VIROLOGY 179, 463–468 (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

MONICA M. PARKER AND PAUL S. MASTERS[1]

Wadsworth Center for Laboratories and Research, New York State Department of Health, Albany, New York 12201–0509

Received May 2, 1990; accepted July 11, 1990

To obtain information about the structure and evolution of the nucleocapsid (N) protein of the coronavirus mouse hepatitis virus (MHV), we determined the entire nucleotide sequences of the N genes of MHV-A59, MHV-3, MHV-S, and MHV-1 from cDNA clones. At the nucleotide level, the N gene sequences of these viral strains, and that of MHV-JHM, were more than 92% conserved overall. Even higher nucleotide sequence identity was found in the 3′ untranslated regions (3′ UTRs) of the five strains, which may reflect the role of the 3′ UTR in negative-strand RNA synthesis. All five N genes were found to encode markedly basic proteins of 454 or 455 residues having at least 94% sequence identity in pairwise comparisons. However, amino acid sequence divergences were found to be clustered in two short segments of N, putative spacer regions that, together, constituted only 11% of the molecule. Thus, the data suggest that the MHV N protein is composed of three highly conserved structural domains connected to each other by regions that have much less constraint on their amino acid sequences. The first two conserved domains contain most of the excess of basic amino acid residues; by contrast, the carboxy-terminal domain is acidic. Finally, we noted that four of the five N genes contain an internal open reading frame that potentially encodes a protein of 207 amino acids having a large proportion of basic and hydrophobic residues.

Coronaviruses are a family of enveloped, single-stranded, positive-sense RNA viruses that are important respiratory, neurologic, and enteric pathogens for humans and domestic animals (*1*). Having the largest genomic coding capacities among RNA viruses (at least 27 kb) as well as a unique strategy of RNA replication, coronaviruses represent very unusual and interesting molecular biological entities (*2, 3*). To gain insight into the roles played by the coronavirus nucleocapsid (N) protein during viral infection, we have been characterizing this protein in the well-studied coronavirus mouse hepatitis virus (MHV).

One approach to understanding protein structure and function is to chart evolutionarily permissible changes among closely related proteins. To this end, we have cloned and sequenced the N genes of four strains of MHV: MHV-A59, MHV-3, MHV-S, and MHV-1. Although closely related, these viruses have distinct histories, most notably, separate times and geographic loci of isolation and different mouse strains of origin (*4–7*). As well, the N proteins of these MHV strains exhibit considerable electrophoretic mobility variation on sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) (data not shown; Refs. (*8–10*)), suggesting differences in protein size or amino acid composition.

