PROTEOMICS: STUDY OF PROTEINS

HLeeYu Jsuico Junsay
Department of Chemistry
School of Science and Engineering
Ateneo de Manila University
Why study proteins?

Proteins are mainly studied because of their ubiquitous role as the cell’s machinery and structural foundations.

How do we study proteins?
How do we study proteins?

• How do we isolate, purify, and analyze proteins?
  ▫ The short answer:
    • Based on their properties!!!
    • Size, charge, polarity, shape, reactivity

• How do we determine the structure of proteins?
  ▫ Primary structure: sequence?
  ▫ 3-D structure: secondary, tertiary, quaternary?

• How do we determine the function of proteins?
  ▫ Activity assays
The first step is usually **extraction**.

**SALTING OUT**

- Competition between protein solubility and salt solubility (What kinds of IFA are in play?)
- Different proteins precipitate at different salt concentrations
The first step is usually **extraction**.

- **DIALYSIS**
  - Removal of small molecule via a semipermeable membrane
Once a crude protein extract is obtained, one can purify the extract to its protein components by means of **chromatography**.

Chromatography is a separation technique based on differences in the rates at which components of a mixture are carried through a fixed or stationary phase (SP) by a gaseous or liquid mobile phase (MP).
Different chromatographic techniques exploit a protein’s different properties.

**Size exclusion chromatography** separates purifies proteins based on **Size. GEL PERMEATION** – larger molecules elute first!
Different chromatographic techniques exploit a protein’s different properties.

**Ion exchange chromatography (IEX)** separates proteins based on charge. Least attracted to column elutes first.

- **Load**
  - Positive charges
  - Neutral charges
  - Negative charges

- **Elute with NaCl**
  - Positive charges
  - Neutral charges
  - Negative charges

- **More NaCl**
  - Positive charges
  - Neutral charges
  - Negative charges

- **Result**
  - Pure positive charges
  - Pure negative charges
Different chromatographic techniques exploit a protein’s different properties.

**Ion exchange chromatography (IEX)** separates proteins based on charge. Least attracted to column elutes first

- **CM-cellulose columns**
  - For cationic proteins

- **DEAE-cellulose columns**
  - For anionic proteins
Different chromatographic techniques exploit a protein’s different properties.

**Affinity chromatography** uses a protein’s specific interaction with its substrate immobilized in the SP. Upon addition of free substrate, proteins are released.

Glucose-binding protein attaches to glucose residues (G) on beads.

Glucose-binding proteins are released on addition of glucose.
For faster separation, automation and computer interfacing, high performance (or high pressure) liquid chromatography (HPLC) may be used.
For faster separation, automation and computer interfacing, high performance (or high pressure) liquid chromatography (HPLC) may be used.
Electrophoresis may also be used to separate proteins. It is the migration of particles across a fluid/matrix due to an electric field.
A technique called **SDS-PAGE** is used to separate proteins by means of size.

Denaturating agents and Sodium dodecyl sulfate (SDS) stretch the proteins and gives them similar charge densities.

A technique called **SDS-PAGE** is used to separate proteins by means of size.

Denaturating agents and Sodium dodecyl sulfate (SDS) stretch the proteins and gives them similar charge densities.
A technique called **SDS-PAGE** is used to separate proteins by means of size.
Denatured proteins are loaded onto a polyacrylamide gel (PAG) and is exposed to an electric field and the proteins are allowed to migrate across the PAG (hence the name PAGE).
A technique called **SDS-PAGE** is used to separate proteins by means of size.
A technique called **SDS-PAGE** is used to separate proteins by means of size.

The Gel is stained to visualize the proteins
A technique called **SDS-PAGE** is used to separate proteins by means of size.

The Gel is stained to visualize the proteins

![Coomassie Blue stain = reversible](image)

![Silver stain](image)
EXAMPLES:

Pure hemoglobin gives you one band in the SDS-PAGE corresponding to 17 kDa. Size exclusion chromatography calibrated for molecular weight gives you 68 kDa for the same sample. What can you infer about the quaternary structure of hemoglobin?
Isoelectric focusing (IEF) is another electrophoresis technique based on isoelectric points of proteins.
SDS-PAGE and IEF may be combined to form a 2D gel. Protein separation is based on two properties: **pI and Size**

(A)
**SDS-PAGE** and **IEF** may be combined to form a **2D gel**. Protein separation is based on two properties: **pI** and **Size**.
EXAMPLES:

1. Calculate the pI and MW of glutathione, a tripeptide, Glu-Cys-Gly.
2. Calculate the pI and MW of C-H-E-M-I-S-T-R-Y
Amino acids are polyproFc acids.

