

Principle of Carbohydrate Analysis

BCMB 8020

April 9th 2009

Lianchun Wang

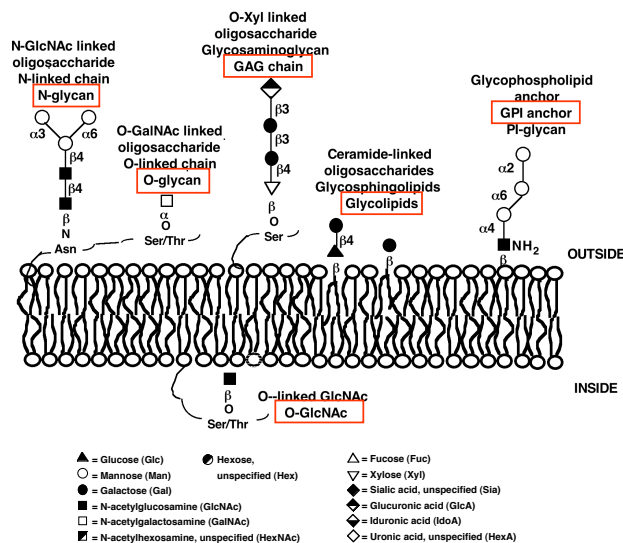
1

Outline

- I. Class of carbohydrate (types of glycosylation)
- II. Why study carbohydrate
- III. Approaches for carbohydrate analysis
- IV. Sample paper

2

I. Species of Carbohydrate Moieties



3

II. Roles of Glycoprotein-associated Carbohydrates

1. Quality Control/folding. (deglycosylation/reglycosylation)
 - glycosyltransferase
2. Solubility
 - Peroxidase
3. Circulating half-life
 - Lutropin (LH), follicle-stimulating hormone (FSH)
4. Cell-cell interactions
 - lymphocyte homing, cell growth, tumor metastasis

4

Significance

Fully understanding the biology of glycoconjugates requires in-depth knowledge of the carbohydrate chains

5

III. Experimental approaches for carbohydrate analysis

Table 1. Primary structural features of a complex carbohydrate

General description	Specific examples
Qualitative and quantitative composition, i.e., nature and number of constituting monosaccharides including absolute configuration (D- or L-) and ring size (pyranose [p] or furanose [f])	Gal, GlcNAc, Man, Fuc in ratio 2:4:3:1 D-Gal, D-GlcNAc, D-Man, L-Fuc Man _p , Gal _f
Positions of glycosidic linkages	α , β
Sequence of monosaccharides, including occurrence of branchpoints (double or triple substitution of a monosaccharide)	Man α (1-6)[Man α (1-3)]Man β (1-4)GlcNAc
Nature, number, and location of appended noncarbohydrate groups: phosphate, sulfate, acetate (peptide, lipid)	IdoA2SO ₃ , Neu5Ac9OAc

6

III. Experimental approaches for carbohydrate analysis

Table 2. Secondary structural features of a complex carbohydrate

Conformational aspect	Example
Precise ring conformation of each monosaccharide	Complete set of H-C-C'-H dihedral angles
Orientation of monosaccharides with respect to each other	Torsional angles ϕ , ψ (ω) around glycosidic bonds and/or interatomic distances
Flexibility of the spatial structure	Dynamics parameters (rotational correlation times, order parameters)

🍏 For most carbohydrates, the secondary and higher-order structures in solution are not readily defined, due to their inherent flexibility.

🍏 The secondary and higher-order structure analysis of carbohydrates is not discussed in further detail in this lecture. Instead, the analytical methods used to determine composition and sequence of carbohydrates are discussed

7

Carbohydrate Analysis Offers Unique Challenges

1. Branched
2. Synthesis is not "template driven"
3. Alternative linkage positions are possible
4. Alternative anomeric configurations are possible
5. Cell-type specific glycosylation
6. Influence of environmental conditions
[Glucose] [NH₃] pH
7. Site-specific glycosylation
8. Microheterogeneity

8

III. General Considerations for Analyzing the Primary Structure of a Carbohydrate

- Due to enormous structural diversity of naturally occurring glycans, their structural analysis requires a flexible approach
- The choice of methodology and the final result expected are dictated by the amount and purity of carbohydrate material available

9

III. General Considerations for Analyzing the Primary Structure of a Carbohydrate

Information is expected to obtain through serial carbohydrate analysis

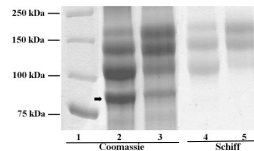
- Presence of carbohydrate moiety ?
- Class of carbohydrate ?
- Quantity and monosaccharide composition ?
- More detailed analysis: linkage and sequence ?

10

IIIA. Presence of Carbohydrates ?

