J. gen. Virol. (1987), **68**, 47–56. *Printed in Great Britain*

Key words: *coronavirus MHV-JHM/nucleotide sequence/surface projection glycoprotein gene*

Nucleotide Sequence of the Gene Encoding the Surface Projection Glycoprotein of Coronavirus MHV-JHM

By IRENE SCHMIDT, MICHAEL SKINNER† AND STUART SIDDELL*

Institute of Virology, University of Würzburg, Versbacher Strasse 7, 8700 Würzburg, F.R.G.

(*Accepted 15 September 1986*)

SUMMARY

Sequences encoding the surface projection glycoprotein of the coronavirus, murine hepatitis virus (MHV), strain JHM, have been cloned into pAT153 using cDNA produced by priming with specific oligonucleotides on infected cell RNA. The regions of three clones pJMS1010, pJS112 and pJS92, which together encompass the surface protein gene have been sequenced by the chain termination method. The sequence of the primary translation product, deduced from the DNA sequence, predicts a polypeptide of 1235 amino acids with a molecular weight of 136600. This polypeptide displays the features characteristic of a group 1 membrane protein; an amino-terminal signal sequence and carboxy-terminal membrane and cytoplasmic domains. There are 21 potential glycosylation sites in the polypeptide and a cysteine-rich region in the vicinity of the transmembrane domain. During maturation proteolytic processing of the polypeptide occurs and at positions 624 to 628 the sequence Arg–Arg–Ala–Arg–Arg is found, which is similar to a number of basic sequences involved in the cleavage of enveloped RNA virus glycoproteins. The fusogenic properties of the MHV surface protein do not appear to correlate with a strongly hydrophobic region at the putative amino terminus of the carboxy-terminal cleavage product.

INTRODUCTION

Coronaviruses are pleomorphic, enveloped viruses which replicate in the cytoplasm of vertebrate cells and are associated with diseases of economic importance (Siddell *et al.*, 1983*a*). In the laboratory, the murine hepatitis virus (MHV) has been extensively used for the study of viral pathogenesis, in particular as a component of a model for demyelinating diseases of man (Knobler *et al.*, 1982; Watanabe *et al.*, 1983; Massa *et al.*, 1986). The MHV genome is a monopartite, positive-stranded RNA of approximately 18 kb. The genome encodes the nucleocapsid (N), membrane (M or E1) and surface (S or E2) proteins of the virion, as well as several non-structural proteins (Sturman & Holmes, 1983; Siddell *et al.*, 1983*b*).

The MHV S protein is synthesized on membrane-bound ribosomes as a co-translationally *N*-glycosylated polypeptide with an apparent mol. wt. of 150000 (Niemann *et al.*, 1982; Holmes *et al.*, 1981; Siddell *et al.*, 1981). The polypeptides synthesized *in vitro* or in tunicamycin-treated cells have mol. wt. of approximately 120000 (Rottier *et al.*, 1981; Siddell, 1983). During transport within the cell, oligosaccharides are trimmed and terminal sugars are added, resulting in a 180000 mol. wt. S polypeptide. Shortly before, or at the time of, virus release a proportion of S is cleaved into two approximately 90000 mol. wt. polypeptides, S_1 and S_2 (Niemann *et al.*, 1982; Sturman *et al.*, 1985). S_1 and S_2 (which are also referred to as 90B and 90A; Sturman *et al.*, 1985) cannot be distinguished by SDS–PAGE but can be separated by hydroxyapatite chromatography. It has been shown that S_2 is acylated (Ricard & Sturman, 1985). The cleavage of the S polypeptide is a host cell-dependent event (Frana *et al.*, 1985) and activates its cell-fusing ability. The S protein is also responsible for the attachment/infectivity of the MHV virion and some monoclonal hybridoma antibodies which react with the S protein are able to mediate virus

† Present address: Department of Microbiology, University of Reading, London Road, Reading RG1 5AQ, U.K.

0000-7380 © 1987 SGM

neutralization *in vitro* and passively protect mice against lethal virus challenge *in vivo* (Collins *et al.*, 1982).

The organization and expression of the MHV genome has been studied in detail (for reviews, see Holmes, 1985; Siddell, 1986) (Fig. 1). Briefly, in MHV-infected cells six subgenomic mRNAs, as well as genome-sized RNA, are produced. These mRNAs form a 3′ co-terminal nested set and each also has a common 5′ leader sequence of about 70 bases (Lai *et al.*, 1984). The available evidence suggests that only the information contained within the 'unique' sequences at the 5′ end of each mRNA (i.e. those absent from the next smallest RNA) is translated into protein (Siddell, 1986). The translation of size-fractionated MHV mRNAs has shown that subgenomic mRNA 3 encodes the S protein (Siddell, 1983).

Previously, we have isolated cDNA clones containing overlapping viral inserts which encompass approximately 4·6 kb at the 3′ end of the MHV-JHM genome (Skinner & Siddell, 1983, 1985; Skinner *et al.*, 1985; Pfleiderer *et al.*, 1986). Using specific oligonucleotide primers we have now isolated two further clones which contain inserts extending to the 5′ end of mRNA 3. The regions of the three clones which together contain the MHV S gene (Fig. 1) have been completely sequenced on both strands. This sequence, together with the predicted amino acid sequence of the S gene product is presented in this paper.

