

7. S. W. Squyres *et al.*, *Science* **306**, 1709 (2004).
8. J.-P. Bibring *et al.*, *Eur. Space Agency Spec. Publ.* **1240**, 37 (2004).
9. A. Chicarro, P. Martin, R. Trautner, *Eur. Space Agency Spec. Publ.* **1240**, 3 (2004).
10. J.-P. Bibring *et al.*, *Science* **307**, 1576 (2005).
11. J. Mustard *et al.*, *Science* **307**, 1594 (2005).
12. J. L. Bandfield, *J. Geophys. Res.* **107**, 5092 (2002).
13. C. M. Pieters, in *Remote Geochemical Analysis: Elemental and Mineralogical Composition*, C. M. Pieters, P. A. J. Englert, Eds. (Cambridge Univ. Press, New York, 1993), pp. 309–340.
14. P. Pinet, S. Chevrel, *J. Geophys. Res.* **95**, 14435 (1990).
15. J. F. Bell III *et al.*, *J. Geophys. Res.* **105**, 1721 (2000).
16. P. Bertelsen *et al.*, *Science* **305**, 827 (2004).
17. F. Poulet *et al.*, *Nature* **438**, 623 (2005).
18. A. Gendrin *et al.*, *Science* **307**, 1587 (2005).
19. Y. Langevin, F. Poulet, J.-P. Bibring, B. Gondet, *Science* **307**, 1584 (2005).
20. R. Arvidson *et al.*, *Science* **307**, 1591 (2005).
21. J. L. Bandfield, T. D. Glotch, P. R. Christensen, *Science* **301**, 1084 (2003).
22. J. C. Bridges *et al.*, *Space Sci. Rev.* **96**, 365 (2001).
23. L. L. Griffith, E. L. Shock, *J. Geophys. Res.* **102**, 9135 (1997).
24. H. E. Newson, *Icarus* **44**, 207 (1980).
25. X. Phillips *et al.*, *Science* **291**, 2587 (2001).
26. S. C. Solomon *et al.*, *Science* **307**, 1214 (2005).
27. R. Kahn, *Icarus* **62**, 175 (1985).
28. J.-P. Bibring *et al.*, *Nature* **428**, 627 (2004).
29. T. Encrenaz *et al.*, *Icarus* **170**, 424 (2004).
30. R. V. Morris *et al.*, *Science* **305**, 833 (2004).
31. See plates in (35).
32. E. Chassefière, F. Leblanc, *Planet. Space Sci.* **52**, 1039 (2004).
33. B. M. Jakovsky, R. J. Phillips, *Nature* **412**, 237 (2001).
34. J. D. Bernal, *The Physical Basis of Life* (Routledge & Kegan Paul, London, 1951).
35. J. F. Bell III *et al.*, *Science* **305**, 800 (2004).
36. The OMEGA instrument was developed with the support of the Centre National d'Etudes Spatiales (CNES), Agenzia Spaziale Italiana (ASI), and Russian Space Agency. The scientific activity is funded by national space and research agencies and universities in France, Italy, Russia, Germany, and the United States. We are very grateful to all of the European Space Agency (ESA) teams who, together with industry, enable this mission. Public OMEGA data are accessible on the ESA Planetary Science Archive Web site. This paper benefited from very valuable comments from S. Murchie.

15 November 2005; accepted 20 March 2006  
10.1126/science.1122659

# Structure and Receptor Specificity of the Hemagglutinin from an H5N1 Influenza Virus

James Stevens,<sup>1\*</sup> Ola Blixt,<sup>1,2</sup> Terrence M. Tumpey,<sup>4</sup> Jeffery K. Taubenberger,<sup>5</sup> James C. Paulson,<sup>1,2</sup> Ian A. Wilson<sup>1,3\*</sup>

The hemagglutinin (HA) structure at 2.9 angstrom resolution, from a highly pathogenic Vietnamese H5N1 influenza virus, is more related to the 1918 and other human H1 HAs than to a 1997 duck H5 HA. Glycan microarray analysis of this Viet04 HA reveals an avian  $\alpha$ 2-3 sialic acid receptor binding preference. Introduction of mutations that can convert H1 serotype HAs to human  $\alpha$ 2-6 receptor specificity only enhanced or reduced affinity for avian-type receptors. However, mutations that can convert avian H2 and H3 HAs to human receptor specificity, when inserted onto the Viet04 H5 HA framework, permitted binding to a natural human  $\alpha$ 2-6 glycan, which suggests a path for this H5N1 virus to gain a foothold in the human population.

The H5N1 avian influenza virus, commonly called “bird flu,” is a highly contagious and deadly pathogen in poultry. Since late 2003, H5N1 has reached epizootic levels in domestic fowl in a number of Asian countries, including China, Vietnam, Thailand, Korea, Indonesia, Japan, and Cambodia, and has now spread to wild bird populations. More recently, the H5N1 virus has spread to infect bird populations across much of Europe and into Africa. However, its spread to the human population has so far been limited, with only 191 documented severe infections, but with a high mortality accounting for 108 deaths in Indonesia, Vietnam, Thailand, Cambodia, China, Iraq, Turkey, Azerbaijan, and Egypt [as of 4 April 2006, see the World Health Organization Web site (1)]. Of these, evidence suggests direct bird-to-human transmission, although indirect transmission, perhaps

through contaminated water supplies, cannot be ruled out.

Of the three influenza pandemics of the last century, the 1957 (H2N2) and 1968 (H3N2) pandemic viruses were avian-human reassortments in which three and two of the eight avian gene segments, respectively, were reassorted into an already circulating, human-adapted virus (2, 3). The origin of the genes of the 1918 influenza virus (H1N1), which killed about 50 million people worldwide (4), is unknown. The extinct pandemic virus from 1918 has recently been reconstructed in the laboratory and was found to be highly virulent in mice and chicken embryos (5, 6). With continued outbreaks of the H5N1 virus in poultry and wild birds, further human cases are likely, and the potential for the emergence of a human-adapted H5 virus, either by reassortment or mutation, is a threat to public health worldwide.

Hemagglutinin (HA), the principal antigen on the viral surface, is the primary target for neutralizing antibodies and is responsible for viral binding to host receptors, enabling entry into the host cell through endocytosis and subsequent membrane fusion. As such, the HA is an important target for both drug and vaccine development. Although 16 avian and mammalian serotypes of HA are known, only three (H1, H2, and H3) have become adapted to the

human population. HA is a homotrimer; each monomer is synthesized as a single polypeptide (HA0) that is cleaved by host proteases into two subunits (HA1 and HA2). HA binds to receptors containing glycans with terminal sialic acids, where their precise linkage determines species preference. A switch in receptor specificity from sialic acids connected to galactose in  $\alpha$ 2-3 linkages (avian) to  $\alpha$ 2-6 linkages (human) is a major obstacle for influenza A viruses to cross the species barrier and to adapt to a new host (7, 8). On H3 and H1 HA frameworks, as few as two amino acid mutations can switch human and avian receptor specificity.

Of the H5N1 viral isolates studied to date, A/Vietnam/1203/2004 (Viet04) is among the most pathogenic in mammalian models, such as ferrets and mice (9, 10). This virus was originally isolated from a 10-year-old Vietnamese boy who died from bird flu. Because of the importance of HA in viral pathogenesis and host response to viral infection, we cloned and expressed the ectodomain (HA0) of its HA gene (fig. S1) in a baculovirus expression system, using the same strategy that led to the crystal structure of the 1918 influenza virus HA0 (11, 12). Viet04 HA0 was cleaved during protein production into its activated form (HA1/HA2) and was crystallized at pH 6.55 (13). Its structure was determined by molecular replacement (MR) to 2.95 Å resolution (table S1) (14). In addition, we have investigated the potential of this H5 HA to acquire human receptor specificity by introducing mutations known to effect such a specificity switch on H1 and H3 frameworks.

