VIROLOGY 182, 475–485 (1991)

# Comparison of Complete Amino Acid Sequences and Receptor-Binding Properties among 13 Serotypes of Hemagglutinins of Influenza A Viruses

E. NOBUSAWA,*[,1,2] T. AOYAMA,† H. KATO,‡ Y. SUZUKI,‡ Y. TATENO,§ AND K. NAKAJIMA†[,2]

*Department of Microbiology, Institute of Public Health, 4-6-1, Shirokanedai, Minato-ku, Tokyo, 108 Japan; †Department of Virology, Institute of Medical Science, University of Tokyo, 4-6-1, Shirokanedai, Minato-ku, Tokyo 108 Japan; ‡The Department of Biochemistry, University of Shizuoka, School of Pharmaceutical Science, Yada, Shizuoka-shi, 422 Japan; and §National Institute of Genetics, Yata, Mishima-shi, 411 Japan

Received August 17, 1990; accepted February 1, 1991

**We determined the sequences of 7 serotypes (H4, H6, H8, H9, H11, H12, and H13) of hemagglutinin (HA) genes, which have not been reported so far. The coding regions consisted of 1692 nucleotides in H4, 1698 in H6, 1695 in H8, 1680 in H9, 1695 in H11, 1692 in H12, and 1698 in H13, and specified 564, 566, 565, 560, 565, 564, and 566 amino acids, respectively. By comparison of amino acid sequences, 13 HA serotypes could be divided into two families, i.e., an H1 group (H1, H2, H5, H6, H8, H9, H11, H12, and H13) and an H3 group (H3, H4, H7, and H10). The relationship was essentially similar to that reported by Air from the comparison of 80 amino-terminal amino acid sequence of 12 HA serotypes (G.M. Air, 1981, *Proc. Natl. Acad. Sci. USA* 78, 7639–7643). Though a considerable amino acid sequence difference exists between certain HA serotypes, several amino acid residues in fusion peptides ($HA_2$1–11) and receptor-binding sites ($HA_1$98, −134, −138, −153, −183, and −195) were shown to be conserved among the 13 HA serotypes. Human H1 and avian H3, H4, H8, and H10 viruses preferentially bound NeuAcα2,3Gal sequences, whereas human H2 and H3 and avian H6 and H9 viruses bound NeuAcα2,6Gal sequences, although the amino acid residues at position 226 of human H2 and avian H6 and H9 serotype HAs are glutamine. These results show that the amino acid residue at position 226 is not necessarily a determinant of receptor specificity for all serotypes.** © 1991 Academic Press, Inc.

## INTRODUCTION

The hemagglutinin (HA) glycoprotein of influenza A virus performs two crucial functions in the early stage of virus infection. First, HA is responsible for binding of the virus to cell surface receptors (Hirst, 1941) and, second, it mediates liberation of the viral genome into the cytoplasm through membrane fusion (Maeda and Ohnishi, 1980; Huang *et al.*, 1981; Lenard and Miller, 1981). Since the three-dimensional structure of H3 HA was determined by X-ray analysis (Wiley *et al.*, 1981; Wilson *et al.*, 1981), functional regions of H3 HA protein, i.e., the receptor-binding site and fusion peptide, have been located on the three-dimensional model. HA is also known to be the major variable antigen (Webster *et al.*, 1983), and influenza A viruses isolated from birds and mammals are subdivided into 13 serotypes by antigenic differences. The primary structures of 6 (H1, H2, H3, H5, H7, and H10) of the 13 serotypes had been determined by 1988 (for references, see Klenk and Rott 1988), and in 1989 those of H4 and H13 were reported (Chambers *et al.*, 1989; Donis *et al.*, 1989). The sequences of the other 5 serotypes (H6, H8, H9, H11, and H12) remained to be determined, even though the sequences of the 3′ end of the viral genomic RNA had been determined (Air, 1981).

The study by Rogers *et al.* (1983) showed that the amino acid residue at position 226 of H3 HA is involved in receptor specificity and sensitivity to neutralization by a glycoprotein inhibitor present in horse serum. Recently, Daniels *et al.* (1987) showed that a deletion mutant ($HA_1$224–230 deleted) retained receptor-binding activity, but lost the ability of discrimination between NeuAcα2,6Gal and NeuAcα2,3Gal linkages. This finding suggested that residue 226 of H3 HA is included in the region which determines linkage specificity but is less directly involved in the receptor binding itself. We showed, however, a significant role of residue 226 in receptor binding for H1 HA (Nobusawa and Nakajima, 1988). A difference in the role of residue 226 between H1 and H3 HA in receptor binding shows a difference in responsibility of amino acid residues located at the same position.

In the present study we determined the nucleotide sequences of genes of 7 serotypes of HA and deduced the amino acid sequences. Amino acid sequences of the receptor-binding sites and its vicinity were compared among 13 serotypes. Receptor specificity was studied for 9 serotypes.

Sequence data from this article have been deposited with the DDBJ, EMBL, and GenBank Data Libraries under Accession Nos. D90302–D90308.

[1] To whom reprint requests should be addressed.

[2] Present address: Department of Virology, Medical School, Nagoya City University, Mizuho-cho, Mizuho-ku, Nagoya-shi, 467 Japan.



0042-6822/91 $3.00
Copyright © 1991 by Academic Press, Inc.
All rights of reproduction in any form reserved.

Petitioner Merck, Ex. 1192, p. 475
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00046

## MATERIALS AND METHODS

### Viruses and virus RNA

The influenza viruses A/Turkey/Ontario/6118/68 (H8N4), A/Duck/Alberta/60/76 (H12N5), and A/Gull/Maryland/704/77 (H13N6) were generously provided by Dr. H. Kida (Hokkaido University). Other viruses used for cDNA cloning included A/Duck/Czechoslovakia/56 (H4N6), A/Shearwater/Australia/1/72 (H6N5), A/Turkey/Wisconsin/66 (H9N2), A/Duck/England/56 (H11N6). A/Duck/Ukraine/1/63 (H3N8) and A/Tern/South Africa/61 (H5N3) were also used for the study of receptor specificity. These viruses were grown in 11-day-old embryonated chicken eggs and were purified by centrifugation through 30–60% (w/v) discontinuous sucrose gradients in a Beckman SW28 rotor. Virion RNA for cloning was extracted by the hot phenol method (Palese and Schulman, 1976).

### Cloning of the HA gene and determination of nucleotide sequences

HA-cDNA was prepared by reverse transcription of virion RNA according to the method of Gubler and Hoffman (Gubler and Hoffman, 1983). Oligonucleotide primer (5′ AGCAAAAGCAGG 3′) complementary to the 3′ terminus of the negative strand RNA was used for cDNA cloning. Double-stranded cDNAs were separated by electrophoresis through agarose gels (0.8% in Tris–EDTA–acetate buffer, pH 8.0) and a cDNA band corresponding to a length of about 1700 bp was eluted. Poly(dC)-tailed cDNA was inserted into poly(dG)-tailed plasmid pBR322 at PstI site. After transfection of *Escherichia coli* C600 with the recombinant DNA, colony hybridization was performed. Several oligonucleotides were synthesized, according to partial sequences of the 3′ end of the negative-strand RNA of HA genes (Air, 1981), and utilized as hybridization probes for screening for a clone of each HA serotype. The oligodeoxyribonucleotides used as probes were 5′ TGGTCCTACATAGTGGA 3′ for H6, H8, and H9; 5′ ATGGTCCTACATAGT 3′ for H11 and H12; and 5′ ATATGCCTGGGACATCA 3′ for H4 and H10. Since not all the virus strains used in this study are the same as those in the study of Air (1981), the synthetic oligonucleotides may not have been completely complementary to cDNA of the respective HA serotype. The conditions for colony hybridization were therefore relaxed initially and filters were washed at increasing temperatures until the probes were dissociated from nonspecific DNA. The colonies showing positive hybridization were purified, and DNAs were isolated from these colonies for subcloning into phage M13 mp18 or mp19. For DNA sequence analysis dideoxy chain termination method was used (Sanger *et al.*, 1977).

### Transfection and generation of recombinant virus stocks

SVHA recombinant DNAs were introduced into subconfluent monolayers of CV-1 cells together with helper virus DNA by DEAE-dextran method (McCutchan and Pagano, 1968) as described previously (Nobusawa *et al.*, 1987). For characterization of the HA expressed from SVHA, subconfluent monolayers of CV-1 cells were infected with virus stock containing SVHA recombinant virus and helper virus dl-1055.