METHODS

cDNA cloning. The isolation and characterizaton of the plasmid pJMS1010 has been previously described (Skinner *et al.*, 1985). The growth of Sac(−) cells, the propagation of MHV-JHM stocks and the isolation of polyadenylated RNA from MHV-JHM-infected cells have also been described (Siddell *et al.*, 1980). The oligonucleotide primers A (3′ GTCGACGACCACACGG 5′), B (3′ GTGTGGGACATTCGGAT 5′) and others were synthesized using the phosphoramidite method on an Applied Biosystems 380 A DNA Synthesizer. cDNA synthesis was carried out using the method of Gubler & Hoffman (1983) with slight modifications. In particular, prior to trailing the double-stranded cDNA with dC residues, potential RNA overhangs were removed by treatment with DNase-free RNase A (10 μg/ml) for 8 min at 37 °C. The tailed ds cDNA was cloned into dG-tailed *Pst*I-cleaved pAT153. This material was used to transform *Escherichia coli* DH1 and selection was made for tetracycline resistance. Clones containing viral inserts were identified by colony hybridization using polynucleotide kinase ^{32}P-labelled, cDNA synthesis primer as probe. The size of viral inserts in plasmids from hybridizing clones was determined by gel electrophoresis of *Pst*I-cleaved DNA. An oligonucleotide (3′ GCATGCCTGCGGTTAG 5′), which corresponds to a region near the 5′ end of the MHV-JHM mRNA leader (Skinner & Siddell, 1983) was used in hybridizations to identify the plasmid pJS92.

Subcloning in M13. Fragments of the viral inserts contained within pJMS1010, pJS112 and pJS92 were generated by a variety of restriction enzymes and were cloned either as mixtures or as single fragments (purified by electroelution from either agarose or acrylamide gel) into the M13 vectors mp8, mp9, mp18 and mp19. Where necessary, specific clones were identified by hybridization to single-stranded DNA probes generated from characterized M13 clones (O'Hare *et al.*, 1983).

DNA sequencing. M13 dideoxynucleotide sequencing was carried out using [α-^{35}S]dATP. The complete sequence was obtained on both strands. To complete the project oligonucleotides complementary to specific MHV sequences were synthesized and used to prime the sequencing reactions. Sequence data were analysed and assembled using the programs of Staden (1982*a*).

Southern/Northern analysis. Northern blot analysis of RNA following electrophoresis in 1% agarose–formaldehyde gels, Southern blot analyses of DNA and nick translations were performed according to Maniatis *et al.* (1982).

RESULTS

The position of the MHV sequences contained within the plasmid pJMS1010 has been previously determined by Northern blot and sequence analysis (Skinner *et al.*, 1985). The viral insert within the plasmid extends from within the M protein gene (which is translated from mRNA 6) to a position approximately 2·6 kb from the 5′ end of mRNA 3 (Fig. 1). A 16-base oligonucleotide, primer A, complementary to a sequence towards the 5′ end of the pJMS1010 insert was used to prime cDNA synthesis from infected cell poly(A)-containing RNA. Plasmid pJS112 obtained from this experiment contained a 2·2 kb insert which hybridized in Northern blots to the MHV mRNAs 3, 2 and 1 (data not shown). Sequence analysis confirmed that the 3′ end of the insert corresponded to the cDNA synthesis primer. In a second cDNA synthesis experiment a 17-base oligonucleotide, primer B, complementary to a sequence towards the 5′

