Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 7977–7981, November 1988
Biochemistry

NH_2-terminal acidic region of the phosphoprotein of vesicular stomatitis virus can be functionally replaced by tubulin

(cloning in pGEM vector/transcription–translation/chimeric protein/transcription activators)

DHRUBAJYOTI CHATTOPADHYAY AND AMIYA K. BANERJEE

Department of Molecular Biology, Research Institute, The Cleveland Clinic Foundation, Cleveland, OH 44106

Communicated by Aaron J. Shatkin, August 1, 1988 (received for review June 28, 1988)

ABSTRACT **The phosphoprotein (NS) of vesicular stomatitis virus is an indispensable subunit of the virion-associated RNA polymerase (L). NS consists of a highly acidic NH_2-terminal domain and a basic COOH-terminal domain. Unlike the latter, the amino acid sequences of the NH_2-terminal regions are highly dissimilar among different viral serotypes, although they share structural similarities. We have cloned an NS gene into the SP6 transcription vector and replaced the 5′-terminal 80% by a full-length gene for β-tubulin, which contains an acidic COOH-terminal domain. Here we present evidence that the chimeric tubulin–NS protein is biologically active and that the acidic region in tubulin directly affects the transcription reaction. These observations indicate that NS probably functions as an activator protein in which the acidic domain stimulates transcription of the viral genes by interacting with the RNA polymerase as observed for eukaryotic cellular transcription activators.**

An RNA-dependent RNA polymerase packaged within the purified virion of vesicular stomatitis virus (VSV) transcribes the negative-strand genome RNA *in vitro* into five distinct mRNAs and a leader RNA (1, 2). This transcription is essential for initiation of infection when VSV infects cells. The RNA polymerase consists of two distinct polypeptide subunits: a large polypeptide, L, of 240 kDa and a small phosphoprotein, NS, of 29 kDa tightly associated with each other and also with the N protein–RNA template (N–RNA template). There are approximately 50 molecules of the L protein and 500 molecules of the NS protein per virion (3). The unique feature of the VSV transcription complex is that each of the reacting components can be separated free from the others and *in vitro* RNA synthesis can be restored effectively when L and NS proteins are added to the N–RNA template (4).

By using the above transcription reconstitution system, it was shown (5) that the L protein, which acts catalytically in the initiation of RNA chains, fails to synthesize full-length RNA unless the NS protein, which acts stoichiometrically, is added to the transcription mixture. These results indicated that the L protein is the RNA polymerase whereas the NS protein is a regulatory protein that plays a role in the transcription process, possibly by interacting with the N–RNA template to facilitate access of L protein to the genome template for RNA synthesis (5, 6). Phosphorylation of the NS protein seems to have important regulatory effect(s) in this process (1).

To study the precise function of the NS protein in the transcription process, we (7) inserted a full-length NS gene (New Jersey serotype) into an SP6 transcription vector. This system enabled us to obtain NS mRNA in amounts that, when subsequently translated *in vitro*, resulted in the synthesis of biologically active protein. The latter activity was demonstrated by transcription reconstitution with purified L protein and the N–RNA template. Systematic deletion mapping studies (7) showed (*i*) that the COOH-terminal basic 27 amino acids (Fig. 1, domain III) are not involved in transcription but are essential for binding to the N–RNA template in the absence of the L protein and (*ii*) that the adjacent 34 amino acids (domain II) interact with the L protein and are directly involved in the transcription process. The phosphorylation of the invariant serine residues (Ser-236 and Ser-242) in domain II plays an important role in its capacity to bind to the L protein leading to RNA synthesis *in vitro* (8).

