JOURNAL OF VIROLOGY, Aug. 1993, p. 5003–5011
0022-538X/93/085003-09$02.00/0

Vol. 67, No. 8

Multiple Binding Sites for Cellular Proteins in the 3′ End of Sindbis Alphavirus Minus-Sense RNA

NATHALIE PARDIGON, EDITH LENCHES, AND JAMES H. STRAUSS*
Division of Biology, California Institute of Technology, Pasadena, California 91125

Received 27 January 1993/Accepted 4 May 1993

The 3′ end of Sindbis virus minus-sense RNA was tested for its ability to bind proteins in mosquito cell extracts, using labeled riboprobes that represented different parts of this region. We found four domains in the first 250 nucleotides that could bind the same 50- and 52-kDa proteins, three with high affinity and one with low affinity, whereas tested domains outside this region did not bind these proteins. The first binding domain was found in the first 60 nucleotides, which represents the complement of the 5′-nontranslated region, the second in the next 60 nucleotides, the third in the following 60 nucleotides, and the fourth between nucleotides 194 and 249 (all numbering is 3′ to 5′). The relative binding constants, K_r, of the first, second, and fourth sites were similar, whereas that of domain 2 was fivefold less. Deletion mapping of the first domain showed that the first 10 nucleotides were critical for binding. Deletion of nucleotides 2 to 4, deletion or replacement of nucleotide 5, or deletion of the first 15 nucleotides was deleterious for binding, deletion of nucleotides 10 to 15, 26 to 40, or 41 to 55 had little effect on the binding, and deletion of nucleotides 15 to 25 increased the binding affinity. We also found that the corresponding riboprobes derived from two other alphaviruses, Ross River virus and Semliki Forest virus, and from rubella virus were also able to interact with the 50- and 52-kDa proteins. The K_r value for the Semliki Forest virus probe was similar to that for the Sindbis virus probe, while that for the Ross River virus probe was four times greater. The rubella virus probe was bound only weakly, consistent with the fact that mosquito cells are not permissive for rubella virus replication. We suggest that the binding of the 50- and 52-kDa proteins to the 3′ end of alphavirus minus-sense RNA represents an important step in the initiation of RNA replication.

Conserved sequence elements in alphavirus RNAs have been proposed as possible promoters for initiation of replication and transcription from viral RNA (22). The different effects of several mutations within these elements upon virus growth in different cell lines also led to the hypothesis that they function as promoters by binding to host cell factors (9, 15, 16) and that therefore these elements should be conserved during viral evolution. Of the four conserved elements that have been described for alphaviruses, the 24-nucleotide-long promoter for transcription of the subgenomic RNA has been best characterized (11, 19). Mutational analyses of the three other elements, the 19 nucleotides at the 3′ end of the genomic RNA, a 51-nucleotide sequence that begins around nucleotide 150 in the genome, and the first 44 nucleotides of the genomic RNA, have confirmed their importance in viral RNA transcription and replication (9, 15, 16).

The 44 residues at the 5′ end of the genomic RNA and the complementary sequence in the minus strand are potentially able to form stem-loop structures, which in the minus strand could be involved in the initiation of plus-strand RNA synthesis. Three deletions tested in this region proved to be lethal for the virus, namely, deletion of nucleotide 1, deletion of nucleotides 2 to 4, or deletion of nucleotide 5 (numbering in the genomic sense) (15). The first two of these would disrupt part of the stem, and the third causes a change in a bulged nucleotide in the stem. Deletion of nucleotide 8, 36, or both shortened the stem and led to temperature-sensitive virus. Interestingly, deletions in this region often affected virus growth differently in cultured cell lines derived from the chicken or mosquito, both of which are natural hosts for the virus. It was concluded that host factors interacting with this sequence element might be involved in viral RNA transcription and replication.

Using the complement of the first 132 bases of the Sindbis virus genome as a probe in binding assays, we demonstrated that this domain could interact with three chicken proteins of 42, 44, and 52 kDa, the first two migrating as a doublet (18). These proteins still interacted with the probe after deletion of residue 5, which is lethal for the virus (16), but surprisingly, the half-life of the protein-RNA complexes with the mutant probe was threefold longer than that of the wild-type complexes, indicating that the chicken proteins bound more tightly to the mutant probe. A truncated probe containing only the 3′-terminal 60 nucleotides, i.e., the complement of the 5′-nontranslated region (NTR), also bound the same three chicken proteins.

We describe here our results with probes representing domains in the 3′-terminal region of Sindbis virus minus-sense RNA upon incubation with extracts from mosquito cells. We found that there are three high-affinity and one low-affinity binding sites for two cellular proteins within the 3′-terminal 249 residues of the minus strand. Deletion analysis of the 3′-terminal binding site showed that the 3′ 10 bases are critical for the binding of the two proteins. We also examined the binding of the complement of the 5′ NTR of two other members of the alphavirus genus, Ross River (RR) and Semliki Forest (SF) viruses, as well as that of the complement of the first 61 bases of rubella virus minus-sense RNA (which includes the 5′ NTR), which is classified as a different genus within the family *Togaviridae*. We found that the alphavirus probes as well as the rubella virus probe interacted with the same mosquito cell proteins but with different affinities and that the affinity of binding is correlated with the ability of the virus to replicate in mosquito cells.

* Corresponding author.

MATERIALS AND METHODS

Preparation of DNA clones. The isolation of cDNA clones containing the sequence of the first 62 (clone pGEX) or 132 (pGEH) bases of the wild-type Sindbis virus genome, of a deletion derivative of pGEH (pGEHd5), and of control plasmid pBEH has been described previously (18). Plasmid pRR64 (10) was obtained from R. Kuhn, and plasmid pSF was from T. Rümenapf.

To prepare derivatives of pGEX containing various deletions, we synthesized in both genomic and antigenomic polarity oligonucleotides that represented the cDNA sequence of the 5′ NTR of the virus genome containing the desired deletion. The Sindbis virus sequence was flanked by an *Eco*RI restriction site at the 5′ end and by an *Xba*I site at the 3′ end. To form double-stranded DNA, the oligonucleotides (1,000 pmol each) were heated to 95°C for 5 min in 10 mM Tris HCl (pH 7.5)–10 mM $MgCl_2$–10 mM 2β-mercaptoethanol–100 mM NaCl and then slowly cooled to room temperature. The double-stranded cDNA was then digested with *Eco*RI and *Xba*I, purified by gel electrophoresis, and inserted into the transcription vector pGEM4 (Promega) prepared by digestion with the same enzymes. In the same manner, synthetic oligonucleotides representing the regions between nucleotides 129 and 196, 552 and 604, and 5252 and 5305 of the Sindbis virus genome and flanked by same restriction sites were cloned into pGEM4. Recombinant plasmids pGEXd(2-4), pGEXd(1-15), pGEXd(10-15), pGEXd(15-25), pGEXd(26-40), and pGEXd(41-55) contain deletions of nucleotides 2 to 4, 1 to 15, 10 to 15, 15 to 25, 26 to 40, and 41 to 55, respectively, while recombinant plasmids pNEX, pGEX194, pGEX552, and pGEX5252 contained the nucleotide sequences between residues 129 and 196, 194 and 249, 552 and 604, and 5252 and 5305 of the Sindbis virus genome, respectively. Deleted sequences of the 5′ NTRs of RR and SF viruses, as well as the sequence of the first 61 bases of rubella virus RNA and of a deletion derivative, all flanked by *Eco*RI and *Xba*I restriction sites, were also cloned into vector pGEM4 [pRRd(1-20), pSFd(1-20), pRUB, and pRUBd(1-20)].

