Protein Structure Determination-Part 2

Dr. Kiran Meena

19/09/2018

Specific learning objectives

- Polyacrylamide Gel Electrophoresis (PAGE): Native and SDS-PAGE
- Electrophoresis Pattern for Plasma Proteins
- Isoelectric Focusing
- Two-Dimensional Electrophoresis
- Sanger sequencing
- Edman degradation
- Summary

Introduction

- Separation of proteins based on migration of charged proteins in an electric field, a process called electrophoresis.

- Estimate quickly the number of different proteins in a mixture or the degree of purity of a particular protein preparation.

- Determination of crucial properties of a protein such as its isoelectric point and molecular weight.

Polyacrylamide Gel Electrophoresis (PAGE)

- Polyacrylamide gel acts as a molecular sieve, slowing the migration of proteins in proportion to their charge-to-mass ratio.

- Electrophoretic mobility of the molecule, μ is the ratio of the velocity of the particle molecule, V, to the electrical potential, E, the force moving the macromolecule is the electrical potential, $\mu=V/E$.

- Electrophoretic mobility is also equal to the net charge of the molecule, Z, divided by the frictional coefficient, f, which reflects in part a protein's shape, $\mu=Z/f$.

- Migration of a protein in a gel during electrophoresis based on a function of its size and its shape.

Cont--

- SDS (Sodium Dodecyl Sulfate) binds to most proteins proportional to the molecular weight of the protein.
- One molecule of SDS for every two amino acid residues.
- Bound SDS contributes a large net negative charge, confer on each protein a similar charge-to-mass ratio.

Cont--

- Native conformation of a protein is altered when SDS is bound, and most proteins attains a similar shape.

- Electrophoresis in the presence of SDS, separates proteins based on mass (Mol Wt.) with smaller polypeptides migrating more rapidly.

- After electrophoresis, the proteins are visualized by adding a dye such as Coomassie blue, which binds to proteins but not to the gel itself.

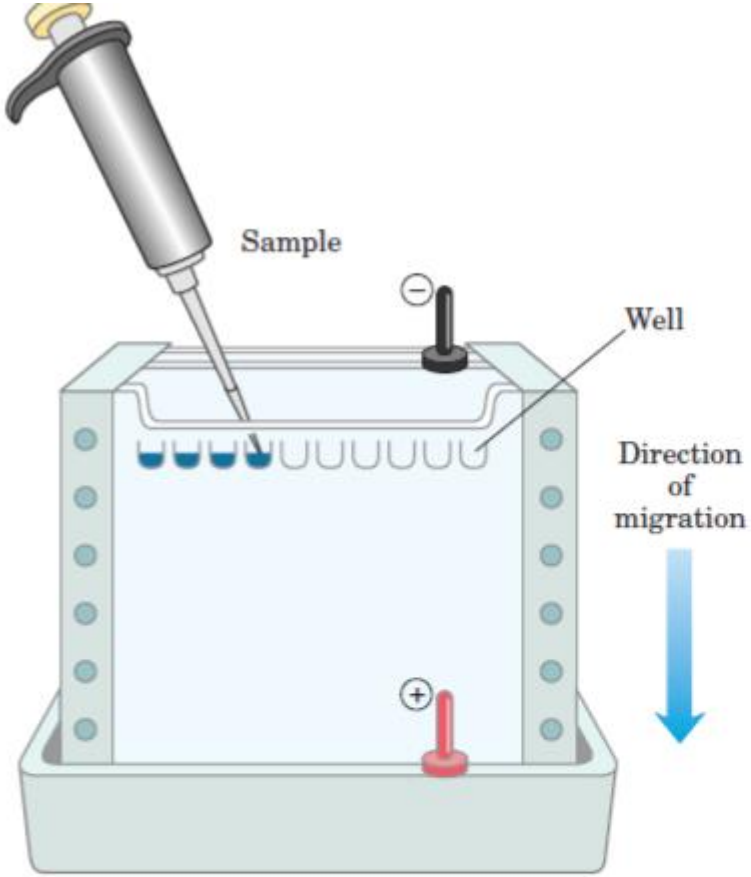


FIGURE 3–19 Electrophoresis. **(a)** Different samples are loaded in wells or depressions at the top of the polyacrylamide gel. The proteins move into the gel when an electric field is applied. The gel minimizes convection currents caused by small temperature gradients, as well as protein movements other than those induced by the electric field. **(b)** Proteins can be visualized after electrophoresis by treating the gel with a stain such as Coomassie blue, which binds to the proteins but not to the gel itself. Each band on the gel represents a different protein (or protein subunit); smaller proteins move through the gel more rapidly than larger proteins and therefore are found nearer the bottom of the gel. This gel illustrates the purification of the enzyme RNA polymerase from *E. coli*. The first lane shows the proteins present in the crude cellular extract. Successive lanes (left to right) show the proteins present after each purification step. The purified protein contains four subunits, as seen in the last lane on the right.

Fig 3.19:Harper's Illustrated Biochemistry, 30th Edition

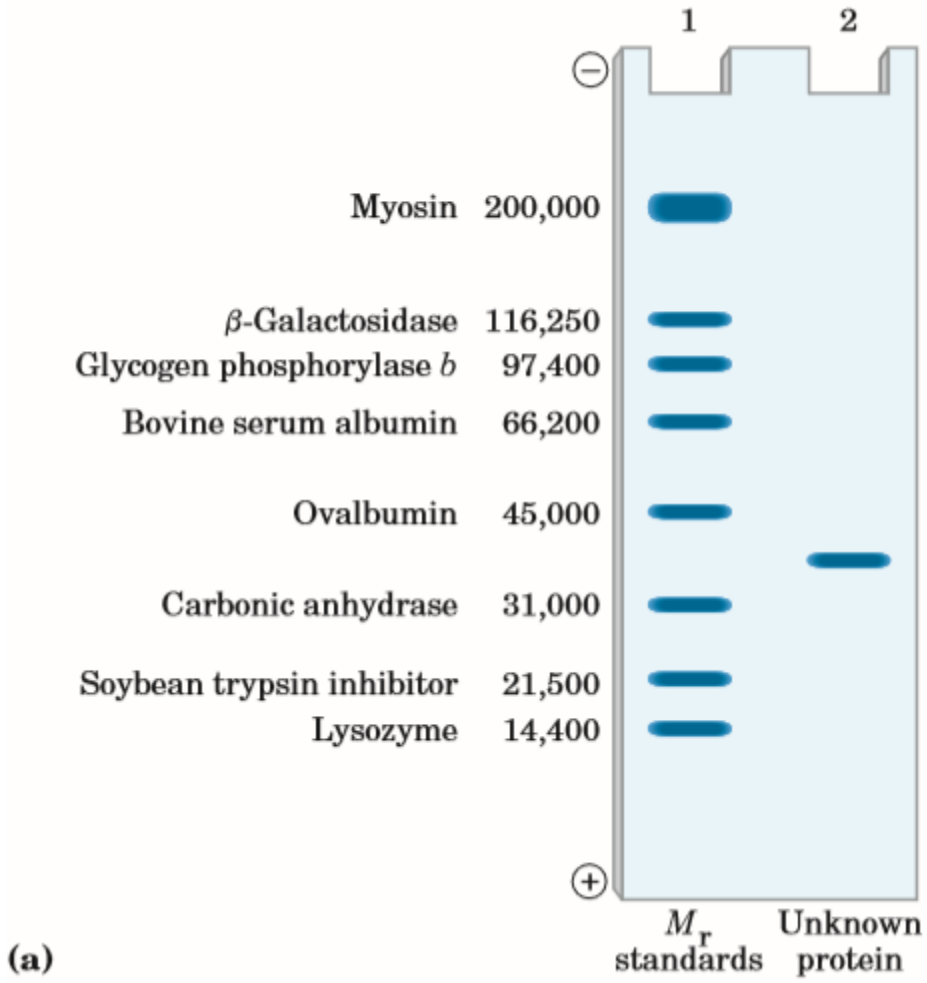
Short Native and SDS-PAGE videos attached with email

How to estimating molecular weight of a protein?

- Standard proteins of known molecular weight (Proteins marker) used to estimate the molecular weight of an unknown protein.

- Position of an unidentified protein provide an measure of its MW compared with the positions to which standard proteins of known molecular weight (MW) migrate in the gel.

- If the protein has two or more different subunits, the subunits separated by SDS treatment and a separate band will appear for each.



