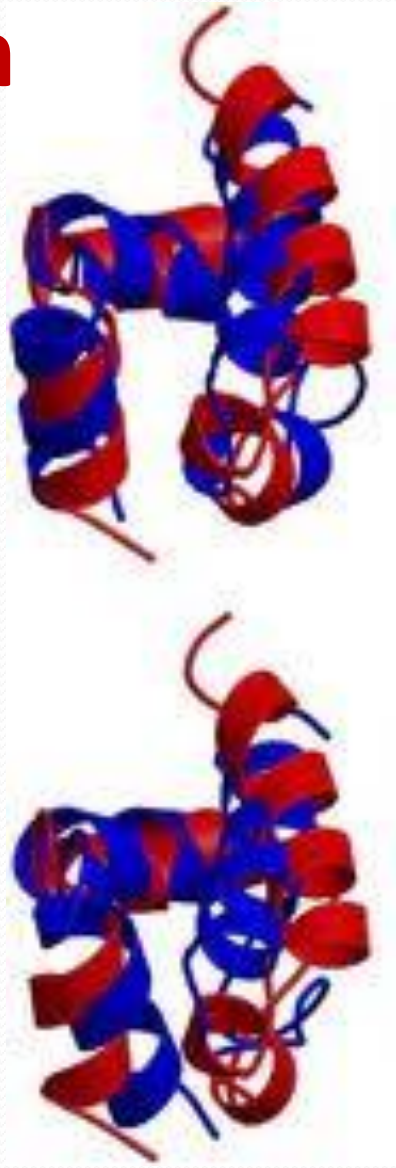
Polypeptide & protein structure

Nafith Abu Tarboush
DDS, MSc, PhD
natarboush@ju.edu.jo
www.facebook.com/natarboush

Protein conformation

- Many conformations are possible for proteins due to flexibility of amino acids linked by peptide bonds

- At least one major conformations has biological activity, and hence is considered the protein's **native conformation or the native protein**



Levels of Protein Structure

- **1°structure**: sequence and number, from N to C
- **2°structure:** the ordered 3-dimensional arrangements (conformations) in localized regions of a polypeptide chain, backbone interactions through hydrogen bonding;
  - e. g., α-helix and β-pleated sheet
- **3° structure**: 3-D arrangement of all atoms
- **4° structure**: multimeric proteins, arrangement of monomer subunits with respect to each other

Primary Structure of Proteins

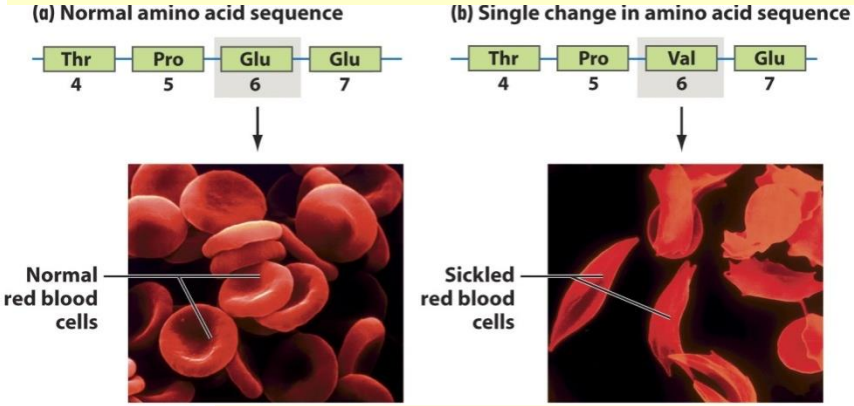
- Zigzag arrangement
- R-groups
- Determination?
- 1° Sequence & 3-D conformation; relation to functional properties (MW& genetic mutations)
- Site-directed mutagenesis and structure function relationship

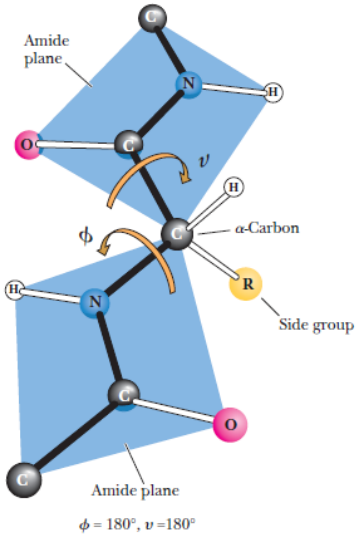
HBB Sequence in Normal Adult Hemoglobin (Hb A):

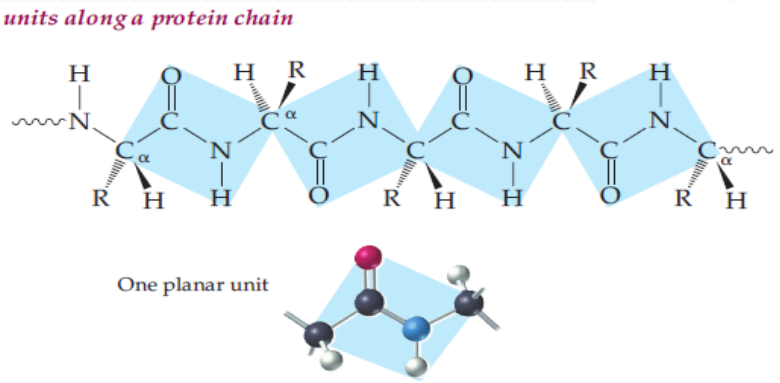
Nucleotide	CTG	ACT	CCT	GAG	GAG	AAG	TCT
Amino Acid	Leu	Thr	Pro	Glu	Glu	Lys	Ser
	3			6			9

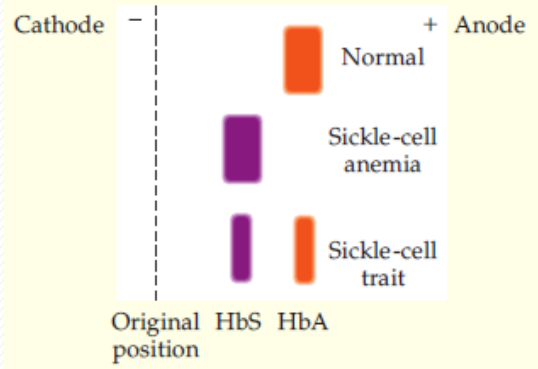
HBB Sequence in Mutant Adult Hemoglobin (Hb S):

Nucleotide	CTG	ACT	CCT	GTG	GAG	AAG	TCT
Amino Acid	Leu	Thr	Pro	Val	Glu	Lys	Ser
	3			6			9



