

**Functional Genomics and Proteomics
National Centre for Biomolecular Research**

National Centre for Biomolecular Research**Faculty of Science · Masaryk University**

CEITEC

Protein characterization by mass spectrometry

C7250

Part I

Zbyněk Zdráhal

RG Proteomics, CEITEC-MU Proteomics CF, CEITEC-MUNCBR FS MUzdrahal@sci.muni.cz

Zbyněk Zdráhal Phone: +420 777 926 602 E-mail: zbynek.zdrahal@ceitec.muni.cz

CF: www.ceitec.eu/proteomics-core-facility/cf95RG: www.ceitec.cz/proteomika-zbynek-zdrahal/rg49

> **CIISB** Czech Infrastructure for Integrative Structural Biology

Aims of course:

- applications/potential of mass spectrometry in proteomics¢
- basic approaches of MS analysis۰
- , interpretation/validity" of MS results \bullet

Proteins are responsible for both the structure and the functions of all living organisms.

Genes are simply the instructions for making proteins.

IT IS PROTEINS THAT MAKE LIFE.

http://www.biology-pages.info/P/Proteomics.html

Proteomics - Why?

- several proteins/proteoforms might form from each gene, not possible to \bullet indicate them by DNA/RNA analysis
- there no direct correlation between mRNA content and final content of proteins
- functionality of protein depends frequently on its interaction with other \bullet proteins or DNA/RNA
- only at protein level epigenetics factors of gene expression regulation are \bullet detectable

The Desperate Man, Gustave Courbet **Prote**

the word "proteome" is derived from PROTEins expressed by a genOME, and it refers to all the proteins produced by an organism

Marc Wilkins in 1994

C7250

the complement of proteins expressed in a cell, tissue, or organism by a genome

the entire complement of proteins found in an organism over its entire life cycle, or in a particular cell type at a particular time under defined environmental conditions.

The entire set of proteins expressed by a genome, cell, tissue or organism at a certain time. More specifically, it is the set of expressed proteins in a given type of cell or organism, at a

given time, under defined conditions.

proteotype

Genome vs Proteome

Genome vs Traskriptome vs Proteome

Estimates for human

http://www.piercenet.com/method/overview-post-translational-modification

C7250

Proteome complexity

Estimates from deep proteomics and transcript profiling suggest that about half the human genome is expressed in proteins at over 20 copies per cell in a given cell type (i.e., about 10,000 of the 20,000 human genes)²⁸. Assuming this expression threshold of 10,000 genes and allowing for detection of ~100 proteoforms for each gene product, one then multiplies these two to arrive at a measurement target of 1,000,000 distinct proteoforms in a given cell type. A 2016 estimate based on trends in databases indicated that the number may be \sim 6 million proteoforms²⁹. Better estimates of this proteoform diversity are needed, and are analogous to the extrapolations of the number of human genes using expressed sequence tags (ESTs) in the year 2000 (ref. 30). Aebersold R. et al, Nat. Chem. Biol., 14, 206-214 (2018)

74 unique histone H4 proteoforms out of still not developed at ally
possible (based on known prechnologies are still not differentiating
here are stingly of technologies are still differentiating human

nature behaves rationally 7115.

Phanstiel D. et al, PNAS, 105, 4093–4098 (2008)

Proteomics

Proteomics is the large-scale study of proteins, particularly their structures and functions.

The first is the more classical definition, restricting the large-scale analysis of gene products to studies involving only proteins.

The second and more inclusive definition combines protein studies with analyses that have a genetic readout such as mRNA analysis, genomics, and the yeast two-hybrid analysis (Pandey A, Mann M Nature. 2000 Jun 15; 405(6788):837-46)

Proteomics has been enabled by the accumulation of :

- \triangleright DNA and protein sequence databases
- \triangleright improvements in mass spectrometry
- \triangleright computer algorithms for database searching.

