Identification of the amino acid sequence in Sindbis virus nsP4 that binds to the promoter for the synthesis of the subgenomic RNA

Mei-Ling Li* and Victor Stollar**

*Department of Molecular Genetics, Microbiology, and Immunology, University of Medicine and Dentistry of New Jersey—Robert Wood Johnson Medical School, Piscataway, NJ 08854; and **Cancer Institute of New Jersey, New Brunswick, NJ 08901

Edited by Wolfgang K. Joklik, Duke University Medical Center, Durham, NC, and approved May 11, 2004 (received for review February 11, 2004)

A gel mobility-shift assay was used to demonstrate the binding of the Sindbis virus transcriptase to the promoter for the synthesis of subgenomic (SG) RNA. The assay made use of a P15 fraction (the cell fraction that is pelleted at 15,000 × g) from cells infected with recombinant vaccinia virions expressing various Sindbis virus nonstructural proteins (nsPs) and a 32P-labeled 24-mer oligoribonucleotide representing the minimal sequence with SG promoter activity. By itself, nsP4, the viral RNA-dependent RNA polymerase, did not bind to the SG promoter; rather, all four nsPs were required for the binding of the transcriptase to the promoter. UV crosslinking of the transcriptase to a thioU-containing SG promoter, followed by V8 protease digestion of the complex, generated a peptide fragment that was bound to the SG promoter. This peptide fragment contained a sequence that corresponded to residues 329–334 of nsP4. This peptide may be in the fingers domain of nsP4. The peptide that was identified contained Arg residues at positions 331 and 332. Another Arg is present at position 327. By changing the amino acid sequence in Sindbis virus nsP4 that binds to the promoter for the synthesis of the subgenomic RNA.

Sindbis virus (SV) is the prototype virus of the family Togaviridae, genus Alphavirus. Its positive-strand 11,703-nt RNA genome has a type 0 cap at the 5′ end (1), and the 3′ end is polyadenylated. The four nonstructural proteins (nsPs) nsP1, nsP2, nsP3, and nsP4 are encoded by the 5′ terminal two-thirds of the viral genome, and the three structural proteins are encoded by a subgenomic (SG) RNA, the sequence of which is identical to the 3′ terminal one-third of the viral genome (2, 3).

Three species of viral RNA are made in SV-infected cells: (i) a positive-strand genomic RNA, (ii) a genome-length negative-strand RNA, and (iii) a positive-strand SG RNA. The genomic RNA serves as the template for the genome-length negative-strand RNA, which in turn serves as the template for both the genomic RNA and the positive-strand SG RNA.

nsP4 has been identified as the viral RNA-dependent RNA polymerase (RDRP) (4, 5). The function of nsP3, the only SV nsP that is phosphorylated (6, 7), is not known. nsP2 has an RNA helicase domain (8, 9) and possesses protease activity. It is responsible for the processing of two large polyproteins, P123 and P1234, from which the four nsPs are derived. nsP1 has both guanylyl transferase and methyltransferase activity and is thus involved in the formation of the 5′ cap on the positive-strand RNAs (10–12).

As suggested by Kao et al. (13), we shall use the term “replicase” to denote the enzyme complex that is responsible for the replication of the genome, i.e., the synthesis of both the positive- and negative-strand genome-length RNAs. Because the SG RNA serves only as messenger RNA, the complex that is responsible for its synthesis will be referred to as the “transcriptase.” Both the replicase and the transcriptase are thought to contain the RDRP (i.e., nsP4), the other virus-coded nsPs, and, possibly, cellular proteins. The designation RDRP will be used to refer only to nsP4, the subunit of the replicase or transcriptase that carries out the polymerization.

