

Key words: coronavirus IBV/spike protein/DNA sequence/mRNA E

Cloning and Sequencing of the Gene Encoding the Spike Protein of the Coronavirus IBV

By MATTHEW M. BINNS,* MICHAEL E. G. BOURSNELL,
DAVID CAVANAGH, DARRYL J. C. PAPPIN¹
AND T. DAVID K. BROWN

Houghton Poultry Research Station, Houghton, Huntingdon, Cambs. PE17 2DA and

¹*Department of Biochemistry, University of Leeds, Leeds LS2 9JT, U.K.*

(Accepted 19 December 1984)

SUMMARY

RNA sequences encoding the surface projection (spike) of the coronavirus infectious bronchitis virus, strain Beaudette, have been cloned into pBR322 using cDNA primed with a specific oligonucleotide. A 5·3 kilobase viral insert in the clone pMB179 has been identified. The region of this clone coding for the spike gene has been sequenced by the chain termination method, and we present here the first report of DNA sequence data for a coronavirus spike protein, the protein which forms the characteristic 'corona' after which the group is named. The amino acid sequence of the primary translation product, deduced from the DNA sequence, predicts a polypeptide of 1162 amino acids with a molecular weight of 127006. This has many interesting features which confirm and extend our knowledge of this recently characterized membrane glycoprotein. The polypeptide is subsequently cleaved to S1 and S2, and partial amino acid analysis of the amino-terminus of the S1 polypeptide has been employed to locate the position of this terminus of S1 within the large open reading frame. The amino acid analysis also reveals the presence of an 18 amino acid putative signal sequence on the primary translation product which is not present on the mature S1 polypeptide.

INTRODUCTION

Infectious bronchitis virus (IBV) causes respiratory disease in the fowl and is of considerable economic importance to the poultry industry. The type species of the Coronaviridae, it possesses a single-stranded RNA genome, approximately 20 kb in length, of positive polarity which specifies the production of three major structural proteins: nucleocapsid protein, membrane glycoprotein, and spike glycoprotein. The spike protein, encoded by mRNA E, has recently been characterized (Cavanagh, 1983*a,b,c*) as comprising two or three copies each of two glycopolypeptides, S1 (90000 mol. wt.) and S2 (84000 mol. wt.). The polypeptide components of the glycopolypeptides S1 and S2 have been estimated after enzymic removal of oligosaccharides to have molecular weights of 64000 and 61000 (Cavanagh, 1983*a*). It appears that the spike protein is attached to the viral membrane by S2 (Cavanagh, 1983*c*). A neutralizing and haemagglutination-inhibiting monoclonal antibody produced against the spike protein binds the S1 glycopolypeptide, an effect which is strain-specific (Mockett *et al.*, 1984).

The organization of the IBV genome and subgenomic mRNAs has been studied in detail (Stern & Kennedy, 1980*a,b*; Stern & Sefton, 1984; Brown *et al.*, 1984; Brown & Bournsnel, 1984) and is summarized in Fig. 1. Using oligo(dT)-primed cDNA synthesis we have previously isolated 3·3 kb of overlapping cDNA clones extending from the 3' poly(A) tract (Brown & Bournsnel, 1984). We report here the use of a specific oligonucleotide to prime cDNA synthesis, which has allowed the isolation of a 5·3 kb viral insert containing the spike gene of IBV. The region of this clone containing the spike gene has been completely sequenced on both strands.

METHODS

cDNA cloning. The isolation of IBV strain Beaudette virion RNA has been described previously (Brown & Bournnell, 1984) as has the synthesis, by the phosphotriester method, of the specific oligonucleotide primer used to prime reverse transcription (Gait *et al.*, 1982; Bournnell *et al.*, 1984). cDNA synthesis was carried out using the method of Gubler & Hoffman (1983) with approximately 20 µg of virion RNA in a final reaction volume of 50 µl. Double-stranded cDNA was tailed with dC residues and cloned into dG-tailed *Pst*I-cleaved pBR322. This material was used to transform (Hanahan, 1983) *Escherichia coli* LE392 and selection made for tetracycline resistance. Clones containing viral inserts were identified by colony hybridization (Grunstein & Hogness, 1975) using polynucleotide kinase ³²P-labelled, alkali-treated IBV genomic RNA as a probe. The plasmid (pMB179) which was isolated from the clone showing the strongest signal in the colony hybrid experiment was studied in more detail.