Increase in knowledge of genomes

protein characterization by MS is in principle based on knowledge of primary sequence

Single cell proteomics

Mass spectrometry technology progress

 \Box +MBR \Box -MBR

 Ω

COO

Cell count

HeLa cells, FACS, LC-MS/MS (TimsTOF Pro)

Brunner A.-D. et al. bioRxiv (2021), https://doi.org/10.1101/2020.12.22.423933

mass spectrometry

enables simultaneous qualitative and quantitative characterization of thousands of proteins Analysis of "proteome"

Cwww.ClipProject.info

Mass spectrometry in proteomics

Intact mass analysis \overline{O} (MW, MALDI-MS profiling)

Protein identification \overline{O}

(incl. protein complexes, de novo sequencing)

Characterization of protein modifications \overline{O}

Protein quantification \overline{O} (relative and absolute quantification)

- MALDI-MS imaging,
- •3D structure

ومبعضا والمعامل والمتحادث والمستعاري والمتعارض والمتحارب والمتحارب والمستعار والمتحدث والمستعمل والمتحارب والمتحارب

Proteomic approaches

- Differential (expression) proteomics
- Functional proteomics
- Structural proteomics

Differential (expression) proteomics

Qualitative and quantitative comparison of proteomes

aim – determination of changes at protein (and their forms (e.g. PTMs)) levels which were induced by internal or extermal stimuli.

stress

control

Rhodotorula ^glutinis

Cooperation with prof. Márová, FCH BUT Brno

Differential proteomicsTargeted approach

Monitoring of quantitative changes of selected proteins (e.g. biomarkers) in sample sets.

Determination of enterotoxin *(S. aureus)*

LC-MS/MS (MRM, PRM)

Functional proteomics

Study of interactions of proteins and their functional context

- -- protein-protein interactions
- architecture of protein complexes
- protein interactions with other types of molecules (RNA, DNA metabolites ...)

Cell Cycle, 14 (6), 920-930 (2015)

A Drosophila Protein Interaction Map

K.G. Guruharsha et al., Cell, 147, 690–703 (2011)

Structural proteomics

Study of higher levels of protein structure (tertiary, quaternary) and relation of ^a structure to protein function.

- •• X-ray crystalography
- •NMR
- cryoEM •
- •• MS (in minority)

Structure is formed by different types of bonds – ion interactions, hydrogen bridges, van der Waals forces or disulfidic bridges.

F**unctional Genomics and Proteomics**
National Centre for Biomolecular Research

National Centre for Biomolecular ResearchFaculty of Science Masaryk University

MS instrumentationandproteins

History of mass spectrometry

J.J. Thomson working with his cathode ray tube

Thomson J.J. (1856 – 1940) On the Masses of the Ions in Gases at Low PressuresPhilosophical Magazine, 1899, 48:295, p.547-567

¹⁹⁰⁶ – Nobel prize for physics for theoretical and experimental investigations on the conduction of electricity by gases

Eight of his students and his son also became Nobel Prize winners

Thomson's investigations into the action of electrostatic and magnetic fields on the nature of so called "anode rays" or "canal rays" would eventually result in the invention of the mass spectrometer (then called a *parabola spectrograph*) by Fr<mark>ancis Aston</mark> (Nobel Prize for Chemistry 1922), a tool which allows the determination of the mass-to-charge ratio of ions and which has since become an ubiquitous research tool in Chemistry.

... By this means there is attained what is known as a mass spectrogram, that is to say a series of lines in which each line corresponds to a certain atomic weight."

