Functional Genomics and Proteomics
National Centre for Biomolecular Research
Faculty of Science Masaryk University

CEITEC

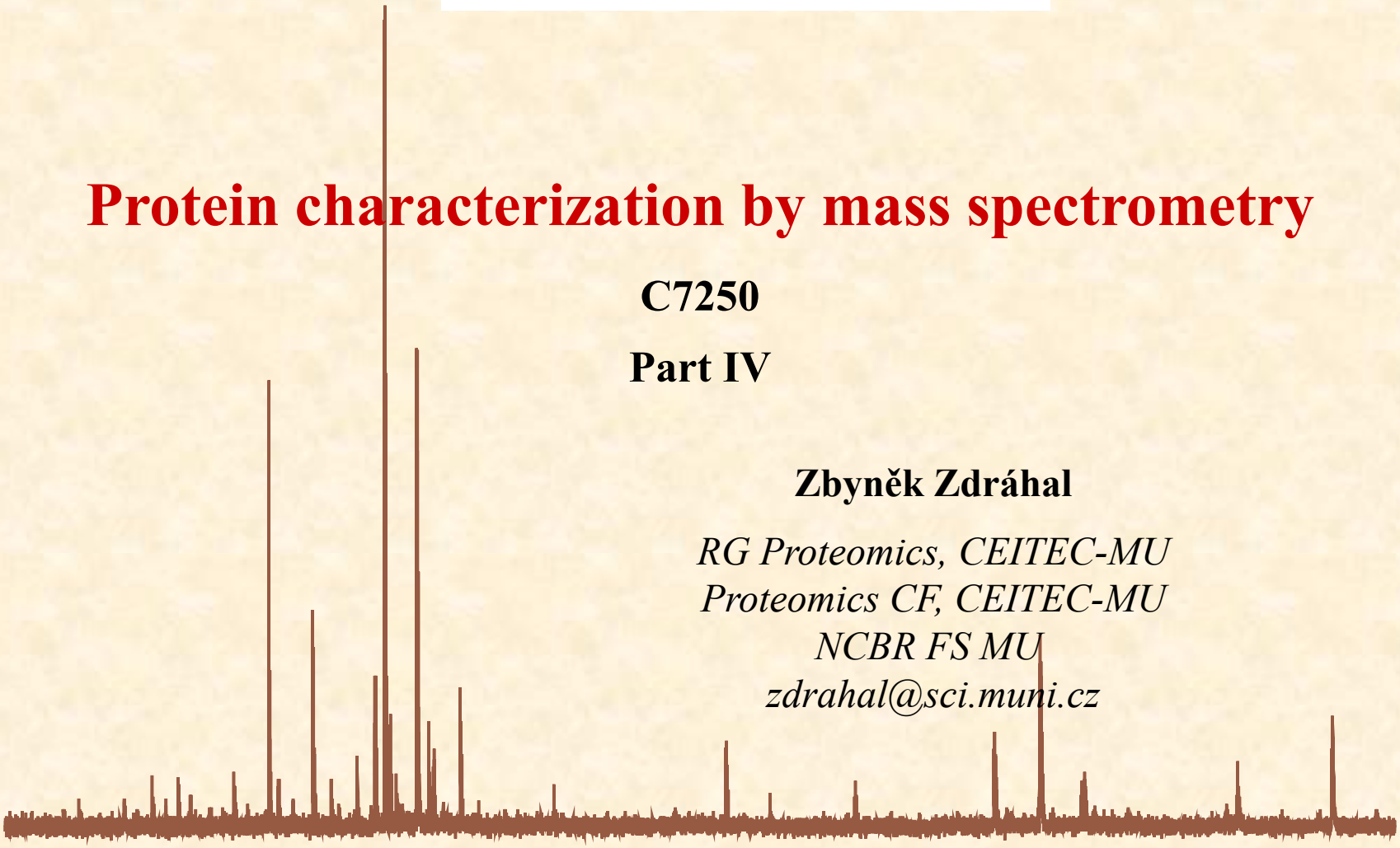
Protein characterization by mass spectrometry

C7250

Part IV

Zbyněk Zdráhal

RG Proteomics, CEITEC-MU
Proteomics CF, CEITEC-MU
NCBR FS MU
zdrahal@sci.muni.cz

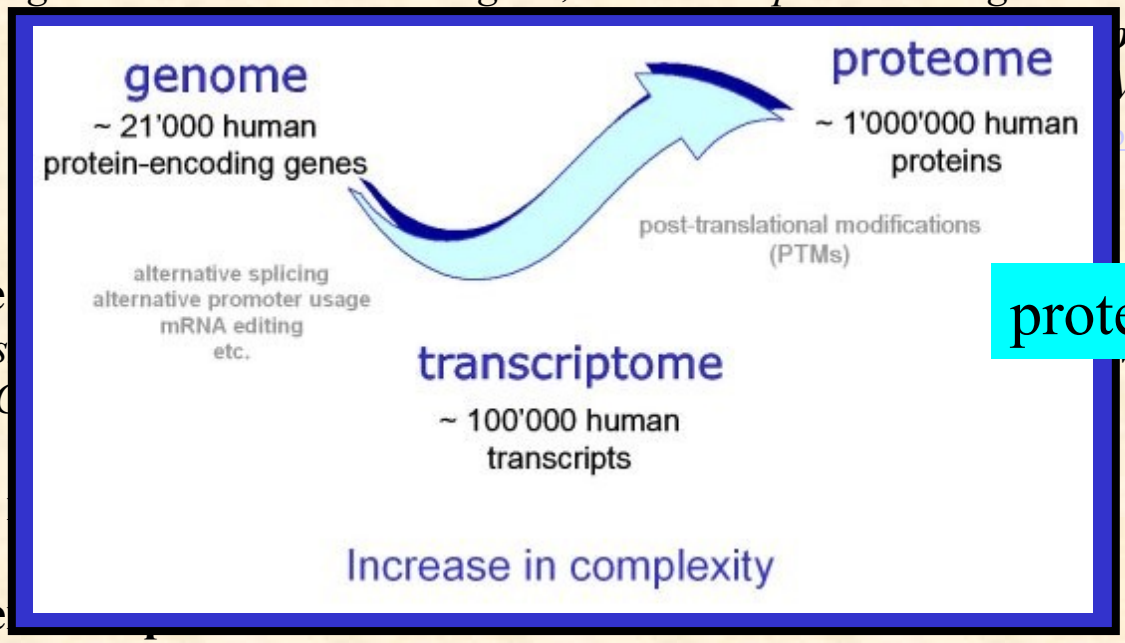


Appropriate sample preparation – key stone of success

conservation of original protein status
conservation of modification status (e.g. phosphatase inhibitors)
removal of contaminants interferring MS analysis
...

Demandingness of proteome analysis

- **protein number exceeds substantially number of genes**
  human genome contains ~21 000 genů, but human proteome might contain
  orms
  Ms)
  rot/hpi/hpi_desc.html)



proteoforms

- **wide**
  neces
  no P
  ones,

- **wide**

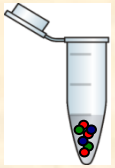
- **prote**
  necessity of protein complex analysis for deeper understanding mechanisms of cellular processes
  about 80% of proteins perform their funcions only as a part of a complex

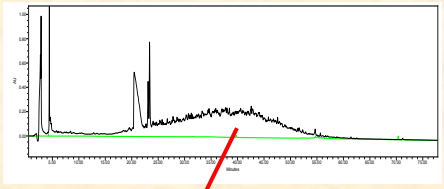
Fractionation/separation

to obtain maximum information

Direct LC-MS/MS analysis of the whole sample

~ 10^5 peptides





1000 peptides/peak

scan rate
time limitation

~ 100 peptides

~ 900 peptides
not measured

Fractionation/separation

the aim: **to simplify extremely complex mixture**
to separate specific group of proteins/peptides (e.g. phosphopeptides)

necessity of combination of separation principles – **multidimensional separation**
selection of appropriate combination for given experiment
(separation dimension might be also selected method of MS analysis)

electrophoretic techniques:

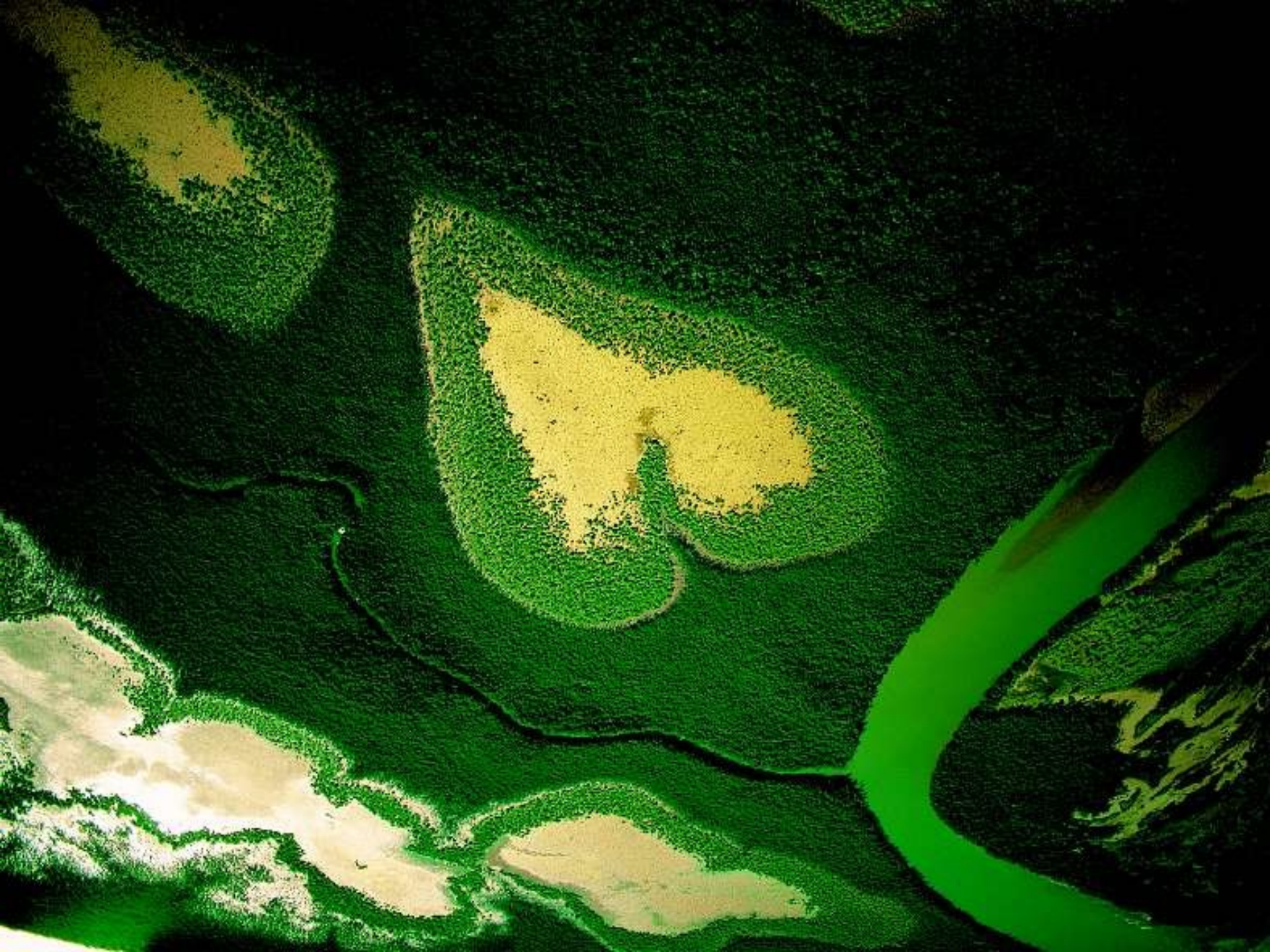
- isoelectric focusing (in-gel, in-liquid)
- SDS PAGE
- 2D gel electrophoresis (DIGE)
- capillary electrophoresis

chromatographic techniques:

immunoprecipitation

- liquid chromatography
  - **reverse phase**
  - ionex
  - molecular sieve
  - affinity (IMAC, MOAC, antibody)
  - HILIC (hydrophilic interaction chromatography)

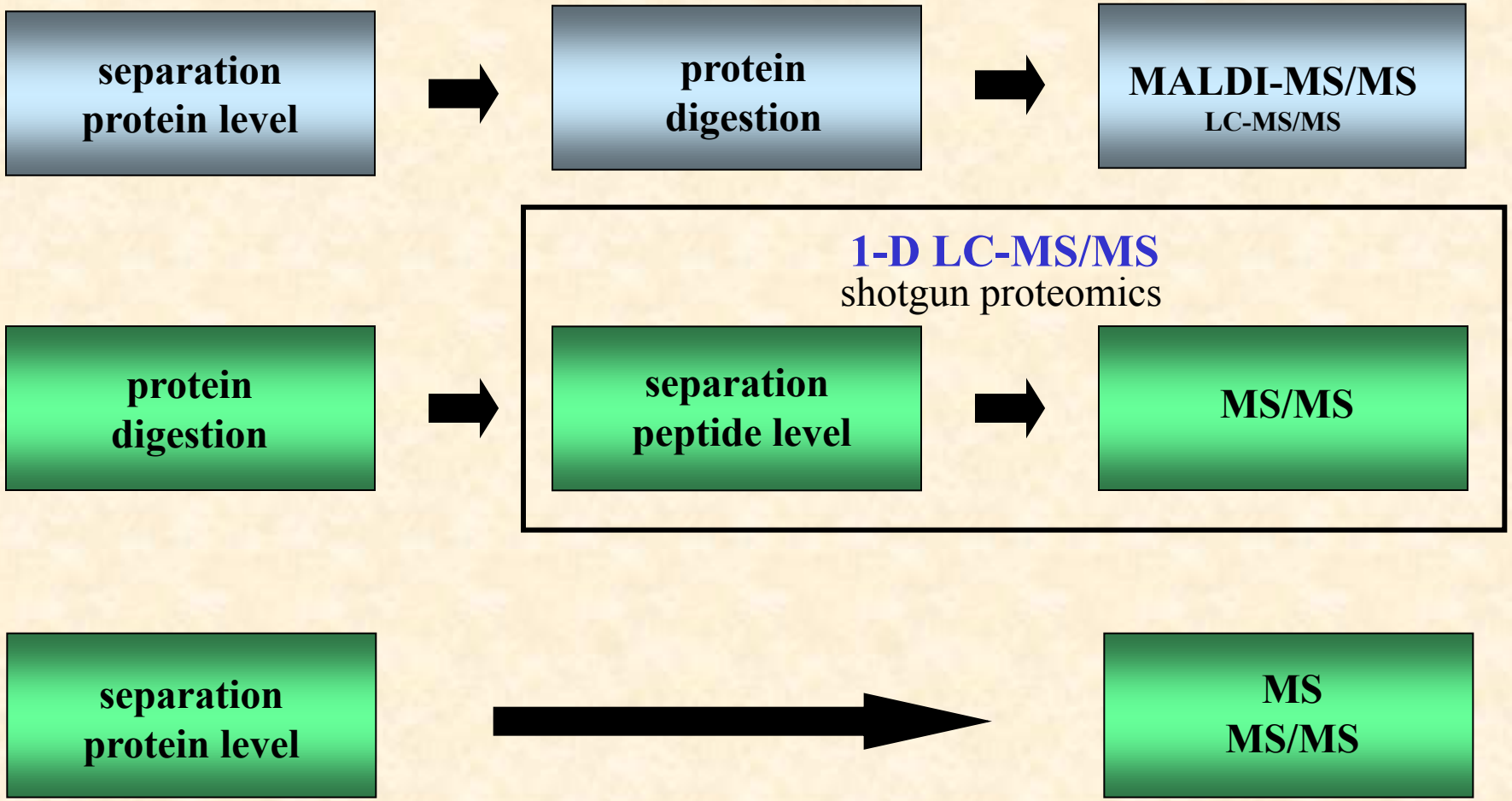
- off-line
- on-line



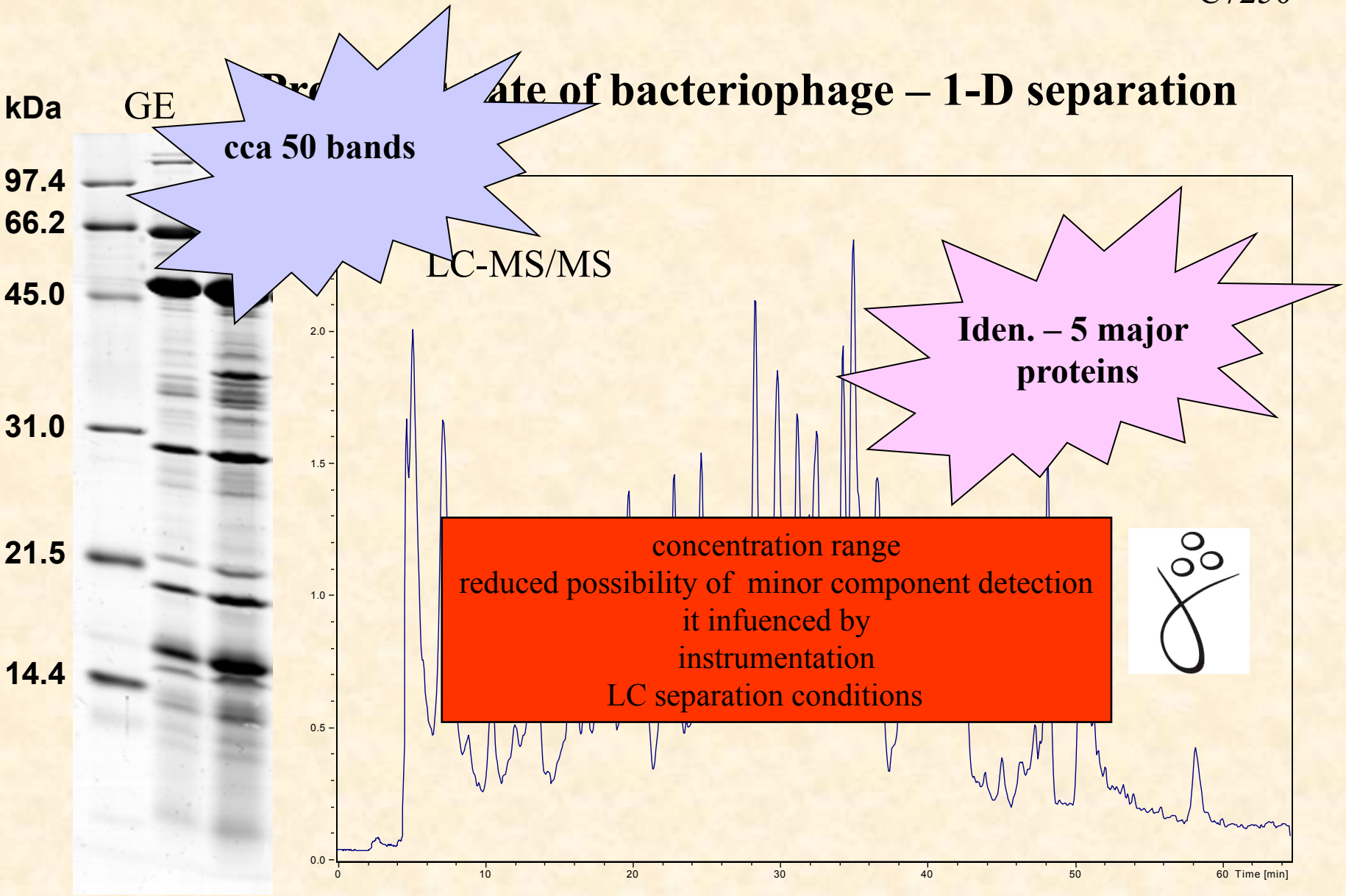
Standard 1-D approaches

„SIMPLE“ MIXTURES

1-D GE



ate of bacteriophage – 1-D separation



Classical 2-D approaches

2-D GE

separation protein level → protein digestion → MALDI-MS / MALDI-MS/MS; LC-MS/MS

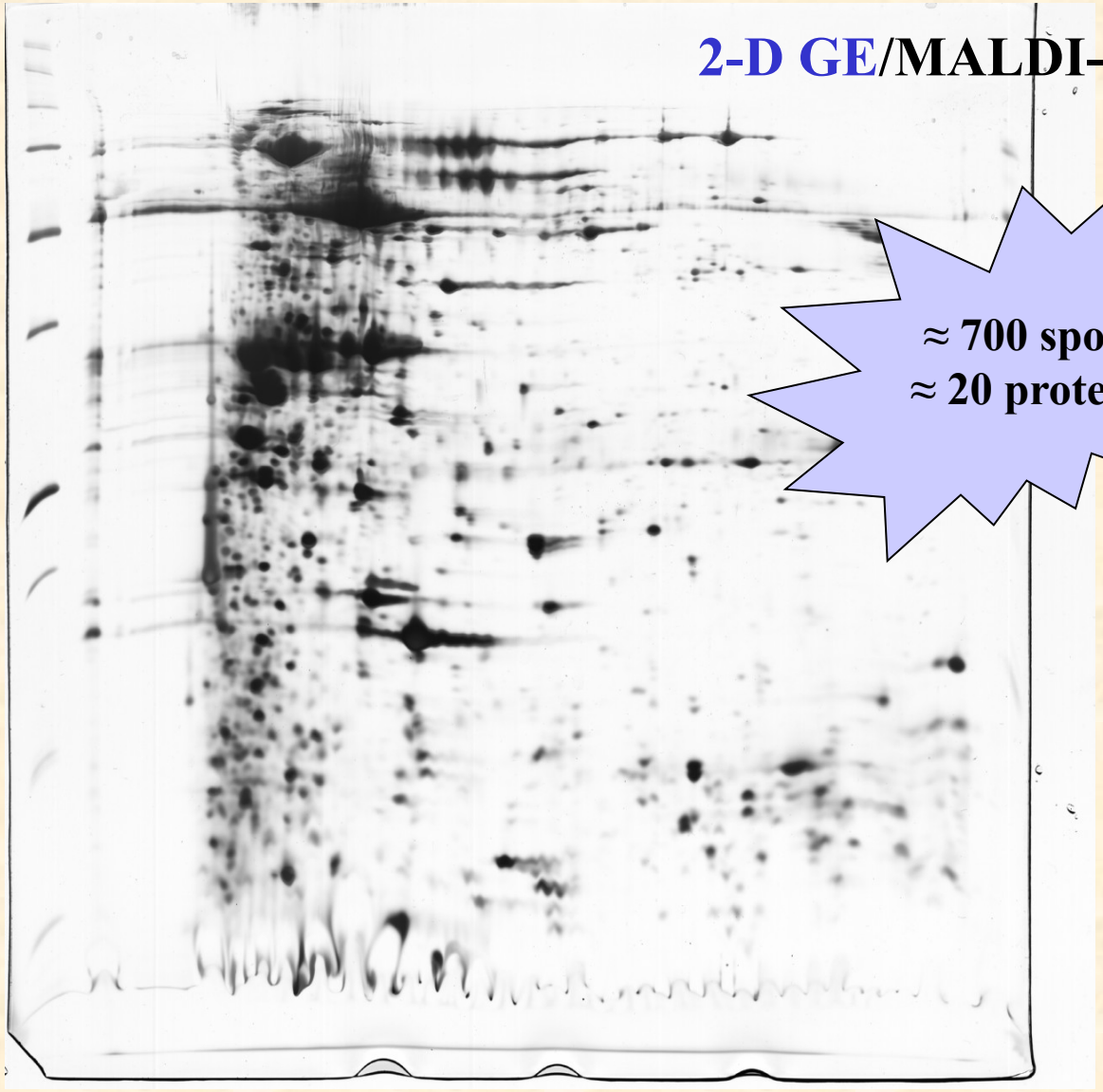
protein digestion → separation peptide level → MS/MS

2-D LC-MS/MS

on-line/off-line

LC-MALDI

Protein isolate of bacteriophage – 2-D separation



Protein mix

2-D LC
peptides

digestion

On-line ("MudPIT")

