JOURNAL OF VIROLOGY, Nov. 1991, p. 6331–6333
0022-538X/91/116331-03$02.00/0

Vol. 65, No. 11

The 5′ End of Coronavirus Minus-Strand RNAs Contains a Short Poly(U) Tract

MARTIN A. HOFMANN† AND DAVID A. BRIAN*

Department of Microbiology, The University of Tennessee, Knoxville, Tennessee 37996-0845

Received 23 May 1991/Accepted 25 July 1991

A radiolabeled oligodeoxynucleotide primer that anneals near the common 5′ end of bovine coronavirus minus-strand RNAs was extended with reverse transcriptase, and a major product suggesting poly(U) tracts of 8 to 20 nucleotides was found. The extended primer molecules were ligated head to tail, amplified by the polymerase chain reaction, cloned, and sequenced, and poly(U) tracts of 9 to 26 nucleotides were found. Poly(A) tails of 100 to 130 nucleotides on the 3′ end of coronavirus plus-strand mRNAs and genome must, therefore, be generated by a mechanism that uses only a short poly(U) template. This pattern contrasts with that of other cytoplasmic, polyadenylated, plus-strand animal RNA viruses which utilize a full-length poly(U) template for poly(A) synthesis.

Coronavirus subgenomic mRNAs and plus-strand genome have been shown to be polyadenylated with tracts of 100 to 130 A residues (14, 34). The origin of the poly(A) tract is unknown, but polyadenylation probably occurs by a cytoplasmic mechanism, since coronavirus replication does not require the cell nucleus (4, 33). Judging from the precedents of other single-stranded RNA viruses of eucaryotes, one possible mechanism for polyadenylation might be synthesis during transcription from a full-length poly(U) template. In the case of coronaviruses, the poly(U) template would be expected to be on the common 5′ end of the minus-strand antigenome and subgenomic mRNAs, since these molecules have been shown to be active in genome and subgenomic mRNA synthesis (11, 23, 26, 27). A full-length poly(U) template is used by picornaviruses and togaviruses, for example, two other positive-strand animal RNA viruses that undergo cytoplasmic replication (6, 22). For these viruses, however, genomic RNAs with experimentally shortened poly(A) tails can give rise by an unknown mechanism to progeny genomes with full-length poly(A) tails when transfected into cells (7, 21, 30). A second possibility for the polyadenylation of coronavirus mRNAs might be synthesis from a less-than-full-length poly(U) template on the minus strand. In this case, poly(A) synthesis might occur by a replicase chattering mechanism of the kind proposed for the polyadenylation of mRNAs of minus-strand rhabdoviruses, paramyxoviruses, and myxoviruses (18, 19, 25). A third possibility, based on cellular mechanisms, might be that there are no terminal U residues on the 5′ end of minus-strand RNAs and that A's are added to the 3′ end of plus strands by nontemplated synthesis in response to a *cis*-acting polyadenylation signal (28, 31, 32). Although this mechanism was recently shown to be used by a member of the *Potexvirus* family, a family of single-stranded, plus-strand cytoplasmic RNA viruses of plants (8), it is unlikely to be used by coronaviruses, since the requisite consensus polyadenylation signal (AAUAAA) is not found in the 3′-terminal sequence of coronavirus mRNAs (1, 2, 5, 12, 13, 16, 24, 29).

