Review
Exploration of Sialic Acid Diversity and Biology Using Sialoglycan Microarrays

Lingquan Deng,1 Xi Chen,2 Ajit Varki1
1 Departments of Medicine and Cellular & Molecular Medicine, Glycobiology Research and Training Center, University of California, San Diego, La Jolla, CA 92093-0687
2 Department of Chemistry, University of California, Davis, CA 95616

Received 2 June 2013; accepted 4 June 2013
Published online 13 June 2013 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/bip.22314

ABSTRACT:
Sialic acids (Sias) are a group of α-keto acids with a nine-carbon backbone, which display many types of modifications in nature. The diversity of natural Sia presentations is magnified by a variety of glycosidic linkages to underlying glycans, the sequences and classes of such glycans, as well as the spatial organization of Sias with their surroundings. This diversity is closely linked to the numerous and varied biological functions of Sias. Relatively large libraries of natural and unnatural Sias have recently been chemically/chemoenzymatically synthesized and/or isolated from natural sources. The resulting sialoglycan microarrays have proved to be valuable tools for the exploration of diversity and biology of Sias. Here we provide an overview of Sia diversity in nature, the approaches used to generate sialoglycan microarrays, and the achievements and challenges arising. © 2013 Wiley Periodicals, Inc. Biopolymers 99: 650–665, 2013.

Keywords: sialic acids; diversity; sialoglycan microarrays

INTRODUCTION
DNA, RNA, proteins, and glycans are the four biopolymers essential to all known life forms.1,2 Glycans differ from the other types of biopolymers in several ways. Glycan biosynthesis is not template-driven and cannot be accurately predicted by any known method. In addition, glycans form branched structures, display numerous modifications, and have far more diversity in overall structure. These and other differences partly account for the historical lag in chemical and biological studies of glycans, compared to the other classes of biopolymers. The development of technologies to elucidate the structures and functions of these complex molecules have recently drawn increasing scientific interest and it is now well appreciated that glycans play vital roles in numerous complex biological processes and systems,3,4 a fact recently recognized by a special report from the US National Academies, urging greater investments in this area.5

Among the various monosaccharide building blocks of glycans, sialic acids (Sias) are rather unusual. They are typically found as terminal residues on the glycan chains of vertebrate glycoconjugates. In contrast to most other common monosaccharides, which are aldoses or ketoses with five or six carbons, Sias are α-keto acids with a nine-carbon backbone (Figure 1). Furthermore, the types of natural modifications found on Sias far exceed that of any other monosaccharides (Figure 1).6–13 Additional differences include the limited occurrence of Sias (primarily in the deuterostome lineage of animals and certain types of bacteria), as well as its unusual form of nucleoside monophosphate-activated sugar donor (CMP-Sia).12 These and other distinct features of Sias contribute to higher structural complexity and the potential for more unique and varied biological functions, in comparison to other monosaccharides.
Accumulating evidence also indicates that evolutionary selection forces play a critical role in the diversification of glycans, which are intrinsically linked to their biological functions. Among various glycan classes, Sias appear to be the most rapidly evolving. In keeping with the recommended style for this contributed article, we will not attempt to cover the whole field; rather we will review our own work within the context of the field, including our personal insights, and speculate on where research is headed in the future. We will begin by briefly discussing the diversity and biology of Sias in nature, with examples from our own labs in the context of work of others. Recent advances in the synthesis of sialosides and the application of sialoglycan microarrays will then be reviewed, from our perspective.

**DIVERSITY AND BIOLOGY OF SIALIC ACIDS IN NATURE**

**Natural Diversity and Biology of Sialic Acid forms and Modifications**

The two core Sia forms are neuraminic acid (Neu) and 2-keto-3-deoxy-D-glycero-D-galacto-nononic-acid (Kdn) (Figures 2a and 2b). More than 50 different types of naturally occurring Sia variants (c.f., Table 1 in Ref. 12) have been found based on modifications to these two forms, including 5-N-acetylation, hydroxylation of 5-N-acetyl group, O-acetylation, O-methylation, O-lactylation, O-sulfation, O-phosphorylation, and intramolecular lactam or lactone formation. Among the diverse naturally occurring Sias, N-acetyleneuraminic acid (Neu5Ac) and N-glycolyneuraminic acid (Neu5Gc) are the two most abundant ones in mammals (Figures 2c and d). The nine-carbon Sia backbone is also shared by other nonulosonic acids (NulOs) that are found in prokaryotes, such as legionaminic acid (Leg) and pseudaminic acid (Pse) (Figures 2e and 2f). This review will focus on the more common Sias, some of which are shared by animals and pathogenic bacteria.

The diversity of Sia forms and modifications are closely associated with biological functions in both intracellular and intercellular environments, from early discovered effects on the physical properties of glycoproteins, to their protective and masking roles, and to more recently established mediation and modulation of numerous biological processes including signaling, fertilization, immunity, growth and differentiation. Among the many Sia modifications, 9-O-acetylation is best studied. When modified by O-acetyl group(s), Sias can be blocked from recognition by intrinsic...
lectins such as Siglecs (Sia-binding immunoglobulin-like lectins).\textsuperscript{24,25} O-Acetylation can also modulate interactions between Sias and microbial proteins, in a manner either benefiting or harming the host.\textsuperscript{22,26–30} O-Methylation has been found to make Sias resistant to sialidases, indicating a mechanism of biological significance.\textsuperscript{31}

One variation at the C-5 position has generated some uniquely interesting questions related to human evolution and disease.\textsuperscript{32} Specifically, hydroxylation of Neu5Ac at the 5-\text{O-acetyl} group forms Neu5Gc, a molecule that differs from Neu5Ac by a single oxygen atom (Figures 2c and 2d). Humans lost the ability to synthesize Neu5Gc due to a 92 bp exon deletion in the \textit{CMAH} gene encoding CMP-Neu5Ac hydroxylase,\textsuperscript{33,34} which occurred \(~3\) million years ago.\textsuperscript{35} However, Neu5Gc can be metabolically incorporated into human cells and tissues, via human consumption of Neu5Gc-rich foods such as red meat,\textsuperscript{36–37} even in the face of circulating anti-Neu5Gc antibodies (Abs) in human blood.\textsuperscript{38,39} The interactions between these “xeno-autoantigens” and “xeno-autoantibodies” are currently postulated to have numerous effects on human diseases, ranging from induction of chronic inflammation, tumor progression/suppression, effects on biotherapeutic molecules and cells,\textsuperscript{40–46} and even to xenotransplantation rejections.\textsuperscript{47–52} Moreover, certain bacterial toxins preferentially recognize glycan epitopes containing metabolically incorporated Neu5Gc.\textsuperscript{53}

