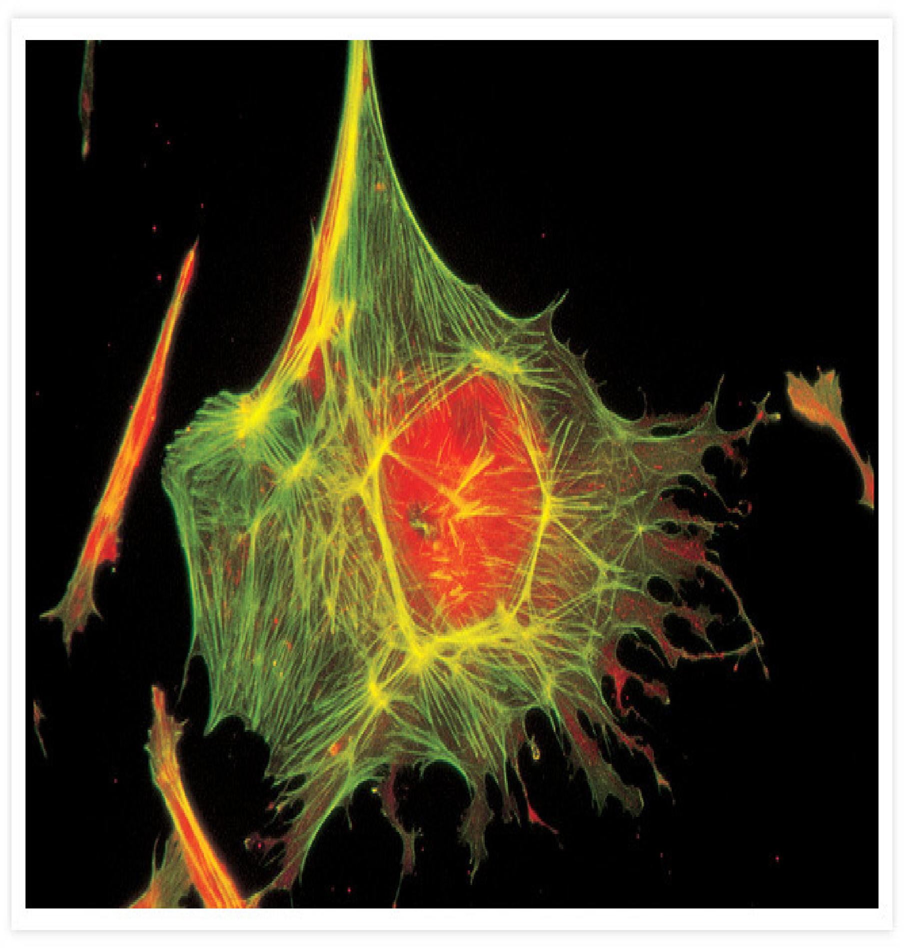
Proteomics/Peptidomics

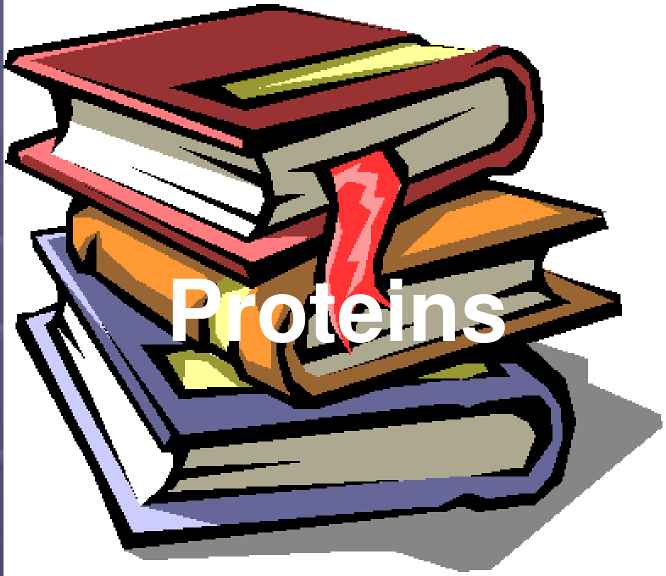
System biology tools and preclinical models for translational research in endometriosis, ESHRE Campus workshop, 4-5 September 2009

E. Waelkens

KATHOLIEKE UNIVERSITEIT
LEUVEN

Proteomics: What?



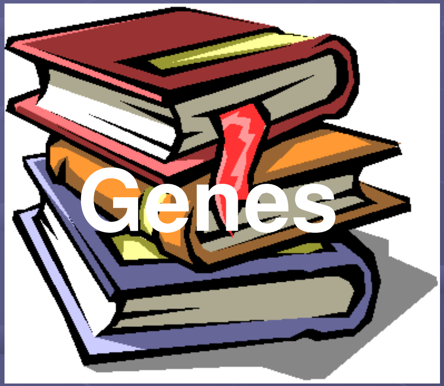


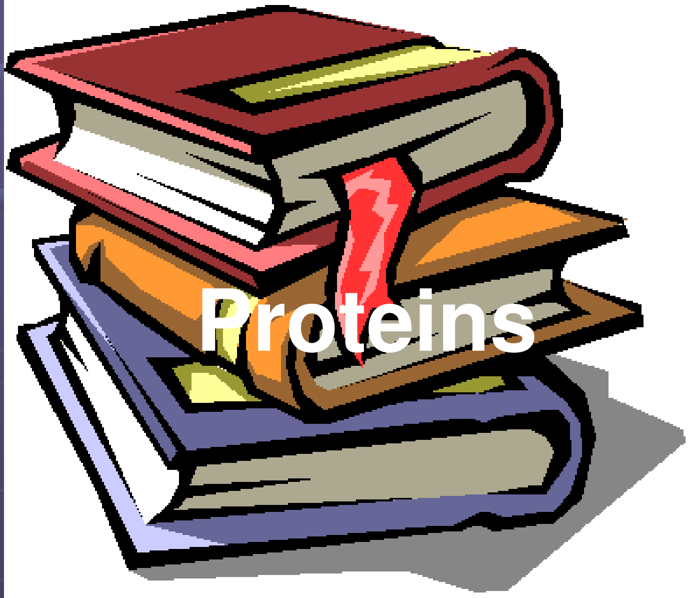
Proteomics = the study of the protein library

- under certain well-defined conditions
- in certain cells/tissues/ organisms

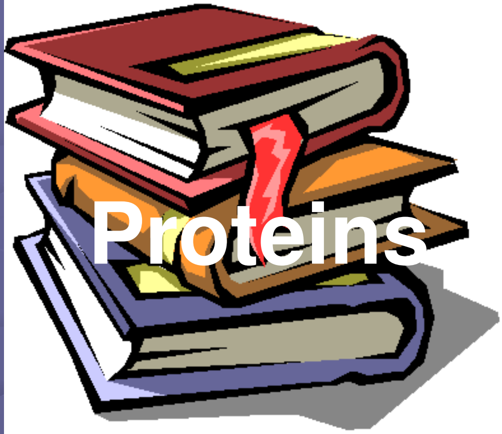
Proteomics: Why ?

Genomics =/=> Proteomics

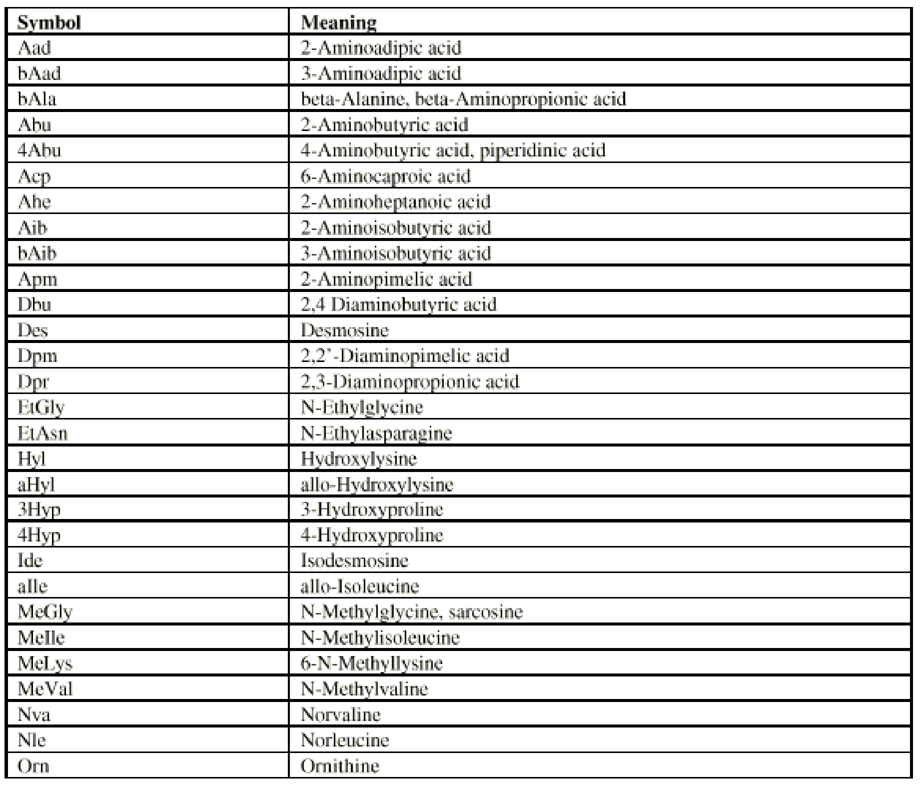




Proteomics



AA level



Symbol	Meaning
Aad	2-Aminoadipic acid
bAad	3-Aminoadipic acid
bAla	beta-Alanine, beta-Aminopropionic acid
Abu	2-Aminobutyric acid
4Abu	4-Aminobutyric acid, piperidinic acid
Acp	6-Aminocaproic acid
Ahe	2-Aminoheptanoic acid
Aib	2-Aminoisobutyric acid
bAib	3-Aminoisobutyric acid
Apm	2-Aminopimelic acid
Dbu	2,4 Diaminobutyric acid
Des	Desmosine
Dpm	2,2'-Diaminopimelic acid
Dpr	2,3-Diaminopropionic acid
EtGly	N-Ethylglycine
EtAsn	N-Ethylasparagine
Hyl	Hydroxylysine
aHyl	allo-Hydroxylysine
3Hyp	3-Hydroxyproline
4Hyp	4-Hydroxyproline
Ide	Isodesmosine
aIle	allo-Isoleucine
MeGly	N-Methylglycine, sarcosine
MeIle	N-Methylisoleucine
MeLys	6-N-Methyllysine
MeVal	N-Methylvaline
Nva	Norvaline
Nle	Norleucine
Orn	Ornithine

Protein level

- •Proteolytic maturation
- •Methionine processing
- •Sugars, Ubiquitination
- •…….

MODIFICATIONS !

Challenges in Proteomics/Peptidomics

- Dynamic changing (not static)
- Labile modifications (eg phosphorylation)
- Handling dependent (sample (pre)treatment !)
- Wide concentration range (no PCR equivalent)
- Localization dependent
- Better technology = more data
- Differential analysis and quantification

Proteomics/Peptidomics: How?

Various ways

examples:

- **(2D)- gel electrophoresis** (old, old fashioned, but it still climbs up the mountains)
- **immunological techniques** [western blot , immunohistochemistry, fluorescence,...]
- **chromatography: UV profiling**
- **fluorescent tagged proteins** [full- length, (sub)cellular localization]
- **Protein (Peptide) arrays**
- **Mass spectrometry**

Proteomics/Peptidomics: How?

Thanks to 2 “soft ionization” techniques
(ESI and MALDI)

Mass spectrometry

Flow Chart: Proteomics/Peptidomics

Separation

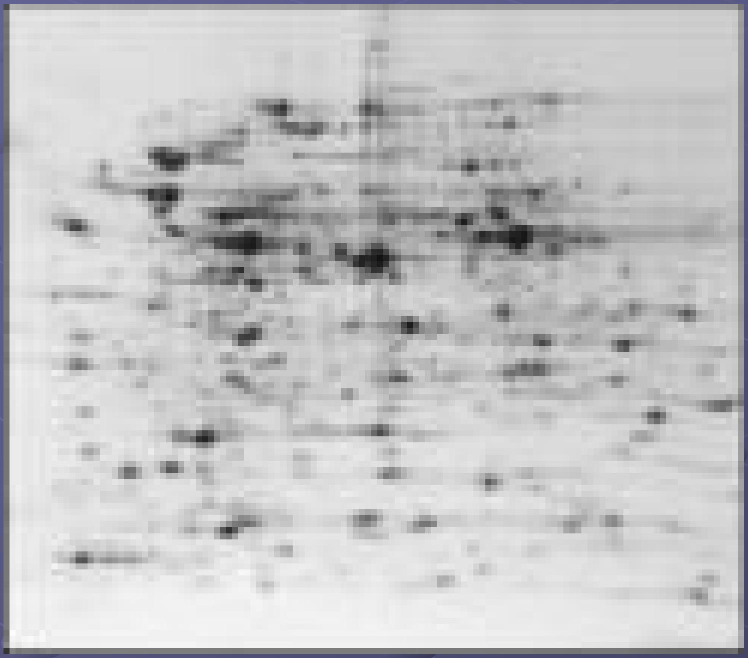
Analysis

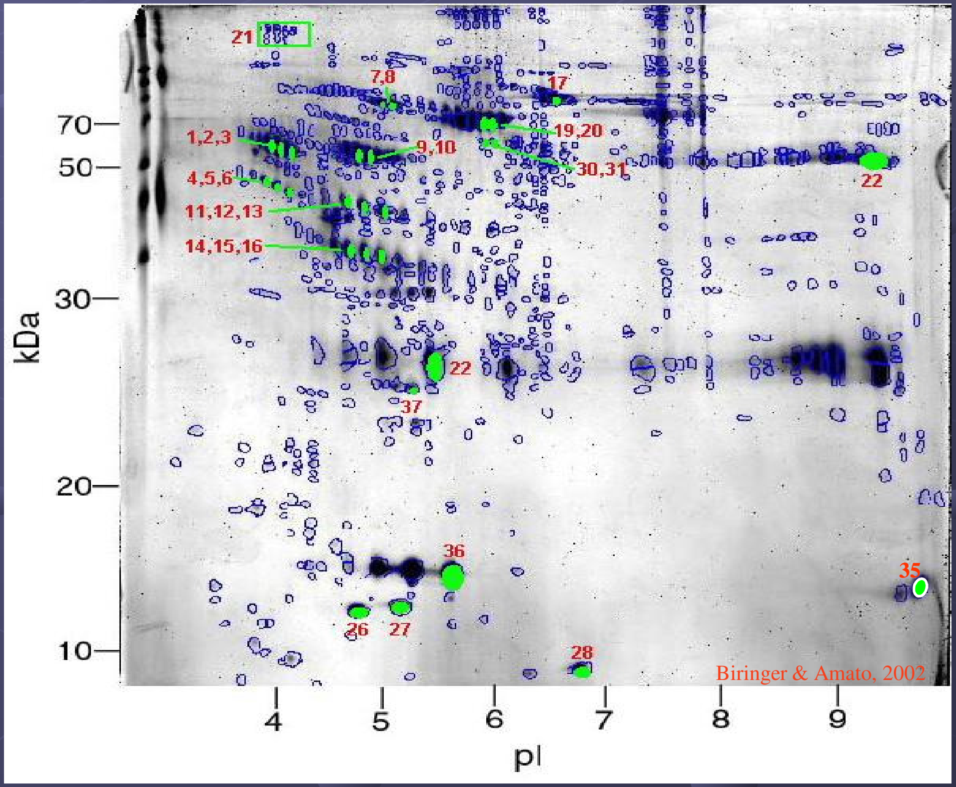
Data processing

Separation

GEL based separations

- 1D Gel (SDS-PAGE)
- 2D GEL -> 2D DIGE



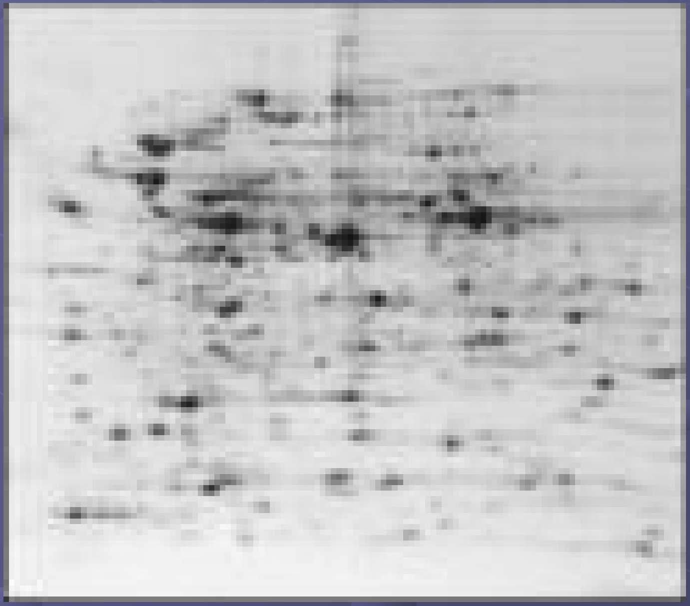


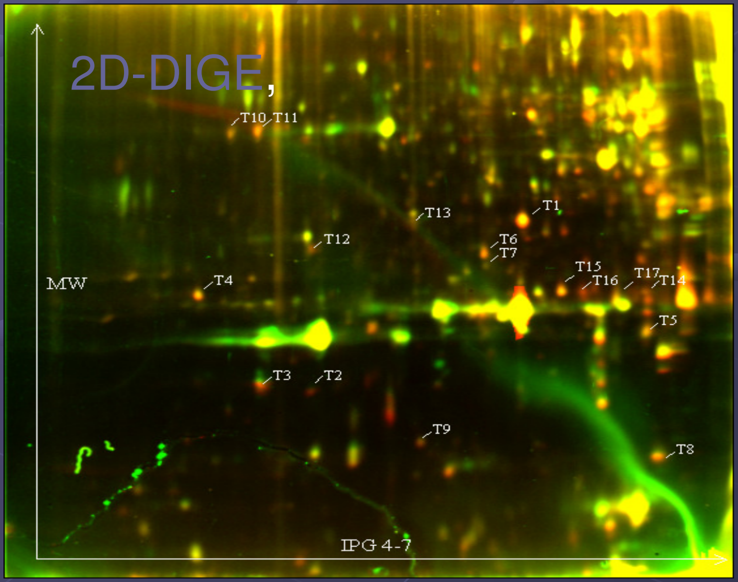
Biringer & Amato, 2002

Separation

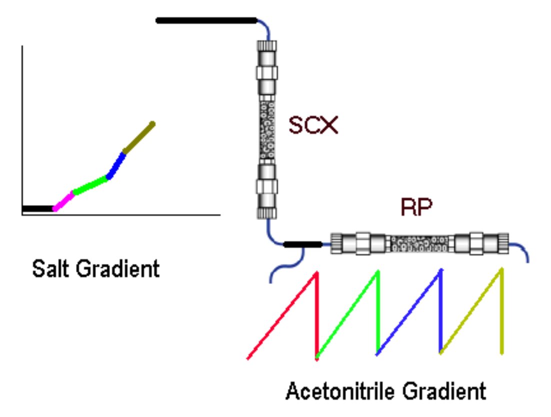
GEL based separations

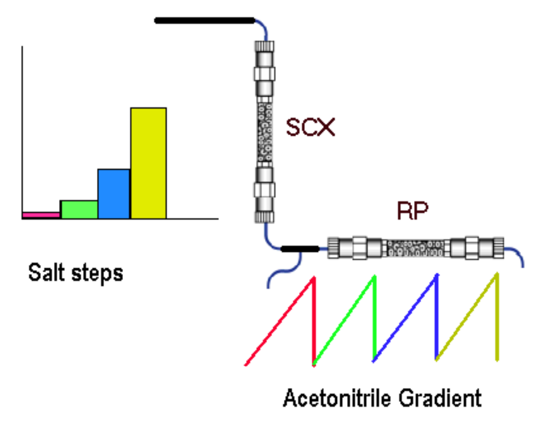
- **2D DIGE = labeling with fluorescing cyanine dyes:**
  Cy2 (green) Cy3 (orange) Cy5 (red)





Separation: 2D LC





- 2D –Gel ←→ 2D LC
- massive amount of data when linked to MS

Separation

- Special affinity based separations (special affinity matrixes: eg for serum samples/ membrane proteins/ phoshorylated peptides, antibodies based separations, dendrimer capture). Popular application: magnetic beads

- solid phase extraction

- new separations in the MS analyser itself: ion mobility

- new LC based separations:

  - Advion Robot
  - UPLC chip based separations
  - robotic MALDI MS spotters
  - mixed functional phase columns

Serum Proteomics

- high dynamic range of protein concentrations
- dominated by a limited number of high-abundance proteins, constituting up 95–99% of the total protein

