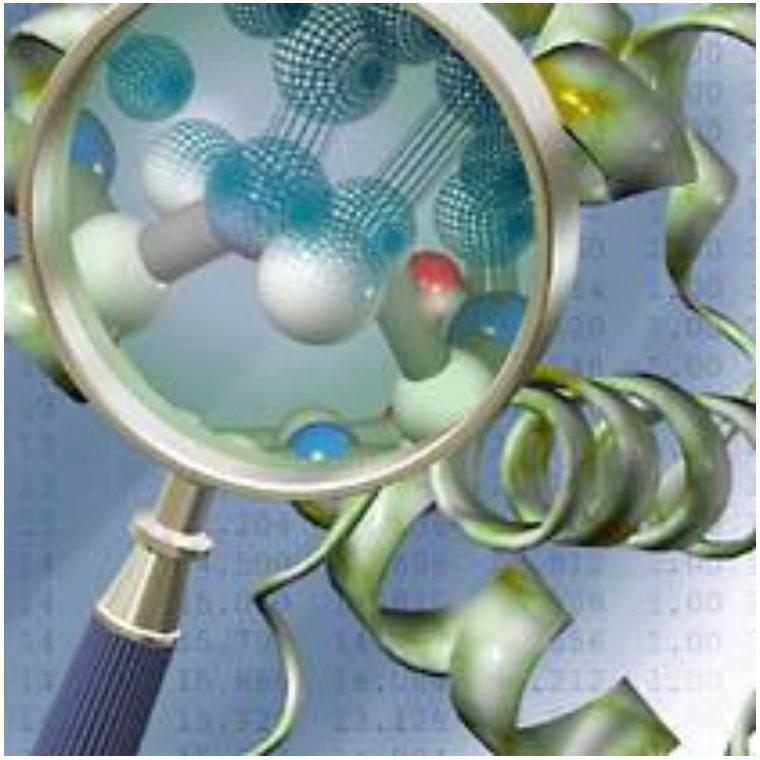
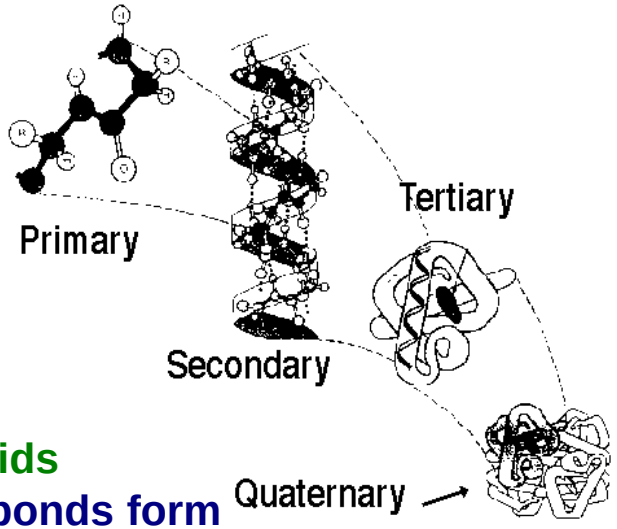
Biochemistry



2. 2 Protein structure and function

Protein structure & function

Proteins are the most versatile macromolecules in living systems, and serve crucial functions in essentially all biological processes



3.1 Proteins are built from a repertoire of 20 amino acids

3.2 Primary structure: amino acids linked by peptide bonds form polypeptide chains

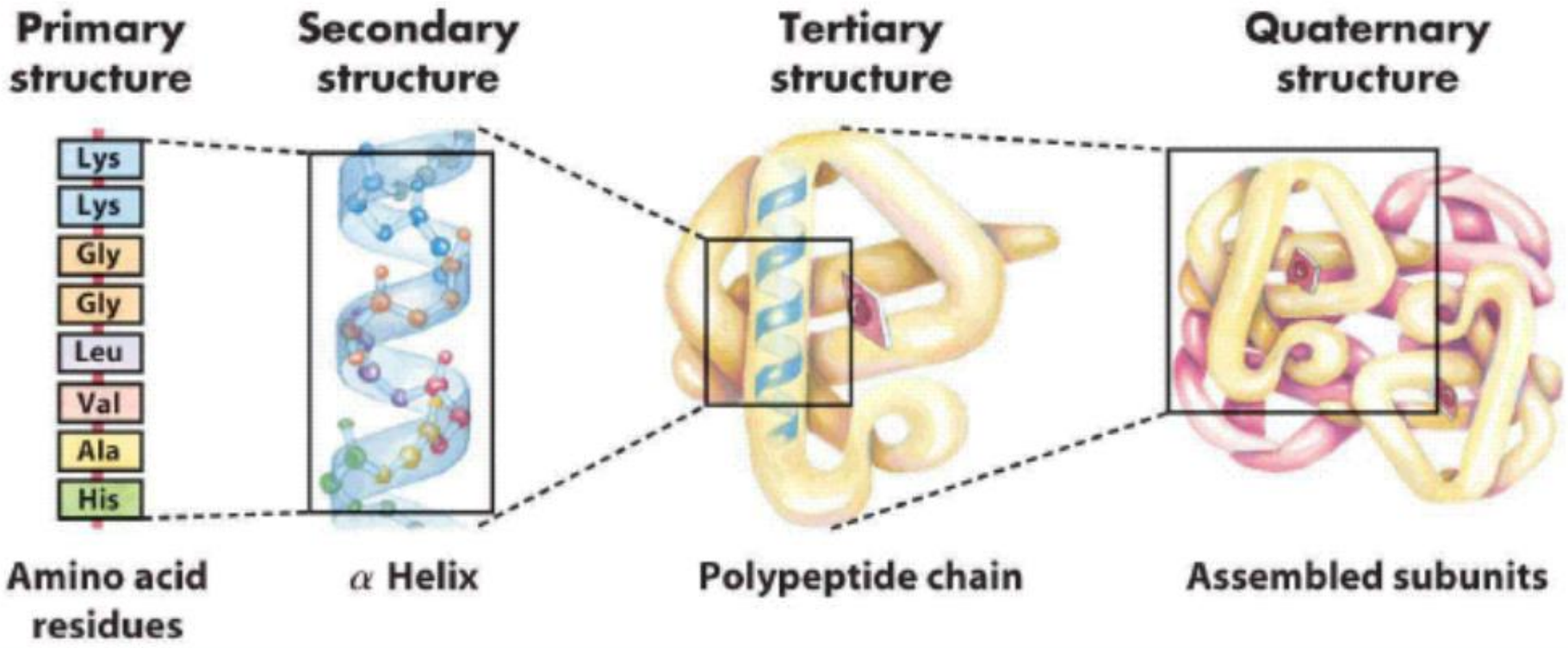
3.3 Secondary structure: polypeptide chains fold into regular structures such as alpha helix, beta sheet, & turns & loops

3.4 Tertiary structure: water-soluble proteins fold into compact structures with nonpolar cores

3.5 Quaternary structure: polypeptide chains can assemble into multisubunit structures

3.6 The amino acid sequence of a protein determines its three-dimensional structure

Protein structure



Proteins - Key properties - a wide range of functions

1. Proteins are linear polymers built of monomer units called amino acids - spontaneously fold into 3-dimensional structures

2. Proteins contain a wide range of functional groups - alcohols, thiols, thioethers, carboxylic acids, carboxamides, & a variety of basic groups - eg. chemical reactivity essential to function of ***enzymes***

3. Proteins can interact with one another, & with other biological macromolecules to form complex assemblies - macromolecular machines

4. Some proteins are quite rigid, whereas others display limited flexibility - structural elements in the cytoskeleton *v* parts that act as hinges, springs, & levers *etc*

Protein functions (test)

- **1) Catalysis** – Enzymes (the proteins that direct and accelerate 1000 biochem. reactions, mild condition, temperature)
- **2) Structure** (structural materials, provide protection and support, specific properties – collagen, elastin, fibroin)
- **3) Movement** (all type of cell movement, actin, tubulin, other cytoskeleton proteins)
- **4) Defense** (protective role, keratin-skin cells, bloodclotting proteins-fibrinogen, thrombin; immunoglobullins)
- **5) Regulation** (peptide hormones insilin, glucagon, growth hormone)
- **6) Transport** (carriers of molecules or ion across membrane, Na+K+ATPase, glucose transporter, hemoglobin, lipoproteins)

Primary structure

charge and the amide nitrogen a partial positive charge, setting up a small electric dipole. Virtually all peptide bonds in proteins occur in this trans configuration; an exception is noted in Figure 4–8b.

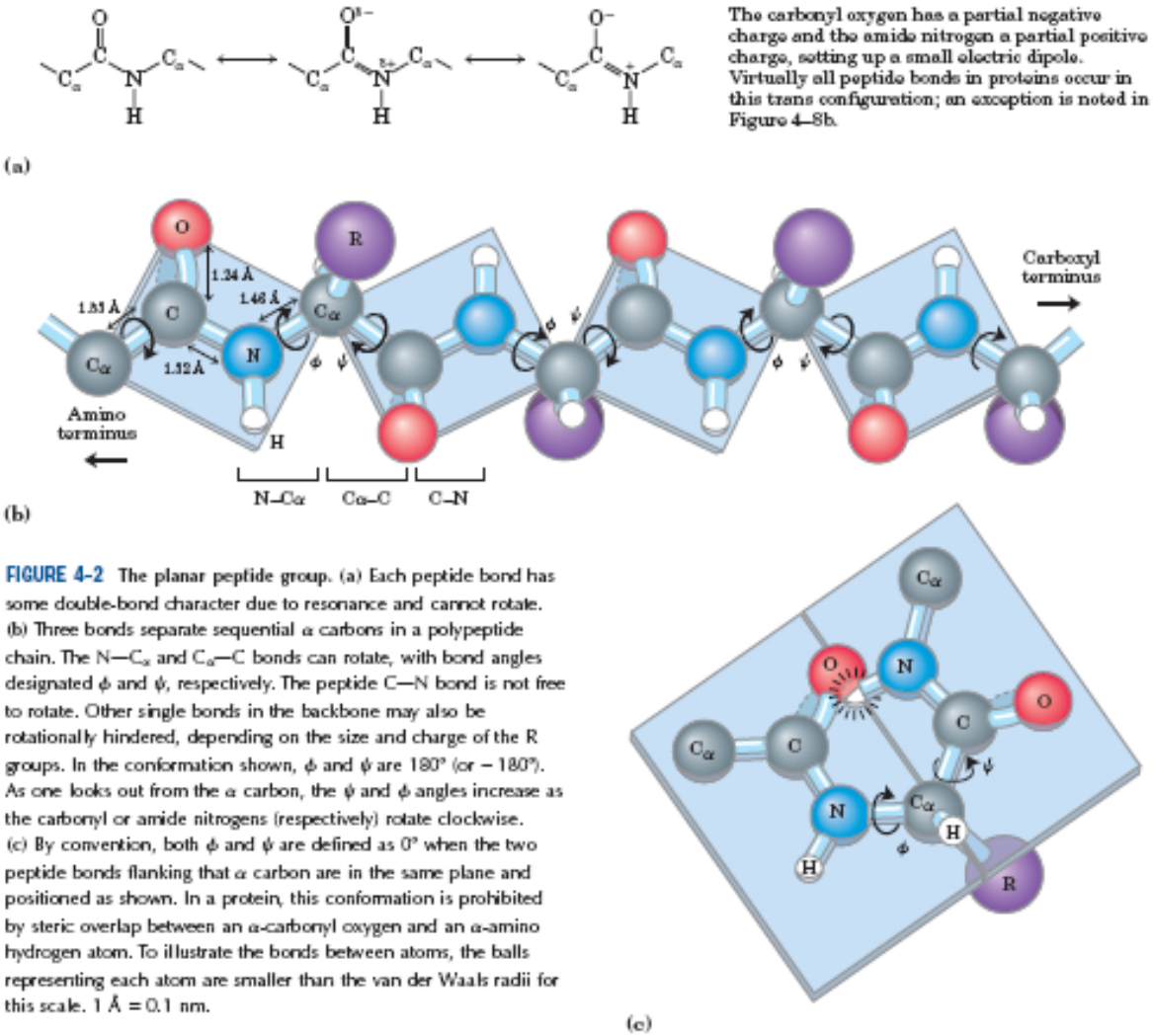
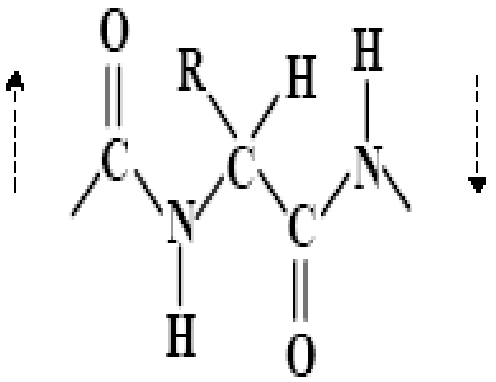


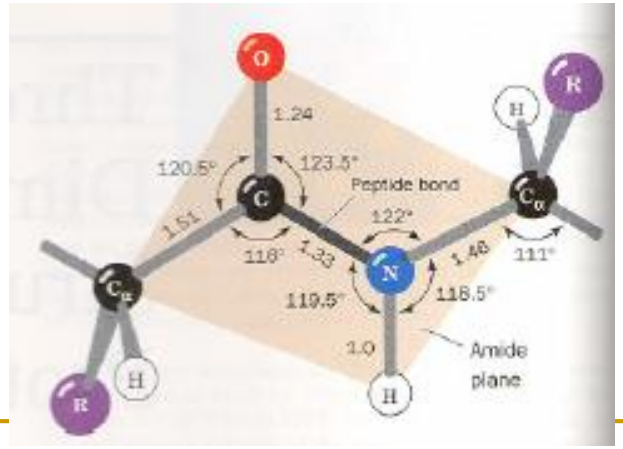
FIGURE 4–2 The planar peptide group. (a) Each peptide bond has some double-bond character due to resonance and cannot rotate. (b) Three bonds separate sequential α carbons in a polypeptide chain. The N—C_α and C_α—C bonds can rotate, with bond angles designated ϕ and ψ, respectively. The peptide C—N bond is not free to rotate. Other single bonds in the backbone may also be rotationally hindered, depending on the size and charge of the R groups. In the conformation shown, ϕ and ψ are 180° (or −180°). As one looks out from the α carbon, the ψ and ϕ angles increase as the carbonyl or amide nitrogens (respectively) rotate clockwise. (c) By convention, both ϕ and ψ are defined as 0° when the two peptide bonds flanking that α carbon are in the same plane and positioned as shown. In a protein, this conformation is prohibited by steric overlap between an α-carbonyl oxygen and an α-amino hydrogen atom. To illustrate the bonds between atoms, the balls representing each atom are smaller than the van der Waals radii for this scale. 1 Å = 0.1 nm.

