. 4*B* were used to reconstitute transcription in the presence of purified L protein and N–RNA complex. The RNA products were analyzed by 20% PAGE for leader RNA and by 5% PAGE for mRNA and then autoradiographed. XC, xylene cyanol. Each lane contained purified L and N–RNA complex in addition to no NS (lanes A and a), virion NS (lanes B and b), *Bam*HI–NS (lanes C and c), *Alu* I–NS (lanes D and d), *Nar* I–NS (lanes E and e), *Rsa* I–NS (lanes F and f), *Bst*NI–NS (lanes G and g), *Bst*EII–NS (lanes H and h).

domain III may have a direct role in binding to the N protein, which is not involved in transcription. On the other hand, removal of the critical 34 amino acid residues from the *Nar* I–NS polypeptides (*Rsa* I–NS) completely abolishes its ability to bind stably to both the L and the N–RNA template and results in total abrogation of RNA synthesis *in vitro*.

DISCUSSION

In the present work, we have used molecular cloning and expression of the phosphoprotein (*NS*) gene of VSV to study its structure and function. A full-length cDNA clone of the *NS* gene of VSV(NJ) was inserted into the SP6 transcription vector pGEM4 downstream of the SP6 promoter. This system enabled us to synthesize adequate quantities of NS mRNA *in vitro*, which can be translated into biologically active NS protein in a cell-free rabbit reticulocyte system (Figs. 2 and 3).

Using various restriction enzymes, it was possible to generate a series of truncated *NS* genes that, when transcribed *in vitro*, synthesized RNAs of corresponding sizes (Fig. 4). Transcription–reconstitution with the *in vitro*-translated polypeptides with the L and the N–RNA complex clearly defined a region within the NS protein (spanning amino acid residues 213–247), which is essential for transcription. Removal of this domain (*Rsa* I–NS) eliminates its ability to bind to the N–RNA complex even in the presence of L protein, resulting in the cessation of RNA synthesis *in vitro* (Figs. 5 and 6) and indicating that the 34-amino acid domain is critical for stable binding to the L protein to form an active RNA polymerase complex.

Surprisingly, the highly homologous region between the NS protein of VSV(NJ) and VSV(IND) (14), domain III, seemed not to be involved in transcription (Fig. 5). However, this domain mediates strong binding with the N–RNA complex, since removal of this domain greatly reduces the ability

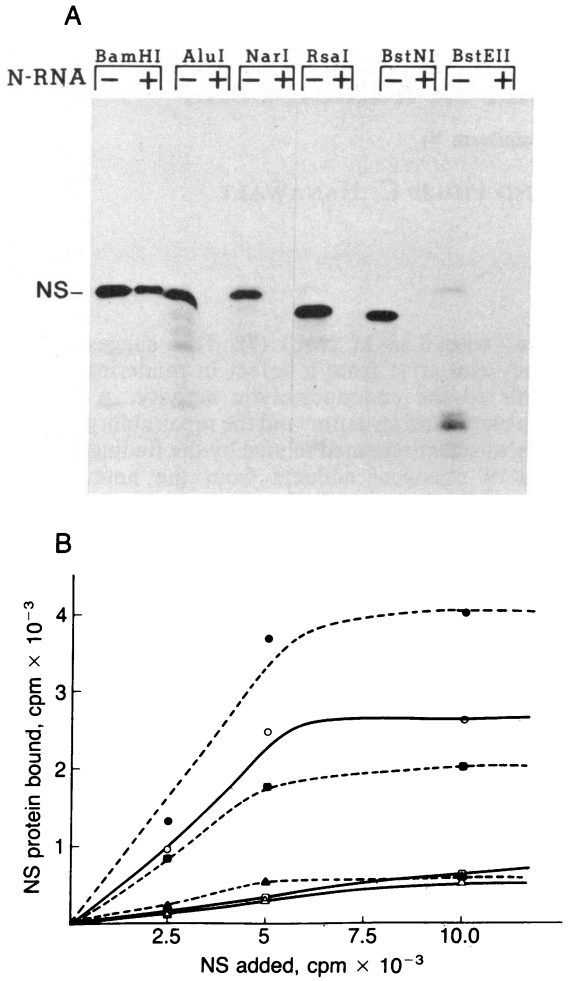

FIG. 6. Binding of truncated NS proteins with purified N–RNA complex. (*A*) An aliquot of each of the ^{35}S-labeled truncated NS proteins was mixed with N–RNA complex and centrifuged. The pellet was analyzed by 10% PAGE and fluorographed. As a control, an identical aliquot of the proteins was analyzed directly by 10% PAGE without binding to the N–RNA complex. (*B*) Different amounts of truncated NS proteins (in cpm) were bound with N–RNA complex in either the presence or the absence of L protein, were recovered by pelleting, and acid-insoluble radioactivity was determined. (*Bam*HI–NS + N–RNA) plus L protein (●), minus L protein (○); (*Alu* I–NS + N–RNA) plus L protein (■), minus L protein (□); (*Rsa* I–NS + N–RNA) plus L protein (▲), minus L protein (△).