The synthesis of the different species of viral RNA is temporally regulated by the processing of the P123 and P1234 polyproteins (2, 14). According to the reports of Lemm et al. (2, 14) and Wang et al. (16), complexes composed of either nsP4 and uncleaved P123 or nsP4, nsP1, and uncleaved P23 are required for the synthesis of negative-strand RNA. These complexes can also initiate genomic and SG RNA synthesis, but only inefficiently. When P123 or P23 is processed (the cleavage of P23 is the critical event), the synthesis of negative-strand RNA ceases and the synthesis of positive-strand RNA becomes more efficient. These events occur relatively early in the viral replication cycle. Thus, during most of the replication cycle, the genomic positive-strand replicase and the transcriptase are composed of fully processed nsP1, nsP2, nsP3, and nsP4. The observation, with certain temperature-sensitive SV mutants, of an inverse relationship at high temperatures between the synthesis of SG RNA and the synthesis of negative-strand RNA suggested to Fata et al. (17) that the switch from the synthesis of the latter to the synthesis of SG RNA was associated with the cleavage of P23. Exactly how the processing of P123 and P1234 regulate the synthesis of viral RNA is not understood.

The SG RNA, which encodes the three viral structural proteins E2, E1, and C, is synthesized by internal initiation on the genome-length negative-strand RNA. In mammalian and avian cells, the fully active promoter sequence for the synthesis of SG RNA (SG promoter) is contained within 112 nt, from positions −98 to +14 relative to the start site of SG RNA transcription (18). The minimal sequence essential for promoter activity extends from nucleotide −19 to +5 (19) and corresponds to the positive-strand sequence from nucleotide 7579 to nucleotide 7602. We will refer to this minimal promoter as the SGprom.min.

No work has been reported describing which component of the viral transcriptase recognizes the SG promoter, although it would appear likely to be nsP4, the RDRP.

We recently described the isolation of a mutant of SV, SVPZF, which can grow in cells with low levels of UTP and CTP (20). Subsequently, we found that SVPZF has a second phenotype: its replication in BHK cells is severely restricted because of a marked reduction in the synthesis of SG RNA and the viral structural proteins (21). Addition of adenosine to SVPZF-infected BHK cells completely reverses all aspects of the restriction phenotype, including the decreased synthesis of SG RNA and the viral structural proteins. Only one of the three SVPZF mutations in nsP4, C7593A, is needed to produce the restricted phenotype. It is notable, however, that in addition to changing...
the amino acid at position 609 of nsP4, this mutation also alters the SG promoter at position 5. We suggest, on the basis of our findings and the demonstration that the addition of adenosine to BHK cells leads to a large increase in the level of ATP, that (i) the mutation in the SG promoter, together with the amino acid change at position 609 of nsP4, decreases the stability of the transcriptase initiation complex that is assembled at the promoter for the synthesis of SG RNA and (ii) because the initiating nucleotide of the SG transcript is A, high levels of ATP serve to stabilize the transcriptase initiation complex and thereby reverse the effects of the C7593A mutation.

Our work with SVFZ led us to investigate the binding of the SV transcriptase to the SG promoter. Here, we report that nsP4 is the protein that recognizes the SG promoter but that, in order for nsP4 to bind to the SG promoter, the other three nsPs must be present. We also identify an amino acid sequence in nsP4 involved in its binding to the SG promoter.

Materials and Methods

Cells and Viruses. BSC40 cells were grown in MEM supplemented with 10% FBS. CV-1 cells were grown in DMEM supplemented with 10% FBS. Recombinant vaccinia viruses encoding SV P123, nsP1, nsP2, nsP3, or T7 polymerase (VTF7-3) were kindly provided by Charles M. Rice and Richard Hardy (The Rockefeller University, New York and Indiana University, Bloomington, IN, respectively) and were propagated in BSC40 cells as described previously (14). The SV P123 we used contained the N614D change in nsP2 that results in more rapid processing of P123 (2, 22) than is observed with the WT P123. The various nsPs were all expressed from a T7 promoter.

