La Protéomique :
Etat de l’art et perspectives

Odile Schiltz

Institut de Pharmacologie et de Biologie Structurale
CNRS, Université de Toulouse, Odile.Schiltz@ipbs.fr

Protéomique et Spectrométrie de Masse des Biomolécules

Workshop "Protéomique et Maladies Rares " - Paris - 25 septembre 2012
What is Proteomics?

“Proteomics includes not only the identification and quantification of proteins, but also the determination of their localization, modifications, interactions, activities, and, ultimately, their function.”


Because proteomes are *dynamic*, proteomics great challenge is to measure accurately qualitative and quantitative changes of intracellular and extracellular protein content under different conditions to understand biological processes and define pathological states.
What can Proteomics do?

- To provide lists of proteins and implement databases:
  Descriptive proteomics

- To characterize and quantify proteins:
  Functional proteomics

- To decipher protein connections on a large scale:
  Systems biology
How does MS-based Proteomics work?

Sample preparation

Mass spectrometry analysis

Bioinformatic data analysis

Global proteomic approach for discovery studies:
in depth, unbiased, and quantitative proteome analysis

Proteins → Peptides → nanoLC-MS/MS analysis

Additional steps: fractionation, enrichment

High speed, high resolution acquisition

10 000’s MS and MS/MS scans

Identification
Quantification
Classification

Reliable data processing using dedicated tools

Session 1, Lydie Lane

Session 3, Yves Vandenbrouck

- Differential quantitative analysis of proteomes
- Biomarker discovery
How does MS-based Proteomics work?

Targeted proteomic approach for validation studies:
quantification of known proteins in many samples

- Hypothesis driven studies: set of known proteins in specific pathways, ...
- Biomarker validation

Minimal sample preparation → Proteins → Peptides → Selective, sensitive targeted acquisition → SRM analysis → Quantification Multiplexing → Reliable data processing using dedicated tools
# Mass spectrometry capabilities for global analyses

<table>
<thead>
<tr>
<th></th>
<th>Q-Star</th>
<th>LTQ-Orbitrap-XL</th>
<th>LTQ-Orbitrap-Velos</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Year</strong></td>
<td>2000</td>
<td>2005</td>
<td>2010</td>
</tr>
<tr>
<td><strong>Resolution</strong></td>
<td>10 000</td>
<td>60 000 – 100 000</td>
<td>60 000 – 100 000</td>
</tr>
<tr>
<td><strong>Sequencing speed</strong></td>
<td>4 MS/MS in 10 s</td>
<td>5 MS/MS in 1 s</td>
<td>20 MS/MS in 1 s</td>
</tr>
<tr>
<td><strong>Sensitivity</strong></td>
<td>4 fmol</td>
<td>0.5 fmol</td>
<td>0.5 fmol</td>
</tr>
<tr>
<td><strong>Identified proteins</strong></td>
<td>50</td>
<td>500</td>
<td>1500</td>
</tr>
</tbody>
</table>
Mass spectrometry diversity

Mass spectrometry instrumentation is improving rapidly and constantly

Instrumentation  Application

Orbitrap  Global proteomics
Triple Quadrupole  Targeted proteomics
High resolution FT  Top down analysis of intact proteins
Ion mobility  Intact protein complexes analysis
MALDI TOF  Peptide/protein imaging

Session 1, Charles Pineau
**Functional proteomics**

**Main objectives**

- To understand molecular mechanisms
- To decipher protein interactions and networks
- To characterize cell signaling pathways
- To discover and validate biomarkers

**Proteomic analysis**

- Quantitative proteomics for relative protein abundance and dynamics
- Analysis of protein complexes including labile and transient partners
- Characterization of post-translational modifications
- Targeted proteomics

*Patterson, S.D. & Aebersold, R.H (2003)*
Quantitative proteomics strategies

**Stable isotope labeling strategy**

- Sample 1 → Light isotope labeling
- Sample 2 → Heavy isotope labeling

*Labeling methods: SILAC, ICAT, iTRAQ, …*

**Label-free strategy**

- Sample 1 → LC-MS/MS analysis
- Sample 2 → LC-MS/MS analysis
- …
- Sample n

*MS peak intensities MS/MS counting*
Technical challenges

- Analysis of complex protein mixtures: *fractionation, resolution, acquisition speed*
- Differential quantitative analysis: *repeatability, accuracy, bioinformatic tools*
- Low-abundance proteins: *enrichment, dynamic range, sensitivity*

Need for dedicated strategies for each biological objective in terms of sample preparation, mass spectrometry acquisition, and data analysis
Protein complexes
Protein complexes analysis by mass spectrometry

Structural biology approach
- Stoichiometry of subunits
- Assembly/2D architecture

Proteomic approach
- Protein subunit/partners identification
- Characterization of subunits (PTMs)
- Quantification and Dynamics
- Labile/transient partners identification

In-depth analysis of protein complexes:
Spatial organization and function

Challenges: Amount, stability, heterogeneity, contaminants
Protein complexes analysis workflow

