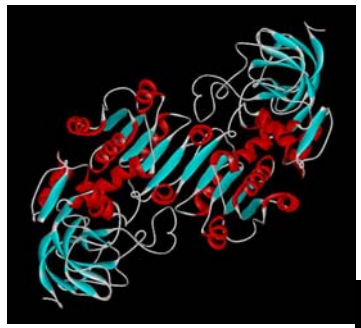
Protein Structure 101

Alexey Onufriev,
Virginia Tech

www.cs.vt.edu/~onufriev

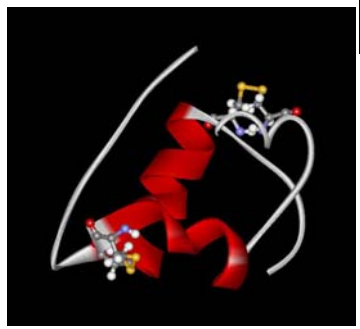
Proteins play key roles in a living system

- Three (out of many many) examples of protein functions
  - Catalysis:
    Almost all chemical reactions in a living cell are catalyzed by protein enzymes.
  - Transport:
    Some proteins transports various substances, such as oxygen, ions, and so on.
  - Information transfer:
    For example, hormones.



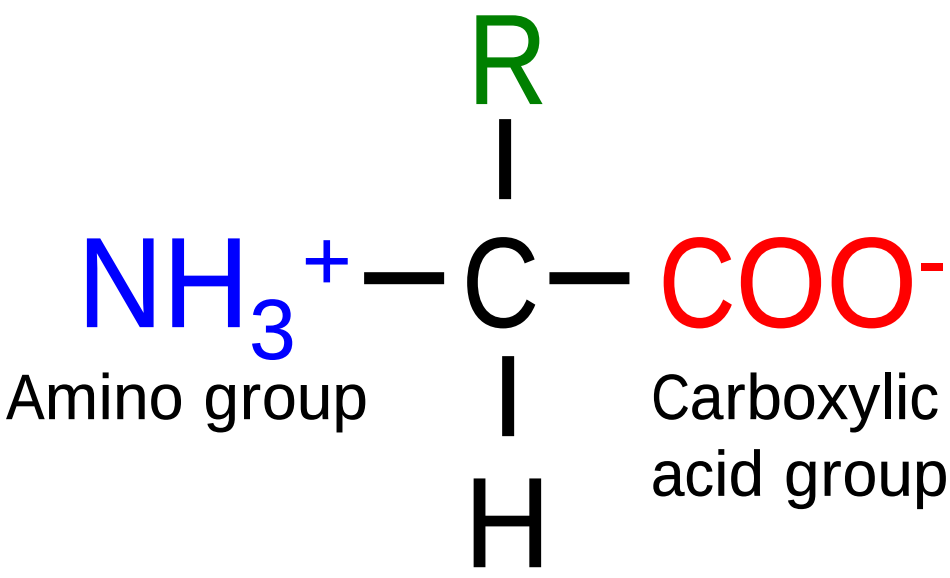
Alcohol dehydrogenase oxidizes alcohols to aldehydes or ketones

Myoglobin stores oxygen



Insulin controls the amount of sugar in the blood

Amino acid: Basic unit of protein



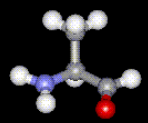
An amino acid

Different side chains, R, determin the properties of 20 amino acids.

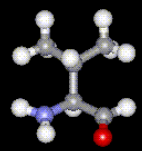
20 Amino acids

Glycine (G)

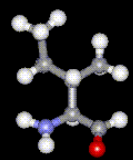
Alanine (A)



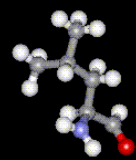
Valine (V)



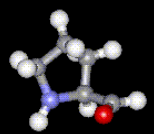
Isoleucine (I)



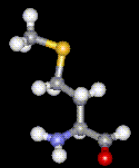
Leucine (L)



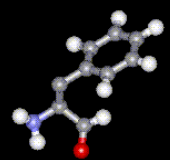
Proline (P)



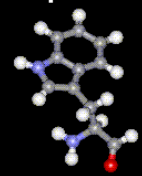
Methionine (M)



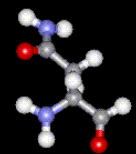
Phenylalanine (F)



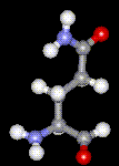
Tryptophan (W)



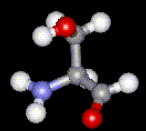
Asparagine (N)



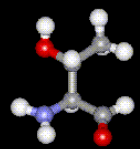
Glutamine (Q)



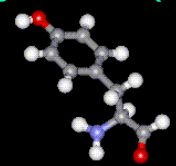
Serine (S)



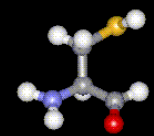
Threonine (T)



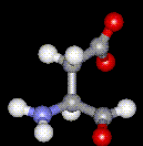
Tyrosine (Y)



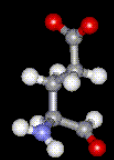
Cysteine (C)



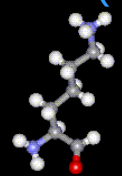
Asparatic acid (D)



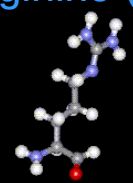
Glutamic acid (E)



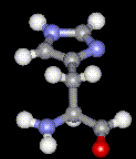
Lysine (K)



Arginine (R)

