VIROLOGY 161, 479–487 (1987)

Primary Structure of the Glycoprotein E2 of Coronavirus MHV-A59 and Identification of the Trypsin Cleavage Site

WILLEM LUYTJES,* LAWRENCE S. STURMAN,† PETER J. BREDENBEEK,* JEROEN CHARITE,* BERNARD A. M. VAN DER ZEIJST,‡ MARIAN C. HORZINEK,* AND WILLY J. M. SPAAN*,[1]

*Institute of Virology, Veterinary Faculty, State University of Utrecht, 3508 TD Utrecht, The Netherlands; †Wadsworth Center for Laboratories and Research, New York State Department of Health, Albany, New York 12201; and ‡Section of Bacteriology, Veterinary Faculty, State University of Utrecht, The Netherlands

Received March 17, 1987; accepted August 3, 1987

The nucleotide sequence of the peplomer (*E2*) gene of MHV-A59 was determined from a set of overlapping cDNA clones. The *E2* gene encodes a protein of 1324 amino acids including a hydrophobic signal peptide. A second large hydrophobic domain is found near the COOH terminus and probably represents the membrane anchor. Twenty glycosylation sites are predicted. Cleavage of the E2 protein results in two different 90K species, 90A and 90B (L. S. Sturman, C. S. Ricard, and K. V. Holmes (1985) *J. Virol.* 56, 904–911), and activates cell fusion. Protein sequencing of the trypsin-generated N-terminus revealed the position of the cleavage site. 90A and 90B could be identified as the C-terminal and the N-terminal parts, respectively. Amino acid sequence comparison of the A59 and JHM E2 proteins showed extensive homology and revealed a stretch of 89 amino acids in the 90B region of the A59 E2 protein that is absent in JHM.

INTRODUCTION

Murine hepatitis viruses (MHV) are coronaviruses which cause a variety of diseases including hepatitis and encephalomyelitis in the natural host (Wege *et al.*, 1982). They are studied extensively, since MHV is a useful animal model for virus-induced demyelination and because coronaviruses possess a unique mode of replication (Siddell *et al.*, 1983).

The infectious genome of MHV consists of a single-stranded RNA of about 20 kb which is associated with a single protein species with a mol wt of 54K in a helical nucleocapsid. Two membrane-associated proteins are present in the virions: the large glycoprotein E2, forming the characteristic surface projections or peplomers, and the smaller membrane glycoprotein E1 (26.5K) (Armstrong *et al.*, 1984a). The peplomer protein, encoded by mRNA 3 (Rottier *et al.*, 1981), is synthesized on ribosomes bound to the rough endoplasmic reticulum (RER) where it is cotranslationally glycosylated (Sturman and Holmes, 1983; Holmes *et al.*, 1984) and subsequently acylated, probably during transport through the Golgi apparatus (Nieman and Klenk, 1981; Sturman *et al.*, 1985). MHV virions bud from the RER and Golgi membranes and are apparently transported to the exterior by the internal secretory apparatus. Two forms of the E2 protein are present on the surface of the virion, the 180K and the 90K species (Sturman and Holmes, 1977). Recently it has been shown that the 90K protein consists of two different species, 90A and 90B, arising from proteolytic cleavage of the 180K protein (Sturman *et al.*, 1985). This cleavage activates cell fusion, and the ratio of 180K to 90K proteins is host dependent (Sturman *et al.*, 1985; Frana *et al.*, 1985). It has been suggested that such host-dependent differences in the processing of E2 may be important for cytopathic effects, virulence, and tissue tropism of the murine coronaviruses (Frana *et al.*, 1985).

The peplomer protein is involved in cell attachment (Collins *et al.*, 1982) and is the target for neutralizing antibodies (Fleming *et al.*, 1983). E2 plays an important role in the pathology of MHV. Buchmeier *et al.* (1984) showed that in MHV-JHM-infected mice passive transfer of neutralizing monoclonal antibodies, recognizing E2, prevented fatal infection by wild-type virus. Instead, a chronic demyelinating disease developed. These changed pathogenic properties seem to be a result of mutations in E2 (Dalziel *et al.*, 1986; Fleming *et al.*, 1986). To understand the biological and pathogenic properties of MHV at the molecular level the primary structure of E2 and data on its processing are essential. Here we report the cDNA cloning and sequence analysis of the gene encoding the E2 protein of MHV-A59. By direct amino acid sequence analysis of the N-terminal part of the 90A species we were also able to identify the trypsin cleavage site. E2 is the main structural protein determining strain differences be-

[1] To whom requests for reprints should be addressed.

0042-6822/87 $3.00

cause it shows significant antigenic polymorphism, in contrast to the other structural proteins (Talbot and Buchmeier, 1985). It will probably reflect the major differences between related coronavirus species. To localize these differences we have compared the *E2* gene sequence of the MHV-A59 strain with that of strain JHM published recently (Schmidt *et al.*, 1987).

MATERIALS AND METHODS

Identification of the trypsin cleavage site by amino acid sequence analysis

Virus purification was carried out with modification as described by Sturman *et al.* (1980, 1985). Purified 90A and 90B proteins were prepared from trypsin-treated virions and the 180K E2 protein was prepared from untreated virus. Following incubation with trypsin, 10 μg/ml, in TMEN, pH 6.5 (50 m*M* Tris–maleate, 1 m*M* EDTA, 100 m*M* NaCl), at 37° for 30 min, soybean trypsin inhibitor, 50 μg/ml, was added for 30 min at 4°. Virus was then sedimented at 24,000 rpm in an SW 28 rotor at 4° for 2.5 hr. E2 was extracted with Triton X-114, and 90A and 90B were separated as described previously by HPLC on HPHT (hydroxyapatite) columns in sodium dodecyl sulfate (SDS) (Ricard and Sturman, 1985). Uncleaved (180K) E2 was separated from 90K species by HPLC size exclusion chromatography with a Bio-Sil TSK (Bio-Rad) guard column and Bio-Sil TSK 400, 7.5 × 300 mm, and Spherogel TSK 4000 (Altex), 7.5 × 300 mm, columns connected in series. SDS was removed from purified proteins by ion pair extraction with acetone:triethylamine:acetic acid:water, 85:5:5:5 (Henderson *et al.*, 1979). Proteins were washed with trifluoroacetic acid, lyophylized, and dissolved in trifluoracetic acid. The amino terminal sequence was determined by automated Edman degradation using an Applied Biosystems gas phase sequencer. Phenylthiohydantoyn (PTH) amino acids were identified by HPLC.

