

Complete sequence of a cDNA clone of the hemagglutinin gene of influenza A/Chicken/Scotland/59 (H5N1) virus: comparison with contemporary North American and European strains

Barun K.De, George G.Brownlee¹, Alan P.Kendal and Michael W.Shaw*

Division of Viral Diseases, Center for Infectious Diseases, Centers for Disease Control, Public Health Services, US Department of Health and Human Services, Atlanta, GA 30333, USA and ¹Sir William Dunn School of Pathology, University of Oxford, Oxford, OX1 3RE, UK
Submitted November 19, 1987

We report here the complete nucleotide and deduced amino acid sequence of the hemagglutinin gene of the avian influenza virus, A/Chicken/Scotland/59 (H5N1) cloned into plasmid pATX (1). The deduced amino acid sequence has been compared to the sequences of related HA genes from avian influenza viruses causing recent outbreaks in the commercial poultry industry (2,3). Sequencing was done by the Sanger dideoxy method (4). When compared to another European strain, A/Turkey/Ireland/83, there is a homology of 88% at the nucleotide level and 91% at the amino acid level compared to 81% nucleotide and 87% amino acid homology with a North American strain, A/Chicken/Pennsylvania/83. Despite the sequence variation, the H5HA gene of A/Chicken/Scotland/59 expressed in vaccinia virus protects 95% of the birds challenged with the otherwise lethal influenza A/Chicken/Pennsylvania/83 virus (De *et al.*, *Vaccine*. In press.).

A:

```

AGCAAAAGCAGGGGTCTGATCTACTAAA   ATG GAG AGA ATA GTG CTT CTT CTT GCA ATA GTC AGT CTT GTC AAA AGT GAC   79
CAG ATT TGC ATT GGT TAC CAT GCA AAC AAA TGG ACA AAA CAG GTT GAC ACA ATA ATG GAG AAG AAT GTG ACT GTC 154
ACA CAC GCC CAA GAT ATA CTT GAA AGG ACA CAC AAC GGG ANG CTC TGC AGT CTA AAT GGG GTG AAA CCT CTC ATC 229
TTG AGG GAC TGT AGT GTA GCT GGA TGG CTC CTG GGA AAC CCC ATG TGT GAC GAG TTC CTA AAT CTG CCA GAA TGG 304
TTA TAC ATA GTA GAA AAG GAC AAT CCA ATC AAC AGC CTC TGC TAC CCA GGG GAT TTT AAT GAC TAC GAA GAA CTG 379
AAA TAC CTA TTG AGC AGT ACA AAC CAT TTT GAG AAA ATT CGA ATC ATC CCC AGG AGC TCT TGG TCC AAT CAT GAT 454
GCT TCA TCA GGA GTG AGC TCT GCA TGT CCA TAC ATT GGG AGG TCC TCC TTT TTA AGA AAT GTG GTA TGG CTG ATC 529
AAA AAG AAC AAT ACA TAC CCA ACA ATA AAG AGG AGT TAC AAC AAT ACC AAC CAA GAG GAT CTT CTA ATA CTG TGG 604
GGG ATC CAT CAT CCT AAT GAC GCA GGG GAA CAA ACA AAG CTC TAT CAA AAT CCG ACC ACT TAC GTT TCT GTC GGA 679
ACA TCA ACA CTG AAC CAG AGA TCA ATT CCA GAA ATA GCC ACT AGG CCT AAA GTA AAT GGG CAG AGC GGA CGA ATG 754
GAA TTC TTC TGG ACA ATT TTA AAG CCA ATC GCA ATC AAT TTT GAG AGT AAT GGG AAC TTC ATA GCT CCA AGA 829
TAT GCA TAC AAA ATC GTC ANG AAG GGG GAC TCG GCA ATC ATG AAA AGC GGC TTG GCA TAT GGT AAC TGT GAT ACT 904
AAG TGT CAA ACT CCA GTG GGT GAG ATA AAC TCC AGT ATG CCC TTC CAC AAT ATA CAC CCT CAC ACC ATT GGG GAA 979
TGC CCT AAA TAC GTG AAA TCA GAT AGG TTG GTC CTT GCA ACT GGA CTC AGA AAC GTC CCT CAA AGG AAG AAA AGA 1054
GGT CTA TTT GGG GCT ATA GCG GGC TTT ATT GAA GGG GGA TGG CAG GGA ATG GTA GAC GGT TGG TAT GGA TAC CAT 1129
CAT AGC AAC GAA CAG GGA AGT GGG TAT GCT GCA GAC AAA GAG TCC ACT CAA AAG GCA ATT GAT GGA ATC ACC AAC 1204
AAG GTC AAT TCA ATC ATT GAC AAA ATG AAC ACC CCG TTT GAG GCC GTC GGA AAG GAA TTT AAT AAT TTG GAG AGG 1279
AGG GTA GAG AAT TTG AAC AAG AAA ATG GAA GAC GGG TTT CTA GAT GTT TGG ACT TAC AAT GTG GAA CTT CTA GTC 1354
CTT ATG GAA AAT GAA AGG ACT CTG GAT TTC CAT GAC TCA AAC GTC AAC AAC CTT TAT GAT AAA GTC CGA CTA CAG 1429
CTT AAG GAC AAC GCG AGG GAA CTG GGA AAT GGT TGT TTC GAG TTC TAC CAC AAA TGT GAC AAT GAA TGT ATG GAA 1504
AGT GTG AGA AAC GGA ACA TAT GAC TAC CCA CAG TAT TCA GAA GAA GCA AGA CTA AAC AGG GAG GAA ATA AGT GGA 1579
GTT AAA TTG GAA TCA ATG GGG GTT TAT CAA ATA CTG TCA ATT TAT TCA ACA GTG GCA AGT TCC CTA GCA CTG GCA 1654
ATC ATG ATA GCT GGT CTA TGT TTT TGG ATG TGT TCC AAT GGG TCA TTG CAG TGC AGA ATT TGC ATT TGA  ATTTA 1728
TTGAGTTCAGATTGTAGTTAAAAACACCCTTGTCTACT 1768

