

SUPPLEMENTAL DATA

Structure and receptor binding of the hemagglutinin from a human H6N1 influenza virus

Netanel Tzarum,¹ Robert P. de Vries,^{2,3,4} Xueyong Zhu,¹ Wenli Yu,¹ Ryan McBride,^{2,3,4}
James C. Paulson,^{2,3,4*} Ian A. Wilson^{1,5*}

¹Department of Integrative Structural and Computational Biology,

²Department of Cell and Molecular Biology,

³Department of Chemical Physiology,

⁴Department of Immunology and Microbial Science,

⁵Skaggs Institute for Chemical Biology,

The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037,

USA

* Correspondence: wilson@scripps.edu, jpaulson@scripps.edu

A

```
Consensus sequence MIAIIVIAILAAAGKSDKICIGYHANNSTTQVDTILEKNVTVTHSVELLENQKEERFCKI
A/Taiwan/2/2013 -----PG-----L-----K-----
A/Chicken/Taiwan/A2837/2013 . . . . . V . . . . . S . . . . . L . . . . . K . . . . .
A/chicken/Taiwan/TC135/2010 -----R-----L-----I-----S-----N-----N-----
A/chicken/Taiwan/ch1006/04 -----L-----I-----K-----
A/chicken/Taiwan/0305/04 -----L-----I-----K-----
A/chicken/Taiwan/0107/02 -----L-----S-----K-----
A/chicken/Taiwan/2838N/00 -----L-----S-----K-----
A/chicken/Taiwan/SP1/00 -----I-----
A/chicken/Taiwan/0705/99 -----I-----
A/chicken/Taiwan/G2/87 -----I-----
A/chicken/Taiwan/0222/02 . . . . . S . . . . . L . . . . . K . . . . .
A/chicken/Taiwan/67/2013 . . . . . V . . . . . S . . . . . L . . . . . K . . . . .
A/chicken/Taiwan/G23/87 . . . . . T T . . . . .
A/chicken/Taiwan/ch1006/04 . . . . . S T . . . . . L . . . . . I . . . . . K . . . . .
A/duck/Guangxi/3333/2006 . . . . . R . . . . .
A/duck/Guangxi/GXd-5/2010 . . . . . F . . . . . T E . . . . . K . . . . .
A/duck/Hong Kong/202/77 . . . . . I . . . . . V . . . . . T . . . . .
A/duck/Hong Kong/3461/99 . . . . . I . . . . .
A/duck/Korea/S17/03 . . . . . I . . . . . S . . . . . I . . . . .
A/duck/Taiwan/0526/72 . . . . . T T . . . . . F . . . . .
A/gray teal/Australia/1/1979 . . . . . I F . . . . . V S T S . . . . . I . . . . . S . . . . . L R V
A/mallard/California/6695/2009 . . . . . M . . . . . T . . . . . R . . . . .
A/mallard/Jiangxi/7787/2003 . . . . . R . . . . .
A/mallard/Sweden/30/2005 . . . . .
```

```
Consensus sequence LNKAPLDLRGCTIEGWILGNPQCDDLQDQSWSYIVERPTAQNIGICYPGVLNEVEELKAL
A/Taiwan/2/2013 M . . . . . K D . . . . . K . . . . . N . . . . . L . . . . . F
A/Chicken/Taiwan/A2837/2013 M . . . . . K D . . . . . K . . . . . N . . . . . L . . . . . F
A/chicken/Taiwan/TC135/2010 M . . . . . D . . . . . K . . . . . N . . . . . L . . . . . F
A/chicken/Taiwan/ch1006/04 M . . . . . D . . . . . K . . . . . N . . . . . L . . . . . F
A/chicken/Taiwan/0305/04 M . . . . . S . . . . . E . . . . . K . . . . . T . . . . .
A/chicken/Taiwan/0107/02 M . . . . . D . . . . . K . . . . . T . . . . .
A/chicken/Taiwan/2838N/00 M K . G . . . . . E . . . . . K . . . . . T . . . . .
A/chicken/Taiwan/SP1/00 M . . . . . S . . . . . E . . . . . K . . . . . A . . . . .
A/chicken/Taiwan/0705/99 M . . . . . S . . . . . E . . . . . K . . . . . A . . . . .
A/chicken/Taiwan/G2/87 M . . . . . D . . . . . K . . . . . I . . . . . N . . . . .
A/chicken/Taiwan/0222/02 M . . . . . K D . . . . . K . . . . . N . . . . . A . . . . . L . . . . . F
A/chicken/Taiwan/67/2013 M . . . . . D . . . . . K . . . . . N . . . . . A . . . . . L . . . . . F
A/chicken/Taiwan/G23/87 M . . . . . D . . . . . K . . . . . N . . . . . A . . . . . L . . . . . F
A/chicken/Taiwan/ch1006/04 M . . . . . D . . . . . K . . . . . N . . . . . A . . . . . L . . . . . F
A/duck/Guangxi/3333/2006 . . . . . G V . . . . . K . . . . . A . . . . .
A/duck/Guangxi/GXd-5/2010 S . . . . . D . . . . . L . . . . . R . G I . . . . . A . . . . . N . R . . . . . I . G A . . . . .
A/duck/Hong Kong/202/77 . . . . . I . . . . .
A/duck/Hong Kong/3461/99 . . . . . A . . . . . R . . . . .
A/duck/Korea/S17/03 . . . . . S . . . . . R . . . . . I . . . . .
A/duck/Taiwan/0526/72 . . . . .
A/gray teal/Australia/1/1979 . . . . . K . . . . . I . . . . . G . . . . . H . . . . . A . . . . . L . . . . .
A/mallard/California/6695/2009 . . . . . A . . . . .
A/mallard/Jiangxi/7787/2003 . . . . . A . . . . .
A/mallard/Sweden/30/2005 . . . . .
```

```
Consensus sequence IGSGERVERFEMFPKSTWAGVDTSSGVNTACPSYTSGSSFYRNLLWIIKTKSAAYPVIK
A/Taiwan/2/2013 . . . . . R . . . . . I D . . . . . V . . . . . V . . . . . D . . . . . T . . . . .
A/Chicken/Taiwan/A2837/2013 . . . . . R . . . . . I D . . . . . V . . . . . V . . . . . D . . . . . T . . . . .
A/chicken/Taiwan/TC135/2010 . . . . . R . . . . . I D . . . . . V . . . . . V . . . . . E . . . . .
A/chicken/Taiwan/ch1006/04 . . . . . R . . . . . I D . . . . . V . . . . . V . . . . . E . . . . .
A/chicken/Taiwan/0305/04 . . . . . G . . . . . N . . . . . D . I . . . . . V . . . . . N . . . . .
A/chicken/Taiwan/0107/02 . . . . . R . . . . . S . I D . . . . . V . . . . . V . . . . . E . . . . .
A/chicken/Taiwan/2838N/00 . E . . . . . R . . . . . D . I D . . . . . V . . . . . G . . . . . E . . . . .
A/chicken/Taiwan/SP1/00 . . . . . I . . . . . V . . . . . N . . . . . N . . . . .
A/chicken/Taiwan/0705/99 . . . . . I . . . . . V . . . . . N . . . . .
A/chicken/Taiwan/G2/87 . . . . . K . . . . . N . . . . . G . . . . .
A/chicken/Taiwan/0222/02 . . . . . R . . . . . I D . . . . . V . . . . . V . . . . . E . . . . .
A/chicken/Taiwan/67/2013 . . . . . R . . . . . I D . . . . . V . . . . . V . . . . . D . . . . . T . . . . .
A/chicken/Taiwan/G23/87 . . . . . R . . . . . K . . . . . N . . . . . G . . . . .
A/chicken/Taiwan/ch1006/04 . . . . . R . . . . . I D . . . . . V . . . . . V . . . . . E . . . . .
A/duck/Guangxi/3333/2006 . . . . . T . . . . . K . . . . . H . . . . .
A/duck/Guangxi/GXd-5/2010 . . . . . N . . . . . S S . . . . . L G N . P . S . . . . . L . . . . . S E . . . . . R . . . . .
A/duck/Hong Kong/202/77 . . . . . K . . . . . N . . . . . M . . . . . P . . . . .
A/duck/Hong Kong/3461/99 . . . . . R . . . . . H . . . . . P . . . . .
```

