

INNOVATION

Glycan microarray technologies: tools to survey host specificity of influenza viruses

James Stevens, Ola Blixt, James C. Paulson and Ian A. Wilson

Abstract | New technologies are urgently required for rapid surveillance of the current H5N1 avian influenza A outbreaks to gauge the potential for adaptation of the virus to the human population, a crucial step in the emergence of pandemic influenza virus strains. Owing to the species-specific nature of the interaction between the virus and host glycans, attention has recently focused on novel glycan array technologies that can rapidly assess virus receptor specificity and the potential emergence of human-adapted H5N1 viruses.

Influenza, an acute viral disease of the respiratory tract, affects millions of people each year. As members of the *Orthomyxoviridae* family, influenza viruses have an envelope and a negative sense, single-stranded and segmented RNA genome. Of the three groups of influenza virus (A, B and C), type A accounts for all known major epidemics and pandemics. Influenza A virus subtypes are named according to their surface antigens: haemagglutinin (HA) and neuraminidase (NA). Currently, all sixteen HA (H1 to H16) and nine NA (N1 to N9) serotypes of influenza A virus circulate in the avian population and, consequently, birds are believed to be an important reservoir for influenza A viruses.

In the past century, influenza A viruses with only three HA (H1, H2 and H3) and two NA (N1 and N2) serotypes have adapted sufficiently to humans to produce pandemic strains: H1N1 in 1918, H2N2 in 1957 and H3N2 in 1968 (REFS 1,2). The antigenic shifts that occurred in 1957 and 1968 were a result of avian-human reassortment: in 1957, three of the eight avian gene segments, and in 1968, two avian gene segments were reassorted into existing, human-adapted viruses. This reassortment event is believed by some to occur in animals such as pigs, which can serve as a mixing vessel, as these animals are

susceptible to both avian and human influenza virus³. The genetic origin of the 1918 pandemic H1N1 virus, which killed about 50 million people worldwide⁴ is not known, and there is an ongoing debate as to whether or not the virus 'jumped' directly to the human population or adapted through reassortment^{1,5-7}.

Since late 2003, circulating H5N1 avian influenza virus strains have reached epizootic levels in both domestic and wild bird populations across Asia. Recently, the H5N1 avian influenza virus has spread into Europe, the Middle East and the African continent. An unprecedented level of global concern surrounds the circulating highly pathogenic H5N1 avian influenza A virus, with fears that the next human influenza pandemic could arise from these strains. So far, the spread of H5N1 avian influenza virus to the human population has been limited; as of 20 June 2006 there have been around 228 well-documented human cases, but these cases have been associated with a high mortality rate of 130 deaths⁸, although the rate might be inflated by the lack of reporting of asymptomatic cases.

The influenza A virus surface is coated with three antigenic proteins: HA, which is responsible for virus binding to receptors on the host cell, internalization and

subsequent membrane fusion events in the endosomal pathway of the infected cell; NA, which is a sialidase involved in the removal of sialic acid from the infected cell surface to allow the escape of progenitor viral particles after budding; and the M2 ion channel protein, which is involved in the acidification of the inside of the viral particle during internalization through the endosomal pathway. Although all three coat proteins are targets for the host immune system, HA is by far the most abundant antigen on the viral surface and, therefore, harbours the primary epitopes for neutralizing antibodies.

Several factors are involved in host-range restriction of influenza viruses, but these will not be discussed here, as they have been reviewed recently². Here, we focus on receptor specificity of influenza virus HA. Influenza viruses use host glycans, which contain sialic acid (*N*-acetylneuraminic acid (NeuAc)) moieties as cell-surface receptors, which vary in structure from species to species. Human-adapted influenza viruses have HAs with a binding preference for cell receptors that contain α 2-6-linked sialic-acid moieties (NeuAc α 2-6Gal), which are found predominantly on epithelial cells of the human upper respiratory tract, whereas avian viruses prefer receptors with α 2-3-linked sialic acid moieties (NeuAc α 2-3Gal)⁹⁻¹¹, which are found on epithelial cells in the intestines and respiratory tract of birds (FIG. 1). This subtle difference in the recognition of glycan structure is a key determinant of the species barrier that prevents avian influenza viruses from easily infecting humans. Another barrier to infection of humans by avian viruses are mucins produced by specialized epithelial cells that lubricate and protect the otherwise exposed upper respiratory tract epithelium. These proteins are rich in α 2-3-linked sialosides (sialylated glycans), and function to trap avian influenza viruses and prevent their attachment to epithelial cells¹². The inefficient binding of avian viruses to the α 2-6-linked sialic-acid moieties on the surface of epithelial cells and their entrapment by mucins provide a barrier to virus infection of the human upper airway.

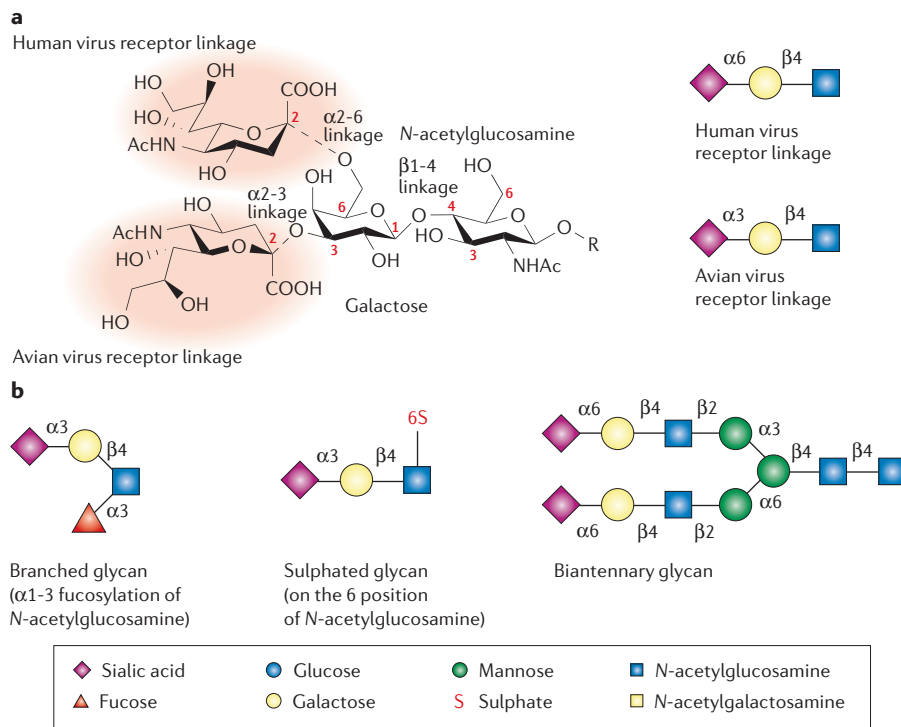


Figure 1 | Glycans and influenza virus specificity. **a** | The two possible positions of the sialic-acid linkage to a vicinal galactose (α 2-6 and α 2-3), which are crucial for recognition by the haemagglutinin (HA) proteins of avian and human viruses, are shown in the left panel. The sialic-acid moiety is highlighted by red shading. Cartoon representations of the typical sialic-acid linkage that is bound by avian and human influenza virus HAs are shown in the right panel. **b** | Cartoon representations of some of the different glycans used in the microarray experiments and discussed in the main text are shown. To simplify glycan descriptions, a symbol and text nomenclature has been developed by the [Nomenclature Committee of the Consortium for Functional Glycomics](#).

