Published in final edited form as:
Science. 2010 April 16; 328(5976): 357–360. doi:10.1126/science.1186430.

Structural basis of pre-existing immunity to the 2009 H1N1 pandemic influenza virus

Rui Xu[1,*], **Damian C. Ekiert**[1,*], **Jens C. Krause**[2], **Rong Hai**[3], **James E. Crowe Jr.**[2], and **Ian A. Wilson**[1,4,†]

[1]Dept. of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA

[2]Dept. of Pediatrics and of Microbiology and Immunology, Vanderbilt University Medical Center, Nashville, TN 37232, USA

[3]Dept. of Microbiology, Mount Sinai School of Medicine, New York, NY 10029, USA

[4]Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA

Abstract

The 2009 H1N1 swine flu is the first influenza pandemic in decades. The crystal structure of the hemagglutinin from the A/California/04/2009 H1N1 virus shows that its antigenic structure, particularly within the Sa antigenic site, is extremely similar to human H1N1 viruses circulating early in the 20th century. The co-crystal structure of the 1918 HA with 2D1, an antibody from a survivor of the 1918 Spanish flu that neutralizes both 1918 and 2009 H1N1 viruses, reveals an epitope that is conserved in both pandemic viruses. Thus, antigenic similarity between the 2009 and 1918-like viruses provides an explanation for the age-related immunity to the current influenza pandemic.

Influenza pandemics in humans tend to occur decades apart and infect a large percentage of the human population with significant mortality. In the 20th century, three pandemics were caused by the emergence of different influenza A subtypes that were antigenically divergent from human viruses circulating at the time : 1918 H1N1 ("Spanish flu"), 1957 H2N2 ("Asian flu") and 1968 H3N2 ("Hong Kong flu") (1). Since April 2009, the outbreak of a novel influenza A H1N1 virus (2009 H1N1) in Mexico has spread globally and developed into the first human influenza pandemic in 40 years. The 2009 H1N1 virus has now infected the human population worldwide, and contributed to at least 16,000 deaths as of February 26, 2010 (2).

The influenza virus envelope protein, hemagglutinin (HA), is the principal surface antigen (3) and the most critical component of flu vaccines (4). At the beginning of a flu pandemic, pre-existing immunity to the HA of the newly emerging virus is generally low, guaranteeing a large pool of susceptible hosts for rapid spread and infection of 10 to 40% of the population world-wide. After a new HA becomes fixed in circulating human viruses, it undergoes gradual changes in its antigenic structure, in a process called "antigenic drift", so as to escape

†To whom correspondence should be addressed, wilson@scripps.edu.
*These authors contributed equally to this work.

Supporting Online Material
Materials and Methods
Fig. S1 to S2
Tables S1 to S4
References

recognition by the human immune system. Such drift leads to loss of immunity and is associated with the frequent flu epidemics that occur during interpandemic periods.

The 2009 pandemic virus HA originated from the swine lineage of H1 HAs and closely resembles current circulating H1 viruses in swine (Fig. 1A) (5–7), whereas seasonal human H1 HAs diverged from the swine lineage early in the 20th century (8). Descendants of the human 1918 H1N1 virus continued to circulate until the 1957 pandemic, when they were replaced by H2N2 viruses. H1N1 reappeared in humans in the late 1970s and H1N1 viruses remain a component of seasonal influenza today. In the meantime, H1 viruses continued to circulate in the swine population, remaining antigenically stable and causing only sporadic human infections, including the 1976 outbreak in Fort Dix, New Jersey (9,10).

To understand the structural basis for the high antigenicity of this emerging pandemic virus, we determined the crystal structure of the HA from the H1N1 A/California/04/2009 (CA04) virus (11). Our structural analysis reveals that the 2009 H1 HA shares conserved antigenic epitopes with human and swine H1 viruses from the early 20th century. This observation is supported by the crystal structure of the 1918 H1 HA in complex with a neutralizing antibody that cross-reacts with both pandemic viruses. These structures shed light on the basis of pre-existing immunity against these new H1N1 pandemic viruses in people born early in the 20th century.