### Mutagenesis

The procedure used for site-specific mutagenesis was described previously (Nobusawa *et al.*, 1985). M13 recombinant phages containing a mutant HA DNA were screened by plaque hybridization. After the presence of desired base substitutions was confirmed by sequencing using dideoxy chain termination method (Sanger *et al.*, 1977), the DNA fragments containing substitutions in the HA cDNA of M13 recombinant phages were exchanged with the corresponding region in the SVHA-2 (Nobusawa *et al.*, 1987).

### Indirect immunofluorescence

Monolayers of CV-1 cells were infected with each recombinant virus stock. At 48 to 60 hr postinfection, infected cells were examined either for cell surface expressed HA protein without fixation or for intracellular HA protein after fixation with ethanol/acetone (1:1). Indirect immunofluorescence staining with rabbit anti-A/USSR/77 HA serum generously provided by Dr. K. Nerome (National Institute of Health, Tokyo, Japan) was performed as described previously (Nobusawa *et al.*, 1985).

### Receptor specificity of influenza virus hemagglutinin for sialyloligosaccharides of gangliosides

The assay system was designed to determine the ability of ganglioside-coated erythrocytes to undergo influenza virus-mediated hemolysis at a low pH at 37° as well as agglutination at pH 7.2 at 4°, as described in previous papers (Suzuki *et al.*, 1986, 1989).

## RESULTS

### Comparison of nucleotides and amino acid sequences among 13 serotype HAs

The nucleotide sequences of the HA genes of H4 (A/Duck/Czechoslovakia/56), H6 (A/Shearwater/Australia/1/72), H8 (A/Turkey/Ontario/6118/68), H9 (A/Turkey/Wisconsin/66), H11 (A/Duck/England/56), H12 (A/Duck/Alberta/60/76), and H13 (A/Gull/Maryland/

Petitioner Merck, Ex. 1192, p. 476
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00046

**TABLE 1**

| Serotype | Amino acid number[a] −15 −10 −1 | +1 |
|---|---|---|
| (a) Comparison of amino acid sequences of signal peptides among the 13 serotype HAs | | |
| H1 | M K A N L L V L L C A L A A A D A | D |
| H2 | M A I I Y L I L L F T A V R G | D |
| H5 | M E E I V L L F A I V S L A R S | D |
| H6 | M I A I I V V A I L A T A G R S | D |
| H11 | M E K T L L F A A I F L C V K A | D |
| H13 | M A L N V I A T L T L I S V C V H A | D |
| H9 | M E T K A I I A A L L M V T A A N A | D |
| H8 | M E K F I A I A T L A S T N A | Y |
| H12 | M E K F I I L S T V L A A S F A | Y |
| H7 | M N T Q I L V F A L V A V I P T N A | D |
| H10 | M Y K V V V I I A L L G A V K G | L |
| H4 | M L S I V I L F L L I A E N S S | Q |
| H3 | M K T I I A L S Y I F C L A L G | Q |

| Serotype | Amino acid number 185 190 200 210 | |
|---|---|---|
| (b) Comparison of amino acid sequences of membrane anchor region among the 13 serotype HAs | | |
| H1 | Q I L A I Y S T V A S S L V L L V S L G A I S F W M C S | (−24.8) |
| H2 | Q I L A I Y A T V A G S L S L A I M M A G I S F W M C S | (−28.7) |
| H5 | Q I L S I Y S T V A S S L A L A I M V A G L S F W M C S | (−27.8) |
| H6 | Q I L A I Y S T V S S S L V L V G L I I A V G L W M C S | (−29.9) |
| H11 | K I L S I Y S C I A S S L V L A A L I M G F M F W A C S | (−28.7) |
| H13 | K A L S I Y S C I A S S V V L V G L I L S F I M W A C S | (−27.1) |
| H9 | K I L T I Y S T V A S S L V L A M G F A A F L F W A M S | (−28.7) |
| H8 | K I L S I Y S T V A A S L C L A I L I A G G L I L G M Q | (−24.2) |
| H12 | K I L S I Y S S V A S S L V L L L M I I G G F I F G C Q | (−26.4) |
| H7 | V I L W F S F G A S C F L L L A I A V − G L V F − I C V | (−36.7) |
| H10 | I I L W F S F G E S C F V L L A V V M − G L V F F − C L | (−34.7) |
| H4 | I I L W I S F S I S C F L L V A L L L A − F I L W A C Q | (−39.8) |
| H3 | W I L W I S F A I S C F L L C − V V L L G F I M W A C Q | (−41.4) |

*Note.* Hydrophobic values are given in parentheses. Dashes are inserted for maximum sequence homology.

[a] The amino acid numbering used for the H3 HA (Wilson *et al.*, 1981) was used. The signal peptidase cleavage site is between the amino acid residues at positions −1 and +1.

704/77) serotype viruses were determined by sequencing cloned DNAs. The coding regions of H4, H6, H8, H9, H11, H12, and H13 serotype HA genes consisted of 1692, 1698, 1695, 1680, 1695, 1692, and 1698 nucleotides and specified 564, 566, 565, 560, 565, 564, and 566 amino acids, respectively. Therefore the complete amino acid sequences are now available for HAs of all 13 serotypes. The nucleotide sequences of H4 and H13 HA genes have recently been determined for the same strains used in this study (Donis *et al.*, 1989; Chambers *et al.*, 1989). Our sequences are different from the above-reported sequences in 14 nucleotides and 4 amino acids for H4, and in 8 nucleotides and 2 amino acids for H13 HA. The discrepancy might be due to different passage histories of the virus strains used.

The amino-terminal signal peptide cleavage site has not been identified directly by amino acid sequencing. By analogy with H1, H2, H5, and H7 (Air, 1981; Kawaoka *et al.*, 1984) we assumed that the amino terminus of HA1 is aspartic acid for H6, H9, H11, and H13 serotype HAs and glutamine for H4 as shown by Donis *et al.* (1989). According to the (−3, −1)-rule of Von Heijne (1986), it was unlikely that the HA1s of H8 and H12 have aspartic acid at its amino terminus as had been suggested (Air, 1981). Using an algorithm which predicts cleavage sites for signal peptidase (Von Heijne, 1986), we presumed that the amino terminus of

Petitioner Merck, Ex. 1192, p. 477
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00046