Depletion methods (2-20 proteins):

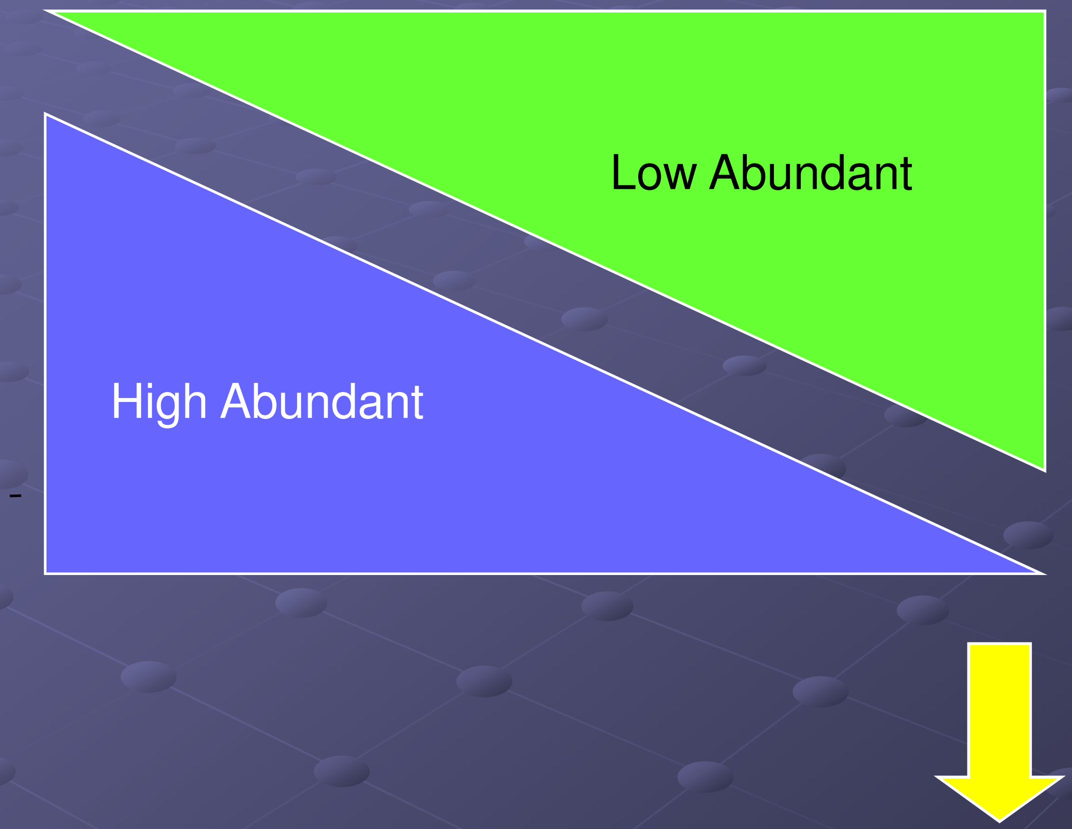
Immunodepletion methods

Affinity methods

- old methods: Cibacron Blue (Blue Sepharose): albumin; Protein G (A): IgG

- new methods: combinatorial chemistry: eg short peptide fragments (hexapeptides)

Serum Proteomics



Up to 6 fold increase in identification !

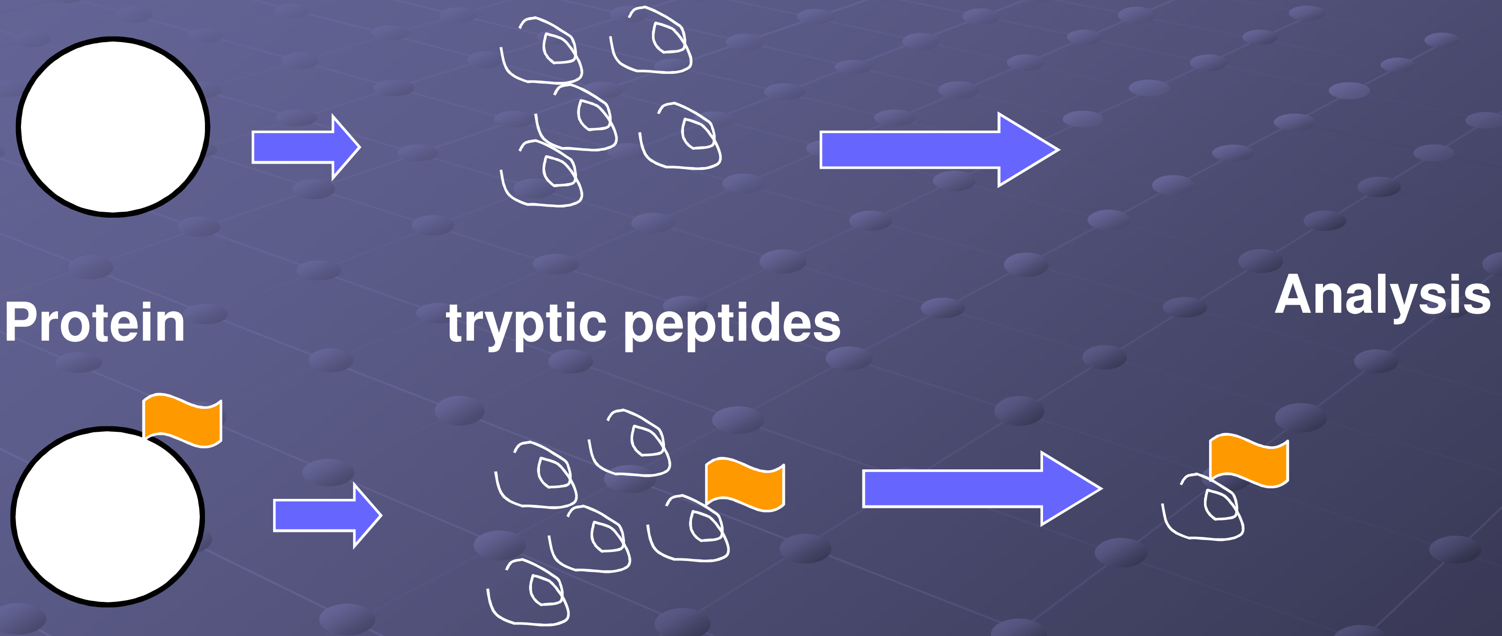
Separation

MORE SEPARATION = MORE IDENTIFICATION

→ **Combine (pre-) fractionation steps: immuno-affinity based, subcellular fractionation, serum depletion kits, ….**

Special Separation

- label proteins at rarely occurring AA’s (eg cysteine or methionine)
- analyze labeled peptides only --> reducing complexity



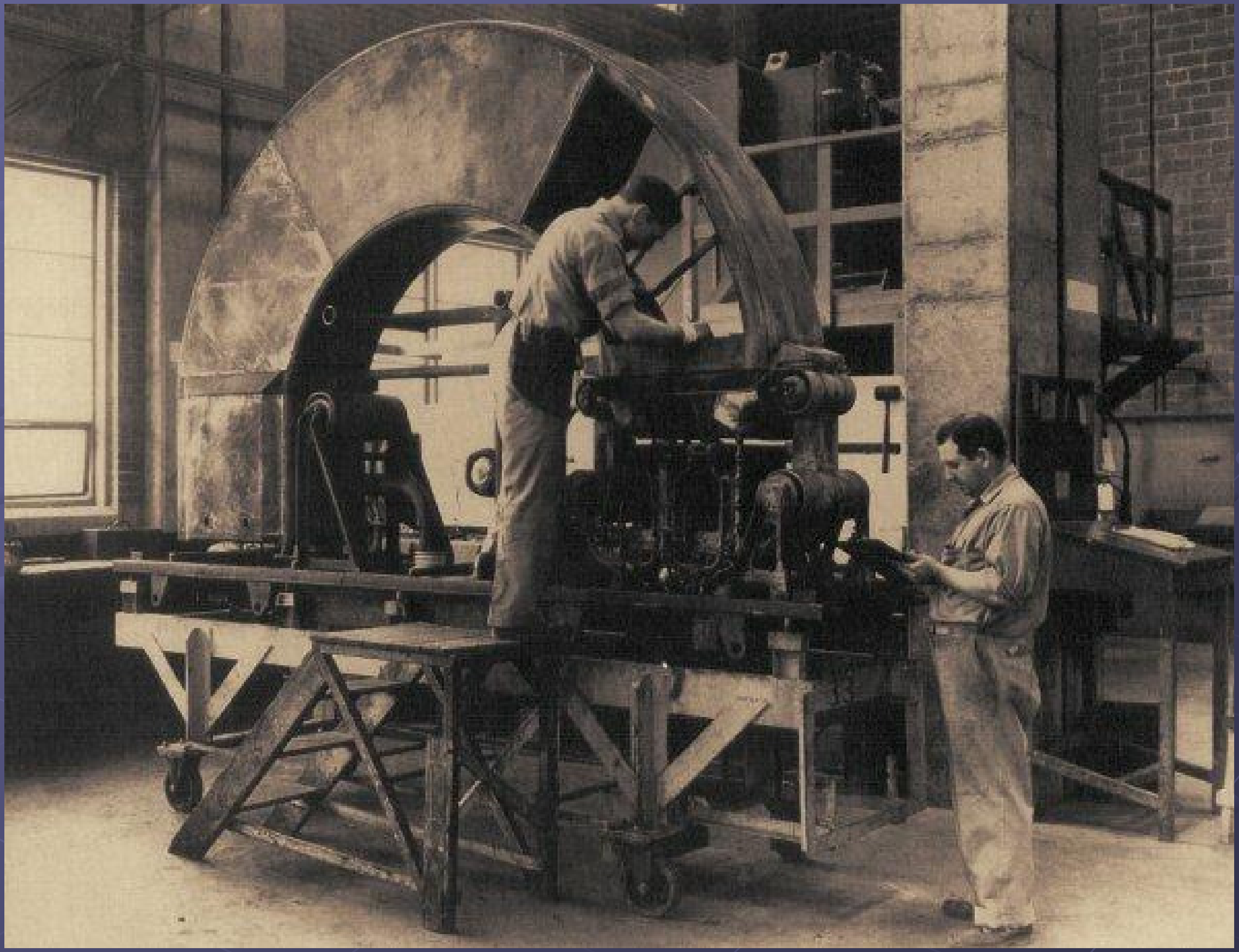
Protein

tryptic peptides

Analysis

ICAT (isotope coded affinity tag) , COFRADIC

Analysis: History



obtained from:
Yergey A.L., Yergey A.K.
Preparative Scale Mass Spectrometry: A Brief History of the Calutron
JASMS, **1997**, V8(N9), p943-953

Analysis

Huge improvement of mass spectrometers:

- end 80’s: Soft ionization methods : ESI and MALDI (Nobel price 2002)
- end 90’s: hybrid MS instruments, nanospray, MALDI TOF/TOF

- begin 21 century: exponential growth of MS equipment:

TOF/TOF with real CID, linear iontrap, orbitrap, 9.4 Tesla FT ICR MS, ECD and ETD dissociation, new hybrids (eg MALDI-Iontrap; 3Q-Iontrap..), ion mobility

- CPU power

Milestones

1897: Sir J.J Thomson, Cavendish Laboratory , University of Cambridge, discovery of the electron

1906: Thomson: Nobel Prize for this studies on the conduction of electricity by gasses (ion movements)

1912: Thompson: First mass spectrometer

1946: William E. Stephens: concept of time-of-flight analyzer

1953-58: Wolfgang Paul: quadrupole analyzer

1983: first commercial ion trap

1989: Paul: Nobel Prize in Physics

1968: Malcolm Dole: first concept of ESI

1974: Comisarow & Marshall: FTMS

1984-88: John Fenn: ESI for biomolecule analysis

Wong, S.F., Meng, C.K. and Fenn, J.B., J. Phys. Chem., 92, 546 (1988)
Meng, C.K., Mann, M. and Fenn, J.B., Z. Phys.D., 10, 361 (1988)

1988: Tanaka (Japan) and Franz Hillenkamp/Michael Karas (Germany): MALDI

M. Karas and F. Hillenkamp, Anal. Chem. 60, 2299-2301 (1988)
K. Tanaka et al., Rapid Commun. Mass Spectrom. 2, 151-153 (1988)

Analysis: MS

Turbo Engine behind Proteomics/Peptidomics = Mass Spectrometry (MS)

What should we know about MS ?

1. create ions : ionization
2. separate the ions : analyzer
3. measure the ions : detector

Basic MS instrument

source

analyzer

detector

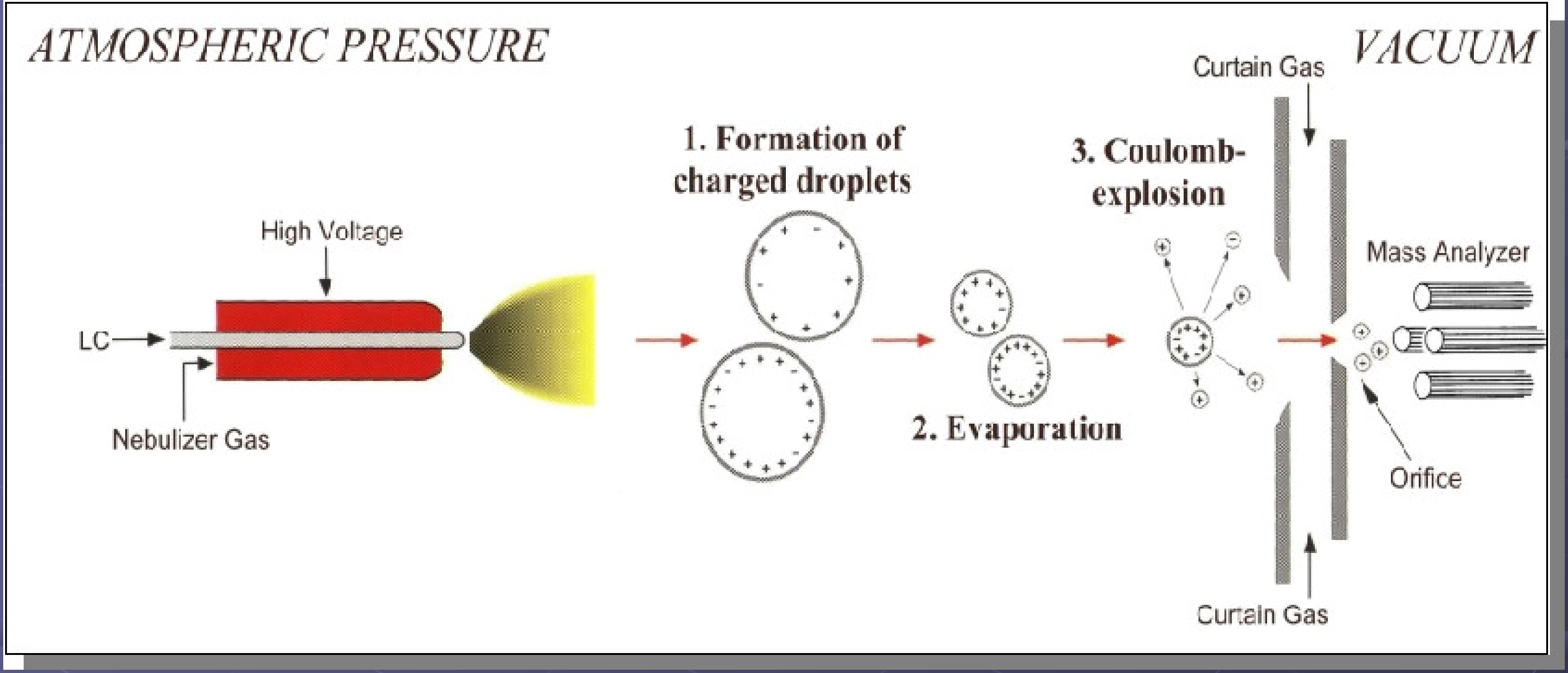
ESI or MALDI

e.g. IonTrap
or TOF (time of flight)

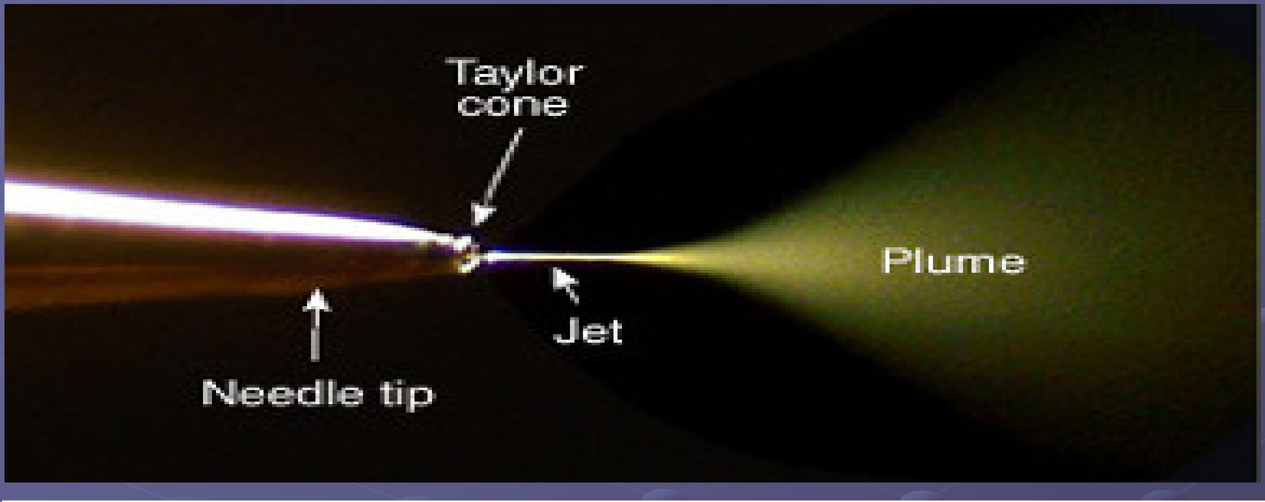
Examples:

- ESI IonTrap
- MALDI TOF

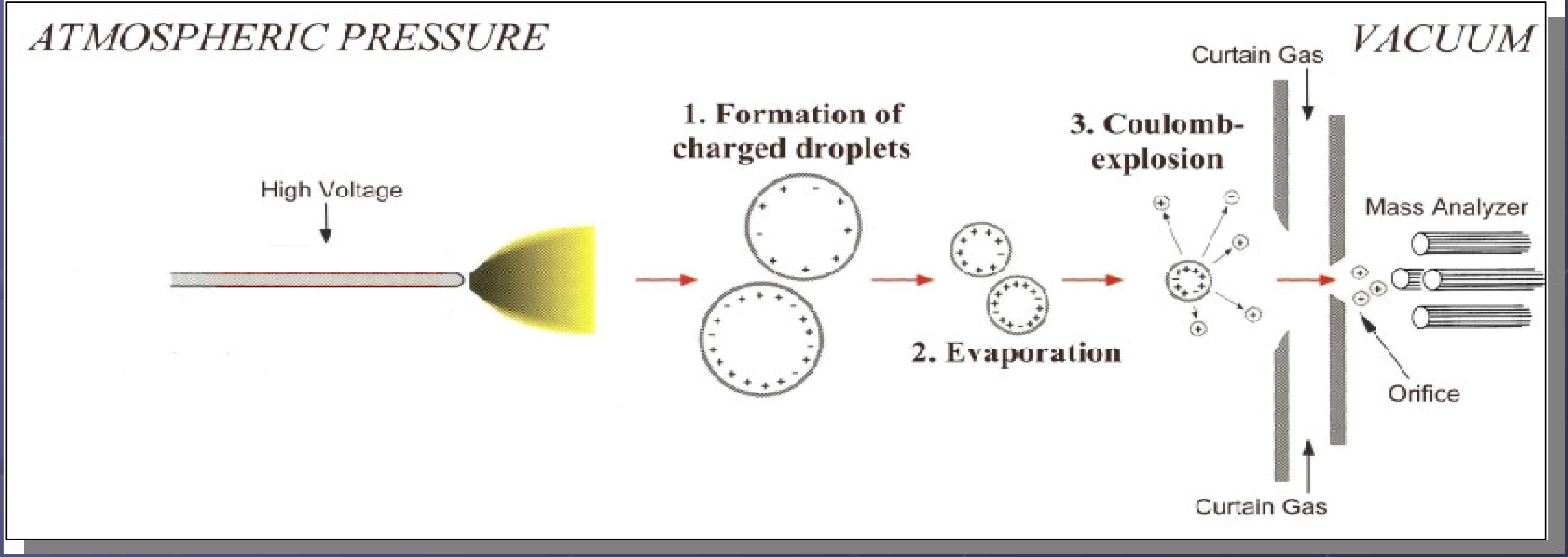
Electrospray Ionization (ESI)



