



Principles of Glycobiology

These suggestions and data are based on information we believe to be reliable. They are offered in good faith, but without guarantee, as conditions and methods of use of our products are beyond our control. We recommend that the prospective user determine the suitability of our materials and suggestions before adopting them on a commercial scale.

Suggestions for use of our products or the inclusion of descriptive material from patents and the citation of specific patents in this publication should not be understood as recommending the use of our products in violation of any patent or as permission to license to use any patents of ProZyme, Inc.

Principles of Glycobiology

Professor Akira Kobata, Tokyo Metropolitan Institute of Gerontology, Tokyo, Japan

Nucleic acids, proteins and glycans are biopolymers widely distributed in living organisms. All of these polymers are composed of covalently linked units but glycans have characteristic features not found in nucleic acids and proteins. Notably the nature of the linkage between the monomer units in glycans is more variable than that found in the other biopolymers and this leads to a much greater variety in the sequence of the biopolymer (Table 1).

The structural diversity of oligosaccharides found on glycoproteins and glycolipids is enormous. This is due to the number of parameters that are varied in a controlled way during assembly of an oligosaccharide as indicated in Figure 1. Those structural elements which are consistent with current knowledge of both mammalian oligosaccharide structure and known biosynthetic pathways are shown in Table 2.

The complexity of this linkage may be illustrated by considering the number of different ways in which one glycan monomer, galactose, may be linked to another monomer, mannose.

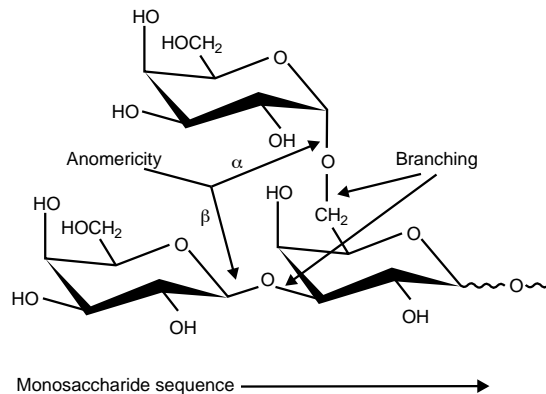


Figure 1. Origins of structural diversity.

Table 1. Number of linear oligomers of length N .

N	DNA	Proteins	Oligosaccharides*	
			O=4	O=8
1	4	20	4	8
2	16	400	128	800
3	64	8000	4096	6.40×10^4
-	-	-	-	-
-	-	-	-	-
6	4096	6.40×10^7	1.34×10^8	3.27×10^{10}
-	-	-	-	-
-	-	-	-	-
10	1.04×10^6	1.28×10^{13}	1.40×10^{14}	1.34×10^{18}

* = Number of monosaccharide types

Table 2. Monosaccharides found in eukaryotic glycoproteins.

Monosaccharide ¹	Abbreviation	Anomer ²	Attachment Point ³
N-Acetyl neuraminic acid	NeuNAc	α	3, 6, 8
N-Glycolyl neuraminic acid	NeuGc	α	3, 6
D-Galactose	Gal	α	3
		β	3, 4, 6
N-Acetyl D-glucosamine	GlcNAc	β	2, 3, 4, 6
N-Acetyl D-galactosamine	GalNAc	α	3
		β	4
D-Mannose	Man	α	2, 3, 6
		β	4
L-Fucose	Fuc	α	2, 3, 4, 6
D-Xylose	Xyl	β	2

1. Only the monosaccharides commonly found on mammalian glycoproteins are included.

2. Refers to the anomericity of the O-glycosidic bond linking the particular monosaccharide to the parent structure.

3. Refers to the location of the hydroxyl on the substituted monosaccharide to which the particular monosaccharide is linked.

Note: Please see page 7.8 for monosaccharide structures.

