

PROTEINS AND ENZYMES 1-2, Bernadette Byrne

pI: pH at electrical neutrality, amino acids exist as zwitterion (R is neutral)- When R is neutral the isoelectric point is midway between pK1 and pK2

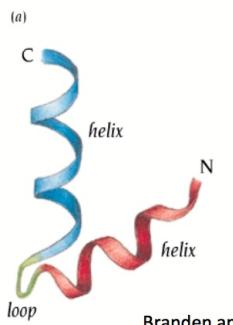
$$pH_{\text{isoelectric}} = (pK_1 + pK_2)/2$$

Charge of Histidine is ca. +0.5

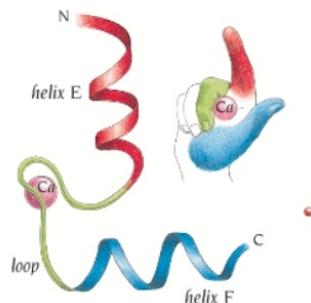
- Alpha helices: Phi/Psi angles of usually 50-60°.
- In beta sheets, H-bonding is interstrand. Phi/Psi angles: +180°.
- Loop regions link regions of secondary structure; phi/psi angle pairs fall within the allowed regions of the Ramachandran plot.
- Secondary structure describes local fold
- Tertiary structure describes overall
- Quaternary structure describes multichain association

Irregularity is important for function of protein; ordered arrangements of secondary structure
 -> motifs (super secondary structure)--> are simple combinations of a few secondary structure elements with a specific geometric arrangements; do not constitute complete structures in themselves; often have a specific function; often have conserved sequence:

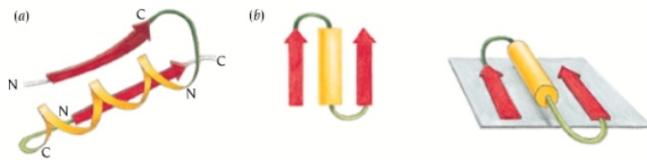
- **Helix-loop-helix** --> simplest motif with a specific function--> motif specific for DNA binding



- **EF-hand motif** --> is a helix-loop-helix but specific for Calcium binding . The loop binds the Ca²⁺, side chain and main chain atoms involved in coordinating the metal



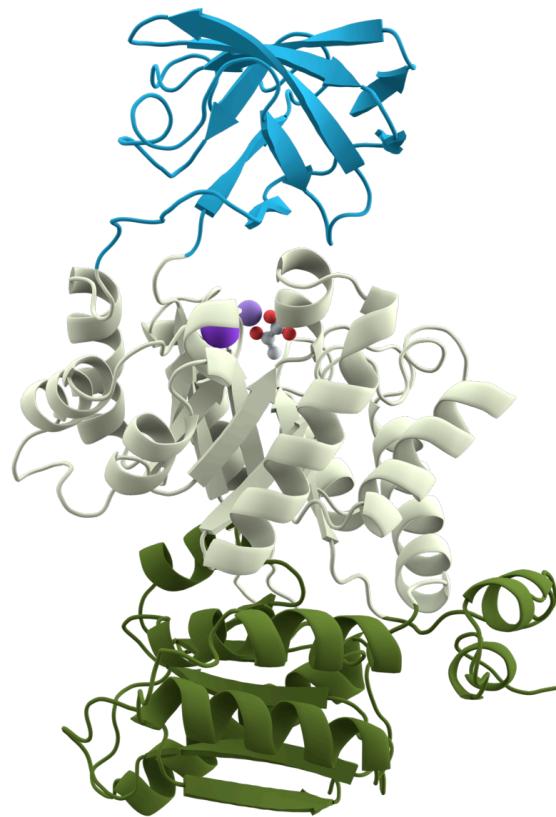
- **β - α - β motif** -> found in almost every structure that contains parallel β sheets; α -helix connects the C-terminus of one β -strand with the N-terminus of a second; β - α - β contains two β -strands, two loop regions and one α -helix; it can be found in larger domains



Proteins can be comprised of different domains plugged together:

Pyruvate Kinase:

- 530 amino acids- 4 different domain
- Molecule is a tetramer-only shown one monomer
- N-terminal domain forms contacts for oligomer formation
- Core functional domain- α/β barrel
- Two regions extend from the core domain, β sheet domain and C-terminal domain



PROTEIN PURIFICATION

Structure is related to function of proteins. Most proteins are produced in the cell--> how to remove the molecules from native environment.

Protein purification is essential, there are various ways to do it. One technique involves differential centrifugation--> obtain fractions containing specific parts of the cell (is also used to determine sub-cellular location of the protein of interest).

First thing is to burst the cell--> using a homogenizer--> squashes the cell up and ruptures the cell membrane. Take the mixture of stuff (cell lysate) then centrifuge at 500G for ten minutes. Pellet is the heaviest material (nuclear fraction). Supernatant contains everything else. Then 10000g for 20 minutes--> pellet out the mitochondria. Then at 100000 for 1 hour--> pellet the microsomal fraction (soluble proteins).

Homogenation is a very gentle approach, works very well with mammalian cell--> does not work for other common types used in labs (*E. coli*). *E. coli* is gram negative, contains 2 membranes. 2 membranes make the *E. coli* much more robust and cannot be disrupted by a cell homogenizer. Easiest thing to do is to use Freeze/Thaw cycle. Freeze helps to lyse the cell--> ice crystals form and this will rupture the membrane. This works very well with small volumes; with larger volumes there is a different approach--> **osmotic shock**. Use a hypo or hyper tonic solution--> so a lot of material will either come out or get in the cell--> wrinkle the cell--> much weaker cell structure (changing the concentration of the solution). Often is combined with a technique called sonication: take the samples and irradiate with high energy soundwaves (Sonic probe). Add molecules called protease inhibitors so they block the ability of proteases to chew the target proteins (because if you lyse the cell proteins that normally would not be in contact would stay together). We do not heat shock too much otherwise the proteins would be not stable--> no higher than 4 degrees because proteins are heat sensitive. Yeast cells have a plasma membrane and then the cell wall--> really tough exterior and so difficult to lyse. Use the french press.

