REGULATION OF mRNA 1 EXPRESSION BY THE 5'-UNTRANSLATED REGION (5'-UTR) OF THE CORONAVIRUS INFECTIOUS BRONCHITIS VIRUS (IBV)

D. X. Liu, H. Y. Xu, and K. P. Lim

Institute of Molecular Agrobiology
National University of Singapore
59A The Fleming
1 Science Park Drive
Singapore 118240

1. ABSTRACT

In this report, we show that expression of the coronavirus IBV mRNA1 is regulated by its 5'-UTR. Evidence presented domonstrates that the IBV sequence from nucleotide 1 to 1904 directs very inefficient synthesis of a product of approximately 43 kDa. Deletion of either the first 362 bp or the whole part of the 5'-UTR, however, dramatically increased the expression of the 43 kDa protein species. The mechanisms involved were investigated by two different approaches. Firstly, translation of the same construct in the presence of $[^3H]$-leucine ruled out the possibility that initiation of small reading frames from non-AUG codons located in the 5'-UTR may compete with the authentic AUG initiation codon, and therefore inhibit the expression of ORF 1a. Secondly, expression and deletion analyses of a dicistronic construct showed that translation of the 43 kDa protein was initiated by ribosome internal entry mechanism. These studies suggest that a 'weak' ribosome internal entry signal is located in the 5'-UTR and is involved in the regulation of mRNA1 expression.

2. INTRODUCTION

Six mRNA species, including the genome-length mRNA (mRNA1) of 27.6 kb and five subgenomic mRNA species (mRNAs 2–6) with sizes ranging from 2 to 7 kb, are produced in cells infected with the prototype virus of the Coronaviridae, avian infectious

bronchitis virus (IBV). Nucleotide sequencing of the genomic RNA has shown that the 5'-terminal unique region of mRNA 1 contains two large ORFs (1a and 1b), with potential to encode a fusion polyprotein of 741 kDa by a ribosomal frameshift (Boursnell *et al.*, 1987; Brierley *et al.*, 1987, 1989). This 1a/1b fusion polyprotein is expected to be cleaved by viral or cellular proteinases to produce functional products associated with viral RNA replication (Gorbalenya *et al.*, 1989).

Examination of the nucleotide sequence of mRNA 1 shows that a 527bp untranslated region is located at its 5' end. During the studies of the expression of mRNA1, it was consistently observed that this messenger RNA was very poorly expressed both in *in vitro* and *in vivo* expression systems unless the majority of the 5'-UTR was deleted from the constructs. In this report, we present evidence demonstrating that the 5'-UTR may contain a 'weak' ribosome internal entry signal, which is likely involved in the regulation of mRNA 1 expression.

3. MATERIALS AND METHODS

3.1. SDS-Polyacrylamide Gel Electrophoresis

SDS-polyacrylamide gel electrophoresis (SDS-PAGE) of virus polypeptides was carried out with a range of polyacrylamide concentrations from 12.5 to 17.5 % (Laemmli, 1970). Labelled polypeptides were detected by autoradiography or fluorography of dried gels.

3.2. Cell-Free Transcription and Translation

Plasmid DNA was transcribed and translated *in vitro* by using the TnT T7-coupled reticulocyte lysate system (TnT system) according to the instructions of the manufacturer (Promega). Reaction products were separated by SDS-PAGE and detected by autoradiography.

3.3. Polymerase Chain Reaction (PCR)

Appropriate primers and template DNAs were used in amplification reactions with Pfu DNA polymerase (Stratagene) under standard buffer conditions with 2 mM $MgCl_2$. The PCR conditions were 30 cycles of 92°C for 30 sec, 56°C for 30 sec and 72°C for 6 min.

