Identification of Two New Polypeptides Encoded by mRNA5 of the Coronavirus Infectious Bronchitis Virus

D. X. LIU* AND S. C. INGLIS*,^{†,1}

* Division of Virology, Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QP United Kingdom; and + Immunology Limited, 184 Cambridge Science Park, Milton Road, Cambridge CB4 4GN United Kingdom

Received June 13, 1991; accepted September 13, 1991

The second smallest subgenomic messenger RNA, mRNA5, of the coronavirus infectious bronchitis virus includes in its "5" unique region" two separate open reading frames (5a and 5b), whose coding function has not so far been established, and thus it may represent a dicistronic messenger RNA. We report here that two polypeptides with the sizes expected for the 5a and 5b products can be synthesised by *in vitro* translation of a single artificial mRNA containing both the 5a and 5b ORFs. To establish whether these polypeptides represent genuine virus gene products, both the 5a and 5b coding sequences were expressed as bacterial fusion proteins, and these were used to raise monospecific antisera. Antisera raised against both the 5a and 5b-specific sequences recognized specifically proteins of the expected size in infectious bronchitis virus-infected chicken kidney and Vero cells, indicating that 5a and 5b do represent genuine virus genes, and suggesting that mRNA5 is indeed functionally dicistronic. © 1992 Academic Press, inc.

Infectious bronchitis virus (IBV), a pathogen of chickens, is the prototype virus of the coronaviridae, a family of enveloped viruses with a large positivestranded RNA genome. Deduction of the complete nucleotide sequence of the IBV genome, recently accomplished through cDNA cloning (4), has indicated that the virion RNA is 27 kilobases (kb) in length, and contains at least 10 separate open reading frames (ORFs) with the capacity to encode proteins of between 6.7 and 440K.

Although the genomic RNA is capped and polyadenylated and is infectious (21, 23), it appears that it does not act as messenger RNA for most of the known virus proteins; available evidence indicates that all but the two large ORFs nearest to the 5' end of the genome are translated from subgenomic mRNA species. Five such subgenomic mRNAs have been identified in virus-infected cells, and these form a 3' co-terminal 'nested' structure (28, 29). Recently these mRNAs have been redesignated mRNA 2 to 6 in decreasing order of size (9), with the genomic mRNA representing mRNA1. For mRNAs 2, 4, and 6, the 5'-"unique" region (i.e., the region which is not present in the next smallest mRNA species) contains a single ORF, suggesting that these mRNAs are functionally monocistronic. This is supported by in vitro translation studies which have shown that mRNAs 2, 4, and 6 encode the major virion structural proteins spike (S), membrane

(M), and nucleocapsid (N), respectively (30). However, the other three mRNAs 1, 3, and 5, all contain more than one ORF in their 5'-"unique" region, suggesting that they may be functionally polycistronic. Previous studies in this laboratory have indicated that the two 5' proximal ORFs present on the genomic RNA are expressed in infected cells (6) and can be translated from a single mRNA through an efficient ribosomal frame-shift mechanism (5, 7). Similarly all three of the ORFs present on mRNA 3 encode proteins in infected cells and can be translated from a single mRNA (20, 26) although by a different mechanism involving independent initiation. As yet however no information is available regarding the coding capacity of mRNA5.

Nucleotide sequence analysis (*3*) has shown that the 5'-"unique" region of mRNA5 contains two ORFs, designated 5a and 5b, which have the potential to encode two polypeptides of molecular weight 7.4 and 9.5K, respectively. We report here the identification of two products encoded by this mRNA in IBV-infected chicken kidney (CK) and monkey kidney (Vero) cells. In addition we show, by *in vitro* translation of synthetic mRNA, that both polypeptides may be synthesized from a single mRNA, suggesting that this mRNA, like mRNA 3, is functionally polycistronic.

As a first step toward identification of 5a and 5b gene products, we tested the coding potential of these ORFs by *in vitro* translation of synthetic mRNA; successful expression of the ORFs would support the idea that they represented genuine gene products, and would further provide markers for identification of 5a

¹ To whom reprint requests should be addressed.



