Molecular Clones Representing Sendai Virus Genes P, NP and M

By KAILASH C. GUPTA, EXEEN M. MORGAN, GEOFFREY KITCHINGMAN AND D. W. KINGSBURY*

Division of Virology and Molecular Biology, St. Jude Children's Research Hospital, 332 N. Lauderdale, Memphis, Tennessee 38101, U.S.A.

(Accepted 21 April 1983)

SUMMARY

DNA copies of segments of Sendai virus genes P, NP and M, obtained by reverse transcription of virus mRNA species extracted from infected cells, were cloned in plasmid pBR322. Genes were identified by hybrid-arrested translation of viral mRNAs in vitro. Hybrid selection of NP mRNA confirmed the identity of an NP gene clone. Partial sequencing of this insert showed that it represents the 5'-terminal region of the gene, containing transcription termination signals. Hybridization of DNA inserts to blots of electrophoretically separated, denatured mRNA species indicated that the P, NP and M messages had sizes of 2400, 2100 and 1500 nucleotides, respectively. Specific T1 ribonuclease-resistant oligonucleotides, previously identified in the NP and M genes, were selected by hybridization to the respective inserts. Although most of the inserts are smaller than 500 base pairs, representing no more than 16% of the P gene and 24% of the NP gene, one of the M gene inserts, comprising 700 base pairs, represents almost half of that gene. These cloned virus gene segments will assist further investigations of the molecular biology of this model paramyxovirus.

INTRODUCTION

Although the negative-strand strategy of replication was first described for a paramyxovirus (Kingsbury, 1966; Bratt & Robinson, 1967), members of this virus family have not been studied as extensively at the molecular level as the rhabdovirus, vesicular stomatitis virus (VSV), because the latter is much easier to produce in quantities suitable for biochemical work (Wagner, 1975). However, recombinant DNA technology has made it possible to obtain virtually unlimited quantities of paramyxovirus genetic information for detailed analysis. So far, the only reported application of this technology to a paramyxovirus concerns measles virus (Gorecki & Rozenblatt, 1980; Rozenblatt et al., 1982). In the present paper, we describe the molecular cloning of segments of three genes from another paramyxovirus, the murine pathogen Sendai virus. We report the application of these cloned sequences to the identification of viral mRNA species and gene-specific oligonucleotides, and we provide the sequence of 65 nucleotides at the 5' terminus of the NP gene, derived from one of the molecular clones.

METHODS

Synthesis of cDNA. At 48 to 72 h after infection, RNA was extracted from chick embryo lung cells and poly(A)-containing mRNAs were isolated by oligo(dT)-cellulose chromatography (Amesse & Kingsbury, 1982). The yield of poly(A)-rich RNA from 100 cell cultures in 100 mm Petri dishes was 30 to 40 µg. This material was used as the template for oligo(dT)12-18-prime DNA synthesis by avian myeloblastosis virus (AMV) reverse transcriptase (Miller et al., 1980; Norgard et al., 1980), in 1-h reactions at 40 °C. Reverse transcriptase was provided by the Division of Cancer Cause and Prevention, National Cancer Institute, Bethesda, Md., U.S.A. The single-stranded cDNA product was isolated and used to self-prime second strand synthesis with fresh AMV reverse transcriptase (Miller et al., 1980; Norgard et al., 1980). About 95% of the product of the second reaction was resistant to S1 nuclease. About 10 µg of double-stranded DNA was obtained, starting with 50 to 75 µg of mRNA template.

Molecular cloning. Double-stranded DNA molecules were trimmed with S1 nuclease, tailed with oligo(dC) (average chain length, 20 nucleotides) by the use of terminal transferase (Roychoudhury & Wu, 1980), and annealed to oligo(dG)10-tailed pBR322 DNA that had been cut by restriction endonuclease PstI (Miller et al.,
Previous work suggested that Sendai virus mRNA species might comprise as much as 10% of the total intracellular mRNA late in infection of CEL cells (Davies et al., 1976). Thus, we anticipated obtaining a high proportion of bacterial clones containing virus genetic information if we used total oligo(dT)-selected RNA from infected cells as the template for cDNA synthesis. Indeed, out of 110 clones that were presumed to contain plasmids bearing cDNA inserts on the basis of antibiotic sensitivities, 12 gave strong positive signals by colony hybridization with a $^{125}$I-labelled virion RNA probe (Commerford, 1971; Grunstein & Hogness, 1975).