**a**

```
H1  ----------DTICIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDSHNGKLCRLKGIAPLQLGKCNIAGWLLGNPECDPL
H2  ----------DQICIGYHANNSTEKVDTNLERNVTVTHAKDILEKTHNGKLCKLNGIPPLELGDCSIAGWLLGNPECDRL
H3  QDLPGNDNSTATLCLGHHAVPNGTLVKTITDDQIEVTNATELVQSSSTGKICN-NPHRILDGIDCTLIDALLGDPHCDVF  79
H4  ----QNYTGNPVICMGHHAVANGTMVKTLADDQVEVVTAQELVESQNLPELCPS-PLRLVDGQTCDIINGALGSPGCD-H
H5  ----------DQICIGYHANNSKKQIDTIMEKNVTVTHAQDILEKKHNGKLCSLKGVKPLILKDCSVAGLVLGNPMCDDF
H6  ----------DKICIGYHANNSTTQIDTILEKNVTVTHSVELLENQKEERFCKILKKAPLDLKGCTIEGWILGNPQCDLL
H7  ----------DKICLGHHAVSNGTKVNTLTERGVEVVNATETVERTNIPKICS-KGKRTTDLGQCGLLGTITGPPQCDQF
H8  ---------YDRICIGYQSNNSTDTVNTLIEQNVPVTQTMELVETEKHPAYCNTDLGAPLELRDCKIEAVIYGNPKCDIH
H9  --------- DKICIGYQSTNSTETVDTLTESNVPVTHTKELLHTEHNGMLCATDLGHPLILDTCTIEGLIYGNPSCDIL
H10 ---------LDRICLGHHAVANGTIVKTLTNEQEEVTNATETVESTNLNKLCM-KGRSYKDLGNCHPVGMLIGTPVCDPH
H11 ----------DEICIGYLSNNSTDKVDTIIENNVTVTSSVELVETEHTGSFCSINGKQPISLGDCSFAGWILGNPMCDEL
H12 ---------YDKICIGYQTNNSTETVNTLSEQNVPVTQVEELVHRGIDPILCGTELGSPLVLDDCSLEGLILGNPKCDLY
H13 ----------DRICVGYLSTNSSERVDTLLENGVPVTSSIDLIETNHTGTYCSLNGVSPVHLGDCSFEGWIVGNPACTSN

H1  LPVRSWSYIVETPNSENGICYPGDFIDYEELREQLSSVSSFERFEIFPKESSWPNHNTTKGVTAACSHAGK~SSFYRNLL
H2  LSVPEWSYIMEKENPRDGLCYPGSFNDYEELKHLLSSVKHFEKVKILPK-DRWTQHTTT-GGSRACAVSGNP-SFFRNMV
H3  ~QNETWDLFVERSKA-FSNCYPYDVPDYASLRSLVASSGTLEFITEGF---TWTGVTQN-GGSNACKRGPG~SGFFSRLN 152
H4  LNGAEWDVFIERPNA-VDTCYPFDVPEYQSLRSILANNGKFEFIAEEF---QWNTVKQN-GKSGACKRANV~DDFFNRLN
H5  LNAPEWSYIVEKNNPINGLCYPGDFNDYEELKHLVSSTNLFEKIRIIPR-NSWTNHDASSGVSSACPHLGR~SSFFRNVV
H6  LGDQSWSYIVERPTAQNGICYPGVLNEVEELKALIGSGERVERFEMFPK-STWTGVDTSSGVTRACPYNSG~SSFYRNLL
H7  L-EFSADLIIERREG-NDVCYPGKFVNEEALRQILRGSGGIDKETMGF---TYSGIRTN-GTTSACRRS-G~SSFYAEME
H8  LKDQGWSYIVERPSAPEGMCYPGSVENLEELRFVFSSAASYKRIRLFDY-SRWNVTR-S-GTSKACNASTGGQSFYRSIN
H9  LGGKEWSYIVERSSAVNGMCYPGNVENLEELRSLFSSAKSYKRIQIFPD-KTWNVTY-S-GTSRAC---SN--SFYRSMR
H10 L-TGTWDTLIERENA-IAHCYPGATINEEALRQKIMESGGISKMSTGF---TYGSSITSAGTTKACMRN-GGDSFYAELK
H11 IGKTSWSYIVEKPNPTNGICYPGTLESEEELRLKFSGVLEFNKFEVFT-SNGWGAVNSGVGVTAACKFG-GSNSFFRNMV
H12 LNGREWSYIVERPKEMEGVCYPGSIENQEELRSLFSSIKKYERVKMFDF-TKWNVTYT--GTSKACNNTSNQGSFYRSMR
H13 FGIREWSYLIEDPAAPHGLCYPGELNNNGELRHLFSGIRSFSRTELIPP-TSWGEVL--DGTTSACRDNTGTNSFYRNLV

H1  WLTEKEG--SYPKLKNSYVNKKGKEVLVLWGIHHPSNSKDQQNIYQNENAYVSVVTSNYNRRFTPEIAERPKVRDQAGRM
H2  WLTKEGS--DYPVAKGSYNNTSGEQMLIIWGVHHPIDETEQRTLYQNVGTYVSVGTSTLNKRSTPEIATRPKVNGQGGRM
H3  WLTKSGS--TYPVLNVTMPNNDNFDKLYIWGIHHPSTNQEQTSLYVQASGRVTVSTRRSQQTIIPNIGSRPWVRGLSSRI 230
H4  WLVKSDGNA-YPLQNLTKINNGDYARLYIWGVHHPSTSTEQTNLYKNNPGRVTVSTKTSQTSVVPDIGSRPLVRGQSGRV
H5  WLIKKNN--VYPTIKRTYNNTNVEDLLILWGIHHPNDAAEQAKLYQNLNAYVSVGTSTLNQRSIPKIATRPKVNGQSGRM
H6  WIIKTKSAA-YSVIKGAYNNTGNQPILYFWGVHHPPDTNEQNTLYGSGDRYVRMGTESMNFAKSPEIAARPAVNGQRGRI
H7  WLLSNTDNASFPQMTKSYKNTRRESALIVWGIHHSGSTTEQTKLYGSGNKLITVGSSKYHQSFVPSPGTRPQINGQSGRI
H8  WLTKKEPD~TYDFNEGAYVNNEDGDIIFLWGIHHPPDTKEQTTLYKNANTLSSVTTNTINRSFQPNIGPRPLVRGQQGRM
H9  WLTHKSN--SYPFQNAHYTNNERENILFMWGIHHPPTDTEQTDLYKNADTTTSVTTEDINRTFKPVIGPRPLVNGQQGRI
H10 WLVSKTKGQNFPQTTNTYRNTDTAEHLIIWGIHHPSSTQEKNDLYGTQSLSISVESSTYQNNFVPVVGARPQVNGQSGRI
H11 WLIHQSG--TYPVIKRTFNNTKGRDVLIVWGIHHPATLTEHQDLYKKDSSYVAVGSETYNRRFTPEINTRPRVNGQAGRM
H12 WLTLKSG--QFPVQTDEYKNTRDSDIVFTWAIHHPPTSDEQVKLYKNPDTLSSVTTVEINRSFKPNIGPRPLVRGQQGRM
H13 WFIKKNT--RYPVISKTYNNTTGRDVLVLWGIHHPVSVDETKTLYVNSDPYTLVSTKSWSEKYKLETGVRPGYNGQRSWM

H1  NYYWTLLKPGDTIIFEANGNLIAPRYAFALSRGFGSGIIT-SNASMHECNTKCQTPLGAINSSLPFQNIHPVTIGECPKY
H2  EFSWTLLDMWDTINFESTGNLIAPEYGFKISKRGSSGIMKT-EGTLENCETKCQTPLGAINTTLPFHNVHPLTIGECPKY
H3  SIYWTIVKPGDVLVINSNGNLIAPRGYFKMRTGKSS-IMRSD-APIDTCISECITPNGSIPNDKPFQNVNKITYGACPKY 308
H4  SFYWTIVEPGDLIVFNTIGNLIAPRGHYKLNNQKKSTILNTA-IPIGSCVSKCHTDKGSLSTTKPFQNISRIAVGDCPRY
H5  EFFWTILRPNDTISFESTGNFIAPEYAYKIVKKGDSAIMRS-ELEYGNCDTKCQTPLVAINSSMPFHNVHPLTIGECPKY
H6  DYYWSILKPGETLNVESNGNLIAPWYAFRFVSTSNKGAVFKSNLPIENCDATCQTVAGVLRTNKTFQNVSPLWIGECPKY
H7  DFHWLILDPNDTVTFSFNGAFIAPNRASFL-RGKSMGIQSDVQ-VDANCEGECYHSGGTITSRLPFQNINSRAVGKCPRY
H8  DYYWGILKRGETLKIRTNGNLIAPEFGYLLKGESYGRIIQNEDIPIGNCNTKCQTYAGAINSSKPFQNASRHYMGECPKY
H9  DYYWSVLKPGQTLRIRSNGNLIAPWYGHVLTGESHGRILK-TDLNNGNCVVQCQTEKGGLNTTLPFHNISKYAFGNCPKY
H10 DFHWTLVQPGDNITFSHNGGLIAPSRVSKL-TGRDLGIQSEA-LIDNSCESKCFWRGGSINTKLPFQNLSPRTVGQCPKY
H11 TFYWKIVKPGESITFESNGAFLAPRYAFEIVSVGNGKLFR-SELNIESCSTKCQTEIGGINTNKSFHNVHRNTIGDCPKY
H12 DYYWAVLKPGQTVKIQTNGNLIAPEYGHLITGKSHGRILKN-NLPMGQCVTECQLNEGVMNTSKPFQNTSKHYIGKCPKY
H13 KIYWSLIHPGEMITFESNGGFLAPRYGYIIEEYGKGRIFQS-RIRMSRCNTKCQTSVGGINTNRTFQNIDKNALGDCPKY

H1  VRSAKLRMVTGLRNIPSIQS----R
H2  VKSEKLVLATGLRNVPQIES----R
H3  VKQNTLKLATGMRNVPEKQT----R 329
H4  VKQGSLKLATGMRNIPEKAS----R
H5  VKSDKLVLATGMRNVPQKKK----R
H6  VKSESLRLATGLRNVPQIET----R
H7  VKQESLLLATGMKNVPEPSKKREKR
H8  VKKASLRLAVGLRNTPSVEP----R
H9  VGVKSLKLPVGLRNVPAVSS----R
H10 VNQRSLLLATGMRNVPEVVQG---R
H11 VNVKSLKLATGPRNVPAIAS----R
H12 IPSGSLKLAIGLRNVPQVQD----R
H13 IKSGQLKLATGLRNVPAI-SN---R
```

FIG. 1. Comparison of amino acid sequences of 13 HA serotypes. The sequences are aligned according to Kawaoka *et al.* (1984) and Feldmann *et al.* (1988). Dashes are inserted to give maximam sequence homology. The conserved amino acid residues are in bold face. The amino acid numbering for H3 serotype is shown on the right side. Numbering is restarted for HA2 in (b). H1; A/PR/8/34 (Winter *et al.*, 1981)); H2; A/Japan/305/57 (Gething *et al.*, 1980); H3; A/Aichi/2/68 (Verhoeyen *et al.*, 1980); H4, A/Duck/Czechoslovakia/56; H5; A/Chick/Pennsylvania/1370/83 (Kawaoka *et al.*, 1984); H6; A/Shearwater/Australia/1/72; H7; A/FPV/Rostock/34 (Porter *et al.*, 1979); H8; A/Turkey/Ontario/6118/68; H9; A/Turkey/Wisconsin/66; H10; A/Chicken/Germany/N/49 (Feldmann *et al.*, 1988); H11; A/Duck/England/56; H12; A/Duck/Alberta/60/76; H13; A/Gull/Maryland/704/77.