**Natural Diversity and Biology of Sia Linkages and Underlying Glycans**

Sias are glycosidically linked via C-2 to underlying glycans (Figure 1). They can be attached to a peripheral galactose (Gal) residue via \(\alpha\)-2,3, \(\alpha\)-2,4, or \(\alpha\)-2-6 linkage, to \(\alpha\)-5-acetylglactosamine (GalNac) via \(\alpha\)-2-6 linkage, to \(\alpha\)-5-acetylglucosamine (GlcNac) via \(\alpha\)-2-4 or \(\alpha\)-2-6 linkage, or to another Sia via \(\alpha\)-2-8 or \(\alpha\)-2-9 linkages.\textsuperscript{13} The structural diversity of the underlying glycan structures and modifications further increase the complexity of Sia-containing molecules. Combinations of the multiple forms of Sias, different sialyl linkages, and diverse underlying glycans generate many thousands of possible Sia presentations.

This diversity of Sia linkages and underlying glycans is known to be biologically relevant. One well-known example is the recognition of different Sia linkages by human and avian adapted influenza A viruses.\textsuperscript{24} Although the human adapted ones bind Sia\(\alpha\)-2-6Gal preferentially, the avian adapted ones prefer Sia\(\alpha\)-2-3Gal.\textsuperscript{27,55–57} The relative expression level and distribution of Sia\(\alpha\)-2-6Gal and Sia\(\alpha\)-2-3Gal in host respiratory tract are believed to play an important role in the infection and transmission of these viruses.\textsuperscript{38,39} Endogenous lectins such as Siglecs also have distinct Sia ligand preferences. Human Siglec 2 (CD22), for example, selectively recognizes \(\alpha\)-2-6-linked Sias,\textsuperscript{60} but human Siglec-1 (sialoadhesin) prefers \(\alpha\)-2-3-linked Sias.\textsuperscript{61} In addition, Sia linkage to the underlying glycans plays a role in the progression and spread of human malignancies. For example, carcinoma behavior can be modulated by \(\alpha\)-2-6-sialylation on N-glycans.\textsuperscript{62–66} Moreover, underlying glycans can also influence or even determine Sia binding events. Sias with an underlying Lewis x or Lewis a structure are preferentially recognized by selectins, and sulfation at the C-6 position of an underlying Gal or GalNac residue further enhances the binding affinity.\textsuperscript{67–70} Additionally, different underlying glycans capped with the same \(\alpha\)-2-6 linked Sia were found to show markedly different binding affinities to CD22.\textsuperscript{71} Numerous other examples could be cited.

**Natural Diversity and Biology of Polysialic Acids**

Sia-containing homopolysaccharides and heteropolysaccharides have been found in both animals and bacteria.\textsuperscript{7,17,72–76} Homo-PolySias are found in a few animal glycoproteins, such as \(\alpha\)-2-8-linked Polysia on the N-glycans of the neural cell adhesion molecule (NCAM) and on O-glycans of some fish egg glycoproteins. The capsular polysaccharides of certain pathogenic bacteria can express \(\alpha\)-2-8-, \(\alpha\)-2-9, or alternating \(\alpha\)-2-8- and \(\alpha\)-2-9-linked Sias polymers. Heteropolysaccharides containing \([\alpha\text{-}4\text{Sia}_{\alpha\text{-}2\text{-}6}\text{Gal}_{\alpha\text{-}1\text{-}2}]\) and \([\alpha\text{-}4\text{Sia}_{\alpha\text{-}2\text{-}6}\text{Glc}_{\alpha\text{-}1\text{-}2}]\) repeats have also been found on the capsular polysaccharides of \textit{N. meningitidis} serotypes W135 and Y, respectively.\textsuperscript{77} PolySia chains based on Neu5Ac, Neu5Gc, Kdn, or Leg building blocks have been reported, and the inter-residue linkages can vary from \(\alpha\)-2-8-, \(\alpha\)-2-9, to \(\alpha\)-2-4, and to \(\alpha\)-2-5-O-glycoyNeu5Gc.\textsuperscript{75} The degree of polymerization and O-acetylation,\textsuperscript{78,79} homo- or heteropolymerization, and the dynamic equilibrium between polylactones and polycarboxylates further add to the complexity.\textsuperscript{80,81} Altogether, there is a considerable range of diversity in polySia chains.

**Natural Diversity and Biology of Sialoglycans in Clustered Saccharide Patches**

The concept of “clustered saccharide patches” was first suggested in the context of understanding selectin ligands,\textsuperscript{67} and was recently updated to take new evidence into account.\textsuperscript{82} Cell surface glycans can be imagined as the outermost aspects of a tropical rainforest canopy, or more accurately, in a more dynamic fashion as a kelp bed or coral reef in the ocean (analogy suggested by P. Gagneux). Viewed in this manner, one can appreciate the importance of considering the glycome as a whole when studying glycan-based interactions. Such clustered saccharide patches are proposed to involve multiple glycan
components and their surroundings. Thus multiple copies of the same glycan or multiple copies of different glycans, or glycans and nonglycan structures including adjacent sulfates or peptide sequences on the same scaffold/carrier or on different ones, can participate to make up clustered saccharide patches. Indeed, clustered saccharide patches comprised of Sia and adjacent sulfates or peptide sequences were reported as ligands for selectins.70,83–85

