Protein Chip Development and Applications

Heng Zhu
Dept. of Pharmacology & the HiT Center
Johns Hopkins University School of Medicine

Background in Proteomics

Protein Chip Development

Application in Basic Research

Applications in Clinical Research

Background in Proteomics

Progresses in Proteomics
Protein profiling
2D-MS, analytical protein chip
High-throughput protein localization
Transposon vs GFT
Biochemical Genomics
Pooling strategy
Large-scale protein interaction mapping
Y2H and protein complex coupled w/ MS
Transcription factor-DNA interaction
ChIP-chip
High throughput biochemistry assays
Functional protein chips

Why Microarrays?
Higher Sensitivity
Much higher throughput
More flexibility
Less sample consumption
Quantitation
Direct target detection

Protein Microarrays
• Protein-Protein Interactions
• Protein Modification and Regulation
• Serum Profiling
• Signaling Pathways
• Drug Discovery
Comparison of Interaction Proteomics

<table>
<thead>
<tr>
<th>Approach</th>
<th>Application</th>
<th>Advantage</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yeast two-hybrid</td>
<td>Protein-protein interactions, DNA interactions</td>
<td>High-throughput and systematic approach to reveal protein interactions</td>
<td>No control over interaction conditions, interactions are usually in the nucleus</td>
</tr>
<tr>
<td>Affinity tagging MS</td>
<td>Dissociating protein complexes</td>
<td>In vivo interactions that involve multiple partners</td>
<td>May miss transient or weak interactions, hard to identify false positives</td>
</tr>
<tr>
<td>Antibody array</td>
<td>Protein profiling, protein detection, clinical diagnostics</td>
<td>Very sensitive and low sample consumption, great potential in biomarker and drug development</td>
<td>Highly restricted by the quantity and quality of available antibody, semi-quantitative protein detection</td>
</tr>
<tr>
<td>Functional protein array</td>
<td>Diverse, e.g., protein-protein, protein-ligand, protein-small molecule, enzyme-substrate interactions as well as drug discovery and posttranslational modifications</td>
<td>Great potential for analyzing biochemical activities of proteins and high-throughput drug and drug target screening</td>
<td>In vitro assays</td>
</tr>
<tr>
<td>Peptide array</td>
<td>Enzyme-substrate interaction and drug discovery</td>
<td>Sensitive and straightforward to identify epitopes</td>
<td>Expensive to fabricate, in vitro assays</td>
</tr>
<tr>
<td>Carbohydrate array</td>
<td>Carbohydrate-mediated molecular recognition and inhibition response</td>
<td>A new and sensitive way to study carbohydrate-mediated molecular events</td>
<td>In vitro assays; tough to acquire carbohydrate molecules in pure forms</td>
</tr>
<tr>
<td>Small molecule array</td>
<td>Protein-small molecule interaction, drug discovery, enzyme specificity profiling</td>
<td>Minimum small molecule consumption and high sensitivity</td>
<td>In vitro assay, necessary to improve throughput to cover 10^9 molecules in a normal combinatorial chemistry library</td>
</tr>
</tbody>
</table>

Protein Chip Development

Surface Structure

- Porous Surface
- 3-D Surface Structure
- Nanowell
- Plain Glass Surface

Chemically Modified Surface

- Non-specific diffusion
- Adsorption/absorption
- Covalent Cross-linking
- Affinity Attachment

Common Surfaces to Immobilize Proteins

- Covalent linkage
- Adsorption absorption
- Affinity-based linkage
Comparison of Surface Chemistry

<table>
<thead>
<tr>
<th>Surface</th>
<th>Attachment</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>PVDF</td>
<td>Adhesion</td>
<td>No protein modification requirement, high protein binding capacity</td>
<td>Non-specific protein attachment in solution</td>
</tr>
<tr>
<td>Noncovalent</td>
<td>Adhesion</td>
<td>No protein modification requirement, high protein binding capacity</td>
<td>Non-specific binding, high background, low density arrays</td>
</tr>
<tr>
<td>Poly-lactic acid</td>
<td>Adsorption</td>
<td>Non-specific adsorption requirement</td>
<td></td>
</tr>
<tr>
<td>Silica-coated</td>
<td>Covalent cross-linking</td>
<td>High density and strong protein attachment, High-resolution detection methods available</td>
<td>Random orientation of surface attached proteins</td>
</tr>
<tr>
<td>Ni-NTA coated</td>
<td>Affinity binding</td>
<td>Strong, specific and high-density protein attachment, low-background</td>
<td>Proteins have to be biotinylated</td>
</tr>
<tr>
<td>Gold-coated</td>
<td>Covalent cross-linking</td>
<td>Strong and high-density protein attachment, low-background. Can be easily coupled with SPR and Mass spectrometry</td>
<td>Random orientation of surface attached proteins. Rough to fabricate, not commercially available</td>
</tr>
<tr>
<td>PDMS nanowell</td>
<td>Covalent cross-linking</td>
<td>Strong and high-density protein attachment, well suited for sophisticated biochemical analyses</td>
<td>Random orientation of surface attached proteins</td>
</tr>
<tr>
<td>2-D gel and agarose gel</td>
<td>Affinity binding</td>
<td>High protein binding capacity, no protein modification requirement</td>
<td>Tough to fabricate, not commercially available</td>
</tr>
<tr>
<td>DNA/RNA coated</td>
<td>Hybridization</td>
<td>Strong, specific and high-density protein attachment, low-background.</td>
<td>Sophisticated in vitro production of labeled proteins</td>
</tr>
</tbody>
</table>

