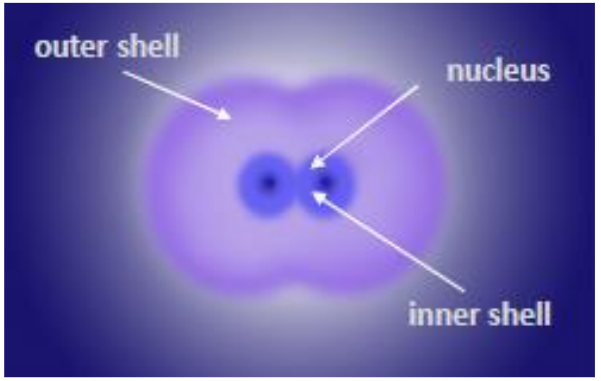
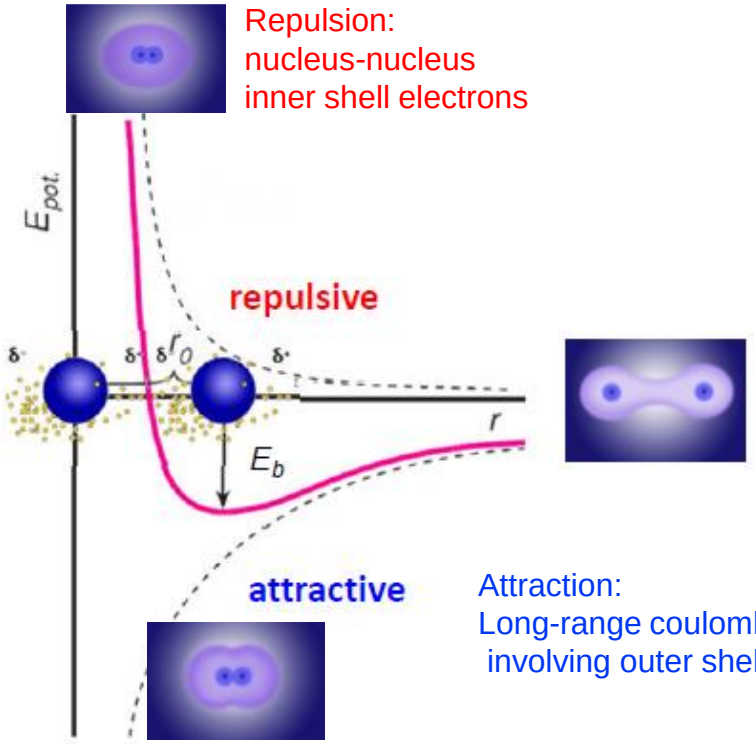
Structural hierarchy of proteins

- G.Schay
- Semmelweis University
- Dept. Biophysics and Radiation Biology

Proteins are LARGE structures consisting thousands of atoms.

The atomic interaction rules apply here too.





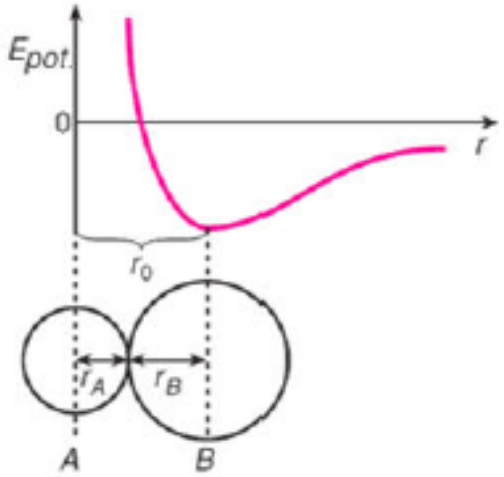
Attraction:
Long-range coulomb interaction
involving outer shell electrons

$$E_{pot} = \frac{-A}{r^n} + \frac{B}{r^m}$$

Power law

Example: 6-12 potential

Van der Waals radius: r_0



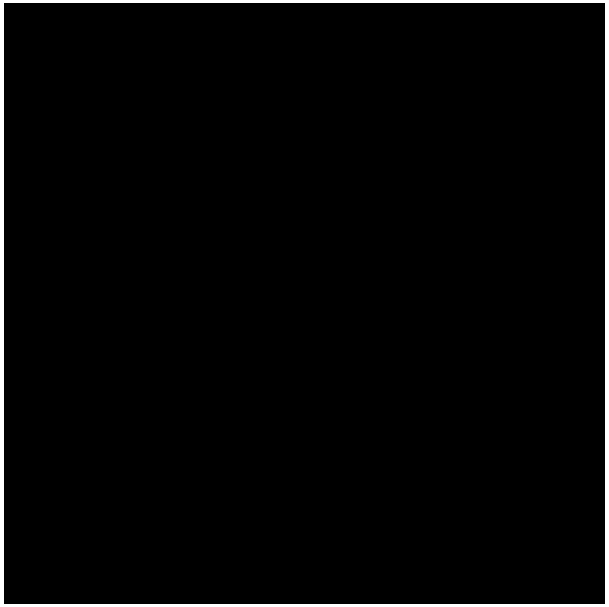
Primary bonds: E_b>1eV

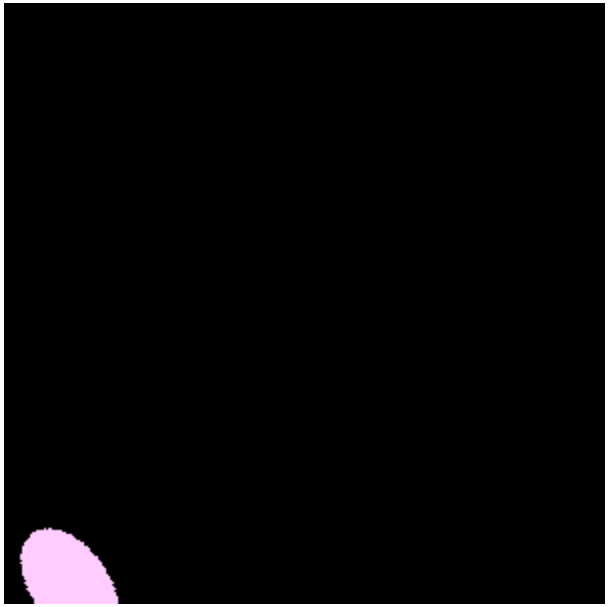
- **covalent:** common electron state around the participating nuclei
- **metallic bond:**multi-atomic system
- **ionic bond:** Coulomb-forces between ions

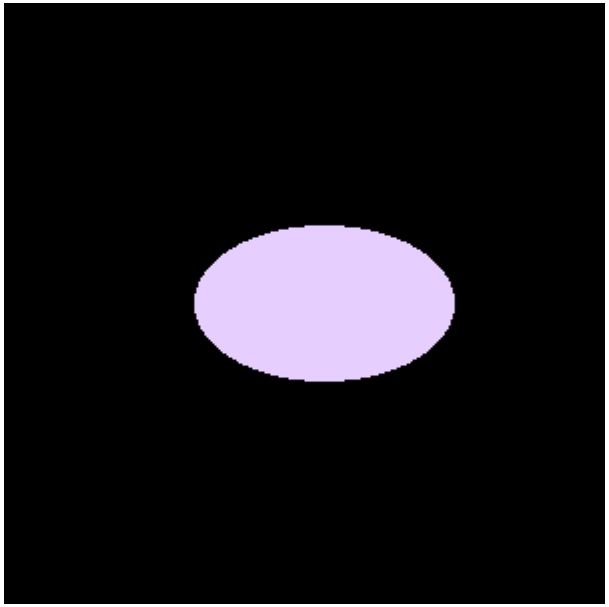
Secondary bonds: **inter**molecular bonds, weaker than **intra**molecular

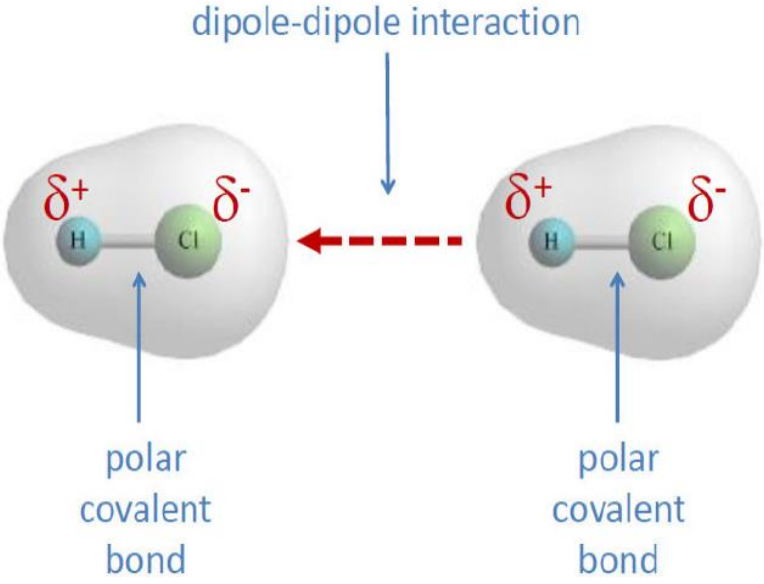
- **Van der Waals:** between two apolar atoms (without permanent dipole moment) where a temporarily created dipole interacts with an apolar molecule or atom thus converting it into a dipole **(induced dipole)**
- Van der Waals radius of atoms can be used (-> “spacefill models, Solvent Accessible Surface Area”)
- Weak: (E_b ~ 0,02 eV)

1eV=1.9·10^{-19} Joule

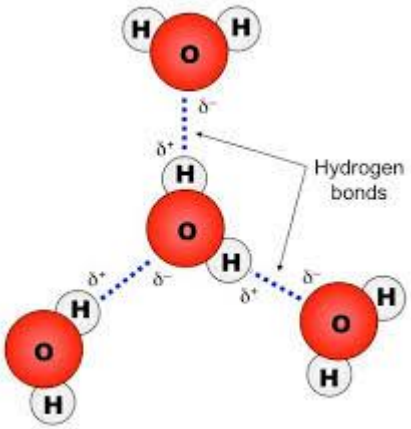


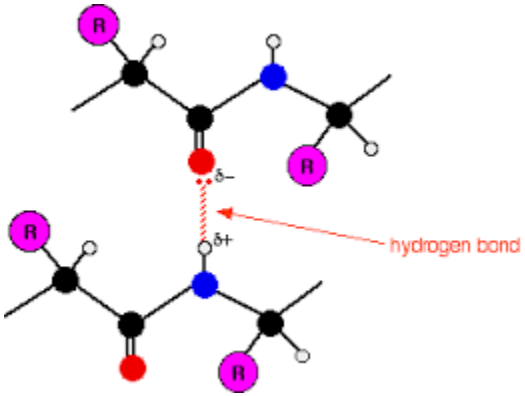






$E_{binding}$ ~ 0.003 – 0.02 eV





$E_{binding}$ ~ 0.2 eV

Classification of proteins

According to the biological function

- Enzymes (E.g.: trypsin, cytochrome-c…)
- Transport proteins (E.g.: hemoglobin...)
- Chaperone proteins
- Toxins (E.g.: ricin, snake venom…)
- Hormones (E.g.: insulin, growth hormone…)
- Contractile proteins (E.g.: myosin, actin…)
- Structure proteins (E.g.: collagen, elastin …)
- Storage proteins (E.g.: ovalbumin, casein, ferritin…)
- Others (E.g.: histone proteins…)

Classification of proteins

According to shape and solubility

- Fibrous proteins (E.g.: collagen…)

  These proteins have a rod like structure. They are not soluble in water.

