Assigned Reading:

**Chapter 34**, Microbial Lectins: Hemagglutinins, Adhesins, and Toxins - *Essentials of Glycobiology, 2nd edition*

**Manuscript**, Structural Evolution of Glycan Recognition by a Family of Potent HIV Antibodies

**Review**, Role of Receptor Binding Specificity in Influenza A Virus Transmission and Pathogenesis
Chapter 34
Microbial Lectins: Hemagglutinins, Adhesins, and Toxins

Essentials of Glycobiology, 2nd edition
Chapter 34  Microbial Lectins: Hemagglutinins, Adhesins, and Toxins

Esko JD, Sharon N.

Many microorganisms exploit host cell-surface glycans as receptors for cell attachment and tissue colonization, and a large number of pathogenic species depend on these interactions for infection. This chapter describes examples of proteins on the surface of microorganisms (adhesins or hemagglutinins), secreted proteins (heat-labile toxins), their glycan partners on mammalian cell surfaces (receptors), and insights into the molecular interactions that take place.

BACKGROUND

Viruses, bacteria, and protozoa express an enormous number of glycan-binding proteins or lectins. Many of these microbial lectins were originally detected based on their ability to aggregate red blood cells (i.e., to induce hemagglutination). The first microbial hemagglutinin identified was in the influenza virus, and it was shown by Alfred Gottschalk in the early 1950s to bind to erythrocytes and other cells though sialic acid residues of cell-surface glycoconjugates. Don Wiley and his associates crystallized the viral hemagglutinin and determined its structure in 1981, and later they solved the structure of cocrystals prepared with sialyllectose. Since then, a number of viral hemagglutinins have been identified and crystallized. These studies set the stage for analyzing other types of microbial lectins produced by bacteria and protozoa.

Nathan Sharon and colleagues first described bacterial surface lectins in the 1970s. These lectins also have hemagglutinating activity, but their primary function is to facilitate attachment or adherence of bacteria to host cells, a prerequisite for bacterial colonization and infection (see Chapter 39). Thus, bacterial lectins are often called adhesins, and the glycan ligands on the surface of the host cells are called receptors. Note that the term “receptor” in this case is equivalent to “ligand” for animal cell lectins. Bacteria also produce toxins, whose actions often depend on glycan-binding subunits that allow the toxin to combine with membrane glyco-conjugates and deliver the functionally active toxic subunit across the plasma membrane. During the last 30 years, many adhesins and toxins have been described, cloned, and characterized. Additionally, adhesins have been discovered on various parasites. The interactions of adhesins with glycan receptors can help determine the tropism of the organism.

Colonization of tissues by microorganisms is not always pathogenic. For example, the normal flora of the lower gastrointestinal tract is determined by appropriate and desirable colonization by beneficial bacteria. The initial events in the formation of nitrogen-fixing nodules in leguminous root tips by species of *Rhizobium* may also involve lectins on the root tip binding to Nod factors generated by the bacterium (see Chapter 37).

Many adhesins contain carbohydrate-recognition domains (CRDs) that bind to the same carbohydrates as endogenous mammalian lectins (see Chapter 26). Like animal cell lectins, some microbial adhesins bind to terminal sugar residues, whereas others bind to internal sequences found in linear or branched oligosaccharide chains. Detailed studies of the specificity of microbial lectins have led to the identification and synthesis of powerful inhibitors of adhesion that may form the basis for therapeutic agents for treating infection (see Chapter 51).

VIRAL GLYCAN-BINDING PROTEINS

By far, the best-studied example of a viral glycan-binding protein is the influenza virus hemagglutinin, which binds to sialic acid–containing glycans. The affinity of this interaction is relatively low, like that of other glycan-binding proteins with their glycan ligands, but the avidity for cell membranes increases because of oligomerization of the hemagglutinin into trimers and the high density of glycan receptors present on the host cell (see Chapter 27). Binding is a prerequisite for fusion of the viral envelope with the plasma membrane and for uptake of the virus into cells. The specificity of the interaction of the hemagglutinin with host glycans varies considerably for different subtypes of the virus. Human strains of influenza-A and -B viruses bind primarily to cells containing Neu5Acα2–6Gal–, whereas chicken influenza viruses bind to Neu5Acα2–3Gal– and porcine strains bind to both Neu5Acα2–3Gal- and Neu5Acα2–6Gal-containing receptors. This linkage preference is due to certain amino acid changes in the hemagglutinin. Influenza-C virus, in contrast, binds exclusively to glycoproteins and glycolipids containing 9-O-acetylated N-acetylneuraminic acid (see Chapter 39).

The specificity of the hemagglutinin correlates with the structures of sialylated glycans expressed on target epithelial cells in animal hosts. For example, trachea epithelial cells in humans express glycans with a preponderance of Neu5Acα2–6Gal linkages, whereas other tissues
contain many more Neu5Ac2–3Gal-terminated glycans. Thus, the specificity of the hemagglutinin determines the tropism of the virus with respect to species and target cells. The hemagglutinin is also the major antigen against which neutralizing antibodies are produced, and antigenic changes in this protein are in part responsible for new viral outbreaks. In addition, it plays a critical role inside the cell where it facilitates pH-dependent fusion of the viral envelope with endosomal membranes after internalization.

The crystal structure of the hemagglutinin shows two disulfide-bonded subunits, HA1 and HA2, which are derived by cleavage of a precursor protein. The monomer consists of a hydrophilic carboxy-terminal domain located inside the viral envelope and a hydrophobic membrane-spanning domain. An elongated triple-helical coiled stem region extends 135 Å from the membrane topped by a globular domain that contains the carbohydrate-recognition domain (Figure 34.1). The crystal structure also revealed a hydrophobic pocket near the CRD, which explains why sialosides containing a hydrophobic aglycone bind the hemagglutinin with greater affinity than simple glycans.

![Figure 34.1](http://www.ncbi.nlm.nih.gov/books/NBK1907/)

**FIGURE 34.1**
Structure of the influenza virus hemagglutinin (HA) ectodomain. (a) A schematic diagram of the trimeric ectodomain of the H3 avian HA from A/duck/Ukr/63 showing residues HA1 9–326 and HA2 1–172. (Gray, red, blue) Modeled carbohydrate side (more...)

In addition to the hemagglutinin, influenza-A and -B virions express a sialidase (traditionally and incorrectly called neuraminidase) that cleaves sialic acids from glycoconjugates. Its functions may include (1) prevention of viral aggregation by removal of sialic acid residues from virion envelope glycoproteins, (2) dissociation of newly synthesized virions inside the cell or as they bud from the cell surface, and (3) desialylation of soluble mucins at sites of infection in order to improve access to membrane-bound sialic acids. In influenza-C virus, a single glycoprotein contains both the hemagglutinin activity and the receptor-destroying activity, which in this case is an esterase that cleaves the 9-O-acetyl group from acetylated sialic acid receptors. Powerful inhibitors have been designed based on the crystal structure of the sialidase from influenza-A virus. Some of these inhibit the enzyme activity at nanomolar concentrations and are in clinical use as antiviral agents (see Chapter 51).

Rotaviruses, the major killer of children worldwide, also can bind to sialic acid residues. These viruses only bind to the intestinal epithelium of newborn infants during a period that appears to correlate with the expression of specific types and arrangements of sialic acids on glycoproteins. Many other viruses (e.g., adenovirus, reovirus, Sendai virus, and polyomavirus) also appear to use sialic acids for infection, and crystal structures are now available for several of their sialic acid–binding domains (Table 34.1).

![Table 34.1](http://www.ncbi.nlm.nih.gov/books/NBK1907/)

**TABLE 34.1**
Examples of viral lectins and hemagglutinins

A number of viruses use heparan sulfate proteoglycans as adhesion receptors (e.g., herpes simplex virus [HSV], foot-and-mouth disease virus, and dengue flavivirus [Table 34.1]). In many cases, the proteoglycans may be part of a coreceptor system in which the microorganisms make initial contact with a cell-surface proteoglycan and later with another receptor. For example, herpesvirus infection is thought to involve viral glycoproteins gC and/or gB binding to cell-surface heparan sulfate proteoglycans, followed by binding of viral glycoprotein gD to one of several cell-surface receptors, including heparan sulfate and one or more members of the Hve (herpesvirus entry) family of receptors. This step leads to fusion of the viral envelope with the host-cell plasma membrane. The coreceptor role of proteoglycan is reminiscent of the formation of ternary complexes required for antithrombin inhibition of thrombin and for fibroblast growth factor (FGF) cell signaling (see Chapter 35). The HSV glycoprotein gB binds heparan sulfate and promotes virus-cell fusion and syncytium formation (cell–cell fusion) as well as adherence. The mechanism by which heparan sulfate facilitates membrane fusion is unknown, but perhaps it acts like a template facilitating the association of fusogenic membrane proteins of the virus or host cell. Although it is clear that cell-surface proteoglycans act as adhesion receptors, their role in invasion and pathogenesis has not been established.

Dengue flavivirus, the causative agent of dengue hemorrhagic fever, binds to heparan sulfate. Modeling the primary sequence of the viral envelope protein on the crystal structure of a related virion envelope protein suggests that the heparan sulfate–binding site may lie along a groove in the protein lined by positively charged amino acids (Figure 34.2). Thus, the glycosaminoglycan (GAG)-binding adhesins may
have a more open structure, consistent with the combining sites of other heparin-binding proteins (see Chapter 35). Foot-and-mouth disease viruses normally bind to epithelial cells through integrins, but after passage in cell culture, they can acquire the capacity to bind heparan sulfate. The combining site for heparin is located in a shallow depression formed by three of the major capsid proteins (Figure 34.3). HIV also can bind to heparan sulfate and other sulfated polysaccharides by way of the V3 loop of the gp120 glycoprotein. Thus, the heparan-sulfate-binding adhesins appear to pick out carbohydrate units within the polysaccharide chains as opposed to binding to terminal sugars. Interestingly, the interaction of the gD glycoprotein of Herpes simplex virus with heparan sulfate shows specificity for a particular substructure in heparan sulfate containing a 3-O-sulfated glucosamine residue, the formation of which is catalyzed by specific isozymes of the glucosaminyl 3-O-sulfotransferase gene family (see Chapter 16). However, it is unclear whether other viruses show a preference for specific heparan sulfate oligosaccharides. A growing body of literature suggests that heparan sulfate and heparin can block the interaction of the viruses with host cells, suggesting a simple therapeutic approach for treating viral infections.

**FIGURE 34.2**
Two views of a putative heparin-binding site in dengue virus envelope protein. The envelope protein monomer is shown in ribbon form, displayed along its longitudinal axis and as an external side view. (Top) Note the alignment of positively charged amino (more...)

**FIGURE 34.3**
Model of foot-and-mouth disease virus (FMDV) in complex with heparin trisaccharides. (a) A schematic depiction of the icosahedral capsid of FMDV. VP1–3 contribute to the external features of the capsid: (blue) VP1; (green) VP2; (red) VP3. The (more...)

**BACTERIAL ADHESION TO GLYCANS**

Bacterial lectins occur commonly in the form of elongated, submicroscopic, multisubunit protein appendages, known as fimbriae (hairs) or pili (threads), which interact with glycoprotein and glycolipid receptors on host cells. The best characterized of these are the mannose-specific type-1 fimbriae, the galabiose-specific P fimbriae, and the N-acetylglucosamine-binding F-17 fimbriae produced by different strains of *Escherichia coli*. Fimbriated bacteria express 100–400 of these appendages, which typically have a diameter of 5–7 nm and can extend hundreds of nanometers in length (Figure 34.4). Thus, pili extend well beyond the bacterial glycocalyx formed from lipopolysaccharide and capsular polysaccharides (see Chapter 20), which can actually interfere with their activity. The carbohydrate-recognition domain of the fimbriae is found in the minor subunit (FimH of type-1 fimbriae, PapG of P fimbriae, and F17G of F17 fimbriae) that is usually located at the tip of the fimbriae. Some of the other bacterial lectins are monomeric or oligomeric membrane proteins. Most bacteria (and possibly other microorganisms) have multiple adhesins with different carbohydrate specificities, which help define the range of susceptible tissues (i.e., the microbe’s ecological niche). Binding is generally of low affinity, but because the adhesins and the receptors often cluster in the plane of the membrane, the resulting avidity can be great. Perhaps an appropriate analogy for adhesin-receptor binding is the interaction of the two faces of Velcro™ strips.

**FIGURE 34.4**
*E. coli* express multiple pili as indicated by the fine filaments surrounding the cells. (Reprinted, with permission of Elsevier, from Sharon N. 2006. *Biochim. Biophys. Acta* 1760: 527–537; courtesy of David L. Hasty, University of Tennessee, Memphis, (more...)

Examination of the high-resolution, three-dimensional structure of the glycan-binding subunit FimH of type-1 fimbriae in complex with bound mannose revealed that although mannose exists in solution as a mixture of α and β anomers, only the former was found in the complex. It is buried at a deep and negatively charged site at the edge of FimH (Figure 34.5). All of the mannose hydroxyl groups, except the anomic one, interact extensively with combining-site residues, almost all of which are situated at the ends of β-strands or in the loops extending from them. Part of the hydrogen-bonding network is identical to that found in mannose complexes of other lectins, such as those of plants.
In the urinary tract, the type-1 fimbriae of *E. coli* mediate binding of the bacteria to a protein called uroplakin Ia. This glycoprotein presents high levels of terminally exposed mannose residues that are capable of specifically interacting with FinH. Anchorage of *E. coli* to the urothelial surface via type-1 fimbriae–uroplakin Ia interactions may play a role in their colonization of the bladder and eventual ascent through the ureters against urine flow to invade the kidneys. Another receptor for these fimbriae is the urinary Tamm–Horsfall glycoprotein, which acts as a soluble inhibitor of the adhesion of the bacteria to the uroplakens and helps to clear the bacteria from the urinary tract. Indeed, mice lacking the Tamm–Horsfall gene are considerably more susceptible to bladder colonization by type-1-fimbriated *E. coli* than normal mice, whereas they are equally susceptible to P-fimbriated *E. coli* that do not bind to the same glycoprotein. Type-1 fimbriae also are instrumental in the attachment of *E. coli* to human polymorphonuclear cells and to human and mouse macrophages. This is often followed by the ingestion and killing of the bacteria, a phenomenon named lectinophagocytosis, which is an early example of innate immunity (see Chapter 39).

Other binding specificities have been described as well (Table 34.2). The specificity of binding can explain the tissue tropism of the organism. The columnar epithelium that lines the large intestine expresses Galβ1–4Gal-Cer, whereas the cells lining the small intestine do not. Thus, *Bacteroides*, *Clostridium*, *E. coli*, and *Lactobacillus* only colonize the large intestine under normal conditions. P-fimbriated *E. coli* as well as different toxins bind specifically to galabiose (Galα1–4Gal) and galabiose-containing oligosaccharides, most commonly as constituents of glycolipids. Binding occurs either to terminal nonreducing galabiose units or to internal ones (i.e., when the disaccharide is capped by other sugars). P-fimbriated *E. coli* adhere mainly to the upper part of the kidney, where galabiose is abundant. *E. coli* K99 provides a striking illustration of the fine specificity of bacterial surface lectins and their relationship to the animal tropism of the bacteria. This organism binds to glycolipids that contain N-glycolylneuraminic acid (Neu5Gc), in the form of Neu5Gco2–3Galβ1–4Glc, but not to those that contain N-acetyleneuraminic acid. These two sugars differ in only a single hydroxyl group, present in the acyl substituent on the amino group at C-5 of N-glycolylneuraminic acid and absent in that of N-acetyleneuraminic acid. N-glycolylneuraminic acid is found on intestinal cells of newborn piglets, but it disappears when the animals develop and grow, and it is not formed normally by humans. This can explain why *E. coli* K99 can cause often lethal diarrhea in piglets but not in adult pigs or humans.

The relationship between microbes and the host can be quite complex. For example, colonization of germ-free mice with *Bacteroides thetaiotaomicron*, a normal resident microbe of the small intestine, can induce an α1–2 fucosyltransferase in the mucosal epithelial cells. The bacteria bind to L-fucose residues and also use it as a carbon source.

