Glycan Analysis by Mass Spectrometry: Instrumentation

- **Overview of Mass Spectrometry**
  - **Ionization sources**
  - **Mass analyzers**
- **Common Configurations for Glycan Analysis**
  - **LC-MS/MS**

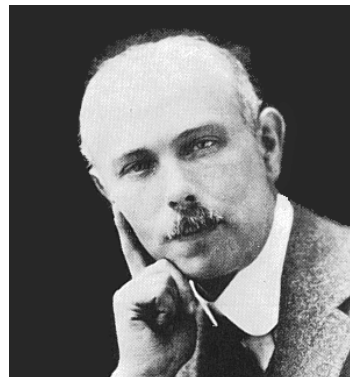**Adam M. Hawkridge**

School of Pharmacy

Spring 2019

Mass Spectrometry is Not a New Technique

Joseph J. Thomson
(1856-1940)

Francis W. Aston
(1877-1945)

- 1906 Nobel Prize in Physics for the discovery of the electron
- Conceptualized the idea of mass spectrometry in 1897
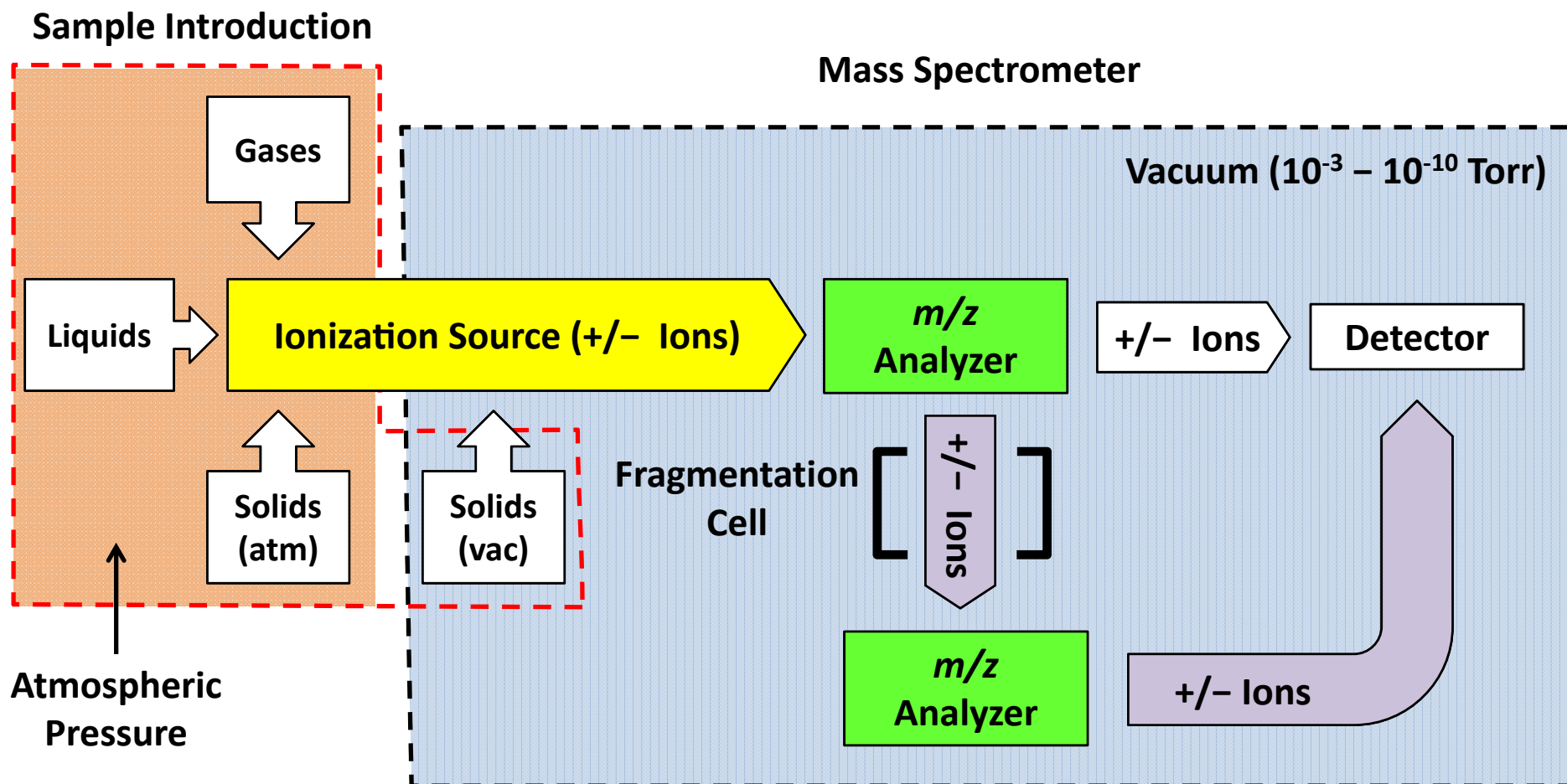- Prophesized that mass spectrometry would be invaluable for chemical analysis

- Became an assistant to J.J. Thomson at the Cavendish Laboratory in 1909
- Developed the first mass spectrograph in 1919 following service in World War I and used it to measure 212 isotopes
- 1922 Nobel Prize in Chemistry

Francis W. Aston's 3rd Generation Mass Spectrograph (1937)

What are the Primary Components of a Mass Spectrometer?

Sample Introduction

Mass Spectrometer

Vacuum ($10^{-3} - 10^{-10}$ Torr)

Gases

Liquids

Ionization Source (+/− Ions)

m/z Analyzer

+/− Ions

Detector

Solids (atm)

Solids (vac)

Fragmentation Cell

[+/− Ions]

m/z Analyzer

+/− Ions

Atmospheric Pressure

Mass Spectrometry was Limited to Volatile Low MW Analytes for 1st 70 Years

Early Ionization Techniques

- 1918 Electron Ionization
- 1966 Chemical Ionization
- 1969 Field Desorption
- 1974 Atmospheric Pressure Chemical Ionization
- 1976 Plasma Desorption
- 1980 Inductively Coupled Plasma MS
- 1981 Fast Atom Bombardment

"Harsh" or "Hard" Ionization

Contemporary Ionization Sources in Biological Mass Spectrometry

...the Proteomics and Metabolomics Era Begins

- 1984 **Electrospray Ionization (ESI)**
- 1987 Matrix-Assisted Laser Desorption Ionization (MALDI)

"Soft" Ionization

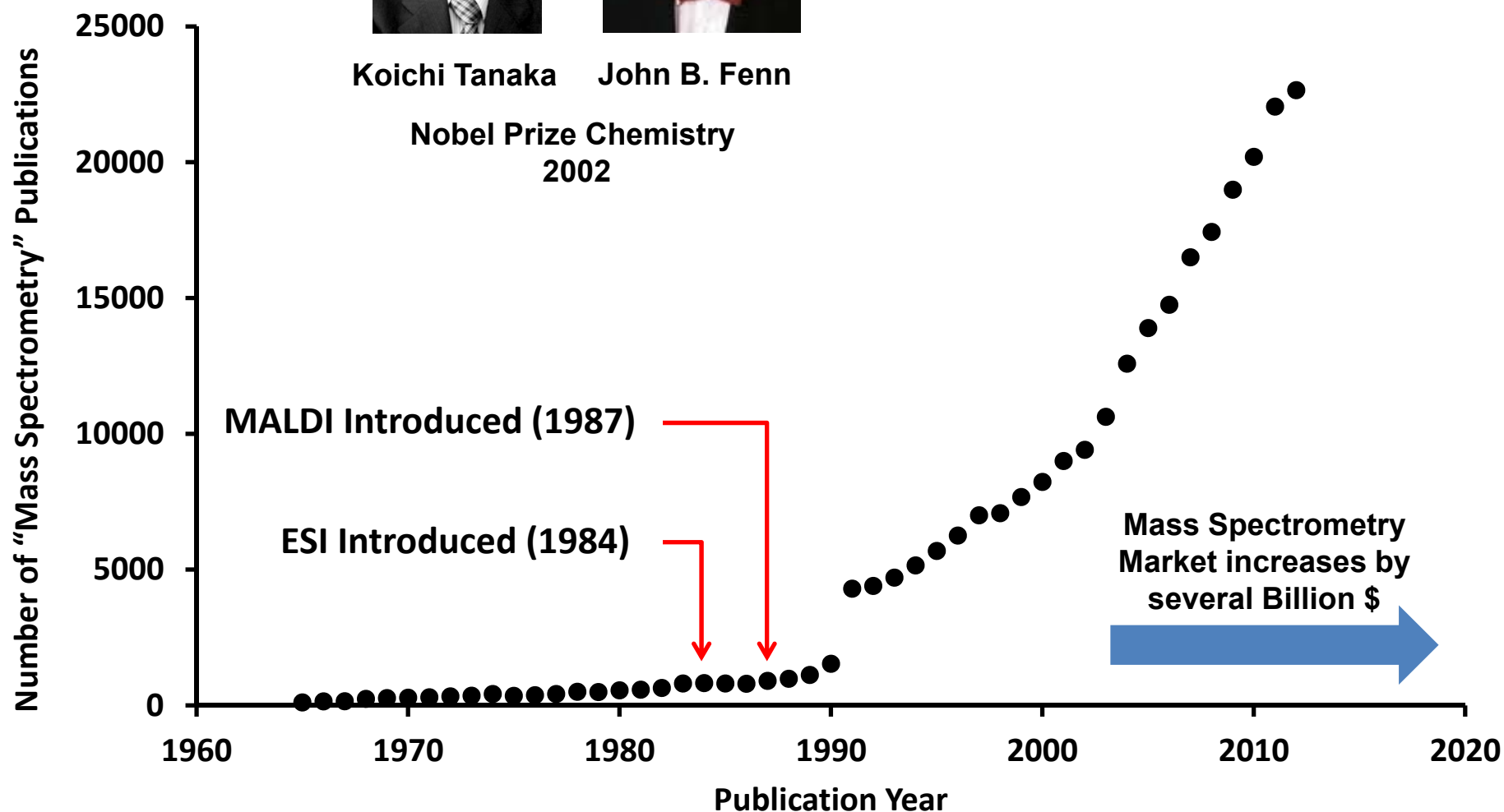ESI and MALDI Have Revolutionized Mass Spectrometry

Historical Development of MALDI

- Starting in the late 1960's, researchers worked for many years to develop laser-based methods to desorb and ionize biomolecules for mass spectral detection. Several techniques were developed such as laser desorption ionization (LDI) and resonance-enhanced multiphoton ionization (REMPI). These techniques were never able to desorb and ionize large biomolecules (>2000 Da).

- **Karas et al., *Analytical Chemistry*, 1985, 57, 2935-2939 and Karas et al., Int. J. Mass Spectrom. Ion Proc. 1987, 78, 53-68.**

  - Started using organic acids to form crystalline matrices for ionizing small organic molecules and biomolecules (<2000 Da).

- **Tanaka et al. *2nd Japan-China Joint Symposium on Mass Spectrometry*, Osaka, Japan, Sept. 15-18, 1987 and Tanaka et al. *RCMS* 1988, 2, 151-153.**

  - Used a glycerol/metal mixture as a matrix to demonstrate the ionization and subsequent mass spectral detection of lysozyme (14.3 kDa) and chymotrypsinogen (25.7 kDa) (experimental approach was referred to as soft laser desorption or SLD).

- **M. Karas and F. Hillenkamp, *Analytical Chemistry*, 1988, 60, 2301-2303.**

  - First demonstration of the use of organic acid matrices for desorbing and ionizing large, intact biomolecules. This approach later came to be known as MALDI.

- **½ the Nobel Prize in Chemistry, 2002 awarded to Tanaka and Fenn.**

Advantages of MALDI

- Amenable to off-line collection of liquid separations and preservation for extended interrogation times (e.g., multiple analyses of the same sample using potentially multiple MS platforms)

- Less susceptible to ionization suppression from salts and surfactants in biological systems relative to ESI.

- Soft ionization conditions

- Amenable to extremely high throughput analysis (>100 samples/hr)

- Practical mass range of up to 300,000 Da

- Low femtomole detection limits are routine
  - Sub-femtomole detection limits possible

- Produces primarily singly charged species thus spreading detected species over wider *m/z* window and making interpretation simpler

MALDI Overview

- Matrix-assisted Laser Desorption/Ionization (MALDI).
- Generally uses a pulsed, UV laser, but other lasers are also used..
- IR MALDI is useful when more sample penetration is needed (TLC plates).
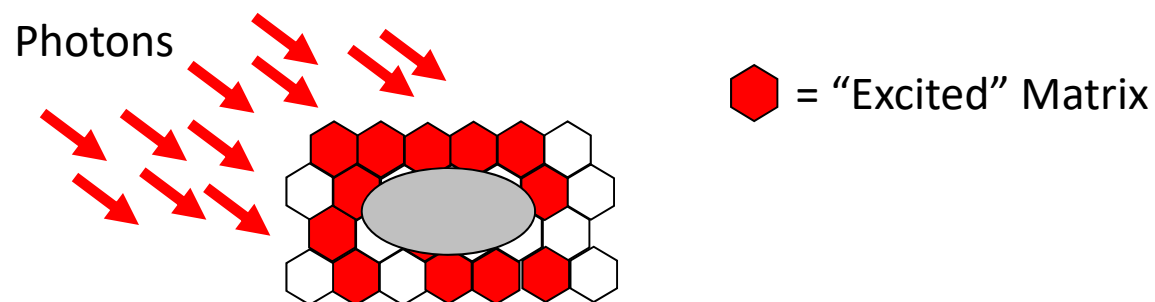- Ions are formed by proton transfer between the excited matrix ions and the analyte.
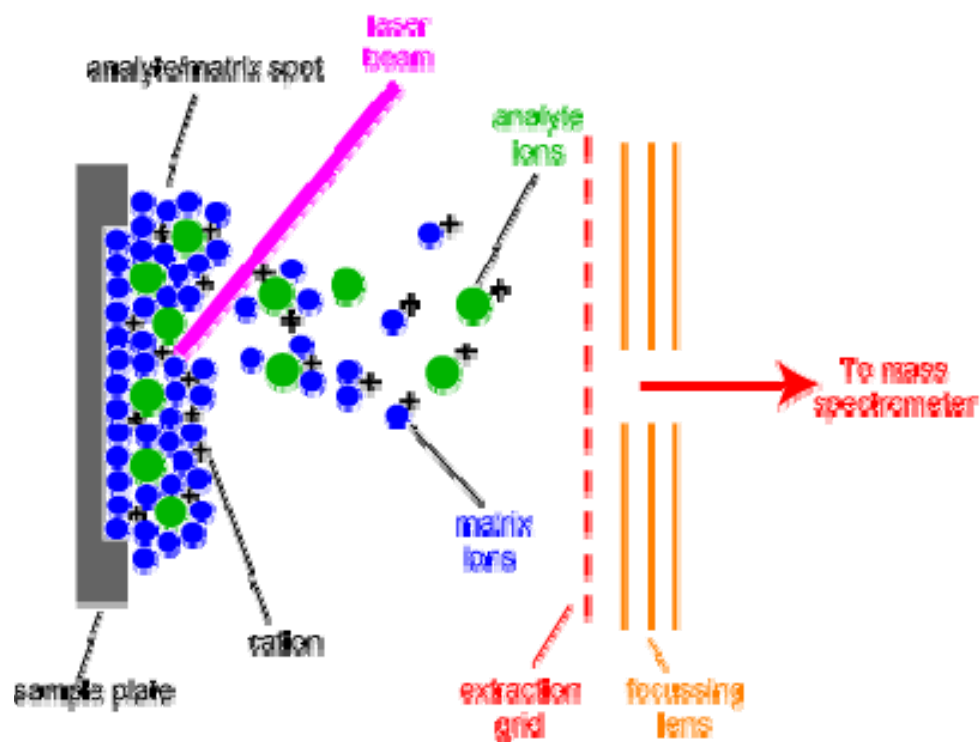
Essential Functions of the Matrix

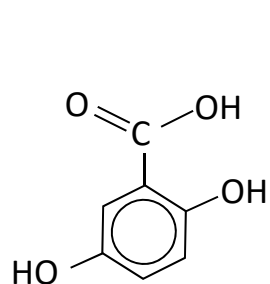- **Isolate and encase the analyte molecules (analogous to a solvent shell)**

= Analyte

= Matrix
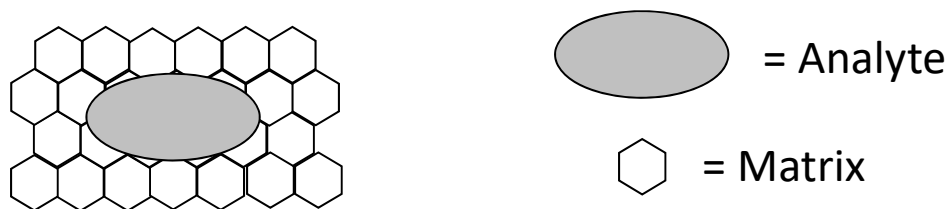
- **Absorb the laser energy via electronic or vibrational coupling/excitation**
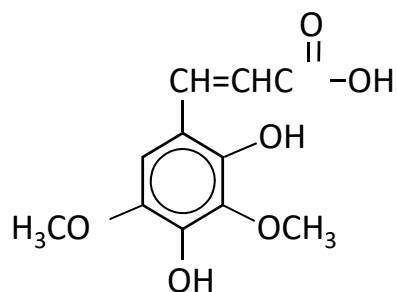
Photons

= "Excited" Matrix

- **Facile desorption from the condensed phase WITH the analyte molecules but WITHOUT destructive heating of the analyte molecules ("softness")**
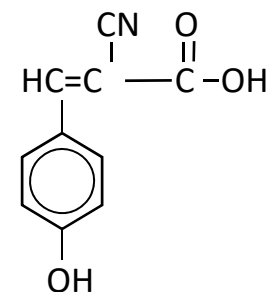
- **Efficient ionization of analyte molecules**

Karas et al. J. Mass Spectrom. 2000, 35, 1-12.
K. Dreisewerd Chem. Rev. 2003, 103, 395-425
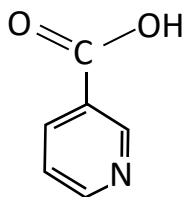
Common Organic Matrices

2,5-Dihydroxybenzoic acid (DHB)
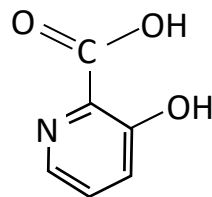**(Gentisic Acid)**

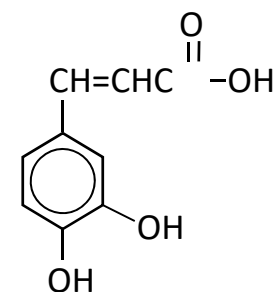**3,5-dimethoxy-4-hydroxycinnamic acid (SA)**
(Sinapinic Acid)

α-cyano-4-hydroxycinnamic acid
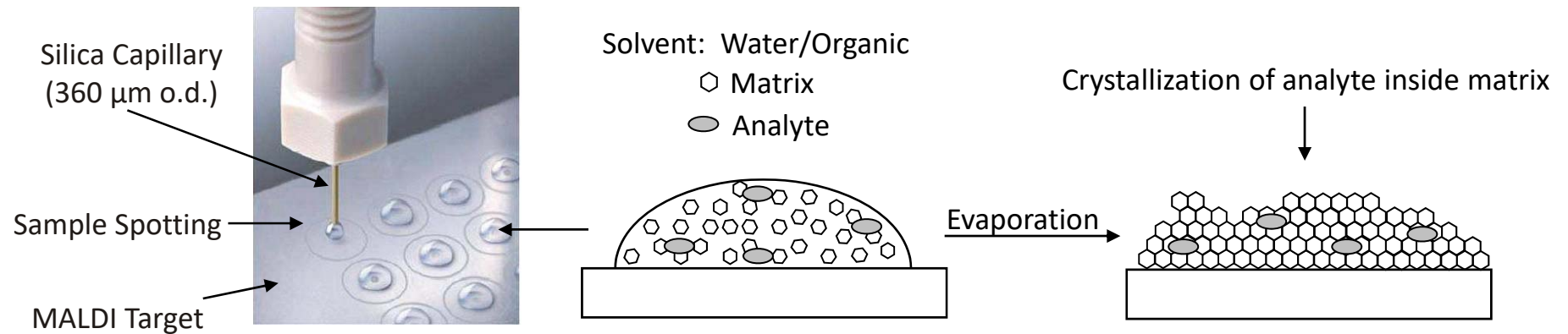**(CHCA)**

Nicotinic acid (NA)

3-hydroxypicolinic acid (HPA)

3,4-dihydroxycinnamic acid
Caffeic acid (CA)

MALDI Sample Preparation

Picture:
Automated MALDI Spotter

Silica Capillary
(360 μm o.d.)

Sample Spotting

MALDI Target

Solvent: Water/Organic

⬡ Matrix

⬭ Analyte
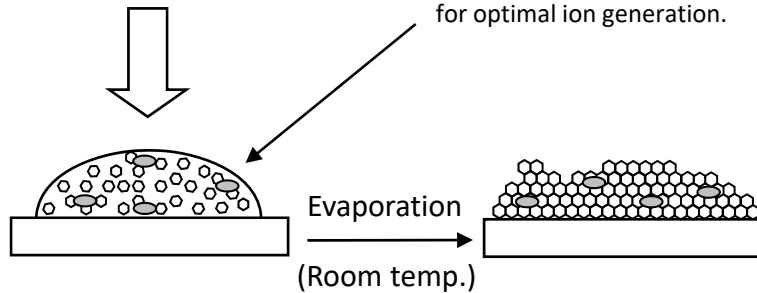
Evaporation

Crystallization of analyte inside matrix

Dried-Droplet Method (Most Common)

Matrix and analyted are pre-mixed typically
in an acidic aqueous/organic mixture (pH < 4)

Analyte:Matrix
(1:1000 to 1:10,000)

Contaminants (salts, surfactants, etc.)
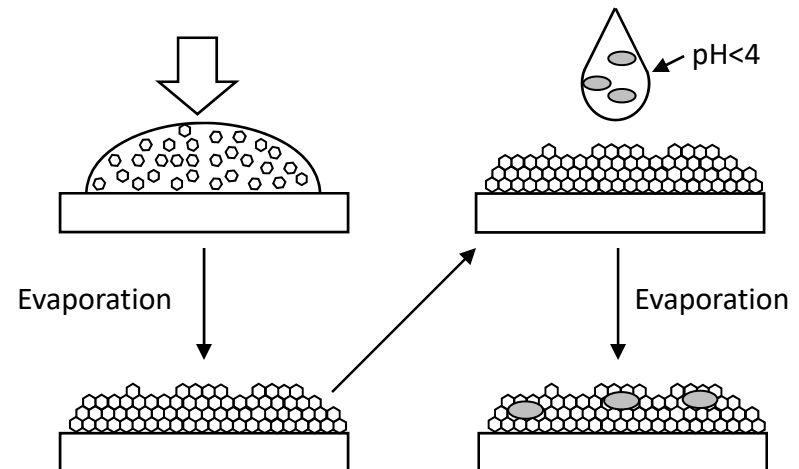should be typically kept below 50 mM
for optimal ion generation.

Evaporation

(Room temp.)

Thin-layer/Fast Evaporation Method

~10 mM matrix in volatile
organic solv. (eg., acetone)

~1μM analyte in aqueous
soln. (organic<30%)

pH<4

Evaporation

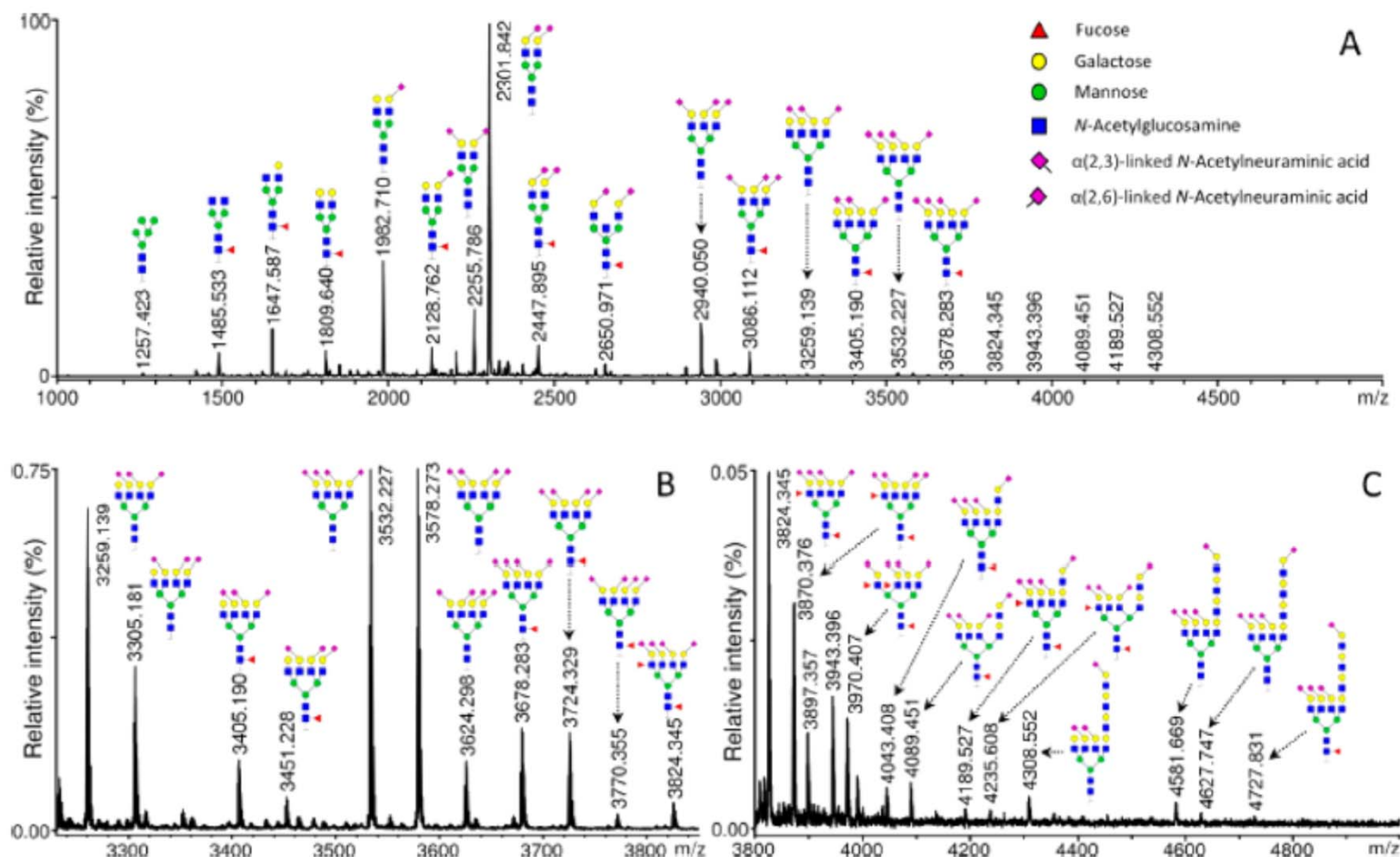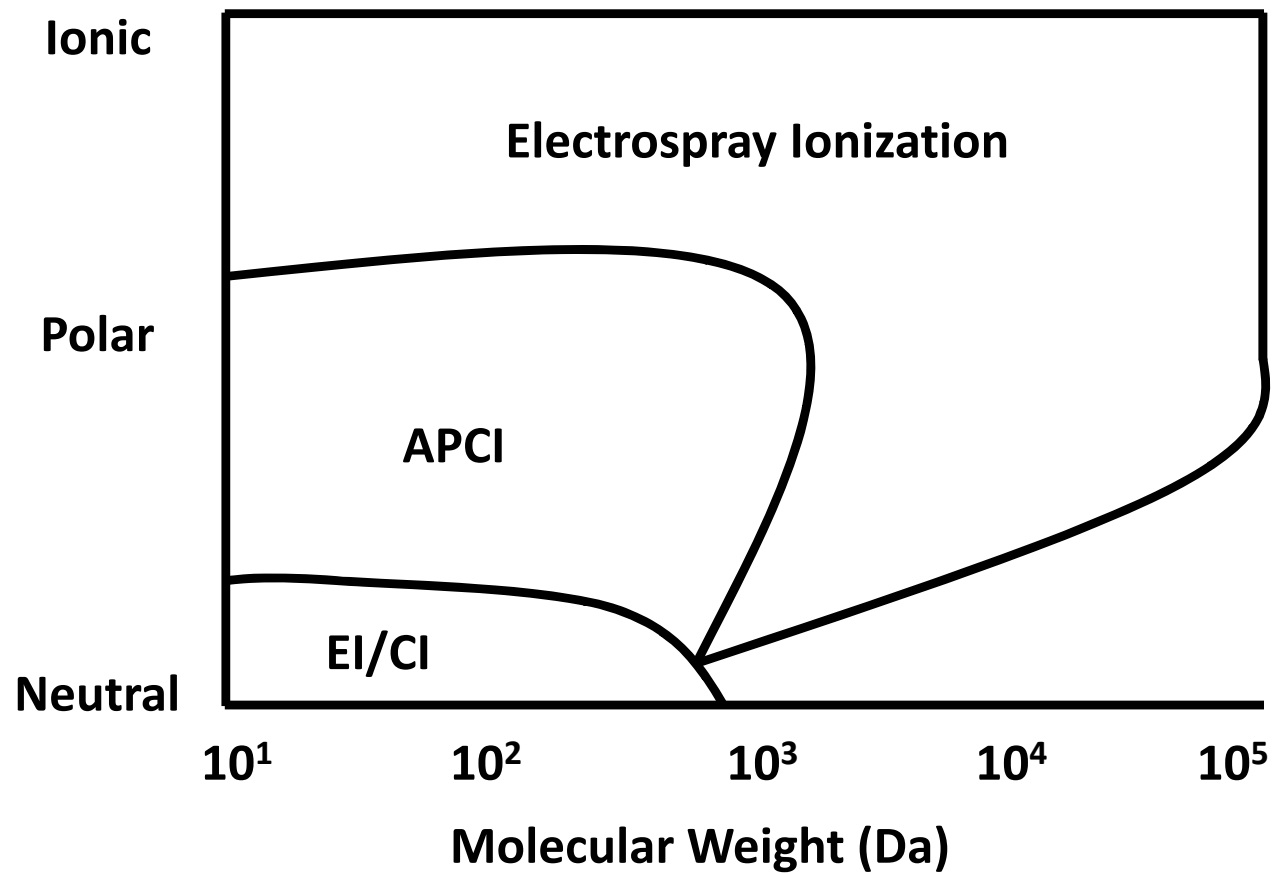Evaporation

MALDI-TOF MS of N-glycans

Figure 12. MALDI-TOF MS spectrum of human plasma N-glycans after linkage-specific sialic acid ethyl esterification. Reprinted with permission from ref 109. Copyright 2014 American Chemical Society.