A/duck/Korea/S17/03K.....N.....K.....-.....S.....
 A/duck/Taiwan/0526/72K.....-.....N.....G.....
 A/gray teal/Australia/1/1979R.....Q.....T.....G.....K.....S.....N.....R.....
 A/mallard/California/6695/2009T.....N.....R.....-.....S.....
 A/mallard/Jiangxi/7787/2003T.....K.....-.....N.....
 A/mallard/Sweden/30/2005K.....-.....N.....

170 180 190 200 210 220
 Consensus sequence TYNNTGNQPILYFWGVHHPDNEQNTLYGSGDRYVRMGTESMNFAKSFEIAARPAVNGQ
 A/Taiwan/2/2013T.....L.....TV.DN.....K.....
 A/Chicken/Taiwan/A2837/2013T.....TV.DN.....K.....
 A/chicken/Taiwan/TC135/2010TV.DN.....
 A/chicken/Taiwan/ch1006/04TV.DN.....
 A/chicken/Taiwan/0305/04S.GL.ES.....E..V.....
 A/chicken/Taiwan/0107/02TV.DN.....
 A/chicken/Taiwan/2838N/00MV.....
 A/chicken/Taiwan/SP1/00N.GL.D.....E..I.....
 A/chicken/Taiwan/0705/99N.GV.D.....E.....D.....E..V.....
 A/chicken/Taiwan/G2/87K.....
 A/chicken/Taiwan/0222/02TV.D.....
 A/chicken/Taiwan/67/2013T.....TV.DN.....K.....
 A/chicken/Taiwan/G23/87K.....
 A/chicken/Taiwan/ch1006/04TV.DN.....
 A/duck/Guangxi/3333/2006S.....K.....G.....
 A/duck/Guangxi/GXd-5/2010F.....DKS.....V.T.EA.....N.....R.....
 A/duck/Hong Kong/202/77S.....S.....
 A/duck/Hong Kong/3461/99S.....S.....I.....V.....
 A/duck/Korea/S17/03S.P.....G.....
 A/duck/Taiwan/0526/72K.....
 A/gray teal/Australia/1/1979S.S.....
 A/mallard/California/6695/2009
 A/mallard/Jiangxi/7787/2003
 A/mallard/Sweden/30/2005

230 240 250 260 270 280
 Consensus sequence RGRIDYYWSVLKPGETLNVESNGNLIAPWYAYKRVSTNKKGAVFKSNLPIENC DATCQTI
 A/Taiwan/2/2013S.....R.....D.....
 A/Chicken/Taiwan/A2837/2013S.....R.....D.....
 A/chicken/Taiwan/TC135/2010S.....D.....
 A/chicken/Taiwan/ch1006/04S.....D.....
 A/chicken/Taiwan/0305/04H.....
 A/chicken/Taiwan/0107/02S.....
 A/chicken/Taiwan/2838N/00S.....
 A/chicken/Taiwan/SP1/00G.....A.....
 A/chicken/Taiwan/0705/99S.....S.....K.....
 A/chicken/Taiwan/G2/87S.....S.....K.....
 A/chicken/Taiwan/0222/02S.....R.....K.....D.....
 A/chicken/Taiwan/67/2013S.....R.....K.....D.....
 A/chicken/Taiwan/G23/87S.....D.....
 A/chicken/Taiwan/ch1006/04S.....D.....
 A/duck/Guangxi/3333/2006I.....N.....
 A/duck/Guangxi/GXd-5/2010F.....I.....R.....NKDSN.....I.R.....T.....
 A/duck/Hong Kong/202/77F.....N.....
 A/duck/Hong Kong/3461/99D.....
 A/duck/Korea/S17/03R.....N.....T.....
 A/duck/Taiwan/0526/72S.....S.....K.....
 A/gray teal/Australia/1/1979I.....I.S.N.....I.....D.....T.A...V.....
 A/mallard/California/6695/2009N.....
 A/mallard/Jiangxi/7787/2003N.....
 A/mallard/Sweden/30/2005R.....N.R.....S.....

