

# Antigen-Antibody docking reveals the molecular basis for cross-reactivity of the 1918 and 2009 Influenza A/H1N1 pandemic viruses

Sarah Cherian, Pratip Shil, Akhilesh Chandra Mishra\*

Bioinformatics and Data management Division, National Institute of Virology, 20 A, Dr. Ambedkar Road, Pune - 411001, India; Akhilesh C Mishra - Email: acm1750@rediffmail.com; Phone : +91 20 26124386; Fax: +91 20 26122669; \*Corresponding author

Received December 16, 2010; Accepted January 06, 2011; Published March 02, 2011

## Abstract:

To understand the reported cross-reactivity of the 2009 H1N1 and the 1918 H1N1 pandemic viruses we docked the crystal structure of 2D1, an antibody derived from a survivor of the 1918 pandemic, to the structures of hemagglutinin (HA) of the 2009 strain and seasonal H1 vaccine strains. Our studies revealed that 2D1 binds to the 2009 HA at antigenic site 'Sa', with stabilizing contacts, similar to that in an available co-crystal structure of 2D1-1918 HA. However, 2D1 failed to bind to the known antigenic sites in the HAs of seasonal strains. Our study thus reveals the molecular basis for pre-existing immunity in elderly people to the 2009 pandemic virus.

**Keywords:** pandemic Influenza, 2009 H1N1, 1918 H1N1, hemagglutinin, antibody, antigenic site, docking, pre-existing immunity

## Background:

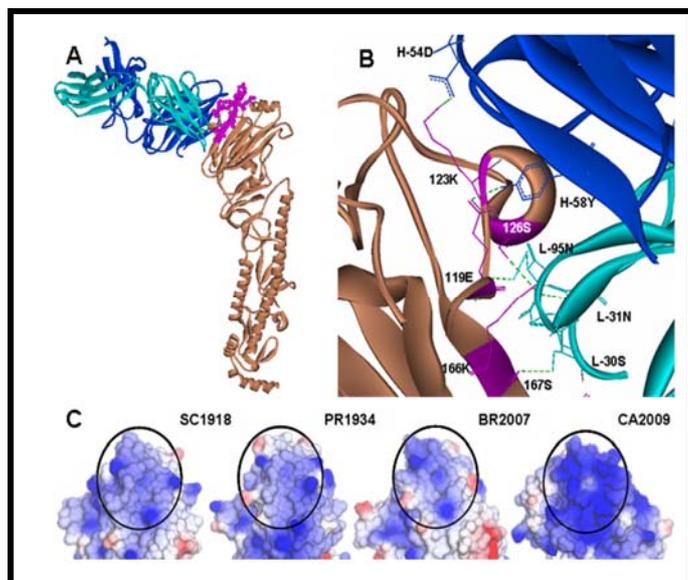
The 1918 H1N1 influenza pandemic has been referred as the 'mother' of all pandemics. Studies on epidemiological features [1, 2], neutralization [3], hemagglutinin (HA) receptor binding and glycosylation sites [4] and antigenic epitopes [5] have provided evidence that the 2009 H1N1 pandemic virus is closer to the 1918 virus than other seasonal H1N1 Influenza viruses. Antigenic analysis has shown that antibodies to the seasonal H1N1 virus do not protect against the pandemic H1N1 virus. However, other studies [3, 6] have shown that a significant percentage of people aged 65 and older do have some immunity against the pandemic virus. This suggests that the presence of pre-existing cross-reactive antibodies provides protection against the 2009 virus in people born early in the 20<sup>th</sup> century. Xu *et al.* (2010) reported a co-crystal structure of an antibody 2D1, derived from a survivor of the 1918 pandemic, and the hemagglutinin of A/South Carolina/1/1918 (SC1918) and also the crystal structure of HA of A/California/04/2009 (CA2009) [7]. Among the known antigenic sites in H1 HAs, namely Sa, Sb, Ca1, Ca2 and Cb [8], the 2D1 antibody largely binds to the Sa antigenic site at the apex of the globular head of the HA protein and also forms contacts with residues outside this site. The binding residues in SC1918 were also conserved in CA2009 except for the two mutations, V169I and S159N. The authors speculated that these two amino acid changes may not adversely affect the antibody binding. They substantiated their results by demonstrating that the antibody 2D1 exhibits strong binding to both 1918 HA and CA2009 HA, but not to HAs of seasonal influenza subtypes, A/Puerto Rico/8/1934 (PR1934) and A/Brisbane/59/2007 (BR2007), as tested in ELISA assays. However, in the absence of the crystallographic structure of

the HA of the 2009 pandemic H1N1 virus bound with an antibody we attempted to determine the nature of the reactivity between the 2D1 antibody and CA2009 HA, PR1934 and BR2007 using *in-silico* docking studies. We compared the binding of 2D1 to the CA2009 HA and the HA of seasonal H1N1 strains, both in terms of the specificity to the antigenic site and the total energy of the complex. This would provide insight into the molecular basis for the cross-reactivity and protection between the 2009 H1N1 and the 1918 pandemic viruses and not other seasonal H1N1 strains.

