



Published in final edited form as:

*Chem Biol.* 2009 August 28; 16(8): 803–814. doi:10.1016/j.chembiol.2009.08.002.

## Context-Specific Target Definition in Influenza A Virus Hemagglutinin – Glycan Receptor Interactions

Zachary Shriver, Rahul Raman, Karthik Viswanathan, and Ram Sasisekharan<sup>||</sup>

Harvard-MIT Division of Health Sciences and Technology, Koch Institute for Integrative Cancer Research, Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge MA 02139.

### Summary

Protein-glycan interactions are important regulators for a variety of biological processes, ranging from immune recognition to anticoagulation. An important area of active research is directed towards understanding the role of host cell surface glycans as recognition sites for pathogen protein receptors. Recognition of cell surface glycans is a widely employed strategy for a variety of pathogens, including bacteria, parasites, and viruses. We present here a representative example of such an interaction: the binding of influenza A hemagglutinin (HA) to specific sialylated glycans on the cell surface of human upper airway epithelial cells, which initiates the infection cycle. We detail a generalizable strategy to understand the nature of protein-glycan interactions both structurally and biochemically, using HA as a model system. This strategy combines a *top-down* approach using available structural information to define important contacts between glycans and HA, with a *bottom-up* approach using data mining and informatics approaches to identify the common motifs which distinguish glycan binders from non-binders. By probing protein-glycan interactions simultaneously through *top-down* and *bottom-up* approaches we can scientifically validate a series of observations. This in turn provides additional confidence and surmounts known challenges in the study of protein-glycan interactions, such as accounting for multivalency, and thus truly defines concepts such as *specificity*, *affinity*, and *avidity*. With the advent of new technologies for glycomics—including glycan arrays, data mining solutions, and robust algorithms to model protein-glycan interactions—we anticipate that such combination approaches will become tractable for a wide variety of protein-glycan interactions.

### Introduction

Complex glycans at the surface of cells and on circulating signaling molecules play a fundamental role in determining how a cell “sees” and responds to external events. In this capacity, through interactions with proteins, complex glycans modulate a variety of biological processes including pathogen recognition, innate and acquired immunity (Alexopoulou, et al., 2007; Crocker, et al., 2007; Rudd, et al., 2004; van Die and Cummings, 2006), glycoprotein targeting (Bhatia and Mukhopadhyay, 1999; Helenius and Aebi, 2004; Varki, et al., 2008), and adhesion (Kawashima, et al., 2005; Lowe, 2002; Taylor and Drickamer, 2007; Wang, et al., 2005) and trafficking (Crocker, 2005; Smithson, et al.,

© 2010 Published by Elsevier Ltd.

<sup>||</sup>Author to whom correspondence should be addressed: 77 Massachusetts Ave, Building 16-561, Cambridge, MA 02139. Phone: 617-258-9494; Fax: 617-258-9409; rams@mit.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

2001). Obtaining a complete picture of a biological process or understanding higher-level organization of a biological system accordingly requires decoding protein-glycan interactions. In recognition of this practical need, there has been a surge in the development of strategies for chemical and chemo-enzymatic synthesis of diverse glycan structures (Blixt and Razi, 2006; Hanson, et al., 2004; Seeberger and Werz, 2007) that represent the common terminal motifs displayed as part of N- and O-linked glycans and glycolipids. To properly present these glycans to their protein partners, synthetic glycan motifs have been anchored to a variety of platforms including polymeric backbones such as polyglutamic acid and polyacrylamide, on dendrimers, and more recently on glycan microarrays (Bovin, et al., 2004; Collins and Paulson, 2004; Gambaryan, et al., 2006; Mammen, et al., 1995; Totani, et al., 2003).

Despite these advances in the development of platforms to study glycan-protein interactions, challenges remain in defining glycan targets for the purpose of bridging the biochemical and biophysical specificity of glycan-protein interactions with the biological functions modulated by these interactions. In the case glycan-protein interactions, one size does not fit all. Strategies which have proven informative in the areas of genomics or proteomics may or may not facilitate our description of glycome biology for several reasons. Glycan-protein interactions, leading to either the activation or inhibition of a biological response, are often not binary but rather involve more subtle mediation of a signaling pathway. In addition, glycan-protein interactions typically involve *multivalency* with regards to both the protein and the glycan, 1:1 monovalent complexes are often weak and display dissociation constants on the order of 1-1000  $\mu\text{M}$ . Biochemical/biophysical descriptions of protein-glycan interactions therefore depend both on context and experimental design. In this framework, the careful description of experimental results, including an understanding of both the strengths and limitations of an approach is appreciated. Finally, it is becoming clear that there are both finer and coarser determinants to the specificity of a given protein-glycan interaction than the simple monosaccharide sequence of a glycan. Describing a glycan-binding sequence by its monosaccharide composition and the linkages between the monosaccharides, while important, in most cases does not afford the same descriptive capacity as it does for other biopolymers.

Ensuring an accurate and complete picture of protein-glycan interactions in spite of these difficulties demands studies that integrate the basic biochemistry and molecular biology of the system with analytical approaches, while simultaneously enabling the appropriate translation to the biology. One of the best-studied systems of glycan-protein interactions modulating a biological function is that of the role of influenza A virus hemagglutinin (HA) – glycan interactions for the pathogenesis of the virus. This review presents an overview of the various technologies that have been used for glycan receptor target definition for HA, their limitations, and how an integrated framework enabled the bridging of HA-glycan interactions with the host adaptation of the virus. This integrated approach could serve as a framework for target definition in other protein-glycan interactions.

## Overview of Influenza A Viruses

Influenza A virus is a negative strain RNA virus with eight gene segments. Three of the genes: hemagglutinin (HA), neuraminidase (NA) and the polymerase (PB) have been shown to be critical for infection and human-to-human transmission (Palese, 2004; Pappas, et al., 2008; Tumpey, et al., 2007) (Figure 1). Within influenza A, five of the genome segments, encoding the nucleocapsid protein (NP), the matrix proteins (M1 and M2), the nonstructural proteins (NS1 and NS2) and polymerase proteins (PB1, PB2 and PA) have maintained a relatively unbroken evolutionary history in humans. In contrast, the two genes encoding the major cell surface proteins (HA and NA) have been subjected to substantial evolutionary

pressure, including mutation (*antigenic drift*) and wholesale reassortment (*antigenic shift*). Due to their variability, strains of influenza virus are identified based on their serotype of HA and NA. There are currently 16 known serotypes of HA and nine of NA.

At the moment there is substantial public health concern surrounding the prospect of another influenza A pandemic and its associated potential global implications. The past century has seen four influenza pandemics. The first and most severe occurred in 1918, involved an H1N1 virus, and led to the death of at least 40 million people worldwide. Less serious pandemics occurred in 1957, 1968, and 1977. Notably, real-time has been the advent of a reassorted H1N1 virus, *viz.*, 2009 H1N1 or “swine” flu, which is antigenically dissimilar from seasonal, circulating H1N1s and which has already been declared a pandemic by the World Health Organization (WHO). Viruses containing the 16 HA and 9 NA serotypes are naturally present in wild aquatic bird populations where they exist commensally without causing disease, allowing birds to become a reservoir for influenza strains. This is of specific concern since, the influenza pandemics of the last century which arose from H2N2 (1957) and H3N2 (1968) were avian-human reassortments which resulted in the humanization of an avian virus and efficient human-to-human transmission. Those genetic reassortments which led to an avian-to-human switch yield a number of important scientific and medical questions, not the least of which is what changes lead away from infectivity and propagation in avian species and towards human transmissibility? In light of a particular influenza strain, H5N1 or the so-called “bird flu”, addressing these questions becomes more critical. Transmission of avian H5N1 influenza viruses to humans thus far has been observed thus far only upon direct contact with infected poultry; the virus has not yet demonstrated human-to-human transmission ability. Given that human infectivity has occurred and appreciating that this virus strain is highly lethal (estimates are as high as ~60% infected individuals), the import of comprehending the mechanism and specificity of viral entry and infection as well as identifying additional strategies for intervention is clear.

## HA-Glycan Interactions: Description and Biological Importance

Influenza A has a three-stage infection cycle (Figure 1). The first stage is the attachment of HA of the virus to complex glycan receptors on the host cells. Following attachment, the virus is internalized by endocytosis where structural changes in HA produce the fusion of the viral membrane with the endosomal membrane. This facilitates activation of the ion channel activity of M1 and transport of the viral RNA to the nucleus. In the nucleus, viral RNA undergoes replication and transcription. Newly synthesized proteins HA and NA are secreted through the Golgi to the cell surface. Other proteins are transported to the nucleus where they associate with the synthesized transcripts of viral RNA to form virions. These virions, with HA and NA on their surface, then bulge from the cell membrane. The action of NA, which cleaves the sialic acid-capped glycan structure and eliminates the interaction of the host cell glycans with the newly formed virus particle, facilitates the release of progeny viruses from the host cells.

In the context of infectivity, the HA-glycan interaction is one of the most critical components governing virus selectivity. HA itself is a homotrimer (Figure 2); each monomer is synthesized as a single polypeptide that contains a proteolytic site, cleaved by host enzymes into two subunits (HA1 and HA2). Numerous crystal structures of different HAs have been solved, both alone and as co-crystals comprising various glycan structures (Eisen, et al., 1997; Gamblin, et al., 2004; Ha, et al., 2003; Ha, et al., 2001; Sauter, et al., 1992; Skehel and Wiley, 2000; Stevens, et al., 2006; Weis, et al., 1988; Yamada, et al., 2006). This body of structural work offers valuable insights on the specificity of HA binding. In tandem, biochemical studies identified a key feature for binding of human-adapted HA to glycans terminated by N-acetyl neuraminic acid linked  $\alpha 2 \rightarrow 6$  to galactose

(Neu5Ac $\alpha$ 2 $\rightarrow$ 6Gal, hereafter referred to as  $\alpha$ 2 $\rightarrow$ 6) present on human respiratory epithelia (Ibricevic, et al., 2006; Russell, et al., 2006; Shinya, et al., 2006; Skehel and Wiley, 2000; van Riel, et al., 2007). These biochemical findings have been correlated with the finding that human respiratory tissue contains epithelial cells with  $\alpha$ 2 $\rightarrow$ 6 sialic acid capped glycans (sites for attachment of human-adapted viruses), whereas cells that primarily express glycans terminated by N-acetyl neuraminic acid linked  $\alpha$ 2 $\rightarrow$ 3 to galactose (Neu5Ac $\alpha$ 2 $\rightarrow$ 3Gal, hereafter referred to as  $\alpha$ 2 $\rightarrow$ 3), such as the alveolar cells, are sites for attachment of avian-adapted viruses (Shinya, et al., 2006). These and other findings suggest that for a virus to crossover from avian species to humans, its HA must switch binding preference from  $\alpha$ 2 $\rightarrow$ 3 sialylated glycan receptors, present in avian species, to  $\alpha$ 2 $\rightarrow$ 6 receptors (Connor, et al., 1994; Glaser, et al., 2006; Kumari, et al., 2007; Matrosovich, et al., 2007; Matrosovich, et al., 2004; Rogers, et al., 1983; Russell, et al., 2006; Tumpey, et al., 2007), present in human upper airway epithelia. Recent developments question the conception that the preference of HA binding to  $\alpha$ 2 $\rightarrow$ 3 or  $\alpha$ 2 $\rightarrow$ 6 alone is sufficient to designate human- *versus* avian-adapted HA (Bewley, 2008).