The entire ectodomain structure of CA04 HA was determined by molecular replacement at 2.6 Å resolution (table S1). Its overall structure is similar to other H1 HA structures from avian (12), swine (13) and human viruses (13,14). The HA homotrimer comprises a long, extended stem region and a membrane-distal, globular cap that includes the receptor-binding domain and the vestigial esterase domain. Whereas the stem region houses the membrane fusion machinery, the membrane-distal domain mediates cell attachment through receptor binding and displays most of the epitopes for antibody recognition.

The antigenic sites of H1 HA are distributed roughly over four conformational epitopes (Fig. 1, B and C) (15–17); the Sa and Sb sites are proximal to the receptor-binding pocket, the Ca site (Ca1 and Ca2) is at the subunit interface, and the Cb site is within the vestigial esterase domain (Figs. 1B and 2A). The majority of the epitopes contain highly variable, protruding loops that can be accessed readily by approaching antibodies. Although the protein sequences are diverse, the CA04 epitopes maintain highly similar backbone conformations to those in other H1 structures (18, 19), with the only exception being the Cb site, in which a flexible loop adopts variable conformations. This structural conservation facilitates comparisons of antigenic structures at the sequence level.

The 2009 pandemic virus displays distinct antigenic properties from the circulating seasonal H1 viruses (6,20). The previous influenza pandemics of last century all were initiated by antigenic shift: the introduction of a new HA subtype into humans. The 2009 swine flu marks the first time that a distant variant of a current circulating HA subtype has triggered a new pandemic (21). Vaccines for current seasonal flu do not elicit any cross-reactivity in humans against the 2009 pandemic virus, highlighting their considerable antigenic divergence (22, 23). Furthermore, the 2009 pandemic exhibits an unusual pattern of age-related morbidity and mortality reminiscent of the 1918 Spanish flu, as it disproportionately affects children and young adults (ages 4–25) (24). The infection rate decreases with age, with the lowest occurrence in the population 65 years and older (25–30). The low infection rate in the elderly, the common victims of seasonal flu, suggests some pre-existing immunity to the 2009 pandemic viruses (22).

The CA04 HA structure reveals the molecular basis for its antigenic properties. Compared with seasonal H1 HAs, substantial amino-acid differences are found in all four antigenic regions

(Fig. 2, B to D). However, among other human viruses, the HA from the 1918 influenza pandemic (A/South Carolina/1/1918, SC1918) is a remarkably close relative of CA04 with only 20% amino-acid differences in the antigenic sites. Moreover, these differences are restricted mainly to the Ca region, with high conservation of the Sa, Sb and Cb epitopes. The largely conserved antigenic surface between SC1918 and CA04 suggests a potential for substantial antibody cross-reactivity between these two pandemic strains. Indeed, serological tests suggest that individuals who likely experienced the 1918 Spanish flu carry the highest titers of neutralizing antibodies against the novel 2009 H1N1 viruses amongst all age groups (20, 22). In contrast, the antigenic sites of all known human H1 HAs from the 1930s to the present are highly divergent from CA04. HAs from A/Puerto Rico/8/1934 (PR8/34) and A/Brisbane/59/2007 (Brisbane07) differ from CA04 by 46% and 50%, respectively, in the residues corresponding to the antigenic sites. Comparison of all available H1 HA sequences with CA04 reveal a steady increase in divergence during the 1930s to about 50% residue differences in the HA antigenic sites in the 1940s and have remained stable at that level since (Fig. 2E and table S2). Thus, exposure to early H1N1 viruses that present conserved patches of antigenic surface would explain production of CA04 cross-reactive antibodies, whereas increased drift in later strains would make generation of such antibodies in the population increasingly rare.

Variation in glycosylation is also used by influenza and other viruses to interfere with surveillance by the host immune system. Acquisition of a glycosylation site masks the protein surface from antibody recognition as the glycans themselves are host-derived and, hence, considered as ‘self’ by the immune system (31). Whereas SC1918 HA lacks any N-glycosylation sites within or near Sa, human H1 HAs have gradually acquired up to 3 such sites in the relatively conserved Sa region from 1930 to 2007 (Fig. 2D and F,table S2) (32), paralleling the evolution of protein sequence differences. What is most significant is that CA04 (Fig. 2B), like SC1918, does not have any glycosylation in or around the Sa site and, hence, the epitope is exposed for antibody recognition.

