# Carbohydrate post-glycosylational modifications

**Hai Yu and Xi Chen***

***Received 2nd January 2007, Accepted 19th January 2007***
***First published as an Advance Article on the web 6th February 2007***
**DOI: 10.1039/b700034k**

Carbohydrate modification is a common phenomenon in nature. Many carbohydrate modifications such as some epimerization, *O*-acetylation, *O*-sulfation, *O*-methylation, *N*-deacetylation, and *N*-sulfation, take place after the formation of oligosaccharide or polysaccharide backbones. These modifications can be categorized as carbohydrate post-glycosylational modifications (PGMs). Carbohydrate PGMs further extend the complexity of the structures and the synthesis of carbohydrates and glycoconjugates. They also increase the capacity of the biological regulation that is achieved by finely tuning the structures of carbohydrates. Developing efficient methods to obtain structurally defined naturally occurring oligosaccharides, polysaccharides, and glycoconjugates with carbohydrate PGMs is essential for understanding the biological significance of carbohydrate PGMs. Combined with high-throughput screening methods, synthetic carbohydrates with PGMs are invaluable probes in structure–activity relationship studies. We illustrate here several classes of carbohydrates with PGMs and their applications. Recent progress in chemical, enzymatic, and chemoenzymatic syntheses of these carbohydrates and their derivatives are also presented.

## Introduction

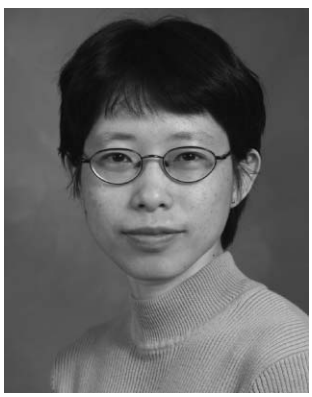
Structural modification of carbohydrates is a common phenomenon in nature. For example, sulfated carbohydrates are presented in a number of proteoglycans (glycosaminoglycans), glycoproteins (*e.g.* selectin ligand, HNK-1 epitope, and oligosaccharides of pituitary hormones *etc.*), and glycolipids (*e.g.* sulfatide and seminolipid *etc.*), and are believed to play important roles in specific molecular recognition processes.[1] Furthermore, modifications of sialic acid monosaccharides, such as sulfation, phosphorylation, methylation, acetylation, and lactylation, lead to the observation of more than 50 different sialic acid forms in nature.[2] These widely observed sialic acid modifications are also considered to be closely related to their biological functions.

In vertebrates, the sulfation of carbohydrates on proteoglycans, glycoproteins, and glycolipids, as well as varied modification of sialic acid including sulfation, methylation, some acetylation, and lactylation are believed to take place after the formation of the

*Department of Chemistry, University of California-Davis, One Shields Avenue, Davis, CA, 95616, USA. E-mail: chen@chem.ucdavis.edu; Fax: 01 530 752 8995; Tel: 01 530 754 6037*

*Dr Hai Yu received his BS in Chemistry from Liaocheng University, China in 1995. He obtained his PhD in Organic Chemistry in 2000 from Dalian Institute of Chemical Physics, Chinese Academy of Sciences, under the supervision of Professor Xiuwen Han and Professor Biao Yu, working on the synthesis of bioactive diosgenyl saponins. After a post-doctoral stay at Tulane University with Professor Harry E. Ensley working on the total synthesis of beta glucans, he joined the research group of Professor Xi Chen as a post-doctoral scholar at the University of California, Davis, in 2003. He is currently an assistant specialist. His research focuses on the chemoenzymatic syntheses of complex biologically important carbohydrates/glycoconjugates and site-specified glycoconjugation of glycoproteins.*

**Hai Yu**



**Xi Chen**

*Professor Xi Chen received her BS in Chemistry from Xiamen University in 1994 and her PhD in Biological/Organic Chemistry from Wayne State University in 2000. She worked at Neose Technologies, Inc. for two and a half years before joining the faculty in the Department of Chemistry at the University of California, Davis in August 2003. Currently, the research interests of her laboratory are: 1) developing chemoenzymatic approaches for synthesizing structurally defined complex carbohydrate-containing structures; 2) studying carbohydrate–protein interactions; 3) collaborative crystal structure studies, enzyme mechanistic studies, and mutagenesis of glycosyltransferases and other carbohydrate biosynthetic enzymes. Professor Chen has received an NSF CAREER Award and a Beckman Young Investigator Award.*

glycosidic bond. These processes can therefore, be described as carbohydrate post-glycosylational modifications (PGMs), analogous to the term of commonly known protein post-translational modifications (PTMs).

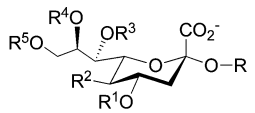
Little is known about the structure–activity relationship (SAR) of carbohydrate PGMs, due to technical difficulties in analyzing natural presence of these structures and challenges in obtaining homogeneous forms of these compounds. With currently available advanced analytical methods, the natural presence of structurally modified carbohydrates is being elucidated more precisely. Much attention has been paid to obtaining structural defined naturally occurring carbohydrates with PGMs and to elucidating the biological significance of these structures.

Chemical, enzymatic, and chemoenzymatic approaches have been developed for the synthesis of structurally defined complex carbohydrates and glycoconjugates, including those with carbohydrate PGMs. These compounds are essential probes for illustrating the structure determinants of the substrates or ligands for carbohydrate-recognizing enzymes or proteins. They are also critical to answer the fundamental questions on the biological significance of carbohydrate modifications in nature, their biosynthetic and degradation pathways in vertebrates and pathogenic bacteria, and their involvement in the normal and pathological physiology of humans and other animals.

## Sialic acids

Sialic acids (2-keto-3-deoxynonulosonic acids) are a family of negatively charged 9-carbon sugars that have been predominantly found as the outermost carbohydrate units on glycoproteins and glycolipids of vertebrates, or as components of polysaccharides in certain types of bacteria.[2]

Sialic acids have a great structural complexity in nature and more than 50 structurally distinct sialic acid forms have been found. Based on three basic sialic acid forms: *N*-acetylneuraminic acid (Neu5Ac), *N*-glycolylneuraminic acid (Neu5Gc), and deaminoneuraminc acid (KDN), single or multiple modifications can take place at the hydroxyl groups on C-4, C-5, C-7, C-8, and/or C-9 positions, including *O*-acetylation and less frequent *O*-lactylation, *O*-methylation, *O*-sulfation, and *O*-phosphorylation (Fig. 1).[2]




$R^1$ = H, or acetyl;
$R^2$ = NHAc, OH, NH-glycolyl, or their derivatives;
$R^3$ = H, or acetyl;
$R^4$ = H, acetyl, sulfate, or methyl;
$R^5$ = H, acetyl, lactyl, or phosphate;
R = molecules containing Gal(NAc) at the non-reducing end.