The negatively charged NH_2-terminal half of the NS protein (domain I) presumably is involved in some step(s) in the RNA synthetic process. This domain is remarkably dissimilar in amino acid sequence yet structurally similar among VSV serotypes (9–11). Recently, we subcloned the region containing domains II and III in the SP6 transcription vector and demonstrated that the encoded polypeptide (termed EB NS protein) binds to the L–N–RNA complex but fails to support transcription. However, it does support transcription when the polypeptide encoded by domain I is added in trans (12). The intriguing observation was that tubulin, which has acidic amino acid residues clustered at the COOH-terminal region (13), was able to replace NS domain I for RNA synthesis. We chose tubulin because it was previously shown (14) to function as a positive transcription factor for VSV as well as for Sendai, another negative-strand RNA virus. In addition, tubulin associates with the L protein in VSV-infected cells and, more important, anti-tubulin antibody virtually abolishes *in vitro* RNA synthesis by purified virions (14). These results indicated that tubulin, perhaps specifically due to its acidic character, has a role in the VSV transcription process. In this communication we demonstrate that a chimeric protein containing β-tubulin fused to domains II and III of the NS protein functions efficiently for *in vitro* transcription. Further, the acidic domain of tubulin in the chimeric polypeptide plays a direct role in the transcription process.

MATERIALS AND METHODS

Construction of pGEM Tubulin Clone (pGEM-Tub). A full-length cDNA clone of β-tubulin in pBR322 plasmid at the unique *Eco*RI site (13) was digested with *Eco*RI (Boehringer Mannheim). The resulting insert was purified by 0.8% agarose gel electrophoresis and the DNA was eluted from the gel. The eluted insert DNA was purified by phenol/chloroform extraction followed by ethanol precipitation and was dissolved in 1 mM Tris·HCl, pH 7.5/0.1 mM EDTA. Two hundred nanograms of the insert was ligated with 200 ng of *Eco*RI-cut, phosphatase-treated plasmid pGEM-4 by using phage T4 DNA ligase (0.1 unit) in ligase reaction buffer [50

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.

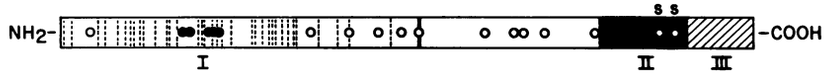

FIG. 1. Diagram of the functional domains of the NS protein of VSV serotype New Jersey. The protein (274 amino acids) encoded by the full-length NS gene (New Jersey serotype) was divided into three domains (7). The dashed lines within domain I represent acidic amino acid residues. Solid circles represent constitutively phosphorylated residues, and open circles represent the potential phosphorylation sites conserved between Indiana and New Jersey serotypes (9, 10). Two invariant serine residues (S) within domain II are shown (8).

mM Tris·HCl, pH 7.6/10 mM $MgCl_2$/5% (wt/vol) polyethylene glycol 8000/1 mM ATP/1 mM dithiothreitol] in a total volume of 20 μl. The reaction mixture was incubated at 8°C for 16 hr. The ligation mixture was then diluted 1:5 and used to transform competent *Escherichia coli* RR1 cells. Positive transformants were confirmed by restriction enzyme digestion and dideoxy sequencing of the junction region according to the Promega Biotec protocol (Promega Biotec, Madison, WI).

Construction of pGEM Chimeric Plasmids. pGEM-Tub (10 μg) was digested completely with *Stu* I and *Sal* I to remove the 300-base-pair segment representing the noncoding part of the tubulin gene and a portion of the pGEM polylinker region. pGEM-EBNS has two *Ava* II sites, at positions 11 and 51. pGEM-EBNS was partially digested with *Ava* II, blunt-ended with Klenow polymerase, and then digested with *Sal* I. Theoretically, two different fragments are possible (see Fig. 4). The large *Stu* I–*Sal* I fragment of pGEM-Tub (300 ng) was then ligated with 10 ng of *Sal* I-cut, blunt-ended smaller fragments of pGEM-EBNS in a 20-μl reaction mixture. The ligation reaction was carried out at 23°C for 4 hr and then stopped by the addition of 1 μl of 0.5 M EDTA. The ligation mixture was diluted 1:5 and then used to transform competent RR1 cells. The positive transformants of pGEM-Teb4 and pGEM-Teb1 were then scored by sequencing the tubulin–NS junction region according to the Promega Biotec dideoxy sequencing protocol. Both pGEM-Teb4 and pGEM-Teb1 contain nucleotides 1–1405 of the tubulin gene. pGEM-Teb4 contains nucleotides 564–894 of the NS gene, which results in a +1 shift in reading frame. pGEM-Teb1 contains nucleotides 604–894 of the NS gene, in frame with the tubulin gene.