Principles of Glycobiology

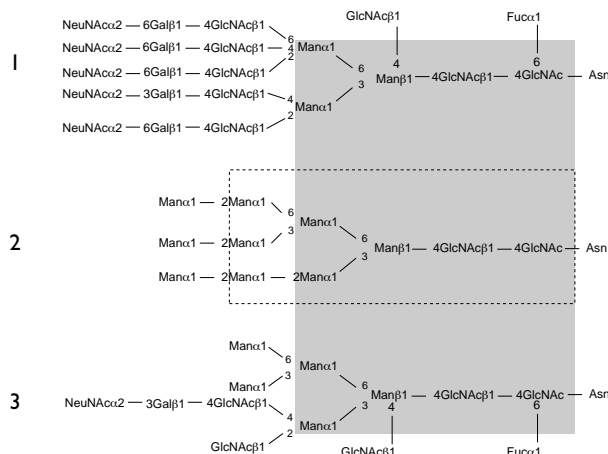


Figure 4. Three subgroups of N-linked sugar chains: (1) complex type sugar chain; (2) high mannose type sugar chain; (3) hybrid type sugar chain. The structure within the shaded box is the trimannosyl core common to all N-linked glycans. The structure enclosed with a dotted line is the common core of high mannose type sugar chains. Structures outside these lines can vary by glycan.

series of structural rules can describe them and variable regions occur in a limited part of their structures. With such rules in mind, elucidation of the glycan chain structures related to a particular biological function comes within the range of laboratory investigation.

The glycans of glycoproteins can be classified into two groups. Those that are called mucin type or O-linked glycans generally contain an N-acetylgalactosamine residue at their reducing termini. This N-acetylgalactosamine is linked to the hydroxyl group of either a serine or a threonine residue of a polypeptide. The glycans that belong to another group called asparagine-linked or N-linked glycans contain an N-acetylglucosamine residue at their reducing termini and are linked to the amide group of an asparagine residue of a polypeptide. Studies on the functional aspects of glycans reveal that they play two major roles. One is to confer particular physicochemical properties on proteins. The other is to act as signals of cell surface recognition phenomena, which are important in multicellular organisms. Generally speaking, O-linked glycans mainly work for the former function and N-linked glycans for the latter function.

N-LINKED OLIGOSACCHARIDES

All N-linked glycans contain the pentasaccharide $\text{Man}\alpha 1-6(\text{Man}\alpha 1-3)\text{Man}\beta 1-4\text{GlcNAc}\beta 1-4\text{GlcNAc}$ as a common core, which will be called the “trimannosyl core”. On the basis of the structure and the location of glycan residues added to the trimannosyl core, N-linked glycans are further classified into three subgroups as shown in Figure 4⁽¹⁾. Glycans of the complex type contain no mannose residues other than those in the trimannosyl core. Outer chains with an N-acetylglucosamine residue at their reducing termini are linked to the two α -mannosyl residues of the trimannosyl core. The presence or absence of the α -fucosyl

residue linked to the C-6 position of the proximal N-acetylglucosamine residue and the N-acetylglucosamine residue linked to the C-4 position of the β -mannosyl residue of the trimannosyl core (bisecting GlcNAc) contributes to the structural variation of complex type glycans. High mannose type glycans contain only α -mannosyl residues in addition to the trimannosyl core. A heptasaccharide with a two branch structure, $\text{Man}\alpha 1-6(\text{Man}\alpha 1-3)\text{Man}\alpha 1-6(\text{Man}\alpha 1-3)\text{Man}\beta 1-4\text{GlcNAc}\beta 1-4\text{GlcNAc}$, is commonly included in this type of glycan chain as shown by the dotted line in Figure 4⁽²⁾. The third group is called hybrid type, because the glycans have the characteristic features of both complex type and high mannose type glycans^(3,4). One or two α -mannosyl residues are linked to the $\text{Man}\alpha 1-6$ arm of the trimannosyl core as in the case of high mannose type, and the outer chains found in complex type glycans are linked to the $\text{Man}\alpha 1-3$ arm of the core. The presence or absence of the α -fucosyl residue and the bisecting GlcNAc linked to the trimannosyl core also produce structural variations of the glycans of the subgroup.

