



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Sequence Analysis Reveals Extensive Polymorphism and Evidence of Deletions within the E2 Glycoprotein Gene of Several Strains of Murine Hepatitis Virus

SUEZANNE E. PARKER, THOMAS M. GALLAGHER, AND MICHAEL J. BUCHMEIER¹

Department of Immunology, Scripps Clinic and Research Foundation, La Jolla, California 92037

Received June 13, 1989; accepted August 25, 1989

Direct RNA sequence analysis of the E2 gene of wild-type MHV-4 and of neutralization resistant, neuroattenuated variants has identified a polymorphic region with respect to deletions. These variants had large deletions of 142 to 159 amino acids mapping to a localized region in the amino-terminal domain of the peplomer glycoprotein. The nucleotide sequence of the E2 gene for wild-type strain MHV-4 was found to be very similar to that of MHV-JHM but had an insertion of 423 nucleotides resulting in the addition of a stretch of 141 unique amino acids in the amino-terminal domain of E2. We propose that deletions reflect a major source of heterogeneity in the E2 protein of MHV. © 1989 Academic Press, Inc.

INTRODUCTION

Mouse hepatitis virus type 4 (MHV-4, strain JHM) is a neurotropic member of the Coronaviridae (Weiner, 1973; Lampert *et al.*, 1973; Haspel *et al.*, 1978; Wege *et al.*, 1982). The virus causes acute, subacute, and chronic infections of the CNS in mice and rats with symptoms ranging from a fatal encephalitis to a chronic demyelinating disease. Intracerebral inoculation of wild-type MHV-4 in susceptible mice usually results in a fatal encephalomyelitis characterized by widespread destruction of CNS neurons with accompanying demyelination. Recovery from the acute infection is rare, and the few mice that survive develop a central nervous system disease characterized by focal demyelinating lesions followed by remyelination (Knobler *et al.*, 1982; Lampert *et al.*, 1973). The demyelination has been attributed to virus infection and destruction of oligodendrocytes (Lampert *et al.*, 1973; Haspel *et al.*, 1978; Dubois-Dalq *et al.*, 1982).

Neuroattenuated variants of MHV-4 developed in this laboratory and others offer a model system for the study of the pathogenesis of MHV infections. The fatal encephalitis usually associated with wild-type MHV-4 is not seen following infection with ts mutants (Haspel *et al.*, 1978), spontaneously arising mutants (Stohlman *et al.*, 1982), or variants selected on the basis of resistance to neutralization by monoclonal antibodies (MAb) against the E2 glycoprotein of the virus (Dalziel *et al.*, 1986; Fleming *et al.*, 1986; Wege *et al.*, 1988). Infection of susceptible mice with these variants results in chronic demyelinating disease with episodes of remy-

elination and demyelination with little neuronal involvement.

Characterization of the genetic alteration(s) responsible for the neuroattenuation would map the determinants which contribute to the neurovirulence of a MHV-4 infection. In the case of ts mutants, it would be difficult to localize these changes as the attenuated virus may contain multiple mutations which may or may not contribute to the attenuated phenotype. Variants that have escaped neutralization by MAbs directed against the peplomer glycoprotein (E2) offer an advantage for mapping of neurovirulence because mutations in E2 are specifically selected. We have previously reported the isolation and characterization of two such groups of variants selected for the ability to resist neutralization by MAb directed against epitopes E2B (MAb 5A13.5) and E2C (MAb 4B11.6), respectively. In contrast to the fatal encephalitis induced by the wild-type MHV-4, both variants were found to be attenuated and caused chronic demyelinating disease with demyelination and remyelination occurring up to 65 days postinfection (Dalziel *et al.*, 1986).

E2 is initially synthesized as a large 180-kDa peripheral membrane glycoprotein (Siddell *et al.*, 1982) which is post-translationally cleaved. Cleavage yields two nonidentical subunits of approximately 90 kDa (Sturman *et al.*, 1985), consisting of an amino-terminal domain (S₁ in the terminology of Cavanagh (1983) or 90B in the terminology of Sturman *et al.* (1985)) and a carboxy-terminal domain (S₂ or 90A). To map these E2 mutations we have sequenced the E2 genes of wild-type and antibody-resistant variants. We have found by direct RNA sequencing that the MAb-resistant variants of MHV-4 have large deletions in the 5' coding region

¹ To whom requests for reprints should be addressed at IMM5.

of the E2 gene. The deletions we describe result in a loss of 142 to 159 amino acids localized to the amino-terminal domain (S₁ or 90B) of the E2 protein. In contrast, the carboxy-terminal domain was found to be largely conserved.

MATERIALS AND METHODS

Virus and cells

MHV-4 was originally obtained from L. P. Weiner and is routinely propagated on SAC- cells as previously described for MHV-A59 (Sturman *et al.*, 1980). A working stock of MHV-4 for RNA sequence analysis was prepared by minimal amplification of an isolate plaque purified in 1986. SAC- cells were grown to confluency in 175-cm² flasks in Dulbecco's modified Eagle's medium (DMEM) supplemented with 8.5% heat-inactivated calf serum. Cells were inoculated with 0.01 to 0.1 PFU per cell and after a one-hour adsorption at 37° the inoculum was removed and replaced with growth medium. Virus was harvested when signs of cytopathic effect were evident (75–90% syncytium formation; 16–18 hr postinfection) by three cycles of freeze-thawing (–70°), followed by a low-speed centrifugation at 300 *g* to remove cellular debris.

Isolation and characterization of antibody-resistant variants V5A13.1 and V4B11.3 have been previously described (Dalziel *et al.*, 1986). For purposes of identification these will be referred to hereafter as V5A13.1(86) and V4B11.3(86). Briefly, a plaque-purified stock of MHV-4 (1986) was subjected to two rounds of neutralization with MAb 5A13.5 or 4B11.6. Virus that escaped neutralization was plaque purified two times. Working stocks of variant virus for RNA sequence analysis were prepared from plaque-purified stocks by propagation on SAC- cells and screened for maintenance of the antibody-resistant phenotype (Dalziel *et al.*, 1986). A second variant, V5A13(88), was selected with MAb 5A13.5 from a stock of MHV-4 that had been passaged several times *in vitro*. A working stock of V5A13(88) was prepared for RNA sequence analysis by propagation on SAC- cells as described. This virus was not subsequently plaque purified, thus the RNA sequence of this variant reflects the consensus sequence of the population of antibody-resistant virus. A derivative of MHV-JHM maintained in Japan, JHM-X (Makino *et al.*, 1984), was kindly provided by Dr. M. M. C. Lai (University of Southern California).

