Molecular analysis of the haemagglutinin gene of an avian H1N1 influenza virus

F. J. Austin,1* Y. Kawaoka2 and R. G. Webster2

1Virus Research Unit, The Medical Research Council of New Zealand, University of Otago, P.O. Box 56, Dunedin, New Zealand and 2Department of Virology and Molecular Biology, St Jude Children's Research Hospital, 332 N. Lauderdale, P.O. Box 318, Memphis, Tennessee 38101, U.S.A.

This study presents the first nucleotide sequence and deduced primary amino acid sequence of a subtype H1 haemagglutinin from the avian influenza virus A/duck/Alberta/35/76 (H1N1). The molecule is structurally, antigenically and molecularly similar to H1 haemagglutinins of human viruses but sequence homology differences indicate that there has not been a recent transfer of haemagglutinin genetic information between them.

The sudden appearance of pandemic strains of influenza A viruses in the human population led to the development of the hypothesis that these 'new' viruses may arise by genetic reassortment between human strains and influenza viruses from other mammals or birds (Laver & Webster, 1979). The close genetic and antigenic relationships between the haemagglutinins (HA) of the early human prototype H3N2 strains A/Hong Kong/68 (A/HK/68) and A/Aichi/68 and those of influenza A(H3) viruses isolated from ducks (Laver & Webster, 1973; Scholtissek et al., 1978; Ward & Dopheide, 1981; Fang et al., 1981) provides strong circumstantial evidence that the HA gene of A/HK/68 was derived from a duck influenza virus. The high degree of nucleotide and deduced amino acid homology between subtype H3 HA genes from human and duck isolates (Fang et al., 1981; Kida et al., 1987) corroborates this.

The evidence for a similar origin of the H1 HA of human viruses is less compelling. Historically, it appears to have emerged as a 'new' antigen on the swine influenza-like virus of the Spanish influenza pandemic which caused more than 20 million deaths in 1918 to 1919 (Palese & Young, 1983). Influenza A viruses of the H1 subtype infect humans, pigs and birds (Easterday, 1975; Webster et al., 1984) and genetic and antigenic relationships have been demonstrated between the subtype H1 HAs of viruses isolated from these hosts (Scholtissek et al., 1983; Hinshaw et al., 1984; Aymard et al., 1985; Austin & Webster, 1986). However, nucleotide and deduced amino acid sequence comparisons between subtype H1 HA genes of isolates from humans, birds and mammals have not been possible because, despite a plethora of sequence data on human isolates (Daniels et al., 1985; Raymond et al., 1986; Stevens et al., 1987; Cox et al., 1989), there is only one published sequence of a swine influenza-like H1 HA (Both et al., 1983) and none at all from avian isolates of this subtype.

In this paper we report the nucleotide and deduced amino acid sequences of the HA of A/duck/Alberta/35/76(H1N1) (A/duck/Alberta) and compare them with those published for subtype H1 human isolates A/NJ/11/76(X-53a), A/PR/8/34 and A/USSR/90/77. In addition, antigenic sites on the primary amino acid sequence of A/duck/Alberta were mapped by comparing sequences of monoclonal antibody-selected neutralization escape mutants with the sequence of the parent virus. These analyses provided information on the extent of sequence similarities in HAs of subtype H1 viruses isolated from different species.

The source of influenza A/duck/Alberta, which has a swine influenza virus-like HA, and the method of selecting neutralization escape mutants with monoclonal antibodies have been described in a previous paper (Austin & Webster, 1986). The viruses were grown in embryonated hen eggs, then concentrated from allantoic fluid and purified by two cycles of sucrose density gradient centrifugation. Viral RNA was isolated by treatment of purified virions with proteinase K and SDS, followed by extraction with phenol–chloroform (1:1) (Bean et al., 1980). Full-length dsDNA copies of the HA genes were prepared and blunt-end ligated into the PstI site of pATX153, as described by Kawaoka et al. (1987). Nucleotide sequences were determined by the chain terminating method of Sanger et al. (1977) using reverse transcriptase with either ds cDNA extracted from plasmids or virion RNA as the template. Oligonucleotide
primers (18-mer) complementary to sections of the viron HA gene were synthesized on an Applied Biosystems 380A DNA synthesizer. Sequence data were stored, edited, compiled and analysed using programs available in the Department of Biochemistry, University of Otago, Dunedin, New Zealand (Fraser et al., 1990). The amino acids have been numbered to maximize homology with the H3 HA sequence (Winter et al., 1981), making them structurally equivalent to the amino acids in the model proposed by Wiley et al. (1981).