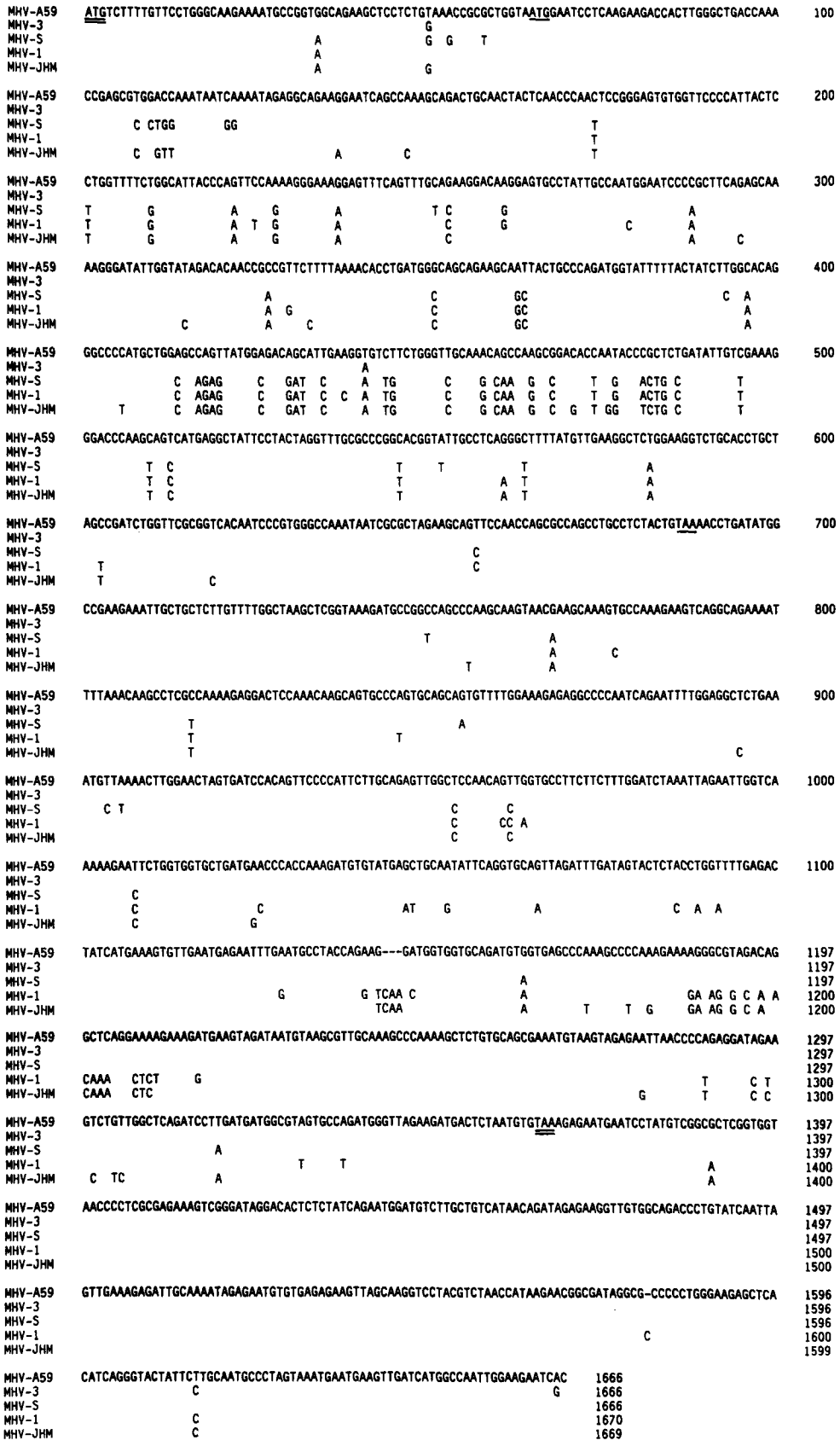
Multiple cDNA clones, prepared from poly(A)-containing RNA from infected mouse 17 clone 1 cells, were used to determine the nucleotide sequences of the N genes of MHV-A59, MHV-1, MHV-3, and MHV-S. MHV-A59 was taken as our reference strain because its sequence had been previously reported (*11, 12*) and because our heat-stable variant of this strain is the parent of a number of temperature-sensitive mutants that we plan to characterize in future work (L. S. Sturman *et al.*, manuscript in preparation). With the exception of the final 71 nt of the 3′ untranslated region (3′ UTR, see Fig. 1), the entire sequence of this N gene was determined in both directions at least once. At all positions where differences occurred between our sequence and the previously reported sequence (nt 441, 784, 1317, 1399, and 1481–1483), we verified the difference on at least four additional independent cDNA clones.

Similarly, the entire N gene sequences of MHV-1, MHV-3, and MHV-S were determined in both directions at least once. At all positions where a difference occurred with respect to our prototypic MHV-A59 sequence (Fig. 1), this change was verified on at least two additional independent cDNA clones. All cDNA clones were in agreement at all positions examined with the following exceptions: nt 1317 of MHV-A59, for which

Nucleotide sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession Nos. M35253 (MHV-3), M35254 (MHV-1), M35255 (MHV-S), and M35256 (MHV-A59).