RNA sequence analysis

Total infected cytoplasmic RNA was prepared by the guanidine isothiocyanate (GTC) extraction method of Chirgwin *et al.* (1979) from SAC- infected cells. Ethanol-precipitated RNA was used directly for sequence

analysis. The viral RNA sequence was determined by a modification of the dideoxy primer extension method of Hamlyn *et al.* (1981).

Initially, synthetic oligonucleotide primers of 20 bases complementary to the cDNA sequence of MHV-JHM mRNA 3 (Schmidt *et al.*, 1987) were synthesized on an Applied Biosystems 380-A oligonucleotide synthesizer and purified by electrophoresis on 16% acrylamide, 8 *M* urea, TBE [0.089 *M* Tris–HCl (pH 8.3), 0.089 *M* boric acid, 0.002 *M* EDTA] sequencing gels (Sanger *et al.*, 1977). As sequence was obtained for the E2 of MHV-4, oligonucleotide primers were made which were complementary to the 5' end of the extending MHV-4 E2 sequence. Oligonucleotide primers were end labeled by incubating 50 pM of purified primer with 0.5 mCi of [γ -³²P]ATP (7000 Ci/mM, ICN Radiochemicals) and 3 U of T4 polynucleotide kinase (Promega) in a final volume of 8 μ l of kinase buffer [50 mM Tris–HCl (pH 7.5), 10 mM MgCl₂, 5 mM DTT, 0.1 mM spermidine, 0.1 mM EDTA] for 30 min at 37°. Radiolabeled primers were extracted with phenol/chloroform, ethanol precipitated, and resuspended in water at a concentration of approximately 5.0×10^3 cpm/pM (3.0×10^5 cpm/ μ l). Template RNA (75–100 μ g in water) was added to 3.5 pM of radiolabeled primer in a final volume of 12.5 μ l of annealing buffer [50 mM Tris–HCl (pH 8.3), 60 mM NaCl, 10 mM DTT, 1 mM EDTA] and incubated for 3 min at 90°, followed by 10 min at 65°, then for 20 min at 45°. The resulting hybrids (2 μ l) were added to four separate sequencing reactions each containing all four deoxynucleotides (0.22 mM in each dNTP), the appropriate dideoxynucleotide (0.04 mM), and reverse transcriptase (1U, AMV Life Sciences) in a final volume of 6 μ l of reverse transcriptase buffer [50 mM Tris–HCl (pH 8.3), 50 mM KCl, 10 mM MgCl₂]. Sequencing reaction mixtures were incubated for 30 min at 45° and then 2 μ l of a deoxynucleotide chase mix (0.1 mM in each dNTP and 0.5 U reverse transcriptase in reverse transcriptase buffer) was added to each reaction and incubation was continued for an additional 15 min at 45°. Dye mixture (10 mM EDTA, 0.1% bromophenol blue, and 0.1% xylene cyanol FF in deionized formamide) was added to each reaction, and the samples were heated for 5 min at 100°. The reaction mixture (3–4 μ l) was loaded onto each lane of a sequencing gel (6% polyacrylamide, 7 *M* urea in TBE, 48.5 \times 20 \times 0.4 mm) (Sanger *et al.*, 1977) and the samples were electrophoresed at 40 W for 2.5–3.5 hr. Gels were fixed for 15 min by immersion in 10% acetic acid and 10% methanol, dried, and subsequently autoradiographed on Kodak XAR or XRP film. By this method it was possible to sequence 100–150 nucleotide bases per oligonucleotide primer. Sequence data were compiled and

analyzed using the University of Wisconsin Genetics Computer Group Sequence Analysis software package (Devereux *et al.*, 1984).

RESULTS

Sequence of the MHV-4 E2 gene

The cDNA sequence of the E2 gene of MHV-JHM (Schmidt *et al.*, 1987) was used as a starting point to design synthetic oligonucleotides at 150-base intervals which served as primers for the direct dideoxy sequence analysis of viral RNA. From primer extension analysis 4149 nucleotides were sequenced comprising a single open reading frame (ORF) of 4131 nucleotides with the potential to encode a protein of 1376 amino acids (Fig. 1). The sequence starts with the 5' TAATCTAAAC intergenic homology sequence and is immediately followed by the initiation codon. A similar intergenic sequence has been found to precede the 5' initiation codons of several MHV RNAs (Skinner *et al.*, 1985; Skinner and Siddel, 1985), including the E2 gene of MHV-JHM (Schmidt *et al.*, 1987) and MHV-A59 (Luytjes *et al.*, 1987). The nucleotide sequence of the E2 gene for MHV-4 revealed one major difference from the sequence previously reported for MHV-JHM (Schmidt *et al.*, 1987). In the MHV-4 gene, there is an insertion of 423 nucleotides relative to MHV-JHM starting at position 1359 and extending through to nucleotide 1781. The inserted nucleotide sequence potentially encodes 141 unique residues starting at amino acid 454 in the amino-terminal domain of E2. The MHV-4 E2 gene contains 6 additional nucleotide substitutions that result in 4 amino acid differences in the predicted sequence of the E2 protein as compared to MHV-JHM (Fig. 1). There is also a lysine to asparagine substitution at amino acid 453 in the E2 gene of MHV-4 at the 5' junction of the MHV-4 insertion.

