Analytical methods in Proteomics

Otaniemi, May 2012
Meilahti Clinical Proteomics Core Facility

Biomedicum Helsinki and Haartman Institute, 00014 University of Helsinki, Finland

E-Mail: marc.baumann@helsinki.fi

(http://research.med.helsinki.fi/corefacilities/proteinchem)
- Proteomic services from Bedside-to-Bench
- High-throughput analysis of Clinical Samples (CSF, Saliva, Urine, Blood, Ascites, Tears etc.)
- Validation of Clinical Samples
- Validation of Clinical Biomarkers

- Basic Proteomics Analyses (1-2D-GE, 1-2D-LC)
- Ion Mobility Structural Analyses

- Array-based epitope mapping (Pep/Prot-Arrays)
- Pep-Array synthesis
- Imaging Mass Spectrometry (IMS) from tissue to cells
- Glycoproteomics
- Glycopeptidomics
- LC-MS/MS2 analysis of carbohydrates
- N-glycopeptide spectra analysis
- Core fucosylation analysis
Over 1200 Researchers in only Medical Research
(Cancer, Genetics, Developmental Medicine, Neuroscience etc.)
Proteomics

What is it all about??
2001: Human Genome Project Reveals

3,000,000,000 base pair nucleotides
= only about 25,000 genes
And that ...

Only 0.1% of each’s persons DNA differs from any other person
So... what makes the Difference?

= Proteins

The study of the Proteins expressed by a Genome

= Proteome
same Genome - different *Proteome*
Proteomic Research today

The main objectives today for clinical and general Proteomics:

- Quantification of all the proteins expressed in a cell or tissue proteome, body fluids e.g. blood, CSF etc. Searching for Biomarkers!

- Functional study of thousands of proteins in parallel, which protein is in contact to another protein and where? Searching for functionality!

For quantification purposes, standard method is 2DE electrophoresis or MudPIT separation followed by MS identification

For protein function studies, microarray based assays are used to study protein-protein and protein-ligand interactions
Proteins Rule
Biotech’s latest mantra is “proteomics,” as it focuses on how dynamic networks of human proteins control cells and tissues
By Carol Ezzell | Wednesday, April 24, 2002 | 0

Move over, human genome, your day in the spotlight is coming to a close. Researchers are now concentrating
The Post-Genome Project
Whether The Human Proteome Will Be Successfully Mapped In Three Years Depends On How You Define "Proteome"
By Karen Hopkin  | August 17, 2001  | 0
A Long Way to the Bedside

Despite many breakthroughs, Proteomics has not translated yet to patient care.
How can we make Proteomics more suitable to the “real” life?

The technology...

2D gel electrophoresis
2D liquid chromatography
Micro arrays
We need a technology to find changes in Proteome

Administration of a drug known to bind to an orphan receptor

- Changes in expression level of 23 proteins

A protein Array
How Proteomics Can Help Drug Development

Finding New Drug Targets
(Here, developing a drug to kill skin cancer melanoma)

1. Gels separate proteins by charge and mass. Comparing them reveals a protein (circled) overproduced by cancer.

2. Protein is isolated and crystallized.

3. X-ray crystallography reveals the protein's structure.

4. Drugs can be designed to block the protein's activity.

Avoiding Drugs with Side Effects
(Here, determining whether an investigational drug prompts production of possibly harmful proteins)

Investigational Drug

Blood Serum

No Drug

Two-dimensional gels show that the drug prompted the production of new proteins—some innocuous (green) and some with potential to cause side effects (red).
Two-dimensional gel electrophoresis (2D) could do it?

- 1st dimension, IEF, Proteins are separated according to their isoelectric point (IP)
- 2nd dimension, SDS-PAGE, Proteins are separated according to their molecular mass
- Efficient: More than a thousand proteins resolved in E-Coli cell lysates and ~8000 in brain lysates
1st Dimension - Isoelectric Focusing

pH4

pH5.5

pH7

neutral

gelsurface

pI - isoelectric point

ready made Gel-strips
2\textsuperscript{nd} Dimension - Isoelectric Focusing

2DE
Mw
Protein Fingerprint: 2-DE

About 2000 proteins
CyDye DIGE Fluor dyes
Minimal labelling dyes

- Label 50 µg of protein
- 3 colors: Cy™2, Cy3, Cy5
- MW matched (~450Da)
- Charge matched (positive)
- Label ε-amino group of lysine
- Sensitivity - 0.025 ng
- Linear dynamic range over 4 orders of magnitude
Achieving accurate quantitative data

Ettan™ DIGE system

- Pooled internal standard label with Cy™2
- Protein extract 1 label with Cy3
- Protein extract 2 label with Cy5

Mix labelled extracts → 2-DE separation → Typhoon™ Variable Mode Imager → DeCyder™ Differential Analysis Software
Overlay of normal and patient protein samples

Ettan™ DIGE
the quantitative approach to do Proteomics

Increased abundance

Equal abundance

Reduced abundance

Normal control = Cy™3 labelled - Blue
Patient A sample = Cy5 Labelled - Red
2D Databases exist!