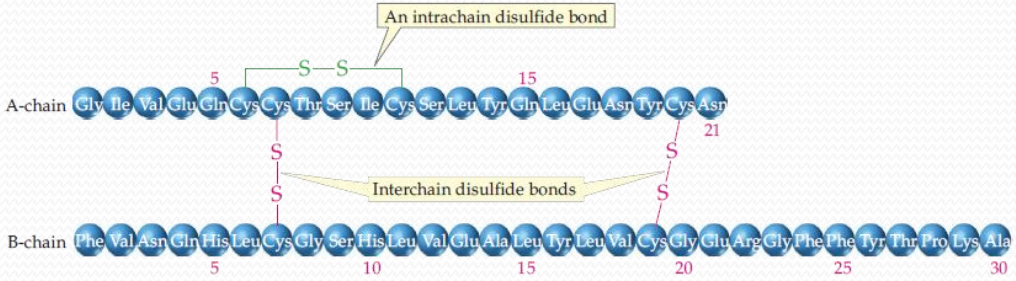


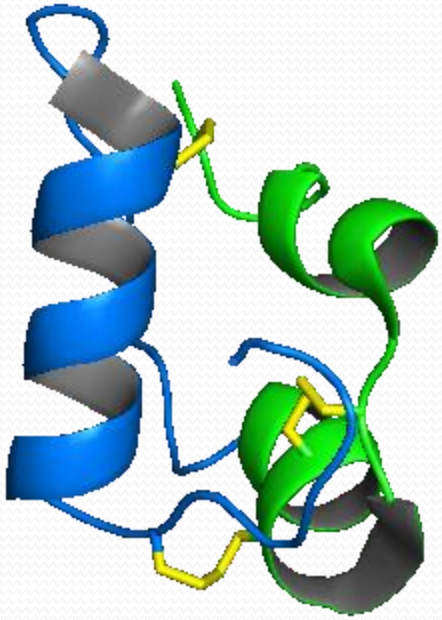


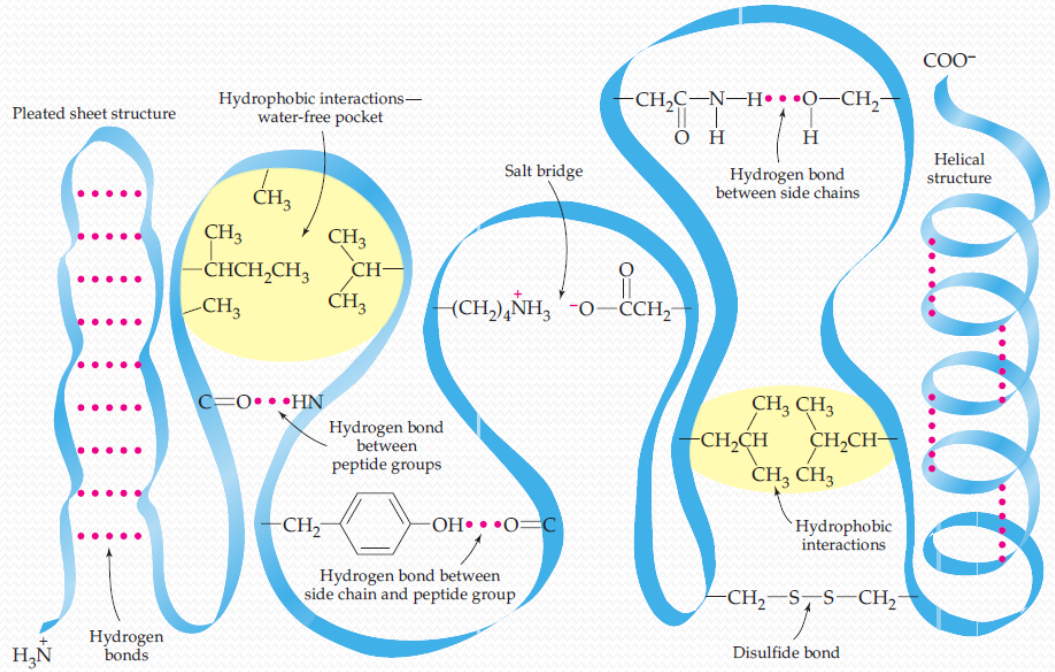


Shape-Determining & stabilizing Interactions in Proteins

- Is it ordered or spaghetti?
- Hydrogen, ionic, covalent, & hydrophobic

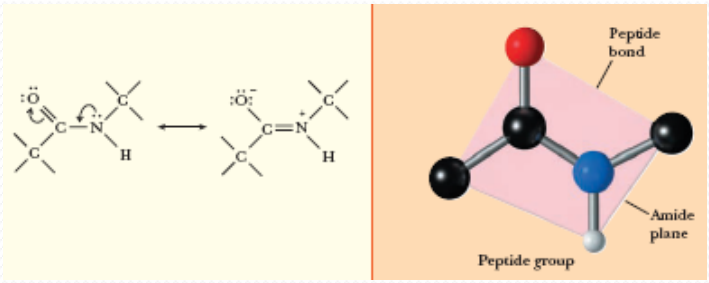


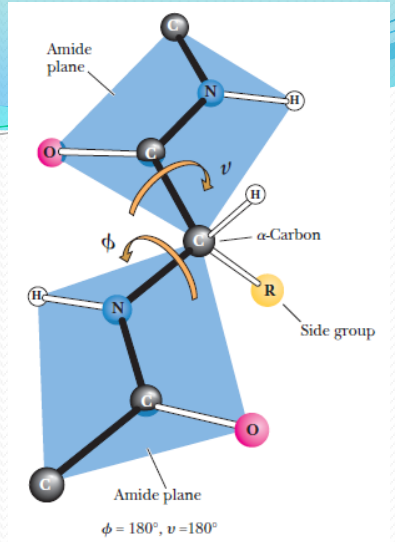




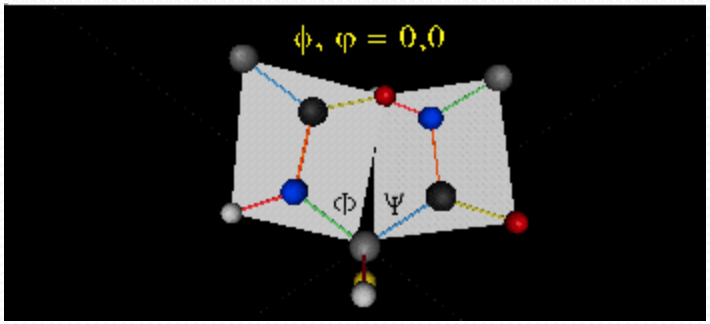
Secondary Structure of Proteins

- What is the secondary structure of proteins? “folding of the backbone”
- What are the bonds that have a free rotation? What is the implication? These angles repeat themselves in regular secondary structures
- Two main kinds: **α-helix** and **β-pleated sheet**
- They are periodic; their features repeat at regular intervals
- Stability of secondary structure

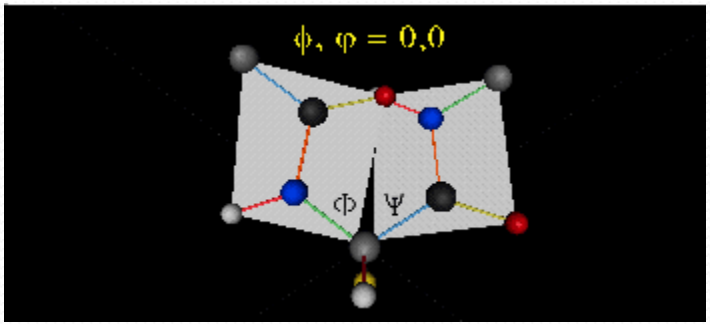




Rotation around phi (φ)

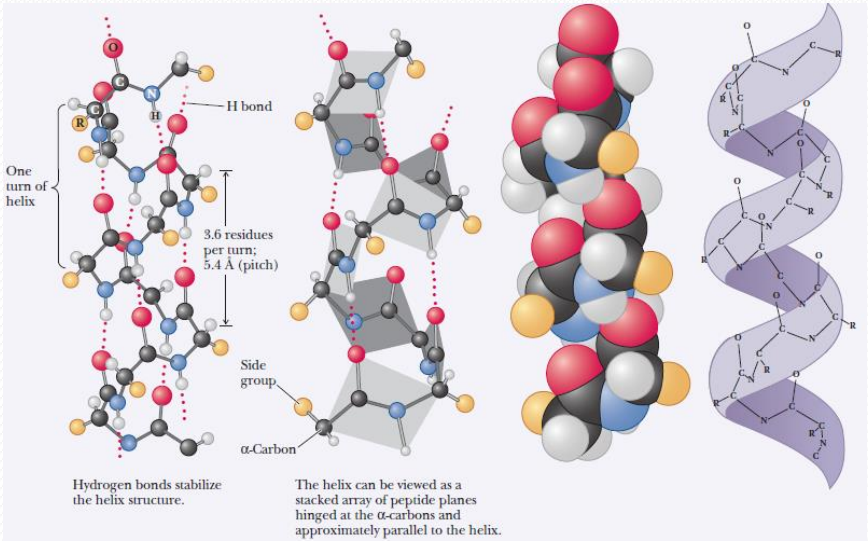


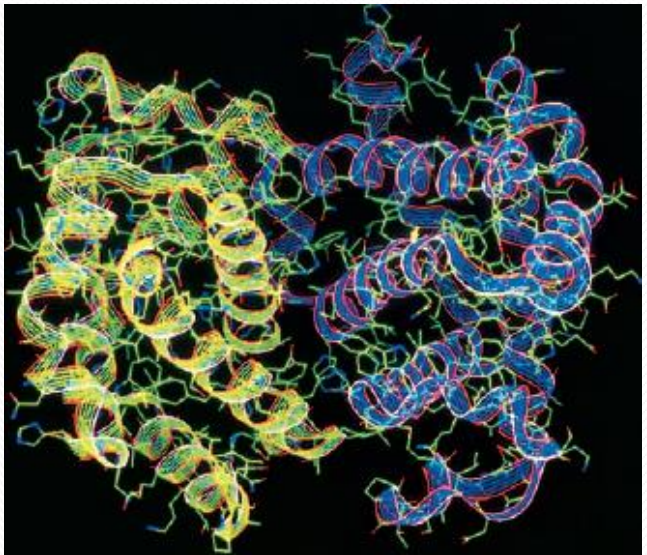
Rotation around psi (ψ)



The α-helix

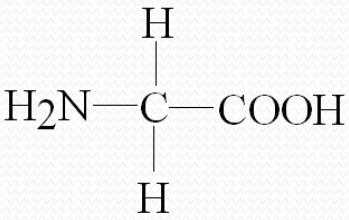
- H-bonds are parallel to the helix axis, same segment
- C=O binds N—H four residues away
- linear arrangement of H-bonds (maximum strength and stability)
- Turns occur every 3.6 residues, right handed, clockwise
- The pitch (linear distance between corresponding points on successive turns) is 5.4 Å
- Proteins have varying amounts of α-helical structures
- What factors affect the helix (specific amino acids, electrostatic repulsion, steric repulsion)





Amino acids NOT found in α-helix

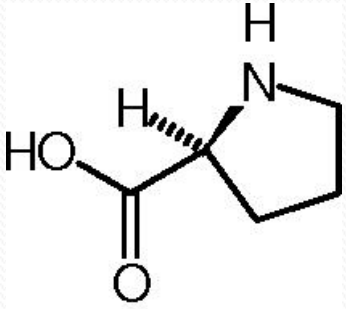
- Glycine: too small & entropically expensive (high flexibility)



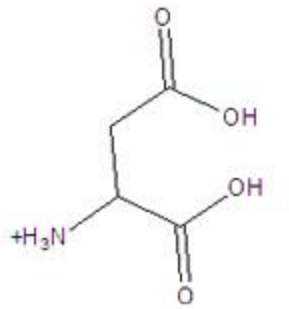
glycine

- Proline
  - No rotation around psi bond
  - No hydrogen bonding of α-amino group

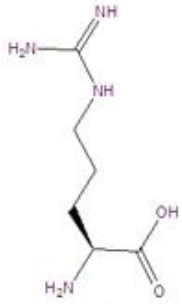
- Close proximity of a pair of charged amino acids with similar charges