1D - SCX
fractionation
step by step

2D - RP

MS/MS

Off-line

1D - SCX
LC

UV

Peptide fraction

MS/MS

2D - RP

Peptide fraction

MudPIT (multidimensional protein identification technology)

2-D LC

peptides

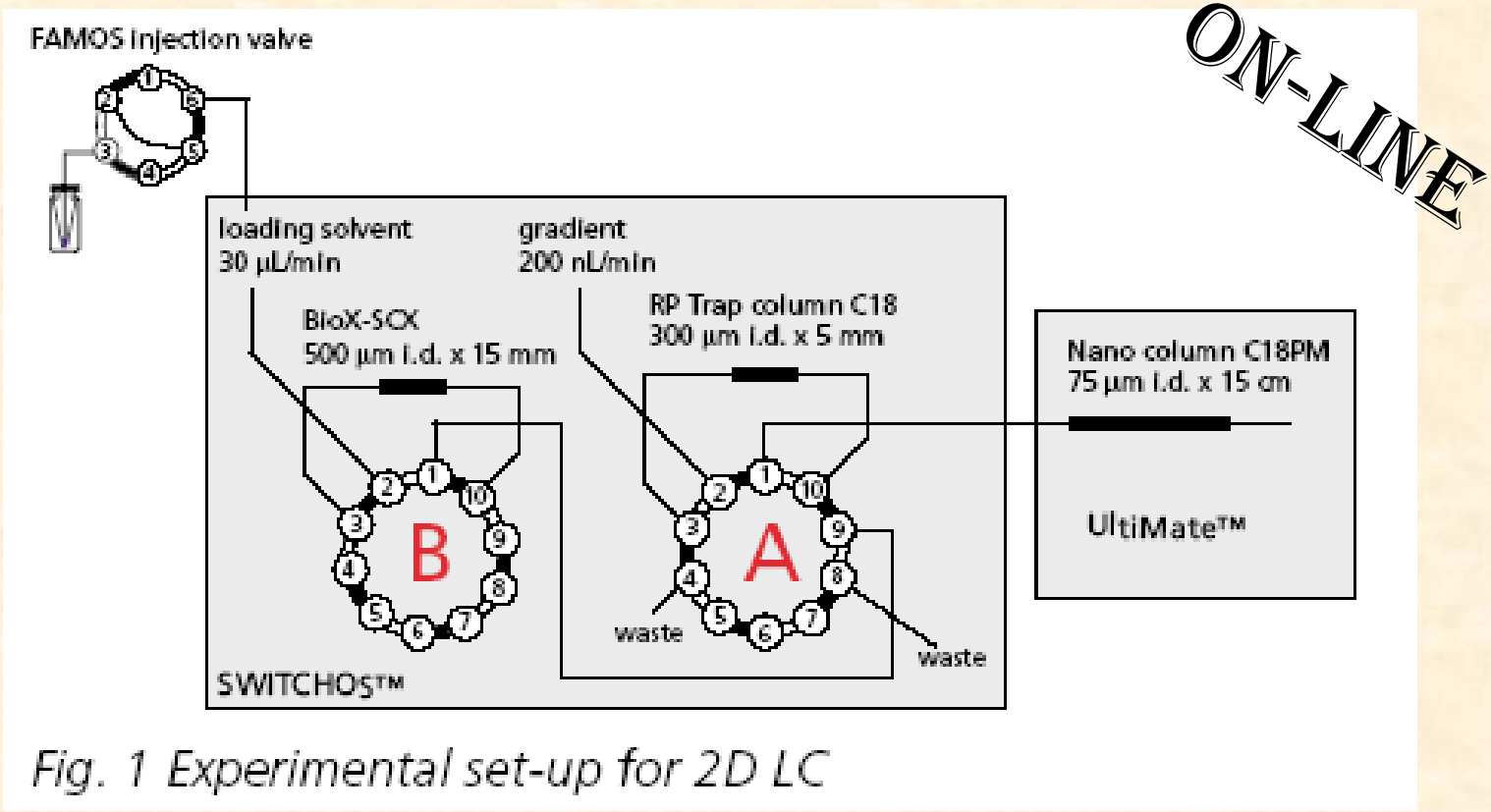


Fig. 1 Experimental set-up for 2D LC

2-D LC (peptides)

ON-LINE

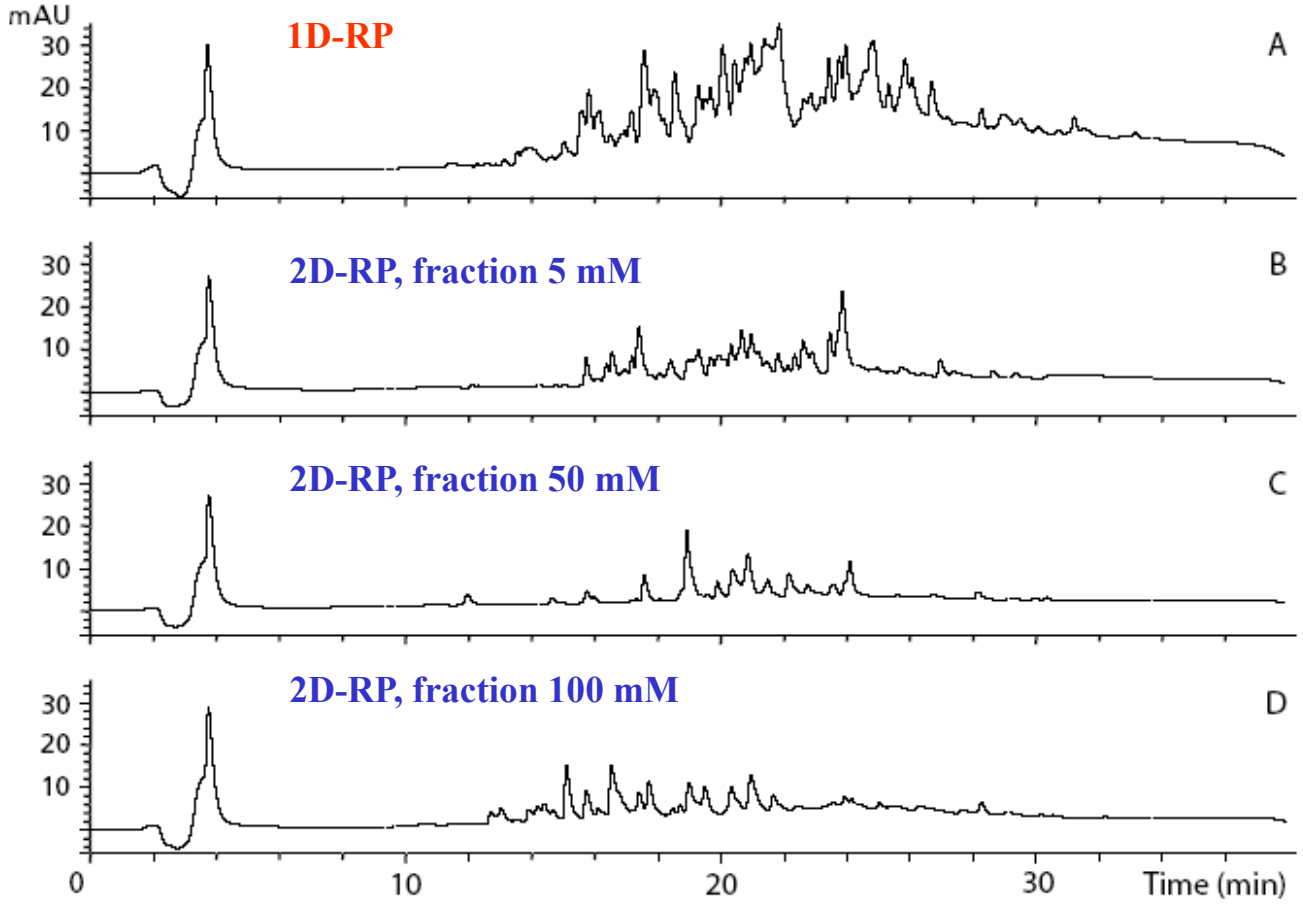


Fig. 2 Chromatograms of the digest mixture. Upper trace is the result of a separation without SCX. The next chromatograms are the result of the 5, 50 and 100 mM fractions. (Not all fractions are shown)

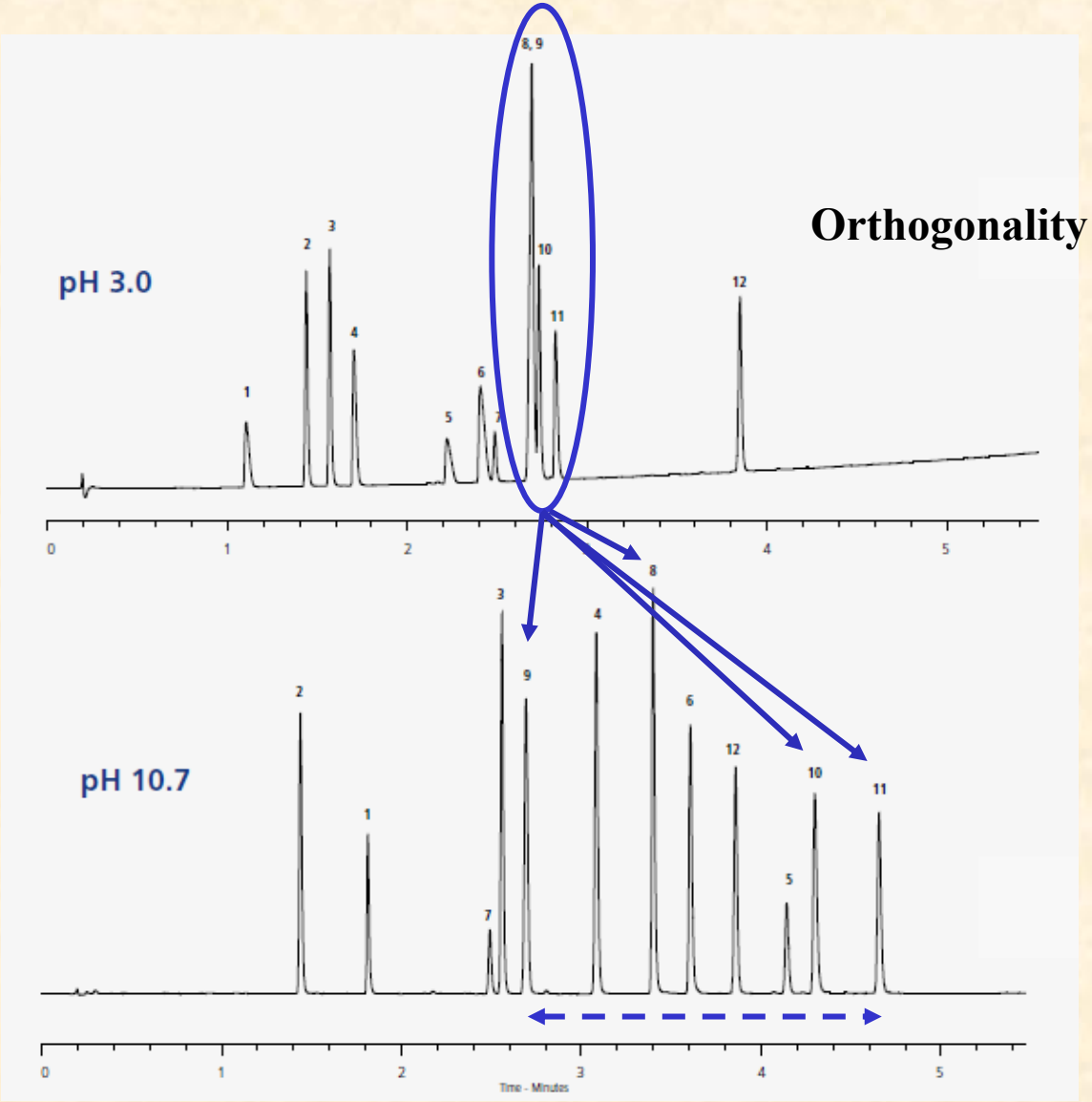
Dionex application note

2-D LC (peptides)

C7250

RP-RP

the same stationary phase – jiné pH



Orthogonality of Separation

www.ace-hplc.com

Characterization of proteome and phosphoproteome of HEK293 cells

cooperation with Assoc. Prof. Bryja group, FS MU

lysis
SDT buffer [SDS+DTT+Tris]

FASP

TMT labeling

direct analysis
w/o and with SPL
LC-MS/MS (pH 3)
(Orbitrap Elite)

Data processing

1D LC separation

pH 10

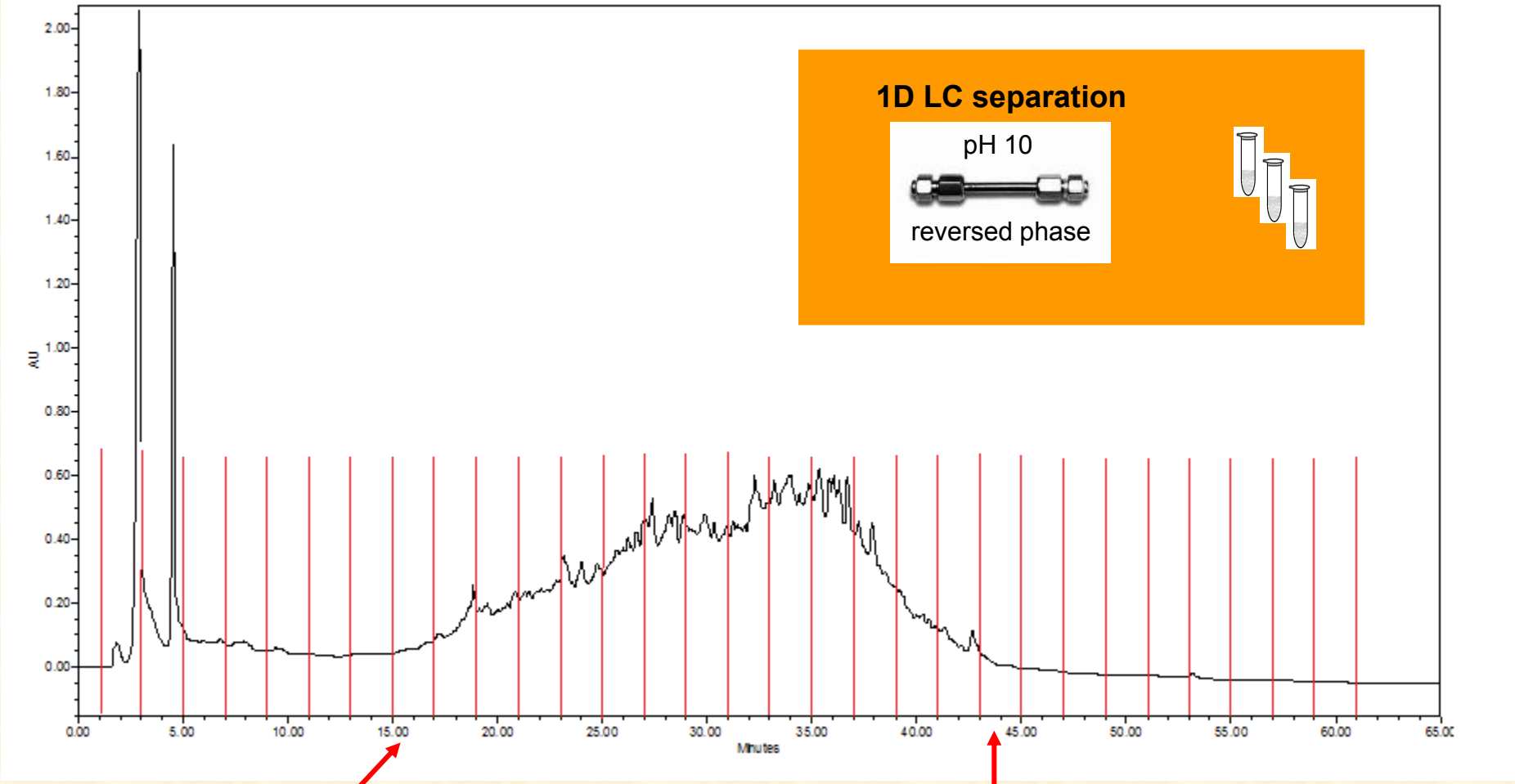
reversed phase

analysis of fractions
LC-MS/MS (pH 3)
(Orbitrap Elite)

Data processing

SPL – „scheduled precursor list“ analysis enables repeated analysis of sample with exclusion of already identified peptides in previous analysis

LC-separation of the digested sample in 1D (high pH)



1 2 3 4 5 6 7 8 **9 10 11 12 13 14 15 16 17 18 19 20 21 22 23** 24 25 26 27 28 29 30

direct analysis
w/o and with SPL
LC-MS/MS (pH 3)
(Orbitrap Elite)

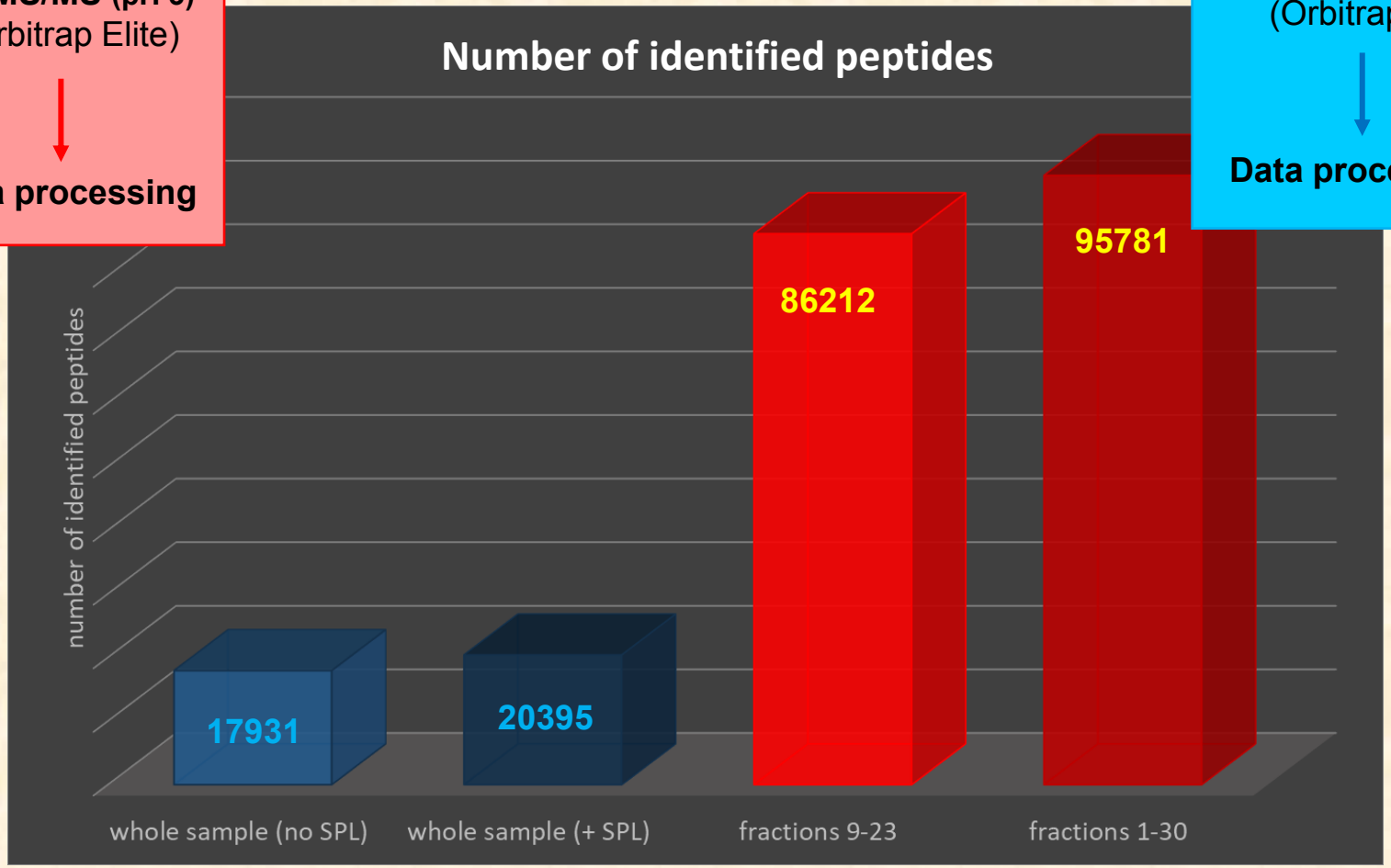
↓

Data processing

analysis of fractions
LC-MS/MS (pH 3)
(Orbitrap Elite)

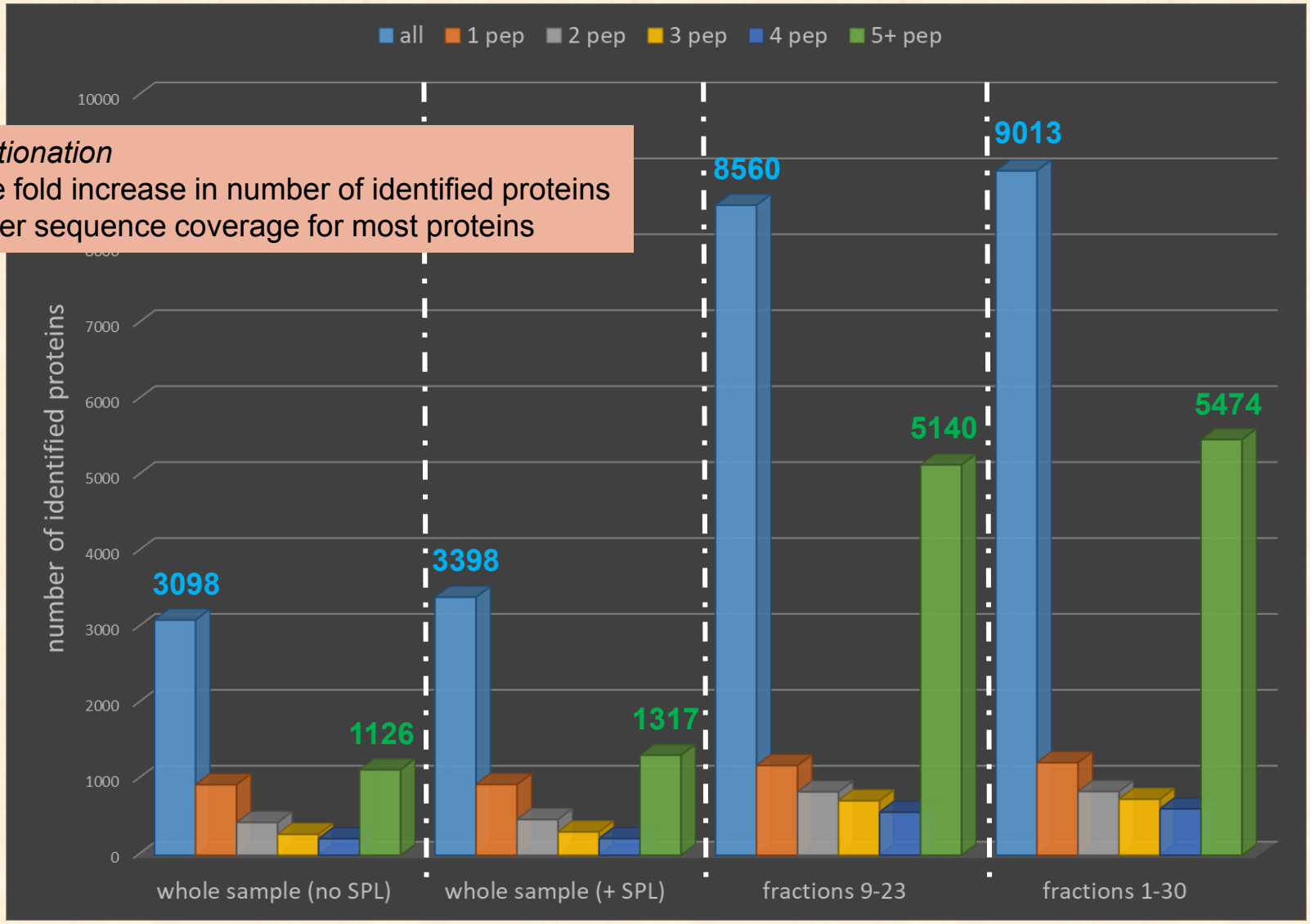
↓

Data processing



Number of identified proteins

Fractionation
three fold increase in number of identified proteins
Higher sequence coverage for most proteins