Recent studies have revealed that some human epithelial cells in the lower respiratory tract contain α 2-3-linked sialic acids and can be directly infected by avian influenza viruses^{10,11,13}. Although humans have been infected by the H5N1 avian influenza virus, most documented cases either lived or worked in close proximity to infected poultry and were, therefore, exposed to a sufficiently high viral load that would increase the chances of the virus escaping the protective mucin layer to infect cells in the lower respiratory tract. The restricted replication of H5N1 avian influenza viruses in the lower respiratory tract in humans, where the virus cannot be readily spread by sneezing and coughing¹¹, is believed to account in part for the inefficient human-to-human transmission of H5N1 viruses so far.

Several other species have been implicated in the origin of new human influenza viruses. The respiratory epithelial cells of pigs have receptors that express both α 2-3- and α 2-6-linked sialic acids and can be infected by human and avian viruses, which indicates that domesticated pigs could be a 'melting pot' for the emergence of new

reassortant viruses⁹ (FIG. 2). There is evidence that some terrestrial avian species, such as quail, also harbour both types of sialic-acid receptor in the trachea and intestine and, therefore, could also support the creation and spread of reassortments among avian and human influenza viruses^{14,15}.

The current H5N1 strains show increased transmissibility to other mammals, including domestic cats¹⁶, captive wild cats (tigers and leopards)¹⁷ and Indonesian pigs, in which the virus causes asymptomatic infection¹⁸, although the report on Indonesian pigs has yet to be confirmed. Importantly, no change has been reported in the epidemiology of H5N1 virus transmission and it still remains firmly an avian virus. However, with the constantly increasing host range for H5N1 viruses, many believe that intermediate hosts might ultimately provide the vehicle for H5N1 influenza viruses to adapt to the human population¹⁹. Furthermore, as described above, in the correct conditions, the H5N1 virus can infect humans directly⁸, and it could subsequently mutate to a strain that can effect human-to-human transmission.

In view of the rapid geographic spread of the H5N1 avian influenza virus in the bird population and the increasing numbers of confirmed human cases, new technologies are urgently required for the surveillance of current influenza outbreaks to gauge the potential for adaptation to the human population, a crucial step in the emergence of a pandemic strain. This Perspective article discusses existing methods that are used to study influenza virus receptor specificity and then introduces recently described glycan microarray technologies that are used to study the glycan-binding properties of several influenza virus HAs, including a current H5 HA.

Methods to assess HA receptor specificity

One important aspect of surveillance is to determine whether a particular avian influenza virus strain is adapting to human receptors. For many years, techniques to study the interactions of the influenza virus with host cell receptors have used whole viruses. This includes haemagglutination assays, which measure the ability of intact influenza virus and other viruses to agglutinate red blood cells (RBCs), and haemagglutination inhibition assays, which detect and quantify antibodies and soluble receptor analogues that inhibit RBC agglutination by intact viruses²⁰⁻²⁴. With the identification and isolation of specific α 2-3-sialidases and α 2-6-sialidases, and advances in technologies for genetic manipulation of the influenza genome, a haemadsorption assay based on desialylation of endogenous sialic acid from RBCs, followed by selective linkage-specific resialylation of RBCs, and subsequent analysis of their binding to viruses or cells transfected with influenza HA, has become routine in many laboratories^{21,24}.

Although such assays potentially reflect the natural binding of a virus to host cell receptors, the outcome is subject to many factors that are often difficult to control. For example, the quality of cell preparations can vary owing to the relative instability of RBCs, and the extensive variability in transfection efficiencies. The activity of the sialyltransferase enzymes might differ substantially between batches and suppliers and lead to varying degrees of enzymatic resialylation, which can skew results and mask any important specificity differences among viral strains. Another factor to consider when using whole-cell assays is the polyvalent cooperative effect of the virus interaction with multiple heterogeneous receptor ligands on the cell surface.

Therefore, these assays reflect only the overall 'global' affinity for α 2-3- or α 2-6-linked sialic-acid receptors and cannot distinguish among the various classes and modifications of sialylated glycans.

Since the early 1990s, new technologies have emerged to overcome the problems of cell-based assays. One of the most successful methods reported in recent years uses an assay that is based on competition for binding to solid-phase immobilized virus between a horseradish-peroxidase-conjugated blood sialyl glycoprotein, called fetuin, and unlabelled α 2-3 and α 2-6 sialyloligosaccharides²⁵. The α 2-3 and α 2-6 sialyloligosaccharides are coupled to polyacrylamide, providing increased valency to compensate for the low HA affinity (mM range) for receptor analogues²⁶⁻²⁸. Such assays have increased our understanding of the diversity of the receptor specificities of influenza viruses from both birds and mammals^{14,29-32}, including H5 influenza viruses^{33,34}. These assays have been used to evaluate details of influenza virus receptor specificity beyond the standard analysis for recognition of NeuAc α 2-3Gal and NeuAc α 2-6Gal linkages, demonstrating that influenza virus receptor specificity is more complex than previously recognized.

However, major drawbacks of these assays are that they are relatively low throughput, and have been optimized for screening with whole viruses, thereby restricting the study of new pathogenic strains to specialized laboratories. In the United States, biosafety level 2 (BSL-2) facilities are recommended for working with contemporary, circulating human influenza strains, low pathogenicity avian influenza strains, and equine and swine influenza viruses³⁵, whereas increased biosafety level 3 (BSL-3) practices are recommended for non-contemporary and highly pathogenic avian influenza (HPAI) viruses. Indeed, all of the reconstructed viruses that contain one or more gene