Peptide bonds- trans configuration

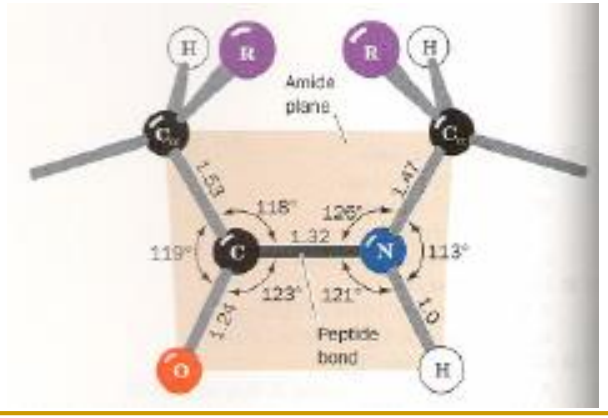
B. Peptidické vazby jsou v trans konfiguraci



Trans

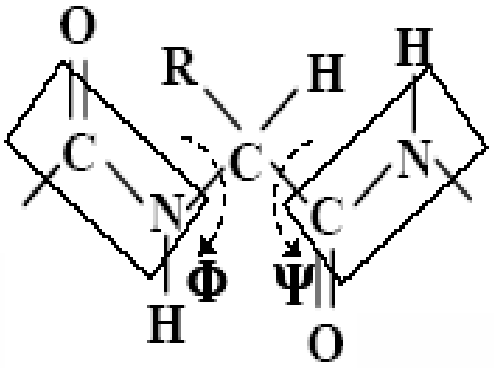


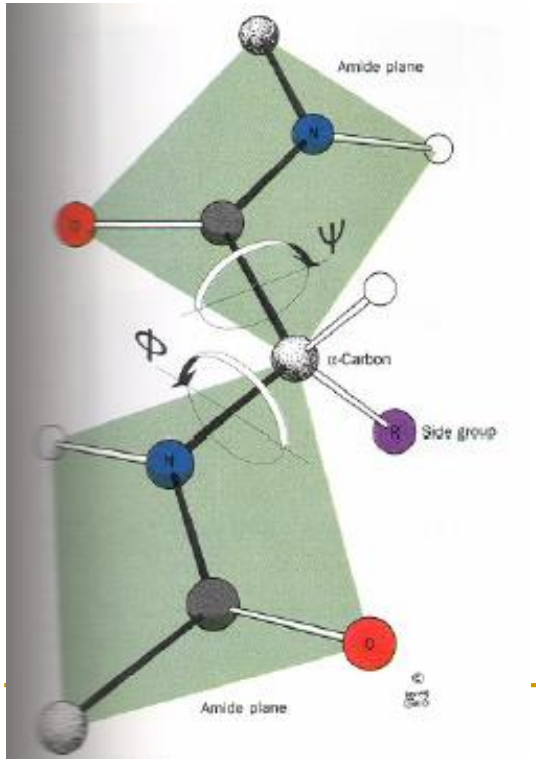
Cis

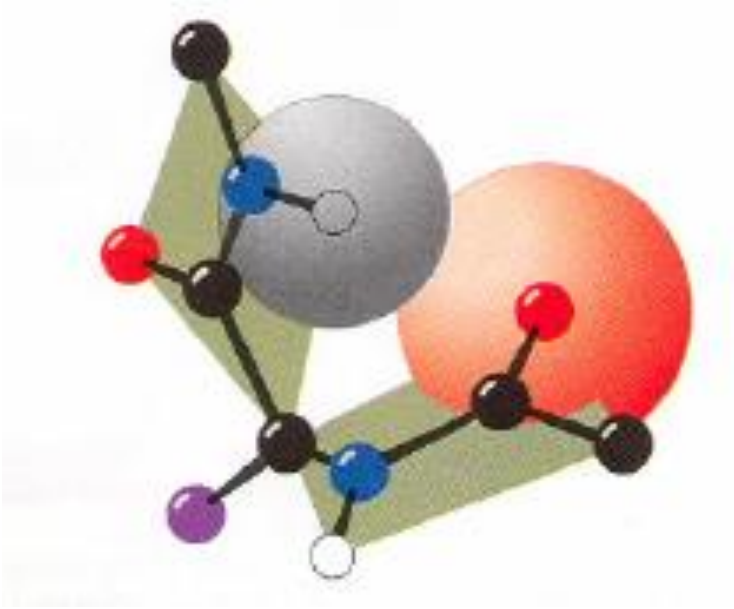


C.Peptidické vazby ležící v rovině mohou svírat určité torzní uhly

ϕ, φ







Secondary structures of proteins

- Refers local conformation of some part of polypeptide

Types:

- α helix
- β sheet
- β loop
- **Hydrogen bond** between carbonyl group and N-H in polypeptide

Linus Pauling, 1901–1994

Robert Corey, 1897–1971

Pauling and Corey
predicted the existence of these secondary structures
in 1951, several years before the first complete protein
structure was elucidated.

α - helix

-Rigid rodlike structure,

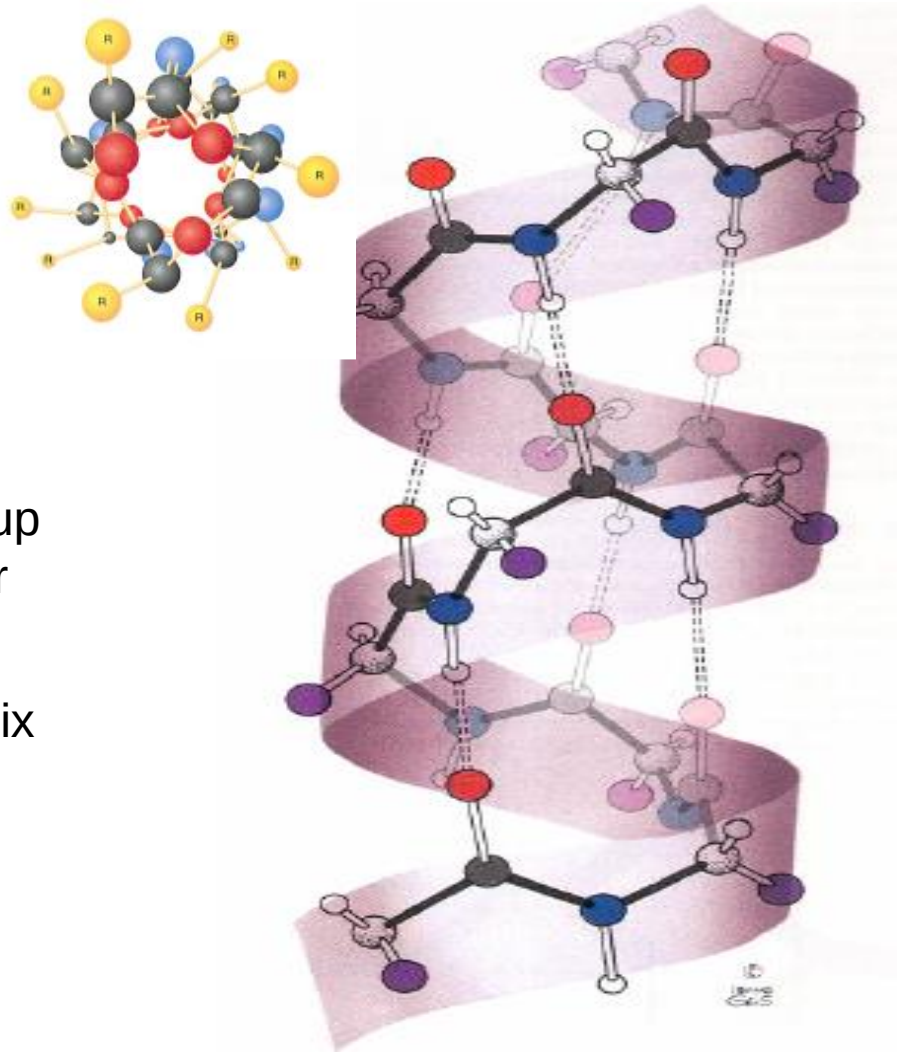
- **right handed** coiled springlike conformation

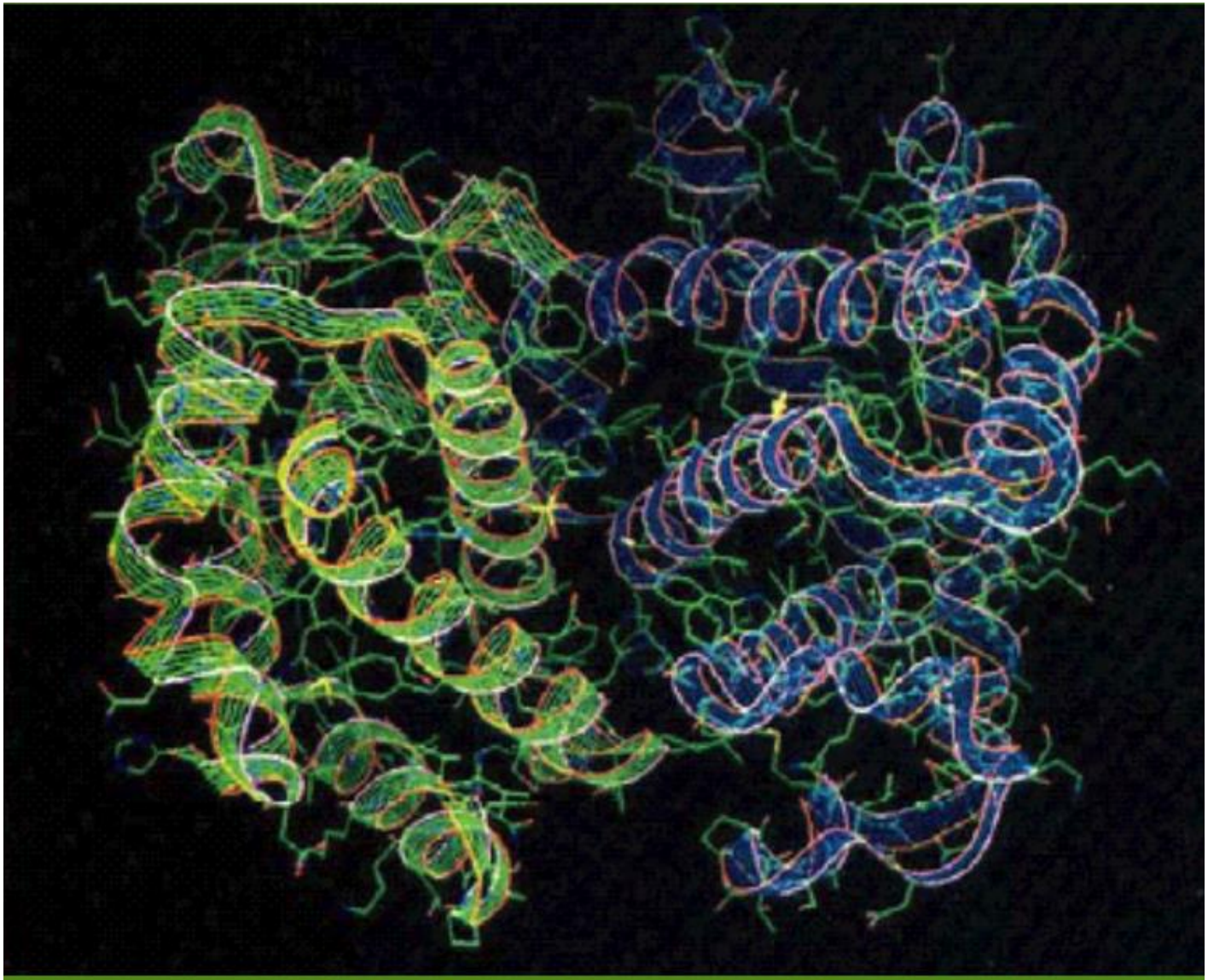
- 3.6 AA per turn of the helix

-Hydrogen bonds between N-H group of each AA and CO group of AA four residues away

- **Collagen** –prolin – left handed helix

-Charged AA and Try incompatible with a-helix



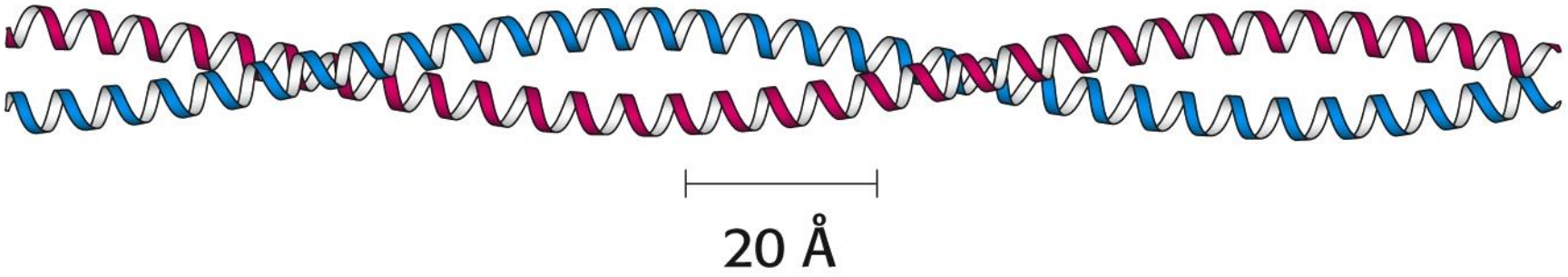


Alfa helixy v hemoglobinu

Super helix: alpha helical coiled coil

Can be as long as 1000 Å, very stable

Helical cables in these proteins serve a mechanical role, forming stiff bundles of fibers



Found in: • **myosin** and tropomyosin in muscle,
• **fibrin** in blood clots,
• **keratin** in hair, quills, claws, hoofs, & horns
• intermediate filaments
(cytoskeleton or internal scaffolding of cells)

β - sheet

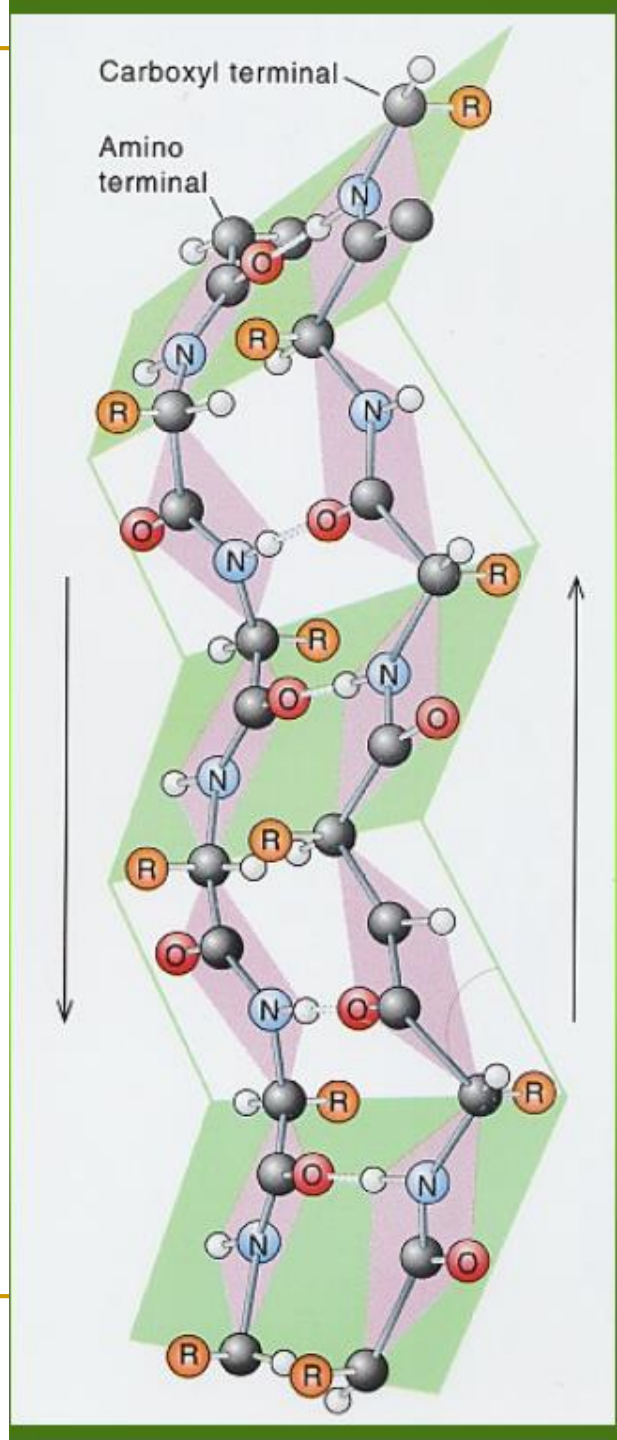
-Two and more polypeptide chains line up side by side,

-Each polypeptide is fully extended

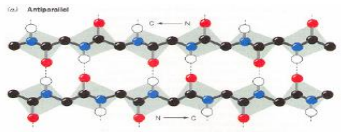
-Hydrogen bonds polypeptide backbone N-H and CO adjacent chains

-Parallel (same directions)

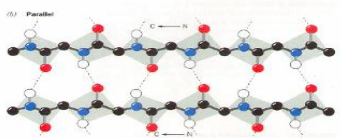
-Antiparallel (opposite direction)



