Nucleotide and amino acid sequence of the S1 subunit of the spike glycoprotein of avian infectious bronchitis virus, strain D3896

G.Koch and A.Kant
Central Veterinary Institute, Department of Virology, PO Box 365, 8200 AJ Lelystad, The Netherlands

Submitted April 5, 1990 EMBL accession no. X52084

The nucleotide sequence of the S1 glycopeptide of the spike protein of variant strain D3896 of avian infectious bronchitis virus (IBV) was determined. Reverse transcriptase was used in the dideoxy chain termination method to synthesize cDNA from genomic RNA. Ten synthetic oligonucleotides with a length of 24–25 nucleotides were used to prime the synthesis of cDNA. The spike protein (S) is proteolytically processed into two parts, the N-terminal S1 and the C-terminal S2 glycopeptide. Neutralising antibodies are directed against the S1 glycopeptide. The start codon of the S1 subunit is underlined in the sequence listing. Although the underlined X at position 1323 was unclear, the amino acid sequence of strain D3896 was unambiguous.

We aligned the S1 nucleotide sequence of strain D3896 with the sequences of the closely related strains D207 and D274 (1, 2), which belong to a distinct IBV serotype (3). D207 had 31–35 nucleotide mutations (depending on uncertainties in the D207 and D3896 sequences) and D274 had 38–39 nucleotide mutations (depending on an uncertainty in the D3896 sequence) (2). The nucleotide mutations caused 18–19 amino acid differences between the S1 protein sequences of strains D3896 and D207 and 20 differences between the sequences of strains D3896 and D207.

```
       ATA GTA TAT TTG ACG TTG CTA AGT TTG ATT TGA AAT TAA AAG CAA CGC CAG 51
TTG TAA ATT TGA AAA CTG AAC ACA AGA CCG ACT TAG TAG TTA ATT TAC TAA GGA ACG 108
                        M  L  E  K  S  L  L  L  V  T  L  L  F
GTA AAT TAT TGG TTA GAG ATG TTG GAG AAG TCA CTG TTA TTA GTG ACT CTT TTG TTT 165
 A  L  C  S  A  N  L  F  G  N  N  S  Y  V  Y  Y  Y  Q  S
GCA CTA TGT AGT GCT AAT CTG TTT GGT AAT AAT TCT TAC GTG TAC TAC TAC CAA AGT 222
 A  F  R  P  P  N  G  W  H  L  H  G  G  A  Y  E  V  V  N
GCC TTC AGA CCA CCT AAT GGT TGG CAT TTA CAT GGT GGT GCT TAT GAA GTA GTA AAT 279
 V  S  T  E  S  S  N  A  G  T  T  E  C  T  A  G  A  I  Y
GTT TCT ACG GAA TCT AGT AAT GCA GGC ACG ACT GAG TGT ACT GCC GGT GCT ATT TAT 336
 W  S  K  N  F  S  A  A  S  V  A  M  T  A  P  Q  N  G  M
TGG AGT AAG AAT TTC AGT GCT GCT TCC GTG GCT ATG ACA GCA CCT CAA AAT GGT ATG 393
 L  W  S  T  A  Q  F  C  T  A  H  C  N  F  T  D  F  V  V
TTA TGG TCT ACT GCG CAA TTT TGT ACG GCT CAC TGC AAT TTT ACT GAT TTT GTA GTA 450
 F  V  T  H  C  Y  K  S  A  S  G  S  C  P  L  T  G  L  I
TTT GTT ACA CAT TGC TAT AAA AGT GCT TCT GGT TCA TGT CCT TTA ACA GGT CTG ATT 507
 P  Q  Y  H  I  R  I  S  A  M  K  N  S  S  L  F  Y  N  L
CCA CAG TAT CAT ATT CGT ATT TCT GCT ATG AAA AAT AGC AGT TTG TTT TAT AAC TTA 564
 T  V  A  V  T  K  Y  P  R  F  K  S  L  Q  C  V  N  N  M
ACA GTT GCT GTG ACT AAA TAT CCT AGA TTT AAG TCG CTT CAG TGT GTT AAT AAT ATG 621
 T  S  V  Y  L  N  G  D  L  V  F  T  S  N  E  T  K  D  V
ACA TCT GTA TAC CTA AAT GGC GAT CTC GTT TTT ACT TCT AAC GAG ACT AAA GAT GTT 678
 S  A  A  G  V  H  F  K  A  G  G  P  I  T  Y  K  V  M  R
AGT GCT GCA GGT GTT CAT TTT AAA GCT GGT GGC CCT ATA ACT TAT AAA GTT ATG CGG 735
 E  V  K  A  L  A  Y  F  V  N  G  T  A  Q  D  V  I  L  C
GAA GTT AAA GCA TTG GCT TAT TTT GTT AAT GGC ACC GCA CAA GAT GTG ATT CTT TGT 792
 D  G  S  P  T  G  L  L  A  C  Q  Y  N  T  G  N  F  S  D
```