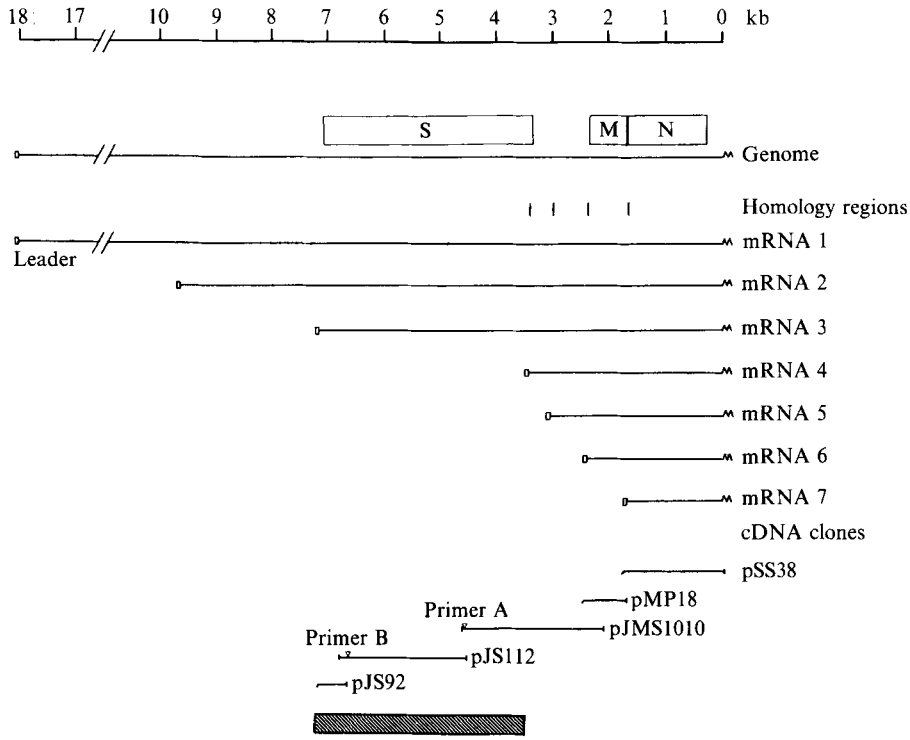


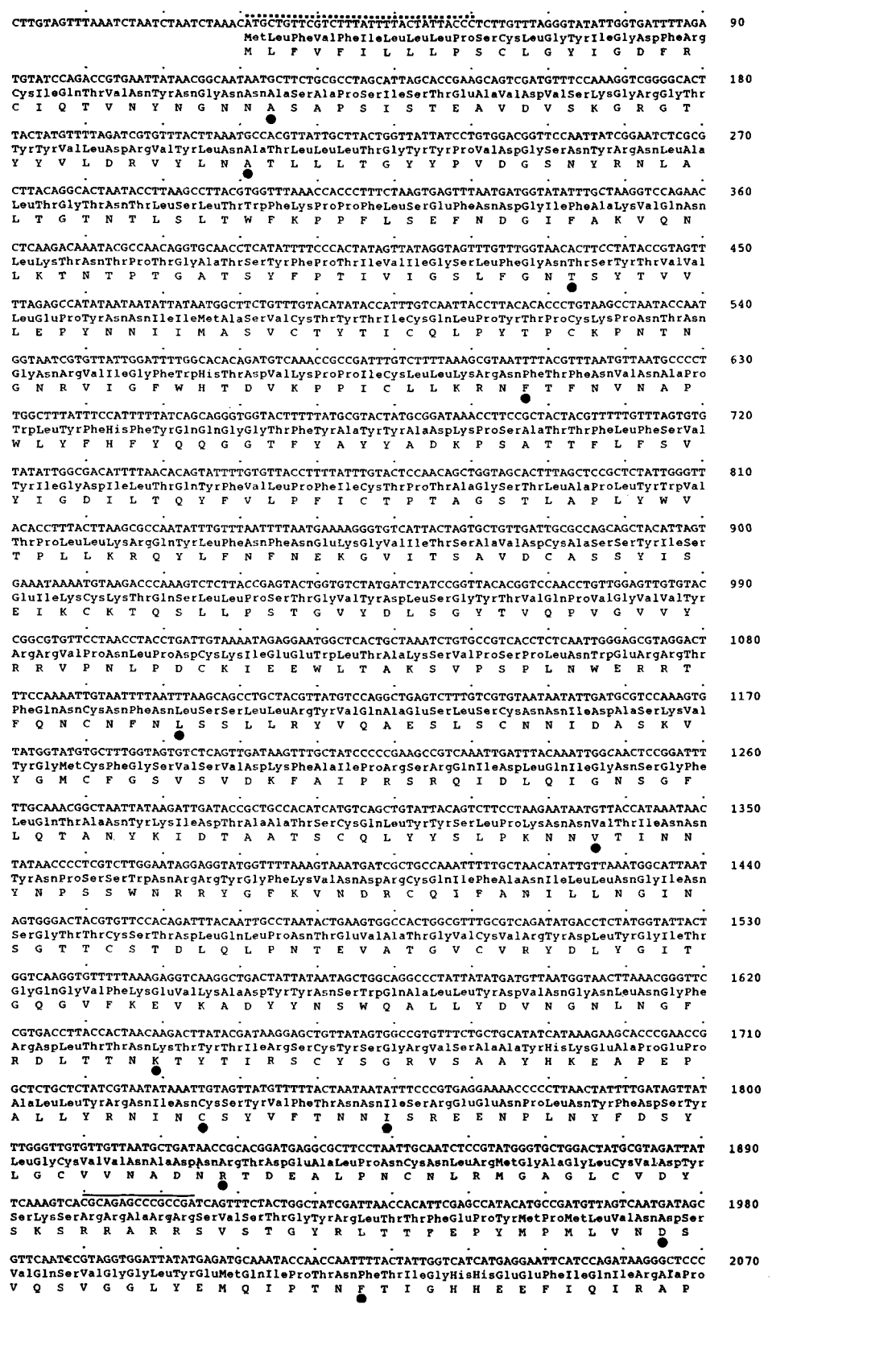
Fig. 1. Genomic organization of murine hepatitis virus. The relationship between the 3′ co-terminal nested set of mRNAs and the viral genome is shown, together with the coding regions for the structural proteins, nucleocapsid (N), membrane (M) and surface (S), specified by mRNAs 7, 6 and 3 respectively. The arrangement of the cDNA clones and the positions of the primers used are shown. The sequences presented in Fig. 2 are represented by the hatched box.

end of the pJS112 insert was used to obtain the plasmid pJS92. pJS92 contained an insert of 530 bp and hybridized in Northern blot analysis to all viral mRNAs. Hybridization of the pJS92 insert to the cDNA synthesis primer and to a primer corresponding to the MHV-JHM leader sequences (see Methods) was confirmed by Southern blot analysis of *Pst*I-cleaved pJS92 DNA (data not shown).

A 3780-base sequence containing the gene encoding the MHV-JHM S propolypeptide (i.e. the predicted primary translation product) is presented in Fig. 2. Immediately preceding the AUG initiation codon is the sequence UCUAAAC. This sequence is identical to genomic sequences preceding the known or presumed 5′ initiation codons of mRNAs 7, 5 and 4 and differs by only one base from the sequence UCCAAAC, preceding the initiation codon of mRNA 6 (Skinner *et al.*, 1985; Skinner & Siddell, 1985; Pfleiderer *et al.*, 1986). It is thought that these sequences, referred to as regions of homology, are involved in regulating the synthesis of MHV mRNAs (Armstrong *et al.*, 1984; Spaan *et al.*, 1983).

