BME 2901-BIOCHEMISTRY

Protein Isolation and Purification

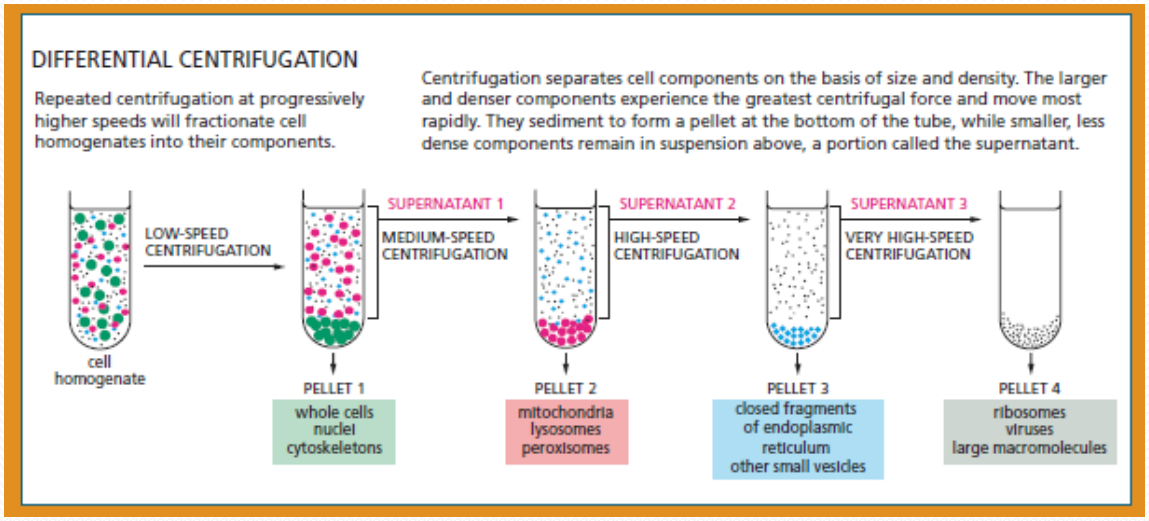
by Assist. Prof. Görke Gürel Peközer

Yıldız Technical University
Biomedical Engineering Department
Fall 2019

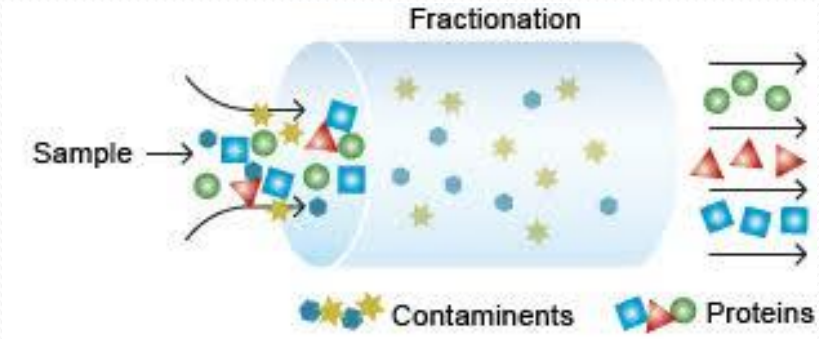
Protein Isolation and Purification

- **Protein purification** is a series of processes intended to isolate one or a few proteins from a complex mixture, usually cells, tissues or whole organisms.
- Protein purification is vital for the characterization of the function, structure and interactions of the protein of interest.
- A pure preparation is essential before a protein's properties and activities can be determined.
- Methods for separating proteins take advantage of properties that vary from one protein to the next, including size, charge, and binding properties.

- The source of a protein is generally tissue or microbial cells.
- If the protein of interest in intracellular, the first step in protein purification procedure is to break open these cells, releasing their proteins into a solution called a **crude extract.**
- If necessary, **differential centrifugation** can be used to prepare subcellular fractions or to isolate specific organelles.
- If the protein of interest is extracellular there is no need for cell disruption. Instead secreted proteins can be collected in the supernatant after centrifugation. Membrane proteins can be released from the lipid bilayer using detergents.

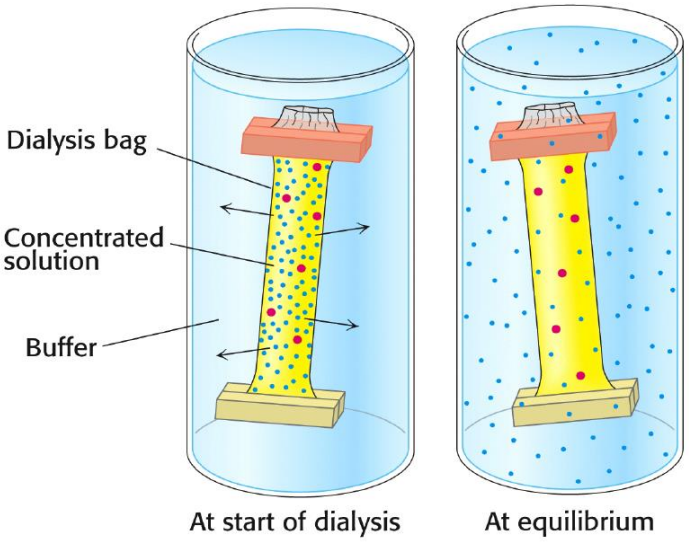


- Once the extract or organelle preparation is ready, various methods are available for purifying one or more of the proteins it contains.
- Commonly, the extract is subjected to treatments that separate the proteins into different fractions based on a property such as size or charge, a process referred to as **fractionation.**
- Early fractionation steps in a purification utilize differences in protein solubility, which is a complex function of pH, temperature, salt concentration, and other factors.
- The solubility of proteins is generally lowered at high salt concentrations, an effect called "**salting out**."
- The addition of a salt in the right amount can selectively precipitate some proteins, while others remain in solution.

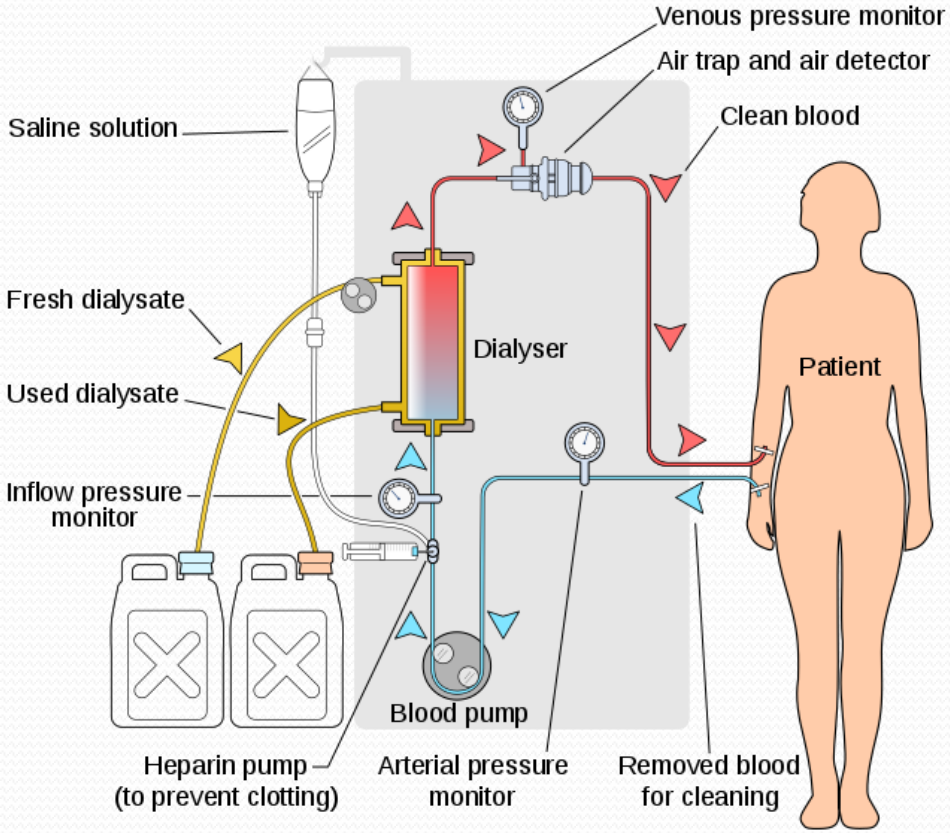


Dialysis

- A solution containing the protein of interest often must be further altered before subsequent purification steps are possible.
- **Dialysis** is a procedure that separates proteins from solvents by taking advantage of the proteins' larger size.
- The partially purified extract is placed in a bag or tube made of a semipermeable membrane.
- When this is suspended in a much larger volume of buffered solution of appropriate ionic strength, the membrane allows the exchange of salt and buffer but not proteins.
- Thus dialysis retains large proteins within the membranous bag or tube while allowing the concentration of other solutes in the protein preparation to change until they come into equilibrium with the solution outside the membrane.