Comparison of Detection Methods

<table>
<thead>
<tr>
<th>Detection</th>
<th>Probe labeling</th>
<th>Data acquisition</th>
<th>Real time</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELISA</td>
<td>Enzyme-linked antibodies</td>
<td>CCD imaging</td>
<td>No</td>
<td>Low</td>
</tr>
<tr>
<td>Isotopic labeling</td>
<td>Radio isolate-labeled analyte</td>
<td>X-ray film or phosphomimager</td>
<td>No</td>
<td>High</td>
</tr>
<tr>
<td>Sandwich</td>
<td>Immunosensor</td>
<td>Fluorescently-labeled antibodies</td>
<td>Laser scanning</td>
<td>No</td>
</tr>
<tr>
<td>SPR</td>
<td>Not necessary</td>
<td>Refractive index change</td>
<td>Yes</td>
<td>Low</td>
</tr>
<tr>
<td>Non-contact AFM</td>
<td>Not necessary</td>
<td>Surface topological change</td>
<td>No</td>
<td>High</td>
</tr>
<tr>
<td>Planar waveguide</td>
<td>Fluorescently-labeled antibodies</td>
<td>CCD imaging</td>
<td>Yes</td>
<td>High</td>
</tr>
<tr>
<td>Silicon biosensor</td>
<td>Fluorescently-labeled antibodies</td>
<td>CCD imaging</td>
<td>Yes</td>
<td>High</td>
</tr>
<tr>
<td>SELDI</td>
<td>Not necessary</td>
<td>Mass spectrometry</td>
<td>No</td>
<td>Low</td>
</tr>
<tr>
<td>Electro-chemical</td>
<td>Metal-coupled analyte</td>
<td>Conductivity measurement</td>
<td>Yes</td>
<td>Medium</td>
</tr>
</tbody>
</table>

Protein Microarrays Are of Two Types

<table>
<thead>
<tr>
<th>Analytical</th>
<th>Functional</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antibody</td>
<td>Protein</td>
</tr>
<tr>
<td>Antigen</td>
<td>Peptide</td>
</tr>
<tr>
<td>Aptamer</td>
<td>Peptide</td>
</tr>
<tr>
<td>Allergen</td>
<td>Drugs</td>
</tr>
</tbody>
</table>

Application in Basic Research

- Serum probes
- Cell lysates
- Living cells

Protein expression level
Protein profiling
Diagnostics

Protein binding properties
Pathway building
Posttranslational modification
Drug discovery

Summary of Protein Chip Applications

- Protein-protein
- Protein-liposome
- Protein-drug
- Enzyme-substrate
- Antigen-antibody

Application of Analytical Microarrays

<table>
<thead>
<tr>
<th>Analytical</th>
<th>Antibody Chip</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antibody</td>
<td>Antigens</td>
</tr>
<tr>
<td>Antigen</td>
<td>Serum probes</td>
</tr>
<tr>
<td>Aptamer</td>
<td>Cell lysates</td>
</tr>
<tr>
<td>Allergen</td>
<td>Living cells</td>
</tr>
</tbody>
</table>

Protein expression level
Protein profiling
Diagnostics
Protein Microarrays Are of Two Types

**Functional**
- Protein
- Peptide
- Drug probes
- Enzymes

**Key points**
- Clone collection
- Cloning strategy
- Yeast, *C. elegans*, humans
- Protein production
- Hosts for making proteins
- Affinity tags
- In vitro system