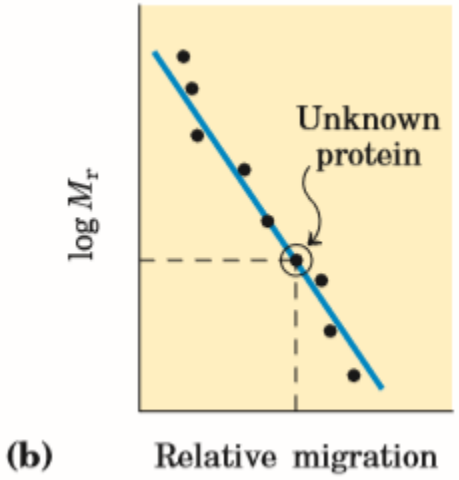


FIGURE 3–20 **Estimating the molecular weight of a protein.** The electrophoretic mobility of a protein on an SDS polyacrylamide gel is related to its molecular weight, M_r. **(a)** Standard proteins of known molecular weight are subjected to electrophoresis (lane 1). These marker proteins can be used to estimate the molecular weight of an unknown protein (lane 2). **(b)** A plot of log M_r of the marker proteins versus relative migration during electrophoresis is linear, which allows the molecular weight of the unknown protein to be read from the graph.

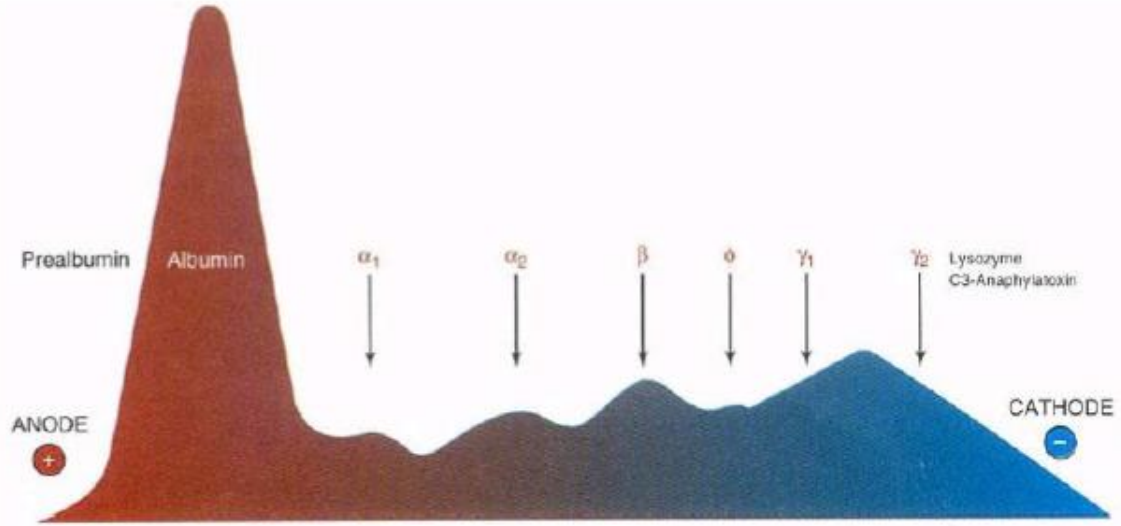
Fig 3.20: Lehninger Principles of Biochemistry by David L Nelson

Electrophoresis Pattern for Plasma Proteins

- Major peaks observed based on their migration are those of albumin, α_1, α_2, and β-globulins, fibrinogen, and γ_1 and γ_2 globulins.

- Some of these peaks represents tens to hundreds of different plasma proteins that have a similar migration rate at pH 8.6.

- Certain proteins predominate in each peak and variation in their relative amounts is characteristic of certain diseases.



α_1	α_1/α_2	α_2	α_2/β_1
α_1-Acid glycoprotein	Thyroxine-binding globulin	Retinol-binding protein	Serum cholinesterase
α_1-T Glycoprotein	Zn-α_2-glycoprotein	α_2 HS glycoprotein	8 S-α_3-glycoprotein
α_1-Antitrypsin	Gc globulin	Histidine-rich 3,8 S-α_2-glycoprotein	4 S-α_2, β_1-glycoprotein
Transcortin	Ceruloplasmin	Haptoglobin	Transcobalamin III
α_1-Antichymotrypsin	Inter-α-trypsin inhibitor	Pregnancy zone protein	
α_1-B Glycoprotein	Antithrombin III	α_2-Macroglobulin	
9.5 S-α_1-glycoprotein	Factor X (Stuart-Prower)	Prothrombin	
Vitamin D-binding protein	Transcobalamin I	Antihemophilic factor	
α_1-Lipoproteins	C9	C1 inactivator	
		C1s	

β_1	β_2	γ_1	γ_2
Hemopexin	β_2-Microglobulin	IgG	IgG
Steroid-binding β-globulin	β_2-Glycoprotein III	IgA	Clq
Transferrin	β_2-Glycoprotein I	IgD	Properdin
Pregnancy-specific β_1-glycoprotein	Fibrinogen	IgE	
Cold insoluble globulin	Factor XI (PTA)	IgM	
Factor V (Accelerin)	Factor XII (Hageman)	Amylase	
Factor VII (Proconvertin)	Factor XIII (FSF)		
Factor IX (Christmas)	C3, C6, C7		
Plasminogen	C3 Activator (β_2I)		
C3 Proactivator			
Transcobalamin II			
β-Lipoproteins			
C1r, C2			
C4, C5			

Figure 2.20
Electrophoresis pattern for plasma proteins at pH 8.6.
Plot shows the order of migration along the horizontal axis with proteins of highest mobility closest to the anode. Height of the band along the vertical axis shows the protein concentration. Different major proteins are designated underneath their electrophoretic mobility peaks.
Reprinted with permission from Heide, K., Haupt, H., and Schwick, H. G. In: F. W. Putnam (Ed.), *The Plasma Proteins*, 2nd ed., Vol. III. New York Academic Press, 1977, p. 545.

Fig. 2.20. Textbook of Biochemistry with Clinical Correlations, 4th edition by Thomas M Devlin

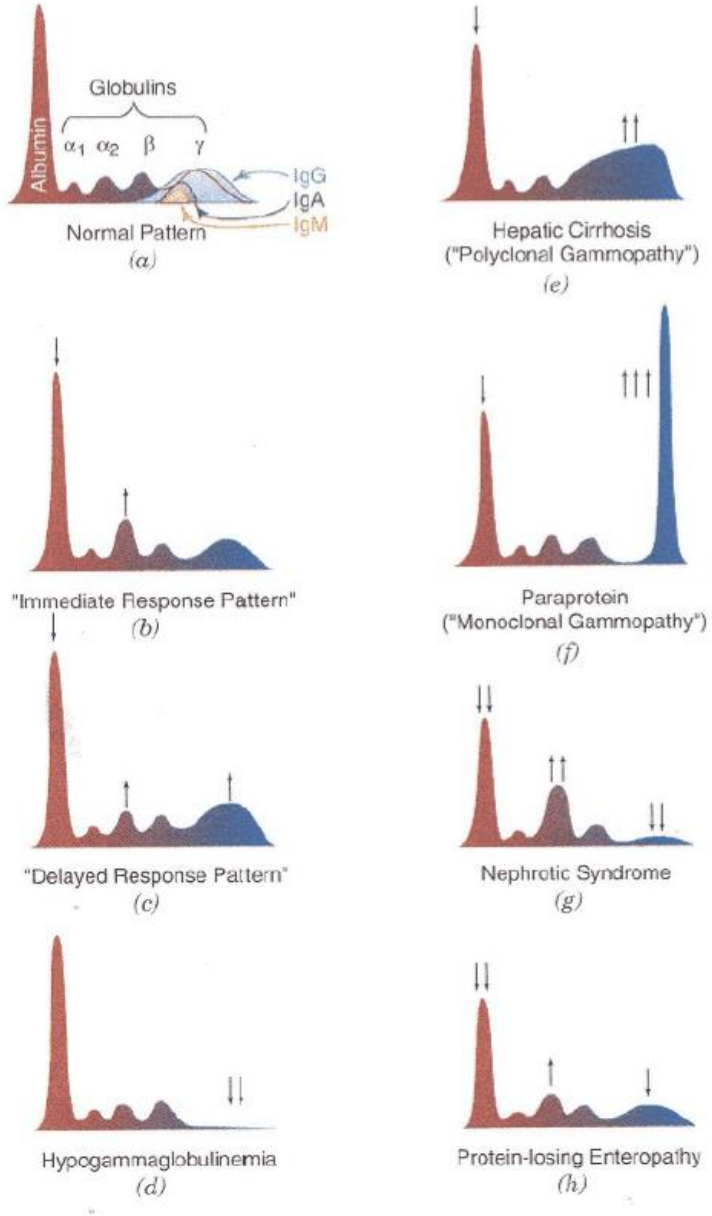
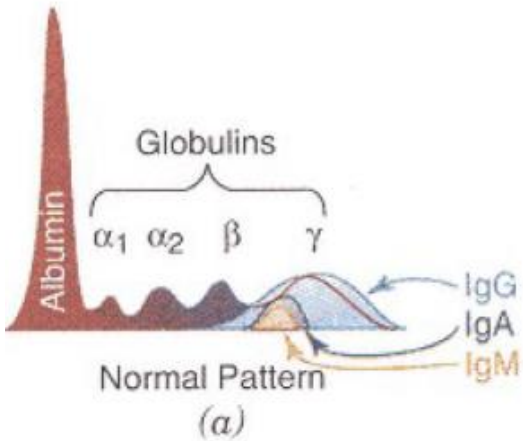


Figure 2.21
Examples of the electrophoretic mobility patterns observed for a normal individual and patients with abnormal concentrations of serum proteins, analyzed by agarose gel electrophoresis.

