CMSC 838T – Lecture 12

**Proteomics**
- Study of the proteome

**Methods**
- 2D electrophoresis
- Mass spectrometry
- Protein arrays & microarrays

Ciphergen Protein Chip Autoloader

Proteomics

- **The study of the proteome (all proteins) of an organism**
  - What proteins are present?
  - What is the 3D structure of each protein?
  - How do proteins interact (networks)?
  - How do proteins correlate with DNA?

- **The proteome is dynamic**
  - Varies with time, in response to environment
  - Proteins are very cell / tissue-specific

- **The proteome is large**
  - 100,000+ human proteins (vs. 30,000+ genes)
  - Due to alternative splicing, post-translation modifications
Proteomics – Experimental Methods

- **2D electrophoresis**
  - Separating expressed proteins
- **Mass spectroscopy**
  - Identifying expressed proteins
- **Protein arrays / microarrays**
  - Identifying expressed proteins & interactions
- **Other techniques**
  - X-ray crystallography
  - NMR spectroscopy
  - Edman degradation
  - Etc…

Proteomics

- **Overview**
  - 2D electrophoresis
  - Mass spectrometry
  - Protein arrays / microarrays
Proteomics – 2D Electrophoresis

◆ Property of proteins
  - Some amino acids are acidic / basic (donate / accept H+)
  - Collection of amino acids in protein determines its pI value
    • pI = pH at which molecular charge = zero

◆ 2D electrophoresis
  - Separate proteins according to both pI & molecular weight

CMSC 838T – Lecture 12

Proteomics – 2D Electrophoresis

◆ Method
  1. Extract & prepare protein sample in solution
  2. Separate proteins (in each dimension)
    I. Based on pH
      ◆ Using isoelectric focusing (IEF)
      ◆ Using immobilized pH gradient (IPG) strips
    II. Based on molecular weight (size)
      ◆ Using gel electrophoresis
  3. Stain proteins to enable visualization

CMSC 838T – Lecture 12
Proteomics – 2D Electrophoresis

2D PAGE (polyacrylamide gel electrophoresis) examples

human liver

human kidney
Proteomics – 2D Electrophoresis

**Observations**
- Able to separate proteins with pI around 0.0025 pH units
- Relative amounts of protein quantified by intensity
- Provides clear pattern of protein expression
- Comparisons can identify protein expression differences
- Good for identifying novel proteins

**Limitations**
- Unable to identify all proteins based on pI / size alone
- Cannot handle extremely acidic / basic proteins
- Misses some large proteins & membrane proteins
- Limited reproducibility of 2D gels
- Relatively slow, expensive process

Proteomics

**Overview**
- 2D electrophoresis
- Mass spectrometry
- Protein arrays / microarrays
Proteomics – Mass Spectrometry (MS)

- **Mass spectrometer**
  - Particles are
    1. Ionized (charge added)
    2. Accelerated through magnetic field
      - Path taken is function of mass / charge ratio
    3. Vary magnetic field
    4. Record quantity of particles for different m/z ratios

Mass Spectrometer
Proteomics – Mass Spectrometry

◆ **Goal**
  - Identify proteins separated by 2D electrophoresis

◆ **Method**
  1. Excise individual dots from 2D electrophoresis
  2. Digest protein into fragments with enzyme (e.g., trypsin)
  3. Ionize protein fragments (without breaking)
     - Matrix Assisted Laser Desorption Ionization (MALDI)
     - Electrospray Ionization (ESI)
  4. Accelerate through mass spectrometer
  5. Produces **peptide mass fingerprint**

**Peptide mass fingerprint**
Proteomics – Mass Spectrometry

- **Identifying peptide mass fingerprint**
  - Compare with
    - Fingerprint for actual protein in database
    - Predicted fingerprint for predicted / hypothetical protein
      - Precompute predicted fingerprints for efficiency
  - May fail to distinguish
    - Post-translation modifications to protein

- **Protein databases / web servers (e.g., SWISS-2D PAGE)**
  - For each protein, record its
    - Protein pI, molecular weight, peptide mass fingerprint...
    - Experimentally determined location in 2D gel
  - Tools to computationally predict same information

