Mass Spectrometry and Proteomics

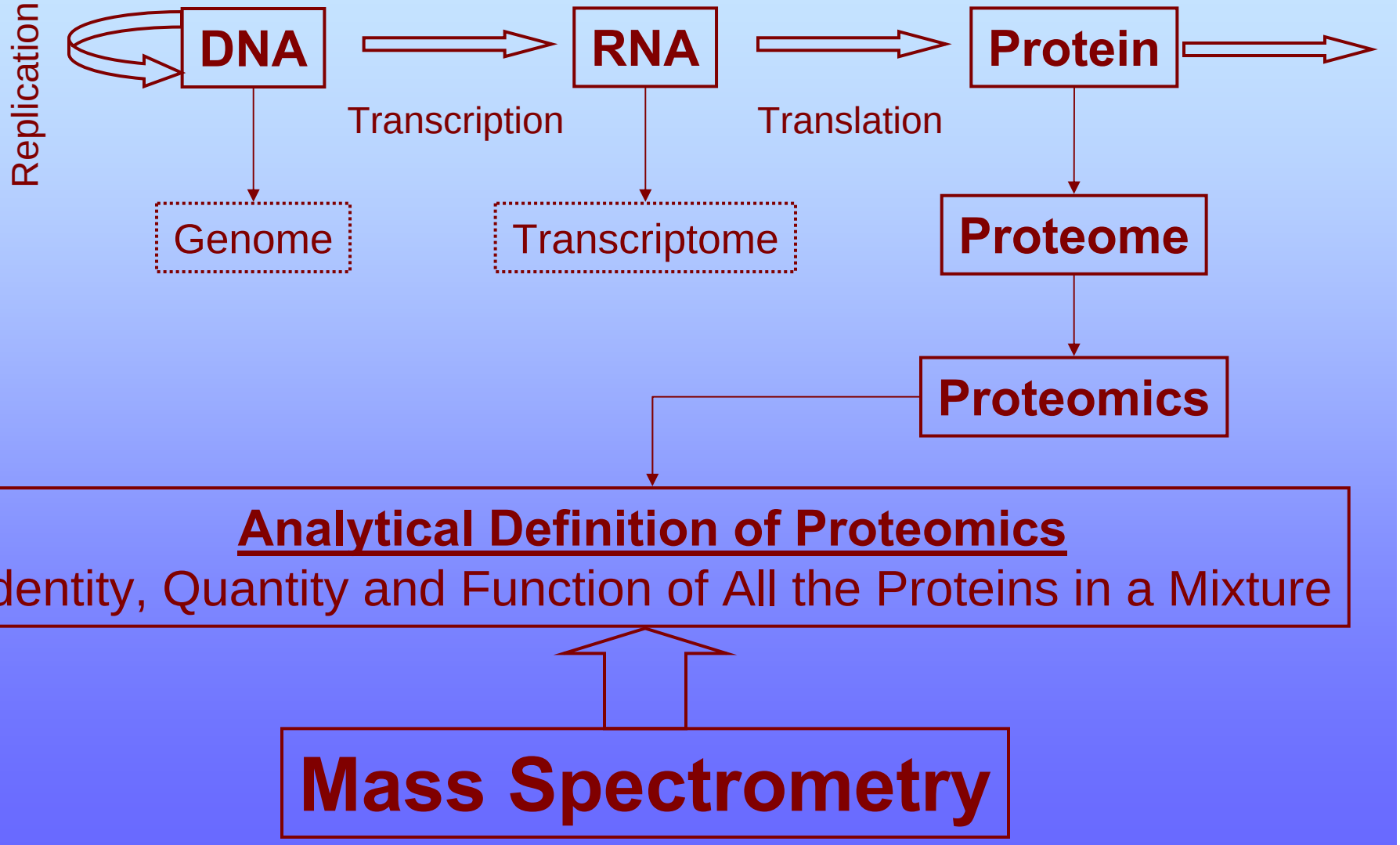
Professor Xudong Yao

Bioanalytical Chemistry

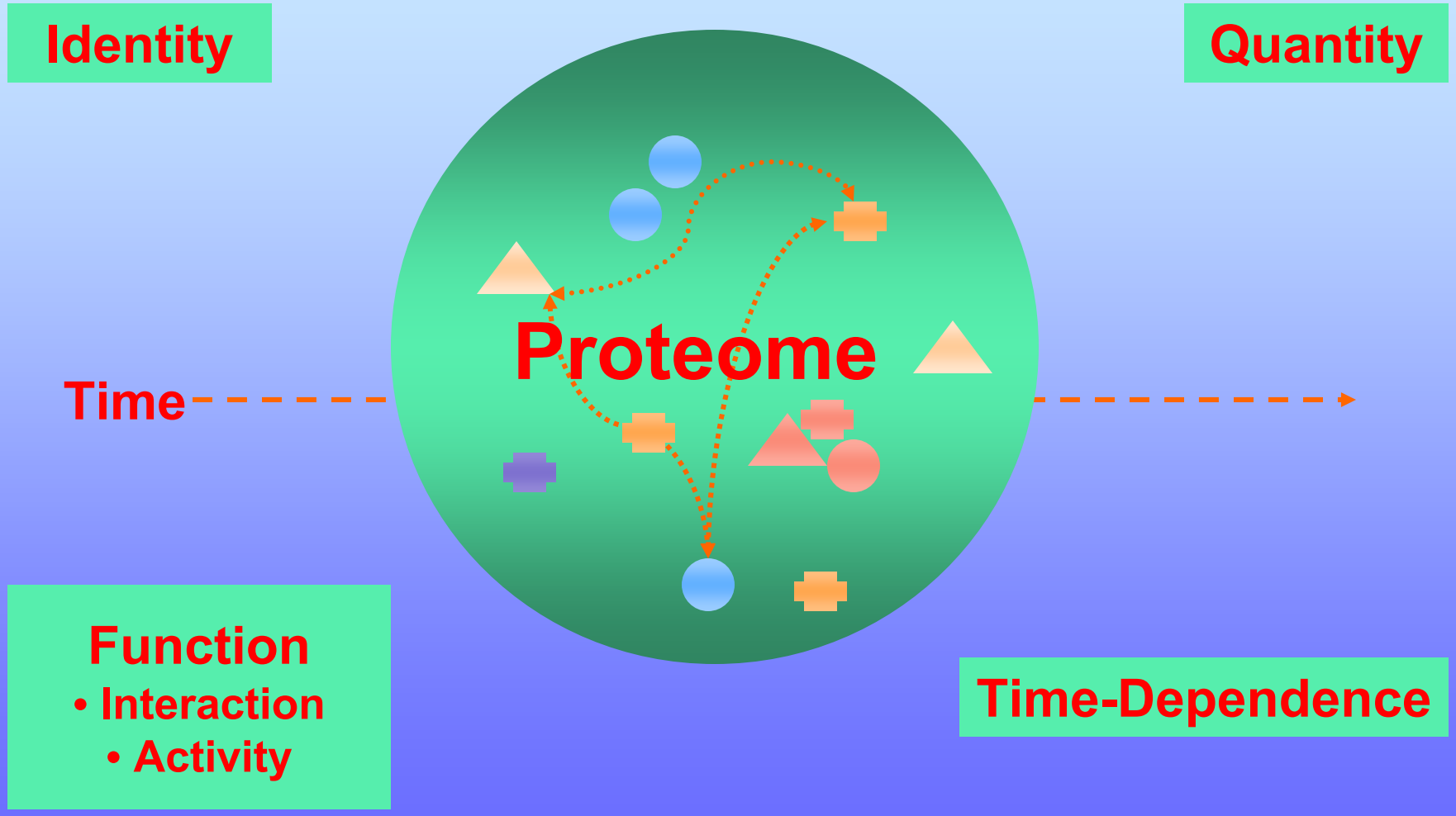
Spring 2007

- Proteomics and “-omics”
- Roles of mass spectrometry
- Comparative proteomics
- Chemical proteomics

Protein, Proteome and Systems Biology



Objectives of Proteomics



Analytical Challenges in Post-genome Research

- Sample complexity
  - “Peak capacity”
  - Multi-dimensional separation
- Collective analysis
  - Not traditional, one-by-one analysis
  - Sensitive, specific and quantitative
  - and the answers is …
- Data treatment, analysis and achieving
  - Hardware
  - Software
- Researchers of multi-disciplinary training

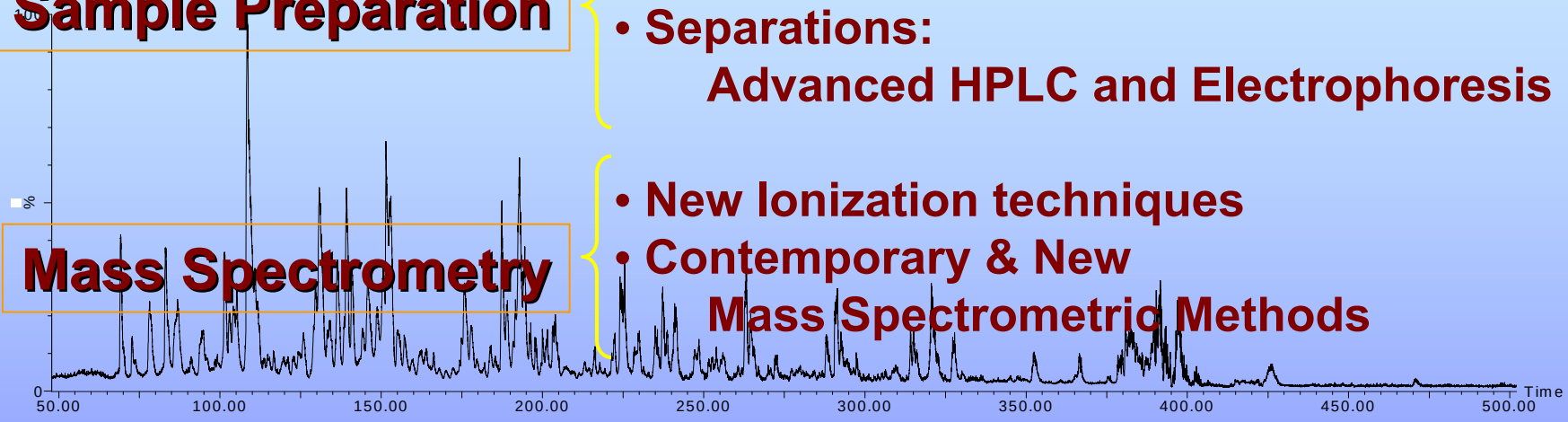
A Typical Proteomics Experiment

Sample Preparation

- Protein/Peptide Chemistry
- Enzymology
- Separations:
  Advanced HPLC and Electrophoresis

Mass Spectrometry

- New Ionization techniques
- Contemporary & New
  Mass Spectrometric Methods

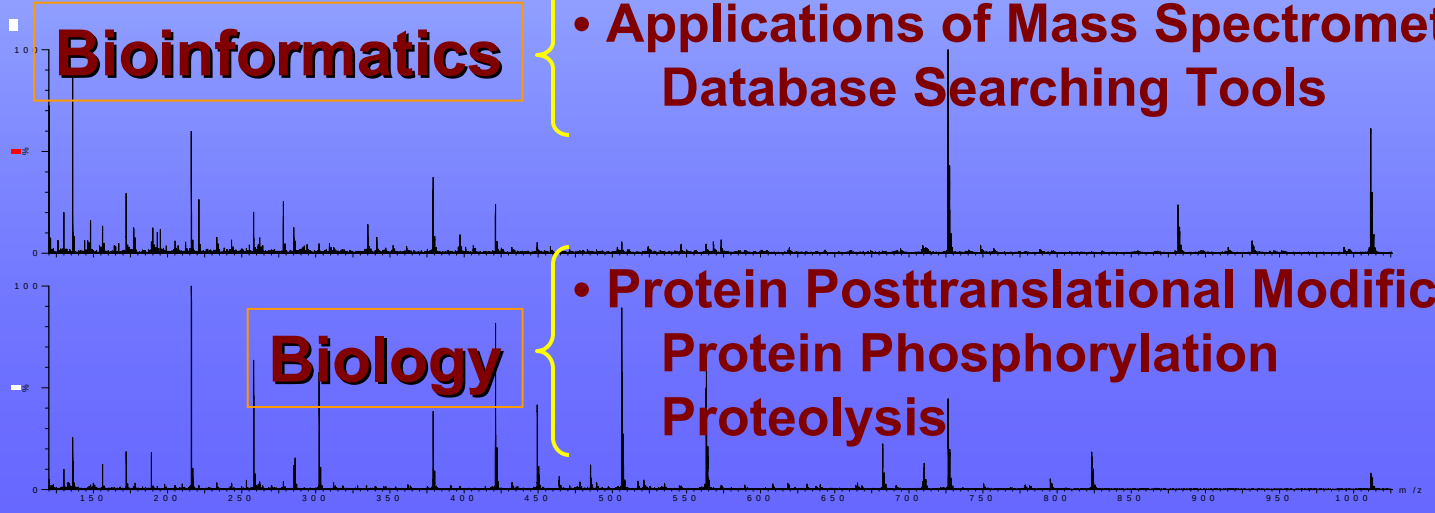


Bioinformatics

- Applications of Mass Spectrometry and
  Database Searching Tools

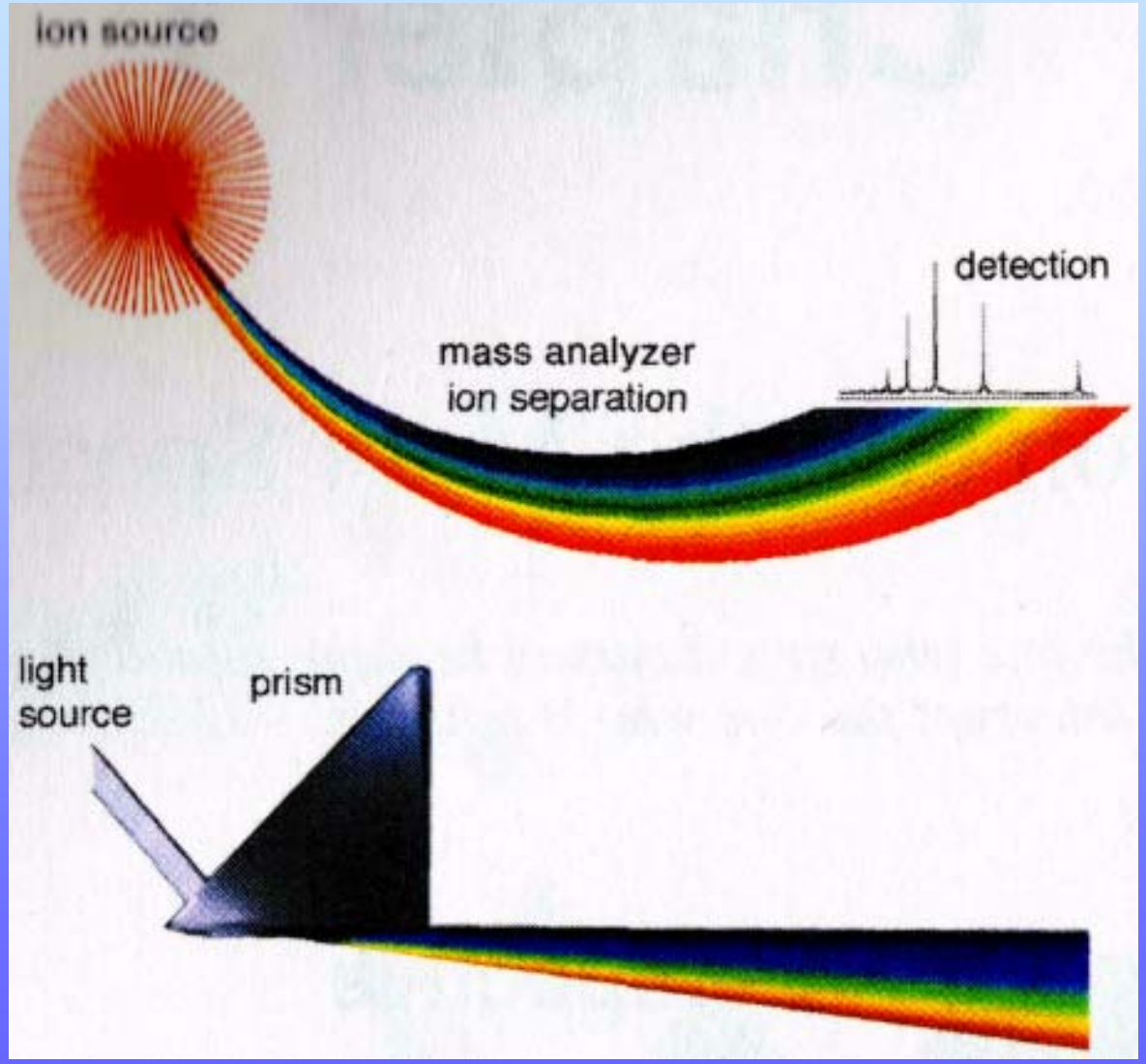
Biology

- Protein Posttranslational Modification
  Protein Phosphorylation
  Proteolysis



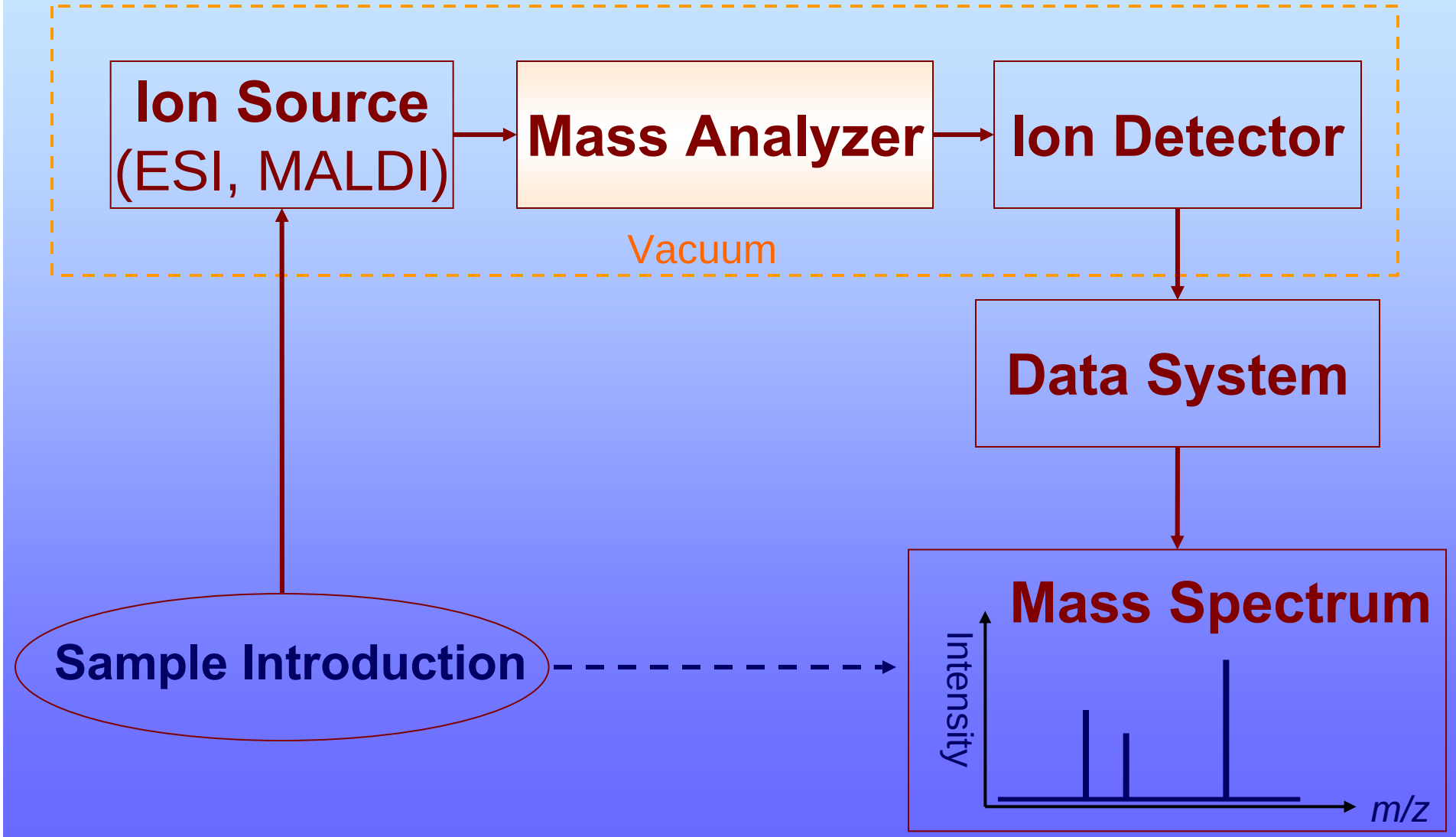
Roles of Mass Spectrometry

The mass analysis process as compared to the dispersion of light by a prism



Siuzdak, 2003

Mass Spectrometry



Mass Analyzer

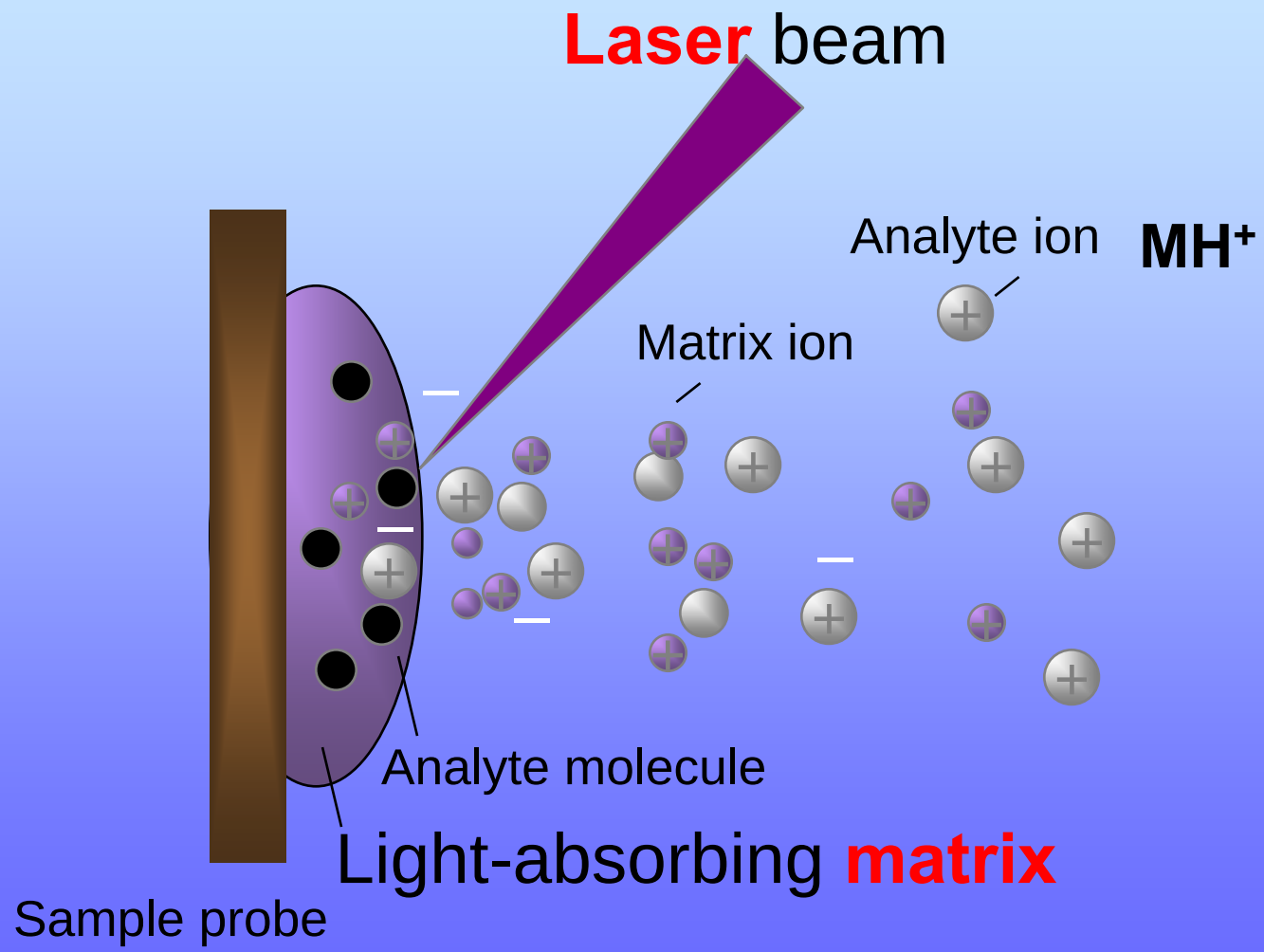
- Separate ions by mass/charge
- Common types
  - Quadrupole mass filter (Q), Time-of-Flight (TOF), Ion Trap (IT), Fourier Transform Ion Cyclotron Resonance (FTICR)
- Tandem mass spectrometry
  - Spatial, such as Q-q-TOF, TOF-TOF, Q-q-Q
  - Temporal, such as IT, FTICR
  - Spatial and Temporal, such as IT-FTICR, Q-q-FTICR, IT-TOF