- Globular proteins (E.g.: hemoglobin, myoglobin...)

  These proteins more or less spherical in nature. Due to their distribution of amino acids (hydrophobic inside, hydrophilic outside) they are very soluble in aqueous solution.

- Membrane proteins (E.g.: rhodopsin…)

  These are protein which are in association with lipid membranes. Those membrane proteins that are embedded in the lipid bilayer have extensive hydrophobic amino acids that interact with the non-polar environment of the bilayer interior. Membrane proteins are not soluble in aqueous solution. Rhodopsin is an integral membrane protein and is embedded in the bilayer.

Classification of proteins

According to secondary structure

- All Alpha helical (E.g.: myoglobin…)

  Proteins that contain only (in some exceptional instances there may be isolated beta-sheets) alpha helical secondary structure.

- Alpha/beta (E.g.: Triose phosphate-isomerase…)

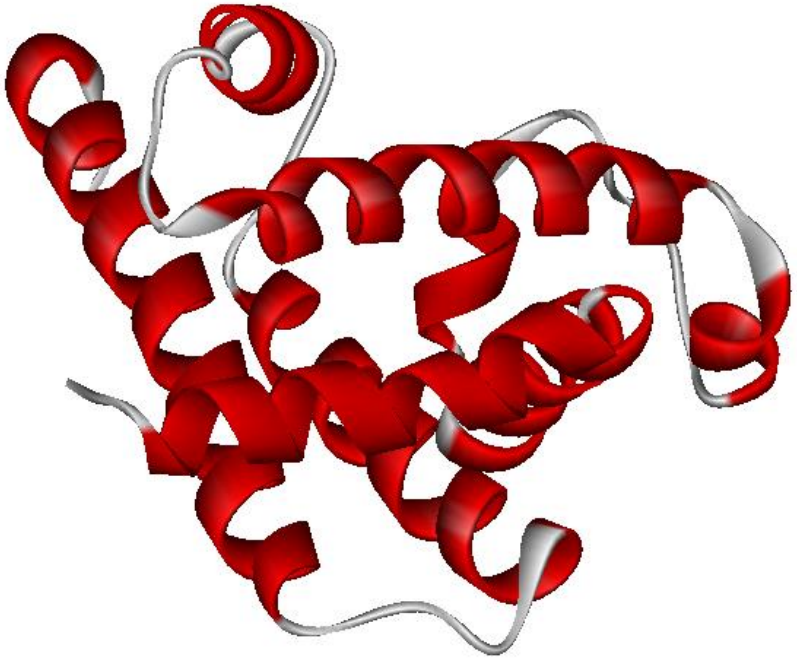
  Proteins that contain alternating alpha-helical and beta-sheet secondary structure elements. The alternating helical and sheet segments may form a closed barrel like structure.

- Alpha+beta (E.g.: ribonuclease…)

  In these proteins the alpha helical and beta sheet regions occur in independent regions of the molecule.

- Beta sheet (E.g.: tenascin…)

  Protein that contain only (in some exceptional instances there may be isolated alpha helices) beta-sheet secondary structure.