***In Vitro* Transcription of Recombinant Plasmids, Translation of mRNAs, and Reconstitution of VSV RNA Synthesis.** The recombinant plasmid pGEM-Tub (10 μg) was linearized by digestion with *Bam*HI or *Sal* I restriction enzyme. For the chimeric plasmids pGEM-Teb1 and pGEM-Teb4 (10 μg), the DNAs were linearized prior to SP6 transcription by digestion with *Sal* I. For control NS protein (*Bam*HI NS), pGEM-NS was digested with *Bam*HI before the *in vitro* transcription reaction. *In vitro* transcription reactions were performed with the Riboprobe system (Promega Biotec) by following exactly the manufacturer's protocol (7). Approximately 2 μg of RNA was obtained per μg of linearized DNA. *In vitro* translation of mRNAs was carried out in rabbit reticulocyte lysate (7). Transcription-reconstitution reactions using *in vitro* synthesized proteins in the presence of purified L protein, N–RNA template, and EB NS protein were carried out as detailed (12).

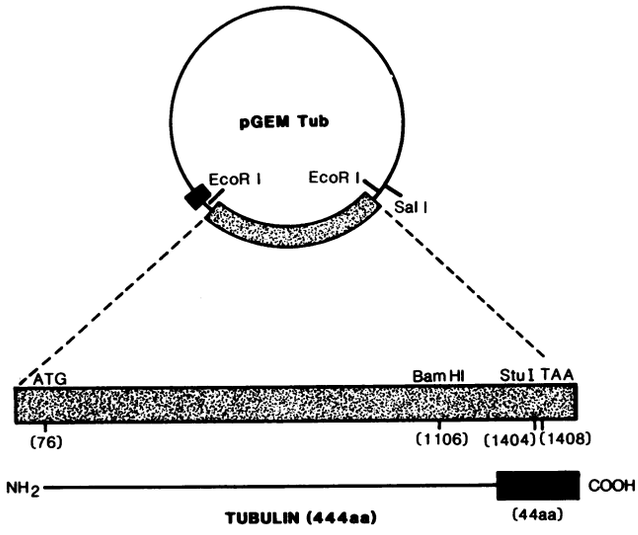

FIG. 2. Schematic representation of the constructed pGEM-Tub plasmid, including the important restriction sites. The construction of the pGEM tubulin clone is detailed in *Materials and Methods*. The plasmid, pGEM-Tub, has a single *Bam*HI site, in the 3′-terminal region of the tubulin coding sequence; cleavage at this site essentially removes the COOH-terminal region containing the negatively charged 44 amino acids (black box). Nucleotide positions of the initiation and termination codons and the *Bam*HI and *Stu* I restriction sites are given below the stippled bar representing the tubulin cDNA.

