Volume 15 Number 7 1987 Nucleic Acids Research

Nucleotide sequence of the human parainfluenza virus 3 matrix protein gene

Kevin Prinoski, Marie-José Côté, C.Y.Kang and Kenneth Dimock

Department of Microbiology and Immunology, University of Ottawa, Health Science Centre, Ottawa,
Ontario KlH 8M5, Canada

Submitted February 25, 1987 Accession no. Y00119

The nucleotide sequence of the matrix (M) gene of human parainfluenza virus 3 (HPIV3), strain 47885, was determined using overlapping cDNA clones derived from mRNA and genomic RNA. The sequence is presented as DNA in the mRNA sense beginning with the first nucleotide of M mRNA. The HPIV3 M gene is 1150 nucleotides long, excluding 5 3'- terminal A residues. A single long open reading frame begins at nucleotide 33 and encodes a polypeptide of 353 amino acids (molecular weight 39499). The HPIV3 M protein is basic (net charge +20.5 at pH 7). Many arginine and lysine residues are conserved among paramyxovirus M proteins (1-4). Two slightly hydrophobic stretches are found (amino acids 173-208 and 264-294); the second within a region, with a predicted secondary structure common among paramyxovirus M proteins, believed to interact with but not span the envelope (4-6). The HPIV3 M protein exhibits considerable homology with the M protein of Sendai virus (∿62%) but less homology with the M proteins of measles virus (∿30%), and NDV (∿18%).

```
AGGATTAAAGAATAAATTAATCCTTGTCCAAA ATG AGT ATA ACT AAC TCT GCA ATA TAC ACA TTC CCA GAG TCA TCA TTC TCT GAG AAT GGT CAT ATA   98
                                 Met Ser Ile Thr Asn Ser Ala Ile Tyr Thr Phe Pro Glu Ser Ser Phe Ser Glu Asn Gly His Ile   22

 GAA CCA TTA CCA CTC AAA GTC AAT GAA CAG AGA AAA GCA GTA CCT CAC ATT AGA GTT GCC AAA ATC GGA AAT CCA CCA AAA CAC GGA TCC  188
 Glu Pro Leu Pro Leu Lys Val Asn Glu Gln Arg Lys Ala Val Pro His Ile Arg Val Ala Lys Ile Gly Asn Pro Pro Lys His Gly Ser   52

 CGG TAT TTG GAT GTC TTC TTA CTC GGC TTC TTC GAG ATG GAA CGA ATC AAA GAC AAA TAC GGG AGT GTG AAT GAT CTT GAC AGT GAC CCG  278
 Arg Tyr Leu Asp Val Phe Leu Leu Gly Phe Phe Glu Met Glu Arg Ile Lys Asp Lys Tyr Gly Ser Val Asn Asp Leu Asp Ser Asp Pro   82

 GGT TAC AAA GTT TGT GGC TCT GGA TCA TTA CCA ATC GGA TTA GCC AAA TAC ACT GGG AAT GAC CAG GAA TTA TTA CAG GCT GCA ACT AAA  368
 Gly Tyr Lys Val Cys Gly Ser Gly Ser Leu Pro Ile Gly Leu Ala Lys Tyr Thr Gly Asn Asp Gln Glu Leu Leu Gln Ala Ala Thr Lys  112

 CTG GAT ATA GAA GTG AGA AGA ACA GTT AAA GCG AAA GAA ATG ATT GTT TAT ACG GTA CAA AAT ATA AAA CCA GAA CTG TAC CCA TGG TCC  458