L-proline

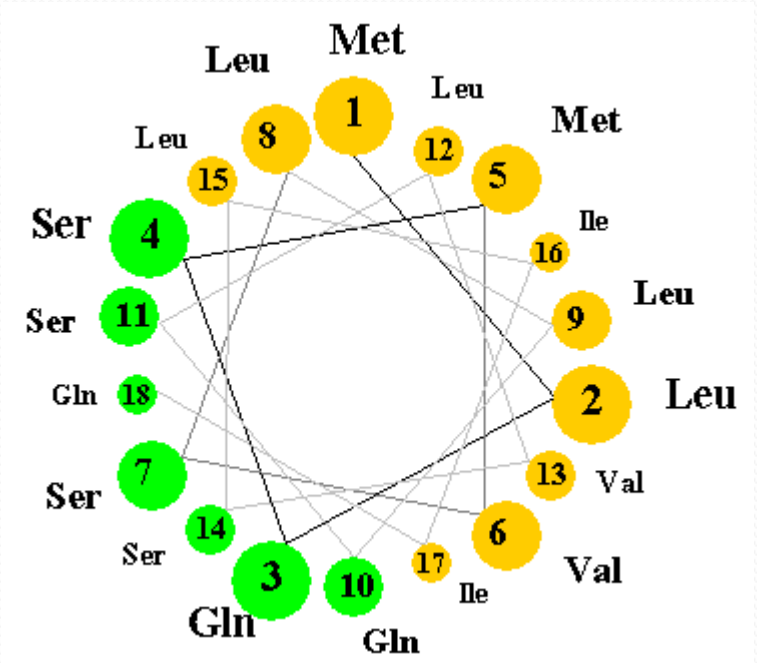


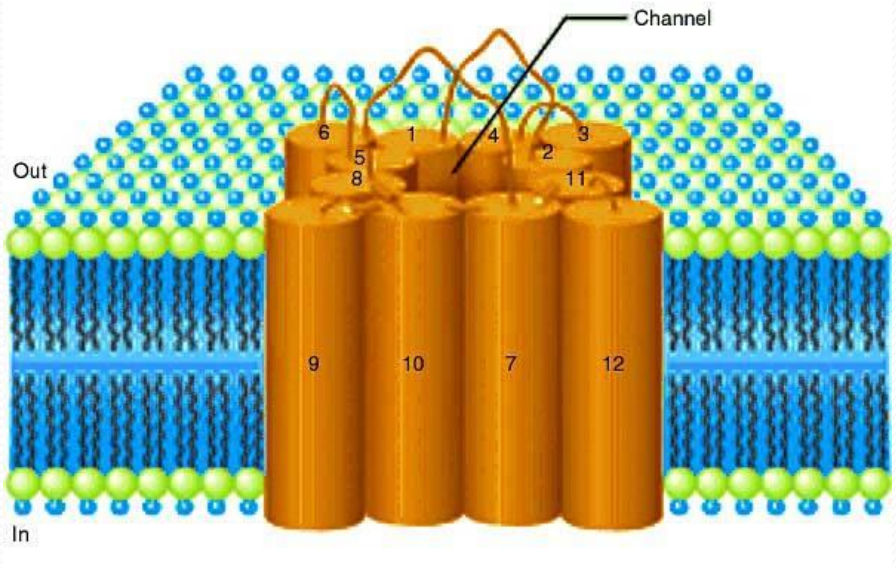
aspartic acid



arginine

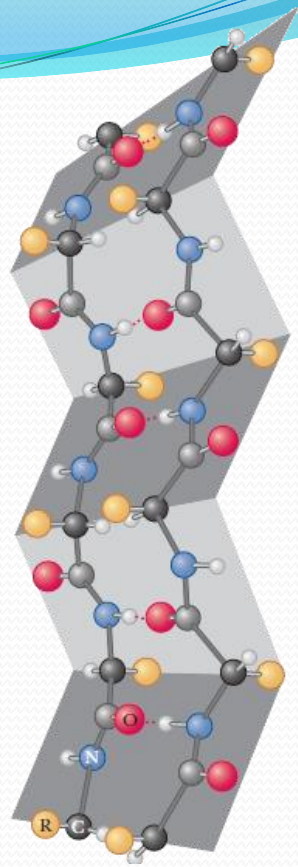
Amphipathic α helices

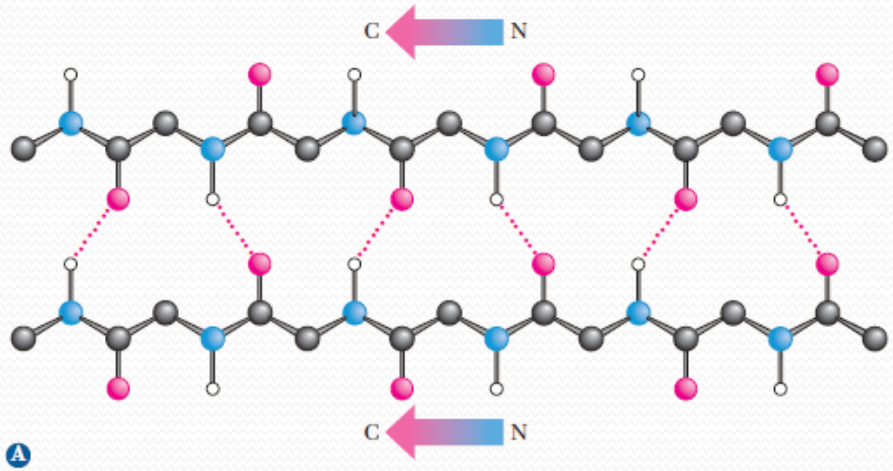


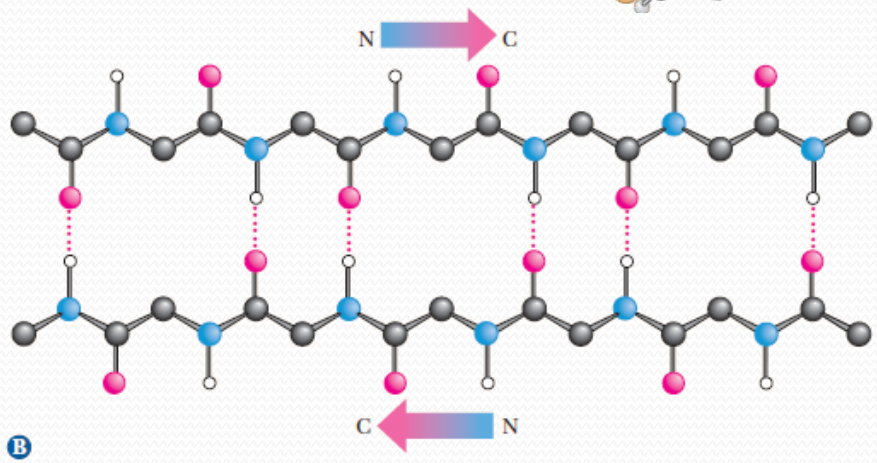


The β-sheets

- Backbone is almost completely extended
- R groups extending above and below the sheet
- H-bonds are *intra-chain or inter-chain bonds*
- Perpendicular to the direction of the protein chain
- Parallel vs. anti-parallel
- Zigzag structure

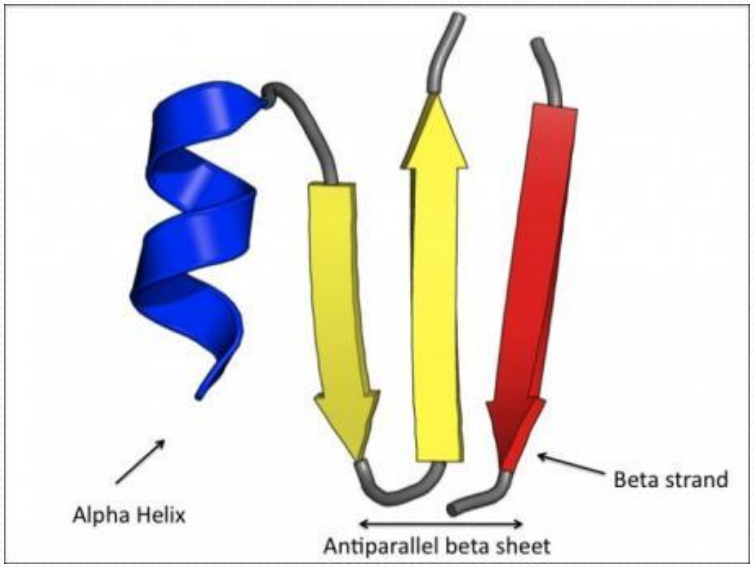






How many β strands can a β sheet have?

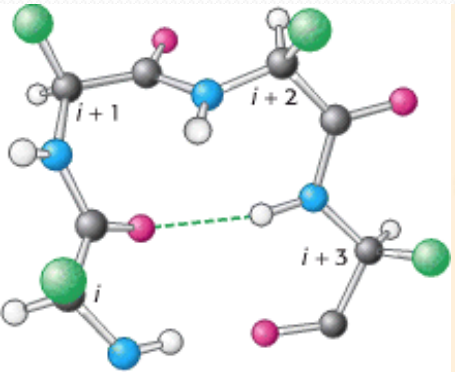
- β sheets can form between many strands, typically 4 or 5 but as many as 10 or more
- Such β sheets can be purely antiparallel, purely parallel, or mixed
- Proline tends to disrupt β strands

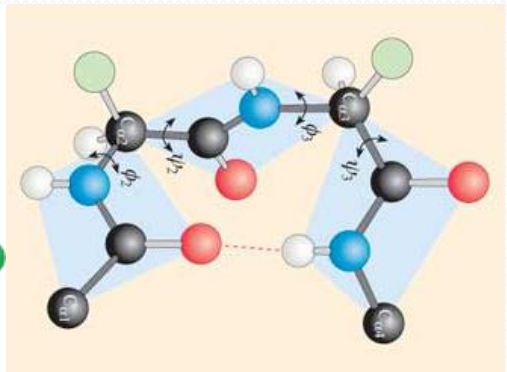


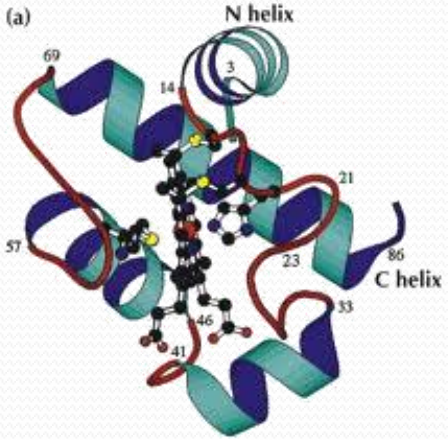
Others regular ones: Turns & loops

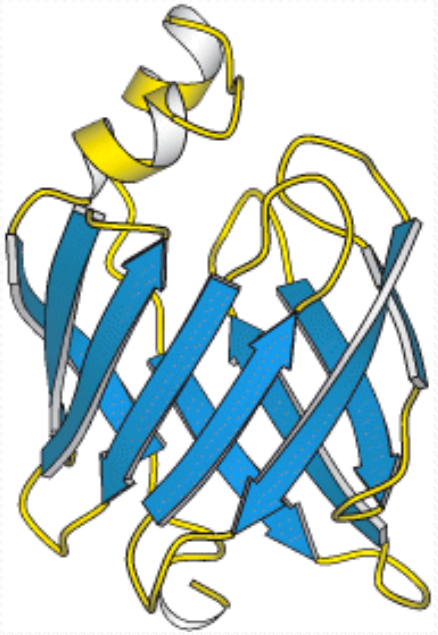
- Turns
  - Compact, U-shaped secondary structures
  - Also known as β turn or hairpin bend
  - What are they used for? How are they stabilized?
  - Involve 4 amino acids (H-bond: C=O of 1 & N–H of 4)
  - Glycine and proline are commonly present in turns

