Advances in the Biology and Chemistry of Sialic Acids

Xi Chen†,* and Ajit Varki‡,*

†Department of Chemistry, University of California, Davis, California 95616 and‡Departments of Medicine and Cellular and Molecular Medicine, Glycobiology Research and Training Center, University of California, San Diego, La Jolla, California 92093-0687

Sialic acids are a subset of nonulosonic acids. Not long ago, it was thought that sialic acids (Sias) were unique inventions of the deuterostome lineage of animals, which emerged around the time of the Cambrian expansion ~530 million years ago, with certain pathogenic bacteria having then “acquired” them from hosts by gene transfer (1, 2). However, the relevant genes of bacterial pathogens were then found to be only distantly homologous to corresponding host genes (3). Meanwhile, work from multiple investigators over the past few decades has shown that the unusual nine-carbon backbone of Sias is in fact shared by a larger family of nonulosonic acids (NulOs), which are much more widely distributed in nature (4−6). Furthermore, the key steps in the biosynthesis of nonulosonic acids share remarkable similarities, and the genes involved are homologous. These aspects have recently been discussed extensively in a phylogenomic evaluation of nonulosonic acids (7). Although the other forms of nonulosonic acids, such as legionaminic acid and pseudaminic acid, have sometimes been called “bacterial sialic acids” (4), we prefer to reserve the term Sia for the nine-carbon sugars found both in the deuterostome lineage of animals and in certain bacteria, which are based on a neuraminic acid (Neu) or a 2-keto-3-deoxy-nonulosonic acid (Kdn) backbone (7) (Figure 1). Thus, this Review will focus only on these “traditional” Sias, emphasizing recent challenges and advances at the interface of chemistry and biology.

NATURAL STRUCTURAL DIVERSITY IN SIALIC ACIDS

Even within this restricted subset of nonulosonic acids of the deuterostome lineage, there is a remarkable array of natural modifications, exceeding that of any other common monosaccharide (1, 3, 8). The reasons for this

ABSTRACT Sialic acids are a subset of nonulosonic acids, which are nine-carbon α-keto aldonic acids. Natural existing sialic acid-containing structures are presented in different sialic acid forms, various sialyl linkages, and on diverse underlying glycans. They play important roles in biological, pathological, and immunological processes. Sialobiology has been a challenging and yet attractive research area. Recent advances in chemical and chemoenzymatic synthesis, as well as large-scale E. coli cell-based production, have provided a large library of sialo-side standards and derivatives in amounts sufficient for structure−activity relationships studies. Sialoglycan microarrays provide an efficient platform for quick identification of preferred ligands for sialic acid-binding proteins. Future research on sialic acid will continue to be at the interface of chemistry and biology. Research efforts not only will lead to a better understanding of the biological and pathological importance of sialic acids and their diversity but also could lead to the development of therapeutics.

*Corresponding authors, chen@chem.ucdavis.edu, a1varki@ucsd.edu.
Many possible modifications and sialosidic linkage varieties found in nature are not shown here. For sialic acids in deuterostome animals, it was once thought that each type of Sia was unique to different ages can be presented on different underlying glycans. Beyond variations within the Sia molecule, additional diversity arises from a variety of glycosidic linkages from C-2 to the underlying glycans. Further complexity arises from the fact that these linkages can be presented on different underlying glycan chains, variations that can further alter the biology of Sias.

Beyond variations within the Sia molecule, additional diversity arises from a variety of glycosidic linkages from C-2 to the underlying glycans. Again, details have been discussed elsewhere (1–3) regarding families of sialyltransferases that synthesize the different linkages. Further complexity arises from the fact that these linkages can be presented on different underlying glycan chains, variations that can further alter the biology of Sias. Combinations of all these possibilities generate a wide variety of presentations of Sias in nature, and we are only beginning to scratch the surface of this diversity from the chemical, synthetic, and biological point of view. A further level of complexity arises because sialo-glycans on cell surfaces can be organized into “clustered saccharide patches” that involve interactions with other glycans, modulating recognition by different Sia-recognition proteins (13).

EVOLUTIONARY PATTERNS OF SIALIC ACID DIVERSITY

For sialic acids in deuterostome animals, it was once thought that each type of Sia was unique to different lineages of deuterostome animals. This is actually not the case, and most of these kinds of Sias are probably present in most vertebrates. However, there are marked cell-type and species-specific differences in the levels of modifications found in nature, as well as instances of complete elimination of some kinds of Sias from the entire lineage. An example is the complete loss of biosynthesis of N-glycolyneuraminic acid in humans (8, 10, 11) and very likely as an independent event in sauropods (birds and reptiles) (14).

SIALIC ACIDS AS KEY COMPONENTS OF BACTERIAL POLYSACCHARIDES

Multiple bacterial polysaccharides and lipo polysaccharides are known to contain Sias either as terminal structures or in the form of polysialic acid. In several cases they can be also modified, particularly with O-acetyl groups (5, 13). Interestingly many of these Sia-producing bacteria seem to have independently invented Sias by convergent evolution in order to dampen immune responses in vertebrates, and the biosynthetic pathways involved are related to those that can synthesize the more distantly related nonulosonic acids (7). In addition to the well-known effects of microbial Sias in negative charge repulsion, inhibiting alternate complement pathway activation by factor H recognition (16), and masking of underlying antigenic residues, it has recently emerged that the Sias on these pathogens “take advantage” of inhibitory Siglecs present on innate immune cells. Thus by engaging these Siglecs, the bacteria send a false “self signal” to the innate cells and avoid attack (17).