In light of the relative conservation of the Sa site in human H1 viruses, Sa-specific antibodies are potentially the major underlying basis of age-related immunity to the 2009 H1N1 virus. Recently, antibody 2D1, isolated from elderly survivors of the 1918 pandemic (33), showed high affinity for the pandemic 2009 H1 HA (CA04) (Fig. 3C), and cross-neutralized the 2009 H1N1 viruses *in vivo* (34). To understand the basis of 2D1 cross-reactivity against these two pandemic viruses, we determined the crystal structure of the 2D1 Fab in complex with the 1918 HA (table S1).

2D1 Fab recognizes an epitope at the apex of the receptor-binding domain (Fig. 3A). The footprint of the Fab on the HA largely coincides with the previously defined Sa region, but extends its boundaries (15,16) (Fig. 3B). Binding of 2D1 to SC1918 HA buries a total surface area of 1506 $Å_2$ (743 $Å^2$ on HA and 763 $Å_2$ on the Fab) with typical, heavy chain-dominant binding (~63% V_H, 37% V_L). The epitope is conformational and consists primarily of HA1 residues 125C-129 and 157–169 (Fig. 3B and table S3). Heavy chain complementarity determining region 1 (HCDR1) adopts an unusual, open loop conformation when compared with other Fab structures that have identical HCDR1 sequences (fig. S2). As a result, HCDR1 makes few contacts with HA and is likely very flexible in solution (35). Instead, HCDRs 2 and 3 dominate the interaction, along with light chain CDRs (LCDRs) 1 and 3. The heavy and light chains recognize separate, nearly discontinuous surfaces on HA, with a small cavity being formed at the junction between V_H, V_L, and the HA. Of the 18 HA residues contacting Fab 2D1, 11 lie within the previously defined Sa site, accounting for 3/3 salt bridges, 7/13 hydrogen bonds, and 92/123 van der Waals contacts. Indeed, the Sa site comprises the centerpiece of the antibody-binding surface for 2D1, with the remaining contact residues lining its borders along the periphery of the Fab-HA interface (Fig. 3B).

The antibody-binding site for 2D1 is well conserved in SC1918 and CA04, but not in seasonal HAs (table S3). A conservative variation at position 169 in HA1 (Val in SC1918 and Ile in CA04) and an S159N substitution are the only differences within the 2D1 footprint between these two pandemic viruses and do not appear to adversely affect 2D1 binding (Figs. 3B and table S3). Val169 lies on the edge of the 2D1 epitope and makes a single van der Waals contact with HCDR1. Ser159 also lies along the periphery and makes minor contacts with the highly flexible HCDR1 and a single contact with HCDR3. The 1934 PR8 isolate carries 3 amino-acid variations from SC1918 in that region, all of which lie within the Sa site, including K166N, which would abolish the salt bridge with Asp93 in the 2D1 light chain. Mutations of Lys166 significantly decrease binding affinity for antibody 2D1 (table S3). Furthermore, escape mutants for 2D1 were also selected at position 166 in 2009 H1N1, as well as in SC1918 human and 1930 swine viruses (33,34). Brisbane07 shows an even more diverse antigenic surface. Seven amino-acid differences from SC1918 (three of which map to Sa), along with two potential N-glycosylation sites in the center of the epitope, reveal why 2D1 does not cross-react with the current seasonal viruses.

This increased divergence in the antigenic surface of human H1N1 viruses, especially the Sa site, directly correlates with decreasing antibody cross-reactivity to CA04 among serum donors when grouped by age (22,23). Thus, the highest titers of cross-reactive antibodies to CA04 are elicited in the population born before the 1930-40s, whereas individuals born after ~1940 generally have lower titers and lack cross-protection from previous exposure to influenza viruses, except for US adults immunized with the swine A/NJ/76 vaccine (22). Current vaccine guidelines place those 65 and above into the “less at risk” group with lower priority for vaccination (25), consistent with the extent of the antigenic variation observed in the H1 HA structures since 1918. Cross-reactive antibodies elicited by infection with H1N1 viruses in the first few decades of the 20th century, along with pre-existing cell-mediated immunity (36,37) contribute to the overall milder symptoms and lower than expected mortality rate in the elderly.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References and Notes