cDNA synthesis and cloning

Viral genomic RNA and poly(A)-containing intracellular RNAs were isolated from purified MHV-A59 virions and infected cells, respectively (Spaan *et al.*, 1981). Procedures for synthesis of cDNA were essentially identical to those described by Dowling (1983) and Gubler and Hoffman (1983). For the synthesis of the single-stranded cDNA, pentanucleotides and specific primers were used. Full details will be presented elsewhere (P. J. Bredenbeek *et al.*, manuscript in preparation).

After homopolymer tailing of the double-stranded cDNA (Peacock, 1981) or digestion with restriction endonucleases, the cDNA was annealed to dG-tailed pUC9 DNA (Pharmacia) or ligated to pEMBL DNA (Dente *et al.*, 1983), respectively. Transformation was carried out by adding the annealed or ligated DNA to *Escherichia coli* strain JM101 or JM109 competent cells (Messing, 1983), prepared by the method described by Hanahan (1983), which were subsequently plated on petri dishes containing 25 μg/ml ampicillin.

Screening and analysis of recombinants

Plasmid DNA from ampicillin-resistant colonies obtained after transformation of the ligated restriction fragments was prepared according to the method described by Birnboim and Doly (1979). The mapping of the cDNA clones on the genome will be described in detail elsewhere (P. J. Bredenbeek *et al.*, manuscript in preparation).

Formaldehyde–agarose gel analysis and hybridization

Poly(A)-containing RNA from MHV-infected cells was denatured in the presence of formaldehyde and separated in an 1.5% agarose–formaldehyde gel (Lehrach *et al.*, 1977). After electrophoresis the gel was dried on Whatmann 3 MM paper and subsequently incubated with a kinase-labeled oligonucleotide probe according to Meinkoth and Wahl (1984). Hybridization and washing temperature was 5–10° below the calculated T_d.

Oligonucleotide synthesis

Oligonucleotides were prepared as described previously (Niesters *et al.*, 1986) or were synthesized using a DNA-synthesizer, Biosearch Model 8600, and subsequently purified by HPLC.

DNA sequence analysis

DNA fragments were prepared by digestion with a variety of restriction enzymes and ligated either as a mixture or as single fragments purified from agarose gels into the M13 vectors mp8 and mp9 (Messing, 1983). White plaques were screened for viral inserts using pentamer-primed probes from cDNA clones (Feinburg and Vogelstein, 1983; Roberts and Wilson, 1985). Single-stranded M13 DNA was isolated and used for sequence analysis using the dideoxynucleotide chain termination procedure of Sanger *et al.* (1977). Sequence data were assembled and analyzed using the computer programs created by Staden (1986).

Protein sequence homology searches

The predicted amino acid sequence was compared to other sequences and to the NBRF Protein Bank using the FASTP program set of Lipman and Pearson (1985) and the DIAGON program of Staden (1982).

RESULTS

cDNA cloning, mapping of recombinant plasmids, and sequence analysis

When we started this sequence study, a number of cDNA clones against cellular RNAs of MHV-A59 was already available. Mapping by hybridization to the viral mRNAs (P. J. Bredenbeek *et al.*, manuscript in preparation) and sequence analysis indicated that the overlapping clones 95, 918, and 85 were positioned around the 5′ end of the *E2* gene (Fig. 1). Clone 853 was mapped beyond the 3′ end in gene *D*. Oligonucleotide 7 (OL 7) complementary to a sequence in the 3′ end of clone 85 and oligonucleotide 8 (OL 8) based upon the sequence of clone 853 were synthesized and used to screen the new random genomic cDNA library. Several positive recombinant DNA clones were isolated and characterized by restriction site mapping. This permitted construction of a continuous map of approximately 5 kb containing the complete unique region of mRNA 3 encoding the E2 protein.

The large insert of clone B24 was isolated and subsequently digested with restriction endonuclease *Hpa*II or *Taq*I. The complete digests were ligated into M13 mp9. Initial selection of subclones overlapping the consensus sequence of clones 95, 918, and 85 was performed by hybridizing a probe from clone B60 to phage DNA. The sequence strategy is summarized in Fig. 1.

