VIROLOGY 198, 355–359 (1994)

A Murine Coronavirus MHV-S Isolate from Persistently Infected Cells Has a Leader and Two Consensus Sequences between the M and N Genes

FUMIHIRO TAGUCHI,*,[1] TOSHIO IKEDA,* SHINJI MAKINO,† AND HIROSHI YOSHIKURA‡

*National Institute of Neuroscience, NCNP, 4-1-1 Ogawahigashi, Kodaira, Tokyo 187, Japan; †Department of Microbiology, University of Texas at Austin, Austin, Texas 78212-1095; and ‡Department of Bacteriology, University of Tokyo, Hongo, Tokyo 113, Japan

Received July 1, 1993; accepted September 14, 1993

A plaque-cloned mouse hepatitis virus mutant, MHV-S No. 8, was isolated from Ki-BALB cells persistently infected with MHV-S. The mRNAs 1 to 6 were larger in the mutant, whereas there was no difference between the two viruses in the size of the smallest mRNA, mRNA 7. Sequence analyses of the genomic RNA, mRNA 6, and mRNA 7 of the two viruses revealed that an additional 111 nt were inserted just upstream of the intergenic consensus sequence preceding the N gene in MHV-S No. 8. The inserted region consisted of two different parts; the 3′-most 30 nt corresponded to nucleotides 28 to 57 of the leader sequence and the 5′-most 81 nt corresponded to nucleotides 58 to 138 of mRNA 7. This structure of No. 8 was most likely generated by RNA–RNA recombination between genomic RNA and subgenomic RNA species. The nucleotide insertion in the intergenic sequence between genes M and N resulted in two consensus sequences separated by 111 nt. Primer extension analysis revealed that the amount of a slightly larger, subgenomic mRNA resulting from initiation of synthesis at the upstream consensus sequence was only 5% of the usual sized mRNA 7 initiated from the downstream consensus sequence. © 1994 Academic Press, Inc.

Mouse hepatitis virus (MHV) is a coronavirus which contains a single-stranded, nonsegmented, positive sense RNA of about 31 kb (*1, 2*). The MHV genome has a 70- to 80-nt-long leader sequence at the 5′ end (*1, 3, 4*). The 3′ end region of the leader sequence contains one to four repeats of the pentanucleotide, UCUAA (*1, 3*). Downstream of the leader are the MHV-specific genes encoding the viral structural and nonstructural proteins in the order of the genes P (polymerase), 2, HE (hemagglutinin esterase), S (spike), 4, 5, M (membrane), and N (nucleocapsid) (*3, 5*). These genes are separated by a short stretch of sequence called the intergenic region. All intergenic regions, which precede genes required for MHV replication, contain the consensus sequence, UCUAAC, or a very similar sequence (*6*).

In MHV-infected cells, seven to eight different species of mRNAs which form a 3′ coterminal nested set are detected (*4, 7*). Though genomic MHV RNA contains only a single copy of the leader sequence, the 5′ end of each subgenomic mRNA has a leader sequence which is identical to that of the genome (*1, 5, 8*). The negative-stranded RNAs complementary to each mRNA exist in coronavirus-infected cells (*9, 10*) and these negative-stranded RNAs have been shown to have antileader sequence (*11*) and a short poly(U) tail (*12*). It was found that subgenomic-size replicative-intermediate RNA is present in MHV-infected cells (*10*) and it has been proposed that subgenomic mRNA species may accumulate by the replication of each subgenomic mRNA species (*9, 10*). Although several models explaining how subgenomic-size RNA is initially synthesized have been proposed (*1*), this unique mechanism remains to be described.

Only one protein, encoded by the 5′-most coding region of each mRNA, is usually translated. The major structural proteins are the N protein of about 60 kDa, the 23- to 25-kDa M protein, and the S protein of about 150 to 200 kDa. They are translated from mRNAs 7, 6, and 3, respectively. Another major protein of some coronaviruses is the HE glycoprotein with an approximate molecular weight of 65 kDa (*13, 14*).

A mutant virus isolated from cells persistently infected with MHV-S showed cold sensitivity for the formation of polykaryocytes and lacked the HE protein (*15*). In the present paper, the genomic structure of that mutant virus, MHV-S No. 8, was compared with the parental (wt) virus, MHV-S. The data suggested that MHV-S No. 8 underwent nonhomologous RNA–RNA recombination, which probably occurred between genomic and subgenomic RNA species.

