

Key words: *IBV/spike precursor/sequence comparison*

Comparison of the Spike Precursor Sequences of Coronavirus IBV Strains M41 and 6/82 with that of IBV Beaudette

By MATTHEW M. BINNS,* MICHAEL E. G. BOURSNELL, FIONA M. TOMLEY AND T. DAVID K. BROWN

Houghton Poultry Research Station, Houghton, Huntingdon, Cambridgeshire PE17 2DA, U.K.

(Accepted 29 August 1986)

SUMMARY

The nucleotide sequences of the spike precursor genes of infectious bronchitis virus strains M41 and 6/82 have been determined and compared with that of the Beaudette strain which we have previously sequenced. The two Massachusetts strains, M41 and Beaudette, were found to be remarkably similar, having only 3·7% of the amino acids different. The situation with 6/82, one of the new field isolates, is quite different and this strain had 13·8% of its amino acids different from Beaudette. The differences identified are discussed in terms of the structural features of the spike protein.

The comparison of the amino acid sequences of viral proteins from different viral strains and serotypes has yielded much useful information. This approach has been used to help locate conserved domains in proteins which might be essential for their structure and function, to identify epitopes involved in virus neutralization (Minor *et al.*, 1983), and to study antigenic drift (Gething *et al.*, 1980).

Infectious bronchitis virus (IBV), a coronavirus, causes an economically important disease of poultry. There are a large number of serotypes of the virus. Recently, a number of strains have been isolated which are not related serologically in neutralization tests to strains of IBV previously isolated in England (Cook, 1984). The purpose of the present study was to see how closely representative strains of IBV are related at the level of the spike protein, the virus structure against which neutralizing antibodies are directed (Mockett *et al.*, 1984).

We have recently reported the nucleotide sequence of the gene encoding the spike precursor protein of IBV strain Beaudette (Binns *et al.*, 1985*a*). The Beaudette strain, which is no longer pathogenic for chickens, is a Massachusetts-derived strain which has been passaged over 250 times in eggs. In this paper we present the spike sequences of the IBV strains M41 and 6/82, and compare them with that of the Beaudette strain. Massachusetts strain M41 has been passaged extensively in chickens and remains fully pathogenic. IBV 6/82 is one of the new serotypes of IBV recently isolated from the field (Cook, 1983). Both M41 and 6/82 have relevance to vaccine production; M41 is currently used as an inactivated vaccine and 6/82 is being assessed for similar use.

cDNA cloning was carried out by two methods on RNA isolated from gradient-purified virus. For IBV M41, a specific oligonucleotide primer 15 bases long, complementary to a viral sequence towards the 5' end of mRNA D was used to prime reverse transcription from genomic RNA as previously described (Binns *et al.*, 1985*a*). Two overlapping clones, pMB250 and pMB276, were isolated which together spanned the spike gene. With IBV 6/82, random priming with calf thymus oligonucleotides (Binns *et al.*, 1985*b*) was used to generate a library of 6/82 clones. Clones containing the spike gene were then identified using radioactively labelled M13mp10 clones containing IBV Beaudette spike sequences as probes. Three overlapping clones, pMB252, pMB253 and pMB277, containing the spike gene were identified and used for further studies.