**Identifying Peptide Mass Fingerprints**

- Enzymatic Cleavage
- MS (ESI-MS/MALDI)
- Mass Selected MS/MS
- Peptide fragment fingerprint - size alone
- Compare
- Protein ID
- MS/MS - actual peptide sequence

- Translate
- A Non-redundant Protein Sequence Database
- Computer Derived Peptide Fingerprint
- n peptides
- n computer-derived MS/MS spectra
- Compare each
Peptide Fragmentation

Proteomics – Mass Spectrometry

**Observations**
- Ionization methods (MALDI, ESI)
  - Developed in 1980’s, significant improvement
  - Combine with capillary electrophoresis for high throughput
- Very precise measure of mass / charge ratios (5 ppm)
- Extremely sensitive (can detect $< 10^{-18}$ moles of protein)

**Tandem mass spectrometry (MS-MS)**
- Feed output of mass spectrometer to 2nd mass spectrometer
- Increases precision of measurements
Proteomics

◆ Overview
  - 2D electrophoresis
  - Mass spectrometry
  - Protein arrays / microarrays

Proteomics – Protein Arrays / Microarrays

◆ Goal
  - High-throughput analysis of protein expression / interaction
  - Adapt approach similar to DNA microarrays
  - Improves on speed vs. 2D electrophoresis

◆ Approach
  - No equivalent of hybridization for proteins
  - Exploit other biochemical binding reactions
    ● Antibody–antigen
    ● Receptor–ligand
    ● DNA–protein...
Proteomics – Antibody / Antigen Binding

- **binding sites**

![](image)

Proteomics – Protein Arrays / Microarrays

- **Method**
  1. Place on glass slide many probes at known locations
     - Chemical probes
       - Ionic, hydrophobic, hydrophilic...
     - Biochemical probes
       - Antibody, receptor, DNA...
  2. Mix protein samples with probes, bind
  3. Wash off remaining proteins
  4. Collect & identify bound proteins with mass spectrometer
     - Surface Enhanced Laser Desorption / Ionization (SELDI)

- **Produces**
  - Protein expression profile
  - Protein interaction with probes
Protein Microarrays – Antibody Probes

- **Using antibody probes**
  - Antibodies recognize & bind to specific antigen
  - Many antibodies generated by shuffling DNA sequences
    - Over 100 million different antibodies
  - Monoclonal antibodies (mAb) can recognize single protein
    - Can even recognize post-translational modifications (phosphorylation, glycosylation, ubiquitination, etc...)

- **Monoclonal antibody protein microarrays**
  - Preliminary experiment at Stanford
  - 1/3 of 120 monoclonal antibodies bound to protein on array
  - Potential approach

---

Monoclonal Antibody Capture Microarray

6566 protein samples representing 5800 unique yeast proteins, using anti-GST antibody probes

[Zhu+ 2001]
Ciphergen Antibody Capture Protein Chip

1) Protein with antigen bound to antibody probe
2) Remainder of protein digested enzyme, leaving peptide antigen
3) Wash away protein fragments
4) SELDI laser ionizes & desorbs epitope binding peptide, sends to mass spectrometer

Protein Microarrays – Antigen Probes

- Using antigen probes
  - Create array with bound antigens
    - Proteins, peptides
  - Mix with serum from diseased & control patients
  - Then add fluorescent labeled antibodies
  - Measure intensity to detect antibody-antigen bindings

- Antigen microarrays
  - Measures immune responses in organism
  - Good for detecting / characterizing auto-immune diseases
    - Lupus, rheumatoid arthritis, multiple sclerosis
Proteomics – Protein Arrays / Microarrays

◆ Issues
  - Protein–probe interaction may not be one-to-one mapping
    ● Multiple proteins may bind to same site
  - Binding kinetics (strength of bond) differ for proteins
    ● Intensity may be due to kinetics, not protein concentration
  - Protein structure / function affected by contact with surface
    ● May not achieve bond as expected

◆ Scale
  - Currently less miniaturization than for DNA microarrays
  - Protein array
    ● 8-16 spots of 96-well spot chips = ~1500 probes
  - Protein microarray
    ● 10,000+ probes

Proteomics – Summary

◆ Proteomic techniques
  - Large variety of approaches
  - Attempt to directly measure
    ● Protein expression
    ● Protein interactions
  - Not at scale / precision of genomic techniques
  - Still in the process of improving