To prepare cDNA clones suitable for transcription of riboprobes from RR and SF viruses, we synthesized RNA transcripts from the 5′ end of the genomes from pRR64 linearized at a *Bss*HI site or pSF linearized at an *Eco*RV site, using SP6 RNA polymerase in a mixture containing 500 μM each ATP, GTP, CTP, and UTP for 15 min at 38°C. The transcripts were then used as templates for cDNA synthesis, with a primer annealing to nucleotides 61 to 88 of RR virus RNA or nucleotides 68 to 95 of SF virus RNA, both primers containing an *Xba*I site. The first-strand cDNA products were amplified by the polymerase chain reaction with a primer containing an *Eco*RI site followed by the first 15 bases of either RR or SF virus genome together with the original cDNA primers. The products were purified by gel electrophoresis after digestion with *Eco*RI and *Xba*I and inserted in *Eco*RI-*Xba*I-digested pGem4. The recombinant plasmids (called pRR and pSF respectively) contained the 5′ NTR of the RR or SF virus genome.

To obtain a cDNA clone corresponding to Sindbis virus residues 59 to 132, we transcribed RNA from Sindbis virus 5NTd(10-14) (16) into cDNA using a primer annealing to nucleotides 126 to 145. Polymerase chain reaction amplification followed with this primer and a primer annealing to nucleotides 46 to 67 that contained a mismatch at position 58 (G instead of C) and a A inserted between nucleotides 57 and 58 so as to introduce an *Xba*I site. After purification of the *Xba*I-*Hin*dIII fragment on a polyacrylamide gel, the cDNA was inserted into vector pGEM4 prepared by digestion with the same enzymes, to give plasmid pGXH.

Plasmids containing the 5′ NTR of Sindbis virus RNA with a deletion of nucleotide 5 (pGEXd5) or the substitution of this nucleotide by G (pGEXG5) were obtained by polymerase chain reaction amplification from pGEX DNA with the primer annealing to nucleotides 46 to 67 that contained mismatches to produce an *Xba*I site (see above) together with a primer containing an *Eco*RI site followed by the first 11 bases of Sindbis virus RNA containing either a deletion or a substitution (A to G) of nucleotide 5. After restriction enzyme digestion and gel purification, the cDNA was inserted into pGEM4.

Transcription of riboprobes. Riboprobes were transcribed from cDNA plasmids with T7 RNA polymerase in the presence of [α-^{32}P]UTP and [α-^{32}P]CTP (8,000 Ci/mmol; Amersham) or of [^{3}H]UTP (Amersham), as described previously (18).

Preparation of cell extracts. Monolayers of mosquito C6/36 cells from *Aedes albopictus* were grown to confluence in 150-mm-diameter dishes as described previously (9). P100 salt wash extracts were prepared as described previously (7). The extracts were concentrated 2.5-fold on a Centricon 10 device (Amicon) according to the manufacturer's protocol. Protein concentration was determined by the Bio-Rad protein assay.

RNA-protein complex formation and gel retardation analysis. Binding reactions and gel retardation analysis were performed as described previously (12, 18). Quantitation of the RNA-protein complexes and of free RNA was performed either by scanning the dried gels with a Phosphorimager scanner (Molecular Dynamics) or by excising the complexes and free probe from the gels and counting the radioactivity of each in a scintillation counter. Data were transformed as previously described (3, 18), and values for P_o and K_r were extracted. Each binding experiment was performed more than once [four times with probes 62S(−), 129/196S(−), 62S(−)d1-15, 62S(−)d2-4, 62S(−)d5, 62S(−)d15-25, 62S(−)d26-40, 62S(−)d41-55; twice with probes 59/132S(−), 62S(−)G5, 62S(−)d10-15, RR(−), SF(−), RUB(−)] except for probe 194/249(−). The values reported for P_o and K_r are the average value obtained. The standard deviation in these values from experiment to experiment was estimated to be ±30 to 50%.

UV-induced cross-linking and analysis of cross-linked proteins were performed as described previously (18).

RESULTS

Synthesis of riboprobes. Radiolabeled minus-sense probes were transcribed from cDNA clones with T7 RNA polymerase. Most probes contained about 60 nucleotides of Sindbis virus sequence. Probes 62S(−), 59/132S(−), 129/196S(−), 194/249S(−), 552/604S(−), and 5252/5305S(−) contained either the first 62 bases of Sindbis virus minus-strand RNA, numbered from 3′ to 5′ (which comprises the complement of the 5′ NTR and the next three bases), or bases 59 to 132, 129 to 196, 194 to 249, 552 to 604, or 5252 to 5305, respectively. The first four probes are illustrated in Fig. 1A and also contained 19, 8, 19, or 17 nonviral bases at the 5′ end and 5, 4, 2, or 2 extra bases at the 3′ end, respectively. A longer probe that contained the first 132 bases of Sindbis virus genome, 132S(−), was also used and contained eight additional nucleotides at the 5′ end and five extra bases at the 3′ end (Fig. 1A).

