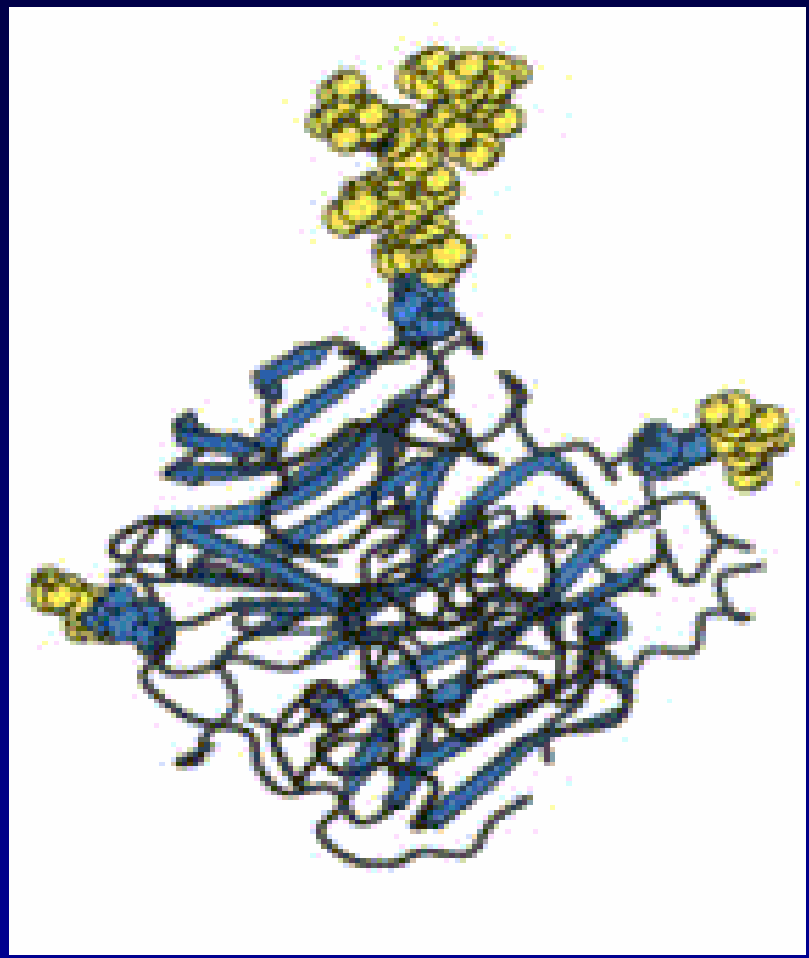
Pražské analytické centrum inovací

Projekt CZ.04.3.07/4.2.01.1/0002 spolufinancovaný ESF a Státním rozpočtem ČR

Analytical Glycobiology

Miloġ NovotnĨ

Indiana State University, Bloom USA



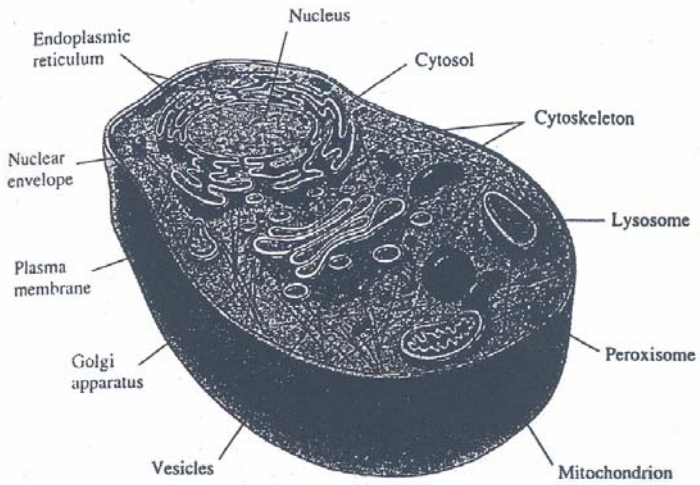
GLYCOBIOLOGY
GLYCOPROTEOMICS
FUNCTIONAL GLYCOMICS

PROTEOMICS

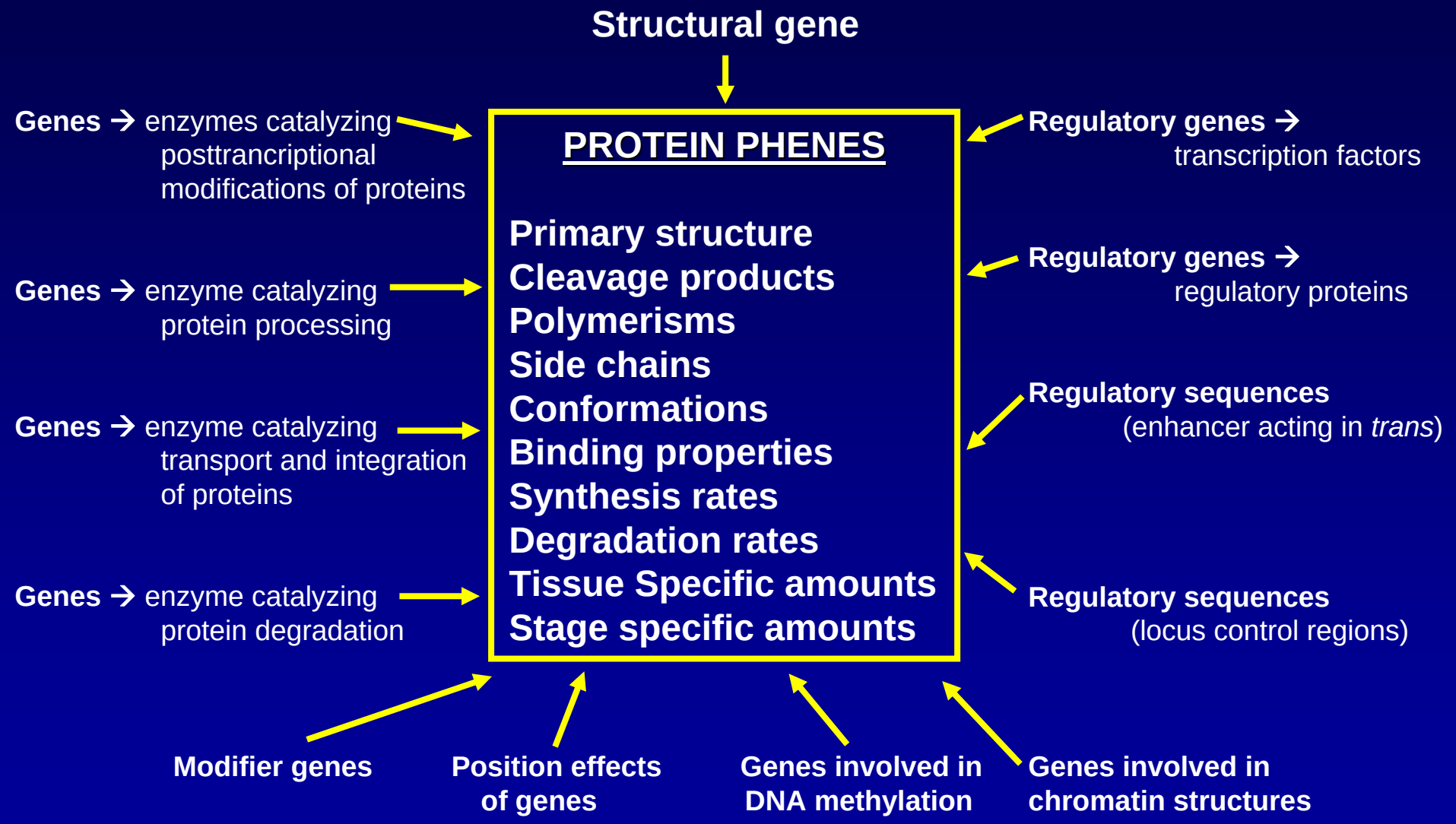
- *Science* **291**, 1221 (2001): "The analysis of complete complements of proteins. Proteomics includes not only the identification and quantification of proteins, but also the determination of their localization, modifications, interactions, activities, and, ultimately their function."

- "The work that has been done with genome sequencing may turn out to have been trivial by comparison with the challenge we now face trying to understand proteins on a grand scale"

  -- Francis S. Collins, Director of the National Human Genome Institute



The Polygenic Nature of Proteins



Klose, J. *Electrophoresis* (1999) **20**, 643-652.

GLYCOBIOLOGY

"…one of the last great frontiers of biochemistry."

G. W. Hart, *Current Opinion in Cell Biology* **4**, 1017 (1992).

"The chemistry and biology of carbohydrates has been a Cinderella field: an area that involves much work but, alas, does not get to show off at the ball with her cousins, the genomes and proteins. What has rescued this Cinderella from the shadows is no fairy godmother but a plethora of new synthetic and analytic methods that a previous generation of researchers would have found nearly magical nonetheless."

Science **291**, 2337 (2001)

"Glycobiology…how sweet it is!"

Csaba Horváth

The cell surface landscape is richly decorated with oligosaccharides anchored to proteins or lipids within the plasma membrane. Cell surface oligosaccharides mediate the interactions of cells with each other and with extracellular matrix components.

The important roles that carbohydrates play in biology and medicine have stimulated a rapid expansion of the field of glycobiology

23 MARCH 2001
"Carbohydrates and Glycobiology"
Vol. 291 (#5512) Pages 2263-2502

Glycosylation

- The most common post-translational modification of proteins: membrane bound receptors, many soluble proteins and even nuclear proteins.
- Oligosaccharides substitution occurring at asparagine residues (N-linked) and threonine or serine residues (O-linked)
- Microheterogeneities: statistical or on purpose?

Some Functions of Oligosaccharides

- Recognition markers

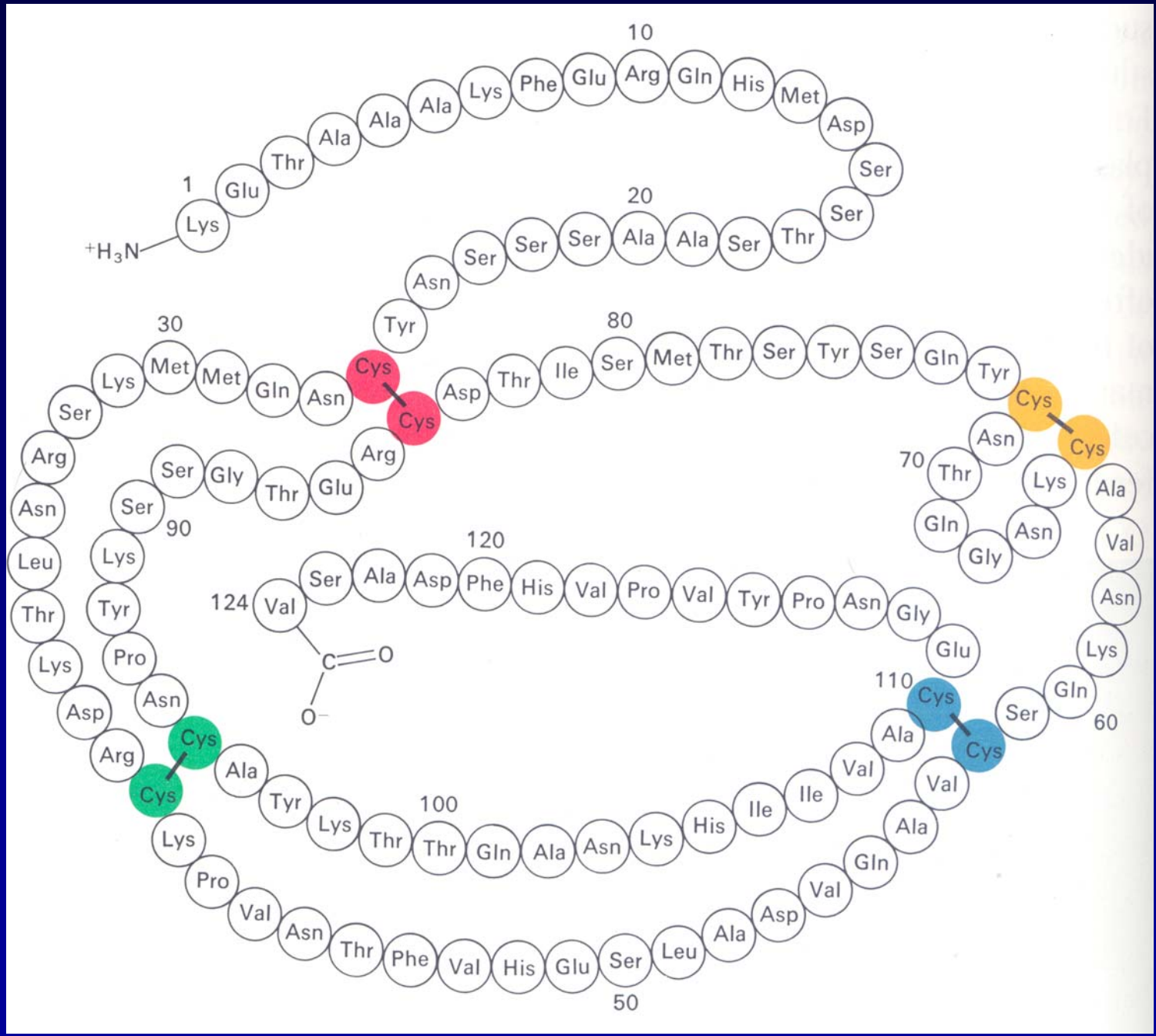
  (“recognition markers” can be put on quite different proteins without coding the information into DNA)
- Increased stability and protease resistance
- Folding and maintaining the advanced structures

GLYCOSYLATION is protein-specific, site- specific and tissue/cell-specific

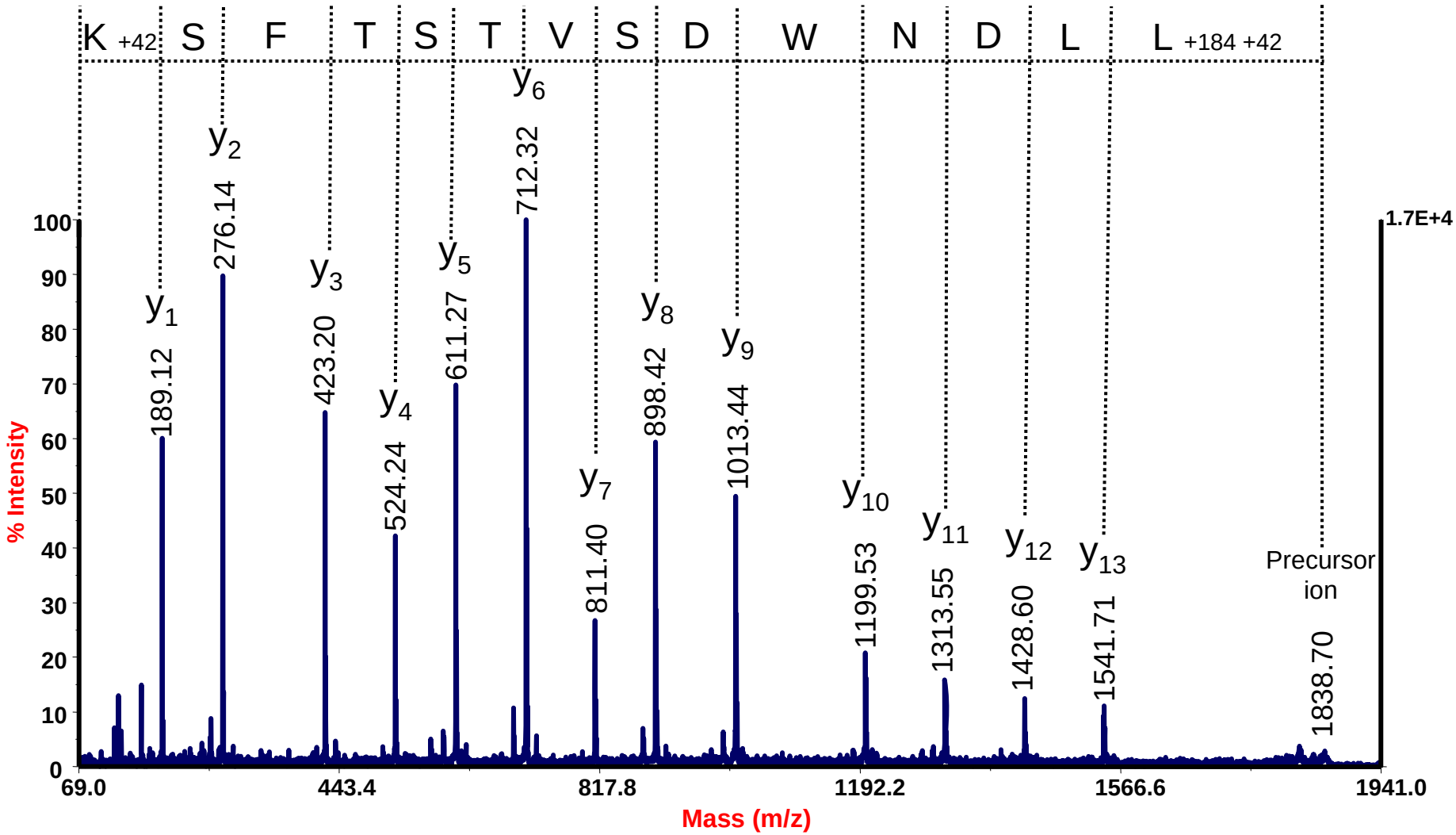
- MASS SPECTROMETRY is the key methodology to address protein and glycoconjugate sequencing, differential quantitative measurements, extent of protein modifications and, to some degree, biomolecular complexation phenomena. We utilize different ionization techniques (ESI, MALDI), different mass separation technologies (ion trap, TOF, ICR, IMS, etc.) and their tandem modes, and different detection techniques to accomplish these tasks.
- MS measurements are greatly assisted by proper combinations (on- or off-line) with modern SEPARATION METHODS, including affinity chromatography, 2-D electrophoresis, capillary LC and electrophoretic methods (CEC and CZE). They provide prefractionation and/or discrete separation of the complex mixtures of biological molecules, enhancing the SENSITIVITY of final measurements
- Extensive measurements and complex analytical data are interpreted through the extensive use of BIOINFORMATICS in both proteomic and glycomic investigations
- MICROCHEMICAL PROCEDURES (e.g., sample derivatizations or enzymatic treatments) are significant in the overall success of glycomic and glycoproteomic measurements

A BRIEF HISTORY OF PROTEIN SEQUENCING

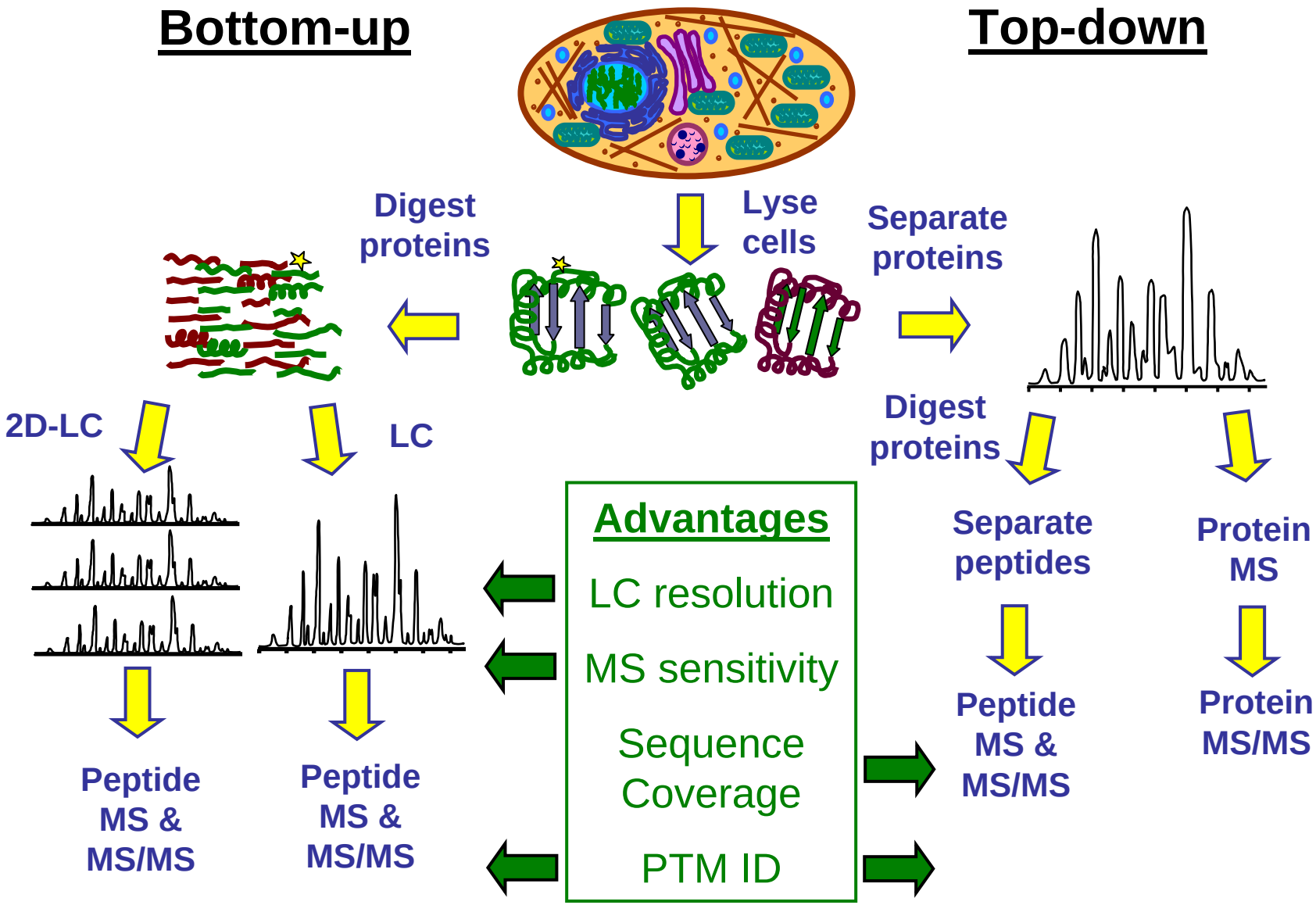
- Per Edman, the 1950s and 60s: chemical approach and chromatography.
- Klaus Biemann, since the 1960s: mass spectrometry of derivatized, small peptides (GC/MS, EI and CI)
- Introduction of FAB during the early 1980s extends the scope
- Inventions of ESI and MALDI dramatically changes our capabilities during the late 1980s.
- Gradual evolution of microseparation techniques, new MS capabilities and bioinformatics continues to this date.


10
Glu Arg Gln His Met Asp Ser Ser Thr Ala Ala Ala Lys Phe
1
Glu Thr Ala Ala Ala Lys
+H3N
Lys
20
Ser Ser Ser Ala Ala Ser Thr Ser Asn Tyr
30
Lys Met Met Gln Asn Cys Cys Asp Thr Ile Ser Met Thr Ser Tyr Ser Gln Tyr
80
Cys Cys
Asn
70
Thr Lys Ala Gln Asn Val Gly Asn
Ser Gly Thr Glu Arg Ser Ser Lys Tyr Pro Asn Cys Cys Arg Asn Leu Thr Lys Asp Arg
90
120
Ser Ala Asp Phe His Val Pro Val Tyr Pro Asn Gly Glu Lys Gln Ser
124
Val
C=O
O⁻
110
Cys Cys Ala Val Val Ala Gln Val Ala Ile Ile His Lys Asn Ala Gln Thr Thr Gln Ala Asn Lys Tyr Ala Cys Cys Lys Pro Val Asn Thr Phe Val His Glu Ser Leu Ala Asp
100
50
60

Sequencing of Peptides

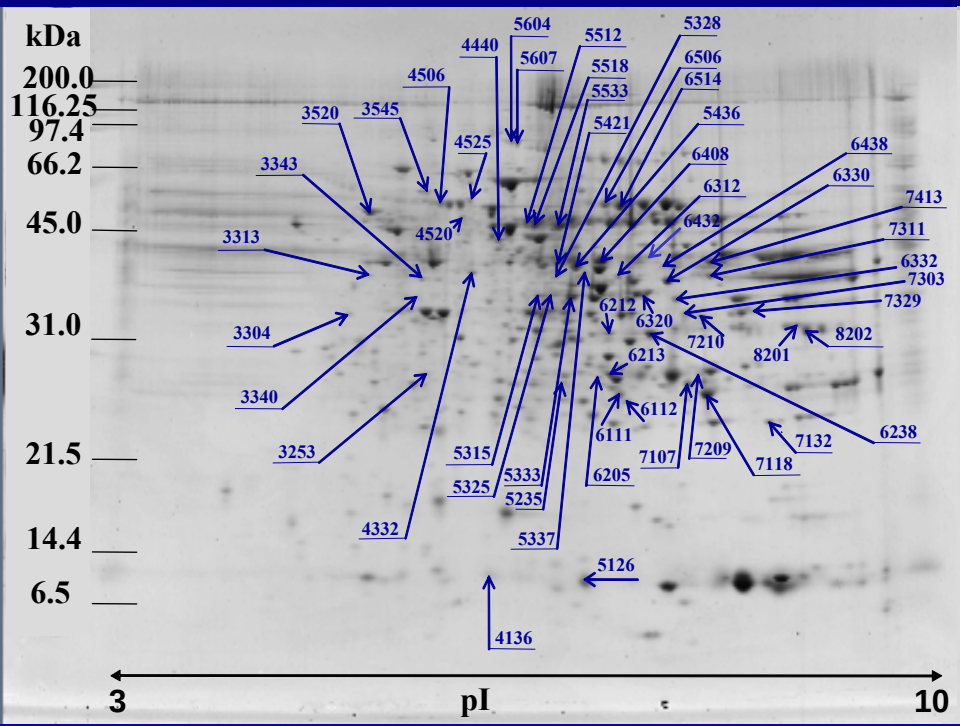


Proteomics Approaches

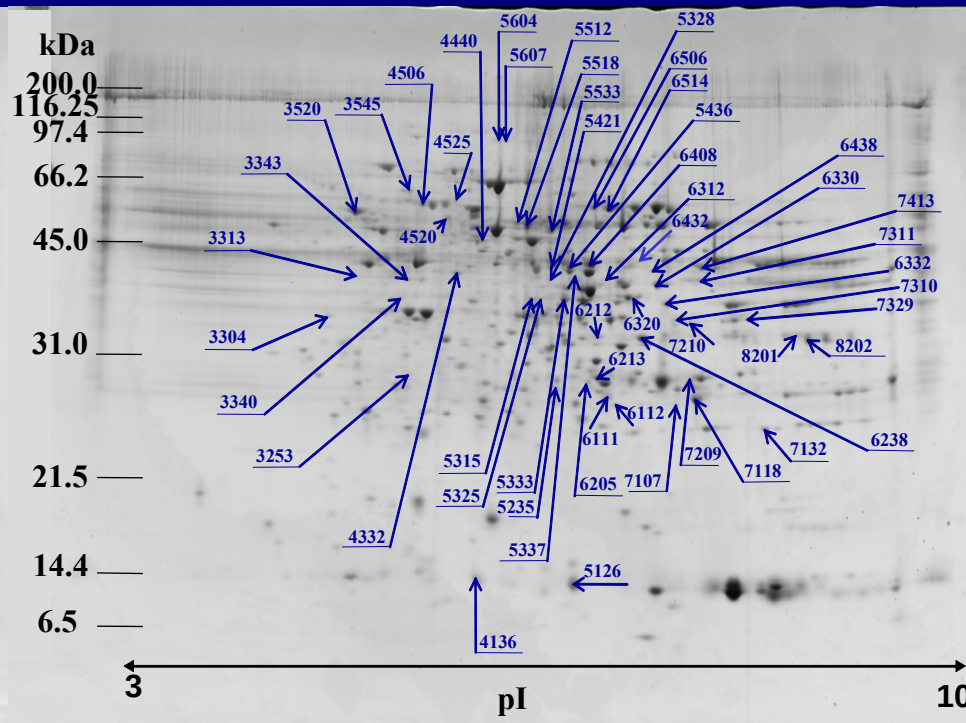


2-D gel images of Extracts from Liver Tissues of Alcohol Naïve and Alcohol Exposed Rats.

117 proteins were up- or down regulated as a result of alcohol exposure

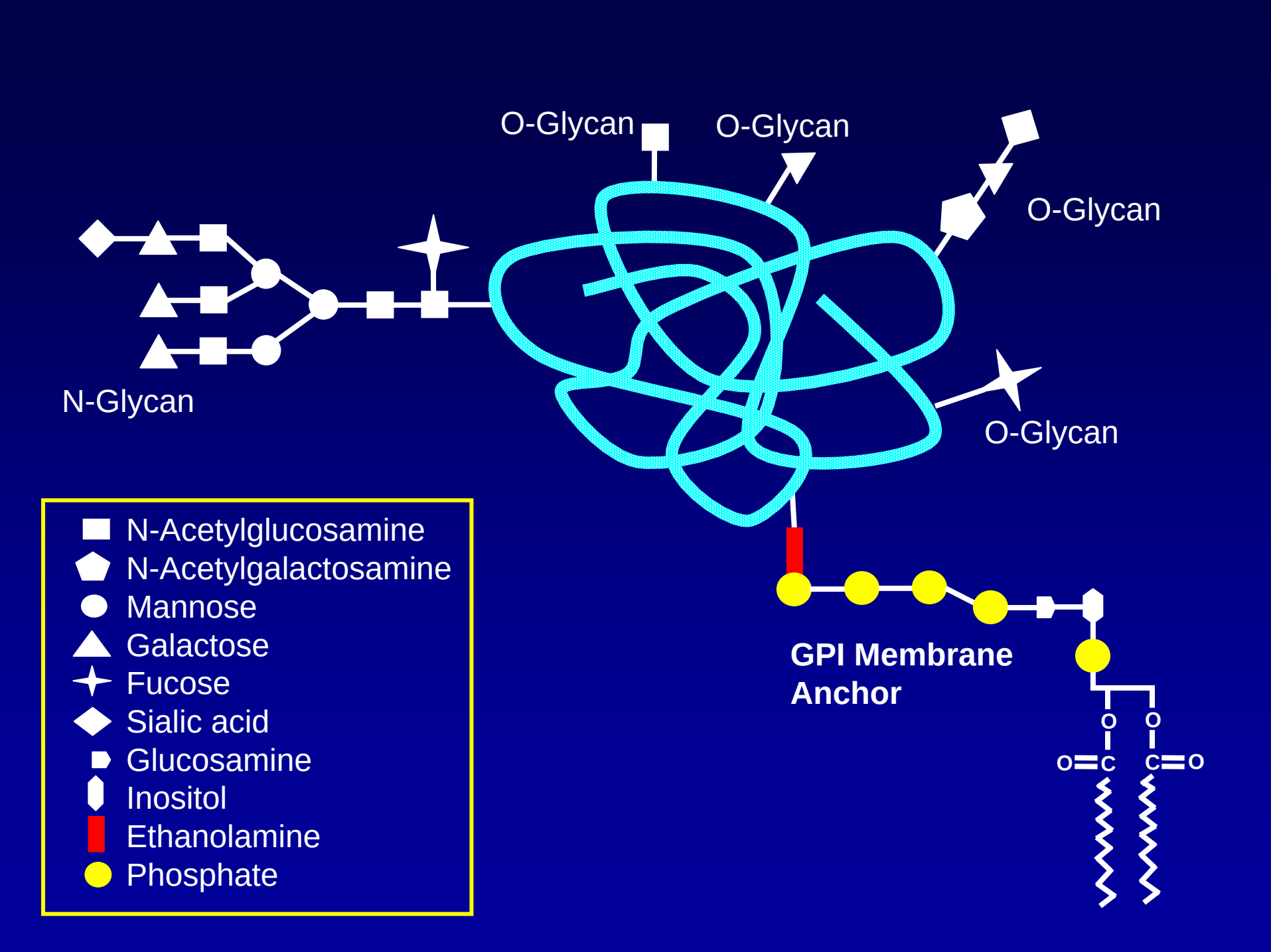


Alcohol Naïve



Alcohol Exposed

I. Klouckova, P. Hrncirova, Y. Mechref, R. J. Arnold, T.-K. Li, W. J. McBride, M. V. Novotny "***Changes in liver protein abundance in inbred alcohol-preferring rats due to chronic alcohol exposure measured through a proteomics approach***" ***Proteomics***, **2006**, 6, 3060-3074.