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

0042-6822/90 $3.00


MHV-A59 ATGTCTTTTGTTCCTGGGCAAGAAAATGCCGGTGGCAGAAGCTCCTCTGTAAACCGCGCTGGTAATGGAATCCTCAAGAAGACCACTTGGGCTGACCAAA 100
MHV-3 G
MHV-S A G G T
MHV-1 A
MHV-JHM A G
MHV-A59 CCGAGCGTGGACCAAATAATCAAAATAGAGGCAGAAGGAATCAGCCAAAGCAGACTGCAACTACTCAACCCAACTCCGGGAGTGTGGTTCCCCATTACTC 200
MHV-3
MHV-S C CTGG GG T
MHV-1 T
MHV-JHM C GTT A C T
MHV-A59 CTGGTTTTCTGGCATTACCCAGTTCCAAAAGGGAAAGGAGTTTCAGTTTGCAGAAGGACAAGGAGTGCCTATTGCCAATGGAATCCCCGCTTCAGAGCAA 300
MHV-3
MHV-S T G A G A T C G A
MHV-1 T G A T G A C G C A
MHV-JHM T G A G A C A C
MHV-A59 AAGGGATATTGGTATAGACACAACCGCCGTTCTTTTAAAACACCTGATGGGCAGCAGAAGCAATTACTGCCCAGATGGTATTTTTACTATCTTGGCACAG 400
MHV-3
MHV-S A C GC C A
MHV-1 A G C GC A
MHV-JHM C A C C GC A
MHV-A59 GGCCCCATGCTGGAGCCAGTTATGGAGACAGCATTGAAGGTGTCTTCTGGGTTGCAAACAGCCAAGCGGACACCAATACCCGCTCTGATATTGTCGAAAG 500
MHV-3 A
MHV-S C AGAG C GAT C A TG C G CAA G C T G ACTG C T
MHV-1 C AGAG C GAT C C A TG C G CAA G C T G ACTG C T
MHV-JHM T C AGAG C GAT C A TG C G CAA G C G T GG TCTG C T
MHV-A59 GGACCCAAGCAGTCATGAGGCTATTCCTACTAGGTTTGCGCCCGGCACGGTATTGCCTCAGGGCTTTTATGTTGAAGGCTCTGGAAGGTCTGCACCTGCT 600
MHV-3
MHV-S T C T T T A
MHV-1 T C T A T A
MHV-JHM T C T A T A
MHV-A59 AGCCGATCTGGTTCGCGGTCACAATCCCGTGGGCCAAATAATCGCGCTAGAAGCAGTTCCAACCAGCGCCAGCCTGCCTCTACTGTAAAACCTGATATGG 700
MHV-3
MHV-S C
MHV-1 T C
MHV-JHM T C
MHV-A59 CCGAAGAAATTGCTGCTCTTGTTTTGGCTAAGCTCGGTAAAGATGCCGGCCAGCCCAAGCAAGTAACGAAGCAAAGTGCCAAAGAAGTCAGGCAGAAAAT 800
MHV-3
MHV-S T A
MHV-1 A C
MHV-JHM T A
MHV-A59 TTTAAACAAGCCTCGCCAAAAGAGGACTCCAAACAAGCAGTGCCCAGTGCAGCAGTGTTTTGGAAAGAGAGGCCCCAATCAGAATTTTGGAGGCTCTGAA 900
MHV-3
MHV-S T A
MHV-1 T T
MHV-JHM T C
MHV-A59 ATGTTAAAACTTGGAACTAGTGATCCACAGTTCCCCATTCTTGCAGAGTTGGCTCCAACAGTTGGTGCCTTCTTCTTTGGATCTAAATTAGAATTGGTCA 1000
MHV-3
MHV-S C T C C
MHV-1 C CC A
MHV-JHM C C
MHV-A59 AAAAGAATTCTGGTGGTGCTGATGAACCCACCAAAGATGTGTATGAGCTGCAATATTCAGGTGCAGTTAGATTTGATAGTACTCTACCTGGTTTTGAGAC 1100
MHV-3
MHV-S C
MHV-1 C C AT G A C A A
MHV-JHM C G
MHV-A59 TATCATGAAAGTGTTGAATGAGAATTTGAATGCCTACCAGAAG---GATGGTGGTGCAGATGTGGTGAGCCCAAAGCCCCAAAGAAAAGGGCGTAGACAG 1197
MHV-3 1197
MHV-S A 1197
MHV-1 G G TCAA C A GA AG G C A A 1200
MHV-JHM TCAA A T T G GA AG G C A 1200
MHV-A59 GCTCAGGAAAAGAAAGATGAAGTAGATAATGTAAGCGTTGCAAAGCCCAAAAGCTCTGTGCAGCGAAATGTAAGTAGAGAATTAACCCCAGAGGATAGAA 1297
MHV-3 1297
MHV-S 1297
MHV-1 CAAA CTCT G T C T 1300
MHV-JHM CAAA CTC G T C C 1300
MHV-A59 GTCTGTTGGCTCAGATCCTTGATGATGGCGTAGTGCCAGATGGGTTAGAAGATGACTCTAATGTGTAAAGAGAATGAATCCTATGTCGGCGCTCGGTGGT 1397
MHV-3 1397
MHV-S A 1397
MHV-1 T T A 1400
MHV-JHM C TC A A 1400
MHV-A59 AACCCCTCGCGAGAAAGTCGGGATAGGACACTCTCTATCAGAATGGATGTCTTGCTGTCATAACAGATAGAGAAGGTTGTGGCAGACCCTGTATCAATTA 1497
MHV-3 1497
MHV-S 1497
MHV-1 1500
MHV-JHM 1500
MHV-A59 GTTGAAAGAGATTGCAAAATAGAGAATGTGTGAGAGAAGTTAGCAAGGTCCTACGTCTAACCATAAGAACGGCGATAGGCG-CCCCCTGGGAAGAGCTCA 1596
MHV-3 1596
MHV-S 1596
MHV-1 C 1600
MHV-JHM 1599
MHV-A59 CATCAGGGTACTATTCTTGCAATGCCCTAGTAAATGAATGAAGTTGATCATGGCCAATTGGAAGAATCAC 1666
MHV-3 C G 1666
MHV-S 1666
MHV-1 C 1670
MHV-JHM C 1669