Soft Ionizations of Biomolecules

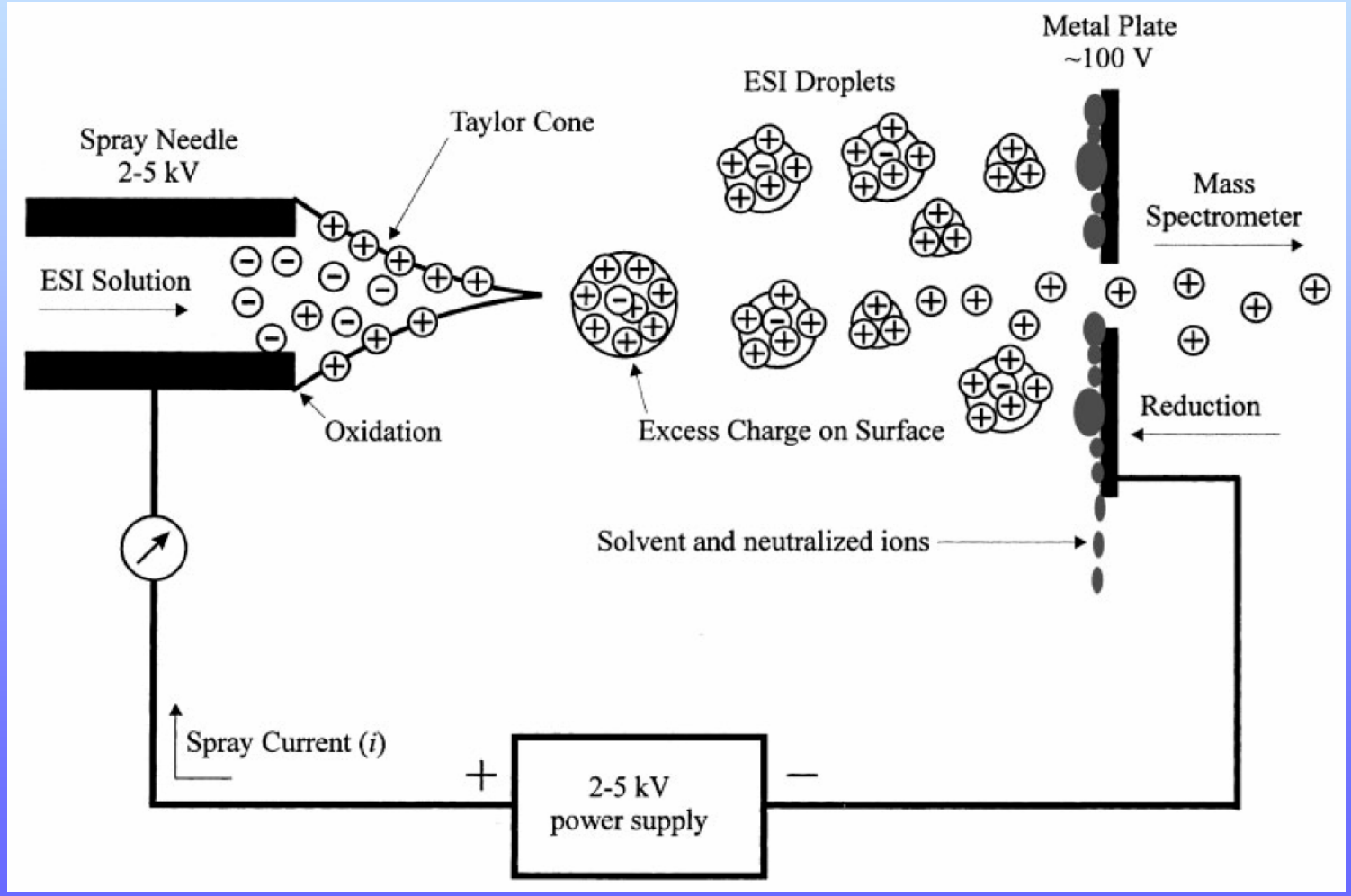
1. Matrix-Assisted Laser Desorption/Ionization (MALDI)
2. Electrospray Ionization (ESI)

Nobel Prize in 2002

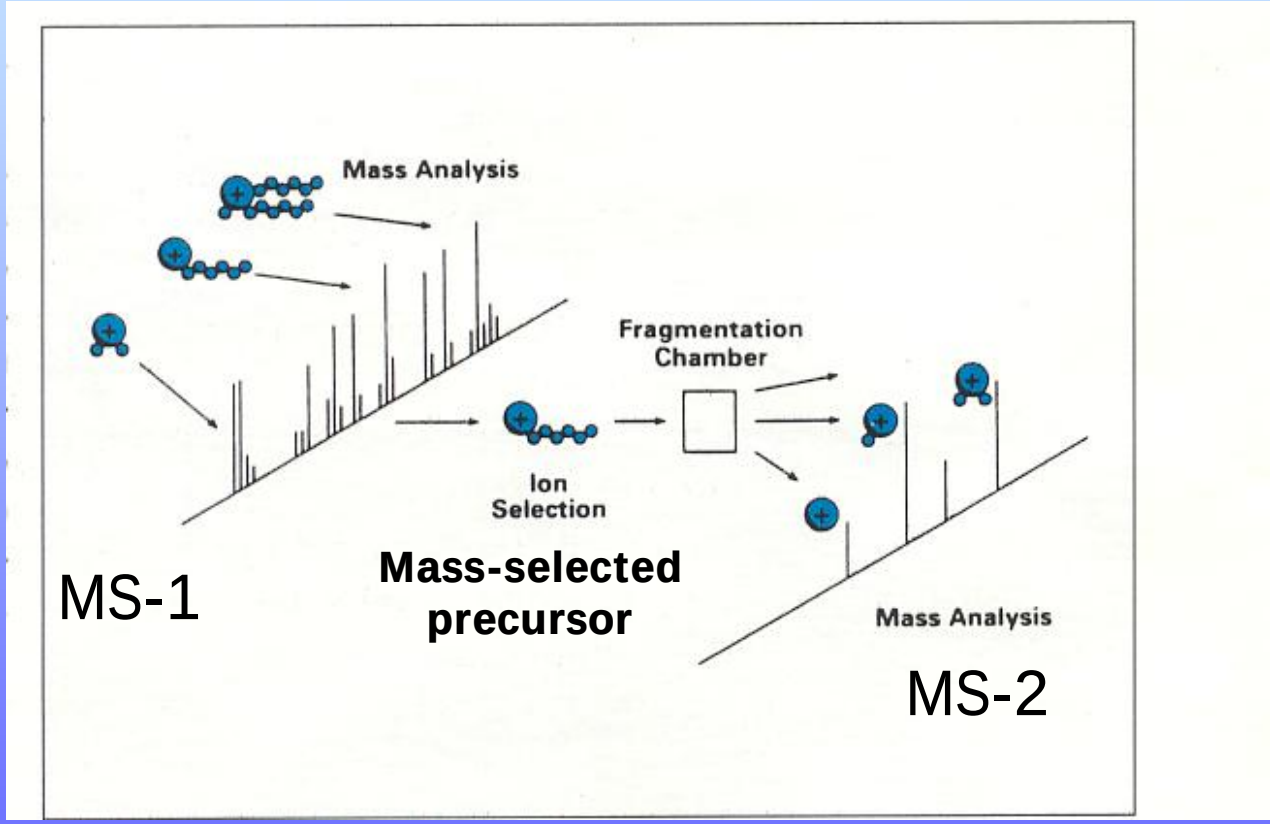
Matrix-Assisted Laser Desorption/Ionization (MALDI) Generates Singly-Charged Ions



Electrospray Ionization (ESI) Generates Multiply-Charged Ions



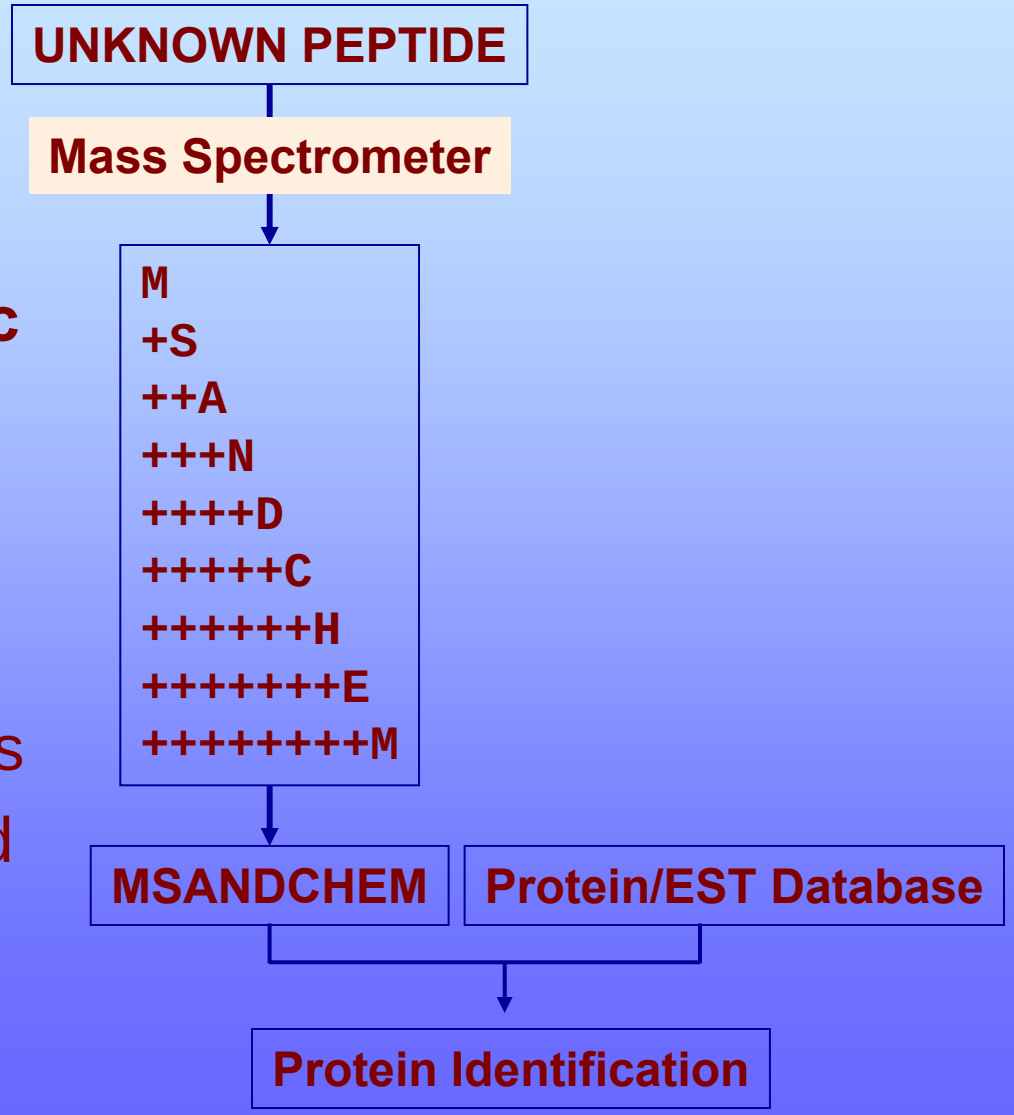
Tandem MS: Basic Concepts



Mass Spectrometry for Proteins and Peptides

- **Structural elucidation**
- **Quantitative analysis**
- **Sensitive and automatic mixture analysis**

- Sequencing proteins
- Identification of post-translational modifications
- High-order structures and dynamics of proteins and protein dynamics

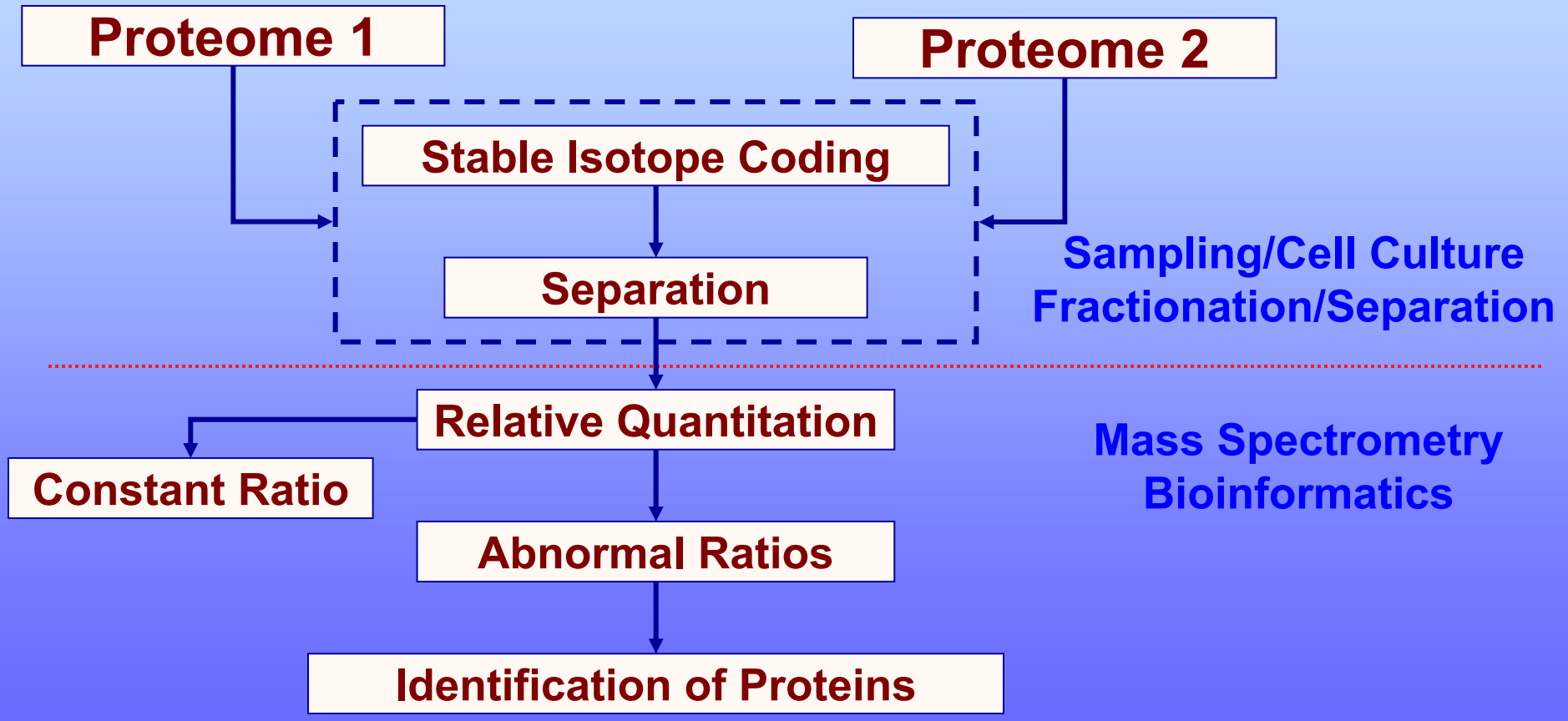


Comparative Proteomics

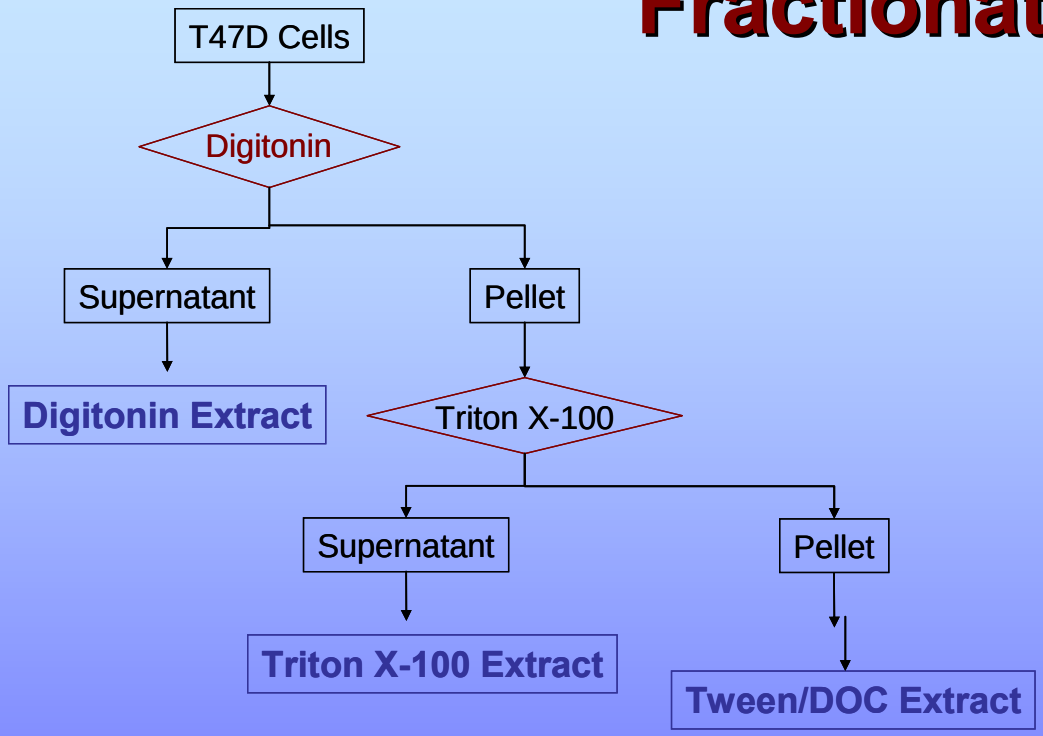
Gel or Non-gel
Label or Non-label

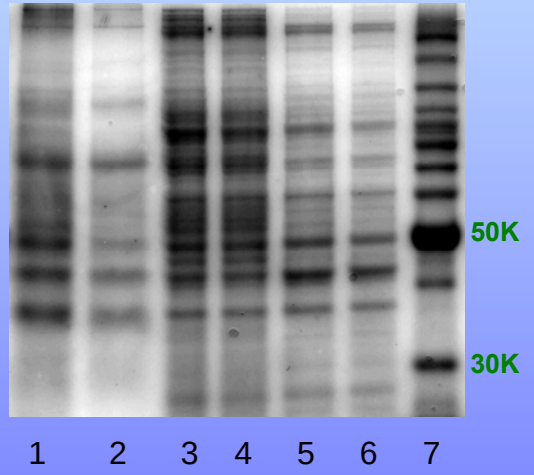
Comparative Proteomics

Relative Changes
in proteins including concentration and composition



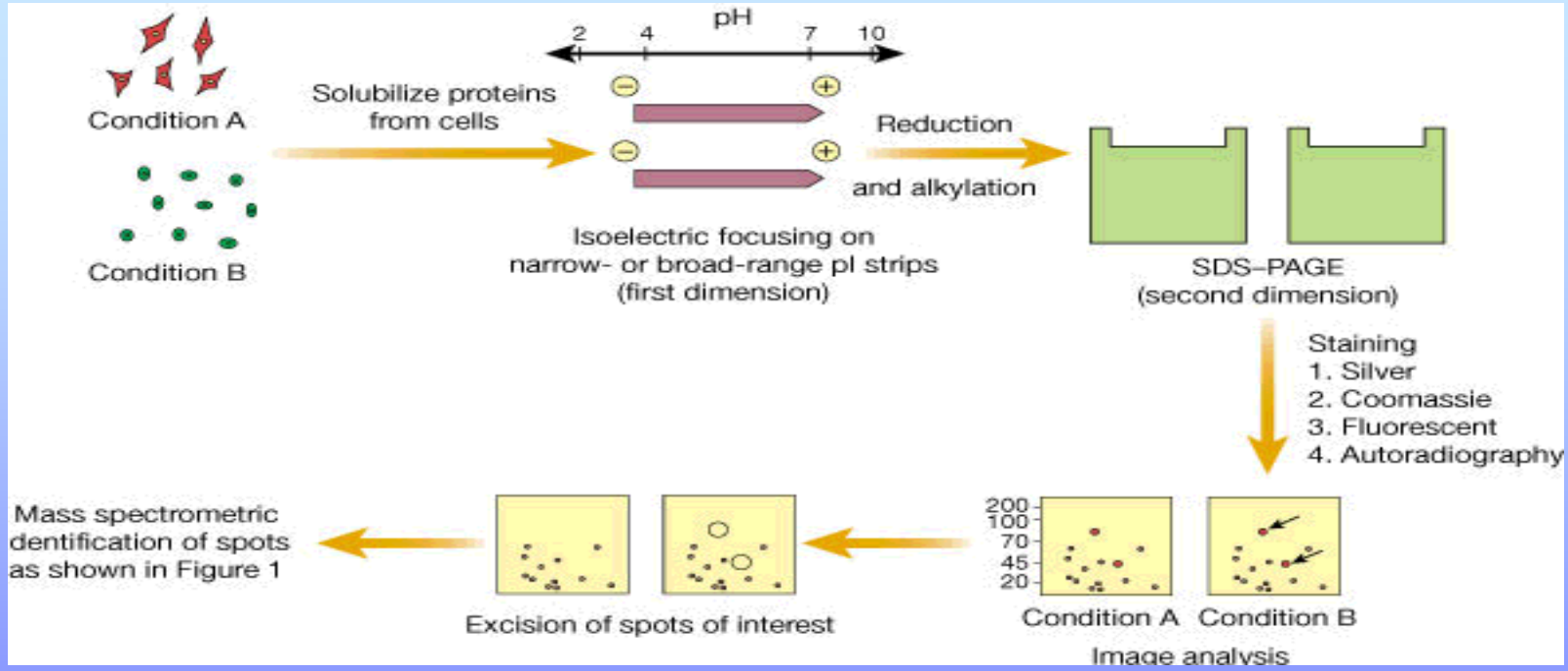
Reduce sample complexity at the protein level: DDF (Differential Detergent Fractionation)