## HA-Glycan Interactions: Answering Questions and Generating Others

The specificity of HA binding to cell surface glycan structures has been addressed not only by crystal structure analysis but also by the advent of powerful biochemical and biological tools which provide significant insight and raise additional questions regarding HA-glycan interactions.

### Tools to study HA-glycan interactions

A variety of biochemical methods have traditionally been used to characterize the specificity of HA-glycan interactions. The earliest method, still in use today, for probing the glycan-binding specificity of a particular virus (through its HA) involves measuring its ability to agglutinate red blood cells (RBCs). RBCs from species such as chicken, turkey, horse, guinea pigs and humans have been used (Connor, et al., 1994; Paulson and Rogers, 1987; Rogers and Paulson, 1983; Tumpey, et al., 2007; Yang, et al., 2007). Equine RBCs primarily contain  $\alpha$ 2 $\rightarrow$ 3 glycans at their cell surface whereas turkey and guinea pig RBCs primarily contain  $\alpha$ 2 $\rightarrow$ 6 glycans. To augment the usefulness and specificity of these models, the agglutination method has been modified to include a step of complete desialylation of species-specific RBCs followed by specific resialylation by either  $\alpha$ 2 $\rightarrow$ 3 or  $\alpha$ 2 $\rightarrow$ 6 sialyltransferase, creating a more specific assay (Carroll, et al., 1981). Agglutination and other traditional hemagglutination assays permit the definition of viruses' specificity according to the sialic acid linkage, as described above. These assays prove useful in the classification of virus specificity and virulence. In its present form this assay does not allow the examination of specificity beyond the sialic acid linkage—probing the fine specificity of HA binding is therefore not possible. There are other drawbacks to the use of RBC agglutination to define glycan-binding specificity of a virus strain. First, there is likely substantial variability in the N- and O-linked cell surface glycans between different batches of RBCs and related potential for variable response. Additionally, the glycan structures on the cell surface of RBCs are not similar to those present on the surface of cells of the upper airway. Viruses may bind to, and elicit agglutination through, glycan receptors that are not present in the upper airway and therefore not physiologically relevant.

Subsequent development of solid-phase fetuin capture assays provided a wealth of information on the glycan binding properties of influenza A viruses. In these assays the viruses are immobilized on fetuin-coated surfaces and their binding to various sialylated glycans (including polyvalent compounds) is evaluated (Gambaryan, et al., 2005; Gambaryan and Matrosovich, 1992; Gambaryan, et al., 1995; Gambaryan, et al., 2008; Matrosovich, et al., 2000). The presentation of the viruses is heterogeneous since the amount

of virus captured on the plate depends on the binding of the viral HA to the sialylated glycans on fetuin. Furthermore, measuring the binding of fixed viruses to glycans in solution is opposite to the physiological event where glycans are less mobile on the cell surface as compared to the viruses.

While not commonly used to investigate HA-glycan interactions, isothermal titration calorimetry (ITC) and surface plasmon resonance (SPR) have been widely employed to determine equilibrium binding affinity constants and thermodynamic parameters for glycan-protein interactions (Dam, et al., 2009; Duverger, et al., 2003; Karamanska, et al., 2008). These methodologies thus offer promising tools for quantifying the binding affinity of HA-glycan interactions.

Recent advances in the chemical and chemo-enzymatic synthesis of glycans have yielded the development of glycan array platforms. These platforms consist of hundreds of synthetic glycan motifs (that are typically present on N- and O-linked glycoproteins and glycolipids) displayed on the surface of the array. Multiple types of arrays have been developed that utilize different strategies including the formation of neoglycolipids (Fukui, et al., 2002), neoglycoproteins (Gildersleeve, et al., 2008; Huang, et al., 2008), or the direct application of glycans to various surfaces (Blixt, et al., 2004; Grun, et al., 2006; Karamanska, et al., 2008; Liang, et al., 2008; Mercey, et al., 2008; Xia, et al., 2005). Several fine reviews outline advances that have occurred recently in the development of array technologies (Blixt, et al., 2004; Feizi, et al., 2003; Houseman and Mrksich, 2002; Liang, et al., 2008; Oyelaran and Gildersleeve, 2007; Seeberger and Werz, 2007; Stevens, et al., 2006). Among the various glycan array platforms, those developed by the Consortium for Functional Glycomics (CFG) are arguably the most accessible. More than two thousand samples have been screened on CFG glycan arrays by the scientific community and data generated on the CFG glycan array platform is disseminated freely to the public via web-based interfaces ([www.functionalglycomics.org/static/consortium/resources/resourcecoreh.shtml](http://www.functionalglycomics.org/static/consortium/resources/resourcecoreh.shtml)). Recent studies have begun to adapt these technologies towards the ultimate presentation of *natural* glycans by harvesting glycans from the surface of cells and imprinting these on a glycan array format (Song, et al., 2009), thus allowing one to probe the glycan repertoire of a biological system. In this manner, one could envision arrays specific to a disease or biological process that permit the interpretation of glycan-binding information as a function of a biological process.

The glycan array platforms developed by the CFG have already been used to screen wild-type and mutant forms (mutations in HA) of intact viruses and recombinant HAs belonging to the H1, H2, H3, H5, H7 and H9 subtypes (Belser, et al., 2008; Kumari, et al., 2007; Stevens, et al., 2008; Stevens, et al., 2006; Stevens, et al., 2006; Wan, et al., 2008). These studies have increased understanding of HA binding to sialylated glycan receptors by mapping the effect of substitutions such as sulfation and fucosylation on the HA-glycan interactions. The scope of most of these studies is intended to serve as a primary screen (of the glycans on the array) where high viral titers or HA concentrations are used to define the binding preference of HA in terms of  $\alpha 2 \rightarrow 3$  and/or  $\alpha 2 \rightarrow 6$  binding.

### Tools to bridge HA-glycan interactions with viral pathogenesis

In 1989 Palese and colleagues first demonstrated the ability to manipulate the influenza virus genome by developing a system that allowed the use of standard recombinant DNA technology to modify the genome of influenza virus to express foreign genes (Luytjes, et al., 1989). Their advance formed the base for the development of methods—termed reverse genetics—which permit synthesis of the whole virus from the cDNAs of individual virus genes (Hoffmann, et al., 2000; Pleschka, et al., 1996) (Figure 3). In addition to the advent and widespread adoption of reverse genetics, the emergence of ferrets as a model system to



study pathogenesis and contact and respiratory droplet modes of transmission (Lowen and Palese, 2007; Tumpey, et al., 2007; van der Laan, et al., 2008) has proven of equal importance to the investigation of various aspects of influenza biology. Studies have demonstrated that ferrets possess similar glycan structures to humans, including a predominance of human-like  $\alpha 2 \rightarrow 6$  glycans in their upper respiratory tract epithelium (Maines, et al., 2006).

The ability to completely reconstruct the pandemic 1918 H1N1 viruses through reverse genetics and test its virulence in ferrets permitted a systematic exploration of the roles for various viral genes in the virulence and transmissibility of influenza A strains (Palese, 2004; Tumpey, et al., 2005). Single gene reassortants of the highly virulent pandemic human H1N1 (A/South Carolina/1/18 or SC18) virus with a contemporary epidemic human H1N1 (A/Texas/36/91 or Tx91) virus showed that the HA of SC18 had the most profound effect on the virulence of the reassorted viruses, followed by the NA and PB1 genes (Pappas, et al., 2008). More recently, gene reassortment and reverse genetics were employed to demonstrate that HA and PB2 are the two critical genes conferring viral transmissibility via respiratory droplets in ferrets (Van Hoeven, et al., 2009).

In one study pertaining to the SC18 virus, either a single point mutation (NY18) or two point mutations (AV18) in HA resulted in a virus that was unable to transmit efficiently via respiratory droplets in the ferret model. This study permitted investigation of the relationship between glycan binding property of the HA of these 1918 H1N1 viruses (the other genes are identical between SC18, NY18 and AV18) and their airborne transmissibility. The RBC agglutination assay described above was used to characterize SC18 and determined that it was an  $\alpha 2 \rightarrow 6$  binder, NY18 was a mixed  $\alpha 2 \rightarrow 6 / \alpha 2 \rightarrow 3$  binder and AV18 was an  $\alpha 2 \rightarrow 3$  binder. The change in binding preference of NY18 and AV18 HA in combination with the invariant NA can potentially influence the tissue tropism and virulence of the viruses. NY18 and AV18 are more likely to infect deep lung tissues which preferentially express  $\alpha 2 \rightarrow 3$  glycans. Given that the NA of these  $\alpha 2 \rightarrow 3$  glycan-binding viruses remains the same as that of SC18, it might be inefficient at releasing the viral particles from the deep lung tissues leading to lower virulence of these mutant viruses. When coupled with the fact that this virus transmitted inefficiently, the binding specificity of NY18 HA leads to the conclusion that loss of  $\alpha 2 \rightarrow 3$  binding is necessary for efficient transmission but that the gain of  $\alpha 2 \rightarrow 6$  is insufficient. In an apparent contradiction to this conclusion, Tx91—also a mixed  $\alpha 2 \rightarrow 3 / \alpha 2 \rightarrow 6$  binding H1N1 virus—is able to transmit efficiently (Tumpey, et al., 2007). In other studies using H7 and H9 viruses (Belser, et al., 2008; Wan, et al., 2008), it was observed that although some of the wild-type or HA mutant viruses showed substantial  $\alpha 2 \rightarrow 6$  binding, none transmitted via respiratory droplets in the ferret model. The control human adapted H3N2 viruses used in these studies which showed similar  $\alpha 2 \rightarrow 6$  binding did however transmit efficiently. These studies suggest that defining glycan receptor targets of HA in terms of  $\alpha 2 \rightarrow 3 / \alpha 2 \rightarrow 6$  alone (using RBC agglutination, glycan arrays, etc.) is inadequate to bridge with the biological role of HA in human adaptation and viral virulence. In light of the critical role of HA with respect to virulence and human adaptation of the virus, it is clear that additional factors govern the specificity of HA binding and determine whether a virus is able to efficiently infect or transmit in humans.