Fig. 2.21. Textbook of Biochemistry with Clinical Correlations, 4th edition by Thomas M Devlin

- An “Immediate response” occur with stress or inflammation caused by infection, injury or surgical trauma in which hapatoglobin in the α_2 mobility are selectively increased.



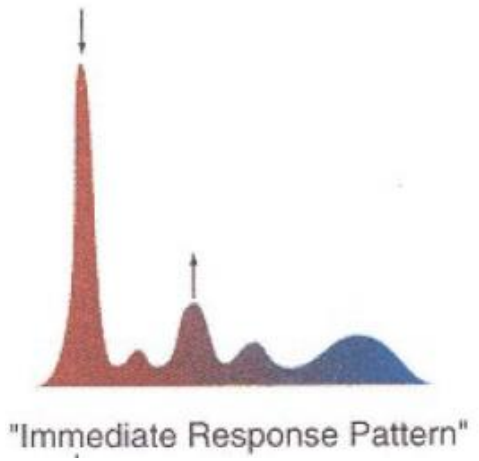
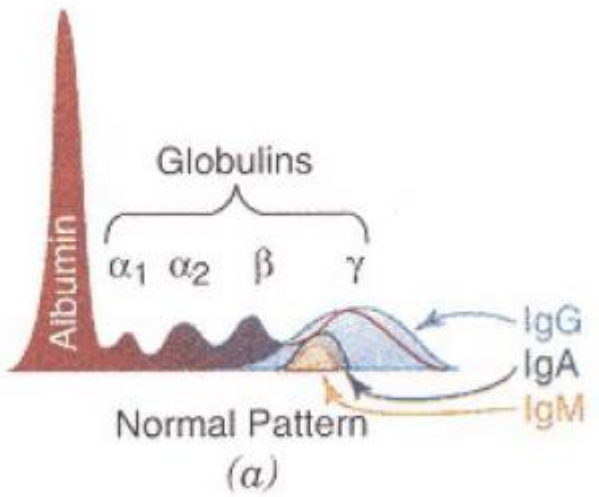


Fig. 2.21. Textbook of Biochemistry with Clinical Correlations, 4th edition by Thomas M Devlin

- “Delayed response” correlated with infection and an increase in the γ-globulin peaks due to increase in immunoglobulins (Igs).



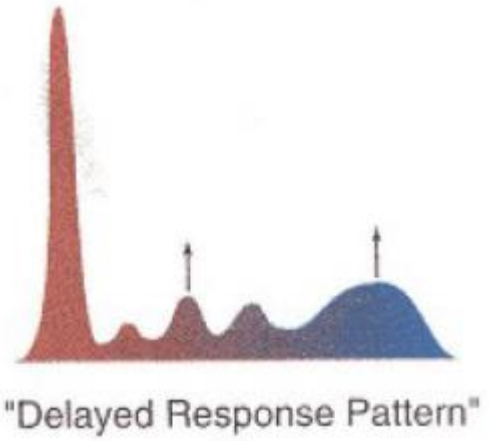


Fig. 2.21. Textbook of Biochemistry with Clinical Correlations, 4th edition by Thomas M Devlin

- Hypogammaglobulinemia due to an immunosuppressive disease.
- Found in autoimmune disorders such as Crohn's disease (inflammatory bowl disease).
- Reduction in all types of γ-globulins.

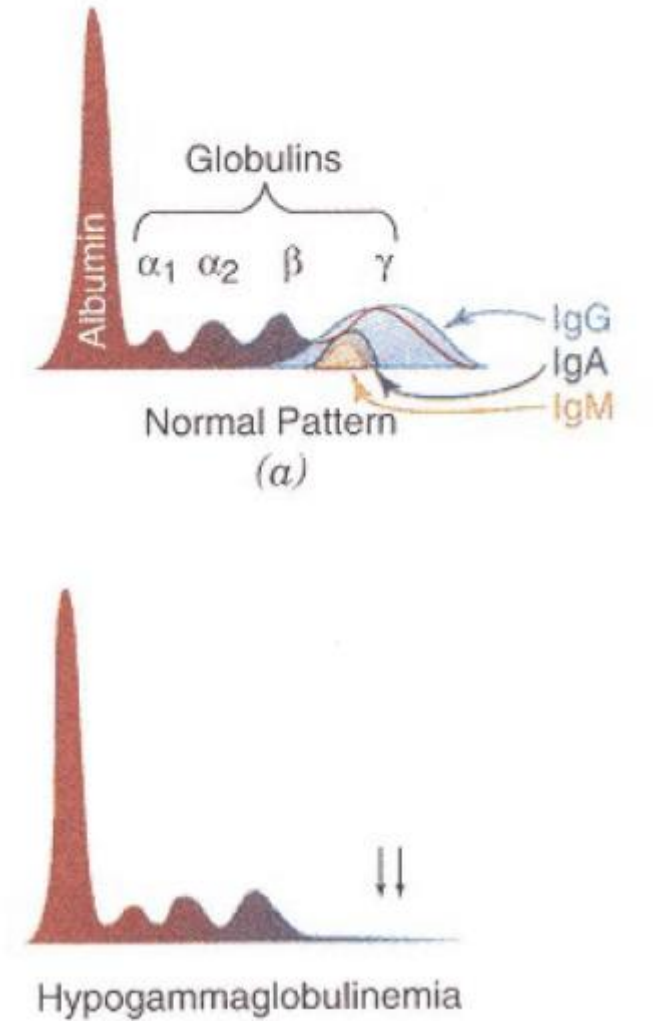


Fig. 2.21. Textbook of Biochemistry with Clinical Correlations, 4th edition by Thomas M Devlin

- In hepatic cirrhosis, elevation of the γ-globulins with reduction of albumin pattern.

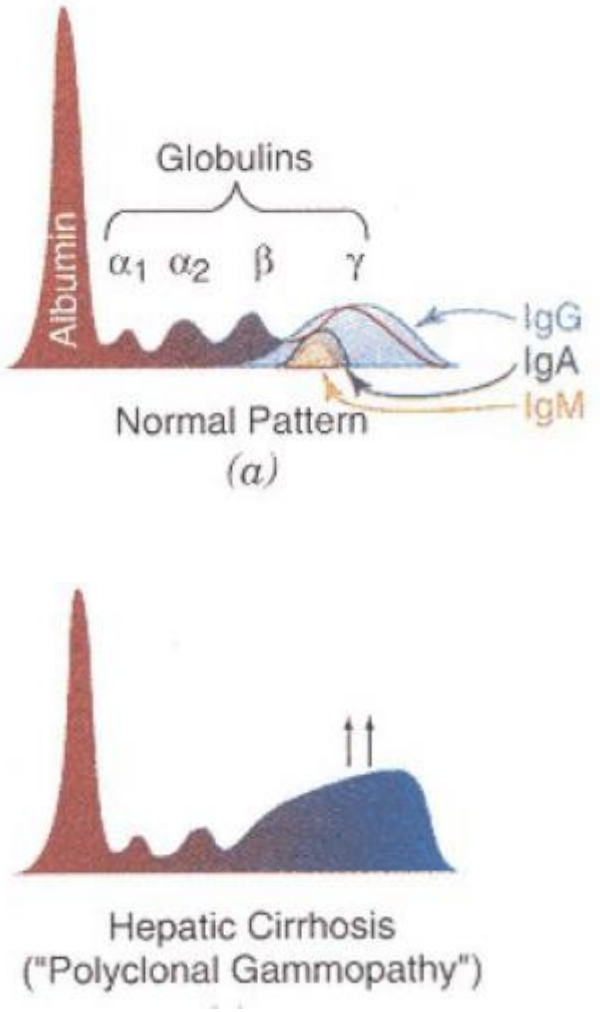


Fig. 2.21. Textbook of Biochemistry with Clinical Correlations, 4th edition by Thomas M Devlin

- Monoclonal gammapathies are due to the clonal synthesis of a unique Ig and give rise to a sharp γ-globulin band pattern.