290 300 310 320 1 10
 Consensus sequence AGVLR TNKTFQNV SPLWIG ECPKYVKSES LRLATGLRNVPQIETRGLFGAIAGFIEGGWT
 A/Taiwan/2/2013T.....A.....I.....
 A/Chicken/Taiwan/A2837/2013K.....A.....I.....
 A/chicken/Taiwan/TC135/2010K.....V.....I.....
 A/chicken/Taiwan/ch1006/04A.....I.....
 A/chicken/Taiwan/0305/04V.....I.....I.....
 A/chicken/Taiwan/0107/02I.....
 A/chicken/Taiwan/2838N/00E.....I.LR.....
 A/chicken/Taiwan/SP1/00K.....I.....K.....I.....

	140	150	160	170
Consensus sequence	GCFEFWHKCDNECIESVKNGTYDYPKYQDES	KLNRQEI	E	
A/Taiwan/2/2013M.....K.....G.....	
A/Chicken/Taiwan/A2837/2013M.....K.....G.....	
A/chicken/Taiwan/TC135/2010R.....	
A/chicken/Taiwan/ch1006/04E.....G.....
A/chicken/Taiwan/0305/04KV.....
A/chicken/Taiwan/0107/02G.....
A/chicken/Taiwan/2838N/00G.....
A/chicken/Taiwan/SP1/00HK.....
A/chicken/Taiwan/0705/99HK.....
A/chicken/Taiwan/G2/87N.....G.....
A/chicken/Taiwan/0222/02E.....G.....
A/chicken/Taiwan/67/2013M.....K.....G.....
A/chicken/Taiwan/G23/87HK.....
A/chicken/Taiwan/ch1006/04E.....G.....
A/duck/Guangxi/3333/2006R.....
A/duck/Guangxi/GXd-5/2010M.....T.....R.....K.....
A/duck/Hong Kong/202/77
A/duck/Hong Kong/3461/99
A/duck/Korea/S17/03
A/duck/Taiwan/0526/72
A/gray teal/Australia/1/1979D.....N.....D.....
A/mallard/California/6695/2009
A/mallard/Jiangxi/7787/2003
A/mallard/Sweden/30/2005

B

H6N1_A/Taiwan/2/2013
H1N1_A/Brevig_Mission/1/1918
H1N1_A/California/04/2009
H2N2_A/Singapore/1/1957
H2N2_A/Japan/305/1957
H3N2_A/Hong_Kong/1/1968
H5N1_A/Vietnam/CL12/2004
H5N1_A/Indonesia/5/2005
H7N9_A/Shanghai/02/2013
H7N9_A/Hangzhou/1/2013

H6N1_A/Taiwan/2/2013
H1N1_A/Brevig_Mission/1/1918
H1N1_A/California/04/2009
H2N2_A/Singapore/1/1957
H2N2_A/Japan/305/1957
H3N2_A/Hong_Kong/1/1968
H5N1_A/Vietnam/CL12/2004
H5N1_A/Indonesia/5/2005
H7N9_A/Shanghai/02/2013
H7N9_A/Hangzhou/1/2013

H6N1_A/Taiwan/2/2013
H1N1_A/Brevig_Mission/1/1918
H1N1_A/California/04/2009
H2N2_A/Singapore/1/1957
H2N2_A/Japan/305/1957
H3N2_A/Hong_Kong/1/1968
H5N1_A/Vietnam/CL12/2004
H5N1_A/Indonesia/5/2005
H7N9_A/Shanghai/02/2013
H7N9_A/Hangzhou/1/2013

H6N1_A/Taiwan/2/2013
H1N1_A/Brevig_Mission/1/1918
H1N1_A/California/04/2009
H2N2_A/Singapore/1/1957
H2N2_A/Japan/305/1957
H3N2_A/Hong_Kong/1/1968
H5N1_A/Vietnam/CL12/2004
H5N1_A/Indonesia/5/2005
H7N9_A/Shanghai/02/2013
H7N9_A/Hangzhou/1/2013

H6N1_A/Taiwan/2/2013
H1N1_A/Brevig_Mission/1/1918
H1N1_A/California/04/2009
H2N2_A/Singapore/1/1957
H2N2_A/Japan/305/1957
H3N2_A/Hong_Kong/1/1968
H5N1_A/Vietnam/CL12/2004
H5N1_A/Indonesia/5/2005
H7N9_A/Shanghai/02/2013
H7N9_A/Hangzhou/1/2013

H6N1_A/Taiwan/2/2013
H1N1_A/Brevig_Mission/1/1918
H1N1_A/California/04/2009
H2N2_A/Singapore/1/1957
H2N2_A/Japan/305/1957
H3N2_A/Hong_Kong/1/1968
H5N1_A/Vietnam/CL12/2004
H5N1_A/Indonesia/5/2005
H7N9_A/Shanghai/02/2013
H7N9_A/Hangzhou/1/2013