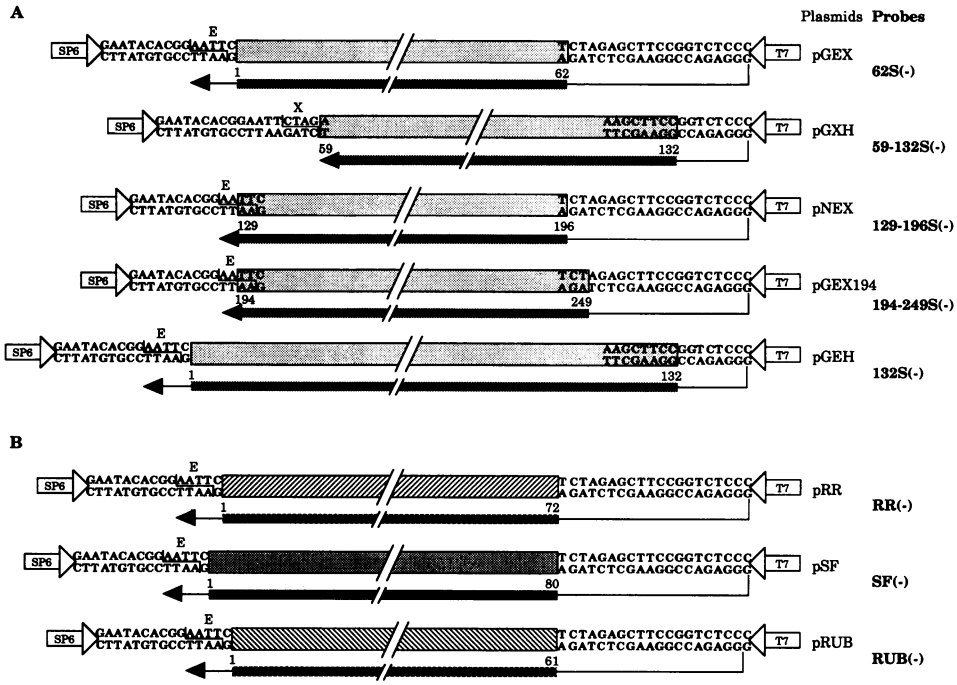

FIG. 1. Schematic diagrams of in vitro transcripts used as probes in binding assays. The shaded boxes represent viral sequences which have been inserted in the polylinker of pGem4 transcription vector. The shaded arrows represent the RNA transcripts synthesized from the corresponding plasmids with T7 RNA polymerase, after linearization at the indicated restriction enzyme site. The name of each plasmid and its corresponding probe is indicated in plain and boldface letters, respectively. Nucleotide numbers are 5′ to 3′ in the genome (plus sense). E, *Eco*RI; X, *Xba*I. (A) Plasmids and corresponding transcripts containing sequences of Sindbis virus RNA. (B) Plasmids and corresponding transcripts containing sequences of RR and SF alphaviruses and rubella virus (RUB).

Probes 62S(−)d2-4, 62S(−)d5, 62S(−)d1-15, 62S(−)d10-15, 62S(−)d15-25, 62S(−)d26-40, and 62S(−)d41-55 are identical to probe 62S(−), but nucleotides 2 to 4, 5, 1 to 15, 10 to 15, 15 to 25, 26 to 40, and 41 to 55, respectively (numbered 3′ to 5′ in the Sindbis virus minus strand), have been deleted. Another mutant probe called 62S(−)G5 contained G instead of A at position 5. Probe 132S(−)d5 is identical to probe 132S(−) except that nucleotide 5 (numbered 3′ to 5′ in the Sindbis virus minus strand) is deleted.

We also produced antisense probes RR(−), SF(−), and RUB(−) which contained the 5′ NTR of RR or SF virus RNA (72 and 80 bases, respectively) or the 5′ NTR (40 bases) and the next 21 bases of rubella virus RNA. These probes are schematically depicted in Fig. 1B; they contained 19 extra bases at the 5′ end and 5 extra bases at the 3′ end. Deletion probes in which the first 20 bases were deleted, RR(−)d1-20, SF(−)d1-20, and RUB(−)d1-20, respectively, were also produced.

As a control, an exact probe corresponding to probe 62S(−) was also produced. A double-stranded oligonucleotide was synthesized that contained the 62 3′-terminal nucleotides of the minus strand downstream of a T7 promoter. RNA transcribed from this synthetic DNA contained only one extra G at the 5′ end (required for initiation of transcription) followed by the 62 Sindbis virus nucleotides.

Mosquito protein binds to the 3′ end of Sindbis virus minus-strand RNA. We have previously shown specific binding of chicken proteins to the 3′ end of Sindbis virus minus-strand RNA using radiolabeled minus-sense probes (18) and hypothesized that the 3′-terminal region of the minus-sense genome is bound by cellular proteins to promote RNA replication. To further explore this hypothesis, we tested mosquito cell extracts for their ability to interact with minus-sense probes, mosquitoes being a natural host for alphaviruses. Formation of complexes between the 132S(−) probe or its mutagenized derivative 132S(−)d5 and proteins in the mosquito cell extracts is shown in Fig. 2. Three retarded bands were observed in mobility shift assays with either the wild-type or the mutant probe: band I migrated the fastest, while bands II and III migrated more slowly (Fig. 2, lanes 2 and 6). A fourth faint band designated with an asterisk migrated slightly slower than band I but was not always present and may be nonspecific. We consistently found that the band I complexes were present in greater quantities when the mutant probe was used, while those migrating as band II were present in smaller quantities.

To assess the specificity of complex formation, we performed competition experiments in which the reaction equilibrium was displaced by adding an excess of unlabeled competitor, either a non-Sindbis virus probe (Bls) or unlabeled 132S(−) or 132S(−)d5 transcripts. Band I was found to be specific for the Sindbis virus probes (either wild type or mutant), since it could be outcompeted in both cases by the corresponding unlabeled competitor but not by the nonspecific Bls competitor (Fig. 2, lanes 3, 4, 7, and 8). However, bands II and III appeared to be specific only for the wild-type probe. They were outcompeted by the specific competitors, but in the case of the mutant probe, they were almost totally outcompeted by the nonspecific Bls competitor as well, whereas these bands were only slightly affected by the nonspecific competitor in the case of the wild-type probe (Fig. 2, lanes 4 and 8). For reasons that remain unclear, the wild-type band III was incompletely blocked by the specific 132S(−) competitor. It is possible that these

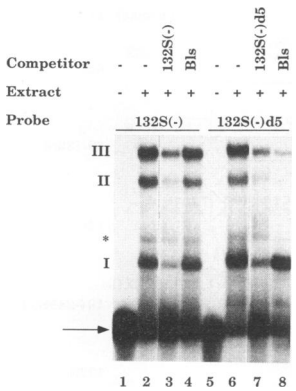

FIG. 2. Formation of complexes with the 132S(−) and 132S(−)d5 probes. Formation of complexes was done as described in Materials and Methods, with either probe 132S(−) (lanes 1 to 4) or 132S(−)d5 (lanes 5 to 8) labeled with ^{32}P and incubated with mosquito cell extract containing 500 ng of protein. Samples were analyzed on a nondenaturing 5% polyacrylamide gel. Above each lane is indicated the presence or absence of mosquito extract or of competitor and the probe used. Competitors Bls (nonspecific) or 132S(−) or 132S(−)d5 (specific) were used at 30 ng. The three major retarded bands are referred as bands I, II, and III, and a minor retarded band is indicated by an asterisk (*). The free probe is indicated by an arrow.

