The Hemagglutinins of Duck and Human H1 Influenza Viruses Differ in Sequence Conservation and in Glycosylation

MICHELLE D. INKSTER,¹ VIRGINIA S. HINSHAW,² AND IRENE T. SCHULZE^{1*}

Department of Molecular Microbiology and Immunology, St. Louis University School of Medicine, 1402 South Grand Boulevard, St. Louis, Missouri 63104-1004,¹ and University of Wisconsin-Madison School of Veterinary Medicine, Madison, Wisconsin 53706²

Received 15 July 1993/Accepted 20 September 1993

We determined the deduced amino acid sequences of two H1 duck influenza A virus hemagglutinins (HAs) and found that the consensus sequence of the HA, determined directly from virus recovered from the intestinal tract, remains unchanged through many generations of growth in MDCK cells and chicken embryos. These two duck viruses differ from each other by 5 amino acids and from A/Dk/Alberta/35/1976 (F. J. Austin, Y. Kawaoka, and R. G. Webster, *J. Gen. Virol.* 71:2471–2474, 1990) by 9 and 12 amino acids, most of which are in the HA1 subunit. They are antigenically similar to each other but different from the Alberta virus. We compared these H1 duck HAs with the HAs of human isolates to identify structural properties of this viral glycoprotein that are associated with host range. By comparison to the human H1 HAs, the duck virus HA sequences are highly conserved as judged by the small fraction of nucleotide differences between strains which result in amino acid substitutions. However, the most striking difference between these duck and human HAs is in the number and distribution of glycosylation sites. Whereas duck and swine viruses have four and five conserved glycosylation sites per HA1 subunit, none of which are on the tip of the HA, all human viruses have at least four additional sites, two or more of which are on the tip of the HA. These findings stress the role of glycosylation in the control of host range and suggest that oligosaccharides on the tip of the HA are important to the survival of H1 viruses in humans but not in ducks or swine.

Fourteen antigenically distinct hemagglutinin (HA) subtypes of influenza A virus have been identified (18), and only three have been isolated from human cases of influenza. The remaining subtypes have been found in a variety of mammalian and avian species, in which they produce infections ranging from asymptomatic to rapidly fatal. Each of the subtypes appears to be largely restricted to one or only a few animal species, but intraspecies transfers have been observed (for a review, see reference 35). Thus, despite the restrictions in host range exhibited by these viruses, infected animals apparently provide reservoirs from which a virus can be introduced into the human population. Aquatic birds appear to be important in this regard. Most influenza A virus strains which grow in wild ducks cause asymptomatic infections in which the virus replicates in cells of the respiratory and intestinal tracts and is excreted into lake water (36). These findings, along with the fact that most, if not all, of the influenza A virus subtypes have been found in wild ducks, have led to the hypothesis that aquatic birds provide an important reservoir from which influenza A viruses can move into the human population, either directly or by way of other animal species (35).

Since the influenza virus genome is replicated by a viral RNA polymerase which apparently lacks proofreading activity, influenza virus populations contain minority components which have genomes that differ from the consensus sequence. These minority components can be amplified when growth conditions favor their replication so that a new consensus sequence can be established by growth of the virus in a new host.

The work described here was initiated to investigate the role of the HA gene in this selection process and is based on the

finding that host cells can exert selective pressures on virus populations both by providing cell-specified sialoglycoconjugates to act as receptors and by altering the binding properties of the virus by cell-specific glycosylation of the HA (7, 12). These two effects together would be expected to select HA variants which are present in virus populations in nature or arise during replication in laboratory hosts. We have attempted to demonstrate this type of selection by growing H1 duck viruses in two different hosts, starting with virus samples taken directly from the intestinal tracts of ducks. We used chicken embryos (CE), a laboratory host reported to select HA variants from some human isolates (16, 17, 27, 34), and Madin-Darby canine kidney (MDCK) cells, a cell line which supports the growth of most human influenza virus isolates. The consensus sequence of the HA gene of laboratory-grown virus populations was then compared to that of virus taken directly from the intestines of infected ducks. Secondly, we compared the HA consensus sequence and antigenic properties of three H1 duck viruses, the two used in this study and A/Dk/Alberta/35/76, the only other H1 duck virus which has been sequenced (2). Our data prompted us to compare the positions at which glycosylation sequons are found in H1 viruses recovered from different species. We present evidence here that the distribution of oligosaccharides on the HA1 subunits is a host-dependent characteristic of H1 viruses.

MATERIALS AND METHODS

Virus strains and virus growth. The virus used in this study was obtained by infecting ducks via the cloaca with fecal samples taken from infected wild ducks in Wisconsin in 1980. Two isolates, A/Dk/WI/1938/80 (WI/1938) and A/Dk/WI/259/80 (WI/259), were investigated. Virus recovered from the feces and from intestinal tissue taken 48 h after infection was examined directly and used to infect CE and MDCK cells.

* Corresponding author. Electronic mail address: IN%“SCHULZEIT@SLUVCA”.

Each virus sample was then carried through 10 serial passages in each host.

Passage of virus in the allantoic cavity of 11-day-old embryonated eggs (SPAFAS) was done by injecting 0.2 ml of a 1:20 dilution of virus in phosphate-buffered saline containing 0.2% bovine serum albumin, 10,000 U of penicillin per ml, 10 µg of streptomycin per ml, and 40 µg of gentamicin per ml. Allantoic fluid harvested after 3 days at 37°C contained 128 to 512 hemagglutinating units of virus per ml after centrifugation to remove cellular debris.

