Chapter 4 Proteins

Except for water, proteins are the most abundant substance in most cells. 10-20% by weight. Humans manufacture at least 22,000 different proteins.

Function	Example
Catalysis	Enzymes
Transport	Hemoglobin
Structure	Collagen
Contractile	Actin
Nutrient	Ovalbumin
Defense	Immunoglobins
Regulatory	Insulin

Molecule	#AA	Mass	#Chains
H_2O	-	18	-
Insulin	51	5,700	2
Lysozyme	129	13,900	1
Hemoglobin	574	64,500	4
Glutamate Dehydrogenase	8300	1 million	40
Titin	26,926	2,993,000	1

Some proteins are simple, **non-conjugated**.

Others require non-AA **cofactors** or **prosthetic groups** for full activity. Cofactors may be inorganic *e.g.* metal, phosphate or organic *e.g.* sugar, lipid, heme, flavin.

They may be covalently or non-covalently attached to the protein.

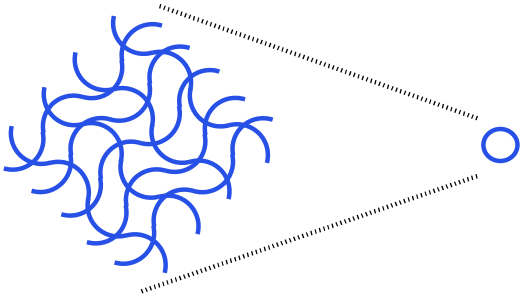
Enzyme cofactors are called **coenzymes**.

Protein Purification: Proteins may be purified on the basis of differences in size, charge, solubility, affinity for materials.

Ion exchange chromatography of AA was discussed earlier. The same principles can be used to separate proteins with different pI's.

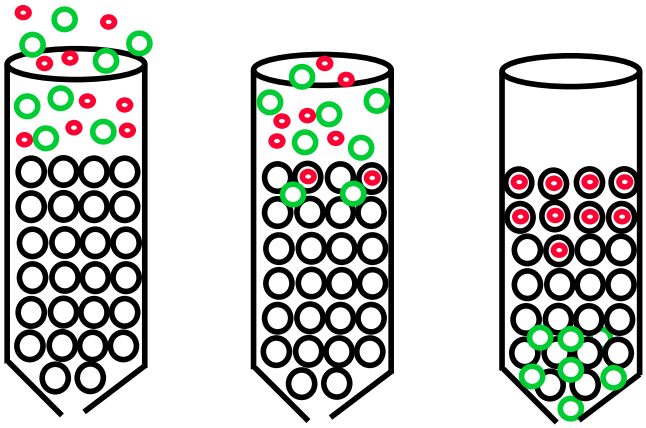
Size-Exclusion or Gel Sieving Chromatography

Small beads of polymerized glucose, agarose, or acrylamide are manufactured with different sizes of pore depending on the degree of cross-linking of the polymer.



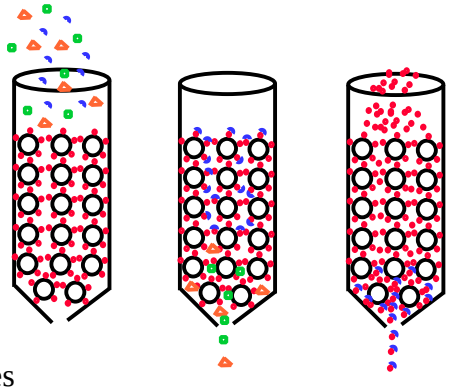
The beads ○ are packed into a cylinder and a mixture of proteins is applied.

Big proteins don't enter the porous beads and run **quickly** through the column. **Small proteins** enter and exit the beads and elute more **slowly**.



Affinity Chromatography A **Ligand** is any molecule that is bound specifically by a protein. *e.g.* **ATP** binds to *hexokinase*.

1. Ligands ⁘ are attached to polymer beads ○ and packed into a column.
2. A mixture of proteins ▲● is applied to the beads.
3. Hexokinase ◗ binds, all others ▲● are washed out.
4. ATP ⁘ is added and competes for the binding site causing pure protein ◗ to unbind and elute from the column. ◗ + ● ⇌ ◗●

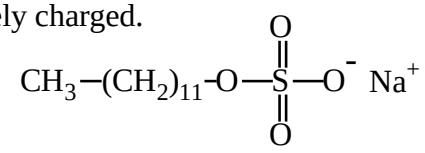


SDS-PAGE

It separates proteins by electrophoresis and then estimates their masses.

1. The gel is a cross-linked **p**oly**a**crylamide **g**el molecular sieve.

2. The detergent **s**odium **d**odecyl**s**ulphate binds to the proteins and makes them highly negatively charged.

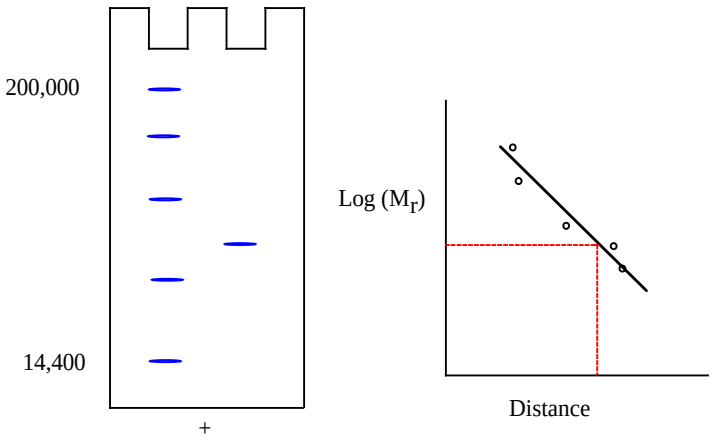


About 1 SDS binds per 2 AA.

3. SDS-coated proteins move through the gel by **e**lectrophoresis when an electric potential is applied. The **small** proteins move **quickly** and easily through the pores. The **big proteins** move **slowly**.

