ORGANIZATION OF THE IBV GENOME

David F. Stern* and Bartholomew Sefton

Molecular Biology and Virology Laboratory
The Salk Institute
P.O. Box 85800
San Diego, CA 92138
*Center for Cancer Research
Massachusetts Institute of Technology
Cambridge, MA 02139

We have investigated how the information contained within the large RNA genome of IBV is expressed. We began by examining the structure of IBV-specified RNAs. Infected chicken embryo kidney (CEK) cells contain at least 6 IBV RNA species. These consist of the viral genome, RNA F, with an estimated complexity of 23 kb[1], and subgenomic RNAs A, B, C, D, and E, which range in size from 2.4 kb to 7.9 kb[2]. We have recently identified an additional IBV-specified RNA species, RNA M, by Northern blot analysis using a hybridization probe containing cloned IBV cDNA sequences (see below). RNA M is intermediate in size between RNAs B and C (Table 1). Structural analysis of RNAs A, B, C, D, E, and F by ribonuclease T_1 fingerprinting revealed that they comprise a 3' coterminal nested set[3]. Synthesis of 3' coterminal RNAs is now known to be a characteristic feature of coronavirus multiplication[4,5].

Since the IBV intracellular RNAs were single-stranded, of positive sense, and polyadenylated, they were likely to function as viral mRNAs. This raised the question of the location of translationally active sequences within each mRNA. It was possible that each is translated over most of its length, yielding polypeptides proportional in size to the mRNAs. Another possibility, suggested by the activities of the overlapping alphavirus mRNAs, was that only the "unique" 5' domain of each mRNA not contained in smaller mRNAs would be translated. The size of the polypeptides produced would be proportional to the difference in size between each mRNA and the next smaller species, rather than to its

P. J. M. Rottier et al. (eds.), *Molecular Biology and Pathogenesis of Coronaviruses*

Table 1. Coding Capacities of IBV mRNAs Predicted from Two Different Models

RNA	length (kb)	total coding capacity (kd)[a]	"non-overlapping" coding capacity (kd)[b]	size of translation product (kd)[c]
A	2.4	80	80	51
B	2.7	90	10	
M	3.2	110	20	
C	3.9	130	20	23
D	4.5	150	20	
E	7.9	260	110	110
F	23.	800	540	

[a]The molecular weight of each RNA was divided by 10 to estimate the size of a polypeptide encoded by the entire RNA.

[b]From the molecular weight of each RNA species the molecular weight of the next smaller species was subtracted, yielding the size of the unique 5' domain. This figure was divided by 10 to determine the coding capacity of that region.

[c]Determined by cell-free translation of fractionated mRNA as described in the text.

absolute size[3]. The sizes of polypeptides predicted according to this non-overlapping translational scheme are listed in Table 1. We describe here the use of cell-free translation to identify the mRNAs that encode the major viral structural proteins.

IBV virions prepared according to our procedures contain P14, which has not been localized within virions, P51, the nucleocapsid protein, and three distinct membrane proteins, GP90, GP84, and the P23 family proteins[8,9,10,11,12,13]. One or both of the large glycoproteins, GP90 and GP84, comprise the large virion surface projections. The P23 family of proteins consists of P23, which is not glycosylated, and glycoproteins GP28, GP31, and GP36. These proteins all contain the same 23 kd core polypeptide and differ in the number and type of N-linked oligosaccharides which they bear[7].

The P23 family of proteins and P51 are synthesized without obvious post-translational proteolytic modification. However, GP90 and GP84 are derived by post-translational cleavage of a cell-associated precursor, GP155[14]. The approximate size of the core polypeptide of GP155 was determined by removal of most of the carbohydrate with endoglycosidase H. This digestion product had an approximate molecular weight of 115 kd[7]. P14 was not reliably detected in infected cells, so we were unable to determine the manner of its synthesis [14]. We were able to conclude that 7 virion proteins are derived by maturation of only three polypeptides, which have molecular weights of 23 kd, 51 kd, and approximately 115 kd.

To determine which mRNAs specify these polypeptides, polyadenylated RNA from infected cells was translated in a messenger-dependent rabbit reticulocyte lysate[15]. Fractionation of the RNA prior to translation permitted the correlation of messenger activity with the presence of particular mRNAs.