3.4. Construction of Plasmids

Plasmid pIBV1a5, which covers the IBV sequence from nucleotide 1 to 1904, was constructed by cloning an ApaI- and EcoRI-digested PCR fragment into ApaI/EcoRI digested pPCRII vector (Invitrogen). The ApaI site is introduced by the upstream PCR primer and EcoRI cuts the IBV sequence at nucleotide 1904. The sequence of the upstream primer is 5'-ACTAGGGCCCACTTAAGATAGATATTAA-3', and the sequence of the downstream primer is 5'-TTCCATATGCAAGCTTCCAGA-3', which is complementary to the IBV sequence from nucleotide 1909 to 1929.

Plasmid pBP5Δ5, which covers the IBV sequence from nucleotide 11306 to 11877, was described previously (Liu *et al.*, 1997).

For construction of pBP5–1a5, a 1904 bp fragment covering the IBV sequence from nucleotide 1 to 1904 was generated by digestion of pIBV1a5 with ApaI, end-repair with Klenow and re-digestion with EcoRI. This fragment was then cloned into pBP5Δ5, which was firstly digested with BamHI (the BamHI site is located immediately downstream of the viral sequence), end-repaired with Klenow and re-digested with EcoRI, giving pBP5–1a5.

Plasmid pBP5–1a5Δ1 was constructed by cloning a BamHI- and EcoRI-digested PCR fragment into BamHI/EcoRI-digested pBP5Δ5. This PCR fragment covers the IBV sequence from nucleotide 140 to 1904 and was generated by using pIBV1a5 as the template and oligonucleotide 5'-CTGGATGGATCCTGGCCACCTG-3' as the upstream primer. The same downstream primer used to clone pIBV1a5 was used to generate this PCR fragment as well as PCR fragments for constructing pBP5–1a5Δ2 and pBP5–1a5Δ3. Plasmid pBP5–1a5Δ2 was constructed by cloning a BamHI- and EcoRI-digested PCR fragment, covering the IBV sequence from nucleotide 528 to 1904, into BamHI/EcoRI-digested pBP5Δ5. The sequence of the upstream primer is 5'-ACGCGGATCCAC-CATGGCTTCAAGC-3'. Plasmid pBP5–1a5Δ3 was constructed by cloning a BamHI- and EcoRI-digested PCR fragment, covering the IBV sequence from nucleotide 288 to 1904, into BamHI/EcoRI-digested pBP5Δ5. The sequence of the upstream primer is 5'-GTGCGAGGATCCTCTGGTTCA-3'. Plasmid pBP5–1a5Δ4 was constructed by digestion of pBP5–1a5Δ1 with BamHI and BstEII (BstEII cuts the IBV sequence at nucleotide 364), end-repair with Klenow and re-ligation with T4 DNA ligase.

Plasmid pIBV1a2, which was formerly called pKT1a2 and covers the IBV sequence from nucleotide 362 to 5753, was described previously (Liu et al., 1995). Plasmid pIBV1a2Δ2, which covers the IBV sequence from nucleotide 528 to 5753, was constructed as follows. A BamHI/MluI digested PCR fragment covering the IBV sequence from nucleotide 528 to 3997 was cloned into BglII/MluI digested pIBV1a2, giving the deletion construct pIBV1a2Δ2. MluI cuts the IBV sequence at nucleotide 3997. The upstream PCR primer is the same one used to construct pBP5–1a5Δ2; the sequence of the downstream primer is 5'-TCCAAATTGACCCAATGAGTGTC-3', which is complementary to the IBV sequence from nucleotide 4117 to 4140.