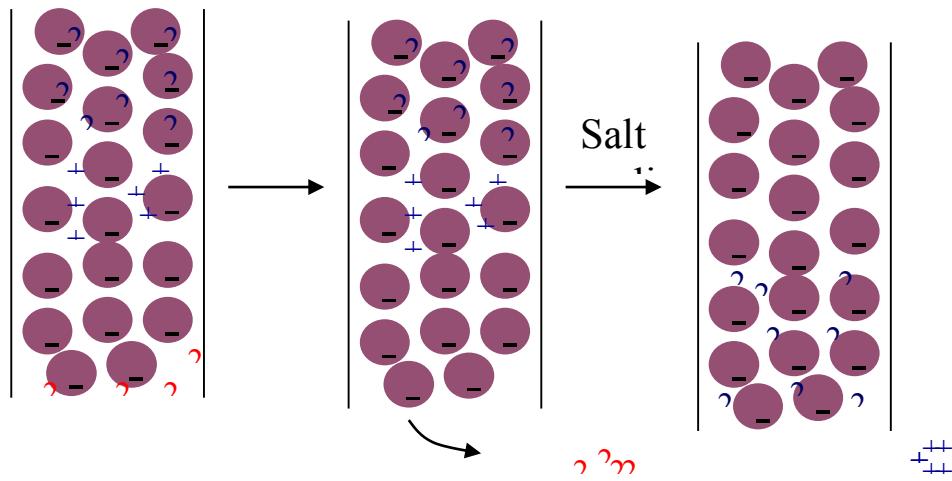
PROTEIN SEPARATION

After cell lysate--> take the supernatant and separate it using chromatography--> separate proteins on the basis of charge, size, solubility and affinity for a particular ion/biological molecule.

- **GEL FILTRATION** (or size exclusion chromatography)--> separates peptides based on their sizes. There is a column packed with porous beads, usually a cross linked polysaccharide material. Small molecules penetrate the beads (slower passage), larger molecules cannot enter the pores and pass around the beads and flow through the column much quicker--> they will elute (come out) much faster.

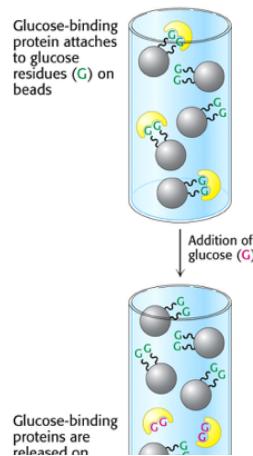
Monitor what is coming through and out with a spectrophotometer. As proteins come out we see increase in absorbance at 280--> where the proteins are as it comes out the column. Large molecules will come out first; then smaller molecules come out. Proteins do not form interaction with the gel, is used extensively as a purification technique but also very good to estimate the size of a protein.

- **High pressure liquid chromatography**--> uses very high pressure and non compressible material and strong metal columns--> get much better resolution/separation of protein beads.
- **ION EXCHANGE CHROMATOGRAPHY**--> separates proteins on the basis of their charge. Separation of peptides in a cationic exchange column. Need a negatively charged resin. At pH 7 some molecules are pushed out of the column--> repelled by the negatively charged resin (so it is negative). Neutral molecules (protein at pI) won't form interaction with the resin beads--> will elute but take a bit longer. The remaining sample are positively charged and form interactions with the negative resin. Then add gradients--> increase concentration slowly over time--> salt will dissociate in sodium ions and chloride--> sodium will compete to form interaction with the resin beads --> compete with proteins positive. Changing the pH also is used to elute proteins (rather than salt gradient). Change pH make proteins neutrals and make the protein fall out from the column. For cationic exchange can be used Carboxymethyl group.

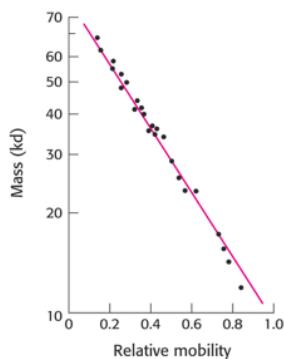


These techniques work only if proteins are already purified. Because there would be lots of different proteins with almost the same charge and size. For this reason it is used affinity chromatography--> based on the affinity of a protein for another molecule (example is glucose-binding protein)

- **AFFINITY CHROMATOGRAPHY**--> Based on the affinity of the protein of interest and some other factor immobilised on a column.



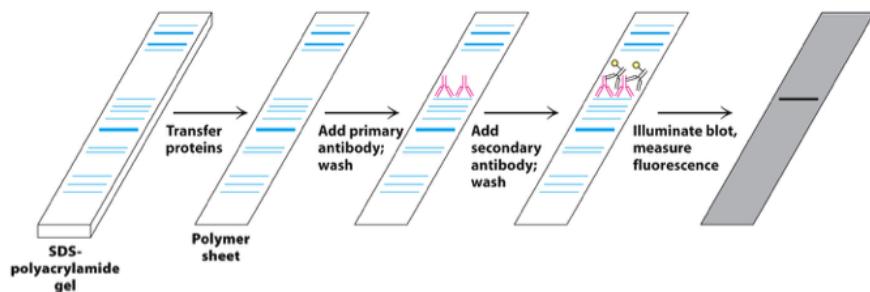
- **ELECTROPHORESIS:** SDS-PAGE (Sodium dodecyl sulphate - poly acrylamide gel electrophoresis) --> it monitors the progress of the purification process. Checks the purity of the final product and gives an idea of the molecular weight of a protein. Separation of the proteins in the samples occurs in a gel matrix that acts as a molecular sieve. Small molecules travel faster in the gel than large ones. Proteins are treated with SDS(detergent) and Beta-mercaptoethanol--> are reducing molecules. Disrupts h bonds and destroys the hydrophobic interactions that maintain the overall fold of the protein.>> binds to the hydrophobic regions of the protein --> reducing agent breaks any disulfide bridge--> overall effect is that the whole molecule is completely unraveled (to a nascent polypeptide chain). SDS also coats the hydrophobic region of the molecule with a relatively even charge . Under these conditions there is a linear relationship between the log of the MW and the relative distance of migration of the protein-SDS micelle



Coomassie blue stain--> must have molecular weight markers. Indication of the size of the protein of interest. Binds specific to charge residues, is a sensitive technique.

