Arch Virol (1995) 140: 571–579

Archives of Virology

© Springer-Verlag 1995
Printed in Austria

Sequence analysis of the nucleoprotein genes of three enterotropic strains of murine coronavirus

Brief Report

F. R. Homberger

Institute of Laboratory Animal Science, University of Zurich, Zurich, Switzerland

Accepted November 15, 1994

Summary. The nucleotide sequences of the nucleoprotein genes of three enterotropic strains of the murine coronavirus mouse hepatitis virus (MHV-Y, MHV-RI and DVIM) were determined and compared with previously reported sequences of three polytropic (respiratory) strains (MHV-A59, MHV-JHM and MHV-S). Greater than 92% homology was found among the six strains by pair-wise comparison at the nucleotide level. The genes encoded proteins of 451 to 455 residues and the deduced amino acid sequences were more than 91% homologous. A unique deletion of twelve nucleotides was found at the carboxy terminus of MHV-Y and a three nucleotide deletion was found in MHV-RI, which corresponded to the one previously reported in MHV-A59 and MHV-S. Two internal open reading frames were found within the coding region of the nucleoprotein, the smaller one was specific for the enterotropic strains. It could potentially encode a truncated version of the hypothetical protein described for MHV-A59 and MHV-S. Sequence relationship of the N gene showed no correlation with tissue tropism and no sequence or even single amino acid change unique to either tropism group was found. This indicates that the nucleoprotein of MHV probably has no part in the determination of the primary tissue tropism of an MHV strain. The role of the potential internal protein warrants further investigation.

*

Mouse hepatitis virus is an important pathogen of the laboratory mouse and the most common virus detected in contemporary mouse colonies throughout the world [17, 21]. Its implications on biomedical research are well documented [2, 11]. MHV is an RNA virus belonging to the coronavirus group. It is highly mutable and consists of a number of antigenically different strains which vary markedly in virulence and tissue tropism. According to their tropism the strains

can be divided into two main groups with different disease patterns [2]. One biotype, the respiratory or polytropic strains, primarily infects the mucosa of the upper respiratory tract, where initial virus replication takes place. Virus disseminates by viremia to different organs depending on strain virulence and host susceptibility factors. Neurotropic strains may spread along the olfactory nerve directly to the brain. Polytropic strains cause hepatitis, encephalitis as well as many other forms of disease [9]. A wasting disease caused by multisystemic infection is found in immunodeficient animals [5]. On the other hand, infections with the enterotropic strains are usually restricted to the intestinal mucosa, the site of their primary replication. They rarely disseminate to other organs even in immunodeficient animals [7], and are mainly associated with diarrhoea in infant mice [8].

Aside from pathology many more differences exist between the polytropic and the enterotropic strains of MHV. Enterotropic strains seem to be more contagious among mice. Both biotypes cause an acute infection in immunocompetent mice, but enterotropic MHV persists for up to 30 days whereas mice clear the polytropic strains within two weeks [8]. Active immunity against polytropic virus is less cross-reactive and its protective effect wanes faster than against enterotropic MHV [4, 16]. Passive immunity is mediated differently, protection against enterotropic strains depends on intraluminal antibody in the intestine of the pup, whereas serum antibody is critical for immunity to polytropic MHV [12, 14].

The mechanism that determines the tissue tropism of an MHV strain is still poorly understood. The antigenic relationships among MHV isolates are merely of academic interest, since they yield no information on tropism or virulence. The polytropic strain MHV-S, for example, is antigenically more closely related to an enterotropic strain MHV-Y than to another polytropic isolate MHV-JHM [3]. The determinant of the primary tissue tropism has not yet been conclusively assigned to any morphological structure of the virion. In this study, the nucleoprotein (N) gene of three enterotropic strains of MHV were sequenced and compared with published sequence information of three polytropic strains, in an attempt to identify a unique structure in either tropism group that may be a determinant of tropism.

