Protein Cell 2010, 1(5): 459–467
DOI 10.1007/s13238-010-0059-1

Protein & Cell

RESEARCH ARTICLE

Crystal structure of the swine-origin A (H1N1)-2009 influenza A virus hemagglutinin (HA) reveals similar antigenicity to that of the 1918 pandemic virus

Wei Zhang[1,2], Jianxun Qi[1], Yi Shi[1,2], Qing Li[1,2,3], Feng Gao[4], Yeping Sun[1], Xishan Lu[1,5], Qiong Lu[1,2], Christopher J. Vavricka[1], Di Liu[1,6], Jinghua Yan[1], George F. Gao[1,2,7] ✉

[1] CAS Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China
[2] Graduate University, Chinese Academy of Sciences, Beijing 100049, China
[3] College of Life Science, University of Science and Technology of China, Hefei 230027, China
[4] Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China
[5] College of Veterinary Medicine, China Agricultural University, Beijing 100094, China
[6] Network Information Center, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China
[7] Beijing Institutes of Life Science, Chinese Academy of Sciences, Beijing 100101, China
✉ Correspondence: gaof@im.ac.cn
Received April 22, 2010 Accepted May 5, 2010

ABSTRACT

Influenza virus is the causative agent of the seasonal and occasional pandemic flu. The current H1N1 influenza pandemic, announced by the WHO in June 2009, is highly contagious and responsible for global economic losses and fatalities. Although the H1N1 gene segments have three origins in terms of host species, the virus has been named swine-origin influenza virus (S-OIV) due to a predominant swine origin. 2009 S-OIV has been shown to highly resemble the 1918 pandemic virus in many aspects. Hemagglutinin is responsible for the host range and receptor binding of the virus and is therefore a primary indicator for the potential of infection. Primary sequence analysis of the 2009 S-OIV hemagglutinin (HA) reveals its closest relationship to that of the 1918 pandemic influenza virus, however, analysis at the structural level is necessary to critically assess the functional significance. In this report, we report the crystal structure of soluble hemagglutinin H1 (09H1) at 2.9 Å, illustrating that the 09H1 is very similar to the 1918 pandemic HA (18H1) in overall structure and the structural modules, including the five defined antiboby (Ab)-binding epitopes. Our results provide an explanation as to why sera from the survivors of the 1918 pandemics can neutralize the 2009 S-OIV, and people born around the 1918 are resistant to the current pandemic, yet younger generations are more susceptible to the 2009 pandemic.

KEYWORDS Influenza virus, pandemic, 2009, 1918, S-OIV, hemagglutinin (HA), structure

INTRODUCTION

In April 2009, severe flu-like human cases were reported in Mexico and the causative virus was soon identified as the influenza A virus (CDC, 2009). The virus spread rapidly to the USA and was recognized as a new reassortant with three genetic lineages, mainly with a swine origin, and was therefore called swine-origin influenza virus (S-OIV) (Chang et al., 2009). Due to its extremely rapid human-to-human transmission rate, 2009 S-OIV was detected throughout the entire world within two months and the World Health Organization (WHO) declared an official pandemic on June 11th, 2009, the first pandemic in the 21st century (WHO, 2009).

Influenza virus is a negative-strand segmented RNA virus classified into *Family Orthomyxoviridae*, which is comprised of three types of influenza virus, A, B and C (Fields et al.,

2007). Among them, influenza A viruses are the major cause of the seasonal flu and occasional pandemics. There are 8 gene segments in the influenza A virus genome, encoding 11 proteins (Fields et al., 2007; Neumann et al., 2009). They are known as HA, NA, NS1, NS2, M1, M2, PA, PB1, PB2 and PB1-F2. Hemagglutinin (HA) is the major influenza virus surface envelope protein with 16 known serotypes and it is responsible for the host-range, receptor binding, stimulating host neutralization-antibody response and some additional functions. Neuraminidase (NA), with 9 subtypes that have been found to date, is the key component functioning in virus release. Gene segment exchanges (gene reassortment) from different virus origins lead to the emergence of new viruses. Current S-OIV (2009 H1N1 influenza virus) is a triple reassortment (in terms of host-origin) of influenza gene segments, with PA and PB2 from avian, PB1 from human and HA, NP, NS, NA and M from classical/Eurasian swine (Dawood et al., 2009; Smith et al., 2009; Gao and Sun, 2010). This current H1N1 virus has been found to have many similar properties to the 1918 pandemic H1N1 virus, including the overall similarity of the genome and characteristics of pathogenesis (Itoh et al., 2009; Shen et al., 2009; WHO, 2009; Yang et al., 2009; Igarashi et al., 2010; Yeping Sun et al., 2010).

Influenza A virus infects the host cell by way of endocytosis mediated by the binding of HA to sialic acid receptors (either 2, 3 or 2, 6 linked (reviewed in Liu et al., 2009)). Following the fusion between the virus and endosomal membrane at low pH, viral RNA is released into the cytoplasm and the virus is able to replicate (Skehel and Wiley, 2000). Binding of HA to its receptor initiates this infection process and is therefore a major target for the host immune system and the major component of influenza vaccine.