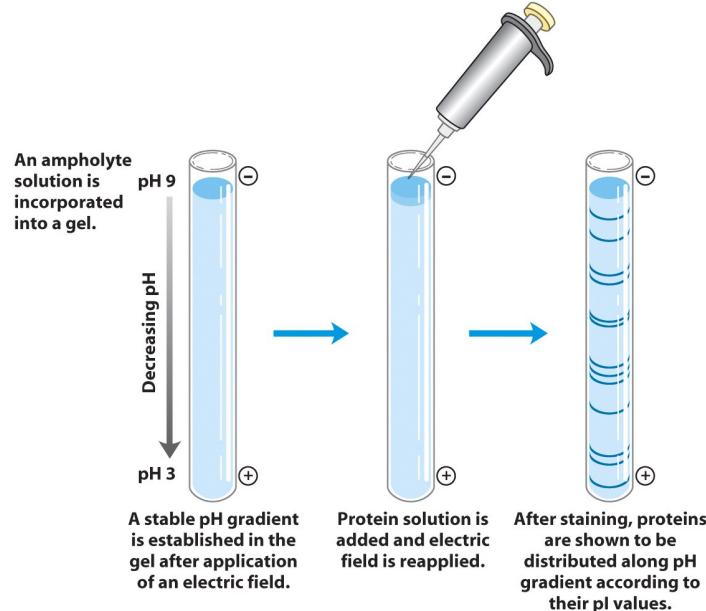
Silver Stain--> all proteins will bind silver ions then add reducing agent in the gel and precipitates silver ions to silver metals.

- **WESTERN BLOTTING:** uses protein specific antibodies. Still need to separate proteins using SDS analysis. The gel is not stained when doing western blot analysis. Transfers protein to a polymer sheet. Then add primary antibody that will recognize very specifically that specific sequence (only bind a specific tag on the gel). Then add a second antibody, that binds primary antibody.



At the end there is a chemiluminescent signal. Must run molecular weight markers. Keep the information intact when doing the transfer in the blot. Allow to specifically detect protein in a very mixed sample. Useful once you have purified sample

Isoelectric Focusing --> electrophoretic technique that divide on the basis of charge. Ampholyte solution is incorporated into a gel, a stable pH gradient is established in the gel after application of an electric field. Protein solution is added and electric field is reapplied. After staining, proteins are shown to be distributed along pH gradient according to their pI values. It is not very specific however because different proteins can have very similar pI values.



2-D GEL ELECTROPHORESIS: Powerful technique for the separation of peptides and proteins Separates on the basis of both charge and mass.

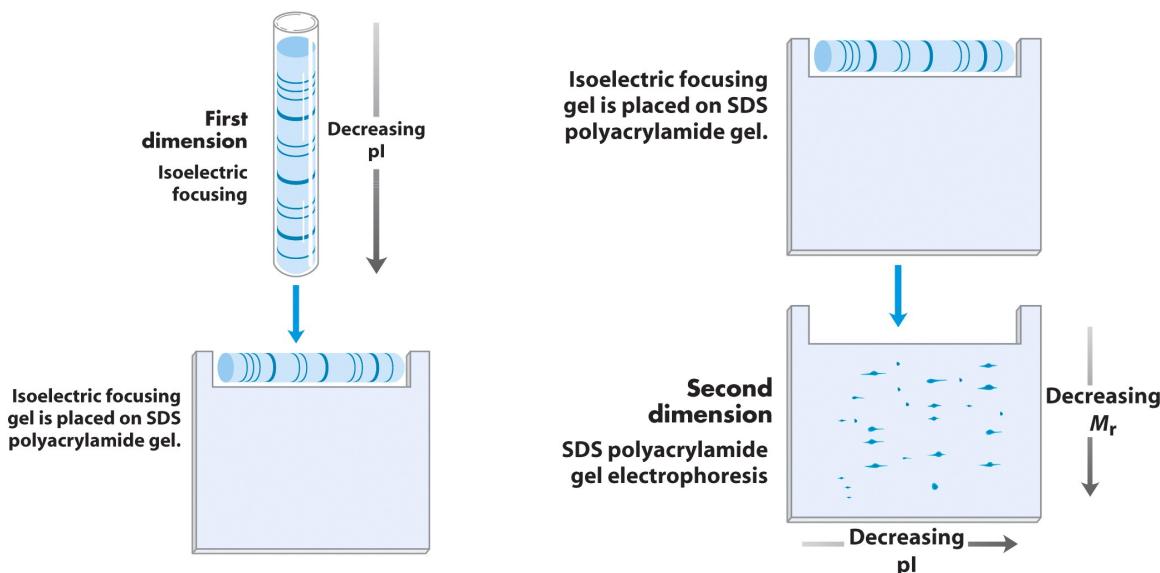
1. Mobility directly proportional to charge
2. Mobility to inversely proportional to mass

Possible to manipulate mobility of peptides by changing pH of buffers

1st dimension separation based on the pI of proteins/ peptides.