Passages of virus in MDCK cells were done by infecting washed monolayers in 75-cm² plastic flasks with 1.0 ml of virus diluted 1:100 in standard basic salt solution (22) containing 0.2% bovine serum albumin. After adsorption for 45 min at room temperature, monolayers were overlaid with a 1:1 mixture of medium 199 and Dulbecco's modified Eagle medium (DMEM; GIBCO) containing 1 µg *N*-tosyl-L-phenylalaninyl chloromethyl ketone (TPCK)-treated trypsin (Sigma) per ml, 1 mM glutamine, 7.5 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid (HEPES), 2,000 U of penicillin per ml, 1 µg of streptomycin per ml, and 4 µg of gentamicin per ml. Infected cultures were incubated at 37°C in the presence of 5% CO₂. Cell culture fluid harvested after 48 h contained 256 to 512 hemagglutinating units of virus per ml after centrifugation to remove cellular debris.

Extraction of RNA. Virus from fecal material, intestinal tissue, allantoic fluid, and cell culture fluid was cleared of debris by centrifugation at 10,000 × *g* for 10 min at 4°C and pelleted by centrifugation at 105,000 × *g* at 4°C for 2 h. The virus was resuspended in 0.1 M NaCl–50 mM Tris (pH 7.2)–1 mM EDTA (STE), and an equal volume of 0.3 M NaCl–0.2 M Tris (pH 7.4)–25 mM EDTA–2% sodium dodecyl sulfate (SDS) was added. After addition of 160 µg of proteinase K (Boehringer Mannheim) and incubation at 37°C for 1 h, 20 µg of glycogen was added and the RNA was extracted with phenol-chloroform (1:1). Following reextraction with chloroform-*n*-butanol (4:1), the RNA was again ethanol precipitated as previously described (25). The purified RNA was stored in ethanol at –70°C until used.

Preparation of DNA complementary to virion RNA (cDNA). Except as noted below, reverse transcription was done after denaturation of the RNA with the reducing agent methylmercuric hydroxide as described by Davis et al. (9). In this procedure, the RNA is pretreated with 0.1 M methylmercuric hydroxide for 10 min at room temperature, after which β-mercaptoethanol is added to a concentration of 25.5 mM, followed 10 min later by the reagents for reverse transcription. cDNA was synthesized in a volume of 55 µl containing 90 mM Tris (pH 8.3), 125 mM KCl, 9 mM MgCl₂, 0.9 mM each deoxynucleoside triphosphate, 60 U of RNasin (Promega), 5 µg of primer vir3', 30 U of avian myeloblastosis virus reverse transcriptase, and virion RNA. (Primer vir3' is complementary to the conserved 12-mer at the 3' end of all virion segments.) After incubation at 42°C for 2 h, EDTA was added to 20 mM and the cDNA was extracted with phenol-chloroform (1:1) and precipitated with ethanol.

HA gene amplification. DNA copies of the entire HA gene were synthesized by the polymerase chain reaction using the procedure previously used for amplification of the HA1 region of virus from clinical samples (25), except that the concentration of each of the deoxynucleoside triphosphates was 0.200 mM and 100 µg of each of primers Dvir3' and 1786D was used (see Results). These primers contained *Bam*HI and *Hind*III restriction sites at their 5' ends. Amplification was carried out for 30 cycles in a Coy model 50 tempeycycler, and each cycle consisted of 1 min of denaturation at 94°C, 2 min of annealing

at 45°C, and 2 min of extension at 72°C, with a final extension of 7 min at 72°C. Reactions were set up in a biohazard hood, and appropriate precautions and controls were used to document template dependence and freedom from cross-contamination.

DNA cloning. HA double-stranded DNA that had been amplified as described above was isolated by electrophoresis on 1.5% agarose and eluted into 15 µl of TE (25) with a GENECLON kit (Bio 101, Inc.). Following digestion with *Bam*HI and *Hind*III, the DNA was ligated into the multiple cloning site of pUC19 and *Escherichia coli* DH5α cells were transformed with the ligation mixture. Transformed cells, selected by growth in the presence of ampicillin, were plated on agar, and clones containing the HA gene were isolated and used for sequence analysis.

Nucleotide sequencing. Plasmid DNA that had been treated with RNase A (Sigma) was denatured with 0.2 N NaOH and 3 mM EDTA (pH 8.0) and sequenced with Sequenase kits (United States Biochemical). The sequences adjacent to the conserved regions at the ends of the HA gene were determined with M13 universal forward and reverse primers. The remainder of the sequence was determined with HA-specific primers.

Hemagglutination inhibition (HI) tests. Virus to be used in HI tests was partially purified by centrifugation through 30% sucrose in STE and resuspended in the same buffer. HI tests were carried out as previously described (23), with monoclonal antibodies (MAbs) against A/Sw/WI/27/86 and A/Sw/IN/1726/88 (31) and against A/Dk/Alberta/35/76 (1). The latter antibodies were kindly provided by R. Webster, St. Jude Children's Research Hospital, Memphis, Tenn.

Western immunoblot analysis. Western blots were done as previously described (3), except that the blots were treated with MAbs in a solution containing 6.5 mM potassium phosphate, 13.5 mM sodium phosphate, 2 mM calcium chloride, and 1% nonfat dried milk (pH 6.28).

Nucleotide sequence accession numbers. The GenBank/EMBL accession numbers for A/Dk/WI/259/80 and A/Dk/WI/1938/80 are L25071 and L25072, respectively.