Hemodialysis



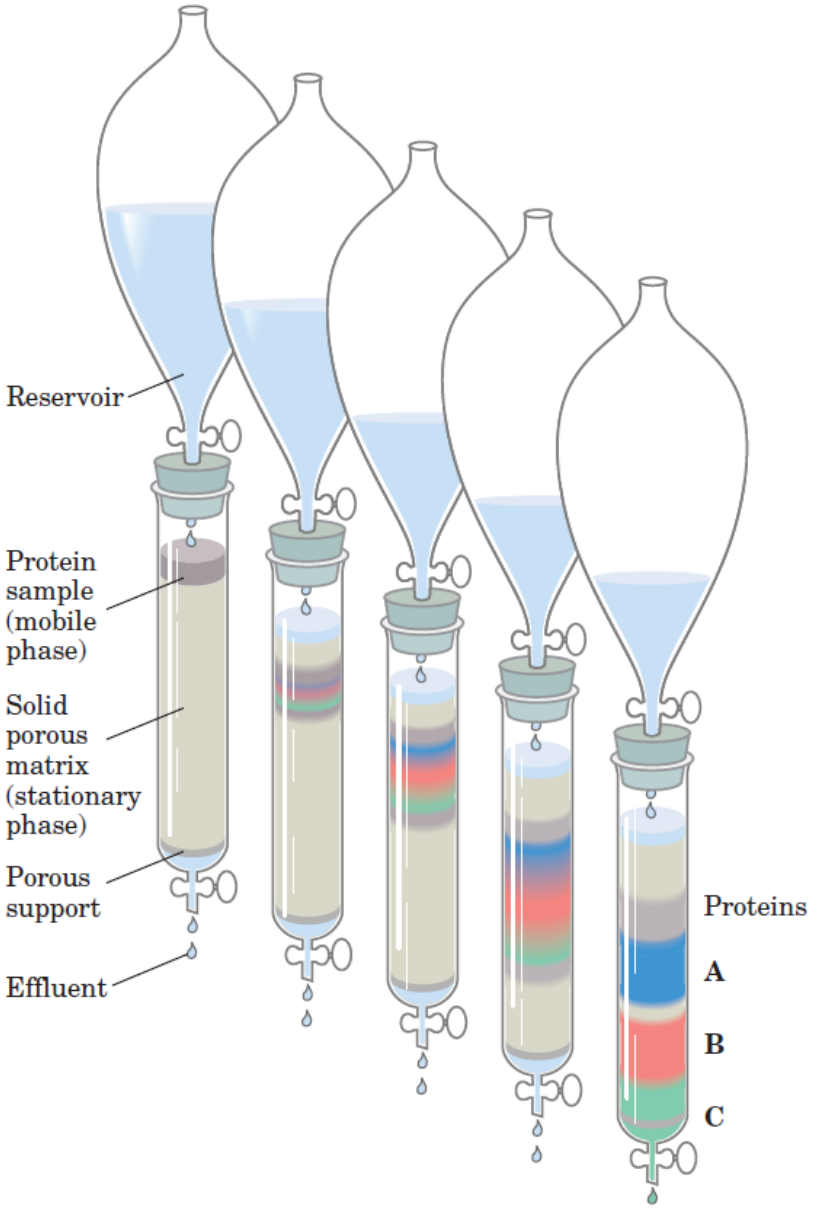
Chromatography

- Chromatography is a method by which a mixture is separated by distributing its components between two phases.
- **The stationary phase** remains fixed in place while **the mobile phase** carries the components of the mixture through the medium being used. The stationary phase acts as a constraint on many of the components in a mixture, slowing them down to move slower than the mobile phase.
- The movement of the components in the mobile phase is controlled by the significance of their interactions with the mobile and/or stationary phases.
- Because of the differences in factors such as the solubility of certain components in the mobile phase and the strength of their affinities for the stationary phase, some components will move faster than others, thus facilitating the separation of the components within that mixture.
- Chromatography steps separate the individual components of a complex mixture into *fractions* based on the properties of the protein—such as size, shape, or electrical charge.

Column Chromatography

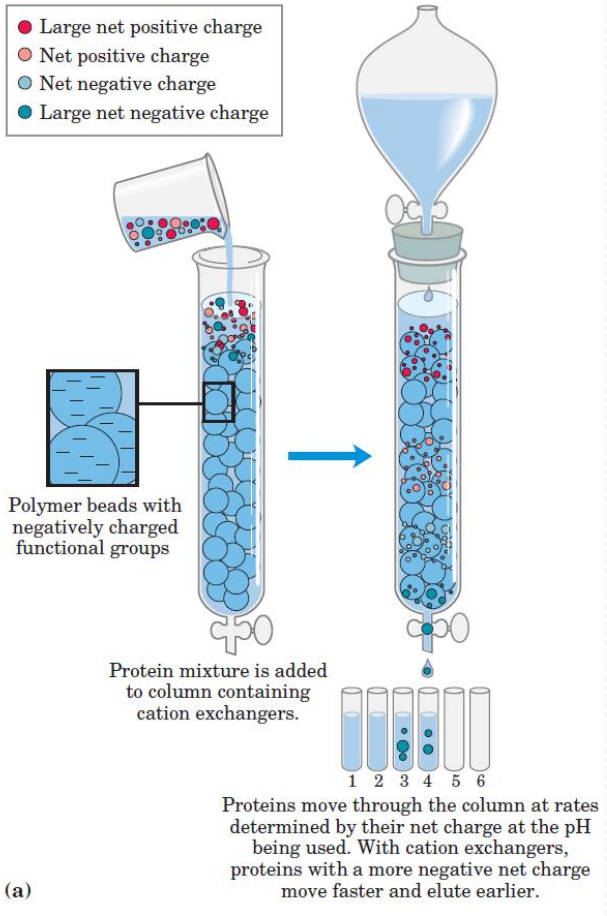
- Chromatography is a method by which a mixture is separated by distributing its components between two phases.
- The most powerful methods for fractionating proteins make use of **column chromatography.**
- A porous solid material with appropriate chemical properties (**the stationary phase**) is held in a column, while buffered solution (**the mobile phase**) carries the molecule of interest through it.

- The stationary phase acts as a constraint on many of the components in a mixture, slowing them down to move slower than the mobile phase.
- The movement of the components in the mobile phase is controlled by the significance of their interactions with the mobile and/or stationary phases.
- Because of the differences in factors such as the solubility of certain components in the mobile phase and the strength of their affinities for the stationary phase, some components will move faster than others, thus facilitating the separation of the components within that mixture.
- Chromatography steps separate the individual components of a complex mixture into *fractions* based on the properties of the protein—such as size, shape, affinity or electrical charge.
- Individual proteins migrate faster or more slowly through the column depending on their properties.