A similar analogy can be applied to the sialome, as a subset of the glycome (Figure 3).86 Five hierarchical levels of Sia structural and spatial complexities can be conceived of, in analogy to the leaves/flowers, stems, branches, trees and the forest, respectively. This theoretical concept is supported by accumulating experimental evidence.82 For example studies of Ab binding to Sia-containing ligands revealed that in certain cases, a complex of several sialylated glycans would generate specific binding, despite the fact that a single glycan from the complex was not adequate for recognition.87,88 Another example involves the “glycosynapse,” where Sia epitopes of the glycans in the cellular surface microdomains play an important functional role. The Sia epitopes cluster with other components in the microdomains and together they mediate cell adhesion and signaling to influence the cellular phenotypes.89 Sia-containing clustered saccharide patches on RBCs of different blood types can also explain the differential binding of Sia-recognizing proteins in relation to ABO blood group polymorphisms in humans.90 Furthermore, evidence for Siglec or Ab recognition of bacterial capsular polysaccharides exists, indicating the presence of Sia-containing clustered saccharide patches on pathogens.91,92

**SIALIC ACID RECOGNIZING PROTEINS**

Many functional roles of glycans are realized or modulated by interactions with glycan-binding proteins (GBPs).93–95 Glycan–glycan interactions also play a critical role in biological processes such as cell–cell recognition and membrane organization, but that aspect will not be discussed here. Interested readers are referred to relevant reviews.96,97

The variety and importance of Sias as ligands in glycan–protein recognition phenomena were extensively reviewed as well.98–102 A large number of Sia-binding proteins have already been identified in viruses, bacteria, plants, invertebrates and vertebrates, and this number continues to grow. Besides exploring their occurrence and identifying their cognate ligands, special emphasis has been placed on studying Sia-binding proteins from an evolutionary point of view, together with their biological significance in human health and disease, especially in immunity, infectious disease and cancer biology.14,23,103,104 In addition to Sia-binding lectins, Abs that recognize Sias have attracted considerable research interest, and the discovery of cancer biomarkers involving Sia and Sia-recognizing Abs has become an active research area.46,105,106 For example, co-existence of Neu5Gc and anti-Neu5Gc Abs in human body represents a unique human phenomenon.38–41,43,44,46 Another class of Sia-recognizing proteins are Sia-recognizing enzymes including sialyltransferases and sialidases. These enzymes are essential to the chemistry and biology of Sias.102 The extensive literature on this subject is not reviewed here. However, examples of their use in sialochemistry will be presented in the following section.

**SYNTHESIS OF SIALOGLYCANS BEARING NATURAL AND UNNATURAL SIAS**

In nature, the majority of modifications on Sias are added after the sialoglycosidic bond formation (so called postglycosylosidal modifications or PGMs).107 One clear exception involves the formation of Neu5Gc from Neu5Ac, which takes place at the CMP-Sia level.108–110 Another exception is the formation of O-acetylated Neu5Ac-containing capsular polysaccharides, possibly by direct transfer of O-acetylated Neu5Ac from O-acetylatedCMP-Neu5Ac.111 It is also likely that one form of mammalian O-acetylation takes place at the CMP-Sia level.112 The gene sequences for most of the enzymes responsible for adding modifications to the Sia residues in sialoglycans have not been identified to date. The few exceptions include an O-acetyltransferase from Campylobacter jejuni,113 an O-acetyltransferase for the capsular polysaccharide of Neisseria meningitidis serogroups C, W-135, and Y,79,114 and a candidate for a vertebrate enzyme.115 Therefore, it is impractical to synthesize the majority of sialosides containing naturally occurring postglycosylational modifications in vitro, by totally following natural biosynthetic pathways. Among various chemical and chemoenzymatic methods developed for the synthesis of sialoglycans bearing natural and unnatural Sias, the one-pot multienzyme (OPME) chemoenzymatic sialylation method116 is the most efficient for generating libraries of sialosides102 that can be used for sialoglycan microarray studies. In this approach, naturally occurring and non-natural Sia forms, or their six-carbon precursors such as derivatives of N-acetylmannosamine (ManNAc) or mannose, can be synthesized chemically or enzymatically. These can then be used as substrates, by a CMP-Sia synthetase and a suitable sialyltransferase in the absence or the presence of a Sia aldolase, for the synthesis of desired sialoglycans. Diverse sialosides can be synthesized due to unusual promiscuities of sialyltransferases and other related sialoside biosynthetic enzymes, including Sia aldolases and CMP-Sia synthetases, toward substrate modifications. Obtaining a broad array of sialyltransferases,117–120 CMP-Sia
Synthetases, and Sia aldolases, with high expression levels, good solubility and high activities from bacterial species has made the preparative-scale and even large-scale preparation of comprehensive libraries of sialoglycans a reality. Designing enzyme mutants based on protein crystal structures also enables the expansion of the available sialoglycan products and renders the synthetic and purification schemes more efficient.

**FIGURE 3** Analogy of the sialome to the canopy of a forest. Five hierarchical levels of structural and organizational complexities are compared. Sia structures are assimilated with tree leaves and flowers (a); Sia linkages are compared to tree stems (b); Underlying glycans are like branches of trees (c); Glycan classes are compared with the tree trunks (d); and finally the spatial organization of glycans and their surroundings are assimilated to the entire forest (e). Figure reproduced with permission.
STUDIES OF SIA DIVERSITY AND BIOLOGY USING SIALOGLYCAN MICROARRAYS

As discussed above, interactions between glycans and GBPs underlie much of the biological significance of glycans. The advent of glycan microarrays has revolutionized the screening of GBP specificities and fueled the discovery of new GBPs, providing invaluable information in a high-throughput manner. During the years after its introduction in 2002,127,128 glycan microarray studies have extended by development of glycan libraries and immobilization chemistries, proof-of-principle demonstrations, discovery of novel GBPs and their binding specificities, as well as applications in vaccine/inhibitor identification and biomarker discoveries.127–137 Here, we focus on discussing the development of a subclass of such arrays, the sialoglycan microarrays, and their versatile utility in studying Sia diversity and biology.