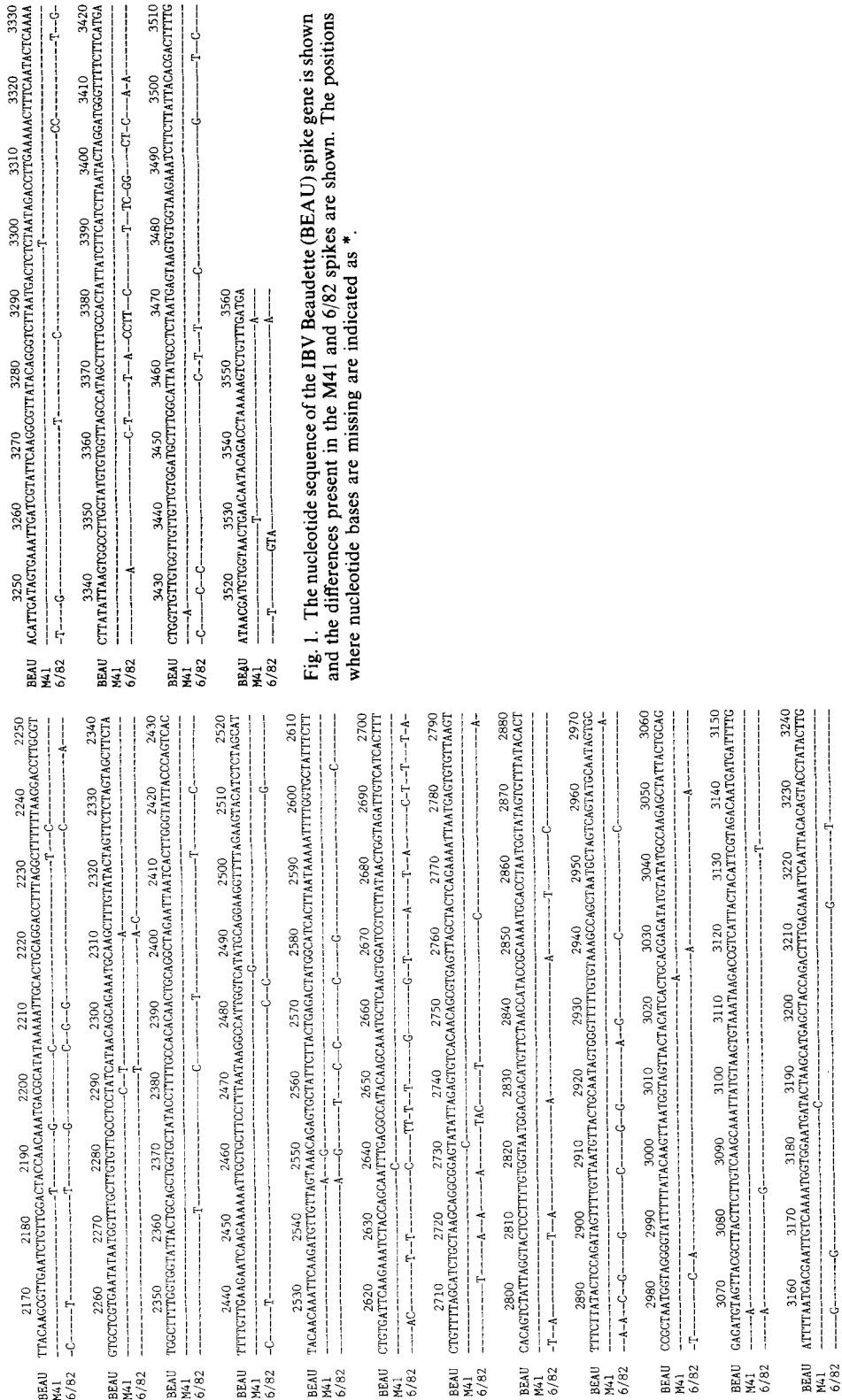


Fig. 1. The nucleotide sequence of the IBV Beaudette (BEAU) spike gene is shown and the differences present in the M41 and 6/82 spikes are shown. The positions where nucleotide bases are missing are indicated as *.

For M13 DNA sequencing, random fragments of pMB250, pMB276, pMB252, pMB253 and pMB277 were generated by sonication (Deininger, 1983), end-repaired, and cloned into *Sma*I-cleaved, phosphatase-treated M13mp10 (Amersham). Clones containing viral inserts were identified by colony hybridization with kinase-labelled, alkali-treated viral probe. In addition *Pst*I and *Rsa*I fragments of pMB250, pMB276, pMB252 and pMB253 were cloned into *Pst*I-digested M13mp11 and *Sma*I-cleaved M13mp10 respectively. DNA sequencing and data analysis were carried out as previously described (Boursnell *et al.*, 1985).