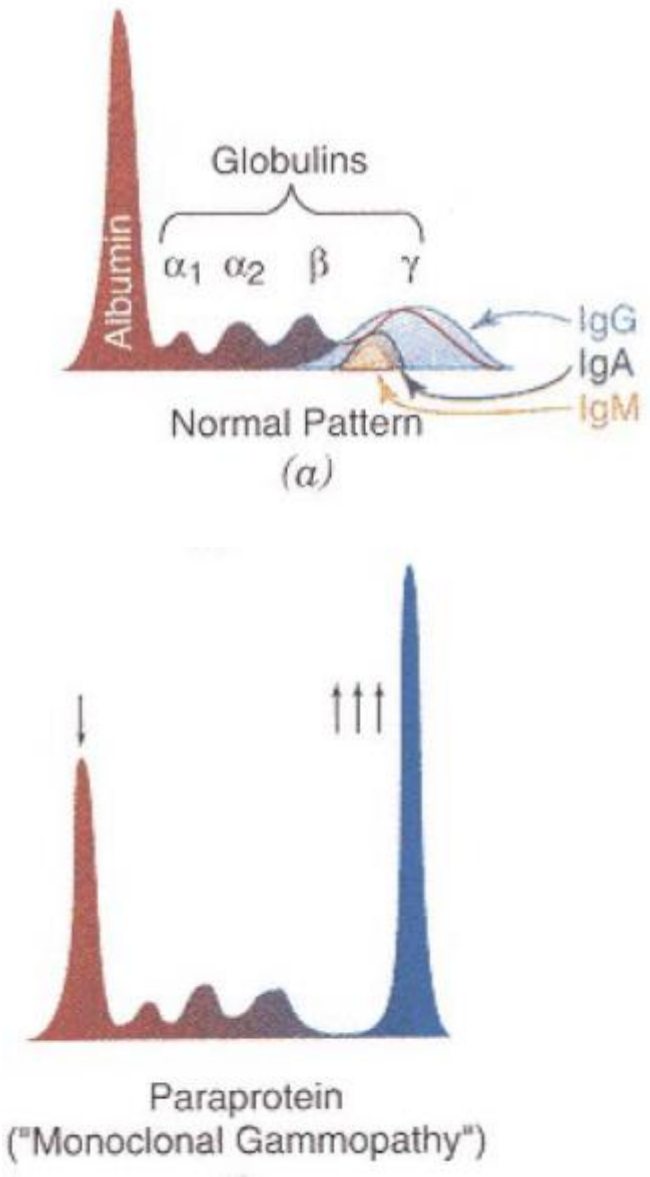


Fig. 2.21. Textbook of Biochemistry with Clinical Correlations, 4th edition by Thomas M Devlin

- Nephrotic syndrome shows selective loss of lower molecular weight proteins from plasma.

- Pattern shows a decrease in albumin (65 kDa), but retention of the bands composed of the higher molecular weight proteins α_2 macroglobulin (725 kDa) and β-lipoproteins (2000 kDa) in α_2 band.

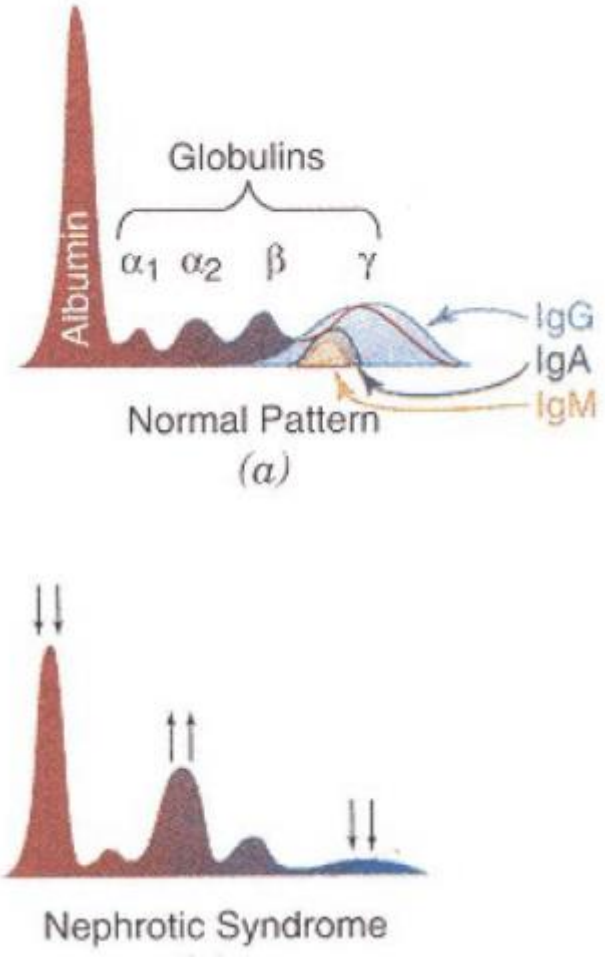
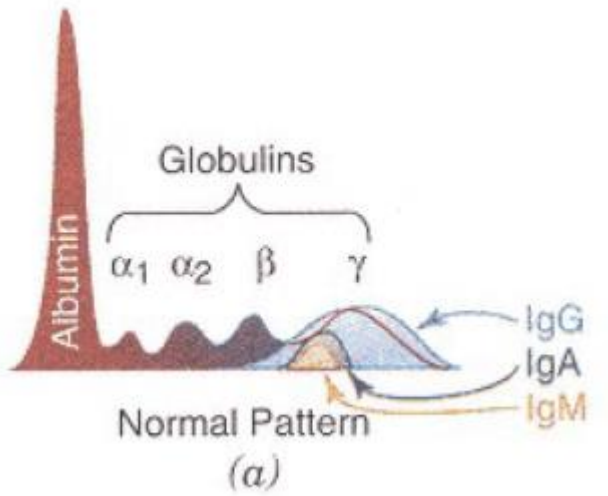


Fig. 2.21. Textbook of Biochemistry with Clinical Correlations, 4th edition by Thomas M Devlin

- This pattern is from a patient with a protein losing enteropathy.
- Slight increase in the α_2 band in this pattern is due to an immediate or late response from a stressful stimulus.



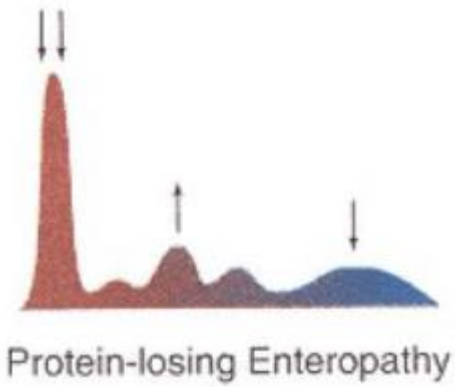


Fig. 2.21. Textbook of Biochemistry with Clinical Correlations, 4th edition by Thomas M Devlin

Electrophoretogram of a normal serum sample showing polyclonal nature of Igs

- Normal serum sample shows polyclonal nature of Igs.
- Ig peak is very broad as it consists of many different proteins each with a different amino acid sequence and different mobility.
- Contrasted with sharp peaks of transferrin and complement C3 each consist of a single protein.

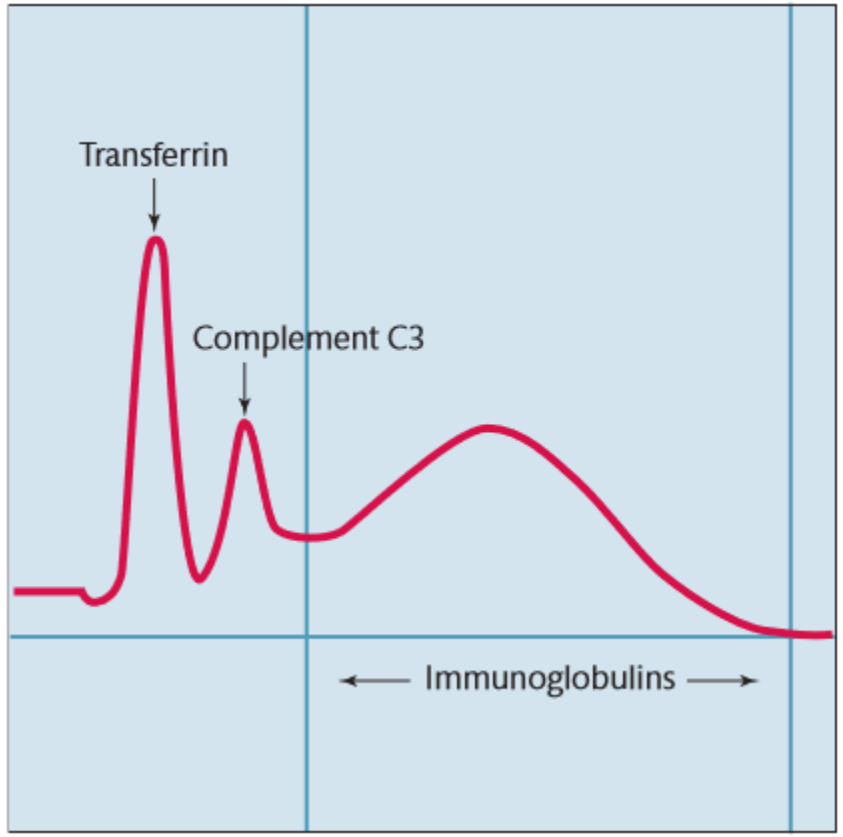


Fig.18.11.Clinical Biochemistry, by Nessar Ahmed

Electrophoretogram of a normal serum sample showing a monoclonal Ig band

- If one particular clone of plasma cells grows large, then it synthesize and secrete a correspondingly large amount of a single Ig. This particular Ig described as monoclonal.

- Presence of monoclonal antibodies in serum samples detected as ‘bands’ using electrophoretic techniques.