To determine whether 5′ poly(U) stretches exist on bovine coronavirus (BCV) minus strands, an oligodeoxynucleotide (primer 1; 5′-CTTGCGGAAGTAATTGCCGA-3′) homologous to plus-strand nucleotides −91 through −110 [beginning at the base of the 3′ poly(A) tail] (Fig. 1A) (16) was used as a radiolabeled primer in an extension reaction with reverse transcriptase, and the product of the extended minus strands was examined by electrophoresis on a denaturing polyacrylamide gel (Fig. 1B). To do this, total cytoplasmic RNA was extracted from BCV-infected human rectal tumor cells at 6 h postinfection, as previously described (11), and 100 μg was digested with RNase A (50 μg/ml; Sigma) or RNase T_1 (100 U/ml; Boehringer Mannheim) in 0.3 M NaCl, as previously described (26), to remove excess strands of plus-strand RNA molecules which are known to be in a 100 to 1,000 molar excess over minus-strand RNA molecules (11) and which interfered with extension reactions in preliminary experiments (data not shown). The rationale for this step is based on the high probability that all or nearly all intracellular genome- and subgenome-length minus strands exist in double-stranded replicative forms and are extractable as double-stranded forms (23). A total of 50 pmol of primer, 5′ end labeled with ^{32}P by the forward reaction, was used with the RNase-digested RNA in the extension reaction by the method of Sambrook et al. (20). Electrophoresis was done on a 6% polyacrylamide sequencing gel containing 50% urea, and the wet gel was exposed to film for observing the distribution of extended products. The only products found that extended into the putative poly(U) region were those of about 118 to 130 nucleotides (nt) [corresponding to potential poly(U) tails of 8 to 20 nt] (Fig. 1B, region 2). Products longer than this were no more abundant from reactions with infected-cell RNA than from reactions with uninfected-cell RNA, regardless of whether RNase A (not shown) or RNase T_1 (Fig. 1B, region 2) had been used in the RNA preparation, suggesting that there were very few or no poly(U) tracts longer than 20 nt.

To characterize the extended sequence on the 5′ end of minus strands, single-stranded cDNA products were eluted from the polyacrylamide gel (regions 1, 2, and 3 in Fig. 1B were eluted and studied separately), ligated head to tail with RNA ligase, amplified by the polymerase chain reaction (PCR), and cloned into a *Sma*I-linearized pGEM3Z plasmid (Promega Biotec) as described elsewhere (10). Only extended cDNA from region 2 in Fig. 1B gave a PCR-amplifi-

* Corresponding author.
† Present address: Institut fuer Viruskrankheiten und Immunoprophylaxe, CH-4000 Basel, Switzerland.

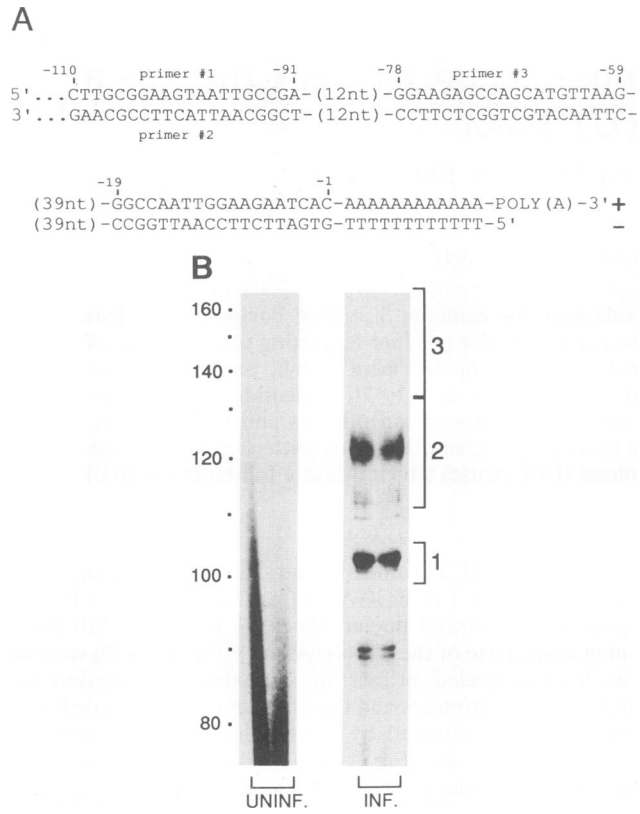

FIG. 1. Estimation of poly(U) tail lengths on BCV minus-strand RNAs. (A) Schematic representation of the 3′ end of the BCV genome (and subgenomic mRNAs) shown as plus-strand DNA with a poly(A) tail and the complementary minus strand shown as DNA with a short poly(T) tract [to represent poly(U)]. The sequences and positions of the primers used for characterization of the poly(U) tract are identified. (B) 5′-end-labeled primer 1 that binds near the 5′ end of BCV minus-strand RNAs was extended with reverse transcriptase and electrophoresed on a 6% DNA sequencing gel together with products from a DNA sequencing reaction that were used as size markers (not shown). The length of extended primer in nucleotides is indicated on the left. Regions identified by brackets were cut from the gel, and the material from each was eluted and studied separately as described in the text. Extended products representing the ends of poly(U) tracts are shown in bracketed region 2. An extended primer of 120 nt represents a poly(U) tail of 10 nt. UNINF. and INF., uninfected-cell and infected-cell mRNA, respectively.