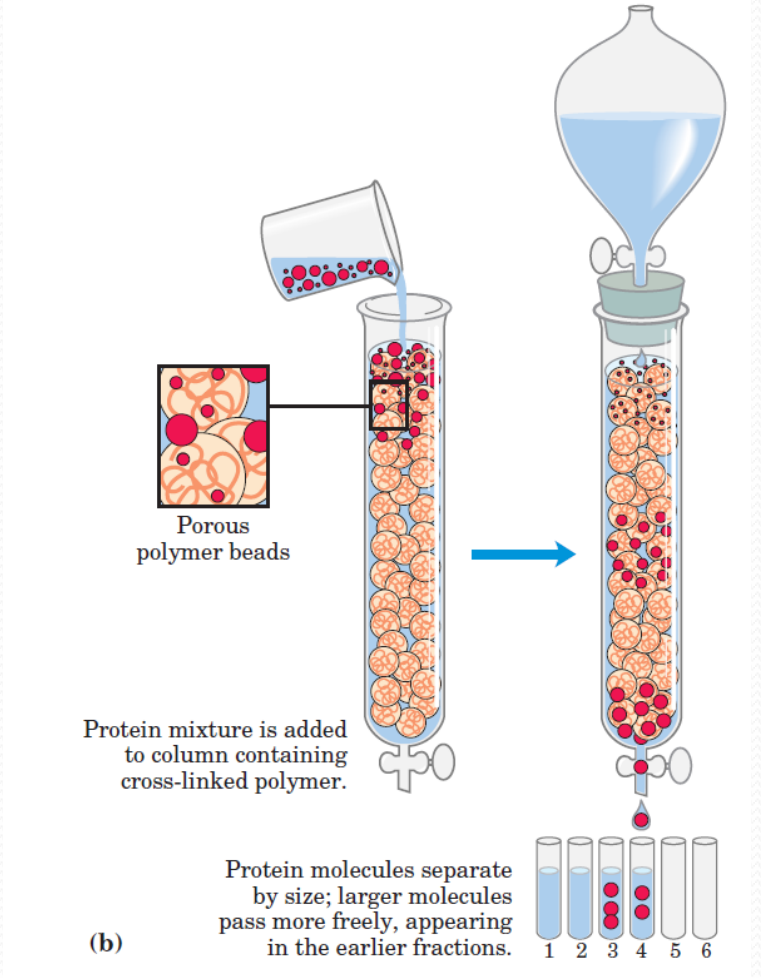
Ion-exchange Chromatography

- In **ion-exchange chromatography**, the solid matrix has either positively or negatively charged groups.
- If the stationary phase is negatively charged, in the mobile phase, proteins with a net positive charge migrate through the matrix more slowly than those with a net negative charge, because the migration of the former is retarded more by interaction with the stationary phase.
- The two types of protein can separate into two distinct bands. The expansion of the protein band in the mobile phase (the protein solution) is caused both by separation of proteins with different properties and by diffusional spreading.
- As the length of the column increases, the resolution of two types of protein with different net charges generally improves.
- However, the rate at which the protein solution can flow through the column usually decreases with column length. And as the length of time spent on the column increases, the resolution can decline as a result of diffusional spreading within each protein band.



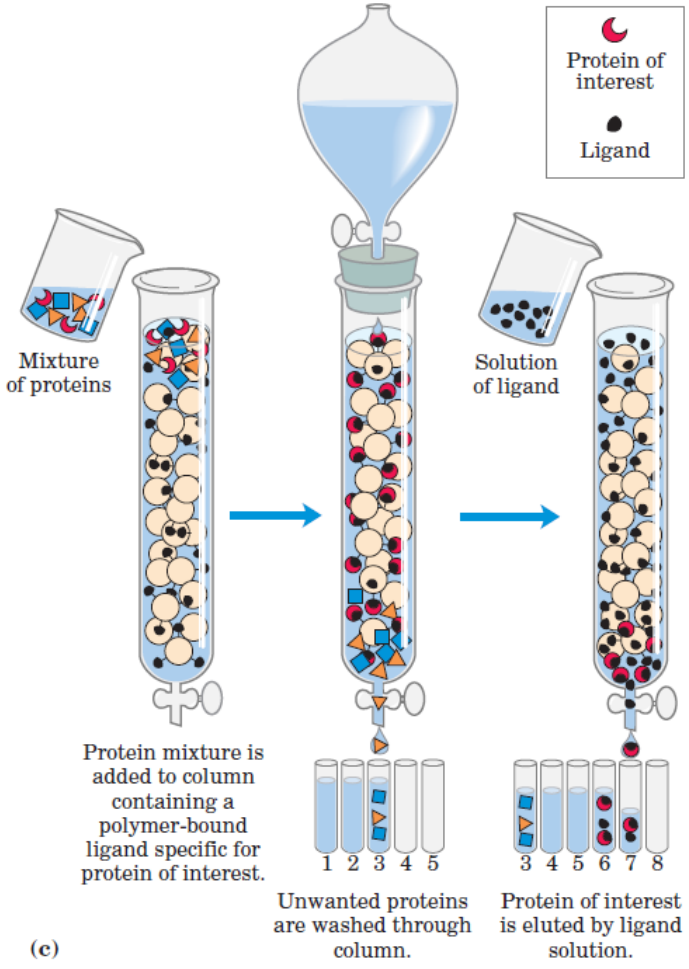
Size Exclusion Chromatography

- **Size exclusion chromatography** separates proteins according to size.
- In this method, large proteins emerge from the column sooner than small ones.
- The solid phase consists of beads with engineered pores or cavities of a particular size.
- Large proteins cannot enter the cavities, and so take a short (and rapid) path through the column, around the beads.
- Small proteins enter the cavities, and migrate through the column more slowly causing their retardation.
- The smaller the protein, the later it elutes from the column.



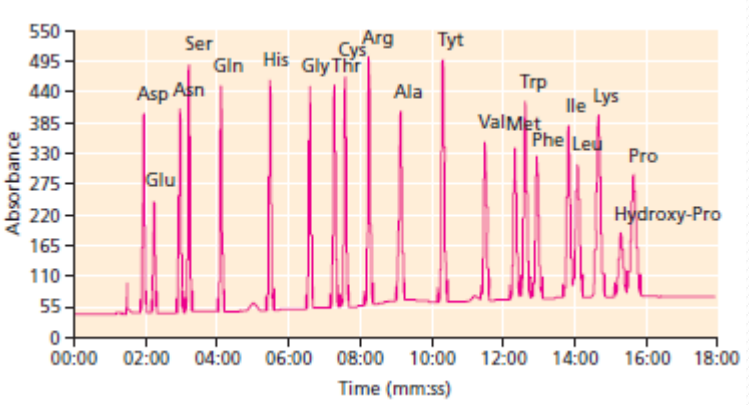
Affinity Chromatography

- **Affinity chromatography** is based on the binding affinity of a protein.
- The beads in the column have a covalently attached chemical group.
- A protein with affinity for this particular chemical group will bind to the beads in the column, and its migration will be retarded as a result.



HPLC

- A modern refinement in chromatographic methods is **HPLC**, or **high-performance liquid chromatography.**
- HPLC makes use of high-pressure pumps that speed the movement of the protein molecules down the column, as well as higher-quality chromatographic materials that can withstand the crushing force of the pressurized flow.
- By reducing the transit time on the column, HPLC can limit diffusional spreading of protein bands and thus greatly improve resolution.



Which method to pick?

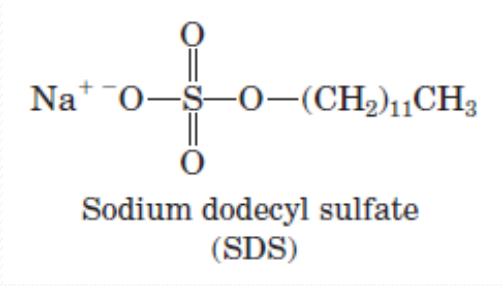
- The approach to purification of a protein that has not previously been isolated is guided both by established precedents and by common sense.
- In most cases, several different methods must be used sequentially to purify a protein completely.
- The choice of method is somewhat empirical, and many protocols may be tried before the most effective one is found.
- Trial and error can often be minimized by basing the procedure on purification techniques developed for similar proteins.
- Published purification protocols are available for many thousands of proteins. Common sense dictates that inexpensive procedures such as salting out be used first, when the total volume and the number of contaminants are greatest.
- Chromatographic methods are often impractical at early stages, because the amount of chromatographic medium needed increases with sample size.
- As each purification step is completed, the sample size generally becomes smaller, making it feasible to use more sophisticated (and expensive) chromatographic procedures at later stages.