T was read on four clones and A on one clone; nt 1279 of MHV-3, T on five clones and C on one clone; nt 69 of MHV-S, A on three clones and T on one clone; nt 293 of MHV-S, C on three clones and T on one clone; nt 416 of MHV-S, C on three clones and T on one clone; nt 638 of MHV-S, A on four clones and G on one clone; and nt 134 of MHV-1, G on two clones and T on one clone. Thus, for these seven nucleotides, the bases given in Fig. 1 represent consensus sequences. The apparent disagreements at these positions most likely reflect the error rate either of the MHV RNA-dependent RNA polymerase, which generated the original transcripts, or of reverse transcriptase, which was used in the construction of the cDNA clones.

An alignment of the four determined MHV nucleotide sequences, together with the previously reported N gene sequence of MHV-JHM (*16*), is presented in Fig. 1. All five N genes are more than 92% homologous. In pairwise comparisons, the two most similar sequences are those of MHV-A59 and MHV-3; the most distant are those of MHV-1 and either MHV-A59 or MHV-3 (Table 1). The greatest densities of nucleotide differences among the N genes are in two regions corresponding to nt 414–486 and nt 1141–1214 of the MHV-A59 sequence. For the most distant strains, 50% of the nucleotide differences are clustered in these segments, which, combined, represent less than 8% of either sequence. By contrast, the most conserved portion of the N genes occurs in the 3′ untranslated regions (UTRs), which diverge by no more than 3 nt over a total span of 301 nt. This degree of sequence identity, which exceeds that of any portion of the N gene coding region, may reflect some functional constraint on the 3′ UTR, which presumably acts as a recognition site for the viral RNA polymerase during negative-strand RNA synthesis.

An alignment of the deduced amino acid sequences of the N proteins of the five MHV strains is shown in Fig. 2. All five N genes encode proteins of 454 or 455 residues, having molecular weights ranging from 49.6 to 49.7 kDa. Thus, the apparent size differences observed among them (data not shown; Refs. (*8–10*)) probably reflect differences in amounts of bound SDS or residual secondary structure under the conditions of SDS–PAGE. Alternatively, the variation in electrophoretic mobilities may indicate different types or extents of post-translational modification.