4. After electrophoresis, the proteins are visualized by staining with **Coomassie Blue** or **Silver**.

5. The migration distance is proportional to the $\log_{10}[M_r]$. The mass of an unknown protein can be determined by interpolation, using the migration distances of proteins of known molecular weight:



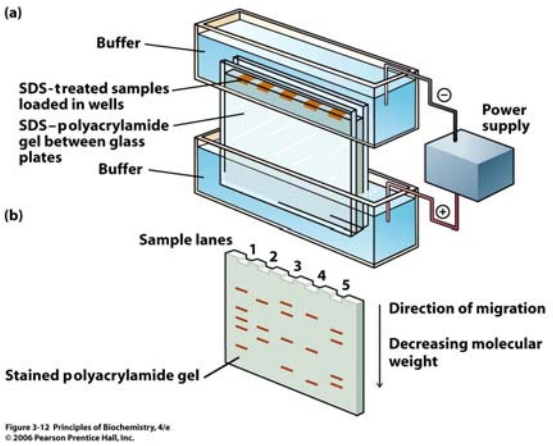


Figure 3-12 Principles of Biochemistry, 4/e
© 2006 Pearson Prentice Hall, Inc.

Protein Structure

1º **Primary**: AA sequence.

2º **Secondary**: regular, repeating backbone conformation.

3º **Tertiary**: 3-dimensional polypeptide conformation.

4º **Quaternary:** association of polypeptides.

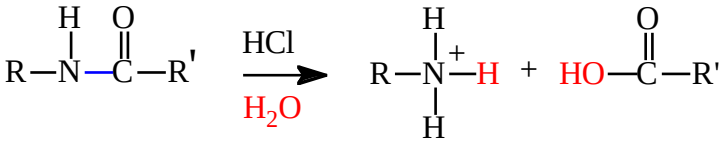
Primary structure determination

Amino Acid Analysis determines the AA composition of a protein, **not** the order of the AA.

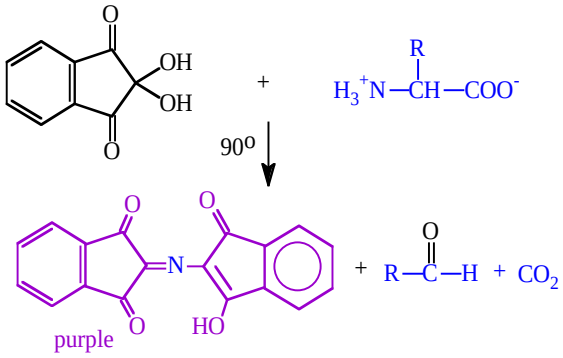
1. Hydrolyze all the peptide bonds in a pure protein using 6M HCl @ 110°C for 24 h.

This is a bond **Cleavage** by **hydrolysis**.

2. Separate the AA by ion exchange chromatography.



3. Quantify the AA by reacting with **Ninhydrin** to produce a purple colour.



4. Measure the absorption of light at 540 nm.

5. Use **Beer's Law** to determine the concentration of each AA.

$A_{540} = \varepsilon cl$

l = path length of light absorption: cm

ε = extinction coefficient: $l \cdot mole^{-1} \cdot cm^{-1}$

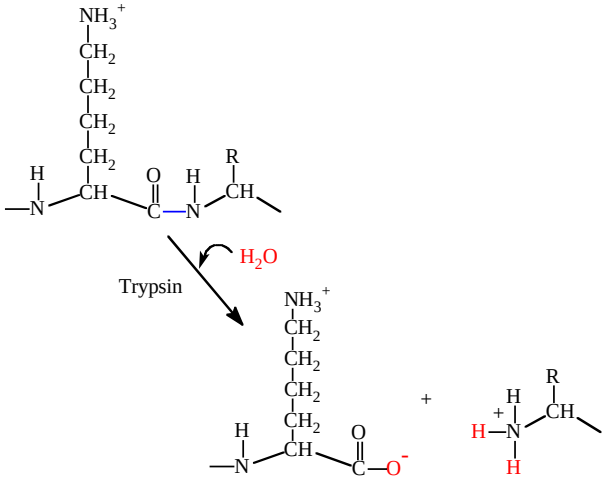
c = concentration: M

Protein Sequencing

1. Cleave the polypeptide into shorter peptides using **proteolytic** enzymes or chemicals, and separate them.
e.g. Trypsin cleaves at the C-end of Lys or Arg.

Chymotrypsin cleaves at the C-end of Tyr, Trp, Phe.

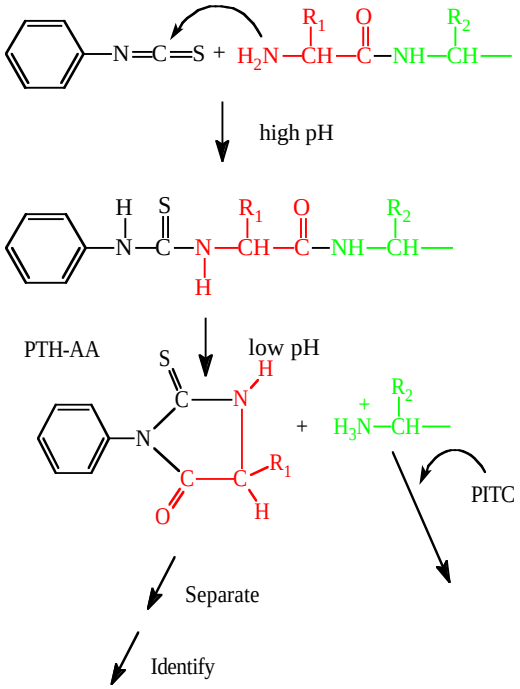
CNBr is a chemical that cleaves at the C-end of Met.