1. Tween/DOC Extraction (Nuclear Fraction) (Prolactin)
2. Tween/DOC Extraction (Nuclear Fraction) (Control)
3. Triton X-100 Extraction (Membrane/Organelle Fraction) (Prolactin)
4. Triton X-100 Extraction (Membrane/Organelle Fraction) (Control)
5. Digitonin Extraction (Cytosolic Fraction) (Prolactin)
6. Digitonin Extraction (Cytosolic Fraction) (Control)
7. Protein Ladder

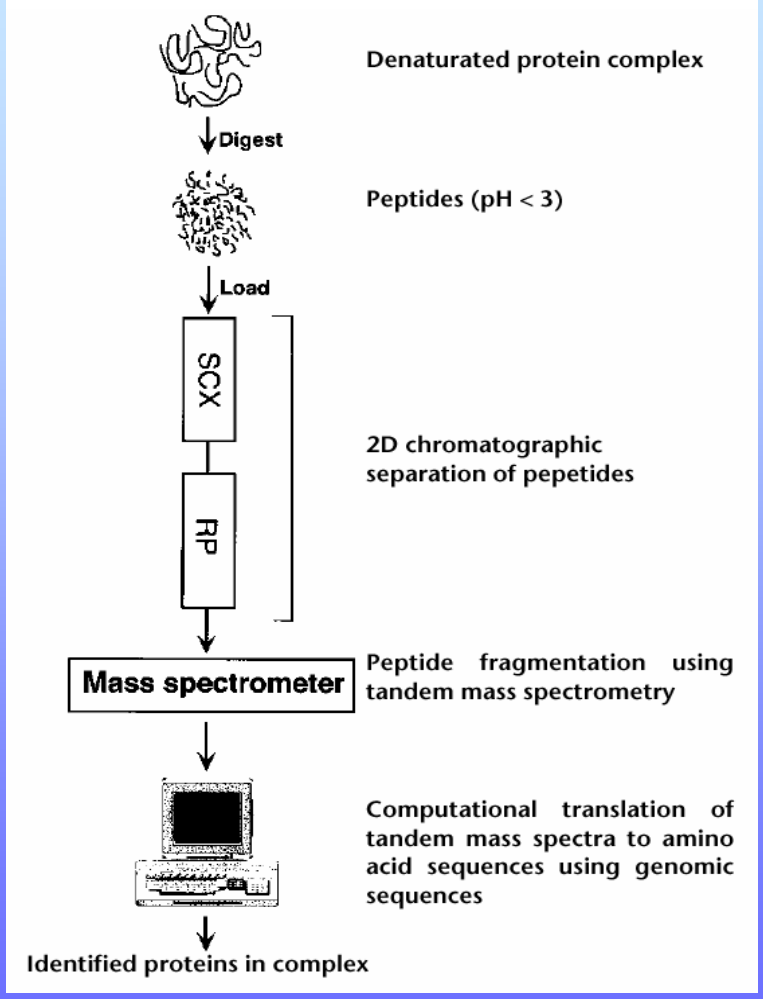
GEL: 2-Dimensional Electrophoresis (2-DE)



A Pandey; M Mann, *Nat. Biotechnol.* **2000**

- **Difficulties with: extreme pI proteins, low abundance, proteins, hydrophobic proteins**
- **Inefficient in-gel digestion of proteins for MS analysis**
- **Labor extensive**

NON-GEL: MudPIT Analysis of Protein Complex

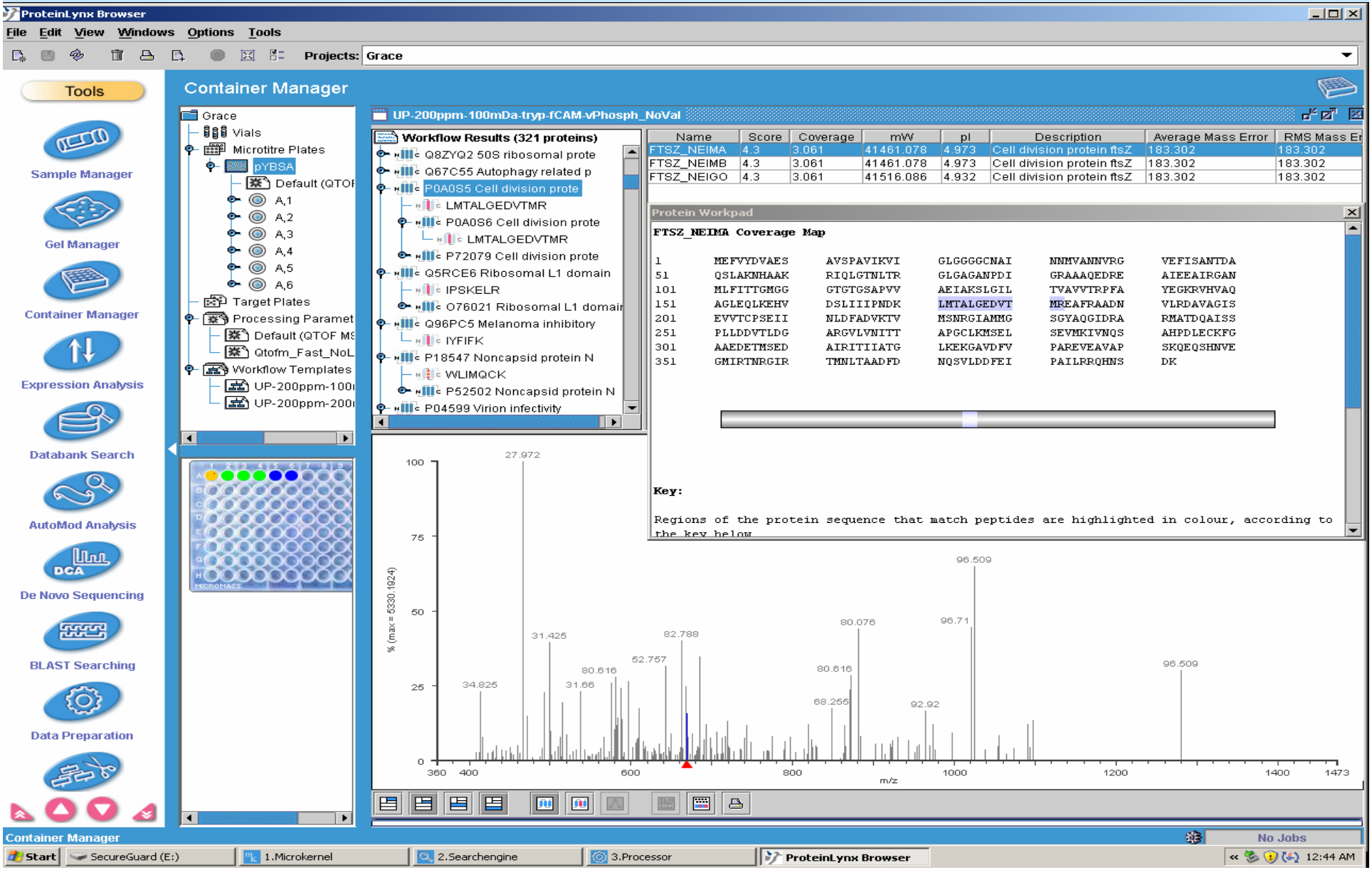


- Easier separation
- Easier automation
- "Problematic" proteins
  - Small, large, hydrophobic, low abundance
- Easier sample preparation for MS analysis

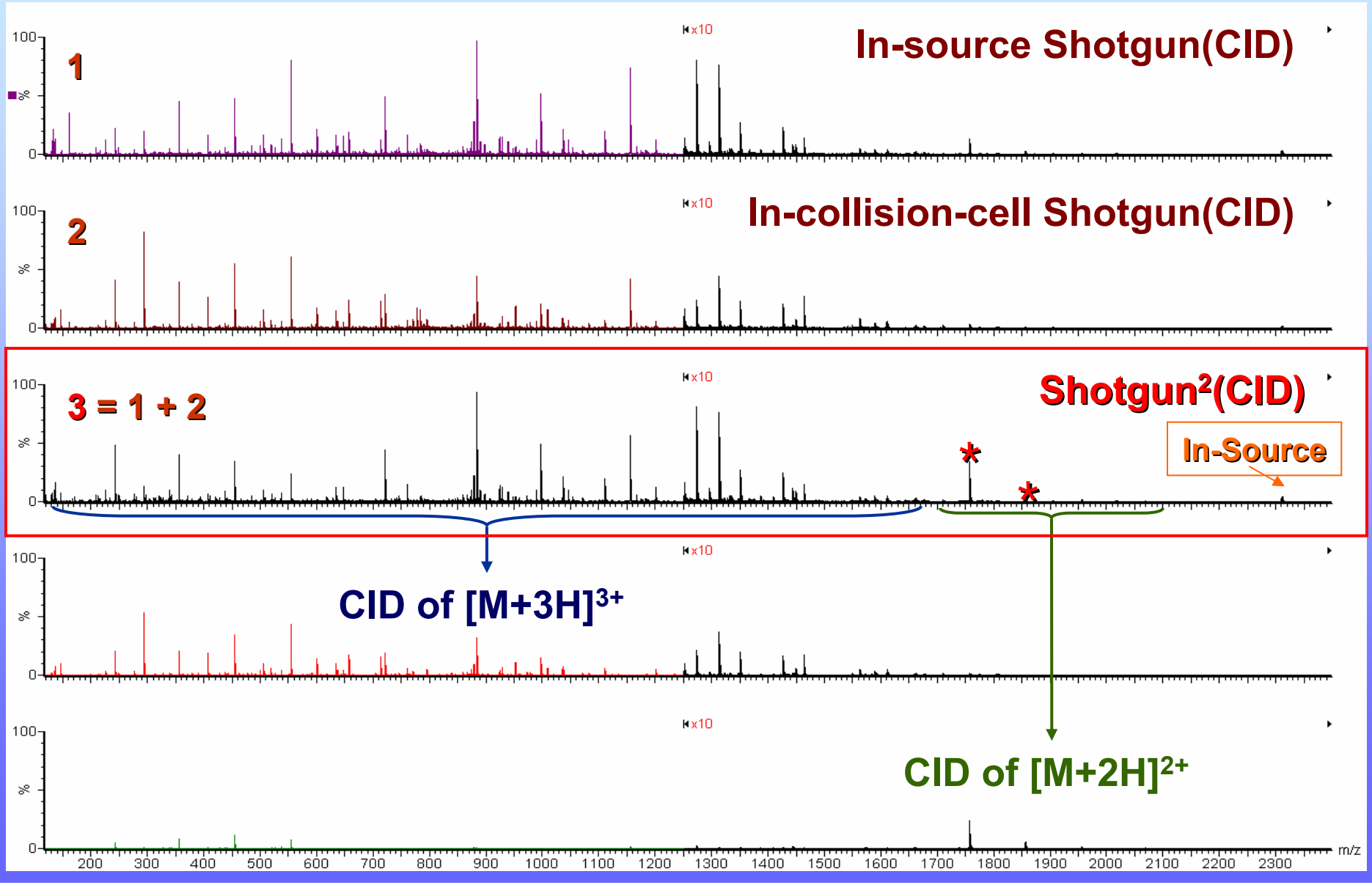
- Computational capacity
- **Quantitative capability**

Link *et al. Nat. Biotechnol.* **1999**

A Glance at Proteomic Bioinformatics

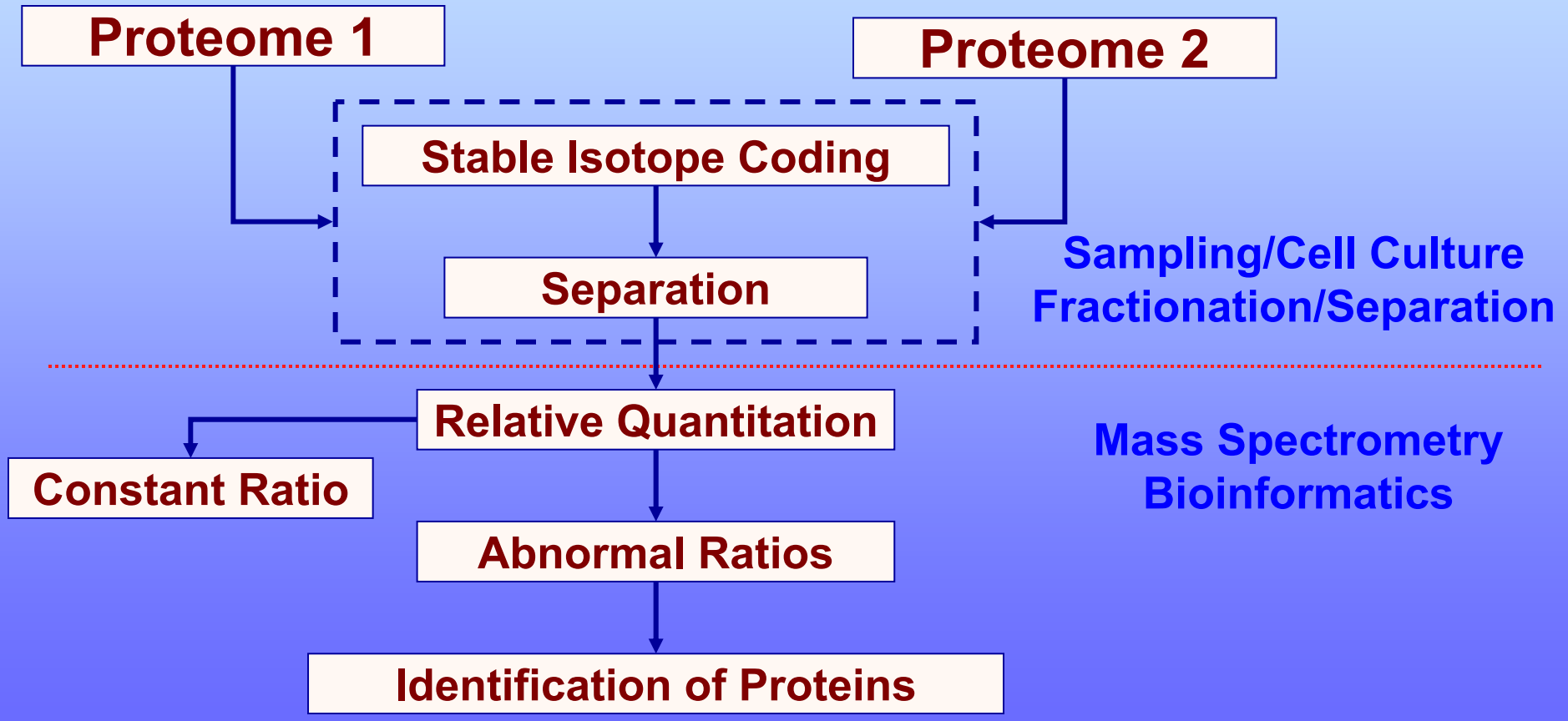


Shotgun^2 Collision Induced Dissociation Mass Spectrometry



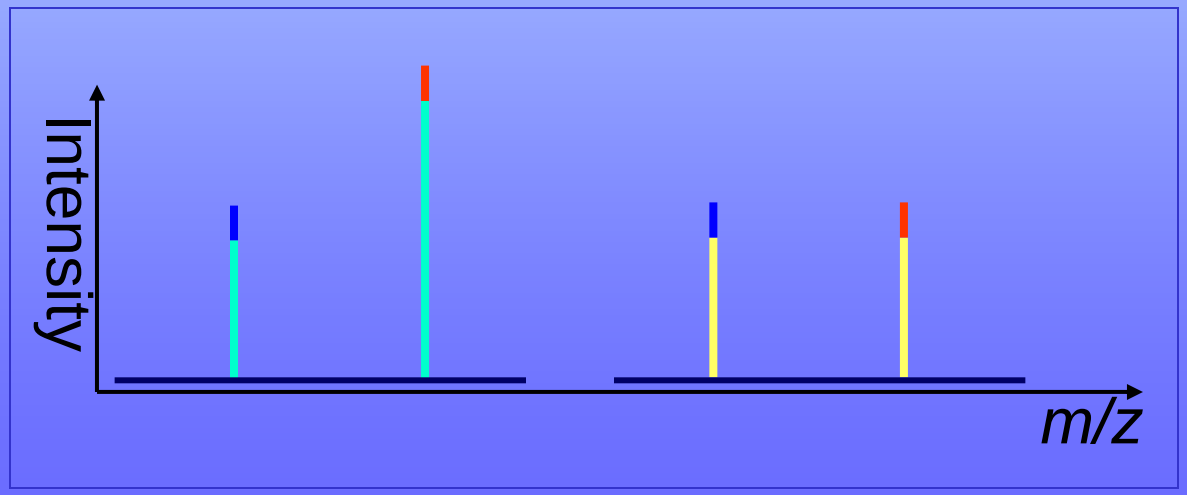
Comparative Proteomics

Relative Changes
in proteins including concentration and composition



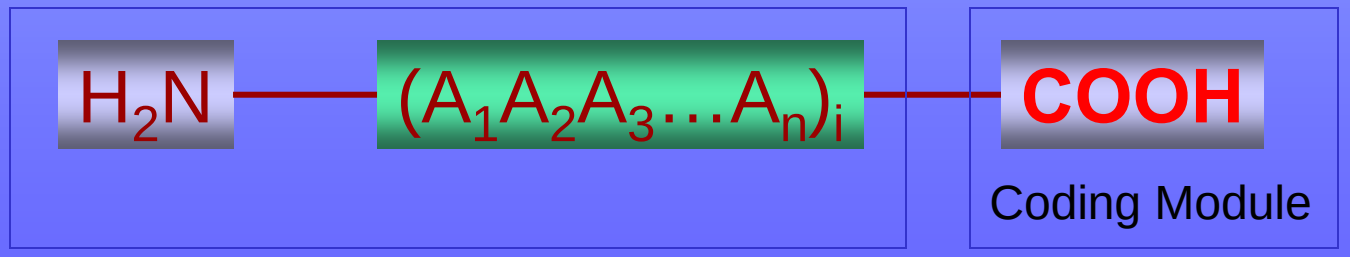
Label: MS-Based Relative Quantitation

- Large differences in concentration
  - Direct ESI/MALDI MS
- Small differences in protein concentration
  - Stable isotope dilution: inherent choice

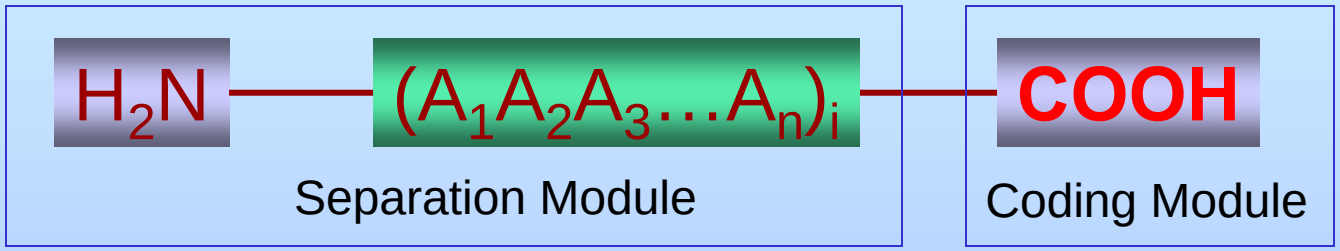


Introduction of Stable Isotopes

- Criteria for isotope internal standards
  - Ideally behaving the same before mass analysis
- Metabolic labeling during biosynthesis/bioprocess
- Post-biosynthesis/bioprocess labeling: chemical and enzymatic
  - Functional groups on side chains: -SH, $-OPO_3H_3$
  - Termini: N-terminal, **C-terminal**
  - Active/Binding sites: Chemical Proteomics



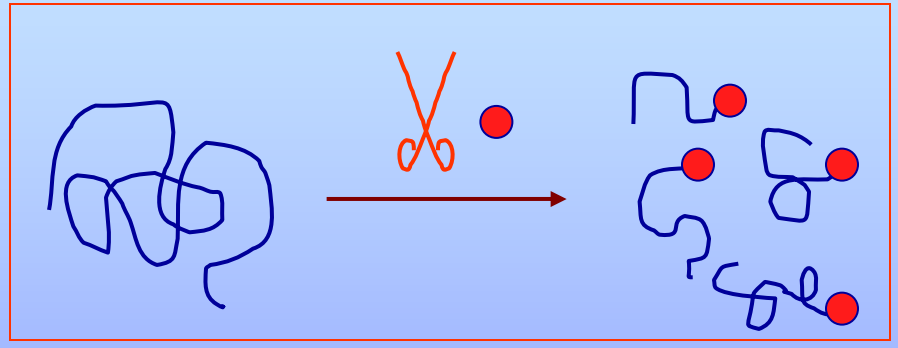
Advantages of Modular Design

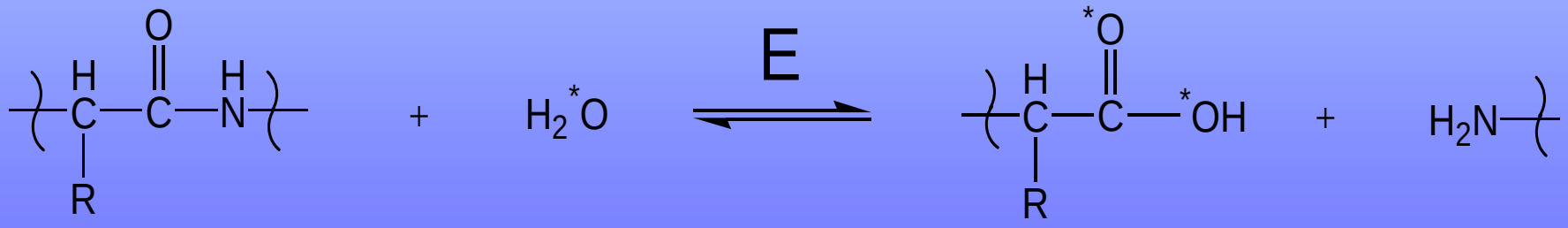


- **Isotope Coding**
  - **Universal**
    - Important to small proteins
  - **Specific**
  - **Efficient**
  - **Minimal Structural Modification**
    - Chromatographic co-elution
  - **Stable during separation**

- **Separation**
  - **Portable to all separation platforms, including affinity separation**