Subcloning for M13 sequencing. Random subclones of pMB179 were generated by cloning either DNase I (Anderson, 1981) or sonicated (Deininger, 1983) fragments into *Sma*I-cut, phosphatase-treated M13mp10 (Amersham). Clones containing viral inserts were identified by colony hybridization with kinase-labelled or reverse-transcribed viral probes. In addition, *Pst*I and *Rsa*I fragments were cloned into *Pst*I-digested M13mp11 and *Sma*I-cut, phosphatase-treated M13mp10 respectively.

DNA sequencing. M13/dideoxynucleotide sequencing (Sanger *et al.*, 1977) was carried out using [α -³⁵S]dATP (Amersham), the complete sequence being obtained on both strands. Reverse sequencing was used to obtain the last sequences required (Hong, 1981). The products of the sequencing reactions were analysed on buffer gradient gels (Biggin *et al.*, 1983). A sonic digitizer (Graf/Bar, Science Accessories Corporation) was used to read data into a BBC microcomputer, and data were analysed on a VAX 11/750, using the programs of Staden (1982, 1984).

Isolation of S1 polypeptide and partial amino acid analysis. Plaque-purified IBV Beaudette was radiolabelled with [³H]serine (Amersham) in chick kidney cells (Stern *et al.*, 1982) and purified as described previously (Cavanagh, 1981). Viral polypeptides were resolved by SDS-polyacrylamide gel electrophoresis in 5 to 10% gels which were fluorographed without fixation. The S1 polypeptide was eluted from the gels by electrophoresis (Welch *et al.*, 1981), extensively dialysed against distilled water containing 0.03% SDS and lyophilized. The powdered protein was dissolved in 200 µl of 0.1 M-sodium bicarbonate containing 4% SDS and added to 100 mg of *p*-phenylenediisothiocyanate-treated glass (17 nm pore size) prepared by the method of Wachter *et al.* (1973). Following incubation for 90 min at 56 °C under nitrogen the glass was washed with water and methanol to remove non-covalently bound material. The glass-coupled peptide was then sequenced by automated solid-phase Edman degradation (Brett & Findlay, 1983).

RESULTS

DNA sequence of the spike protein

A 13 base oligonucleotide complementary to a sequence towards the 5' end of clone C5.136 (see Fig. 1) was used to prime cDNA synthesis from viral RNA. Clone pMB179 obtained from this experiment contained a 5.3 kilobase viral insert and in Southern blot analysis (Southern, 1975) hybridized to a small clone pMB172 which had previously been shown to contain mRNA E sequences by Northern blot analysis (data not presented). The DNA sequence analysis indicated that the 3' end of the clone was within 12 bases of the 5' end of the oligonucleotide used to prime DNA synthesis. 3645 bases of sequence containing the gene encoding the spike precursor protein are presented in Fig. 2. It is of note that 50 bases upstream from the AUG initiation codon is a sequence, AACTGAACAAAA, which resembles the homology regions that we have identified on the genome at the positions corresponding to the 5' ends of the bodies of mRNAs A, B and C (Brown & Bournnell, 1984; Bournnell *et al.*, 1984). This homology region maps approximately 8 kilobases from the 3' end of the viral genome, which is in good agreement with the size estimates for mRNA E of 7.9 kb (Stern & Kennedy, 1980*a*) and 7.8 kb (Bournnell *et al.*, 1984) as measured by gel electrophoresis. A similar sequence, AACTGAACAATA, is present at the predicted 5' end of the body of mRNA D and both sequences are underlined in Fig. 2. The sequence containing an open reading frame of 3486 bases and the primary amino acid sequence of the 127006 mol. wt. protein deduced from it are presented in Fig. 2 using the single-letter amino acid code (Commission on Biochemical Nomenclature, 1968). Spike precursor synthesis is initiated at the 5'-proximal AUG of mRNA E although the sequence GNNAUGU occurs rarely amongst functional eukaryotic initiator sequences (Kozak, 1983). The 3486 base pair open reading frame is followed by two UGA termination codons.