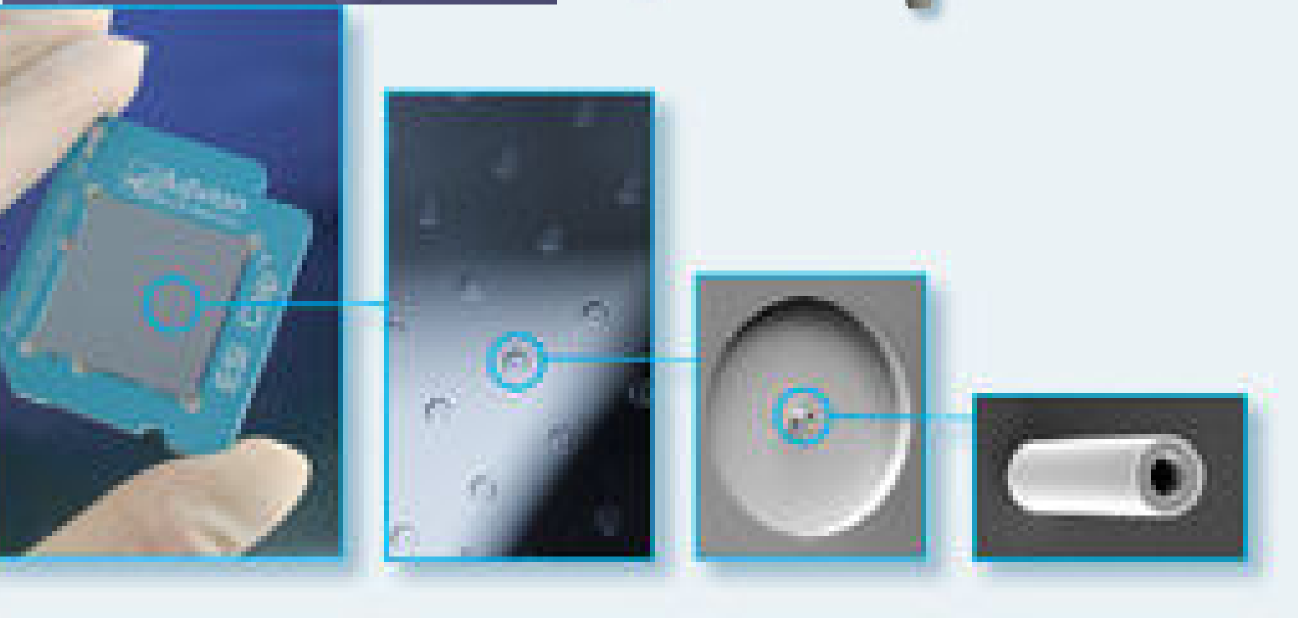
Special ESI: Nanospray



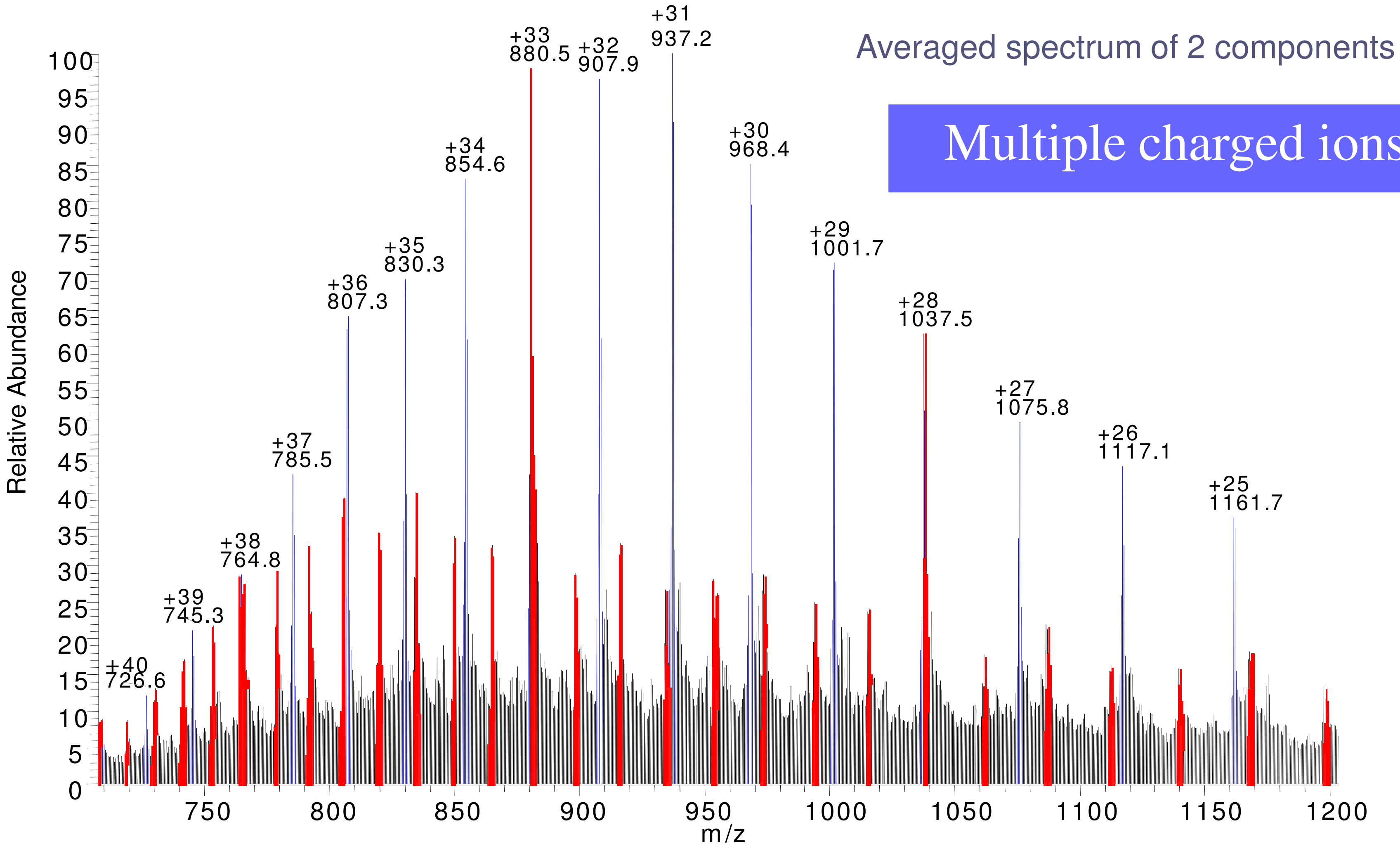
Wilm, 1991



Special ESI: Nanospray

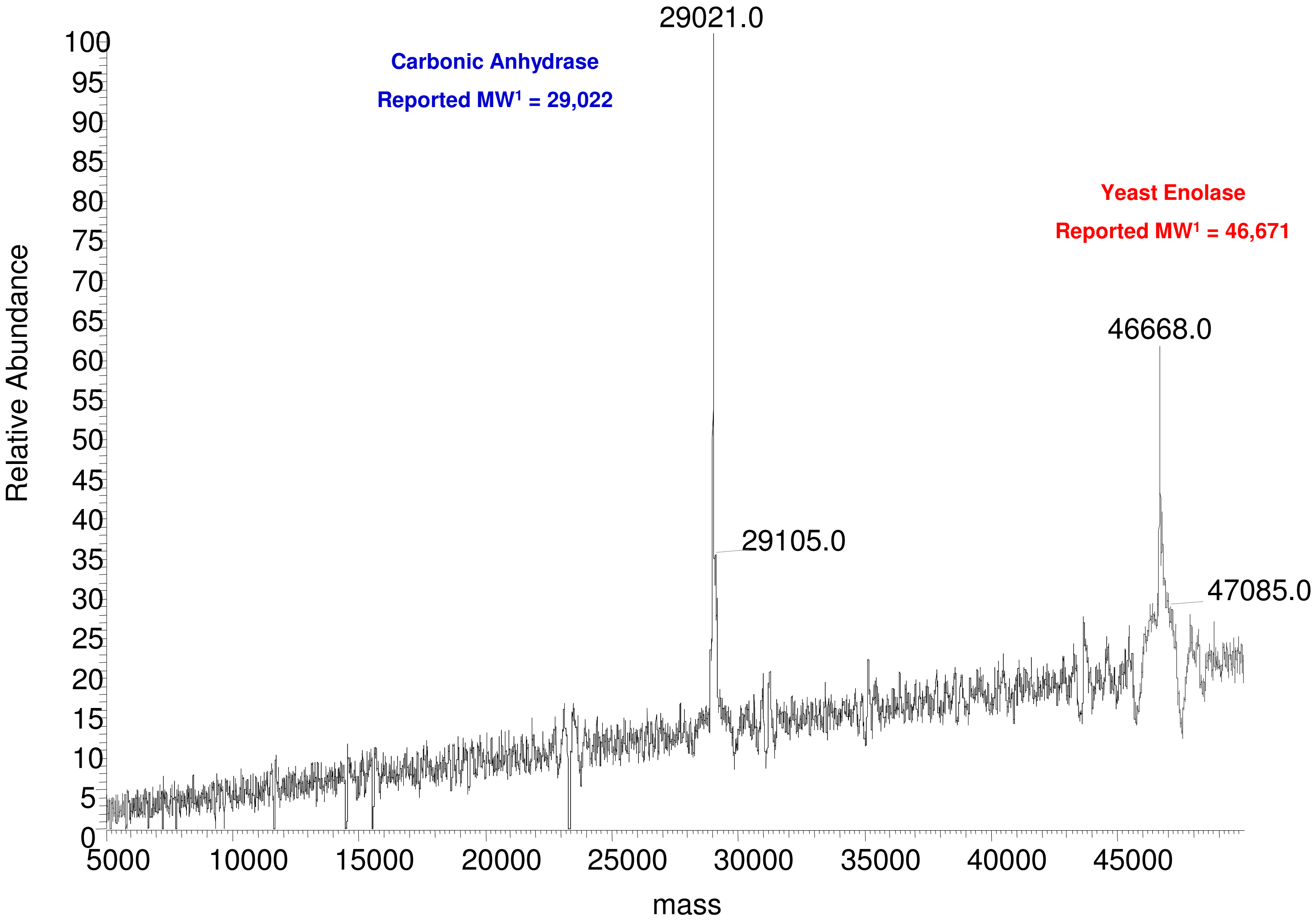


Electrospray Ionization (ESI)