Sequence analysis of the E2 gene of MHV-4 variants

Direct RNA sequencing of the viral RNA encoding the E2 gene of V5A13.1(86) and V4B11.3(86) demonstrated that both variants had large deletions localized within the 5' half of the gene (Fig. 2). The sequence of the entire E2 gene of V5A13.1(86) was determined and a large deletion of 426 nucleotides starting at nucleotide 1298 and extending through to nucleotide 1723 of

the parental MHV-4 was found. This sequence predicts a truncated E2 protein containing 1234 amino acids compared with the 1376 amino acids as was found in MHV-4. The deleted domain corresponds to amino acids 434–575 in the amino-terminal half of the protein. In addition, as a consequence of the deletion, a lysine → asparagine substitution occurred at position 433 (the 5' boundary of the deletion) in the variant E2. Having localized a deletion in V5A13.1(86) to the amino-terminal half of E2, the coding region for this domain (nucleotides 1–2333) was subsequently sequenced for V4B11.3(86). This E2 gene was found to have deleted nucleotides 1285 through 1761 of the MHV-4 parental sequence which results in an ORF with the potential to encode 1217 amino acids representing a loss of residues 429–586 relative to wild-type. V4B11.3(86) also had a glutamine → lysine substitution at amino acid 404 as a result of a single base substitution in the V4B11.3(86) gene. However, this substitution was found in only one of three 4B11.6 variants which were sequenced in this region and thus does not correlate with resistance to neutralization.

Given that the deletions in two of the variants were found to localize to a specific region in the 5' portion of the gene, we concentrated subsequent analysis in this region. The RNA from eight additional V5A13(86) variants and 12 additional V4B11(86) variants was sequenced from nucleotides 1228 through 1788 (relative to the MHV-4 sequence). All of the variants selected in 1986 by MAb 5A13.5 were found to contain the identical 426-nucleotide deletion found in V5A13.1(86), and all 1986 4B11.6 variants were found to contain the identical 477-nucleotide deletion found in V4B11.3(86). We were surprised to find that these 20 additional variants fall into only two selection groups typified by V5A13.1(86) and V4B11.3(86). This likely reflects the selection process by which the original variants were isolated. The V5A13(86) and V4B11(86) variants presumably represent independent plaque isolates of two deletion mutants which were present in the 1986 stock of wild-type virus that were selected in the first round of neutralization. To determine whether these two deletions represented reproducible as opposed to chance events we selected an additional 5A13.5 variant from a multiply passaged stock derived from the 1986 plaque-purified stock of MHV-4 termed

Fig. 1. Nucleotide and predicted amino acid sequence of the MHV-4 E2 gene. The numbering starts at the ATG codon. Open circles indicate the N-terminal signal sequence (calculated according to the algorithm of Von Heijne (1986) as described by Fazakerley and Ross (1989)) and closed circles indicate the hydrophobic carboxy-terminal transmembrane domain. Potential glycosylation sites are indicated by boxed asparagine residues (Neuberger *et al.*, 1972). The intergenic homology sequence TAATCTAAAC is boxed. The putative proteolytic cleavage site between the amino-terminal and carboxy-terminal domain is indicated by an arrowhead. The MHV-JHM amino acid sequence (Schmidt *et al.*, 1987) is indicated where differences occur. The sequence unique to MHV-4 and deleted from MHV-JHM is marked with asterisks.

TAATCTAAAC

1
M L F V F I L L L P S C L G Y I G D F R C I Q T V N Y N G N **N** A S A P S I S T E 40
ATGCTGTTTCGCTTTATTTACTATTACCCCTCTGTTTAGGGTATATTGGTGATTTAGATGTATCCAGACCGTGAATTATAACGGCAATAATGCTTCTGCGCCTAGCATTAGCACCAGAA 120
O O O O O O O O R O O O O O O O
A V D V S K G L G T Y Y V L D R V Y L **N** A T L L L T G Y Y P V D G S N Y R N L A 80
GCAATCGATGTTCCAAAGGCTGGGCACTTACTATGTTTAGATCGTGTACTTAAATGCCACGTTATTGCTTACTGTTATTATCCTGTGGACGGTCCAAATATCGGAATCTCGCG 240
L T G T N T L S L T W F K P P P L S E F N D G I F A K V Q N L K T N T P T G A T 120
CTTACAGGCACTAATACCTTAAGCCTTACGTGGTTAAACCAACCCCTTCTAAGTGAGTTAATGATGGTATATTGCTAAGGTCAGAACCTCAAGACAAATACGCCAACAGGTGCAACC 360
S Y F P T I V I G S L F G **N** T S Y T V V L E P Y N N I I M A S V C T Y T I C Q L 160
TCATATTTCCCACTATAGTTATAGGTAGTTGTTTGGTAACACTTCCATACCGTAGTTTATAGACCATATAATAATATTATAATGGCTTCTGTTTGTACATATACCATTGTCAATTA 480
P Y T P C K P N T N G N R V I G F W H T D V K P P I C L L K R **N** F T F N V N A P 200
CCTTACACCCCTGTAAGCCTAATACCAATGGTAATCGTGTATTGGATTTTGGCACACAGATGTCAAACCGCGATTGTCTTTAAAGCGTAATTTTACGTTAATGTTAATGCCCTT 600
W L Y F H P Y Q Q G G T F Y A Y Y A D K P S A T T F L F S V Y I G D I L T Q Y F 240
TGGCTTATTTCCATTTTATCAGCAGGGTGGTACTTTTTATGCGTACTATGCGGATAAACCTTCCGCTACTACGTTTGTGTTAGTGTGTATATTGGCGACATTTAAACACAGATTTT 720
A
V L P P I C T P T A G S T L L P L Y W V T P L L K R Q Y L F N F N E K G V I T S 280
GTGTTACCTTTTATTGTACTCCAACAGTGGTAGCACTTTACTGCGCTCTATTGGGTTACACCTTACTTAAAGCGCAATATTGTTTAAATTTAATGAAAAGGGTGTCACTACTAGT 840
A V D C A S S Y I S E I K C K T Q S L L P S T G V Y D L S G Y T V Q P V G V V Y 320
GCTGTGATTGGCCAGCAGCTACATTAGTAAATAAAATGAAGACCCAAAGTCTCTTACCGAGTACTGGTGTCTATGATCTATCCGGTTACAGGTTCCAACTGTTGGAGTTGTGTAC 960
R R V P N L P D C K I E E W L T A K S V P S P L N W E R R T F Q N C N F N L S S 360
CGGCGTGTCCTAACCTACCTGATTGTAATAAGAGGAATGGCTCACTGCTAAATCTGTGCCGTCACTCTCAATTTGGGAGCGTAGGACTTTCCAAAATGTAATTTAATTTAAGCAGC 1080
L L R Y V Q A E S L S C N N I D A S K V Y G M C F G S V S V D K F A I P R S R Q 400
CTGCTACGTTATGTCAGGCTGAGTCTTGTGCGTAAATAATTTGATGCGTCAAAGTGTATGGTATGCTTGGTAGTGTCTCAGTTGATAAGTTTGTATCCCCGAAGCCGCTCAA 1200
I D L Q I G N S G F L Q T A N Y K I D T A A T S C Q L Y Y S L P K N **N** V T I N N 440
ATTGATTTACAAATGGCAACTCCGGATTTTGGCAACCGCTAATTAAGATTGATACCGCTGCCACATCATGTCAGCTGATTACAGTCTTCTAAGAATAATGTTACATAAATAAC 1320
K
Y N P S S W N R R Y G F N D A G V F G K S K H D V A Y A Q Q C F T V R P S Y C P 480
TATAACCCCTCGTCTTGGAAATAGGAGGTATGTTTAAATGATGCTGGTGTGTTTGGCAAAAGTAAACATGATGTTGCTACCGCCAGCAATGTTTACTGTGCGACCTAGCTATTGTCCG 1440