Of note, the great majority of Sia-producing bacteria that have been reported to date are human-specific pathogens or commensals. One theory to explain this is the fact that no microbe has reinvented the ability to express the sialic acid N-glycolyneuraminic acid (Neu5Gc). This could be because the conversion of an N-acetyl group to and N-glycolyl group is a difficult chemical barrier. Regardless of the reasons, when humans lost Neu5Gc and ended up with an excess of the precursor Neu5Ac, our Siglecs then apparently underwent evolutionary adjustments to recognize this change (10). This may have opened up a niche for bacteria that can express Neu5Ac and take advantage of the Neu5Ac-prefering human Siglecs. In keeping with this,
Neu5Ac has been “re-invented” multiple times by various human-specific pathogens.

**SIALOBIОLOGY AS A CHALLENGING AREA FOR ANALYTICAL AND SYNTHETIC CHEMISTRY**

Sialic acids have many major biological roles, ranging from embryogenesis to neural plasticity to pathogen interactions, and these are detailed elsewhere (8, 10, 18). At the present time, the methodologies to study these modifications and their biological roles remain somewhat limited. In comparison to progress in studies of other monosaccharides, much improvement is still needed in analytical approaches to Sias, as well in their synthesis. Many of the difficulties relate to the unusual nine-carbon backbone of Sias, the relative acid lability of its glycosidic linkages, and the instability of some of its modifications. Examples of major difficulties in analysis compared to other monosaccharides include the migration and loss of O-acetyl esters, the difficulty in obtaining stable derivatized molecules for mass spectrometric analyses, and the instability of CMP-sialic acids.

Meanwhile, chemical sialylation has been considered as one of the most challenging glycosylation reactions as a result of the hindered and disfavored tertiary anomic center, the presence of an electron-withdrawing carboxyl group linked to the anomic carbon, and the lack of a neighboring participating group in sialic acids to regulate the stereochemistry outcome of the sialyl products. Some of these N-protecting groups also allowed access to different sialic acid forms (e.g., Neu5Gc) by deprotection of the glycosylated product followed by derivatization at C-5 amino group (20). Recently developed donors include C-4-aminated sialyl hemiketal donor (31) (Figure 2, panel b), 5-N,4-O-carbonyl (32, 33) (Figure 2, panel c), and N-acetyl-N,4-O-carbonyl (34–36) (Figure 2, panel d) protected sialyl donors. C2-Hemiketal sialyl donors with a C-4 cyclic secondary amine auxiliary (Figure 2, panel b) improved α-stereoselectivity in the formation of sialyl linkages with primary (90–98% yields with α/β ratios from 91:9 to 98:2 in comparison to 94–98% yields with α/β ratios from 1:2 to 1.5:1 for similar donors with no C-4 auxiliary) and secondary alcohols (90–95% yields with α/β ratios varied from 91:9 to 96:4) (31). The α-selectivities in the formation of 2–6-linked sialyl galactosides or sialyl glucosides were dependent on substrates with α/β ratios from 3:2 to α only while yields were quite consistent (89–91%) (31). In addition, N,N-dimethylglycolamide was used as a new C-1 auxiliary in sialyl donors including sialyl chlorides, sulfoxides, and phosphites (Figure 2, panel e) for chemical sialylation reactions to improve the stereoselective formation of α2–3- and α2–6-sialyl linkages (37). Compared to a sialyl donor with a traditional methyl ester protecting group at C-1, the donor with the N,N-dimethylglycolamide auxiliary at C-1 showed improved α-selectivity for the formation of sialyl linkages (α/β ratios improved from 1:10–5:1 to 2:1–13:1 with similar or improved yields). The C-1 auxiliary neighboring group participation strategy was general and applicable for common sialylation donors including sialyl chlorides, sulfoxides, and phosphites (37). The strategy was also applicable to optimize dehydrative chemical sialylation conditions when C2-sialyl hemiketals were used as sialyl donors (Figure 2, panel f) (38). Sialyl N-phenyltrifluoroacetimidate donors (Figure 2, panel g) were shown to be effective donors for direct sialylation when a catalytic amount of trimethylsilyl trifluoromethanesulfonate (TMSOTf) was used as a promoter (39). Some of these sialyl donors, including C-5 N-trifluoroacetyl (N-TFA) (Figure 2, panel a) (26) and 5-N,4-O-carbonyl protected (Figure 2, panels c and d) (32, 33) sialyl donor as well as
the 1,5-lactame derivative of sialic acid (40, 41), were used for the synthesis of more challenging α2–8-linked disialylated oligosaccharides in moderate yields (23, 26, 30). The reactivity of the C-8 hydroxyl of sialic acid acceptors for the synthesis of α2–8-linked disialylated oligosaccharides was also improved using the C-5 N-TFA protecting group (26).

Another useful strategy for efficient chemical synthesis of complex sialosides was to use chemically synthesized sialyloligosaccharide building blocks for further glycosylations to produce more complex structures (42, 43). Despite recent advances, current chemical synthesis of sialosides remains a time-consuming process and requires skillful expertise.