To construct a recombinant vaccinia virus containing a mutant nsP4 gene, the desired mutation was introduced by PCR into an nsP4 gene that was in the plasmid transfer vector pTM3. The entire mutant nsP4 gene was then sequenced to rule out the presence of extraneous mutations. Recombinant vaccinia viruses were generated by marker rescue on CV-1 cells (23), then identified and purified by the gpt selection method (24) and grown in BSC40 cells.

Expression and Preparation of Recombinant SV nsPs. BSC40 monolayers in T75 tissue culture flasks were infected with recombinant vaccinia virus vectors encoding the desired nsPs of SV (P123 and nsP1, nsP2, nsP3, and nsP4, as indicated in specific experiments) and T7 RNA polymerase (VTF7-3), each at a multiplicity of infection of one plaque-forming unit per cell. In the ensuing 6-7 h, an infectious SV (P123 and nsP1, nsP2, nsP3, and nsP4, as indicated in specific experiments) and T7 RNA polymerase (VTF7-3), each at a multiplicity of infection of one plaque-forming unit per cell. In the ensuing 6-7 h, an infectious SV transcriptase complex. Because the P123 sequence contains the region of the protein that is involved in the binding, because nsP4 is the SV RDRP, we assumed that it is the protein involved in the recognition of the SG promoter; however, this has never been clearly demonstrated. Thus, we began by determining the necessary and optimum conditions for the binding of the nsPs to the SGprmtr.min.

Preparation of Labeled Oligoribonucleotides. Twenty-four-mer oligoribonucleotides representing the negative-strand region of the genome were synthesized by Dharamco Research (Lafayette, CO). As already noted, this is the minimal sequence that has SG promoter activity. Where indicated, thiodine (thioU) was substituted for U at the +13 position of the oligoribonucleotide, giving SGprmtr.min.thioU. Oligoribonucleotides were labeled at their 5' ends by using polynucleotide kinase and [γ-32P]ATP.

Assay for Binding the SV Transcriptase to the SG Promoter. A P15 fraction (2.5 μg of protein) prepared from BSC40 cells infected with recombinant vaccinia virus expressing T7 polymerase and various SV nsPs was incubated for 30 min at 25°C with one of the 32P-end-labeled oligoribonucleotides (10 ng, 1 × 105 cpm) as described above. The reaction was carried out in binding buffer (see below), and the final volume of the reaction mixture was 10 μl. A gel mobility-shift assay was then carried out to monitor the binding of nsP4 to the SG promoter (25).

Identification of the Amino Acid Sequence in the nsP4 Region That Binds to the SG Promoter. A reaction mixture was set up containing 25 μg of a P15 fraction prepared from BSC40 cells infected with vaccinia virus recombinants expressing T7 polymerase, P123, and nsP4, and 50 ng (1 × 105 cpm) of 32P-end-labeled SGprmtr.min.thioU. The mixture (final volume of 300 μl), which contained binding buffer (10 mM Hepes, pH 7.5/150 mM KCl/0.5 mM EGTA/2 mM MgCl2/1 mM DTT/1 unit RNasin/10% glycerol), was incubated for 30 min at 25°C, then placed on ice and exposed to UV light for 30 min (366 nm at a distance of 2 cm) (25). The crosslinked protein–oligoribonucleotide complexes were separated from free oligoribonucleotide by electrophoresis on 10% denaturing gels. The gel slices containing the crosslinked complexes were diazoy overnight to achieve equilibration with the buffers used for protease digestion. The proteins in the gel slices were then digested with chymotrypsin, trypsin, or V8 protease (0.3 μg) for 4 h at 25°C, after which the gel slices were placed on a 20% SDS polyacrylamide gel and fractionated by electrophoresis. After determining that only the V8 protease yielded a single peptide-probe band, the reaction mixture was scaled up: 750 μg of the P15 fraction prepared from BSC40 cells expressing T7 polymerase, P123, and nsP4 was incubated with 80 ng (1 × 105 cpm) of 32P-end-labeled SGprmtr.min.thioU and irradiated. After electrophoresis, the shifted band (see Results) was digested with V8 protease, and the products were fractionated by SDS/PAGE. The band containing the labeled SGprmtr.min.thioU and the bound peptide that remained after protease digestion was cut out and eluted. The peptide was microsequenced by the Protein Chemistry Section of the Howard Hughes Medical Institute Biopolymer Facility and W.M. Keck Foundation Biotechnology Resource Laboratory at Yale University by using automated Edman degradation.