1. Palese P. Nat. Med 2004;10:S82. [PubMed: 15577936]
2. World Health Organization. http://www.who.int
3. Skehel J, Wiley D. Annu. Rev. Biochem 2000;69:531. [PubMed: 10966468]
4. Most neutralizing antibodies that result from viral infection or vaccination specifically target the HA. The current seasonal flu vaccine formulation includes representative strains from the circulating H1N1, H3N2 and influenza B viruses.
5. Smith GJ, et al. Nature 2009;459:1122. [PubMed: 19516283]
6. Garten RJ, et al. Science 2009;325:197. [PubMed: 19465683]
7. Three H1 genes from the human lineage are included. They respresent human H1 HAs from the 1918 influenza pandemic (A/South Carolina/1/18) and later seasonal strains (A/Puerto Rico/8/34 and A/Brisbane/59/2007). A/Brisbane/59/2007 is the current seasonal vaccine strain. In the swine lineage, classical swine HAs (A/swine/Iowa/15/1930 and A/swine/Ohio/511445/2007) are listed along with swine-origin human viruses (A/New Jersey/1976 and A/California/04/2009).
8. Smith GJ, et al. Proc. Natl. Acad. Sci. USA 2009;106:11709. [PubMed: 19597152]
9. Zimmer SM, Burke DS. N. Engl. J. Med 2009;361:279. [PubMed: 19564632]
10. Gaydos JC, Top FH Jr, Hodder RA, Russell PK. Emerg. Infect. Dis 2006;12:23. [PubMed: 16494712]
11. Materials and methods are avaliable as supporting material on *Science* Online.
12. Lin T, et al. Virology 2009;392:73. [PubMed: 19628241]
13. Gamblin S, et al. Science 2004;303:1838. [PubMed: 14764886]

14. Stevens J, et al. Science 2004;303:1866. [PubMed: 14764887]
15. Caton A, Brownlee G, Yewdell J, Gerhard W. Cell 1982;31:417. [PubMed: 6186384]
16. Brownlee GG, Fodor E. Phil. Trans. R. Soc. Lond. B 2001;356:1871. [PubMed: 11779386]
17. Residue numbering based on H3 HA, as in Wilson IA, et al. Nature 1981;289:366. [PubMed: 7464906]
18. The RMSD between the membrane-distal domain of CA04 to other H1 HA structures (as PDB IDs): 1RD8 (A/South Carolina/1/18, 0.58Å); 1RUZ (A/South Carolina/1/18, 0.76Å); 1RU7 (A/Puerto Rico/8/34, 0.63Å); 1RUY (A/swine/Iowa/15/1930, 0.62Å); 3HTO (A/wild duck/Jiangxi/12416/2005, 0.60Å).
19. In the crystal structures of H1 HA, the three peptide segments comprising the Sa site are all anchored at both ends by aromatic residues whose side chains are buried inside the HA1 receptor-binding domain. These residues are absolutely conserved in all H1 HAs, and the numbers of residues between these hydrophobic anchors remain the same, thus maintaining the structural integrity of the Sa site. Across different subtypes, however, these aromatic residues are less well conserved and slight variations in the length of the component peptide segments may occur.
20. Itoh Y, et al. Nature 2009;460:1021. [PubMed: 19672242]
21. A similar situation may presumably have happened before the 1900s for which we have insufficient data.
22. Hancock K, et al. N. Engl. J. Med 2009;361:1945. [PubMed: 19745214]
23. MMWR Morb. Mortal. Wkly Rep 2009;58:521. [PubMed: 19478718]
24. Taubenberger JK, Morens DM. Emerg. Infect. Dis 2006;12:15. [PubMed: 16494711]
25. Centers for Disease Control and Prevention. http://www.cdc.gov/h1n1flu/
26. Jain S, et al. N. Engl. J. Med 2009;361:1935. [PubMed: 19815859]
27. Webb SA, et al. N. Engl. J. Med 2009;361:1925. [PubMed: 19815860]
28. Chowell G, et al. N. Engl. J. Med 2009;361:674. [PubMed: 19564633]
29. Cao B, et al. N. Engl. J. Med 2009;361:2507. [PubMed: 20007555]
30. Fisman DN, et al. N. Engl. J. Med 2009;361:2000. [PubMed: 19907052]
31. Schulze IT. J. Infect. Dis 1997;176:S24. [PubMed: 9240690]
32. Due to limited surveillance during this time period, it is difficult to estimate exactly when most H1 viruses acquired glycosylation sites in the vicinity of Sa. Between 1933 and 1943, some isolates have predicted glycans in the region and others do not, suggesting that variants with and without Sa glycans co-circulated during this time. The last isolate without any predicted glycans in the Sa site was A/Weiss/1943 (Fig. 2F, table S2).
33. Yu X, et al. Nature 2008;455:532. [PubMed: 18716625]
34. Krause JC, et al. J. Virol 2010;84:3127. [PubMed: 20042511]
35. Indeed, the electron density is weak in this region and is probably only present at all due to a weak crystal contact to an adjacent HA molecule.
36. Xing Z, Cardona CJ. Emerg. Infect. Dis 2009;15:1847. [PubMed: 19891882]
37. Greenbaum JA, et al. Proc. Natl. Acad. Sci. USA 2009;106:20365. [PubMed: 19918065]