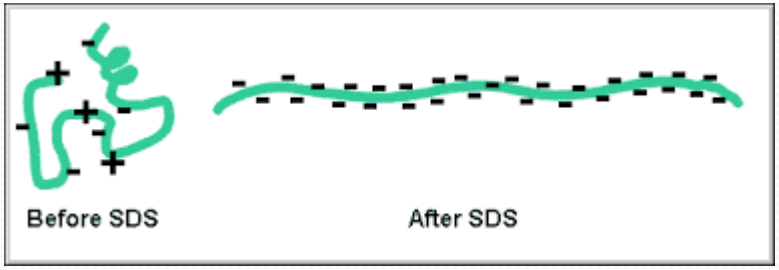
Electrophoresis

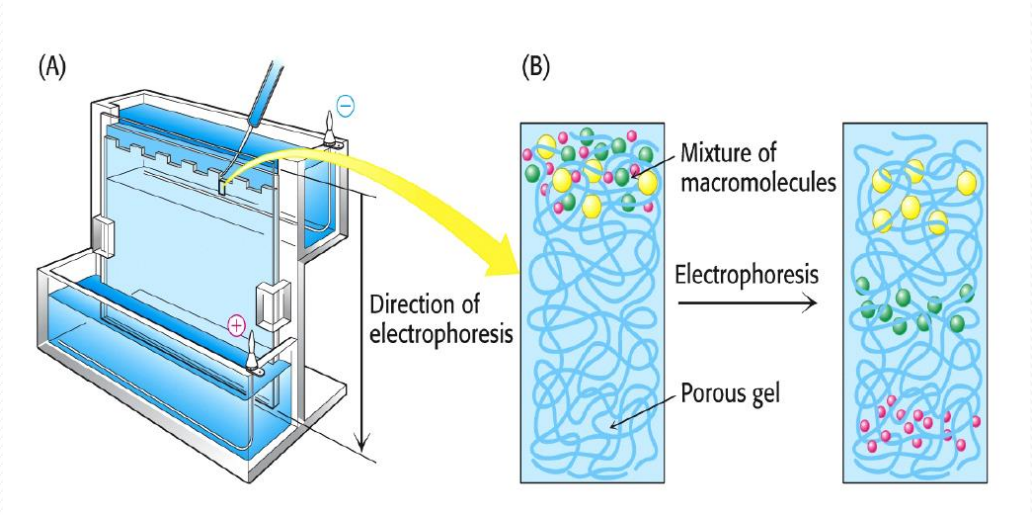
- Another important technique for the **separation** of proteins is based on the migration of charged proteins in an electric field, a process called **electrophoresis.**
- It is **not used to purify** proteins in large amounts, because simpler alternatives are usually available and electrophoretic methods often adversely affect the structure and thus the function of proteins.
- Its advantage is that proteins can be visualized as well as separated, permitting a researcher to estimate quickly the number of different proteins in a mixture or the degree of purity of a particular protein preparation.
- It also allows determination of crucial properties of a protein such as its isoelectric point and approximate molecular weight.

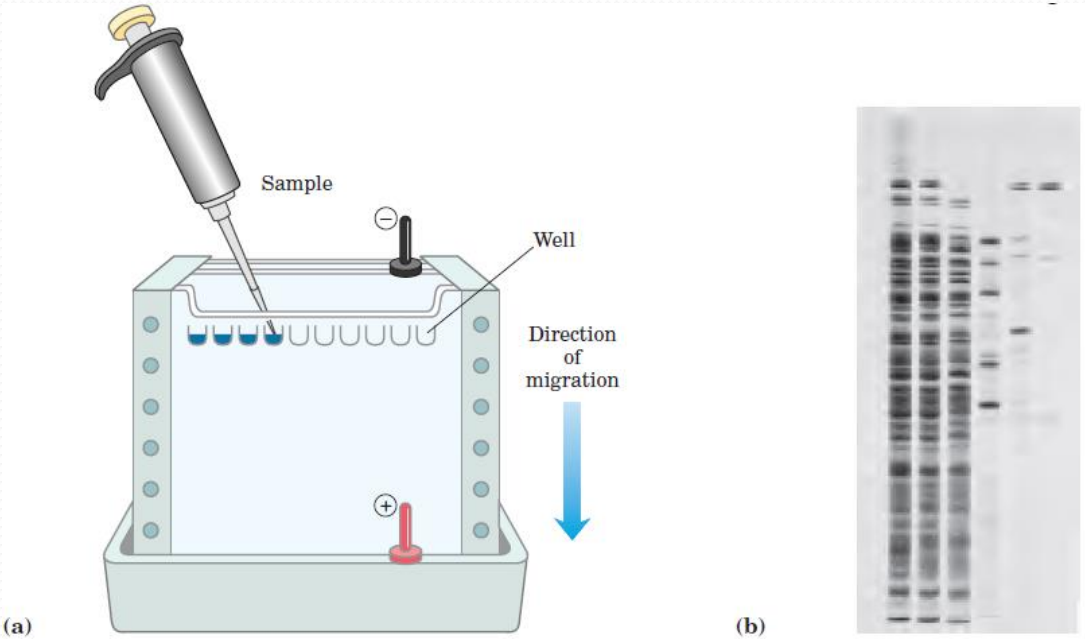
- Electrophoresis of proteins is generally carried out in gels made up of the cross-linked polymer polyacrylamide.
- The polyacrylamide gel acts as a molecular sieve, slowing the migration of proteins approximately in proportion to their charge-to-mass ratio.
- Migration may also be affected by protein shape.

- The migration of a protein in a gel during electrophoresis is therefore a function of its size and its shape.
- An electrophoretic method commonly employed for estimation of purity and molecular weight makes use of the detergent **sodium dodecyl sulfate (SDS).**
- SDS binds to most proteins in amounts roughly proportional to the molecular weight of the protein, about one molecule of SDS for every two amino acid residues. The bound SDS contributes a large net negative charge, rendering the intrinsic charge of the protein insignificant and conferring o each protein a similar charge-to-mass ratio.
- In addition, the native conformation of a protein is altered when SDS is bound, and most proteins assume a similar shape. Electrophoresis in the presence of SDS therefore separates proteins almost exclusively on the basis of mass (molecular weight), with smaller polypeptides migrating more rapidly.

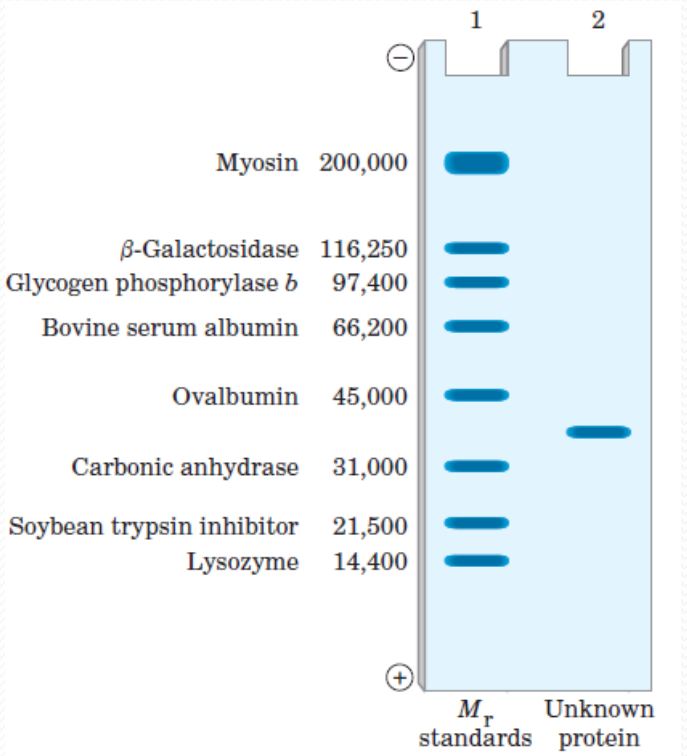





(A)
(B)
Direction of electrophoresis
Mixture of macromolecules
Electrophoresis
Porous gel