C A Q P D I V S A C T S Q T K P M S A Y C P T G T I H R E C S L W N G P H L R S 520
TGTCACAACCGGACATAGTTAGCGCTTGCACTAGTCAGACCAAACCATGTCTGCTTATTGCCCCACAGGCACAATTCATCGTGAGTGTCTCTTGGAAATGGGCCCATTTGCGCTCG 1560

A R V G S G T Y T C E C T C K P N P F D T Y D L R C G Q I K T I V N V G D H C E 560
GCACGTGTAGGTTCCGGCACGTACAGTGTGAGTGCACTTGTAAACCAATCCATTGATACGTATGATCTCCGCTGTGGGCAAATTAATACTATTGTTAATGTGGGCGATCATTTGTGAA 1680

G L G V L E D K C G N S D P H K G C S C A **N** D S F I G W S H D T C L V N D R C Q 600
GGTCTGGGTGTTTAGAAGATAAATGTGGCAATAGCGATCCACATAAGGGCTGTTCTGTGCCAATGATTCTTTATCGGATGGTCACATGACACTGTTTAGTAAATGATCGCTGCCAA 1800
I F A N I L L N G I N S G T T C S T D L Q L P N T E V A T G V C V R Y D L Y G I 640
ATTTTGTCTAACATATTGTTAAATGGCATTAAATAGTGGGACTACGTGTTCCACAGATTACAATTGCCTAATACTGAAGTGGCCACTGGCGTTTGGTGCAGATAGACCTCTATGGTATT 1920
T G Q G V F K E V K A D Y Y N S W Q A L L Y D V N G N L N G F R D L T T **N** K T Y 680
ACTGGTCAAGGTGTTTAAAGAGGTCAAGGCTGACTATTATAATAGCTGGCAGGCCCTATTATATGATGTTAATGGTAACTTAAACGGGTTCCGTGACCTTACCCTAACAAGACTTAT 2040
T I R S C Y S G R V S A A Y H K E A P E P A L L Y R N I **N** C S Y V F T N **N** I S R 720
ACGATAAGGAGCTGTTATAGTGGCCGTGTTTCTGTGCATATCATAAAGAAGCACCCGAACCGGCTCTGCTCTATCGTAATATAAATTTAGTATTGTTTACTAATAATATTTCCCCT 2160
N
E E N P L N Y F D S Y L G C V V N A D **N** R T D E A L P N C D L R M G A G L C V D 760
GAGGAAAACCCCTTAACTATTTGATAGTTAATTTGGGTTGTTGTTAATGCTGATAACCGCACGGATGAGGCGCTCTCAATTTGGATCTCCGTATGGGTGCTGGCACTATGCGTAGAT 2280
Y S K S R R A R R S V S T G Y R L T T F E P Y M P M L V **N** D S V Q S V G G L Y E 800
TATTCAAAGTCAAGCAGAGCCCGGATCAGTTTCTACTGGCTATCGAATTAACCACTCGAGCCATACATGCGGATGTTAGTCAATGATAGCGTTCAATCCGTAGGTGGATTATATGAG 2400
M Q I P T **N** F T I G H H E E F I Q I R A P K V T I D C A A P V C G D N A A C R Q 840
ATGCAATACCAACCAATTTACTATTGGTCACTCATGAGGAATTCATCCAGATAAGGGCTCCAAAGTGACTATAGATTGTGCTGCATTGTTTGGTGATAACCGTGCATGCAGACAG 2520
Q L V E Y G S F C D N V N A I L N E V N N L L D N M Q L Q V A S A L M Q G V T I 880
CAGTTGGTGGATGCGCTTTTTGTGATAATGTTAATGCCATTCTAATGAGGTTAATAACCTCTGGATAAATATGCAATTACAAGTTGTAGTGCATTAATGCAGGGTGTACTATA 2640

