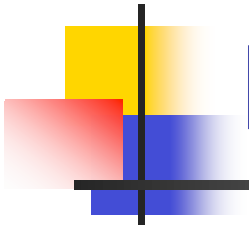


# Protein structure, intro

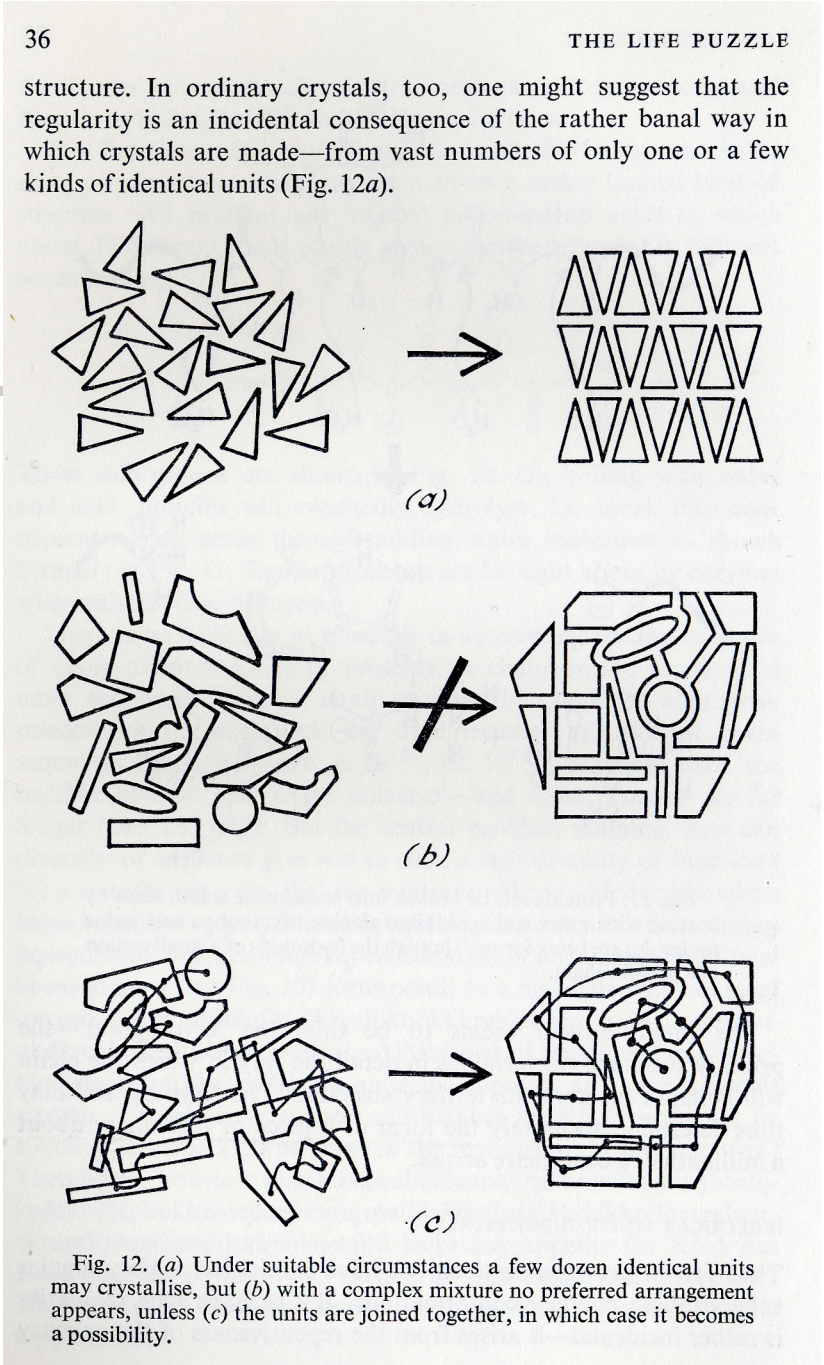
- Linear polymer of amino acids
- AA sequence encodes information for 3D folding of protein
- Sequence dictates structure -- structure dictates function
- Cooperative folding

# Structure formation

Arrangement of a diverse set of molecular shapes along a linear polymer allows formation of specific, stable 3D structures.

The Life Puzzle

A.G. Cairns-Smith



36 THE LIFE PUZZLE

structure. In ordinary crystals, too, one might suggest that the regularity is an incidental consequence of the rather banal way in which crystals are made—from vast numbers of only one or a few kinds of identical units (Fig. 12*a*).

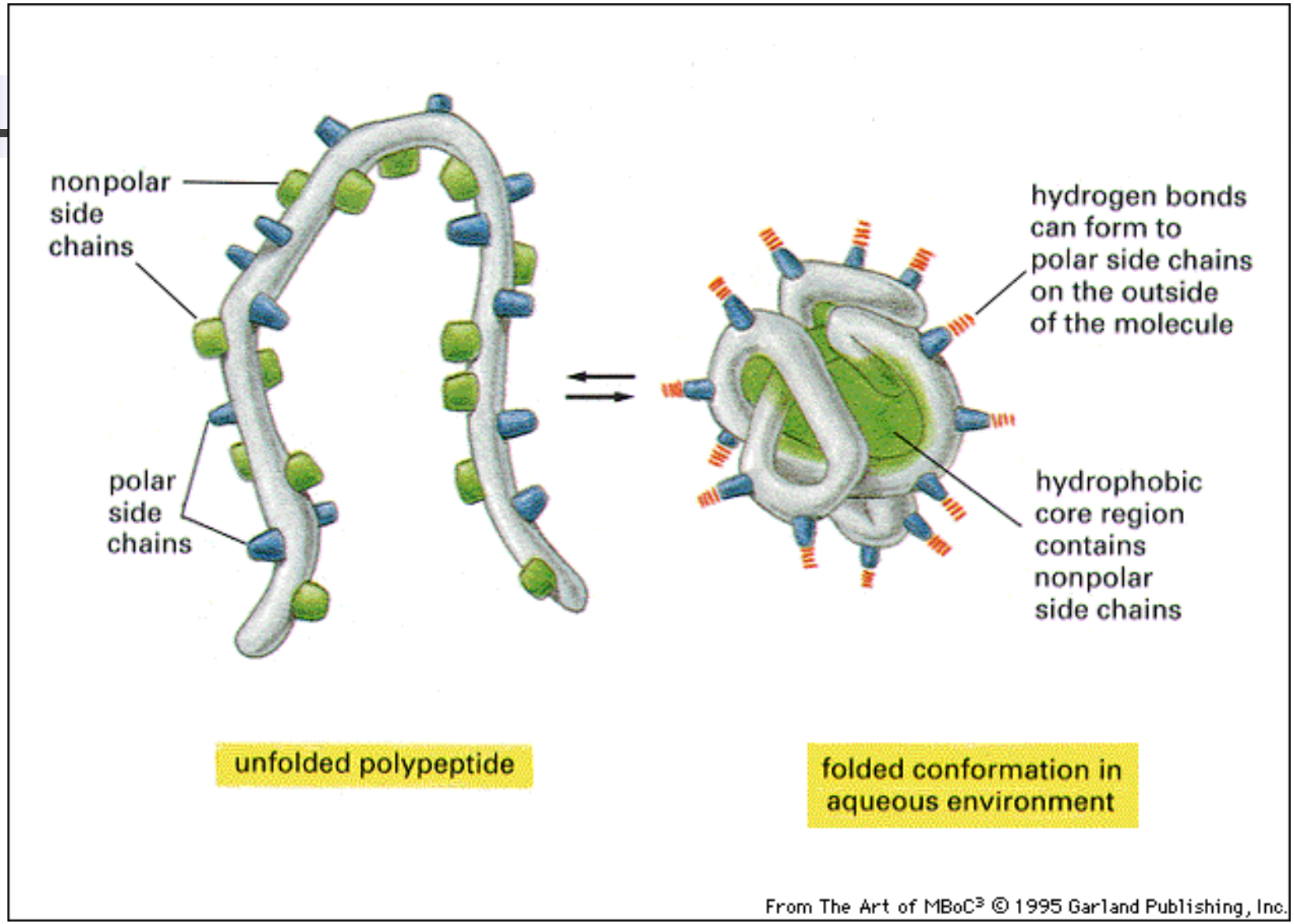
(*a*)

(*b*)

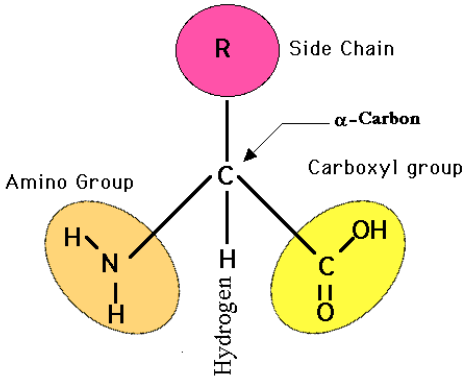
(*c*)

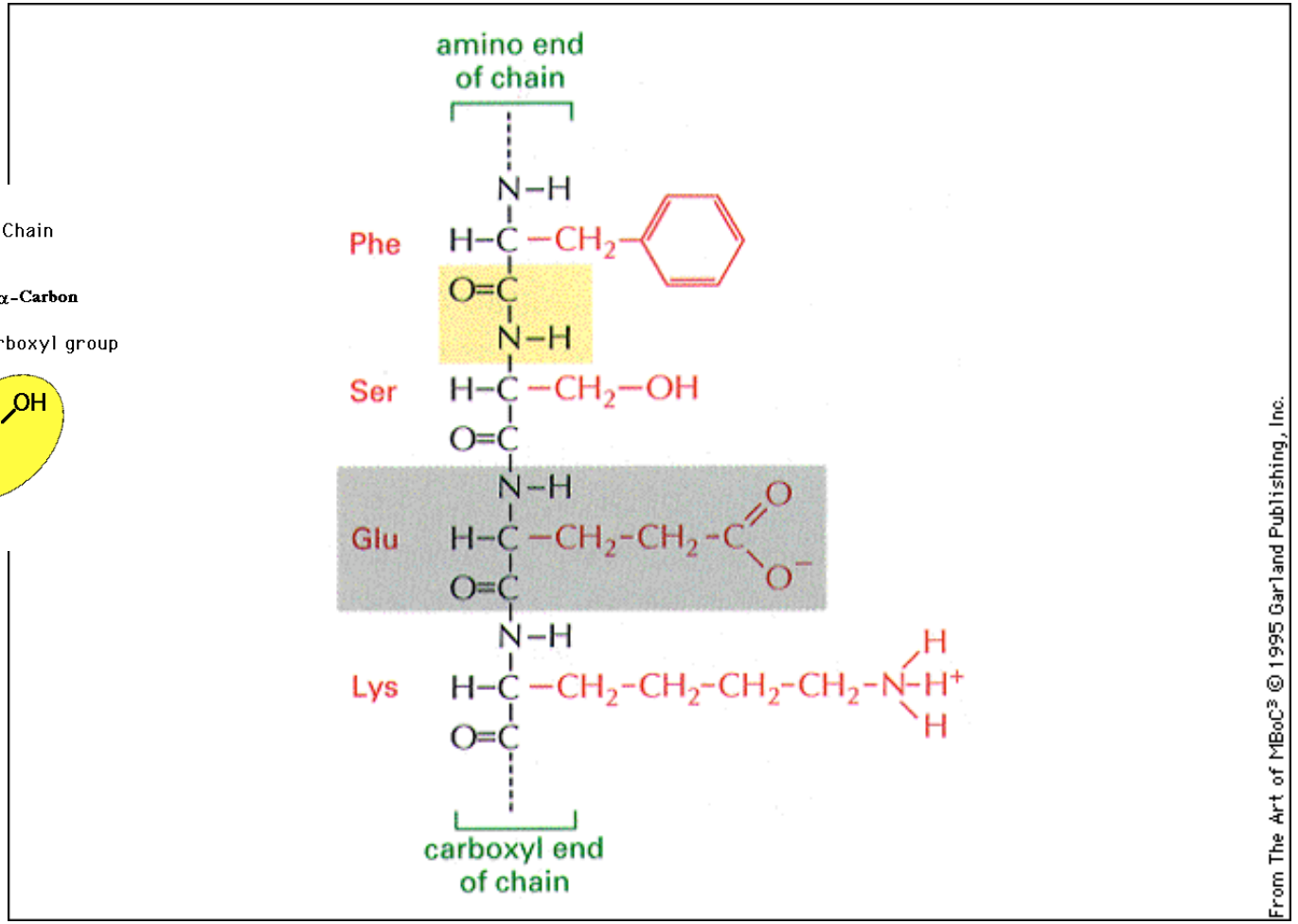
Fig. 12. (*a*) Under suitable circumstances a few dozen identical units may crystallise, but (*b*) with a complex mixture no preferred arrangement appears, unless (*c*) the units are joined together, in which case it becomes a possibility.