Antiparallel

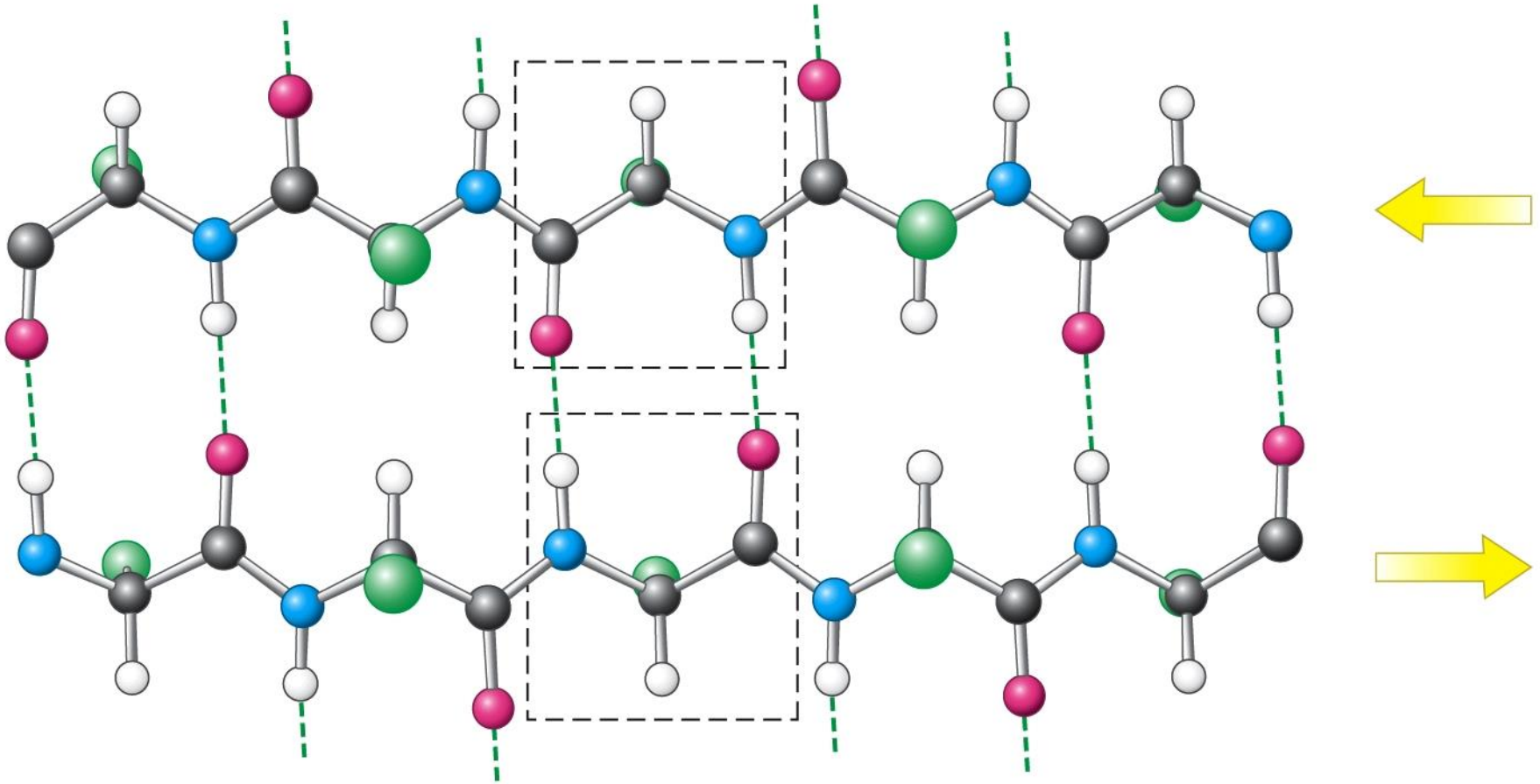


parallel



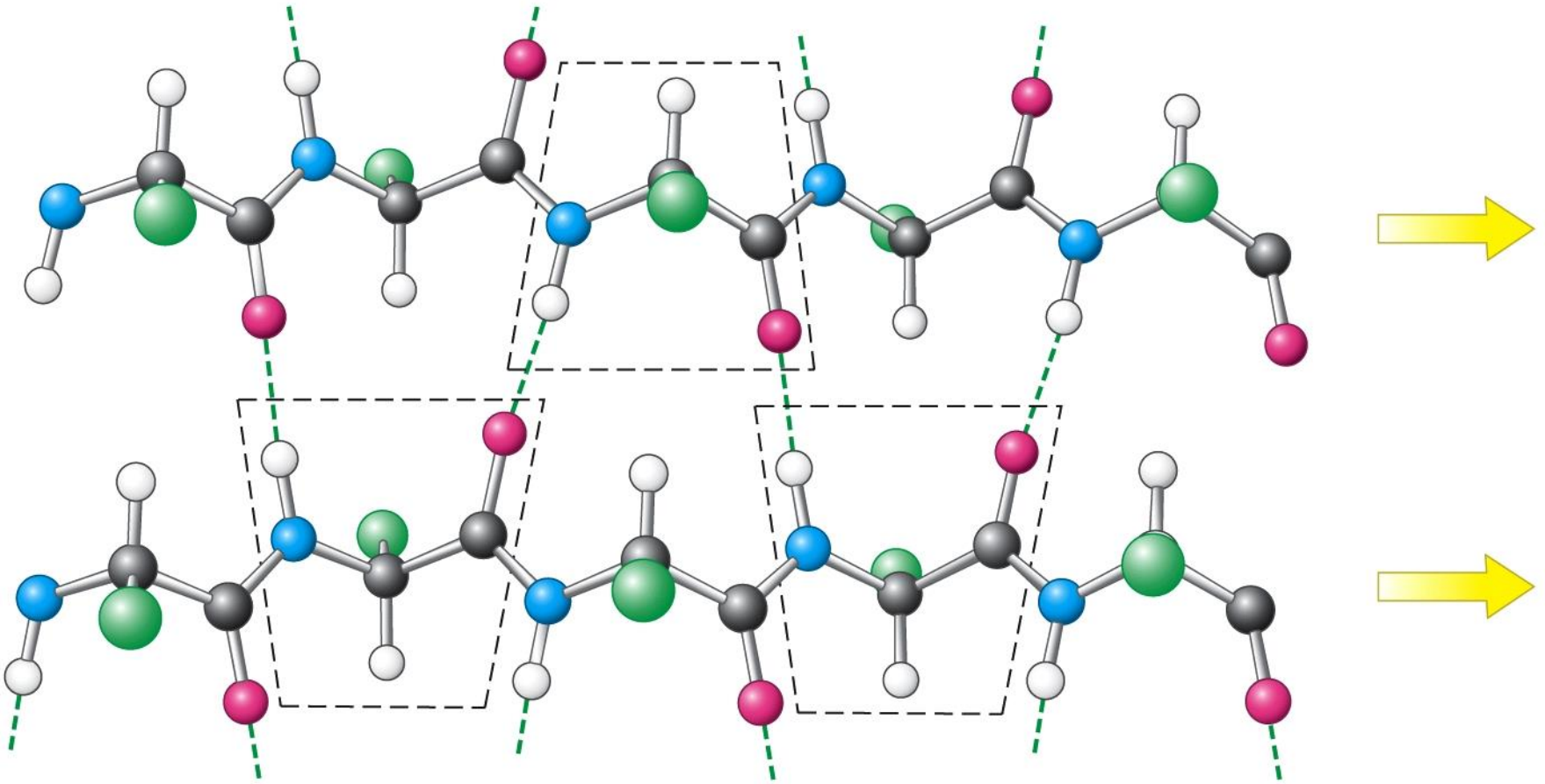
Antiparallel beta sheet

Strands linked by H-bonding between opposite amino acids

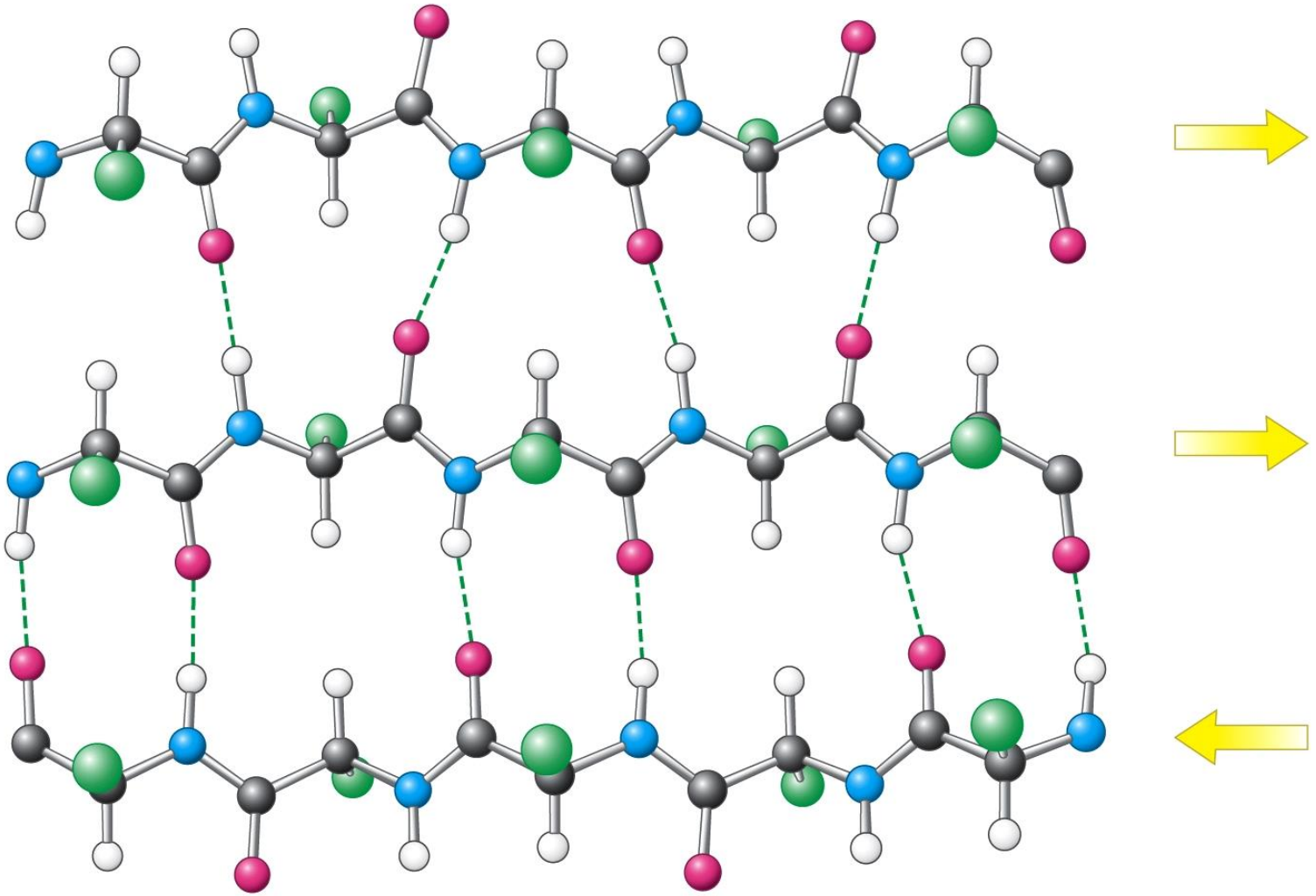


Parallel beta sheet

Strands linked by H-bonding of an aa on one strand to two different aa on the adjacent strand

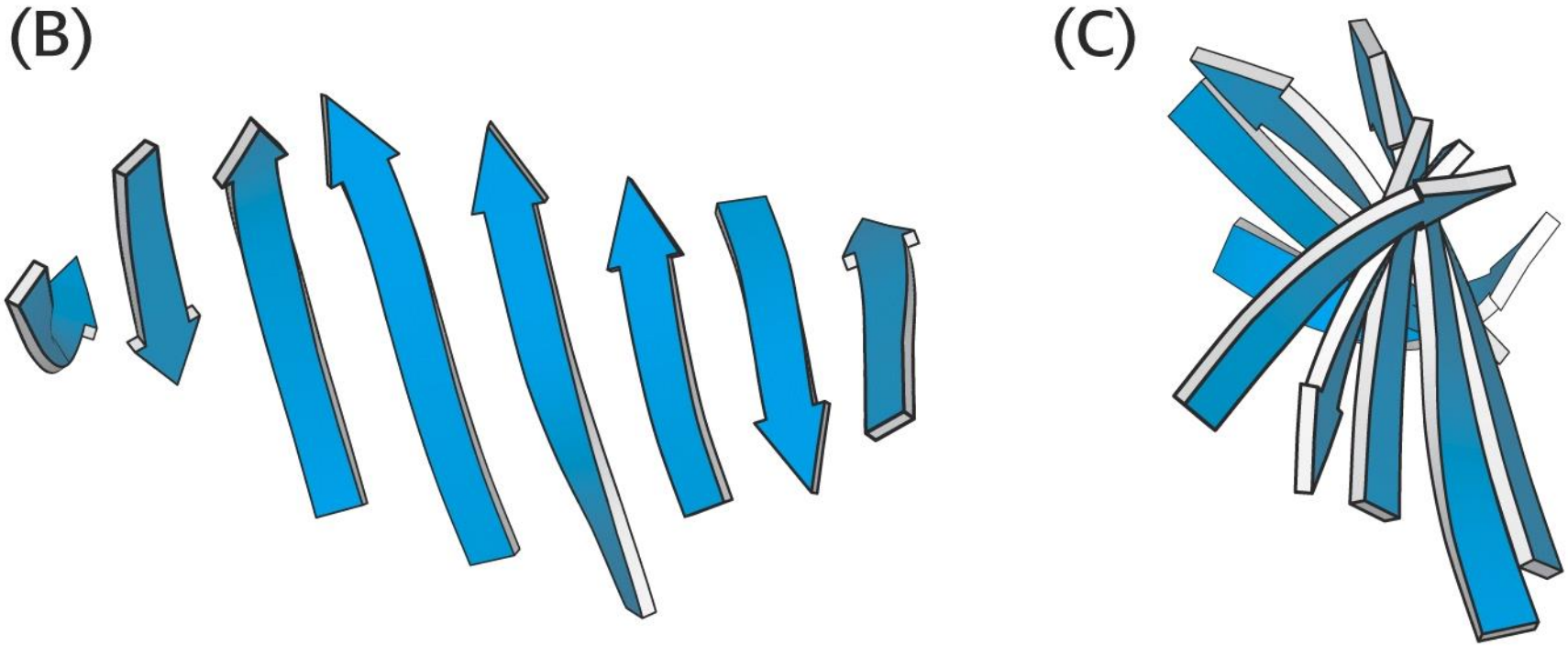


Structure of mixed beta sheet



A twisted beta sheet, schematic model

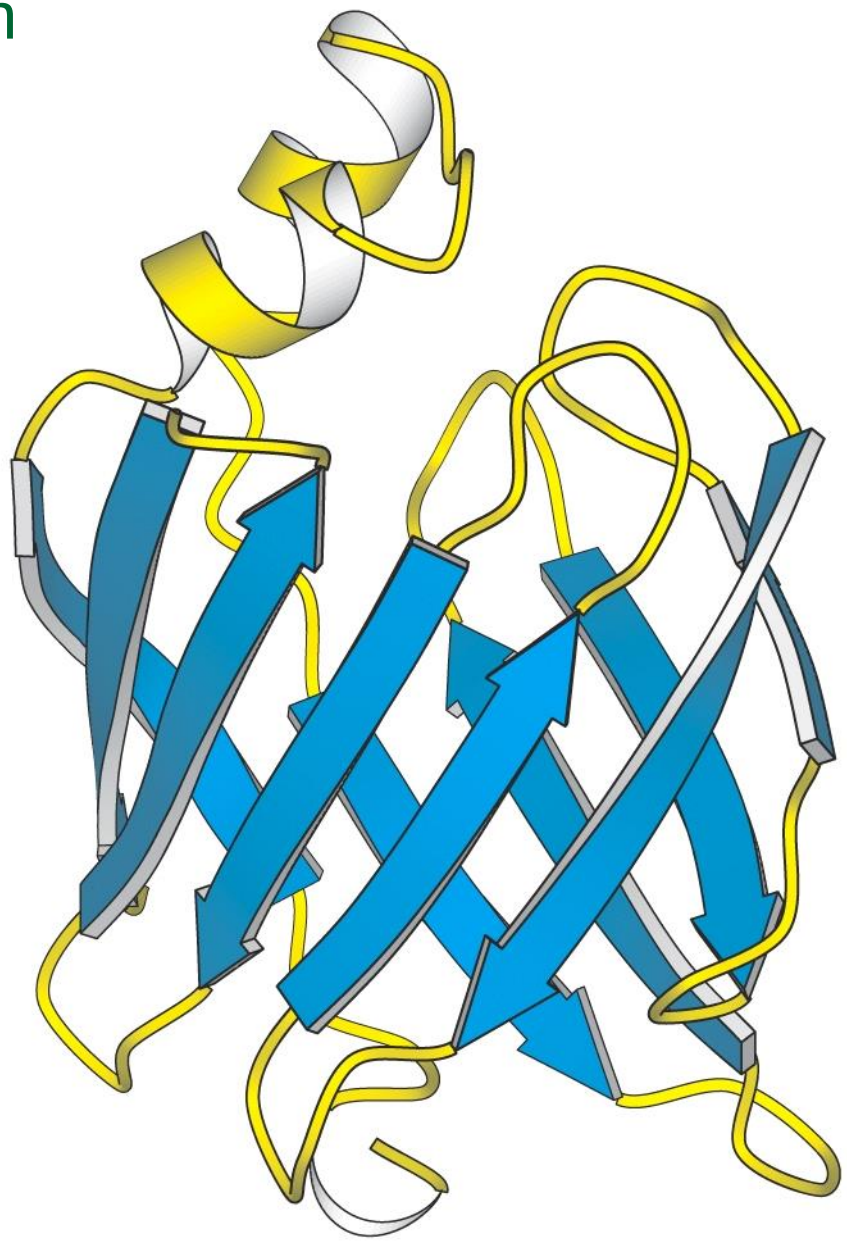
Rotated 90 degrees



Fatty acid-binding protein

Rich in beta sheets

Arrow pointing to carboxyl-terminal end



Relative frequency of aa in secondary structures

TABLE 3.3 **Relative frequencies of amino acid residues in secondary structures**