---

**Protein binding properties**
- Pathway building
- Posttranslational modification
- Drug discovery

---

**Nanowell Chip**
- Round shape wells
- 1.4 mm diameter
- 300 micron depth
- 2.0 mm pitch
- 300 nl volume

**Nanowell Chips and Protein Attachment**
- Laser Milling Tool
- Etched Mold
- Pour PDMS on Surface
- Nanowells
- Activated by Crosslinker
- Substrate Attached
- Kinase Assay

---

**Modified GST Expression Vector pEGH**
- GST
- 6xHis
- Polylinker
- Yeast ORFs
- ATG
- STOP
- 2µ Plasmid Vector

---

**Kinase-Substrate Assays on Nanowell Chips**
- Substrate
- ATP
- ADP
- Kinase
- 30°C
- Wash
**Autophosphorylation Assays on Nanowell Chips**

- Kinase Proteins
- ATP
- 30°C
- Wash
- ATP
- ADP

**Kinase Assays Using Protein Chips**

- Gic2

**Swi6p as a Substrate**

- Activity
- SI

**Yeast Proteome**

- 6282 Protein Coding Genes
- 4042 Characterized
- 2244 Uncharacterized
- 334 Homologs
- 1910 Unique

**Cloning Strategy**

- Yeast ORFs
- Transformed into Yeast
- Restriction digestion and sequencing
- Plasmid preps
- Rescued into E. coli

**96-Well Yeast Protein Purification**

- Induce cells
- Supernatant
- Filtrate
- Glutathione Sepharose
- Combine
- High Salt Wash
- Low Salt Wash
- Elute
- Aliquot
- Vortex to lyse
- Spin
Producing the Yeast Proteome

~80% full-length proteins
5800 expression clones 93.7%

GST-His6::ORF

KD 250 175 105 75 60 55 35 20

GST: P1      GST: P2    GST: P3

The Yeast Proteome Chip

95% above background
>90% spots: 50 to 950 fgs

Types of Assays Developed

• Protein-protein
• Protein-antibody
• Protein-lipid
• Protein-DNA/RNA
• Protein-drug
• Protein-small molecule
• Phosphorylation
• Acetylation
• Ubiquitinylation
• Glycosylation

Antibody Probing of the Yeast Proteome Microarray

<table>
<thead>
<tr>
<th>Antibody</th>
<th># of +s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monoclonal (3 Yeast + 3 Control)</td>
<td>α-Sed3, α-Cox4</td>
</tr>
<tr>
<td></td>
<td>α-Pep12</td>
</tr>
<tr>
<td>Anti-Peptide Polyclonal (6)</td>
<td>α-Hda1</td>
</tr>
<tr>
<td></td>
<td>α-Mad2</td>
</tr>
<tr>
<td>Anti-FL Protein Polyclonal (2)</td>
<td>α-Nap1</td>
</tr>
<tr>
<td></td>
<td>α-Cdc11</td>
</tr>
</tbody>
</table>

Calmodulin-Binding Proteins

• 12 Known or Suspected Targets
• 33 New Binding Proteins
• Derived New Consensus Binding Site

Calmodulin
Validation of Bmh1,2 Targets

- ~140 in vitro targets
- 4 of 5 targets verified co-IP.
- All 4 IP preferentially with Bmh1

Drug Discovery and Target Validation

- Small Mole Library
- Phenotype Screening
- Drug Candidates
- Candidate Targets
- Proteome Microarrays
- Genetics Cell Biology Expression Profiling
- Protein Targets?

SMIR3 & 4 Function in Tor1/2 Pathway

Identification of Drug Targets
**COX1 Expression Is Regulated by Arg5,6**

![](image)

**Posttranslational Modification**
- Phosphorylation
- Dephosphorylation
- Acetylation
- Ubiquitinylation
- Glycosylation

**Network Biology**

**Kinase Assays on Protein Chips**

*Han et al., Nature 430:88-93, 2004*

*MacBeath et al.*
Identification of Kinase Substrates

Kinase Assays Are Specific

Average No. = 47, ranging from 1 to 256
73% substrates were recognized by fewer than 3 kinases

Kinases Often Recognize Functional Classes of Protein Substrates

Related Kinases Recognize Different Substrates

Previously Known Kinase-Substrate Interactions in Yeast

Phosphorylation Network in Yeast
New Regulatory Modules Are Revealed

- Module 1: 1,563 P = 0
- Module 2: 2,448 P = 0
- Module 3: 25 P = 0.001
- Module 4: 145 P = 0
- Module 5: 92 P = 0
- Module 6: 147 P = 0

New Regulatory Modules Are Revealed

Versatile Ubiquitin – Different Functions For Different Length And Position

Mono-Ub: protein sorting, protein-protein interaction, virus budding

Poly-Ub: K48: protein degradation (26S proteasome)
K63: IKK activation
K29: protein degradation (26S proteasome)