**Structural overview.** The overall fold of the Viet04 HA trimer (Fig. 1, A and B) is very similar to other published HAs, as expected, with a globular head containing the receptor binding domain (RBD) and vestigial esterase domain, and a membrane proximal domain with its distinctive, central  $\alpha$ -helical stalk and HA1/HA2 cleavage site (essential for viral pathogenicity). Although Viet04 HA and the only other avian H5 HA structure, Sing97 [A/Duck/Singapore/3/1997; Protein Data Bank (PDB) entry 1j5m (15)], are closely related in sequence (HA1, 90%; HA2, 98%), the best molecular replacement (MR) solutions were sur-

<sup>1</sup>Department of Molecular Biology, <sup>2</sup>Glycan Array Synthesis Core-D, Consortium for Functional Glycomics, <sup>3</sup>Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA. <sup>4</sup>Influenza Branch, Division of Viral and Rickettsial Diseases, Centers for Disease Control and Prevention, Atlanta, GA 30333, USA. <sup>5</sup>Department of Molecular Pathology, Armed Forces Institute of Pathology, Rockville, MD 20306, USA.

\*To whom correspondence should be addressed. E-mail: wilson@scripps.edu (I.A.W.) and jstevens@scripps.edu (J.S.)

prisingly achieved by using the 1918 H1 structure (sequence identity: HA1, 58%; HA2, 85%) as a search model (16). Superimposition of human, avian, and swine HA structures by using their HA2 domains (table S2) or individual domains (table S3) confirms that the Viet04 HA is more closely related to human 1918 H1 HA [root mean square deviation (RMSD) 1.2 Å] than to Sing97 H5 HA (RMSD 1.7 Å). For example, an interhelical loop between the two major helices in HA2 is stabilized by a hydrogen bond between HA2 Arg<sup>68</sup> and HA2 Asn<sup>81</sup>, resulting in its having an overall conformation much more akin to the 1918 H1 loop than to that of Sing97 or H3 (Fig. 1C).

The amino acid sequence of Viet04 HA predicts seven possible glycosylation sites per monomer, although one is in the cytoplasmic tail and unlikely to be glycosylated. Interpretable electron density is observed at 16 of the possible 54 glycosylation sites in the asymmetric unit (nine monomers), which represents carbohydrates at two sites, Asn<sup>34</sup> and Asn<sup>169</sup> in HA1 (17).

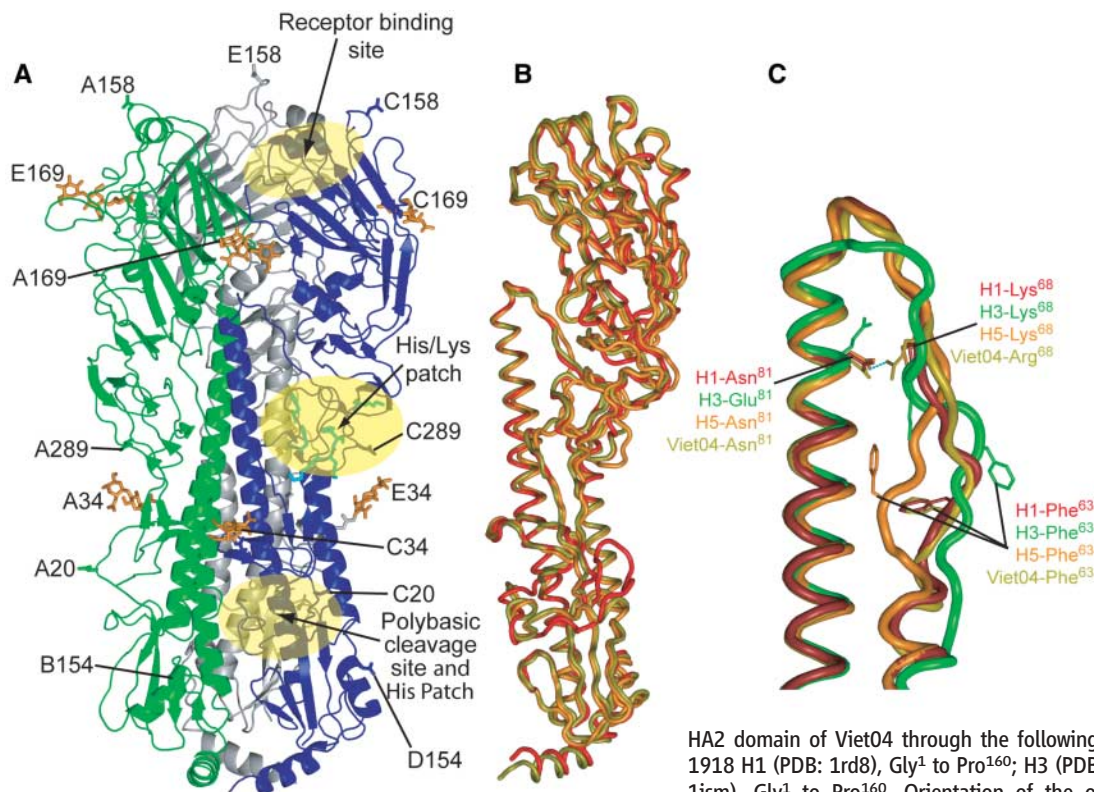
Hemagglutinin is synthesized as a single-chain precursor (HA0) in the endoplasmic reticulum, where it is assembled as a trimer, and is then exported to the cell surface via the Golgi network. On the cell surface, HA0 is cleaved by specific host proteases, such as trypsin Clara (18), into HA1 and HA2 (19). For the majority of HAs, the specific cleavage site (Q/E-X-R)

(20) and the narrow tissue distribution of the relevant proteolytic enzymes restricts infection to the lung in mammals. However, for H5 and H7 subtypes, a polybasic sequence has been associated with high virulence in birds (21), because of enhanced cleavage susceptibility by a broader range of cellular proteases, as seen with our baculovirus-expressed Viet04 HA (fig. S1) (22). Consequently, the tissue tropism for H5 viruses in mammals is not restricted to the lungs, but extends to other organs, including the brain (10). In the Viet04 structure, the C-terminal HA1 cleavage site region could be interpreted only as far as Pro<sup>324</sup> and does not account for the remaining QRERRRKKR residues before Gly<sup>1</sup> at the N terminus of HA2 (fig. S3). As in other HAs, the HA2 N terminus is stabilized within an electronegative cavity by hydrogen bonds from its backbone amide groups to Asp<sup>112</sup> and to Ser<sup>113</sup> of the adjacent HA2 (fig. S3).

From our previous 1918 HA0 structure, we proposed that a pH-sensitive histidine patch (His<sup>A18</sup>, His<sup>A38</sup>, and His<sup>B111</sup>) (14), together with the adjacent HA2 Trp<sup>B21</sup>, could play a role in fusion peptide destabilization and release (Fig. 1A) (11). This structural feature is conserved in other avian and human H1, H2, and H5 serotypes, as well as in Viet04 HA (fig. S3). In 1918 HA0, a second patch of four exposed histidines within the vestigial esterase domain (Fig. 1A and fig. S4A), together with a nearby lysine, was also implicated in pathogenicity via

enhanced membrane fusion (11). Of the five HA1 residues in this basic patch (His<sup>47</sup>, Lys<sup>50</sup>, His<sup>275</sup>, His<sup>285</sup>, and His<sup>298</sup>), only three are conserved in avian H5 structures (His<sup>47</sup>, Lys<sup>50</sup>, and His<sup>298</sup>) (fig. S4, B to D), but Viet04 and Sing97 HAs have an additional lysine (Lys<sup>45</sup>) and histidine (His<sup>295</sup>) (fig. S4, B, C, and E). Furthermore, Viet04 has yet another lysine (Lys<sup>46</sup>), which renders this patch even more basic and is found in two strains (1203/1204) that were isolated from the same patient (10) (fig. S5). The contribution of this region to virulence, if any, is as yet unknown, but is worthy of further investigation.

