Advanced biochemistry and its methods
Lecture 4

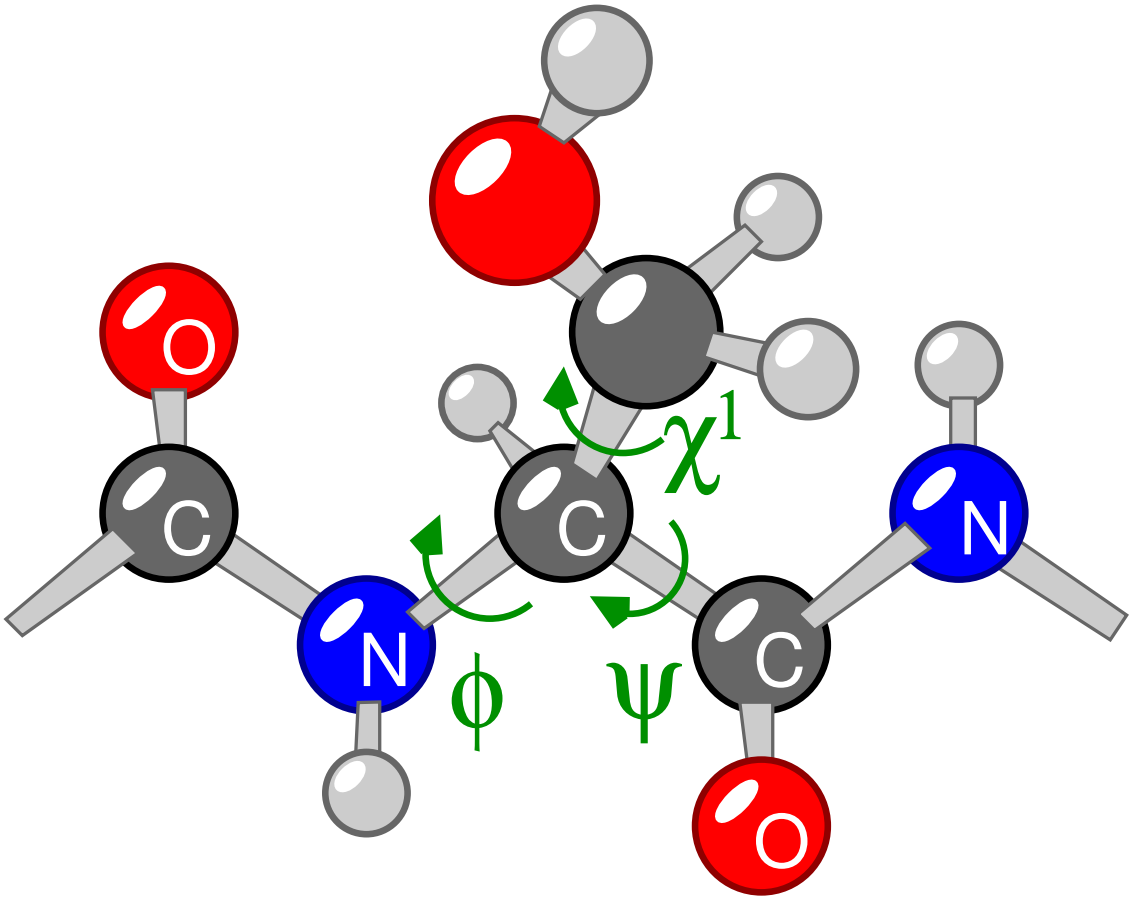
Lukáš Žídek

lzidek@chemi.muni.cz

Finkelstein and Ptitsyn: Protein Physics, Academic Press 2002
Daune: Molecular Biophysics, Molecular Biophysics, Oxford University Press 1999
Žídek: Strukturní biochemie (skripta k prednášce C9530), kapitoly 2, 6, A

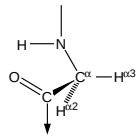
Proteins

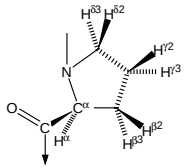
Amino acids connected by peptide bonds

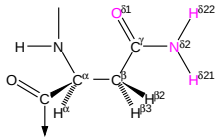


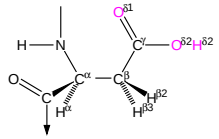
Protein structure = conformation
defined by torsion angles $(\phi, \psi, \chi^1, \ldots)$

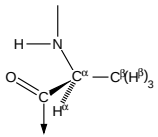
Amino acids

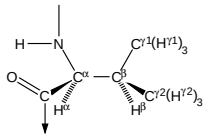

Gly (G)

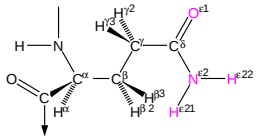

Pro (P)

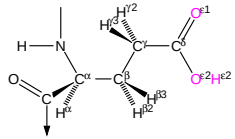

Asn (N)

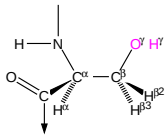

Asp (D)