**TOXINS THAT BIND GLYCANS**

A number of secreted bacterial toxins also bind to glycans (Table 34.3). The best-studied example is the toxin from *Vibrio cholera* (cholera toxin), which consists of A and B sub-units in the ratio AB₅. Its crystal structure shows that the B subunits bind to the Galβ1–3GalNAc moiety of GM1 ganglioside receptors through CRDs located on the base of the subunits (Figure 34.6). The A subunit is loosely held above the plane of the B subunits, with a single α-helix penetrating through a central core created by the pentameric B subunits. Upon binding to membrane glycolipids through the B subunits, the A subunit is delivered to the interior of the cell by an unknown mechanism. In cholera toxin, the affinity of the B subunit for its glycan ligand (GM1) is unusually high compared to the binding of related AB₅ toxins from *Shigella dysenteria*, *Bordetella pertussis*, and *E. coli* to their glycan receptors (Kₐ in the nanomolar range for cholera toxin B subunit vs. millimolar range for Shiga toxin B subunit binding to α determinants). In both examples, the formation of the pentameric complex greatly
increases the avidity of the interaction. This phenomenon is being exploited to make oligovalent glycan ligands as protective agents against the toxin.

**TABLE 34.3**
Examples of receptors for bacterial toxins

![Image](image1)

**FIGURE 34.8**
Crystal structure of cholera toxin B-subunit pentamer with bound GM1 pentasaccharide shown from the bottom (a) and from the side (b). (Redrawn, with permission, from Merritt E.A., Sarfaty S., van den Akker F., L’Hoir C., Martial J.A., and Hol (more...)

_**Bacillus thuringiensis**, which lives in the soil, produces crystal toxins (Bt toxins) that can kill larval stages of plant-pathogenic insects, but they are harmless to most other organisms, including humans. Bt toxins are used in crop protection by spraying plants or by genetically engineering crops to express the toxins. Recent work demonstrates that Bt toxins act by binding to glycolipids that line the gut in nematodes and presumably other invertebrates (see Chapter 23). These receptors belong to the arthrosesies of glycolipids and include in their structure the characteristic ceramide-linked, mannose-containing core tetrasaccharide GalNAcβ1–4GlcNAcβ1–3Manβ1–4Glcβ1Cer, which is found only in invertebrates and is conserved between nematodes and insects but is absent in vertebrates (see Chapter 24). Thus, the specificity of Bt toxins determines their tissue and species sensitivity.

Shiga toxin produced by _Shigella dysenteriae_ and the homologous Shiga-like toxins of _E. coli_ (also called verotoxins) will bind to Galα1–4Gal determinants on both glycolipids and glycoproteins, but only the binding to glycolipids results in cell death. The apparent preference of most toxins for glycolipids may be related to the juxtaposition of glycolipid glycans to the membrane surface, compared to the more distal location of glycans attached to glycoproteins and proteoglycans. Binding of a toxin or bacterium to a glycolipid also might increase the likelihood of further interactions with the membrane (e.g., binding to another receptor or membrane intercalation).

**PARASITE LECTINS**

In addition to viruses and bacteria, a number of parasites also utilize glycans as receptors for adhesion (Table 34.4). _Entamoeba histolytica_ expresses a 260-kD heterodimeric lectin that binds to terminal galactose/N-acetylglactosamine residues on glycoproteins and glycolipids. Binding may have a role in attachment, invasion, and cytolyis of intestinal epithelium, and it may function in binding the amoeba to bacteria as a food source. The lectin is heterodimeric, with a transmembrane subunit of 170 kDa and a glycosylphosphatidylinositol (GPI)-anchored subunit of 35 kDa. The glycan-binding site is located in a cysteine-rich domain. The importance of this receptor in virulence has been established by antisense silencing of the adhesin and by expression of a dominant-negative form of the light subunit. Thus, the adhesin is a potential target to manage _E. histolytica_ infection (see Chapter 40).

**TABLE 34.4**
Examples of glycan receptors for parasites

![Image](image2)

The interaction of _Plasmodium falciparum_ (malaria) merozoites with red blood cells depends on sialic acids present on the host cell, in particular on the major erythrocyte membrane protein glycophorin. In this organism, attachment is mediated by a family of specific sialic acid–binding adhesin on merozoites, the most prominent of which is called EBA-175 (erythrocyte-binding antigen-175). The adhesin shows specificity for the type of sialic acid, with preference for Neu5Ac, rather than for 9-O-acetyl-Neu5Ac or Neu5Gc. Soluble Neu5Ac and Neu5Aca2–6Gal-containing oligosaccharides do not competitively inhibit the binding of EBA-175 to erythrocytes, but Neu5Aca2–3Gal-containing oligosaccharides are effective inhibitors, indicating that the adhesin is sensitive to the linkage of the sialic acid to the underlying galactose. Binding to erythrocytes leads to invasion and eventual production of additional merozoites. Other organisms expressing sialic
acid adhesins can also bind to erythrocytes (e.g., influenza virus), but these interactions do not lead to productive infections in these nonnucleated cells. Thus, under these circumstances, the erythrocyte might be considered to be a clearance mechanism for these agents (see Chapter 3).

*Plasmodium*-infected erythrocytes also express GAG-binding proteins that are thought to facilitate adherence of infected cells to tissues. The circumsporozoite form of *P. falciparum* binds to heparan sulfate in a tissue-specific manner, with preferred binding to the basal surface of hepatocytes and the basement membrane of kidney tubules. The carboxyl terminus of the protein contains positively charged residues. Clustering of the circumsporozoite protein on the surface of the organism may generate a high concentration of positively charged residues that facilitate binding. Recent studies show that the extent of sulfation of heparan sulfate that the parasite encounters determines whether it migrates or productively invades host cells.

**MICROBIAL GLYCAN LIGANDS FOR ANIMAL CELL LECTINS**

Like mammalian cells, bacteria, viruses, and parasites are covered by glycans. The enveloped viruses contain virally encoded membrane glycoproteins and may pick up host glycoproteins and glycolipids during the budding process. Bacteria contain a cell wall composed of peptidoglycan, teichoic acids (Gram-positive organisms), and lipopolysaccharide (Gram-negative organisms) and also, in some species, a capsule (see Chapter 20). All of these glycans can potentially interact with host-cell lectins. Thus, it is not surprising that interactions between microbial glycans and host lectins can also aid in infection and colonization.

*Trypanosoma cruzi* has developed an interesting strategy of molecular camouflage in which a parasite-encoded trans-sialidase transfers sialic acid from serum glycoproteins in the host to membrane proteins on its own surface. The primary function of this reaction is most likely to cover surface glycans as a way of preventing host immune reactivity, although the trans-sialidase may also act as an adhesin. *Neisseria gonorrhoeae* uses low levels of tissue CMP-sialic acid to cover itself with sialic acid residues, making it resistant to the alternative pathway of complement. It also appears likely that these and other sialylated pathogens interact with host via the Siglec family of sialic acid–binding lectins (see Chapter 32). Schistosomes, which are parasitic filarial worms, contain the Lewis^a^ antigen that is also found on human leukocytes. Because Lewis^a^ is recognized by selectins, the presence of these carbohydrates may provide a mechanism for attachment or transcellular migration of the parasite. However, these glycans also generate a massive anti-Lewis^a^ antibody response in the host.

In a similar way, the capsules and lipopolysaccharide that surround bacteria, as well as yeast cell walls, contain glycans that may be recognized by mammalian cell lectins. For example, yeast mannans are recognized by both soluble and macrophage mannose-binding proteins, which have an important role in innate immune defense during the preimmune phase in infants. Highly virulent forms of pathogenic bacteria such as *Streptococcus* contain a hyaluronan capsule, which may interact with hyaluronan-binding proteins (such as CD44) present on host-cell surfaces and facilitate colonization. The structure and biology of these types of glycans and glycan-binding proteins are discussed in Chapters 20, 21, and 40.

**ROLE IN INFECTIOUS DISEASE**

As pointed out earlier, the major function of the microbial lectins is to mediate adhesion of the organisms to host cells or tissues, which is a prerequisite for infection to occur. This has been extensively demonstrated both in vitro (in studies with isolated cells and cell cultures) and in vivo (in experimental animals) and is supported in some cases by clinical data. For example, lectin-deficient microbial mutants often lack the ability to initiate infection. In addition, type-1 fimbrial expression is associated with the severity of urinary tract infection in infants. Moreover, mono- or oligosaccharides have been shown to protect against infection by lectin-carrying bacteria in experimental models.

Glycans recognized by microbial surface lectins block the adhesion of the organisms to animal cells not only in vitro, but also in vivo, and thus protect animals against infection by such organisms. For example, coadministration of methyl α-mannoside with type-1-fimbriated *E. coli* into the urinary bladder of mice reduces significantly the rate of urinary tract infection, whereas methyl α-glucoside, which is not inhibitory to the fimbriae, has no effect. The protective effect of antiadhesive sugars has been demonstrated in a variety of studies with different pathogenic bacteria and animals, from rabbits to monkeys. Multivalent ligands should prove even more potent.

The ability of exogenous heparin and related polysaccharides to inhibit viral replication suggests that this approach might lead to polysaccharide-based antiviral pharmaceutical agents. As more crystal structures become available, the ability to custom design small-molecule inhibitors that fit into the carbohydrate-recognition domains of adhesins should improve. Already, the structures of influenza hemagglutinin and sialidase have suggested numerous ways to modify sialic acid to fit better into the active sites. Some of these compounds are presently in use to limit the spread of virus (see Chapters 50 and 51).
FURTHER READING


Publication Details

Author Information

Jeffrey D Esko and Nathan Sharon.

Copyright

Copyright © 2009, The Consortium of Glycobiology Editors, La Jolla, California.

Publisher

Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY)

NLM Citation

Structural Evolution of Glycan Recognition by a Family of Potent HIV Antibodies
Structural evolution of glycan recognition by a family of potent HIV antibodies

Fernando Garces¹,²,³, #, Devin Sok²,³,⁴, #, Leopold Kong¹,²,³, Ryan McBride⁵, Helen J. Kim¹, Karen F. Saye-Francisco²,³,⁴, Jean-Philippe Julien¹,²,³, Yuanzi Hua¹, Albert Cupo⁷, John P. Moore⁷, James C. Paulson⁵, Andrew B. Ward¹,²,³, Dennis R. Burton²,³,⁴,⁸, *, and Ian A. Wilson¹,²,³,⁶, *

¹Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, California 92037, USA
²International AIDS Vaccine Initiative Neutralizing Antibody Center, The Scripps Research Institute, La Jolla, California 92037, USA
³Scripps Center for HIV/AIDS Vaccine Immunology & Immunogen Discovery, The Scripps Research Institute, La Jolla, California 92037, USA
⁴Department of Immunology and Microbial Science, The Scripps Research Institute, La Jolla, California 92037, USA
⁵Department of Cell and Molecular Biology, The Scripps Research Institute, La Jolla, California 92037, USA
⁶Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, California 92037, USA
⁷Department of Microbiology and Immunology, Weill Medical College of Cornell University, New York, NY 10021, USA
⁸Ragon Institute of MGH, MIT and Harvard, Cambridge, MA 02139, USA

SUMMARY

© 2014 Elsevier Inc. All rights reserved.

*co-corresponding authors (to whom correspondence should be addressed. wilson@scripps.edu; burton@scripps.edu).
#co-first authors

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.

Accession Codes
Coordinates and structure factors for Fab PGT124 and its complex with gp120 and CD4 are deposited with the PDB under accession codes 4R26 and 4R2G, respectively. The Fab PGT124:SOSIP.664 trimer EM reconstruction is deposited in the EMDB under accession code EMD-2753.

SUPPLEMENTAL DATA
Supplemental Data includes 7 Figures and 6 Tables.

Author Contributions: Project design by F.G., D.S., A.B.W., D.R.B. and I.A.W.; X-ray work and analysis by F.G. and L.K.; EM work by H.K. and A.B.W.; glycan array work by R.M. and J.C.P.; mutational work by D.S. and K. F. S.; manuscript written by F.G, D.S., L.K., A.B.W., D.R.B. and I.A.W. All authors were asked to comment on the manuscript. This is manuscript # 26090 from The Scripps Research Institute.
The HIV envelope glycoprotein (Env) is densely covered with self-glycans that should help shield it from recognition by the human immune system. Here we examine how a particularly potent family of broadly neutralizing antibodies (Abs) has evolved common and distinct structural features to counter the glycan shield and interact with both glycan and protein components of HIV Env. The inferred germline antibody already harbors potential binding pockets for a glycan and a short protein segment. Affinity maturation then leads to divergent evolutionary branches that either focus on a single glycan and protein segment (e.g. Ab PGT124) or engage multiple glycans (e.g. Abs PGT121-123). Furthermore, other surrounding glycans are avoided by selecting an appropriate initial antibody shape that prevents steric hindrance. Such molecular recognition lessons are important for engineering proteins that can recognize or accommodate glycans.

INTRODUCTION

The HIV-1 envelope glycoprotein (Env) trimer is the sole target of the neutralizing antibody response and the primary platform for vaccine design. However, variable loops on gp120 mediate antibody escape and extensive N-linked glycosylation shields much of the Env protein surface from immune recognition. Additionally, many antibodies against monomeric gp120 bind without measurable glycan involvement or show enhanced binding following deglycosylation (Binley et al., 1998; Koch et al., 2003; Ma et al., 2011). Notwithstanding, a number of potent broadly neutralizing antibodies (bnAbs) have recently been discovered that bind to a heavily glycosylated region around the base of the V3 loop that we have termed a supersite of vulnerability (Kong et al., 2013). These bnAbs include antibody families from different germline lineages such as PGT121-123/PGT133-134/10-1074, PGT125-128/PGT130-131 and PGT135-137 (Julien et al., 2013c; Kong et al., 2013; Mouquet et al., 2012; Pejchal et al., 2011; Walker et al., 2011). Crystal structures of PGT128 and PGT135 in complex with gp120 outer domain and with gp120 core, respectively, and PGT122 in complex with the soluble, cleaved BG505 SOSIP.664 gp140 trimer (SOSIP.664) (Julien et al., 2013a; Julien et al., 2013b; Kong et al., 2013; Pejchal et al., 2011) have enabled molecular characterization of their glycan-dependent bnAb epitopes. Although these bnAbs are derived from different germline lineages, they all interact with the Asn332 (N332) glycan that is highly conserved across the majority of HIV-1 isolates. In addition, PGT128 binds the glycan at Asn301 (N301) and the base of the gp120 V3 loop, PGT135 interacts with glycans at Asn386 (N386) and Asn392 (N392) and an extensive β-sheet motif on the gp120 outer domain, and PGT122 contacts glycans at N301, Asn137 (N137), and Asn156 (N156), as well as protein components of the V1 and V3 loops. A family of trimer-prefering antibodies, PG9/PG16, also recognize N156 in V1 but interact with a glycan in V2, Asn160 (N160) at the trimer apex (Julien et al., 2013a; McLellan et al., 2011). A common feature of these antibodies is interaction with multiple glycans and protein components to achieve high affinity. Indeed, these same bnAbs generally have low or undetectable affinity to single glycans (McLellan et al., 2011; Mouquet et al., 2012). For carbohydrate binding lectins, high affinities that are relevant in vivo are achieved through interaction with multiple glycans (Dam et al., 2000). Only one HIV-1 antibody 2G12 has been able to attain high affinity for glycans alone by using multivalency through domain swapping of the variable heavy chain (V_H) domains, whereby two tightly linked Fabs then bind multiple glycans in the N332 high mannose patch (Calarese et al., 2003). To achieve
high affinity binding without multivalency, a combination of glycan and protein interactions would appear to be a more general solution.