10 20 30 40 50
KSDKICIGYHANNSTTQVDLLEKNVTVTHSVLLENQKEKRFCKIMNKA
NADTICIGYHANNSTDTVDTVLEKNVTVTHSVNLEEDSHNGKLCCKLKGIA
NADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLEEDKHNGKLCCKLRGVA
RGDQICIGYHANNSTEKVDITLERNVTVTHAKDILEKTHNGKLCCKLNGIP
RGDQICIGYHANNSTEKVDITLERNVTVTHAKDILEKTHNGKLCCKLNGIP
STATLCLGHHAVPNGTLVKTITDDQIEVTNATELVQSSSTGKICNNP-HR
KSDQICIGYHANNSTEQVDITIMEKNVTVTHAQDILEKTHNGKLCCKLDGVK
KSDQICIGYHANNSTEQVDITIMEKNVTVTHAQDILEKTHNGKLCCKLDGVK
NADKICLGHHAHSVNGTKVNTLTERGVEVNVNATEVVERTNIPRICSKG-KR
NADKICLGHHAHSVNGTKVNTLTERGVEVNVNATEVVERTNIPRICSKG-KR

60 70 80 90 100
PLDLKDCITIEGWILGNPKCDLLLDGQSWSYIVERPNAQNGICYPGVLNLEL
PLQLGKCNIAAGWLLGNPECDLLLTASSWSYIVETSNSENGTCYPGDFIDY
PLHLGKCNIAAGWLLGNPECELSSTASSWSYIVETPSSDNGTCYPGDFIDY
PLELGDCSIAGWLLGNPECDRLLSVPEWSYIMEKENPRDGLCYPGSFNDY
PLELGDCSIAGWLLGNPECDRLLSVPEWSYIMEKENPRDGLCYPGSFNDY
ILDGIDCTLIDALLGDPHCDVFO-NETWDLFVERSKAFSN-CYYPDVPDY
PLILRDCSVAGWLLGNPMCEDEFINVPWSYIVEKANPVNDLCYPGDFDDY
PLILRDCSVAGWLLGNPMCEDEFINVPWSYIVEKANPTNDLCYPGSFNDY
TVDLGQCGLLGTITGPPQCDQFL-EFSADLIERREGSDV-CYPGKVFNE
TVDLGQCGLLGTITGPPQCDQFL-EFSADLIERREGSDV-CYPGKVFNE

110 120 130 140 150
EELKAFIGSGERVERFEMFP-KSTWAGVDTSRGVTNACPSYITIDSSFYRN
EELREQLSSVSSFEKFEIFPKTSSWPNHETTKGVTAAACSYAGA-SSFYRN
EELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTAAACPHAGA-KSFYKN
EELKHLSSVVKHFKEVKILP-KDRWTQH-TTTGGSRACAVSGN-PSFFRN
EELKHLSSVVKHFKEVKILP-KDRWTQH-TTTGGSRACAVSGN-PSFFRN
ASLRSLVASSGTLEFITEGF---TWTGV-TONGGSSNACKRPGG-SGFFSR
EELKHLSSRINHFEKIQIIP-KSSWSSHEASLGVSSACPYQGG-KSSFFRN
EELKHLSSRINHFEKIQIIP-KSSWSDHEASSGVSSACPYLGS-PSFFRN
EALRQILRESGGIDKEAMGF---TYSGI-RTNGATSACRR-SG-SSFYAE
EALRQILRESGGIDKEAMGF---TYSGI-RTNGATSACRR-SG-SSFYAE

160 170 180 190
LVWIVKTDG--ATYPIKGTYNNTGTQPIILYFWGVHHPDITVQDNLYGS
LLWLTKK-G--SSYPKLSKSYVNNKGEVLLVWGVHHPPTGTDDQSLYQN
LIWLTKK-G--NSYPKLSKSYINDKGEVLLVWGVHHPSTADQSLYQN
MVWLTKK-G--SNYPVAKGSYNNTSGEQMLI IWGVHHPNDEKEQRTLYQN
MVWLTKK-G--SDYPVAKGSYNNTSGEQMLI IWGVHHPIDVETRTLYQN
LNWLTKK-G--STYPLVNVTMPNNDNFDKLYI IWGVHHPSTNQEQTSLYVQ
VVWLTKK-N--STYPTIKRSYNNNTQEDLLVMWGIHHPNDAAEQTKLYQN
VVWLTKK-N--STYPTIKRSYNNNTQEDLLVWGIHHPNDAAEQTRLYQN
MKWLLSN-TDNAAFPQMTKSYKNTRKSPALIVWGIHHSVSTAEQTKLYGS
MKWLLSN-TDNAAFPQMTKSYKNTRKSPALIVWGIHHSVSTAEQTKLYGS

200 210 220 230 240
GDKYVRMGTESMNFSAKSPETIARPAVNGQSRIDYVWVLRPGETLNVES
ADAYVSVGSSKYNNRFTPEIARPKVRDQAGRMNYWYTLLEPGDITFEA
ADTYVSVGSSRYSKKFKPEIARPKVRDQAGRMNYWYTLLEPGDKITFEA
VGTYSVGTSTLNKRSTPDIATRPKVNGLGRMEFSWTLDDMWDITNFES
VGTYSVGTSTLNKRSTPDIATRPKVNGLGRMEFSWTLDDMWDITNFES
ASGRVTVSTRSQTIIPNIGSRPWRGLSSRSISYWTIVKPGDVLVINS
PTYYSVGTSTLNQRLVPRDIATRPKVNGLGRMEFFWTILKPNDAINFES
PTYYSIGTSTLNQRLVPRDIATRPKVNGLGRMEFFWTILKPNDAINFES
GNKLVTVGSSNYQQSFVPSPGARPVNGLSGRIDFHWLMLNPNDITVTFSS
GNKLVTVGSSNYQQSFVPSPGARPVNGLSGRIDFHWLMLNPNDITVTFSS