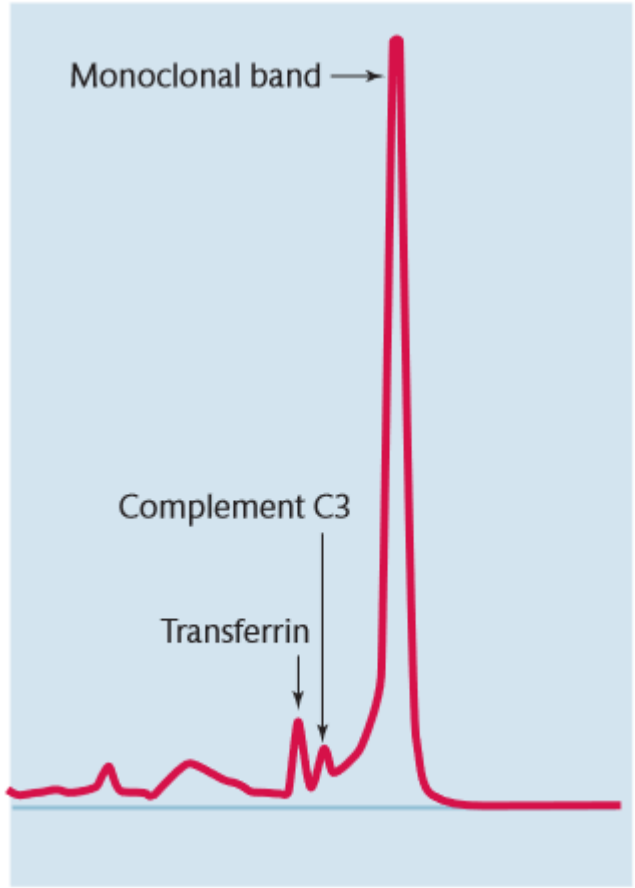


Fig.18.12.Clinical Biochemistry, by Nessar Ahmed

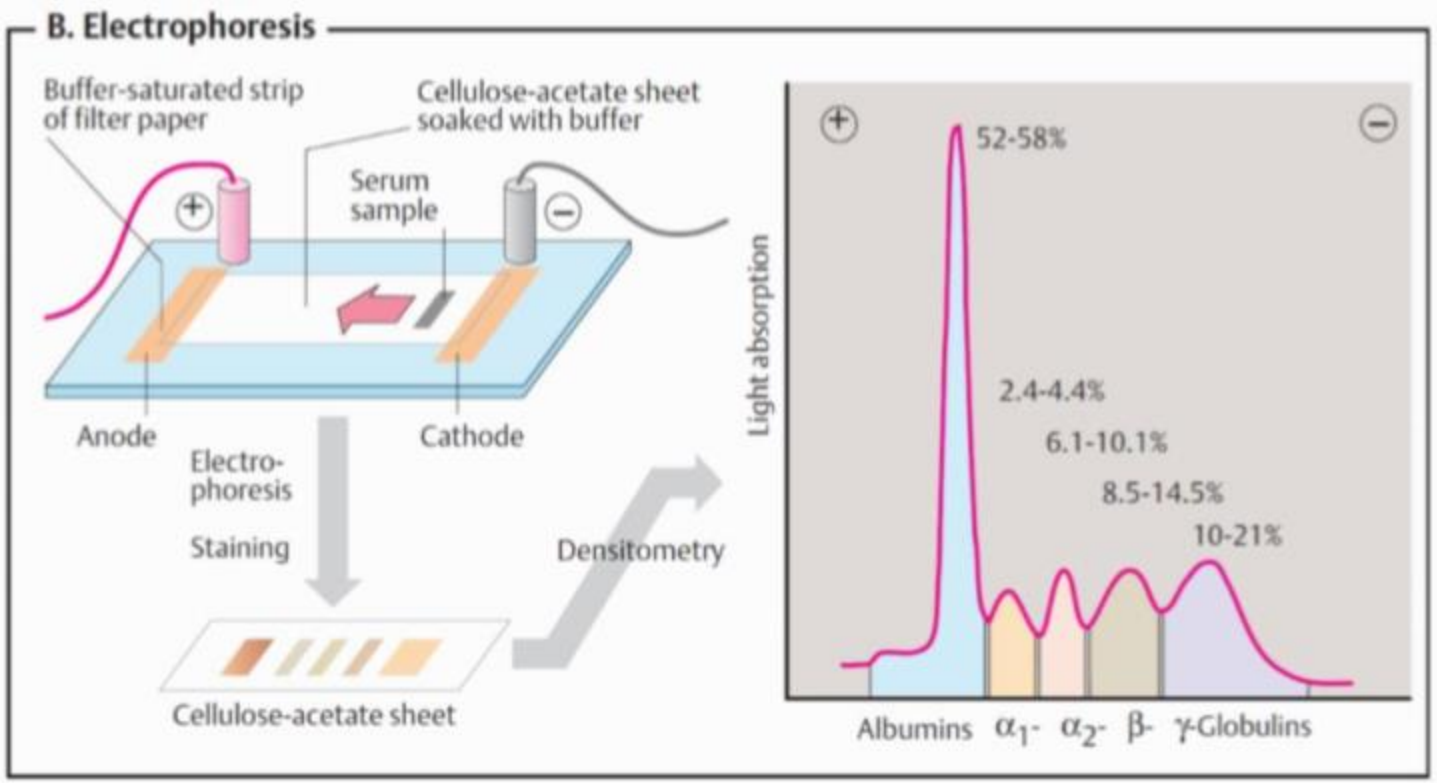
Cellulose Acetate Electrophoresis of Serum Sample

- This technique utilizes native protein charge to separate proteins based on their isoelectric point.
- A sample protein is dotted on the marked center of a cellulose acetate strip.
- Strip is placed in barbital buffer of a desired pH and voltage is applied across the strip.
- Proteins that migrate towards the anode have a pI greater than the pH of the buffer while proteins that migrate towards the cathode have a pI less than the pH of the buffer.

Cont--

- Positively charged proteins migrate towards the cathode while negatively charged proteins migrate toward the anode.
- It is readily visible that albumin is major protein constituent.
- α_1 fraction is mainly α_1-antitrypsin, while the α_2 fraction includes hapatoglobin.
- B-globulin fraction includes a number of lipoproteins, transferrin, plasminogen and complement proteins.

Cellulose Acetate Electrophoresis of Serum Sample



UG Practical Manual 2018, AIIMS, Delhi

Serum Protein Electrophoresis (SPEP)

- In this technique serum is applied on a support medium and exposed to an electric current.
- Different fractions of serum proteins separate usually into 5 bands, as- albumin, α1, α2 , β, and γ-globulin fractions.
- Interpretation of SPEP to the gamma region, because it mainly composed of Ig.
- Increase in gamma region , shows homogenous spike like a peak in the gamma globulin zone, in case of monoclonal gammapathies.

Cont--

- Monoclonal gammapathies result from the proliferation of a single, malignant clone of plasma cells which produce either a single class of intact Igs, heavy and light chains or both.

- These proteins are called para proteins or M (monoclonal) proteins. The M protein detected as a sharp symmetric spike (M spike) with an α2, β, or a γ mobility.

- Multiple myeloma found at multiple sites within the bone marrow (myelo-), with the accumulation of the tumour (oma) cells.

Cont--

- Normally, plasma cells constitute 1% of cells in the bone marrow, but as disease progress, tumor load in bone marrow increases up to 80%, depends upon disease severity.

- Malignant plasma cells synthesize monoclonal antibodies which are released into the circulation and its level increases in serum.

- In clinical practice, serum protein electrophoresis is used to identify multiple myeloma and other serum protein disorders.

- Many specialists include SPEP as a screening test in the initial evaluation for many clinical conditions.