**b**

```
H1  GLFGAIAGFIEGGWTGMIDGWYGYHHQNEQGSGYAADQKSTQNAINGITNKVNSVIEKMNIQFTAVGKEFNKLEKRMENL
H2  GLFGAIAGFIEGGWQGMVDGWYGYHHSNDQGSGYAADKESTQKAFDGITNKVNSVIEKMNTQFEAVGKEFGNLERRLENL
H3  GLFGAIAGFIENGWEGMIDGWYGFRHQNSEGTGQAADLKSTQAAIDQINGKLNRVIEKTNEKFHQIEKEFSEVEGRIQDL  80
H4  GLFGAIAGFIENGWQGLIDGWYGFRHQNAEGTGTAADLKSTQAAIDQINGKLNRLIEKTNDKYHQIEKEFEQVEGRIQDL
H5  GLFGAIAGFIEGGWQGMVDGWYGYHHINGQGSGYAADKKSTQKAIDGITNKVNSIIDKMNTQFEAVGREFNNLERRIENL
H6  GLFGAIAGFIEGGWTGMIDGWYGYHHENSQGSGYAADRESTQKAVDGITNKVNSIIDKMNTQFEAVDHEFSNLERRIDNL
H7  GLFGAIAGFIENGWEGLVDGWYGFRHQNAQGEGTAADYKSTQSAIDQITGKLNRLIEKTNQQFELIDNEFTEVEKQIGNL
H8  GLFGAIAGFIEGGWSGMIDGWYGFHHSNSEGTGMAADQKSTQEAIDKITNKVNNIVDKMNREFEVVNHEFSEVEKRINMI
H9  GLFGAIAGFIEGGWPGLVAGWYGFQHSNDQGVGMAADKGSTQKAIDKITSKVNNIIDKMNKQYEVIDHEFNELEARLNMI
H10 GLFGAIAGFIENGWEGMVDGWYGFRHQNAQGTGQAADYKSTQAAIDQITGKLNRLIEKTNTEFESIESEFSETEHQIGNV
H11 GLFGAIAGFIEGGWPGLINGWYGFQHRDEEGTGIAADKESTQKAIDQITSKVNNIVDRMNTNFESVQHEFSEIEERINQL
H12 GLFGAIAGFIEGGWPGLVAGWYGFQHQNAEGTGIAADRDSTQRAIDNMQNKLNNVIDKMNKQFEVVNHEFSEVESRINMI
H13 GLFGAIAGFIEGGWPGLINGWYGFQHQNEQGTGIAADKESTQKAIDQITTKINNIIDKMNGNYDSIRGEFNQVEKRINML

H1  NKKVDDGFLDIWTYNAELLVLENERTLDFHDSNVKNLYEKVKSQLKNNAKEIGNGCFEFYHKCDNECMESVRNGTYDYP
H2  NKRMEDGFLDVWTYNAELLVLMENERTLDFHDSNVKNLYDKVRMQLRDNVKELGNGCFEFYHKCDDECMNSVKNGTYDYP
H3  EKYVEDTKIDLWSYNAELLVALENQHTIDLTDSEMNKLFEKTRRQLRENAEEMGNGCFKIYHKCDNACIESIRNGTYDHD 160
H4  ENYVEDTKIDLWSYNAELLVALENQHTIDVTDSEMNKLFERVRRQLRENAEDKGNGCFEIFHKCDNNCIESIRNGTYDHD
H5  NKNLEDGFIDVWTYNAELLVLMENERTLDLHDSNVKNLYDKVRLQLRDNAKEWGNGCFEFYHKCDNECMESVRNGTYNYP
H6  NKRMEDGFLDVWTYNAELLVLLENERTLDLHDANVKNLYERVKSQLRDNAMILGNGCFEFWHKCDDECMESVKNGTYDYP
H7  INWTKDFITEVWSYNAELLVAMENQHTIDLADSEMNKLYERVRKQLRENAEEDGTGCFEIFHKCDDDCMASIRNNTYDHS
H8  NDKIDDQIEDLWAYNAELLVLLENQKTLDEHDSNVKNLFDEVKRRLSANAIDAGNGCFDILHKCDNECMETIKNGTYDHK
H9  NNKIDDQIQDIWAYNAELLVLLENQKTLDEHDANVNNLYNKVKRALGSNAVEDGNGCFELYHKCDDQCMETIRNGTYDRQ
H10 INWTKDSITDIWTYNAELLVAMENQHTIDMADSEMLNLYERVRKQLRQNAEEDGKGCFEIYHTCDDSCMESIRNNTYDHS
H11 SKHVDDSVVDIWSYNAQLLVLLENEKTLDLHDSNVRNLHEKVRRMLKDNAKDEGNGCFTFYHKCDNKCIERVRNGTYDHK
H12 NSKIDDQITDIWAYNAELLVLLENQKTLDEHDANVRNLHDRVRRVLRENAIDTGDGCFEILHKCDNNCMDTIRNGTYNHK
H13 ADRIDDAVTDIWSYNAKLLVLLENDKTLDMHDANVKNLHEQVRRELKDNAIDEGNGCFELLHKCNDSCMETIRNGTYDHT

H1  KYSEESKLNREKVDGVKLESMGI-YQILAIYSTVASSLVLLVSLGAISFWMCS-NGSLQCRICI
H2  KYEEESKLNRNEIKGVKLSSMGV-YQILAIYATVAGSLSLAIMMAGISFWMCS-NGSLQCRICI
H3  VYRDEALNNRFQIKGVELKS-GYKDWILWISFAISCFLLC-VVLLGFIMWACQR-GNIRCNICI 221
H4  IYRDEAINNRFQIQGVKLTQ-GYKDIILWISFSISCFLLVALLLA-FILWACQ-NGNIRCQICI
H5  KYSEESKLKRKEIDGIKWESMG-TYQILSIYSTVASSLALAIMVAGLSFWMCS-NGSLQCRICI
H6  KYQDESKLNRQEIESVKLESLG-VYQILAIYSTVSSSLVLVGLIIAVGLWMCS-NGSMQCRICI
H7  KYREEAMQNRIQIDPVKLSS-GYKDVILWFSFGASCFLLLAIAV-GLVF-ICVKNGNMRCTICI
H8  EYEEEAKLERSKINGVKLEEN-TTYKILSIYSTVAASLCLAILIAGGLILGMQ-NGSCRCMFCI
H9  KYQEESRLERQKIEGVKLESEG-TYKILTIYSTVASSLVLAMGFAAFLFWAMS-NGSCRCNICI
H10 QYREEALLNRLNINPVKLSS-GYKDIILWFSFGESCFVLLAVVM-GLVFF-CLKNGNMRCTICI
H11 EFEEESKINRQEIEGVKLDSSGNVYKILSIYSCIASSLVLAALIMGFMFWACS-NGSCRCTICI
H12 EYEEESKIERQKVNGVKLEENS-TYKILSIYSSVASSLVLLLMIIGGFIFGCQ-NGNVRCTFCI
H13 EYAEESKLKRQEIDGIKLKSEDNVYKALSIYSCIASSVVLVGLILSFIMWACS--GNCRFNVCI
```

FIG. 1—*Continued*

HA1 is at Tyr-16 for H8 and Tyr-17 for H12 HA (Table 1a).

With the above seven HA serotypes, proteolytic cleavage of HA0 to HA1 and HA2 was thought to take place at the carboxyl side of arginine at position 329 (H3 numbering) as with other HA serotypes. There is a single arginine at the cleavage site between HA1 and HA2 (Fig. 1). The HA2 amino terminus starts at glycine residue following arginine. Mature HAs, therefore, consist of the HA1 subunit containing 327 (H6), 326 (H4 and H8), 325 (H11), 324 (H12 and H13), and 319 (H9) amino acids and the HA2 subunit containing 223 (H11 and H13), 222 (H6, H8, H9, and H12), and 221 (H4) amino acids.