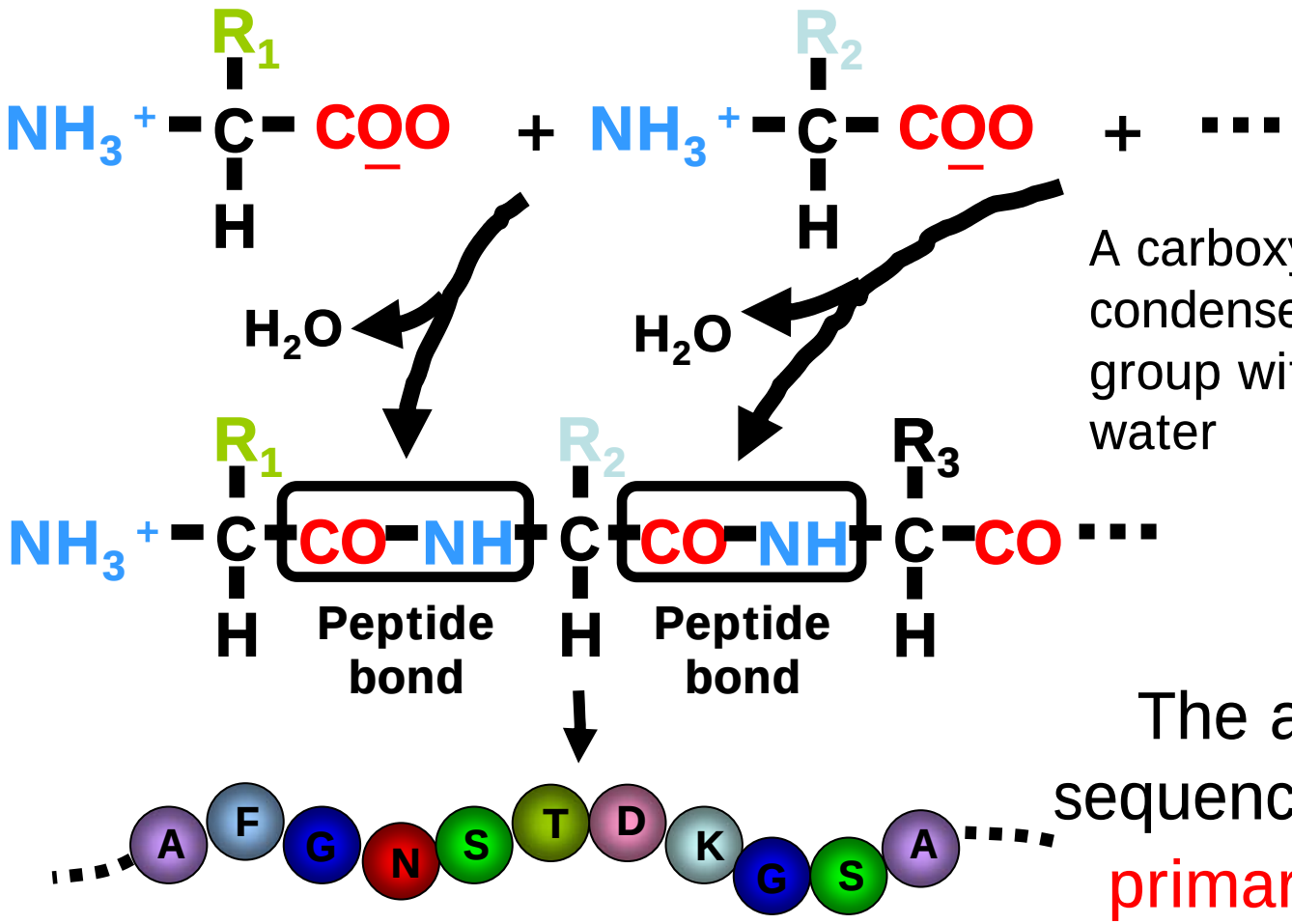


Histidine (H)