PGT122 is a member of the PGT121 family of bnAbs, which are among the most potent antibodies identified to date. Passively administered PGT121 protects against mucosal SHIV (chimeric simian HIV) challenge in macaques at serum concentrations achievable by vaccination and causes a dramatic and sustained lowering of viral load in established SHIV infection (Moldt et al., 2012; Barouch et al., 2013). The crystal structure of BG505 SOSIP. 664 with PGT122 revealed how an affinity-matured antibody from the PGT121 family recognizes gp120 in the context of the Env trimer (Julien et al., 2013a). PGT124 is a newly discovered bnAb from the same germline lineage as PGT121, but represents an alternative branch in the antibody maturation process and is 89% identical in VH amino-acid sequence to 10-1074 (Figure S1C) (Mouquet et al., 2012; Sok et al., 2014; Sok et al., 2013).

Structural comparison of representatives from these two different evolutionary branches therefore provides an opportunity to investigate the evolution of high affinity recognition of an epitope involving glycans and protein surfaces. Molecular details of the PGT124-gp120 interaction were determined by X-ray crystallography and EM, with additional functional and biochemical insights from isothermal titration calorimetry (ITC), next-generation sequencing, virus neutralization and glycan arrays. Surprisingly, PGT124 engages in a distinct mode of interaction with the N332 supersite of vulnerability that involves a single glycan whereas PGT122 contacts multiple glycans.

RESULTS

Structural characterization of PGT124 in complex with gp120 and CD4

To identify the PGT124 epitope, we co-crystallized PGT124 Fab in complex with a gp120 core containing a mini V3 loop (mV3) (Pejchal et al., 2011). The JRCSF gp120 mV3 core was produced in a GnTI−/− mutant cell line that yields Man5,9GlcNAc2 oligomannose glycans (Chang et al., 2007; Reeves et al., 2002). The complex was deglycosylated by Endoglycosidase H (EndoH) to remove glycans not protected by PGT124 binding and two-domain soluble CD4 was added to facilitate crystallization. The crystal structure was determined at 3.3 Å with four ternary complexes in the asymmetric unit (Table S1 and S6).

The most prominent interaction of PGT124 with gp120 is with a single Man8GlcNAc2 glycan (Man8) linked to N332 and a small protein segment at the V3 loop base (Figure 1A, B). Structurally, both protein and glycan components make comparable contributions to the PGT124 epitope, as the bnAb buries 428 Å² of gp120 polypeptide and 480 Å² of N332 Man8 glycan (Figure 1C). While the N332 glycan was protected from EndoH deglycosylation by the antibody prior to crystallization, most other glycans appear to be cleaved, as indicated by electron density for only single N-acetylgalactosamine (GalNAc) moieties at N262, N276, N295, N301, N339, N355, N386, N392, N411 and N448. Thus, it is unlikely that the N301 glycan near the base of V3 makes productive contacts with PGT124. However, additional electron density at N295 and N339 could be assigned to two core GlcNAcs (data not shown) in one of the four copies in the asymmetric unit, suggesting partial protection from cleavage by EndoH. Electron density for gp120 was mostly
unambiguous except for the mV3 tip, which is disordered. However, the V3 base is highly ordered as PGT124 makes its only contacts with the gp120 polypeptide in this region (Figure 2A). Superposition of the unliganded PGT124 structure at 2.5 Å (Table S1) with its bound form gave an all-atom root mean square deviation (RMSD) of 0.6 Å, indicating no major changes in the antibody on gp120 binding (Figure S1A).

To characterize the PGT124 interaction in the context of the SOSIP.664 Env trimer (Depetris et al., 2012; Sanders et al., 2002), we derived negative-stain electron microscopy (EM) images at 17.8 Å resolution. Three Fabs are bound per trimer with each oriented diagonally relative to the trimer axis (Figure S2A). The PGT122:SOSIP.664 crystal structure (PDB ID:4NCO) fits extremely well into the PGT124:SOSIP.664 EM reconstruction, indicating both bnAbs approach Env at a similar angle (Figure 1D).

Molecular basis of PGT124 recognition of gp120 polypeptide

PGT124 recognizes the $^{324}$GDIR$^{327}$ region at the V3 loop base. As other N332-dependent antibodies, including PGT128, bind this same sequence, this motif may be a common anchor point. For PGT124, Ser$^{30}$ and Ser$^{93}$ in CDR L1 and L3, respectively, interact with the gp120 Asp325 carboxylate (Figure 2A), which is also involved in stacking interactions with CDR H3 Tyr$^{100B}$. Remarkably, PGT124 utilizes three different CDR loops to target Asp325. Furthermore, a salt bridge is formed between CDR H3 Glu$^{100I}$ and gp120 Arg327. However, these side-chain mediated PGT124 interactions contrast with the backbone-mediated recognition of the same $^{324}$GDIR$^{327}$ region by PGT128, thus highlighting substantial differences in how antibodies from different lineages recognize the same region in an epitope (Figure 2B). Furthermore, Asp325 can be substituted by Asn without affecting PGT124 interaction as HIV-1 isolates such as CNE20 and CAAN5342.A2 with $^{324}$GNIR$^{327}$ can be neutralized. This $^{325}$D/NxR$^{327}$ motif is 94.1% conserved across HIV-1 strains, as calculated using 3,045 aligned gp120 sequences (Los Alamos HIV Database). As expected, mutation of GDIR to GAIA completely abrogated PGT124 binding to gp120 in an ITC assay (Figure S3A, B). A JRCSF pseudovirus variant containing GAIA was poorly neutralized by PGT124 compared to wild type, while the same mutations had no effect on neutralization by PGT128 (Figure 2D). Thus, PGT124 interacts with the highly conserved V3 base in a distinct manner from PGT128. Accordingly, successful elicitation of these two types of bnAbs may require differently designed immunogens.

Further comparison between the PGT124:gp120 core-mV3 and PGT128:ODmV3 (PDBID: 3TYG) structures confirmed the GDIR motif is structurally similar (1.0 Å Cα RMSD) although significant flexibility is observed in the overall V3 loop conformation (2.4 Å Cα RMSD) where the observed ends of the V3 loop shift by 10 Å and 5 Å (Figure 2C). The PGT128 V3 conformation is likely affected by antibody binding as extensive interactions are also made with the N301 glycan within the mV3.

The PGT124:core-mV3 crystal structure was also compared with the PGT122:SOSIP.664 trimer structure (PDB ID:4NCO), where superposition indicates that the V3 base, and the GDIR motif in particular, adopts similar but not identical, conformations (Figure S4), suggesting binding does not substantially perturb the gp120 conformation in monomer or trimer context.
Man_8\textsuperscript{\textbf{9}} linked to Asn332 is the primary PGT124 glycan contact in the N332 supersite of vulnerability

The PGT124 structure reveals the molecular basis for the highly focused N332 glycan interaction. The antibody engages carbohydrate primarily at the junction of the light (LC) and heavy (HC) chains (Figure 3A) and makes multiple interactions to gp120 via CDR L2, light chain framework 3 (FR L3) and the 24-residue CDR H3 that forms an extended β-hairpin (Figure 3A). The long H3 loop penetrates the glycan shield to reach the gp120 protein surface and its extended backbone packs along the entire length of the N332 glycan. Additional glycan interactions are made with a pocket formed by CDR L2, FR L3 and CDR H3, where the C-mannose moiety forms hydrogen bonds and van der Waals interactions (Figure 3D). Site-directed alanine mutagenesis confirms that Asn^L51 and Arg^H100, which directly contact the glycan, moderately affect PGT124 neutralization, whereas Ile^L66b has a weaker effect (Figure S5). Together, these extensive glycan contacts, in combination with interactions with the GDIR motif, appear to be sufficient to attain high affinity binding, without engaging other glycans, as with PGT128 and PGT135. Nevertheless, although PGT124 contacts the three mannose moieties on the D1 arm, only two seem to impact neutralization (Figure S5). Moreover, PGT124 can neutralize viruses grown in different cell lines or in the presence of glycosidase inhibitors (Figure 3E), presumably because the N332 glycan composition is not significantly impacted as it is protected from modifying enzymes by the dense oligomannose patch on the gp120 outer domain (Bonomelli et al., 2011).

Co-crystal structures of gp120 with a variety of N332-dependent bnAbs (PGT128 (PDB ID: 3TYG), PGT135 (PDB ID: 4JM2), and PGT124) allow assessment of the structure and conformational flexibility of the N332 glycan. First, superimposition of their gp120 protein components shows that the D1, D2 and D3 arms of the N332 glycan generally project in the same direction. However, while the PGT135- and PG128-bound N332 glycans varying by only 8° in an arc from each other, the PGT124-bound N332 deviates by 20-25° from its corresponding orientations in the PGT128 and PGT135 complexes, bending closer towards the V3 loop (Figure 3A-C and Figure S6A, B). Otherwise, the N332 glycan seems to have a rather rigid architecture. Superimposition only on the N332 glycan reveals remarkably good correspondence of the glycan conformation with only slight deviations (Figure S6C) as reflected in all-atom RMSD’s of 2.5 Å and 2.9 Å for N332-PGT124 vs N332-PGT135 and N332-PGT124 vs N332-PGT128 structures, respectively. Thus, the N332 glycan has a consistent similar general orientation, but antibodies can fix some of the apparently limited conformational heterogeneity. Finally, PGT124 binds a different face of the glycan than PGT135, which interacts with the N332 D2 and D3 arms rather than D1. In contrast, PGT124 contacts the same glycan face as PGT128, which binds the D1 and D3 arms. Consequently, this variation in N332 glycan recognition and binding modes creates specific footprints for each antibody.

Although derived from the same putative germline precursor, PGT124 and PGT122 make different sets of glycan contacts. PGT124 contacts only N332 on the JRCSF gp120 core, while PGT122 contacts four different glycans on the BG505 SOSIP.664 trimer (Julien et al., 2013a). Indeed, in the absence of the N332 glycan, PGT122 requires N137 and N301 to neutralize the BG505 virus (Julien et al., 2013a). To determine if PGT124 also depends on
these glycans, we eliminated single glycan sites in the V1/V2 loops and in the high-mannose patch of Env; these changes had little to no effect on PGT124 neutralization (Table S2). Consistent with the structural data, only N332 glycan removal abrogated PGT124 neutralization across a number of isolates (Sok et al., 2014). In contrast, PGT121 exhibits isolate-specific effects in the presence or absence of N332 (Table S2) (Sok et al., 2013), likely reflecting its ability to utilize alternative glycans at either N137 or N301 (Sok et al., 2014). Removal of these glycan sites individually did not affect or even enhance PGT121 activity, implying that each glycan interaction is relatively weak. Thus, it appears that the glycans are recognized in combination with the protein component via a cooperative interaction to achieve a high-affinity interaction. Although PGT124 and PGT121/PGT122 derive from the same germline genes and have high structural homology, they have diverged significantly in how they interact with their epitopes centered on the N332 glycan. Nevertheless, the entire constellation of glycans likely dictates the approach angle of these antibodies to the Env trimer where either several (PGT122) or one (PGT124) glycans are contacted and others are avoided.

We previously reported that removal of the N332 glycan abrogates PGT124 neutralization for the overwhelming majority of isolates (Sok et al., 2014). To determine how PGT124 neutralizes HIV in the absence of N332, we investigated a small set of representative isolates (Table S3). Neighboring glycan sites were removed in combination with N332 by alanine mutagenesis, and the double mutants were tested to see whether PGT124 was dependent on them for neutralization. Loss of neutralization was found for removal of N295, N301, N339, and N392 glycan sites in combination with N332 removal, but little to no effect was seen with single mutations (Table S3). Accordingly, for a small subset of viruses, PGT124 is capable of utilizing nearby glycans when N332 is not present (Table S3).

**PGT124 binding does not appear to be dependent on other surrounding glycans on the HIV-1 trimer**

PGT124 primarily contacts N332 whereas PGT122 also contacts N137, N156, and N301 (Figure 1) (Julien et al., 2013a). Therefore, we analyzed sequences in this family in the context of known structures to assess the evolution of their binding sites during antibody maturation. Residues in common in PGT122 and PGT124 CDRs mediate contacts to N332 and GDIR (Figure 4). Indeed, almost all of these contact residues are also present in the germline (GL) and retained in various intermediates (3H, 32H, 9H for the heavy chain and 3L for the light chain) as well as in the affinity-matured antibodies from different branches of the phylogenetic tree (PGT121-124). Thus, N332 glycan and GDIR recognition appears to be the key event in initiating the antibody response.

Comparison of HIV neutralization by inferred intermediates and affinity-matured antibodies against various glycan mutants reveals considerable complexity in glycan epitope recognition. We therefore focused on the four glycan sites (N137, N156, N301, and N332) that mediate binding of PGT122 to the BG505 isolate. We tested neutralization of antibody 3H+3L, the least-mutated inferred intermediate for this lineage, as well as 32H+3L that is an inferred intermediate of PGT124, and 9H+3L as an inferred intermediate of PGT121-123. First, we observed loss of neutralization for 3H+3L upon removal of individual glycan sites.
at N301 and N332, as previously reported (Sok et al., 2013), implying use of both glycan sites early in this lineage (Table S4). Neutralization was reduced after N301 removal for 32H+3L and 9H+3L with complete loss after N332 removal, suggesting a somewhat reduced dependence on N301 during affinity maturation relative to N332. A similar pattern was observed for other isolates (Sok et al., 2013). Strikingly, N137 removal resulted in enhanced neutralization for both inferred intermediate antibodies and for the affinity-matured antibodies (Table S4), suggesting the N137 glycan plays a largely obstructive role. However, in the absence of the N332 glycan, it appears that the N137 glycan can play a positive binding role. If N332 and N137 are both eliminated, antibodies from the PGT121 family fail to neutralize virus. The overall conclusion is that the N332 glycan is preferentially used by all antibodies of the PGT121 lineage but the N137 glycan can be utilized by PGT121-123 if N332 is absent (Table S5). Moreover, this effect is also observed to a lesser degree for N301 and N156 glycans that do not generally contribute to binding in the presence of N332, but become important for PGT121-123 in the absence of N332 (Table S4). In contrast, for the PGT124 branch, the N332 glycan appears to be sufficient for neutralization and contact with N137 is avoided (Table S4 and S5). Thus, markedly different epitope recognition is observed for antibodies derived from the same germline lineage. With regard to the N301 glycan, PGT124 CDR H2 is shifted away from its equivalent position in PGT122 (Figure 5C) and towards N301 glycan (Figure 5D). In fact, PGT124 would clash with N301 if it were to maintain the same orientation it adopts in the PGT122 complex. Thus, PGT124 appears to avoid this glycan consistent with apparent cleavage by EndoH in the presence of PGT124 and no visible electron density.

As previously observed, glycan arrays indicated that PGT121 binds bi-antennary complex glycans containing terminal sialic acid moieties with α-6 linkages (Julien et al., 2013c; Mouquet et al., 2012), but not high-mannose glycans (Figure 6 A-C). PGT128 and 2G12 only bind high-mannose glycans on glycan arrays (Calarese et al., 2003; Pejchal et al., 2011). In contrast, PGT124 does not bind any glycans on these arrays (Figure 6 A-C). The lack of complex glycan binding is consistent with the structural evidence that PGT124 does not interact with V1/V2 glycans, which are likely complex. Indeed, similar data were reported for 10-1074, a closely related somatic variant of PGT124 (Figure S1B) (Mouquet et al., 2012). However, lack of detectable high-mannose binding to PGT124 on the arrays is more puzzling but consistent with weak interaction with the N332 glycan that requires cooperative interaction with the GDIR motif to attain high affinity binding. For WEAU, one of the few HIV-1 strains in which N332A mutation does not abrogate PGT124 neutralization, virus produced in the presence of kifunensine, which results in only Man₈₋₉GlcNAc₂ glycoforms, is more sensitive to PGT124 neutralization than wild-type virus, suggesting complex glycans are not required for PGT124 neutralization even in the absence of N332, in contrast to PGT121 (Figure S7).

PGT121-123 allosterically inhibit CD4 binding to gp120 by interacting with the V1/V2 loop (Julien et al., 2013c). To test whether this mechanism also applies to PGT124, we performed sequential ITC binding experiments in which CD4 was titrated into a solution containing full-length gp120 pre-complexed with PGT124 or vice-versa. We found no difference in PGT124 or CD4 binding whether or not the other was precomplexed with full-length gp120.
Thus, PGT124 does not inhibit gp120-CD4 binding, either allosterically or otherwise, which further highlights differences in epitope recognition compared to PGT121 (Figure S3E).