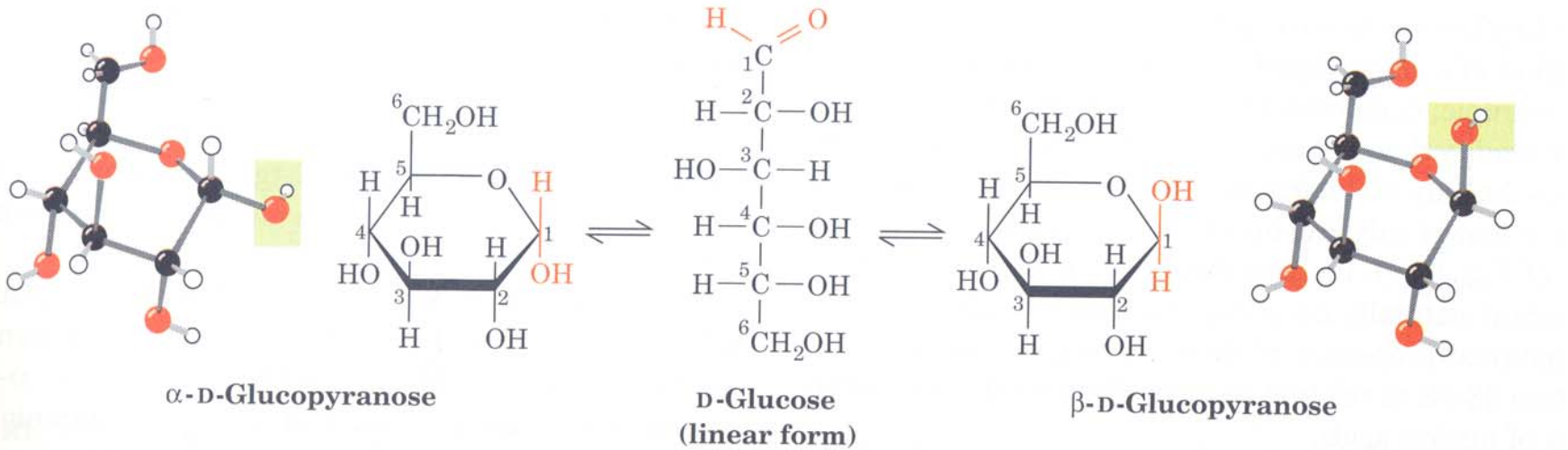

O-Glycan
O-Glycan
O-Glycan
O-Glycan
N-Glycan
N-Acetylglucosamine
N-Acetylgalactosamine
Mannose
Galactose
Fucose
Sialic acid
Glucosamine
Inositol
Ethanolamine
Phosphate
GPI Membrane Anchor

CARBOHYDRATES: The Most Abundant Class of Biological Molecules

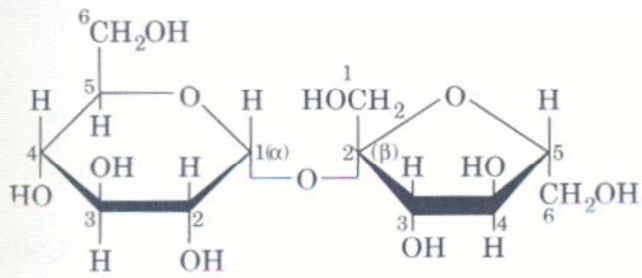
- In various polysaccharides (cellulose, starch, pectin, chitin, glycogen, hyaluronic acid, etc.), aiding architecture of living cells and serving in energy storage and metabolism
- As oligosaccharides in association with proteins (glycoproteins) and lipids (glycolipids), serving various biological function (recognition determinants)
- As glycosaminoglycans, the highly charged, unbranched polysaccharides of alternating uronic acid and hexosamine residues, which are the essential components of proteoglycans, the giant molecules of glycobiology
- As the essential components of a rapidly growing number of biocompatible materials derived from the chemical modifications of polysaccharides, and a number of biotechnologically-derived and synthetic compounds

BASIC TERMS AND NOMENCLATURE

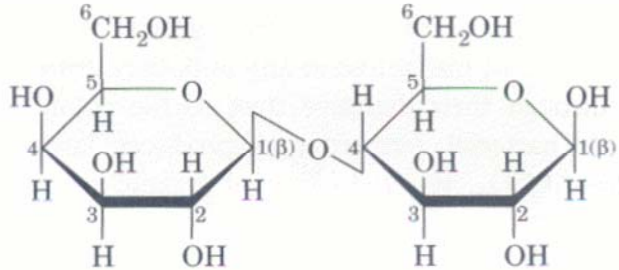
- Monosaccharides (basic units) → oligosaccharides (linear or branched structures) → polysaccharides
- Glycans and glycoconjugates
- Anomericity and different linkages between the monosaccharide units
- Stereochemical relationships (D-stereochemistry preserved from D-glyceraldehyde through hexose sugars)
- Different branched structures (e.g., biantennary, triantennary, tetra-antennary glycans)
- Different core structures (consensus sequences) for differently linked glycans



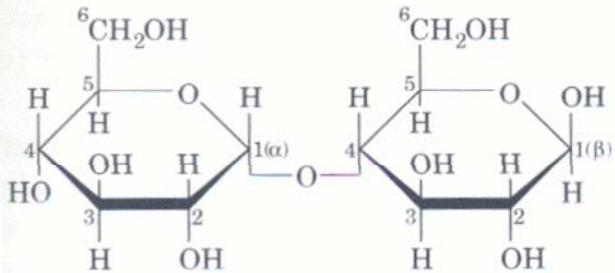
The anomeric monosaccharides α-D-glucopyranose and β-D-glucopyranose, drawn as both Haworth projections and ball-and-stick models. These pyranose sugars interconvert through the linear form of D-glucose and differ only by the configuration about their anomeric carbon atoms, C(1).



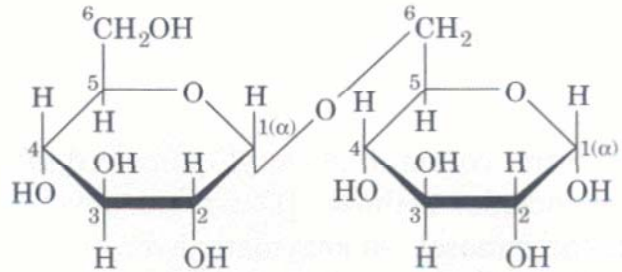
Sucrose



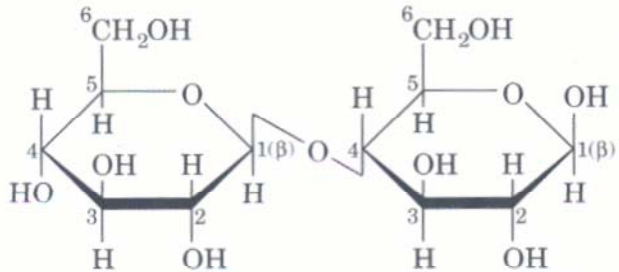
Lactose



Maltose

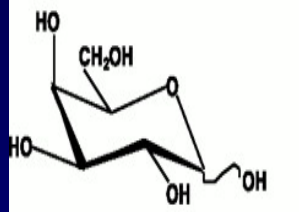


Isomaltose

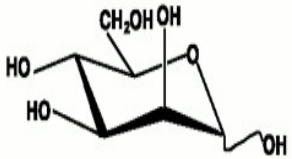


Cellobiose

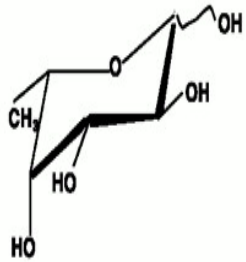
MONOSACCHARIDES COMMONLY FOUND IN GLYCOPROTEINS



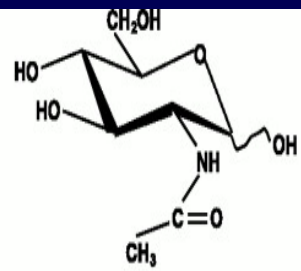
D-Galactose (Gal)



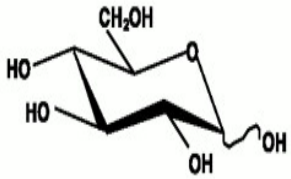
D-Mannose (Man)



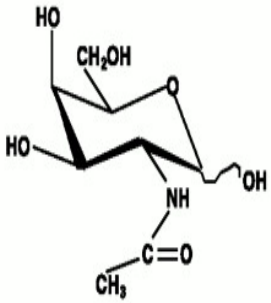
L-Fucose (Fuc)



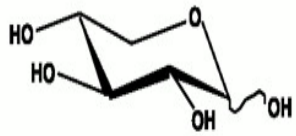
N-acetyl-D-glucosamine (GlcNAc)



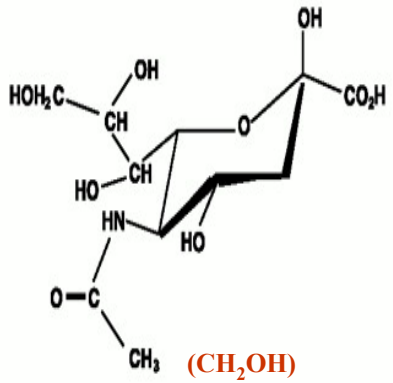
D-Glucose (Glc)



N-acetyl-D-galactosamine (GalNAc)



D-Xylose (Xyl)



(CH_2OH)

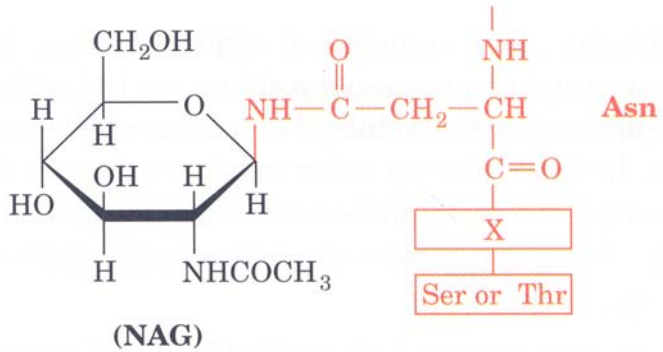
N-acetylneuraminic acid (NeuAc)

N-glycolylneuraminic acid (NeuGc)

Glycans Heterogeneity

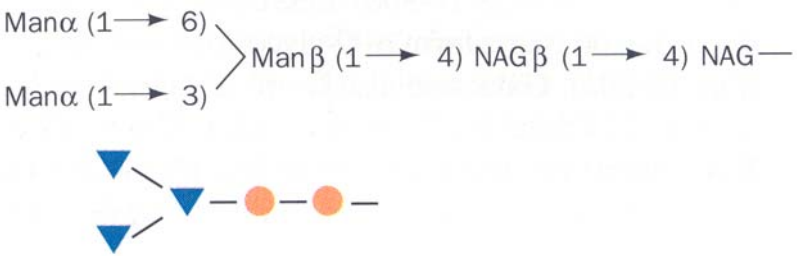
- A linear oligomer of DNA consisting of 3 monomers could have a maximum of 64 isomers.
- A protein of the same length could have 8000 isomers.
- An oligosaccharide could have **64 000** isomers.

(a)

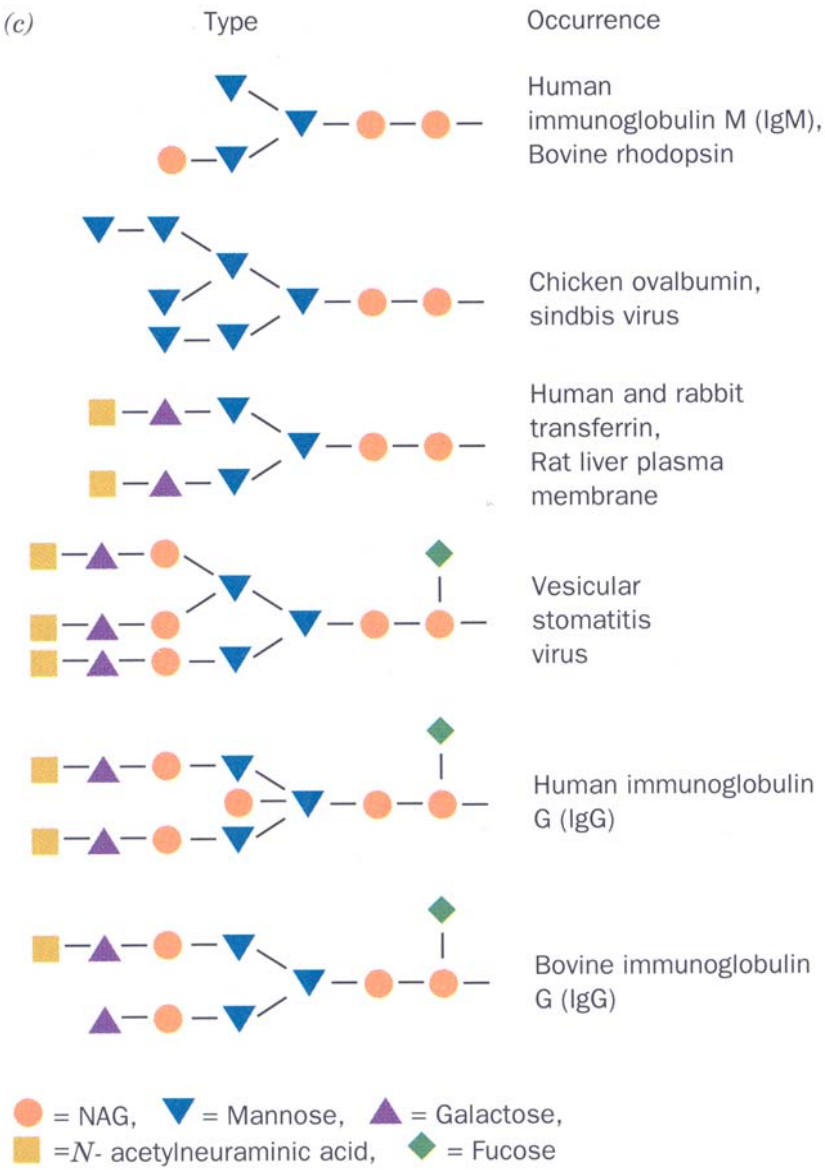

CH2OH
O
NH
H
O
NH—C—CH2—CH
Asn
H
H
OH
H
C=O
HO
H
X
H
NHCOCH3
Ser or Thr

(NAG)

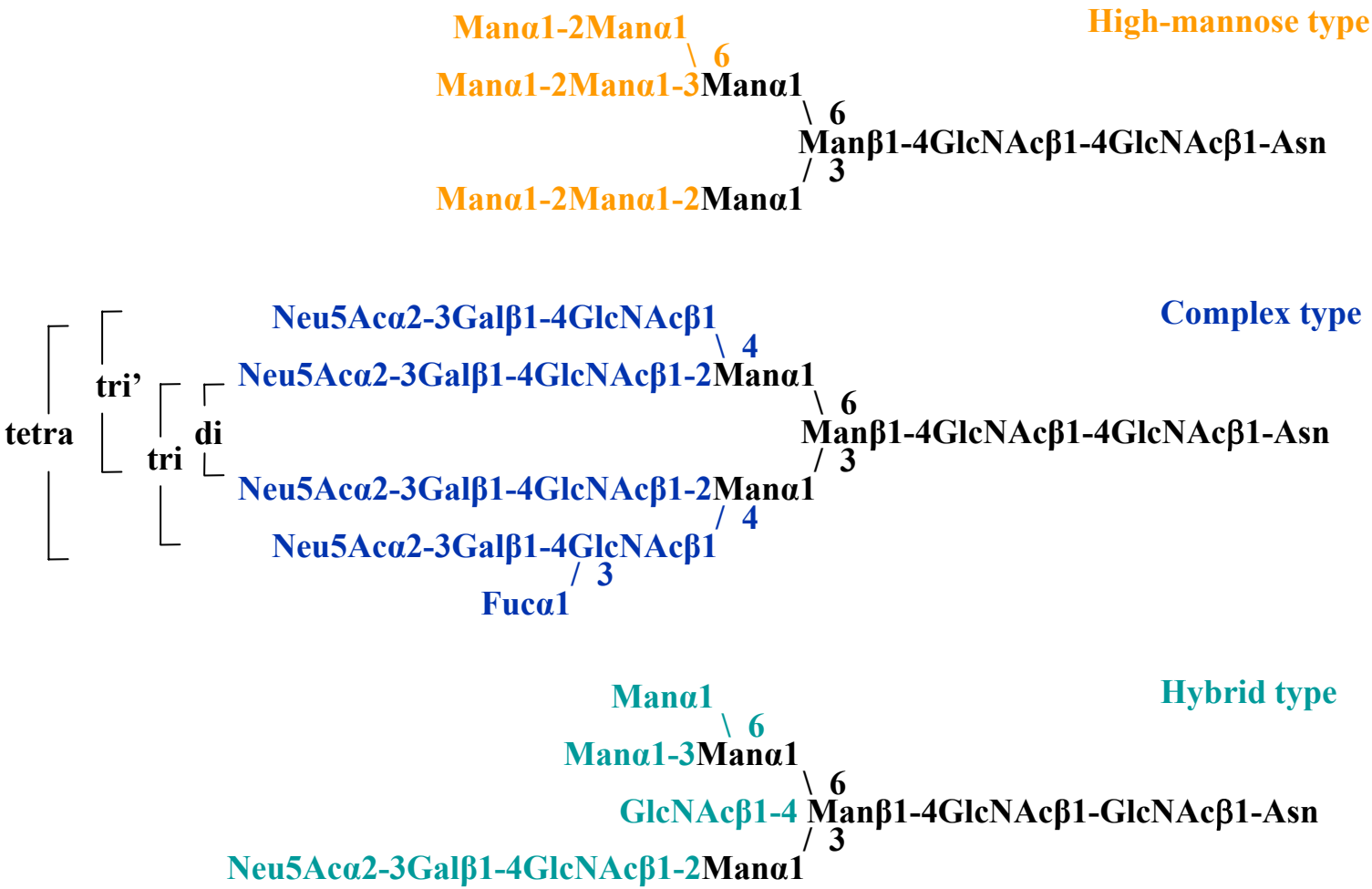
(b)


Manα (1→ 6)
Manβ (1→ 4) NAGβ (1→ 4) NAG—
Manα (1→ 3)

(c)


Type
Occurrence
Human immunoglobulin M (IgM), Bovine rhodopsin
Chicken ovalbumin, sindbis virus
Human and rabbit transferrin, Rat liver plasma membrane
Vesicular stomatitis virus
Human immunoglobulin G (IgG)
Bovine immunoglobulin G (IgG)
= NAG, = Mannose, = Galactose,
=N- acetylneuraminic acid, = Fucose

The Three Types of N-Linked Glycans



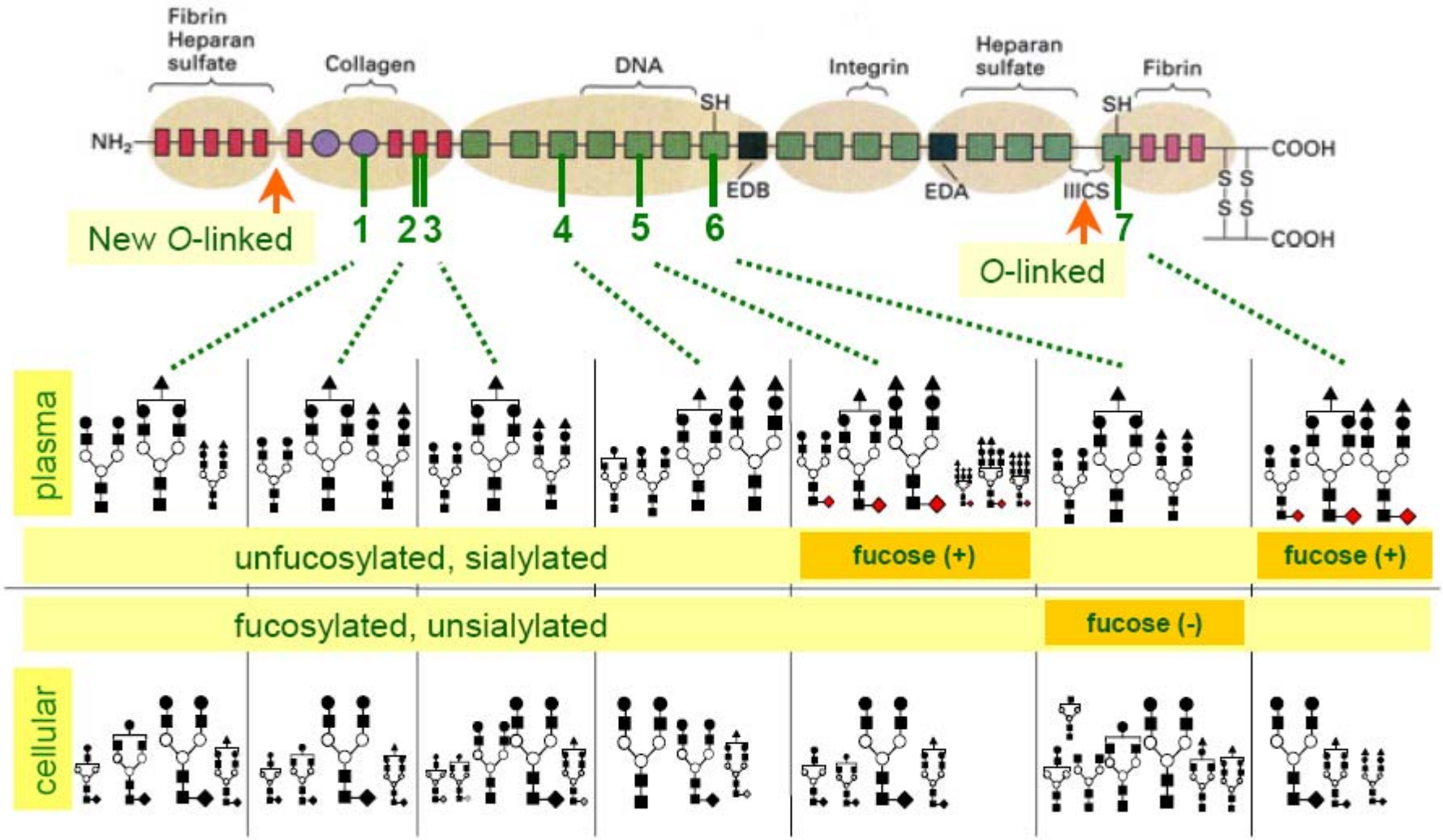
Core Structures of Mucin Type O-Linked Glycans

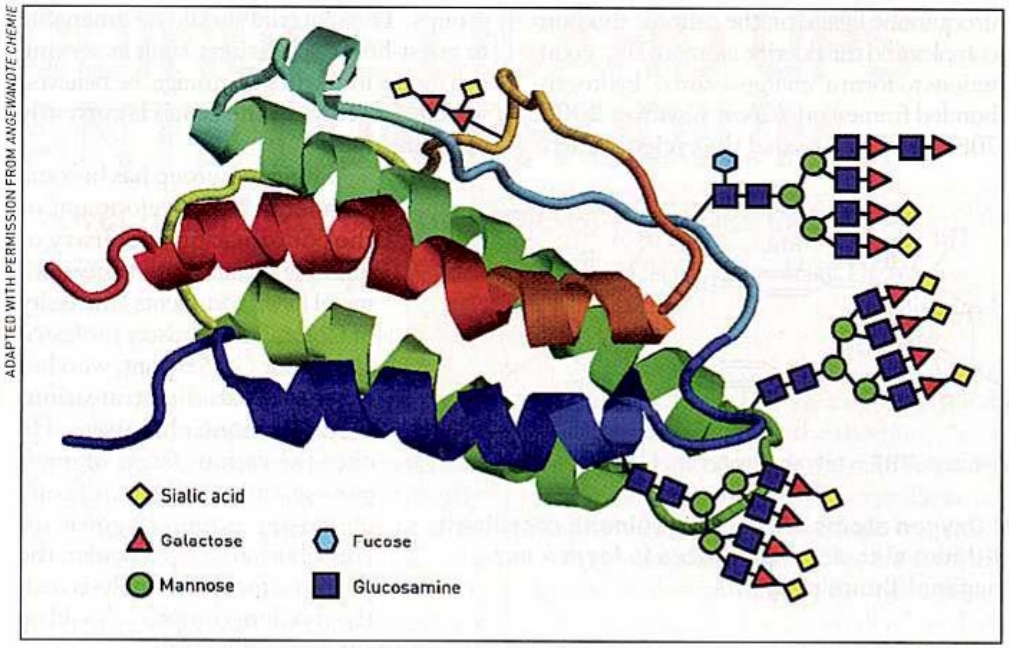
GalNAcα-Ser (Thr)

Type	Structure	Type	Structure
Core 1	Galβ1-3GalNAc	Core 4	GlcNAcβ1-6 \ GlcNAcβ1-3GalNAc
Core 2	GlcNAcβ1-6 \ Galβ1-3GalNAc	Core 5	GalNAcα1-3GalNAc
Core 3	GlcNAcβ1-3GalNAc	Core 6	GlcNAcβ1-6 \ GalNAc

Glycan heterogeneity in the *N*-glycosylation sites of plasma and cellular fibronectins

(2350 residues and seven *N*-glycosylation sites)





ADAPTED WITH PERMISSION FROM *ANGEWANDTE CHEMIE*

IN PURSUIT OF SYNTHETIC EPO

Complex erythropoietin-like glycopeptides are made from scratch

Introduction
Structural Information Needed

- Site of glycosylation
- Site occupancy
- Sequences
- Definition of branching
- Linkages and configuration
- Distinction of isobaric structures

Fragmentation of Glycans

- Fragmentation of glycans observed in MALDI/MS is similar to that observed in FAB/MS and ESI/MS and is dependent on factors such as
  - ion formation,
  - its charge state,
  - the energy deposited into an ion, and
  - the time available for fragmentation.

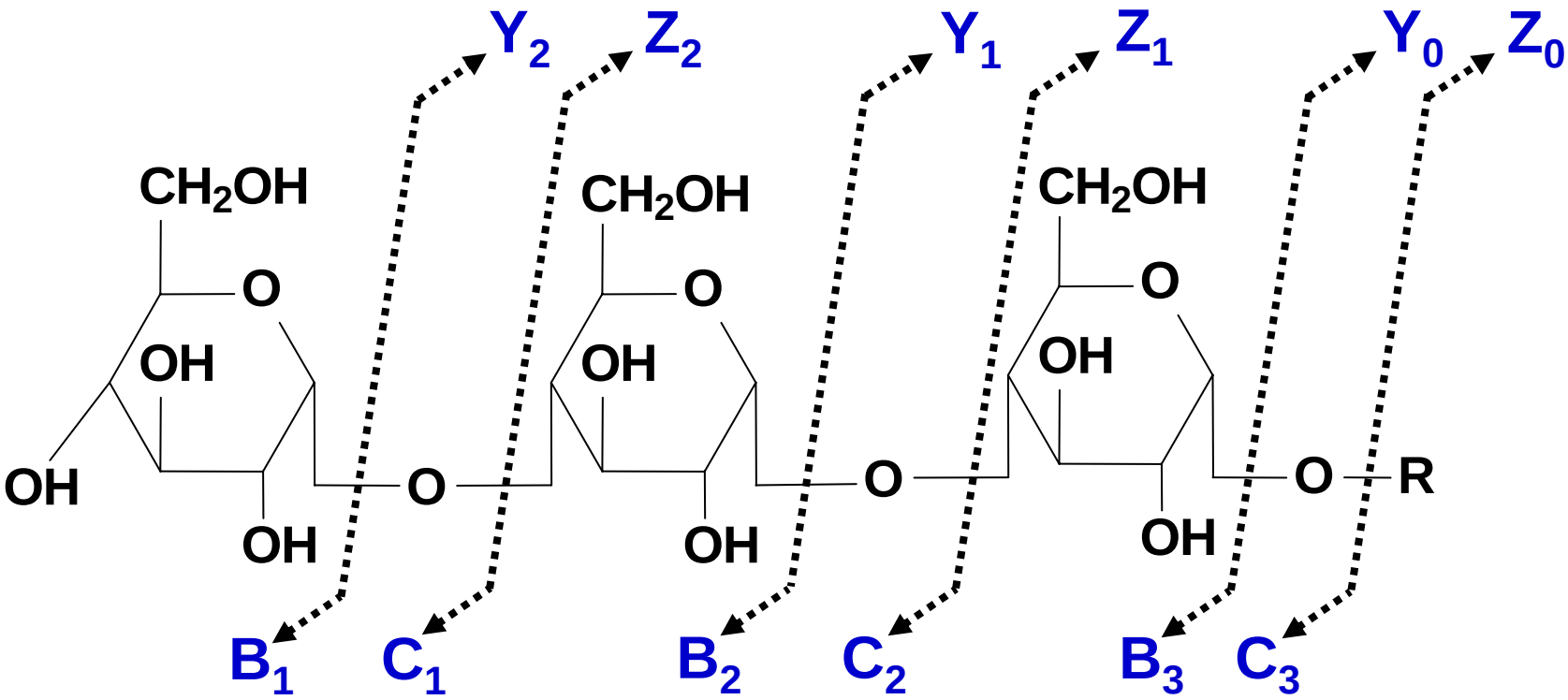
Fragmentation of Glycans in MALDI-MS

- Fragmentations in MALDI/MS can result from
  - (a) the post-source decay (PSD) which designates the fragments formed after ion extraction from the ion source,
  - (b) in-source-decay (ISD) which designates the fragments formed within the ion source, and
  - (c) collision-induced dissociation (CID) which designates the fragments formed in a collision cell filled with a gas.

Glycomic Analysis

Fragmentation of Glycans in MALDI-MS

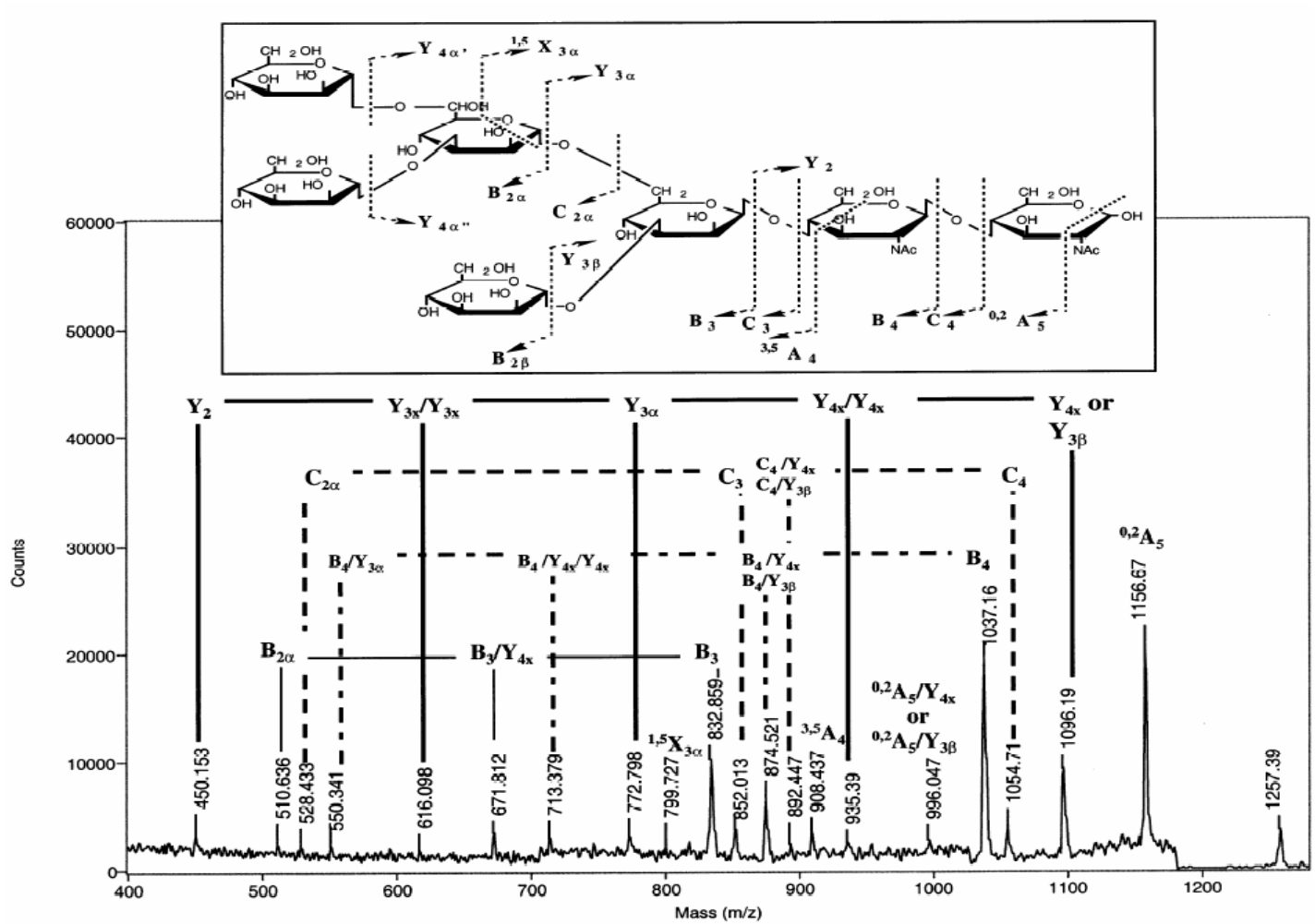
A. Glycosidic Fragmentation



Domon, B.; Costello, C. E. *Glycoconjugate J.* **1988**, *5*, 397-409.