# Protein, hydropathic folding



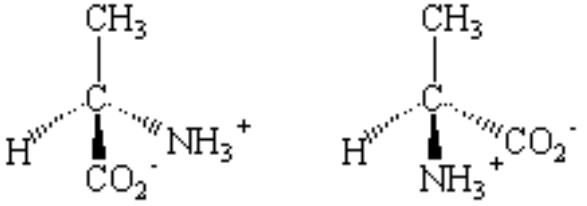
# Protein, primary structure



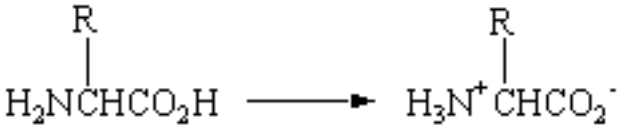



# Amino acids, properties

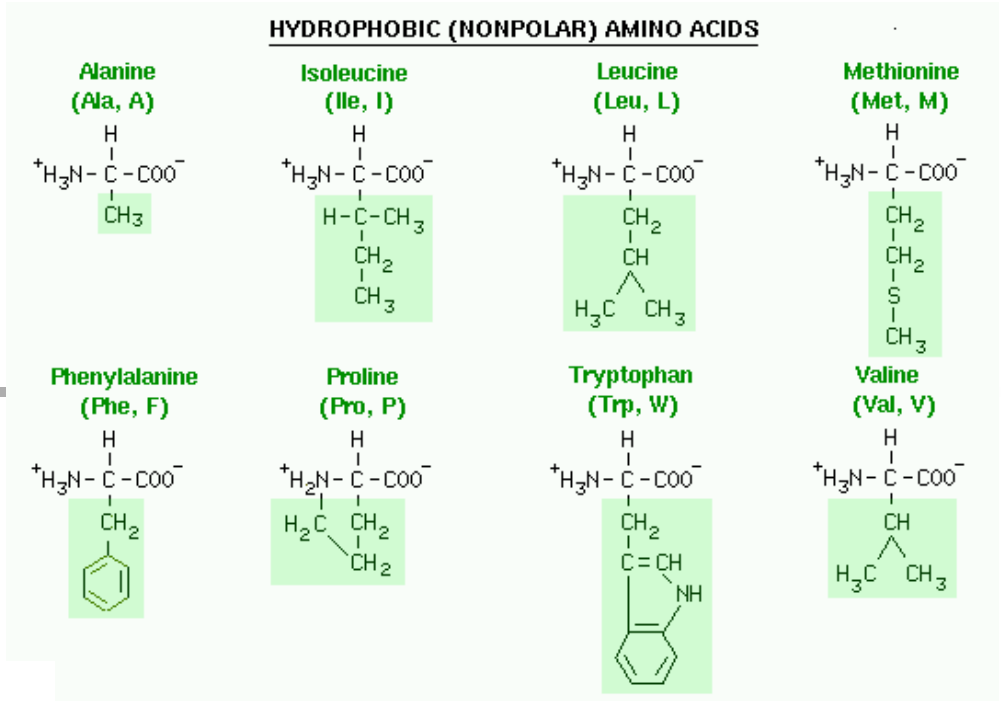
- Stereoisomers D-alanine L-alanine

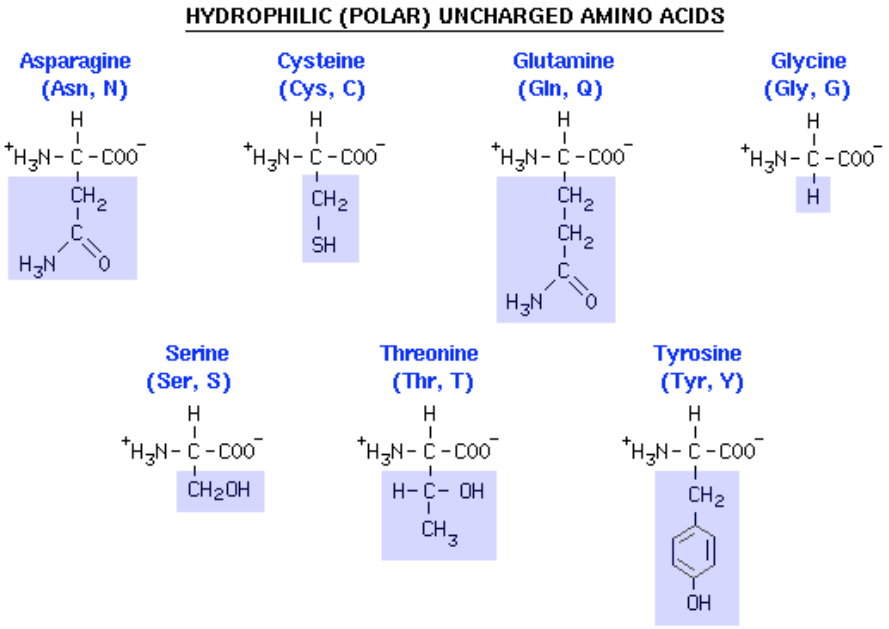


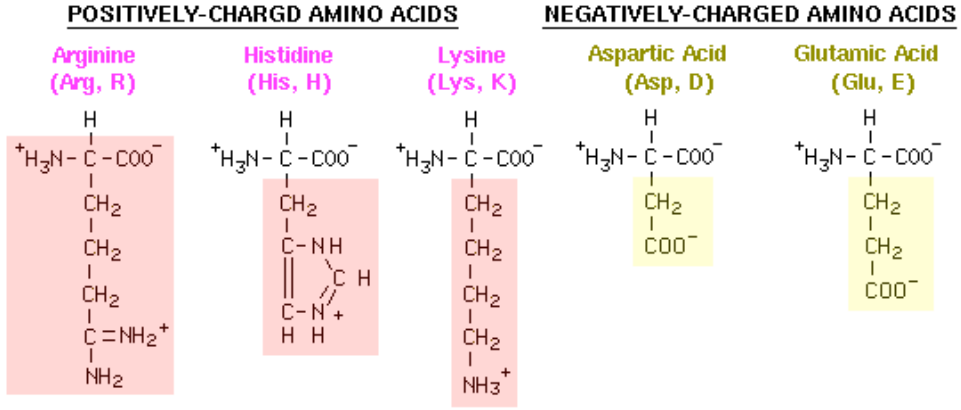

- Zwitterionic form



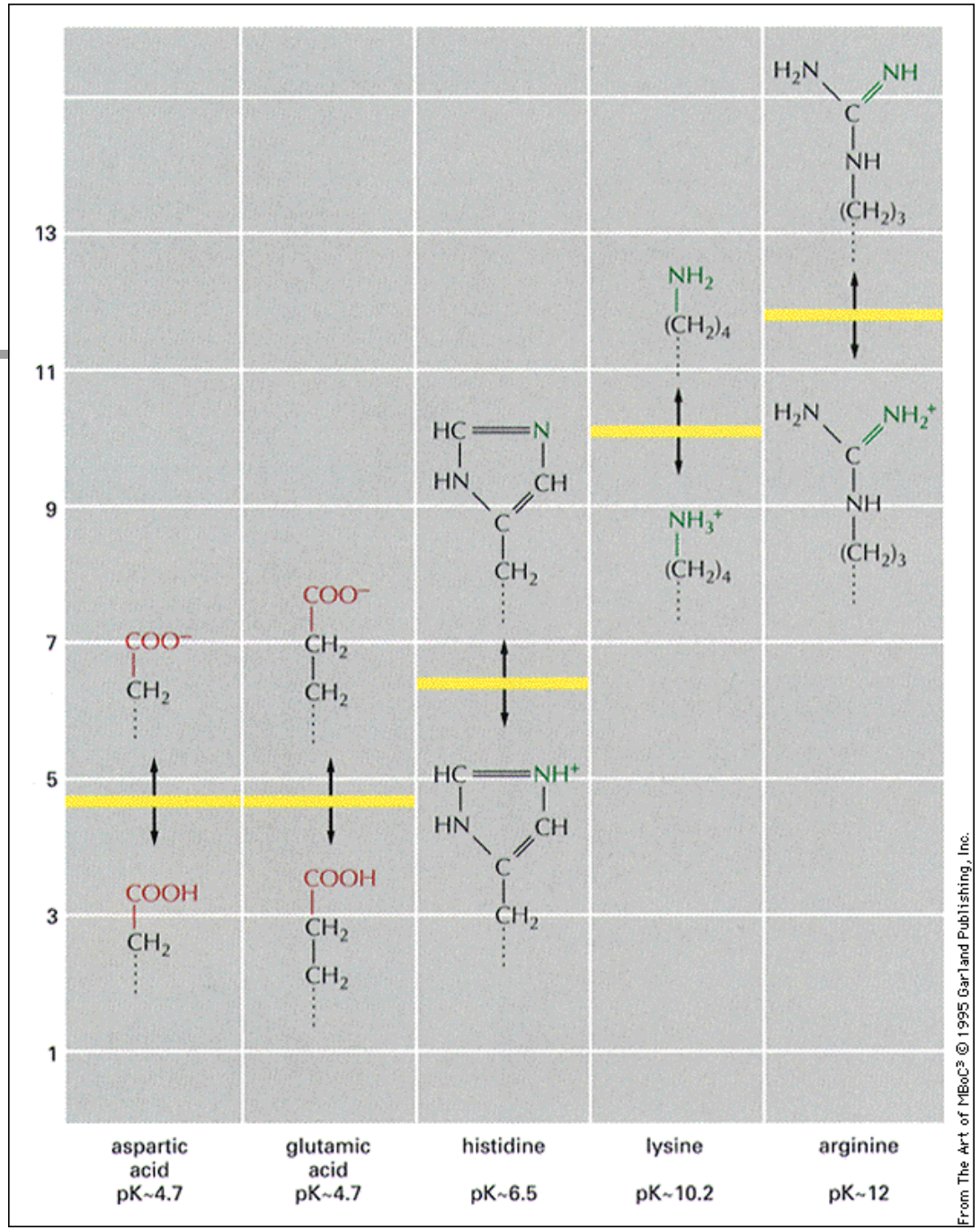
# 20 common amino acids



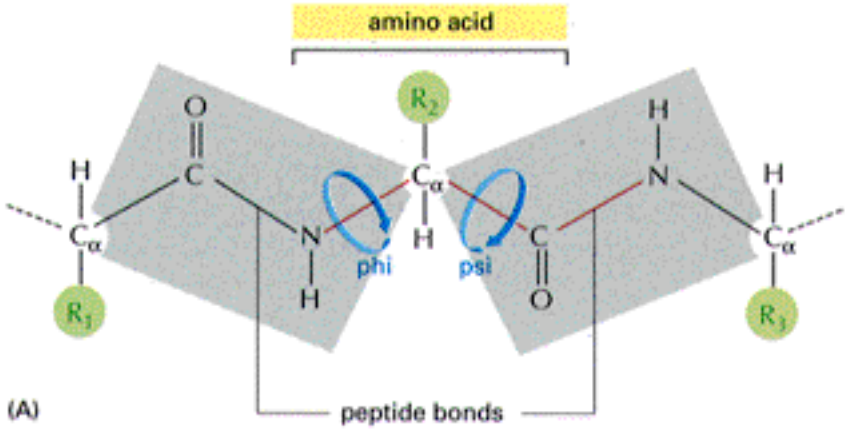



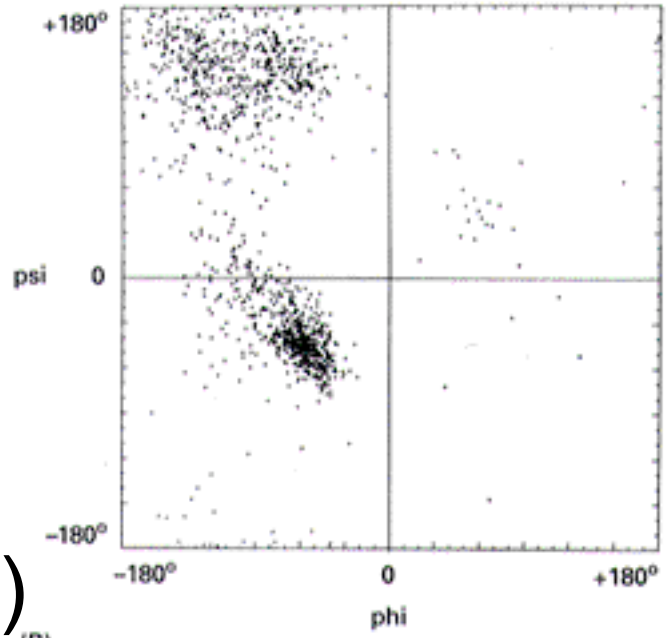
# AA pKa


13
11
9
7
5
3
1
aspartic acid pK~4.7
glutamic acid pK~4.7
histidine pK~6.5
lysine pK~10.2
arginine pK~12
From The Art of MBoC3 © 1995 Garland Publishing, Inc.

