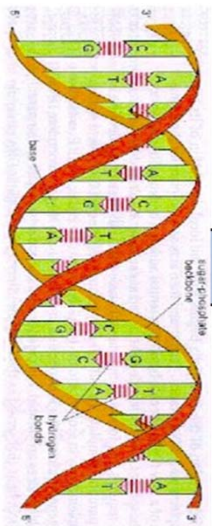
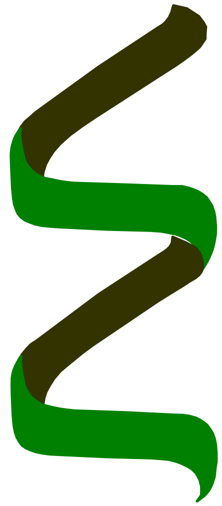
What Are proteins? How are they made?

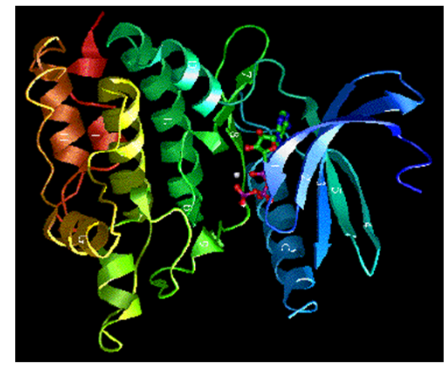
Central Dogma



Transcription



Translation



Gene (DNA)

mRNA

Protein

Hair (Keratins)

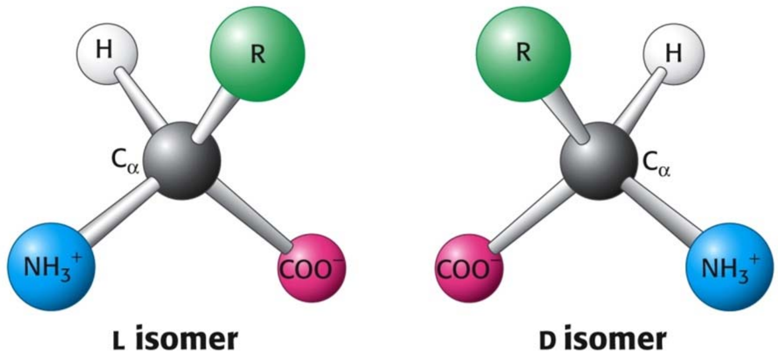
Silk (fibroin)

Enzymes

Proteins

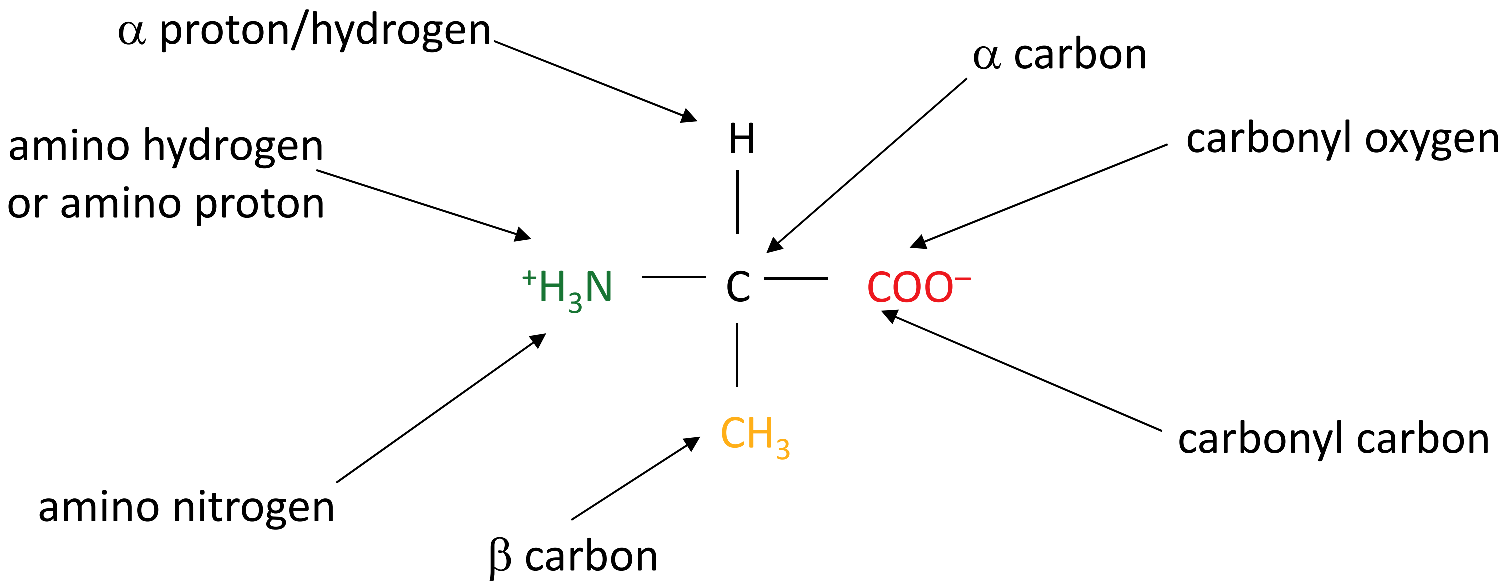
Amino Acids

Amino acids: Basic constituent of proteins



- Proteins are composed of amino acids
- All 20 amino acids have central carbon atom ($C\alpha$) to which are attached a hydrogen atom, an amino group (NH_2) group and a carboxyl group (COOH).The L amino acids are so designated if one looks downward from hydrogen atom to $C\alpha$ then **CO,R and N** substituents are in clockwise direction.
- L amino acids are found in most of the naturally occurring proteins.

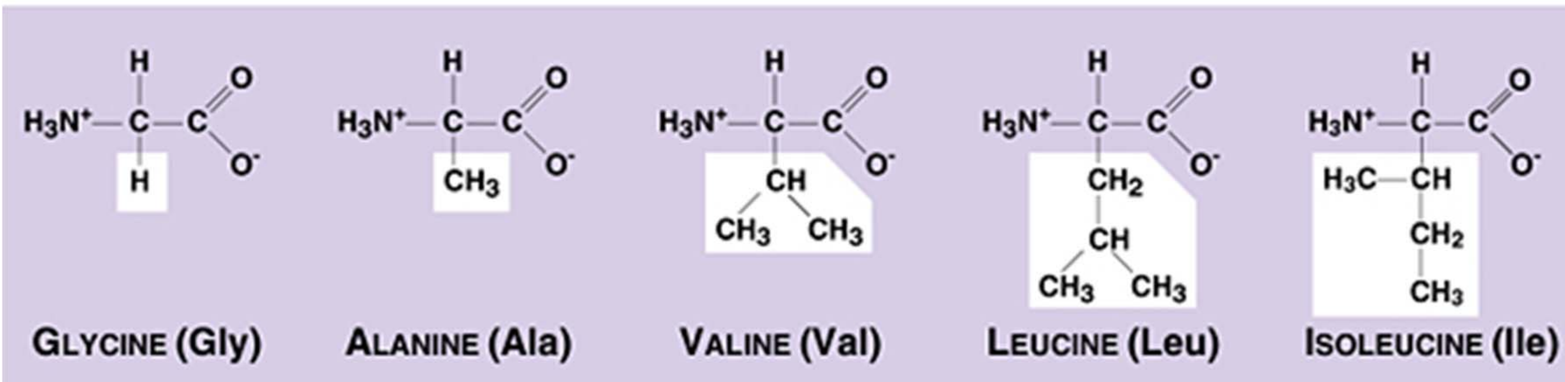
Naming amino acids



alanine
(a polar amino acid)
R = CH_3
(a methyl functional group)