MHV-Y was originally isolated in NCTC-1469 cells during a natural outbreak of typhlocolitis in infant mice [6] and MHV-RI from a nude mouse by passage in CMT-93 cells [7]. Both virus strains were propagated in infant CD-1 mice. The MHV strain DVIM was provided by Kathryn V. Holmes (Uniformed Services University of the Health Sciences, Bethesda, MD) and passaged in NCTC-1469 cells. Infant CD-1 mice were inoculated with the three MHV strains and the intestine was collected two days later at the peak of infection [16]. Homogenates (10% w/v) of the intestines were prepared in tissue culture medium and used as virus stocks.

In a first step RNA was extracted from intestinal homogenates pretreated with an RNase inhibitor (40 units of RNasin [Promega, Madison, WI] per 100 µl). Samples were treated with sodium dodecyl sulphate (SDS) and phenol-

extracted. RNA was precipitated in 70% ethanol and resuspended in TE buffer (10 mM Tris, 1 mM EDTA). Using the isolated RNA as template, first-strand cDNA was synthesised with avian reverse transcriptase (Promega) at pH 8.4 and 42 °C. The reaction was primed by oligo $(dT)_{12-18}$ (Pharmacia, Piscataway, NJ) [15]. The complete N genes of the three virus strains were amplified by polymerase chain reaction (PCR) as described [15]. One of the primers used in the PCR reaction (5′-ACGTACCCTCTCAACTC-3′) was derived from the leader sequence common to all MHV mRNAs, synthesised based on sequence information from MHV-JHM [24]. The anti-sense primer (5′-ATGACAGCA-AGACATCC-3′), located approximately 100 nucleotides downstream of the termination codon of the N gene in the 3′ noncoding region, was synthesised based on sequence information obtained from MHV-A59 [23] (Microsynth AG, Windisch, Switzerland). Prior to the blunt end ligation into the *Sma* I site of pUC 18 by T4 DNA ligase, the 3′ overhangs of the PCR fragments were removed by Klenow fragment of DNA polymerase I and the DNA fragments were phosphorylated by T4 polynucleotide kinase (SureClone, Pharmacia) [13]. For each virus strain three different clones were constructed, each clone originating from a separate PCR reaction.

Plasmid DNA was grown in *E. coli* (DH5α; Life Technologies, Gaithersburg, MD) and extracted with phenol-chloroform. Denaturing of DNA and degradation of RNA was then carried out at the same time in an alkaline solution by boiling [31]. The precipitation of the DNA with ethanol was followed immediately by the annealing of the sequencing primer. Primers used for the sequencing reaction were two universal primers T7 and -40 (Life Technologies) corresponding to the sequence of pUC 18 on both sides of the multiple cloning site, as well as a number of internal primers synthesised on the basis of MHV-A59 sequence information. Sequencing reaction was performed according to the Sanger dideoxy-mediated chain-termination method [26] using Sequenase 2.0 (United States Biochemical, Cleveland, OH) and ^{35}S-labelled dATP (Amersham, Arlington Heights, IL) according to the manufacturers protocol.

Three independent clones of the N gene from each of the three virus strains were sequenced at least once in both directions. When the sequencing results of the three clones of each virus strain were compared among themselves, five differences were found among the MHV-Y clones, four differences among the MHV-RI clones and two among the DVIM clones (data not shown). These apparent differences in nucleotide sequence were most likely due to errors introduced by either the MHV RNA-dependent RNA polymerase during the replication in vivo, by the avian reverse transcriptase during cDNA synthesis or by the *taq* DNA polymerase during PCR. Each aberrant nucleotide was found only on one clone whereas the corresponding nucleotides on the other two clones were identical. The consensus results of the N gene sequences have been deposited in the GenBank nucleotide sequence bank and were used in the following analysis.

