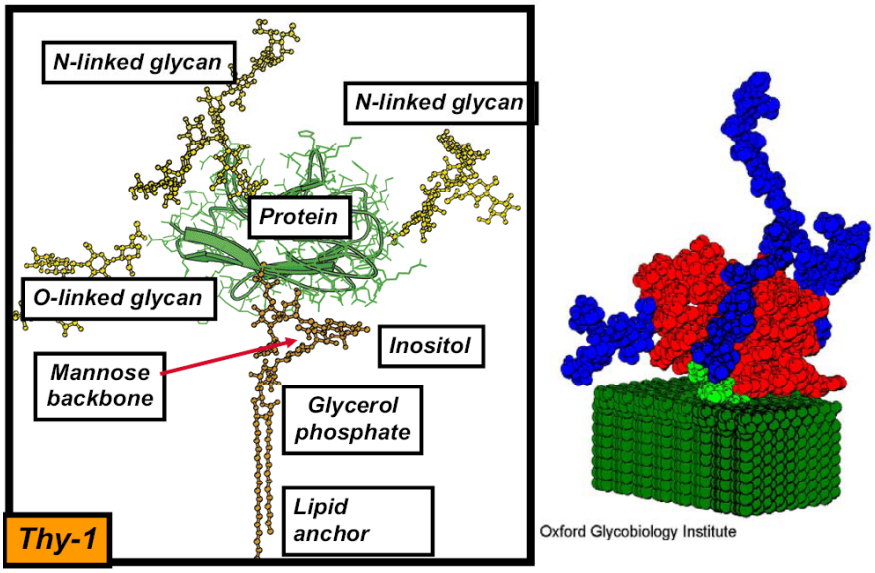
Glycoproteins
Release and Analyze

Ron Orlando

Glycobiology

Glycoconjugates
- Glycolipids
- GPI anchors
- Proteoglycans
- Protein N- and O-linked glycosylation
  - o O-linked GlcNAc
    - Nucleus, Cytoplasm
    - Associated with Phosphorylation
  - o Asn-linked
    - Asn-Xxx-Ser/Thr Consensus Sequence
    - Complexity Increases with that of cell
  - o Ser/Thr-Linked
    - Consensus? (proline, other indicators, not reliable)
    - Hard to release, no good endoglycosidases

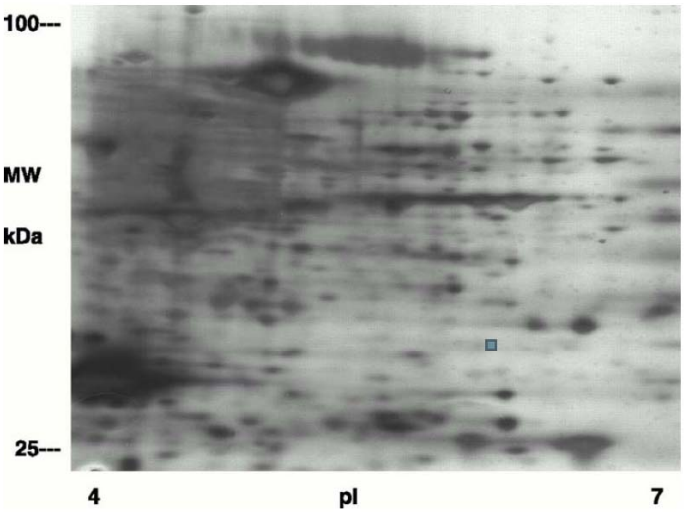
Glycans Modify Proteins in Many Ways



Importance of the Carbohydrate Chains attached to Glycoproteins

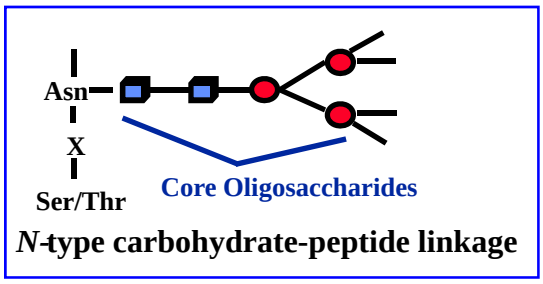
- 50-90% of proteins are glycosylated
- Often required for biological activity
- Required for proper protein folding
- Protect against proteolysis and thermal denaturation
- Participate in the immune response
- Change with condition of the cell/tissue

Silver Stained 2D Gel of O-GlcNAc-Containing Nucleocytoplasmic Proteins from HeLaS3 Cells

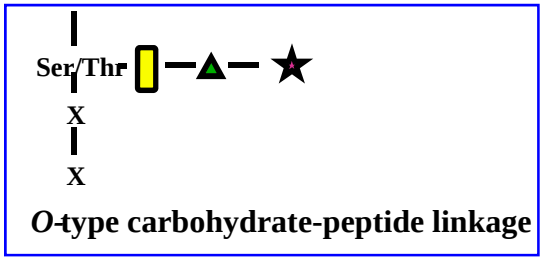


Wells et al. Science, 2001.

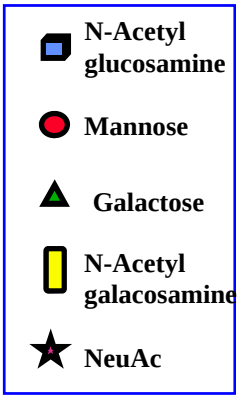
N- and *O*-linked Carbohydrate Chains of Glycoproteins



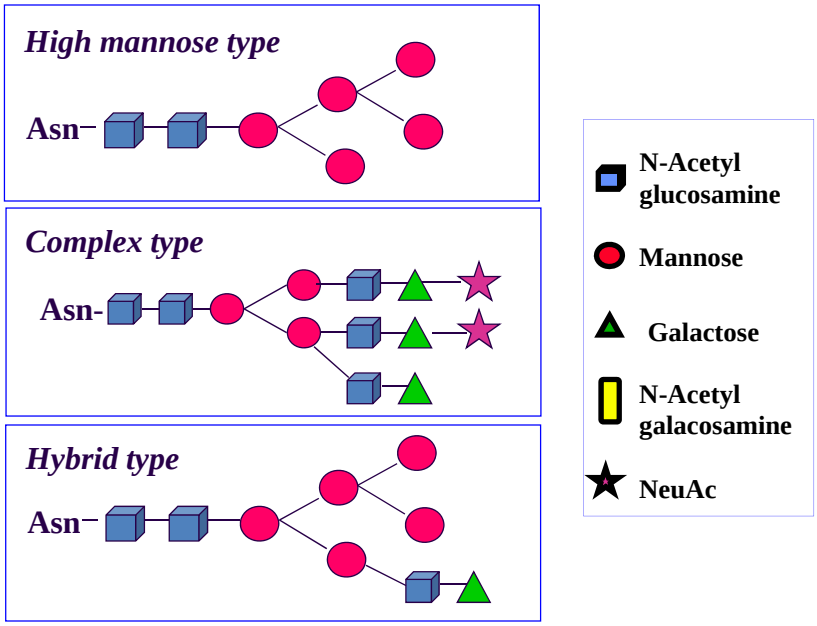
***N*-type carbohydrate-peptide linkage**



***O*-type carbohydrate-peptide linkage**

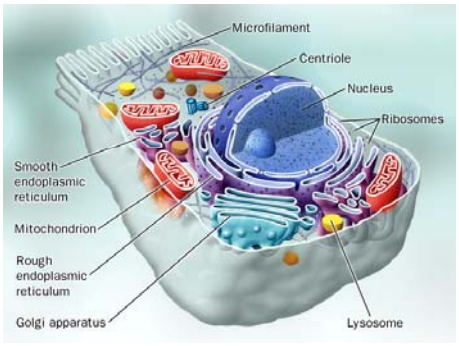


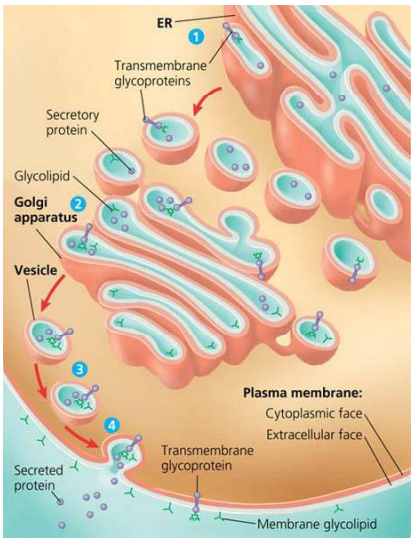
Typical Structures of *N*-linked Carbohydrate Chains



Glycoprotein Synthesis, Folding, Modification and Transport

Protein glycosylation begins in the ER and continues in the Golgi. The type and specific glycan structure attached to proteins is highly dependent on cellular architecture, enzyme expression, and chaperone proteins.





Synthesis and maturation of Asn-linked oligosaccharides

- All eukaryotes have covalently bound carbohydrates attached through an amide linkage to Asn residues
- Many aspects of the pathway are conserved from yeast to plants and animals
- Common features:
  - » Similar synthesis of the lipid-linked precursor of the protein-bound oligosaccharide
  - » Similar transfer to protein acceptor sequence motif: Asn-X-Ser/Thr
  - » Trimming of all of the glucose residues
  - » Trimming of some of the mannose residues
  - » Extension of the trimmed oligosaccharide by sugar addition in the Golgi complex
  - » Similar mechanism for acquiring nucleotide sugars into the Golgi complex

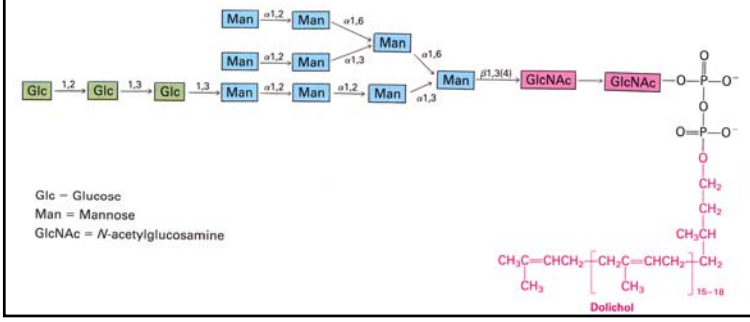
Synthesis of lipid-linked precursor

N-linked Oligosaccharide precursor synthesized on a lipid substrate

Transferred to protein as it emerges on the lumenal side of the ER membrane

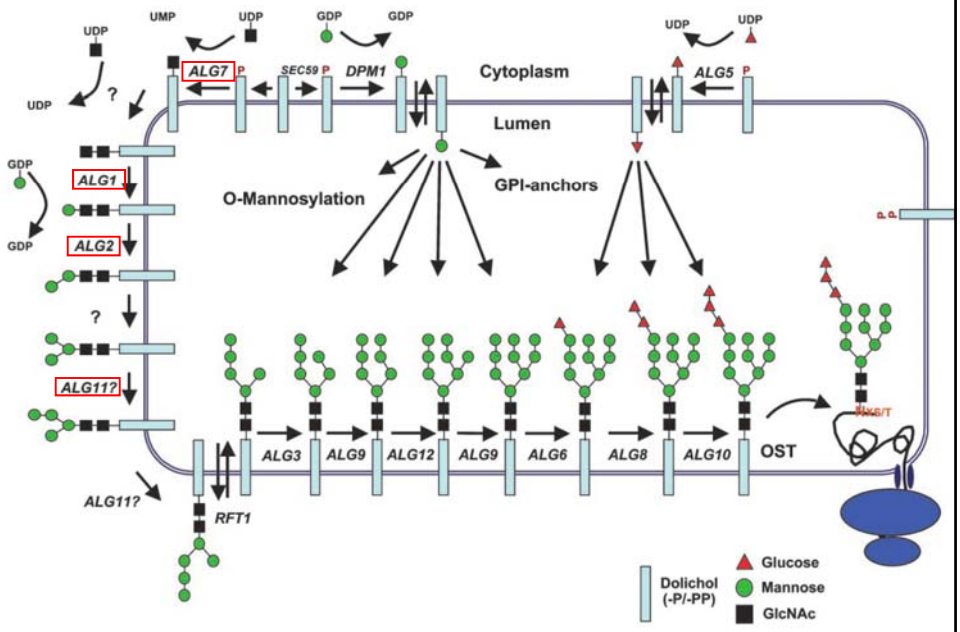
Sugar donors: UDP-GlcNAc, GDP-Man, UDP-Glc