# Backbone torsion angles



(A)

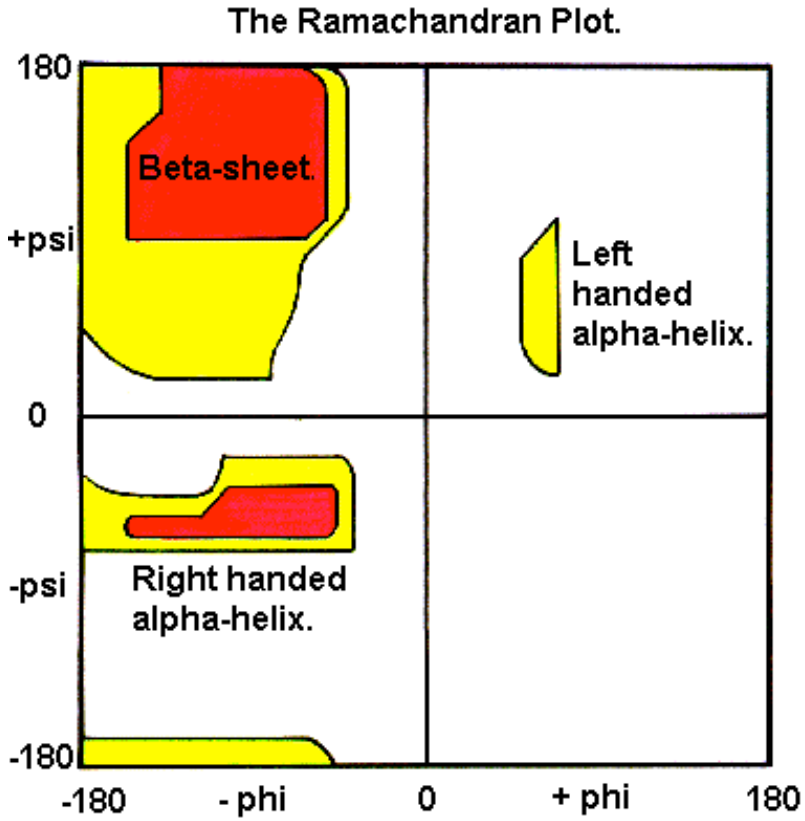


(B)

- Preferred angles (Pauling)
- Ramachandran Plot
- Repeated angles = Sec. Struct.

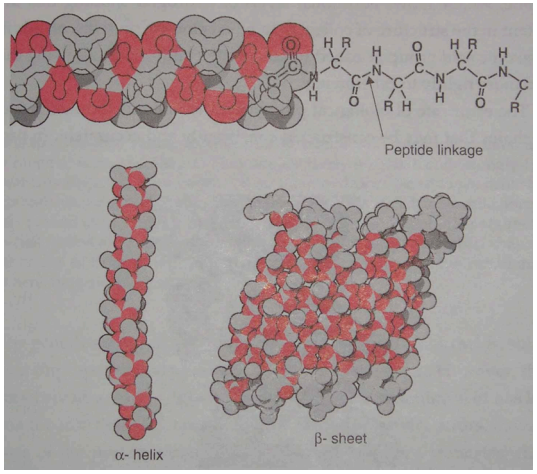
# Ramachandran Plot

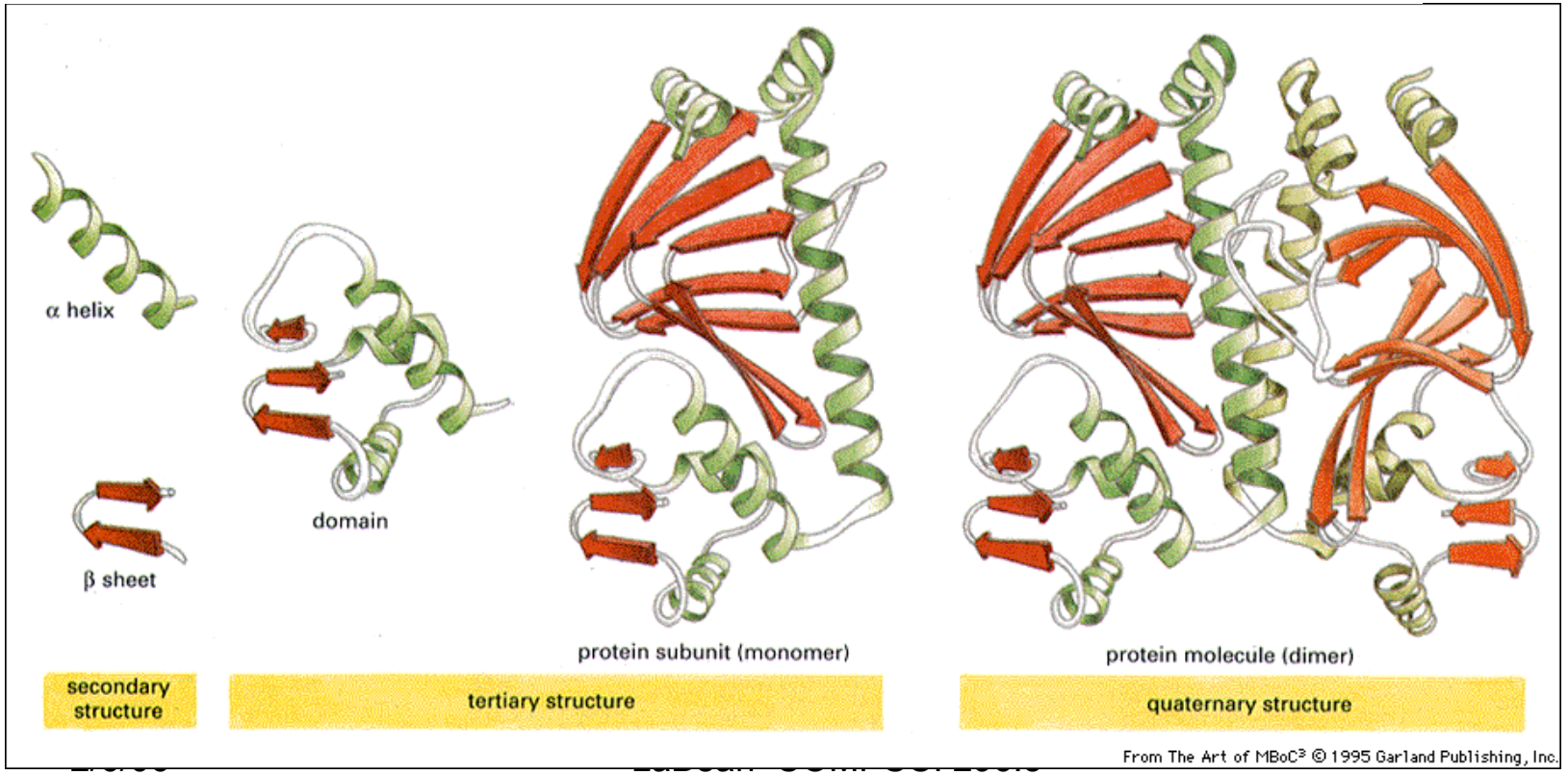
Red regions are conformations with no steric clashes. Yellow areas are allowed regions if slightly shorter van der Waals radi are used. Thus the left-handed alpha-helix region appears. L-amino acids cannot form extended regions of left-handed helix but occassionally individual residues adopt this conformation. These residues are usually glycine but can also be asparagine or aspartate where the side chain forms a hydrogen bond with the main chain and therefore stabilizes this otherwise unfavorable conformation. The 3(10) helix occurs close to the upper right of the alpha-helical region and is on the edge of allowed region indicating lower stability.



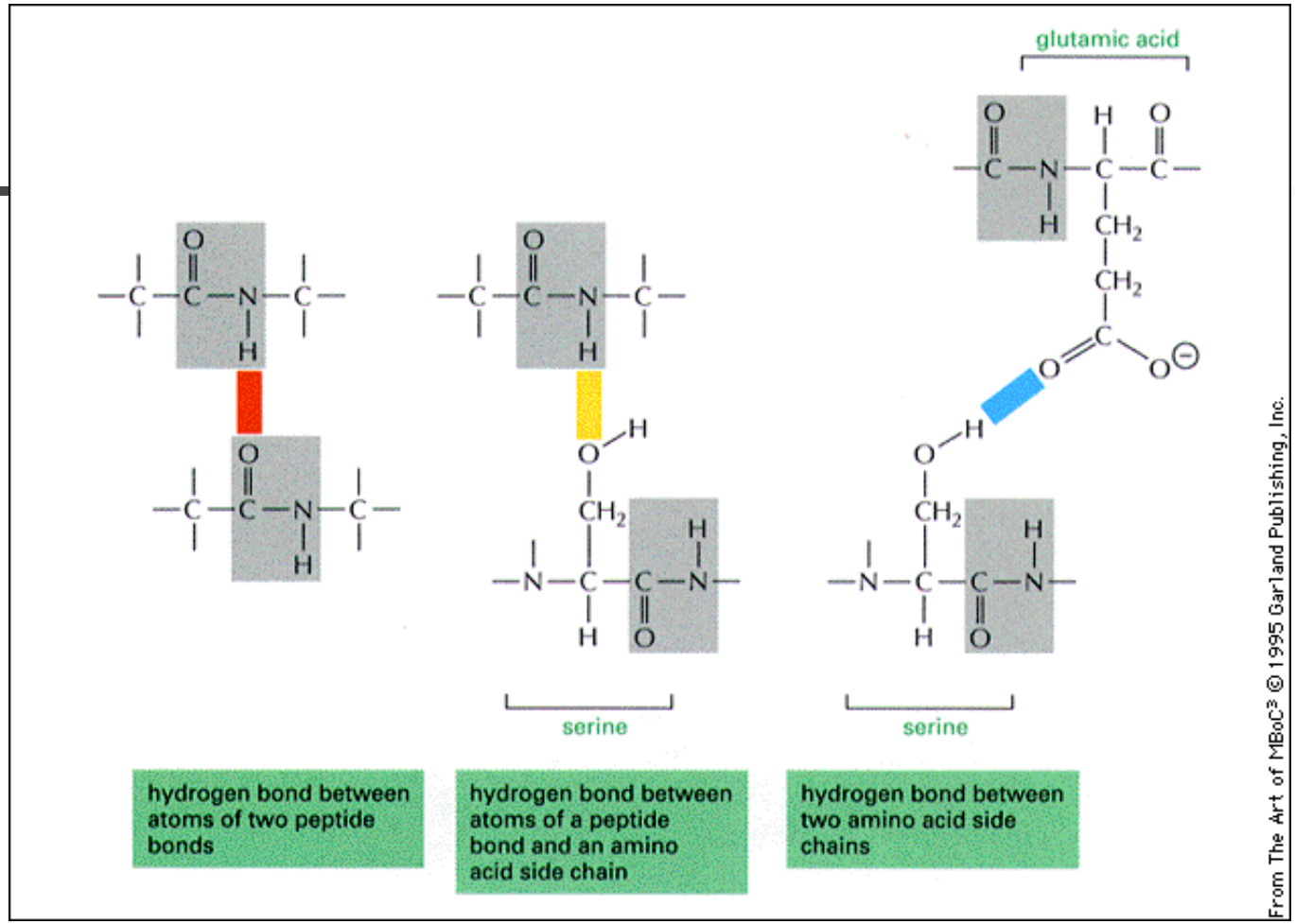
# Protein, structural levels

- Primary
  - AA sequence, covalent chemical structure
- Secondary
  - Local chain folding (helix, sheet, turn)
- Tertiary
  - Global chain folding (domain formation)
- Quaternary
  - Association of well-folded units