(109) Reiding, K. R.; Blank, D.; Kuijper, D. M.; Deelder, A. M.; Wuhrer, M. High-throughput profiling of protein N-glycosylation by MALDI-TOF-MS employing linkage-specific sialic acid esterification. Anal. Chem. 2014, 86, 5784−5793.
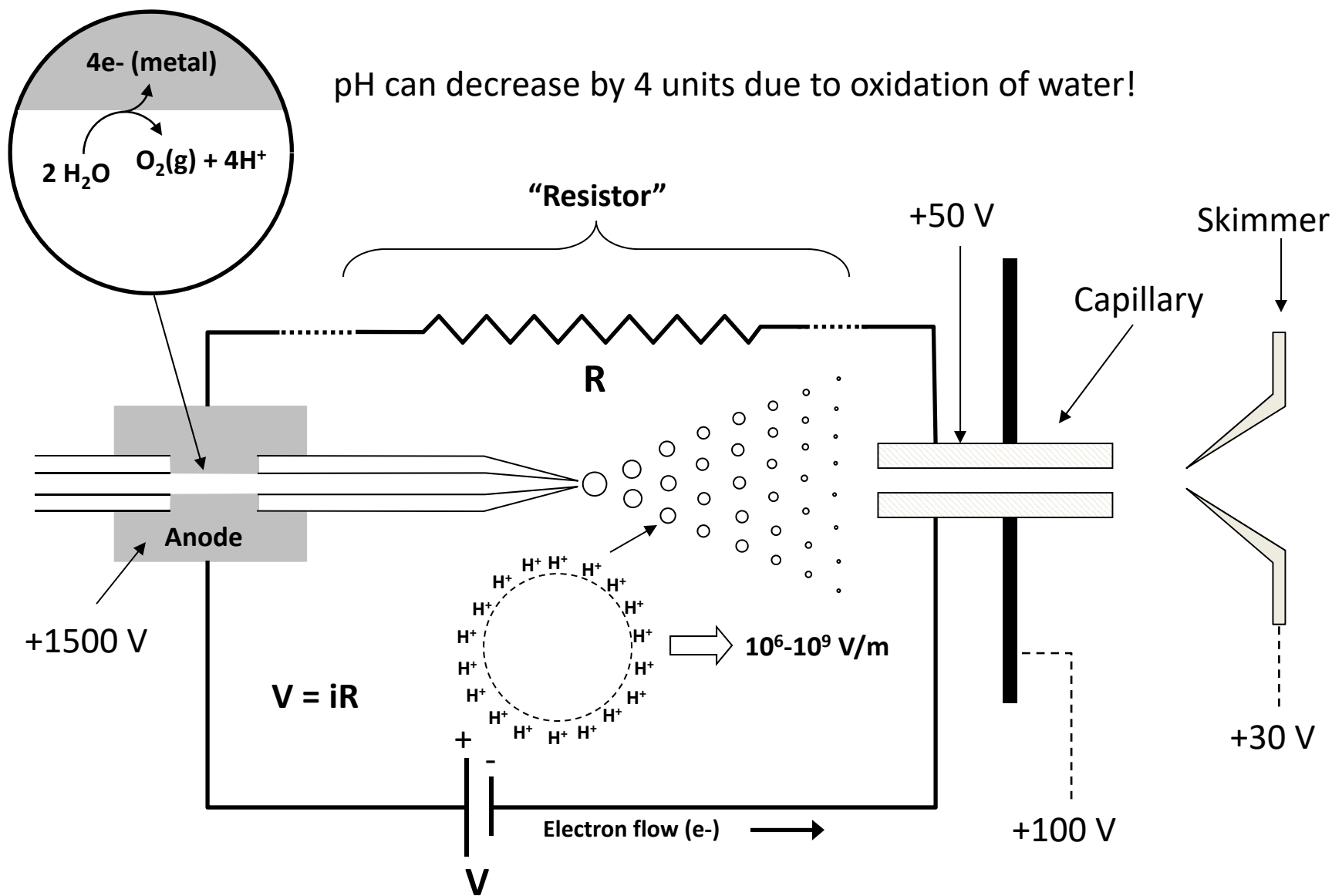
Advantages of Electrospray Ionization

- Allows for the direct coupling of liquid separations to mass spectrometer

- Multiple-charging extends the mass range of an analyzer by a factor equal to z.

- Soft(est) ionization technique which allows for the analysis of non-covalent complexes

- Ions produced are conducive to structural analysis

- Practical mass range up to 100 kDa but MegaDalton species can be detected….intact viruses have been electrosprayed!

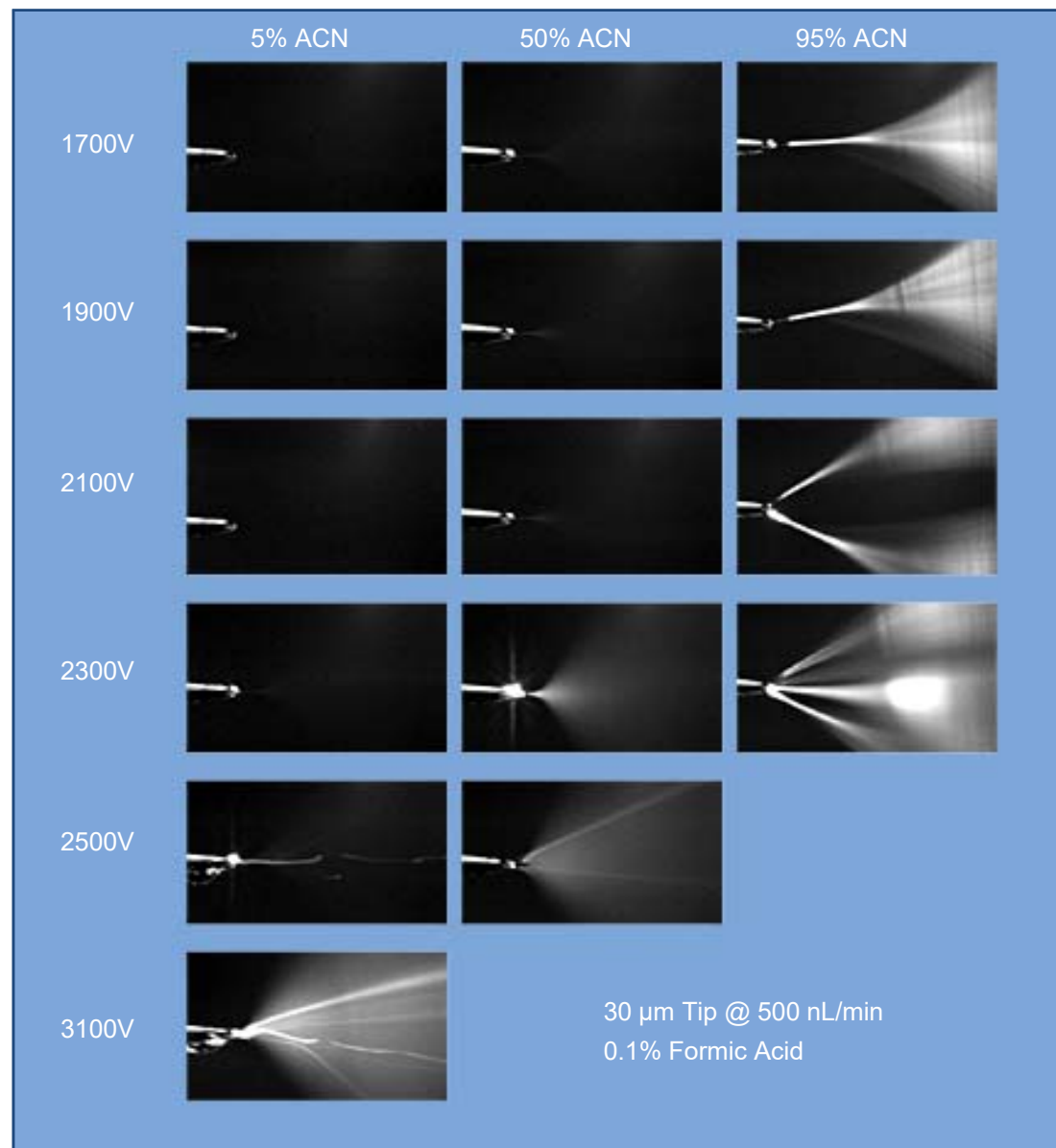- Good detection limits – femtomole to attomole routinely achieved

Ionization of Different Classes of Molecules

Electrospray: An Electrolytic Cell

4e- (metal)

2 H₂O → **O₂(g) + 4H⁺**

pH can decrease by 4 units due to oxidation of water!

"Resistor"

+50 V

Skimmer

R

Capillary

Anode

+1500 V

10⁶-10⁹ V/m

V = iR

+30 V

+100 V

Electron flow (e-) →

V

H⁺ (multiple, surrounding droplet)

Spray Morphology for Different Voltages and Solution Compositions

5% ACN   50% ACN   95% ACN

1700V
1900V
2100V
2300V
2500V
3100V

30 µm Tip @ 500 nL/min
0.1% Formic Acid

Electrospray Ionization (Ion Evaporation Model)

ESI Emitter Tip
10-30 µm i.d.

Taylor Cone

< 5 µm

Distance → MS Inlet
1-5 mm

Progeny Droplets

200-500 nL/min

Repeats for subsequent generations of charged droplets

< 5 µm

Solvent
Evaporation

80-100%
Rayleigh Limit

Fission → Progeny
Droplets

Ion Evaporation

Eyers and Gaskell (Eds) Quantitative Proteomics; A. M. Hawkridge *"Practical Considerations and Current Limitations in Quantitative Mass Spectrometry-Based Proteomics, Ch. 1"* **2014.**

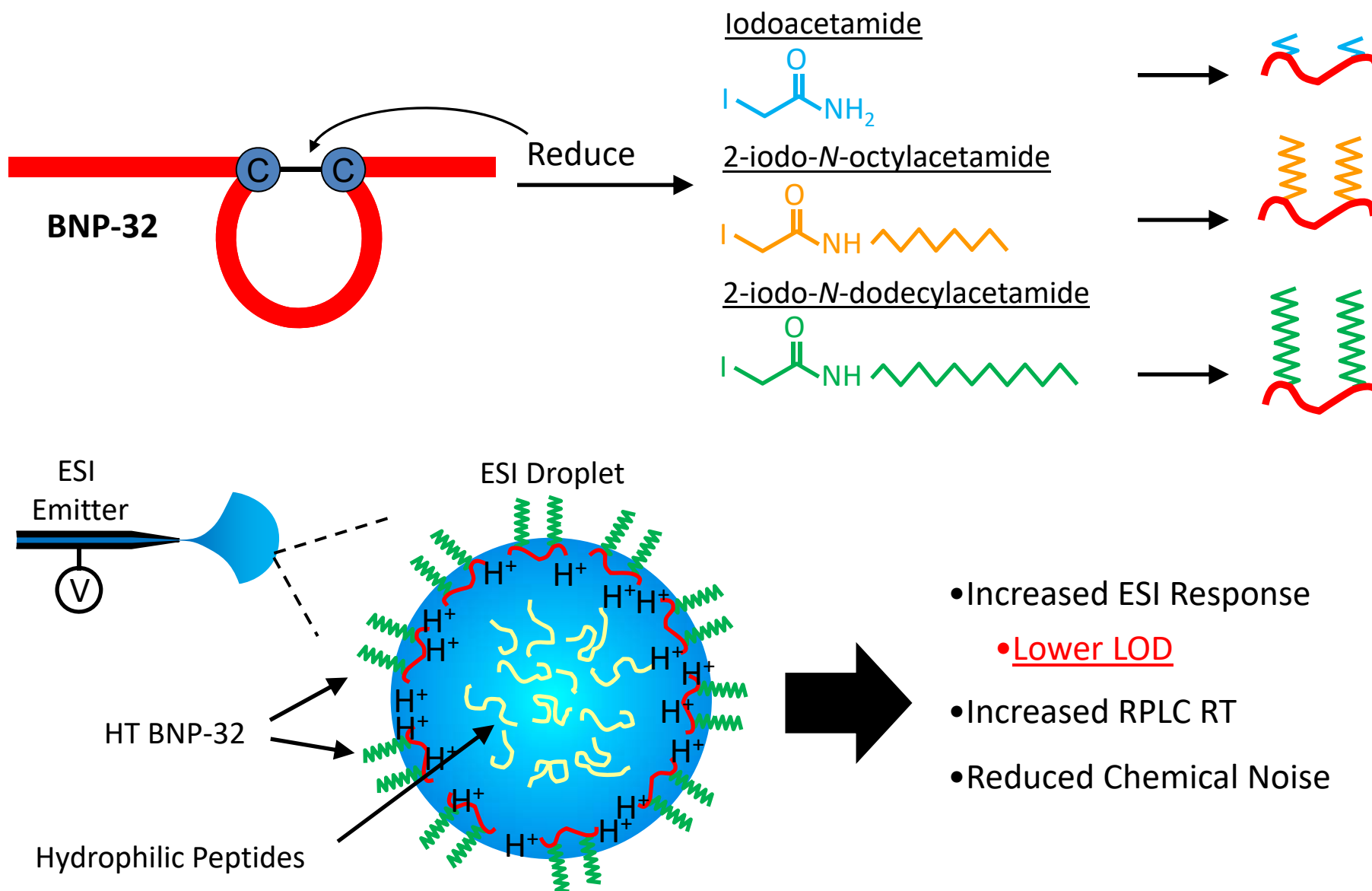# Response Observed in ESI and The Role of Hydrophobicity
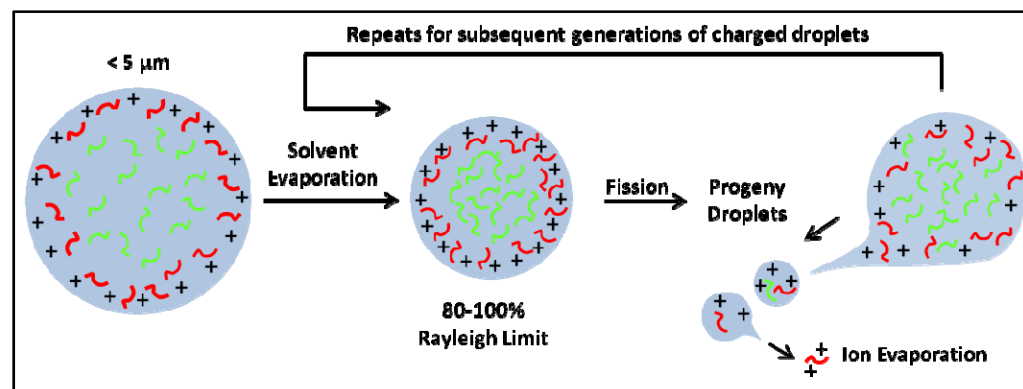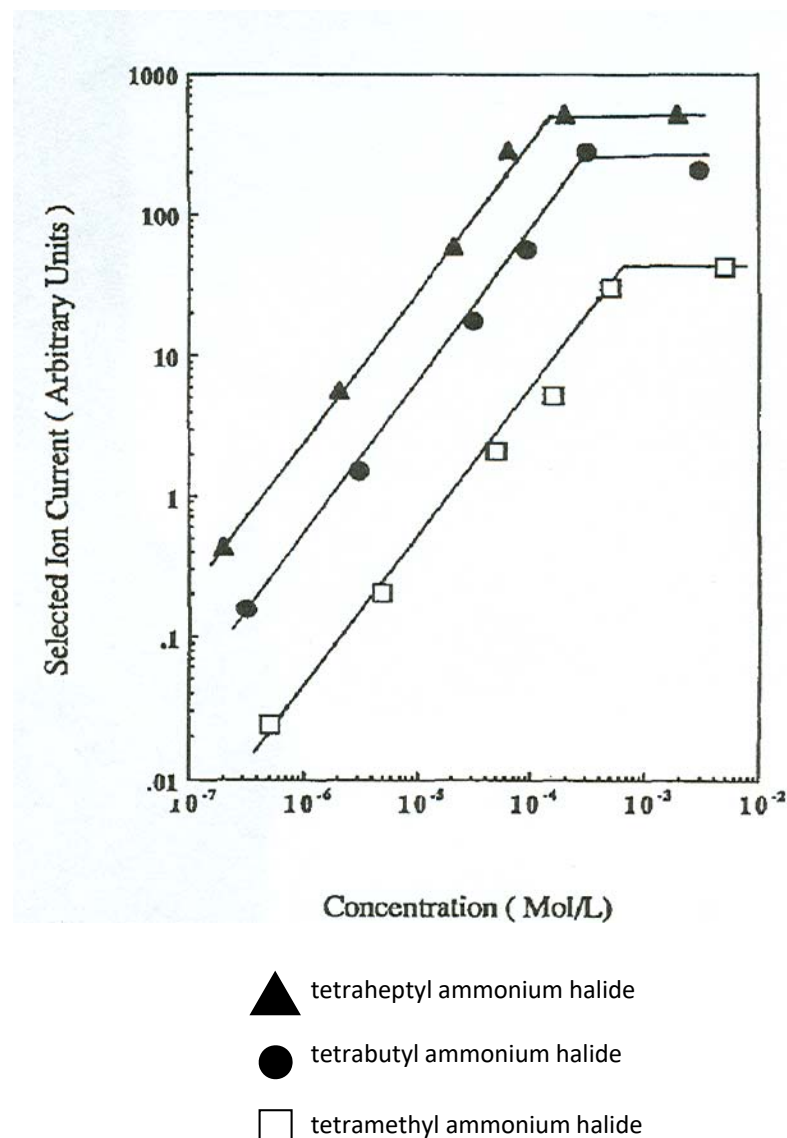
$$N_{iz} = 3Ae^{(-\Delta G_{iz}^{\circ}/RT)}(N_i/r)e^{(\Delta rzQ/4\pi\varepsilon_0 RTr^2)}$$

- N_{iz} = ESI response (ion flux) of an ion (i) with z charges
- A = proportionality constant that relates bulk concentration to surface activity for a given analyte
- ΔG°_{iz} = free energy of solvation
- R = gas constant
- → T = temperature
- N_i = moles of analyte
- → r = radius of the droplet
- → Δr = the distance the ions must travel to become desolvated
- Q = excess charge
- ε_0 = the gas permittivity constant

▲ tetraheptyl ammonium halide

● tetrabutyl ammonium halide

□ tetramethyl ammonium halide

J.B. Fenn, *J. Am. Soc. Mass Spectrom.*, **1993**, 4, 524-535.

Hydrophobic Tagging of BNP-32

BNP-32

Reduce

Iodoacetamide

2-iodo-*N*-octylacetamide

2-iodo-*N*-dodecylacetamide

ESI Emitter

ESI Droplet

HT BNP-32

Hydrophilic Peptides

- Increased ESI Response
- Lower LOD
- Increased RPLC RT
- Reduced Chemical Noise
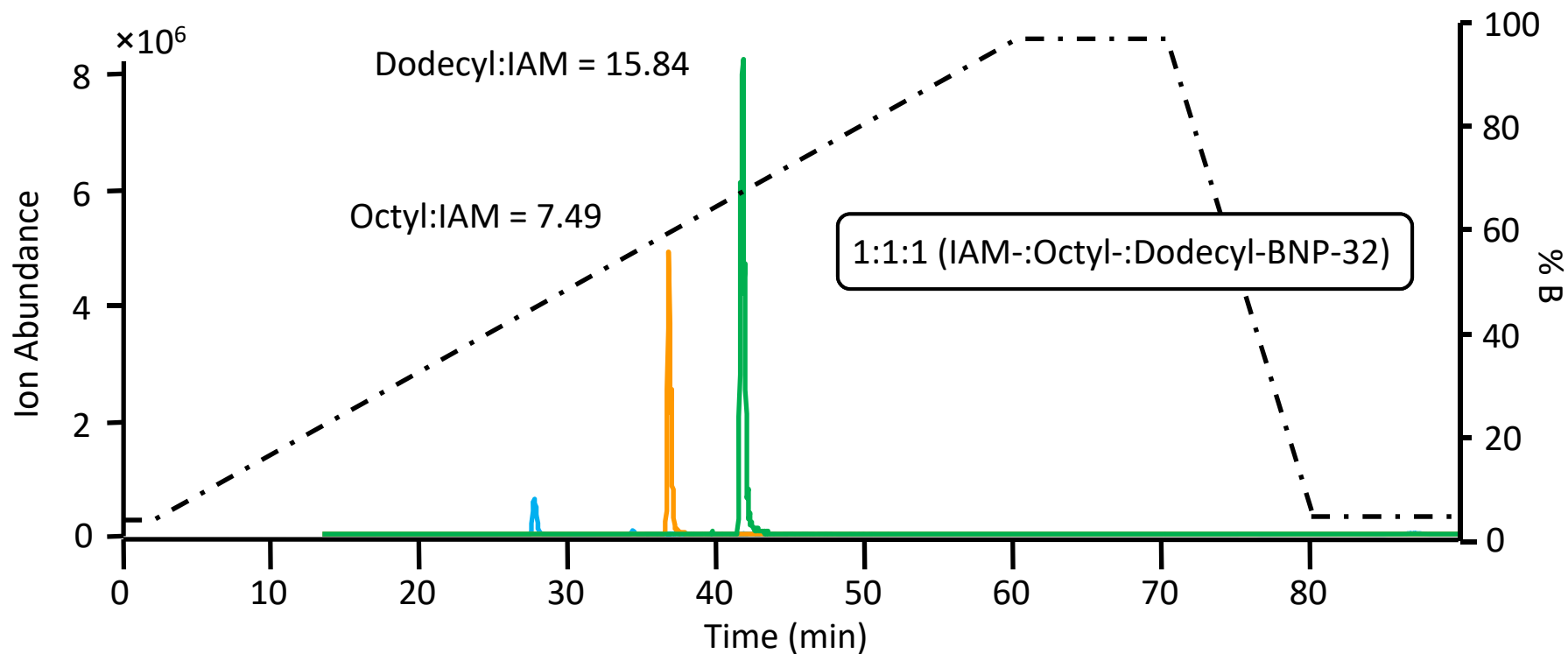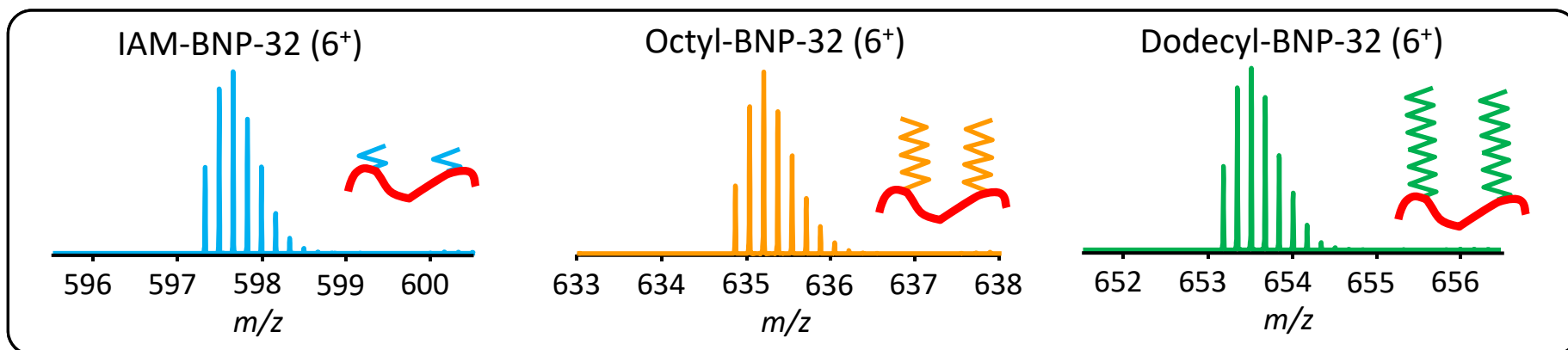
nLC-MS/MS ESI Response of Tagged E-76 Peptide

butyl-benzyl-idoacetamide

[M+2H]$^{2+}$
m/z_{exp}=1296.1152
+0.89 ppm

MS/MS

1295.5   1299.5
m/z

y_{13}^{2+}

b_{17}^{2+}

b_9

y_{14}^{2+}
b_{16}^{2+}
b_7

y_9
y_{11}
y_{10}

400   800   1200   1600   2000
m/z

Tag Does Not Fragment

[M+2H]$^{2+}$
m/z_{exp}=1164.0194
-0.67 ppm
*Abundance x100

Iodoacetamide

1163.5   1167.5
m/z

×100

E-76 Peptide
Coagulation VIIIa Inhibitor

b_7 b_9 b_{16} b_{17}

Ac-ALCDDPRVDRWYCQFVEG-NH$_2$

y_{13} y_{11} y_9
y_{14} y_{10}

0   15   30   45   60   75   90
Time (min)
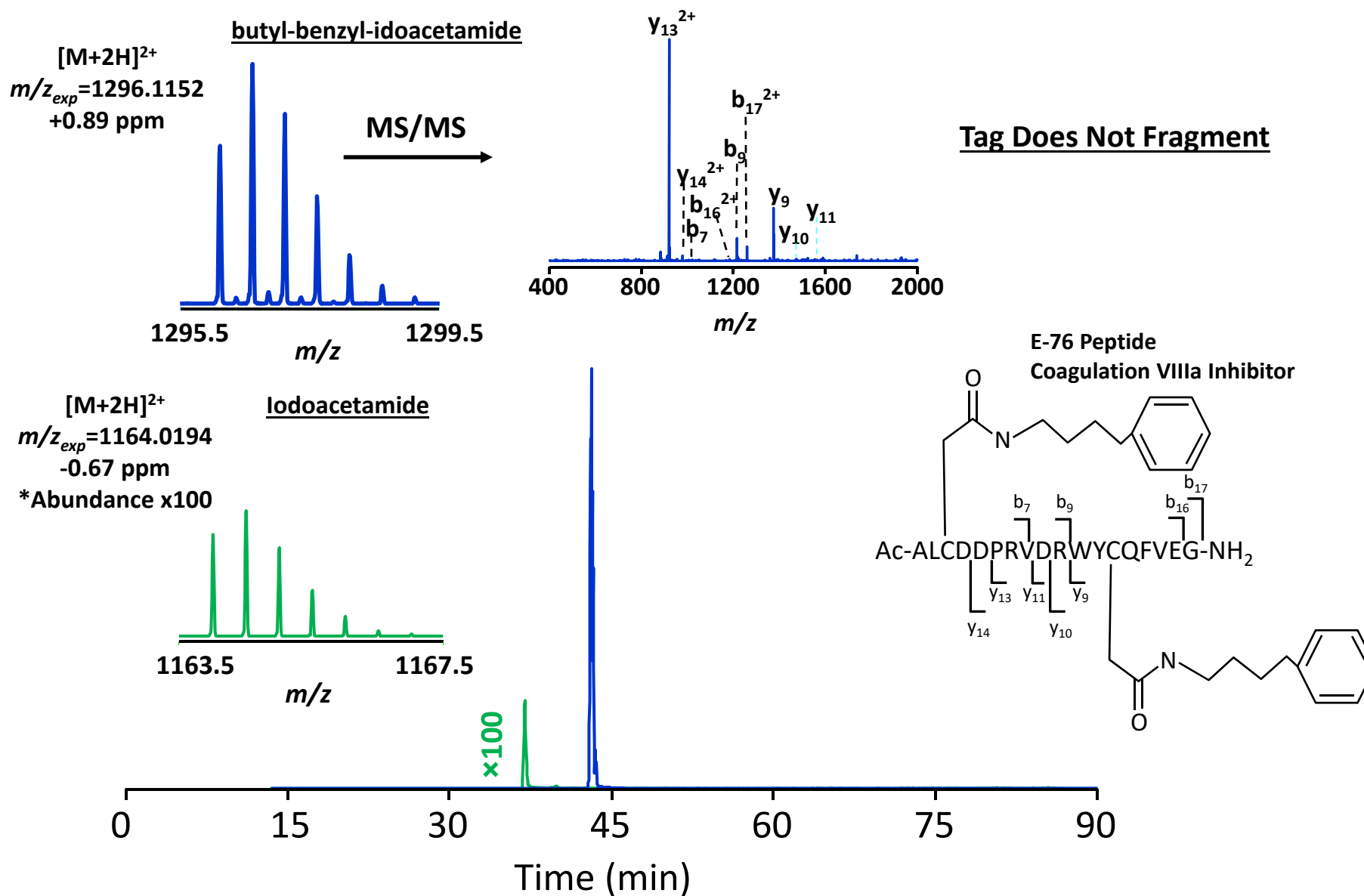
Tailoring non-polar surface area of reagents to influence ESI response
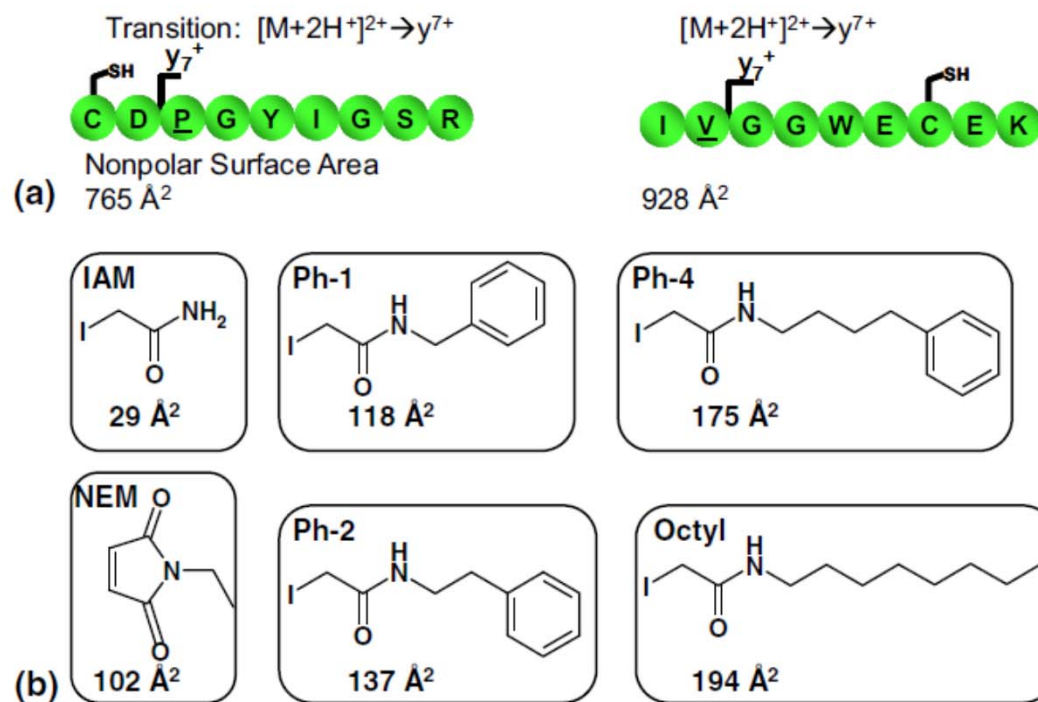