<table>
<thead>
<tr>
<th>Human Tissue/Cell Type</th>
<th>Image</th>
<th>Image</th>
<th>Image</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colorectal epithelia cells</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
</tr>
<tr>
<td>Cerebrospinal Fluid</td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
</tr>
<tr>
<td>Colorectal adenocarcinoma cells</td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
<td><img src="image9.png" alt="Image" /></td>
</tr>
<tr>
<td>Erythroleukemia Cell</td>
<td><img src="image10.png" alt="Image" /></td>
<td><img src="image11.png" alt="Image" /></td>
<td><img src="image12.png" alt="Image" /></td>
</tr>
<tr>
<td>HepG2 Secreted Proteins</td>
<td><img src="image13.png" alt="Image" /></td>
<td><img src="image14.png" alt="Image" /></td>
<td><img src="image15.png" alt="Image" /></td>
</tr>
<tr>
<td>HepG2</td>
<td><img src="image16.png" alt="Image" /></td>
<td><img src="image17.png" alt="Image" /></td>
<td><img src="image18.png" alt="Image" /></td>
</tr>
<tr>
<td>Promyelocytic leukemia cells</td>
<td><img src="image19.png" alt="Image" /></td>
<td><img src="image20.png" alt="Image" /></td>
<td><img src="image21.png" alt="Image" /></td>
</tr>
<tr>
<td>Kidney</td>
<td><img src="image22.png" alt="Image" /></td>
<td><img src="image23.png" alt="Image" /></td>
<td><img src="image24.png" alt="Image" /></td>
</tr>
<tr>
<td>Liver</td>
<td><img src="image25.png" alt="Image" /></td>
<td><img src="image26.png" alt="Image" /></td>
<td><img src="image27.png" alt="Image" /></td>
</tr>
<tr>
<td>Lymphoma</td>
<td><img src="image28.png" alt="Image" /></td>
<td><img src="image29.png" alt="Image" /></td>
<td><img src="image30.png" alt="Image" /></td>
</tr>
<tr>
<td>Plasma</td>
<td><img src="image31.png" alt="Image" /></td>
<td><img src="image32.png" alt="Image" /></td>
<td><img src="image33.png" alt="Image" /></td>
</tr>
<tr>
<td>Platelet</td>
<td><img src="image34.png" alt="Image" /></td>
<td><img src="image35.png" alt="Image" /></td>
<td><img src="image36.png" alt="Image" /></td>
</tr>
</tbody>
</table>
The simple way to test “real” Proteomics?

An example from the clinics…
Stem cells are located in the lower crypt poles, colonocytes are migrating to the surface of the crypt.

Continually replacement in the toxic and mechanically stressing environment of the gut.

Hyperproliferation is believed to precede adenoma formation.

Loss of APC function leads to deregulated crypt homeostasis and is thought to be the initiating event in adenoma formation.
Correlation of Histological Progression and Molecular Changes in Colorectal Cancer
Samples

- Polyp
- Normal
- Carcinoma
- Liver metastasis
Tissue microdissection
Why Tissue Microdissection?

You would like to isolate only the targeted diseased tissue
Laser-based microdissection:
Laser Capture Microdissection (Arcturus Autopix)
Capturing of wanted tissue section to the slide
MALDI MS analysis directly from the tissue captured on the cap membrane
Sample enrichment/preparation

1st Dimension

2nd Dimension

PC based matching

2-DE Algorithm

Mass spectrometrical identification

Statistics
Deviating Proteins (n=112)
Identification of 72

N=26

benign
maligne

Controls: normal liver tissue
HCT116 and Lovo cell lines

N=46

benign
maligne
Principal Component Analysis (t1,t2,t3)
Intra-individual expression differences of Cytokeratin 20 in patient 14.