White: Hydrophobic, Green: Hydrophilic, Red: Acidic, Blue: Basic

Proteins are linear polymers of amino acids

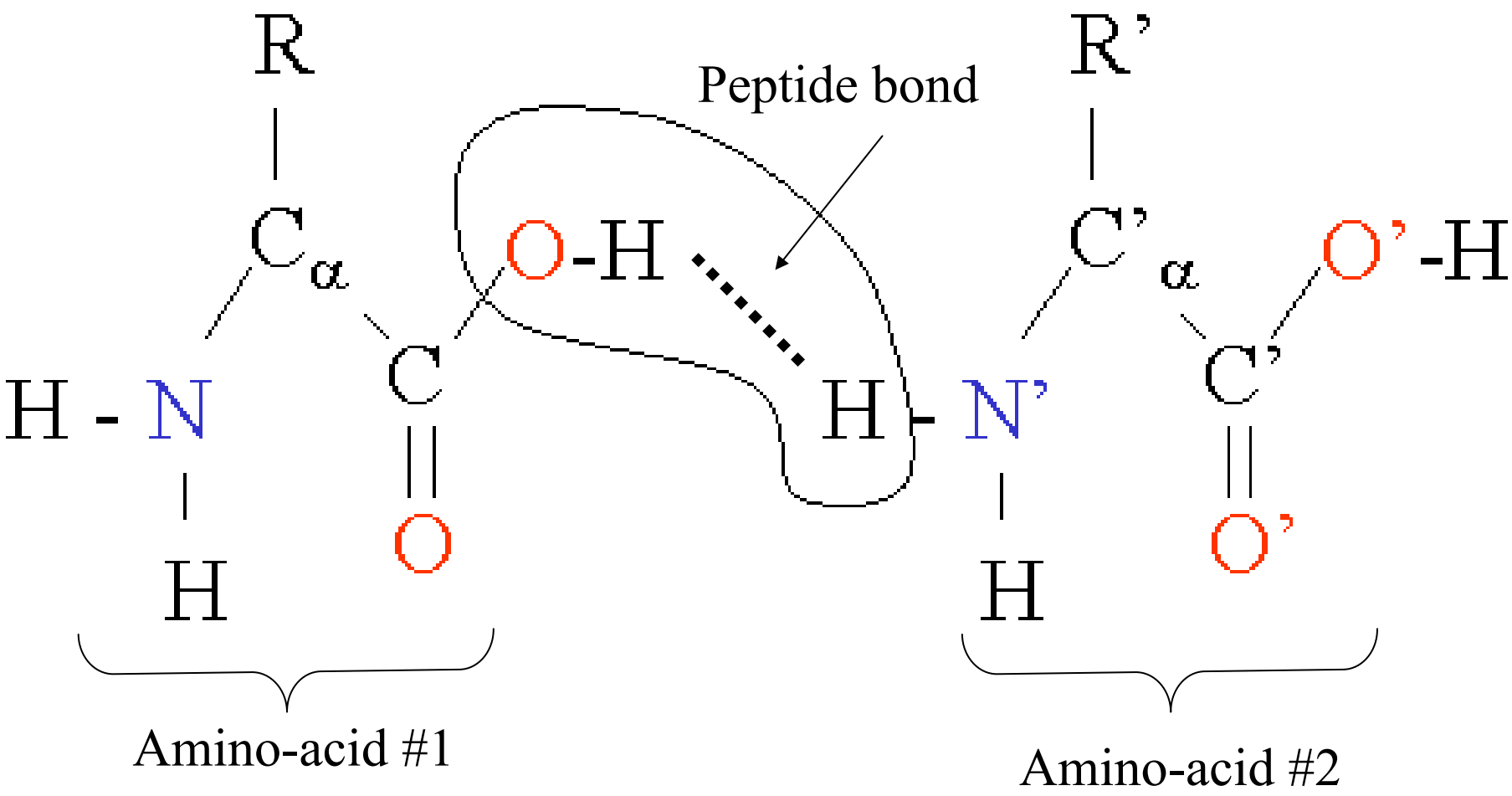


A carboxylic acid condenses with an amino group with the release of a water

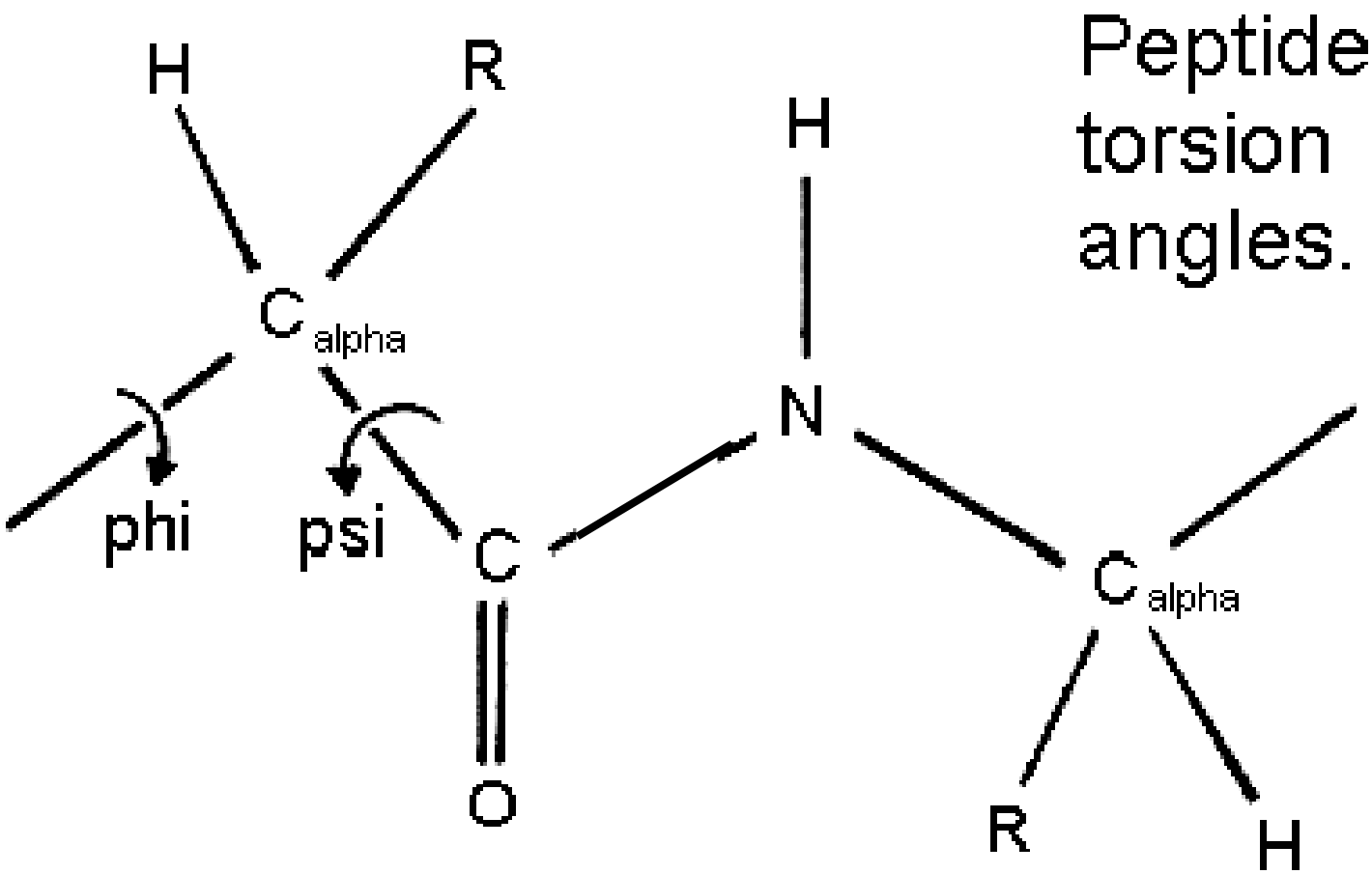
The amino acid sequence is called as primary structure

Protein Structure in 3 steps.