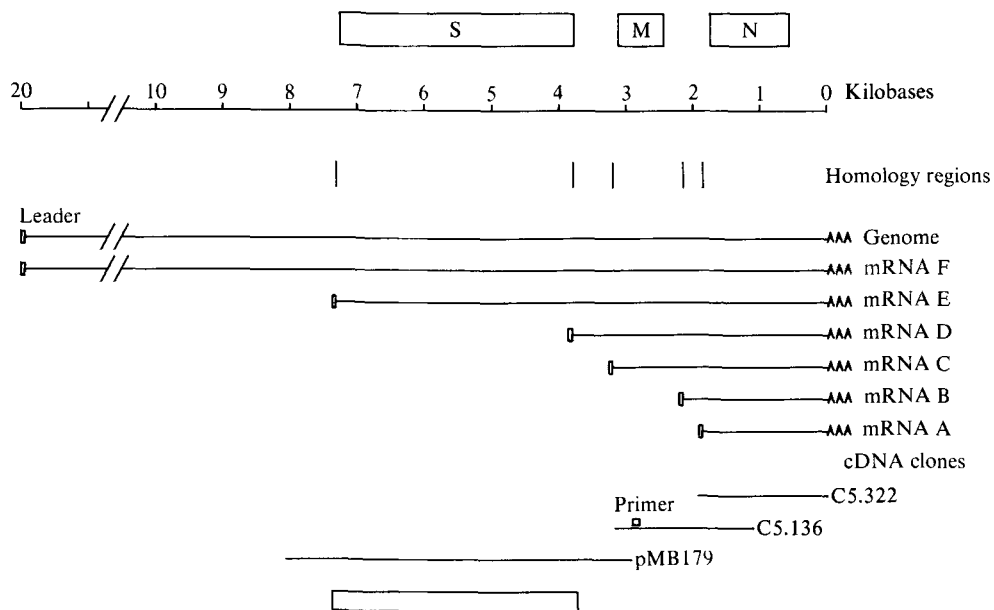


Fig. 1. Genomic organization of infectious bronchitis virus. The relationship between the 3' co-terminal nested set of mRNAs and the viral genome is indicated, as are the coding regions for the structural proteins, spike (S), membrane (M) and nucleocapsid (N) specified by mRNAs E, C and A respectively. The 'homology regions' are sequences present in the genome at positions corresponding to the 5' termini of the mRNA bodies and are thought to be involved in subgenomic RNA synthesis. The arrangement of the cDNA clones and the position of the primer used to generate clone pMB179 are shown. The sequence data presented in Fig. 2 are represented by the box below pMB179.

Partial amino acid sequence analysis of S1

To locate the position of the S1 polypeptide within the open reading frame, and to look for potential signal sequences, partial amino acid sequence analysis of the amino terminus of S1 was undertaken. The results indicated the presence of serine residues at positions 5, 6, 7, 14 and 20 in S1. These results unambiguously identified the N-terminal amino acids of S1 within the predicted sequence. The amino acid data indicated that an 18 amino acid signal sequence with a typical hydrophobic core and small neutral residues, alanine and cysteine, at positions -1 and -3 from the cleavage site (Von Heijne, 1984), is cleaved from S1. The positions of the N-terminal amino acids of S1, and of the proposed signal sequence are shown in Fig. 2.

