The Mumps Virus Fusion Protein mRNA Sequence and Homology among the Paramyxoviridae Proteins

By NARAYANASAMY ELANGO,*† TAMAS M. VARSANYI, JAN KÖVAMEES AND ERLING NORRBY

Department of Virology, School of Medicine, Karolinska Institute, Stockholm S-105 21, Sweden

(Accepted 29 November 1988)

SUMMARY

The complete nucleotide sequence of the fusion protein (F) mRNA of the virulent SBL-1 strain of mumps virus has been determined by sequencing cDNA clones and mRNA and confirmed by partially sequencing the genomic RNA. The mRNA was 1721 nucleotides long excluding the poly(A) sequence and had one long open reading frame which encoded a protein of 538 amino acids with a calculated Mₐ of 58 791. The predicted amino acid sequence had a proteolytic cleavage/activation site, Arg Arg His Lys Arg, cleavage at which yields proteins F₂ and F₁. The uncleaved protein contained three highly hydrophobic regions: (i) the amino-terminal signal peptide, (ii) the amino-terminal region of F₁ and (iii) the carboxy-terminal membrane anchorage domain. There were seven potential N-glycosylation sites, two in F₂ and five in F₁. Comparison of the virulent strain F protein sequence with that of an avirulent strain of mumps virus showed a difference of 14 amino acids. Among paramyxoviruses, mumps virus fusion protein shows the highest degree of homology with the fusion proteins of simian virus 5 and Newcastle disease virus.

INTRODUCTION

Mumps virus, an enveloped RNA virus, is a member of the paramyxovirus family, which contains viruses of human and veterinary importance. The virus envelope contains a non-glycosylated matrix protein and two glycoproteins, a haemagglutinin–neuraminidase protein (HN) and a fusion protein (F), which form spike-like projections on the outer surface of the virus particle. Inside the envelope is the ribonucleoprotein complex containing the genome of approximately 15000 nucleotides of negative-sense RNA, a nucleocapsid protein (NP), a polymerase-associated protein (P) and a large protein (L) (Wolinsky & Server, 1985). Mumps virus has recently been shown to encode a small 6·7K hydrophobic protein (SH) (Elango et al., 1988).

The fusing strains of mumps virus are neurovirulent, whereas the non-fusing strains are not (McCarthy et al., 1980). The fusion proteins of both virulent and avirulent strains of mumps virus are synthesized as a precursor, F₀, which is cleaved in vivo; however, the fusion proteins of avirulent strains do not cause cell-to-cell fusion (Merz et al., 1983). To find out whether there are any differences in the primary structure of the fusion proteins of virulent and avirulent strains of mumps virus and to understand the fusion mechanism further, we have sequenced the fusion protein mRNA of the virulent SBL-1 strain of mumps virus and compared the deduced amino acid sequence with those of the avirulent RW strain (Waxham et al., 1987), simian virus 5 (SV5) (Paterson et al., 1984) and Newcastle disease virus (NDV) (McGinnes & Morrison, 1986).

* Present address: Biotechnology Research Institute, National Research Council of Canada, 6100 Avenue Royal Mount, Montreal, Quebec, Canada H4P 2R2.
METHODS

Cells and virus. Vero cells were grown in Eagle's minimum essential medium supplemented with 5% heat-inactivated foetal calf serum. The SBL-1 strain of mumps virus (Orvell, 1978) was plaque-purified three times and grown in Vero cells.

Isolation of mRNA and genomic RNA. Isolation of RNA from virus-infected and actinomycin D-treated Vero cells and purification of poly(A)-containing RNA have been described (Elango et al., 1988).

cDNA clones. The construction and identification of mumps virus-specific cDNA clones has been described (Elango et al., 1988). Three clones, designated pMF1, pMF2 and pMF3, with inserts which hybridized to two mRNA species of 2000 and 2500 nucleotides in Northern blot analysis, were selected for further analysis.

DNA sequencing. Plasmid DNA was purified by the NaOH-SDS lysis method (Maniatis et al., 1982). cDNA inserts were digested with appropriate restriction enzymes, the fragments were subcloned into M13 phage vectors (mp18 and mp19) and sequenced by the dideoxynucleotide method (Sanger et al., 1977) using the Sequenase Kit from the United States Biochemical Corporation. Wherever necessary, synthetic oligonucleotides were used as primers to sequence regions not covered by restriction sites.

RNA sequencing. Oligonucleotides were synthesized using a Pharmacia Gene Assembler. Labelling of the 5' end of gel-purified oligonucleotides and dideoxynucleotide sequencing of RNA were carried out as described (Elango, 1989).

Sequence analysis. Computer-assisted sequence analysis was accomplished with Beckman's Microgenie software.