38. The work was supported in part by NIH grants AI058113 (I.A.W., J.E.C. and Peter Palese), and AI057157 (J.E.C.), predoctoral fellowships from the Achievement Rewards for College Scientists Foundation and the NIH Molecular Evolution Training Program GM080209 (D.C.E.), and the Skaggs Institute for Chemical Biology. X-ray diffraction datasets were collected at the Stanford Synchrotron Radiation Lightsource beamline 9-2 and at the Advanced Photon Source beamline 23ID-B (GM/CA-CAT). The GM/CA CAT 23-ID-B beamline has been funded in whole or in part with federal funds from National Cancer Institute (Y1-CO-1020) and NIGMS (Y1-GM-1104). Use of the Advanced Photon Source (APS) was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under contract no. DE-AC02-06CH11357. We thank P. Palese (Mount Sinai School of Medicine) for providing the CA04 clone and helpful comments, X. Dai and M.A. Elsliger (The Scripps Research Institute) and C. J. Huffman (Vanderbilt University) for expert technical assistance; H. Tien and D. Marciano of the Robotics Core at the JCSG for automated crystal screening; M. Becker and the staff of the APS GM/CA CAT 23-ID-B for beamline support. This is publication 20395 from The Scripps Research Institute. Coordinates and structure factors are deposited in the Protein Data Bank (PDB) (3LZG for CA04 HA (2 trimers) and 3LZF for 1918 HA/2D1 complex).