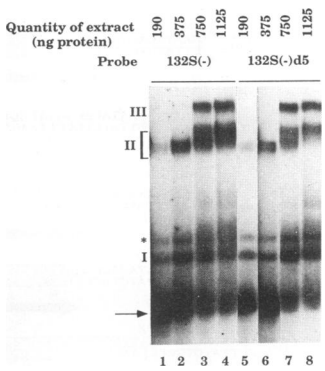

FIG. 3. Formation of complexes with increasing concentrations of mosquito cell extract. Formation of complexes was done as described in Materials and Methods, with either probe 132S(−) (lanes 1 to 4) or 132S(−)d5 (lanes 5 to 8) labeled with ^{32}P and incubated with increasing amounts of mosquito cell extract. Samples were analyzed on a nondenaturing 5% polyacrylamide gel. The amount of extract protein is indicated above each lane. The three major retarded bands are referred as bands I, II, and III. A minor retarded band is indicated by an asterisk (*). The position of free probe is indicated by an arrow.

complexes require a greater quantity of competitor to be outcompeted or that a small portion of the complexes present in this band are formed by nonspecific interactions. Cross-competition experiments demonstrated that when the labeled probe was wild type, the unlabeled mutant probe was able to outcompete band I but not band II or III, whereas the unlabeled wild-type probe could outcompete bands I, II, and III formed with the mutant probe (data not shown). Thus, a stable and specific complex I can be formed with either the wild-type or mutant probes, and this complex appears to represent a complex similar to that formed with proteins from chicken extracts. Complexes II and III are stable and specific when formed with the wild-type probe, but are unstable when formed with the mutant probe.

To study further the formation of the higher-molecular-weight complexes in bands II and III, we added increasing quantities of mosquito cell extract to binding reactions with a constant amount of either the wild-type or mutant probe (Fig. 3). We found that with the wild-type probe, the complexes present in band I were maximal when 375 ng of extract protein was added (Fig. 3, lane 2) and thereafter decreased with increasing amount of extract, whereas the quantity of complexes present in band II increased, suggesting that with increasing protein concentration the band I complexes were converted into band II complexes (Fig. 3, lanes 3 and 4), presumably by binding additional protein. (Note that, depending on the batch of extract, band II sometimes appeared as a single band, as in Fig. 2, for example, or as multiple closely spaced bands as in Fig. 3; these differences did not affect the nature of the conclusions reached.) Band III, on the other hand, was formed only at high protein concentration (Fig. 3, lanes 3 and 4), suggesting that band II may be converted to band III. The mutant probe needed twice as much extract (750 ng of protein) to reach saturation of band I, and these appeared to be only inefficiently converted to bands II and III (Fig. 3, lanes 5 to 8). Fewer higher-molecular-weight complexes were formed with the mutant probe and they were less stable than those formed with the wild-type probe.

Binding of mosquito protein to the 62S(−) probe. We had previously shown that a truncated probe containing only the complement of the 5′ NTR, probe 62S(−), formed a complex with chicken proteins (18). Upon incubation of mosquito cell extracts with radiolabeled 62S(−), two major retarded bands were formed (Fig. 4A, lane 2, bands I and III), which comigrated with bands I and III detected with the 132S(−) probe (data not shown). Two minor bands were also visible in the experiment shown in Fig. 4A, one of which is marked with an asterisk; these minor bands were not consistently observed and depended on the batch of extract used. Only band I appeared to be specific with the 62S(−) probe. Band I could be outcompeted by unlabeled 62S(−) but not by nonspecific competitor Bls (Fig. 4A, lanes 3 and 4), whereas band III as well as the minor bands were outcompeted by the nonspecific probe as well as by 62S(−). Furthermore, the quantity of complexes present in band III was more variable with the 62S(−) probe than with the 132S(−) probe. Band III formed with 62S(−) could represent the nonspecific component of band III formed with 132S(−) described above.

To assess the possible importance of the extra bases located at the extremities of the 62S(−) transcripts in the binding process, we produced labeled transcripts that represented the exact complement of the 5′ NTR of Sindbis virus genome in which one extra G was present at the 5′ end of the transcript but whose 3′ end was precise. For this, a synthetic double-stranded oligonucleotide was transcribed in vitro, as described above. In binding assays with proteins from mosquito cell extracts, these exact transcripts behaved exactly as did the 62S(−) probe (data not shown). The complexes formed were the same and the same proteins were labeled after UV cross-linking (see below). Cross-competition experiments demonstrated that the exact probe and the 62S(−) probe competed for the same proteins. Thus, the extra nucleotides at the 5′ and 3′ ends of the probes used do not affect the binding results.

Identification of mosquito proteins bound by the antisense probes. To identify the proteins that interact with probe

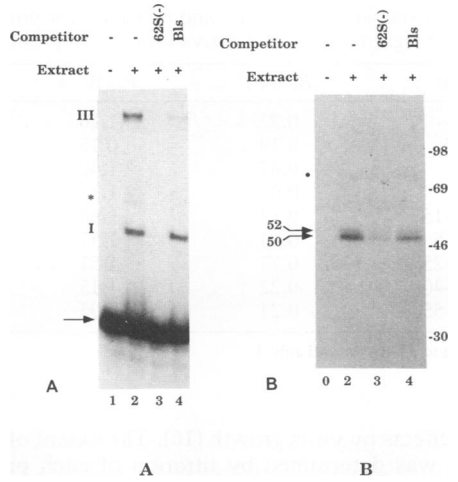

FIG. 4. Formation of complexes with the 62S(−) probe and identification of the proteins involved in formation of complexes. (A) Formation of complexes was done as described in Materials and Methods, with probe 62S(−) labeled with ^{32}P and incubated with 500 ng of mosquito cell protein. Samples were analyzed on a nondenaturing 5% polyacrylamide gel. The presence or absence of extract or competitor is indicated above each lane. The two major retarded bands are referred as bands I and III, and a minor retarded bands is indicated by an asterisk (*). Only band I is specific for the 62S(−) probe (see text). The position of free probe is indicated by an arrow. (B) Analysis of the proteins involved in the formation of the Sindbis virus specific complexes. After ^{32}P-labeled probe 62S(−) was incubated with a mosquito cell extract, the complexes were UV irradiated and the RNA probe removed with RNase. Samples were analyzed on an SDS-14% polyacrylamide gel. Lanes are as in panel A. 50 and 52 refer to two cellular proteins specifically labeled by the 62S(−) probe that have apparent molecular masses of 50 and 52 kDa, respectively. • indicates the position of a protein of 76 kDa that interacts with the 62S(−) probe nonspecifically.

62S(−), we irradiated the complexes with UV light to cross-link the probe to the bound proteins. The probes were digested with RNase, and the proteins were separated by electrophoresis on sodium dodecyl sulfate (SDS)-containing polyacrylamide gels. At least three proteins were labeled by transfer of radiolabel from the probe, with apparent masses of 50, 52, and 76 kDa (Fig. 4B, lane 2). The 76-kDa protein was only weakly labeled and could not be carefully studied, although the results suggested that this protein interacts nonspecifically with the Sindbis virus probe. The 50- and 52-kDa proteins, which migrated as a doublet, were found to be specifically bound to the 62S(−) probe from their behavior in competition experiments, in which labeling could be abolished by competition with the specific competitor but not with the nonspecific competitor (Fig. 4B, lanes 3 and 4).

Similar results were obtained when the 132S(−) or 132S(−)d5 probe was cross-linked to bound proteins. Only the 50- and 52-kDa proteins (and to a lesser extent the 76-kDa protein) were detectable on polyacrylamide-SDS gels after UV irradiation (data not shown). This result suggests that more copies of the 50- and 52-kDa proteins bind to the longer probes to form the higher-molecular-weight complexes (bands II and III) or that some proteins in the larger complexes cannot be cross-linked to the probe (see also below).