- Loops do not have regular structures



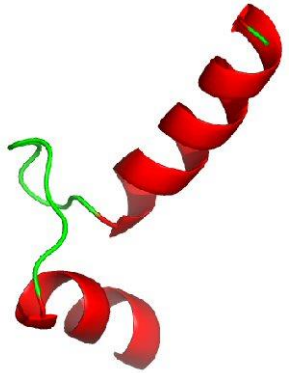


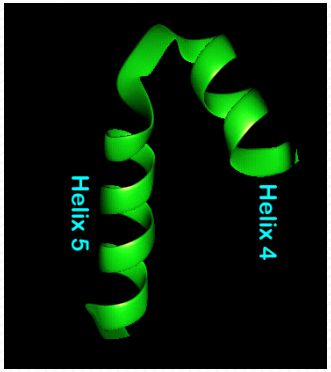


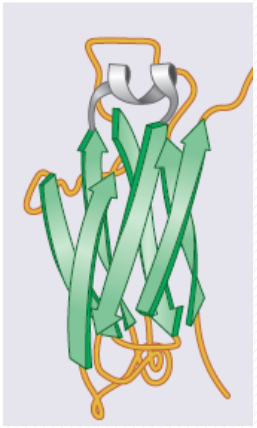


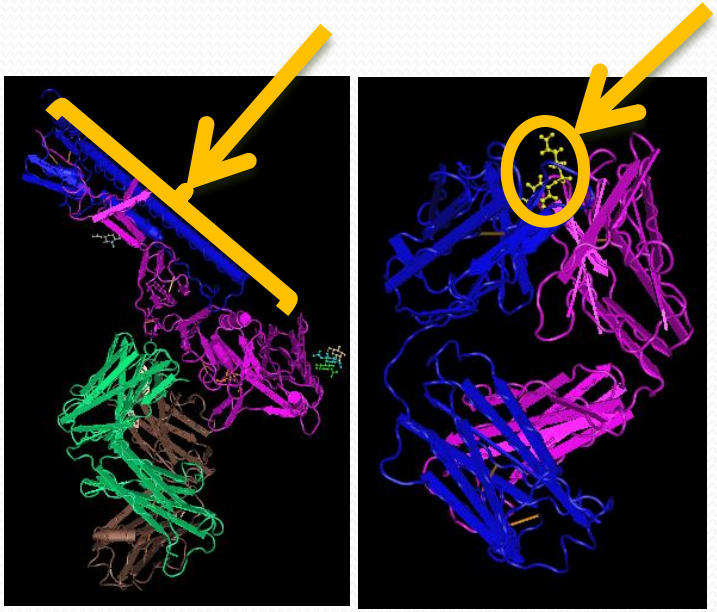
Super-secondary structures: Motifs & Domains

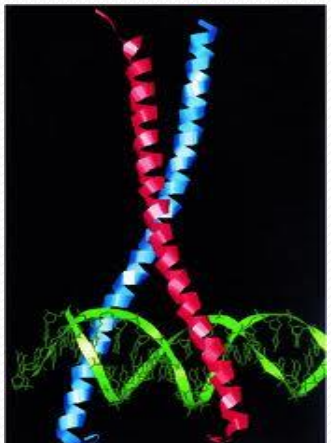
- A **motif:** a small portion of a protein (typically less than 20 amino acids)
  - In general, motifs may provide us with information about the folding of proteins, but no biological function
- **Domains**; protein conformations with similar functions, 100–200 residues, fold independently of the rest of the protein
  - leucine zipper
  - Immunoglobulin fold





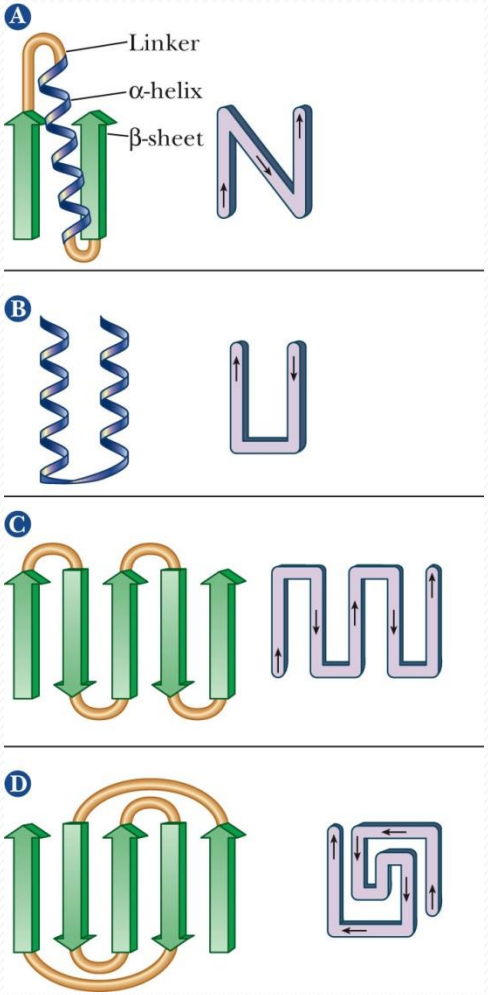


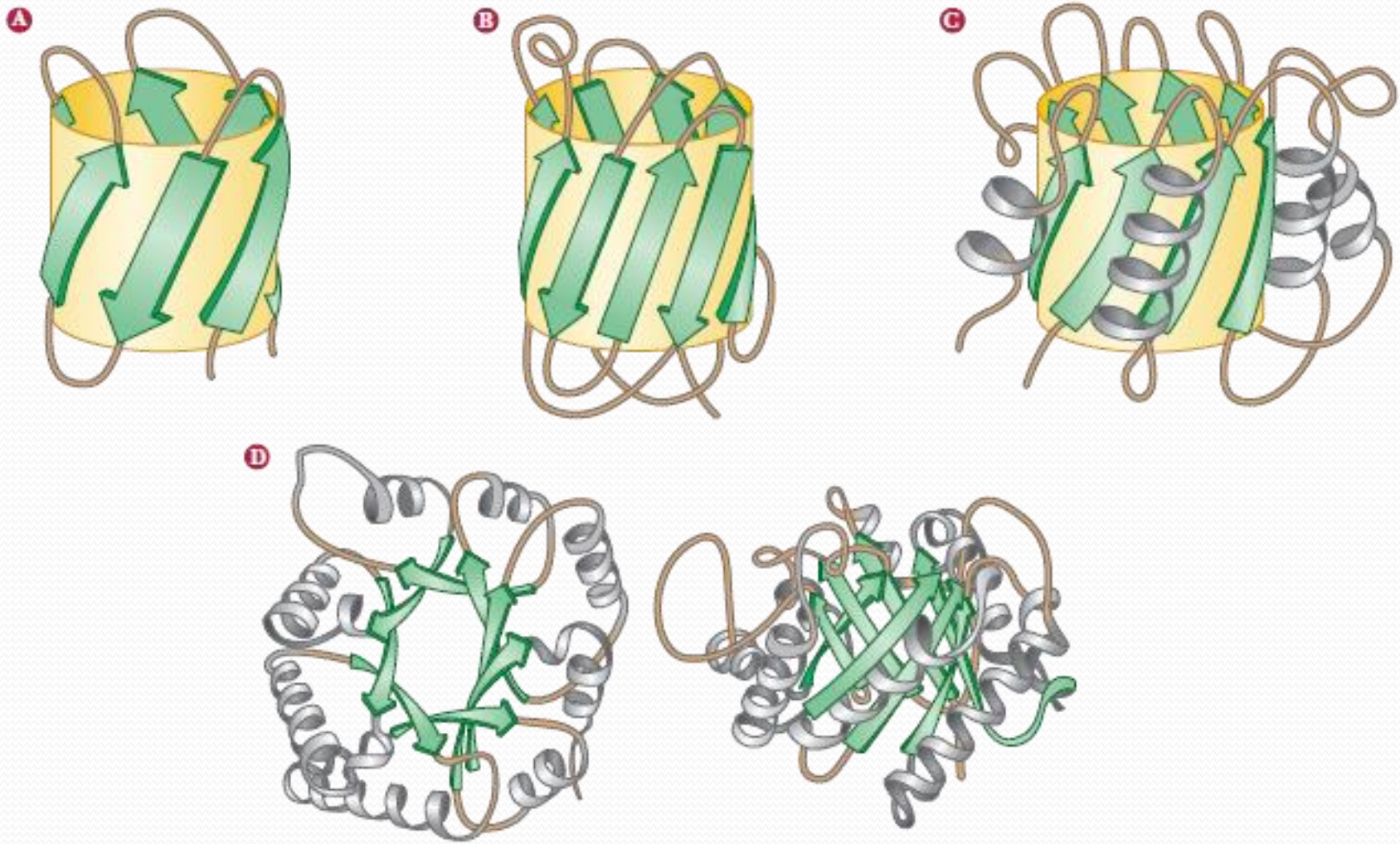




α-Helices and β-Sheets

- **Supersecondary structures**: a combination of α- and β-sections
  - **βαβ unit**: (parallel)
  - **αα unit**: (helix-turn-helix), anti-parallel
  - **β-meander**: an anti-parallel sheet formed by a series of tight reverse turns connecting stretches of a polypeptide chain
  - **Greek key**: a repetitive super-secondary structure formed when an anti-parallel sheet doubles back on itself
  - **β-barrel**: created when β-sheets are extensive enough to fold back on themselves




A
B
C
D

Fibrous Proteins

- Contain polypeptide chains organized approximately parallel along a single axis:
  - Consist of long fibers or large sheets
  - Mechanically strong
  - Insoluble
  - play an important structural role
- Examples are
  - Keratin
  - Collagen
  - fibroin

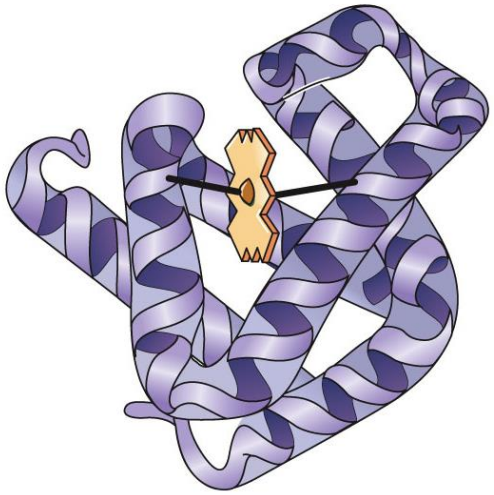
Fibroin, β-sheets, alternating glycine and alanine

α-keratins, bundles of α-helices

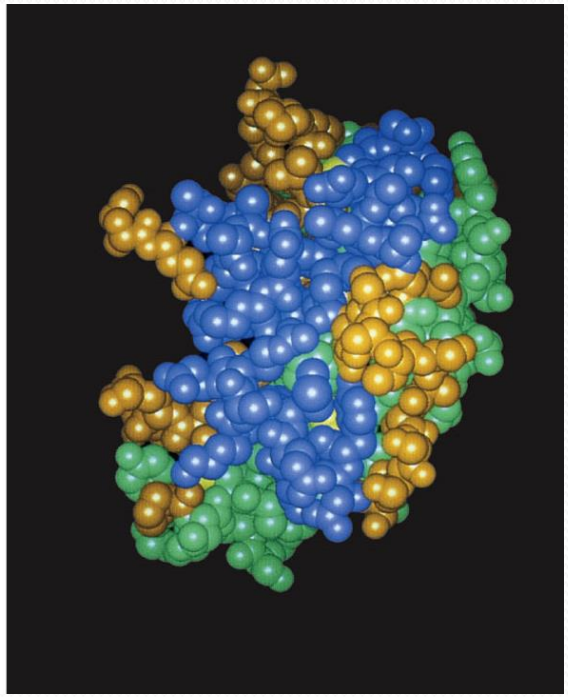
Globular Proteins

- Folded to, a more or less, spherical shape
  - Soluble
  - Polar vs. non-polar, exterior vs. interior
  - Most of them have substantial sections of α-helix and β-sheet