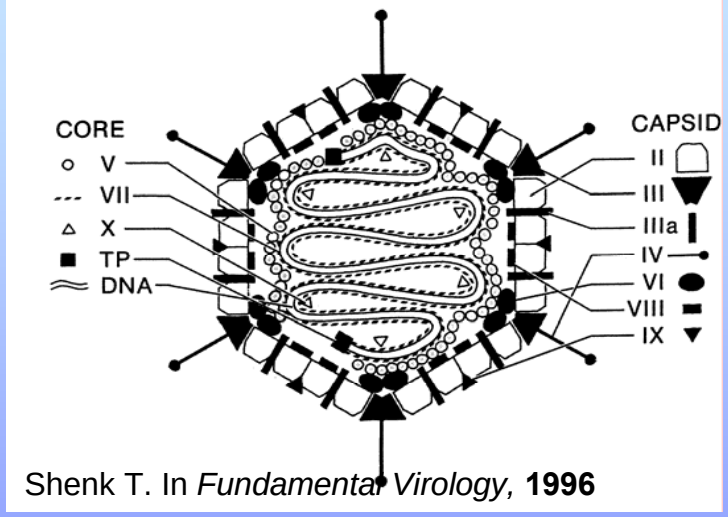
Tryptic Incorporation of Two ^{18}O Atoms into Peptide C-Terminus



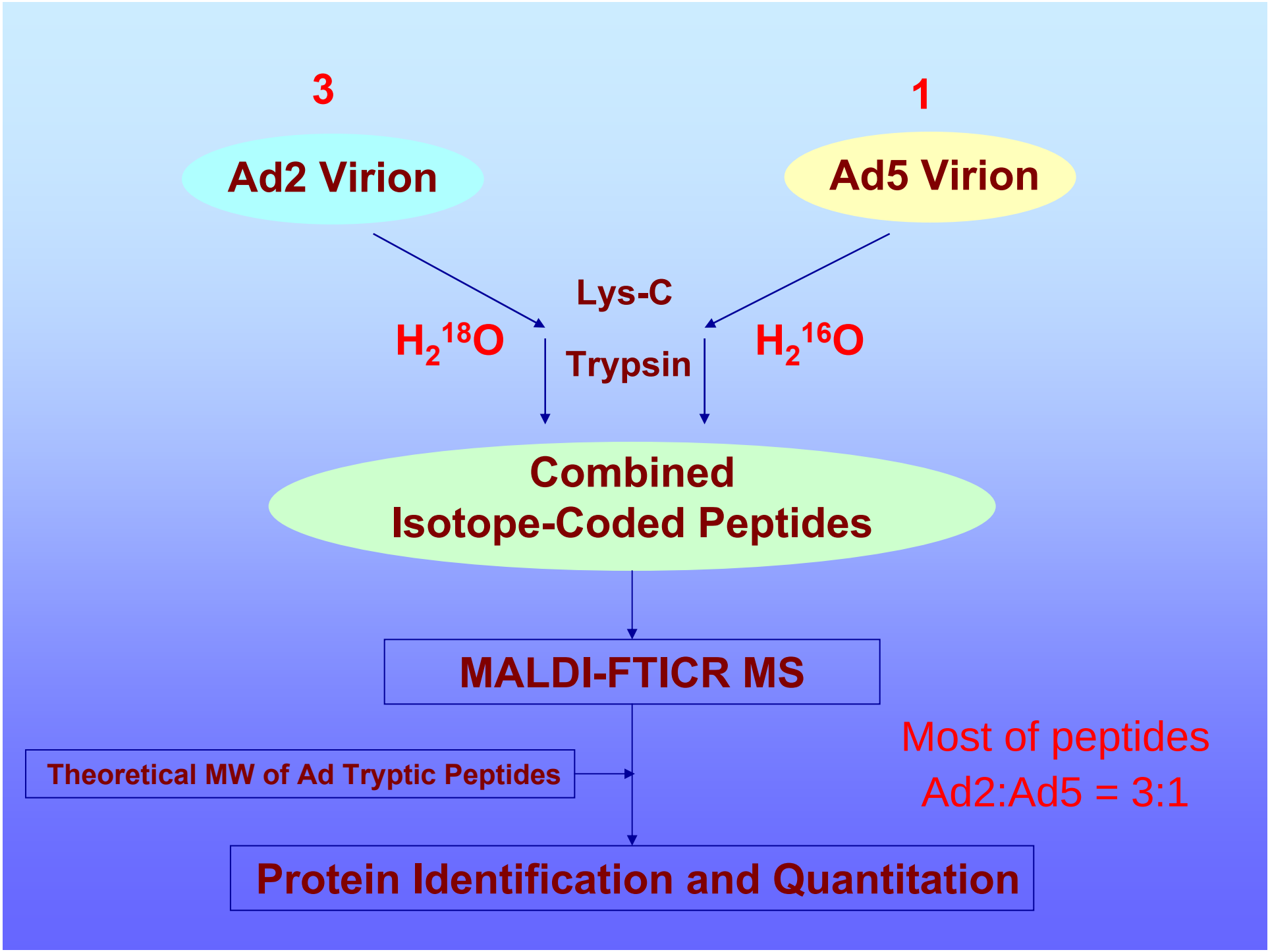


Adenovirus as Model System

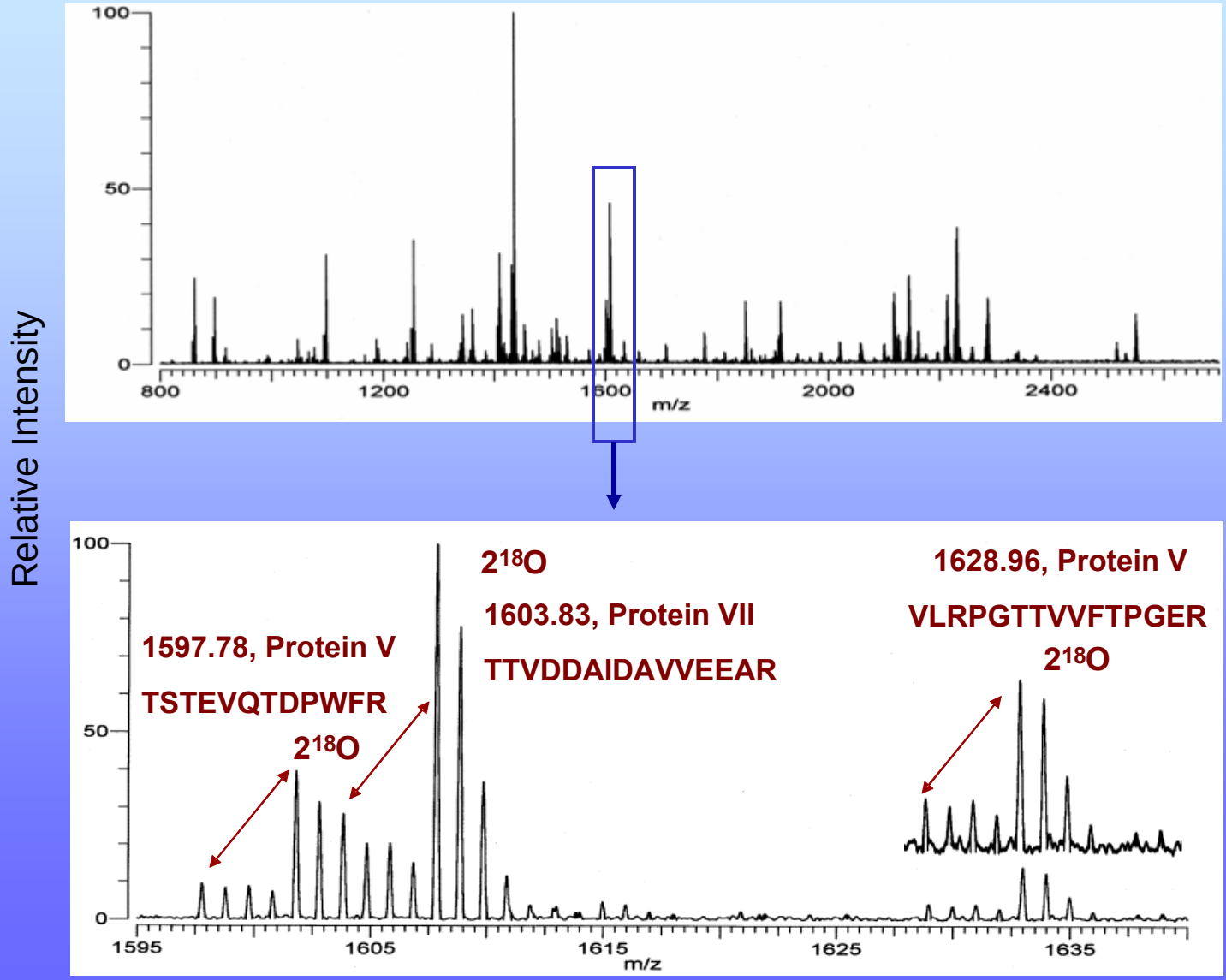
- Minimal real system
  - DNA + Proteins
- Known genome/proteome
  - Predictable proteolytic peptides
- Defined architecture
  - Predictable protein expression
    - Comparative quantitation of Ad2/Ad5 proteins
    - Dynamic range of 600-fold
  - Capsid protein modeling membrane/hydrophobic proteins
- Mutations modeling post-translational modifications



Shenk T. In *Fundamental Virology*, **1996**


3
1
Ad2 Virion
Ad5 Virion
Lys-C
$H_2{}^{18}O$
Trypsin
$H_2{}^{16}O$
Combined
Isotope-Coded Peptides
MALDI-FTICR MS
Theoretical MW of Ad Tryptic Peptides
Most of peptides
Ad2:Ad5 = 3:1
Protein Identification and Quantitation

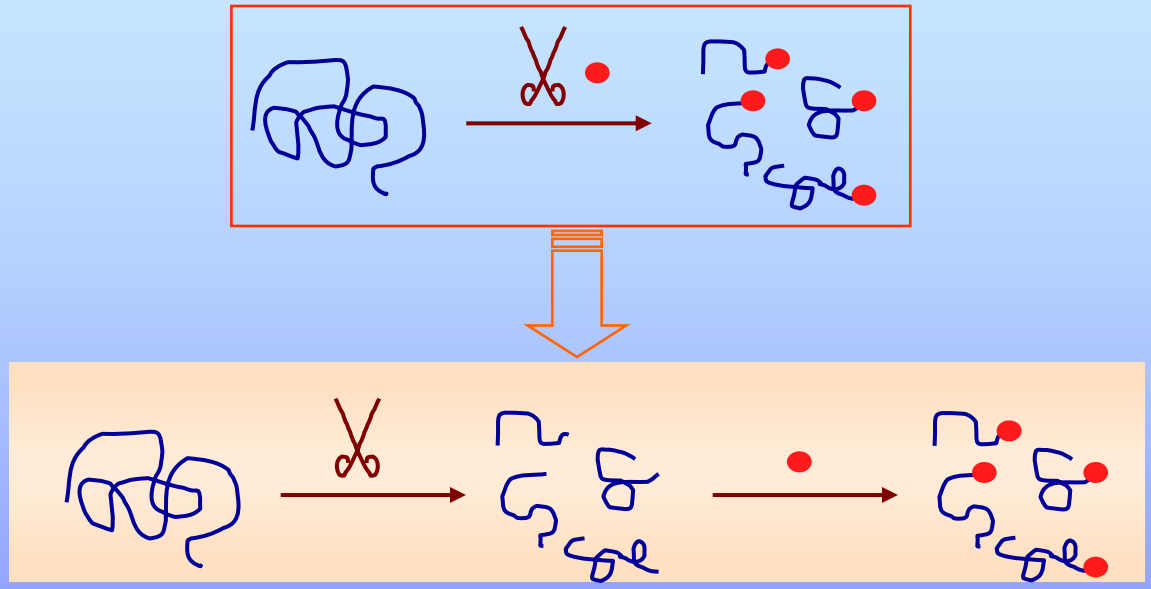
MALDI-FTICR Mass Spectrum of Combined Digests



Controversies and Challenges in Proteolytic Labeling

- Reported controversies in tryptic ^{18}O labeling
  - One ^{18}O incorporation for K-terminated peptides
  - Low efficient incorporation of two ^{18}O for short peptides
  - Two ^{18}O in each new peptides
- Capabilities of endoproteases for ^{18}O labeling
  - Two ^{18}O incorporation by trypsin, Lys-C, and Glu-C only
  - One ^{18}O incorporation by chymotrypsin…
- Challenges for automated, large-scale application
  - Amount and cost of $H_2{}^{18}O$

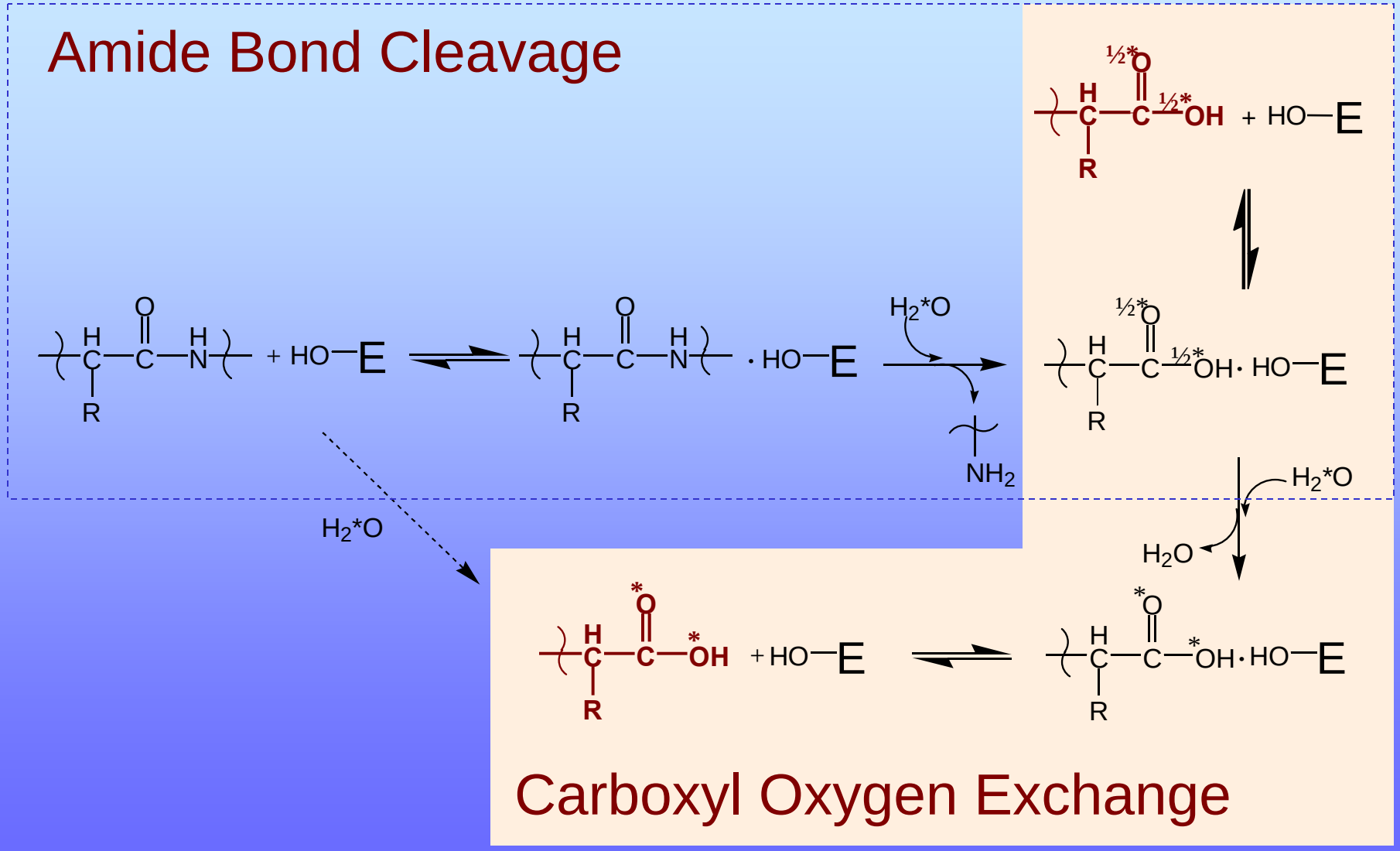
Decoupling Proteolytic ^{18}O Labeling from Protein Digestion



Proteins in solution prior to digestion
Peptide labeling in small volume of $H_2{}^{18}O$
Separate optimization of digestion and labeling

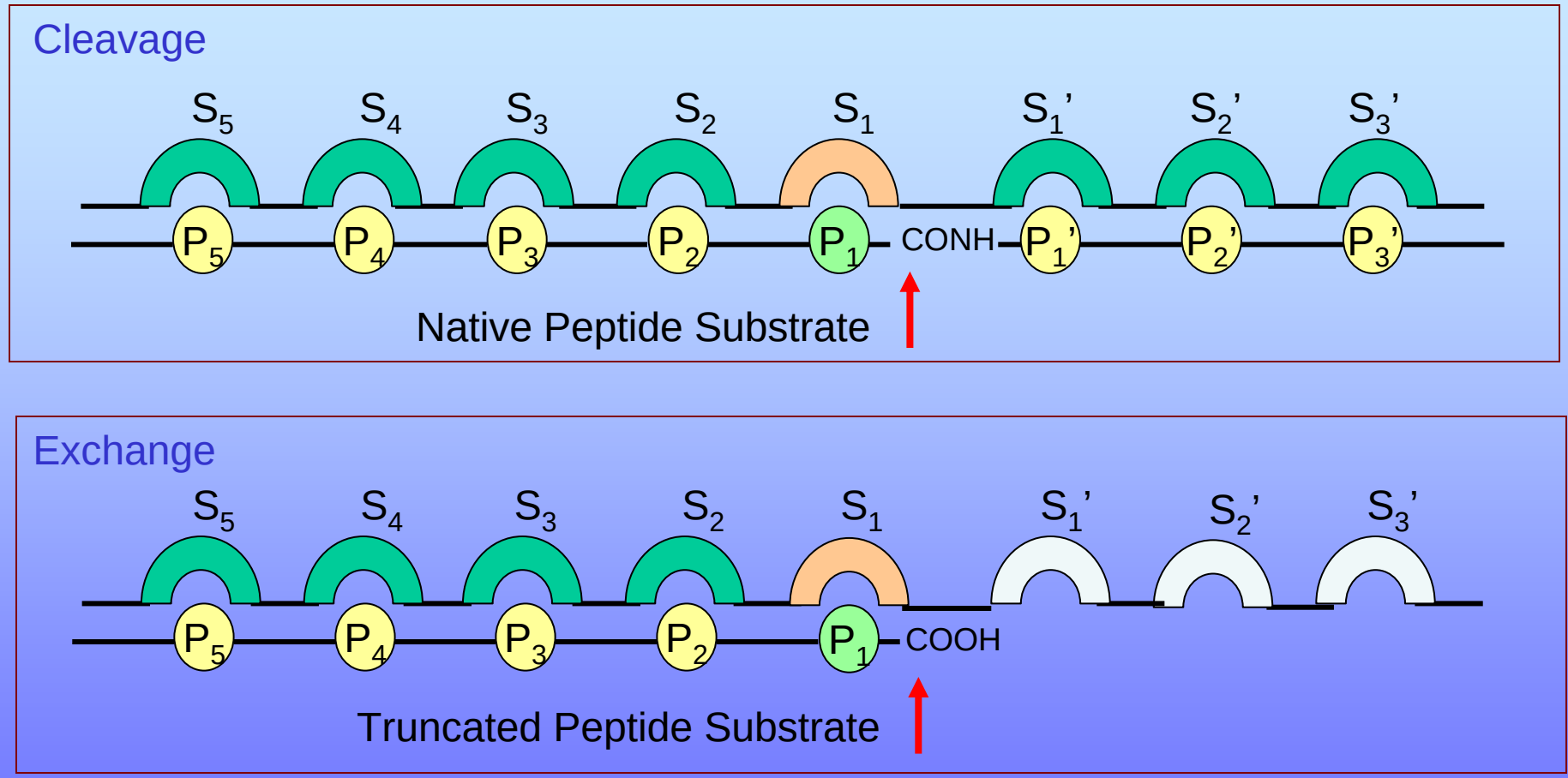
Automatic, high-throughput, large-scale applications

Dissection of Proteolytic Incorporation of Two ^{18}O



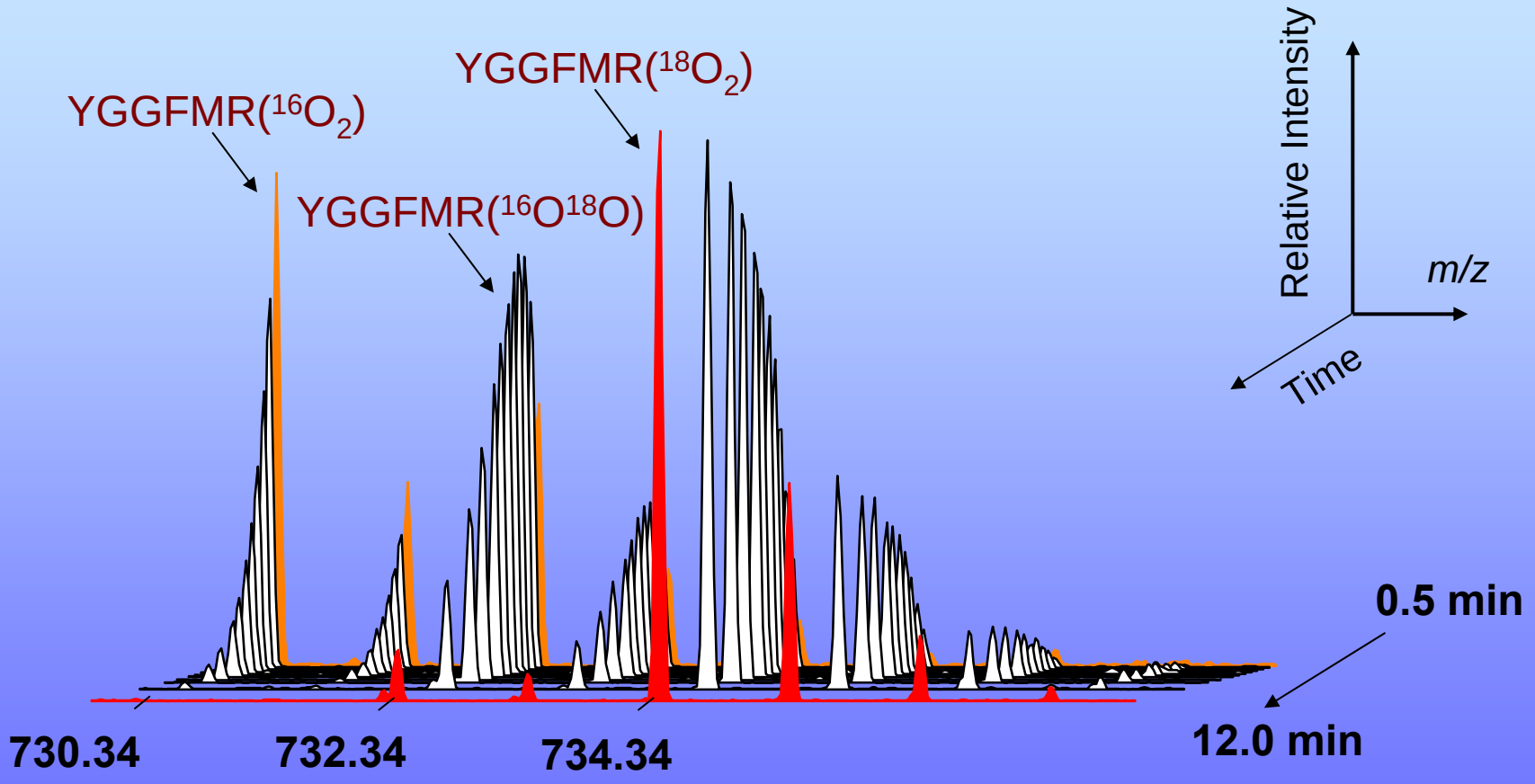
Yao, Afonso, Fenselau. *J. Proteom. Res.* **2003**, *2*, 147.