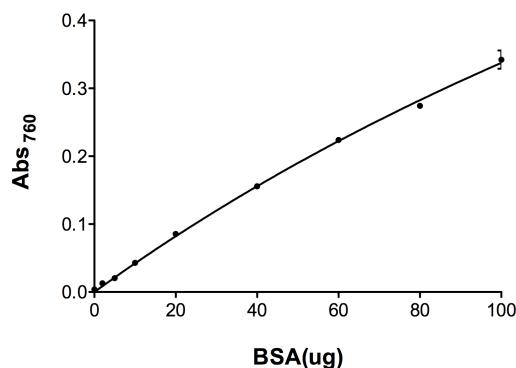
Run the sample on a gel with a pH gradient.

In a pH gradient under the influence of an electrical field a protein will move to the position in the gradient where its net charge is zero.



LOWRY ASSAY:

- Relies on chelation of Cu^{2+} ions by peptide bonds
- Reduced to Cu^+
- Allows reduction of phosphomolybdic-tungstic acid (in Lowry reagent)
- Distinctive blue colour
- Quantified by measuring absorbance at 760 nm
- Under alkaline conditions the divalent copper ion forms a complex with peptide bonds in which it is reduced to a monovalent ion. Monovalent copper ion and the radical groups of tyrosine, tryptophan, and cysteine react with Folin reagent to produce an unstable product that becomes reduced to molybdenum/tungsten blue.



Standard Curve:
Example

$$\begin{aligned} A_{760} &= 0.25 \text{ for } 100 \mu\text{l of a } 5 \text{ ml solution} \\ \text{From the curve} &= 70 \mu\text{g}/100 \mu\text{l} \\ &= 700 \mu\text{g/ml} \\ \text{Total protein} &= 3.5 \text{ mg} \end{aligned}$$

AMINO ACID COMPOSITION:

Determination of amino acid composition:

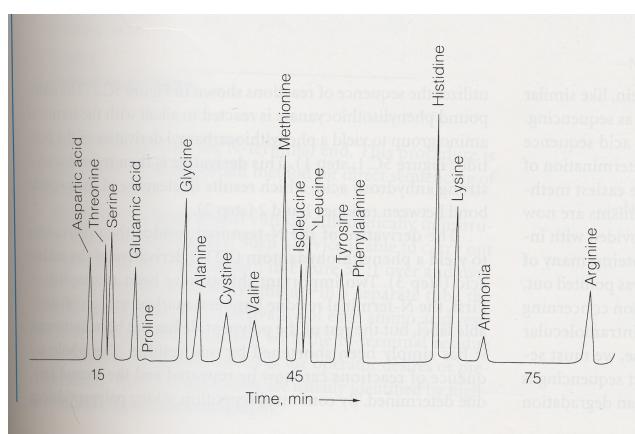
- i) hydrolysis of the protein to its constituent amino acids
- ii) separation of the amino acids in the mixture
- iii) quantitation of the individual amino acids

Method:

Protein is dissolved in 6 M HCl and the solution sealed in an evacuated ampoule.

Heated at 105-110°C for about 24 hrs (peptide bonds between the residues are completely hydrolysed)

Sample separated into the constituent amino acids using a cation exchange column

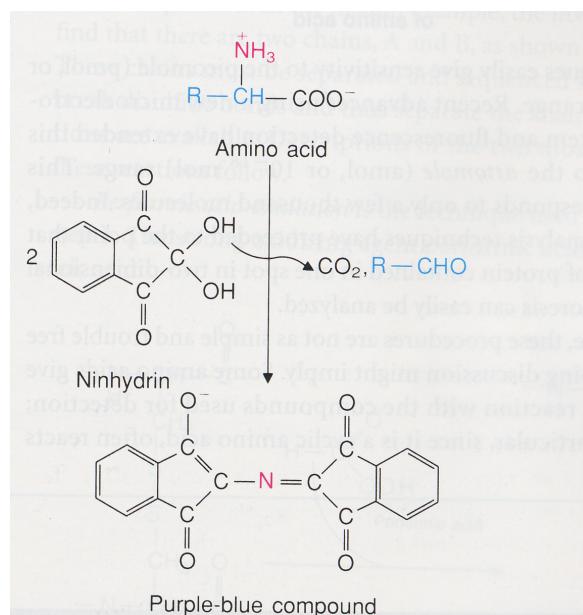


Elution order of Amino acids

DETECTION AND QUANTIFICATION OF AMINO ACIDS:

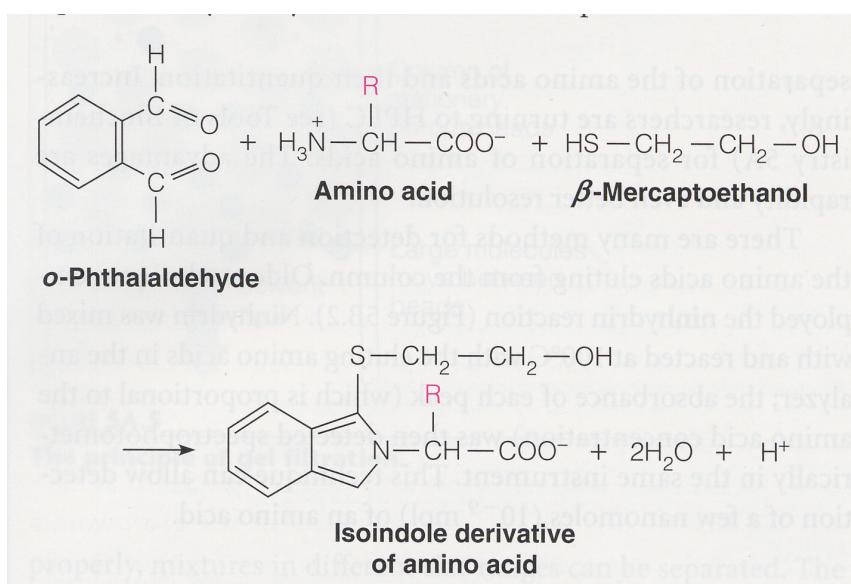
- **Ninhydrin reaction**

Ninhydrin mixed and reacted at 100°C with each peak. The absorbance of each peak (proportional to the amino acid concentration) detected spectrophotometrically



- **Fluorescence detection**

Amino acids reacted with a fluorescent compound e.g *o*-phthalaldehyde to give a fluorescent complex (Very sensitive method, can detect amol amounts)



PROBLEMS:

- Some amino acids have problems reacting with the compounds used for detection.
- Some amino acids tend to be destroyed by the harsh hydrolysis method.
- Tryptophan must be detected using a separate method based on OD_{280}
- Ser, Thr and Tyr also tend to be degraded, can be protected using phenol or thiol.