To compare the size of subgenomic mRNAs of wt MHV-S and those of the mutant No. 8, total RNA was isolated from DBT cells infected with wt or No. 8 as previously described (*16*). The virus-specific RNAs were detected by hybridization with ^{32}P-labeled cDNA as previously reported (*16, 17*). It was reported that mRNA 4 of MHV-S was not synthesized while the smaller mRNA 5, which encoded the small membrane protein, was synthesized (*18*) (Fig. 1A). The comparison of the mRNA pattern of MHV-S and that of No. 8 revealed that there were differences in the sizes of mRNAs between wt and No. 8. The mRNAs 6 and 5 of No. 8 were about 100 nt longer than those of wt. This

[1] To whom reprint requests should be addressed.

0042-6822/94 $5.00
Copyright © 1994 by Academic Press, Inc.
All rights of reproduction in any form reserved.

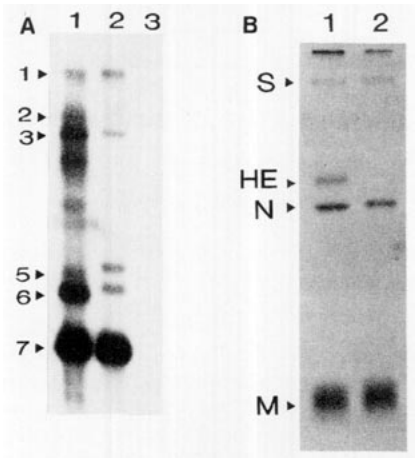

FIG. 1. (A) Northern blot analysis of MHV-S and No. 8 mRNAs. RNAs extracted from DBT cells infected with MHV-S (1), No. 8 (2), and from mock-infected cells (3) were electrophoresed in agarose gel in MOPS buffer and transferred onto nitrocellulose paper, and virus-specific mRNAs were detected by hybridization with ^{32}P-labeled cDNA to JHM N gene. (B) SDS–polyacrylamide gel electrophoresis analysis of virion proteins of MHV-S (1) and No. 8 (2). Virions labeled with [^{3}H]leucine were purified from culture fluids of infected DBT cells and analyzed by electrophoresis on a 10% polyacrylamide gel.

difference in size was also observed in mRNAs larger than mRNA 3, although the difference was not as pronounced as mRNAs 6 and 5. In contrast, mRNA 7 size was the same for both viruses. The data suggested that a sequence has been inserted between an area upstream of the intergenic region preceding gene 7 (N gene) and an area downstream of the intergenic region preceding gene 6 (M gene). The comparison of the mRNA also showed that the amounts of mRNA 6 of No. 8 was extremely low as compared with that of wt.

To test whether or not an insertion existed within the open reading frame (ORF) of the M gene, the size of the No. 8 M protein was compared with that of wt by sodium dodecyl sulfate (SDS)–polyacrylamide gel electrophoresis using radiolabeled purified virions (*19, 20*). As shown in Fig. 1B, there was no detectable difference in the size of M protein or the other structural proteins. It was also found that No. 8 lacked the 65-kDa protein, as reported previously (*15*). These data suggested that the larger mRNA 6 of the mutant contained no additional sequence within the M gene ORF.

The location of the inserted sequence in No. 8 was investigated by comparing the structures of the M gene and the 5′-region of gene N with those of MHV-S. The mRNA 6 sequences of both viruses were determined from cDNA clones. The clones were constructed using a synthetic oligonucleotide primer, Viro II (5′ GGCATTTTCTTGCCCAG 3′), which is complementary to the sequence 97 to 113 nt from the 5′-end of MHV–JHM mRNA 7 (*21, 17*). To examine the sequence at the intergenic region between genes M and N of both viruses, direct sequencing of the reverse transcriptase–polymerase chain reaction (RT-PCR) products from genomic RNA was performed (*23*). The RT-PCR product was obtained using primers N280 (5′ AATGCCCGAAAACCAAGAGT 3′), which corresponds to the complementary sequence of mRNA 7 (nt 280 to 299) of MHV–JHM (*17*), and MR6-P1 (5′ AAGGTAGACGGTGTTAGCGG 3′), corresponding to nt 644 to 663 from the 5′-end of MHV–JHM M gene (*22*). The mRNA 7 5′-region was examined by direct sequencing of RT-PCR product which was made using primers N280 and Lea-1 (5′ TATAAGAGTGATTGGCG 3′), which corresponds to the 5′-end of the leader sequence of mRNA7 of MHV–JHM (nt 1 to 17) (*17*).