RESULTS

Determination of the sequence of the HA gene directly from virus in feces and after growth in MDCK cells and CE. Before the HA sequences of native and laboratory-grown viruses could be compared, conditions which would provide accurate sequence data directly from duck feces had to be determined. To do this, normal duck feces to which known amounts of the WSN strain of influenza A virus (H1N1) had been added were used as starting material. These trials indicated that intact WSN HA RNA could be recovered from feces and that DNA copies of the HA1 region of the HA gene could be obtained, amplified by polymerase chain reaction, and sequenced from aliquots of fecal material containing 50 hemagglutinating units of virus. However, attempts to obtain full-length HA DNA from the duck virus by this procedure were unsuccessful, even when large amounts of RNA obtained from CE-grown or MDCK cell-grown virus were used. As shown in Fig. 1, lanes 1 and 2, treatment of the duck virus virion RNA with methylmercuric hydroxide (see Materials and Methods), an agent which reduces secondary structure in RNA (29), was required for the synthesis and subsequent amplification of full-length HA DNA. Since the primers used in these experiments were complementary to the conserved 12-mer and 13-mer at the ends of all of the virion RNAs (Fig. 1, legend), this amplified DNA contained, in addition to the HA DNA, large amounts of a smaller DNA (lane 2) which proved to be copies of the NS

FIG. 1. Agarose gels showing the effect of RNA denaturation and primer specificity on synthesis of HA DNA by reverse transcription and polymerase chain reaction amplification. Lanes M, molecular size markers. Lane 1, duck virus virion RNA not treated prior to cDNA synthesis. Primers: B₁vir3' (5'-dGCGTGGATCCAGCAAAAGCAGGGGA-3') and 1786 (5'-dCGCGAAGCTTAGTAGAAACAAGGGTGT-3'). Lane 2, duck virus virion RNA treated with methylmercuric hydroxide prior to cDNA synthesis as described in Materials and Methods. Primers: B₁vir3' and 1786. Lane 3, WSN positive control. Lane 4, duck virus virion RNA treated with methylmercuric hydroxide prior to cDNA synthesis. Primers: Dvir3' (5'-dGCGTGGATCCAGCAAAAGCAGGGGATA) and 1786D (5'-dCGCGAAGCTTAGTAGAAACAAGGGTGT-3'). Underlined bases represent cloning sites, and bases in boldface type represent the conserved 12-mer and 13-mer at the ends of the virion RNA segments.

gene. NS DNA was eliminated from the amplified pool by using primers that have HA-specific sequences at their 3' ends (lane 4).

The sequence of the HA gene segment was determined from cloned DNA as described in Materials and Methods. This procedure was used to avoid the large number of amplification cycles which would have been required to obtain enough DNA to determine the sequence of the entire HA gene directly from the amplified pool. Five clones from each amplified DNA pool were sequenced to ensure that the sequence obtained was an accurate representation of the consensus sequence.

To determine whether growth of the duck virus in laboratory hosts would lead to selection of HA variants, the sequence of the HA genes of WI/1938 was determined directly from the feces, from passages 2 and 10 in MDCK cells and passages 3 and 10 in CE. Duplicate passages were done in both laboratory hosts, so that 10 clones of each virus at each passage were examined. Identical nucleotide sequences were obtained from the original sample and all of the CE-grown and MDCK cell-grown viruses, except that all five clones obtained from the original fecal sample had thymidine (T) residues at positions 797 and 1716 whereas all other clones had cytidine (C) at both of those positions. Neither of these variations altered the amino acid sequence; C versus T at position 797 does not change the codon, and nucleotide 1716 is downstream from the stop codon. Analysis of the sequence differences between

individual clones indicated (i) that there were four single-base differences from the consensus sequence among the 10 clones from passage 3 in CE and from passage 10 in MDCK cells, indicating a misincorporation rate for each of those samples of 1 in 4,392 bases, and (ii) that there was no deviation from the consensus sequence in the virus taken directly from feces (misincorporation rate, less than 1 in 8,785 bases) or from passage 10 in CE or passage 2 in MDCK cells (misincorporation rate of each sample, less than 1 in 17,570 bases). Thus, the overall misincorporation rate due to the combined effect of the reverse transcriptase, the *Taq* enzyme, and the influenza virus RNA-dependent RNA polymerase was 1 in 9,800 bases. The data indicate that the original sample did not consist of a mixed population, as might have been expected from published data (13). Selection of minority forms with HA sequences which varied from the original consensus sequence had not occurred during prolonged growth in either laboratory host.

Comparison of the sequences and antigenic properties of three duck virus HAs. The HA gene of a second 1980 isolate, WI/259, was then sequenced so that the amount of variation between these two 1980 isolates and an H1 virus isolated in 1976, Alb/35 (2, 14), could be compared. Figure 2 presents a diagram of the HA gene in which the relative positions at which each of the Wisconsin isolates differs from Alb/35 are illustrated by lines perpendicular to the length of the gene. The diagram extends from the first nucleotide in the coding region at the 5' end of the mRNA through the stop codon at nucleotide 1713. Differences were found at 128 positions and were distributed throughout the gene. However, when only differences which change the amino acid sequence were considered, 13 differences between the two Wisconsin isolates and Alb/35 were found, 7 of which are clustered in a region which constitutes 12% of the HA sequence and encodes amino acids 163 to 233 of HA1. In this region of the gene, 44% of the nucleotide differences between the Wisconsin strains and Alb/35 result in amino acid differences, compared with 16% in the HA gene as a whole. The data indicate the presence of a variable region in an otherwise highly conserved sequence.

The deduced amino acid sequence of this region is shown in Fig. 3, which compares the sequences of the three duck viruses and A/SW/IN/1726/88, an H1 swine virus known to be antigenically related to H1 isolates from ducks (31). The two Wisconsin isolates differ from each other at three positions in this region and from Alb/35 at four and five.