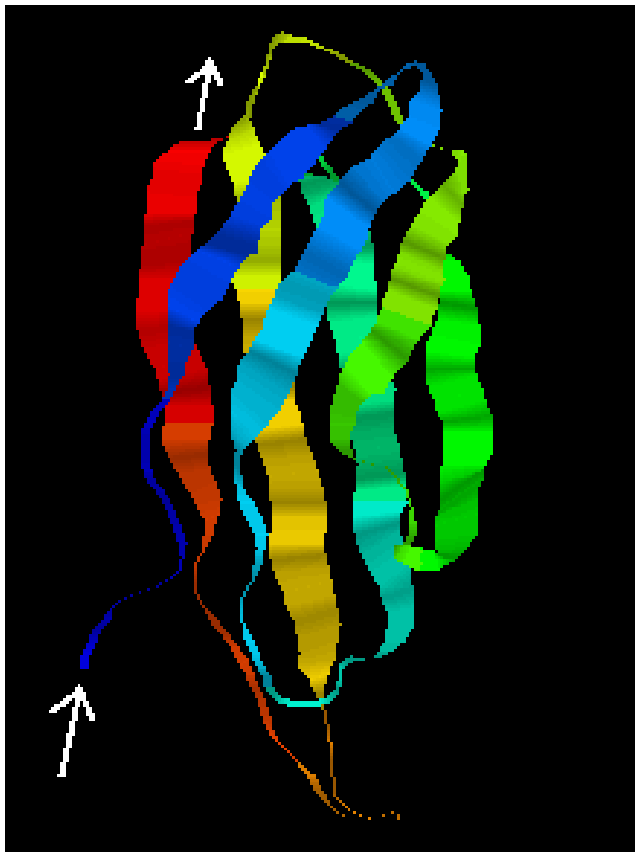


Myoglobin

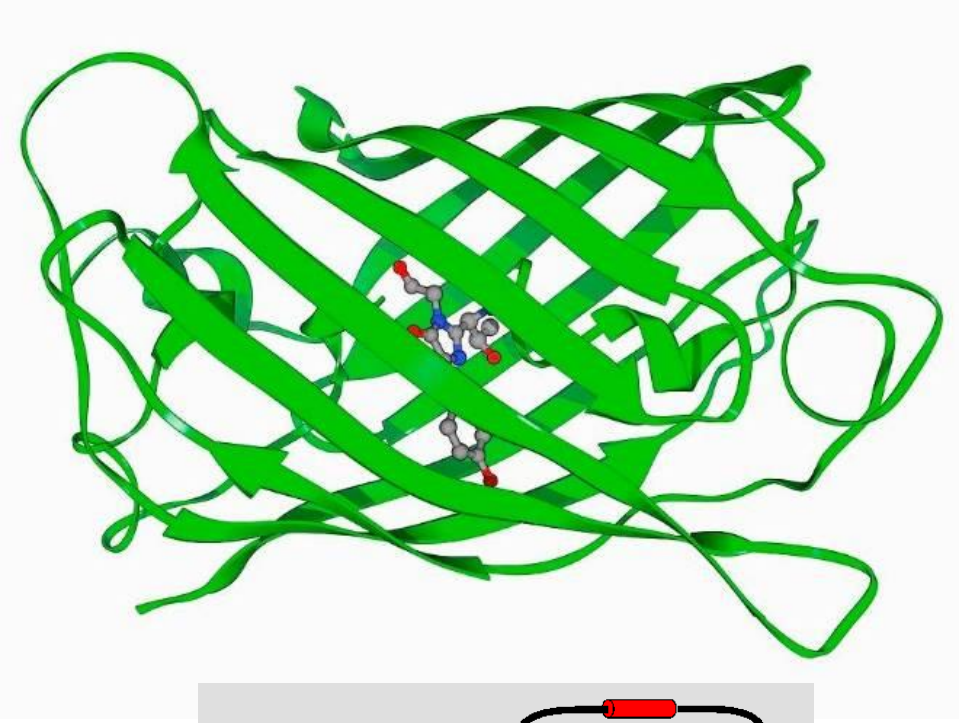
All alpha

Tenascin

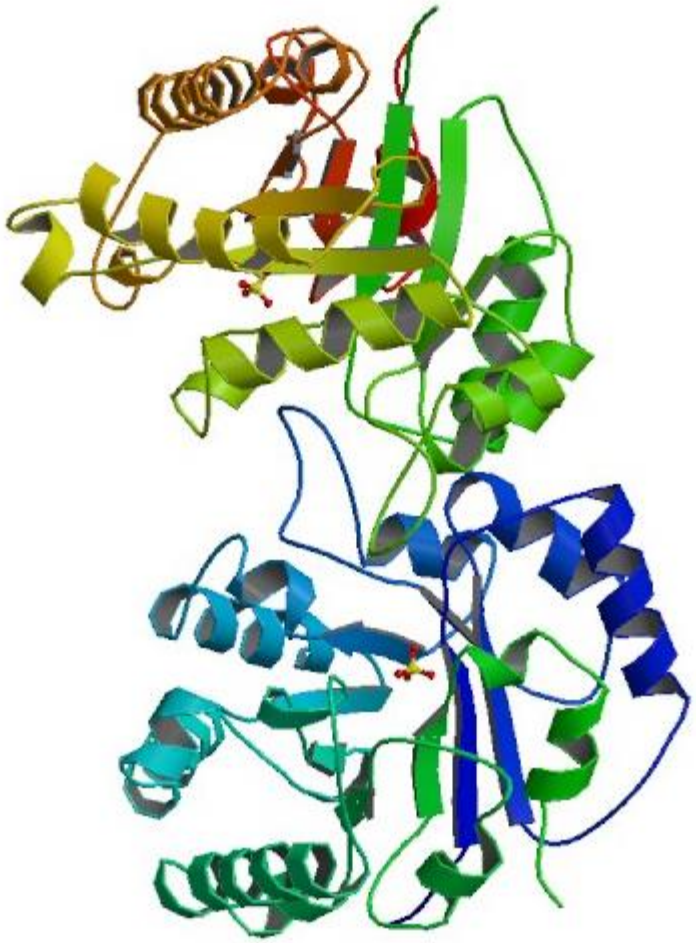
Beta

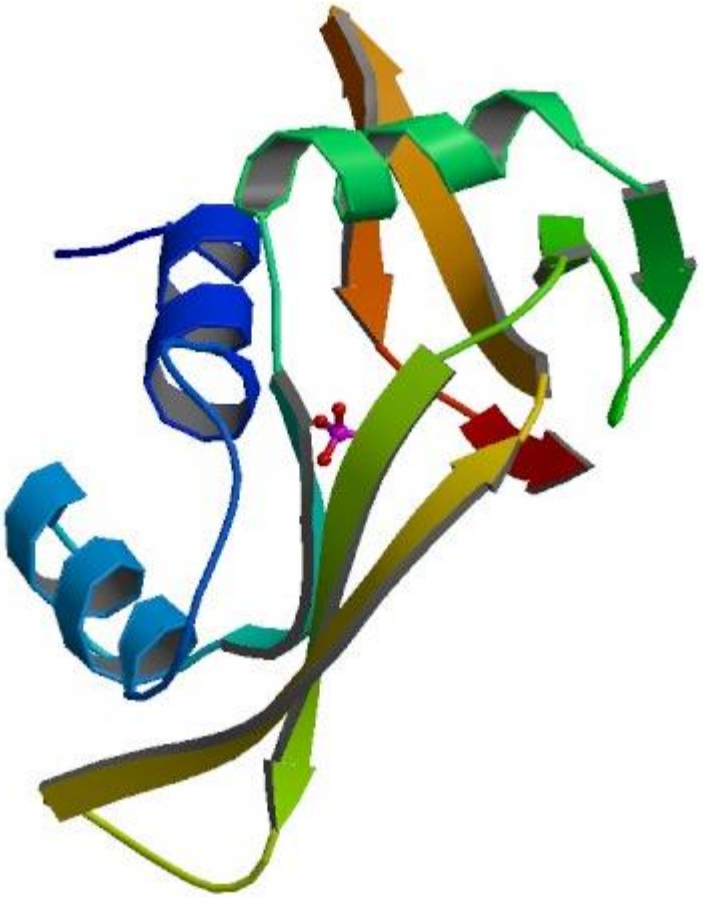


Beta sandwich



Beta barrel (GFP)





Alpha/beta structure (Triose phosphate isomerase)

Alpha+beta (ribonuclease)

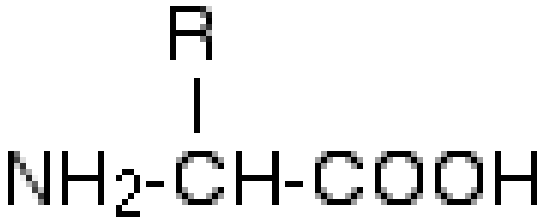
Structural hierarchy

- Primary structure
- Secondary structure
- Tertiary structure
- Quaternary structure

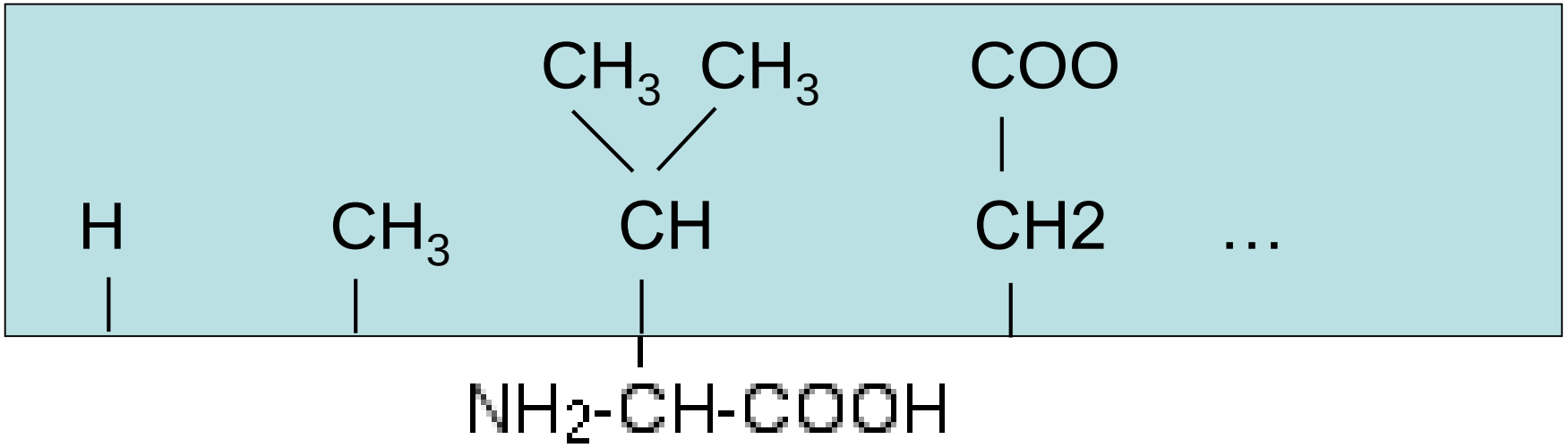
Supramolecular organizations

Amino acids: the building blocks of proteins

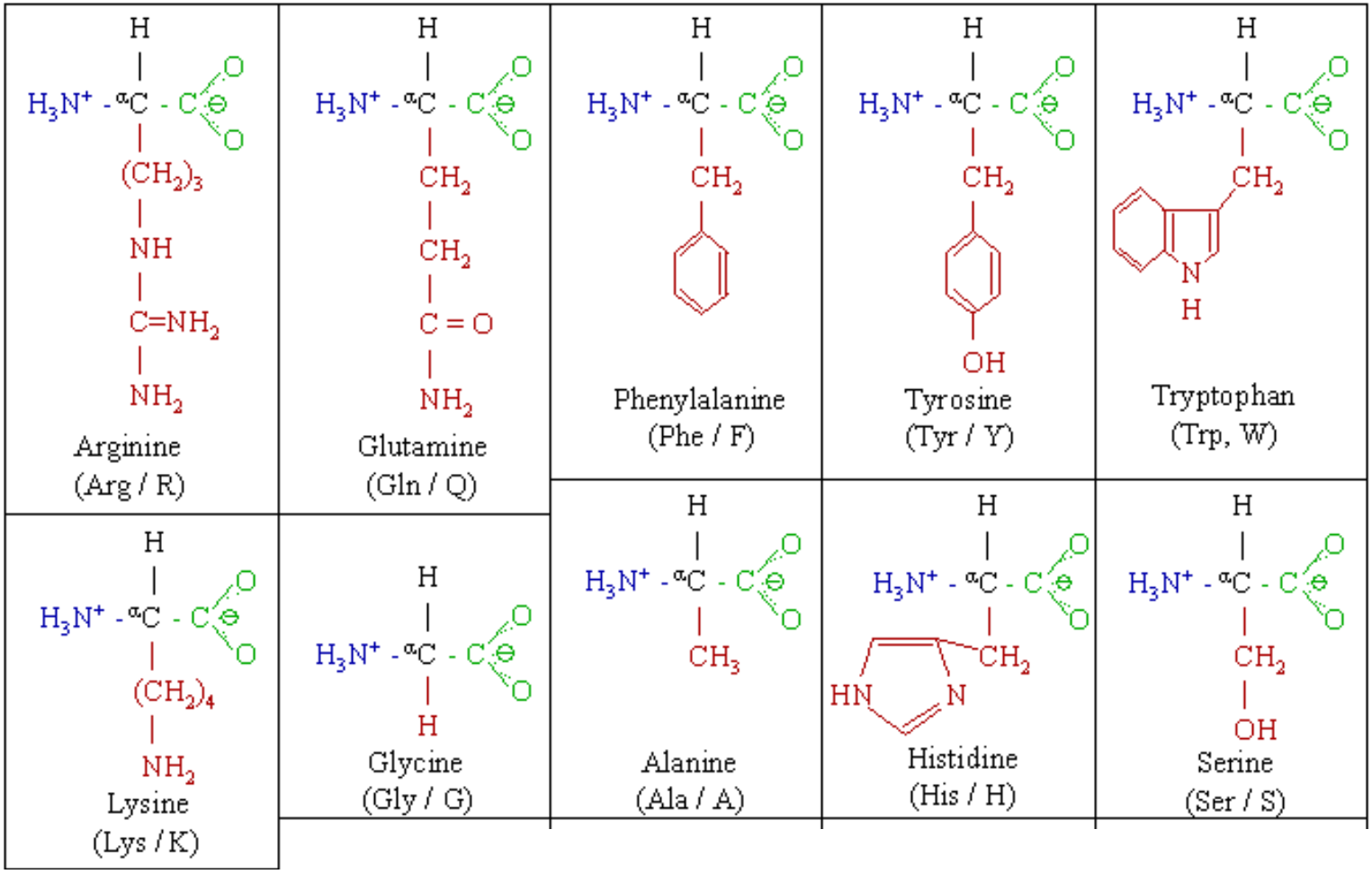
General structure of amino acids:



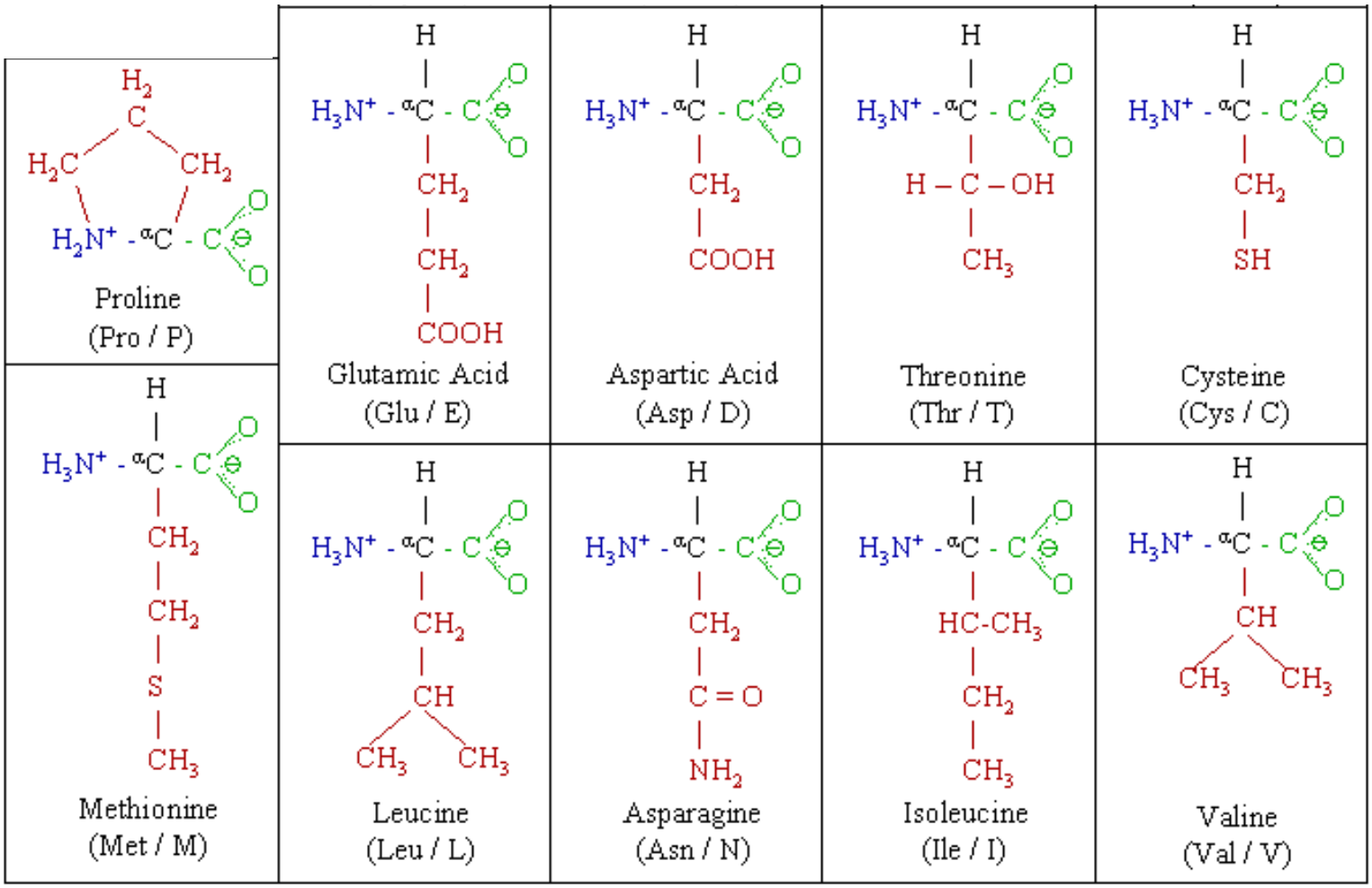
Structural variability:



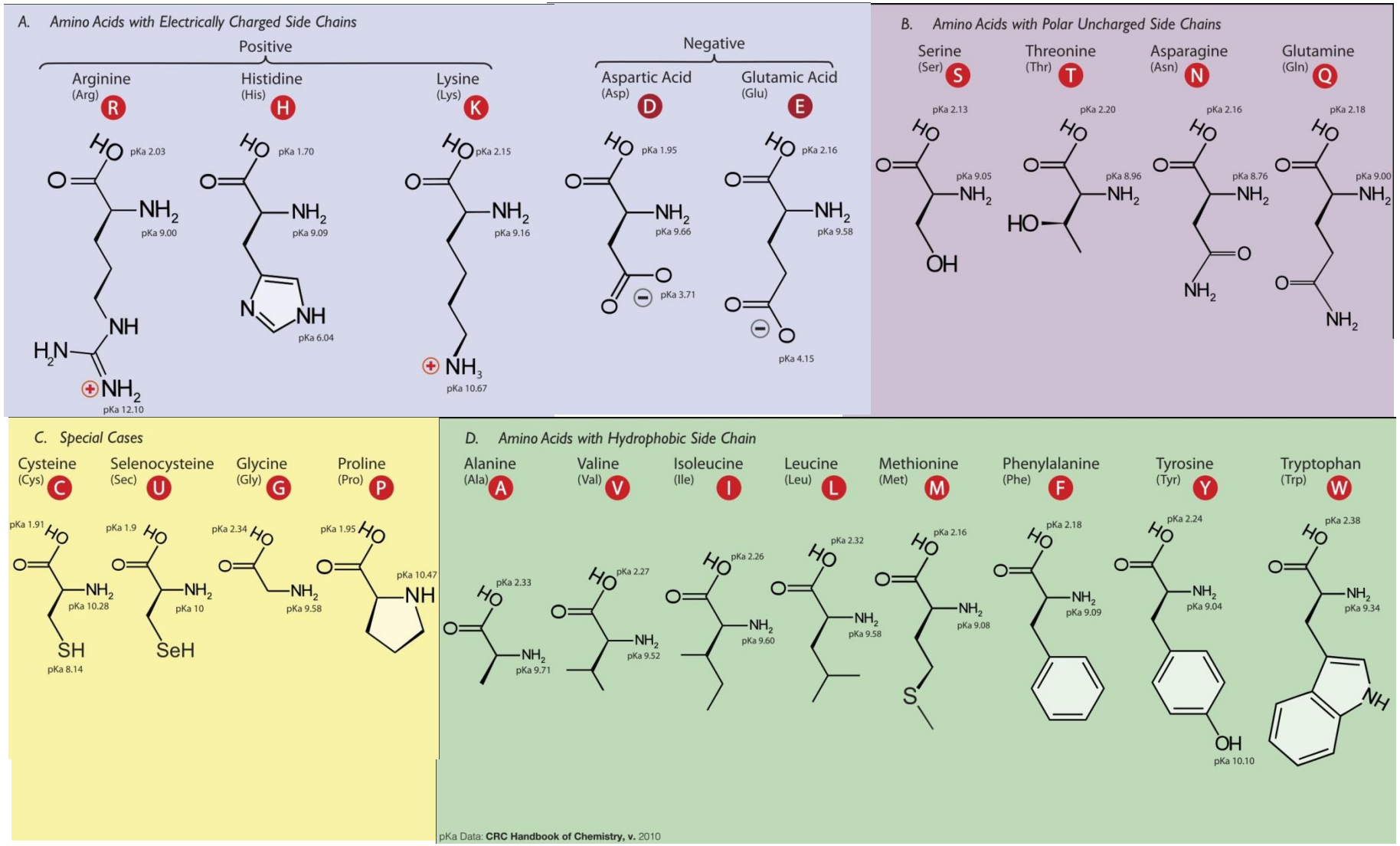
Amino acids of the proteins



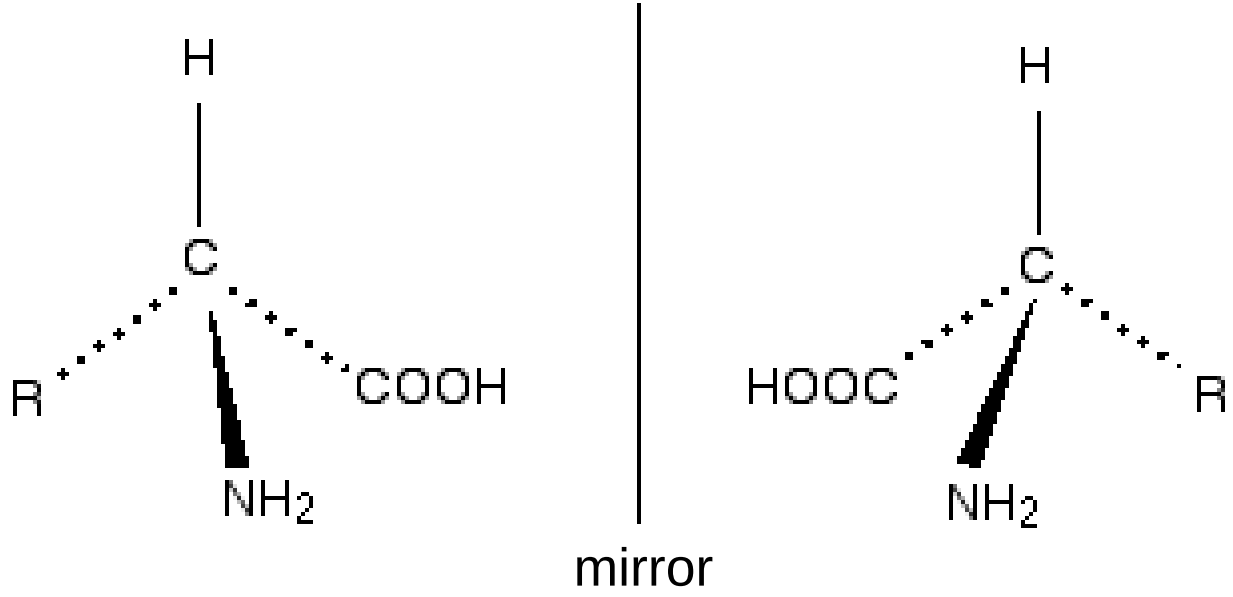
Amino acids of the proteins



Properties of amino acids



Chirality

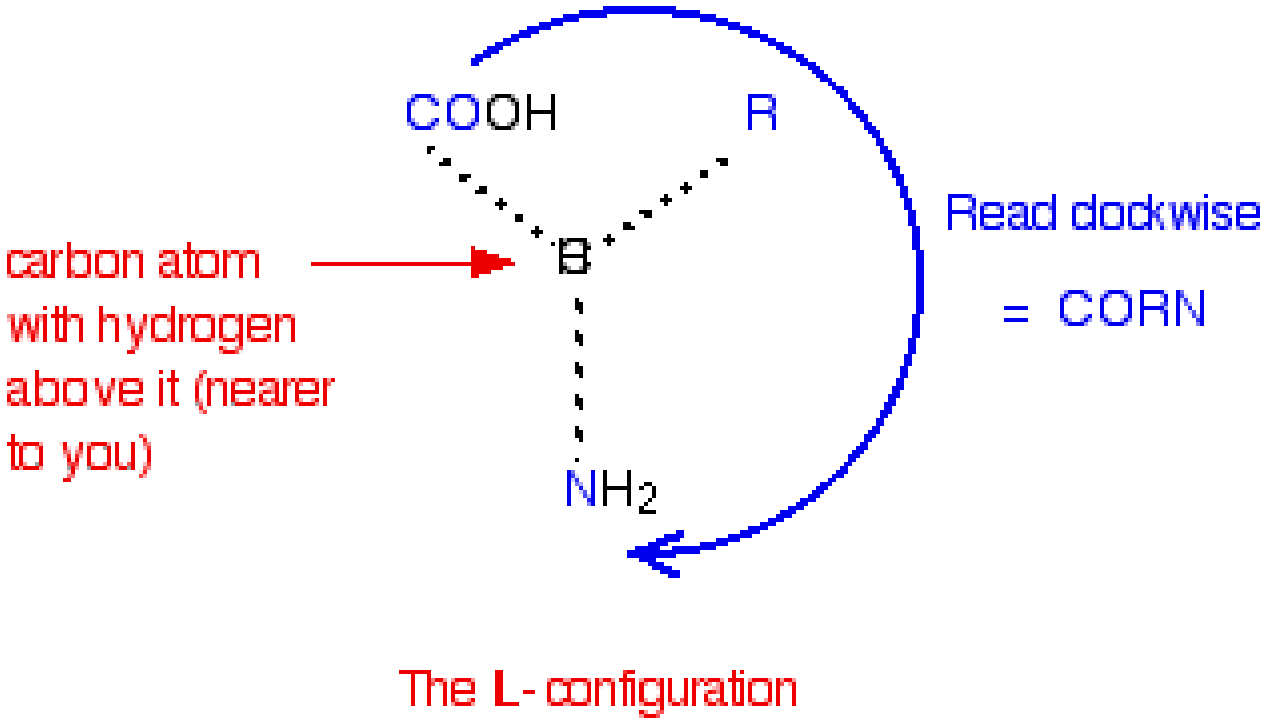


Chirality center:
a carbon atom, connected to four different atoms.

⇨ Optical activity
(rotation of the plane of polarization of linearly polarized light)

Hand:

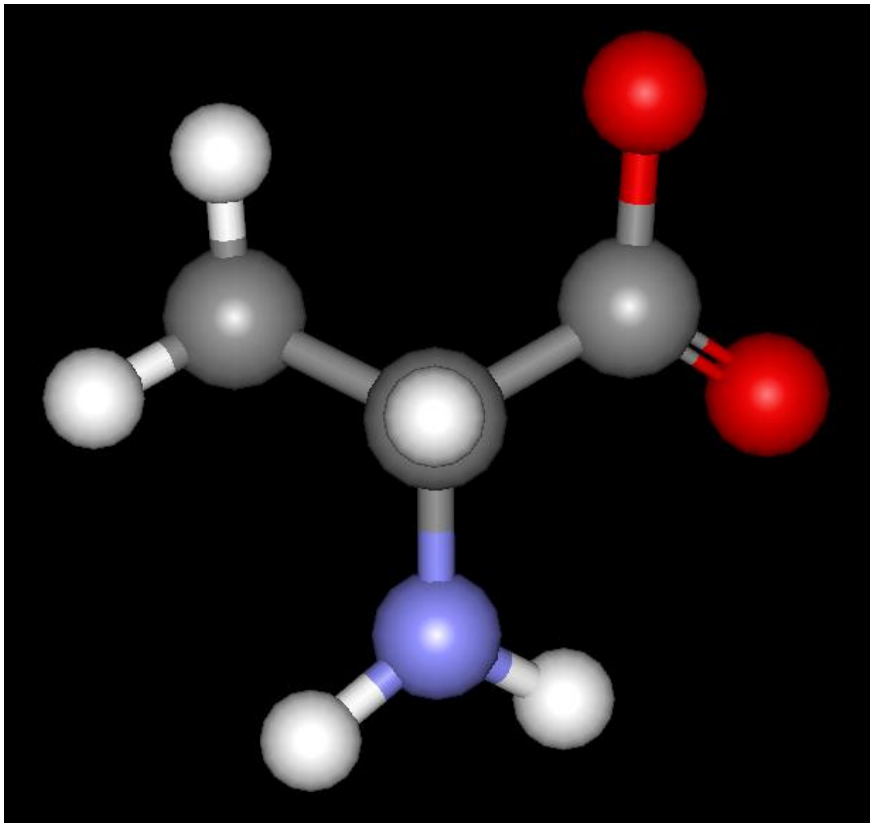
D and L enantiomers



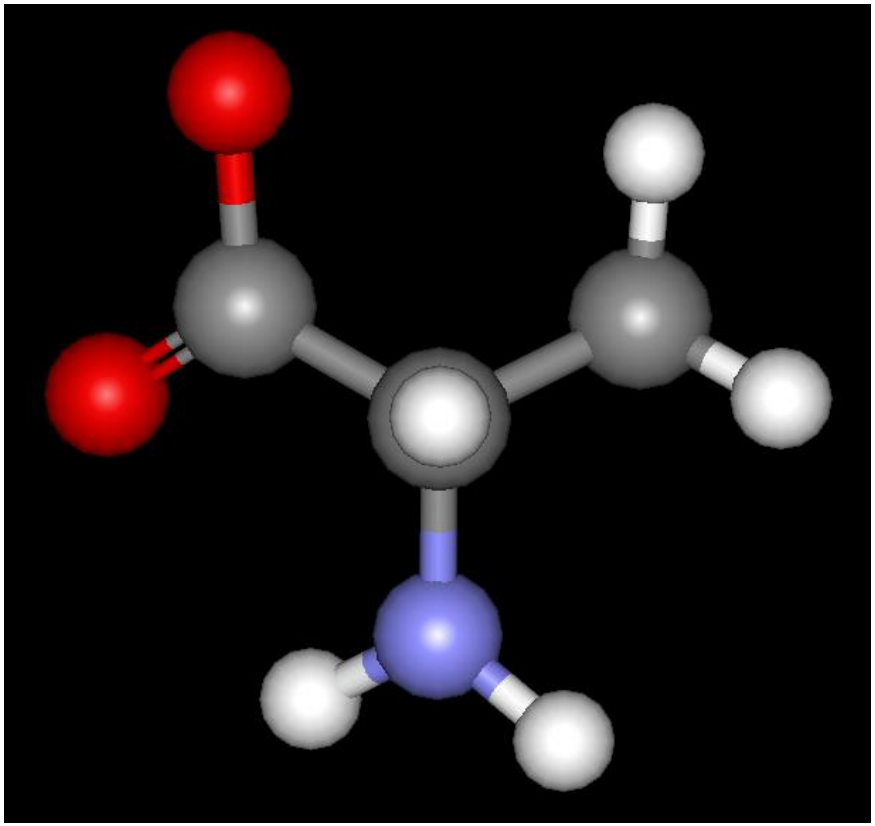
In living systems: L type !

No direct connection between the L-D enantiomer forms an the direction of the rotation of the polarization plane.
E.g.: (+)alanine (-)cysteine (-)tyrosine (+)valine

D and L enantiomers of alanine

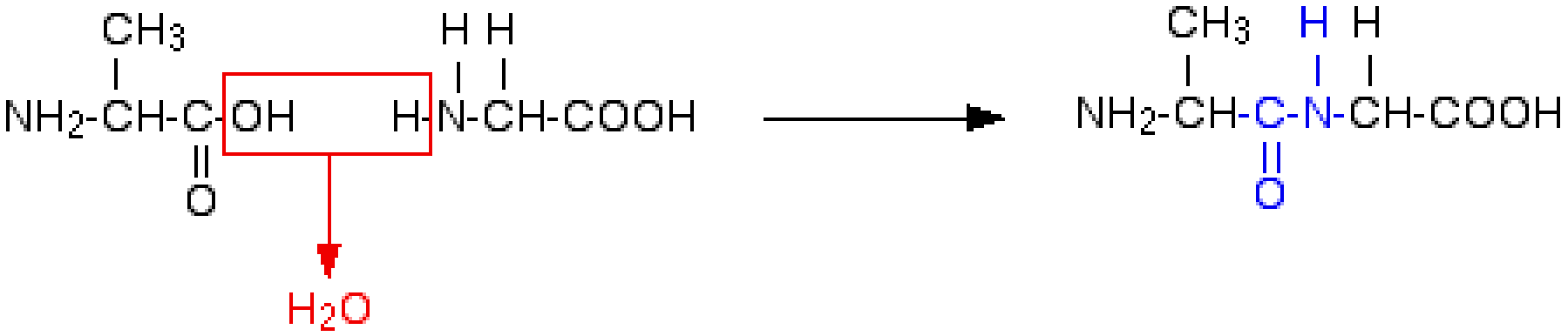


D



L

Connection of the amino acids: the peptide bond



Peptide 2.. c.a 20 amino acids

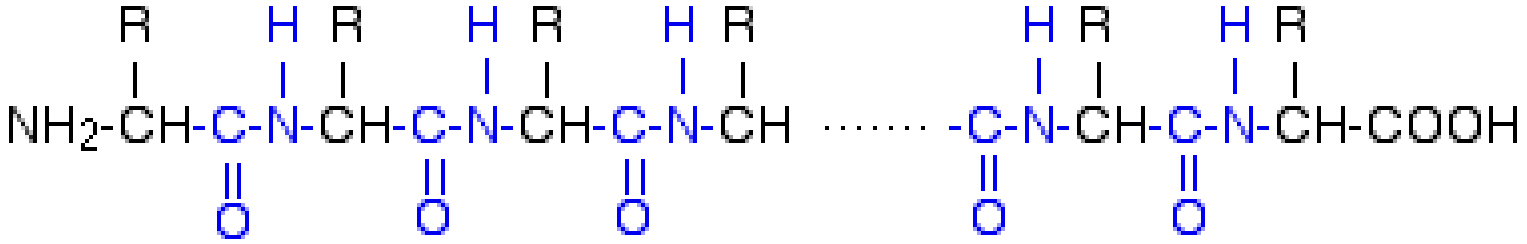
Protein: more than 20 amino acids

The primary structure

Primary structure:

the amino acid sequence in the polypeptide chain

Which direction? N terminal -> C-terminal



E.g.: (myoglobin, 1YMB)