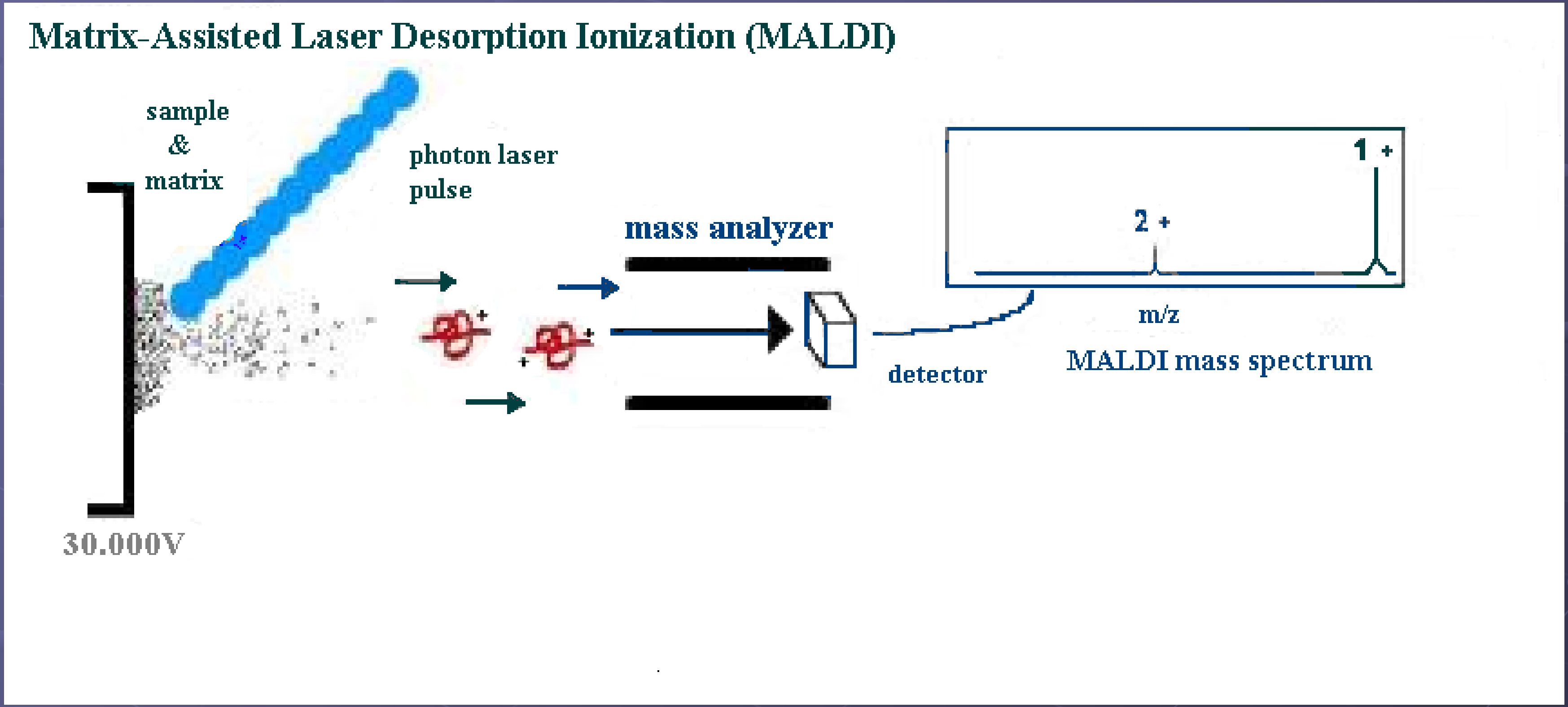


Multiple charged ions

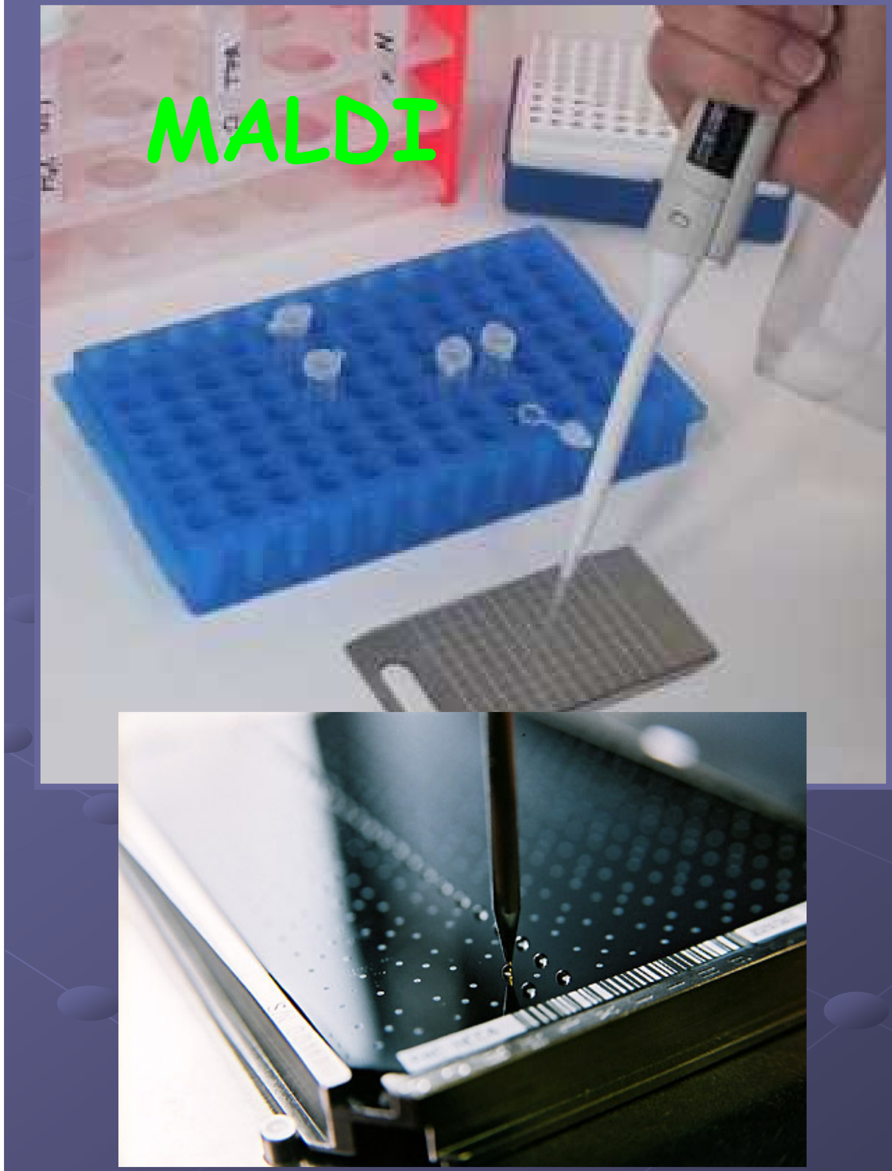
Deconvolution



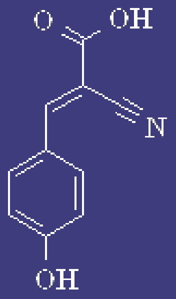
MALDI ionization



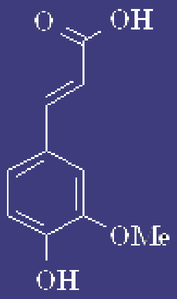
MALDI ionization



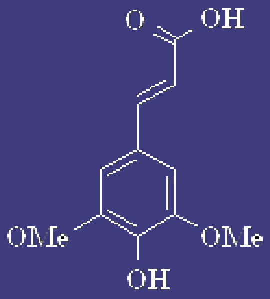
α–cyano-4-hydroxycinnamic acid



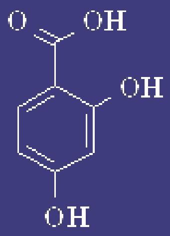
3-methoxy-4-hydroxycinnamic acid



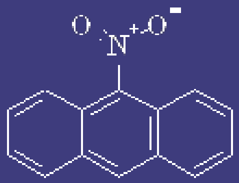
3,5-dimethoxy-4-hydroxycinnamic acid



2,5-dihydroxybenzoic acid



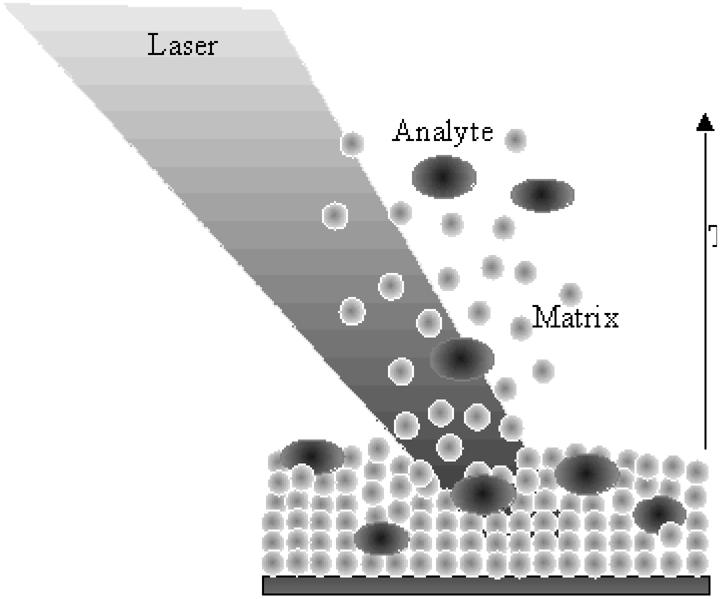
9-nitroanthracene

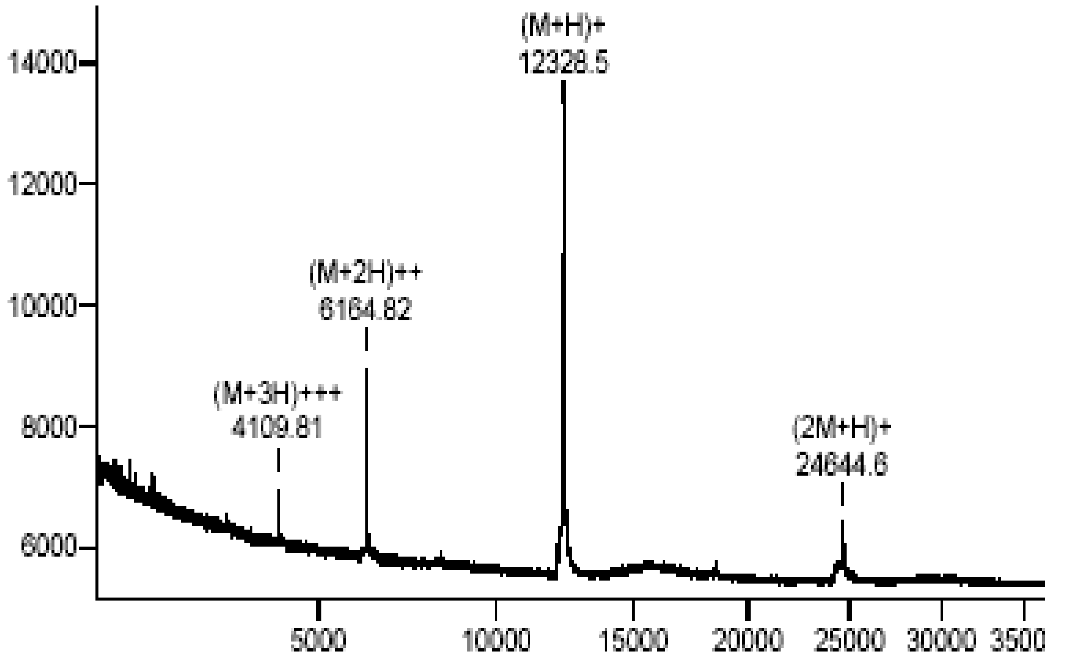


nicotinic acid

MALDI ionization

MALDI

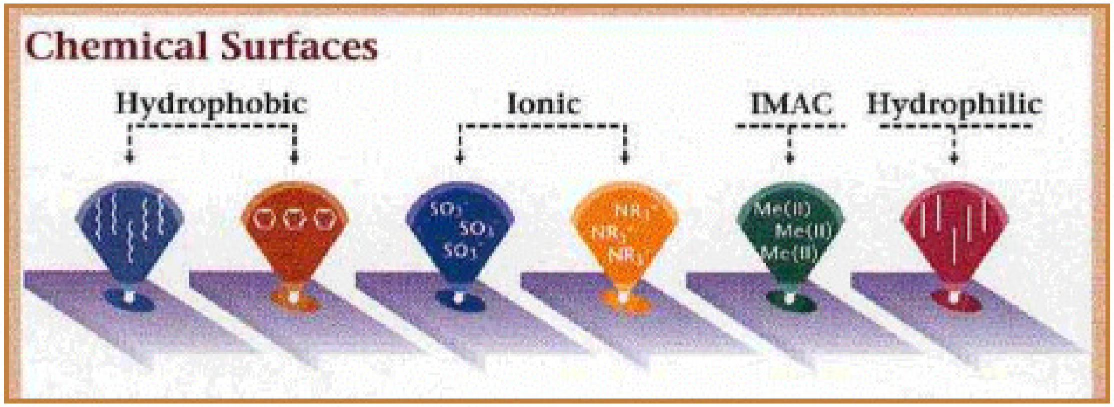




Ions are preferentially +1

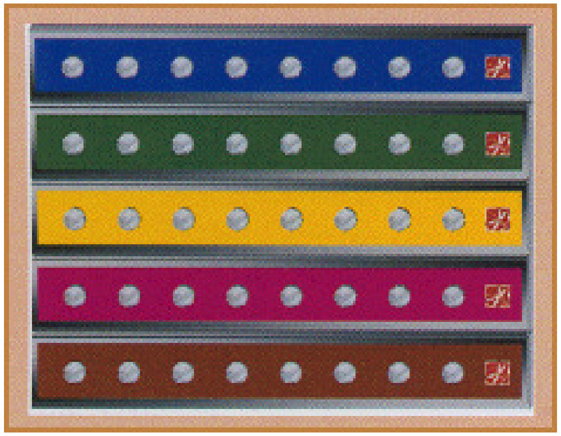
Special MALDI: SELDI

Surface **E**nhanced **L**aser **D**esorption **I**onization



←Surface Chemistries

Each chip binds a specific set of proteins based on the chromatographic surface of the ProteinChip®.



←Protein Chips

Each spot on the chip will contain sera from a control- or toxicant-treated animal. The spots are analyzed separately and a mass spectra is created for each spot representing the proteins bound to the chip surface.

Analyser: separate ions

source **analyzer** **detector**

ESI or MALDI

e.g. IonTrap
or TOF (time of flight)

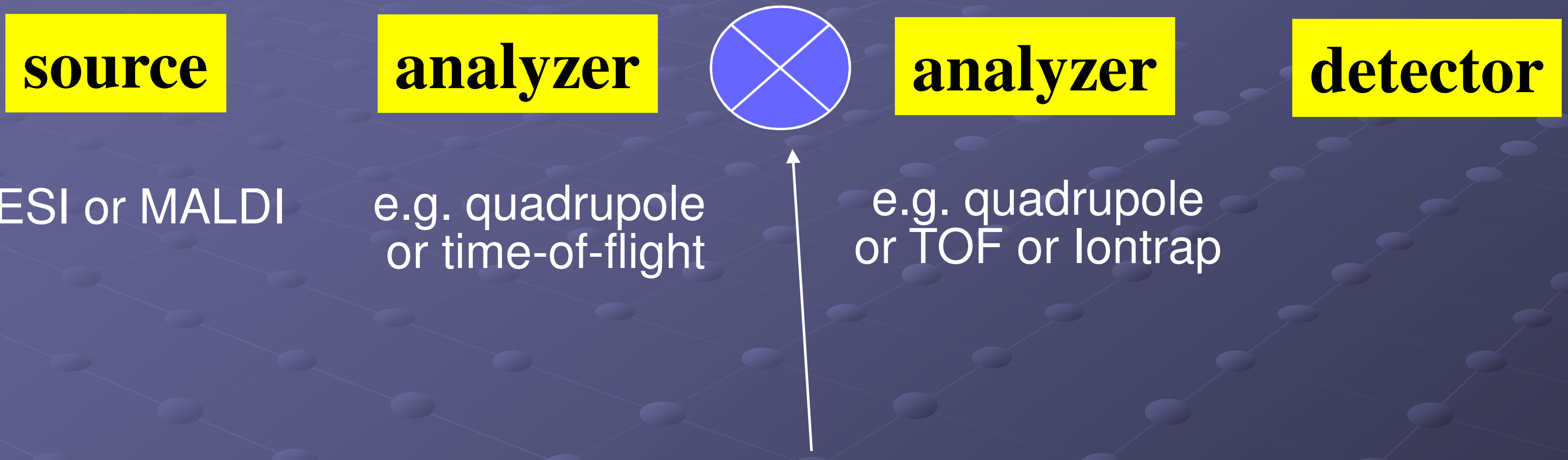
Examples:

- ESI IonTrap
- MALDI TOF

Analyzer: separate ions

- Ion-trap MS
- Quadrupole MS
- Time-of-flight MS
- Fourier-transform MS
- Magnetic-sector MS

More complex MS instruments



source

analyzer

analyzer

detector

ESI or MALDI

e.g. quadrupole
or time-of-flight

e.g. quadrupole
or TOF or Iontrap

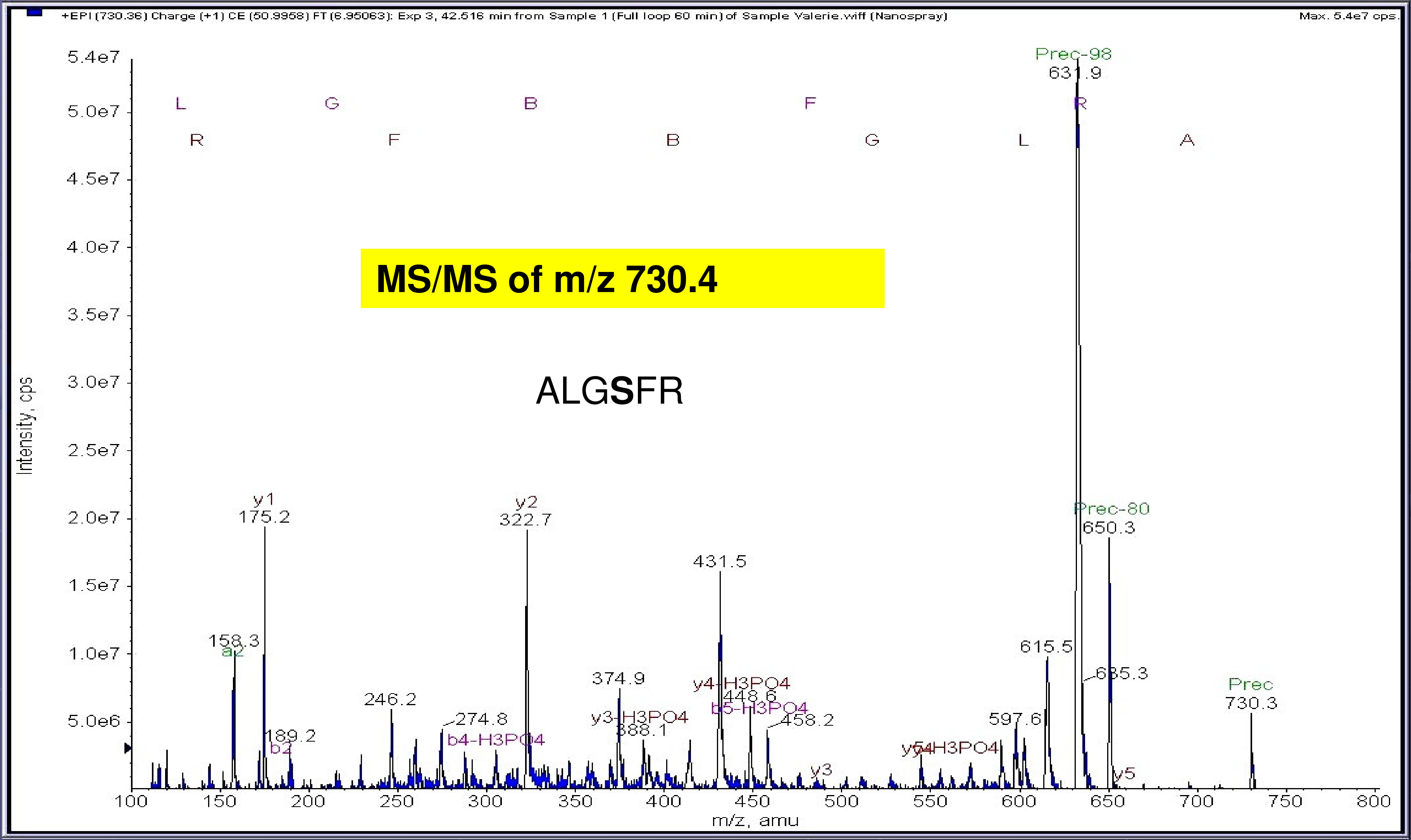
Collision cell

Examples:

- ESI Q TOF
- ESI Q Iontrap
- MALDI Iontrap
- MALDI TOF TOF

MS/MS: AA sequence

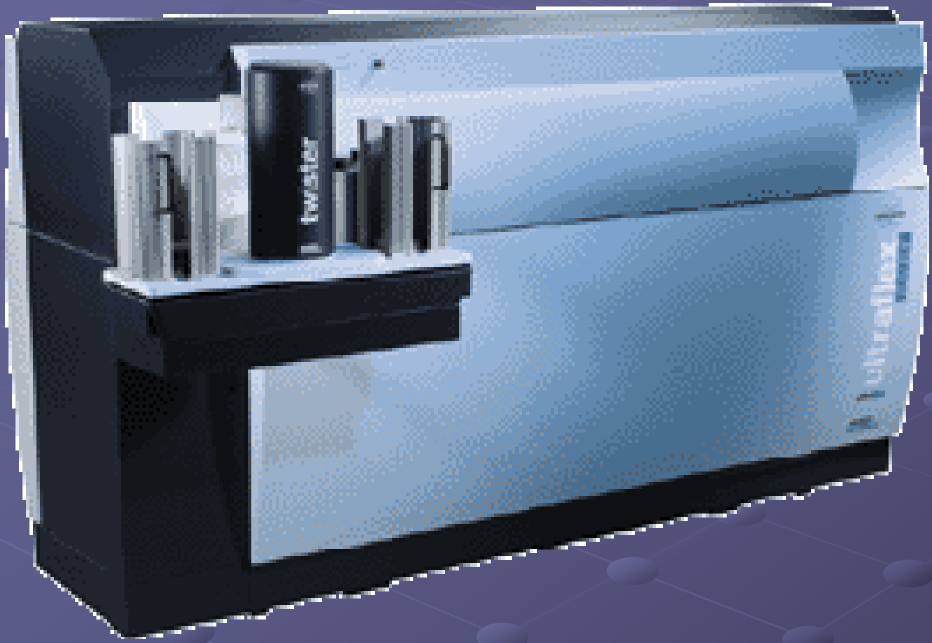
Note: manual interaction or *de novo* sequence



MS/MS of m/z 730.4

ALG**S**FR

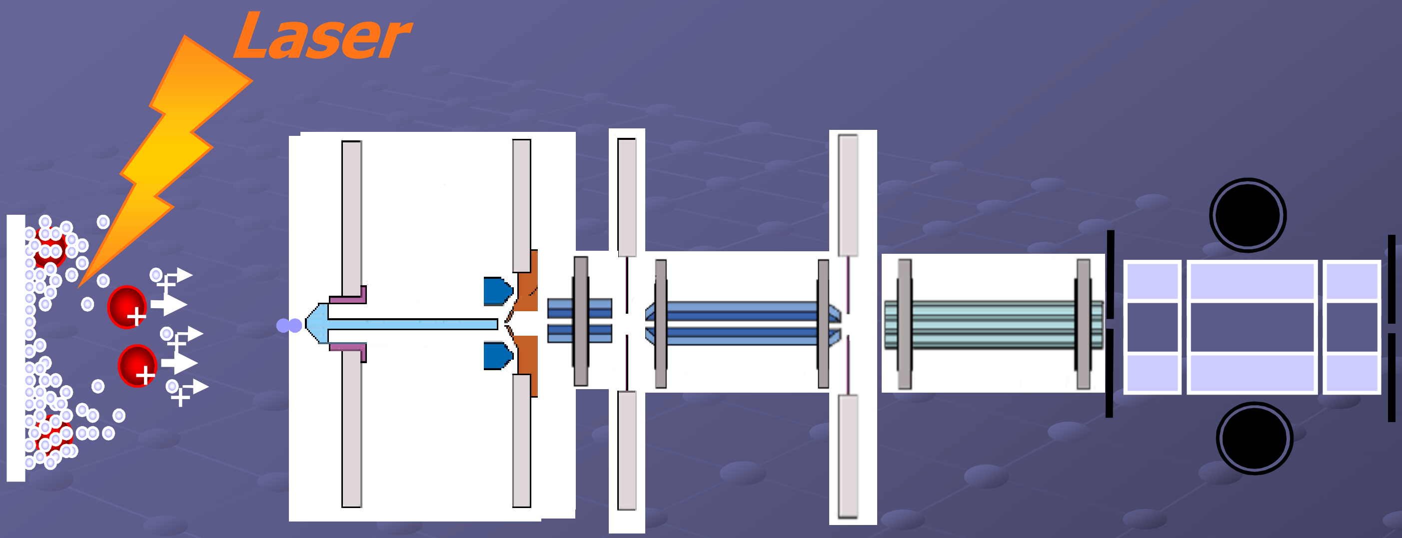
MALDI-TOF/TOF



Triple Q - IonTrap

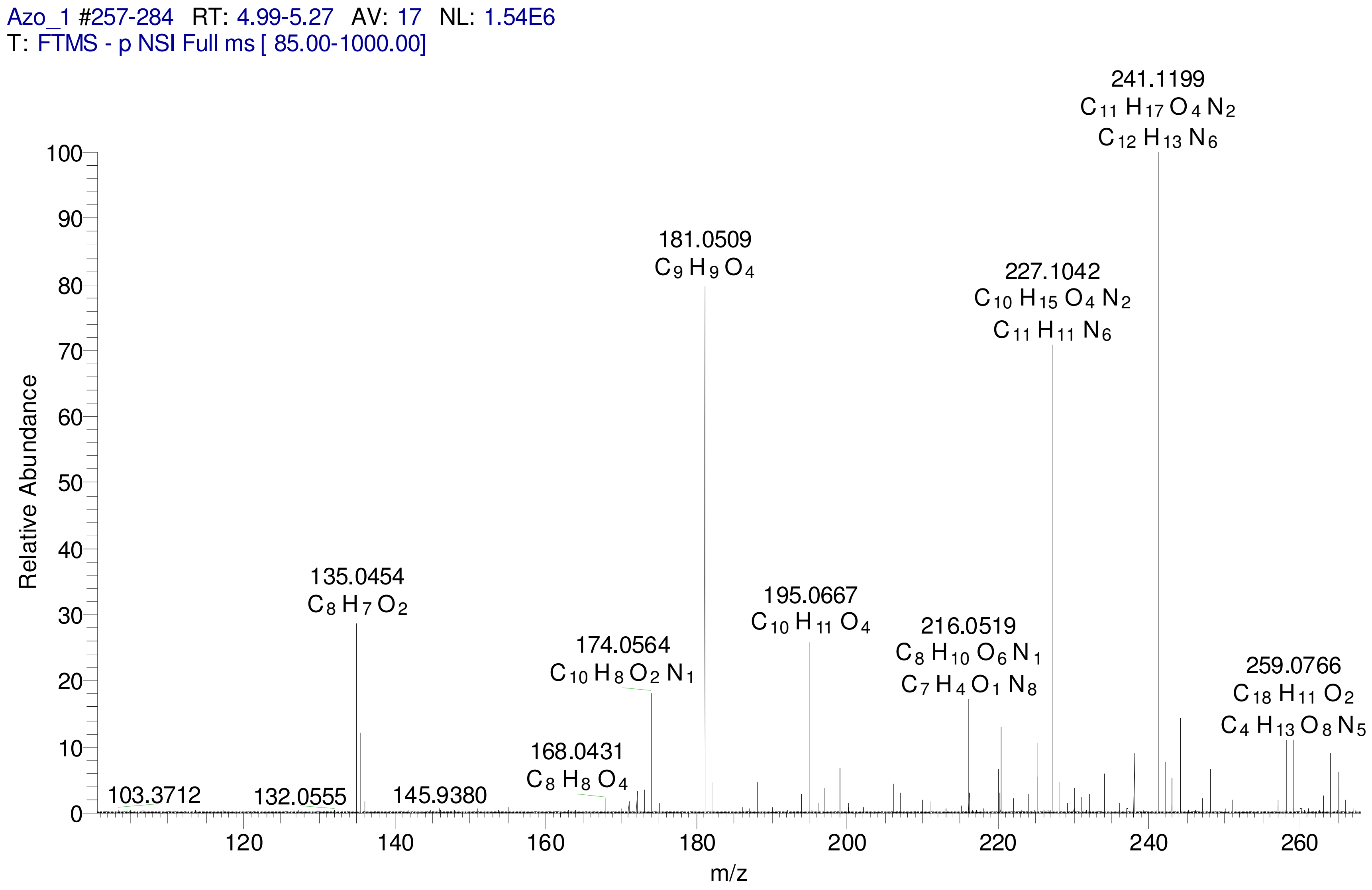
Q - TOF

MALDI-IonTrap



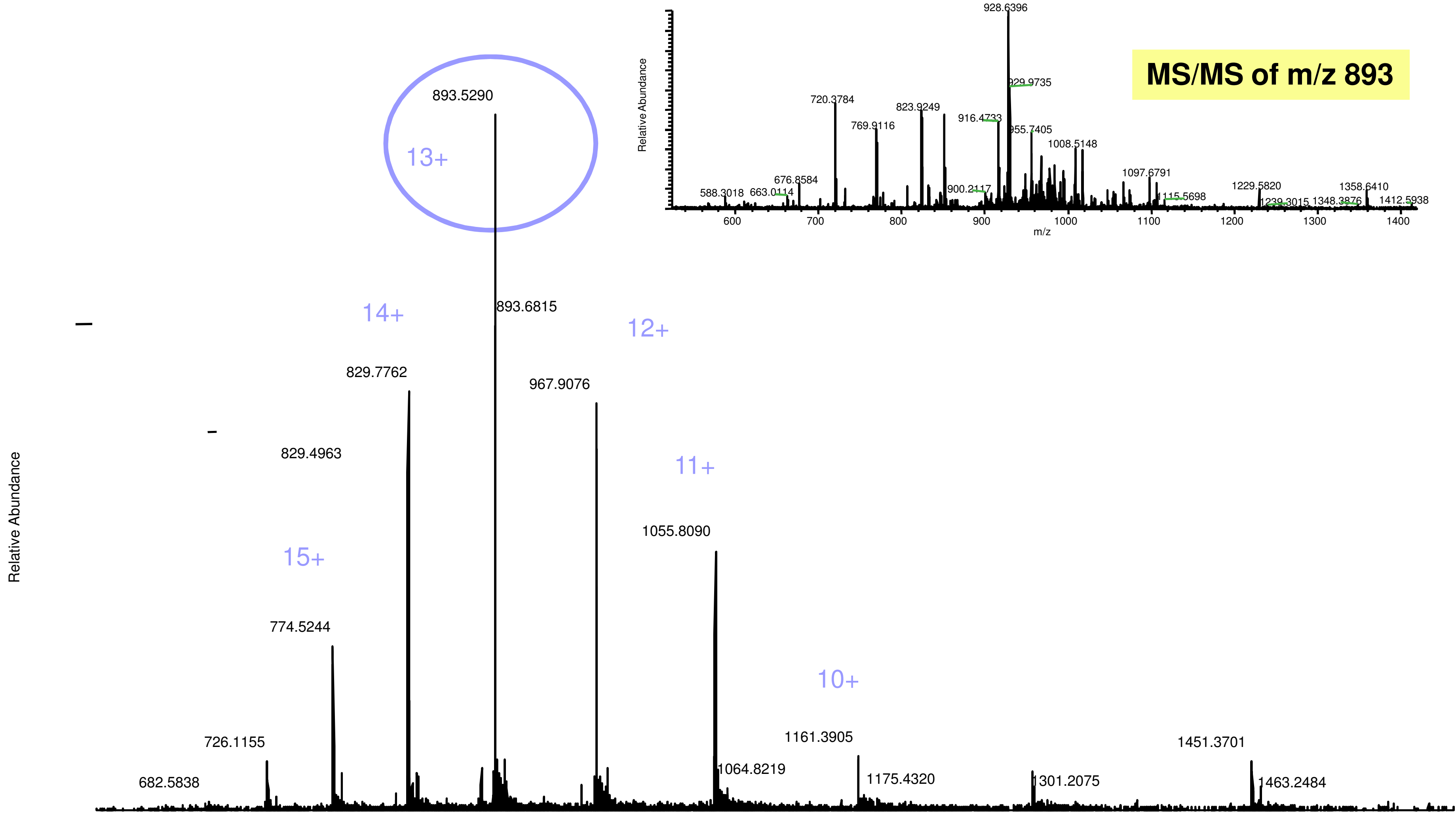
FT-MS

Empirical formula (composition)