Sample preparation maintaining protein-protein interactions

Protein complex enrichment: Immuno-affinity purification

Quantitative MS analysis

Proteins

Non-specific binding to the beads
Non-specific binding to the antibody
Specific interactions
Non-specific binding to the complex

Beads
Bait

Non-specific binding to the beads
Non-specific binding to the antibody

PROTEINS

TEST

CONTROL

Proteins

T/C Ratio >> 1
Specific protein partners
T/C Ratio ≈ 1
Non-specific protein interactions

Session 3, Romain Roncagalli
Protein complexes analysis

Perspectives and challenges:

- Quantitative MS analysis to study protein complexes dynamics
- Crosslinking combined to MS analysis
- Emerging new MS technologies (ion mobility MS)
- Combining complementary MS approaches
- Interaction networks on a large scale
Analysis of modified proteomes
Analysis of post-translational modifications

- Phosphorylation cascades are involved in many signalling pathways.
- Various modifications regulate microtubule function.
- Plasma-membrane proteins can be linked to the membrane by a GPI anchor.
- Plasma-membrane proteins can carry N-glycans.
- The histone code controls many nuclear processes.
- Polyubiquitylation can induce protein degradation.
- Nuclear and cytoplasmic proteins can carry O-glycans.

Analysis of post-translational modifications

Technical challenges

- Low amount of modified proteins: *enrichment, sensitivity*
- Transient state of the modification: *inhibitors to freeze the system when available*
- Stability of the modification: *appropriate buffers and pH, MS analysis conditions*
- Localization of the modification: *sequence coverage and appropriate MS/MS*

Need to adapt analytical strategies to the modification of interest
Workflow for the analysis of phosphoproteomes

Enrichment for pTyr-containing peptides

Enrichment for all phosphopeptides: mainly pSer- and pThr-containing peptides

Leitner et al., Analytica Chimica Acta, 2011.
# Examples of phosphoproteomes studies

Combination of quantitative approaches and phosphoproteome analyses to study signaling pathways on a global scale

<table>
<thead>
<tr>
<th>Starting material</th>
<th>Objective</th>
<th>Method</th>
<th>Total # of phospho sites</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human Jurkat T cells (100 million cells)</td>
<td>TCR signaling</td>
<td>SILAC TiO₂ and pY enrichments</td>
<td>10,665 (696 regulated) from multiple runs</td>
<td>Mayya et al., Science Signaling, 2009</td>
</tr>
<tr>
<td>Human B cells (630 million cell nuclei)</td>
<td>DNA damage response</td>
<td>SILAC + time points TiO₂ enrichment</td>
<td>7,043 (594 regulated)</td>
<td>Bennetzen et al., Mol Cell Proteomics, 2010</td>
</tr>
</tbody>
</table>
| Human embryonic stem cells (50 million cells) | Differentiation           | SILAC + time points TiO₂ enrichment  | 15,004 from multiple runs  
10,066 (4,504 regulated)  
11,104 (3,380 regulated) | Rigbolt et al., Science Signaling, 2011 |
| Mouse liver (10 mg proteins)           | Insulin signaling          | Spike-in SILAC (Hepa1-6 cells) TiO₂ enrichment | 14,857 (1,000 regulated)                  | Monetti et al., Nature Methods, 2011          |

Remaining challenges:

Phosphorylation sites localization, stoichiometry, interplay with other PTMs
Biomarkers
Process flow for the development of biomarkers

Proteomic analysis of biological fluids

Technical challenges:
- Protein concentration dynamic range
- Few proteins are highly abundant

Protein concentration dynamic range: \(10^{10} - 10^{12}\)

Mass spectrometry dynamic range: \(10^{3} - 10^{4}\)

Equalization

Proteins of interest

Depletion

Session 3,
Anne Gonzalez de Peredo
SRM analysis of PSA biomarker in patients serum

**Experimental workflow**

**External calibration**
- 100 µl blank female serum spiked with PSA
- Immunodepletion of albumin
- Trypsin digestion
- Solid Phase Extraction on HLB cartridges
- Peptides fractionation by SPE on MCX cartridge
- Vacuum drying
- Peptides resuspended in 200 µL of CH$_3$CN/H$_2$O (3:97, v/v)/0.1 % formic acid
- LC-SRM/MS analysis
  - Injection of 50 µl (2.1 mm inner diameter column)

**Clinical samples**
- 100 µl patient serum
- Immunodepletion of albumin
- Trypsin digestion
- Solid Phase Extraction on HLB cartridges
- Peptides fractionation by SPE on MCX cartridge
- Vacuum drying
- Peptides resuspended in 200 µL of CH$_3$CN/H$_2$O (3:97, v/v)/0.1 % formic acid
- LC-SRM/MS analysis
  - Injection of 50 µl (2.1 mm inner diameter column)

**SRM and ELISA correlation for PSA quantification**

LOQ: low ng/ml range

Proteomics community

France

Europe

International

http://www.sfeap.fr/

http://www.eupa.org/

http://www.hupo.org/
Acknowledgements

"Proteomics and Mass Spectrometry of Biomolecules"
"Proteomics Infrastructure of Toulouse" - http://proteomique.ipbs.fr/
"Proteomics French Infrastructure (ProFI)"