Glycomic Analysis

Fragmentation of Glycans in MALDI-MS

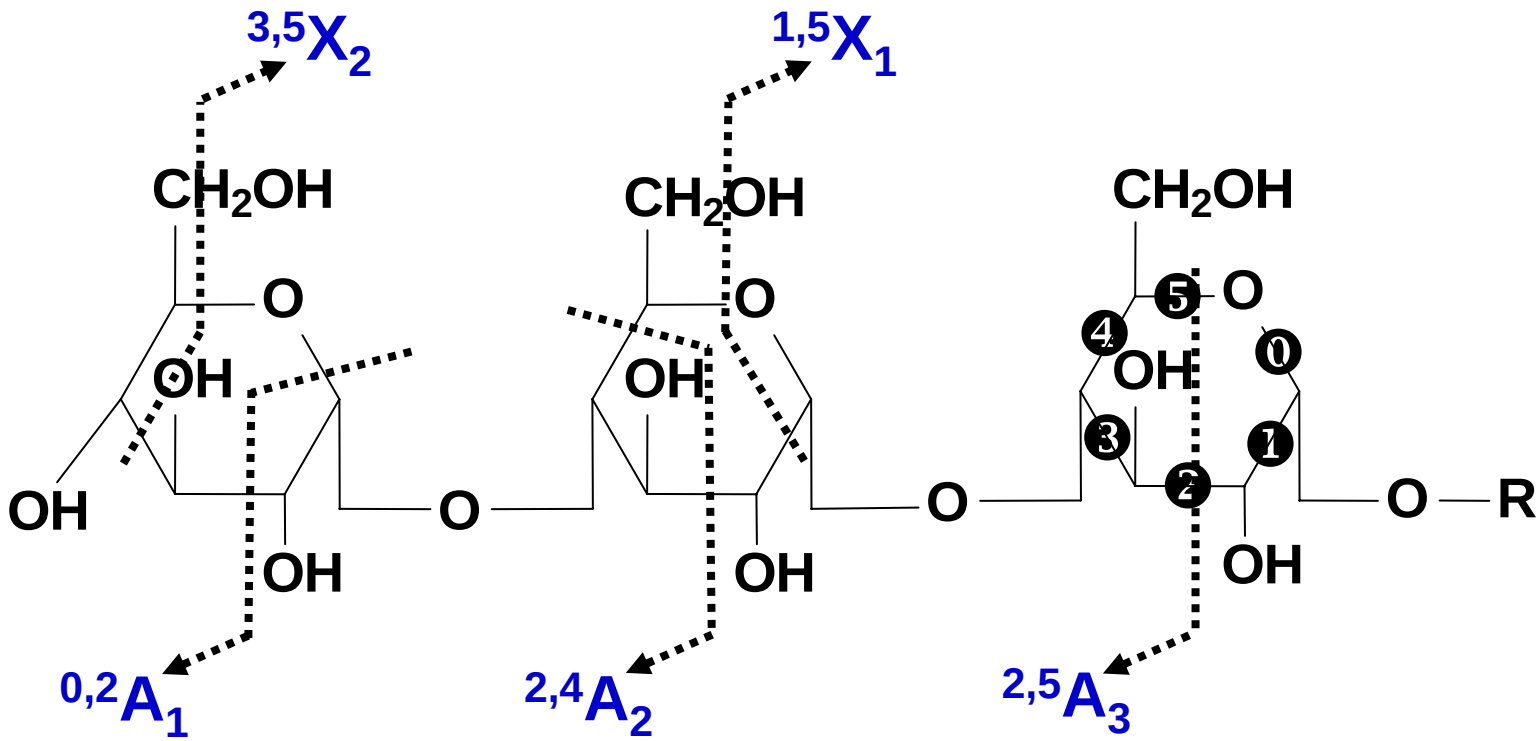


Y. Mechref, A. G. Baker, M. V. Novotny, ***Carbohydrate Res.***, 313 (**1998**): 145–155.

Glycomic Analysis

Fragmentation of Glycans in MALDI-MS

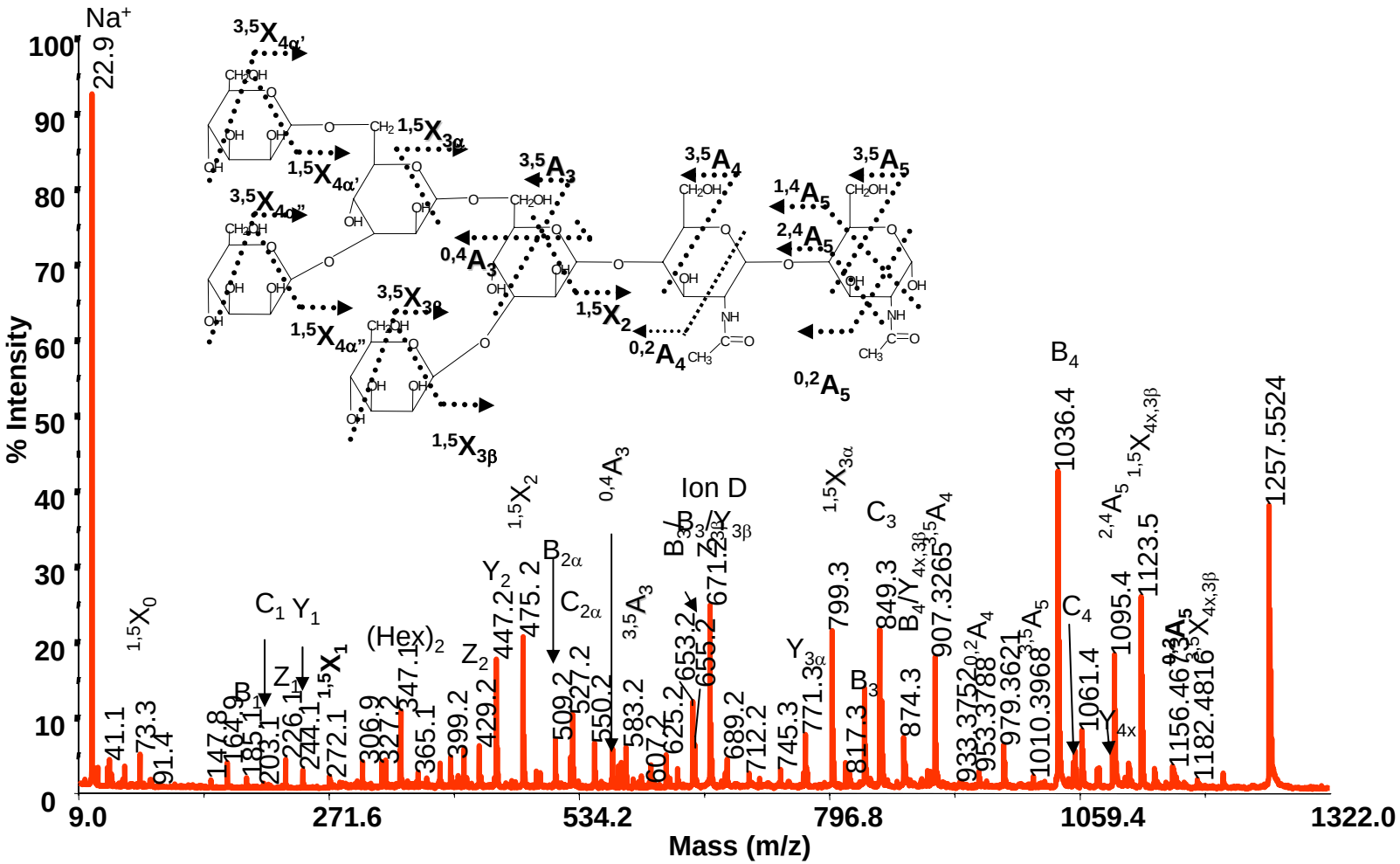
B. Cross-ring Fragmentation



Domon, B.; Costello, C. E. *Glycoconjugate J.* **1988**, *5*, 397-409.

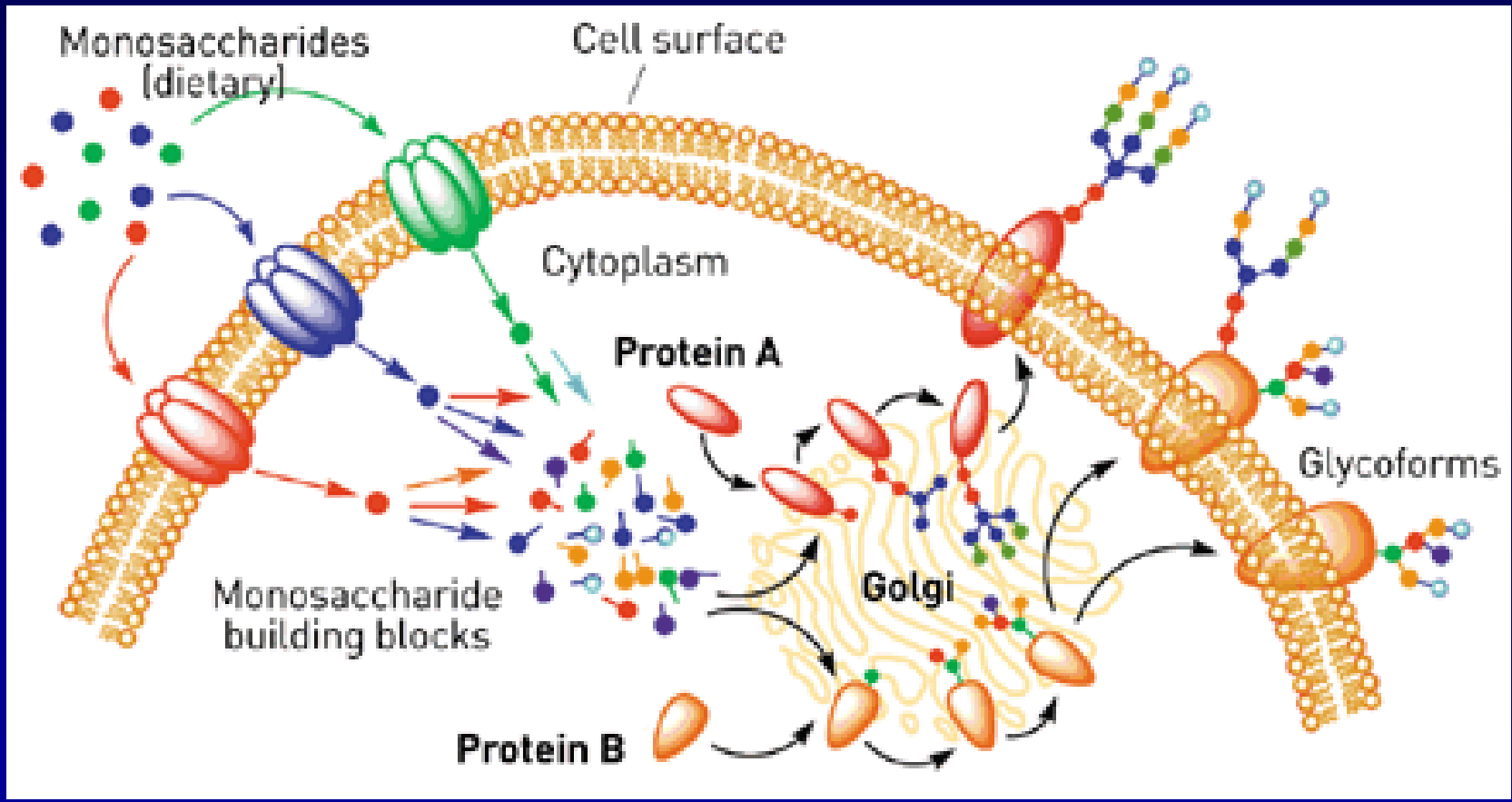
Glycomic Analysis

Fragmentation of Glycans in MALDI-MS

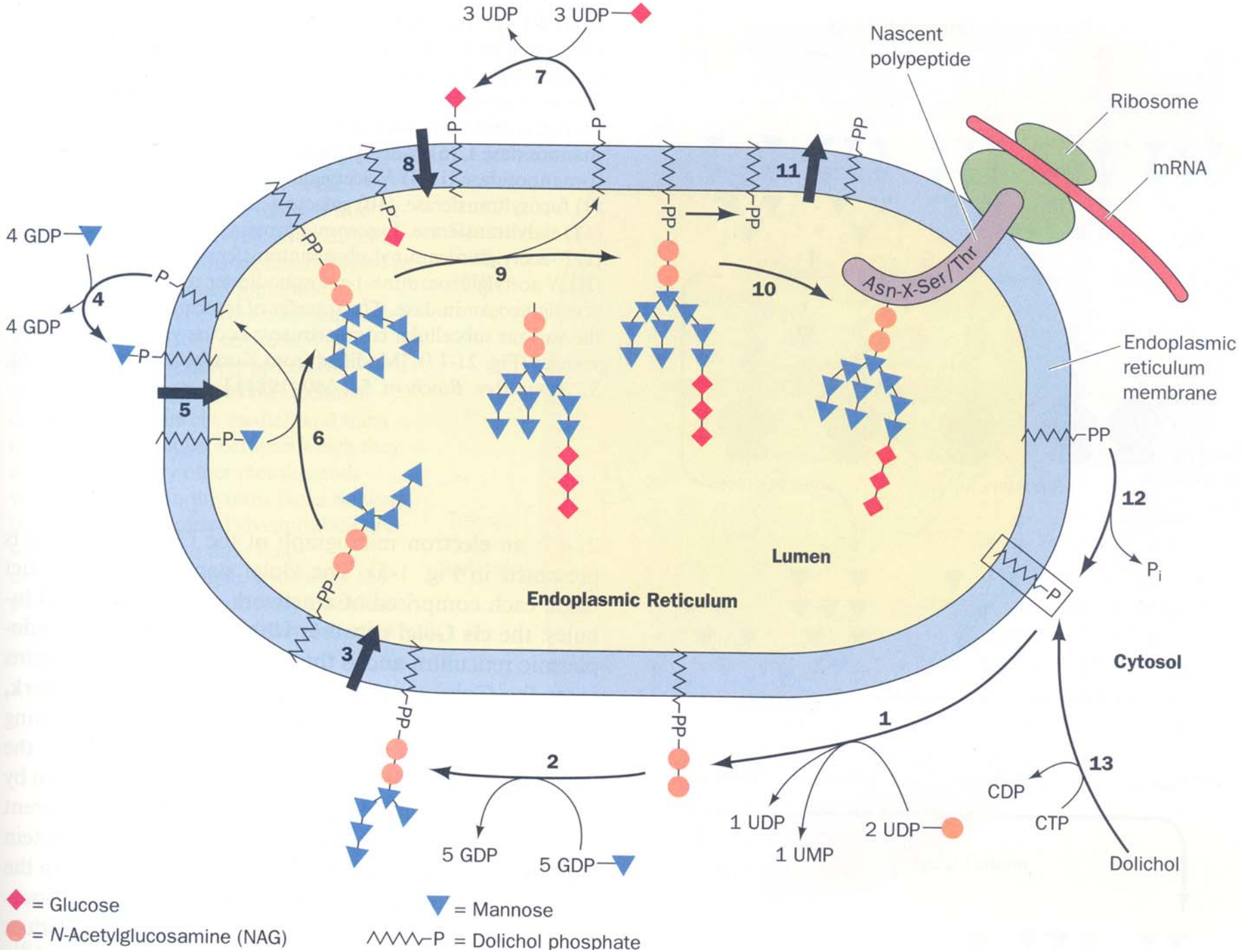


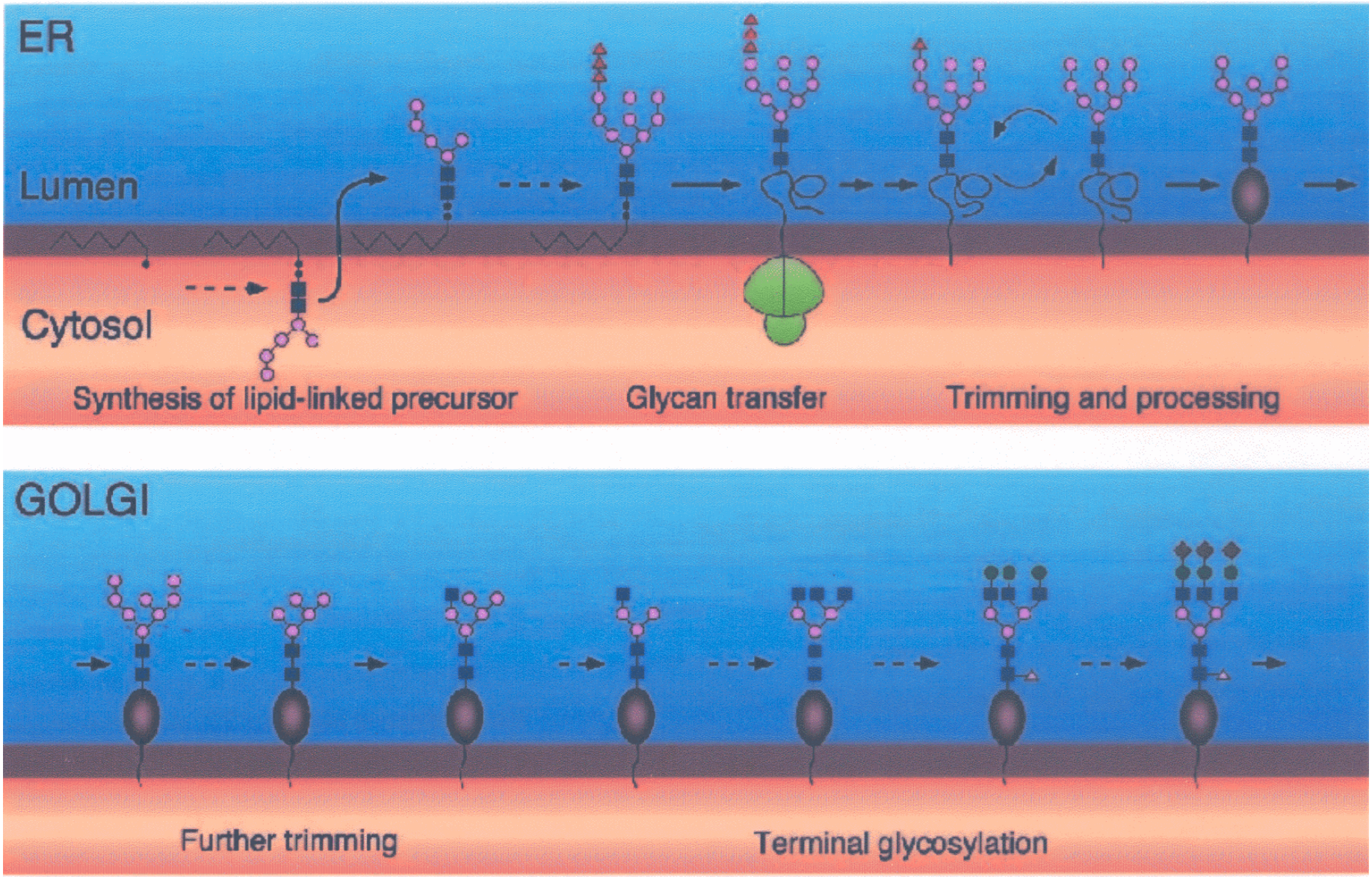
Yehia Mechref, Milos V. Novotny and Cheni Kirshnan, ***Anal. Chem.***, ***Anal. Chem.***, 75 (**2003**), 4895-4903.

Glycoconjugate Biosynthesis

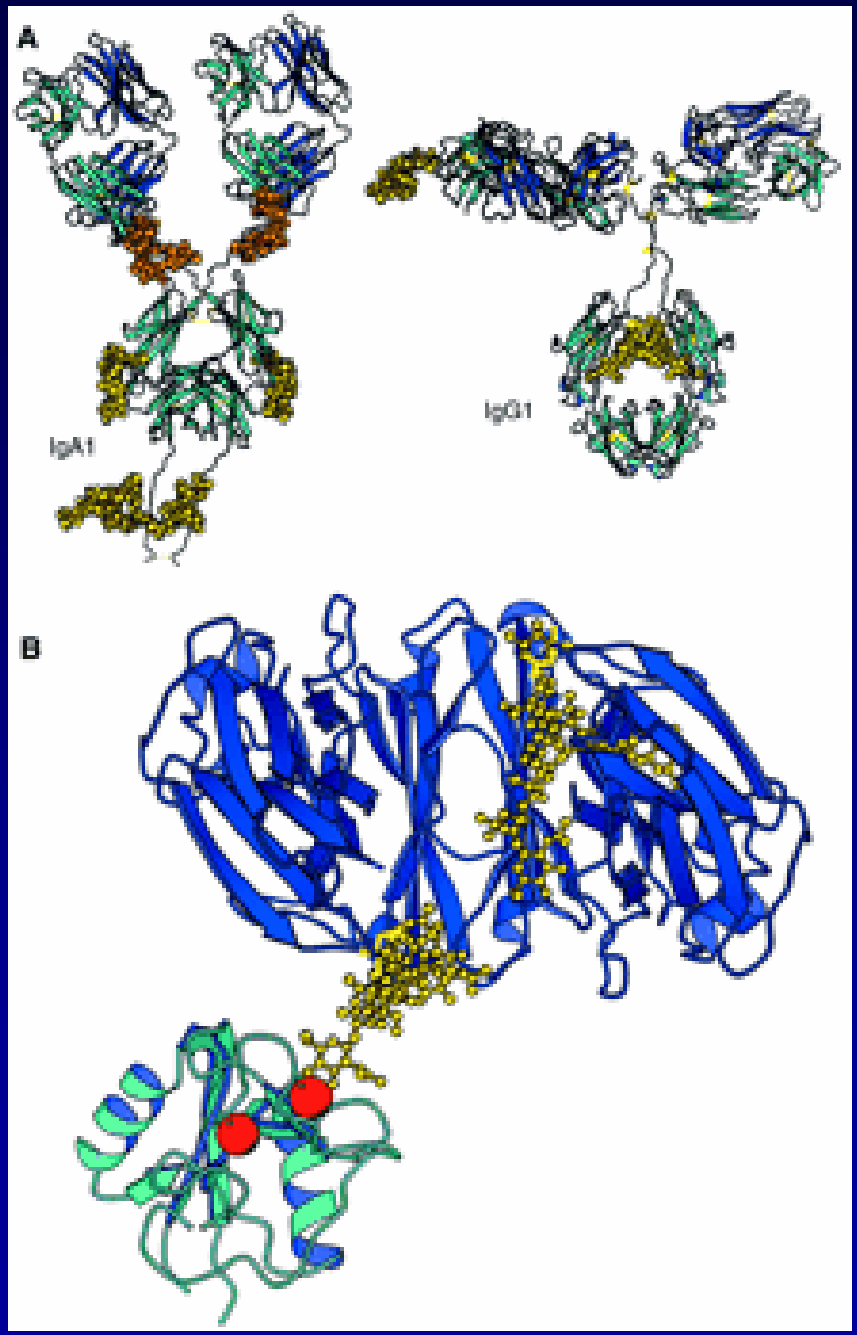


Carolyn R. Bertozzi and Laura L. Kiessling, *Science*, 291 (2001) 2357-2364


3 UDP
3 UDP
7
Nascent polypeptide
Ribosome
mRNA
8
11
4 GDP
4 GDP
4
9
10
Asn-X-Ser/Thr
Endoplasmic reticulum membrane
5
6
12
Lumen
Endoplasmic Reticulum
Cytosol
3
1
2
13
5 GDP
5 GDP
1 UDP
1 UMP
2 UDP
CDP
CTP
Dolichol
= Glucose
= N-Acetylglucosamine (NAG)
= Mannose
= Dolichol phosphate



A. Helenius and M. Aebi, *Science* **291**, 2365 (2001)



Yellow color = N-linked glycans
Orange color = O-linked glycans

Model of the interaction of IgA1 with a carbohydrate recognition domain of a different protein

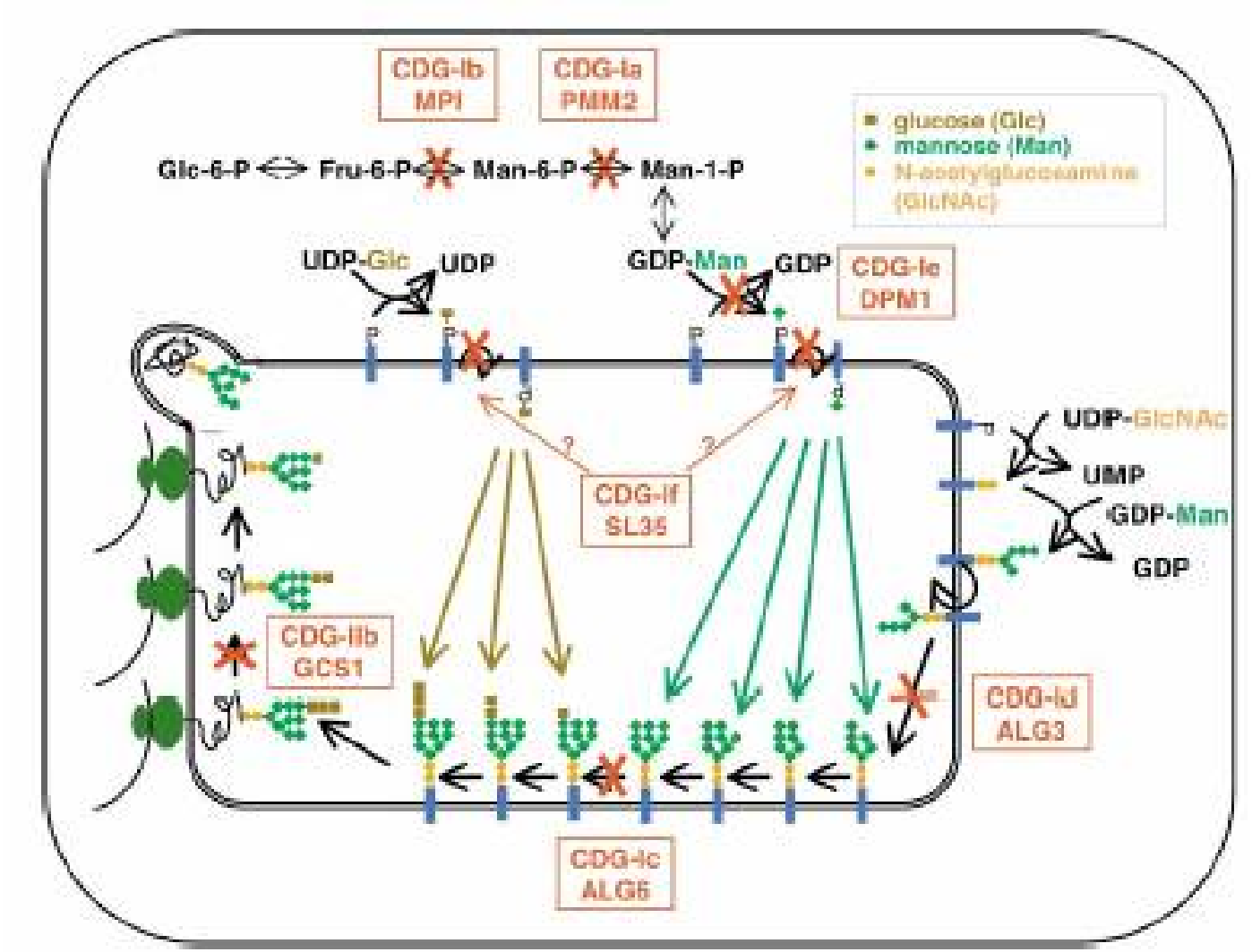
P.M. Rudd, T. Elliott, P. Cresswell, I.A. Wilson and R.A. Dwek, *Science* **291**, 2370 (2001)

Some Glycoconjugate Disease Associations

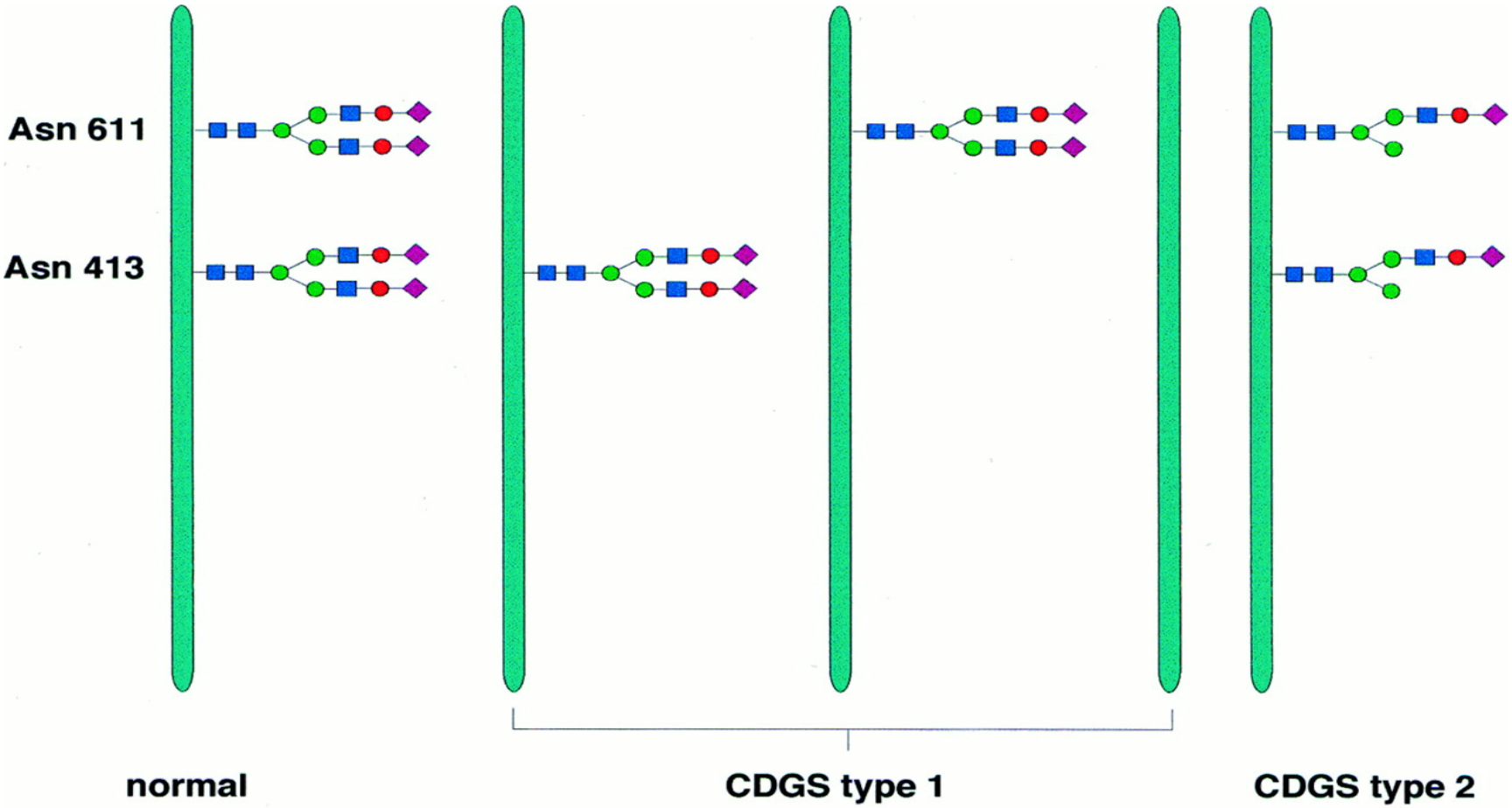
- immunity to infectious diseases, including HIV
- rheumatoid arthritis (altered composition of IgG and levels of the serum mannose-binding protein)
- prion diseases
- congenital disorders of glycosylation (rare, usually resulting in CNS impairment)
- oral pathologies
- cystic fibrosis
- heart pathologies
- cancer

CONGENITAL DISORDERS OF GLYCOSYLATION (CDG) AS EXAMPLE

- Related to errors in fundamental glycosylation machinery; also, relevant to various aspects of developmental biology where biosynthesis, monosaccharide additions or processing may differ
- More than 20 CDGs now identified (most, but not all, involve N-glycans)
- In some CDGs, multiple pathways seem involved; disruption of traffic in Golgi glycosylation machinery
- At their recent meeting in Osaka, Japan, Human Disease Glycomics/Proteome Initiative (HGPI) directors have suggested that CDGs be employed as a model for unifying methodologies in functional glycomics



Jaeken and Matthijs, *Annu. Rev. Genomics Hum. Genet*. 2: 129-51 (2001)

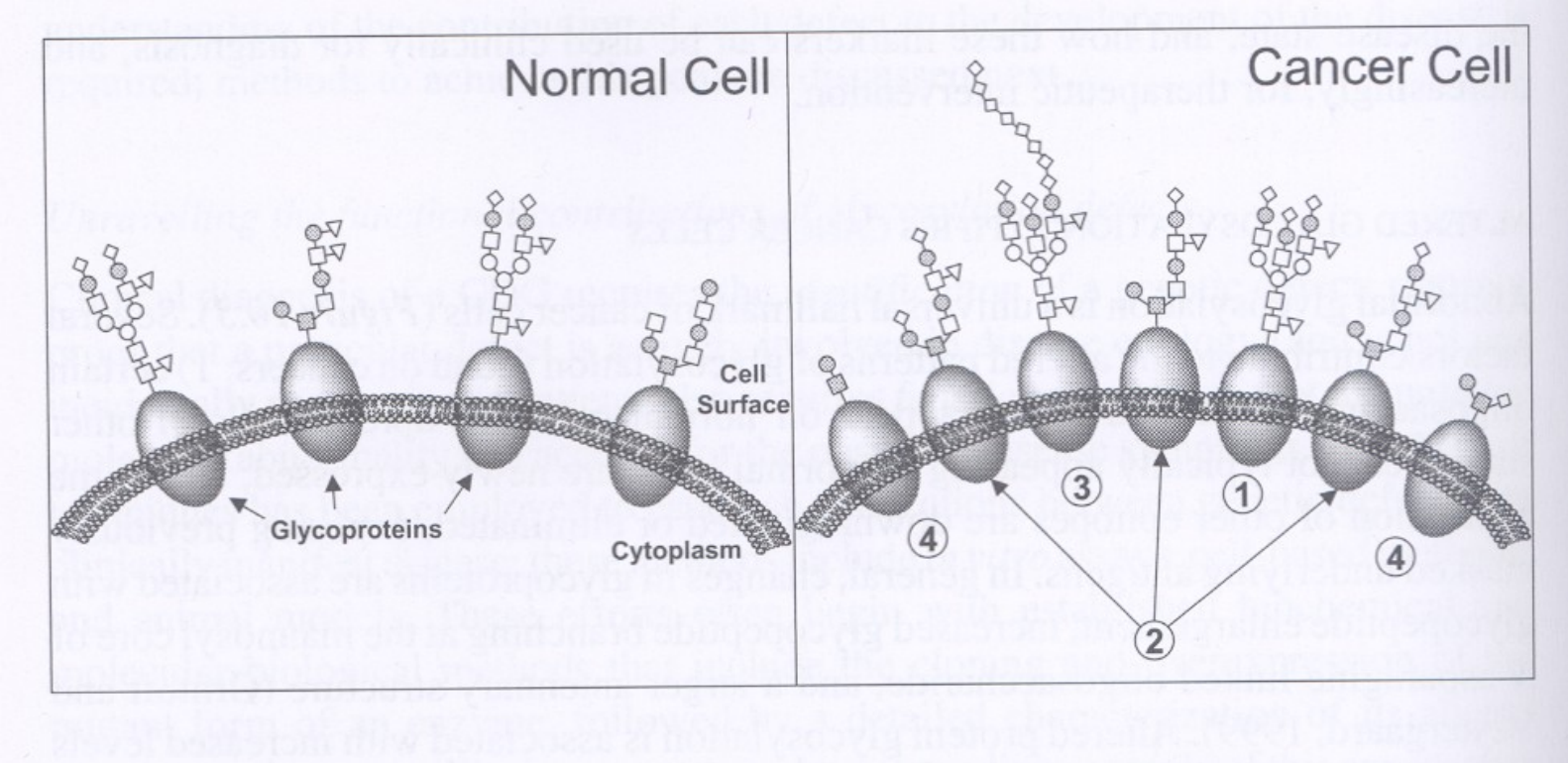


A. Dell and H. Morris, *Science* 291, 2001

GLYCOSYLATION PERTAINING TO CANCER: REPRESENTATIVE CHANGES

- Increased branching and sialylation of N-linked glycans; fucosylation
- Increased sialylation of O-linked glycans
- Occurrence of polysialic acid
- Truncation of some O-linked structures

Most current emphasis on O-linked structures = methodological challenges



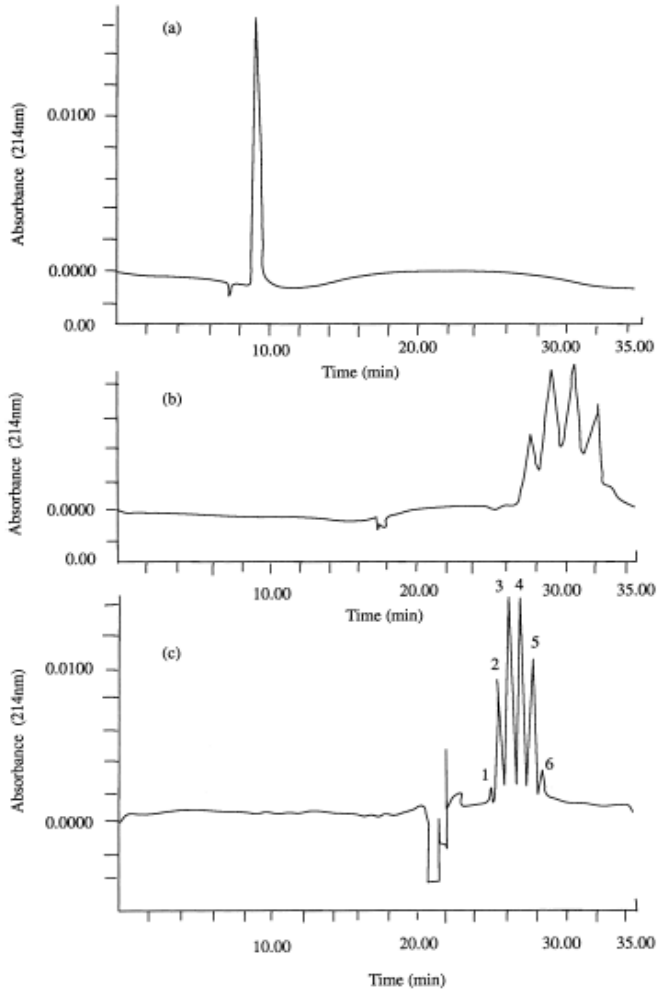
L. Tong, et al. in Biotechnology & Genetic Engineering Reviews, Vol 20, p. 214 (2003)

Isolation of Glycoproteins

- Resolution of isoforms in gels and CE
- Isolation schemes
- Use of lectins
- Examples of isolation/structural determinations

Analysis of Glycoproteins
Capillary Electrophoresis

Effect of additives on the CZE separation of rHuEPO: Sample: 1 mg/ml; fused-silica capillary (50 cm_75 mm i.d.); voltage: 10 kV. Buffers at pH 6.2: (A) 10 mM tricine/10 mM NaCl; (B) 10 mM tricine/10 mM NaCl/2.5 mM 1,4-diaminobutane; (C) 10 mM tricine/10 mM NaCl/2.5 mM 1,4-diaminobutane/7 M urea. UV detection at 214 nm.