**Hybrid-arrested translation.** Virus-specific DNA inserts were separated from plasmid vector DNA by PstI digestion and electrophoresis in 5% polyacrylamide gels (Maxam & Gilbert, 1980). About 5 μg of each insert was annealed with 0-4 μg of oligo(dT)-selected mRNA from virus-infected cells for 2 h at 45 °C in 80% formamide, 0-002 M-EDTA, 0-4 M-NaCl, 0-1 M-PIPES pH 6-4 (Paterson et al., 1977; Lamb et al., 1980). Then, 25 μg of wheat germ tRNA was added to each sample as carrier, the nucleic acids were precipitated with ethanol and the precipitates were dissolved in water. Half of each sample was heat-denatured for 1 min at 100 °C as a control and both halves were incubated in a rabbit reticulocyte lysate (Bethesda Research Laboratories) at 30 °C for 1 h with [35S]methionine as the labelled precursor. The polypeptide products were separated by electrophoresis in 9% discontinuous SDS–polyacrylamide gels (Davies et al., 1976).

**RNA blotting.** Oligo(dT)-selected mRNA from infected cells was denatured with glyoxal, separated by electrophoresis in 1.25% agarose gels, and transferred to nitrocellulose paper (Thomas, 1980). Cloned DNA was labelled with [32P]deoxyribonucleotides by nick translation (Rigby et al., 1977) and its hybridization to RNA on nitrocellulose was detected by autoradiography.

**Selection of T1-oligonucleotides.** Oligonucleotides obtained by complete digestion of Sendai virus genomic RNA were labelled at their 5' termini by polynucleotide kinase-catalysed transfer of [32P]phosphate from gamma-[32P]ATP (Maxam & Gilbert, 1980) and added to 7 mm nitrocellulose filter discs containing 10 μg of denatured, PstI-cleaved plasmid DNA (Kafatos et al., 1979). Samples were annealed for 2 to 4 h at 37 °C in 0.4 M-NaCl, 0.25% SDS, 0.002 M-EDTA, 0.05 M-Tris–HCl pH 7.5. Filters were washed extensively and bound oligonucleotides were eluted into water and identified by two-dimensional gel electrophoresis (Amesse & Kingsbury, 1982).

**Hybrid selection of mRNA.** Filters containing plasmid DNA were prepared as described above and prehybridized with 200 μg of mixed E. coli 3RNAs (Sigma) per ml of 0.4 M-NaCl, 0.002 M-EDTA, 0.2% SDS, 0.01 M-PIPES pH 6.4, 70% (v/v) formamide at 50 °C for 1 h. The liquid was replaced by fresh buffer lacking tRNA, but containing 600 μg of unfractionated RNA from Sendai virus-infected chick embryo lung (CEL) cells per ml. After annealing at 30 °C for 4 h, filters were washed four times at 45 °C with 0.4 M-NaCl, 0.002 M-EDTA, 0.2% SDS, 0.05 M-Tris–HCl pH 7.5. mRNA was eluted by boiling in H2O and collected by ethanol precipitation in the presence of 10 μg tRNA.

**RESULTS**

**Identification of cloned gene segments**

Examples of the hybrid-arrested translation results that identified five of the six clones are shown in Fig. 1. We could identify inserts representing the genes for RNA polymerase-associated protein P (clones pP1 and pP2), nucleocapsid structure unit NP (clones pNP1 and pNP2) and non-glycosylated envelope protein M (clone pM1), because these genes specify proteins that are produced abundantly when a cell-free translation system is programmed by a mixture of virus mRNA species from infected cells. (The prominent protein band appearing between NP and M is an endogenous product of the reticulocyte extract; see below.) Surprisingly, clone pM2 was inactive in hybrid-arrested translation (data not shown); however, it was identified as M-gene specific by other means (Fig. 4 and 6). Of the remaining virus gene products, L has not yet been identified in a cell-free translation system, and the two glycoproteins, HN and F₀, are translated **in vitro** with very low efficiency (Davies et al., 1976). Therefore, we were unable to identify inserts representing these genes by hybrid-arrest.