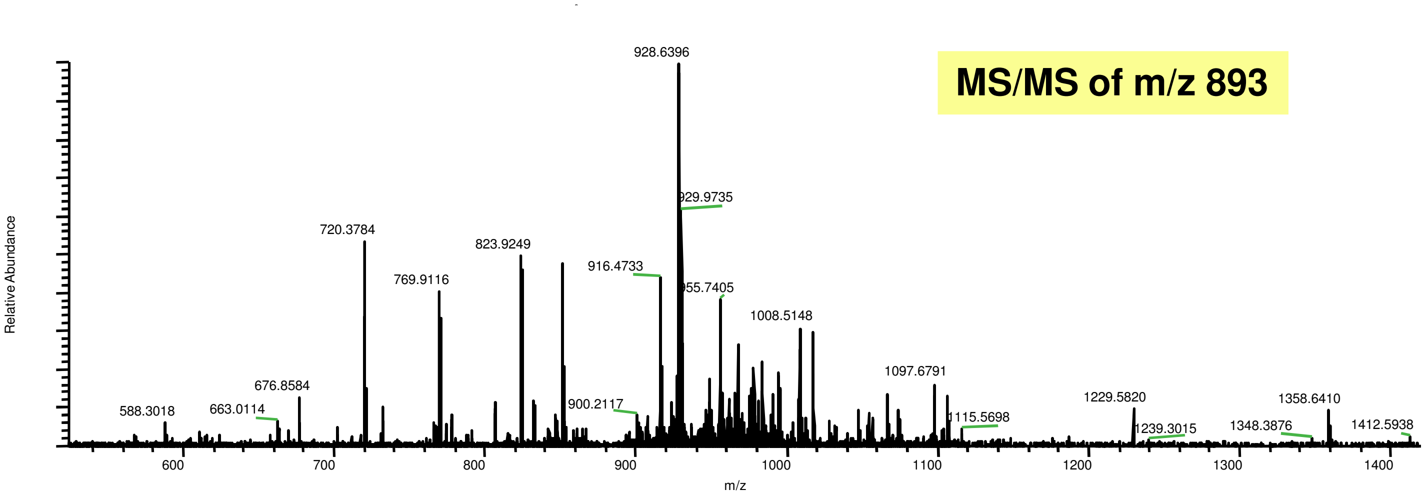


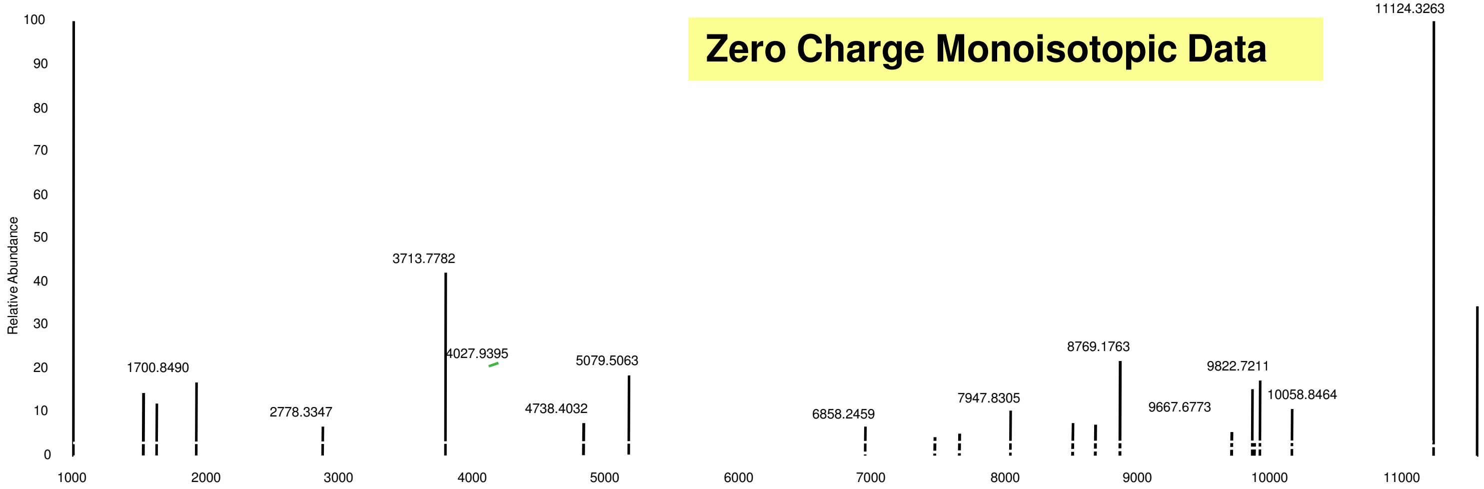
Top Down Analysis

Sequence Analysis of m/z 893



Top Down Analysis





Data Processing

- Protein/peptide identification = search engine to interrogate database(s)
- Find the differences: Biomarker discovery

Clinical Proteomics: Biomarkers

- Diagnostic
- Prognostic
- Therapeutic

Clinical Proteomics: Biomarkers

How to find differences ?

2 different strategies:

Labeling

Label free

Clinical Proteomics: Biomarkers

Labeling

Based on stable isotopes (not radioactive): $C^{12} \rightarrow C^{13}$

$N^{14} \rightarrow N^{15}$

Examples:

SILAC (stable isotope labelling)
ICAT
ITRAQ
O^{18} labelling

} unbiased

AQUA (=targetted approach)

Labeling: ICAT

ICAT (Isotope-coded affinity tag)

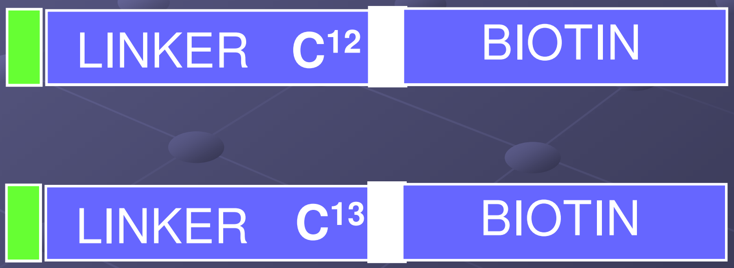
= Duplex Labeling:

2 tags (“light” and “heavy” linker, 9x $\mathbf{C}^{12}$ and 9x $\mathbf{C}^{13}$, Δ mass= 9 Da)

has a biotin tag

labeling = at the protein level (at cysteines)

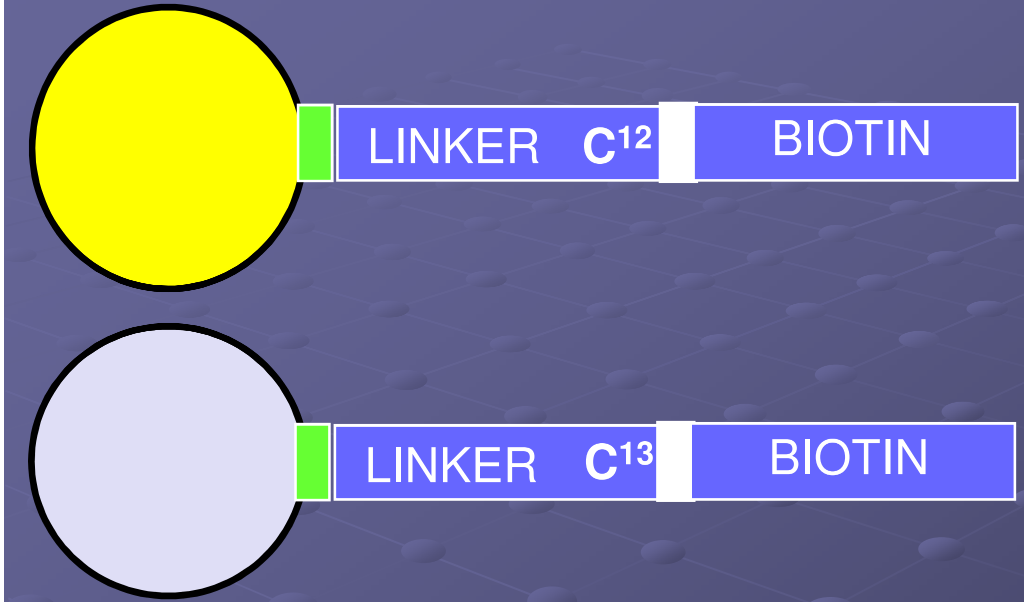
analysis = only labeled peptides → reduced complexity



LINKER $\mathbf{C}^{12}$ BIOTIN

LINKER $\mathbf{C}^{13}$ BIOTIN

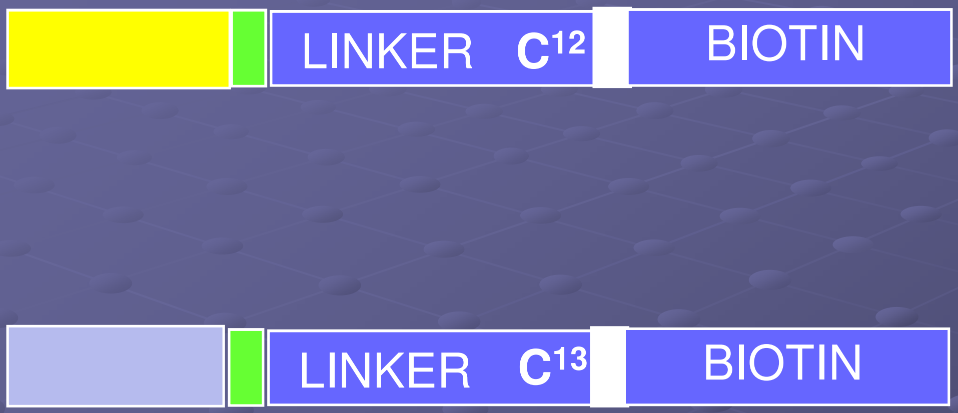
ICAT



Step 1: label proteins

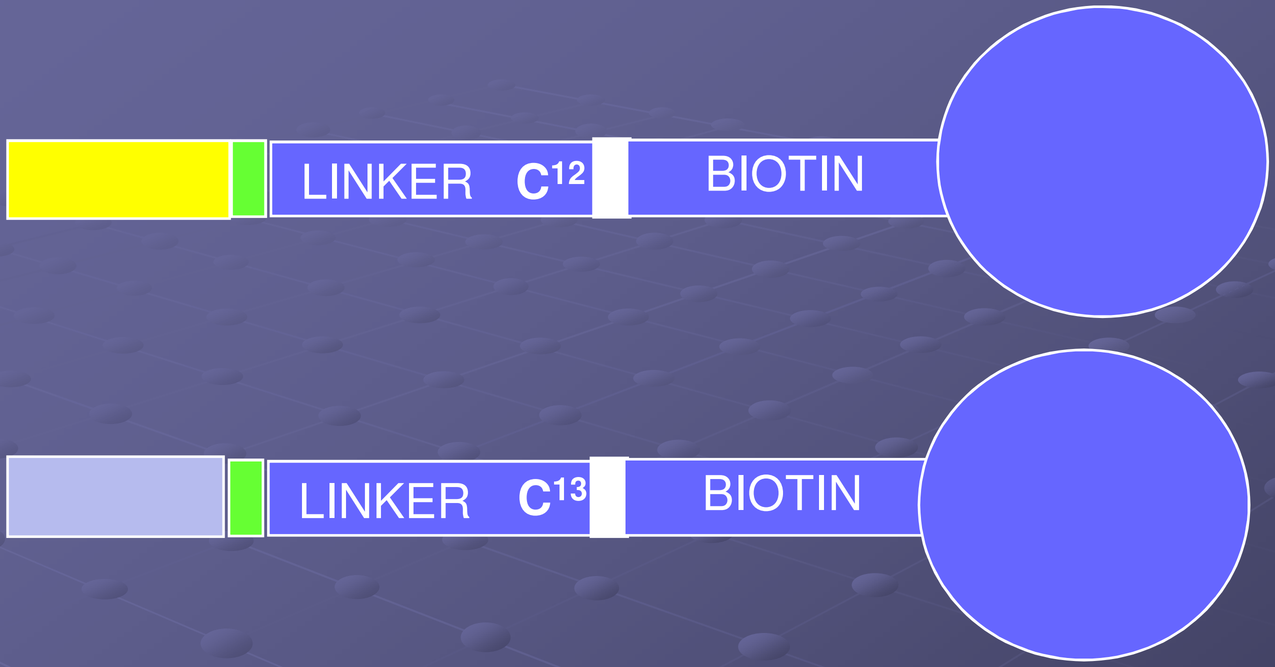
Step 2: digest proteins

ICAT



Step 2: digest proteins

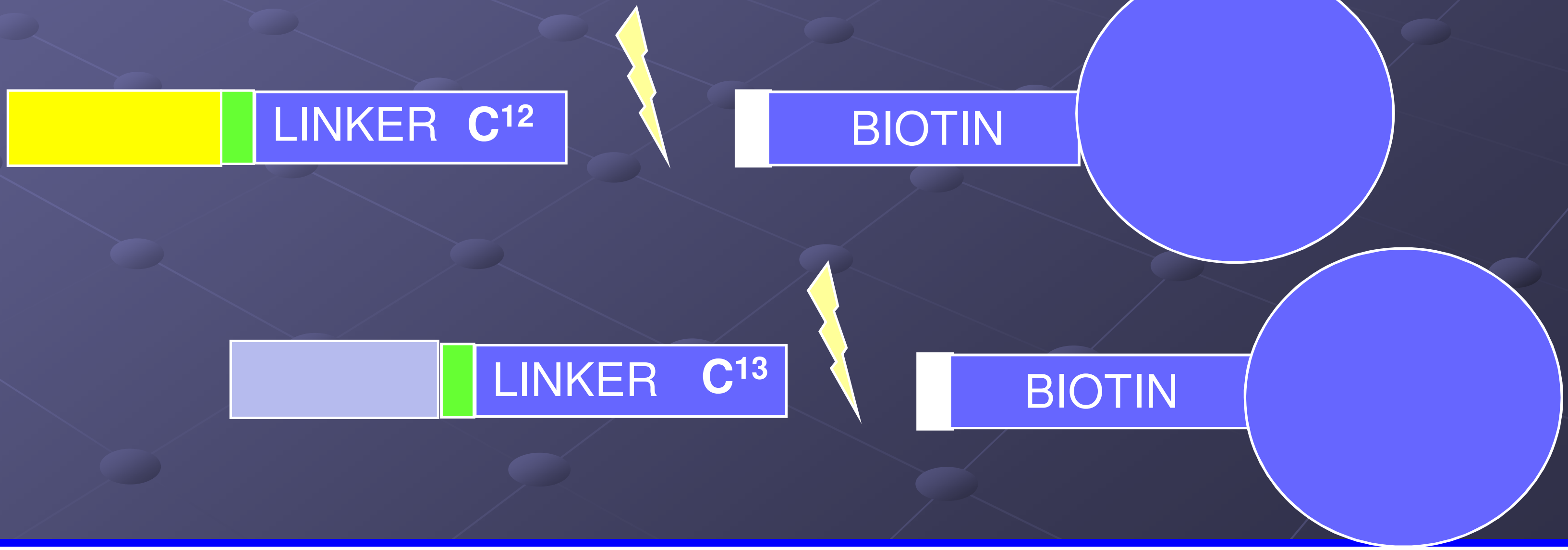
ICAT



Step 3: isolate labeled peptides

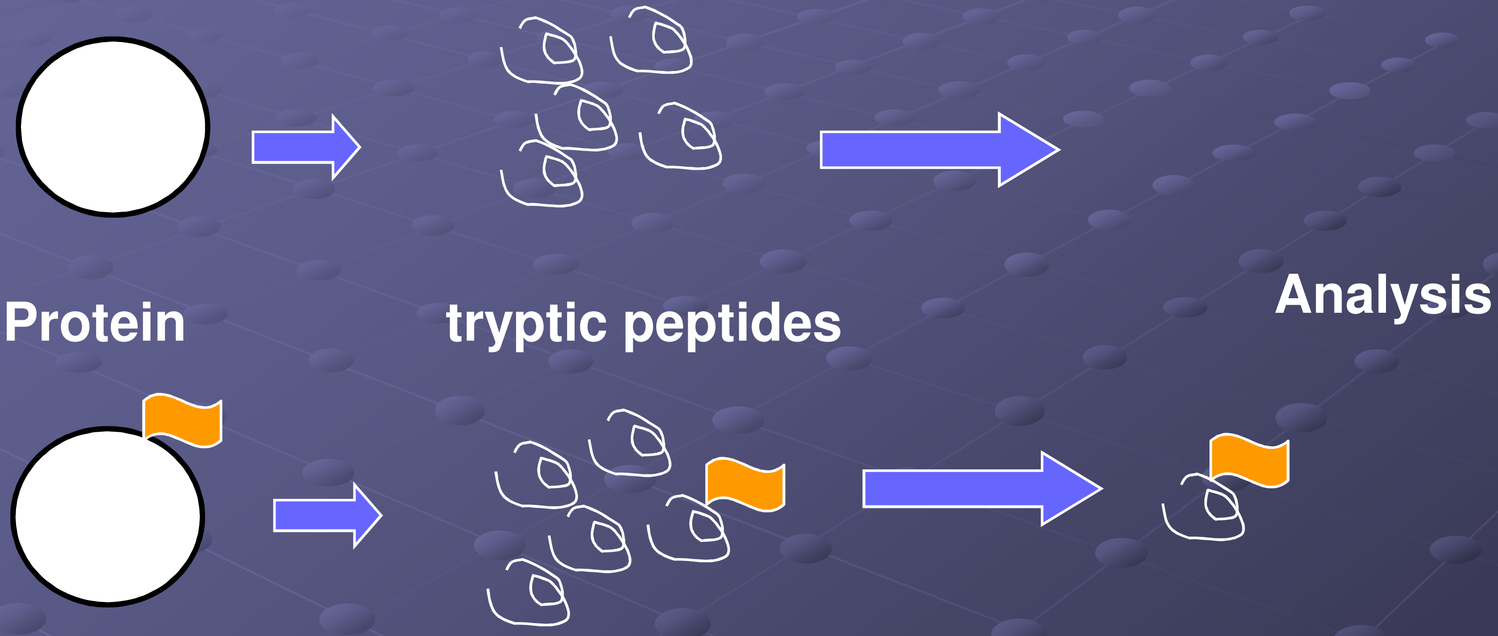
ICAT

Step 4: remove Biotin tag (= cleavable Isotope-Coded Affinity Tag (cICAT))



Labeling for Separation

- label proteins at rarely occurring AA’s (eg cysteine or methionine)
- analyze labeled peptides only --> reducing complexity