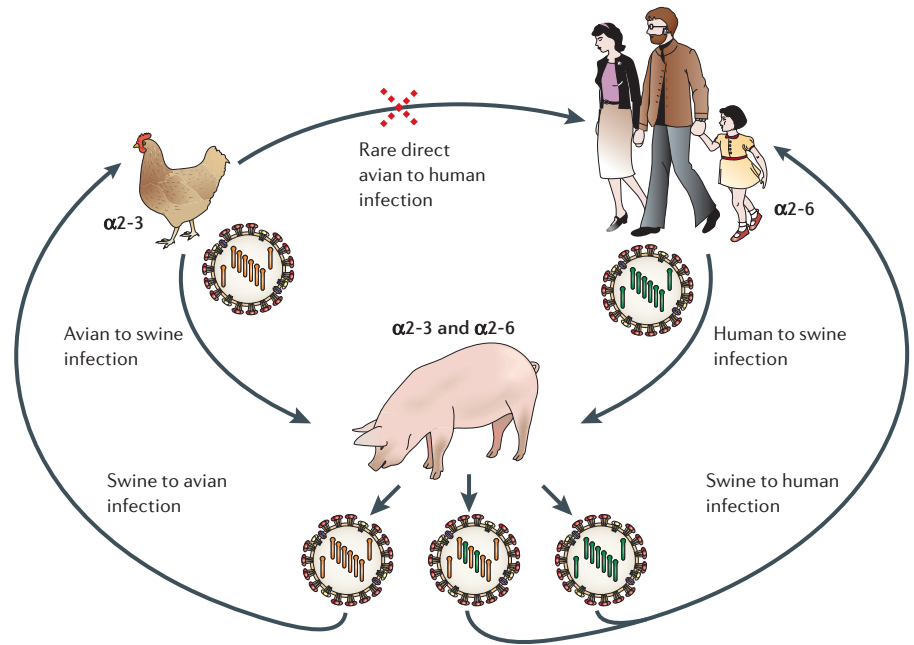


Figure 2 | Known mechanisms for the emergence of pandemic influenza A virus strains. Two of the three pandemic influenza A virus strains during the past century, H2N2 in 1957 (which caused Asian flu) and H3N2 in 1968 (which caused Hong Kong flu), arose from genetic reassortment, in which gene segments from an avian virus were mixed with genetic material from a co-infecting human virus, probably through an intermediate host, such as a pig (the mixing vessel theory). The haemagglutinin (HA) of human influenza A viruses has a binding preference for cell receptors that contain α 2-6-linked sialic-acid moieties, whereas avian viruses bind preferentially to α 2-3-linked sialic acids moieties. The respiratory epithelial cells of pigs have receptors that express both α 2-3- and α 2-6-linked sialic-acid moieties, and can be infected by both human and avian viruses. The resulting viral progeny will either be intact avian or human virus, which can only infect their respective hosts or, if reassortment yields functional virus (usually swapped HA, PB2 and/or neuraminidase (NA)), a new pandemic strain might emerge with the ability to retain efficient human-to-human transfer, but be sufficiently different (for example, by a species change in HA) to reduce the ability of the host to mount an effective immune response. PB2, polymerase basic-2 protein.

segments from the 1918 influenza virus were generated and handled under BSL-3 enhanced (BSL-3+) laboratory conditions³⁶, in accordance with the guidelines from the National Institutes of Health and the CDC³⁵. Any study that analyses potential humanization of the current H5 HPAI strains by mutagenesis of the viral genome would require extreme care and a minimum of BSL-3 facilities (TABLE 1).

Glycan microarray technologies

Since 2002, studies of the interactions of glycan-binding proteins with their glycan ligands have advanced significantly through the development of glycan microarray technologies³⁷⁻⁴³. Such technologies allow investigation of glycan-binding protein interactions on a single chip to which hundreds of different glycan structures can be coupled⁴⁰. Arrays have already been used for

Table 1 | Comparison of glycan microarray analysis of whole-influenza virus versus recombinant haemagglutinin protein

	Whole-virus assay	Recombinant HA assay
Biosafety requirements	BSL-2, 3 or 3+ (or requires inactivation)	None
Sample preparation	Days to weeks	~1 month
Assay considerations	May be influenced by NA	No NA activity
Mutational analysis	Straightforward, but obvious biosafety issues with unknown viruses	Straightforward, and no biosafety concerns
Binding	High valency, high avidity	Low valency, HA binding sites weaker avidity (12 per complex)
Detection	Unknown as yet whether high valency and avidity will distinguish between weak and strong binders	Can distinguish between weak and strong binders

BSL, biosafety level; HA, haemagglutinin; NA, neuraminidase.

the analysis of a number of glycan-binding proteins, such as glycan-specific antibodies against tumours and HIV^{40,44–46}, plant and microbial lectins^{37,40,47}, human glycan-binding proteins involved in both the innate and adaptive immune system^{40,48–57}, virus glycan-binding proteins^{40,58} and even the binding of whole cells⁵⁹.

A glycan microarray that is well suited for the analysis of influenza virus receptor specificity was developed by the Consortium for Functional Glycomics (CFG), funded by the National Institute of General Medical Sciences⁴⁰. The microarray comprises a library of structurally defined sugars with amino-terminal linkers that are robotically printed on amino-reactive *N*-hydroxysuccinimide-activated glass slides to yield a covalent amide bond. The library of more than 260 glycans range in size

from monosaccharides to decasaccharides, 61 of which contain NeuAc, the type of sialic-acid moiety found in humans. Of these, 38 contain NeuAc in α 2-3 linkages, 16 have α 2-6 linkages and 7 have α 2-8 linkages, and a further 16 contain sialic acids found in other species (for example, *N*-glycolylneuraminic acid), sialic-acid analogues or have β -linkages. Therefore, the array offers an unprecedented opportunity to simultaneously assess the details of the specificity of influenza virus HAs towards numerous glycans that have not been previously studied.

From studies conducted so far, the array seems to have value for analysis of both whole virus and recombinant HA (FIG. 3). Simultaneous binding of recombinant HAs to the different glycans is detected by fluorescent antibodies

that bind to a histidine (His) tag on the recombinant HA (FIG. 3). Details of the general properties, use and design of this technology have been well described elsewhere^{42,43}, and a regularly updated database of over 300 published and unpublished examples of plant, microbial and mammalian glycan-binding proteins analysed on the CFG glycan array can be found online (see **Further information**). Here, we focus specifically on the applications of glycan microarrays to characterize influenza–host-receptor interactions.

Recombinant HA assays. While the CFG glycan microarray was being developed, a research programme was initiated to structurally characterize the HA from the 1918 influenza virus. The usual method for HA production is bromelain release from intact viruses, followed by sucrose density gradient centrifugation and anion-exchange chromatography⁶⁰. However, owing to the increased biosafety requirements for making intact 1918 viruses, a new recombinant route was sought.

HA is normally expressed on the influenza A virus as a homotrimeric, membrane-bound glycoprotein. Each monomer is synthesized as a single polypeptide (HA0) that is cleaved by specific host proteases into two subunits, a membrane-proximal HA2 chain and a membrane-distal HA1 chain, that forms the receptor-binding pocket for binding to complex glycans with terminal sialic acids.

To form functional trimers, HA was produced in a baculovirus expression system⁶¹ with the ectodomain engineered as a fusion construct to a carboxy-terminal, trimerizing ‘foldon’ sequence from the bacteriophage T4 fibritin⁶². After removal of the trimerization domain by thrombin cleavage, the intact HA0 was crystallized and its structure was solved⁶¹. This important methodological advance for the production of the 1918 pandemic HA has recently been applied to determine the crystal structure of an avian influenza H5 HA⁶³ and a homotrimeric viral coat protein from the parainfluenza virus⁶⁴, demonstrating its usefulness for the production of other homotrimeric proteins for structural and immunological analysis.

For glycan microarray analysis, glycan-binding proteins have to be directly or indirectly labelled with a fluorescent tag⁴³. The HA cloning strategy introduces a His tag at the carboxyl terminus of the HA2 chain to enable rapid purification from the culture media. With the commercial