The putative membrane anchor regions at the carboxy terminus of HA2 lie between $HA_2$185 and $HA_2$211 (H3 numbering; Table 1b) and the nearly entire stretch is hydrophobic, as are the signal peptides. The hydrophobicity value of each membrane anchor peptides in the 13 serotype HAs is shown in parentheses in Table 1b. These values show that the membrane anchor regions of the H3 group is more hydrophobic than that of the H1 group. The significance of higher hydrophobic value of the H3 group remains unknown. $HA_2$187 leucine, (H3 numbering) is conserved all through the 13 serotypes, while several amino acid residues are conserved within each of the H1 or H3 groups [Ile-186, Ile-189, Tyr/Ser-190, Phe-191, Ser-194, Cys-195, Ser/Phe-196, Leu-199, Cys-210 (H3 numbering)] and among some serotypes within each group [Gln/Lys-185, Phe/Ile-189, Leu/Phe-205 (H3 numbering), etc.] as shown in Table 1b.

There are 4, 8, 10, 7, 7, 9, and 7 potential glycosylation sites (Asn–X–Ser/Thr) on H4, H6, H8, H9, H11, H12, and H13 HAs, respectively. The sites conserved among two or more members of the 13 HA serotypes are the residues Asn-20,21 (H1, H2, H6, H8, and H11), Asn-21 (H9 and H13), or Asn-22 (H3, H4, H7, and H10) of HA1 and the residue 154 of HA2 (H3 numbering).

### Homology of amino acid sequences among the 13 serotype HAs

The deduced amino acid sequences of the coding regions for the 13 serotype HAs are aligned with each

Petitioner Merck, Ex. 1192, p. 479
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00046

other (Fig. 1). Thirteen cysteine residues, 12 of which may compose six disulfide bonds and the 18 glycine and 6 proline residues are conserved among the 13 serotypes. Moreover, fusion peptides (14 amino acids in length) at the amino terminus of HA2 are strictly conserved except for the residue at position 12 (Fig 1b).

For the purpose of phylogenetic analysis of all HA serotypes, the degree of amino acid sequence homology among the 13 serotypes was calculated using the Biosequence Analysis System of the Institute of Medical Science, University of Tokyo, run on a VAX-11/750 computer system. After transformation of the sequence homology into amino acid differences and application of Kimura's (1983) correction method to the differences, phylogenetic analysis was carried out by the modified Farris method (Tateno *et al.,* 1982) and neighbor joining method (Saitou and Nei, 1987). Both methods gave essentially the same phylogenetic tree topology. The result obtained by the modified Farris method is shown in Fig. 2. The phylogenetic tree is similar to those suggested by Air (1981) and Hinshaw (1982), except that the relationships among the H8, H9, and H12 are different between ours and theirs. The 13 serotypes are arranged in Table 2 in the same order as in the phylogenetic tree (Fig. 2). Comparison of the entire amino acid sequence homology of the 13 serotypes confirmed the validity of the previous division of 12 serotypes into two groups from the comparison of 80 amino-terminal amino acid sequence (Air, 1981). In this study, the 13 serotypes are divided into an H1 group comprising H1, H2, H5, H6, H8, H9, H11, H12, and H13 serotypes and an H3 group comprising H3, H4, H7, and H10 serotypes. H13 HA is included in the

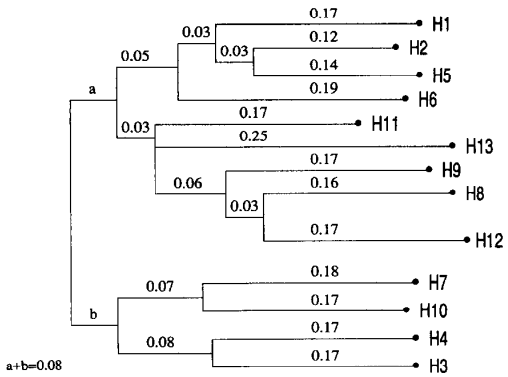


**FIG. 2.** Phylogenetic tree for the 13 serotype HAs. The distance matrix was calculated from amino acid identity values shown in Table 2 by the method of Dayhoff (1972) and the dendrogram was constructed from using the modified Farris method (Tateno *et al.,* 1982). The root of the dendrogram could not be estimated by this method. Branch length represents genetic distance. Viruses in this figure are the same as those in Fig. 1.

H1 group and has 59.3% homology with H11 HA and 52.2% homology with H9 HA, but has less than 50% homology with the other serotype HAs in the H1 group. Amino acid sequence homology within each group ranges from 47.5 to 74.1% (group H1) and 49.6 to 66.5% (group H3) (Table 2). In contrast, similarity ranges from 39.8 to 44.4% between serotypes belonging to different groups. However, the similarity of the HA2 subunit is about 50% even between groups and this shows that the HA2 subunit is more conserved than the HA1 subunit. In group H1, the closest pair is H2 and H5 (74.1%) followed by the pairs of H1–H2 (67.2%), H8–H12 (66.7%), H1–H5 (65.5%), and H9–H12 (65.1%). In group H3, the closest pairs are H3 and H4 (66.5%) and H7 and H10 (65.3%). Other pairs in group H3 have about 50% homology. Comparison of entire amino acid sequences shows that the regions conserved among several closely related serotypes are not localized in a particular region, but the extent of homology is roughly uniform throughout the HA protein.

### Comparison of the receptor-binding site among the 13 serotypes

Examination of amino acid sequences involved in receptor binding in the 13 serotype HAs revealed the following. Several amino acids known to form the receptor-binding cavity and the "second shell" in H3 HA (H3 numbering: Tyr-98, Gly-134, Ala-138, Trp-153, and His-183, Tyr-195) (Weis *et al.,* 1988) are conserved all through the 13 serotypes (Table 3). Other amino acids relatively well conserved are Ser-136 (Thr in H1, H6, H7, H10, H11, and H13), Leu-194 (Ile in H1), Gly-225 (Asp in H1), Gly-228 (Ser in H3 and H13), and Arg-229 (Trp in H13). Additionally, the 7 serotype HAs sequenced in this study possessed Gln-226, which is known to be involved in differential receptor specificity in the H3 serotype.

### Receptor-binding specificity for several HA serotype influenza isolates

Receptor-binding specificity was assayed with gangliosides containing sugar chains with either *N*-acetylneuraminic acid (NeuAc) or *N*-glycolylneuraminic acid (NeuGc) and with either NeuAc$\alpha$2,3Gal or NeuAc$\alpha$2,6Gal linkage for the HA of human and avian origin, and of various serotypes (H1 through H10 with the exception of H7) (Table 4). All viruses bound more strongly to the *N*-acetylneuraminic acid species of sialic acid than to those of the *N*-glycolyl type. It is worth noting that A/PR/8/34 (H1N1), A/Duck/Ukraine/1/63 (H3N8), A/Duck/Czechoslovakia/56 (H4N6), A/Tern/South Africa/61 (H5N3), A/Turkey/Ontario/6118/68(H8N4), and A/Chicken/Germany "N"/49 (H10N7)

Petitioner Merck, Ex. 1192, p. 480
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00046

preferentially bound to NeuAcα2,3Gal-containing oligosaccharides, whereas, A/Japan/305/57(H2N2), A/Aichi/2/68 (H3N2), A/Shearwater/Australia/1/72(H6N5), and A/Turkey/Wisconsin/66(H9N2) bound to NeuAcα2,6Gal-containing oligosaccharides.

### Comparison of the structure of the receptor-binding site and its vicinity

Comparison of amino acid sequences of the receptor-binding site among the 13 serotypes showed that there are several amino acid residues which compose the receptor-binding site but are not conserved. A variant of H3 HA from which residues $HA_1$224–230 consisting the left edge of the receptor-binding site are deleted retained receptor-binding activity (Daniels *et al.*, 1987). To compare the structural requirement of H1 and H3 HAs for receptor-binding activity, we prepared a similar mutant H1 HA protein expressed from the SVHA recombinant virus. A stretch of wild-type H1 (A/USSR/90/77) HA gene in the SVHA-2 (Nobusawa *et al.*, 1988) was replaced by a mutant sequence. In the