<table>
<thead>
<tr>
<th>Name</th>
<th>pK₁</th>
<th>pK₂</th>
<th>pKᵣ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glycine</td>
<td>2.4</td>
<td>9.8</td>
<td></td>
</tr>
<tr>
<td>Alanine</td>
<td>2.3</td>
<td>9.7</td>
<td></td>
</tr>
<tr>
<td>Valine</td>
<td>2.3</td>
<td>9.6</td>
<td></td>
</tr>
<tr>
<td>Leucine</td>
<td>2.4</td>
<td>9.6</td>
<td></td>
</tr>
<tr>
<td>Isoleucine</td>
<td>2.4</td>
<td>9.7</td>
<td></td>
</tr>
<tr>
<td>Methionine</td>
<td>2.3</td>
<td>9.2</td>
<td></td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>1.8</td>
<td>9.1</td>
<td></td>
</tr>
<tr>
<td>Proline</td>
<td>2.0</td>
<td>10.6</td>
<td></td>
</tr>
<tr>
<td>Serine</td>
<td>2.1</td>
<td>9.2</td>
<td></td>
</tr>
<tr>
<td>Threonine</td>
<td>2.6</td>
<td>10.4</td>
<td></td>
</tr>
<tr>
<td>Cysteine</td>
<td>1.8</td>
<td>10.8</td>
<td>8.3</td>
</tr>
<tr>
<td>Asparagine</td>
<td>2.0</td>
<td>8.8</td>
<td></td>
</tr>
<tr>
<td>Glutamine</td>
<td>2.2</td>
<td>9.1</td>
<td></td>
</tr>
<tr>
<td>Tyrosine</td>
<td>2.2</td>
<td>9.1</td>
<td>10.9</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>2.4</td>
<td>9.4</td>
<td></td>
</tr>
<tr>
<td>Aspartate</td>
<td>2.0</td>
<td>10.0</td>
<td>3.9</td>
</tr>
<tr>
<td>Glutamate</td>
<td>2.2</td>
<td>9.7</td>
<td>4.3</td>
</tr>
<tr>
<td>Histidine</td>
<td>1.8</td>
<td>9.2</td>
<td>6.0</td>
</tr>
<tr>
<td>Lysine</td>
<td>2.2</td>
<td>9.2</td>
<td>10.8</td>
</tr>
<tr>
<td>Arginine</td>
<td>1.8</td>
<td>9.0</td>
<td>12.5</td>
</tr>
</tbody>
</table>
Primary structure may be elucidated using sequencing techniques.

1. Relative content of different AA
   a. Hydrolysis
   b. Chromatography
   c. Dyeing
2. Sequencing techniques
   a. EDMAN degradation/Dansylation
   b. Proteolytic cleavage via enzymes
Concentrated acids will destroy all amide bonds.
Free amino acids can react with derivatizing solutions which creates a derivative that has quantifiable properties.
Free amino acids can react with derivatizing solutions which creates a derivative that has quantifiable properties.

1-Fluoro-2,4-dinitrobenzene (Sanger’s reagent)

2,4-Dinitrophenylamino acid (yellow)
EXAMPLES:

You have an unknown hexapeptide. You did the following analyses:

1) Heat in 6 M HCl at 110° (24 hrs)
2) Perform Ion Exchange Chromatography
3) React each fraction (from IEC) with ninhydrin

Given the chromatogram on the next page, how many of each AA is present?
EXAMPLES:

ELUTION PROFILE OF PEPTIDE HYDROLYSATE

Asp
Thr
Ser
Glu
Pro
Gly
Ala
Cys
Val
Met
Ile
Leu
Tyr
Phe
Lys
His
NH₃
Arg

ELUTION PROFILE OF STANDARD AMINO ACIDS

pH 3.25
0.2 M Na citrate

pH 4.25
0.2 M Na citrate

pH 5.28
0.35 M Na citrate

Elution volume →
Primary structure may be elucidated using sequencing techniques.