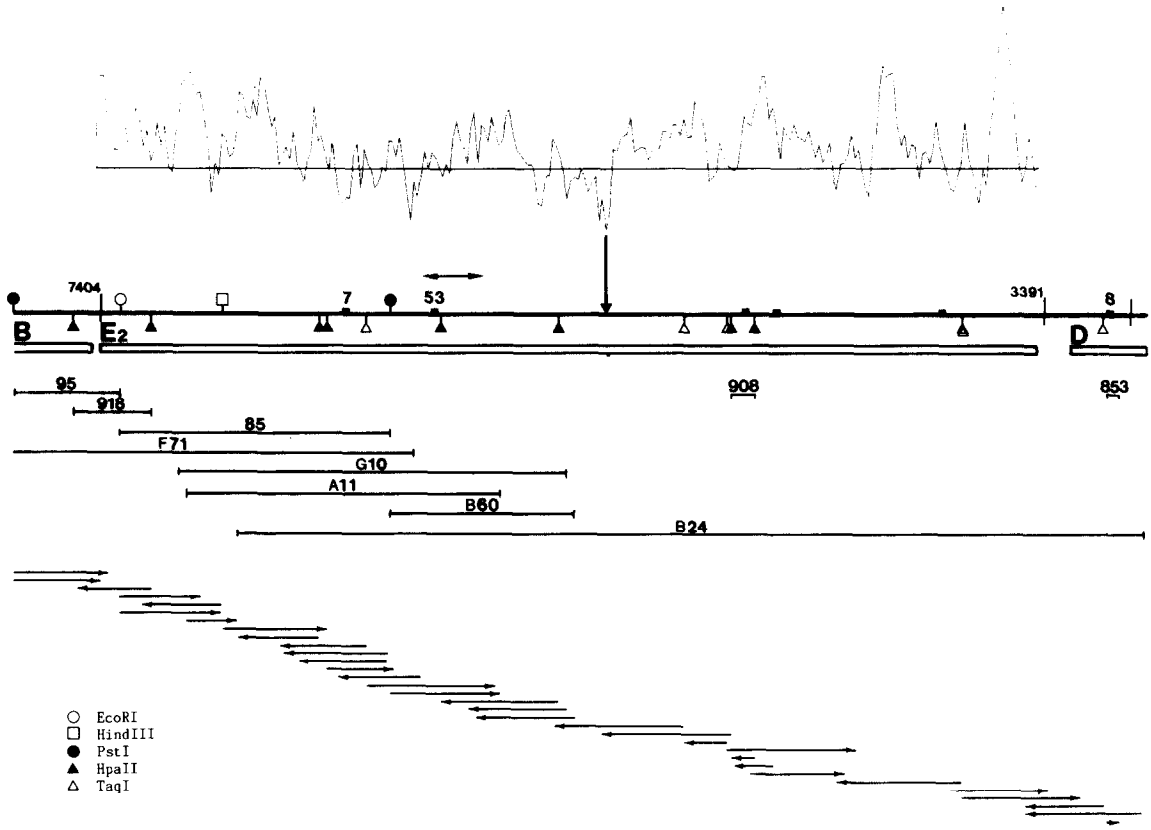

FIG. 1. Cloning and sequencing strategy of the MHV-A59 *E2* gene and hydrophobicity pattern of the predicted amino acid sequence. Open boxes represent open reading frames in the coding regions of RNA 2 (B), RNA 3 (E2), and RNA 4(D) transposed on the genome. Vertical bars indicate homology regions in the intergenic sequences of MHV-A59. Numbers represent the nucleotide distance to the start of the 3′ poly(A) tail. The vertical arrow points at the trypsin cleavage site. The double arrow marks the region in MHV-A59 absent in strain JHM. Small boxes represent the synthetic oligonucleotides used in cloning, sequencing, and hybridizations, numbered when referred to in the text. cDNA clones are indicated by horizontal lines. Extent and direction of sequencing is shown by means of the arrows below. Symbols indicating restriction sites are explained in the figure. The hydrophobicity pattern was generated using the HYDROPLOT program created by Staden (1986), modified with hydrophobicity data from Eisenberg *et al.* (1982). Above the line is hydrophobic.