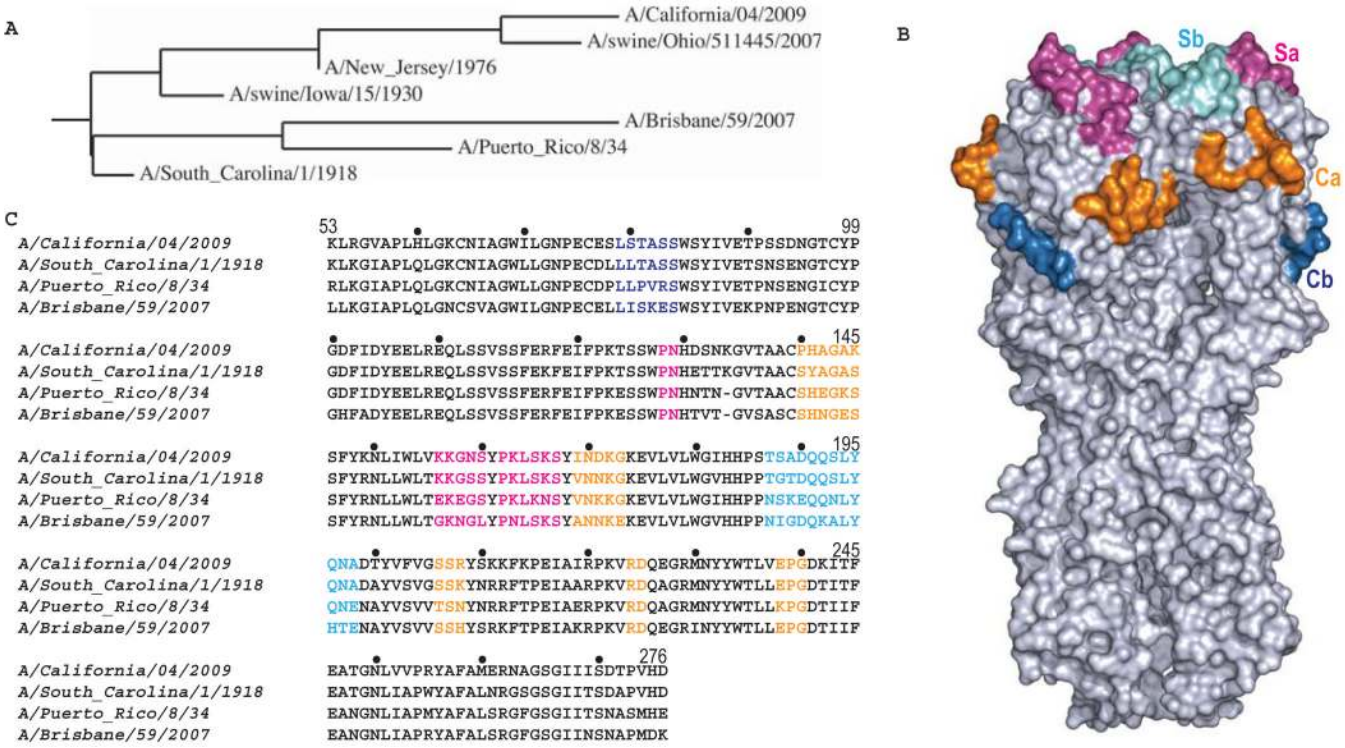

Fig. 1.

Crystal structure, phylogeny and antigenic variation in influenza A 2009 H1N1 HA. (**A**) Phylogenetic tree of selected H1 HAs in swine and human. (**B**) Antigenic structure of CA04 HA from the 2009 H1N1 pandemic virus. A trimer complex is shown in surface representation with the antigenic sites highlighted: Sa site in magenta; Sb site in cyan; Ca site in orange; and Cb site in blue. Sa and Sb sites are located near the receptor-binding site. The Ca site straddles the subunit interface in the trimer. (**C**) Sequence alignment of membrane-distal domains from representative H1 HAs. Antigenic epitopes are color-coded as in **B**.