The AUG codon at position 31 initiates an open reading frame (ORF) of 3705 bases encoding a polypeptide of 1235 amino acids with a predicted mol. wt. of 136600. This ORF ends with a single UGA termination codon. The sequence context of the initiating codon, AAAC*AUG*C, is frequently found amongst functional eukaryotic initiator sequences (Kozak, 1983).

A number of structural features of the S propolypeptide are noteworthy. Firstly, within the MHV-S propolypeptide sequence there are 21 potential *N*-glycosylation sites of the type Asn–X–Thr/Ser (assuming that X is not Pro) (Fig. 2). The distribution of these sites is also shown in Fig. 3. It is clear that at least one cluster of potential glycosylation sites occurs in the carboxy-terminal region of the polypeptide, between amino acids 1092 and 1158. Secondly, a hydropathicity plot of the amino acid sequence of the S propolypeptide, determined using the



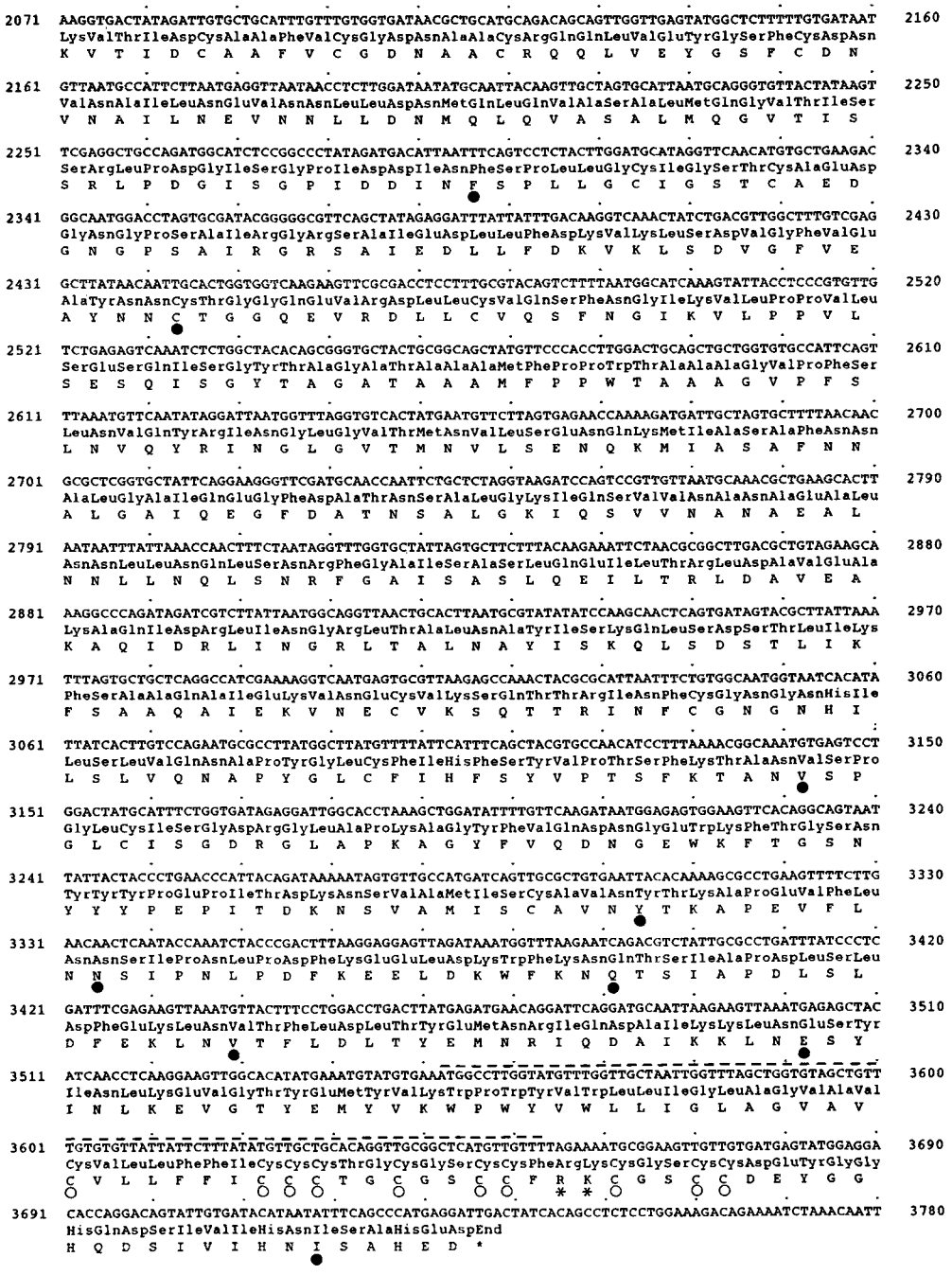


Fig. 2. Nucleotide sequence of the MHV surface protein gene and the predicted amino acid sequence of the surface protein precursor. The amino-terminal signal sequence (*****), the carboxy-terminal transmembrane domain (---), the charge cluster (**), the cysteine-rich region (○), potential glycosylation sites (●) and the putative proteolytic cleavage site (——) are indicated.

procedure of Kyte & Doolittle (1982), reveals two regions of striking hydrophobicity (Fig. 3). At the amino terminus, the initiator methionine is followed by nine non-polar amino acids and this hydrophobic core precedes a number of small neutral residues (e.g. Ser-11 Gly-14) which are characteristically found at the signal peptidase recognition site (Fig. 2) (Von Heijne, 1984). At