**Fig. 1** Naturally occurring sialosides.

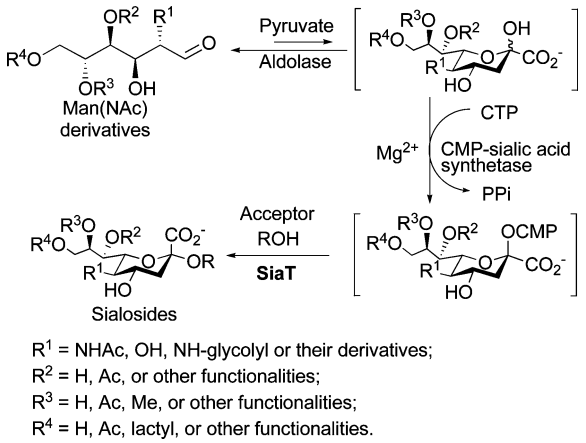
Sialic acid-containing structures play vital roles in a variety of physiological and pathological processes in vertebrates, such as cellular recognition and communication.[2] They are also believed to be important virulence factors in bacteria, used by bacteria to mimic sialylated host cell surface carbohydrate structures to evade detection and attacking by the immune defense mechanisms of the host. Modifications of sialic acids are species- and tissue-specific, they are developmentally regulated and believed to be closely related to their biological functions, such as immunogenicity, inflammation, bacterial or viral infection, tumor growth, and metastasis. For example, *O*-acetylation of bacterial polysaccharides seems to alter host immune responses[3] and may have great impact on host innate response, immunogenicity, and disease pathogenesis.[3b] Acetylation at C-9 OH group of sialic acid has been shown to enhance the activation of the alternate complement pathway of complement.[3b,4] The 9-*O*-acetylated sialic acids on the host cell surface are necessary for the binding and subsequent invasion of influenza C viruses[5] but prevent the attachment of malaria parasites[6] and that of influenza A and B viruses.[7] Mouse hepatitis virus strain S is specific to 4-*O*-acetylated Neu5Ac ($Neu4,5Ac_2$).[2a,8] Gangliosides, especially GD3, with $Neu5,9Ac_2$ as terminal sugar are considered as tumor-associated antigens or differentiation markers.[6,8,9] 4-*O*-Acetyl-GM3 has also been detected in human colon carcinomas.[10] Loss of *O*-acetylation of $sLe^x$ in human colon cancer facilitates metastasis.[2a] Modifications on sialic acids usually lead to the reduction or even resistance of sialidases or *trans*-sialidases.[11] Most sialic acid modifications reported to date have been found in vertebrates and only a few have been found in bacteria.[12]

Sialic acid modifications, such as *O*-acetylation, *O*-methylation, *O*-lactylation, and *O*-sulfation, are believed to take place after the synthesis of sialoglycoconjugates in mammalian,[3b,13] Group C *meningococci*,[14] and *E. coli*,[15] although *O*-acetylation is believed to occur on free Neu5Ac in the biosynthesis of GBS (Group B *Streptococcus* or *Streptococcus agalactiae*) capsular polysaccharide (a major virulence factor and the active principle of vaccines in phase II trials).[3b] Only a few enzymes involved in the sialic acid modifications have been discovered. For example, 9-*O*-acetyltransferase has been found in rat liver,[13] bovine submandibular gland,[16] and *E. coli* K1;[2b,17] 4-*O*-acetyltransferase has been found in guinea pig liver[18] and equine submandibular gland;[19] the enzyme catalyzing the transfer of methyl groups from S-adenosylmethionine to the C8 hydroxyl group of sialic acid (sialic acid 8-*O*-methyltransferase) has been identified in starfish *A. rubens.*[20] Recently, the gene encoding polysialic acids *O*-acetyltransferase *neuO* has been identified in *E. coli* K1.[21] A soluble sialate 9-*O*-acetyltransferase having high specificity for terminal α2,8-linked sialic acid has been cloned from *Campylobacter jejuni.*[22] Gene *neuD* in Group B *Streptococcus* (GBS) and *E. coli* K1 has also been identified to encode a sialic acid *O*-acetyltransferase which is also involved in the sialylation of capsular polysaccharides,[23] although this enzyme may acetylate free sialic acid monosaccharide[24] and the process cannot be considered as carbohydrate PGM.

As a result of the increased recognition of the biological importance of sialic acid containing structures in both eukaryotes and prokaryotes, considerable efforts have been placed on chemical, enzymatic, and chemoenzymatic synthesis of structurally defined sialic acid-containing carbohydrates and their derivatives. Earlier attempts, however, have been focused on obtaining non-natural sialic acid derivatives and a limited number of naturally occurring sialic acid-containing structures with or without sialic acid modifications (*e.g.* Neu5Ac, Neu5Gc, KDN, Neu5,9Ac, Neu5Ac9Lt, Neu5Ac8Me, and $Neu5Gc8SO_3$).[25] Recently, our group has started the systematic synthesis of libraries of α2,3- and α2,6-linked sialosides containing naturally occurring sialic

acid modifications using a highly efficient and convenient one-pot three-enzyme chemoenzymatic approach established in our laboratory.[26a,b] In this method, sialic acid modifications can be chemically or enzymatically introduced at an early stage, onto ManNAc and mannose which can be considered as six carbon precursors for sialic acids. These ManNAc and mannose derivatives can then be directly converted to naturally occurring sialosides in one-pot using three enzymes, including a sialic acid aldolase, a CMP-sialic acid synthetase, and an α2,3- or an α2,6-sialyltransferase (Scheme 1). Taking advantage of the relaxed substrate specificity of all the enzymes involved in the synthesis, the approach can be used for convenient preparation of structurally defined naturally occurring sialosides and their non-natural derivatives. ManNAc derivatives have also been synthesized and used to introduce non-natural sialic acid derivatives onto the cell surface by cultured cells and living animals.[26c–f] This metabolic engineering approach provides an efficient method to introduce chemical handles for detecting or further elaborating sialic acid-containing molecules in living systems.




**Scheme 1** One-pot three enzyme synthesis of naturally occurring sialosides with sialic acid modifications.

## Glycosaminoglycans

Glycosaminoglycans (GAGs), including hyaluronic acid (HA), chondroitin sulfate (CS), dermatan sulfate (DS), keratan sulfate (KS), and heparan sulfate (HS)/heparin (HP), are linear polysaccharides composed of modified or non-modified repeating disaccharide units containing a hexosamine.[27] They play important roles such as regulation of development, angiogenesis, axonal growth, cancer progression, microbial pathogenesis, and anticoagulation.[27b] Except for HA which contains non-modified repeating units of GlcAβ1,3GlcNAcβ1,4 disaccharide, the formation of all other GAGs involves post-glycosylational modifications (PGMs) of polysaccharide backbones formed by glycosyltransferase-catalyzed glycosylation reactions, including single or multiple sulfation and/or epimerization (Table 1).