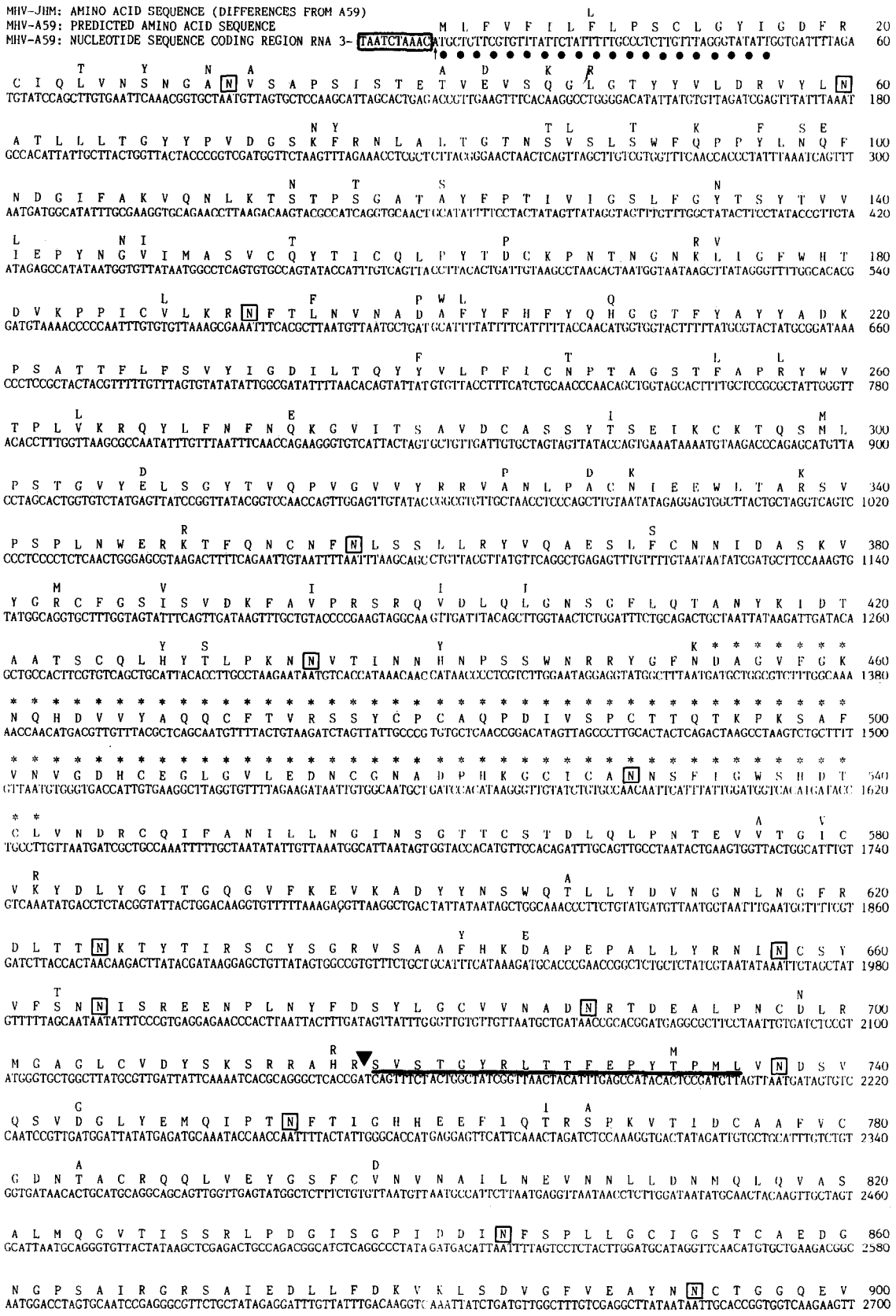

FIG. 2. Nucleotide and predicted amino acid sequence of the MHV-A59 *E2* gene. Numbering starts at the ATG codon (arrow) at position −7403 from the poly(A) tail. Dots mark the N-terminal signal sequence and the C-terminal membrane anchor. The trypsin cleavage generated N-terminal amino acid sequence of 90A as analyzed by Edman degradation is underlined. The cleavage site between 90B and 90A is indicated by an arrowhead. Potential glycosylation sites are indicated by boxed asparagine residues. The MHV–JHM amino acid sequence (Schmidt *et al.*, 1987) is printed where differences with MHV-A59 occur. Asterisks represent deletions. Intergenic homology regions are boxed.

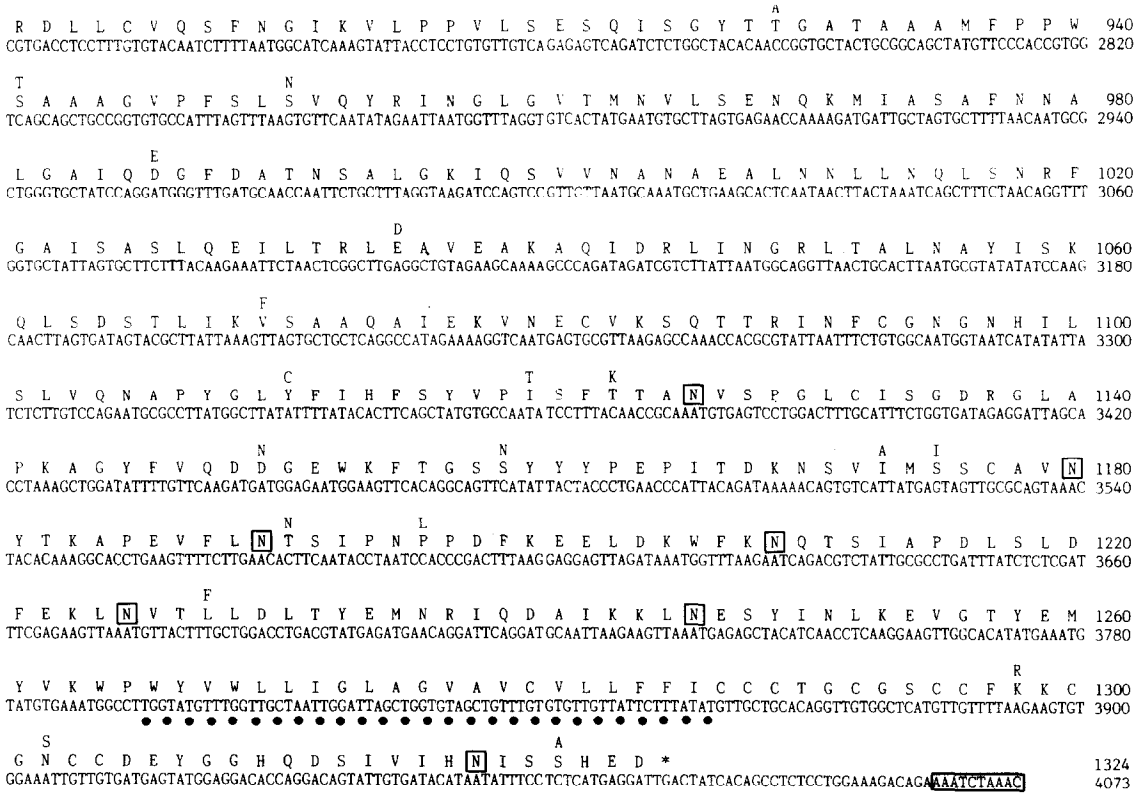

FIG. 2—*Continued.*

Nucleotide and amino acid sequence

The consensus nucleotide sequence shows an open reading frame (ORF) of 3972 nucleotides stretching from position -7403 to -3429 from the poly(A) tail. The initiation codon lies immediately adjacent to a short sequence which is fully compatible with the intergenic homology sequence 5′-(A/T)AATC(T/C)AAAC-3′ (Bredenbeek *et al.*, 1987). A similar sequence is found 28 nucleotides downstream from the end of the ORF (Fig. 2). There were no alternative ORFs longer than 60 amino acids in the unique region of mRNA 3. The large ORF is therefore identified as the coding sequence of the E2 protein.

The ORF encodes a protein of 1324 amino acids with some typical features. The N-terminal region (Fig. 2) contains a stretch of amino acids consistent with a signal sequence (Von Heyne, 1986). Another region of high hydrophobicity is found at the C-terminus and probably represents a membrane anchor. In the hydrophobicity plot this region appears as a strong symmetrical peak (Fig. 1). It starts with a series of nonpolar amino acids spanning the membrane and ends with a cluster of cysteine residues; it is followed by a number of charged residues which are probably located at the interior of the virion.