4. RESULTS

4.1. Very Inefficient Expression of Constructs Containing the 5'-Most 2 kb Part of mRNA1

During the course of studying the expression of IBV mRNA1, it was consistently observed that constructs containing the 527 bp 5'-UTR of mRNA1 exhibited very low expression efficiency. Figure 1 shows an example. As can be seen, expression of EcoRI-digested plasmid pIBV1a5, which covers the IBV sequence from nucleotide 1 to 1904, resulted in very inefficient expression of a polypeptide of approximately 43 kDa. Only a trace amount of the 43 kDa protein was observed after prolonged exposure of the gel (Figure 1). Deletion of the 5'-UTR, however, dramatically increased the expression of the 43 kDa protein species. As shown in Figure 1, expression of plasmids pIBV1a2 and pIBV1a2Δ2, which contain deletions of the IBV sequences from nucleotide 1 to 346 and from 1 to 526, respectively, led to much more efficient expression of the 43 kDa protein. It was of interest to investigate the mechanisms.
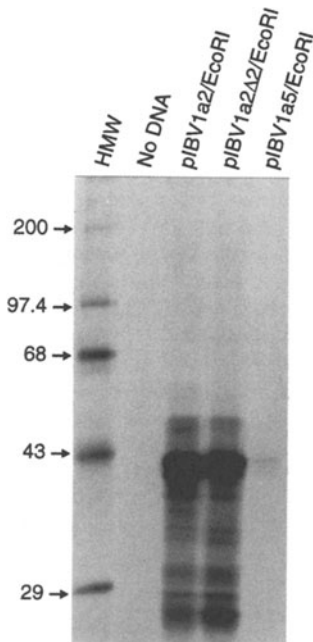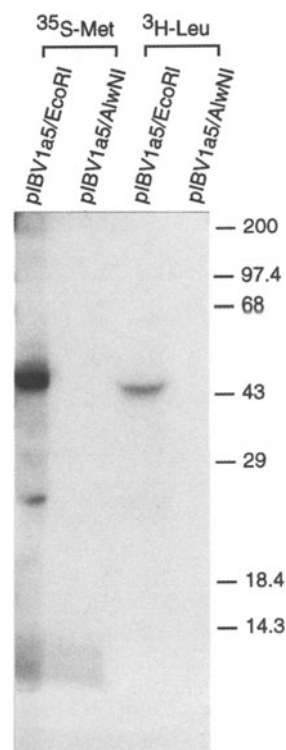
Figure 1. Analysis of *in vitro* translation products of RNAs obtained by *in vitro* transcription of EcoRI-digested pIBV1a2, pIBV1a2Δ2 and pIBV1a5 with T7 RNA polymerase. RNA was added to reticulocyte lysate as indicated above each lane. [^{35}S] methionine-labelled translation products were separated on an SDS-12.5% polyacrylamide gel and detected by fluorography. HMW-high molecular weight markers (numbers indicate molecular mass in kilodaltons).

4.2. No Initiation of Small Reading Frames from Non-AUG Initiation Codons Located in the 5'-UTR

Examination of the 5'-UTR sequence shows that several small reading frames (over 50 amino acids) are located in the IBV sequence from nucleotide 1 to 450. A prominent example is a 261 bp reading frame from nucleotide 138 to 398. Translation of this reading frame would result in the synthesis of a polypeptide with 86 amino acid residues, if a CUG triplet (from nucleotide 138 to 140) could be used as a translation initiation codon. This possibility was investigated by *in vitro* translation of pIBV1a5 in the presence of [^{3}H]-Leucine. For this purpose, pIBV1a5 was linearised by digestion with AlwNI and EcoRI, which cut the IBV sequences at nucleotide positions 486 and 1904, respectively, and was translated in reticulocyte lysate in the presence of either [^{35}S]-methionine or [^{3}H]-leucine. As shown in Figure 2, expression of EcoRI-linearised pIBV1a5 led to the detection of the 43 kDa protein species. No specific product was detected from expression of AlwNI-digested pIBV1a5. This result rules out the possibility that the inefficient expression of mRNA1 is caused by expression of small reading frames located in the 5'-UTR.