Dr. H.G. Soderhaum - ¹⁹²²

C7250

History of mass spectrometry

The first Czech mass spectrometer - V. Čermák, V. Hanuš, Č. Jech, J. Cabicar ¹⁹⁵³

by courtesy of dr. M. Polášek

Mass spectrometry- today

Benchtop LC-MS/MS systemion trap LTQ Velos (Thermo)

C7250

High resolution hybrid mass spectrometer Orbitrap Fusion™ Lumos™ Tribrid™ (Thermo)

Mass spectrometry

principle:

measurement of ratio of relative molecule mass and charge number (m/z) of ions of analyzed compounds

 $m - i$ on mass

C7250

```
basic steps:
```
- ionization of molecules of analyzed compounds
- \bullet ion separation according to their m/z
- *<u>ion</u>* detection

Why vacuum in MS

to prevent ions from unwanted collisions during their way from ion source to detector

The mean free path of a molecule

the average distance the molecule travels between two consecutive collisions with other moving particles

example: N_2 molecules

C7250

adopted from presentation of dr. M. Polášek

Functional Genomics and Proteomics

National Centre for Biomolecular Research Faculty of Science Masaryk University

Ionization techniques

-

 \circ

-

-

 $\mathcal{L}_{\mathcal{A}}$

New "soft" ionization techniques - ESI and MALDI basic prerequisite for wide use of MS in biomolecule analysis(Nobel prize 2002)

MALDI matrix-assisted laser desorption/ionization most often in combination with time-of-flight mass analyzer - TOF(MALDI-MS, MS/MS)

ESI

electrospray ionization

 usually in combination with ion trap and hybrid mass spectrometers(IT, QQQ, QTOF, IT-Orbitrap, IT-ICR etc.) $\mathcal{L}^{\text{max}}_{\text{max}}$

MALDIDesorption-ionization process

Sample embedded inlight-absorbing matrix LASER-excitation ofmatrix molecules

Sample desorption andprotonation

C7250

matrix is low mass compound capable to absorb laser radiation e.g. Dihydroxybenzoic acid (for UV laser)

 \triangleright Soft ionization without fragmentation

- \triangleright Simple spectra
- Sample storage on sample target

pictures by courtesy of Dr. Sauerland (Bruker)

Sample target for MALDI-MS Ultraflextreme (Bruker)

Sample spot with DHB matrix

MALDI – ionization

MALDI-TOF spectrum of myoglobin (16 951 Da)

Salt content influence on MALDI ionization

Electrospray Ionization

M. Wilm, Principles of Electrospray Ionization, MCP, 2011; 10 (7)

ESI-z spray

ESI-ionization

(orthogonal geometry)

ESI spectrum of myoglobin (16 951 Da)

F**unctional Genomics and Proteomics**
National Centre for Biomolecular Research

National Centre for Biomolecular ResearchFaculty of Science Masaryk University

Mass analyzers

MS and MS/MS

m/z of the whole ion

m/z of the fragments derived from the ion

Quadrupole analyzer (Q)

- **Mass filter**
A limited me
- \div limited mass range ($m/z < 4000$)
- ❖ low resolution
- discrimination of high mass ions
- ❖ MS/MS not possible

Full Scan

Q scan

Selected ion monitoring

Q fix

Ion trap (IT)

- \triangleright limited mass range ($m/z < 6000$, 20000)
- \triangleright low resolution
- \triangleright enables MS/MS (MSⁿ up to 10)

IT-MS

ം Ω

Ion trap

MS scan

- ion capture \bullet
- sequential ion ejecting out from the trap according to m/z \bullet
- ion detection \bullet

MS/MS scan
• ion captur

- ion capture
- isolation of ions with selected *m/z* (precursors) \bullet
- excitation and fragmentation of isolated ions \bullet
- fragment detection (product ions) \bullet

ETD in the HCTultra a C7250

- 1. Electrospray ion accumulation
- 2. Precursor ion isolation
- 3. Reactant anionaccumulation(nCI source)
- 4. ETD fragmentation
- 5. Scan

by courtesy of Dr. Arnd Ingendoh (Bruker)