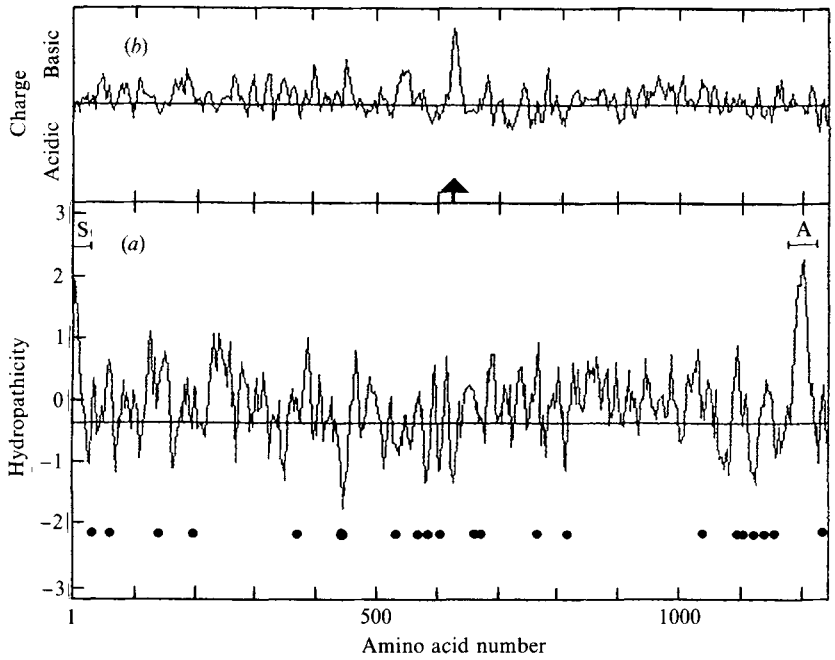


Fig. 3. (*a*). Hydropathicity analysis of the S propolypeptide according to the method of Kyte & Doolittle (1982). The vertical scale is the average hydropathicity (+3 to −3) for a frame of seven amino acids. The midpoint line represents the average hydropathicity of the 20 amino acids. Hydrophobic sequences appear above the base line. The putative signal (S) and anchor (A) domains, as well as the potential glycosylation sites (●) are shown. (*b*) Charge analysis of the MHV S propolypeptide. The analysis sums charge values over a window of nine amino acids. The putative cleavage site is indicated (↑).

the carboxy terminus of the propolypeptide a sequence of 34 predominantly hydrophobic and neutral amino acids is followed by two positively charged amino acids (Arg-1209, Lys-1210), which are situated 25 amino acids from the carboxy terminus. Moreover, these 'charge cluster' residues are flanked by an unusual distribution of cysteine residues, which constitute 50% of the residues between positions 1198 to 1215. Not only an unusual distribution, but also a specific sequence, Cys–Gly–Ser–Cys–Cys, is found on either side of the 'charge cluster' (Fig. 2). Finally, the distribution of charged amino acids in the S propolypeptide (Fig. 3) indicates a particularly striking domain of basic residues, Arg–Arg–Ala–Arg–Arg, at positions 624 to 628.

DISCUSSION

The DNA sequences presented in Fig. 2 encompass the entire 'unique' region of MHV-JHM mRNA 3. In common with the mRNAs for the virion proteins, N and M, the 5′ unique region of mRNA 3 is used to encode a single polypeptide, the precursor to the surface protein. In each of the mRNAs encoding MHV virion structural proteins, translation is initiated at the first AUG codon within the mRNA and the expressed ORF does not overlap with any downstream ORF present in the 3′ co-terminal sequences. In mRNAs 3, 6 and 7 the initiating codon is found in a preferred context and follows closely (1, 4 and 8 bases respectively) the 'region of homology' sequences which define the 5′ ends of the mRNA bodies (Skinner & Siddell, 1983; Pfleiderer *et al.*, 1986). These data further support the model of a non-overlapping translation strategy for coronavirus gene expression, at least for the virion structural proteins (Siddell, 1986; Brown *et al.*, 1986).

The predicted amino acid sequence of the S propolypeptide, derived from the DNA sequence, reveals several interesting features. The predicted mol. wt. of the S propolypeptide is 136600. This agrees approximately with the size of the polypeptide found in tunicamycin-treated cells or

in vitro translation (Siddell, 1983). If the average mol. wt. of mannose-rich viral glycoprotein carbohydrate side chains is assumed to be 2000 to 3000, it seems likely that a considerable number of the 21 potential glycosylation sites are utilized, in order to account for the approximately 40000 to 50000 apparent mol. wt. difference between the glycosylated and non-glycosylated S polypeptides (Siddell, 1982).

The hydropathicity analysis indicates that the MHV S polypeptide belongs to the group 1 membrane proteins (Garoff, 1985). These proteins are inserted across the endoplasmic reticulum membrane starting from their amino terminus using the same mechanism as secretory proteins. They therefore carry a typical amino-terminal hydrophobic signal sequence which is not present in the mature protein. A putative signal sequence is present in the predicted MHV S propolypeptide. The small neutral glycine residue at position 14 appears to be a possible signal peptidase cleavage site (Von Heijne, 1984).