Chemical Reaction

Periodic acid-Schiff (PAS) reaction: Based on the susceptibility of sugars to periodate oxidation. Sensitivity: 5-10 ng of glycoprotein



SDS-PAGE analysis of the S-layer glycoprotein of *G. stearothermophilus* NRS 2004/3a. Lane 1, molecular mass standard; lanes 2 & 4, S-layer glycoprotein from continuous culture; lanes 3 and 5, S-layer glycoprotein from batch culture; lanes 2 and 3, Coomassie blue staining; lanes 4 and 5, periodic acid-Schiff staining. Steiner K. et al. *J. Bacteriol.* 2006

Radioactive labeling

- Proteins metabolically labeled with radioactive sugar precursors. After purification, the presence of carbohydrate moiety can be detected by radioactivity. Sometimes, the ability to label proteins with specific radioactive precursors may suggest the presence of specific types of glycoprotein or carbohydrate structure.
- Label endogenous glycan by radioactive sugar nucleotides by glycosyltransferase

Lectin binding

Lectins are carbohydrate-binding proteins. Sensitivity: 5-10 ng. Based on the specificity, the lectin binding assay may also facilitate to determine the type of carbohydrate

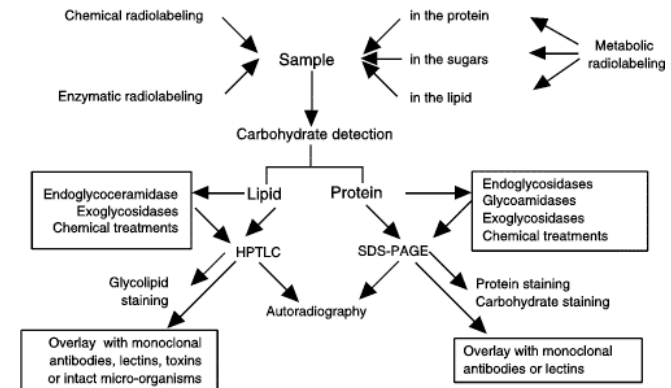
Enzyme digestion

Carbohydrate degrading enzymes

11

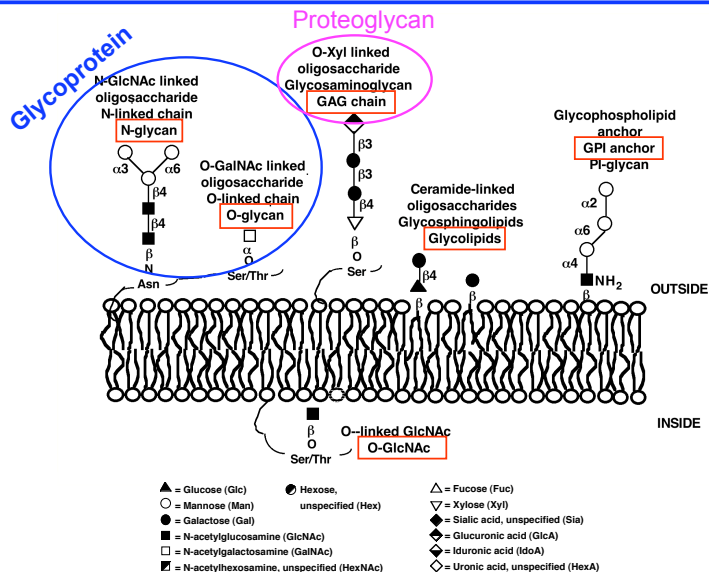
IIIA. Presence of Carbohydrates ?

Basic Strategies for the detection of carbohydrates in glycoconjugates



12

IIIB. Class of Carbohydrate ?



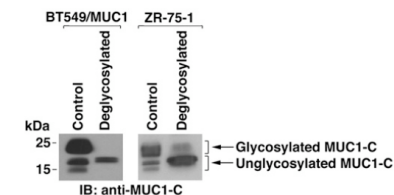
Glycoproteins and Proteoglycans ?

🍏 **Glycoprotein** carbohydrate moieties contain N-glycan and O-glycans.

🍏 **Proteoglycan** carbohydrate moieties contain glycosaminoglycan (heparan sulfate, chondroitin sulfate, dermatan sulfate, keratan sulfate and hyaluronan).