Cell-free translation of total polyadenylated RNA from IBV-infected cells yielded products which comigrated with virion proteins P23 and P51 (Fig. 1, lanes I and V). These proteins were not produced by translation of RNA from uninfected cells (Fig. 1, lane M). Maps of methionine-labelled tryptic peptides demonstrated the identity of the 23 kd and 51 kd translation products (Fig. 2, panels E and A, respectively) with their counterparts purified from infected cells (Fig. 2, panels F and C). A series of proteins just larger than P51 was also produced by translation of RNA from infected cells (Fig. 1, lane I). Two-dimensional tryptic peptide mapping (Fig. 2, panel B) and one-dimensional partial proteolytic mapping (data not shown) of these proteins showed that they are closely related to P51. The reason for the polymorphism of P51 synthesized in vitro is not clear. It could, in principle, result from the use of multiple sites for initiation or termination of translation.

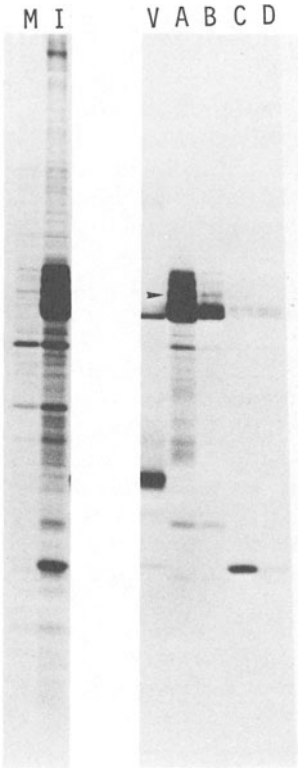


Fig. 1. In vitro translation of gel-purified IBV RNAs. Polyadenylated RNA labeled biosynthetically with $^{32}P_i$ was purified from cells. The RNA was fractionated on a preparative 2% acrylamide-0.1% bisacrylamide gel. IBV intracellular RNAs were eluted and translated in a messenger-dependent rabbit reticulocyte lysate in the presence of ^{35}S-methionine. The products were analyzed by SDS-polyacrylamide gel electrophoresis. Lane M, translation products of RNA from uninfected cells; lane I, translation products of nonfractionated RNA from infected cells. Lane V, virion proteins; lanes A, B, C, and D, translation products of gel-purified IBV RNAs A, B, C, and D, respectively.