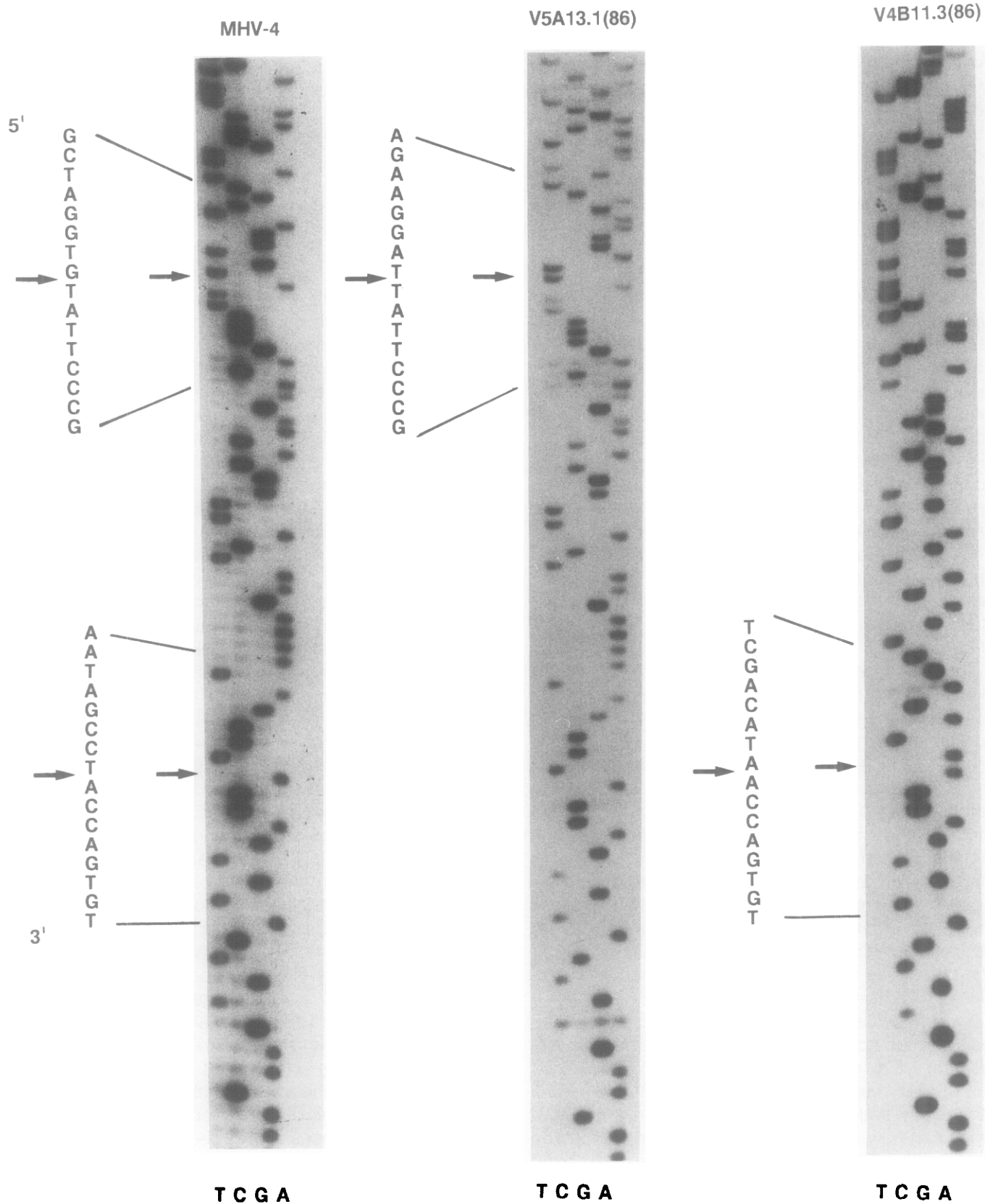


Fig. 2. Nucleotide sequence analysis from the relevant regions of sequencing gels comparing wild-type MHV-4, V5A13.1(86), and V4B11.3(86). The sequence shown for MHV-4 represents nucleotides 1699–1777. Sequences were determined using the chain terminating dideoxy sequencing procedure as described. Lanes are labeled T, C, G, and A for the appropriate chain terminating dideoxynucleotide. Arrows indicate the position of the deletions with the nucleotide sequence flanking the deletions shown.

tive to the MHV-4 sequence) for several strains of MHV. MHV-4 and MHV-A59 contain a run of relatively conserved amino acids (90% homology) beginning at position 454 (454–500 and 553–594 relative to the MHV-4 sequence) that are not found in JHM. In addition, MHV-4 contains an additional stretch of 52 unique amino acids starting at position 501 that are not found in A59.

DISCUSSION

In this study we have analyzed the E2 gene of wild-type MHV-4 and of three independently selected MAb-resistant variants V4B11.3(86), V5A13.1(86), and V5A13(88). By direct RNA sequencing all three variants were found to have deletions of 426 to 477 nucleotides localized to the 5' coding region of the gene (nucleotides 1285 through 1761 of the parental E2 sequence; Fig. 1). The variant E2 genes have the potential to encode truncated E2 proteins with deletions in the amino-terminal region ranging from 142 amino acids in the case of V5A13.1(86) to 159 amino acids in the case of V4B11.3(86) (Table 1). These results confirm and extend previous analyses of viral RNA and protein (Gallagher *et al.*, submitted for publication) which demonstrate deletions in the amino-terminal cleavage product of E2.

The sequence of the E2 gene of MHV-4 was nearly identical to that of MHV-JHM reported by Schmidt *et al.* (1987) with the notable exception that MHV-4 has an additional 423 nucleotides in the coding region of E2 (Fig. 1). The E2 gene of our MHV-4 strain potentially encodes a larger protein than JHM (1376 amino acids as compared to the 1235 amino acids of MHV-JHM) with a unique stretch of 141 amino acids in the amino-terminal domain of E2. These data clearly demonstrate that MHV-JHM and MHV-4, which were previously considered interchangeable, are in fact distinct viruses. Given the otherwise near identity between the E2 genes of the two viruses and the relative ease of selection of deletion mutants, it might be speculated that MHV-JHM arose as a deletion mutant of MHV-4.

It was previously observed that the selection for MAb resistance to epitope E2B also resulted in resistance to epitope E2C and the reciprocal was true for selection of MAb resistance to epitope E2C (Dalziel *et al.*, 1986). This acquisition of resistance to multiple epitopes can now most easily be reconciled by the sequence data showing that the antibody-resistant variants have large overlapping deletions which presumably result in the loss of both epitopes. While we cannot ignore the possibility that these deletions affect protein folding and conformation at distant sites, it appears clear that the deleted sequence contains determinants critical for the structure of both epitopes.