2. **Edman Degradation:**

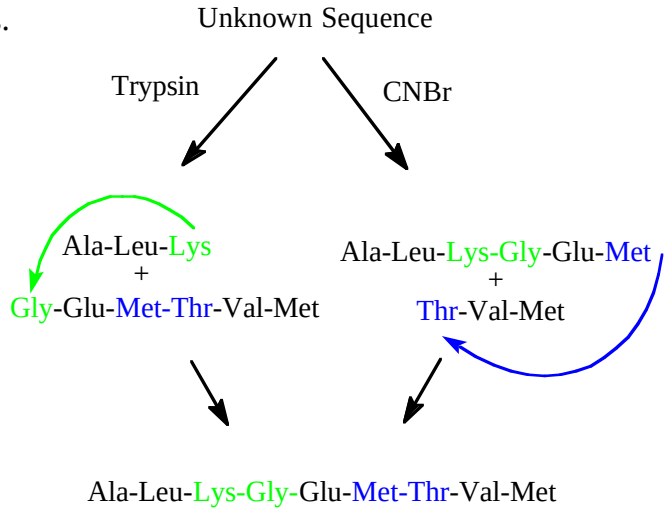
React the N-end of the peptide with PITC which weakens the first peptide bond.

A PTH-AA is released and can be detected by absorption at 254 nm.



3. This is repeated up to 50X, each AA released and detected 1 at a time.

4. The protein sequence is deduced by sequencing overlapping peptides.



AA sequences can also be determined by **Mass Spectrometry.**

They can also be deduced from the sequence of the **gene**.

Notes on Protein Sequences:

1. The linear polypeptide contains the information to direct its folding into a 3D conformation that determines the activity.

Inactive Active

2. AA sequences sometimes contain signal sequences that determine cellular location or export, chemical modifications, protein 1/2 life, *etc*.

3. Incorrect AA incorporation can lead to a loss or alteration of protein activity or folding and thus disease.

e.g. A single Glu → Val **mutation** in the hemoglobin β-chain results in Sickle Cell Anemia.

Of 1400 Human Genetic Diseases 1/3 are due to a single AA change.

4. About 30% of human proteins are **polymorphic**. *i.e.* in the human population slightly different sequences may be present but there is little or no difference in activity.

5. Proteins that perform the same function in 2 different species usually have similar sequences.

e.g. pig and cow insulin have been used to treat human diabetics.

6. The greater the **phylogenetic** (evolutionary history) difference between 2 species, the greater the # of AA differences in their proteins.

Sequences of cytochrome c from a wide diversity of organisms.

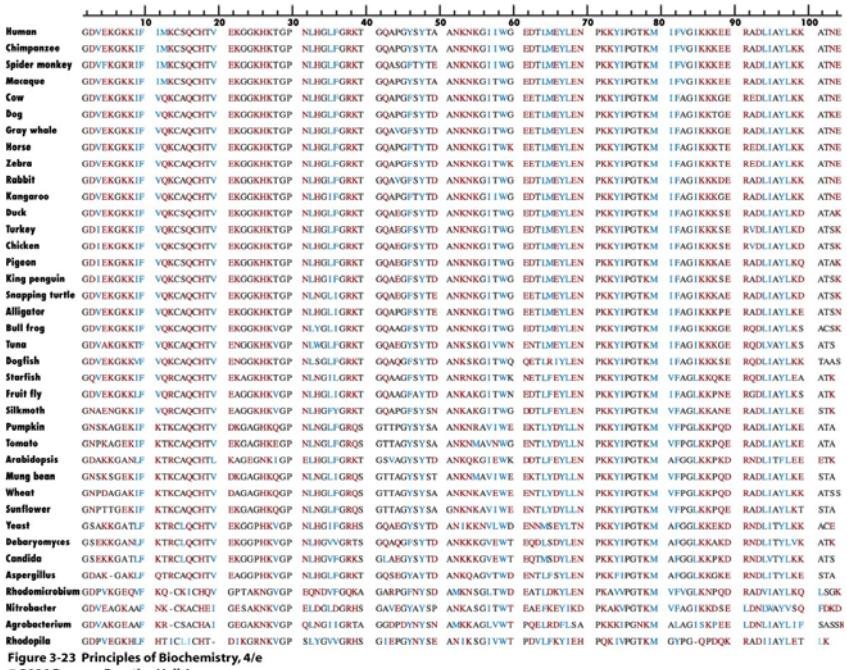


Figure 3-23 Principles of Biochemistry, 4/e
© 2006 Pearson Prentice Hall, Inc.

Hydrophobic residues are **blue** and hydrophilic residues are **red.**

Note that humans and chimpanzees have identical sequences and spider monkeys are only different at 2 positions but human and yeast are very different.

Evolutionary trees or **cladograms** based on protein sequences generally agree with trees constructed on the basis of anatomy and physiology.

Here is a tree based on the sequence of cytochrome c.

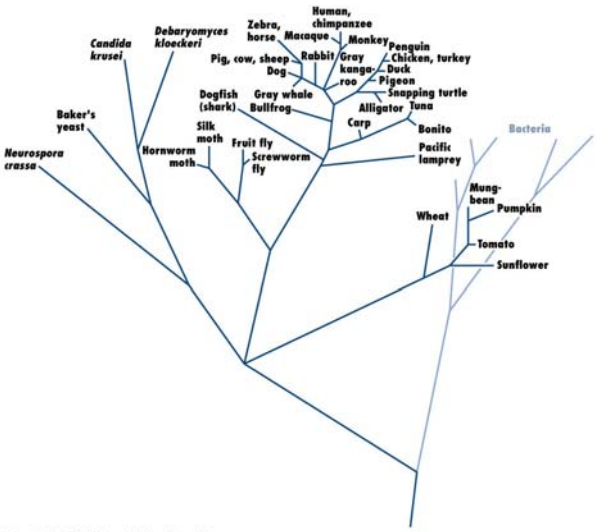


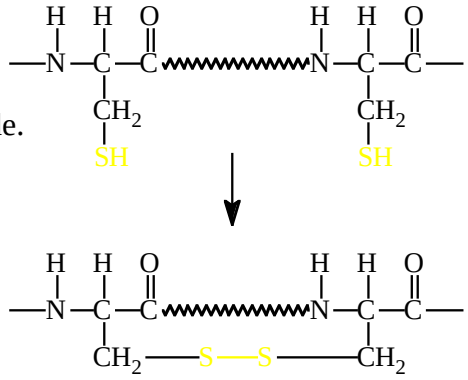
Figure 3-24 Principles of Biochemistry, 4/e
© 2006 Pearson Prentice Hall, Inc.