able product. For the PCR amplification, oligodeoxynucleotide 2 (primer 2; 5′-TCGGCAATTACTTCCGCAAG-3′), complementary to primer 1, and oligonucleotide 3 (primer 3; 5′-GGAAGAGCCAGCATGTTAAG-3′), homologous to genomic bases −59 through −78 (Fig. 1A) (16), were used. Clones were screened initially by colony hybridization to radiolabeled primers 2 and 3 and then by size determination of excised inserts by agarose gel electrophoresis. Inserts of more than 110 nt [110 nt is equivalent to no poly(U) tract] were sequenced by dideoxynucleotide sequencing of asymmetrically amplified DNA as previously described (9).

A total of 48 clones were sequenced, and of these, 29 contained poly(U) tracts. The remainder contained artifactual inserts of primer sequences. No clone representing an intact 5′ end of a minus strand, judging by the clones' complementarity to the 3′ noncoding sequences on the genome, was devoid of a poly(U) tract. Poly(U) tracts ranged from 9 to 26 U's, and a majority of these (21 of 29) were 9 to 13 nt.

Attempts to determine by Northern (RNA) analysis whether each individual subgenomic minus-strand species contains the poly(U) tract were unsuccessful, apparently because of the presence of a large number of cytoplasmic oligo(dA)-binding nucleic acid species that anneal under the conditions of low stringency required for this assay. Although minus strands were readily detectable with intragenomic minus-strand-binding oligodeoxynucleotide probes (11; also our unpublished data), γ-^{32}P-end-labeled oligo(dA)$_{18}$ probes detected only heterogeneous RNAs of a wide size range in both uninfected and infected cells (data not shown).

Three kinds of experimental evidence suggest that the poly(A) tails on BCV plus-strand RNAs are considerably longer than the poly(U) tracts identified on minus strands. First, when cDNA clones of BCV genomic RNA were prepared by priming first-strand cDNA synthesis with oligo(dT)$_{18}$, clones with poly(A) tails of up to 68 nt were obtained (16). Second, in experiments in which BCV RNA was digested to completion with RNase T$_1$ and analyzed by two-dimensional gel electrophoresis, a large spot that migrated to the same relative position on gels as the poly(A) spot from digested mouse hepatitis virus RNA (15) was obtained. Since no homopolymeric stretches of bases other than A have been identified in fully sequenced coronavirus genomes (3, 17), it is unlikely that the RNase T$_1$-resistant spot in BCV RNA is anything other than the 3′ poly(A) tail. These data together suggest that BCV poly(A) tails, like those on mouse hepatitis virus RNAs (14), are 100 to 130 nt. Finally, when the lengths of the BCV poly(A) tails were measured by electrophoresis of PCR-generated products that had been initiated by random priming along the poly(A) tails, tail homopolymeric lengths of up to 140 nt were obtained (unpublished data).

Short poly(U) tracts on minus strands could have arisen from the use of RNases in the preparation of minus-strand templates for primer extension, but we view this as unlikely, since it would have required a coincidence of two separate unlikely events, one for RNase A and the other for RNase T$_1$. First, short poly(U) tracts could have been generated by RNase A digestion if there had been breathing of the poly(U) · poly(A) heteroduplex in the 0.3 M NaCl. Inasmuch as poly(U) · poly(A) heteroduplexes of up to 90 nt are protected in 0.2 M NaCl (22), such breathing in 0.3 M NaCl would be extremely unusual. Second, short poly(U) tracts could have been generated by RNase T$_1$ if there had been one or more G residues positioned 9 to 26 nt from the 5′ ends of the minus-strand RNAs and within longer poly(U) tracts. Since such G's would most probably have C counterparts in the poly(A) tail, and since these have not been found in sequence analyses of cloned BCV poly(A) tails (16), digestion of poly(U) tracts with RNase T$_1$ would not be expected. Short poly(U) tracts on minus strands are, in addition, consistent with our finding that minus strands almost universally electrophoretically migrate faster than their plus-strand counterparts on denaturing agarose gels (11, 27).