Discussion

Following HIV-1 infection, a proportion of individuals develop potent broadly neutralizing antibody responses over time. Typically, sets of monoclonal antibodies can be isolated that consist of somatic variants arising from a single germline precursor but individual variants can differ considerably in sequence, neutralization profiles and breadth. Investigation of these somatic variants can aid in understanding what to expect from a vaccine-induced polyclonal antibody response to a given epitope and highlight features that might be gainfully incorporated into a vaccine immunogen and in vaccination strategies designed to elicit antibodies to that epitope.

To this end, we performed a structural study of members of the highly potent PGT124 family of bnAbs. PGT121, for example, protects against virus challenge at exceptionally low serum titers in macaque models (Moldt et al., 2012). Previous work also noted differences in glycan site recognition between somatic variants (Sok et al., 2014) and phylogenetic studies reported divergence in affinity maturation between PGT121-123 and 10-1074/PGT124 (Mouquet et al., 2012; Sok et al., 2013). We therefore investigated these phenotypic differences by comparing epitope recognition between somatic variants PGT122 and PGT124 in atomic detail.

The overall binding site architectures of the PGT124/10-1074, PGT121-123, and germline precursor antibodies are strikingly similar; all contain a closed face (Julien et al., 2013c; Mouquet et al., 2012) on one side of the paratope for binding the N332 glycan, and similarly oriented long CDR loops that reach through to the protein surface below (Figure S1B). All antibodies also have an equivalent open face with a U-shaped depression; for PGT122, this face contacts surrounding glycans, whereas for PGT124, it enables avoidance of neighboring glycans. These observations suggest that, from the initial recombination event, the overall shape of the germline precursor is compatible with penetration of the glycan shield on the Env trimer in the region of N332 (Figure S1B). Indeed, the germline precursor already contains 8 of the 11 residues that are used by PGT124 to recognize the N332 glycan (5 residues) and the GDIR peptide sequence (6 residues) (Figure 4). A three amino-acid insertion in FR L3, which occurred early in the maturation process, makes key contacts to the N332 glycan, highlighting its importance for the binding and neutralization by this antibody family (Sok et al., 2013).

However, the structural analyses, as well as neutralization and binding assays, revealed a remarkable divergence in glycan epitope recognition by this antibody lineage that could have been driven by changes in the glycan sites over the course of chronic infection, although this hypothesis cannot be tested directly as longitudinal samples are not available. However, enhanced neutralization by all members of this lineage in the absence of the N137 glycan suggests that the initial encounter with this glycan was unfavorable, if it was present, but likely contributed to selecting the overall antibody shape required to bind the N332...
glycan, penetrate through the glycan shield to access the V3 base, and avoid unfavorable contacts with other neighboring glycans. Following this initial event, antibody maturation appears to have proceeded along two different paths. Via the PGT124 route, affinity maturation appears to have focused on recognition on the V3 base and the N332 glycan. For PGT121-123, affinity maturation appears to have expanded the recognition to up to four glycans, including N137, depending upon isolate context, which likely could counter potential virus escape through deletion of the N332 glycan (Figure 7). Even though the PGT124 epitope is substantially smaller than PGT121-123, PGT124’s potency and breadth match the other antibodies (Sok et al., 2014). One possible explanation is that interactions with the protein component substitute for binding to multiple glycans. Accordingly, the GDIR motif is just as important as the N332 glycan for binding and neutralization for most isolates (Figures 2D, 3E and Figure S3B-C). Thus, the single glycan at N332 is necessary but not sufficient since the GDIR motif is also required for PGT124 binding and neutralization.

Our observations also highlight the inherent flexibility in how glycans can serve as binding sites for antibodies as well as for other proteins. Furthermore, N137 glycan deletion in the presence of the N332 glycan enhances neutralization of BG505, indicating that this glycan can restrict antibody recognition. In contrast, deletion of both N137 and N332 totally abrogates PGT121-123 neutralization against the BG505 strain (Figure 7 and Table S4). Thus, the same glycan can participate in or shield a neutralizing epitope. For example, deletion of the N156 glycan has no effect on neutralization by any PGT121 lineage antibody except when N332 is absent, where some dependency is now observed (PGT122 and PGT123) (Table S4). In contrast, N332 makes productive contacts with all antibodies, confirming its role in the epitope core. The N301 glycan is important for neutralizing activities of the 3H precursor and, to a lesser extent, 9H and 32H precursors (Table S4). In contrast, in the absence of N332, PGT121 and particularly PGT122 show some dependency on N301 while deletion of both glycans completely abrogates binding of PGT123 (Table S4). In summary, several glycans may be involved in initiating an early PGT121 family response but different maturation pathways result in divergent glycan recognition profiles. This complexity in glycan recognition may be generalized to other antibodies such as PG9 and PG16 (Pancera et al., 2013; Walker et al., 2009), which both recognize N156 and N160 in V1/V2, but PG9 has evolved a preference for N160 while PG16 favors N156 (McLellan et al., 2011; Pancera et al., 2013).

These results also have implications for protein-glycan recognition and vaccine design. Broader recognition of glycoproteins or glycopeptides can be achieved by focusing recognition on the conserved core of otherwise variable glycans. Secondly, because of such conserved features, neighboring glycans may be able to substitute for some variation in glycan location. Thirdly, affinity can be enhanced by binding to multiple glycans even within a single binding site and also by latching on to the conserved regions of the protein, albeit even very short segments. For vaccine design, further considerations must be taken into account. For the PGT121 family, immunogens would first be designed to select for germline antibodies with relatively long CDR H3 lengths (26 amino acids) that could insert between the glycans and reach the underlying protein surface below. An appropriate antibody shape would be required for intimate binding of the N332 glycan (closed face in

Cell. Author manuscript; available in PMC 2015 September 25.
the PGT121 family), and another cavity (open face) to contact (PGT121-123) or avoid (PGT124) surrounding glycans. To facilitate initial engagement, an immunogen without the N137 glycan may enhance epitope recognition as illustrated for engagement of germline antibodies to the CD4 binding site by removal of a blocking glycan (Jardine et al., 2013). It is not yet clear whether removal of a blocking glycan would compromise selection of an appropriately shaped antibody. Following this initial selection event, modified immunogens would be used to drive clonal diversity of the selected antibody germline family that could include adding back or altering glycans to emulate events that likely occurs during virus evolution under immune pressure. To achieve a productive angle of approach of the antibody in the context of the Env trimer, the immunogen would likely require a native glycan shield or native-like Env structure to encourage accommodation of potentially obstructive neighboring glycans and drive somatic mutations to productively utilize neighboring glycans. Thus, these results provide new insights into protein recognition of glycosylated proteins and demonstrate the remarkable ability of the human immune system to find and evolve multiple solutions for high affinity antibody binding of complex epitopes so as to neutralize highly variable and highly glycosylated viruses, such as HIV.

**EXPERIMENTAL PROCEDURES**

**Expression and purification of proteins**

CD4, was expressed in *E. coli* BL-21 DE3* cells in LB media and purified by affinity chromatography followed by size exclusion chromatography (SEC). PGT124 Fab was produced in FreeStyle™ 293F cells (Invitrogen) and purified by affinity chromatography followed by cation exchange chromatography and SEC. JRCSF gp120 core was produced in FreeStyle™ 293S. Protein was purified using a GN Lectin column, followed by SEC. BG505 SOSIP.664 trimer was expressed in FreeStyle™ 293F or 293S cells and purified using a 2G12-coupled affinity matrix followed by SEC (Supplemental Experimental Procedures).

**Formation of protein complexes, crystallization and data collection**

Complexes were formed by combining gp120:ligands in a 1:1.2 molar ratio, followed by deglycosylation with Endoglycosidase H before being purified by SEC. Crystals of PGT124-gp120corenV3-CD4 ternary complex were obtained from 2.4 M (NH₄)₂SO₄, 0.1 M Tris pH 8.0 and 13% glycerol that diffracted to 3.3Å resolution (Table S1). Unliganded PGT124 Fab crystals appeared from 20% (w/v) PEG 4000, 0.2M MgCl₂, 0.1 M Tris-HCl pH 8.5 and diffracted to 2.5 Å (Supplemental Experimental Procedures and Table S1).

**Structure determination and refinement**

The unliganded PGT124 structure was solved by molecular replacement using Phaser with the PGT122 Fab structure (PDB 4JY5) as the initial model, in which the CDR loops had been deleted. For the ternary complex, multiple components were used for phasing: CD4-gp120 (PDB ID:4JM2) and the high-resolution unliganded PGT124 Fab. Model building and refinement were carried out using Coot (Emsley and Cowtan, 2004) and Phenix (Adams et al., 2010), respectively. Table S6 shows the components that were included in the final model. R<sub>cryst</sub> and R<sub>free</sub> for the unliganded structure are 23.5% and 26.7%, and 21.0% and
26.3% for the complex (Table S1). Buried molecular surface areas were analyzed with the Molecular Surface Package (Connolly, 1993) using a 1.7 Å probe radius. The Fab residues were numbered according to Kabat (Martin, 1996) and gp120 is numbered following the Hxbc scheme (Ratner et al., 1987).

Isothermal titration calorimetry

ITC binding experiments were performed using a MicroCal Auto-iTC200 instrument (GE). All proteins were extensively dialyzed against a buffer containing 20 mM Tris, 150 mM NaCl, pH 7.4 and protein concentrations were adjusted and confirmed using calculated extinction coefficients and absorbance at 280 nm. In the syringe, ligand was either PGT124 Fab or sCD4 at concentrations ranging between 60-125 μM. The gp120 monomer was in the cell at concentrations ranging between 3.5-10 μM. Two-protein binding experiments were performed as follows: cell at 25°C, 16 injections of 2.5 μl each, injection interval of 180 s, injection duration of 5 s, and reference power of 5 μcals. To perform sequential binding experiments, the mixed sample from the first titration was left in the cell and the concentration of the HIV-1 component was recalculated based on the dilution from the first experiment (approximately ~88% of the initial concentration). Subsequently, either PGT124 Fab or sCD4 was added in a second titration. To calculate the affinity constants (K_D), the molar reaction enthalpy (ΔH) and the stoichiometry of binding (N), Origin 7.0 software was used by fitting the integrated titration peaks using a single-site binding model.

Antibody and envelope substitutions

Substitutions in the PGT124 paratope, the HIV-1 Env glycoprotein, and the JRCSFmV3 construct were introduced using QuikChange site-directed mutagenesis (Stratagene, La Jolla, CA). Substitutions were verified by DNA sequencing (Retrogen, San Diego, CA).

Neutralization assays

Neutralization activity of antibodies against pseudovirus in TZM-bl cells was determined as described previously (Li et al., 2005).

Electron microscopy (EM)

PGT124 Fab in complex with BG505 SOSIP.664 gp140 produced in HEK 293S cells were analyzed by negative stain EM. A 3 μL aliquot of 10 μg/ml of the complex was applied for 15s onto a glow discharged, carbon coated 400 Cu mesh grid and stained with 2% uranyl formate for 20s. Grids were imaged using a FEI Tecnai T12 electron microscope operating at 120 kV using 52,000 × magnification and electron dose of 25 e⁻/Å², resulting in a pixel size of 2.05 Å at the specimen plane. Images were acquired with a Tietz 4k × 4k CCD camera in 5° tilt increments from 0° to 55° at a defocus of 1000 nm using LEGINON (Suloway et al., 2005).

Image processing

Particles were picked automatically by using DoG Picker and put into a particle stack using the Appion software package (Lander et al., 2009; Voss et al., 2009). Initial reference-free 2D class averages were calculated using unbinned particles via the Xmipp Clustering 2D
Alignment and sorted into 400 classes (Sorzano et al., 2010). Particles corresponding to the complexes were selected into a substack and another round of reference-free alignment was carried out using Xmipp Clustering 2D alignment and IMAGIC softwares (van Heel et al., 1996). To generate an ab initio 3D starting model, a template stack of 120 images of 2D class averages was used with imposing C3 symmetry. This starting model was refined against 12,245 raw particles for 30 cycles using EMAN (Ludtke et al., 1999). The resolution of the final reconstruction was calculated to be 17.8 Å using an FSC cut-off of 0.5 (Figure S2).

**Glycan array production**

Glycan arrays were custom printed on a MicroGridII (Digilab) contact microarray robot equipped with StealthSMP4B microarray pins (Telechem) as previously described. Briefly, samples of each glycan were diluted to 100uM in 150 NaPO4 buffer, pH 8.4. Aliquots of 10uL were loaded in 384-well plates and imprinted onto NHS-activated glass slides (SlideH, Schott/Nexterion) with 7 arrays, each containing 6 replicates of each sample at both concentrations. Printed slides were humidified post-print for 1h and desiccated overnight. Remaining NHS-ester residues were quenched by immersing slides in 50mM Ethanolamine in 50mM borate buffer, pH 9.2, for 1h. Blocked slides were washed with water, spun dry and stored at room temperature until use.

**Glycan array screening and analyses**

Antibodies were diluted to 30ug/mL + 15ug/mL anti-human-IgG-RPE (JacksonImmuno) in PBS + 0.05% Tween-20. The prepared mixture was incubated for 30 min on ice and incubated on the array surface in a humidified chamber for 1 hour. Slides were subsequently washed by successive rinses with PBS-T, PBS and deionized H2O. Washed arrays were dried by centrifugation and immediately scanned for FITC and R-PE signal on a Perkin-Elmer ProScanArray Express confocal microarray scanner. Fluorescent signal intensity was measured using Imagene (Biodiscovery) and mean intensity minus mean background was calculated and graphed using MS Excel.

**Supplementary Material**

Refer to Web version on PubMed Central for supplementary material.