E. Watson, F. Yao, ***J. Chromatogr.*** 630 (1993) 442.

Isolation of Glycoproteins

General Considerations

- Glycoproteins are often encountered in minute quantities in biological materials such as cellular extracts and physiological fluids
- There is high demands placed on both the measurement sensitivity and proper isolation procedures.
- A combination of orthogonal separation techniques and the use of affinity principles are the most commonly practiced isolation/fractionation strategies.
- Miniaturization of these separation and isolation/fractionation methodologies represents a general trend in glycoanalysis.

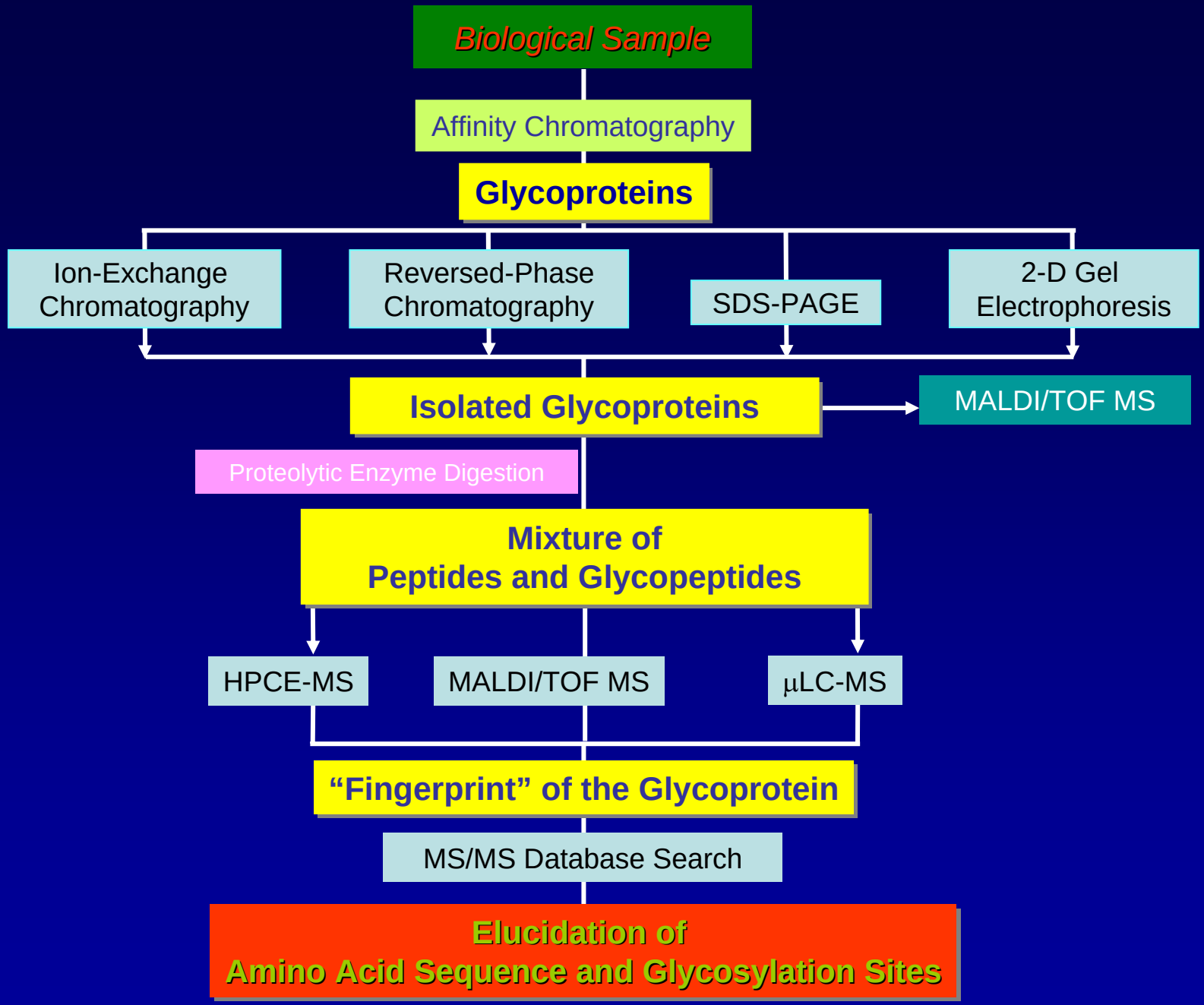
Isolation of Glycoproteins

General Considerations

- Glycoproteins at the low-microgram scale, while becoming measurable with the modern instrumental techniques, can easily be adsorbed on the surface of glassware before such measurements.
- Sample loss during ultrafiltration, dialysis, lyophilization, etc., can easily become a bottleneck of the entire analysis.
- Another problem with working at such a reduced scale is contamination (dust, solvent, reagent impurities, etc.).
- It is thus crucial to minimize the number of handling and transfer steps during the analysis.
- Miniaturized forms of separation, in terms of
  - reduced column diameters,
  - solvent flow-rates and
  - the overall surface area that a glycoprotein sample may encounter during analysis, are becoming significant in high-sensitivity work.

Isolation of Glycoproteins
General Considerations

- Glycoproteins can be purified by most conventional protein separation methodologies, including
  - gel electrophoresis
  - various forms of HPLC (ion-exchange, size exclusion, reversed phase using C18, C8 or C4 columns, hydrophobic interaction, and affinity).
- A most useful, specific isolation principle is the use of lectins that are immobilized on chromatographic resins

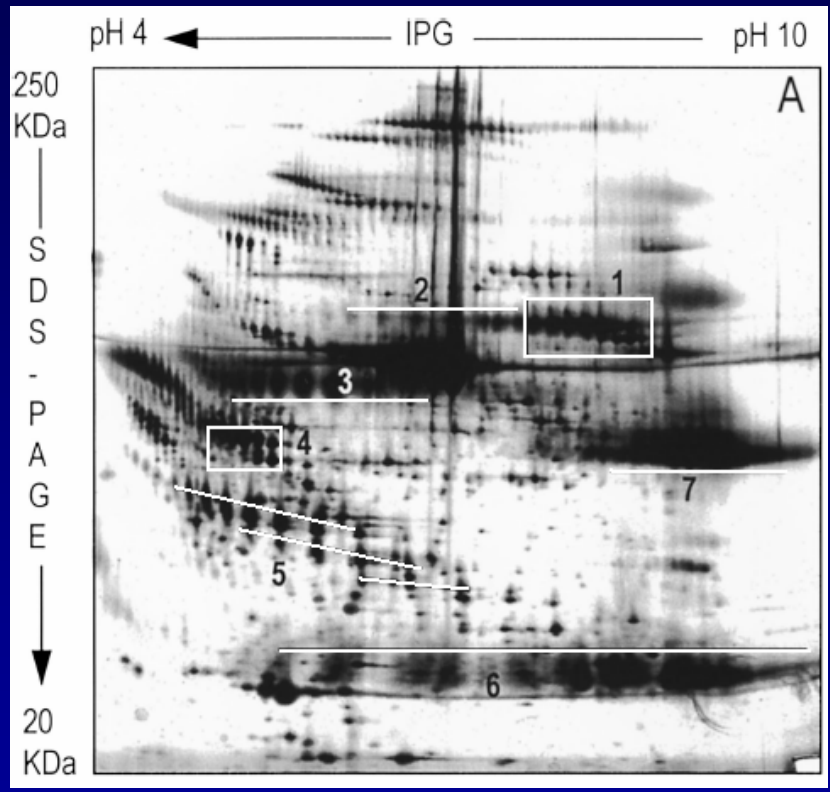


Y. Mechref and M.V. Novotny, *Chem. Rev.*, **102**, 321, (2002).

Analysis of Glycoproteins

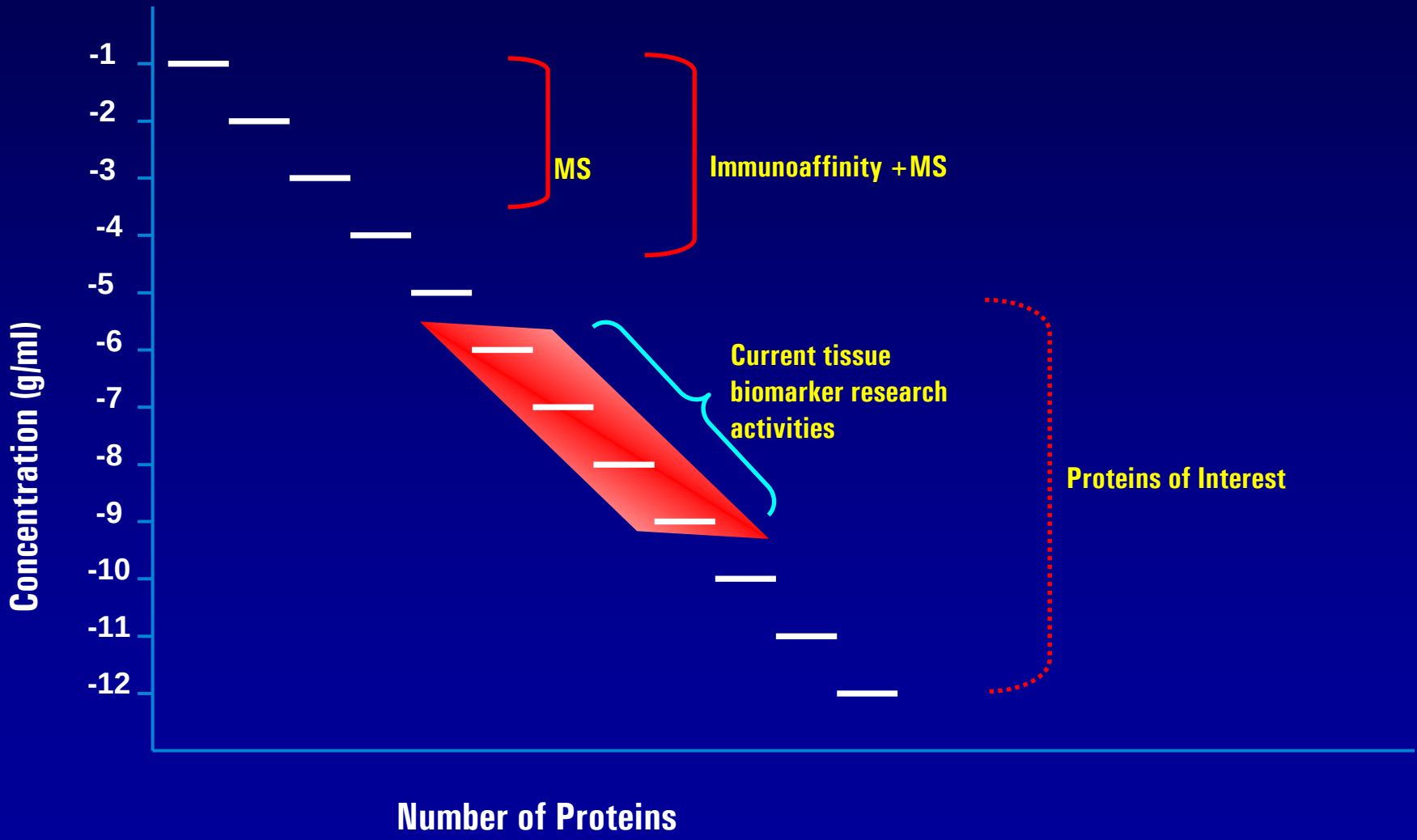
2-DE

- In 2-DE separations, glycoproteins tend to be translocated into “trains” of spots, reflecting their differences in both the molecular mass and isoelectric points.
- Silver-stained serum/plasma proteins after 2-DE.
  - (1) transferrin,
  - (2) IgM m chain,
  - (3) IgA a chain,
  - (4) α1-antitrypsin,
  - (5) haptoglobin b chain and haptoglobin cleaved b chain,
  - (6) Ig light chains,
  - (7) IgG g chain.



Gravel, P.; Walzer, C.; Aubry, C.; Balant, L. P.; Yersin, B.; Hochstrasser, D. F.; Guimon, J., *Biochem. Biophys. Res. Commun.* **1996,** 220, 78-85.

The Challenge of Protein Identification in Human Sera



Lectins have in the past been regarded by many scientists as curious proteins of uncertain structure and specificity that bind to carbohydrates of dubious significance themselves. All this is rapidly changing. The functional importance of glycosylation in cell-cell and cell-pathogen interactions, as well as intracellular events, has been recognized by the explosion of the science of glycobiology. This has been paralleled by the realization that lectins, once they have been well characterized, can be extremely useful tools for examining structural changes in glycosylation and their functional consequences for human pathophysiology.

Preface to “Lectin Methods and Protocols,” by Jonathan M. Rhodes and Jeremy D. Milton, Humana Press, 1998

Isolation of Glycoproteins

Lectin Affinity Chromatography

- Lectins are specialized proteins that have been isolated from various plants and animal sources.
- Lectins have been widely used to isolate, purify and characterize glycoproteins and glycolipids in various modes of affinity chromatography.
- These techniques are based on a reversible biospecific interaction of certain glycoproteins with the lectins immobilized to a solid support.

Isolation of Glycoproteins

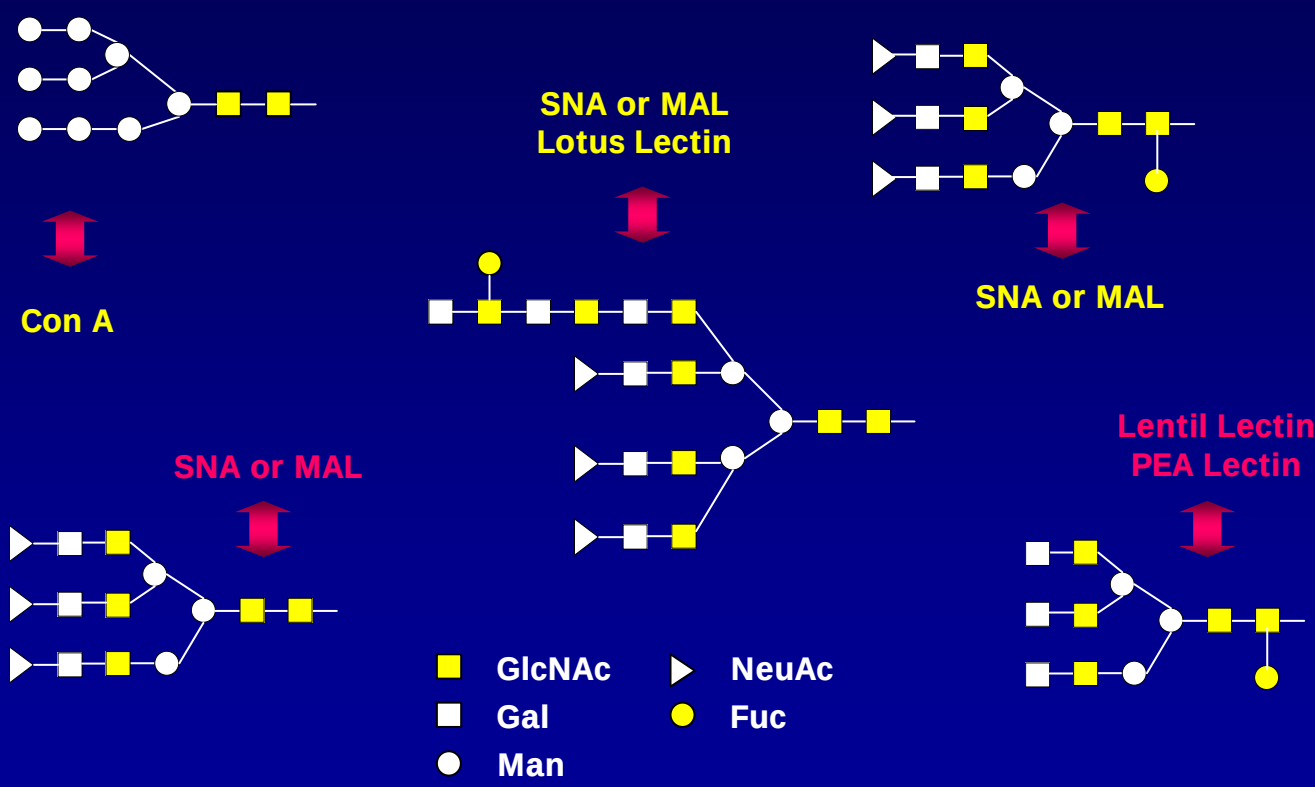
Lectin Affinity Chromatography

Table 1. Commonly used agarose-based lectins.

Name or abbreviation	Source	Amount of immobilized lectin [mg/mL]	Specificity
Con A (Concanavalin A)	*Canavalia ensiformis*	8 - 10	Glc, Man
SNA-I	*Sambucus nigra*	2 - 3	NeuAc α(2,6)
MAA	*Maackia amurensis*	2 -3	NeuAc α(2,3)
UEA-I	*Ulex europaeus*	4 - 5	Fuc α(1,2)
Jacalin	*Artocarpus integrifolia*	2 - 3	Gal
PHA-L	*Phaseolus vulgaris*	4 - 5	complex
Lotus	*Lotus tetragonolobus*	4 -5	Fuc α(1,2)
HPA	*Helix pomatia*	1 -2	GalNAc
WGA (Wheat germ)	*Triticum vulgaris*	4 -5	GlcNAc
RCA-I (Ricin)	*Ricinus communis*	4 -5	Gal, GalNAc
LcH (Lentil)	*Lens culinaris*	4 - 5	Man

EY Laboratories (St. Mateo, CA)

Lectin Specificity



Lectins

Con A – *Canavalia ensiformis*
Lotus Lectin – *Tetragonolobus purpureas*

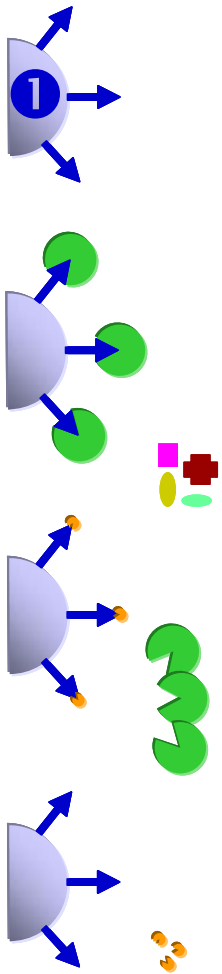
SNA – *Sambucus nigra*
PEA Lectin – *Phaseolus vulgaris*

MAL – *Maackia amurensis*
Lentil Lectin – *Lens culinaris*

R.D.Cummings, ***Methods in Enzymology***, Vol. 230 (**1994**) 66-86

Isolation of Glycoproteins

Lectin Affinity Chromatography

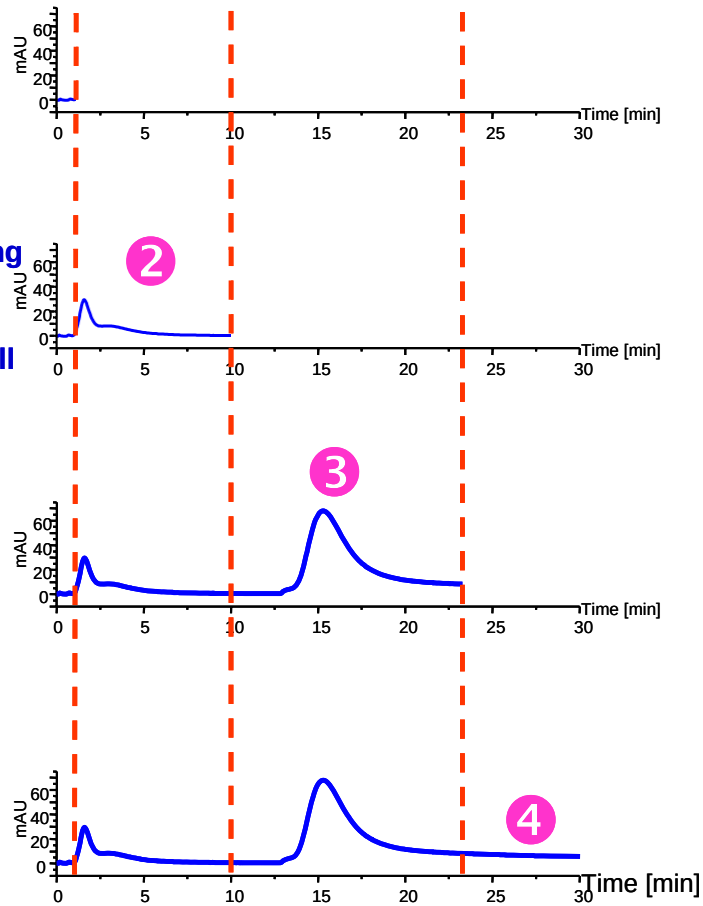


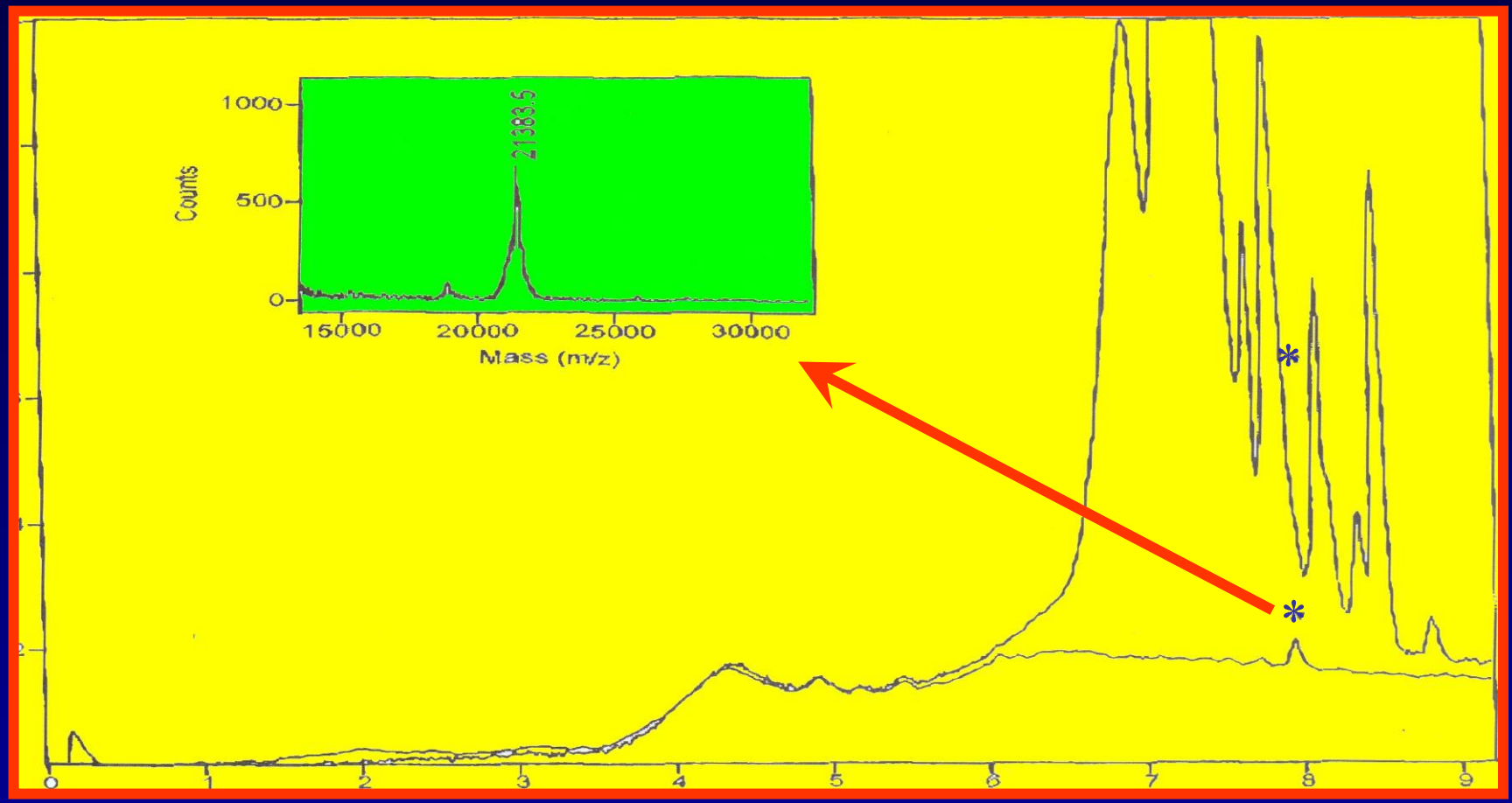
Equilibrating with binding buffer

Applying sample prepared in binding buffer, prompting specific, but reversible, binding of target substances, while other material will wash through

Elution of bound material, using a competitive ligand which displaces target substances

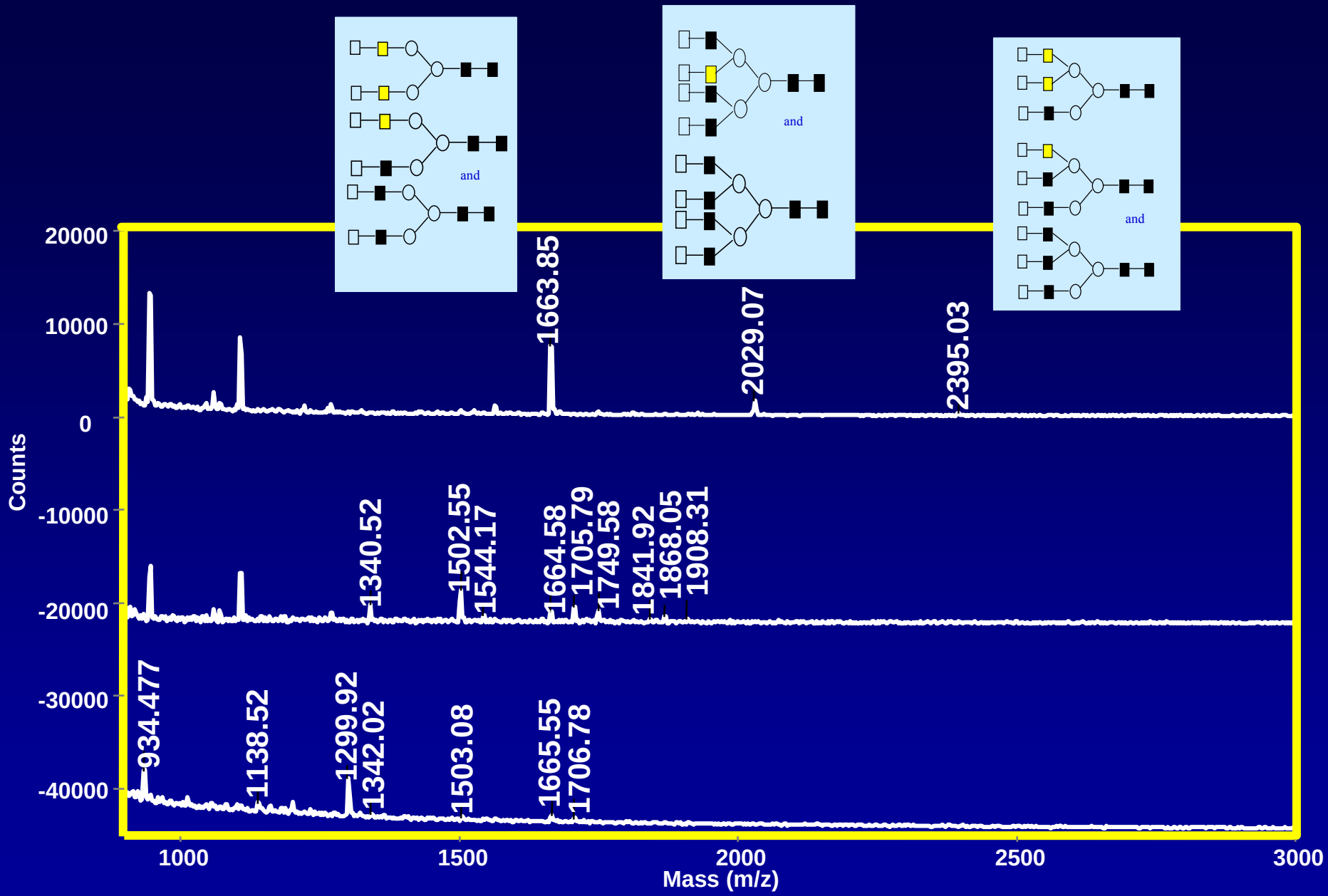
Re-equilibrating with binding buffer





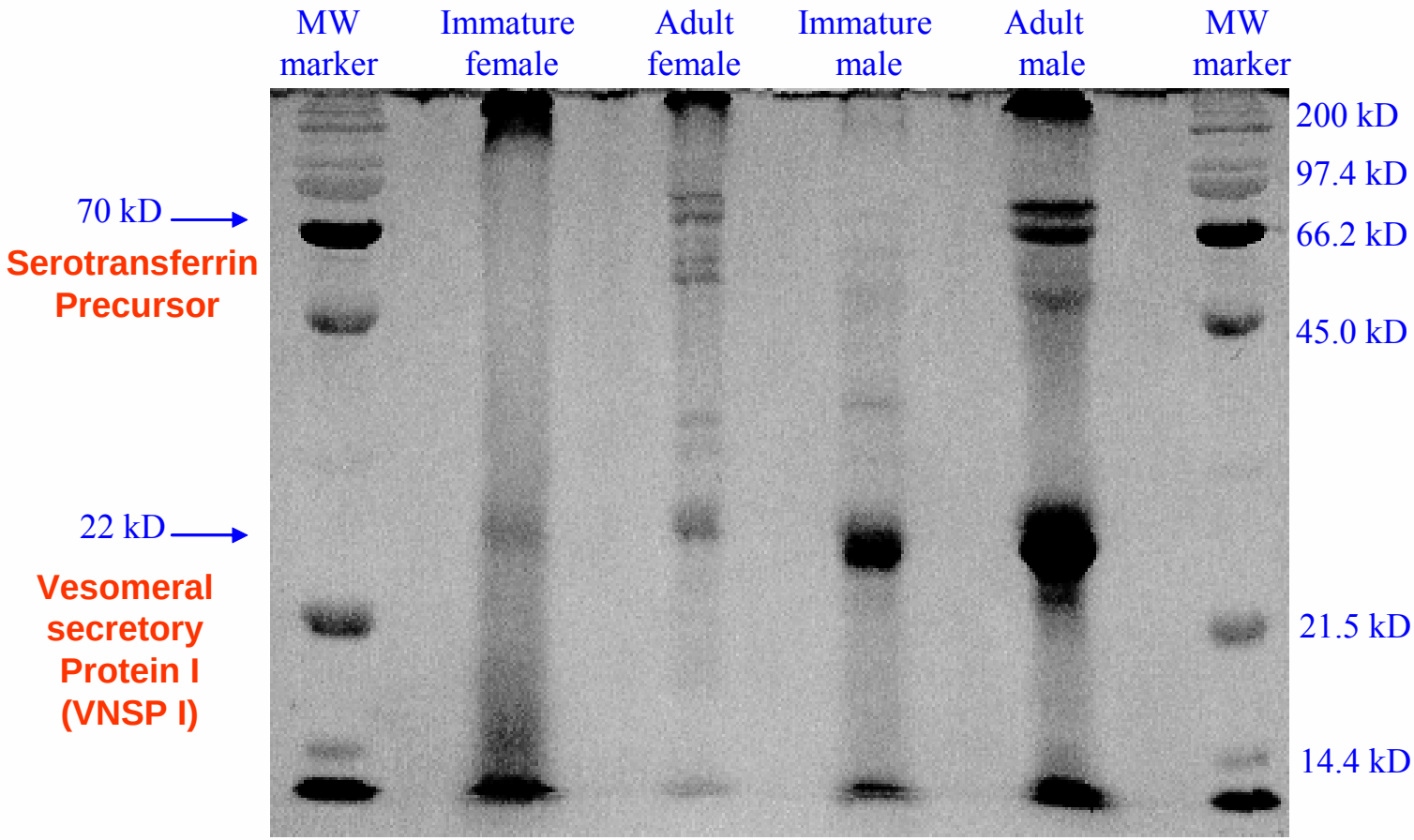
Anion-exchange chromatogram of the isolated MUP components (upper trace) and the Concanavalin A bound fraction (lower trace); inset is the mass spectrum of the isolated glycoprotein, as indicated with the asterisk.