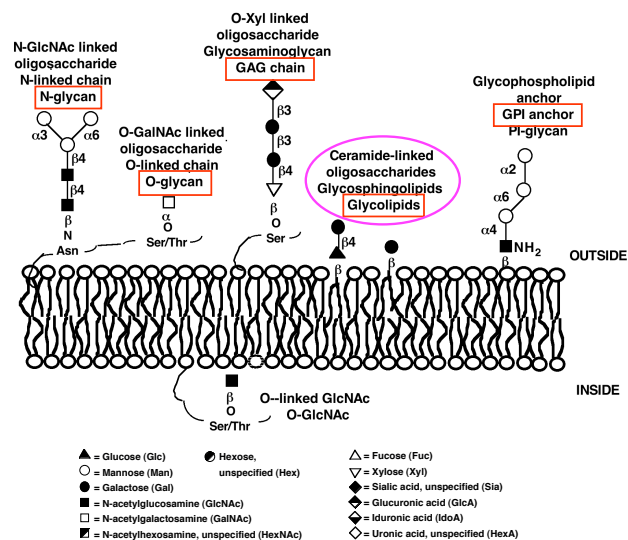
🍏 A glycosylated protein typically presents diffused or multiple bands at electrophoresis gel due to heterogeneity of carbohydrate moiety.

- Carbohydrate staining (Periodate-Schiff assay, PAS)
- Degrade carbohydrate by enzymes, leading to mobility change



Lysates from tumor cells were immunoprecipitated with anti-MUC1. The precipitated MUC1 were left untreated or digested with N-glycosidase and then immunoblotted. Ramasamy S., et al. *Molecular Cell* 27, 992-1004 (2007).

Class of Carbohydrate ?



Presence of Carbohydrates: Glycolipids

Glycolipid (glycosphingolipid): an oligosaccharide attaches via glucose or galactose to the terminal primary hydroxyl group of the lipid moiety ceramide

- Fractionation by Thin-layer Chromatography (TLC)
- Low abundant fractions need prepurification
- Carbohydrate staining (PAS)
- Degrade carbohydrate by enzymes or chemical method leading to mobility change
- Detect with monoclonal antibody, lectin and others.

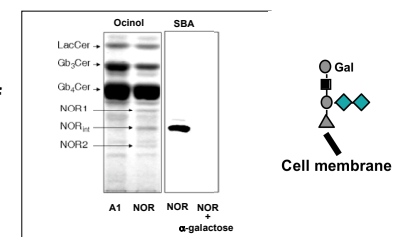
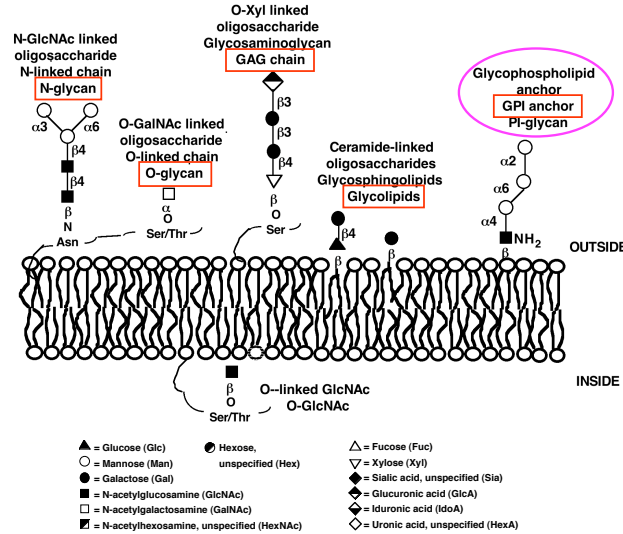


Fig. 1. Characterization of NOR2 glycolipid of erythrocytes on the HPTLC plates. Left panel: orcinol staining. Lanes A1 and NOR, total neutral glycolipid from control A1 and NOR erythrocytes, respectively. Right panel: SBA staining. 1, purified NOR2; 2, NOR2 treated with α -galactosidase. Duk M., et al. *Glycobiology* 17:304-12, 2007

HPTLC: high-performance thin-layer chromatography Soybean agglutinin (SBA): GalNAc/Gal-specific

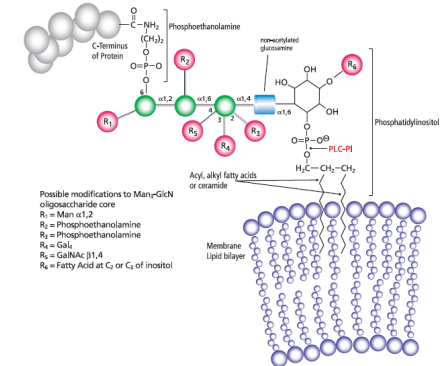
Class of Carbohydrate ?