Figure 1. (a) Illustrates the amino acid sequences of the two peptides used in this study, the transition used by the mass spectrometer, and the nonpolar surface area for the given peptide. The underlined amino acid shows the stable isotope labeled form for in the internal standard. (b) The alkylating reagents utilized in this study are shown along with their abbreviation used in this manuscript and their nonpolar surface area.
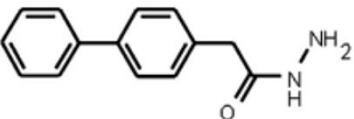
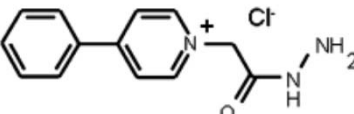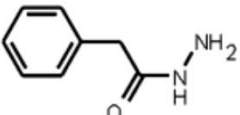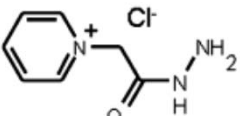# ESI response is affected by the analyte and tag
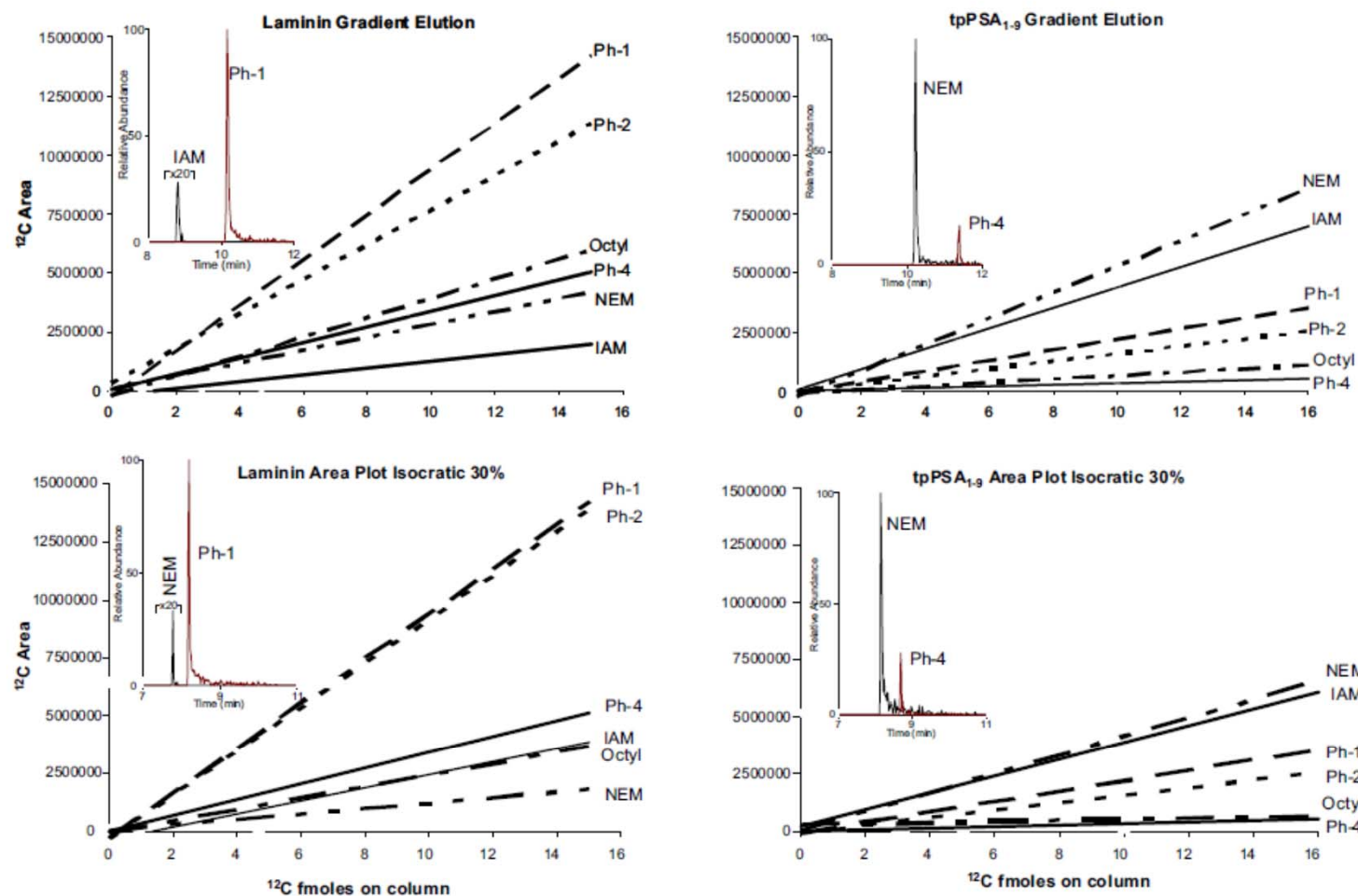
Figure 3. Sensitivity is shown to differ for both peptides depending on which alkylating reagent is utilized. Abbreviations denote which alkylating reagent was used. Inset is the extracted ion chromatogram for the best and worst performing peptide and alkylating reagents for the denoted experimental set.

Chemical tagging of N-glycans commonly used for LC-MS/MS

Table 1. Reagents, Abbreviations, And Non-Polar Surface Areas of the Tags Used in This Experiment

Abbreviation	Structure	NPSA[a]
Phenyl-GPN		181
Phenyl-GP		180
GPN		109
GP		108

[a] Non-Polar Surface Area.

Table 4. Retention Times and Relative Abundance Data for the NA2 Equimolar Mixture

molecule	retention time (min)	fold increase[a]
phenyl-GPN+NA2[b]	22.8	18
GPN + NA2[b]	25.5	11
phenyl-GP+NA2[c]	29.2	7
free NA2	29.1	1
GP + NA2[c]	32.23	<1

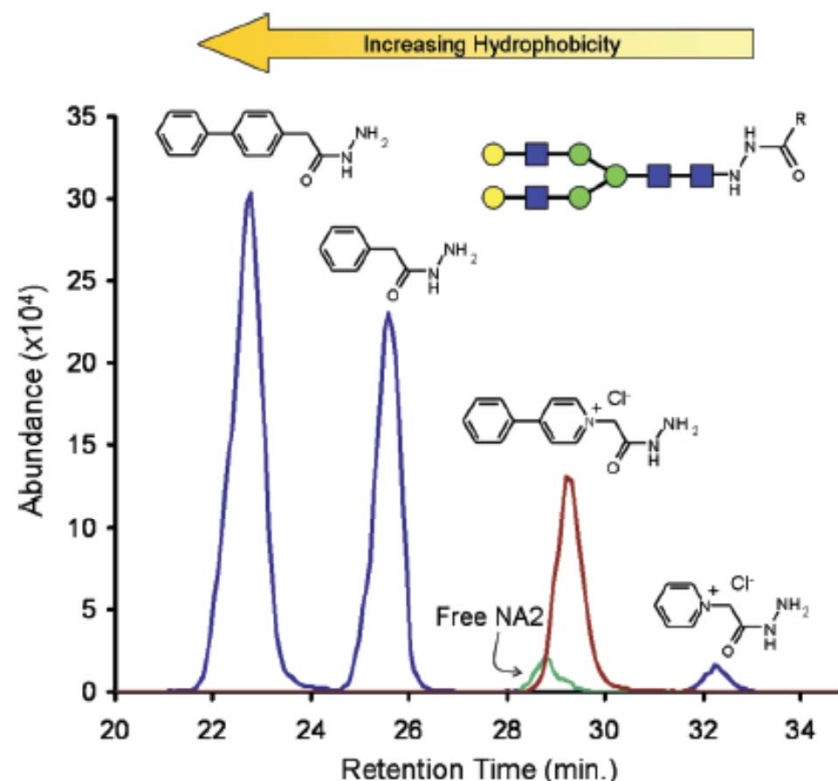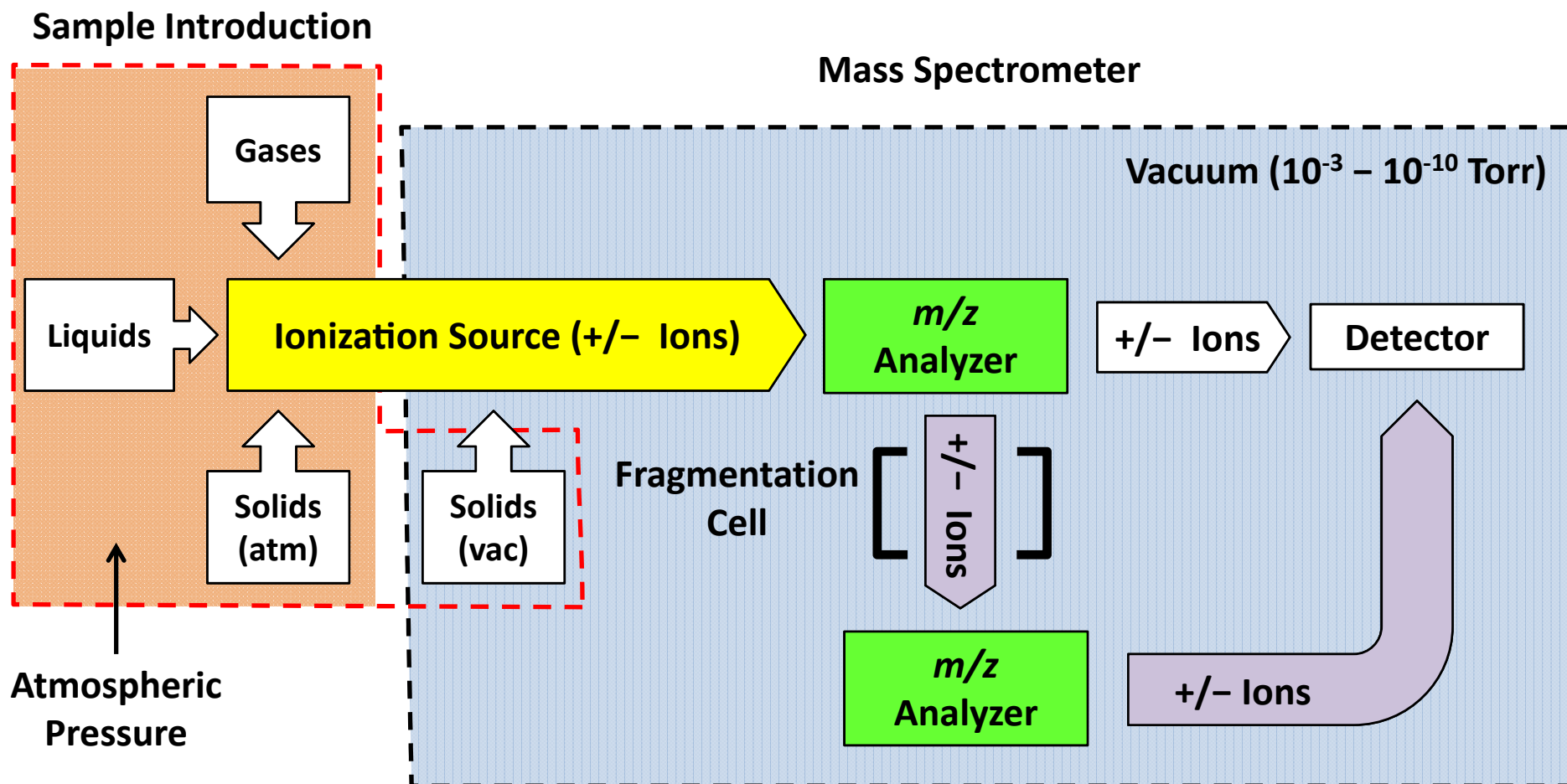[a] Relative to free glycan. [b] Neutral reagent. [c] Charged reagent.

Figure 5. The extracted ion chromatogram of the equimolar mixture made from the NA2 glycan with each tag. The phenyl-GP glycan and the free glycan EIC's are overlaid to show overlapping retention, and the phenyl-GP glycan out-competes the free glycan for excess charge in the electrospray droplet.

Key Points

- **MALDI and ESI are essential for the MS analysis of carbohydrates. The following summary highlights some general strengths and weaknesses for each ionization technique with respect to carbohydrates**
- **MALDI**
  - High-throughput, pulsed, and amenable to off-line separations
  - Less susceptible to salt contaminants compared to ESI
  - Carbohydrate ionization and sample preservation is highly dependent on choice of matrix. Labile glycan groups (e.g., sialic acid, sulfates) can be lost if matrix and laser conditions are not optimized.
  - Quantitation is not routine
- **ESI**
  - Continuous and therefore amenable to on-line separations (U/HPLC)
  - Lower throughput but can be made high-throughput with multiple HPLC systems.
  - Gentler/softer than MALDI but more susceptible to salts
  - Typically used for quantification

What are the Primary Components of a Mass Spectrometer?

Sample Introduction

Mass Spectrometer

Gases

Vacuum ($10^{-3} - 10^{-10}$ Torr)

Liquids

Ionization Source (+/− Ions)

m/z Analyzer

+/− Ions

Detector

Solids (atm)

Solids (vac)

Fragmentation Cell

[+/− Ions]

Atmospheric Pressure

m/z Analyzer

+/− Ions

Mass analyzers discussed in this lecture

- **Quadrupole**
  - The most common mass analyzer in use today used primarily for m/z isolation

- **Two dimensional (2D) linear quadrupole ion trap**
  - Used for ion isolation/trapping and full-scan applications

- **Time-of-flight**
  - Most common high-resolving power mass analyzer (RP = 10k-70k)
  - Used primarily for accurate mass measurements (3-10 ppm)

- **Orbitrap**
  - Highest resolving power mass analyzer (RP = 70k-1M) on the market (except for less common FT-ICR)
  - Used primarily for accurate mass measurements (≤ 3ppm)

- **Importantly, all of these analyzers are used in combination with each other to form what are referred to as *'tandem mass spectrometers'* (MS/MS).**
  - triple quadrupoles (QqQ)
  - quadrupole + linear ion trap (Q-trap)
  - linear ion trap + Orbitrap (LTQ-Orbitrap)
  - quadrupole + time-of-flight (Q-TOF)
  - quadrupole + Orbitrap (Q Exactive)
  - time-of-flight^2 (TOF-TOF)
  - quadrupole + linear ion trap + Orbitrap (Fusion Lumos).

Milestones in *m/z* analyzer development

1919 First mass spectrometer (MS)

1942 First commercial magnetic sector instrument

1953 Quadrupole and ion trap (3D) developed

1955 Time-of-Flight Instrument (TOF) Developed

1958 Time-of-flight (TOF) commercialized

1968 Quadrupole mass spectrometer commercialized

1974 Fourier transform ion cyclotron resonance (FT-ICR) developed

1978 Triple quadrupole MS (QqQ) developed

1982 Triple quadrupole commercialized

1995 Quadrupole time-of-flight instrument (Q-TOF) commercialized

2000 Orbitrap is developed

2002 Linear ion trap (2D) developed

2006 Orbitrap is commercialized

2010 Quadrupole Orbitrap is commercialized

Time-of-flight (TOF)

Advantages
- Simple and robust
- High transmission
- Detects ion simultaneously (vs. scanning analyzers such as quads and traps)
- High mass measurement accuracy (3-10 ppm)
- Excellent resolving power
  - Linear: 2000-10,000
  - Reflectron: 10,000-70,000
- High duty cycle (>50 Hz)
- Directly compatible with continuous (ESI) and pulsed (MALDI) ionization sources
- Low limits of detection/quantification (femtomole)

Disadvantages
- Modest Linear dynamic range (10^3 -10^4)
- Requires ultrahigh vacuum (UHV) conditions (10^{-9}-10^{-12} Torr)
- Cost can exceed $500k

Common configuration(s)
- Identification and Quantification
  - TOF
  - Q-TOF
  - TOF-TOF

Time-of-flight

- Ions are start from the same point (e.g., ESI source or MALDI plate) with a small amount of kinetic energy
  - The ion packet has an initial kinetic energy distribution (e.g., 10-50 eV)
- Ions are then accelerated to a larger kinetic energy (>1 keV) which is dependent on their mass (m) and charge (z):

$$zeV = \frac{1}{2}mv^2 \longrightarrow v = \sqrt{\frac{2zeV}{m}}$$

z = charge
eV = acceleration energy
m = mass
v = velocity

- At this point, the kinetic energy (zeV) for all ions is 'equivalent' but the velocities are different....this creates the time-of-flight (tof) effect that is measured:

$$tof = \frac{L}{v} = L\sqrt{\frac{m}{2zeV}}$$

L = distance between the accelerator plates and detector

Time-of-flight

$$tof = \frac{L}{v} = L\sqrt{\frac{m}{2zeV}}$$

v = velocity
z = charge
eV = acceleration energy
m = mass
L = distance

lower v: larger m/z
larger v: lower m/z
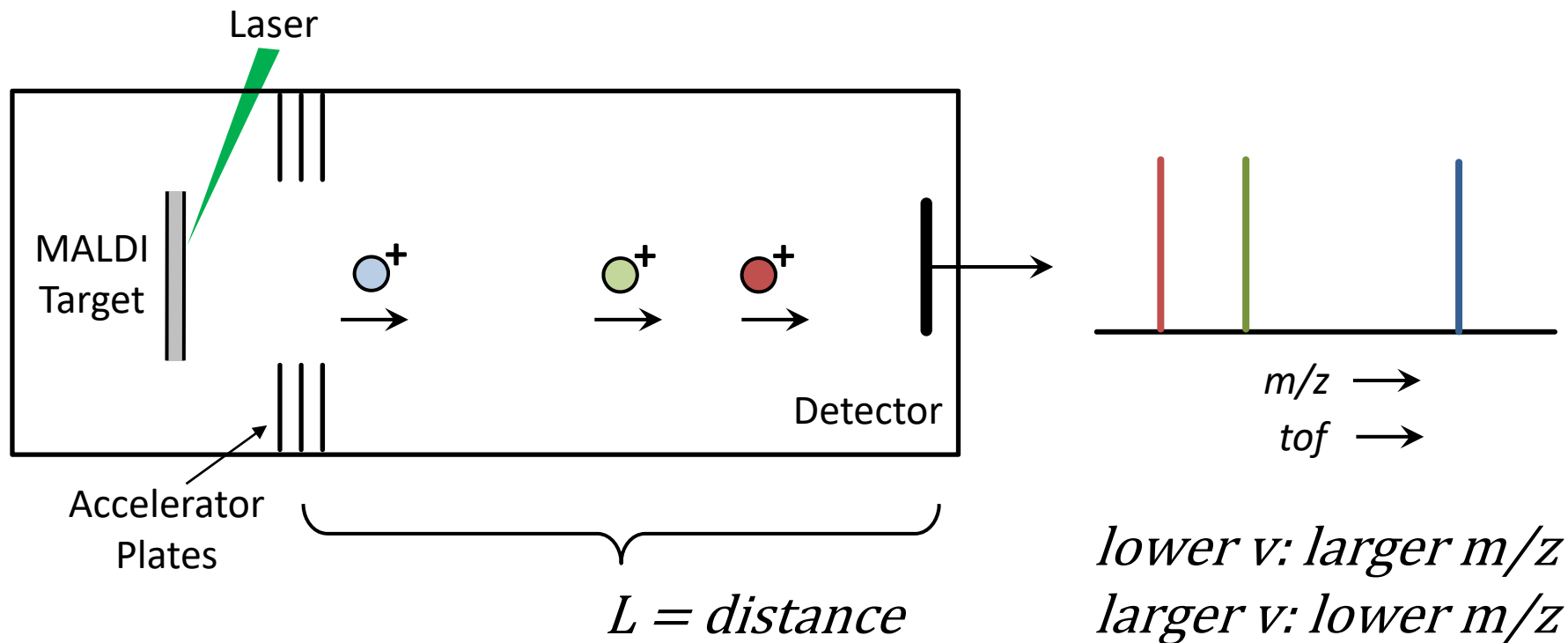
Effect of TOF on analysis time and RP as a function of *m/z*

$$tof = \frac{L}{v} = L\sqrt{\frac{m}{2zeV}}$$
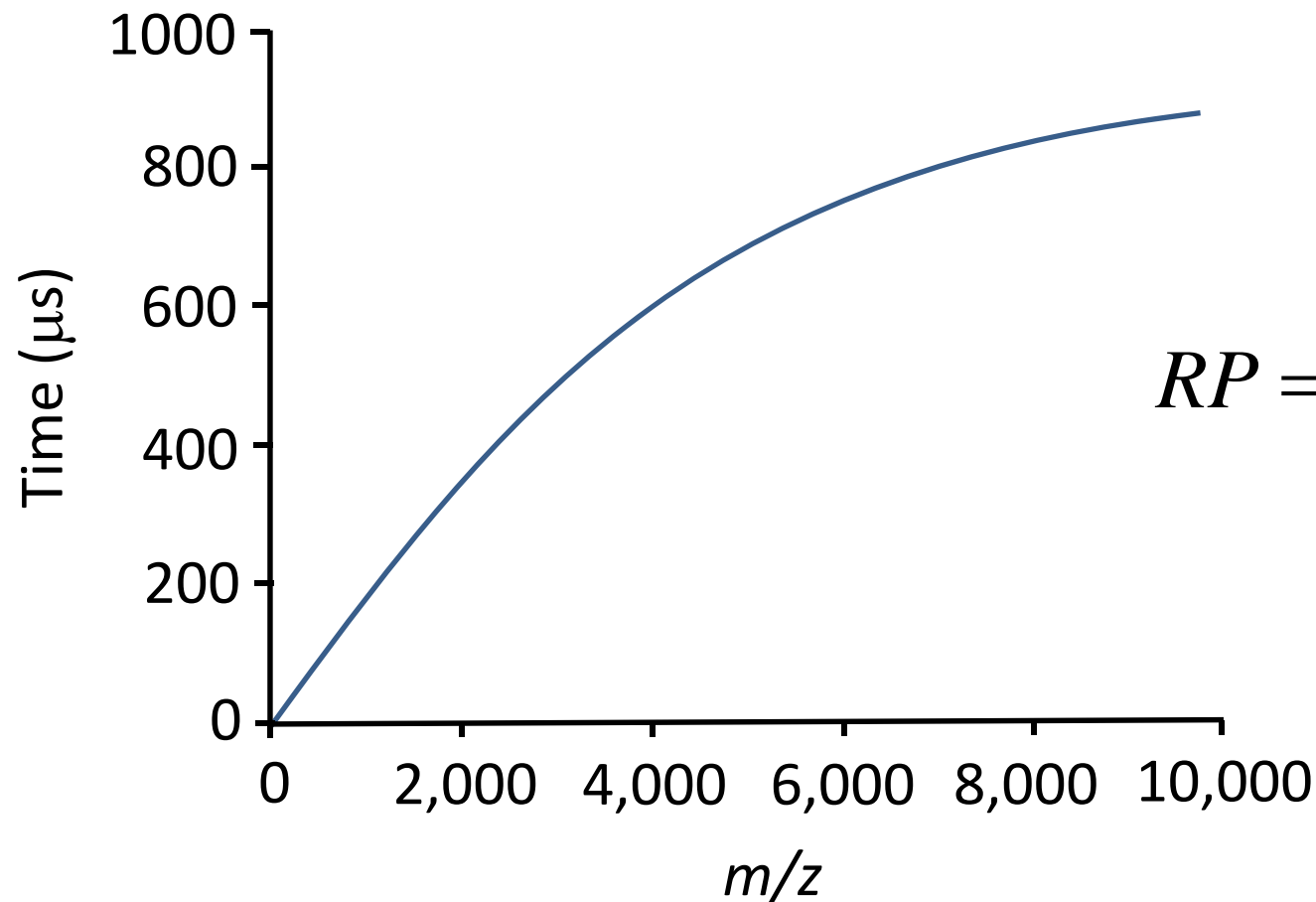
$$RP = \frac{m}{\Delta m} = \frac{t}{2\Delta t}$$

Effect of Reflectrons in Time-of-flight analyzers

$E_1 > E_2$; *m/z* values are equal (single isotope)