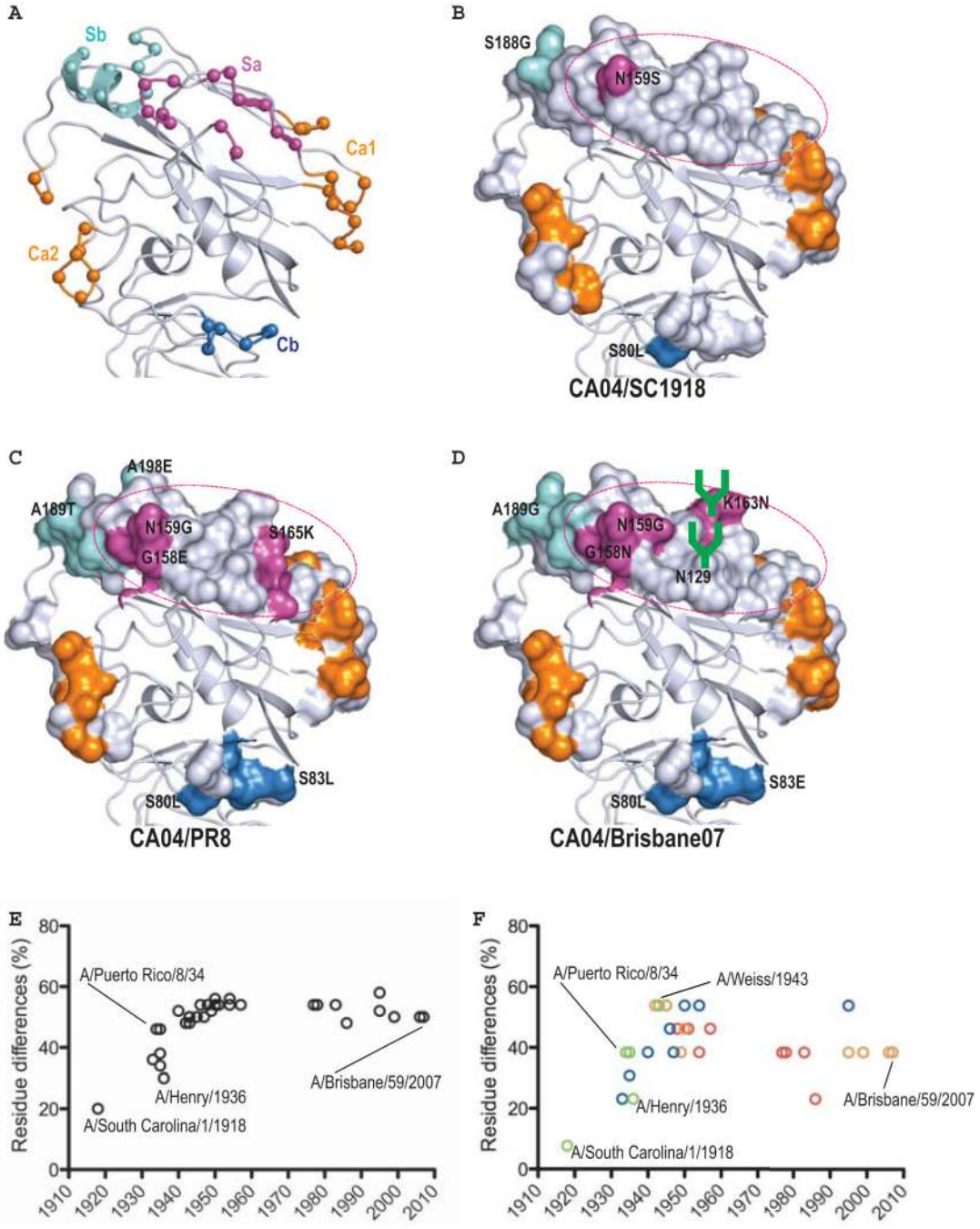

Fig. 2.

Antigenic sites and antigenic variation in H1N1 HAs. (**A**) Localization of H1 antigenic sites in the HA1 membrane-distal domain in close view. Antigenic sites are color-coded as in Fig. 1. (**B–D**) Antigenic variation of seasonal flu H1N1 HAs compared to CA04. Antigenic sites of three H1 HAs are displayed in molecular surface [(**B**) SC1918; (**C**) PR8/34; (**D**) Brisbane2007 (current seasonal flu vaccine strain)]. Residual differences between CA04 and selected H1 HAs are highlighted based on antigenic regions (Sa, magenta; Sb, cyan; Ca, orange; Cb, blue). The SC1918 HA antigenic surface is highly conserved in CA04. In later years (1930–2007), the H1N1 HA antigenic surface has become more variable through mutations and importantly, acquired additional N-glycosylation sites (shown as a cartoon with branched sticks

in green) that mask the surface from recognition by neutralizing antibodies. (**E**) Increased variation over these time periods is shown in the plot of residual differences between CA04 and selected human H1 HAs in the antigenic sites over time. Sequences of H1 HA in the early half of 20th century are selected from the NCBI Influenza Virus Resource (http://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html). After 1977, only the nine H1N1 vaccine strains are shown. The sequences and antigenic site variation of the HAs analyzed in panels **E** and **F** are listed in table S2. (**F**) Similar plot for the Sa site color-coded by the number of potential N-glycosylation sites in the Sa antigenic region (green, 0 glycans; blue, 1 glycan; orange, 2 glycans; red, 3 glycans).