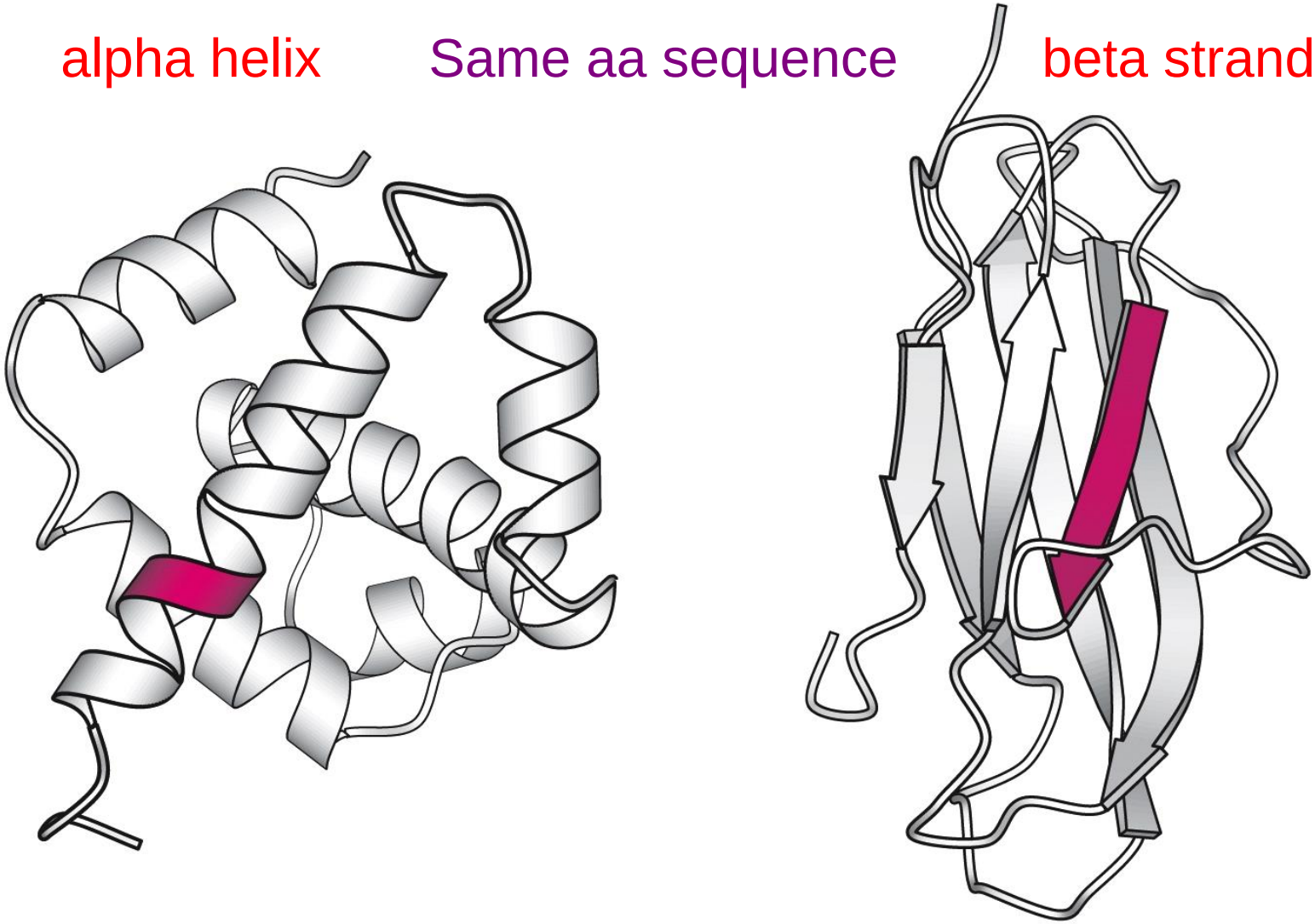
Amino acid	α helix	β sheet	Turn
Ala	**1.29**	0.90	0.78
Cys	**1.11**	0.74	0.80
Leu	**1.30**	1.02	0.59
Met	**1.47**	0.97	0.39
Glu	**1.44**	0.75	1.00
Gln	**1.27**	0.80	0.97
His	**1.22**	1.08	0.69
Lys	**1.23**	0.77	0.96
Val	0.91	**1.49**	0.47
Ile	0.97	**1.45**	0.51
Phe	1.07	**1.32**	0.58
Tyr	0.72	**1.25**	1.05
Trp	0.99	**1.14**	0.75
Thr	0.82	**1.21**	1.03
Gly	0.56	0.92	**1.64**
Ser	0.82	0.95	**1.33**
Asp	1.04	0.72	**1.41**
Asn	0.90	0.76	**1.28**
Pro	0.52	0.64	**1.91**
Arg	0.96	0.99	0.88

Note: The amino acids are grouped according to their preference for α helices (top group), β sheets (second group), or turns (third group). Arginine shows no significant preference for any of the structures.
After T. E. Creighton, *Proteins: Structures and Molecular Properties*, 2d ed. (W. H. Freeman and Company, 1992), p. 256.

Alternative conformations: context

Tertiary interactions (between residues far apart) affect secondary structures

alpha helix Same aa sequence beta strand



Five themes of 3D structure of proteins

1. 3D is determined by AA sequences
2. Function of proteins depends on its structure
3. Isolated protein 1 or small number of stable struc.forms
4. The most important fources- noncovalent interaction
5. Common structural patterns

Every protein has a three-dimensional structure that reflects its function.

Protein conformation

- Spatial arrangement of atoms in protein
- Rotation about single bound
- Thermodynamically the most stable
- Lower Gibson free energy
- **Native proteins**

• Proteins are stabilized by multiple WEAK interactions

• Hydrophobic interactions are the major contributors – globular forms of most soluble proteins

•Hydrogen bonds and ionic interactions are optimized in the specific structures - Thermodynamically the most stable

Protein conformation

the three-dimensional structure

- **Proteins conformation** (covalent bond, free rotation, unlimitied number of conformations)
- each protein has a specific chemical or structural function, strongly suggesting that each has a **unique** three-dimensional structure
    - Spatial arrangement of atoms in protein, Rotation about single bound, Thermodynamically the most stable, Lower Gibson free energy
    - **Native proteins**

• Proteins are stabilized by multiple **WEAK interactions**

• **Hydrophobic interactions** are the major contributors – globular forms of most soluble proteins

• **Hydrogen bonds** and **ionic interactions** are optimized in the specific structures - Thermodynamically the most stable

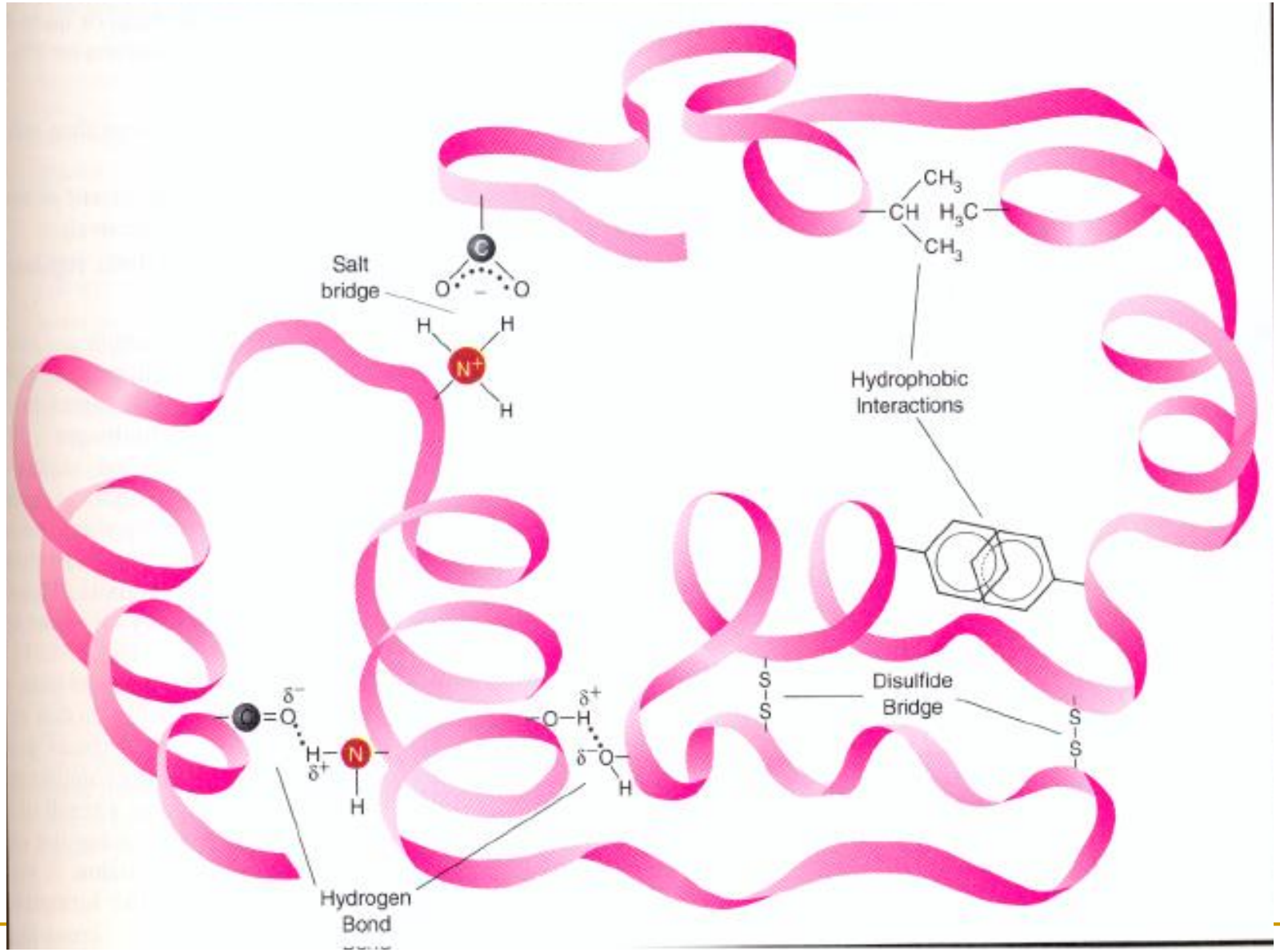
Proteins tertiary and quartery structures

- Complete 3D structure of polypeptides
- 2 general classes: fibrous and globular
- **Fibrous proteins** – structural roles, simple repeating elements of secondary structures
- **Globular proteins** – complicated, several types of 2nd structure, myoglobin (1st proteins X-ray)
- **Domains**- region of proteins which can fold stably and **independently**
- **Quartery structures**: interaction between subunit, promoters

Interaction –stabilized tertiary structure

- Hydrophobic interaction (hydrophobic R groups –close proximity, exclusion water, folding of globular proteins)
- Electrostatic interactions- between ionic groups – salt bridge
- Hydrogen bond
- Covalent bond (disulfide bridges)

TEST


Salt
bridge
Hydrophobic
Interactions
Disulfide
Bridge
Hydrogen
Bond

Fibrous and globular proteins

- **FIBROUS** – High proportions of regular secondary structures (α-helix, β-sheeds)

- long rod-shaped, sheetlike molecules,

- **insoluble in water, physically tough, keratins (skin, hair, nails)**
- α-keratin, colagen, silk fibroin
- **STRUCTURAL, PROTECTIVE FUNCTION**

- **GLOBULAR**

  - compact spherical molecules, usually water soluble,

  - DYNAMIC function : ENZYMES, immunoglobulin's, TRANSPORT

  - **cavities, clefts-complementary to LIGAND**
  - **Hemoglobin-----**

α-keratin

- Hair, wool, skin, fingernails
- A-helical polypeptides
- 3 a helical chains –left handed supercoiled structure – protofibril
- Microfibril
- Macrifibril
- AA- no prolin, ala, leu; R outside- wather insoluble; hard keratins- disulfide (oxidizing)…

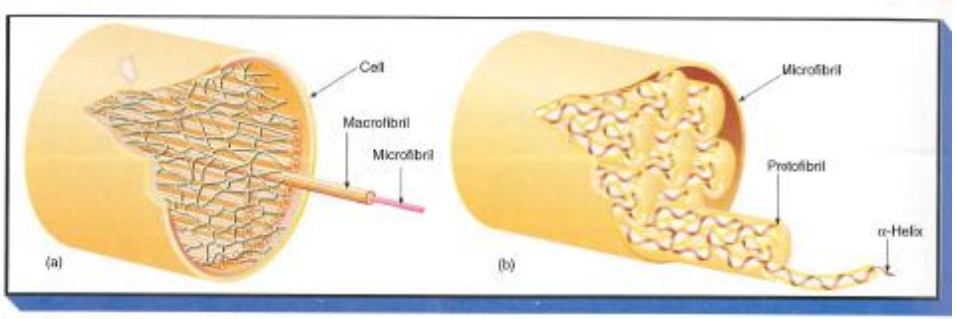
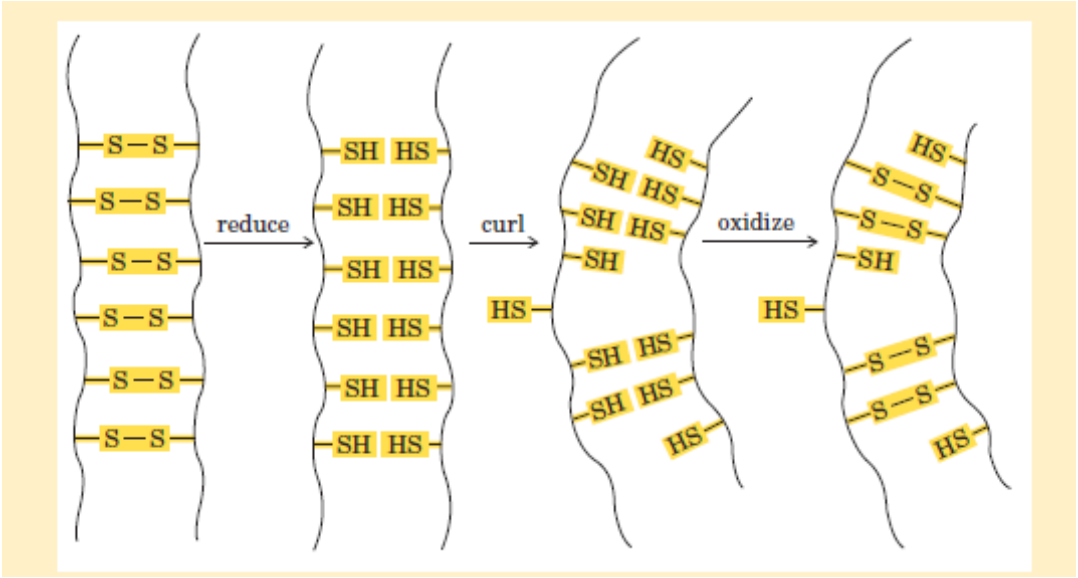


FIGURE 5.22
Hair Structure.
Each hair consists of several dead cells packed with macrofibrils. (a) Macrofibrils are constructed from microfibrils, each of which contains 11 protofibrils. (b) Each protofibril contains 3 α-keratin molecules.