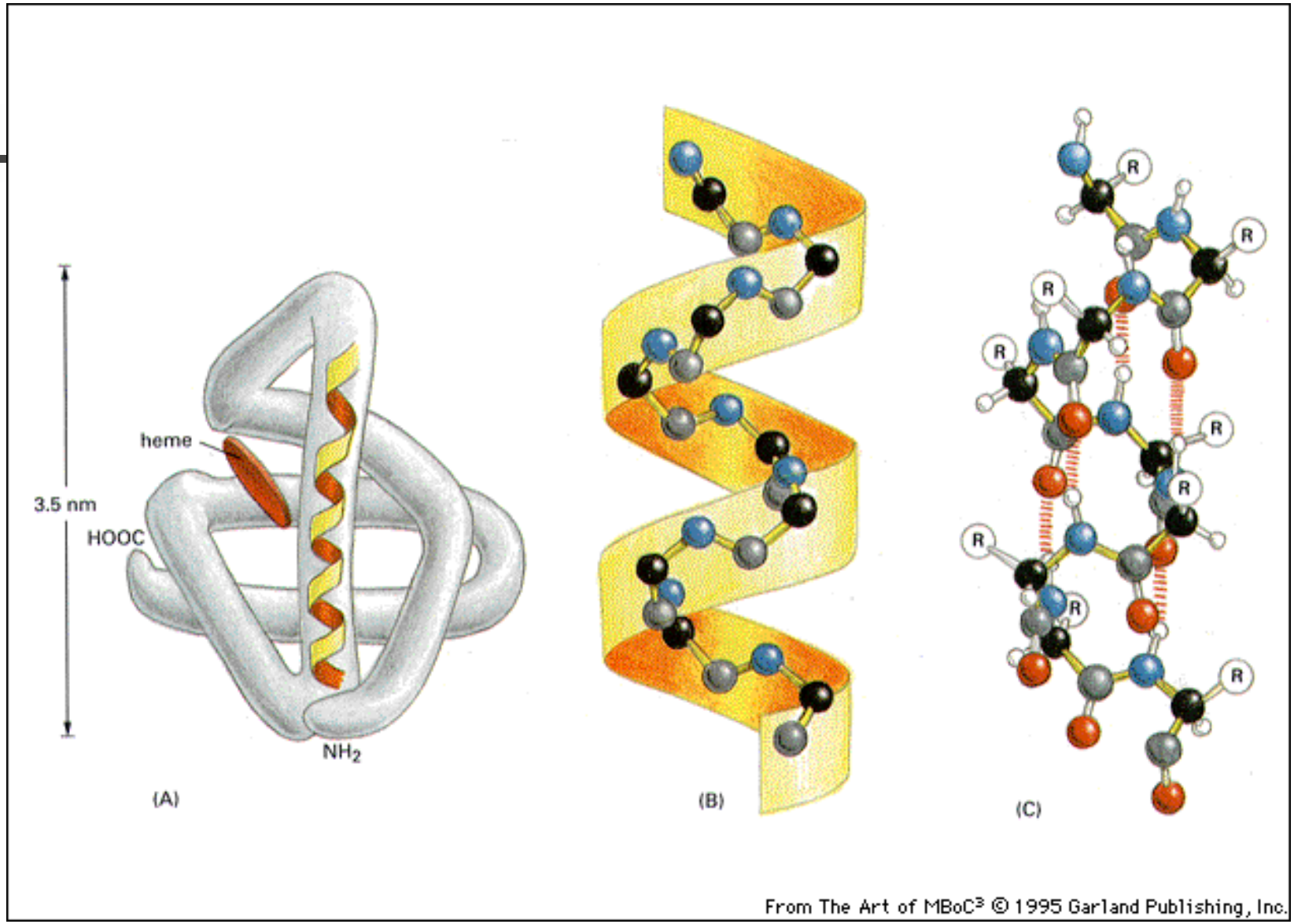



# Protein, H-bonds

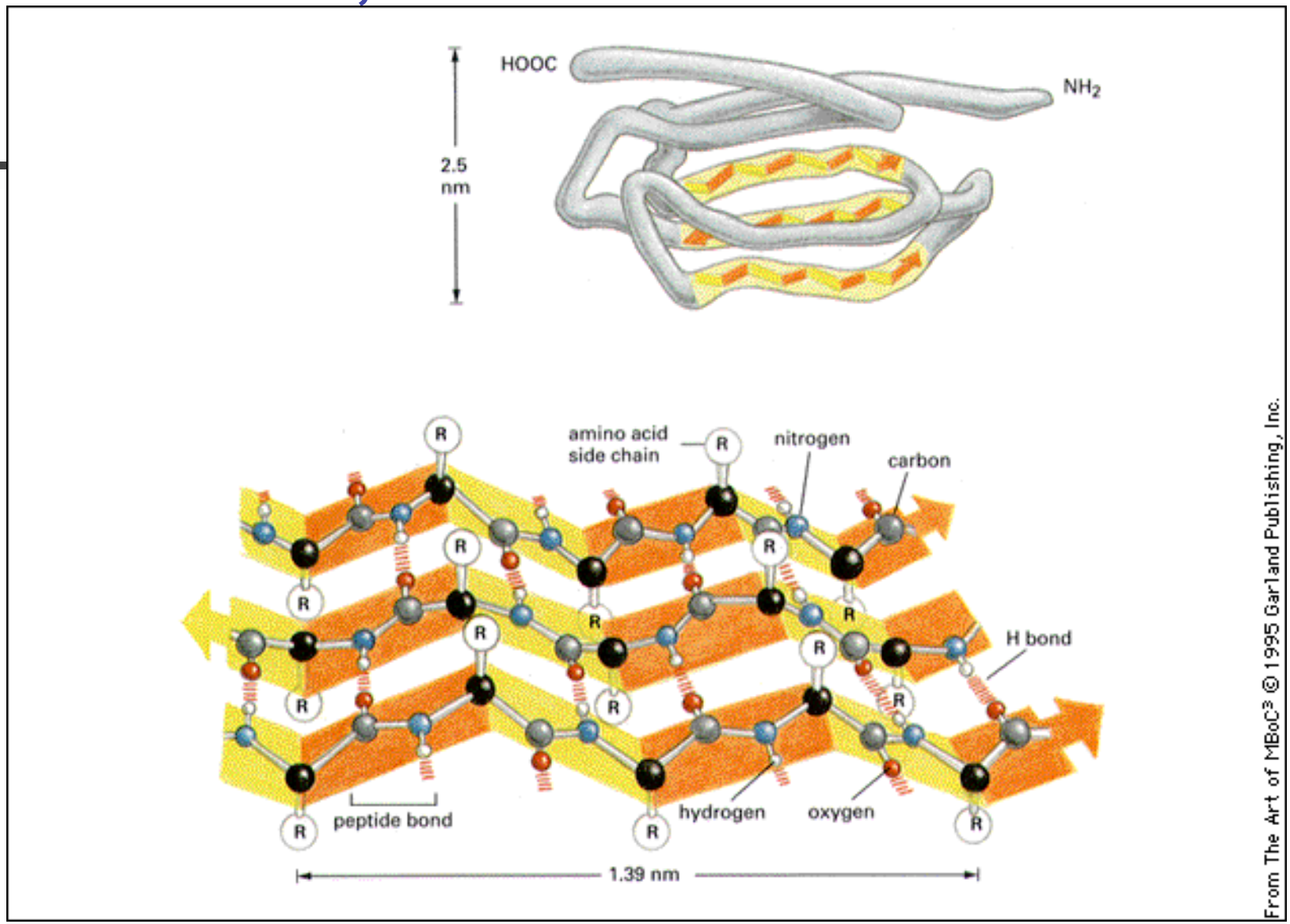



Mainchain-mainchain mainchain-sidechain sidechain-sidechain

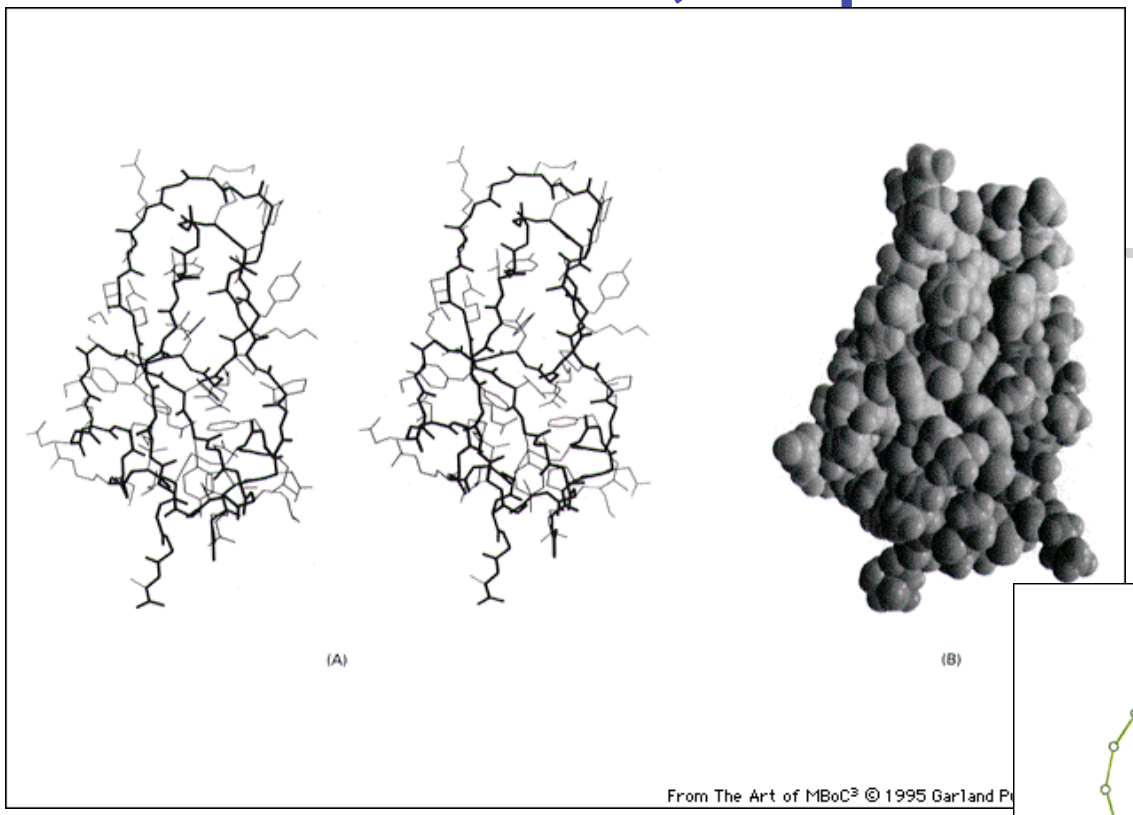
# Protein, alpha-helix

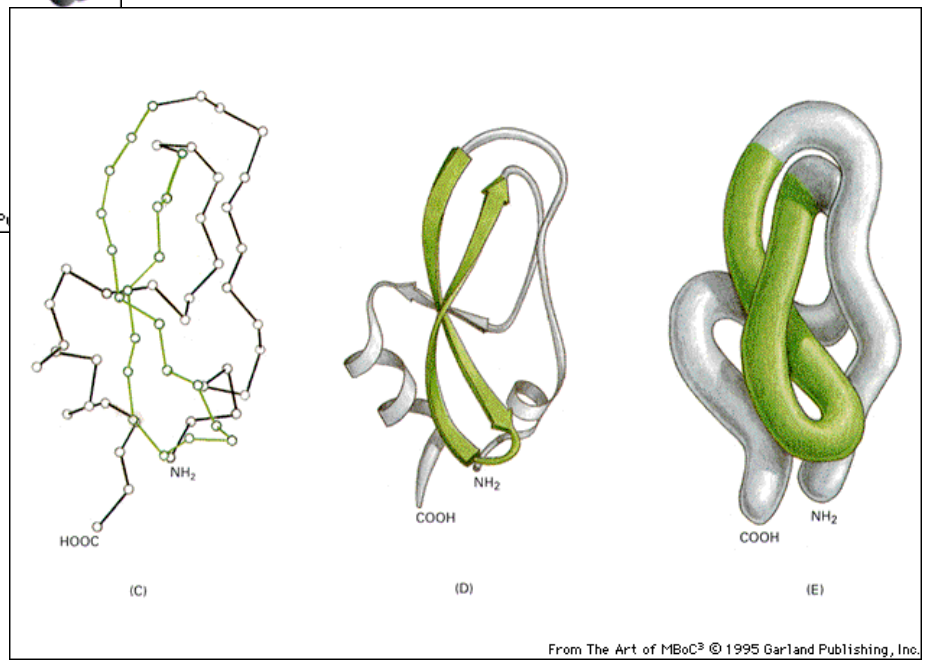


# Protein, beta-sheet

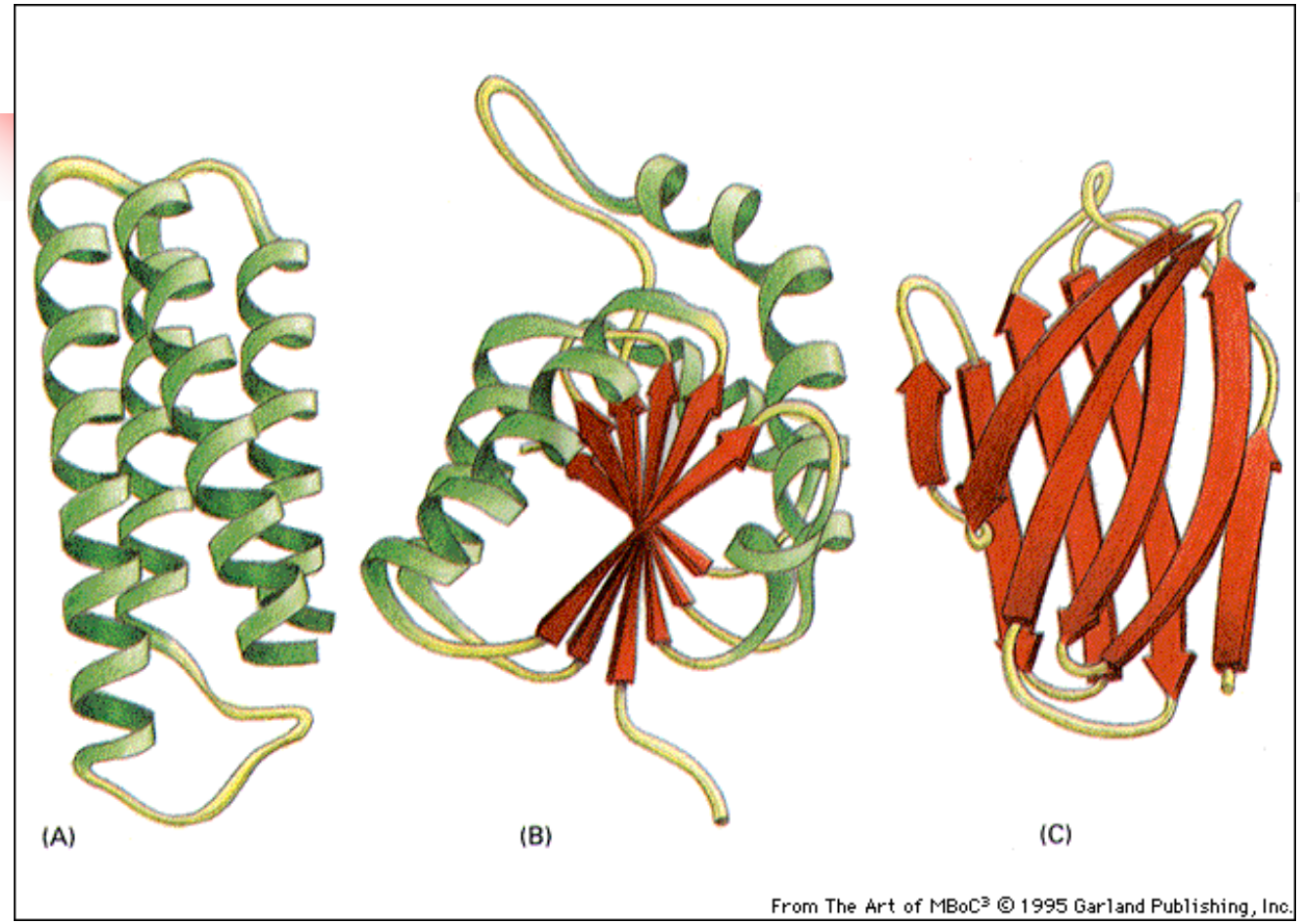


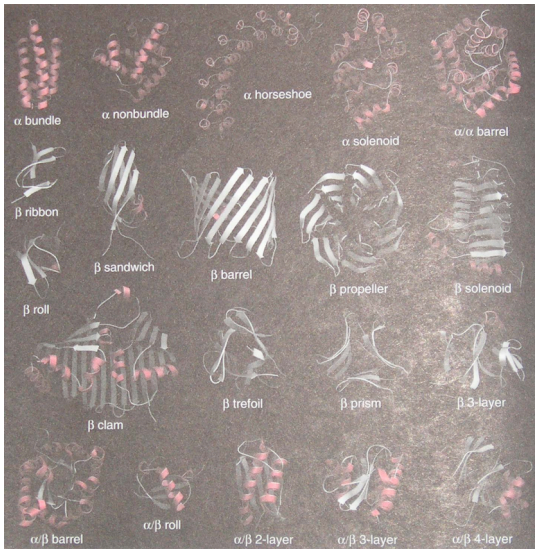
# Protein, representations



# Protein, structural families

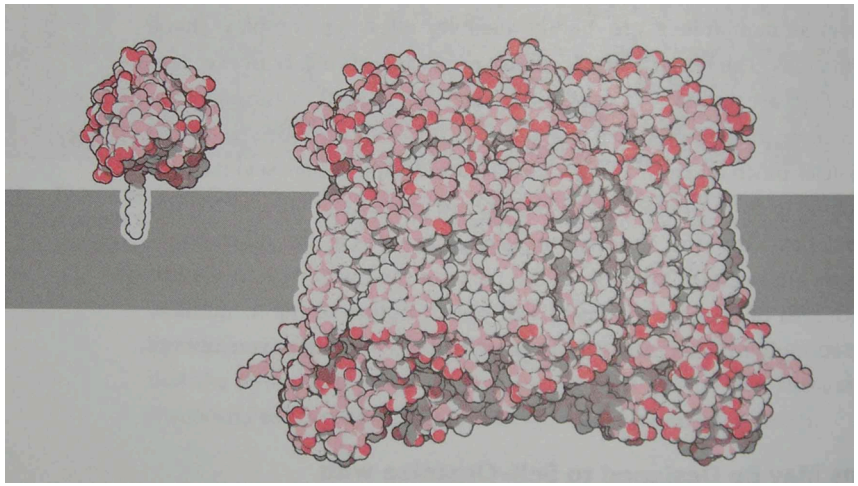


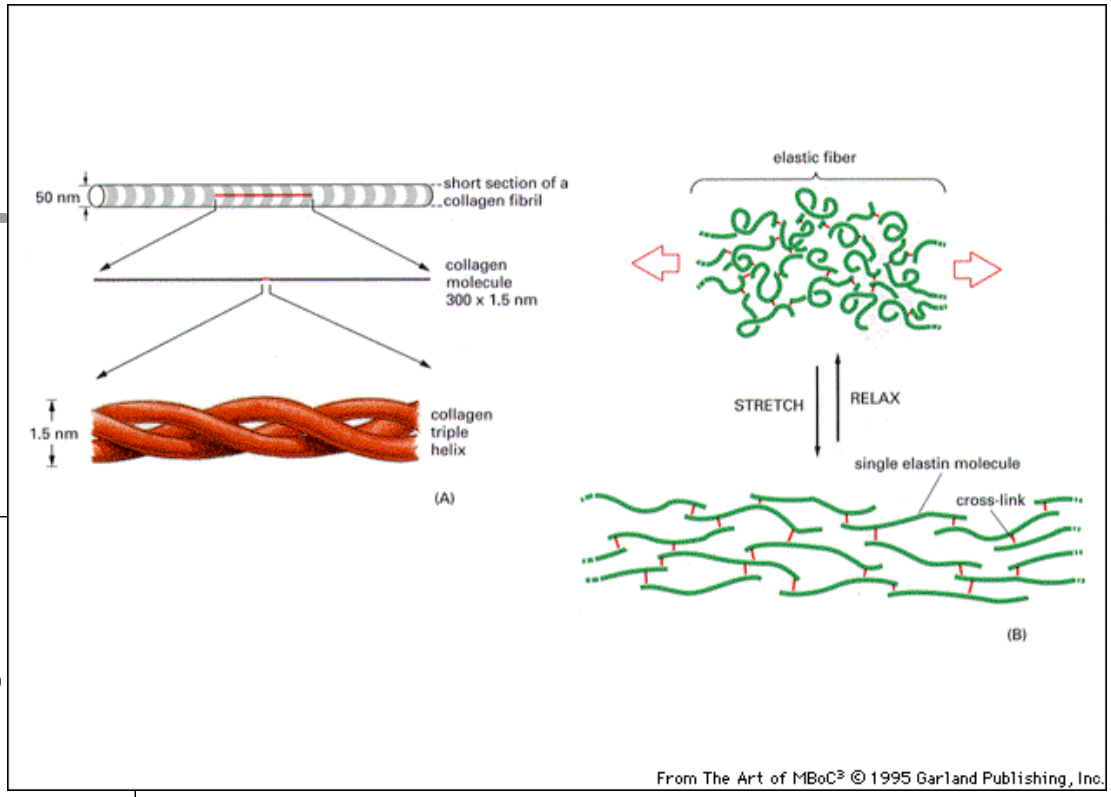

See: Anatomy & Taxonomy of Protein Structure