Diameter of HPLC column *vs* sensitivity

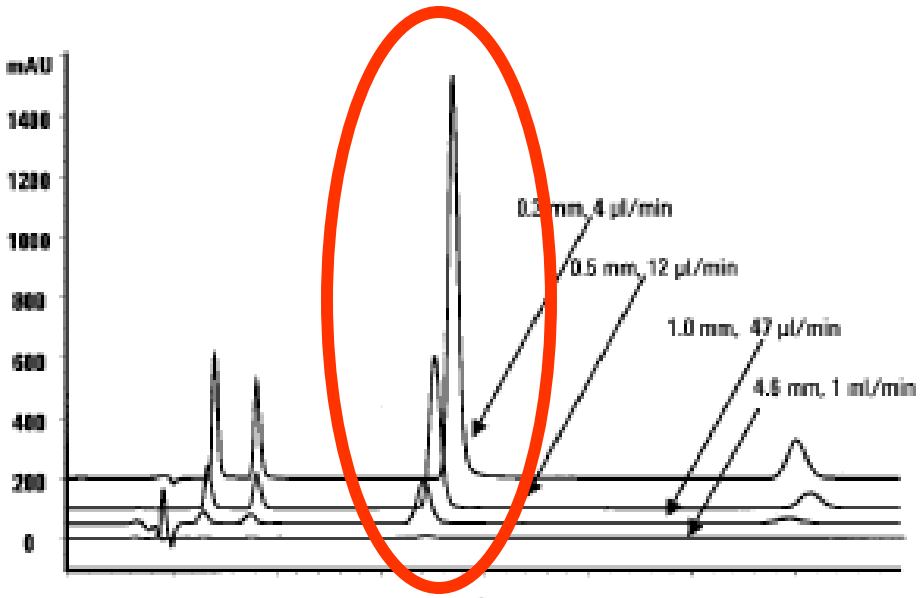


Figure 5 *Mass sensitivity benefit. Injection of the same sam - ple amount on HPLC columns with decreasing internal di - ameter. Stationary phase: ZORBAX® SB-C18; length: 150 mm; solvent: water/acetonitrile, 40/60; flow rate: see dia - gram; sample: isocratic checkout sample; injection volume: 0.1 µL; third peak: biphenyl, 200 ng; temperature: 25 °C; detection wavelength: 230 nm.*

*„Sensitivity increases with a decrease in column diameter because the **same sample mass (amount)** is eluted in a smaller volume. Therefore the concentration of the eluting peak is higher and the detection signal is stronger.“*

*Amer. Lab., 2001, **33** (10), 26-38.*

Capillary and nano columns

- Increase in sensitivity
- reduction of injected sample amounts
- reduced consumption of solvents

Table 27. Sensitivity Increase

COLUMN I.D. (mm)	TYPICAL FLOW RATE (µL/min)	THEORETICAL SENSITIVITY INCREASE[1]
4.6	1000	1
1.0	40	21
0.5	10	85
0.3	3	235
0.1	0.5	2100
0.075	0.3	3760

[1]For same sample mass

2-D LC

peptides

sorbents

1-D: ionex
reverse phase
HILIC
IMAC (phospho)
affinity (e.g. lectin – glyco)

2-D: **reverse phase**

On-line *vs* Off-line

automation

flexibility
optimalization
continuous collection of fractions

LC –MALDI

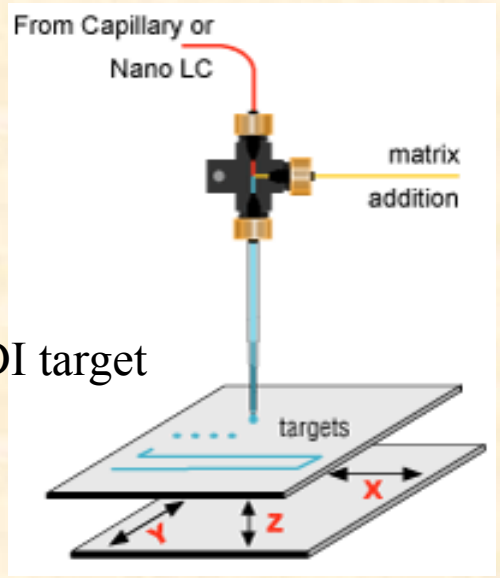
(peptides)

Peptide fraction

nano LC (RP)

on-line

ESI-MS/MS



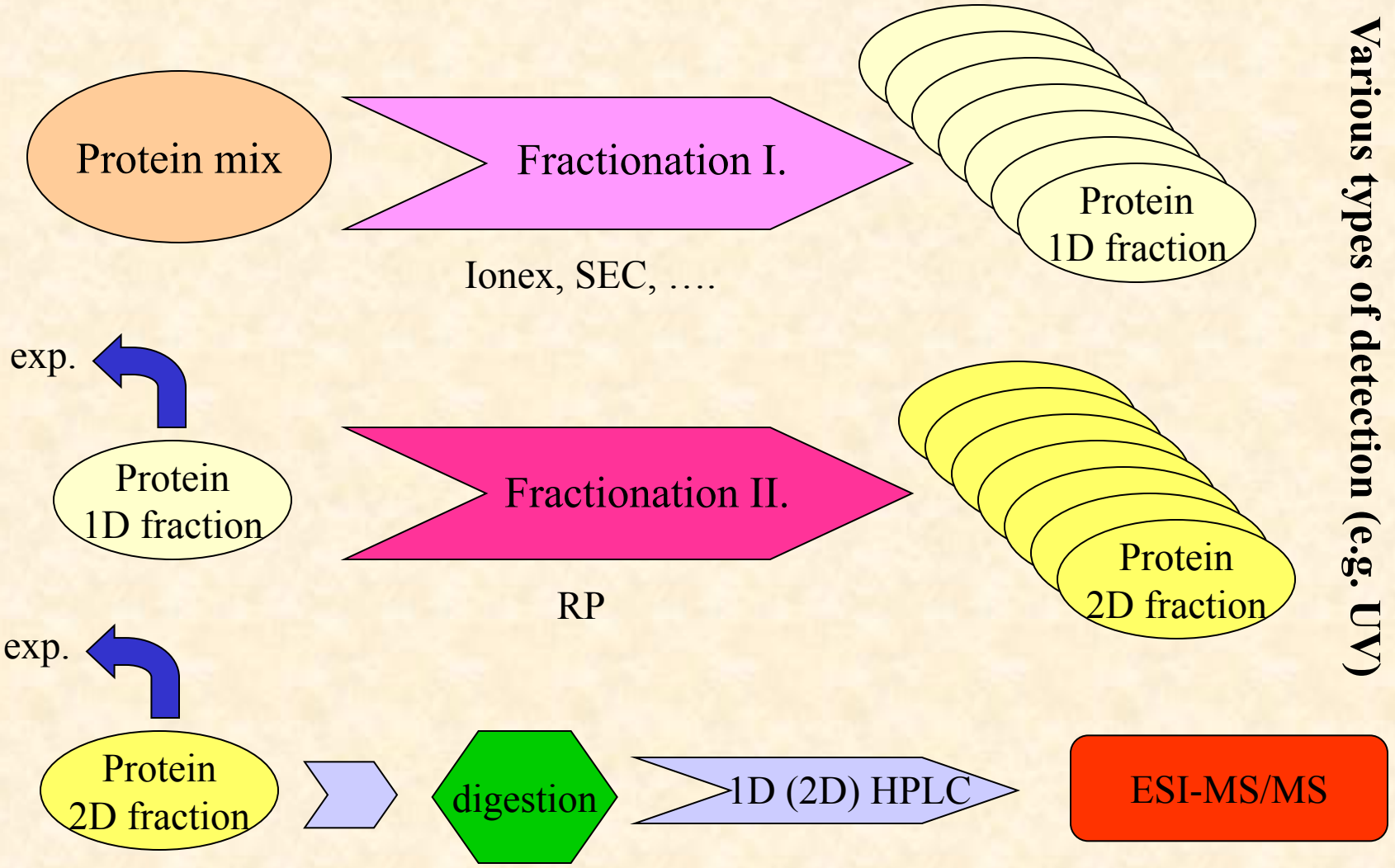
Deposition on MALDI target

off-line

MALDI TOF/TOF MS

Sample storage

LC separation of complex protein mixtures



LC separation of complex protein mixtures

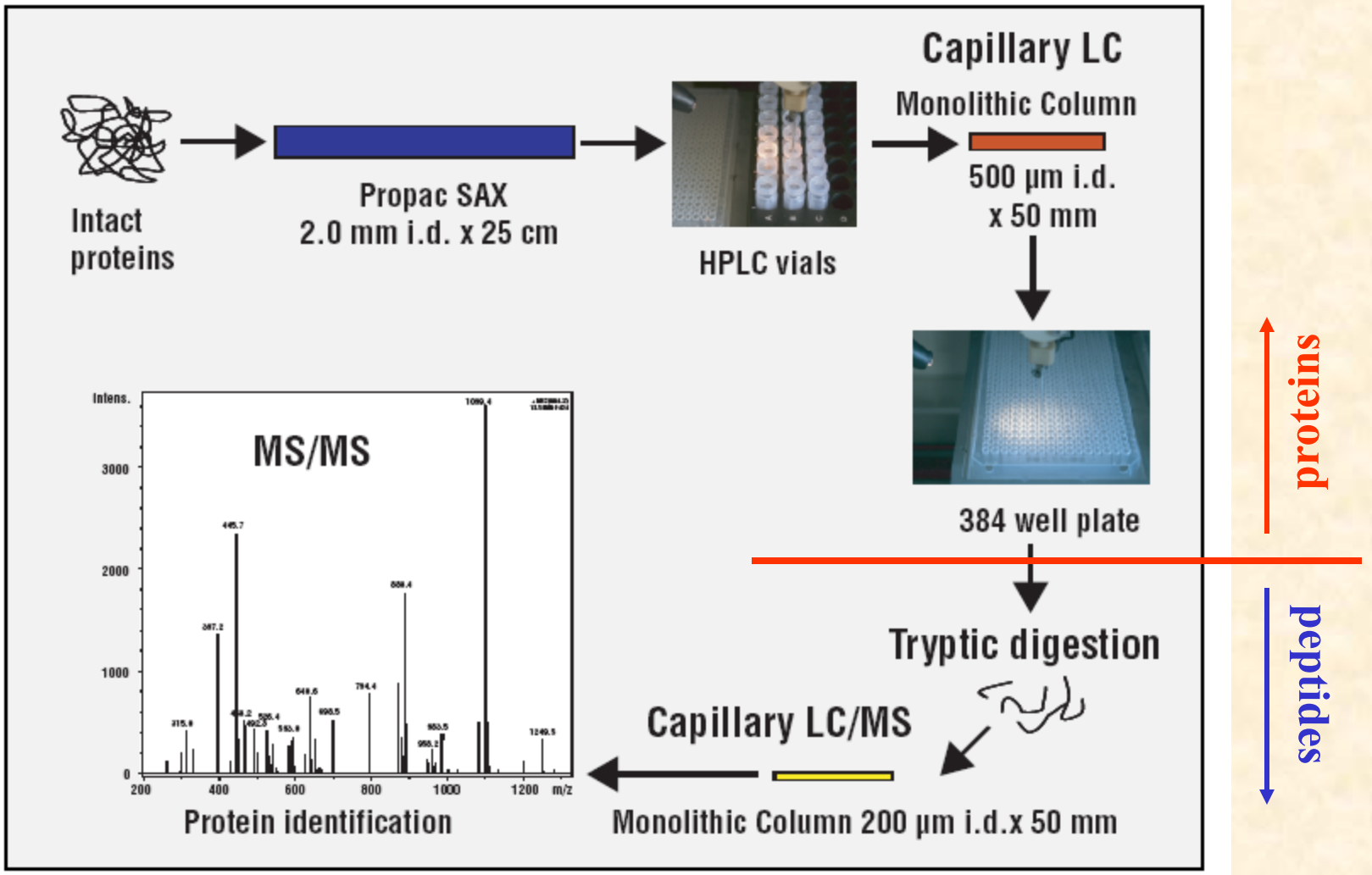
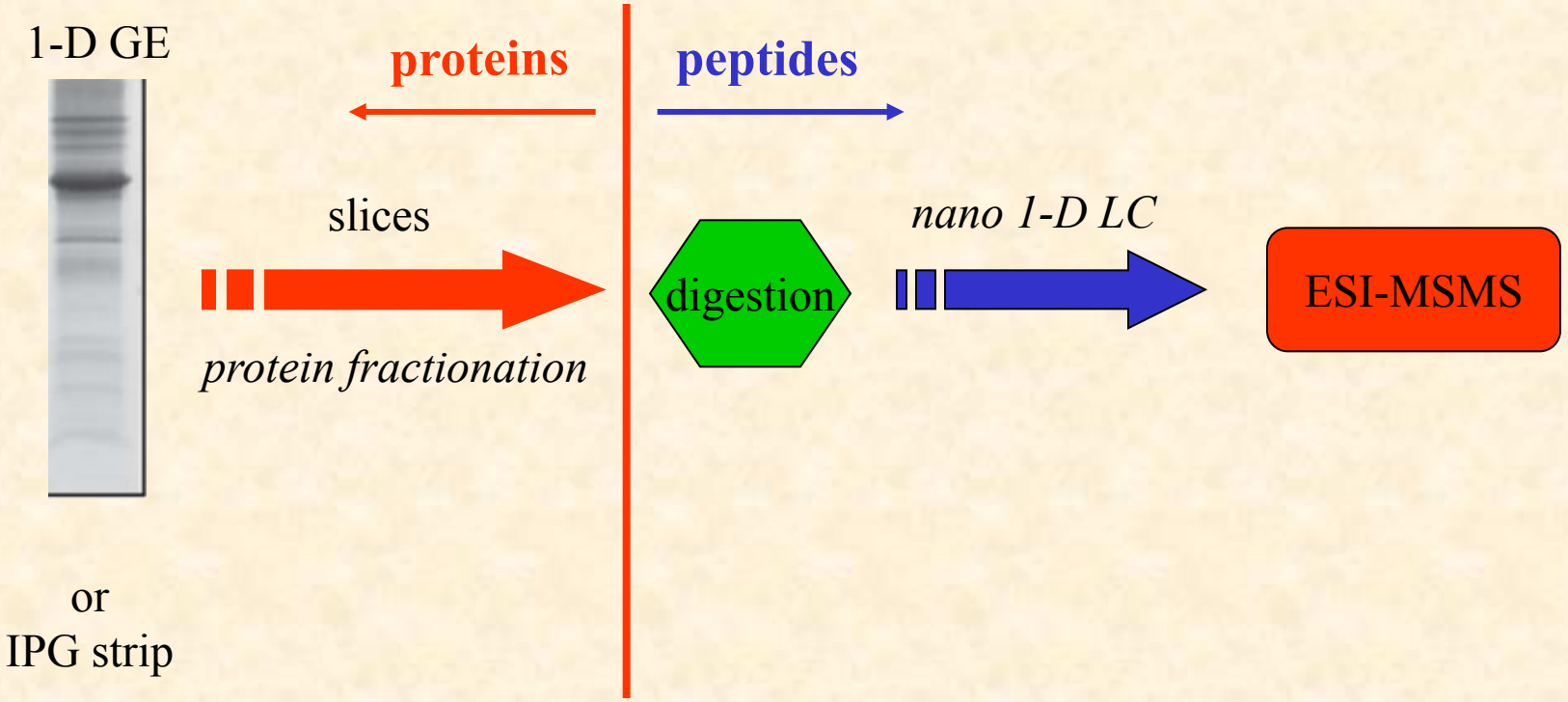
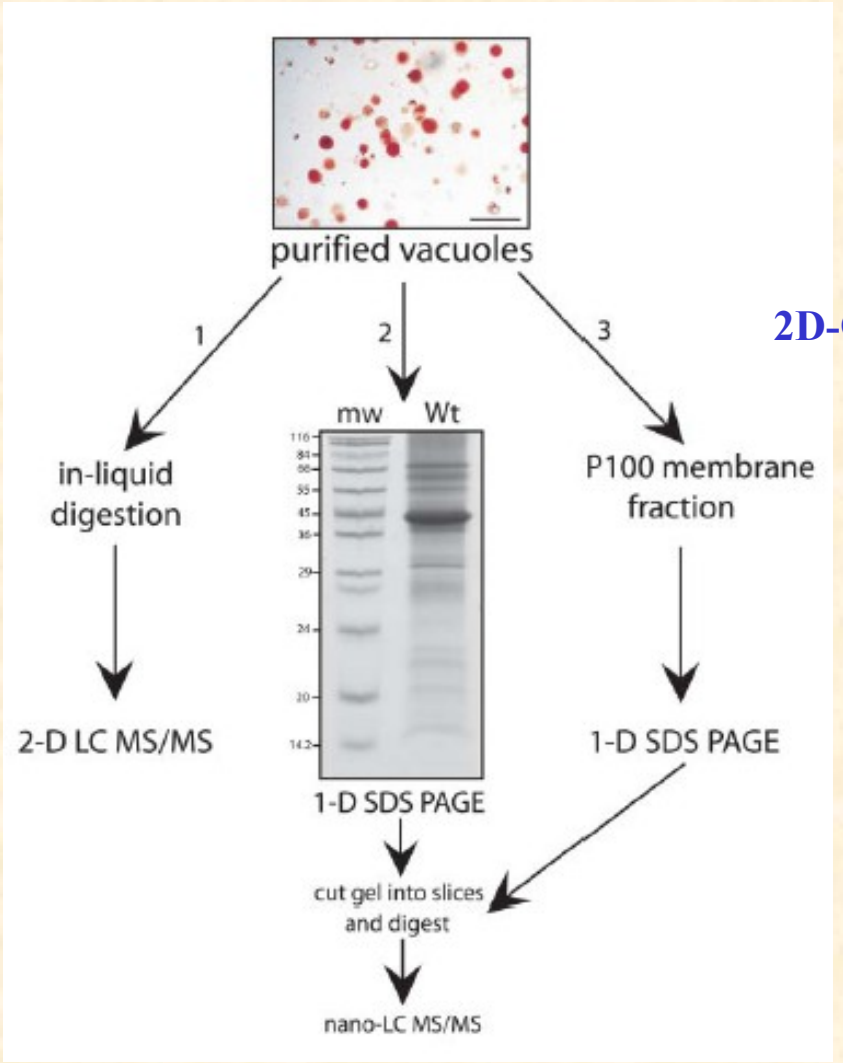


Figure 1: Multidimensional LC work flow.

Combination of GE a LC separation





2D-On-line

2D-Off-line (3D ?)

*from Carter et. al. The Plant Cell, 2004, **16**, 3285–3303.*

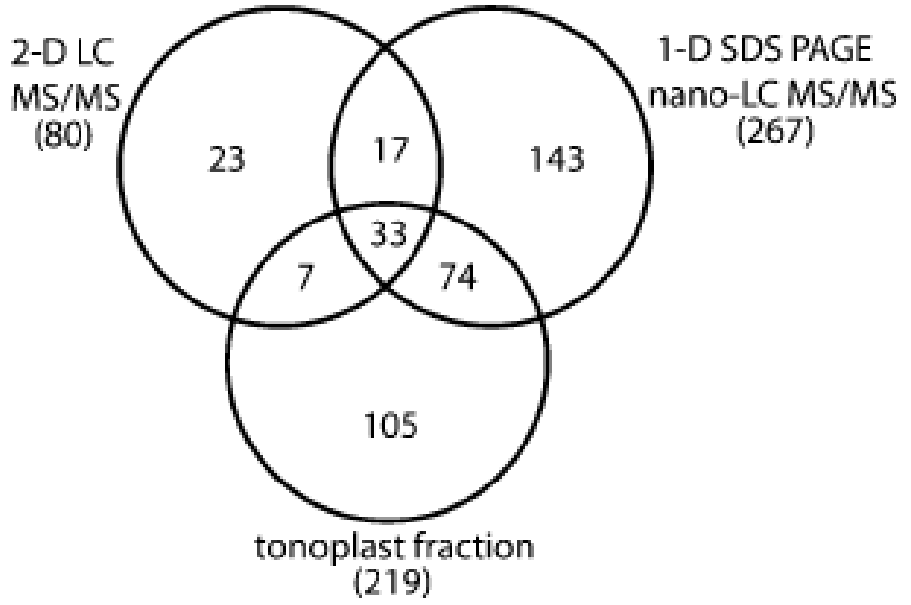


Figure 2. Distribution of Identified Proteins by Different Methods.

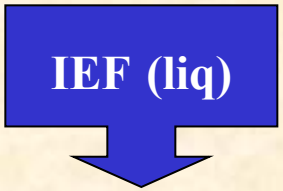
Overlap of the different protein sets is shown. Numbers in parentheses indicate the total number of proteins found by a particular method.

*from Carter et. al. The Plant Cell, 2004, **16**, 3285–3303.*

Example of multidimensional proteome analysis (screening)

Depleted blood plasma
(3500 – 9000 proteins ??)