A second hydrophobic region at the carboxy terminus of the MHV S propolypeptide is characteristic of a transmembrane domain and is delineated from the hydrophilic cytoplasmic domain by the positively charged Arg/Lys residues at positions 1209/1210. Garoff and colleagues (Cutler & Garoff 1986; Cutler *et al.*, 1986) have tested the postulate that the hydrophobic transmembrane domain, together with the 'charge cluster' of the cytoplasmic domain make up a membrane binding region of group 1 proteins, which acts as a 'stop transfer signal' to arrest translocation and prevent secretion. Their results, however, indicate that the 'charge cluster' is necessary only for stabilization of the protein–membrane interaction and the hydrophobic stretch alone is able to arrest translocation. Surrounding the charge cluster in the MHV S polypeptide are an unusual number of cysteine residues, which occur in a specific sequence context. It seems possible that these sequences are involved in the acylation of the S_2 polypeptide, because acylation of the vesicular stomatitis virus G protein is believed to occur at cysteine residues in the vicinity of the hydrophobic transmembrane domain (Rose *et al.*, 1984; McGee *et al.*, 1984).

During virus maturation the MHV S polypeptide is cleaved by a host cell enzyme to yield the S_1 and S_2 polypeptides. The charge analysis of the MHV S polypeptide reveals a sequence Arg–Arg–Ala–Arg–Arg at positions 624 to 628 which is very similar to a number of basic sequences involved in the cleavage of several other enveloped, RNA virus glycoproteins (White *et al.*, 1983). Sturman & Holmes (1977) have shown that the MHV-A59 S polypeptide can be cleaved by trypsin and it appears likely that coronaviruses, in common with many enveloped RNA viruses, utilize a cellular trypsin-like endoprotease activity to achieve proteolytic processing. It is not yet known whether a second, carboxypeptidase, enzyme plays a role in the maturation of the coronavirus S protein, as has been shown for the influenza virus haemagglutinin (Garten & Klenk, 1983).

Following cleavage the fusion properties of the MHV S protein are activated (Sturman *et al.*, 1985). Examination of the MHV S polypeptide sequence shows that cleavage at the site mentioned above would not result in a strongly hydrophobic domain at the amino terminus of S_2. This would contrast with the fusogenic myxovirus proteins, where the hydrophobicity of the amino terminus of, for example, the influenza virus HA_2 protein is essential to the fusion process. Possibly, the mechanism of coronavirus-induced cell fusion involves either a less hydrophobic amino-terminal domain on S_2 or other, as yet unidentified, hydrophobic domains on the S protein.

Finally, it is of interest to compare the nucleic acid and predicted amino acid sequences of the MHV S protein gene reported here with the recently determined sequence of the avian infectious bronchitis coronavirus (IBV) S protein gene (Binns *et al.*, 1985). At the nucleic acid level the sequences appear to be essentially unrelated. However, a comparison of the predicted amino acid sequences using the dot matrix program DIAGON (Staden, 1982*b*), which looks not only for identical residues, but also for residues with similar properties, reveals a striking degree of similarity in the S_2 polypeptide, but little similarity in the S_1 region (Fig. 4). To some extent the similarities in the S_2 region represent recognizable features, for example, the transmembrane domain, the cysteine-rich region, or the putative cleavage site. Additionally however, there are regions of amino acids with similar properties, and also specific sequences of identical

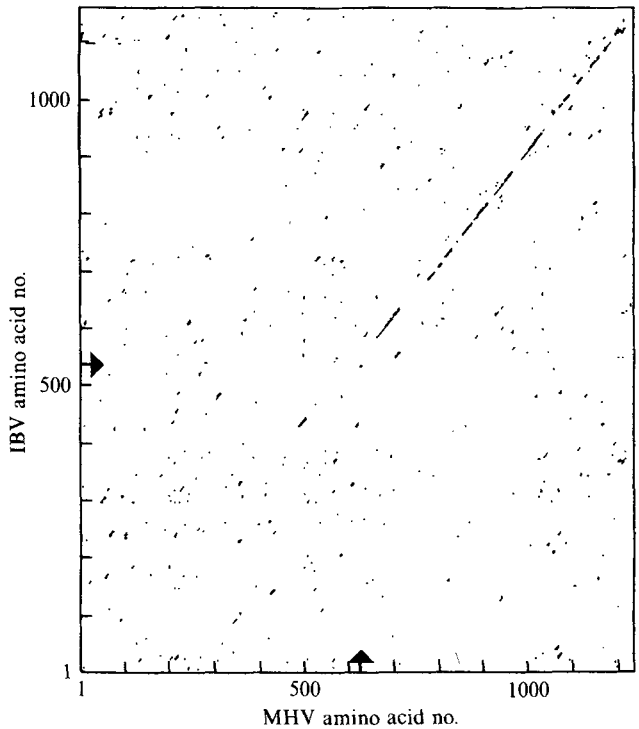


Fig. 4. DIAGON analysis (Staden, 1982*b*) of the homologies between the MHV-JHM S propolypeptide amino acid sequence and the IBV Beaudette S propolypeptide sequences (Binns *et al.*, 1985). Points represent matches at the 1% significance level. The putative MHV cleavage site as well as the known cleavage site (Cavanagh *et al.*, 1986) are indicated (↑).

amino acids (for example, the sequence Trp–Pro–Trp–Tyr–Val–Trp–Leu, positions 1184 to 1191 for MHV; 1092 to 1099 for IBV), the significance of which remains to be determined.