Non-polar amino acids

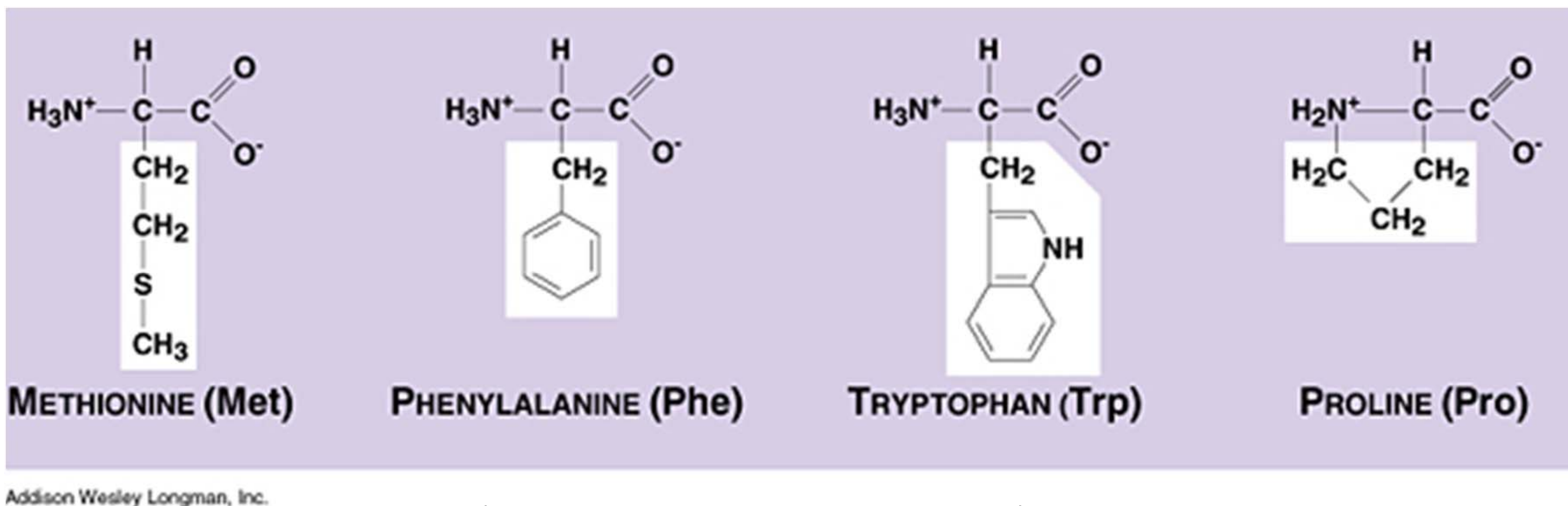
- R group consists of carbon chains



Leucine and isoleucine
are structural isomers

Nonpolar amino acids

- R group consists of carbon chains



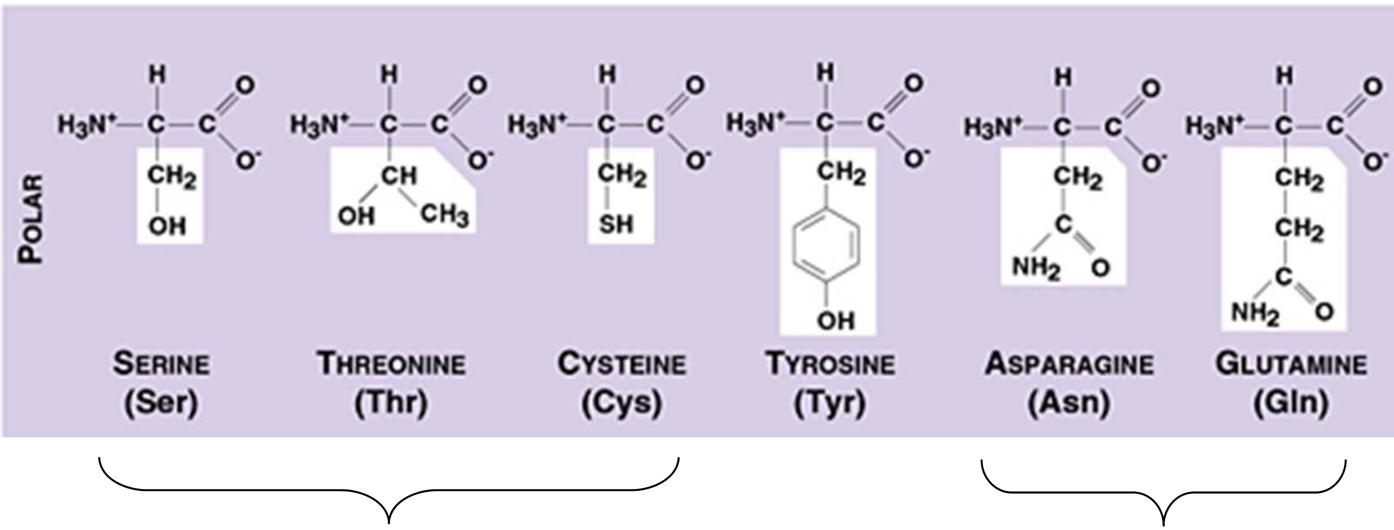
Methionine has a sulphur atom in its sidechain

Phenylalanine and tryptophan have aromatic rings

Proline has its R group bound to the amino nitrogen to form a ring network

Polar amino acids

- R group consists of carbon, oxygen and nitrogen atoms together they make the sidechain more hydrophilic

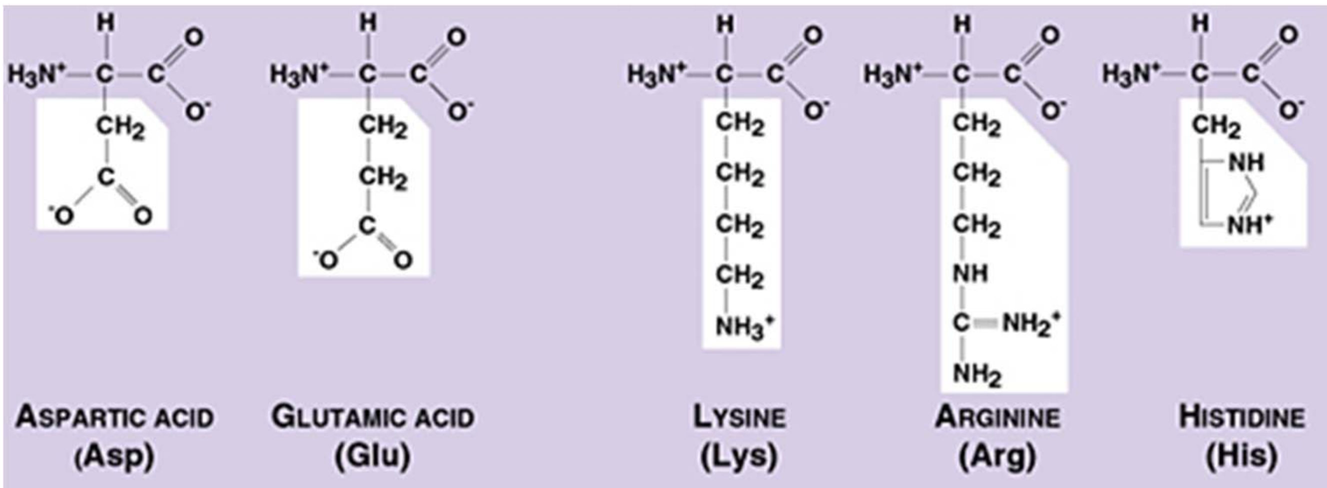


Ser and thr are a mix of carbon chains and hydroxyl functional groups (-OH). Cysteine has a thiol group (-SH) which is otherwise structurally similar to serine but not chemically similar

Asn and gln have an amide functional group

Charged amino acids

- R group has a charge at physiological pH (7.4). pK of the charged groups vary



carboxyl group | carboxyl group | amino group | guanidinio group | imidazole group