http://kinemage.biochem.duke.edu/~jsr/

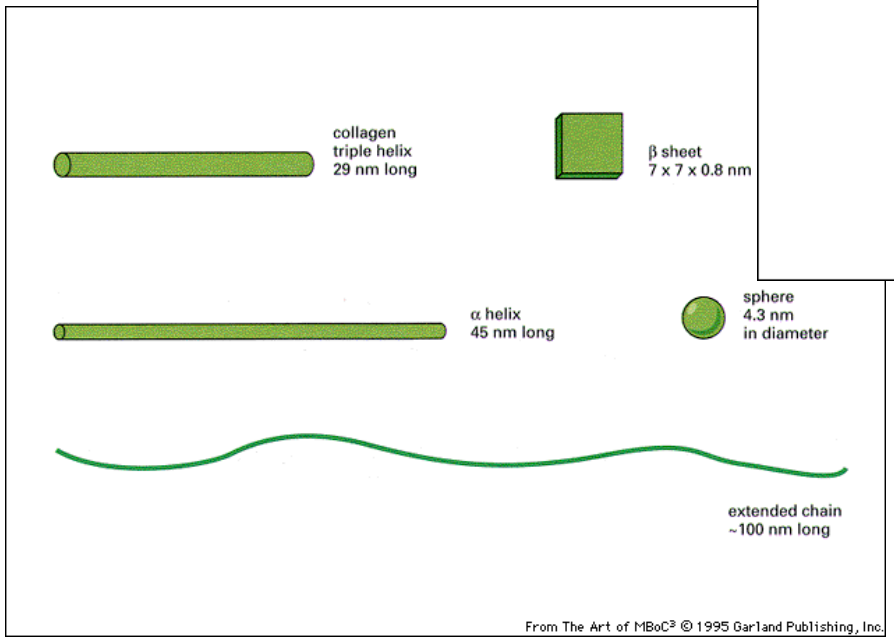
# Protein, classes

- Globular (soluble)
- Fibrous (structural)
- Membrane



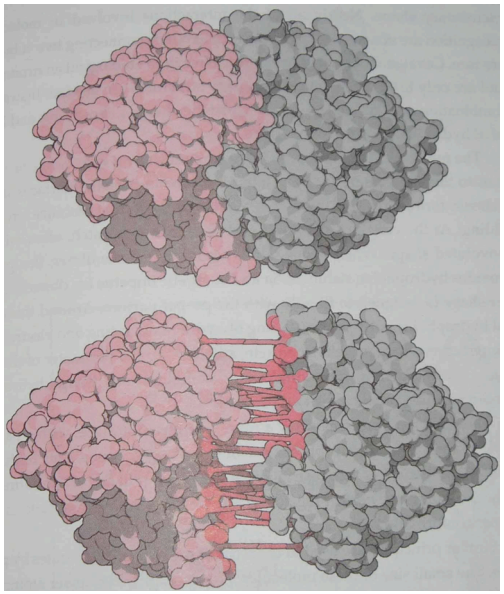
# Structural Proteins



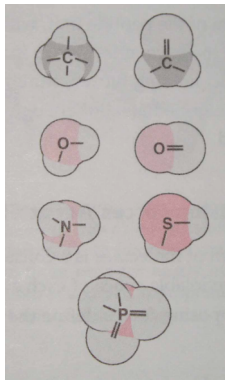
# Molecular recognition

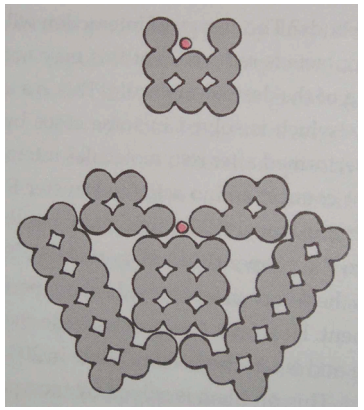
- Crane design concepts
  - High specificity requires multiple and weak contacts.
  - Matching particles must have geometric complementarity.



Goodsell

- Also:
  - Large overhead of surrounding infrastructure can position a few key sites.