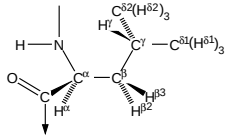

Ala (A)

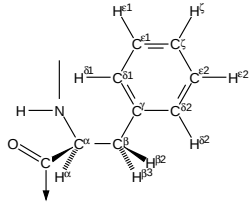

Val (V)

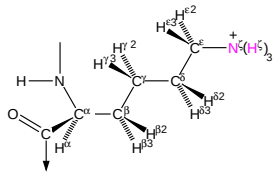

Gln (Q)

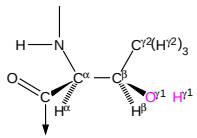

Glu (E)

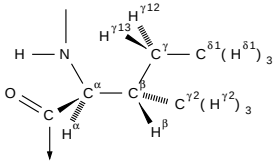

Ser (S)

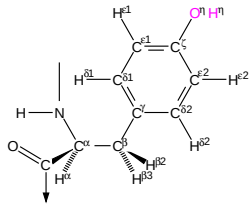

Leu (L)

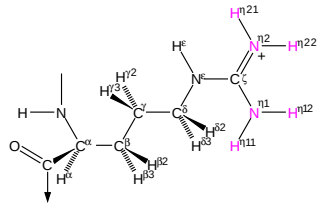

Phe (F)

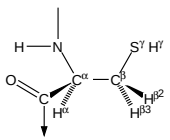

Lys (K)

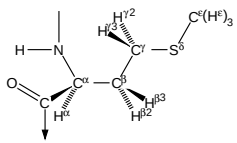

Thr (T)

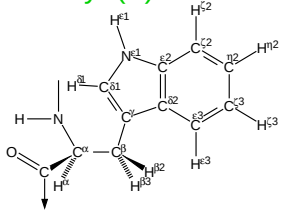

Ile (I)

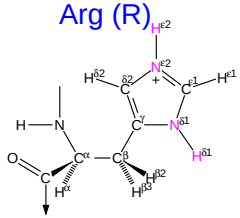

Tyr (Y)


Arg (R)


Cys (C)


Met (M)


Trp (W)


His (H)

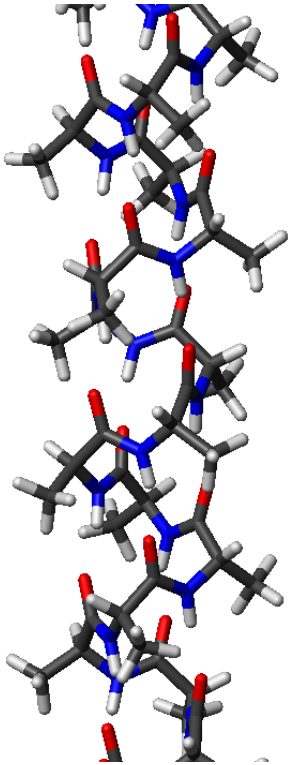
Primary structure

Amino acid sequence

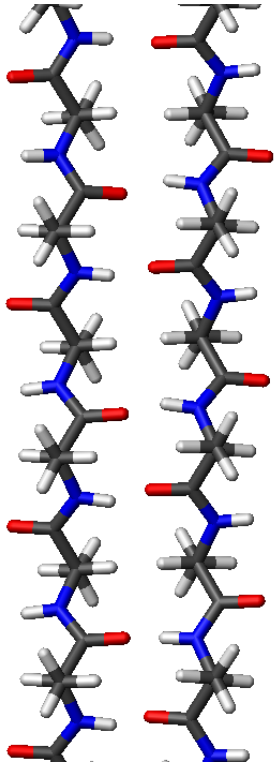
SAKIIHLTDDSFDTDVLKAILVDFW
AEWCGPCKMIAPILDEIADEYQGKL
TAPKYGIRGIPTLLLFKNGEVAATK
VGALSKGQLKEFLDANLA

Secondary structure

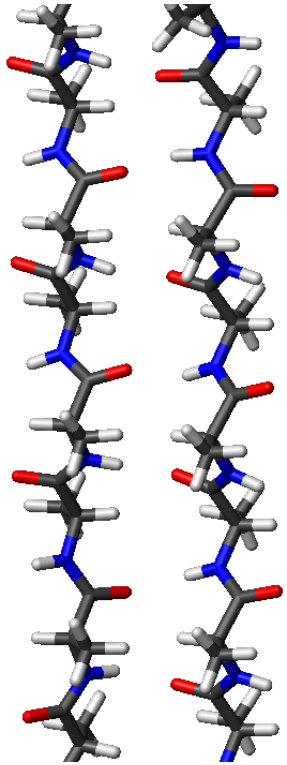
Conformation of protein backbone
regular universal repetitive motifs