Natural and Synthetic Sialoglycan Libraries

To date, one of the most widely used glycan libraries is the one developed by the Consortium for Functional Glycomics (CFG). It consists of over 760 glycan structures (610 mammalian glycans and 153 microbial glycans) of various types, including a total of around 170 Sia-containing structures (Table I) (http://www.functionalglycomics.org). However, quite a few Sia-containing structures in this sialoglycan library have an overlapping or identical terminal di-/tri-/tetrasaccharide, some varying merely in spacer structures. Nevertheless, sialoglycan microarrays comprised of ~50–90 of these structures have been readily produced and applied in various studies.138–142 During the past few years, Paulson’s group has been developing chemical and chemoenzymatic methods to create a large library of sialoside analogues (Table I).144–146 In 2008, Blixt et al.144 utilized 9-azido-Neu5Ac intermediates and synthesized a library of 44 sialoside analogues bearing unnatural acyl substituents at the C-9 position, among which 16 were α2-3-linked and 28 α2-6-linked. In 2012, Rillahan et al. took advantage of click chemistry and further developed the idea of using a minimum quantity of synthetic sialoside analogues on microarrays to screen for high affinity Siglec-binding ligands.145 This time, a considerably larger library was generated thanks to the facile copper catalyzed azide-alkyne cycloaddition (CuAAC) chemistry. Besides modifications at the C-9 position, C-5 modified Sias were also produced. In total, 224 Sia-analogues with either α2-3 or α2-6 linkages, and with lactose as underlying glycan, were produced. It is noteworthy that these compounds were printed on glycan microarrays for screening without column purifications, resulting in higher throughput than the previous effort.144 However, most of the Sias on these arrays are unnatural ones. Another sialoglycan
library comes from the Feizi group, which consists of more than 700 naturally derived and then chemically tagged neoglycolipids. Among these, there are about 120 sialoglycans with different Sia forms and linkages, and underlying glycans (Table I). \(^{148-151}\) This library of sialoglycans differs from most others by the nature of glyccan source, that is, they are derived from natural sources instead of chemically built as in most other cases. It also contains a few \(\alpha2-9\) linked Sias that other major sialoglycan libraries lack.\(^{150}\) Our own labs have been studying the diversity and biology of naturally occurring Sias and are especially interested in the chemistry and biology of the nonhuman Sia molecule, Neu5Gc, as well as various Sia O-acetylations (c.f., sections and relevant references above, and additional Refs. 30,78,152–156). A library of over 70 synthetic sialoglycans was thus produced, which is unique in that it contains various pairs of Neu5Ac and Neu5Gc counterparts bearing the same glycosidic linkages, underlying glycans and spacers, that is, matched structures differing only by a single oxygen atom. Moreover, many pairs were further diversified by 9-O-acetylation on the core Sia structures (Table I).\(^{46,157–160}\) This library is thus well suited for comparing protein recognition properties between Neu5Ac- and Neu5Gc-containing glycans, as well as for exploring the effects of Sia O-acetylation on binding. Furthermore, the number of sialoglycans in this library is actively increasing. The Cummings and Smith group also prepared a sialoglycan library in collaboration with one of us. The library has 77 sialoglycans with 16 different Sia modifications based on three core Sia structures (Neu5Ac, Neu5Gc, and Kdn), either \(\alpha2-3\) or \(\alpha2-6\) linked, and four different underlying glycan structures (Table I).\(^{158,161,162}\) Among the 16 modified Sias, 13 are naturally occurring and three have not yet been found in nature. This sialoglycan library expanded the Sia diversity formerly covered in glycan microarray systems by adding the methylated and lactylated Sia structures. The Bovin lab and Wong lab have two other libraries of sialoglycans, containing about 40 and 30 different structures, respectively. Both are synthetically produced, and most of them are Neu5Ac-bearing structures (Table I).\(^{163–166}\) Besides monovalent sialoglycans, the Bovin lab also produced sialoglycans on PAA polymer backbones and with additional functionalities such as biotin or fluorescein labels. The Institute for Glycomics in Griffiths University also has a library of ~20 sialoglycans.\(^{167}\)

**Current Platforms for Sialoglycan Microarrays and the Immobilization Chemistries**

Based on the libraries discussed above, standard sialoglycan microarray platforms have been developed. Typically, functional group modified glass substrates are used to accommodate robotically printed microarrays that contain hundreds to thousands of tiny sialoglycan solution spots. After incubation to ensure proper attachment, the glycan microarrays are blocked and are then ready for high-throughput screening studies. Many immobilization chemistries have been developed in the past decade.\(^{134,135}\) Although elegant proof-of-principle studies regarding each chemistry have been extensively demonstrated, most of them have not yet been tested broadly for more advanced applications. In the case of sialoglycan microarrays, only three immobilization methods are widely used and applied for various studies (Table I). These are chemistries involving reactions of an amine with NHS-ester, or with epoxide, and noncovalent hydrophobic attachment. Occasionally, the high affinity noncovalent biotin and streptavidin interaction was utilized for immobilizing sialoglycans, but mostly in a multiwell plate format.\(^{157,168,169}\)

Besides the standard sialoglycan microarrays, a few mini-sialoglycan array platforms appeared to couple with other attractive techniques/technologies and provide interesting potential. In 2008, the Cheng group in collaboration with one of us reported a sialoglycan biosensor platform based on surface plasmon resonance (SPR) spectroscopy, to study Sia-lectin interactions in a label-free format and in real-time.\(^{170}\) The biotin-neuraminidin interaction was used to immobilize four biotinylated Sias, Neu5Acx2-3Lac, Kdnx2-3Lac, Neu5Acx2-6Lac, and Kdnx2-6Lac on the sensor chip and the system was interrogated with a number of lectins. Interaction kinetics and affinity data were obtained, and the effects of Sia structure and linkage variations on lectin binding proved to be detectable by the system. In 2009, a follow-up SPR imaging study was carried out for more detailed characterizations of the sensor surface chemistry and lectin interactions with additional techniques including fluorescence microscopy and atomic force microscopy (AFM).\(^{171}\) In 2011, the Flitsch group presented a platform to generate sialoglycans in situ on surfaces in an array format and the synthesis was monitored by MALDI-TOF (matrix-assisted laser desorption ionization time-of-flight) mass spectrometry (MS). After completion of the in situ chemoenzymatic reactions, the sialoglycan surfaces were directly used for studying interactions by cells bearing recombinant Siglces.\(^{172}\) The authors demonstrated that a self-assembled monolayer (SAM) surface was suitable for studying whole mammalian cell interactions with limited nonspecific interactions. Combination of glycan microarray and MALDI-TOF techniques was also exploited to assess enzymatic activities and specificities of influenza neuraminidases from whole viruses.\(^{173}\) However, this study used different surfaces and immobilization chemistry. Their sialoglycan array took advantage of a DNA/DNA hybridization method and glass slides bearing microreactors were produced to anchor the surface DNA strand. In 2012, the Sun group reported two interesting...
Studying Sia-Recognizing Proteins by Sialoglycan Microarrays