TABLE 1

LOCATION OF DELETIONS IN E2

Variant	Nucleotide deletion	Amino acid deletion
V5A13.1 (86)	426 (1298–1723)	142 (434–575)
V4B11.3 (86)	477 (1285–1761)	159 (429–586)
V5A13 (88)	447 (1307–1753)	149 (436–585)
JHM-X	458 (1336–1794)	153 (446–598)

It is interesting to note that although V5A13.1(86) and V5A13(88) were selected by the same MAb, these variants contain unique deletions (Table 1). V5A13.1(86) was isolated from a minimally passaged stock of MHV-4 and the E2 sequence of this variant represents the sequence of an individual plaque-purified isolate. In contrast, V5A13(88) was isolated from a stock of MHV-4 which had been passed multiple times *in vitro*. As V5A13(88) was not plaque purified, the E2 sequence for this variant represents the predominant sequence of the population of antibody-resistant variants present in the multiply passaged stock of MHV-4.

The deletions we have observed all map to a hydrophilic domain in the amino-terminal half (aa 429–598) of E2 (Fig. 3) indicating that neutralization epitopes and determinants associated with neurovirulence map to this region of the protein. This is in contrast to a previous study where neutralization epitopes and neurovirulence determinants were mapped to the carboxy-terminal domain of E2 using recombinants between MHV-JHM and a ts mutant of MHV-A59 (Makino *et al.*, 1987). Recombinants which had lost the MHV-JHM neutralizing epitopes defined by MAbs J7.2 and J2.2 were analyzed by T₁ oligonucleotide fingerprinting and the crossover site in these recombinants was localized to a region 1.5 kb from the 3' end of the E2 gene; thus, it was concluded that these epitopes mapped to the carboxy-terminal domain of the protein. However, MAbs J7.2 and J2.2 react with our wild-type MHV-4 but fail to react with any of our antibody-resistant variants (Dalziel *et al.*, 1986), nor do they react with JHM-X (Taguchi and Fleming, 1989). Our variants and JHM-X all have large deletions in the amino-terminal half of E2. These data strongly suggest that MAbs J7.2 and J2.2 detect epitopes in the amino-terminal and not the carboxy-terminal domain of E2. This apparent discrepancy in mapping of epitopes recognized by MAbs J7.2 and J2.2 could most easily be reconciled if both recombination and deletion occurred in the cells coinfecting with MHV-A59 and MHV-JHM. Alternatively, the A59-JHM recombinant viruses may have arisen as a result of a recombination between MHV-A59 and E2 deletion variants preexisting in the MHV-JHM parent population.

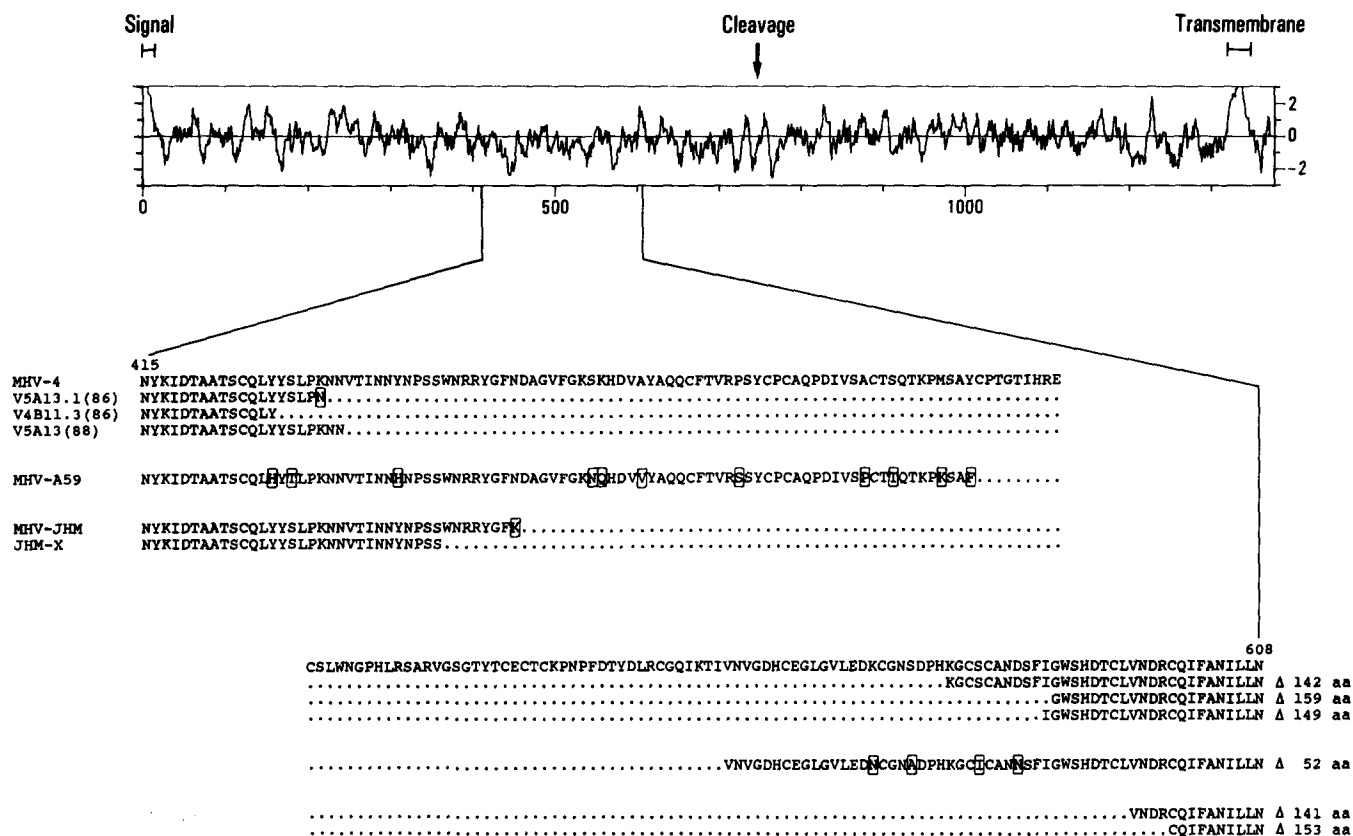


Fig. 3. Hydropathicity plot of the MHV-4 E2 glycoprotein according to the analysis of Kyte and Doolittle (1982) and localization of the deletions in the amino-terminal domain of MHV-4 E2. The vertical scale is the average hydropathicity (+2 to -2) index for each residue over a window of 9 amino acids. Hydrophobic sequences appear above the midline and hydrophilic sequences appear below the midline. Below is an alignment of amino acid residues 415 through 608 from MHV-4 E2 with the comparable amino acid sequences from MHV-JHM (Schmidt *et al.*, 1987), MHV-A59 (Luytjes *et al.*, 1987), JHM-X, and MHV-4 variants V8A13.1(86), V4B11.3(86), and V5A13(88), all determined in this laboratory. Dots indicate deletions and boxed amino acids indicate sequence changes from MHV-4 E2. The signal sequence, putative cleavage site, and transmembrane domains of E2 are indicated.