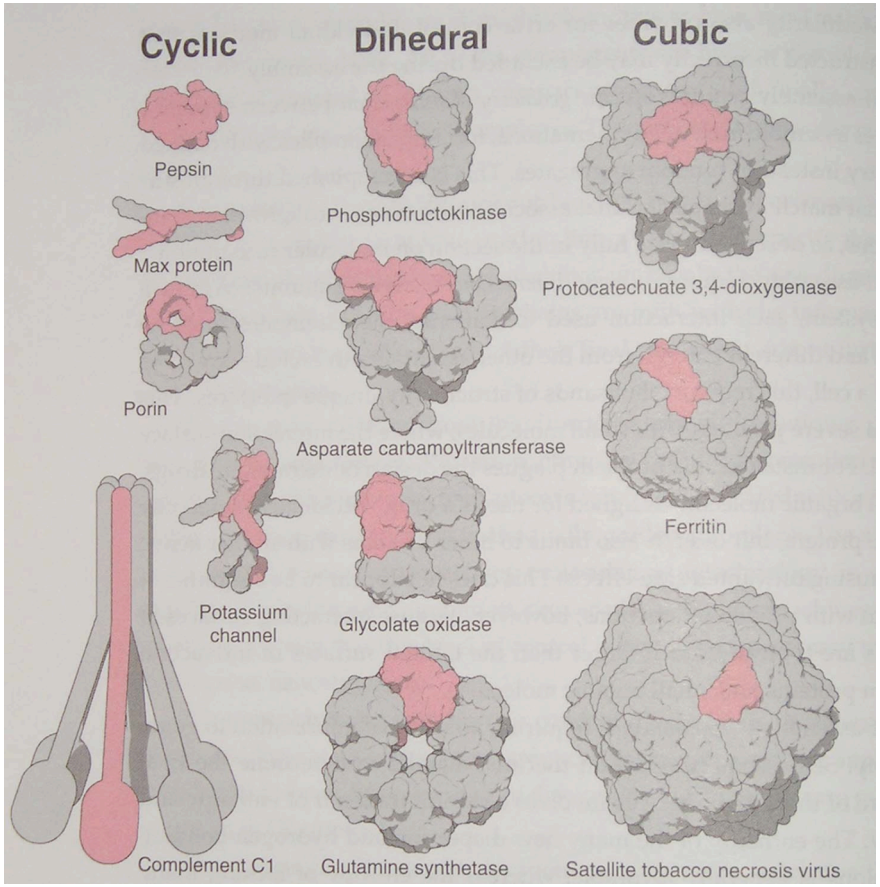


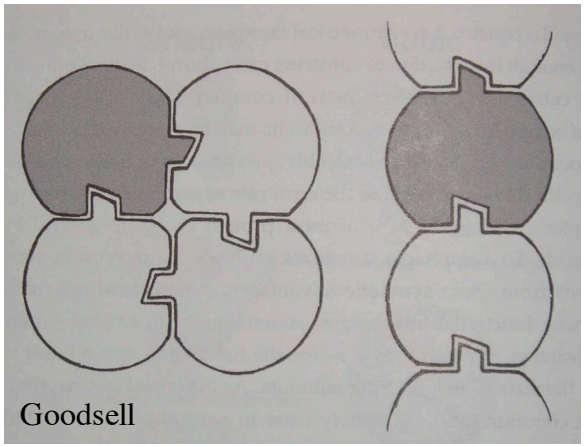



Goodsell

# Symmetry

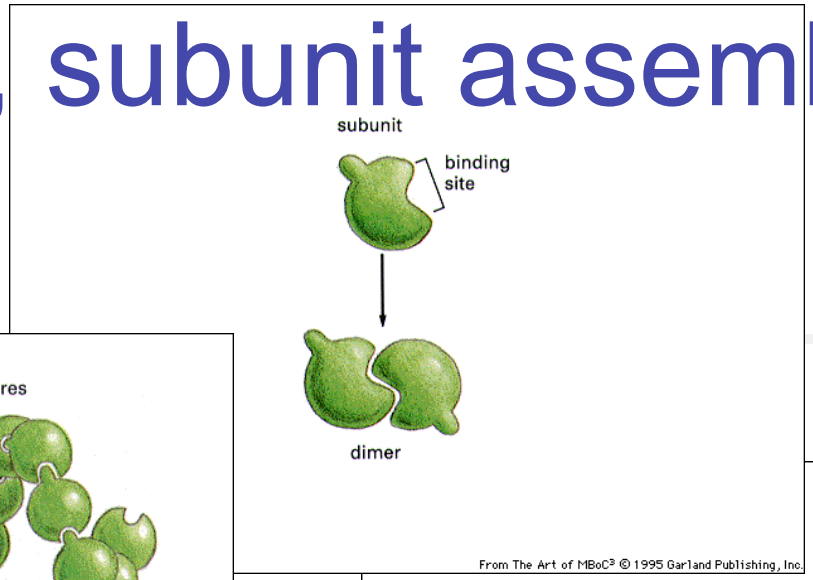
- Hierarchical structures of defined size.