Y. Mechref, L. Zidek, W. Ma, and M.V. Novotny, ***Glycobiology***, 10 (**2000**) 231-235.

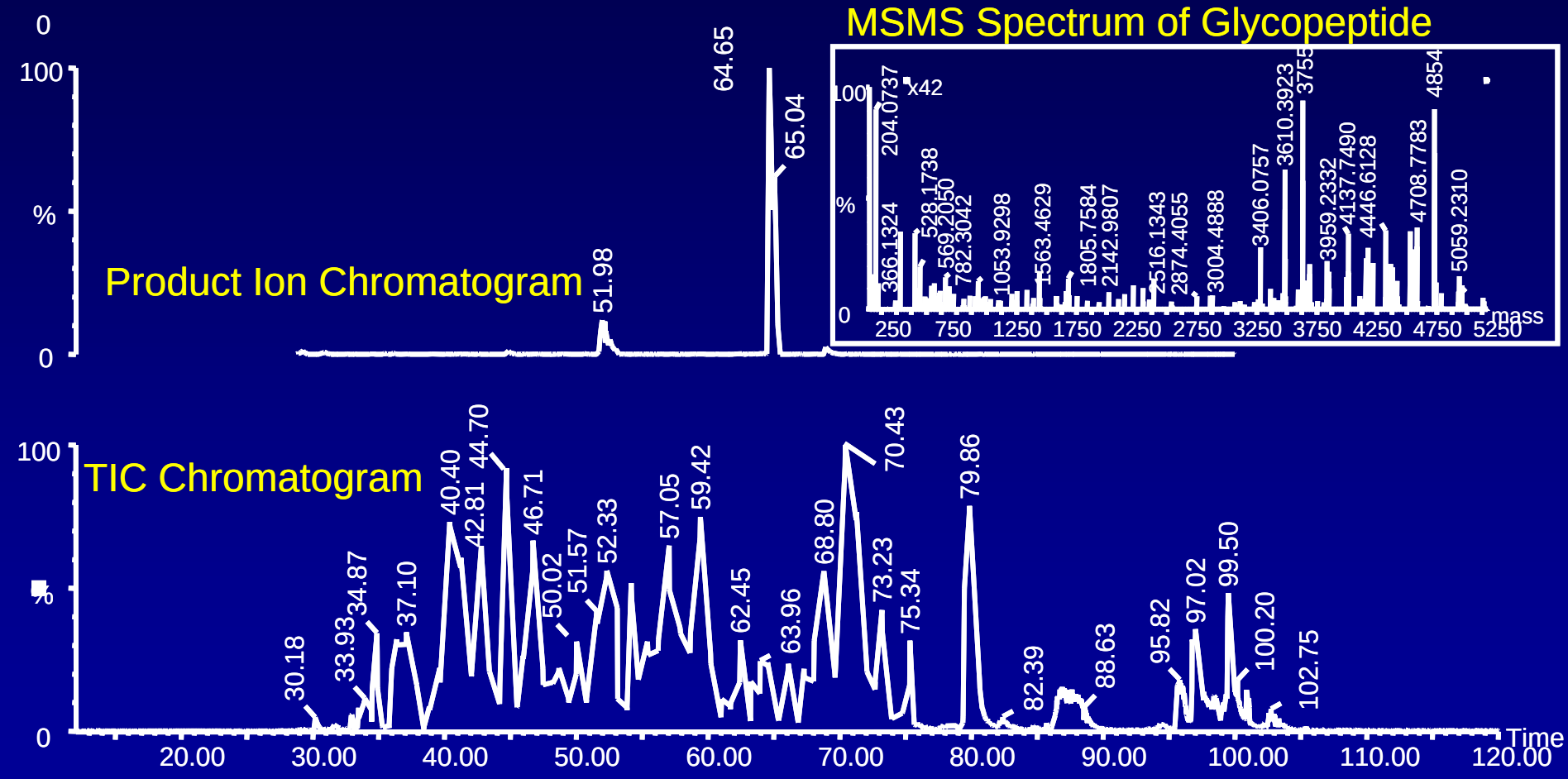


Y. Mechref, L. Zidek, W. Ma, and M.V. Novotny, ***Glycobiology***, 10 (**2000**) 231-235.

SDS-PAGE, Water Soluble Glycoproteins in the VNO of Mice

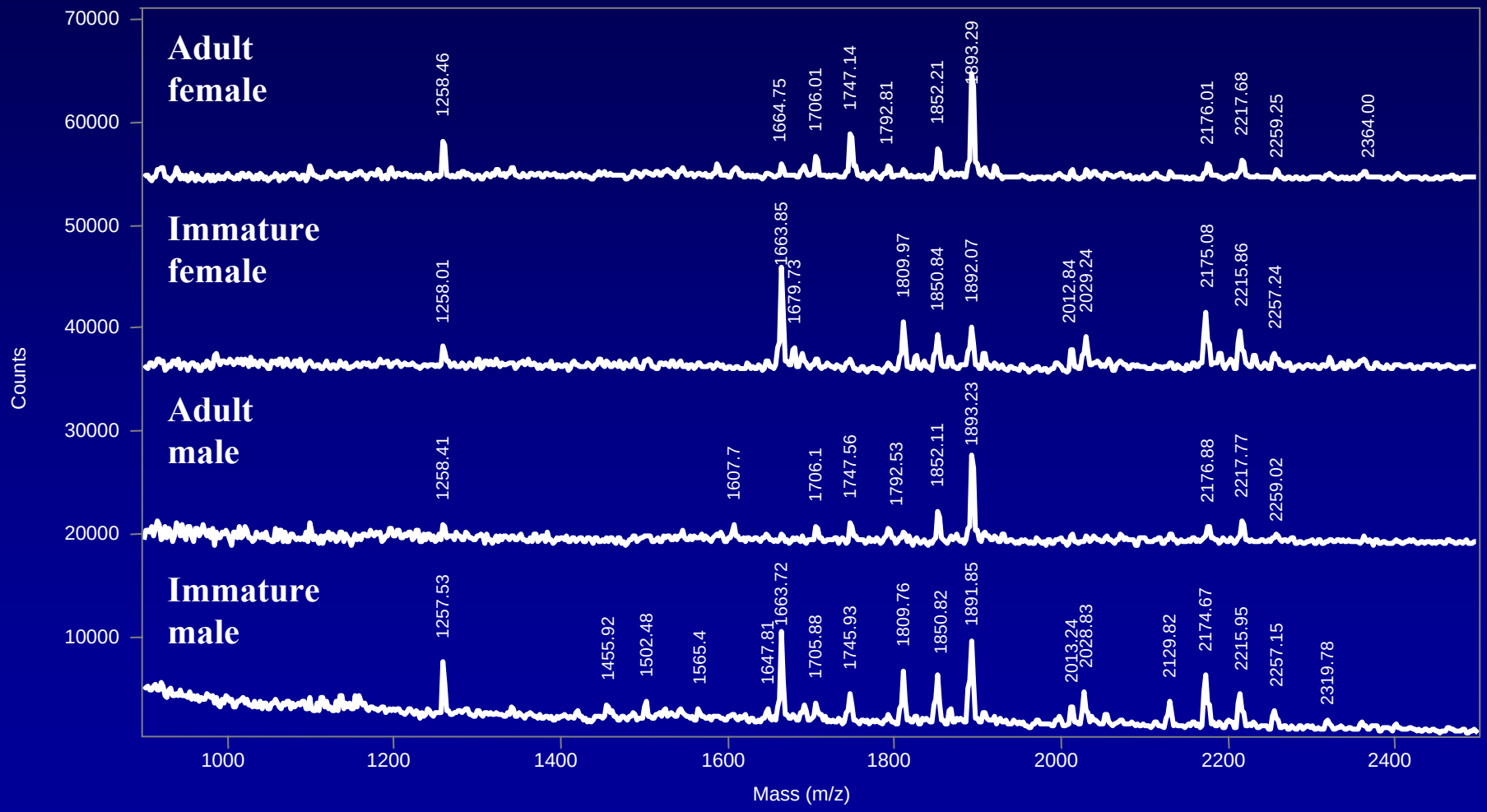


Glycosylation Site of VNSP I



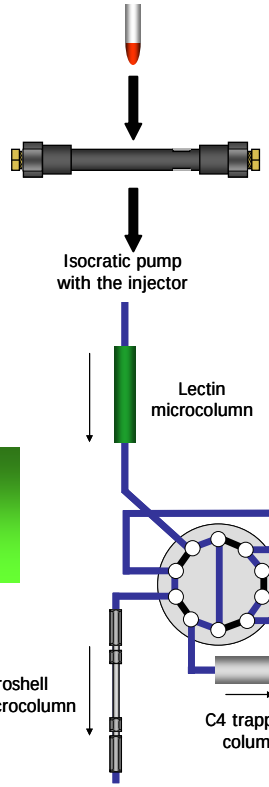
Glycosylation site [LAFNNGNFSGK]

We have recently found that certain proteins expressed in the olfactorily active tissues of mice (*Mus domesticus*) exhibit sexually dimorphic glycosylation (attachment of N-linked oligosaccharides). An example is shown for VNSP I protein isolated from the vomeronasal organ:



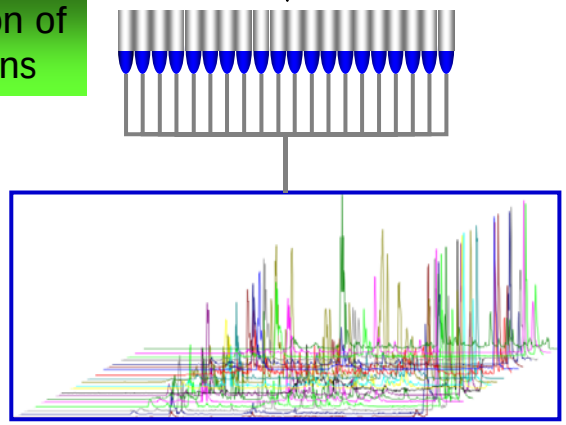
Sample

MARS Column



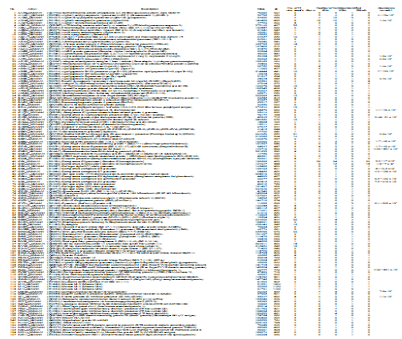
On-line microcolumn lectin affinity chromatography with high-temp RPHPLC

Collection of fractions



Nano LC-MS/MS

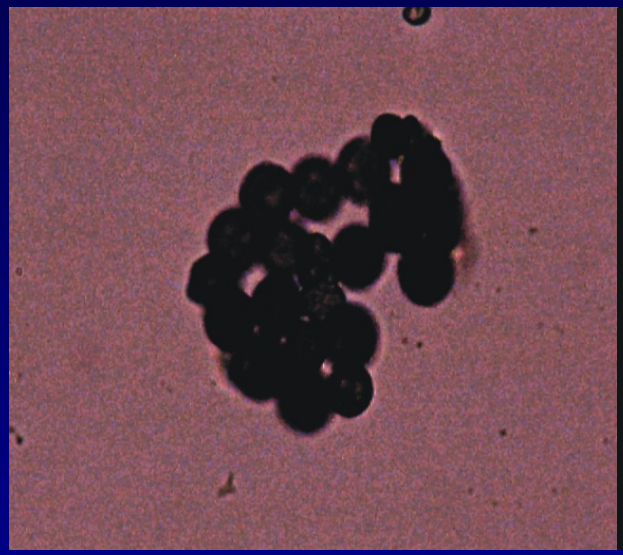
Database search



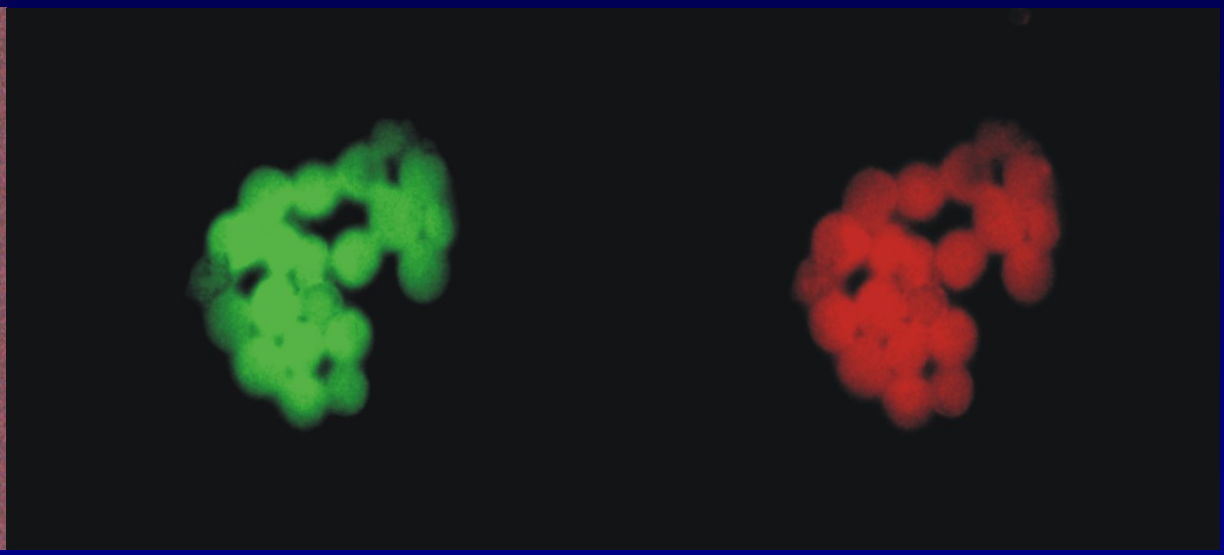
Results

M. Madera, Y. Mechref, I. Klouckova, M. V. Novotny “Semiautomated High-sensitivity Profiling Of Human Blood Serum Glycoproteins Through Lectin Preconcentration And Multidimensional Chromatography/Tandem Mass Spectrometry” ***J. Proteome Res.***, in press.

Silica-Based Con A



Silica Resins



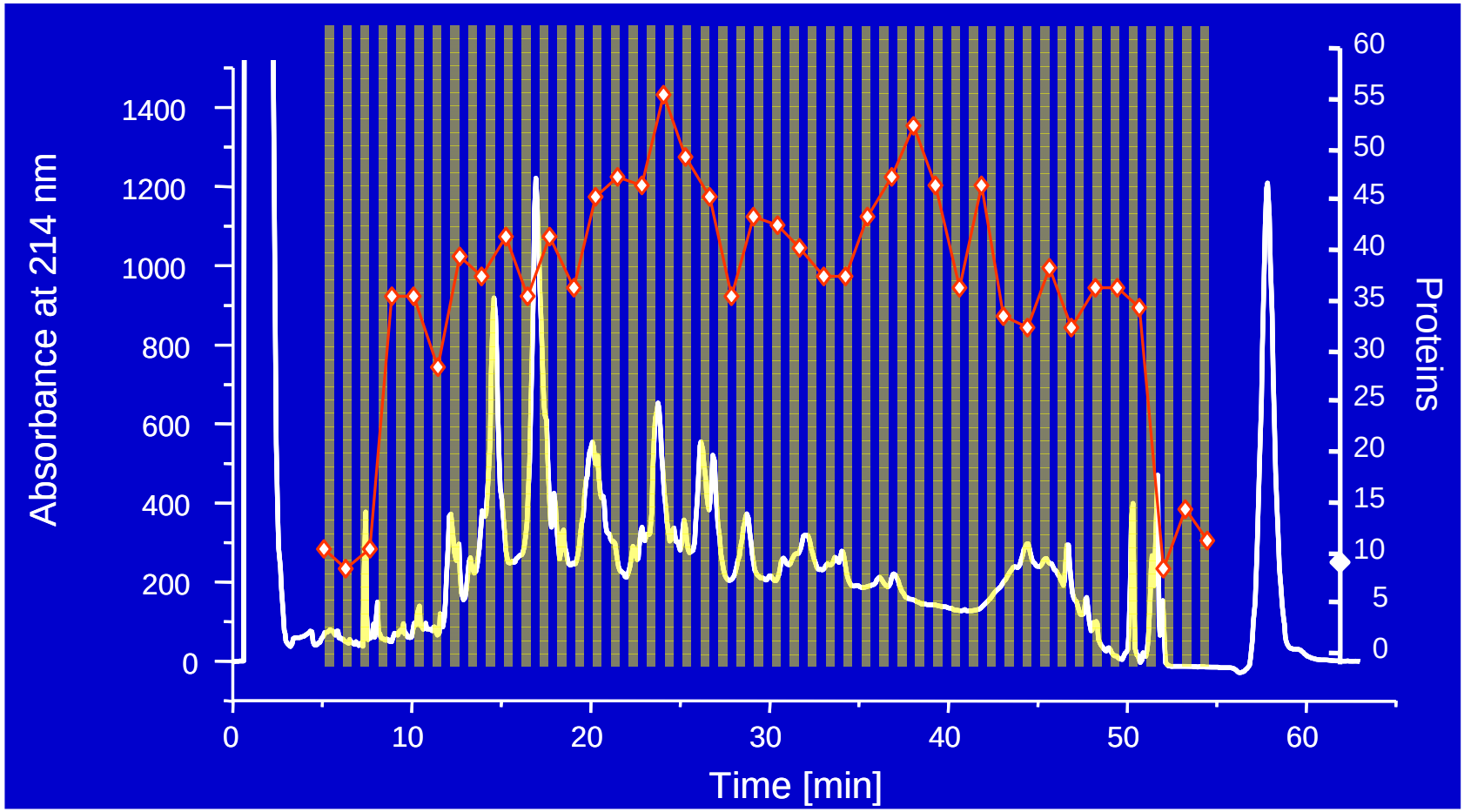
FITC-Con A Immobilized

Texas Red Ovalbumin Bound to FITC-Con A

Milan Madera, Yehia Mechref and Milos V. Novotny “Combining Lectin Microcolumns With High-Resolution Separation Techniques For Enrichment Of Glycoproteins And Glycopeptides”, ***Anal. Chem.***, 77(13) (2005) 4081-4090.

Analysis of Glycoproteins

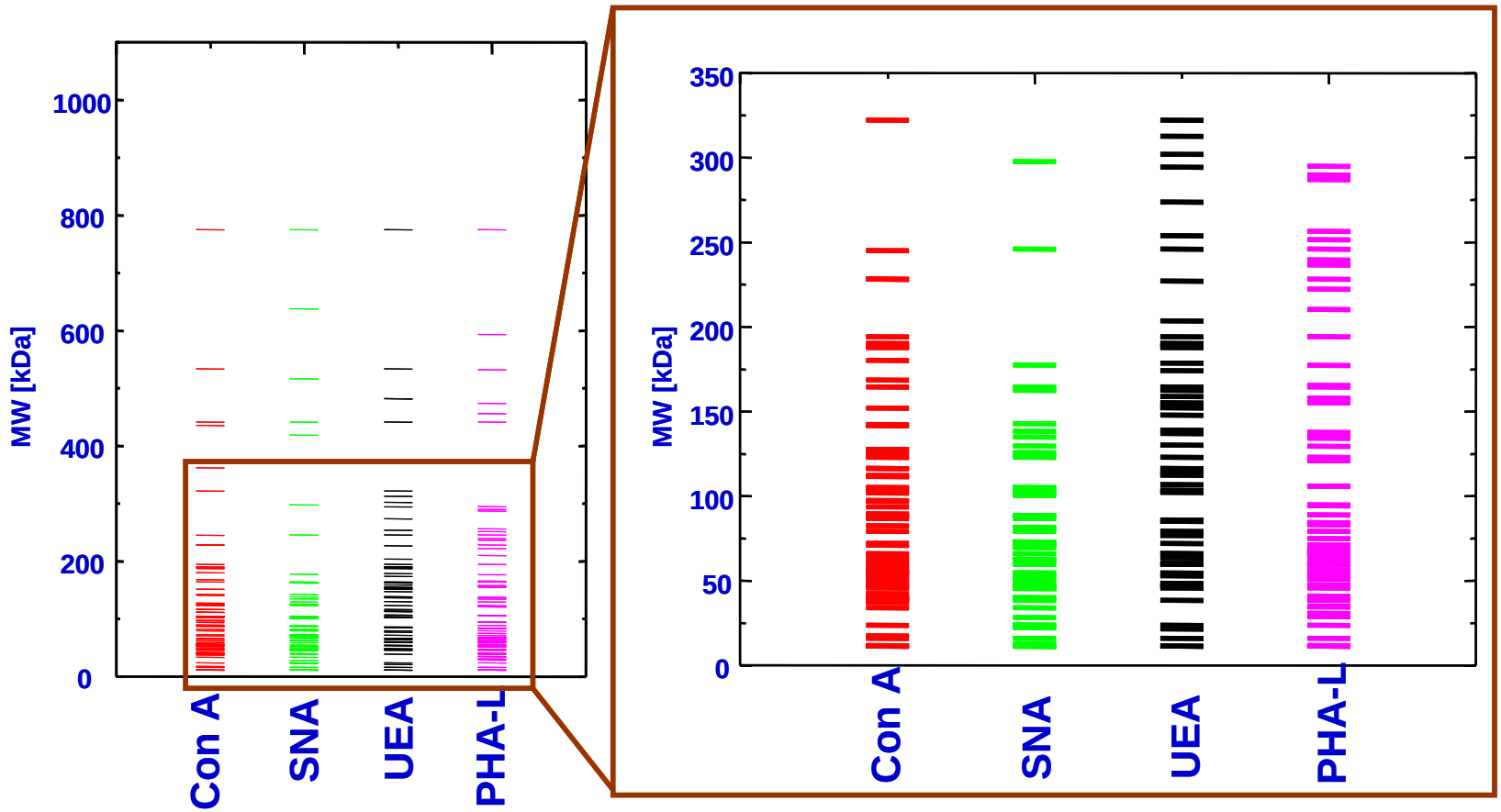
Multidimensional Approach



M. Madera. Y. Mechref, I. Klouckova, M.V. Novotny, ***J. Proteome Res.***, in press.

Analysis of Glycoproteins

Multidimensional Approach



M. Madera. Y. Mechref, I. Klouckova, M.V. Novotny, ***J. Proteome Res.***, in press.

Glycoprotein Enrichment

Approach	Sample Volume	Glycoproteins Identified	Detection Limits	Mass Spectrometer
MLAC[a]	100 μl	50	NA	Proteome X
MLAC[b]	100 μl	150	NA	LTQ-FT MS
Hydrazide Attachment to Beads[c]	800 μl	**303**	200 pg/ml	LTQ-FT MS
Silica-based Lectin with Poroshell[d]	20 μl	271	200 pg/ml	XCT plus
Silica-based Lectin (no Depletion)[e]	20 μl	108	low ng/ml	LCQ DECA Plus
Agarose-based (4) Lectins with mRP[f]	600 μl	**380**	Sub 200pg/ml	**XCT plus**
Agarose-based Con A, SNA, PHA, and UEA with mRP[g]	450 μl	**740**	Sub 200 pg/ml	**LTQ-FT MS**

[a] Z. Yang, W. S. Hancock, ***J. Chromatogr. A*** 1053 (2004) 79-88

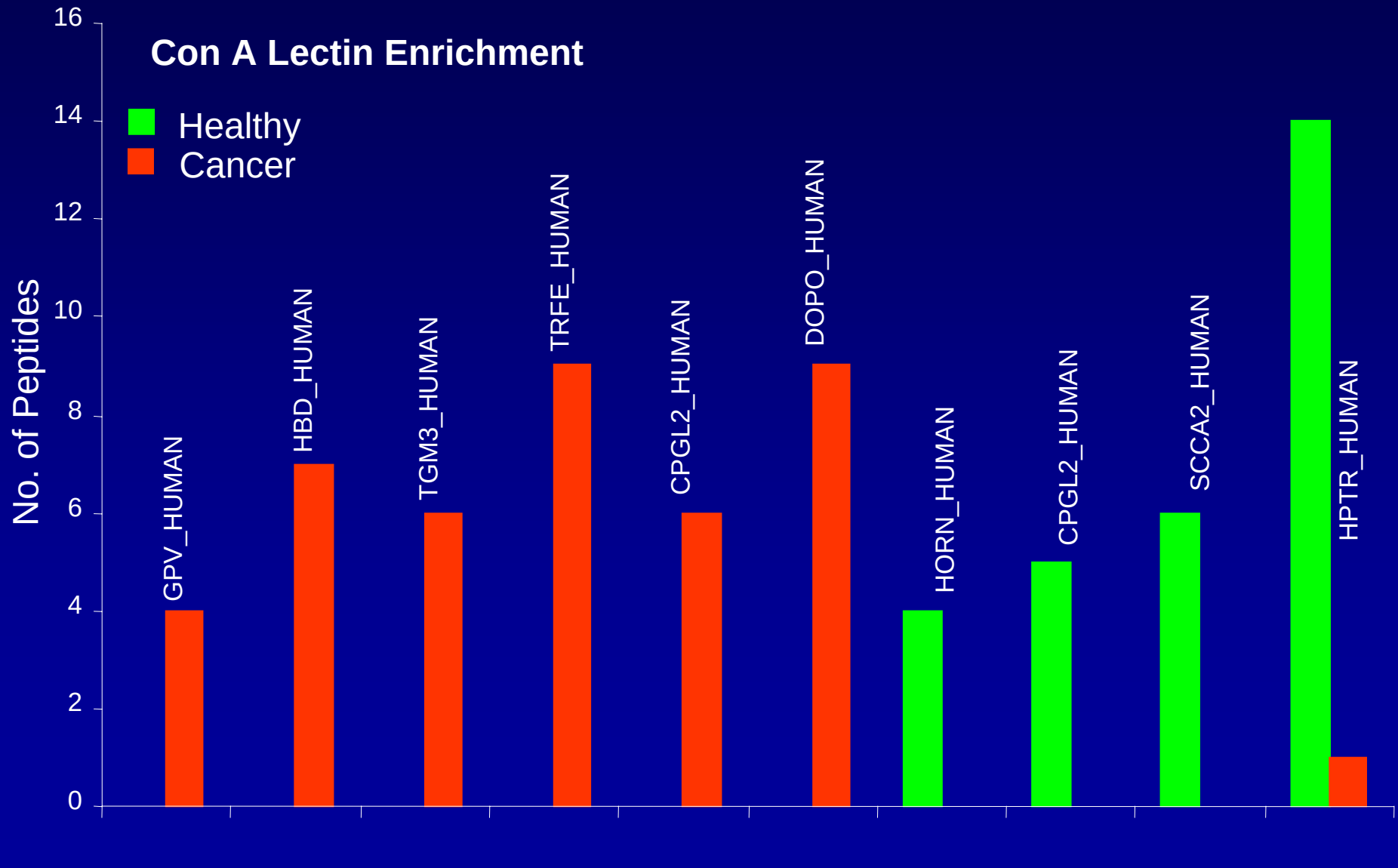
[b] Z. Yang, W. S. Hancock, T. R. Chew, L. Bonilla, ***Proteomics*** 5 (2005) 3353-3366

[c] T. Liu, W-J. Qian, M. A. Gritsenko, D. G. Camp II, M. E. Monroe, R. J. Moore, R. D. Smith, ***J. Proteome Res.*** 4 (2005) 2070-2080.

[d] M. Madera, Y. Mechref, I. Klouckova, M. V. Novotny "Semiautomated High-sensitivity Profiling Of Human Blood Serum Glycoproteins Through Lectin Preconcentration And Multidimensional Chromatography/Tandem Mass Spectrometry" *J. Proteome Res.*, in press.

[e-g] work submitted for publication

Glycoproteome Changes in the Blood Serum of Healthy and Breast Cancer Patient



Analysis of Glycopeptides

General Considerations

- It is most common that glycosylation of a particular protein is investigated through a chemical or enzymatic release of glycans and their subsequent characterization such as sequencing and linkage analysis.
- While this information may be highly significant, there are additional structural aspects that must be addressed.
- For each site of glycosylation, there are possible structural variations (extent of substitution, site-specific microheterogeneities, a site-specific accessibility for particular glycosyltransferases, etc.), which can all have important biochemical consequences.
- Investigating protein glycosylation at the level of glycopeptide is at least as important as the investigation of released glycan structures.

Analysis of Glycopeptides

General Considerations

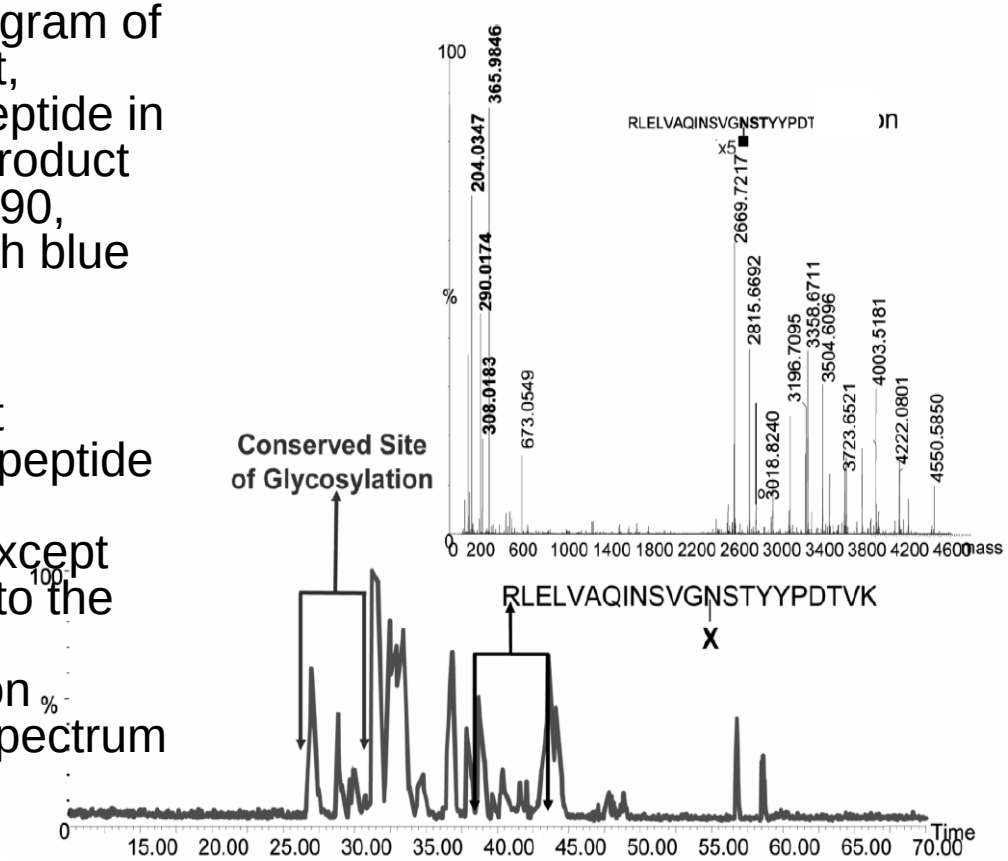
- It is difficult to directly identify glycopeptides in a complex protein digest by MS.
- This is partly due to the low sensitivity of the detection of glycopeptides caused by
  - site heterogeneity and/or ion adduct formation.
  - Glycopeptide signals are often suppressed in the presence of other peptides, especially if the glycans are terminated with the negatively charged sialic acid moiety.
- Due to the glycan heterogeneity and a frequent multiple adduct formation, the overall glycopeptide signal distributes into several peaks resulting in weak signals detected by MS.

Analysis of Glycopeptides

ESI/MS

- Base peak intensity chromatogram of tryptically digested mAb. Inset, MS/MS spectrum of a glycopeptide in which the glycan-diagnostic product ions with *m/z* values of 204, 290, 308, and 366 are depicted with blue lines.

- The signal labeled in the inset represent the ion of this glycopeptide resulting from the loss of all monosaccharides residues, except the GlcNAc residue attached to the peptide backbone. This ion is commonly the most intense ion observed in the tandem MS spectrum of a glycopeptide.