Reactant Anion Production

Presentation 17th-Oct-2005

M

negative Chemical Ionization (nCI) Source

1st Step:

generation of low energy electrons by EI of $CH₄$ (Mediator)

$$
M + e^{-*} \rightarrow M^{\frac{1}{4}} + 2e
$$

2nd Step:

electron attachment

to **Nouranthene**

 $A + e^- \rightarrow A^{\dagger} (= \bullet)$

by courtesy of Dr. Arnd Ingendoh (Bruker)

Combining Electron-Transfer and Higher-Energy CollisionEThcD

Frese et al., Anal. Chem. 2012, 84, 9668−9673

Linear ion trap

- \triangleright displays advantages of quadrupole and ion trap
- \triangleright limited mass range ($m/z < 6000$)
- \triangleright increased ion capacity by order \blacksquare sensitivity increase
- \blacktriangleright enables MS/MS (MSⁿ)

Trapping Forces in a Linear Ion Trap

Dual pressure ion trap

- \bullet HP chamber –- increase of efficiency of ion traping and their fragmentation
- LP chamber improves resolution and scan speed

Time-of-Flight analyzer (TOF)

 $E = 1/2mv^2$

Detector 1

- \blacktriangleright , unlimited" mass range ($m/z < 1000000$)
- \triangleright fast scanning
- \triangleright high resolution (R až 60 000)
- \triangleright enables MS/MS by PSD (post source decay) not used today

MALDI

Protein (BSA, 66.4 kDa, \approx 15 pmol on spot) C7250

Peptide (\approx 50 fmol on spot) C7250

Ion cyclotron resonance with Fourier transformationFT- ICR MS

 \triangleright ultimate resolution (R >10 000 000)

determination of elemental composition

- \triangleright enables MSⁿ, top-down approaches
- disadvantage (magnet up to 15T, liquid He), high operational costs

In an simplified way,

Significance of high resolution

Sufficient resolution results in possibility to deduce elemental compositionIn case of compounds above we need resolution of 1185

M5 Smallest mass difference ∆m between two equal magnitude peaks such that the valley between them is a specified fraction of the peak height MU, 10/27/2009

Significance of high resolution

Significance of high resolution

Orbitrap

 \blacktriangleright high resolution in MS and MS/MS $\left(\text{up to } 1\text{ 000 000, but w/o magnet}\right)$ \triangleright limited mass range $m/z < 4000$

> ESI

Full Scan – High Resolution, Zoom-in C7250

with permission Ing. Petr Verner (Thermo)

Simulations at different resolution settings

with permission Ing. Petr Verner (Thermo)

Functional Genomics and Proteomics

National Centre for Biomolecular Research Faculty of Science Masaryk University

Basic approaches of data acquisition in MS

DDA – Data Dependent Acquisition One precursor selected for MS/MS at a time

DIA –– Data Independent Acquisition Set of precursors is fragmented simultaneously

Acquisition modes

SWATH MSQ-TOF, MS/MS < 10 ppm

Figure 3. Performance profiles comparing technical advantages and disadvantages of shotgun proteomics, SRM and SWATH MS. In the radar chart, analytical variables are presented on axes staring from the same point and each variable is represented by a spoke. The length of a spoke indicates the magnitude of the variables. Note that SWATH-MS combines the strengths of shotgun and SRM technologies; however, requires more powerful bioinformatic tools for data analysis. LOQ: Limit of quantification; MRM: Multiple reaction monitoring; MS: Mass spectrometry; SRM: Selected reaction monitoring.

Liu et al, Expert Rev. Mol. Diagn., 13(8), 811 (2013)

Unleash the power DIA combined with short gradients (evosep.com)

F**unctional Genomics and Proteomics**
National Centre for Biomolecular Research

National Centre for Biomolecular ResearchFaculty of Science Masaryk University

Hybrid systems

Triple quadrupole, 3-Q

(original design, not hybrid)

- <u>❖</u> robust
- * quantification
- \div limited mass range $(m/z < 4000)$
 \div enables MS/MS (MS²)
- * enables MS/MS (MS²)
- ❖ variety of scan modes
- low resolution
- *◆* ESI