**KEYWORDS**

**Biology:** The science of life and living organisms, including their structure, function, growth, origin, evolution, distribution, and taxonomy.

**Carbohydrates:** Polyhydroxyaldehydes, polyhydroxyketones, and their derivatives, or compounds that produce polyhydroxyaldehydes, polyhydroxyketones, and their derivatives after hydrolysis. Historically, carbohydrates were viewed as hydrates of carbon.

**Chemistry:** The science of substances, including their composition, structure, properties, isolation, and the reactions that change them into other substances.

**CMP-sialic acid synthetase:** An enzyme (EC 2.7.7.43) that catalyzes the formation of cytidine 5'-monophosphate-sialic acid (CMP-sialic acid) from cytidine 5'-triphosphate (CTP) and sialic acid.

**Glycan:** A generic term used for any oligosaccharide, polysaccharide, or the carbohydrate portion of a glycoconjugate.

**LARGE-SCALE SYNTHESIS OF SIALOSIDES USING GENETICALLY ENGINEERED MICROORGANISMS**

**Whole Cell Approach.** Large-scale production of Neu5Acα2–3Lac was achieved by researchers in Kyowa Kakko Kogyo Co., Ltd. in Japan (44) using permeablized [by treating cell pellets with polyoxyethylene octadecylamine (Nymeen S-215) and dimethylbenzenes (xylene)] whole cells of a Corynebacterium ammoniagenes strain and three recombinant *E. coli* strains as catalysts. UTP was produced from inexpensive orotic acid by the *C. ammoniagenes* DN510 cells, which also converted CMP to CDP. The UTP and CDP obtained were used for producing CTP by *E. coli* NM294 cells with a plasmid containing *pyrG* gene for *E. coli* K12 CTP synthetase. The CTP produced reacted with Neu5Ac to form CMP-Neu5Ac catalyzed by *E. coli* NM522 cells with a plasmid contain-
ing neuA gene for E. coli K1 CMP-Neu5Ac synthetase. The CMP-Neu5Ac and lactose were used by E. coli NM522 cells to produce a plasmid for Neisseria gonorrhoeae α2–3-sialyltransferase for the formation of 3′-sialyllactose Neu5Acα2–3Lac. The CMP byproduct of the sialyltransferase reaction can be recycled back to CDP by the Corynebacterium ammoniagenes cells. A 36% yield (0.99 g) was obtained from a 30 mL reaction carried out in a 200 mL beaker at 32 °C for 11 h, and a 44% yield (72 g) was achieved from a 2 L reaction carried out in a 5 L fermentor at 32 °C for 11 h from lactose, Neu5Ac, and orotic acid.

**Living Factory Approach.** Genetically engineered living E. coli cells have been used to produce sialosides in gram scales by the Samain group. The most significant advantage is that nucleotides, simple monosaccharides, and some sugar nucleotides (e.g., UDP-GlcNAc, UDP-Gal, UDP-Glc, GDP-Fuc) are provided by living bacterial cells’ own metabolic machinery from simple carbon and energy source such as glycerol. The method uses the β-galactosidase-negative (lacZ-) E. coli strain JM107, which is genetically engineered to delete the sialic acid aldolase gene nanA and add plasmids containing sialoside biosynthetic genes.

Early attempts achieved the synthesis of 3′-sialyllactose (Neu5Acα2–3Galβ1–4Glc) by feeding engineered (lacZ- and nanA-) living cells transformed with two plasmids containing a Neisseria meningitidis CMP-Neu5Ac synthetase and an N. meningitidis α2–3-sialyltransferase, respectively, with exogenous Neu5Ac and lactose (46). Neu5Ac was transported into the cells by permease NanT, and lactose was incorporated by β-galactoside permease LacY, both enzymes being endogenous to the E. coli host cells. By incorporating plasmids containing additional glycosyltransferase genes into the genetically engineered (lacZ- and nanA-) E. coli K12 strain JM107, the method has been used for the gram-scale synthesis of carbohydrate portions GalNAcβ1–4( Neu5Acα2–3)Galβ1–4Glc and Galβ1–3GalNAcβ1–4( Neu5Acα2–3)Galβ1–4Glc of gangliosides GM2 and GM1, respectively (46). For the synthesis of GM2 oligosaccharide, in addition to a plasmid containing a CMP-Neu5Ac synthetase gene and an α2–3-sialyltransferase gene from N. meningitidis, the plasmid pACT3cgta carrying a C. jejuni cgta gene for a β1–4GalNAc transferase and a plasmid carrying a P. aeruginosa wbpP gene for UDP-GlcNAc C4-epimerase, respectively, were introduced. GM2 oligosaccharide was produced in both the intracellular fraction and extracellular fraction from lactose and Neu5Ac after 20 h of isopropyl-1-thio-β-o-galactopyranoside (IPTG) induction. Improved efficiency in large-scale synthesis of GM2 oligosaccharide (71%) was achieved by using high cell density cultivation and an increase in the amount of lactose and Neu5Ac. By introducing the additional glycosyltransferase gene C. jejuni cgta for a β1–3GalT downstream of the cgta in the plasmid pACT3cgta in the GM2 biosynthetic cells, the engineered cells were used to produce GM1 oligosaccharide in intracellular fraction and side-products sialyllactose and GM2 in both intracellular and extracellular fractions (47).