The biosynthesis of CS involves the formation of the polysaccharide backbone containing GlcAβ1,3GalNAcβ1,4 (GlcA, glucuronic acid; GalNAc, *N*-acetylgalactosamine) repeating disaccharide units followed by sulfation at the hydroxyl group at C-4, C-6, or both C-4 and C-6 on GalNAc and/or sulfation at C-2 hydroxyl group of GlcA.[28]

The polysaccharide backbone formation of DS is the same as that for CS. An additional C-5 epimerization step after the glycosylation that epimerizes GlcA to IdoA, however, is included in the biosynthesis of DS. The epimerization also results in the linkage change from GlcAβ1,3 to IdoAα1,3. Sulfation occurs afterwards to give C-4 sulfated GalNAc with or without C-2 sulfation of IdoA.[29]

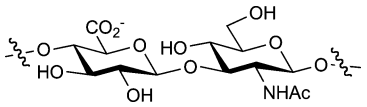
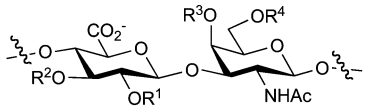
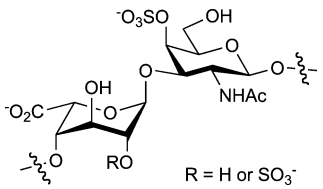
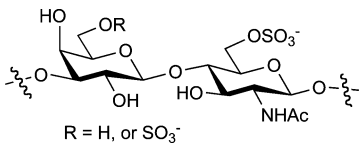
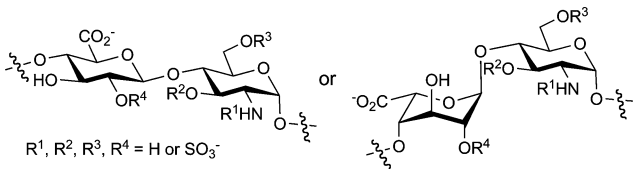
KS biosynthesis is achieved by the formation of polysaccharide backbone (poly *N*-acetyllactosamine) consisting of Galβ1,4GlcNAcβ1,3 (Gal, galactose; GlcNAc, *N*-acetylglucosamine) repeating disaccharide units, followed by single or di-sulfation to form Galβ1,4GlcNAc(6S)β1,3 or Gal(6S)β1,4GlcNAc(6S)β1,3 structures.[30]

The biosynthetic processes for heparan sulfate (HS) and heparin are essentially the same except that heparin has a higher level of sulfation per disaccharide repeating unit and a higher iduronic acid content. Their biosynthesis is complex and several enzymes are involved sequentially in modifying the polysaccharide backbone $(GlcA\beta1,4GlcNAc\alpha1,4)_n$ formed by glycosyltransferase-catalyzed glycosylations. These PGMs include: 1) *N*-deacetylation and *N*-sulfation of GlcNAc residues to form *N*-sulfoglucosamine (GlcNS) residues catalyzed by bifunctional *N*-deacetylase/*N*-sulfotransferases (four isoforms have been found); 2) C5-epimerization of GlcA residues (at the non-reducing neighbor to GlcNS) to form IdoA residues catalyzed by a GlcA C5-epimerase; 3) 2-*O*-sulfation of IdoA and GlcA residues to form IdoA(2S) and GlcA(2S) respectively catalyzed by a 2-*O*-sulfotransferase; 4) 6-*O*-sulfation of GlcNAc and GlcNS catalyzed by 6-*O*-sulfotransferases (three isoforms have been found) and occasionally 3-*O*-sulfation of GlcN and GlcNS residues catalyzed by 3-*O*-sulfotransferases (seven isoforms have been found).[31] All of these post-glycosylational modification processes in the heparan sulfate/heparin biosynthesis modify only partial residues in the polysaccharide backbone, thus lead to the complexity and heterogenecity of the produced HS and heparin molecules.

The inherent chemical heterogeneity and diversity of GAGs challenge the clear understanding of their structure–activity relationship and accurate defining of their chemical structures. Nevertheless, great progress has been made in recent years, including the development of chemical and enzymatic synthetic strategies, efficient and high-throughput assay formats, and analytic methodologies.

Extensive studies have been carried out for HS/heparin due to their important roles in regulating cancer growth, blood coagulation, inflammation, assisting viral and bacterial infections, and cell differentiation. Heparin, the most commonly used anticoagulant drug, can be considered as a special form of HS with higher levels of sulfation and iduronic acid content.[31b] Currently available evidence indicates that the interaction of HS/heparin and different biologically important proteins (*e.g.* growth factors, antithrombin) is specific and is closely related to the structure, such as sulfation pattern, sequence, as well as IdoA and sulfate contents, of the HS/heparin. For example, a pentasaccharide sequence HexA-GlcN(NS)-HexA-GlcN(NS)-IdoA(2S) (Fig. 2a) has been identified to be specific for fibroblast growth factor 2 (FGF2) binding,[32] another pentasaccharide sequence of heparin GlcN(6S)-GlcA-GlcN(NS,3S)-IdoA(2S)-GlcN(NS) (Fig. 2b) is responsible for antithrombin III (AT-III, a serine protease inhibitor that blocks thrombin and factor Xa in the coagulation cascade) binding,[33]

**Table 1** GAGs and their repeating units

| GAGs | Repeating units | Structures of repeating units |
|---|---|---|
| a) Hyaluronic acid (HA) | GlcAβ1,3GlcNAcβ1,4 |  |
| b) Chondroitin sulfate (CS) | GlcA(no S or 2S)β1,3GalNAc(4S, 6S, or 4,6$S_2$)β1,4 |  **CS-A** $R^1 = R^2 = R^4 = H$, $R^3 = SO_3^-$ **CS-E** $R^1 = R^2 = H$, $R^3 = R^4 = SO_3^-$ **CS-C** $R^1 = R^2 = R^3 = H$, $R^4 = SO_3^-$ **CS-K** $R^1 = R^4 = H$, $R^2 = R^3 = SO_3^-$ **CS-D** $R^2 = R^3 = H$, $R^1 = R^4 = SO_3^-$ |
| c) Dermatan sulfate (DS) | IdoA(no S or 2S)α1,3GalNAc(4S)β1,4 |  R = H or $SO_3^-$ |
| d) Keratan sulfate (KS) | Gal(no S or 6S)β1,4GlcNAc(6S)β1,3 |  R = H, or $SO_3^-$ |
| e) Heparan sulfate(HS)/ heparin (HP) | GlcA(no S or 2S)β1,4[or IdoA (no S or 2S)α1,4]GlcNAc(or GlcNS)(with 3S, 6S, or 3,6$S_2$)α1,4 |  $R^1$, $R^2$, $R^3$, $R^4$ = H or $SO_3^-$ |