Molecular Basis for Cleavage and Exchange

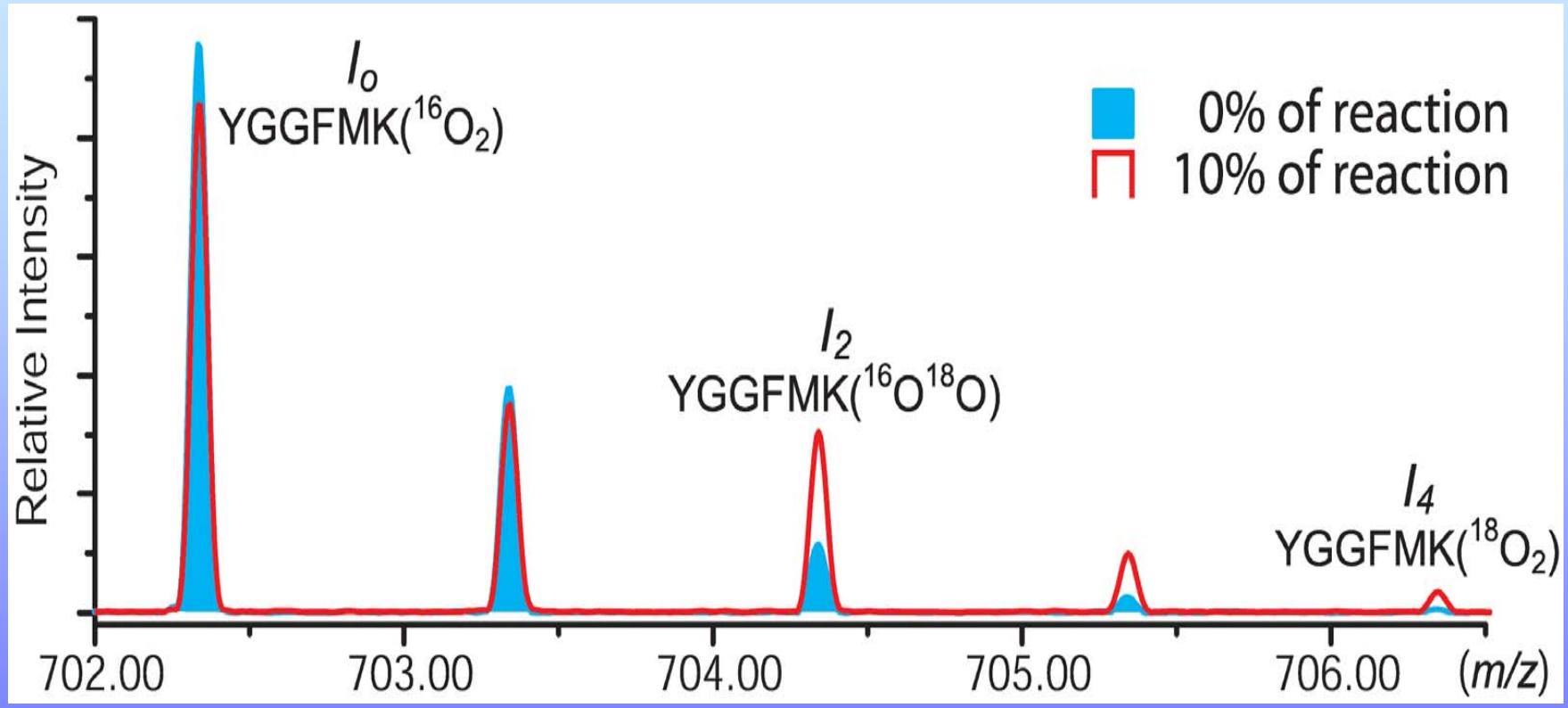


Protease catalyzes exchange → TWO ^{18}O INCORPORATION

^{16}O-to-^{18}O Exchange Studied by MALDI-FTICR MS

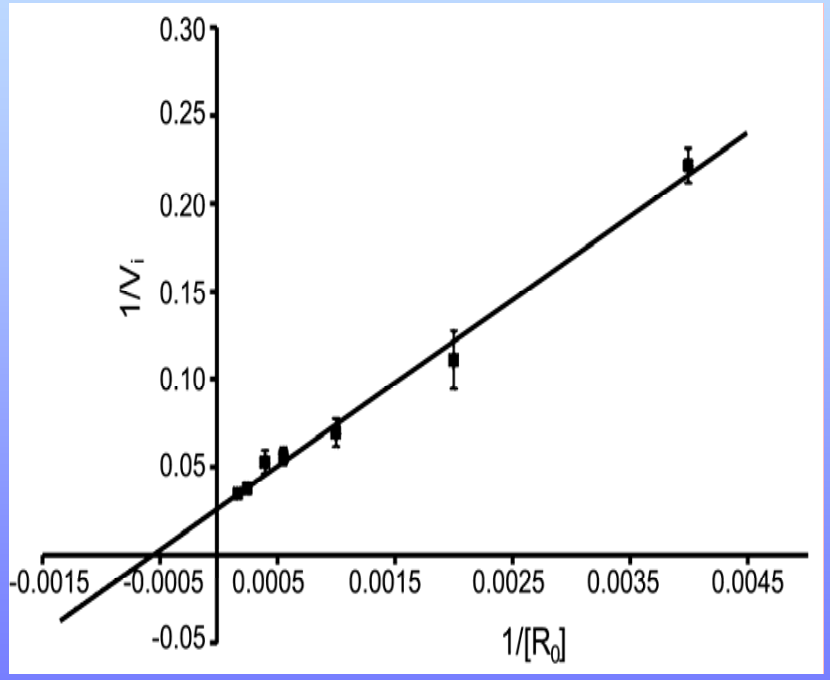


Determination of Reaction Initial Rates



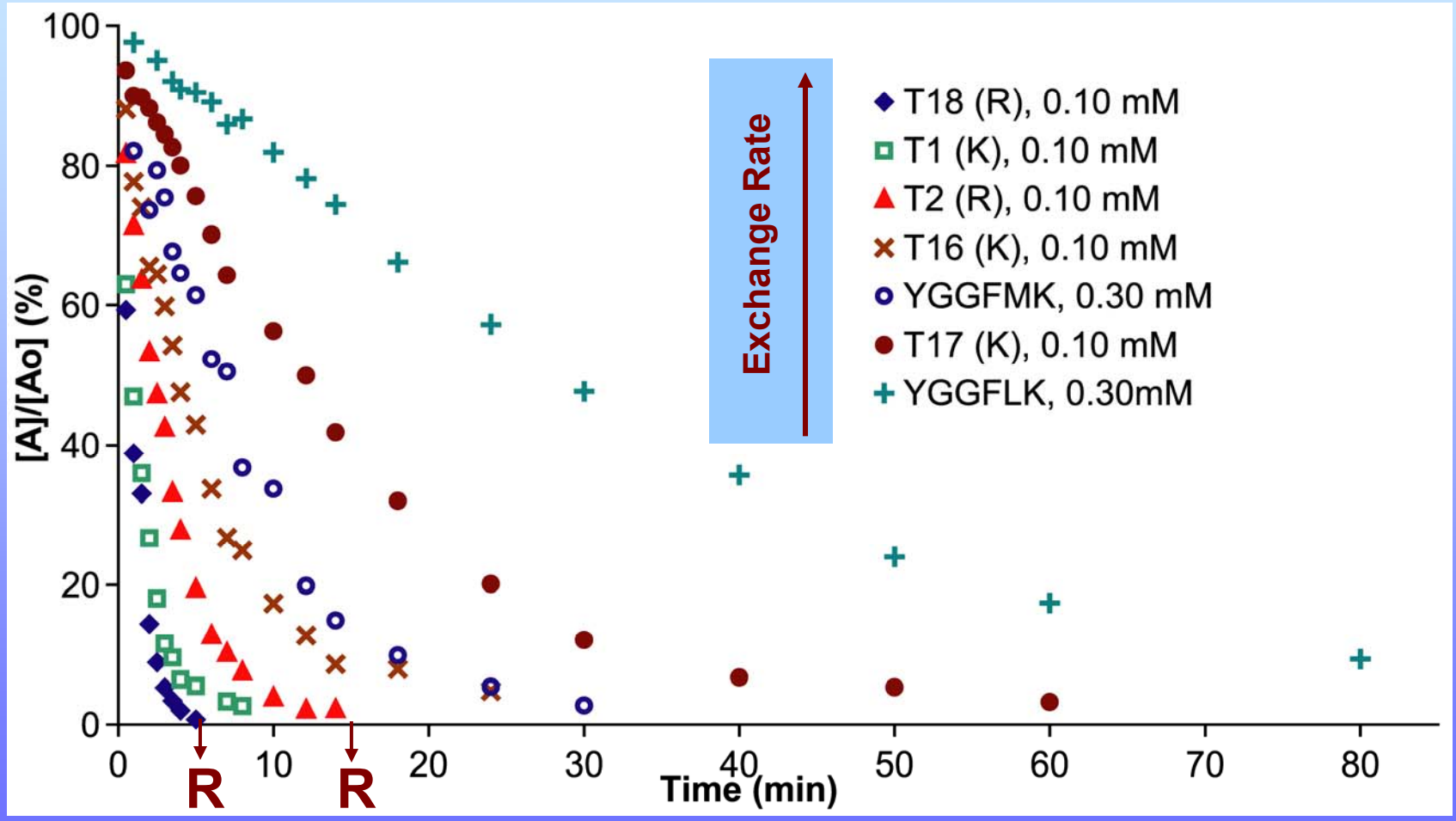
$$[R] = [R_o] \frac{I_o}{I_{total}(I_o, I_2, I_4, M_o, M_2, M_4)}$$

Kinetics Comparison in R- and K-Peptides



	YGGFMR	YGGFMK
K_{cat} (min^{-1})	3500±500	2800±300
K_M (μM)	1300±300	4400±700
k_{cat}/K_M ($\mu M^{-1}min^{-1}$)	2.6±0.9	0.64±0.14

Simultaneous Mass Spectrometric Determination of Kinetics for Trypsin-Catalyzed ^{16}O-to-^{18}O Exchange



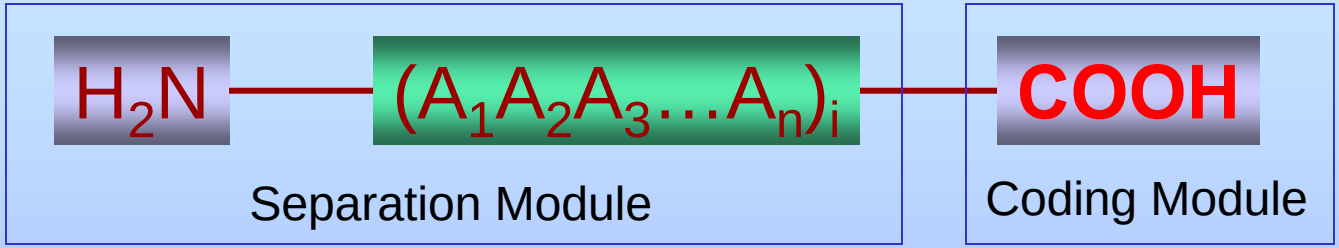
Complete Exchange for Mixture

Enzymatic ^{18}O Labeling

- Universal two ^{18}O labeling of proteolytic peptides by protease-catalyzed exchange
  - Both K- and R-terminated peptides
  - Chymotrypsin and pepsin for two ^{18}O labeling, in addition to trypsin, Lys-C, Glu-C, …
  - Both short and long peptides

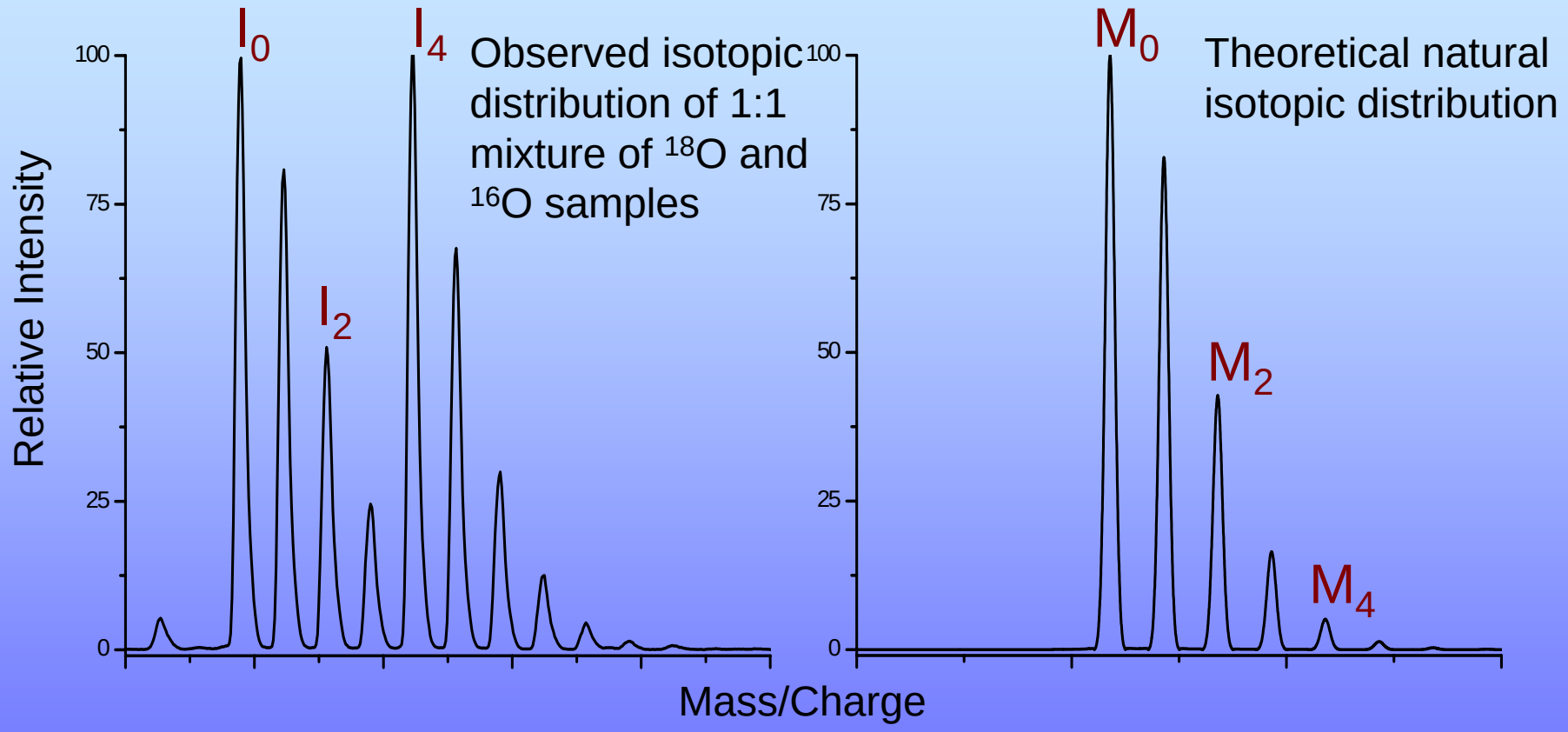
- 4 Da mass increase at the C-terminus of proteolytic peptides to be differentiated in mass spectrometry

Mass Spectrometry of Peptide-$^{16}O_2/^{18}O_2$ Pairs



- ^{18}O-labeling enabled mass spectrometric quantitation
- Effect of peak resolution on quantitation
- Analysis on different mass analyzer configurations
- More than relative quantitation from differential oxygen labeling

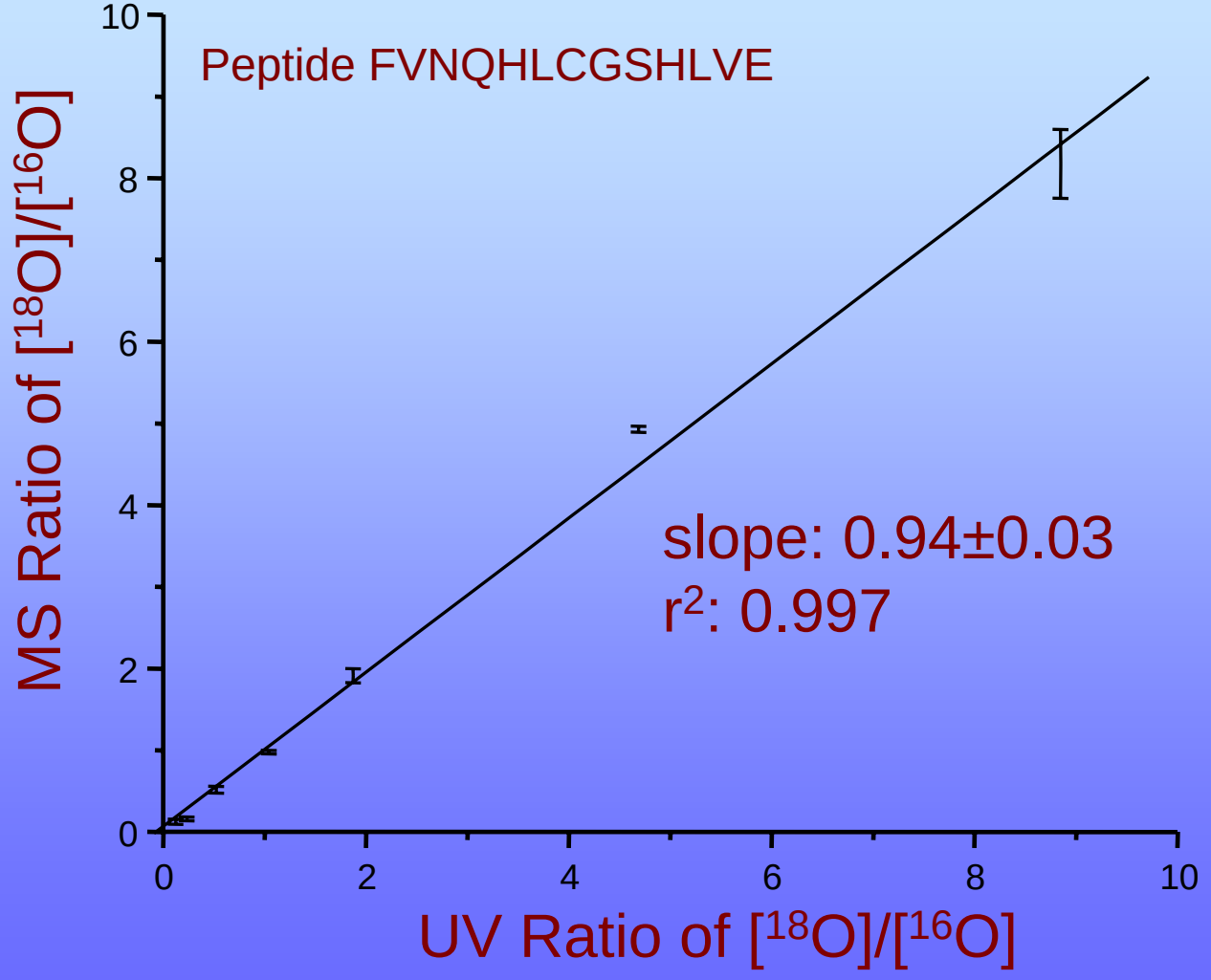
Relative Quantitation Using Paired Isotope Clusters



$$\text{Ratio} = \frac{I_4}{I_o} + \left(1 - \frac{M_2}{M_o}\right)\frac{I_2}{I_o} + \left[\left(\frac{M_2}{M_o}\right)^2 - \frac{M_2}{M_o} - \frac{M_4}{M_o}\right]$$

$$\frac{I_5}{I_1}$$

Correlation of ESI Quantitation with Peptide UV Quantitation



Reynolds, Yao, Fenselau. *J. Proteom. Res.* **2003**, *1*, 27.

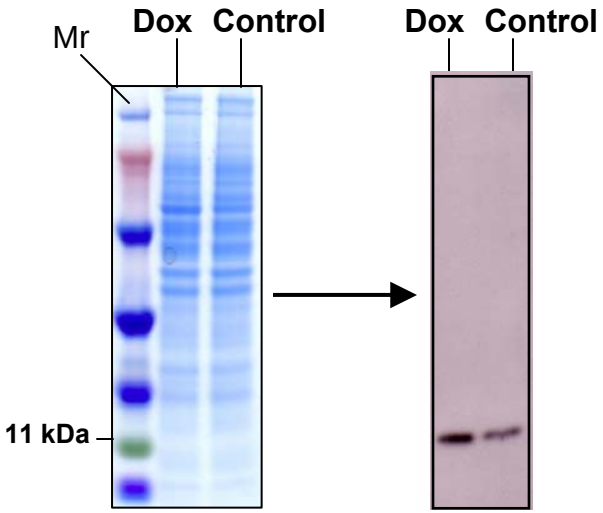
Labeling Consistency

BSA Peptide Sequence	Sequence Position	Unlabeled (I_0) Elution Time (minutes)[a]	Labeled (I_4) Elution Time (minutes)[a]	Ratio 1[b]	I_4/I_0[c]
ACFAVE	589-594	46.48	46.48	0.95	0.87
KKFWGKYLYE	155-164	52.17	52.09	0.84	0.85
TYVPKAFDE	519-527	52.93	52.93	0.99	0.96
DKDVCKNYQE	335-344	58.43	58.43	0.99	0.97
DKGACLLPKIE	196-206	57.36	57.36	0.88	0.87
KQIKKQTALVE	544-554	67.70	67.70	0.90	0.88
LLYYANKYNGVFQE	177-190	75.32	75.32	0.99	1.07
YAVSVLLRLAKE	364-375	77.56	77.48	0.91	0.91
AKDAFLGSFLYE	345-356	88.78	88.78	0.86	0.81
DYLSLILNRLCVLHE	474-488	109.11	109.03	0.85	0.96
			Average	0.92	0.92
			Standard Deviation	0.06	0.08

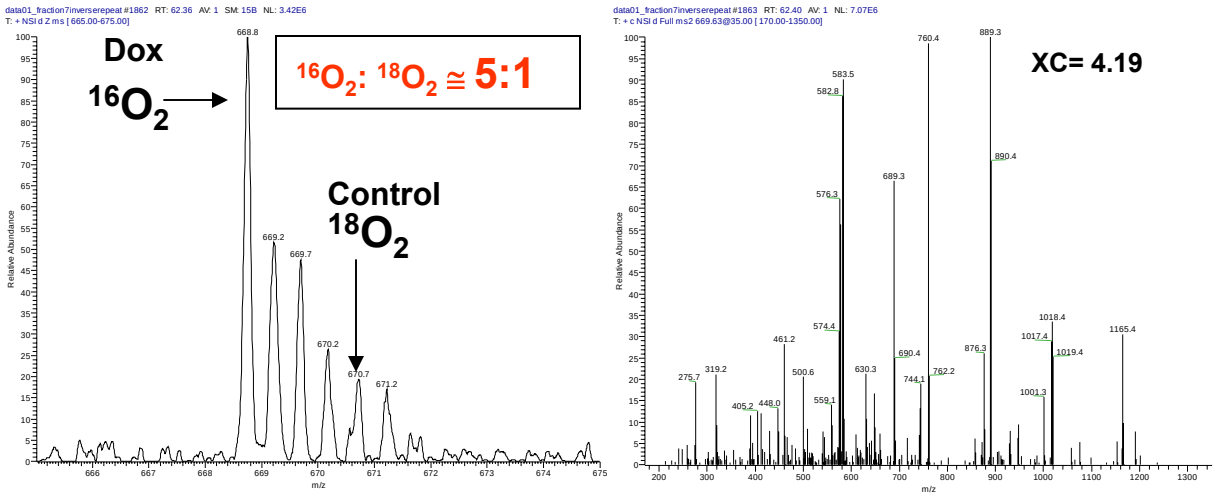
Reynolds, Yao, Fenselau. *J. Proteom. Res.* **2003**, *1*, 27.