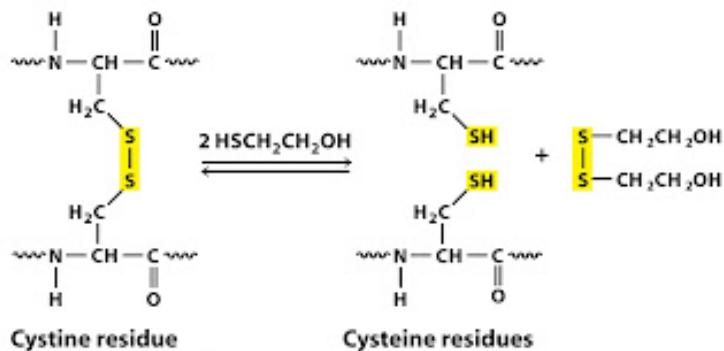
DISULPHIDE BRIDGES CLEAVAGE:

Reductive alkylation: Cleavage of disulphides and irreversible denaturation of the protein ready for further study

1. First step: Reduction

Most common reducing agent is β -mercaptoethanol

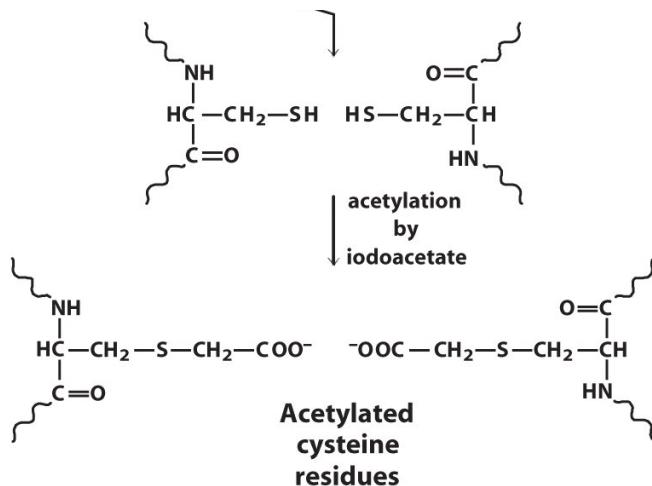
The equilibrium constants are near unity and normally a large excess of thiol is required to force the reaction to completion.



2. Second step: **Alkylation**

Leaves free sulphhydryl groups, potential to reoxidise, reforming the disulphide bond.

Prevented by immediate alkylation using most commonly iodoacetic acid (blocking group). Iodoacetic acid must be added to the protein solution in excess



ENZYMATIC METHOD FOR FRAGMENTATION OF PROTEINS:

Downstream analysis often requires that proteins are chopped into small bits

Proteases cleave proteins at specific sites

TRYPSIN: Serine protease, mwt=28 kDa

Cuts the protein after Lys-, Arg- in the sequence

Optimum activity at pH 8.0 at a enzyme:substrate of 1:50.

Trypsin prepared from bovine pancreas can be contaminated with other enzymes, particularly CHYMOTRYPSIN-can complicate cleavage studies

Chymotrypsin cleaves after Tyr, Trp, Phe and Leu.

Generally this problem is dealt with by purifying the trypsin through a soya bean trypsin inhibitor column:

Key characteristic: peptides produced by tryptic cleavage all have C-terminal Lys or Arg.
Except.....the peptide which was originally at the C-terminus of the protein:

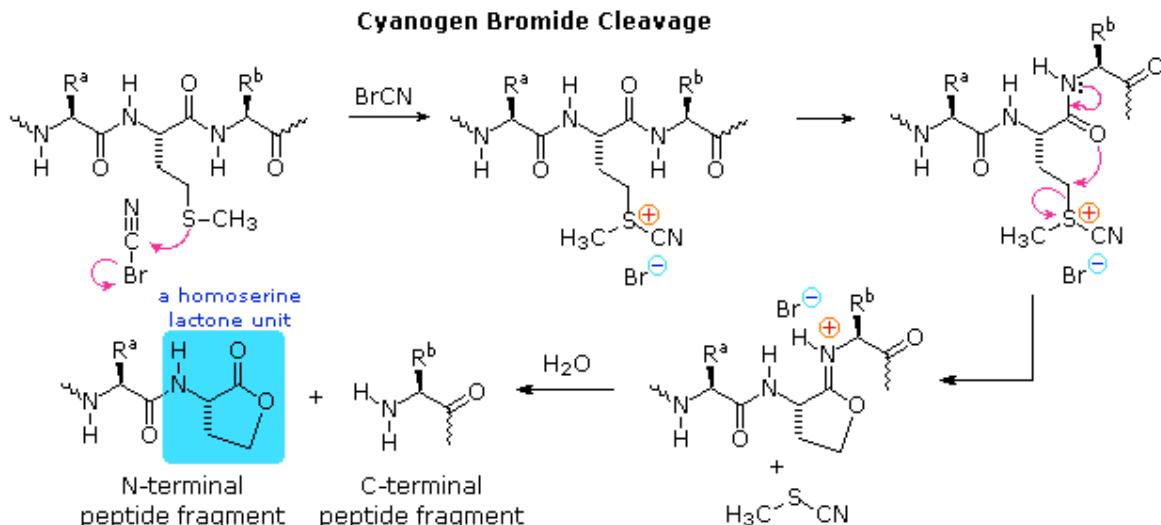
Rate of hydrolysis of various peptide bonds dependent upon steric and charge factors

1. Lys or Arg followed by negatively charged Aas such as Asp or Glu are cleaved relatively slowly.
2. Lys or Arg followed by Pro are not cleaved at all (v.v. slowly over 24 hours)

CHEMICAL METHOD FOR FRAGMENTATION OF PROTEINS:

Cyanogen bromide cleaves at the C-terminal side of Met, converting methionine to homoserine lactone.