The DNA sequences of the spike genes of M41 and 6/82 assembled from the appropriate clones are presented in Fig. 1 where they are compared with that of the Beaudette strain. There are 71 nucleotide differences between IBV Beaudette and M41, and 496 differences between Beaudette and 6/82 over a region of 3554 nucleotides. In addition IBV 6/82 has six extra nucleotides at positions 420 to 425 and is missing three nucleotides at positions 249 to 251 (see Fig. 1). While sequencing pMB250 we identified a sequence of approximately 1250 nucleotides within the spike sequence which did not correspond with the Beaudette spike sequence. Comparison of this sequence with the EMBL database revealed that this was an insertion sequence, IS5. It had presumably transposed from the chromosome of *Escherichia coli* LE392 onto the recombinant plasmid shortly after transformation. Insertion resulted in the duplication of four bases at the insertion site but we have confirmed the sequence in this region from the overlapping clone pMB276.

The amino acid sequences of the spikes from the three strains are compared in Fig. 2. There are 44 amino acid differences between Beaudette and M41, and in addition the last nine amino acids at the COOH terminus of S2 are missing due to the presence of a new stop codon in the homology region of M41. This new stop codon results in the coding region of the M41 spike being contained wholly within mRNA E, with no overlap into mRNA D as is seen with Beaudette. This region encodes mainly charged amino acids and is at the end of the potential hydrophobic anchor present at the COOH terminus of the S2 component of the spike. The M41 result is unusual however, as 6/82 possesses the same stop codons as Beaudette. Furthermore, Niesters *et al.* (1986), in presenting the sequence of mRNA D of M41 showed that the strain of M41 they are using lacks the stop codon in the homology region but has the same stop codons as Beaudette. In order to confirm the presence of the stop codon within the homology region in our IBV M41 preparation a second independent cDNA clone, pMB233, spanning this region was sequenced. This clone also contained the stop codon in the homology region suggesting that the result from pMB250 is not due to a cloning artefact. It is also of note that the stop codon regenerates the homology sequence present at the boundaries of mRNAs A, B, C and the genome of IBV Beaudette (Brown *et al.*, 1986).

There are 161 amino acid differences between Beaudette and 6/82. This is a much lower proportion of amino acid differences to nucleotide differences (161/496) than is seen in M41 (44/71). In 6/82 there are more third base changes not altering the amino acid, and also several instances where two or three bases are altered in one triplet codon. IBV 6/82 also has two extra amino acids at positions 120 and 121 and is missing one amino acid at position 63 (see Fig. 2).

The distribution of the amino acid differences reveals a number of interesting features of the spike proteins. In IBV 6/82 five out of 18 amino acids are altered in the signal sequence of the protein, supporting the idea that it is the overall hydrophobicity of the signal that is important, not the exact sequence. In Fig. 2 it can be clearly seen that most of the amino acid differences are in the S1 polypeptide, to which a neutralizing monoclonal antibody binds (Mockett *et al.*, 1984), and that there are three regions which show considerable variation. These are from residues 37 to 81, 117 to 160 and 269 to 298. The region 37 to 81 contains two hydrophilic areas, region 117 to 160 is hydrophobic and 269 to 298 contains a strongly hydrophilic region. Previous studies have shown that major antigenic sites on proteins tend to lie in hydrophilic regions and that they are frequently associated with the termini of proteins due to their greater flexibility (van Regenmortel, 1986). There is a single change in 6/82, from a phenylalanine to a serine, in the connecting peptide region between S1 and S2. It is interesting to compare this with influenza virus haemagglutinin where the number of basic residues present at the cleavage site varies