Multiple protein-binding sites in Sindbis virus minus-strand RNA. The fact that the 132S(−) probe was able to form larger stable complexes with proteins from mosquito cells than the shorter 62S(−) probe led us to search for the existence of other protein-binding sites at the 3′ terminus of the Sindbis virus minus-sense RNA. For this purpose, we used Sindbis virus RNA transcripts containing the complement of nucleotides 59 to 132, 129 to 196, 194 to 249, 552 to 604, and 5252 to 5305 as probes in binding experiments. The first three probes were able to form specific complexes with mosquito cell proteins which were similar to those observed with the 62S(−) probe, although probe 129/196S(−) bound only weakly. Cross-competition experiments showed that these probes and probe 62S(−) all competed for the same protein(s) for complex formation, although probe 129/196S(−) did not completely outcompete complex formation with labeled 62S(−), 59/132S(−), or 194-249S(−) probes (data not shown). UV cross-linking experiments confirmed that the same 50- and 52-kDa proteins interacted with these three probes (data not shown). Probes 552/604S(−) and 5252/5305S(−) were unable to form specific complexes with mosquito cell proteins (data not shown).

To compare the binding affinity of the different probes, we determined the extent of complex formation with the different probes upon titrating a constant amount of cellular extract with increasing concentrations of each probe. The proportion of free and complexed transcripts was quantitated after gel retardation analysis, and estimates of the two parameters K_r and P_o were obtained after data transformation as described previously by Calzone et al. (3). K_r is the ratio of the affinity constant for the specific binding reaction (K_s) to the affinity constant for the nonspecific binding reaction with the large excess of poly(dI-dC) present during the reaction (K_n), and P_o represents the molar concentration of the binding proteins in the binding reaction. An example of the transformed data is shown in Fig. 5, where the parameters have been determined for the 62S(−) probe. These two parameters estimated for various probes are presented in Table 1. Also shown in Table 1 are the correlation coefficients r^2 from the least-squares fit of the data transformed as in Fig. 5A, which serve as a measure of the reliability of the constants determined. Probe 59/132S(−) had the highest K_r value (0.39×10^6) and probe 129/196S(−) the lowest (0.04×10^6). The K_r values of probes 62S(−) and 194/249S(−) were somewhat less than that of 59/132S(−) but much greater than that of probe 129/196S(−). Thus, there are three high-affinity binding sites and one low-affinity binding site for the 50- and 52-kDa proteins in the 3′-terminal 249 nucleotides of the minus strand. The P_o values found for all four probes were approximately equivalent, between 1×10^{-5} and 1.8×10^{-5} M, consistent with the finding that all interact with the same set of proteins.

These results confirmed the hypothesis that the 132S(−) probe contains two contiguous binding sites such that the formation of higher-molecular-weight complexes with this probe probably arises because proteins can bind to both sites. It is interesting that when formation of complexes with labeled 132S(−) probe is inhibited with an excess of unlabeled 62S(−) or 59/132S(−) probe, band I was completely outcompeted but band II was only partially outcompeted (data not shown). This suggests that there may be some cooperativity in binding when both sites in probe 132S(−) are occupied such that the binding affinity is greater. When the 62S(−) and 59/132S(−) probes were both present in a binding reaction, only band I was produced (data not shown).

Deletion analysis of the 3′ binding site. We wished to analyze the effect of deletions within the 3′ binding site on formation of complexes with mosquito proteins. Derivatives

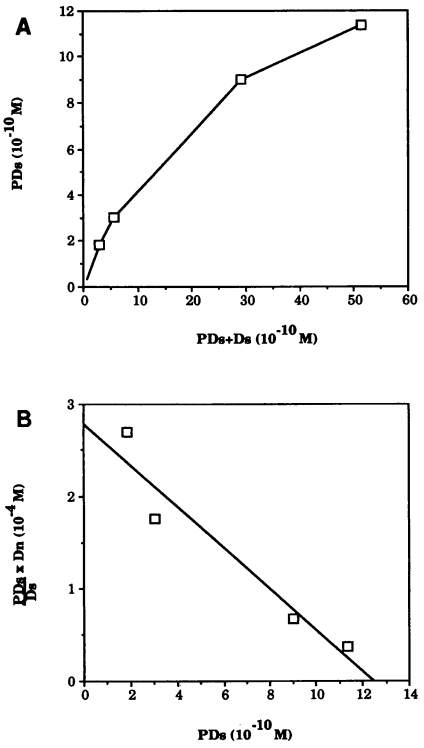

FIG. 5. Formation of complexes as a function of concentration of the 62S(−) probe. Formation of complexes was done as described in Materials and Methods, with increasing concentrations of ^{32}P-labeled 62S(−) probe (0.06, 0.3, 1.5, and 3 ng) incubated with 1,200 ng of mosquito cell extract. Samples were analyzed by gel retardation on a nondenaturing 5% polyacrylamide gel. The amount of radioactivity present in the free RNA and in the RNA-protein complexes migrating as band I was quantitated as described in Materials and Methods. (A) Plot of the concentration of specific complexes formed (PDs) versus the concentration of specific complexes (PDs) plus free RNA (Ds). (B) Transformation of the data by the method of Calzone et al. (3), where Dn is the concentration of nonspecific binding sites. The slope of the line is $-K_r$, and the intercept is P_oK_r, where P_o is the total concentration of the binding protein in the assay.

of the wild-type 62S(−) probe were constructed that lacked nucleotides 2 to 4, 5, 1 to 15, 10 to 15, 15 to 25, 26 to 40, or 41 to 55 or that contained G instead of A at position 5 (numbered 3′ to 5′ in the minus strand). These mutations were chosen in part because in vivo studies of mutants carrying these deletions have shown that some of them have

TABLE 1. Comparison of the P_o and K_r values for four Sindbis virus probes[a]

Probe name	K_r (10^6)	P_o (10^{-5})	r^2
62S(−)	0.22	1.26	0.92
59/132S(−)	0.39	1.83	0.99
129/196S(−)	0.04	0.98	0.99
194/249S(−)	0.35	1.44	0.91

[a] The data were plotted as in Fig. 5, and P_o and K_r values were extracted. P_o is the molar concentration of the binding proteins in the binding reaction, while K_r is the ratio of the affinity constant K_s for the specific binding reaction to the affinity constant K_n for the nonspecific binding reaction with a large excess of nonspecific polynucleotide [poly(dI-dC)] present during the reaction ($K_r = K_s/K_n$). r^2 is the correlation coefficient for the least-squares line.