Among the three subgroups of N-linked glycans, the complex type has the largest structural variation. This variation is caused mainly by two factors. As shown in Figure 5a, from one to five outer chains are linked to the trimannosyl core by different linkages, resulting in the formation of mono-, bi-, tri-, tetra- and penta-antennary glycans⁽⁵⁾. Two isomeric triantennary glycans containing either the $\text{GlcNAc}\beta 1-4(\text{GlcNAc}\beta 1-2)\text{Man}\alpha 1-3$ group or the $\text{GlcNAc}\beta 1-6(\text{GlcNAc}\beta 1-2)\text{Man}\alpha 1-6$ group are found. These isomeric triantennary glycans are called 2,4-branched and 2,6-branched triantennary glycans, respectively. Various structures are found in the outer chain moieties of complex type glycans as shown in Figure 5b. The combination of different antennary structures and various outer chains forms a large number of different complex type glycans.

O-LINKED OLIGOSACCHARIDES

In contrast to N-linked glycans, O-linked glycans have fewer structural rules. They do not share a common core structure but are based on a number of different cores. So far they can be categorized into at least four groups according to different core structures, Figure 6. In addition, O-linked glycans with the $\text{GlcNAc}\beta 1-6\text{GalNAc}$ core and the $\text{GalNAc}\beta 1-3\text{GalNAc}$ core are found on a limited number of glycoproteins. Although the glycans are often linked to serine or threonine residues through GalNAc, the linkage may occasionally be through other residues e.g. fucose. It is also recognized that single glycans such as GlcNAc or fucose may be O-linked to the peptide backbone.

OTHER GLYCOCONJUGATES

In addition to glycoproteins there are a number of other types of glycoconjugate. The oligosaccharide moieties present on polypeptides and proteins are discrete, specific and conserved structures, but this is not always the case on these other glycoconjugates.