250 260 270 280 290
NGNLIAPWYAYKIVKSTNKKGAVFKSDLPIEN-CDATCQITGVLRNKTFT
TGNLIAPWYAFALNRG-SGSGIITSDAPVHD-CNTKQCTPHGAINSSLFP
TGNLIVPRYAFAMERN-AGSGIISDTPVHD-CNTKQCTPHGAINSSLFP
TGNLIAPYGFYKISKR-GSSGIMKTEGTLEN-CETKQCTPLGAINSSLFP
TGNLIAPYGFYKISKR-GSSGIMKTEGTLEN-CETKQCTPLGAINSSLFP
NGNLIAPRYGFKMRTG--KSSIMRSDAPIDT-CISECITPNGSIPNDKFP
NGNFIAPYAYKIVK-KGSTIMKSELEYGN-CNTKQCTPMGAINSSMPF
NGNFIAPYAYKIVK-KGSTAIMKSELEYGN-CNTKQCTPMGAINSSMPF
NGAFIAPDRASFLRGK---SMGIQSGVQVDANCEGDCYHSGGTIISNLPF
NGAFIAPDRASFLRGK---SMGIQSGVQVDANCEGDCYHSGGTIISNLPF

Figure S1, related to Figure 1. Sequence Alignment of Taiwan2 H6 HA with Other H6 HAs and with HA Subtypes that Have Infected Humans.

(A) Sequence alignment of representative full-length H6N1 HA sequences deposited in NCBI influenza virus database. The Taiwan 2 HA sequence is marked in blue. Residues that contribute to receptor binding are highlighted in bold (H3 numbering). (B) Alignment of the sequence of H6 HA of the Taiwan2 virus isolated from a human (in blue) with sequences of HAs of human pandemic viruses (H1N1, H2N2 and H3N2) and avian zoonotic viruses H5N1 and H7N9. Residues that contribute to receptor binding are highlighted in bold (H3 numbering). The alignment scores are shown in Table S2.

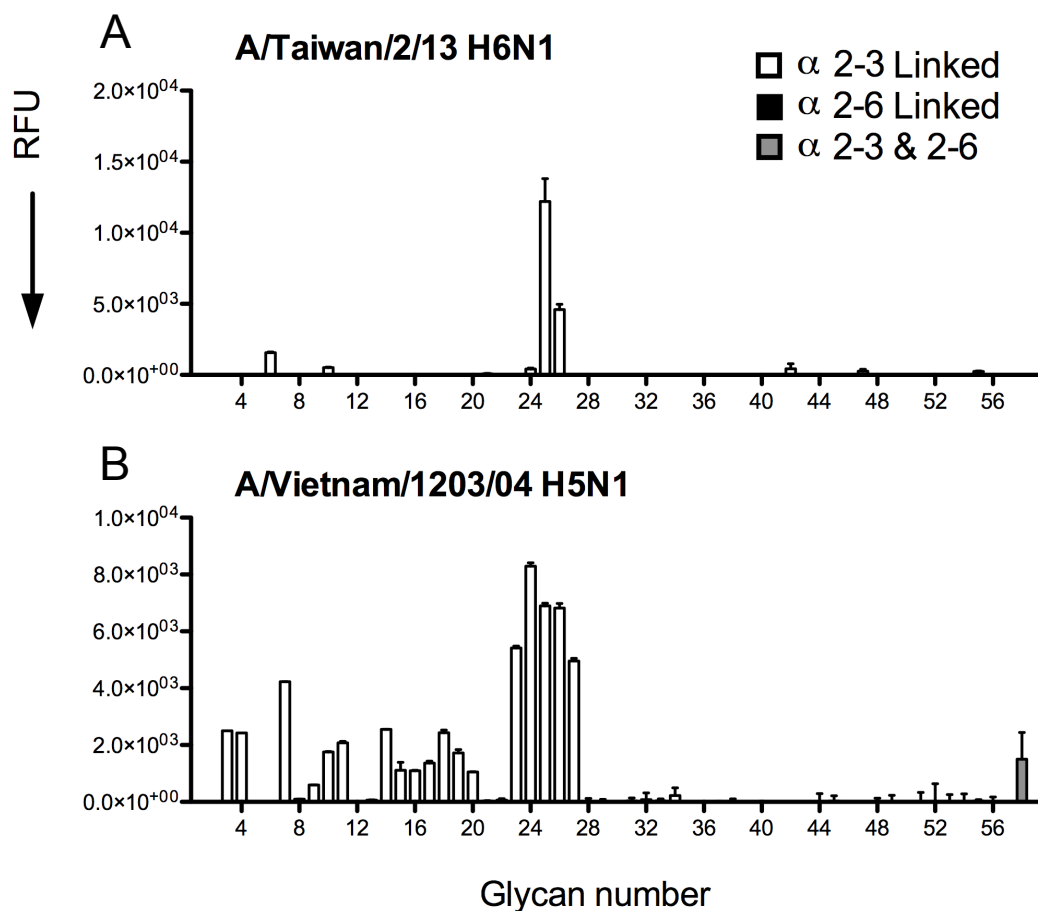
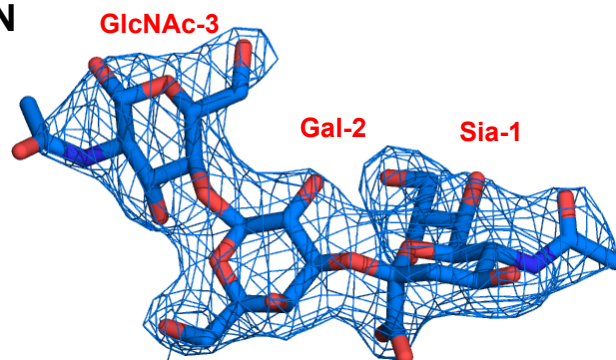


Figure S2, related to Figure 2. Receptor Binding Specificity of Taiwan2 H6N1 HA

(A-B) Glycan microarray analysis of recombinant Taiwan2 H6 (A) and A/Vietnam/1203/04 H5N1 (B) HAs, expressed in mammalian cells, indicates specific binding to α 2-3 sialosides. The mean signal and standard error were calculated from six independent replicates on the array. Nonsialylated control glycans are glycans 1 and 2 on the x axis, α 2-3 linked sialosides are in white bars (glycans 3 to 35), α 2-6 linked sialosides in black bars (glycans 36 to 56), and mixed biantennary glycans containing both α 2-3 and α 2-6-linked sialylated glycans are in gray bars (glycans 57 and 58). Glycans imprinted on the array are listed in Table S3.