The identity of the pNP2 insert was confirmed by hybrid selection using nitrocellulose filters containing the plasmid. NP was the only viral protein synthesized by reticulocyte extracts
Fig. 1. Polyacrylamide gel electrophoresis of the products of cell-free translation of Sendai virus mRNAs after hybridization with cloned inserts. In parts 1, 2 and 3, lane (a) is the experimental sample, lane (b) is an identical sample, heat-denatured before translation, as a control. 1, pP2; 2, pNP2; 3, pM1. Letters designate Sendai virus proteins. Arrowheads designate bands of reduced intensity indicating hybrid-arrest. *, a new band appearing after arrest of NP translation by pNP2 DNA.

Fig. 2. Cell-free translation after mRNA selection by hybridization to plasmid pNP2. Autoradiograms are shown of [35S]methionine-labelled proteins separated by polyacrylamide gel electrophoresis. Lane 1, proteins from purified virions. Lanes 2 to 4, cell-free translation products: 2, no added mRNA; 3, mRNA selected by clone pNP2 was added; 4, unfractionated mixture of mRNA species from infected cells was added. X is a cellular gene product. The remaining letters designate virus proteins. Arrowhead: NP protein.

Table 1. Sizes of Sendai virus gene inserts in pBR322

<table>
<thead>
<tr>
<th>Plasmid*</th>
<th>Base pairs†</th>
<th>Per cent of gene‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>pP1</td>
<td>135</td>
<td>5-6</td>
</tr>
<tr>
<td>pP2</td>
<td>375</td>
<td>16</td>
</tr>
<tr>
<td>pNP1</td>
<td>440</td>
<td>21</td>
</tr>
<tr>
<td>pNP2</td>
<td>495</td>
<td>24</td>
</tr>
<tr>
<td>pM1</td>
<td>295</td>
<td>20</td>
</tr>
<tr>
<td>pM2</td>
<td>700</td>
<td>47</td>
</tr>
</tbody>
</table>

* The capital letters designate the virus gene represented by the insert. Gene assignments were made by hybrid-arrested translation (Fig. 1), except for pM2 (by hybridization to mRNA, see text).

† Approximate number of base pairs in the insert determined by PstI digestion of the plasmid and electrophoresis (Fig. 3).

‡ Approximate numbers of bases in the genes are: P, 2400; NP, 2100; M, 1500. These figures are based on the electrophoretic migration of the respective mRNA species (Fig. 4) relative to ribosomal RNA species in denaturing conditions (Gupta & Kingsbury, 1982).
Fig. 3. Electrophoresis of Sendai virus gene inserts. Plasmids were cleaved with PstI and subjected to electrophoresis in a 5% polyacrylamide gel (see Methods). The figure shows a photograph of a gel stained with ethidium bromide. Lanes: 1, pNP1; 2, pNP2; 3, pP1; 4, pP2; 5, pM1; 6, pM2; 7, pBR322; 8, pBR322 cleaved with Aval. Numbers on the right are sizes, in base pairs, of the pBR322 fragments. The arrowhead marks the pP1 insert.

Fig. 4. Hybridization of DNA inserts to Sendai virus mRNA species separated by electrophoresis. Shown is an autoradiogram of 32P-labelled nick-translated plasmid DNA hybridized to mRNA on nitrocellulose paper. Lanes: C, viral mRNAs labelled in vivo with [32P]orthophosphate; P, pP3; NP, pNP2; M, pM2.
Sendai virus clones

receiving the selected mRNA (Fig. 2). The band labelled X in Fig. 2 (also seen in Fig. 1) was
synthesized by extracts that did not receive viral mRNA. We did not succeed in attempts to
select M or P mRNA species, probably because these species were not sufficiently abundant in
the RNA mixture extracted from infected cells.

The electrophoretic migration of the six virus-specific inserts cut from the pBR322 vector by
PstI is shown in Fig. 3. Comparisons of the deduced molecular weights of these inserts with our
best estimates of the molecular weights of the genes that they represent revealed that each of the
NP gene clones was about one-fourth the size of that gene, that 5% and 16% of the P gene was
present in cloned inserts, and that the two M gene inserts contained 20% and 47% of that gene
(Table 1).