**H5N1 antigenic variation.** Phylogenetic analysis of H5 HA genes from 2004 and 2005 has revealed two distinct lineages, termed clades 1 and 2 (23); Viet04 belongs to the Indochina peninsula lineage (clade 1). Comparison of their amino acid sequences identified 13 positions of antigenic variation that are mainly clustered around the receptor-binding site; the rest are within the vestigial esterase domain (Fig. 2). Escape mutants of H5 HAs (24, 25) can be clustered into three epitopes (24), as follows: site 1, an exposed loop (HA1 140 to 145) that overlaps with antigenic sites A (26) of H3 (27) and Ca2 of H1 (28); site 2, HA1 residues 156 and 157, which correspond to antigenic site B in H3 serotypes; and site 3, HA1 129 to 133, which is restricted to the Sa site in H1 HAs (28) and H9 serotypes (29). Thus, natural variation



**Fig. 1.** Crystal structure of Viet04 HA and comparison with 1918 human H1, duck H5, and 1968 human H3 HAs. (A) Overview of the Viet04 trimer, represented as a ribbon diagram. For clarity, each monomer has been colored differently. Carbohydrates observed in the electron-density maps are colored orange, and all the asparagines that make up a glycosylation site are labeled. Only Glu<sup>20</sup>, Glu<sup>289</sup>, and Phe<sup>154</sup> are not labeled, as these are on the back of the molecule. The location of the receptor binding, cleavage, and basic patch sites are highlighted only on one monomer. All the figures were generated and rendered with the use of MacPymol (66).

(B) Structural comparison of the Viet04 monomer (olive) with duck H5 (orange) and 1918 H1 (red) HAs. Structures were first superimposed on the

HA2 domain of Viet04 through the following residues: Viet04, Gly<sup>1</sup> to Pro<sup>160</sup>; 1918 H1 (PDB: 1rd8), Gly<sup>1</sup> to Pro<sup>160</sup>; H3 (PDB: 2hmg), Gly<sup>1</sup> to Pro<sup>160</sup>; H5 (PDB: 1j5m), Gly<sup>1</sup> to Pro<sup>160</sup>. Orientation of the overlay approximates to the blue monomer in (A). (C) Superimposition of the two long  $\alpha$ -helices of HA2 for 1918

(yellow in Fig. 2), as well as escape mutants (blue in Fig. 2, green in both 2004 and 2005 viral isolates), suggests continued evolution of the virus that impacts decisions on which strain should be considered for a bird flu vaccine. One mutation that has alanine at residue 160 replaced by threonine (A160T), which is present in all 2004–05 strains, introduces a new glycosylation site at Asn<sup>158</sup>, consistent with a strategy commonly used by influenza viruses to mask and unmask antigenic sites from the immune system (30, 31). This glycosylation likely results in steric hindrance to antigenic site 2 (around residues 156 and 157), thus reducing the ability of the host to mount an effective immune response to these more recent H5N1 viruses.

**Receptor binding domain.** The RBD is at the membrane distal end (HA1) of each HA monomer (Fig. 1A) and binds to its sialic acid-

containing receptors with very weak (millimolar) affinity (32). However, influenza virus can increase its avidity to host cells through multivalent binding via a high density of HA trimers on the virus surface. Avian viruses bind to sialosides with an  $\alpha$ -2-3 linkage in the intestinal tract, whereas human-adapted viruses are specific for the  $\alpha$ -2-6 linkage in the respiratory tract (7), although H5 viruses have also been reported in human intestine (33). A switch from  $\alpha$ -2-3 to  $\alpha$ -2-6 receptor specificity is a critical step in the adaptation of avian viruses to a human host and appears to be one of the reasons why most avian influenza viruses, including current avian H5 strains, are not easily transmitted from human to human after avian-to-human infection.

All HA structures, including Viet04 (Fig. 3A), have similarly configured RBDs. The binding site comprises three structural elements, namely

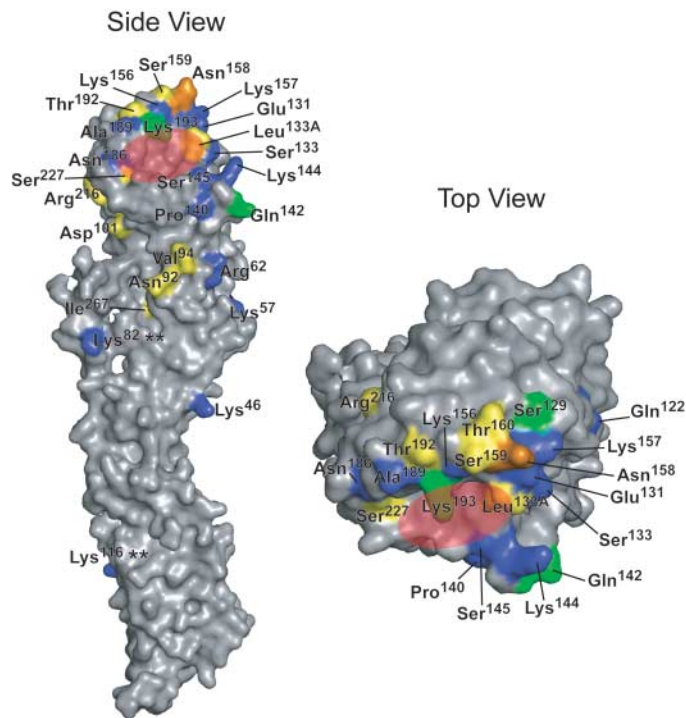
an  $\alpha$ -helix (190-helix, HA1 188 to 190) and two loops (130-loop, HA1 134 to 138, and 220-loop, HA1 221 to 228) (Fig. 3A). A number of conserved residues are involved in receptor binding, including Tyr<sup>98</sup>, Trp<sup>153</sup>, and His<sup>183</sup> (Table 1) (19). Superimposition of the RBD structural elements of Viet04 with Sing97 H5 reveals a very close relation (RMSD 0.3 Å) (Fig. 3B). Indeed, all key residues implicated in receptor specificity [reviewed in (19)] (Table 1) are conserved between structures, although loop 210 to 221 is displaced  $\sim 1$  Å from its equivalent in Sing97 (Fig. 3B). Otherwise, only two RBD residues differ between these two H5 HAs (Viet04, Arg<sup>216</sup> and Ser<sup>221</sup>; Dk97, Glu<sup>216</sup> and Pro<sup>221</sup>). Thus, the question arises as to how a current H5 virus could adapt its HA for binding to human receptors.

#### Receptor binding specificity of Viet04 HA.

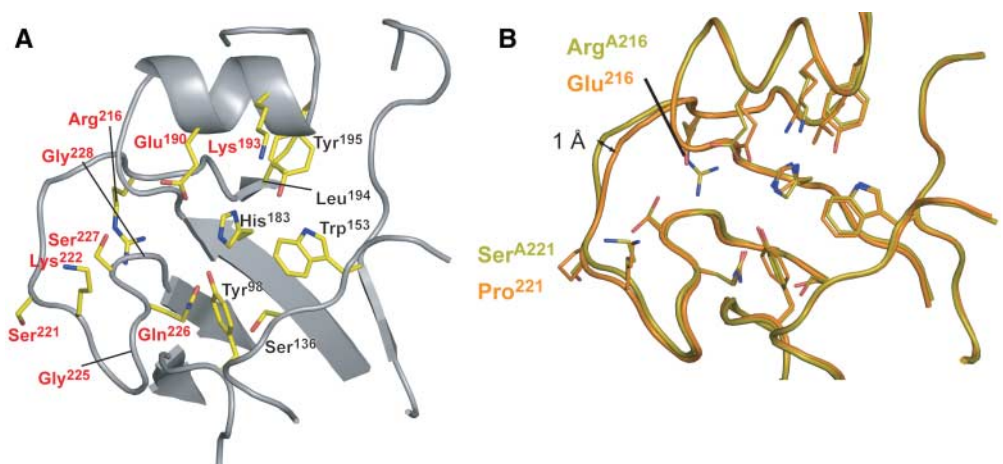
Our cloning and expression strategy produces HA with a His-tag at the C terminus, which facilitates receptor-binding studies using a glycan microarray (34–37). Glycan binding analyses of Viet04 HA reveal an avian  $\alpha$ -2-3 specificity in which the highest affinity is for glycans with sulfate on the 6 position of the *N*-acetylglucosamine (GlcNAc) residue at the third position in the glycan chain (Fig. 4A and table S4) (38, 39). Considerable binding to only one  $\alpha$ -2-6-linked sialoside was observed (6'-sialyllactose, no. 49), but this glycan is only found in milk and is not a receptor candidate for influenza (40). We also expressed and investigated the glycan-binding properties of A/Duck/Singapore/Q-F119-3/1997 (Dk97), whose sequence is identical to that of Sing97, for correlation with its structure (15). Binding of glycoproteins (nos. 1 to 6) and sulfated glycans was comparable to those of Viet04, but binding to other  $\alpha$ -2-3 sialosides was reduced relative to Viet04 (Fig. 4B).

