Molecular Cloning and Sequence Analysis of the Mumps Virus Gene Encoding the P Protein: Mumps Virus P Gene Is Monocistronic

By KAORU TAKEUCHI,* MICHIKO HISHIYAMA, AKIO YAMADA AND AKIRA SUGIURA

SUMMARY
The nucleotide sequence of the P (phosphoprotein) gene of two strains of mumps virus has been determined from overlapping cDNA clones. The P gene contained a single open reading frame coding for a protein of 391 amino acids with a calculated Mr of 41587, in good agreement with the value (40K to 45K) estimated from electrophoretic mobility on SDS-polyacrylamide gels. No open reading frame analogous to the C gene of other paramyxoviruses existed in the mumps virus P gene region. Comparison of the amino acid sequence of the mumps virus P protein with that of Newcastle disease virus showed a limited sequence homology.

INTRODUCTION
Mumps virus, a member of the paramyxovirus family, causes parotitis and occasionally acute aseptic meningitis, encephalomyelitis and orchitis (Wolinsky & Server, 1985). Mumps virus contains six structural proteins encoded by a single-stranded negative-sense RNA. The nucleocapsid protein (NP) and two nucleoprotein-associated proteins (P and L) are believed to be involved in RNA transcriptase activity and, in association with viral RNA, to form the nucleocapsid. A non-glycosylated membrane protein (M) and two glycosylated proteins (HN and F), responsible for haemagglutination and cell fusion, respectively, are constituents of the viral membrane (Wolinsky & Server, 1985). In addition to these structural proteins it has been reported that mumps virus synthesizes non-structural protein(s) in infected cells. The previous studies are not fully in accord, however, as to whether the non-structural proteins are related to the P protein or not (Herder & Compans, 1982; Merz et al., 1983; Simpson et al., 1984). It is known that many paramyxoviruses such as Sendai virus (Giorgi et al., 1983; Shioda et al., 1983), human parainfluenza virus 3 (PF-3) (Galinski et al., 1986; Luk et al., 1986; Spriggs & Collins, 1986), measles virus (Bellini et al., 1985) and canine distemper virus (Barrett et al., 1985) utilize different reading frames present in the P protein mRNA to code for the non-structural C proteins. By sequencing cloned cDNA of the P gene region, we attempted to determine whether a bicistronic coding strategy is also used in the mumps virus P gene. Curran et al. (1985) have previously obtained cDNA clones of the mumps virus P gene, but their clones did not cover the entire length of the gene. In this paper, we report the nucleotide sequence of the mumps virus P gene, demonstrating the absence of an open reading frame (ORF) analogous to the C gene of other paramyxoviruses.

METHODS

Cells. LLC-MK2 and Vero cells were grown in Eagle’s minimum essential medium supplemented with 5% calf serum and antibiotics.

Viruses. The attenuated vaccine strain Miyahara, obtained from The Chemo-Sero Therapeutic Research Institute (Kumamoto, Japan), and the Vero cell-adapted Enders strain were propagated in Vero cells as reported previously (Tsurudome et al., 1984).
Preparation of nucleocapsid RNA. LLC-MK₂ cells were infected with the Miyahara strain at 0.1 to 0.2 p.f.u./cell. Five days after infection, cells were collected and lysed in lysis buffer (0.15 M-NaCl, 0.05 M-Tris-HCl pH 8.0, 0.6% NP40), and nuclei were removed by centrifugation. Intracellular viral nucleocapsids were purified using CsCl step gradient centrifugation and CsCl equilibrium centrifugation according to Kolakofsky (1976). Viral RNA was extracted from nucleocapsids with proteinase K and water-saturated phenol and then precipitated with ethanol.

Preparation of mRNA. LLC-MK₂ and Vero cells were infected with the Miyahara and Enders strains, respectively, at 0.1 to 0.2 p.f.u./cell and incubated at 37°C for 5 days. After addition of actinomycin D to infected cells at a final concentration of 5 μg/ml, they were further incubated at 37°C for 6 h. mRNA was prepared by the guanidinium/CsCl method (Maniatis et al., 1982) and purified using an oligo(dT)-cellulose column.

cDNA cloning from mRNA. cDNA was synthesized according to Okayama & Berg (1982) and introduced into *Escherichia coli* DH1. To identify virus-specific cDNA, transformant colonies were screened by colony hybridization (Grunstein & Hogness, 1975) using a ³²P-labelled probe synthesized by reverse transcription of mumps virus nucleocapsid RNA in the presence of random primers, prepared from calf thymus DNA, and [α-³²P]dCTP (New England Nuclear; 800 Ci/mmol).

cDNA cloning from nucleocapsid RNA. The first cDNA strand was synthesized from nucleocapsid RNA using reverse transcriptase (Bio-Rad) and pentadexynucleotide (Pharmacia). After purification of the cDNA-RNA hybrid using a Sephadex G-100 column, the second cDNA strand was synthesized using RNase H and DNA polymerase I according to Gubler & Hoffman (1983). The double-stranded cDNA was then inserted into the *PstI* site of pBR322 by dG-dC tailing.