RP = 10,000

Sample

E_1

E_2

L_1

R

Linear Detector

Reflecting Detector

L_2

11980.0

11975 11980 11985 11990
m/z

L_1 - First Field-Free Drift Region
L_2 - Second Field-Free Drift Region
R - Reflectron

$$\Delta x = L_1 + L_2 + R$$

11980.0354

RP = 30,000

11975   11980   11985   11990
m/z
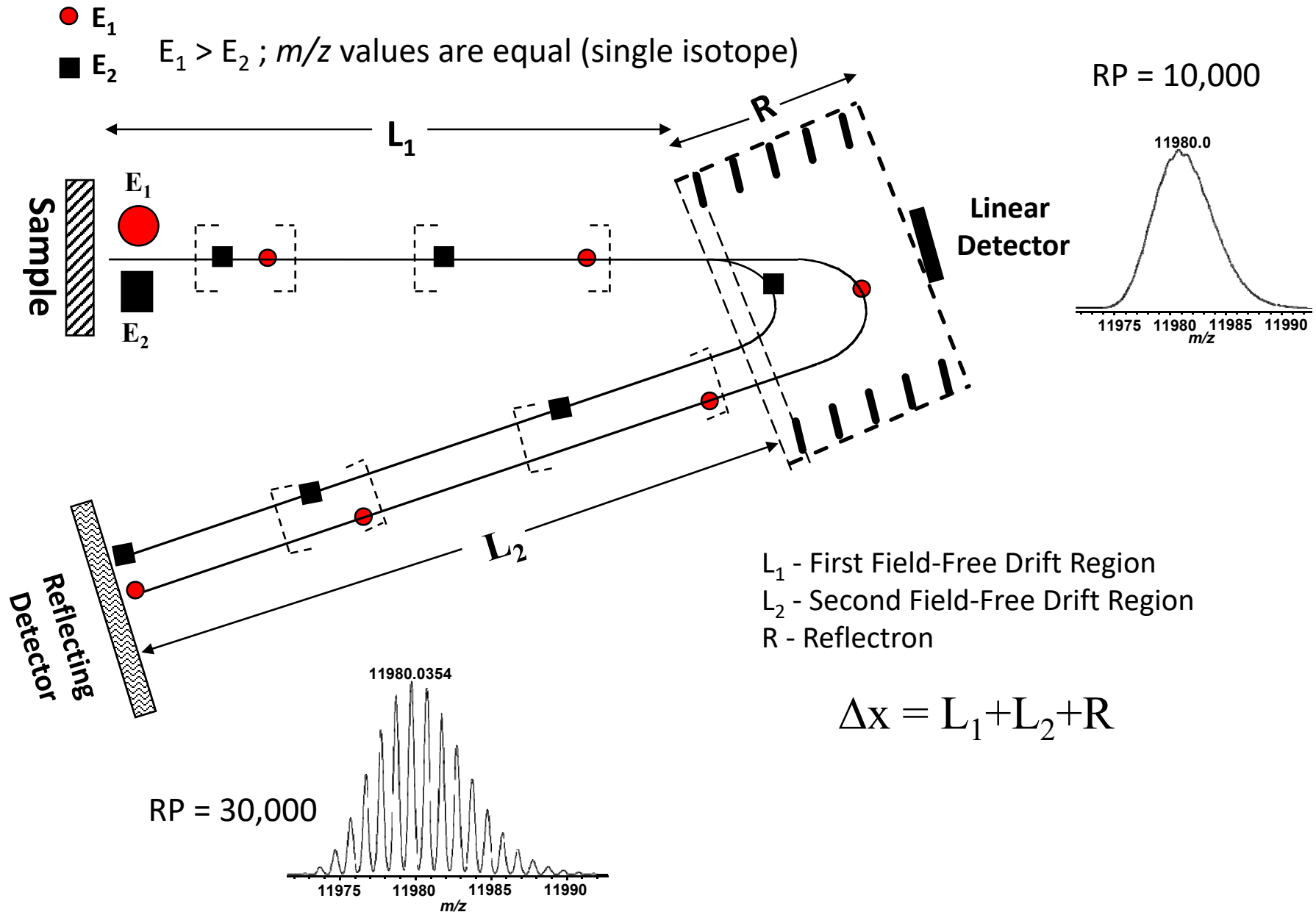
Orbitraps

Advantages
- High mass measurement accuracy (<3 ppm)
- Highest resolving power available
  - Low field: 140,000
  - High field: ≤ 500,000
  - Ultra-high field: 1,000,000
- High duty cycle (10-30 Hz)
- Directly compatible with continuous (ESI) and pulsed (MALDI) ionization sources but ESI dominates
- Low limits of detection/quantification (attomole-femtomole)
- Available in multiple tandem configurations

Disadvantages
- Linear dynamic range (10^3 -10^4)
- Requires ultrahigh vacuum (UHV) conditions (10^{-12} Torr)
- Cost for upper level system can exceed $1M

Common configuration(s)
- Identification and Quantification
  - Orbitrap
  - Q-Orbitrap (Q Exactive)
  - LTQ-Orbitrap
  - Q/LTQ-Orbitrap (Fusion/Fusion Lumos)

Orbitrap Principles: Ion injection (C-trap) and trapping (Orbitrap)

C-trap

CID ← ← Quadrupole/LTQ

+ **Focused ion packet**

r

φ

z

-15 kV

time
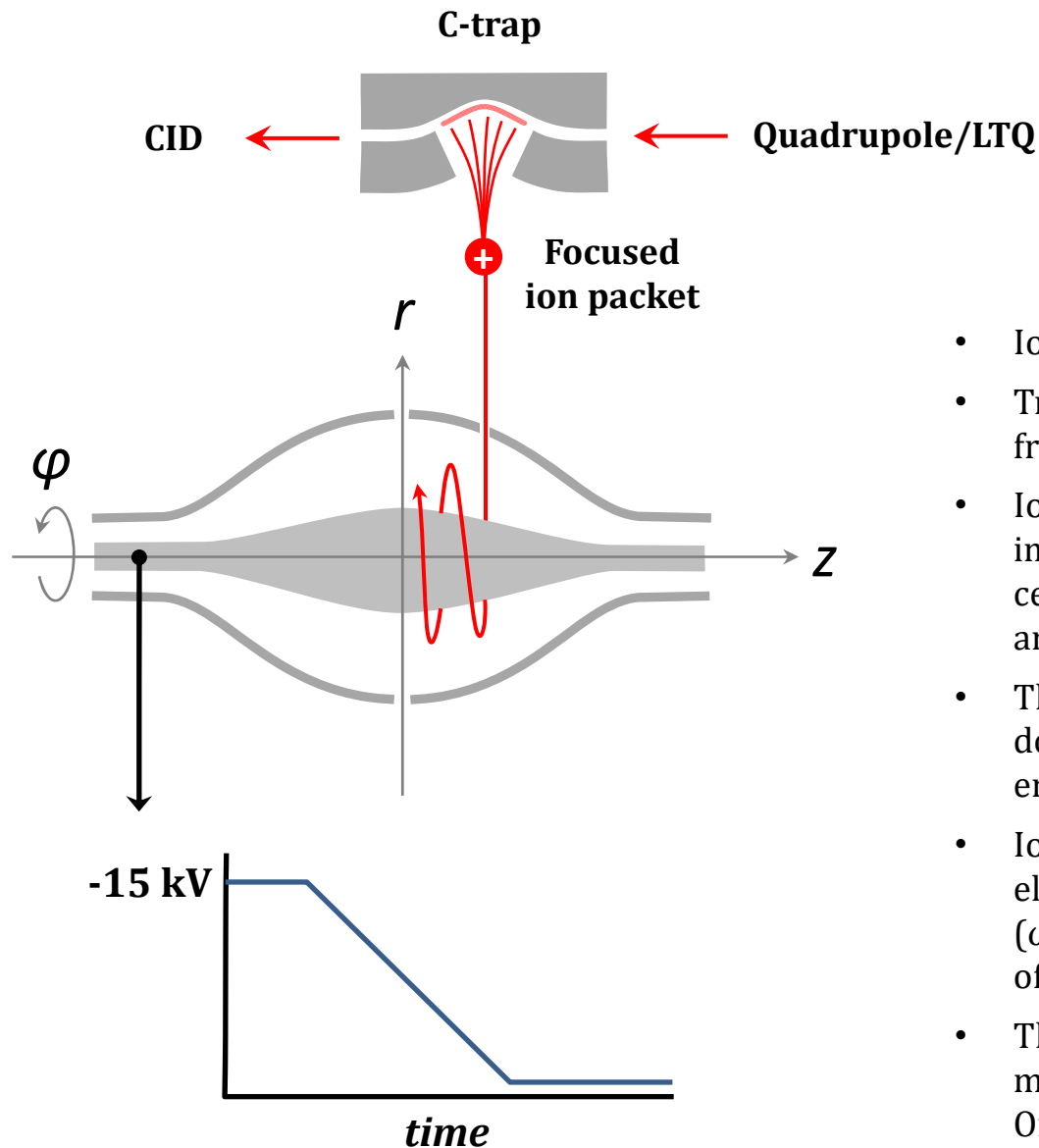
- Ions trapped in the C-trap

- Trapped ions are ejected as a focused packet from the C-trap toward the Orbitrap entrance.

- Ion packet enters the Orbitrap where they are immediately redirected via the high voltage center electrode into a rotational motion around the central electrode (ϕ).

- The voltage on the central electrode is ramped down as the ion packets are compressed and enter into an 'Orbitrap' motion.

- Ion packets oscillate along the central electrode in the z-direction with a frequency (ω) inversely proportional to the square root of the *m/z*.

- The frequencies of each ion packet is measured as an image current on the outer Orbitrap electrodes.

Orbitrap Principles: Frequency → *m/z*

Ion motion frequency components in an Orbitrap:
- Rotation (ω_φ)
- Radial (ω_r)
- Axial (ω_z) ← **Axial frequency is used to determine *m/z***
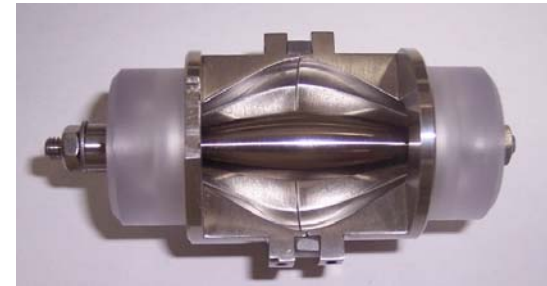
Relationship between axial frequency and *m/z*

$$\omega_z = \sqrt{\dfrac{k}{m/z}}$$

Signal Amplifier

Image Current
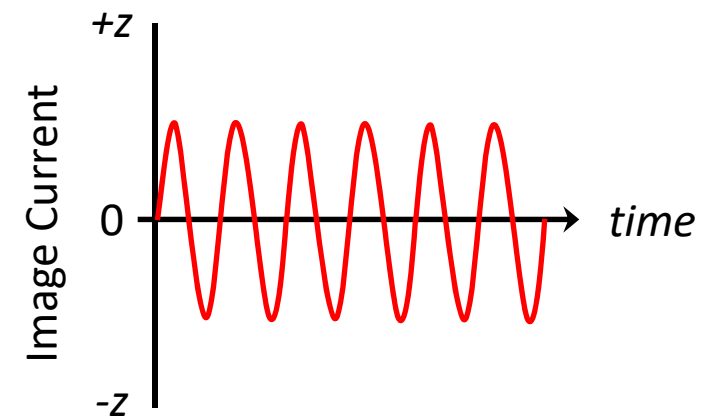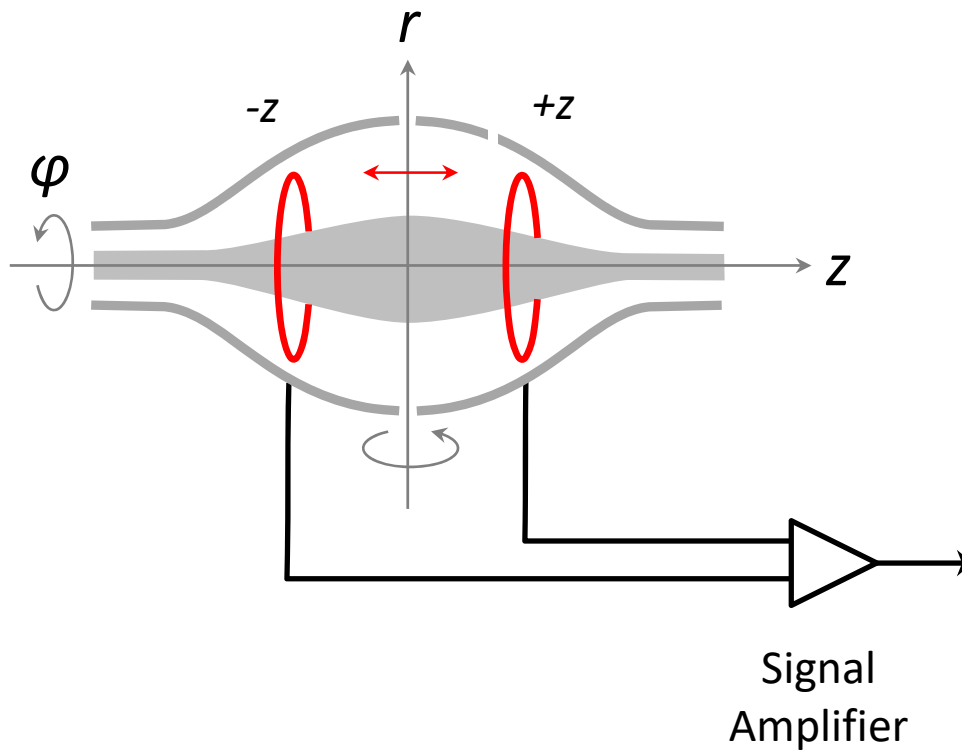
Resolving power in an Orbitrap

Resolving power (RP) is proportional to the detection time (T_{det}) divided by the oscillation period (T).

$$RP \propto \frac{T_{det}}{T} \qquad T \propto \sqrt{\frac{m}{z}}$$
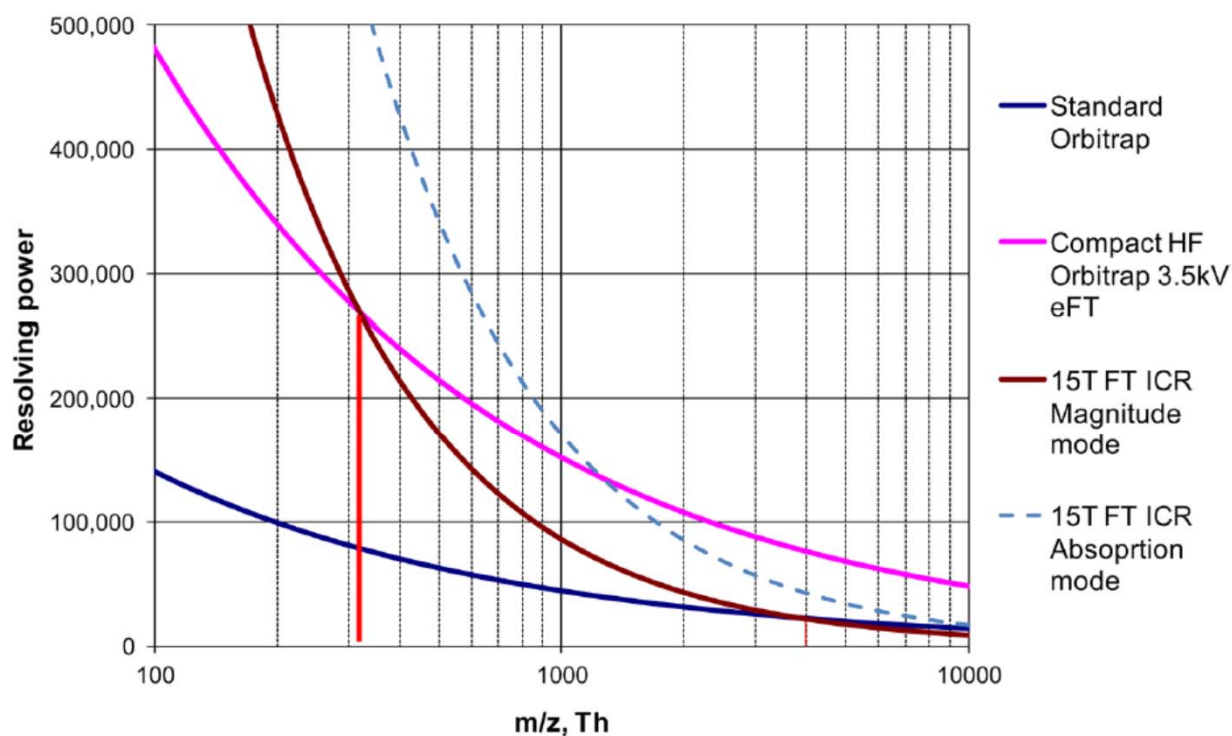
Figure 4. Dependence of resolving power on *m/z* for the following analyzers (all data are shown for a 0.76 s scan): (i) standard trap (magnitude mode, 3.5 kV on central electrode), (ii) compact high-field trap (eFT, 3.5 kV on central electrode), (iii) FTICR (magnitude mode, 15 T), (iv) FTICR (absorption mode, 15 T).

Zubarev and Makarov, *Anal Chem* **2013**, 85, 5288-5296

Separation Science and Mass Spectrometry

- **The introduction of ESI and MALDI significantly broadened bioanalytical mass spectrometry, including the analysis of glycans**

- **Critical to these efforts has been separation science**

- **Separation science covers a broad spectrum of approaches**
  - liquid-liquid extraction
  - Gel electrophoresis
  - Capillary electrophoresis
  - Solid phase extraction
  - High- and ultra-high liquid chromatography (HPLC and UHPLC/UPLC)

- **ESI is generally used to directly couple HPLC/UPLC to a mass spectrometer (e.g., LC-MS or LC-MS/MS)**

- **MALDI is generally used to directly analyze native (no separation) or separated species (e.g. glycans).**
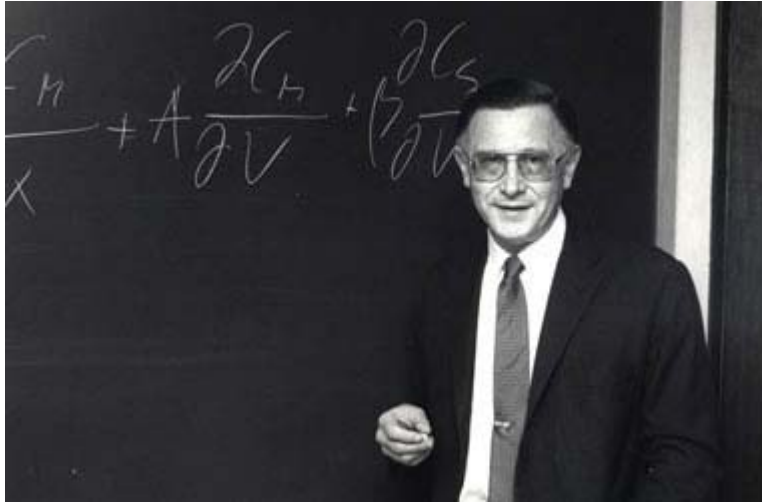
Inventor of HPLC

Professor Csaba Horváth
1930-2004

1952 Graduated with a degree in chemical engineering from the Budapest Institute of Technology.

1962 Graduated with a Ph.D. in Chemistry from J.W. Goethe University in Frankfurt

1964 Joined the Medical School at Yale and built the first high performance liquid chromatograph in 1964 to demonstrate the feasibility and potential of HPLC in bioseparation sciences.

1972 Became a faculty member in the Department of Chemical Engineering at Yale where he spent the remainder of his career.

Separation Science: on-line and off-line coupling to Mass Spectrometry

On-line (ESI: LC-MS)

ESI

Mass Spectrometer

Mixer

Flow

HPLC Column

Pump A

Injector

Pump B

Ion Abundance

Time

Off-line (MALDI)

MALDI Spotter

Silica Capillary (360 μm o.d.)

Sample Spotting

MALDI Target

Solvent: Water/Organic

⬡ Matrix

⬭ Analyte

Evaporation

Crystallization of analyte inside matrix
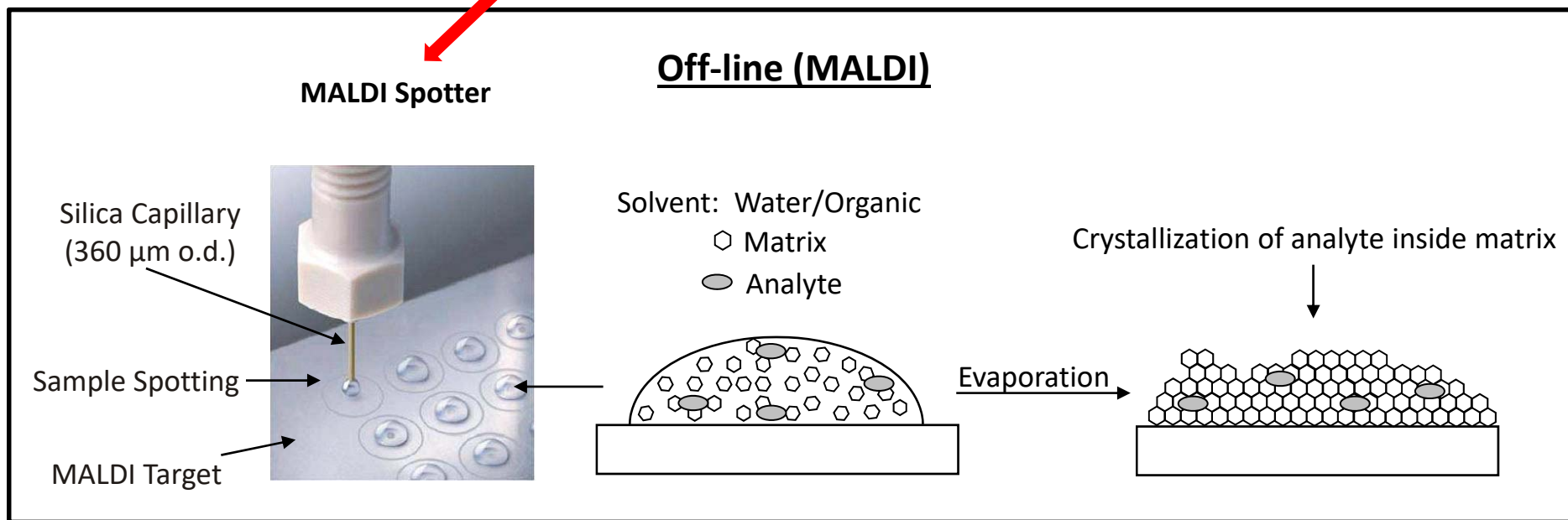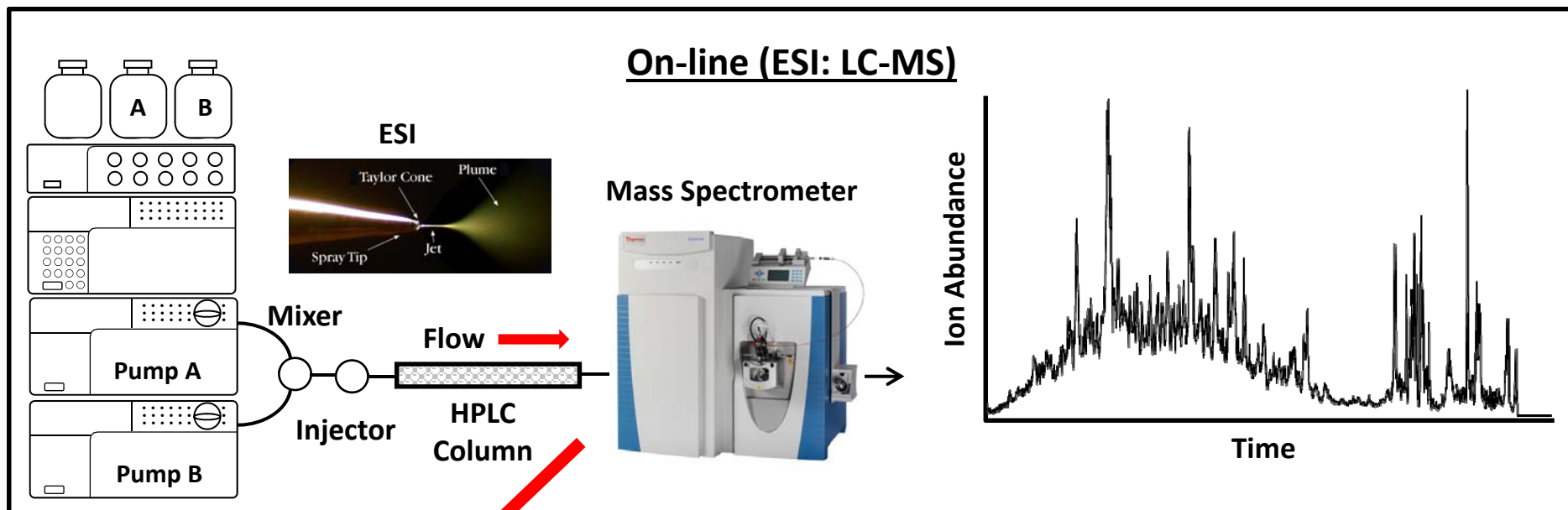
Important modes of high-performance liquid chromatography

- Reverse-phase chromatography (RP)
    - By far the most commonly used form of HPLC
    - Retention mechanism is based on hydrophobicity of analyte

- Normal-phase (NP) and Hydrophilic Interaction liquid chromatography (HILIC)
    - Becoming more common form of HPLC
    - Retention mechanism is based on acid/base interactions and hydrophilicity

- Ion exchange chromatography (IEX)
    - Very useful for separating charged species such as drugs
    - Retention mechanisms are based on acid/base character and strengths of solutes (e.g., drugs)
    - Requires the use of high salt concentrations (10-1000 mM) which are incompatible with ESI (and MALDI for highest salt concentrations)

- Graphitized carbon
    - Widely used stationary phase for glycans
    - Has been shown to be able to separate isomeric forms of glycans

- Size exclusion chromatography (SEC)
    - Important for large biomolecules such as proteins
    - Separation mechanism is based on molecular weight
    - Requires the use salts making it incompatible with ESI.

HPLC Stationary Phases Are Primarily Porous Silica

Porous silica stationary phases provide high surface area substrates which are amenable to a variety of chemical functionalities. Although pH stability can be an issue, silica-based columns are by far the most commonly used in HPLC.

Porous Silica Particle (5-1.8 μm)

OH

CH_3

O—Si—R

CH_3

OH

Typical pH range 2-8 (1-10 for newer phases)

Reverse-Phase (Non-polar)

C18 CH_3

C8 CH_3

C4 CH_3

Hydrophobicity

Normal-Phase/HILIC

Amino NH_2

Cyano $C≡N$
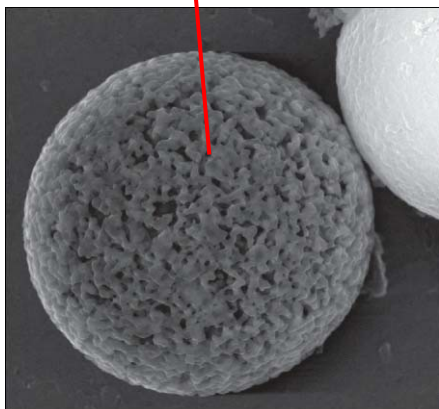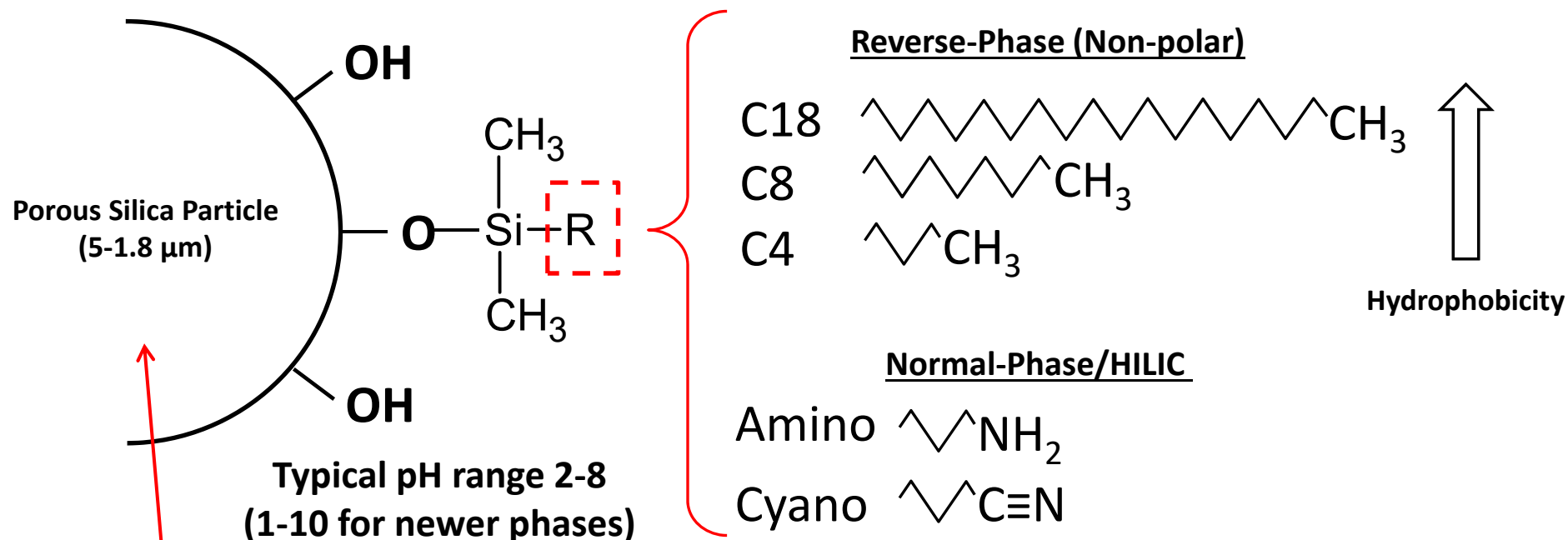
Typical porous silica properties:
Surface area ~170 m^2/g
~8 μmol SiOH groups/m^2
~4 μmol R groups/m^2
% Carbon (C18) on porous silica ~10-15%