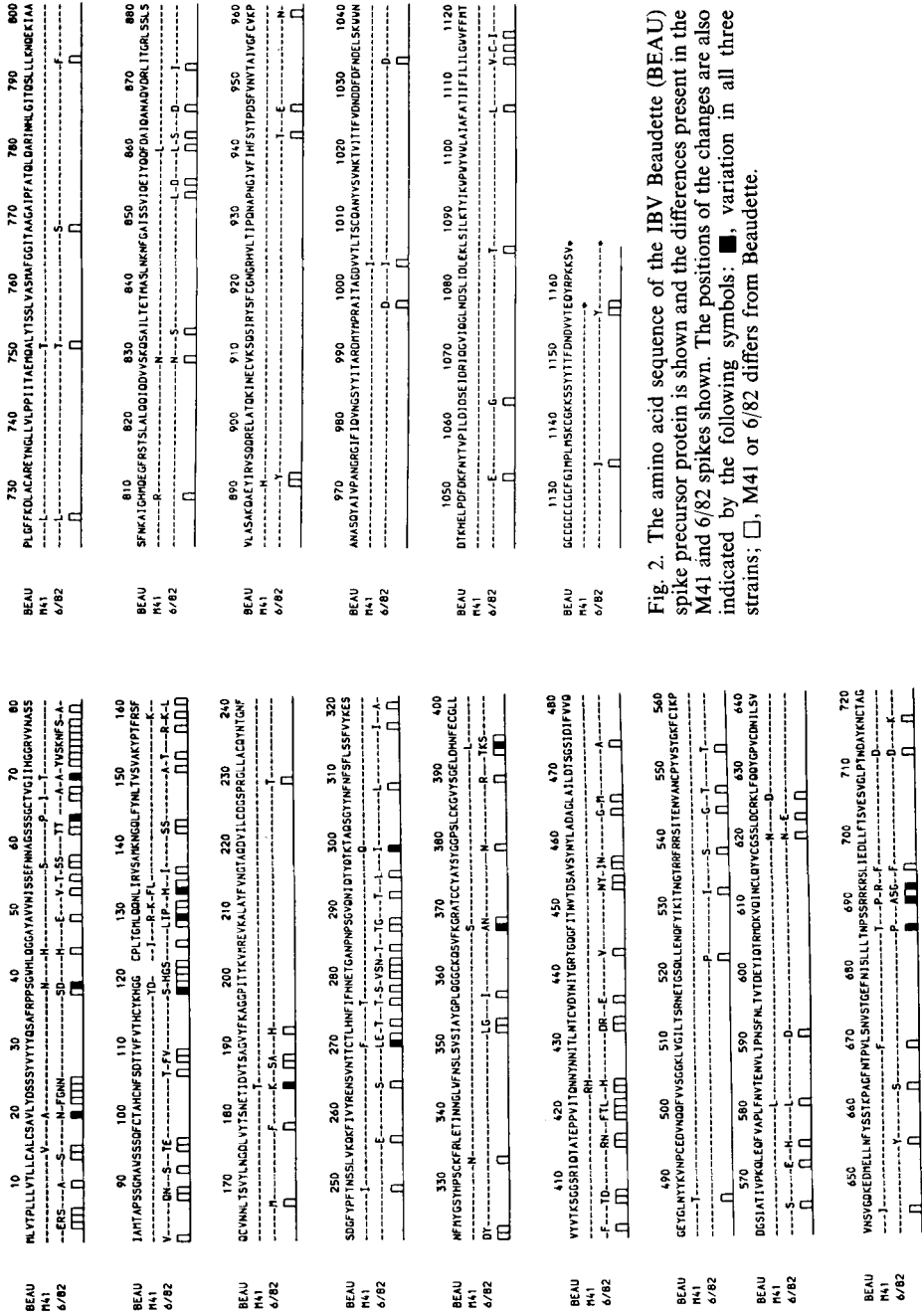


Fig. 2. The amino acid sequence of the IBV Beaudette (BEAU) spike precursor protein is shown and the differences present in the M41 and 6/82 spikes shown. The positions of the changes are also indicated by the following symbols: ■, variation in all three strains; □, M41 or 6/82 differs from Beaudette.

between strains and is thought to affect the host range of the virus (Gething *et al.*, 1980; Porter *et al.*, 1979).

Two structural features of the spike protein have been highly conserved, the potential glycosylation sites and the positions of the cysteine residues in the protein. IBV Beaudette contains 28 potential glycosylation sites and these are all present in M41, in which there is also an additional site present in S1. In 6/82 one potential glycosylation site is missing from S1 and one from S2, but there are five new sites in S1 and one new site in S2. How many of these potential glycosylation sites are actually glycosylated is unknown but from the molecular weights of glycosylated and non-glycosylated spikes a large proportion of them must be. The 38 cysteine residues present in Beaudette are all conserved in M41 and 6/82 (and one extra cysteine residue is present in 6/82 at the COOH terminus of S2) maintaining the overall secondary structure of the spikes.