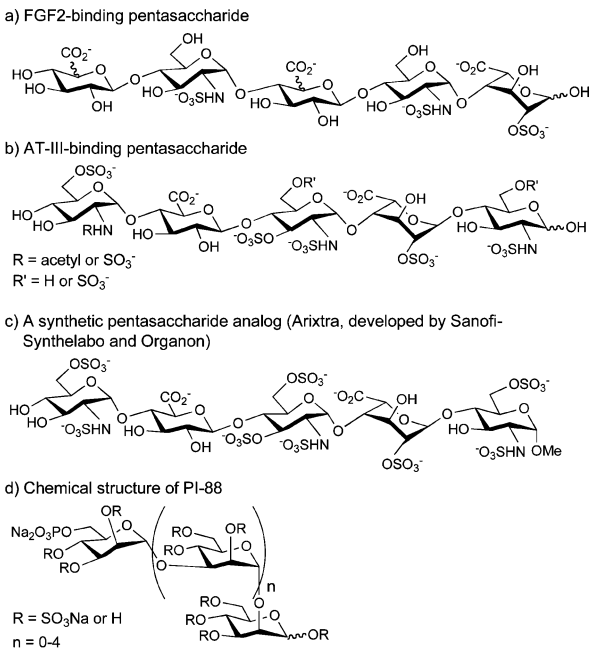



**Fig. 2** Heparin sequences of therapeutic significance.

although a longer sequence (14–20 saccharide units) is required to accelerate the AT-III–thrombin interaction and inhibit thrombin activity.[34] Based on the structure of the AT-binding pentasaccharide, a new synthetic pentasaccharide anticoagulant under the trade name of Arixtra was approved by the FDA (Fig. 2c).[33b] A synthetic hexadecasaccharide (SR123781) that displays excellent anticoagulant activity against thrombin *in vitro* and *in vivo* was produced (Fig. 3).[35] A low molecular weight heparin mimetic, PI88 (Fig. 2d), is under phase II clinical trials to treat herpes simplex viral infections.[36] It is also a potential anticancer drug that inhibits the activity of heparanase in facilitating tumor metastasis.[31b,37] Nevertheless, the detailed structures and specific sequences of HS/heparin that interact with many other proteins, such as other growth factors, chemokines, and annexins, are currently unclear.[38]

Traditional methods of purifying oligosaccharides after enzymatic digestion of GAG chains have provided useful quantities of HS/heparin oligosaccharides for early structure–activity relationship studies.[39]

Information on biochemical characterization and understanding of HS/heparin biosynthetic enzymes and the availability of their gene sequences led to the successful enzymatic synthesis of some active HS/heparin sequences using *Escherichia coli* K5 capsular polysaccharide.[40] More recently, multimilligram scale

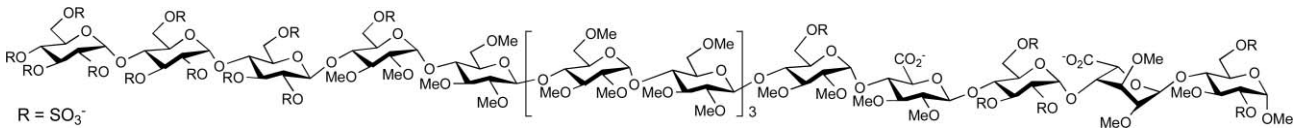


**Fig. 3** Structure of synthetic hexadecasaccharide SR123781A.

enzymatic synthesis of bioactive HSs that can bind to antithrombin, FGF2, or herpes simplex virus glycoprotein D was achieved, using completely desulfated and *N*-sulfated heparin followed by multiple *O*-sulfation catalyzed by immobilized recombinant HS *O*-sulfotransferases with the regeneration of activated sulfo donor 3′-phosphoadenosine 5′-phosphosulfate (PAPS).[41]

A chemoenzymatic approach has also been developed as an alternative for the synthesis of heparin-like polysaccharide with anticoagulant activity using purified *Escherichia coli* K5 capsular polysaccharide followed by C-5 epimerization, chemical persulfation, and subsequently selective desulfonation.[42]

Although challenging, a variety of chemical methods have been developed for the synthesis of structurally defined HS/heparin oligosaccharides and their analogs.[43] These include target-oriented synthesis,[33b,44] modular synthesis,[45] and solid-phase synthesis.[46] Libraries of chemically synthesized HS/heparin oligosaccharides have been used to generate carbohydrate microarrays for the studies of HS/heparin–protein interactions.[47] A novel di-sulfated monosaccharide IdoA(2S,4S) has been found to be able to bind FGF-1. This information may help to design novel angiogenic inhibitors with anticancer properties.[48]

Currently available methods provide the access to a range of defined HS/heparin oligosaccharide sequences, but not to all possible oligosaccharide sequences. Furthermore, the chemical synthetic yields decrease dramatically with the increase of the length of the target molecules. Obtaining defined structures longer than hexasaccharides remains to be a major challenge for chemical synthesis. Enzymatic approaches relying on polysaccharide modification inevitably lead to inherent heterogenecity of the products formed. An effective method to obtain structurally defined polysaccharides or large oligosaccharides is still lacking.

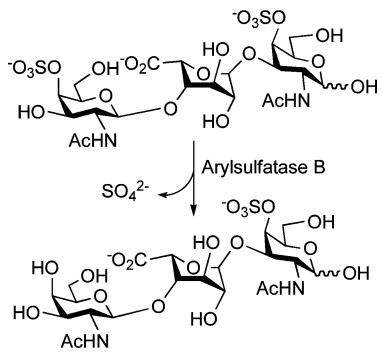
CS proteoglycans are located on cell surfaces and distributed in various human tissues, having various important biological activities such as cell migration, recognition, morphogenesis, and signal transduction. CSs have been classified into five categories (CS-A, C, D, E, K) according to the sulfation patterns of the repeating disaccharide units (Table 1b). CSs with different sulfation patterns have been known to participate in specific physiological functions. For example, the squid cartilage CS-E [GlcAβ1,3GalNAc(4S,6S)β1,4] has been found to interact with various heparin-binding growth factors.[49] CS octa or larger oligosaccharides with at least three continuous CS-E sequences have been shown to bind to type V collagen.[50]

Purified squid *N*-acetylgalactosamine 4-sulfate 6-*O*-sulfotransferase (GalNAc4S-6ST) has been used in enzymatic synthesis of CS-E from CS-A [GlcAβ1,3GalNAc(4S)β1,4], converting half of GalNAc(4S) residues to GalNAc(4S,6S) residues.[51] This enzyme can also catalyze the synthesis of oversulfated DS IdoAα1,3GalNAc(4S,6S)β1,4 units when DS IdoAα1,3GalNAc(4S)β1,4 is used as the substrate.