Protein

tryptic peptides

Analysis

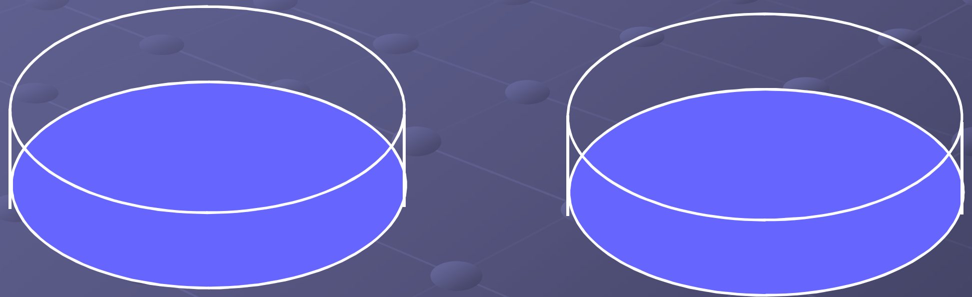
ICAT (isotope coded affinity tag) , COFRADIC

Labeling: SILAC

Stable isotope labeling with amino acids in cell culture (SILAC)

eg: **^{13}C labeled L-lysine** :

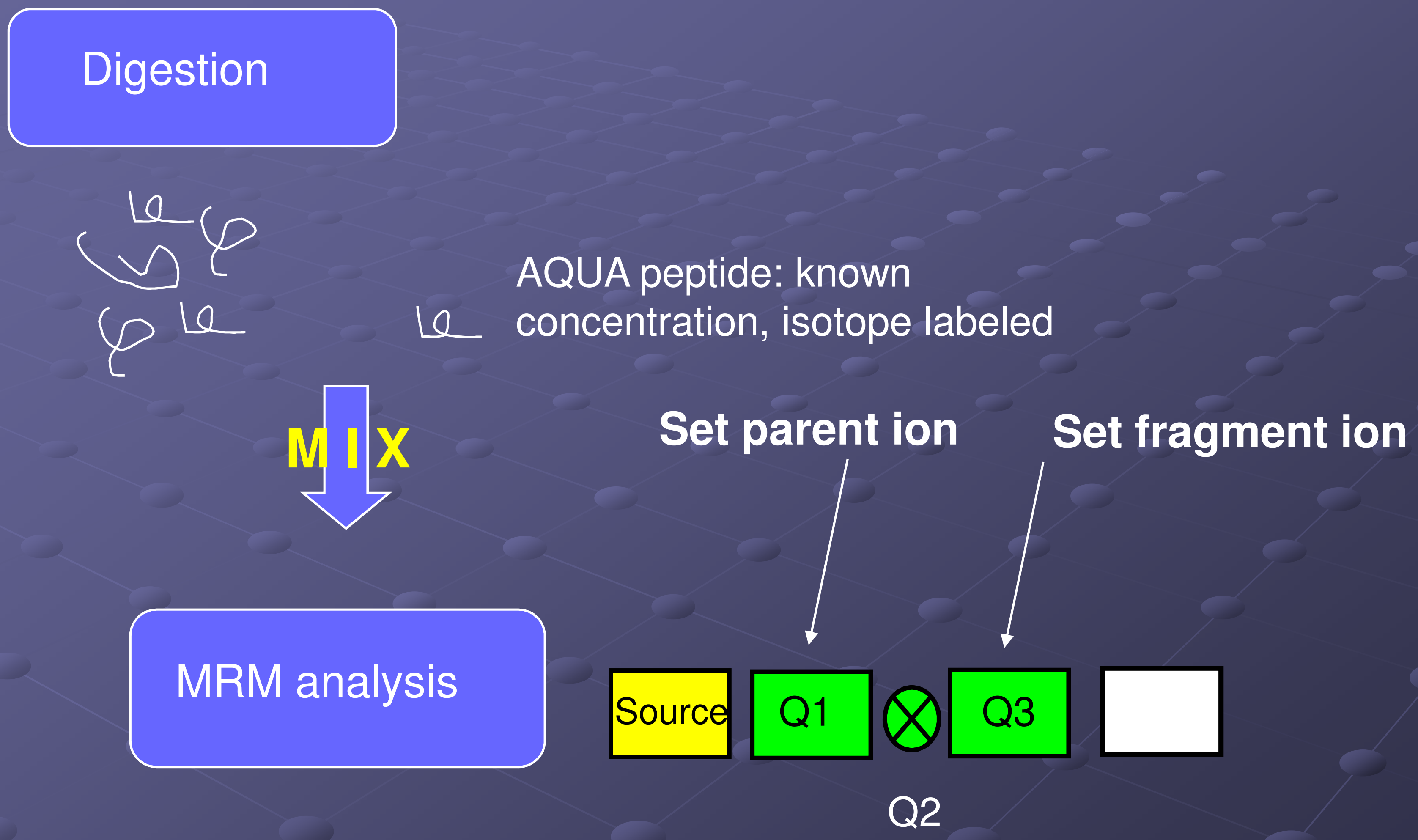
- 2 populations ("light" and "heavy" amino acid in cell culture)



^{12}C L-lysine **^{13}C labeled L-lysine**

Duplex Labeling technique
Labeling *in vivo*

Labeling: AQUA



Digestion

AQUA peptide: known concentration, isotope labeled

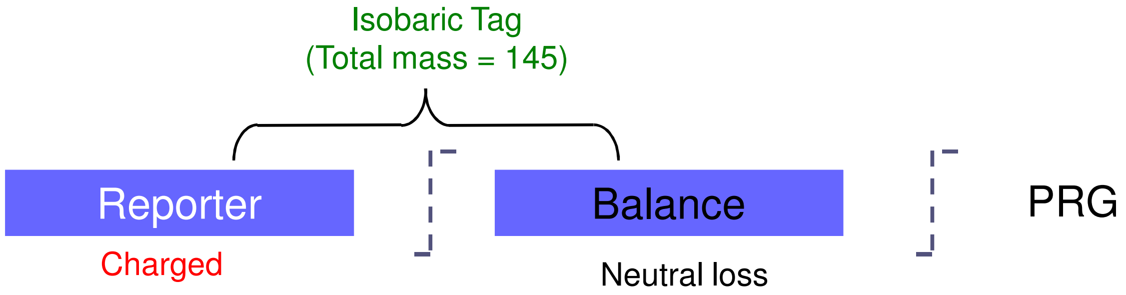
MIX

MRM analysis

Set parent ion

Set fragment ion

Labeling: iTRAQ



Isobaric Tag
(Total mass = 145)

Reporter

Charged

Balance

Neutral loss

PRG

- ✓ **Gives strong signature ion in MS/MS**
- ✓ **Gives good b- and y-ion series**
- ✓ **Maintains charge state**
- ✓ **Maintains ionization efficiency**
- ✓ **Signature ion masses lie in quiet region**

Reporter
(Mass = 114 thru 117)

Balance
(Mass = 31 thru 28)

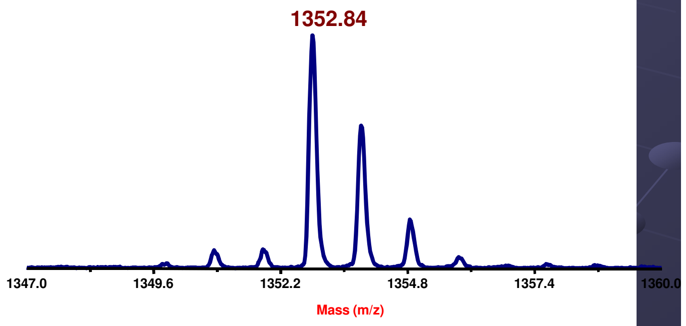
Peptide Reactive
Group

MS/MS Fragmentation Site

iTRAQ: Multiplexing

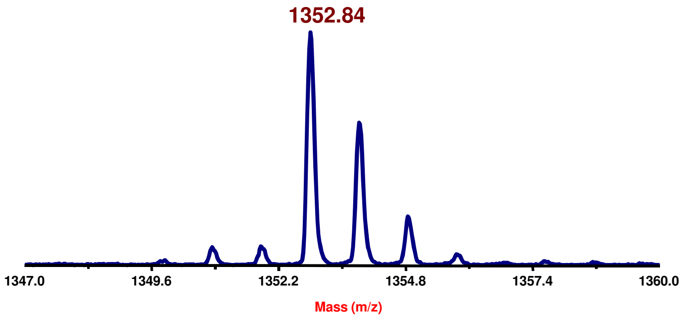
114	31	PEPTIDE
115	30	PEPTIDE
116	29	PEPTIDE
117	28	PEPTIDE

MIX → MS



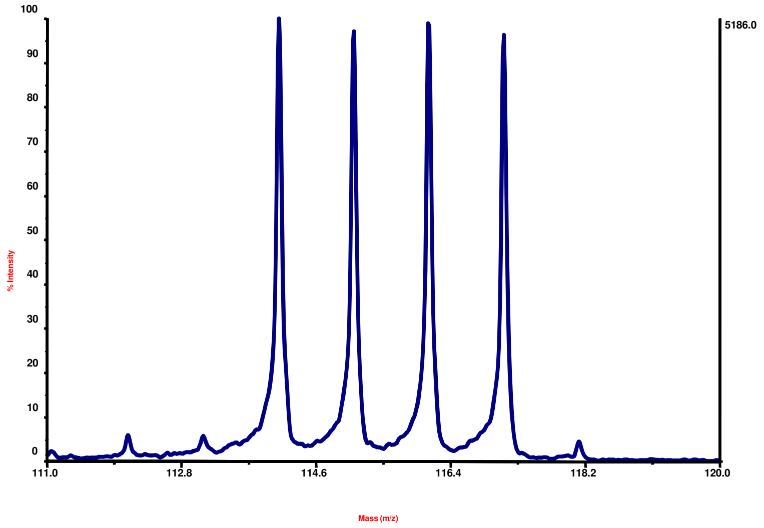
1 M/Z peak !

iTRAQ: Multiplexing



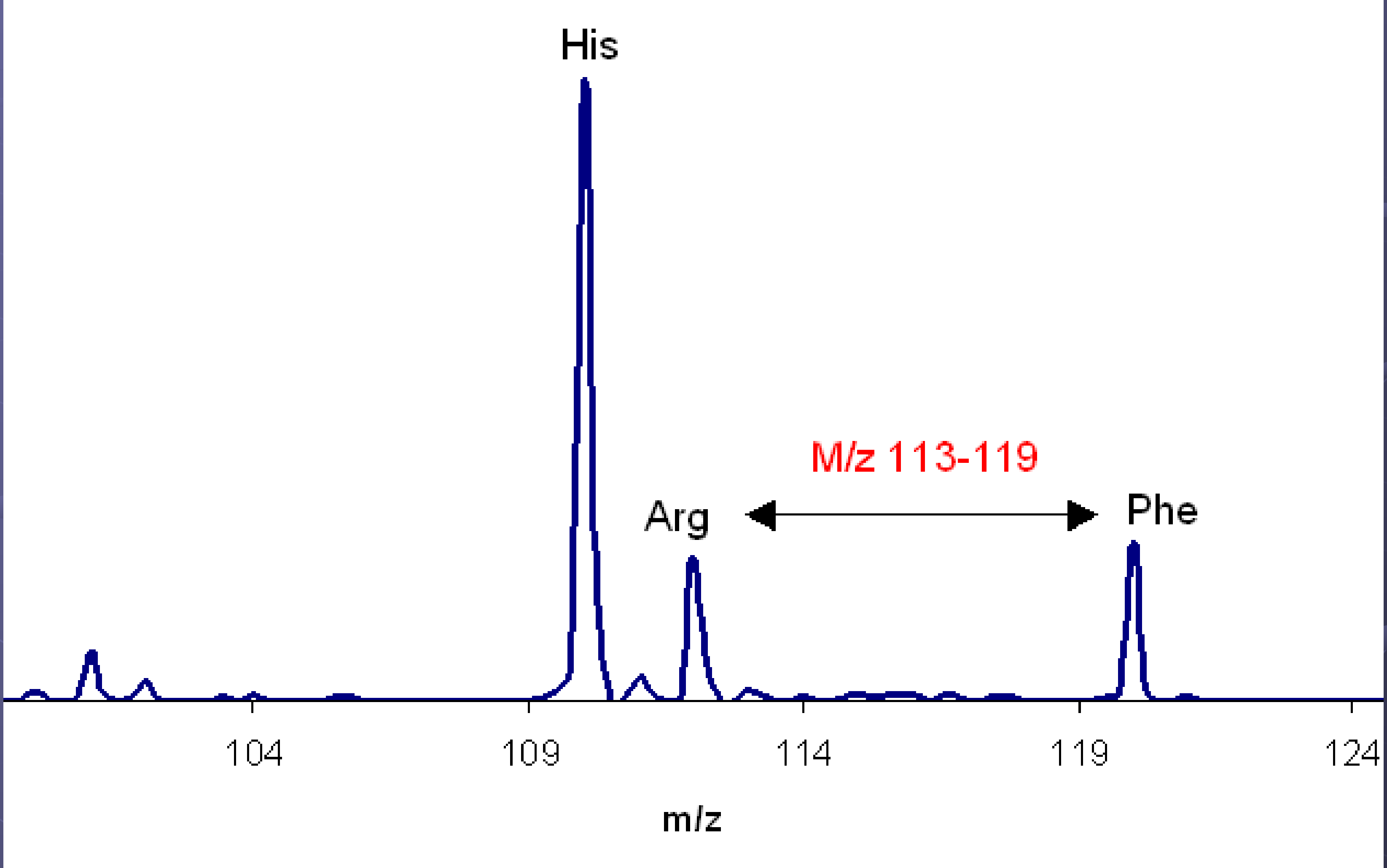
1 M/Z peak !

MS/MS

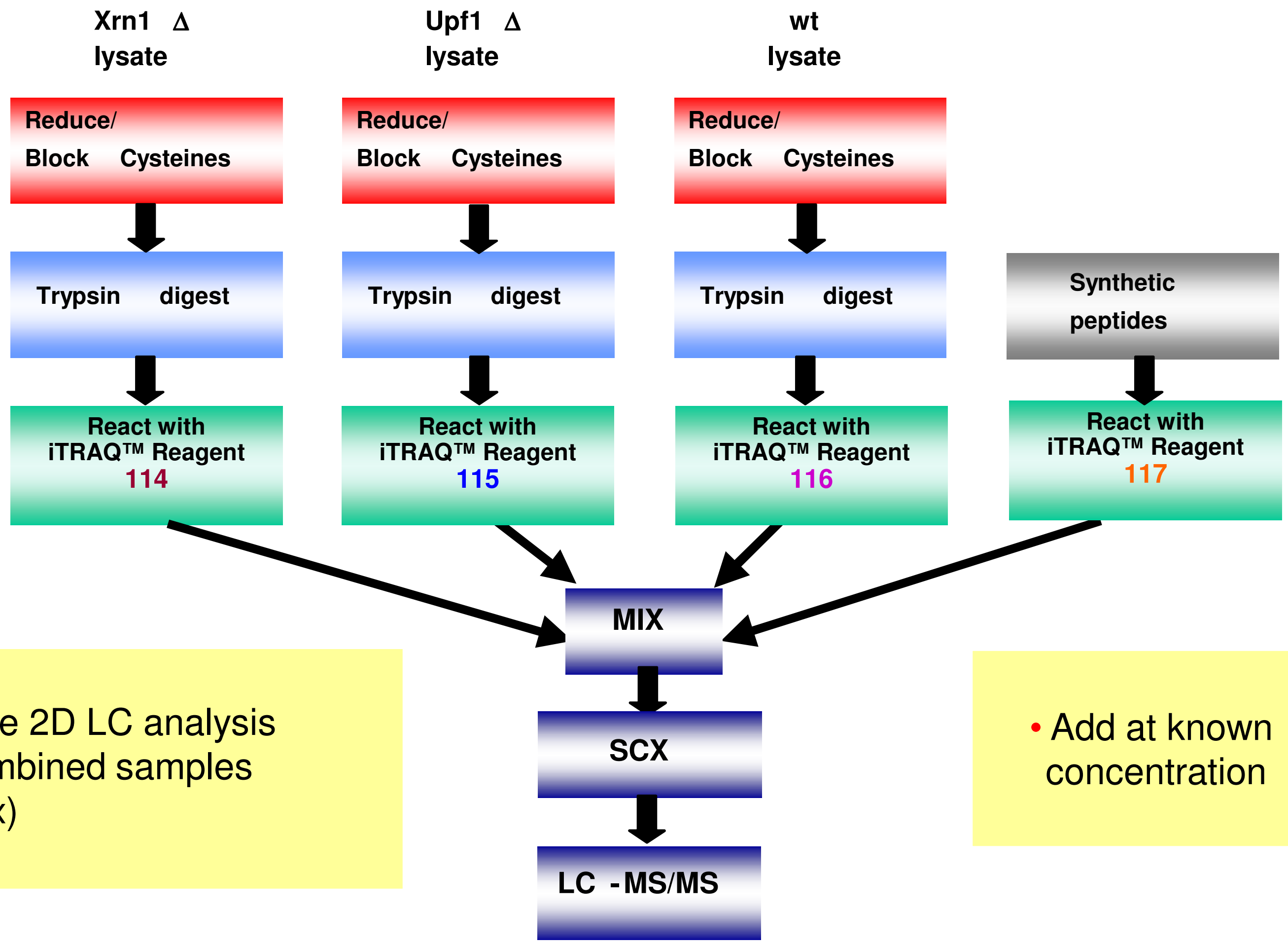


4 # fragments !

iTRAQ



iTRAQ: Absolute Quantification



Xrn1 Δ lysate

Upf1 Δ lysate

wt lysate

Reduce/ Block Cysteines

Trypsin digest

React with iTRAQ™ Reagent 114

Reduce/ Block Cysteines

Trypsin digest

React with iTRAQ™ Reagent 115

Reduce/ Block Cysteines

Trypsin digest

React with iTRAQ™ Reagent 116

Synthetic peptides

React with iTRAQ™ Reagent 117

MIX

SCX

LC - MS/MS

- Single 2D LC analysis for combined samples (4-plex)

- Add at known concentration

Clinical Proteomics: Biomarkers

Labeling

- At protein level or at peptide level
- relative quantification and absolute quantification (using standards)
- allows MULTIPLEXING

-

Note:

2D DIGE = also multiplexing and relative quantification

Clinical Proteomics: Biomarkers

Label-free

NO MULTIPLEXING

DATA Processing !!!!