**Mutational analysis of the RBD.** Previous studies using whole virus identified a number of key RBD mutations that were implicated in avian-human receptor specificity switching in H1, H2, and H3 serotypes. However, adaptation of avian H1 and H2/H3 serotypes for human receptor binding occurs by different mecha-

**Fig. 2.** Antigenic variation in recent H5N1 viruses mapped onto the Viet04 structure. (Left) Side view of the Viet04 structure in which natural mutations identified by comparison of 2005 with 2004 isolates (23) are colored yellow; escape mutants (24, 25) are blue; and those that overlap in both analyses are green. All of the 2004 and 2005 strains have a new potential glycosylation site at position 158 in the HA1 chain (orange). The receptor binding site is highlighted with a red oval. (Right) Top view looking down onto the globular membrane distal end of the trimer around the RBD showing that the mutations mainly cluster around the RBD.



**Fig. 3.** Analysis of Viet04 receptor binding site. (A) The Viet04 receptor-binding domain (RBD) with the side chains of key residues for receptor binding labeled. The binding site comprises three structural elements: an  $\alpha$ -helix (190-helix) and two loops (130-loop and 220-loop). Residues mutated in this study are labeled red. (B) Overlay of the RBDs of Viet04 with Sing97 structure (PDB: 1j5m) reveals a similar RBD. The most divergent part of the pocket is the loop made up of residues 210 to 221, in which the Viet04 loop is displaced  $\sim 1$  Å farther away from the binding pocket compared with the 1997 avian H5. Only two residues, at position 216 and 221, differ in these two RBDs.



nisms. For H2 and H3, mutation of Gln<sup>226</sup> and Gly<sup>228</sup> in avian strains to Leu<sup>226</sup> and Ser<sup>228</sup> in human viruses correlates with a shift to human receptor specificity (41, 42). In H1 serotypes, the avian Gln<sup>226</sup> and Gly<sup>228</sup> framework is maintained and a Glu<sup>190</sup> to Asp<sup>190</sup> mutation now appears critical for adaptation to human  $\alpha$ 2-6 receptors (43, 44). Indeed, glycan microarray and cell-based assays revealed that the 1918 HA

could be readily converted from classic  $\alpha$ 2-6 receptor specificity to classic avian  $\alpha$ 2-3 specificity by only two mutations (D190E and D225G) (35, 45). Here, the reverse experiment was performed with an avian H1 virus [A/Duck/Alberta/35/1976 (Dk76)] in which the same two residues were mutated to the “human” sequences (E190D and G225D), which completely converted Dk76 to exclusive  $\alpha$ 2-6 specificity, similar

to that seen for the South Carolina 1918 virus (Figs. 4C and 5, A to C; and table S4) (11, 46).

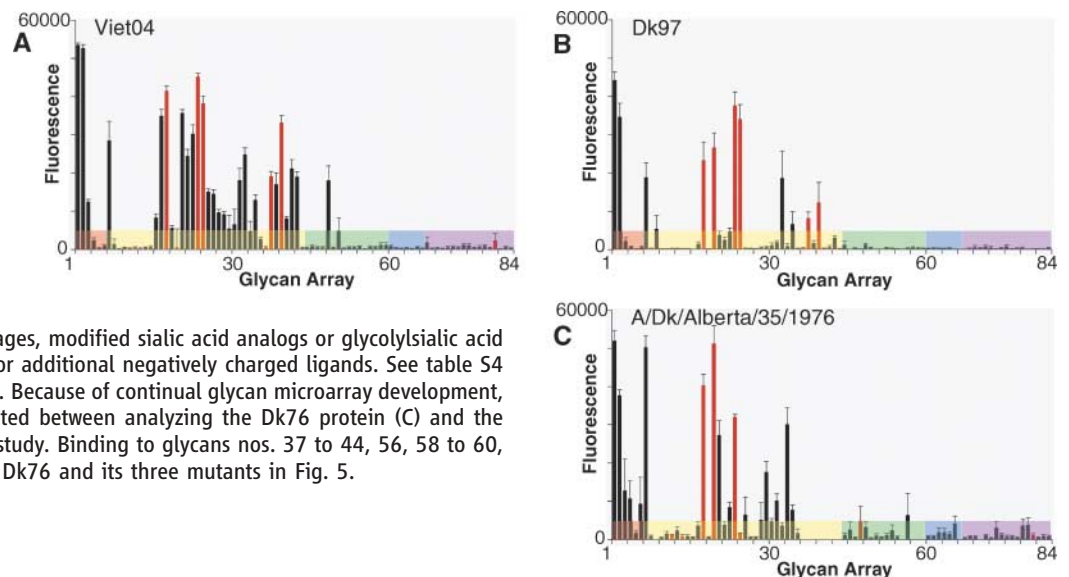
However, which mutations are likely to modulate receptor specificity in the H5 serotype is not so obvious. Based on sequence similarity, H5 is in the same clade as H1, H2, and H6 serotypes (47). So, to address that issue, we analyzed glycan binding of Viet04 HA (Fig. 5 and fig. S6) by generating a panel of mutants

**Table 1.** Conserved residues within the RBDs of H1 and H5 serotypes that are implicated in receptor specificity. Accession numbers for each wild-type HA are listed in supporting online material. Residues mutated in this study are highlighted in gray. The last two columns give a qualitative assessment of  $\alpha$ 2-3/ $\alpha$ 2-6 binding preferences for each mutant with the glycan array.

Viral strain	Amino acid position														Specificity	
	98	136	153	183	190	193	194	216	221	222	225	226	227	228	$\alpha$ 2-3	$\alpha$ 2-6
<b>H1 serotype</b>																
A/Duck/Alberta/35/1976	Y	T	W	H	E	S	L	E	P	K	G	Q	A	G	✓✓	O
A/Duck/Alberta/35/1976 (E190D)	Y	T	W	H	D	S	L	E	P	K	G	Q	A	G	✓	O
A/Duck/Alberta/35/1976 (G225D)	Y	T	W	H	E	S	L	E	P	K	D	Q	A	G	O	O
A/Duck/Alberta/35/1976 (E190D,G225D)	Y	T	W	H	D	S	L	E	P	K	D	Q	A	G	O	✓✓✓
<b>H5 serotype</b>																
A/Duck/Singapore/Q-F119-3/1997	Y	S	W	H	E	K	L	E	P	K	G	Q	S	G	✓✓	O
A/Vietnam/1203/2004	Y	S	W	H	E	K	L	R	S	K	G	Q	S	G	✓✓✓	✓
A/Vietnam/1203/2004 (E190D)	Y	S	W	H	D	K	L	R	S	K	G	Q	S	G	✓✓	O
A/Vietnam/1203/2004 (G225D)	Y	S	W	H	E	K	L	R	S	K	D	Q	S	G	✓✓✓	✓
A/Vietnam/1203/2004 (E190D,G225D)	Y	S	W	H	D	K	L	R	S	K	D	Q	S	G	O	O
A/Vietnam/1203/2004 (Q226L)	Y	S	W	H	E	K	L	R	S	K	G	L	S	G	✓	O
A/Vietnam/1203/2004 (S227N)	Y	S	W	H	E	K	L	R	S	K	G	Q	N	G	✓✓✓	✓
A/Vietnam/1203/2004 (G228S)	Y	S	W	H	E	K	L	R	S	K	G	O	S	S	✓✓✓	✓✓✓*
A/Vietnam/1203/2004 (Q226L,G228S)	Y	S	W	H	E	K	L	R	S	K	G	L	S	S	✓✓	✓✓✓*
A/Vietnam/1203/2004 (R216E)	Y	S	W	H	E	K	L	E	S	K	G	Q	S	G	ND	ND
A/Vietnam/1203/2004 (S221P)	Y	S	W	H	E	K	L	R	P	K	G	Q	S	G	✓✓✓	✓
A/Vietnam/1203/2004 (R216E,S221P)	Y	S	W	H	E	K	L	E	P	K	G	Q	S	G	✓✓✓	✓

\*Although Viet04 mutants (G228S and Q226L,G228S) only bound a limited number of  $\alpha$ 2-6 ligands, they bound strongly to these glycans and were, therefore, assessed as ✓✓ for  $\alpha$ 2-6 specificity. No binding is represented by “O”; ND indicates binding to the array was not determined.