Fig. 1. (a) Diagram of portions of the plasmids pIBS6 and pIBS7, showing the SP6 promoter site, the restriction enzyme sites used to linearise the plasmids for *in vitro* transcription, and the sizes of the expected *in vitro* translation products from transcribed RNA. In order to construct pIBS6, a 1016-bp cDNA fragment containing the 5b ORF, together with the beginning of the downstream N ORF, was excised from clone c5.136 (*3*) by digestion with *Accl*, end-repaired with DNA *Pol*I, redigested with *Pst*I, and cloned into *Smal/Pst*I digested pSP65 (*18*). To construct pIBS7, a 533-bp cDNA fragment containing the 5a and 5b ORFs was excised from clone c5.136 by digestion with *Ava*II, end-repaired with DNA *Pol*I, and cloned into *Smal/Pst*I digested pSP65. (b) Analysis of cell-free translation products of mRNA obtained by *in vitro* transcription from *Hind*III-digested pIBS6 and pIBS7, using SP6 RNA polymerase. RNA was transcribed from the plasmids as indicated, recovered by phenol-chloroform extraction and ethanol precipitation, and further purified by gel-filtration on Sephadex G-50, before addition to the wheat germ cell-free system, as indicated above each lane, at approximately 100 μ g/mI. Translation products were labeled with [³⁵S] methionine (0.75 μ Ci/ μ I), analyzed by SDS–PAGE on a 22% polyacrylamide gel, and detected by fluorography. HMW, molecular weight markers. *In vitro* transcription and translation was carried out as described previously (*5*, *15*).

and 5b proteins in virus-infected cells. For this purpose, two plasmids (pIBS6 and pIBS7) were constructed. Plasmid pIBS6 contains the 5b ORF adjacent to the SP6 phage RNA polymerase promoter, and plasmid pIBS7 is identical to pIBS6 except that it contains both the 5a and 5b ORFs (with 5a adjacent to the SP6 promoter). Details of these plasmid constructions are given in the legend to Fig. 1a.

As shown in Fig. 1a, plasmids pIBS6 and pIBS7 may be transcribed *in vitro* using the SP6 phage RNA polymerase into mRNA containing either both ORFs (pIBS7) or 5b alone (pIBS6). Transcripts were therefore prepared from *Hin*dIII-digested pIBS6 and pIBS7, incorporating the dinucleotide ^{7m}GpppG to provide a 5' cap structure (10), purified as detailed in Fig. 1b, and then translated in a cell-free system derived from wheat germ (15). The results of this experiment are shown in Fig. 1b. Messenger RNA from pIBS6 directed the synthesis of a single major translation product, with an apparent molecular weight of about 10K, which was consistent with the predicted MW of 9.5K for the product of 5b. Two major products were produced by pIBS7-derived transcripts. The larger one comigrated exactly with the 10K protein synthesized in response to pIBS6-derived RNA, identifying it as the product of 5b. The more abundant product however was not present among those directed by pIBS6 transcripts and migrated with the expected molecular weight for 5a (7.5K), strongly suggesting that it does indeed represent the product of the 5a ORF.

These data are therefore consistent with the idea that both the 5a and 5b ORFs could encode polypeptides in the virus-infected cells. Furthermore the observation that both 5a and 5b-encoded proteins could be synthesized *in vitro* in response to a single mRNA (i.e., the pIBS7-derived transcripts) raises the possibility that the natural mRNA5 may be functionally dicistronic. To investigate these possibilities further we therefore set out to examine virus-infected cells for the presence of 5a- and 5b-encoded polypeptides.

In order to identify the products of the 5a and 5b ORFs in virus-infected cells, we first sought to express the appropriate sequences in bacteria and to raise monospecific antisera against the resulting protein products. The bacterial expression system chosen (based on the PEX series of plasmids) was that developed by Stanley and Luzio (27). In this system, foreign sequences are inserted in the correct reading frame, at the C terminus of a β -galactosidase gene, which is itself fused to the promoter, operator and N terminal region of the crogene of bacteriophage λ . Synthesis of the β -galactosidase fusion protein in bacteria carrying these plasmids is therefore repressed in cells carrying a functional λ repressor, but can be induced in *Esche*richia coli POP2136 cells (which carry a temperaturesensitive repressor) simply by increasing the culture temperature to 42°. We have used this system previously to identify three proteins encoded by the IBV mRNA3 (20, 26).