HA is initially translated as a precursor (HA0) and is cleaved by host proteases into two subunits, HA1 and HA2, which form the functional protein. The structural basis for the function of the HA trimer in both its HA0 and HA1/HA2 forms has previously been solved by Wiley and colleagues (Wilson et al., 1981; Chen et al., 1998). To elucidate the molecular characteristics of HA from the 2009 S-OIV (09H1) in comparison with other HAs, especially the 1918 HA (18H1), we took the advantage of soluble transmembrane-removed recombinant HA for structural studies. Here we report the 09H1 structure at 2.9 Å in comparison with 1918 H1. The results not only provide the structural basis of the suspected similarity between the 09H1 and 18HA but also reveal some 09H1 distinct characteristics including the basic patches and *N*-glycosylation of N279.

RESULTS

The overall structure of soluble 09H1

The soluble 09H1 (A/California/04/2009) was prepared based on the method described by Stevens et al. in a baculovirus expression system (Stevens et al., 2004, 2006; Ekiert et al., 2009). The construct incorporated a C-terminal thrombin cleavage site, a trimerizing sequence ('foldon') and a hexa His-tag at the extreme C terminus of the construct to enable protein purification. Although protein was theoretically expressed as a trimer (Fig. 1), only monomers were purified by gel filtration after foldon removal by the thrombin cleavage step, judged by gel-filtration molecular weight estimation (Fig. 1). However, monomer preparations were still able to reform trimers in the crystal state and the HA1 and HA2 have been cleaved in our observed structure although over 90% of the protein preparations are uncleaved HA0 during the protein preparation process (Fig. 1).

Crystals of the 09H1 protein were grown using the hanging-drop diffusion method at pH 6.5 and 18°C. The structure was

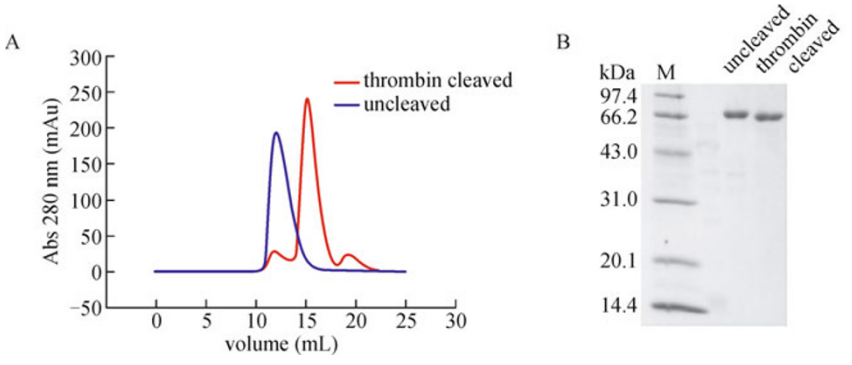

Figure 1. Gel filtration and SDS-PAGE analysis of 09H1. (A) The initially expressed 09H1 protein eluted as a trimer (blue line) by size exclusion chromatography on a Superdex-200 10/300 GL column (GE Healthcare), after thrombin cleavage, the protein behaved as a monomer (red trace) but yielded quality crystals as a trimer. (B) Uncleaved HA protein ran as a 70 kDa protein on an SDS-PAGE gel while thrombin cleaved 09H1 ran as a 65 kDa protein. The decrease in size was due to the removal of the foldon trimerizing domain from the C terminus.

solved at 2.9 Å by molecular replacement using the 1934-HA (PDB code 1RU7) as a search model. Crystallographic statistics are given in Table 1. The overall structure of the 09H1 (Fig. 2) is highly similar to all the solved HA structures to date (Yeping Sun et al., 2010) which is in accordance to the soluble 18H1 (A/South Carolina/1/18) preparation reported earlier (Stevens et al., 2004). The crystal packing is the same for the 09HA as it is in all other previously reported HA structures and the average RMSD value among the six HA molecules in an asymmetric unit is 0.489 Å (calculated with LSQMAN).

Striking similarity of the 5-defined antibody recognizing epitopes between 09H1 and 18H1

To reveal the antigenicity of the 09H1, we studied the defined antibody-recognizing sites (Sa, Sb, Ca1, Ca2 and Cb) described earlier (Caton et al., 1982), which is important for the escape of influenza virus from host immune system surveillance. We compared these five antigenic sites with all the known HA structures of the H1 serotype. They are 09H1, 18H1 and 30H1 (swine A/swine/Iowa/30), 34H1 (human A/Puerto Rico/8/34) and 05H1 (avian A/WDK/JX/12416/2005) (Gamblin et al., 2004; Lin et al., 2009). As shown in Fig. 3, human 34H1 is unique, in comparison with all other four animal-origin H1s, which further confirms the antigenicity difference between human- and animal-origin HAs. This may be explained by differences in the immune pressures encountered in animals and humans. All the four animal-origin H1s have structurally similar antigenic sites. In detail, between 18H1 and 30H1, the following mutations occurred: S173N in Sa site, G202S and S207N in Sb site, D239G and S154P in Ca2 site, A90V in Cb site (totally 6 mutations); between 18H1 and 09H1, these mutations occurred: S173N in Sa site, G202S and T203A in Sb site, N185D, V183I and K223R in Ca1 site, S159K, Y155H and S154P in Ca2 site, L88S in Cb site (totally 10 mutations); between 18H1 and 05H1, these mutations occurred: S173T in Sa site, G202T,