**Fig. 4.** Glycan microarray analyses of (A) Viet04, (B) Dk97, and (C) an avian H1, Dk76. The Dk97 HA sequence is identical to that in the published structure of duck virus Sing97, so a direct structural comparison can be made. Binding to different types of glycans on the array are highlighted where orange represents glycoproteins; yellow,  $\alpha$ 2-3 ligands; green,  $\alpha$ 2-6 ligands; blue,  $\alpha$ 2-8 ligands; and purple, other ligands such as  $\beta$ -linkages, modified sialic acid analogs or glycolysialic acid glycans. Red bars indicate sulfated or additional negatively charged ligands. See table S4 for list and tabulated binding results. Because of continual glycan microarray development, a number of new ligands were printed between analyzing the Dk76 protein (C) and the remaining samples reported in this study. Binding to glycans nos. 37 to 44, 56, 58 to 60, 67, and 70 was not determined for Dk76 and its three mutants in Fig. 5.

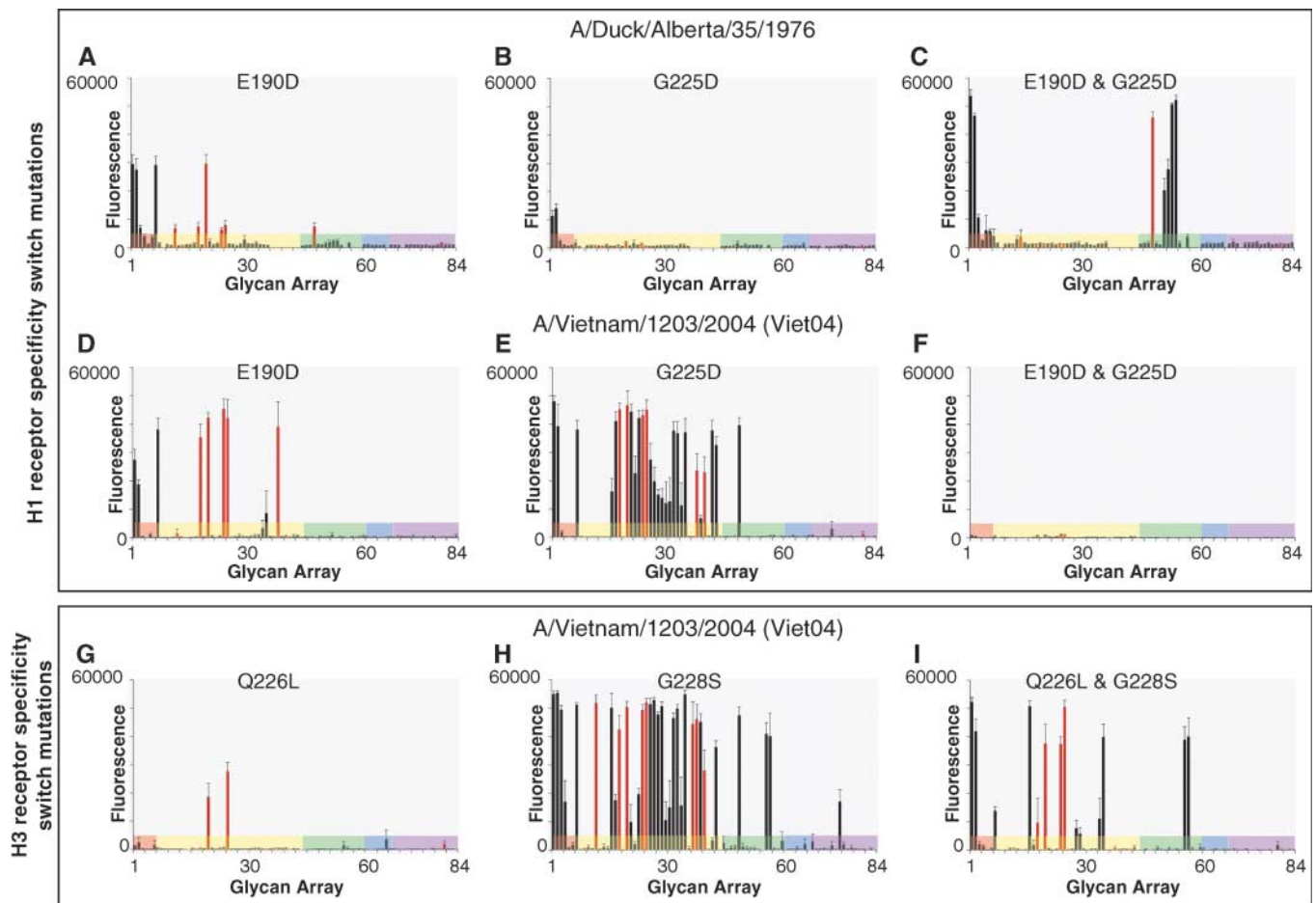


(Fig. 3A and Table 1) in and around the RBD to explore whether this H5 HA can readily become adapted to humans through mutations that are known to change receptor specificity in H1 and H3 serotypes. Mutations at positions 190 and 225 did not reveal any adaptation of Viet04 to human receptor analogs (Fig. 5, D to F) (48), in contrast to H1 Dk76 (Fig. 5, A to C) and 1918 HAs (35). Indeed, the single E190D mutation on the Viet04 framework reveals markedly reduced affinity to  $\alpha$ 2-3 sialosides (Fig. 5D), whereas the double mutant (E190D,G225D) did not interact at all with the glycan microarray (Fig. 5F) (49, 50). However, sulfated glycans bound equally well to the single E190D mutant and to the wild type (Figs. 4A and 5D), which suggests that other residues within the Viet04 RBD, such as Lys<sup>193</sup> or Lys<sup>222</sup> (Fig. 3A), may enhance interaction with charged glycans.

Mutation of residues 226 and 228, which enable H3 viruses to switch from avian to human specificity, was also evaluated as a potential

route for H5 viruses to acquire human receptor specificity. Although a dramatic switch to a classic  $\alpha$ 2-6 human receptor binder was not observed (51), the double mutant (Q226L,G228S) showed substantially reduced affinity to  $\alpha$ 2-3 sialosides, as noted for mutants of the H3 A/Hong Kong/156/1997 virus (52). But it was notable that significant binding to a natural, branched  $\alpha$ 2-6 biantennary glycan (nos. 56 and 57) was observed for both the double mutant and the single G228S mutant (Fig. 5H). Although the glycan composition of lung epithelia have not been analyzed in detail, the mammalian sialyltransferase that produces  $\alpha$ 2-6-linked structures on many human tissues (53, 54) is found in lung epithelial cells (55–57). Thus, these two effects could offer advantages for an H5N1 virus to adapt to a human host. Decreased binding to  $\alpha$ 2-3-linked glycans would help circumvent the inhibitory effects of respiratory mucins (58), whereas increased binding to biantennary N-linked glycans with  $\alpha$ 2-6-linked

sialic acids would allow the virus to attach to the surface of epithelial cells that express this carbohydrate receptor (55–57). In this regard, human H1 viruses before 1957 were reported to bind sialic acid receptors with both  $\alpha$ 2-3 and  $\alpha$ 2-6 linkages; post 1957 viruses were specific only for  $\alpha$ 2-6 linkages (37). These binding patterns suggest that, once a foothold in a new host species is made, the virus HA optimizes its specificity to the new host. It is noteworthy that, of the HAs tested on the array, the humanized avian H1 (Dk76) double mutant (E190D,G225D) (Fig. 5C) and the human H3 HA (A/Moscow/10/1999) (35) did not bind  $\alpha$ 2-6 biantennary glycans, in contrast to 1918 South Carolina H1 HA and human H1, A/Texas/36/1991 (35). Therefore, the HAs of some viruses may be able to increase avidity through interaction with such bivalent structures on N-linked glycans, whereas, for others, the geometry of the bivalent structure appears to restrict binding to linear sequences containing  $\alpha$ 2-6