Acceptor: Dolichol-phosphate (α-unsaturated polyisoprenoid))



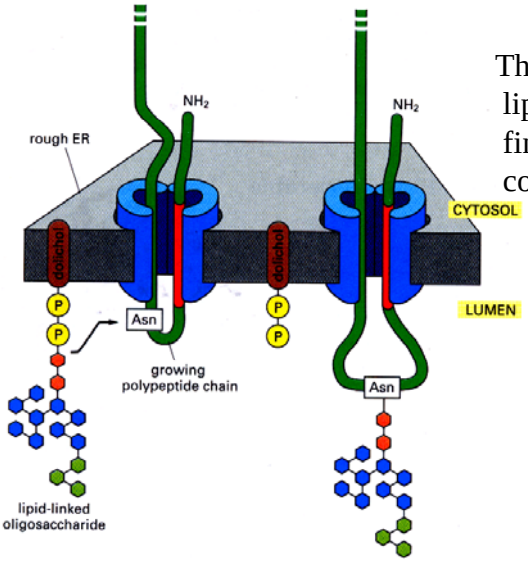
◀ Figure 17-28 Structure of the dolichol pyrophosphoryl oligosaccharide precursor of *N*-linked oligosaccharides. Dolichol is strongly hydrophobic and long enough to span a phospholipid bilayer membrane four or five times. This oligosaccharide lipid is located on the rough ER, and the sugar residues face the lumen of the organelle.

Synthesis of the lipid-linked precursor for Asn-linked glycosylation



Helenius, A. and Aebi, M. (2004) Ann. Rev Biochem 73, 1019-49

Co-translational glycosylation is the rule from yeast to plants and animals.



The steps in the synthesis of the lipid-linked precursor and the final structure are highly conserved.

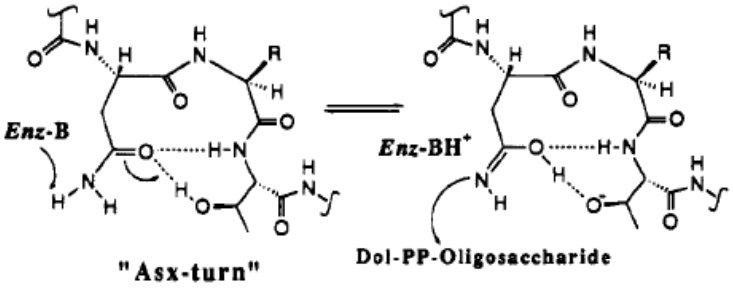
Even the exceptions are indicative of a highly conserved process:

(From Alberts el al (1994) Molecular Biology of the Cell, 3rd ed.)

Consensus sequence for N-linked glycosylation

N-X-S/T - X cant be P

Recent evidence suggests S/T-X-N can also "work"

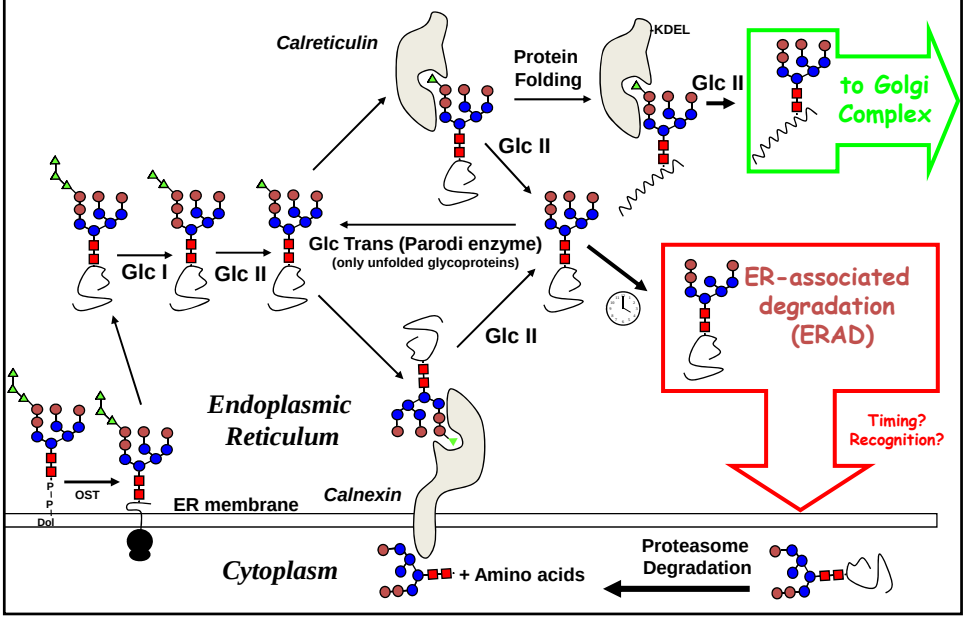


A Mechanistic Proposal for Asparagine-Linked Glycosylation

B. Imperiali,* K. L. Shannon, M. Unno, and K. W. Rickert

J. Am. Chem. Soc. **1992**, *114*, 7944–7945

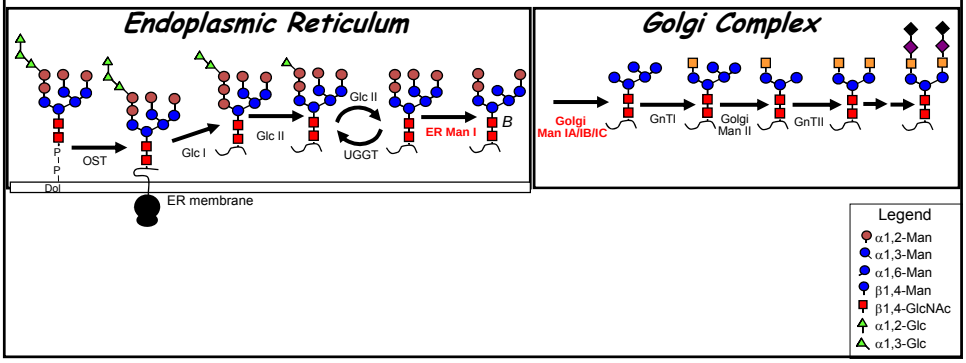
Carbohydrate structures influence glycoprotein folding in the ER



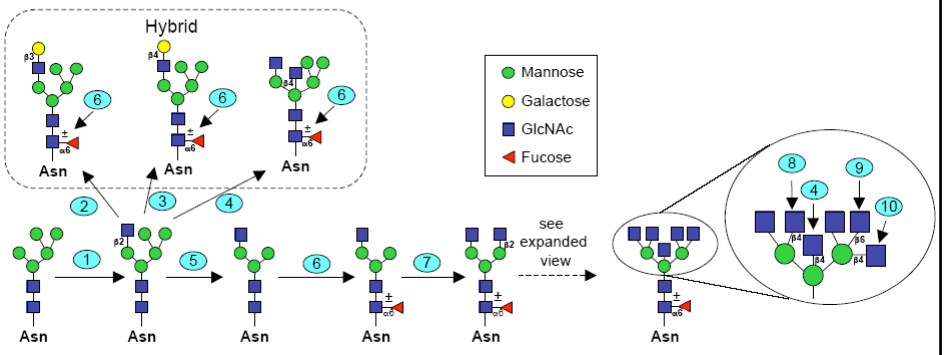
Asn-linked Glycoprotein Maturation in the ER and Golgi

Early steps in N-glycan processing:

- **Many of the early steps in the ER are highly conserved in eukaryotes**
- **Play roles in chaperone-mediated folding, quality control, and disposal of terminally unfolded intermediates**
- **Critical for maturation of N-glycans on cell surface and secreted glycoproteins**

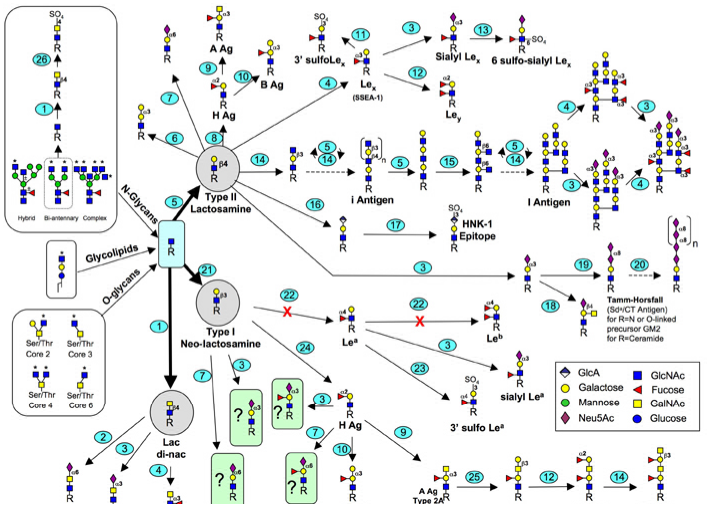


N-Glycan Branching Reactions



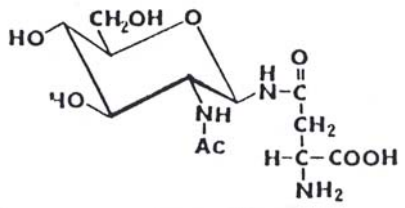
http://www.ccrc.uga.edu/~moremen/glycomics/