The cloning and sequencing of the MHV S protein gene is an important step in studies on the molecular biology of MHV and especially the interaction between the virus, the host cell and the immune system during infection. Experiments to investigate these interactions at the molecular level can now be undertaken.

We would like to thank Barbara Schelle-Prinz for skilful technical assistance, Helga Kriesinger for typing the manuscript and Professor V. ter Meulen for continuous support. We would also like to thank R. Staden for computer programs. This work was supported by the Deutsche Forschungsgemeinschaft (SFB 165).

REFERENCES

ARMSTRONG, J., SMEEKINS, S., SPAAN, W., ROTTIER, P. & VAN DER ZEIJST, B. (1984). Cloning and sequencing the nucleocapsid and E1 genes of coronavirus. *Advances in Experimental Medicine and Biology* **173**, 155–162.

BINNS, M. M., BOURSNELL, M. E. G., CAVANAGH, D., PAPPIN, D. J. C. & BROWN, T. D. K. (1985). Cloning and sequencing of the gene encoding the spike protein of the coronavirus IBV. *Journal of General Virology* **66**, 719–726.

BROWN, T. D. K., BOURSNELL, M. E. G., BINNS, M. M. & TOMLEY, F. M. (1986). Genomic organization of avian coronavirus IBV. *The Molecular Biology of the Positive Strand RNA Viruses*. Edited by D. J. Rowlands *et al.* New York: Academic Press (in press).

CAVANAGH, D., DAVIS, P. J., PAPPIN, D. J. C., BINNS, M. M., BOURSNELL, M. E. G. & BROWN, T. D. K. (1986). Coronavirus IBV: partial amino-terminal sequencing of spike polypeptide S2 identifies the sequence ARG–ARG–PHE–ARG–ARG at the cleavage site of the spike precursor propolypeptide of IBV strains Beaudette and M41. *Virus Research* **4**, 133–143.

COLLINS, A. R., KNOBLER, R. L., POWELL, H. & BUCHMEIER, M. J. (1982). Monoclonal antibodies to murine hepatitis virus 4 (strain JHM) define the viral glycoprotein responsible for attachment and cell-cell fusion. *Virology* **119**, 358–371.

CUTLER, D. F. & GAROFF, H. (1986). Mutants of the membrane-binding region of Semliki Forest virus E2 protein. I. Cell surface transport and fusogenic activity. *Journal of Cell Biology* **102**, 889–901.

CUTLER, D. F., MELANCON, P. & GAROFF, H. (1986). Mutants of the membrane-binding region of Semliki Forest virus E2 protein. II. Topology and membrane binding. *Journal of Cell Biology* **102**, 902–910.

FRANA, M. F., BEHNKE, J. N., STURMAN, L. S. & HOLMES, K. V. (1985). Proteolytic cleavage of the E2 glycoprotein of murine coronavirus: host-dependent differences in proteolytic cleavage and cell fusion. *Journal of Virology* **56**, 912–920.

GAROFF, H. (1985). Using recombinant DNA techniques to study protein targeting in the eucaryotic cell. *Annual Review of Cell Biology* **1**, 403–445.

GARTEN, W. & KLENK, H.-D. (1983). Characterization of the carboxypeptidase involved in the proteolytic cleavage of the influenza haemagglutinin. *Journal of General Virology* **64**, 2127–2137.

GUBLER, U. & HOFFMAN, B. J. (1983). A simple and very efficient method for generating cDNA libraries. *Gene* **25**, 263–269.

HOLMES, K. V. (1985). Replication of coronaviruses. In *Virology*, pp. 1331–1343. Edited by B. N. Fields, D. M. Knipe, R. M. Chanock, J. L. Melnick, B. Roizman & R. E. Shope. New York: Raven Press.

HOLMES, K. V., DOLLER, E. W. & STURMAN, L. S. (1981). Tunicamycin resistant glycosylation of coronavirus glycoprotein; demonstration of a novel type of glycoprotein. *Virology* **115**, 334–344.

KNOBLER, R. L., LAMPERT, P. W. & OLDSTONE, M. B. A. (1982). Virus persistence and recurring demyelination produced by a temperature-sensitive mutant of MHV-4. *Nature, London* **298**, 279–280.

KOZAK, M. (1983). Comparison of initiation of protein synthesis in procaryotes, eucaryotes and organelles. *Microbiological Reviews* **47**, 1–45.

KYTE, J. & DOOLITTLE, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology* **157**, 105–132.

LAI, M. M. C., BARIC, R. S., BRAYTON, P. R. & STOHLMAN, S. A. (1984). Characterization of leader RNA sequences on the virion and mRNAs of mouse hepatitis virus, a cytoplasmic RNA virus. *Proceedings of the National Academy of Sciences, U.S.A.* **81**, 3626–3630.

McGEE, A. I., KOYAMA, A. H., MALTER, C., WEN, D. & SCHLESINGER, M. J. (1984). Release of fatty acids from virus glycoproteins by hydroxylamine. *Biochimica et biophysica acta* **798**, 156–166.

MANIATIS, T., FRITSCH, E. F. & SAMBROOK, J. (1982). *Molecular Cloning: A Laboratory Manual.* New York: Cold Spring Harbor Laboratory.

MASSA, P., DÖRRIES, R. & TER MEULEN, V. (1986). Virus particles induce Ia antigen expression on astrocytes. *Nature, London* **320**, 543–546.