Chemoenzymatic synthesis of structurally defined CS has been achieved by hyaluronidase-catalyzed polymerization of synthetic sulfated disaccharide oxazoline. For example, hyaluronidase from *ovine tests* or *bovine tests* afforded synthetic CS of molecular weights ranging from 4000 to 18400 with exclusively GalNAc(4S) units.[52]

Chemical synthesis of structurally defined CS oligosaccharides has been reported.[53] More recently, a convergent synthetic approach to CS-E [GlcAβ1,3GalNAc(4S,6S)β1,4], CS-C [GlcAβ1,3GalNAc(6S)β1,4], CS-A [GlcAβ1,3GalNAc(4S)β1,4], and non-natural CS-R [GlcAβ1,3GalNAc(2S,3S)β1,4] oligosaccharides was described.[54]

DS plays a similar role as CS in cellular processes and it serves as a co-factor for several growth factors and a mediator for FGF-2 response. The hexasaccharide sequence in DS that specifically binds to heparin co-factor II has been postulated as $[IdoA(2S)\alpha1,3GalNAc(4S)\beta1,4]_3$.[55] Lately a longer nonasaccharide DS fragment containing four disacccharide sequences of IdoA(2S)α1,3GalNAc(4S)β1,4 has also been identified as the active site of DS for heparin co-factor II.[56] Enzymatic approaches using *N*-acetylgalactosamine 4-sulfatase (arylsulfatase B) have been applied in the de-sulfonation of 4S of the non-reducing terminal GalNAc residue of DS disulfated trisaccharide to afford monosulfated trisaccharide, which can be further converted to monosulfated disaccharide IdoAα1,3GalNAc(4S) using *N*-acetylhexosaminidase to provide a substrate for α-L-iduronidase (Scheme 2).[57] Chemical synthesis of DS oligosaccharides has also been reported in the last decade.[58] As L-IdoA is considered as a poor acceptor in the glycosylation with D-GalN donors, the strategy of using L-IdoA derivatives instead of L-IdoA as acceptors is often adopted in the synthesis of DS oligosaccharides.[58]




**Scheme 2** Enzymatic de-sulfation of oligosaccharide with arylsulfatase B.

Found in the cornea, various types of cartilage and brain, KS plays important roles in corneal transparency, nerve growth cone guidance, and cell adhesion.[59] Structural changes in KS are found in response to some human diseases. For instance, the lack and reduction of sulfation in corneal KS are closely associated with type I and type II macular corneal dystrophies.[60] Also, the change of sulfation contents in synovial fluid KS is clearly correlated with the severity of osteoarthritis.[61] KS differs from

other GAGs by containing a nonacidic residue. This simplifies its chemical synthesis since there is no demand of elaborating C-6 to generate a carboxylate group. Chemical synthesis of defined KS oligosaccharides has been achieved.[62] A chemoenzymatic approach has also been developed for the synthesis of KS-derived di- and tetra-saccharides by keratanase II-digestion of shark cartilage followed by subsequent chromatography and selective desulfonation.[63]

## Other carbohydrates with PGMs

Other than as important modifications on some sialic acid structures and most glycosaminoglycans, sulfation of carbohydrates has also been observed in some glycoproteins and glycolipids as PGMs. Carbohydrate structures with 3′-sulfo-galactose have been found in both *N*- and *O*-linked glycans of glycoproteins. Compared to their un-sulfated sialylated structures, 3′-sulfo-Le$^x$ Gal(3S)β1,4(Fucα1,3)GlcNAc-R and 3′-sulfo-Le$^a$ Gal(3S)β1,3(Fucα1,4)GlcNAc-R (Fig. 4a) have been shown to be more potent ligands for both L- and E-selectin.[64] Furthermore, 6-*O*-sulfo sialyl Le$^x$ Neu5Acα2,3Galβ1,4(Fucα1,3)GlcAc(6S) has been identified by staining lymph nodes with specific monoclonal antibodies G72 and G152.[65] Enzymatic syntheses of 3′-sulfated Lewis epitopes have been achieved by the recently cloned sulfotransferases including Gal3ST-2 and Gal3ST-3.[66] Several 6-*O*-sulfotransferases that transfer sulfate from PAPS to the C-6 hydroxyl group of GlcNAc have been identified.[67] Total syntheses of sulfated Le$^x$ and Le$^a$-type oligosaccharide selectin ligands have also been obtained by chemical approaches.[68] Three sulfated variants of the sialyl Le$^x$ sequence (Fig. 4b) have been chemically synthesized and used in selectin binding assays.[69] Among them, 6-sulfo-sialyl Le$^x$ Neu5Acα2,3Galβ1,4(Fucα1,3)-GlcAc(6S) is the preferred ligand for L-selectin; 6′-sulfo-sialyl Le$^x$ Neu5Acα2,3Gal(6S)β1,4(Fucα1,3)GlcAc does not show a signal of binding; and 6′,6-disulfo-sialyl Le$^x$ Neu5Acα2,3Gal(6S)-β1,4(Fucα1,3)GlcAc(6S) shows an intermediate signal of binding.[70]

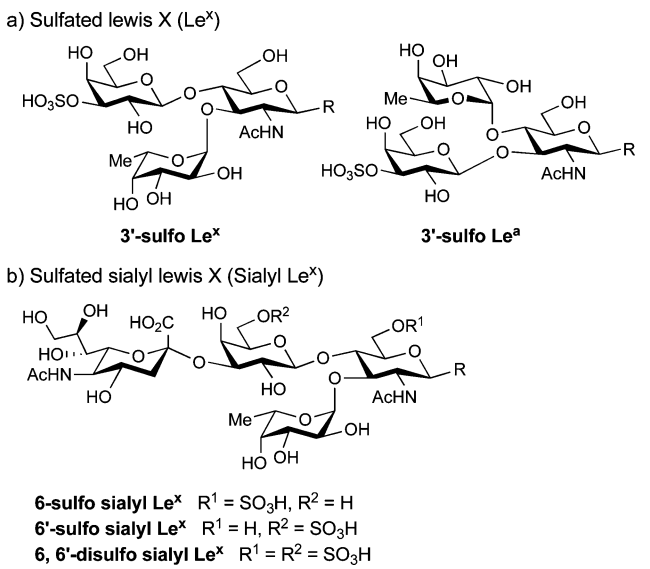



**Fig. 4** Structures of sulfated sialyl Le$^x$ and sulfated Le$^x$/Le$^a$.

Epimerization can also be considered as one type of PGM. Other than the glucuronic acid C5-epimerase activity mentioned in the biosynthesis of some glycosaminoglycans such as DS and HS/heparin, C5-mannuronan epimerase catalyzing the formation of α-L-guluronic acid residues from β-D-mannuronic acid residues is involved in the biosynthesis of linear polysaccharide alginate in brown seaweeds and certain bacteria. Both these C5-epimerization processes take place at the polymer level. A family of seven $Ca^{2+}$-dependent epimerases (AlgE1-7) has been identified in *Azotobacter vinelandii.* Different enzymes can introduce different distribution patterns of guluronic acid residues in the alginate polymer, contributing to different degrees of flexibility of the polysaccharides.[71]

## Summary and perspective

The diversity of carbohydrate structure and function is truly extraordinary. Carbohydrate post-glycosylational modifications (PGMs), including sulfation, acetylation, phosphorylation, methylation, lactylation, epimerization, and others represent a major source of carbohydrate variation and extend the complexity and the information that is controlled by carbohydrate structures. Understanding the mechanism and the significance of naturally occurring carbohydrate structural diversity may facilitate the discovery and development of new therapeutics for human diseases such as cancer, inflammatory, infectious, and others.