TABLE 1

N Gene Nucleotide and Amino Acid Sequence Differences

MHV strains	Nucleotide differences[a]		Amino acid differences[a]
	Coding region	3′ UTR	
A59 and 3	2	2	1
A59 and S	73	0	18
A59 and 1	102	3	28
A59 and JHM	101	2	30
3 and S	71	2	17
3 and 1	102	3	29
3 and JHM	99	2	29
S and 1	65	3	25
S and JHM	60	2	25
1 and JHM	53	1	20

[a] Gaps in the sequence alignments are counted as differences.

All five N proteins possess at least 94% sequence identity in pairwise comparisons (Table 1). All have the salient features noted previously for N of MHV-A59 and MHV-JHM: a large excess of basic residues over acidic residues (calculated pI's of 10.4–10.6); numerous serine residues, some of which are potential phosphorylation targets (*17, 18*); and an acidic carboxy terminus, in contrast to the rest of the molecule (*11, 16*).

As with the nucleotide sequences, the divergences among the amino acid sequences are clustered in two regions, corresponding to amino acids 140–162 and

FIG. 1. Nucleotide sequence comparison of the N genes of five strains of MHV. The heat-stable strain of MHV-A59 used in this study was obtained from Dr. Lawrence Sturman, Wadsworth Center for Laboratories and Research. MHV-3 was from Dr. Kathryn Holmes, Uniformed Services University of the Health Sciences, and, in turn, had been obtained from Dr. Abigail Smith, Yale University. MHV-S and MHV-1 were originally from Dr. John Parker, Microbiological Associates. Libraries of cDNA clones were generated from poly(A)-containing infected cell RNA by a modification of the procedure of Gubler and Hoffman (*13*) using the vector pMG5, as described previously (*14*). DNA sequencing was carried out by a variation of the dideoxy chain termination method of Sanger *et al.* (*15*) using modified T7 DNA polymerase (Sequenase, U.S. Biochemical). The synthetic oligodeoxynucleotide primers used for sequencing corresponded to nt 77–93, 328–345, 577–597, 827–847, 1077–1094, and 1307–1323, or were complementary to nt 280–297, 580–597, 730–747, 880–897, 1030–1047, 1180–1197, 1330–1347, and 1632–1649 of the MHV-A59 sequence. To obtain sequences of the 5′ and 3′ extremes of genes, cDNA inserts were subcloned into pGEM vectors (Promega), and sequencing was primed with oligodeoxynucleotides corresponding to the SP6 or T7 RNA polymerase promoters. The MHV-A59, -3, -S, and -1 sequences were determined in this work, except for nt 1596–1666 of MHV-A59, which is taken from Armstrong *et al.* (*11*). The MHV-JHM N sequence is from Skinner and Siddell (*16*). Spaces indicate positions for which the nucleotide is identical to that of MHV-A59. Nucleotides are numbered from the first base of the N protein initiation codon; 3′ polyadenylate tails are omitted. Hyphens indicate gaps introduced to maximize the alignment of sequences. The N protein initiation and termination codons are double-underlined. The initiation and termination codons of the major internal open reading frame are single-underlined.

```
MHV-A59   MSFVPGQENA GGRSSSVNRA GNGILKKTTW ADQTERGPNN QNRGRRNQPK QTATTQPNSG   60
MHV-3                    G
MHV-S                S   GS S                  AG   G
MHV-1                S
MHV-JHM              S   G                      L        K

MHV-A59   SVVPHYSWFS GITQFQKGKE FQFAEGQGVP IANGIPASEQ KGYWYRHNRR SFKTPDGQQK  120
MHV-3
MHV-S                            VQ
MHV-1                            Q
MHV-JHM                          Q                Q

MHV-A59    LLPRWYFYY LGTGPHAGAS YGDSIEGVFW VANSQADTNT RSDIVERDPS SHEAIPTRFA  180
MHV-3
MHV-S                           E    D    V    SQ    K   TA
MHV-1                           E    D D  V    SQ    K   TA
MHV-JHM                  Y      E    D    V    SQ  E R   SA

MHV-A59   PGTVLPQGFY VEGSGRSAPA SRSGSRSQSR GPNNRARSSS NQRQPASTVK PDMAEEIAAL  240
MHV-3
MHV-S
MHV-1
MHV-JHM                                P

MHV-A59   VLAKLGKDAG QPKQVTKQSA KEVRQKILNK PRQKRTPNKQ CPVQQCFGKR GPNQNFGGSE  300
MHV-3
MHV-S
MHV-1
MHV-JHM                                                                  P

MHV-A59   MLKLGTSDPQ FPILAELAPT VGAFFFGSKL ELVKKNSGGA DEPTKDVYEL QYSGAVRFDS  360
MHV-3
MHV-S                           A
MHV-1                           PS                   D              I
MHV-JHM                         A                    G

MHV-A59   TLPGFETIMK VLNENLNAYQ K-DGGADVVS PKPQRKGRRQ AQEKKDEVDN VSVAKPKSSV  419
MHV-3                                                                        419
MHV-S                                                                        419
MHV-1                    D      DQA                RGTK QKAL G              420
MHV-JHM                         NQ                 RGTK QKAQ                420

MHV-A59   QRNVSRELTP EDRSLLAQIL DDGVVPDGLE DDSNV  454
MHV-3                                          454
MHV-S                                          454
MHV-1                                          455
MHV-JHM                                        455
```