4.3. *In Vitro* Translation of Capped and Uncapped Transcripts Containing the 5'-UTR

In vitro synthesised capped and uncapped RNAs were then translated *in vitro* in reticulocyte lysate to compare the translatability of the two transcripts. For this purpose, pIBV1a5 was linearised with EcoRI and transcribed *in vitro* with T7 bacteriophage RNA
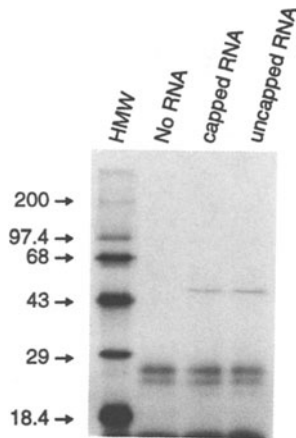
Figure 2. Analysis of *in vitro* translation products of RNAs obtained by *in vitro* transcription of EcoRI- and AlwNI-digested pIBV1a5 with T7 RNA polymerase. RNA was added to reticulocyte lysate and translated in the presence of [^{35}S]-methionine and [^{3}H]-leucine as indicated above each lane. The radio-labelled translation products were separated on an SDS-17.5% polyacrylamide gel and detected by fluorography. Numbers indicate molecular mass in kilodaltons.

polymerase. Equal amount of *in vitro* synthesised RNAs were used to translate *in vitro* in reticulocyte lysate. As shown in Figure 3, both transcripts direct inefficient expression of the 43 kDa protein. Furthermore, the amounts of the 43 kDa protein species produced from both transcripts were very similar. This result indicates that the cap structure of mRNA1 may not be essential for its expression and suggests that translation of mRNA1 may be initiated by cap-independent mechanism.

4.4. Cap-Independent Translation of mRNA1