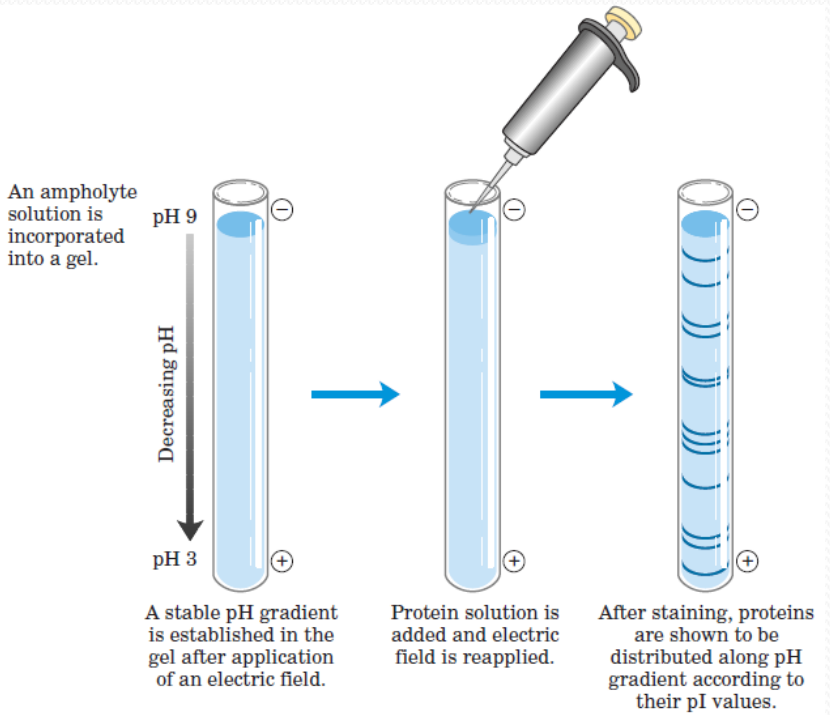

Sample
Well
Direction of migration
(a)
(b)

- After electrophoresis, the proteins are visualized by adding a dye such as Coomassie blue, which binds to proteins but not to the gel itself.
- Thus, a researcher can monitor the progress of a protein purification procedure as the number of protein bands visible on the gel decreases after each new fractionation step.
- When compared with the positions to which proteins of known molecular weight migrate in the gel, the position of an unidentified protein can provide an excellent measure of its molecular weight.
- If the protein has two or more different subunits, the subunits will generally be separated by the SDS treatment and a separate band will appear for each.

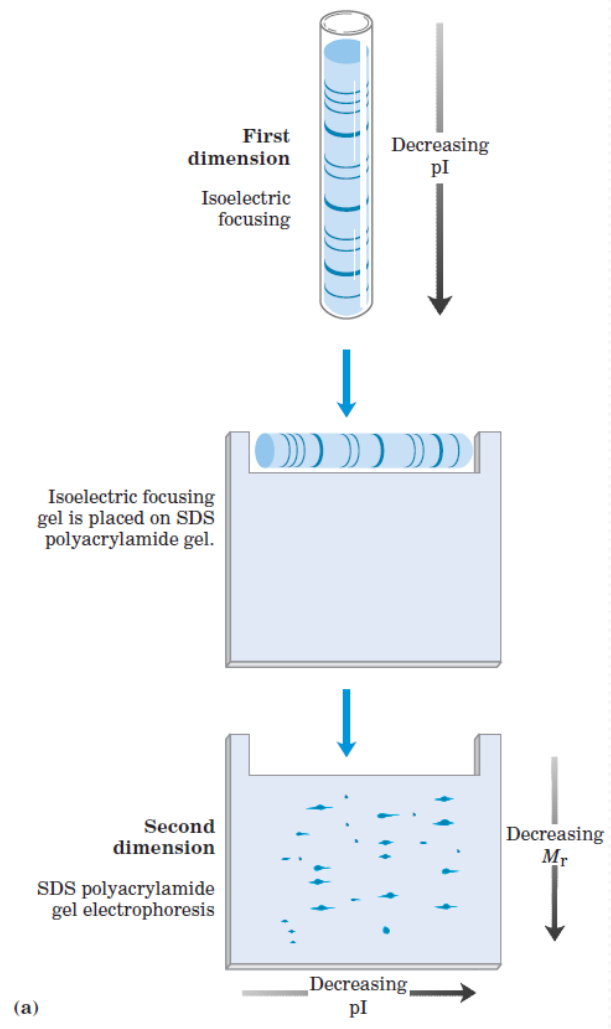


Isoelectric focusing

- **Isoelectric focusing** is a procedure used to determine
- the isoelectric point (pI) of a protein.
- A pH gradient is established by allowing a mixture of low molecular weight organic acids and bases (ampholytes) to distribute themselves in an electric field generated across the gel.
- When a protein mixture is applied, each protein migrates until it reaches the pH that matches its pI.
- Proteins with different isoelectric points are thus distributed differently throughout the gel.

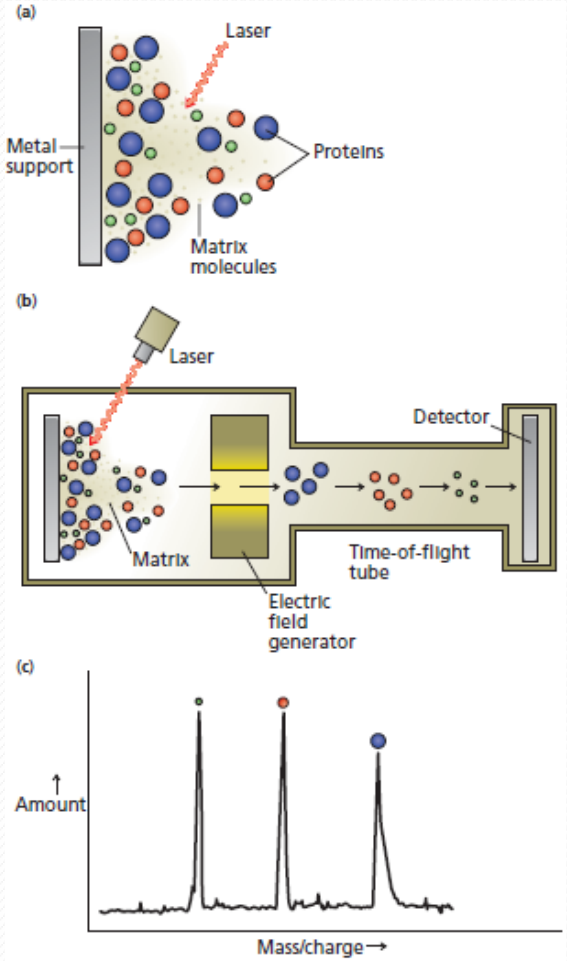


- Combining isoelectric focusing and SDS electrophoresis sequentially in a process called **two-dimensional electrophoresis** permits the resolution of complex mixtures of proteins.
- This is a more sensitive analytical method than either electrophoretic method alone.
- Two-dimensional electrophoresis separates proteins of identical molecular weight that differ in pI, or proteins with similar pI values but different molecular weights.