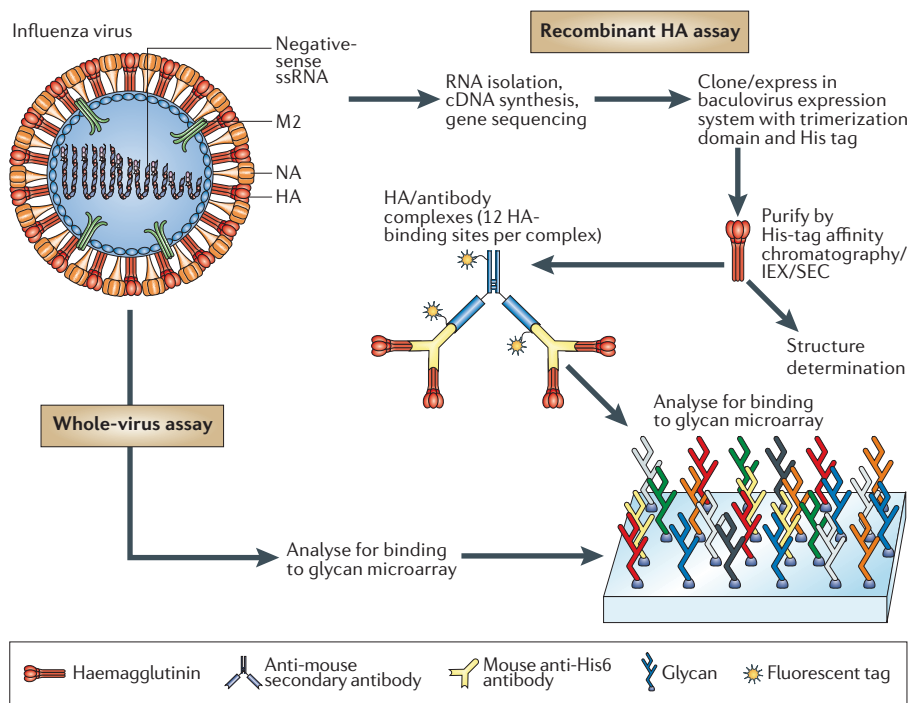


Figure 3 | Schematic representation of the two assays for analysis of the influenza A virus receptor-binding domain. Recombinant haemagglutinin (HA) analysis involves cloning, expression and purification of recombinant viral HA, with the incorporation of a histidine (His) tag. Approximately 200 different glycans are bound to a glass slide, and the binding of recombinant HA proteins to the different glycans is detected by fluorescent antibodies that bind to a His tag on the recombinant HA. Preliminary experiments⁶⁵ used His-tagged recombinant HA pre-complexed to a mouse fluorescent anti-penta-His antibody (six HA-binding sites per complex), but this failed to detect binding to the array owing to the weak HA affinity (mM range) for receptor analogues²⁶. Binding to a secondary fluorescent anti-mouse-IgG1 antibody (resulting in twelve HA-binding sites per complex) allowed the detection of fluorescence. It was concluded that the additional multivalency enabled detection through avidity effects, comparable to the natural interaction of HA trimers on the viral envelope with the host cell. The intensity of the fluorescent signal determines whether the binding is recorded as strong, weak or absent. The HA produced by this method can also be used for crystallization and structural studies. Whole-virus analysis involves glycan microarray analysis with intact virus. IEX, ion-exchange chromatography; NA, neuraminidase; M2, M2 ion-channel protein; SEC, size exclusion chromatography.

availability of anti-His-tag antibodies, the HA can be pre-complexed with fluorescent primary and/or secondary antibodies (FIG. 3). Because the affinity of HA for its glycan ligand is low (mM range)²⁶, the use of such complexes has proved highly advantageous as the secondary antibodies provide increased valency to obtain detectable signals. Another advantage of this technology is the small amount of protein that is required for analysis (~1.5–15 µg), so that the bulk of the expressed protein can be used for crystallization trials.

So far, several viral HAs from human and avian H1, H3 and H5 serotypes have been produced and analysed on the array, including variants from the 1918 pandemic influenza virus, and from the highly pathogenic H5 avian influenza strain Viet04 (A/Vietnam/1203/2004)^{63,65}. Striking differences in the receptor specificity of the different HAs were revealed not only for the type of glycan sialic-acid linkage, that is, either an α 2-3 and/or an α 2-6 sialic-acid linkage, but also in the fine specificity of HAs for other glycan modifications, such as fucosylation, sulphation and additional sialylation^{63,65} (FIG. 4). For example, a single amino-acid difference between the two variants of the 1918 pandemic influenza virus was sufficient to switch the HA from a classic human α 2-6-linked sialoside-binding pattern (Asp²²⁵), as seen with the A/South Carolina/1/1918 (SC) HA, to mixed specificity with a significant reduction in binding to α 2-6-linked sialosides and an increased affinity for α 2-3-linked sialosides, particularly for glycans with an additional negative charge, such as a sulphate or sialic acid (FIG. 4a,b) for the A/New York/1/1918 (NY) strain HA (Gly²²⁵)⁶⁵. From these results, it was proposed that the extra cavity arising from Gly²²⁵ in the 1918 NY strain HA might allow more flexibility in its glycan binding. This would enable *N*-acetylglucosamine (GlcNAc) to adopt an orientation within the HA receptor-binding site, that would allow Lys²²², which is maintained in nearly all HAs from avian and human H1, H2 and H5 serotypes, to interact preferentially with the 6-sulphated moieties. Although the biological significance of binding to sialylated and sulphated glycans is not clear at present, human mucins have been shown to contain 6-sulphated glycans, demonstrating that the necessary enzymes are present in the human epithelium and, therefore, that sialylated sulphated structures might be of importance in virus–host interactions^{33,66}.

Whereas binding of both the SC and the H3 (A/Moscow/10/1999) human viruses are highly restricted to α 2-6-linked sialosides (FIG. 4a,e), the array data clearly showed that not all human HAs have identical ligand preferences. For example, the human H1 (A/Texas/36/1991) HA has a strong preference for α 2-6 human receptor analogues, but also binds weakly to a range of α 2-3-linked sialosides⁶⁵ (FIG. 4c). It is of interest that this relatively recent H1 strain also has the Asp¹⁹⁰ Gly²²⁵ combination, similar to the 1918 NY strain (FIG. 4b), yet results from the glycan microarray show that it is better adapted to human receptors with a higher preference for α 2-6 sialosides, presumably owing to other as-yet-unidentified residues.

Both the avian H5 (A/Vietnam/1203/2004) and H3 (A/Duck/Ukraine/1/1963) serotype HAs reveal a classic preference for α 2-3 avian receptor analogues (FIG. 4d,f). Comparison between these two avian subtypes shows the power of the array in differentiating HA binding to glycan ligands that contain additional fucose linkages to GlcNAc. Ligands 25–29 on the array are all fucosylated glycans and it is clear that the human H3 HA does not bind to such ligands (FIG. 4f), whereas the H5 HA does bind them, albeit weakly (FIG. 4d). The use of such analyses allows the ready identification of a receptor footprint for each virus. Therefore, it is clear that, although human and avian virus HAs have a primary specificity for either α 2-6- or α 2-3-linked sialosides, each virus might use a different range of glycan receptors for cell entry, and the ability of a virus to infect a host cell might depend on features in the glycan structure, other than simply the type of sialic-acid linkage.

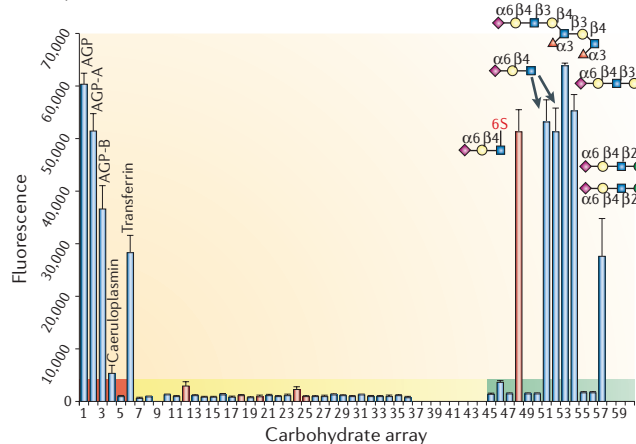
It has been suggested that the use of recombinant HAs produced in baculovirus expression systems in these microarray studies might not reflect the situation of natural virus HA expression. Like mammalian cells, insect cells can assemble, transfer and trim *N*-glycans to produce high mannose or paucimannose products⁶⁷. But insect cells differ from mammalian cells in that they are unable to process these glycan products any further to produce complex glycans that contain terminal galactose and/or sialic acids. However, the presence of influenza NA, a sialidase, ensures that complex glycans of influenza HAs usually terminate only in galactose and, therefore, the size of the *N*-glycans that are elaborated by insect cells approximate the size of the complex *N*-glycans in mammalian host cells. Indeed, our studies

using insect-cell-produced HAs, and cell-based assays with whole viruses show that glycans produced in both systems can affect receptor binding through steric hindrance if the *N*-glycan is positioned close enough to the receptor-binding site of the HA^{65,68}. In addition, results for both avian H3 (A/Duck/Ukraine/1/1963) and H1 (A/Duck/Alberta/35/1976) HAs^{63,65} were in agreement with previous whole-viral studies^{21,23}.