**TABLE 2**

AMINO ACID SEQUENCE HOMOLOGY[a] AMONG HA MOLECULE OF VIRUSES REPRESENTING EACH OF THE 13 SEROTYPES

| | H1 | H2 | H5 | H6 | H11 | H13 | H9 | H8 | H12 | H7 | H10 | H4 | H3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| H1 | | 67.2<br>(58.7)<br>[79.6] | 65.5<br>(55.7)<br>[79.8] | 60.8<br>(50.1)<br>[76.4] | 55.5<br>(51.1)<br>[61.9] | 48.6<br>(42.7)<br>[57.2] | 53.0<br>(45.1)<br>[64.5] | 52.6<br>(46.5)<br>[61.5] | 49.8<br>(43.9)<br>[58.4] | 41.6<br>(33.8)<br>[53.0] | 43.5<br>(36.4)<br>[53.9] | 41.8<br>(35.4)<br>[51.1] | 42.6<br>(35.2)<br>[53.4] |
| H2 | | | 74.1<br>(66.7)<br>[84.9] | 62.2<br>(50.3)<br>[79.5] | 55.0<br>(50.6)<br>[61.4] | 47.5<br>(42.9)<br>[54.2] | 53.2<br>(46.5)<br>[63.0] | 49.7<br>(43.0)<br>[59.5] | 48.2<br>(43.6)<br>[54.9] | 43.1<br>(36.1)<br>[53.3] | 43.9<br>(37.7)<br>[53.0] | 40.4<br>(35.2)<br>[48.0] | 41.7<br>(36.0)<br>[50.0] |
| H5 | | | | 61.2<br>(50.0)<br>[77.6] | 56.2<br>(51.2)<br>[63.5] | 49.2<br>(43.4)<br>[57.7] | 51.5<br>(42.6)<br>[64.5] | 50.3<br>(41.9)<br>[62.2] | 48.3<br>(41.3)<br>[58.5] | 44.4<br>(36.8)<br>[55.5] | 43.2<br>(35.8)<br>[54.0] | 41.1<br>(35.0)<br>[50.0] | 43.2<br>(37.3)<br>[51.8] |
| H6 | | | | | 55.5<br>(50.2)<br>[63.2] | 49.6<br>(43.0)<br>[59.2] | 54.7<br>(48.3)<br>[64.0] | 52.8<br>(46.8)<br>[61.6] | 51.0<br>(45.1)<br>[59.8] | 42.4<br>(35.6)<br>[52.3] | 41.9<br>(34.5)<br>[52.7] | 40.8<br>(33.6)<br>[51.3] | 41.5<br>(35.8)<br>[49.8] |
| H11 | | | | | | 59.3<br>(50.9)<br>[71.6] | 53.9<br>(45.8)<br>[65.7] | 50.9<br>(41.8)<br>[64.2] | 51.2<br>(41.4)<br>[65.5] | 41.0<br>(34.5)<br>[50.5] | 43.3<br>(36.7)<br>[52.9] | 42.3<br>(34.0)<br>[54.4] | 43.8<br>(36.3)<br>[54.8] |
| H13 | | | | | | | 52.2<br>(42.6)<br>[66.2] | 49.3<br>(40.8)<br>[61.7] | 50.5<br>(41.1)<br>[64.2] | 39.8<br>(33.0)<br>[49.7] | 40.2<br>(32.8)<br>[51.0] | 40.3<br>(30.9)<br>[54.0] | 39.9<br>(31.5)<br>[52.2] |
| H9 | | | | | | | | 64.7<br>(61.0)<br>[70.1] | 65.1<br>(62.3)<br>[69.2] | 42.5<br>(34.3)<br>[54.5] | 43.3<br>(36.3)<br>[53.5] | 43.1<br>(37.9)<br>[50.7] | 41.7<br>(35.7)<br>[50.5] |
| H8 | | | | | | | | | 66.7<br>(61.0)<br>[75.0] | 40.9<br>(33.6)<br>[51.6] | 42.6<br>(36.3)<br>[51.8] | 43.5<br>(37.2)<br>[52.7] | 42.6<br>(35.2)<br>[53.4] |
| H12 | | | | | | | | | | 43.9<br>(36.3)<br>[55.0] | 43.2<br>(35.2)<br>[54.9] | 43.6<br>(36.6)<br>[53.8] | 42.7<br>(35.3)<br>[53.5] |
| H7 | | | | | | | | | | | 65.3<br>(53.8)<br>[82.1] | 49.6<br>(36.5)<br>[68.7] | 49.6<br>(38.0)<br>[66.5] |
| H10 | | | | | | | | | | | | 51.7<br>(42.1)<br>[65.7] | 52.2<br>(42.0)<br>[67.1] |
| H4 | | | | | | | | | | | | | 66.5<br>(53.8)<br>[85.0] |
| H3 | | | | | | | | | | | | | |

*Note.* H1:A/PR/8/34; H2:A/Jap/305/57; H3:A/Aichi/2/68;H4:A/Duck/Czechoslovakia./56; H5:A/Chick/Penn/1370/83; H6:A/Shearwater/Australia/1/72; H7:A/FPV/34; H8:A/Turkey/Ontario/6118/68; H9:A/Turkey/Wisconsin/66; H10:A/Chick/Germany "N"/49; H11:A/Duck/Eng/56; H12:A/Duck/Alberta/60/76; H13:A/Gull/Maryland/704/77.

[a] Figures not parenthesized are percent homology between whole HA molecule including the connecting peptide or arginine residue between the $HA_1$ and $HA_2$ subunits, but without a signal peptide. Percentages of amino acid identity among HA1 and HA2 subunits are shown in the upper or lower parentheses, respectively. Percent identity was calculated, deletions or insertions being counted as mismatches.

Petitioner Merck, Ex. 1192, p. 481
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00046

TABLE 3

COMPARISON OF AMINO ACID SEQUENCES AROUND THE RECEPTOR-BINDING SITE AMONG THE 13 SEROTYPE HAs

| | Amino acid number[a] | | | | | | |
|---|---|---|---|---|---|---|---|
| Serotype | 98 | 153 | 155 | 183 | 190 | 194 | 195 |
| | (a) Receptor-binding site | | | | | | |
| H1 | Y | W | T | H | D | I | Y |
| H2 | Y | W | T | H | E | L | Y |
| H5 | Y | W | I | H | E | L | Y |
| H6 | Y | W | I | H | E | L | Y |
| H11 | Y | W | I | H | E | L | Y |
| H13 | Y | W | I | H | E | L | Y |
| H9 | Y | W | T | H | E | L | Y |
| H8 | Y | W | T | H | E | L | Y |
| H12 | Y | W | T | H | E | L | Y |
| H7 | Y | W | L | H | E | L | Y |
| H10 | Y | W | V | H | E | L | Y |
| H4 | Y | W | V | H | E | L | Y |
| H3 | Y | W | T | H | E | L | Y |

| | Amino acid number[a] | | | | | |
|---|---|---|---|---|---|---|
| Serotype | 224 | 225 | 226 | 227 | 228 | 229 |
| | (b) Left edge of receptor-binding pocket | | | | | |
| H1 | R | D | Q | A | G | R |
| H2 | N | G | Q | G | G | R |
| H5 | N | G | Q | S | G | R |
| H6 | N | G | Q | R | G | R |
| H11 | N | G | Q | A | G | R |
| H13 | N | G | Q | R | S | W |
| H9 | N | G | Q | Q | G | R |
| H8 | R | G | Q | Q | G | R |
| H12 | R | G | Q | Q | G | R |
| H7 | N | G | Q | S | G | R |
| H10 | N | G | Q | S | G | R |
| H4 | R | G | Q | S | G | R |
| H3 | R | G | L | S | S | R |

| | Amino acid number[a] | | | | |
|---|---|---|---|---|---|
| Serotype | 134 | 135 | 136 | 137 | 138 |
| | (c) Right edge of receptor-binding site | | | | |
| H1 | G | V | T | A | A |
| H2 | G | G | S | R | A |
| H5 | G | V | S | S | A |
| H6 | G | V | T | R | A |
| H11 | G | V | T | A | A |
| H13 | G | T | T | S | A |
| H9 | G | T | S | R | A |
| H8 | G | T | S | K | A |
| H12 | G | T | S | K | A |
| H7 | G | T | T | S | A |
| H10 | G | T | T | K | A |
| H4 | G | K | S | G | A |
| H3 | G | G | S | N | A |

[a] The amino acid numbering used for the H3 HA (Wilson *et al.*, 1981) was used.

mutant seven residues are deleted: 224–230(SVHA-del-2). This mutant HA protein was expressed in CV-1 cells from the recombinant virus vector SVHA-del-2. Expression of the HA protein on the infected cell surface was determined by indirect immunofluorescence staining using rabbit anti-A/USSR/77 serum. While the mutant HA of SVHA-del-2 was demonstrated on the infected cell surface by immunofluorescence staining, hemadsorption and fusion activity could not be detected (data not shown). These findings suggest that the role of the left edge of the receptor-binding site in H1 HA is different from that in the H3 HA.

## DISCUSSION

We determined the nucleotide sequences of H4, H6, H8, H9, H11, H12, and H13 serotype HAs of influenza A viruses in the present study. Amino acid sequences are now available for all 13 serotypes. Alignment of amino acid sequences was carried out with conserved cysteine residues. It showed that the proportion of conserved amino acid residues all through the 13 serotype HAs is 12% (42/346) for HA1 and 30% (67/225) for HA2. All conserved amino acid residues including 13 cysteines, 18 glycines, and 6 prolines seem to be important in maintaining the basic structure of the HA protein.