- trvalá ondulace

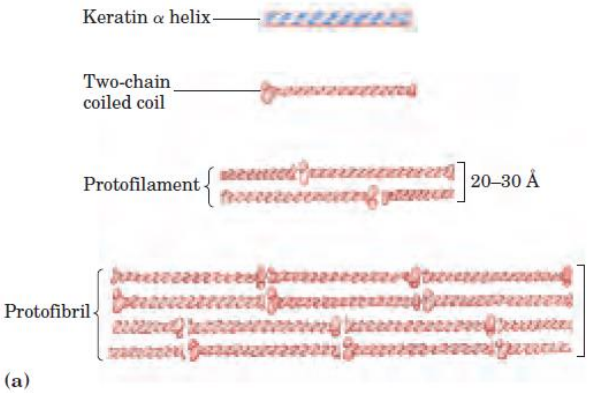
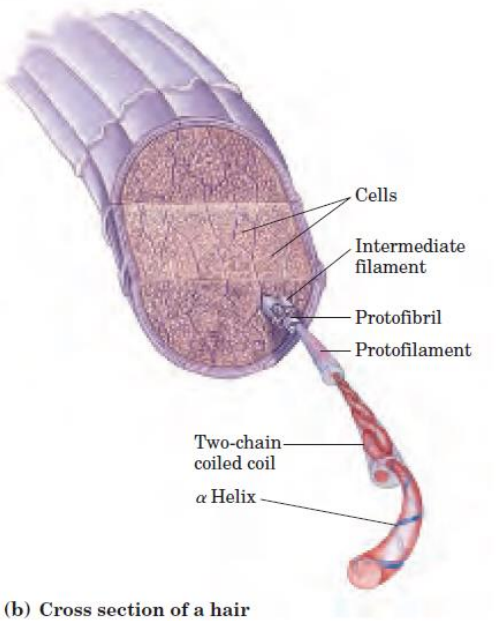
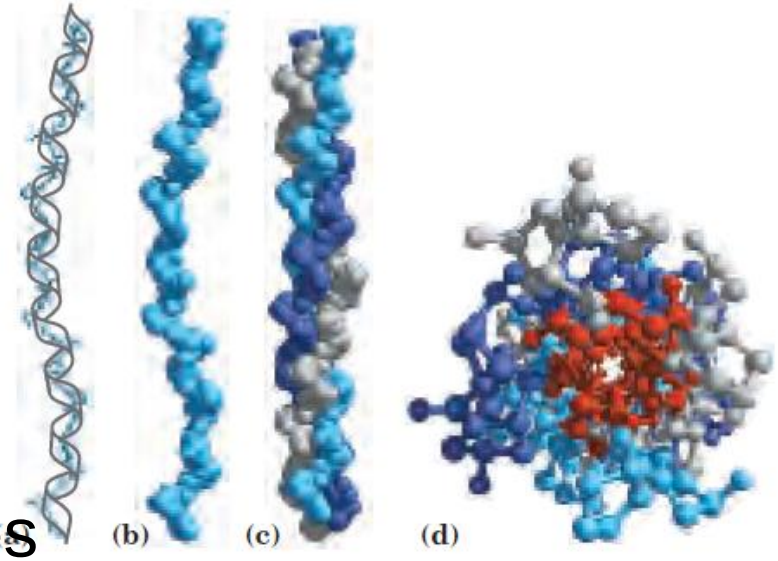


FIGURE 4–11 **Structure of hair.** (a) Hair α-keratin is an elongated α helix with somewhat thicker elements near the amino and carboxyl termini. Pairs of these helices are interwound in a left-handed sense to form two-chain coiled coils. These then combine in higher-order structures called protofilaments and protofibrils. About four protofibrils—32 strands of α-keratin altogether—combine to form an intermediate filament. The individual two-chain coiled coils in the various substructures also appear to be interwound, but the handedness of the interwinding and other structural details are unknown. (b) A hair is an array of many α-keratin filaments, made up of the substructures shown in (a).



Collagen



- Abundant protein in vertebrates
- Connective tissue cells, secreted to extracellular matrices
- In structures: skin, bones, tendors, blood vessels
- 3 left handed polypep. helices, twisted around each other – righ handed superhelix
- AA – 30% glycine, 30% prolin, 4-hydroxyprolin
- ER hydroxylation of pro, lys
- Repeating triplets Gly-X-Y (X,Y often Pro, Hydro-Pro), Y – hydroxy-Lys

Simple, conjugated proteins

- SIMPLE proteins – albumin, keratin (only AA)
- CONJUGATED proteins:

- **Prosthetic group-** nonprotein component
- **Apoprotein** – without prosthetic group
- **Holoprotein** (apo+ prost)

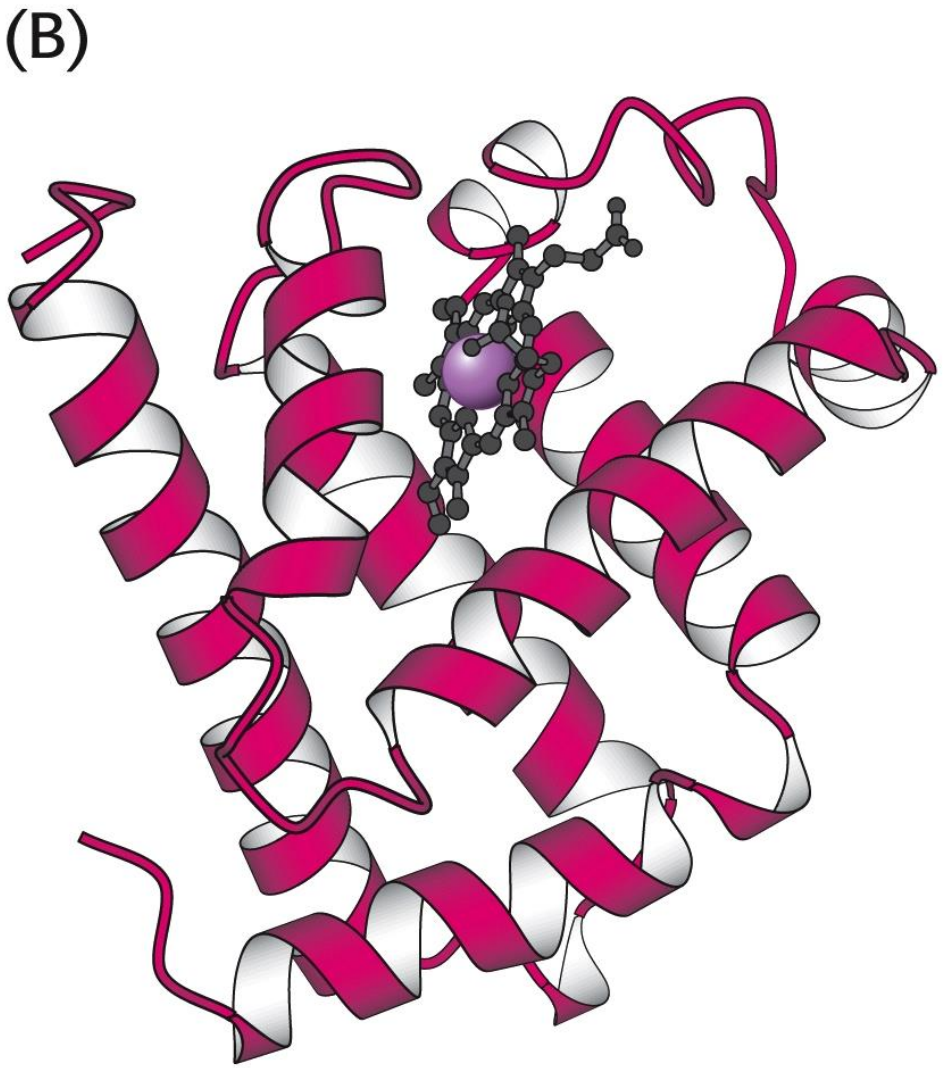
Types: glycoproteins, metaloproteinss, lipoproteins, phosphoproteins, hemoproteins

Primary structure and evolution:

Homologues

Conservative, variable

Tertiary structure, myoglobin, schematic



Mainly alpha helices,
total = 8 helices
(75% of main chain)

Prosthetic (helper) group to bind O_2

Heme group is **protoporphyrin IX, & central iron atom**

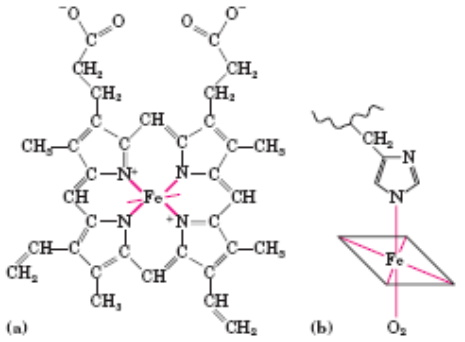
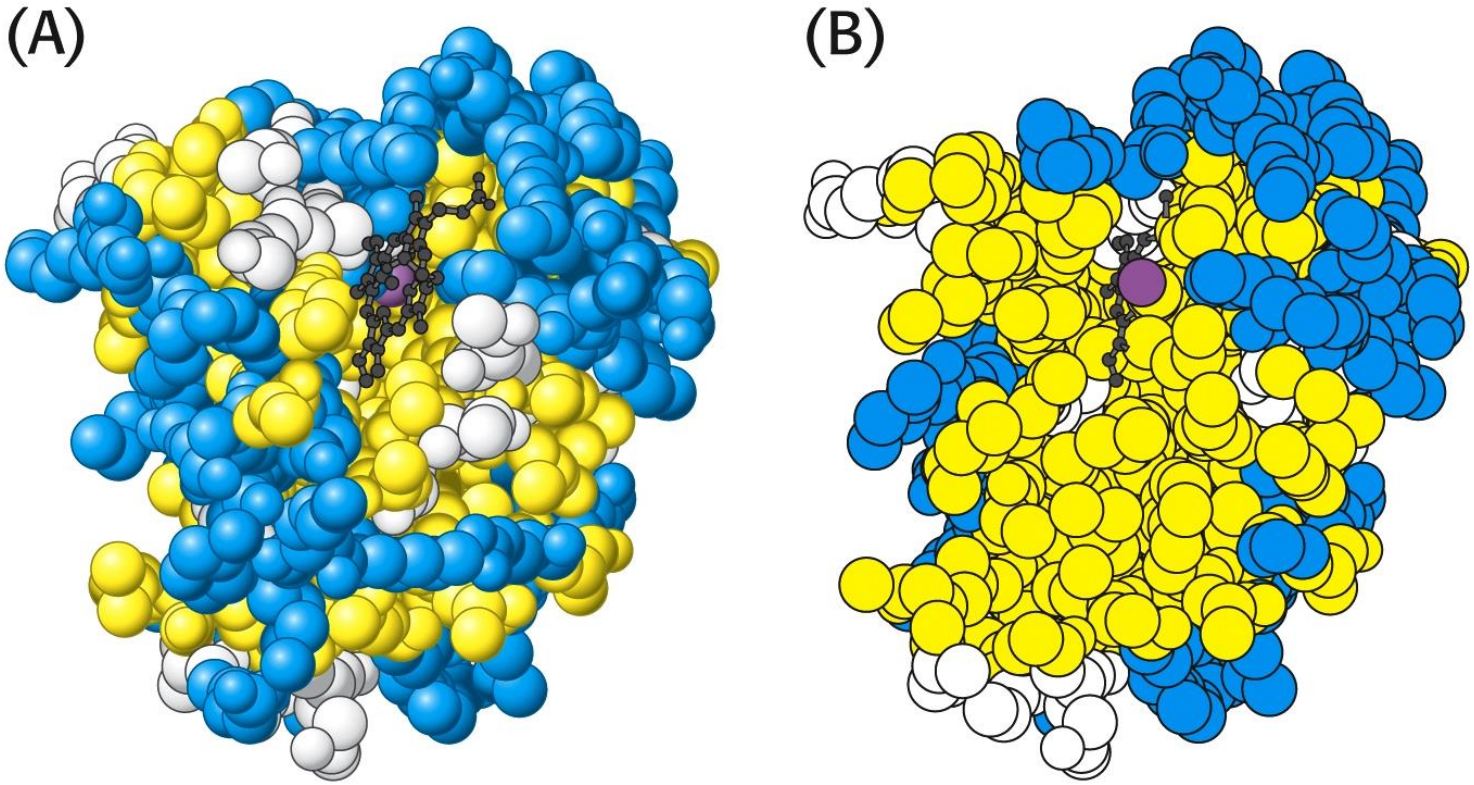


FIGURE 4–17 The heme group. This group is present in myoglobin, hemoglobin, cytochromes, and many other heme proteins. (a) Heme consists of a complex organic ring structure, protoporphyrin, to which is bound an iron atom in its ferrous (Fe^{2+}) state. The iron atom has six coordination bonds, four in the plane of and bonded to the flat por-

Distribution of aa in myoglobin



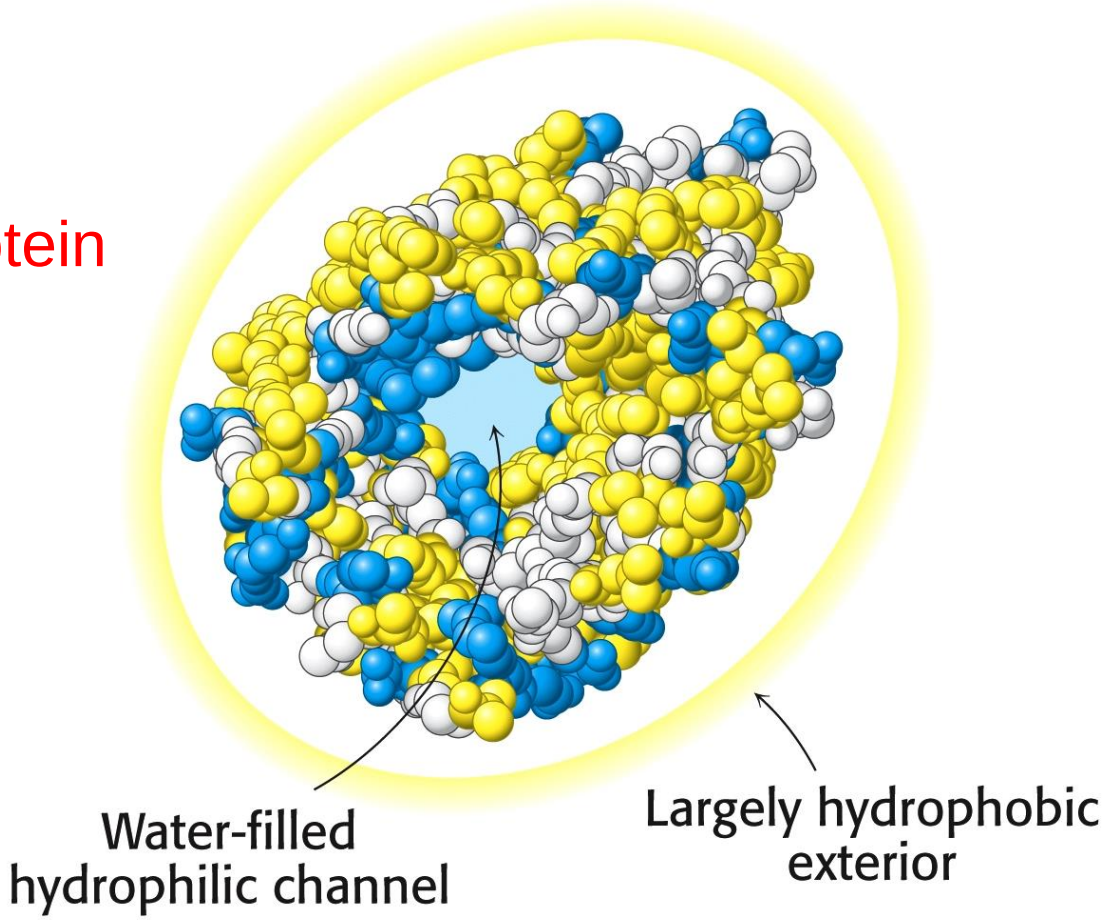
Surface, mainly charged aa. **Interior, mainly hydrophobic aa**

Yellow: hydrophobic aa
Blue: charged aa
White: other aa

Cross-section

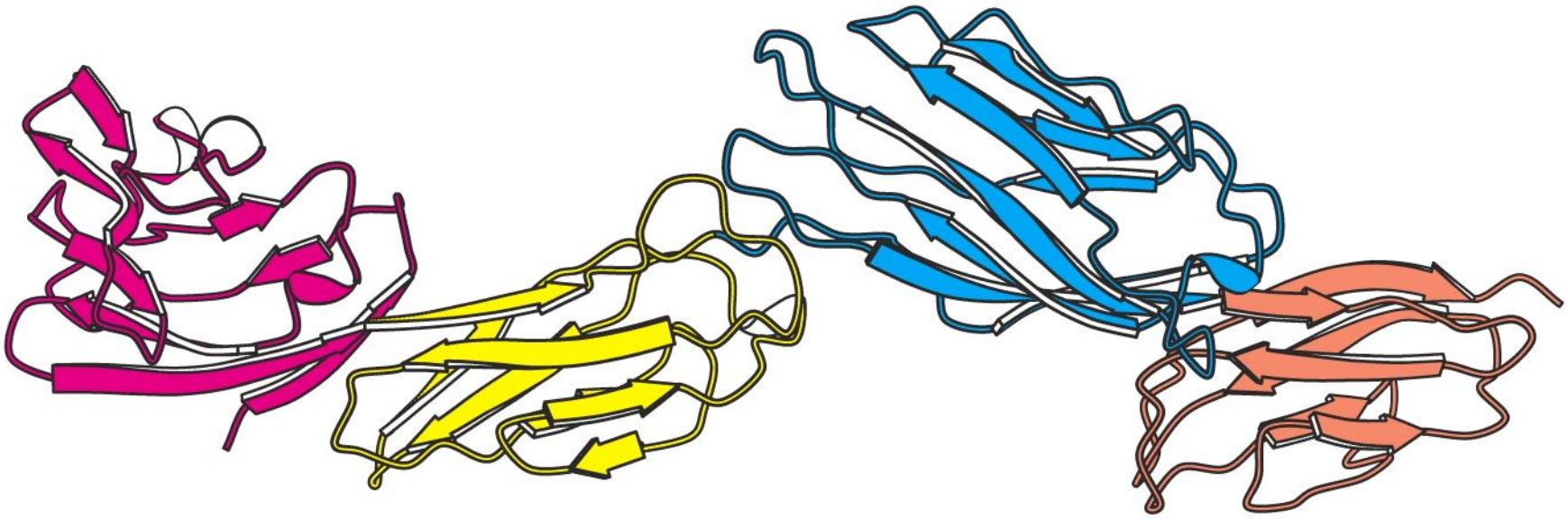
Porin: “inside out”