It is evident that Sia binding lectins can be found in many microbes, plants and animals. Studies of these Sia-recognizing proteins by sialoglycan microarrays are rapidly growing. In part due to a joint effort between the CFG and the Centers for Disease Control (CDC), and partly because of the global public health concern/high importance of the topic, the proteins from influenza viruses of various origins, subtypes, clades, and strains have been extensively studied by sialoglycan microarrays. These include influenza virus HAs from various subtypes such as H5N1, H2N2, H1N1, H17N10 and H3N2 as well as the N10 neuraminidase-like protein from the same origin. 

The system was interrogated by a range of lectins, in the presence or absence of free Sia competitors. An SPR biosensor system was also developed based on the same immobilization chemistry to study lectin and influenza hemagglutinin (HA) binding to the attached sialoglycan polymers. The authors claim this system better mimics the 3D nature of cell surface Sia presentation and should provide considerable advantages for relevant studies. In the other study, Ma et al. fabricated a liposomal sialoglycan microarray by the Staudinger ligation method. Gangliosides GM1 and GM3 inserted liposomes were immobilized on glass slides and remained intact, after which lectin and bacterial toxin bindings were evaluated. This system was claimed to present sialoglycans on a surface mimicking native cell membranes, and could thus be deemed an important tool for various applications.

All of the above mentioned mini-sialoglycan array platforms provide interesting additional functionality to the standard sialoglycan microarrays and show considerable promise. However, most of these examples included only a few sialoglycans, and some systems were designed in a well-based array format. Thus, expansion of the sialoglycan libraries and miniaturization of the array spots in these systems are yet to be explored, so that truly high-throughput evaluation of sialoglycan-involved interactions can be realized.

Wang et al. took a different approach to study the influenza HAs using sialoglycan microarrays. They examined the effects of HA glycosylation on its host receptor binding and found that removal of certain glycan structures on viral surface glycoproteins could result in higher-affinity Abs elicited and better neutralization activities of these Abs. More recently, both the influenza HA homologue from bat influenza virus (H17N10) and the N10 neuraminidase-like protein from the same origin were found to show no binding to ligands on the CFG sialoglycan microarrays. In contrast, a neuraminidase mutant from human H3N2 influenza virus avidly bound to Sia ligands on the sialoglycan microarrays. The observed neuraminidase binding was of much higher affinity (µM range) compared to HA bindings (usually low mM). Other Sia-binding proteins from different types of viruses have been studied by sialoglycan microarrays, including human JC polyomavirus, serotype 1 reovirus, bovine coronavirus and canine adenovirus, as well as parasites like Toxoplasma gondii and Eimeria tenella.

The sialoglycan microarrays derived from more extensive glycan libraries (e.g., CFG and the Feizi group) have been used to study specific proteins or microbes directly. However, the two more recently established sialoglycan microarrays were tested extensively against a wide range of proteins and microbes, from common plant lectins to viruses.

Because of their important roles in human immunity, Siglecs are another type of favored candidates for sialoglycan microarray studies. The Paulson lab and the Chen lab utilized different chemical and enzymatic methods to generate relatively large libraries of unnatural and natural Sias in a high-throughput manner. These libraries of Sias could be directly used for array fabrication and Siglec ligand identification. A few of the unnatural Sia analogues generated by Rillahan et al. have proved to be of high affinity to a range of Siglecs. Besides the C-5 and C-9 position modifications on the Sia core structure, the Kelm lab very recently showed that modifications at the Sia C-4 position could act synergistically with C-9 modification and also enhance Siglec binding. Higher affinity Sia analogues could be potentially used as Siglec inhibitors or for targeted therapeautic delivery to treat, for example, B cell leukemia. However, Siglecs like CD22 are recently noted to be expressed not only on B cells, but also on dendritic cells and gastrointestinal eosinophils. Thus, precautions and extra investigations are needed to successfully pursue this strategy.

Sialoglycan microarrays have also been used to identify cancer biomarkers. Taking advantage of its unique feature, the
paired Neu5Gc/Neu5Ac sialoglycan microarray platform was utilized in our labs to study the unusual anti-Neu5Gc Abs in humans. Sera from cancer and noncancer patients were characterized, and Abs against Neu5Gc,Gc or Neu5Gc,NAc were found to be more prominently present in patients with carcinomas than with other diseases. Furthermore, the corresponding patient sera and purified polyclonal Abs which showed strong anti-GcSTn reactivity both proved capable of killing human tumors expressing GcSTn, via either complement-dependent cytotoxicity or Ab dependent cellular cytotoxicity. This study was followed up by a detailed LC-MS analysis of the polyclonal human anti-Neu5Gc Abs, and all four IgG subclasses of Abs were confirmed to present after the human immune response to the xeno-autoantigen Neu5Gc. The same sialoglycan microarray platform was also used to detect anti-Neu5Gc Ab responses in patients with Kawasaki Disease. Other glycan microarray platforms (including microarrays with various non-Sia glycan epitopes), have also been used to study complex samples like human sera and for biomarker discoveries. Aside from studying Sia-recognizing lectins and Abs, sialoglycan microarrays can also be applied to measure enzymatic activities of various neuraminidases/sialidases. In addition, sialyltransferase reactions have been monitored directly on glycan microarrays.