α-helix

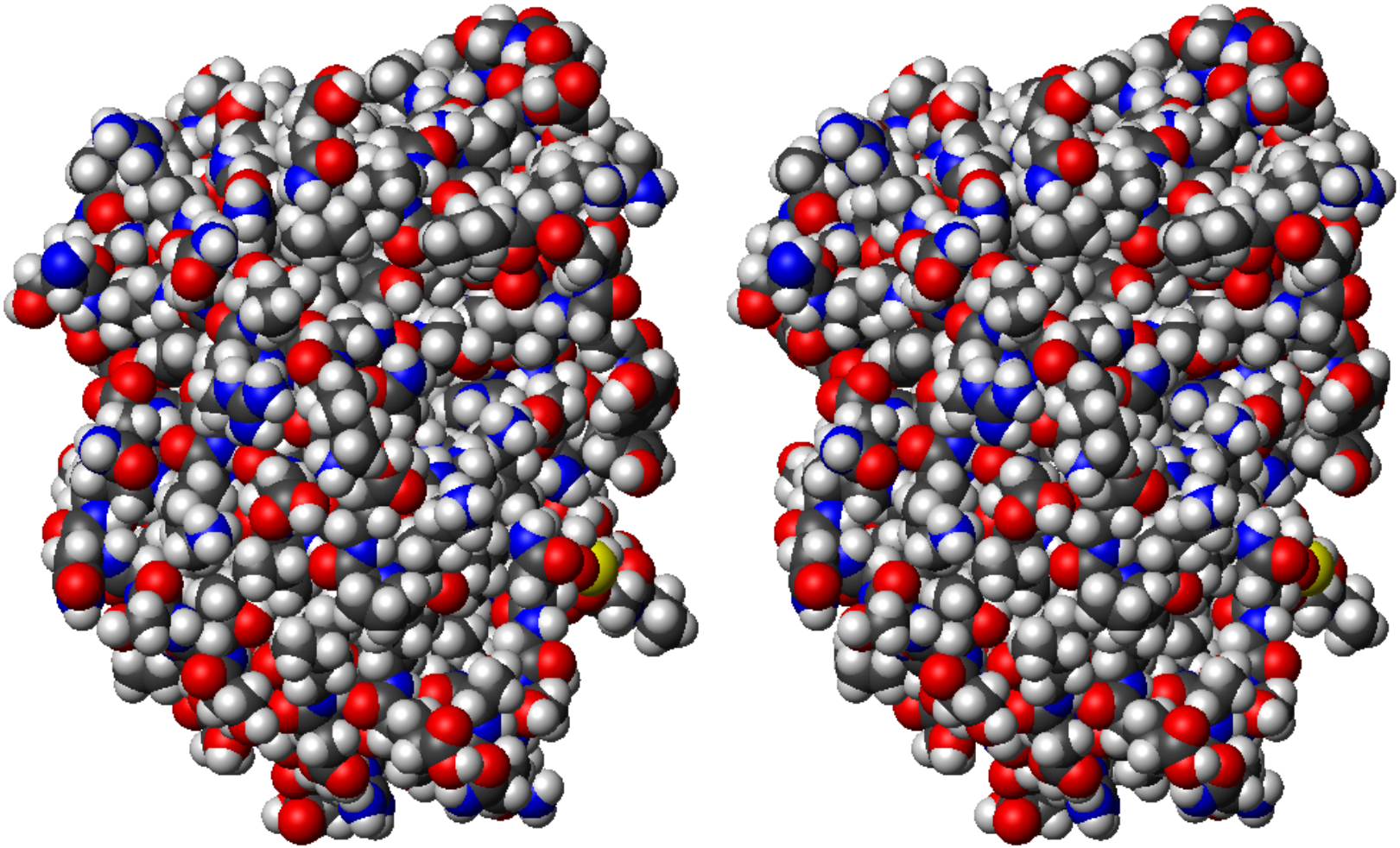


antiparallel β-sheet

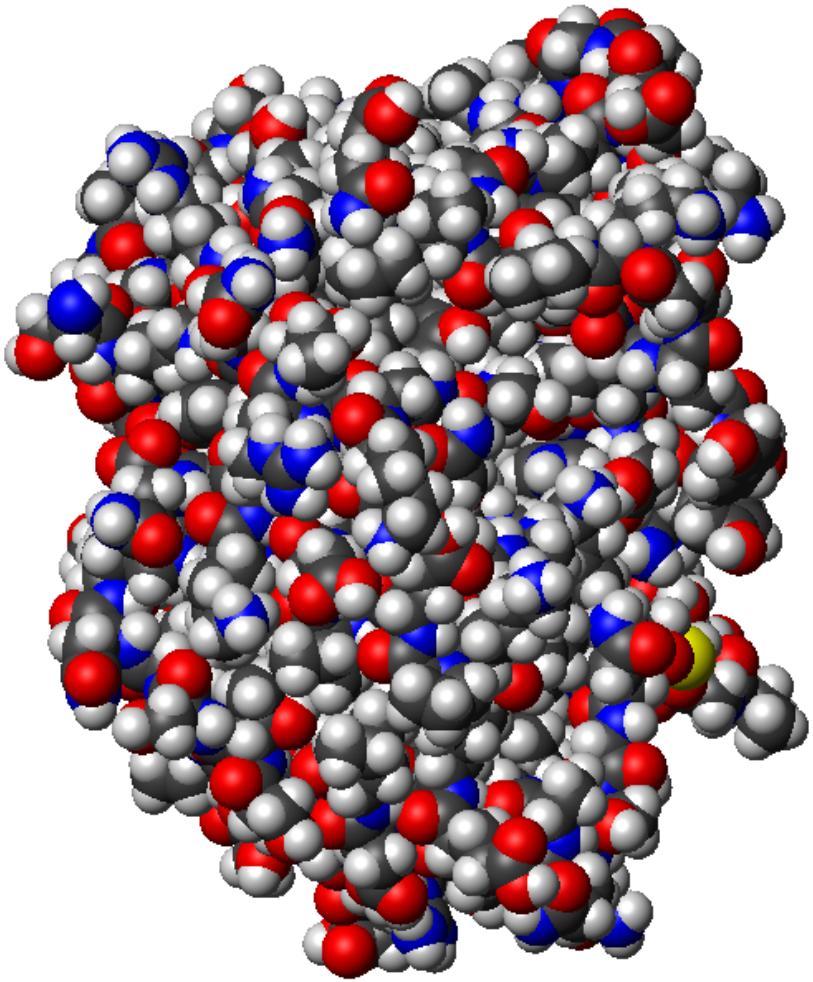


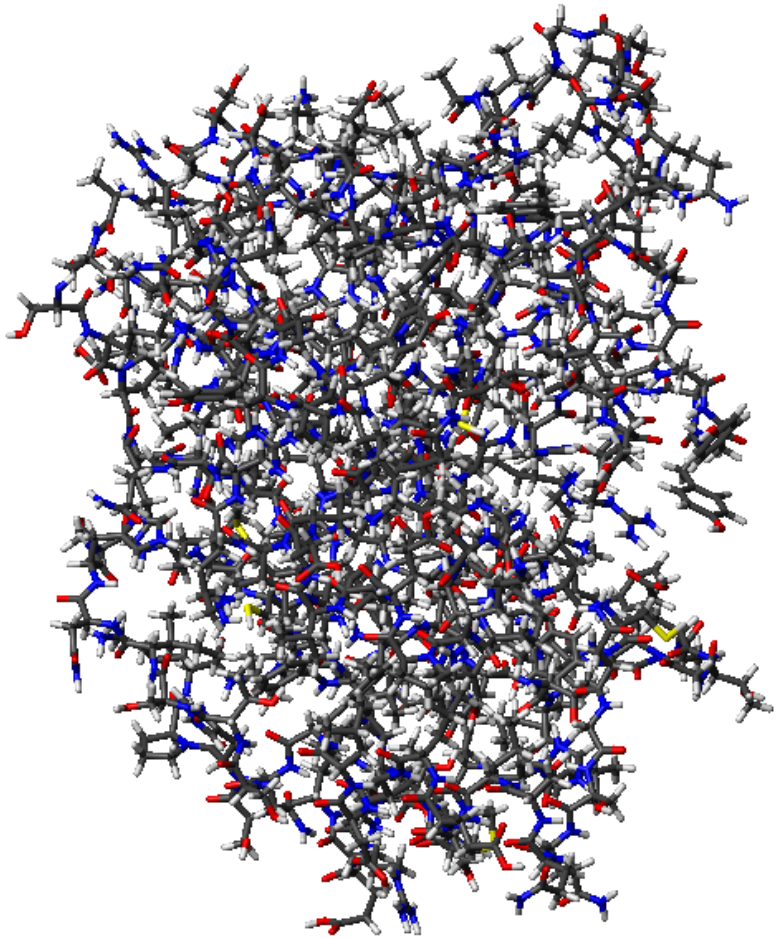
parallel β-sheet

Tertiary structure

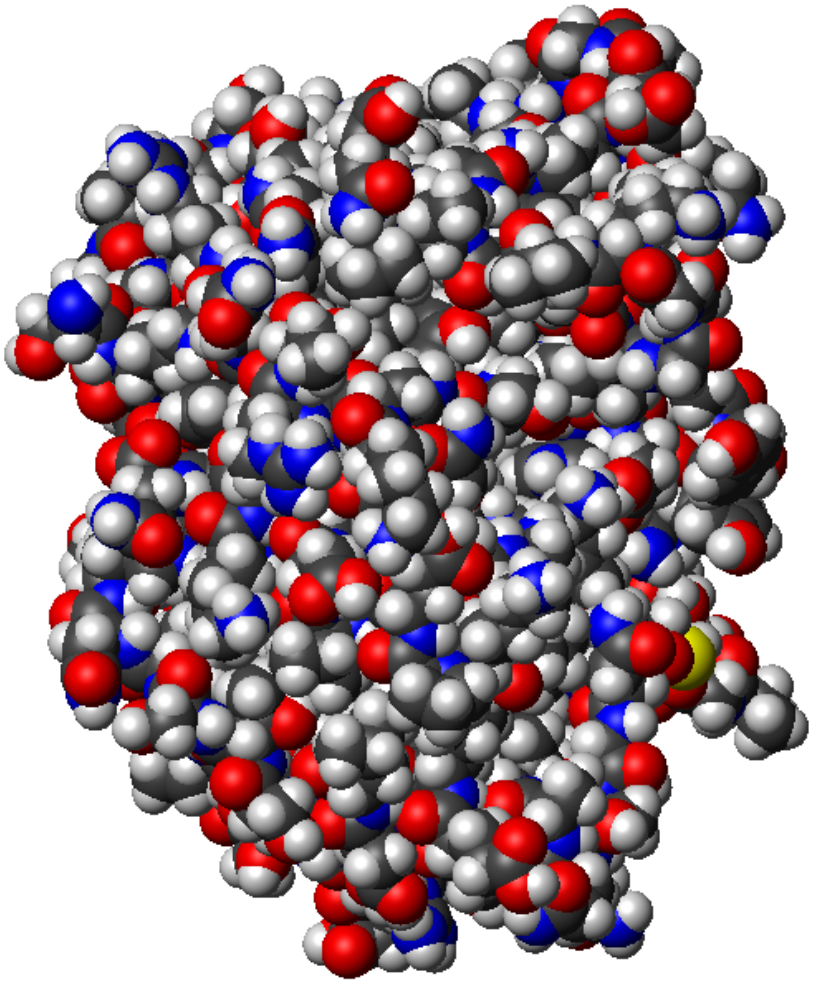


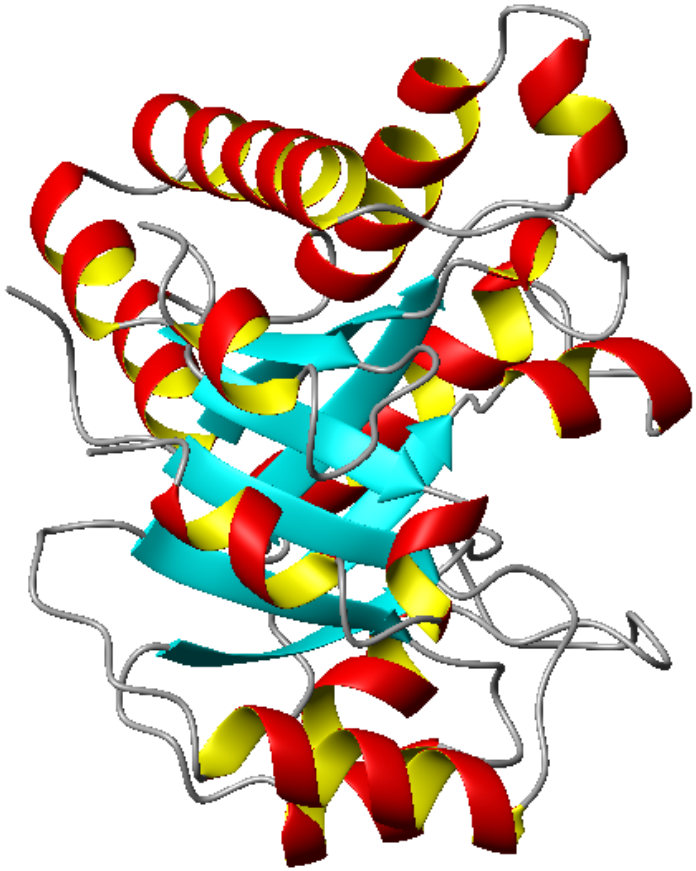
Tertiary structure





Tertiary structure





Thermodynamics

Protein samples in biochemistry:
many molecules with multiple possible conformational states in thermal equilibrium $\Rightarrow$ (statistical) thermodynamics

Energy U:

First law: $\Delta U = \underbrace{Q}_{\text{heat}} + \underbrace{W}_{\text{work}}$

Second law: $T\Delta S \geq Q$

Entropy $S = R \ln \Omega$ ($\Omega =$ number of microstates, combinations)

Taken together, $\Delta U - T\Delta S \leq 0$ if $W = 0$, including work due to expansion ($p\Delta V = 0$)

$A = U - TS$ (Helmholtz free energy) has minimum at equilibrium at constant temperature & volume $\mathrm{d}T = 0, \mathrm{d}V = 0$.

Enthalpy $H = U + pV$:

$G = H - TS$ (Gibbs free energy) has minimum at equilibrium at constant temperature & pressure $\mathrm{d}T = 0, \mathrm{d}p = 0$

Boltzmann's law:
numbers of molecules in states 1 and 2 of the most probable macrostate (with the highest number of microstates):

$$\frac{n_1}{n_2} = \mathrm{e}^{-(U_1 - U_2)/RT}$$

"Small" energy is $\ll RT \approx 2\,500\,\mathrm{J/mol}$ at $300\,\mathrm{K}$ (room temp.)