The appearance of a new protein species migrating between NP and M after hybrid-arrest by
the pNP2 insert (Fig. 1, lane 2a) may represent a prematurely terminated NP polypeptide, due
to hybridization of the insert to the 3' end of the mRNA (Lamb
et al., 1980). We have not
examined this new band for homology with the NP protein, but, as discussed below, the pattern
of genomic T1 RNase-resistant oligonucleotides hybridizing to pNP2 and the nucleotide
sequence of one end of the insert revealed that it represents the 3' end of the mRNA.

Corresponding mRNAs

Glyoxal-denatured Sendai virus mRNAs migrated as three components on electrophoresis in
an agarose gel (Gupta & Kingsbury, 1982). The mRNA specifying the M protein had the most
rapid migration under fully denaturing and partially denaturing conditions (Gupta &
Kingsbury, 1982; Jones
et al., 1978), and was easily distinguished from the rest. After glyoxal
treatment, it migrated as a polynucleotide of 1500 bases, relative to ribosomal RNA (Amesse
et al., 1982). However, the messages for the other four abundantly expressed Sendai virus proteins,
P, NP, F0 and HN, segregated into two bands, labelled 1 and 2, with apparent sizes of 2400 and
2100 bases, respectively (Fig. 4). By hybridization of nick-translated plasmid DNAs to nitro-
cellulose blots of these electrophoretically separated mRNAs (Thomas, 1980), we found that the
mRNA for protein P migrated in band 1, and that for NP in band 2 (Fig. 4). Hybridization to the
M band mRNA confirmed the identity of clone pM1 (data not shown) and showed that clone
pM2 was M gene-specific. The positions of the P and NP messages in the electropherogram
were congruent with the molecular weights of the protein products (Glazier
et al., 1977). The
locations of the F0 and HN mRNAs remain undetermined, but we expect that they reside in
bands 2 and 1, respectively (Gupta & Kingsbury, 1982).

Genome oligonucleotides selected by inserts

We have previously used RNase T1-oligonucleotide maps to identify unique sequences in
Sendai virus genes (Amesse & Kingsbury, 1982). Although non-overlapping oligonucleotide
maps were obtained, the gene assignments were tentative, as they were based primarily on
molecular weight correlations with viral proteins. It was therefore of interest to learn which
oligonucleotides would be selected by the P, NP and M inserts from a RNase T1 digest of the
Sendai virus genome. For comparison, Fig. 5 is a more extensive oligonucleotide map of the
virus genome than those previously published (Amesse & Kingsbury, 1982); it displays more of
the lower molecular weight oligonucleotide spots at the upper part of the two-dimensional
polyacrylamide gel.

As shown in Fig. 6, distinct populations of T1 oligonucleotides hybridized to DNA from two
clones, representing the NP and M genes. In the NP pattern, only spot 2 had been identified
previously (Amesse & Kingsbury, 1982); the other three large oligonucleotides that were
recovered from the NP gene, numbers 10, 20, and 21, were not selected by pNP2 DNA.
Selection of oligonucleotide number 2 by pNP2 is consistent with our evidence that the plasmid
contains the 5' end of the NP gene, (Fig. 1, 7 and 8), since oligonucleotide 2 mapped in the 5'
region of the NP gene in analyses of Sendai virus defective interfering RNA species (Amesse
et al., 1982). In the case of the M gene, oligonucleotide 33 was the only characteristic spot
identified previously (Amesse & Kingsbury, 1982); this spot was prominent among the group of
oligonucleotides selected by hybridization with pM2. In each case, an oligonucleotide spot was
Fig. 5. Two-dimensional electrophoresis of RNase T1-resistant oligonucleotides from the Sendai virus genome. After complete digestion with RNase T1, oligonucleotides were labelled at their 5' termini (Maxam & Gilbert, 1980) before separation in polyacrylamide gels. The origin is at the lower left and the first dimension is horizontal. Oligonucleotides 1 to 45 were designated previously (Amesse & Kingsbury, 1982).

seen that we could not locate in the map of the Sendai virus genome (arrows, Fig. 6). These species may have arisen by artefactual cleavage by a contaminating nuclease during the hybridization reaction or they may represent sequences from genome-length positive strands present in the virion RNA preparation (Kolakofsky & Bruschi, 1975). In both maps, a spot appeared at position 56 (Fig. 6). This probably reflects the existence of two distinct oligonucleotides with similar mobilities, since spot 56 occupied a larger than average area in the map of the genome (Fig. 5).