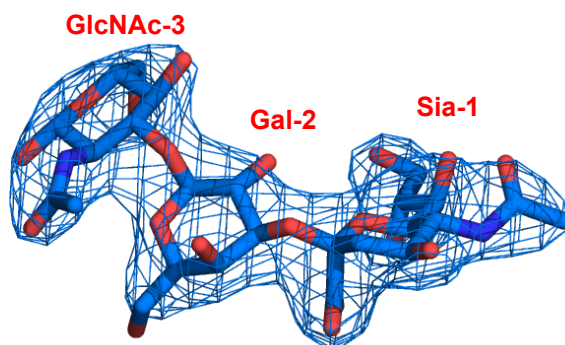
A

3'-SLN



B

LSTa



C

6'-SLN

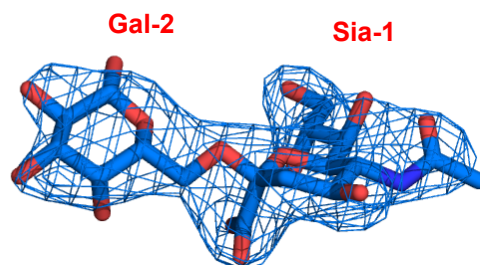


Figure S3, related to Figures 3 and 4. Electron Density Maps for the Receptor Analogs in Taiwan2 H6 HA Crystal Structures

(A-C) Electron density maps (2Fo-Fc) of avian receptor analogues 3'-SLN (A) and LSTa (B) and human receptor analog 6'-SLN (C) in crystal structures of Taiwan2 H6 wild type HA. Density corresponding only to the receptor analog is shown at a 1 σ level.

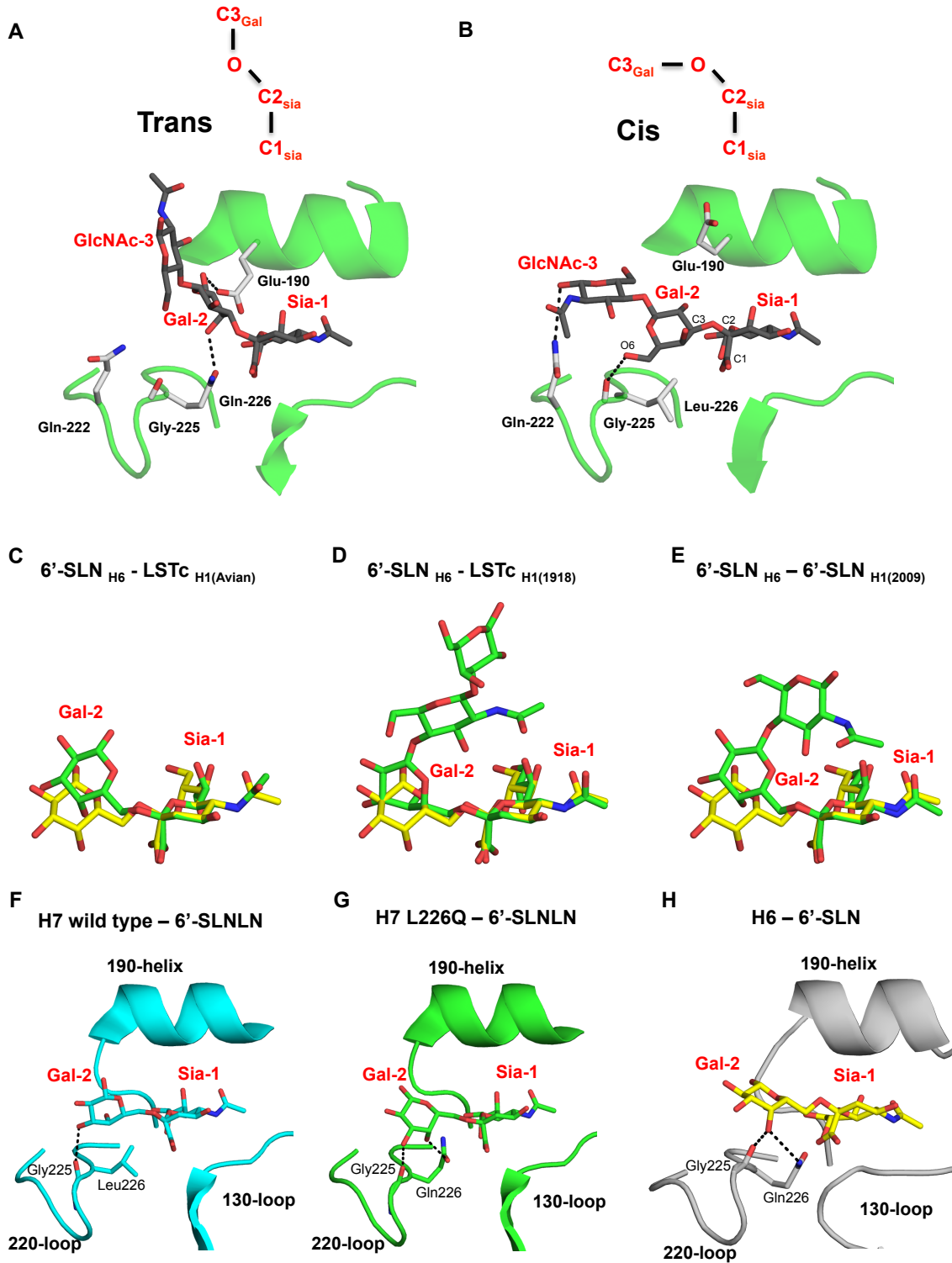


Figure S4, related to Figures 3 and 4. Conformation of Avian and Human Receptor Analogs in HA Structures