Ideal gas: $V_\mathrm{m} = 0.0224\,\mathrm{m}^3$, $p_\mathrm{atm} = 10^5\,\mathrm{Pa} \Rightarrow p_\mathrm{atm} V_\mathrm{m} = 2\,240\,\mathrm{J/mol}$
Liquid water: $V_\mathrm{m} = M_\mathrm{r}/\rho = 1.8 \times 10^{-5}\,\mathrm{m}^3 \Rightarrow p_\mathrm{atm} V_\mathrm{m} = 1.8\,\mathrm{J/mol}$

$U \approx H$, $A \approx G$ in biochemistry

Energy

Chemistry: electromagnetic force only

Coulomb's law:

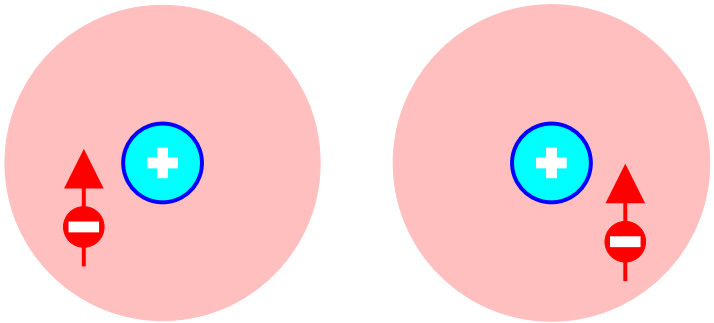
$$F = \frac{1}{4\pi\epsilon_0}\frac{Q_1 Q_2}{r^2}$$

$$U = \int_{r_{\text{ref}}}^{r} F\mathrm{d}r' = \int_{\infty}^{r} F\mathrm{d}r' = \frac{Q_1 Q_2}{4\pi\epsilon_0}\int_{\infty}^{r}\frac{1}{r^2}\mathrm{d}r' = \frac{1}{4\pi\epsilon_0}\frac{Q_1 Q_2}{r}$$

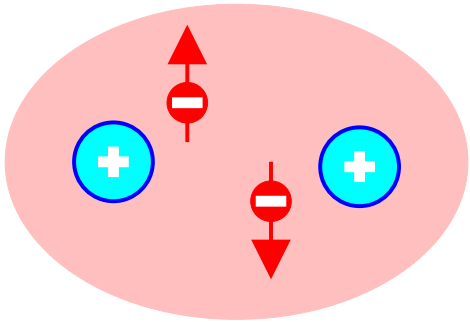
- Force is a vector: $\vec{F} = \frac{1}{4\pi\epsilon_0}\frac{Q_1 Q_2}{r^2}\cdot \underbrace{\frac{\vec{r}}{r}}_{\text{unit vector}}$
- Electric intensity: $\vec{E} = \frac{1}{4\pi\epsilon_0}\frac{Q}{r^2}\cdot\frac{\vec{r}}{r}$
- $U = \frac{N_A}{4\pi\epsilon_0}\frac{Q_1 Q_2}{r}$ if expressed in kJ/mol