Table 1 Data collection and refinement statistics (molecular replacement)

	09HA
data collection	
space group	P1
cell dimensions	
a, *b*, *c* (Å)	66.02, 115.19, 114.99
α, β, γ (°)	62.31, 77.94, 81.05
resolution (Å)	50–2.90 (2.90–3.00)[a]
R_{sym} or R_{merge}	0.105 (0.501)
I/σ (I)	12.9 (1.91)
completeness (%)	98.0 (90.7)
redundancy	3.7 (2.7)
refinement	
resolution (Å)	24.2–2.9
No. reflections	64,796
R_{work}/R_{free}	24.6/27.0
No. atoms	
protein	22,860
ligand/ion	431
water	321
B-factors	
protein	81.2
N-glycosylation sites	130.6
water	68.3
r.m.s. deviations	
bond lengths (Å)	0.004
bond angles (°)	0.971

[a] Values in parentheses are for highest-resolution shell.

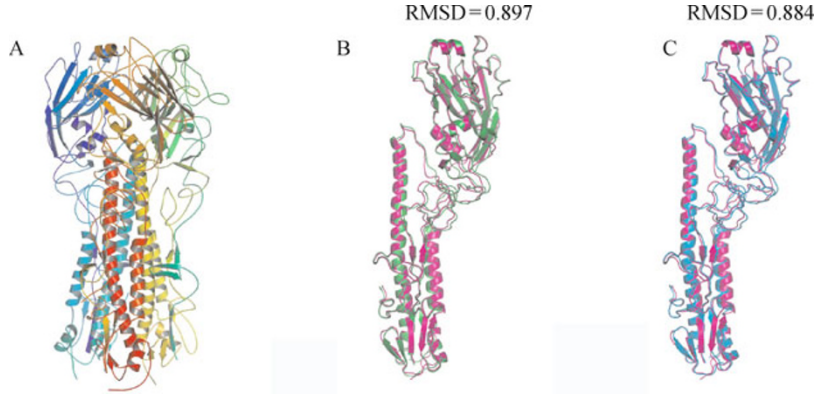

Figure 2. The 09H1 trimer structure and the comparison between 09H1 , 18H1 and 34H1. (A) Overall view of the 09H1 trimer structure. (B and C) Comparison of the monomers between the 09H1 (magenta) and 18H1 (green) or 34H1 (blue). The RMSD values for both two comparisons are very small (0.90 and 0.88), which shows that 09H1 has a very similar structure to 18H1 and 34H1 (seasonal flu).