Describing amino acids

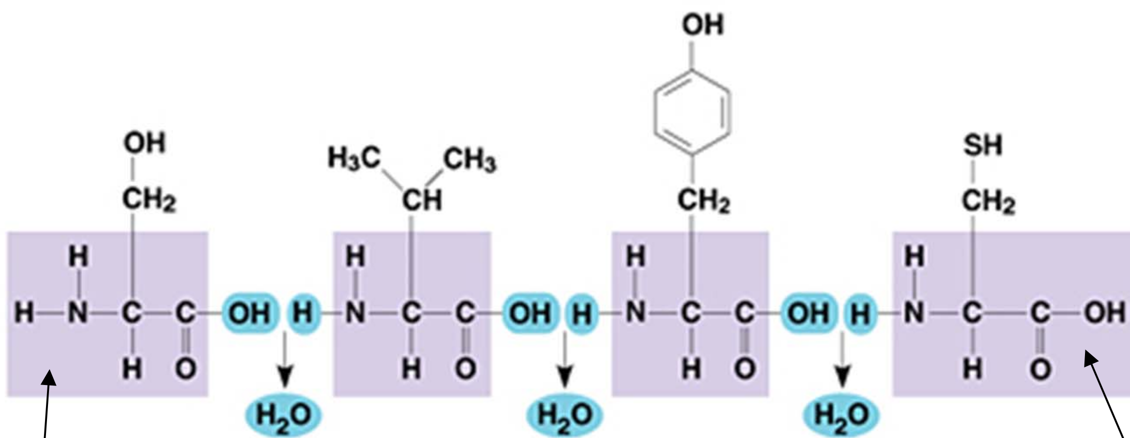
- amino acids have a full name (glycine), a short three-letter name (gly) and an even shorter one-letter name (G)

nonpolar	
polar	
acidic (negative charge)	
basic (positive charge)	

A	ala	alanine
C	cys	cysteine
D	asp	asparatic acid
E	glu	glutamic acid
F	phe	phenylalanine
G	gly	glycine
H	his	histidine
I	ile	isoleucine
K	lys	lysine
L	leu	leucine
M	met	methionine
N	asn	asparagine
P	pro	proline
Q	gln	glutamine
R	arg	arginine
S	ser	serine
T	thr	threonine
V	val	valine
W	trp	tryptophan
Y	tyr	tyrosine

Joining amino acids together

- In a cell, a complex set of proteins called a ribosome catalyze a dehydration reaction (loss of water) to join amino acids together



chain extends in this direction

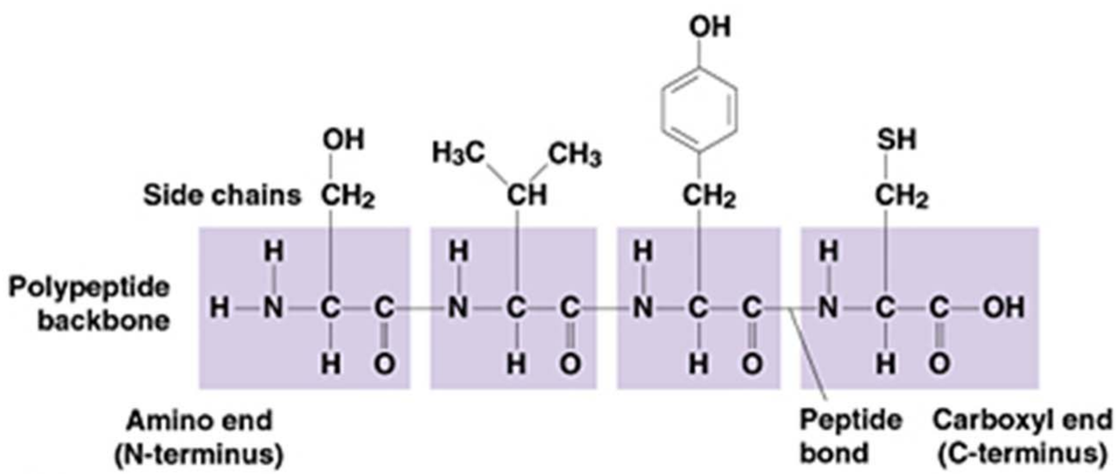
The ribosome does not join an amino acid to this end (the amino end)

loss of water

The ribosome only joins new amino acids to this end (the carboxy end)

Joining amino acids together

• In a cell, a complex set of proteins called a ribosome catalyze a dehydration reaction (loss of water) to join amino acids together



• a peptide bond (like an amide bond C-O-N) joins each amino acid

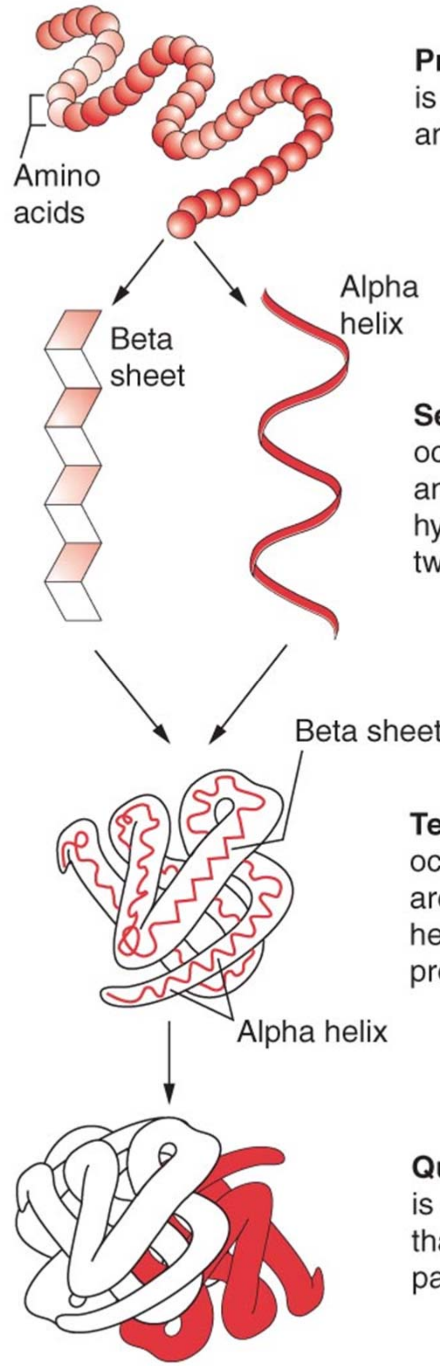
• the invariant purple part of the polypeptide is generally called the backbone

• it's the sidechains that give a protein its unique chemical and structural character

Protein Structures

- Structural Arrangement – four levels
  - **Primary structure** is the sequence in which amino acids are linked together
  - **Secondary structure** occurs when chains of amino acids fold or twist at specific points
    - Alpha helices and beta sheets
  - **Tertiary structures** are formed when secondary structures combine and are bound together
  - **Quaternary structures** are unique, globular, three-dimensional complexes built of several polypeptides

Protein Structures



Primary protein structure
is the sequence of a chain of amino acids

Secondary protein structure
occurs when the sequence of amino acids is linked by weak hydrogen bonds into one or two structural forms

Tertiary protein structure
occurs when certain attractions are present between alpha helixes and beta sheets producing a unique 3D structure

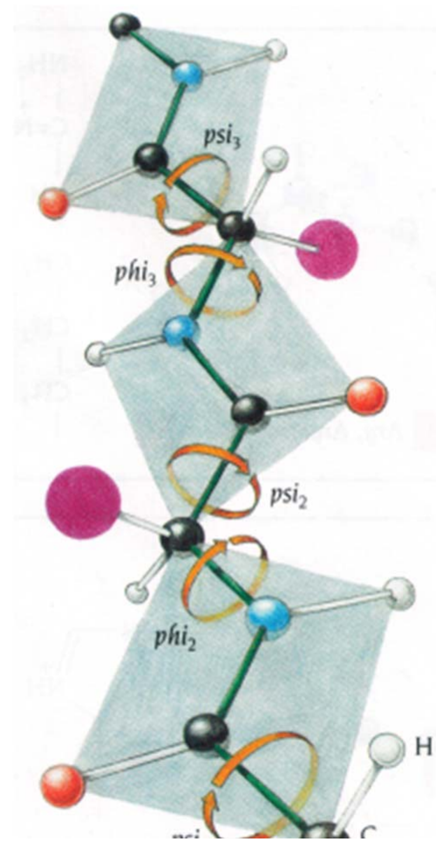
Quaternary protein structure
is a protein consisting of more than one amino acid chain, each participating in the final 3D shape