Membrane protein



Protein domains (single polypeptide)

CD4: cell surface protein (immune system), four similar domains

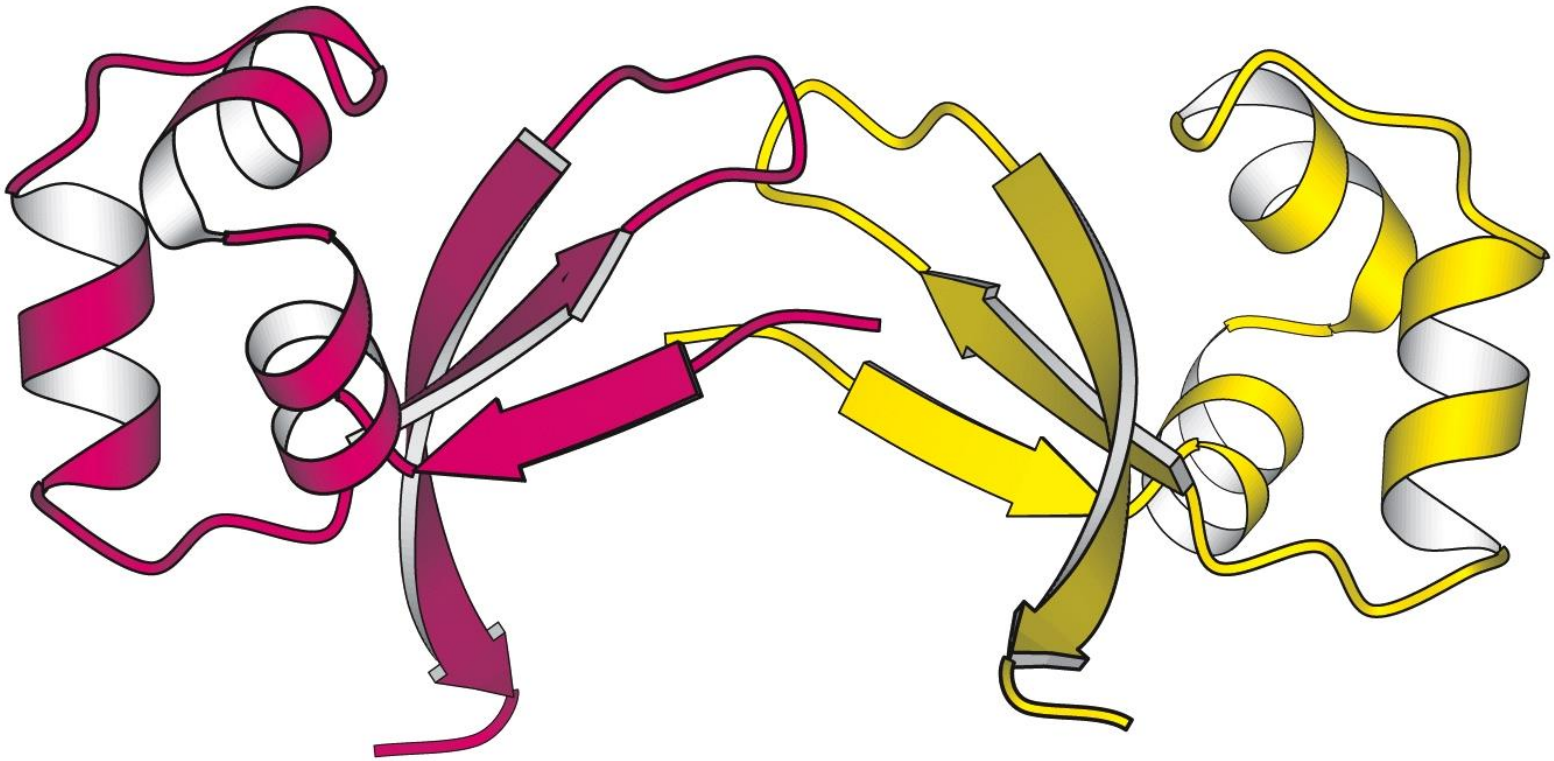


Protein to which HIV attaches

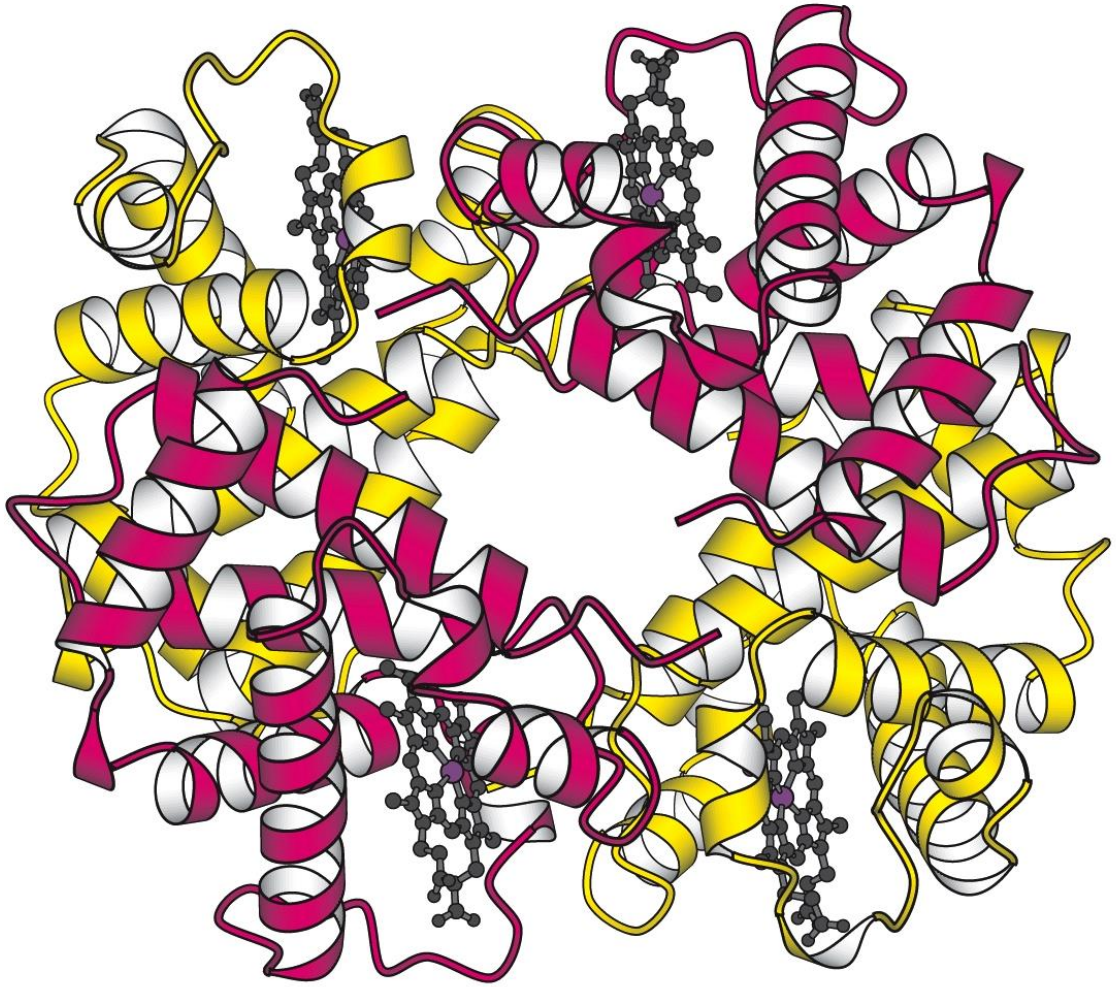
Quaternary structure, dimer

Cro protein of bacteriophage lambda

Dimer of identical subunits



Quaternary structure, tetramer



Human **hemoglobin,** two alpha(red) two beta(yellow) subunits,

4 heme groups

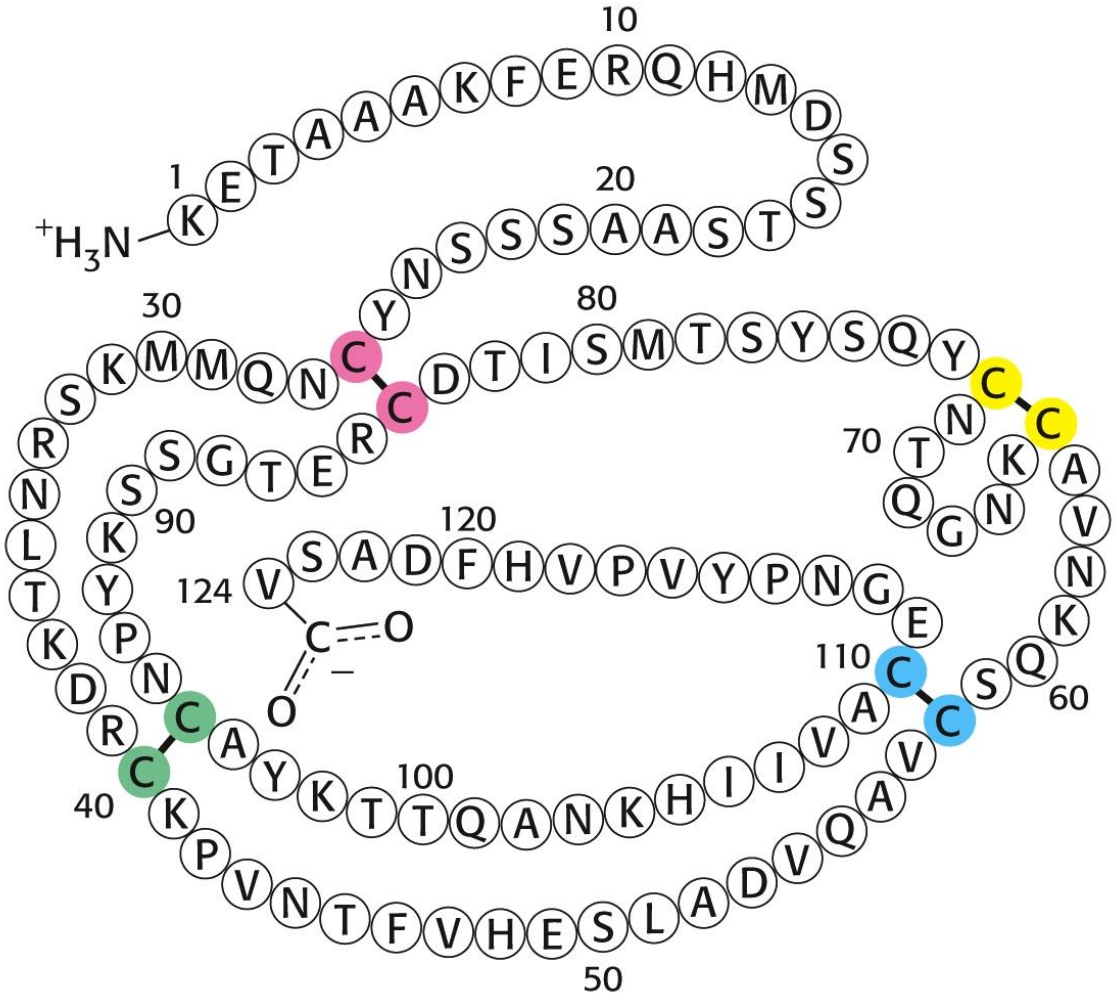
Protein denaturation and folding

- **Denaturation** – loss of 3D structure
- - partially folded state
- **Physical condition** – T, heat, mechanical stress (foam – egg white)
- **Chemical condition** :
- extreme pH (strong acids and bases),
- organic solvents (alcohols, acetone),
- certain solutes – urea, guanidine chloride,
- detergents
- Salt concentration, heavy metals (Pb, Hg-anemia)
- Mild treatment- no covalent bonds
- Need not be equivalent

Amino acid sequence determines 3D-structure

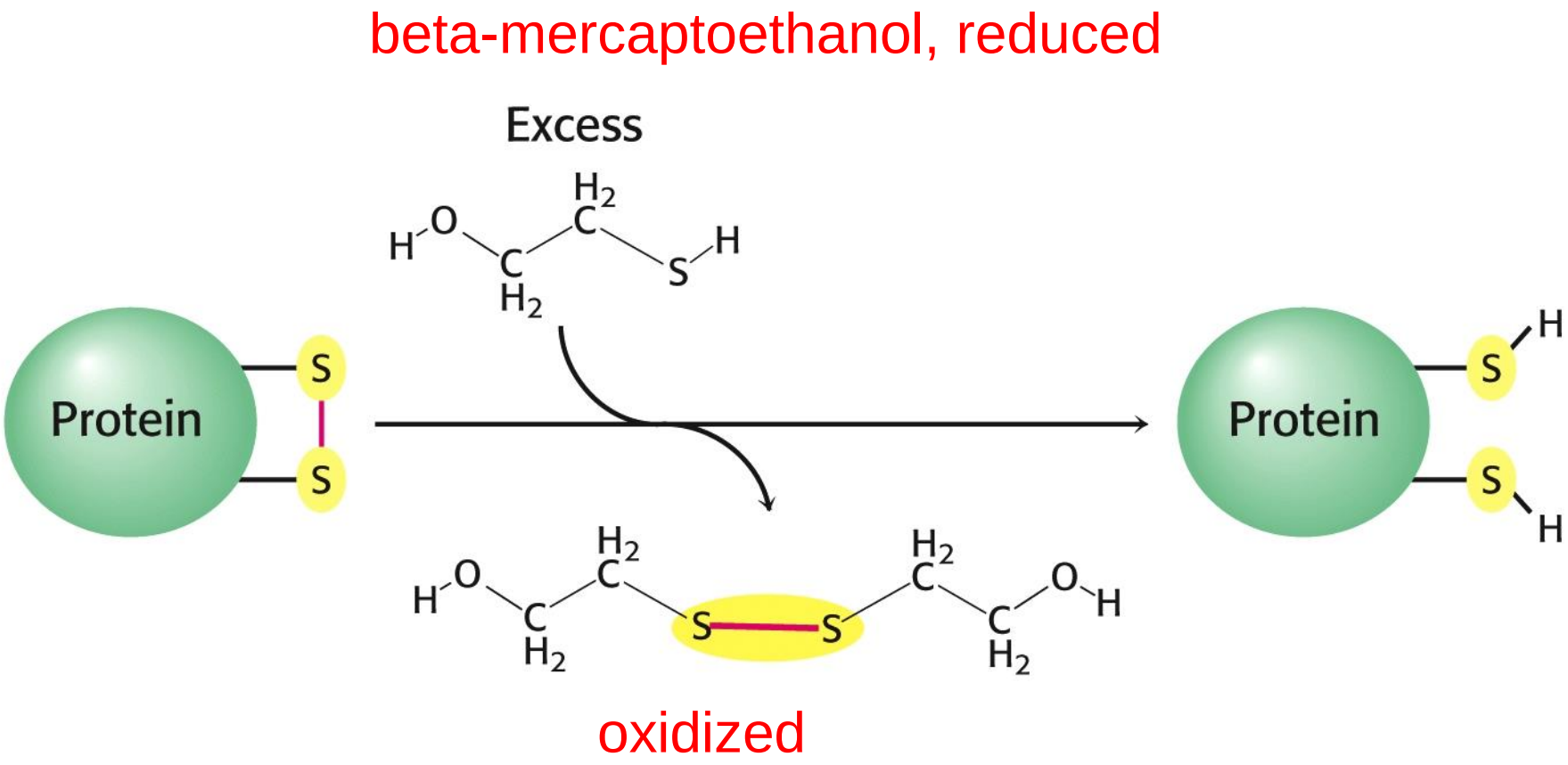
Bovine ribonuclease, 1950, C. Anfinsen work