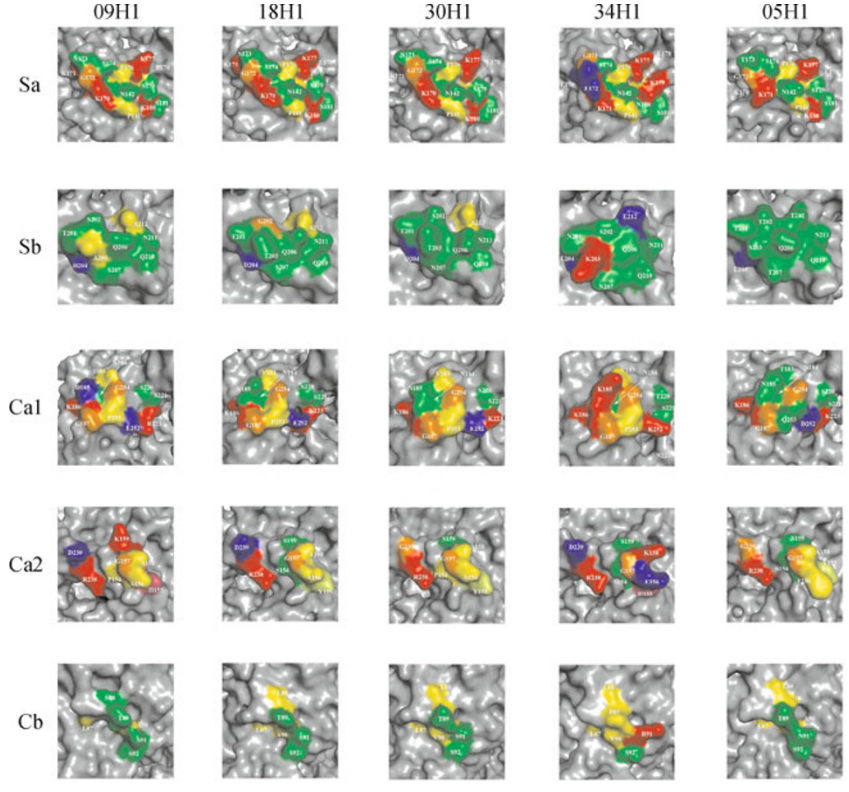

Figure 3. Highlights of the five defined antibody-epitopes (Sa, Sb, Ca1, Ca2, Cb) of the 09H1 in comparison with 18H1, 30H1, 34H1 and 05H1, all the known crystal structures of the H1 serotypes to date. Amino acids are colored by the customized ClustalX color scheme: Ala, Leu, Ile, Pro, and Val (yellow); Lys and Arg (red); Asp and Glu (blue); Ser, Thr, Gln and Asn (green); His (pink); Tyr (limon); Gly (orange). It is clearly apparent that the 09H1 and 18H1 are most likely related among these H1 structures. All the pictures (including the pictures in other figures) were generated by PYMOL based on the crystal structures of 09H1, 18H1, 30H1, 34H1 and 05H1 (PDB codes: 1RUZ, 1RUY, 1RU7, 3HTO and 3LYJ, respectively).

T203S, D204E, S207T and A212T in Sb site, V183T, E252D and P253Q in Ca1 site, A156F and D239G in Ca2 site, and S91N in Cb site (totally 12 mutations). Thus, from 1918 to 2009, in animals like swine and avian, the antibody-mediated

selective pressure may be much lower than in humans due to less complex immune systems, and the amino acid residues in the antigenic sites would not change frequently which is consistent with previous studies (Sugita et al., 1991).

However, in the more complex human immune system, which has also been exposed to vaccination since the 1918 pandemic flu, the HA molecule of seasonal H1N1 flu would benefit from frequent mutation of antigenic sites (antigenic drift) as a way to escape from immune system surveillance, just as was observed in the 34H1 structure. It is noteworthy that 09H1 and 18H1 structures most closely resemble each other (Fig. 3) with only ten amino acid differences in the five epitopes (most of them result from similar amino acid substitutions).

Strong basic patches and their implications for virus fusion

There are two basic patches in the 09HA structure (Fig. 4). The first one is near the cleavage site, which contains three basic histidine residues (His18 and His38 from the HA1 chain, and His111 from the HA2 chain) and this basic patch is conserved in human H1, H2 and H5 sequences (Stevens et al., 2004, 2006). While the second basic patch is a defined structural module initially found in the 18H1 structure, later found in other H1 HA structures and one H5 HA structures, but not observed in other HA structures and thought to be important for virus membrane fusion and infection (Stevens et al., 2004, 2006). Thus, we examined the second basic patch on the HA1 chain adjacent to the vestigial esterase domain. This very-basic patch is composed of four HA1 histidines (His47, His275, His285, His298) and a lysine (Lys50), in the four H1 HA molecules (18H1, 30H1, 34H1 and 05H1), as shown in Fig. 3. By contrast, the area of the basic patches (blue regions) is obviously larger in 09H1 than any other H1, including 18H1, due to its N/S46K and H285K substitutions. Although there has been no direct evidence till now, the increased basicity of the basic patch has been proposed to be related to enhanced virus membrane fusion and subsequent pathogenicity as proposed by Stevens et al. (2004, 2006), and experimental testing of such a hypothesis is required.

Extra *N*-glycosylation site near the second basic patch region and Cb antigenic site

The primary sequence of the soluble 09H1 in this analysis predicts 6 possible glycosylation sites (NXS/T) per monomer. Interpretable electron density in the complex is observed at all the potential *N*-glycosylation sites in each 09H1 monomer (36 sites in total of the six molecules).Therefore we have determined that all potential *N*-glycosylation sites are occupied in the 09H1, at least in our *in vitro* baculovirus-expressed form, with typical *N*-glycosylation structures. Compared with 18H1 and other H1 molecules (Fig. 5), a new glycosylation site at NTT (residue N279, H3 numbering) is observed near the second basic patch we mentioned above. This glycosylation likely results in steric hindrance to the second basic patch, thereby reducing the ability of the basic enzymes to interact with the basic patch, which may help the membrane-fusion of virus. Furthermore, this new glycosylation site may also interfere with antibody recognition of the Cb antigenic site. Clearly, it deserves further investigation in the near future.

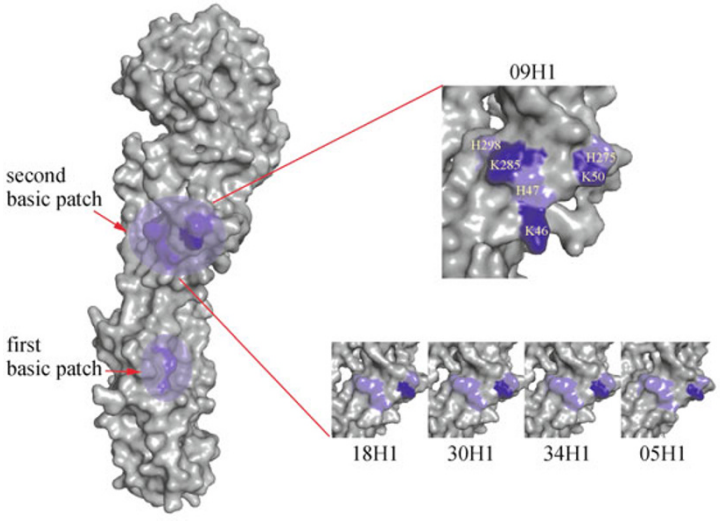

Figure 4. Comparison of the basic patches of 09H1 with other H1 HAs. Overview of the two basic patches (blue region) is shown on the left, and a detailed comparison of different HAs is shown on the right. Lysine is colored with dark blue, and histidine is shown in light blue.