Significant progress has been made over the past two decades in obtaining homogeneous structurally defined carbohydrates with PGMs. Challenges still exist. Future work in this area should be devoted to further characterizing natural occurrence of carbohydrates with PGMs and developing novel synthetic strategies, including chemical, enzymatic, and chemoenzymatic approaches, for the synthesis of complex, structure-defined carbohydrates with PGMs. Assay development, including the development of a high-throughput screening method, for understanding the biological roles of carbohydrate PGMs is equally important. Furthermore, identifying, obtaining, and characterizing carbohydrate biosynthetic enzymes with flexible substrate specificity and enzymes that are responsible for PGMs will greatly facilitate the chemoenzymatic and enzymatic synthesis of naturally occurring carbohydrates with PGMs. Biochemical characterization and protein X-ray crystal structure based studies will also offer a better understanding of enzyme mechanism and supply a foundation for site-directed mutagenesis. Directed evolution is an alternative strategy to provide tailor made enzymes for chemoenzymatic and enzymatic synthesis.

## Acknowledgements

We thank the National Institutes of Health, the National Science Foundation, the Arnold and Mable Beckman Foundation, Mizutani Foundation for Glycoscience, the University of California Cancer Research Coordinating Committee, and the Regents of the University of California for financial support.

## References

1 K. Honke and N. Taniguchi, *Med. Res. Rev.*, 2002, **22**, 637.

2 (*a*) R. Schauer, *Glycoconjugate J.*, 2000, **17**, 485; (*b*) T. Angata and A. Varki, *Chem. Rev.*, 2002, **102**, 439.
3 (*a*) D. S. Berry, F. Lynn, C. H. Lee, C. E. Frasch and M. C. Bash, *Infect. Immun.*, 2002, **70**, 3707; (*b*) A. L. Lewis, V. Nizet and A. Varki, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 11123; (*c*) T. B. McNeely, J. M. Staub, C. M. Rusk, M. J. Blum and J. J. Donnelly, *Infect. Immun.*, 1998, **66**, 3705.
4 V. Sharma, M. Chatterjee, G. Sen, C. A. Kumar and C. Mandal, *Glycoconjugate J.*, 2000, **17**, 887.
5 (*a*) G. N. Rogers, G. Herrler, J. C. Paulson and H. D. Klenk, *J. Biol. Chem.*, 1986, **261**, 5947; (*b*) G. Herrler, R. Rott, H. D. Klenk, H. P. Muller, A. K. Shukla and R. Schauer, *EMBO J.*, 1985, **4**, 1503.
6 R. Schauer, H. Schmid, J. Pommerencke, M. Iwersen and G. Kohla, in *Molecular immunology of complex carbohydrates 2*, ed. A. M. Wu, Plenum, New York, 2001, pp. 325.
7 (*a*) A. Gottschalk, *Biochim. Biophys. Acta*, 1957, **23**, 645; (*b*) E. Klenk, H. Faillard and H. Lempfrid, *Hoppe-Seyler's Z. Physiol. Chem.*, 1955, **301**, 235.
8 G. Regl, A. Kaser, M. Iwersen, H. Schmid, G. Kohla, B. Strobl, U. Vilas, R. Schauer and R. Vlasak, *J. Virol.*, 1999, **73**, 4721.
9 (*a*) A. Gocht, G. Rutter and B. Kniep, *Histochem. Cell Biol.*, 1998, **110**, 217; (*b*) G. Kohla, E. Stockfleth and R. Schauer, *Neurochem. Res.*, 2002, **27**, 583.
10 H. Higashi, Y. Hirabayashi, Y. Fukui, M. Naiki, M. Matsumoto, S. Ueda and S. Kato, *Cancer Res.*, 1985, **45**, 3796.
11 (*a*) A. P. Corfield, R. W. Veh, M. Wember, J. C. Michalski and R. Schauer, *Biochem. J.*, 1981, **197**, 293; (*b*) A. Varki and S. Diaz, *Anal. Biochem.*, 1984, **137**, 236.
12 (*a*) S. Hara, Y. Takemori, M. Yamaguchi, M. Nakamura and Y. Ohkura, *Anal. Biochem.*, 1987, **164**, 138; (*b*) A. Klein, S. Diaz, I. Ferreira, G. Lamblin, P. Roussel and A. E. Manzi, *Glycobiology*, 1997, **7**, 421; (*c*) G. Reuter and R. Schauer, *Methods Enzymol.*, 1994, **230**, 168.
13 C. Butor, S. Diaz and A. Varki, *J. Biol. Chem.*, 1993, **268**, 10197.
14 W. F. Vann, T. Y. Liu and J. B. Robbins, *J. Bacteriol.*, 1978, **133**, 1300.
15 H. H. Higa, C. Butor, S. Diaz and A. Varki, *J. Biol. Chem.*, 1989, **264**, 19427.
16 V. Vandamme-Feldhaus and R. Schauer, *J. Biochem. (Tokyo)*, 1998, **124**, 111.
17 H. H. Higa and A. Varki, *J. Biol. Chem.*, 1988, **263**, 8872.
18 M. Iwersen, V. Vandamme-Feldhaus and R. Schauer, *Glycoconjugate J.*, 1998, **15**, 895.
19 J. Tiralongo, H. Schmid, R. Thun, M. Iwersen and R. Schauer, *Glycoconjugate J.*, 2000, **17**, 849.
20 A. Kelm, L. Shaw, R. Schauer and G. Reuter, *Eur. J. Biochem.*, 1998, **251**, 874.
21 E. L. Deszo, S. M. Steenbergen, D. I. Freedberg and E. R. Vimr, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 5564.
22 R. S. Houliston, H. P. Endtz, N. Yuki, J. Li, H. C. Jarrell, M. Koga, A. van Belkum, M. F. Karwaski, W. W. Wakarchuk and M. Gilbert, *J. Biol. Chem.*, 2006, **281**, 11480.
23 A. L. Lewis, M. E. Hensler, A. Varki and V. Nizet, *J. Biol. Chem.*, 2006, **281**, 11186.
24 S. M. Steenbergen, Y. C. Lee, W. F. Vann, J. Vionnet, L. F. Wright and E. R. Vimr, *J. Bacteriol.*, 2006, **188**, 6195.
25 (*a*) H. Ando, Y. Koike, S. Koizumi, H. Ishida and M. Kiso, *Angew. Chem., Int. Ed.*, 2005, **44**, 6759; (*b*) C. Auge, S. David, C. Gautheron, A. Malleron and B. Cavaye, *New J. Chem.*, 1988, **12**, 733; (*c*) G. S. Dufner, R. Schworer, B. Muller and R. R. Schmidt, *Eur. J. Org. Chem.*, 2000, 1467; (*d*) H. H. Higa and J. C. Paulson, *J. Biol. Chem.*, 1985, **260**, 8838; (*e*) M. J. Kiefel and M. von Itzstein, *Chem. Rev.*, 2002, **102**, 471; (*f*) J. L. C. Liu, G. J. Shen, Y. Ichikawa, J. F. Rutan, G. Zapata, W. F. Vann and C. H. Wong, *J. Am. Chem. Soc.*, 1992, **114**, 3901.