## Methodology:

The co-crystal structure of the antibody (Fab 2D1) and the HA protein of A/South Carolina/1/1918 (3LZF.pdb) and the HA proteins of A/California/04/2009 (3LZG.pdb), A/Puerto Rico/8/1934 (1RU7.PDB) were obtained from the Protein Data Base (PDB). A homology model for residues 18 to 508 of the HA of A/Brisbane/59/2007 strain was built using the template, chain B of A/Japan/305/1957 (H2N2), using the Swiss Model automated server (<http://swissmodel.expasy.org>). Sequence identity and similarity with template (2WRE.pdb) were found to be 64.97% and 79.8% respectively. Sequence comparisons were performed as per methods stated elsewhere [9]. Computational docking, a method which predicts the preferred stable orientation of one molecule when bound to a second one, was carried out to determine the binding site and the residues interacting within the antigen-antibody complex through hydrogen bonds (H-bonds) and salt bridge contacts. In silico antigen-antibody docking was performed by using the ZDOCK online server [10] using default parameters [11]. The antibody Fab 2D1 was subjected

to docking against the HA proteins of CA2009, BR2007 and PR1934 using 'blind' (non-targeted) docking protocol. From the ten solutions returned by ZDOCK, the best solution was the one which satisfied the following conditions: i) the complementarity determining regions (CDRs) of the antibody interacting with the antigen at the antigen-antibody interface and ii) the value of minimized energy of the complex being the least or energetically favourable. For the energy minimization calculations, a GROMOS96 force-field application in SwissPDBViewer (SPDBV) [12] was considered. The structural divergence was evaluated by structural superimposition and determination of the root mean square deviation (RMSD) using SPDBV. For structure visualization and hydrogen bond analyses, the Accelrys' Discovery Studio package (Accelrys Inc., USA) was utilized. NOC [13] was used for determining the electrostatic surface potential of the HA protein structures studied. The details of contacts between the amino acids of antigen and antibody for each complex were obtained by the Contacts of Structural Units (CSU) analyses online facility [14]. The numbering of amino acid residues in all the structures considered in this study is as per the HA nomenclature by Brownlee *et al.* [15].

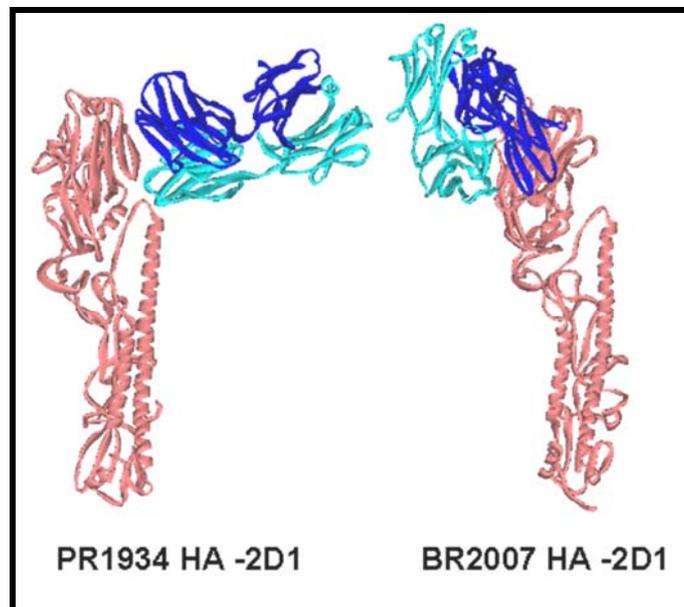


**Figure 1:** (A) Antibody 2D1 docked in the CA2009 HA monomer. 2D1 light and heavy chains in cyan and blue respectively recognizes 'Sa' site indicated in pink. (B) Interacting residues in 2D1-CA2009 HA complex, with H-bonds shown in green dashes (C) Comparison of the electrostatic surface charge distribution in the 'Sa' region (circled) of the HAs of SC1918, PR1934, BR2007 and CA2009 viruses. Blue, white and red colors indicate basic, neutral, and acidic charge distribution respectively.