TABLE 2. Comparison of the P_o and K_r values for probe 62S(−) and eight mutagenized derivatives of this probe[a]

Probe name	K_r (10^6)	P_o (10^{-5})	r^2
62S(−)	0.22	1.26	0.92
62S(−)d2-4	0.19	0.19	0.95
62S(−)d5	0.07	1.48	0.99
62S(−)G5	0.07	1.36	0.91
62S(−)d10-15	0.19	1.64	0.93
62S(−)d1-15	0.03	1.57	0.98
62S(−)d15-25	0.77	1.51	0.97
62S(−)d26-40	0.22	1.15	0.99
62S(−)d41-55	0.27	0.95	0.94

[a] P_o, K_r, and r^2 are as in Table 1.

dramatic effects on virus growth (16). The extent of complex formation was determined by titration of each probe, and estimates of K_r and P_o were determined as above. The values of these parameters as well as the corresponding correlation coefficients r^2 are presented in Table 2.

Probes 62S(−)d10-15, 62S(−)d26-40, and 62S(−)d41-55 all had K_r and P_o values approximately the same as that for the wild-type 62S(−). Thus, these deletions have only a slight effect on the binding of the mosquito proteins. Probes 62S(−)d5 and 62S(−)G5, on the other hand, were bound with a K_r only one-third that of the wild-type probe, although the P_o was essentially unchanged, indicating the importance of this nucleotide for binding. Probe 62S(−)d1-15 displayed a binding affinity only one-seventh that of the wild-type probe, and although above background (P_o is unchanged), this probe is very poorly bound by mosquito proteins.

The results with probe 62S(−)d2-4 were not consistent with results with the other probes. Its apparent K_r value (0.19×10^6) was similar to that of the 62S(−) probe, but its P_o value (0.19×10^5 M) was about seven times smaller. One possible explanation is that this mutant is bound by only a subset of the host proteins active with the wild-type probe. A second possibility is that the deletion affects the secondary structure, locking the structure in an unfavorable conformation so that the site is hidden or much less accessible to protein binding than is the wild-type structure. The on-rate for the binding reaction would then be much longer with the mutant probe such that the reaction does not reach equilibrium in the time used. A third possibility is that the K_r value obtained for the mutant probe results from nonspecific binding, since band I was barely detectable on the retardation gels, and that this deletion abolishes or greatly reduces the specific binding.

Finally, the deletion of nucleotides 15 to 25 gave the surprising result that the binding affinity increased more than threefold ($K_r = 0.77 \times 10^6$), with P_o unchanged or only slightly elevated. These results indicate that the first 10 residues are crucial for the binding of the 50- and 52-kDa proteins and that the residues between 15 and 25 down-regulate the binding of these proteins.

It is of note that (nonspecific) band III was not formed when the 62S(−)d26-40 probe was used in the binding reaction, but was formed with all other probes (data not shown). In cross-competition experiments, probes 62S(−), 62S(−)d15-25, 62S(−)d26-40, 62S(−)d41-55, 62S(−)d10-15, 62S(−)d5, and 62S(−)G5 behaved similarly, whereas probes 62S(−)d1-15 and 62S(−)d2-4 proved to be very poor competitors (data not shown). The reason why the low-affinity probes 62S(−)d5 and 62S(−)G5 are good competitors is unclear.

TABLE 3. Comparison of P_o and K_r for the probe containing the 3′ end of the minus strand of three different alphavirus RNAs or of rubella virus RNAs[a]

Probe name	K_r (10^6)	P_o (10^{-5})	r^2
62S(−)	0.22	1.26	0.92
RR(−)	0.87	1.04	0.99
SF(−)	0.23	0.91	0.94
RUB(−)	0.06	1.08	0.84

[a] P_o, K_r, and r^2 are as in Table 1.

Binding of 3′ ends of other alphavirus and rubella virus RNAs. Having demonstrated that the 3′ end of Sindbis virus minus-strand RNA contained binding sites for cellular proteins, we wished to examine the 3′ ends of minus-strand RNA of other alphaviruses and chose RR and SF alphavirus RNAs. We also tested the 3′ end of rubella virus minus-strand RNA, as this virus shares similarities with the alphaviruses and is believed to be related to them, although sequence comparisons fail to show extensive sequence similarity. We used RNA transcripts containing the complement of the wild-type 5′ NTR for the alphaviruses, probe 62S(−), RR(−), or SF(−), or transcripts containing the complement of the first 61 nucleotides of the rubella virus RNA which includes the 40-nucleotide 5′ NTR and the following 21 residues, probe RUB(−). We found that formation of complex (band I) occurred with each of these probes upon incubation with the mosquito cell extracts, although the RUB(−) probe formed significantly less complex than the alphavirus probes (data not shown). Nonspecific band III was observed only with the Sindbis virus probe. Cross-competition experiments showed that the RR(−), SF(−), and 62S(−) unlabeled probes competed similarly, whereas the RUB(−) probe competed less efficiently (data not shown). UV cross-linking experiments confirmed that the 50- and 52-kDa proteins were involved (data not shown). We estimated the values of the two parameters K_r and P_o in titration experiments as above, and these values are presented in Table 3. The SF(−) probe displayed the same affinity for the mosquito proteins ($K_r = 0.23 \times 10^6$) as did the Sindbis 62S(−) probe ($K_r = 0.22 \times 10^6$). The RR(−) probe displayed a fourfold-higher affinity (0.87×10^6), however, while the RUB(−) probe had a fourfold-lower affinity (0.06×10^6). The P_o value for each probe was found to be about the same as for the Sindbis virus probe and in line with previous estimates.

In an attempt to assess the importance of the 3′ nucleotides of the minus strand of the RR virus, SF virus, and rubella virus RNAs, we used derivatives of probes RR(−), SF(−), and RUB(−) that lacked the 20 first residues [RR(−)d1-20, SF(−)d1-20, and RUB(−)d1-20 probes, respectively] in binding assays with mosquito cell extracts. Surprisingly, we observed only a 10 to 20% decrease in the binding ability (data not shown). We have not further explored this topic, but it is possible that there is more than one binding site within the probes used, analogous to the situation with Sindbis virus.

DISCUSSION

Mosquito cell proteins bind to the 3′ end of Sindbis virus minus-sense RNA. We demonstrated that four different domains within the first 249 nucleotides at the 3′ end of the Sindbis virus minus-sense RNA interact with two mosquito proteins of 50 and 52 kDa. Three domains bound with high affinity ($K_r = 0.2 \times 10^6$ to 0.4×10^6) and one with low affinity ($K_r = 0.04 \times 10^6$). We consistently found that of the three high-affinity sites, the site between nucleotides 59 and 132 had the highest affinity and the site between nucleotides 1 to 62 had the lowest affinity. The probe containing the low-affinity site also contains a conserved sequence element, the 51-nucleotide domain that is capable of forming two stable stem-loop structures, which has been postulated to have a function in RNA synthesis distinct from that of the 3′-terminal 44 nucleotides that form the terminal stem-loop structure (22). It is unknown if the 51-nucleotide element functions in the genomic RNA or in the minus-strand RNA or both. If it functions in the minus strand by binding with high-affinity cellular proteins distinct from the 50- and 52-kDa proteins, then the fact that it is flanked by three high-affinity sites for the 50- and 52-kDa proteins indicates that the 3′ sequence binds several proteins that could interact with one another and prime RNA for initiation of plus-strand synthesis on the minus-strand template. Our results that longer probes can bind more than one protein molecule and that the bound proteins may interact to stabilize the interactions are consistent with this model.