Several amino acid residues forming the receptor-binding cavity and the second shell in H3 HA are conserved all through the 13 serotype HAs. Basic structure of the receptor-binding site is, therefore, similar among the 13 serotype HAs. While crucial amino acids for receptor binding are conserved, only Gly-134 is conserved on the right edge (134–138) of the receptor-binding site all through the 13 serotype HAs. Ala-138 is conserved in the strains studied here, but not in other strains (A/USSR/77, A/WSN/34). Several amino acid substitutions were found in the right edge of the pocket, but the atom which is presumed as interacting with sialic acid in the H3 HA (Weis *et al.*, 1988) is conserved in either the main chain or the side chains of a substituting amino acid residue at the same position in the other serotype HA as in the H3 serotype. For instance, the hydroxyl group in the side chain of the residue Ser-136 presumed as interacting with sialic acid in H3 is conserved on the side chain of the substitute residue threonine in other serotypes (H6, H7, H10, H11, and H13). On the other hand, the significance of the left edge for the receptor-binding activity might be different from that of the right edge, since not all the amino acid residues on the left edge in H3 HA seem to interact with sialic acid (Weis *et al.*, 1988). Moreover, the left edge appears to be one of the determinants of receptor specificity (Rogers *et al.*, 1983), but does not appear to be pivotal for receptor binding in H3 serotype

Petitioner Merck, Ex. 1192, p. 482
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00046

HA, because the residue 224–230 constituting the left edge is deleted in a variant of H3N2 virus (Daniels *et al.*, 1987). By contrast, H1 serotype HA expressed from SVHA-del-2 with the same amino acid deletion lost its receptor-binding activity. The left edge, therefore, seems to be pivotal for receptor binding for H1 HA. The significance of the left edge of the pocket for receptor-binding activity might be different between the two HA serotypes.

The receptor specificity of the HA has been considered to be one of the determinants of the tissue tropism and host range in influenza virus. It was reported that all human isolates of the H3 serotype agglutinated the erythrocytes containing the NeuAcα2,6Gal linkage but not those bearing the NeuAcα2,3Gal linkage, whereas antigenically similar isolates from avian and equine species preferentially bound the erythrocytes containing the NeuAcα2,3Gal linkage (Rogers and Paulson, 1983). The amino acid residue at position 226 in H3 serotype HA is known to be a determinant of receptor specificity; i.e., H3 HA that recognizes NeuAcα2,6Gal linkage has leucine, while that which is NeuAcα2,3Gal-specific has glutamine at position 226 (Rogers *et al.*, 1983). We investigated then the relationship between the amino acid residues at position 226 and receptor specificity with H1, H2, H3, H4, H6, H8, H9, and H10 HAs. In contrast to H3 HA, H2, H6, and H9, HAs having glutamine at position 226 preferentially recognized NeuAcα2,6Gal linkage (Table 4). These findings suggested that the amino acid residue at position 226 is not necessarily the determinant of receptor specificity. Our previous paper showed that H1 serotype HA with leucine at position 226 expressed from SVHA-330 has no hemadsorption activity (Nobusawa *et al.*, 1988). Suzuki *et al.* (1989) showed recently that a single amino acid substitution of (Ser-205 to Thr) altered the receptor specificity of H3 HA. In this study, however, the amino acid residues at position 205 were valine (H1), glycine (H2, H5, H6, and H7), serine (H3, H4), threonine (H8, H9), and glutamic acid (H10), and no correlation could be found between the amino acid residue at position 205 and receptor specificity.

In this study, the receptor specificity with H1, H2, H3, H4, H5, H6, H8, H9, and H10 HAs was examined by hemolysis assay with gangliosides containing either NeuAcα2,3Gal or NeuAcα2,6Gal. In the previous paper, receptor specificity was examined by adsorption to derivatized human erythrocytes containing either NeuAcα2,6Gal or NeuAcα2,3Gal linkage. The amount of virus adsorbed was quantitated by the neuraminidase activity (Paulson and Rogers, 1986). In this study, on the other hand, ganglioside–erythrocytes were prepared by incorporating gangliosides into desialylated chicken erythrocytes and the adsorbed virus was measured by virus-mediated hemolysis. By these two methods, the degree of receptor specificity examined for certain strains was not always the same, though similar preference in the linkage specificity was observed. In the case of A/Aichi/2/68, Rogers and Paulson (1983) observed no virus adsorbed to derivatized erythrocytes containing NeuAcα2,3Gal, but we detected a small amount of virus bound to NeuAcα2,3Gal-containing erythrocyte. This discrepancy may due to the different amount of virus used for the assay; i.e., Suzuki *et al.* (1986) used $2^7$–$2^{10}$ HA units, in contrast to $2^2$ HA units used by Rogers and Paulson (1983). The differences in species of erythrocyte and in the method of reconstructing erythrocytes may also cause a difference in the amount of virus bound to reconstructed erythrocytes.

Subdivision of serotypes of influenza A viruses was based on the antigenic relationships in double immuno-

**TABLE 4**

RECEPTOR SPECIFICITY OF INFLUENZA VIRUSES (HUMAN AND AVIAN ISOLATES)

| Chicken erythrocytes desialylated and then coated with | Receptor specificity[a] | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Human isolates | | | Avian isolates | | | | | | |
| | H1N1 | H2N2 | H3N2 | H3N8 | H4N6 | H5N3 | H6N5 | H8N4 | H9N2 | H10N7 |
| Sialylparagloboside | | | | | | | | | | |
| NeuAcα2,3Gal | 63 ± 5 | 51 ± 4 | 49 ± 6 | 78 ± 7 | 78 ± 10 | 77 ± 8 | 52 ± 7 | 61 ± 5 | 41 ± 3 | 82 ± 10 |
| NeuAcα2,6Gal | 29 ± 4 | 83 ± 9 | 85 ± 7 | 53 ± 7 | 65 ± 6 | 62 ± 5 | 79 ± 4 | 28 ± 3 | 61 ± 5 | 51 ± 6 |
| NeuGcα2,3Gal | 0 | 0 | 3 ± 1 | 4 ± 1 | 5 ± 2 | 5 ± 1 | 6 ± 2 | 2 ± 1 | 5 ± 1 | 4 ± 1 |

*Note.* H1, A/PR/8/34; H2, A/Jap/305/57; H3, A/Aichi/2/68, A/Duck/Ukraine/63; H4, A/Duck/Czechoslovakia/56; H5, A/Tern/South Africa/61; H6, A/Shearwater/Australia/1/72; H8, A/Turkey/Ontario/6118/68; H9, A/Turkey/Wisconsin/66; H10, A/Chick/Germany "N"/49.

[a] The receptor specificity was expressed by virus-mediated hemolysis at low-pH of desialylated and ganglioside-coated chicken erythrocytes. The recovery of hemolysis by ganglioside coated erythrocytes were calculated by (absorbance at 540 nm of the supernatant of virus-mediated hemolysis of ganglioside coated erythrocytes—absorbance at 540 nm of the supernatant of virus-mediated hemolysis of desialylated erythrocytes)/pmol ganglioside attached to $10^6$ erythrocytes.

Petitioner Merck, Ex. 1192, p. 483
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00046

diffusion tests (WHO Memorandum, 1980). In contrast to the absense of antigenic relationship among 13 HA serotypes by double immunodiffusion tests, certain degrees of amino acid homology were found to exist. A homology as high as 74% (66.7% in HA1) was found between H2 and H5 serotypes. After submission of this manuscript Kawaoka *et al.* (1990) reported the nucleotide sequence of a new serotype, H14. They showed that H14 is closely related to H4 in amino acid sequence homology (68.5% for HA1, 92.3% for HA2). This is even higher than the homology found between the thus far most closely related H2 and H5 serotypes. They discussed that near identity in HA2 between H4 and H14 argues for very recent divergence of these two serotypes. On the other hand, the amino acid sequence homology within a serotype is 80% or more even between virus strains readily distinguishable by polyclonal antibodies (Nestorowicz *et al.,* 1987; Kawaoka *et al.,* 1989; Donis *et al.,* 1989). As substitution of a single amino acid out of five to six residues which usually make up an antigenic determinant may profoundly alter the antigenicity, antigenic change is not predicted from amino acid sequence change. Amino acid sequence homology of 68.5% or less for the HA1 subunit may serve, however, as a rough indication for different serotypes.

## ACKNOWLEDGMENTS

We are greatly indebted to Dr. Akira Sugiura for helpful suggestions and valuable discussion. We also thank Dr. Nobuyuki Miyajima for the helpful suggestion about computer analysis and Dr. Tadayuki Takeda who kindly synthesized the oligonucleotides. This work was partially supported by a scientific research grant from the Ministry of Education.