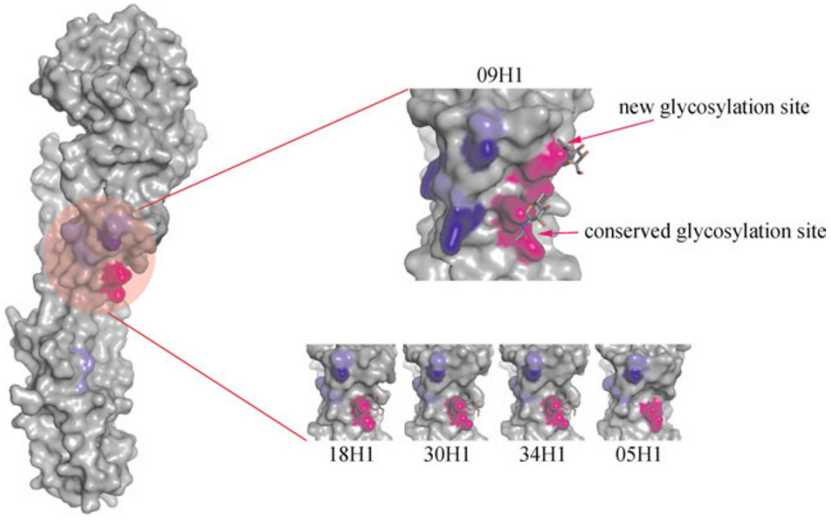

Figure 5. Comparison of the *N*-glycosylation sites near the upper basic patch. A new glycosylation site is found in the 09H1 molecule, and the glycosylation sites are shown in magenta with the glycans represented as stick patterns.

Lack of the glycosylation sites on the receptor binding proximal region (receptor binding apex) and its implications for the virus neutralizations

Influenza virus can escape targeting by the host immune system by several ways. Two examples are discussed here. First, the virus can change its antigenic surface, which is called antigen drift. Second, the virus can evolve new glycosylation sites that allow sugars derived from the host cell to attach to the antigenic sites. These sugars can mask the antigenic sites in order to avoid the antibodies to recognize the virus. As we mentioned above, the HA molecule of the 2009 S-OIV pandemic virus has highly similar antigenic sites with 1918 pandemic virus and other animal-origin influenza viruses, and is divergent with HA molecules of seasonal flu which are under high pressure in the host body. With exception to the considerable pandemic divergence, the seasonal flu has evolved gradually to acquire up to three glycosylation sites near the receptor binding site (Fig. 6), and the sugars attached to these glycosylation sites can affect the ability of antibodies to recognize the amino acids within the Sa and Sb antigenic sites. As a result, the antibodies in the body of the young people whose immune systems are adapted to the seasonal flu tend to attack the other three antigenic sites (Ca1, Ca2 and Cb). Ca1 and Ca2 sites are at the subunit interface which is unfavorable for the antibody recognition. However, the Cb site is within the vestigial esterase domain full of protruding loops that can be accessed readily by the antibodies. Thus, we propose that the new glycosylation site near the vestigial esterase domain in the 09H1 may help the swine-origin 2009 influenza virus to escape antibody recognition, partly responsible for the pandemic. Of course this requires further confirmation in the near future.

DISCUSSION

Our crystal structure of the soluble 09HA, together with a recent publication during our preparation of this manuscript (Xu et al., 2010), demonstrates that 09H1 has very similar antigenic sites with 18H1, but is distinct from the seasonal flu viruses, indicating that the pandemic (H1N1) influenza virus (2009 S-OIV H1N1) can be neutralized by the antibody from people who have been infected with 1918 pandemic influenza virus. This might explain why younger people are more susceptible to the current virus, but the people who experienced the 1918 pandemic are more resistant to the current pandemic virus (Itoh et al., 2009). The lack of the cross-antigenicity with the decades' seasonal flu HAs might also explain how the 2009 S-OIV can spread so rapidly to a pandemic level. We can predict that in the future, under the pressure of the host immunity, the 2009 S-OIV will evolve to similar seasonal flu just as the 1918 influenza virus did from 1918 to 2009.

A stronger basic patch may increase the infectivity of the influenza virus, while the new *N*-glycosylation site (N279) near the stronger basic patch may reduce its function. If the infectivity of the influenza virus is too high, it may kill the host before it is able to spread. Therefore, we speculate that the stronger basic patch and new glycosylation site may be an evolutionary balance for the infectivity of the 2009 S-OIV and might help the virus to survive in the host.