Sequence analysis of cDNA clones and RT-PCR products revealed that there were no nucleotides inserted within the ORF of the No. 8 M gene. Also, no

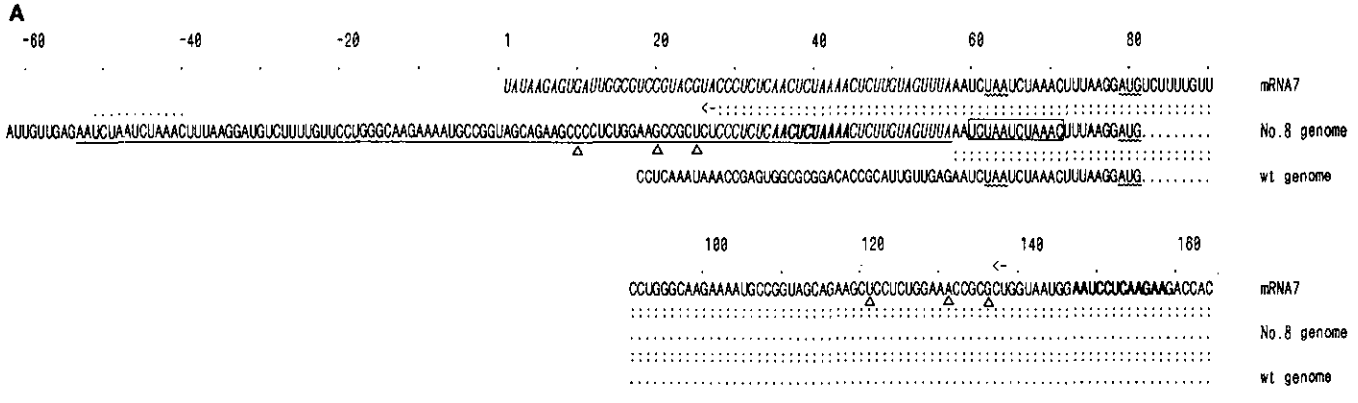

FIG. 2A. Nucleotide sequences of genomic RNA and mRNA 7, 5′-ends of MHV-S and No. 8. Nucleotide numbering is the same as the JHM N gene (*16*). Identical sequences downstream of the initiation codon of the N gene are shown by dots. The leader sequence is italicized and the 111 additional nucleotides found in No. 8 genome are underlined. The upstream consensus sequence is shown by overhead dots and the downstream by a box. Homologous sequences in the leader and N gene are shown by the italicized and nonitalicized bold letters, respectively. The initiation codon of the N gene and termination codon of the M gene are underlined with wavy lines. Nucleotides number 27 and 138, shown by arrows, mark the sites where nearly identical sequences start. The three triangles in these similar sequences mark three mismatches.