Principles of Glycobiology

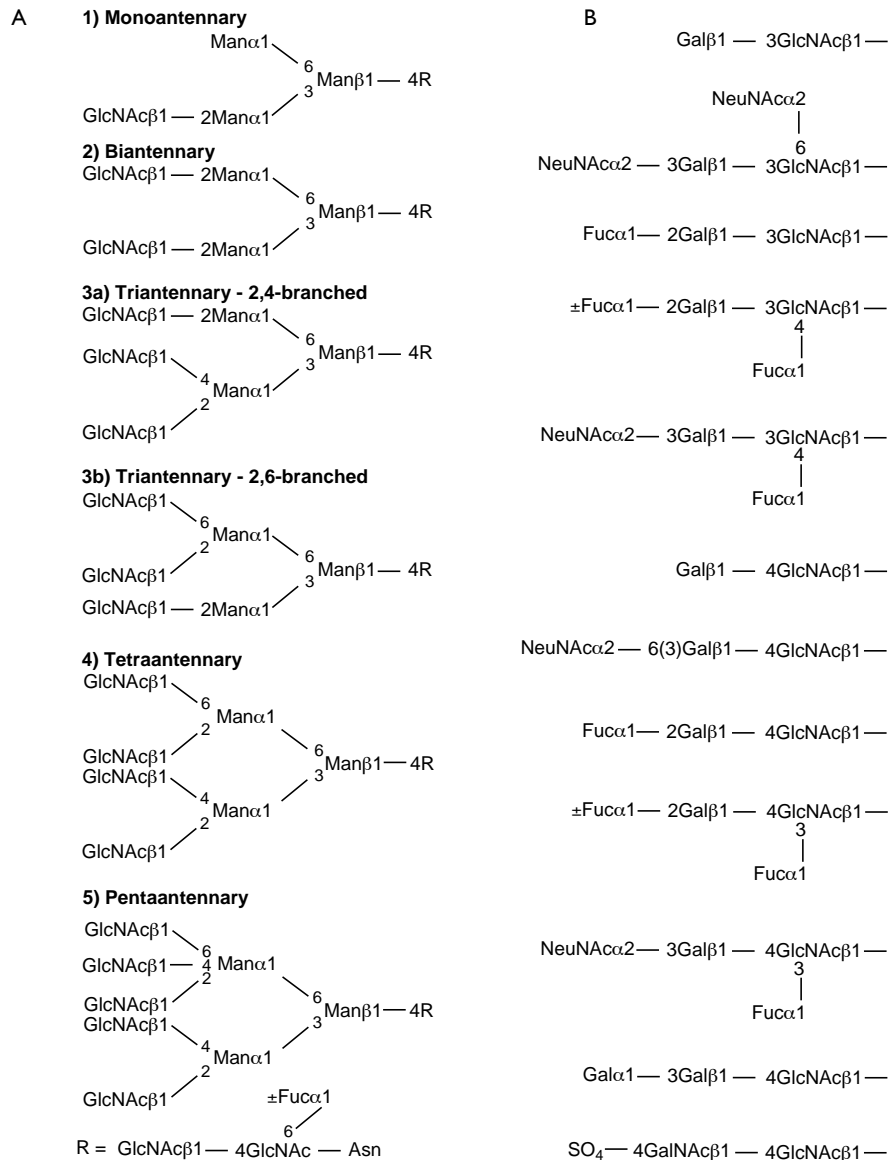


Figure 5. Two major elements that form the various structures of complex type sugar chains: (A) branching of complex type sugar chains; (B) various outer chain structures found in complex type sugar chains.

Glycolipids More than 200 glycosphingolipids, which are different in their glycan moieties, have been isolated from mammalian tissues. Most of these glycosphingolipids fall into five groups by their tetrasaccharide core structures (Figure 7). In addition to these major groups, minor groups containing trisaccharides, disaccharides and galactose as common cores are found. These glycosphingolipids are formed by a complicated biosynthetic network which has not yet been fully elucidated.

Proteoglycans Although proteoglycans should also be considered as glycoproteins in that glycans are linked to protein cores, their glycans called glycosaminoglycans are much longer (100 ~ 200 monosaccharide residues) than regular N- and O- linked glycans, and contain many anionic residues such as uronic acid and sulfate. Furthermore, their structures are unique in that their basic construction is disaccharide repeats. Based on the disaccharide structures, glycosaminoglycans have long been classified into six different groups: hyaluronic acids, chondroitin 4-sulfates, chondroitin 6-sulfates, dermatan sulfates, heparan sulfates and keratan sulfates.

Principles of Glycobiology

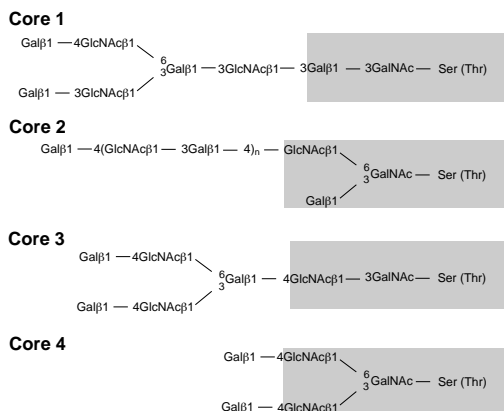
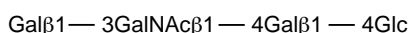


Figure 6. Four types of core structures are among those found in O-linked glycans. The cores are enclosed in the tinted box.

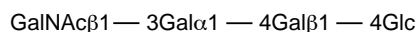
However, detailed study of the structures of glycosaminoglycans revealed that many microheterogeneities occur in the glycans and it is rather hard to discriminate chondroitin 4-sulfates, chondroitin 6-sulfates and dermatan sulfates. Accordingly, glycosaminoglycans are currently classified into four groups as shown in Figure 8.

Elucidation of the detailed structures of glycosaminoglycans is opening a new age for proteoglycan study. Heparin was found more than 70 years ago as an anti-blood coagulation material. A recent finding shows that a pentasaccharide structure in a heparin molecule binds specifically to anti-thrombin, elucidating the molecular mechanism of the anti-coagulating activity of this proteoglycan. With this finding as a turning point, many bioactive segments of glycosaminoglycans have been determined. Since blood clots can be formed on a healthy blood vessel wall, some anti-coagulation material may be included in the wall surface.