Normal and Monoclonal Gammapathy Pattern in SPEP

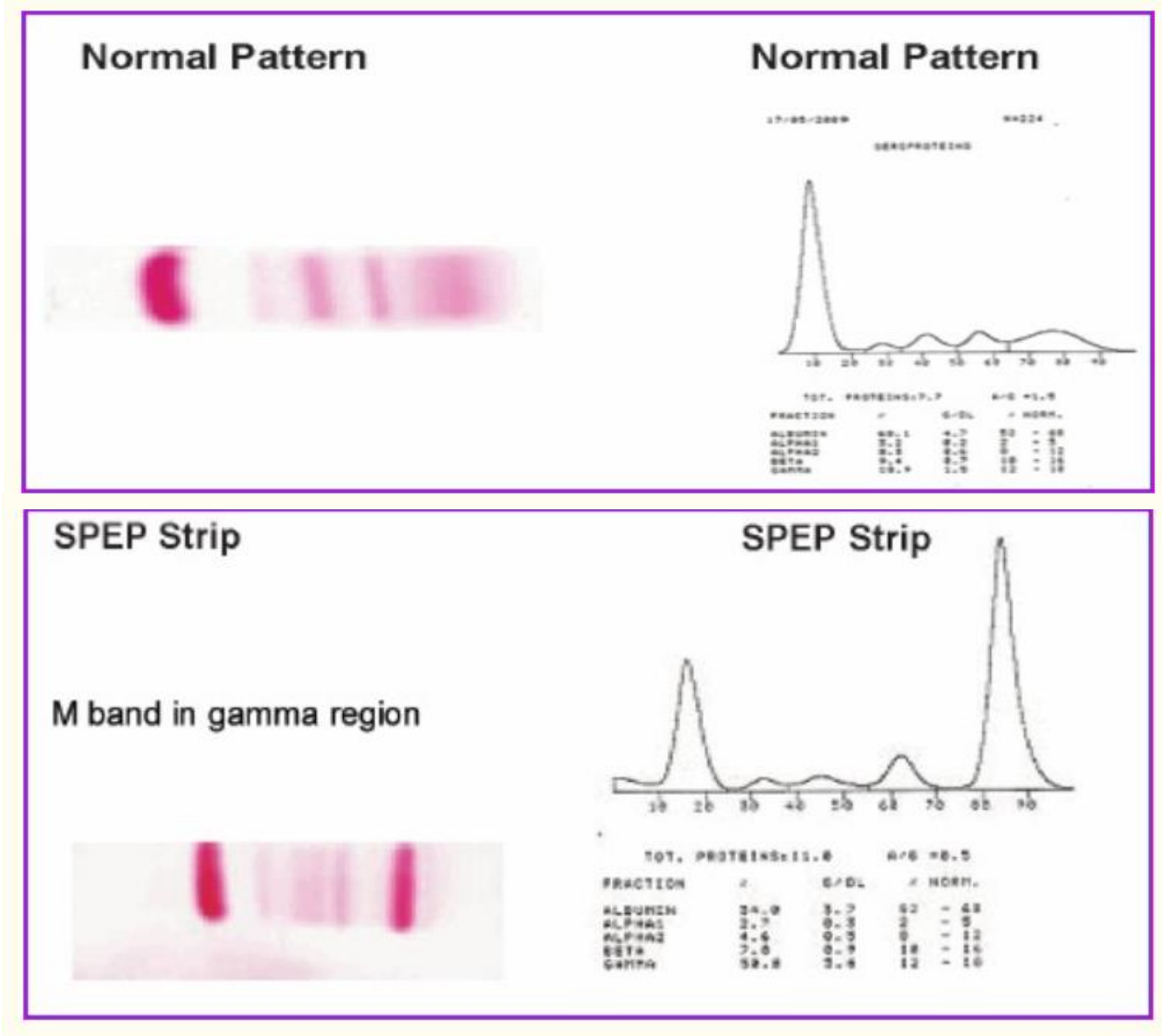


Fig.2 & 3: Tripathy S et al 2012

TABLE 18.6 Monoclonal gammopathies.

Condition	Ig affected	Features
MGUS	G, A, M	Symptomless
Multiple myeloma	G, A (rarely D, extremely rarely M, E)	Bone pain/lesions Anaemia Hypercalcaemia Renal insufficiency
Waldenstrom's macroglobulinaemia	M	Enlarged liver and spleen Anaemia, hyperviscosity
Solitary plasmacytoma	G, A	Bone pain/lesions

Table.18.6.Clinical Biochemistry, by Nessar Ahmed

Proteins Separation Based on pI Values

Isoelectric Focusing

- Used to determine the isoelectric point (pI) of a protein.
- A pH gradient obtained by allow a mixture of low molecular weight organic acids and bases to distribute themselves in an electric field generated across the gel.
- Each protein migrates until it reaches the pH that matches its pI.
- Proteins with different pI are distributed differently throughout the gel.

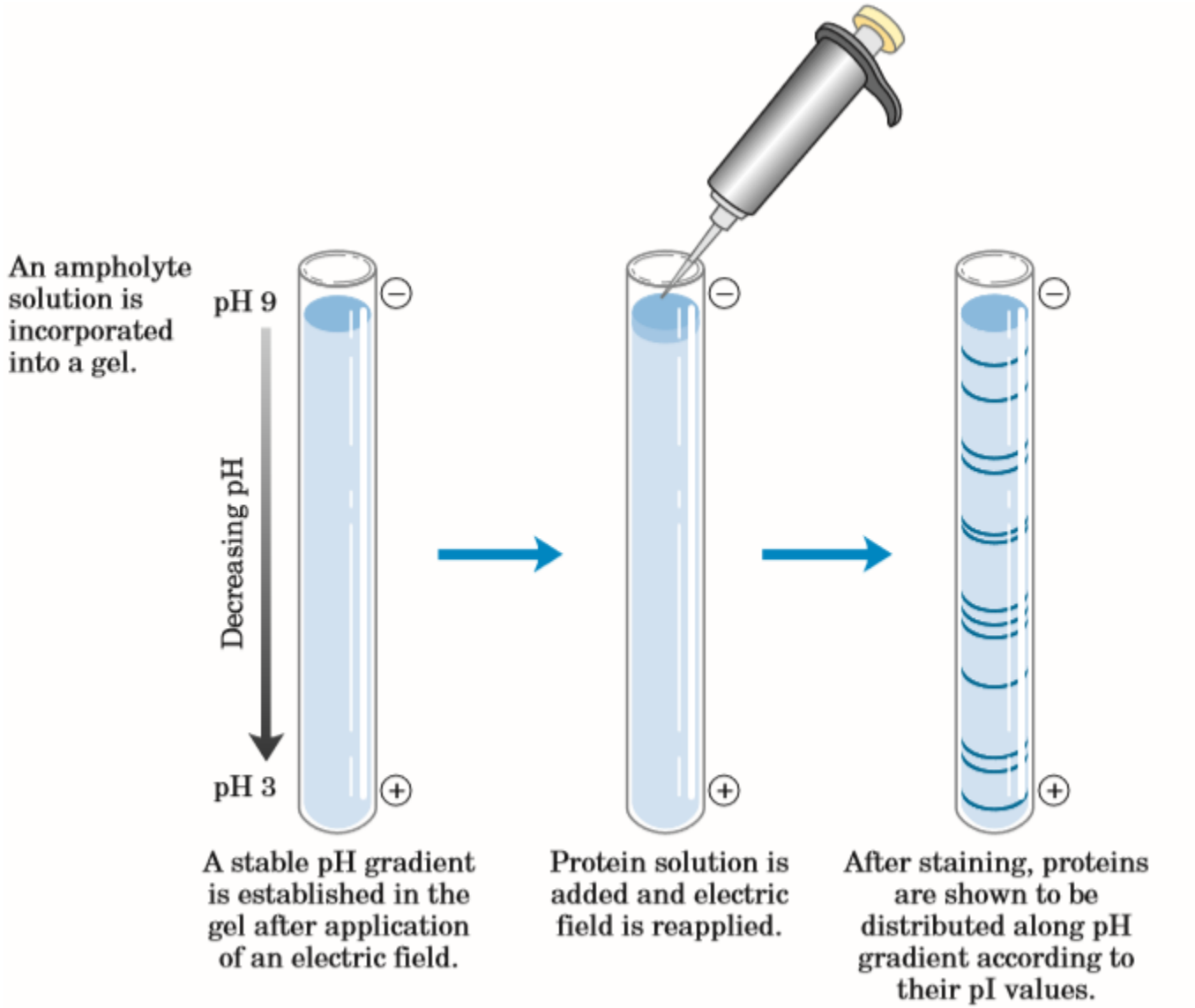


FIGURE 3–21 **Isoelectric focusing.** This technique separates proteins according to their isoelectric points. A stable pH gradient is established in the gel by the addition of appropriate ampholytes. A protein mixture is placed in a well on the gel. With an applied electric field, proteins enter the gel and migrate until each reaches a pH equivalent to its pI. Remember that when pH = pI, the net charge of a protein is zero.

Fig 3.21: Lehninger Principles of Biochemistry by David L Nelson

Two-Dimensional Electrophoresis

- Sequential combination of isoelectric focusing and SDS electrophoresis in a process called two-dimensional (2-D) electrophoresis permits the resolution of complex mixtures of proteins.

- 2-D electrophoresis separates proteins of identical molecular weight that differ in pI, or proteins with similar pI values but different molecular weights.