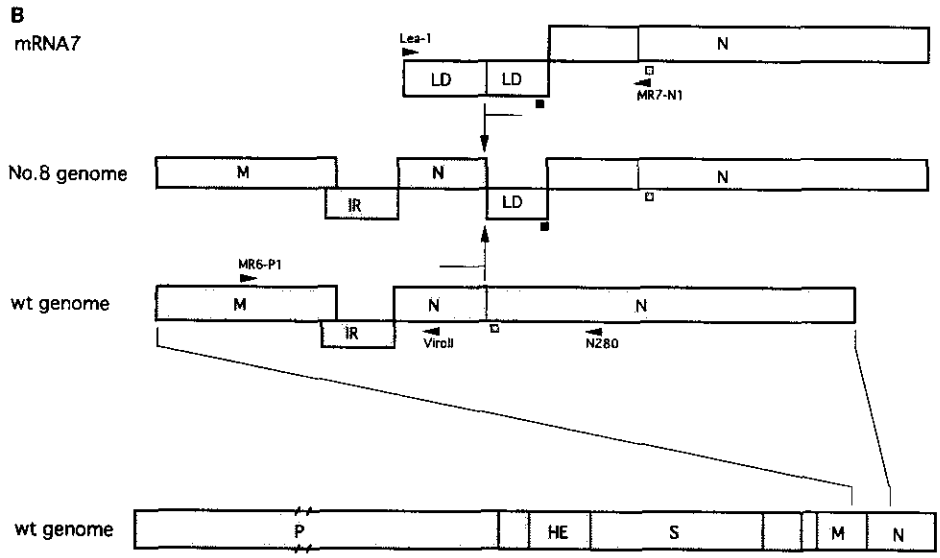

FIG. 2B. Comparison of No. 8 genome structure with the wt genome and mRNA 7. The structure of No. 8, in the intergenic region between genes M and N, indicates that the genome was constructed from the MHV-S genome lacking most of the coding region of the N gene (upstream shown by arrow) fused with mRNA 7 minus the leader sequence 5′ end (downstream shown by arrow). Arrows show the site of fusion. ■ and □ indicate the homologous sequences. IR, intergenic region; LD, leader sequence. Arrowheads indicate the position of the primers described in the text.

sequence difference was found in the 5′-most 280 nt of mRNA 7. However, No. 8 contained an additional 111 nt just upstream of the N gene intergenic consensus sequence (Fig. 2A). The inserted region had two different parts; the 3′-most 30 nt corresponded to leader sequence nt 28 to 57 and the 5′-most 81 nt corresponded to mRNA 7 nt 58 to 138. Therefore, the genome of mutant No. 8 consisted of a partial N gene, deleted from nt 139 to the 3′-end, fused to a partial mRNA 7 lacking the first 27 nt from the 5′-end (Fig. 2B). This inserted 111 nt contained three substitutions at nt positions 10, 20, and 25 as compared to the corresponding regions of the wt sequence (Fig. 2A). The insertion resulted in two intergenic regions between M and N genes separated by 111 nt.

Northern blot analysis of MHV-S and No. 8 revealed that the mRNA 7 sizes of both viruses were the same (Fig. 1), suggesting that the intergenic region which is present at the 5′-end of the inserted 111 nt in No. 8 was transcriptionally inactive. It was possible that small quantities of a "large mRNA 7" were not detectable at the level of sensitivity of Northern blotting, therefore primer extension analysis was used to look for this putative RNA. The primer extension procedure was described previously (*24, 25*). A ^{32}P-end-labeled oligonucleotide, MR7-N1 (5′ GAGGATTCCAGGACCAGCGC 3′), which binds to mRNA 7 at 139 to 158 nt from the 5′-end, was used as a primer and the total intracellular RNA species from wt- and No. 8-infected cells were used as templates. The primer extension products were analyzed by 6% denaturing polyacrylamide gel electrophoresis. There was one 165-nt-long primer extension product in wt-infected cells and two products in the No. 8-infected cells (Fig. 3A). The products in the No. 8-infected cells were 165-nt long and 276-nt long; the latter was expected from the mRNA with 111 extra nucleotides. The amount of RNA in the 276-nt-long

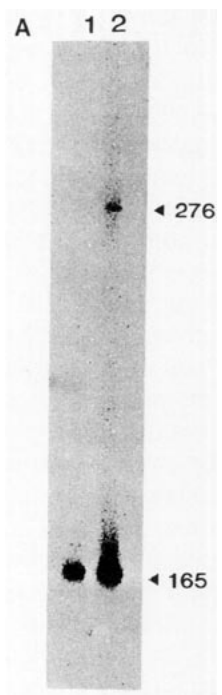

FIG. 3A. Primer extension analysis of RNA extracted from MHV-S (1) or No. 8 (2) infected cells. RNAs extracted from DBT cells infected with MHV-S or No. 8 were reversed transcribed with ^{32}P-end-labeled oligonucleotide, MR7-N1, as a primer. The products were analyzed by the electrophoresis on a 6% denaturing polyacrylamide gel. To see the precise length of the products, sequencing reactions with M13 DNA were also electrophoresed.