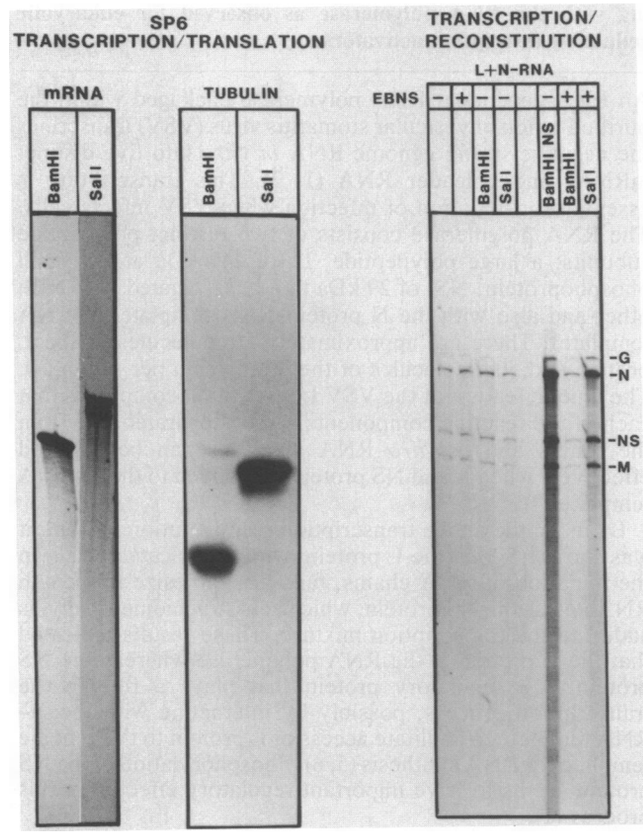

FIG. 3. Transcription (*Left*) and translation (*Center*) of full-length and truncated tubulin mRNAs and *in vitro* transcription reconstitution (*Right*). The tubulin (*Sal* I) and its truncated version (*Bam*HI) were used to reconstitute transcription in the presence of purified L protein, N–RNA template, and EB NS protein in a total volume of 200 μl, with [α-^{32}P]CTP as the labeled precursor as detailed previously (8). The RNA products synthesized in 2 hr were analyzed by 5% PAGE, with 7 M urea for mRNA, and autoradiography. Migration positions of mRNAs coding for VSV G, N, NS, and M proteins are shown.

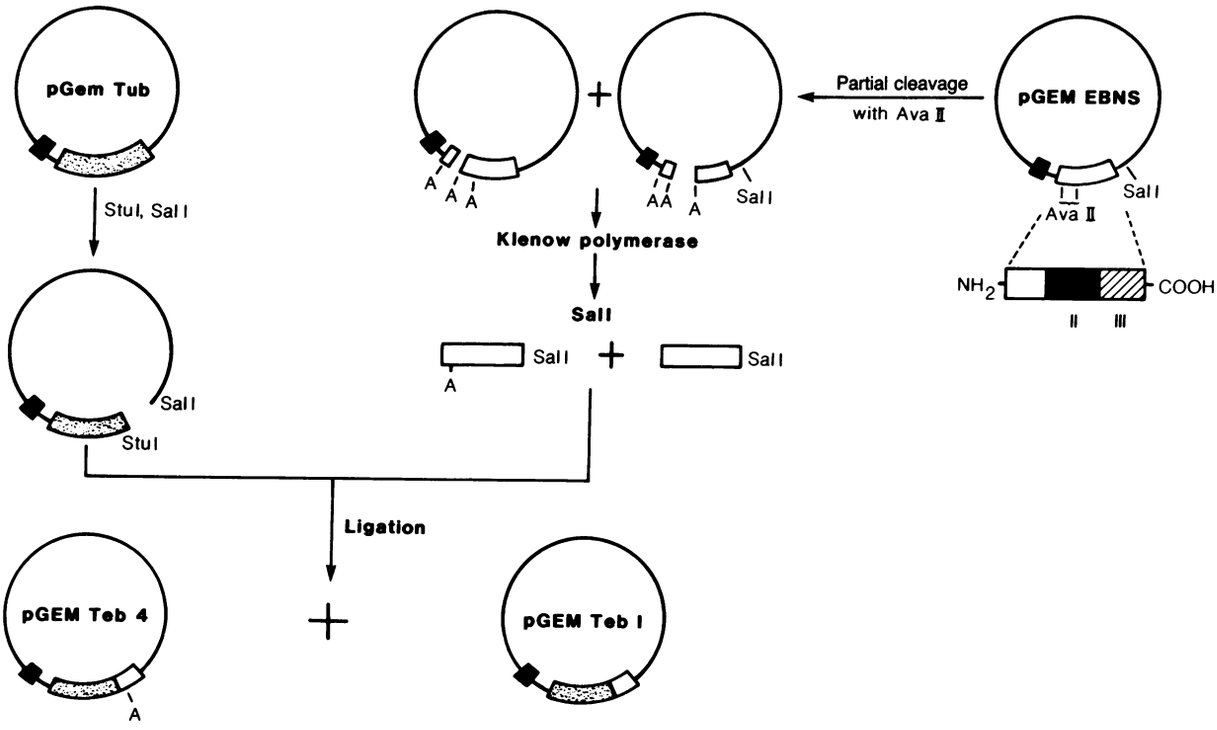

FIG. 4. Construction of pGEM plasmids containing part of the tubulin gene and the region of the NS gene encoding protein spanning domains II and III. Details of the strategy for the construction of the tubulin–NS chimeric gene are described in *Materials and Methods*.