We have found that the E2 protein of MHV-4 is very heterogeneous with respect to deletions in the amino-terminal domain (Fig. 3). Heterogeneity in terms of the size of the E2 gene has also been reported to result from both the *in vivo* and *in vitro* passage of MHV-JHM. Taguchi *et al.* have reported isolation of variants of JHM from the brains of rats and from primary rat neuronal cultures with a larger E2 (ca. 15 kDa) than that of the parental JHM wild-type (Taguchi *et al.*, 1985, 1986). In contrast, Morris *et al.* (1989) have isolated a variant from the spinal cord of MHV-JHM infected rats which had a smaller E2 than the wild-type strain used to inoculate the animals. It will be of interest to determine whether these differences represent insertions or deletions of the polymorphic sequences we have described.

The mechanism of generation of diversity in E2 remains an open question. It has been suggested that genetic diversity among murine coronaviruses arises as a result of high-frequency RNA-RNA recombination between heterologous strains (Lai *et al.*, 1985; Makino

et al., 1986; Keck *et al.*, 1988) and even between coronaviruses and other enveloped RNA viruses (Luytjes *et al.*, 1988). Recombination is proposed to occur as a consequence of the discontinuous, nonprocessive mechanism of transcription involved in MHV replication. According to this model, RNA synthesis frequently stops at sites of secondary structure on template RNA releasing the transcriptional complex which then rejoins the template RNA and resumes transcription (Baric *et al.*, 1987). This process involving a jumping polymerase may also result in deletions whereby the disassociated transcriptional complex reinitiates and resumes transcription at a distant site on the same template or on a sister strand. If this transfer involves RNA base pairing, one might expect to find homologous sequences flanking the deletion sites that would provide site specificity for the nascent RNA strand. However, we were unable to find any homologous or consensus sequences flanking the deletions in the E2 gene (Table 2). In addition, a computer analysis of the nucleotide sequence of E2 did not reveal any unique

TABLE 2
NUCLEOTIDE SEQUENCE OF MHV-4 IN REGIONS FLANKING DELETIONS IN E2

Virus	5' Flanking sequence	3' Flanking sequence
V5A13.1 (86)	AGTCTTCCTAAGAATAATGT . . .	AGCGATCCACATAAGGGCTG
V5A13 (88)	AAGAATAATGTTACCATAAA . . .	AATGATTCTTTTATCGGATG
V4B11.3 (86)	TCAGCTGTATTACAGCTTCC . . .	TTTTATCGGATGGTCCACATG
MHV-JHM	ATGGTTTTAATGATGCTGGT . . .	ACACTTGTTTAGTAAATGAT
JHM-X	CCCCTCGTCTTGAATAGGT . . .	AAATGATCGCTGCAAAATTTT

^a 5' Endpoint of deletion.

^b 3' Endpoint of deletion.

sites of secondary structure in the polymorphic region of the gene. Equivalent regions of predicted secondary structure were found throughout the entire gene (data not shown). Thus perhaps deletions occur randomly throughout the genome during replication but deletions in most regions may be lethal. It appears clear that the amino-terminal domain of E2 can accommodate large deletions while retaining functions necessary for growth *in vitro*. In contrast, viable recombination in other genes or in other sites on E2 may require site-specific recombination to occur without the accompanying loss of genetic information.

We have localized an area in the amino-terminal region of the E2 glycoprotein of MHV-4 which is highly heterogeneous and subject to deletion mutation. Determinants localized within this region, while dispensable for *in vitro* growth and replication, play an important role in determining the outcome of an *in vivo* infection. This could be reflected at the level of receptor binding and hence determine the tissue tropism of the virus or at the level of events occurring during the replication and maturation of the virus and hence determine the ability of the virus to amplify and spread. We are presently pursuing the functional significance of this domain of the MHV-4 E2 in order to evaluate its role in tissue tropism and pathogenesis *in vivo*.

ACKNOWLEDGMENTS

The authors thank Maria Salvato and Elaine Shimomaye for advice on RNA sequencing and Joseph O'Neill for technical assistance. We are indebted to Jim Johnston for manuscript preparation. This work was supported by NIH Grants AI 25913 and NS 12428. S.E.P. was the recipient of a fellowship from the National Multiple Sclerosis Society. T.M.G. was supported by NIH Training Grant 5-T32-AG00080. This is Publication 5931-IMM from the Department of Immunology, Research Institute of Scripps Clinic.

REFERENCES

- BARIC, R. S., CHIEN, C.-K., STOHLMAN, S. A., and LAI, M. M. C. (1987). Analysis of intracellular small RNAs of mouse hepatitis virus: Evidence for discontinuous transcription. *Virology* **156**, 342-354.
- CAVANAGH, D. (1983). Coronavirus IBV: Structural characterization of the spike protein. *J. Gen. Virol.* **64**, 2577-2583.
- CHIRGWIN, J. M., PRZYBYLA, A. E., MACDONALD, R. J., and RUTTER, W. J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* **18**, 5294-5299.
- DALZIEL, R. G., LAMPERT, P. W., TALBOT, P. J., and BUCHMEIER, M. J. (1986). Site-specific alteration of murine hepatitis virus type 4 peplomer glycoprotein E2 results in reduced neurovirulence. *J. Virol.* **59**, 463-471.
- DEVEREUX, J., HABERLI, P., and SMITHIES, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**, 387-395.
- DUBOIS-DALCO, M. E., DOLLER, E. W., HASPEL, M. V., and HOLMES, K. W. (1982). Cell tropism and expression of mouse hepatitis virus (MHV) in mouse spinal cord cultures. *Virology* **119**, 317-331.
- FAZAKERLEY, J. F., and ROSS, A. M. (1989). Computer analysis suggests a role for signal sequences in processing polyproteins of enveloped RNA viruses and as a mechanism of viral fusion. *Virus Genes* **2**, 219-235.
- 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.
- HAMLIN, P. H., GAIT, M. J., and MILSTEIN, C. (1981). Complete sequence of an immunoglobulin mRNA using specific priming and the dideoxynucleotide method of RNA sequencing. *Nucleic Acids Res.* **9**, 4485-4494.
- HASPEL, M. V., LAMPERT, P. W., and OLDSTONE, M. B. A. (1978). Temperature sensitive mutants of mouse hepatitis virus produce a high incidence of demyelination. *Proc. Natl. Acad. Sci. USA* **75**, 4033-4036.
- KECK, J. G., MATSUSHIMA, G. K., MAKINO, S., FLEMING, J. O., VANNIER, D. M., STOHLMAN, S. A., and LAI, M. M. C. (1988). *In vivo* RNA-