NIEMANN, H., BOSCHEK, B., EVANS, D., ROSING, M., TAMURA, T. & KLENK, H.-D. (1982). Post-translational glycosylation of coronavirus glycoprotein E1, inhibition by monensin. *EMBO Journal* **1**, 1499–1504.

O'HARE, K., LEVIS, R. & RUBIN, G. M. (1983). Transcription of the *white* locus in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences, U.S.A.* **80**, 6917–6921.

PFLEIDERER, M., SKINNER, M. A. & SIDDELL, S. G. (1986). Coronavirus MHV-JHM: nucleotide sequence of the mRNA that encodes the membrane protein. *Nucleic Acids Research* **14**, 6338.

RICARD, C. S. & STURMAN, L. S. (1985). Isolation of the subunits of the coronavirus envelope glycoprotein E2 by hydroxyapatite high-performance liquid chromatography. *Journal of Chromatography* **326**, 191–197.

ROSE, J. K., ADAMS, G. A. & GALLIONE, C. J. (1984). The presence of cysteine in the cytoplasmic domain of the vesicular stomatitis virus glycoprotein is required for palmitate addition. *Proceedings of the National Academy of Sciences, U.S.A.* **81**, 2050–2054.

ROTTIER, P. J. M., SPAAN, W. J. M., HORZINEK, M. C. & VAN DER ZEIJST, B. A. M. (1981). Translation of three mouse hepatitis virus (MHV-A59) subgenomic RNAs in *Xenopus laevis* oocytes. *Journal of Virology* **38**, 20–26.

SIDDELL, S. G. (1982). Coronavirus JHM: tryptic peptide fingerprinting of virion proteins and intracellular polypeptides. *Journal of General Virology* **62**, 259–269.

SIDDELL, S. (1983). Coronavirus JHM: coding assignments of subgenomic mRNAs. *Journal of General Virology* **64**, 113–125.

SIDDELL, S. (1986). The organization and expression of coronavirus genomes. In *The Molecular Biology of the Positive Strand RNA Viruses*. Edited by D. J. Rowlands *et al.* New York: Academic Press (in press).

SIDDELL, S. G., WEGE, H., BARTHEL, A. & TER MEULEN, V. (1980). Coronavirus JHM. Cell-free synthesis of structural protein p60. *Journal of Virology* **33**, 10–17.

SIDDELL, S., WEGE, H., BARTHEL, A. & TER MEULEN, V. (1981). Coronavirus JHM: intracellular protein synthesis. *Journal of General Virology* **53**, 145–155.

SIDDELL, S. G., ANDERSON, R., CAVANAGH, D., FUJIWARA, K., KLENK, H.-D., MACNAUGHTON, M. R., PENSAERT, M., STOHLMAN, S. A., STURMAN, L. & VAN DER ZEIJST, B. A. M. (1983*a*). Coronaviridae. *Intervirology* **20**, 181–189.

SIDDELL, S., WEGE, H. & TER MEULEN, V. (1983*b*). The biology of coronaviruses. *Journal of General Virology* **64**, 761–776.

SKINNER, M. A. & SIDDELL, S. G. (1983). Coronavirus JHM: nucleotide sequence of the mRNA that encodes nucleocapsid protein. *Nucleic Acids Research* **11**, 5045–5054.

SKINNER, M. A. & SIDDELL, S. G. (1985). Coding sequence of coronavirus MHV-JHM mRNA 4. *Journal of General Virology* **66**, 593–596.

SKINNER, M. A., EBNER, D. & SIDDELL, S. G. (1985). Coronavirus MHV-JHM mRNA 5 has a sequence arrangement which potentially allows translation of a second, downstream open reading frame. *Journal of General Virology* **66**, 581–592.

SPAAN, W., DELIUS, H., SKINNER, M., ARMSTRONG, J., ROTTIER, P., SMEEKINS, S., VAN DER ZEIJST, B. A. M. & SIDDELL, S.G. (1983). Coronavirus mRNA synthesis involves fusion of non-contiguous sequences. *EMBO Journal* **2**, 1839–1844.

STADEN, R. (1982*a*). Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing. *Nucleic Acids Research* **10**, 4731–4751.

STADEN, R. (1982*b*). An interactive graphics program for aligning nucleic acid and aminoacid sequences. *Nucleic Acids Research* **10**, 2951–2961.

STURMAN, L. S. & HOLMES, K. V. (1977). Characterization of a coronavirus. II. Glycoproteins of the viral envelope: tryptic peptide analysis. *Virology* **77**, 650–660.

STURMAN, L. S. & HOLMES, K. V. (1983). The molecular biology of coronaviruses. *Advances in Virus Research* **28**, 35–112.

STURMAN, L. S., RICARD, C. S. & HOLMES, K. V. (1985). Proteolytic cleavage of the E2 glycoprotein of murine coronavirus: activation of cell-fusing activity of virions by trypsin and separation of two different 90K cleavage fragments. *Journal of Virology* **56**, 904–911.

VON HEIJNE, G. (1984). How signal sequences maintain cleavage specificity. *Journal of Molecular Biology* **173**, 243–251.

WATANABE, R., WEGE, H. & TER MEULEN, V. (1983). Adoptive transfer of EAE-like lesions from rats with coronavirus-induced demyelinating encephalomyelitis. *Nature, London* **305**, 150–153.

WHITE, J., KIELIAN, M. & HELENIUS, A. (1983). Membrane fusion proteins of enveloped animal viruses. *Quarterly Reviews of Biophysics* **16**, 151–195.

(Received 28 July 1986)