Mass Spectrometry

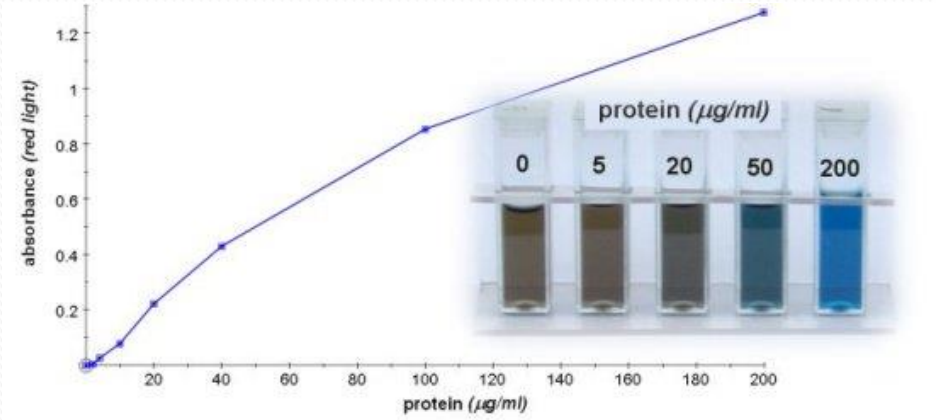
- **Mass spectrometry**, is a technique that determines the mass of a molecule. The most basic type of mass spectrometer measures the time that it takes for a charged gas phase molecule to travel from the point of injection to a sensitive detector.
- This time depends on the charge of a molecule and its mass and the result is reported as the mass/charge ratio.
- The technique has been used in chemistry for almost 100 years but its application to proteins was limited because, until recently, it was not possible to disperse charged protein molecules into a gaseous stream of particles.
- This problem was solved in the late 1980s with the development of two new types of mass spectrometry. In **electrospray mass spectrometry** the protein solution is pumped through a metal needle at high voltage to create tiny droplets. The liquid rapidly evaporates in a vacuum and the charged proteins are focused on a detector by a magnetic field.
- The second new technique is called **matrix-assisted laser desorption ionization (MALDI)**. In this method, proteins are placed in a matrix which is matrix is a small organic molecule that absorbs light at a particular wavelength. With a short pulse of laser light, the proteins are ionized and then desorbed from the matrix into the vacuum system. They fly directed to the detector. When time-of-flight (TOF) is measured, the technique is called MALDI–TOF.



Mass spectrometry is very sensitive and highly accurate. Often the mass of a protein can be obtained from picomole (10-12 mol) quantities that are isolated from an SDS–PAGE gel. The correct mass can be determined with an accuracy of less than the mass of a single proton.

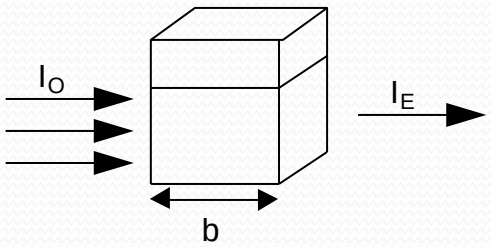
Protein Quantification

- Total protein amount in a solution can be identified using Bradford method.
- The principle of this assay is that the binding of protein molecules to Coomassie Blue dye under acidic conditions results in a color change from brown to blue.
- This method actually measures the presence of the basic amino acid residues, arginine, lysine and histidine, which contributes to formation of the protein-dye complex.
- The assay is based on the observation that the absorbance maximum for an acidic solution of Coomassie Blue shifts from 465 nm to 595 nm when binding to protein occurs.



Beer-Lambert Law

- Total protein concentration in the solution can then be calculated using the absorbance values obtained by the Bradford assay in Beer-Lambert equation.
- This law states that "the proportion of light absorbed by a medium is independent of the intensity of incident light" and "the absorbance of light is directly proportional to the concentration of the absorbing medium and the thickness or path length of the medium".



I_O = Intensity of incident light
I_E= Intensity of exiting light
b = path length of sample

$$T = I_o / I_E$$

Where:
T: Transmittance

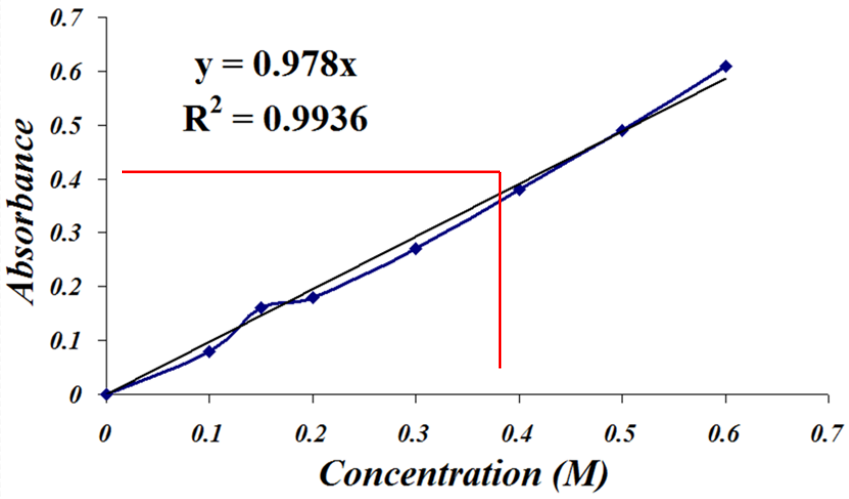
Beer-Lambert Equation

$$A = \varepsilon \times b \times c$$

Where:
A: absorbance
Ɛ : molar absorptivity (molar extinction coefficient)
b: length of the path
c: concentration

Beer-Lambert Law

- Thus, according to Beer-Lambert's Law absorbance of a solution is proportional to its concentration and absorbance vs. concentration is a straight line.
- A calibration curve is constructed using known concentrations of the sample from which an unknown concentration can be derived using its absorbance value.



Protein Quantification

- To purify a protein, it is essential to have a way of detecting and quantifying that protein in the presence of many other proteins at each stage of the procedure.
- Often, purification must proceed in the absence of any information about the size and physical properties of the protein or about the fraction of the total protein mass it represents in the extract.
- For proteins that are enzymes, the amount in a given solution or tissue extract can be measured, or assayed, in terms of the catalytic effect the enzyme produces—that is, the *increase* in the rate at which its substrate is converted to reaction products when the enzyme is present.
- For this purpose one must know (1) the overall equation of the reaction catalyzed, (2) an analytical procedure for determining the disappearance of the substrate or the appearance of a reaction product, (3) whether the enzyme requires cofactors such as metal ions or coenzymes, (4) the dependence of the enzyme activity on substrate concentration, (5) the optimum pH, and (6) a temperature zone in which the enzyme is stable and has high activity. Enzymes are usually assayed at their optimum pH and at some convenient temperature within the range 25 to 38 °C.
- Also, very high substrate concentrations are generally used so that the initial reaction rate, measured experimentally, is proportional to enzyme concentration.

Activity-Specific Activity

- 1.0 unit of enzyme activity is defined as the amount of enzyme causing transformation of 1.0 µmol of substrate per minute at 25 °C under optimal conditions of measurement.
- The term **activity** refers to the total units of enzyme in a solution.
- The **specific activity** is the number of enzyme units per milligram of total protein. The specific activity is a measure of enzyme purity: it increases during purification of an enzyme and becomes maximal and constant when the enzyme is pure.

Specific Activity = Activity / Total Protein Amount

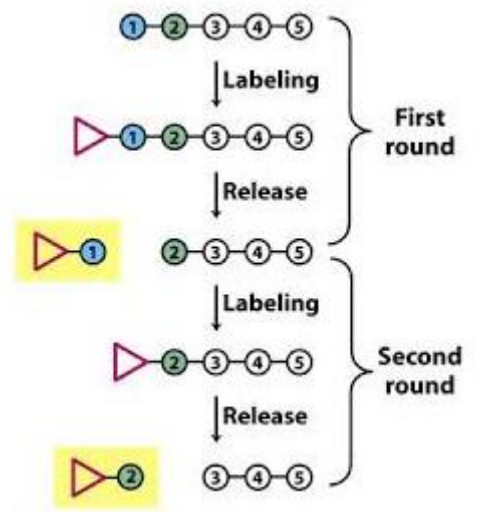
- After each purification step, the activity of the preparation (in units of enzyme activity) is assayed, the total amount of protein is determined independently, and the ratio of the two gives the specific activity.
- A protein is generally considered pure when further purification steps fail to increase specific activity and when only a single protein species can be detected (for example, by electrophoresis).