Recently, it has been suggested that the focus of immune system response can be changed by modification of glycosylation sites in the HA protein (Wei et al., 2010). The addition of a new *N*-glycosylation site (N279) within the vestigial esterase domain, which is not observed in the 1918 H1 HA, may facilitate world-wide transmission of the swine-

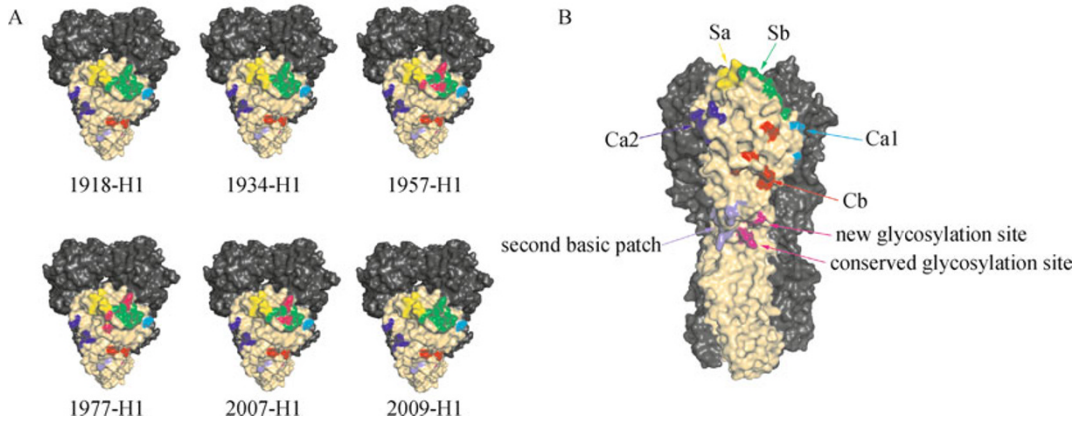

Figure 6. Analysis of 09HA *N*-glycosylation sites. (A) Comparison of the *N*-glycosylation sites on the head of HA molecules between swine-origin flu and seasonal flu from 1918 to 2009. Glycosylation sites on the head of HA molecules are not present in 09H1 and 18H1, while in HA from the seasonal flu (1957-H1, 1977-H1, 2007-H1) there are up to three glycosylation sites (with exception to 1934-H1). Only the structures of 1957-H1 (A/Denver/57), 1977-H1 (A/USSR/90/1977), and 2007-H1 (A/Brisbane/59/2007) were modeled for comparison, as structures for all the H1 HAs are very similar. Glycosylation sites on the head are marked in hot pink color. (B) Closer view of the the N279 *N*-glycosylation site, the second basic patch and the five antigenic sites of 09H1. The recognition of the antibody specific to the Cb sites can be affected by the presence of glycans in this region.

origin 2009 pandemic virus by interfering with antibody recognition specific to the Cb site. Thus, people who are adapted to seasonal HA will likely be less resistant to the 2009 pandemic virus than influenza strains originating from the 1918 virus, which do not possess this *N*-glycosylation site.

In conclusion, our structural analyses of the 2009 pandemic HA show its true similarity to the 1918 pandemic HA. The new *N*-glycosylation site and strong second basic patch of the 2009 HA might contribute to the rapid human-to-human transmission and high virulence, which deserves further experimental work in the future.

MATERIALS AND METHODS

Cloning, expression and purification of 09H1

Methods for the cloning, expression and purification of secreted 09H1 were based on those reported by Stevens et al. (2004, 2006) and Ekiert et al. (2009). Briefly, cDNA encoding amino acid residues 11–505 (11–329 (HA1) and 1–176 (HA2)) from the A/California/04/2009 HA ectodomain was amplified and cloned into the baculovirus transfer vector pAcGP67-B (BD Biosciences), containing a thrombin cleavage site, a trimerizing 'foldon' sequence and a his-tag.

The recombinant transfer vector and BaculoGold™ linearized viral DNA (BD Biosciences) were co-transfected into log phase Sf9 cells using CellfectinR Reagent (Invitrogen). Sf9 cells were grown in suspension using Sf-900II SFM serum-free media (Invitrogen) at 28°C and 120 RPM. A high-titer viral stock of pure recombinant virus was generated through amplification in suspension cultured Sf9 cells.

Sf9 suspension cultures were transfected with high tighter baculovirus and grown for 3 d. Cells and debris were removed from the sample by centrifugation at 4000 g for 40 min and the subsequent growth media was applied to a HisTrap FF 5-mL column (GE Health). After elution with 300 mM imidazole, combined HA fractions were dialyzed against 10 mM Tris-HCl, 50 mM NaCl, pH 8.0 and then further separated using a Mono-Q 4.6/100 PE column (GE Health).