Structural features of the IBV spike protein

In addition to the presence of a signal sequence at the amino terminus of S1, two other interesting structural features of the spike precursor protein were revealed by analysis of the predicted amino acid sequence. Firstly, the sequence contains 28 potential sites for *N*-glycosylation (assuming that Asn-Pro-Thr and Asn-Pro-Ser are not used; Neuberger *et al.*, 1972) which are shown in Fig. 2 and 3. Secondly, a hydrophilicity plot (Kyte & Doolittle, 1982) of the amino acid sequence (see Fig. 3) shows the presence of a hydrophobic region which contains 44 non-polar amino acids preceding charged amino acids at the carboxy-terminus of S2. This structure may anchor the spike protein to the viral envelope, as has been proposed for similar structures on human influenza virus and fowl plague virus haemagglutinins (Gething *et al.*, 1980; Porter *et al.*, 1979).

O E G F R S T S L A L Q Q I Q D V V S K Q S A I L T E T M A
 CGCAGGAGTTTATAGACTACTCTAGCATTTACACAAATTCAGAGTGTGTATGTAACAGAGTGCCTATCTTCTACTGACACTATGCC 2600
 2530 2540 2550 2560 2570 2580 2590 2600
 S L N K N F G A I S S V I Q E I Y Q Q F D A I O A N A Q V D
 ATCACTTATATAAAATTTGTGCTGATTTCTGTGATTTGAGAAATCTACACGCAATTTGACGCCAATACAGCAAAATGCTCTCAAGTCSGA 2650
 2620 2630 2640 2650 2660 2670 2680 2690 2700
 R L I T G R L S S L S V L A S A K Q A E Y I R V S Q O R E L
 TOSGTTATATACGCTAGATTTCTATCACTTCTTCTTTATGATCTCTATAGACAGCGGAGGATATATAGAGTGTCAACAGAGCGTGGT 2750
 2720 2730 2740 2750 2760 2770 2780 2790
 A T O K I N E C V K S Q S I R Y S F C G N G R H V L T I P Q
 AGCTCTCAGAAAATTAATAGTGTATAGTCCACGCTCTTATGAGTACTCTTTATGAGTATGSGAGCAATGCTCTTAACCAATACGCCCA 2800
 2810 2820 2830 2840 2850 2860 2870 2880
 N A P N G I V F I H F S Y T P D S F V N V T A I V G F C V K
 AAATGCACTTAATGCTATGATCTTCTTCTTACTCCAGATATGTTTCTTAAATGCTTACTCCATGAGTGGSTTTTGTGTA 2900
 2890 2900 2910 2920 2930 2940 2950 2960
 P A N A S Q Y A I V P A N G R G I F I Q V N G S Y Y I T A R
 GCCAGCTTAATGCTAGTCCATGATGAGTGGCCGCTAATGSGAGGATTTTATACAGTATGATGCTACTGACTACATCACTCCAGC 2990
 2980 2990 3000 3010 3020 3030 3040 3050
 D M Y M P R A I T A G D V V T L T S C Q A N Y V S V N K T V
 AGATATGATATCCAGAGACTTACTTCGAGGAGATGATGCTTACTCTCTTCTCAAGCAATTTATGATGCTGATTAATGAGCCST 3070
 3080 3090 3100 3110 3120 3130 3140 3150
 I T F V D N D D F N D E L S K W N D T K H E L P D F
 CATTTACTACATCTGATGACATGATGATTTGATTTTATGACCAATTTGTCAAAATGCTGGAAATGCTGATGACTACTGAGCTACAGAGCTT 3160
 3170 3180 3190 3200 3210 3220 3230 3240
 D K F N Y T V P I L D I D S E I D R I Q G V I Q G L N D S L
 TGCAAAATTCACAGTACCTTACTGCAATTTGATGCAATTTGATGCTTATCAAGCGTATATACAGGCTCTTATATGATGATCTCT 3250
 3260 3270 3280 3290 3300 3310 3320 3330
 I D L E K L S I L K T Y I K W P W Y V L A I A F A T I I F
 AATAGCCTTGAAGAAATTTTCATGACTTATATATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3340
 3350 3360 3370 3380 3390 3400 3410 3420
 I L I L G W V F F M T G C C G C C C G C F G I M P L M S K K C
 CXTCTTATATAGAGATGCT 3430
 3440 3450 3460 3470 3480 3490 3500 3510
 G K K S S Y Y T T F D N D V V T E Q Y R P K K S V *
 TCGTAAGAAATCTTATGACAGCTTATGATACAGTGTATGCAACATACACACCTTAAAGCTCTCTTTCATGATCCAAAGT 3520
 3530 3540 3550 3560 3570 3580 3590 3600
 CCGAGTCTCTCCATATGATTTATCTCTTCTGCTGTAACCTT 3610
 3620 3630 3640