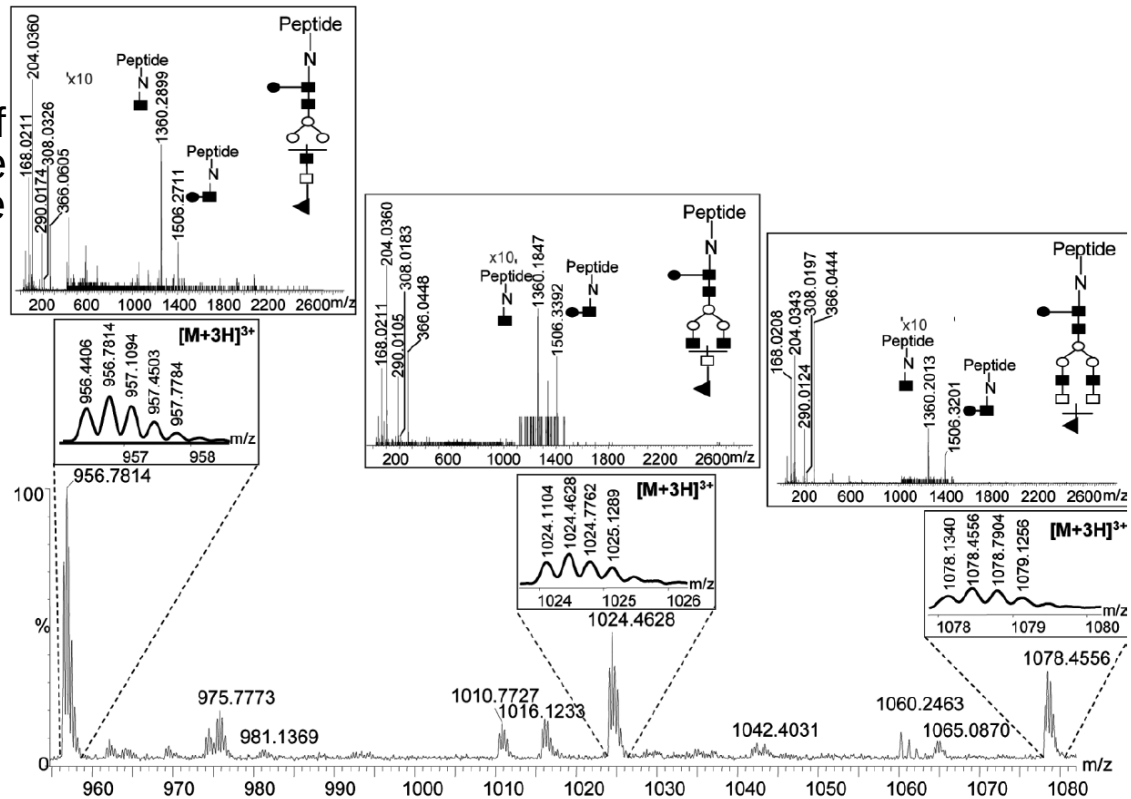


Y. Mechref, J. Muzikar, M.V. Novotny, ***Electrophoresis***, 26 (**2005**) 2034–2046.

Analysis of Glycopeptides

ESI/MS

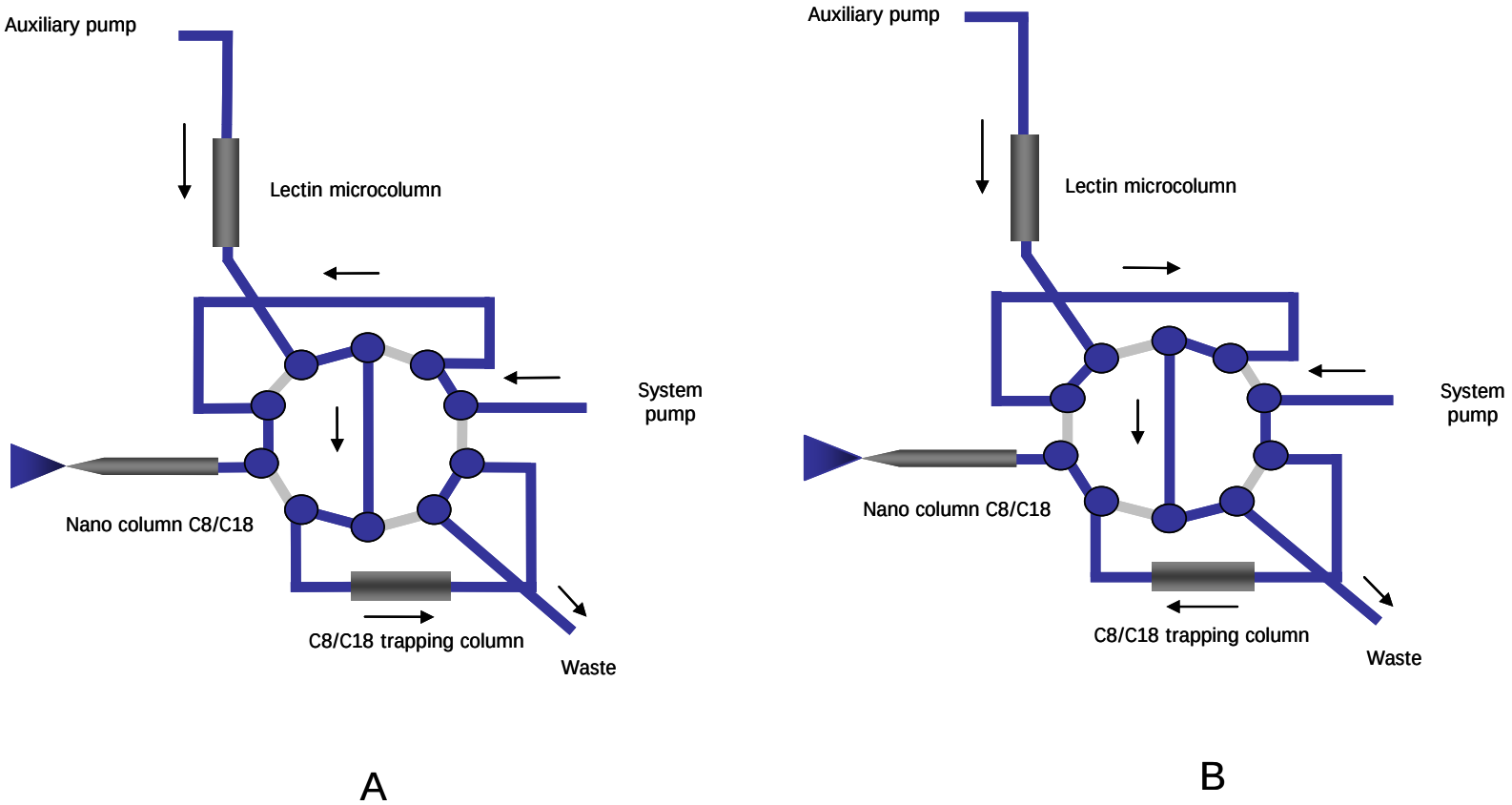
MS spectrum of sialylated glycopeptides associated with the Fc region of the antibody. The lower inset represent the charge state of different glycopeptides, while the upper inset represent the tandem MS spectra of each of the glycopeptides. Structures of the glycan attached to the peptide backbones are represented by the following symbols: (■) GlcNAc; (●) fucose; (◯) mannose; (□) galactose; (▲) *N*-acetylglycolylneuraminic acid.



Y. Mechref, J. Muzikar, M.V. Novotny, ***Electrophoresis***, 26 (**2005**) 2034–2046

Analysis of Glycopeptides

ESI/MS (Lectin Enrichment)

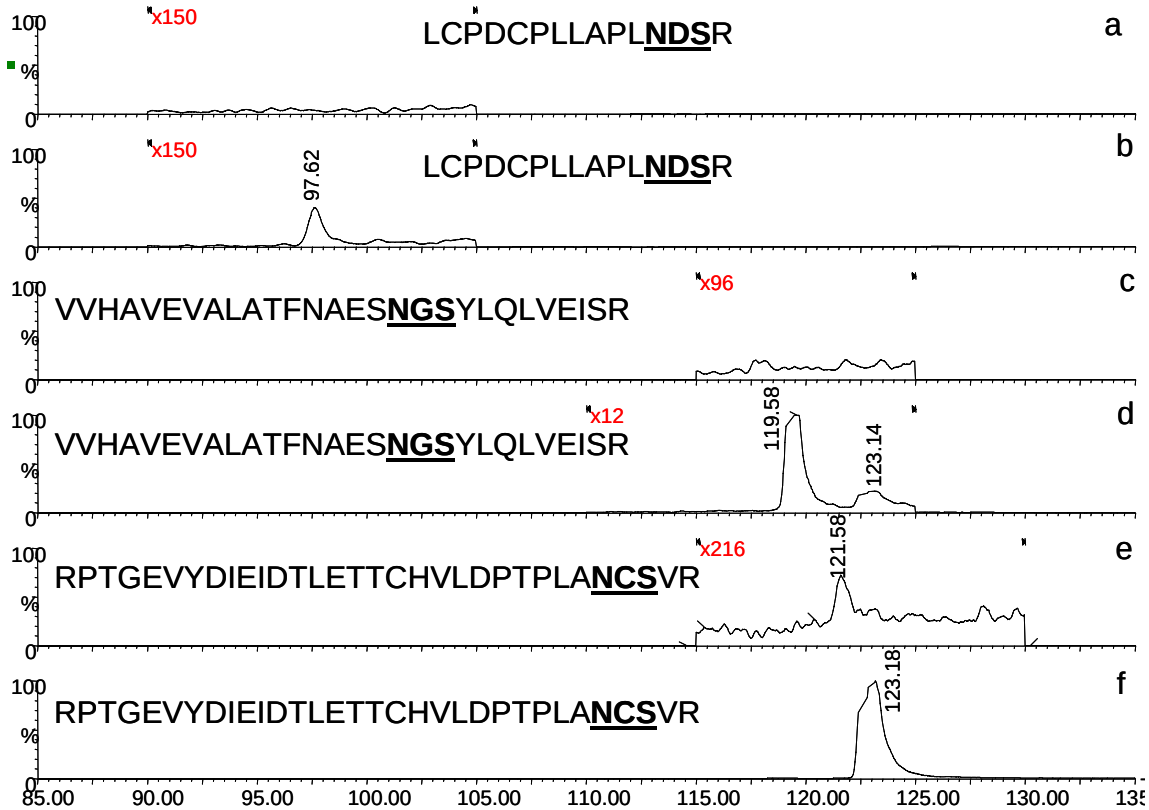


High-performance affinity chromatography setup coupled on-line to electrospray mass spectrometry, including a system for loading of proteins/glycoproteins (A); and elution and on-line analysis of proteins/glycoproteins (B).

M. Madera, Y. Mechref, M.V. Novotny, ***Anal. Chem.*** 77(13) (**2005**) 4081-4090.

Analysis of Glycopeptides

ESI/MS (Lectin Enrichment)



Extracted ion chromatograms of fetuin glycopeptides analyzed by on-line SNA-lectin trapping and LC/MS. Lectin unbound fractions (a, c, and e), and lectin bound fractions (b, d, and f).

M. Madera, Y. Mechref, M.V. Novotny, ***Anal. Chem.*** 77(13) (**2005**) 4081-4090.

Glycan Release

Enzymatic

Table 2. Glycoprotein oligosaccharide-releasing enzymes.

Highlighted enzymes are most commonly used

Enzyme	Source	Susceptible N-glycans[a]
Endo D	*Diplococcus pneumoniae*	Some high mannose Some (few) hybrids
Endo H	*Streptomyces plicatus*	High mannose Most hybrids
Endo C_I	*Clostridium perfringens*	Similar to endo D
Endo C_{II}	*Clostridium perfringens*	Similar to endo H (narrower)
Endo F_1	*Flavobacterium meningosepticum*	Similar to endo H (narrower)
Endo F_2	*Flavobacterium meningosepticum*	High mannose Biantennary complex
Endo F_3	*Flavobacterium meningosepticum*	Biantennary complex Triantennary complex
Endoglycosidase	*Arthrobacter protophormiae*	Similar to endo C_{II}
Endoglycosidase	*Bacillus circulans*	High mannose[b]
Endoglycosidase	*Bacillus alvei*	Similar to endo C_{II}
Endo S	*Dictyostelium discoideum*	Similar to endo C_{II}
Endo M	*Mucor Hiemalis*	Similar to endo F_2
Endo B	*Sporotrichum dimorphosporum*	Similar to endo F_2
Endoglycosidase	*Aspergillus oryzae*	Similar to endo H
Endoglycosidase	Fig (enzyme 1)	Similar to endo C_{II}
Endoglycosidase	Fig (enzyme 2)	Similar to endo D
Endoglycosidase	Jack bean	Similar to endo C_{II}
Endoglycosidase	Human kidney (enzyme 1)	Mostly high mannose
Endoglycosidase	Human kidney (enzyme 2)	High mannose and complex (free oligosaccharides only)
Endoglycosidase	Rat liver	Similar to endo F_2
Glycoamidase A	Almond	Similar to Glycoamidase F
Glycoamidase F	*Flavobacterium meningosepticum*	All N-linked (except 1,3-core Fuc)
Glycoamidase	Jack bean	Similar to Glycoamidase F)
EndoGalNacase D	*Diplococcus*	Gal-β-1,3-GalNAc only
EndoGalNacase A	*Alcaligenes*	Gal-β-1,3-GalNAc only
EndoGalNacase S	*Streptomyces*	Gal-β-1,3-GalNAc (plus larger structures)

[a] N-Glycan substrate specificities given are brief. Further details of most specificities are given in the text
[b] Substrate specificity not fully characterized.

R.A. O'Neill ***J. Chromatogr. A*** 720 (**1996**) 201-215

Glycan Release

Enzymatic

- Unlike N-glycans, no endoglycosidases are reliably available for the release of O-linked oligosaccharides, with the partial exception of endo-α-N-acetylgalactosaminidase, permitting the release of unsubstituted Core-1 O-glycans.
- However, this highly specific enzyme has very limited use, as it does not cover the other core structures. At this time, chemical release methods provide the only universal means for O-linked glycans.

Glycan Release

Chemical

- hydrazinolysis
  - Takasaki and Kobata, Methods Enzymol., 50 (**1978**) 50.
- improved hydrazinolysis
  - Patel *et al.*, ***Biochemistry***, 32 (**1993**) 679.
- Carlson β-elimination
  - Carlson and Blackwell, ***J. Biol. Chem.***, 243 (**1968**) 616.
- improved β-elimination
  - Y. Huang, T. Konse, Y. Mechref, and M. V. Novotny, ***Rapid Commun. Mass Spectrom.***, 16 (**2002)** 1199-1204.
- ammonia-based β-elimination
  - Y. Huang, Y. Mechref and M.V. Novotny, Anal. Chem., 73 (**2001**) 6063.

Glycan Release

Chemical (Hydrazinolysis)

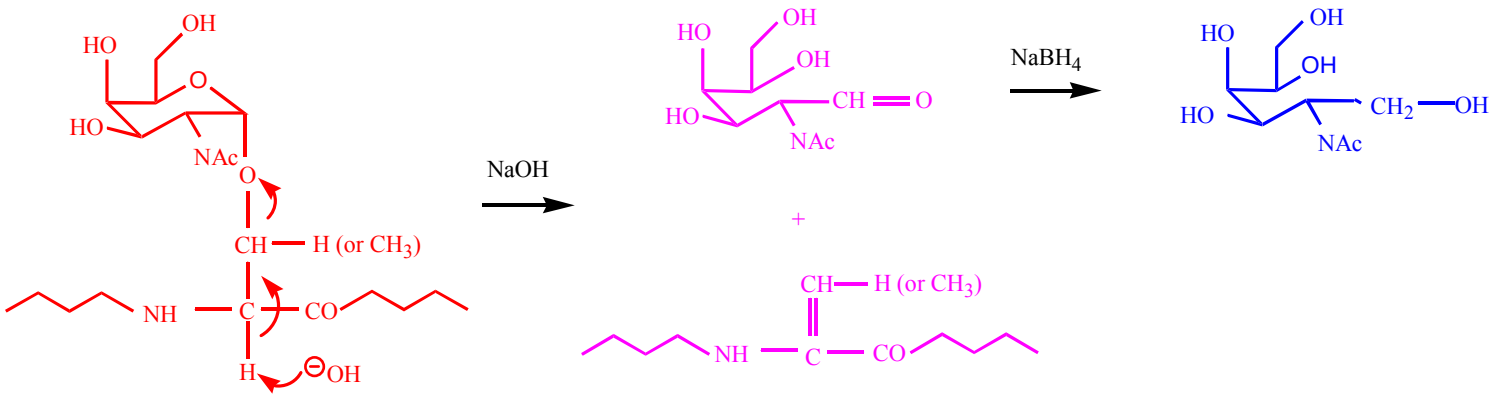
- N- and O-glycans can be chemically cleaved from glycoproteins with hydrazine (hydrazinolysis).
- O-glycans are claimed to be specifically released at 60^{o}C, while 95^{o}C is needed to release N-linked oligosaccharides.
- This chemical release approach suffers from several major disadvantages:
  - the reagent cleaves amidic bonds, including the linkage between the N-glycans and asparagine, the samples are destroyed. Consequently, any information is lost pertaining to the site of glycosylation and the extent to which it occurs.
  - the acyl groups of N-acetylamino sugars and sialic acids are hydrolyzed, calling for a reacetylation step, assuming that a sialic acid possessed originally an acetyl group rather than any other substitution.
  - the residual hydrazide or amino groups are often incorporated to the reducing terminus of some glycans.
  - a loss of the reducing terminal GlcNAc is commonly observed as a result of conducting the reaction at high temperatures.
  - It is essential to maintain strictly anhydrous conditions, which may not be always feasible.

Glycan Release

Chemical (β-Elimination)

Reductive β-Elimination of O-Glycans

Standard reagent: 1 M $NaBH_4$, 0.05 M NaOH

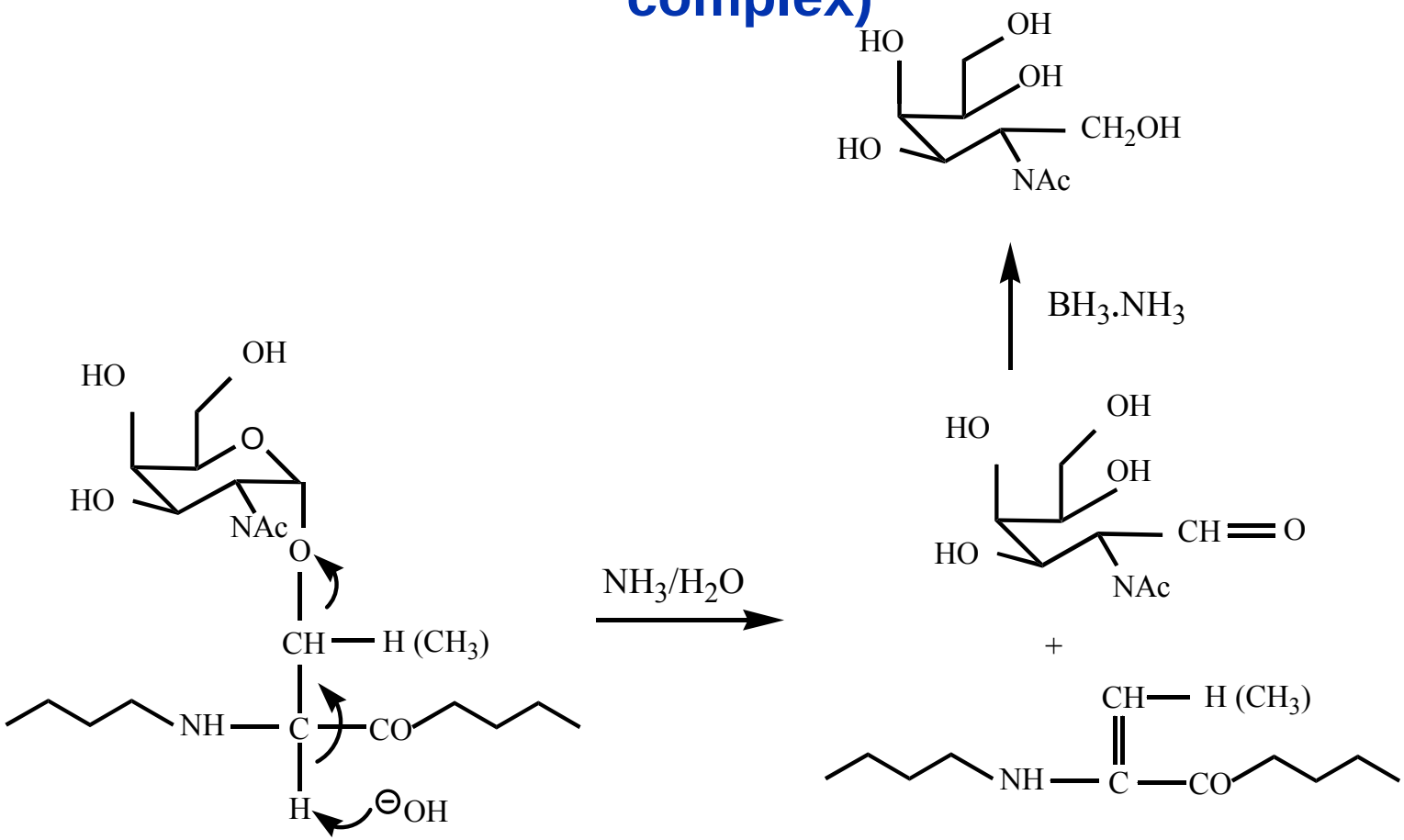


- **Drawbacks:**
  - **Oligosaccharides are reduced to alditols: lack of a reducing end**
    - **not feasible for chromatography**
  - **Desalting is necessary: high concentration of salts**

Glycan Release

Chemical (β-Elimination with $BH_3.NH_3$)

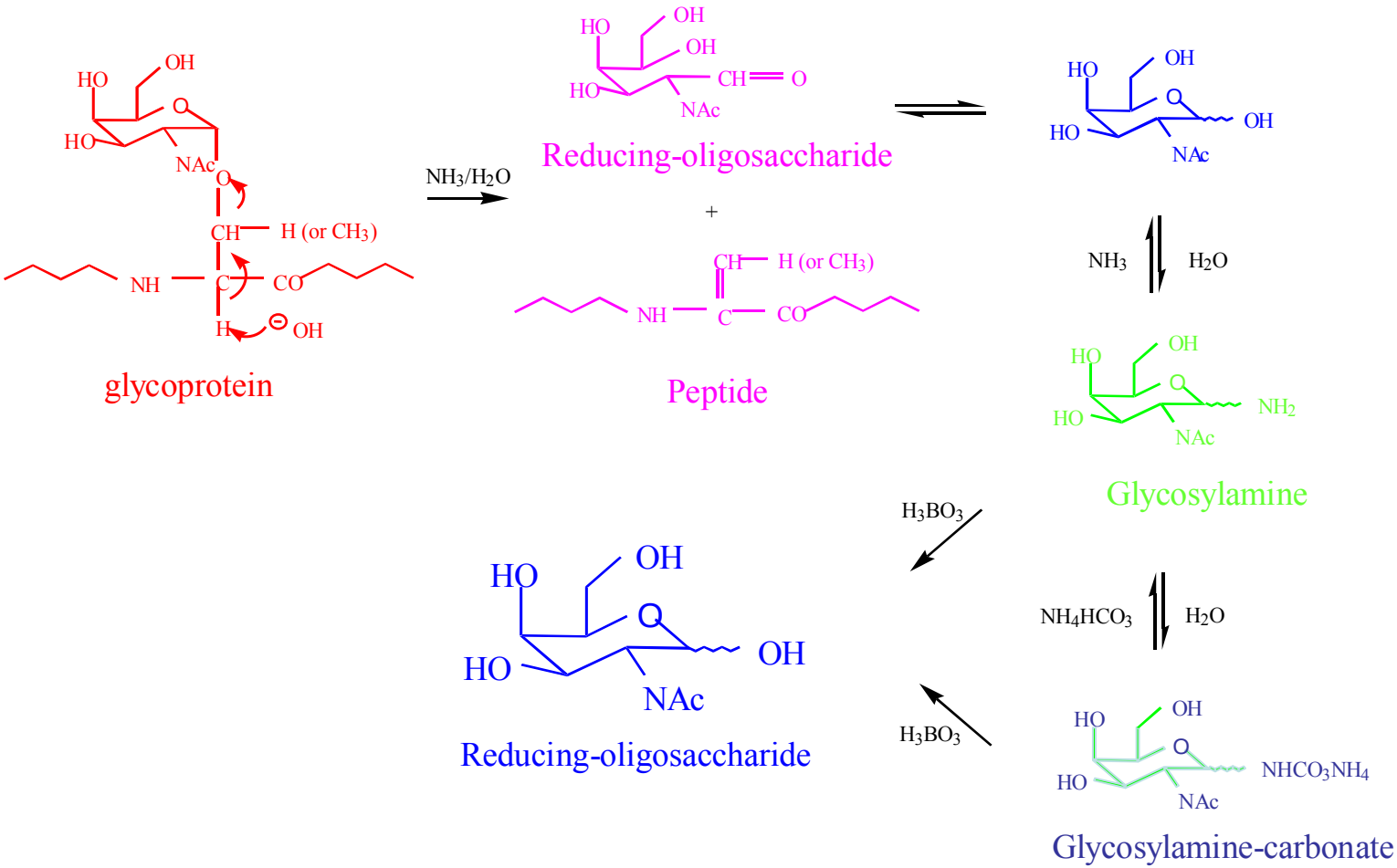
Reductive β-Elimination with $BH_3.NH_3$ (ammonia borane complex)



Y. Huang, T. Konse, Y. Mechref, and M. V. Novotny, ***Rapid Commun. Mass Spectrom.***, 16 (**2002**) 1199-1204.

Glycan Release

Chemical (ammonia-based β-Elimination)



Y. Huang, Y. Mechref and M.V. Novotny, Anal. Chem., 73 (**2001**) 6063.

Glycomic Analysis

Sample Preparation

- Since the ion yield and crystal formation in MALDI/MS analysis are adversely influenced by the presence of salts and buffers, their prior removal becomes desirable.
- Carbohydrates are generally less tolerant than proteins to salts and other compounds. This is despite the fact that small amounts of sodium or other alkali metals are required for efficient ionization.
- Many methods have been developed for the removal of salts and buffers.
  - drop dialysis
  - Nafion-117 membranes
  - synthetic membranes (polyethylene and polypropylene)
  - ion-exchange or hydrophobic resins packed pipette tips
  - Hydrophobic resins

Glycomic Analysis

MS-based Approaches (MALDI-MS)

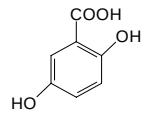
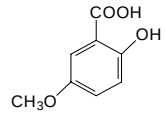
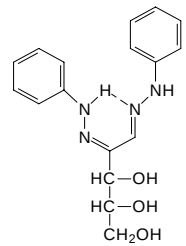
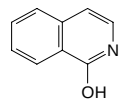
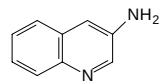
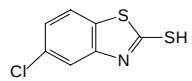
- DHB (most commonly used matrix for MALDI-MS analysis of glycans) typically crystallizes as long needle-shaped crystals that originate at the periphery of the spot and project toward the center when a mixture of acetonitrile or methanol and water is used.
- An amorphous mixture of the analyte, contaminants and salts are present in the central region of the spot.
- It is known that in a mixture of glycans and glycoproteins, glycans were fractioned in the central region of the spot while the glycoproteins were in the periphery, as concluded from the acquired spectra.
- Therefore, a more even film of crystals is produced by re-dissolving the spot in dry ethanol and allowing it to recrystalize.
- In addition to producing a thin and even film of crystals, this technique also increased sensitivity by an order of magnitude as a result of more efficient mixing of matrix and analyte from a single solvent.
- Even film of crystals is also attained by drying under vacuum.

Glycomic Analysis

MS-based Approaches (MALDI-MS)

Matrices

Table 2. Matrices and additives utilized for the analysis of N-linked glycans.

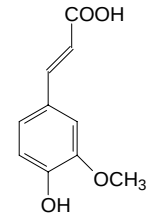
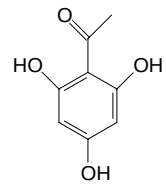
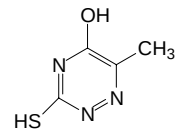
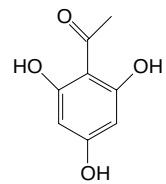
matrix	abbreviation	structure
2,5-dihydroxybenzoic acid	2,5-DHB	
2-hydroxy-5-methoxybenzoic acid	___	
arabinosazone	ara	
1-hydroxyisoquinoline	HIQ	
3-aminoquinoline	3-AQ	
5-chloro-2-mercaptobenzothiazole	CBMT	

Glycomic Analysis

MS-based Approaches (MALDI-MS)

Matrices

Table 2. Continued.

matrix	abbreviations	structure
ferulic acid	—	
2,5-dihydroxyacetophenone	DHA	
6-aza-2-thiothymine	ATT	
2,4,6-trihydroxyacetophenone	THAP	

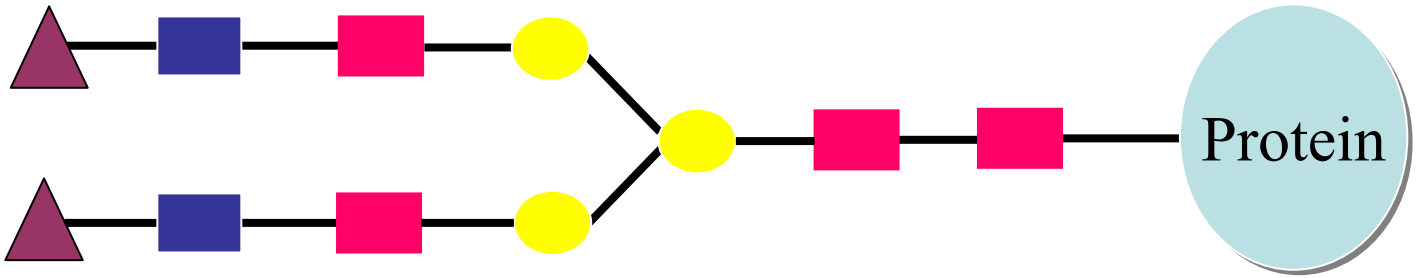
Glycomic Analysis

MS-based Approaches (MALDI-MS)

- While mass determination through MALDI/MS can often lead to compositional data (in terms of isobaric monosaccharides) additional information must be secured through other methodologies.
- Monosaccharide sequences, branching and, in some cases, linkages can be determined through fragmentation that a glycan may experience in either a post-source decay (PSD) or a collision-induced dissociation.
- The combination of MALDI/MS and enzymatic sequencing using exoglycosidases provides the necessary information related to sequence, branching and linkage of a glycan.

Glycomic Analysis

Enzymatic Sequencing (MALDI-MS)



PNGase F

Neuraminidase

β-Galactosidase

N-Acetyl-β-D-glucosaminidase

Sialic acid

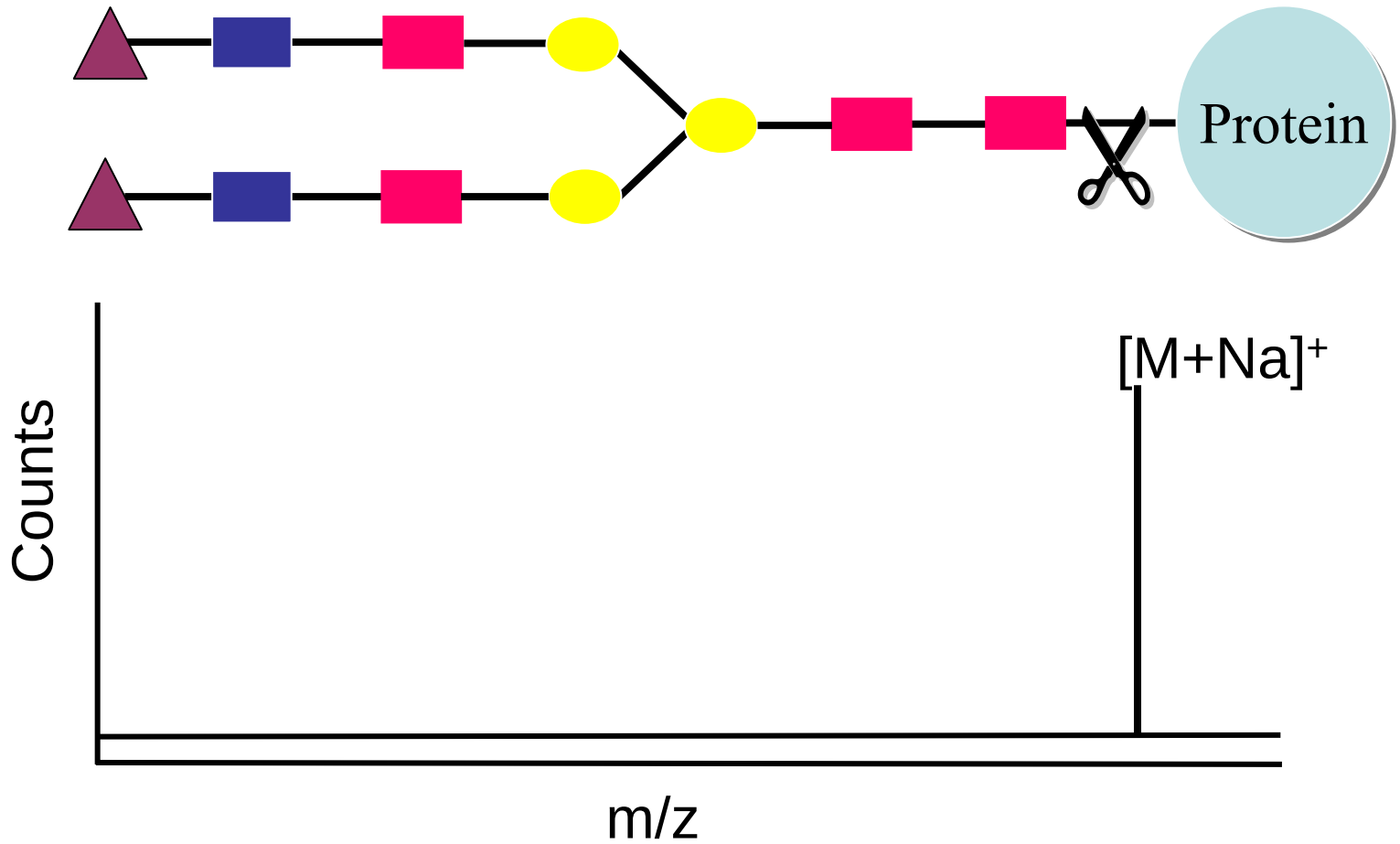
Galactose

N-Acetylglucosamine

Mannose

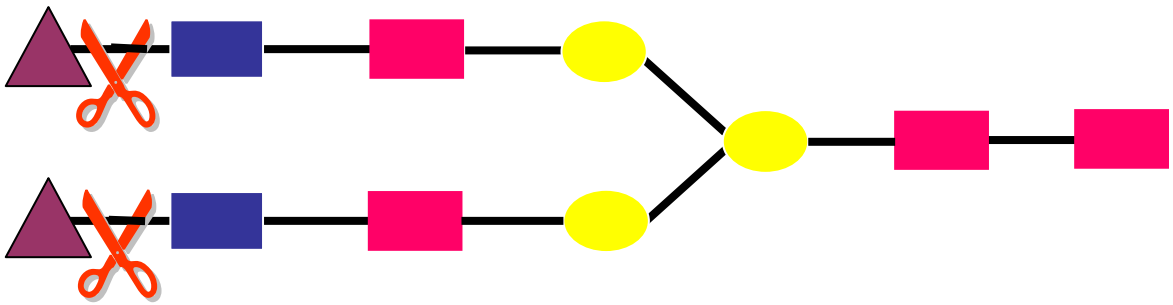
Glycomic Analysis

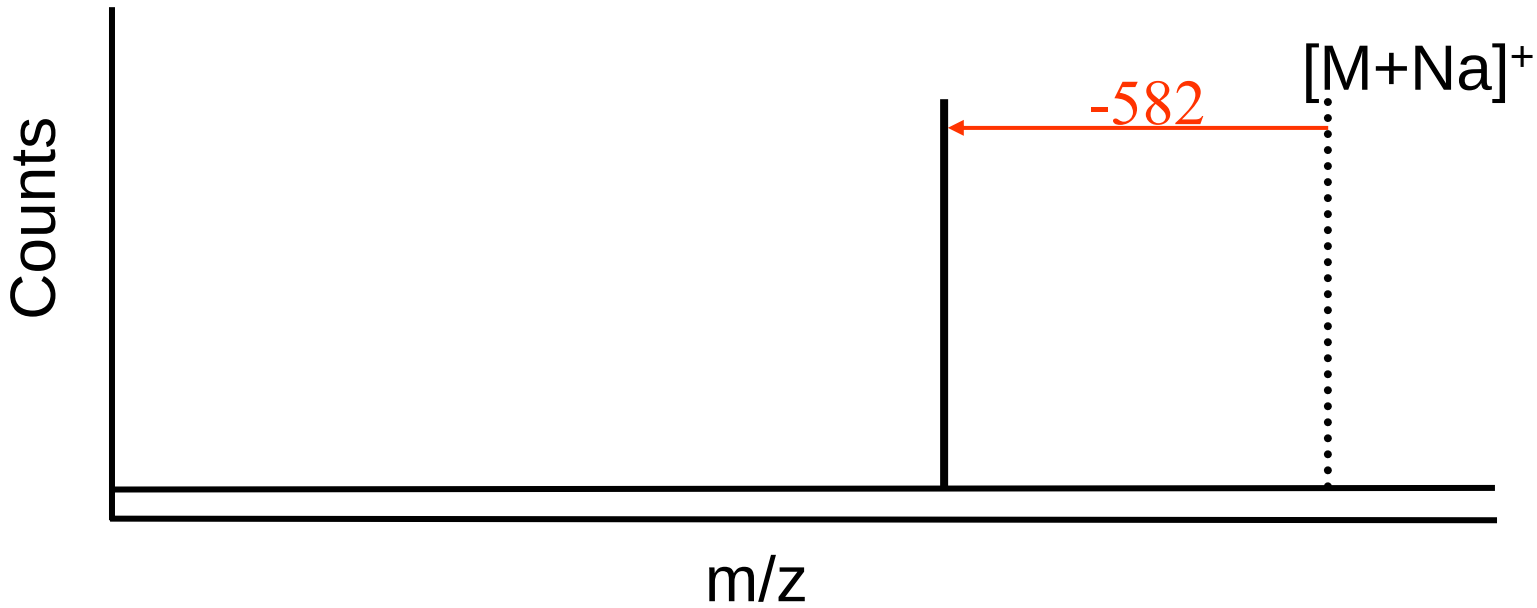
Enzymatic Sequencing (MALDI-MS)