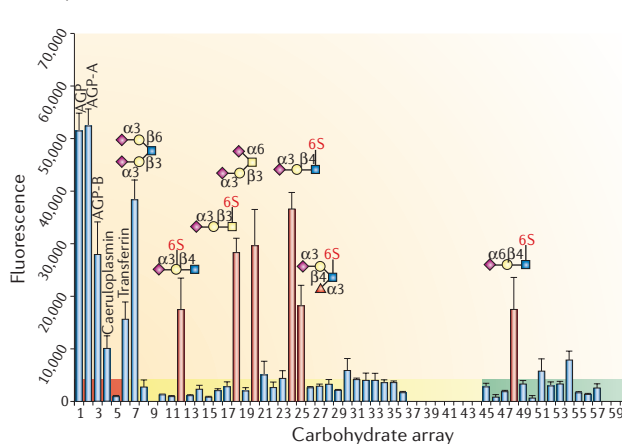
A key advantage of the baculovirus expression of recombinant HA is the ability to introduce specific mutations in the HA receptor-binding site to assess their effect on receptor specificity using the array. For example, different mutations known to switch the receptor specificity of pandemic H1, H2 and H3 viruses from an avian to a human preference were introduced into the H5N1 viral isolate, Viet04, one of the most pathogenic viruses studied so far^{69,70}. In pandemic H1, H2 and H3 viruses, two different combinations of two mutations effectively switched receptor specificity from α 2-6 to α 2-3-linked glycans and *vice-versa*. Although no clear switch in receptor specificity was observed for the H5 HA, the mutations that switched the specificity of H3 viruses from avian to human (at residues 226 and 228) not only reduced overall avian specificity for the H5 HA, but also increased specificity/avidity towards biantennary sialosides⁶³, which might function as receptors on human lung epithelial cells. These combined effects could allow an H5N1 virus to escape entrapment by mucins and increase binding to susceptible lung epithelial cells and so might provide a possible route by which H5N1 viruses could become established in the human population. Current studies are ongoing to determine whether mutations of any other residues around the HA binding pocket can cause a more obvious switch in specificity. Therefore, incorporation of mutations in and around the receptor-binding site into the recombinant HA can assess the effects on binding specificity without the need to generate potentially dangerous mutant viruses.

Whole-virus assays. Although the glycan microarray technologies are still in their infancy, the use of this technology for the analysis of whole viruses is an important goal, and the preliminary results so far look promising. Experiments with A/Puerto Rico/8/1934 virus reveal both α 2-3 and α 2-6 receptor specificity, which is in agreement with experiments from cell-based

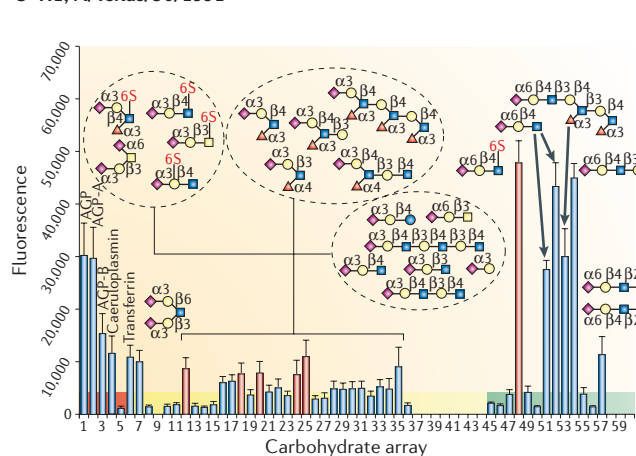
a H1; A/South Carolina/1/1918



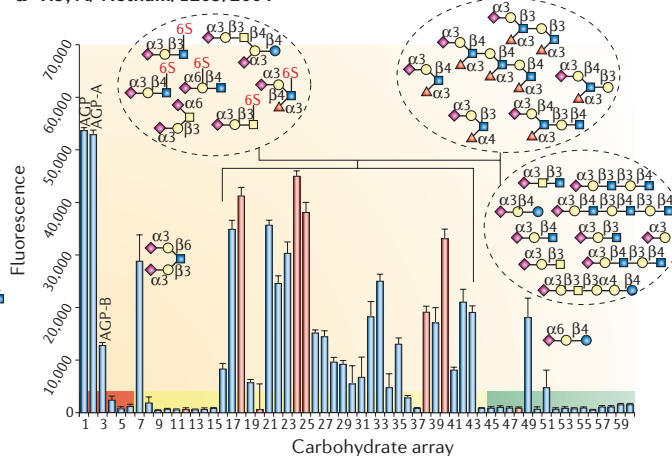
b H1; A/New York/1/1918



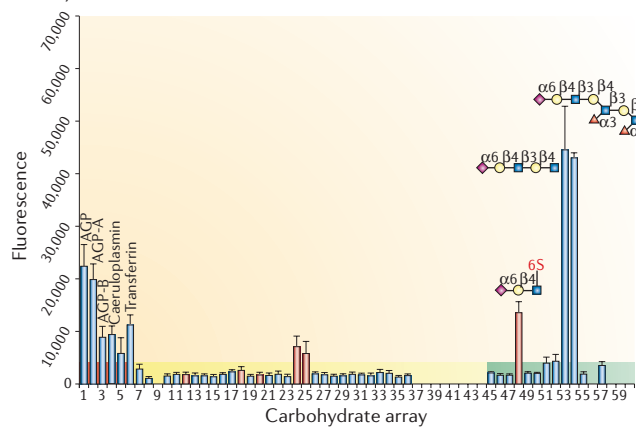
c H1; A/Texas/36/1991



d H5; A/Vietnam/1203/2004



e H3; A/Moscow/10/1999



f H3; A/Duck/Ukraine/1/1963

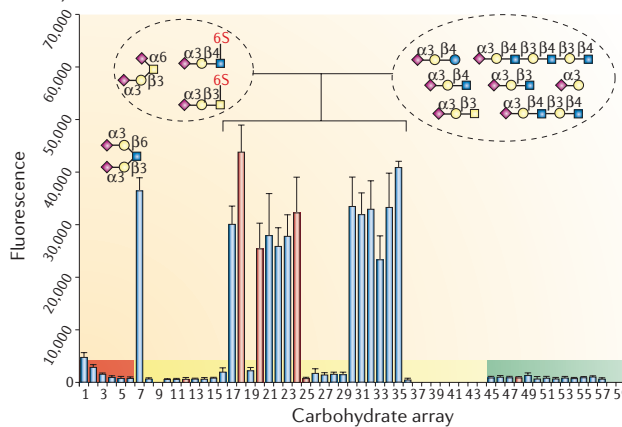


Figure 4 | Glycan microarray analysis of human H1, human H3, avian H5 and duck H3 influenza virus haemagglutinins. Glycan binding analyses as measured by fluorescence intensity are presented for natural haemagglutinins (HAs) from influenza viruses circulating during the 1918 pandemic (A/South Carolina/1/1918 (a) and A/New York/1/1918 (b)), a human H1 (A/Texas/1991) (c), an avian H5 originally isolated from a 10-year-old Vietnamese boy who died from bird flu (A/Vietnam/1203/2004) (d), a human H3 (A/Moscow/10/1999) (e) and a duck H3 (A/Duck/Ukraine/1/1963), the probable progenitor of the pandemic 1968 Hong Kong virus (f). Binding results for glycoproteins