**Fig. 5.** Glycan microarray analysis of mutants of Viet04 and Dk76. Mutations of an avian H1, Dk76: (A) E190D, (B) G225D, and (C) E190D and G225D were generated and subjected to glycan microarray analysis. Both positions were reported to be important for conversion of  $\alpha$ 2-6 receptor specificity of the human 1918 virus HA to avian  $\alpha$ 2-3 specificity (35, 45). These mutations did indeed result in exclusive  $\alpha$ 2-6 specificity for this avian H1 HA. (D to F) Consequently, Viet04 mutations were generated at the same positions, but did not result in a switch of receptor specificity,

except to 6'-sialyllactose, although they did result in decreased  $\alpha$ 2-3 binding, particularly to nonsulfated glycans (compare Fig. 4A). (G to I) Viet04 was mutated at positions 226 and 228, known to be important for H3 HA  $\alpha$ 2-6 receptor adaptation. Again, no clear switch in receptor specificity was observed, although binding to biantennary  $\alpha$ 2-6 moieties was observed, as well as reduced  $\alpha$ 2-3 binding in the double and single (Q226L) mutant. Graphs are generated as described in the legend for Fig. 4 and labels to the introduced mutations.

linkages. Thus, although human viral HAs have a primary specificity for  $\alpha$ 2-6 linkages, each may use a different spectrum of glycan receptors for cell entry.

All key residues within the RBD are conserved in the majority of H5 strains that have infected humans (fig. S5). However, two A/Hong Kong/2003 (HK2003) isolates acquired a S227N mutation within the binding site, whereas a double mutation (E216R,P221S) in the 220-loop is observed in all 2003–05 isolates (fig. S5). The possible effect of these natural mutations on Viet04 HA binding specificity (Table 1) was, therefore, assessed. The S227N mutation had comparable specificity to that of Viet04, with the exception of increased binding, particularly for branched  $\alpha$ 2-3 fucosylated glycans (nos. 26 to 29) and for 6-sialylated *N*-acetylgalactosamine (GalNAc) (no. 20) (fig. S6A) (59, 60), contrary to previous reports that HK2003 isolates had increased affinity toward  $\alpha$ 2-6 analogs, but decreased affinity toward  $\alpha$ 2-3 analogs (39). However, in a previous study from a 1997 isolate, such changes were also not observed (52), although Viet04 differs at a number of other positions around the RBD compared with the Hong Kong isolates that could account for this difference (61) (Fig. 3A). Reverse R216E and S221P mutants were also generated, as well as the double mutant (R216E,S221P), but the R216E mutant expressed poorly and could not be analyzed. However, only the double mutant is found in natural isolates, suggesting a pressure to select for both mutations, which possibly are related to the HA stability. Whereas Viet04 HA binds to branched fucosylated sialosides (nos. 26 to 29) (Fig. 4A), the S221P mutation showed weaker binding, whereas the double mutant abrogated binding to all branched fucosylated glycans unless sulfated (no. 25) (fig. S6, B and C). In the Viet04 HA structure, these residues hydrogen bond to an adjacent monomer in the trimer (Arg<sup>216</sup> with Asn<sup>210</sup> and Ser<sup>221</sup> with Asp<sup>241</sup>) (15) and stabilize the displaced 210 to 229 loop (Fig. 3B), which, therefore, could possibly enhance binding to branched fucosylated glycans.

So how might H5 avian HA adapt to human receptors? Knowledge of genetic changes in circulating viral isolates (39) by themselves obviously cannot be used to predict the impact on receptor specificity, let alone predict the effect of future mutations. Here, we use a completely recombinant system for structural and functional analyses that enables such investigation in the laboratory. Our conclusion is that the mutations that cause a shift from the avian-type to human-type specificity on the H1 and H3 frameworks do not cause an equivalent shift in specificity on the H5 framework of the Viet04 isolate. However, the mutations that give rise to  $\alpha$ 2-6 specificity in H3 HAs do in fact reduce avidity to  $\alpha$ 2-3 sialosides and increase specificity for  $\alpha$ 2-6-linked biantennary N-linked glycans that could serve as receptors for the virus on lung

epithelial cells. These combined effects could allow the Viet04 virus to escape entrapment by mucins and increase the likelihood of binding to and infection of susceptible epithelial cells (52). Thus, such mutations provide one possible route by which H5 viruses could gain a foothold in the human population, although it is possible that other, as yet unidentified, mutations may allow the H5N1 virus to effect a switch in receptor specificity.

This glycan microarray technology can, therefore, be used to analyze not only existing viral HAs, but as we show here, to identify mutations that enable adaptation of the remaining influenza serotypes into the human population. Monitoring such changes in the “receptor binding footprint” in the field on whole viruses using the glycan microarray could be invaluable in the identification of emerging viruses that could cause new pandemics or epidemics.