Although each of the P gene DNA inserts selected a unique set of oligonucleotides, most of these were of low molecular weight (data not shown), and none could be clearly identified in any of the previously reported gene-specific oligonucleotide patterns (Amesse & Kingsbury, 1982).

Partial sequence of the pNP2 insert

We used the Maxam & Gilbert (1980) method to sequence both ends of the pNP2 insert. The insert was excised from the plasmid with PstI and terminally labelled with [α-32P]dCTP by means of bacteriophage T4 DNA polymerase (Maniatis et al., 1982). After heating at 70 °C for 5 min to inactivate the polymerase, the insert was digested with HinfI and two labelled fragments were separated by electrophoresis in a 6% polyacrylamide gel.

The sequence of one fragment terminated in poly(A), followed by (5')TAAG (Fig. 7), the DNA complement of the tetranucleotide (3')AUUC that we had previously identified as the...
Fig. 6. Oligonucleotides from the Sendai virus genome selected by hybridization to plasmid DNA molecules containing inserts. Two-dimensional electrophoresis in polyacrylamide gels was performed as described for Fig. 5. The DNA species used were pNP2 and pM2; top and bottom sections, respectively. The arrows designate oligonucleotides not seen in the genomic RNA map of Fig. 5 (see text).

Fig. 7. Sequence of one terminus of the pNP2 insert determined by the Maxam & Gilbert (1980) method. Letters above the lanes designate the nucleotide specificities of the chemical cleavage reactions.

terminating signal of the five most abundantly transcribed Sendai virus genes (Gupta & Kingsbury, 1982). This shows that the pNP2 insert represents the 5' end of the NP gene.

In Fig. 8, we give the sequence of 65 nucleotides at the NP gene terminus, deduced from further sequencing of the complementary strand in pNP2. At the 5' side of the (3')AUUC termination signal, the pentanucleotide (3')UUUUU had been found in all of the most abundantly transcribed genes; this is evidently the site for initiation of poly(A) addition to the 3' terminus of a Sendai virus mRNA (Gupta & Kingsbury, 1982). As a reference point, we have designated the 3'-terminal uracil residue of this oligo(U) tract as position 0 (Fig. 8).

We had obtained the five gene-terminal sequences by analyses of T1 RNase-resistant oligonucleotides selected from digests of virus genomic RNA. One of these oligonucleotides, number 48 (Gupta & Kingsbury, 1982), corresponds most closely to the 5' terminus of the sequence shown in Fig. 8. However, we had previously identified a uracil in position 8 of the sequence; we now see that the correct assignment is cytosine.
Fig. 8. Deduced Y-terminal sequence of the Sendai virus NP gene, based on the results in Fig. 7 and similar (not shown) data. The 5'-terminal sequence of the VSV N gene (Gallione et al., 1981) is shown for comparison. The known termination anticodon of the N gene and possible termination anticodons of the NP gene are underlined. Dots signify an adenine- and uracil-rich sequence in the N gene noted by Rose (1980). The gene-terminal consensus tetranucleotides and oligo(U) tracts have been highlighted by spaces.

Also shown in Fig. 8 is the 5' terminus of the N gene of the rhabdovirus VSV. The rhabdoviruses have a replication strategy very similar to that of paramyxoviruses, and the N gene is the rhabdovirus analogue of NP. A comparison of the 5'-terminal regions of these genes is of interest, since similar sequences might represent signals that regulate transcription. For example, we have already described the striking similarities between the poly(A) initiation sequences and adjacent nucleotides in the two virus genomes (Gupta & Kingsbury, 1982). In the case of VSV, the poly(A) initiation site comprises seven uracil residues, compared to the five of Sendai virus. Immediately upstream in the VSV genome is the consensus tetranucleotide (3')AUAC that differs from the Sendai virus (3')AUUC by a single base.