Two plasmids were constructed for bacterial expression of the 5a and 5b ORFs. One of these, pEX1/5a, contains the 5a ORF fused in frame with the β -galactosidase gene of the vector, while the other, pEX2/5b, contains the 5b ORF fused in frame with the β -galactosidase gene of the vector. Details for the construction of these plasmids are given in the legend to Fig. 2. Following heat induction, bacteria harboring the pEX1/ 5a and pEX2/5b plasmids produced fusion proteins that were larger than the wild-type β -galactosidase by the expected amount (data not shown), indicating successful expression of the viral coding sequences. These fusion proteins were purified by electroelution. and then inoculated into rabbits to produce monospecific antisera. The specificity and reactivity of the monospecific antisera generated by this route were then tested by immunoprecipitation studies using radiolabeled 5a- and 5b-encoded proteins synthesized by in vitro translation of synthetic mRNA transcribed from pIBB4 using the T7 phage RNA polymerase. The results of this experiment (Fig. 2) indicated that both sera are indeed capable of recognising specifically the appropriate target sequence, although the relative efficiency of precipitation of the two proteins suggested that the avidity of the anti-5b serum may be considerably higher.

The two specific antisera were next used to immunoprecipitate 5a and 5b ORF-related proteins from [³⁵S]methionine labeled IBV-infected CK and Vero cell extracts. For this purpose, newly confluent monolayers of CK cells and Vero cells were infected with the Beaudette strain of IBV, labeled with [³⁵S]methionine, and harvested as previously described (*20*). Immunoprecipitation was carried out as previously described (*20*). The results of these experiments (Figs. 3a and 3b)



Fig. 2. Immunoprecipitation of in vitro synthesized 5a and 5b polypeptides with anti-5a and anti-5b antibodies. [35S]methionine-labeled in vitro translation products of mRNA derived from HindIII digested pIBB4 were analysed by electrophoresis on a 20% polyacrylamide gel and detected by fluorography either before (lane plBB4/HindIII) or after immunoprecipitation with anti-5a antiserum, anti-5b antiserum, or control rabbit antiserum (raised against a B-galactosidase/influenza NP fusion protein). LMW, molecular weight markers. In vitro transcription and translation was carried out as before (5, 15). Anti-5a and anti-5b antibodies were prepared by immunisation of rabbits with purified *β*-galactosidase fusion proteins expressed from the plasmids pEX1/5a and pEX2/5b, respectively (described in the text), and were affinity purified from crude sera by affinity chromatography as described (26). Construction of the plasmids was as follows: For insertion of a cDNA fragment containing the 5a ORF into a pEX-based plasmid, it was necessary to introduce a suitable restriction enzyme site immediately upstream of the 5a coding sequence. To construct a plasmid suitable for this manipulation, an EcoRI fragment containing the 5a and 5b ORFs was excised from pIBS7, end-repaired with DNA Poll, redigested with HindIII, and cloned into pIBT4 (20) which had been digested with Bg/II, end-repaired with DNA Poll, and redigested with HindIII. This plasmid, pIBB4, contains the 5a and 5b ORFs adjacent to the phage T7 RNA polymerase promoter, and additionally the intergenic sequence of bacteriophage f1. Plasmid pIBB4 was then modified by site-directed mutagenesis to create a site for the restriction enzyme Smal immediately upstream of the 5a initiation codon using the synthetic oligonucleotide 5'-CTTAACAAACCCGGGCGATGAAATG-3'. The resulting plasmid, pIBM3, was then digested with Smal and Pstl to obtain a 422-bp cDNA fragment containing the 5a and 5b ORFs, which was then cloned into Smal and Pstl digested pEX1 (26) to generate the plasmid pEX1/5a. Plasmid pIBM3 was further digested with Accl to release a fragment of 249-bp containing the 5b ORF which was then end-repaired with DNA Po/I, redigested with BamHI, and cloned into Smal and BamHI digested pEX-2, to generate the plasmid pEX2/5b.