#### References and Notes

- WHO ([www.who.int/en/](http://www.who.int/en/)).
- C. Scholtissek, W. Rohde, V. Von Hoyningen, R. Rott, *Virology* **87**, 13 (1978).
- Y. Kawaoka, W. J. Bean, R. G. Webster, *Virology* **169**, 283 (1989).
- N. P. Johnson, J. Mueller, *Bull. Hist. Med.* **76**, 105 (2002).
- J. K. Taubenberger *et al.*, *Nature* **437**, 889 (2005).
- T. M. Tumpey *et al.*, *Science* **310**, 77 (2005).
- C. R. Parrish, Y. Kawaoka, *Annu. Rev. Microbiol.* **59**, 553 (2005).
- Y. Suzuki *et al.*, *J. Virol.* **74**, 11825 (2000).
- E. A. Govorkova *et al.*, *J. Virol.* **79**, 2191 (2005).
- T. R. Maines *et al.*, *J. Virol.* **79**, 11788 (2005).
- J. Stevens *et al.*, *Science* **303**, 1866 (2004).
- Materials and Methods are available as supporting material on *Science* Online.
- Viet04 HA at a concentration of 9 mg/ml was used to grow crystals in sitting drops with a precipitant solution of 22% polyethylene glycol 2000 and 0.1 M Hepes, pH 6.55 (see also supporting online material).
- 1918 H1 HA0 (PDB: 1rd8), truncated to remove residues around the cleavage site, was used as the initial MR model. The final  $R_{\text{cryst}}$  and  $R_{\text{free}}$  values are 26.9 and 31.9% respectively, at 2.9 Å resolution. The crystal asymmetric unit contains nine hemagglutinin monomers (six HA monomers in two noncrystallographic trimers and three HA monomers that each form one-third of three crystallographic trimers) with an estimated solvent content of 57% based on a Matthews' coefficient ( $V_m$ ) of 2.9 Å<sup>3</sup>/dalton (fig. S2). For comparison with previous structures, the Viet04 sequences are numbered as for the H3 subtype. A, C, E, G, I, K, M, O, and Q refer to the nine HA1 subunits in the asymmetric unit, and B, D, F, H, J, L, N, P, and R refer to the nine HA2 subunits; e.g., His<sup>A18</sup> refers to HA1 residue 18 in the A subunit and His<sup>B11</sup> refers to HA2 residue 111 in the B subunit of the same monomer. Insertions in Viet04 relative to H3 are labeled by the preceding residue with a letter (e.g., Asn<sup>19A</sup>).
- Y. Ha, D. J. Stevens, J. J. Skehel, D. C. Wiley, *EMBO J.* **21**, 865 (2002).
- Scores from the molecular replacement program PHASER revealed superior scores for the 1918 H1 structure ( $Z$  score: 37.2; and log-likelihood gain, 3412), as compared with the Sing97 structure ( $Z$  scores, 33.8; and log-likelihood gain, 768).
- Two *N*-acetyl glucosamines were interpretable at 13 of these sites (Asn<sup>A34</sup>, Asn<sup>C34</sup>, Asn<sup>G169</sup>, Asn<sup>F34</sup>, Asn<sup>G34</sup>, Asn<sup>I34</sup>, Asn<sup>L169</sup>, Asn<sup>K34</sup>, Asn<sup>K169</sup>, Asn<sup>M34</sup>, Asn<sup>M169</sup>, Asn<sup>O34</sup>, Asn<sup>O169</sup>), but an additional mannose residue could be interpreted at a further three sites (Asn<sup>A169</sup>, Asn<sup>E169</sup>, Asn<sup>G169</sup>). The glycans are stabilized at Asn<sup>34</sup> by a neighboring residue (Gln<sup>24</sup>) in the same chain, whereas
- at Asn<sup>169</sup>, an additional mannose was visualized because of stabilization with Lys<sup>56</sup> and main-chain amide of Val<sup>57</sup>, in a symmetry-related monomer.
- H. Kido *et al.*, *J. Biol. Chem.* **267**, 13573 (1992).
- J. J. Skehel, D. C. Wiley, *Annu. Rev. Biochem.* **69**, 531 (2000).
- Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; X, any amino acid; and Y, Tyr.
- D. J. Hulse, R. G. Webster, R. J. Russell, D. R. Perez, *J. Virol.* **78**, 9954 (2004).
- H. D. Klenk, W. Garten, *Trends Microbiol.* **2**, 39 (1994).
- WHO Global Influenza Program Surveillance Network, *Emerg. Infect. Dis.* **11**, 1515 (2005).
- N. V. Kaverin *et al.*, *J. Gen. Virol.* **83**, 2497 (2002).
- M. Philpott, C. Hioe, M. Sheerar, V. S. Hinshaw, *J. Virol.* **64**, 2941 (1990).
- Regions of antigenic variation have been identified in H1 and H3 serotypes. For H1, these sites were designated Sa, Sb, Ca, and Cb; for H3, sites were designated A, B, C, and D.
- D. C. Wiley, I. A. Wilson, J. J. Skehel, *Nature* **289**, 373 (1981).
- A. J. Caton, G. G. Brownlee, J. W. Yewdell, W. Gerhard, *Cell* **31**, 417 (1982).
- N. V. Kaverin *et al.*, *J. Virol.* **78**, 240 (2004).
- M. L. Perdue, D. L. Suarez, *Vet. Microbiol.* **74**, 77 (2000).
- S. J. Baigent, J. W. McCauley, *Virus Res.* **79**, 177 (2001).
- N. K. Sauter *et al.*, *Biochemistry* **31**, 9609 (1992).
- J. H. Beigel *et al.*, *N. Engl. J. Med.* **353**, 1374 (2005).
- O. Blixt *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 17033 (2004).
- J. Stevens *et al.*, *J. Mol. Biol.* **355**, 1143 (2006).
- HA binding can be analyzed not only for sialic acid-linkage preference, but also for additional features, such as charge; glycan length; or additional sulfation, fucosylation, and sialylation. Of the 265 glycans currently imprinted on the array, 6 are glycoproteins; 38 have sialic acids with  $\alpha$ 2-3 linkages; 16 have  $\alpha$ 2-6 linkages; 7 have  $\alpha$ 2-8 linkages; and a further 16 are  $\beta$ -linkages, modified sialic acid analogs, or glycolysialic acid glycans. (See table S4 for the glycans analyzed in this study. Of the  $\alpha$ 2-6 sialosides, only natural full-sized N-linked glycans represented on the array are the biantennary structures (nos. 56 and 57). The remaining sialosides are fragments or terminal sequences found on glycoproteins. For full information on the array, contact the Consortium for Functional Glycomics (62). Previous binding data using this technology and cell-based assays with whole viruses show that N-linked glycans close to the receptor-binding site can affect receptor binding through steric hindrance (35, 63). Insect cells do not produce complex glycans containing terminal galactose and/or sialic acids, as seen in mammalian cells, although high-mannose glycans are produced (64). However, because of the presence of the influenza sialidase, complex glycans of influenza HAs usually terminate only in galactose, and thus the size of the N-glycans elaborated by insect cells approximate to the size of the complex N-glycans in mammalian host cells. Thus, any importance of complex glycans for HA function is still unknown. Indeed, results for the avian H3 HA (A/Duck/Ukraine/1/1963), published recently (35), are in agreement with previous whole viral studies (65). However, independent studies are ongoing to develop the array for whole-virus analyses so that a direct comparison can be made. Such initial experiments are promising, because the strict  $\alpha$ 2-3 specificity observed here for Dk76 is also seen with whole-virus studies (37) and preliminary experiments with A/Puerto Rico/8/1934 virus that reveal both  $\alpha$ 2-3 and  $\alpha$ 2-6 specificity (34), in agreement with experiments from cell-based assays (37).
- G. N. Rogers, B. L. D'Souza, *Virology* **173**, 317 (1989).
- Whole-virus studies, including those for Viet04 virus, also revealed  $\alpha$ 2-3 specificity with a preference for sulfation in current H5N1 strains (39). However, this assay used only seven ligands (one  $\alpha$ 2-6 and six  $\alpha$ 2-3), which is considerably fewer than the 84 sialosides, sialoside analogs, and glycoproteins analyzed here. In our glycan array,