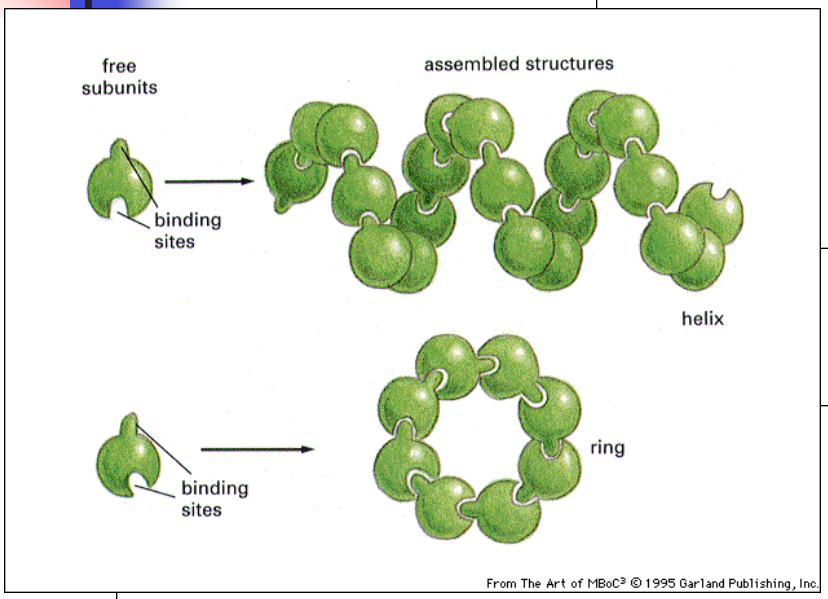


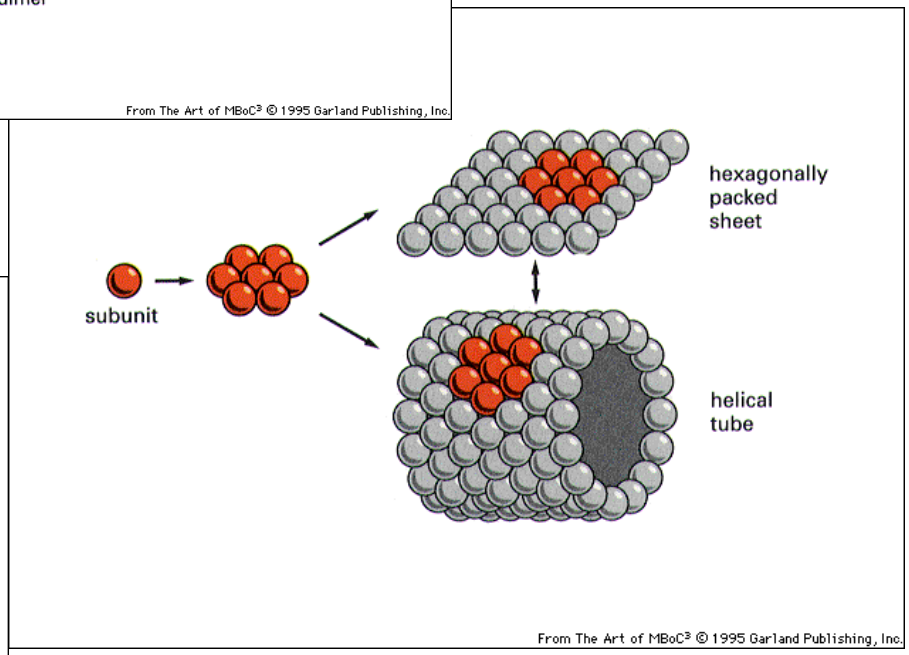



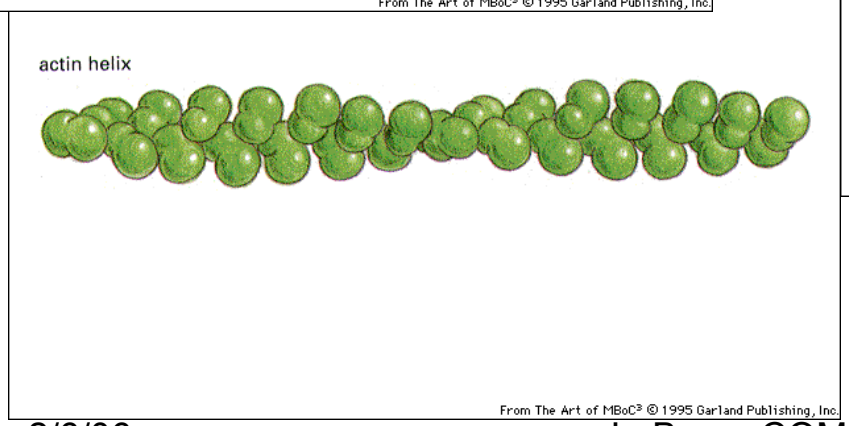

Various plans for placement of interaction surfaces…

# Protein, subunit assembly

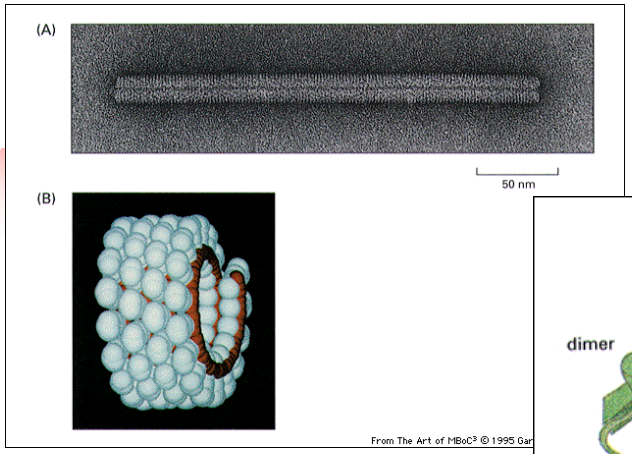


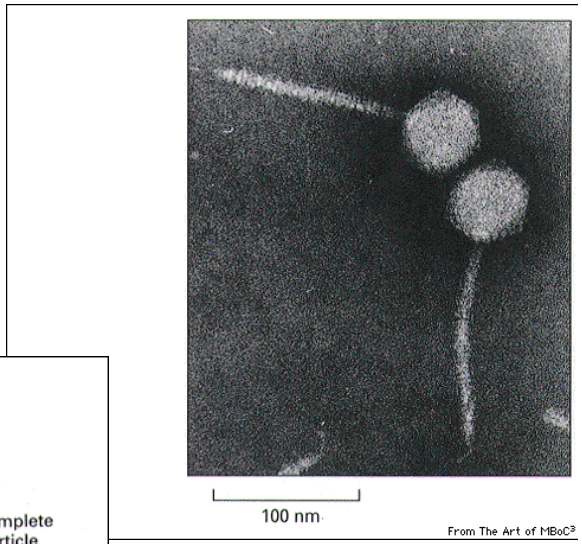



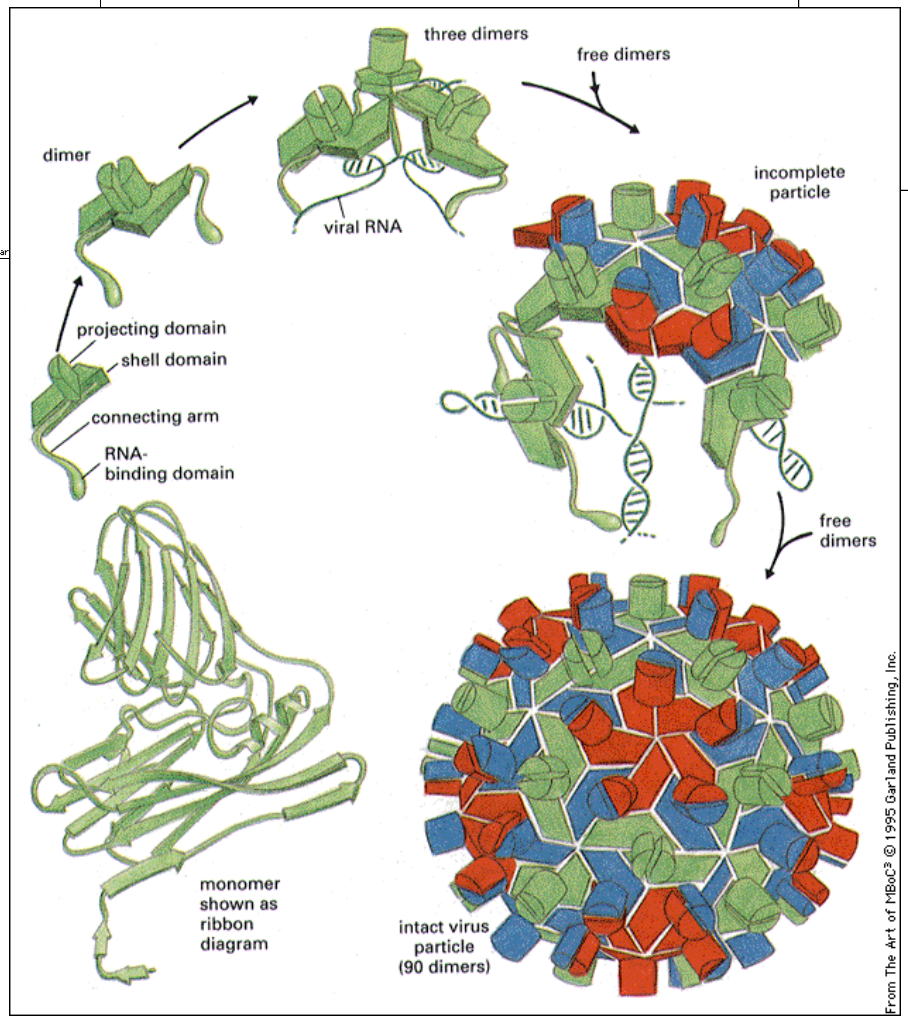



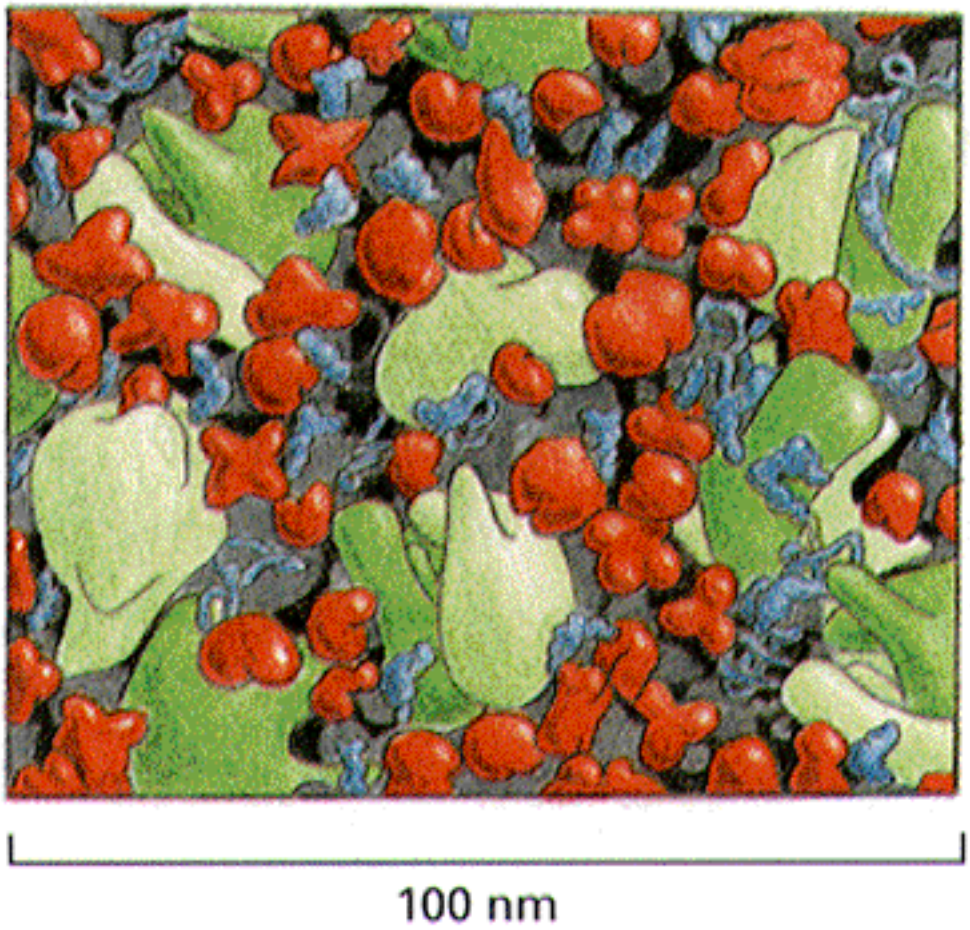
# Viral Assembly



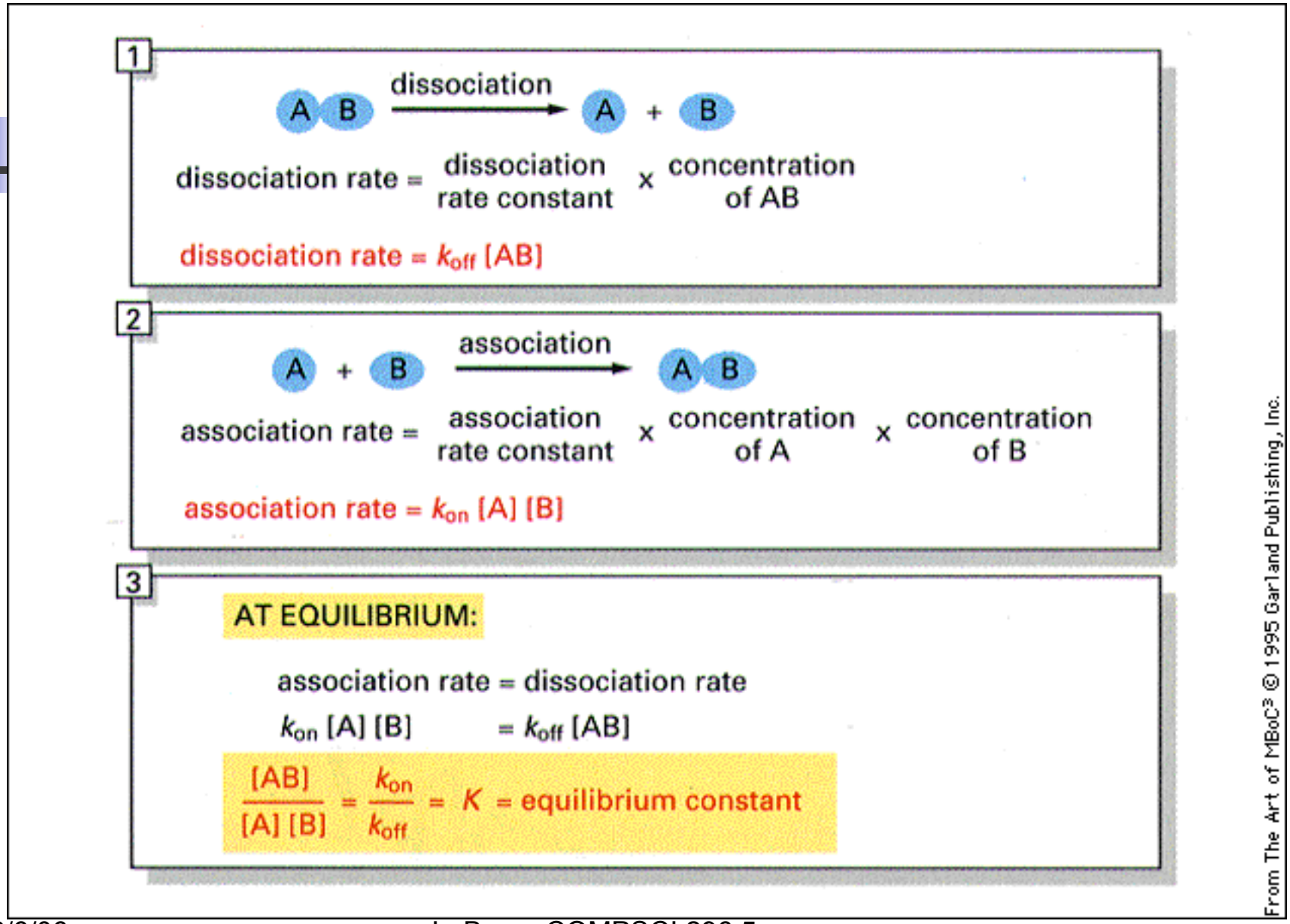



# Protein, bioconcentration



# Protein, interaction kinetics 1/2



# Protein, interaction kinetics 2/2

EXAMPLE:

The concentration of a molecule present in only one copy in a typical mammalian cell (volume of 2000 $\mu m^3$) is about $10^{-12}$M.
If such a cell contains $10^4$ copies of protein molecule A and $10^6$ copies of protein molecule B,

$$[A] = 10^{-8}\ M \quad \text{and} \quad [B] = 10^{-6}\ M$$

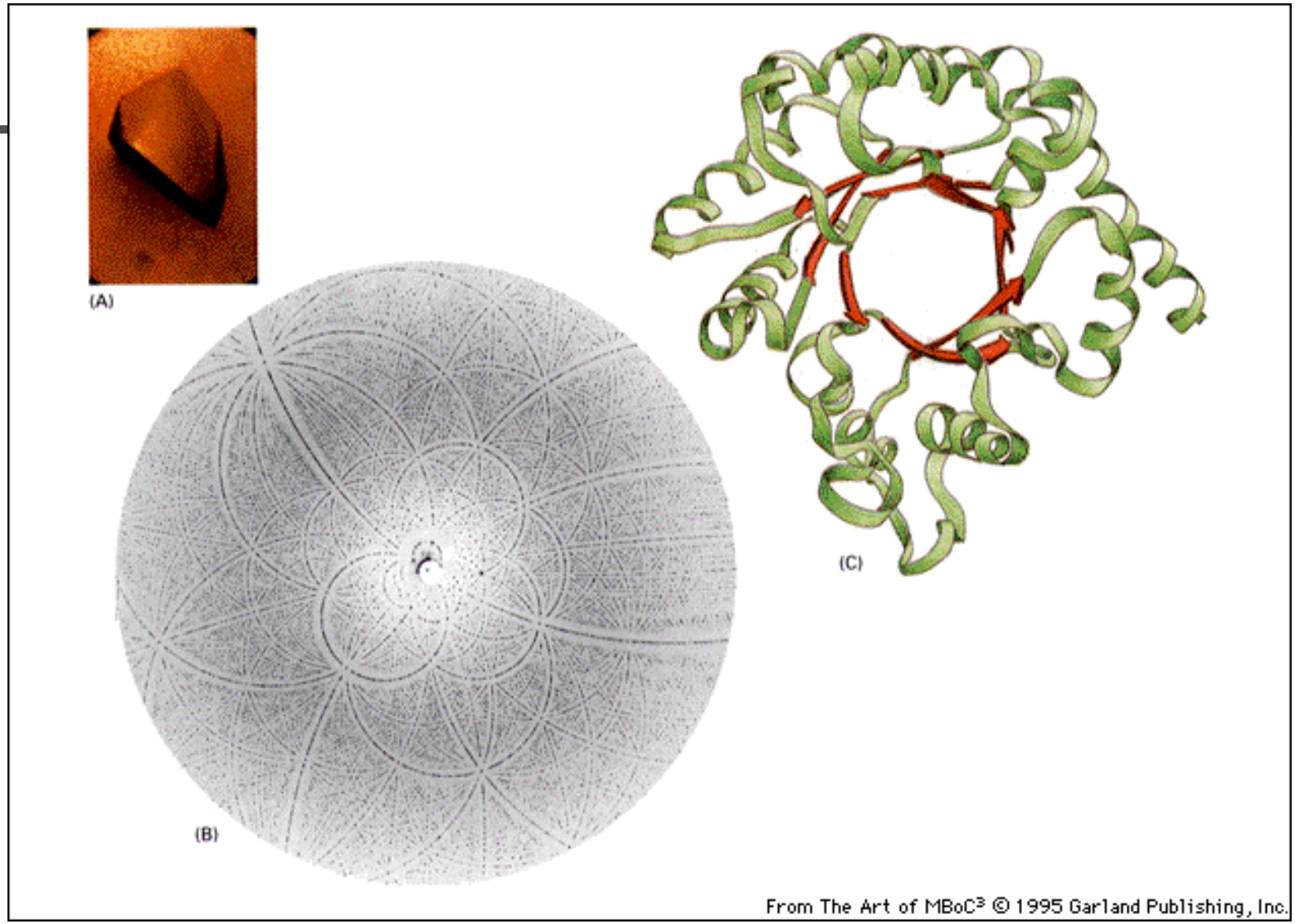
Suppose that protein A binds to protein B with $K = 10^7 M^{-1}$. The ratio of bound to unbound A will be [AB]/[A] and since