Fig. 2. Nucleotide sequence of the IBV spike gene and the predicted amino acid sequence of the spike precursor protein. The homology regions at the ends of mRNA bodies D and E have been underlined and the 18 amino acids of the spike signal sequence are boxed. Potential glycosylation sites (N-X-S or N-X-T) on the spike precursor polypeptide are indicated by ● over the middle residue, and the potential anchor region of non-polar amino acids at the carboxy terminus by dotted underlining.

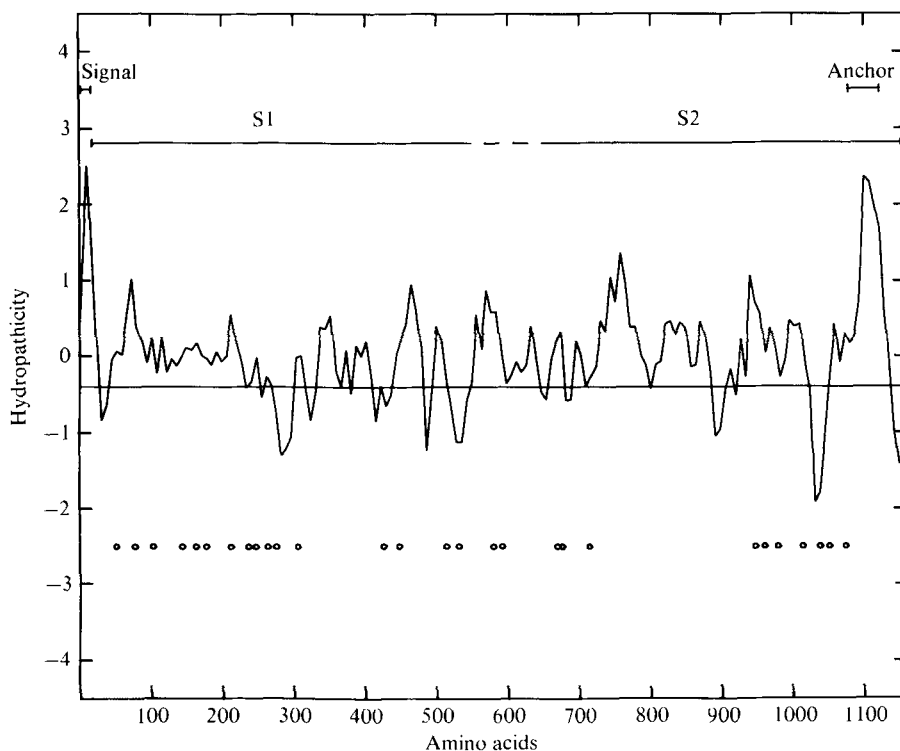


Fig. 3. Hydropathicity profile of the predicted amino acid sequence of the spike polypeptide. Positive values indicate hydrophobic regions and negative values indicate hydrophilic regions. The midpoint line represents a grand average of the hydropathy of the amino acid compositions of a large number of sequenced proteins (Kyte & Doolittle, 1982). Each point on the graph represents the average hydropathy of a span of 19 residues. The putative signal and anchor sequences are shown, as are the approximate regions of the gene encoding S1 and S2. The circles below the plot show the positions of potential glycosylation sites.