The nucleotide sequences of the N genes of three enterotropic strains of MHV, MHV-Y, MHV-RI and DVIM, were compared with the previously

reported sequences of three polytropic prototype strains MHV-S, MHV-A59 [1, 23] and MHV-JHM [24, 27]. A 12-base deletion was found in MHV-Y but not in any of the other strains, either entero- or polytropic, at the position corresponding to nt 1339–1350 of the MHV-JHM sequence. In addition to this a three nucleotide deletion (nt 1164–1166) was found in MHV-RI corresponding to a similar deletion reported in MHV-S and MHV-A59. This deletion was not found in MHV-Y and DVIM nor has it been described for MHV-JHM. All other variations between the strains were only nucleotide substitutions, no frameshifts or nonsense mutations were observed. In the following homology comparison, all deletions were counted as single mutations regardless of their size. The nucleotide sequences of the N genes of all six MHV strains were more than 92% homologous (Table 1). The greatest homology was found between MHV-Y and MHV-RI with 47 nucleotide differences (96.6%). The three enterotropic strains MHV-Y, MHV-RI and DVIM together with MHV-A59 seem to form a closely related group with 96.6–94.6% homology (47–74 nucleotide differences). The N gene sequence of MHV-JHM differs more significantly from these four strains (92.9- 92.1% homology, 99-108 nucleotide differences) whereas MHV-S seems to hold a midway position, being as close to MHV-Y (68 nucleotide differences, 95.1% homology) as to MHV-JHM (59 nucleotide differences, 95.7% homology).

The N genes of the six MHV strains presented in this work encode proteins of 451 to 455 residues (Fig. 1). The comparison of the deduced amino acid sequences reveals greater than 91% homology between the six proteins (Table 1). Again the three enterotropic strains and MHV-A59 are most closely related (13–20 amino acid differences, 97.1–95.6% homology) and MHV-JHM is most distant to these four (30–37 amino acid differences, 93.3–91.9% homology). MHV-S however, on the amino acid level seems to be closer to MHV-Y (19 amino acid differences, 95.8% homology) than to MHV-JHM (26 amino acid differences, 94.3% homology).

Table 1. Nucleotide and amino acid sequence homology between the nucleoprotein genes and the deduced nucleoproteins of three enterotropic strains (MHV-Y, MHV-RI and DVIM) and two respiratory strains (MHV-A59 and MHV-JHM) of MHV. Deletions were counted as single mutations

aa nt	MHV-Y	MHV-RI	DVIM	MHV-A59	MHV-S	MHV-JHM
MHV-Y		47 (96.6)[a]	61 (95.5)	53 (96.2)	68 (95.1)	99 (92.9)
MHV-RI	16 (96.4)[b]		74 (94.6)	66 (95.3)	80 (94.3)	108 (92.1)
DVIM	18 (96.0)	20 (95.6)		68 (95.1)	93 (93.2)	102 (92.5)
MHV-A59	20 (95.6)	14 (96.9)	13 (97.1)		73 (94.8)	100 (92.6)
MHV-S	19 (95.8)	17 (96.3)	25 (94.4)	18 (96.0)		59 (95.7)
MHV-JHM	37 (91.9)	31 (93.0)	32 (92.9)	30 (93.3)	26 (94.3)	

[a] Number of different nucleotides (homology %)
[b] Number of different amino acids (homology %)