Sequence-Structure-Function Relationship

- Each type of protein has a unique amino acid sequence.
- The amino acid sequence plays a fundamental role in determining the three-dimensional structure of the protein, and ultimately its function.
- How do we know? One-third of proteins that were found to be defective in genetic diseases are defective because of a single change in their amino acid sequence; hence, if the primary structure is altered, the function of the protein may also be changed.
- However, An estimated 20% to 30% of the proteins in humans are **polymorphic**, having amino acid sequence variants in the human population. Many of these variations in sequence have little or no effect on the function of the protein.
- Also, the amino acid sequence in some regions of the primary structure might vary considerably without affecting biological function, but most proteins contain crucial regions that are essential to their function and whose sequence is therefore conserved

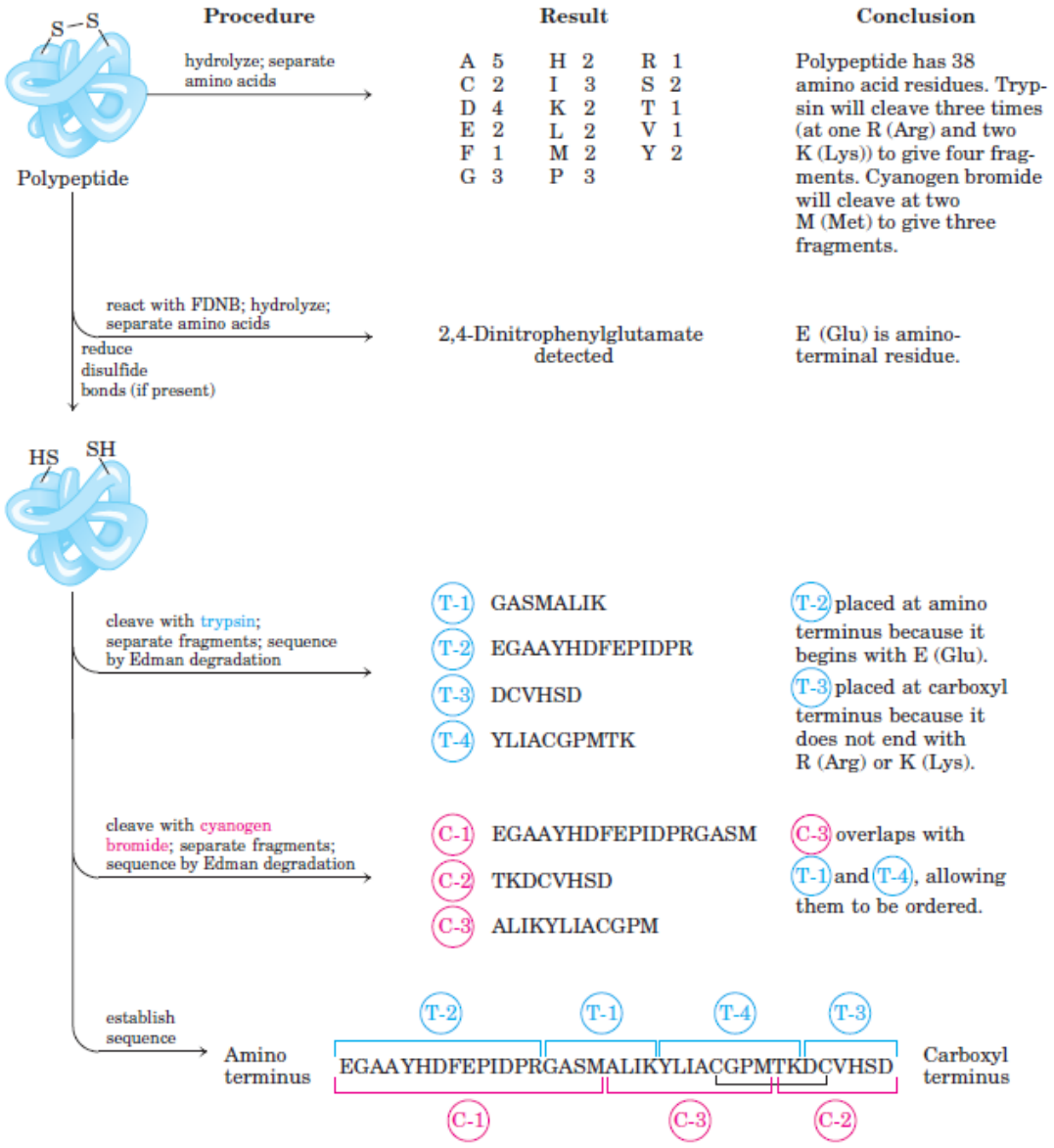
Protein Sequencing

- Amino acid sequence of the protein, thus, gives informations about 3D structure and function of the protein as well as its cellular location, and evolution.
- To sequence an entire polypeptide, a chemical method devised by Pehr Edman is usually employed.
- The **Edman degradation** procedure labels and removes only the amino-terminal residue from a peptide, leaving all other peptide bonds intact.
- After removal and identification of the aminoterminal residue, the *new* amino-terminal residue exposed can be labeled, removed, and identified through the same series of reactions. This procedure is repeated until the entire sequence is determined.



Sequencing Large Proteins

- The overall accuracy of amino acid sequencing generally declines as the length of the polypeptide increases. The very large polypeptides found in proteins must be broken down into smaller pieces to be sequenced efficiently.
- Steps for sequencing large protein:
    1. Cleave S-S bonds
    2. Separate subunits
    3. Determine N-terminus of protein
    4. Determine amino acid composition
    5. Use cleavage agents to digest protein into smaller fragments
    6. Amino acid composition and sequence of fragments
    7. Use overlapping fragments to get full sequence

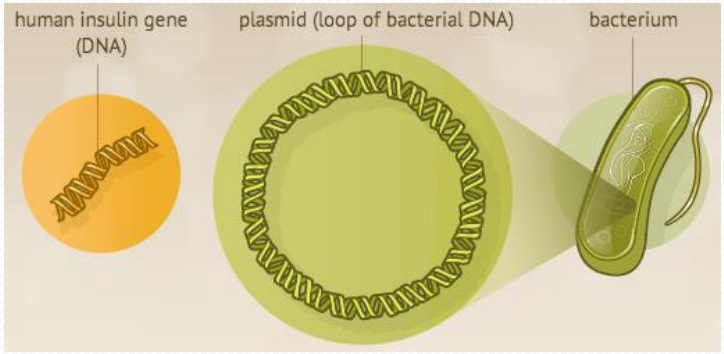

Procedure
Result
Conclusion
S—S
hydrolyze; separate amino acids
A 5 H 2 R 1
C 2 I 3 S 2
D 4 K 2 T 1
E 2 L 2 V 1
F 1 M 2 Y 2
G 3 P 3
Polypeptide has 38 amino acid residues. Trypsin will cleave three times (at one R (Arg) and two K (Lys)) to give four fragments. Cyanogen bromide will cleave at two M (Met) to give three fragments.
Polypeptide
react with FDNB; hydrolyze; separate amino acids
2,4-Dinitrophenylglutamate detected
E (Glu) is amino-terminal residue.
reduce disulfide bonds (if present)
HS SH
cleave with trypsin; separate fragments; sequence by Edman degradation
T-1 GASMALIK
T-2 EGAAYHDFEPIDPR
T-3 DCVHSD
T-4 YLIACGPMTK
T-2 placed at amino terminus because it begins with E (Glu).
T-3 placed at carboxyl terminus because it does not end with R (Arg) or K (Lys).
cleave with cyanogen bromide; separate fragments; sequence by Edman degradation
C-1 EGAAYHDFEPIDPRGASM
C-2 TKDCVHSD
C-3 ALIKYLIACGPM
C-3 overlaps with T-1 and T-4, allowing them to be ordered.
establish sequence
Amino terminus
T-2 T-1 T-4 T-3
EGAAYHDFEPIDPRGASMALIKYLIACGPMTKDCVHSD
C-1 C-3 C-2
Carboxyl terminus