of the NS protein to bind to the complex (Fig. 6*A*). Nevertheless, the NS protein lacking this domain binds efficiently to the L protein and functions appropriately as the RNA polymerase complex (Fig. 5). Thus, the site(s) on the NS protein that interacts with the L protein and with the N protein–RNA template during RNA synthesis does not seem to reside in domain III. However, increased and efficient binding of the NS protein containing domain III with the N–RNA template indicates that the domain may be directly linked to strong association with the N protein. Thus, it appears that the phosphoprotein contains at least three functional domains: (*i*) the highly charged NH_2-terminal half of the polypeptide (14), (*ii*) the 34-amino acid region essential for transcription, and (*iii*) the COOH-terminal 28-amino acid region (*Nar* I–NS) not involved directly in transcription but that mediates strong binding with the N–RNA template. It should be emphasized, however, that since the activity of a polypeptide is determined by its properly folded structure (20), removal of a crucial segment may have an adverse effect on the three-dimensional structure of the polypeptide, resulting in loss of activity. Thus, the 34-amino acid segment may be vital for maintaining the integrity of the protein without being the direct site of L protein binding.

Finally, a closer examination of the L protein binding region (Fig. 4*A*) reveals that it contains two invariant serine residues. It is tempting to speculate that one or both of the serine residues are directly involved in secondary phosphorylation (9) and, consequently, in the transcription process *in vitro*. Using oligonucleotide-directed mutagenesis and recombinant DNA technology, it will now be possible to determine precisely the location of the phosphorylation site(s) involved in transcription, as well as the function(s) of the highly charged NH_2-terminal half and the constitutive phosphorylation (21–23) sites within the phosphoprotein molecule.

D.S.G. and D.C. contributed equally to this work.

1. Emerson, S. U. & Yu, Y.-H. (1975) *J. Virol.* **15,** 1348–1356.
2. Banerjee, A. K., Masters, P. S. & Gill, D. S. (1986) in *The Biology of Negative Strand Viruses*, eds. Mahy, B. & Kolakofsky, D. (Elsevier Biomedical, Amsterdam), in press.
3. Wagner, R. R. (1975) in *Comprehensive Virology*, eds. Fraenkel-Conrat, H. & Wagner, R. R. (Plenum, New York), pp. 1–93.
4. Banerjee, A. K., Abraham, G. & Colonno, R. J. (1977) *J. Gen. Virol.* **34,** 1–8.
5. De, B. P. & Banerjee, A. K. (1985) *Biochem. Biophys. Res. Commun.* **126,** 40–49.
6. Clinton, G. M., Burge, B. W. & Huang, A. S. (1979) *Virology* **99,** 84–94.
7. Hsu, C.-H. & Kingsbury, D. W. (1982) *J. Virol.* **42,** 342–345.
8. Kingsford, L. & Emerson, S. U. (1980) *J. Virol.* **33,** 1097–1105.
9. Hsu, C.-H., Morga, E. M. & Kingsbury, D. W. (1982) *J. Virol.* **43,** 104–112.
10. Watanabe, Y., Sakuma, S. & Tanaka, S. (1974) *FEBS Lett.* **41,** 331–341.
11. Witt, D. J. & Summers, D. F. (1980) *Virology* **107,** 34–49.
12. Talib, S. & Banerjee, A. K. (1981) *Biochem. Biophys. Res. Commun.* **98,** 875–883.
13. Sánchez, A., De, B. P. & Banerjee, A. K. (1985) *J. Gen. Virol.* **66,** 1025–1036.
14. Gill, D. S. & Banerjee, A. K. (1985) *J. Virol.* **55,** 60–66.
15. Melton, D. A., Kreig, P. A., Rebagliati, M. R., Maniatis, T., Zin, K. & Green, M. R. (1984) *Nucleic Acids Res.* **12,** 7035–7056.
16. Messing, J. (1983) *Methods Enzymol.* **101,** 20–78.
17. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 329–333.
18. De, B. P. & Banerjee, A. K. (1984) *J. Virol.* **51,** 628–634.
19. Laemmli, U. K. (1970) *Nature (London)* **227,** 680–685.
20. Rossmann, M. G. & Argos, P. (1981) *Annu. Rev. Biochem.* **50,** 497–532.
21. Hsu, C.-H. & Kingsbury, D. W. (1985) *J. Biol. Chem.* **260,** 8990–8995.
22. Bell, J. C. & Prevec, L. (1985) *J. Virol.* **54,** 697–702.
23. Marnell, L. L. & Summers, D. F. (1984) *J. Biol. Chem.* **259,** 13518–13524.