To study further if expression of mRNA1 is cap-independent, the IBV sequence from nucleotide 1 to 1,904 was cloned downstream of a reporter gene covering the IBV sequence from nucleotide 11306 to 11877, giving rise to a dicistronic construct pBP5–1a5 (Figure 4a). Translation of plasmids pIBV1a5 and pBP5Δ5 in reticulocyte lysate led to the detection of two proteins with apparent molecular weights of 43 kDa and 25 kDa, respectively, representing the full-length products encoded by the two constructs (Figure 4b). Expression of EcoRI-linearised pBP5–1a5 also resulted in the detection of two protein species of approximately 25 kDa and 43 kDa. The 25 kDa protein co-migrates with the 25 kDa protein translated from pBP5Δ5 and the 43 kDa product co-migrates with the 43 kDa protein expressed from pIBV1a5 (Figure 4b), suggesting that they may represent the prod-
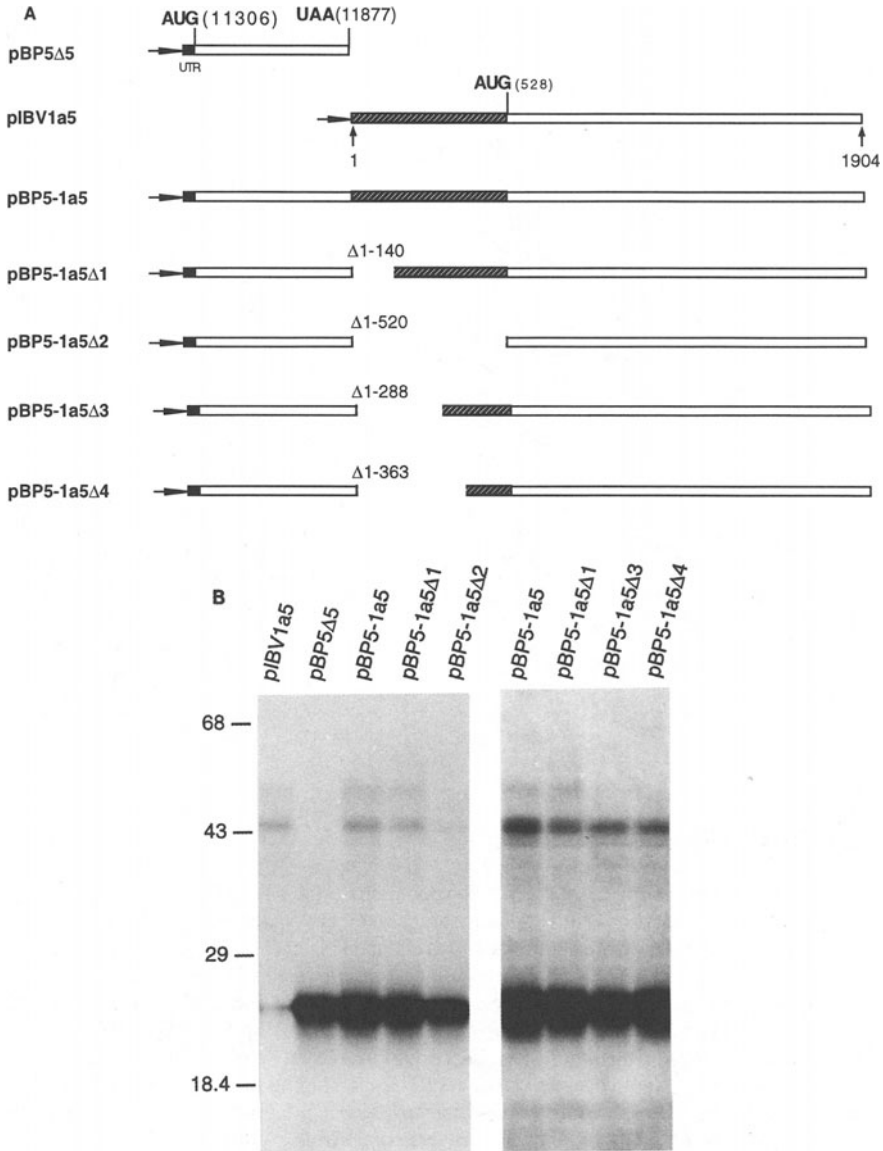
Figure 3. Analysis of *in vitro* translation products of capped and uncapped RNAs obtained by *in vitro* transcription of EcoRI-digested pIBV1a5 with T7 RNA polymerase. The capped RNA was prepared by incorporating the dinucleotide ^{7m}GpppG to provide a 5' cap structure (Contreras *et al.*, 1982). RNA was added to reticulocyte lysate and translated in the presence of $[^{35}S]$-methionine. The radio-labelled translation products were separated on an SDS-12.5% polyacrylamide gel and detected by fluorography. HMW-high molecular weight markers (numbers indicate molecular mass in kilodaltons).