Glycan Capping Reactions

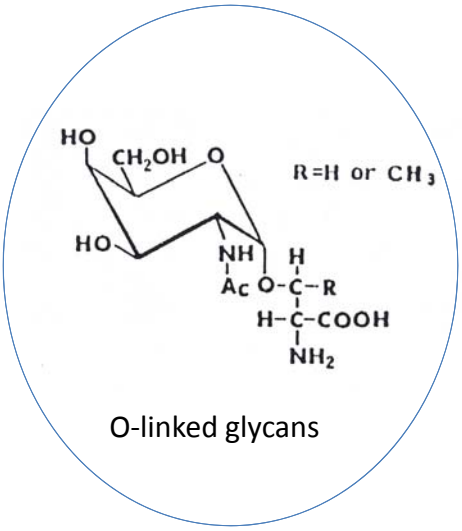


Questions should be directed to Dr. Alison Nairn(avnairn@uga.edu) or Dr. Kelley Moremen(moremen@uga.edu)
COMPLEX CARBOHYDRATE RESEARCH CENTER | 315 RIVERBEND ROAD
THE UNIVERSITY OF GEORGIA | ATHENS | GEORGIA 30602-4712 | USA
1-706-542-1705 | FAX: 1-706-542-1759

http://www.ccrc.uga.edu/~moremen/glycomics/

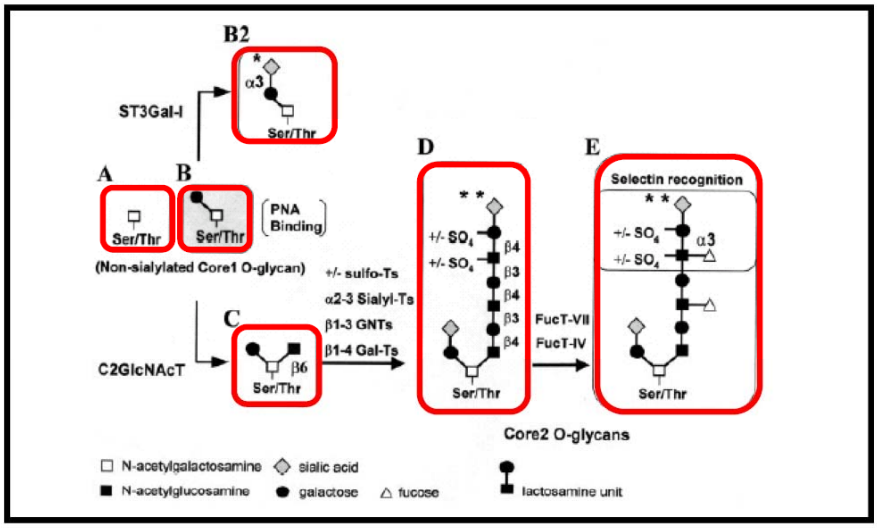


N-linked glycans



O-linked glycans

O-Linked Protein Glycosylation

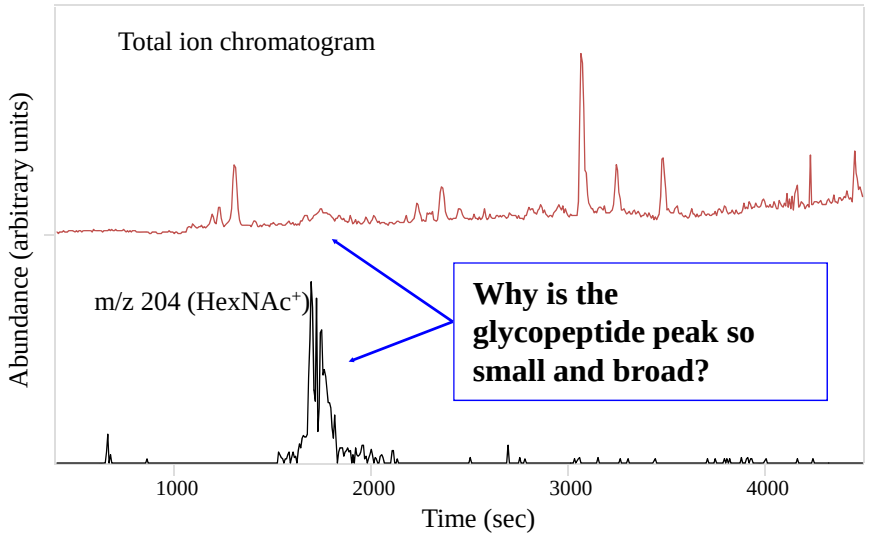


N. Grabie et al. Eur. J. Immunol. 2002. *32*: 2766–2772

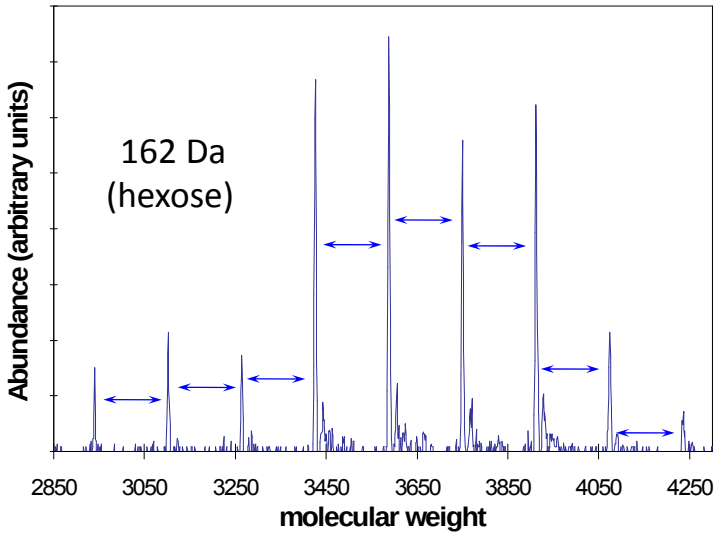
Glycoprotein Microheterogeneity

- Glycosylation is a co-/post-translational modification
  - Glycans dependent on cellular architecture and individual enzymes
  - Cell-specific
  - Heterogeneous "glycoforms" at each site of glycosylation
- A glycoprotein purified to "homogeneity" is generally still a distribution of many combinations of glycoforms at each site
- "Microheterogeneity" of glycosylation a mechanism for modulation of protein activity, circulatory half-life, etc.

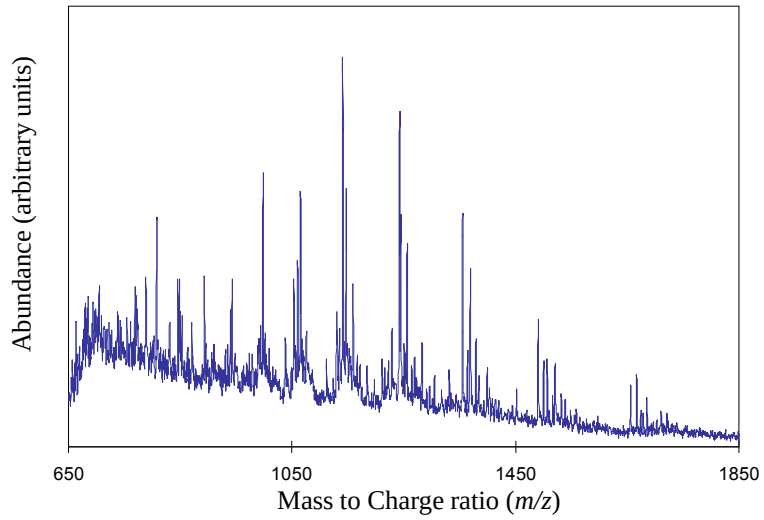
LC-MS Data for EPG I



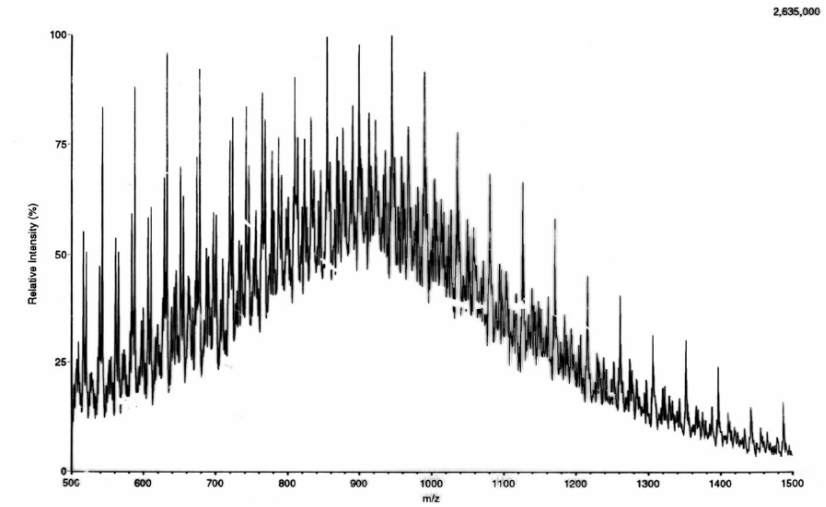
Deconvoluted Spectrum of "Pure" Glycopeptide From EPG I: Carbohydrate Heterogeneity



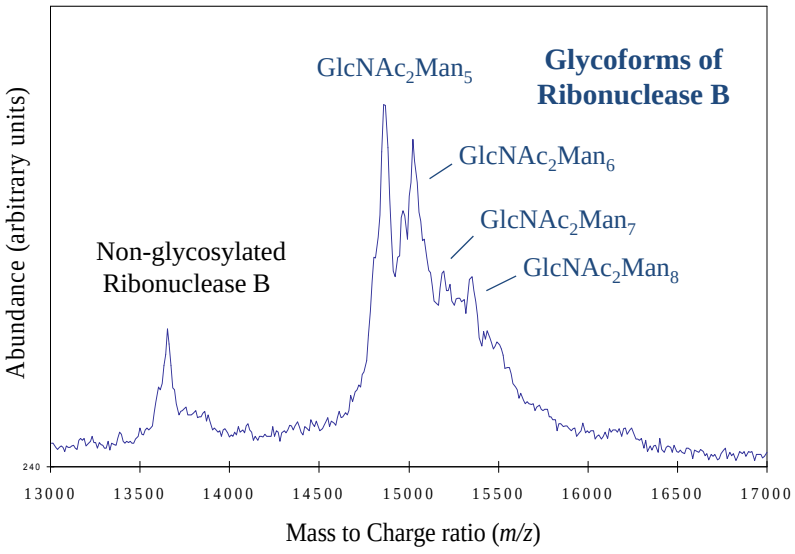
Analysis of Ribonuclease B by ESI-MS



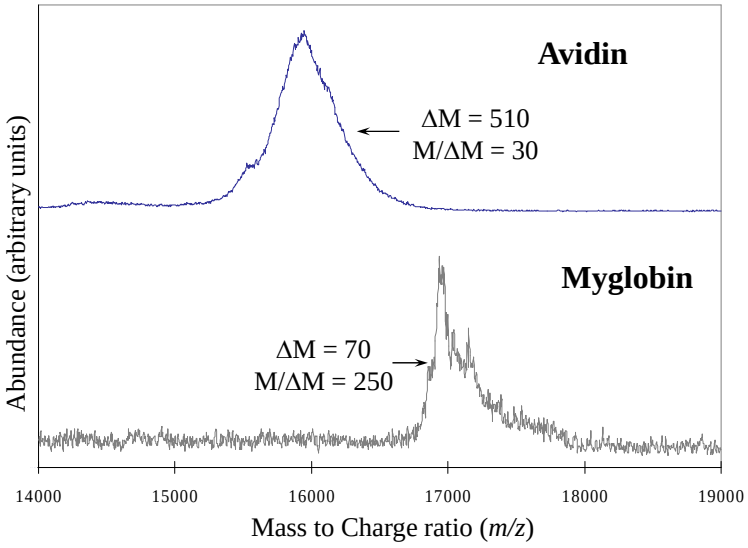
Analysis of a more heterogeneous glycoprotein by ESI-MS