1. Relative content of different AA  
   a. Hydrolysis  
   b. Chromatography  
   c. Dyeing  
2. Sequencing techniques  
   a. EDMAN degradation/Dansylation  
   b. Proteolytic cleavage via enzymes
EDMAN Degradation can give you the identity of the N-terminal AA
Dansyl chloride can give you a similar result as Edman.

Figure 3-9d Concepts in Biochemistry, 3/e © 2006 John Wiley & Sons
Enzymes that cleave proteins are called **proteases** or **peptidases**. They only cleave **SPECIFIC** amide bonds.

![Enzyme Diagram]

Figure 3-17a Concepts in Biochemistry, 3/e
© 2006 John Wiley & Sons

- **Trypsin cleavage site**
  - Arginine (Arg) or lysine (Lys)

- **Chymotrypsin cleavage site**
  - Phenylalanine (Phe), tyrosine (Tyr), or tryptophan (Trp)

Figure 3-17b Concepts in Biochemistry, 3/e
© 2006 John Wiley & Sons
Enzymes that cleave proteins are called **proteases** or **peptidases**. They only cleave SPECIFIC amide bonds.
Enzymes that cleave proteins are called **proteases** or **peptidases**. They only cleave SPECIFIC amide bonds

- Dansylation or Edman Degradation: N-terminus
- Carboxypeptidase: C-terminus except P
- CNBr: After Met
- Trypsin: After RK (basic except H), not if P is next to RK
- Chymotrypsin: After WYF (aromatic), not if P is next to WYF
- V-8 Protease: After DE (acidic), not if P is next to DE
EXAMPLES:

• Given the following peptides after Tryptic and chymotryptic cleavage, identify the original peptide sequence.

<table>
<thead>
<tr>
<th>Tryptic peptides</th>
<th>Chymotryptic peptide</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala—Ala—Trp—Gly—Lys</td>
<td>Val—Lys—Ala—Ala—Trp</td>
</tr>
<tr>
<td>Thr—Phe—Val—Lys</td>
<td></td>
</tr>
</tbody>
</table>
EXAMPLES:

A mutant form of polypeptide hormone angiotensin II has the amino acid composition \((D, R, I, M, F, P, Y, V)\)

The following observations are made:

1. Trypsin yields a dipeptide containing D and R, and a hexapeptide
2. Cyanogen bromide cleavage yields a dipeptide containing F and P, and a hexapeptide
3. Chymotrypsin cleaves the hormone into two tetrapeptides of composition:
   \((D, R, Y, V)\) and \((I, M, F, P)\)
4. The dipeptide of composition \((P, F)\) cannot be cleaved by carboxypeptidase.
EXAMPLES:

A pentapeptide contains K, W, D, F, and E. Treatment with trypsin yields a free AA and a tetrapeptide. Treatment with chymotrypsin yields a dipeptide and a tripeptide that gives free tryptophan upon carboxypeptidase A treatment. Treatment with V8 protease yields a free AA and two pairs of dipeptides. What is the peptide sequence?
EXAMPLES:

Solve the sequence of a nonapeptide whose amino acid composition is E, F, G, K, R, T, V(2), Y.

A. Dansylation gave dansyl-E, while T was released first by carboxypeptidase.

B. Trypsin cleavage of the original peptide gave 3 peptides T-1 (a tripeptide), T-2 (a dipeptide), and T-3 (a tetrapeptide). T-1 had amino acid composition of E, R, Y; while T-2 had composition K, V. The N-terminal residue of T-3 was F.

C. Chymotrypsin cleavage of the original peptide gave 3 peptides C-1 (a tripeptide), C-2 (a dipeptide), and C-3 (a tetrapeptide). C-1 had composition G, T, V; and its N-terminal residue was G. C-2 had composition E, Y; while the N-terminal residue of C-3 was R.
Mass Spectrometry is gaining popularity in protein sequencing.