```

B:

```

MRKIV LLLAI VSLVK SDQIC IGYHA NKSTK QVDTI MEKNV TVTHA QDILE RTHNG KLCSL NGVKP LILRD CSVAG YLLGN PNCDE 85
      N K I          KK          K          K          LV  D
      E  F  AR      N          K          K          W

FLNVP EWLVI VEKDN PINSI CYPGD FNDYE ELKYL LSSTN HFEKI RIIPR SSSYN HDASS GVSSA CPYIG RSSFL RNVVV LIKKN 170
      A  S  N  G          H  V  L          N  T          HL  F
      S          VG          H  C  K          D  P  E  L          N  F

NTYPT IKRSY NTINQ EDLLI LMGIH HPNDA ABQTK LYQNP TTYVS VGTST LNQRS IPEIA TRPKV NGQSG RMEFF WTILK PNDAI 255
      V  T  V          A  L  NA  -          K          R  T
      A  S  L          K          L          S  T

NPNES GNPIA PRYAY KIAKK GDSAI NKSGL AYGNK DTKCQ TPVGE INSSH PFHNI HPHTI GBCPK YVKSQ RLVLA TGLRN VPKRK 340
      S  T          E  V          R  E  E          LVA          V  L          K  H  K
      E  V          E  N  IA          L          T

-K-RG LFGAI AGFIE GQWQG MVDGN YGYHH SNEQG SGYAA DKEST QKAID GITNK VNSII DKMNT RFEAV GKEFN NLERR VENLN 425
      I  G          K          Q  R  I
      R  K  H          Q          I

KKMED GFLDV WTYNV ELLVL MENER TLDPH DSNVN NLYDK VRLQL KDNAR ELGNG CFEFY HKCDN RCHES VREMT YDYPQ YSEEA 10
      DL  I  A          L  K          R  K  W          N  K  S
      IL          A          A  K  S          S          I  N

RLNRE EISGV KLESH GYVQI LSIYS TVASS LALAI MIAGL SFWMC SNGLS QCRIC I 566
      K  K  D  I  W  T          V
      I  I
    
```

Figure 1. A. Nucleotide sequence (mRNA sense) of the cDNA clone of the hemagglutinin gene of influenza A/Chicken/Scotland/59 (H5N1) virus. B. Comparison of the deduced amino sequences of the hemagglutinin genes of A/Chicken/Scotland/59 (top), A/Turkey/Ireland/83 (middle), and A/Chicken/Pennsylvania/83 (bottom) viruses. The signal peptide and those residues involved in the cleavage of HA into HA1 and HA2 are underlined. Potential glycosylation sites are indicated by double underlines.

ACKNOWLEDGEMENTS:

The authors thank Renee Black, Brian Holloway, and Vroni Knott for technical assistance. This work was also supported by the MRC, U.K. (Dr. G.G. Brownlee).

*To whom correspondence should be addressed

REFERENCES:

1. Huddleston, J.A. and Brownlee, G.G., (1982), Nucl. Acids Res. 10, 1029-1038.
2. Kawaoka, Y., Naeve, C.W., and Webster, R.G., (1984), Virology 139, 303-316.
3. Kawaoka, Y., Nestorowicz, A., Alexander, D.J., and Webster, R.G., (1987), Virology 158, 218-227.
4. Chen, E.Y. and Seeberg, P.H., (1985), DNA 165-171.