Step 1. Two amino-acids together (di-peptide)

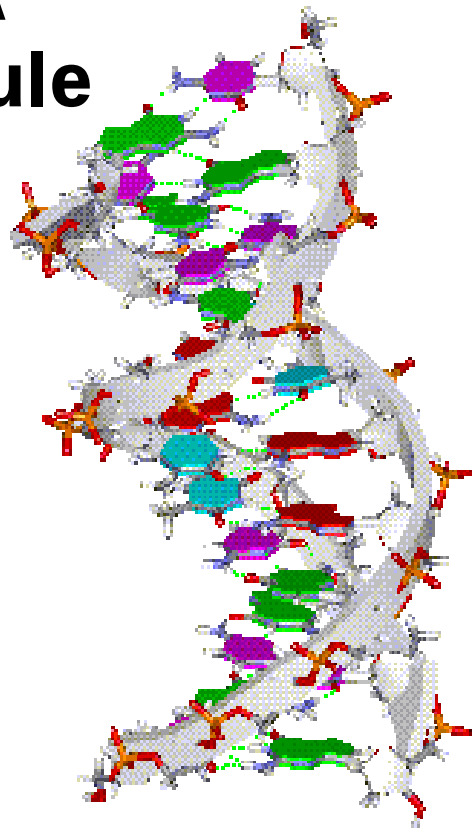


Step 2: Most flexible degrees of freedom:



Amino acid sequence is encoded by DNA base sequence in a gene

DNA molecule

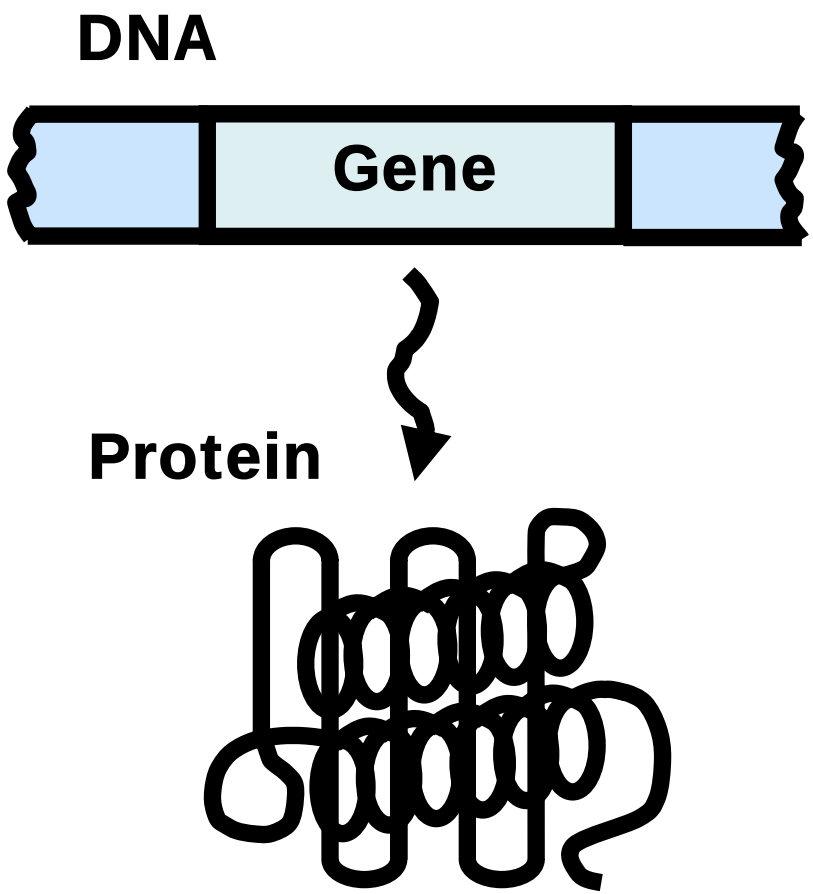


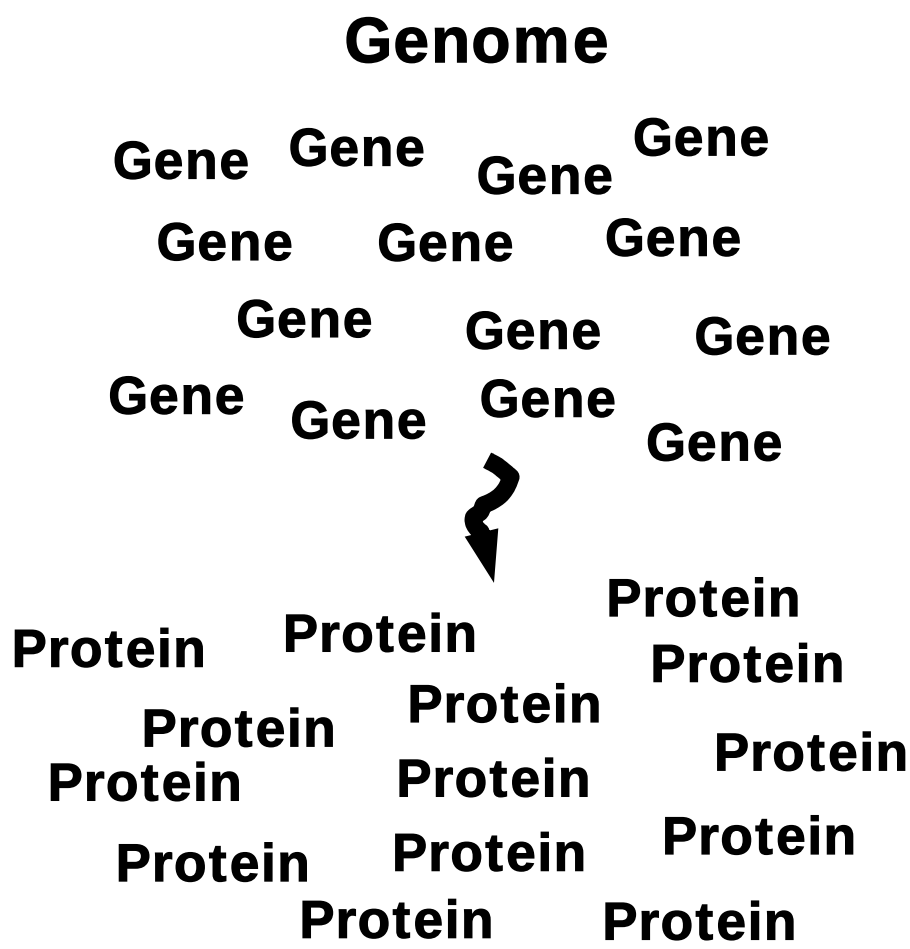
=

DNA base sequence

G ≡ C
C ≡ G
G ≡ C
C ≡ G
T = A
T = A
A = T
A = T
G ≡ C
C ≡ G
G ≡ C
C ≡ G

Gene is protein's blueprint, genome is life's blueprint

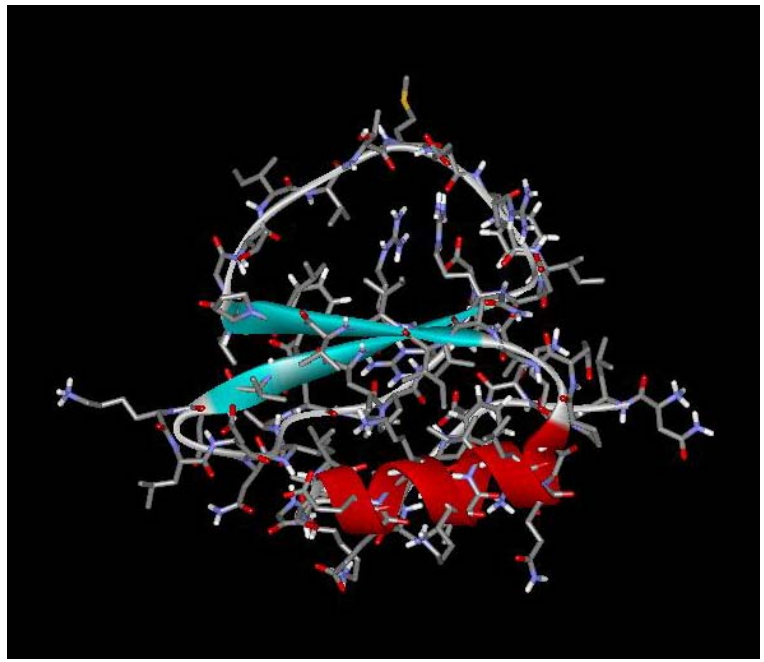




Each Protein has a unique structure

Amino acid sequence

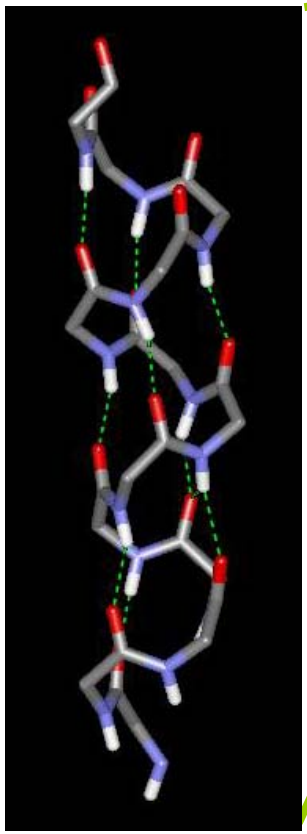
NLKTEWPELVGKSVEE
AKKVILQDKPEAQIIVL
PVGTIVTMEYRIDRVR
LFVDKLDNIAEVPRVG

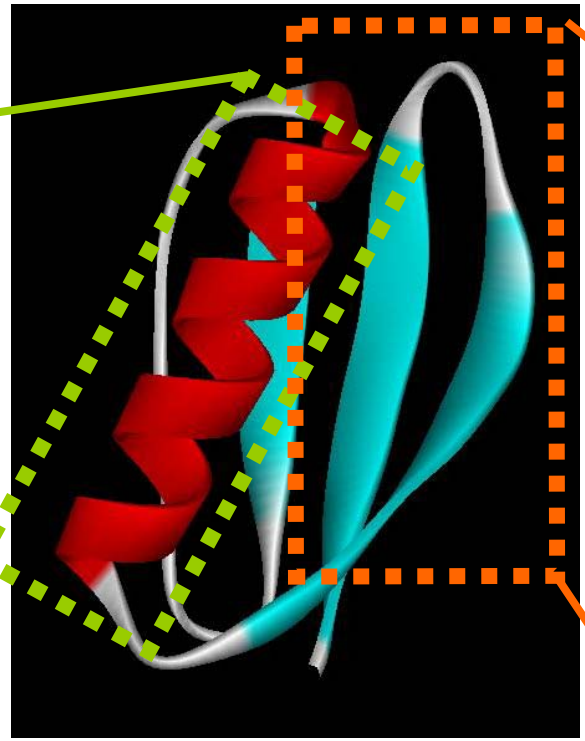


Folding!

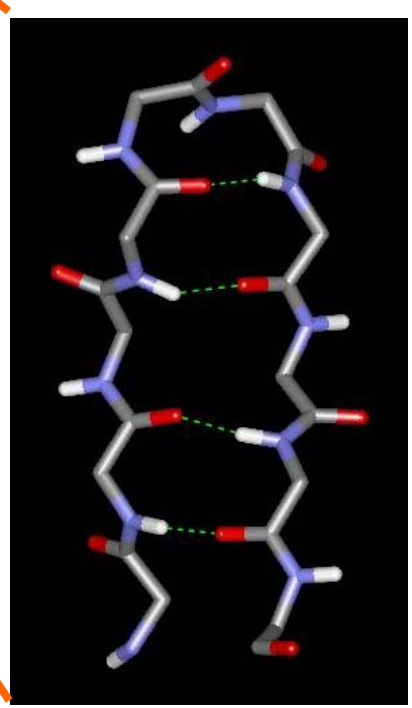
Basic structural units of proteins: Secondary structure