The ORF potentially codes for an apoprotein with a mol wt of 146K, which is in the range reported by several authors (see Siddell *et al.*, 1983; Repp *et al.*, 1985). Based on the assumption that Asn-X-Thr and Asn-X-Ser (X not being Pro) signals can be glycosylated and assuming that the extreme C-terminal site is located in the interior and thus unlikely to be used, we could identify 20 potential sites for N-glycosylation (Neuberger *et al.*, 1972). These are enough to add the extra 35K needed to reach the M_r of 180K required for the E2 protein. Acylation of E2 has been reported (Sturman *et al.*, 1985) but little is known about acylation signals; we could therefore not determine its contribution to the weight of the protein.

Identification of the trypsin cleavage site

The two 90K cleavage products, designated 90A and 90B, can be separated by SDS–hydroxyapatite chromatography (Ricard and Sturman, 1985). The location of the trypsin cleavage site and the relationship of 90A and 90B to the uncleaved protein was determined by comparison of the amino terminal sequence identified by Edman degradation with the sequence deduced from analysis of cDNA.

The amino terminal sequence of 90A, Ser–Val–Ser–Thr–Gly–Tyr–Arg–Leu–Thr–Thr–Phe–Glu–Pro–Tyr–Thr–Pro–Met–Leu, is identical to the sequence underlined in Fig. 2. The trypsin cleavage site can thus be positioned between residues 717 and 718 in the amino acid sequence. The 90B and 180K species appear to possess blocked amino termini as no definitive amino terminal sequences could be determined.

The identification of a signal sequence and of a membrane anchor, the determination of the amino terminal sequence of 90A, and the finding that 90A but not 90B is acylated (Sturman *et al.*, 1985) allows us to conclude that the structure of E2 is NH_2–90B–90A–COOH. The cleavage products 90A and 90B have ORFs with lengths of 606 and 717 amino acids, respectively, corresponding with coding capacities for apoproteins of 66K and 79K.

Comparison of the peplomer protein sequences of MHV strains A59 and JHM

Considerable polymorphism has been seen on the E2 glycoprotein of coronaviruses (Talbot and Buchmeier, 1985). To localize the differences we have compared the predicted amino acid sequences of the E2 protein of MHV strains A59 and JHM (Fig. 2).

The two proteins are highly conserved: there is an overall homology of 93% and 90A is more conserved than 90B (96 and 89%, respectively). However, there is a remarkable difference: starting at amino acid (aa) 454 we find a stretch of 89 aa (267 nucleotides) that is not present in the E2 sequence of JHM. To rule out the possibility that this additional sequence is the product of cDNA cloning artifacts, we isolated and sequenced several independent cDNA clones covering the region (G10, A11, and B60, Fig. 1). They all contained the additional sequence. We then synthesized an oligonucleotide (OL 53, Fig. 1) complementary to nucleotide position 1423 to 1442 in the A59 sequence and hybridized it to MHV-A59 poly(A)-selected messengers separated by electrophoresis. It is clear from Fig. 3 that the A59 "insertion" is an actual genomic feature, as it is found in mRNAs 3, 2, and 1. The extra bands in the gel can not be accounted for but have been found with other MHV probes (data not shown) and possibly represent leaderless RNAs.

The fact that the sequences of both strains can be perfectly aligned (when we exclude the additional sequence) allows a nucleotide to nucleotide comparison and the creation of a mutation table (Table 1). The sequences of the genes coding for the nucleocapsid (N) and the matrix (E1) protein were included because their products show little antigenic variation (Talbot and Buchmeier, 1985) and may thus be used as references. The ratio of nonsilent to silent (N/S) mutations

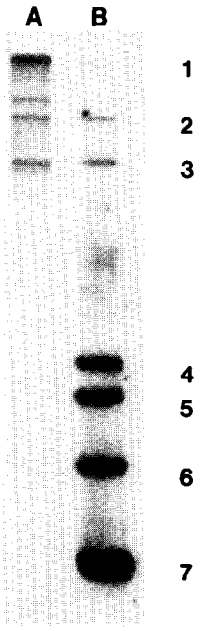

FIG. 3. Hybridization of a synthetic oligonucleotide from the MHV-A59 E2 region absent in strain JHM to the MHV-A59 messenger RNAs. Lane A, hybridization with OL 53 (Fig. 1) specific for the additional sequence of MHV-A59. Lane B, hybridization with an oligonucleotide complementary to part of the A59 leader sequence (Spaan *et al.*, 1984). RNAs are numbered according to Spaan *et al.* (1981).

can be interpreted as an indication of mutation selection. In random mutated sequences, when no selection mechanism is involved, this ratio will be about 3. Lower ratios will reflect selection against mutation whereas higher values indicate positive selection. For functional genes, however, this ratio ranges from 0.2 to 1.7, since many mutations will be lethal and therefore not found (Hewett-Emmett *et al.*, 1982). The ratios N/S for the coronaviral proteins are indeed in this range (Table 1). When we consider the *N* and *E1* genes as less susceptible to selective pressure we can understand the lower ratio found for the 90A species—about half of that of the other proteins (Table 1)—as an indication of a negative selection, i.e., supression of amino acid mutations.

DISCUSSION

The unique region of MHV-A59 mRNA 3 contains the information for the viral peplomer protein E2 (Rottier *et al.*, 1981). The nucleotide and derived amino acid sequence of the gene presented in this paper allows us to position several functional domains of the coronaviral peplomer protein in the sequence. The predicted signal sequence at the N-terminus is consistent with the finding that E2 is translated on ribosomes bound to the rough endoplasmic reticulum (Holmes *et al.*, 1984).