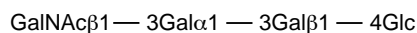
Gangliotetraose (ganglio-series)



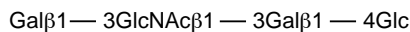
Globotetraose (globo-series)



Isoglobotetraose (isogloblo-series)



Lacto-N-tetraose (lacto-series)



Lacto-N-neotetraose (neolacto-series)

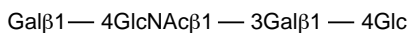
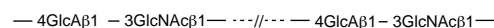


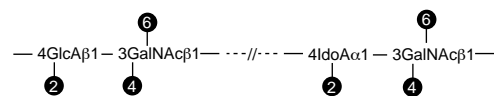
Figure 7. Structures of major tetrasaccharide cores found in animal glycosphingolipids.

Based on the observation that perfusion of blood vessels by heparinase solution prominently decreases its anti-coagulation activity, occurrence of heparin-like glycans in the endoderm has been suspected. Actually, a heparan sulfate-proteoglycan, which has a different core protein from that of heparin-proteoglycan is synthesized by the cultured endoderm obtained from bovine and rat artery. Another candidate for anti-coagulation factor is thrombomodulin found on the surface of endoderm. This protein binds to thrombin and suppresses its fibrin clotting activity. Thrombomodulin-thrombin complex also enhanced degradation of factors Va and Vlla by activating protein C. Accordingly, thrombomodulin protects the effect of thrombin overproduction. By chondroitinase ABC digestion, thrombomodulin was revealed to be a proteoglycan containing chondroitin sulfate. An interesting finding is that the activity of thrombin-dependent suppression of fibrin clotting is prominently decreased by treating thrombomodulin with chondroitinase ABC, while thrombin-dependent protein C activation is not affected. This information will be effective in developing new oligosaccharide drugs in the near future.

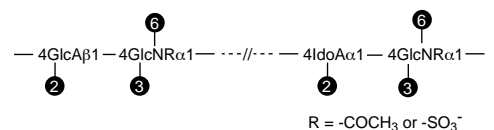
Hyaluronic acid type



Chondroitin/dermatan sulfate type



Heparan sulfate/heparin type



Keratan sulfate type

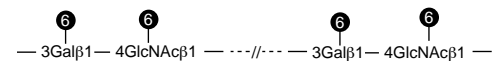


Figure 8. Classification of glycosaminoglycans and their microheterogeneities. Except for hyaluronic acid, epimerization at the C-5 position of uronic acids and N- and O- sulfation (locations indicated by circled numbers) are the sources of microheterogeneities.

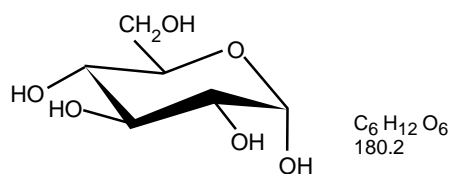
Cartilage is a tissue rich in extracellular matrix. Approximately 50% of the dry weight of cartilage is occupied by a proteoglycan called aggrecan. The structure of the core protein of aggrecan was elucidated by successful cloning of its cDNA. A lethal recessive mouse mutant, which is cartilage-matrix-deficient (cmd/cmd), is an effective model to investigate the functional role of aggrecan. The skeleton of the embryo of this mutant is strikingly different from that of age paired normal mouse embryo, and

the limbs are drastically shortened and deformed. It was found that the mutant mouse is genetically deficient in the biosynthesis of the core protein of aggrecan. When the mesenchymal cells obtained from normal mouse were cultured, they formed nodules rich in extracellular matrix molecules. The cells within the nodule are round. Culture of the mesenchymal cells of a cmd mouse also form nodules. However, these nodules are devoid of extracellular matrix and are simple clusters of cells with various types and sizes. When aggrecan is added to the culture medium, the cells from a cmd mouse form nodules similar to those observed in the normal cells cultured. Immunohistochemical observation of the apparently normalized nodules with use of anti-aggrecan and anti-type II collagen antibodies revealed that these two molecules coexist in the extracellular matrix as in the case of nodules formed by normal cells. The level of fibronectin mRNA in the cultured cmd cells was 4 ~ 8 times higher than in the cultured normal cells, and abnormal accumulation of fibronectin was observed in the nodule. This abnormal formation of fibronectin in cmd cells was suppressed to normal levels by addition of aggrecan. This data indicated that the proteoglycan is playing an essential role for the histogenesis of cartilage.