The tolerance of β-galactoside permease LacY in transporting lactoside derivatives into E. coli cells was also explored. Lactosides with an allyl or a propargyl aglycon were readily internalized into the cell, and GM2 and GM3 ganglioside oligosaccharides were successfully synthesized using the living factory approach. In contrast, N-allyl acetamide β-lactoside could not be internalized into the cell, and an azido ethyl lactoside gave poor results (48).

To avoid adding exogenous relatively expensive Neu5Ac as one of the starting materials, 3′-sialyllactose biosynthetic E. coli K12 cells (45) were engineered further by deleting ManNAc kinase nanK gene and incorporating plasmids containing Campylobacter jejuni neuC and neuB genes encoding N-acetylgalactosamine-6-phosphate-epimerase and sialic acid synthase, respectively, for producing Neu5Ac from endogenous UDP-GlcNAc (49). Using this improved engineered bacterial strain, 3′-sialyllactose was obtained at a much higher concentration (25 g L⁻¹) compared to that obtained previously (2.6 g L⁻¹) (46). In addition, the new system does not require the addition of exogenous Neu5Ac, thus decreasing the cost of production further.

Both whole-cell-based approaches discussed above take the advantage of microorganisms’ own metabolic machinery for the synthesis of sialosides from inexpensive materials without adding nucleotides. One of the limitations of the living factory approach for large-scale synthesis of sialosides is that the glycosyltransferase acceptors that can be used are restricted to those that can be internalized by the cell. In addition, further improvements will be needed for both whole-cell-based approaches to be applicable in obtaining sialosides containing different forms of sialic acid including those contain the labile O-acetyl or O-lactyl groups.
CHEMOENZYMATIC SYNTHESIS

Chemoenzymatic synthesis combines the flexibility of chemical synthesis and the highly efficient stereo- and regioselective enzymatic approaches. It is considered one of the most effective ways to generate naturally occurring and non-natural sialosides with great diversity. Most chemoenzymatic methods involve chemical or chemoenzymatic synthesis of sialyltransferase acceptors and sialic acid derivatives or precursors followed by enzyme-catalyzed formation of sialoside products. However, an alternative chemoenzymatic method using enzymatically synthesized sialyloligosaccharides as building blocks for chemical synthesis of more complex sialosides has also been explored for the synthesis of sialyl galactosides (50, 51) and sialyl Lewis x tetrasaccharide (52, 53).

Bacterial Sialoside Biosynthetic Enzymes. Recent discovery of many bacterial sialoside biosynthetic enzymes with substrate promiscuity greatly expands the scope of applying enzymatic methods in synthesizing naturally occurring and non-natural sialosides. A key finding is that enzymes from different bacterial sources have significant diversity in tolerating substrate modifications even if they catalyze the same reaction or share sequence similarity.

Sialic acids and their derivatives are commonly synthesized from N-acetylmannosamine (ManNAc), mannoside, or their derivatives as the six-carbon precursors of Neu5Ac, Kdn, or their analogues, respectively. Although bacterial sialic acid synthases are the enzymes responsible for the formation of sialic acids in nature, a reversed reaction catalyzed by sialic acid degrading enzymes (sialic acid aldolases) has been used more commonly in the large-scale synthesis of sialic acids and their derivatives. This is because of the low cost of the pyruvate used by the sialic acid aldolase-catalyzed formation of sialic acids, compared to the phosphoenol pyruvate required by the sialic acid synthase. Other than the most commonly used recombinant E. coli sialic acid aldolase (54), a recently cloned Pasteurella multocida sialic acid aldolase has shown a higher expression level and a more promiscuous substrate tolerance than the E. coli enzyme (55).

Sialic acids and their derivatives can be activated by CMP-sialic acid synthetases (CSSs) to form CMP-sialic acids, the sugar nucleotide donors required by sialyltransferase-catalyzed reactions for the synthesis of sialosides. Among three CMP-sialic acid synthetases cloned from E. coli, Streptococcus agalactiae, and N. meningitidis, the N. meningitidis enzyme (NmCSS) is the best with regard to expression level, activity, and tolerance toward substrate modifications (54). It has become an important catalyst for the synthesis of sialosides. In addition, a heat-stable CSS has been cloned from Clostridium thermocellum (56).