Glycomic Analysis

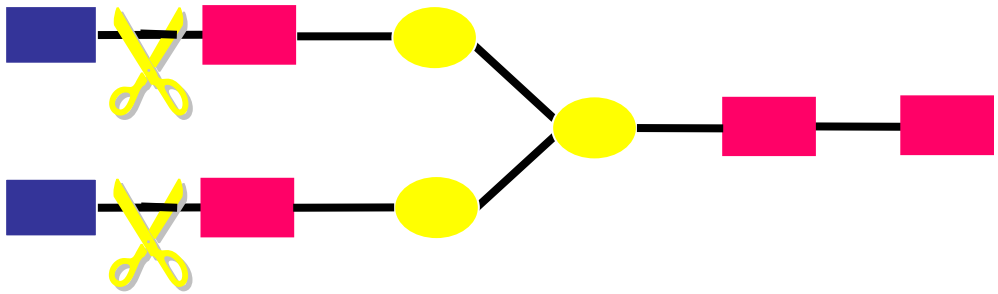
Enzymatic Sequencing (MALDI-MS)

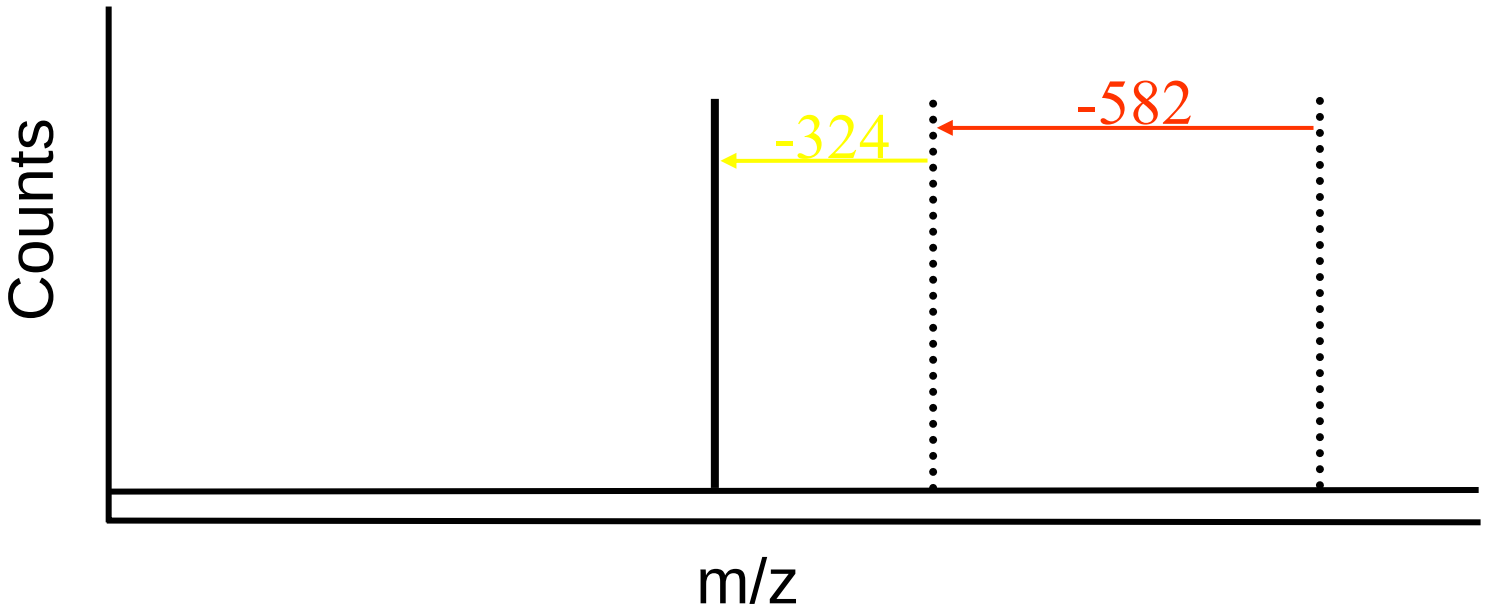




Glycomic Analysis

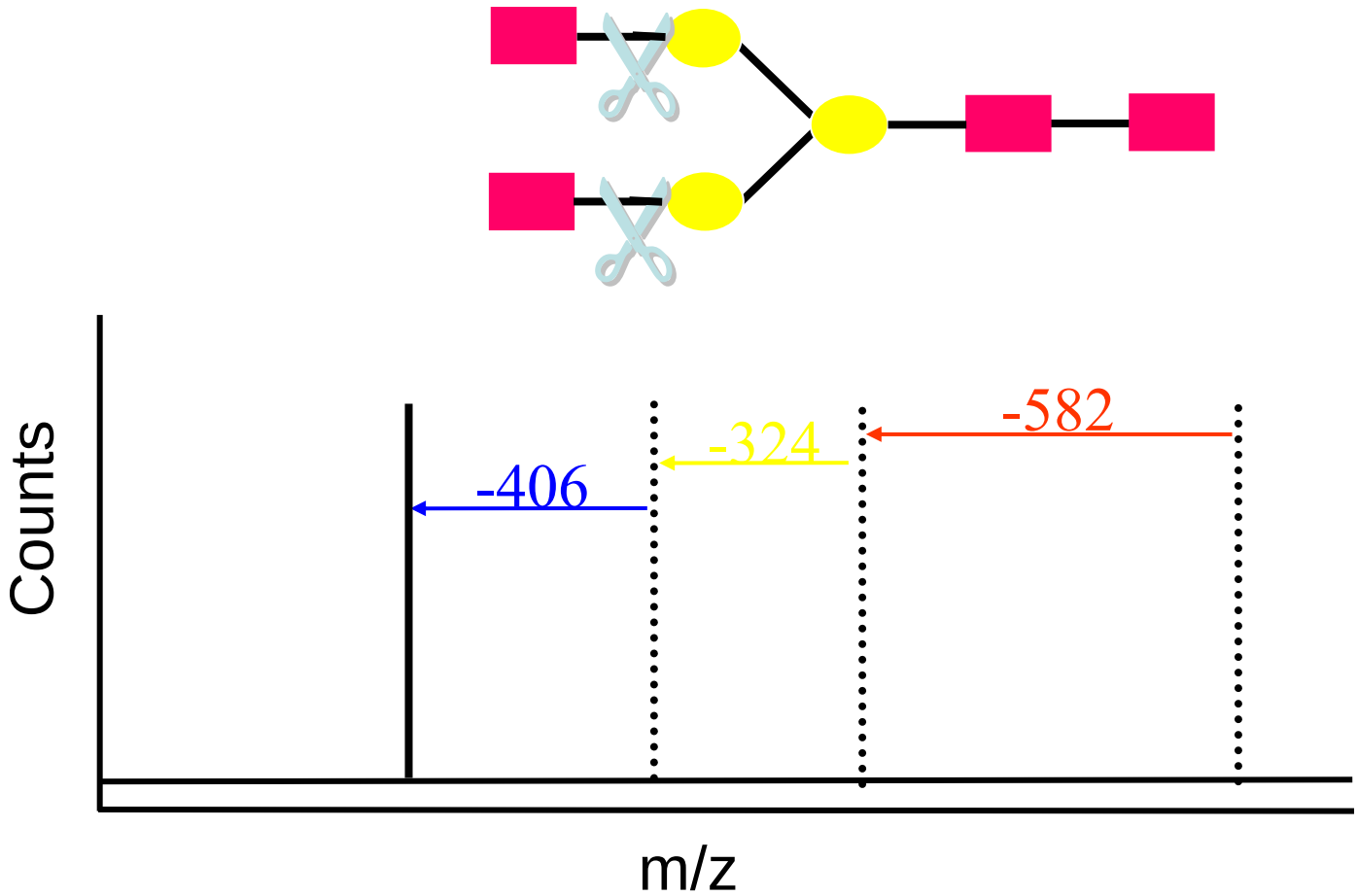
Enzymatic Sequencing (MALDI-MS)





Glycomic Analysis

Enzymatic Sequencing (MALDI-MS)



Glycomic Analysis

Enzymatic Sequencing (MALDI-MS)

Endo- and Exoglycosidase Specificity and Optimum pH

Enzyme	Specificity	Optimum pH
N-Glycosidase F, Recombinant (PNGase F)	Cleaves high mannose structures, hybride structures and complex structures	5.0-7.0
β-Galactosidase from *Diplococcus pneumoniae*	Hydrolyzes terminal galactose residues which are β(1-4)-linked to GlcNAc	6.0-6.5
N-Acetyl- β-D-glucosaminidase from *Diplococcus pneumoniae*	Cleaves Terminal GlcNAc residues that are β(1-2)-linked to Man	5.0
Neuraminidase From *Arthrobacter ureafaciens*	Cleaves terminal sialic acid residues that are α(2-3)-, α(2-6)- or α(2-8)-linked to Gal , GlcNAc .	5.0-5.5

Glycomic Analysis

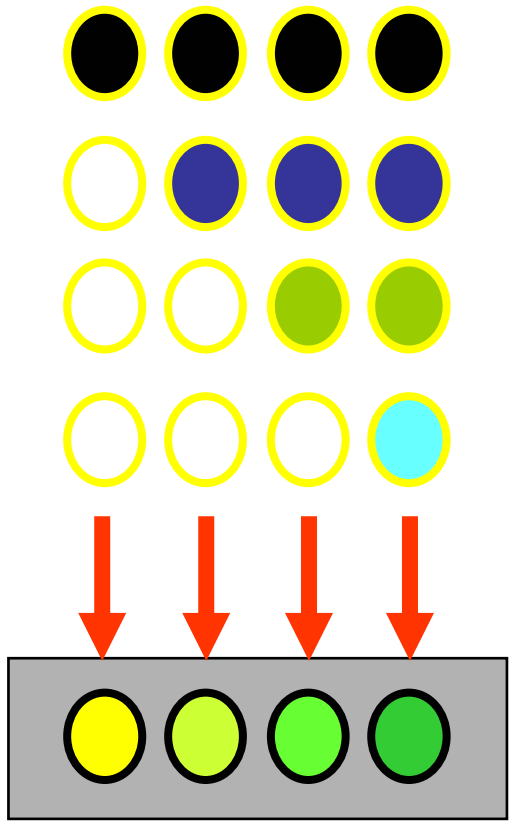
Enzymatic Sequencing (MALDI-MS)

N-glycosidase F

Neuraminidase

N-Acetyl-β-D-glucosaminidase

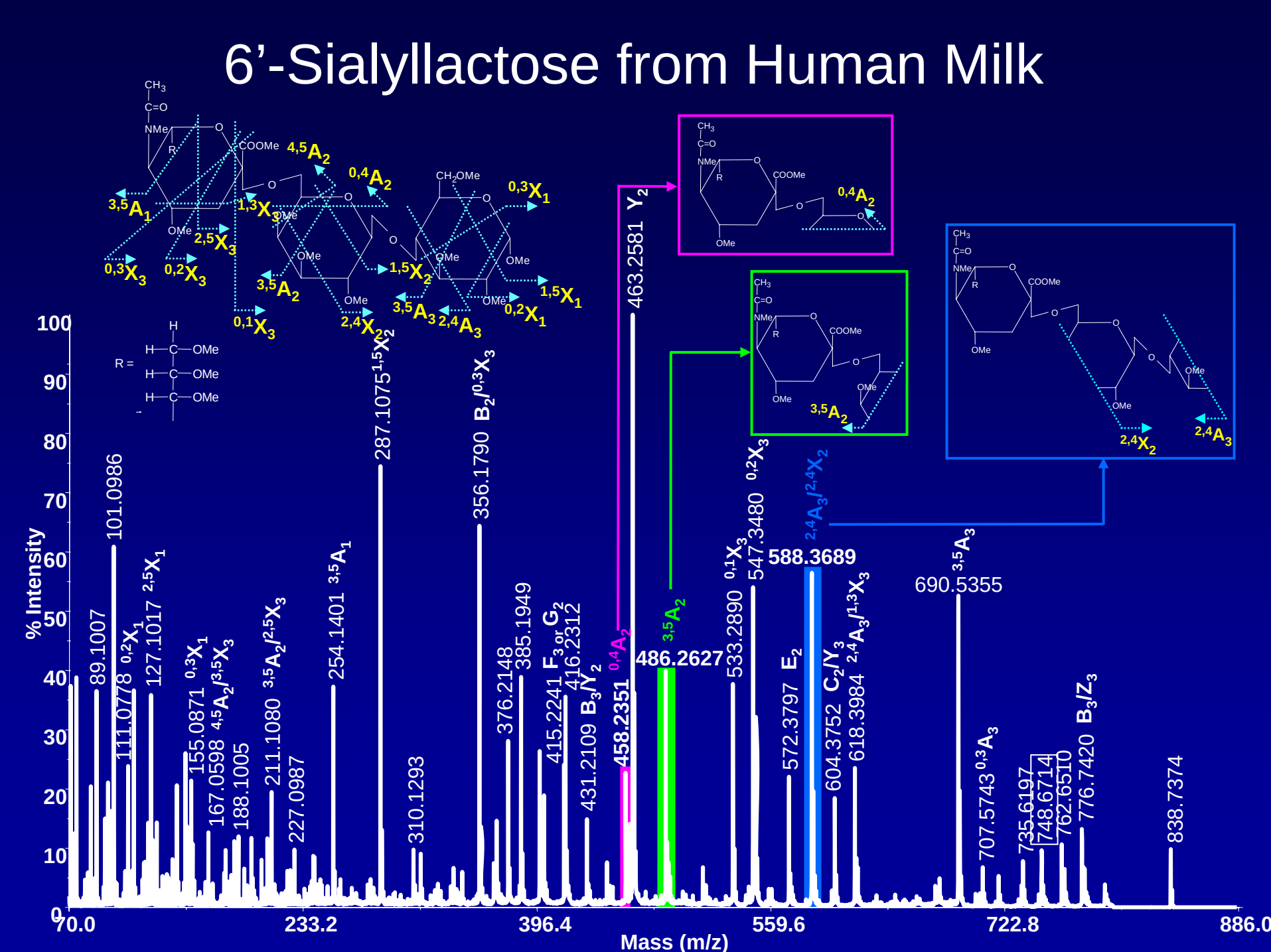
β-Galactosidase



MALDI PLATE

Y. Mechref, M.V. Novotny, ***Anal. Chem.***, 70 (**1998**) 455-463.

6’-Sialyllactose from Human Milk



Glycomic Analysis

Permethylation

Permethylation of Oligosaccharides for MS Analysis --

- allows simultaneous analysis of neutral and sialylated structures
- permits reversed-phase LC separation of permethylated structures
- enhances MSMS
- Simplifies MSMS interpretation

Glycomic Analysis

Permethylation

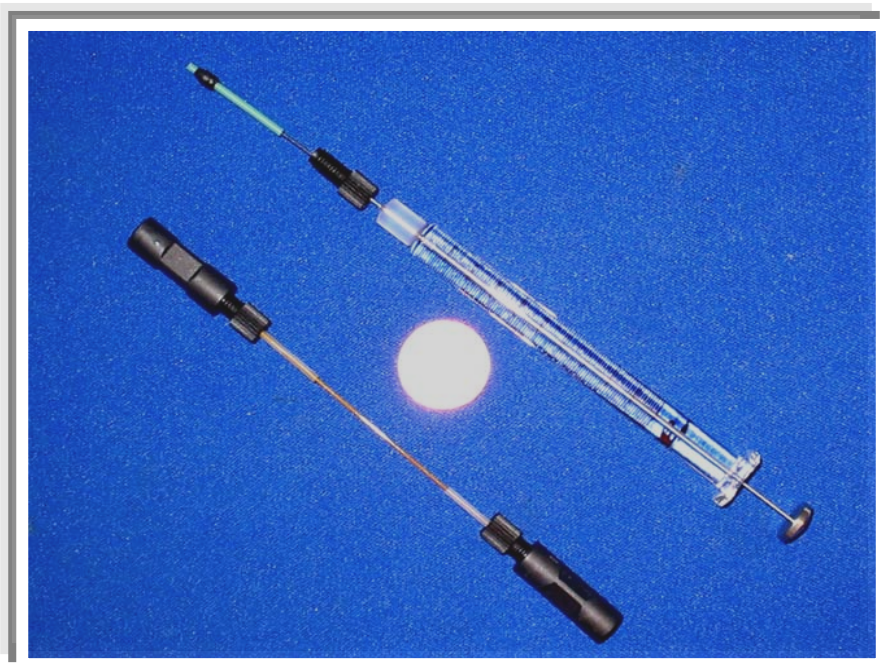
- Methylation of carbohydrates in dimethyl sulfoxide by mixing with powdered sodium hydroxide and methyl iodide

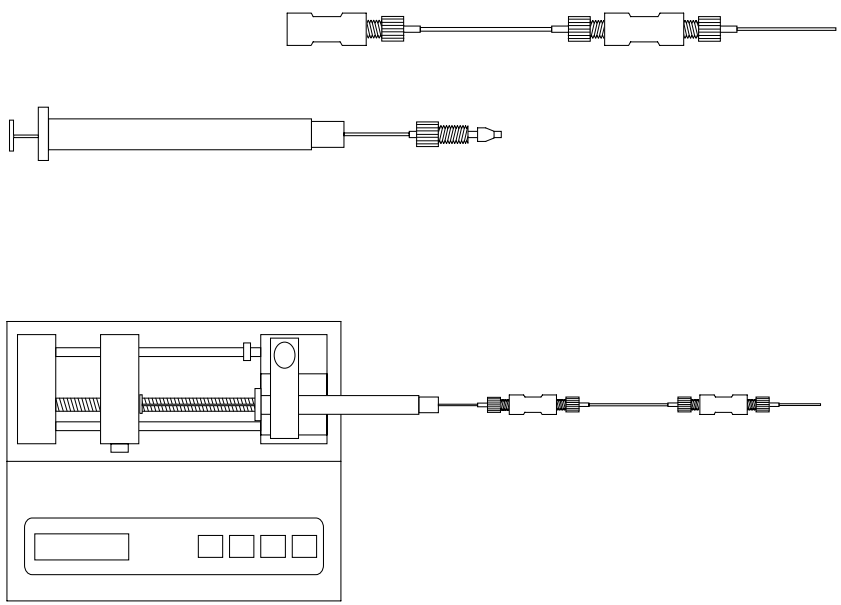
  (Ciucanu I, Kerek F *Carbohydr. Res.* **1984**, *131*, 209-217*)*

- Oxidative degradation and peeling reactions due to the high pH resulting from dissolving sodium hydroxide powder prior to liquid-liquid extractions

Glycomic Analysis

Permethylation





P. Kang, Y. Mechref, I. Klouckova M. V. Novotny, ***Rapid Commun. Mass Spectrom***, 19 (**2005**), 3421-3428.

Glycomic Profiling of Human Blood Serum Collected from Healthy Individuals and Breast Cancer Patients

Human Blood Serum

Protein Extraction

Reduction & Alkylation with TFE

Glycoproteins

Tryptic Digestion

Glycopeptides

PNGase F Digestion

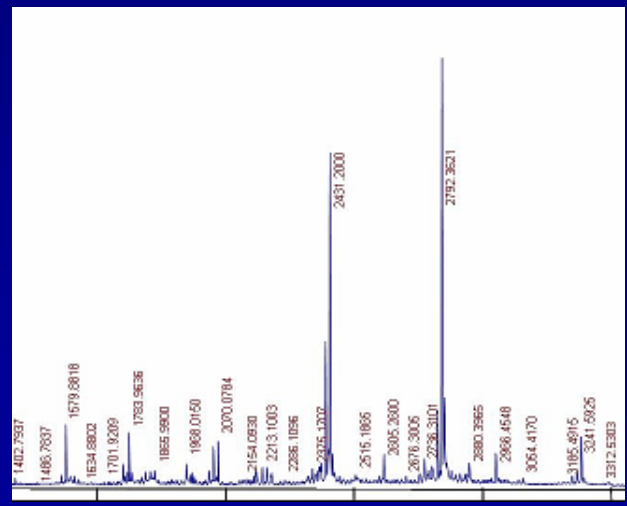
N-Glycans

Solid Phase Extraction

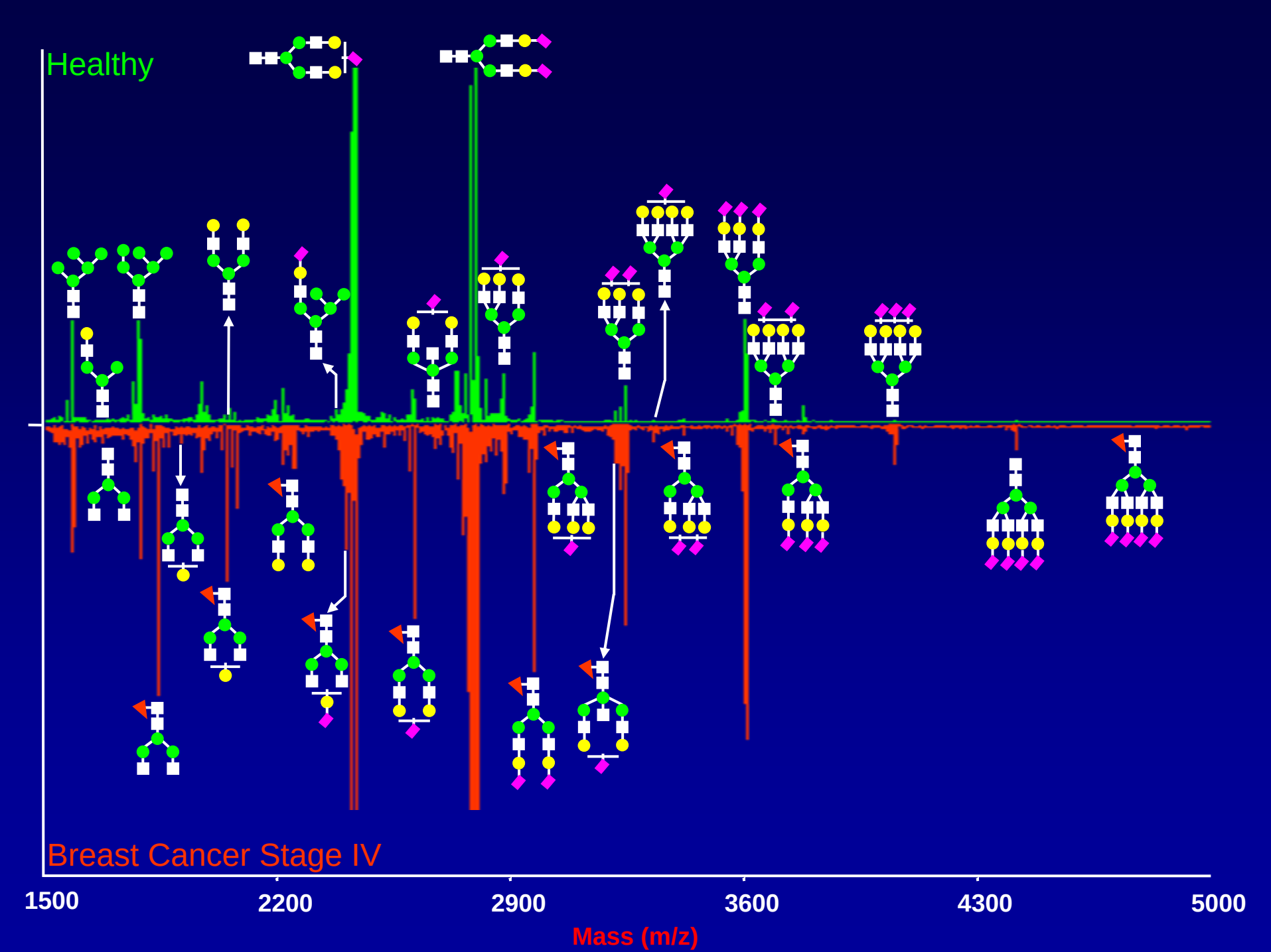
N-GLYCANS

Permethylation

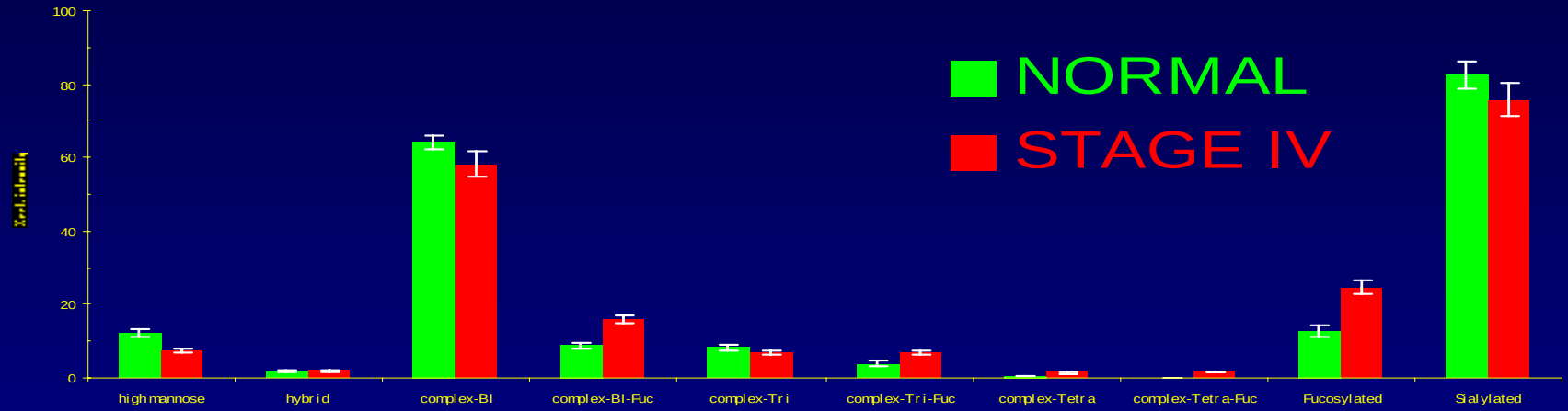
MALDI TOF / MS

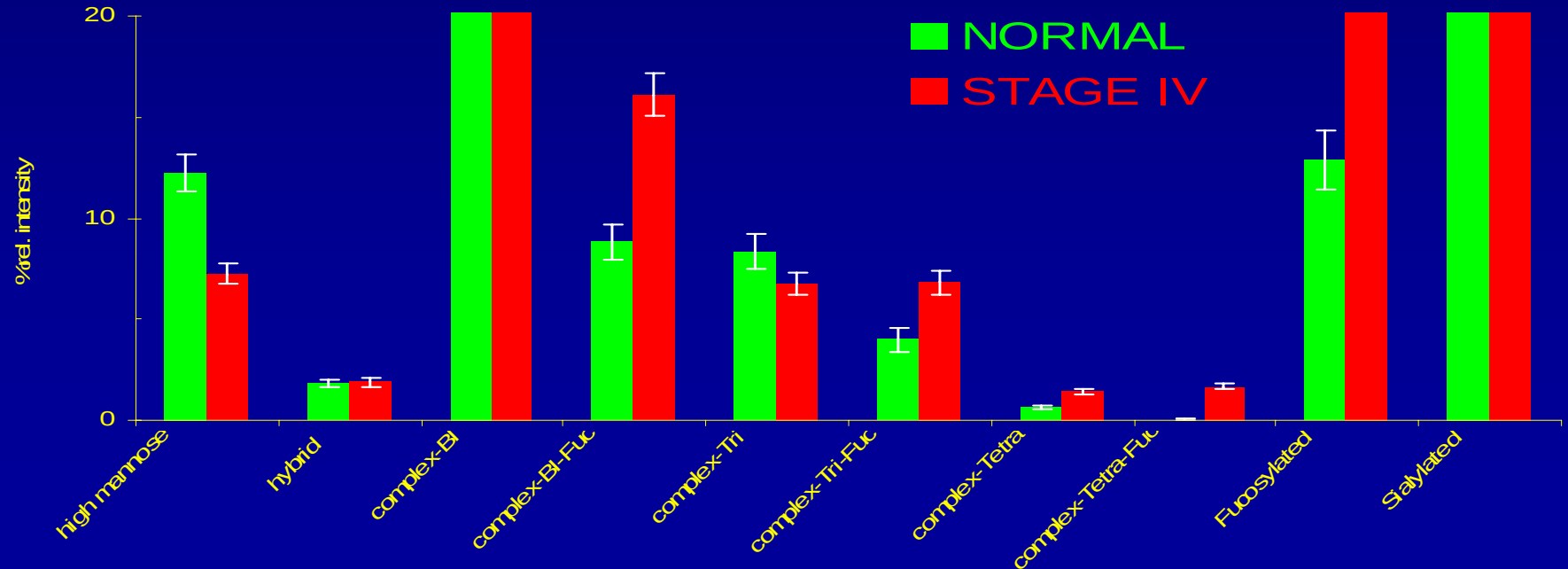


GLYCOMIC MAPS

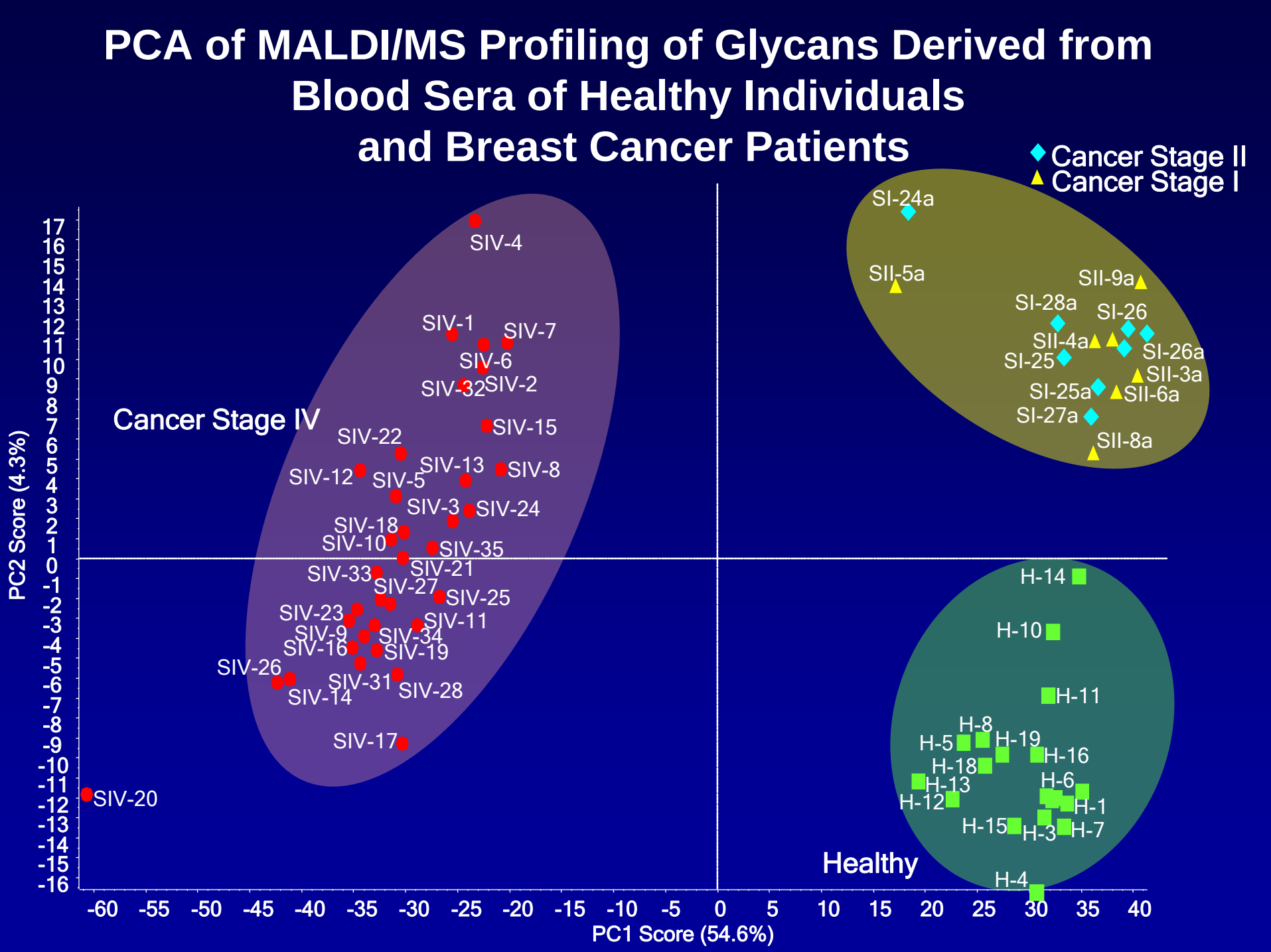

Healthy
Breast Cancer Stage IV
1500
2200
2900
3600
4300
5000
Mass (m/z)

Comparing Glycomic Profiles of Normal (n=27) vs. Stage IV of Breast Cancer (n=50)





PCA of MALDI/MS Profiling of Glycans Derived from Blood Sera of Healthy Individuals and Breast Cancer Patients