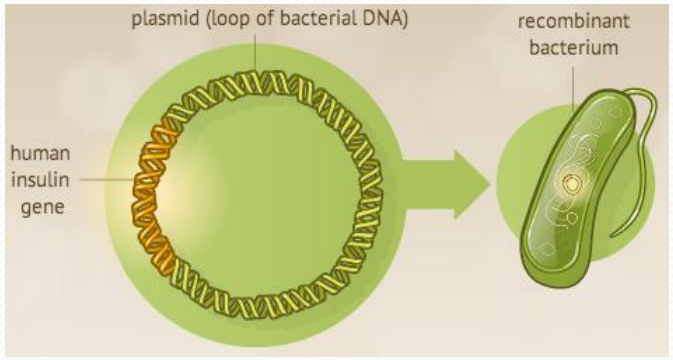
Protein Synthesis

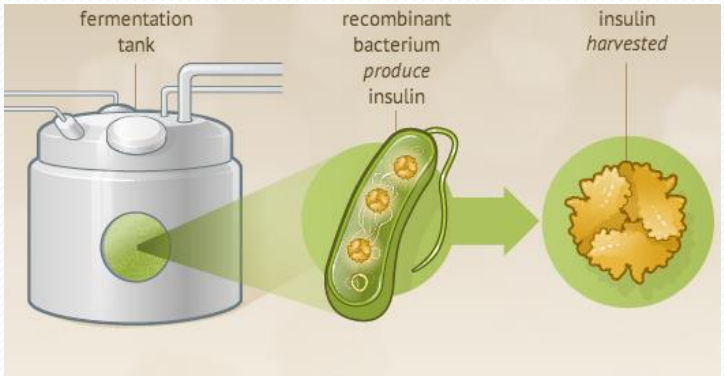
- Proteins are synthesized in the cell by ribosomes using mRNAs transcribed from the genes.
- Proteins can also be obtained by:
  (1) purification from tissue
  (2) genetic engineering using recombinant DNA technology
  (3) direct chemical synthesis

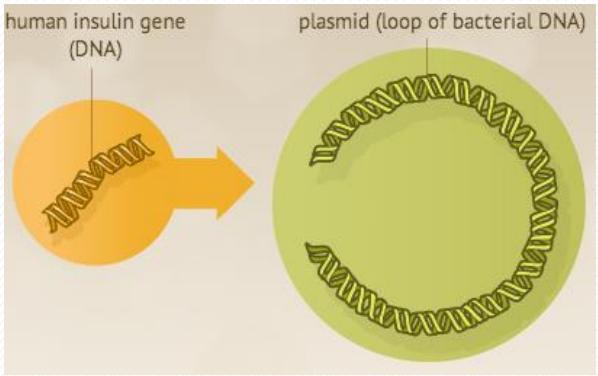
Recombinant DNA Technology

- Protein of interest can be synthesized by genetic recombination in bacteria and then isolated and pırified for further use.









Chemical Synthesis of Proteins

- Merrifeld proposed in 1962 to synthesize a peptide while keeping it attached at one end to a solid support. The support is an insoluble polymer (resin) contained within a column.
- The peptide is built up on this support one amino acid at a time using a standard set of reactions in a repeating cycle.
- At each successive step in the cycle, protective chemical groups block unwanted reactions.
- Peptides up to 100 aa residue can be synthesized efficiently using this method.

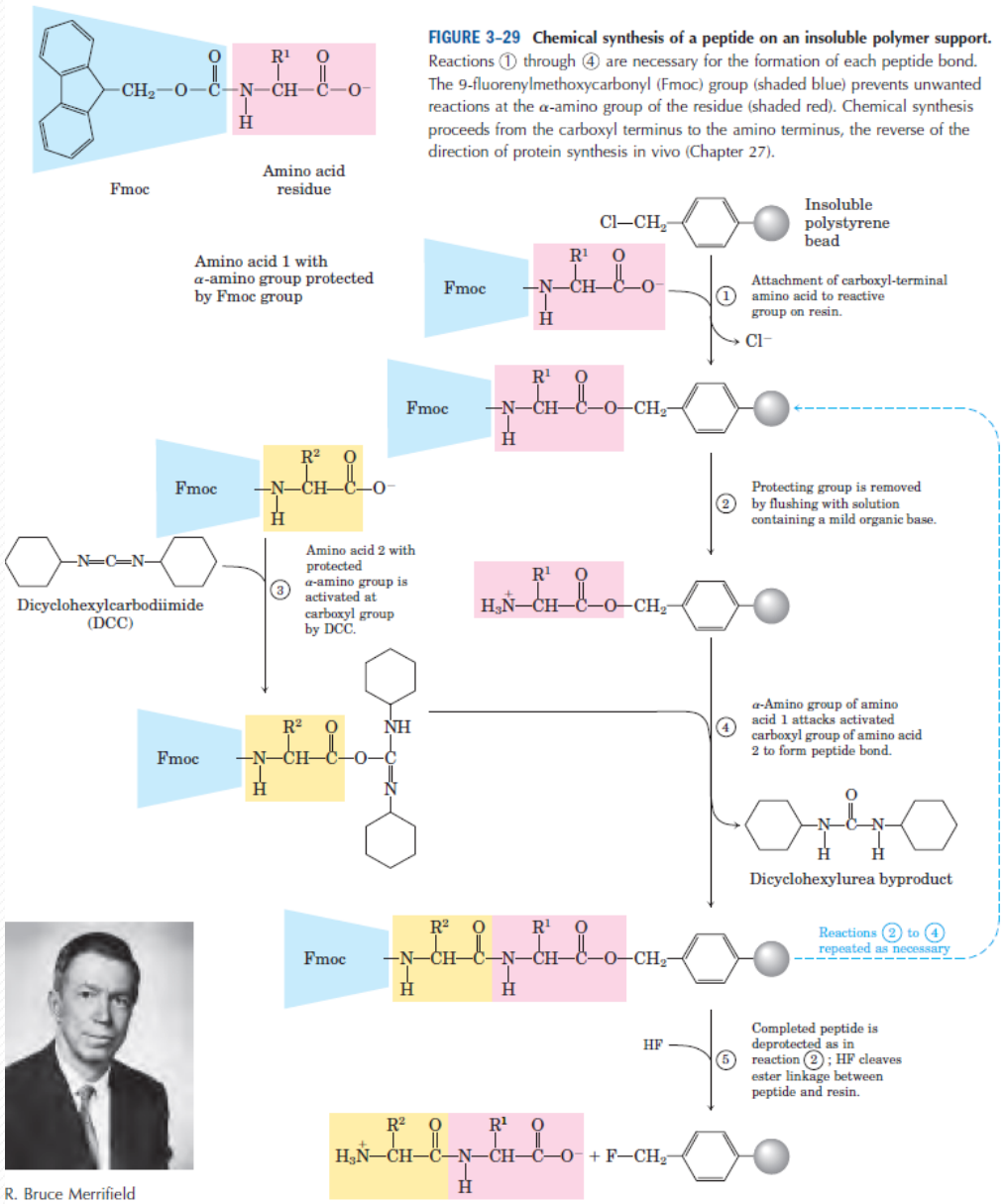


FIGURE 3–29 **Chemical synthesis of a peptide on an insoluble polymer support.** Reactions ① through ④ are necessary for the formation of each peptide bond. The 9-fluorenylmethoxycarbonyl (Fmoc) group (shaded blue) prevents unwanted reactions at the α-amino group of the residue (shaded red). Chemical synthesis proceeds from the carboxyl terminus to the amino terminus, the reverse of the direction of protein synthesis in vivo (Chapter 27).

R. Bruce Merrifield

Studying 3D structure of Proteins

- X-ray
- NMR