26 (*a*) H. Yu, H. Chokhawala, R. Karpel, H. Yu, B. Y. Wu, J. B. Zhang, Y. X. Zhang, Q. Jia and X. Chen, *J. Am. Chem. Soc.*, 2005, **127**, 17618; (*b*) H. Yu, S. S. Huang, H. Chokhawala, M. C. Sun, H. J. Zheng and X. Chen, *Angew. Chem., Int. Ed.*, 2006, **45**, 3938; (*c*) S. T. Laughlin, N. J. Agard, J. M. Baskin, I. S. Carrico, P. V. Chang, A. S. Ganguli, M. J. Hangauer, A. Lo, J. A. Prescher and C. R. Bertozzi, *Methods Enzymol.*, 2006, **415**, 230; (*d*) J. A. Prescher, D. H. Dube and C. R. Bertozzi, *Nature*, 2004, **430**, 873; (*e*) L. K. Mahal, K. J. Yarema and C. R. Bertozzi, *Science*, 1997, **276**, 1125; (*f*) E. Saxon and C. R. Bertozzi, *Science*, 2000, **287**, 2007.
27 (*a*) P. L. DeAngelis, *Glycobiology*, 2002, **12**, 9R; (*b*) R. Sasisekharan, R. Raman and V. Prabhakar, *Annu. Rev. Biomed. Eng.*, 2006, **8**, 181.
28 G. Sugumaran, M. Katsman, P. Sunthankar and R. R. Drake, *J. Biol. Chem.*, 1997, **272**, 14399.
29 H. H. Hannesson, A. Hagner-McWhirter, K. Tiedemann, U. Lindahl and A. Malmstrom, *Biochem. J.*, 1996, **313**(Pt 2), 589.
30 J. L. Funderburgh, *Glycobiology*, 2000, **10**, 951.
31 (*a*) J. D. Esko and U. Lindahl, *J. Clin. Invest.*, 2001, **108**, 169; (*b*) J. Liu and L. C. Pedersen, *Appl. Microbiol. Biotechnol.*, 2007, **74**, 263.
32 M. Maccarana, B. Casu and U. Lindahl, *J. Biol. Chem.*, 1993, **268**, 23898.
33 (*a*) W. Li, D. J. Johnson, C. T. Esmon and J. A. Huntington, *Nat. Struct. Mol. Biol.*, 2004, **11**, 857; (*b*) M. Petitou and C. A. van Boeckel, *Angew. Chem., Int. Ed.*, 2004, **43**, 3118.
34 G. M. Oosta, W. T. Gardner, D. L. Beeler and R. D. Rosenberg, *Proc. Natl. Acad. Sci. U. S. A.*, 1981, **78**, 829.
35 J. M. Herbert, J. P. Herault, A. Bernat, P. Savi, P. Schaeffer, P. A. Driguez, P. Duchaussoy and M. Petitou, *Thromb. Haemostasis*, 2001, **85**, 852.
36 K. Nyberg, M. Ekblad, T. Bergstrom, C. Freeman, C. R. Parish, V. Ferro and E. Trybala, *Antiviral Res.*, 2004, **63**, 15.
37 V. Ferro and R. Don, *Australas. Biotechnol.*, 2003, **13**, 38.
38 D. R. Coombe and W. C. Kett, *Cell. Mol. Life Sci.*, 2005, **62**, 410.
39 R. E. Hileman, A. E. Smith, T. Toida and R. J. Linhardt, *Glycobiology*, 1997, **7**, 231.
40 B. Kuberan, D. L. Beeler, M. Lech, Z. L. Wu and R. D. Rosenberg, *J. Biol. Chem.*, 2003, **278**, 52613.
41 J. Chen, F. Y. Avci, E. M. Munoz, L. M. McDowell, M. Chen, L. C. Pedersen, L. Zhang, R. J. Linhardt and J. Liu, *J. Biol. Chem.*, 2005, **280**, 42817.
42 U. Lindahl, J. P. Li, M. Kusche-Gullberg, M. Salmivirta, S. Alaranta, T. Veromaa, J. Emeis, I. Roberts, C. Taylor, P. Oreste, G. Zoppetti, A. Naggi, G. Torri and B. Casu, *J. Med. Chem.*, 2005, **48**, 349.
43 C. Noti and P. H. Seeberger, *Chem. Biol.*, 2005, **12**, 731.
44 L. Poletti and L. Lay, *Eur. J. Org. Chem.*, 2003, 2999.
45 (*a*) J. C. Lee, X. A. Lu, S. S. Kulkarni, Y. S. Wen and S. C. Hung, *J. Am. Chem. Soc.*, 2004, **126**, 476; (*b*) H. A. Orgueira, A. Bartolozzi, P. Schell, R. E. Litjens, E. R. Palmacci and P. H. Seeberger, *Chem.–Eur. J.*, 2003, **9**, 140; (*c*) D. B. Werz and P. H. Seeberger, *Chem.–Eur. J.*, 2005, **11**, 3194.
46 (*a*) C. M. DreefTromp, H. A. M. Willems, P. Westerduin, P. vanVeelen and C. A. A. van Boeckel, *Bioorg. Med. Chem. Lett.*, 1997, **7**, 1175; (*b*) R. Ojeda, O. Terenti, J. L. de Paz and M. Martin-Lomas, *Glycoconjugate J.*, 2004, **21**, 179.
47 (*a*) J. L. de Paz, C. Noti and P. H. Seeberger, *J. Am. Chem. Soc.*, 2006, **128**, 2766; (*b*) C. Noti, J. L. de Paz, L. Polito and P. H. Seeberger, *Chem.–Eur. J.*, 2006, **12**, 8664.
48 T. Karoli, L. Liu, J. K. Fairweather, E. Hammond, C. P. Li, S. Cochran, K. Bergefall, E. Trybala, R. S. Addison and V. Ferro, *J. Med. Chem.*, 2005, **48**, 8229.
49 S. S. Deepa, Y. Umehara, S. Higashiyama, N. Itoh and K. Sugahara, *J. Biol. Chem.*, 2002, **277**, 43707.
50 K. Takagaki, H. Munakata, I. Kakizaki, M. Iwafune, T. Itabashi and M. Endo, *J. Biol. Chem.*, 2002, **277**, 8882.
51 O. Habuchi, R. Moroi and S. Ohtake, *Anal. Biochem.*, 2002, **310**, 129.
52 S. Fujikawa, M. Ohmae and S. Kobayashi, *Biomacromolecules*, 2005, **6**, 2935.
53 (*a*) J. Tamura, K. W. Neumann, S. Kurono and T. Ogawa, *Carbohydr. Res.*, 1997, **305**, 43; (*b*) A. Lubineau and D. Bonnaffe, *Eur. J. Org. Chem.*, 1999, 2523; (*c*) N. Karst and J. C. Jacquinet, *Eur. J. Org. Chem.*, 2002, 815.
54 C. I. Gama, S. E. Tully, N. Sotogaku, P. M. Clark, M. Rawat, N. Vaidehi, W. A. Goddard, 3rd, A. Nishi and L. C. Hsieh-Wilson, *Nat. Chem. Biol.*, 2006, **2**, 467.
55 M. M. Maimone and D. M. Tollefsen, *J. Biol. Chem.*, 1990, **265**, 18263.
56 G. Mascellani, L. Liverani, A. Prete, G. L. Bergonzini, P. Bianchini, L. Silvestro, G. Torri, A. Bisio, M. Guerrini and B. Casu, *J. Carbohydr. Chem.*, 1995, **14**, 1165.
57 F. Dasgupta, R. I. Masada, C. M. Starr, B. Kuberan, H. O. Yang and R. J. Linhardt, *Glycoconjugate J.*, 2000, **17**, 829.
58 (*a*) F. Goto and T. Ogawa, *Bioorg. Med. Chem. Lett.*, 1994, **4**, 619; (*b*) P. Bourhis, F. Machetto, P. Duchaussoy, J. P. Herault, J. M. Mallet, J. M. Herbert, M. Petitou and P. Sinay, *Bioorg. Med. Chem. Lett.*, 1997, **7**, 2843; (*c*) N. Barroca and J. C. Jacquinet, *Carbohydr. Res.*, 2002, **337**, 673.
59 G. W. Conrad, in *Carbohydrates in Chemistry and Biology*, ed. B. Ernst, G. W. Hart and P. Sinay, Wiley-VCH, New York, 2000, pp. 717.