**Acknowledgements**

We are very grateful to N. Laursen and X. Dai for assistance with data processing; A. Irinia, and A. Sarkar for assistance with glycan refinement; N. Laursen and R. Stanfield for helpful discussions; M. Elsliger for computer support; M. Deller and H. Tien for crystallization screening; C. Corbaci for help with preparing figures; and J.P. Verenini for manuscript formatting. X-ray data sets were collected at the Stanford Synchrotron Radiation Lightsource (SSRL), a Directorate of the SLAC National Accelerator Laboratory and an Office of Science User Facility operated for the U.S. Department of Energy (DOE) Office of Science by Stanford University. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research; NIH’s National Center for Research Resources, Biomedical Technology Program (P41RR001209); and the National Institute of General Medical Sciences (NIGMS). Electron microscopy data were collected at the National Resource for Automated Molecular Microscopy (NRAMM) at the Scripps Research Institute, which is supported by US National Institutes of Health (NIH) through the National Center for Research Resources’ P41 program (RR017573). This work was supported by the International AIDS Vaccine Initiative Neutralizing Antibody Center; by the Center for HIV/AIDS Vaccine Immunology and Immunogen Discovery (CHAVI-ID U1ID AI00665) (A.B.W., D.R.B., I.A.W.); by the HIV Vaccine Research and Design (HIVRAD) program (P01 AI832362 and R37 Garces et al. Page 12
References


Highlights

- Potent broadly neutralizing HIV antibody PGT124 contacts both glycan and protein
- PGT124 contrasts with other family members by contacting only a single glycan
- Inferred germline antibody incorporates features important for protein and glycan recognition
- Antibody maturation diversifies modes of glycan recognition
Figure 1. Crystal structure of PGT124 in complex with HIV-1 gp120 and CD4

(A) The ternary complex of PGT124 (purple and brown for heavy and light chains respectively) and 2-domain CD4 (gray) bound to gp120 (light blue) is shown as a ribbon representation. PGT124 contacts the N332 glycan and the base of V3. The glycan is represented in ball-and-stick, with green carbons, red oxygens, and blue nitrogens, and covered with its molecular surface. (B) CDR H3, L1, L2, L3 loops, and the light chain FR3 make extensive contacts with both the Man₈ N332 glycan and V3 (in red) of gp120. (C) The surface areas (Å²) buried on the gp120 protein and N332 glycan by individual CDR loops and framework regions on PGT124 are presented on the lower right. (D) Fitting of the PGT122-BG505 SOSIP.664 gp140 crystal structure into the PGT124:BG505 SOSIP EM reconstruction at 17.8 Å resolution. The PGT122 fitting is excellent with a correlation of 0.9, indicating that PGT122 and PGT124 share a similar angle of approach for interaction with Env (see also Figures S1 and S2 and Tables S1 and S6).
Figure 2. Interaction between PGT124 and the base of the V3 loop

(A) The interaction between PGT124 CDR loops (purple ribbons) and the base of the V3 loop (red ribbon). Red dots represent residues at the tip of the mini V3 loop that are not visible in the electron density. The interaction with contacting residues is displayed in ball-and-stick representation below. Contacts between side chains of the $^{324}$GDIR$^{327}$ motif on the V3 loop and the CDR loops of PGT124 are indicated with solid lines for van der Waals’ interactions and dotted lines for hydrogen bonds. (B) For comparison, the interaction between PGT128 CDR loops (pale green) and the base of the V3 (gray) (PDB 3TYG) for comparison with (A). The PGT128 antibody-antigen interactions with the V3 base are all H-bonds of the Fab main chain with V3 main chain. (C) The V3 loop structures bound to PGT124 or PGT128 are superimposed for comparison. (D) The effect on the neutralization activity of PGT124 and PGT128 by single and double mutations of the V3 loop GDIR motif. Values are shown as IC$_{50}$ in μg/mL (see also Figures S3, S4 and S5).
Figure 3. Interaction between PGT124 and the N332 glycan

(A) Interaction of the CDR loops of PGT124 in purple with the oligomannose glycan at the N332 position in red ball-and-stick representation with the D1, D2 and D3 arms labeled.  
(B) For comparison, the interaction of the PGT128 CDR loops (light green) is shown with the N332 glycan (gray ball-and-stick representation).  
(C) The N332 glycan structures bound to PGT128 or PGT124 are shown after superposition of the gp120 protein components. Conformational differences are highlighted by comparative distances between equivalent mannose moieties (D) Schematic diagram of the Man$_6$GlcNAc$_2$ glycan with the glycan moiety nomenclature. (E) The effect on PGT124 neutralization activity of differences in gp120 glycosylation arising from different cell expression systems and removal of the N332 glycan. Values are shown as IC$_{50}$ in μg/mL (see also Figures S3, S5 and S6 and Tables S2 and S3).
Figure 4. Epitope core of the PGT121 germline lineage

(A) Cartoon representation of the interaction between mV3-gp120 core and the PGT124 variable region (left panel) and close-up views of the PGT124 epitope core with the $^{324}$GDIR$^{327}$ motif at the V3 base in blue spheres and the N332 glycan as pale green ball-and-stick (middle and right panels). The PGT124 residues that contact the N332 glycan and the V3 protein motif are represented in red and orange as a ball-and-stick, respectively, and covered with its molecular surface. (B) Sequence alignment of heavy and light chains of PGT121-124 Abs and germline precursors (* the germline CDR L3 and CDR H3 sequences are not used for alignment because of ambiguity associated with D gene assignment). Red and orange boxes indicate residues that contact the N332 glycan and the GDIR motif, respectively, on the epitope core of PGT124, are also largely conserved in PGT121-123 and, importantly, in the germline sequence (see also Figure S1).
Figure 5. PGT122 contacts glycans at N137 and N301 but PGT124 accommodates these glycans without forming contacts

(A) The PGT122 epitope is composed of V1 and V3 regions from gp120, as well as four glycans at N137, N156, N301 and N332. The CDR loops and FR3 are shown in cartoon representation (in gray) (PDB ID:4NCO). (B) Superposition of PGT124 (in purple) on the PGT122 epitope. (C) Although PGT124 seems capable of accommodating a glycan in the open face formed by H2 and H3, those loops appear to be shifted towards the center of the epitope, when compared to the PGT122 paratope, thereby avoiding contacts with the N137 glycan. (D) However, PGT124 is predicted to clash with the 301 glycan and presumably pushes this glycan away to access the protein surface below.
Figure 6. Evolution of recognition of complex glycans by the PGT121 family

(A) Key residues on PGT121 (Kabat numbering) that contact the complex glycan attached at N137 of gp120 V1 are represented in red as a stick-and-ball and covered with its molecular surface on the ribbon representation of the antibody variable regions (left). Residues at equivalent positions on PGT124 are highlighted in orange (right). (B) Sequence alignment for heavy chain variable region for the germline precursor (GL) (* the germline CDR H3 sequence is not used for the alignment because of ambiguity associated with D gene assignment), intermediate precursors (3H, 32H and 9H) and mature antibodies of the PGT121 family (PGT121-124). Red boxes indicate residues involved in N137 glycan binding by PGT121 while orange boxes indicate the residues at equivalent positions on PGT124. (C) Glycan arrays illustrate that PGT124, in contrast to PGT121, PGT128 or 2G12, does not bind either complex or high mannose glycans with any demonstrable affinity (see also Figures S1, S3 and S7).
Figure 7. Summary of the contributions of the N137, N156, N301 and N332 glycans to recognition of the BG505 isolate by antibodies of the PGT121 lineage

Early precursors are highly dependent on the N301 and N332 glycans for neutralization, and negatively impacted by the presence of the N137 glycan (* N137 glycan can play a double role in glycan recognition: first it can inhibit antibody binding by shielding the epitope, here represented in blue; second, although it can still inhibit antibody binding to some extent, the N137 glycan can also be used by PGT121-3 to neutralize BG505 in the absence of N332, here represented in blue/red). Affinity maturation appears to create two distinct lineages that evolve separate dependencies on glycans that surround N332. N332 is the only critical glycan in the PGT124 epitope, while the N137 glycan contributes negatively to recognition by this antibody. Fold changes in neutralization IC\(_{50}\) for single and double glycan mutants (Table S4) were used to determine whether a glycan positively contributes to the epitope (decrease in neutralization IC\(_{50}\) upon removal) or contributes to shielding of the epitope (increase in neutralization IC\(_{50}\) upon removal). A relative scale for both positive and negative glycan contributions to the epitope was determined based on the extent of the change in neutralization IC\(_{50}\). Interestingly, for PGT121-123, the N137 glycan contributes to shielding in the presence of N332, but is a critical part of the epitope in the absence of N332 (see also Table S4 and S5).
Role of Receptor Binding Specificity in Influenza A Virus Transmission and Pathogenesis
Role of receptor binding specificity in influenza A virus transmission and pathogenesis

Miranda de Graaf & Ron A M Fouchier*

Abstract

The recent emergence of a novel avian A/H7N9 influenza virus in poultry and humans in China, as well as laboratory studies on adaptation and transmission of avian A/H5N1 influenza viruses, has shed new light on influenza virus adaptation to mammals. One of the biological traits required for animal influenza viruses to cross the species barrier that received considerable attention in animal model studies, in vitro assays, and structural analyses is receptor binding specificity. Sialylated glycans present on the apical surface of host cells can function as receptors for the influenza virus hemagglutinin (HA) protein. Avian and human influenza viruses typically have a different sialic acid (SA)-binding preference and only few amino acid changes in the HA protein can cause a switch from avian to human receptor specificity. Recent experiments using glycan arrays, virus histochemistry, animal models, and structural analyses of HA have added a wealth of knowledge on receptor binding specificity. Here, we review recent data on the interaction between influenza virus HA and SA receptors of the host, and the impact on virus host range, pathogenesis, and transmission. Remaining challenges and future research priorities are also discussed.

Keywords A/H5N1 influenza virus; A/H7N9 influenza virus; airborne transmission; pathogenesis; receptor binding specificity

DOI 10.1002/embj.201387442 | Received 19 November 2013 | Revised 17 February 2014 | Accepted 17 February 2014 | Published online 25 March 2014

EMBO J (2014) 33, 823–841

Influenza virus zoonoses and pandemics

Influenza A viruses can infect a wide range of hosts, including humans, birds, pigs, horses, and marine mammals (Webster et al., 1992). Influenza A viruses are classified based on the antigenic properties of the major surface glycoproteins hemagglutinin (HA) and neuraminidase (NA). To date, 18 HA and 11 NA subtypes have been described. All subtypes have been found in wild aquatic birds except for the recently discovered H17N10 and H18N11 viruses, which have only been detected in bats (Webster et al., 1992; Fouchier et al., 2005; Tong et al., 2012, 2013). Throughout recent history, avian-origin influenza viruses have crossed the species barrier and infected humans. Some of these zoonotic events resulted in the emergence of influenza viruses that acquired the ability to transmit between humans and initiate a pandemic. Four pandemics were recorded in the last century: the 1918 H1N1 Spanish pandemic, the 1957 H2N2 Asian pandemic, the 1968 H3N2 Hong Kong pandemic, and the 2009 H1N1 pandemic (pH1N1) that was first detected in Mexico (reviewed in Sorrell et al., 2011). Various other influenza A viruses of pig and avian origin (e.g., of subtypes H5, H6, H7, H9, and H10) have occasionally infected humans—sometimes associated with severe disease and deaths—but these have not become established in humans.

The H5N1 avian influenza virus that was first detected in Hong Kong in 1997 has frequently been reported to infect humans and cause serious disease. As of October 8, 2013, WHO has been informed of 641 human cases of infection with H5N1 viruses, of which 380 died (www.who.int/influenza/human_animal_interface/en/). The majority of these cases occurred upon direct or indirect contact with infected poultry. Due to the high incidence of zoonotic events, the enzootic circulation of the virus in poultry, and the severity of disease in humans, the H5N1 virus is considered to pose a serious pandemic threat. Fortunately, sustained human-to-human transmission has not been reported yet (Kandun et al., 2006; Wang et al., 2008).

A more recent example of a major zoonotic influenza A virus outbreak started in China in early 2013. This outbreak, caused by an avian H7N9 virus, resulted in 137 laboratory-confirmed human cases of infection and 45 deaths (www.who.int/influenza/human_animal_interface/en/). Although one case of possible human-to-human transmission was described, thus far, this outbreak also has not lead to sustained transmission between humans (Gao et al., 2013b; Qi et al., 2013). Some mild cases of infection were reported, but the Chinese H7N9 influenza viruses frequently caused severe illness, characterized by severe pulmonary disease and acute respiratory distress syndrome (Gao et al., 2013a,b). Although influenza viruses of the H7 subtype have sporadically crossed the species barrier in the past, with outbreaks reported, for example, in the United Kingdom, Centers for Disease Control and Prevention (2004), Canada (Tweed et al., 2004), the Netherlands (Fouchier et al., 2004), and Italy (Puzelli et al., 2005; www.who.int/influenza/human_animal_interface/en/), the 2013 H7N9 virus appeared to

© 2014 The Authors
jump the species barrier more easily and was generally associated with more serious disease in humans. Interestingly, the internal genes of the H7N9 virus belong to the same genetic lineage as the internal genes of the H5N1 virus; both are derived from H9N2 viruses (Guan et al., 1999; Lam et al., 2013). The H7N9 virus is thought to have emerged upon four reassortment events between H9N2 viruses and avian-origin viruses of subtype H7 and N9 (Lam et al., 2013). The H9N2 virus itself has been shown to have zoonotic potential as well, resulting in relatively mild human infections upon contact with poultry. H9N2 viruses remain enzootic in domestic birds in many countries of the Eastern Hemisphere (Peiris et al., 1999).

For all influenza pandemics of the last century, the animal-origin viruses acquired the ability to transmit efficiently via aerosols or respiratory droplets (hereafter referred to as “airborne transmission”) between humans (Sorrell et al., 2011). To evaluate airborne transmission in laboratory settings, ferret and guinea pig transmission models were developed, in which cages with uninfected recipient animals are placed adjacent to cages with infected donor animals. The experimental setup was designed to prevent direct contact or fomite transmission, but to allow airflow from the donor to the recipient ferret (Lowen et al., 2006; Maines et al., 2006). Using the ferret transmission model, pandemic and epidemic viruses isolated from humans and animals are generally transmitted efficiently via the airborne route, whereas avian viruses are generally not airborne transmissible (Sorrell et al., 2011; Belser et al., 2013b).

Why some animal influenza viruses frequently infect humans, whereas others do not, and how viruses adapt to become airborne transmissible between mammals have been key questions in influenza virus research over the last decade. Studies on H1N1, H2N2, and H3N2 viruses from the 1918, 1957, and 1968 pandemics, respectively, revealed that interaction of HA with virus receptors on host cells was a critical determinant of host adaptation and airborne transmission between ferrets (Tumpey et al., 2007; Pappas et al., 2010; Roberts et al., 2011). Sorrell and colleagues showed that a wild-type avian H9N2 virus was not airborne transmissible in ferrets, but reassortment with a human H3N2 virus and subsequent adaptation in ferrets yielded an airborne transmissible H9N2 virus, primarily due to changes in the H9N2 HA and NA surface glycoproteins (Sorrell et al., 2009). Several laboratories have studied the transmissibility of avian H5N1 viruses and their potential to become airborne, and four studies recently described airborne transmission of laboratory-generated H5N1 influenza viruses. Three of these studies—two using ferrets and one using guinea pigs—used H5 reassortant viruses between human pH1N1 or H3N2 viruses and H5N1 virus (Chen et al., 2012; Imai et al., 2012; Zhang et al., 2013d). Herfst et al. (2012) were the first to show mammalian adaptation of a fully avian H5N1 virus to yield an airborne transmissible virus in ferrets. Although avian H7 influenza viruses are generally not transmitted via the airborne route between ferrets (Belser et al., 2008), H7N9 strains from the Chinese 2013 outbreak were shown to be transmitted via the airborne route without further adaptation, albeit less efficiently as compared to human seasonal and pandemic viruses (Belser et al., 2013a; Richard et al., 2013; Watanabe et al., 2013; Xu et al., 2013; Zhang et al., 2013a; Zhu et al., 2013). No transmission was observed from pigs to ferrets (Zhu et al., 2013).

It has thus become increasingly clear that animal influenza viruses beyond the H1, H2, and H3 subtypes—specifically H5, H7, and H9—can acquire the ability of airborne transmission between mammals. To date, the exact genetic requirements for animal influenza virus to cross the species barrier and establish efficient human-to-human transmission remain largely unknown, but receptor binding specificity is clearly one of the key factors.

**HA as determinant of receptor binding specificity**

As a first step to entry and infection, influenza viruses attach with the HA protein to sialylated glycan receptors on host cells. The influenza virus HA protein is a type I integral membrane glycoprotein, with a N-terminal signal sequence. Post-translational modifications include glycosylation of the HA and acylation of the cytoplasmic tail region. Cleavage of the HA (HA0) by cellular proteases generates the HA1 and HA2 subunits, which form a disulfide bond-linked complex. The HA protein forms trimers, with each monomer containing a receptor binding site (RBS) capable of engaging a sialylated glycan receptor, a vestigial esterase subdomain, and a fusion subdomain (Fig. 1).

Hirst (1941) was the first to demonstrate the ability of influenza viruses to agglutinate and elute from red blood cells. Treatment of these cells with *Vibrio cholerae* neuraminidase (VCNA) revealed that the ability to agglutinate and elute was dependent on sialic acid (SA; Gottschalk, 1957). Early comparisons of influenza viruses isolated from different species revealed differences in their abilities to agglutinate red blood cells from various species. Using red blood cells and cells expressing only a certain type of SA, it was discovered that human and animal influenza viruses displayed differences in the receptor specificity of HA (Rogers & Paulson, 1983). Avian influenza viruses were found to have a preference for SA that are linked to the galactose in an α2,3 linkage (α2,3-SA; Rogers & Paulson, 1983; Nobusawa et al., 1991). In contrast, human influenza viruses (H1, H2, and H3) attached to SA that are linked to galactose in an α2,6 linkage (α2,6-SA; Gambaryan et al., 1997; Matrosovich et al., 2000). Pig influenza viruses either attached to both α2,3-SA and α2,6-SA, or exclusively to α2,6-SA (Rogers & Paulson, 1983). These gross differences in receptor binding properties were found to be important determinants of virus host range and cell and tissue tropism. A wide range of assays has been developed to assess HA binding specificity and affinity to glycans containing α2,3-SA and α2,6-SA (Paulson et al., 1979; Sauter et al., 1989; Takemoto et al., 1996; Stevens et al., 2006a; Chutinimitkul et al., 2010b; Lin et al., 2012; Matrosovich & Gambaryan, 2012). Binding assays such as glycan arrays, which allow the investigation of hundreds of different sialylated glycan structures, have demonstrated that influenza virus receptor specificity also involves structural modifications of the SA and overall glycan (Gambaryan et al., 2005; Stevens et al., 2006b).