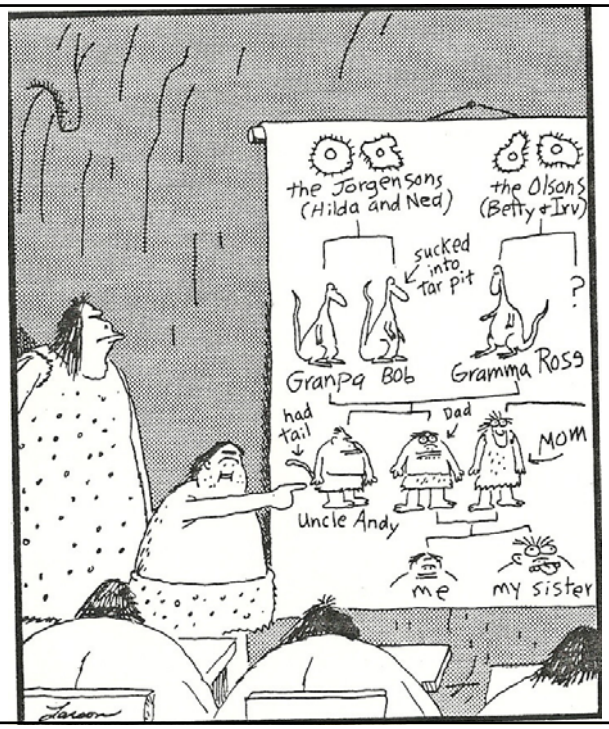
Disulphide Bond

Oxidation of two cysteines gives cystine.

The bonds may be intrapolypeptide or interpolypeptide.

They can hold two distant parts of the polypeptide close together.





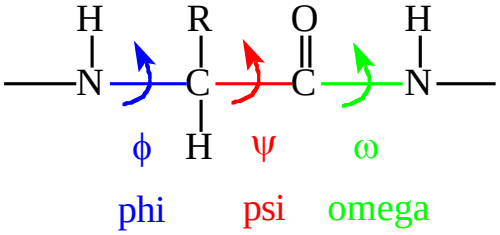
Secondary Structures of Proteins

Biologically active proteins are folded into a well-defined 3D conformation.

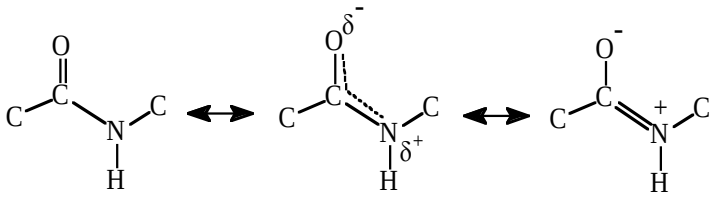
Unfolding / denaturation of a protein eliminates activity.

The conformation can be described by the torsion angles about the single bonds.

Backbone torsion angles:



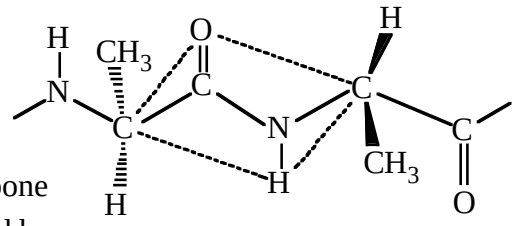
ω is almost always = 180°
because the peptide has partial double bond
character due to delocalization of the N lone pair electrons.



Resonance Structures

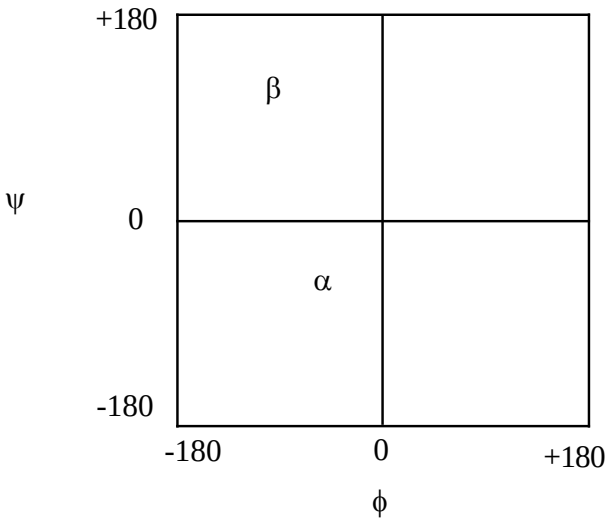
So the peptide bond is planar, trans, and rigid.

Notice that all 6 backbone atoms lie in a plane.



Since ω is fixed, the backbone conformation is determined by the ϕ and ψ angles.

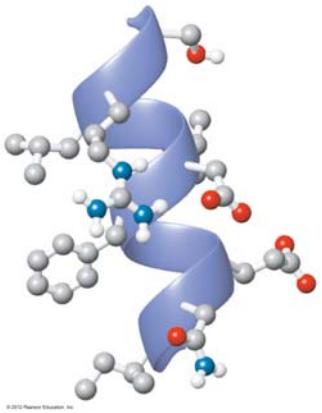
Many combinations of ϕ and ψ are not allowed because of **steric interactions.** This is indicated in a **Ramachandran Diagram**.



1. **α-Helix**

$\phi = \psi \quad -60^{o}$ right-hand twist most

$\phi = \psi \quad +60^{o}$ left-hand twist few



The C=O point to the C-terminus. The N-H point to the N-terminus.