GLY LEU SER ASP GLY GLU TRP GLN GLN VAL LEU ASN VAL…

... ALA LYS TYR LYS GLU LEU GLY PHE GLN GLY

Example: Myoglobin

Primary structure with 3-letter code (153 aa.):

GLY LEU SER ASP GLY GLU TRP GLN GLN VAL LEU ASN VAL
TRP GLY LYS VAL GLU ALA ASP ILE ALA GLY HIS GLY GLN
GLU VAL LEU ILE ARG LEU PHE THR GLY HIS PRO GLU THR
LEU GLU LYS PHE ASP LYS PHE LYS HIS LEU LYS THR GLU
ALA GLU MET LYS ALA SER GLU ASP LEU LYS LYS HIS GLY
THR VAL VAL LEU THR ALA LEU GLY GLY ILE LEU LYS LYS
LYS GLY HIS HIS GLU ALA GLU LEU LYS PRO LEU ALA GLN
SER HIS ALA THR LYS HIS LYS ILE PRO ILE LYS TYR LEU
GLU PHE ILE SER ASP ALA ILE ILE HIS VAL LEU HIS SER
LYS HIS PRO GLY ASP PHE GLY ALA ASP ALA GLN GLY ALA
MET THR LYS ALA LEU GLU LEU PHE ARG ASN ASP ILE ALA
ALA LYS TYR LYS GLU LEU GLY PHE GLN GLY

Example: Myoglobin

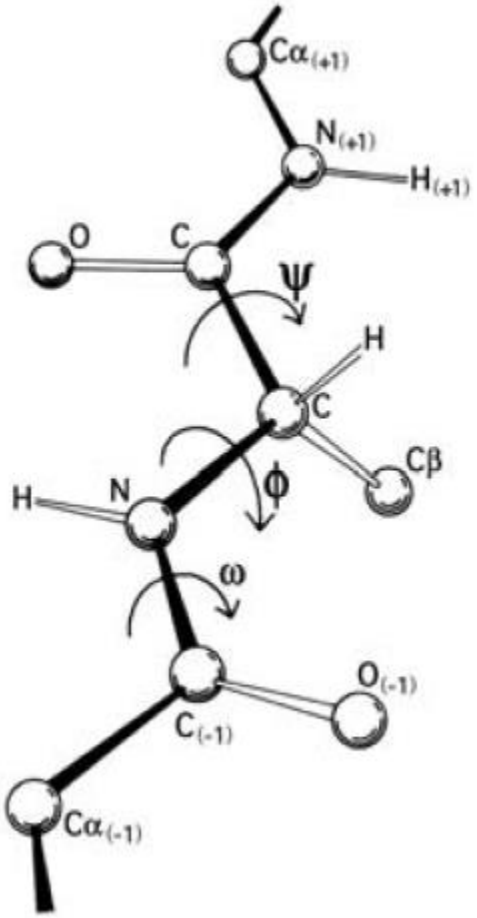
Primary structure with one letter code (153 aa.):

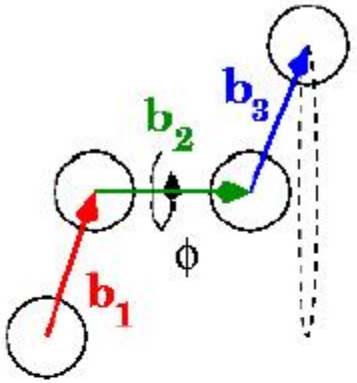
```
>1YMB:A|PDBID|CHAIN|SEQUENCE
GLSDGEWQQVLNVWGKVEADIAGHGQEVLIRLFTGHPETLEKFDKFKHLKTEAE
MKASEDLKKHGTVVLTALGGILKKKGHHEAELKPLAQSHATKHKIPIKYLEFIS
DAIIHVLHSKHPGDFGADAQGAMTKALELFRNDIAAKYKELGFQG
(FASTA format)
```

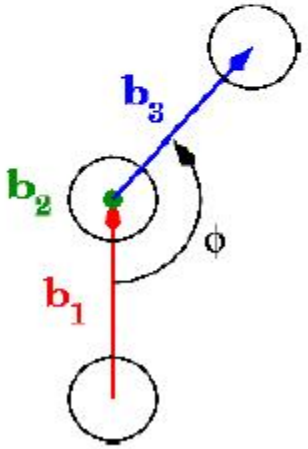
Rotation around the chemical bonds of the backbone

3 backbone bonds for each AA

- 1 rigid (delocalization)
- 2 rotation possible:
  - Φ, Ψ dihedral angles







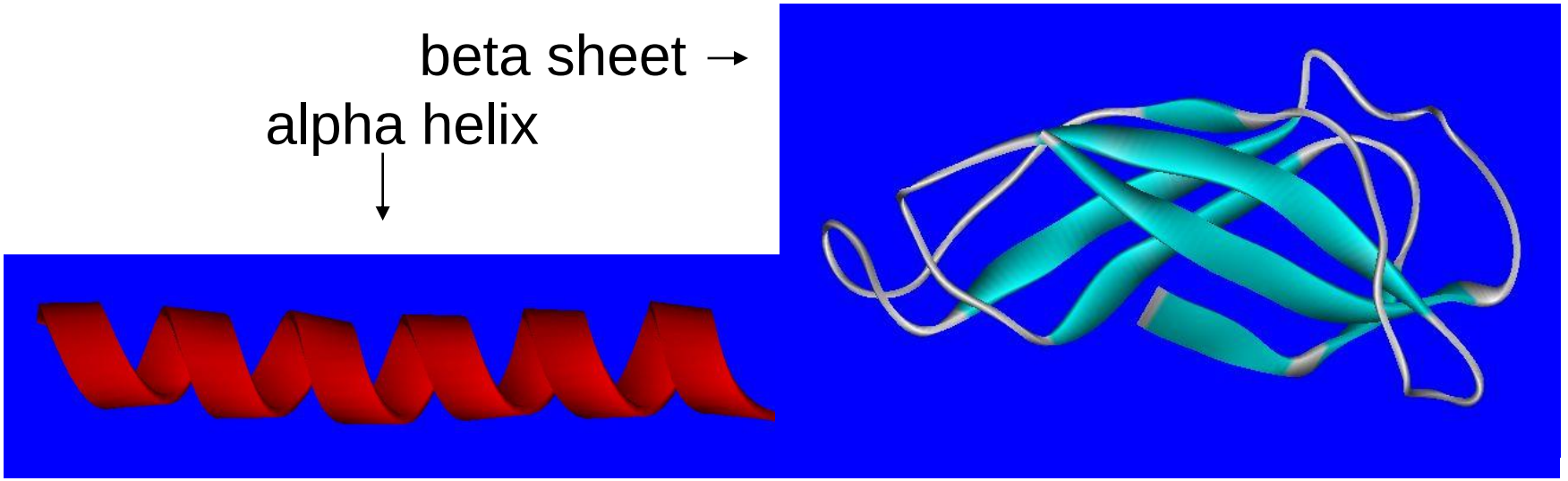
2N rotational degree of freedom

Secondary structure

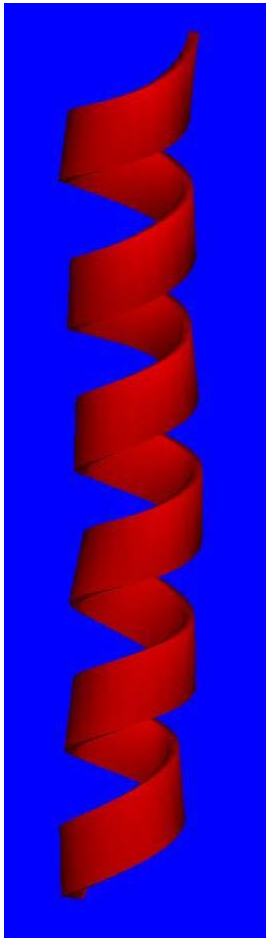
Regular local structures of linear segments of polypeptide chains.

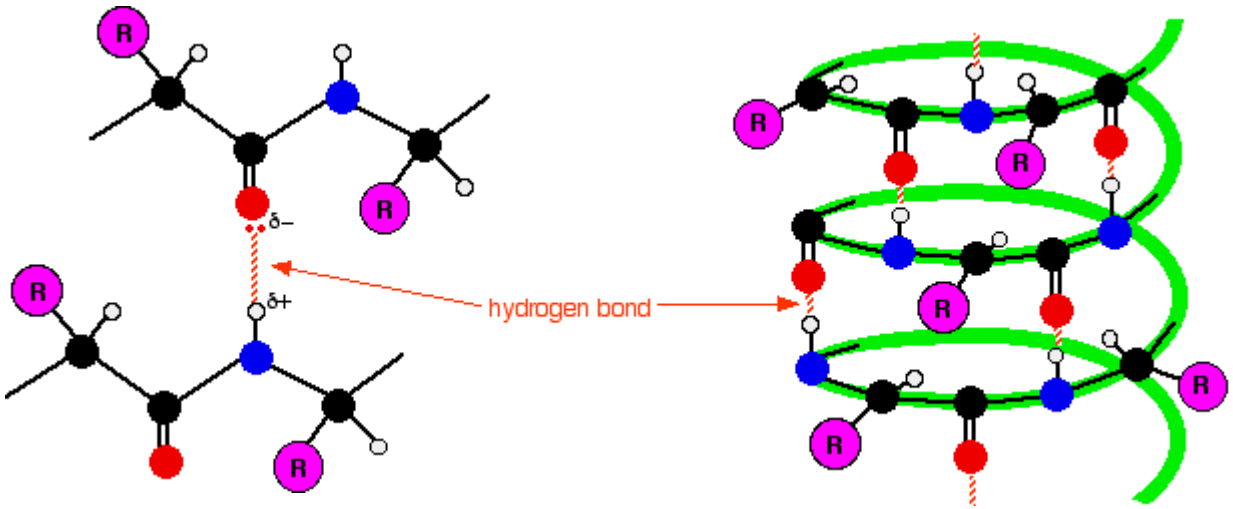
They are stabilized by H-bonds

Typical forms:

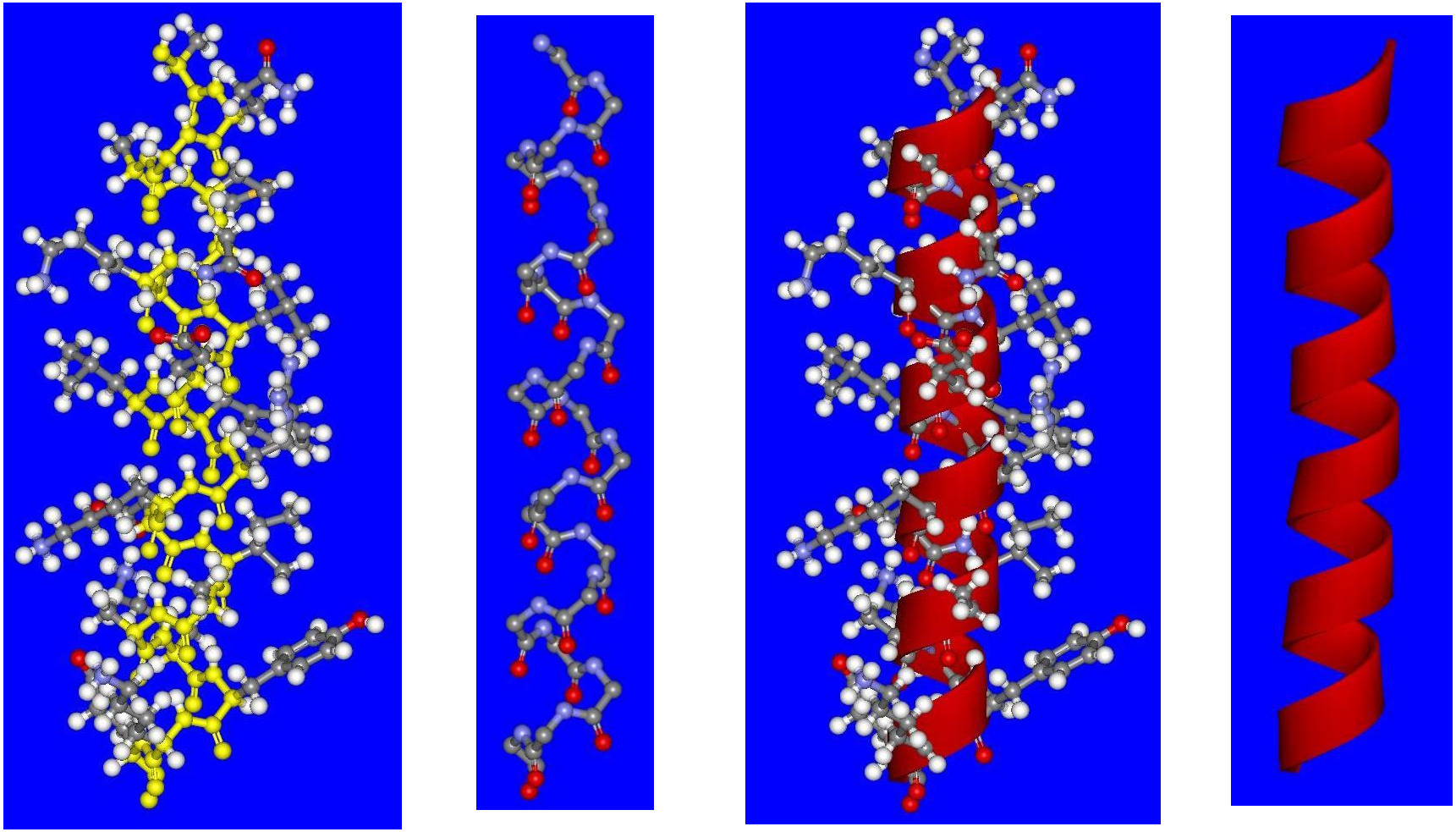


Alpha helix

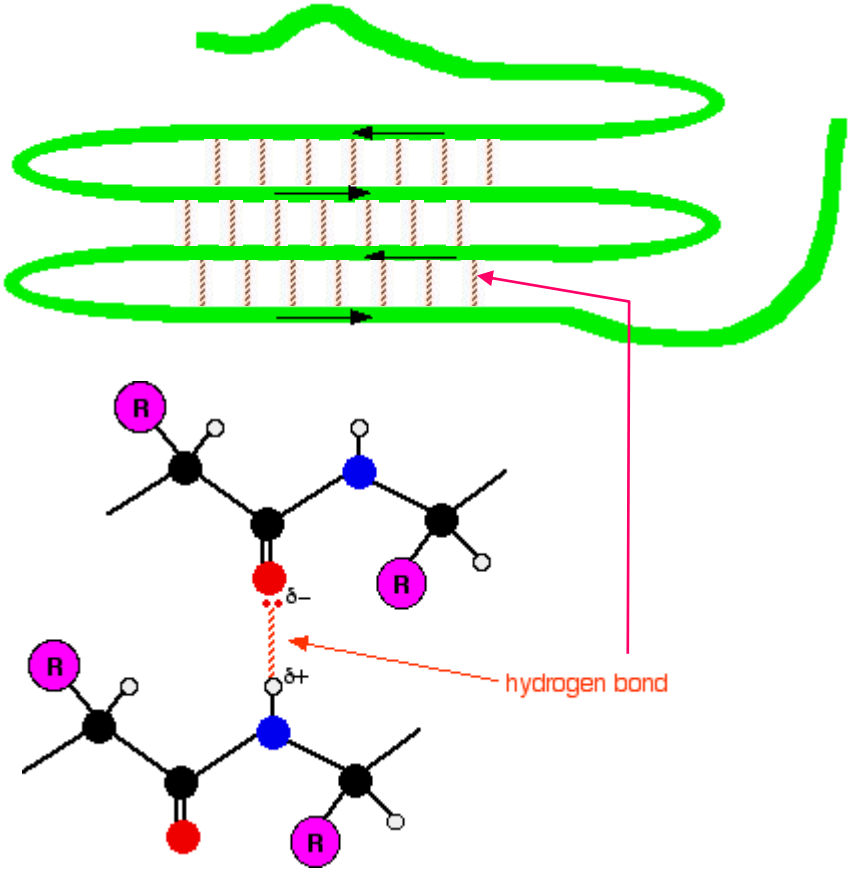


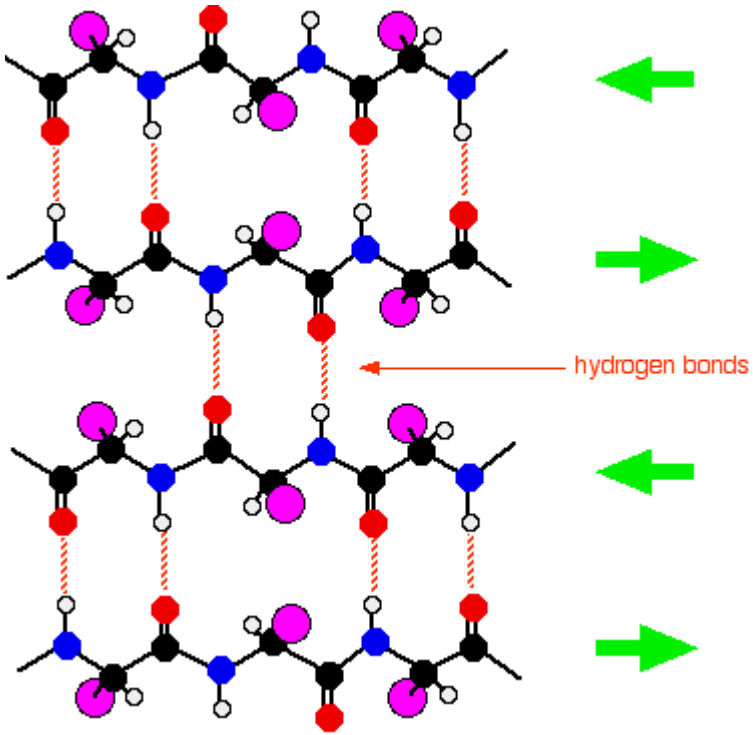


Alpha helix



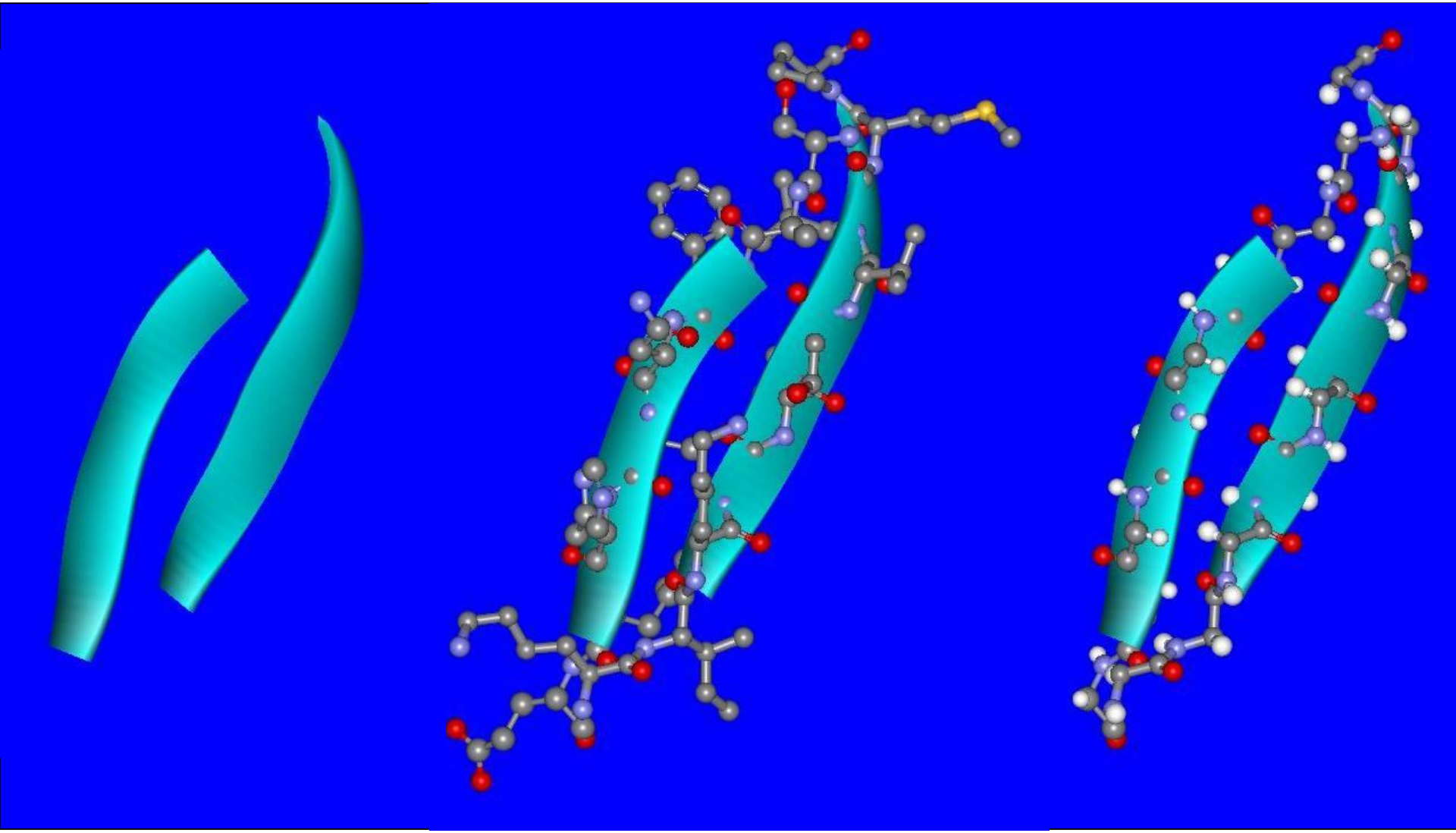
Beta sheet



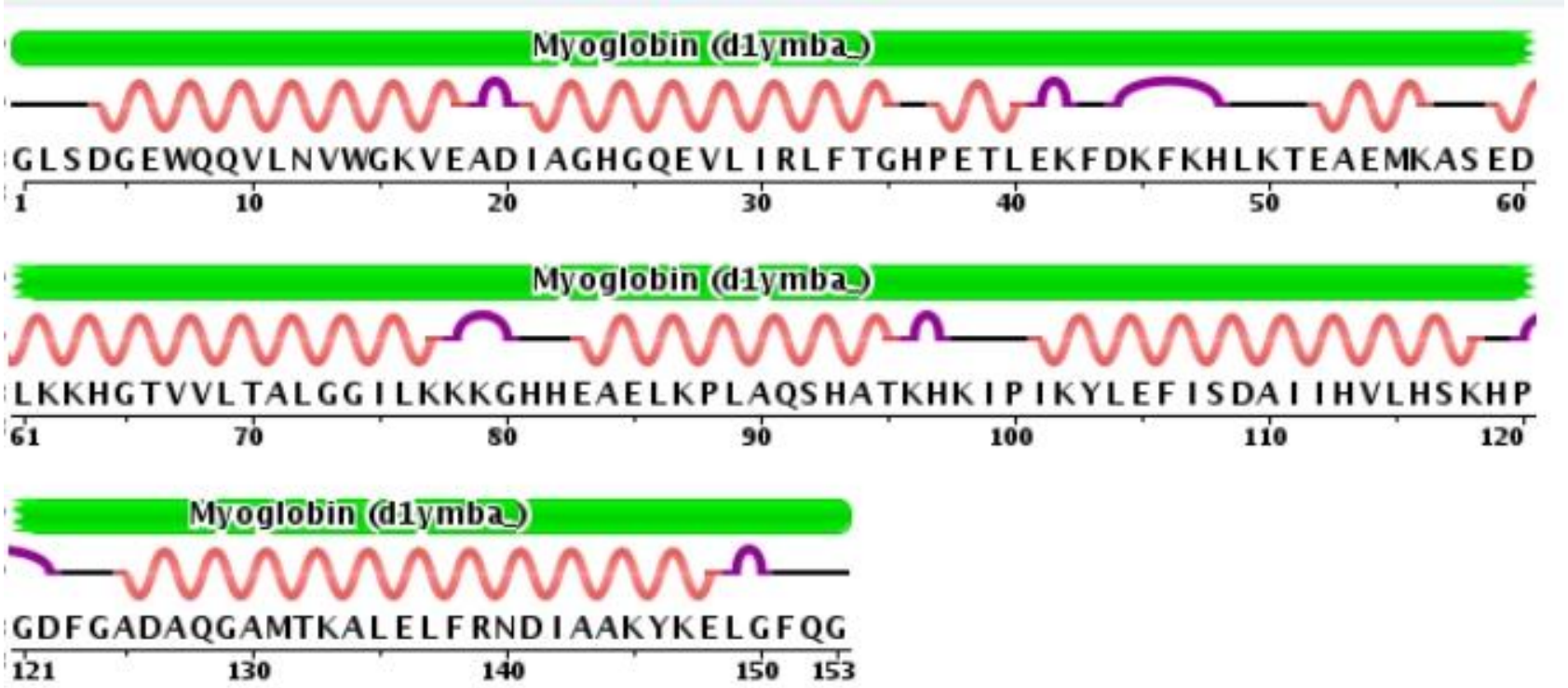


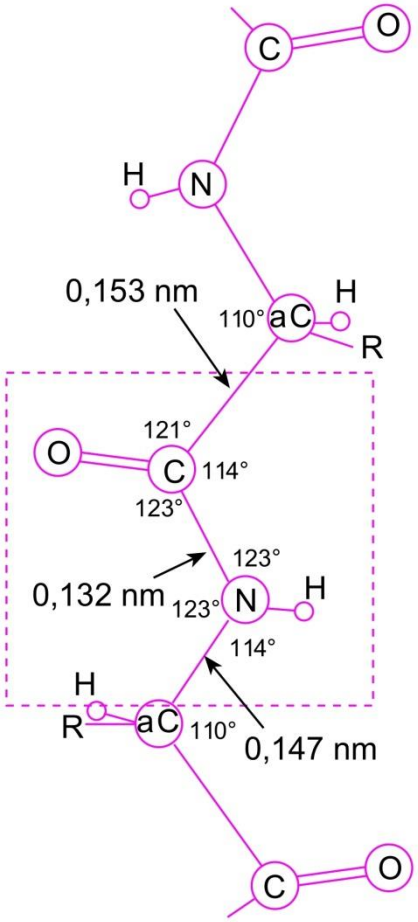
antiparallel

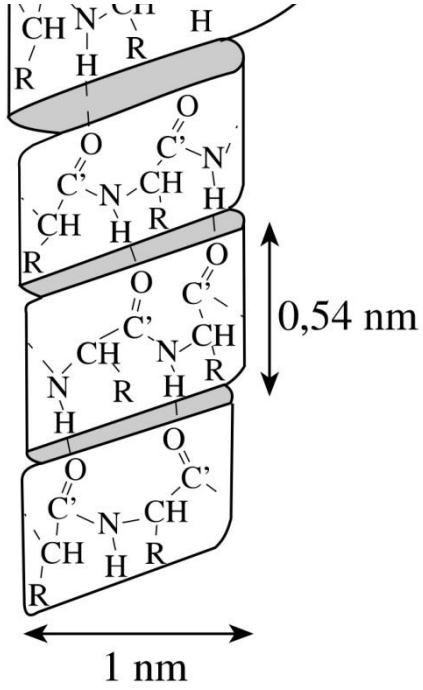
Beta sheet



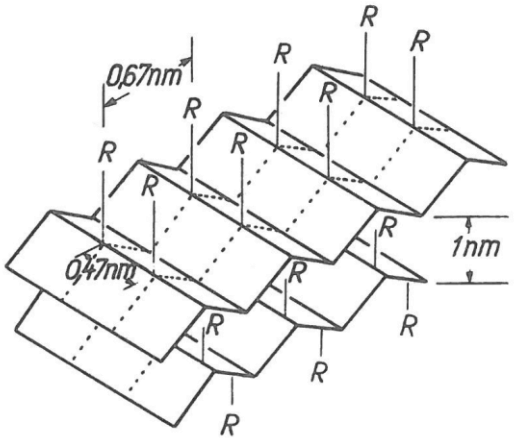