Reverse-phase HPLC Stationary Phases-Mobile Phase Partitioning

Reverse-Phase HPLC: Isocratic

A = H₂O
B = Acetonitrile

A = H_2O
B = Acetonitrile

20% B

Pump A

Pump B

Inject

t = 0 min

t = 5 min

t = 12 min

t = 24 min

UV-vis

Detector Response→

% B

100

80

60

40

20

0

20% B

0   5   10   15   20   25   30   35

Time (minutes)
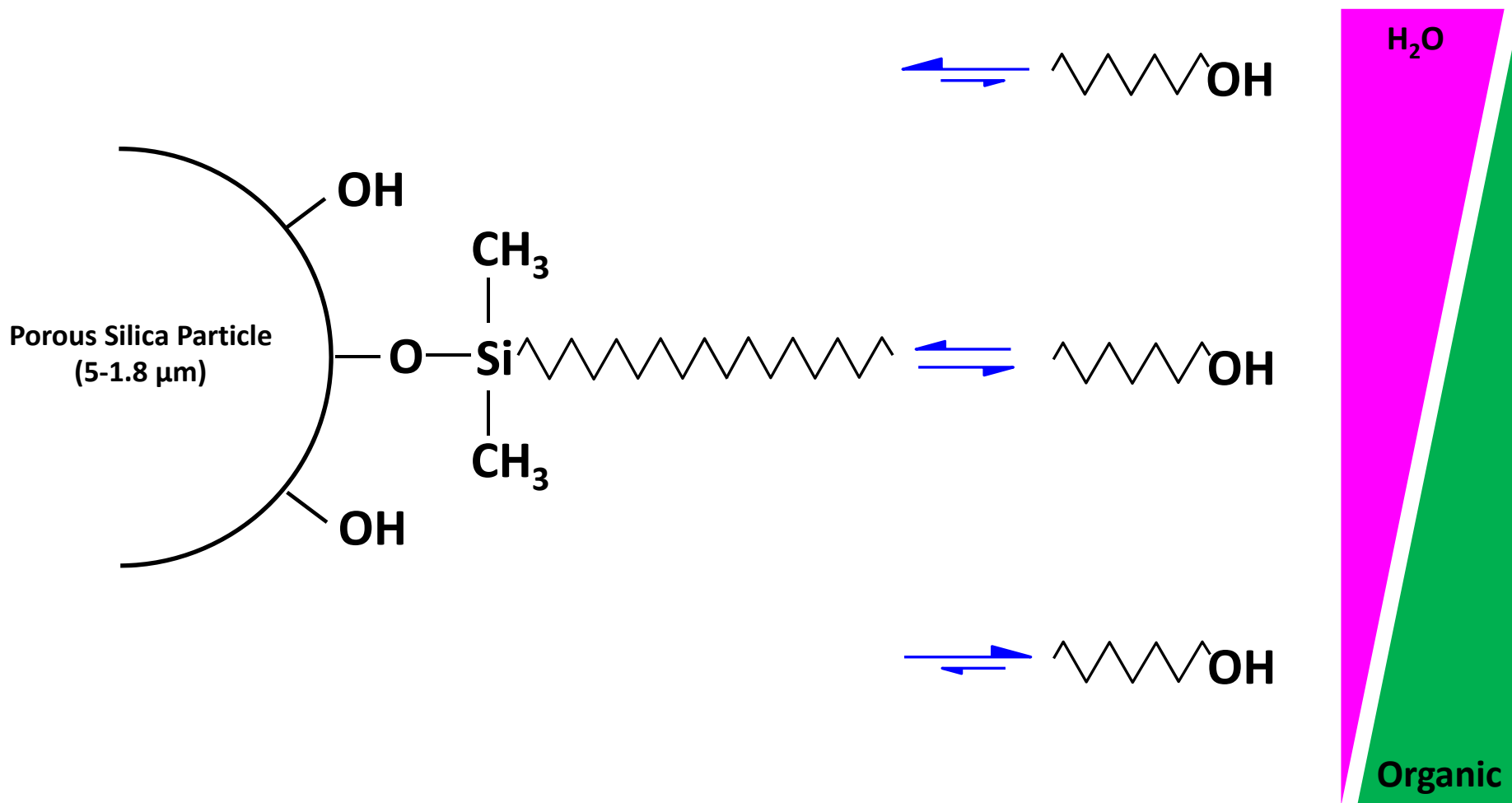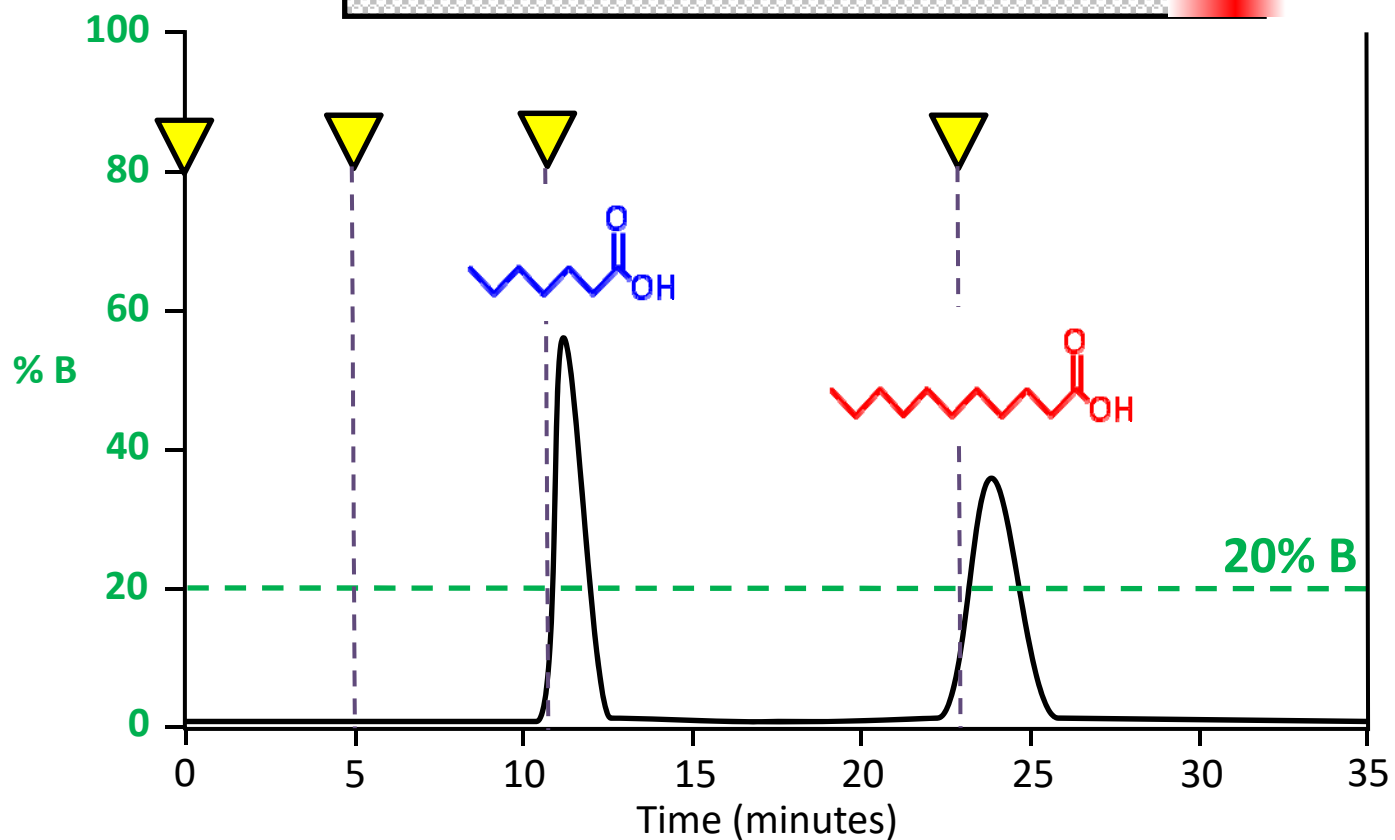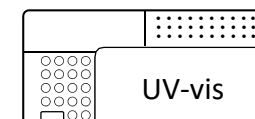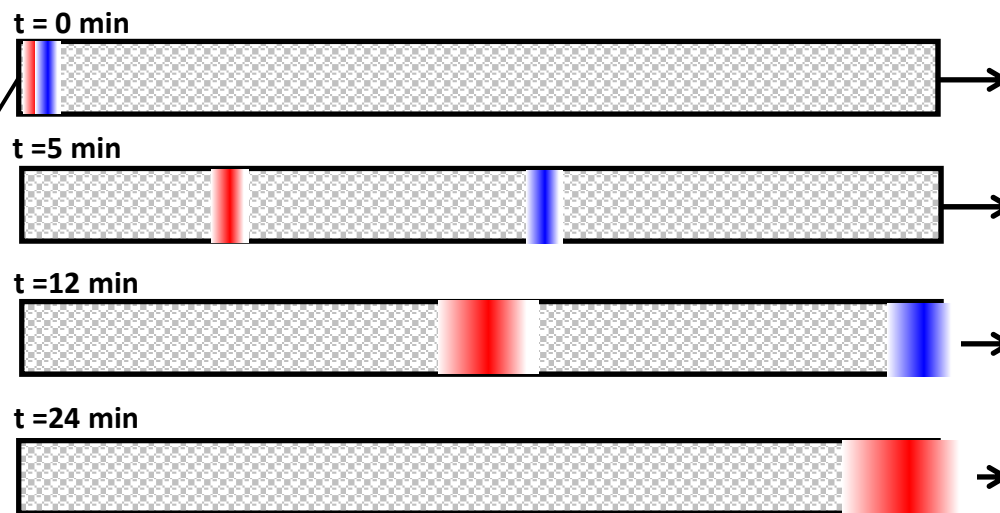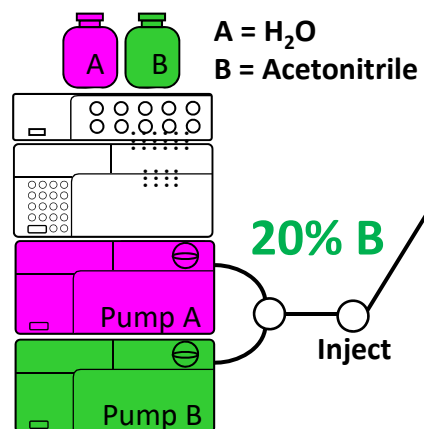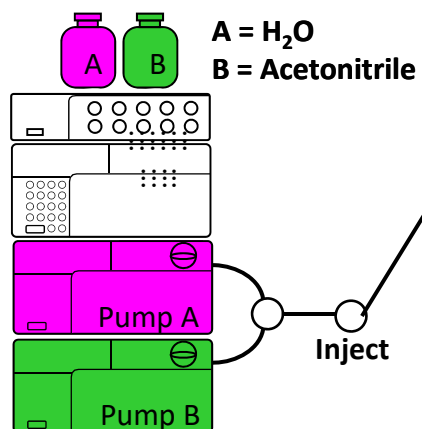
Reverse-Phase HPLC: Gradient
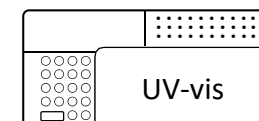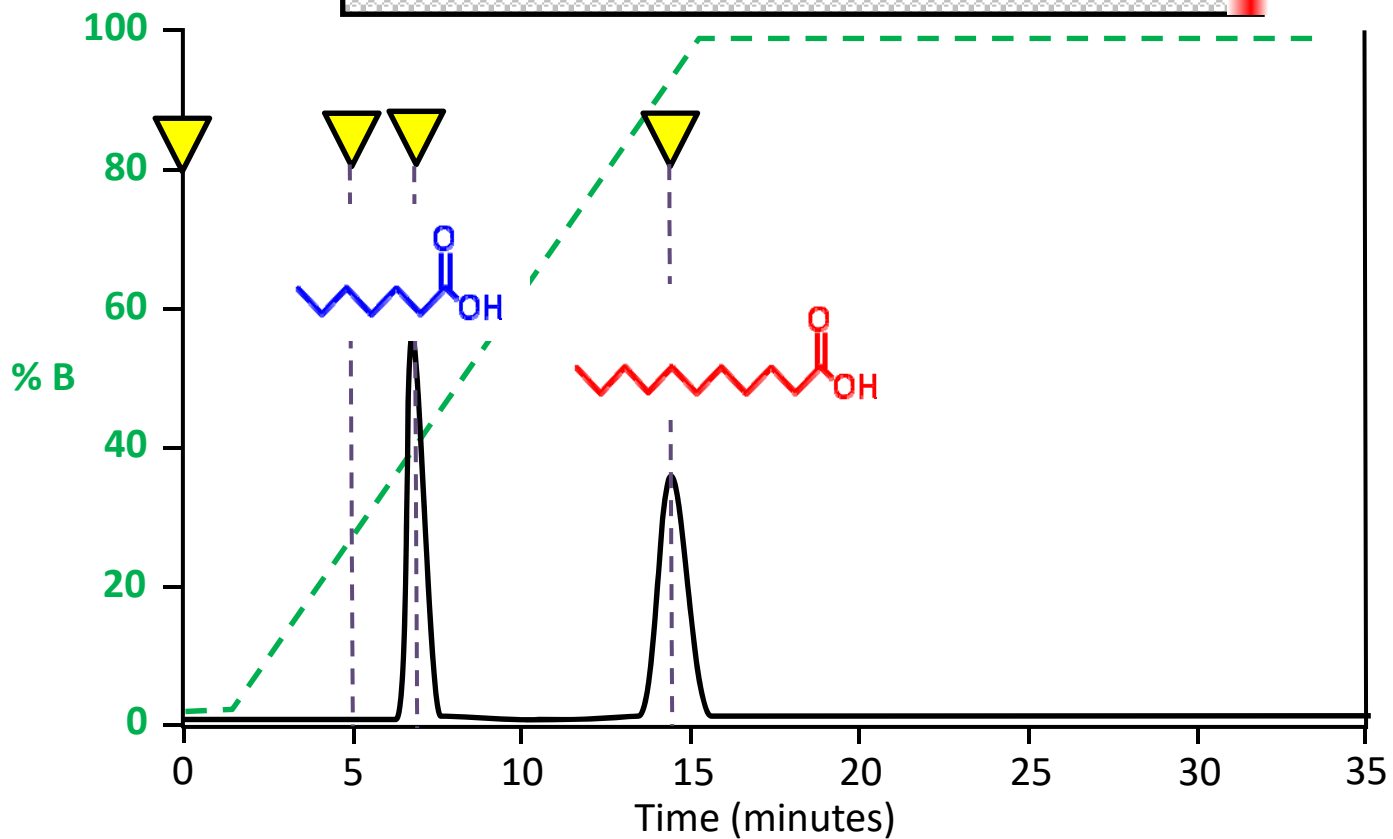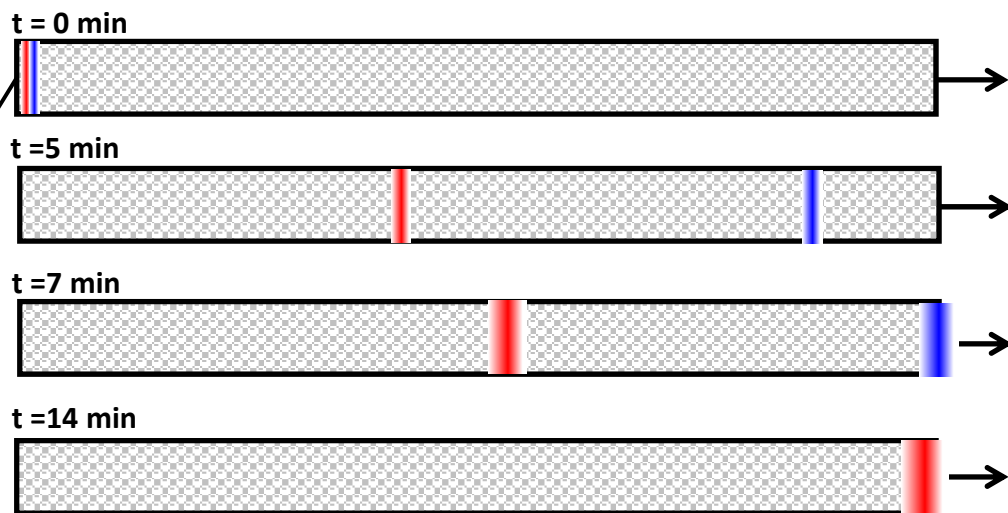
Isocratic Elution of Two Species with Different Hydrophobicities: C-18

A = H$_2$O
B = Acetonitrile

C-6 Carboxylic Acid
"less hydrophobic"

C-10 Carboxylic Acid
"more hydrophobic"

Pump A

Pump B

Inject

FLOW

C-18 Reverse-phase HPLC Column

UV-vis

0% B

30% B

60% B

Retention Time

0% B

30% B

60% B

Mobile Phase
(MP)

$$k = \frac{t_R - t_M}{t_M} = \frac{t_S}{t_M}$$

$$k = \frac{\text{time-SP}}{\text{time-MP}}$$

Stationary Phase
(SP)

C-18

C-18

C-18
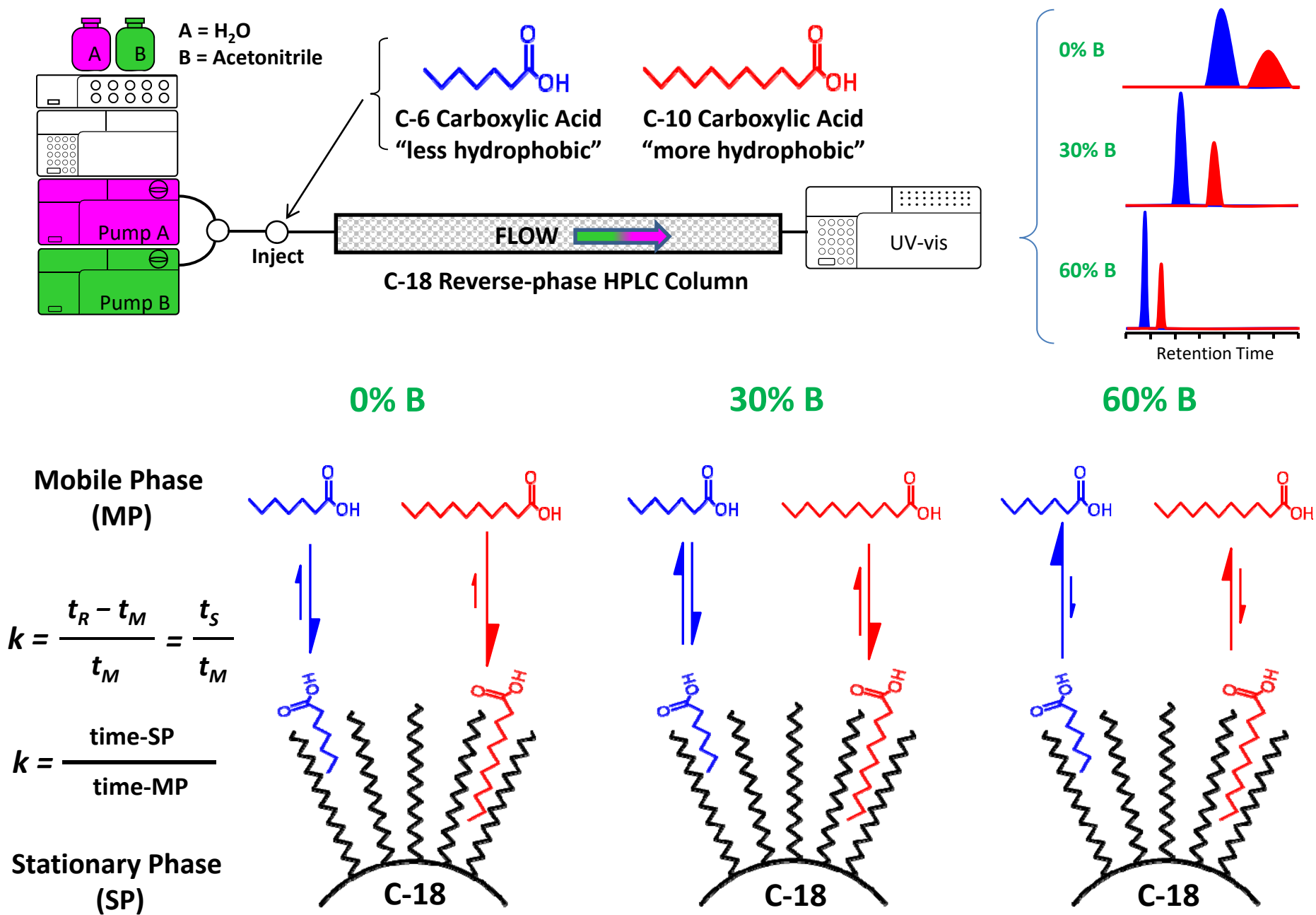
Glycan Analysis by Mass Spectrometry Examples

- Overview of glycans
  - N-linked and O-linked glycans
  - Glycosaminoglycans
- Standard methods for preparing glycan samples for mass spectral analysis
- Derivatization strategies
- Representative LC-MS/MS data for N-glycans and glycosaminoglycans

Adam M. Hawkridge

School of Pharmacy

MEDC 691

Spring 2018

Analysis of Glycans: Overview

- **Complex carbohydrates are challenging biomacromolecules to characterize**
  - Unlike DNA, RNA, and proteins; carbohydrates are <u>non-template driven</u>
  - Carbohydrates are structurally (e.g., branched, linear) and chemically (e.g., sugar monomers, sialic acid, sulfate) heterogeneous
  - The number of possible unique carbohydrates (i.e., glycome) is unknown
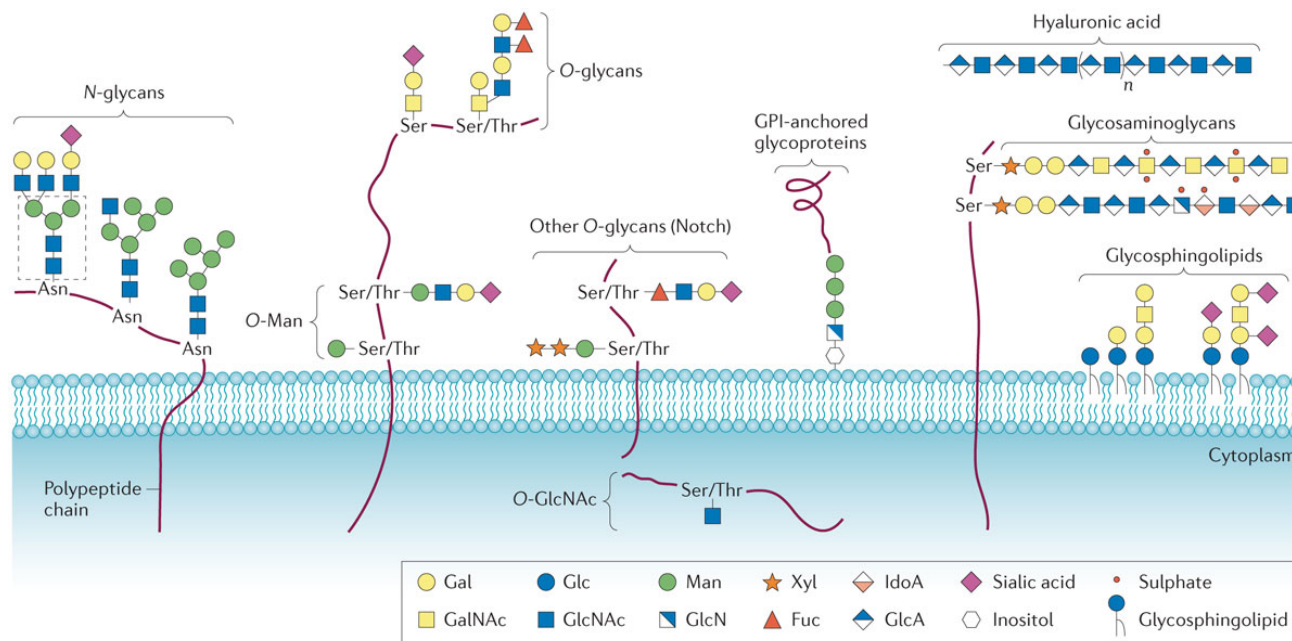
Nature Reviews | Cancer

Major classes of glycans

- **N-linked glycans**
- **O-linked glycans**
- **Glycosaminoglycans (GAGs)**
  - Hyaluronic acid
  - Heparin
  - Heparan sulfate
  - Chondroitin sulfate
  - Keratin sulfate
  - Dermatan sulfate
- **Glycolipids**

Aberant glycosylation and glycan modification are common in disease

- **Altered glycome in cancer**
  - Glycoproteins in cancer are modified in at least one or more of the following ways:
    - N-glycan and O-glycan structure via glycotransferase dysregulation
    - Differential modification of sugar monomers (e.g., sialic acid, sulfation)
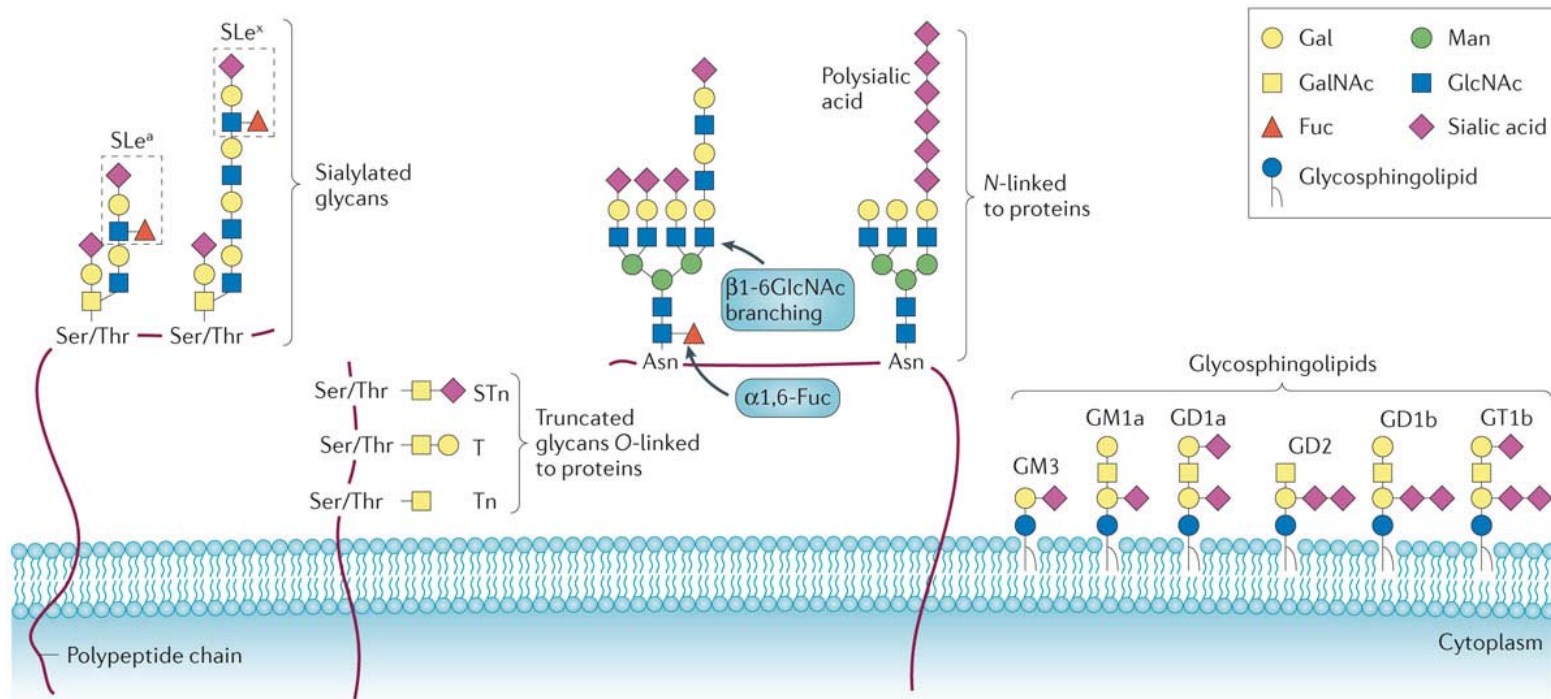
The size of the glycome is unknown: between $7 \times 10^3 - 2 \times 10^{11}$

TABLE 1. Diversity space of oligonucleo-tides, peptides, and mammalian oligo-saccharides[a]

Oligomer size	Numbers of different oligomers		
	Nucleotides	Peptides	Carbohydrates
1	4	20	20
2	16	400	1360
3	64	8,000	126,080
4	256	160,000	13,495,040
5	1024	3,200,000	1,569,745,920
6	4096	64,000,000	192,780,943,360

[a]The numbers for the mammalian oligosaccharides are based on the 10 mammalian monosaccharides: D-Glc [4], D-Gal [4], D-Man [4], D-Sia [4], D-GlcNAc [3], D-Gal-NAc [3], L-Fuc [3], D-Xyl [3], D-GlcA [3], and L-IdoA [3]. The number of substitutable OH groups (excluding the ano-meric one) is given in square brackets. Commonly, only the pyranose ring forms and not the furanose ring forms of the above-mentioned monosaccharides are found in mammals (35, 36).

ACS Chemical Biology, 2007, 2(10), 685-691.

Estimates for the Number of Glycan Determinants

Oligomer Size = 6

↓

~200 billion possibilities

↓

Filter based on:
Glycan-binding proteins (N=700)
(e.g., lectins, enzymes, toxins, etc.)