The positions on the HA subunit of the residues which differ in the three duck viruses are illustrated in Fig. 4. Residue 163 (H3 numbering [38]) is on the surface of the subunit, and residues 187, 188, and 189 are on the tip, close to the receptor binding pocket. Thus, these viruses differ in regions of the HA which are known to be important in receptor binding and in which single amino acid substitutions permit the virus to escape neutralization by antibodies. MAbs which had been

FIG. 2. Nucleotide differences in the coding regions of three H1 duck HAs. The HA gene is depicted as a solid bar for HA1 (residues 33 to 1059) and a stippled bar for HA2 (residues 1060 to 1713). Silent nucleotide differences are represented by short vertical lines. Coding changes are represented by elongated lines that extend through the bar at positions at which the two Wisconsin HA genes have the same nucleotide. The amino acid differences are indicated with the single-letter code and H3 numbering (38); the first letter indicates the amino acid in Alb/35.

FIG. 3. Deduced amino acid sequences of four H1 HAs extending from residues 150 to 239. H3 numbering is used. Symbols: ■, amino acids found in the Sa antigenic site; ●, amino acids in the Sb site; ▲, amino acids in the Ca site; ◇, five of the six residues which constitute the receptor binding pocket (37).

characterized by the ability to react with this region of H1 viruses from swine and ducks were therefore used to characterize the Wisconsin viruses. Experiments were first done with WI/1938 and pooled MAb's in a Western blot analysis. One pool consisted of two MAb's, 7B1b and 1-6B2, which react with residues 157 and 158 in the Sa region of the swine HA (31), and the other pool consisted of two MAb's, 71 and 88, against CE-grown Alb/35 (1), one of which, MAb 88, has been mapped to amino acid 188 of HA1 (1a). As expected from the sequence differences shown in Fig. 3, the two swine virus MAb's reacted weakly with both Alb/35 and WI/1938, whereas the two duck virus MAb's reacted strongly with Alb/35 and very weakly with WI/1938 (data not shown).

The antigenic differences between the three duck HAs were further investigated by HI tests with virus grown in both CE and MDCK cells. Experiments were initially carried out with both whole virus and ether-extracted HA, since earlier findings indicated that ether-extracted HA was required for detection of HI activity with some duck viruses (20). With the three viruses studied here, HI titers obtained with whole virus were reproducible from experiment to experiment whereas those obtained with ether-disrupted virus varied widely, apparently depending on the specific hemagglutinating activity of the extracted HA protein.

Table 1 presents data obtained with the MABs described above, as well as with three additional MABs against Alb/35 which showed a wide range of reactivity with H1 isolates from avian, swine, and human sources classified as group I (1). Compared with the Alberta strain, both of the Wisconsin viruses showed low levels of reactivity with all of the MABs used in the study. With virus from CE, HI titers of 100 or less were obtained with both of the Wisconsin viruses, whereas with MDCK cell-grown virus, higher titers were observed with all of the MABs except 71 and 88. Since the amino acid consensus sequences of the CE-grown and MDCK cell-grown Wisconsin isolates are identical, this host effect is apparently a consequence of the properties of the oligosaccharides attached to the HA by the two host cells.

On the basis of the results obtained with MDCK cell-grown viruses, the two Wisconsin viruses showed a fourfold difference in the HI titer in response to MA b 1-6B2, which recognizes

FIG. 4. Alpha-carbon tracing of the HA subunit showing positions at which the two Wisconsin viruses differ from Alb/35. The arrow indicates the position of the receptor binding pocket. H3 numbering is used.

antigenic site Sa (5), and MAb 24 (epitope unknown). The difference in reactivity with MAb 1-6B2 is probably due to the fact that WI/259 has threonine at amino acid 163 whereas WI/1938 and Alb/35 both have lysine at that position (Fig. 3). The data in Table 1 indicate that the reactivity of MDCK cell-grown WI/1938 with MAb 1-6B2 is, in fact, highly similar to that of Alb/35.

Distribution of glycosylation sites on the HA of H1 viruses. Our sequence data show that both of the Wisconsin strains resemble the Alberta virus in having four glycosylation sequons in HA1, counting the two overlapping glycosylation sequons (Asn-Asn-Ser-Thr) found in all H1 strains at amino acids 20 and 21 as one. Evidence that all four of these sites are glycosylated by both host cells was obtained by comparing the size of the HA1 subunits of WI/1938 to those from two variants of the WSN strains, which are composed of 337 amino acids (1 less than in the duck virus HA1 subunit). These two variants differ from each other by an Asp-to-Asn substitution at residue 129 in the HA1 subunit, and as a consequence, the one variant (C_f) has three glycosylation sites per HA1 subunit whereas the other (F_n) has four, and analysis of the glycopeptides from

TABLE 1. Comparison of the antigenic properties of three H1N1 avian viruses with MABs

Virus	Cells used for growth	HI titer ^a with MAB ^b						
		7B1b	1-6B2	5	10	24	71	88
Alb/35	CE	32	32	32	8	32	64	128
	MDCK	32	32	16	16	32	128	128
WI/259	CE	<1	1	<1	<1	1	<1	<1
	MDCK	2	4	4	2	2	<1	<1
WI/1938	CE	<1	1	<1	<1	4	<1	<1
	MDCK	<1	16	4	4	8	<1	<1

^a Reciprocal of the highest dilution (10²) of antibodies which inhibited 4 hemagglutinating units of the virus.