 Leu Asp Ile Glu Val Arg Arg Thr Val Lys Ala Lys Glu Met Ile Val Tyr Thr Val Gln Asn Ile Lys Pro Glu Leu Tyr Pro Trp Ser  142

 AGT AGA CTA AGA AAA GGA ATG TTG TTC GAT GCC AAC AAA GTT GCT CTT GCT CCT CAA TGT CTT CCA CTA GAT AGG AGC ATA AAA TTC AGA  548
 Ser Arg Leu Arg Lys Gly Met Leu Phe Asp Ala Asn Lys Val Ala Leu Ala Pro Gln Cys Leu Pro Leu Asp Arg Ser Ile Lys Phe Arg  172

 GTA ATC TTC GTT AAT TGT ACG GCA ATT GGA TCA ATA ACC TTG TTC AAA ATT CCT AAG TCA ATG GCA TCA CTA TCT CTA CCC AGC ACA ATA  638
 Val Ile Phe Val Asn Cys Thr Ala Ile Gly Ser Ile Thr Leu Phe Lys Ile Pro Lys Ser Met Ala Ser Leu Ser Leu Pro Ser Thr Ile  202

 TCA ATC AAT CTT CAG GTA CAC ATC AAA ACA GGG GTT CAG ACT GAT TCT AAA GGG ATA GTT CAA ATT TTG GAT GAG AAG GGT GAA AAA TCA  728
 Ser Ile Asn Leu Gln Val His Ile Lys Thr Gly Val Gln Thr Asp Ser Lys Gly Ile Val Gln Ile Leu Asp Glu Lys Gly Glu Lys Ser  232

 CTG AAT TTC ATG GTC CAT CTC GGA TTG ATC AAA AGA AAA GTA GGC AGA ATG TAC TCT GTC GAG TAC TGT AAA CAG AAA ATC GAG AAA ATG  818
 Leu Asn Phe Met Val His Leu Gly Leu Ile Lys Arg Lys Val Gly Arg Met Tyr Ser Val Glu Tyr Cys Lys Gln Lys Ile Glu Lys Met  262

 AGA TTG ATA TTT TCT TTG GGA TTA GTT GGA GGA ATC AGT CTT CAT GTC AAT GCA ACT GGA TCT ATA TCA AAA ACA CTA GCA AGT CAG CTG  908
 Arg Leu Ile Phe Ser Leu Gly Leu Val Gly Gly Ile Ser Leu His Val Asn Ala Thr Gly Ser Ile Ser Lys Thr Leu Ala Ser Gln Leu  292

 GTA TTC AAA AGG GAG ATT TGT TAT CCC TTA ATG GAT CTA AAT CCA CAT CTC AAT CTA GTT ATC TGG GCT TCA TCA GTA GAG ATT ACA AGA  998
 Val Phe Lys Arg Glu Ile Cys Tyr Pro Leu Met Asp Leu Asn Pro His Leu Asn Leu Val Ile Trp Ala Ser Ser Val Glu Ile Thr Arg  322

 GTG GAT GCA ATT TTC CAA CCT TCT TTA CCT GGC GAG TTC AGA TAC TAT CCT AAC ATT ATT GCA AAA GGA GTT GGG AAA ATC AAA CAA TGG 1088
 Val Asp Ala Ile Phe Gln Pro Ser Leu Pro Gly Glu Phe Arg Tyr Tyr Pro Asn Ile Ile Ala Lys Gly Val Gly Lys Ile Lys Gln Trp  352

AAC TAG TAATCTCTATTTTGATCTGGATATATCTATTAAGCCAAAGCAAATAAGAGATAATCAAAAA                                                 1155
Asn ---                                                                                                                   353
```

1. Blumberg, B.M. et al. (1984) J. Virol. 52, 656-663. 2. Hidaka, Y. et al. (1984) Nucl. Acids Res. 12, 7965-7972. 3. Bellini, W.J. et al. (1986) J. Virol. 58, 408-416. 4. Chambers, P. et al. (1986) Nucl. Acids Res. 14, 9051-9061. 5. Garnier, J. et al. (1978) J. Mol. Biol. 120, 97-120. 6. Kyte, J. and Doolittle, R.F. (1982) J. Mol. Biol. 157, 105-132.