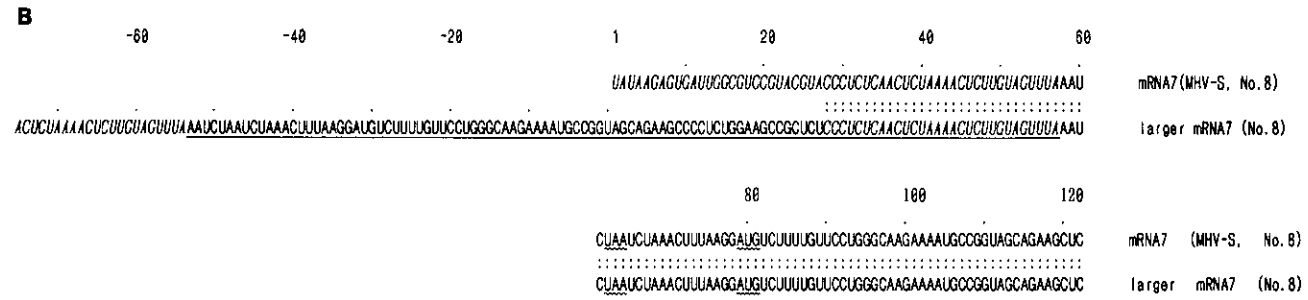

FIG. 3B. Nucleotide sequences of the 5′ end of mRNA 7 and the "larger mRNA 7." The 165-nt bands from MHV-S and No. 8 as well as the 276-nt band of No. 8 were cut out from gels and amplified by PCR. The PCR products were directly sequenced. The additional 111 nt found in the larger mRNA 7 are underlined. The leader sequence is shown in italics and the termination codon of the M gene and the initiation codon of the N gene are underlined with wavy lines.

band was only about 5% of that in the 165-nt-long band. The structure of the primer extension products were examined; the three different DNA bands were cut out from gels and amplified by PCR with primers, MR7-N1 and Lea-1. Direct sequencing of the PCR products showed that the 165-nt fragment corresponded to the 5′-end sequence of mRNA 7 and that the 276-nt fragment contained an additional 111 nt identical to those found in genomic RNA as shown in Fig. 3B. These data clearly showed that the two different consensus sequences found in the intergenic region between the M and N genes of No. 8 were used for mRNA synthesis, whereas the upstream intergenic consensus sequence demonstrated significantly reduced transcriptional activity.

The structure of No. 8 strongly suggested that nonhomologous RNA recombination occurred between different-sized RNA species. It is possible that during the elongation of negative-stranded RNA on the mRNA 7 template, nascent negative-stranded RNA switched its template from mRNA 7 to mRNA 1. It should be noted that there is some nucleotide homology between the leader sequence, at nt 35 to 44, and mRNA 7, at nt 147 to 158 (Figs. 2A and 2B). It is likely that the template RNA switching was facilitated by this sequence complementarity. Alternatively, No. 8 genomic structure might be the result of nonhomologous RNA recombination where an elongating mRNA 1 switched from mRNA 1 minus-stranded template to mRNA 7 minus-stranded template. It is reported that the genome of human coronavirus OC43 contains 47 nt from the 3′-end of the leader sequence between the S and M genes (*26*). Probably, a similar mechanism is involved in the generation of MHV-S No. 8 and OC43 genomic RNA.

It is interesting that the amount of the "larger mRNA 7," which contained an extra 111 nt, was extremely low as compared with the normal-sized mRNA 7 in No. 8-infected cells. Analysis of cloned MHV DI RNAs demonstrated that the downstream and upstream flanking sequences of the intergenic region preceding N gene do not play an important role in DI subgenomic RNA synthesis (*27*). It is possible that the presence of the 3′-half of the leader sequence downstream of the intergenic region may suppress subgenomic RNA synthesis at the upstream intergenic region. This possibility is currently under investigation.

ACKNOWLEDGMENTS

We thank A. Yamada for valuable suggestions on the direct sequencing of the PCR products, S. G. Siddell for valuable suggestions on cDNA cloning, and J. K. Makino for reading the manuscript. This work was supported by a grant from the Ministry of Education, Science and Culture of Japan and Public Health Service Grant AI29984 (to S.M.) from the National Institutes of Health, U.S.A.