Peptides as building blocks of protein structure

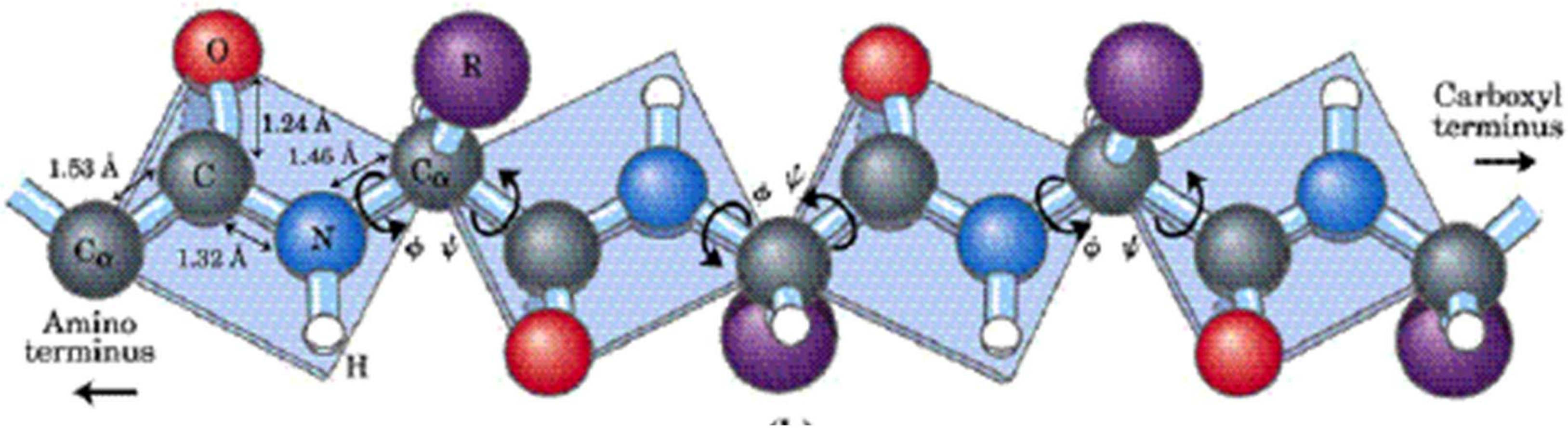
- The polypeptide chain can be divided into peptide units going from one Cα atom to the next Cα atom.

- The peptide units are rigid groups linked to the chain by covalent bonds at the Cα atom, the only degree of freedom is rotation around these bonds.

.



Peptides as building blocks of protein structure

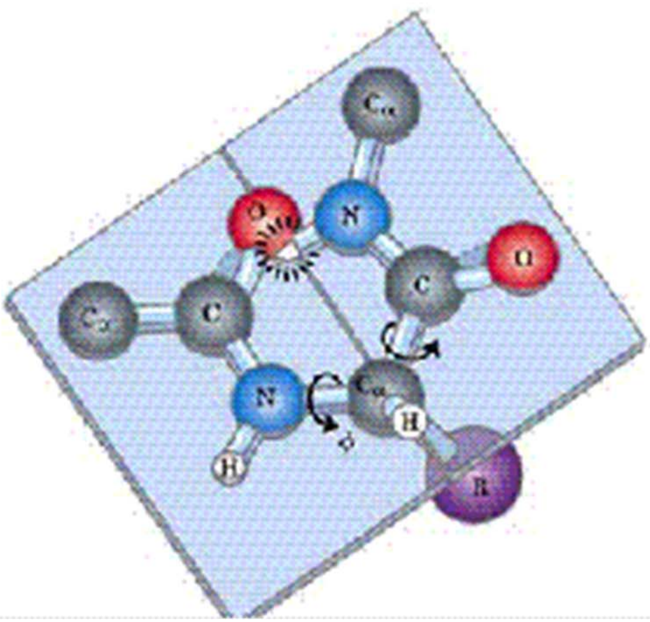


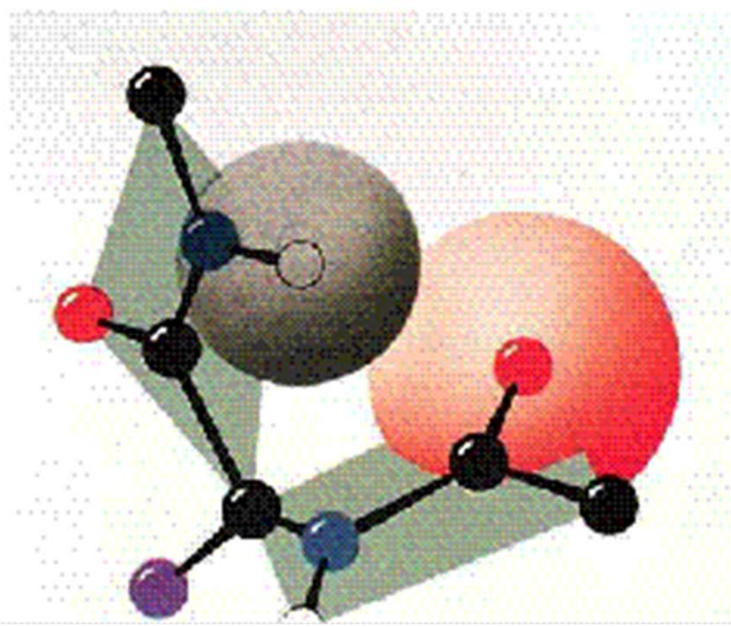
- These bonds are designated as psi (φ) for rotational angle between Cα and N of the amino group and phi (ϕ) for rotational angle between Cα and C of the carboxyl group.

- Each amino acid residue is associated with two conformational or dihedral angles and besides this is there slight contribution by side chain groups also. If all the dihedral angles for each amino acid residue is defined with high accuracy then the conformation of the main chain is completely determined

Steric hindrance

Steric hindrance (van der Waals volume overlap) of back bone and side chain atoms limits back bone flexibility



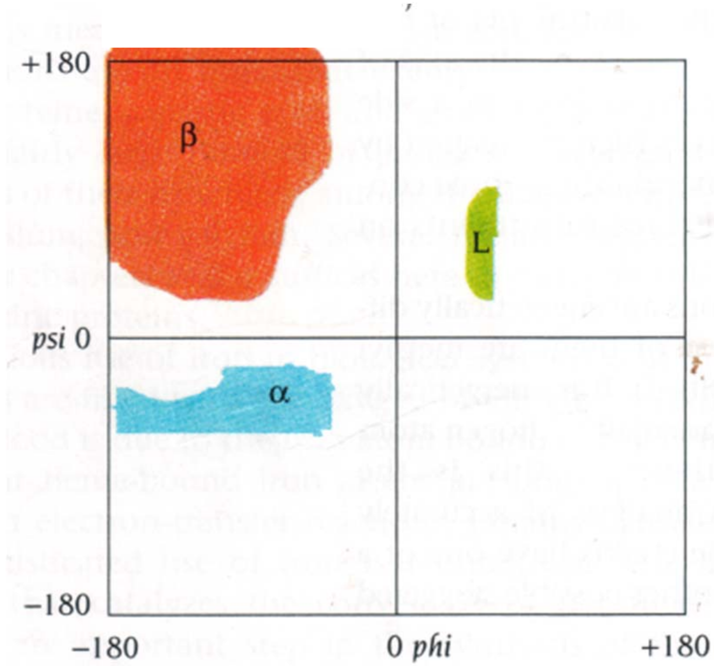


Ramachandran Plot

- Most of the dihedral angles φ and θ are not allowed for an amino acid because of the steric collisions between side chain and main chain.