Bacterial sialyltransferases are often expressed by pathogenic bacteria and are believed to be associated with bacterial virulence (3, 57–59). Bacterial sialyltransferases that have been cloned and expressed in E. coli include an N. meningitidis α2–3-sialyltransferase (Lst) (60), an N. gonorrhoeae α2–3-sialyltransferase (Lst) (60), an α2–8/2–9-polsialyltransferase from E. coli K92 (61), three sialyltransferases from Haemophilus influenzae (encoded by siaA, lic3A, and IsqB, respectively) (62, 63), a Pasteurella multocida multifunctional sialyltransferase (PmST1) (64), two sialyltransferases from C. jejuni (Cst-I and Cst-II) (58, 65), two sialyltransferases from Haemophilus ducreyi (66, 67), a Vibrio sp. JT-FJ-16 α2–3-sialyltransferase (68), and marine bacterial sialyltransferases including Photobacterium damselae α2–6-sialyltransferase (bst or Pd2,6ST) (69, 70), Photobacterium phosphoreum α2–3-sialyltransferase (71), and Photobacteria leiognathi JT-SHIZ-145 α2–6-sialyltransferase (72). Among these enzymes, many (e.g., PmST1, Pd2,6ST, and CstII) have high expression levels and have been used in the efficient synthesis of α2–3- (64), α2–6- (73), and α2–8-linked sialosides (74). Another interesting feature of bacterial sialyltransferases is that many of them have multifunctionality. For example, α2–3-sialyltransferase, α2–6-sialyltransferase, α2–3-sialidase, and α2–3-trans-sialidase activities have been identified for PmST1 (64). Pd2,6ST has α2–6-sialidase and α2–6-trans-sialidase activities (75) in addition to its α2–6-sialyltransferase activity (69, 70, 73). In addition to α2–3-sialyltransferase and α2–8-sialyltransferase activities reported before for CstII (58), α2–8-sialidase and α2–8-trans-sialidase activities

KEYWORDS

Sialic acid: An α,keto aldonic acid with a nine-carbon backbone with the core structure of neuraminic acid (Neu) or ketodeoxynululosonic acid (Kdn).
Sialic acid aldolase: Or N-acetylenuraminic pyruvate lyase (EC 4.1.3.3) is an enzyme that catalyzes the reversible reaction of breaking down (or producing) sialic acid (N-acetnylenuraminic acid) to (or from) an N-acetylmannosamine and pyruvic acid.
Sialidase: Or neuraminidase (EC 3.2.1.18) is an enzyme that catalyzes the cleavage of sialic linkages.
Sialoside: An oligosaccharide, a polysaccharide, or a glycoconjugate containing at least one sialic acid residue at the non-reducing end.
Sialyltransferase: An enzyme (EC 2.4.99.7) that catalyzes the transfer of sialic acid from cytidine 5’-monophosphate-sialic acid (CMP-sialic acid) for the formation of sialosides.

168 VOL.5 NO.2 • 163–176 • 2010 CHEN AND VARKI www.acschemicalbiology.org
have been found (65). The presence of competing sialyltransferase and sialidase/trans-sialidase activities in the same enzyme may contribute to dynamic presentation of sialic acids or other nonulosonic acids by pathogenic bacteria.

Chemoenzymatic Synthesis of Sialosides. The application of sialyltransferases and other related sialoside biosynthetic enzymes in the preparative-scale (19, 76–79) or large-scale synthesis of sialosides (80, 81) with or without in situ cofactor regeneration has been much explored in the past two decades. Engineered whole cells or enzymes entrapped in calcium pectate–silica gel beads have also been used for the synthesis of CMP-sialic acid (82). Nevertheless, systematic synthesis of a large library of sialosides containing diverse sialic acid forms was not achieved until recently via the combination of chemical synthesis of sialic acid precursors and the use of recombinant bacterial sialoside biosynthetic enzymes with promiscuous substrate specificity.

One-Pot Multi-Enzyme Approaches. Efficient one-pot multienzyme systems have been developed for the synthesis of sialosides with diverse sialic acid forms, various sialyl linkages, and a variety of underlying glycans without the purification of intermediates. Such a method has been proven very efficient in the synthesis of sialoside libraries of great diversity. To do this, different sialic acid forms or their six-carbon precursors (N-acetylmannosamine, mannose, and their derivatives) can be chemically or enzymatically synthesized as sialyltransferase substrate precursors. These compounds can then be converted by CMP-sialic acid synthetase with or without a sialic acid aldolase to CMP-sialic acid, which will be used by sialyltransferases to transfer sialic acid to acceptors for the formation of naturally occurring sialosides or sialylglycoconjugates. Depending on the different types of sialyltransferases, naturally existing α2–3-, α2–6-, or α2–8-linked sialosides can be obtained (64, 65, 73, 83, 84). Typical yields for the preparative-scale (>20 mg) one-pot multienzyme synthesis are higher than 60%, and many reactions can achieve more than 90% yields. Many sialosides synthesized have an alkyl azido aglycon that can be conveniently reduced to an amido group for efficient conjugation to proteins (85) or biotin (86) for biological studies of the importance of sialosides. Some have a p-nitrophenyl aglycon for high-throughput substrate specificity studies of sialidases (84).

So far, sialosides containing 12 naturally occurring sialic acid forms (Figure 3) have been successfully synthesized from 12 sialic acid precursors including N-acetylmannosamine (ManNAc), mannose, and their derivatives using the one-pot three-enzyme system. In addition, sialosides containing many non-natural sialic acid forms have also been obtained.