TABLE 1

MUTATIONAL DIFFERENCES BETWEEN MHV STRAINS A59 AND JHM

Gene	N/S	Total number of mutated	
		Nucleotides (%)	Amino acids (%)
90B	0.46	253 (13.4)	67 (10.7)
90A	0.26	122 (6.7)	25 (4.1)
E1	0.50	21 (3.1)	7 (3.1)
N	0.53	98 (7.2)	29 (6.4)

Note. Mutations were scored from aligned sequences. The ratio nonsilent to silent (N/S) mutations was calculated based upon a method described by Nei and Gojobori (1986). The scores for 90B are obtained by excluding the stretch in A59 that is absent in JHM. 90B, the N-terminal cleavage product of the peplomer protein; 90A, the C-terminal part; E1, matrix protein; N, nucleocapsid protein. Data are from Armstrong *et al.* (1984b), Skinner and Siddell (1983), Pfleiderer *et al.* (1986), and Schmidt *et al.* (1987).

The MHV trypsin cleavage site has been determined by analyzing the cleavage-generated amino terminus and localizing it in the protein sequence. Cleavage of E2 by trypsin is required for activation of the cell-fusing activity of the coronavirus (Sturman *et al.*, 1985). Cleavage activation of cell fusion is also found with the HA and F glycoproteins of myxo- and paramyxoviruses where a hydrophobic amino terminus is involved in cell fusion (Gething *et al.*, 1978; Richardson *et al.*, 1980). The amino terminal sequence of 90A shows no homology with analogous regions of HA2 and F1 of myxo- and paramyxoviruses and does not have a similar highly hydrophobic character, because it contains two charged residues (Arg and Glu). Moreover, there is no sequence homology at the amino terminus of the trypsin cleavage site between the spike proteins of MHV and infectious bronchitis virus (IBV; Binns *et al.*, 1985), although their positions are similar. Cleavage of E2 by thermolysin, which has a specificity different from that of trypsin, also activates MHV-induced cell fusion (Baker and Sturman, manuscript in preparation). This suggests that proteolytic cleavage of E2 may expose a functionally important domain that is internal rather than adjacent to the cleavage site. The sequence upstream of the cleavage site resembles the consensus sequences of trypsin cleavage sites of several other glycoproteins (Cavanagh *et al.*, 1986).

Proteolytic cleavage of E2 appears to be an important determinant of MHV pathogenesis. Investigations are in progress to identify the host- and strain-dependent differences in the processing of E2.

At its C-terminus 90A contains the highly hydrophobic potential membrane anchor of the peplomer protein. A feature of this sequence is that it starts with a stretch of eight residues: Lys–Trp–Pro–Trp–Tyr–Val–Trp–Lys which appears to be identical in coronaviruses MHV-A59, MHV-JHM, IBV-M41 (Niesters *et al.*, 1986), IBV-M42 (Binns *et al.*, 1985), feline infectious peritonitis virus (FIPV; R. J. De Groot *et al.*, manuscript in preparation), and transmissible gastroenteritis virus (TGEV; Jacobs *et al.*, manuscript in preparation). This sequence apparently represents a structural signal associated with membrane anchoring. Both E2 cleavage products in virions have an apparent mol wt of 90K as determined by SDS–PAGE (Sturman *et al.*, 1985) but the ORFs of 90A and 90B differ in length and coding capacity. In comparison in MHV-JHM the E2 cleavage products are also of an equal apparent mol wt of 98K (Siddell *et al.*, 1981), yet in this strain the lengths of the 90A and 90B ORFs are similar. Even if we take into consideration the inaccuracy of electrophoretic size estimation due to different SDS binding capacities of the cleavage products, we cannot exclude the possibility of extra or different processing of the A59 cleavage products compared to JHM.

It is not clear whether the additional sequence is deleted in JHM in the course of evolution or inserted into A59, but it is important to notice that it starts in an eight nucleotide stretch 5′-TTAATGAT-3′ (Fig. 2) that is repeated at the point where the sequences of both strains are in step again. This repeat is possibly involved in the creation of the genetic difference between A59 and JHM.

Apparently the 90B part of the peplomer protein can undergo radical changes without losing its function. This is also reflected by the fact that 90B shows the highest relative number of mutations. In contrast 90A is less mutated—but more important—shows a much lower ratio of nonsilent to silent mutations. This indicates a selection against sequence changes. De Groot *et al.* (1987) compared the peplomer protein sequences of coronaviruses from three different antigenic clusters and found that the C-terminal parts were conserved whereas the N-terminal parts were not. They demonstrated that the C-terminal sequence contained sequence patterns that could explain the typical elongated form of the coronaviral spike. The negative selection in 90A may therefore reflect preservation of structural features.

The fact that the ratio of nonsilent to silent mutations in 90B is comparable to that in the nucleocapsid and *E1* gene suggests that there is no stronger positive selection mechanism—favoring escape mutations—in this part of the protein. Talbot and Buchmeier (1985) tested a panel of neutralizing monoclonal antibodies to MHV-JHM E2 on strain A59 and demonstrated that two conformation-dependent antigenic determinants were not shared by JHM and A59 whereas a third conformation-independent determinant was found on

both strains. From our data we suggest that the conformation-dependent epitopes are on the more variable 90B part; the SDS-stable site is probably situated on the structurally important and higher conserved 90A part of the MHV peplomer protein. Experiments are in progress to localize these epitopes in the predicted amino acid sequence.

ACKNOWLEDGMENTS

The authors thank Dr. J. A. Lenstra for technical advice, Dr. A. Maagdenberg for help in computer analysis, C. Ricard for excellent technical assistance, Donna Atherton of the Rockefeller University Protein Sequencing Facility for the protein sequence analyses, Dr. G. Weeda for initial sequence analysis and Dr. Diana Ruyzendaal for oligonucleotide synthesis. This work was supported in part by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO), by Grant GM 31698 from the National Institutes of Health, and by funds provided by the U.S. Army Research Office for the purchase of equipment.