RESULTS

Trans-activation of Transcription by *in Vitro* Translated β-Tubulin. To ascertain the exact role of the polypeptide encoded by domain I, a recombinant plasmid was constructed containing the β-tubulin gene linked at its 3′ end with the NS gene sequence coding for domains II and III. However, before we embarked on the construction of a chimeric gene, it was important to demonstrate that clonally expressed tubulin functions in trans as previously demonstrated for β-tubulin purified from bovine brain (12, 14). The full-length cDNA clone of mouse β-tubulin isotype Mβ5, derived from a cDNA library prepared from mRNA isolated from mouse testis (13), was subcloned into pGEM-4 vector at the *Eco*RI site behind the SP6 promoter (Fig. 2). The plasmid pGEM-Tub, containing the full-length β-tubulin gene, would yield tubulin polypeptide (444 amino acids) upon transcription followed by translation of the mRNA *in vitro*. Digestion with *Bam*HI, however, would result in truncation of the gene and loss of the region coding for the negatively charged COOH-terminal domain of tubulin. The plasmid pGEM-Tub was linearized by digestion with *Bam*HI or *Sal* I and transcribed. The resulting mRNAs, analyzed by PAGE, migrated at the expected rates (Fig. 3, *Left*). The runoff transcription products were then translated *in vitro* with rabbit reticulocyte lysate to synthesize sufficient amounts of full-length and truncated β-tubulin for further studies (Fig. 3, *Center*). These proteins were then used in place of domain I in a transcription-reconstitution reaction containing purified L protein and N–RNA template (Fig. 3, *Right*). The full-length tubulin (from the *Sal* I-treated template) stimulated RNA synthesis by 10-fold over background (second lane from left) and was 30% as active in supporting transcription as the control reaction containing *in vitro* expressed NS protein (*Bam*HI NS). The extent of RNA synthesis with the control expressed protein was ≈70% of that when NS purified from virions was used under identical condition (7). In striking contrast, when the negatively charged carboxyl-terminal 44 amino acids were removed (*Bam*HI tubulin), the polypeptide was virtually inactive in the transcription reaction. These results demonstrate that pGEM-Tub-derived polypeptide is functionally active and that the acidic part of domain I plays a major role in stimulating transcription of the genome RNA.

Construction of Tubulin–NS Chimeric Gene. We next constructed a chimeric clone containing the β-tubulin gene linked to the sequence encoding domains II and III of NS (pGEM-EBNS) (Fig. 4). The plasmid, pGEM-Teb1, contains the sequences of domains II and III of the NS gene linked to the tubulin gene such that the expressed chimeric protein (Teb1) will be initiated at the 5′ AUG codon of the tubulin mRNA and terminate in frame. On the other hand, in the pGEM-Teb4 construct the NS polypeptide region (domains II and III) is out of frame in the chimeric protein (Teb4). The expressed chimeric proteins were then tested for biological activity by transcription-reconstitution reactions (Fig. 5). Teb1 chimeric polypeptide stimulated RNA synthesis by 20-fold over the reaction mixture containing L protein plus N–RNA only. It supported transcription at 65% with respect to the control reaction containing the full-length *in vitro* expressed NS protein (*Bam*HI NS). In contrast, Teb4 chimeric polypeptide failed to support RNA synthesis *in vitro*. These results support the contention that domain I of the NS protein can be replaced by an unrelated but similarly charged polypeptide without abolishing its activity. Poly(glutamic acid), but not polylysine, can also support transcription, albeit less effectively (20% of control), when it replaces domain I of the NS protein (Fig. 5). By contrast, domains II and III of the NS protein govern the specificity of its interaction with the template and L protein. From these results it appears that the acidic property of domain I plays a crucial role in allowing the RNA polymerase complex to transcribe the N–RNA template.