Product Ion Scan

- quadrupole Q1 transmits into collisional cell **only ions with selected** *m/z*
- quadrupole Q3 analyzes all fragments formed in collisional cell by CID (originated from selected ions (precursors) transmitted by Q1)

 $M1$

M1 Product ion scan consists of selecting a precursor ion of a chosen mass-to-charge ratio and determining all of the product ions resulting from collision-induced-dissociation (CID) MU, 10/27/2009

Selected reaction monitoring, SRMMultiple reaction monitoring, MRM

- 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)

improved sensitivitydetection of low abundant components (e.g. biomarkers in complex samples)

ABI documents

Snímek 82

M4 Selected reaction monitoring consists of selecting a fragmentation reaction. For this scan Q1 and Q3 are focused on selected masses There is no scan. The method is analogous to selected ion monitoring in MS but here the ions selected by Q1 are only detected if they produced a given fragment. The absence of scanning allows one to focus on the prcursor and fragment ions over longer times, increasing the sensitivity as for SIM but this sensitivity is now associated with a high increase in selectivity. MU, 10/27/2009

Parallel reaction monitoring (PRM)Z2

Q-Orbitrap MS

 $Z1$

- similar to SRM/MRM
- all fragments of selected precursor are deected with high resolution/accuracy for determination of m/z

- Simplier method adjustment
- improved precursor identification and quantification securing high sensitivity

https://www.creative-proteomics.com/services/parallel-reaction-monitoring-prm.htm

- Z1 PRM is based on Q-Orbitrap as the representative quadrupole-high resolution mass spectrum platform. Unlike the SRM, which performs one transition at a time, the PRM performs a full scan of each transition by a precursor ion, that is, parallel monitoring of all fragments from the precursor ion. First, the PRM uses the quadrupole (Q1) to select the precursor ion, and the selection window is usually m/z≤2; then, the precursor ion is fragmented in the collision cell (Q2); finally, Orbitrap replaces Q3, scans all product ions with high resolution and high accuracy. Therefore, PRM technology not only has the SRM/MRM target quantitative analysis capabilities, but also have the qualitative ability. (1) The mass accuracy can reach to ppm level, which can eliminate the background interference and false positive better than SRM / MRM, and improve the detection limit and sensitivity in complex background effectively; (2) Ful scan of product ions, without the need to select the ion pair and optimize the fragmentation energy, easier to establish the assay; (3) a wider linear range: increased to 5-6 orders of magnitudeZbynek, 2/21/2019
- Z2 Parallel reaction monitoring (PRM) is an ion monitoring technique based on high-resolution and high-precision mass spectrometry. The principle of this technique is comparable to SRM/MRM, but it is more convenient in assay development for absolute quantification of proteins and peptides. It is most suitable for quantification of multiple proteins in complex sample with an attomole-level detection.Zbynek, 2/21/2019

3-Q(LIT instead of Q3)

- Steed sensitivity (enhanced scans)
- * enables MS/MS (MSⁿ)
- low resolution
- *❖* ESI

Linear ion trap

Collisional cell

accumulation of ions during LIT scan (reduced loss of ions)

Q-TOF(TOF instead of Q3)

Q-TOF

Q-Orbitrap(Orbitrap instead of Q3)

C7250

\checkmark high resolution up to 480 000 $\frac{m}{z}$ range ≤ 6000 HCD, ETD

Q-Orbitrap

\mathbf{p} BoxCar acquisition method (extended dynamic range)

- •limitation – C-trap capacity (1 000 000 charges)
- •• capacity is often saturated by high abundant ions
- • BoxCar collects ions in C-trap in narrow m/z segments allowing to accumulate low abundant ions increasing number of identified peptides and extending dynamic range
- •MS/MS scans are restricted
- •• creation of peptide library is required for identification (mostly done at MS level)

Meier F. et al., Nature Methods, 15, 440-448 (2018)