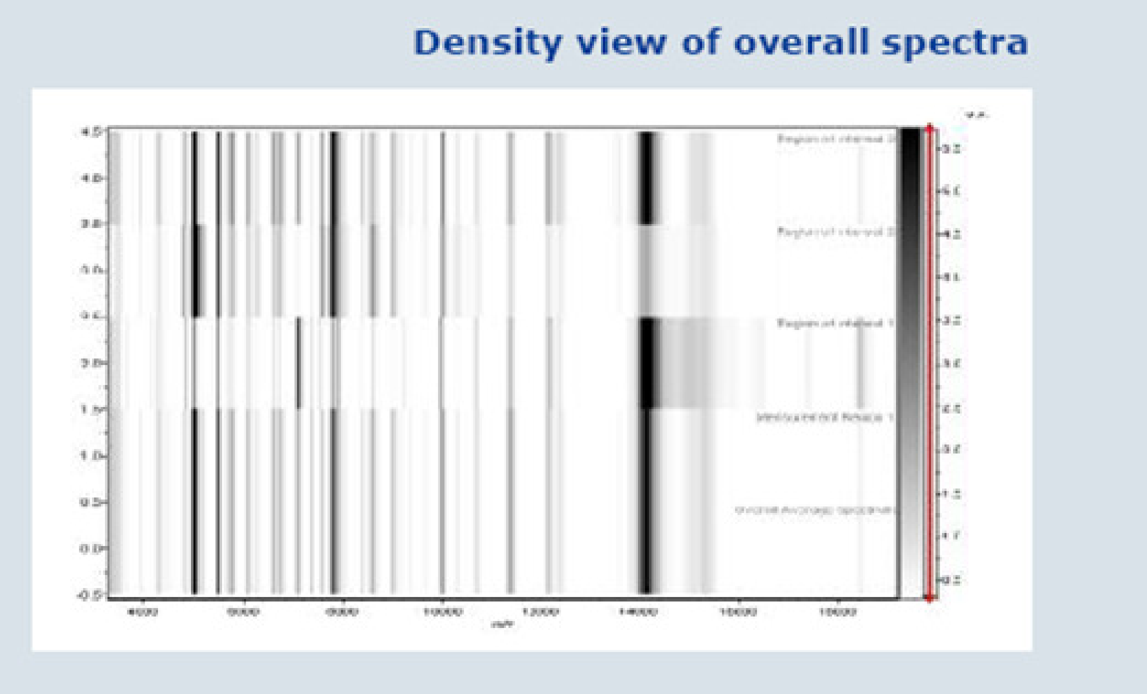
-

Clinical Proteomics: Biomarkers

Label-free Profiling

"GEL" view matching approach

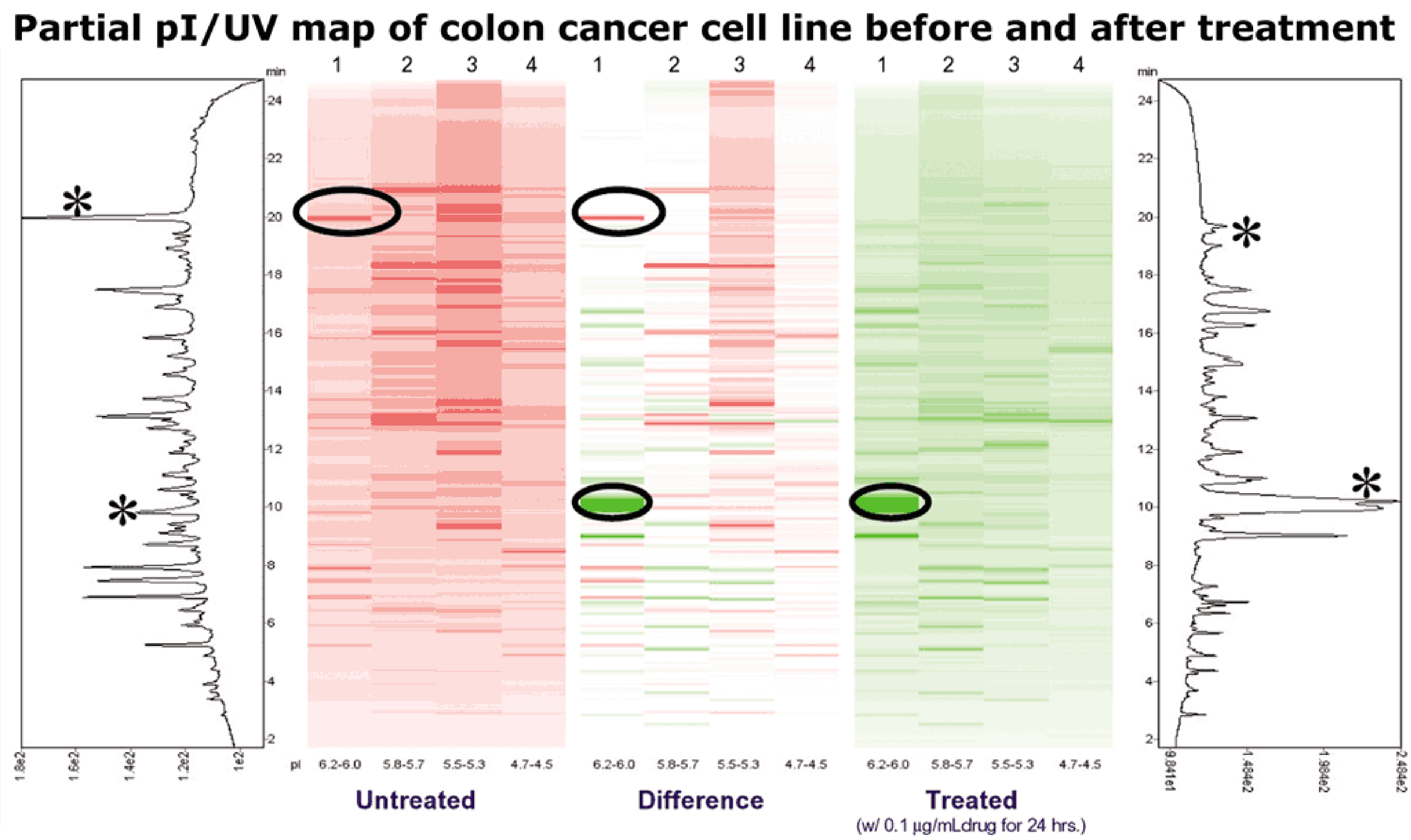
1D GEL view:



Density view of overall spectra

1D Gel view

Using UV spectra: PF2D

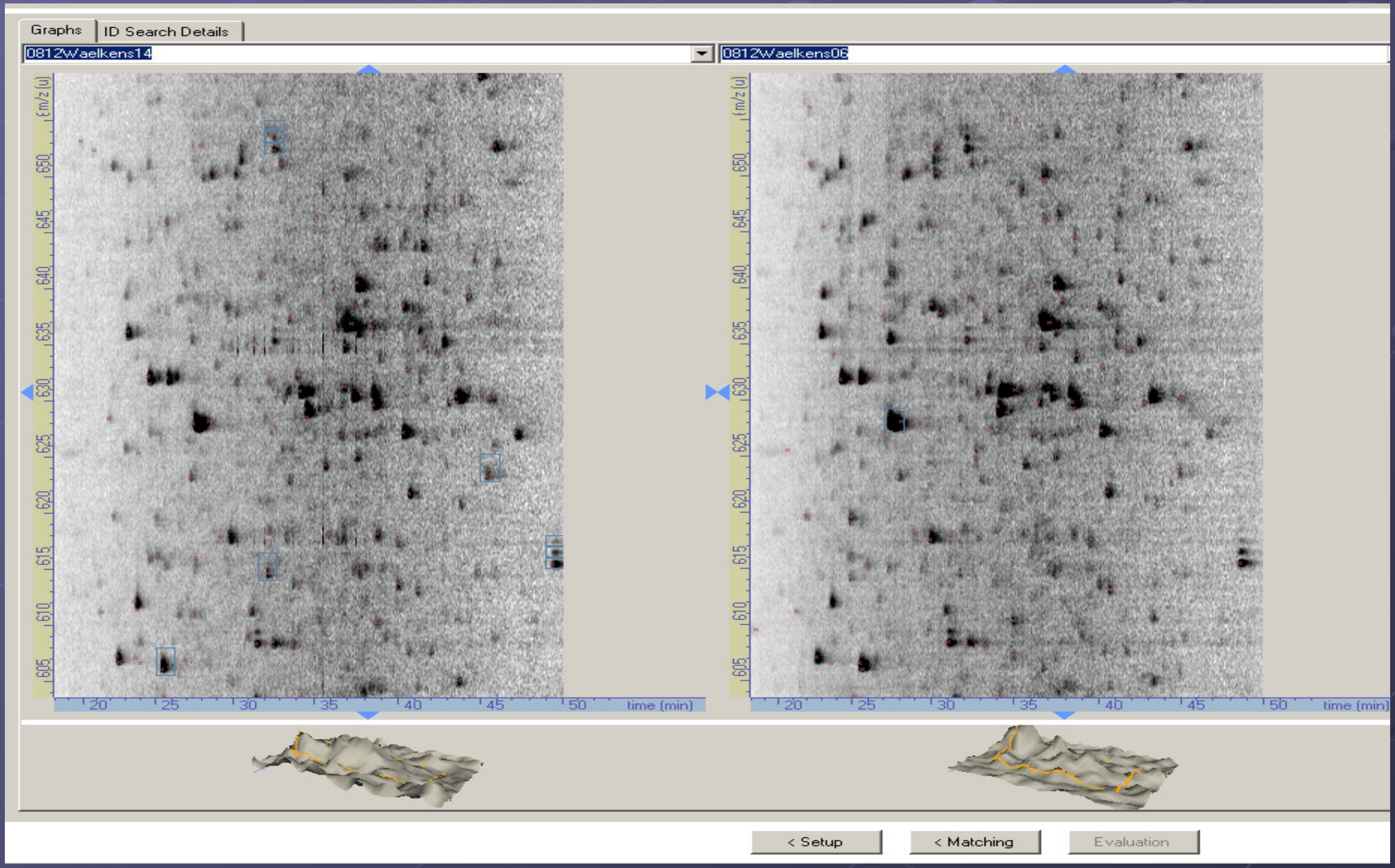


Partial pI/UV map of colon cancer cell line before and after treatment

2D Gel view

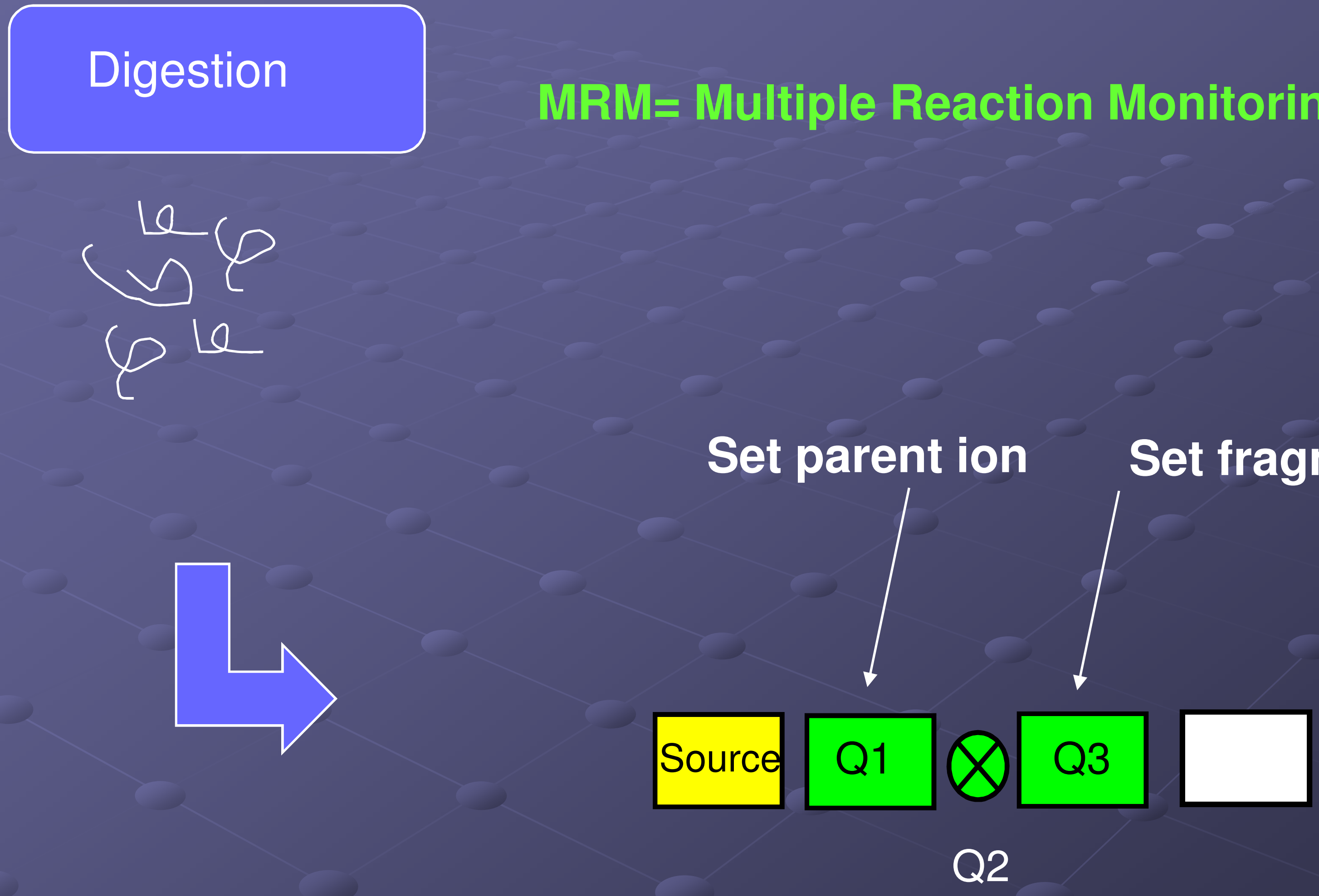
Label-free “GEL” view matching approach

2D GEL view: mapping using retention time and m/z value


0812Waelkens14

0812Waelkens06

Label-free: MRM approach



Digestion

MRM= Multiple Reaction Monitoring

Set parent ion

Set fragment ion

Clinical Proteomics: Biomarkers

Label-free

- At protein level or at peptide level
- only relative quantification (using standards)
- no MULTIPLEXING
- Cost effective and easy

-

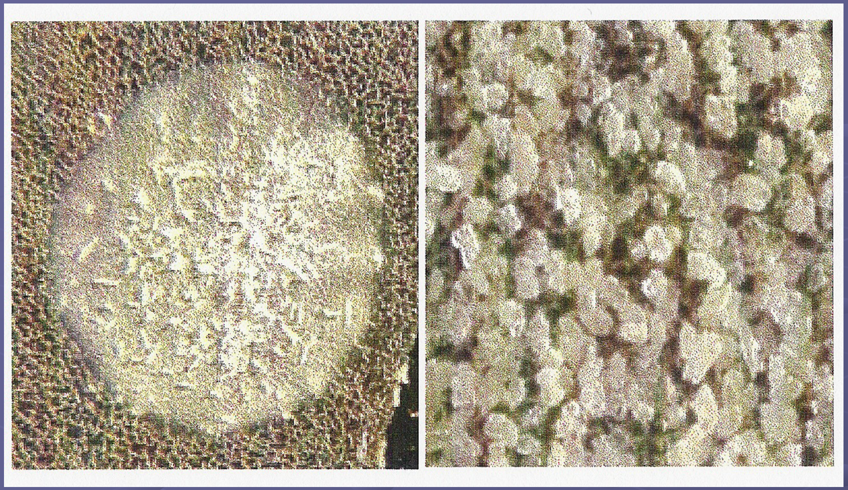
Examples:

SELDI-MS

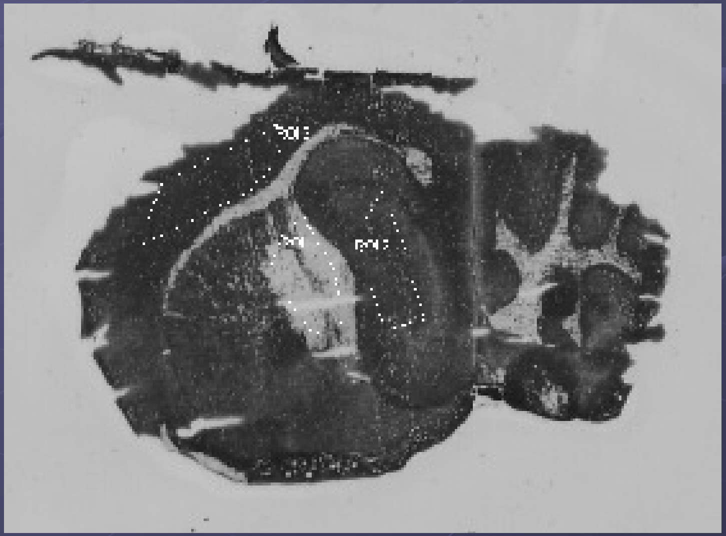
2D LC MS/MS with mapping of the data using retention time and m/z value

MRM methods

Appetizer: Imaging Mass Spectrometry

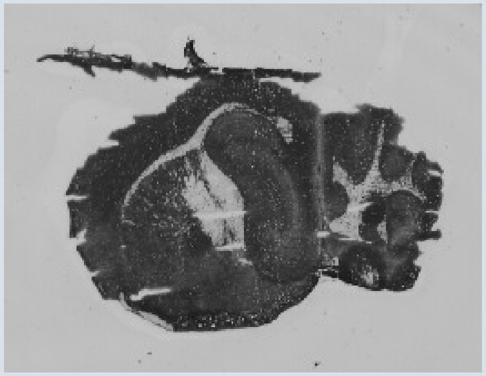


Matrix Coating on Tissue

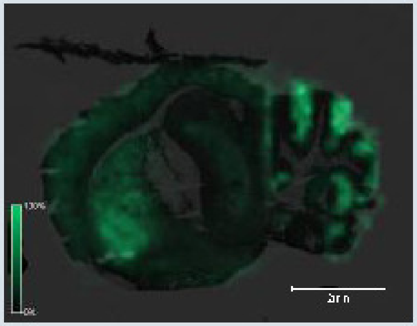


Optical Image

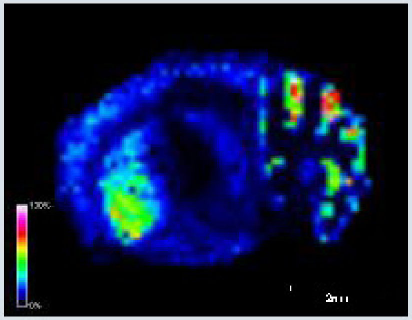
Appetizer: Imaging Mass Spectrometry



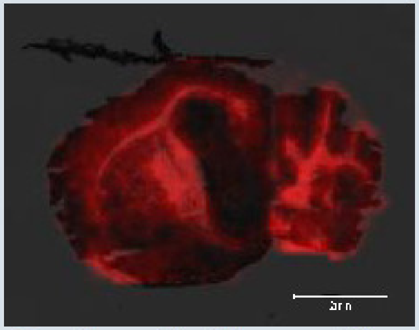
Optical image



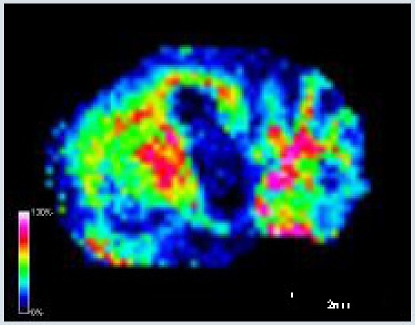
Overlay optical image with distribution of peak m/z = 6723 Da



Intensity distribution of peak m/z = 6723 Da



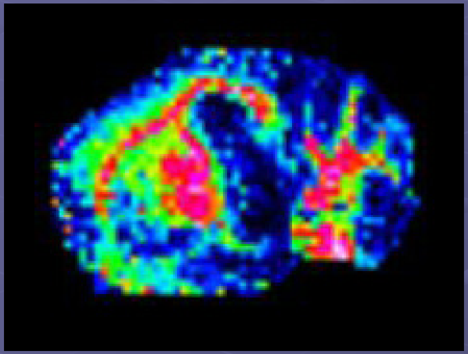
Overlay optical image with distribution of peak m/z = 7069 Da

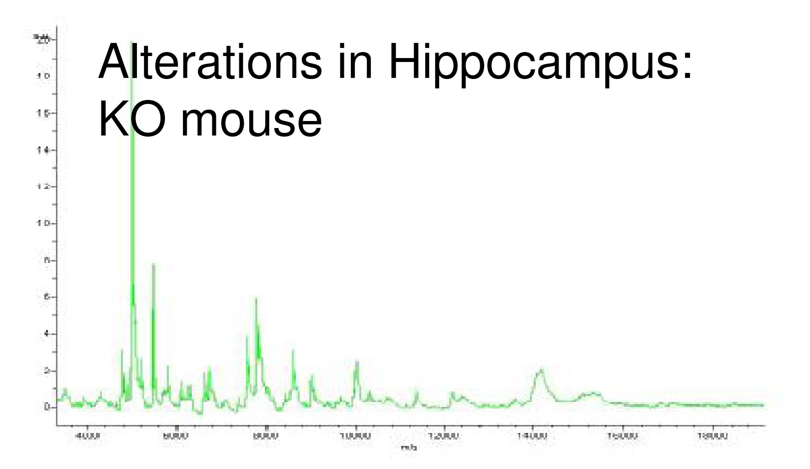


Intensity distribution of peak m/z = 7069 Da

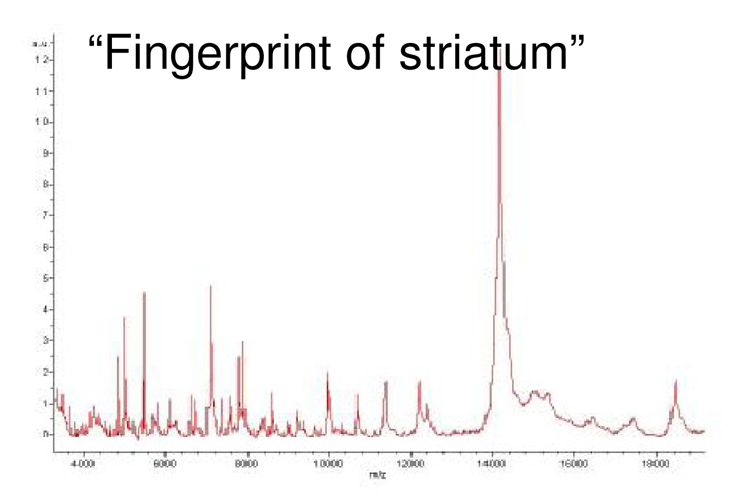
Appetizer: Tissue Proteomics using IMS

Tissue Biomarker Discovery

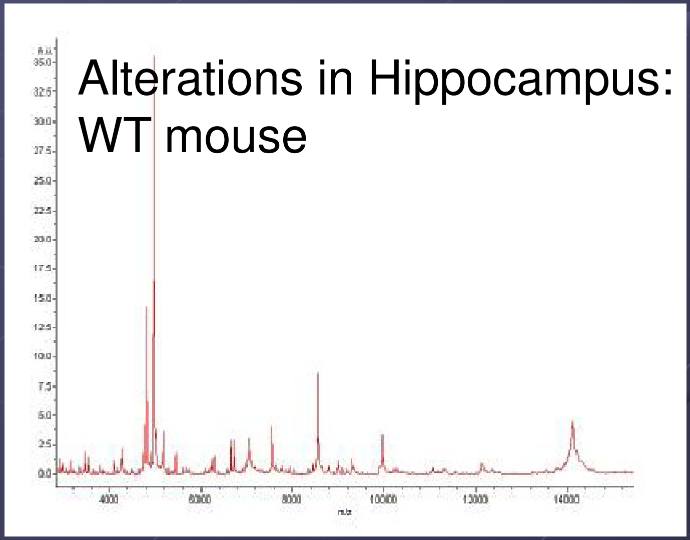




Alterations in Hippocampus: KO mouse



“Fingerprint of striatum”