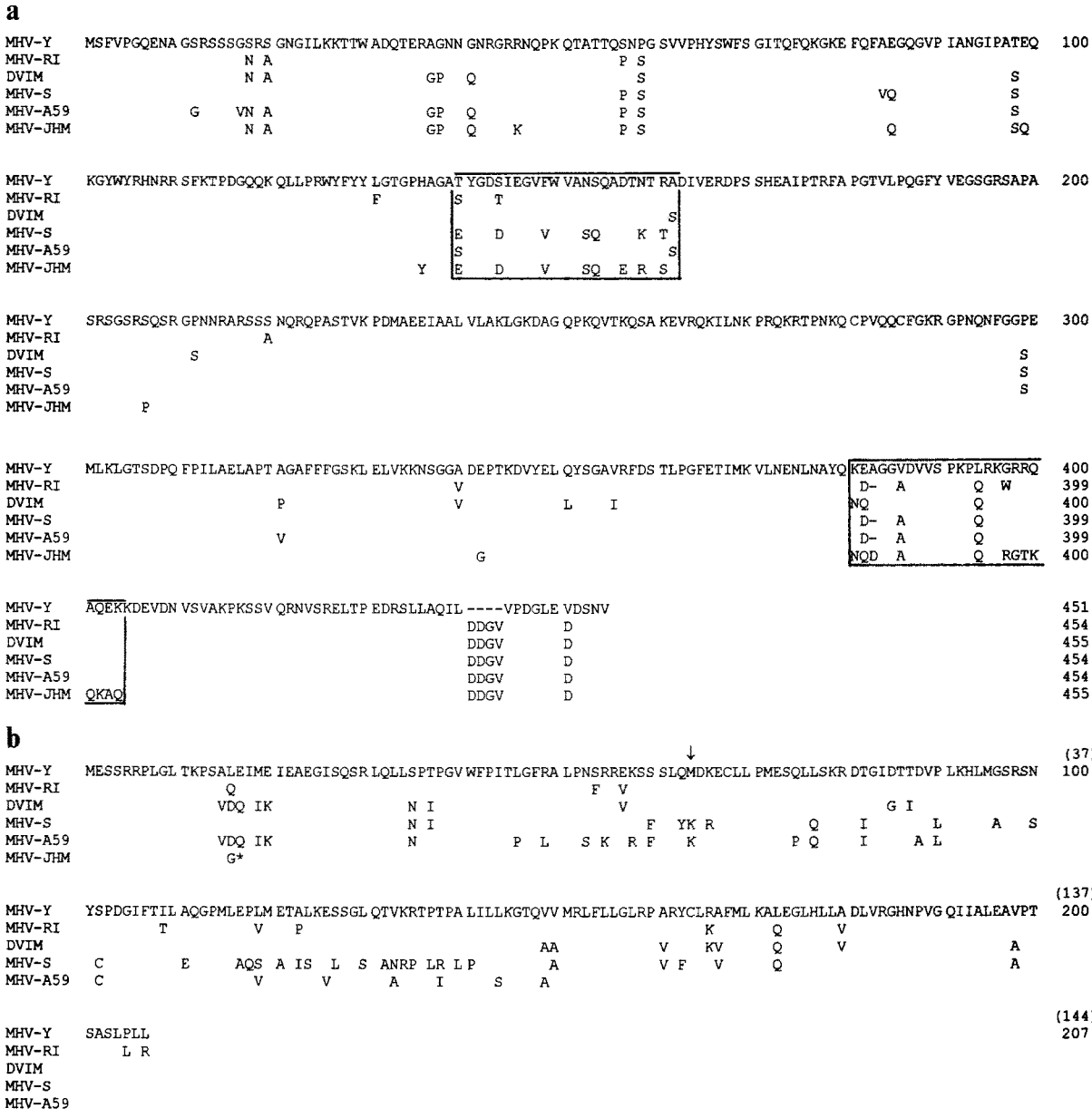

Fig. 1. Comparison of the deduced amino acid sequences of all open reading frames (orf) found within the N gene region of three enterotropic strains (MHV-Y, MHV-RI, DVIM) and three respiratory strains (MHV-A59, MHV-S, MHV-JHM) of MHV. **a** The main orf encoding the nucleoproteins. The variable regions are boxed, deletions are represented by dashes. **b** The hypothetical products of the two internal orfs. The asterix represents the stop codon in MHV-JHM. The arrow marks the beginning of the specific internal orf of the enterotropic strains, its amino acids are numbered in parenthesis