Finally, we wished to determine the efficiency of transcription activation by tubulin when added in trans compared to the tubulin–NS chimeric protein. Stimulation of *in vitro* transcription increased with increasing amounts of either type of protein (Fig. 6); however, the chimeric protein Teb1 stimulated RNA synthesis to 65% of control, whereas RNA

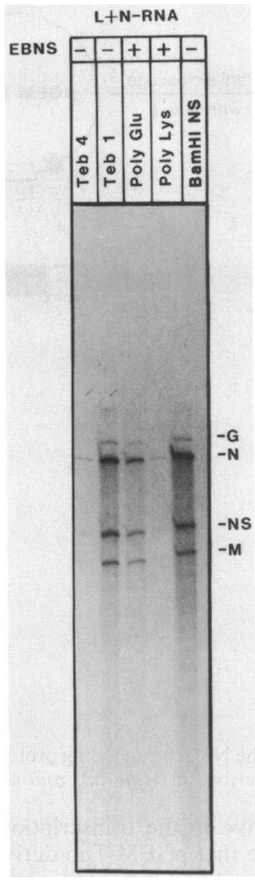

FIG. 5. Transcription reconstitution *in vitro* with chimeric proteins. The tubulin–NS chimeric proteins encoded in pGEM-Teb4 and pGEM-Teb1 were synthesized *in vitro* and used to reconstitute transcription in the presence of purified L protein and N–RNA template in a total volume of 200 μl (7). Poly(glutamic acid) (10 μg) or polylysine (10 μg) was also used to reconstitute transcription in presence of EB NS, L protein, and N–RNA template. The [α-^{32}P]CMP-labeled RNA synthesized in 2 hr was analyzed as described (12). Migration positions of mRNAs coding for VSV G, N, NS, and M proteins are shown.

synthesis reached only 30% of the control value for tubulin added in trans. These results indicate that the acidic domain functions more efficiently when linked to the NS-protein domains II and III, which correspond to the binding sites for the template and polymerase, respectively.

DISCUSSION

The data presented here provide a better understanding of the structure and possible function(s) of the highly acidic NH_2-terminal region of the NS phosphoprotein of VSV. By itself, this segment of the polypeptide is inactive in the transcription process; however, when added in trans with the polymerase- and template-binding regions (domains II and III), the biological activity of the phosphoprotein is effectively restored (12). Here, we demonstrate that the acidic domain can be intramolecularly substituted by tubulin, a similarly charged protein (Figs. 4 and 5). Thus, a tubulin–NS chimeric protein that contains only 18% of the NS sequence retains most of its biological activity (Fig. 5). These results indicate that the nucleotide sequence of the NS gene encoding domain I can undergo substantial alteration and mutation during evolution so long as it retains a negative charge character in the polypeptide. This probably explains the high mutability of the NS domain I within vesiculoviruses (11). By contrast, the carboxyl-terminal sequence encoding domains II and III has withstood evolutionary pressure, presumably because it imparts specificity for interaction of NS protein with the template and L protein. This is reflected in the high degree of amino acid sequence identity within the COOH-terminal portion among the NS genes of the different serotypes (9–11). It is noteworthy that tubulin–NS chimeric protein is more effective in supporting transcription (65% of control) than tubulin added in trans with NS domains II and III (30%; Fig. 6). These results indicate that the NS gene has evolved to link domain I with the COOH-terminal conserved domain for efficient transcription of the genome RNA. The entire structure of the polypeptide must interact with the reacting components to enhance the transcription process.

An important observation is that poly(glutamic acid) can also activate transcription, although poorly, when added in trans (Fig. 5). Similarly, the COOH-terminal 44 amino acids of tubulin, containing clusters of acidic residues, appear to be directly involved in the transcription process, because their removal abrogates RNA synthesis *in vitro* (Fig. 3). Removal of that domain from tubulin did not alter the net negative charge of the protein (pI change, 5.14 to 5.15). This indicates that a stretch or cluster of acidic amino acids of variable length in domain I provides the necessary and sufficient condition to impart a trans-activating property to the NS protein. The stimulation of *in vitro* RNA synthesis by tubulin observed by Moyer *et al.* (14) may have been due to interaction of the acidic domain of tubulin with the N–RNA template. Microtubule-associated proteins have also been shown to act as transcription factors in VSV RNA synthesis