0. dimension



1. dimension

20 fractions

2. dimension

1600 fractions

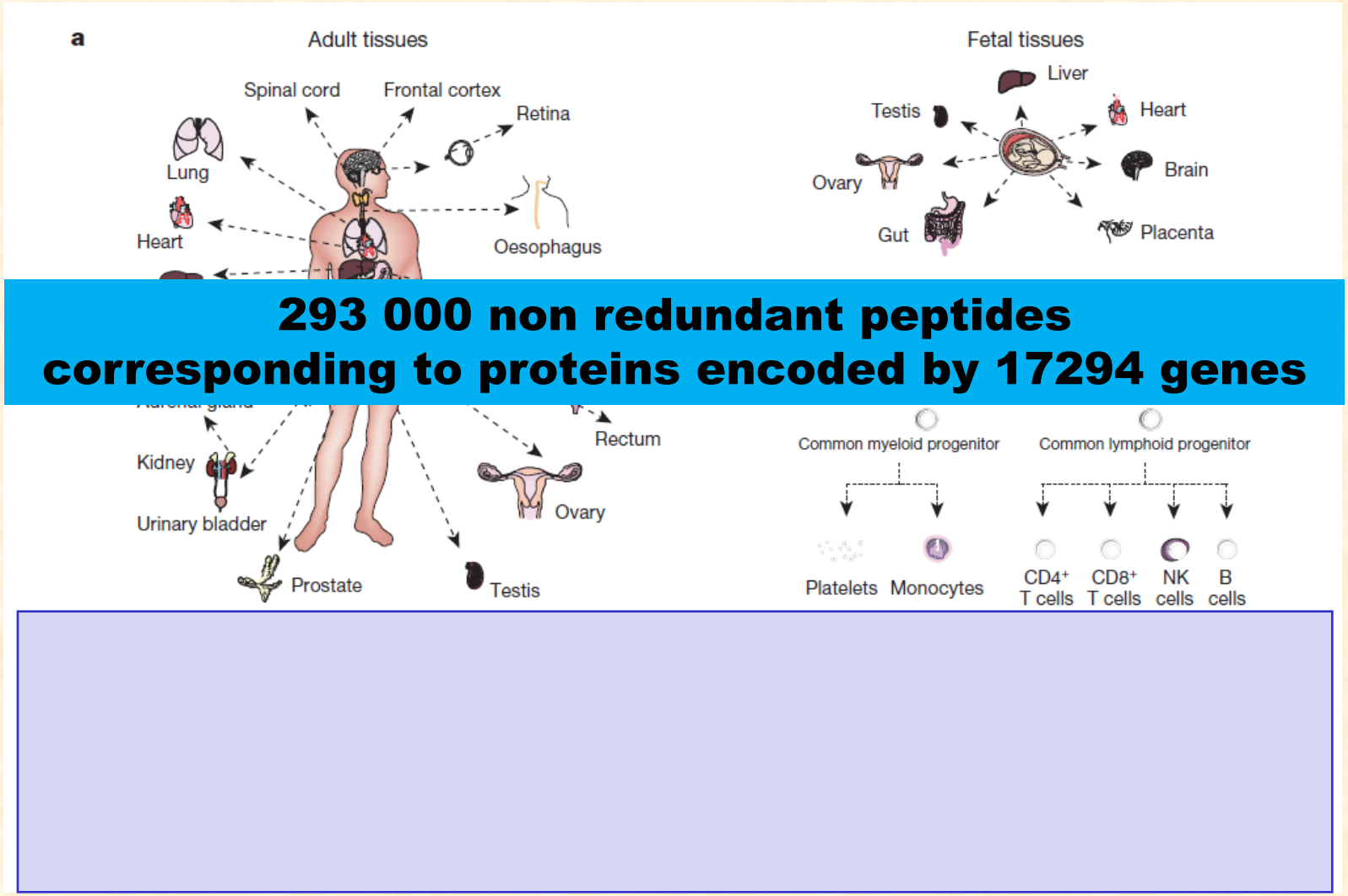
3/4. dimension

“ ∞ “ fractions

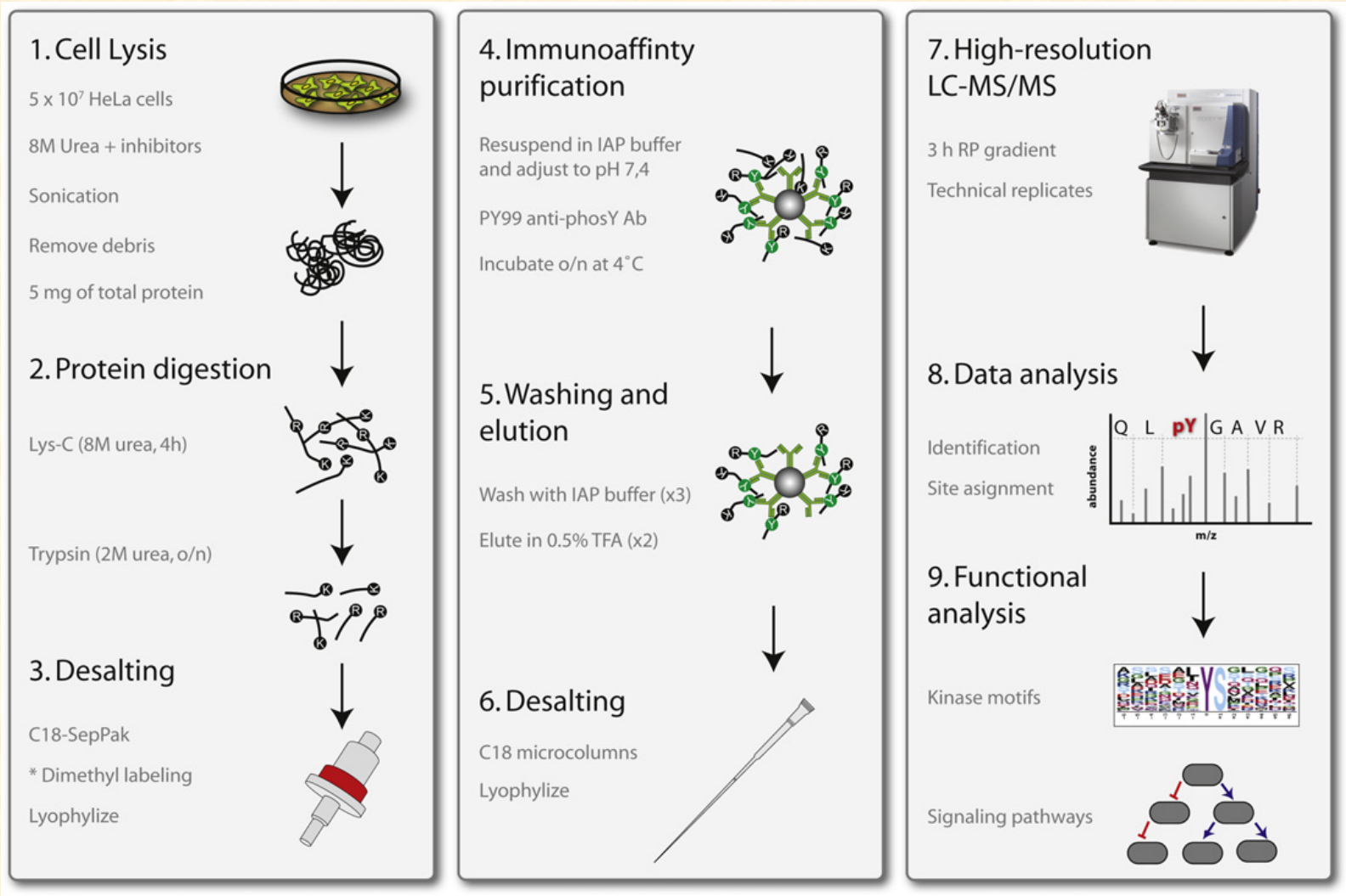
*from H. Wang, Molecular & Cellular Proteomics, 2005, **4**, 618–625.*

A draft map of the human proteome

Min-Sik Kim et al., Nature 509, 575-581 doi:10.1038/nature13302

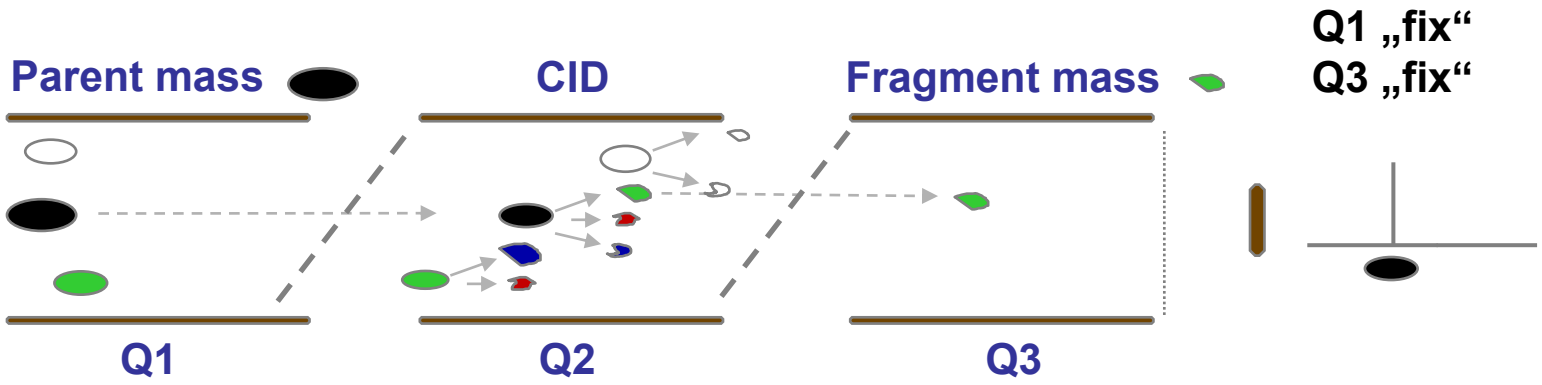


Targeted separation - immunoaffinity fractionation (Y(phos))



from A.D. Zoumaro-Djayoon et al. / Methods 56 (2012) 268–274

Targeted MS analysis of selected proteins



- quadrupole **Q1** and **Q3** are fixed to selected values of *m/z* (Q1-precursor and Q3- selected fragment), **only precursors displaying production of selected fragment** during fragmentation in collisional cell **are recorded**
- enables to follow tens of reactions (transitions) during analytical run (MRM)

from K. Bluemlein et al. / Nature protocols 6 (2011) 859 –869

High throughput

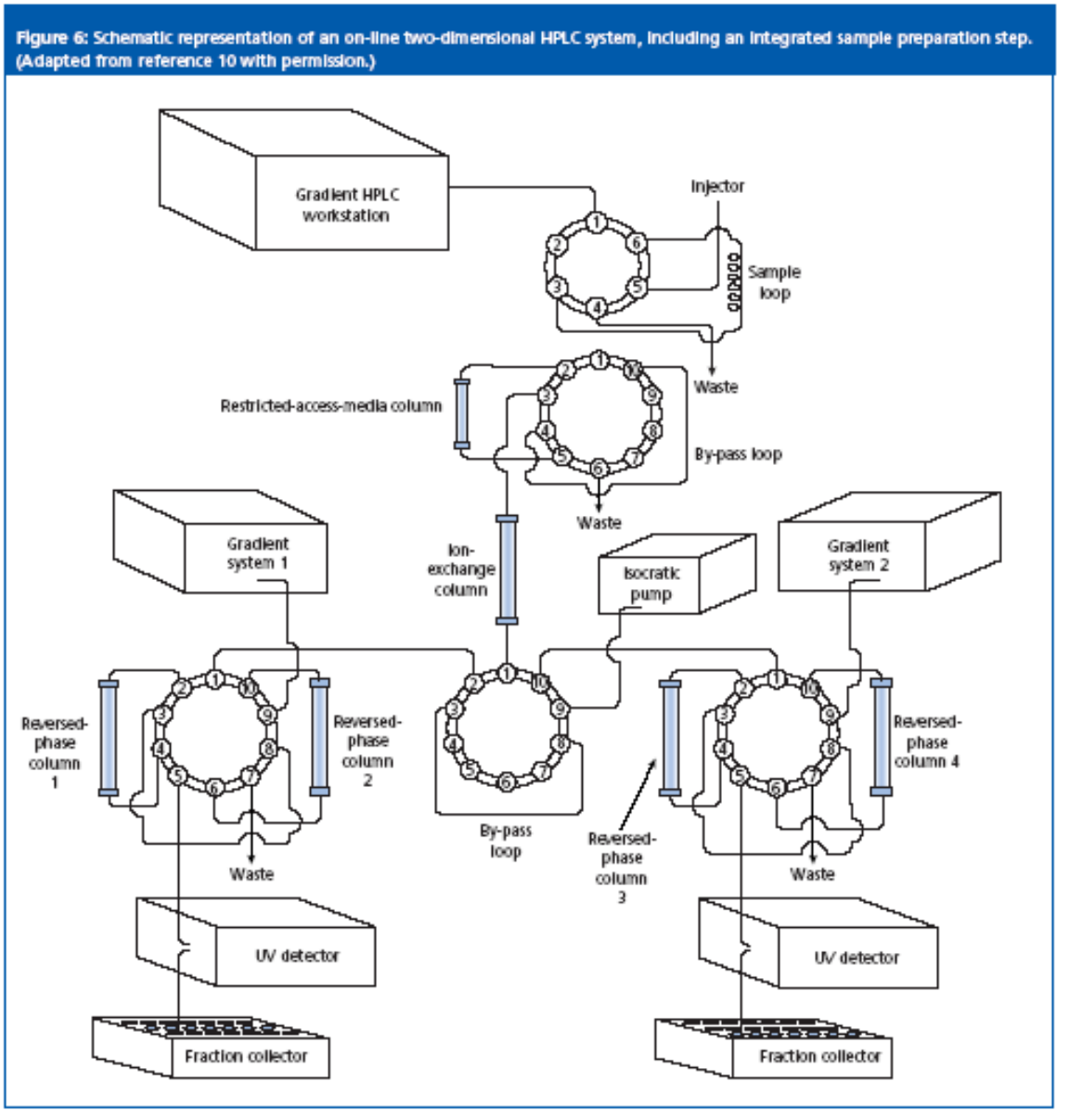
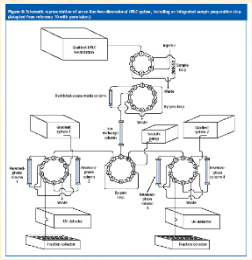


Figure 6: Schematic representation of an on-line two-dimensional HPLC system, including an integrated sample preparation step. (Adapted from reference 10 with permission.)

+ Miniaturization - chip technology



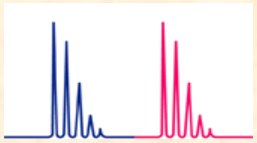
Functional Genomics and Proteomics
National Centre for Biomolecular Research
Faculty of Science Masaryk University

Protein quantification by MS

Protein quantification by MS

Approaches:

using isotopically different tags



- **Absolute quantification**
  determination of protein concentration (amount) by addition of corresponding standard with known amount (AQUA, PSAQ)

- **Relative quantification**
  evaluation of relative changes of the protein content in compared samples

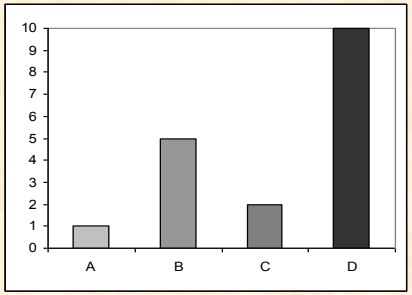
label free

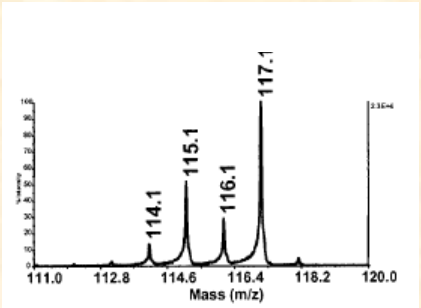
metods of absolute and relative quantification based on statistical processing of MS, or MS/MS data
advantage of this approach is possibility of comparison of unlimited number of samples and absence of derivatization reaction or isotopically labeled standards

Relative quantification approaches

Protein sample A

Protein sample B

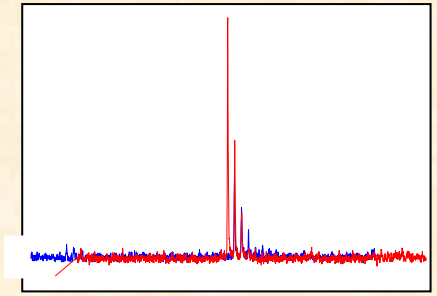




MS/MS

Separated digestion of samples

iTRAQ, $^{16}O/\ ^{18}O$

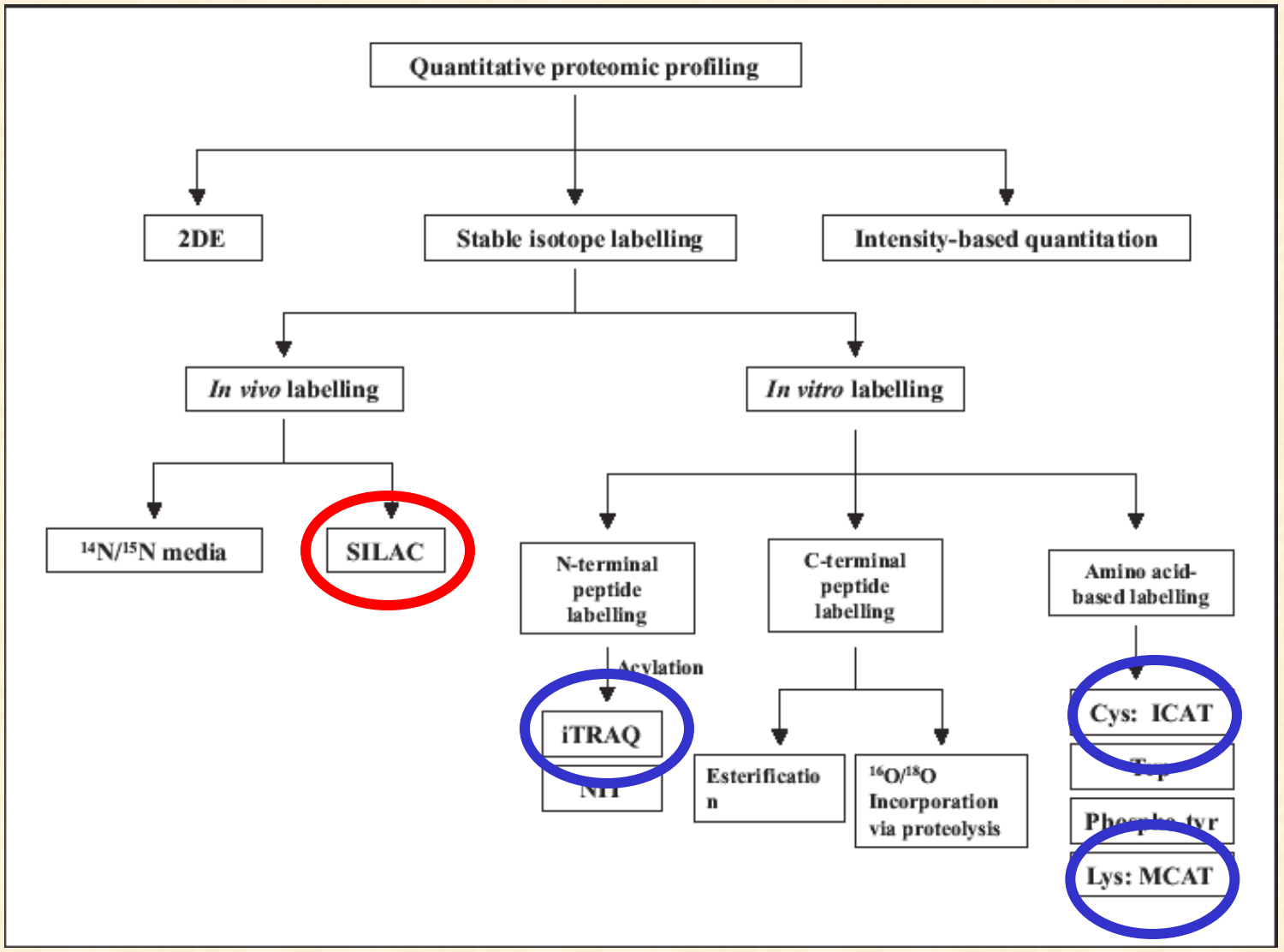


LC-MS

LC-MS/MS (iTRAQ)

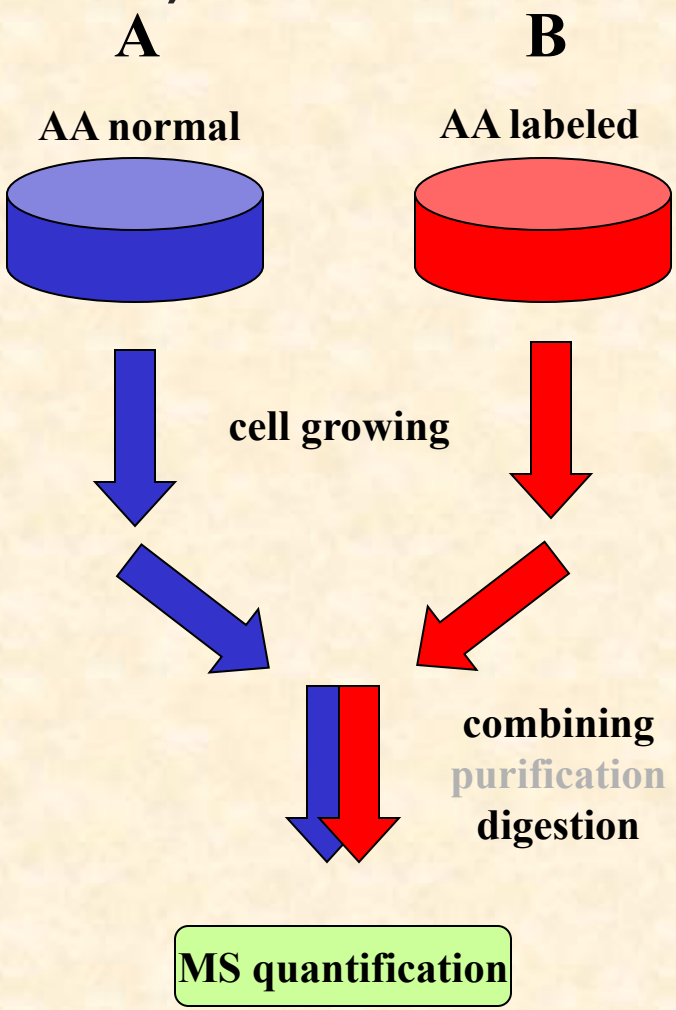
Identical peptides in MS mode are not distinguished

Overview of relative quantification methods

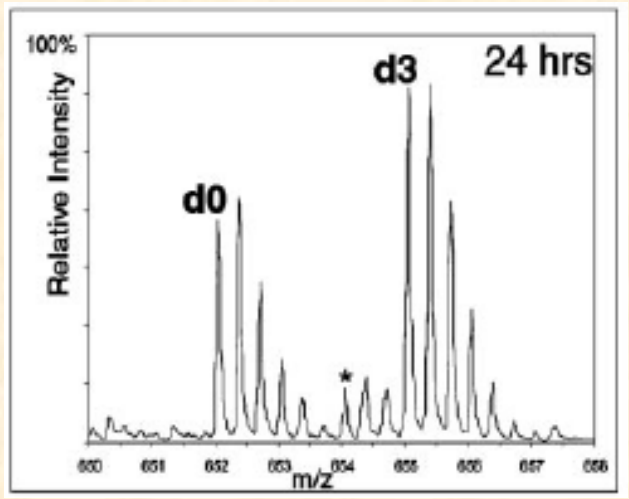


W. Yan, S.S. Chen, *Briefings in functional genomics and proteomics* **4** (1), 1–12, (2005)

Stable Isotope Labelling with Amino acids in Cell culture (SILAC)



- *in vivo*
- proteins are labeled by growing cells in media containing isotopically labeled amino acids
  (e.g. ^{2}H-Leu, ^{13}C-Lys, ^{13}C-Tyr,^{13}C-Arg, ^{13}C/^{15}N-Arg)