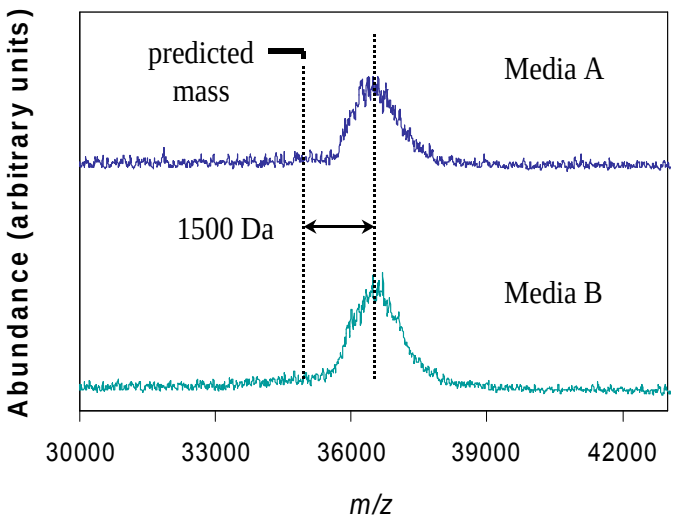
Analysis of Ribonuclease B by MALDI-MS



More Glycan Heterogeneity Problems



MALDI-TOF Data for EPG I



Glycoprotein Microheterogeneity

Electrospray mass spectrometry of an intact glycoprotein generally yields a spectrum **too complex** for conventional deconvolution or interpretation

MALDI-TOF/MS generally **too low** resolution to reveal any details of glycoprotein structure, but average masses of some value

Generally unusual to say anything useful about protein glycosylation from MS of intact material

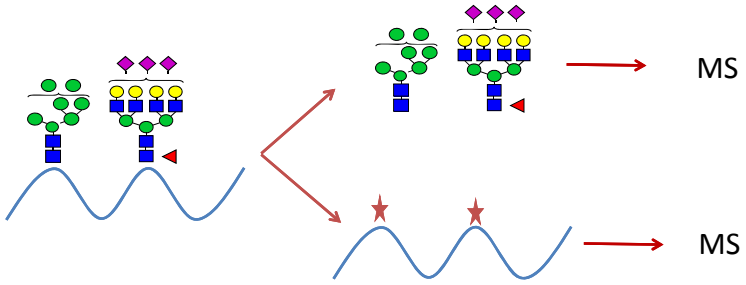
- Exceptions: single sites of glycosylation, very low heterogeneity

Analytical Strategy for Glycoprotein Characterization

Two questions are central to glycoprotein analysis:

1. Where on the protein are glycan chains attached?
2. What are the structures of the glycan chains?

Release and Analyze

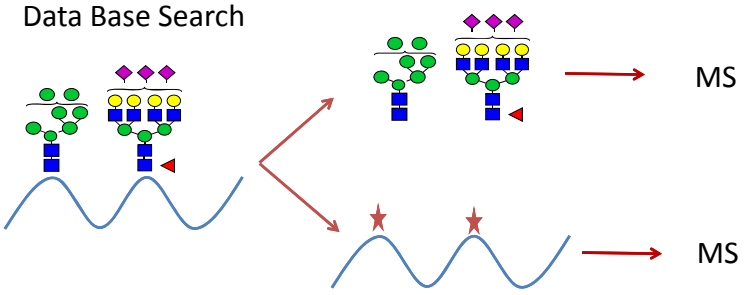


Glycosylation Site Identification

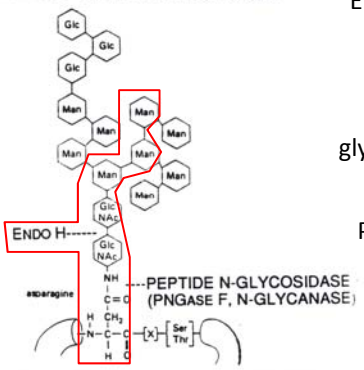
Release Glycans and Tag Sites

For this type of site mapping one wants to release glycans but "tag" the sites of glycosylation, in this manner making all sites look the same

- Release glycans
- MS or MS^n
- Data Base Search



SPECIFICITY OF ENDOGLYCOSIDASES

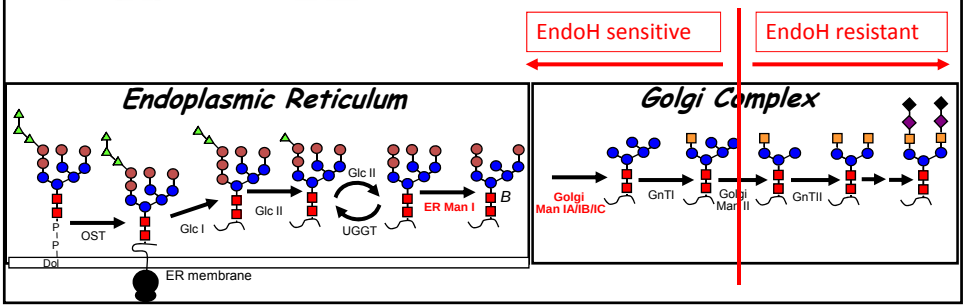


Endoglycosidases can be used to identify location of a glycoprotein in a cell or release glycans for structural analysis

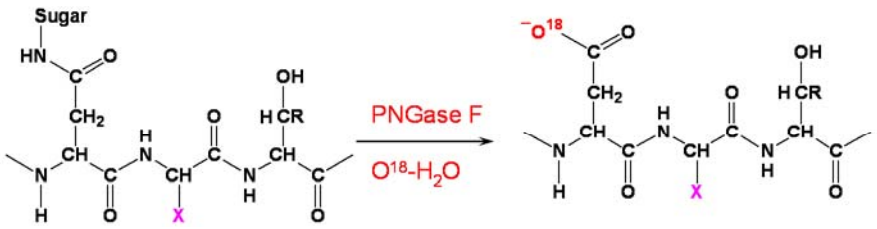
PNGAse F most common, but it will not cleave glycans with Fuc linked to the 3 position of the core GlcNAc (present in plants)

PNGase A will but its not commercially available

Endoglycosidase work much better after proteolytic digestion



PNGaseF (Peptide-N-Glycosidase F) Catalyzed Mechanism

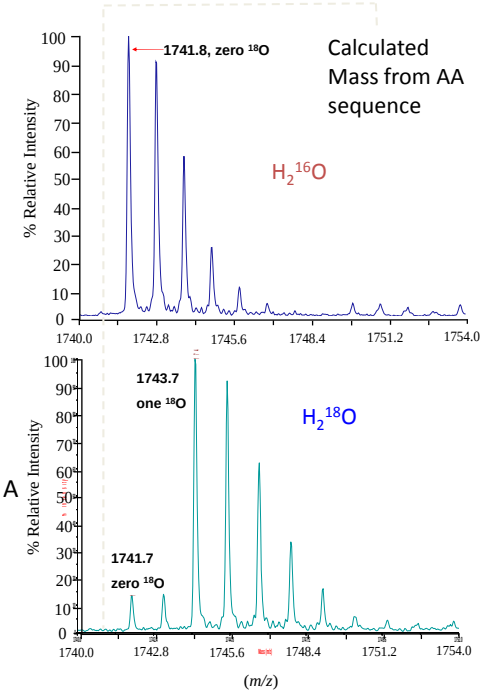


Products of PNGase F digestion

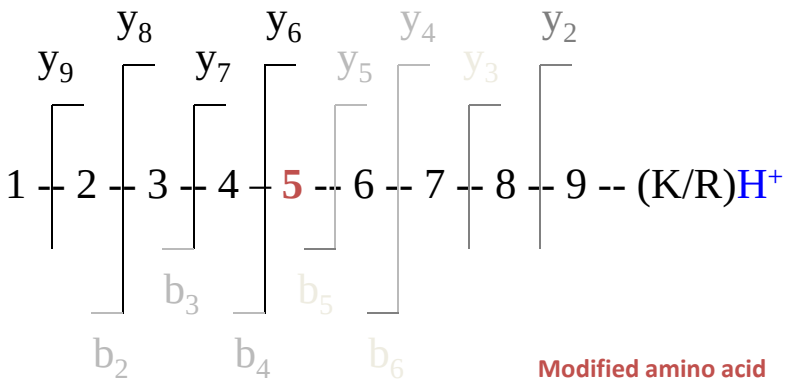
Deglycosylation causes deamidation of glycosylated N producing D, which results in either a 1 or a 3 Da shift if the reaction is performed in $H_2{}^{16}O$ or $H_2{}^{18}O$

The + 3 Da shift is easier to identify and eliminates confusion with naturally deamidated N residues

This can also be performed with PNGase A



Fragment Ions Observed upon MS/MS Analysis of Tryptic Peptides



Incorrect Assignment of Glycosylation Sites

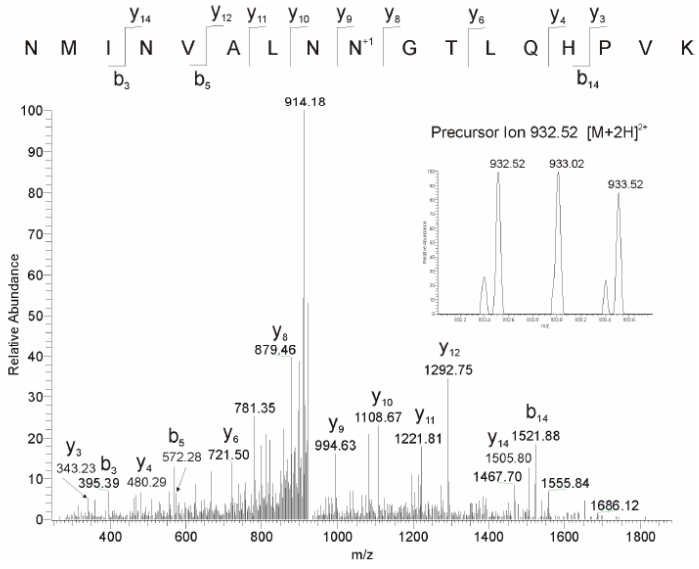
- Naturally occurring sites of deamidation
- Poor mass accuracy
- Data base search routines
- Trypsin

N-linked glycosylation sites identified in a proteomic experiment on *E. coli*, which does not glycosylate

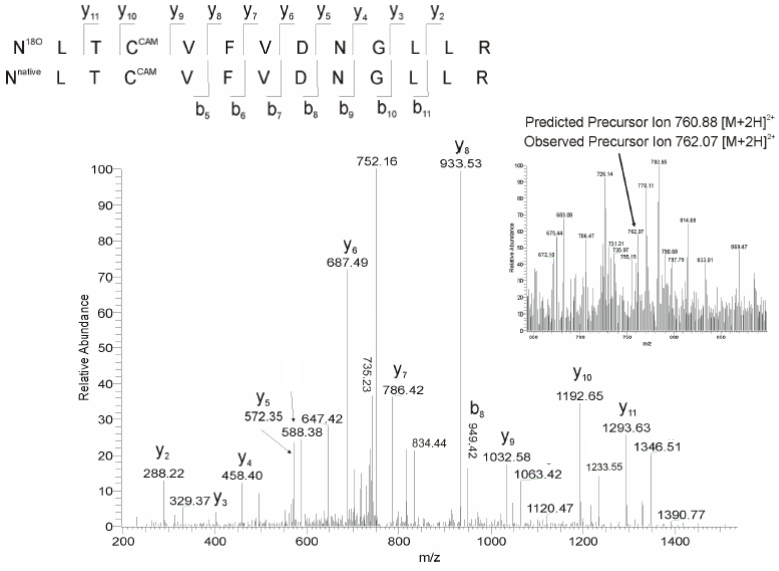
1% FDR	LTQ		LTQ-FTICR	
	Deamidation ^{16}O	Deamidation ^{18}O	Deamidation ^{16}O	Deamidation ^{18}O
Unique glycopeptides	519	312	65	19
Unique glycopeptides containing motif (NXS/T)	55	33	12	**0**
Unique protein	271	270	317	334
Unique glycoproteins	212	170	58	18
Unique glycoproteins containing motif (NXS/T)	43	25	10	**0**