Denature & renature

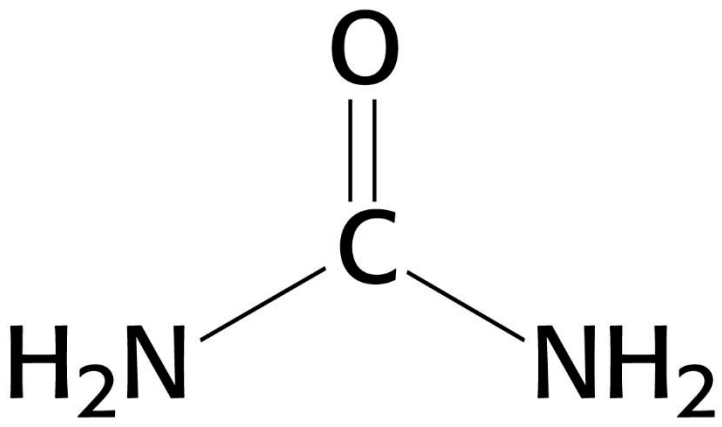


4 disulfide bonds
124 amino acids

Reducing disulfied bonds

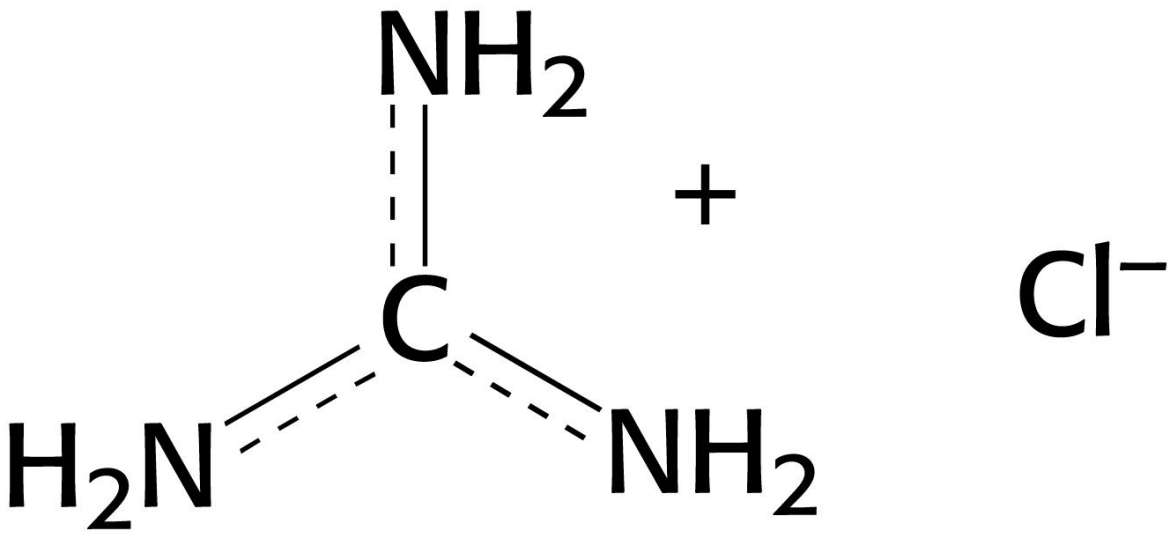


Denaturing agent, urea



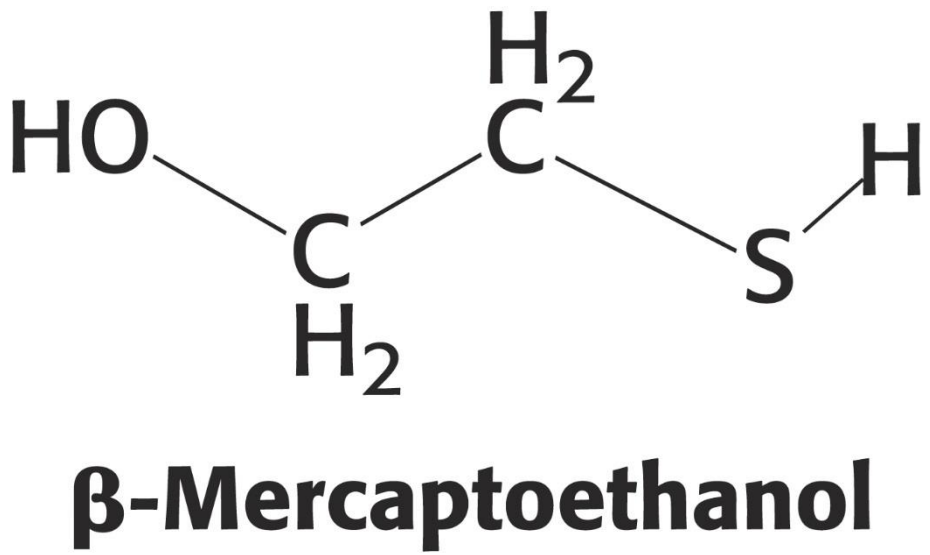
Urea

Denaturing agent, guanidinium chloride

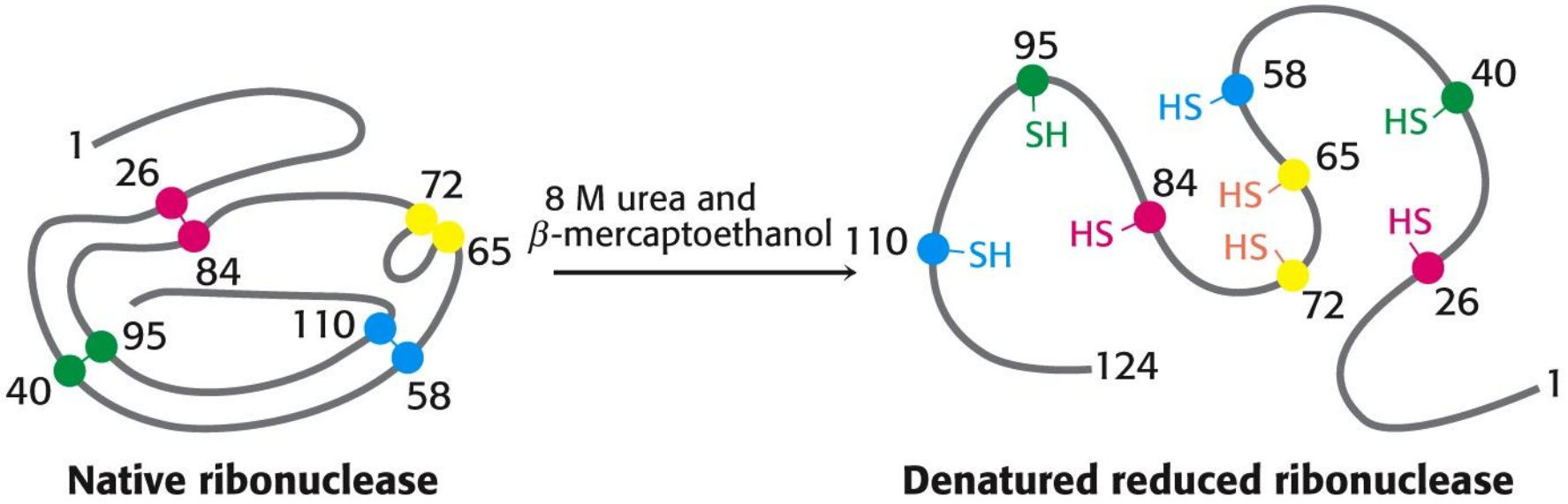


Guanidinium chloride

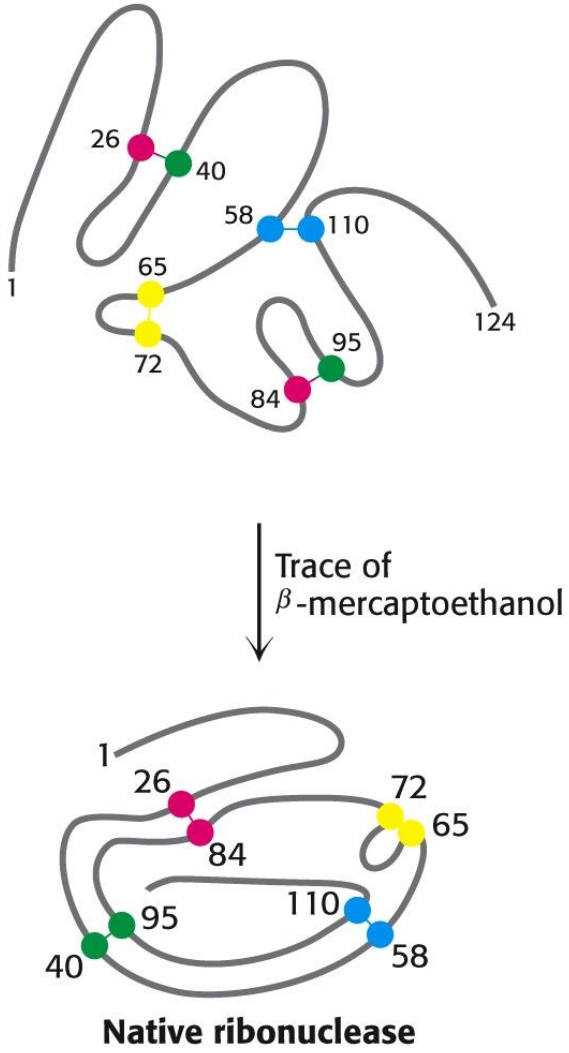
Denaturing agent, beta mercaptoethanol



Ribonuclease: reduction & denaturation



Reestablishing correct disulfide pairing

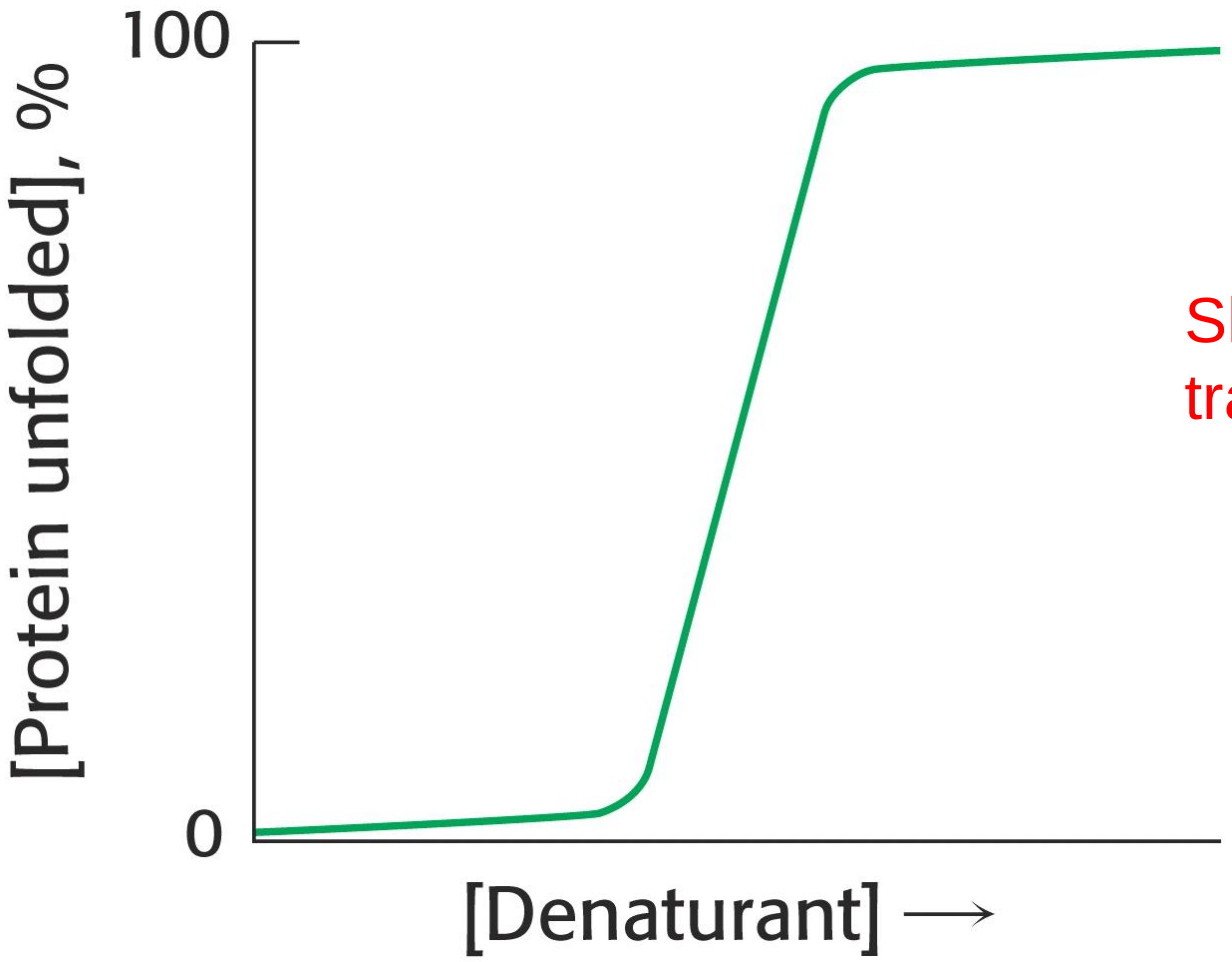


Scrambled conformation, from oxidation in 8 M urea, only 1% activity, (105 possible pairings)

Urea removed before trace of mercaptoethanol added, full activity restored.

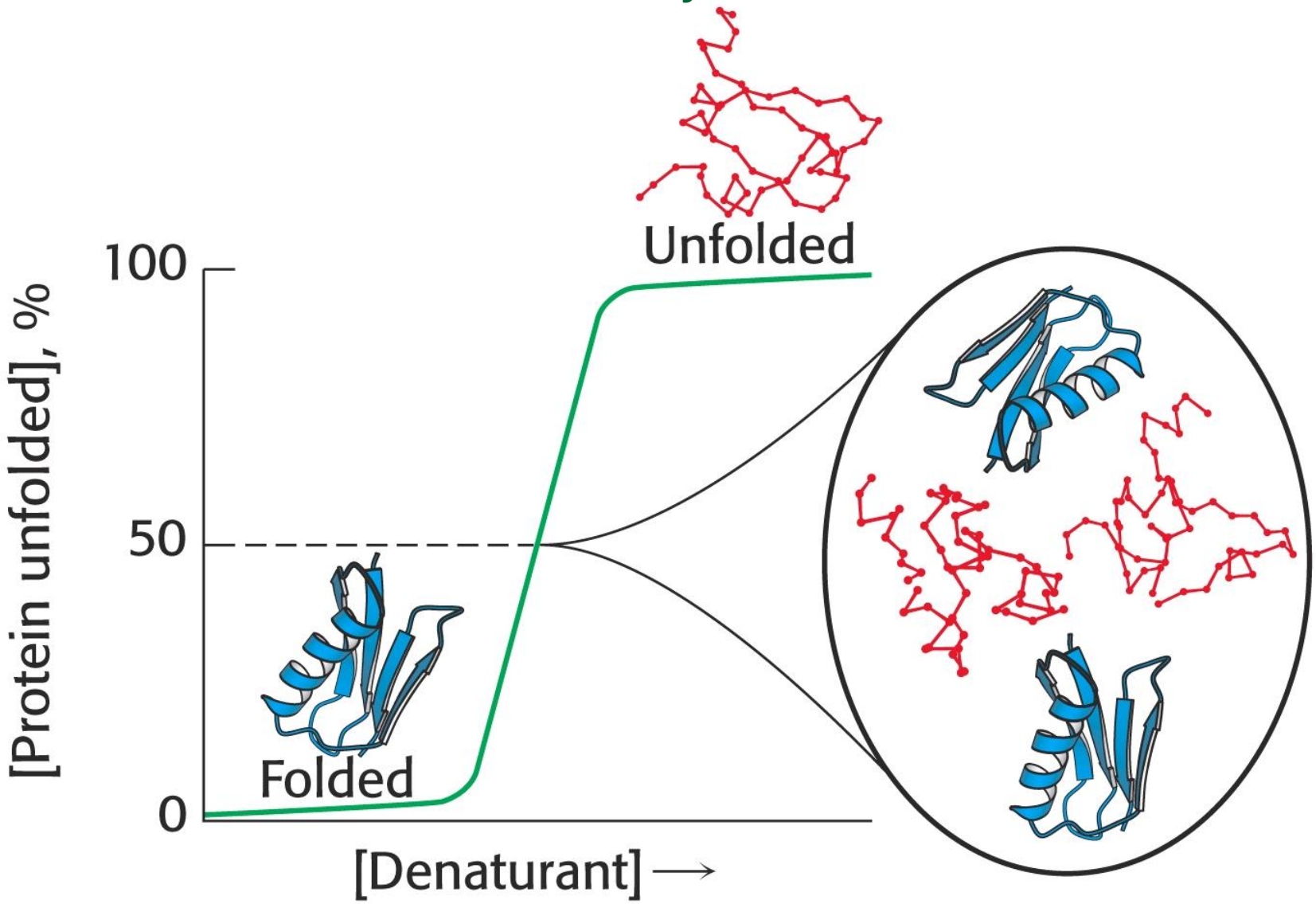
Process driven by decrease in free energy

Transition: folded to unfolded



Sharp transition

50 / 50 mixture at halfway



Patological consequences of perturbation of protein conformation

- Prions
- Alzheimer disease
- Beta-Thalassemias

- A **prion** in the Scrapie form (PrPSc) is an infectious agent composed of protein in a misfolded form.