Filament
(four right-hand
twisted protofilaments)



Myoglobin, a globular protein

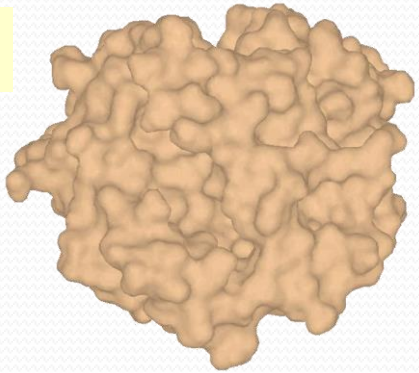


3° Structure

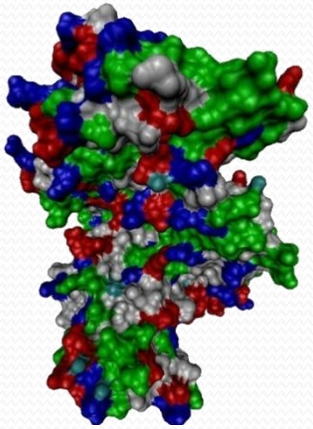
- The 3-dimensional arrangement of all amino acids in a protein
- The overall conformation of a polypeptide chain
- The spatial arrangement of amino acid residues that are far apart in the sequence
- Simple vs. conjugated

How to look at proteins…

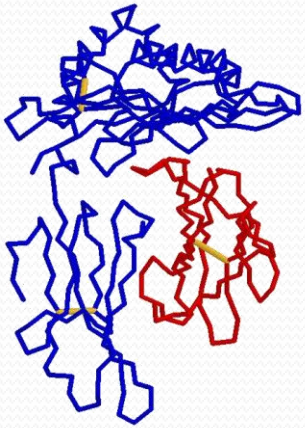
Protein surface map



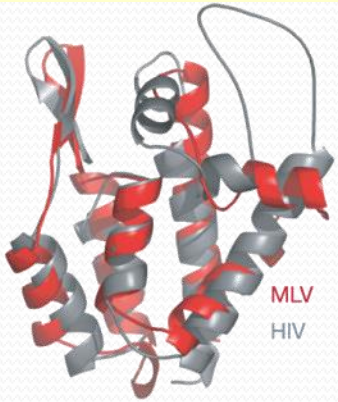
Space filling structure



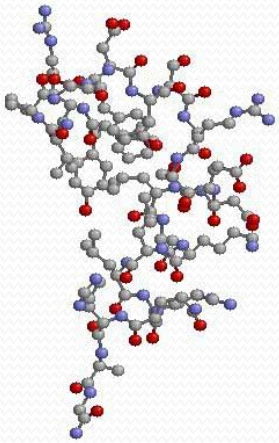
Trace structure



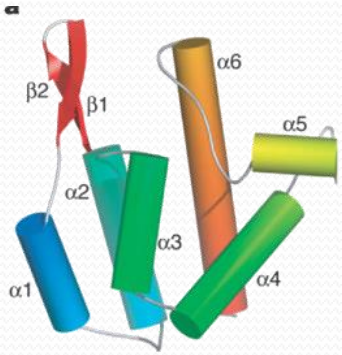
Ribbon structure



Ball & stick structure

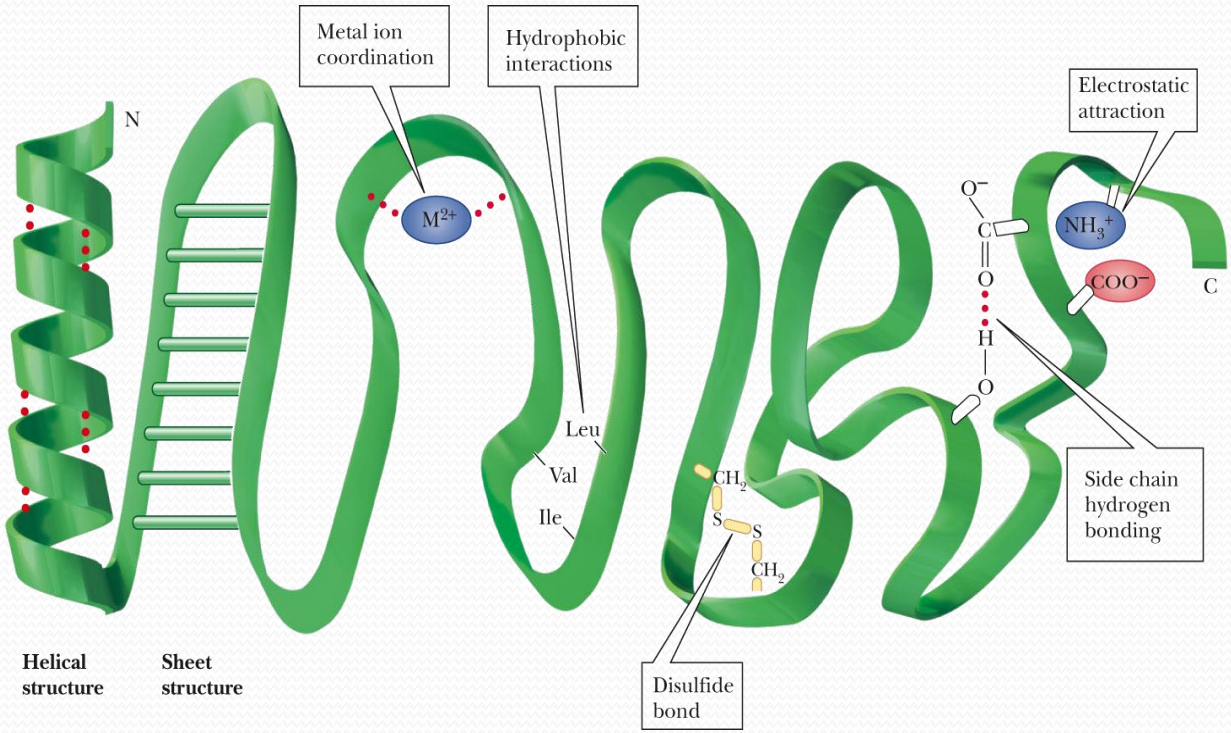


Cylinder structure



Forces That Stabilize Protein Structure

- Backbone H-bonding
- Side chain H-bonding
- Hydrophobic interactions
- Electrostatic attraction
- Electrostatic repulsion
- Metal coordination

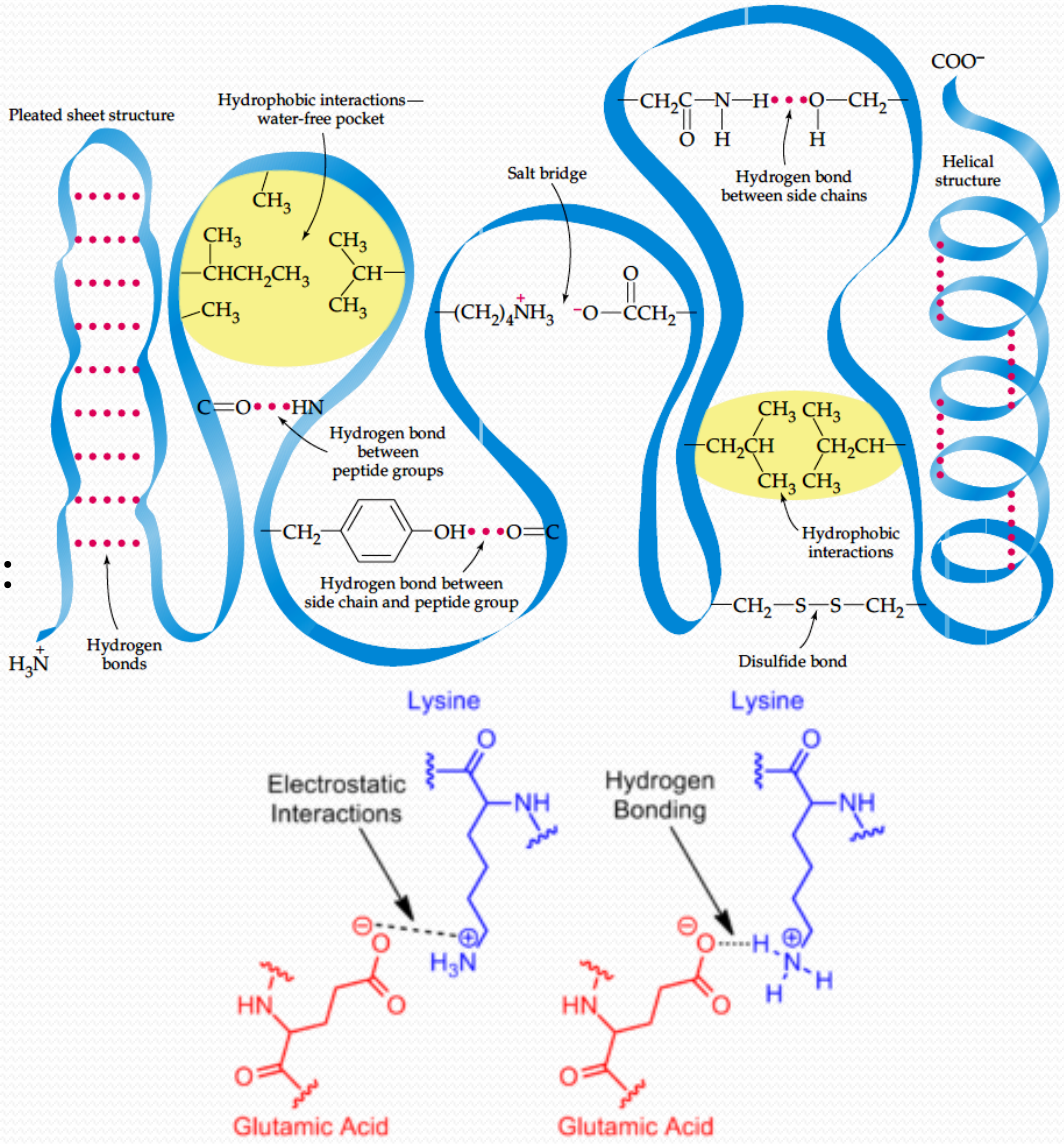


- Not every protein have all kinds of interactions (myoglobin & hemoglobin; no S-S) (trypsin & chymotrypsin, no metal complexes)