MHV contains three to four structural proteins [9]. The N protein together with the genomic RNA forms the nucleocapsid, a long helical structure in the centre of the virion. It is surrounded by the envelope which contains two or three glycoproteins. One of them, the membrane (M) protein, serves as a bridge

between the nucleocapsid and the viral envelope while another, the spike (S) protein, forms the characteristic petal-shaped spikes that give the *Coronaviridae* their name [9]. Some MHV strains possess a third glycosylated protein, the hemagglutinin/esterase (HE). DVIM is the only strain known with a functional HE-protein and while JHM possesses a truncated inactive version it has not been found in MHV-A59, -S, -Y or -RI [19, 30, 32]. Of these proteins HE most likely plays no part in the determination of the tropism since it is not present in all strains. It has previously been shown that the most conserved structural protein M shows no sequence unique to enterotropic strains [13]. A comparison of the most variable of the coronavirus structural proteins, the spike-protein, of MHV-Y and -RI revealed high diversity between these two enterotropic strains. They had only three single amino acids in common which could not be found in the deduced amino acid sequence of either the MHV-A59 or -JHM S protein and could therefore be considered to be unique to the enterotropic strains [19].

The N protein is antigenically well conserved among different strains of MHV [10]. The present study confirmed this and showed that the N gene sequences of the enterotropic strains lay well within the variation found among polytropic isolates. Most changes found were conservative and were located within the two variable regions previously described for the polytropic strains except for the four amino acid deletion at the carboxy terminus unique to MHV-Y. Aside from this, the N proteins of the enterotropic strains exhibited the same three-domain structure proposed for the prototype MHV [23]. The deduced proteins had an overall excess of basic residues concentrated in the central domain and an acidic carboxy terminus.

The N protein together with the genomic RNA forms the nucleocapsid, a long flexible, helical structure in the centre of the virion. The N protein associates with the genomic RNA by specific binding to the leader sequence as well as unspecific binding [25, 28]. The specific RNA binding domain of the protein is believed to be contained between amino acids 176 to 230 [22]. A monoclonal antibody binding in this region is able to block the specific interaction between the N protein and the leader sequence of the genome [29]. This RNA binding domain is highly conserved among different polytropic strains and the present study demonstrated that this homology extends to the enterotopic strains as well.

Two internal open reading frames (orf) within the N gene sequence of all three entropic MHV strains were found. The longer one with a start codon at position 83 and a stop codon at position 706 potentially encodes a protein with 207 residues. An orf of the same size was described for MHV-A59 and MHV-S [23]. This hypothetical protein has not yet been detected and seems to be non-essential for the replication of the virus, since it is truncated in MHV-JHM by a stop codon after only 16 amino acids [23]. It is therefore unlikely that this protein has any influence on the primary tropism of a strain. The second orf is in the same reading frame and starts at position 272 ending in the stop codon at position 706. This orf is found in all three enterotropic

strains examined but not in any of the prototype strains. It has, however, been described in the Nu67 strain of MHV [18]. MHV-Nu67 was isolated from a nude mouse in Japan and seems to be an enterotropic strain. The second internal orf potentially encodes a protein of 144 residues which would, however, be a truncated version of the product of the larger orf (Fig. 1). While this second orf found within the N gene is a feature specific to enterotropic strains, its potential product has not been detected yet and would be very similar to the larger hypothetical protein postulated for MHV-A59 and MHV-S [23]. Considering all these factors the role of this protein in the determination of the primary tissue tropism of enterotropic MHV seems doubtful, but must not be discarded without further investigation.