Fig. 3. (a) Detection of polypeptides encoded by the 5a and 5b ORFs in IBV-infected and mock-infected Vero cells by immunoprecipitation using antisera raised against bacterial fusion proteins. Cells were infected with IBV, labeled with [³⁵S] methionine, lysates prepared, and polypeptides either analyzed directly, or immunoprecipitated with the antisera indicated above each lane as described previously (*5*). Lanes labeled LMW and HMW represent low and high molecular weight markers, respectively. Polypeptides were separated on a 22% SDS-polyacry-lamide gel and detected by fluorography. *In vitro* translation products from pIBB4-derived RNA (Fig. 4c) were also included as markers for the 5a and 5b products. (b) Detection of polypeptides encoded by the 5a and 5b ORFs in IBV-infected and mock-infected chicken kidney (CK) cells by immunoprecipitation. Cell lysates were prepared and the polypeptides were analyzed as indicated above.

indicated that antisera raised against the 5a-specific fusion protein could specifically recognize in each labeled lysate, although faintly, a polypeptide with the expected MW for 5a (i.e., 7.4K) and with the same electrophoretic mobility as the *in vitro* translation product of the 5a ORF. Likewise the anti-5b antisera was able to precipitate specifically a protein of the size expected for the 5b product (9.5K), as judged both by gel mobility and by comparison with *in vitro* translated 5b. Thus it appears that both the 5a and 5b ORFs are indeed expressed in IBV-infected cells.

To provide additional evidence for the presence of products from the 5a and 5b ORFs, and to investigate the subcellular location of the proteins in virus-infected cells, indirect immunofluorescence was carried out on IBV-infected Vero cells using anti-5a and anti-5b antibodies purified by affinity chromatography as previously described (*26*). As shown in Fig. 4, both antisera gave weak but clear positive staining in infected cells at 18 hr postinfection (p.i.) (Figs. 4a and 4c) but not in mock-infected cells (Figs. 4b and 4d). Using the anti-5a antibodies (Fig. 4a) diffuse fluorescence was observed throughout the cell, with some apparent concentration around the nucleus. With the anti-5b antibodies (Fig. 4c), the staining was also distributed over the whole cell, but in this case the pattern of fluorescence appeared more granular and was more obviously perinuclear in location.

Thus the results reported here clearly indicate that IBV-infected CK and Vero cells contain polypeptides that are recognized specifically by antisera raised against bacterial fusion proteins containing either 5a or 5b amino acid sequences. These polypeptides are of the sizes expected for the products of the 5a and 5b ORFs, and each comigrates during SDS-gel electrophoresis with its in vitro-translated counterpart. Thus both the 5a and 5b ORFs are expressed in infected cells. The function of these polypeptides however remains unclear. The deduced amino acid sequence of 5a is unusual in that 26% (17 out of 65) of its residues are leucines (3), and its obvious hydrophobic nature suggests that it may be membrane-associated, although it does not appear to contain a conventional "membrane-spanning" domain. Such an association would be consistent with the general pattern of staining observed in our immunofluorescence studies using anti-5a antibodies, and preliminary cell fractionation experiments suggest that it is indeed associated with



Fig. 4. Indirect immunofluorescence of IBV-infected (a and c) and mock-infected (b and d) Vero cells at 18 hr p.i. using anti-5a (a and b) and anti-5b (c and d) antisera. Cells were permeabilized with 1% Triton X100, and staining was carried out as described previously (26).

the membranes of infected cells (data not shown), but it has not so far proved possible to detect the protein in purified virions. The deduced amino acid sequence of the 5b protein however betrays no obvious clue as to its function, although our immunofluorescence studies suggest that it concentrates in the perinuclear region of infected cells.

Available evidence suggests that the 5a and 5b polypeptides are expressed from a single subgenomic mRNA, mRNA5, which consists of a 65 nucleotide leader sequence corresponding to the 5' end of the genomic RNA, fused to a ''body'' which initiates 25459 nucleotides from the genomic 5' end, and runs through to its extreme 3' terminus. This is based on the observation that the next smallest subgenomic mRNA, mRNA6, which encodes the viral nucleocapsid protein, does not contain the 5a and 5b ORFs, and the next largest, mRNA4, encodes the viral membrane protein, but not the products of 5a and 5b ORFs as indicated by *in vitro* translation of viral mRNA (*30*).