## Discussion:

Docking of 2D1 onto the crystal structure of CA2009 HA was carried out to determine the binding site and the residues interacting within the antigen-antibody complex through hydrogen bonds (H-bonds) and salt bridges. The 2D1 antibody docked into the Sa antigenic site of CA2009 HA. However, in comparison with SC1918 HA-2D1 co-crystal, some changes were observed in the orientation of the antibody (Figure 1A). The antibody formed three H-bonds and one salt bridge with the HA residues within the Sa site, and five hydrogen bonds and one salt bridge outside the Sa site in the 2D1-CA2009 HA complex (Table 1 see Supplementary material). When compared with the SC1918 HA-2D1 co-crystal, within the Sa site, the H-bonds formed by 166K and 167S with the 2D1 light chain and the salt bridge between 166K and 93D in the 2D1 light chain were retained in the CA2009 HA-2D1 complex (Figure 1B, Table 1 see Supplementary material). Outside the Sa site, the H-bond between 126S and 93D was retained. The changed orientation in CA2009 HA-2D1 complex enabled formation of new H-bonds between 119E and 58Y in the heavy chain complementarity determining region 2 (HCDR2) and 95D in the light chain CDR3 (LCDR3) and 171D with 94S in LCDR3. Salt bridges were formed by 123K and 54D in HCDR2; 166K and 93D in LCDR3. The mutation N171D in CA2009 facilitated the new H-bond formed by 171D. On the other hand, when the docking of 2D1 on to the HA of BR2007 and PR1934 was

carried out, the antibody could not recognize antigenic site Sa or other known antigenic sites (Figure 2). In case of BR2007 the docking program returned a complex wherein the antibody formed three H-bonds between 246E and 57Y of the 2D1 heavy chain, 170N and 32T (in 2D1 light chain) and 125S and 94S (in 2D1 light chain). In case of PR1934, the antibody formed three H-bonds between 215P and G100, 212R and 54D and 201Y and 54D all in the 2D1 heavy chain.



**Figure 2:** Docking positions of 2D1 antibody with HA protein monomers from PR1934 and BR2007 obtained using the ZDOCK program. These positions are physically unrealistic orientations because under normal physiological conditions the interfacing residues on the antigen are embedded in the core of HA trimer. The HA proteins are shown in Orange color, Heavy and light chains are colored in Blue and Cyan respectively.

1	SC1918	HEARLIVLLCAFAATNADTICIGYHANNSTDIVDITVLEKGVIVTVHNSHLEDSDHGRKCLKGTAQLQKGNHAGWLLG	80
	PR1934	.K.N.....L..AD.....R.....	
	BR2007	.KVK.....T.T..Y.....N.....L.....H.SV...I..	
	CA2009	.K.I.V...YT..TA..L.....K.....R.V..H.....I..	
81	SC1918	NPECDLLLTASSWSYIVETSNSHNGICYPDFIDYEELREQLSSVSEFEKELFPKTSVSHDDETTKQVTAACSAYAGASS	160
	PR1934	.....P..PVR.....I.....R.....E.....N.NG.....E.....	
	BR2007	.....E..ISKE.....KP.P.....H.A.....R.....E.....TV.G..S.S..PH..K..	
	CA2009	.....E.S.S.....PS.D.....R.....R.....DSN.....PH..K..	
161	SC1918	FYRNLMLTAKDSSYFKSKSYVHKRQEVLVGVVHPPITGTDQQLSYQHADAVSVGSSKYNRRTPETIAARPKVHQ	240
	PR1934	.....F..DQ.....KQ.....K.....HSKE..N.....EN.....VT.....E.....	
	BR2007	.....M.G..N.A..HTEN.....V.....S.K.....K.....	
	CA2009	.....K..I..V.....K.D.....I..S.SA.....F.....SKK.K.....I.....	
241	SC1918	AGRIQVYVTLLEDDTITFEATGHLIAPVYAFALNRGSGSGLIITSDAPVHDQHTKQTPHGAHNSLFPQHHPVTIGEC	320
	PR1934	.....I..N.....M.....S.F.....H.SM.E.....L.....Y.....	
	BR2007	E..I.....I..N.....R.....S.F.....H.N.H..MDK.DA..Q.....V.....	
	CA2009	.....V..K.....VV.R.....ME.NA.....I..T.....T.....K.....T.....I..K..	
321	SC1918	PKYVRSKLEHATGLRHIPSIQSRGLFGAIAQFIEGQNTGMIDGWYGDQHQNEQGSYZAADQKSTQNALDGI DTKVNSVI	400
	PR1934	.....A.....V.....V.....V.....N.....T.....	
	BR2007	.....A.....V.....V.....V.....N.....	
	CA2009	.....K.....L.....L.....V.....L.....E.....	
401	SC1918	EQHQTQF TAUGKEFHNLEERRIEHLMKQDDGFLDITWYNAELLVLIEMERTLDFHDSHURHLVEKVKSQLKDDAKETGHC	480
	PR1934	.....I.....K..K.M.....K.....	
	BR2007	.....K..M.....I.....K.....	
	CA2009	.....H..K.....Y.....K.....R.....	
481	SC1918	CFEYFKCDDAQHESVNRGTIDYPKYSEESKLHREEIDGVKLESHGVYQILAIYSTVASSLVLLVSLGAISSFMCSSHGSLS	560
	PR1934	.....NE.....KV.....	
	BR2007	.....H.E.....K.....K.....	
	CA2009	.....HT.....K.....A.....TRI.....V.....	
561	SC1918	QRCRICI	
	PR1934	.....	
	BR2007	.....	
	CA2009	.....	