Glycomic Analysis

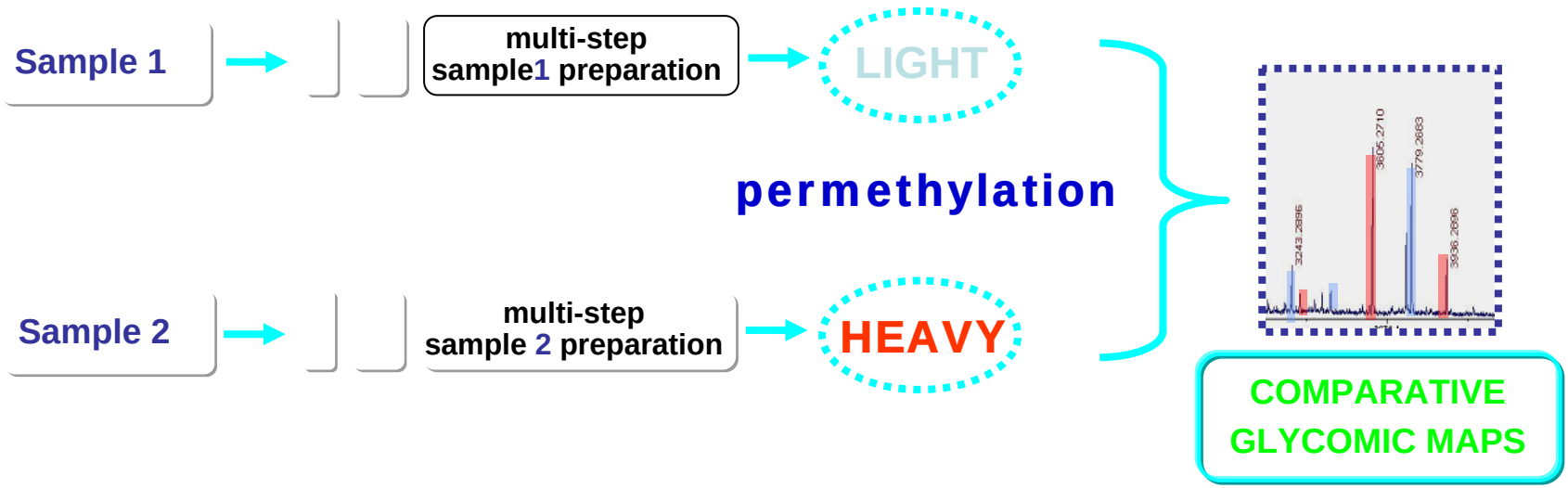
Comparative Glycomic Mapping (C-GlycoMAP)

– isotope labeling with **deuteromethyl iodide**
  - allows **comparative glycomic profiles** of two samples analyzed simultaneously.
  - determines changes in abundances of glycans

– Therefore, isotopic labeling will provide more precise comparison

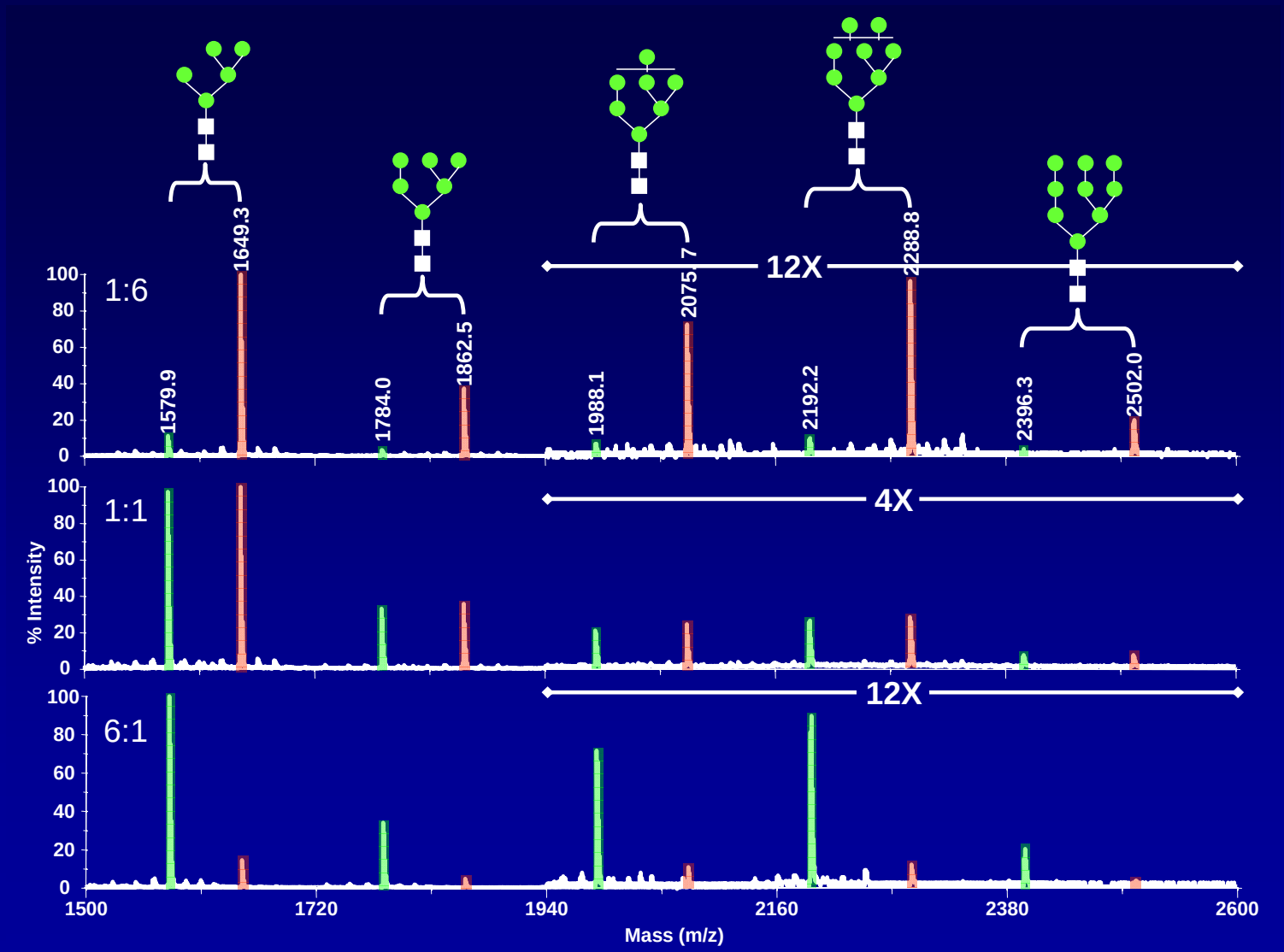
Glycomic Analysis

Comparative Glycomic Mapping (C-GlycoMAP)



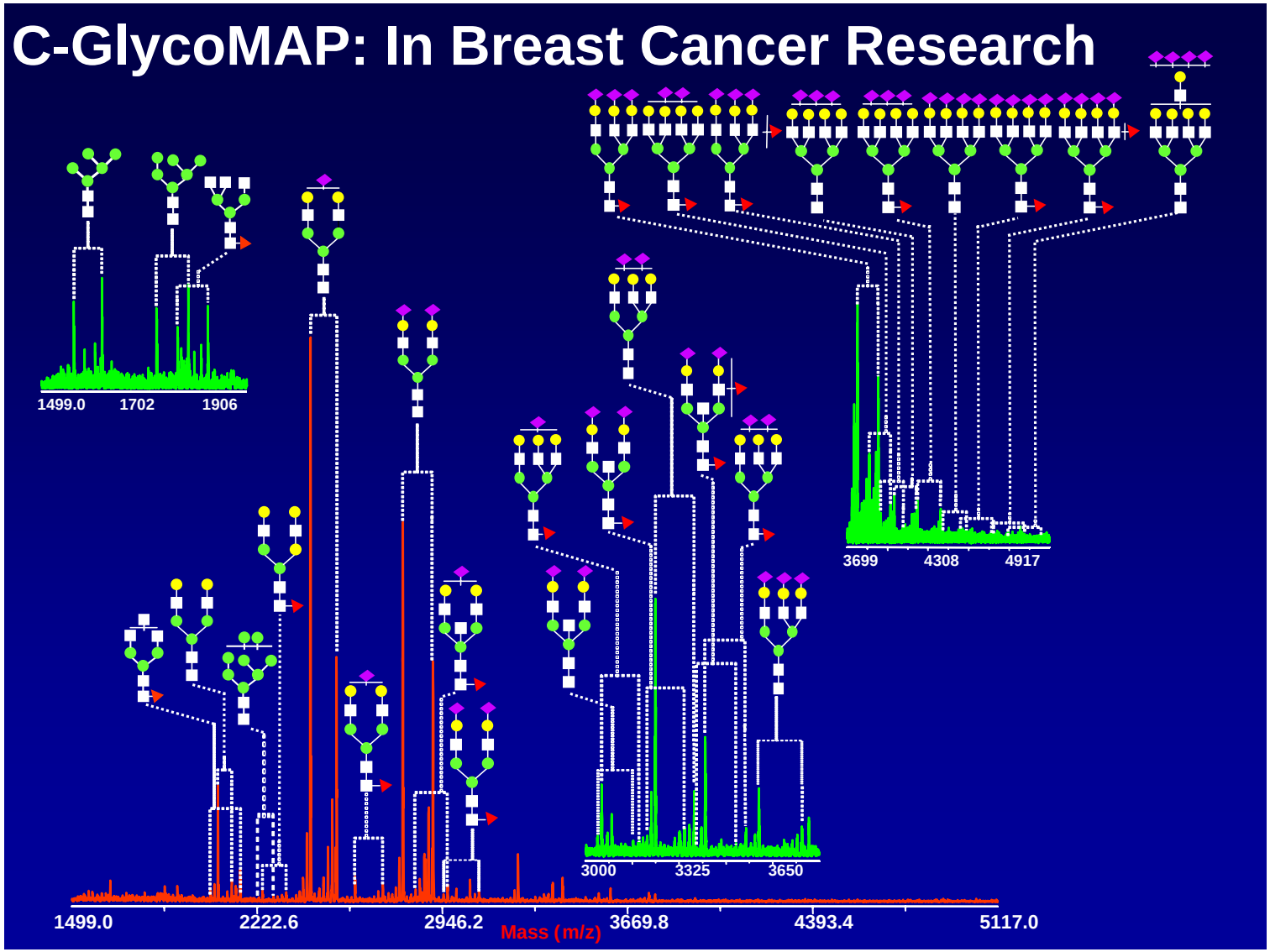
Glycomic Analysis

Comparative Glycomic Mapping (C-GlycoMAP)



Glycomic Analysis

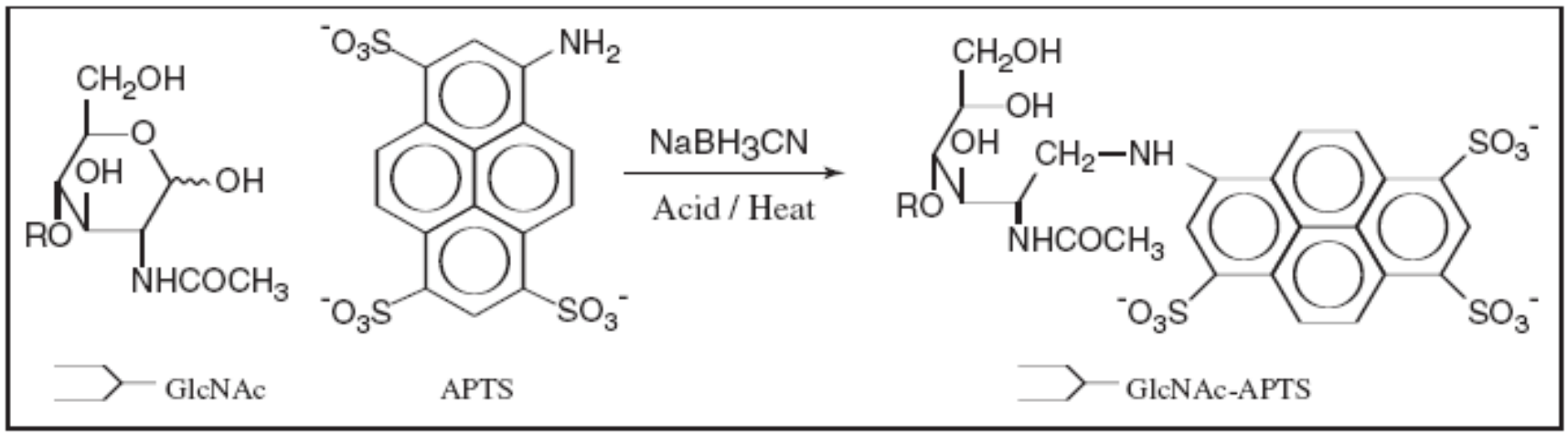
Comparative Glycomic Mapping (C-GlycoMAP)



Glycomic Analysis

Electrokinetically-driven Approaches (CE)

Labeling of Glycans by Reductive Amination

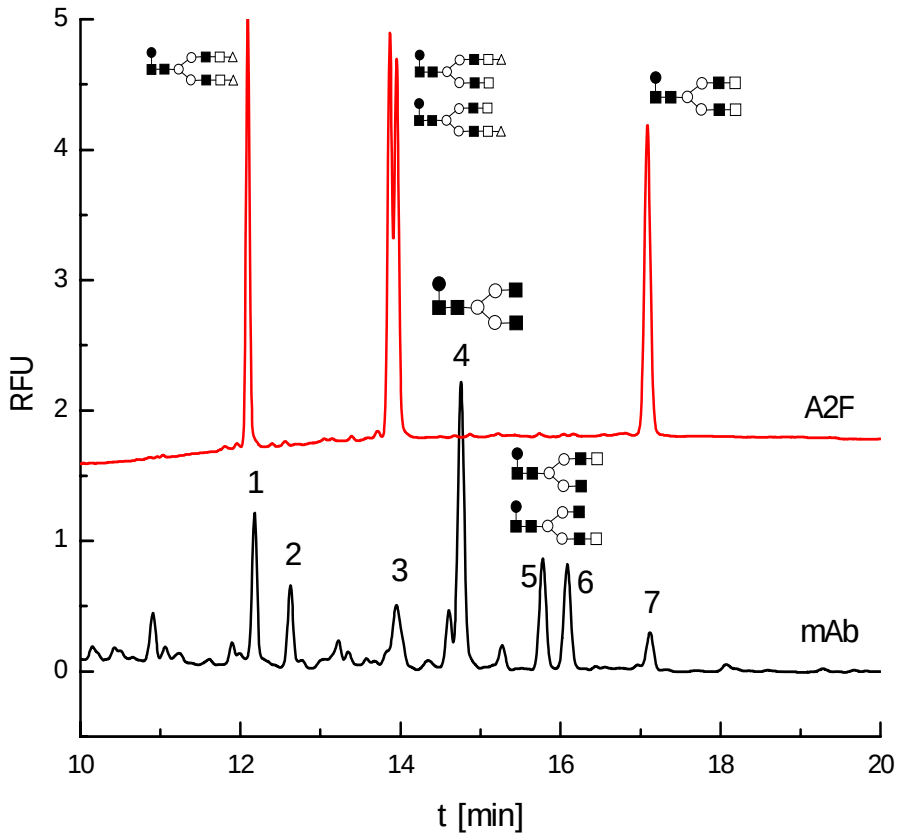


Reductive amination of the oligosaccharide with APTS (8-aminopyrene-1,3,6-trisulfonate).

Glycomic Analysis

Electrokinetically-driven Approaches (CE)

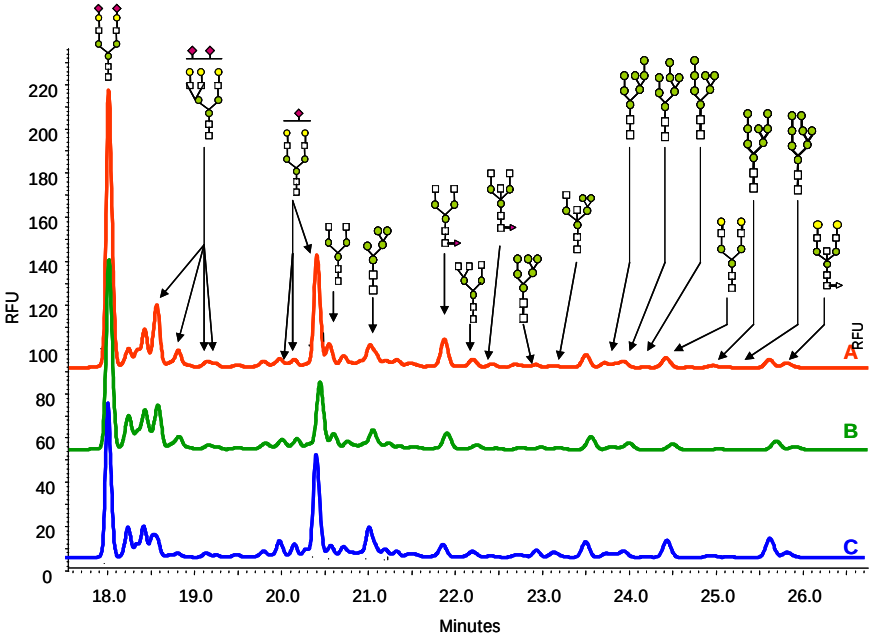
- CE profile of APTS-labeled glycans derived from mAb. The upper trace represents standard core-fucosylated biantennary/disialylated, monosialylated, and asialylated glycans. Conditions: column, polyacrylamide-coated 50/365 mm ID/OD; length, 50.5 cm total, 40.5 cm effective length; temperature, 257C; injection pressure, 0.5 psi for 5.0 s; voltage, 15 kV anodic electroosmotic flow; lex 488 nm, lem 520.



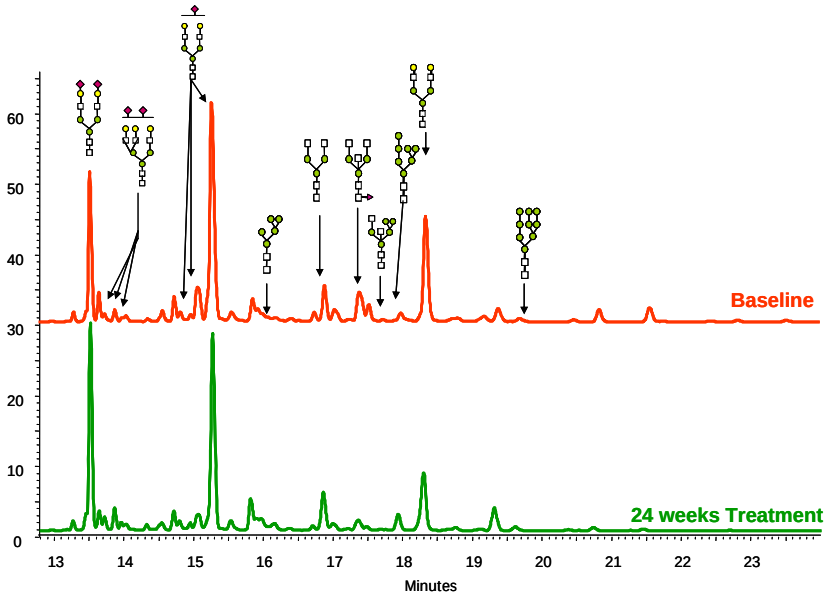
Y. Mechref, J. Muzikar, M.V. Novotny, ***Electrophoresis***, 26 (**2005**) 2034–2046

Glycomic Analysis

Electrokinetically-driven Approaches (CE)

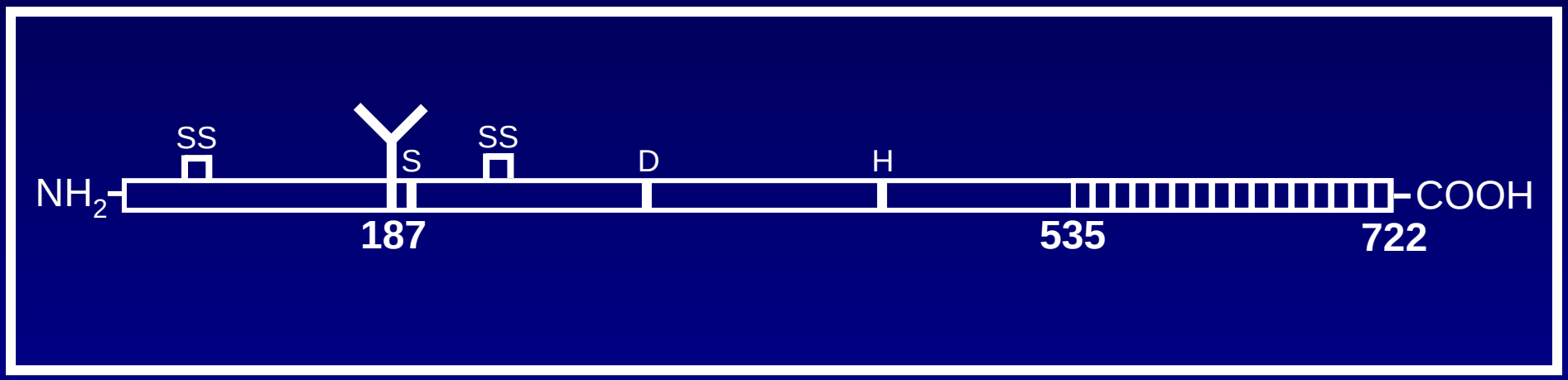


CE-LIF of APTS-labled glycans derived from human blood serum of stage IV breast cancer patients with short (A), intermediate (B) and long (C) survival time



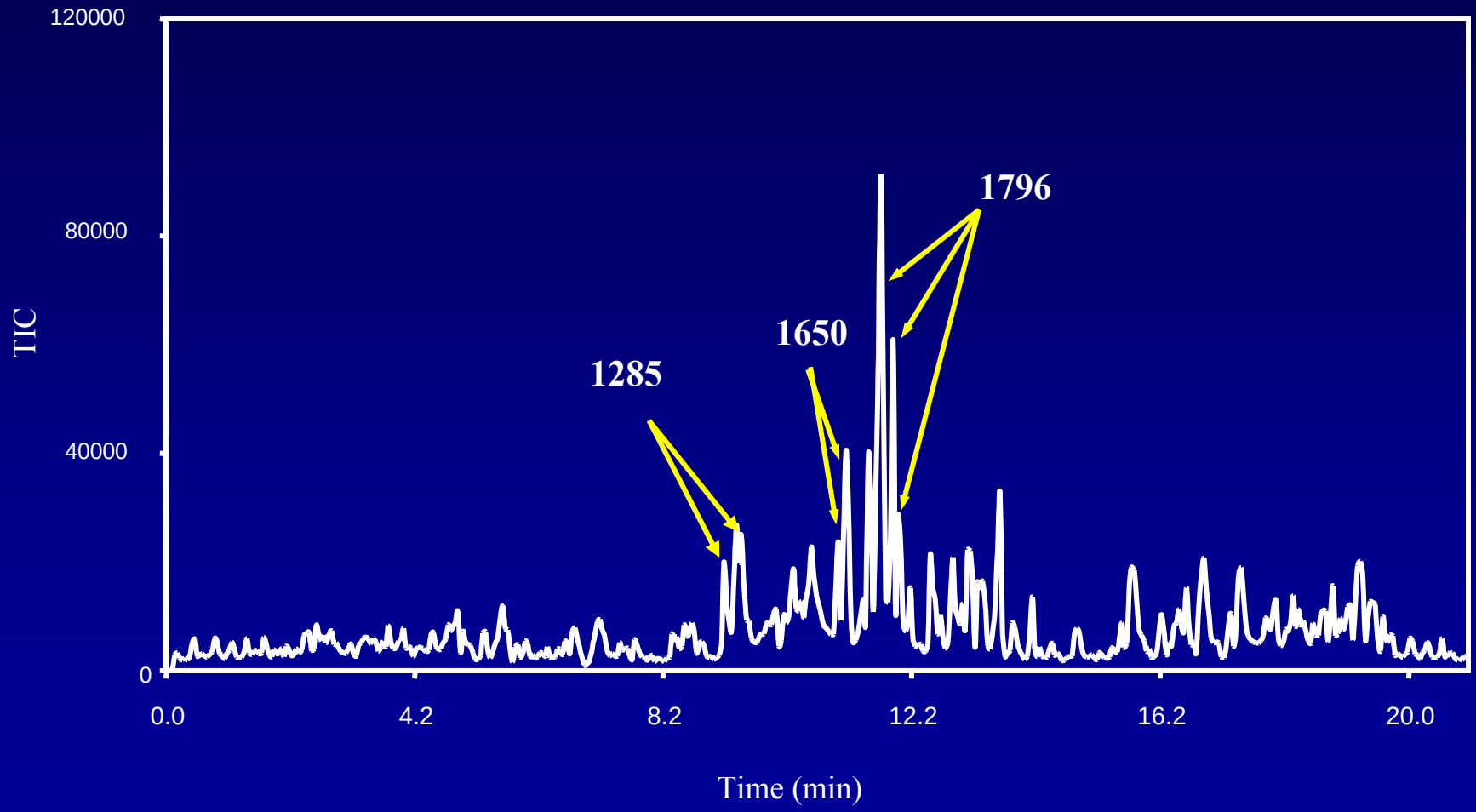
CE-LIF of APTS-labled glycans derived from human blood serum of prostate cacner patients before (upper trace) and after (lower trace) treatment .

Schematic of human BSSL



S serine, D aspartate, H histidine, Y position of N-linked Glycosylation site, SS disulfide bridge

CEC SEPARATION OF N- AND O-GLYCANS DERIVED FROM BSSL

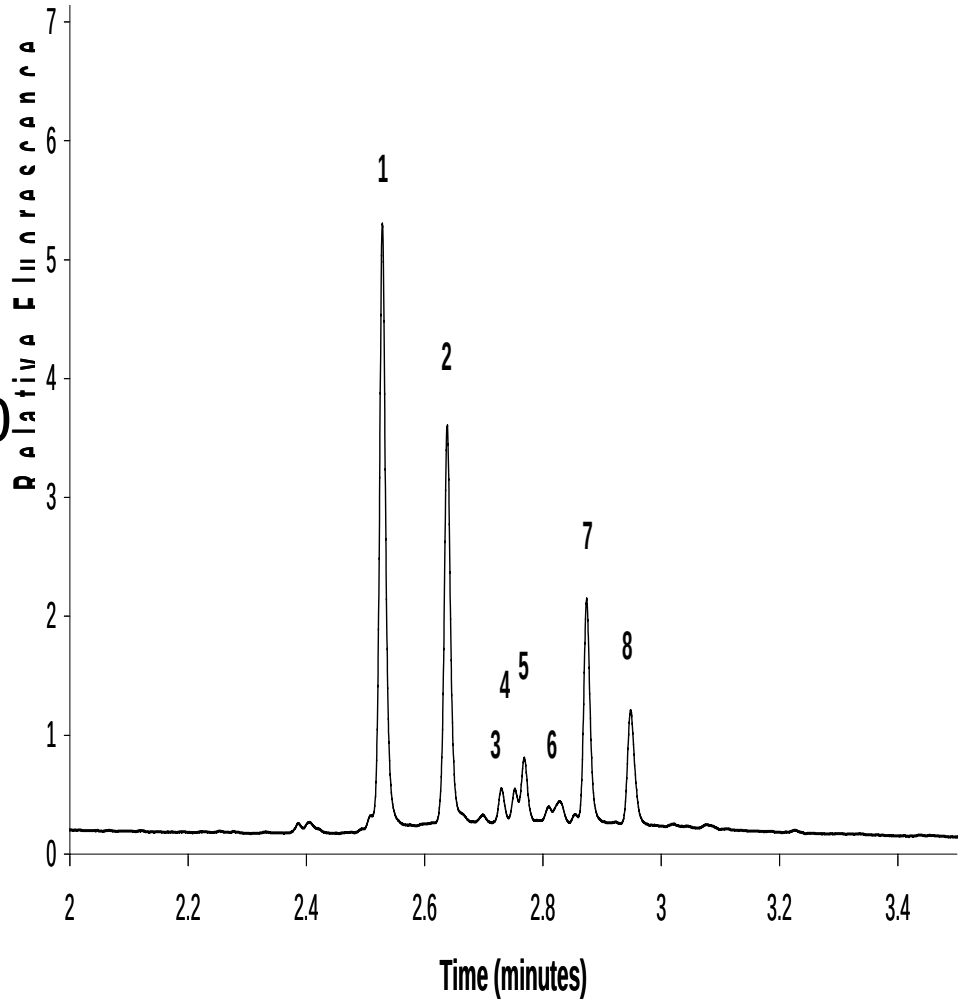


A.H. Que, M.V. Novotny, *Anal. Bioanal. Chem.*, **375**, 599-608, (2003).

Glycomic Analysis

Electrokinetically-driven Approaches (CE)

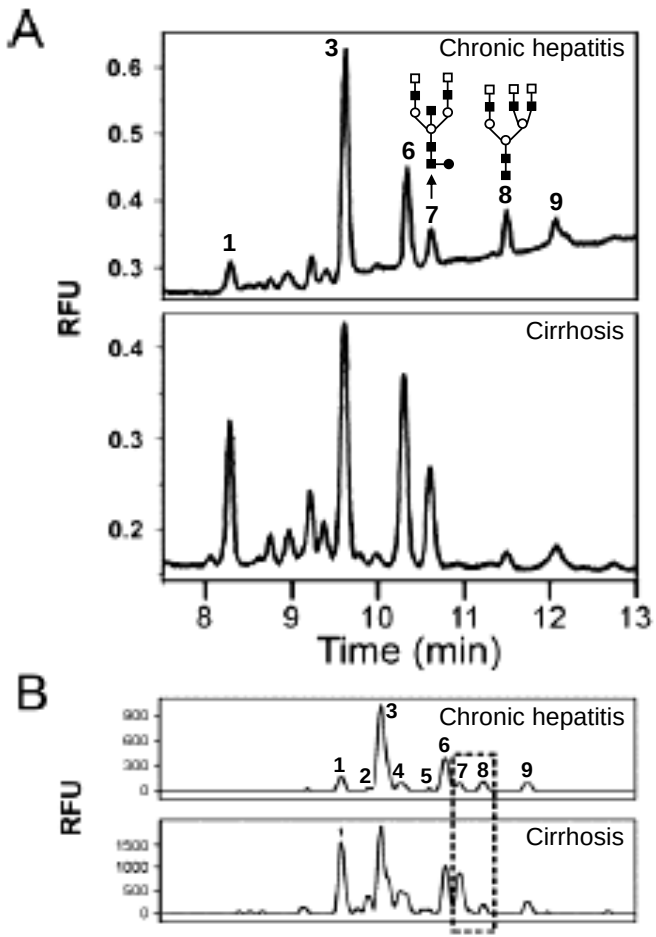
- **Conditions:**
  22 cm effective length, 31 cm total length; 20 mM Hepes, 1 mM Phosphate buffer (pH 6.5); Field Strength, 750 V/cm; Peak Identities;
  1) Man 5,
  2) Man 6,
  3-5) Man 7,
  6-7) Man 8,
  8) Man 9



Glycomic Analysis

Chip-based Approaches (CE)

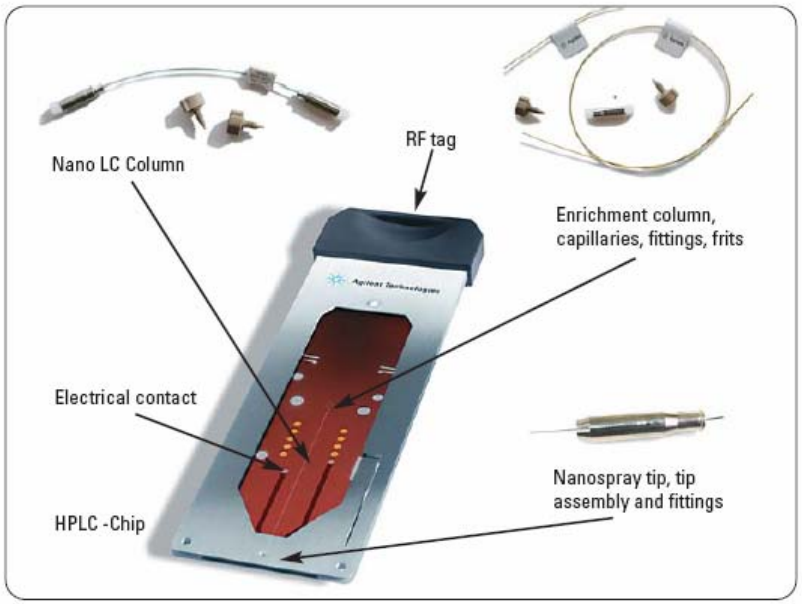
- (A) ME profiling of serum samples from a noncirrhotic chronic hepatitis patient (upper trace) and cirrhotic patient (lower trace). (B) Profiling of the same samples using the ABI377 gel-based DNA-sequencer. Symbols: , *N*-acetylglucosamine; ○, mannose; □, galactose; ●, fucose.



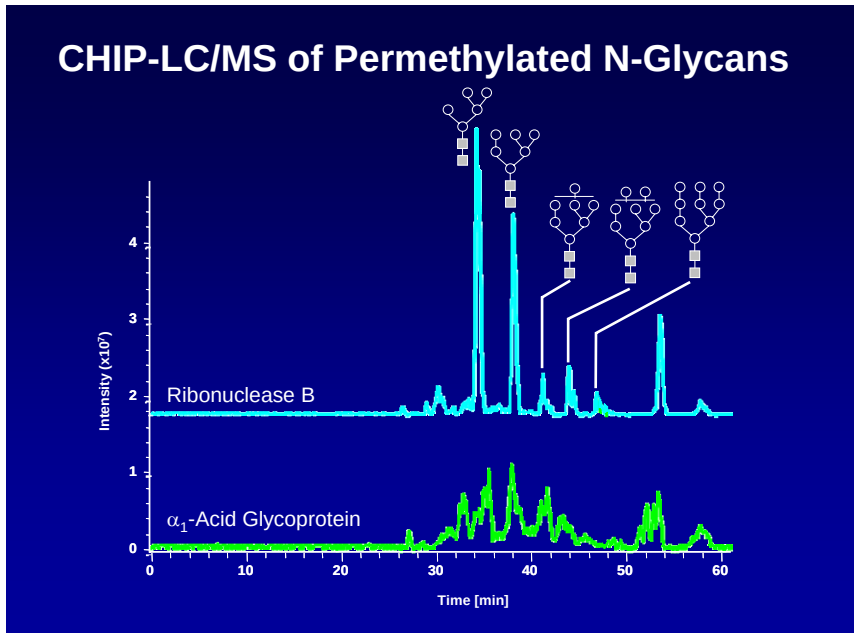
N. Callewaert, H. van Vlierberghe, A. van Hecke, W. Laroy, J. Delanghe, R. Contreas, ***Nature Med.*** 10 (2004) 429.

Glycomic Analysis

Chip-based Approaches (LC)



Integrated HPLC-Chip eliminates tedious and complex connections and delivers uncompromised chromatographic separations.



CONCLUSIONS

- Structural analysis of glycoproteins has come a long way during the last decade in terms of sensitivity and information content. It is still a multimethodological task.
- The current emphasis is on working with complex mixtures (biological fluids and tissue extracts) on-line – building complete analytical platforms for functional glycomics and glycoproteomics.
- In various searches for biomarkers, the glycomic approach could be an easier route to diagnostic and prognostic information, but high-sensitivity glycoproteomics is needed for explanations in terms of biochemistry and physiology.