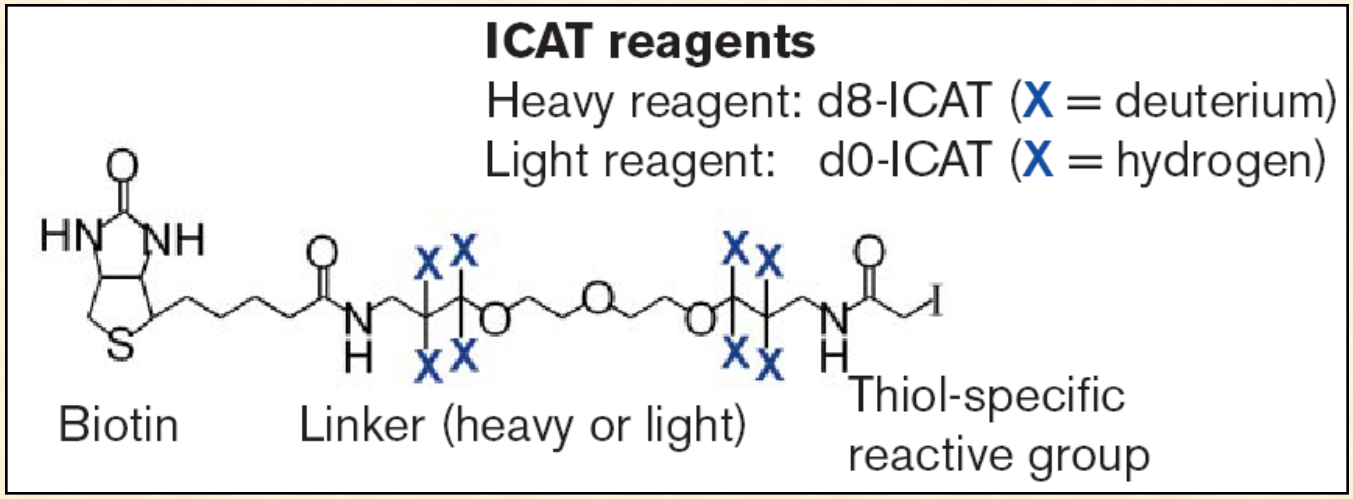


picture from Ong et al.: *MCP* **1**(2002), 376

ICAT ... Isotope-Coded Affinity Tags

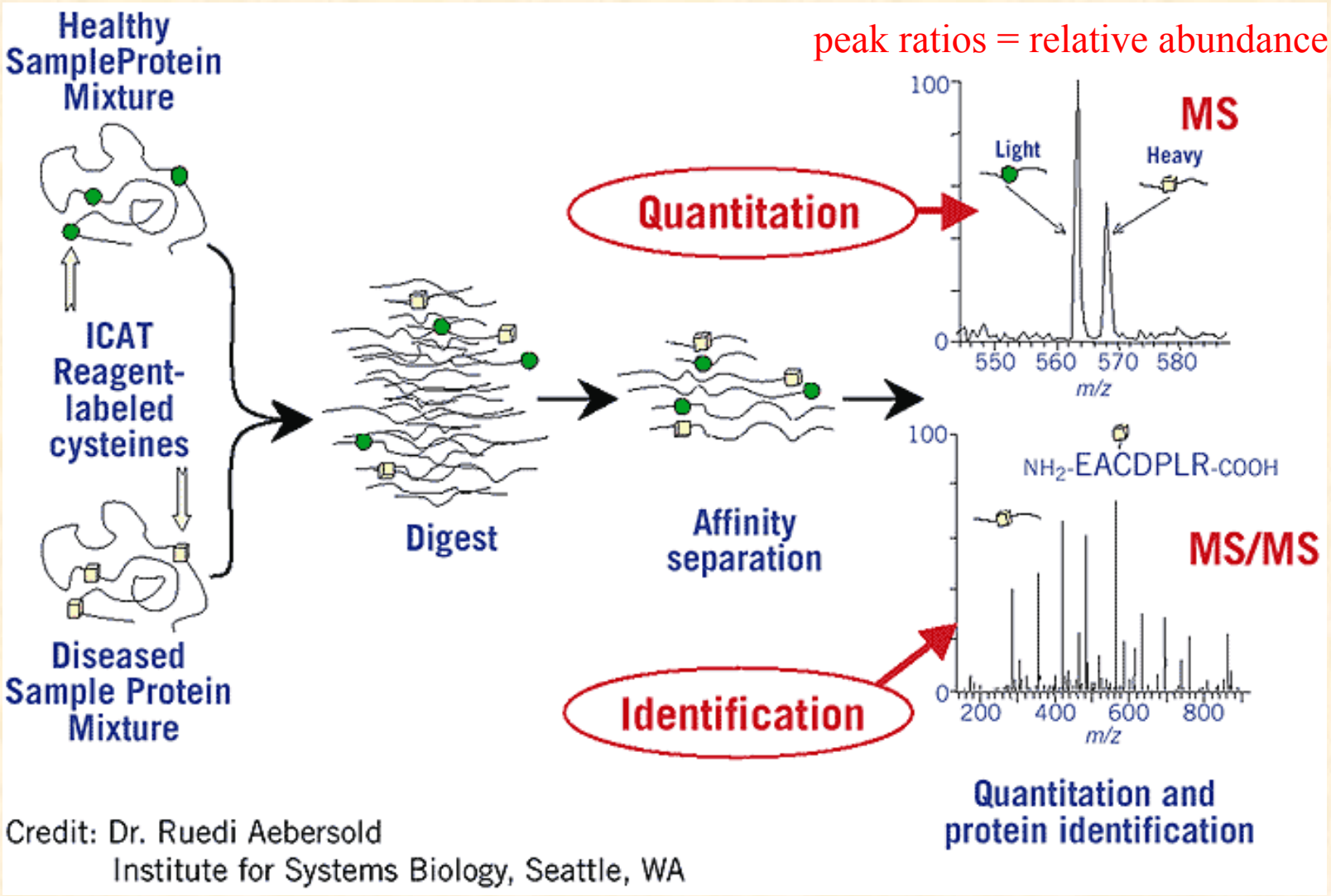
technology for protein expression analysis

- ✓ improved quantitation of a wider range of proteins
- ✓ overcomes limitations of 2-D gel method (e.g. membrane, low abundant proteins)

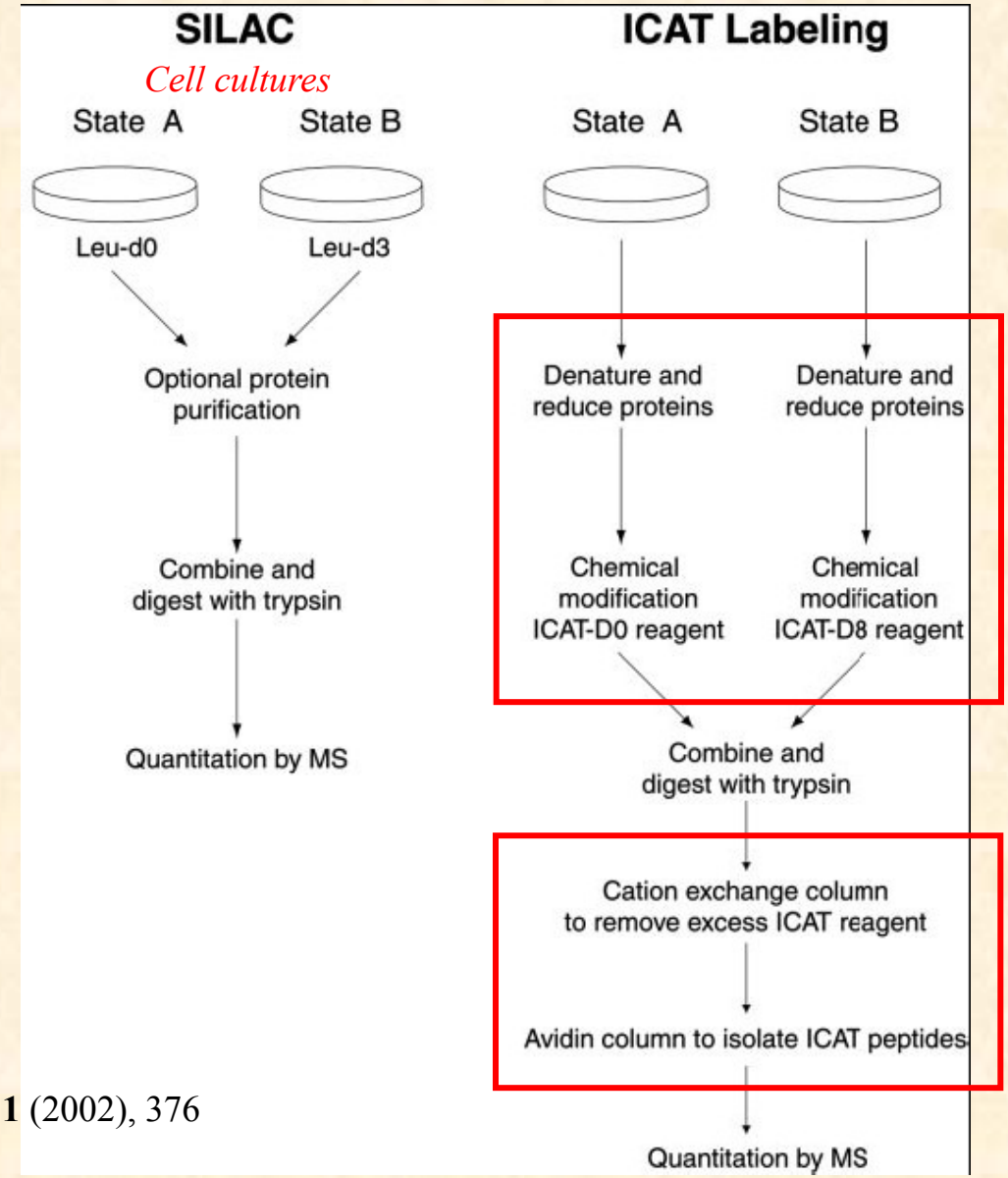


- ➢ tags specific for cysteine-containing peptides (reduction of sample complexity)
- ➢ easy automation of a procedure

ICAT analysis



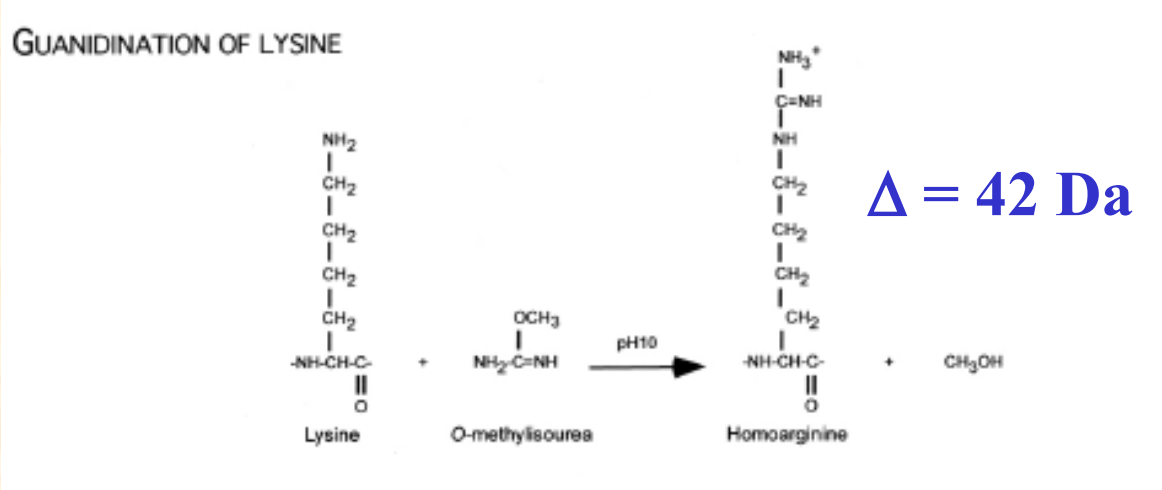
Comparison of *in vivo* and *in vitro* quantification methods (SILAC vs ICAT)



from Ong et al.: *MCP* **1** (2002), 376

Mass Coded Abundance Tagging (MCAT)

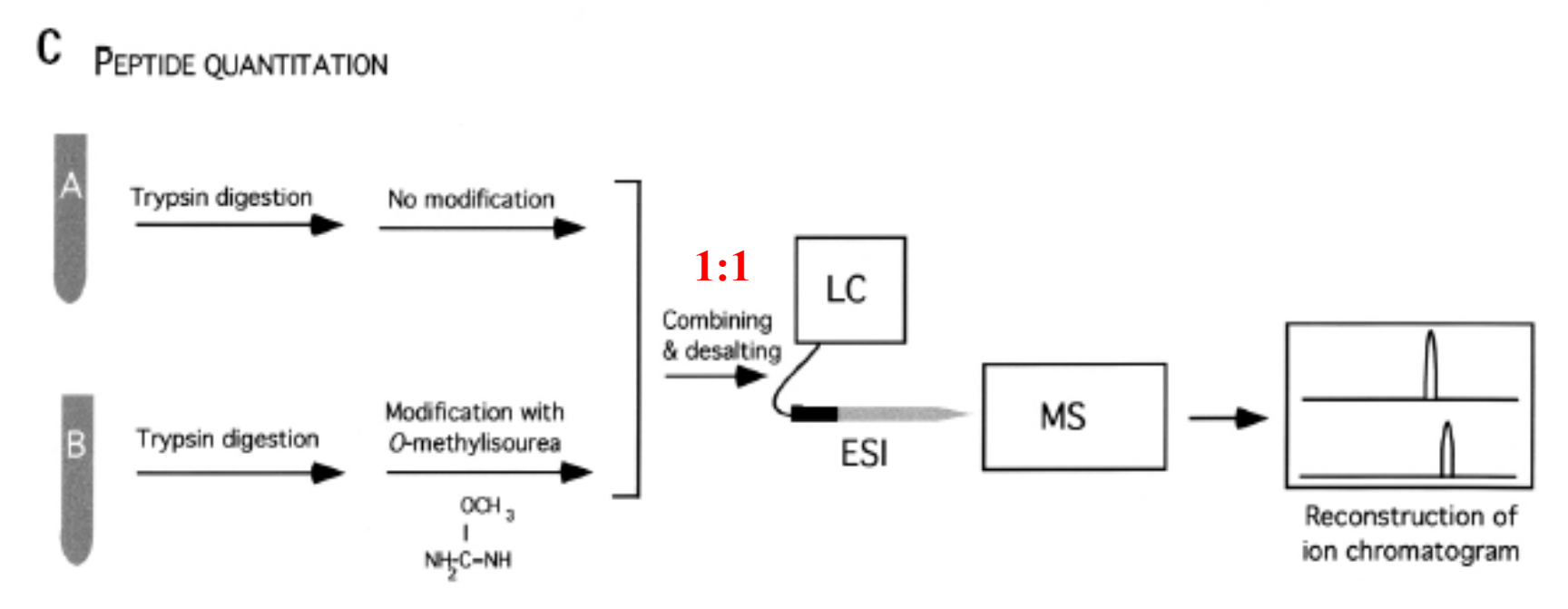
- tryptic digestion
- modification of digest of selected sample (K)



- mixing **nemodif/modif** in ratio 1:1

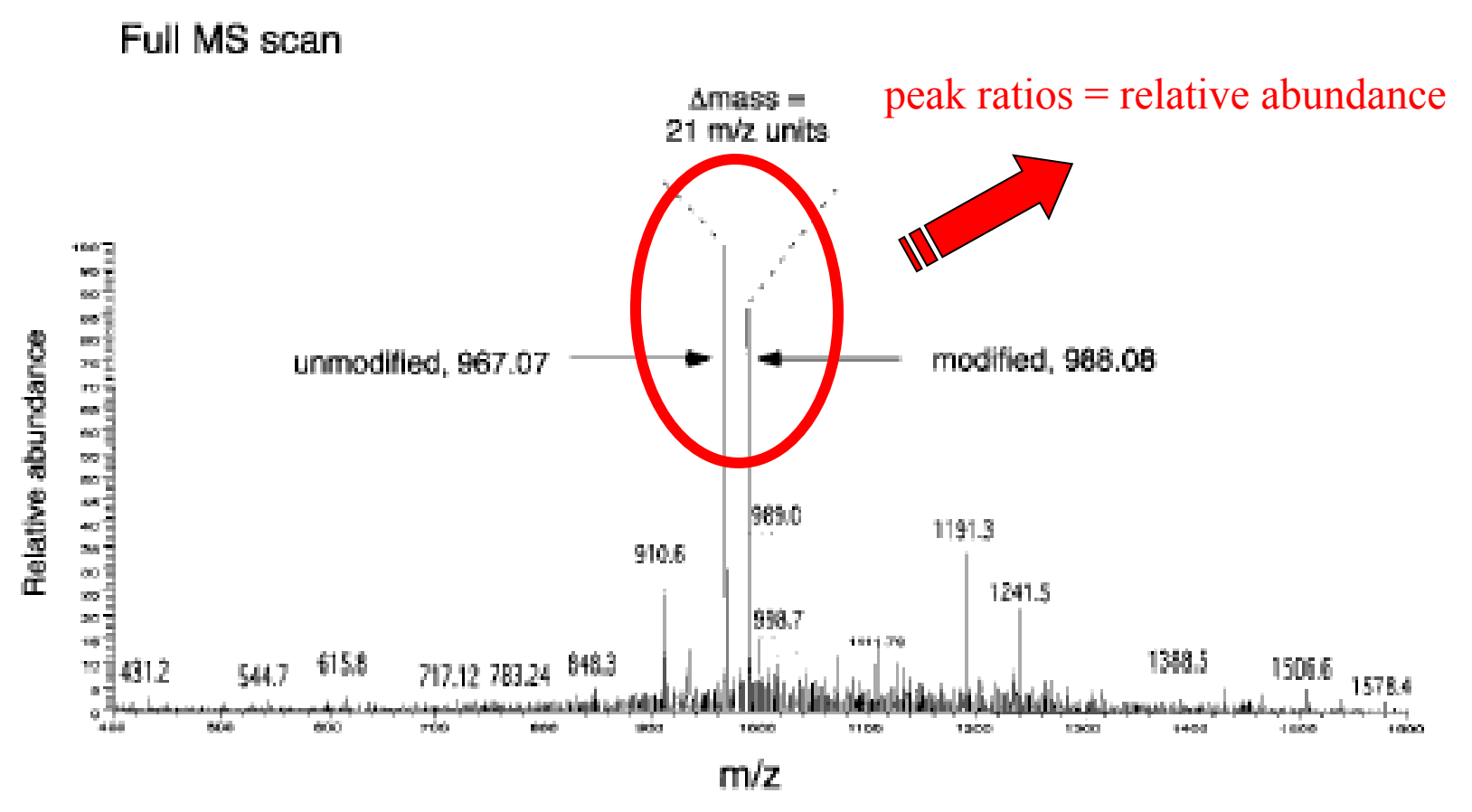
Cagney G., Emili A.: *Nature Biotechnol* **20** (2002), 163-170

MCAT



Cagney G., Emili A.: *Nature Biotechnol* **20** (2002), 163-170

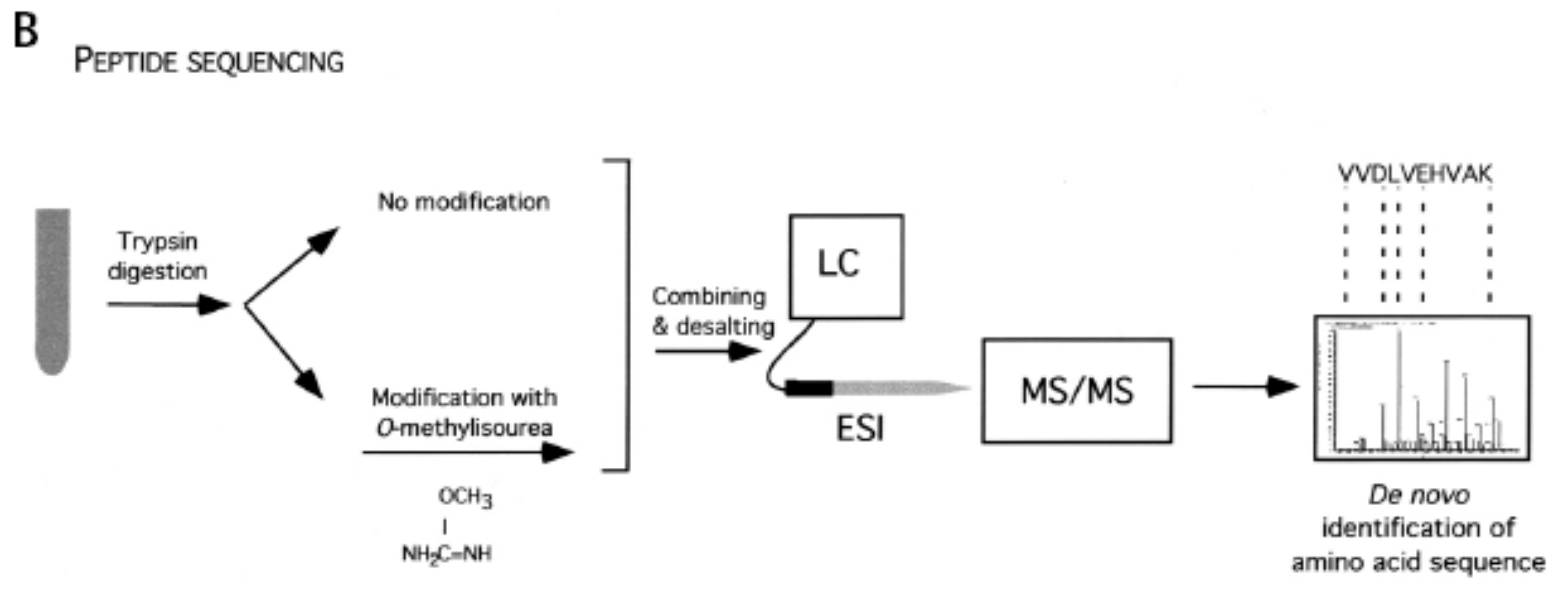
MCAT



Cagney G., Emili A.: *Nature Biotechnol* **20** (2002), 163-170

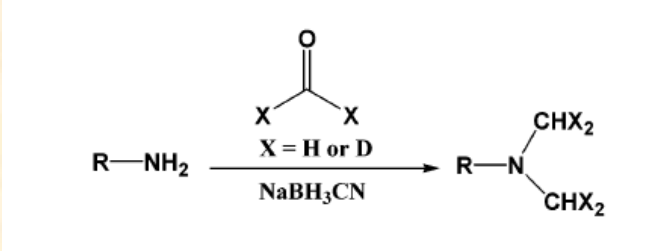
MCAT

possibility of utilization of derivatization for *de novo* sequencing
b ions unchanged , y ions in doublets (42 Da)



Reductive alkylation - dimethylation

- lysine and N-term of peptide
- isotopically labeled formaldehyde (D, ^{13}C)



Hsu et al., Anal.Chem. 2003

Reductive alkylation - dimethylation

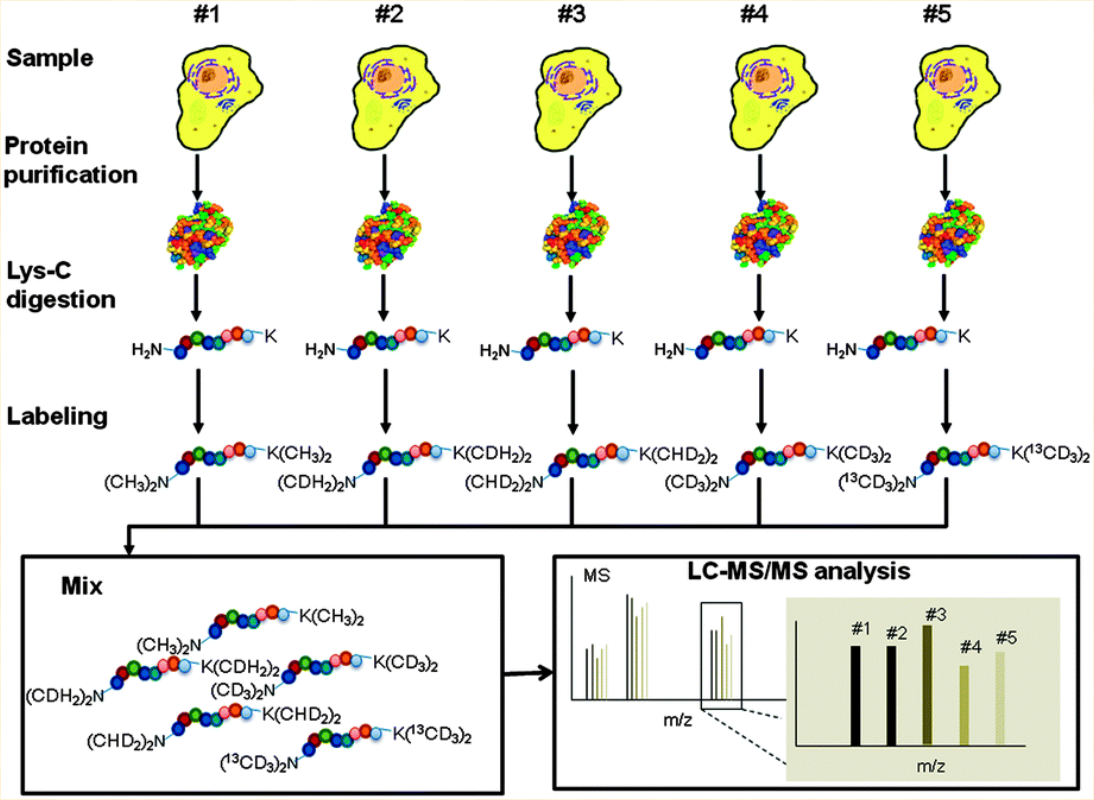


Table 1 The combination of different isotopic reagents in the five-plex isotope dimethyl labeling method