RESULTS

Identification of F-specific cDNA clones

Five different groups of mumps virus-specific cDNA clones were obtained from a cDNA library (Elango et al., 1988). One group contained the three cDNA clones designated pMF1, pMF2 and pMF3 with insert sizes of 900 bp, 900 bp and 600 bp, respectively. These three inserts hybridized, in Northern blots, to two mRNAs of 2000 and 2500 nucleotides indicating that either the 2500 nucleotide mRNA is polycistronic or the cDNA inserts were from polycistronic mRNA. Limited sequencing showed that each of these three inserts contained poly(A) and had the same sequence adjacent to the poly(A) end. When the mumps virus genomic RNA was sequenced with an mRNA-sense oligonucleotide homologous to a region adjacent to the poly(A) end of the three inserts, the SH gene sequence was obtained (Elango et al., 1988). By analogy with SV5 (Hiebert et al., 1985), it was reasoned that the F gene should precede the SH gene and that the clones pMF1, pMF2 and pMF3 therefore represented the F mRNA.

Nucleotide sequence of the F mRNA and its predicted amino acid sequence

The presence of poly(A) in all three clones suggested that the 3' end of the mRNA had been cloned. The cDNA insert from pMF1 was sequenced completely by the dideoxynucleotide method (Sanger et al., 1977). To obtain the missing 5' mRNA sequence an oligonucleotide complementary to nucleotides 880 to 864 of the mRNA (Fig. 1) was synthesized using the 5' end of the pMF1 insert and used to 'walk' on the F mRNA. Another oligonucleotide was synthesized from the new sequence and mRNA walking was continued until the 5' terminus of the mRNA was sequenced.

Fig. 2 shows mapping of the 5' end of the F mRNA. Consistent with dideoxynucleotide sequencing of the 5' terminus of the RNA, two bands that differ by one nucleotide were seen in all four lanes. The 5' sequence of the F mRNA is 5' NAGCCUAGAAG. The terminal nucleotide was found to be an A by sequencing the genomic RNA. The sequence obtained for the 5' end of the mRNA was identical to that obtained when the genome was sequenced with an M mRNA-sense oligonucleotide sequence close to the poly(A) region of M mRNA (Elango et al., 1988). The sequence seen beyond the 5' terminus of the F mRNA shows the 3' end of M mRNA, coming from the M–F dicistronic mRNA, confirming the gene order.

The complete nucleotide sequence of the F mRNA, exclusive of poly(A), is shown in Fig. 1. The mRNA is 1721 nucleotides long and has a single, long open reading frame starting at the first AUG (nucleotides 79 to 81) and ending at UAA (nucleotides 1678 to 1680). The sequence around the initiation codon, AUCAUGA, partially conforms to the consensus sequence,
Fig. 1. Complete nucleotide sequence of the mumps virus F mRNA and the predicted amino acid sequence of the F protein. The signal sequence, the F l polypeptide amino-terminal hydrophobic region and the membrane anchorage domain are underlined. The proteolytic cleavage site is indicated by broken lines and the potential glycosylation sites are boxed.
Fig. 2. Sequencing the 5' end of the F mRNA. 5' End-labelled, genome-sense oligonucleotide was used to sequence the 5' end of the F mRNA by the dideoxynucleotide method. Sequencing reaction products were analysed on a 6% sequencing gel. The deduced sequence (genome sense) is shown. Sequence above the strong stops represents the M mRNA sequence of the M-F dicistronic mRNA.

δXXAUGG, which surrounds most eukaryotic functional translation initiation codons (Kozak, 1984). The predicted protein contains 538 amino acids with a calculated Mr of 58791.

**DISCUSSION**

The F mRNA of a fusion-inducing strain of mumps virus has been sequenced. The mRNA is 1721 nucleotides long and the deduced amino acid sequence is 538 residues long. The calculated Mr (58791) of the predicted F protein is slightly greater than that of the unglycosylated protein (53K) identified in tunicamycin-treated mumps virus-infected cells (Herder & Compans, 1983). The protein has three highly hydrophobic regions: the signal peptide (amino acid residues 1 to 19), the amino-terminal region (amino acid residues 103 to 128) and the membrane anchorage region (amino acid residues 483 to 512) of F1. The activation of F0 involves a proteolytic cleavage after the last Arg residue of the cleavage site, Arg Arg His Lys Arg (amino acid residues 98 to 102). This cleavage site is identical to that present in the measles virus fusion protein (Richardson et al., 1986). Whether these five amino acids are removed from the cleavage site or remain as part of F2 is not known. There are two potential N-glycosylation sites in F2 and five sites in F1. The F0 has 13 Cys residues, two in the signal peptide, one in F2 and 10 in F1, out of which one is in the membrane anchor region. The single Cys residue of F2 must be involved in the disulphide bond formation between F1 and F2.