- A **prion** is an infectious agent that is composed primarily of protein. To date, all such agents that have been discovered propagate by transmitting a mis-folded protein state;
- the protein itself does not self-replicate and the process is dependent on the presence of the polypeptide in the host organism.
- The mis-folded form of the prion protein has been implicated in a number of diseases in a variety of mammals, including bovine spongiform encephalopathy (BSE, also known as "mad cow disease") in cattle and Creutzfeldt-Jakob disease (CJD) in humans.
- All known prion diseases affect the structure of the brain or other neural tissue, and all are currently untreatable and are always fatal. In general usage, **prion** refers to the theoretical unit of infection. In scientific notation, PrPC refers to the endogenous form of prion protein (PrP), which is found in a multitude of tissues, while PrPSc refers to the misfolded form of PrP, that is responsible for the formation of amyloid plaques and neurodegeneration.

Post-translational modification of proteins

•After synthesis of a protein is often attached to the molecule non-protein component.

Some of the OH group of the side chain (Ser, Thr, ...) is phosphorylated.

Despite nitrogen (Asn) or oxygen (Ser, Thr) is attached oligosaccharide (glycoproteins)

It is connected acyl fatty acids (lipoprotein) or isotrenová group (anchoring protein in the membrane)

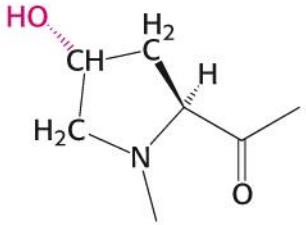
It is attached prosthetic group required for catalytic function (an organic molecule, metal ion ...)

Partial proteolysis (insulin, zymogens (pepsinogen, chymotrypsinogen ..., viral proteins)

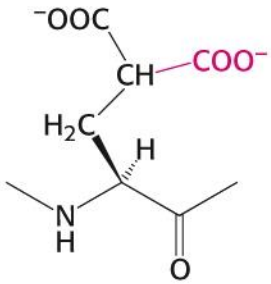
Zajímavé linky:
3D modely proteinů
http://www.ncbi.nlm.nih.gov/structure
např. enzym aldolasa
http://www.ncbi.nlm.nih.gov/Structure/mmdb/mmdbsrv.cgi?uid=69559
Nutné je mít nainstalován plug-in modul Cn3D:
http://www.ncbi.nlm.nih.gov/Structure/CN3D/cn3d.shtml
Aplikace umožňující porovnávat sekvence proteinů
http://www.jalview.org/examples/applets.html
http://www.wehi.edu.au/education/wehitv/

Finishing touches: covalent modifications

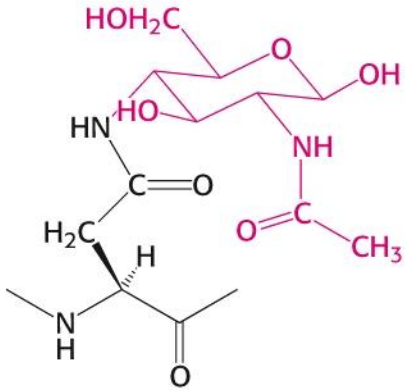
Proteins covalently modified to augment function



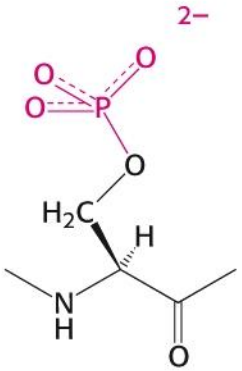
Hydroxyproline



γ-Carboxyglutamate



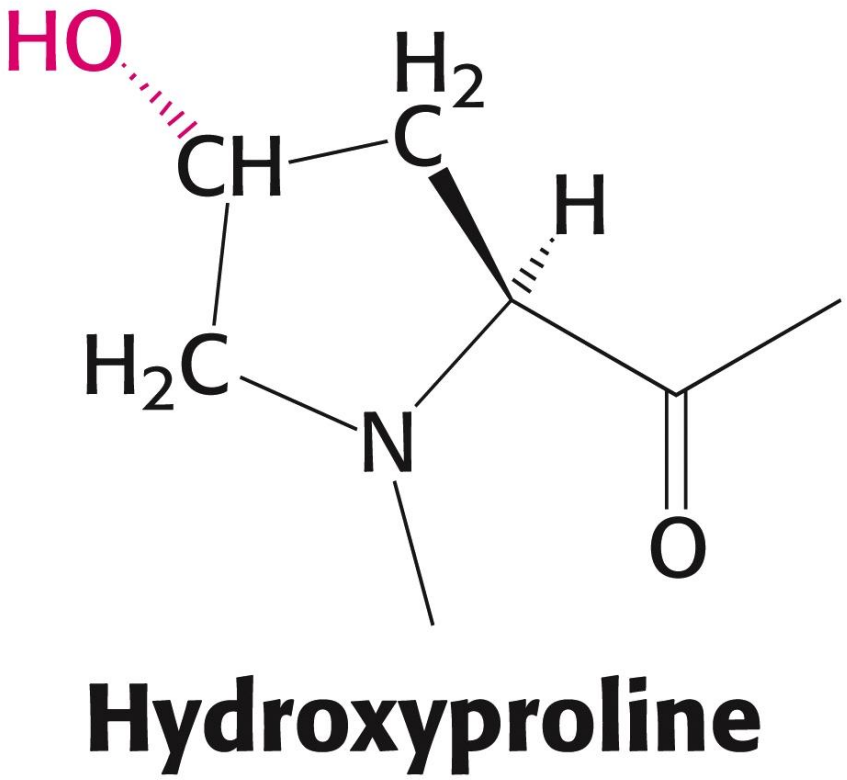
Carbohydrate–asparagine adduct



Phosphoserine

Hydroxyproline

Hydroxylation of proline residues in polypeptide

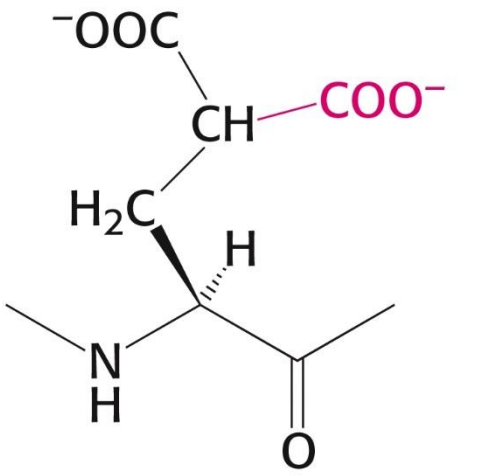


Stabilizes fibers of collagen in bone & connective tissue.

Scurvy: vitamin C deficiency, leads to insufficient hydroxylation

gamma-Carboxyglutamate

Carboxylation of glutamate residues in polypeptides

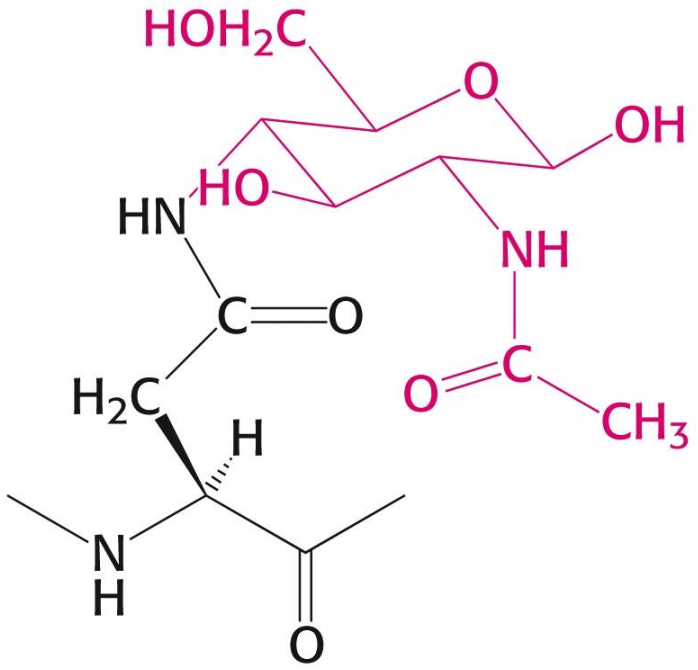


γ-Carboxyglutamate

Carboxylation of glutamate in prothrombin (clotting protein)

Vitamin K deficiency leads to insufficient carboxylation, and hemorrhage

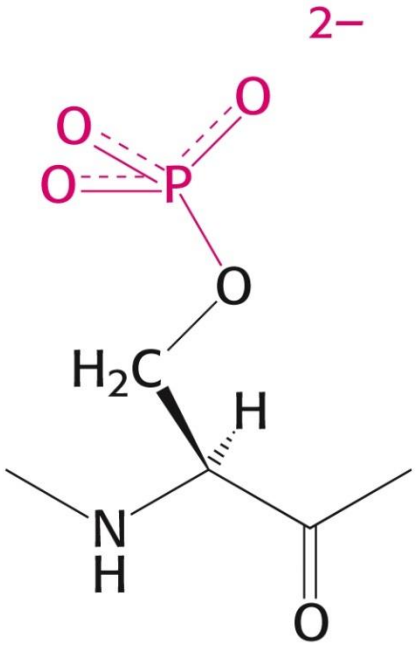
Carbohydrate to asparagine residues



Carbohydrate–asparagine adduct

Addition of sugars makes proteins more hydrophilic, and more interactive with other proteins

Phosphorylation of serine, threonine, & tyrosine



Phosphoserine

Triggered by hormones, and growth factors.

Phosphorylation is reversible, thus acts as, reversible switches for regulating cellular processes

Assisted Folding

- Molecular chaperones- correct folding
- HSP70

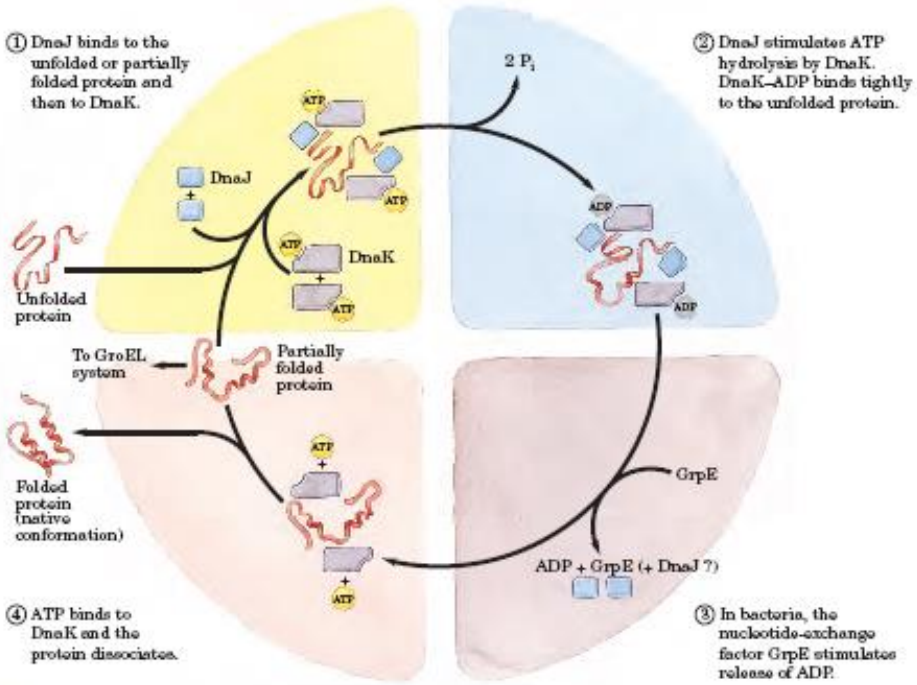
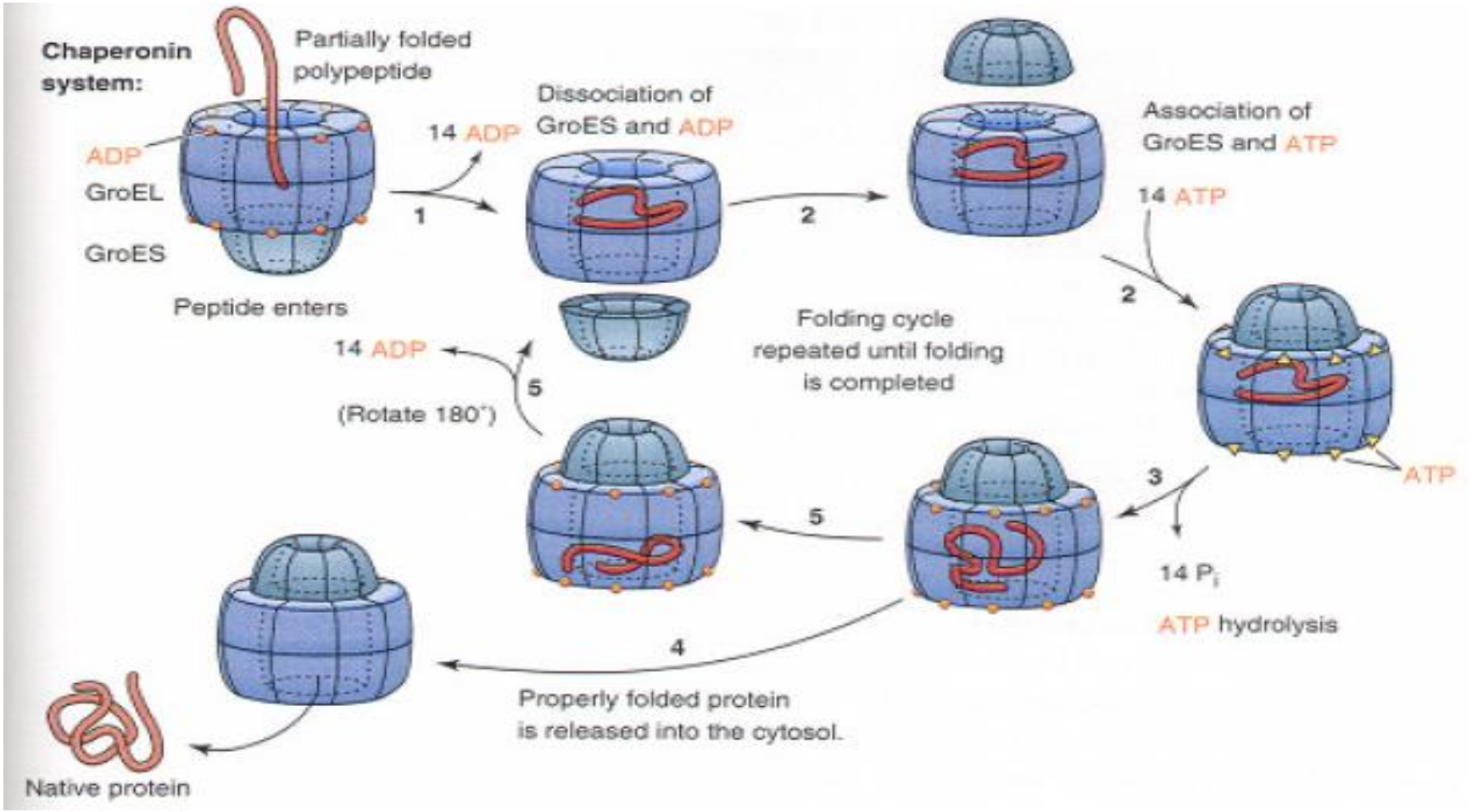


FIGURE 4–30 Chaperones in protein folding. The cyclic pathway by which chaperones bind and release polypeptides is illustrated for the *E. coli* chaperone proteins DnaK and DnaJ, homologs of the eukaryotic chaperones Hsp70 and Hsp40. The chaperones do not actively promote the folding of the substrate protein, but instead prevent aggregation of unfolded peptides. For a population of polypeptides, some fraction of the polypeptides released at the end of the cycle are in the native conformation. The remainder are rebound by DnaK or are diverted to the chaperonin system (GroEL; see Fig. 4–31). In bacteria, a protein called GrpE interacts transiently with DnaK late in the cycle (step ③), promoting dissociation of ADP and possibly DnaJ. No eukaryotic analog of GrpE is known.

Chaperony


Chaperonin system:
Partially folded polypeptide
ADP
GroEL
GroES
Peptide enters
14 ADP
1
Dissociation of GroES and ADP
2
Association of GroES and ATP
14 ATP
2
ATP
3
14 P_i
ATP hydrolysis
5
14 ADP
(Rotate 180°)
5
Folding cycle repeated until folding is completed
4
Properly folded protein is released into the cytosol.
Native protein