- Interactions between side chains also plays a role

Shape determining forces

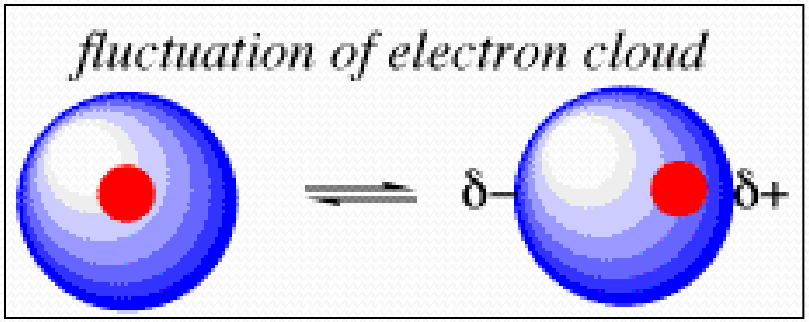
- Non-covalent interactions
  - Hydrogen bonds: amino acids, aqueous medium
  - Charge-charge interactions (salt bridges)
  - Charge-dipole interactions: charged R groups with partial charges of water

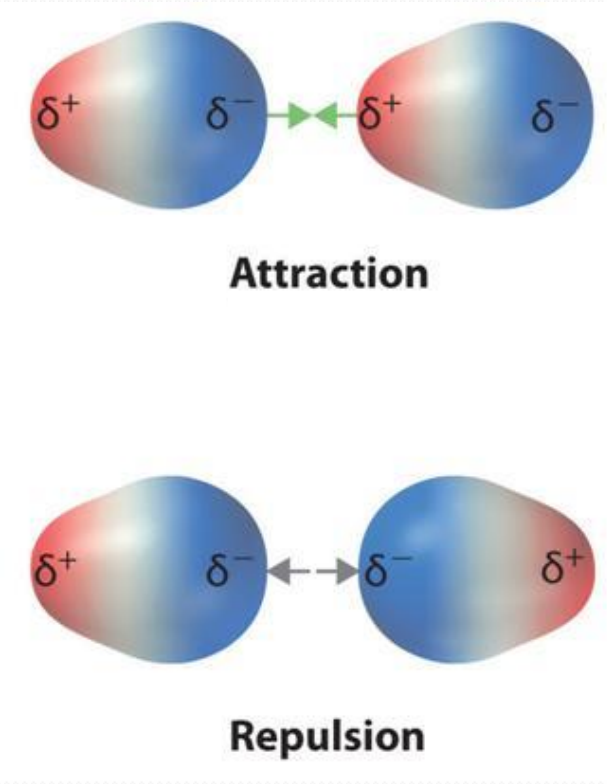


Van der Waals Forces

- Attractive & repulsive forces control protein folding

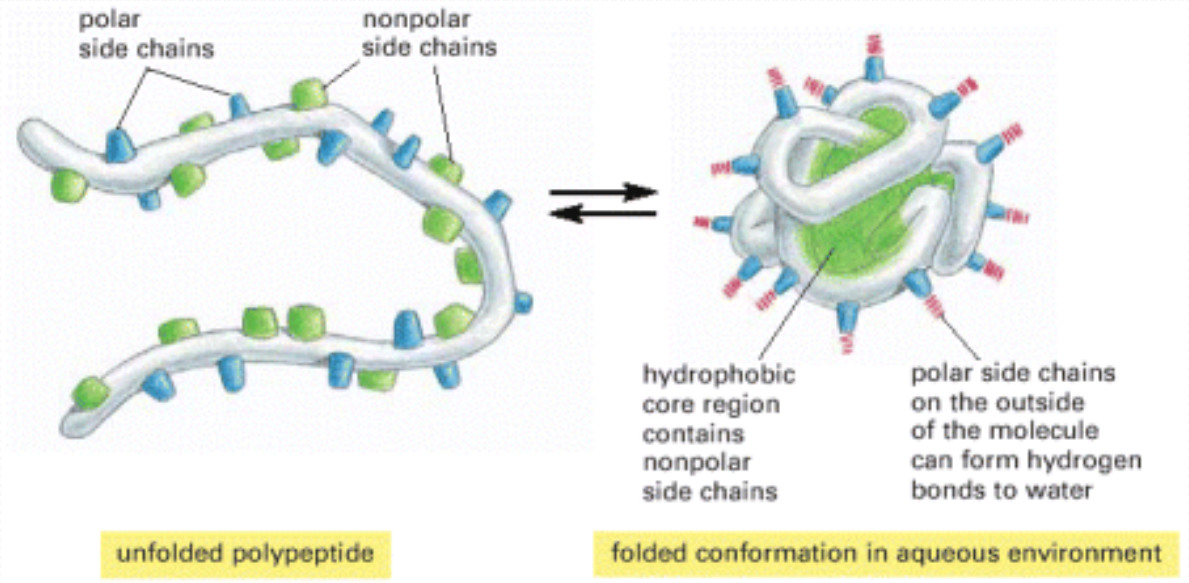
- Extremely weak (2-4 kJ/mol/atom pair), but significant!



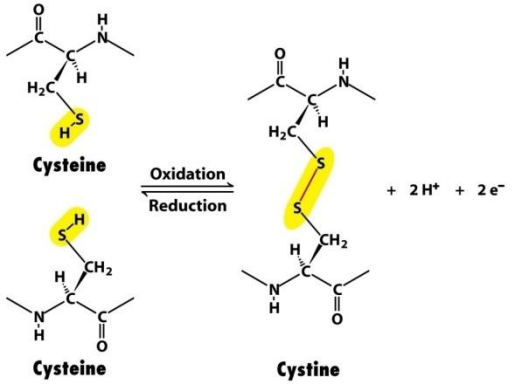


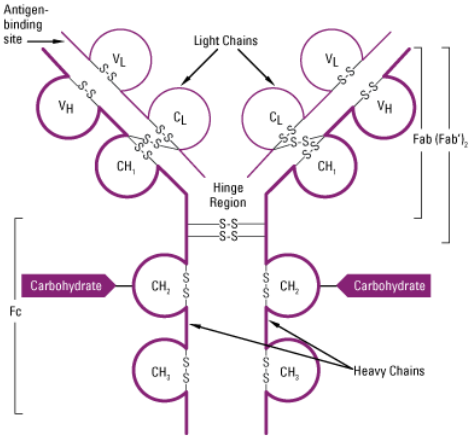
Hydrophobic interactions

- A system is more thermodynamically (energetically) stable when hydrophobic groups are clustered together rather than extended into the aqueous surroundings
- Can polar amino acids be found in the interior?
  - H-bonds to other amino acids (side chain or backbone)
  - Play important roles in the function of proteins

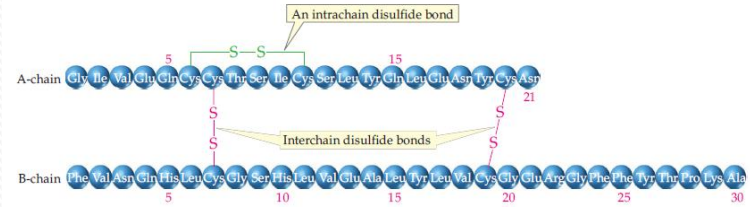


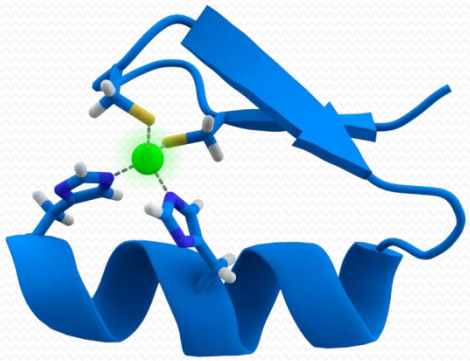
Factors that stabilize protein structures

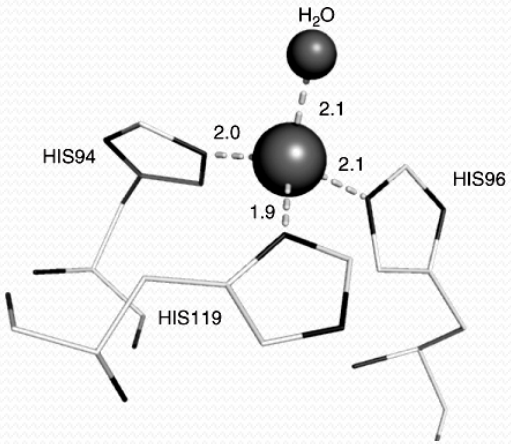


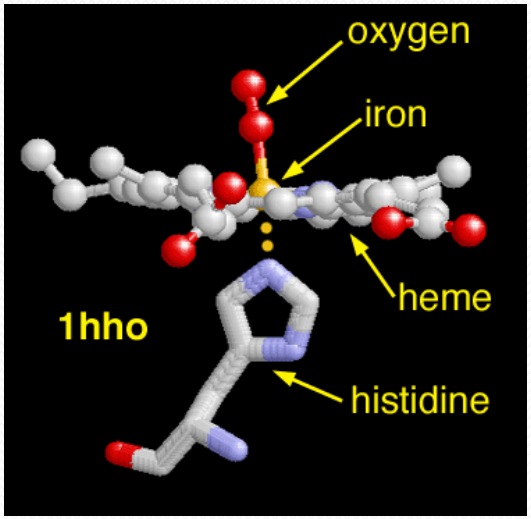


- Do not determine the three-dimensional structure of proteins, but stabilizes it:
  - Disulfide bonds (redox)
  - Metal ions
    - Covalent interaction (myoglobin)
    - Salt bridges (carbonic anhydrase)