- Lower m/z accuracy of LTQ leads to more false positives
- Glycosylation sites identified by LTQ-FT in ^{16}O were either sites of deamidation or selection of the $1^{13}C$ peak as the precursor
- ^{18}O labeling reduces incorrect assignments by reducing mass accuracy demands and eliminating natural deamidation being assigned as a glycosylation site

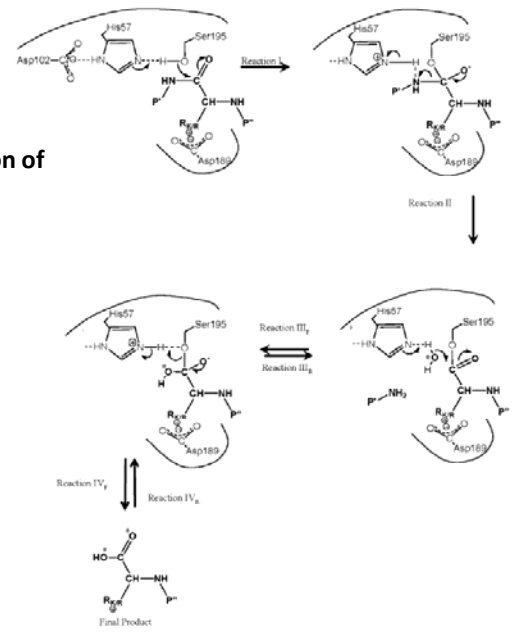
Example of natural deamidation being identified as a site of N-glycosylation



Example of mass error causing the incorrect assignment of an N-linked glycosylation site



Trypsin can also lead to the addition of ^{18}O during PNGase F digestion



Peptides identified as containing N-linked glycosylation sites Care was not taken to prevent tryptic activity during PNGase F deglycosylation in ^{18}O enriched water

Sites "identified" as being glycosylated are denoted with as N*, and the consensus sequences for N-linked glycosylation (N-X-S/T) are underlined.

Denoted N-linked glycosylation sites	Charge	XCorr	ΔC_n
IYGSIPVEFTQLN*FQFLN*VSYN*R(L)	3	5.46	0.71
IYGSIPVEFTQLNFQFLN*VSYN*R(L)	2	4.58	0.10
LQSFDEYSYFHN*R(C)	2	3.48	0.18
NKLEGDASVIFGLN*K(T)	3	4.22	0.12

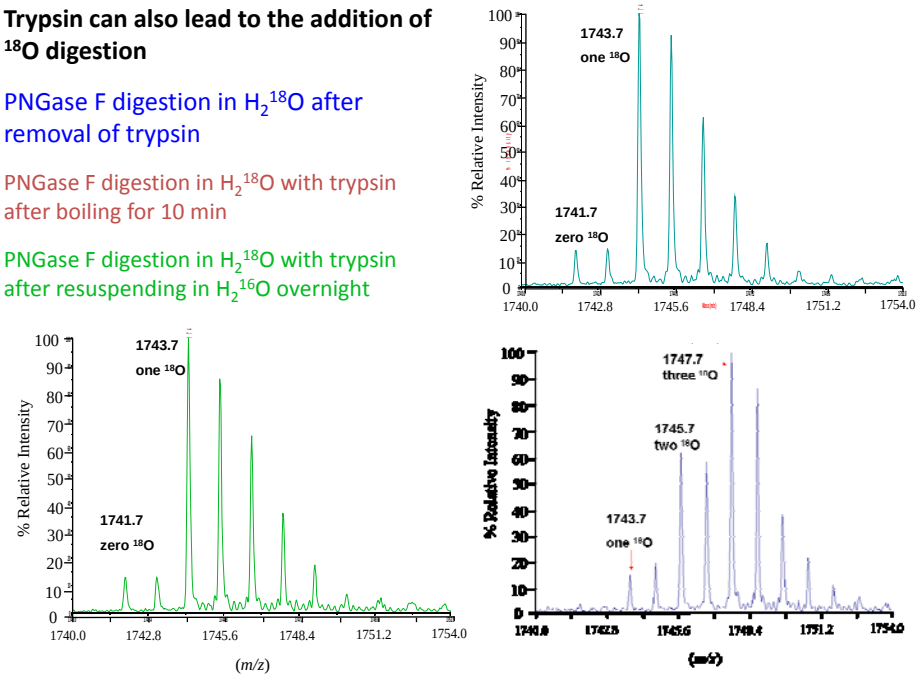
5 of the 7 sites identified must be incorrect

Trypsin can also lead to the addition of ^{18}O digestion

PNGase F digestion in $H_2{}^{18}O$ after removal of trypsin

PNGase F digestion in $H_2{}^{18}O$ with trypsin after boiling for 10 min

PNGase F digestion in $H_2{}^{18}O$ with trypsin after resuspending in $H_2{}^{16}O$ overnight



What we do

- Perform deglycosylation site (N-linked) experiments with $H_2{}^{18}O$.
- Re-suspend in $H_2{}^{16}O$ for 24 hours after deglycosyalation
- Use a high resolution/accurate mass instruments (Q-ToF, Orbitrap, FTMS)
- Accept only automated assignments with a false discover rate of 1% or better if visual inspection reveals at least 1 fragment ion containing the modified residue.

Release of O-linked glycans: Enzymatic Cleavage

O-glycanase (*endo*-a-N-acetyl-D-galactosaminidase) from *Diplococcus pneumoniae* – hydrolyzes O-glycosidic linkage between GlcNAc and Ser or Thr of the disaccharide Galα1-3GalNAc.

Appears to require unsubstituted residues; so must desialylate first by mild acid hydrolysis or neuraminidase.

Endo-a-N-Acetylgalactosaminidase (*Diplococcus pneumoniae*)

(O-Glycanase™)

Galß1-3GalNAcα1-Ser/Thr-(Peptide)

↓

Galß1-3GalNAc + Ser/Thr-(Peptide)

Limitation: Substitution with sialic acid or other saccharides blocks activity

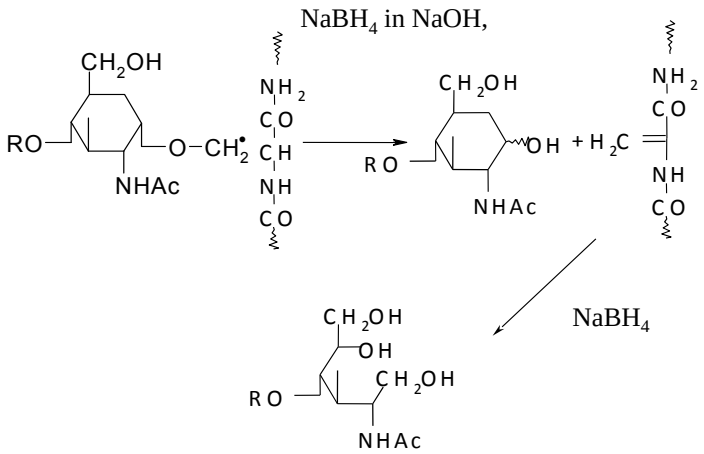
Release of O-linked glycans: β-Elimination

The conventional alkaline ß-elimnation has been used in occasion where there are mg quantity of material present. Using this method, the base labile O-glycosidic linkages between the GlaNAc and the Ser/Thr residues of the protein are cleaved under mild alkaline conditions. Involves treatment of glycopeptides or glycoproteins with mild alkaline borohydride

Complete release of *O*-linked oligosaccharides via β-elimination.

Yields stable sugar alditols with destruction of the peptide.

Alkali-labile O-linked oligosaccharides



Release of O-linked glycans: Ammonia-based ß-elimination

- Alternative ß-elimination using ammonia-based medium has recently been proposed (Yunping Huang , Rapid Commun. Mass Spectrom, 2002, 16, 1199-1204).
- Aqueous ammonia has been substituted instead of the NaOH used in the conventional method.
- The ammonia keeps the pH at around 11 which is suitable for ß-elimnation of O-glycans. It is also easily removed by evaporation.
- Borane-ammonia complex (BH_3 NH_3) has replaced the $NaBH_4$ reducing agent which was the cause of much of the salt.
- The elimination of the $NaBH_4$ allowed the use of only small amount (40ul) of cation exchange resin for removal of the borane.ammonia complex.

Labeling sites of O-glycosylation

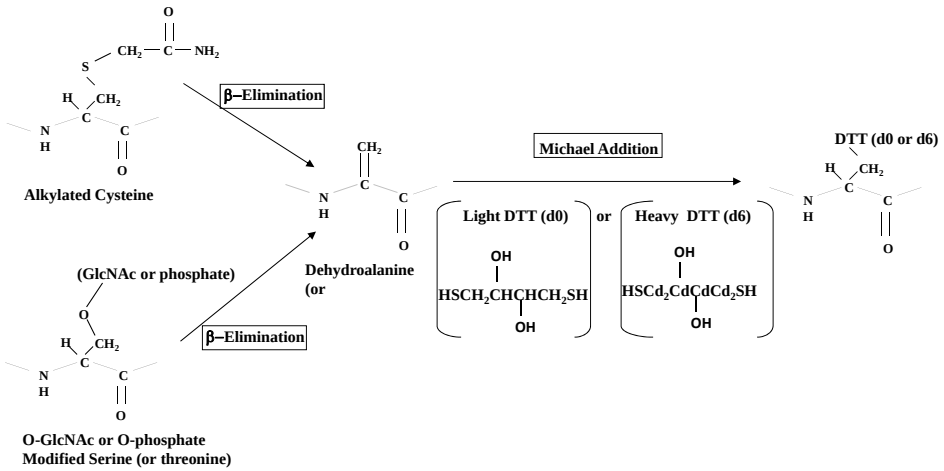
The draw back for both the ß-elimination and the hydrazinolysis method has been that the integrity of the protein component is not retained and no information can be obtained on the actual sites of O-glycosylation.

A recent based catalyzed ß-elimination method uses the addition of ammonia to the unsaturated amino acid, i.e. OH group is replaced by a NH_2 group. The problem is that the difference of 1 mass unit is often difficult to observe.

A base catalyzed ß-elimination is attractive with being able to label the original Ser/Thr residue so that the site of O-glycosylation can be determined.

This can be followed by alkylamine labeling by the addition of methylamine or ethylamine

Differential isotopic tagging of both cysteine and post-translationally modified ser/thr through β-elimination/Michael addition with light (d0) and heavy (d6) DTT.