Antisera. Rabbit antisera specific for each of the four SV nsPs were kindly provided by Jim and Ellen Strauss (California Institute of Technology, Pasadena, CA).

Results

Binding of the SV Transcriptase to the SG Promoter Requires the Presence of All Four nsPs. We began these studies with the goal of determining which component of the SV transcriptase recognizes and binds to the SG RNA promoter and identifying the region of the protein that is involved in the binding. Because nsP4 is the SV RDRP, we assumed that it is the protein involved in the recognition of the SG promoter; however, this has never been clearly demonstrated. Thus, we began by determining the necessary and optimum conditions for the binding of the nsPs to the SGprmtr.min.

A 24-mer oligoribonucleotide representing the SGprmtr.min was synthesized (see Materials and Methods), labeled with 32PO4 at its 5' end, and incubated with P15 fractions from BSC40 cells in which various SV nsPs were expressed by infection with recombinant vaccinia virions. Binding of one or more proteins to the SGprmtr.min would be recognized by a slower migration of the labeled oligoribonucleotide, i.e., a gel mobility shift.

As shown in Fig. 1, lane 2, the P15 extract from cells infected with vaccinia virus expressing only the T7 RNA polymerase did not cause a gel mobility shift and served as a control. Nor was a mobility shift seen when nsP4 was expressed alone (Fig. 1, lane 10) or with any one or two of the other nsPs. Only when all four nsPs were expressed (Fig. 1, lane 3) was a mobility shift observed. Thus, we conclude that binding to the SG promoter by the SV
transcriptase complex requires the participation of all four SV nsPs.

We next wished to know whether the shifted band seen in the gel mobility-shift assay contained only nsP4 or the entire complex of the four nsPs. Accordingly, a gel mobility-shift experiment was carried out as in Fig. 1: the band was cut out, subjected to SDS/PAGE, and then analyzed by Western blotting with antibodies to nsP1, nsP2, and nsP3. As shown in Fig. 2, nsP1, nsP2, and nsP3 were all present in the transcriptase–SGprmtr.min complex. Because nsP4 is the RDRP, it was assumed that nsP4 was also present in this complex; that this is indeed the case is shown below by the immunoprecipitation of the nsP complex by anti-nsP4 serum (see the experiment illustrated in Fig. 7).

We also carried out the binding assay by using an unrelated 31-mer oligoribonucleotide (kindly provided by Gary Brewer, Robert Wood Johnson Medical School, Piscataway, NJ) as the labeled probe to assess the specificity of the interaction between the SV nsPs and the SG promoter. As expected, expression of T7 polymerase by itself did not cause a mobility shift when incubated with either SGprmtr.min or the unrelated 31-mer (Fig. 3, lanes 2 and 5). Nor was a mobility shift seen when the four SV nsPs were incubated with the 31-mer oligoribonucleotide (Fig. 3, lane 6). However, as shown in Fig. 1, a mobility shift was observed after the four nsPs were incubated with SGprmtr.min (Fig. 3, lane 3). Thus, the binding of the SV nsPs to the SG promoter is sequence-specific.