How could these two products be translated from a single messenger RNA? The most obvious possibility is by a "leaky scanning" mechanism, in which ribosomes bind at the capped 5' end of the mRNA and scan in a 3' direction until an initiation codon is en-

countered; occasionally however they may fail to recognize the initiation codon for the upstream 5a ORF and continue onward to the 5b ORF (17). The seguence context around the initiation codon of the upstream 5a ORF (CGGACGAUGG) conforms well to that preferred for functional eukaryotic initiation codons (16) with an A at the -3 position and G at +4, suggesting that translational initiation of this ORF should occur efficiently. Our in vitro cell free translation results using synthetic mRNAs bearing the 5a and 5b ORFs are consistent with this idea; these mRNAs directed efficient synthesis of a product corresponding to the 5a polypeptide (Fig. 3b). Expression of the 5b ORF from such synthetic dicistronic mRNAs in vitro was by contrast relatively inefficient, which would be expected if its synthesis relies on leaky scanning (17). Furthermore synthesis of both proteins from the dicistronic mRNA was sensitive to inhibition by the cap analogue ^{7m}GTP in vitro (data not shown) suggesting that translation of each required ribosome entry at the 5' end. However 5b synthesis in vitro appeared to be inefficient even when the 5b ORF was placed proximal to the 5' end of a synthetic mRNA (Fig. 1b), in spite of the fact that sequences around the 5b initiation codon (GCUGG-CAUGA) would seem to present a favorable context for

translation with a G at position -3 and A at +4. It is perhaps surprising therefore that expression of 5b was relatively easy to detect in virus-infected cells in comparison to that of 5a. This could represent a genuine difference in the relative expression ratio of the two proteins in vivo from that observed by in vitro translation, as has been reported previously for the products of other di-cistronic mRNAs (11, 12, 32). If this were the case, it would undermine somewhat the argument for leaky scanning as a mechanism to account for 5b translation in vivo. However, the relative ease of 5b detection may simply reflect differing avidities of our mono-specific antisera; indeed, the anti-5b serum clearly recognized its target protein more efficiently than did the anti-5a serum when tested against in vitro translated material (Fig. 2). At present then the mechanism by which mRNA5 can function dicistronically in infected cells remains unclear, although our in vitro translation data would favor a leaky scanning hypothesis.

Most eukaryotic mRNAs studied to date encode only a single protein, but in recent years, more and more animal viruses have been found to encode di- or polycistronic mRNAs. These include adenovirus (the E1B mRNA encodes two tumor antigens starting at different AUG triplets of two overlapping ORFs (2)), the papovaviruses (the late 19S RNA of Simian Virus 40 expresses both the VP2 and VP3 proteins (24)), the paramyxoviruses (for example, the measles virus and Sendai virus P/C mRNAs (1, 13)), and the influenza viruses (the influenza B virus RNA segment 6 encodes both the NB and NA glycoproteins in different overlapping reading frames (31)). Among the coronaviruses, there are also some examples of functionally polycistronic mRNAs. One of the other IBV subgenomic mRNAs, mRNA3, can function as a tricistronic messenger (20), encoding a 6.7K proteins (3a), a 7K protein (3b), and a membrane-associated virion protein (3c). Murine hepatitis virus also appears to encode a dicistronic mRNA. Messenger RNA5 contains two ORFs in its 5'-unique region, and a product from the second of these has been detected in virus infected cells (8, 19, 25). It seems therefore that the translation of small, often briefly overlapping ORFs from di- or polycistronic mRNAs may not be an uncommon feature of coronaviruses.

ACKNOWLEDGMENTS

This work was supported by a U.K. AFRC Link grant (LRG171) awarded to SCI. D. X. Liu is supported by a Cambridge Overseas Trust Studentship.