$$\frac{[AB]}{[A]} = K[B] = (10^7\ M^{-1})(10^{-6}\ M) = 10$$

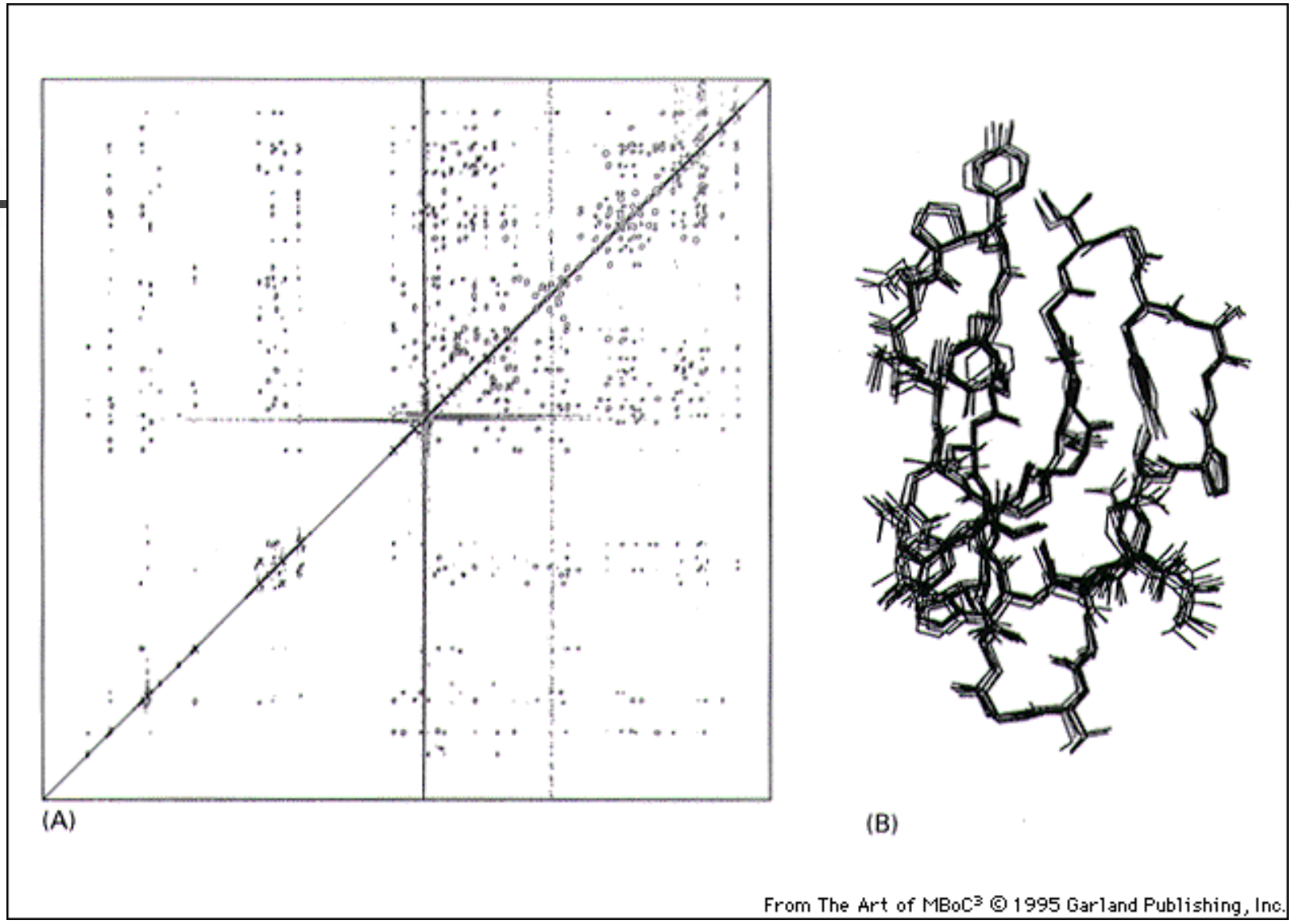
we expect one molecule of A to be free for every 10 molecules of A that are bound to B inside the cell.

Repeating the calculation for $K = 10^4 M^{-1}$ shows that only about one molecule of A in a hundred would be bound to B.

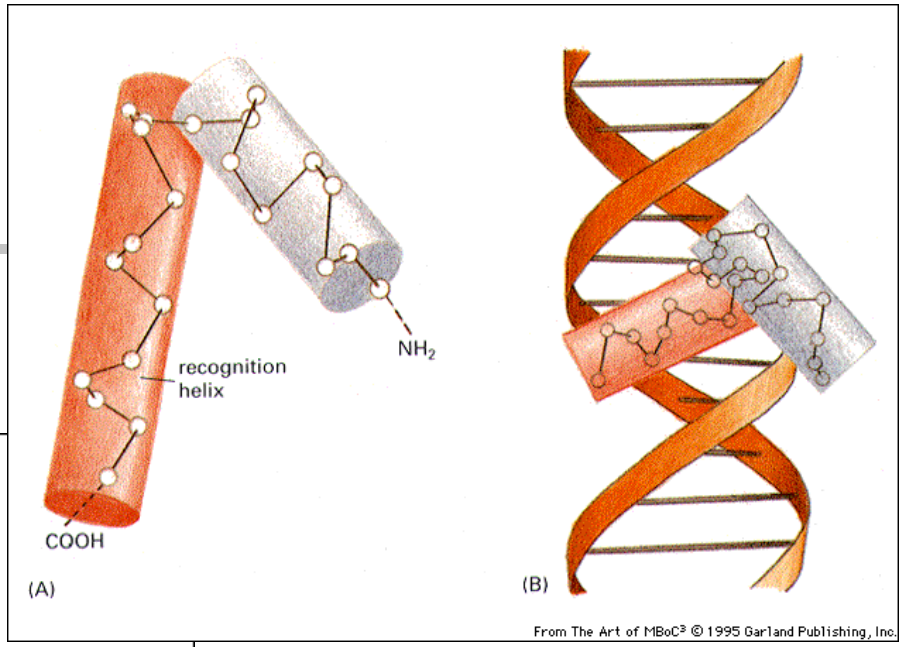
# Protein, crystallography

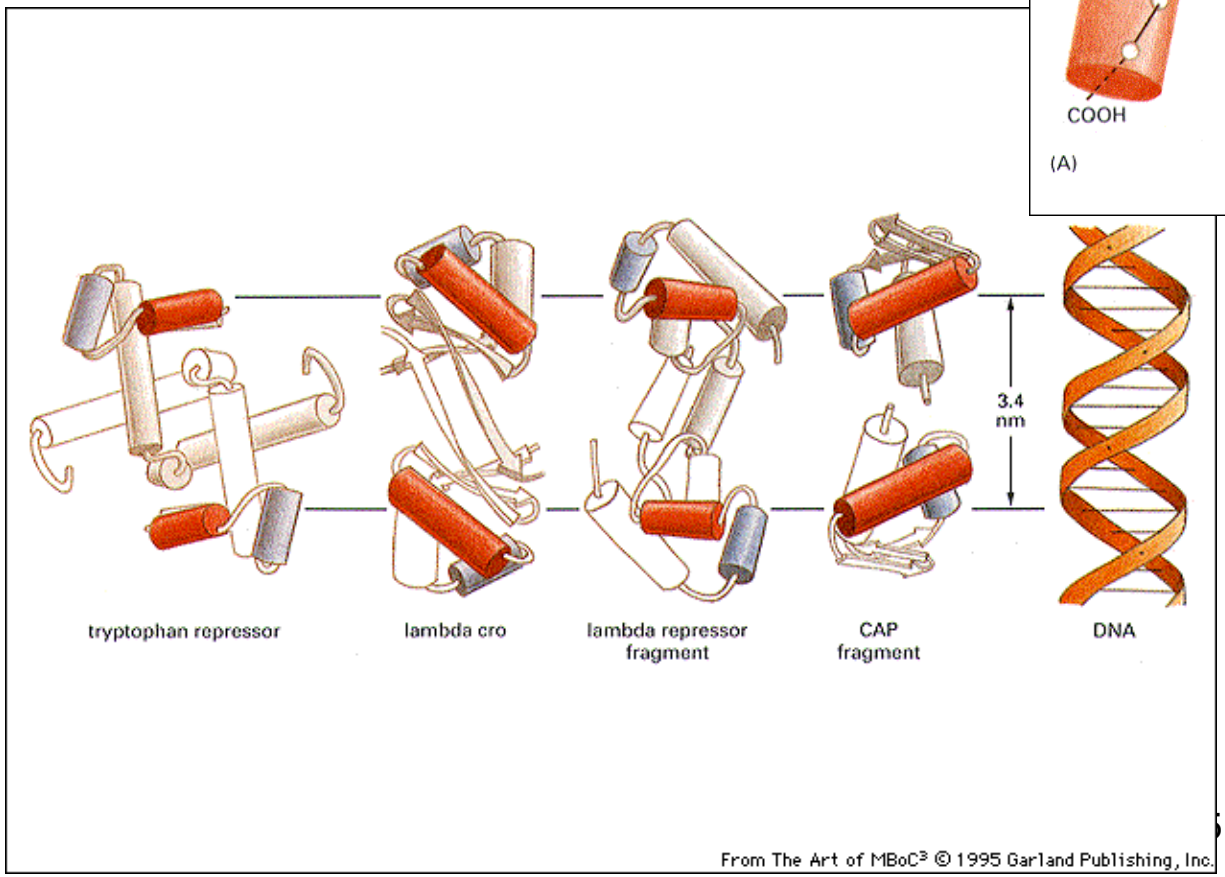


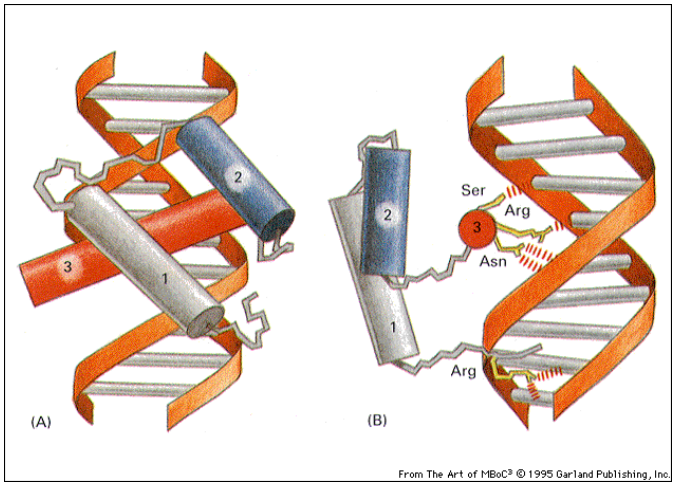
# Protein, NMR



# DNA binding proteins

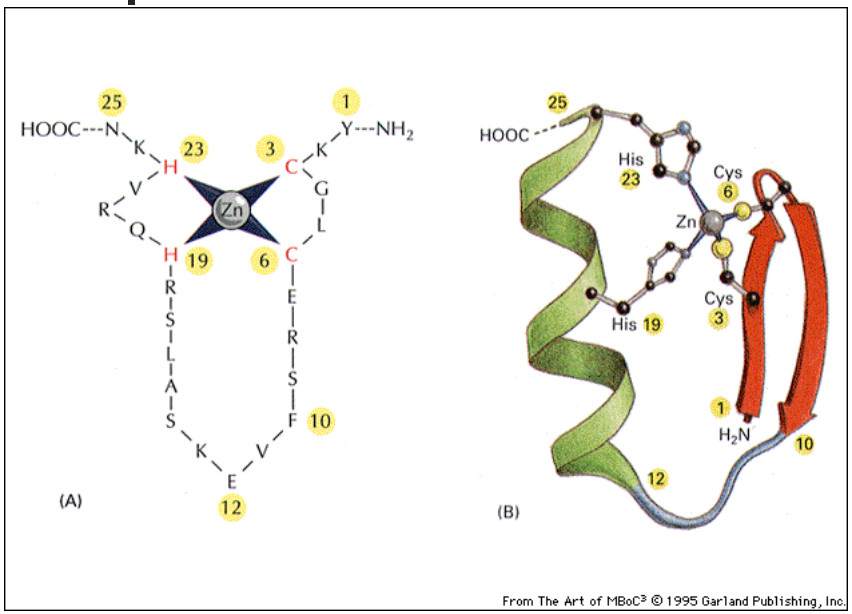


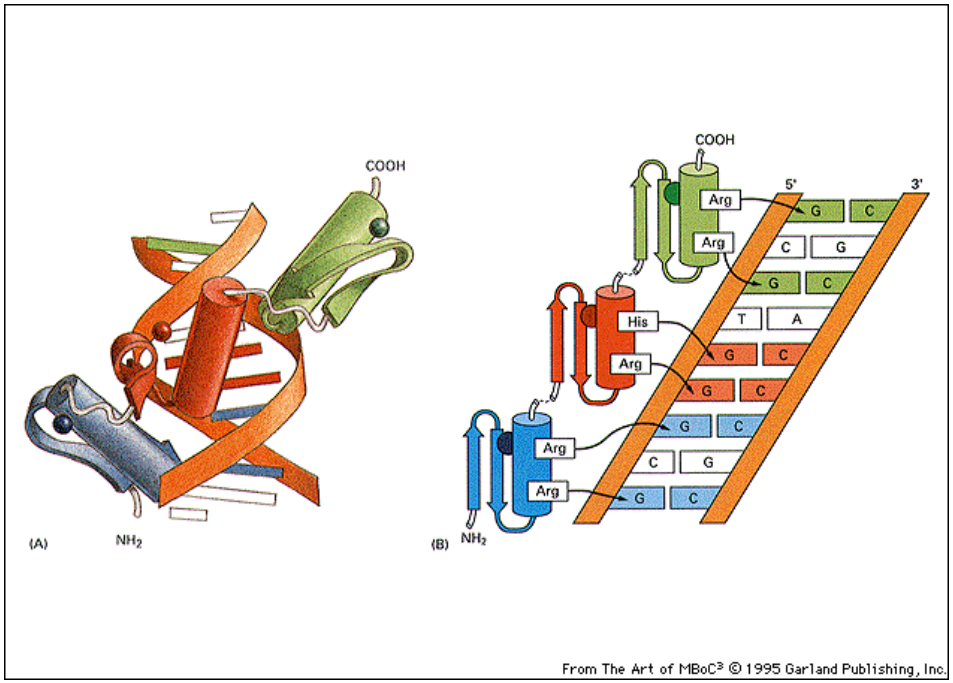



# DNA binding proteins

## zinc fingers



# DNA binding proteins

## other