Most specific chemical method for peptide bond cleavage, because only one codon corresponds to Met the proportion of Met in a protein is usually quite low.



IDENTIFICATION OF PROTEINS:

Two methods for protein identification

1. Mass spectrometry -fragments DO NOT need to be separated
 2. N-terminal sequencing -fragments DO need to be separated
- Many genomes have been completely sequenced
 - Possible to map the trypsin cleavage sites within the coded proteins
 - Predict the number and size of fragments for a particular protein
 - A protein which is then produced and purified can then be identified by trypsin cleavage followed by mass spec.

If mass spec is not suitable for analysis of the protein then N-term sequencing must be used

EDMAN DEGRADATION:

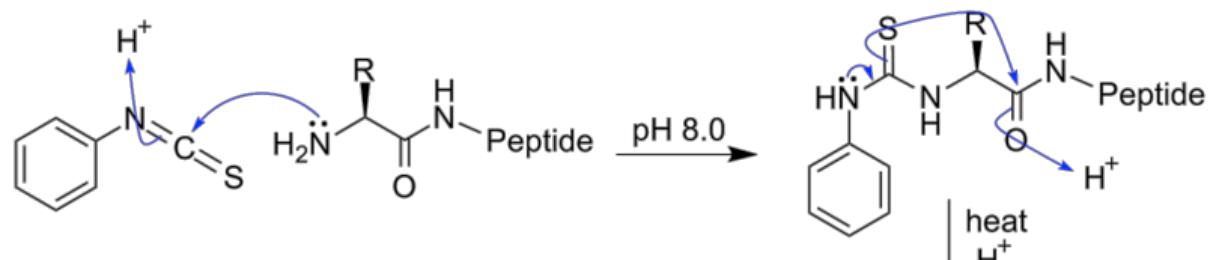
Stepwise procedure:

- i) Specific removal of the N-terminal amino acids of the peptide/protein
- ii) Identifying the amino acid derivative formed
- iii) Carrying out another cycle of removal of the N-terminal amino acid

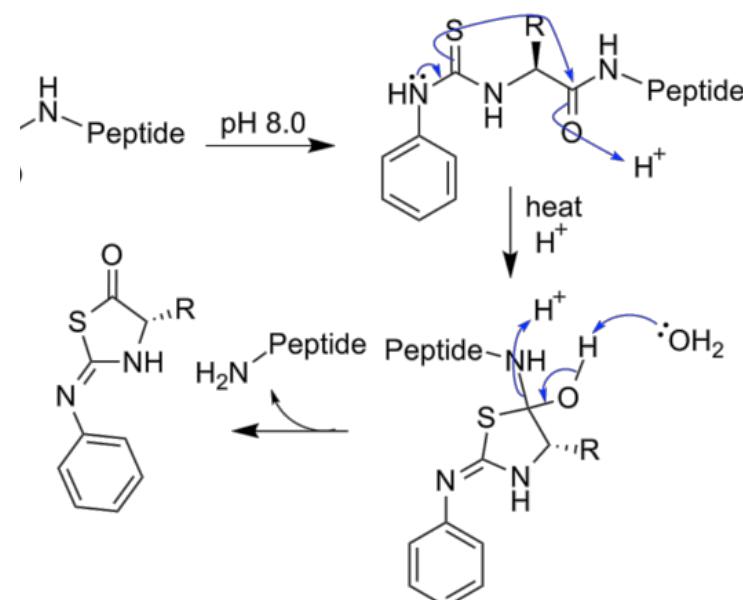
Theoretically continues until the C-terminus of the protein is reached

Manual or automated N-terminal sequencing possible.

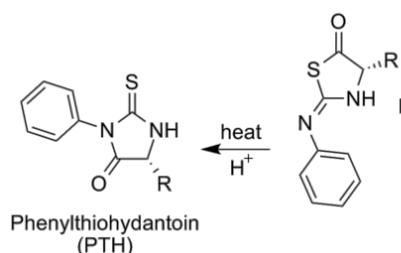
1. Phenylisothiocyanate (PTC) is reacted in alkali with the terminal amino group to yield a phenylthiocarbomyl peptide derivative.



2. The phenylthiocarbamyl derivative is then treated with anhydrous trifluoroacetic acid, which results in cleavage of the peptide bond between residues 1 and 2



The derivative of the N-terminal amino acid residue is extracted using organic solvents
Converted to a more stable phenylthiohydantoin (PTH) derivative in the presence of acid. At this point the amino acid can be detected



Two important things have been accomplished at this stage:

- The N-terminal residue has been marked with an identifiable label
- The rest of the polypeptide has not been destroyed it has simply been shortened by one residue
- Whole sequence of reactions can now be repeated.
- The PTH derivative of each individual amino acid is collected in a separate tube and can be analysed by HPLC.