Different Functions for Different Length and Position

- Mono-Ub: protein sorting, protein-protein interaction, virus budding
- Poly-Ub: K48: protein degradation (26S proteasome), K63: IKK activation, K29: protein degradation (26S proteasome)

Strategy to Identify HECT substrates

- Uba1/Ubc4/E3 (K48) + biotin-Ub + anti-GST
- Cy5-Streptavidin Cy3-2' antibodies

Ubiquitinylation by Rsp5 + K63 Ub

- Red: Cy5-streptavidin
- Green: anti-GST
- Protein amounts

Cecile Pickart, Cell 116:181-90, 2004


Ubiquitinylation Requires Three Enzymes Sequentially

- A. HECT domain E3a
- B. RING domain E2s

Bioinformatics Strategy to Identify HECT substrates

- Uba1/Ubc4/E3 (K48) + biotin-Ub + anti-GST
- Cy5-Streptavidin Cy3-2' antibodies
Many Substrates of Rsp5 and Hul5 Are Shared in K63 Linkage

Few Substrates of Rsp5 and Hul5 Are Shared in K48 Linkage

Top Candidate Substrates – Rsp5 With Both Forms of Ubiquitin

<table>
<thead>
<tr>
<th>Gene</th>
<th>Protein</th>
<th>Domain</th>
<th>Function</th>
<th>Binding with Rsp5</th>
</tr>
</thead>
<tbody>
<tr>
<td>YDL070W</td>
<td>Bdi2</td>
<td>bromodomain</td>
<td>transcription factor (predicted)</td>
<td></td>
</tr>
<tr>
<td>YHR097C</td>
<td>Arrestin</td>
<td>unknown</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YJL084C</td>
<td>Bub1</td>
<td>Ub-dependent protein degradation</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>YMR275C</td>
<td>Dia1</td>
<td>unknown</td>
<td></td>
<td>yes</td>
</tr>
<tr>
<td>YOR042W</td>
<td>Cue5</td>
<td>monoubiquitin binding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YPR030W</td>
<td>Csr2</td>
<td>Galactose transport?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Top Candidate Substrates – Rsp5 With Ub K63 Only (Specific?)

<table>
<thead>
<tr>
<th>Gene</th>
<th>Protein</th>
<th>Domain</th>
<th>Function</th>
<th>Binding with Rsp5</th>
</tr>
</thead>
<tbody>
<tr>
<td>YJL031C</td>
<td>Bet4</td>
<td>CAAX-protein GG-transferase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YML013W</td>
<td>Se1</td>
<td>UBX protein sorting, Ub-dependent protein degradation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YMR140W</td>
<td>Slp5</td>
<td>unknown</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YNL094W</td>
<td>App1</td>
<td>actin cytoskeleton assembly</td>
<td></td>
<td>no, but bind to Rvs167</td>
</tr>
<tr>
<td>YPR154W</td>
<td>Pin3</td>
<td>SH3</td>
<td>actin cytoskeleton assembly</td>
<td></td>
</tr>
</tbody>
</table>
Applications in Clinical Research

- Antibody-carbohydrate
- Carbohydrate chip
- Sera

- Antigen-antibody
- Antibody chip
- Antigen chip

- Allergen-IgE
- Allergen chip
- IgE

- Antibody-carbohydrate
- Carbohydrate chip
- Sera

Applications in Clinical Diagnostics

- Antibodies
- Serum

- Antigen-antibody
- Antigen chip

- IgE
- Allergen-IgE

Serum Profiling

Normal Serum vs Patient Serum

Serum Profiling

α-IgG

α-IgM

α-IgA

Merged

Allergen Microarray

MHC Chips to Profile T Cells


Activation of OT-1 Lymphocytes on an MHC Array

SARS Coronavirus

RNA virus
spherical in shape
club-shaped peplomers
HCV:
- SARS
- 229E
- OC43

Experimental design

• Cloning:
  SARS-CoV: Human
  229E: Human
  OC43: Human
  FIPV: Cats
  MHVA59: Mouse

• Expression:
  Yeast and E. coli.

Serum Probing on Coronaviral Chips

SARS Patients Tested

Three Datasets:
- China I (56): Sera from recovered patients
- China II (150): Fever patients
- Toronto (350): Fever patients

Toronto Dataset:
- 521 probings
  493 good
  262 normal
  231 SARS
  28 bad – will be repeated

- 65 unique features
  protein fragments from 5 viruses

Hierarchical Clustering