DNA sequencing. Plasmid DNAs were cleaved into fragments with restriction endonucleases and ligated with M13 phage DNAs or pUC plasmid DNAs (Yanisch-Perron et al., 1985) by standard protocols (Maniatis et al., 1982). Bal 31 nuclease (Takara Syuzo, Japan) was used to remove the poly(A:T) stretch in the Okayama-Berg cloning vector. DNA sequences were determined by the dideoxy method (Sanger et al., 1977).

Northern hybridization. mRNAs were denatured with glyoxal and electrophoresed on a 1% agarose gel in 0.01 M-sodium phosphate buffer pH 7.2. The resolved RNAs were transferred to nylon membranes (Pall Ultrafine Filtration, U.S.A.) and hybridized with ³²P-labelled probes prepared by nick translation.

Hybrid-arrested translation. mRNA prepared from actinomycin D-treated virus-infected cells was hybridized to the double-stranded cDNA fragments isolated from specific plasmids (Paterson et al., 1977). The resulting mRNA-DNA hybrids were translated in wheatgerm extracts (Amersham) in the presence of L-[³⁵S]methionine and [α-³²P]dCTP (New England Nuclear; 1000 Ci/mmol), and the products were analysed by SDS-PAGE following immunoprecipitation with rabbit anti-mumps virus serum as described previously (Yamada et al., 1984).

Computer analysis. GENETYX (Software Development, Japan) was used to analyse the nucleotide and amino acid sequences.

RESULTS

A cDNA library was initially constructed from mRNA isolated from LLC-MK₂ cells infected with the Miyahara strain. Mumps virus-specific cDNA clones were selected by colony hybridization using ³²P-labelled probes synthesized by reverse transcription of nucleocapsid RNA. In order to identify P gene-specific cDNA clones, hybrid-arrested translation was carried out using mRNA extracted from mumps virus-infected cells and the double-stranded restriction fragments of mumps virus-specific cDNA clones. Fig. 1 shows typical results of hybrid-arrested translation. The NP and P proteins were readily detected after immunoprecipitation of *in vitro* translation products although, for an unknown reason, the ratio of the NP and P proteins varied from one experiment to another. Peptide mapping analysis showed that the NP and P gene products synthesized *in vitro* varied from one experiment to another. Peptide mapping analysis showed that the NP and P gene products synthesized *in vitro* were indistinguishable from the authentic polypeptides synthesized in infected cells (data not shown). Both the 150 bp *BamHI-BamHI* fragment and the 300 bp *BamHI-EcoRV* fragment arrested the synthesis of the P protein (Fig. 1, lanes 3 and 4). The latter fragment also arrested the synthesis of a 25K product (Fig. 1, lane 4). Peptide mapping after partial digestion with V8 protease revealed that the 25K protein was related to the 45K P protein (data not shown).

In order to confirm that pMMP9 contained a mumps virus-specific sequence, the 300 bp *BamHI-EcoRV* fragment of pMMP9 was nick-translated and hybridized to Northern blots of mRNA isolated from uninfected and mumps virus-infected LLC-MK₂ cells. The results are shown in Fig. 2. The fragment hybridized to a 1.2 kb mRNA species contained in virus-infected cells. Using pMMP9 as a probe, we selected six other cDNA clones from cDNA libraries constructed from nucleocapsid RNA of the Miyahara strain and mRNA of the Miyahara and Enders strains. Restriction endonuclease maps of seven overlapping cDNA clones are shown in
Fig. 1. Hybrid-arrested translation analysis using a wheatgerm lysate. mRNAs extracted from cells infected with the Miyahara strain of mumps virus were hybridized with cDNA fragments and translated in the wheatgerm extract in the presence of L-[35S]methionine. In vitro translation products were immunoprecipitated with rabbit anti-mumps virus serum and separated on a 10% polyacrylamide gel. (a) Restriction endonuclease map of pMMP9. (b) Lane 1, mumps virus polypeptides immunoprecipitated from infected cell extract with rabbit anti-mumps serum. Lane 2, no cDNA fragment added. Lanes 3 and 4 represent the results of a hybrid-arrested translation reaction using the clone pMMP9; for lane 3, the 150 bp BamHI-BamHI fragment (A) was used for hybridization, and for lane 4, the 300 bp BamHI-EcoRV fragment (B) was used for hybridization.

Fig. 2. Northern blot analysis using the 300 bp BamHI–EcoRV fragment of pMMP9 32P-labelled by nick translation. Northern blots were prepared as described in Methods. Lane 1, mRNA extracted from uninfected cells. Lane 2, mRNA extracted from cells infected with Miyahara strain virus.
Fig. 3. Restriction endonuclease maps and sequencing strategies of cDNA clones of the mumps virus P mRNA and mumps virus intracellular nucleocapsid RNA. Arrows indicate the direction and extent of sequence determination. AAAAA represents the poly(A:T) stretch in the Okayama–Berg cloning vector. The inserts of pMMP9 and pMMP10 were synthesized from mRNA prepared from cells infected with Miyahara strain virus. Those of pMGP8, pMGP15 and pMGP19 were synthesized from intracellular nucleocapsid RNA from cells infected with Miyahara strain virus. Those of pEMP4 and pEMP5 were synthesized from mRNA prepared from cells infected with Enders strain virus.