^b MABs 7B1b and 1-6B2 were made against A/SW/IN/1726/88 and A/SW/WI/27/86, respectively (31). The remaining MABs were made against A/Dk/Alberta/35/76 (1) and were a gift from R. Webster, St. Jude Children's Research Hospital, Memphis, Tenn.

these two HAs has indicated that all of the glycosylation sites on each HA1 subunit are occupied (30). The sizes of the HA1 subunits of these WSN substrains obtained from both MDCK cells and CE fibroblasts is shown in Fig. 5A. As indicated by their mobility in SDS-polyacrylamide gel electrophoresis, the size of these HA1 subunits is affected by both the number of glycosylation sites per subunit (compare lanes 1 and 2 or 3 and 4) and by the cells in which the virus was grown (compare lanes 1 and 3 or 2 and 4). In Fig. 5B, we compared the size of the HA1 subunits of Dk/WI/1938 to that of the MDCK cell-grown WSN substrain, which has three glycosylation sites per HA1 subunit. As indicated in Fig. 5B, lanes 1 and 2, the HA1 subunit of the duck virus is about 2 kDa larger than that of the WSN

FIG. 5. Sizes of the HA1 subunits of WI/1938 grown in CE and MDCK cells. Following SDS-polyacrylamide gel electrophoresis and transfer of the viral proteins to nitrocellulose, the HA1 subunits of WSN variants F₀ and C₁ in part A were detected by Western blot analysis with MAb Ad2 as previously described (3). In Part B, the HA1 subunits of Dk/WI were first stained with a pool of MABs to Alb/35. Following color development, the blots were reprobed with MAb Ad2 to visualize the C₁ HA. Lanes M contained molecular size markers.

HA obtained from the same host cell, suggesting that the duck virus HA1 subunit contains four oligosaccharides per HA1 subunit. The positions of the HA1 bands in lanes 2 and 3 also indicate that these subunits from MDCK cell-grown virus are larger than those from CE-grown virus. Since the amino acid sequences of the duck HAs from the two sources are the same, this difference reflects the size of the oligosaccharides attached to the HA by the two cells.

Since each of the duck HAs has only four glycosylation sites per HA1, whereas recent human isolates have eight or nine, we used the method of Winter et al. (38) to apply H3 numbering to all sequences and compared the distribution of glycosylation

TABLE 2. Distribution of glycosylation sites on HA1 subunits of duck, swine, and human H1 isolates

Virus source and designation or yr(s) of isolation	No. of strains	Presence of glycosylation at site(s):												
		20, 21	33	63	81	94A	129	131	158	163	225	271	278	289
Ducks														
Alb/35/76 ^a	1	+	+			+								+
WI/259/80	1	+	+			+								+
WI/1938/80	1	+	+			+								+
Swine														
1976 ^{b,c}	2	+	+			+							+ ^d	+
1982 ^c	1	+	+			+							+	+
1988 ^c	2	+	+			+							+	+
1992 ^c	1	+	+			+					+		+	+
Humans														
1951–1952 ^f	2	+	+		+	+ ^d		+ ^d	+	+ ^d		+ ^d		+
1954–1957 ^f	3	+	+		+ ^d	+		+	+	+				+
1977–1978 ^f	3	+	+			+		+	+	+		+		+
1979–1980 ^{f,g}	3	+	+			+		+	+	+ ^d		+		+
1982–1983 ^{f,h}	10	+	+			+		+	+	+ ^d		+		+
1986–1987 ^{i,j,k}	5	+	+	+		+	+		+ ^d	+		+		+
1988–1989 ^{j,m}	12	+	+	+		+	+		– ⁿ	+		+		+
1990–1991 ^{j,m}	8	+	+	+		+	+			+		+		+

^a Austin and Webster (2).

^b GenBank accession no. K00992.

^c Luoh et al. (21).

^d One strain lacks this site.

^e GenBank accession no. L09063.

^f Raymond et al. (26).

^g GenBank accession no. M38353.

^h GenBank accession no. M59324.

ⁱ Cox et al. (6).

^j Xu et al. (39).

^k Rajakumar et al. (25).

^l GenBank accession no. D13574.

^m Rajakumar and Schulze (24a).

ⁿ One strain has this site.

TABLE 3. Distribution of glycosylation sites on HA1 subunits of duck, swine, and human H1 isolates

Virus source	No. of strains	No. of strains glycosylated at site(s):											
		20, 21	33	63	81	94A	129 or 131 ^a	158	163	225	271	278	289
Ducks	3	3	3	0	0	3	0	0	0	0	0	0	3
Swine	6	6	6	0	0	6	0	0	0	1	0	5	6
Humans	46	46	46	25	4	45	45	25	43	0	42	0	46

^a Sites 129 and 131 are considered one site, since only one or the other can be glycosylated.

sites on the HA1 subunits of H1 viruses from these two sources and from swine. As shown in Tables 2 and 3, the duck virus HAs have glycosylation sites at positions 20 and 21, 33, 94a, and 289. They do not have the sites at positions 63, 81, 129 or 131, 158, 163, and 271 found on human H1 viruses. The duck virus HAs are, in fact, more similar to H1 HAs from swine viruses than to those from human viruses. Except for the site at position 94a, which is present on the HAs from all three sources, and the site at position 225 found on one of the swine

viruses, there are no glycosylation sites on the globular head of either the duck or the swine HAs (Fig. 6). In contrast, all of the human strains except one isolated in 1952 (Table 2) have, in addition to the conserved sites, two to five sites on the globular head of the HA (see the number of strains glycosylated at each site in Table 3). Sites 129 and 131 and site 158 are on the tip of the HA, and site 163 is on the surface of the globular head. All of these sites are close to the receptor binding pocket and/or are located in antigenic sites (Fig. 3). In addition, the sites at positions 63, 81, and 271 on the human virus HAs are on the globular head of the HA. The data reveal that the distribution of glycosylation sequons on the HA is a host-associated property of H1 viruses.