α-helix





β-sheet

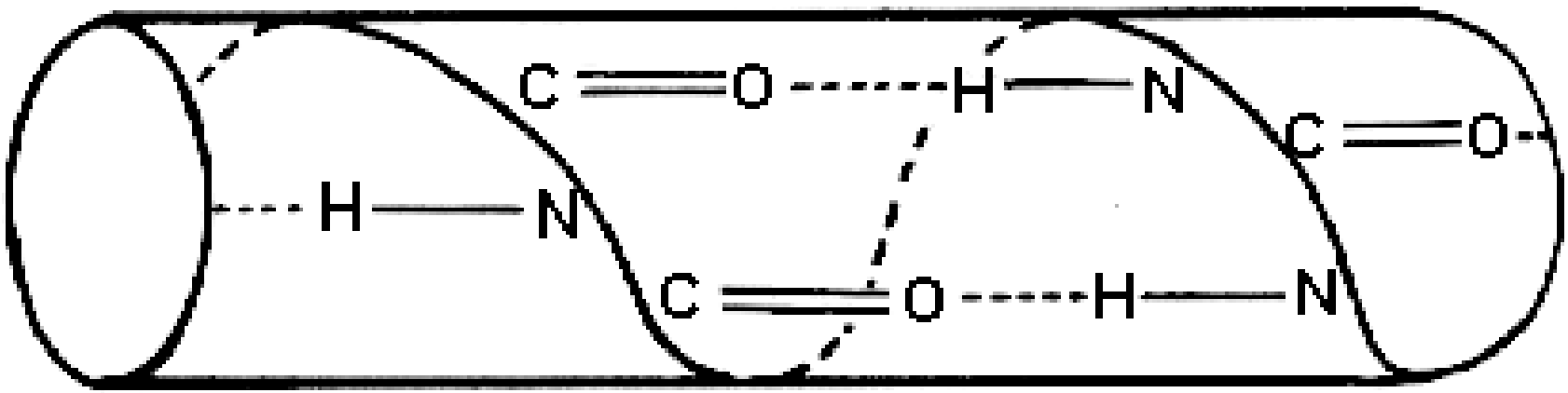


Secondary structures, α-helix and β-sheet, have regular hydrogen-bonding patterns.

Protein Structure in 3 steps.

Sometimes, polypeptide chain forms helical structure:

Toilet roll representation of the main chain hydrogen bonding in an alpha-helix.



Hydrogen Bonding

- Involves three atoms:
  - Donor electronegative atom (**D**)
    (Nitrogen or Oxygen in proteins)
  - Hydrogen bound to donor (**H**)
  - Acceptor electronegative atom (**A**) in close proximity

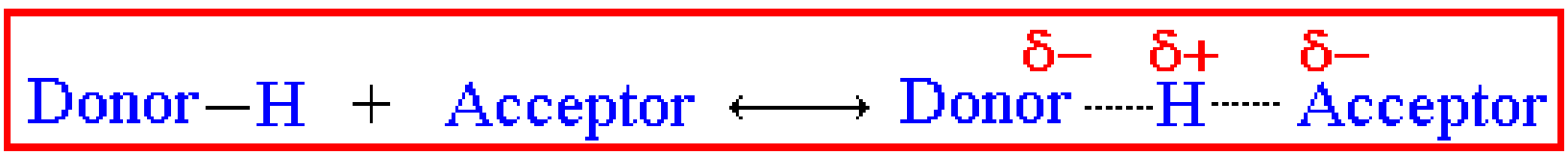
D – H A

D-H Interaction

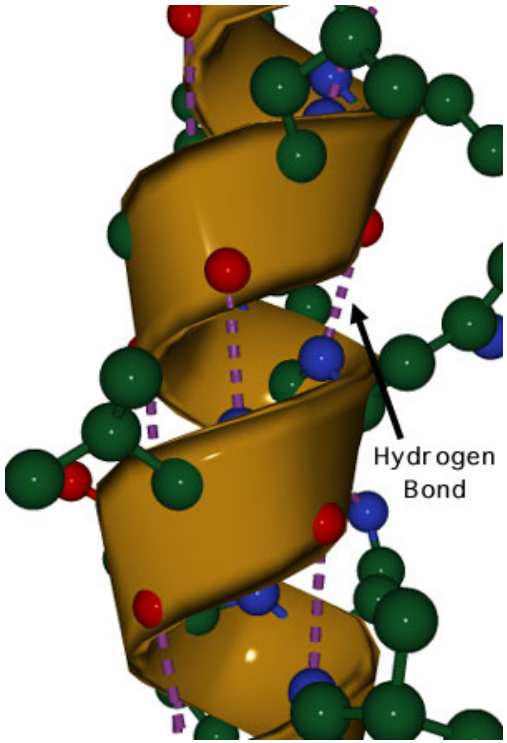
- Polarization due to electron withdrawal from the hydrogen to D giving D partial negative charge and the H a partial positive charge
- Proximity of the Acceptor A causes further charge separation

$$\overset{\delta-}{D}-\overset{\delta+}{H}\quad\overset{\delta-}{A}$$

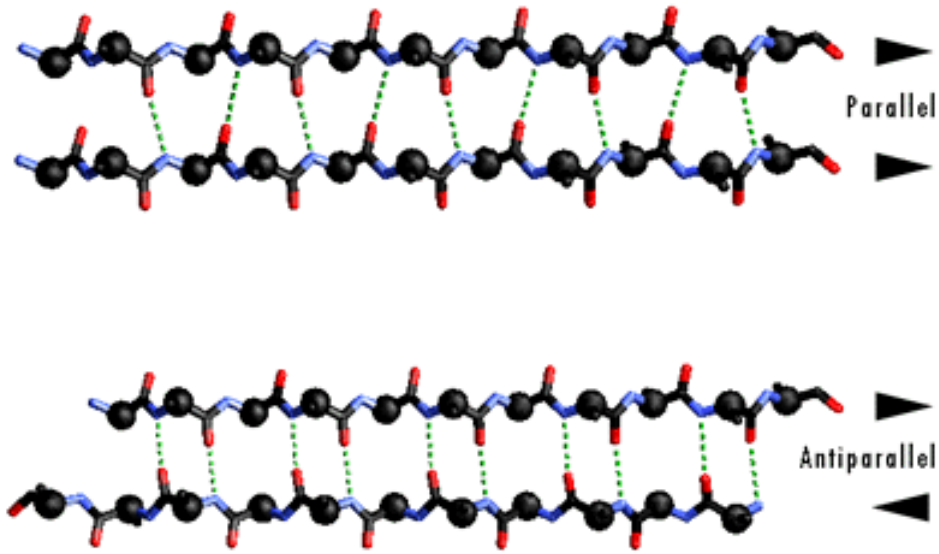
- Result:
  - Closer approach of A to H
  - Higher interaction energy than a simple van der Waals interaction