It has been described that the M gene sequences of DVIM differs somewhat from that of MHV-Y, MHV-RI and MHV-A59 [13]. This study, however, showed that the N genes of these four viruses were closely related. This seeming discrepancy might be explained by a recombination event which had occurred during the evolution of DVIM somewhere in the vicinity of the intergenic region between M and N and had created the virus from an MHV-A59-like and an unrelated parental strain. Similar recombination events between MHV strains have been observed repeatedly in vitro as well as in vivo [20] and are considered to be responsible for the great antigenic diversity of this group of viruses. It has been postulated that the N gene of MHV-S as well had resulted from a naturally occurring recombination between two strains similar to MHV-JHM and A59 [23]. Moreover, this study together with previous findings [13] show that MHV-RI may also be a result of a recombination between the two above-mentioned parental strains. The RI M and N sequence is highly homologous with A59 white its S sequence closely resembles that of JHM. The recombination event seems to have taken place somewhere within the non-structural protein genes between the S and M.

No unique sequence to enterotropic MHV could be identified in this study. The N protein does not seem to be important for the determination of the primary tissue tropism of an MHV strain. Possibly a single amino acid change like one of those identified on the S protein [19] or multiple factors located on several different genes could be responsible for the tropism of an isolate. Together with the above mentioned findings this could suggest that the evolution of the enterotropic strains was caused by point mutations rather than recombinations. This question, however, will have to be answered by using targeted recombination to produce new MHV strains carrying specific mutations. Alternatively the inclusive or exclusive involvement of the nonstructural proteins has to be considered.

Acknowledgements

This work was supported by grant #31-39728.93 from the Swiss National Science Foundation in Berne, Switzerland and by a grant from the "Stiftung fuer wissenschaftliche Forschung" of the University of Zurich. I thank Martin Schwyzer, Stephen W. Barthold and Susan R. Compton for helpful discussions. The technical assistance of Kirsten Rappold

is appreciated. The nucleotide sequence data generated in this study has been deposited in the GenBank nucleotide sequence bank (accession numbers L37758–L37760).

References

1. Armstrong J, Niemann H, Smeekens S, Rottier P, Warren G (1984) Sequence and topology of a model intracellular membrane protein, E1 glycoprotein, from a coronavirus. Nature 308: 751–752
2. Barthold SW (1986) Mouse hepatitis virus biology and epidemiology. In: Bhatt PN, Jacoby RO, Morse AC III, New AE (eds) Viral and mycoplasmal infections of laboratory rodents: effect on biomedical research. Academic Press, Orlando, pp 571–601
3. Barthold SW, Smith AL (1987) Mouse hepatitis virus strain-related patterns of tissue tropism in suckling mice. Arch Virol 81: 103–112
4. Barthold SW, Smith AL (1989) Duration of challenge immunity to coronavirus JHM in mice. Arch Virol 104: 187–196
5. Barthold SW, Smith AL (1990) Duration of mouse hepatitis virus infection: studies in immunocompetent and chemically immunosuppressed mice. Lab Anim Sci 40: 133–137
6. Barthold SW, Smith AL, Lord PFS, Bhatt PN, Jacoby RO, Main AJ (1982) Epizootic coronaviral typhlocolitis in suckling mice. Lab Anim Sci 32: 376–383
7. Barthold SW, Smith AL, Povar AL (1985) Enterotropic mouse hepatitis virus infection in nude mice. Lab Anim Sci 35: 613–618
8. Barthold SW, Beck DS, Smith AL (1993) Enterotropic coronavirus (mouse hepatitis virus) in mice: influence of host age and strain on infection and disease. Lab Anim Sci 43: 276–284
9. Compton SR, Barthold SW, Smith AL (1993) The cellular and molecular pathogenesis of coronaviruses. Lab Anim Sci 43: 15–28
10. Fleming JO, Stohlman SA, Harmon RC, Lai MMC, Frelinger JA, Weiner LP (1983) Antigenic relationships of murine coronaviruses: analysis using monoclonal antibodies to JHM (MHV-4) virus. Virology 131: 296–307
11. Hamm TE Jr, ed (1986) Complications of viral and mycoplasmal infections in rodent toxicology research and testing. Hemisphere Washington
12. Homberger FR (1992) Maternally-derived passive immunity to enterotropic mouse hepatitis virus. Arch Virol 122: 133–141
13. Homberger FR (1994) Nucleotide sequence comparison of the membrane protein genes of three enterotropic strains of mouse hepatitis virus. Virus Res 31: 49–56
14. Homberger FR, Barthold SW (1992) Passively acquired challenge immunity to enterotropic coronavirus in mice. Arch Virol 126: 35–43
15. Homberger FR, Smith AL, Barthold SW (1991) Detection of rodent coronaviruses in tissues and cell cultures by using polymerase chain reaction. J Clin Microbiol 29: 2789–2793
16. Homberger FR, Barthold SW, Smith AL (1992) Duration and strain specificity of immunity to enterotropic mouse hepatitis virus. Lab Anim Sci 42: 347–351
17. Kraft V, Meyer B (1990) Seromonitoring in small laboratory animal colonies. A five year survey: 1984–1988. Z Versuchstierk 33: 29–35
18. Kunita S, Tedara E, Goto K, Kagiyama N (1992) Sequence analysis and molecular detection of mouse hepatitis virus using the polymerase chain reaction. Lab Anim Sci 42: 593–598
19. Kunita S, Zhang L, Homberger FR, Compton SR (1995) Characterization of two Enterotropic Murine Coronavirus Strain. Virus Res (in press)
20. Lai MMC (1992) RNA recombination in animal and plant viruses. Microbiol Rev 56: 61–79