**Studying Viruses, Bacteria, and Whole Mammalian Cells by Sialoglycan Microarrays**

No matter what one finds using purified glycan-binding proteins, there is always the possibility that the binding specificity will be different when studying the intact organism that expresses the same protein. Thus there is a need to study interactions of intact organisms with arrays. To date, the most studied examples are influenza viruses. Receptor specificities of human, avian, and porcine influenza viruses have been examined and comparisons among these strains have yielded insights into the molecular basis for their receptor specificity, transmissibility, as well as virulence. In most of these studies, intact or biotinylated viruses were applied to the sialoglycan microarrays and detected by fluorochrome-labeled virus-binding Abs or streptavidin-fluorochromes, respectively. However, fluorochrome-conjugated viruses could be directly used for glycan microarray studies. Other types of viruses were also studied by sialoglycan microarrays. In 2007, a neoglycolipid microarray was used to study receptor specificity of simian virus 40 (SV40), finding that the N-glycolyl GM1 ganglioside was a preferred receptor. Cell studies and molecular modeling further supported the finding. Preferential binding to Neu5Gc over Neu5Ac was also observed for some other viruses and bacterial toxins. There are many more such proteins and microbes that await exploration. The potential impact of this type of preferential binding in host–pathogen interactions on human health and disease warrants further research, and sialoglycan microarrays containing pairs of Neu5Gc and Neu5Ac glycan will serve as a novel platform for such investigations.

Surprisingly, there are very few reports to date on studying intact bacteria using sialoglycan microarrays. In 2004, Seeberger and coworkers tested *Escherichia coli* binding on a different type of glycan microarray comprised of mannose (Man), glucose (Glc), N-acetyl-glucosamine (GlcNAc), galactose (Gal), or fucose (Fuc) structures. Unexpectedly, the printed glycan microarrays with oligomannose structures showed no increase in bacterial binding capacity compared to mono-mannose arrays. The authors attributed this phenomenon to the possibility that the tested bacterial strain only required a single mannose residue for recognition and the multivalency of mannose and stereochemistry of the intermammose linkages played little role. However, a more recent study from the same group showed that trimannose and nonamannose structures attached to a cantilever array sensor tip did show differential binding affinities to the same *E. coli* strain tested earlier, indicating an increased multisite and multivalent binding for the nonmannose structure compared to less complex mannose structures. Taken together, these studies showed a need to improve the glycan microarray platform for studying intact bacteria. In this regard we have recently embarked on using sialoglycan microarrays to investigate a group of Gram-positive bacteria (Deng and Varki, et al., unpublished observations). Bacterial adhesins/mutants and corresponding intact isogenic strains are being tested on printed sialoglycan microarrays, and after careful optimization, the data show interesting correlations between ligand spectra/affinity and virulence. The effect of shear force on bacterial binding also needs investigation.

Testing whole mammalian cells on sialoglycan microarrays is also an emerging endeavor. Although whole cell bindings were demonstrated in a millimeter-spot-sized microarray system as early as in 2004, robotically printed micrometer-sized glycan arrays have not been used to study whole mammalian cells until very recently. Rillahan et al. demonstrated binding of Siglec-bearing Chinese hamster ovary (CHO) cells and CD22-expressing human B-cells on their Sia analogue microarrays. In another report, Arndt et al. tested a range of human cancer cell lines directly on their glycan microarrays (various types of glycans, not completely sialoglycans) and on a lectin microarray, and they found an inverse relationship between how many glycans the tested cells could recognize and how many types of intrinsic glycans were expressed on those cells. This study also pointed to the importance of characterizing/evaluating cell surface glycosylation status when using those common laboratory cell
lines for cell-based assays. Well-plate based glycan arrays have also been developed for studies of glycan-cell interactions recently. Initial efforts to pair the screening of binding specificity of secreted proteins with corresponding protein-bearing cells are currently undertaken as well.

Decoding Clustered Saccharide Patches by Sialoglycan Microarrays

Current glycan microarray platforms are usually composed of individual spots each displaying a single glycan structure. However, cell surfaces present glycans in a highly heterogeneous fashion. Thus there is also a need to try to explore glycan–protein interactions in a manner more similar to the natural state. Although difficult to prove conclusively, the hypothesis of clustered saccharide patches is worth testing. The information gained could further enhance our understanding of glycan–protein interactions and improve drug design strategies. This concept can be to some extent tested using glycan microarrays.

In 2010, Wu and coworkers published a glycan microarray design of such a nature. They spotted the glass slides with mixed glycans, for example, SSEA4/Gb5, Globo H/Gb5, Gb4/Gb5, Gb3/Gb5, Gb2/Gb5, Bb2/Gb5, in 1:1 molar ratio and interrogated the slides with anti-Gb5 Ab. Interestingly, the SSEA4/Gb5 mixed glycan spot consistently showed higher binding than the Gb5-alone glycan spot. On the other hand, neighboring glycans could also exert a negative effect on the Gb5 and anti-Gb5 interaction, possibly via steric hindrance, as evidenced by lower binding of the other mixed glycan spots compared to the Gb5-alone spot. Further tests were done by mixing glycans in varied concentrations and by using synthetic oligomannose dendrimers containing different ratios of Man4 and Man9 structures. Both experiments demonstrated Ab binding could be affected by the density and structures of neighboring glycans. In parallel, the Willison lab combined a few techniques and recently produced polyvinylidene difluoride membranes (PVDM) affixed glass slides for the generation of sialoglycan (ganglioside) microarrays. They called this setup a “combinatorial glycoarray,” and used it primarily to study neuropathy-associated anti-glycolipid autoantibodies. These Abs specifically recognize complexes of glycolipid pairs, but fail to interact with either of the glycolipids alone. Other exciting developments include that relatively stable lipid bilayer coated glycan microarray slides have been readily made. This is promising due to its better mimicry of the cell surface features. It can be envisaged that this type of surface may serve as an ideal platform for decoding the effects of clustered saccharide patches on protein and cell bindings.