## **Integrating a Bottoms-Up and a Top-Down Approach to Bridge Structure to Biology**

Addressing the above issues required an integration of information from complementary approaches that permit a move from the biological to the structural space in an iterative and transitive fashion (Figure 4). Each endeavor addressed a series of specific questions and collectively the data provides complementary, over-lapping sets of information which define

the underlying specificity of influenza A HA. The *bottom-up approach* commenced with an examination of the structure of glycans present on the cell surface of human airway epithelial cells and then using this information to quantitatively probe the affinity of HA to individual glycan species.

Because human upper respiratory epithelia are the primary targets for human-to-human transmission of influenza A virus, it is essential to identify the types of sialylated glycans present on these tissues. The characterization of the glycan repertoire for a given tissue or cell type is challenging due to the complexity associated with accessing and elucidating these structures. It is recognized that sialylated glycans of epithelial cells are the receptor for HA (and accordingly virus) binding and internalization but epithelial cells themselves possess a broad array of structurally diverse glycan structures. Addressing this diversity requires orthogonal biochemical and biological tools which are specific, sensitive, and robust. One approach that has been employed is to bridge lectin staining of different tissue sections with glycan profiling of representative human cell lines using MALDI-MS and MS/MS (TOF-TOF) analyses (Comelli, et al., 2006; Uematsu, et al., 2005). Other information-rich techniques have also been employed, including capillary HPLC, LC-MS, and NMR (Manzi, et al., 2000; Norgard-Sumnicht, et al., 2000). Regardless of the particular analytical techniques employed, the use of these complementary approaches provides a superior quality of information because (1) data sets can be crosschecked with one another and (2) no single analytical approach, no matter how powerful, has equally descriptive power for all structures, especially glycans. Notably, in one approach, *Sambucus nigra* agglutinin (SNA-I) a prototypic lectin that specifically binds to  $\alpha 2 \rightarrow 6$  glycans and Jacalin and Concanavalin A (Con A), which bind to specific motifs in O-linked (-Gal $\beta$ 1-3GalNAc $\alpha$ - and -GlcNAc $\beta$ 1-3GalNAc $\alpha$ -) and N-linked (trimannosyl core) glycans, respectively are used alone or in combination to probe tissue samples (Chandrasekaran, et al., 2008). Using this lectin matrix in an iterative manner provides detailed information about glycan distribution. In this case, it was shown that there is a widespread distribution of N-linked  $\alpha 2 \rightarrow 6$  (on ciliated cells) glycans and localized distribution of O-linked  $\alpha 2 \rightarrow 6$  glycans on goblet cells in human tracheal epithelia (Chandrasekaran, et al., 2008). This analysis was then confirmed and extended using detailed MALDI analysis of glycans derived from representative human upper epithelia cells to identify glycan composition, and MS/MS fragmentation to determine their structural features, including the presence of a lactosamine extension. Thus, MS and MS/MS information is interpreted in the context of the information obtained from lectin analysis (Chandrasekaran, et al., 2008). This is but one example/strategy of the application of different, overlapping analytical techniques to obtain complementary data sets; other, equally valid, strategies can be devised.

Importantly, this coupled approach can provide a powerful complement to the construction of natural arrays and increase their utility to identify glycan-binding partners to proteins. Alternatively, such analysis can provide an important framework for the development of quantitative array and array-like binding assays using selected synthetic structures which are representative of cell surface glycans. For an analysis of this nature to be valid, three key variables must be addressed and controlled to ensure accurate interpretation. First, as is the case in biological systems, glycan-protein interactions are typically multivalent and the strength of such a contact is described based on its avidity. Numerous experimental setups have been proposed to measure avidity, Kiessling and colleagues provide an excellent overview of many of these (Kiessling, et al., 2000). Regardless of platform, the most useful quantitative information is obtained when  $N$ , the number of binding sites is known, though even without  $N$ , useful thermodynamic information can be determined (Mammen, et al., 1998) (Figure 5). The spatial arrangement of glycans and the glycan-binding sites on proteins influence the structural valency (i.e. number of binding sites on a protein that are occupied by the glycan motifs) of glycan-protein interactions (Dam and Brewer, 2008; Dam,

et al., 2009). The spatial arrangement of glycan motifs depends on the extent of branching (bi, tri, tetra, penta-antennary N-linked glycans) and spatial distribution of glycosylation sites (clustered, linear and globular) in different glycoprotein structures. The spatial arrangement of glycan-binding sites in a protein instead depends on the quaternary association of individual domains. In the case of HA, there is a homotrimeric association of identical domains resulting in three glycan-binding sites per trimeric HA unit. Second, similar to biological systems, the glycan structures on the surface should be in excess. If this is not the case ambiguous results may be obtained, particularly if this situation is not made explicit when plotting and calculating an association constant. Finally, the response should be measured at various concentrations to determine dissociation constant ( $K_d$  which is inverse of the affinity constant,  $K^N_{poly}$ ).

In most of the earlier studies which focused on screening different HAs on glycan arrays, the binding event was designated as a single point “on” or “off”. This designation, while potentially useful, necessarily misses the context of the interaction and the relative biological importance of the interaction. More recently, biochemical assays have been designed to screen HA-glycan interactions over an entire range of HA concentrations (Srinivasan, et al., 2008) that take into consideration the aforementioned factors including spacing of the glycan motifs in an array platform and the spatial arrangement of glycan-binding sites. These assays have permitted quantification of the relative binding affinity of HA to different glycan motifs. Using such quantitative assays, it has been demonstrated that the human adapted HAs share a high binding affinity to  $\alpha 2$ -6 glycans (particularly those that have multiple lactosamine repeats) that in turn correlate with their efficient airborne transmissibility (Maines, et al., 2009; Srinivasan, et al., 2008). Viruses of avian or swine origin although known to infect humans could be distinguished from the human-adapted viruses based on their quantitative  $\alpha 2$ -6 binding affinity of their respective HAs (Chandrasekaran, et al., 2008; Maines, et al., 2009).

Analysis employing the *top-down* approach originates from consideration of the molecular and structural aspects of HA-glycan interactions and correlates these aspects with the biochemical binding affinities. Using glycan conformational analysis, examination of HA-glycan co-crystal structures demonstrated that glycan topology (or three-dimensional shape of the glycan) plays a critical role in distinguishing binding of  $\alpha 2 \rightarrow 3$  and  $\alpha 2 \rightarrow 6$  glycans to avian and human adapted HAs (Bewley, 2008; Chandrasekaran, et al., 2008). Analysis of the various crystal structures indicated that a highly conserved set of amino acids Tyr98, Ser/Thr136, Trp153, His183, Leu/Ile194 (numbered based on H3 HA) across different HA subtypes are involved in anchoring the sialic acid. The specificity of HA to either  $\alpha 2 \rightarrow 3$  or  $\alpha 2 \rightarrow 6$  is governed by an extended range of interactions within the glycan-binding site—not only with the sialic acid, but also with the glycosidic oxygen atom and monosaccharides beyond sialic acid.

The ensemble of conformations sampled by the  $\alpha 2 \rightarrow 3$  and  $\alpha 2 \rightarrow 6$  glycans in the binding site of HA was described using a shape-based topological description. In the case of  $\alpha 2 \rightarrow 3$  glycans, the conformations sampled by the Neu5Ac $\alpha 2 \rightarrow 3$ Gal linkage (keeping the Neu5Ac anchored) and the sugars beyond this linkage (at the reducing end) span a region on the binding surface of HA that resembles a “cone.” The assembly of these conformations is therefore described by the term *cone-like* topology (Figure 6).

When contrasted with the Neu5Ac $\alpha 2 \rightarrow 3$ Gal linkage, the presence of the C6-C5 bond within the Neu5Ac $\alpha 2 \rightarrow 6$ Gal linkage, provides additional conformational flexibility. The different conformations sampled by Neu5Ac $\alpha 2 \rightarrow 6$ Gal linkage (keeping the Neu5Ac anchored) and the sugars beyond this linkage (at the reducing end) thus span a wider region on the HA binding surface. One part of this wider region is similar to the cone-like surface



and the other part resembles a space that is readily described by the opening of an umbrella from a fully folded to a fully open form. In contrast to the cone-like topology, the set of conformations that sample this other region is better described using the term umbrella-like topology (Figure 6). In this case, the stem of the umbrella is occupied by the Neu5Aca2-6Gal- motif and the spokes of the umbrella (that are the flexible part causing the opening and closing) are occupied by the sugars at the reducing end of Gal.

The defining characteristic of glycan conformations that span a cone-like topology is that the majority of the interactions with the HA are made by a three-sugar (or trisaccharide)  $\alpha 2 \rightarrow 3$  (Neu5Aca2 $\rightarrow$ 3Gal $\beta 1 \rightarrow 3/4$ GlcNAc-) or  $\alpha 2 \rightarrow 6$  (Neu5Aca2 $\rightarrow$ 6Gal $\beta 1 \rightarrow 4$ GlcNAc-) motif. On the other hand, the glycan conformations that sample the umbrella-like topology are such that longer oligosaccharides (beyond a trisaccharide) make substantial contacts with the binding site on HA.

Using these shape-based definitions of the flexible glycan conformation, it was shown that the umbrella-like topology is predominantly adopted by  $\alpha 2 \rightarrow 6$  glycans which possess at least 4 sugars including the Neu5Ac, for example, poly-lactosamine branches terminated by  $\alpha 2 \rightarrow 6$  linked Neu5Ac (long  $\alpha 2 \rightarrow 6$ ). The cone-like topology can instead be adopted by both  $\alpha 2 \rightarrow 3$  and  $\alpha 2 \rightarrow 6$  glycans. In the case of  $\alpha 2 \rightarrow 6$  glycans, those with a three-sugar  $\alpha 2 \rightarrow 6$  motif such as N-linked glycans having single lactosamine branches terminated by  $\alpha 2 \rightarrow 6$  linked Neu5Ac (short  $\alpha 2 \rightarrow 6$ ), are more likely to adopt cone-like topology as compared to the long  $\alpha 2 \rightarrow 6$  branch.