- sulfation on the second galactose was not tolerated (no. 37) for Viet04, although binding was apparent for sialosides with Gal in either  $\beta$ 1-3 or  $\beta$ 1-4 linkage to a GlcNAc or GalNAc (nos. 21 to 23, 32, 33), as well as to fucosylated glycans (nos. 26 to 29).
39. A. Gambaryan *et al.*, *Virology* **344**, 432 (2006).
  40. H. Debray, D. Decout, G. Strecker, G. Spik, J. Montreuil, *Eur. J. Biochem.* **117**, 41 (1981).
  41. R. J. Connor, Y. Kawaoka, R. G. Webster, J. C. Paulson, *Virology* **205**, 17 (1994).
  42. G. N. Rogers *et al.*, *Nature* **304**, 76 (1983).
  43. M. Matrosovich *et al.*, *J. Virol.* **74**, 8502 (2000).
  44. E. Nobusawa, H. Ishihara, T. Morishita, K. Sato, K. Nakajima, *Virology* **278**, 587 (2000).
  45. L. Glaser *et al.*, *J. Virol.* **79**, 11533 (2005).
  46. Avian H1 bound only to nonbranched glycans and to sulfated and/or negatively charged glycans (Figs. 4C and 5, A to C). The single E190D mutation reduced binding to most  $\alpha$ 2-3 glycans, except to sulfated sialosides (Fig. 5A). These results suggest mutation at both 190 and 225 positions is always a requirement for H1 serotypes to adapt to a human host.
  47. R. A. Fouchier *et al.*, *J. Virol.* **79**, 2814 (2005).
  48. The E190D mutation (Fig. 5D) reduced overall binding of  $\alpha$ 2-3 ligands and glycoproteins, except for the sulfated and/or negatively charged glycans (nos. 18, 20, 24, 25, and 38). The G225D mutation (Fig. 5E) appeared to have little effect on the binding profile, in contrast to avian H1, where binding was not detected (Fig. 5B). The double mutant (E190D,G225D) did not bind to any glycan on the array (see Fig. 5F).
  49. S. J. Gamblin *et al.*, *Science* **303**, 1838 (2004).
  50. For the human H1 HA from A/Puerto Rico/8/1934, the longer side chain of Glu<sup>190</sup> can form hydrogen bonds to sialic acid of both  $\alpha$ 2-6 and  $\alpha$ 2-3 sialosides, whereas for structures of A/Swine/Iowa/1930, H1 HA bound to human receptor analogs, the shorter side chain of Asp<sup>190</sup> can only interact with the GlcNAc to stabilize the  $\alpha$ 2-6 conformation (49). Binding data, with the 1918 South Carolina H1 HA (35) and the Dk76 double mutation (E190D,G225D) (Fig. 5C), show that some sulfated glycans with  $\alpha$ 2-6 sialic acid linkages can bind. However, this situation does not arise for the Viet04 double mutant. Although the G225D mutation would have been expected to enhance  $\alpha$ 2-6 specificity, the additional stabilizing influence of the E190D mutation toward the GlcNAc may not be possible because of the neighboring Lys<sup>193</sup>, which could inhibit interaction of Asp<sup>190</sup> with the glycan either by steric hindrance or by direct interaction with Asp<sup>190</sup>. Experiments are in progress to test this notion.
  51. The Q226L mutation eliminated binding to the microarray, except for two negatively charged  $\alpha$ 2-3 glycans [with either an extra sialic acid on the 6-position of a GalNAc (no. 20) or 6-sulfation on GlcNAc with a branched fucose (no. 25)]. The G228S mutation did not have any significant effect compared with Viet04, except that sialosides with sulfation on the 6-position of the galactose, with or without branched fucosylation on the GlcNAc (nos. 12, 37) were tolerated. Stronger binding was observed for fucosylated glycans (nos. 26 to 29), and reduced binding was observed for sialosides with  $\beta$ 1-3 linkages between the galactose and GlcNAc/GalNAc (nos. 21 to 23) (Fig. 5H). In addition to 6'-sialyllactose (no. 49), as seen for Viet04, binding was observed for  $\alpha$ 2-6 biantennary structures (nos. 56 and 57). The double mutant (Q226L,G228S) showed reduced binding to  $\alpha$ 2-3 sialosides. Only sulfated and long-chain glycans were tolerated (nos. 16, 20, 24, 25, 35), but binding to  $\alpha$ 2-6 biantennary structures (nos. 56 and 57), as with the G228S mutation, was also maintained.
  52. R. Harvey, A. C. Martin, M. Zambon, W. S. Barclay, *J. Virol.* **78**, 502 (2004).
  53. D. H. Joziase *et al.*, *J. Biol. Chem.* **262**, 2025 (1987).
  54. See glycan structure database ([www.functionalglycomics.org](http://www.functionalglycomics.org)).
  55. L. G. Baum, J. C. Paulson, *Acta Histochem. Suppl.* **40**, 35 (1990).
  56. P. Gagneux *et al.*, *J. Biol. Chem.* **278**, 48245 (2003).
  57. M. N. Matrosovich, T. Y. Matrosovich, T. Gray, N. A. Roberts, H. D. Klenk, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 4620 (2004).
  58. G. Lamblin *et al.*, *Glycoconj. J.* **18**, 661 (2001).
  59. Attenuated viruses with a S227N mutation led to higher hemagglutinin inhibition titers in ferrets (60). Thus, enhanced binding to  $\alpha$ 2-3 ligands, especially to 6-sulfated GalNAc, could lead to an increased uptake into antigen-presenting cells and subsequent antibody production.
  60. E. Hoffmann, A. S. Lipatov, R. J. Webby, E. A. Govorkova, R. G. Webster, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 12915 (2005).
  61. The 2003 isolates contain Ala<sup>160</sup>, Arg<sup>193</sup>, Lys<sup>216</sup> and Asn<sup>227</sup>, whereas Viet04 has Thr<sup>160</sup> (which introduces a glycosylation site at Asn<sup>158</sup>), Lys<sup>193</sup>, Arg<sup>216</sup>, and Ser<sup>227</sup>.
  62. Consortium for Functional Glycomics ([www.functionalglycomics.org](http://www.functionalglycomics.org)).
  63. H. D. Klenk, R. Wagner, D. Heuer, T. Wolff, *Virus Res.* **82**, 73 (2002).
  64. T. A. Kost, J. P. Condeary, D. L. Jarvis, *Nat. Biotechnol.* **23**, 567 (2005).
  65. G. N. Rogers, J. C. Paulson, *Virology* **127**, 361 (1983).
  66. W. L. Delano (2002); ([www.pymol.org](http://www.pymol.org)).
  67. The work was supported in part by National Institute of Allergy and Infectious Diseases grant AI058113 (I.A.W., T.T., J.K.T.); National Institute of General Medical Sciences grants GM062116 (to J.C.P., I.A.W.) and GM060938 (to J.C.P.); and partial support from NIH grants to I.A.W. (CA55896 and AI42266). We thank P. Palese and L. Glaser (Mount Sinai School of Medicine, New York) for providing the full-length clone of A/Vietnam/1203/2004; the staff of the Advanced Light Source Beamline 8.2.2 for the beamline assistance; X. Dai, S. Ferguson, P. Carney, and J. Vanhansy (The Scripps Research Institute) for expert technical assistance; and R. Stanfield and M. Elsliger (The Scripps Research Institute) for helpful discussions. This is publication 17916-MB from The Scripps Research Institute. Coordinates and structure factors have been deposited in the Protein Data Bank (code 2FK0) and will be released on publication.

#### Supporting Online Material

[www.sciencemag.org/cgi/content/full/1124513/DC1](http://www.sciencemag.org/cgi/content/full/1124513/DC1)  
Materials and Methods  
Figs. S1 to S6  
Tables S1 to S4  
References

3 January 2006; accepted 28 February 2006  
Published online 16 March 2006;  
10.1126/science.1124513  
Include this information when citing this paper.

## REPORTS

# Ultrafast Laser-Driven Microlens to Focus and Energy-Select Mega-Electron Volt Protons

Toma Toncian,<sup>1</sup> Marco Borghesi,<sup>2</sup> Julien Fuchs,<sup>3</sup> Emmanuel d'Humières,<sup>3,4</sup> Patrizio Antici,<sup>3</sup> Patrick Audebert,<sup>3</sup> Erik Brambrink,<sup>3</sup> Carlo Alberto Cecchetti,<sup>2</sup> Ariane Pipahl,<sup>1</sup> Lorenzo Romagnani,<sup>2</sup> Oswald Willi<sup>1\*</sup>

We present a technique for simultaneous focusing and energy selection of high-current, mega-electron volt proton beams with the use of radial, transient electric fields ( $10^7$  to  $10^{10}$  volts per meter) triggered on the inner walls of a hollow microcylinder by an intense subpicosecond laser pulse. Because of the transient nature of the focusing fields, the proposed method allows selection of a desired range out of the spectrum of the polyenergetic proton beam. This technique addresses current drawbacks of laser-accelerated proton beams, such as their broad spectrum and divergence at the source.

The recent development of ultra-intense laser pulses (1) has opened up opportunities for applications in many areas, including particle acceleration (2–5), inertial fusion energy (6), generation of intense x-ray

pulses (7), laser-driven nuclear physics (8), and laboratory astrophysics (9). In particular, the acceleration of mega-electron volt ions from the interaction of high-intensity laser-pulses with thin solids has major applicative prospects

because of the high beam quality of these ion bursts (10, 11). Such proton beams are already applied to produce high-energy density matter (12) or to radiograph transient processes (13), and they offer promising prospects for tumor therapy (14), isotope generation for positron emission tomography (15), fast ignition of fusion cores (16), and brightness increase of conventional accelerators. However, because these proton beams are polyenergetic and divergent at the source, reduction and control of their divergence and energy spread are essential requirements for most of these applications.

<sup>1</sup>Heinrich Heine Universität Düsseldorf, D-40225 Düsseldorf, Germany. <sup>2</sup>School of Mathematics and Physics, The Queen's University of Belfast, Belfast BT7 1NN, Northern Ireland, UK. <sup>3</sup>Laboratoire pour l'Utilisation des Lasers Intenses, UMR 7605 CNRS-CEA-Ecole Polytechnique-Université Paris VI, 91128 Palaiseau, France. <sup>4</sup>Centre de Physique Théorique, UMR 7644 CNRS-Ecole Polytechnique, 91128 Palaiseau, France.

\*To whom correspondence should be addressed. E-mail: [oswald.willi@laserphy.uni-duesseldorf.de](mailto:oswald.willi@laserphy.uni-duesseldorf.de)