(highlighted in red), α 2-3-linked sialosides (highlighted in yellow) and α 2-6-linked sialosides (highlighted in green) are shown. The bars corresponding to signals generated by binding to sulphated or additional negatively charged glycans are coloured red. Owing to continual glycan microarray development, a number of new ligands were printed for the analysis of A/Vietnam/1203/2004 HA. Therefore, binding to glycans 37–44 and 58–60 was not determined, except for the A/Vietnam/1203/2004. Cartoon representations of the glycans bound to different spots on the array (numbered 1 to 60) are shown. AGP, α ₁-acid glycoprotein. Data adapted from REF. 63 and REF. 65.

assays^{23,40}. Studies that compare viruses with their equivalent recombinant HAs are ongoing but have to take into account the increased valency of the virus, which might enhance binding to weak ligands. Furthermore, activity of the viral NA can also influence the outcome by destroying glycans on the array; or the NA itself could mediate binding of the virus, although the use of NA inhibitors could provide a partial solution for both scenarios.

Categorizing influenza virus receptor specificity into two main types based on preferential recognition of $\alpha 2-6$ and $\alpha 2-3$ linkages continues to be useful for understanding the adaptation of influenza strains to the individual receptors that are present on the epithelial cells of different host species. Most recently, staining of human airway epithelium by plant lectins documented for the first time that human upper airway cells have predominantly $\alpha 2-6$ -linked glycans, whereas the lower airway surprisingly has increasing numbers of cells that express $\alpha 2-3$ -linked glycans^{11–13,71}. As avian viruses have now been shown to infect human epithelial cells that have $\alpha 2-3$ -linked glycans^{10,11,13}, this observation has been proposed to account for the propensity of the H5N1 virus to localize to the lower lung during human infection and, as mentioned previously, the subsequent lack of human-to-human transmission by virus shedding^{11,13}.

However, the ability to screen for the receptor specificity of influenza virus for a highly diverse set of sialic-acid-containing glycans has revealed that the situation is more complex, with different virus strains binding preferentially to novel structures (such as sulphated and sialylated glycans) and markedly different binding patterns to glycan structures in a given class^{33,63,65}.

Conclusions and future outlook

The biological relevance of this new binding data is not yet fully understood, in part owing to a lack of knowledge about the range of glycan structures on the surface of human and avian epithelial cells and the protective mucin secretions⁶⁶. Additional information about the glycan structures produced by the natural hosts of influenza combined with the emerging data from glycan array analysis is certain to provide new biological insights relevant to modes of influenza adaptation across species barriers.

A major goal for worldwide influenza surveillance is the continuing development of new technologies that can rapidly characterize the evolution of strains that

emerge in humans and in avian populations. Currently, viruses are collected, shipped to the WHO or other secure facilities, such as the CDC, where the viruses are serotyped and undergo partial sequence analysis, before informed guesses are made as to whether any identified mutations signify antigenic drift that is relevant for incorporation into the next influenza vaccine or, in the case of avian viruses, signal significant adaptation to the human population. This powerful glycan microarray technology can now facilitate mapping of the coarse and fine specificity of previous, current and emerging influenza viruses.

A key advantage of using the baculovirus expression system described here is that we can now obtain specificity data directly from pathological specimens without the need for prior growth in eggs or other laboratory hosts. This issue is important because several earlier observations show that HA receptor variants of human viruses can be selected by growth in eggs, which could bias subsequent binding studies⁷². The use of such a recombinant system reduces the probability of unwanted mutations that can happen when maintaining whole viruses in a laboratory environment.

The technology described here, however, is limited by its reliance on existing WHO surveillance methods to correctly identify the type and subtype of each virus before any analysis can be done. The recombinant assay requires knowledge of the DNA sequence of the HA prior to cloning and expression studies, whereas the whole-cell assay requires a prior knowledge of the virus subtype to select a suitable reactive antibody. Of interest are recent advances in DNA array technologies for influenza A typing: current methods for identification require 3–7 days, but a recent report describes the development of a low-density oligonucleotide array technology (FluChip) for rapid identification of H1N1, H3N2 and H5N1 influenza A virus subtypes within 12 hours (with an accuracy of >72%)^{73,74}. Commercial products are already available, such as an influenza microarray from **CombiMatrix Corporation** that requires <5 hours for completion, and that is claimed to be able to identify all known strains of influenza A, including differentiating between highly pathogenic and less pathogenic strains⁷⁵.

By working together, future development of DNA and glycan microarray technologies could enable rapid assessment of emerging viruses to detect virus

subtypes and receptor specificities that increase the risk to the human population, and to eliminate guesswork from sequence analyses alone. A recent and exciting development is an agreement between the CDC and the CFG that will provide a customized influenza glycan microarray to allow the CDC to assess its usefulness for screening influenza field isolates. Although such technological developments still require improvement in both their accuracy and sensitivity of detection, in the future, portable machines might be used to conduct such tests in the field, adding a new dimension to characterizing and assessing avian influenza outbreaks. Such knowledge could aid decisions for deployment of emergency services to contain a potential outbreak.

James Stevens, Ola Blixt, James C. Paulson and Ian A. Wilson are at the Department of Molecular Biology; Ola Blixt and James C. Paulson are also with the Glycan Array Synthesis Core-D, Consortium for Functional Glycomics, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA.
e-mails: jstevens@scripps.edu; olablixt@scripps.edu; jpaolson@scripps.edu; wilson@scripps.edu

doi:10.1038/nrmicro1530

Published online 2 October 2006

1. Taubenberger, J. K. *et al.* Characterization of the 1918 influenza virus polymerase genes. *Nature* **437**, 889–893 (2005).
2. Neumann, G. & Kawaoka, Y. Host range restriction and pathogenicity in the context of influenza pandemic. *Emerg. Infect. Dis.* **12**, 881–886 (2006).
3. Scholtissek, C., Burger, H., Kistner, O. & Shorridge, K. F. The nucleoprotein as a possible major factor in determining host specificity of influenza H3N2 viruses. *Virology* **147**, 287–294 (1985).
4. Johnson, N. P. & Mueller, J. Updating the accounts: global mortality of the 1918–1920 'Spanish' influenza pandemic. *Bull. Hist. Med.* **76**, 105–115 (2002).
5. Antonovics, J., Hood, M. E. & Baker, C. H. Molecular virology: was the 1918 flu avian in origin? *Nature* **440**, E9 (2006).
6. Gibbs, M. J. & Gibbs, A. J. Molecular virology: was the 1918 pandemic caused by a bird flu? *Nature* **440**, E8 (2006).
7. Taubenberger, J. K. *et al.* Molecular virology: Was the 1918 pandemic caused by a bird flu? (Reply). *Nature* **440**, E9–E10 (2006).
8. WHO. *Cumulative number of confirmed human cases of avian influenza A(H5N1) reported to WHO* [online]. http://www.who.int/csr/disease/avian_influenza/country/cases_table_2006_08_23/en/index.html (2006).
9. Ito, T. *et al.* Molecular basis for the generation in pigs of influenza A viruses with pandemic potential. *J. Virol.* **72**, 7367–7373 (1998).
10. Matrosovich, M. N., Matrosovich, T. Y., Gray, T., Roberts, N. A. & Klenk, H. D. Human and avian influenza viruses target different cell types in cultures of human airway epithelium. *Proc. Natl Acad. Sci. USA* **101**, 4620–4624 (2004).
11. Shinya, K. *et al.* Avian flu: influenza virus receptors in the human airway. *Nature* **440**, 435–436 (2006).
12. Couceiro, J. N., Paulson, J. C. & Baum, L. G. Influenza virus strains selectively recognize sialyloligosaccharides on human respiratory epithelium; the role of the host cell in selection of hemagglutinin receptor specificity. *Virus Res.* **29**, 155–165 (1993).
13. van Riel, D. *et al.* H5N1 virus attachment to lower respiratory tract. *Science* **312**, 399 (2006).