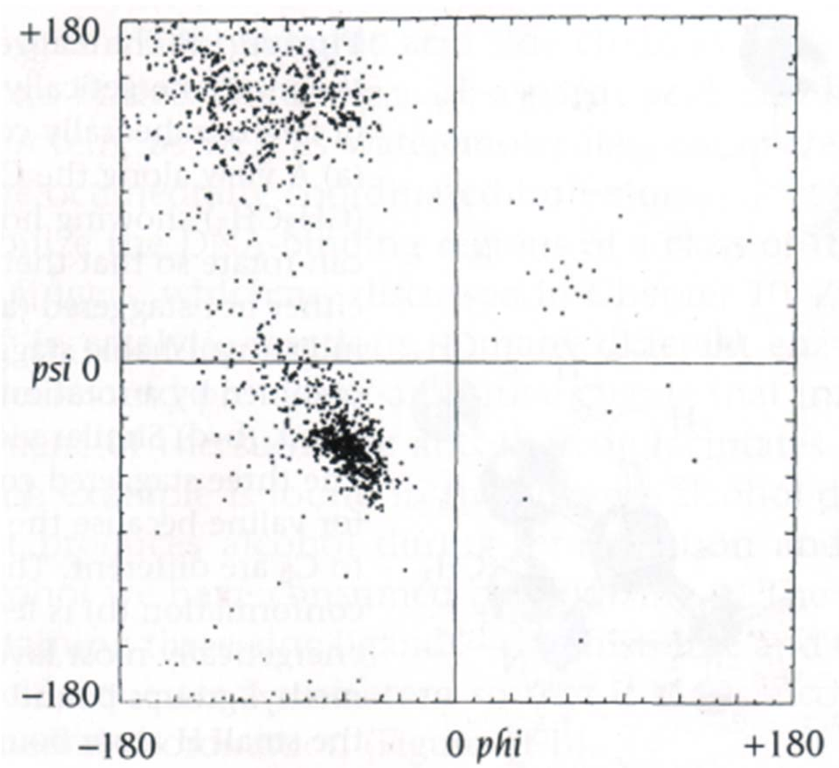
- Based on the Van der waals radius of the amino acids and side chains the allowed combination of dihedral angles for amino acids can be calculated.

- Ramachandran plot depicting sterically allowed dihedral angles as colored areas. the areas denoted as α, β and L correspond to conformational angle found for the usual right-handed α helices, β-strands and left-handed α helices

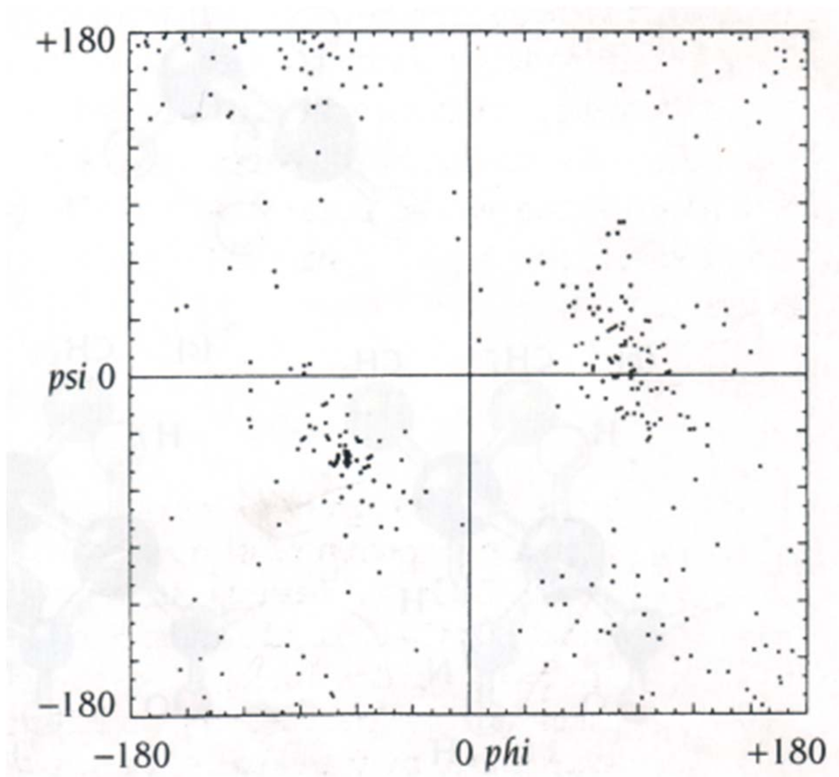


Ramachandran Plot

Observed values of dihedral angles for all residues except glycine in high resolution X-ray structure



Observed values of dihedral angles for glycine



Each point represents φ and θ value for an amino acid residue

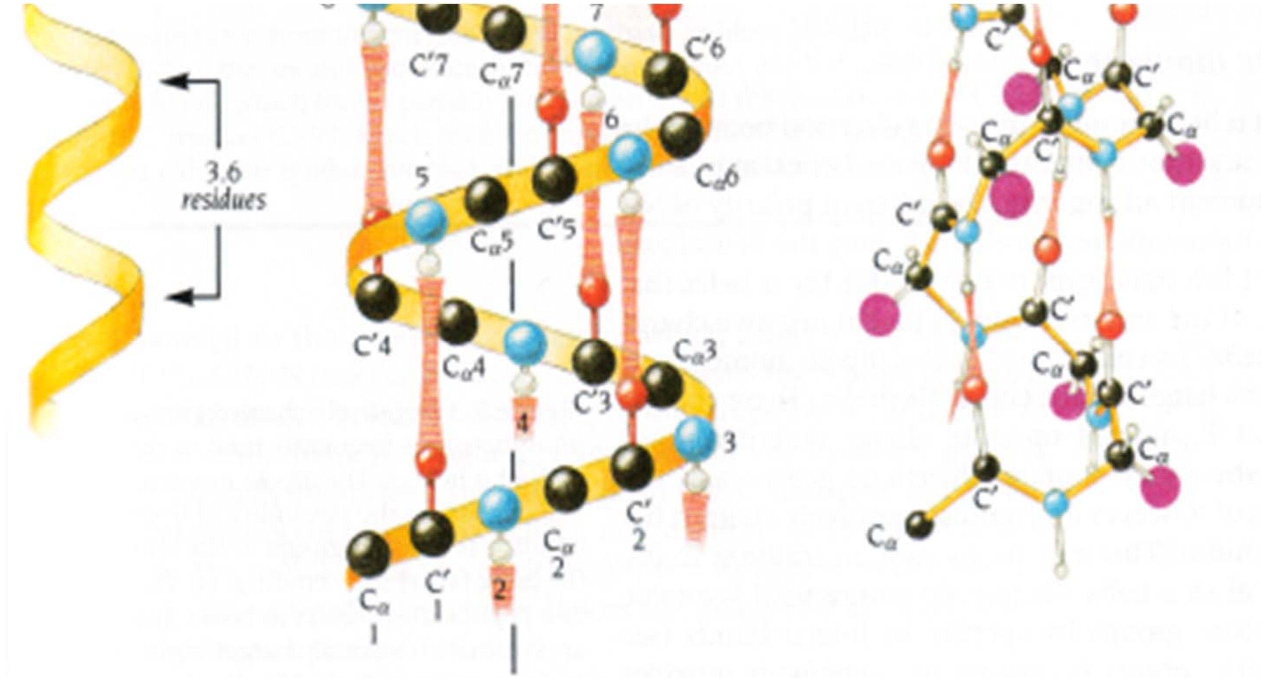
Secondary Structure of Proteins

- Regular arrangements of amino acids
- Allow hydrogen bonding propensities of backbone N-H and C=O groups to be satisfied in the absence of water
- Held together by hydrogen bonds

Two major classes:

- Alpha helices
- Beta sheets

α-Helix



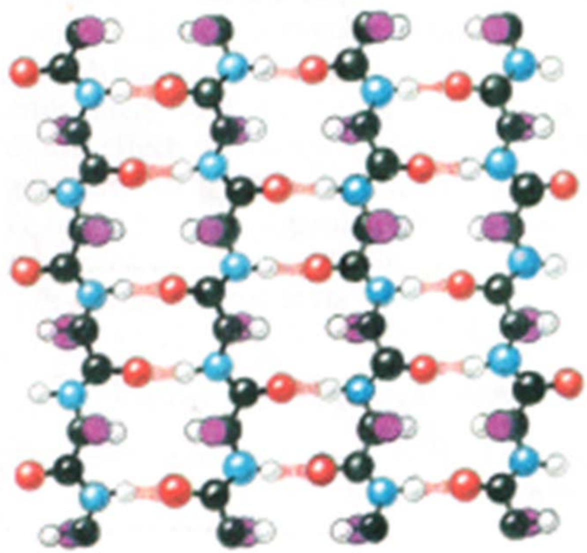
• α-Helix conformation was first predicted by Linus Pauling to be energetically favorable and stable in protein.