The HA protein was then subjected to thrombin digest (Sigma, 3 units/mg HA; overnight at 4°C) followed by gel filtration chromatography using a Superdex-200 10/300 GL column (GE Healthcare) with 20 mM Tris-HCl, 50 mM NaCl, pH 8.0 as running buffer. After final processing with thrombin, the recombinant HA includes additional amino acid residues (RLVPR) at the C-terminal. The protein also contains additional N-terminal residues (ADLGSR) encoded by the plasmid. Fractions containing highly pure HA were pooled and concentrated to 5 mg/mL using a membrane concentrator with a molecular weight cutoff of 10,000 (Millipore).

Crystallization and data collection

09H1 crystals were grown using the hanging-drop diffusion method with a reservoir solution (0.4 mL) of 10% PEG 6000, 5% MPD, 0.1 M MES, pH 6.5. Crystallization drops were set up at 18°C, with 1.5 μL of protein mixed with an equal volume of reservoir solution. Crystals appeared in approximately one month. The resulting crystals were cryoprotected by soaking in 2 μL well solution mixed with 1 μL 50% PEG 6000, then flash-cooled in a cold nitrogen-gas stream at 100 K. X-ray diffraction data were collected on Shanghai Synchrotron Radiation Facility (SSRF) beamline 17U at a wavelength of 0.9795 Å.

Data processing, molecular replacement and structure refinement

The data were processed and scaled by using the HKL2000 (Otwinowski and Minor, 1997). The structure of 09H1 was solved at 2.9 Å by the molecular replacement method using Phaser (Read, 2001) from the CCP4 program suite (CCP4, 1994) (Collaborative Computational Project, Number 4) with the structure of 34H1 (PDB ID 1RU7) as the search model. Extensive model building and restrained

refinement were performed using COOT (Emsley and Cowtan, 2004) and REFMAC5 (Murshudov et al., 1997). The Further rounds of refinement were performed using the Phenixrefine program implemented in the PHENIX package (Adams et al., 2002) with isotropic ADP refinement and bulk solvent modeling with NCS restriction. The final model has an R_{work} of 24.6% and an R_{free} of 27.0%. The stereochemical quality of the final model was assessed with the program PROCHECK (Laskowski et al., 1993). 81.3% of the peptide bonds were present in the most favored region of the Ramachandron plot, 18.4% in the additionally allowed region and 0.3% in the generally allowed region.

PROTEIN DATA BANK ACCESSION CODES

Atomic coordinates and structure factors for the reported crystal structure have been deposited with the Protein Data Bank accession number 3LYJ. Correspondence and requests for reprints and materials should be addressed to G.F.G. (gaof@im.ac.cn).

ACKNOWLEDGMENTS

We thank Ms. Hui Niu (NIBS), Dr. Jijie Chai (NIBS), Dr. Jinsong Liu (GIBH, CAS), Dr. Zhijie Liu (IBP, CAS) and Dr. Maojun Yang (Tsinghua University) for help in performing the experiments. Assistance by the staff (esp. Dr. Sheng Huang and Dr. Jianhua He) at Shanghai Synchrotron Radiation Facility (SSRF-beamline 17U) is acknowledged. This work is supported by the intramural grant of the Chinese Academy of Sciences (Grant No. KSCX2-YW-R-158), the National Basic Research Program (973 Program) (Grant Nos. 2010CB534004 and 2005CB523001). G.F.G. is a distinguished young investigator of the NSFC (Grant No. 30525010). Dr. Christopher Vavricka is, partly, supported by the Fellowship for Young International Scientists of the Chinese Academy of Sciences (Grant No. 2009Y2BS2).

ABBREVIATIONS

05H1, 2005 H1; 09H1, 2009 H1; 18H1, 1918 H1; 34H1, 1934 H1; HA, hemagglutinin; NA, neuraminidase; S-OIV, swine-origin influenza virus

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.

Caton, A.J., Brownlee, G.G., Yewdell, J.W., and Gerhard, W. (1982). The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). Cell 31, 417–427.

Centers for Disease Control and Prevention (CDC). (2009). Outbreak of swine-origin influenza A (H1N1) virus infection- Mexico, March-April 2009. MMWR Morb Mortal Wkly Rep 58, 467–470.

Chang, L.Y., Shih, S.R., Shao, P.L., Huang, D.T., and Huang, L.M. (2009). Novel swine-origin influenza virus A (H1N1): the first pandemic of the 21st century. J Formos Med Assoc 108, 526–532.

Chen, J., Lee, K.H., Steinhauer, D.A., Stevens, D.J., Skehel, J.J., and Wiley, D.C. (1998). Structure of the hemagglutinin precursor cleavage site, a determinant of influenza pathogenicity and the origin of the labile conformation. Cell 95, 409–417.

Collaborative Computational Project, Number 4. (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr 50, 760–763.

Dawood, F.S., Jain, S., Finelli, L., Shaw, M.W., Lindstrom, S., Garten, R.J., Gubareva, L.V., Xu, X., Bridges, C.B., and Uyeki, T.M., and the Novel Swine-Origin Influenza A (H1N1) Virus Investigation Team. (2009). Emergence of a novel swine-origin influenza A (H1N1) virus in humans. N Engl J Med 360, 2605–2615.