On the basis of precedent, we can envisage two possible mechanisms by which coronaviruses might synthesize long poly(A) tails from a short poly(U) template. First, poly(A) tails might be synthesized by slippage or chattering mechanisms in the replicase similar to those described for rhabdoviruses, paramyxoviruses, and myxoviruses in the polyadenylation of mRNAs (for a review, see reference 18). In these instances, poly(A) tails of more than 100 nt are generated from an oligo(U) template of 4 to 8 nt. Second, it is

conceivable that coronaviruses could use a cellular cytoplasmic poly(A) polymerase that normally functions to extend poly(A) tails on cellular mRNAs in the cytoplasm (28, 31, 32). In these cases, the AAUAAA polyadenylation signal required for polyadenylation in the nucleus or cytoplasm is not required if the mRNA species already possesses a short tract of at least 10 A's (28). Coronavirus mRNAs and genome may, therefore, use short a poly(U) tract for templated synthesis of short poly(A) tails by the viral replicase (or transcriptase) and then use the cellular cytoplasmic poly(A) polymerase and the AAUAAA-independent mechanism for nontemplated addition to form long poly(A) tails.

We thank Phiroze B. Sethna for many helpful discussions.

This work was supported by Public Health Service grant AI14367 from the National Institutes of Health and in part by a grant from the Upjohn Company, Kalamazoo, Michigan. M.A.H. was supported by a postdoctoral fellowship from the Swiss National Science Foundation.