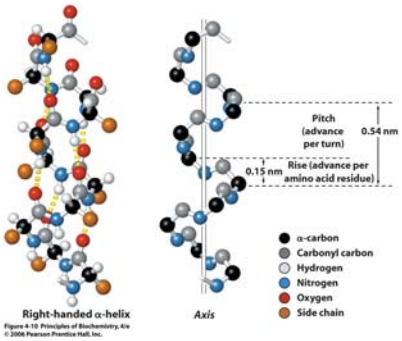
Except at the ends of the helix, all the C=O of AA_i are H-bonded to the H-N of AA_{i+4}. See the dashed yellow line. This gives the helix stability.

Each turn of helix contains 3.6 AA.

The height of 1 turn is 0.54 nm.

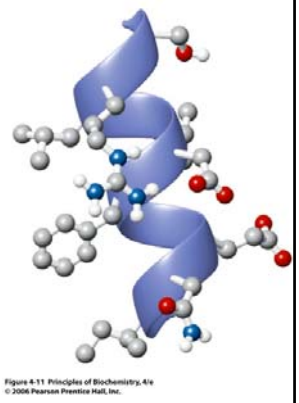
The rise per residue is 0.15 nm.

Side chains (**orange**) protrude from the sides of the helix.

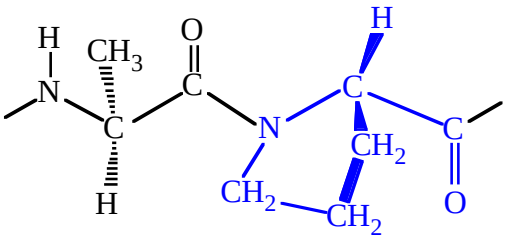


The helical conformation is very **compact**. A *D*-AA will not fit into a right-handed helix because of steric interactions.

Only certain combinations of AA form helices.
Ala, Met, Glu favour helix.



Pro puts bends in helices:



Its side-chain clashes with the preceding C=O.
It has no NH for H-bonding.

Clusters of **Lys / Arg** and **Glu / Asp** are not helical due to charge repulsion in the compact helix.

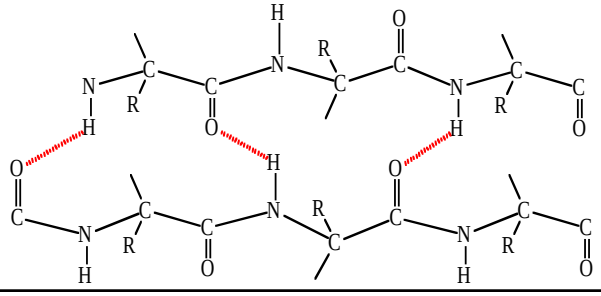
Clusters of bulky AA are not helical for steric reasons: Ile, Trp ...
Glycine adds flexibility because of its small side-chain.

2. **β-Strand** $\phi = -120^{o}$ $\psi = +120^{o}$

The polypeptide is in an **extended** conformation.

β-strands often associate by H-bonding to an adjacent β-strand, either (a) parallel or (b) antiparallel, forming a β-pleated sheet.

(a)



Side-chains are alternately above and below the plane of the sheet.

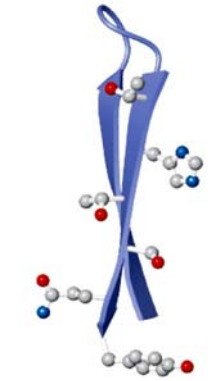


Figure 4-16 Principles of Biochemistry, 4/e
© 2006 Pearson Prentice Hall, Inc.

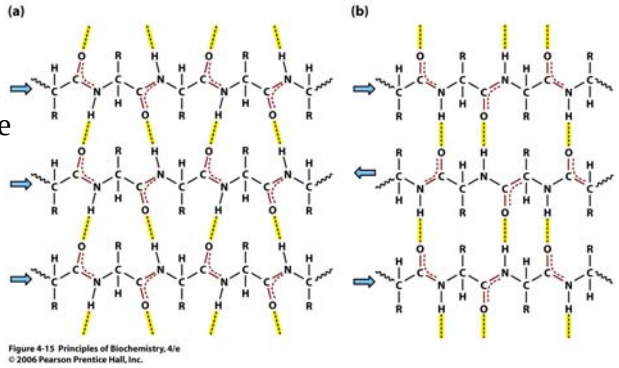


Figure 4-15 Principles of Biochemistry, 4/e
© 2006 Pearson Prentice Hall, Inc.

Val, Ile, Phe, Tyr are found more frequently than Asp, Pro, Glu, Gly in β-sheets.

3. **Turns**: β-turns are 4 AA connected to 2 antiparallel β-strands.

They produce sharp 180^{o} turns in the polypeptide and often are found at the protein surface.

Proteins also contain **irregular** structure such as loops, important for binding metals and substrates.

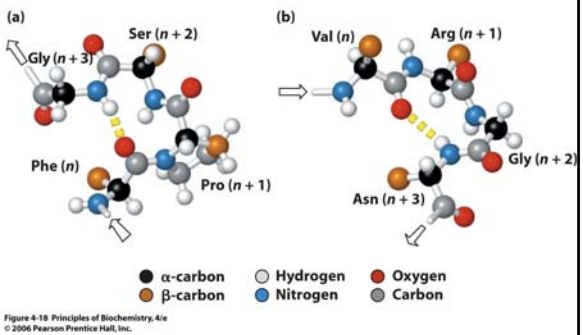


Figure 4-18 Principles of Biochemistry, 4/e
© 2006 Pearson Prentice Hall, Inc.

In addition to H-bonding, van der Waal's, ionic interactions, disulphide bonds and metal binding can stabilize secondary structure.

Tertiary Structure
The arrangement of 2º structures to form the 3D structure. It is dominated by long-range interactions and side-chain interactions.

X-ray Diffraction is used to determine the 3D structures of **crystalline** protein.