Correlation between MS and Western-Blot Quantification of Thioredoxin in Doxorubicin-Treated HeLa Cells

SDS-PAGE Analysis of Whole Cell-Extracts:



ESI-IT-MS Analysis of Anion-X Fraction 7:
(Labeling: Dox Treated ^{16}O, Control ^{18}O)



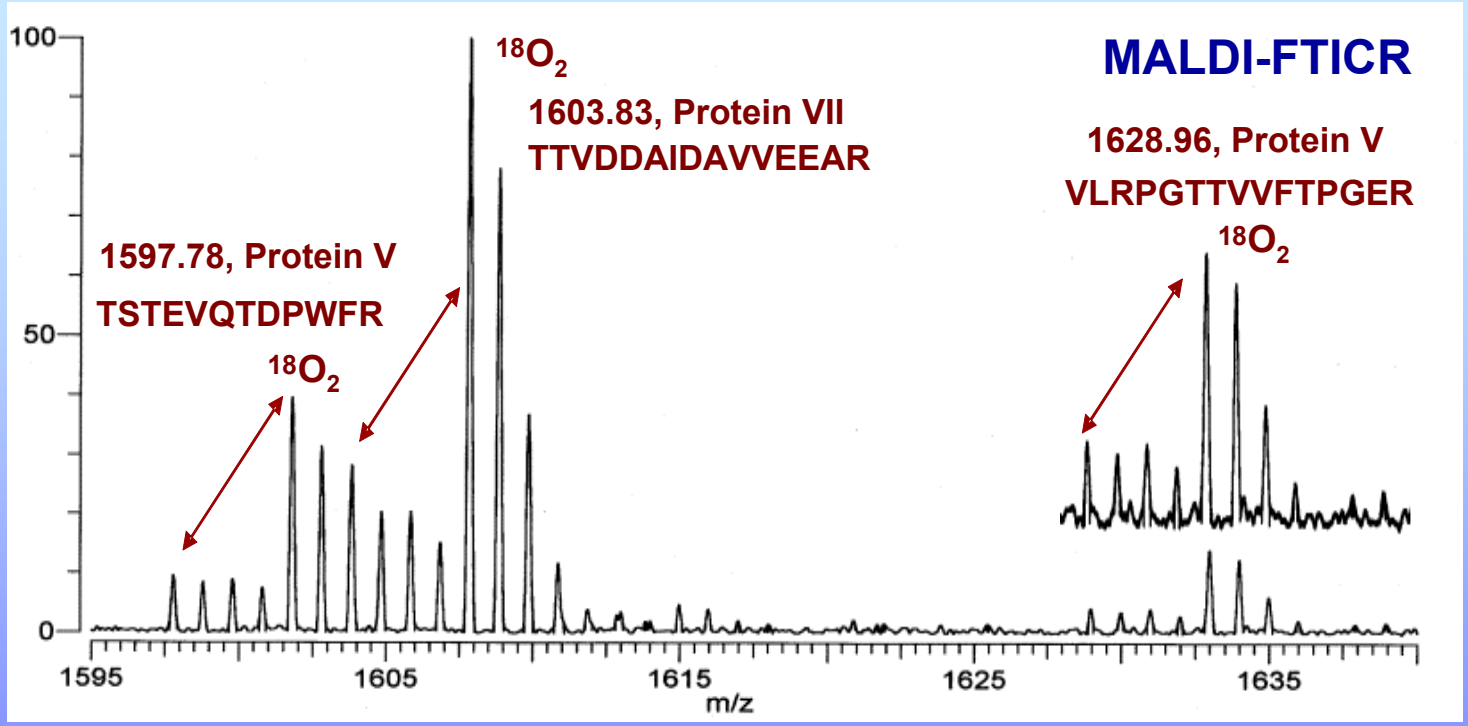
Zoom Scan: Relative Quantification

$[M+2H]^{+2}$=668.8
$[M+2H +{}^{18}O_2]^{+2}$=670.7

$[M+H]^{+}$=1337.42
K.TAFQEALDAAGDK.L
Thioredoxin

Courtesy of Dan Clark of Stratagene

Effect of Peak Resolution on $^{18}O/^{16}O$ Ratio (I)



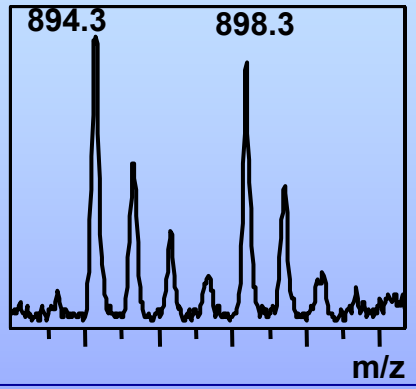
Protein	II	VII	IX	VIII	Terminal
% wt	60	14	3	0.3	0.1
Relative Intensity	2.9±0.3	3.7±0.4	3.0±0.5	3.3±0.5	2.6 ±0.3

Yao, Freas, Ramirez, Demirev, Fenselau. *Anal. Chem.* **2001**, *73*, 2836.

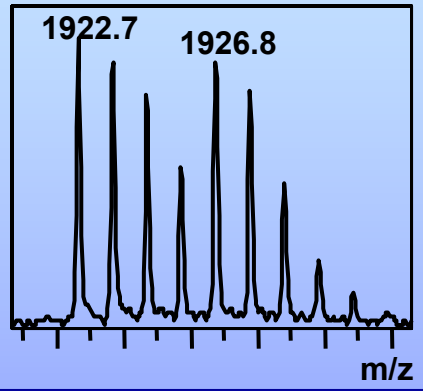
Effect of Peak Resolution on $^{18}O/^{16}O$ Ratio (II)

MALDI-TOF

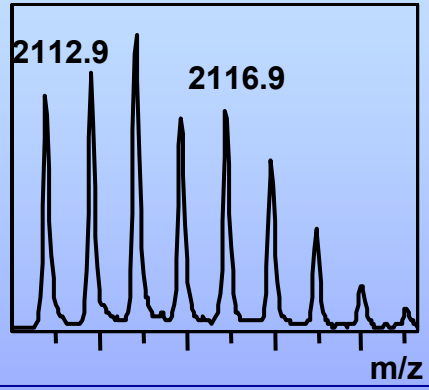
APC2_Human: 0.91



APC3_Human: 0.94

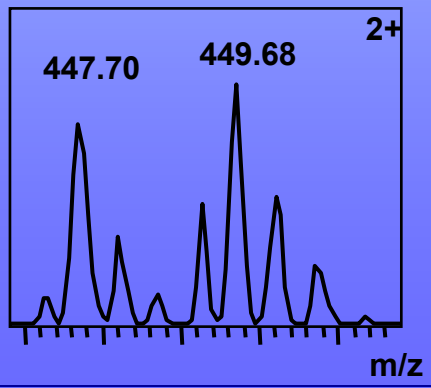


A1AH_Human: 0.94

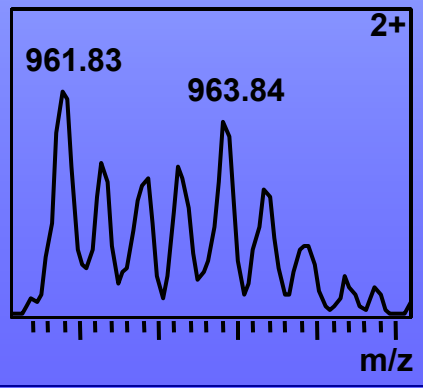


ESI-IT

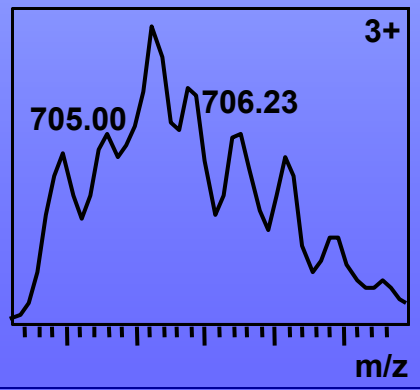
APC2_Human: 1.2



APC3_Human: 0.87

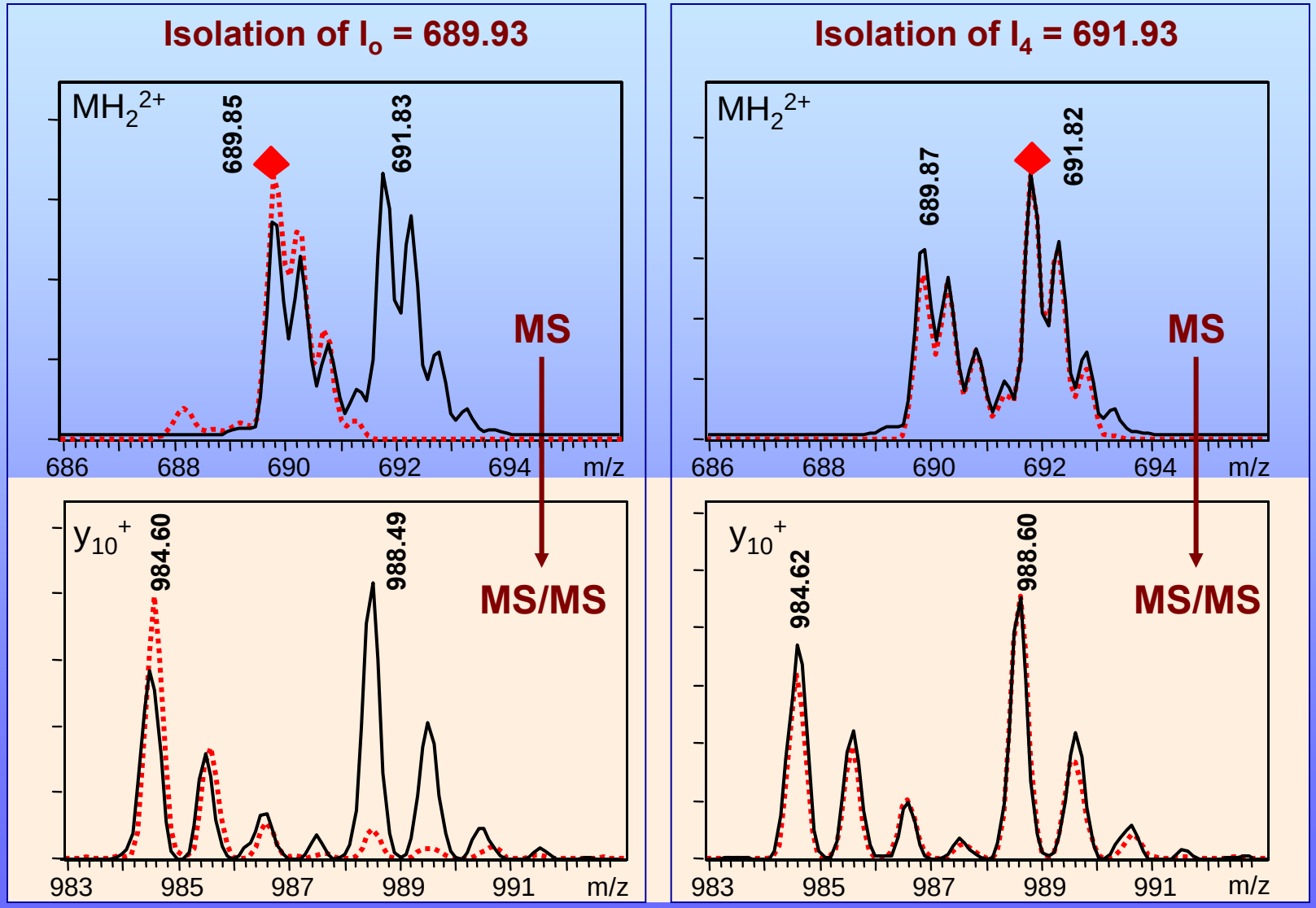


A1AH_Human: 1.3



Heller, Mattou, Menzel, Yao. *JASMS* **2003**, *14*, 704.

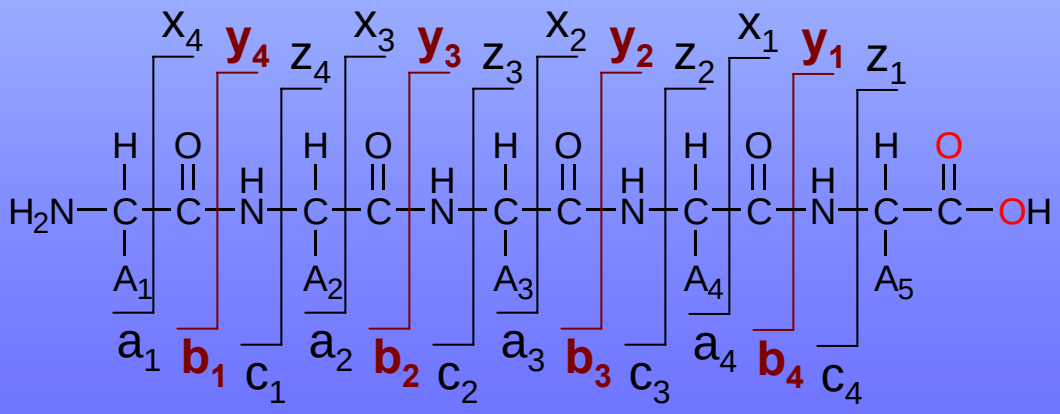
Effect of Isolation Window Width on Quantitation Using $^{16}O/^{18}O$ y-Ion Pairs on IT MS



Heller, Mattou, Menzel, Yao. *JASMS* **2003**, *14*, 704.

Protein Sequence Ions Generated by Tandem Mass Spectrometry

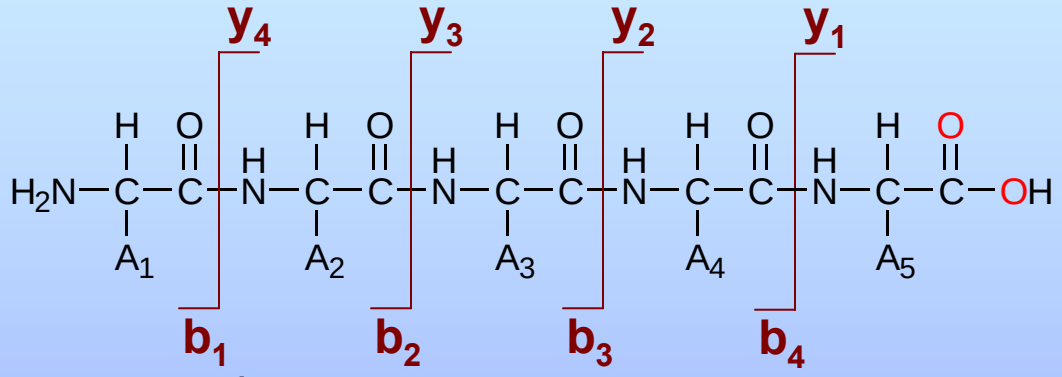
Mass Analyzer 1 → **Fragmenting the Selection Ion** → **Mass Analyzer 2**

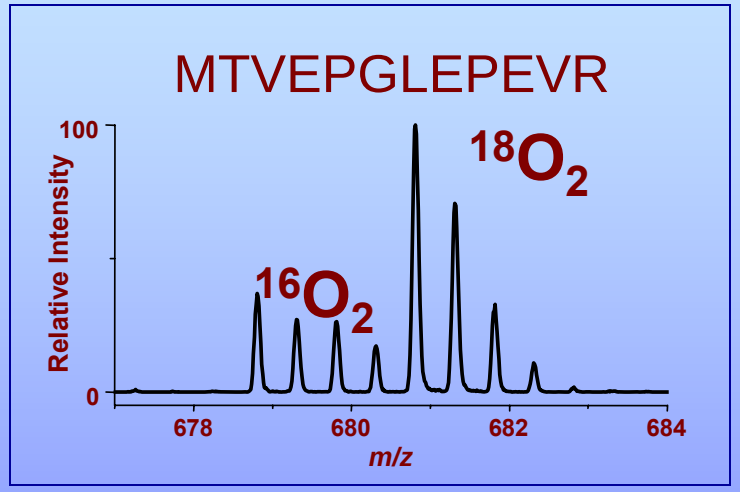


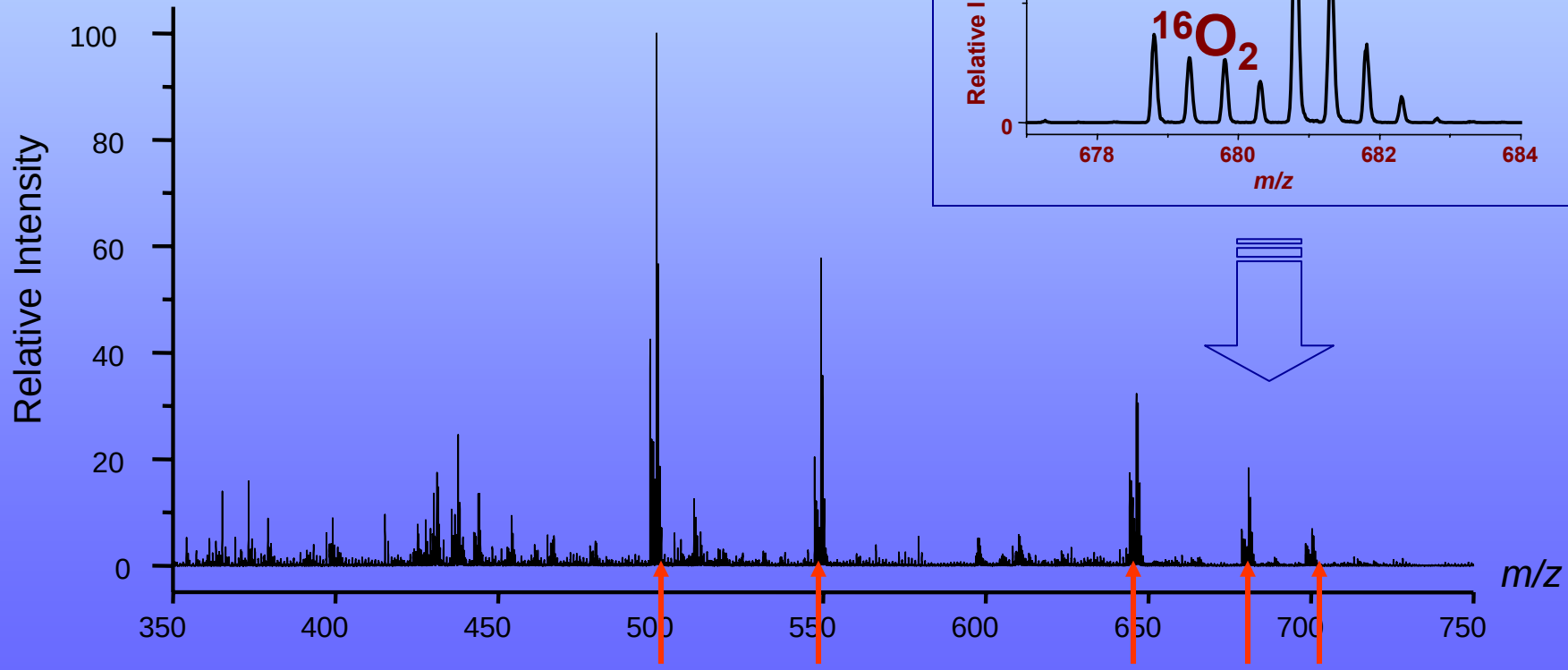
Masses of Amino Acid Residues

Amino acid	Symbol		Residue Mass (Da)	Immonium Ion (Mass (Da)	Side-Chain Mass (Da)
Alanine	Ala	A	71.04	44	15
Arginine	Arg	R	156.10	129	100
Asparagine	Asn	N	114.04	87	58
Aspartic acid	Asp	D	115.03	88	59
Cysteine	Cys	C	103.01	76	47
Glutamic acid	Glu	E	129.04	102	73
Glutamine	Gln	Q	128.06	101	72
Glycine	Gly	G	57.0	30	1
Histidine	His	H	137.06	110	81
Isolcucine	Ile	I	113.08	86	57
Leucine	Leu	L	113.08	86	57
Lysine	Lys	K	128.10	101	72
Methionine	Met	M	131.04	104	75
Phenylalanine	Phe	F	147.07	120	91
Proline	Pro	P	97.05	70	–
Serine	Ser	S	87.03	60	31
Threonine	Thr	T	101.05	74	45
Tryptophan	Trp	W	186.08	159	30
Tyrosine	Tyr	Y	163.06	136	107
Valine	Val	V	99.07	72	43

Quantitation Based on MS/MS Spectrum (y-Ions)

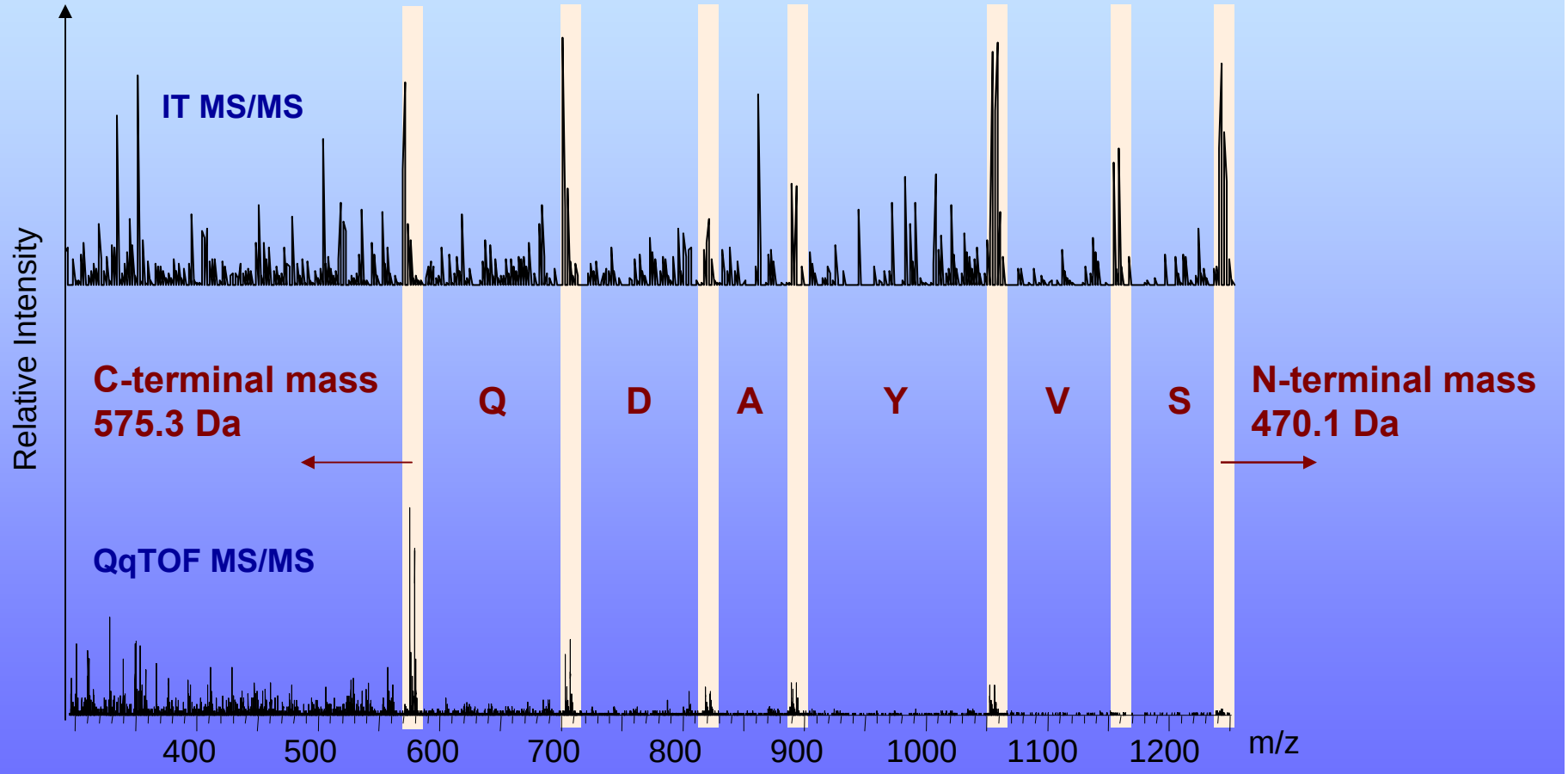




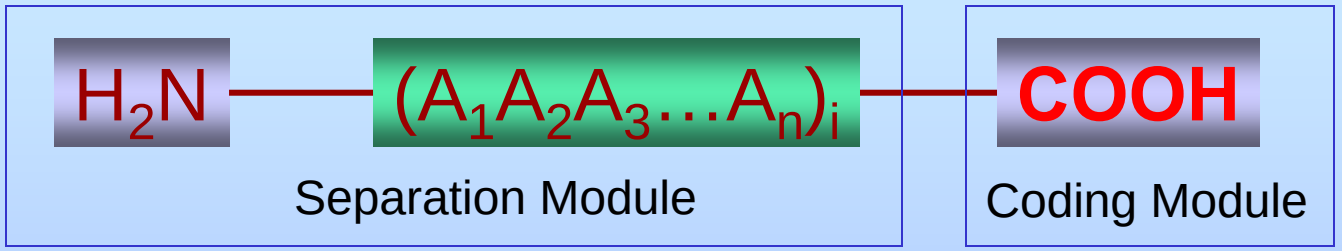


Yao, Freas, Ramirez, Demirev, Fenselau. *Anal. Chem.* **2001**, *73*, 2836.