Several DNA promoter and enhancer elements containing multiple binding sites have been reported, including the CpG island HTF9 (21) and the simian virus 40 early promoter (1). In the simian virus 40 early promoter, six adjacent sites operate independently to activate viral transcription (1). The CpG island HTF9 was also found to contain multiple protein-binding sites, but only a small subset of elements appeared to be required to activate the transcription of the HTF9 gene (21). We do not know whether the three high-affinity sites at the 3′ end of the Sindbis virus minus-sense genome must be occupied for efficient initiation of viral replication, but the redundant architecture of binding elements may play a role in the adaptation of a promoter to different cellular backgrounds, since Sindbis virus replicates in mosquitoes, birds, and mammals.

We previously reported that chicken cell extracts contained proteins of 42 and 44 kDa that bound to the 3′ end of the Sindbis virus minus-strand RNA with a K_r of 0.3×10^6 to 0.4×10^6, very similar to the K_r found for the mosquito proteins (18). In that study, we did not explore the possibility of multiple independent binding sites for these proteins, but we have found that there are at least two sites, one within the first 62 nucleotides and a second between nucleotides 59 and 132 (17). We presume that the chicken proteins and mosquito proteins are homologs, but isolation and characterization of these proteins will be required to establish this. The similarity in binding affinities and in the P_o values obtained for chicken and mosquito cell extracts, despite the obvious difference in size of the proteins, suggests that binding of these proteins does play an important role in virus replication. We do not know what proportion of the binding proteins are present in a ribosome-bound form. Recent results have shown that the ribosomal supernatant contains significant amounts of the binding proteins, but the exact amount has not been quantitated (17). The P_o values obtained for the high-salt wash of the ribosomal fraction represents about 5,000 molecules of binding protein per cell.

It is not known whether the 50- and 52-kDa mosquito cell proteins involved in the formation of complexes are two distinct proteins or two forms of a single polypeptide. Several other proteins that bind nucleic acids have been found to migrate as doublets that are suspected to derive from a single polypeptide (5, 8). It is noteworthy that the

chicken proteins that bind to the 3′ end of the Sindbis virus minus strand also migrate as a doublet of 42 and 44 kDa (18).

Mapping of residues important for protein binding. Our results with deleted probes demonstrate that the first 10 residues are crucial for the binding to the 3′-terminal site. Deletion of nucleotides 2 to 4, deletion or substitution of nucleotide 5, or deletion of the first 15 nucleotides resulted in greatly reduced binding. Deletion of nucleotides 10 to 15, 26 to 40, or 41 to 55 had no effect on binding, and deletion of nucleotides 15 to 25 resulted in a large increase in the affinity of binding. It is intriguing that deletion of nucleotides 2 to 4 or nucleotide 5 was found to be lethal for the virus (16), suggesting that the binding reported here represents an important step in virus replication. This reduced binding cannot be completely responsible for the mutant phenotype, however, because substitution of nucleotide 5 by G also resulted in reduced binding but was not lethal for the virus, although it did result in reduced growth, especially in mosquito cells (16). Conversely, deletion of nucleotides 10 to 14 led to a serious impairment of virus growth in mosquito cells, whereas binding to the mosquito 50- and 52-kDa proteins was not affected.

We previously reported that deletion of nucleotide 5 in the probe 132S(−)d5 resulted in tighter binding of the probe to chicken proteins leading to a significantly decreased off-rate ($t_{1/2}$ = 16 min for the mutant probe compared with 5.5 min for the wild-type probe) (18). In contrast, binding of 132S(−)d5 and 132S(−) to the 50- and 52-kDa mosquito proteins was indistinguishable, and both complexes dissociated with a $t_{1/2}$ of 5 min (17). We suggest that these differences in binding are due to interactions between binding sites 1 and 2 and that these interactions are different when chicken proteins are bound than when mosquito proteins are bound. This is consistent with the observation that with mosquito extracts, we readily observed formation of larger complexes formed presumably by binding to both sites 1 and 2, whereas with chicken extracts, we did not observe such larger complexes. Together with the results from other deletion probes and the results of Niesters and Strauss (16), this suggests that the linear sequence of the RNA is not solely responsible for the binding of cellular proteins but that the secondary structure in this region of the RNA is also important for binding. Certainly the finding that deletion of nucleotides 15 to 25 led to a greater than threefold increase in the K_r suggests that binding is moderated by secondary structure.

The increased binding observed upon deletion of nucleotides 15 to 25 is also of interest because deletion of these nucleotides in a defective interfering RNA was found to be lethal (23), whereas the deletion in viral RNA results in viable virus (16). This suggests that in a nondefective genome the deletion may be partially compensated for by interactions with proteins bound to downstream sequences, allowing the survival of the virus, although with an impaired growth.

Secondary structures are important for the replication of RNA viruses from different families, including picornaviruses, retroviruses, and alphaviruses. In some cases, mutational analysis of these regions has highlighted the importance of secondary structures in the binding of cellular factors. Mutations in the proximal stem-loop structure of the ribosome landing pad of poliovirus type 2 abolished internal initiation, as did deletions predicted to disrupt stem-loop structures V and VI, whereas deletion of the second stem-loop reduced internal initiation by 50% (14). Interestingly, cellular factors have been shown to bind to the proximal stem-loop structure of poliovirus (4). Mutational analysis has also demonstrated that two cellular factors could bind two separate sites in the upper portion of the transactivation-responsive RNA hairpin of human immunodeficiency virus type 1 (20). In SF alphavirus, it has been demonstrated that the efficiency of recognition of the mRNA by initiation factors was determined by a secondary structure close to the 5′ cap structure (2). Increasing the stability of the secondary structure in the 5′ NTR by treating the RNA with increasing amounts of K^+ led to the decrease of binding of eIF-4B and eIF-4E to both the 42S and 26S RNAs. Interestingly, the decrease was more pronounced for the 42S RNA than for the 26S RNA, confirming the computer prediction that the 5′ NTR of 26S RNA contained less secondary structure than that of the 42S RNA and was therefore more efficiently recognized by the cellular translation machinery.