DISCUSSION

These studies were initially undertaken to determine whether host cell selection of HA variants would occur during growth of duck viruses in laboratory hosts. Our data indicate that viruses taken directly from a duck intestinal tract can grow for many generations in both MDCK cells and CE without evidence of host-determined selection of HA variants. Thus, with the WI/1938 virus, a change in the consensus sequence of the HA was not required for the virus to replicate in cells of either avian or mammalian origin.

With respect to the replication of these H1 avian viruses in nature, our sequence data suggest that selection by antibodies, receptor binding, or a combination of the two is involved in determination of the HA consensus sequence. Whereas nucleotide differences between the Alberta and Wisconsin isolates were found to be distributed throughout the HA gene, as would be expected from random misincorporation by the viral RNA-dependent RNA polymerase, 10 of the 13 amino acid differences between the Wisconsin and Alberta viruses are in the HA1 subunit and 4 of them are in recognized antibody-binding sites. Given these differences, it is not surprising that the two Wisconsin isolates are distinct from the Alberta virus in antigenic properties as measured by both Western blots and HI tests.

We found that MDCK cell-grown virus consistently gave slightly higher HI titers than did CE-grown virus, even though the HAs from the two sources have identical amino acid sequences. The basis for this difference remains to be determined. The data suggest that either the affinity of the HA for chicken erythrocytes is reduced or its affinity for MAb is increased by the oligosaccharides put onto the HA by MDCK cells. Of these two possibilities, lower affinity of MDCK cell-grown virus for erythrocytes seems to be the more plausible explanation. Since the MAbs used in these experiments were made with CE-grown virus, it is difficult to envision how oligosaccharides put onto the HA by MDCK cells would increase the affinity of the HA for MAbs to two or more antigenic sites. In addition, the ability of certain substrains of WSN to bind to cellular receptors is reduced when these viruses are grown in cells which add large oligosaccharides to

FIG. 6. Alpha-carbon tracing of the HA subunit showing the positions of glycosylation sites on the HAs of H1 viruses from duck, swine, and human sources. Symbols: ●, glycosylation sites found on HAs from all three sources; ○, those found only on swine HAs; ▲, those found only on human HAs.

the HA1 subunit (3, 7, 11), and as shown in Fig. 5, the HA1 subunits of MDCK cell-grown WI/1938 are larger than those from the CE-grown virus. Our previous work (23), as well as that of others (8, 40), has indicated that differences in HI and virus neutralization titers can be observed with influenza virus strains which do not differ antigenically but do differ in receptor-binding properties.

The data on the distribution of amino acid differences between H1 strains presented here differ from those obtained from the H3 avian viruses isolated at about the same time from migrating ducks in Hokkaido, Japan (19). In those studies, the amino acid sequence differences were found to be randomly distributed throughout the HA gene rather than concentrated in regions of the HA1 subunit that would be expected to alter antibody binding. It was concluded from those studies that, in contrast to what has been found with human H3 strains, antigenically distinct H3 HAs are conserved and cocirculate in ducks. Only three duck H1 isolates have been sequenced, an insufficient number from which to determine strain lineage. However, the data suggest that the duck H1 and H3 HAs show different patterns of evolution. Whether this is indeed the case will be determined as more sequence data from H1 strains become available.

Sequence data from the three H1 duck viruses have revealed that, as with the H3 viruses (19), a large fraction of the nucleotide changes are silent. With HAs of both subtypes recovered from ducks, only 10 to 20% of the nucleotide differences result in amino acid sequence differences. With H1 viruses recently isolated from human sources, 44 to 49% of the nucleotide changes were expressed as amino acid sequence changes (39). Thus, despite the apparent difference in the patterns of evolution between the two subtypes, the duck virus HAs of both subtypes exhibit a high level of protein structure conservation compared with human virus HAs. The reasons for these host-related differences are unknown. It has been postulated that the duck virus is sufficiently adapted to grow within its natural host so that most amino acid substitutions are detrimental to the survival of the virus (35).

Austin et al. (2) have pointed out that the HA of Alb/35 shows less homology with human H1 strains than has been observed between some avian and human H3 strains. They conclude from this low level of homology (77% for nucleotides and 83 to 86% for amino acids) that there has been no recent exchange of HA genes between human and avian H1 viruses. The data presented here support that conclusion; the two Wisconsin isolates described here show about the same level of homology with human isolates as that reported for Alb/35.

Another finding which supports the concept that duck and recently isolated human H1 viruses are of independent lineages is the striking difference in the distribution of the glycosylation sites on the HA1 subunits of viruses from the two sources (Tables 2 and 3 and Fig. 6). These host-associated differences argue in favor of the idea that oligosaccharides on the tip of the HA are important to the survival of H1 viruses in nature in humans but not in either ducks or swine.

These oligosaccharides are not, however, required for growth of the human isolates in embryonated eggs or in cell cultures. It is interesting in this regard to compare the HA1 glycosylation patterns shown in Fig. 6 with those of early H1 isolates such as A/WSN, a neurotropic variant of the WS strain isolated in 1933, and A/PR/8/34. These strains, isolated some 60 years ago, have been grown in a number of different laboratory hosts and do not show the distribution of glycosylation sites seen on recent human isolates. Both of these viruses have the site at position 20 or 21, but each lacks one or more of the sites at positions 33, 94a, and 289. Both were

originally found to be glycosylated at position 129 or 131 (15, 38), but substrains of each which have lost this glycosylation site have been isolated (5, 10). In the WSN strain, loss of this site actually fosters growth of the virus in MDBK cells. One plausible explanation for the presence of oligosaccharides on the tip of the HA is that they promote growth of the virus by masking antigenic sites (32). If this is the explanation, it remains to be determined why this operates in humans but not in ducks and swine.