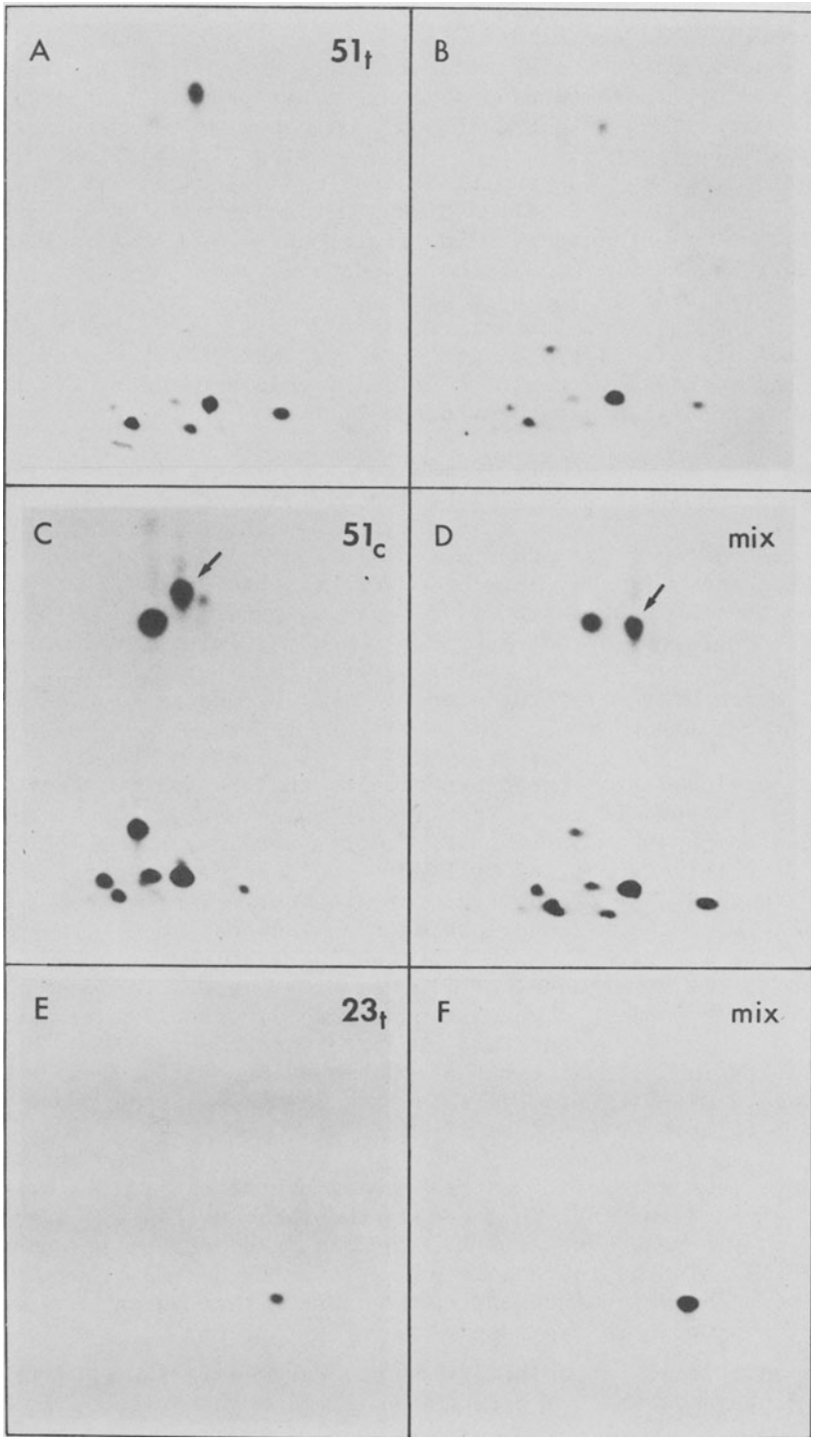
RNAs A, B, C, and D were purified by preparative acrylamide gel electrophoresis and translated in the reticulocyte lysate. P51 and P23 were produced by translation of RNAs A and C, respectively (Fig. 1, lanes A and C). There were no obvious virus-specific translation products associated with RNAs B and D. Because we were concerned about possible degradation of the mRNA during elution from the preparative gel, we also purified viral mRNA by a gentler technique. Infected cell RNA was labelled biosynthetically with ^{3}H-uridine in the presence of actinomycin D and fractionated by velocity sedimentation. The RNA in each fraction was analyzed by agarose gel electrophoresis (Fig. 3, upper panel) and compared with the corresponding cell-free translation products (Fig. 3, lower panel). Production of P51 and P23 again correlated with the presence of RNAs A and C. However, the resolution of this experiment was not sufficient to distinguish RNAs A and B.

The mRNA encoding GP155 remained to be identified. RNA E was an obvious candidate. RNA E, labelled biosynthetically with ^{3}H-uridine was purified by velocity sedimentation (Fig. 4, lane a). Translation of this RNA preparation yielded a 110 kd polypeptide designated P110 (Fig. 4, lane b). Comparison of maps of methionine-labelled tryptic peptides of P110 (Fig. 5, panel A) and GP155 (Fig. 5, panel B) demonstrated that they are closely related. Production of P51 by translation of this RNA preparation can perhaps be attributed to activation of the internal gene for P51 by degradation of the RNA during translation.

We concluded from these experiments that RNAs A, C, and E encode 51 kd, 23 kd, and 110 kd polypeptides, respectively, which are processed to produce P51, the P23 family proteins, and GP155. The sizes of proteins specified by RNAs C and E do not support a model in which each RNA is translated over its entire length, but are compatible with the non-overlapping translational model (Table 1). A map of the IBV genome based upon the non-overlapping scheme is depicted in the upper panel of Fig. 6. Viral genes are demarcated by loci corresponding to the 5' ends of the intracellular RNAs. The predicted sizes of unidentified products encoded by RNAs B, M, D, and F are listed in Table 1. A map of the mouse hepatitis virus (MHV) genome, determined by others[16,17,18,19] is shown in the lower panel of Fig. 6 for comparison.