Hydrogen Bonding And Secondary Structure

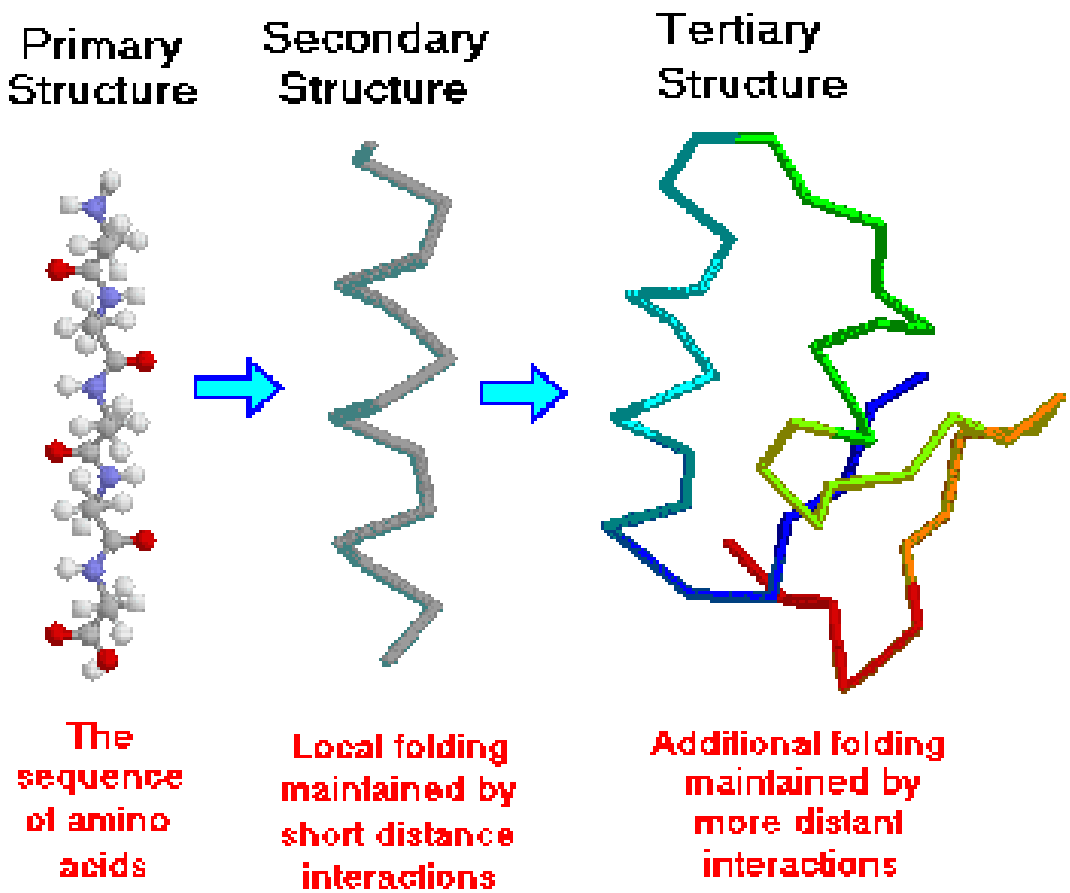


alpha-helix

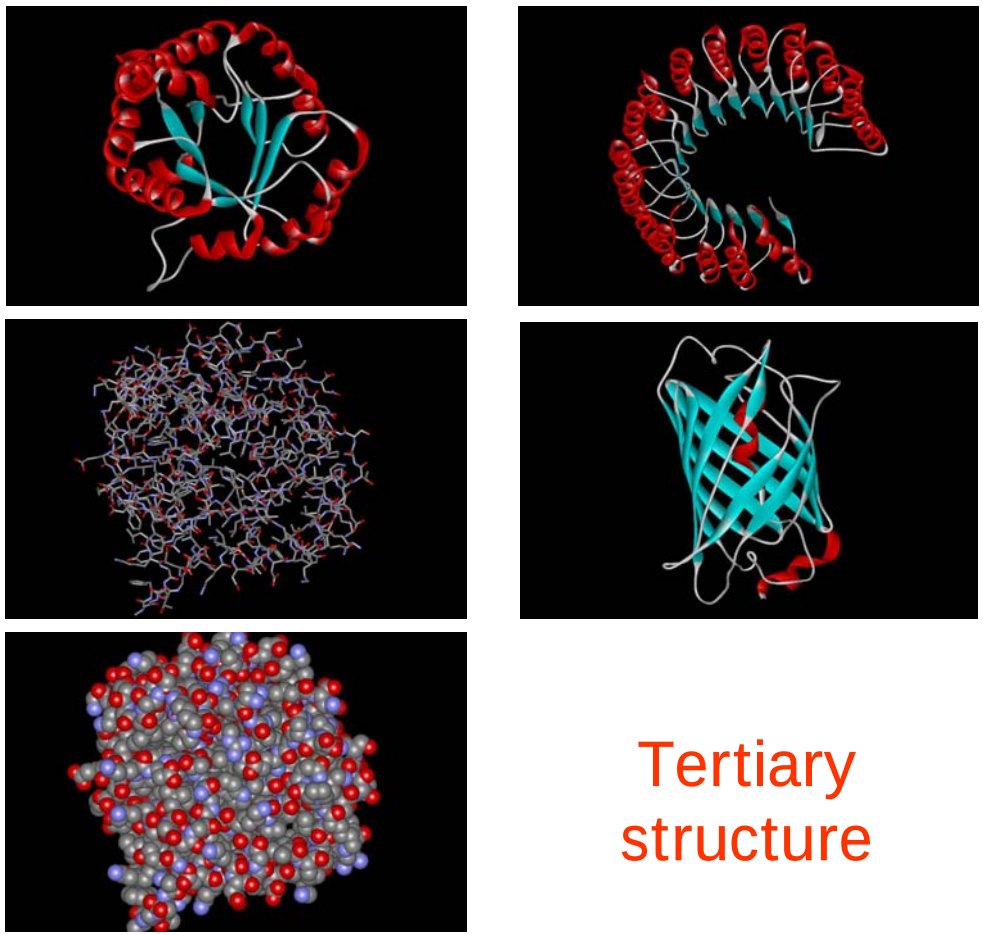


beta-sheet

Protein Structure



Three-dimensional structure of proteins



Tertiary structure

Quaternary structure

Hierarchical nature of protein structure

Primary structure (Amino acid sequence)