↓ ↓

3000 N/O-glycans

4000 GAGs
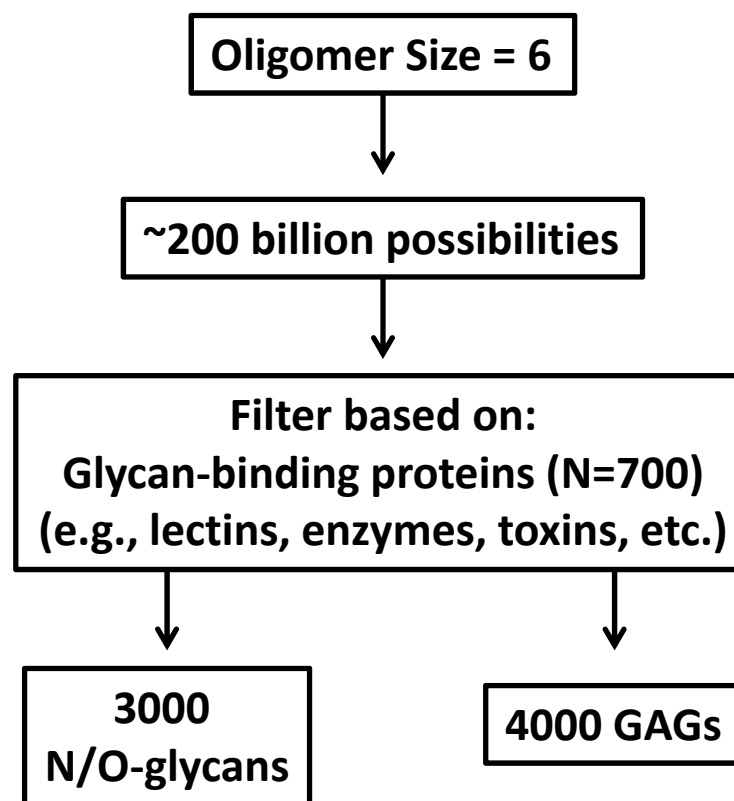
Mol. BioSyst., 2009, 5, 1087–1104.

Analytical challenges related to mass spectrometry analysis of glycans

- **Over 50% of all expressed proteins are glycosylated**

- **The site(s) of glycosylation and the physiochemical structure of the glycan at each site can and does exert important biological functions**

- **Glycans are generally one of the most hydrophilic species in nature. When coupled to proteins, you are presented with a complex biomacromolecules containing both amino acid sequences and complex glycan structures. Furthermore, chemical modification to the glycans (e.g., sialylation, sulfation) make these biomolecules even more complex.**
  - Separation and analysis of these complex species is immense

- **We are only just beginning to unravel this complexity with the introduction of new separation strategies coupled (directly or indirectly) with advanced mass spectrometry.**

Multiple glycoanalytical techniques are used to study glycoproteins
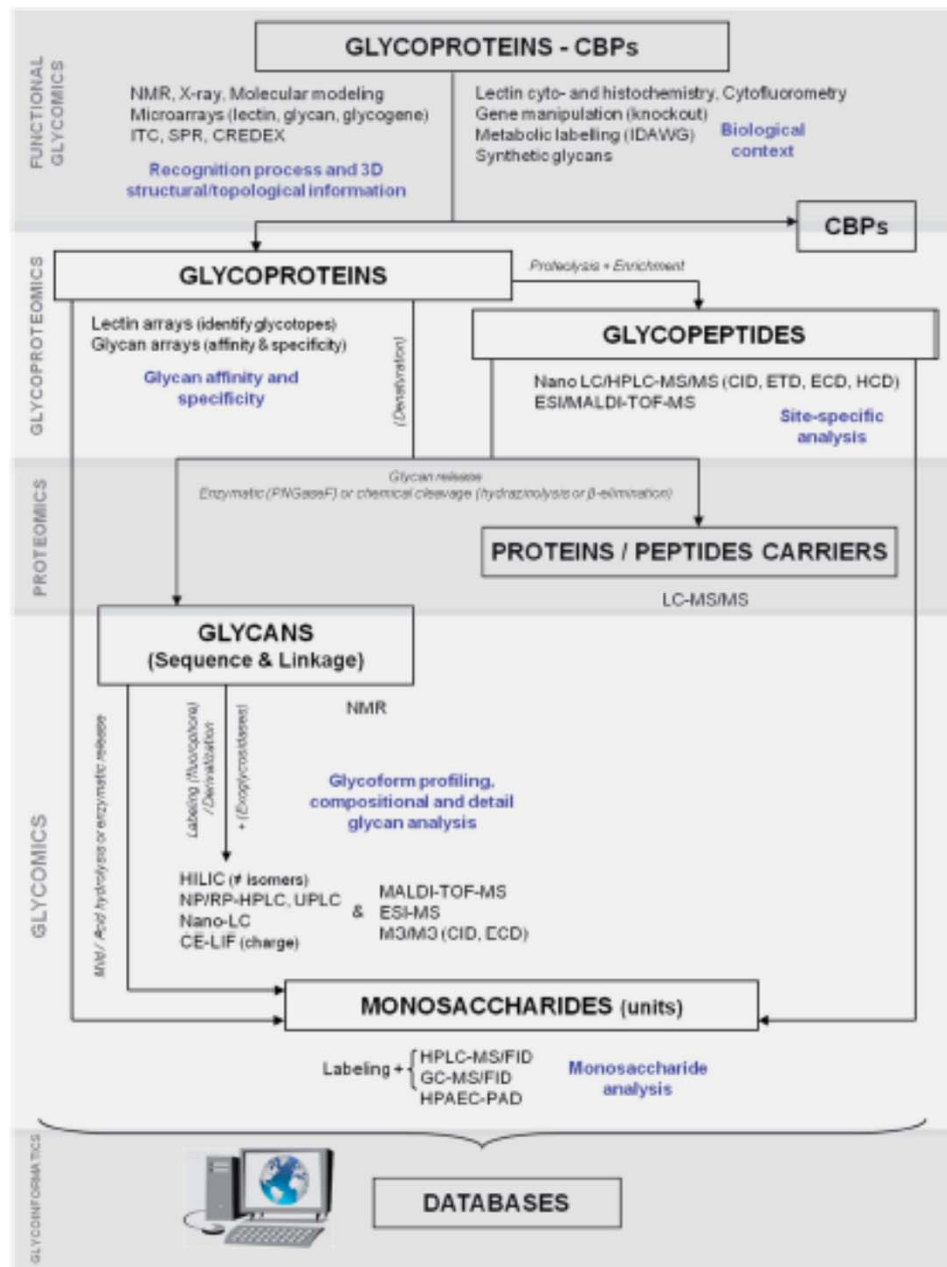
Mass spectrometry (MS) plays a central role in all aspects of glycoprotein analysis. To begin with, we group the types of MS used in glycoprotein analysis into three types based on their ionization method:

1) Matrix-assisted laser desorption ionization (MALDI-MS)

2) Electrospray ionization (ESI-MS) or liquid chromatography (LC-MS)

3) Gas chromatography (GC-MS)

Each of these 'types' of MS systems offer inherent advantages and disadvantages toward glycoprotein analysis. Importantly, there is no single 'best' MS platform although LC-MS is the dominant approach for reasons we will discuss.

Analyst, 2014, 139, 2944-2967.

General Workflow for MS-characterization of glycoproteins
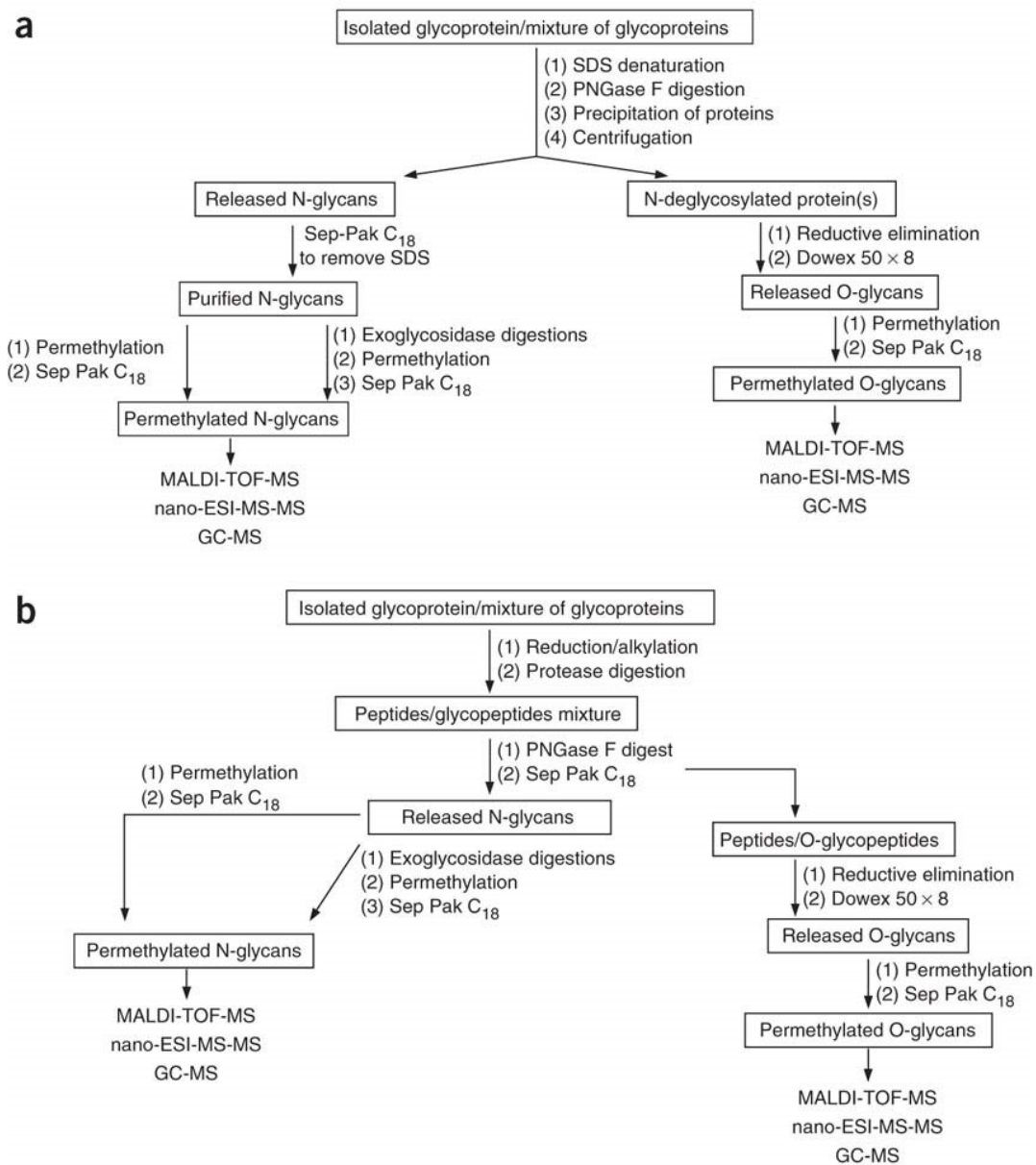
MS-based N-linked and O-linked glycan analysis

N-linked and O-linked glycans are typically removed from the proteoglycan via enzymatic and/or chemical methods prior to MS analysis

- N-glycan de-glycosylation of proteins can be accomplished enzymatically using endoglycosidases *PNGase F* or *Endo H*
  - *PNGase F* releases the N-glycan from the Asn sidechain allowing the N-glycans to be further processes for MS analysis
  - *Endo H* cleaves at the GlcNAc residue allowing for N-glycan site occupancy
  - Following cleavage from the proteo/peptideglycan, the N-glycans can be analyzed in their native state or derivatized to facilitate better separation and/or quantification.

- Complete O-glycan de-glycosylation of proteoglycans is not possible with a single endoglycosidase.
  - O-glycosidase (Endo-α-*N*-Acetylgalactosaminidase) releases some O-glycans
  - Beta-elimination is required for full deglycosylation of O-glycans. The downside of this is that the base required for this reaction degrades the protein.

De-glycosylation Animation

N-linked glycans

- N-linked glycans (N-glycans) are linked to proteins at asparagine residues located within a consensus sequence of **Asn-X-Ser/Thr** where X = any amino acid except proline.
- N-glycans contain a 'conserved core' that allows linear and branched linkages to extend outward from the protein.
- There are estimated to be between 3000+ biologically relevant N-glycans in humans
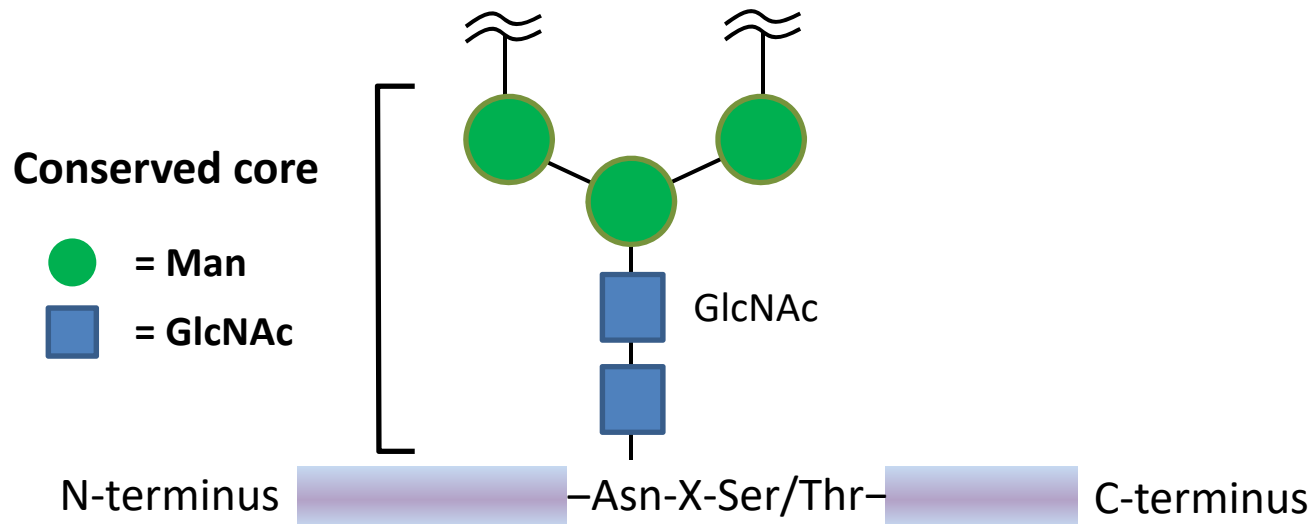
General Classifications of N-linked glycans

Conserved core

Precursor glycan   **High mannose**   **Complex**   **Hybrid**

GlcNAc
Glc
Man
Gal
Neu5Ac
Fuc
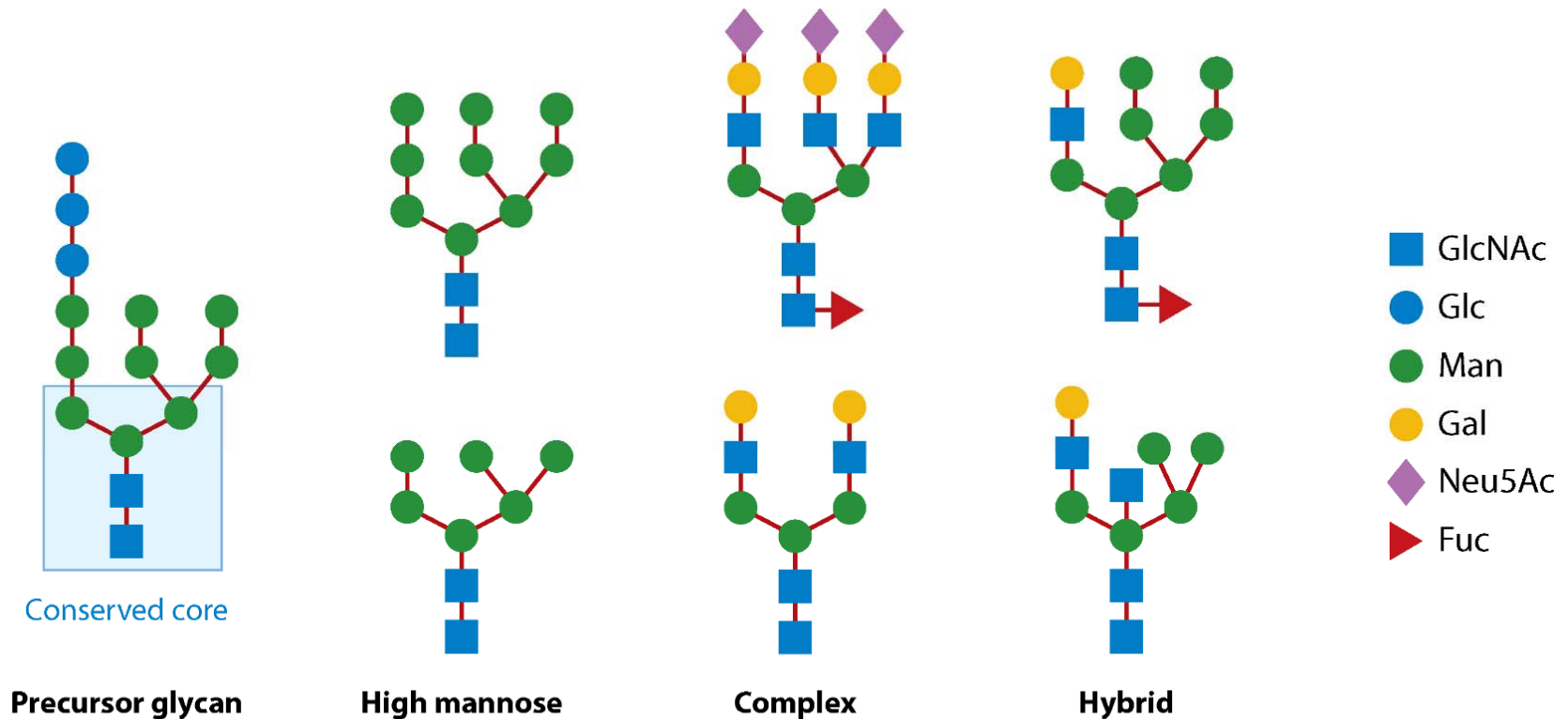
Zhu Z, Desaire H. 2015.
Annu. Rev. Anal. Chem. 8:463–83

Enzymatic Release of N-linked Glycans

Conserved core

Ser/Thr

X

Asn

Endo H

PNGase F

PROTEIN BACKBONE

C-terminus

$C=O$

$HO-CH_2-CH$ Ser (or Thr)

NH

$C=O$

$X-CH$ **Any AA**
(except Pro)

NH

$C=O$

HOH_2C HOH_2C HOH_2C

O O O

O O $NH-C-CH_2-CH$ Asn

HO $NHAc$ HO $NHAc$ HO $NHAc$

NH

N-terminus
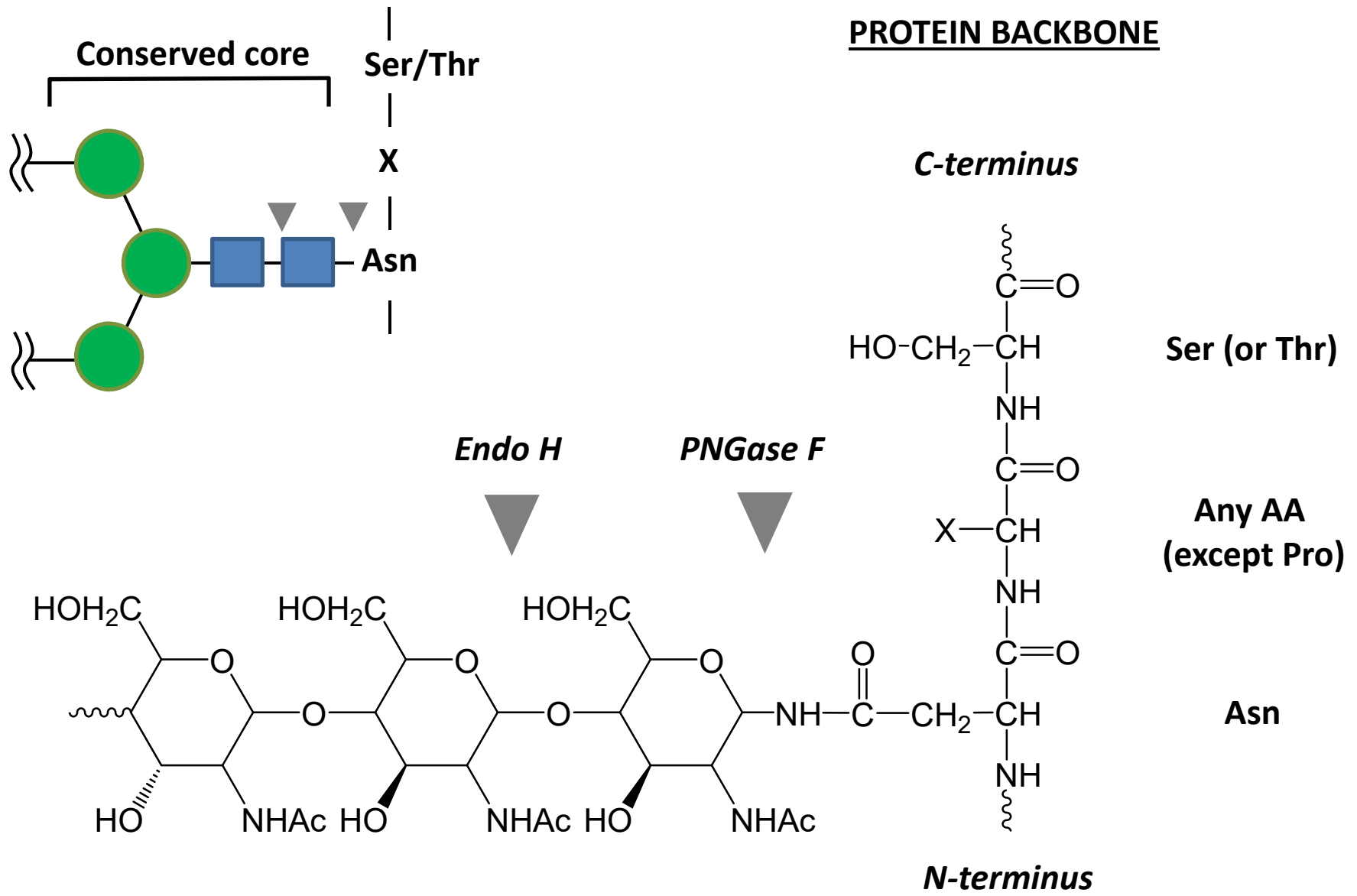
Enzymatic Release of N-linked Glycans: PNGase F
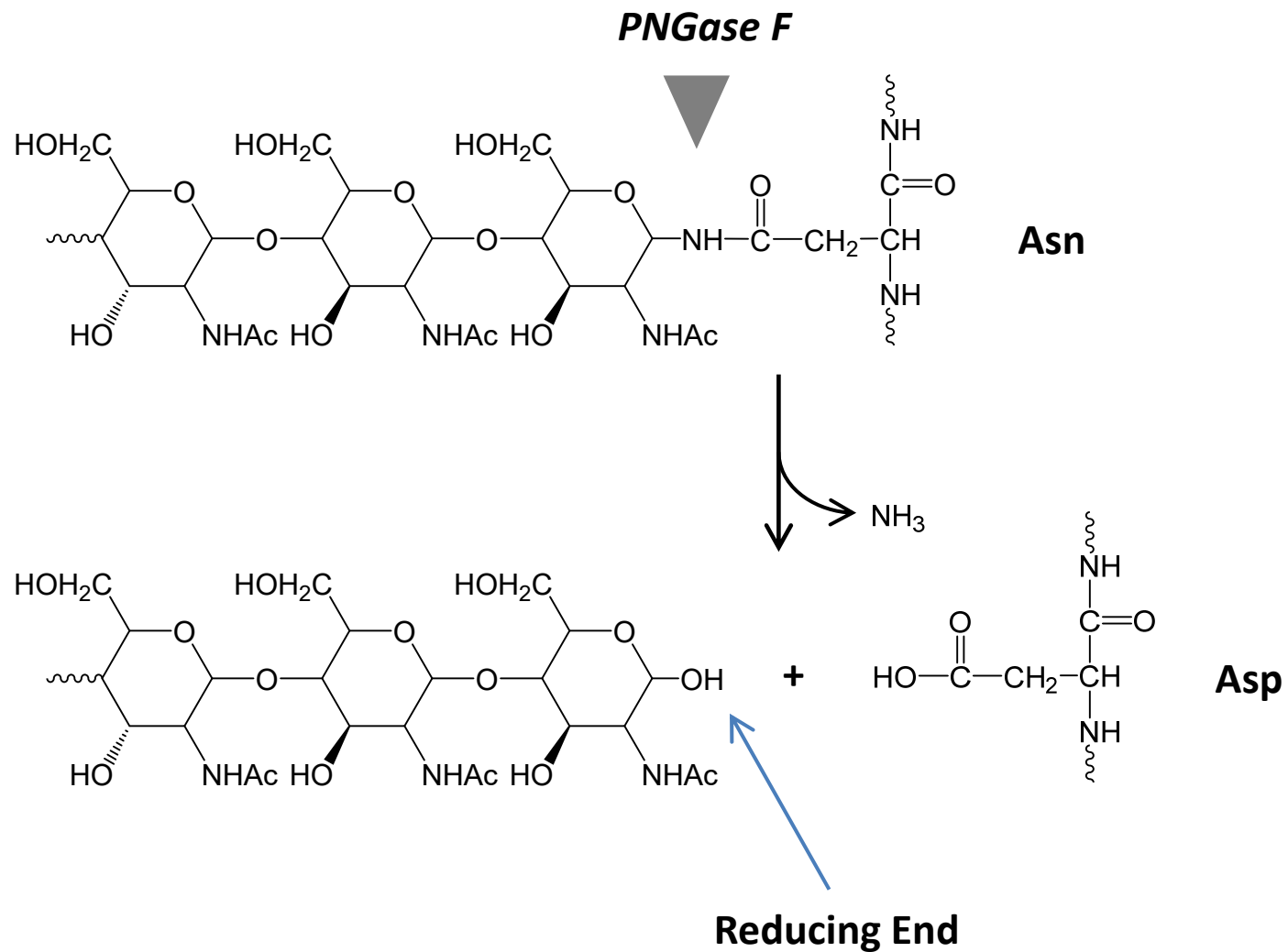
HILIC LC-MS/MS of N-glycans derived from human plasma
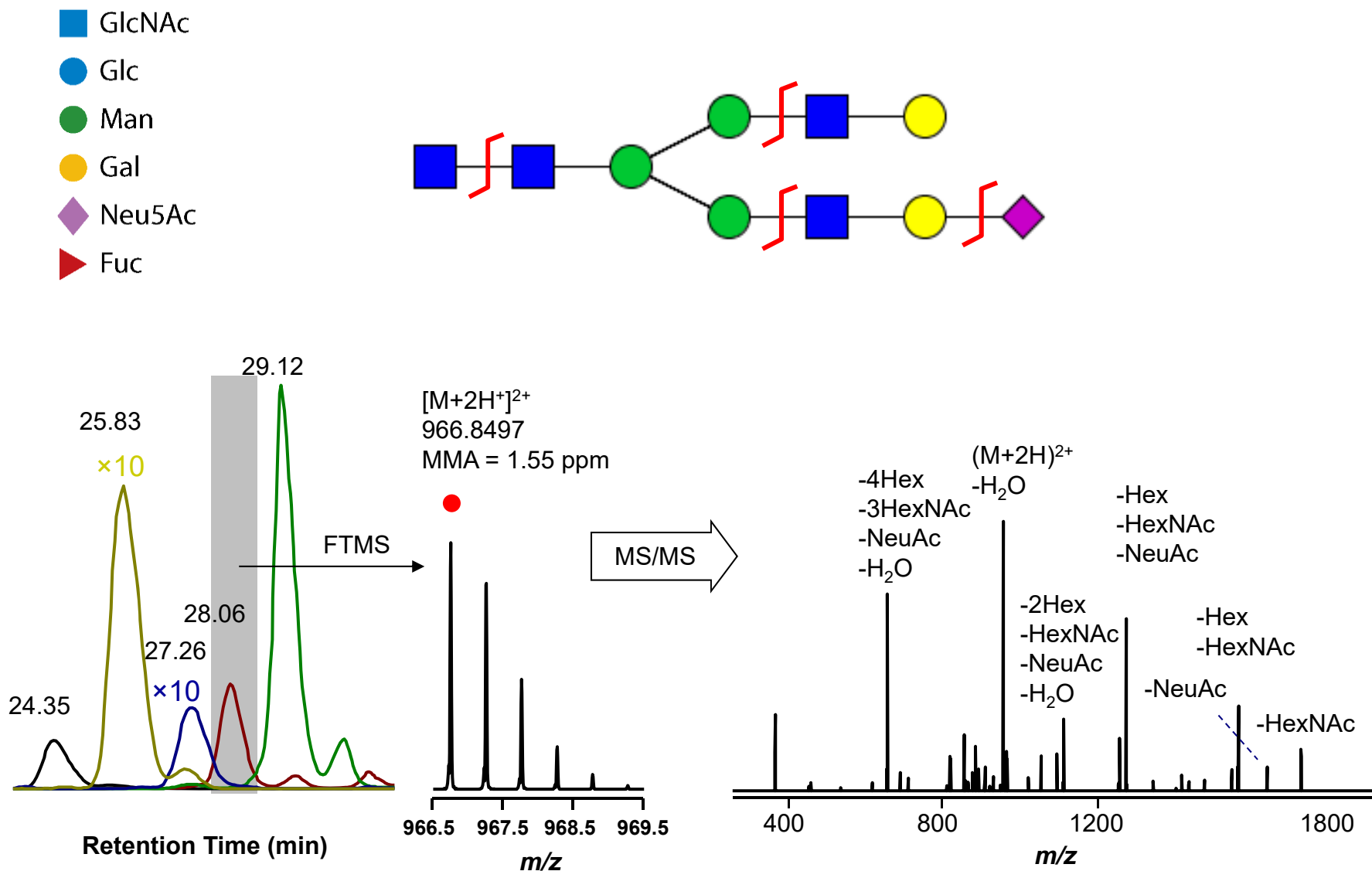
O-linked glycans