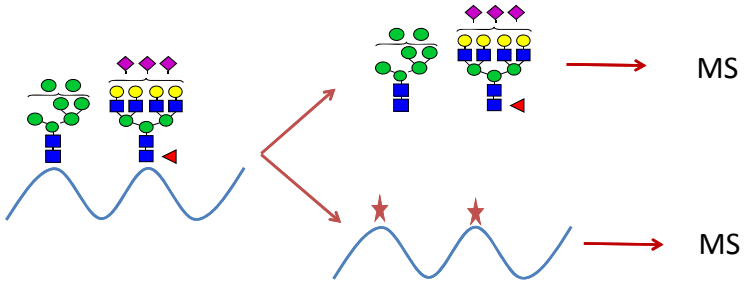


Analytical Strategy for Glycoprotein Characterization

Two questions are central to glycoprotein analysis:

1. Where on the protein are glycan chains attached?
2. **What are the structures of the glycan chains?**

Release and Analyze



Analysis of Released Glycans

- Common derivatizations/sample prep.
- MS
- MS/MS and MS^n
- Exoglycosidase digestions

Why Permethylate the Oligosaccharide?

It increase the sensitivity of oligosaccharides for subsequent MS analysis. "Equalizes" the MS response for different glycans

The mass increase is not too much to shift the mass to higher mass range and decrease sensitivity.

It allows for diagnostic molecular ions which are easier to interpret than the native oligosaccharides.

Stabilizes negatively charged sugars (sialic acids for example)

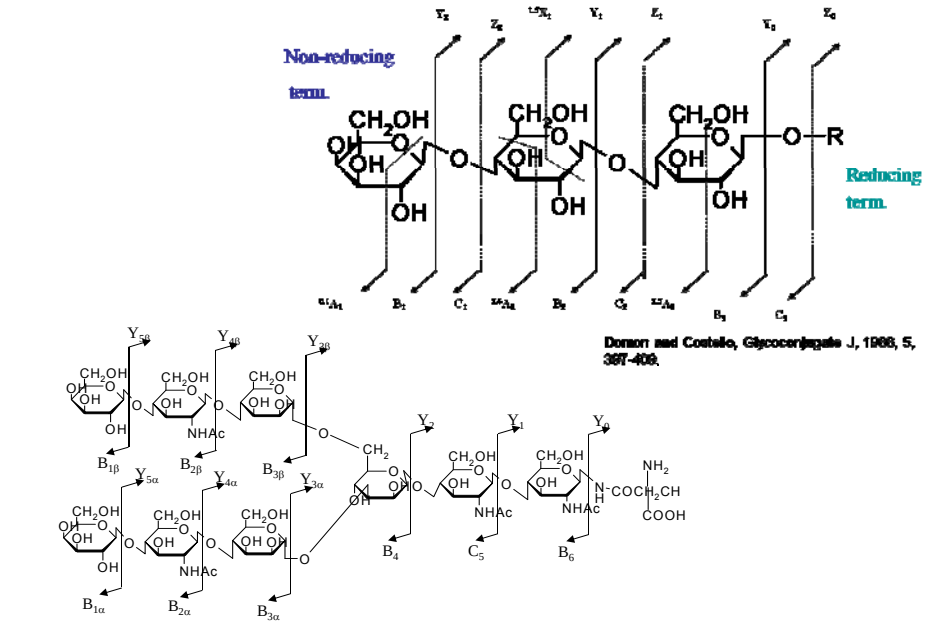
Makes tandem mass spectra more interpretable

MS/MS Analysis of Glycans

These derivatives give diagnostic MS/MS or CID spectra, which are important in obtaining sequence information on the oligosaccharide.

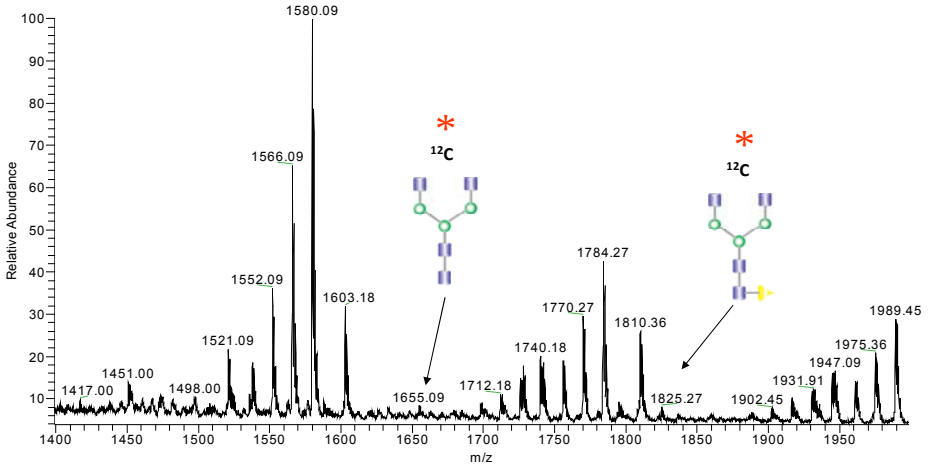
The cross-ring cleavages are more prominent in the spectra of Li+, Na+ or K+ adducts rather protonated molecular ions.

CID Fragment ions for N-glycan Linked Glycans



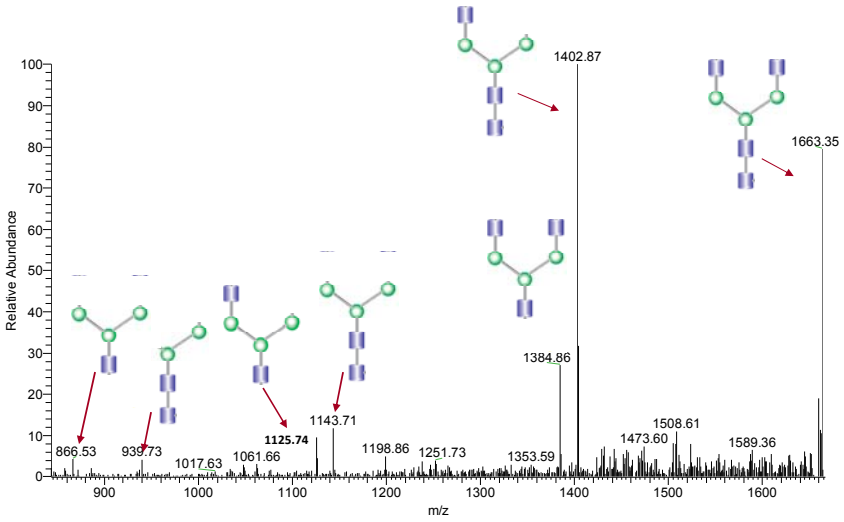
ESI-MS analysis of released N-linked glycans from *T. cruzi*

Show MS/MS capabilities of low abundance ions – glycans fragment very well



ESI-MS analysis of released N-linked glycans from *T. cruzi*

Show MS/MS capabilities of low abundance ions – glycans fragment very well

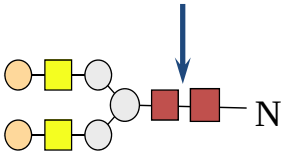


MS/MS of Glycans

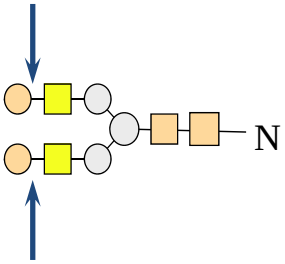
- Provides non stereochemical sequence and branch points
- Can at times provide linkage information based on the appearance of fragments arising from cross-ring cleavage
- Information of stereochemistry and anomericity are inferred from database – biosynthetic pathway (may not always be correct)

Glycosidase Digestions

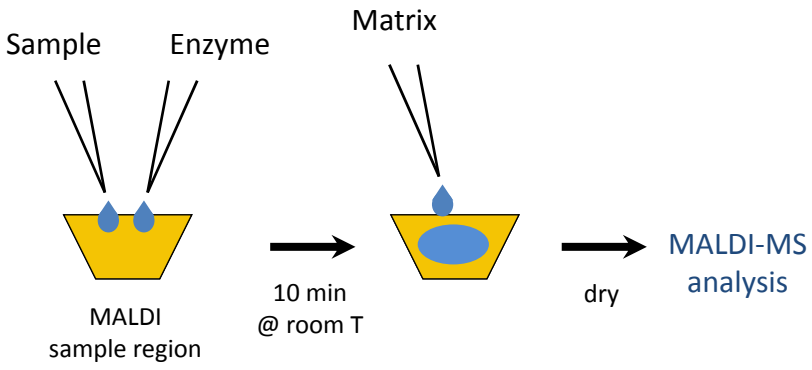
*Endo*glycosidases



*Exo*glycosidases

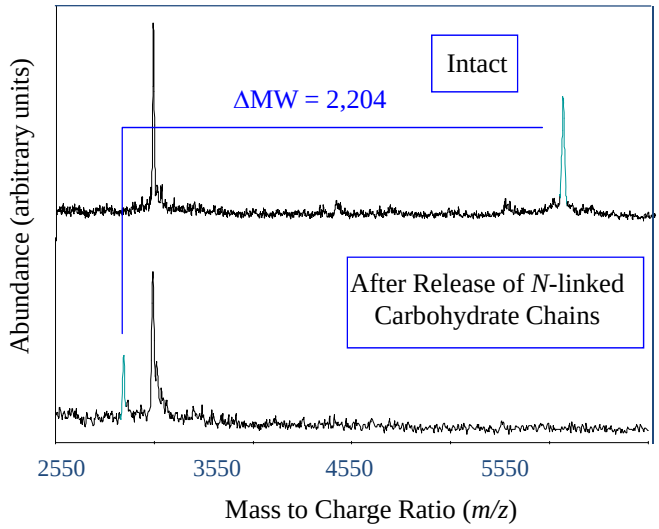


On-target Glycosidase Digestion Procedure

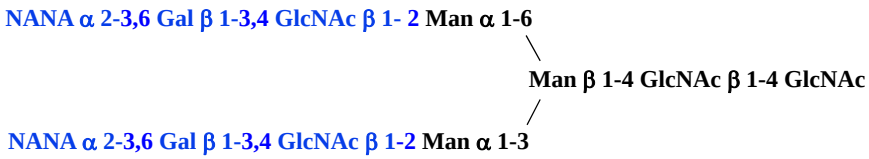


J. Colangelo, **R. Orlando**, *Anal. Chem.*, **1999**, *71*, 1479-1482.
J. Colangelo, R. Orlando, *Rapid. Commun. Mass Spectrom*, **2001**, *15*, 2284-2289.