In conclusion, the present study shows that two Massachusetts strains of IBV, M41 and Beaudette, despite having widely different passage histories, still appear remarkably similar with only 44 amino acid differences out of 1162 (3.7%). Of these differences 26 are in S1 and must include the change which prevents a neutralizing monoclonal antibody raised against M41 from neutralizing Beaudette. It is, however, not possible to determine from these studies whether it is a difference in the spike protein which is responsible for the very different pathogenicities which M41 and Beaudette exhibit. 6/82, a representative of one of the new field isolates, shows greater variation with 13.8% of its amino acids different from Beaudette. RNA viruses can mutate rapidly, partly due to the absence of a proof-reading mechanism during replication such as operates in DNA replication (Holland *et al.*, 1982). In this way RNA viruses can evolve rapidly to avoid host defence mechanisms which in turn may explain the existence of the large number of serotypes seen. This also suggests that new serotypes will continue to emerge under the pressure of vaccination.

We thank Bridgette Britton for excellent technical assistance.

REFERENCES

- BINNS, M. M., BOURSNEILL, M. E. G., CAVANAGH, D., PAPPIN, D. J. C. & BROWN, T. D. K. (1985*a*). Cloning and sequencing of the gene encoding the spike protein of the coronavirus IBV. *Journal of General Virology* **66**, 719–726.
- BINNS, M. M., BOURSNEILL, M. E. G., FOULDS, I. J. & BROWN, T. D. K. (1985*b*). The use of a random priming procedure to generate cDNA libraries of infectious bronchitis virus, a large RNA virus. *Journal of Virological Methods* **11**, 265–269.
- BOURSNEILL, M. E. G., BINNS, M. M., FOULDS, I. J. & BROWN, T. D. K. (1985). Sequences of the nucleocapsid genes from two strains of avian infectious bronchitis virus. *Journal of General Virology* **66**, 573–580.
- BROWN, T. D. K., BOURSNEILL, M. E. G., BINNS, M. M. & TOMLEY, F. M. (1986). Cloning and sequencing of 5' terminal sequences from avian infectious bronchitis virus genomic RNA. *Journal of General Virology* **67**, 221–228.
- COOK, J. K. A. (1983). Isolation of a new serotype of infectious bronchitis-like virus from chickens in England. *Veterinary Record* **112**, 104–105.
- COOK, J. K. A. (1984). The classification of new serotypes of infectious bronchitis virus isolated from poultry flocks in Britain between 1981 and 1983. *Avian Pathology* **13**, 733–741.
- DEININGER, P. L. (1983). Random subcloning of sonicated DNA: application to shotgun DNA sequence analysis. *Analytical Biochemistry* **129**, 216–223.
- 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.
- HOLLAND, J., SPINDLER, K., HORODYSKI, F., GRABAU, E., NICHOL, S. & VANDEPOL, S. (1982). Rapid evolution of RNA genomes. *Science* **215**, 1577–1585.
- MINOR, P. D., SCHILD, G. C., BOOTMAN, J., EVANS, D. M. A., FERGUSON, M., REEVE, P., SPITZ, M., STANWAY, G., CANN, A. J., HAUPTMANN, R., CLARKE, L. D., MOUNTFORD, R. C. & ALMOND, J. W. (1983). Location and primary structure of a major antigenic site for poliovirus neutralization. *Nature, London* **301**, 674–679.
- 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.
- NIESTERS, H. G. M., ZIJDERVELD, A. J., SEIFERT, W. F., LENSTRA, J. A., BLEUMINK-PLUYM, N. M. C., HORZINEK, M. C. & VAN DER ZEIST, B. A. M. (1986). Infectious bronchitis virus RNA D encodes three potential translation products. *Nucleic Acids Research* **14**, 3144.

- 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.
- VAN REGENMORTEL, M. H. V. (1986). Which structural features determine protein antigenicity? *Trends in Biochemical Sciences* **11**, 36–39.

(Received 21 July 1986)