Ekiert, D.C., Bhabha, G., Elsliger, M.A., Friesen, R.H., Jongeneelen, M., Throsby, M., Goudsmit, J., and Wilson, I.A. (2009). Antibody recognition of a highly conserved influenza virus epitope. Science 324, 246–251.

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

Fields, B.N., Knipe, D.M., and Howley, P.M. (2007). Fields virology, 5th edn (Philadelphia, Wolters Kluwer Health/Lippincott Williams & Wilkins).

Gamblin, S.J., Haire, L.F., Russell, R.J., Stevens, D.J., Xiao, B., Ha, Y., Vasisht, N., Steinhauer, D.A., Daniels, R.S., Elliot, A., *et al*. (2004). The structure and receptor binding properties of the 1918 influenza hemagglutinin. Science 303, 1838–1842.

Gao, G.F., and Sun, Y.P. (2010). It is not just AIV: from avian to swine-origin influenza virus. Sci China C Life Sci 53, 151–153.

Igarashi, M., Ito, K., Yoshida, R., Tomabechi, D., Kida, H., Takada, A., and Belshaw, R. (2010). Predicting the antigenic structure of the pandemic (H1N1) 2009 influenza virus hemagglutinin. PLoS ONE 5, e8553.

Itoh, Y., Shinya, K., Kiso, M., Watanabe, T., Sakoda, Y., Hatta, M., Muramoto, Y., Tamura, D., Sakai-Tagawa, Y., Noda, T., *et al*. (2009). *In vitro* and *in vivo* characterization of new swine-origin H1N1 influenza viruses. Nature 460, 1021–1025.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: A program to check the stereochemical quality of protein structures. J Appl Cryst 26, 283–291.

Lin, T., Wang, G., Li, A., Zhang, Q., Wu, C., Zhang, R., Cai, Q., Song, W., and Yuen, K.Y. (2009). The hemagglutinin structure of an avian H1N1 influenza A virus. Virology 392, 73–81.

Liu, D., Liu, X., Yan, J., Liu, W.J., and Gao, G.F. (2009). Interspecies transmission and host restriction of avian H5N1 influenza virus. Sci China C Life Sci 52, 428–438.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 53, 240–255.

Neumann, G., Noda, T., and Kawaoka, Y. (2009). Emergence and pandemic potential of swine-origin H1N1 influenza virus. Nature 459, 931–939.

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

Read, R.J. (2001). Pushing the boundaries of molecular replacement with maximum likelihood. Acta Crystallogr D Biol Crystallogr 57, 1373–1382.

Shen, J., Ma, J., Wang, Q., and Martin, D.P. (2009). Evolutionary trends of A(H1N1) influenza virus hemagglutinin since 1918. PLoS ONE 4, e7789.

Skehel, J.J., and Wiley, D.C. (2000). Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. Annu Rev Biochem 69, 531–569.

Smith, G.J., Vijaykrishna, D., Bahl, J., Lycett, S.J., Worobey, M., Pybus, O.G., Ma, S.K., Cheung, C.L., Raghwani, J., Bhatt, S., *et al*. (2009). Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. Nature 459, 1122–1125.

Stevens, J., Corper, A.L., Basler, C.F., Taubenberger, J.K., Palese, P., and Wilson, I.A. (2004). Structure of the uncleaved human H1 hemagglutinin from the extinct 1918 influenza virus. Science 303, 1866–1870.

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

Sugita, S., Yoshioka, Y., Itamura, S., Kanegae, Y., Oguchi, K., Gojobori, T., Nerome, K., and Oya, A. (1991). Molecular evolution of hemagglutinin genes of H1N1 swine and human influenza A viruses. J Mol Evol 32, 16–23.

Sun, Y., Shi, Y., Zhang, W., Li, Q., Liu, D., Vavricka, C., Yan, J., and Gao, G.F. (2010). In-silicon characterization of the functional and structural modules of the haemagglutinin protein from the swine-origin influenza virus A (H1N1)-2009. Sci China C Life Sci.

Wei, C.J., Boyington, J.C., Dai, K., Houser, K.V., Pearce, M.B., Kong, W.P., Yang, Z.Y., Tumpey, T.M., and Nabel, G.J. (2010). Cross-neutralization of 1918 and 2009 influenza viruses: role of glycans in viral evolution and vaccine design. Sci Transl Med 2, 24ra21.

WHO (2009). World now at the start of 2009 influenza pandemic.

Wilson, I.A., Skehel, J.J., and Wiley, D.C. (1981). Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. Nature 289, 366–373.

Xu, R., Ekiert, D.C., Krause, J.C., Hai, R., Crowe, J.E. Jr, and Wilson, I.A. (2010). Structural Basis of Preexisting Immunity to the 2009 H1N1 Pandemic Influenza Virus. Science 328, 357–60.

Yang, Y., Sugimoto, J.D., Halloran, M.E., Basta, N.E., Chao, D.L., Matrajt, L., Potter, G., Kenah, E., and Longini, I.M. Jr. (2009). The transmissibility and control of pandemic influenza A (H1N1) virus. Science 326, 729–733.