Combinatorial Chemoenzymatic Synthesis of Sialosides. The one-pot multienzyme chemoenzymatic approach described above is very efficient in obtaining structurally defined sialosides. However, the products have to be individually purified before being used in the functional studies. The purification step is tedious and time-consuming and is not necessary for initial ligand screening of sialoside-binding proteins. To avoid the product purification in generating a large library of sialosides, the Chen group developed a combinatorial chemoenzymatic approach, which can be combined directly with high-throughput screening without product purification for evaluating the ligand specificity of sialic acid-binding proteins by changing the structure of sialic acid, the type of glycosidic linkages, and the underlying glycan structures (87). The strategy was to use biotinylated sialyltransferase acceptors and carry out the one-pot multienzyme reactions in microtiter plates. A hexaethylene glycol linker was used between the glycan and the biotin in the biotinylated glycans to minimize nonspecific binding in the protein binding studies. After the completion of the enzymatic reactions, the reaction mixtures were transferred to NeutrAvidin-coated microtiter plates for protein binding assays. Each sample was plated in two sets of triplicates. One set was used for acceptor-binding protein assays to determine the reaction yields. The other set was used for sialic acid-binding protein assays. The data obtained can be adjusted by the yields obtained from the acceptor-binding protein assays to reflect the more accurate comparison of the sialoside structure-related protein binding. This approach for combined synthesis and high-throughput screening of a sialoside library to identify the preferred ligands for sialic acid-binding proteins has been demonstrated by the ligand specificity studies of Sambucus nigra agglutinin (SNA) and human Siglec 2 (CD22) using a library of 72 biotinylated α2–6-linked sialosides obtained from 18 sialic acid precursors and 4 biotinylated sialyltransferase acceptors (87).
STUDIES OF SIALIC ACID-RECOGNIZING PROTEINS

Siglecs (Sialic Acid-Binding Lectins of the Ig Superfamily). Siglecs have emerged as the largest family of Sia-binding proteins thus far found in vertebrate systems, with 16 primate genes and 8 rodent genes being recognized to date (9, 10, 88). Since their initial discovery and formal naming about a decade ago there has been increasing interest in these molecules, as evidenced by the increasing numbers of recent literature references. The signature feature of all Siglecs is a Sia-binding domain contained in the amino terminal V-set Ig like domain. This binding pocket for Sias is well recognized to include a critical arginine residue that forms a salt bridge with the carboxylate of the Sia ligand. Siglecs also generally recognize the exocyclic C7–9 side chain of Sias. Other details of binding preference are more difficult to generalize, not only because they can vary greatly between Siglecs but also because they tend to evolve rapidly, e.g., between primates and rodents (9, 10, 88). Recent evidence has confirmed the hypothesis that engagement of the Sia-binding pocket by sialoglycans can modulate the signaling function of Siglecs (9, 10, 88). Thus, understanding the nature of the natural ligands for Siglecs is well recognized to include a critical arginine residue that forms a salt bridge with the carboxylate of the Sia ligand. Siglecs also generally recognize the exocyclic C7–9 side chain of Sias. Other details of binding preference are more difficult to generalize, not only because they can vary greatly between Siglecs but also because they tend to evolve rapidly, e.g., between primates and rodents (9, 10, 88). Recent evidence has confirmed the hypothesis that engagement of the Sia-binding pocket by sialoglycans can modulate the signaling function of Siglecs (9, 10, 88). Thus, understanding the nature of the natural ligands for Siglecs (as opposed to just understanding their binding preferences on a microarray) becomes even more important. In this regard, recent evidence indicates that sulfate esters on the underlying glycan chains can also modify sialic acid recognition by the Siglecs (89–91).

Selectins. These are well-known Sia-binding proteins, which seem to only mainly require the negative charge of Sias, and even this can sometimes be replaced by sulfate ester at the same position of the underlying galactose residue (92–94). The real key to their binding properties is the presence of a Fuc residue in motifs such as sialyl Lewis x and sialyl Lewis a. Specific binding of selectins is further enhanced by sulfate esters on the C-6 position of an underlying Gal or GlcNAc residue (93). In at least one instance the binding of selectins requires not only the optimal cognate sialoglycans but also adjacent sulfate esters on the polypeptide, i.e., PSGL-1 recognition by P-Selectin (95). The goal of making synthetic ligands that can mimic and/or interrupt selection recognition function is being actively pursued, and recent evidence indicates that this may be possible in vivo.

Microbial Sialic Acid-Binding Proteins. The list of microbes, toxins, and lectins that recognize sialoglycans grows almost by the day, and a full listing will not be attempted here. Relatively soon after the discovery of Sias, it was found that the influenza virus hemagglutinin binds Sias, and later that different linkages of Sias are recognized differentially by different influenza viruses (96). There is now further evidence that several such viruses recognize not only the linkage but also the types of Sias and that these differences can determine the species preference of various viruses (97). However as with all Sia-binding phenomena, it should be kept in mind that the specifics are far more complex than a simple statement of which Sia type and linkage is preferred. Indeed, our recent work suggests that Sia recognition on cell surfaces can even be affected by nonsialylated glycans such as the ABH(O) blood groups (13), presumably because of carbohydrate–carbohydrate interactions between glycans, generating “clustered saccharide patches”. The extensive literature on the role of neuraminidase inhibitors in influenza (98) will not be reviewed here.