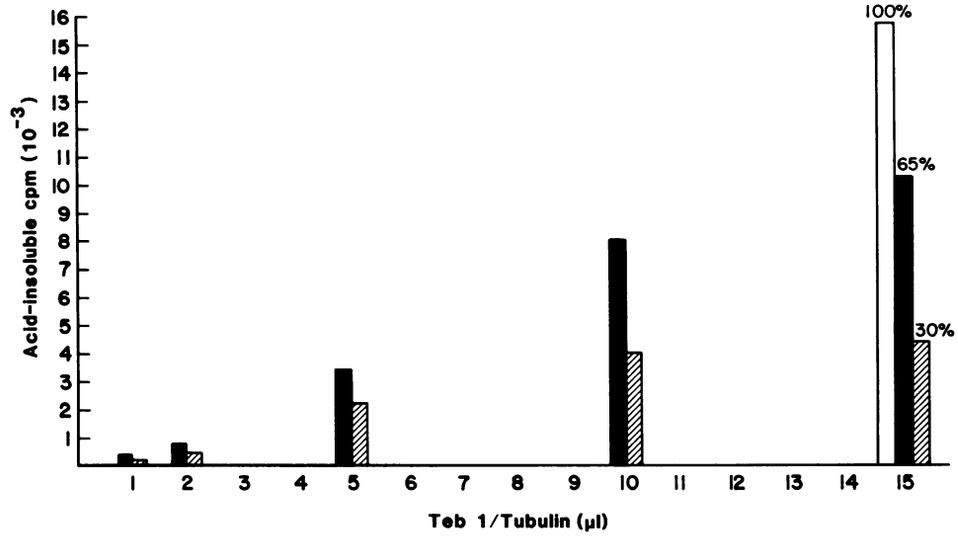

FIG. 6. Transcription reconstitution *in vitro* with tubulin or tubulin-NS protein. Transcription reconstitution in standard reaction mixtures was carried out in the presence of purified L protein, N–RNA template, 60 pg of EB NS protein, and various concentrations of β-tubulin (100 pg/μl of reticulocyte lysate). Acid-insoluble radioactivity was determined after 2 hr of RNA synthesis at 30°C (hatched bars). Similar reactions were carried out with various concentrations of Teb1 protein (120 pg/μl of reticulocyte lysate) (solid bars). RNA synthesis for 2 hr by virion-purified L protein, *in vitro* expressed NS protein (*Bam*HI NS), and N–RNA template (open bar) was taken as 100%. Reaction conditions were as described (8).

in vitro (15). Whether this multiprotein complex, like tubulin, functions as negatively charged proteins in the activation of VSV RNA polymerase or has another role in the life cycle of VSV remains to be seen. Involvement of the cytoskeleton complex in the biosynthesis of VSV proteins has also been documented (16–18).

The above structural and functional features of the NS protein bear a striking resemblance to those of the yeast transcriptional regulator proteins GAL4 and GCN4 (19, 20). These proteins contain two distinct functional domains: one binds DNA and the other is an acidic region that functions as the transcription activator. The former domain binds at sequence-specific sites on the DNA template (upstream activator sequence) but fails to activate transcription, whereas the latter, acidic fragment activates transcription even when substituted by unrelated but acidic *E. coli* proteins or synthetic amphipathic peptide (21, 22). Recently, the DNA-binding portion of *E. coli* LexA repressor protein was fused to v-*fos* and c-*fos* oncogene products (which are acidic), and the resulting chimeric proteins activated yeast gene expression only when the LexA operator was inserted next to the promoter sequence (23). Interestingly, yeast trans-activators have also been shown to activate mouse mammary tumor virus promoter in mammalian cells when GAL4 binding sites are present before the promoter (24). In addition, chimeric protein consisting of the human estrogen receptor DNA-binding domain and either GAL4 or GCN4 activating acidic regions can activate a promoter region controlled by an estrogen-responsive enhancer (25). All of the above-mentioned acidic trans-activators characteristically contain stretches of α-helical structures (22).