21. Lindsey JR (1986) Prevalence of viral and mycoplasmal infections in laboratory rodents. In: Bhatt PN, Jacoby RO, Morse AC III, New AE (eds) Viral and mycoplasmal infections of laboratory rodents: effect on biomedical research. Academic Press, Orlando, pp 801–808
22. Nelson GW, Stohlman SA (1993) Localization of the RNA binding domain of MHV nucleocapsid protein. J Gen Virol 74: 1975–1979
23. Parker MM, Masters PS (1990) Sequence comparison of the N genes of five strains of the coronavirus mouse hepatitis virus suggests a three domain structure for the nucleocapsid protein. Virology 179: 463–468
24. Pfleiderer M, Skinner MA, Siddell SG (1986) Coronavirus MHV-JHM: nucleotide sequence of the mRNA that encodes the membrane protein. Nucleic Acids Res 14: 6338
25. Robbins SG, Frana MF, McGowan JJ, Boyle JF, Holmes KV (1986) RNA-binding proteins of coronavirus MHV: detection of monomeric and multimeric N protein with an RNA overlay-protein blot assay. Virology 150: 402–410
26. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
27. Skinner MA, Siddell SG (1983) Coronavirus JHM: nucleotide sequence of the mRNA that encodes nuleocapsid protein. Nucleic Acids Res 11: 5045–5054
28. Stohlman SA, Baric R, Nelson GN, Soe LH, Welter LM, Dean RJ (1988) Specific interaction between coronavirus leader RNA and nucleocapsid. J Virol 62: 4288–4295
29. Stohlman SA, Bergmann C, Cua D, Wege H, van der Veer R (1994) Location of antibody epitopes within the mouse hepatitis virus nucleocapsid protein. Virology 202: 146–153
30. Sugiyama K, Amano Y (1980) Hemagglutination and structural polypeptides of a new coronavirus associated with diarrhoea in infant mice. Arch Virol 66: 95–105
31. Yie Y, Wei Z, Tien P (1993) A simplified and reliable protocol for plasmid DNA sequencing: fast miniprep and denaturation. Nucleic Acids Res 21: 361
32. Yokomori K, Banner LR, Lai MMC (1991) Heterogeneity of gene expression of the hemagglutinin-esterase (HE) protein of murine coronavirus. Virology 183: 647–657

Authors' address: Dr. F. R. Homberger, Institute of Laboratory Animal Science, University of Zurich-Irchel, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland.

Received August 26, 1994