There is a major difference between the transcription of MHV and IBV. IBV mRNA E and MHV RNA 3 encode the large virion glycoproteins of the two viruses. However, MHV produces an additional larger sub-genomic RNA which has no homologue in IBV. IBV and MHV thus differ in the means by which the gene upstream from that which encodes the large glycoprotein is expressed.

An additional potential difference between IBV and MHV is the location of genes which encode the small membrane protein. In MHV,


A
51_t
B
C
51_c
D
mix
E
23_t
F
mix

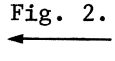


Fig. 2. Tryptic peptide maps of P51 and P23 synthesized in vitro. P23, P51, and a larger member of the P51 series, produced by cell-free translation of RNA from IBV-infected cells in the presence of ^{35}S-methionine, were purified in a preparative gel. The proteins were eluted, digested with trypsin, and peptides were resolved in two dimensions on thin-layer cellulose plates[6]. Electrophoresis in the first dimension was from left to right. Ascending chromatography in the second dimension was from bottom to top. The large member of the P51 series is marked in Fig. 1, lane A, with an arrow. Cell-associated forms of P51 and GP31 labeled with ^{35}S-methionine were mapped for comparison. A map of the preparation of GP31 used in the mixture in panel F has been published[14]. Panel A, P51 synthesized in vitro; panel B, larger member of the P51 series, synthesized in vitro; panel C, P51 synthesized in vivo; panel D, mixture of peptides from P51 synthesized in vitro and in vivo; panel E, P23 synthesized in vitro; panel F, mixture of peptides from P23 synthesized in vitro and GP31 synthesized in vivo.

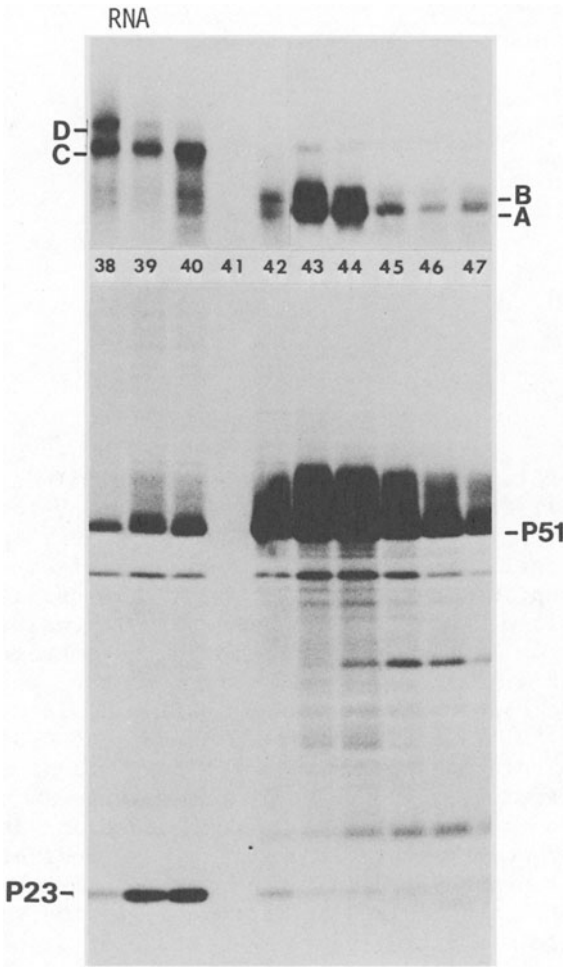


Fig. 3. Translation in vitro of gradient-fractionated RNA. Polyadenylated RNA was purified from IBV-infected cells which had been labeled with ^{3}H-uridine in the presence of actinomycin D from 1 to 8 h post-infection. The RNA was fractionated by sedimentation in an 11.6 ml 15-30% sucrose-TLES gradient at 150,000 x g for 4.5 h at 10°C in a SW41 rotor. 56 fractions, numbered consecutively from bottom to top, were collected. A portion of the RNA from each fraction was analyzed by agarose gel electrophoresis (upper panel). Another portion was translated in vitro and the products analyzed by SDS-polyacrylamide gel electrophoresis (lower panel). RNAs and corresponding translation products from gradient fractions 38 to 47 are shown. RNA in fraction 41 was evidently lost during handling.