↓

Secondary structure (α-helix, β-sheet)

↓

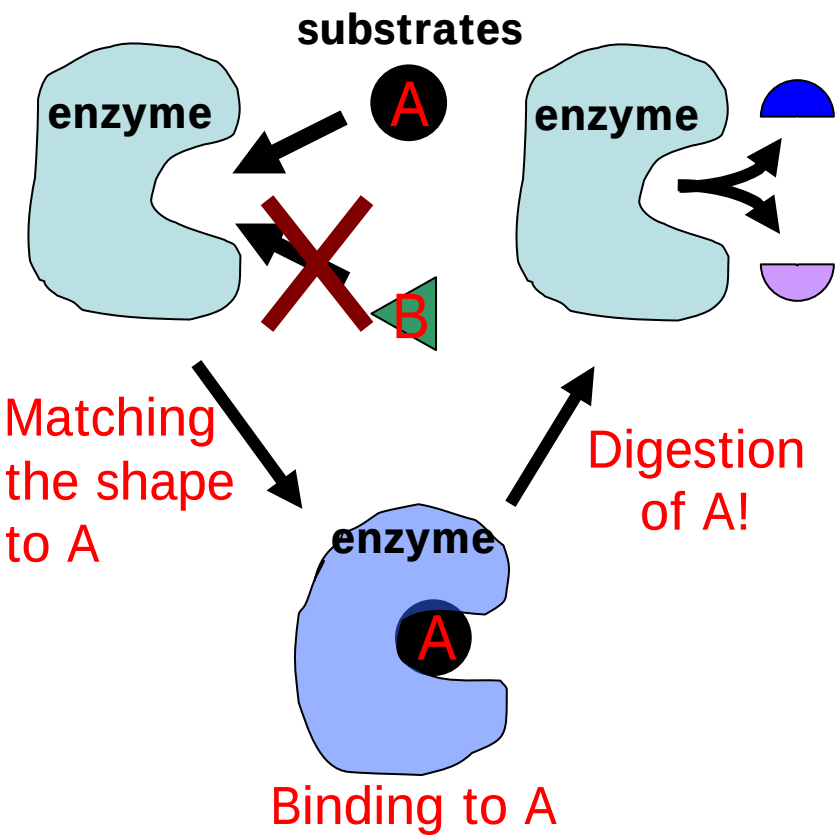
Tertiary structure (Three-dimensional structure formed by assembly of secondary structures)

↓

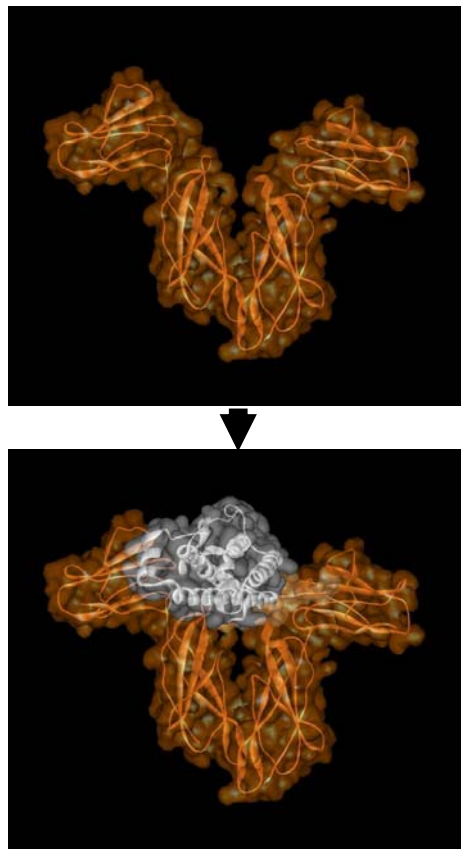
Quaternary structure (Structure formed by more than one polypeptide chains)

Close relationship between protein structure and its function

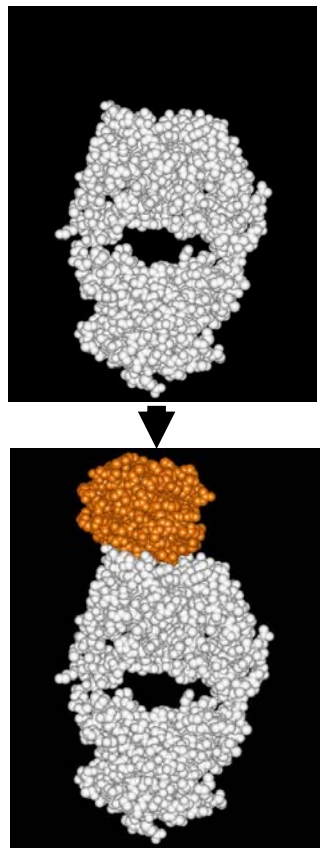
Example of enzyme reaction



Hormone receptor



Antibody



Protein structure prediction has remained elusive over half a century

"Can we predict a protein structure from its amino acid sequence?"

Still virtually impossible at atomic level accuracy (but there are some notable exceptions). Possible in some cases if a rougher structure is acceptable.

So where do we get the high quality protein structures to work with?

- THE PDB (Protein Data Bank. ~30,000 structurs)
- PDB

Summary

- Proteins are key players in our living systems.
- Proteins are polymers consisting of 20 kinds of amino acids.
- Each protein folds into a unique three-dimensional structure defined by its amino acid sequence.
- Protein structure has a hierarchical nature.
- Protein structure is closely related to its function.