Applying the topology-based description of glycan conformation to analyze HA-glycan co-crystal structures revealed that the umbrella-like topology was characteristic of long  $\alpha 2 \rightarrow 6$  motifs interacting with human adapted H1 and H3 HAs and the cone-like topology was characteristic of  $\alpha 2 \rightarrow 3$  glycans interacting with avian HA. Since the avian HA binding pocket is well-suited for maximum contacts with cone-like topology, it is probable that  $\alpha 2 \rightarrow 6$  glycans, particularly those with short  $\alpha 2 \rightarrow 6$  motifs will adopt cone-like topology in the glycan binding site of avian HA. More recently, the role of glycan topology in molecular HA-glycan interactions was investigated using molecular dynamics (MD) simulations (Xu, et al., 2009). Findings from this study support the notion that the umbrella-like topology is energetically preferred by glycans upon binding to human adapted HA. The adoption of umbrella-like topology by a long  $\alpha 2 \rightarrow 6$  motif in an avian H5 HA binding site was associated with a high conformational entropy penalty in comparison with the adoption of this topology in the human adapted H3 HA binding site. These studies have further extended the definition of glycan topology using other parameters such as glycosidic torsion angles of sugars beyond the terminal sialic acid linkage and volume occupied by the different glycan topologies.

The above molecular and structural analyses of HA-glycan interactions offer a framework for the interpretation of glycan array data (Figure 4). Data mining tools have been developed to identify patterns among binders (candidate ligands identified on the array) and non-binders. These patterns are defined using glycan features that are abstracted from the glycan structures tested, for example as part of a glycan array. The combination of patterns from binders and non-binders provide rules, or classifiers, that define glycan binding motifs for a given HA. These classifiers can be used to corroborate the structural aspects of glycan-HA interactions obtained from analysis of the co-crystal structures. For instance, the  $\alpha 2 \rightarrow 3$  classifiers define motifs in terms of substitutions around a trisaccharide Neu5Aca2 $\rightarrow$ 3Gal $\beta 1 \rightarrow 3/4$ GlcNAc $\beta 1 \rightarrow$  whereas  $\alpha 2 \rightarrow 6$  classifiers define length-dependent motif that terminates in Neu5Aca2 $\rightarrow$ 6Gal $\beta 1 \rightarrow 4$ GlcNAc $\beta 1 \rightarrow$  structure.

Additionally, from an analysis of crystal structures and docking of glycan structures, it is possible to estimate,  $\Delta G_{mono}$ , or the strength of a monovalent glycan-protein interaction (Figure 5). Several advances have improved the force fields used to model glycans and protein-glycan interactions (Case, et al., 2005; Kirschner, et al., 2008). These theoretical methods have become remarkably effective at predicting carbohydrate 3D structures, since the advent of accurate carbohydrate-specific force fields (Woods, 1998) and the advances in timescales that are accessible to molecular dynamics (MD) simulations. All commonly employed biomolecular simulation packages (AMBER, CHARMM, GROMOS, etc) now include high-quality carbohydrate force fields. There are currently several methods to calculate binding free energies ( $\Delta G$ ). The most accurate of these computes relative binding affinities for two related ligands by mutating one molecule onto another. Such “computational alchemy” can be performed by thermodynamic integration, but is limited to predicting relative binding energies for structurally similar ligands. While the predicted interaction energies from these docking methods may only be qualitative, the resultant structures of the ligand-protein complexes may well be chemically accurate.

In conclusion, the recent development of biochemical and biophysical tools truly enables the scientific community to rapidly and thoroughly address some fundamental aspects of *describing* glycan-protein interactions. By combining quantitative binding affinity information obtained from the array with molecular simulations, it is possible to correlate changes in the enthalpic and entropic contributions of the monovalent glycan-protein interaction with the quantitative differences in relative glycan-binding affinity. By identifying biochemical systems where multivalency can be taken into account (and hence define  $N$  in a rigorous manner), it is possible to define *avidity* and provide additional context to the glycan-protein interactions, which can be carried through to biological studies.

## Framework for Role Of Ha-Glycan Interactions in Influenza A Virus Biology

The intersection of a *bottom-up* and *top-down* approach highlighted the fact that, for a virus to efficiently infect humans, it must bind to glycans that can adopt the umbrella-like topology (for example sialylated structures containing a lactosamine repeat, *i.e.* long  $\alpha 2 \rightarrow 6$ ). The structural topology of the glycan and not just the linkage appears to govern HA binding specificity. This framework takes into account not only glycan structure in the context of protein binding but also resolves the apparent inconsistencies between the binding and transmission data for SC18, NY18, AV18, and Tx91. SC18 and Tx91, both efficient transmitters, bind with high affinity to glycans that adopt an umbrella-like topology, whereas NY18 and AV18, which are inefficient transmitters, bind with higher affinity to cone-like glycans, including glycans that either have  $\alpha 2 \rightarrow 3$  or  $\alpha 2 \rightarrow 6$  sialic acid. This framework demonstrates that description of HA-glycan interactions based on *trans* and *cis* conformations (adopted by  $\alpha 2 \rightarrow 3$  and  $\alpha 2 \rightarrow 6$  linkages respectively) alone does not fully capture the structural features and conformational flexibility of the diverse sialylated glycans observed in human tissues. However, cone-like and umbrella-like classifications are able to fully capture the conformational plurality of these glycans and their binding to HA. Significantly, these classifications are able to distinguish the  $\alpha 2 \rightarrow 3$  and  $\alpha 2 \rightarrow 6$  binding of avian HAs from that of the  $\alpha 2 \rightarrow 6$  binding of human adapted H1 and H3 HAs. Moreover, H1, H5, H9 are of different structural clades and hence have distinct spatial arrangement of residues in the glycan-binding sites. The conformational flexibility of the  $\alpha 2 \rightarrow 6$  permits different types of umbrella-like topologies in the glycan binding sites to accommodate the diverse constraints imposed by the different structural clades of HA. These observations suggest that it is challenging to design mutations in H5 and H7 HA based simply on the characteristic changes in a few residues in H1 and H3 HAs that lead to their human adaptation. A set of mutations must instead occur which accommodate umbrella-like

glycans in the context of the binding pocket, which is subtly distinct for each HA structural clade.

## Conclusion

Lessons learned from the study of glycans, and particularly the HA example presented here highlight: (1) the need to employ multiple biophysical, bioanalytical and biological approaches to ultimately define the binding specificity of HA and (2) the fact that as different cell types have a distinct cell surface glycan repertoire, studies employed using cell culture or any animal model must be carefully interpreted, as glycan structures present on cells in culture or in a given animal, may not be reflective of the structures present on the surface of primary cells within the tissue. For example, the use of mice (from above), and Madin-Darby canine kidney (MDCK) cells to characterize virus infectivity (Hatakeyama, et al., 2005; Stray, et al., 2000) may not be ideal owing to differences in their glycan receptors present in them as compared to physiologically relevant cells.

The development of a systematic understanding of the receptor binding specificity of the HA from various influenza strains was anticipated to help address a number of critical questions, including what defines an avian influenza strain versus one that has become humanized, and what mutations in HA enable the conversion of a strain from an avian virus to one capable of efficient human-to-human transmission. Understanding the receptor specificity of HA as well as the set of mutations which allow a virus to gain the ability to recognize human-like glycans of the upper respiratory tract, provides insight into the epidemic and pandemic potential of various strains. Addressing these questions is not only of great scientific significance but also is immediately pertinent in light of the rise of H5N1 and the specter that it could become fully “humanized” through mutations. The application of approaches outlined in this review, as well as others, should illuminate novel strategies for the development of vaccines and/or therapeutics.

Recent developments in the glycomics analysis of influenza in many respects, offer an example for future studies in glycobiology. The role of glycans in cellular events, where modulation tends to be more a function of avidity and presentation (context) rather than a simple on/off event, integration of biochemical/structural/data mining/ and *in vivo* studies is critical to ensure accurate interpretation and extension of findings. With the advent of analytical, synthetic, and computational tools to interrogate protein-glycan interactions, it is possible to design a framework such as that presented here for HA, to study many such systems. We anticipate that the availability of such tools, through the efforts of large-scale glycomics research initiatives as well as of individual researchers, will dramatically increase our understanding of the glycome and its role in fundamental biology.

## Acknowledgments

The authors would like to acknowledge support from National Institute of General Medical Sciences of the National Institutes of Health (GM 57073 and U54 GM62116 to RS) and the Singapore– Massachusetts Institute of Technology Alliance for Research and Technology (SMART). The authors thank Dr. V. Sasisekharan for critical reading of the manuscript and constructive discussion.

## Abbreviations Used

<b>HA</b>	Hemagglutinin
<b>NA</b>	Neuraminidase
<b>RBC</b>	Red Blood Cells

<b>Neu5Ac</b>	N-acetyl-D-neuraminic acid
<b>Gal</b>	D-Galactose
<b>GlcNAc</b>	N-acetyl-D-glucosamine
<b>MALDI-MS</b>	matrix assisted laser desorption ionization mass spectrometry
<b>LC-MS</b>	liquid chromatography-mass spectrometry

## References

- Alexopoulou AN, Mulhaupt HA, Couchman JR. Syndecans in wound healing, inflammation and vascular biology. *Int J Biochem Cell Biol.* 2007; 39:505–528. [PubMed: 17097330]
- Belser JA, Blixt O, Chen LM, Pappas C, Maines TR, Van Hoeven N, Donis R, Busch J, McBride R, Paulson JC, et al. Contemporary North American influenza H7 viruses possess human receptor specificity: Implications for virus transmissibility. *Proc Natl Acad Sci U S A.* 2008; 105:7558–7563. [PubMed: 18508975]
- Bewley CA. Illuminating the switch in influenza viruses. *Nat Biotechnol.* 2008; 26:60–62. [PubMed: 18183018]
- Bhatia PK, Mukhopadhyay A. Protein glycosylation: implications for in vivo functions and therapeutic applications. *Adv Biochem Eng Biotechnol.* 1999; 64:155–201. [PubMed: 9933978]
- Blixt O, Head S, Mondala T, Scanlan C, Huflejt ME, Alvarez R, Bryan MC, Fazio F, Calarese D, Stevens J, et al. Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc Natl Acad Sci U S A.* 2004; 101:17033–17038. [PubMed: 15563589]
- Blixt O, Razi N. Chemoenzymatic synthesis of glycan libraries. *Methods Enzymol.* 2006; 415:137–153. [PubMed: 17116472]
- Bovin NV, Tuzikov AB, Chinarev AA, Gambaryan AS. Multimeric glycotherapeutics: new paradigm. *Glycoconj J.* 2004; 21:471–478. [PubMed: 15750788]
- Carroll SM, Higa HH, Paulson JC. Different cell-surface receptor determinants of antigenically similar influenza virus hemagglutinins. *J Biol Chem.* 1981; 256:8357–8363. [PubMed: 6167577]
- Case DA, Cheatham TE 3rd, Darden T, Gohlke H, Luo R, Merz KM Jr. Onufriev A, Simmerling C, Wang B, Woods RJ. The Amber biomolecular simulation programs. *J Comput Chem.* 2005; 26:1668–1688. [PubMed: 16200636]
- Chandrasekaran A, Srinivasan A, Raman R, Viswanathan K, Raguram S, Tumpey TM, Sasisekharan V, Sasisekharan R. Glycan topology determines human adaptation of avian H5N1 virus hemagglutinin. *Nat Biotechnol.* 2008; 26:107–113. [PubMed: 18176555]
- Collins BE, Paulson JC. Cell surface biology mediated by low affinity multivalent protein-glycan interactions. *Curr Opin Chem Biol.* 2004; 8:617–625. [PubMed: 15556405]
- Comelli EM, Sutton-Smith M, Yan Q, Amado M, Panico M, Gilmartin T, Whisenant T, Lanigan CM, Head SR, Goldberg D, et al. Activation of murine CD4+ and CD8+ T lymphocytes leads to dramatic remodeling of N-linked glycans. *J Immunol.* 2006; 177:2431–2440. [PubMed: 16888005]
- Connor RJ, Kawaoka Y, Webster RG, Paulson JC. Receptor specificity in human, avian, and equine H2 and H3 influenza virus isolates. *Virology.* 1994; 205:17–23. [PubMed: 7975212]
- Crocker PR. Siglecs in innate immunity. *Curr Opin Pharmacol.* 2005; 5:431–437. [PubMed: 15955740]
- Crocker PR, Paulson JC, Varki A. Siglecs and their roles in the immune system. *Nat Rev Immunol.* 2007; 7:255–266. [PubMed: 17380156]
- Dam TK, Brewer CF. Effects of clustered epitopes in multivalent ligand-receptor interactions. *Biochemistry.* 2008; 47:8470–8476. [PubMed: 18652478]
- Dam TK, Gerken TA, Brewer CF. Thermodynamics of multivalent carbohydrate-lectin cross-linking interactions: importance of entropy in the bind and jump mechanism. *Biochemistry.* 2009; 48:3822–3827. [PubMed: 19292456]