Identification of the nsP That Binds to the SG Promoter. Having determined the necessary conditions for the recognition of the SG promoter by the SV transcriptase, we next wished to determine which of the nsPs is involved in recognition of the SG promoter. The approach we used to do this was to crosslink the transcriptase to the SGprmtr.min and then identify the protein bound to the promoter. For this experiment we took advantage of the very high efficiency with which thioU residues crosslink with amino acids after exposure to UV light (26, 27). Accordingly, a 24-nt oligoribonucleotide was synthesized, similar to that used in the preceding experiments except that a thioU was substituted for U at position −13 of the negative-strand SGprmtr.min (corresponding to nucleotide 7585 on the positive-strand RNA; see Fig. 4). This oligoribonucleotide is referred to as SGprmtr.min.thioU. Extracts were prepared from BSC40 cells expressing T7 polymerase, P123, and nsP4, and gel mobility-shift assays were carried out. Lanes 2 and 4 in Fig. 4 show the gel mobility shifts with the labeled SGprmtr.min and the SGprmtr.min.thioU, respectively, and demonstrate that the substitution of a thioU for U at the −13 position of the oligonucleotide did not affect recognition by the SV transcriptase.

Having demonstrated that the SV transcriptase can bind to SGprmtr.min.thioU, which contains the indicated thioU-for-U substitution, we incubated the SGprmtr.min.thioU with a P15 extract from cells expressing all four SV nsPs and then exposed the mixture to UV light. As expected, gel electrophoresis of the reaction mixture showed the presence of a labeled band that migrated more slowly than the free oligonucleotide (Fig. 5, lane 2). The proteins in the band shown in lane 2 were then digested with trypsin, chymotrypsin, or V8 protease, and the resulting products were analyzed by gel electrophoresis (data not shown). Trypsin digestion did not reveal any distinct band that migrated more slowly than SGprmtr.min.thioU. Digestion with chymotrypsin produced two labeled bands that migrated more slowly than labeled SGprmtr.min.thioU, but the V8 protease digestion produced only a single labeled band that migrated more slowly than the labeled SGprmtr.min.thioU. This band was taken to represent a small
V8 protease digestion product covalently attached to the labeled SGprmtr.min.thioU.

The preceding experiment was then repeated on a larger scale; this time the crosslinked shifted protein–SGprmtr.min.thioU band was digested only with V8 protease, and the peptide fragment obtained was sequenced. Six amino acids, NH2-LVRRLT, were identified with certainty, and two more amino acids, AV, were identified as probable. The sequence of the first six amino acids listed above matched a sequence in nsP4 extending from amino acid 329 to 334. This result identified nsP4 as the component of the viral transcriptase that recognizes and binds to the SG promoter. Furthermore, it localized the sequence in nsP4 that recognizes and binds to the SG promoter. According to the molecular model of nsP4 that we had generated (20), this sequence is located in the fingers domain of the protein.

Two Args in the 327–346 aa Sequence of nsP4 Are Essential for Binding the Polymerase to the SG Promoter.

The amino acid sequence of nsP4 extending from residues 327 to 346 contains three Args, R327, R331, and R332 (Fig. 6A), all of which are potential candidates for interacting with the negatively charged SGprmtr.min. To determine whether any of these Args were required for RNA binding, each Arg codon was individually replaced with an Ala codon by site-directed mutagenesis, and the resulting mutated nsP4 gene was inserted into a vaccinia virus vector (28). BSC40 cells were then coinfected with vaccinia virus vectors expressing the WT P123 and either the WT nsP4 or one of the mutant nsP4 proteins. The 32P-end-labeled SGprmtr.min was incubated either alone (lane 1) or with P15 extracts from cells expressing only T7 polymerase (lane 2), or T7 polymerase, P123, and either WT nsP4 (lane 3) or mutant nsP4 with the indicated changes (lanes 4–6). The protein-promoter complexes were then analyzed on a 4% nondenaturing gel. T7 pol., T7 polymerase.
more of the other nsPs and thus only indirectly interfered with the binding of the transcriptase to the SG promoter. If this were the case, the nsPs would not form the proper multiprotein complex required for binding to the SG promoter. To test this possibility, we expressed the four nsPs in BSC40 cells with either WT nsP4 or nsP4 with one of the R→A changes at positions 327, 331, or 332. Anti-nsP4 was then used to precipitate the nsP complex, and the proteins in the complex were identified by Western blotting. As seen in Fig. 7, the R→A substitution in nsP4 at positions 331 and 332 had no effect on the composition of the nsP complex. Thus, it seems that these changes in nsP4 did not interfere with complex formation by the four nsPs.