DISCUSSION

The DNA sequence presented in Fig. 2 contains the complete unique region present in IBV mRNA E. This messenger RNA has been found to specify production of the spike precursor in a translation system *in vitro* (Stern & Sefton, 1984). The sequence predicts a primary translation product of 1162 amino acids with a molecular weight of 127006, which is close to that estimated for the polypeptide components of S1 and S2. Translation of mRNA E *in vitro* had indicated that the non-glycosylated spike precursor had a molecular weight of 110000 (Stern & Sefton, 1984), and estimates of the combined molecular weight of S1 and S2 after the removal of oligosaccharides by endoglycosidase H were 115000 (Stern & Sefton, 1982) and 125000 (Cavanagh, 1983c). In addition, partial amino acid sequence analysis of the amino terminus of S1 has unambiguously identified the position of S1 within the predicted primary translation product of the spike gene.

The sequence presented has sequences AACTGAACAAAA towards the 5' end and AACTGAACAATA towards the 3' end (underlined in Fig. 2). Their high homology with sequences which have previously been found at the 5' ends of the bodies of IBV mRNAs A, B and C, referred to in Fig. 1 as 'homology regions' (Brown & Bournnell, 1984; Bournnell *et al.*, 1984) suggests that these sequences represent the position of the 5' ends of the bodies of mRNAs E and D. This is confirmed by mRNA length measurements. It is interesting to note then that the coding sequences for the spike gene are not completely contained within the 'unique' region of mRNA E but extend for approximately 32 bases beyond the predicted 5' terminus of the body

of mRNA D. A similar arrangement may be the case at the boundary of mRNAs A and B where an open reading frame predicting a 9500 mol. wt. polypeptide extends considerably into mRNA A (Bournnell & Brown, 1984). In both cases the homology regions appear to lie within coding regions and this may influence the exact sequence of these homology regions. The homology region at the 5' end of mRNA D differs from that present at the 5' end of mRNAs A, B and C in the presence of a G instead of a T (CTGAACAA rather than CTTAACAA) and it is interesting to note that the presence of a T would have generated an in-frame termination codon which would have eliminated nine amino acid residues, four of which are charged, from the carboxy terminus of the polypeptide.

Analysis of the predicted amino acid sequence reveals three interesting structural features of the spike protein. Firstly the results demonstrate the presence of a typical hydrophobic signal sequence which is not present on the mature protein. This is commonly found in proteins which must pass through membranes, and is of interest because the other surface protein of IBV, the membrane protein, which is believed to span the membrane, does not undergo substantial post-translational processing and contains no obvious signal sequence (Bournnell *et al.*, 1984). It has been proposed in this case that an internal signal sequence may be present in the membrane protein. Secondly, 28 potential sites for N-linked glycosylation are present which reflects the very high level of glycosylation which this protein is known to undergo. It is probable that the majority of these sites are glycosylated in order to account for the approximately 50000 difference in molecular weight observed between glycosylated and unglycosylated spike polypeptides. Mannose-rich viral glycoprotein carbohydrate side chains have molecular weights of approximately 2000 (Klenk & Rott, 1980). The third feature is a long stretch of non-polar amino acids close to the carboxy terminus of the S2 polypeptide which may serve as an anchor attaching the protein to the viral membrane. This agrees well with the observation (Cavanagh, 1983c) that treatment of virions with urea resulted in the removal of S1 but not S2. Similar 'anchor' structures have been proposed for a number of viral proteins.

The cloning and characterization of the spike gene of IBV has confirmed and extended previous observations on the surface glycoprotein of IBV. The availability of cloned spike sequences also represents an important step in attempts to develop a novel vaccine against IBV, as this viral component is thought to be involved in the induction of immunity against the disease.