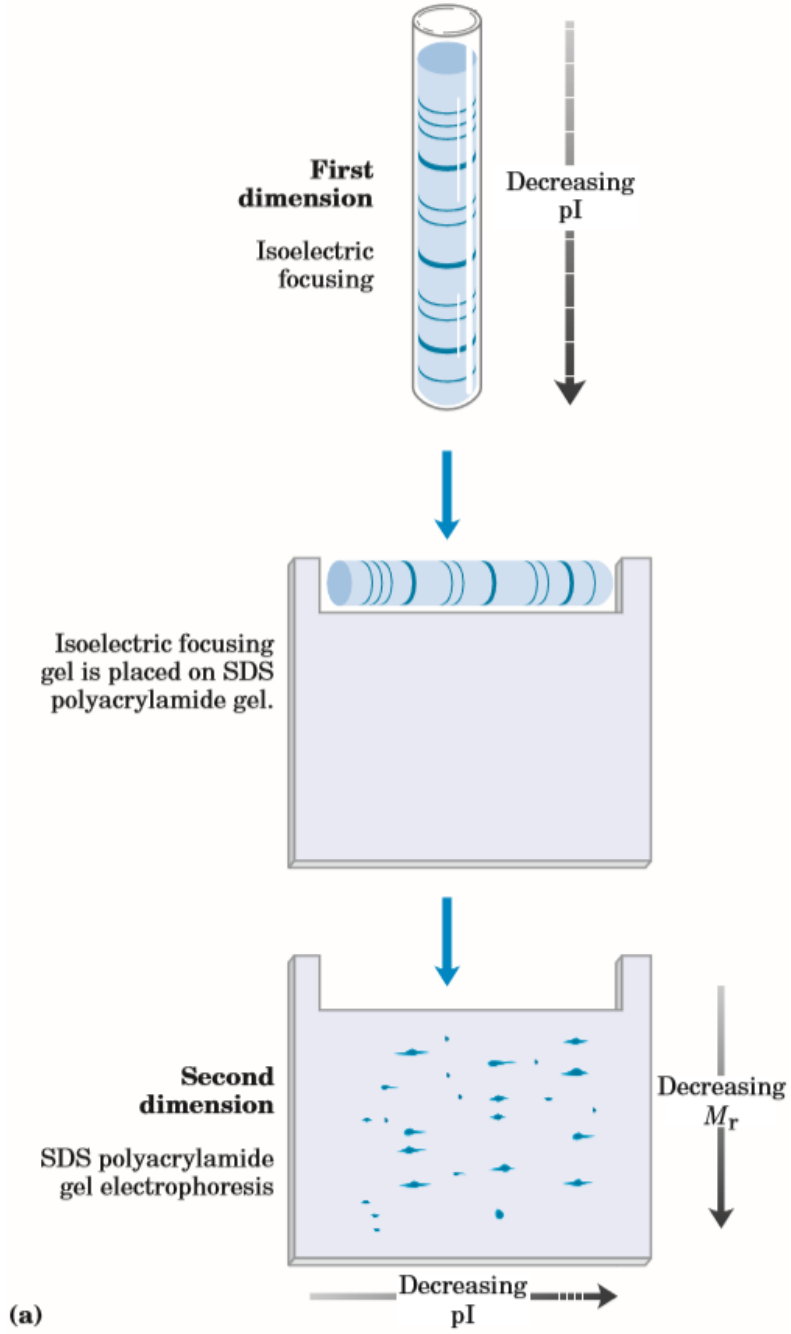


FIGURE 3–22 **Two-dimensional electrophoresis.** **(a)** Proteins are first separated by isoelectric focusing in a cylindrical gel. The gel is then laid horizontally on a second, slab-shaped gel, and the proteins are separated by SDS polyacrylamide gel electrophoresis. Horizontal separation reflects differences in pI; vertical separation reflects differences in molecular weight. **(b)** More than 1,000 different proteins from *E. coli* can be resolved using this technique.

Fig 3.22 (a): Lehninger Principles of Biochemistry by David L Nelson

Short 2-D video attached with email

(b)

FIGURE 3–22 Two-dimensional electrophoresis. **(a)** Proteins are first separated by isoelectric focusing in a cylindrical gel. The gel is then laid horizontally on a second, slab-shaped gel, and the proteins are separated by SDS polyacrylamide gel electrophoresis. Horizontal separation reflects differences in pI; vertical separation reflects differences in molecular weight. **(b)** More than 1,000 different proteins from *E. coli* can be resolved using this technique.

Fig 3.22 (b): Lehninger Principles of Biochemistry by David L Nelson

Sanger Determine Sequence of a Polypeptide

- Mature insulin consists of 21-residue A chain and 30-residue B chain linked by disulfide bonds.

- Frederick Sanger reduced the disulfide bonds separated the A and B chains, and cleaved each chain into smaller peptides using trypsin, chymotrypsin, and pepsin.

- Resulting peptides were isolated and hydrolyzed into a mixture of smaller peptides by treatment with acid.

Cont--

- Each peptide in the mixture was isolated and treated with 1-fluoro-2,4-dinitrobenzene (Sanger reagent FDNB), which reacts with the exposed α-amino groups of the amino-terminal residues.

- Amino acid content of each peptide was determined and amino-terminal amino acid identified.

- The ε-amino group of lysine also reacts with Sanger reagent; but since an amino-terminal lysine reacts with 2 mol of Sanger reagent, it is readily distinguished from a lysine from the interior of a peptide.

- Sanger was able to reconstruct the complete sequence of insulin, an accomplishment for which he received a Nobel Prize, in 1958.

- Sanger, who received his second Nobel prizes for his development of techniques for DNA sequencing, died in 2013 at the age of 95.

TABLE 3–7 The Specificity of Some Common Methods for Fragmenting Polypeptide Chains

*Reagent (biological source)**	*Cleavage points*[†]
Trypsin (bovine pancreas)	Lys, Arg (C)
Submaxillarus protease (mouse submaxillary gland)	Arg (C)
Chymotrypsin (bovine pancreas)	Phe, Trp, Tyr (C)
Staphylococcus aureus V8 protease (bacterium *S. aureus*)	Asp, Glu (C)
Asp-*N*-protease (bacterium *Pseudomonas fragi*)	Asp, Glu (N)
Pepsin (porcine stomach)	Phe, Trp, Tyr (N)
Endoproteinase Lys C (bacterium *Lysobacter enzymogenes*)	Lys (C)
Cyanogen bromide	Met (C)

Table 3.7: Lehninger Principles of Biochemistry by David L Nelson

Edman Reaction Enables Peptides & Proteins to be Sequenced

- Sanger developed the reagent 1-fluoro-2,4dinitrobenzene (FDNB) used to label the amino-terminal residue.

- After the amino-terminal residue is labeled with one of these reagents, the polypeptide is hydrolyzed to its constituent amino acids and the labeled amino acid is identified.

- Because the hydrolysis stage destroys the polypeptide, this procedure cannot be used to sequence a polypeptide beyond its amino-terminal residue.

Cont--

- However, it can help determine the number of chemically distinct polypeptides in a protein, provided each has a different amino-terminal residue. For example, two residues—Phe and Gly—would be labeled if insulin were subjected to this procedure.

- 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.

Cont--

- Peptide is reacted with phenylisothiocyanate under mild alkaline conditions, which converts the aminoterminal aa to a phenylthiocarbamoyl (PTC) adduct.

- Peptide bond next to the PTC adduct is cleaved in a step carried out in anhydrous trifluoroacetic acid, with removal of the amino-terminal aa as an anilinothiazolinone derivative.

- Derivatized amino acid is extracted with organic solvents, converted to more stable phenylthiohydantoin derivative by treatment with aqueous acid, and then identified.

Cont--

- Use of sequential reactions carried out under first basic and then acidic conditions provides control over the entire process.

- Each reaction with the aminoterminal aa can go essentially to completion without affecting any of the other peptide bonds in the peptide.

- After removal and identification of the aminoterminal residue, the new amino-terminal residue so exposed can be labeled, removed, and identified through the same series of reactions. This procedure is repeated until the entire sequence is determined.