$^{16}O/^{18}O$ Paired Peptides Facilitate and Validate Peptide Sequencing



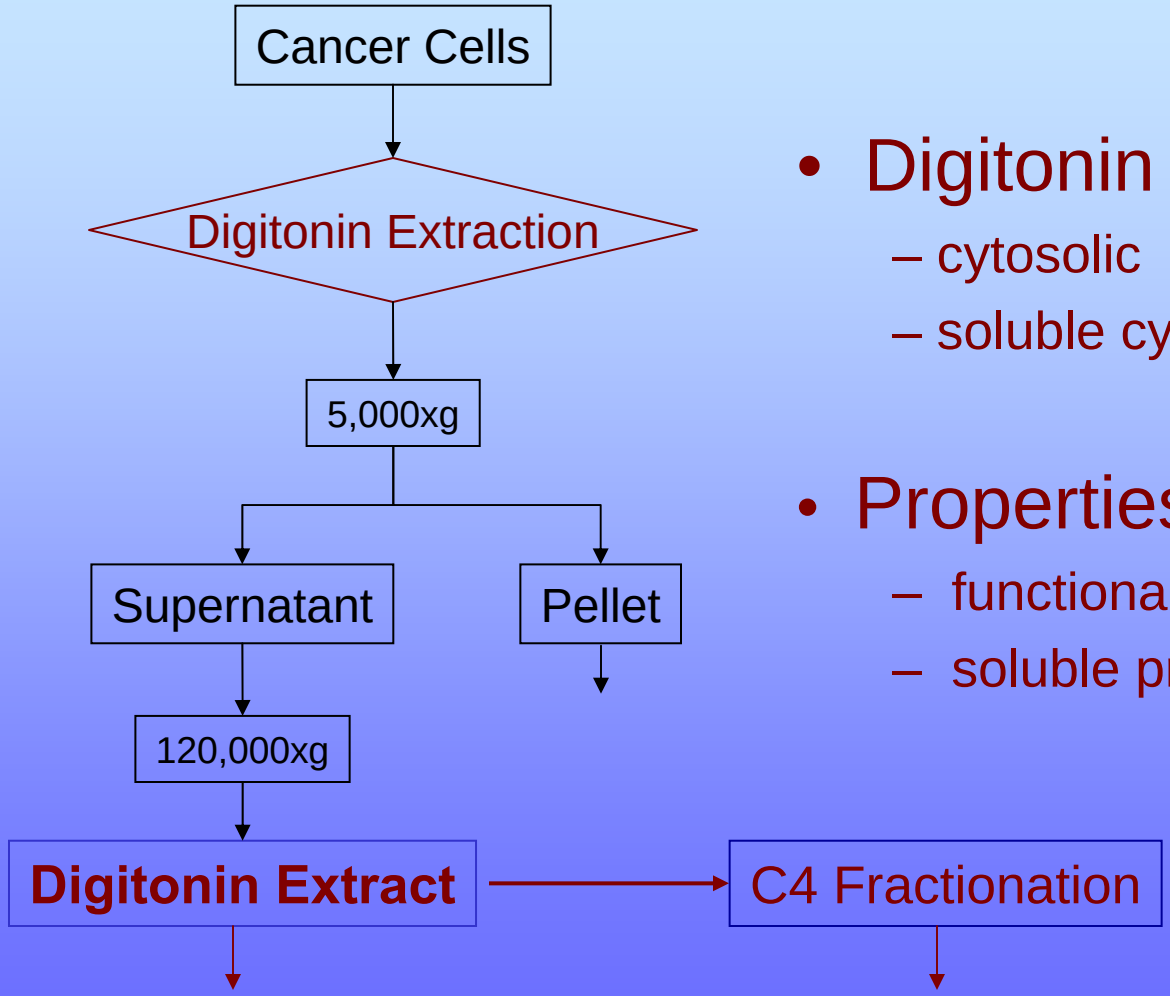
Advantages of Modular Design



- **Isotope Coding**
  - **Universal**
    - Important to small proteins
  - **Specific**
  - **Efficient**
  - **Minimal Structural Modification**
    - Chromatographic co-elution
  - **Stable during separation**

- **Separation**
  - **Portable to all separation platforms: LC/LC, solution IEF, affinity separation**

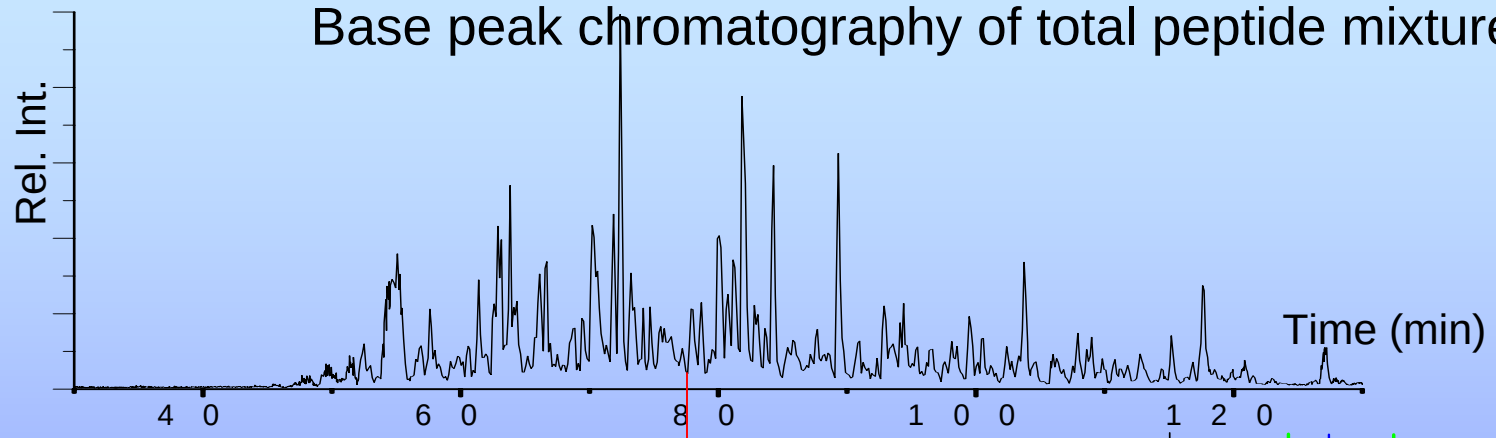
Protein Pools of Digitonin Extract of MCF-7 Cells

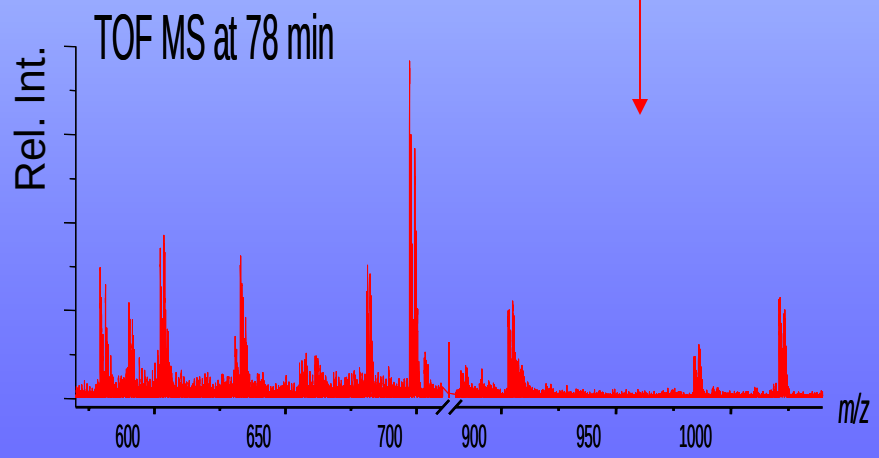


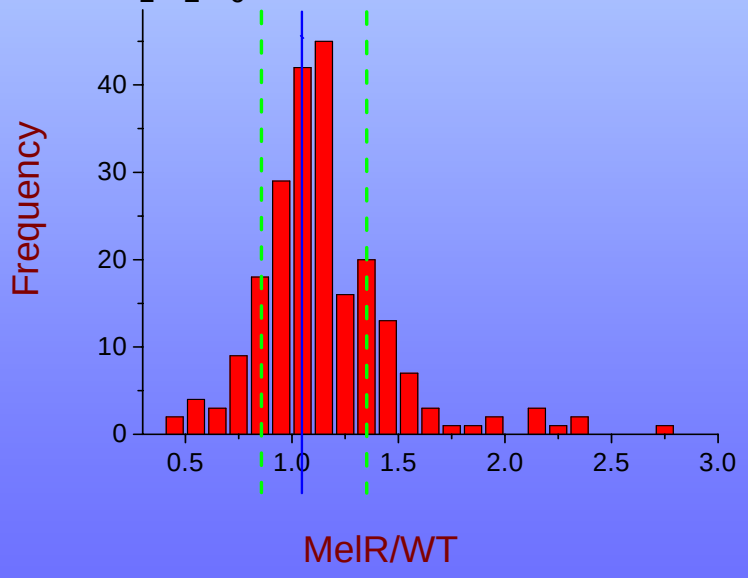
- Digitonin fraction
  - cytosolic
  - soluble cytoskeletal proteins
- Properties
  - functional proteins
  - soluble proteins

Modified according to “Ramsby, Makowski *Method Mol. Biol., 112,* 1999.

LC-MS of Peptides from MCF-7 Digitonin Fraction







- Protein Ratio (MelR/WT) = 1.1
- I_0/I_4 Ratio (MelR/WT) = 1.1 ± 0.3
- 83% (184/223) peptides in 1.1 ± 0.3

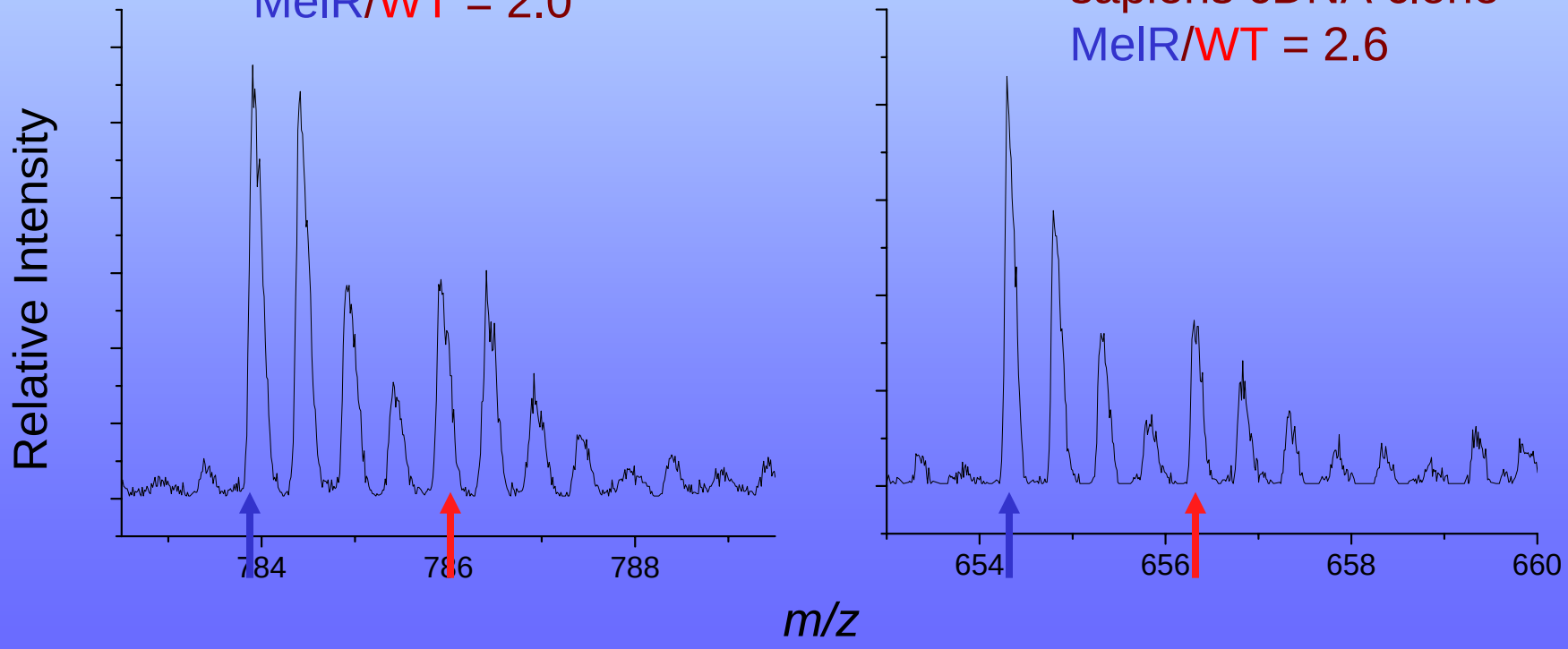
Yao, Fenselau. *ASMS Annual Conference, 2001*

Protein Expression Changes in MCF-7 Cells Upon Acquisition of Melphalan Resistance

[Most Proteins in a Ratio (MelR/WT) of 1.1]

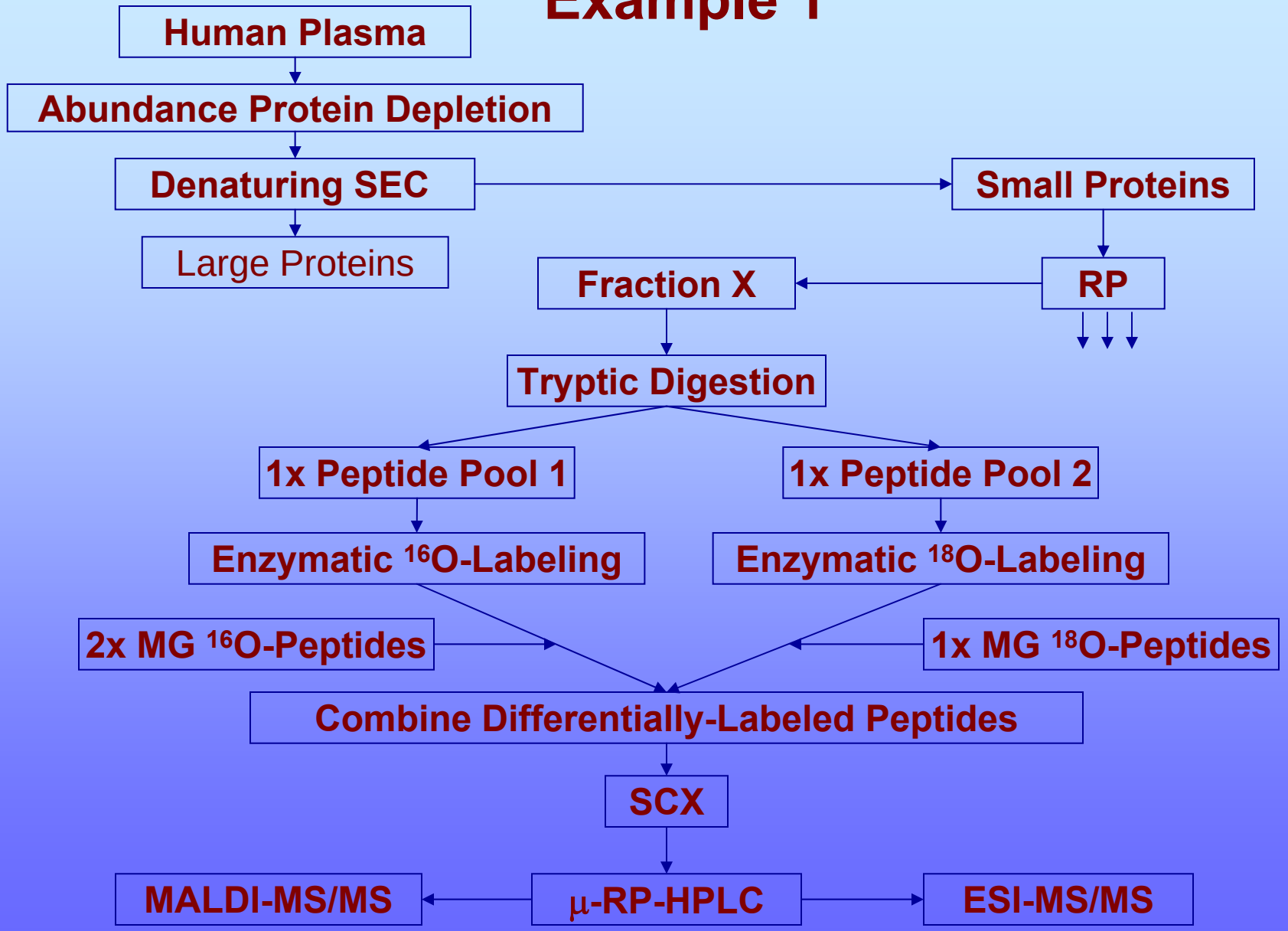
(K)LLPQLTYLDGYDR(E)
PHAPI2b /April protein
pI = 4.0
MelR/WT = 2.0

(R)GIVTNWDDMEK(I)
602308605F1
NIH_MGC_88 Homo sapiens cDNA clone
MelR/WT = 2.6



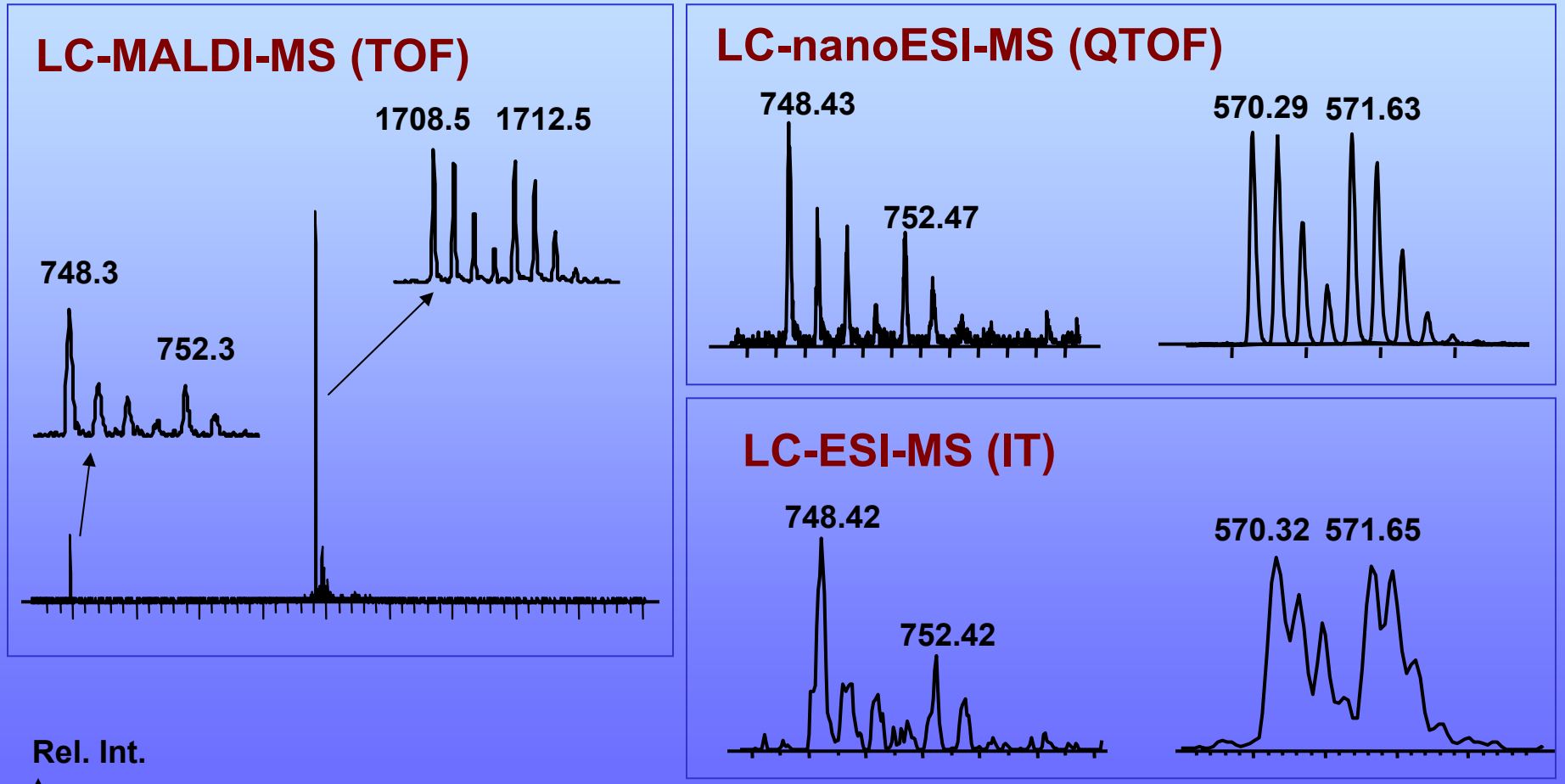
Yao, Fenselau. *ASMS Annual Conference, 2001*