**Figure 3:** Multiple alignment of Hemagglutinin protein sequences. Epitopes Sa, Sb, Ca1, Ca2 and Cb are highlighted in Red, Cyan, Green, Pink and Grey respectively. The putative glycosylation sites are indicated in blue-lined boxes. The residue number 1, as referred in the text, corresponds to the residue number 14 in the sequence alignment (Brownlee GG, *et al.* Phil. Trans. R. Soc. Lond. B (2001), 356:1871-1876).

In order to understand the above docking results, we compared the amino acids 61 to 260, forming the head region of the HA proteins of the 1918 and 2009 pandemic H1N1 as well as seasonal H1N1 viruses, for sequence and structural divergence. Between the 1918 and 2009 HAs in the head region, the amino acid identity was 81% while the root mean square deviation (RMSD), a measure of structural divergence, was estimated to be 0.92 Å over the backbone atoms. In CA2009, only a single mutation S159N was observed in the Sa site (**Figure 3**). However, the RMSD between the Sa sites of the two HA proteins was 0.64 Å. Perhaps, S159N and other mutations in the flanking region may account for the structural deviation. Surface charge distribution over the surface of HA may also affect antibody binding. The electrostatic surface potential in the Sa region ranged from +0.6 to +1.2 eV in SC1918 HA and from +0.68 to +1.36 eV in CA2009 HA (**Figure 1C**). The varied surface charge distribution along with the conformational deviation in the Sa site may be responsible for the altered orientation of 2D1 in the CA2009 HA-2D1 complex. The minimized total energy of the SC1918 HA-2D1 co-crystal was -38,910.35 KJ/mol while of the CA2009 HA-2D1 complex it was -38,436.36 KJ/mol. When the Sa site residues of the SC1918 strain were compared to the equivalent site in BR2007 and PR1934, it was seen that there was a 38.5% amino acid divergence in both cases (**Figure 3**). The electrostatic charge distribution on the surface of HA of BR2007 and PR1934 in the Sa region was also compared (**Figure 1C**). It was noted that the potential varied from -0.7 to +1.40 eV in case of PR1934 HA and from -0.56 to +1.11 eV in case of BR2007 HA, indicating that the Sa region in the HAs of the seasonal strains was more acidic with respect to that of the 1918 HA.

The BR2007 HA possessed three additional glycosylation sites when compared to SC1918, of which two, at amino acid positions 129 and 160 fall in the 'Sa' site (**Figure 3**). The presence of these two glycosylation sites at the 'Sa' epitope of BR2007 may provide protection against neutralization by the antibody and thus evasion of the immune system. Further, though the docking program generated a complex of 2D1 bound to HA of strain BR2007, such a complex would not be feasible under natural physiological conditions. This is because, some of the amino acids on the HA forming H-bonds with the antibody actually remained partially buried in the HA pre-fusion trimer and hence are inaccessible for antibodies. PR1934 HA has no additional glycosylation site associated with any of the antigenic sites when compared to SC1918. However, several of the amino acid replacements in the 'Sa' site of PR1934 (**Figure 3**) were non-homologous and included substitutions K156E, G158E, S165K and K166N which may have affected the overall charge distribution in the Sa region. This probably resulted in the docking of the antibody into a totally different region of HA. This region is again not fully exposed in the trimeric structure and therefore the docked complex is not