REFERENCES

ARMSTRONG, J., NIEMANN, H., SMEEKENS, S., ROTTIER, P., and WARREN, G. (1984a). Sequence and topology of a model intracellular membrane protein, E1 glycoprotein, from a coronavirus. *Nature (London)* **308,** 751–752.

ARMSTRONG, J., SMEEKENS, S., SPAAN, W., ROTTIER, P., and VAN DER ZEIJST, B. (1984b). Cloning and sequencing the nucleocapsid and E1 genes of coronavirus. *In* "Molecular Biology and Pathogenesis of Coronaviruses" (P. J. M. Rottier, B. A. M. Van der Zeijst, W. J. M. Spaan, and M. C. Horzinek, Eds.), *Adv. Exp. Med. Biol.* **173,** 155–162. Plenum, New York/London.

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

BIRNBOIM, H. C., and DOLY, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7,** 1513–1523.

BREDENBEEK, P. J., CHARITE, J., NOTEN, J. F. H., LUYTJES, W., HORZINEK, M. C., VAN DER ZEIJST, B. A. M., and SPAAN, W. J. M. (1987). Sequences involved in the replication of coronaviruses. *Adv. Exp. Med. Biol.,* Plenum, New York/London, *in press.*

BUCHMEIER, M. J., LEWICKI, H. A., TALBOT P. J., and KNOBLER, R. L. (1984). Murine hepatitis virus-4 (strain JHM) induced neurologic disease is modulated *in vivo* by monoclonal antibody. *Virology* **132,** 261–270.

CAVANAGH, D., DAVIS, P.J., PAPPIN, D. J. C., BINNS, M. W., BOURSNELL, M. E. G., and 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 Res.* **4,** 133–143.

COLLINS, A. R., KNOBLER, R. L., POWELL, H., and 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.

DALZIEL, R. G., LAMPERT, P. W., TALBOT, P. J., and BUCHMEIER, M. J. (1986). Site specific alteration of murine hepatitis virus type-4 (MHV-4) peplomer glycoprotein E2 results in reduced neurovirulence. *J. Virol.* **59,** 463–471.

DE GROOT, R. J., LENSTRA, J. A., LUYTJES, W., NIESTERS, H. G. M., HORZINEK, M. C., VAN DER ZEIJST, B. A. M., and SPAAN, W. J. M. (1987). Sequence and structure of the coronavirus peplomer protein. *Adv. Exp. Med. Biol.,* Plenum, New York/London, *in press.*

DENTE, L., CESARENI, G., and CORTESE, R. (1983). pEMBL: A new family of single stranded plasmids. *Nucleic Acids Res.* **11,** 1645–1655.

DOWLING, P. C., GIORGI, C., ROUX, L., DETHLEFSEN, L. A., GALANTOWICZ, M. E., BLUMBERG, B. M., and KOLAKOFSKY, D. (1983). Molecular cloning of the 3′-proximal third of Sendai virus genome. *Proc. Natl. Acad. Sci. USA* **80,** 5213–5216.

EISENBERG, D., SCHWARZ, E., KOMAROMY, M., and WALL, R. (1984). Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol.* **179,** 125–142.

FEINBERG, A. P., and VOGELSTEIN, B. (1983). DNA probes by random priming with Klenow synthesis. *Anal. Biochem.* **132,** 6–13.

FLEMING, J. O., STOHLMAN, S. A., HARMON, R. C., LAI, M. M. C., FRELINGER, J. A., and WEINER, L. P. (1983). Antigenic relationships of murine coronaviruses: Analysis using monoclonal antibodies to JHM (MHV-4) virus. *Virology* **131,** 296–307.

FLEMING, J. O., TROUSDALE, M. D., EL-ZAATARI, F. A. K., STOHLMAN, S. A., and WEINER, L. P. (1986). Pathogenicity of antigenic variants of murine coronavirus JHM selected with monoclonal antibodies. *J. Virol.* **58,** 869–875.

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

GETHING, M. J., WHITE, J. M., and WATERFIELD, M. D. (1978). Purification of the fusion protein of Sendai virus: Analysis of the NH2-terminal sequence generated during precursor activation. *Proc. Natl. Acad. Sci. USA* **75,** 2737–2740.

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

HANAHAN, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166,** 557–580.

HENDERSON, L. E., OROSZLAN, S., and KONIGSBERG, W. (1979). A micromethod for complete removal of dodecyl sulfate from proteins by ion-pair extraction. *Anal. Biochem.* **93,** 153–157.

HEWETT-EMMETT, D., VENTA, P. J., and TASHIAN, R. E. (1982). Features of gene organization and expression that are providing unique insights into molecular evolution and systematics. *In* "Macromolecular Sequences in Systematics and Evolutionary Biology" (M. Goodman, Ed.), pp 357–405, Plenum, New York/London.

HOLMES, K. V., FRANA, M. F., ROBBINS, S. G., and STURMAN, L. S. (1984). Coronavirus maturation. *In* "Molecular Biology and Pathogenesis of Coronaviruses" (P. J. M. Rottier, B. A. M. Van der Zeijst, W. J. M. Spaan, and M. C. Horzinek, Eds.), *Adv. Exp. Med. Biol.* **173,** 37–52. Plenum, New York/London.

LEHRACH, H., DIAMOND, D., WOZNEY, J. M., and BOEDTKER, H. (1977). RNA molecular weight determinations by gel electrophoresis under denaturing conditions, a critical reexamination. *Biochemistry* **16,** 4743–4750.

LIPMAN, D. J., and PEARSON, W. R. (1985). Rapid and sensitive protein similarity searches. *Science* **227,** 1435–1441.

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

MEINKOTH, J., and WAHL, G. (1984). Hybridization of nucleic acids immobilized on solid supports. *Anal. Biochem.* **138,** 267–284.