- RNA recombination of coronavirus in mouse brain. *J. Virol.* **62**, 1810–1813.
- KNOBLER, R. L., TUNISON, L. A., LAMPERT, P. W., and OLDSTONE, M. B. A. (1982). Selected mutants of mouse hepatitis virus type 4 (JHM strain) induce different CNS diseases. *Amer. J. Pathol.* **109**, 157–168.
- KYTE, J., and DOOLITTLE, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–132.
- LAI, M. M. C., BARIC, R. S., MAKINO, S., KECK, J. G., EGBERT, J., LEIBOWITZ, J. L., and STOHLMAN, S. A. (1985). Recombination between nonsegmented RNA genomes of murine coronaviruses. *J. Virol.* **56**, 449–456.
- LAMPERT, P. W., SIMS, J. K., and KNIAZEFF, A. J. (1973). Mechanism of demyelination in JHM virus encephalomyelitis: Electron microscopic studies. *Acta Neuropathol.* **24**, 76–85.
- LUYTJES, W., STURMAN, L. S., BREDEBEEK, P. J., CHARITE, J., VAN DER ZEIJST, B. A. M., HORZINEK, M. C., and SPAAN, W. J. M. (1987). Primary structure of the glycoprotein E2 of coronavirus MHV-A59 and identification of the trypsin cleavage site. *Virology* **161**, 479–487.
- LUYTJES, W., BREDEBEEK, P. J., NOTEN, A. F. H., HORZINEK, M. C., and SPAAN, W. J. M. (1988). Sequence of mouse hepatitis virus A59 mRNA: Indications for RNA recombination between coronaviruses and influenza C virus. *Virology* **166**, 415–422.
- MAKINO, S., TAGUCHI, F., and FUJIWARA, K. (1984). Defective interfering particles of mouse hepatitis virus. *Virology* **133**, 9–17.
- MAKINO, S., KECK, J. G., STOHLMAN, S. A., and LAI, M. M. C. (1986). High frequency of RNA recombination of murine coronaviruses. *J. Virol.* **57**, 729–737.
- MAKINO, S., FLEMING, J. O., KECK, J. G., STOHLMAN, S. A., and LAI, M. M. C. (1987). RNA recombination of coronaviruses: Localization of neutralizing epitopes and neuropathogenic determinants on the carboxy terminus of peplomers. *Proc. Natl. Acad. Sci. USA* **84**, 6567–6571.
- MORRIS, V. L., TIESZER, C., MACKINNON, J., and PERCY, D. (1989). Characterization of coronavirus JHM variants isolated from Wistar Furth rats with a viral induced demyelinating disease. *Virology* **169**, 127–136.
- 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.
- 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 SIDDEL, S. (1987). Nucleotide sequence of the gene encoding the surface projection glycoprotein of coronavirus MHV-JHM. *J. Gen. Virol.* **68**, 47–56.
- SIDDEL, S., WEGE, H., and TER MEULEN, V. (1982). The structure and replication of coronaviruses. *Curr. Topics Microbiol. Immunol.* **99**, 131–163.
- SKINNER, M. A., EBNER, D., and SIDDEL, S. G. (1985). Coronavirus MHV-JHM mRNA 5 has a sequence arrangement which potentially allows a translation of a second downstream open reading frame. *J. Gen. Virol.* **66**, 581–592.
- SKINNER, M. A., and SIDDEL, S. G. (1985). Coding sequence of coronavirus MHV-JHM mRNA 4. *J. Gen. Virol.* **66**, 593–596.
- STOHLMAN, S. A., BRAYTON, P. R., FLEMING, J. O., WEINER, L. P., and LAI, M. M. C. (1982). Murine coronaviruses: Isolation and characterization of two plaque morphology variants of the JHM neurotropic strain. *J. Gen. Virol.* **63**, 265–275.
- 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.
- TAGUCHI, F., SIDDEL, S. G., WEGE, H., and TER MEULEN, V. (1985). Characterization of a variant virus selected in rat brains after infection by coronavirus mouse hepatitis virus JHM. *J. Virol.* **54**, 429–435.
- TAGUCHI, F., MASSA, P. T., and TER MEULEN, V. C. (1986). Characterization of a variant virus isolated from neural cell culture after infection of mouse coronavirus JHMV. *Virology* **155**, 267–270.
- TAGUCHI, F., and FLEMING, J. O. (1989). Comparison of six different murine coronavirus JHM variants by monoclonal antibodies against the E2 glycoprotein. *Virology* **169**, 233–235.
- 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., SIDDEL, S., and TER MEULEN, V. (1982). The biology and pathogenesis of coronaviruses. *Curr. Topics Microbiol. Immunol.* **99**, 165–200.
- WEGE, H., WINTER, J., and MEYERMANN, R. (1988). The peplomer protein E2 of coronavirus JHM as a determinant of neurovirulence: Definition of critical epitopes by variant analysis. *J. Gen. Virol.* **69**, 87–98.
- WEINER, L. P. (1973). Pathogenesis of demyelination induced by a mouse hepatitis virus (JHM virus). *Arch. Neurol.* **28**, 298–303.