Alterations in Hippocampus: WT mouse

General considerations

Clinical Proteomics/Peptidomics

- Many possible approaches
- More separation = more identification
- Serum/plasma proteomics: use depletion methods
- Biomarker discovery: start with the target sample
- Bioinformatics !
- Future perspectives:
  genomics + proteomics + peptidomics + metabolomics + …
  → **Personalized medicine**

Acknowledgements

Prof R. Derua

K. Schildermans

S. Vandoninck

Prof B. De Moor

R. Van de Plas

Prof T. D’Hooghe

Prof I. Vergote

Prof D. Timmerman

KATHOLIEKE UNIVERSITEIT
LEUVEN

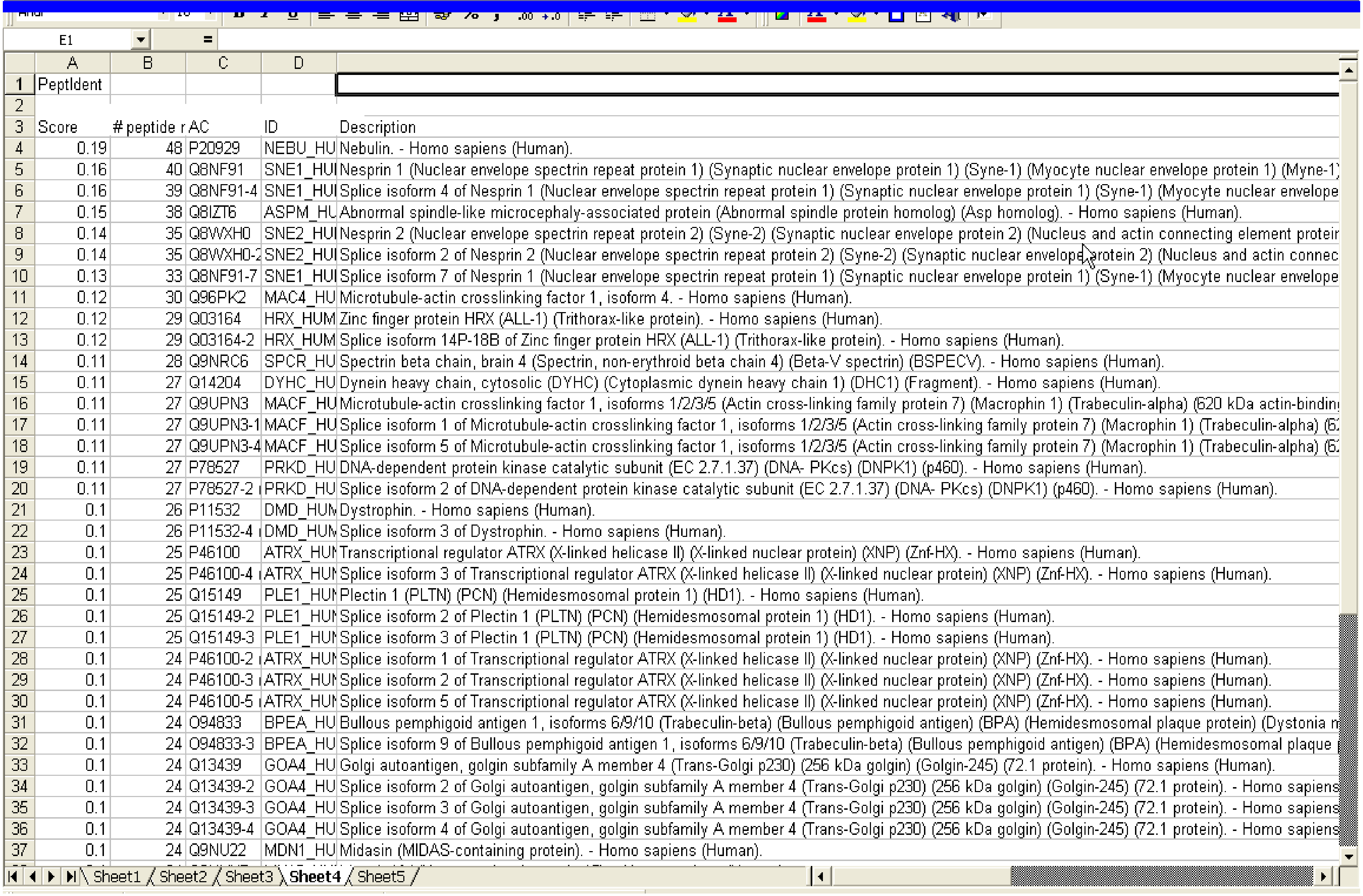
Thank You !

www.Prometa.be

BioMacS.kuleuven.be

KATHOLIEKE UNIVERSITEIT
LEUVEN

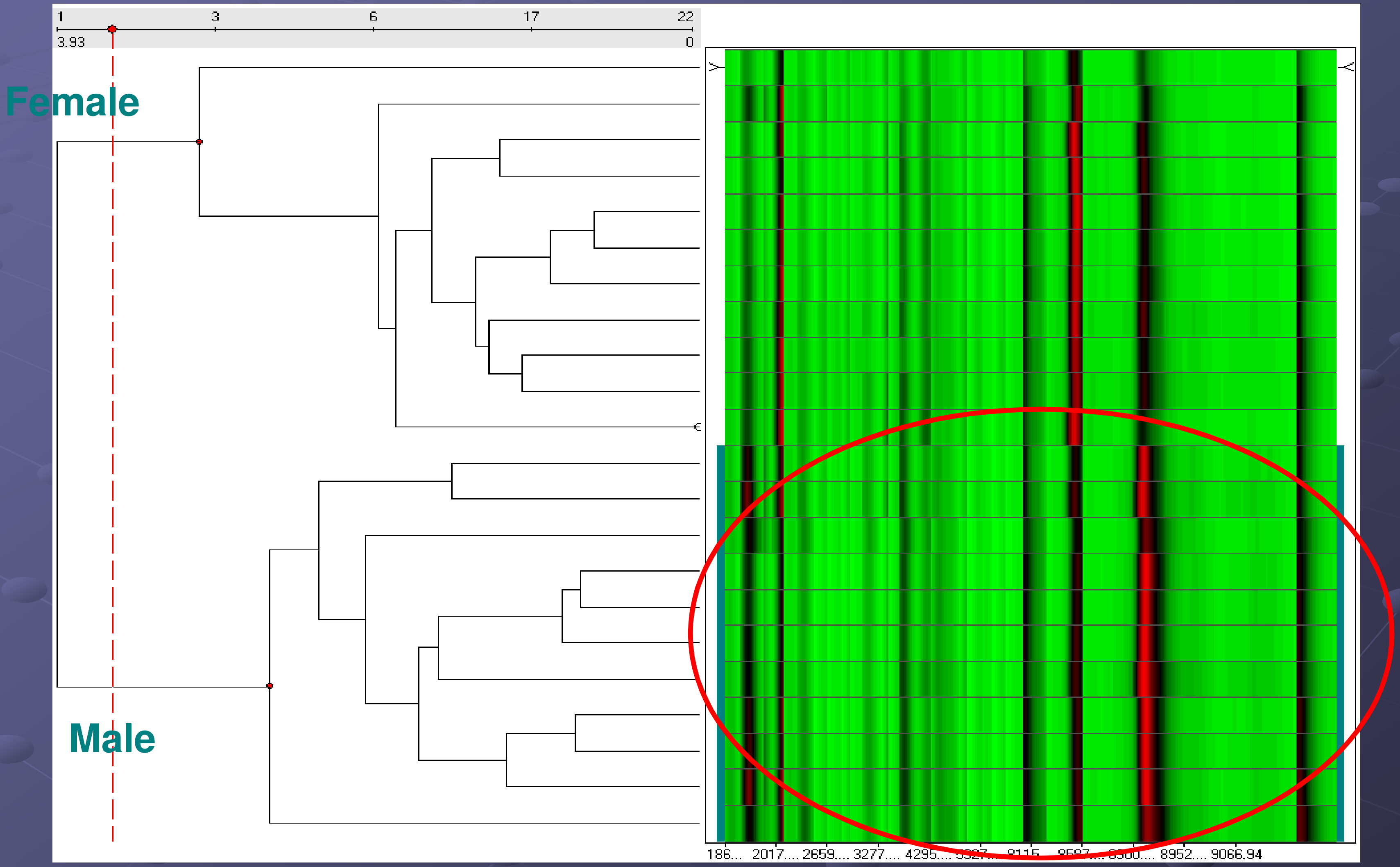
Advanced Data processing



	A	B	C	D	
1	PeptIdent				
2					
3	Score	# peptide r	AC	ID	Description
4	0.19	48	P20929	NEBU_HU	Nebulin. - Homo sapiens (Human).
5	0.16	40	Q8NF91	SNE1_HUI	Nesprin 1 (Nuclear envelope spectrin repeat protein 1) (Synaptic nuclear envelope protein 1) (Syne-1) (Myocyte nuclear envelope protein 1) (Myne-1)
6	0.16	39	Q8NF91-4	SNE1_HUI	Splice isoform 4 of Nesprin 1 (Nuclear envelope spectrin repeat protein 1) (Synaptic nuclear envelope protein 1) (Syne-1) (Myocyte nuclear envelope
7	0.15	38	Q8IZT6	ASPM_HU	Abnormal spindle-like microcephaly-associated protein (Abnormal spindle protein homolog) (Asp homolog). - Homo sapiens (Human).
8	0.14	35	Q8WXH0	SNE2_HUI	Nesprin 2 (Nuclear envelope spectrin repeat protein 2) (Syne-2) (Synaptic nuclear envelope protein 2) (Nucleus and actin connecting element protein
9	0.14	35	Q8WXH0-2	SNE2_HUI	Splice isoform 2 of Nesprin 2 (Nuclear envelope spectrin repeat protein 2) (Syne-2) (Synaptic nuclear envelope protein 2) (Nucleus and actin connec
10	0.13	33	Q8NF91-7	SNE1_HUI	Splice isoform 7 of Nesprin 1 (Nuclear envelope spectrin repeat protein 1) (Synaptic nuclear envelope protein 1) (Syne-1) (Myocyte nuclear envelope
11	0.12	30	Q96PK2	MAC4_HU	Microtubule-actin crosslinking factor 1, isoform 4. - Homo sapiens (Human).
12	0.12	29	Q03164	HRX_HUM	Zinc finger protein HRX (ALL-1) (Trithorax-like protein). - Homo sapiens (Human).
13	0.12	29	Q03164-2	HRX_HUM	Splice isoform 14P-18B of Zinc finger protein HRX (ALL-1) (Trithorax-like protein). - Homo sapiens (Human).
14	0.11	28	Q9NRC6	SPCR_HU	Spectrin beta chain, brain 4 (Spectrin, non-erythroid beta chain 4) (Beta-V spectrin) (BSPECV). - Homo sapiens (Human).
15	0.11	27	Q14204	DYHC_HU	Dynein heavy chain, cytosolic (DYHC) (Cytoplasmic dynein heavy chain 1) (DHC1) (Fragment). - Homo sapiens (Human).
16	0.11	27	Q9UPN3	MACF_HU	Microtubule-actin crosslinking factor 1, isoforms 1/2/3/5 (Actin cross-linking family protein 7) (Macrophin 1) (Trabeculin-alpha) (620 kDa actin-bindin
17	0.11	27	Q9UPN3-1	MACF_HU	Splice isoform 1 of Microtubule-actin crosslinking factor 1, isoforms 1/2/3/5 (Actin cross-linking family protein 7) (Macrophin 1) (Trabeculin-alpha) (6
18	0.11	27	Q9UPN3-4	MACF_HU	Splice isoform 5 of Microtubule-actin crosslinking factor 1, isoforms 1/2/3/5 (Actin cross-linking family protein 7) (Macrophin 1) (Trabeculin-alpha) (6
19	0.11	27	P78527	PRKD_HU	DNA-dependent protein kinase catalytic subunit (EC 2.7.1.37) (DNA- PKcs) (DNPK1) (p460). - Homo sapiens (Human).
20	0.11	27	P78527-2	PRKD_HU	Splice isoform 2 of DNA-dependent protein kinase catalytic subunit (EC 2.7.1.37) (DNA- PKcs) (DNPK1) (p460). - Homo sapiens (Human).
21	0.1	26	P11532	DMD_HUN	Dystrophin. - Homo sapiens (Human).
22	0.1	26	P11532-4	DMD_HUN	Splice isoform 3 of Dystrophin. - Homo sapiens (Human).
23	0.1	25	P46100	ATRX_HUI	Transcriptional regulator ATRX (X-linked helicase II) (X-linked nuclear protein) (XNP) (Znf-HX). - Homo sapiens (Human).
24	0.1	25	P46100-4	ATRX_HUI	Splice isoform 3 of Transcriptional regulator ATRX (X-linked helicase II) (X-linked nuclear protein) (XNP) (Znf-HX). - Homo sapiens (Human).
25	0.1	25	Q15149	PLE1_HUI	Plectin 1 (PLTN) (PCN) (Hemidesmosomal protein 1) (HD1). - Homo sapiens (Human).
26	0.1	25	Q15149-2	PLE1_HUI	Splice isoform 2 of Plectin 1 (PLTN) (PCN) (Hemidesmosomal protein 1) (HD1). - Homo sapiens (Human).
27	0.1	25	Q15149-3	PLE1_HUI	Splice isoform 3 of Plectin 1 (PLTN) (PCN) (Hemidesmosomal protein 1) (HD1). - Homo sapiens (Human).
28	0.1	24	P46100-2	ATRX_HUI	Splice isoform 1 of Transcriptional regulator ATRX (X-linked helicase II) (X-linked nuclear protein) (XNP) (Znf-HX). - Homo sapiens (Human).
29	0.1	24	P46100-3	ATRX_HUI	Splice isoform 2 of Transcriptional regulator ATRX (X-linked helicase II) (X-linked nuclear protein) (XNP) (Znf-HX). - Homo sapiens (Human).
30	0.1	24	P46100-5	ATRX_HUI	Splice isoform 5 of Transcriptional regulator ATRX (X-linked helicase II) (X-linked nuclear protein) (XNP) (Znf-HX). - Homo sapiens (Human).
31	0.1	24	O94833	BPEA_HU	Bullous pemphigoid antigen 1, isoforms 6/9/10 (Trabeculin-beta) (Bullous pemphigoid antigen) (BPA) (Hemidesmosomal plaque protein) (Dystonia m
32	0.1	24	O94833-3	BPEA_HU	Splice isoform 9 of Bullous pemphigoid antigen 1, isoforms 6/9/10 (Trabeculin-beta) (Bullous pemphigoid antigen) (BPA) (Hemidesmosomal plaque
33	0.1	24	Q13439	GOA4_HU	Golgi autoantigen, golgin subfamily A member 4 (Trans-Golgi p230) (256 kDa golgin) (Golgin-245) (72.1 protein). - Homo sapiens (Human).
34	0.1	24	Q13439-2	GOA4_HU	Splice isoform 2 of Golgi autoantigen, golgin subfamily A member 4 (Trans-Golgi p230) (256 kDa golgin) (Golgin-245) (72.1 protein). - Homo sapiens
35	0.1	24	Q13439-3	GOA4_HU	Splice isoform 3 of Golgi autoantigen, golgin subfamily A member 4 (Trans-Golgi p230) (256 kDa golgin) (Golgin-245) (72.1 protein). - Homo sapiens
36	0.1	24	Q13439-4	GOA4_HU	Splice isoform 4 of Golgi autoantigen, golgin subfamily A member 4 (Trans-Golgi p230) (256 kDa golgin) (Golgin-245) (72.1 protein). - Homo sapiens
37	0.1	24	Q9NU22	MDN1_HU	Midasin (MIDAS-containing protein). - Homo sapiens (Human).

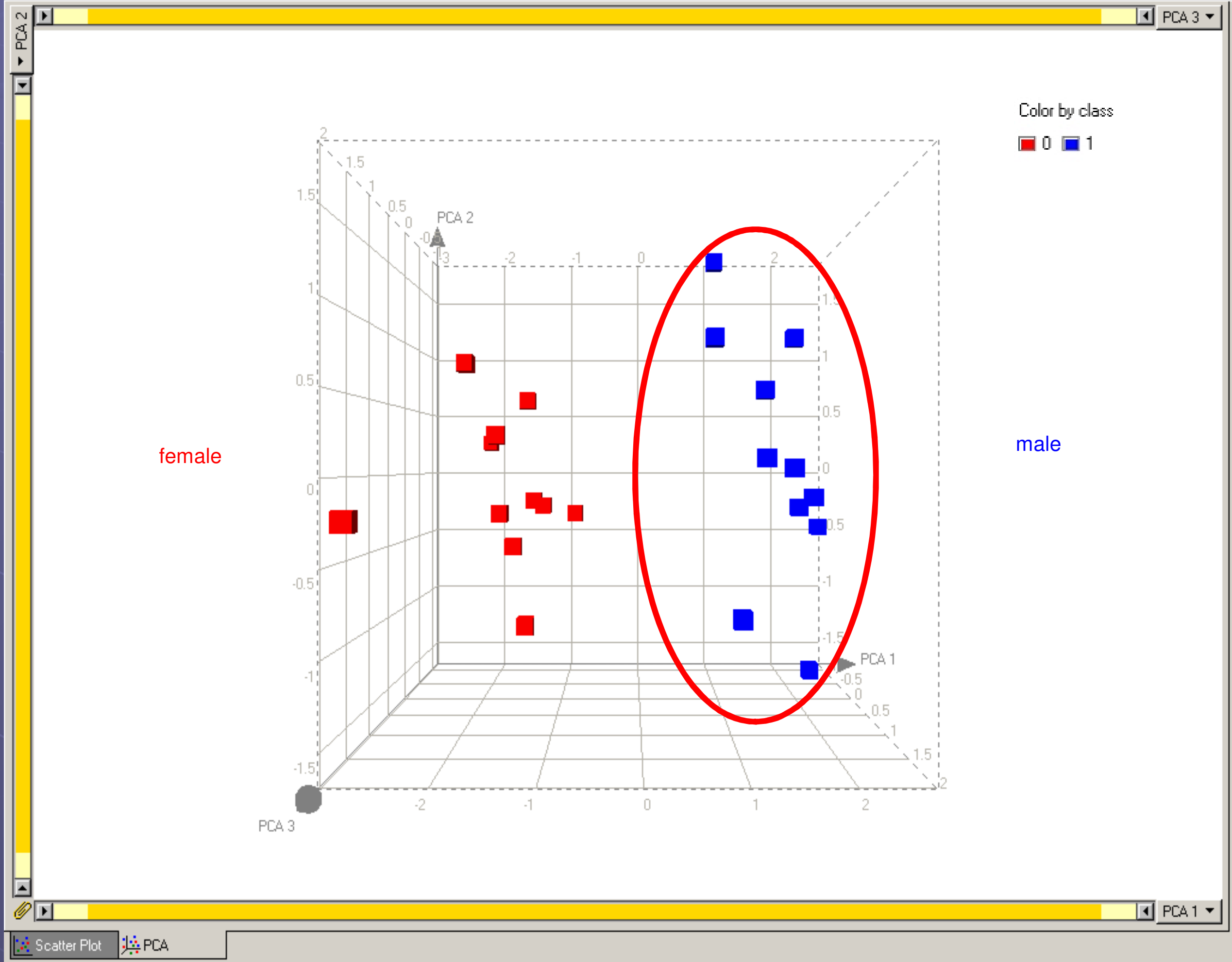
Advanced Data Processing

Hierarchical Clustering



Advanced Data Processing

PCA Classification



And finally…

- Many candidates
- Validation of potential candidates
- Design of “kits”