(A and B) Binding of avian receptor analog 3'-SLN to the RBS of the H7 HA subtype in both cis and trans conformations with respect to the Sia-Gal linkage. Avian analogs (3'-SLN and LSTa) can bind to the RBS in two conformations, trans (A, PDB 4BSI) and cis (B, PDB 4BSD), resulting in two modes of binding. (A) In the trans conformation, Gal-2 hydrogen bonds with the side chains of Gln226 and Glu190. (B) In the cis conformation, Gal-2 hydrogen bonds to the main chain of Gly225 and makes hydrophobic interactions with the Leu226 side chain. In addition, GlcNAc-3 hydrogen bonds with Gln222. The H7 complexes are used as representative of the two different binding conformations of avian receptors in the HA RBS.

(C-E) Superposition of 6'-SLN receptor analog from the H6 HA complex (yellow) compared to LSTc or 6'-SLN human analogs from structures with HAs from (A) avian H1N1 (PDB 3HTQ), (B) 1918 pandemic H1N1 (PDB 2WRG) and (C) 2009 pandemic H1N1 (PDB 3UBN), which are colored in green as representative of other HA subtypes. The superposition indicates conformational changes arising from rotation around the linkage between Sia-1 and Gal-2 (ϕ angle changes from ~ 60 - 70° in H1 structures to $\sim 120^\circ$ for H6). The superposition was done on Sia-1.

(F-H) Structures of complexes of AH H7N9 wild type HA (F, cyan, PDB entry 4KON) and AH H7N9 L226Q mutant HA (G, green, PDB entry 4LKK) with the human analog 6'-SLN (yellow) indicating changes in the linkage of Sia-1 and Gal-2 (from 63° in the wild type to 21° in the mutant). In the mutant complex, Gal-2 hydrogen bonds with Gly225 and Gln226 similar to the interaction between 6'-SLN Gal-2 and H6 HA (F).

Table S1, related to Figure 1. Data Collection and Refinement Statistics for H6 HA and Glycan Complexes

Data collection	Apo	3'-SLN	LSTa	6'-SLN
Beamline	SSRL11-1	APS-23	APS-23	APS-23
Wavelength (Å)	0.97945	1.03320	1.03320	1.03320
Space Group	C2	C2	C2	C2
Unit cell parameters (Å)	a=184.4, b=98.9, c=132.8 $\beta=126.1^\circ$	a=185.5, b=99.2, c=133.4 $\beta=126.3^\circ$	a=185.7, b=99.2, c=133.4 $\beta=126.4^\circ$	a=185.5, b=98.9, c=133.6 $\beta=126.6^\circ$
Resolution range (Å)	50.00-2.50 (2.54-2.50) ^a	50.00-2.35 (2.39-2.35)	50.00-2.45 (2.49-2.45)	50.00-2.25 (2.29-2.25)
Observations	240,934	238,063	249,206	290,644
Unique reflections	66,521 (3186)	78,465 (3652)	71,296 (3296)	90,738 (4375)
Completeness (%)	98.3 (94.4)	98.2 (92.7)	99.1 (93.0)	97.5 (94.9)
I/ σ (I)	11.9 (1.4)	19.7 (1.4)	16.6 (2.1)	21.2 (3.6)
CC _{1/2}	0.78	0.73	0.77	0.87
R _{sym} ^b	0.11 (0.53)	0.12 (0.75)	0.12 (0.51)	0.10 (0.37)
R _{pim} ^c	0.06 (0.28)	0.06 (0.44)	0.06 (0.29)	0.05 (0.20)
Redundancy	3.6 (3.2)	3.0 (2.3)	3.5 (2.9)	3.2 (3.1)
Refinement Statistics				
Resolution (Å)	50.00-2.50 (2.54-2.50)	50.00-2.35 (2.39-2.35)	50.00-2.45 (2.49-2.45)	50.00-2.25 (2.29-2.25)
No. reflections	66,484	78,319	71,268	90,698
R _{cryst} ^d /R _{free} ^e	0.207/0.257	0.230/0.274	0.210/0.253	0.208/0.253
No. atoms				
Protein	11,897	11,909	11,894	11,903
Carbohydrate	84	130	130	116
Water	247	111	140	422
Wilson B (Å ²)	45.2	55.3	42.2	33.5
Average B value (Å ²)				
Overall	41.7	63.5	47.3	40.4
Ligand		78.5	74.3	69.5
R.m.s deviation from ideal geometry				
Bond length (Å)	0.016	0.015	0.011	0.011
Bond angle (°)	1.5	1.5	1.6	1.6
Ramachandran Plot (%)^f				
Favored	97.3	96.7	96.3	97.0
Outliers	0.1	0	0	0
PDB ID	4XKD	4XKE	4XKF	4XKG

^aParentheses refer to outer shell statistics.

^b $R_{\text{sym}} = \sum_{hkl} \sum_i |I_{hkl,i} - \langle I_{hkl} \rangle| / \sum_{hkl} \sum_i I_{hkl,i}$, where $I_{hkl,i}$ is the scaled intensity of the i^{th} measurement of reflection h, k, l , and $\langle I_{hkl} \rangle$ is the average intensity for that reflection.

^c $R_{\text{pim}} = \sum_{hkl} (1/(n-1))^{1/2} \sum_i |I_{hkl,i} - \langle I_{hkl} \rangle| / \sum_{hkl} \sum_i I_{hkl,i}$, where n is the redundancy.

^d $R_{\text{cryst}} = \sum_{hkl} |F_o - F_c| / \sum_{hkl} |F_o| \times 100$, where F_o and F_c are the observed and calculated structures factors.

^e R_{free} was calculated as for R_{cryst} , but on a test set of 5% of the data excluded from refinement.