The similarity in glycosylation between the swine and duck HAs shown in Fig. 6 may facilitate intraspecies transfer and be, in part, responsible for the finding that H1 duck viruses have been implicated in swine epizootics (24) whereas there is no evidence that these viruses cause disease in humans. Furthermore, they suggest that mutations which put glycosylation sites onto the tip of the HA would be required for H1 viruses to replicate efficiently and to exhibit a high level of transmissibility in the human population following transfer from either a duck or a swine reservoir. Studies have indicated that human volunteers are largely refractory to infection with avian viruses (4). In addition, although severe natural infections of humans by swine viruses have occasionally occurred (28, 33), spread of the virus from these individuals or from individuals with occupational exposure to infected swine has not resulted in epidemic disease. A limited ability to replicate in most human hosts because of the lack of oligosaccharides on the tip of the H1 HAs may be responsible for the inability of such transfers to initiate epidemic disease. These findings stress the importance of obtaining the sequences of more H1 HAs from ducks, both to determine whether glycosylation sites are ever found on the tip of the HA and to identify such strains, if they do exist, so that their potential as human pathogens can be investigated.

ACKNOWLEDGMENTS

This work was supported in part by grants AI23520 (to I.T.S.) and AI24920 (to V.S.H.) from the National Institutes of Health and grants 3101 and 3414 (to V.S.H.) from the U.S. Department of Agriculture.

We thank Joseph Eigner for providing starting material for some of these experiments. We thank Judith Standish for skilled secretarial assistance.

REFERENCES

1. Austin, F., and R. G. Webster. 1986. Antigenic mapping of an avian H1 influenza virus haemagglutinin and interrelationships of H1 viruses from humans, pigs, and birds. *J. Gen. Virol.* **67**:983-992.
- 1a. Austin, F. J. Personal communication.
2. Austin, F. J., Y. Kawaoka, and R. G. Webster. 1990. Molecular analysis of the haemagglutinin gene of an avian H1N1 influenza virus. *J. Gen. Virol.* **71**:2471-2474.
3. Aytay, S., and I. T. Schulze. 1991. Single amino acid substitutions in the haemagglutinin can alter the host range and receptor binding properties of H1 strains of influenza A virus. *J. Virol.* **65**:3022-3028.
4. Beare, A. S., and R. G. Webster. 1991. Replication of avian influenza viruses in humans. *Arch. Virol.* **119**:37-42.
5. Caton, A. J., G. G. Brownlee, J. W. Yewdell, and W. Gerhard. 1982. The antigenic structure of the influenza virus A/PR/8/34 haemagglutinin (H1 subtype). *Cell* **31**:417-427.
6. Cox, N. J., R. A. Black, and A. P. Kendal. 1989. Pathways of evolution of influenza A (H1N1) viruses from 1977 to 1986 as determined by oligonucleotide mapping and sequencing studies. *J. Gen. Virol.* **70**:299-313.
7. Crecelius, D. M., C. M. Deom, and I. T. Schulze. 1984. Biological properties of a haemagglutinin mutant of influenza virus selected by host cells. *Virology* **139**:164-177.
8. Daniels, R. S., S. Jeffries, P. Yates, G. C. Schild, G. N. Rogers, J. C.