Label	#1	#2	#3	#4	#5
Formaldehyde isotope	H_2CO	H_2CO	D_2CO	D_2CO	$D_2{}^{13}CO$
Cyanoborohydride isotope	$NaBH_3CN$	$NaBD_3CN$	$NaBH_3CN$	$NaBD_3CN$	$NaBD_3CN$
ΔMass (Da, one active site)	28.0313	30.0439	32.0564	34.0690	36.0757
ΔMass (Da, two active sites)	56.0626	60.0878	64.1128	68.1380	72.1514

Wu et al., Chem. Commun. 2014

Reductive alkylation - dimethylation

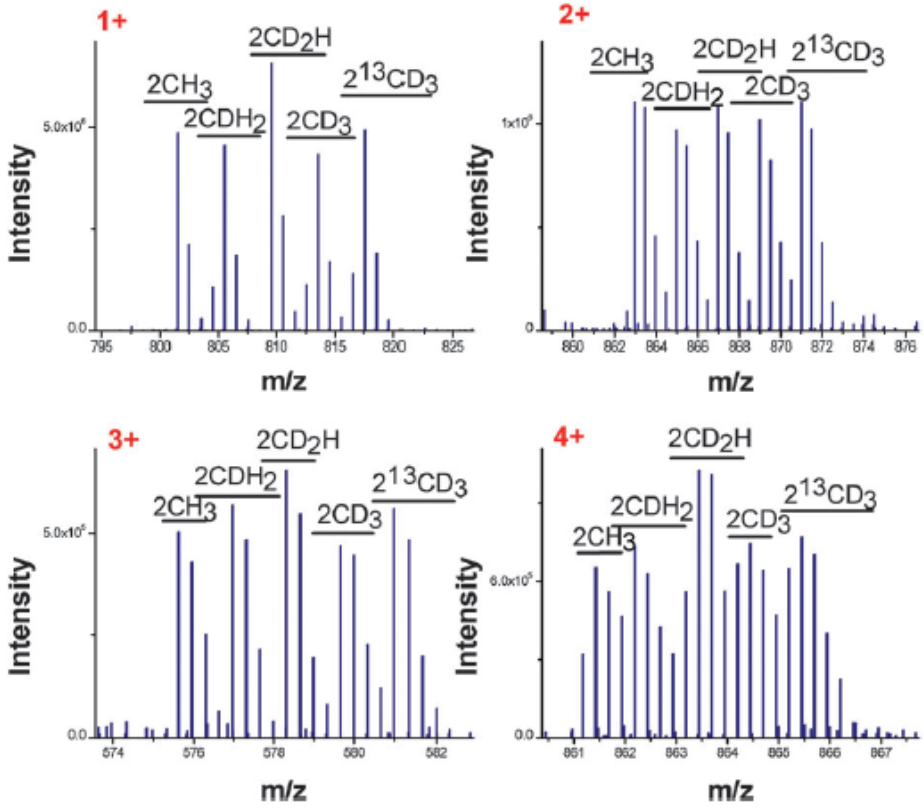
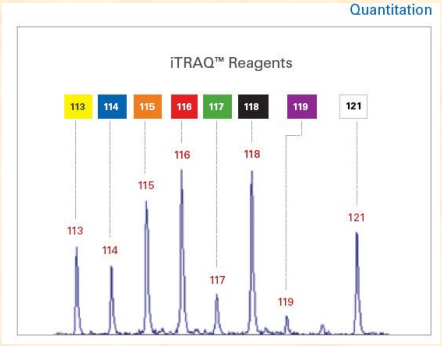


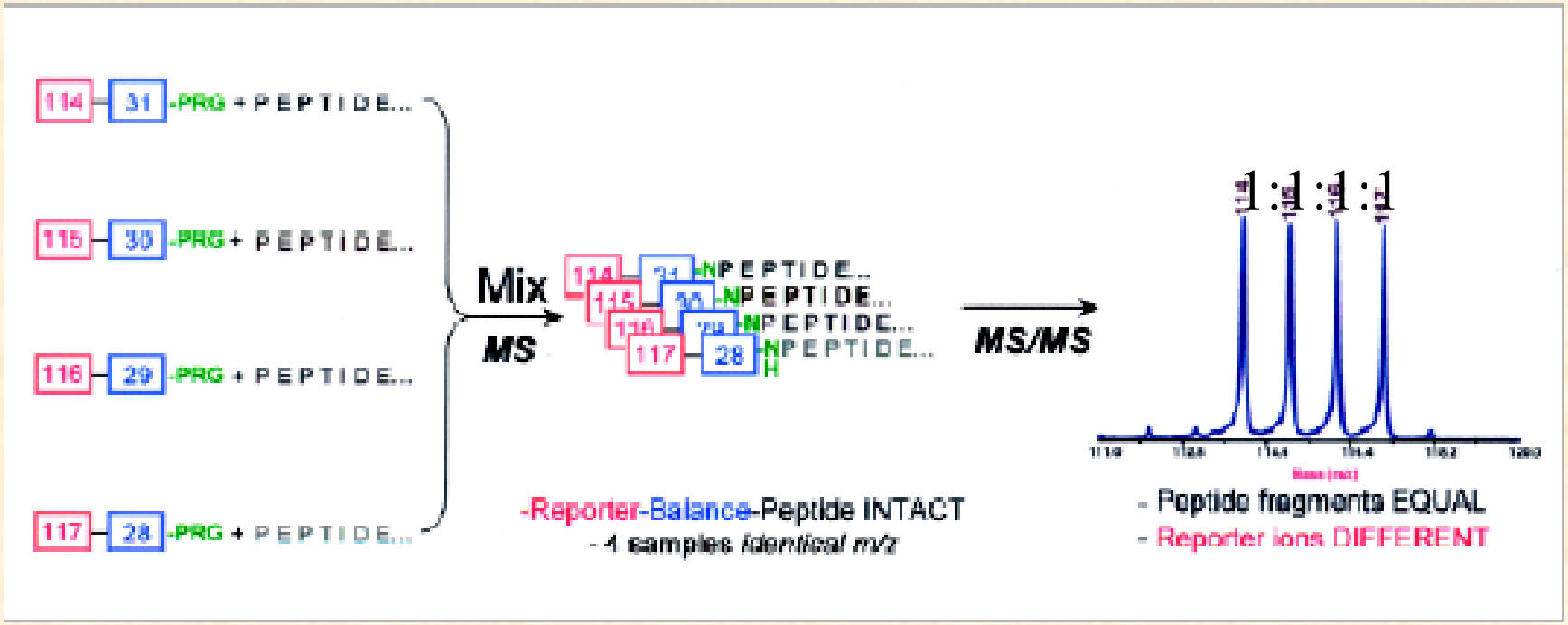
Fig. 2 Mass resolution of five-plex isotopically labeled peptides. MS spectra of Lys-C digested peptides with different charge states (+1, +2, +3 and +4) (1 : 1 : 1 : 1 : 1 ratio).

Wu et al., Chem. Commun. 2014

iTRAQ



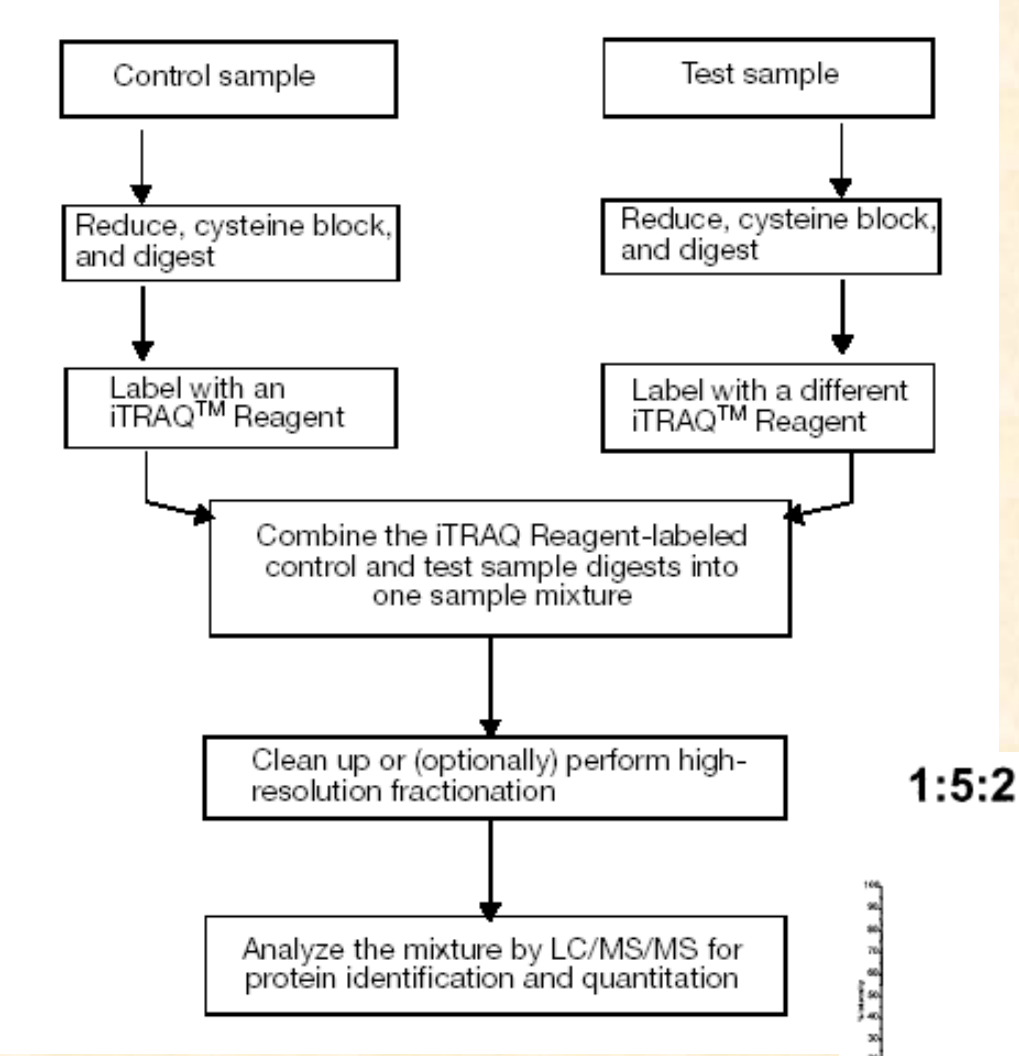
- isobaric tags (4, 8), preferentially on Lys
- labeled samples have the same behavior during LC separation and MS analysis
- quantification based on intensity ratios of reporter ions after MS/MS

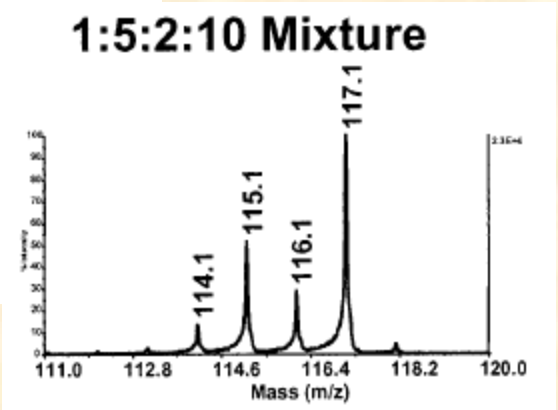


Applied Biosystems

similarly TMT tags (Thermo Fisher Scientific) – 6/10 tags (see later)

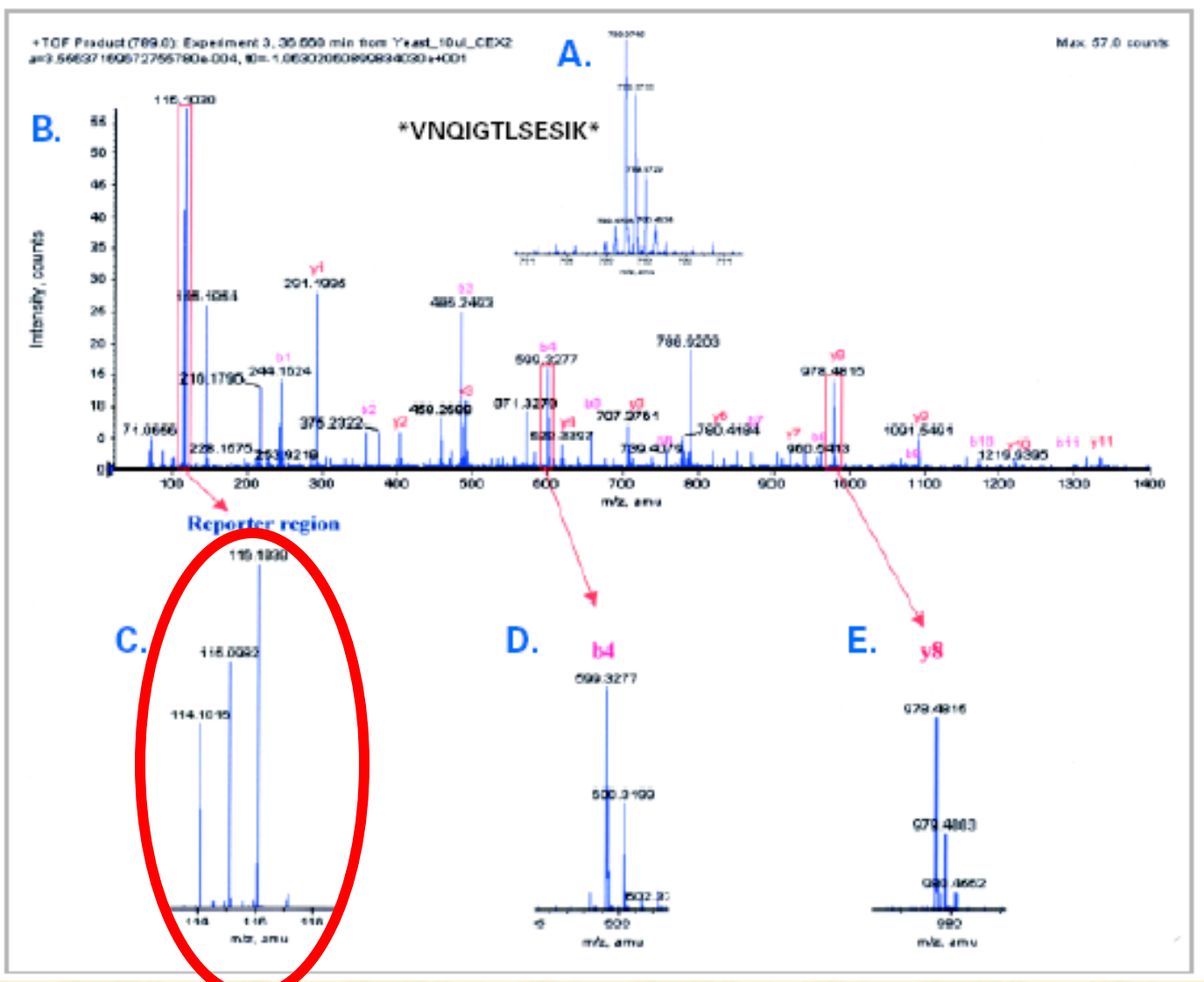
iTRAQ





Applied Biosystems

iTRAQ


A.
B.
VNQIGTLSESIK
Intensity, counts
m/z, amu
Reporter region
C.
D.
b4
E.
y8

iTRAQ

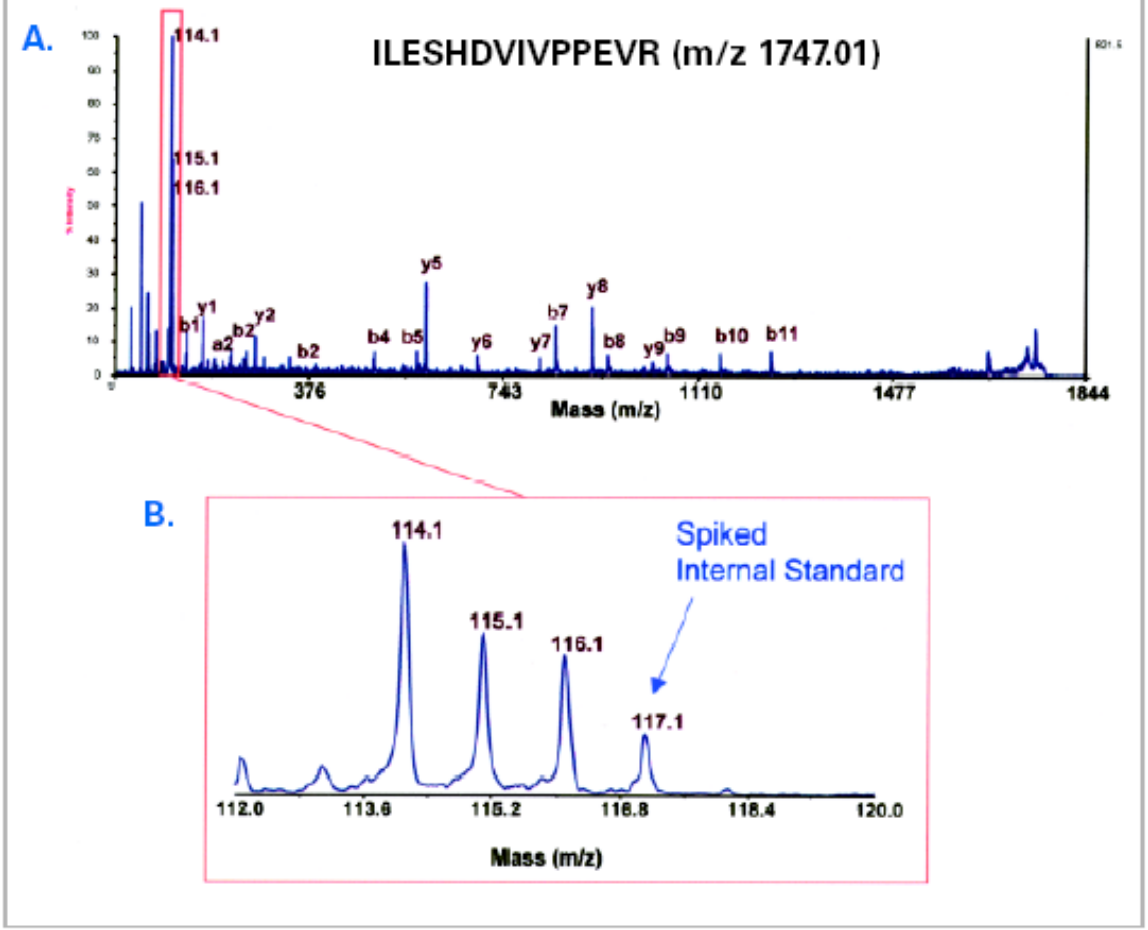
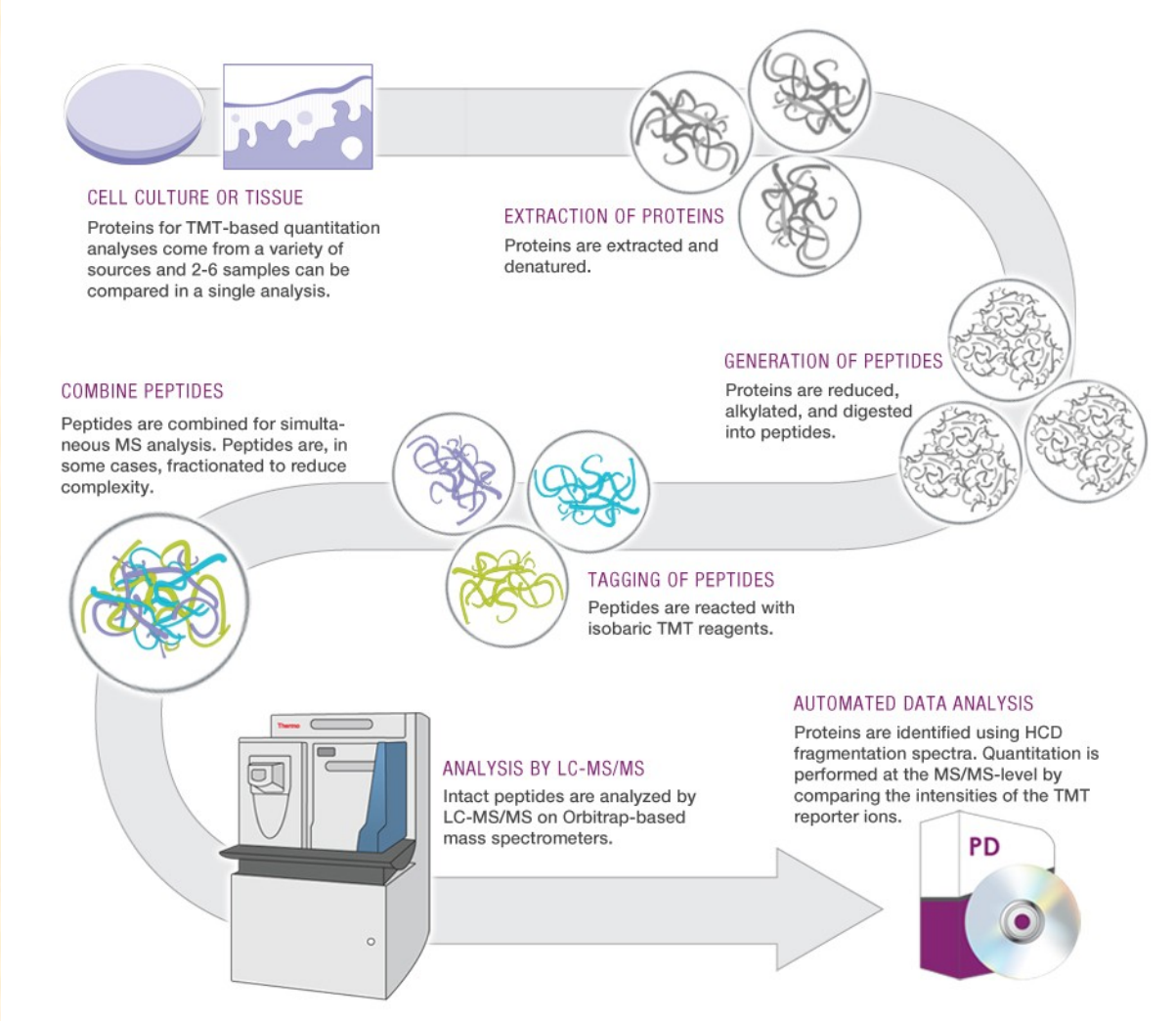


Figure 8. Identification and Quantitation of iTRAQ™ Reagent labeled peptide, ILESHDVIVPPEVR, from Carbamoyl-phosphate synthetase, which is up-regulated in both Xrn1Δ and Upf1Δ mutants. Illustrated in panel A is the TOF MS Spectrum and the reporter ion region is expanded in panel B. The diagnostic reporter ions of 114.1, 115.1, and 116.1 are those for the Xrn1Δ, Upf1Δ and wild-type *S. cerevisiae* strains, respectively. The 117.1 peak is from a specific amount of spiked-in synthetic peptide idenpently labeled with the iTRAQ Reagent 117.