**Influenza virus receptors**

**Types of glycan structures**

Mammalian cells are covered by a glycoalyx, which consists of glycolipids, glycoproteins, glycolipid anchors and proteoglycans (Varki & Varki, 2007; Varki & Sharon, 2009). Glycans represent
one of the fundamental building blocks of life and have essential roles in numerous physiological and pathological processes (reviewed in Hart, 2013). Glycans that are exposed on the exterior surface of cells play an important role in the attachment of toxins and pathogens like influenza A viruses.

Several types of glycans exist, including N-glycans, O-glycans, and glycolipids (reviewed in Brochhausen et al., 2009; Schnaat et al., 2009; Stanley et al., 2009). The N-glycans can be further divided into different types such as “complex,” “hybrid,” and “oligomannose.” There is a wide variety of O-glycans, with eight different core structures. Cores 1 and 2 are common structures that are present on glycoproteins (Fig 2). This particular classification of glycan structures is not necessarily relevant for binding of influenza viruses. For O-glycans and N-glycans, there is extensive structural diversification at the termini of glycan chains, which is potentially more important than differences in the core structures.

Receptor-mediated entry of influenza viruses is reduced in the absence of sialylated N-glycans (Chu & Whittaker, 2004; de Vries et al., 2012). However, specific roles of sialylated N-glycans, O-glycans, and glycolipids for binding and entry of influenza viruses are not fully understood. Mouse cells that are deficient in the production of glycolipids can be infected efficiently with human H5N2 virus, indicating that sialylated glycolipids are not essential for infection (Ichikawa et al., 1994; Ablan et al., 2001). Hamster cells that are deficient in the production of sialylated N-glycans cannot be efficiently infected with influenza viruses in the presence of serum (Robertson et al., 1978; Chu & Whittaker, 2004). However, when the production of sialylated N-glycans was restored, H1N1 virus was able to bind and infect these cells (Chu & Whittaker, 2004). Furthermore, it was shown that internalization of influenza via macropinocytic endocytosis is dependent on N-glycans, but that for clathrin-mediated endocytosis, N-glycans are not required (de Vries et al., 2012). It should be noted that most studies are based on in vitro work with human viruses; whether the same principles hold true in vivo and for avian viruses is not known.

**Types of sialic acid**

SAs share a nine-carbon backbone and are among the most diverse sugars found on glycan chains of mammalian cell surfaces. Common SAs found in mammals are N-acetylneuraminic acid (Neu5Ac) and N-glycolylneuraminic acid (Neu5Gc; Fig 3; reviewed in Varki & Varki, 2007). Neu5Gc has been detected in pigs, monkeys, and a number of bird species, but humans and some avian species like chickens or other poultry, do not contain Neu5Gc or only minor quantities (Chou et al., 1998; Muchmore et al., 1998; Schauer et al., 2009; Walther et al., 2013).

Lectins *Maackia amurensis* agglutinin (MAA) and *Sambucus nigra* agglutinin (SNA) recognize α2,3-SA and α2,6-SA, respectively, and are often used to study the distribution and expression of influenza virus receptors in tissues and cells. There are two different isotypes of MAA: MAA-1 (also known as MAL or MAM) and MAA-2 (also known as MAH). MAA-1 binds primarily to SAα2,3Galβ1-4GlcNAc that is found in N-glycans and O-glycans. MAA-2 preferentially binds to SAα2,3Galβ1-3GlcNAc, which is found in O-glycans. Unfortunately, both MAA-1 and MAA-2 also bind strongly to sialylated glycan structures without SA. The exact specificities of MAA-1 and MAA-2 are reviewed comprehensively by Geisler and Jarvis (2011). To determine which glycan structures present α2,3-SA and α2,6-SA, MAA and SNA can be used in combination with the lectins Concanavalin and Jacalin to detect N-glycan and O-glycan structures, respectively. Unfortunately, lectin staining provides little information on the differences in overall glycan structures, including the number of “antennas”, fucosylation, and the presence of N-acetylactosamine repeats, all of which can influence influenza virus receptor binding. This encouraged scientists to elucidate the glycome of the respiratory tissue of several influenza virus host species.

**SA receptor distribution in humans, pigs, birds, and ferrets**

In humans, α2,6-SAs are expressed more abundantly in the upper respiratory tract compared to the lower respiratory tract (Fig 4). In the nasopharynx of humans, α2,6-SA and α2,3-SA are detected on ciliated cells as well as mucus-producing cells. In the bronchus,
α2,3-SA and α2,6-SA are heterogeneously distributed in the epithelium with no clear distinction between ciliated and non-ciliated cells. Bronchial epithelium contained a higher percentage of α2,3-SA compared to α2,6-SA (Nicholls et al., 2007a). Staining for α2,3-SA (MAA-II) was most abundant in cells lining the alveoli, most likely type II cells (Shinya et al., 2006; Nicholls et al., 2007a). The respiratory tract of young children expresses more α2,3-SA and a lower level of α2,6-SA compared to adults (Nicholls et al., 2007a).

To identify the structures of sialylated glycans present in the human respiratory tract, glycomic analysis of human bronchus, lung, and nasopharynx was performed (Walther et al., 2013). Heterogeneous mixtures of bi-, tri-, and tetra-antennary N-glycans with varying numbers of N-acetyllactosamine repeats were detected, as well as significant levels of core fucosylated structures and high mannose structures. Compared to the glycan composition of the bronchus and lung, the nasopharynx displayed a smaller spectrum of glycans, with fewer N-acetyllactosamine repeats. Both α2,6-SA and α2,3-SA receptors were detected, but with a higher percentage of α2,3-SA receptors in the bronchus compared to the lungs. Sialylated core-1 and core-2 O-glycans were present.

The glycan composition of the human respiratory tract was also compared with the glycans present on available glycan arrays (Table 1). The O-glycan composition of the human respiratory tract was represented well on glycan arrays, but the complex N-glycan structures were underrepresented. On the other hand, a large number of glycans that are present on the array were not detected in the human respiratory tract and can therefore be largely ignored when studying human influenza virus receptor specificity.

Pigs are often regarded as a potential “mixing vessel” of human and avian influenza viruses, because pigs express both α2,3-SA and α2,6-SA in the respiratory tract (Scholli, 1990; Ito et al., 1998; Nelli et al., 2010; Van Poucke et al., 2010; Fig 4). Given the recent data on SA presence in the human respiratory tract (see above), this role as a mixing vessel may not be unique to pigs. The α2,6-SA receptors were dominant in the ciliated epithelia along the upper respiratory lining of the trachea and bronchus. There was a gradual increase in α2,3-SA receptors toward the lower respiratory lining. The mucosa of the respiratory tract predominantly contained α2,3-SA, with α2,6-SA mainly confined to mucous/serous glands. The α2,3-SA receptors were more highly expressed in the epithelial lining of the lower respiratory tract (bronchioles and alveoli; Nelli et al., 2010).

Primary pig respiratory epithelial cells, trachea, and lung tissue were described to contain high mannose structures and complex glycans with core fucosylated bi-, tri-, and tetra-antennary structures. In the pig respiratory epithelial cells, structures bearing N-acetyllactosamine repeats were expressed, but there was a relative lack of these structures in pig trachea and lung tissue (Table 1). Both Neu5Gc and Neu5Ac were present, with a higher abundance of the latter (Bateman et al., 2010; Sriwilaijaroen et al., 2011; Chan et al., 2013b). Pig lung contains both core 1 and core 2 O-glycan structures. For N-glycans, there was a higher abundance...
of α2,6-SA in the lungs and trachea compared to α2,3-SA (Chan et al., 2013b). The pig respiratory tract presents a smaller diversity of sialylated N-glycan structures compared to the glycome of the human respiratory tract (Sriwilaijaroen et al., 2011; Walther et al., 2013).

Avian species have α2,3-SA and α2,6-SA in both the respiratory and intestinal tract, although there are differences in the abundance of these receptors between different species (Franca et al., 2013). Chickens and common quails express α2,3-SA and α2,6-SA on a large range of different epithelial cells in the respiratory tract and intestine (Nelli et al., 2010; Trebbien et al., 2011; Costa et al., 2012; Fig 4). There is conflicting data for the presence of α2,6-SA in the intestine of mallards: Some groups found expression of α2,6-SA, whereas others did not (Kuchipudi et al., 2009; Costa et al., 2012; Franca et al., 2013).

Chicken red blood cells are frequently used for influenza virus agglutination assays. Analysis of its glycome revealed that chicken red blood cells present a plethora of N-glycans, both di-, tri-, and tetra-antennary glycan structures with a mixture of α2,3-SA and α2,6-SAs, but compared to the glycome of the human respiratory tract, an absence of repeating N-acetylglactosamine units was observed (Table 1).

Ferrets are widely used as an animal model for influenza virus infections because they have a similar distribution of α2,6-SA and α2,3-SAs throughout the respiratory tract as humans (Leigh et al., 1995; Kirkeby et al., 2009). In the ferret trachea and bronchus, α2,6-SA was abundantly expressed by ciliated cells and submucosal glands (Fig 4). The presence of α2,3-SA was detected in the lamina propria and submucosal areas, but not on ciliated cells. Alveoli expressed both α2,3-SA and α2,6-SA although staining for the latter was more abundant (Jayaraman et al., 2012).

Structural determinants of receptor specificity

In 1981, the first crystal structure of influenza A virus HA was described by Wilson et al. (1981). These and other structural studies showed that HA is expressed as a trimer, with a site for SA binding in the membrane-distal tip of each of the monomers. Each site comprises a pocket of conserved amino acids that is edged by the 190-helix, the 130-loop, and the 220-loop (Fig 5A). A set of conserved residues, 98-Tyr, 153-Trp, 183-His, and 195-Tyr (numbering throughout the manuscript is based on H3 HA), forms the base of the RBS (Skehel & Wiley, 2000; Ha et al., 2001). Although sequence differences in the 190-helix, the 130-loop, and the 220-loop exist between viruses of different subtypes and from different hosts (Fig 5B), the conformation of each of these elements is similar. Structural studies have demonstrated that the RBS of human and pig influenza viruses that bind α2,6-SA is “wider” than the RBS of avian influenza viruses (Ha et al., 2001; Stevens et al., 2004).

To study interactions of HA with receptors, HAs are crystallized and soaked in human or avian receptor analogs. Commonly, linear sialylated pentasaccharides LSTc (α2,6-SA) and LSTa (α2,3-SA) are used as human and avian receptor analogs, respectively (Fig 6). Among the major difficulties of structural studies are the flexible...
Table 1. Glycan structures of the human respiratory tract that are represented on glycan arrays (adapted from Walther et al, 2013)

<table>
<thead>
<tr>
<th>Glycan arrays</th>
<th>Array A</th>
<th>Array B</th>
<th>Array C</th>
<th>Array D</th>
<th>Array E</th>
<th>Human</th>
<th>Ck</th>
<th>Pig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glycan structures</td>
<td>α2,6 α2,3 Both</td>
<td>α2,6 α2,3 Both</td>
<td>α2,6 α2,3 Both</td>
<td>α2,6 α2,3 Both</td>
<td>α2,6 α2,3 Both</td>
<td>Lung</td>
<td>Bronchus</td>
<td>Tonsil and NP</td>
</tr>
</tbody>
</table>
| Glycan structures detected in the pig lung (Chan et al, 2013b), and chicken (Ck) red blood cells (Aich et al, 2011), that were both present in the human respiratory tract and represented by one or multiple arrays were included.
nature of both the HA protein and the glycan structure, which is largely undetectable in crystallography, and the lack of knowledge of the SA receptors that are relevant for infection in vivo (see above). LSTA in complex with the HA of avian H1, H2, H3, and H5 viruses is usually bound in a trans conformation of the α2,3 linkage. In contrast, LSTc is bound in a cis conformation of the α2,6 linkage by HA of H1, H2, and H3 human influenza viruses (Gamblin & Skehel, 2010; Fig 6A).

HA of pandemic influenza viruses

The H2N2 and H3N2 pandemics were caused by viruses containing HAs of avian origin (Scholtissek et al, 1978; Kawaoka et al, 1989). Two mutations, Gln-226-Leu and Gly-228-Ser in the RBS, were sufficient for avian H2 and H3 viruses to switch from a 2,3-SA to a 2,6-SA specificity (Connor et al, 1994; Matrosovich et al, 2000; Fig 5B). The Leu at position 226 caused a shift of the 220-loop, resulting in a widening of the RBS, favorable binding to the human receptor, and abolished binding to the avian receptor (Liu et al, 2009; Xu et al, 2010b). This shift and the widening of the RBS were maximal if both 226-Leu and 228-Ser were present. In contrast to human H2 and H3, the HAs of human H1N1 viruses retained 226-Gln and 228-Gly, but bound α2,6-SA nevertheless (Rogers & D’Souza, 1989). The α2,6-SA preference of human H1N1 virus was found to be determined primarily by 190-Asp and 225-Asp (Matrosovich et al, 2000; Glaser et al, 2005). Structural differences in the 220-loop resulted in a lower location of 226-Gln, enabling H1 HAs to bind the human receptor analog despite the presence of 226-Gln (Gamblin et al, 2004). Upon binding to the human receptor, interactions between the receptor and 222-Lys and 225-Asp in the 220-loop are formed, and a hydrogen bond with 190-Asp (Gamblin et al, 2004; Zhang et al, 2013c).

HA of airborne H5N1 influenza viruses

Imai et al (2012) reported airborne transmission of a pH1N1 reassortant virus that contained the H5 HA of A/Vietnam/1203/04 (VN1203) with amino acid substitutions Asn-224-Lys and Gln-226-Leu, known to result in a switch in receptor specificity, Asn-158-Asp that resulted in the removal of a glycosylation site, and Thr-318-Ile in the stalk region which stabilized the trimer structure (Fig 1). The effect of these substitutions on receptor binding and HA structure was investigated in the context of autologous VN1203 and the closely related A/Vietnam/1194/04 (VN1194) HA (Lu et al, 2013; Xiong et al, 2013a). Xiong et al reported a small increase in binding to α2,6-SA for VN1194 with the mutations described by Imai et al (VN1194mut), similar to what was described in the original work (Imai et al, 2012). Both VN1194mut and VN1203mut displayed a substantially decreased affinity for α2,3-SA. Overall, the affinity for human and avian receptors of the mutant H5 HA was 5- to 10-fold lower than for human H3 HA (Xiong et al, 2013a). VN1194mut HA acquired the ability to bind α2,6-SA in the same cis conformation like human H1, H2, and H3 HAs. In particular, the Gln-226-Leu substitution facilitated binding to α2,6-SA, while restricting binding to α2,3-SA. For both the VN1203mut and VN1194mut HA, the RBS between the 130- and 220-loops had increased in size by approximately 1–1.5 Å. It was further suggested that the glycosylation at residue 158 might sterically block HA binding to cell surface SAs (Lu et al, 2013; Xiong et al, 2013a).
Herst et al (2012) described fully avian airborne transmissible H5N1 viruses based on A/Indonesia/5/05 (INDO5), of which the HA consistently contained substitution His-110-Tyr with a so far unknown effect, Thr-160-Ala that removes the same glycosylation site as Asn-158-Asp, and Gln-226-Leu and Gly-228-Ser that were shown to affect receptor specificity (Fig 1). The effect of these substitutions on receptor binding and HA structure was also investigated (INDO5mut). INDO5mut had a approximately ninefold reduced affinity for α2,6-SA compared to wild-type HA as measured by surface plasmon resonance. Overall, the affinity for avian and human receptors of the INDO5mut HA was still relatively low. INDO5 bound LSTc in trans conformation, which is different from human HAs. INDO5mut bound this human receptor analog in cis conformation, in an orientation similar as for pandemic viruses of other virus subtypes. The 226-Leu created a favorable environment for the non-polar portion of the human receptor analog and makes a tighter interaction with the receptor analog and LSTc in a similar conformation, which is different from human HAs. INDO5mut had a approximately ninefold increased affinity for α2,6-SA as compared to wild-type HA, as measured by surface plasmon resonance. Overall, the affinity for avian and human receptors of the INDO5mut HA was still relatively low. INDO5 bound LSTc in trans conformation, which is different from human HAs. INDO5mut bound this human receptor analog in cis conformation, in an orientation similar as for pandemic viruses of other virus subtypes. The 226-Leu created a favorable environment for the non-polar portion of the human receptor analog and makes a tighter interaction between the RBS and receptor analog. Substitutions Gln-226-Leu and Gly-228-Ser resulted in a RBS that was approximately 1 Å wider between the 130- and 220-loops compared to that of wild-type HA, as reported for human HAs (Zhang et al, 2013b).