- Duverger E, Frison N, Roche AC, Monsigny M. Carbohydrate-lectin interactions assessed by surface plasmon resonance. *Biochimie*. 2003; 85:167–179. [PubMed: 12765786]
- Eisen MB, Sabesan S, Skehel JJ, Wiley DC. Binding of the influenza A virus to cell-surface receptors: structures of five hemagglutinin-sialyloligosaccharide complexes determined by X-ray crystallography. *Virology*. 1997; 232:19–31. [PubMed: 9185585]
- Feizi T, Fazio F, Chai W, Wong CH. Carbohydrate microarrays - a new set of technologies at the frontiers of glycomics. *Curr Opin Struct Biol*. 2003; 13:637–645. [PubMed: 14568620]
- Fukui S, Feizi T, Galustian C, Lawson AM, Chai W. Oligosaccharide microarrays for high-throughput detection and specificity assignments of carbohydrate-protein interactions. *Nat Biotechnol*. 2002; 20:1011–1017. [PubMed: 12219077]
- Gambaryan A, Tuzikov A, Pazynina G, Bovin N, Balish A, Klimov A. Evolution of the receptor binding phenotype of influenza A (H5) viruses. *Virology*. 2006; 344:432–438. [PubMed: 16226289]
- Gambaryan A, Yamnikova S, Lvov D, Tuzikov A, Chinarev A, Pazynina G, Webster R, Matrosovich M, Bovin N. Receptor specificity of influenza viruses from birds and mammals: new data on involvement of the inner fragments of the carbohydrate chain. *Virology*. 2005; 334:276–283. [PubMed: 15780877]
- Gambaryan AS, Matrosovich MN. A solid-phase enzyme-linked assay for influenza virus receptor-binding activity. *J Virol Methods*. 1992; 39:111–123. [PubMed: 1430058]
- Gambaryan AS, Piskarev VE, Yamskov IA, Sakharov AM, Tuzikov AB, Bovin NV, Nifant'ev NE, Matrosovich MN. Human influenza virus recognition of sialyloligosaccharides. *FEBS Lett*. 1995; 366:57–60. [PubMed: 7789517]
- Gambaryan AS, Tuzikov AB, Pazynina GV, Desheva JA, Bovin NV, Matrosovich MN, Klimov AI. 6-sulfo sialyl Lewis X is the common receptor determinant recognized by H5, H6, H7 and H9 influenza viruses of terrestrial poultry. *Virology*. 2008; 378:5–15. [PubMed: 18652681]
- Gamblin SJ, Haire LF, Russell RJ, Stevens DJ, Xiao B, Ha Y, Vasishth N, Steinhauer DA, Daniels RS, Elliot A, et al. The structure and receptor binding properties of the 1918 influenza hemagglutinin. *Science*. 2004; 303:1838–1842. [PubMed: 14764886]
- Gildersleeve JC, Oyelaran O, Simpson JT, Allred B. Improved procedure for direct coupling of carbohydrates to proteins via reductive amination. *Bioconjug Chem*. 2008; 19:1485–1490. [PubMed: 18597509]
- Glaser L, Zamarin D, Acland HM, Spackman E, Palese P, Garcia-Sastre A, Tewari D. Sequence analysis and receptor specificity of the hemagglutinin of a recent influenza H2N2 virus isolated from chicken in North America. *Glycoconj J*. 2006; 23:93–99. [PubMed: 16575526]
- Grun CH, van Vliet SJ, Schiphorst WE, Bank CM, Meyer S, van Die I, van Kooyk Y. One-step biotinylation procedure for carbohydrates to study carbohydrate-protein interactions. *Anal Biochem*. 2006; 354:54–63. [PubMed: 16713984]
- Ha Y, Stevens DJ, Skehel JJ, Wiley DC. X-ray structure of the hemagglutinin of a potential H3 avian progenitor of the 1968 Hong Kong pandemic influenza virus. *Virology*. 2003; 309:209–218. [PubMed: 12758169]
- Ha Y, Stevens DJ, Skehel JJ, Wiley DC. X-ray structures of H5 avian and H9 swine influenza virus hemagglutinins bound to avian and human receptor analogs. *Proc Natl Acad Sci U S A*. 2001; 98:11181–11186. [PubMed: 11562490]
- Hanson S, Best M, Bryan MC, Wong CH. Chemoenzymatic synthesis of oligosaccharides and glycoproteins. *Trends Biochem Sci*. 2004; 29:656–663. [PubMed: 15544952]
- Hatakeyama S, Sakai-Tagawa Y, Kiso M, Goto H, Kawakami C, Mitamura K, Sugaya N, Suzuki Y, Kawaoka Y. Enhanced expression of an alpha2,6-linked sialic acid on MDCK cells improves isolation of human influenza viruses and evaluation of their sensitivity to a neuraminidase inhibitor. *J Clin Microbiol*. 2005; 43:4139–4146. [PubMed: 16081961]
- Helenius A, Aebi M. Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem*. 2004; 73:1019–1049. [PubMed: 15189166]
- Hoffmann E, Neumann G, Kawaoka Y, Hobom G, Webster RG. A DNA transfection system for generation of influenza A virus from eight plasmids. *Proc Natl Acad Sci U S A*. 2000; 97:6108–6113. [PubMed: 10801978]



- Houseman BT, Mrksich M. Carbohydrate arrays for the evaluation of protein binding and enzymatic modification. *Chem Biol.* 2002; 9:443–454. [PubMed: 11983333]
- Huang W, Ochiai H, Zhang X, Wang LX. Introducing N-glycans into natural products through a chemoenzymatic approach. *Carbohydr Res.* 2008
- Ibricevic A, Pekosz A, Walter MJ, Newby C, Battaile JT, Brown EG, Holtzman MJ, Brody SL. Influenza virus receptor specificity and cell tropism in mouse and human airway epithelial cells. *J Virol.* 2006; 80:7469–7480. [PubMed: 16840327]
- Karamanska R, Clarke J, Blixt O, Macrae JI, Zhang JQ, Crocker PR, Laurent N, Wright A, Flitsch SL, Russell DA, et al. Surface plasmon resonance imaging for real-time, label-free analysis of protein interactions with carbohydrate microarrays. *Glycoconj J.* 2008; 25:69–74. [PubMed: 17574526]
- Kawashima H, Petryniak B, Hiraoka N, Mitoma J, Huckaby V, Nakayama J, Uchimura K, Kadomatsu K, Muramatsu T, Lowe JB, et al. N-acetylglucosamine-6-O-sulfotransferases 1 and 2 cooperatively control lymphocyte homing through L-selectin ligand biosynthesis in high endothelial venules. *Nat Immunol.* 2005; 6:1096–1104. [PubMed: 16227985]
- Kiessling LL, Gestwicki JE, Strong LE. Synthetic multivalent ligands in the exploration of cell-surface interactions. *Curr Opin Chem Biol.* 2000; 4:696–703. [PubMed: 11102876]
- Kirschner KN, Yongye AB, Tschampel SM, Gonzalez-Outeirino J, Daniels CR, Foley BL, Woods RJ. GLYCAM06: a generalizable biomolecular force field. *Carbohydrates. J Comput Chem.* 2008; 29:622–655. [PubMed: 17849372]
- Kumari K, Gulati S, Smith DF, Gulati U, Cummings RD, Air GM. Receptor binding specificity of recent human H3N2 influenza viruses. *Virol J.* 2007; 4:42. [PubMed: 17490484]
- Liang PH, Wu CY, Greenberg WA, Wong CH. Glycan arrays: biological and medical applications. *Curr Opin Chem Biol.* 2008; 12:86–92. [PubMed: 18258211]
- Lowe JB. Glycosyltransferases and glycan structures contributing to the adhesive activities of L-, E- and P-selectin counter-receptors. *Biochem Soc Symp.* 2002:33–45. [PubMed: 12655772]
- Lowen AC, Palese P. Influenza virus transmission: basic science and implications for the use of antiviral drugs during a pandemic. *Infect Disord Drug Targets.* 2007; 7:318–328. [PubMed: 18220963]
- Luytjes W, Krystal M, Enami M, Parvin JD, Palese P. Amplification, expression, and packaging of foreign gene by influenza virus. *Cell.* 1989; 59:1107–1113. [PubMed: 2598262]
- Maines TR, Chen LM, Matsuoka Y, Chen H, Rowe T, Ortin J, Falcon A, Nguyen TH, Mai le Q, Sedyaningsih ER, et al. Lack of transmission of H5N1 avian-human reassortant influenza viruses in a ferret model. *Proc Natl Acad Sci U S A.* 2006; 103:12121–12126. [PubMed: 16880383]
- Maines TR, Jayaraman A, Belser JA, Wadford DA, Pappas C, Zeng H, Gustin KM, Pearce MB, Viswanathan K, Shriver ZH, et al. Transmission and pathogenesis of swine-origin 2009 A(H1N1) influenza viruses in ferrets and mice. *Science.* 2009; 325:484–487. [PubMed: 19574347]
- Mammen M, Choi SK, Whitesides GM. Polyvalent interactions in biological systems: Implications for design and use of multivalent ligands and inhibitors. *Angewandte Chemie-International Edition.* 1998; 37:2755–2794.
- Mammen M, Dahmann G, Whitesides GM. Effective inhibitors of hemagglutination by influenza virus synthesized from polymers having active ester groups. Insight into mechanism of inhibition. *J Med Chem.* 1995; 38:4179–4190. [PubMed: 7473545]
- Manzi AE, Norgard-Sumnicht K, Argade S, Marth JD, van Halbeek H, Varki A. Exploring the glycan repertoire of genetically modified mice by isolation and profiling of the major glycan classes and nano-NMR analysis of glycan mixtures. *Glycobiology.* 2000; 10:669–689. [PubMed: 10910972]
- Matrosovich M, Matrosovich T, Uhlendorff J, Garten W, Klenk HD. Avian-virus-like receptor specificity of the hemagglutinin impedes influenza virus replication in cultures of human airway epithelium. *Virology.* 2007; 361:384–390. [PubMed: 17207830]
- Matrosovich M, Tuzikov A, Bovin N, Gambaryan A, Klimov A, Castrucci MR, Donatelli I, Kawaoka Y. Early alterations of the receptor-binding properties of H1, H2, and H3 avian influenza virus hemagglutinins after their introduction into mammals. *J Virol.* 2000; 74:8502–8512. [PubMed: 10954551]