- O-linked glycans (O-glycans) are linked to proteins at serines (Ser) and threonines (Thr). There is no conserved amino acid sequence for O-glycans
- O-glycans contain 8 core structures that can then branch off into linear and branched linkages to extend outward from the protein.
- There are estimated to be between 3000+ biologically relevant O-glycans in humans

**Core
(8 possible structures)**

■ = GlcNAc

● = Gal

■ = GalNAc

N-terminus ▬▬▬ –Ser/Thr– ▬▬▬ *C-terminus*
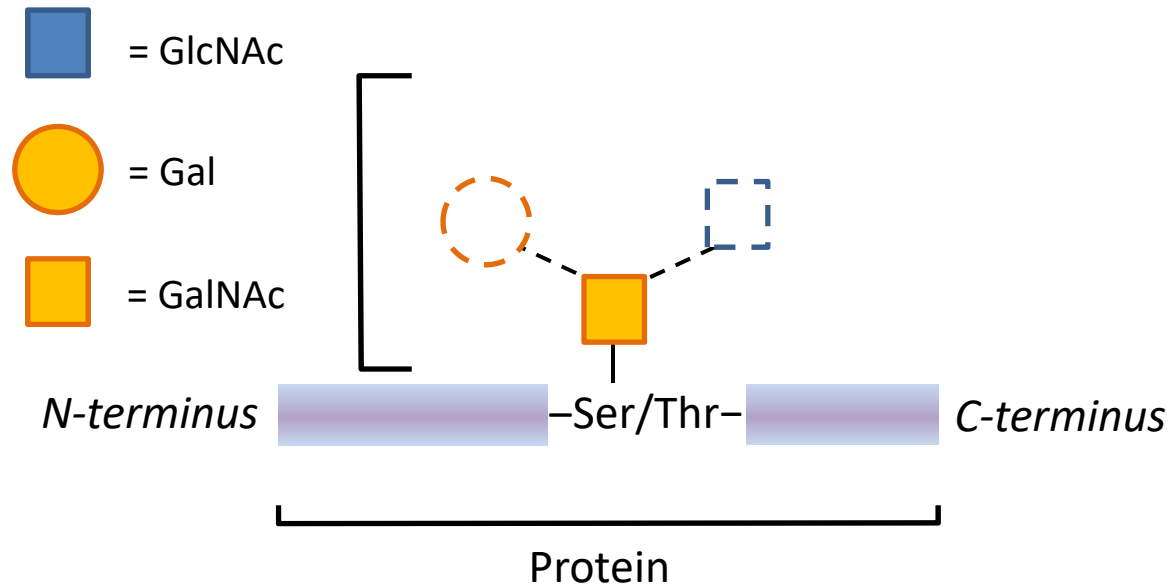
Protein

O-linked glycan core structures

Enzymatic release of O-linked glycans
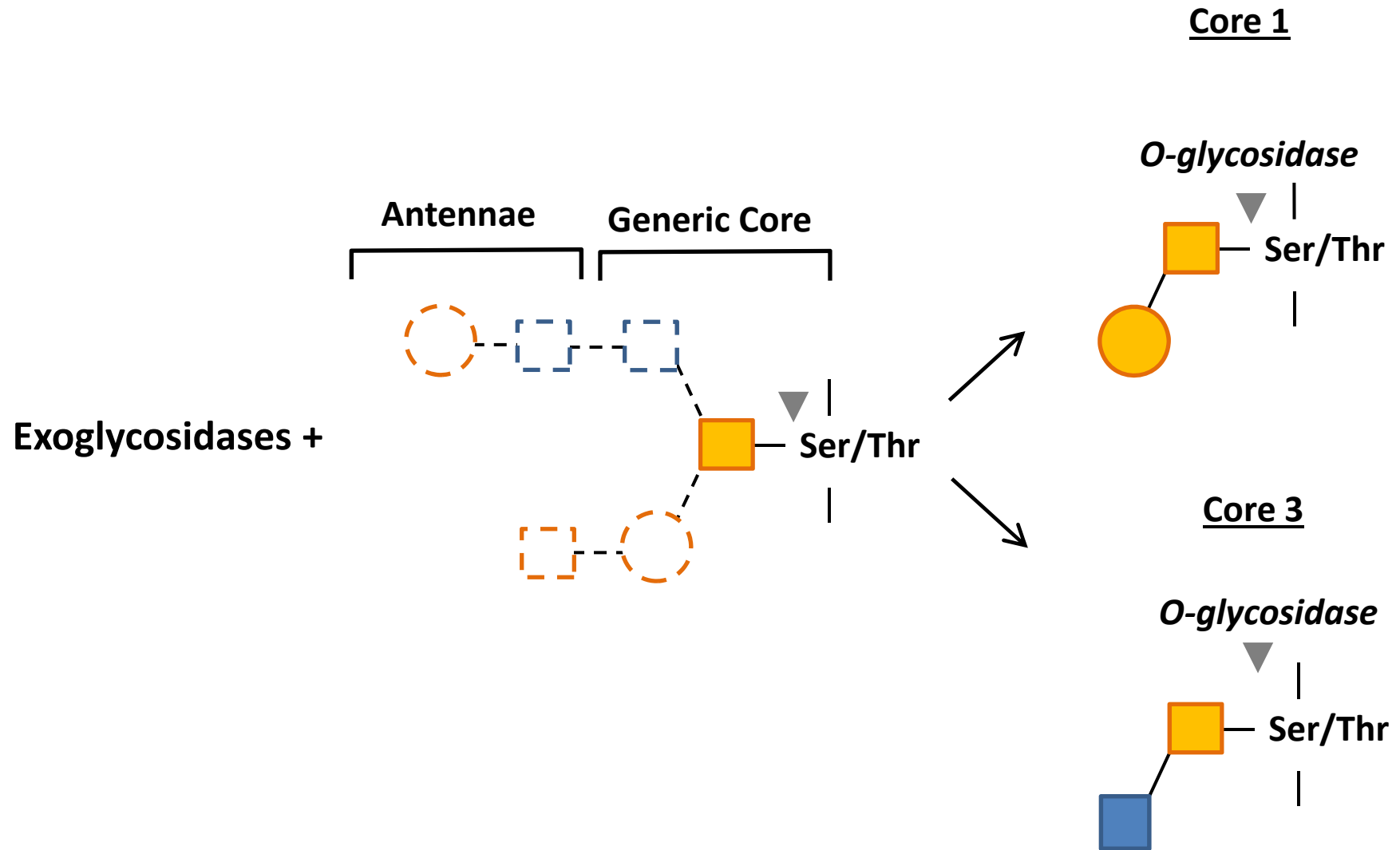
Chemical Structures of Core 1 and 3 O-linked glycans

PROTEIN BACKBONE

Core 1

O-glycosidase

▼

□—● Ser/Thr

O-glycosidase

▼

$\sim\!\!\sim\!OH_2C$

HO

O

$O-CH_2-CH$ Ser (or Thr)

$\sim\!\!\sim\!O$ NHAc

NH
C=O
NH

O-glycosidase

Core 3

O-glycosidase

▼

□—□ Ser/Thr

O-glycosidase

▼

$\sim\!\!\sim\!OH_2C$

HO

O

$O-CH_2-CH$ Ser (or Thr)

$\sim\!\!\sim\!O$ NHAc

NH
C=O
NH
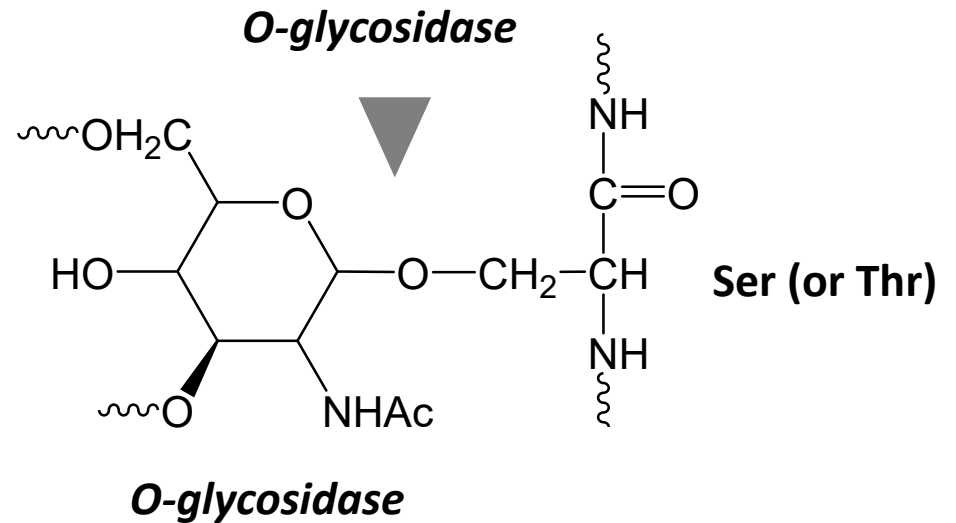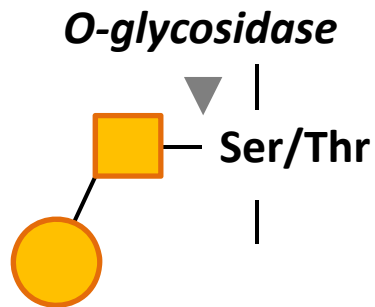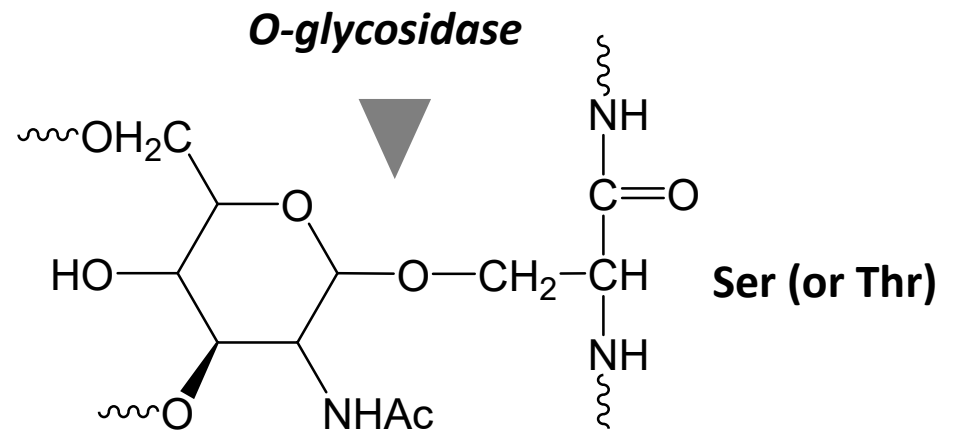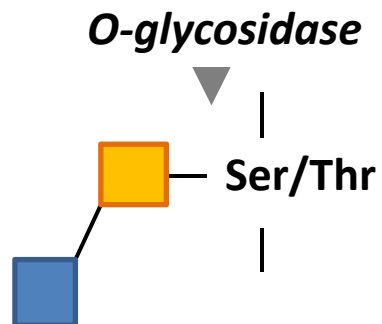
Removal of O-glycans from proteins/peptides via β-elimination
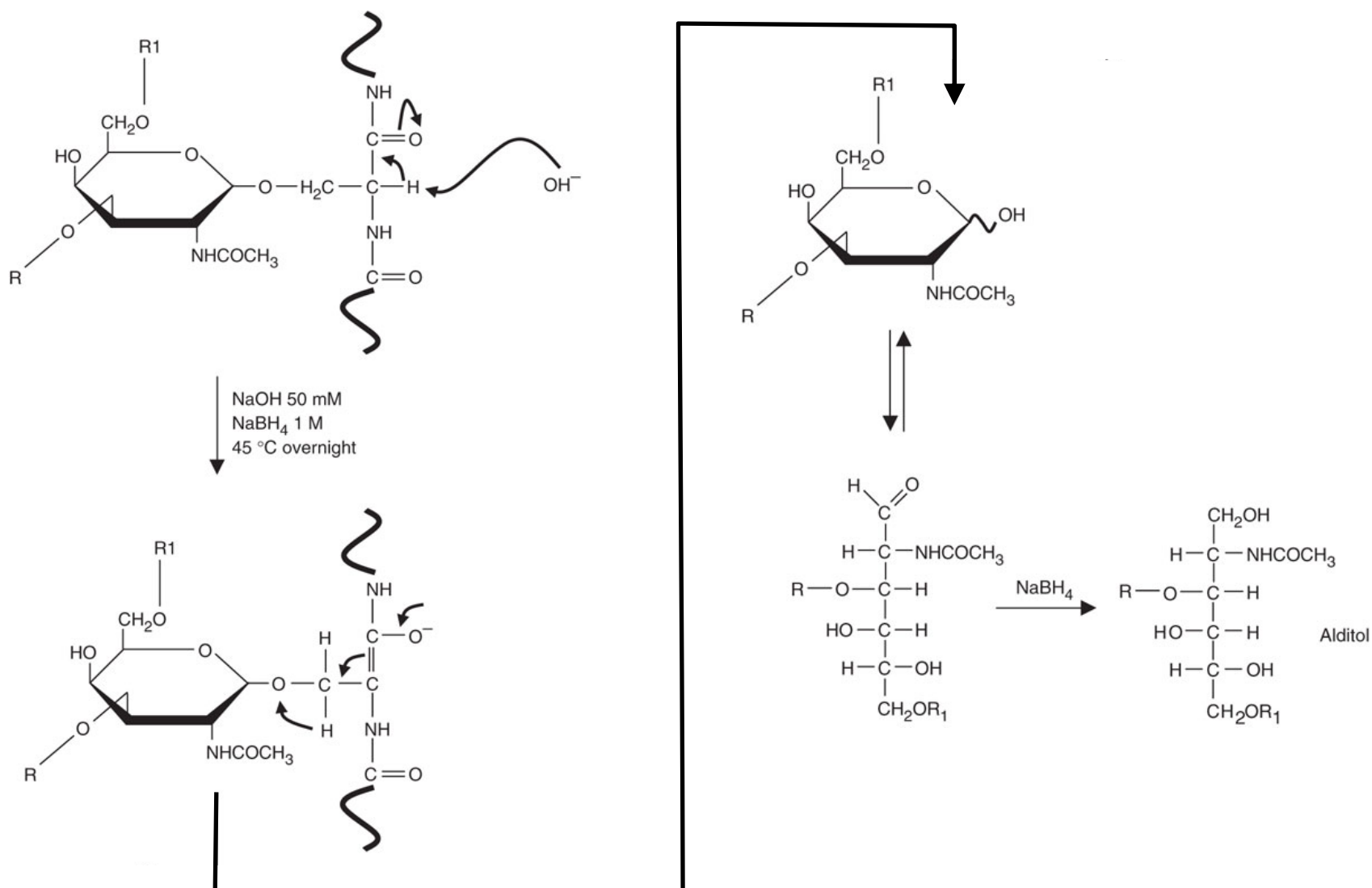
Linear glycans (glycosaminoglycans-GAGs)

- **Glycosaminoglycans (GAGs)**
  - Hyaluronic acid
  - Heparin
  - Heparan Sulfate
  - Chondroitin Sulfate
  - Keratin Sulfate

DS standards:

I-, II-, III-, IV-A
I-, II-, III-, IV-H
II-, III-, IV-S

R^2 = H, SO_3
R^6 = H, SO_3
Y = H, $COCH_3$, SO_3
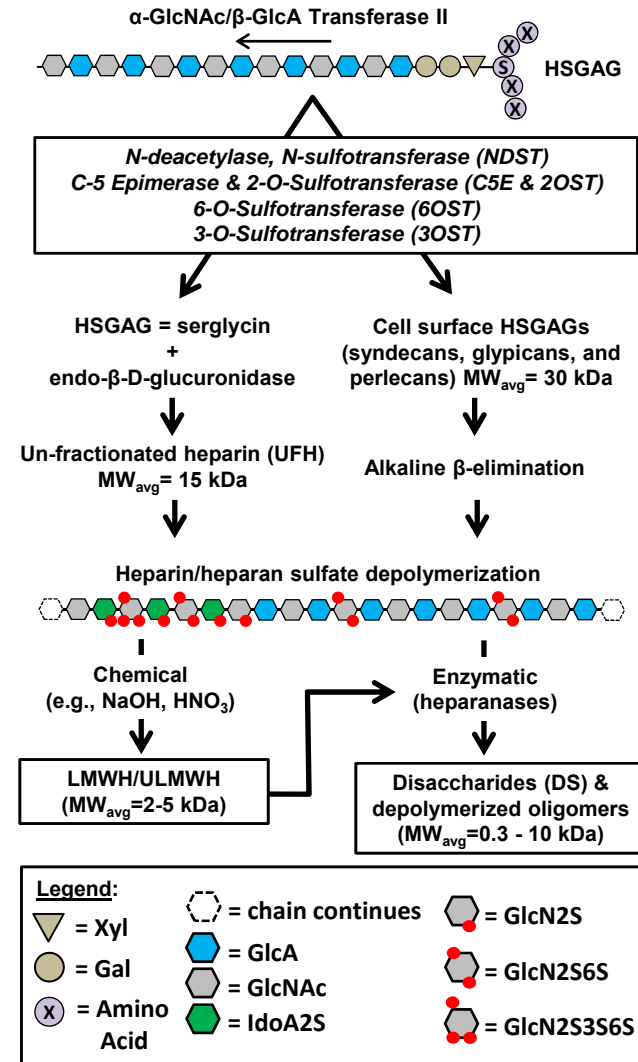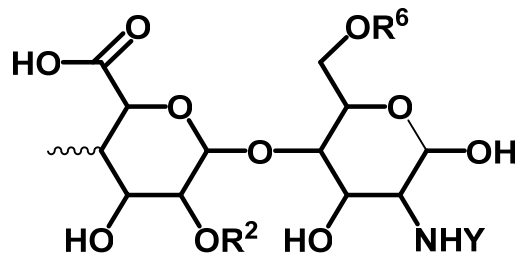
Figure 1. Heparin and heparan sulfate glycosoaminoglycan (GAG) biosynthesis and chemical/enzymatic processing overview.

Characterization of GAGs involves enzymatic digestion

Heparin and heparan sulfate general structure (n = 0 – 40)
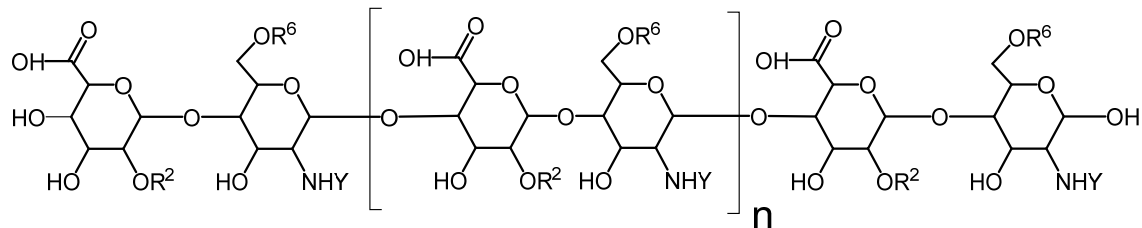
Heparinase I/II/III digestion

I-IV A   I-IV H   I-IV S
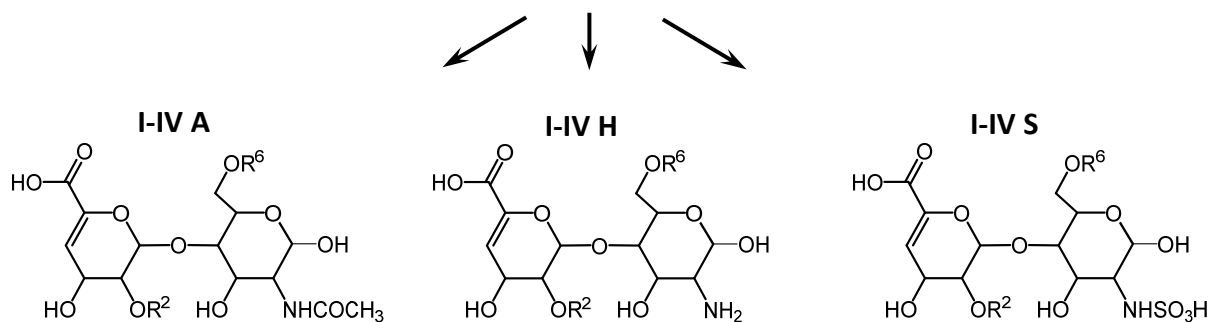
Nomenclature	R^2	R^6	Y	Free DS Mass	DS-INLIGHT
I-A	SO_3	SO_3	$COCH_3$	539.0253	774.1492 [1-]
II-A	H	SO_3	$COCH_3$	459.0088	694.1923 [1-]
III-A	SO_3	H	$COCH_3$	459.0088	694.1923 [1-]
IV-A	H	H	$COCH_3$	379.112	614.2355 [1-]
I-H	SO_3	SO_3	H	497.0151	732.1386 [1-]
II-H	H	SO_3	H	417.0583	652.1818 [1-]
III-H	SO_3	H	H	417.0583	652.1818 [1-]
IV-H	H	H	H	337.1014	572.2249 [1-]
I-S	SO_3	SO_3	SO_3	576.9719	405.5441 [2-]
II-S	H	SO_3	SO_3	497.0151	732.1386 [1-]
III-S	SO_3	H	SO_3	497.0151	732.1386 [1-]
IV-S	H	H	SO_3	417.0583	652.1818 [1-]

Chemical derivatization of the reducing end of glycans

11 DS standards:
I-, II-, III-, IV-A
I-, II-, III-, IV-H
II-, III-, IV-S

$R^2 = H, SO_3$
$R^6 = H, SO_3$
$Y = H, COCH_3, SO_3$

R can be any number of functionalities:
- Improve separation efficiency
- Improve detection
  - UV-vis and fluorescence
  - ESI
- Introduce stable isotope label for LC-MS quantitation
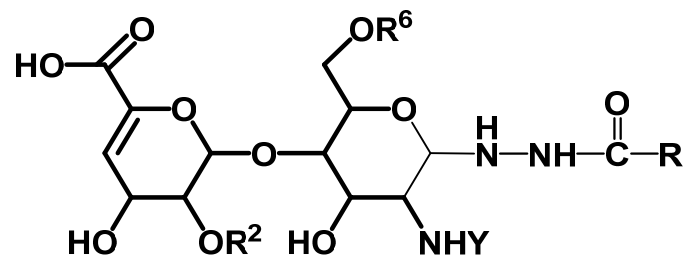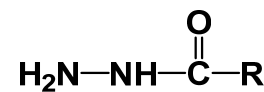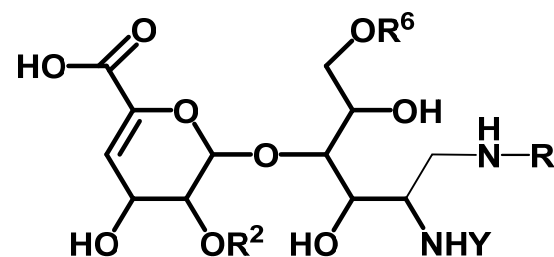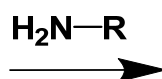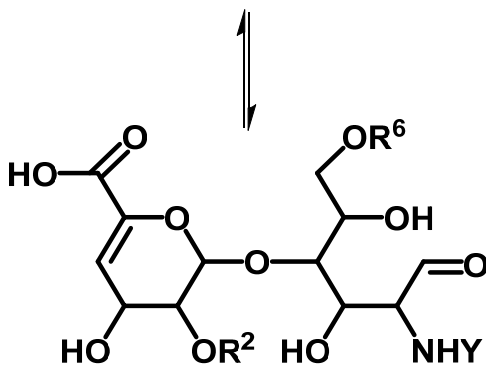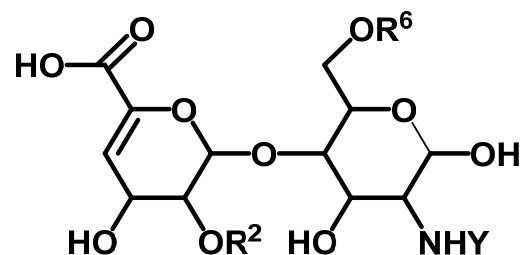
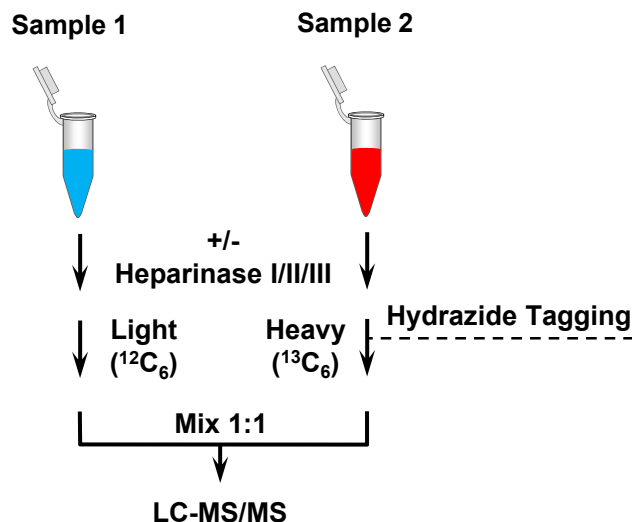# Separation of 2-AB labeled glycans using two HILIC columns
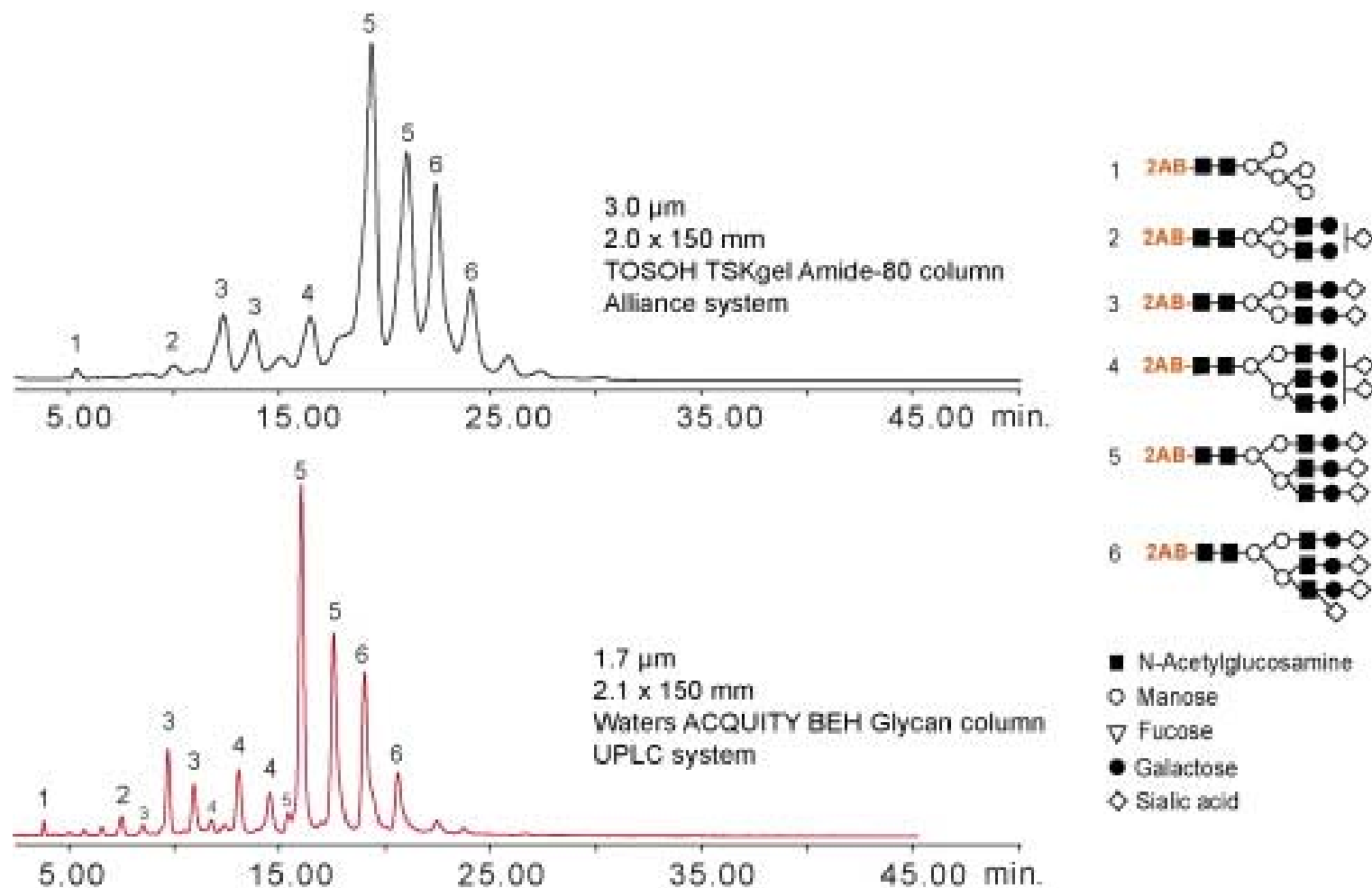
Fig. 3. Glycan separation comparison in 3.0 μm HPLC column and 1.7 μm sorbent. 10 pmol of fetuin 2-AB labeled glycans was separated using 3.0 μm TSKgel Amide-80 column in Alliance 2695 (top) and 1.7 μm BEH Glycan column in UPLC system (bottom). Sialylated biantennary and triantennary glycans including positional isomers were baseline resolved using 1.7m column in UPLC system in 45 min gradient time. The UPLC separation was done in gradient 65–55% B in 45 min at 0.5 mL/min using 2.1mm×150mm and the HPLC separation was done in 65–55% B in 50min at 0.45 mL/min using 2.0mm×150mm. The column temperature was at 40 °C on both runs. The peaks labeled with same numbers indicate the isomers.