Fig. 3 along with the sequencing strategy. Recognition sites for Scal, Stul and EcoRV were absent from cDNA clones pEMP4 and pEMP5 which were derived from the Enders strain. Fig. 4 shows the nucleotide sequence and deduced amino acid sequence of the P genes of the Miyahara and Enders strains. The homology between these strains is 95% and 96% at the nucleotide and amino acid levels, respectively. The stop codon usage map of the Miyahara strain presented in Fig. 5 shows the presence of a single long ORF in frame 2, containing 1173 nucleotides and encoding a protein with a predicted size of 391 amino acids. No other ORF capable of coding for a protein of more than 56 amino acids was found, suggesting that only the P protein is coded for by this gene.
Fig. 4. Nucleotide sequence and deduced amino acid sequence of the P genes of the Miyahara and Enders strains of mumps virus. The top line is the nucleotide sequence of the Miyahara strain numbered from A of the first ATG codon. The middle line is the nucleotide sequence of the Enders strain. A dash indicates the same nucleotide as in the Miyahara strain. The bottom line is the deduced amino acid sequence of the Miyahara strain, with only amino acid substitutions in the Enders strain shown below.

Fig. 5. Termination codon usage in the message-sense strand for all three reading frames.
DISCUSSION

We have determined the nucleotide sequence of the P gene of mumps virus from overlapping cDNA clones. The P gene contains a single ORF which is 1173 nucleotides long and encodes a protein of 391 amino acids. The nucleotide sequences of the P genes of the Miyahara and Enders strains were very similar, but 58 nucleotide changes (5%) were found in the coding region, resulting in 17 amino acid substitutions (4%). The deduced Mr of the P protein of the Miyahara strain was calculated to be 41 587, which is in good agreement with the apparent value (40K to 45K) of the mumps virus P protein estimated from electrophoretic mobility on SDS-polyacrylamide gels. When the amino acid sequence of the mumps virus P protein was compared with those of Newcastle disease virus (NDV) and Sendai virus, we found a limited sequence homology with the NDV P protein. Residues 189 to 208 of the mumps virus P protein showed 60% homology with residues 195 to 214 of the NDV P protein and five consecutive amino acids of residues 301 to 305 of the mumps virus P protein were identical to amino acid residues 307 to 311 of the NDV P protein (not shown). The amino acid composition of the mumps virus P protein was similar to that of NDV (data not shown) (Sato et al., 1987). In contrast, no homology with the Sendai virus P protein was found. It has been reported previously that there is very little homology among the P proteins of paramyxoviruses (Sato et al., 1987).

A prominent feature of the mumps virus P gene found in this study is that no ORF analogous to the C gene of some other paramyxoviruses exists in the mumps virus P gene region. There are 12 ATG codons in the mumps virus P gene ORF and all the sequences preceding ATG codons were in a preferred context according to Kozak’s rule (PuXXATG) (Kozak, 1984) except the last one, suggesting the possible synthesis of truncated P proteins in infected cells. A 25K product detected by in vitro translation was presumed to be one of the truncated P proteins. The observation that no ORF corresponding to the C gene exists in the P gene region together with the possibility of synthesis of truncated P proteins in infected cells seem to support previous observations that non-structural proteins related to the P protein were induced in mumps virus-infected cells (Herrler & Compans, 1982; Simpson et al., 1984). We cannot rule out the possibility, however, that non-structural proteins unrelated to the P proteins are encoded by a gene other than the P gene. The 22K non-structural protein which is unrelated to the P protein (Merz et al., 1983) may be such a protein.

It has been suggested that paramyxoviruses may be divided into two groups by their coding strategy for the P and non-structural protein genes (Spriggs & Collins, 1986). The P and non-structural proteins are encoded in two different reading frames of the P gene in Sendai, PF-3 and measles viruses, whereas they are encoded in the same reading frame in simian virus 5 (SV5), mumps virus and NDV. With respect to NDV, however, Sato et al. (1987) reported the presence of a second small ORF which overlapped the reading frame coding for the P protein. The predicted Mr of the product of the new ORF was calculated to be 11230, but no proteins of this Mr have been identified in NDV-infected cells. Two non-structural proteins of 33K and 36K found in NDV-infected cells were shown to be related to the P protein (Collins et al., 1982). Similarly, V is the only non-structural protein that is also a P protein-related product that has been found in SV5-infected cells (Paterson et al., 1984). Whether truncated P proteins of the latter group of viruses perform the same function as the C protein of the former group must await the elucidation of the role of these non-structural proteins.

REFERENCES


Mumps virus P gene sequence


(Received 2 March 1988)