- Matrosovich MN, Matrosovich TY, Gray T, Roberts NA, Klenk HD. Human and avian influenza viruses target different cell types in cultures of human airway epithelium. *Proc Natl Acad Sci U S A*. 2004; 101:4620–4624. [PubMed: 15070767]
- Mercey E, Sadir R, Maillart E, Roget A, Baleux F, Lortat-Jacob H, Livache T. Polypyrrole oligosaccharide array and surface plasmon resonance imaging for the measurement of glycosaminoglycan binding interactions. *Anal Chem*. 2008; 80:3476–3482. [PubMed: 18348577]
- Norgard-Sumnicht K, Bai X, Esko JD, Varki A, Manzi AE. Exploring the outcome of genetic modifications of glycosylation in cultured cell lines by concurrent isolation of the major classes of vertebrate glycans. *Glycobiology*. 2000; 10:691–700. [PubMed: 10910973]
- Oyelaran O, Gildersleeve JC. Application of carbohydrate array technology to antigen discovery and vaccine development. *Expert Rev Vaccines*. 2007; 6:957–969. [PubMed: 18377358]
- Palese P. Influenza: old and new threats. *Nat Med*. 2004; 10:S82–87. [PubMed: 15577936]
- Pappas C, Aguilar PV, Basler CF, Solorzano A, Zeng H, Perrone LA, Palese P, Garcia-Sastre A, Katz JM, Tumpey TM. Single gene reassortants identify a critical role for PB1, HA, and NA in the high virulence of the 1918 pandemic influenza virus. *Proc Natl Acad Sci U S A*. 2008; 105:3064–3069. [PubMed: 18287069]
- Paulson JC, Rogers GN. Resialylated erythrocytes for assessment of the specificity of sialyloligosaccharide binding proteins. *Methods Enzymol*. 1987; 138:162–168. [PubMed: 3600320]
- Pleschka S, Jaskunas R, Engelhardt OG, Zurcher T, Palese P, Garcia-Sastre A. A plasmid-based reverse genetics system for influenza A virus. *J Virol*. 1996; 70:4188–4192. [PubMed: 8648766]
- Rogers GN, Paulson JC. Receptor determinants of human and animal influenza virus isolates: differences in receptor specificity of the H3 hemagglutinin based on species of origin. *Virology*. 1983; 127:361–373. [PubMed: 6868370]
- Rogers GN, Paulson JC, Daniels RS, Skehel JJ, Wilson IA, Wiley DC. Single amino acid substitutions in influenza haemagglutinin change receptor binding specificity. *Nature*. 1983; 304:76–78. [PubMed: 6191220]
- Rudd PM, Wormald MR, Dwek RA. Sugar-mediated ligand-receptor interactions in the immune system. *Trends Biotechnol*. 2004; 22:524–530. [PubMed: 15450746]
- Russell RJ, Stevens DJ, Haire LF, Gamblin SJ, Skehel JJ. Avian and human receptor binding by hemagglutinins of influenza A viruses. *Glycoconj J*. 2006; 23:85–92. [PubMed: 16575525]
- Sauter NK, Hanson JE, Glick GD, Brown JH, Crowther RL, Park SJ, Skehel JJ, Wiley DC. Binding of influenza virus hemagglutinin to analogs of its cell-surface receptor, sialic acid: analysis by proton nuclear magnetic resonance spectroscopy and X-ray crystallography. *Biochemistry*. 1992; 31:9609–9621. [PubMed: 1327122]
- Seeberger PH, Werz DB. Synthesis and medical applications of oligosaccharides. *Nature*. 2007; 446:1046–1051. [PubMed: 17460666]
- Shinya K, Ebina M, Yamada S, Ono M, Kasai N, Kawaoka Y. Avian flu: influenza virus receptors in the human airway. *Nature*. 2006; 440:435–436. [PubMed: 16554799]
- Skehel JJ, Wiley DC. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annu Rev Biochem*. 2000; 69:531–569. [PubMed: 10966468]
- Smithson G, Rogers CE, Smith PL, Scheidegger EP, Petryniak B, Myers JT, Kim DS, Homeister JW, Lowe JB. Fuc-TVII is required for T helper 1 and T cytotoxic 1 lymphocyte selectin ligand expression and recruitment in inflammation, and together with Fuc-TIV regulates naive T cell trafficking to lymph nodes. *J Exp Med*. 2001; 194:601–614. [PubMed: 11535629]
- Song X, Xia B, Stowell SR, Lasanajak Y, Smith DF, Cummings RD. Novel fluorescent glycan microarray strategy reveals ligands for galectins. *Chem Biol*. 2009; 16:36–47. [PubMed: 19171304]
- Soundararajan V, Tharakaraman K, Raman R, Raguram S, Shriver Z, Sasisekharan V, Sasisekharan R. Extrapolating from sequence--the 2009 H1N1 'swine' influenza virus. *Nat Biotechnol*. 2009; 27:510–513. [PubMed: 19513050]
- Srinivasan A, Viswanathan K, Raman R, Chandrasekaran A, Raguram S, Tumpey TM, Sasisekharan V, Sasisekharan R. Quantitative biochemical rationale for differences in transmissibility of 1918

- pandemic influenza A viruses. *Proc Natl Acad Sci U S A*. 2008; 105:2800–2805. [PubMed: 18287068]
- Stevens J, Blixt O, Chen LM, Donis RO, Paulson JC, Wilson IA. Recent avian H5N1 viruses exhibit increased propensity for acquiring human receptor specificity. *J Mol Biol*. 2008; 381:1382–1394. [PubMed: 18672252]
- Stevens J, Blixt O, Glaser L, Taubenberger JK, Palese P, Paulson JC, Wilson IA. Glycan microarray analysis of the hemagglutinins from modern and pandemic influenza viruses reveals different receptor specificities. *J Mol Biol*. 2006; 355:1143–1155. [PubMed: 16343533]
- Stevens J, Blixt O, Paulson JC, Wilson IA. Glycan microarray technologies: tools to survey host specificity of influenza viruses. *Nat Rev Microbiol*. 2006; 4:857–864. [PubMed: 17013397]
- Stevens J, Blixt O, Tumpey TM, Taubenberger JK, Paulson JC, Wilson IA. Structure and receptor specificity of the hemagglutinin from an H5N1 influenza virus. *Science*. 2006; 312:404–410. [PubMed: 16543414]
- Stray SJ, Cummings RD, Air GM. Influenza virus infection of desialylated cells. *Glycobiology*. 2000; 10:649–658. [PubMed: 10910970]
- Taylor ME, Drickamer K. Paradigms for glycan-binding receptors in cell adhesion. *Curr Opin Cell Biol*. 2007; 19:572–577. [PubMed: 17942297]
- Totani K, Kubota T, Kuroda T, Murata T, Hidari KI, Suzuki T, Suzuki Y, Kobayashi K, Ashida H, Yamamoto K, et al. Chemoenzymatic synthesis and application of glycopolymers containing multivalent sialyloligosaccharides with a poly(L-glutamic acid) backbone for inhibition of infection by influenza viruses. *Glycobiology*. 2003; 13:315–326. [PubMed: 12626382]
- Tumpey TM, Basler CF, Aguilar PV, Zeng H, Solorzano A, Swayne DE, Cox NJ, Katz JM, Taubenberger JK, Palese P, et al. Characterization of the reconstructed 1918 Spanish influenza pandemic virus. *Science*. 2005; 310:77–80. [PubMed: 16210530]
- Tumpey TM, Maines TR, Van Hoeven N, Glaser L, Solorzano A, Pappas C, Cox NJ, Swayne DE, Palese P, Katz JM, et al. A two-amino acid change in the hemagglutinin of the 1918 influenza virus abolishes transmission. *Science*. 2007; 315:655–659. [PubMed: 17272724]
- Uematsu R, Furukawa J, Nakagawa H, Shinohara Y, Deguchi K, Monde K, Nishimura S. High throughput quantitative glycomics and glycoform-focused proteomics of murine dermis and epidermis. *Mol Cell Proteomics*. 2005; 4:1977–1989. [PubMed: 16170054]
- van der Laan JW, Herberts C, Lambkin-Williams R, Boyers A, Mann AJ, Oxford J. Animal models in influenza vaccine testing. *Expert Rev Vaccines*. 2008; 7:783–793. [PubMed: 18665776]
- van Die I, Cummings RD. Glycans modulate immune responses in helminth infections and allergy. *Chem Immunol Allergy*. 2006; 90:91–112. [PubMed: 16210905]
- Van Hoeven N, Pappas C, Belser JA, Maines TR, Zeng H, Garcia-Sastre A, Sasisekharan R, Katz JM, Tumpey TM. Human HA and polymerase subunit PB2 proteins confer transmission of an avian influenza virus through the air. *Proc Natl Acad Sci U S A*. 2009
- van Riel D, Munster VJ, de Wit E, Rimmelzwaan GF, Fouchier RA, Osterhaus AD, Kuiken T. Human and Avian Influenza Viruses Target Different Cells in the Lower Respiratory Tract of Humans and Other Mammals. *Am J Pathol*. 2007
- Varki, A.; Cummings, RD.; Esko, JD.; Freeze, HH.; Hart, GW.; Etzler, ME. *Essentials of Glycobiology*. Cold Spring Harbor Laboratory Press; Cold Spring Harbor, New York: 2008.
- Wan H, Sorrell EM, Song H, Hossain MJ, Ramirez-Nieto G, Monne I, Stevens J, Cattoli G, Capua I, Chen LM, et al. Replication and transmission of H9N2 influenza viruses in ferrets: evaluation of pandemic potential. *PLoS ONE*. 2008; 3:e2923. [PubMed: 18698430]
- Wang L, Fuster M, Sriramarao P, Esko JD. Endothelial heparan sulfate deficiency impairs L-selectin- and chemokine-mediated neutrophil trafficking during inflammatory responses. *Nat Immunol*. 2005; 6:902–910. [PubMed: 16056228]
- Weis W, Brown JH, Cusack S, Paulson JC, Skehel JJ, Wiley DC. Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid. *Nature*. 1988; 333:426–431. [PubMed: 3374584]
- Woods RJ. Computational carbohydrate chemistry: what theoretical methods can tell us. *Glycoconj J*. 1998; 15:209–216. [PubMed: 9579797]