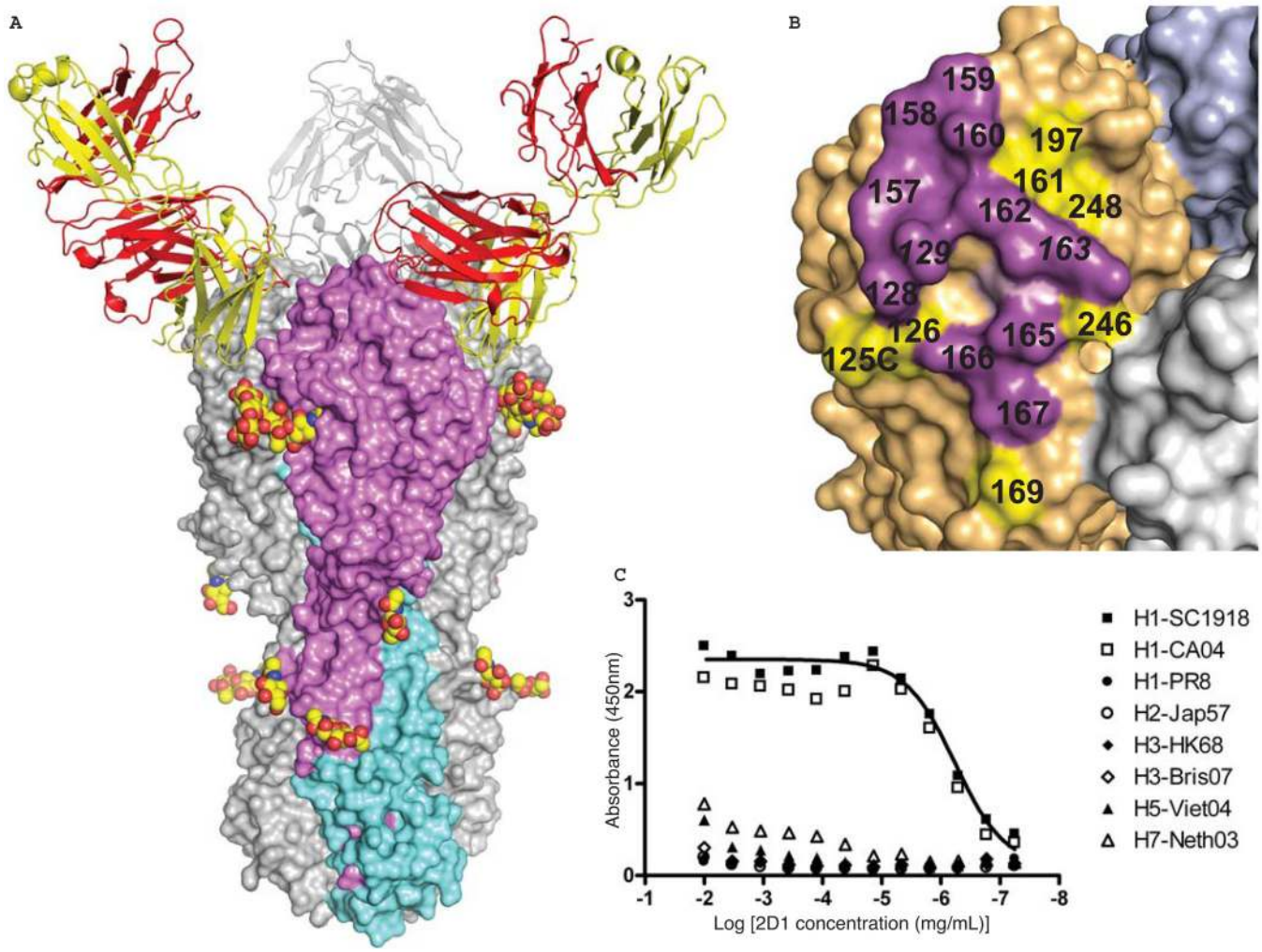

Fig. 3.

Crystal structure of cross-neutralizing antibody 2D1 in complex with SC1918. (**A**) Antibody 2D1, with the light and heavy chains of the Fab in yellow and red, respectively, recognizes the Sa site of SC1918 HA, where HA1 is shown in magenta and HA2 in cyan. (**B**) Footprint of antibody 2D1 on the SC1918 HA shows the central role of the Sa site residues for 2D1 binding. The interacting surface contributed by residues from the Sa site is colored in magenta, whereas residues that contribute to the epitope, but are outside of the 'canonical' Sa site, are colored in yellow. Sa site residues not in contact with 2D1 are shown in pink. SC1918 and CA04 do not have N-glycosylation sites in the Sa region, whereas the vaccine strain Brisbane07, like many other seasonal H1 HAs, have acquired potential glycosylation sites at positions 129 and 163 (italics on surface). (**C**) Antibody 2D1 exhibits strong binding to both 1918 HA and CA04 HA, but not to PR8 and HAs of other influenza subtypes, as tested in ELISA assay.