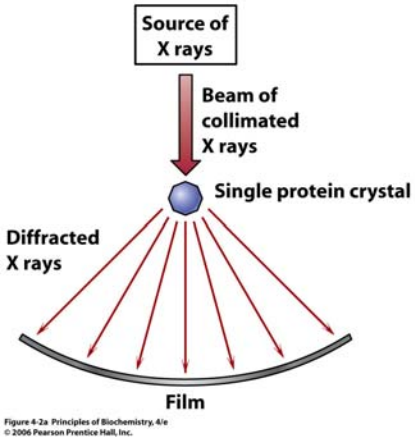
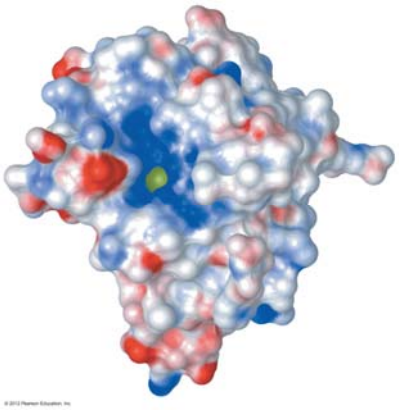
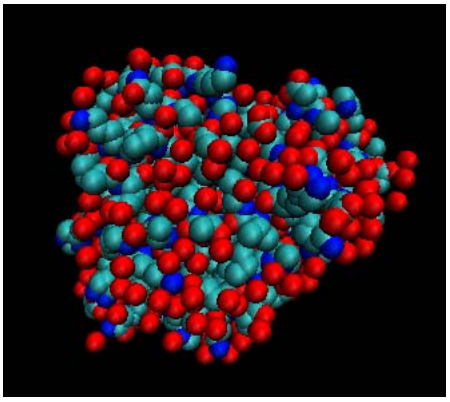


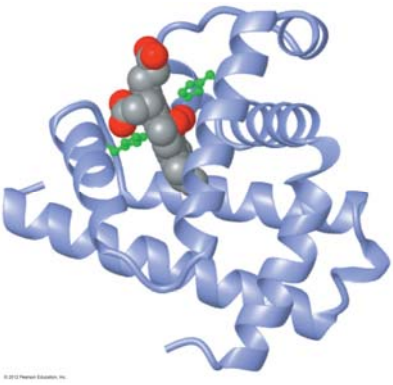
Figure 4-2a Principles of Biochemistry, 4/e
© 2006 Pearson Prentice Hall, Inc.

Here is a picture of the enzyme *Superoxide Dismutase* determined by X-ray diffraction. The positive and negative charges on the surface of the protein are indicated in blue and red. The copper atom in the active site is green.
The substrate O_2^- is negatively charged and attracted to the positively charged binding site.



Here are some pictures of myoglobin determined by X-ray diffraction.





PDB ID: 1U7R

Nuclear Magnetic Resonance (NMR) Spectroscopy can determine the structures of small proteins (M_r< 50,000) dissolved in water.

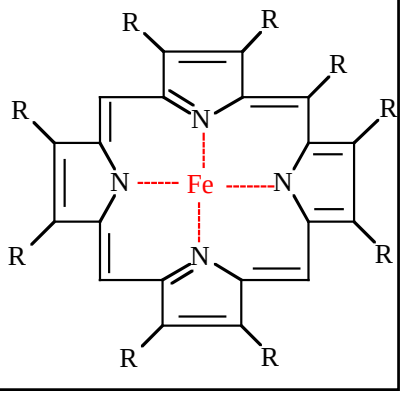
Some examples of Protein Structures:

Myoglobin functions in O_2 storage in muscle.

153 AA + **heme** cofactor

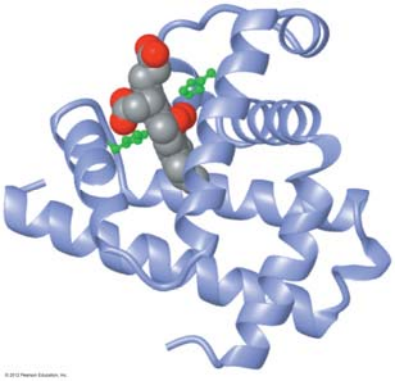
4 N of heme bind Fe^{2+} in the ring plane. 1N of a protein His binds Fe^{2+} from below the plane.

O_2 binds from above to the Fe^{2+}.



Myoglobin has 0 β-strands and 8 α-helices, 7 - 26 AA long, linked by short, irregular loops.

The helices pack tightly to form a compact **globular** structure with **no holes**. The non-polar heme is almost completely buried.



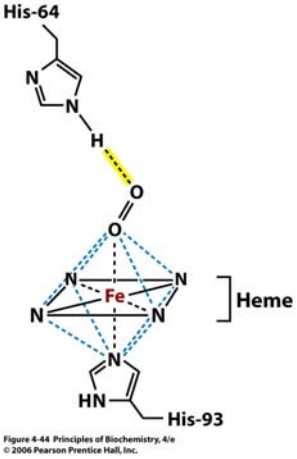
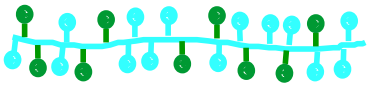
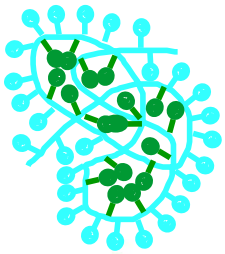


Figure 4-44 Principles of Biochemistry, 4/e
© 2006 Pearson Prentice Hall, Inc.

Hydrophobic side-chains point into the interior and are **"buried"**.

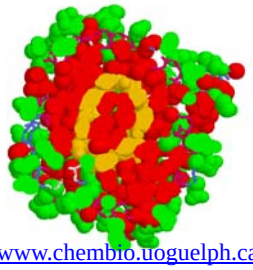
Hydrophilic side-chains are on the outside in contact with water.





This is common to most globular proteins.

Here is a picture of a cross-section of the enzyme **triose-phosphate isomerase**. The yellow and red atoms on the inside represent **hydrophobic AA**; the green atoms represent **hydrophilic AA**.



http://www.chembio.uoguelph.ca/educat/phy456/Protstr4.htm

Cytochrome c: It functions in electron transfer and also binds heme.

39% α, 0% β.