- Xia B, Kowar ZS, Ju T, Alvarez RA, Sachdev GP, Cummings RD. Versatile fluorescent derivatization of glycans for glycomic analysis. *Nat Methods*. 2005; 2:845–850. [PubMed: 16278655]
- Xu D, Newhouse EI, Amaro RE, Pao HC, Cheng LS, Markwick PR, McCammon JA, Li WW, Arzberger PW. Distinct glycan topology for avian and human sialopentasaccharide receptor analogues upon binding different hemagglutinins: a molecular dynamics perspective. *J Mol Biol*. 2009; 387:465–491. [PubMed: 19356594]
- Yamada S, Suzuki Y, Suzuki T, Le MQ, Nidom CA, Sakai-Tagawa Y, Muramoto Y, Ito M, Kiso M, Horimoto T, et al. Haemagglutinin mutations responsible for the binding of H5N1 influenza A viruses to human-type receptors. *Nature*. 2006; 444:378–382. [PubMed: 17108965]
- Yang ZY, Wei CJ, Kong WP, Wu L, Xu L, Smith DF, Nabel GJ. Immunization by avian h5 influenza hemagglutinin mutants with altered receptor binding specificity. *Science*. 2007; 317:825–828. [PubMed: 17690300]

### Glycan-Binding Properties of the 2009 A/H1N1 'Swine Flu' Hemagglutinin [Text Box]

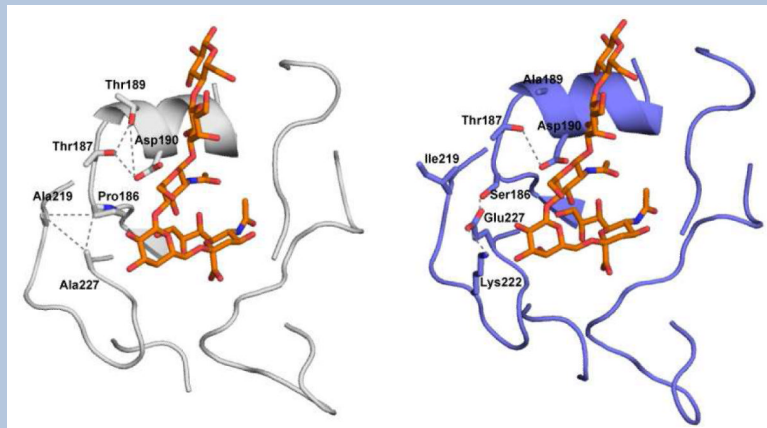
The integration of the *top-down* and *bottoms-up* approach was recently employed to shed key insights into the glycan-binding properties of HA from the 2009 'swine flu' H1N1 influenza A viruses (henceforth referred to as 2009 A/H1N1) and their correlation with the transmissibility of these viruses in ferrets (Maines, et al., 2009). One of the characteristics of a human-adapted virus is in its ability to achieve efficient transmission between humans via respiratory droplets (airborne) (Tumpey, et al., 2007). Comparison of the respiratory droplet transmissibility of representative 2009 A/H1N1 viruses and a recent seasonal H1N1 influenza outbreak strain (A/Brisbane/59/07 or Bris07) in ferrets showed that the 2009 A/H1N1 viruses transmitted less efficiently (Maines, et al., 2009).

It was demonstrated previously (Srinivasan, et al., 2008; Tumpey, et al., 2007) that the efficiency of respiratory droplet transmission in ferrets correlates with the  $\alpha$ 2-6 binding affinity of the viral HA. In fact, a single amino acid mutation in HA of the efficiently transmitting SC18 virus led to a virus (NY18) that transmitted inefficiently. The  $\alpha$ 2-6 binding affinity of NY18 HA was substantially lower than that of SC18 HA (Srinivasan, et al., 2008). In a *bottom-up* approach using a quantitative dose-dependent direct glycan receptor-binding and human lung tissue binding assays (Maines, et al., 2009) the binding pattern of a representative 2009 A/H1N1 HA (A/California/04/09; or CA/04) was found to be similar to that of the pandemic SC18 H1N1 HA. Like SC18 HA, CA/04 showed substantial binding (in a dose-dependent fashion) to the single 6' SLN-LN  $\alpha$ 2-6 oligosaccharide motif. However, the binding affinity of CA/04 HA to 6' SLN-LN was substantially lower than that of SC18 HA. CA/04 HA also showed binding to the apical surface of human tracheal tissue section (in a sialic acid specific manner) and this binding pattern correlates with the predominant distribution of  $\alpha$ 2-6 sialylated glycans on the apical surface of the tracheal tissue (Chandrasekaran, et al., 2008) and the  $\alpha$ 2-6 binding of CA/04 HA in the direct glycan receptor-binding assay.

Through a *top-down* approach involving a combination of sequence analysis and homology based structural modeling of CA/04 HA, it was predicted that this HA would have a substantially higher  $\alpha$ 2-6 as compared to  $\alpha$ 2-3 binding (Soundararajan, et al., 2009). This is consistent with the observed binding pattern in the *bottom-up* approach. Building on this study, a more recent investigation offered a possible structural rationale for the observed reduced  $\alpha$ 2-6 binding affinity of CA/04 HA (in comparison to that of SC18 HA) (Maines, et al., 2009). The receptor-binding site (RBS) of CA/04 HA shares many similar or analogous residues with that of SC18 HA providing an explanation for the similarity in their binding pattern (Maines, et al., 2009; Soundararajan, et al., 2009). These analogous residues include Asp190 and Asp225, which are 'signature' amino acids of human adapted H1N1 HAs that make optimal contacts with the  $\alpha$ 2-6 glycans (Gamblin, et al., 2004). The main differences in RBS between SC18 and CA/04 HA are at positions 145, 186, 189, 219 and 227. The structural analysis of the inter-residue and residue-glycan contacts involving these residue positions pointed to some key differences between CA/04 and SC18 HA (Figure B1). Specifically, in SC18, a network of interactions involving hydrophobic contacts between Ala219, Ala227, Pro186 and ionic contacts between Thr187, Thr189 and Asp190 were involved in positioning Asp190 for optimal contacts with  $\alpha$ 2-6 glycans. On the other hand, the unique combination of residues Ile219 and Glu227 in CA/04 HA had a disrupting effect on the network of interactions given that the contacts between these residues are neither hydrophobic nor ionic. This disrupting effect on the network was implicated to have a negative effect on the positioning of Asp190 and hence offering a possible structural explanation for the observed lower  $\alpha$ 2-6 binding affinity of CA/04 in comparison with that of SC18 HA.

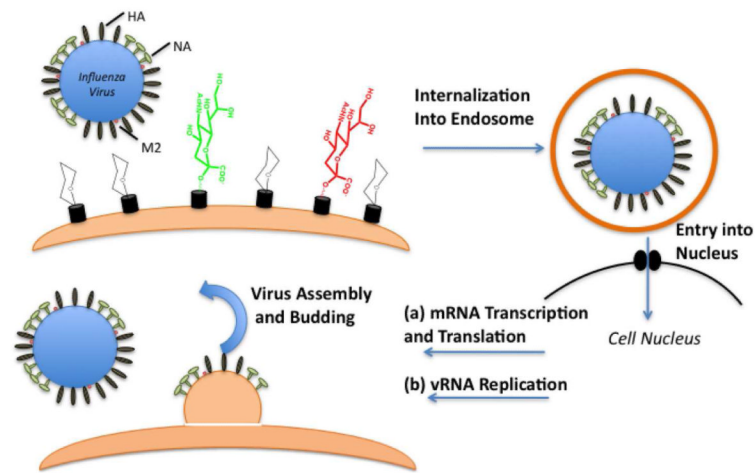


Taken together, the substantially lower  $\alpha$ 2-6 binding affinity of CA/04 HA as compared to that of SC18 HA correlates with its less efficient respiratory droplet transmission in ferrets.



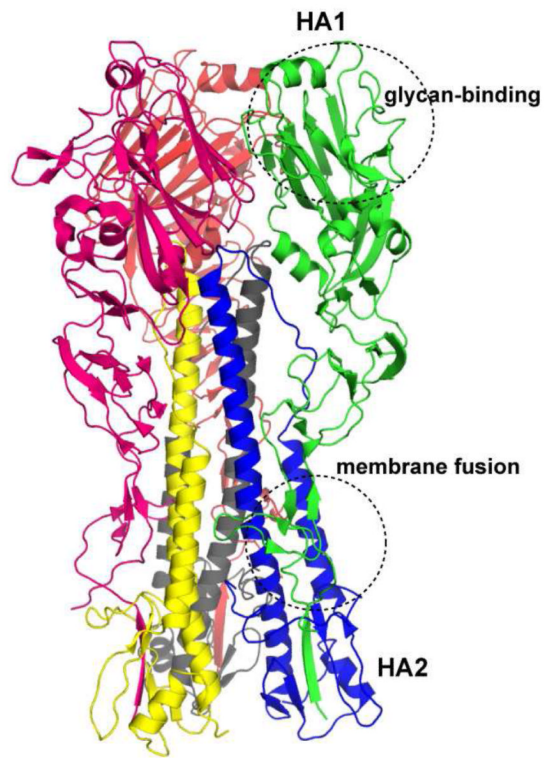
**Figure B1. Structural rationale for  $\alpha$ 2-6 binding affinity of 2009 A/H1N1 HA**

Shown on the *left* are residues in positions 186, 187, 189, 190, 219 and 227 of SC18 HA (cartoon representation in *gray*) in complex with an  $\alpha$ 2-6 oligosaccharide (Neu5Ac $\alpha$ 2-6Gal $\beta$ 1-4GlcNAc $\beta$ 1-3Gal $\beta$ 1-4Glc) (Srinivasan, et al., 2008). The oligosaccharide is shown in stick representation colored orange (carbon atoms). The network of interactions is shown in dotted gray lines. Shown on the *right* are the same residue positions in structural complex of CA/04 HA (cartoon representation in *violet*) with the same  $\alpha$ 2-6 oligosaccharide (Maines, et al., 2009). The presence of unique combination of Ile219 and Glu227 is not favorable for the optimal positioning of Asp 190 (as seen on the *left*) for contacts with the  $\alpha$ 2-6 oligosaccharide. Lys222 is also shown as it is positioned to interact with Glu227.