17

Presence of Carbohydrates: Glycophospholipid (GPI) Anchor

GPI anchor: a glycan bridge between phosphatidylinositol and a phosphoethanolamine in amide linkage to the carboxyl terminus of a protein

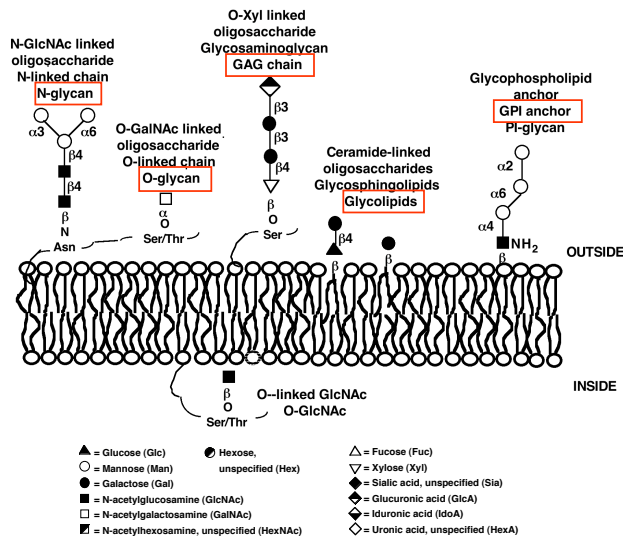


- Triton X-114 extraction at low temperature and partition at warm temperature: GPI anchor and other amphiphilic proteins associated with the detergent-enriched phase.

- GPI-specific phospholipases alter the partitioning between the phases and motility shift in electrophoresis.

18

Class of Carbohydrate ?



19

IIIC. Quantity and Composition of Carbohydrate

🍏 **Total carbohydrate content:** Colorimetric reactions to determine the total amount of hexose, hexuronic acid, or hexosamine

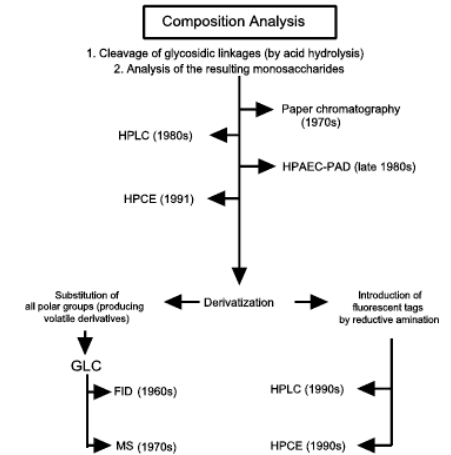
🍏 **Composition analysis:**

- Application:

1. Abundance of carbohydrate moiety in glycoconjugates
2. Precise molar ratio of individual monosaccharides, suggesting for the presence of specific oligosaccharide class

- Experimental steps :

1. Cleavage glycosidic linkage
2. Fractionation of resulting monosaccharide
3. Detection of each monosaccharide
4. Quantification



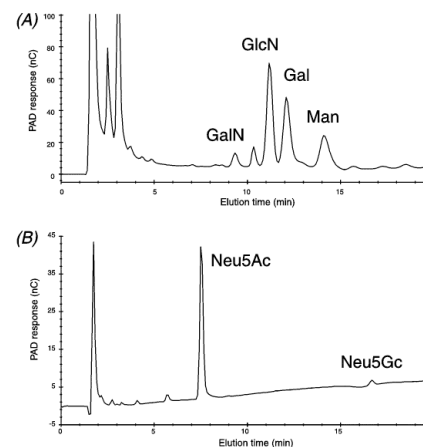
Methods for Composition Analysis of Carbohydrate

Method	Derivatization	Instrumentation and materials required	Limit of detection	Expertise needed	Information obtained
GLC-FID	Complete derivatization to produce volatile compounds	GLC instrument with FID	nmole	Medium	Type and quantity of monosaccharides
	Derivatization with chiral aglycone	Capillary column			Absolute configuration(d/l)
GLC-MS	Complete derivatization to produce volatile compounds	GLC instrument coupled to a MS with EI source Capillary column	10 pmoles < 100 fmole	High	Type and quantity of monosaccharides
HPLC	Fluorescent tagging of the reducing end	HPLC instrument with on-line fluorescence detector HPLC column	fmole-pmole	Medium	Type and quantity of monosaccharides
HPAEC-PAD	Not needed	HPLC instrument coupled to PAD CarboPac PA-1 or PA-10 column	5 – 50 pmoles	Medium	Type and quantity of monosaccharides

(EI) Electron impact; (FID) flame ionization detection; (GLC) gas-liquid chromatography; (HPAEC) high pH anion-exchange chromatography; (MS) mass spectrometry; (PAD) pulsed amperometric detection.

21

Monosaccharide composition analysis of bovine fetuin by HPAEC-PAD



(A) Neutral and amino sugars; (B) Sialic acids. Two 50- μ g samples of fetuin were submitted to strong (A) and mild (B) acid hydrolysis, respectively.