*Endo*glycosidase/MALDI-MS analysis of GP-II (CST)

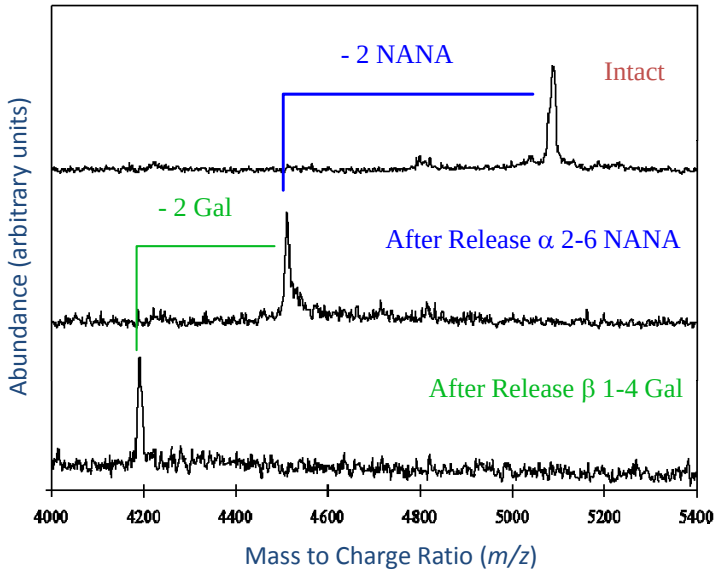


Known Glycoprotein Glycans with a MW Equal to that Found on GP-II (CST)



Non-variable **Variable** **Core**

*Exo*glycosidase/MALDI-MS analysis of GP-II (CST)



Deduced Structure of the Carbohydrate Chain Attached to GP-II of Canine Serum Transferrin

NANA α 2-6 Gal β 1-4 GlcNAc β 1-2 Man α 1-6
****
Man β 1-4 GlcNAc β 1-4 GlcNAc
/
NANA α 2-6 Gal β 1-4 GlcNAc β 1-2 Man α 1-3

Advantages of Exoglycosidase

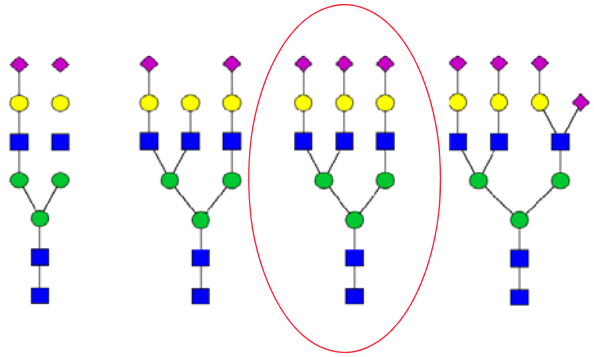
- Can provide linkage, stereochemical, and anomeric configuration
- Small amounts of material – enough for 3-10 additional MS experiments
- Fast – can completely characterize a glycan in an hour or so

Disadvantages of Exoglycosidase digestion/MS for Analyzing Glycoprotein Glycans

- Limited availability of *exo*glycosidases
- Cannot always provide complete glycan structures
- Does not work well with mixtures
- Need prior information (database)

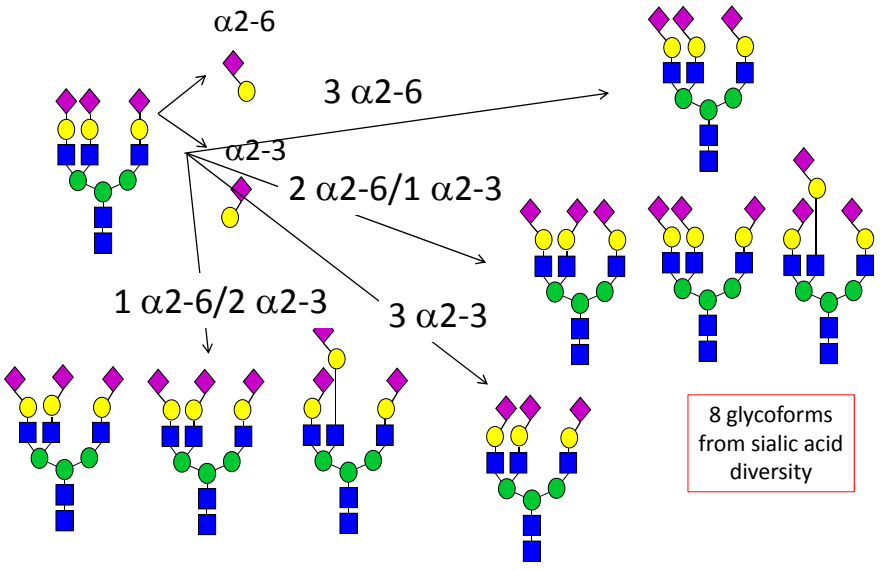
Problem: Identification of Individual Glycoforms

Ex. Major glycans attached to bovine fetuin



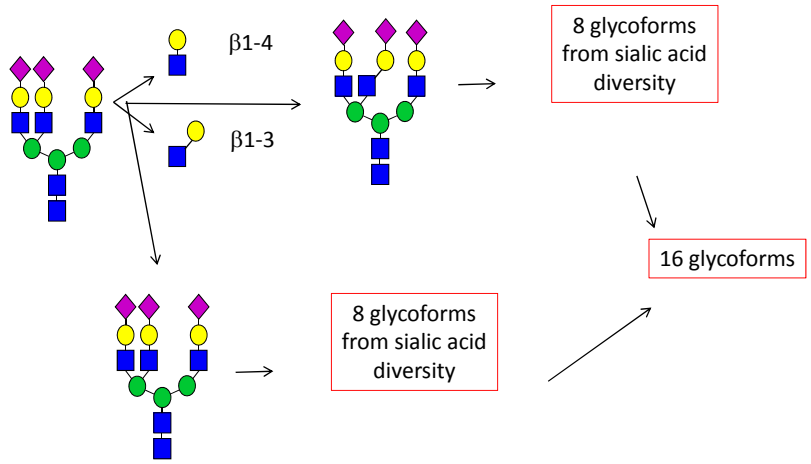
Problem: Identification of Individual Glycoforms

Ex. Fully sialated triantennary glycans from fetuin



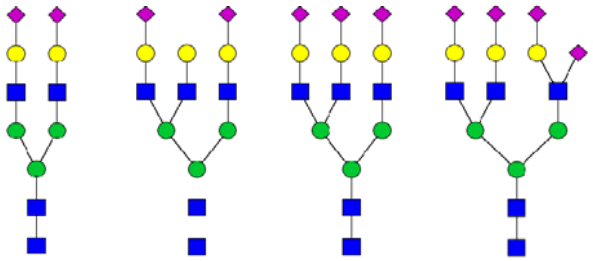
Problem: Identification of Individual Glycoforms

Ex. Fully sialated triantennary glycans from fetuin



Problem: Identification of Individual Glycoforms

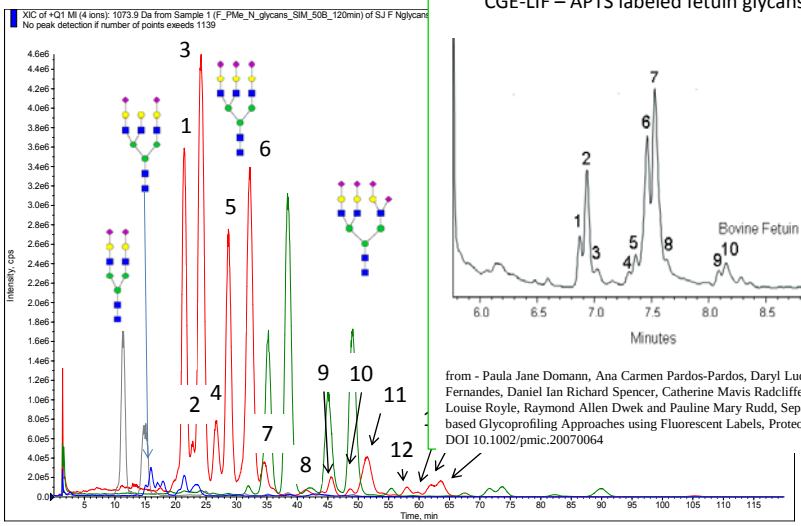
Ex. Major glycans attached to bovine fetuin



# of biosynthetically possible glycoforms	4	14	16	48

RP-LC/MS of per-methylated glycans released from fetuin

Glycoform separation

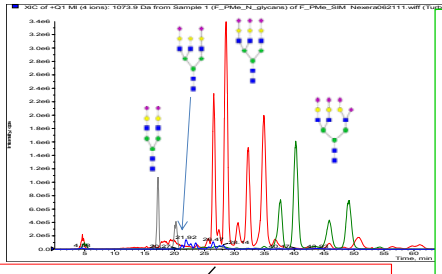


CGE-LIF – APTS labeled fetuin glycans

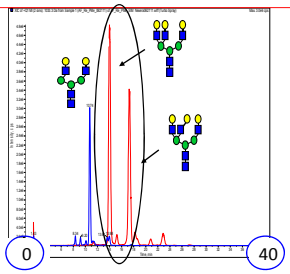
from - Paula Jane Domann, Ana Carmen Pardos-Pardos, Daryl Ludger Fernandes, Daniel Ian Richard Spencer, Catherine Mavis Radcliffe, Louise Royle, Raymond Allen Dwek and Pauline Mary Rudd, Separation-based Glycoprofiling Approaches using Fluorescent Labels, Proteomics DOI 10.1002/pmic.20070064

RP-LC/MS of per-methylated glycans released from fetuin

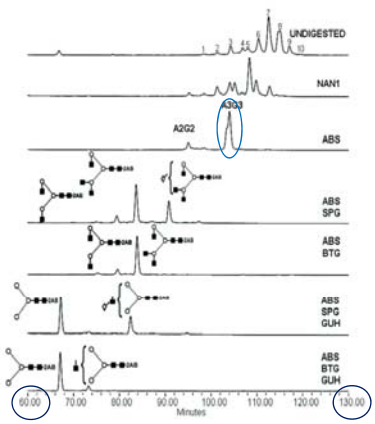
Removal of sialic acids decreases chromatographic complexity



Chemical release of sialic acids



NP LC – fluorescence , 2-AB labeled fetuin glycans

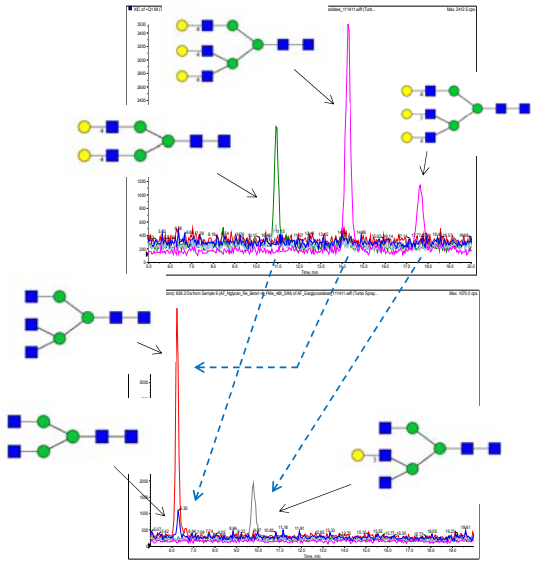


from - Paula Jane Domann, Ana Carmen Pardos-Pardos, Daryl Ludger Fernandes, Daniel Ian Richard Spencer, Catherine Mavis Radcliffe, Louise Royle, Raymond Allen Dwek and Pauline Mary Rudd, Separation-based Glycoprofiling Approaches using Fluorescent Labels, Proteomics DOI 10.1002/pmic.20070064