The VSV phosphoprotein NS, like yeast trans-activator proteins, contains two distinct regions: a specific template- and polymerase-binding region (domains II and III) and an acidic region (domain I) that constitutes the major portion of the polypeptide. Based on the structural and functional similarities between the NS protein and the eukaryotic trans-activators, it is tempting to speculate that the NS protein also functions as a transcriptional activator in the RNA-dependent RNA transcription process in the VSV system. NS plays a crucial role in orchestrating the formation of an active transcription initiation complex with the L protein and the N–RNA template. The basic COOH-terminal region (domains II and III) binds to the L protein and links at a specific site(s) on the N–RNA template. However, this complex remains transcriptionally inactive until activated by the acidic domain I. The precise nature of this complex interaction remains to be understood. The acidic domain I, like the eukaryotic transcriptional activators, contains stretches of α-helical structures (18, 26). These structures probably interact with the helical structures of the N–RNA template and transiently displace the RNA-bound N protein, allowing the L protein to contact the genome RNA. The putative RNA binding sites of the NS protein on the genome template remain to be precisely determined, although a high-affinity binding site at the 3′ end of the genome RNA has been reported (27). Thus, the molecular mechanism underlying transcriptional activation may be conserved in a large variety of eukaryotic species including RNA viruses.

We are grateful to Dr. N. Cowan (New York University Medical Center, New York) for the β-tubulin clone.

1. Banerjee, A. K. (1987) *Microbiol. Rev.* **51,** 66–87.
2. Banerjee, A. K. (1987) *Cell* **48,** 363–364.
3. Thomas, D., Newcomb, W. W., Brown, J. C., Wall, J. S., Hainfeld, J. F., Trus, B. L. & Steven, A. C. (1985) *J. Virol.* **54,** 598–607.
4. Emerson, S. U. & Yu, Y.-H. (1975) *J. Virol.* **15,** 1348–1356.
5. De, B. P. & Banerjee, A. K. (1985) *Biochem. Biophys. Res. Commun.* **26,** 40–49.
6. Hudson, L. D., Condra, C. & Lazzarini, R. A. (1986) *J. Gen. Virol.* **67,** 1571–1579.
7. Gill, D. S., Chattopadhyay, D. & Banerjee, A. K. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 8873–8877.
8. Chattopadhyay, D. & Banerjee, A. K. (1987) *Cell* **49,** 407–414.
9. Gill, D. S. & Banerjee, A. K. (1985) *J. Virol.* **55,** 60–66.
10. Rae, B. P. & Elliott, R. M. (1986) *J. Gen. Virol.* **67,** 1351–1360.
11. Masters, P. S. & Banerjee, A. K. (1987) *Virology* **157,** 298–306.
12. Chattopadhyay, D. & Banerjee, A. K. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 8932–8936.
13. Wang, D., Villasante, A., Lewis, S. A. & Cowan, N. (1986) *J. Cell Biol.* **103,** 1903–1910.
14. Moyer, S. A., Baker, S. C. & Lessard, J. L. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 5405–5409.
15. Hill, V. M., Harmon, S. A. & Summers, D. F. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 5410–5413.
16. Hsu, C.-H., Kingsbury, D. W. & Murti, K. G. (1979) *J. Virol.* **32,** 309–313.
17. Dillon, P. J. & Gupta, K. (1988) *J. Virol.* **62,** 1582–1589.
18. Hsu, C.-H., Morgan, E. M. & Kingsbury, D. W. (1982) *J. Virol.* **43,** 104–112.
19. Ptashne, M. (1986) *Nature (London)* **312,** 697–701.
20. Struhl, K. *Cell* **49,** 295–297.
21. Ma, J. & Ptashne, M. (1987) *Cell* **52,** 179–184.
22. Giniger, E. & Ptashne, M. (1987) *Nature (London)* **330,** 670–672.
23. Lech, K., Anderson, K. & Brent, R. (1988) *Cell* **51,** 113–119.
24. Kakidani, H. & Ptashne, M. (1988) *Cell* **52,** 161–167.
25. Webster, N., Jin, J. R., Green, S., Hollis, M. & Chambon, P. (1988) *Cell* **52,** 169–178.
26. Rae, B. P. & Elliott, R. M. (1986) *J. Gen. Virol.* **67,** 2635–2643.
27. Keene, J. D., Thornton, B. J. & Emerson, S. U. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 6191–6195.