FIG. 2. Amino acid sequence comparison of the N proteins of five strains of MHV. The deduced MHV-A59, -3, -S, and -1 N sequences are from this work. The deduced MHV-JHM N sequence is taken from Skinner and Siddell (*16*). Spaces indicate positions for which the amino acid is identical to that of MHV-A59. The hyphen indicates a gap introduced to maximize the alignment of sequences. The two clustered regions of amino acid differences are boxed.

381–405 of the MHV-A59 sequence. For the most divergent pair of proteins, those of MHV-A59 and MHV-JHM, 63% of the amino acid differences are concentrated in these two portions of N, which together make up only 11% of the molecule. This distribution of residue changes, shown graphically in Fig. 3, suggests a model for the MHV N protein in which three conserved structural domains (basic, basic, and acidic) are tethered to each other by two regions of variable amino acid composition (designated A and B). We suggest that A and B have less constraint on their amino acid sequences and principally serve as spacers connecting the three conserved domains. In contradistinction, domains I, II, and III appear to tolerate few amino acid changes, implying that most changes in these regions impair the functioning of the molecule.

This model is supported by two further observations. First, we have characterized a temperature-sensitive N protein mutant of MHV-A59 that has a deletion almost exactly coincident with spacer B, indicating that, at least at the permissive temperature, the presence of this region is not absolutely required for N protein function (C. A. Koetzner *et al.*, unpublished results). Second, in an *in vitro* assay system, domains I and III were found to be dispensible for the binding of N protein to RNA, suggesting that the RNA-binding characteristic of N resides in domain II (P. S. Masters, manuscript in preparation). Thus, the domains inferred from our amino acid sequence comparison may be functionally separable as well as structurally distinct.

It is noteworthy that the nonconserved residues in spacers A and B tend to vary among a limited set of