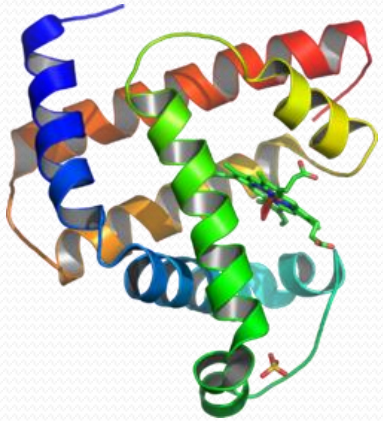


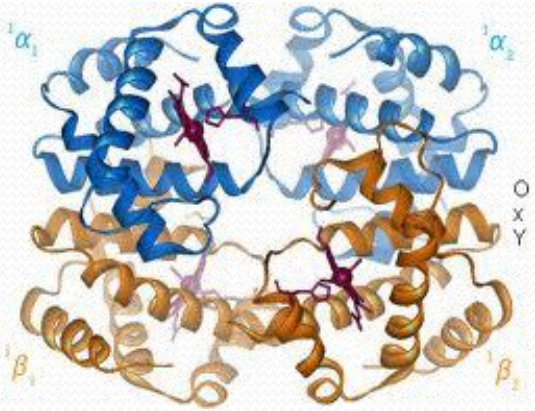




3°& 4°Structure

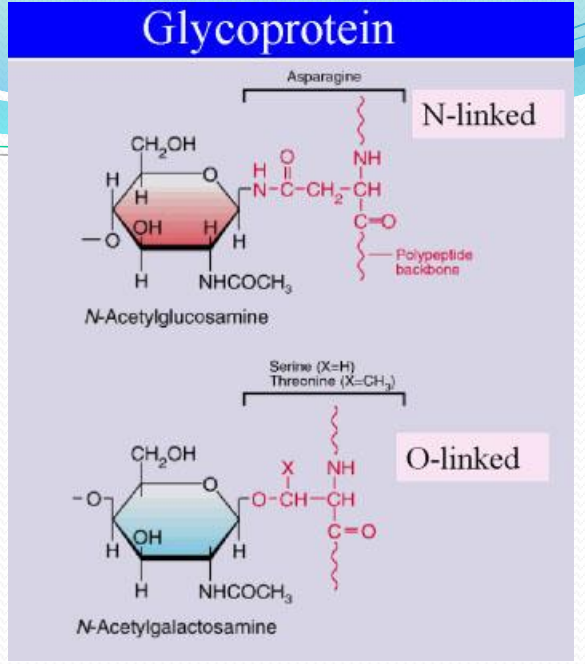
- **Tertiary (3°) structure**: the arrangement in space of all atoms in a polypeptide chain
  - **It is not always possible to draw a clear distinction between 2°and 3°structure**
- **Quaternary (4°) structure**: the association of polypeptide chains into aggregations called "subunits" (dimers, trimers, tetramers,…etc).
- Simple or conjugated (holo vs. apo)
- Homo vs. hetero
- Interactions:
  - Mainly: Non-covalent
  - Sometimes: covalent (S-S)



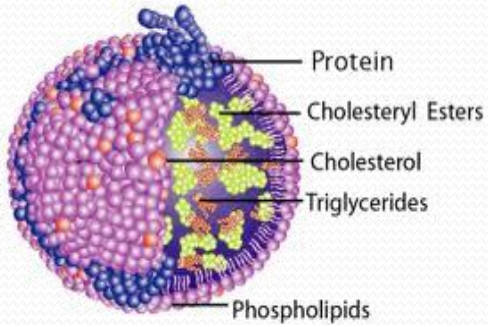


Complex Protein Structures

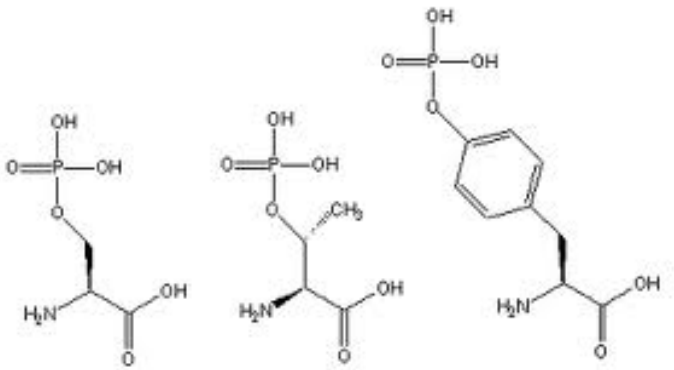
- Carbohydrates (glycoproteins):
  - Covalent conjugation
    - *N*-linked (-N of Asn)
    - *O*-linked (-OH of Ser or Thr) & occasionally to –OH of hydroxy-lysine



- Lipids (lipoproteins):
  - Non-covalent
  - Store & transport lipids & cholesterol



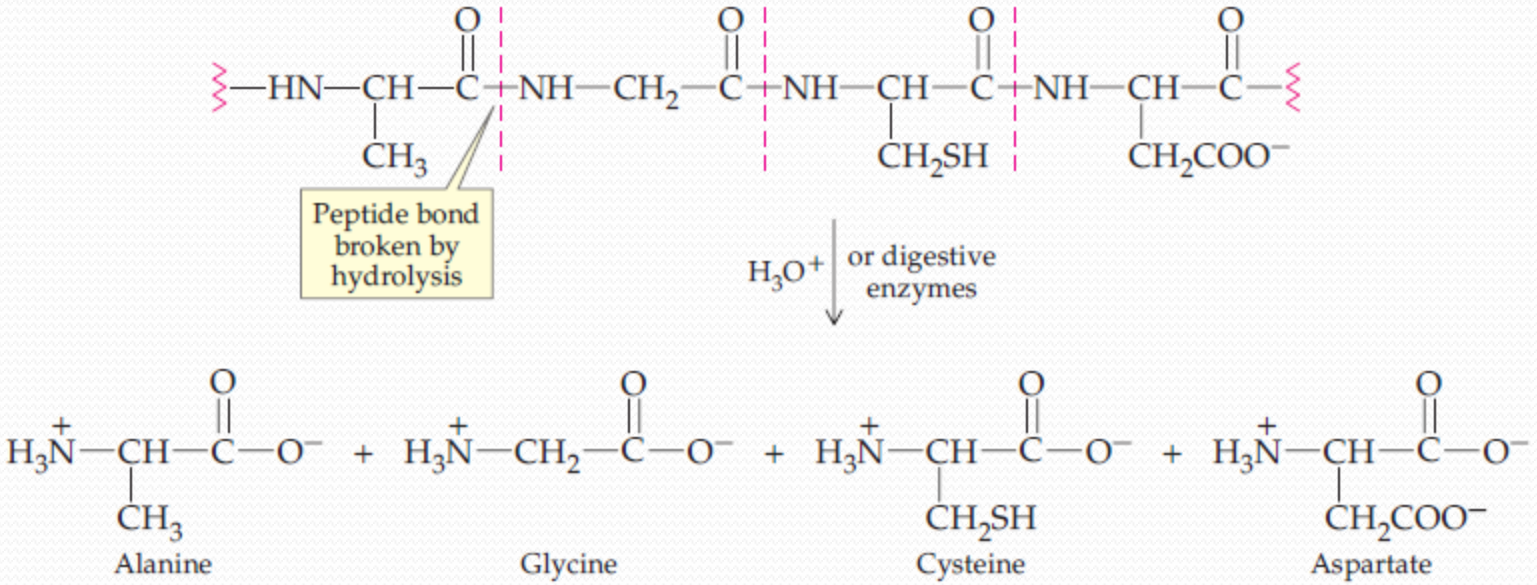
- Phosphates (Phosphoproteins):
  - Esterified to Ser, Thr, or Tyr
  - Usually regulates protein function



Chemical Properties of Proteins

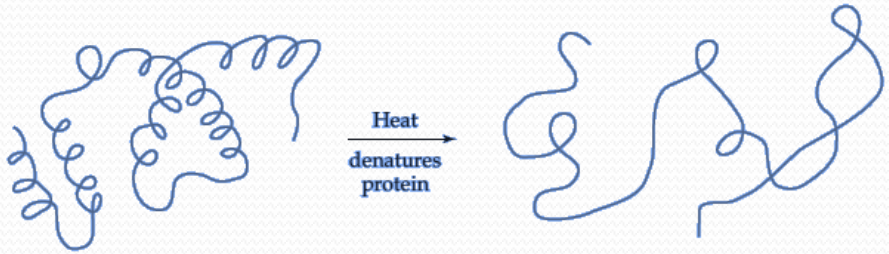
1. Protein Hydrolysis

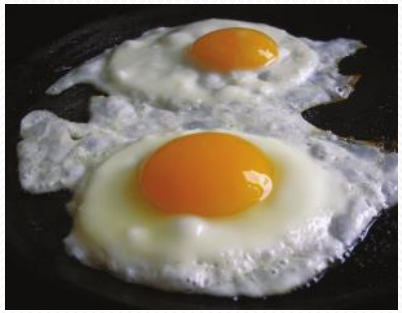
- The reverse of protein synthesis
- Digestion of proteins is hydrolyzing peptide bonds
- Takes place in the stomach and small intestine



2. Protein Denaturation

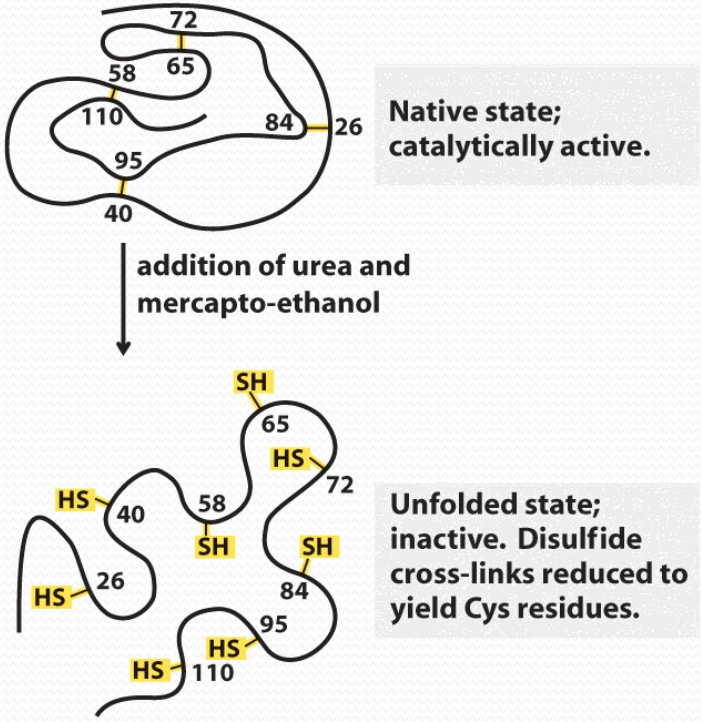
- How the protein preserve its shape?
- What is denaturation? It affects physical, chemical, and biological properties, such as enzymes
- Solubility decreased
- Causes:
  - Heat (≈≥50 °C): low-energy van der Waals forces & H-bonding
  - Mechanical agitation
  - Detergents: hydrophobic forces
    - Triton X-100 (nonionic, uncharged)
    - Sodium dodecyl sulfate (SDS, anionic, charged) - also electrostatic interactions