The left gel segment (a) is zoomed from the normal mucosa,
gel b represents the patients polyp and
gel c is the corresponding segment of the same patients adenocarcinoma
Micro gel devices

gel reservoir: 25x32x0.3/0.5 mm PMMA, silica

Running time 10 minutes
Automated 2D devices

The compress system
2-D map of IEF standards

Repeatability of 2-DE runs

<table>
<thead>
<tr>
<th>Rf values (%)*</th>
<th>pI position errors(%)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>STDV</td>
<td>6,1</td>
</tr>
<tr>
<td>max</td>
<td>15</td>
</tr>
<tr>
<td>min</td>
<td>0,8</td>
</tr>
<tr>
<td>STDV</td>
<td>2,5</td>
</tr>
<tr>
<td>max</td>
<td>6</td>
</tr>
<tr>
<td>min</td>
<td>0,6</td>
</tr>
</tbody>
</table>

* comparison of 3 gels

- 2-DE separation completed in **approx. 80 min**
- Limit of detection is **approx. 65 ng**

Running time 20-30 minutes
Examples for the Clinics

- Performance
  - Native IEF and native PAGE
    - 5 variants of hemoglobin
    - pH 6.7 -7.7
  - Native IEF and SDS-PAGE
    - standard IEF proteins
    - pH 3-10
  - Denatured IEF and SDS-PAGE
    - GFAP protein variants expression differences
    - in control and Alzheimer diseased patients
    - pH 4-6
2D-PAGE Direct In-situ Digest

100 fmol
running the gel
- staining each spot of interest (➔)
- excise
- *in-gel* digestion

➔ PEPTIDE ANALYSIS

Trypsin digest
The MALDI AutoPrep Robot System

Automated MALDI sample preparation of robot systems - including the mixing, washing or purification with Millipore 8-channel liquid handler.

MAP II and MAP II/8 use disposable plates. Samples can be placed in capped plates. Plate formats are supported and can be used as MALDI target.

One single sample spreadsheet controls all steps from sample preparation, measurement, control and the FLEX III MALDI-TOF guiding the path from sample preparation to data generation.
Prespotted AnchorChip targets
dedicated for Clinical Proteomics

> no cross contamination

> no memory effects

> Easy storing and archiving of targets

> Re-visiting of samples

> One calibrant, four matrix spots for samples

> 10-100 fold increased sensitivity, in situ sample purification
MALDI TOF/TOF
High Sample Throughput for the Post-Genomic Era

MALDI-TOF-MS

The TWISTER
Typical mass spectrum from a protein band trypsin digest
Automated Database Search (MASCOT)

Number 1 match: tumor necrosis factor type 1 receptor associated protein TRAP-1 (Mr): 76030.27

Total coverage: 33.4%
Even a single peptide is enough for Protein identification!
MALDI MS TOF/TOF or Quadrupole MS TOF/TOF

Fragmentation of a single peptide in the Mass Spectrometer

>>>> Database search

>>>> Identification of the whole protein
## Automated Database Search (MASCOT)

Number 1 match: tumor necrosis factor type 1 receptor associated protein TRAP-1 (Mr): 76030.27

| 1  | 51 | RALPRAPALA  | AVPGKPiLC  | PRETTAQLGP  | RRNPAAWSLQA  | GRLESTQTAE  |
| 101 | 151 | DHEEPLHSII  | SSV تسقط | SKIEFQAETK  | KLLDIVARSL  | YSEKEVFIRE  |
| 201 | 251 | EELVSNLGTI  | ARSSGKAFLD | ALOMQARFAS  | RWIQQFVGIF  | YSAEVADRV  |
| 301 | 351 | SEARVRDVVT  | KYSNFVSPPL  | YLNGRRMNLT  | QAISMDPKD  | VGWQHEEFY  |
| 401 | 451 | RVQAQHKDP  | RYTLHYKTDGA  | PLNIRSFYYV  | PDKPSMFVDV  | SPELGSSVAL  |
| 501 | 551 | YSRKVLQITK  | ATDIILPKLR  | FIRGVVDSED  | IPLNLSEELL  | QESALIRLKL  |
| 601 | 651 | DRSFAAECLS  | EMEETELMAN  | MNRLGSRVT  | NVKVTNLRTD  | HDAMVTVLEM  |

**Sequence successfully identified**
Applications

Top-Down UPLC-ETD sequencing of Proteins
3 Protein Mixture: Ubquitin (5μM), Myoglobin (10μM) and Calbindin (20μM)

3.5μm Symmetry 300 C4 (75μm x 100mm)
Normal Protein MS
Ubiquitin (5μM) MS
Intact Tandem MS of Bovine Ubiquitin
ETD of m/z 714, [M+12H]$^{12+}$
UBIQ BOVIN  Mass: 8560  Score: 181  Queries matched: 1
Ubiquitin - Bos taurus (Bovine)

Monoisotopic mass of neutral peptide Mr(calc): 8559.6167
Ions Score: 181  Expect: 6.6e-14
Matches (Bold Red): 151/300 fragment ions using 1893 most intense peaks