14. Gambaryan, A. S., Webster, R. & Matrosovich, M. N. Differences between influenza virus receptors on target cells of duck and chicken. *Arch. Virol.* **147**, 1197–1208 (2002).
15. Wan, H. & Perez, D. R. Quail carry sialic acid receptors compatible with binding of avian and human influenza viruses. *Virology* **346**, 278–286 (2006).
16. Kuiken, T. *et al.* Avian H5N1 influenza in cats. *Science* **306**, 241 (2004).
17. Keawcharoen, J. *et al.* Avian influenza H5N1 in tigers and leopards. *Emerg. Infect. Dis.* **10**, 2189–2191 (2004).
18. Cyranoski, D. Bird flu spreads among Java's pigs. *Nature* **435**, 390–391 (2005).
19. Kuiken, T., Fouchier, R., Rimmelzwaan, G., Osterhaus, A. & Roeder, P. Feline friend or potential foe? *Nature* **440**, 741–742 (2006).
20. Levinson, B., Pepper, D. & Belyavin, G. Substituted sialic acid prosthetic groups as determinants of viral hemagglutination. *J. Virol.* **3**, 477–485 (1969).
21. Rogers, G. N. & Paulson, J. C. Receptor determinants of human and animal influenza virus isolates: differences in receptor specificity of the H5 hemagglutinin based on species of origin. *Virology* **127**, 361–373 (1983).
22. Suzuki, Y. *et al.* The hemagglutinins of the human influenza viruses A and B recognize different receptor microdomains. *Biochim. Biophys. Acta* **903**, 417–424 (1987).
23. Rogers, G. N. & D'Souza, B. L. Receptor binding properties of human and animal H1 influenza virus isolates. *Virology* **173**, 317–322 (1989).
24. Glaser, L. *et al.* A single amino acid substitution in 1918 influenza virus hemagglutinin changes receptor binding specificity. *J. Virol.* **79**, 11533–11536 (2005).
25. Gambaryan, A. S. & Matrosovich, M. N. A solid-phase enzyme-linked assay for influenza virus receptor-binding activity. *J. Virol. Methods* **39**, 111–123 (1992).
26. Sauter, N. K. *et al.* Binding of influenza virus hemagglutinin to analogs of its cell-surface receptor, sialic acid: analysis by proton nuclear magnetic resonance spectroscopy and X-ray crystallography. *Biochemistry* **31**, 9609–9621 (1992).
27. Collins, B. E. & Paulson, J. C. Cell surface biology mediated by low affinity multivalent protein-glycan interactions. *Curr. Opin. Chem. Biol.* **8**, 617–625 (2004).
28. Bovin, N. V. *et al.* Synthesis of polymeric neoglycoconjugates based on *N*-substituted polyacrylamides. *Glycoconj. J.* **10**, 142–151 (1993).
29. Gambaryan, A. S. *et al.* Specification of receptor-binding phenotypes of influenza virus isolates from different hosts using synthetic sialylglycopolymers: non-egg-adapted human H1 and H3 influenza A and influenza B viruses share a common high binding affinity for 6'-sialyl(*N*-acetyl)lactosamine. *Virology* **232**, 345–350 (1997).
30. Matrosovich, M. N. *et al.* Avian influenza A viruses differ from human viruses by recognition of sialyloligosaccharides and gangliosides and by a higher conservation of the HA receptor-binding site. *Virology* **233**, 224–234 (1997).
31. Matrosovich, M. N. *et al.* Early alterations of the receptor-binding properties of H1, H2, and H3 avian influenza virus hemagglutinins after their introduction into mammals. *J. Virol.* **74**, 8502–8512 (2000).
32. Gambaryan, A. S. *et al.* Receptor specificity of influenza viruses from birds and mammals: new data on involvement of the inner fragments of the carbohydrate chain. *Virology* **334**, 276–283 (2005).
33. Gambaryan, A. S. *et al.* H5N1 chicken influenza viruses display a high binding affinity for Neu5Ac α 2–3Gal β 1–4[6-HSO $_3$]GlcNAc-containing receptors. *Virology* **326**, 310–316 (2004).
34. Gambaryan, A. S. *et al.* Evolution of the receptor binding phenotype of influenza A (H5) viruses. *Virology* **344**, 432–438 (2006).
35. CDC-NIH. Interim CDC-HIH recommendation for raising the biosafety level for laboratory work involving noncontemporary human influenza viruses. *CDC* [online], <http://www.cdc.gov/flu/h2n2bsl3.htm>
36. Tumpey, T. M. *et al.* Characterization of the reconstructed 1918 Spanish influenza pandemic virus. *Science* **310**, 77–80 (2005).
37. Wang, D., Liu, S., Trummer, B. J., Deng, C. & Wang, A. Carbohydrate microarrays for the recognition of cross-reactive molecular markers of microbes and host cells. *Nature Biotechnol.* **20**, 275–281 (2002).
38. Disney, M. D. & Seeberger, P. H. The use of carbohydrate microarrays to study carbohydrate-cell interactions and to detect pathogens. *Chem. Biol.* **11**, 1701–1707 (2004).
39. Feizi, T. & Chai, W. Oligosaccharide microarrays to decipher the glyco code. *Nature Rev. Mol. Cell Biol.* **5**, 582–588 (2004).
40. Blixt, O. *et al.* Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc. Natl Acad. Sci. USA* **101**, 17033–17038 (2004).
41. Dyukova, V. I. *et al.* Hydrogel glycan microarrays. *Anal. Biochem.* **347**, 94–105 (2005).
42. Dyukova, V. I., Shilova, N. V., Galanina, O. E., Rubina, A. Y. & Bovin, N. V. Design of carbohydrate multiarrays. *Biochim. Biophys. Acta* **1760**, 603–609 (2006).
43. Paulson, J. C., Blixt, O. & Collins, B. E. Sweet spots in functional glycomics. *Nature Chem. Biol.* **2**, 238–248 (2006).
44. Manimala, J. C., Li, Z., Jain, A., VedBrat, S. & Gildersleeve, J. C. Carbohydrate array analysis of anti-Tn antibodies and lectins reveals unexpected specificities: implications for diagnostic and vaccine development. *Chembiochem.* **6**, 2229–2241 (2005).
45. Huflejt, M. E. *et al.* Glycan array identifies specific signatures of anti-glycan autoantibodies in sera of breast cancer patients: diagnostic, prognostic and therapeutic opportunities. *Breast Cancer Res. Treat.* **94**, S85 (2005).
46. Lawrie, C. H. *et al.* Cancer-associated carbohydrate identification in Hodgkin's lymphoma by carbohydrate array profiling. *Int. J. Cancer* **118**, 3161–3166 (2006).
47. Alvarez, R. A., Lee, A., Davis, C., Hoffmann, J. & Blixt, O. Defining the binding specificity of commercially available plant lectins using a printed glycan array. *Glycobiology* **15**, 1207 (2005).