The nucleotide and the deduced amino acid sequences of the F protein of the avirulent RW strain of mumps virus have been published (Waxham et al., 1987); the exact position of the 5'
terminus of the mRNA was not determined and the sequence contained a part of the M gene sequence. As we have determined the exact 5' terminus of the F mRNA, only the corresponding mRNA sequence was considered for comparison. Comparison of the mRNA sequences of the SBL-1 strain and RW strain shows 95.5% homology at the nucleotide sequence level (data not shown) and 97.4% homology at the amino acid level (Fig. 3). The predicted fusion proteins are of the same size, but with a mismatch of 14 amino acids. Of the 14 amino acid changes, three are in the signal peptide, two are in F2, eight are in the extracellular part of F1, and one is in the membrane anchor region (Fig. 3). The changes in F2 are Val to Ile (position 49) and Thr to Ala (position 91). The change from Thr to Ala implies the loss of a potential glycosylation site. None of the eight amino acid changes in the extracellular part of F1 of the RW strain is in the fusion peptide and they are not clustered. It is not known whether the 14 amino acid changes in the fusion protein of the RW strain make the virus non-fusing or if there are additional factor(s) or proteins involved in the fusion activity of mumps virus.

Using fusion protein mutants of Sendai virus (Scheid & Choppin, 1976; Tashiro et al., 1988) and NDV (Toyoda et al., 1987), it has been shown that the cytopathic effect and virulence of paramyxoviruses are in part due to the fusion activity of the F protein and that this fusion activity depends on the proteolytic cleavage of F0 by host proteases. Using a Sendai virus ts mutant containing <5% of the HN found in wild-type virus, Gibson et al. (1988) showed that the HN protein is unlikely to mediate any fusion reactions.

In contrast, the F0 of the non-virulent strains of mumps virus contains a proteolytic cleavage site identical to that of the virulent SBL-1 strain and is cleaved in vivo, but does not cause cell-to-cell fusion (Merz et al., 1983). After observing a more active neuraminidase in avirulent strains than in virulent strains of mumps virus, Merz & Wolinsky (1981) suggested that active neuraminidase promotes the rapid release of the progeny virus so that no cell-to-cell fusion occurs, whereas less active neuraminidase prolongs the association of progeny virus with the infected cell surface thereby increasing the likelihood of cell-to-cell fusion. Later, Merz & Wolinsky (1983) showed that selective proteolysis of the HN could convert non-fusing mumps virus infections to fusing infections. Waxham & Wolinsky (1986) obtained a fusing variant of the non-fusing OTake strain of mumps virus by growing the non-fusing virus under the selective pressure of a competitive inhibitor of neuraminidase, and showed that the fusing variant has no detectable neuraminidase activity whereas the parental OTake strain has. These studies indicate that the neuraminidase of the mumps virus HN modulates the fusion activity of mumps virus.

The HN gene sequences of the avirulent RW strain (Waxham et al., 1988) and virulent SBL-1 strain (Kövamees et al., 1989) have been published recently. Comparison of both sequences showed a homology of 95% at the amino acid level and the HN of strain RW differs from that of strain SBL-1 by 29 amino acids. Of the 29 amino acid changes five are in the cytoplasmic region and the rest are in the extracellular region of the RW HN. None of the changes is in the putative neuraminidase region (Jorgensen et al., 1987) except the change Ala Asn to Thr Asp (positions 265 and 266) and both proteins contain the conserved areas present in all paramyxovirus HN proteins (Kövamees et al., 1989).

Direct evidence to demonstrate that the RW F has fusing activity could be obtained by expressing it in an expression system and then studying its fusion activity.

There is a considerable homology between the fusion proteins of mumps virus, SV5 (Paterson et al., 1984) and NDV (McGinnes & Morrison, 1986). Fig. 3 shows the alignment of the three fusion proteins. Homology is seen throughout the molecule and is greater between the F proteins of mumps virus and SV5. There are two areas of homology which stand out: the hydrophobic sequence at the amino-terminal region of F1 and an area of 10 amino acids of complete identity. Comparison of the mumps virus F protein sequence with that of measles virus, canine distemper virus, parainfluenza virus type 3 and Sendai virus did not show any highly conserved areas except that at the amino-terminal region of F1 (data not shown). The only other feature shared between the F proteins of these paramyxoviruses is the conservation of nine Cys residues present in the extracellular region of F1.
Fig. 3. Comparison of the F protein sequences of a virulent (V) strain of mumps virus (SBL-1 strain), a non-virulent (NV) strain of mumps virus (RW strain), ST5 and NDV. Amino acids identical to those of SBL-1 are marked by asterisks. Boxed areas show regions conserved among all the four viruses.

This work was supported by project B88-I6X-00116-24B of the Medical Research Council, Sweden and a grant from the Wallenberg Foundation, Stockholm, Sweden.

REFERENCES


Mumps virus fusion protein sequence


(Received 5 October 1988)