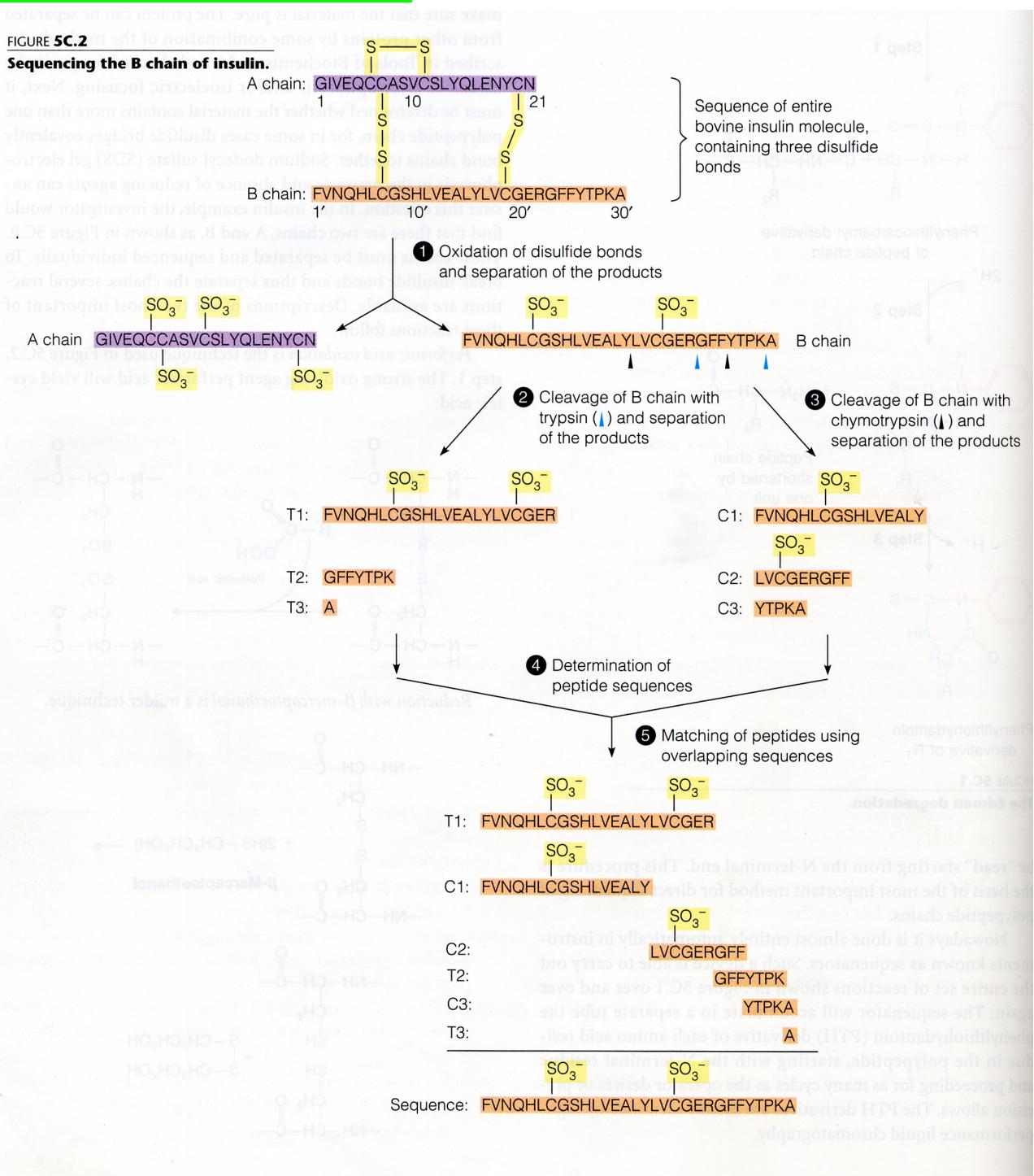
LIMITATIONS OF EDMAN DEGRADATION:

- Cyclical derivitisation doesn't go to completion
- Proportion of the peptide which does not release amino acid
- If efficiency of release is 98%-then proportion of "correct" amino acid released after 60 rounds would be $(0.98)^{60}$ or 0.3
- Can get information for small regions of peptide
- Repetitive yield problem is reason proteins must be cleaved into small peptide fragments prior to sequence analysis.
- **N-terminal amino group must be available**
- **Proline an issue**
- **Post-translation modification (glycosylation, acetylation, phosphohrylation)**
- **Disulphide bridges**
- 10 amino acids is a usual yield and this is enough information to identify a known protein.
- This now the main way in which the Edman cycle is employed
- Although the method was initially developed as a means of completely sequencing a protein

C-TERMINAL SEQUENCING:

- N-terminal methods are rarely successful in providing the complete sequence
- C-terminal sequencing, 1-4 amino acid residues will give enough information to construct the whole protein.
- 3 methods:
- Hydrazinolysis
- Repetitive degradation (Stark Method)
- Carboxypeptidases (Enzyme Method)
- **Cyanogen Bromide** treatment yielded the following fragment:
Asp-Leu-Val-Thr-Lys-Gly-Pro-Ser-His
- **Tryptic Cleavage** yielded the following fragment:
Glu-Val-Phe-His-Met-Asp-Leu-Val-Thr-Lys
- **Chymotryptic** cleavage yielded the following fragment
Ala-Pro-Glu-Thr-Gly-Arg-Glu-Val-Phe

SEQUENCE DETERMINATION OF INSULIN: Frederick Sanger, 1958 Nobel Prize



USES OF SEQUENCE INFORMATION:

- Classification and possible identification of proteins
- Assess homology of protein with related proteins from other organisms
- Model structure
- Reveal how proteins evolved
- Find related proteins in genome sequences

STRUCTURE DETERMINATION: 3-D structure of a protein provides information

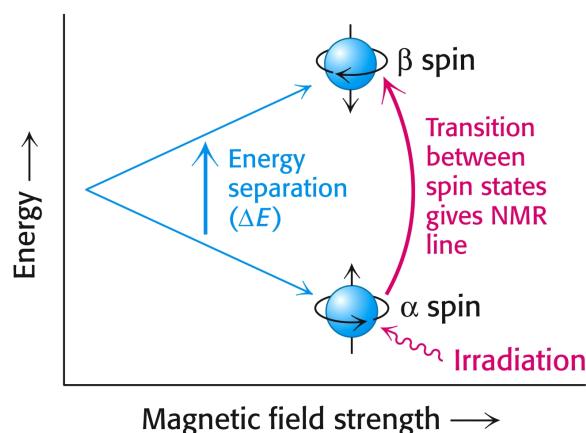
- proteins folding
- recognise other proteins
- catalyse chemical reaction
- help us understand disease states