48. van Vliet, S. J. *et al.* Carbohydrate profiling reveals a distinctive role for the C-type lectin MGL in the recognition of helminth parasites and tumor antigens by dendritic cells. *Int. Immunol.* **17**, 661–669 (2005).
49. Tateno, H., Crocker, P. R. & Paulson, J. C. Mouse Siglec-F and human Siglec-8 are functionally convergent paralogs that are selectively expressed on eosinophils and recognize 6'-sulfo-sialyl Lewis X as a preferred glycan ligand. *Glycobiology* **15**, 1125–1135 (2005).
50. Palma, A. S. *et al.* Ligands for the β -glucan receptor, Dectin-1, assigned using 'designer' microarrays of oligosaccharide probes (neoglycolipids) generated from glucan polysaccharides. *J. Biol. Chem.* **281**, 5771–5779 (2006).
51. McGreal, E. P. *et al.* The carbohydrate-recognition domain of Dectin-2 is a C-type lectin with specificity for high mannose. *Glycobiology* **16**, 422–430 (2006).
52. Coombs, P. J., Graham, S. A., Drickamer, K. & Taylor, M. E. Selective binding of the scavenger receptor C-type lectin to Lewisix trisaccharide and related glycan ligands. *J. Biol. Chem.* **280**, 22993–22999 (2005).
53. Bochner, B. S. *et al.* Glycan array screening reveals a candidate ligand for Siglec-8. *J. Biol. Chem.* **280**, 4307–4312 (2005).
54. Collins, B. E. *et al.* Masking of CD22 by cis ligands does not prevent redistribution of CD22 to sites of cell contact. *Proc. Natl Acad. Sci. USA* **101**, 6104–6109 (2004).
55. van Vliet, S. J. *et al.* Carbohydrate profiling reveals a distinctive role for the C-type lectin MGL in the recognition of helminth parasites and tumor antigens by dendritic cells. *Int. Immunol.* **17**, 661–669 (2005).
56. Crocker, P. R. Siglecs in innate immunity. *Curr. Opin. Pharm.* **5**, 431–437 (2005).
57. Galustian, C. *et al.* High and low affinity carbohydrate ligands revealed for murine SIGN-R1 by carbohydrate array and cell binding approaches, and differing specificities for SIGN-R3 and langerin. *Int. Immunol.* **16**, 853–866 (2004).
58. Nam, H. J. *et al.* Identification of the sialic acid structures recognized by minute virus of mice and the role of binding affinity in virulence adaptation. *J. Biol. Chem.* **281**, 25670–25677 (2006).
59. Nimrichter, L. *et al.* Intact cell adhesion to glycan microarrays. *Glycobiology* **14**, 197–203 (2004).
60. Ha, Y., Stevens, D. J., Skehel, J. J. & Wiley, D. C. H5 avian and H9 swine influenza virus haemagglutinin structures: possible origin of influenza subtypes. *EMBO J.* **21**, 865–875 (2002).
61. Stevens, J. *et al.* Structure of the uncleaved human H1 hemagglutinin from the extinct 1918 influenza virus. *Science* **303**, 1866–1870 (2004).
62. Frank, S. *et al.* Stabilization of short collagen-like triple helices by protein engineering. *J. Mol. Biol.* **308**, 1081–1089 (2001).
63. Stevens, J. *et al.* Structure and receptor specificity of the hemagglutinin from an H5N1 influenza virus. *Science* **312**, 404–410 (2006).
64. Yin, H. S., Wen, X., Paterson, R. G., Lamb, R. A. & Jardetzky, T. S. Structure of the parainfluenza virus 5 F protein in its metastable, prefusion conformation. *Nature* **439**, 38–44 (2006).
65. Stevens, J. *et al.* Glycan microarray analysis of the hemagglutinins from modern and pandemic influenza viruses reveals different receptor specificities. *J. Mol. Biol.* **355**, 1143–1155 (2006).
66. Lo-Guidice, J. M. *et al.* Sialylation and sulfation of the carbohydrate chains in respiratory mucins from a patient with cystic fibrosis. *J. Biol. Chem.* **269**, 18794–18813 (1994).
67. Kost, T. A., Condreay, J. P. & Jarvis, D. L. Baculovirus as versatile vectors for protein expression in insect and mammalian cells. *Nature Biotech.* **23**, 567–575 (2005).
68. Klenk, H. D., Wagner, R., Heuer, D. & Wolff, T. Importance of hemagglutinin glycosylation for the biological functions of influenza virus. *Virus Res.* **82**, 73–75 (2002).
69. Govorkova, E. A. *et al.* Lethality to ferrets of H5N1 influenza viruses isolated from humans and poultry in 2004. *J. Virol.* **79**, 2191–2198 (2005).
70. Maines, T. R. *et al.* Avian influenza (H5N1) viruses isolated from humans in Asia in 2004 exhibit increased virulence in mammals. *J. Virol.* **79**, 11788–11800 (2005).
71. Gagneux, P. *et al.* Human-specific regulation of α 2-6-linked sialic acids. *J. Biol. Chem.* **278**, 48245–48250 (2003).
72. Rogers, G. N. *et al.* Host-mediated selection of influenza virus receptor variants. Sialic acid- α 2,6Gal-specific clones of A/duck/Ukraine/1/63 revert to sialic acid- α 2,3Gal-specific wild type *in ovo*. *J. Biol. Chem.* **260**, 7362–7367 (1985).
73. Mehlmann, M. *et al.* Robust sequence selection method used to develop the FluChip diagnostic microarray for influenza virus. *J. Clin. Microbiol.* **44**, 2857–2862 (2006).
74. Townsend, M. B. *et al.* Experimental evaluation of the FluChip diagnostic microarray for influenza virus surveillance. *J. Clin. Microbiol.* **44**, 2863–2871 (2006).
75. Lodes, M. J. *et al.* Use of semiconductor-based oligonucleotide microarrays for influenza A virus subtype identification and sequencing. *J. Clin. Microbiol.* **44**, 1209–1218 (2006).

Acknowledgements

The work was supported in part by grants from the National Institute of Allergy and Infectious Diseases, and the National Institute of General Medical Sciences. The authors thank Ruben Donis for suggestions. This is publication 18250 from the Scripps Research Institute.

Competing interests statement

The authors declare no competing financial interests.

DATABASES

The following term in this article is linked online to:
 Entrez Genome: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=genome>
 A/Puerto Rico/8/1934

FURTHER INFORMATION

The Scripps Research Institute: http://www.scripps.edu/e_index.html
 CFG: <http://www.functionalglycomics.org>
 CombiMatrix Corporation: <http://www.combimatrix.com>
 Nomenclature Committee of the Consortium for Functional Glycomics: <http://glycomics.scripps.edu/CFGnomenclature.pdf>
 Access to this links box is available online.