Sialyltransferases. Sialyltransferases (STs) (EC 2.4.99.X) are key enzymes in the biosynthesis of sialo-
sides (sialic acid-containing oligosaccharides) and sialoglycoconjugates (sialic acid-containing glycoconjugates) (12, 99). They catalyze the reaction that transfers a sialic acid (N-acetylneuraminic acid or Neu5Ac) residue from its activated sugar nucleotide donor cytidine 5'-monophosphate sialic acid (CMP-sialic acid or CMP-Neu5Ac) to an acceptor, usually a structure terminated with a galactose, an N-acetylgalactosamine (GalNAc), or another sialic acid residue. All sialyltransferases reported so far have been classified into five CAZy glycosyltransferase (GT) families including GT29, GT38, GT42, GT52, and GT80 based on their amino acid sequence similarities (CAZy, Carbohydrate-Active Enzyme database, http://www.cazy.org/) (100, 101). Most mammalian sialyltransferases belong to the GT29 family. Like other mammalian glycosyltransferases, mammalian sialyltransferases are type II membrane proteins with a transmembrane domain of 16–20 residues. They are retained in Golgi lumen, and many have a stem region of variable length (20–300 amino acids). The C-terminal catalytic domains have three major conserved sequence motifs called L, S, and VS sialylmotifs participating in binding to the sugar donor CMP-Neu5Ac, binding to both the donor and acceptor substrates, and catalysis, respectively. Recently, a new four-residue sialyl motif HY/F/W(D/E) was identified and shown to be essential for activity (102). The protein crystal structures of a porcine α2-3-sialyltransferase ST3Gal-I with or without an acceptor substrate analogue and/or CMP were newly reported (103). The structures belong to glycosyltransferase-A (GT-A) (variant 2) structure group which has a single Rossmann domain (mixed αβ fold). A conserved histidine residue (His319) in sialyl motif VS was identified as the catalytic base critical for the enzyme activity (103).

On the other hand, bacterial sialyltransferases belong to GT38, GT42, GT52, and GT80 families, and none of these display sequence homology to mammalian sialyltransferases. Crystal structures of four bacterial sialyltransferases have been reported including a multifunctional Campylobacter jejuni (OH4384) CstII (104) and a Campylobacter jejuni α2-3-sialyltransferase (Cstl) (105) both belonging to CAZy family GT42, a multifunctional sialyltransferase (PmST1) from Pasteurella multocida strain P-1059 (106, 107) and an α2-6-sialyltransferase (Psp2,6ST) from Vibrionaceae Photobacterium sp. (108), both belonging to CAZy family GT80. While the structures of both CstII and Cstl belong to glycosyltransferase-A (or glycosyltransferase-A-like) (GT-A) structure group (109) consisting of a single α/β/α sandwich resembling a Rossmann fold and a smaller lid-domain, the structures of both PmST1 and Psp2,6ST belong to glycosyltransferase-B (GT-B) structure group consisting of two Rossmann-like domains separated by a deep substrate-binding cleft (109).

Sialidases. Sialidases, or neuraminidases (EC 3.2.1.18), are sialic acid-releasing exoglycosidases that catalyze the removal of terminal sialic acids from sialoglycoconjugates in nature (110). Many pathogens express sialidases either as receptor-destroying enzymes, e.g., the influenza virus, or to release cell surface Sias, either for nutritional purposes or to uncover underlying receptors (110). Again, the extensive literature on this subject will not be reviewed here. A special class of sialidases is trans-sialidases, which catalyze the cleavage of an existing sialic acid bond and the formation of a new sialic acid bond simultaneously. Trans-sialidase from Trypanosoma cruzi has been characterized in details (111, 112). Sialidases from influenza A and B viruses belong to CAZy glycoside hydrolase family 34 (GH34), whereas most bacterial and human sialidases, Trypanosoma rangeli sialidase, as well as T. cruzi trans-sialidase are grouped into CAZy glycoside hydrolase family 33 (GH33). In addition, some bacterial sialyltransferases are multifunctional and have trans-sialidase and sialidase activities (63, 64). Endosialidases or endo-N-acetylenuraminidases (EC 3.2.1.129) have also been found from E. coli and Enterobacteriaceae phages and have been grouped into CAZy glycoside hydrolase family 58 (GH58).

NEW STRATEGIES FOR SIALOBIOLGY
Metabolic Engineering of Sialoglycoconjugates in Vitro and in Vivo. Metabolic engineering of sialoglycoconjugate in living cells and animals has become an efficient chemical glycobiology approach for exploring potential roles of sialic acids in biological systems. These have been achieved by metabolic incorporation of sialic acid derivatives and their precursors into cells, microorganism, and vertebrates and presentation as non-natural sialic acid in the glycoconjugates mainly on the cell surface (113–117). In some cases, non-natural sialic acids were used as terminators for polysialic acid chain elongation (118). In other cases, they were used to modulate and enhance the immunogenicity of the glycan-based cancer vaccines (119–122). Some of the
compounds with a bioorthogonal functional group (123) can be used as efficient probes to visualize glycan expression changes in living cells and animals (116, 124–126). A noticeable example is the azido group, which is a common functional group introduced into Neu5Ac or ManNAc to produce glycan probes N-azidoacetylneuramic acid (Neu5Az) or N-azidoacetylmannosamine (ManNAz). They are commonly per-acetylated for efficient incorporation into cells. After feeding the cells, microorganisms, or vertebrates with the per-acetylated Neu5Az or ManNAz, the resulting cell surface modified sialic acid can be selectively probed with fluorescent or other tags for imaging or further analysis (125, 127).