One dimensional representation of the secondary structure

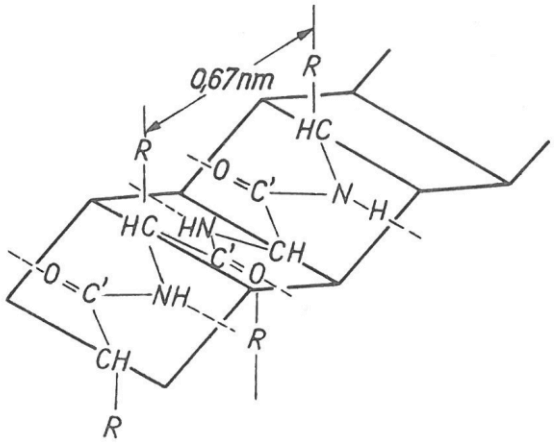






3,6 AA/turn
i -> i+4





Stabilization by H-bonds

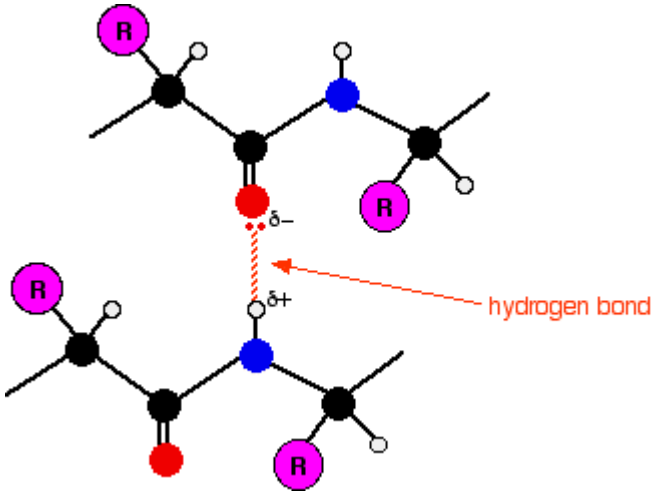
12-30 kJ/mol

Cf: Covalent bond: 200 kJ/mol

van der Wals:1-2 kJ/mol

thermal energy (RT):

2.5 kJ/mol (T=300K)



Boltzmann factor: $e^{-\frac{\Delta E}{RT}} = 0.000335 = \frac{1}{2981} \approx \frac{1}{3000}$

(ΔE=20kJ/mol)