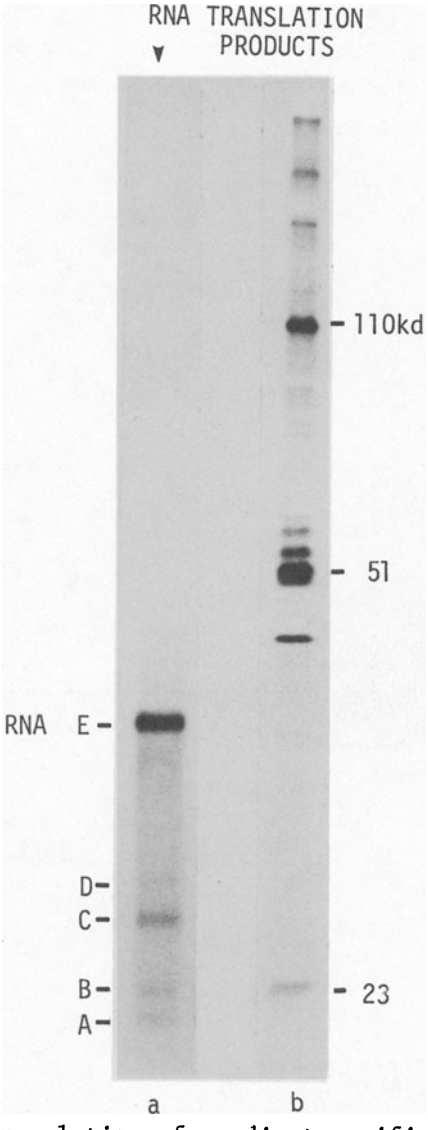


Fig. 4. In vitro translation of gradient-purified RNA E. Polyadenylated RNA, labeled with ^{3}H-uridine, was isolated from infected cells. The RNA was fractionated under the conditions described in the legend to Fig. 3. A portion of the RNA was analyzed by agarose gel electrophoresis and fractions containing RNA E were pooled. RNA in this pool was analyzed by agarose gel electrophoresis (lane a) and the in vitro translation products were analyzed by SDS-polyacrylamide gel electrophoresis (lane b).