Exoglycosidase Digestion of Asialofetuin Glycans with β1-4 Galactosidease

Before

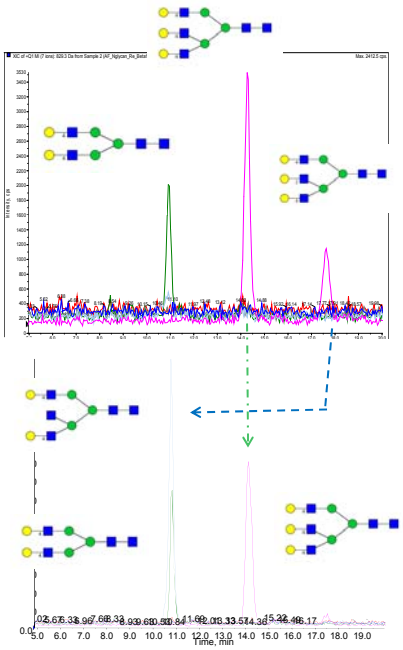
After



Exoglycosidase Digestion of Asialofetuin Glycans with β1-3 Galactosidease

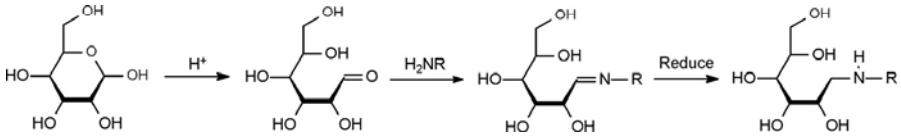
Before

After



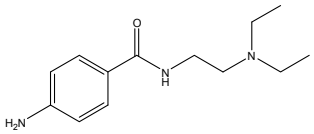
HILIC Analysis of Released and Labeled Glycans

Enzymatic deglycosylation of proteins (or peptides, etc.) using PNGase F releases glycans to yield a free reducing terminus (alditol) that is readily labeled by amines via the formation and reduction of a Schiff's base



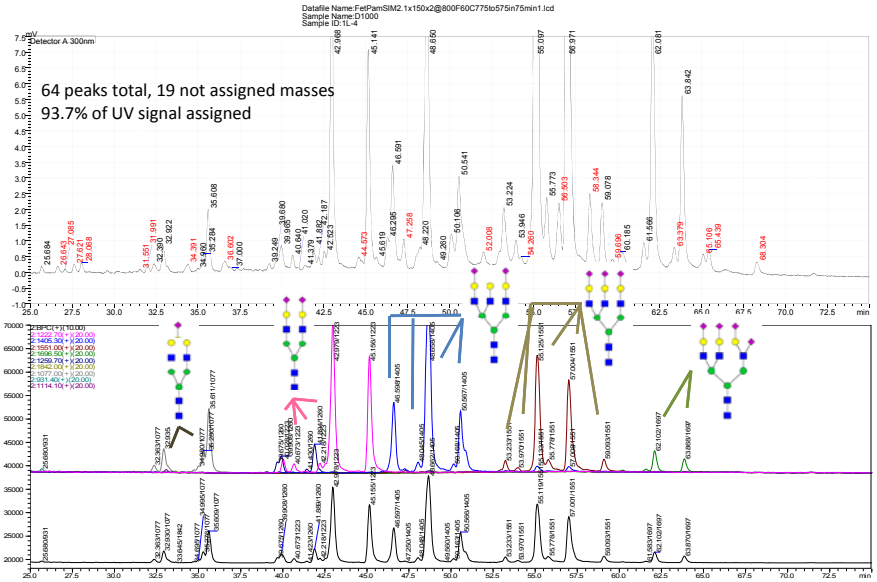
Many amines have been applied to labeling glycans, in the current work Procainamide is favored.

Standard Analysis Conditions
2.1 x 150 Penta-HILIC, 80%B to 55%B in 25 min
B:100%AcN
A: 50 mM Ammonium Formate pH 4.45 (FA titration)
0.6 mL/min, 60C
400-2000 @0.33s/0.1s each, +4.0 kV/12.5L/min, 25^{0}C DL



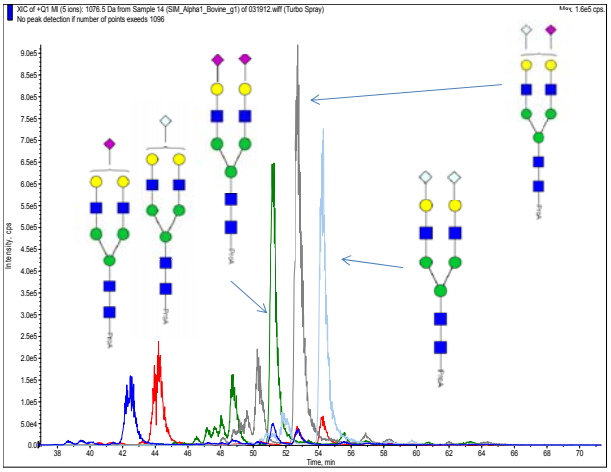
Mass: glycan + (235.325 – "O")
= Glycan + 219.32

HILIC-LC/MS of tagged glycans released from Fetuin



NeuAc vs NeuGC

HILIC-LC/MS of tagged glycans released from α1 acid glycoprotein (bovine)



Conclusions – LC-MS

HILIC of glycans tagged at the reducing termini and **RP** of per-methylated glycans both appear to be capable of resolving some of isomeric glycoforms

Additional work is needed to identify the components that are being separated.

Glycomics

- Identification from all glycans released from cells, tissue, organism, etc.
- Everything we just talked about for analyzing glycoproteins can be used to analyze mixtures of glycoproteins

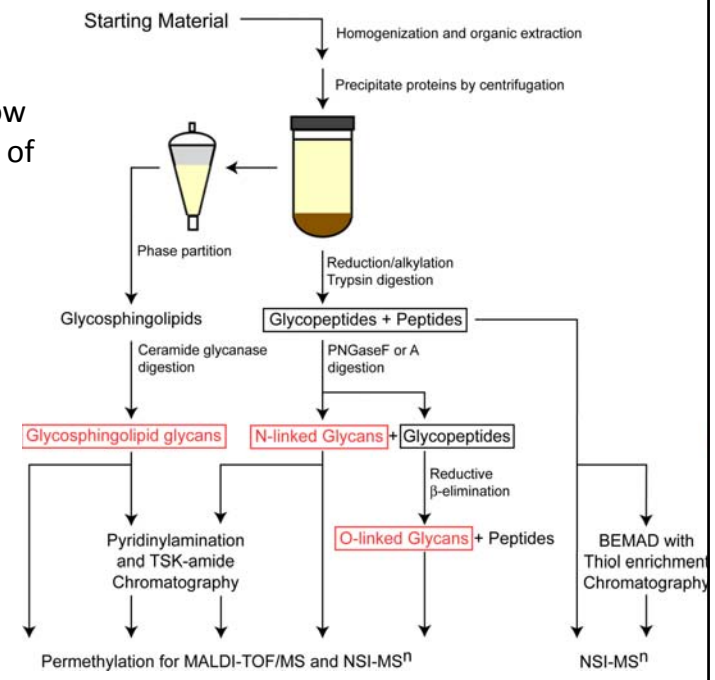
Glycomics necessitates development of an entire analytical and informatics scheme parallel to that of proteomics:

Proteomics methods are still largely inadequate but we have high-throughput analytical methods and informatics tools.

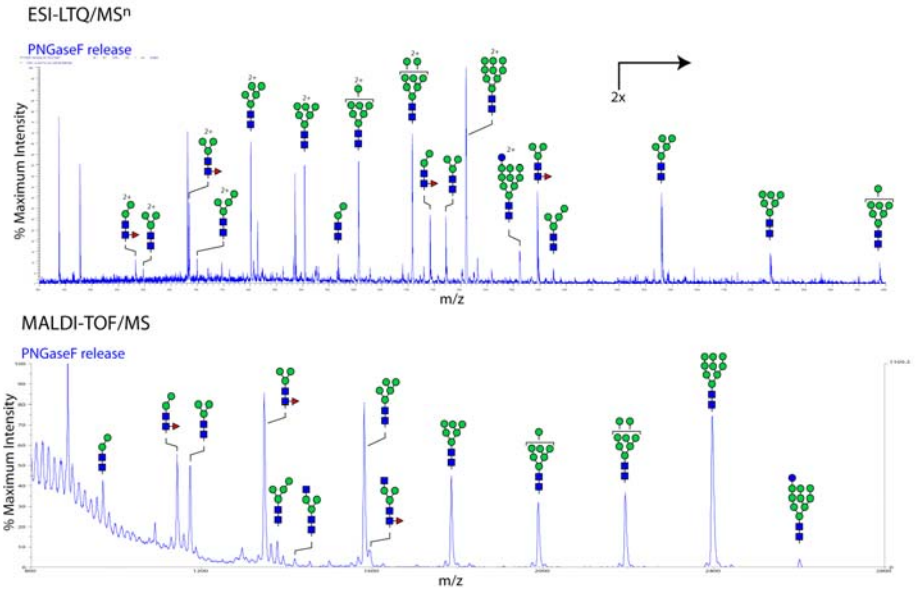
For glycomics, we still lack:

1. Adequate methods for specific detection of glycoproteins or glycopeptides in complex mixtures
2. High-throughput methods for sequence analysis of glycans
3. Adequate glycan or glycan-binding-protein microarrays
4. Automation of glycan or glycopeptide MS spectra
5. An adequate carbohydrate structure database or ontology
6. Data warehousing and curation tools for CH2O-related data
7. Database search capability for MS identification of glycans
8. Informatics tools for linkage of CH2O data with proteomics

Example of a typical workflow for the analysis of glycoprotein glycans



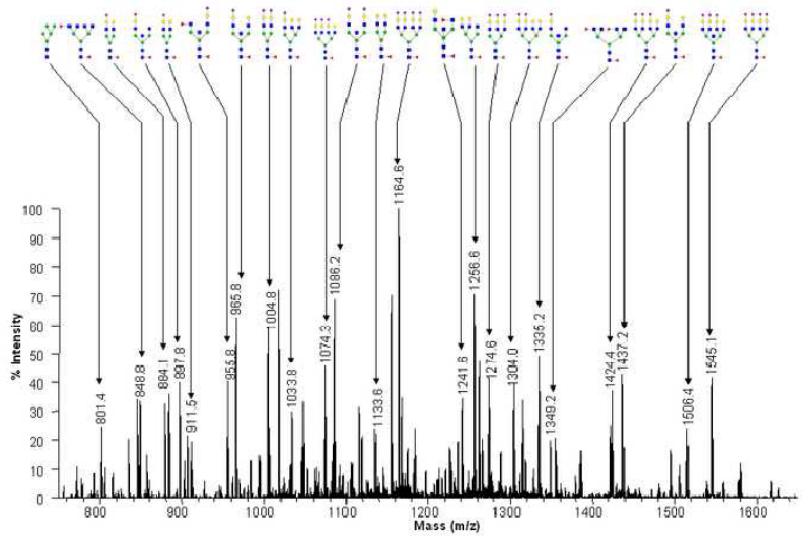
Permethylated Glycans from wildtype Drosophila embryos



K. Aoki, M. Perlman, J. Lim, R. Cantu, L. Wells, M. Tiemeyer, Journal of Biological Chemistry, 2007, 282, 9127-9142.

LTQ-FT MS/MS spectrum of N-glycans cleaved with PNGase F from rhOVGP1

Fig. 3.



Questions ?