REFERENCES

- BELLINI, W. J., ENGLUND, G., ROZENBLATT, S., ARNHEITER, H., and RICHARDSON, C. D., J. Virol. 53, 908–919 (1985).
- BOS, J. L., POLDER, L. J., BERNARDS, R., SCHRIER, P. I., VAN DEN ELSEN, P. J., VAN DER EB, A. J., and VAN ORMONDT, H., *Cell* 27, 121–131 (1981).
- 3. BOURSNELL, M. E. G., and BROWN, T. D. K., Gene 29, 87-92 (1984).
- BOURSNELL, M. E. G., BROWN, T. D. K., FOULDS, I. J., GREEN, P. F., TOMLEY, F. M., and BINNS, M. M., *J. Gen. Virol.* 68, 57–77 (1987).
- BRIERLEY, I., BOURSNELL, M. E. G., BINNS, M. M., BILIMORIA, B., BLOK, V. C., BROWN, T. D. K., and INGLIS, S. C., *EMBO. J.* 6, 3779–3785 (1987).
- BRIERLEY, I., BOURSNELL, M. E. G., BINNS, M. M., BILIMORIA, B., ROLLEY, N. J., BROWN, T. D. K., and INGLIS, S. C. (1990). *In* "Coronaviruses and their Diseases" (D. Cavanagh and T. D. K. Brown, Eds.), Vol. 276, pp. 175–281. Advances in Experimental Medicine and Biology, Plenum Press, New York.
- 7. BRIERLEY, I., DIGARD, P., and INGLIS, S. C., Cell 57, 537–547 (1989).
- BUDZILOWICZ, C. Z., and WEISS, S. R., Virology 157, 509–515 (1987).
- CAVANAGH, D., BRIAN, D. A., ENJUANES, L., HOLMES, K. V., LAI, M. M. C., LAUDE, H., SIDDELL, S. G., SPAAN, W., TAGUCHI, F., and TALBOT, P. J., *Virology* **176**, 306–307 (1990).
- CONTRERAS, R., CHEROUTRE, H., DEGRAVE, W., and FIERS, W., Nucleic Acids Res. 10, 6353–6362 (1982).
- CURRAN, J. C., RICHARDSON, C. D., and KOLAKOFSKY, D. (1986). J. Virol. 57, 684–687.
- DORNER, A. J., SEMLER, B. L., JACKSON, R. J., HANECAK, R., DU-PREY, E., and WIMMER, E., J. Virol. 50, 507–514 (1984).
- GIORGI, C., BLUMBERG, B. M., and KOLAKOFSKY, D., Cell 35, 829– 836 (1983).
- 14. HERMAN, R. C., Trends Biol. Sci. 14, 219-222 (1989).
- 15. INGLIS, S. C., McGEOCH, D. J. and МАНҮ, B. W. J., *Virology* **78**, 522–536 (1977).
- 16. KOZAK, M., Nucleic Acids Res. 15, 8125-8149 (1987).
- 17. KOZAK, M., J. Mol. Biol. 108, 229-241 (1989).
- KRIEG, P. A. and MELTON, D. A., Nucleic Acids Res. 12, 7057– 7071 (1984).
- LEIBOWITZ, J. L., PERLMAN, S., WEINSTOCK, G., DEVRIES, J. R., BUD-ZILOWICZ, C., WEISSEMANN, J. M., and WEISS, S. R., Virology 164, 156–164 (1988).
- LIU, D. X., CAVANAGH, D., GREEN, P., and INGLIS, S. C. Virology 184, 531–544 (1991).
- 21. LOMNICZI, B., J. Gen. Virol. 36, 531-533 (1977).
- 22. RUSSEL, M., KIDD, S., and KELLEY, M. R., Gene 45, 333–338 (1986).
- SCHOCHETMAN, G., STEVENS, R. H., and SIMPSON, R. W., Virology 77, 772–782 (1977).
- 24. SEDMAN, S. A., and MERTZ, J. E., J. Virol. 62, 954-961 (1988).
- 25. SKINNER, M. A., EBNER, D., and SIDDELL, S. G., J. Gen. Virol. 66, 581–592 (1985).
- 26. SMITH, A. R., BOURSNELL, M. E. G., BINNS, M. M., BROWN, T. D. K., and INGLIS, S. C., J. Gen. Virol. 71, 3–11 (1990).
- 27. STANLEY, K. K., and LUZIO, J. P. (1984). EMBO J. 3, 1429-1434.
- 28. STERN, D. F., and KENNEDY, S. I. T., J. Virol. 34, 665–674 (1980).
- 29. STERN, D. F., and KENNEDY, S. I. T., J. Virol. 36, 440-449 (1980).
- 30. STERN, D. F., and SEFTON, B. M., J. Virol. 50, 22-29 (1984).
- 31. WILLIAMS, M. A., and LAMB, R. A., J. Virol. 63, 28-35 (1989).
- 32. WONG, T. C., WIPF, G., and HIRANO, A., *Virology* **157**, 497–508 (1987).