Covalent bonds

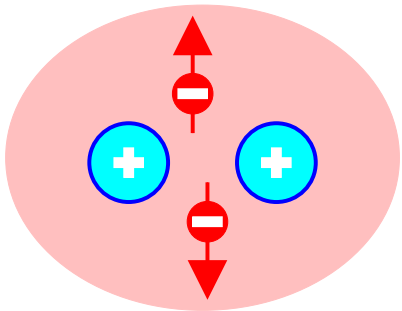
Quantum mechanics



reference energy

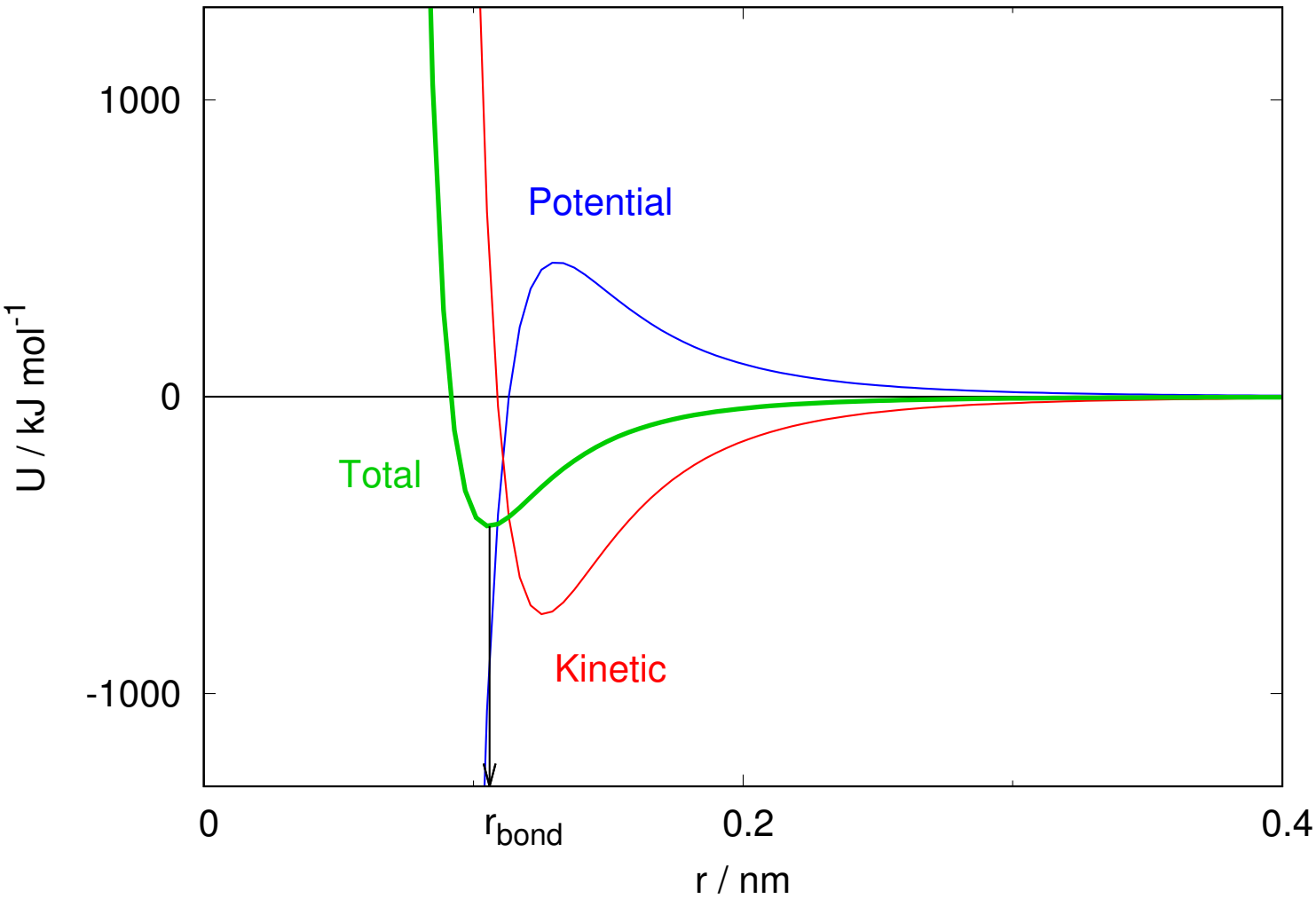


lower energy



higher energy

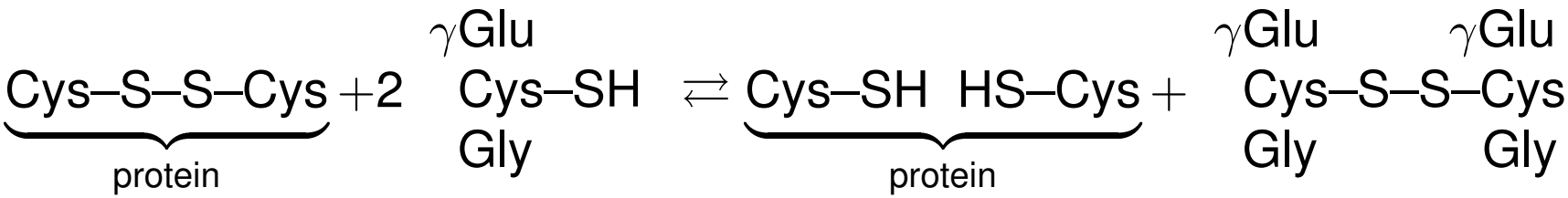
Covalent bonds



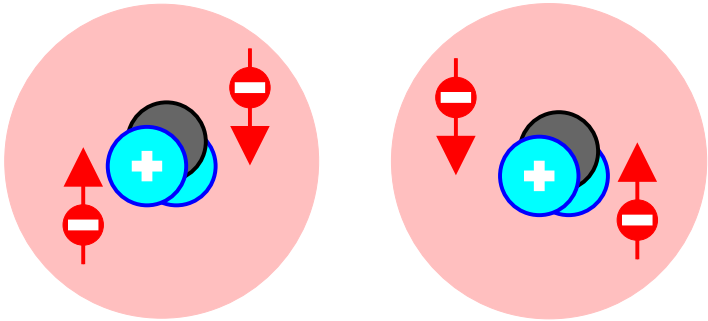
Covalent bonds in proteins

- Define primary structure
- Covalent bonds defining tertiary structure:
  - Metal coordination
  - Disulfide bridges