Analysis of Human Plasma Sample: Example 1



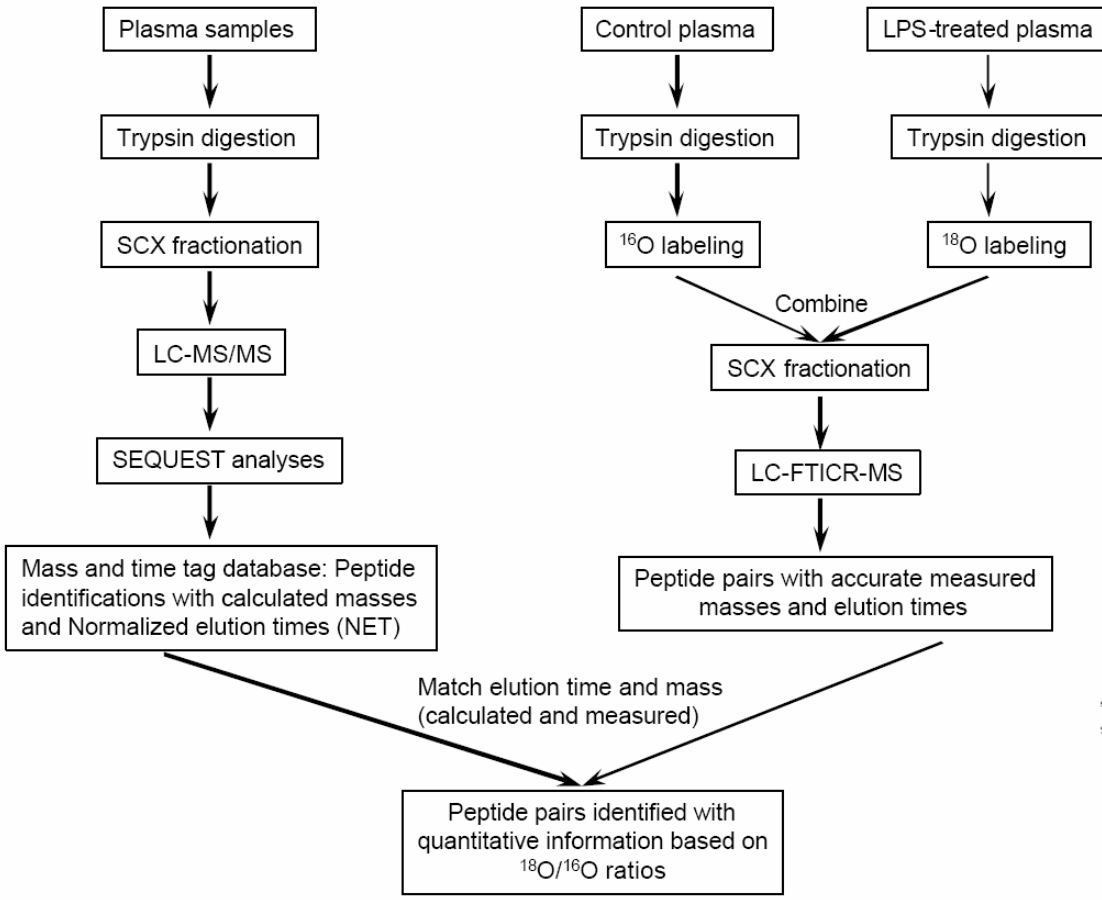
Heller, Mattou, Menzel, Yao. *JASMS* **2003**, *14*, 704.

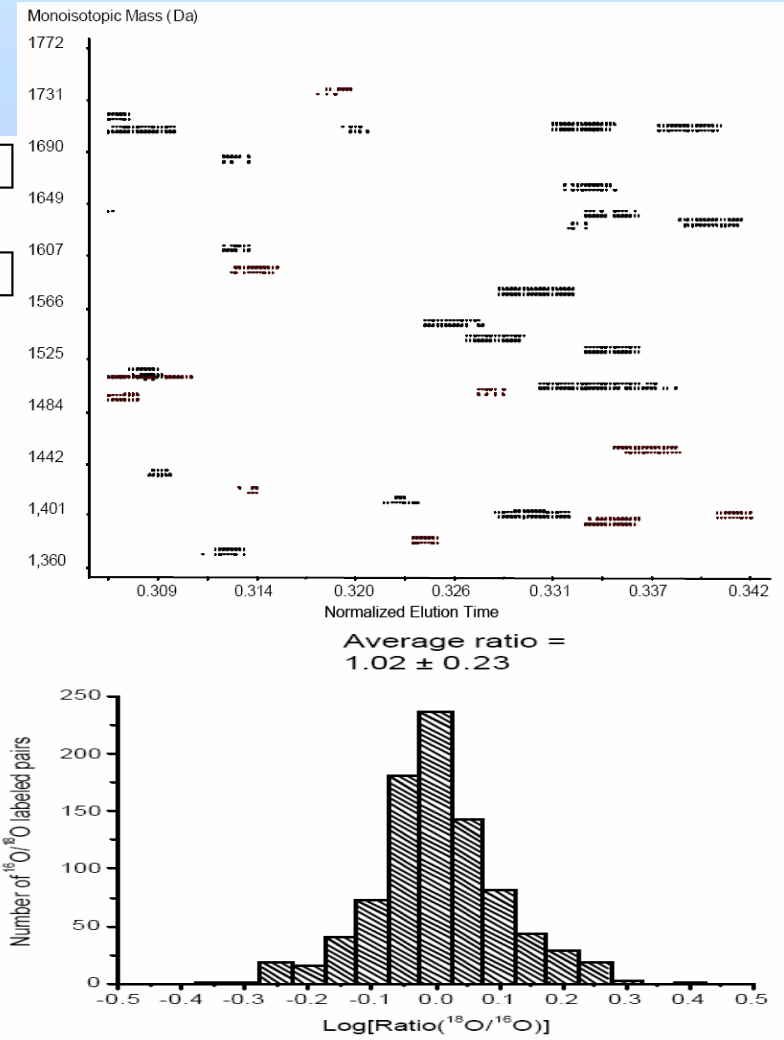
LC-MALDI & LC-ESI MS Analysis of Differentially $^{18}O/^{16}O$-Labeled Peptides Present in Human Plasma



Heller, Mattou, Menzel, Yao. *JASMS* **2003**, *14*, 704.

Analysis of Human Plasma Sample: Example 2



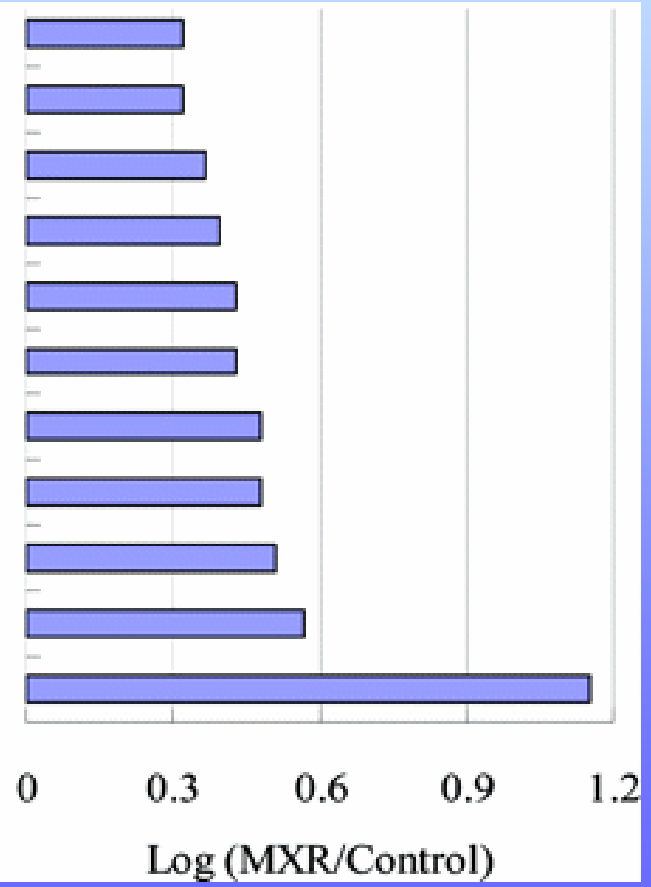


Qian *et. al.*, Smith MCP **2005**, March 7

Coupling $^{16}O/^{18}O$-Labeling and Solution Isoelectric Focusing for Peptide Analysis

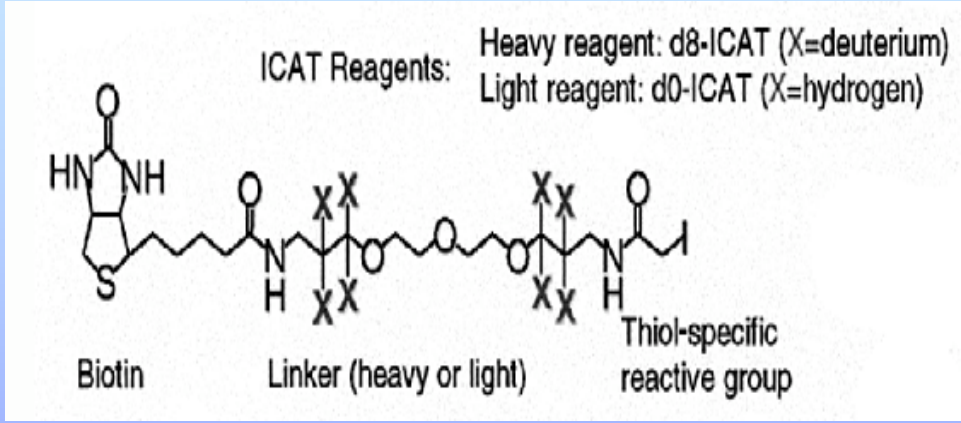
- Comparative investigation of insoluble nuclear subproteome
- Sample separation after combining differentially-labeled peptides

DEAH(Asp-Glu-Ala-His) box polypeptide 9
hnRNP A2/B1
ATP synthase alpha
ubiquinol-cytochrome-c reductase complex core protein
ATP synthase beta
nucleophosmin
heat shock protein 70
hypothetical protein FJL34423
hnRNP H'
hnRNP F
nuclear mitotic apparatus protein 1



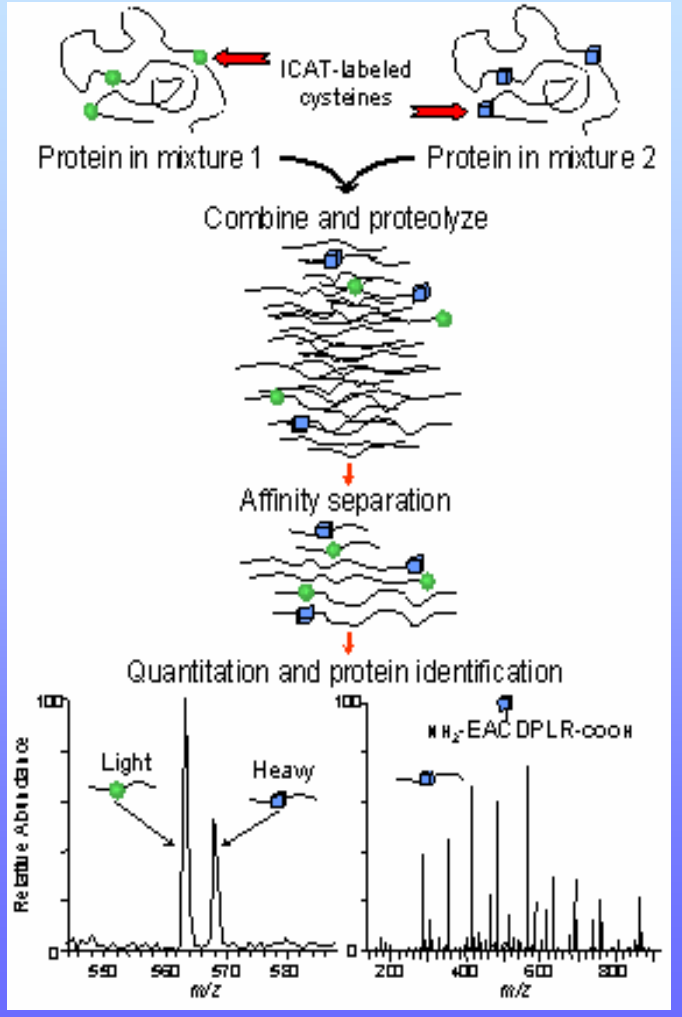
An, Fu, Gutierrez, Fenselau. *J. Proteom. Res.* **2005**, *4*, 2126.

Isotope-Coded Affinity Tag (ICAT)

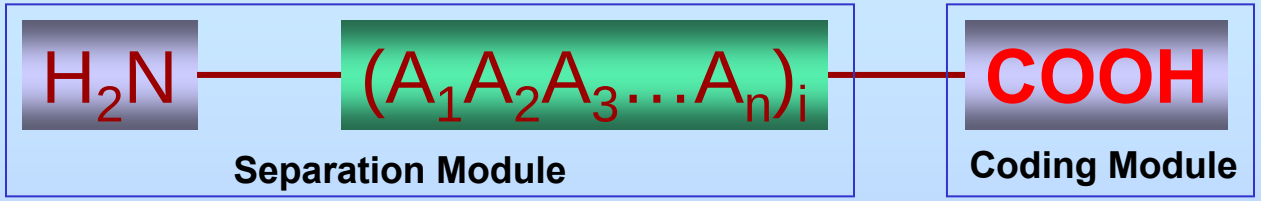


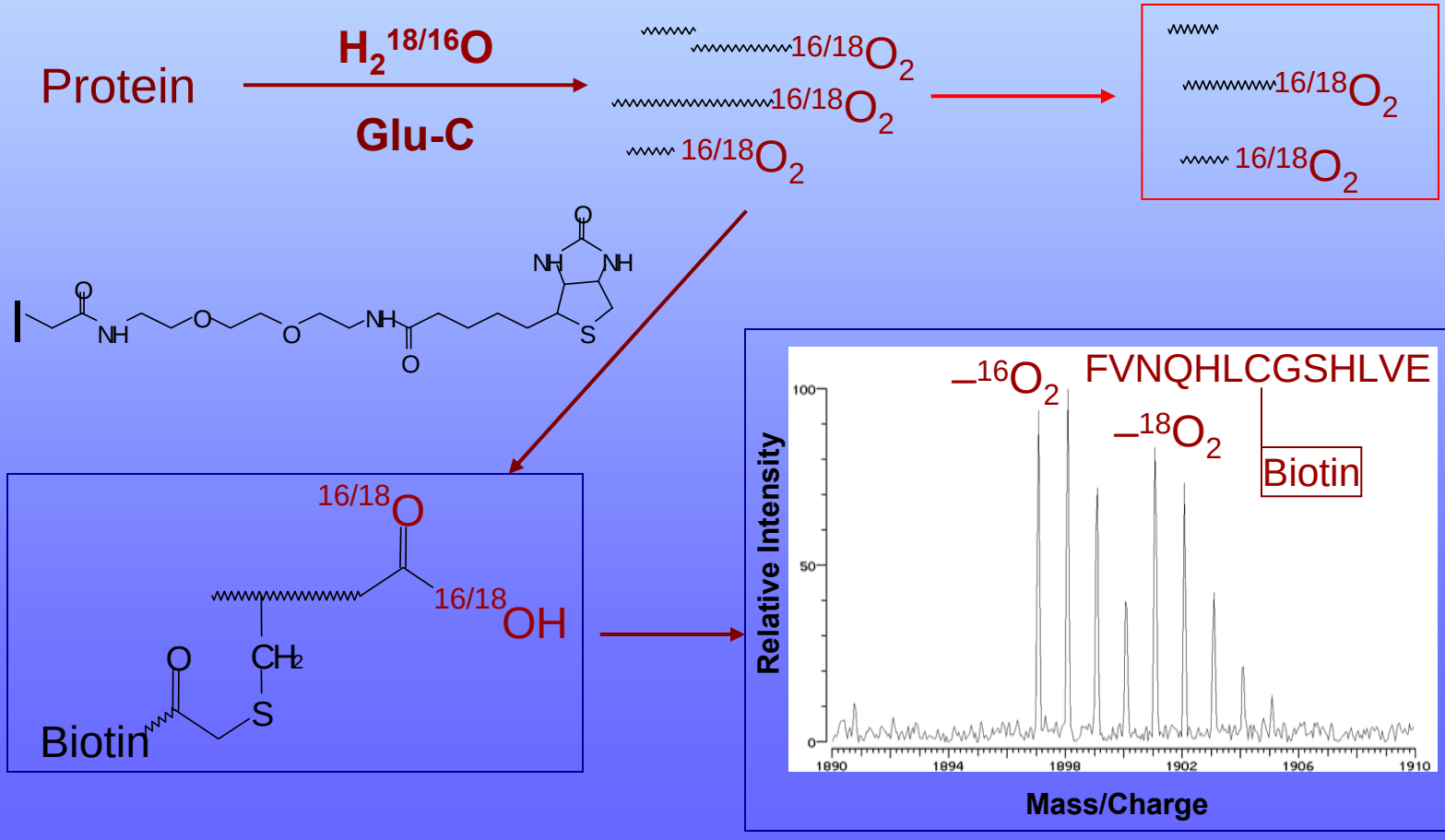
- Unique Chemistry for -SH
- Affinity Tag
- Isotope-Coded Linker

Gygi *et al. Nat Biotechnol.* **1999**

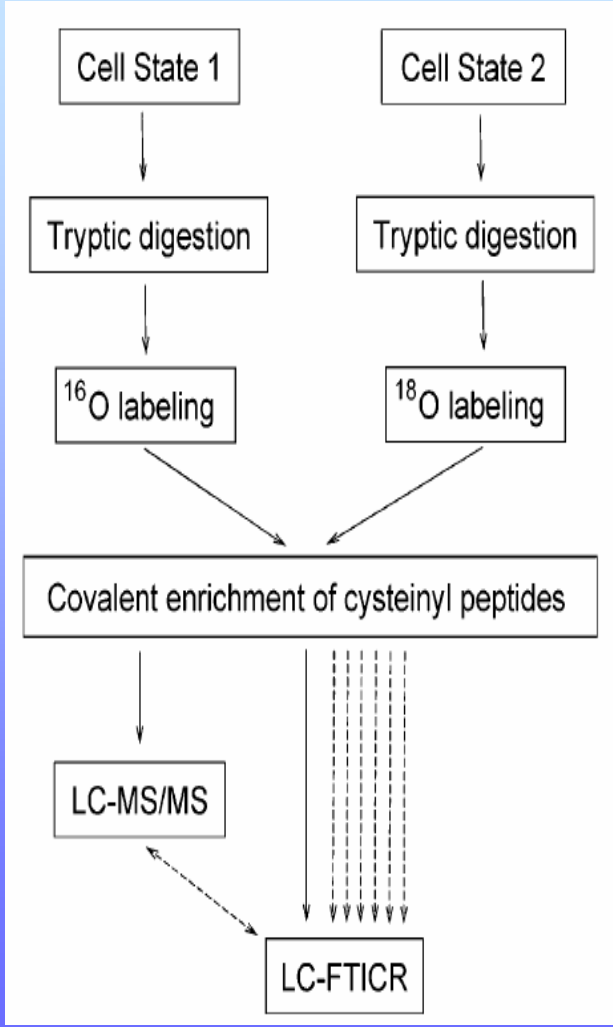


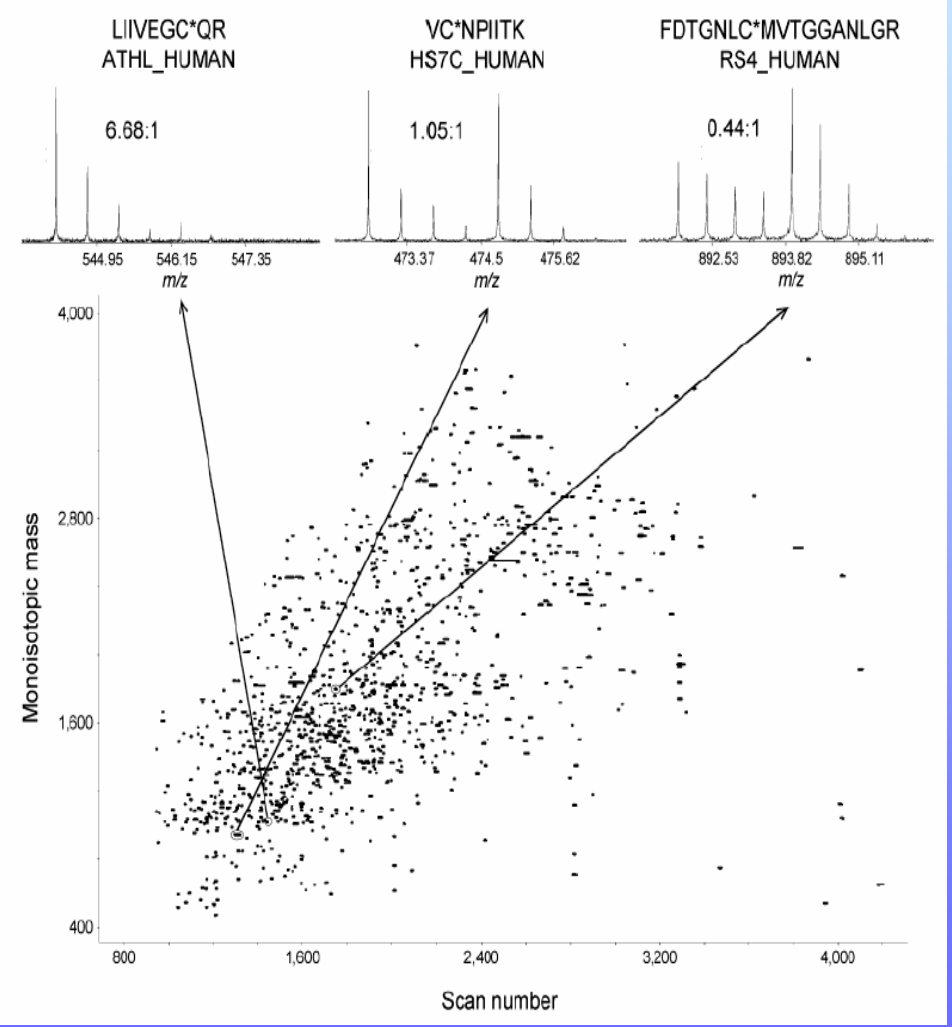
Assembling Separation Module and Coding Module





$^{16}O/^{18}O$-Labeling and Affinity Enrichment to Quantitate Proteins





Liu, Qian, Strittmatter, Camp, Anderson, Thrall, Smith. *Anal. Chem.* **2004**, *76*, 5345.