22

IIID. Detailed Characterization of Carbohydrate

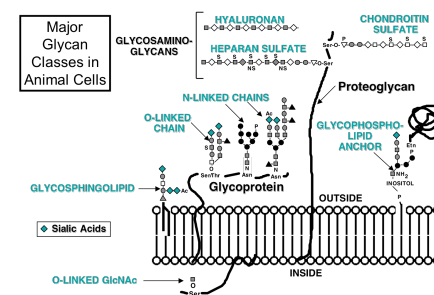
Analytical procedures:

1. Release
2. Profiling: number, relative quantities, and type of oligosaccharide structure present
3. Linkage
4. Sequence

23

IIID.1. Release of Oligosaccharide from Glycoconjugates

- Glycosphingolipid:**
 - enzyme (endoglyceramidase)
 - chemical treatment (ozonolysis)
- Glycoprotein:**
 - Hydrazinolysis (O- and N-glycan)
 - β -elimination (alkaline borohydride treatment, O-glycan)
 - PNGase F or PNGase A (N-glycan)
 - Endo H (high mannose and hybrid type structure)
- GPI anchor:**
 - Phospholipases, proteolysis, deacylation and dephosphorylation treatment
- Proteoglycan:**
 - nitrous acid deamination
 - enzymatic lyase

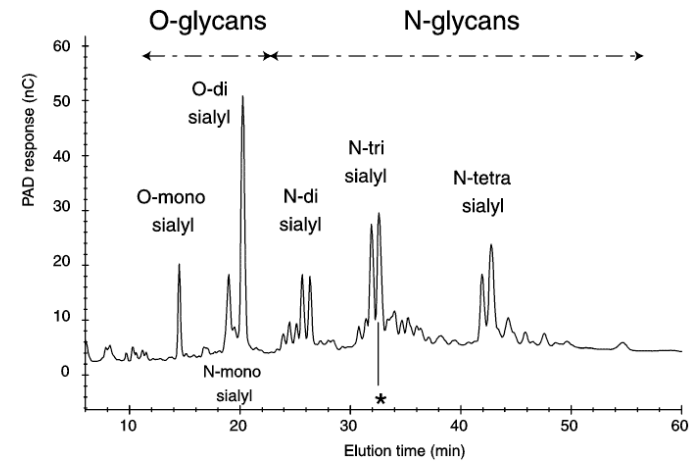


24

IIID.2. Profiling Released Oligosaccharide

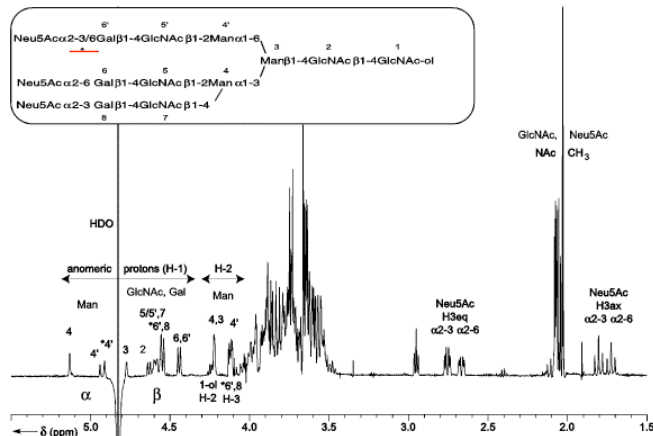
- 🍏 Microheterogeneity:
 - Different glycan class
 - Same class of glycan may be structurally different
- 🍏 Chromatographic profiles are the major readout
 - the number
 - relative quantities
 - type of oligosaccharides
- 🍏 Profiling strategies:
 - chosen based on the quantity of sample available
 - tagging with radioactive isotope or fluorescence dye to increase assay sensitivity
 - If a sufficient quantity of individual glycans available, physicochemical approaches (MS and NMR) can be applied to completely delineate the structure

25



HPAEC-PAD profile of the N- and O-glycans obtained from bovine fetuin (250 μ g) by automated hydrazinolysis. The glycan mixture was dissolved in water, injected onto a CarboPac PA-1 anion-exchange HPLC column, and eluted with a linear gradient (20–250 mM) of sodium acetate in 100 mM sodium hydroxide. The elution positions of the various glycans correspond to mono- and disialyl O-glycans and a range of mono-, di-, tri- and tetrasialyl N-glycans. Within a group of glycans with the same charge, differences in elution times reflect different degrees of branching, and/or a different linkage position and/or branch location of a sialic acid and/or galactose residue. (The glycan fraction marked by an asterisk was reduced with borohydride and analyzed by ^1H NMR spectroscopy in the following talking)

26



^1H NMR spectrum of a mixture of two trisialyl triantennary N-type oligosaccharide-alditols obtained from bovine fetuin by hydrazinolysis, followed by purification on HPAEC and subsequent reduction with sodium borohydride. The spectrum was recorded at 500 MHz, using a solution of 100 μ g of the glycan mixture in 0.7 ml of D_2O in a 5-mm NMR probe at pH 6.5 and 23°C. The structures of the two glycans are given in the inset. Please note that they differ only in the linkage position of sialic acid to galactose in the $\text{Man}\alpha 1-6$ branch. The numbers in the spectrum refer to the corresponding residues in the structures. Assignments are given for the structural-reporter group signals, including those of the anomeric protons (H-1 signals), the Man H-2, Gal H-3, Neu5Ac H-3eq and H-3ax, and N-acetyl amino sugar methyl signals; signals marked by an asterisk (for residues 4' and 6') refer to the component with Neu5Ac in $\alpha 2-3$ linkage to Gal-6'. The H-1 signals to the left of the residual water (HDO) peak are indicative of the presence of α -linked monosaccharides in the glycan(s), whereas those to the right of HDO originate from β -linked monosaccharide residues in the structure(s).