Note added in proof. The nucleotide sequence data reported in this paper have been deposited with the GSDB, DDBJ, EMBL, and NCBI nucleotide sequence databases under Accession nos. D21872, D21873, D21874, and D21875.

REFERENCES

1. LAI, M. M. C., *Annu. Rev. Microbiol.* **44**, 303–333 (1990).
2. LEE, H., SHIEH, C., GORBALENYA, A. E., LOONIN, R. V., LAMONICA, N., TULER, J., BAGDZHADZHYAN, A., and LAI, M. M. C., *Virology* **180**, 567–582 (1991).
3. SPAAN, W., CAVANAGH, D., and HORZNEK, M. C. *J. Gen. Virol.* **69**, 2939–2952 (1988).
4. LAI, M. M. C., PATTON, C. D., BARIC, R. S., and STOHLMAN, S. A. *J. Virol.* **46**, 1027–1033 (1983).
5. SIDDELL, S., WEGE, H., and TER MEULEN, V. *J. Gen. Virol.* **64**, 761–776 (1983).
6. SHIEH, C.-K., SOE, L. H., MAKINO, S., CHANG, M.-F., STOHLMAN, S. A., and LAI, M. M. C., *Virology* **156**, 321–330 (1987).
7. LAI, M. M. C., BRAYTON, P. R., ARMEN, R. C., PATTON, C. D., PUGH, C., and STOHLMAN, S. A., *J. Virol.* **39**, 823–834 (1981).
8. LAI, M. M. C., BARIC, R. S., BRAYTON, P. R., and STOHLMAN, S. A., *Proc. Natl. Acad. Sci. USA* **81**, 3626–3630 (1984).
9. SETHNA, P. B., HUNG, S.-L., and BRIAN, D. A., *Proc. Natl. Acad. Sci. USA* **86**, 5626–5630 (1989).
10. SAWICKI, S. G., and SAWICKI, D. L., *J. Virol.* **64**, 1050–1056 (1990).
11. SETHNA, P. B., HOFMANN, A. M., and BRIAN, D. A., *J. Virol.* **65**, 320–325 (1991).
12. HOFMANN, M. A., and BRIAN, D. A., *J. Virol.* **65**, 6331–6333 (1991).
13. CAVANAGH, D., BRIAN, D. A., ENJUANES, L., HOLMES, K., LAI,

M. M. C., LAUDE, H., SIDDELL, S. G., SPAAN, W., TAGUCHI, F., and TALBOT, P. J., *Virology* **176,** 306–307 (1990).
14. KING, B. B., POTTS, B. J., and BRIAN, D. A., *Virus Res.* **2,** 53–59 (1985).
15. YOSHIKURA, H., and TEJIMA, S., *Virology* **113,** 503–511 (1981).
16. TAGUCHI, F., SIDDELL, S., WEGE, H. H., and TER MEULEN, V. *J. Virol.* **44,** 487–492 (1985).
17. SKINNER, M. A., and SIDDELL, S. G. *Nucleic Acid Res.* **5,** 5045–5054 (1983).
18. YOKOMORI, K., and LAI, M. M. C., *J. Virol.* **65,** 5605–5608 (1991).
19. MAKINO, S., TAGUCHI, F., HAYAMI, M., and FUJIWARA, K., *Microbiol. Immunol.* **27,** 445–454 (1983).
20. TAGUCHI, F., HAYASHI, T., YAMADA, A., and FUJIWARA, K., *Jpn. J. Exp. Med.* **48,** 369–371 (1978).
21. GUBLER, U., and HOFMAN, B. J., *Gene* **25,** 263–269 (1983).
22. PFLEIDERER, M., SKINNER, M. A., and SIDDELL, S. G., *Nucleic Acid Res.* **14,** 6338 (1986).
23. WINSHIP, P. R., *Nucleic Acid Res.* **17,** 1266 (1989).
24. MAKINO, S., and LAI, M. M. C., *Virology* **57,** 227–232 (1989).
25. SAMBROOK, J., FRITSCH, E. F., and MANIATIS, T., "Molecular Cloning," 2nd ed., pp. 7.79–7.82. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989.
26. MOUNIR, S., and TALBOT, P. J., *Virology* **192,** 355–360 (1993).
27. MAKINO, S., and JOO, J., *J. Virol.* **67,** 3304–3311 (1993).