COMPARISONS OF SIALOGLYCAN MICROARRAY PLATFORMS

In the DNA microarray field, the microarray quality control (MAQC) consortium has been established and major efforts toward standardization of processing and reporting microarray data have been made. As the number of new glycan microarray platforms continues to grow and applications using these platforms expand, the need for cross-comparison among platforms and established guidelines for glycan microarray experiments has become increasingly relevant and urgent.

In 2012, two labs took the initiative and cross-compared their newly developed sialoglycan microarray platforms, both developed using the same chemistries. These two glycan microarrays presented comparable sialoglycans produced by the same synthetic strategy from the same group. However, the immobilization chemistries used for attaching the sialoglycans on the two platforms were different, and the linkage monosaccharide ring was opened in one of the two arrays. These microarrays were reciprocally tested against various Sia-binding proteins and analyzed in the two labs. This comparison yielded a lot of useful information but also identified major challenges in the standardization of glycan microarray experiments. For instance, immobilization methods appeared to strongly influence some binding results. Also, the close-ring immobilization method was evidently critical for some glycan-recognition events, and an open-ring immobilization could sometimes lead to the absence of detectable binding. This is in keeping with another intragroup comparison of glycan microarrays differing only in the immobilization method, where it was found that the ring-closed immobilization method was critical for many glycan-recognition events. Lower signal-to-noise ratio bindings based on open-ring immobilization method was also observed in another independent array comparison study. Besides these, binding data with various glycans from the same source but arrayed at different locations were also compared and reported. More recently, a larger scale glycan microarray comparison involving five different research groups has been carried out (personal communication from Dr. Lara K. Mahal, NYU).

CONCLUSIONS AND PERSPECTIVES

The diversity and biology of Sias represent part of the vast information and knowledge gained in the field of glycosciences. It is evident that sialoglycan microarrays are playing an increasingly important role in elucidating this diversity and biology. During the past decade, the advancement of glycan microarray development and applications is indeed highly laudable and exemplary; however, there is still plenty of room for further developments and improvements.
1. Larger libraries of sialoglycans, consisting of either natural structures or synthetic analogues, are still greatly desired. Just by comparing currently available sialoglycan libraries and the diversity of Sia structures occurring in nature, one can realize how big the gap is, let alone the unlimited possibilities of synthetic Sia analogues. For example, it will be interesting to see sialoglycan microarrays made of 4-O-acetylated Sias or polySias. These structures are largely un-explored by sialoglycan microarrays. In addition, current efforts are mainly focused on elucidating binding of terminal glycans. However, binding effects resulting from intact natural sialoglycan classes and underlying peptides/lipids have to be further evaluated.

2. Sialoglycan microarrays have shown to be a suitable platform to demonstrate the concept of clustered saccharide patches. Further studies such as mixing more components in a single glycan spot, controlling spacing between glycans in the spot, and testing their effects on protein binding, are worthwhile.

3. Ligand specificity of many more Sia-recognizing proteins, viruses, bacteria and whole mammalian cells are worthy of investigation in detail. This would require more extensive collaborations between glycoscientists (chemists and biologists) and researchers from other fields, such as microbiologists. The joint effort will certainly benefit all parties involved.

4. To date, more than 20 different types of immobilization methods for glycan microarray fabrication have been developed. For example, the immobilization of glycans by photochemistry is a highly versatile and efficient method. This can be realized by producing photoprobe-derivatized glycans, direct attachment of derivatized glycans, or a “photo-click” hybrid method. Photogenerated glycan microarrays have also been successfully applied to identify immunogenic glycan antigens specific to certain bacteria. However, most of the immobilization methods developed to date were only demonstrated in proof-of-concept type of studies. The exceptional versatility and efficiency of the photochemical immobilization methods warrant further development and broader applications. Cross-comparison of these immobilization methods and insight into their merits and disadvantages are needed for further microarray standardization and applications.

5. Technical issues such as Sia contamination of the most widely used glycan microarray blocking reagent bovine serum albumin (BSA) (our observations, unpublished result), and sample dilution effects, and so forth, are also important considerations in conducting glycan microarray experiments and analyzing glycan microarray data.

6. Initial efforts have been undertaken, but the cross-comparison and standardization of sialoglycan microarray experiments and microarray data processing has a long way to go. Guiding rules for the interpretation of glycan microarray data are also to be established. Moreover, building common glycan microarray databases and further enabling relevant glycobioinformatics would be very helpful and beneficial to the entire glycoscience community.

7. Combination of glycan microarrays with other methods and techniques will, on the one hand, help validate the glycan microarray data, and on the other hand, often add another dimension to the findings. For example, real-time monitoring of the binding events and additional binding kinetics data can be obtained with biosensors like SPR and quartz crystal microbalance (QCM). Isothermal titration calorimetry (ITC) can provide thermodynamic information, and detailed single molecular interaction data can be acquired by using AFM. Combination of microarrays with nanotechnology can greatly enhance the sensitivity of detection. Furthermore, comparisons with other assays like enzyme-linked immunosorbent assay (ELISA) and dot blots and combination with MS, computational modeling, X-ray crystallography, etc., have been and will continue to play major roles in the elucidation of glycan–protein interactions.

8. After glycan microarray screening of glycan–protein interactions and complementary methods to validate the bindings, follow-up biomedical and/or biological studies are critical. These will ultimately move the initial findings into more clinically and/or biologically relevant and more profound discoveries/applications to combat diseases and benefit human health.

9. This review focused on studies of Sia diversity and biology by using sialoglycan microarrays, but the concepts, trends, and ideas discussed in this incomplete survey are readily applicable to other types of glycans and glycan microarrays.

The authors thank Miriam Cohen, Jerry Fong, and Stevan Springer for helpful comments.

REFERENCES

1. Marth, J. D. Nat Cell Biol 2008, 10, 1015–1016.
2. Varki, A. Cold Spring Harb Perspect Biol 2011, 3, 1–14; doi: 10.1101/cshperspect.a005462.
3. Varki, A.; Cummings, R. D.; Esko, J. D.; Freeze, H. H.; Stanley, P.; Bertozzi, C. R.; Hart, G. W.; Etzler, M. E. Essentials of Glycobiology; Cold Spring Harbor Laboratory Press: Plainview, NY, 2009.