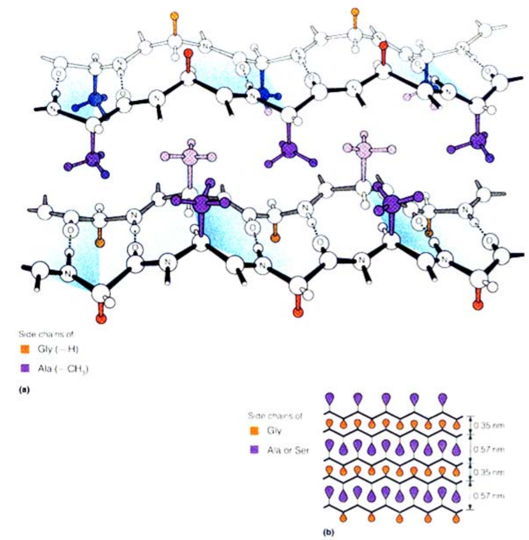
• α-Helices are found in proteins when a stretch of consecutive aa residues all have ϕ. φ angle pair approximately -60° and -50° (upper portion of bottom left quadrant in Ramachandran plot).

α-Helix

- An α-helix has 3.6 residues per turn with hydrogen bonding between C=O of residue n and NH of residue n+4
- In an α-helix, hydrogen bonding potential of all groups is satisfied except for first NH and last CO group.This makes the ends of α-helix polar and consequently are always found on the surface of proteins.
- α-Helix varies from more than 40 residues to 5 residues though average of 10 residues (three turns) are seen.The rise per residue of an α-helix is 1.5 angstrom

β-sheet

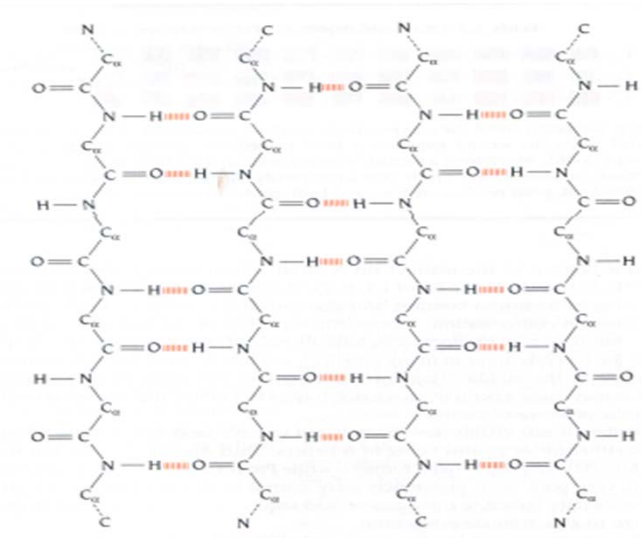




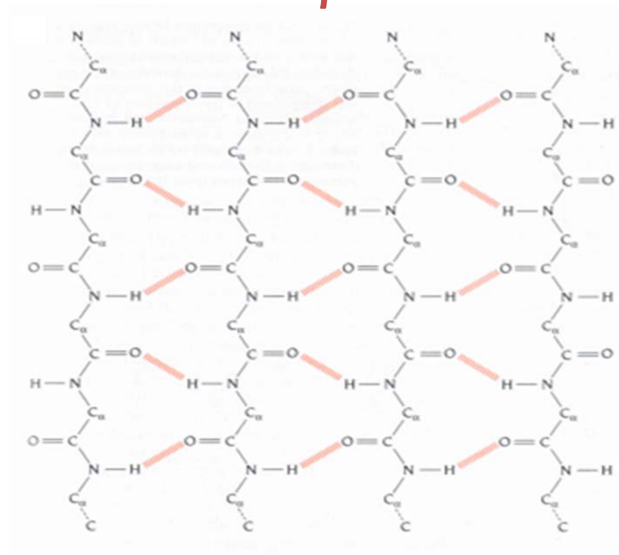
- β-sheet is composed of stacks of β-strands which are usually from 5 to 10 residues long and are aligned adjacent to each other allowing hydrogen bonding between C’=O group of one β-strand and NH group on an adjacent β-strand and vice versa.
- The β-sheet is said to be pleated because Cα successively are little below and little above the plane of β-sheet.
- The side chains follow the pattern of Cα atoms and therefore they point alternately above and below the β-sheet.

β-sheet

Anti-parallel β-sheet



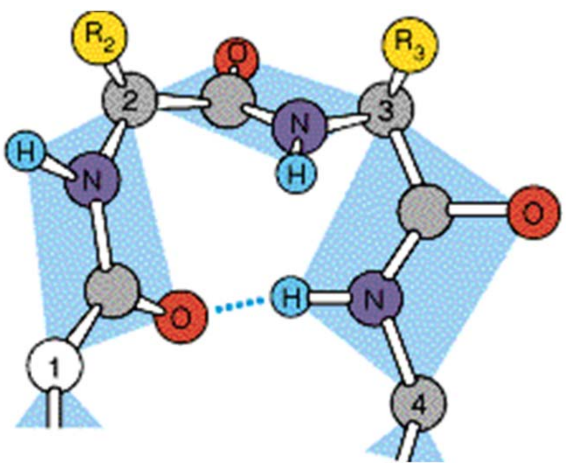
Parallel β-sheet



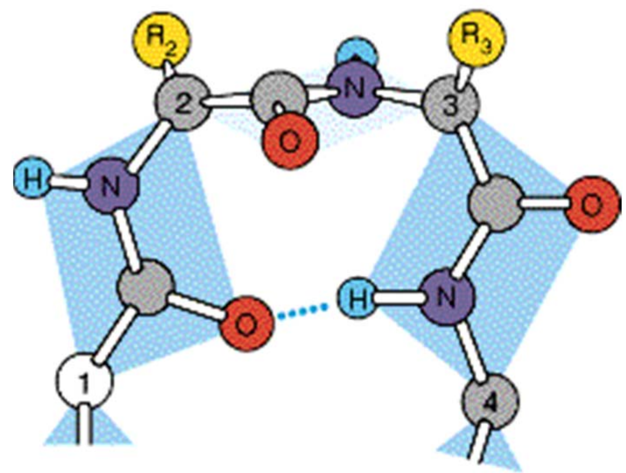
- β-Strands interacts with other β-strands running in same direction or opposite direction to form parallel or anti- parallel β-sheet respectively.
- The anti-parallel β-sheet has a narrowly spaced hydrogen bonding alternating with widely spaced hydrogen bonding.
- Parallel β-sheet have evenly spaced hydrogen bonding but at an angle.