Glycan primers containing a hydrophobic aglycon can also be used as glycosyltransferase acceptor decoys to perturb the production of cellular glycans (128–130) and therefore are useful probes to study the functions of different glycans and glycosyltransferases.

**Sialoglycan Microarrays.** In recent years there have been many reports of glycan microarrays, in which large numbers of glycans of different structures are arrayed on slides and probed with various putative or known glycan-binding proteins (8, 10). While most of these arrays include sialylated glycans, the vast majority carry only the common human Sia Neu5Ac, with occasional representations of Neu5Gc (131, 132) and no examples of Sia O-acetylation or other modifications. Many glycans currently on arrays could potentially be modified by each of many kinds of Sias and in many possible types of linkages. Thus, if the natural diversity of Sias and their linkages were taken into account, representative arrays would have to consist of thousands of possible underlying glycans. Two such arrays are currently under study. In this regard, one practical problem is that Sia O-acetyl esters are somewhat labile and can be partially eliminated during the process of coupling and processing of slides. Thus, a given spot on an array may be presenting both the O-acetylated and non-O-acetylated versions, and of course all of these derivatives only represent C-9 O-acetylation. The more evanescent 7-O-acetyl ester that can migrate from C-7 to the C-9 (133, 134) is at the present time almost impossible to synthesize and very difficult to study in isolation (135).

**REFERENCES**


**PERSPECTIVES AND FUTURE DIRECTIONS**

This has been a brief and incomplete survey of a rapidly advancing field at the interface of chemistry and biology. The future appears bright for studies of sialic acids, especially at this interface. We need to expand and probe existing sialoside microarrays to better represent the diversity found in nature. It would also be worthwhile to generate arrays of both natural and non-natural sialo-sides. Combinations of such arrays can tell us much about the binding specificities of the many sialic acid-binding proteins that have been identified to date. The apparent roles of Neu5Gc and anti-Neu5Gc antibodies in cancer and other diseases (136, 137) also deserve further exploration, as do the roles of sialic acids as targets for various pathogens.

**Acknowledgment:** X.C. is grateful for financial support from NIH R01GM076360, NIH U01CA128442, NSF CHE0548235, and the Arnold and Mabel Beckman Foundation. X.C. is an Alfred P. Sloan Research Fellow, a Camille Dreyfus Teacher-Scholar, and a UC-Davis Chancellor’s Fellow. A.V. is grateful for financial support from NIH grant P01HL57345 and U01CA128442, and from the Mathers Foundation of New York.
11. Varki, A. (2009) Multiple changes in sialic acid biology during hu-
man evolution, Glycoconjugate J. 26, 231–245.

related genes: a phylogenetic approach, Glycobiology 15, 805–
817.

Blood 114, 3668–3676.

14. Schauer, R., Srinivasan, G. V., Coddeville, B., Zanetta, J. P., and 
in birds and reptiles and its absence in the platypus, Carbohydr. 
Res. 344, 1494–1500.

ization of sialic acid O-acetylation in group B Streptococcus, Proc. 

16. Ram, S., Sharma, A. K., Simpson, S. D., Quillen, D. P., 
site on factor H mediates serum resistance of sialylated Neisse-

Varki, A. (2009) Molecular mimicry of host sialylated glycans al-
ows a bacterial pathogen to engage neutrophil Siglec-9 and 
dampen the innate immune response, Blood 113, 3333–3336.


in technology for glycosylation with sialic acid, J. Carbohydr. 
Chem. 21, 723–768.


synthesis of sialic acid derivatives and sialyl/limetics as biological 
probes, Chem. Rev. 102, 471–490.

glycosyl donor for the preparation of glycosides of 

N-acetylgalactosamine glycosides, Org. Lett. 8, 3479–3482.

cient dehydration sialylation of C4-aminated sialyl-hemiketal do-


26. De Meo, C., Demchenko, A. V., and Boons, G. J. (2001) A stereo-
66, 5490–5497.

production of sialoside donors for use in chemical synthesis, 
Synlett 109, 1376–1379.

methyl and selective cleavage of the oxazolidinone ring, J. Org. 
Chem. 68, 4033–4035.

Lactamized sialyl acceptors for various disialoside syntheses: 
new method for the synthesis of glycan portions of Hp-s6 and 

30. Hanashima, S., Castagner, B., Esposito, D., Okada, T., and See-
berger, P. H. (2007) Synthesis of a sialic acid donor toward variant 
GM3 gangliosides by engineered E. coli, Angew. Chem., Int. Ed. 
46, 7072–7081.

production of CMP-NeuAc and sialylated oligosaccharides through 


duction of human milk oligosaccharides by metabolically engi-
neered bacteria, Glycobiology 12, 235–240.

34. Antoine, T., Priem, B., Heyraud, A., Greffe, L., Gilbert, M., Wark-
synthesis of the carbohydrate moieties of gangliosides GM1 and 
GM2 by metabolically engineered Escherichia coli, ChemBioChem 
4, 606–612.

35. Fort, S., Birliki, G., Dubois, M. P., Antoine, T., Samain, E., and 
Driguez, H. (2005) Biosynthesis of conjugatable saccharidic mo-
eties of GM2 and GM3 gangliosides by engineered E. coli, Chem. 
Commun. (Cambridge) 2558–2560.

Escherichia coli for the economical production of sialylated 

Lett. 2, 751–753.