REFERENCES

1. **Armstrong, J., S. Smeekens, and P. Rottier.** 1983. Sequence of the nucleocapsid gene from murine coronavirus MHV-A59. Nucleic Acids Res. **11**:883–891.
2. **Boursnell, M. E. G., M. M. Binns, I. J. Foulds, and T. D. K. Brown.** 1985. Sequences of the nucleocapsid genes from two strains of avian infectious bronchitis virus. J. Gen. Virol. **66**:573–580.
3. **Boursnell, M. E. G., T. D. K. Brown, I. J. Foulds, P. F. Green, F. M. Tomley, and M. M. Binns.** 1987. Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. J. Gen. Virol. **68**:57–77.
4. **Brayton, P. R., R. G. Ganges, and S. A. Stohlman.** 1981. Host cell nuclear function and murine hepatitis virus replication. J. Gen. Virol. **56**:457–460.
5. **DeGroot, R. J., A. C. Andeweg, M. C. Horzinek, and W. J. M. Spaan.** 1988. Sequence analysis of the 3′ end of the feline coronavirus FIPV 79-1146 genome: comparison with the genome of porcine coronavirus TGEV reveals large insertions. Virology **167**:370–376.
6. **Dorsch-Häsler, K., Y. Yogo, and E. Wimmer.** 1975. Replication of picornaviruses. I. Evidence from in vitro RNA synthesis that poly(A) of the poliovirus genome is genetically coded. J. Virol. **16**:1512–1527.
7. **Frey, T. K., and J. H. Strauss.** 1978. Replication of Sindbis virus. VI. Poly(A) and poly(U) in virus-specific RNA species. Virology **86**:494–506.
8. **Guilford, P. J., D. L. Beck, and R. L. S. Forster.** 1991. Influence of the poly(A) tail and putative polyadenylation signal on the infectivity of white clover mosaic potexvirus. Virology **182**:61–67.
9. **Hofmann, M. A., and D. A. Brian.** 1991. Sequencing PCR DNA amplified directly from a bacterial colony. BioTechniques **11**: 16–17.
10. **Hofmann, M. A., and D. A. Brian.** 1991. A PCR-enhanced method for sequencing the 5′-prime end of messenger RNAs. PCR Methods Appl. **1**:43–45.
11. **Hofmann, M. A., P. B. Sethna, and D. A. Brian.** 1990. Bovine coronavirus mRNA replication continues throughout persistent infection in cell culture. J. Virol. **64**:4108–4114.
12. **Kamahora, T., L. H. Soe, and M. M. C. Lai.** 1989. Sequence analysis of nucleocapsid gene and leader RNA of human coronavirus OC43. Virus Res. **12**:1–9.
13. **Kapke, P. A., and D. A. Brian.** 1986. Sequence analysis of the porcine transmissible gastroenteritis coronavirus nucleocapsid protein gene. Virology **151**:41–49.
14. **Lai, M. M. C., P. R. Brayton, R. C. Armen, C. D. Patton, C. Pugh, and S. A. Stohlman.** 1981. Mouse hepatitis virus A59: mRNA structure and genetic localization of the sequence divergence from hepatotropic strain MHV-3. J. Virol. **39**:823–834.
15. **Lapps, W., and D. A. Brian.** 1985. Oligonucleotide fingerprint of antigenically related bovine coronavirus and human coronavirus OC43. Arch. Virol. **86**:101–108.
16. **Lapps, W., B. G. Hogue, and D. A. Brian.** 1987. Sequence analysis of the bovine coronavirus nucleocapsid and matrix protein genes. Virology **157**:47–57.
17. **Lee, H.-J., C.-K. Shieh, A. E. Gorbalenya, E. V. Koonin, N. L. Monica, J. Tuler, A. Bagdzhadzhyan, and M. M. C. Lai.** 1991. The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. Virology **180**:567–582.
18. **Masters, P. S., and A. K. Banerjee.** 1988. Replication of nonsegmented negative strand RNA viruses. *In* E. Domingo, J. J. Holland, and P. Ahlquist (ed.), RNA genetics, vol. 1. RNA-directed virus replication. CRC Press, Inc., Boca Raton, Fla.
19. **Robertson, J. S., M. Schubert, and R. A. Lazzarini.** 1981. Polyadenylation sites for influenza virus mRNA. J. Virol. **38**: 157–163.
20. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
21. **Sarnow, P.** 1989. Role of 3′-end sequences in infectivity of poliovirus transcripts made in vitro. J. Virol. **63**:467–470.
22. **Sawicki, D. L., and P. J. Gomatos.** 1976. Replication of Semliki Forest virus: polyadenylate in plus-strand RNA and polyuridylate in minus-strand RNA. J. Virol. **20**:446–464.
23. **Sawicki, S. G., and D. L. Sawicki.** 1990. Coronavirus transcription: subgenomic mouse hepatitis virus replicative intermediates function in RNA synthesis. J. Virol. **64**:1050–1056.
24. **Schreiber, S. S., T. Kamahora, and M. M. C. Lai.** 1989. Sequence analysis of the nucleocapsid protein gene of human coronavirus 229E. Virology **169**:142–151.
25. **Schubert, M., J. D. Keene, R. C. Herman, and R. A. Lazzarini.** 1980. Site on the vesicular stomatitis virus genome specifying polyadenylation and the end of the L gene mRNA. J. Virol. **34**:550–559.
26. **Sethna, P. B., M. A. Hofmann, and D. A. Brian.** 1991. Minus-strand copies of replicating coronavirus mRNAs contain antileaders. J. Virol. **65**:320–325.
27. **Sethna, P. B., S.-L. Hung, and D. A. Brian.** 1989. Coronavirus subgenomic minus-strand RNAs and the potential for mRNA replicons. Proc. Natl. Acad. Sci. USA **86**:5626–5630.
28. **Sheets, M., and M. Wickens.** 1989. Two phases in the addition of a poly(A) tail. Genes Dev. **3**:1401–1412.
29. **Skinner, M. A., and S. G. Siddell.** 1983. Coronavirus JHM: nucleotide sequence of the mRNA that encodes nucleocapsid protein. Nucleic Acids Res. **11**:5045–5054.
30. **Spector, D. H., and D. Baltimore.** 1974. Requirement of 3′-terminal poly (adenylic acid) for the infectivity of poliovirus RNA. Proc. Natl. Acad. Sci. USA **71**:2983–2987.
31. **Wickens, M.** 1990. How the messenger got its tail: addition of poly(A) in the nucleus. Trends Biochem. Sci. **15**:227–281.
32. **Wickens, M.** 1990. In the beginning is the end: regulation of poly(A) addition and removal during early development. Trends Biochem. Sci. **15**:320–324.
33. **Wilhelmsen, K. C., J. L. Leibowitz, C. W. Bond, and J. A. Robb.** 1981. The replication of murine coronaviruses in enucleated cells. Virology **110**:225–230.
34. **Yogo, Y., N. Hirano, S. Hino, H. Shibuta, and M. Matumoto.** 1977. Polyadenylate in the virion RNA of mouse hepatitis virus. J. Biochem. **82**:1103–1108.