4. Hart, G. W.; Copeland, R. J. Cell 2010, 143, 672–676.

5. National Research Council, Transforming Glycoscience: A Roadmap for the Future; The National Academies Press: Washington, DC, 2012.

6. Schauer, R. Adv Carbohydr Chem Biochem 1982, 40, 131–234.

7. Troy, F. A. Glycobiology 1992, 2, 5–23.

8. Varki, A. Glycobiology 1992, 2, 25–40.

9. Kelm, S.; Schauer, R. Int Rev Cytol 1997, 175, 137–240.

10. Zanetta, J. P.; Pons, A.; Iversen, M.; Mariller, C.; Leroy, Y.; Timmerman, P.; Schauer, R. Glycobiology 2001, 11, 663–676.

11. Schauer, R. Glycoconjugate J 2000, 17, 485–499.

12. Angata, T.; Varki, A. Chem Rev 2002, 102, 439–469.

13. Angata, T.; Varki, A. Chem Rev 2002, 102, 439–469.

14. Varki, A.; Cummings, R. D.; Esko, J. D.; Freeze, H. H.; Stanley, N.; Varki, A.; Muchmore, E. Proc Natl Acad Sci USA 2003, 100, 12045–12050.

15. Banda, K.; Gregg, C. J.; Chow, R.; Varki, N. M.; Varki, A. J Biol Chem 2012, 287, 28852–28864.

16. Padler-Karavani, V.; Yu, H.; Cao, H.; Chokhawala, H.; Karp, F.; Varki, N.; Chen, X.; Varki, A. Glycobiology 2008, 18, 818–830.

17. Taylor, R. E.; Gregg, C. J.; Padler-Karavani, V.; Ghaderi, D.; Yu, H.; Huang, S.; Sorensen, R. U.; Chen, X.; Inostroza, J.; Nizet, V.; Varki, A. J Exp Med 2010, 207, 1637–1646.

18. Mahly, K. N.; Schauer, R.; Shaw, L. Biochimie 2001, 83, 623–634.

19. Nguyen, D. H.; Tangvoranuntakul, P.; Varki, A. J Immunol 2005, 175, 228–236.

20. Hedlund, M.; Tangvoranuntakul, P.; Takematsu, H.; Long, J. M.; Housley, G. D.; Kozutsumi, Y.; Suzuki, A.; Wynshaw-Boris, A.; Ryan, A. E.; Gallo, R. L.; Varki, N.; Varki, A. Mol Cell Biol 2007, 27, 4340–4346.

21. Hedlund, M.; Padler-Karavani, V.; Nizet, V.; Varki, A. Proc Natl Acad Sci USA 2008, 105, 18936–18941.

22. Pham, T.; Gregg, C. J.; Karp, F.; Chow, R.; Padler-Karavani, V.; Cao, H.; Chen, X.; Witztum, J. L.; Varki, N. M.; Varki, A. Blood 2009, 114, 5225–5235.

23. Ghaderi, D.; Taylor, R. E.; Padler-Karavani, V.; Diaz, S.; Varki, A. Nat Biotechnol 2010, 28, 863–867.

24. Padler-Karavani, V.; Hurtado-Ziola, N.; Pu, M.; Yu, H.; Huang, S.; Muthana, S.; Chokhawala, H. A.; Cao, H.; Secrest, P.; Friedmann-Morvinski, D.; Singer, O.; Ghaderi, D.; Verma, I. M.; Liu, Y. T.; Messer, K.; Chen, X.; Varki, A.; Schwab, R. Cancer Res 2011, 71, 3352–3363.

25. Park, J. Y.; Park, M. R.; Bui, H. T.; Kwon, D. N.; Kang, M. H.; Oh, M.; Han, J. W.; Cho, S. G.; Park, C.; Shim, H.; Kim, H. M.; Kang, M. J.; Park, J. K.; Lee, J. W.; Lee, K. K.; Kim, J. H. Cell Reprogram 2012, 14, 353–363.

26. Padler-Karavani, V.; Varki, A. Xenotransplantation 2011, 18, 1–5.

27. Yeh, P.; Ezzelarab, M.; Bovin, N.; Hara, H.; Long, C.; Tomiyama, K.; Sun, E.; Ayares, D.; Awwad, M.; Cooper, D. K. Xenotransplantation 2010, 17, 197–206.

28. Basnet, N. B.; Ide, K.; Tahara, H.; Tanaka, Y.; Ohdan, H. Xenotransplantation 2010, 17, 440–448.

29. Ezzelarab, M.; Ayares, D.; Cooper, D. K. Immunol Cell Biol 2005, 83, 396–404.

30. Miwa, Y.; Kobayashi, T.; Nagasaka, T.; Liu, D.; Yu, M.; Yokoyama, I.; Suzuki, A.; Nakao, A. Xenotransplantation 2004, 11, 247–253.

31. Byres, E.; Paton, A. W.; Paton, J. C.; Lofling, J. C.; Smith, D. F.; Wilce, M. C.; Talbot, U. M.; Chong, D. C.; Yu, H.; Huang, S.; Chen, X.; Varki, N. M.; Varki, A.; Rossjohn, J.; Beddoe, T. Nature 2008, 456, 648–652.

32. Paulson, J. C.; Sadler, J. E.; Hill, R. L. J Biol Chem 1979, 254, 2120–2124.

33. Daniels, P. S.; Jeffries, S.; Yates, P.; Schild, G. C.; Rogers, G. N.; Paulson, J. C.; Wharton, S. A.; Douglas, A. R.; Skehel, J. J.; Wiley, D. C. EMBO J 1987, 6, 1459–1465.

34. Chandrasekaran, A.; Srinivasan, A.; Raman, R.; Viswanathan, K.; Raguram, S.; Tumpey, T. M.; Sassekharan, V.; Sassekharan, R. Nat Biotechnol 2008, 26, 107–113.