Q-Orbitrap \mathbf{p} $\qquad \qquad \mathbf{C7250}$ BoxCar acquisition method

Meier F. et al., Nature Methods, 15, 440-448 (2018)

Q-OrbitrapBoxCar acquisition method

a 1. Mass-to-charge binning 2. Transmission determination signal Relative abundance Summed time Retention t 1,000 1.200 400 1,000 400 600 800 600 800 1.200 Common m/z grid m/z

MaxQuant assembles the full scan (black traces) and BoxCar scans (red traces) of an acquisition cycle to a single high-dynamic-range scan. First, all spectra are transformed to a common high-resolution *m/z* grid, and the signals from each scan (here **one full scan and two BoxCar scans**) are integrated over the entire LC
olution time (step 1). Erem the integrated signals, the shape of the guadrupele transmission function for eac elution time (step 1). From the integrated signals, the shape of the quadrupole transmission function for each BoxCar scan is globally determined by a pointwise comparison to the full scan (step 2). The resulting relativetransmission factors for each *m/z* bin in each BoxCar scan are used as weights for calculating the average i signal intensity from the full scan and the BoxCar scans. These hybrid spectra are taken as a replacement for
atandard full access in all subsequent are seening stans without further adjustments (stan 3). standard full scans in all subsequent processing steps without further adjustments (step 3).

Meier F. et al., Nature Methods, 15, 440-448 (2018)

Q-OrbitrapBoxCar acquisition method

- •generation of a peptide library
- analysis of individual samples•

Meier F. et al., Nature Methods, 15, 440-448 (2018)

 Ω <u>ം</u>

Orbitrap Fusion™ Lumos Tribrid

 low resolution Resolution Orbitrap 15 000–1 000 000 (FWHM) at m/z 200

> ETD HD – high dynamic range ETD providing significantly increasedfragment ion coverage

Orbitrap – high resolution

TOF-TOF

- \triangleright relatively high resolution (< 30 000)
- \triangleright enables MS/MS (MS²)
- > MALDI
- \triangleright enables off-line connection with LC (LC-MALDI approach)

AB Sciex

ion mobility + mass spectrometry

Trapped Ion Mobility SpectrometrytimsTOFTM

- •coupling ion mobility separation with mass spectrometry
- •ion mobility brings additional separation dimension
- • ion mobility allows separation of structural isomers (in general, compounds with the same or close m/z differing in ccs)

The TIMS analyzer is a segmented rf ion guide wherein ions are mobility-analyzed using an electric field that holds ions stationary against a moving gas, unlike conventional drift tube ion mobility spectrometry where the gas is stationary. Ions are initially trapped, and subsequently eluted from the TIMS analyzer over time according to their mobility (K).

TIMS theory – Michelmann K. et al.: J. Am. Soc. Mass Spectrom. 26,14-24 (2015)

Trapped Ion Mobility SpectrometrytimsTOFTM

Ion mobility separates compounds based on their collisional cross section (CCS), which is primarily a
function of these dimensional shape. function of three-dimensional shape.

Trapped Ion Mobility SpectrometrytimsTOFTM

Trapped Ion Mobility SpectrometrytimsTOFTM

resolution of coeluting compounds with overlapping isotopic patterns

Trapped Ion Mobility SpectrometrytimsTOFTM

- •analysis of HeLa digest (200 ng)
- •120 min LC run

Average number of protein group
 $\frac{1}{2}$ Number of proteins quantified in N out identifications in a single run $(N=4)$ with different TIMS settings of four replicates.

Field Asymmetric Ion Mobility Spectrometry FAIMS

If a mixture of ions of different sizes and types is introduced between two metal plates, the application of high voltage in an appropriate waveform to the plates will create a condition where some types of ions drift and hit the metal plates while other types of ions remain between the plates.

basic principle: http://www.faims.com/howpart1.htm

Analysis of tryptic digest of HeLa cell lysate

FAIMS – Thermo Fisher

The end