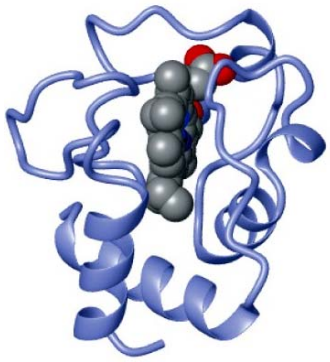


Figure 4-21 Principles of Biochemistry, 4/e
© 2006 Pearson Prentice Hall, Inc.

Below are the structures of (a, b) tuna, (c) rice, (d) yeast, (e) bacterial cytochromes c. Although the sequences are quite different, the proteins fold into essentially the same conformation.

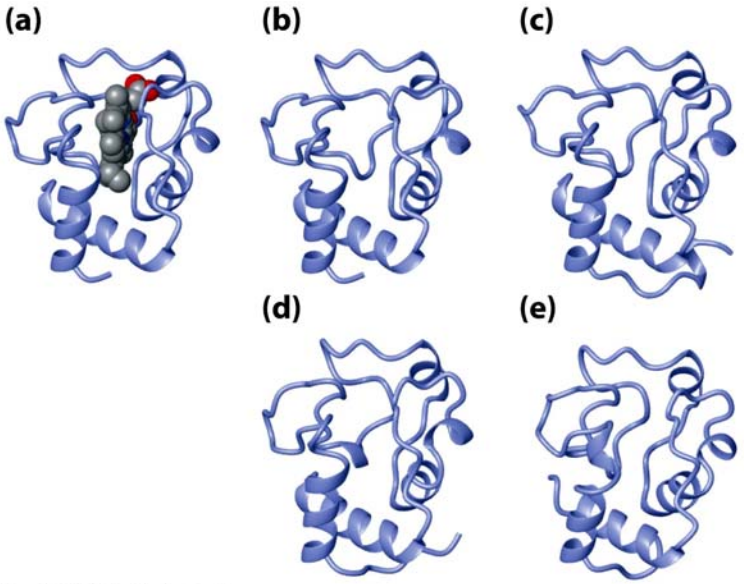


Figure 4-21 Principles of Biochemistry, 4/e
© 2006 Pearson Prentice Hall, Inc.

Immunoglobulin Fold: It is involved in vertebrate defense against foreign invaders.

47% β 5% α.

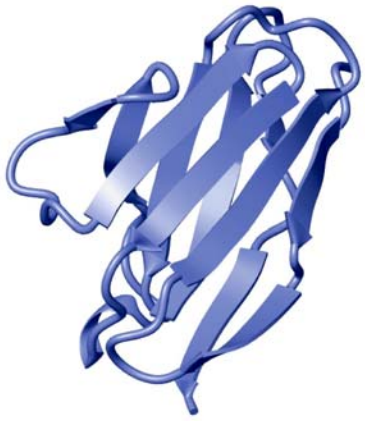
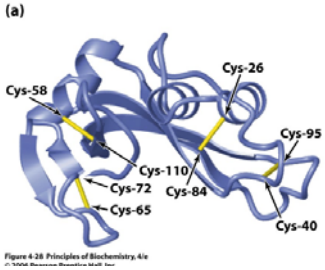


Figure 4-53 Principles of Biochemistry, 4/e
© 2006 Pearson Prentice Hall, Inc.

Ribonuclease is secreted by the pancreas into the small intestine where it hydrolyses ribonucleic acids (RNA).

26% α, 35% β.



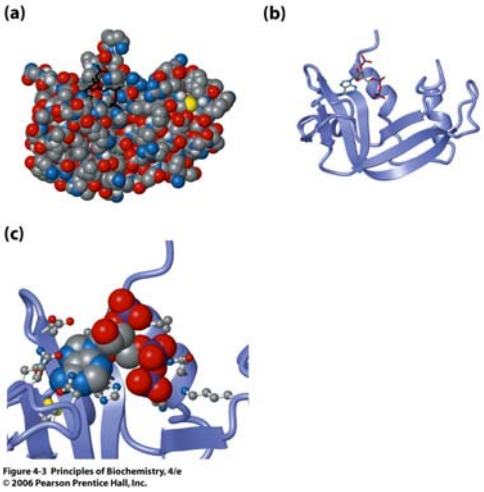


Figure 4-3 Principles of Biochemistry, 4/e
© 2006 Pearson Prentice Hall, Inc.

Many different protein sequences fold into a similar conformation giving rise to a small # of **structural families**. *e.g.* myoglobin & hemoglobin belong to the same family.

Different functions arise from subtle differences in conformation and / or critical AA.

The folded state is maintained by 4 **Weak Interactions**:

1. Hydrophobic (water entropy) effect.
2. H-bonding.
3. Electrostatic Interactions.
4. van der Waals Interactions

Plus

5. Disulphide bonds - strong covalent.
6. Binding of metals and other ligands.

Proteins are marginally stable:
ΔG unfolding ~ 20 - 70 kJ / mol = difference in free energy between folded and unfolded protein.

This means that breaking 4-20 H-bonds is enough to unfold a protein.

Proteins function *via* **changes in conformation**. They are dynamic, not static. Their atoms are constantly in motion.

Protein Denaturation:

The disruption of the 2^o, 3^o, and 4^o structure leading to loss of biological activity.

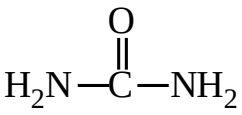
1. **Heat** breaks the weak interactions.

2. **Cold** – Water is more ordered at low temperatures so the entropy difference between the native and unfolded protein is lower at low T.

3. **pH** - Charge repulsion at pH extremes. *e.g.* curdling of milk

4. **Mechanical**: *e.g.* beat egg whites

5. 8M **Urea** disrupts H-bonding and hydrophobic interactions.



6. **Detergents** interact with the hydrophobic AA preventing their burial. *e.g.* SDS

7. **Organic** solvents, ethanol – see 6 above.