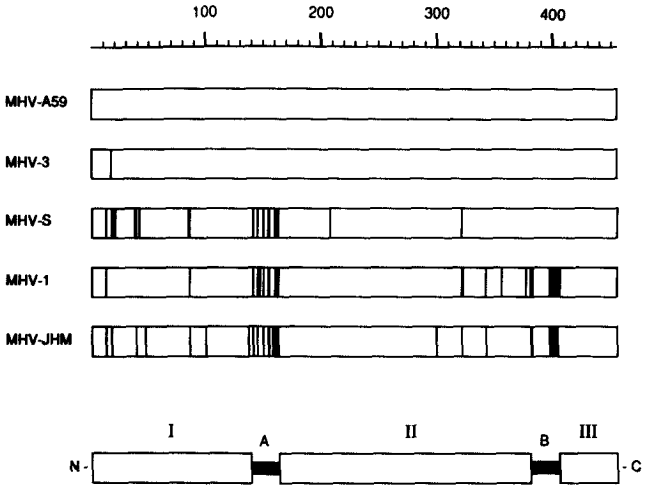


FIG. 3. Schematic representation of amino acid differences among the N proteins of five strains of MHV and a three-domain model of the MHV N protein. The line at the top indicates numbers of amino acid residues. In the rectangle for each MHV strain, a vertical line represents an amino acid difference with respect to the prototype MHV-A59. At the bottom is shown a model for the MHV N protein with three domains separated by two spacer regions (A and B).

two or three alternatives (Fig. 2). This might have suggested that these two regions are required to vary coordinately: i.e., an "A59-like" spacer A must always pair with an A59-like spacer B and a "JHM-like" spacer A must always pair with a JHM-like spacer B. However, the N protein of MHV-S clearly rules out this possibility, since this N protein has a JHM-like spacer A and an A59-like spacer B (Figs. 2 and 3). Thus, the MHV-S N gene is likely to have arisen from a recombination event between two ancestral viruses: one having an N gene more similar to MHV-A59 and MHV-3 and the other having an N gene more similar to MHV-1 and MHV-JHM. RNA recombination among murine coronaviruses has been shown to occur both in tissue culture and in the brains of doubly infected animals (*19, 20*). All five N gene sequences compared here, then, appear to be accounted for by either drift or recombination plus drift from two prototype genes.

Four of the five MHV genes in Fig. 1 contain a potentially significant internal open reading frame (ORF) in the +1 reading frame relative to the N protein ORF, beginning at nt 65 and terminating at nt 688. In each case, the protein encoded by this ORF is 207 residues in length (22.6–22.9 kDa) and is distinguished by a large excess of basic residues (calculated pI's of 10.6–11.1) as well as a relatively high (17%) leucine content (Fig. 4). The MHV-JHM N gene contains a very similar ORF in the same position, but this is interrupted by a stop codon following the 16th amino acid residue. For all of the N genes, the start codon for the internal ORF occurs in a strong context for translation initiation, whereas the N protein start codon (nt 1–3) and an intervening start codon (nt 26–28) both fall in suboptimal contexts. Thus, it is possible that the internal ORF may be translated by means of a "leaky scanning" mechanism (*21*). Leucine-rich internal ORFs also have been noted within the N genes of bovine coronavirus (*22*) and human coronavirus 229E (*23*). The significance of these potential polypeptides awaits determination of

```
MHV-A59   MESSRRPLGL TKPSVDQIIK IEAEGISQSR LQLLNPTPGV WFPITPGFLA LPSSKRERSF   60
MHV-3
MHV-S                    ALE ME                      I         L  R    N R  K
MHV-1                                                I         L  R    NFR  K
MHV-JHM                  AG*         R  P            I         L  R    N R  K

MHV-A59   SLQKDKECLL PMESPLQSKR DIGIDTTAVL LKHLMGSRSN YCPDGIFTIL AQGPMLEPVM  120
MHV-3
MHV-S       Y  R        Q              D        A  S           S E     AQS
MHV-1       H  R     T  Q              DG       A  S              E     AQS
MHV-JHM     H           Q H        T   D P      A  S              E     AQS

MHV-A59   ETALKVSSGL QTAKRTPIPA LILSKGTQAV MRLFLLGLRP ARYCLRAFML KALEGLHLLA  180
MHV-3         E
MHV-S     A ISKEL  S  ANRP LR L P  L    VA            V F   V       Q
MHV-1     A ISTEL  S  ANRP LR L P  L    VA            V     KV      Q       V
MHV-JHM   A ISKEL  S  ANRPRLG L P  L    VA            V     KV      Q       V

MHV-A59   DLVRGHNPVG QIIALEAVPT SASLPLL  207
MHV-3
MHV-S                      A
MHV-1                      A
MHV-JHM
```

FIG. 4. Amino acid sequence comparison of the major internal open reading frames of the N genes of five strains of MHV. The deduced MHV-A59, -3, -S, and -1 sequences are from this work. The corresponding region of MHV-JHM is deduced from Skinner and Siddell (*16*). Spaces indicate positions for which the amino acid is identical to that of MHV-A59. The asterisk in the MHV-JHM sequence indicates a stop codon.

whether any of them are actually synthesized in coronavirus-infected cells.