60 A. H. Plaas, L. A. West, E. J. Thonar, Z. A. Karcioglu, C. J. Smith, G. K. Klintworth and V. C. Hascall, *J. Biol. Chem.*, 2001, **276**, 39788.
61 H. Yamada, S. Miyauchi, M. Morita, Y. Yoshida, Y. Yoshihara, T. Kikuchi, O. Washimi, Y. Washimi, N. Terada, T. Seki and K. Fujikawa, *J. Rheumatol.*, 2000, **27**, 1721.
62 M. Kobayashi, F. Yamazaki, Y. Ito and T. Ogawa, *Carbohydr. Res.*, 1990, **201**, 51.
63 Y. Kariya, S. Watabe, H. Mochizuki, K. Imai, H. Kikuchi, K. Suzuki, M. Kyogashima and T. Ishii, *Carbohydr. Res.*, 2003, **338**, 1133.
64 C. T. Yuen, K. Bezouska, J. O'Brien, M. Stoll, R. Lemoine, A. Lubineau, M. Kiso, A. Hasegawa, N. J. Bockovich and K. C. Nicolaou, *J. Biol. Chem.*, 1994, **269**, 1595.
65 C. Mitsuoka, M. Sawada-Kasugai, K. Ando-Furui, M. Izawa, H. Nakanishi, S. Nakamura, H. Ishida, M. Kiso and R. Kannagi, *J. Biol. Chem.*, 1998, **273**, 11225.
66 (*a*) N. Ikeda, H. Eguchi, S. Nishihara, H. Narimatsu, R. Kannagi, T. Irimura, M. Ohta, H. Matsuda, N. Taniguchi and K. Honke, *J. Biol. Chem.*, 2001, **276**, 38588; (*b*) A. Suzuki, N. Hiraoka, M. Suzuki, K. Angata, A. K. Misra, J. McAuliffe, O. Hindsgaul and M. Fukuda, *J. Biol. Chem.*, 2001, **276**, 24388; (*c*) F. M. El-Fasakhany, K. Uchimura, R. Kannagi and T. Muramatsu, *J. Biol. Chem.*, 2001, **276**, 26988.
67 (*a*) A. Bistrup, S. Bhakta, J. K. Lee, Y. Y. Belov, M. D. Gunn, F. R. Zuo, C. C. Huang, R. Kannagi, S. D. Rosen and S. Hemmerich, *J. Cell Biol.*, 1999, **145**, 899; (*b*) N. Hiraoka, B. Petryniak, J. Nakayama, S. Tsuboi, M. Suzuki, J. C. Yeh, D. Izawa, T. Tanaka, M. Miyasaka, J. B. Lowe and M. Fukuda, *Immunity*, 1999, **11**, 79; (*c*) K. G. Bowman, S. Hemmerich, S. Bhakta, M. S. Singer, A. Bistrup, S. D. Rosen and C. R. Bertozzi, *Chem. Biol.*, 1998, **5**, 447; (*d*) C. L. de Graffenried, S. T. Laughlin, J. J. Kohler and C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 16715.
68 (*a*) K. C. Nicolaou, N. J. Bockovich and D. R. Carcanague, *J. Am. Chem. Soc.*, 1993, **115**, 8843; (*b*) T. Tsukida, M. Yoshida, K. Kurokawa, Y. Nakai, T. Achiha, T. Kiyoi and H. Kondo, *J. Org. Chem.*, 1997, **62**, 6876; (*c*) D. H. Live, Z. G. Wang, U. Iserloh and S. J. Danishefsky, *Org. Lett.*, 2001, **3**, 851; (*d*) M. R. Pratt and C. R. Bertozzi, *Org. Lett.*, 2004, **6**, 2345.
69 S. Komba, H. Ishida, M. Kiso and A. Hasegawa, *Bioorg. Med. Chem.*, 1996, **4**, 1833.
70 T. Feizi and C. Galustian, *Trends Biochem. Sci.*, 1999, **24**, 369.
71 S. Valla, J. Li, H. Ertesvag, T. Barbeyron and U. Lindahl, *Biochimie*, 2001, **83**, 819.