We thank Bridgette Britton, Penny Gatter, Ann Foulds, Phil Davis and Ian Foulds for excellent technical assistance. This work was supported by Research Contract No. GBI-2-011-UK of the Biomolecular Engineering Programme of the Commission of the European Communities.

REFERENCES

- ANDERSON, S. (1981). Shotgun DNA sequencing using cloned DNase I-generated fragments. *Nucleic Acids Research* **9**, 3015-3027.
- BIGGIN, M. D., GIBSON, T. J. & HONG, G. F. (1983). Buffer gradient gels and ³⁵S label as an aid to rapid DNA sequence determination. *Proceedings of the National Academy of Sciences, U.S.A.* **80**, 3963-3965.
- BOURSNELL, M. E. G. & BROWN, T. D. K. (1984). Sequencing of coronavirus IBV genomic RNA: a 195-base open reading frame encoded by mRNA B. *Gene* **29**, 87-92.
- BOURSNELL, M. E. G., BROWN, T. D. K. & BINNS, M. M. (1984). Sequence of the membrane protein gene from avian coronavirus IBV. *Virus Research* **1**, 303-313.
- BRETT, M. & FINDLAY, J. B. C. (1983). Isolation and characterisation of the CNBr peptides from the proteolytically derived N-terminal fragment of ovine opsin. *Biochemical Journal* **211**, 661-670.
- BROWN, T. D. K. & BOURSNELL, M. E. G. (1984). Avian infectious bronchitis virus genomic RNA contains sequence homologies at the intergenic boundaries. *Virus Research* **1**, 15-24.
- BROWN, T. D. K., BOURSNELL, M. E. G. & BINNS, M. M. (1984). A leader sequence is present on mRNA A of avian infectious bronchitis virus. *Journal of General Virology* **65**, 1437-1442.
- CAVANAGH, D. (1981). Structural polypeptides of coronavirus IBV. *Journal of General Virology* **53**, 93-103.
- CAVANAGH, D. (1983a). Coronavirus IBV glycopolypeptides: size of their polypeptide moieties and nature of their oligosaccharides. *Journal of General Virology* **64**, 1187-1191.
- CAVANAGH, D. (1983b). Coronavirus IBV: further evidence that the surface projections are associated with two glycopolypeptides. *Journal of General Virology* **64**, 1787-1791.
- CAVANAGH, D. (1983c). Coronavirus IBV: structural characterization of the spike protein. *Journal of General Virology* **64**, 2577-2583.

