# Resurrecting the Dead - Structural Analysis of Hemagglutinin from the 1918 Influenza Pandemic Strain

James Stevens[1] and Ian A. Wilson[1,2]

*[1]Department of Molecular Biology and [2]Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037*

In 1918, a Great Flu Pandemic swept the world, killing an estimated 20-40 million people, thus making it the largest and most destructive outbreak of any infectious disease in recorded history. This virus was unusual not only in its higher than average mortality rate, but also in the age groups that it attacked. Unlike other outbreaks it targeted not only the very young and old, but also young adults (20-40 year olds). Why the 1918 virus was so devastating is still a mystery. A thorough molecular understanding is now sought to offer reasons why this influenza virus was so pathogenic and how it managed to evade the immune system so effectively. Unfortunately, no intact virus survived, since the pandemic struck before viruses were even identified as the causative agent. Only fragments of the viral genome survived in both Alaskan victims buried in the permafrost and in fixed and archived autopsy material and these have been used to sequence and assemble a number of genes from this virus.

One of the most important viral proteins is the virus coat protein, hemagglutinin (HA). This surface glycoprotein is responsible for virus binding to host cell receptors, and subsequent membrane fusion events within the endosomal pathway in the infected cell. HA is also the most abundant antigen on the surface of the virus and harbors the primary neutralizing epitopes for antibodies. In line with its importance, this viral antigen has been the initial focus of our research on the 1918 flu.

The ectodomain of the HA gene from the 1918 influenza virus[1] A/South Carolina/1/18 was cloned and expressed in a baculovirus expression system as the pre-infective HA0 form (see Figure 1 legend). Protein was produced as trimers using a trimerizing sequence ('foldon') from the bacteriophage T4 fibritin. HA0 crystallized at pH 5.5 and its structure was recently solved by

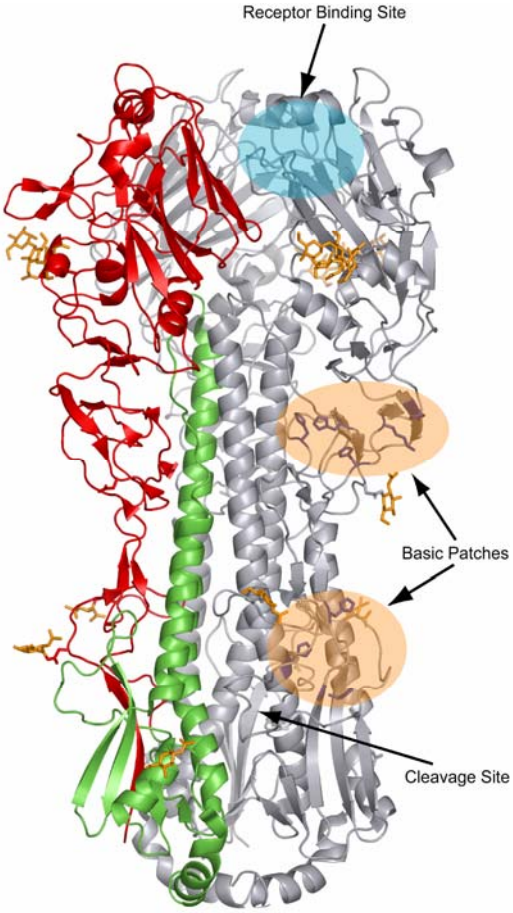


**Figure 1**. Ribbon representation of the hemagglutinin HA0 trimer from the 1918 influenza virus. Each monomer possesses two important sites: 1) the 'Receptor binding site' (blue shade) for virus attachment to the host lung epithelial cells via sialic acid containing host cell receptors. 2) the 'Cleavage site' where for full infectivity, the single chain (HA0) is cut into two chains (HA1 colored red and HA2 colored green). At the N-terminal end of the HA2 chain is the fusion peptide which is critical for subsequent membrane fusion events that lead to infection. The 1918 HA monomer also possesses 2 basic patches (orange shade) which may have contributed to the increased infectivity observed during the pandemic.

molecular replacement (MR) to 3.0 Å resolution from data collected at beamline 9-2 at SSRL (pdb:1RD8)[2]. 18HA0 is ~135 Å in length with two distinct domains (Figure 1). The cylindrical trimer has a tightly intertwined 'stem' domain at its membrane proximal base. The dominant feature of this stalk region is a long triple-stranded coiled-coil. This region also contains the cleavage site where host enzymes cleave HA0 to its infective HA1/HA2 form. The membrane-distal domain consists of a globular 'head' which contains the host receptor binding site and major epitopes for neutralizing antibodies.

Although phylogenetic analyses place the 1918 HA sequence at the base of the evolutionary tree of human viruses, analysis of this structure revealed that it is more closely akin to avian forms. A number of species- and serotype-specific features have been identified:

1. A narrow, avian-like receptor-binding site predominates in the 1918 HA. Within this pocket, the only difference between 1918 HA and known swine-avian adapted H1 viruses is an $Glu^{190}Asp$ mutation, which leads to a minimal increase in the pocket size that could perhaps increase affinity for human receptors.
2. Two previously unobserved histidine-rich basic patches may enhance the pH-dependent viral-fusion event within the endosomal pathway, required for influenza infection. One patch, unique to both human and avian H1, H2 and H5 subtypes, is adjacent to the cleavage site, and may be involved in either trimer destabilization or expulsion of the fusion peptide prior to the membrane fusion event. The second patch situated at the base of the HA1 globular domain is found only in avian H1 subtypes, providing tantalizing evidence of a direct jump of this virus from birds to the human naïve population.

Thus, these and other as yet unidentified features may have contributed to the extraordinarily high infectivity and mortality rates observed during 1918.


1. Reid, A. H., Fanning, T. G., Hultin, J. V. & Taubenberger, J. K. (1999). Origin and evolution of the 1918 "Spanish" influenza virus hemagglutinin gene. *Proc Natl Acad Sci U S A* 96, 1651-6.
2. Stevens, J., Corper, A. L., Basler, C. F., Taubenberger, J. K., Palese, P. & Wilson, I. A. (2004). Structure of the uncleaved human H1 hemagglutinin from the extinct 1918 influenza virus. *Science* 303, 1866-1870.



SSRL is supported by the Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research, and by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program, and the National Institute of General Medical Sciences.