IIID.3. Linkage Analysis

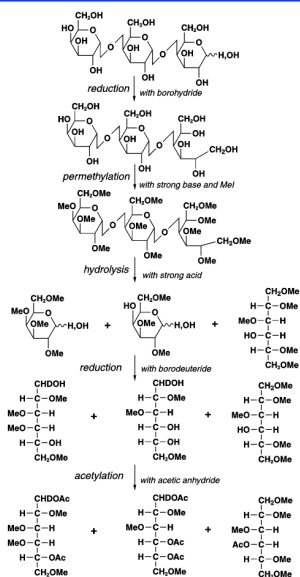
Approaches

Linkage position: Methylation analysis

Linkage anomericity: Specific enzyme & NMR

28

Linkage Position: Methylation



Principle:

1. Introduce on each free -OH group of the native oligosaccharide a stable substituent (methyl group)
2. Leave glycosidic linkage, producing individual monosaccharide residues with new free -OH that appear at the positions that were previously involved in a linkage
3. Derivatize to be volatile molecule and analyzed by GLC-MS

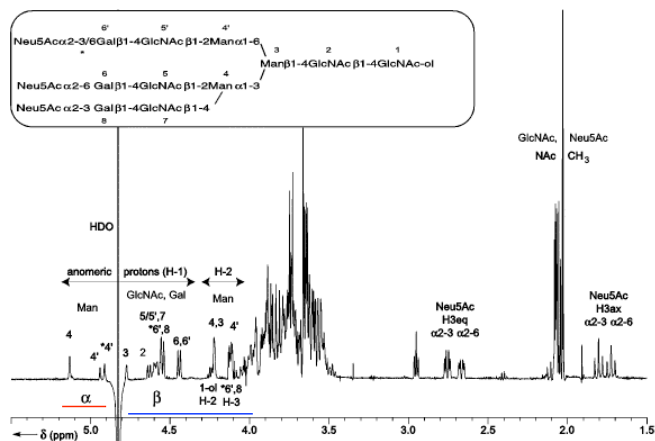
🍏 Methylation analysis could not indicate how residues are attached each other and the linkage configuration

29

Linkage Anomerism

- 🍏 Sequential exoglycosidase digestion, e.g. α - or β -exoglycosidase
- 🍏 Specific endoglycosidase
- 🍏 NMR profiling (if sufficient quantities, ^1H NMR is more reliable)

30



^1H NMR spectrum of a mixture of two trisialyl triantennary N-type oligosaccharide-alditols obtained from bovine fetuin by hydrazinolysis, followed by purification on HPAEC and subsequent reduction with sodium borohydride. The spectrum was recorded at 500 MHz, using a solution of 100 μg of the glycan mixture in 0.7 ml of D_2O in a 5-mm NMR probe at pH 6.5 and 23°C. The structures of the two glycans are given in the inset. Please note that they differ only in the linkage position of sialic acid to galactose in the $\text{Man}\alpha 1-6$ branch. The numbers in the spectrum refer to the corresponding residues in the structures. Assignments are given for the structural-reporter group signals, including those of the anomeric protons (H-1 signals), the Man H-2, Gal H-3, Neu5Ac H-3eq $\alpha 2-3$ $\alpha 2-6$, and N-acetyl amino sugar methyl signals; signals marked by an asterisk (for residues 4' and 6') refer to the component with Neu5Ac in $\alpha 2-3$ linkage to Gal-6'. The H-1 signals to the left of the residual water (HDO) peak are indicative of the presence of α -linked monosaccharides in the $^3\text{1}$ glycan(s), whereas those to the right of HDO originate from β -linked monosaccharide residues in the structure(s).