The nucleotide sequence and deduced amino acid sequence of the entire coding region are shown in Fig. 1. Analysis of the amino acid sequence predicted a protein structurally similar to those of other influenza A virus HA molecules as indicated by the hydropathy profile (not shown), the conservation of the single arginine amino acid which comprises the receptor binding pocket (Tyr 98, His 183, Trp 194) and the second stabilizing shell (Cys 97, Pro 99, Cys 139, Phe 147, Tyr 195, Arg 229) (Wiley & Skehel, 1987). There were seven potential glycosylation sites on the molecule, five on HA1 and two on HA2, but it is not known which are glycosylated.

The nucleotide sequence homology between A/duck/Alberta and the published sequences of A/NJ/11/76, A/PR/8/34 and A/USSR/90/77 was about 77%, and the amino acid homology between 83 to 86 of (Fig. 2). These values are significantly less than the 91 to 97% amino acid homology and the 95 to 98% amino acid homology between the sequences of H1 HA genes of human and duck viruses (Fang et al., 1981; Kida et al., 1987). It was this high homology between the H3 HA sequences which established that changes in sequence between a human virus and an avian one such as A/HK/68 received its HA gene by genetic reassortment between the H1 HAs indicates that there has been no recent exchange of HA genes between these viruses (Austin & Webster, 1986; F. J. Austin, unpublished results) that the HA of A/duck/Alberta HA gene were synthesized on an Applied Biosystems 380A DNA synthesizer. Sequence data were stored, edited, compiled and analysed using programs available in the Department of Biochemistry, University of Otago, Dunedin, New Zealand (Fraser et al., 1990). The amino acids have been numbered to maximize homology with the H3 HA sequence (Winter et al., 1981), making them structurally equivalent to the amino acids in the model proposed by Wiley et al. (1981).

In order to map the antigenic regions on the primary amino acid sequence of the HA gene, the HA

Fig. 1. Nucleotide and deduced amino acid sequences of the coding region of the A/duck/Alberta HA gene. The sequence is shown in the mRNA sense and numbering is from the first nucleotide in the non-coding region at the 3' end of the gene. Underlined amino acids indicate potential glycosylation sites. Synthetic oligonucleotides used as sequencing primers were derived from nucleotides 16 to 35, 186 to 205, 326 to 353, 516 to 533, 723 to 740, 892 to 909, 1047 to 1064, 1231 to 1248 and 1429 to 1447.
The signal peptide. The failure to detect regions equivalent to the Sa or Sb immunodominant regions of the HA molecule of A/PR/8/34 (Caton et al., 1986) was sequenced and compared with that of the parent virus. All selected escape mutants are indicated by arrows. Potential glycosylation sites are underlined.

Fig. 2. The deduced amino acid sequence of A/duck/Alberta compared with published sequences of A/NJ/11/76 (Both et al., 1983), A/PR/8/34 (Winter et al., 1982) and A/USSR/90/77 or the antigenically related H1 HA of A/NJ/11/76. However, because variants of avian influenza viruses co-circulate, further sampling may reveal avian H1 viruses with HAs which are more closely related to those of human strains.

We thank Daniel Channel and Cathy McCaughan for providing excellent technical assistance. This work was supported by the Medical Research Council of New Zealand, U.S. Public Health Service contract AI 52586, grant AI 08831 from the National Institute for Allergy, Infectious Diseases Core grant CA 21765 and American Lebanese Syrian Associated Charities.

References


(Received 21 February 1990; Accepted 5 June 1990)