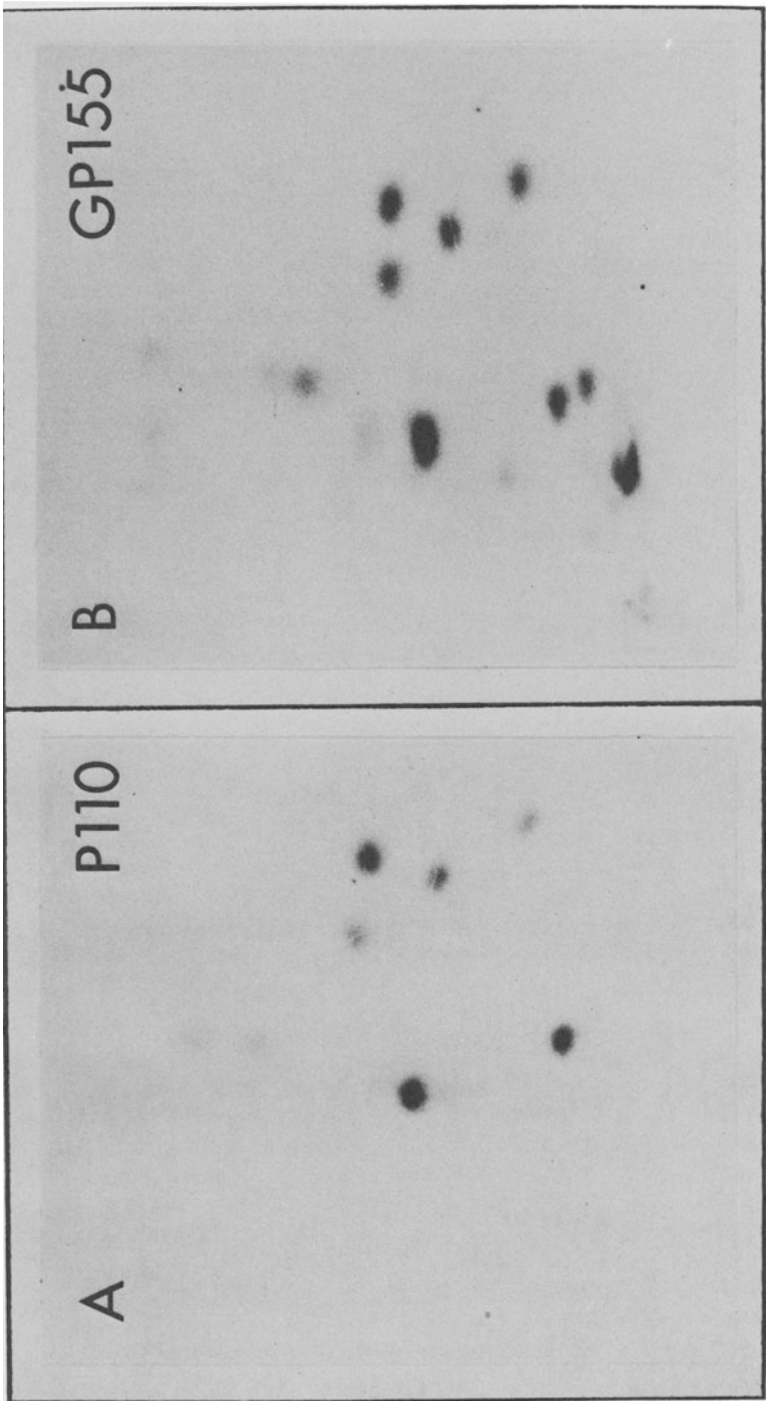


Fig. 5. Tryptic peptide maps of P110 and GP155. The RNA preparation shown in Fig. 4 was translated in vitro and P110 was purified. The map of GP155 labeled biosynthetically with ^{35}s-methionine has been published previously[14]. Panel A, P110; panel B, GP155.

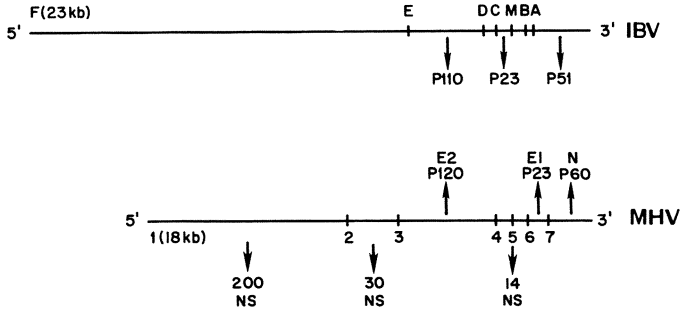


Fig. 6. Genetic maps of IBV and MHV. The genomes of IBV and MHV are depicted with the points corresponding to the 5' ends of the intracellular RNAs marked. It was assumed that the 5' boundary of each viral gene is near a point corresponding to the 5' terminus of a mRNA. The polypeptide encoded by each gene is designated by its molecular weight in kd. The map of MHV is based upon the work of others[16,17,18,19]

genes encoding E1 and the nucleocapsid protein are contiguous and expressed by translation of the two smallest mRNAs. In IBV, however, the locations of promoters for RNA B and RNA M suggest that two genes of unknown function separate the genes encoding P23 and the nucleocapsid protein.

Of the IBV virion proteins that are known to be structurally unique, only P14 remains unassigned to a mRNA. Completion of the genetic map of IBV will therefore depend upon identification of IBV nonstructural proteins.

We have obtained molecular clones of IBV sequences. These will greatly facilitate the detailed study of IBV multiplication. cDNA copies of the IBV genome were synthesized with virion RNA as a template and an oligo(dT) primer. Synthesis of the second strand, primed by spontaneous formation of hairpin loops, was catalyzed by the Klenow fragment of DNA polymerase I. The hairpin loop was opened with nuclease S_1 and deoxycytidylate tails were added with terminal transferase. The DNA was annealed to pBR322 that had been linearized by cleavage with Pst I (interrupting the ampicillin-resistance gene) and tailed with deoxyguanidylate. The recombinant plasmids were introduced into E. coli strain C600 by transformation. Plasmids were purified from 144 tetracycline-resistant, ampicillin-sensitive colonies and compared according to size. The largest, pIBV5, contained 3.3 kb of inserted sequences. In Northern blotting experiments, the insert hybridized to IBV mRNAs but not to RNA from uninfected cells. This confirms the viral origin of the inserted sequences. The presence of a deoxyadenylate tract at one end of the insert suggests that this terminus corresponds to the exact 3' end of the IBV genome. Preliminary data indicate that the 5'-most sequences contained in pIBV5 consist of an open reading frame which extends beyond the 5' end of the clone. This reading frame encodes an extremely hydrophobic protein which is almost certainly P23.