physiologically feasible. Overall, docking studies of 2D1 onto the CA2009 HA revealed that the antibody 2D1 recognizes the similar region on HA of strain CA2009. Compared to the co-crystal of SC1918 HA - 2D1, we noted a loss of few contacts and a gain of a few in the CA2009 HA - 2D1 docked complex. Notably, the residues 126S, 166K and 167S common to the HA proteins of both SC1918 and CA2009 strains formed H-bonds with the antibody 2D1. The formation of the new H-bonds and salt bridges might have contributed to the stability of the CA2009 HA-2D1 complex so that the minimized total energy of the SC1918 HA-2D1 co-crystal was comparable to that of the CA2009 HA-2D1 complex. On the other hand, the substantial amino acid differences and the resulting altered surface charge distribution in the Sa epitope region of the seasonal H1 HAs can explain the reason for the antibody failing to recognize the Sa site.

## Conclusion:

Our docking studies suggest that the binding of the 1918 H1N1 antibody to the 2009 pandemic H1N1 hemagglutinin is comparable to that with the 1918 virus, both in terms of the specificity to the antigenic site as well as the total energy of the antigen-antibody complex. The structural basis for the antibody not recognizing the HAs of seasonal strains could be explained. The results of the study therefore, reveal the molecular basis for reported cross-reactivity of the 1918 and 2009 pandemic viruses and pre-existing immunity in elderly people.

## References:

- [1] Fraser C *et al. Science* 2009 **324**(5934): 1557 [PMID: 19433588].
- [2] Mishra AC *et al. PLoS ONE* 2010 **5**(5): e10540 [PMID: 20479875]
- [3] Hancock K *et al. N Engl J Med.* 2009 **361**(20): 1945 [PMID: 19745214]
- [4] Krause JC *et al. J Virol* 2010 **84**(6): 3127 [PMID: 20042511]
- [5] Chih-Jen W *et al. Sci Trans Med.* 2009 **2**: 24
- [6] Xing Z & Cardona CJ. *Emerg Infect Dis.* 2009 **15**(11): 1847 [PMID: 19891882]
- [7] Xu R *et al. Science* 2010 **328** (5976): 357 [PMID: 20339031]
- [8] Caton AJ *et al. Cell* 1982 **31**: 417 [PMID: 6186384]
- [9] Potdar VA *et al. PLoS One* 2010 **5**(3): e9693 [PMID: 20300625]
- [10] <http://zdock.bu.edu/>
- [11] Weihe K *et al. Proteins* 2007 **69**: 719
- [12] <http://spdbv.vital-it.ch/>
- [13] <http://noch.sourceforge.net>
- [14] <http://bip.weizmann.ac.il/oca-bin/lpccsu>
- [15] Brownlee GG & Fodor E. *Philos Trans R Soc Lond B Biol Sci.* 2001 **356**: 1871 [PMID:11779386]

Edited by P Kanguane

Citation: Cherian *et al.* Bioinformation 6(1): 35-38 (2011)

**License statement:** This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited.

## Supplementary material:

**Table 1:** Comparison of the H-bonds and salt bridge contacts between interacting residues in the docked CA2009 HA - 2D1 complex and those in the 1918SC HA-2D1 co-crystal. Residues within the 'Sa' antigenic site are indicated in bold font. Residue numbers of the heavy and light chain are prefixed with H and L respectively.

HA 1918SC-2D1				CA2009 HA -2D1			
HA1 residue	2D1 residue	Type	Distance (Å)	HA1 residue	2D1 residue	Type	Distance (Å)
125S	L-91	H-bond	3.0	119E	H-58	H-bond	3.0
126S	L-93	H-bond	3.2	119E	L-95	H-bond	3.2
<b>129N</b>	H-97	H-bond	3.2	123K	H-54	H-bond	2.2
<b>157N</b>	H52	Salt	3.7	123 K	H-54	Salt	2.2
<b>157N</b>	H54	Salt	3.0	126S	L-93	H-bond	3.2
160S	H-97	H-bond	3.0	<b>166K</b>	L-31	H-bond	2.1
160S	H-99	H-bond	2.5	<b>166K</b>	L-93	H-bond	1.9
161Y	H-99	H-bond	3.1	<b>166K</b>	L-93	Salt	1.9
161Y	H100	H-bond	3.1	<b>167S</b>	L-30	H-bond	3.2
<b>163K</b>	H-100	H-bond	2.7	171D	L-94	H-bond	2.7
<b>163K</b>	H-100	Salt	2.7				
<b>166K</b>	L-31	H-bond	2.9				
<b>166K</b>	L-93	H-bond	2.7				
<b>166K</b>	L-93	Salt	2.7				
<b>167S</b>	L-30	H-bond	3.1				