^fCalculated using MolProbity (Davis et al., 2007)

Table S2, related to Figure 1. Comparison of H6 HA with Other Subtypes

HA strain name	PDB code	RMSD values of C α atoms (Å)				Alignment score (%) ^b
		Monomer	HA1	HA2	RBS subdomain ^a	
H1N1_A/California/04/2009	3AL4	1.1 (480) ^c	1.0 (319)	0.5 (162)	0.6 (142)	58
H2N2_A/Japan/305/1957	3KU5	0.8 (479)	0.8 (310)	0.5 (169)	0.7 (134)	60
H3N2_A/Hong_Kong/1/1968	4FNK	2.4 (464)	1.6 (300)	0.9 (163)	0.8 (121)	42
H5N1_A/Vietnam/CL17/2004	2FKO	1.3 (490)	1.5 (390)	0.8 (151)	0.8 (145)	59
H7N9_A/Shanghai/02/2013	4N5J	2.5 (480)	2.0 (312)	1.0 (167)	0.8 (126)	42

^a The receptor binding subdomain (RBS subdomain) is defined as residues 117-265 of the HA1 subunit (H3 numbering).

^b Calculated for the HA sequence (Figure S1) using ClustalW2 (ENBL-EBI) (Larkin et al., 2007).

^c Number in parentheses is the number of residues included in the comparison.

Table S3, related to Figure 2. Glycans Imprinted on the Microarray. Purple diamonds represent NeuAc, yellow circles represent Gal, blue circles represent Glc, green circles represent Man, yellow squares represent GalNAc, blue squares represent GlcNAc, red triangles represent Fuc and light blue diamonds represent NeuGc.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Structural Determination of the H6 HA. Diffraction data were collected at the Advanced Photon Source (APS) and at the Stanford Synchrotron Radiation Lightsource (SSRL). Data were integrated and scaled using HKL2000 (Otwinowski and Minor, 1997). The initial H6 apo structure was solved by molecular replacement method using Phaser (McCoy et al., 2005) with an H1 HA structure (PDB entry 4M4Y) as a search model. The H6 HA apo structure was then used as the starting model for structure determination of the H6 HA-glycan complex structures. Structure refinement was carried out in Phenix (Adams et al., 2002) and model building with COOT (Emsley and Cowtan, 2004). Final data collection and refinement statistics are summarized in Table S1.

Expression and Purification of H6 and H5 HAs in Baculovirus Expression System for Glycan Microarray Analyses. The HA expression was similar to that used for the crystallization experiments. The expressed HA0s were purified through a His-tag affinity purification step and dialyzed against 20 mM Tris-HCl, 50 mM NaCl, pH 8.0 overnight at 4 °C. Proteins were concentrated to 1 mg/ml prior to binding assays.

Expression and Purification of H1, H5 and H6 HAs from a Mammalian Expression System for Glycan Microarray Analyses. Codon-optimized H1, H5 and H6 encoding cDNAs (Genscript, USA) of A/Kentucky/07 (Accession; CY028163), A/Vietnam/1203/04 (Accession; EF541403) and A/Taiwan/2/13 (Accession; EPI459855) were cloned into the pCD5 expression as described previously (Xu et al., 2013). The HA proteins were expressed in HEK293S GnTI^{-/-} cells and purified from the cell culture supernatants as described previously (de Vries et al., 2010).

Glycan Microarray Analysis of HAs Expressed in Insect and Mammalian Cells.

Purified, soluble trimeric HA was pre-complexed with horseradish peroxidase (HRP)-linked anti-Strep-tag mouse antibody and with Alexa647-linked anti-mouse IgG (4:2:1 molar ratio) prior to incubation for 15 min on ice in 100 μ l PBS-T, and incubated on the array surface in a humidified chamber for 90 minutes. Slides were subsequently washed by successive rinses with PBS-T, PBS, and deionized H₂O. Washed arrays were dried by centrifugation and immediately scanned for FITC signal on a Perkin-Elmer ProScanArray Express confocal microarray scanner. Fluorescent signal intensity was measured using Imogene (Biodiscovery) and mean intensity minus mean background was calculated and graphed using MS Excel. For each glycan, the mean signal intensity is calculated from 6 replicates spots. The highest and lowest signals of the 6 replicates are removed and the remaining 4 replicates are used to calculate the mean signal, standard deviation (SD), and standard error measurement (SEM). Bar graphs represent the averaged mean signal minus background for each glycan sample and error bars are the SEM value. A list of glycans on the microarray is included in Table S3.

REFERENCES

Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K., and Terwilliger, T.C. (2002). PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.* *58*, 1948-1954.

Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B., 3rd, Snoeyink, J., Richardson, J.S., *et al.* (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* *35*, W375-383.

de Vries, R.P., de Vries, E., Bosch, B.J., de Groot, R.J., Rottier, P.J., and de Haan, C.A. (2010). The influenza A virus hemagglutinin glycosylation state affects receptor-binding specificity. *Virology* *403*, 17-25.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta. Crystallogr. D Biol. Crystallogr.* *60*, 2126-2132.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., *et al.* (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* *23*, 2947-2948.

McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C., and Read, R.J. (2005). Likelihood-enhanced fast translation functions. *Acta Crystallogr. D Biol. Crystallogr.* *61*, 458-464.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Meth. Enzymol.* *276*, 307-326.

Stevens, J., Blixt, O., Tumpey, T.M., Taubenberger, J.K., Paulson, J.C., and Wilson, I.A. (2006b). Structure and receptor specificity of the hemagglutinin from an H5N1 influenza virus. *Science* *312*, 404-410.

Xu, R., de Vries, R.P., Zhu, X., Nycholat, C.M., McBride, R., Yu, W., Paulson, J.C., and Wilson, I.A. (2013). Preferential recognition of avian-like receptors in human influenza A H7N9 viruses. *Science* *342*, 1230-1235.