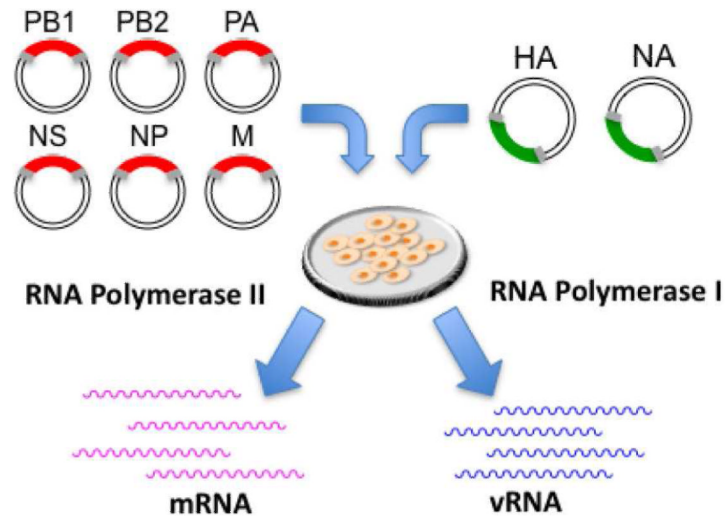
**Figure 1. The influenza infection cycle**

HA on the surface of the virus binds to sialic acid of the correct linkage (shown in green), initiating fusion of the virus with the host cell. This interaction is highly specific for certain types of sialic acid; HA does not bind with equal affinity (avidity) to sialic acid with a different linkage (shown in red) or to several other glycan structures. The other key viral surface proteins are the neuraminidase (NA) and the ion channel protein M2. Once internalized in the cell, fusion occurs, initiated by HA, and complexes of RNA and proteins, termed viral ribonucleoprotein complexes (vRNPs) are transported into the nucleus of the host cell. There, transcription, to mRNA, and, after export, protein synthesis occurs. Also within the nucleus there is transcription of the RNA genome. Assembly of progeny vRNPs then occurs, with export, assembly of the virus progeny and finally budding of the newly formed virus particles.

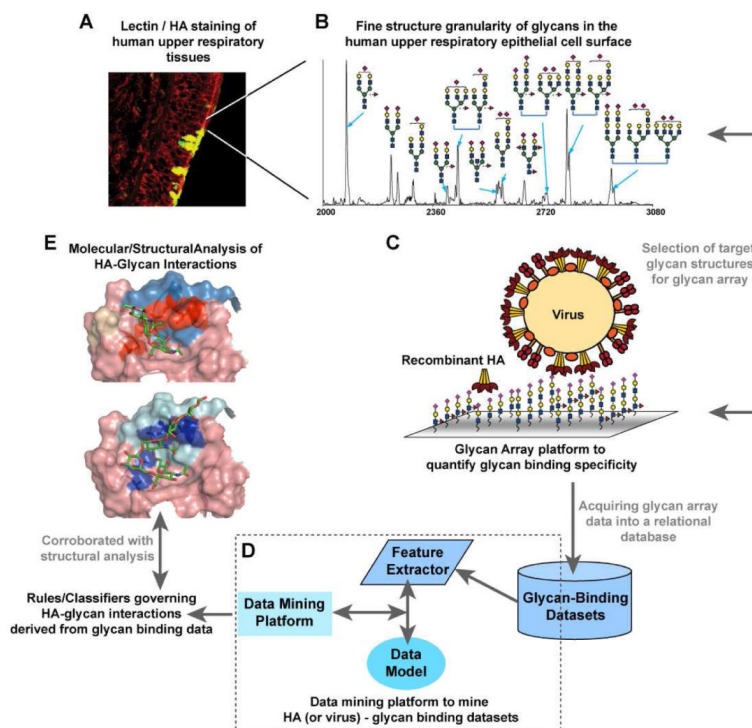


**Figure 2. Structure of trimeric HA**

The cartoon rendering of a representative H1N1 HA (PDB ID: 1RU7) is shown where each chain (HA1 or HA2) in each of the monomer is colored distinctly. The region in HA1 involved in sialylated glycan receptor binding and the region between HA1 and HA2 involved in membrane fusion are also shown.



**Figure 3. The eight-plasmid polymerase(pol) I– II system for the generation of influenza A virus** Expression plasmids containing the eight viral cDNAs, inserted between the human pol I promoter and the pol II promoter are transfected into cells. Because each plasmid contains two different promoters, both cellular pol I and pol II will transcribe the plasmid template, which will result in the production of both viral mRNAs and vRNAs. Synthesis of the viral polymerase complex proteins (PB1, PB2, PA, and nucleoproteins) initiates the viral replication cycle. Ultimately, the assembly of viral molecules results in the interaction of all synthesized molecules to generate infectious influenza A virus.



**Figure 4. Bridging glycan-binding specificity with biological function of influenza A virus**

**A**, Lectin staining of human upper respiratory tissues provides a high-level picture of  $\alpha$ 2-3- and  $\alpha$ 2-6- N- and O-linked glycans. **B**, Finer granularity on the physiological glycans is obtained by analyzing glycans derived from upper respiratory epithelial cells using a combination of analytical tools. **C**, The biochemical specificity of HA-glycan interactions is characterized based on binding of recombinant HA or whole viruses to glycan structures presented using a glycan array platform. The array platform is designed to incorporate target structures based on their predominant expression in the upper respiratory tissues. **D**, The high throughput data on binding of HA or virus to hundreds of glycans on the array is captured into a relational database and this data is mined using data mining methods to obtain rules or classifiers that govern the glycan-binding specificity of WT and mutant HAs. **E**, The rules obtained from data mining are corroborated using X-ray co-crystal structures of HA-glycan complexes and molecular simulation of HA-glycan interactions. The comprehensive knowledge of the key determinants of HA-glycan interactions obtained using this integrated framework provides a much better handle to correlate with the biological function of host adaptation of the influenza A virus.



Many glycan-protein interactions are multivalent and thus are the product of multiple interactions between a glycan structure and a protein. Qualitative description of the strength of this interaction is termed *avidity* and an apparent equilibrium constant is used to quantitatively describe the interaction. Whitesides and others have explicitly dealt with the relationship between *affinity* ( $K^{mono}$ ), a description of a 1:1 complex and *avidity* ( $K_N^{poly}$ ).

$$K_N^{poly} = (K^{mono})^{\alpha N}$$

In this case,  $\alpha$ , is a measure of the cooperativity of the system. Moving from a description of equilibrium constants to a measure of  $\Delta G$ ,  $\Delta G_{avg}^{poly}$  is defined as the average change in free energy of the interaction:

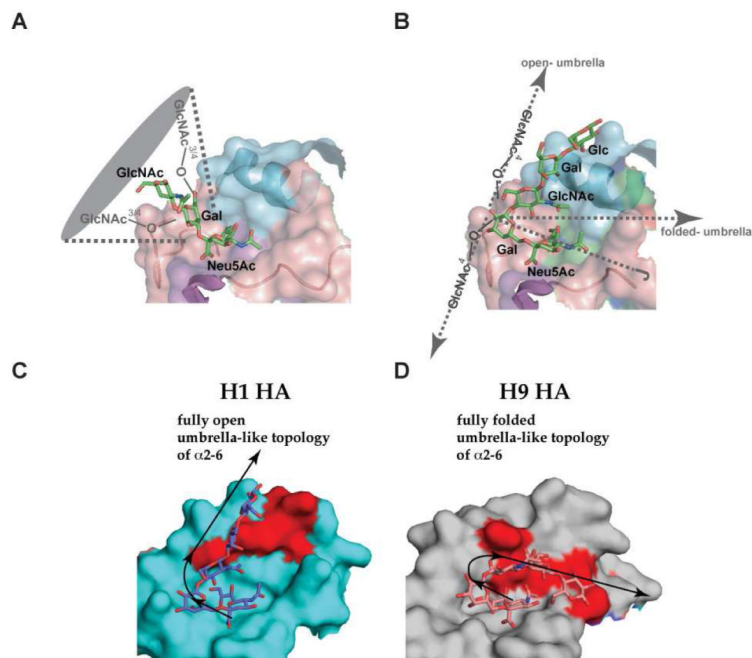
$$\Delta G_{avg}^{poly} = \Delta G_N^{poly} / N$$

$$\Delta G_{avg}^{poly} = \alpha \Delta G^{mono}$$

Where  $\Delta G_{avg}^{poly} > \Delta G^{mono}$ , the degree of cooperativity,  $\alpha$ , in such systems is greater than one. Accordingly, in the case where  $\Delta G_{avg}^{poly} = \Delta G^{mono}$ ,  $\alpha$  is 1, and in the case where  $\Delta G_{avg}^{poly} < \Delta G^{mono}$ , the system experiences negative cooperativity.

#### Figure 5. Defining the thermodynamics of glycan-protein interactions

In many cases, glycan-protein interactions are multivalent, involving multiple 1:1 interactions where the surfaces can either be two cells, or in the case of influenza A, the virus particle and a host cell. From our understanding of the thermodynamic relationships above, to define a glycan-protein interaction, it is critical to define  $N$ , as only then can the precise relationship be determined between  $\Delta G_{poly}$ ,  $K_{poly}$ , and  $\alpha$ . Development of biochemical and array systems where  $N$  can be defined thus have been and will continue to be a priority. Alternatively, if  $N$  is unknown, information can still be derived, but precise relationships, including cooperativity, cannot be determined.



**Figure 6. Three-dimensional glycan topology influences molecular HA-glycan interactions**

**A**, Interactions of HA with *cone-like* topology that is characteristic of avian HA binding to  $\alpha$ 2-3 and short  $\alpha$ 2-6 (such as multiantennary N-linked glycans with single lactosamine branches terminated with  $\alpha$ 2-6). **B**, Interactions of HA with *umbrella-like* topology that is characteristic of polylactosamine branch terminated with  $\alpha$ 2-6. **C**, Contacts between H1 HA with a fully open umbrella-like topology glycan. **D**, Contacts between H9 HA and fully folded umbrella like topology glycans. Due to the larger surface on HA spanned by umbrella-like topology, the amino acids involved in making contacts with umbrella-like topology glycans (shown in *red* in **C** and **D**) are very different for different HAs.