pIBV5 should prove useful in investigation of many problems pertaining to coronavirus multiplication. The nucleotide sequence of the plasmid may reveal open reading frames which indicate the structure of proteins encoded by RNAs B and M. These proteins could then be isolated with antisera directed against synthetic peptides homologous with the predicted sequence. Purification of viral RNA by hybrid selection and mapping cell-free translation products to the genome by hybrid arrest translation will also aid identification of additional viral gene products.

Recent reports indicate that the mechanism of coronavirus RNA synthesis differs from that of other RNA viruses. MHV RNAs contain identical 5' terminal leader sequences which are evidently joined to the bodies of mRNAs by transcription of non-adjacent sequences[20,21]. The use of primers derived from pIBV5 for determining the 5' terminal sequences of RNAs A, B, and M will permit their direct comparison to

genomic sequences and reveal the structure of the leader-body junction. Further insight into IBV RNA synthesis may be gained by comparison of the promoters for synthesis of mRNAs A, B, and M, and of the negative-stranded template for mRNA synthesis, all of which should be contained in pIBV5.

REFERENCES

1. B. Lomniczi and I. Kennedy, J. Virol. 24:99-107 (1977).
2. D.F. Stern and S.I.T. Kennedy, J. Virol. 34:665-674 (1980).
3. D.F. Stern and S.I.T. Kennedy, J. Virol. 36:440-449 (1980).
4. S. Siddell, H. Wege, and V. ter Meulen, Curr. Topics in Microbiol. and Immunol. 99:131-163 (1982).
5. L.S. Sturman and K.V. Holmes, Adv. Virus Res. 28, in press (1983).
6. D.F. Stern, L. Burgess, and B.M. Sefton, J. Virol. 42:208-219 (1982).
7. D.F. Stern and B.M. Sefton, J. Virol. 44:804-812 (1982).
8. D. Cavanagh, J. Gen. Virol. 53:93-103, (1981).
9. J.A. Lanser and C.R. Howard, J. Gen. Virol. 46:349-361 (1980).
10. M.R. MacNaughton, M.H. Madge, H.A. Davies, and R.R. Dourmashkin, J. Virol. 24:821-825 (1977).
11. J.A. Lanser and C.R. Howard, J. Gen. Virol. 46:349-361 (1980).
12. H.A. Davies, R.R. Dourmashkin, and M.R. MacNaughton, J. Gen. Virol. 53:67-74 (1981).
13. C.N. Wadey and E.G. Westaway, Intervirology 15:19-27 (1981).
14. D.F. Stern and B.M. Sefton, J. Virol. 44:794-803 (1982).
15. H.R.B. Pelham and R.J. Jackson, Eur. J. Biochem. 67:247-256 (1976).
16. S. Siddell, H. Wege, A. Barthel, and V. ter Meulen, J. Virol. 33:10-17 (1980).
17. P.J.M. Rottier, W.J.M.Spaan, M.C. Horzinek, and B.A.M. Van der Zeijst, J. Virol. 38:20-26 (1981).
18. J.L. Leibowitz, S.R. Weiss, E. Paavola, and C.W. Bond, J. Virol. 43:905-913, (1982).
19. S. Siddell, J. Gen. Virol. 64:113-125 (1983).
20. M.C. Lai, C.D. Patton, R.S. Baric, and S.A. Stohlman, J. Virol. 46:1027-1033 (1983).
21. W. Spaan, P. Rottier, S. Smeekens, B.A.M. Van der Zeijst, H. Delius, J. Armstrong, M. Skinner, and S.G. Siddell, submitted for publication.