Collectively, these studies on receptor binding and structure of HA of airborne transmissible H5N1 viruses thus showed a binding preference for human over avian receptors (although with overall low affinity), widening of the RBS as a consequence of Gln-226-Leu, and binding to the human receptor analog LSTc in a similar conformation as shown for human pandemic HAs.

Both airborne H5 HAs and a third one described in Zhang et al lost the same glycosylation site near the RBS that was associated with increased binding affinity, potentially as a result of reduced steric hindrance of interactions between HA and the receptor (Herfst et al, 2012; Imai et al, 2012; Zhang et al, 2013d). H5 and H7 influenza viruses from domestic birds generally have more glycosylation sites on the globular head than wild-bird influenza viruses, which has been associated with reduced binding affinity (Matrosovich et al, 1999). The glycosylation state of HA was previously shown to affect both receptor binding efficiency and immunogenicity (Skehel et al, 1984; Schulze, 1997). During the evolution of influenza viruses, glycosylation sites are removed or introduced, resulting in evasion of host immune responses. In addition, changes in glycosylation may cause functional differences, for example, by affecting the stability of the RBS. Removal of N-glycans by means of glycosidase F treatment was shown to enhance hemadsorbing activity of HA (Ohuchi et al, 1997). The importance of proper glycosylation is further evident from studies on recombinant soluble HA trimers. Different expression systems can result in differences in the glycosylation of the HA, and the presence of, for example, sialylated N-glycans on HA can narrow its receptor binding range, which can be increased upon the removal of SA by the addition of exogenous NA (de Vries et al, 2010). Thus, differences in glycosylation of the HA can have major implications for receptor specificity and affinity.

**HA of H7N9 influenza viruses**

During the outbreak in China in 2013, several residues that are associated with changes in the receptor specificity were observed in HA of the circulating H7N9 viruses. A/Shanghai/1/13 contained 186-Val, which was shown previously to enhance binding to α2,6-SA of pig H5N1 viruses (Nidom et al, 2010). A/Anhui/1/13 and A/Shanghai/2/13 possessed 226-Leu, which enhances binding to α2,6-SA in H7 avian influenza viruses, and 186-Val, which was also reported to influence receptor binding of H7 (Gambaryan et al, 2012; Srivinivas et al, 2013; Yang et al, 2013). However, none of the H7N9 viruses contained 228-Ser. The H7 genetic lineage from which the H7N9 HA was derived naturally lacks a glycosylation site at the tip of HA, the one that was lost in the airborne H5 viruses.
Multiple research groups have tested the receptor specificity of H7N9 viruses, most of which reported dual receptor specificity (Ramos et al., 2013; van Riel et al., 2013; Shi et al., 2013; Watanabe et al., 2013; Xiong et al., 2013b; Zhou et al., 2013). Overall, A/Anhui/1/2013 was found to bind better to α2,6-SA than A/Shanghai/1/2013. In glycan arrays, A/Anhui/1/2013 bound to bi-antennary structures as well as structures containing long N-acetyllactosamine repeats (Belser et al., 2013a; Watanabe et al., 2013). H7N9 influenza viruses in general had higher affinity for α2,6-SA compared to other H7 influenza viruses. However, in most studies, the H7N9 viruses maintained preference for avian over human receptors, which is in contrast to the pandemic and airborne transmissible H5 viruses.

The H7N9 HA structure of A/Anhui/1/2013 in complex with human and avian receptors was solved and compared with those of avian H7 HAs. Both LSTc and LSTa were bound to HA in a cis conformation (Shi et al., 2013; Xiong et al., 2013b). LSTc was shown to exit the RBS of H7N9 HA in a different direction from that seen with all pandemic virus HAs and the airborne transmissible H5 HA. The H7 HAs have a subtype-specific insertion in the 150-loop, causing this loop to protrude closer to the RBS. The function of this 150-loop is not yet known (Russell et al., 2006). The Gln-226-Leu substitution in H7N9 HA provided a non-polar-binding site accommodating the human receptor, and substitution Gly-186-Val increased the hydrophobicity of the binding site to further favor interactions with this receptor. The space between the 220-loop and the 130-loop was widened in H7N9 HA, but only marginally so, compared to pandemic virus HA and airborne H5 HA (Xiong et al., 2013b).

Virus attachment, replication, pathogenesis, and transmission

Much of the available data on tropism of influenza viruses in the human respiratory tract has been derived from few autopsy studies of fatal cases. Although highly informative to identify the cell types that are infected in vivo and to understand pathogenesis, these data are generally restricted to late phases of the infection. In attempts to compensate for the lack of data in humans, the effect of receptor specificity on tropism, pathogenesis, and transmission has also been studied in a number of animal models like ferrets, pigs, guinea pigs, and mice, and in primary cell lines of human and ferret airway epithelium, and explants of ferret, pig, or human origin (reviewed in Chan et al., 2013a). However, it should be noted that—despite the recent increase in detailed knowledge on fine specificities of HA interacting with different glycans in vitro (see above)—insight on the relevance of interactions between HA and various glycans in vivo is still largely limited to the coarse discrimination between α2,6-SA and α2,3-SA alone.

Virus attachment

Virus histochemistry provides a method to investigate and compare attachment patterns of influenza viruses to cells of various tissues from different hosts, without a requirement of detailed prior knowledge on the specific receptors. For virus histochemistry, host tissue sections are incubated with inactivated, FITC-labeled influenza viruses, much like immunohistochemistry employs FITC-labeled antibodies. The FITC signals can be enhanced, resulting in a bright red precipitates where the viruses (or antibodies in immunohistochemistry) attach. Attachment patterns of human and avian influenza viruses in the (human) respiratory tract have been shown to be a good indicator for virus cell and tissue tropism. Attachment patterns of a number of avian influenza viruses and human or mutant influenza viruses that are transmissible between humans or ferrets are summarized in Table 2.

Broadly speaking, avian influenza viruses with a preference for α2,3-SA were shown to attach to cells of the lower respiratory tract, to type II pneumocytes and non-ciliated cells. In contrast, viruses that can transmit between mammals and have a preference for α2,6-SA were shown to attach to cells of the upper respiratory tract, primarily to ciliated cells. These patterns of virus attachment were in agreement with the receptor distribution in human respiratory tissue as determined upon staining with various lectins (see above).

The dual receptor specificity of the 2013 H7N9 viruses was also reflected by their attachment patterns; H7N9 viruses were shown to attach to both cells in the upper and the lower respiratory tract of humans, and to both type I and II pneumocytes (van Riel et al., 2013; Table 2). Some discrepancies, however, were noted in attachment studies using recombinant HA proteins. Dortmans et al. (2013) primarily observed attachment in the submucosal glands of the human upper respiratory tract, and Tharakaraman et al. (2013a)...

<table>
<thead>
<tr>
<th>Virus</th>
<th>SA</th>
<th>Nasal turbimates</th>
<th>Trachea</th>
<th>Bronchus</th>
<th>Bronchiole</th>
<th>Alveoli</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Score</td>
<td>Cell type</td>
<td>Score</td>
<td>Cell type</td>
<td>Score</td>
<td>Cell type</td>
</tr>
<tr>
<td>H6N2-Hu</td>
<td>α2,6</td>
<td>++</td>
<td>C</td>
<td>++</td>
<td>C/G</td>
<td>++</td>
</tr>
<tr>
<td>H6N1-Hu</td>
<td>α2,6</td>
<td>++</td>
<td>C/G</td>
<td>++</td>
<td>C/G</td>
<td>++</td>
</tr>
<tr>
<td>pH1N1-Hu</td>
<td>α2,6</td>
<td>+</td>
<td>C</td>
<td>++</td>
<td>C/G</td>
<td>++</td>
</tr>
<tr>
<td>H7N7-Av</td>
<td>α2,3</td>
<td>+/−</td>
<td>nd</td>
<td>+/−</td>
<td>nd</td>
<td>+/−</td>
</tr>
<tr>
<td>H7N9-Av</td>
<td>α2,6/α2,3</td>
<td>++</td>
<td>C</td>
<td>++</td>
<td>C/G</td>
<td>++</td>
</tr>
<tr>
<td>H5N1-Av</td>
<td>α2,3</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>H5N1-Mu</td>
<td>α2,6</td>
<td>C/G</td>
<td>C/G</td>
<td>C/G</td>
<td>+</td>
<td>C</td>
</tr>
</tbody>
</table>

Attachment of human (Hu), avian (Av), and avian influenza with mutations (Mu) that change the receptor specificity (SA; van Riel et al., 2007, 2010, 2013; Chutinimitkul et al., 2010b). The mean abundance of cells to which virus attached was scored as follows: −, no attachment; +/−, attachment to rare or few cells; +, attachment to a moderate number of cells; ++, attachment to many cells. The predominant cell type to which the virus attached is indicated as follows: C, ciliated cells; N, non-ciliated cells; G, goblet cells; I, type I pneumocytes; II, type II pneumocytes; nd, not determined.
only observed limited attachment to the apical epithelial cells of the human trachea. Although influenza viruses that bind 2,3-SA or 2,6-SA thus display distinct attachment patterns to cells of the human respiratory tract, small—but potentially important—differences in attachment patterns of viruses with similar receptor specificity have also been observed. For example, H5N1 influenza viruses with the mutations Asn-182-Lys or Gln-226-Leu and Gly-228-Ser both displayed a preference for 2,6-SA, yet their attachment patterns were not identical; while both viruses attached to nasal turbinate tissue, only H5N1 with Asn-182-Lys attached to the trachea of ferrets (Chutinimitkul et al., 2010b). The 1918 H1N1 and 1957 H2N2 pandemic influenza viruses were also both shown to bind 2,6-SA, yet displayed differences in attachment patterns; whereas H2N2 virus attached to the glycoalyx, goblet cells, and submucosal glands, H1N1 virus showed predominant staining of goblet cells in the human trachea (Jayaraman et al., 2012). The correlation between virus attachment to cells throughout the respiratory tract and receptor preference beyond 2,6-linkage or 2,3-linkage alone is not fully understood and warrants further investigation.

**Virus replication in primary epithelial cell cultures**

Although influenza virus attachment studies have shown great value, not all cells to which influenza viruses can bind are subsequently also infected (Rimmelzwaan et al., 2007), and potentially not all infected cells produce and release infectious virus particles. Moreover, mucus present in the respiratory tract may hamper influenza virus binding and infection. Mucus from the human respiratory tract predominately contains 2,3-SA receptors (Couceiro et al., 1993; Lamblin & Roussel, 1993), although some 2,6-SA receptors were also detected (Breg et al., 1987). Incubation of different influenza viruses with a range of receptor specificities revealed that particularly viruses with 2,3 specificity were inhibited by human mucus (Couceiro et al., 1993).

Human and ferret differentiated airway epithelial cells are frequently used to study the effect of mammalian adaptation mutations on infection and replication of influenza viruses. Differentiated airway epithelial cells form a pseudo-stratified epithelium that closely resembles respiratory epithelium, including the presence of cilia and secretion of mucus. In human differentiated airway epithelial cells, avian influenza viruses infected ciliated cells and to a lesser extent non-ciliated cells, whereas human influenza viruses infected non-ciliated cells and to a lesser extent ciliated cells. This is in contrast with virus attachment studies using human tissues without culturing, which revealed binding of human influenza viruses primarily to ciliated cells. Apparently, ciliated cells primarily expressed 2,3-SA and non-ciliated cells expressed 2,6-SA upon culturing (Matrosovich et al., 2004; Thompson et al., 2006; Wan & Perez, 2007), although the presence of 2,6-SA on ciliated cells has been reported in some studies (Ibricivric et al., 2006; Schrauwen et al., 2013).

In contrast to human differentiated airway epithelial cells, ciliated differentiated airway epithelial cells from ferrets were shown to express 2,6-SA, while non-ciliated cells expressed 2,3-SA. Both human and avian influenza viruses replicated efficiently in differentiated airway epithelial cells from ferrets, but avian viruses had a lower initial infection rate. Using electron microscopy, it was shown that human influenza viruses were predominantly released from the ciliated cells, and release of avian viruses was only sporadically detected and only from non-ciliated cells, in agreement with the distribution of 2,3-SA and 2,6-SA (Zeng et al., 2013).

Although differentiated airway epithelial cell cultures are quite valuable for various studies, it should thus be noted that cultures of human and ferret origin may behave differently and that receptor expression upon culture may not be representative for the in vivo human respiratory tract tissues.

**Virus replication in animal models**

Unlike seasonal human influenza viruses, H5N1 viruses generally do not attach to the upper respiratory tract of ferrets or humans (van Riel et al., 2006, 2010; Chutinimitkul et al., 2010b). However, in contrast to the results from attachment studies, H5N1 virus was reported to replicate both in human upper respiratory tract tissue ex vivo (Nicholls et al., 2007b), and in the ferret upper respiratory tract in vivo (Maines et al., 2005, 2006). It is possible that the high virus titers found in the ferret upper respiratory tract are explained in part by virus replication in the olfactory epithelium, which is related to the nervous system and lines a large part of the nasal cavity (Schrauwen et al., 2012). It has also been reported that H5N1 influenza viruses with mutations that change receptor specificity from 2,3-SA to 2,6-SA did not replicate more efficiently in the upper respiratory tract of ferrets than their wild-type counterparts (Wang et al., 2010; Maines et al., 2011; Herfst et al., 2012). This may be explained by a requirement for additional adaptive mutations in HA to accommodate the mutations responsible for 2,6-SA binding, but more research is needed to identify such “fitness-enhancing” mutations.

The dual receptor specificity of the H7N9 virus (noted above) was reflected in the attachment patterns to the upper and lower respiratory tract of humans (van Riel et al., 2013). In agreement with receptor specificity and patterns of virus attachment, the H7N9 virus was shown to replicate in both trachea and lung explants (with higher titers in the lung explants) and in human lung cell cultures (Belser et al., 2013a; Knepper et al., 2013; Zhou et al., 2013). In ferrets, the H7N9 virus replicated both in the upper and lower respiratory tract (Kreijtz et al., 2013; Zhang et al., 2013a; Zhu et al., 2013).

**Receptor specificity and pathogenesis**

The site of influenza virus replication is potentially one of the most critical determinants of pathogenesis. H5N1 viruses were shown to replicate efficiently in the lower respiratory tract in type II pneumocytes (Gu et al., 2007), probably related to the specificity for 2,3-SA, and contributing to pathogenicity. For pH1N1 virus, amino acid substitution Asp-225-Gly that resulted in increased binding to 2,6-SA was also linked to enhanced disease (Chutinimitkul et al., 2010a; Kilander et al., 2010; Watanabe et al., 2011). Potentially, the dual receptor specificity of the H7N9 virus could be responsible for the severe disease and high mortality observed in humans. Indeed, in various mammalian model systems, the H7N9 virus was shown to be more pathogenic than typical human influenza viruses, which was probably related to efficient virus replication throughout the respiratory tract (Belser et al., 2013a; Kreijtz et al., 2013; Watanabe et al., 2013). However, it should be noted that not all avian influenza viruses (with specificity for 2,3-SA) are pathogenic in humans; human
volunteers that were infected with various avian influenza viruses only displayed mild symptoms (Beare & Webster, 1991), thus emphasizing that other viral (e.g., NS1, polymerase, NA) and host factors (e.g., adaptive and innate immune responses) also critically determine pathogenesis.