Two major methods

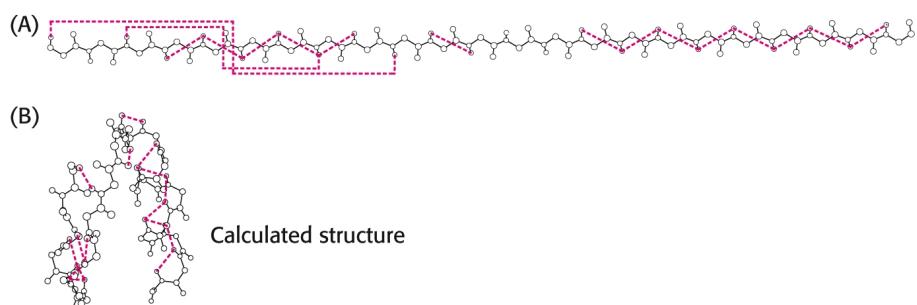
- i. **NMR**
- ii. **X-ray crystallography**

1. NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY:

- Does not require crystals
- Can obtain high resolution dynamic information in solution
- Requires high concentrations of protein (1mM)
- NMR uses the magnetic properties of atomic nuclei
- Certain nuclei have spin
- Relies on absorption of energy to produce a transition between two states
- Difference between energy states is low
- Difference in energy measured by resonance technique
- Sample irradiated by radiation (RF) of frequency matching energy difference



- Energy differences give information about the chemical surroundings of individual spin nuclei.
- Flow of electrons around a magnetic nucleus generates a small local magnetic field opposing applied field.
- The degree of shielding depends on electron density
- Nuclei in different environments will resonate at slightly different RFs
- Also look at the ways in which spin nuclei affect neighbouring nuclei
- If enough information can be obtained for different sets of neighbouring nuclei- obtain a 3D structure

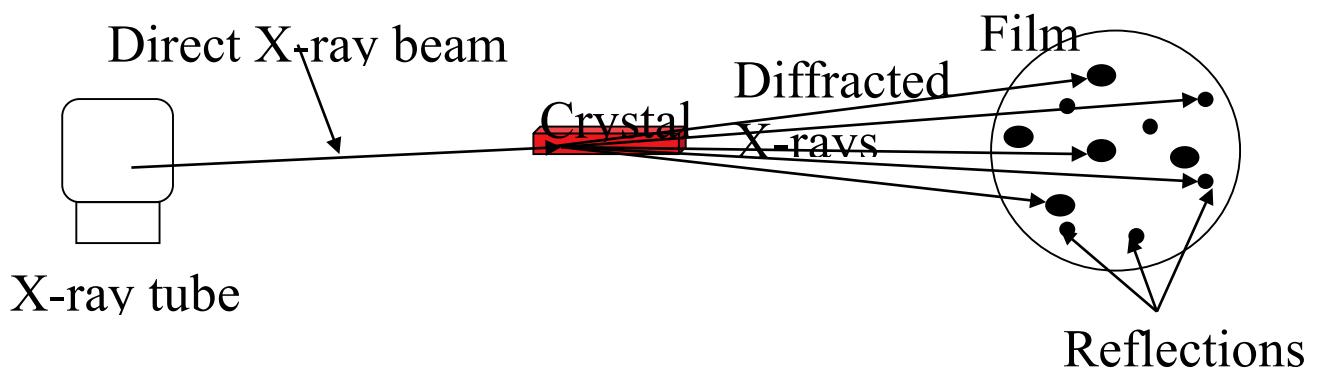


2. X-RAY CRYSTALLOGRAPHY:

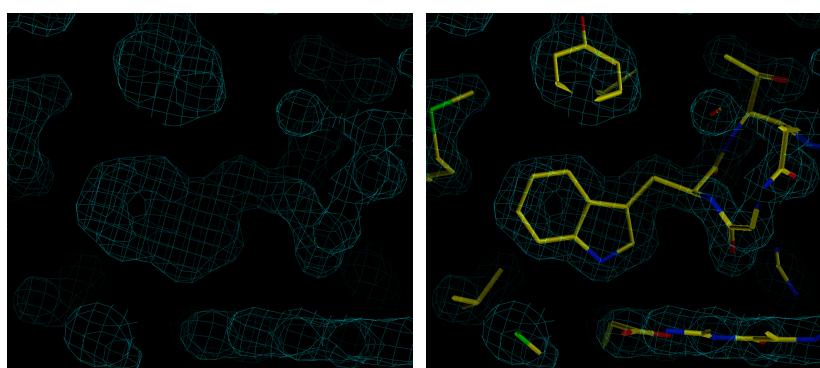
- Best method for obtaining high resolution structural data
- Structures deposited in PDB
- Wavelength of X-rays is about the same length as covalent bonds
- Disadvantages: Lots of protein required
- Difficult to obtain protein crystals

Electrons in the protein scatter the X-ray

The amplitude of the wave scattered by an atom is proportional to the number of electrons
Each molecule in the crystal adds to the signals produced-amplification. Electrons that surround every atom scatter the x ray. Sulfur atom diffracts differently than carbon or oxygen... a single molecule should produce a pattern of reflection used to measure diffraction... however thousand of molecule are needed to combine and give a distinct diffraction pattern.



- An electron density is calculated. Then fit individual bits of amino acids in the electron density. The higher the resolution, the better the diffraction ,the better the model.



SUMMARY:

- Proteins can be purified on the basis of charge, size, solubility and specific interaction with another protein or ion.
- Proteins can be analysed using SDS-PAGE and detected by Western blot.
- It is possible to modify a protein by breaking the disulphide bridges
- Proteins can be specifically cleaved using enzymatic and chemical methods
- Proteins can be identified using mass spectroscopy or N-term sequencing
- Protein structure can be determined by NMR and X-ray crystallography