ucts translated from the upstream and downstream ORFs, respectively. These results indicate that both the upstream and downstream ORFs were expressed from the dicistronic construct. Deletion of the 5'-UTR from pBP5–1a5, however, abolished the expression of downstream ORF expression. As can be seen from Figure 4b, only the 43 kDa protein was observed from transcription and translation of pBP5–1a5Δ2. Deletion of the 5'-most 140 bp of the 5'-UTR, however, did not affect the expression of the downstream ORF. Only slight reduction of the 43 kDa protein expression was observed from transcription and translation of pBP5–1a5Δ1 (Figure 4b). These results confirm that expression the 43 kDa protein from the dicistronic construct is dependent on the presence of the 5'-UTR.

Two more deletion constructs were made to define the sequence required for the internal translation of the 43 kDa protein. Plasmids pBP5–1a5Δ3 contains a deletion of the IBV sequence from nucleotide 1 to 288, and pBP5–1a5Δ4 contains a deletion from nucleotide 1 to 364 (Figure 4a). Expression of both constructs showed that very similar amounts of the 43 kDa protein was expressed (Figure 4b).

4.5. Expression of the Full-Length Product Encoded by the 5'-Most 5753 bp Part of ORF1a Is Increased by the Presence of the Second Half of the 5'-UTR

Data presented above showed that a weak internal initiation signal is located in the 5'-UTR. In an attempt to increase the expression efficiency of the ORF 1a product, plasmid pIBV1a2Δ2 was constructed. This construct covers the IBV sequence from nucleotide 528 to 5753 and therefore excludes the whole 5'-UTR. Expression of the construct in reticulocyte lysate, however, resulted in the synthesis of much less full-length product than that from pIBV1a2. As can be seen from Figure 5, when equal amount of MluI- and BamHI-digested pIBV1a2 and pIBV1a2Δ2 were expressed *in vitro*, both plasmids can direct synthesis of the full-length products of approximately 180kDa and 250 kDa, respectively. However, significantly more efficient expression of the 180 kDa and 250 kDa products was observed from translation of pIBV1a2 than that from pIBV1a2Δ2 (Figure 5). Meanwhile, when EcoRI-digested pIBV1a2 and pIBV1a2Δ2 were expressed, no obvious difference in the synthesis of the 43 kDa products was detected.

Figure 4. (A). Diagram showing the structures of plasmids pBP5Δ5, pIBV1a5, pBP5–1a5, pBP5–1a5Δ1, pBP5–1a5Δ2, pBP5–1a5Δ3 and pBP5–1a5Δ4. (B). Analysis of *in vitro* translation products of RNAs obtained by *in vitro* transcription of EcoRI-digested pIBV1a5, pBP5Δ5, pBP5–1a5, pBP5–1a5Δ1, pBP5–1a5Δ2, pBP5–1a5Δ3 and pBP5–1a5Δ4 with T7 RNA polymerase. RNA was added to reticulocyte lysate and translated in the presence of [³⁵S]-methionine. The radio-labelled translation products were separated on an SDS-12.5% polyacrylamide gel and detected by fluorography. Numbers indicate molecular mass in kilodaltons.
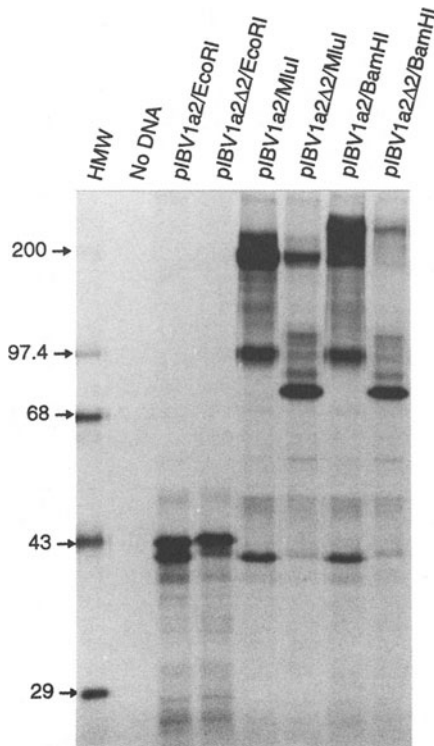
Figure 5. Analysis of *in vitro* translation products of RNAs obtained by *in vitro* transcription of EcoRI-, MluI- and BamHI-digested pIBV1a2 and pIBV1a2Δ2 with T7 RNA polymerase. RNA was added to reticulocyte lysate and translated in the presence of [^{35}S]-methionine. The radio-labelled translation products were separated on an SDS-10% polyacrylamide gel and detected by fluorography. HMW-high molecular weight markers (numbers indicate molecular mass in kilodaltons).

5. DISCUSSION

The genome-length mRNA, mRNA 1, contains a long 5'-UTR. In this report, we show that a 'weak' ribosome internal entry signal may be located in this region and is involved in the regulation of IBV mRNA 1 expression. This conclusion is based on the following observations. Firstly, *in vitro* translation of the IBV sequence from nucleotide 1 to 1904 resulted in very inefficient synthesis of a 43 kDa protein. Expression of this protein, however, was dramatically increased by deletion of either 362 bp from the 5'-UTR (from nucleotide 1 to 362) or the whole 5'-UTR (from nucleotide 1 to 527). Secondly, *in vitro* translation of capped and uncapped artificial RNAs containing the whole 5'-UTR showed no obvious difference in the translation efficiency between the two RNAs, indicating that the cap structure of mRNA 1 may be not essential for the initiation of translation of ORF1a and further suggesting that ribosomes may bypass the cap structure and bind to some internal sites of the 5'-UTR. Furthermore, construction and expression of a dicistronic plasmid showed that the 5'-UTR can direct the expression of the downstream ORFs from the dicistronic constructs. Deletion studies confirmed that the expression of the downstream ORF was dependent on the presence of the 5'-UTR.