Loop regions connectors of secondary structures

Hairpin loop Type I



Hairpin loop Type 2



- Loop regions are often seen as connectors of secondary structures
- The loop regions that connect the two anti parallel adjacent

  β-strands are known as hairpin loops or reverse turns.
- Loop regions are seen at the surface of the molecule therefore are rich in

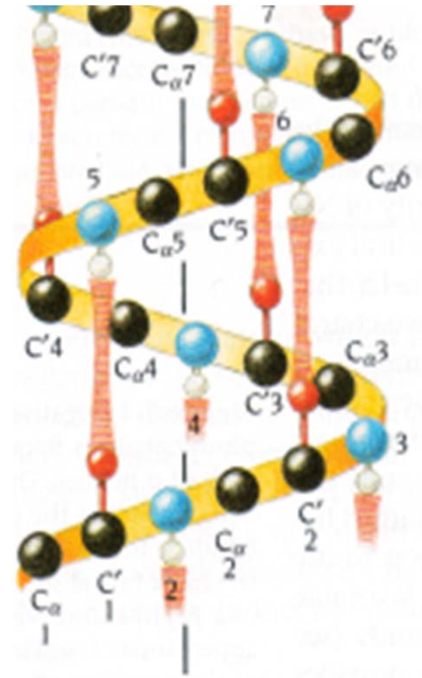
  hydrophilic amino acids.

Introduction to Proteins as Biotech Products

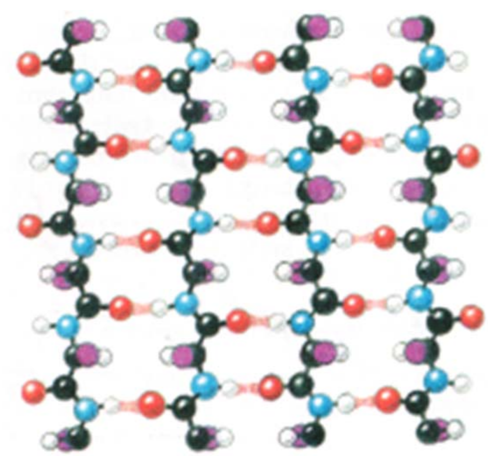
- **Proteins** – large molecules that are required for the structure, function, and regulation of living cells

Protein folding

- Folding of the peptide chain(s) confers structure to the protein
- Certain amino acids (aa) placed at specific intervals along the chain established how chain folds
  - alpha helix
  - beta sheet
    - proline aa inserts a ”kink” in the chain



Alpha helix
Formed by H-bonding between **CO** of residue **n** & **NH** of residue **n+4**



Beta sheets have H-bonding between CO group of one β-strand and NH group on an adjacent β-strand and vice versa

Protein Structures

- **Proteins**
  - Are complex molecules built of chains of amino acids
  - Have electrical charge that causes them to interact with other atoms and molecules
    - *Hydrophilic* – water loving
    - *Hydrophobic* – water hating

Protein Structures

- Protein Folding
  - The structure and function of a protein depends on protein folding
  - If protein is folded incorrectly, desired function of a protein is lost and a misfolded protein can be detrimental
  - Structures are fragile; hydrogen bonds are easily broken
  - tangled "plaques" of Alzheimer's disease may be the result in errors in protein folding

Diseases related to incorrectly folded proteins

- Cystic fibrosis
- mad cow disease (bovine spongiform encephalitis)

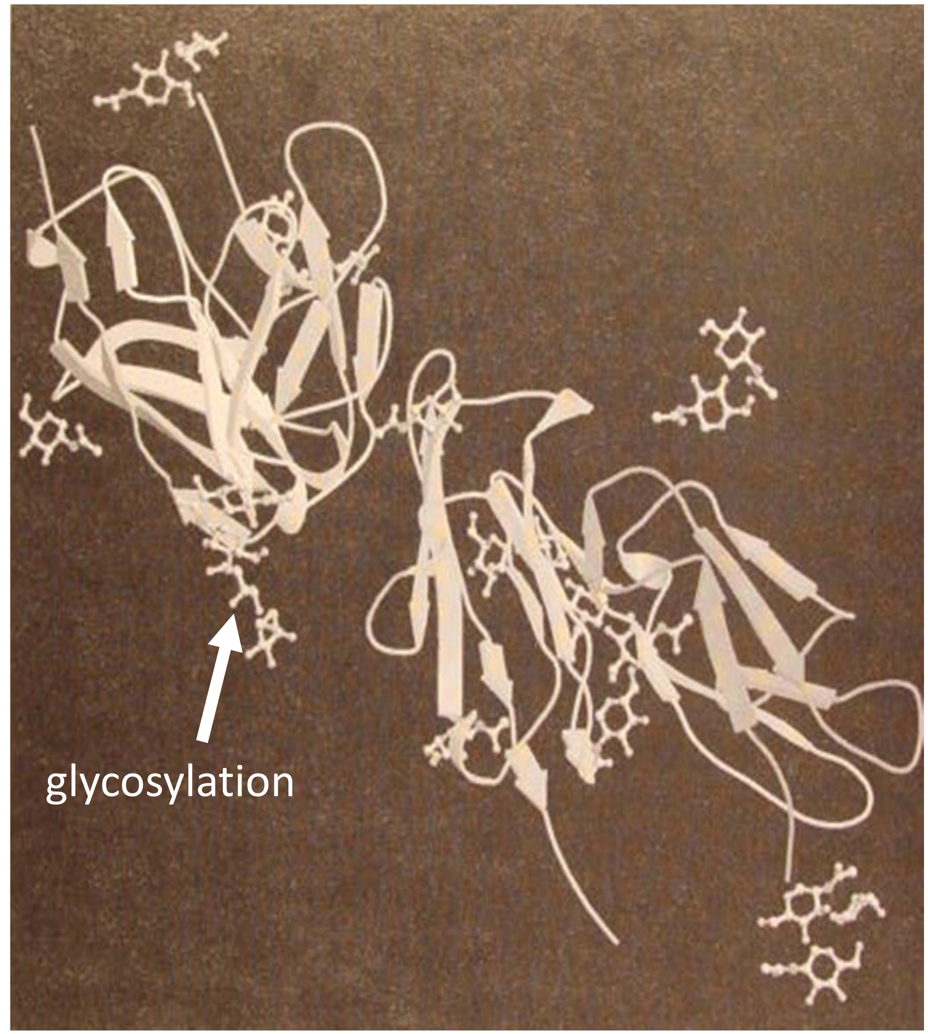
- A substitution of one amino acid for another along the peptide chain can affect how the protein folds.

Protein folding

Post translational
modifications

- Some proteins have sugar residues attached to the peptide chain
  - glycosylation
  - affects activity of protein
  - found in eucaryotic cells

More than 100 post-translational modifications occur



Proteins as Biotechnology Products

- Use of proteins in manufacturing is a time-tested technology
  - Beer brewing and winemaking
  - Cheese making
- Recombinant DNA technology made it possible to produce specific proteins on demand
  - Enzymes – proteins that speed up chemical reactions
  - Hormones
  - Antibodies

Classes of Proteins

- Enzymes
  - Depolymerization
    - Enzymes that degrade large complex molecules such as lignin, starch and cellulose to smaller subunit compounds
      - Amylase acts on starch to produce glucose
      - What is the degradation product of cellulose?
      - Lipases
      - Proteases
- Hormones (chemical messengers)
  - insulin
- Antibodies

Industrial Application of Some Enzymes

Enzyme	Application
Amylases	Digest starch in fermentation and processing
Proteases	Digest proteins for detergents, meats/leather, cheese, brewing/baking, animal/human digestive aids
Lipases	Digest lipids (fats) in dairy and vegetable oil products
Pectinases	Digest enzymes in fruit juice/pulp
Lactases	Digest milk sugar
Glucose isomerase	Produce high-fructose syrups
Cellulases/hemicellulases	Produce animal feeds, fruit juices, brewing converters
Penicillin acylase	Produces penicillin

Proteins as Biotechnology Products

- Applications of Proteins in Industry
  - Medical applications
  - Food processing
  - Textiles and leather goods
  - Detergents
  - Paper manufacturing and recycling
  - Adhesives: natural glues
  - Bioremediation: treating pollution with proteins

Proteins as Biotechnology Products

- Making a Biotech Drug
  - Produced through microbial fermentation or mammalian cell culture
  - Complicated and time-consuming process
  - Must strictly comply with FDA regulations at all stages of the procedure

Proteins in medical applications

- Insulin
  - Recombinant insulin produced in E. coli avoids adverse reactions by our immune system in response to foreign substances (insulin recovered from pigs and cattle).
- Pharmaceutical industry has invested in this area because they can charge so much $$ for use in treatments

Biosynthetic Corneas

- Can help regenerate and repair damaged tissue and improve vision.
- Made from synthetically cross-linked recombinant human collagen.
- Collagen was produced in yeast cells and chemically cross-linked.