IIID.4. Sequence Analysis

Methods for Carbohydrate Sequence Analysis

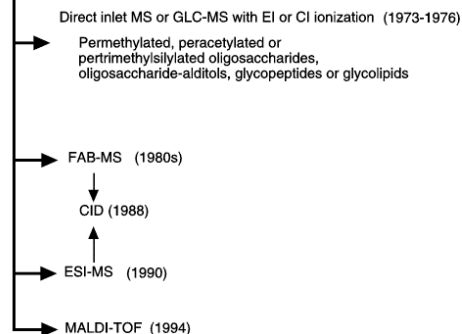
Method	Derivatization	Instrumentation and materials required	Limit of detection	Expertise needed	Information obtained
MS	May or may not be needed depending on quantity available	MS with MALDI source and TOF analyzer	pmole	High	Sequence deduced from the mass decrements produced by exoglycosidases
	Complete derivatization and/or tagging	MS with FAB, LSIMS, or ESI source and MS/MS capabilities	pmoles for ESI	Very high	Sequence deduced from the fragmentation patterns of daughter ions through several generations
NMR	Not necessary	NMR spectrometer	nmoles	High	Sequence

(MALDI) Matrix-assisted laser desorption ionization; (TOF) time-of-flight.

32

MS approaches to analyze intact, underivatized or derivatized oligosaccharades

Glycan sequencing by mass spectrometry

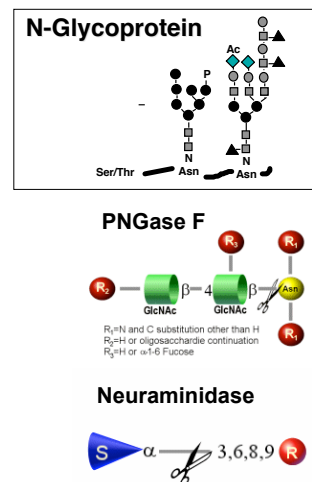
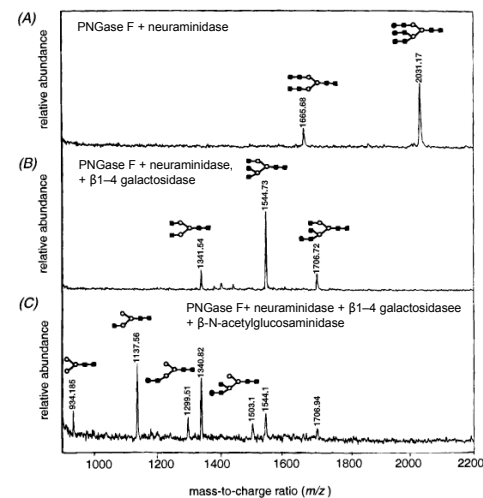


33

III. General Considerations for Analyzing the Primary Structure of a Carbohydrate

- Carbohydrate structure analysis requires a flexible approach
- The choice of methodology and the final result expected are dictated by the amount and purity of carbohydrate material available
- Primary structure analysis is expected to know: the presence, content, relative quantity, class, linkage and sequence of carbohydrate

35



MALDI-TOF mass spectra of N-glycan mixtures obtained from 5 μ g of bovine fetuin after treatment at pH 7.5 with (A) PNGase F and neuraminidase; (B) PNGase F, neuraminidase, and β 1-4 galactosidase; (C) PNGase F, neuraminidase, β 1-4 galactosidase, and β -N-acetylglucosaminidase. The m/z values measured for the various peaks are compatible with the monosodium adducts of oligosaccharides with the given schematic structures.

IV. Sample Paper



Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Analytical Biochemistry 354 (2006) 43–53

ANALYTICAL
BIOCHEMISTRY

www.elsevier.com/locate/yabio

Comprehensive glycan analysis of recombinant *Aspergillus niger* endo-polygalacturonase C

Bryan Woosley¹, Min Xie^{1,2}, Lance Wells, Ron Orlando, Derek Garrison³, Daniel King³, Carl Bergmann^{*}

Complex Carbohydrate Research Center and Department of Chemistry, University of Georgia, Athens, GA 30602, USA

36

I. Is endo-polygalacturonase C (PGC) glycosylated ?

Approach 1: Potential glycosylation sites ?

** * * *	11	21	31	41
ATTC TfSG SE	GASKASKSKT	SCSTIYLSDV	AVPSGTTLDL	SDLNDGTHVI
FQGETTFGYE	EWEGPLVRV S *	GTDITVEGES	DAVLNGDGR	WWDGEGGNGG
KTKPKFFYAH	DL S S T IKSI	YIENSPQVF	SIDGSTDLTM	TDITVDNTDG
DTDDLAANTD	GFDIGESTYI	TITGAEIYNQ	DDCVAINSGE	NIYFSASVCS
GGHGLSIGSV	GGRDDNTVK N @	VTFYDENVLK	SQQAIRIKTI	YGDTGSVSEV
TYHEIAFSDA	TDYGIVIEQN	YDDTSKTPTT	GVPITDFVLE	NIVGTCEDDD
CTEVYIACGD	GSCSDWTWTG	VSVTGGVSD	DCLNVPSGIS	CDL