- Paulson, S. A. Wharton, J. J. Skehel, A. R. Douglas, and D. C. Wiley. 1987. The receptor-binding and membrane-fusion properties of influenza virus variants selected using anti-hemagglutinin monoclonal antibodies. *EMBO J.* **6**:1459–1465.
9. Davis, L. G., M. D. Dibner, and J. F. Battey. 1986. Basic methods in molecular biology, p. 199–202. Elsevier Press, New York.
10. Deom, C. M., A. J. Caton, and I. T. Schulze. 1986. Host cell-mediated selection of a mutant influenza A virus that has lost a complex oligosaccharide from the tip of the hemagglutinin. *Proc. Natl. Acad. Sci. USA* **83**:3771–3775.
11. Deom, C. M., and I. T. Schulze. 1985. Oligosaccharide composition of an influenza virus hemagglutinin with host determined binding properties. *J. Biol. Chem.* **260**:14771–14774.
12. Hearing, J., M.-J. Gething, and J. Sambrook. 1989. Addition of truncated oligosaccharides to influenza virus hemagglutinin results in its temperature-conditional cell-surface expression. *J. Cell Biol.* **108**:355–365.
13. Hinshaw, V. S., W. J. Bean, and G. Sriram. 1980. Genetic reassortment of influenza A viruses in the intestinal tract of ducks. *Virology* **102**:412–419.
14. Hinshaw, V. S., R. G. Webster, and B. Turner. 1978. Novel influenza A viruses isolated from Canadian feral ducks: including strains antigenically related to swine influenza (Hsw1N1) viruses. *J. Gen. Virol.* **41**:115–127.
15. Hiti, A. L., A. R. Davis, and D. P. Nayak. 1981. Complete sequence analysis shows that the hemagglutinins of the H0 and H2 subtypes of human influenza virus are closely related. *Virology* **111**:113–124.
16. Katz, J. M., C. W. Naeve, and R. G. Webster. 1987. Host cell-mediated variation in H3N2 influenza viruses. *Virology* **156**:386–395.
17. Katz, J. M., M. Wang, and R. G. Webster. 1990. Direct sequencing of the HA gene of influenza (H3N2) virus in original clinical samples reveals sequence identity with mammalian cell-grown virus. *J. Virol.* **64**:1808–1811.
18. Kawaoka, Y., S. Yamnikova, T. M. Chambers, D. K. Lvov, and R. G. Webster. 1990. Molecular characterization of a new hemagglutinin, subtype H14, of influenza A virus. *Virology* **179**:759–767.
19. Kida, H., Y. Kawaoka, C. W. Naeve, and R. G. Webster. 1987. Antigenic and genetic concentration of H3 influenza virus in wild ducks. *Virology* **159**:109–119.
20. Lu, B.-L., R. G. Webster, and V. S. Hinshaw. 1982. Failure to detect hemagglutination-inhibiting antibodies with intact avian influenza virions. *Infect. Immun.* **38**:530–535.
21. Luoh, S.-M., M. W. McGregor, and V. S. Hinshaw. 1992. Hemagglutinin mutations related to antigenic variation in H1 swine influenza viruses. *J. Virol.* **66**:1066–1073.
22. Mandel, B. 1958. Studies on the interactions of poliomyelitis virus, antibody and host cells in a tissue culture system. *Virology* **6**:424–447.
23. Nohinek, B., W. Gerhard, and I. T. Schulze. 1985. Characterization of host cell binding variants of influenza virus by monoclonal antibodies. *Virology* **143**:651–656.
24. Pensaert, M., K. Otis, J. Vandeputte, M. M. Kaplan, and P. A. Bachmann. 1981. Evidence for the natural transmission of influenza A virus from wild ducks to swine and its potential importance to man. *Bull. WHO* **59**:75–78.
- 24a. Rajakumar, A., and I. T. Schulze. Unpublished data.
25. Rajakumar, A., E. M. Swierkosz, and I. T. Schulze. 1990. Sequence of an influenza virus hemagglutinin determined directly from a clinical sample. *Proc. Natl. Acad. Sci. USA* **87**:4154–4158.
26. Raymond, F. L., A. J. Canton, N. J. Fox, A. P. Kendal, and G. G. Brownlee. 1986. The antigenicity and evolution of influenza H1 haemagglutinin, from 1950–1957 and 1977–1983: two pathways from one gene. *Virology* **148**:275–287.
27. Robertson, J. S., J. S. Bootman, R. Newman, J. S. Oxford, R. S. Daniels, R. G. Webster, and G. C. Schild. 1987. Structural changes in the haemagglutinin which accompany egg adaptation of an influenza A (H1N1) virus. *Virology* **160**:31–37.
28. Rota, P. A., E. P. Rocha, M. W. Harmon, V. S. Hinshaw, M. G. Sheerar, Y. Kawaoka, N. J. Cox, and T. F. Smith. 1989. Laboratory characterization of a swine influenza virus isolated from a fatal case of human influenza. *J. Clin. Microbiol.* **27**:1413–1416.
29. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed., p. 7.39–7.40. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
30. Schulze, I. T., S. Aytay, A. Rajakumar, A. R. Steward, and I. D. Manger. In W. Doerfler (ed.), Hemagglutinin polymorphism and influenza virus host range. Virus strategies, in press, VCH Verlagsgesellschaft mbH, Weinheim, Germany.
31. Sheerar, M. G., B. C. Easterday, and V. S. Hinshaw. 1989. Antigenic conservation of H1N1 swine influenza viruses. *J. Gen. Virol.* **70**:3297–3303.
32. Skehel, J. J., D. J. Stevens, R. S. Daniels, A. R. Douglas, M. Knossow, I. A. Wilson, and D. C. Wiley. 1984. A carbohydrate side chain on hemagglutinins of Hong Kong influenza viruses inhibits recognition by a monoclonal antibody. *Proc. Natl. Acad. Sci. USA* **81**:1779–1783.
33. Smith, T. F., E. O. Burgert, Jr., W. R. Dowdle, G. R. Noble, R. J. Campbell, and R. E. Van Scoy. 1976. Isolation of swine influenza virus from autopsy lung tissue of man. *N. Engl. J. Med.* **294**:708–710.
34. Underwood, P. A., J. J. Skehel, and D. C. Wiley. 1987. Receptor-binding characteristics of monoclonal antibody-selected antigenic variants of influenza virus. *J. Virol.* **61**:206–208.
35. Webster, R. G., W. J. Bean, O. T. Gorman, T. M. Chambers, and Y. Kawaoka. 1992. Evolution and ecology of influenza A viruses. *Microbiol. Rev.* **56**:152–179.
36. Webster, R. G., M. A. Yakhno, V. S. Hinshaw, W. J. Bean, and K. G. Murti. 1978. Intestinal influenza: replication and characterization of influenza viruses in ducks. *Virology* **81**:268–278.
37. Weis, W., J. H. Brown, S. Cusack, J. C. Paulson, J. J. Skehel, and D. C. Wiley. 1988. Structure of the influenza virus hemagglutinin complexed with its receptor, sialic acid. *Nature (London)* **333**:426–431.
38. Winter, G., S. Fields, and G. G. Brownlee. 1981. Nucleotide sequence of the hemagglutinin gene of a human influenza virus H1 subtype. *Nature (London)* **292**:72–75.
39. Xu, X., E. P. Rocha, H. L. Regenry, A. P. Kendal, and N. J. Cox. 1993. Genetic and antigenic analyses of influenza A (H1N1) viruses, 1986–1991. *Virus Res.* **28**:37–55.
40. Yewdell, J. W., A. J. Caton, and A. Gerhard. 1986. Selection of influenza A virus adsorptive mutants by growth in the presence of a mixture of monoclonal antihemagglutinin antibodies. *J. Virol.* **57**:623–628.