## REFERENCES

AIR, G. M. (1981). Sequence relationships among the hemagglutinin genes of 12 subtypes of influenza A virus. *Proc. Natl. Acad. Sci. USA* **78,** 7639–7643.

CHAMBERS, T. M., YAMNIKOVA, S., KAWAOKA, Y., LVOV, D. K., and WEBSTER, R. G. (1989). Antigenic and molecular characterization of subtype H13 hemagglutinin of influenza virus. *Virology* **172,** 180–188.

DANIELS, R. S., JEFFRIES, S. YATES, P., SCHILD, G. C., ROGERS, G. N., PAULSON, J. C., WHARTON, S. A., DOUGLAS, A. R., SKEHEL, J. J., and WILEY, D. C. (1987). The receptor-binding and membrane-fusion properties of influenza virus variants selected using anti-haemagglutinin monoclonal antibodies. *EMBO J.* **6,** 1459–1465.

DAYHOFF, M. O. (1972). "Atlas of Protein Sequence and Structure," Vol. 5. Natl. Biomed. Res. Found., Silver Spring, MD.

DONIS, R. O., BEAN, W. J., KAWAOKA, Y., and WEBSTER, R. G. (1989). Distinct lineages of influenza virus H4 hemagglutinin genes in different regions of the world. *Virology* **169,** 408–417.

FELDMANN, H., KRETZSCHMAR, E., KLINGEBORN, B., ROTT, R., KLENK, H. D., and GARTEN, W. (1988). The structure of serotype H10 hemagglutinin of influenza A virus: Comparison of an apathogenic avian and a mammalian strain pathogenic for mink. *Virology* **165,** 428–437.

GETHING, M. J., BYE, J., SKEHEL, J., and 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.

GUBLER, U., and HOFFMAN, B. J. (1983). A simple and very efficient method for generating cDNA libraries. *Gene* **25,** 263–269.

HINSHAW, V. S., AIR, G. M., GIBBS, A. J., GRAVES, L., PRESCOTT, B., and KARUNAKARAN, D. (1982). Antigenic and genetic characterization of a novel hemagglutinin subtype of influenza A viruses from gulls. *J. Virol.* **42,** 865–872.

HIRST, G. K. (1941). The agglutination of red cells by allantoic fluid of chick embryos infected with influenza virus. *Science* **94,** 22–23.

HUANG, R. T. C., ROTT, R., and KLENK, H.-D. (1981). Influenza viruses cause hemolysis and fusion of cells. *Virology* **110,** 243–247.

KAWAOKA, Y., NAEVE, C. W., and WEBSTER, R. G. (1984). Is virulence of H5N2 influenza viruses in chickens associated with loss of carbohydrate from the hemagglutinin? *Virology* **139,** 303–316.

KAWAOKA, Y., BEAN, W. J., and WEBSTER, R. G. (1989). Evolution of the hemagglutinin of equine H3 influenza viruses. *Virology* **169,** 283–292.

KAWAOKA, Y., YAMUNICOVA, S. CHAMBERS, T. M., LVOV, D. K., and WEBSTER, R. G. (1990). Molecular characterization of a new hemagglutinin, subtype H14, of influenza A virus. *Virology* **179,** 759–767.

KIMURA, M. (1983) "The Neutral Theory of Molecular Evolution." Cambridge Univ. Press, Cambridge.

KLENK, H. D., and ROTT, R. (1988). The molecular biology of influenza virus pathogenicity. *Adv. Virus Res.* **34,** 247–281.

LENARD, J., and MILLER, D. K. (1981) pH-dependent hemolysis by influenza, Semliki-Forest virus, and Sendai virus. *Virology* **110,** 479–482.

MCCUTCHAN, J. H., and PAGANO, J. S. (1968). Enhancement of the infectivity of simian virus 40 deoxyribonucleic acid with diethylaminoethyl-dextran. *J. Natl. Cancer Inst.* **41,** 351–357.

MAEDA, T., and OHNISHI, S. (1980). Activation of influenza virus by acidic media causes hemolysis and fusion of erythrocytes. *FEBS Lett.* **122,** 283–287.

NESTOROWICZ, A., KAWAOKA, Y., BEAN, W. J., and WEBSTER, R. G. (1987). Molecular analysis of the hemagglutinin genes of Australian H7N7 influenza viruses: Role of passerine birds in maintenance or transmission? *Virology* **160,** 411–418.

NOBUSAWA, E., NAKAJIMA, K., and NAKAJIMA, S. (1985). Expression of haemagglutinin gene. *Vaccine* **3,** 172–174.

NOBUSAWA, E., NAKAJIMA, K., and NAKAJIMA, S. (1987). Determination of the epitope 264 on the hemagglutinin molecule of influenza H1N1 virus by site-specific mutagenesis. *Virology* **159,** 10–19.

NOBUSAWA, E. and NAKAJIMA, K. (1988). Amino acid substitution at 226 of the hemagglutinin molecule of influenza (H1N1) virus affects receptor binding activity but not fusion activity. *Virology* **167,** 8–14.

PALESE, P., and SCHULMAN, J. L. (1976). Differences in RNA patterns of influenza A virus. *J. Virol.* **17,** 876–884.

PAULSON, J. C., and ROGERS, G. N. (1986). Resialylated erythrocytes for assessment of the specificity of sialyloligosaccharide binding proteins. *In* "Methods in Enzymology" (V. Ginsburg, Ed.), Vol. 138, pp. 162–168. Academic Press, San Diego.

PORTER, A. G., BARBER, C., CAREY, N. H., HALLEWELL, R. A., THRELFALL, G., and EMTAGE, J. S. (1979). Complete nucleotide sequence of an influenza virus haemagglutinin gene from cloned DNA. *Nature (London)* **282,** 471–477.

ROGERS, G. N., and PAULSON, J. C. (1983). Receptor determinants of human and animal influenza virus isolates: Differences in receptor

specificity of the H3 hemagglutinin based on species of origin. *Virology* **127,** 361–373.

ROGERS, G. N., PAULSON, J. C., DANIELS, R. S., SKEHEL, J. J., WILSON, I. A., and WILEY, D. C. (1983). Single amino acid substitution in influenza haemagglutinin change receptor binding specificity. *Nature (London)* **304,** 76–78.

SAITOU, N., and NEI, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4,** 406–425.

SANGER, F., NICKLEN, S., and COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.

SUZUKI, Y., NAGAO, Y., MATSUMOTO, M., NEROME, K., NAKAJIMA, K., and NOBUSAWA, E. (1986). Human influenza A virus hemagglutinin distinguishes sialyloligosaccharides in membrane-associated gangliosides as its receptor which mediates the adsorption and fusion processes of virus infection. *J. Biol. Chem.* **261,** 17,057–17,061.

SUZUKI, Y., KATO, H., NAEVE, C. W., and WEBSTER, R. G. (1989). Single-amino acid substitution in an antigenic site of influenza virus hemagglutinin can alter the specificity of binding to cell membrane-associated gangliosides. *J. Virol.* **63,** 4298–4302.

TATENO, Y., NEI, M., and TAJIMA, F. (1982). Accuracy of estimated phylogenetic tree from molecular data. *J. Mol. Evol.* **18,** 387–404.

VERHOEYEN, M., FANG, R., MIN JOU, W., DEVOS, R., HUYLEBROECK, D., SAMAN, E., and FIERS, W. (1980). Antigenic drift between the haemagglutinin of the Hong Kong influenza strains A/Aichi/2/68 and A/Victoria/3/75. *Nature (London)* **286,** 771–776.

VON-HEIJNE, G. (1986). A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.* **14,** 4683–4690.

WHO Memorandum (1980). A revision of the system of nomenclature viruses. *Bull. WHO* **58,** 585–591.

WEBSTER, R. G., LAVER, W. G., and AIR, G. M. (1983). Antigenic variation among type A influenza viruses. *In* "Genetics of Influenza Viruses" (P. Palese and W. Kingsbury, Eds.), pp. 127–168. Springer Verlag, Wien.

WEIS, W., BROWN, J. H., CUSACK, S., PAULSON, J. C., SKEHEL, J. J., and WILEY, D. C. (1988). Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid. *Nature (London)* **333,** 426–431.

WILEY, D. C., WILSON, I. A., and SKEHEL, J. J. (1981). Structure identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. *Nature (London)* **289,** 373–378.

WILSON, I. A., SKEHEL, J. J., and WILEY, D. C. (1981). Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3Å resolution. *Nature (London)* **289,** 366–373.

WINTER, G., FIELDS, S., and BROWNLEE, G. G. (1981). Nucleotide sequence of the haemagglutinin of a human influenza virus H1 subtype. *Nature (London)* **292,** 72–75.

Petitioner Merck, Ex. 1192, p. 485
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00046