Other alphavirus minus-sense RNAs bind the same cellular proteins. Our results demonstrate that 50- and 52-kDa proteins from mosquitoes bind to the 3′ end of Sindbis virus minus-strand RNA and that these same two proteins also bind the 3′ end of minus-strand RNA of two other alphaviruses, SF and RR. The K_r values are comparable, although the RR virus probe had a fourfold-higher K_r than did the Sindbis virus and SF virus probes. We have not explored in detail the number and positions of binding sites for these proteins in RR virus and SF virus, but assume that like Sindbis virus multiple binding sites are present. We found that a probe derived from rubella virus RNA also bound the same 50- and 52-kDa proteins, but poorly. Rubella virus is classified as a member of the *Togaviridae* family (genus *Rubivirus*). There are similarities in the structure of the virus and the viral genome and limited sequence identities in the RNAs (6), and it is believed that these two groups have descended from a common ancestor. Nakhasi and colleagues (13) have described the specific binding of two Vero cell proteins to the stem-loop structure at the 3′ end of rubella virus and of Sindbis virus minus-sense RNAs. Preliminary results from binding experiments with either the 62S(−) or the RUB(−) probe suggest that they also share a similar binding ability for two HeLa proteins (17). We hypothesize that the binding of these cellular proteins to these viral RNAs is required for replication of the viral RNA and that the poor binding by mosquito proteins of rubella virus RNA results from adaptation of rubella virus to mammalian cells. Mosquitoes are not a host for rubella virus, but there may remain a residual binding of mosquito cell proteins to the rubella virus RNA elements because rubella virus and alphaviruses continue to use common proteins in mammalian cells for RNA replication.

ACKNOWLEDGMENTS

We are grateful to E. Davidson for helpful discussions about the analysis of the RNA binding profiles, to J. Moore for his considerable efforts in computer analysis of the binding constants, to M. Harrington for assistance and advice on quantitation of the gels, and to E. Strauss for critical reading of the manuscript.

This work was supported by grant DMB-9104054 from the National Science Foundation.

REFERENCES

1. **Barrera-Saldana, H., K. Takahashi, M. Vigneron, A. Wildeman, I. Davidson, and P. Chambon.** 1985. All six GC-motifs of the SV40 early upstream element contribute to promoter activity in vivo and in vitro. EMBO J. **4**:3839–3849.
2. **Berben-Bloemheuvel, G., M. A. M. Kasperaitis, H. van Heugten, A. A. M. Thomas, H. van Steeg, and H. O. Voorma.** 1992. Interaction of initiation factors with the cap structure of chimae-

ric mRNA containing the 5'-untranslated regions of Semliki Forest virus RNA is related to translational efficiency. Eur. J. Biochem. **208**:581–587.

3. **Calzone, F. J., N. Theze, P. Thiebaud, R. L. Hill, R. J. Britten, and E. D. Davidson.** 1988. Developmental appearance of factors that bind specifically to *cis*-regulatory sequences of a gene expressed in the sea urchin embryo. Genes Dev. **2**:1074–1088.
4. **Del Angel, R. M., A. G. Papavassilliou, C. Fernandez-Tomas, S. J. Silverstein, and V. R. Racaniello.** 1989. Cell proteins bind to multiple sites within the 5' untranslated region of poliovirus RNA. Proc. Natl. Acad. Sci. USA **86**:8299–8303.
5. **Della Seta, F., I. Treich, J. M. Buhler, and A. Sentenac.** 1990. ABF1 binding sites in yeast RNA polymerase genes. J. Biol. Chem. **265**:15168–15175.
6. **Dominguez, G., C.-Y. Wang, and T. K. Frey.** 1990. Sequence of the genome RNA of rubella virus: evidence for genetic rearrangement during togavirus evolution. Virology **177**:225–238.
7. **Huismans, H., A. A. van Dijk, and A. R. Bauskin.** 1987. In vitro phosphorylation and purification of a nonstructural protein of bluetongue virus with affinity for single-stranded RNA. J. Virol. **61**:3589–3595.
8. **Jang, S. K., and E. Wimmer.** 1990. Cap-independent translation of encephalomyocarditis virus RNA: structural elements of the internal ribosomal entry site and involvement of a cellular 57 kDa RNA-binding protein. Genes Dev. **4**:1560–1572.
9. **Kuhn, R. J., Z. Hong, and J. H. Strauss.** 1990. Mutagenesis of the 3' nontranslated region of Sindbis virus RNA. J. Virol. **64**:1465–1476.
10. **Kuhn, R. J., H. G. M. Niesters, Z. Hong, and J. H. Strauss.** 1991. Infectious RNA transcripts from Ross River virus cDNA clones and the construction and characterization of defined chimeras with Sindbis virus. Virology **182**:430–441.
11. **Levis, R., S. Schlesinger, and H. V. Huang.** 1990. The promoter for Sindbis virus RNA-dependent subgenomic RNA transcription. J. Virol. **64**:1726–1733.
12. **Meerovitch, K., J. Pelletier, and N. Sonenberg.** 1989. A cellular protein that binds to the 5' noncoding region of poliovirus RNA: implications for internal translation initiation. Genes Dev. **3**:1026–1034.
13. **Nakhasi, H. L., X.-Q. Cao, T. A. Rouault, and T.-Y. Liu.** 1991. Specific binding of host cell proteins to the 3'-terminal stem-loop structure of rubella virus negative-strand RNA. J. Virol. **65**:5961–5967.
14. **Nicholson, R., J. Pelletier, S.-Y. Le, and N. Sonenberg.** 1991. Structural and functional analysis of the ribosome landing pad of poliovirus type 2: in vivo translation studies. J. Virol. **65**:5886–5894.
15. **Niesters, H. G. M., and J. H. Strauss.** 1990. Defined mutations in the 5' nontranslated sequence of Sindbis virus RNA. J. Virol. **64**:4162–4168.
16. **Niesters, H. G. M., and J. H. Strauss.** 1990. Mutagenesis of the conserved 51-nucleotide region of Sindbis virus. J. Virol. **64**: 1639–1647.
17. **Pardigon, N.** Unpublished data.
18. **Pardigon, N., and J. H. Strauss.** 1992. Cellular proteins bind to the 3' end of Sindbis virus minus-strand RNA. J. Virol. **66**:1007–1015.
19. **Raju, R., and H. V. Huang.** 1991. Analysis of Sindbis virus promoter recognition in vivo, using novel vectors with two subgenomic mRNA promoters. J. Virol. **65**:2501–2510.
20. **Sheline, C. T., L. H. Milocco, and K. A. Jones.** 1991. Two distinct nuclear transcription factors recognize loop and bulge residues of the HIV-1 TAR RNA hairpin. Genes Dev. **5**:2508–2520.
21. **Somma, M. P., C. Pisano, and P. Lavia.** 1991. The housekeeping promoter from the mouse CpG island HTF9 contains multiple protein-binding elements that are functionally redundant. Nucleic Acids Res. **19**:2817–2824.
22. **Strauss, E. G., and J. H. Strauss.** 1986. Structure and replication of the alphavirus genome, p. 35–90. *In* S. Schlesinger and M. J. Schlesinger (ed.), The togaviridae and flaviviridae. Plenum Publishing Corp., New York.
23. **Tsiang, M., B. G. Weiss, and S. Schlesinger.** 1988. Effects of 5'-terminal modifications on the biological activity of defective interfering RNAs of Sindbis virus. J. Virol. **62**:47–53.