Biomarker Proteins

- For rapidly screening molecules associated with disease.
- Eg 28 proteins are predicted which may help in early diagnosis of breast cancer.
- Proteins produced by cancers can trigger the body to produce antibodies not produced by healthy individuals – “autoantibodies”.
- Autoantibodies can be measured in blood to detect cancer.
- Autoantibody biomarkers can also detect other forms of cancer – ovarian, prostate, and lung.

Recombinant proteins produced by the pharmaceutical industry

Protein	Application
Erythropoietins	Treatment of anemia
Interleukin 1,2,3,4	Cancer, AIDS
Human growth factor	Treatment of growth deficiency in children
Tissue plasminogen factor	Heart attack and stroke treatment
Vaccines	Hepatitis B, herpes, malaria

Proteins in Food Processing Industry

Cheese Production

Isolation of prochymosin mRNA from calf cells

↓

Transcription of mRNA to cDNA (using reverse transcriptase)

Determination of the amino acid sequence of prochymosin

↓

Biochemical synthesis of prochymosin DNA

↓

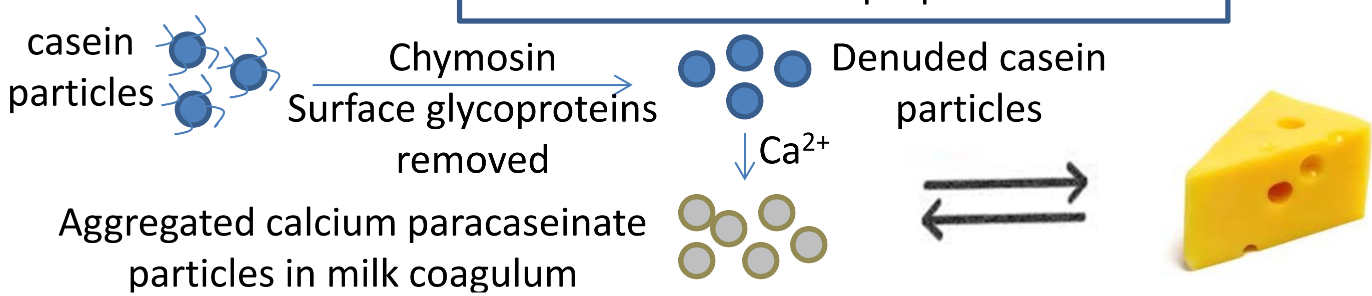
Cloning of cDNA into an appropriate vector

Introduction of vector into a suitable microorganism

Expression and excretion of prochymosin into culture

Conversion of prochymosin into active chymosin at low pH

Purification of chymosin and formulation of commercial preparations



Examples of proteins used in other commercial applications

- Textile industry
  - Enzymes have replaced toxic chemical
    - “stone-washed denim”
    - Spider silk : proteins are responsible for the amazing properties.
- Home products industry
  - Enzymes added to “Drano” and “Liquid Plumber” replaces caustic chemicals to dissolve hair fibers, grease, etc. that plug drains

Adhesives

- Mussel adhesive protein
  - Underwater glue
  - Industries like 3M are trying to design and synthesize proteins in *E. coli* bacterium that mimic the natural proteins excreted by mussels

- Detergents
  - Proteases, lipases, and amylases to dissolve stains in cooler water.
  - Better cleaning and more biodegradable

Bioremediation

- Treating pollution with proteins
- Enzymes can be used to digest organic wastes produced from feedlots, homes, and industries.
- Neutralizing heavy metal pollutants such as mercury and cadmium. Some microbes have a sticky coat of metallothioneins, proteins that capture heavy metals.

- Protein Engineering
  - Introducing specific, predefined alterations in the amino acid sequence through a process known as **directed molecular evolution technology**
  - Creating entirely new protein molecules
  - induce mutations randomly into genes and then select organisms (bacteria) with protein products (enzyme) that, when assayed using an enzyme assay, gives the highest activity.
    - Led to recovery of an enzyme that functioned in presence of high cyanide concentration.

Protein engineering

- Directed molecular evolution
  - Induce mutation in a specific gene and select best protein from mutant gene for our application irrespective of its value to organism. (example: insulin production in *E. coli*)

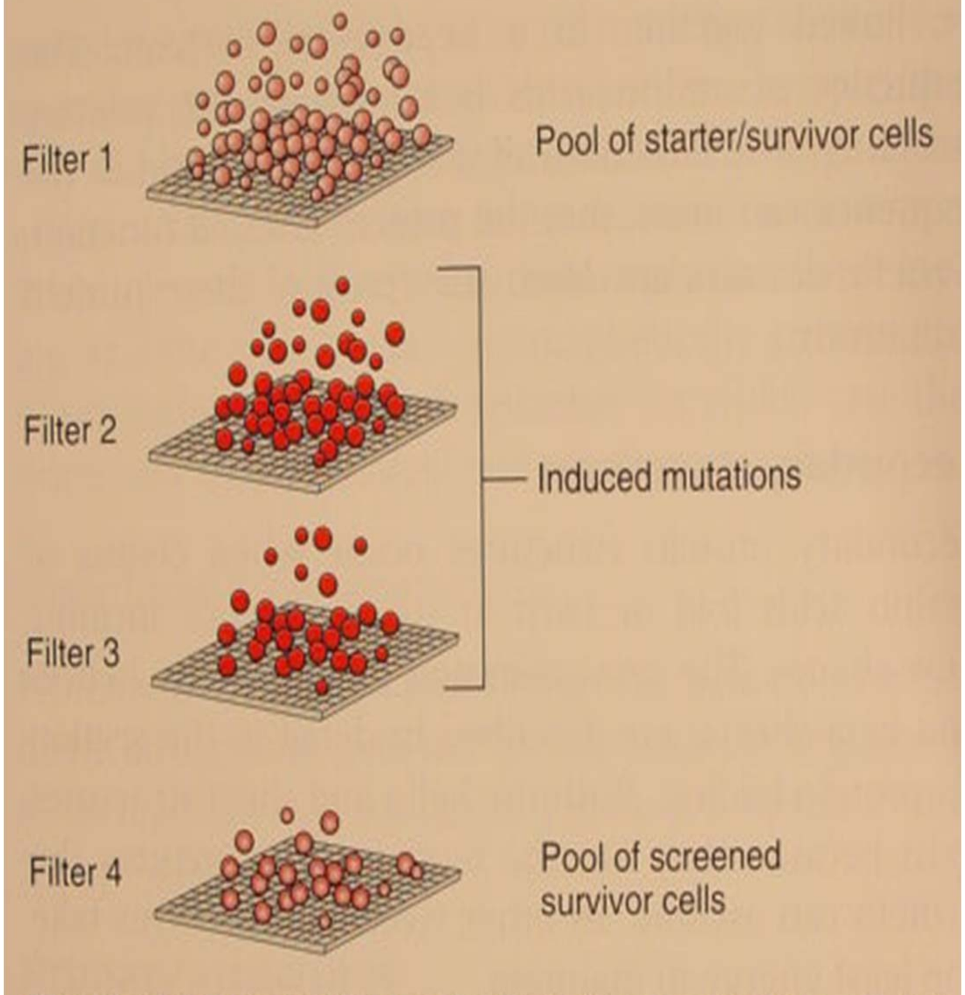


Figure 4.5 Directed Molecular Evolution Technology Genes