- COMMISSION ON BIOCHEMICAL NOMENCLATURE (1968). A one-letter notation for amino acid sequences. *European Journal of Biochemistry* **5**, 151–153.
- DEININGER, P. L. (1983). Random subcloning of sonicated DNA: application to shotgun DNA sequence analysis. *Analytical Biochemistry* **129**, 216–223.
- GAIT, M. J., MATTHES, H. W. D., SINGH, M., SPROAT, B. & TITMAS, R. C. (1982). Rapid synthesis of oligodeoxyribonucleotides. VII. Solid phase synthesis of oligodeoxyribonucleotides by a continuous flow phosphotriester method on a kieselguhr-polyamide support. *Nucleic Acids Research* **10**, 6243–6254.
- GETHING, M. J., BYE, J., SKEHEL, J. & WATERFIELD, M. (1980). Cloning and DNA sequence of double-stranded copies of haemagglutinin genes from H2 and H3 strains elucidates antigenic shift and drift in human influenza virus. *Nature, London* **287**, 301–306.
- GRUNSTEIN, M. & HOGNESS, D. S. (1975). Colony hybridisation: a method for the isolation of cloned DNAs that contain a specific gene. *Proceedings of the National Academy of Sciences, U.S.A.* **72**, 3961–3965.
- GUBLER, U. & HOFFMAN, B. J. (1983). A simple and very efficient method for generating cDNA libraries. *Gene* **25**, 263–269.
- HANAHAN, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology* **166**, 557–580.
- HONG, G. F. (1981). A method for sequencing single-stranded cloned DNA in both directions. *Bioscience Reports* **1**, 243–252.
- KLENK, H. D. & ROTT, R. (1980). Cotranslational and posttranslational processing of viral glycoproteins. *Current Topics in Microbiology and Immunology* **90**, 19–48.
- KOZAK, M. (1983). Comparison of initiation of protein synthesis in procaryotes, eucaryotes and organelles. *Microbiological Reviews* **47**, 1–45.
- KYTE, J. & DOOLITTLE, R. F. (1982). A simple method for displaying the hydrophobic character of a protein. *Journal of Molecular Biology* **157**, 105–132.
- MOCKETT, A. P. A., CAVANAGH, D. & BROWN, T. D. K. (1984). Monoclonal antibodies to the S1 spike and membrane proteins of avian infectious bronchitis coronavirus strain Massachusetts M41. *Journal of General Virology* **65**, 2281–2286.
- NEUBERGER, A., GOTTSCHALK, A., MARSHALL, R. D. & SPIRO, R. G. (1972). Carbohydrate-peptide linkages in glycoproteins and methods for their elucidation. In *Glycoproteins: Their Composition, Structure and Function*, pp. 450–490. Edited by A. Gottschalk. Amsterdam: Elsevier.
- PORTER, A. G., BARBER, C., CAREY, N. H., HALLEWELL, R. A., THRELFALL, G. & EMTAGE, J. S. (1979). Complete nucleotide sequence of an influenza virus haemagglutinin gene from cloned DNA. *Nature, London* **282**, 471–477.
- SANGER, F., NICKLEN, S. & COULSON, A. R. (1977). DNA sequencing with chain terminating inhibitors. *Proceedings of the National Academy of Sciences, U.S.A.* **74**, 5463–5467.
- SOUTHERN, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* **98**, 503–517.
- STADEN, R. (1982). Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing. *Nucleic Acids Research* **10**, 4731–4751.
- STADEN, R. (1984). Graphic methods to determine the function of nucleic acid sequences. *Nucleic Acids Research* **12**, 521–538.
- STERN, D. F. & KENNEDY, S. I. T. (1980a). Coronavirus multiplication strategy. I. Identification and characterisation of virus-specified RNA. *Journal of Virology* **34**, 665–674.
- STERN, D. F. & KENNEDY, S. I. T. (1980b). Coronavirus multiplication strategy. II. Mapping the avian infectious bronchitis virus intracellular RNA species to the genome. *Journal of Virology* **36**, 440–449.
- STERN, D. F. & SEFTON, B. M. (1982). Coronavirus proteins: structure and function of the oligosaccharides of the avian infectious bronchitis virus glycoproteins. *Journal of Virology* **44**, 804–812.
- STERN, D. F. & SEFTON, B. M. (1984). Coronavirus multiplication: the locations of the genes coding for the virion proteins on the avian infectious bronchitis virus genome. *Journal of Virology* **50**, 22–29.
- STERN, D. F., BURGESS, L. & SEFTON, B. M. (1982). Structural analysis of virion proteins of the avian coronavirus infectious bronchitis virus. *Journal of Virology* **42**, 208–219.
- VON HEIJNE, G. (1984). How signal sequences maintain cleavage specificity. *Journal of Molecular Biology* **173**, 243–251.
- WACHTER, E., MACHLEIDT, W., HOFNER, H. & OTTO, J. (1973). Amino profile glass and its *p*-phenylenediisothiocyanate derivative: a new support in solid-phase Edman degradation of peptides and proteins. *FEBS Letters* **35**, 97–102.
- WELCH, W. J., SEFTON, B. M. & ESCH, F. S. (1981). Amino-terminal sequence analysis of alphavirus polypeptides. *Journal of Virology* **38**, 968–972.

(Received 25 October 1984)