In contrast, alignment of the next 61 nucleotides in the two genes reveals no other obvious similarities. In particular, the adenine- and uracil-rich sequence at positions 13 to 22 of the VSV N gene, which was considered a possible promoter for initiation of the adjacent NS gene transcript (Rose, 1980), has not even a fragmentary counterpart anywhere in the Sendai virus sequence. The terminating anticodon of the N gene is the ACU at positions 45 to 47 (Gallione et al., 1981). None of the four candidate termination anticodons of the Sendai virus NP gene (underlined in Fig. 8) occupies the same position, but there are two close by, at positions 39 to 41 and 50 to 52. More extensive sequencing will be necessary to define the correct reading frame of the NP sequence and learn which, if any, of these triplets terminates the protein-coding region of the gene. In any case, the existence of similarities only in the oligo(U) tract and adjacent tetranucleotide suggests that there are no other transcriptional control signals at the 5' end of either gene.

DISCUSSION

None of the cloned Sendai virus gene segments identified here represents a complete gene, but they have enabled us to locate P and NP mRNAs among the species separated by electrophoresis and to confirm our previous assignments of genomic RNase T1-oligonucleotides to the NP and M genes (Amesse & Kingsbury, 1982). As we continue our efforts to obtain inserts containing the complete sequences of the P, NP, and M genes, to be employed in advanced genetic analyses and nucleotide sequencing, the available clones will be useful probes.

That pNP2 should contain the 5' end of the NP gene is not surprising, since all of the inserts were made by reverse transcription using an oligo(dT) primer, assuring that the cDNA products were initiated at the 3' ends of the mRNA templates. However, if second strand DNA synthesis terminates prematurely, such sequences will not be recovered in the clones. Knowing which sequence is represented in pNP2 opens the way to using the insert or a 5'-terminal restriction fragment as a primer for reverse transcription of viral genomic RNA, enabling us to determine the sequence of the 3' end of the adjacent gene, which may be F0 or P (Glazier et al., 1976;
Collins et al., 1980) and to clone that region of the virus genome. We expect to find equally useful 5'-terminal gene segments in at least some of the remaining inserts that we have studied.

The other six clones in our collection may represent one or more of the remaining virus genes, since all of them hybridized with virion RNA, but not with any of the six clones listed in Table 1. Furthermore, none of them was active in hybrid-arrested translation and the reticulocyte lysate was useful only with proteins P, NP and M (Fig. 2). (Combinatorial calculations indicate a 0.44 probability of obtaining all six of the major virus genes and a 0.89 probability of obtaining five of them from twelve randomly picked clones.) The addition of microsomal membranes to the translation system may yield more abundant quantities of glycoproteins HN and F0, allowing us to identify inserts representing these genes. On the other hand, the L mRNA is produced in relatively small amounts and the complete translation of such a large protein in vitro is technically difficult, so other means will be necessary to identify L gene clones. This might be accomplished by hybridization with Sendai virus defective interfering RNA species, all of which contain 5'-terminal portions of the L gene (Amesse et al., 1982).

In the reported characterization of measles virus cDNA clones, inserts were identified representing the G (major glycoprotein), N (major nucleocapsid protein) and M (non-glycosylated membrane protein) genes (Gorecki & Rozenblatt, 1980; Rozenblatt et al., 1982). The apparent molecular weights of the corresponding mRNAs were closely grouped, with estimates of 1750 nucleotides for the G and N messages and 1550 nucleotides for the M message. This contrasts with the greater separation that we observed between Sendai virus NP and M messages (Fig. 3) and is surprising, since NP and N have similar molecular weights, as do the M proteins of the two viruses (Glazier et al., 1977; Gorecki & Rozenblatt, 1980; Rozenblatt et al., 1982). Therefore, it will be interesting to learn if the genes of these viruses contain markedly different amounts of nucleotide sequences that are not involved in protein coding.

Elizabeth Ann Lyne and Karen Rakestraw provided skillful technical help. We are grateful to Robert A. Lamb for advice on hybrid-arrested translation. This work was supported by research grant RG-1142 from the National Multiple Sclerosis Society, by U.S. Public Health Service research grants AI 05343 and AI 17654 from the National Institute of Allergy and Infectious Diseases, by Cancer Center Support Grant CA 21765 from the National Cancer Institute, and by American Lebanese Syrian Associated Charities.

REFERENCES


Virology 118, 8-16.


Journal of General Virology 33, 117-123.


Proceedings of the National Academy of Sciences, U.S.A. 72, 3961-3965.


Virology 120, 518-523.


Nucleic Acids Research 7, 1541-1552.


Journal of Molecular Biology 18, 204-214.


*(Received 18 October 1982)*