**Discussion**

This project was undertaken to clarify the nature of the transcriptase initiation complex at the promoter for the synthesis of SG RNA. Two important findings have emerged from our studies. Whereas it was not surprising that the component of the transcriptase that bound to the SGpromtr.min was nsP4, it was not predicted that this promoter would be recognized by nsP4 only in the presence of the other three nsPs. This finding supports the idea that as components of the viral transcriptase, the four nsPs function as a tightly knit complex and that the activity and conformation of the RDRP is strongly influenced by the other nsPs. That the four alphavirus nsPs exist as a multiprotein complex was first shown by Barton et al. (29). Using antibodies specific for each of the nsPs, they observed that in immunoprecipitation experiments each of the antibodies was able to pull down not only the nsP against which it was directed but also the other three nsPs. Subsequently, in a report of the first in vitro template-dependent and template-specific synthesis of SV RNA, Lemm et al. (14) showed that nsP4 alone lacked polymerase activity and that only when all four nsPs were present was the template (a positive-strand RNA in this case) copied. Our finding that binding of the SV transcriptase to the SG promoter required the presence of all four nsPs is consistent with the preceding observations.

The second important finding in this report is the identification of an amino acid sequence in nsP4 that is critical for recognition and binding to the SG promoter. As shown in Fig. 6, replacement of either R331 or R332 by Ala eliminated the binding of nsP4 to the SG promoter. These results highlight the importance of this sequence for binding to the SG promoter. According to our molecular model, this sequence is on a β-strand in the fingers domain. A more accurate localization awaits the actual determination of the structure of nsP4. In that regard, it appears from our results and those of Barton et al. (29) and Lemm et al. (14) that the structure of nsP4 in its functional state would best be determined in the presence of the other nsPs. This suggestion may apply to other positive-strand RNA viruses.

Experiments similar to those described in this report have been done with influenza virus (25). In those experiments, the labeled probe was a 16-mer oligoribonucleotide representing the 5′ terminus of one of the viral RNA segments, and the proteins involved were those in the polymerase complex, PB1, PB2, and PA. As with the SV nsPs, to achieve binding to the probe all three of the influenza proteins had to be present. When the influenza virus polymerase complex was incubated with the probe and UV-crosslinked, a peptide fragment was obtained after protease digestion, the sequence of which was identical to a sequence in PB1. Furthermore, this protease fragment contained four Arg residues, at positions 560, 563, 571, and 572 in PB1. Site-directed mutagenesis indicated that only the Arg residues at positions 571 and 572 were required for binding to the probe. Thus, as with SV nsP4, adjacent Arg residues were critical for binding to the influenza virus probe, suggesting that adjacent Arg residues may be a common feature of RDRPs. Whether R331 and R332 of SV nsP4 are also involved in the binding of the nsP complex to the promoters for the synthesis of positive- and negative-strand genome-length RNA remains an open question.

Several cellular proteins have been reported to bind to conserved sequence elements in the 3′ UTR of SV RNA (30). These proteins, one of which has been identified as the mosquito homolog of the La protein, are presumably involved in the replication of viral RNA and might be components of the viral replicase. An interesting question arises as to whether any of these proteins also plays a role in the functioning of the viral transcriptase. A related question is whether the site on nsP4 that recognizes the SG promoter may also be involved in the recognition of the promoters for the synthesis of positive- and negative-strand viral RNA; however, these promoters have not yet been precisely defined.

We would like to thank Dr. Stefan Sarafianis for helpful discussion and Yan Hu for technical assistance. This work was supported by U.S. Public Service Grant AI-49273.