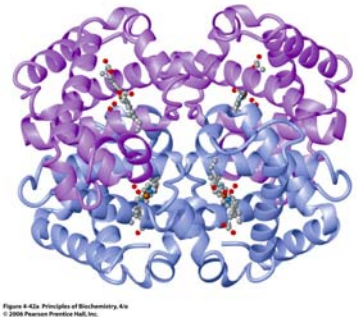
Some proteins will re-fold following denaturation. Others precipitate from solution.

In vivo, some proteins require **molecular chaperones** to prevent aggregation and increase the efficiency of folding.

Quaternary Structure

1. For biological activity many proteins require two or more polypeptide chains. *e.g.* insulin is a dimer (AB), hemoglobin (to the right) is a tetramer $\alpha_2\beta_2$.

The subunits usually associate by weak interactions.



Fibrous Proteins

Elongated molecules built from a single type of 2^o structure.

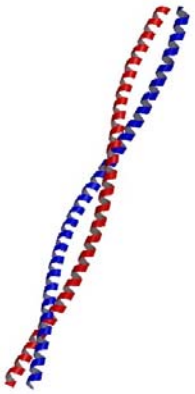
1. **α-Keratin** is found in hair, feathers, nails and is a tough insoluble material. It also forms the outer layer of human skin.

http://www.wellesley.edu/

2 right hand α-helical chains form a left hand **supercoiled** rope.

The coiled-coils organize into supramolecular structures called **protofilaments** consisting of many coiled coils.
Pairs of protofilaments combine to form **protofibrils.**

4 protofibrils form a **microfibril**.
Microfibrils combine to form **macrofibrils** that pass through and around hair cells.



PDB ID: 1D7M

Strength is added by cross-linking the proteins by **disulphide bonds**.
Human hair is 14% **Cys**.

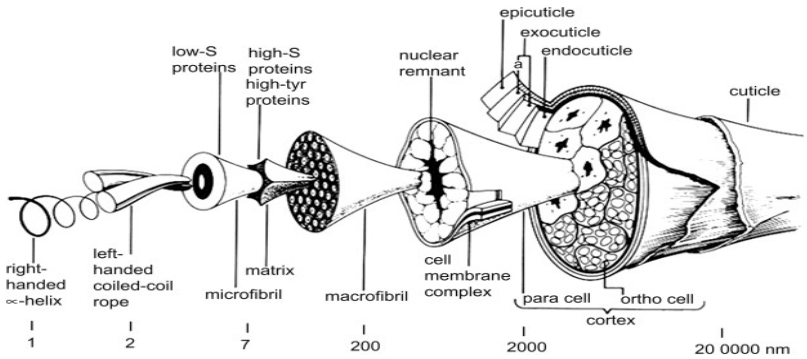


Fig.?1 Schematic of a wool fiber drawn by Bruce Fraser and Tom MacRae [Copyright CSIRO Australia 1996. Reproduced with permission from The Lennox Legacy (Rivett DE, Ward SW, Belkin LM, Ramshaw JAM and Wilshire JFK). Published by CSIRO PUBLISHING,...

Paulina Hill , Helen Brantley , Mark Van Dyke

Some properties of keratin biomaterials: Kerateines

Biomaterials Volume 31, Issue 4 2010 585 - 593

http://dx.doi.org/10.1016/j.biomaterials.2009.09.076

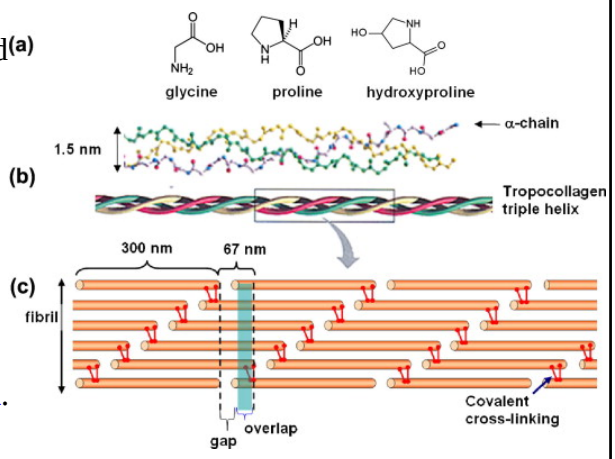
2. **Collagen** is found in tendons and bone matrix and has a high tensile strength.

(Gly-Xxx-Pro) is repeated over and over.

The conformation is **extended**.

$\phi = -60^{o}$ $\psi = +160^{o}$

Collagen helix:
3 AA / turn - **left-handed**.



Three staggered polypeptides form a supercoiled right-handed triple helix called tropocollagen.

Gly is in the middle of the 3 helices.

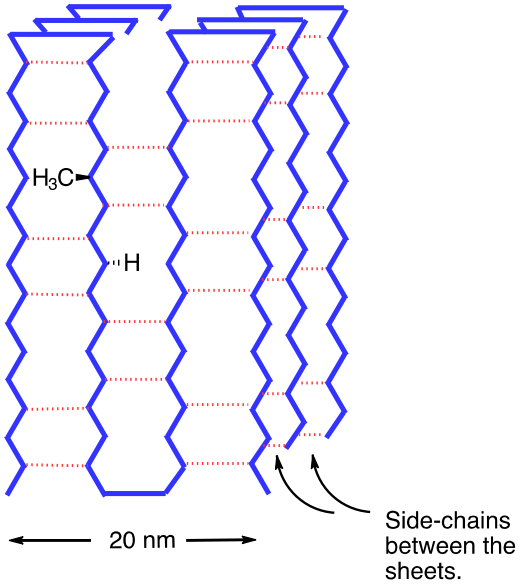
Collagen fibers are formed from staggered arrays of tropocollagen. Covalent cross-links occur between lysines at the N-terminus.

Collagen has a greater tensile strength than steel for fibers of equal weight.

The figure on the previous slide is from an article by Po-Yu Chen[a,], Joanna McKittrick[b, c], Marc André Meyer in Progress in Materials Science Volume 57, Issue 8, November 2012, Pages 1492–1704

3. **Silk Fibroin** is made from stacked antiparallel β-sheets rich in Gly and Ala that permit close packing of the sheets.

It is soft and flexible.



Silk Fibroin