MESSING, J. (1983). New M13 vectors for cloning. *In* "Methods in Enzymology" (R. Wu, L. Grossman, and K. Moldave, Eds.), Vol. 101, pp 20–79. Academic Press, New York.

NEI, M., and GOJOBORI, T. (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* **3,** 418–426.

NEUBERGER, A., GOTTSCHALK, A., MARSHALL, R. O., and SPIRO, R. G. (1972). Carbohydrate–peptide linkages in glycoproteins and methods for their elucidation. *In* "The Glycoproteins: Their Composition, Structure and Function" (A. Gottschalk, Ed.), pp. 450–490. Elsevier, Amsterdam.

NIEMAN, H., and KLENK, H. D. (1981). Coronavirus glycoprotein E1, a new type of viral glycoprotein. *J. Mol. Biol.* **153,** 993–1010.

NIESTERS, H. G. M., LENSTRA, J. A., SPAAN, W. J. M., ZIJDERVELD, A. J., BLEUMINK-PLUYM, N. M. C., HONG, F., VAN SCHARRENBURG, G. J. M., HORZINEK, M. C., and VAN DER ZEIJST, B. A. M. (1986). The peplomer protein sequence of the M41 strain of coronavirus IBV and its comparison with Beaudette strains. *Virus Res.* **5,** 253–263.

PEACOCK, S. L., MCIVER, C. M., and MONOHAN, J. J. (1981). Transformation of *E. coli* using homopolymer linked plasmid chimeras. *Biochim. Biophys. Acta* **655,** 243–250.

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

REPP, R., TAMURA, T., BOSCHEK, C. B., WEGE, H., SCHWARZ, R. T., and NIEMANN, H. (1985). The effects of processing inhibitors of N-linked oligosaccharides on the intracellular migration of glycoprotein E2 of mouse hepatitis virus and the maturation of coronavirus particles. *J. Biol. Chem.* **260,** 15,873–15,879.

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

RICHARDSON, C. D., SCHEID, A., and CHOPPIN, P. W. (1980). Specific inhibition of paramyxovirus and myxovirus replication by oligopeptides with amino acid sequences similar to those at the N-termini of the F1 or HA2 viral polypeptides. *Virology* **105,** 205–222.

ROBERTS, C., and WILSON, G. L. (1985). DNA probes by random priming with Klenow synthesis. *Focus* **7-3,** 16.

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

SANGER, F., NICKLEN, S., and COULSON, A. R. (1977). DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.

SCHMIDT, I., SKINNER, M., and SIDDELL, S. (1987). Nucleotide sequence of the gene encoding the surface projection glycoprotein of coronavirus MHV-JHM. *J. Gen. Virol.* **68,** 47–56.

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

SIDDELL, S., WEGE, H., BARTHEL, A., and TER MEULEN, V. (1981). Coronavirus JHM: Intracellular protein synthesis. *J. Gen. Virol.* **53,** 145–155.

SIDDELL, S., WEGE, H., and TER MEULEN, V. (1983). The biology of coronaviruses. *J. Gen. Virol.* **64,** 761–776.

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

SMITH, H. O., and BIRNSTIEL, M. L. (1976). A simple method for DNA restriction site mapping. *Nucleic Acids Res.* **3,** 2387–2395.

SPAAN, W. J. M., ROTTIER, P. J. M., HORZINEK, M. C., and VAN DER ZEIJST, B. A. M. (1981). Isolation and identification of virus-specific mRNAs in cells infected with mouse hepatitis virus (MHV-A59). *Virology* **108,** 424–434.

SPAAN, W., DELIUS, H., SKINNER, M. A., ARMSTRONG, J., ROTTIER, P., SMEEKENS, S., SIDDELL, S. G., and VAN DER ZEIJST, B. (1984). Transcription strategy of coronaviruses: Fusion of non-contiguous sequences during mRNA synthesis. *In* "Molecular Biology and Pathogenesis of Coronaviruses" (P. J. M. Rottier, B. A. M. Van der Zeijst, W. J. M. Spaan, and M. C. Horzinek, Eds.), *Adv. Exp. Med. Biol.* **173,** 173–196. Plenum, New York/London.

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

STADEN, R. (1986). The current status and portability of our sequence handling software. *Nucleic Acids Res.* **14,** 217–233.

STURMAN, L. S. (1977). Characterization of a coronavirus. I. Structural proteins: Effects of preparative conditions on the migration of protein in polyacrylamide gels. *Virology* **77,** 637–649.

STURMAN, L. S., and HOLMES, K. V. (1977). The molecular biology of coronaviruses. II. Glycoproteins of the viral envelope: Tryptic peptide analysis. *Virology* **77,** 650–660.

STURMAN, L. S., and HOLMES, K. V. (1983). The molecular biology of coronaviruses. *Adv. Virus Res.* **28,** 35–112.

STURMAN, L. S., and HOLMES, K. V. (1985). The novel glycoproteins of coronaviruses. *Trends Bio. Sci.* **10,** 17–20.

STURMAN, L. S., HOLMES, K. V., and BEHNKE, J. (1980). Isolation of coronavirus envelope glycoproteins and interaction with the viral nucleocapsid. *J. Virol.* **33,** 449–462.

STURMAN, L. S., RICARD, C. S., and 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. *J. Virol.* **56,** 904–911.

TALBOT, P. J., and BUCHMEIER, M. J. (1985). Antigenic variation among murine coronaviruses: Evidence for polymorphism on the peplomer glycoprotein, E2. *Virus Res.* **2,** 317–328.

VON HEIJNE, G. (1986). A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.* **14,** 4683–4690.

WEGE, H., SIDDELL, S., and TER MEULEN, V. (1982). The biology and pathogenesis of coronaviruses. *Curr. Top. Microbiol. Immunol.* **99,** 165–200.