MALDI-TOF MS – Matrix Assisted Laser Desorption Ionization – Time of Flight Mass Spectrometry

ESI-MS – Electrospray Ionization Mass Spectrometry

FTICR-MS – Fourier Transform Ion Cyclotron Resonance Mass Spectrometry
Mass Spectrometry is gaining popularity in protein sequencing.
2°, 3° and 4° structures may be elucidated using X-ray crystallography and Nuclear Magnetic Resonance

X-ray Crystallography

- Looks at how the electrons of the protein in crystal diffract X-rays
- The main experimental result is an electron density map, which can be fitted with a model of the protein
- Useful for molecules of just about any size, as long as they can be properly crystallized
2°, 3° and 4° structures may be elucidated using **X-ray crystallography** and **Nuclear Magnetic Resonance**

**NMR**
- Nuclear Magnetic Resonance Spectroscopy
- Reveals the structure of the protein in solution with up to 100 a.a. residues
- **NOESY**
  - Nuclear Overhauser enhancement spectroscopy
  - Graphically displays pairs of protons that are in close proximity, even if they are not close together in the primary structure (5 Å apart)
2°, 3° and 4° structures may be elucidated using X-ray crystallography and Nuclear Magnetic Resonance.
Protein of interest usually come with specific functions, thus to screen for proteins, we can look at **functional assays**.

Measure the absorption of the sample in a given period of time
Another screening process involves the use of **antibodies**. These are proteins that bind to specific molecules!

1. Fab (Fragment, antigen-binding)
2. Fc (Fragment, crystallizable)
3. heavy chain (consist of VH, CH1, hinge, CH2 and CH3 regions: from N-term)
4. light chain (consist of VL and CL regions: from N-term)
5. antigen binding site
6. hinge regions
Another screening process involves the use of **antibodies**. These are proteins that bind to specific molecules!
Another screening process involves the use of **antibodies**. These are proteins that bind to specific molecules!

- **MONOCLONAL ANTIBODIES**
  - Identical antibodies produced by one type of immune cell (a clone line) recognizing a single epitope

- **POLYCLONAL ANTIBODIES**
  - A mixture of antibodies secreted against a specific antigen, each recognizing a different epitope
Antibodies may be used as a probe when conjugated to another molecule that serves as a signal transducer.
ELISA (enzyme-linked immunosorbent assay)

- Enzyme reacts with colorless substrate to produce a colored product
- Enzyme is covalently bound to a specific antibody that recognizes a target antigen
ELISA (enzyme-linked immunosorbent assay)
HOMEWORK:

• ELISA: Virtual Lab
  ▫ Visit: http://www.hhmi.org/biointeractive/vlabs
  ▫ CLICK on “Immunology Lab”
  ▫ Perform the virtual experiment correctly and answer the worksheet (Shockwave flash needed)
Just added info!

• See virtual lab from MIT:
  • Day 1: Isolation of beta-galactosidase from E. coli
    –  [http://mit.edu/7.02/virtual_lab/PBC/PBC1virtuallab.html](http://mit.edu/7.02/virtual_lab/PBC/PBC1virtuallab.html)
  • Day 2: Desalting and Qualitative Activity Assay
    –  [http://mit.edu/7.02/virtual_lab/PBC/PBC2virtuallab.html](http://mit.edu/7.02/virtual_lab/PBC/PBC2virtuallab.html)
  • Day 4A: SDS-PAGE running
    –  [http://mit.edu/7.02/virtual_lab/PBC/PBC4Avirtuallab.html](http://mit.edu/7.02/virtual_lab/PBC/PBC4Avirtuallab.html)
  • Day 5: Western Blot
    –  [http://mit.edu/7.02/virtual_lab/PBC/PBC5virtuallab.html](http://mit.edu/7.02/virtual_lab/PBC/PBC5virtuallab.html)
Bioinformatics Resources for Proteins

- NCBI: www.pubmed.org
- PDB: www.pdb.org
- EXPASY (SwissProt/UniProt): = www.expasy.org
- Viewers
  - Online: CHIME (www.mdl.com)
  - Rasmol (http://www.umass.edu/microbio/rasmol/index2.htm)
  - DeepView (www.expasy.org)
SPECIAL TOPICS IN PROTEINS
We will not discuss special topics... (due to time constraints)

• NO CLASSES on MONDAY, May 3, 2010
• Special topics include
  – How healthy is vegan?
  – Glutathione: THE antioxidant
  – Homology: how are humans different from animals and other organisms at the protein level?
  – When good proteins go bad: Sickle-cell anemia and Prions
• YOU MAY USE THESE AS TOPICS FOR YOUR MAGAZINE PROJECT