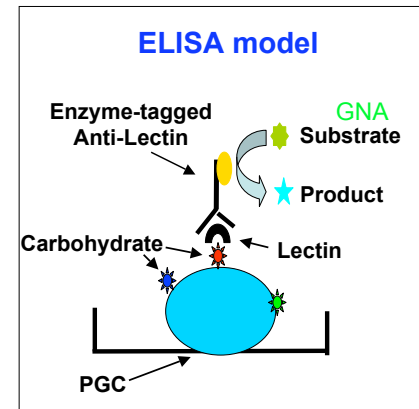
Fig.1 The amino acid sequence of the mature PGC protein (which begins at Ala41 of the preproprotein) containing an N-linked glycosylation site at Asn220 and seven O-linked sites. Site mapping of glycosylation is indicated with larger, bold letters at the site of modification and is designated N-linked glycosylation or O-linked glycosylation by the symbols @ and *, respectively

Result: The presence of T, S and N residues suggests that PGC may be O- and N-glycosylated.

37

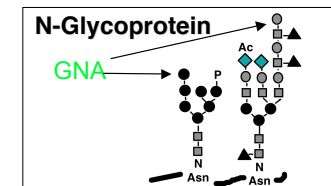
II. Presence and What type of glycosylation ?

Approach 2: ELISA with lectin as probe



Result: Five lectins tested, only GNA showed a strong positive binding, indicating the protein is N-glycosylated.

GNA: binds terminally linked mannose of high-mannose or hybrid-type N-linked glycan



38

II. Presence and What type of glycosylation ?

Approach 3: MS + Endo-H digestion

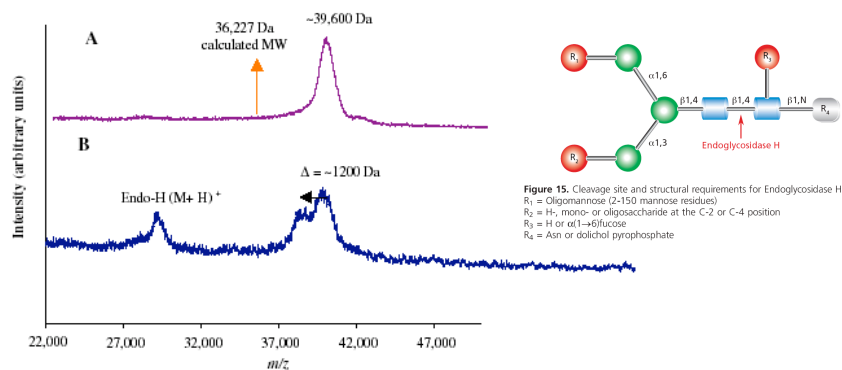


Figure 15. Cleavage site and structural requirements for Endoglycosidase H. R₁ = Oligomannose (2-150 mannose residues) R₂ = H, mono- or oligosaccharide at the C-2 or C-4 position R₃ = H or α 1-6) fucose R₄ = Asn or sialohol pyrophosphate

Fig. 2. MALDI-TOF MS spectra of (a): intact PGC protein and (b): the Endo-H digested PGC. The molecular weight of intact PGC was approximately 39600 Da, 3300 Da larger the molecular weight calculated from amino acid sequence. After Endo-H on target digestion for 30min, a new peak of approximately 1200 Da less appeared. However, the intact protein still was the major peak which suggested the digestion was not complete. The additional peak around 29000 Da shows the Endo-H molecular ion peak.

Result: PGC is N-glycosylated.

39

III. What is the composition of N-glycosylation ?

Approach 3: MS + trypsin digestion

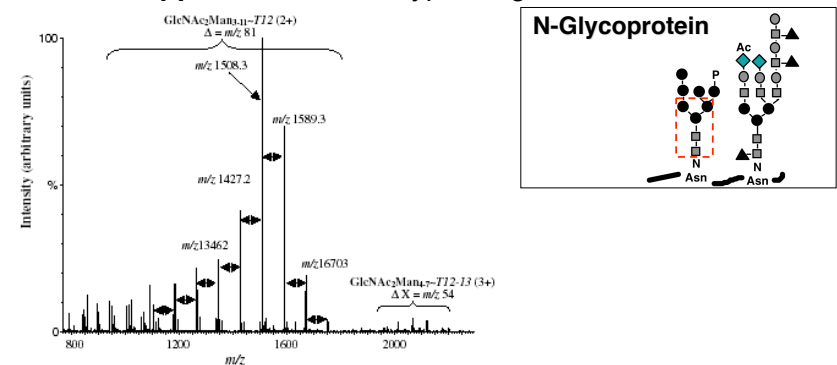


Fig. 4. Averaged spectrum of N-linked glycopeptide. The 2+ charged peaks had incremental spacing of m/z 81 between peaks, and the 3+ charged peaks had incremental spacing of m/z 54. Each species in both spectra was separated by a single hexose residue.

The PGC is N-glycosylated with a composition of GlcNAc₂Man₃₋₁₁ 40