ACKNOWLEDGMENTS

We are grateful to Lawrence Sturman for many valuable discussions and to David Anders for critically reading the manuscript. We thank Scott Goebel for much useful advice on sequencing. This work was supported in part by Grant GM31698 from the National Institutes of Health.

REFERENCES

1. STURMAN, L. S., and HOLMES, K. V., *Adv. Virus Res.* **28,** 35–111 (1983).
2. LAI, M. M.-C., *BioEssays* **5,** 257–260 (1986).
3. SPAAN, W., CAVANAGH, D., and HORZINEK, M. C., *J. Gen. Virol.* **69,** 2939–2952 (1988).
4. GLEDHILL, A. W., and ANDREWES, C. H., *Brit. J. Exp. Pathol.* **32,** 559–568 (1951).
5. DICK, G. W. A., NIVEN, J. S. F., and GLEDHILL, A. W., *Brit. J. Exp. Pathol.* **37,** 90–98 (1956).
6. MANAKER, R. A., PICZAK, C. V., MILLER, A. A., and STANTON, M. F., *J. Natl. Cancer Inst.* **27,** 29–51 (1961).
7. ROWE, W. P., HARTLEY, J. W., and CAPPS, W. I., *Proc. Soc. Exp. Biol. Med.* **112,** 161–165 (1963).
8. BOND, C. W., LEIBOWITZ, J. L., and ROBB, J. A., *Virology* **94,** 371–384 (1979).
9. CHELEY, S., MORRIS, V. L., CUPPLES, M. J., and ANDERSON, R., *Virology* **115,** 310–321 (1981).
10. LAI, M. M.-C., and STOHLMAN, S. A., *J. Virol.* **38,** 661–670 (1981).
11. ARMSTRONG, J., SMEEKENS, S., and ROTTIER, P., *Nucleic Acids Res.* **11,** 883–891 (1983).
12. ARMSTRONG, J., SMEEKENS, S., SPAAN, W., ROTTIER, P., and VAN DER ZEIJST, B., *Adv. Exp. Med. Biol.* **173,** 155–162 (1984).
13. GUBLER, U., and HOFFMAN, B. J., *Gene* **25,** 263–269 (1983).
14. MASTERS, P. S., and BANERJEE, A. K., *Virology* **157,** 298–306 (1987).
15. SANGER, F., NICKLEN, S., and COULSON, A. R., *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467 (1977).
16. SKINNER, M. A., and SIDDELL, S. G., *Nucleic Acids Res.* **11,** 5045–5054 (1983).
17. STOHLMAN, S. A., and LAI, M. M.-C., *J. Virol.* **32,** 672–675 (1979).
18. SIDDELL, S. G., BARTHEL, A., and TER MEULEN, V., *J. Gen. Virol.* **52,** 235–243 (1981).
19. MAKINO, S., KECK, J. G., STOHLMAN, S. A., and LAI, M. M.-C., *J. Virol.* **57,** 729–737 (1986).
20. KECK, J. G., MATSUSHIMA, G. K., MAKINO, S., FLEMING, J. O., VANNIER, D. M., STOHLMAN, S. A., and LAI, M. M.-C., *J. Virol.* **62,** 1810–1813 (1988).
21. KOZAK, M., *J. Cell Biol.* **108,** 229–241 (1989).
22. LAPPS, W., HOGUE, B. G., and BRIAN, D. A., *Virology* **157,** 47–57 (1987).
23. SCHREIBER, S. S., KAMAHORA, T., and LAI, M. M.-C., *Virology* **169,** 142–151 (1989).