TMT labels

Tandem Mass Tags

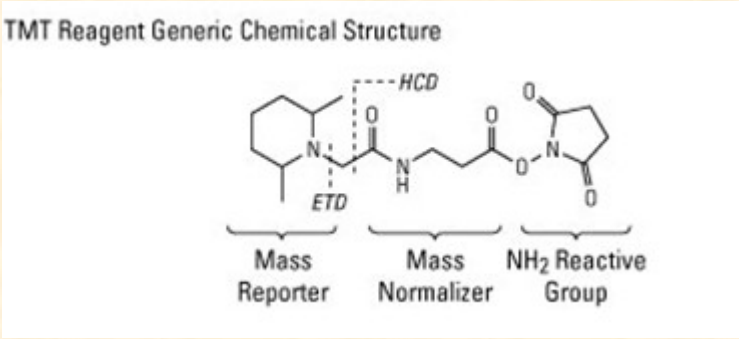


from http://planetorbitrap.com

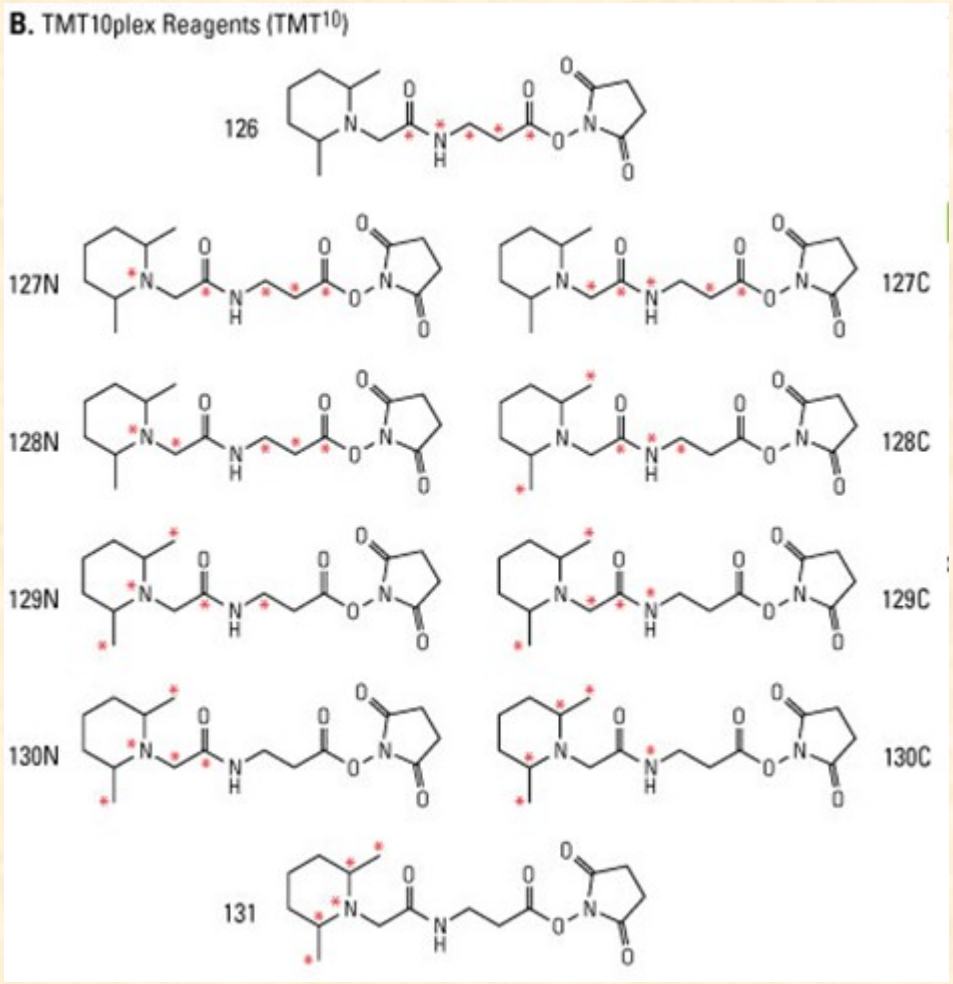
TMT

C7250

- isobaric labels (up to 16-plex)
- MS/MS



lysine labeling



- Cysteine-reactive mass tags (6-plex)

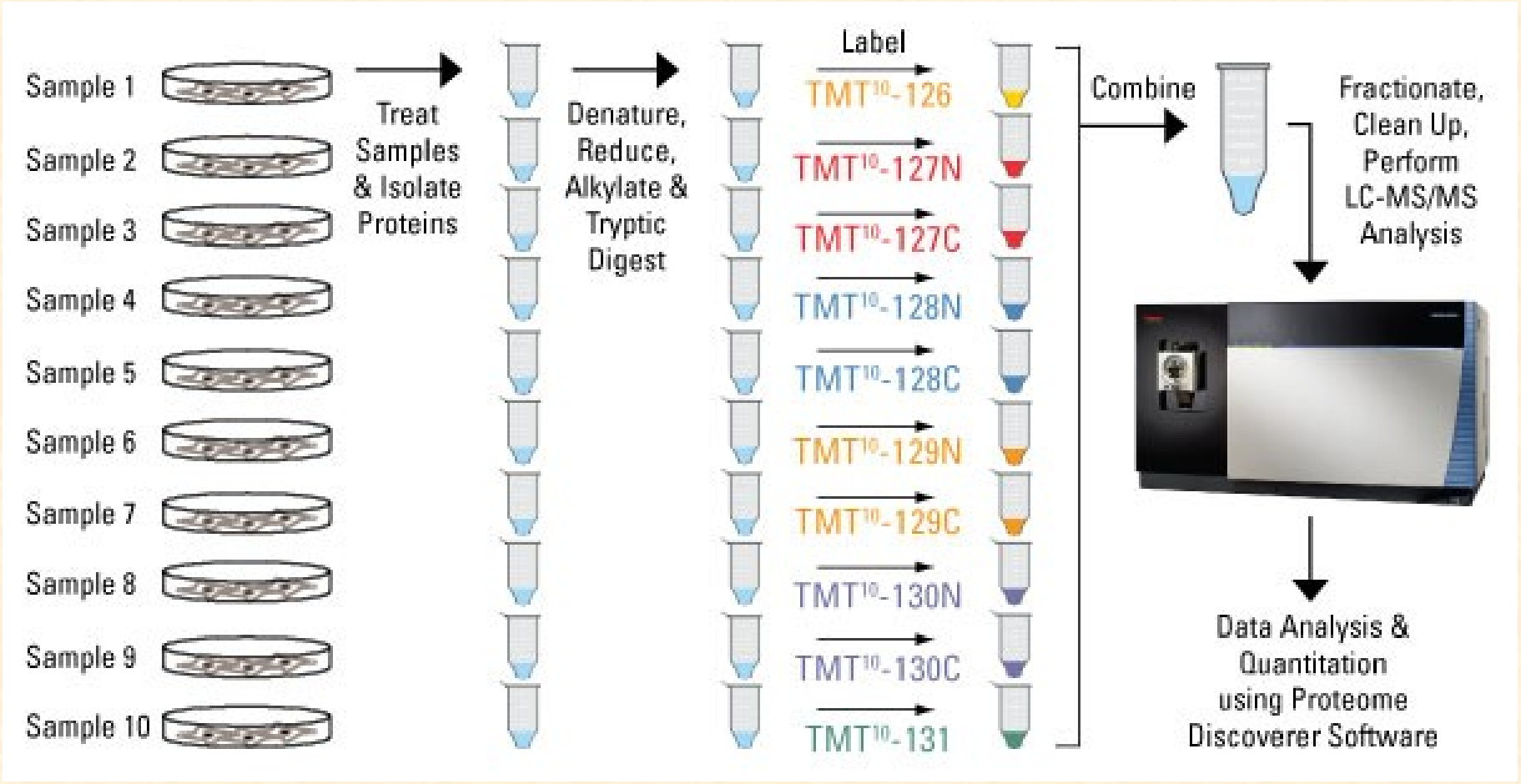
  quantitation of the relative abundance of cysteine modifications, such as S-nitrosylation, oxidation and disulfide bonds

- Carbonyl-reactive mass tags (6-plex)

  glycan, steroids, or oxidized proteins quantification

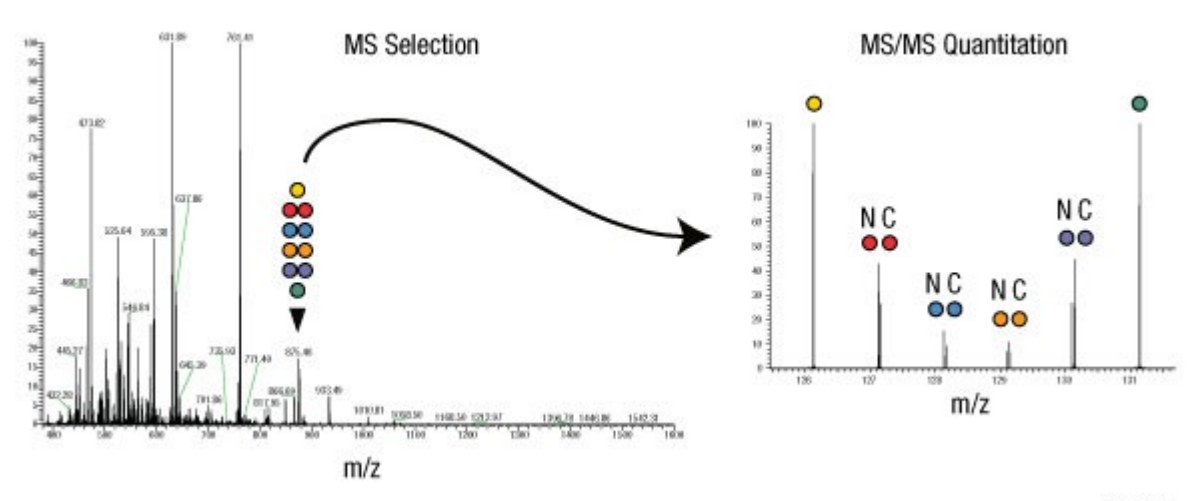
Thermo Fischer Scientific

TMT

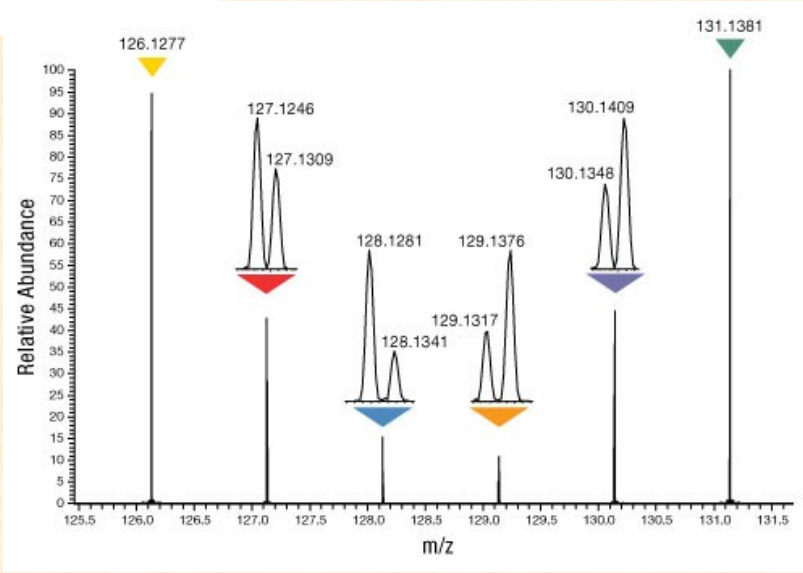


Thermo Fischer Scientific

TMT



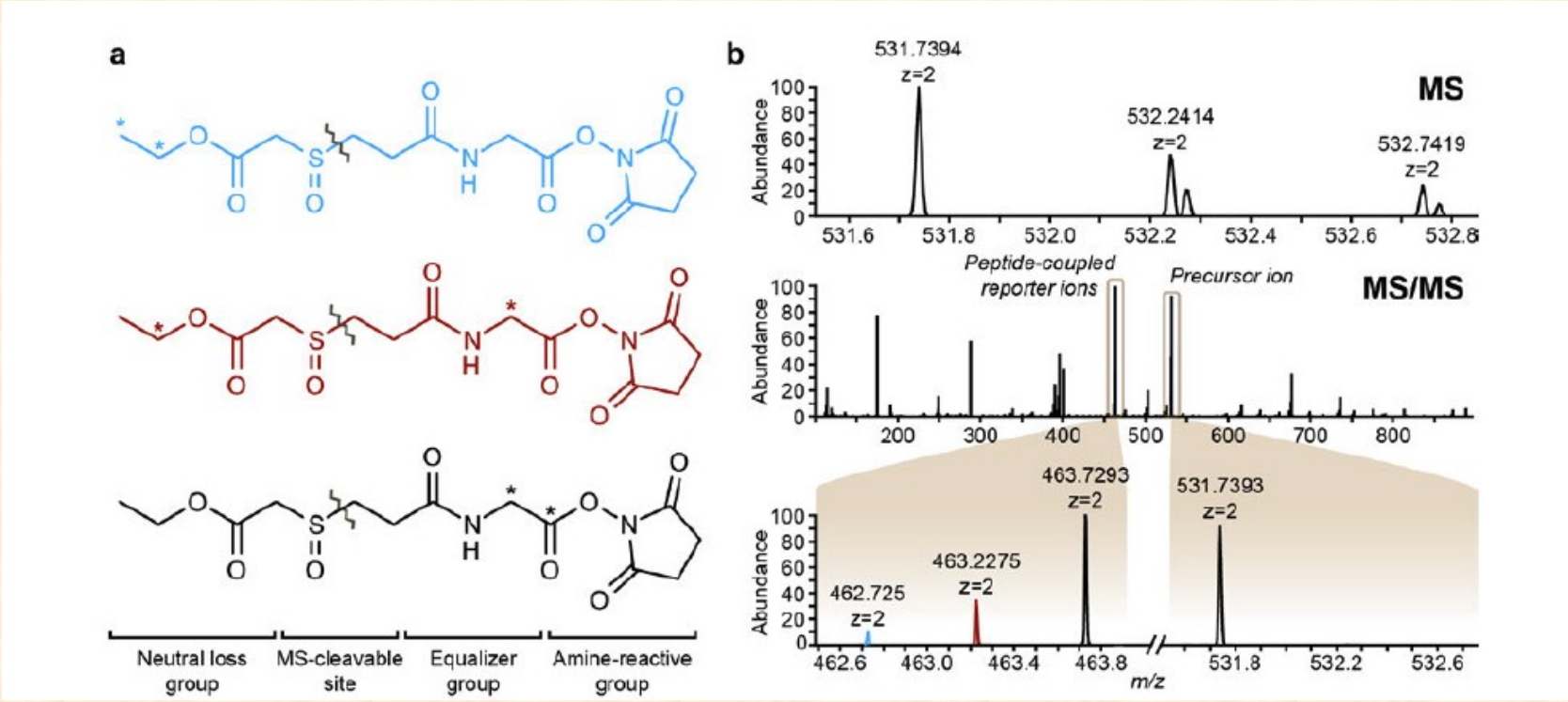
Reagents contain different numbers and combinations of 13C and 15N isotopes in the mass reporter. The different isotopes result in a 10-plex set of tags that have mass differences in the reporter that can be detected using **high resolution** Orbitrap MS instruments.



Thermo Fischer Scientific

EASI-tag

Easily Abstractable Sulfoxide-based Isobaric tag

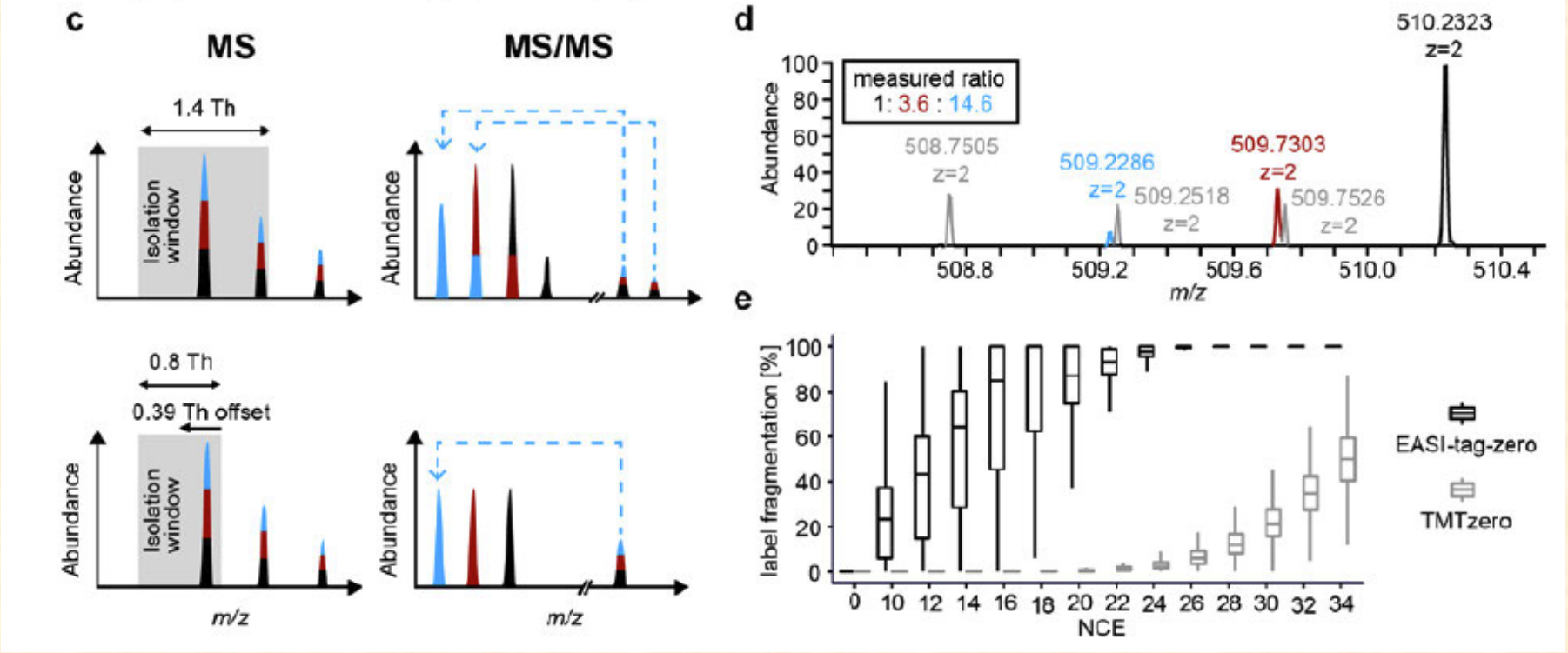


(**a**) Molecular structures of the triplex version of EASI-tag. The isobaric labeling reagents are composed in a modular way of **four functional groups** and feature a central sulfoxide moiety, which introduces an asymmetric, low-energy cleavage site (zig-zag lines indicate fragmentation site). The stable-isotope labeled positions of the neutral loss and equalizer group for multiplexing are indicated by asterisks. Standard labeling protocols can be applied to couple peptides via **the amine-reactive moiety**.
(**b**) Mass spectra of an EASI-tag-labeled yeast peptide mixed in a ratio of 1:3:10. HCD fragmentation of the doubly charged precursor ion abstracts the neutral loss group and yields **the peptide-coupled reporter ion cluster**.

bioRxiv preprint first posted online Nov. 27, 2017; doi: http://dx.doi.org/10.1101/225649

EASI-tag

Easily Abstractable Sulfoxide-based Isobaric tag



(**c**) Co-isolation of the natural isotope cluster in a standard isolation window centered on the precursor ion (upper panel) convolutes the relative abundance of peptide coupled reporter ions. An asymmetric isolation window (lower panel) that suppresses the signal from adjacent isotope peaks and enables direct quantification of reporter ions.
(**d**) The precursor mass information is retained in the peptide-coupled reporter ions for EASI-tag labeled peptides. Colored peaks indicate the peptide-coupled reporter ions from an identified yeast peptide in a two proteome experiment (mixing ratios: 1:3:10 for yeast & 1:1:1 for human). Grey peaks are peptide-coupled reporter ions from a co-isolated peptide.
(**e**) EASI-tag- and TMT-labeled HeLa peptides were fragmented with normalized collision energies between 10 and 34. (N = 17,565 precursors for EASI-tag & 20,610 for TMT)

bioRxiv preprint first posted online Nov. 27, 2017; doi: http://dx.doi.org/10.1101/225649

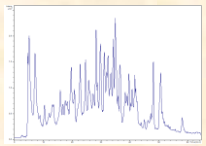
Label – free approach

- No labels
- Samples measured individually
- Comparison of „unlimited“ number of samples

Identification based on MS/MS data, connection of identity to individual signals (intensity, area)

LC-MS/MS

→



noise removal
peak detection
intensity/area calculation
for individual peaks

→

normalization
imputation of missing values

→

calculation of fold changes for individual proteins among samples

↓

Statistics evaluation

↓

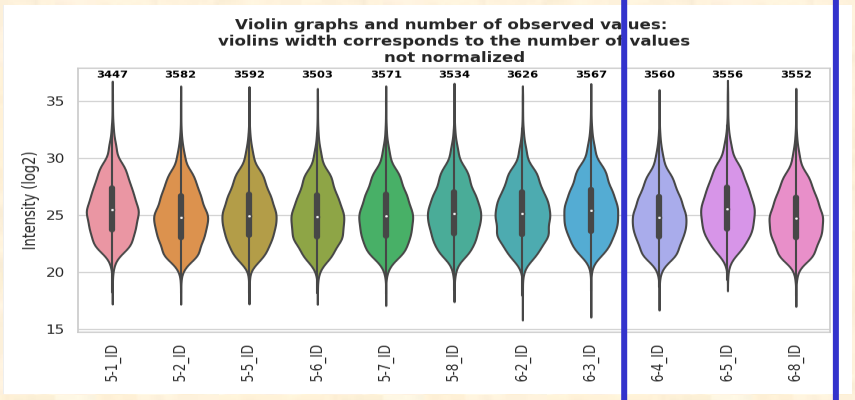
significant changes in proteome induced by stimuli under study

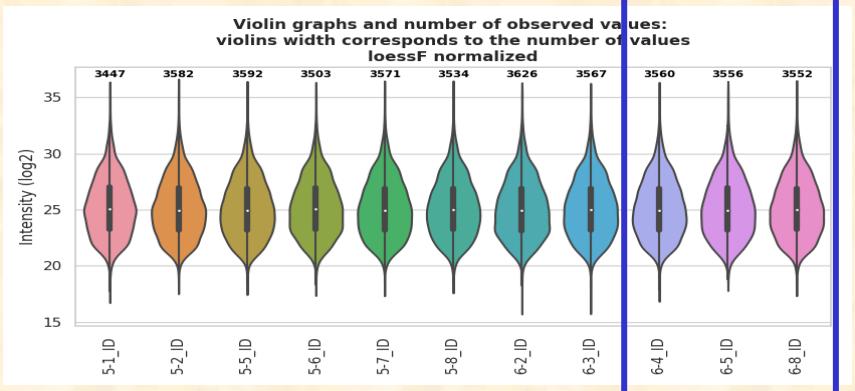
Critical steps:

- data normalization
- imputation of missing values (DIA – reduction of number of missing values)