Like most other eukaryotic mRNAs, IBV mRNA1 is naturally capped. In this report, we show that this cap structure may be not essential for translation initiation of the

mRNA1 expression. Instead of using the conventional cap-dependent translation initiation mechanism, expression of this mRNA may be initiated by ribosome internally binding to the 5'-UTR. Cap-independent translation initiation of other naturally capped mRNAs of coronaviruses was also reported. These include translation of the third ORF of IBV mRNA3 and the second ORF of mouse hepatitis virus mRNA5 (Liu and Inglis, 1992; Thiel and Siddell, 1994). It is currently unclear why coronaviruses utilise both cap-dependent and cap-independent mechanisms to initiate the translation of their individual messenger RNAs, but this may reflect a subtle regulatory mean employed by virus to control the individual gene expression during the virus life cycle.

It is intriguing to observe that when the IBV sequence up to nucleotide 5753 was expressed *in vitro*, much more full-length product was synthesised from the construct containing the 3' end 160 bp part of the 5'-UTR (pIBV1a2) than that from the construct without the 5'-UTR (pIBV1a2Δ2). However, very similar amount of the full-length product was synthesised from both constructs when transcripts covering the IBV sequence only up to nucleotide 1904 was expressed. These results indicate that there is no obvious differences in the stage of translation initiation between the two constructs. It is therefore likely that the presence of the 160 bp leader sequence in the run-off transcripts transcribed from pIBV1a2 may stabilize the *in vitro* synthesised RNA, and consequently increase the synthesis of the full-length product. Further investigations are underway to address this possibility.

REFERENCES

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.

Brierley, I., M. E. G. Boursnell, M. M. Binns, B. Bilimoria, V. C. Blok, T. D. K. Brown, and S. C. Inglis., 1987, An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV, *EMBO. J.* **6**:3779–3785.

Brierley, I., P. Digard, and S. C. Inglis., 1989, Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot, *Cell* **57**:537–547.

Contreras, R., H. Cheroutre, W. Degrave, and W. Fiers., 1982, Simple efficient *in vitro* synthesis of capped RNA useful for direct expression of cloned DNA, *Nucleic Acids Res.* **10**: 6353–6362.

Gorbalenya, A. E., E. Y. Koonin, A. P. Donchenko, and V. M. Blinov., 1989, Coronavirus genome: prediction of putative functional domains in the non-structural polyprotein by comparative amino acid sequence analysis, *Nucleic Acids Research* **17**:4847–4860.

Laemmli, U. K., 1970, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature (London)* **227**:680–685.

Liu, D. X., and S. C. Inglis., 1992, Internal entry of ribosomes on a tricistronic mRNA encoded by infectious bronchitis virus, *J. Virol.* **66**: 6142–6154.

Liu, D. X., K. W. Tibbles, D. Cavanagh, T. D. K. Brown, and I. Brierley., 1995, Identification, expression and processing of an 87 kDa polypeptide encoded by ORF1a of the coronavirus infectious bronchitis virus, *Virology* **208**: 48–57.

Liu, D. X., and T. D. K. Brown. 1997. Proteolytic processing of the coronavirus infectious bronchitis virus 1a polyprotein: identification of a 10-kilodalton polypeptide and determination of its cleavage sites, *J. Virol.* **71**: 1814–1820.

Thiel, V., and S. G. Siddell. 1994. Internal ribosome entry in the coding region of murine hepatitis virus mRNA 5, *J. Gen. Virol.* **75**: 3041–3046.