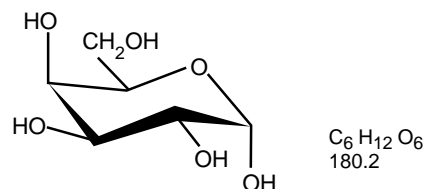
REFERENCES

1. Kobata, A. (1980), 27th International Congress of Pure and Applied Chemistry, Pergamon Press: Oxford and New York, p.185
2. Ito, S., Yamashita, K., Spiro, R.G. & Kobata, A. (1977) *J. Biochem.* Tokyo **81**, 1621
3. Tai, T., Yamashita, K., Ito, S. & Kobata, A. (1977) *J. Biol. Chem.* **252**, 6687
4. Yamashita, K., Tachibana, Y. & Kobata, A. (1978) *J. Biol. Chem.* **253**, 3862
5. Yamashita, K., Kamerling, J. P. & Kobata, A. (1982) *J. Biol. Chem.* **257**, 12609
6. Ashwell, G. & Morell, A. G. (1974) *Adv. Enzymol.* **41**, 99

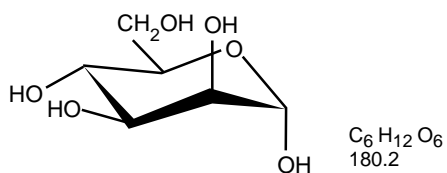
Principles of Glycobiology



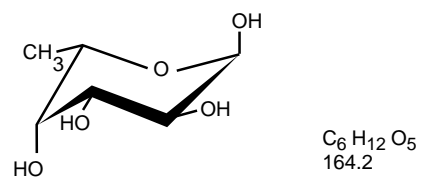
α -D-Glucose



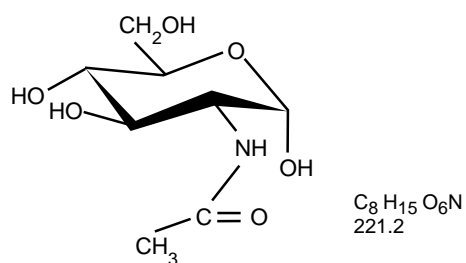
α -D-Galactose



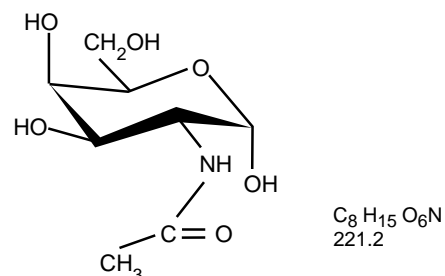
α -D-Mannose



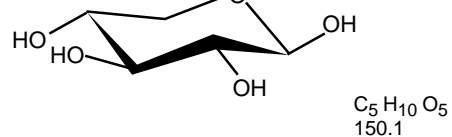
α -L-Fucose



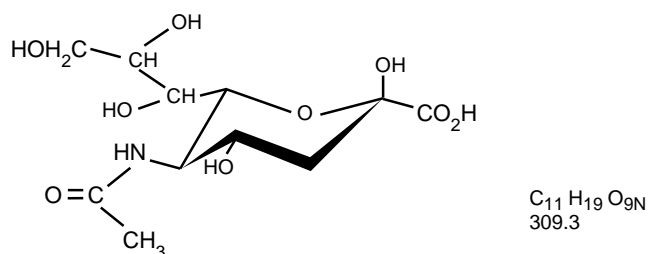
N-Acetyl- α -D-glucosamine



N-Acetyl- α -D-galactosamine



β -D-Xylose



α -N-Acetylneuraminic acid

Structures of monosaccharides commonly found in eukaryotic glycoproteins.



1933 Davis Street, Suite 207
San Leandro, CA 94577-1258

TOLL FREE (800) 457-9444
PHONE (510) 638-6900
FAX (510) 638-6919

E-MAIL info@prozyme.com
WEB www.prozyme.com