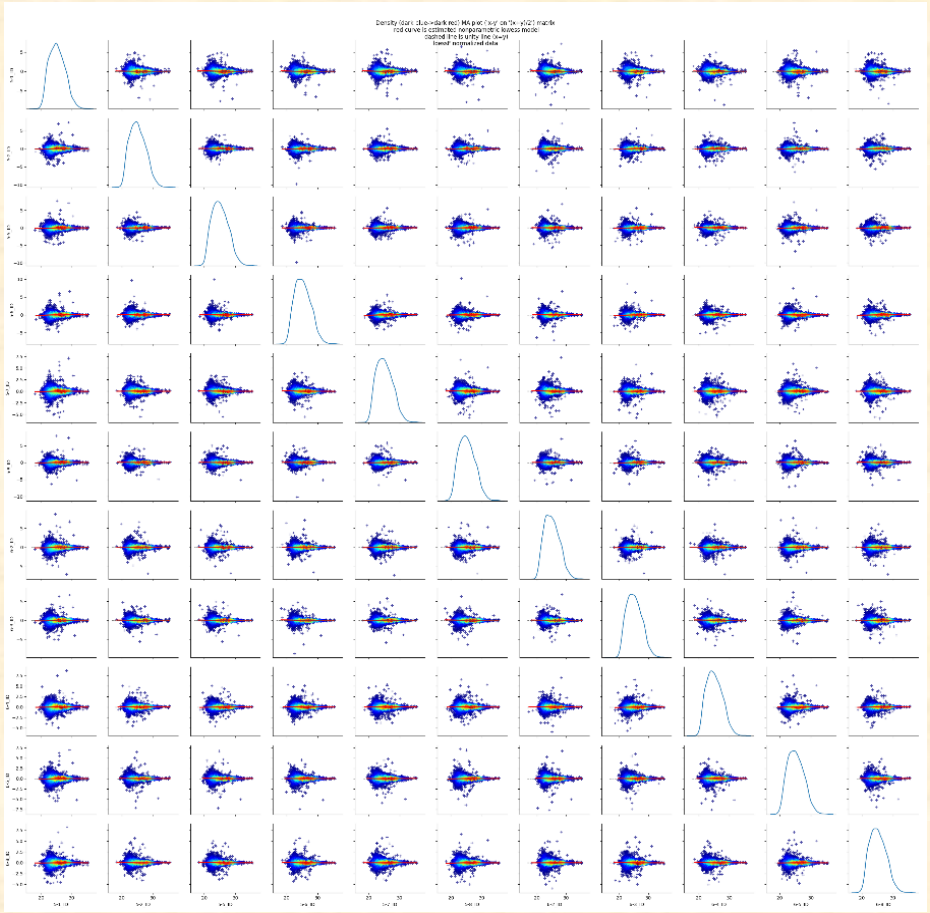
C7250

Label – free approach

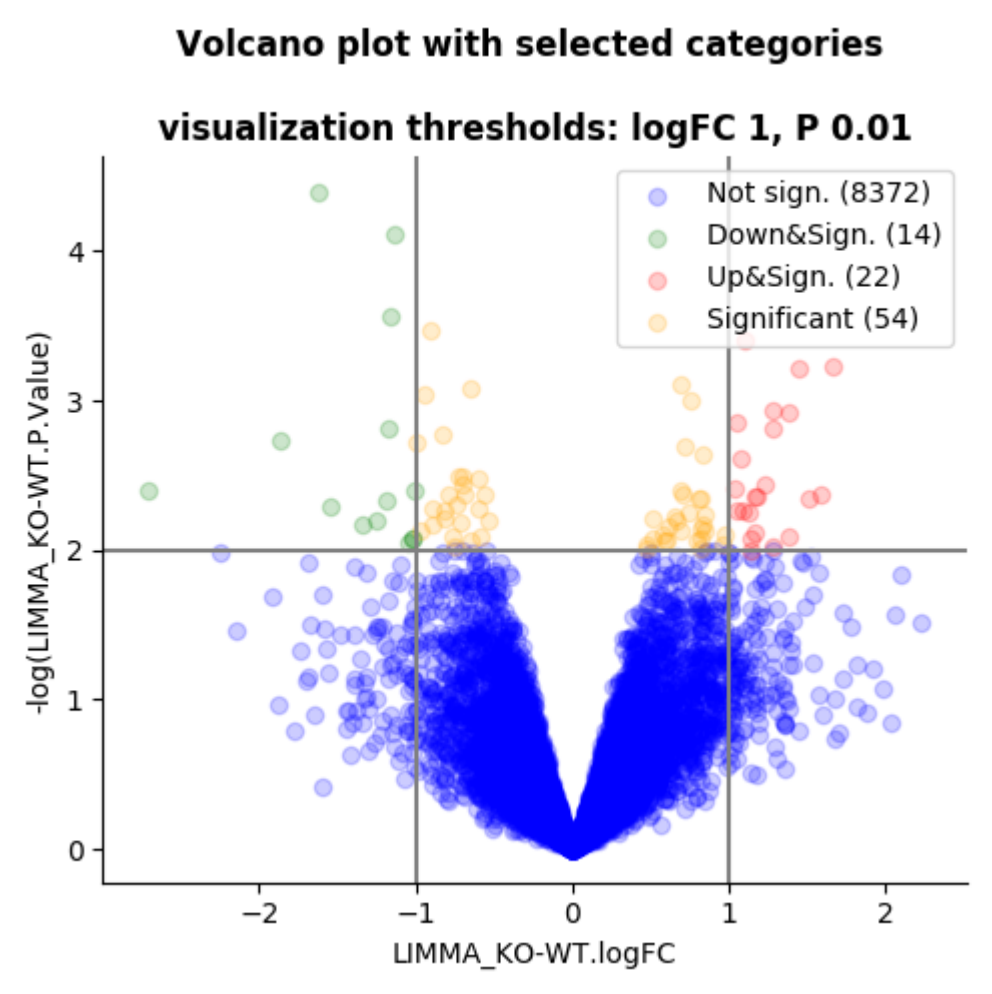




MA plots - dependence of (x-y) on (x+y)/2



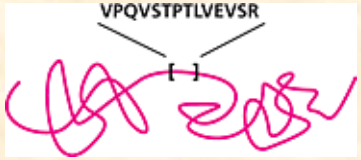
Label – free approach



Absolute quantification using AQUA peptides

- **AQUA Peptide Selection**

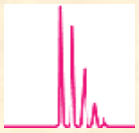
Select an optimal tryptic peptide and stable isotope amino acid from the sequence of your protein of interest



- **Order selected peptide labeled (^{15}N, ^{13}C)**

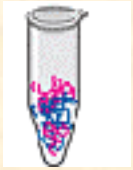
Price !!!

Optimize LC-MS/MS separation protocol for quantitation

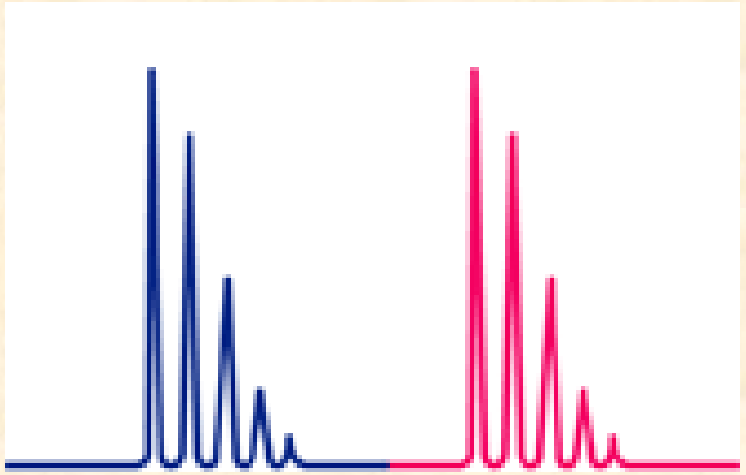


- **Adding labeled peptide to protein mix**

- **Digest**



- **Analyze by LC-MS/MS to quantitate protein of interest**



Only for selected protein

Absolute quantification

AQUA peptides

synthesis of isotopically labeled proteotypic peptides

addition of known amount into sample

digestion

MS analysis

QconCAT

construction of artificial gene for expression of proteotypic peptides originated from up to 20 proteins

expression in *E.coli* using labeled medium

purification of artificial protein

addition of known amount into sample

digestion

MS analysis

Rivers at al., MCP 6, 1416 (2007)

PSAQ

expression of the whole izotopically labeled protein (plus tag for purification)

protein purification

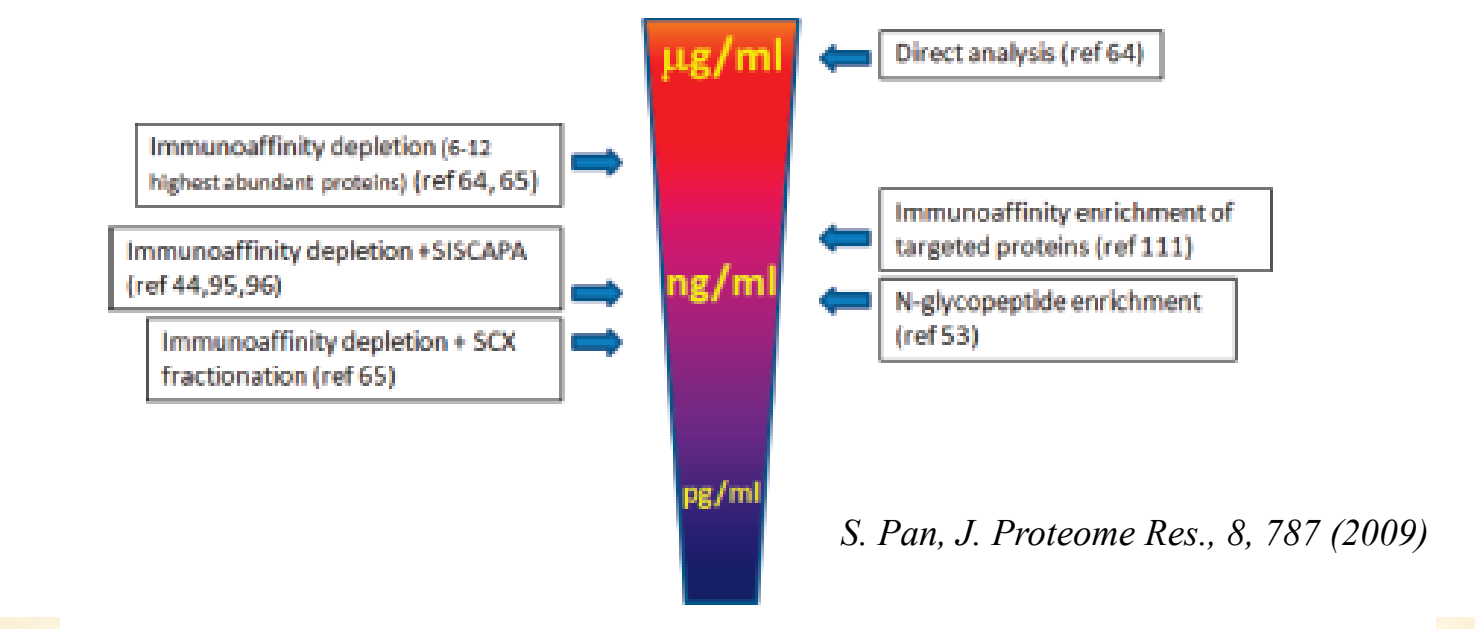
addition of known amount into sample

digestion

MS analysis

Brun et al., MCP 6, 2139 (2007)

Absolute quantification
Stable Isotope Standards and Capture by Anti-Peptide Antibodies (SISCAPA)



S. Pan, J. Proteome Res., 8, 787 (2009)

Table 1. Plasma Protein Targets

protein identity	Swiss-Prot accession no.	protein mass (daltons)	normal range concentration in human serum
interleukin-6 (IL-6)	P05231	23 718	<10 pg/mL
hemopexin (hx)	P02790	51 676	0.5−1.15 mg/mL
α_1−antichymotrypsin (AAC)	P01011	47 650	0.3−0.6 mg/mL
tumor necrosis factor-α (TNF-α)	P01375	25 644	<10 pg/mL

N.L. Anderson, J. Proteome Res., 3, 235 (2004)

Targeted MS/MS analysis of selected proteins

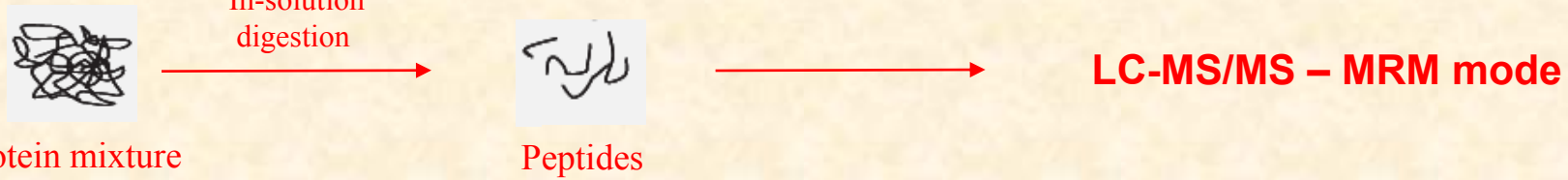
relative /absolute quantification

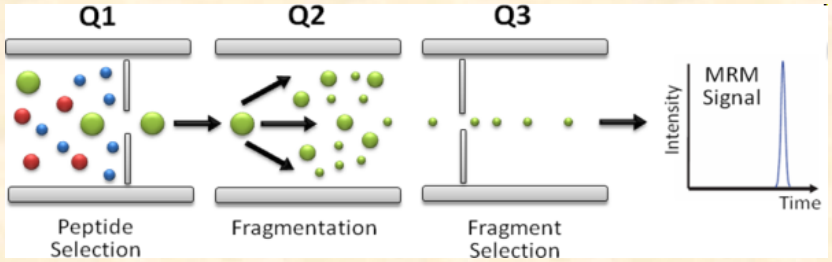
multiple reaction monitoring (MRM)

screening – selection of candidate protein

method establishment (selection of MRM transitions – peptide + selected fragment)

final analysis and data processing



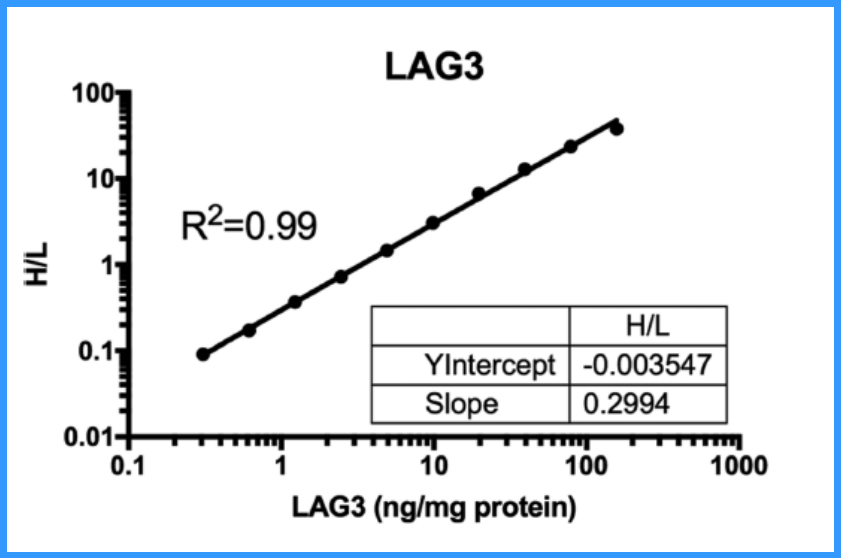


Similarly PRM

http://www.srmatlas.org

Multiplex Immuno-Liquid Chromatography–Mass Spectrometry–Parallel Reaction Monitoring (LC–MS–PRM)

Quantitation of immune markers CD8A, CD4, LAG3, PD1, PD-L1, and PD-L2 in Frozen Human Tissues



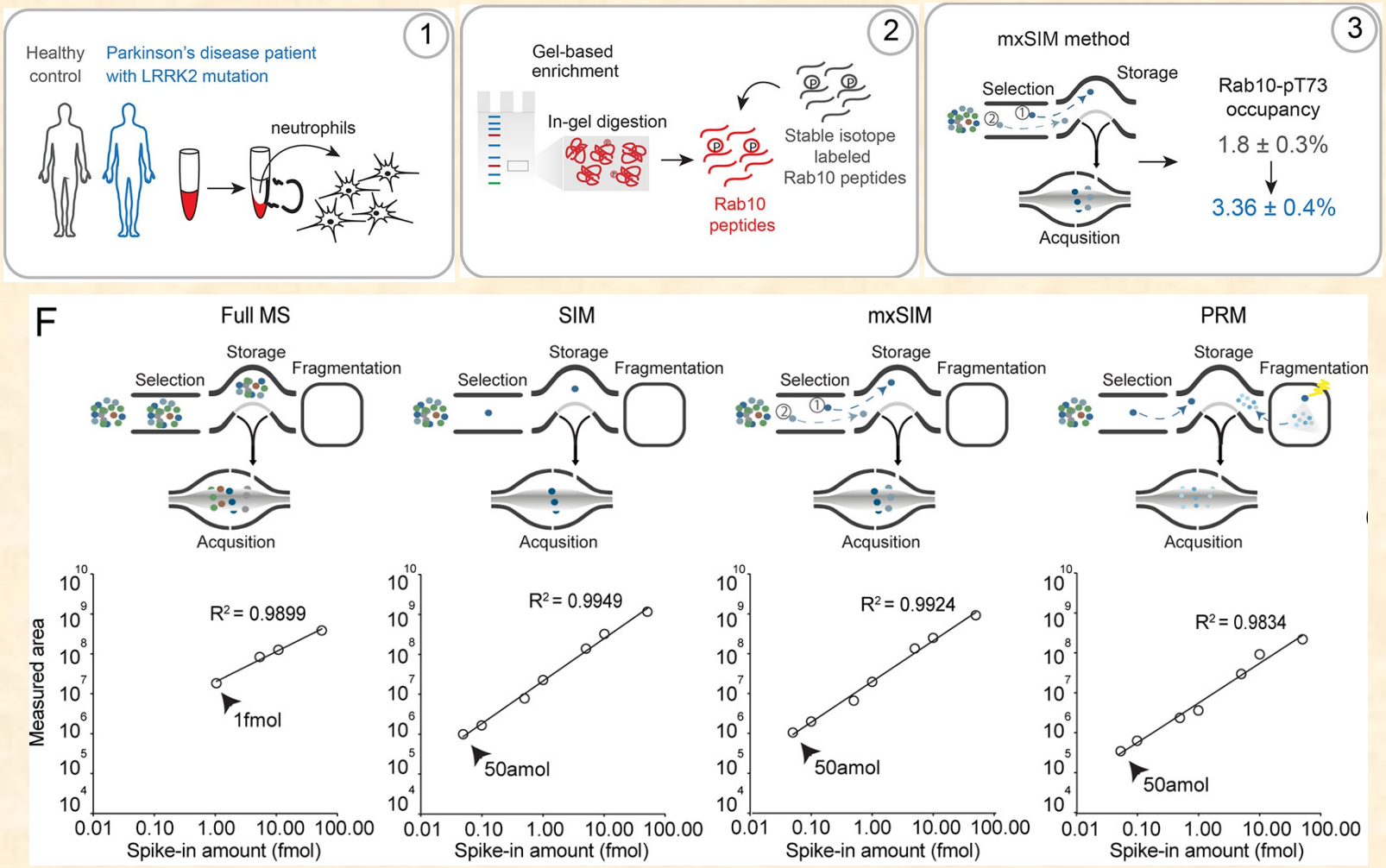
protein	MW (kDa)	monitored peptide	parent m/z	fragment m/z	HCD energy	LLOQ (ng/mg total protein)
LAG3	57.4	FVWSSLDTPSQR	711.85	990.48 (y9), 588.31 (y5)	23	0.15

Zhang et al., J. Proteome Res. 2018, 17, 3932−3940

Accurate MS-based Rab10 Phosphorylation Stoichiometry

Karayel et al., Mol Cell Proteomics (2020) 19(9) 1546–1560

assay to measure increased phospho Rab levels using synthetic stable isotope-labeled analogues for both phosphorylated and non-phosphorylated tryptic peptides surrounding Rab10-Thr73



Limit of detection (LOD) of SIL Rab10-pThr73 tryptic peptide (FHpTITTSYYR) with various acquisition methods; full MS, SIM, mxSIM and PRM.

Absolute Quantification of over 1800 Yeast Proteins

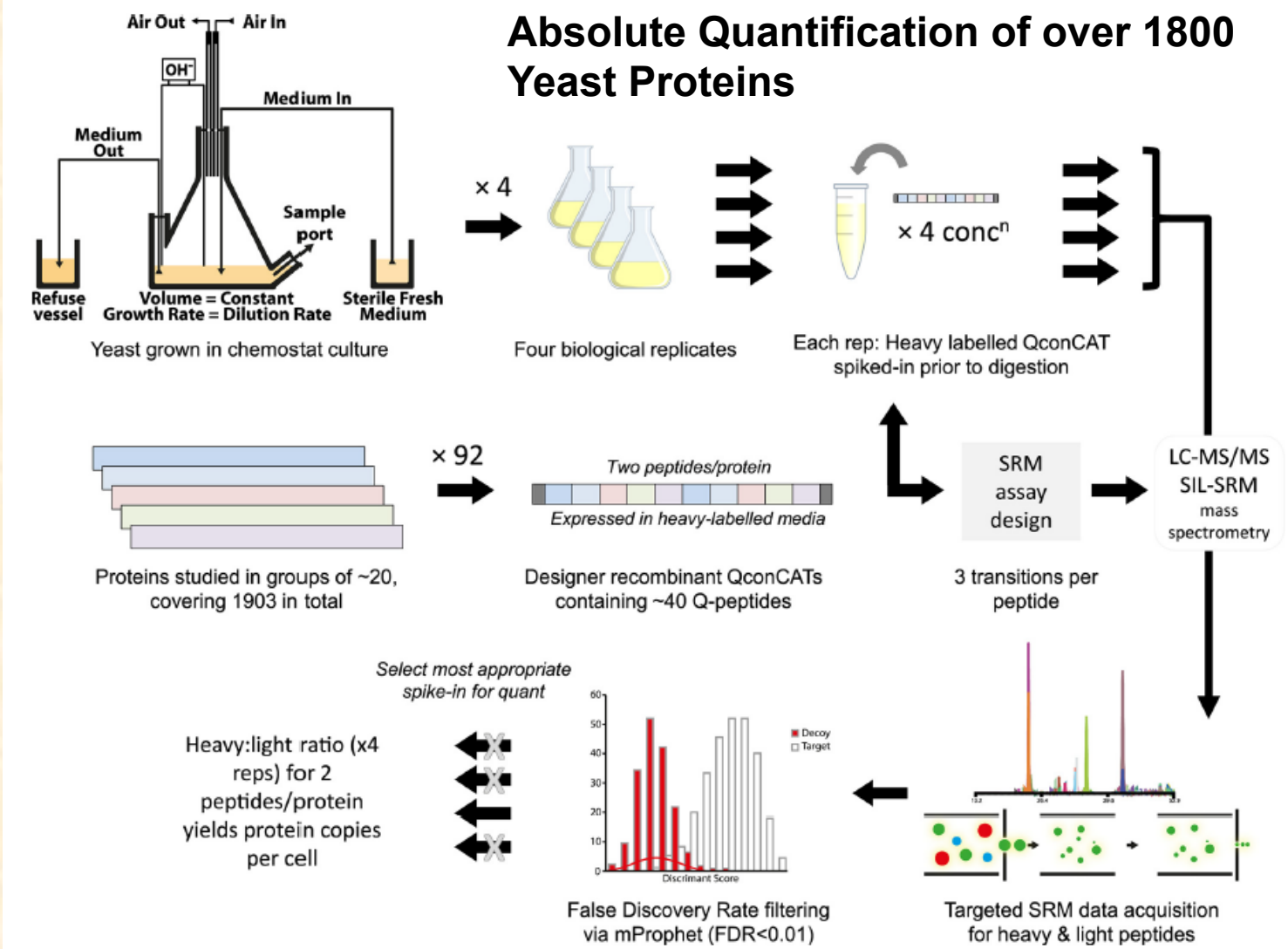


FIG. 1. **Schematic overview of QconCAT-based quantification of the yeast proteome using SIL-SRM methodology.** The experimental workflow is depicted in schematic form, showing how chemostat grown yeast samples are extracted, using four biological replicates, for analysis. These samples were combined with designer QconCAT proteins, containing surrogate quantotypic peptides, expressed in a stable-isotope labeled media. SRM assays, designed using a digest of the expressed QconCATs to generate Q-peptides, were then used to quantify the parent proteins. Mixtures of purified QconCAT and yeast proteins were mixed at four concentrations (one of which contained yeast but no QconCAT) and analyzed by SRM-MS to yield SRM chromatogram peak groups for both light (endogenous yeast) and heavy (Q-) peptides. Subsequent quality control by signal:noise cutoffs and mProphet FDR (estimated from decoy transitions) yielded peptide-level copies per cell values, which were then integrated to the protein level for a final quantification.

Lawless C., MCP,15, 1309–1322, 2016.


The end