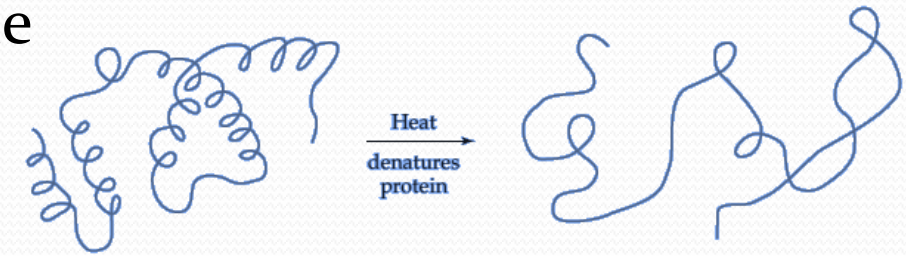


2. Protein Denaturation

- Causes:
  - Organic compounds: acetone, ethanol, bacterial proteins
  - pH change: disrupt salt bridges & H-bonding
    - Urea and guanidine hydrochloride
  - Reducing agents: disulfide bonds
    - β-mercaptoethanol (βME) and dithiothreitol (DTT)



- Most denaturation is irreversible (renaturation)

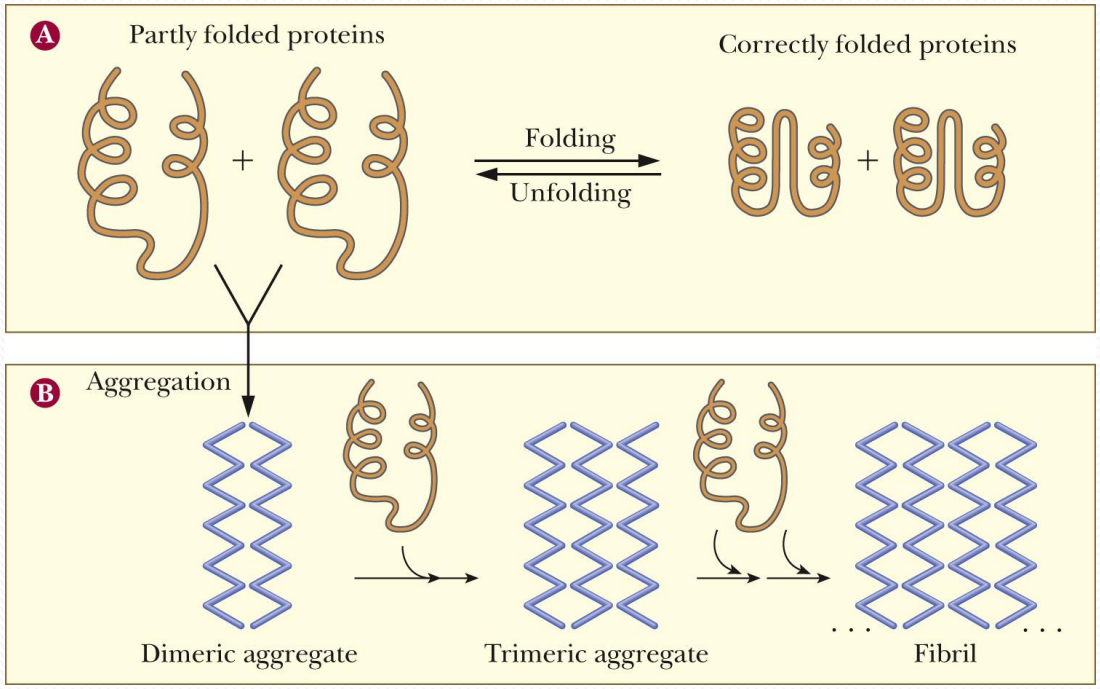


Factors that determine protein structure

- The least amount of energy needed to stabilize the protein. This is determined by:
  - The amino acid sequence (the primary structure), mainly the internal residues - hydrophobic
  - The proper angles between the amino acids
    - The different sets of weak noncovalent bonds that form between the atoms in the polypeptide backbone and in the amino acid side chains
    - Non-protein molecules
  - Chaperones

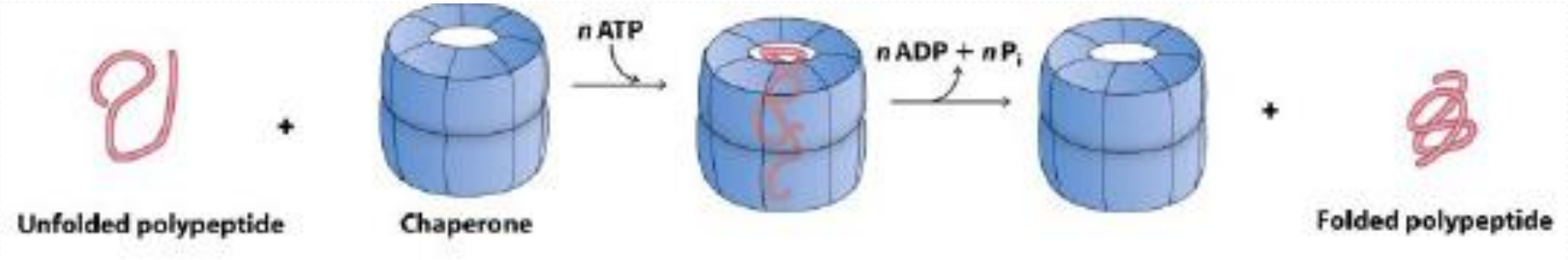
The problem of misfolding

- Hydrophobic interactions are spontaneous
- When proteins do not fold correctly, their internal hydrophobic regions become exposed and interact with other hydrophobic regions on other molecules, and form aggregates



Problem solvers: chaperones

- Aid in correct & timely folding of many proteins
- Exist in organisms from prokaryotes to humans
- hsp70 were the first chaperone proteins discovered
- Function:
  - Help them fold with the most energetically favorable folding pathway
  - Prevent the hydrophobic regions in newly synthesized protein chains from associating with each other to form protein aggregates

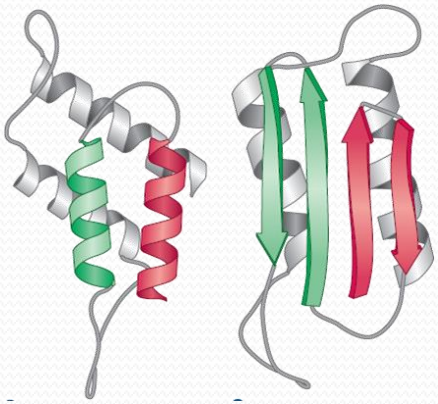


Folding & Diseases

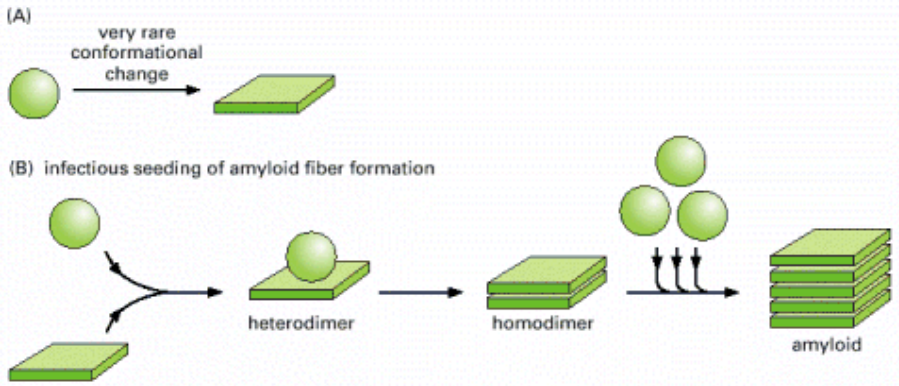
Outcome of protein misfolding

- Partly folded or misfolded polypeptides or fragments may associate with similar chains to form aggregates
- Aggregates vary in size from soluble dimers and trimers up to insoluble fibrillar structures (amyloid)
- Both soluble and insoluble aggregates can be toxic to cells

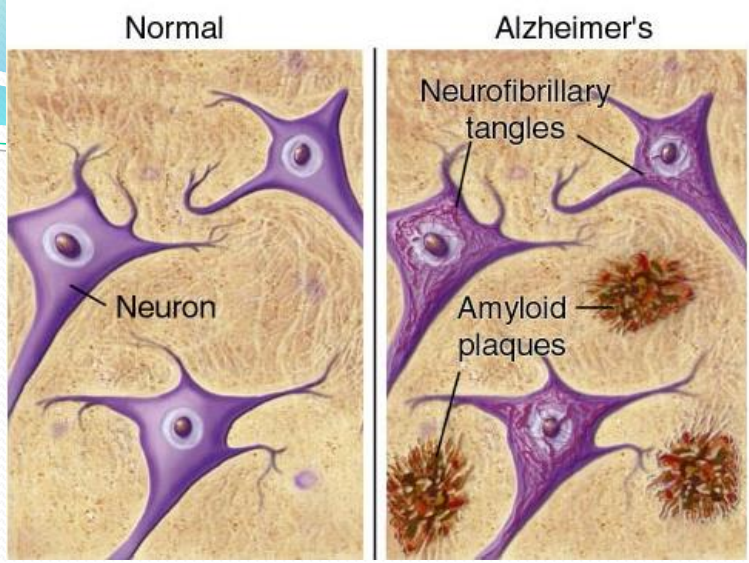
Prion disease



- Prion diseases:
  - Creutzfeldt-Jacob disease (in humans)
  - Mad cow disease (in cows)
  - Scrapie (in sheep) - قعاص الغنم أو الراعوش
- Prion protein (PrP, 28 kDa) is misfolded into an incorrect form called PrPsc - (Met_{129})
- PrPC has a lot of α-helical conformation, but PrPsc has more β strands forming aggregates
- Abnormal protein can be acquired by:
  - Infection
  - Inheritance
  - Spontaneously



Alzheimer's Disease



- Not transmissible between individuals
- Aβ (≈40 a.a) is a short peptide derived from a larger protein (amyloid precursor protein, APP)
- Extracellular plaques of protein aggregates of a protein called *tau* & another known as amyloid peptides (Aβ) damage neurons