Hydrazide derivatization of DS for relative quantitation

Comparative Analysis:
- **Unfractionated Heparin: Lot 1 vs. Lot 2**
- **Low-molecular weight heparin: Trademark vs. Generic**
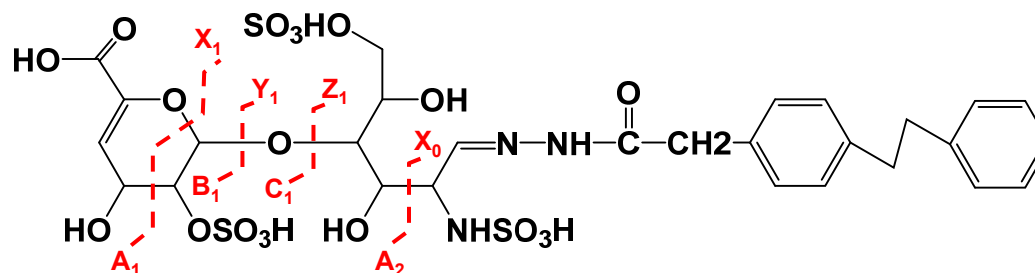- **Heparan sulfate: Healthy vs. Disease Tissue**
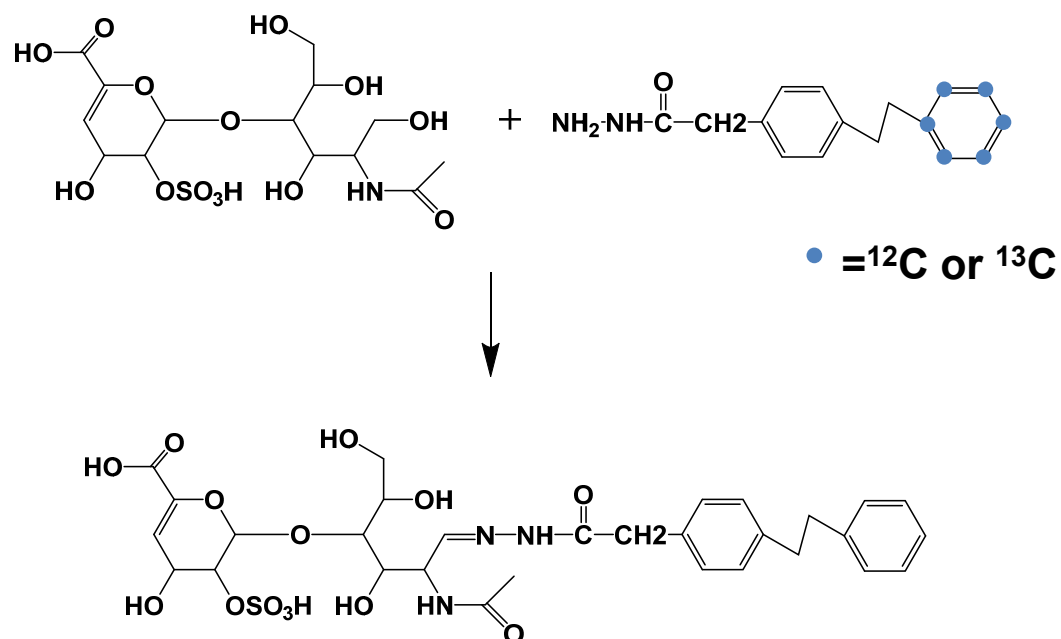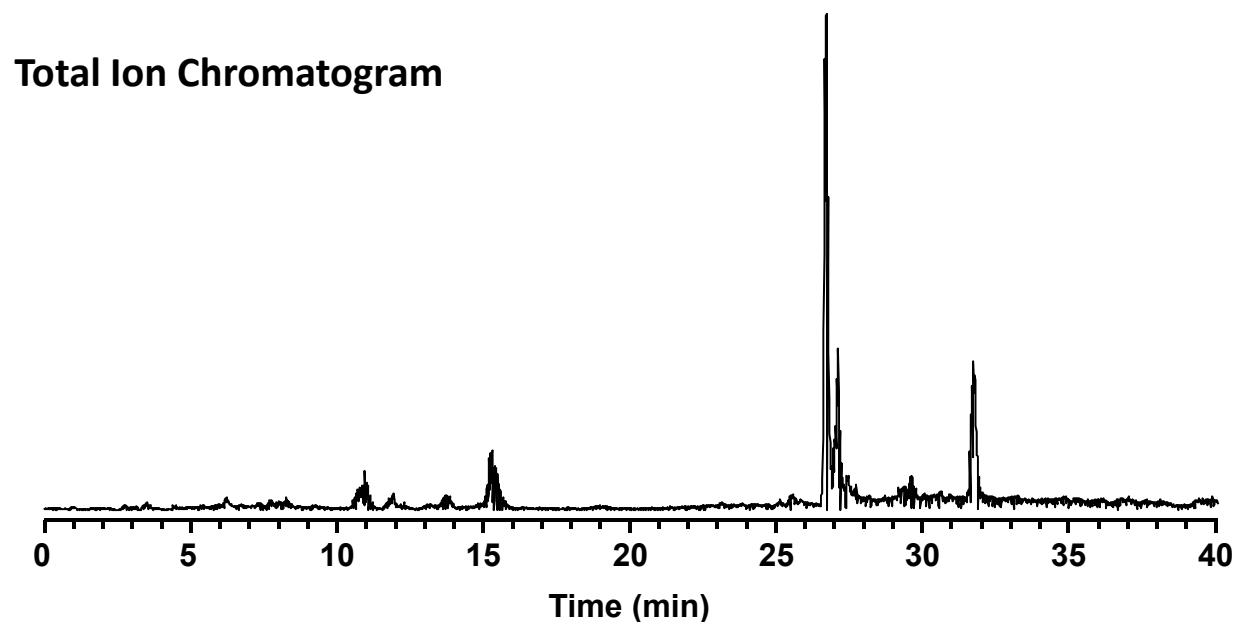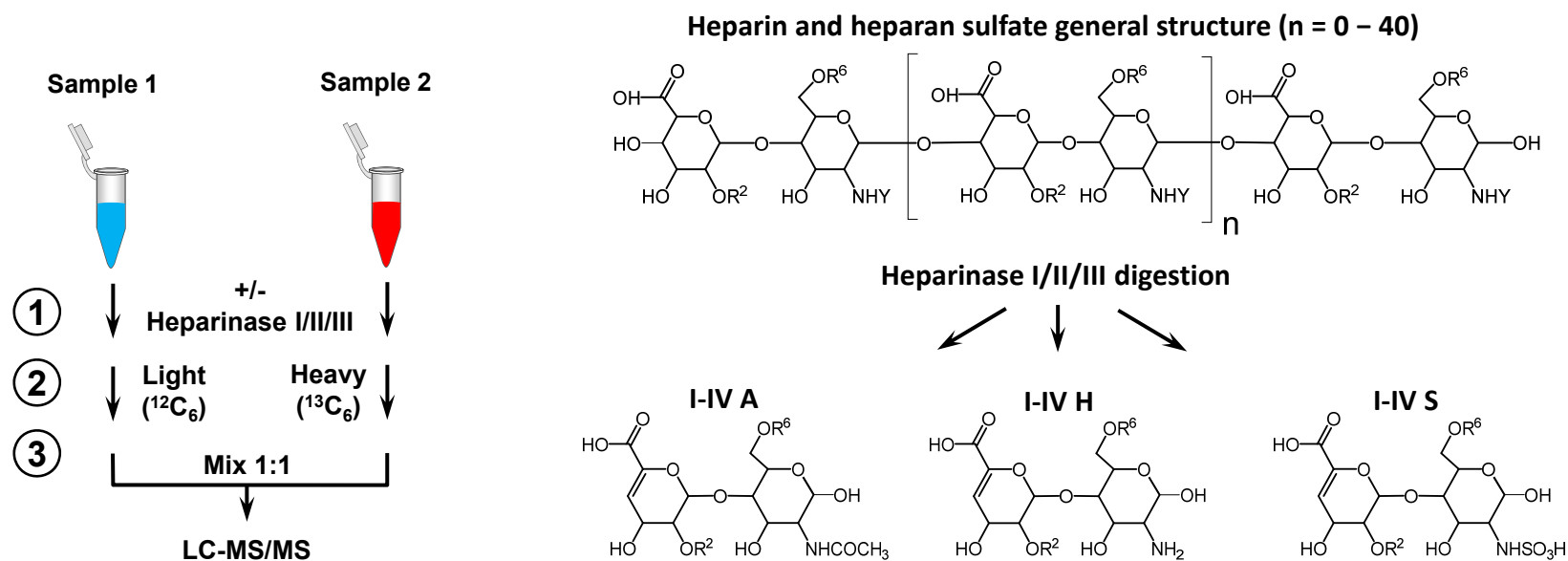
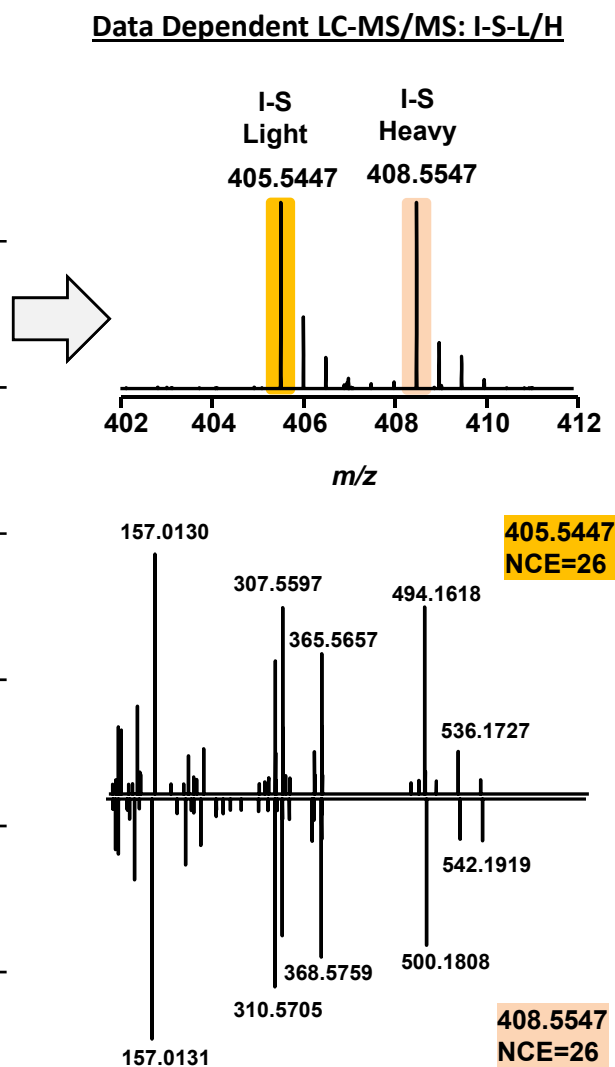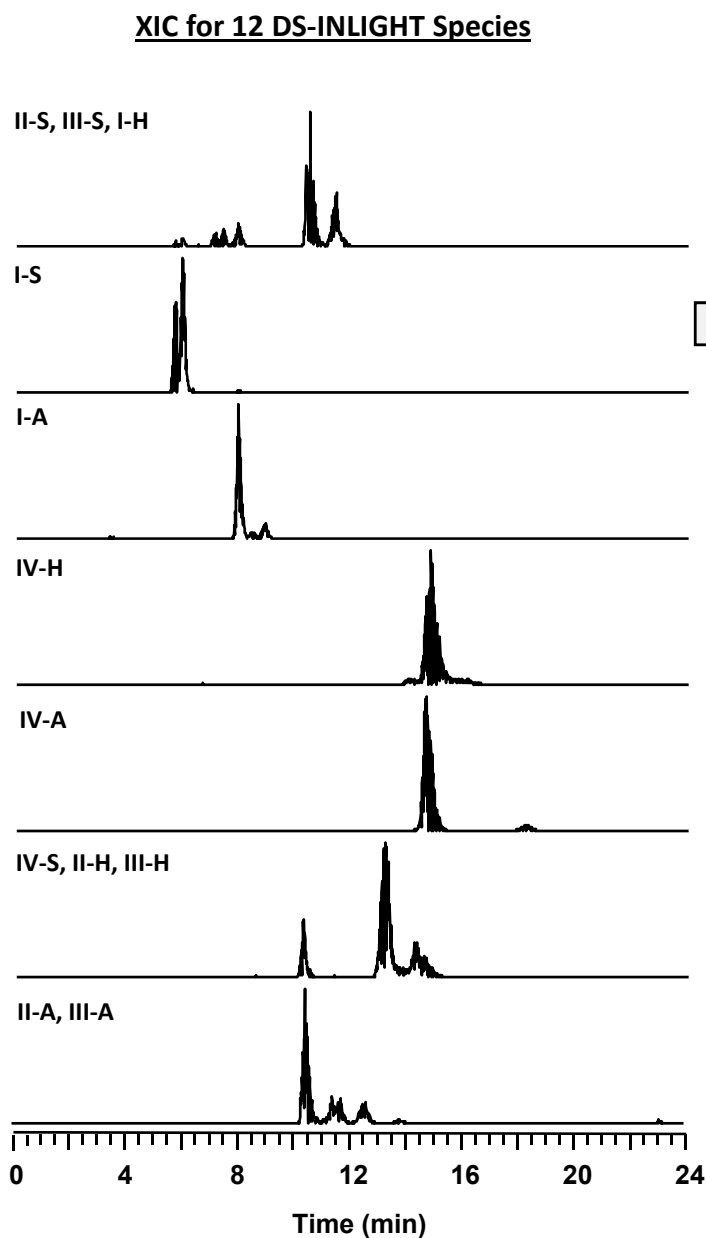# Representative ESI-MS/MS data for INLIGHT-tagged DS-IS
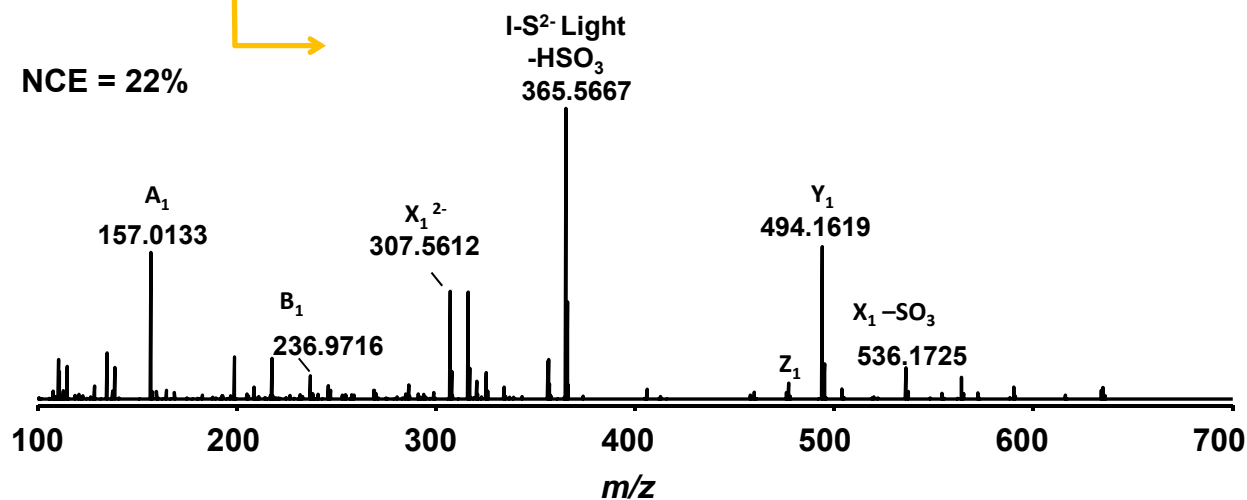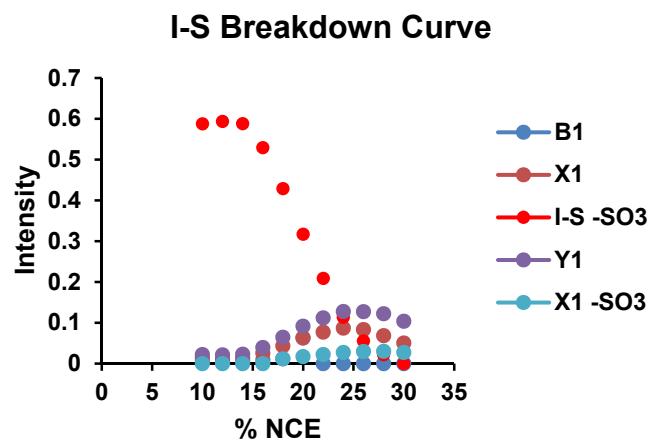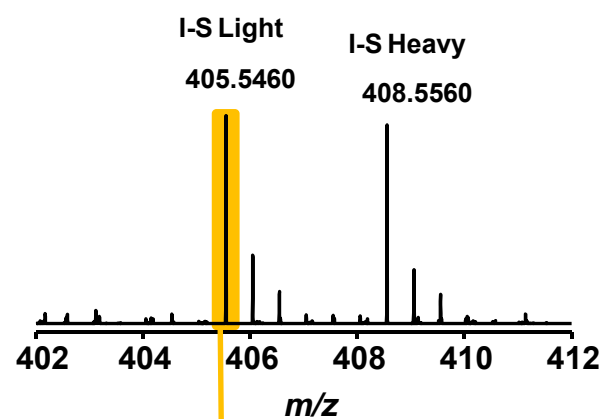
ESI-MS/MS of INLIGHT-tagged octasaccharide (DP8)
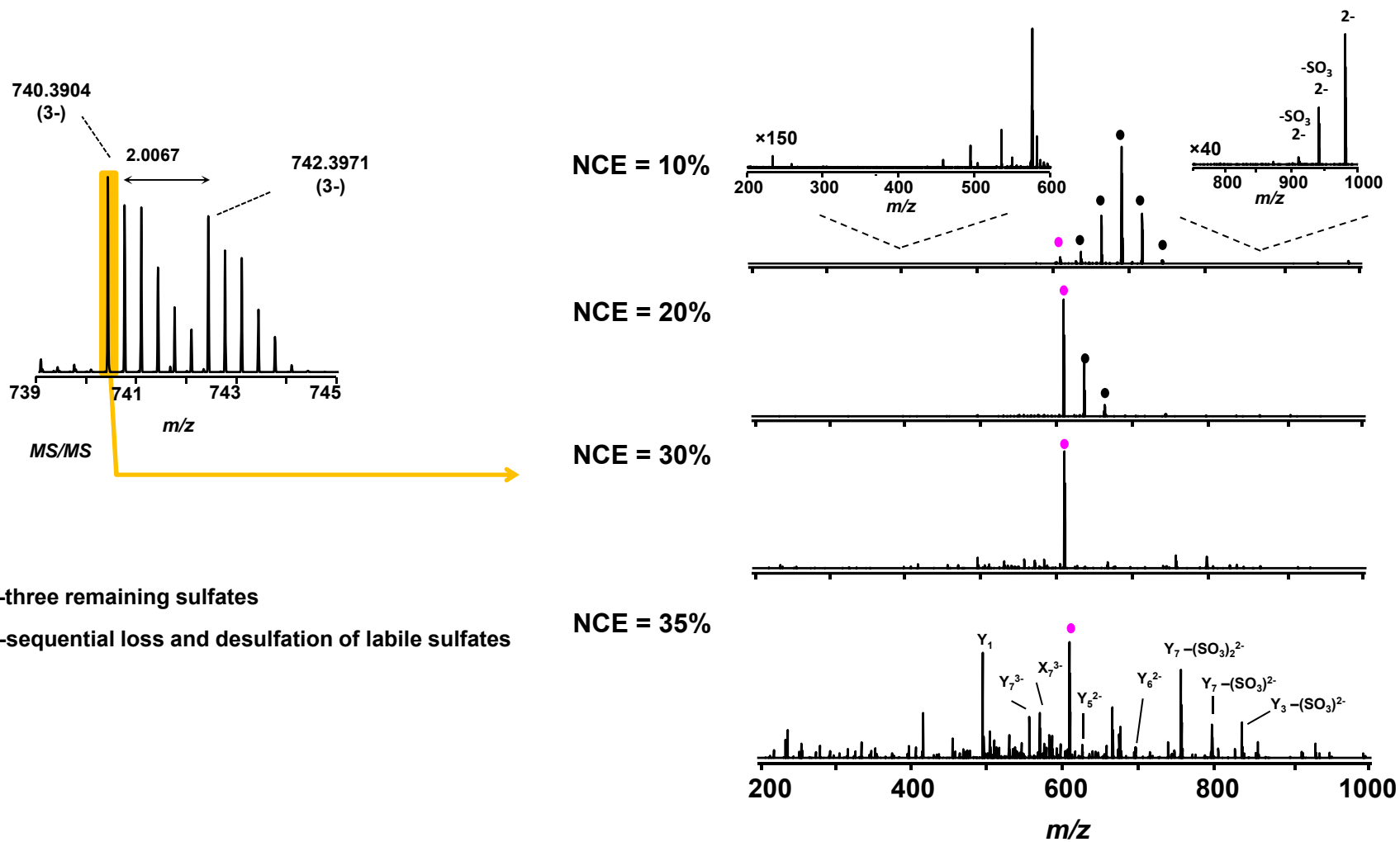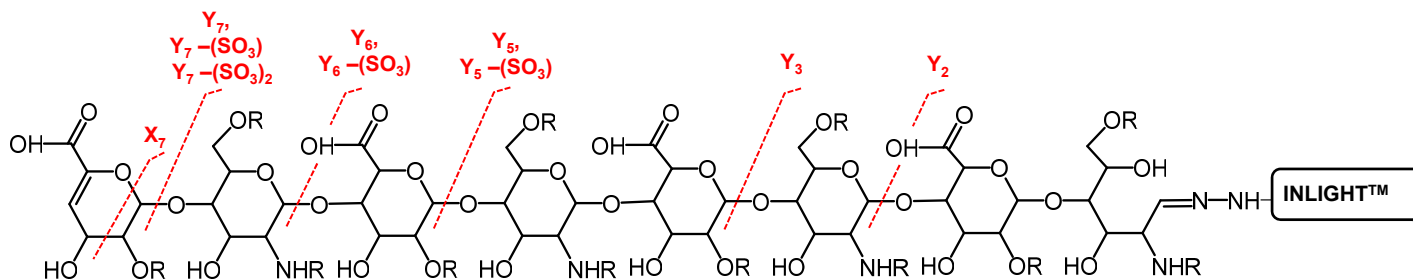
# INLIGHT derivatization of digested heparin for RP LC-MS/MS analysis

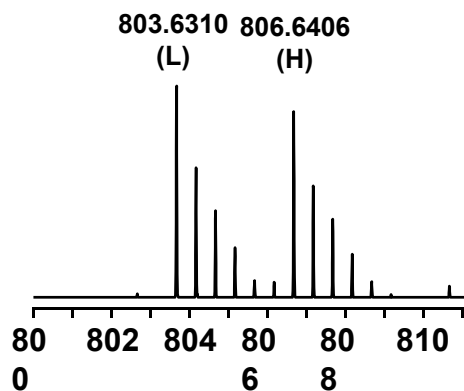# 12 DS extracted ion chromatograms for INLIGHT-tagged Heparin digest

XIC for 12 DS-INLIGHT Species

Data Dependent LC-MS/MS: I-S-L/H

Additional higher molecular weight INLIGHT-tagged heparin oligomers

Selected Extracted Ion Chromatograms and High MW Oligomers

803.6294 (2-) XIC
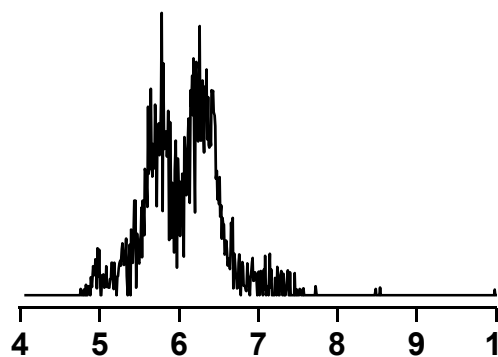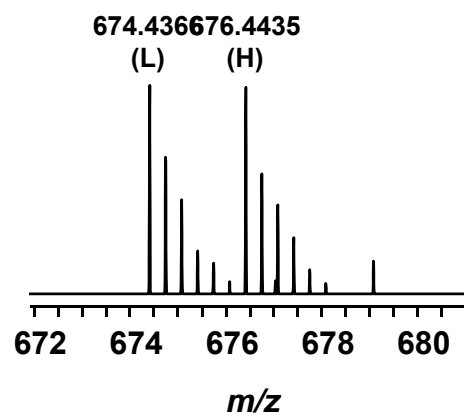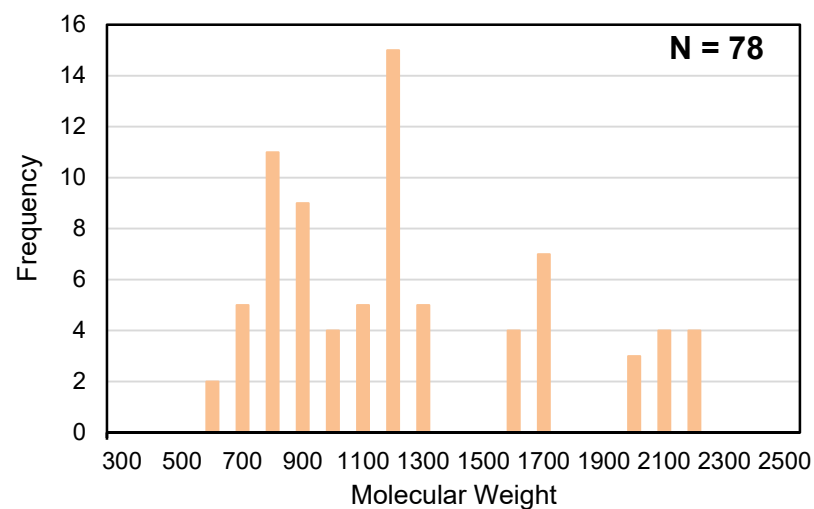
"DP6"

803.6310 (L)   806.6406 (H)

674.4366 (3-) XIC

"DP8"

674.4366 (L)   676.4435 (H)
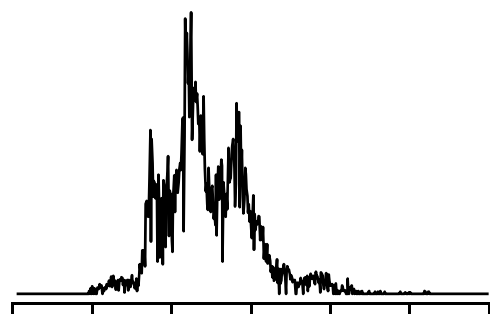
Time (min)

m/z

Molecular Weight Distribution of Detected INLIGHT-tagged Depolymerized Heparin

N = 78

Molecular Weight

Summary Points

- The field of glycoanalytical mass spectrometry continues to expand with the development new tools and approaches.

- Reverse phase and HILIC stationary phases will remain the most attractive stationary phases because they are compatible with LC-MS/MS measurements (i.e., ESI).

- Improved separation efficiencies of glycans can be achieved with chemical derivatization of the reducing ends.

- Relative quantification of glycans from different samples via chemical derivatization will also remain important until synthetic carbohydrate chemists can produce micro-milligram quantities of pure samples to serve as internal standards (this is not expected to be a reality in the near term).

- Ultra-high pressure liquid chromatography (UPLC) will continue to replace traditional HPLC provides better peak shape and higher separation efficiency.

Summary Points (cont.)

- **New gas-phase dissociation techniques are in development for comprehensive characterization of glycopeptides *without deglycosylation.***
  - *Traditional collision induced dissociation (CID) combined with electron transfer dissociation (ETD) on Orbitrap-based systems could be a 'game changer'.*
  - *The Orbitrap-based system from ThermoFisher Scientific that can perform these types of analyses is called the Lumos and it is gaining popularity among pharmaceutical companies who are developing biologics.*

- **Another approach which may impact glycan analysis is ion mobility.**
  - Essentially is like HPLC but in the gas phase
  - Can separate isomeric species for simple standards
  - Typically coupled to Q-TOF systems given the need for high duty cycles.

- **Stay Tuned!**