REFERENCES

1. Kusters,J.G., Niesters,H.G.M., Lenstra,J.A., Horzinek,M.C. and van der Zeijst,B.A.M. (1989) *Virology* **169**, 217–221.
2. Jordi,B.J.A.M., Kremers,A.W.M., Kusters,H.G. and van der Zeijst,B.A.M. (1989) *Nucl. Acids Res.* **17**, 6726.
3. Davelaar,F.G., Kouwenhoven,B. and Burger,A.G. (1984) *Vet. Q.* **6**, 114–120.

```
GAC GGT TCA CCT ACA GGT TTA CTA GCA TGT CAG TAT AAT ACT GGT AAT TTT TCA GAT 849
 G   F   Y   P   F   T   N   S   S   L   V   K   E   K   F   I   V   Y   R
GGC TTT TAT CCT TTT ACT AAT AGT AGT TTA GTT AAG GAA AAG TTT ATT GTT TAT CGT 906
 E   S   S   V   N   T   T   L   E   L   T   N   F   T   F   S   N   V   S
GAA AGT AGT GTT AAC ACT ACT TTG GAG TTA ACT AAT TTC ACT TTT TCT AAT GTA AGT 963
 N   A   N   P   N   T   G   G   V   H   T   I   Q   L   Y   Q   T   S   T
AAT GCT AAC CCT AAC ACA GGG GGT GTC CAT ACC ATA CAA TTA TAT CAA ACC AGC ACA 1020
 A   Q   S   G   H   Y   N   F   N   F   S   F   L   S   S   F   T   Y   K
GCT CAG AGT GGT CAT TAT AAT TTT AAT TTC TCC TTT CTG AGT AGT TTT ACC TAT AAG 1077
 E   S   D   Y   M   Y   G   S   Y   H   P   S   C   K   F   R   L   E   T
GAG TCT GAT TAT ATG TAT GGG TCT TAC CAC CCA AGT TGT AAG TTT AGA CTA GAA ACT 1134
 I   N   N   G   L   W   F   N   S   L   S   V   S   L   G   Y   G   P   I
ATT AAT AAT GGT TTG TGG TTT AAC TCA CTT TCC GTC TCT CTT GGT TAC GGA CCT ATT 1191
 Q   G   G   C   K   Q   S   V   F   Q   N   R   A   T   C   C   Y   A   Y
CAA GGT GGT TGT AAG CAA TCT GTG TTC CAA AAT AGG GCA ACT TGT TGT TAT GCC TAC 1248
 S   Y   N   G   P   P   L   C   K   G   V   Y   R   G   E   L   T   K   S
TCG TAT AAT GGA CCT CCC CTT TGT AAA GGT GTT TAT AGA GGT GAG TTA ACA AAA AGC 1305
 F   E   C   G   L   L   V   F   V   T   K   T   D   G   S   R   I   Q   T
TTT GAA TGT GGA TTG CTX GTT TTT GTG ACT AAG ACT GAT GGT TCC CGT ATA CAA ACC 1363
 R   N   E   P   F   T   L   T   Q   H   N   Y   N   N   I   T   L   D   R
AGA AAT GAA CCA TTT ACG TTA ACC CAG CAC AAT TAT AAT AAT ATT ACT TTA GAT AGA 1419
 C   V   E   Y   N   I   Y   G   R   V   G   Q   G   F   I   T   N   V   T
TGT GTT GAG TAT AAT ATA TAT GGT AGA GTC GGA CAA GGT TTT ATT ACT AAT GTA ACT 1476
 N   Y   A   I   N   Y   N   Y   L   A   D   G   G   M   A   I   L   D   T
AAC TAT GCC ATT AAT TAT AAT TAT TTA GCT GAT GGT GGT ATG GCT ATT TTA GAT ACA 1533
 S   G   A   I   D   I   F   V   V   Q   G   E   Y   G   L   N   Y   Y   K
TCT GGC GCC ATA GAC ATC TTC GTT GTA CAA GGT GAA TAT GGT CTT AAT TAT TAC AAG 1590
 V   N   F   C   E   D   V   N   Q   Q   F   V   V   S   G   G   K   L   V
GTC AAC CCT TGT GAG GAT GTT AAT CAG CAG TTT GTA GTT TCT GGT GGT AAA TTA GTA 1647
 G   I   L   T   S   R   N   E   T   G   S   Q   P   L   E   N   Q   F   Y
GGT ATT CTT ACG TCA CGT AAT GAG ACT GGC TCG CAG CCT CTT GAA AAC CAG TTC TAT 1704
 I   K   I   I   N   G   T   R   R   S   R   R   S   I   T   G   N   V   T
ATC AAA ATC ATT AAT GGA ACT CGT CGT TCT AGA CGT TCT ATT ACT GGG AAT GTT ACA 1761
 N   C   P   Y   V
AAT TGC CCT TAT GTT 1776
```

The X at position 1323 is T, A, C, or G.