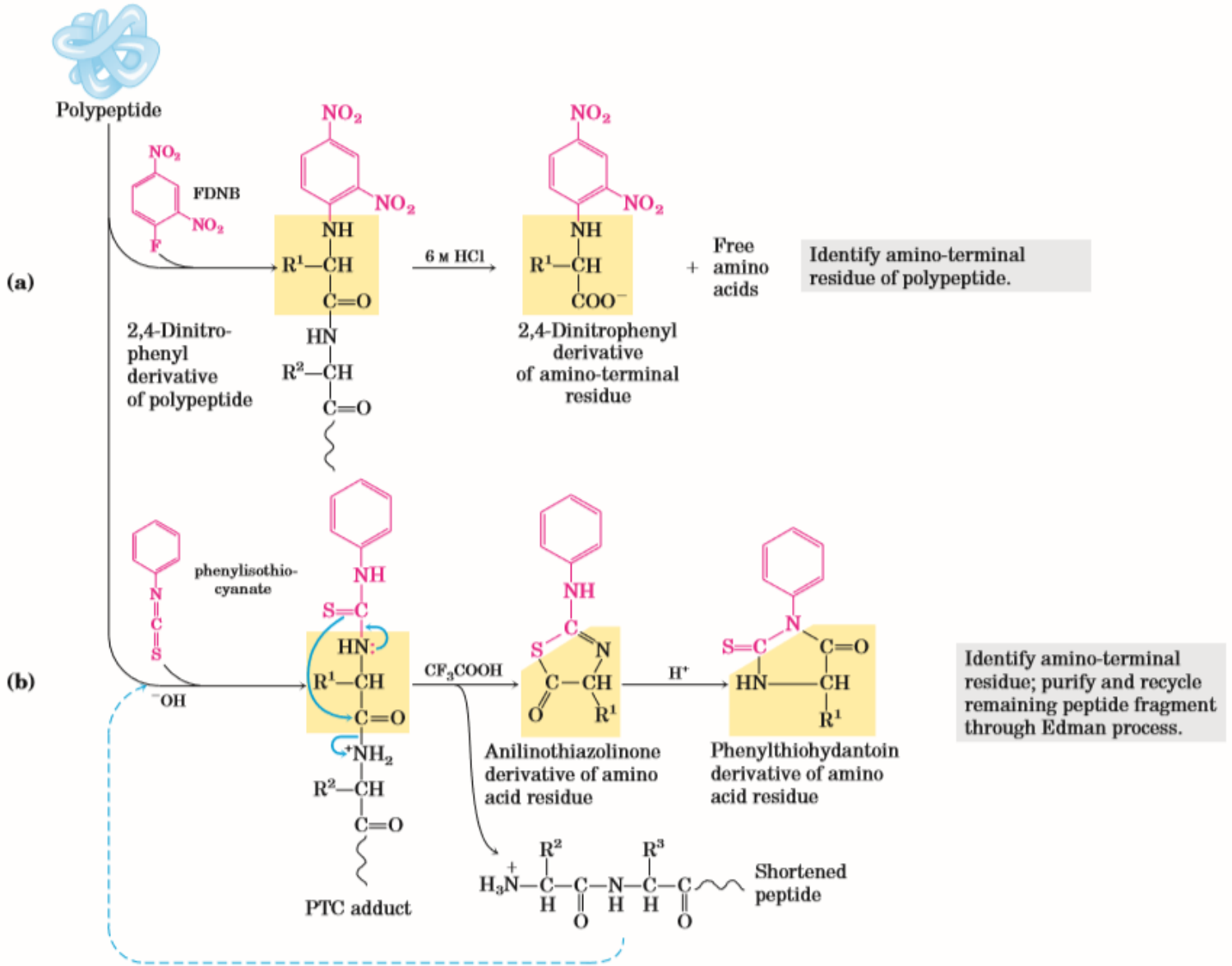


FIGURE 3–25 Steps in sequencing a polypeptide. (a) Identification of the amino-terminal residue can be the first step in sequencing a polypeptide. Sanger's method for identifying the amino-terminal residue is shown here. **(b)** The Edman degradation procedure reveals the entire sequence of a peptide. For shorter peptides, this method alone readily yields the entire sequence, and step **(a)** is often omitted. Step **(a)** is useful in the case of larger polypeptides, which are often fragmented into smaller peptides for sequencing (see Fig. 3–27).

Fig 3.25: Lehninger Principles of Biochemistry by David L Nelson.

Large Proteins Sequenced in Smaller Segments

- Very large polypeptides found in proteins broken down into smaller pieces to be sequenced efficiently. There are several steps in this process:

1. Protein is cleaved into a set of specific fragments by chemical or enzymatic methods.
2. If any disulfide bonds are present, they must be broken.
3. Each fragment is purified, then sequenced by the Edman procedure.
4. Order in which the fragments appear in the original protein is determined and disulfide bonds (if any) are located.

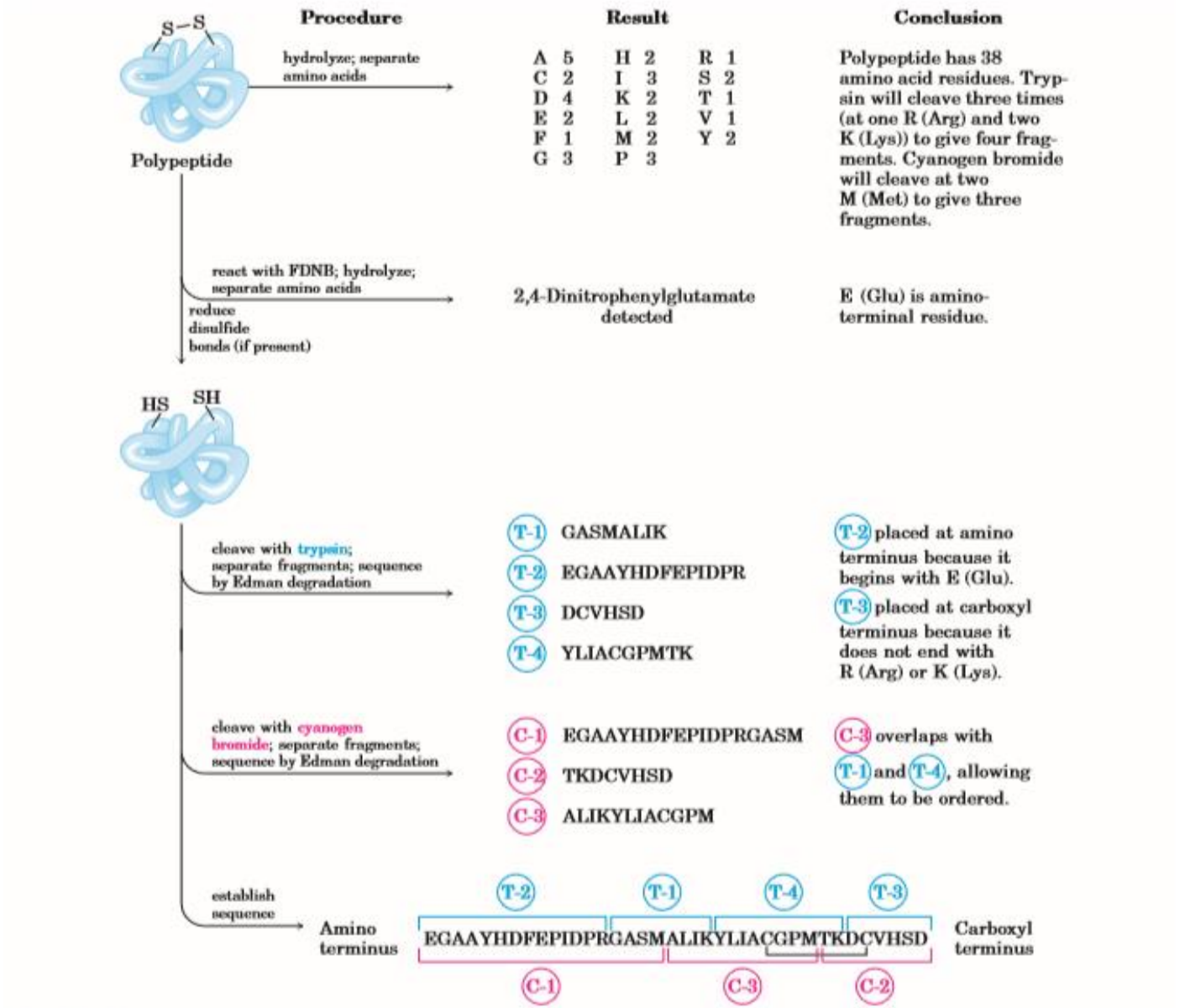


FIGURE 3–27 Cleaving proteins and sequencing and ordering the peptide fragments. First, the amino acid composition and amino-terminal residue of an intact sample are determined. Then any disulfide bonds are broken before fragmenting so that sequencing can proceed efficiently. In this example, there are only two Cys (C) residues and thus only one possibility for location of the disulfide bond. In polypeptides with three or more Cys residues, the position of disulfide bonds can be determined as described in the text. (The one-letter symbols for amino acids are given in Table 3–1.)

Fig 3.27: Lehninger Principles of Biochemistry by David L Nelson.

Summary

- Polyacrylamide gels (PAG) provide a porous matrix for separating proteins on the basis of their mobility in an applied direct current electrical field.

- Constant ratio at which the anionic detergent SDS binds proteins enables SDS-PAGE to separate polypeptides on the basis of relative size.

- Measurements of a number of serum proteins such as albumin, α1-antitrypsin, haptoglobin, ceruloplasmin, transferrin, and C-reactive protein are of value in investigation of certain disease state.

Cont--

- Low concentrations of serum antibodies can result from secondary or primary causes and lead to susceptibility to bacterial infections.

- High serum concentrations of antibody can be polyclonal or monoclonal in nature.

- Presence of monoclonal serum Ig is not uncommon in healthy elderly people, but if accompanied by disease associated symptoms, suggests a diagnosis of myeloma.

Group Discussion

- Subtopics of previous and today's classes discussed in group discussion.

Reference Books

1) Harper's Illustrated Biochemistry-30th edition.
2) Lehninger Principles of Biochemistry by David L Nelson.
3) Textbook of Biochemistry with Clinical Correlations. 4th edition. Thomas M. Devlin.
4) Clinical Biochemistry by Nessar Ahmed.
5) UG Practical Manual 2018, AIIMS, Delhi.
6) You Tube videos on Native, SDS-PAGE and 2-D electrophoresis.
7) Research article: Tripathy S et al 2012.The role of serum protein electrophoresis in the detection of multiple myeloma: an experience of a corporate hospital. J Clin Diagn Res. 2012 Nov;6(9):1458-61.

Thank you