**Virus transmission**

It has been suggested that (i) inefficient virus attachment to (upper) respiratory tissues, (ii) virus replication at only low levels in these tissues, and (iii) poor virus release and aerosolization of virus particles may provide an explanation why influenza viruses with 2,3-SA preference are not transmitted via the airborne route between mammals (Sorrell et al., 2011). In addition, it was proposed that for a virus that requires 2,3-SA receptors, a higher dose is needed to infect a new mammalian host, which makes it less likely to transmit (Roberts et al., 2011). Tumpey et al. reported that 1918 H1N1 influenza viruses with 2,3-SA receptor specificity were not transmitted via the airborne route between ferrets, while viruses with dual receptor specificity were transmitted, albeit less efficiently as compared to viruses with 2,6-SA receptor specificity. Based on these results, it was postulated that a loss of binding to 2,3-SA is important for airborne transmission (Tumpey et al., 2007). While a loss of binding to 2,3-SA can result in more efficient airborne transmission, it should be noted that early H2N2 and H3N2 pandemic viruses (that spread efficiently between humans) still had dual receptor specificity (Matrosovich et al., 2000). Both the airborne transmissible H5N1 viruses and the 2013 H7N9 virus displayed dual receptor specificity and were transmitted between ferrets, but for these viruses transmission was less efficient compared to pandemic and seasonal human influenza viruses (Chen et al., 2012; Herfst et al., 2012; Imai et al., 2012; Belser et al., 2013a; Richd et al., 2013; Watanabe et al., 2013; Xu et al., 2013; Zhang et al., 2013a,d; Zhu et al., 2013). The current common belief is that for airborne transmission between humans or ferrets, influenza viruses require human over avian receptor binding preference; viruses with exclusive 2,3-SA binding are generally not transmitted, viruses with dual receptor specificity are often transmitted inefficiently, while influenza viruses with full-blown airborne transmission between ferrets or humans generally had a strong 2,6-SA binding preference. It is important to identify whether a fine specificity in receptor binding beyond 2,3-SA/2,6-SA is required for airborne influenza virus transmission between mammals.

**Determinants of airborne transmission beyond receptor specificity**

It is clear from airborne transmission studies on pandemic influenza and avian H5N1, H9N2, and H7N9 influenza viruses that a change of receptor specificity is a necessity, but by itself is not enough to result in airborne transmission between mammals. But what are other adaptations of the HA or other viral proteins required for airborne transmission between mammals?

**HA stability**

Upon virus attachment to SA receptors on the cell surface and internalization into endosomes, a low-pH-triggered conformational change in HA mediates fusion of the viral and endosomal membranes to release the virus genome in the cytoplasm. The switch of influenza virus HA from a metastable non-fusogenic to a stable fusogenic conformation can also be triggered at neutral pH when the HA is exposed to increasing temperature. This conformational change in HA is biochemically indistinguishable from the change triggered by low pH (Haywood & Boyer, 1986; Carr et al., 1997).

Generally, the threshold pH of fusion for HAs of human influenza viruses is lower than that of avian influenza isolates (Galloway et al., 2013). Possibly, the optimal fusion pH of a virus depends on the original host. For avian influenza viruses that replicate in the intestinal tract of birds and are transmitted via surface waters, fusion at a high pH might be more optimal. In contrast, efficient transmission through the air to and from the relatively acidic human nasal cavity as site of replication could require a difference in pH threshold (England et al., 1999; Washington et al., 2000).

Lowering the optimal fusion pH for viruses with an avian H5 HA resulted in enhanced replication in the ferret upper respiratory tract and for some viruses in more efficient contact transmission (Shelton et al., 2013; Zarak et al., 2013). The HA of the airborne transmissible H5N1 virus described by Imai et al. (2012) contained the substitution Thr-318-Ile, which is located proximal to the fusion peptide. The wild-type H5N1 virus (VN1203) had a pH 5.7 fusion threshold and introduction of mutations that change receptor specificity increased the fusion threshold to pH 5.9. However, subsequent introduction of Thr-318-Ile reduced the threshold of fusion to pH 5.5. In agreement with this increased pH stability, Thr-318-Ile also increased the thermostability of this HA (Imai et al., 2012). Structural studies revealed that the Thr-318-Ile mutation stabilized the fusion peptide within the HA monomer (Xiong et al., 2013a). Notably, substitution His-110-Tyr in a recombinant protein based on the airborne transmissible virus described by Herfst et al. also appeared to increase the stability of the HA at high temperatures (de Vries et al., 2014). His-110-Tyr is located at the trimer interface and forms a hydrogen bond with 413-Asn of the adjacent monomer, thereby stabilizing the trimeric HA protein (Zhang et al., 2013b). Whether this substitution has the same effect on pH stability is still unknown.

The 2013 H7N9 viruses that showed limited airborne transmission between ferrets were shown to have a relatively unstable HA, with conformational changes induced already at relatively low temperatures and a threshold pH for membrane fusion at pH 5.6–5.8 (Richard et al., 2013). Given the observed differences in threshold pH of fusion and thermostability between HAs of human and avian influenza viruses, and the acquired stability mutations in HA of airborne H5N1 viruses, HA stability as an adaptation marker of human influenza viruses clearly warrants further study. It is important to note that the observed changes in pH stability and thermostability may merely be surrogate markers for other stability phenotypes, for example stability in aerosols, in mucus, or upon exposure to air.

**HA-NA balance**

As a consequence of the switch in HA receptor binding preference when avian influenza viruses adapt to mammals, subsequent adaptation of NA may be required to maintain an optimal balance between attachment to SA and cleavage of SA. Low NA activity can result in inefficient SA cleavage and subsequent aggregation of virus particles, which can result in low infectious virus titers despite high genome copy numbers (Palese et al., 1974; Liu et al., 2014).
Polymerase proteins

In addition to the HA and NA proteins, the influenza virus polymerase proteins were shown to require adaptation for efficient replication to occur upon transmission from animals to humans (reviewed in Manz et al., 2013). Herfst et al introduced the well-known adaptive mutation Gly-627-Lys in the PB2 gene of an H5N1 virus that subsequently accumulated numerous additional substitutions upon ferret passage in polymerase proteins and nucleoprotein to yield airborne transmissible virus. The Lys at position 627 of PB2 was previously described to contribute to successful replication in mammalian cells at lower temperature (Almond, 1977; Subbarao et al., 1993; Hatta et al., 2001; Massin et al., 2001; Yamada et al., 2010), and airborne transmission between mammals (Steel et al., 2009; Van Hoeven et al., 2009). Efficient replication at lower temperature is thought to be important for adaptation to humans, since the temperature in the human upper respiratory tract is approximately 33°C, which is much lower compared to the temperature in the avian intestinal tract (~41°C), the site of replication of avian viruses.

Amino acid substitutions at positions 590/591 and at position 701 of PB2 were also shown to result in more efficient replication in mammalian cells (Steel et al., 2009; Yamada et al., 2010). Of the H7N9 virus genome sequences available from public databases, 65% contained 627-Lys, 9% contained 701-Asn, and 8% contained 627-Glu, in contrast to avian H7N9 viruses that retained 627-Glu, whereas human influenza virus NA s cleave both 2,3-SA and 2,6-SA. It is thus likely that the NA of avian viruses needs to co-adapt to the receptor specificity of the HA, for efficient virus attachment and entry, virus release, and potentially transmission (Baum & Paulson, 1991; Kobasa et al., 1999). In this light, it is important to note that several airborne transmissible H5 viruses contained NA derived from human influenza viruses (Chen et al., 2012; Imai et al., 2012; Zhang et al., 2013d), but that the fully avian H5N1 viruses described by Herfst et al did not require any changes in the NA gene for airborne transmission (Herfst et al., 2012).

Challenges and future outlook

Recent work on virus attachment, replication, pathogenesis, and transmission and follow-up work on structural determinants of HA receptor specificity and stability have shed important new light on influenza virus adaptation to mammalian hosts. Imai et al and Herfst et al described different amino acid substitutions in H5N1 virus HA that represented remarkably similar phenotypic traits leading to airborne virus transmission. Some of these substitutions were also noted in 2013 H7N9 viruses with limited airborne transmissibility, in an airborne H9N2 virus, and in the pandemic viruses of the last century. Some of the substitutions of the laboratory-derived airborne H5N1 viruses and combinations of such substitutions were already detected in naturally circulating H5N1 viruses (Russell et al., 2012). However, it should be noted that these mutations do not necessarily result in the same phenotype in all H5N1 strains or in strains of other influenza virus subtypes. Moreover, it is possible that there are evolutionary constraints to the accumulation of genetic changes required to yield an airborne transmission phenotype, making the emergence of such viruses possibly an unlikely event (Russell et al., 2012). Further identification of the genetic and phenotypic changes required for host adaptation and transmission of influenza viruses remains an important research topic to facilitate risk assessment for the emergence of zoonotic and pandemic influenza.

In currently circulating strains of clade 2.2.1 H5N1 containing a deletion in the 130-loop and lacking a glycosylation site at position 160, the substitutions 224-Asn and 226-Leu resulted in a switch of receptor specificity (Tharakaraman et al., 2013b). We know that there are many more ways by which H5N1 viruses can adapt to attach to 2,6-SA. Some natural human H5N1 virus isolates bind to both 2,3-SA and 2,6-SA due to substitutions in and around the RBS (Yamada et al., 2006; Crusat et al., 2013). In addition, mutational analyses of H5 HA based on analogies with other influenza virus subtypes identified several substitutions that caused a switch in receptor specificity and virus attachment to upper respiratory tract tissues of humans. Although none of these human virus isolates and mutant viruses were capable of airborne transmission between mammals (Maines et al., 2006; Stevens et al., 2006b, 2008; Chutinimitkul et al., 2010b; Paulson & de Vries, 2013), it is clear that there are multiple ways by which circulating H5N1 strains can change receptor specificity and adapt to replication in mammals.

Yang et al (2013) have shown that introduction of substitution 228-Ser in H7N9 HA that already had 226-Leu resulted in overall increased binding to 2,3-SA and 2,6-SA, but not in a switch in receptor preference. Based on structural studies, it was anticipated that Gly-228-Ser would optimize interactions with both 2,3-SA and 2,6-SA. Gly-228-Ser resulted in increased binding of the H7N9 HA to the apical surface of tracheal and alveolar tissues of humans (Tharakaraman et al., 2013a). At present, it is unclear whether other amino acid substitutions would further fine-tune the receptor binding preference of the H7N9 viruses and could potentially increase the airborne transmissibility under laboratory or natural conditions.

Although our knowledge of receptor specificity and its implications for transmission and pathogenesis has improved since the initial characterization of the influenza virus RBS, there are still many unknowns. The lack of knowledge of which cells are the primary target of airborne transmissible viruses and the type and distribution of receptors they express hampers the interpretation of receptor binding data and structural studies. Major current interest is focused on receptor binding assays such as glycan arrays, solid-phase binding assays, and agglutination assays using modified erythrocytes. But without knowledge on which receptors are relevant for influenza virus attachment, replication, and transmission in vivo, it is impossible to assess which binding assays best represent influenza virus attachment in the respiratory tract. As a consequence, studies should focus more on identification of the relevant functional glycans in humans and animal models. The
possibility that a plethora of receptors in the human respiratory tract can be utilized for the attachment of influenza viruses and are relevant for replication and transmission should be considered. Based on studies of the glycome of human and animal hosts, it is clear that a large number of glycan structures are not represented on glycan arrays (Walther et al., 2013). A potential way forward would be the use of shotgun glycan arrays that employ receptors of relevant tissues and cells of the human respiratory tract rather than random synthetic glycans (Yu et al., 2012). At present, such studies are limited since they generally employ all glycans expressed in the respiratory tract rather than just those receptors expressed on relevant cell types at the apical surface of the respiratory tract that are accessible upon influenza virus infection. Future work should clearly focus on improvement of the methods available in glycobiology. Similarly, X-ray crystal structures of human and avian receptor analogs in complex with HAs of human and influenza viruses have revealed crucial structural determinants for binding to s2,3-SA and s2,6-SA receptor analogs. But potentially these structural studies could also be complemented with influenza HAs in complex with a variety of sialylated glycan structures, to increase understanding of the structural determinants of HA receptor binding.

There is a clear need for phenotypic assays to predict which animal influenza viruses can infect humans and transmit between humans. One major unknown is to what extent the transmissibility of influenza viruses in ferret and guinea pig transmission models can be extrapolated to humans. Although thus far there has been good correspondence between the transmissibility of avian and human influenza viruses between humans and the ferret transmission model, it is not certain that all viruses that transmit between ferrets will also transmit between humans. It is also not clear how the ferret and guinea pig models compare with respect to airborne transmission of animal influenza viruses. Better understanding about the behavior of airborne transmissible viruses in vitro may allow replacement of some animal models with in vitro methods and contribute to reduction, replacement, and refinement in animal experiments (Russell & Burch, 1959; Belser et al., 2013b). In addition, such phenotypical assays may complement ongoing surveillance studies to identify biological traits of circulating influenza viruses that could trigger immediate action or attention. So far, when data obtained with attachment studies, infection studies using primary epithelial cell lines explants, and animal studies were compared, the correspondence was not always obvious, and therefore, more work in this field is also needed. While it may never be feasible to perfectly predict the adaptation of animal influenza viruses to humans, increased knowledge from animal studies, in vitro assays, and structural analyses collectively will almost certainly improve such predictions.

Acknowledgements
The authors are supported by EU FP7 programs ANTICONE and EMPERIE, NIAID/NIH contract HHSN26200700010C, and an Intra European Marie Curie Fellowship (PIEF-GA-2009-237505). We thank C.H. Hokke, B. Mänz, D.F. Burke, J.A. Schrauwen, D. van Riel, E. de Vries, C.A. de Haan, and S. Herfst for critically reading the manuscript and their helpful comments.

Author contributions
MG and RAMF wrote the manuscript.

Conflict of interest
The authors declare that they have no conflict of interest.

References
Carr CM, Chaudhry C, Kim PS (1997) Influenza hemagglutinin is spring-loaded by a metastable native conformation. Proc Natl Acad Sci USA 94: 14306–14313
Chan RW, Chan MC, Nicholls JM, Malik Peiris JS (2013a) Use of ex vivo and in vitro cultures of the human respiratory tract to study the tropism and host responses of highly pathogenic avian influenza A (H5N1) and other influenza viruses. Virus Res 178: 133–145

affects receptor binding.
characterisation of viral genome.
cell N-linked glycoprotein.
Sci USA
Fouchier RA, Schneeberger PM, Rozenaal FW, Broekman JM, Kemink SA, Munster V, Kuiken T, Rimmelzwaan GF, Schutten M, Van Doornum GJ, Koch G, Bosman A, Koopmans M, Osterhaus AD (2004) Avian influenza A virus (H7N7) associated with human conjunctivitis and a fatal case of acute respiratory distress syndrome. Proc Natl Acad Sci USA 101: 1356 – 1361
Geisler C, Jarvis DL (2011) Effective glycoanalysis with Maackia amurensis lectins requires a clear understanding of their binding specificities. Glycobiology 21: 988 – 993
Guan Y, Shortridge KF, Krauss S, Webster RG (1999) Molecular characterization of H9N2 influenza viruses: were they the donors of the “internal” genes of H5N1 viruses in Hong Kong? Proc Natl Acad Sci USA 96: 9963 – 9967


Role of HA in influenza A virus transmission
Miranda de Graaf & Ron A M Fouchier