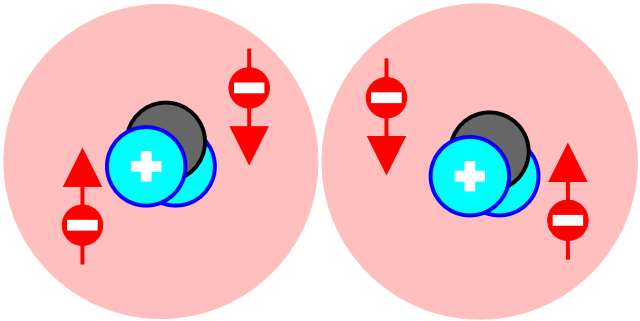
S–S bridges important (and frequent) in extracellular proteins but play marginal structural role in intracellular proteins: Exchange with glutathione ($\Delta G \approx 0$)



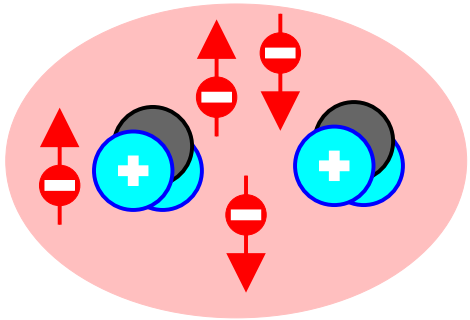
Interactions of nonpolar molecules



reference energy

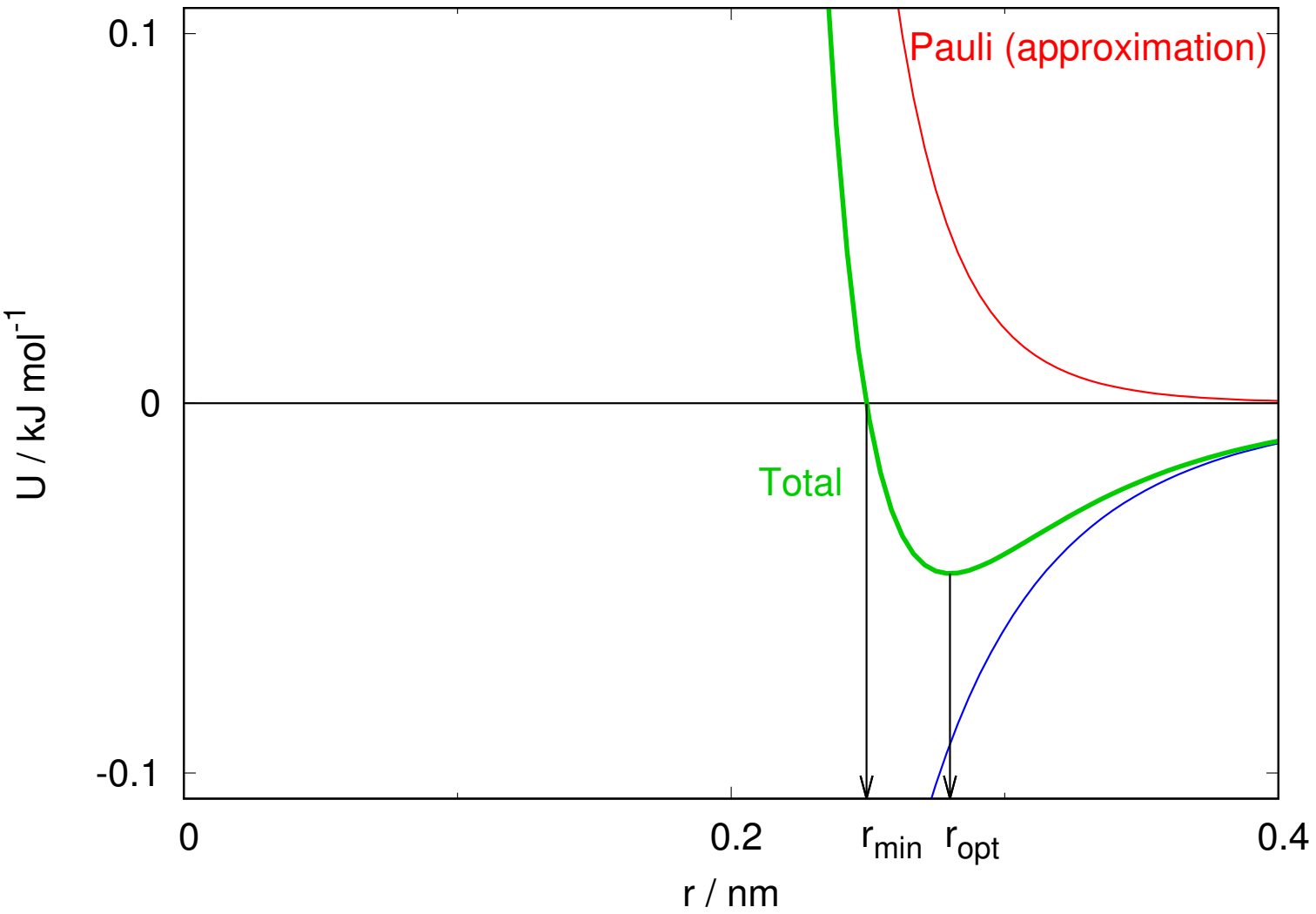


lower energy

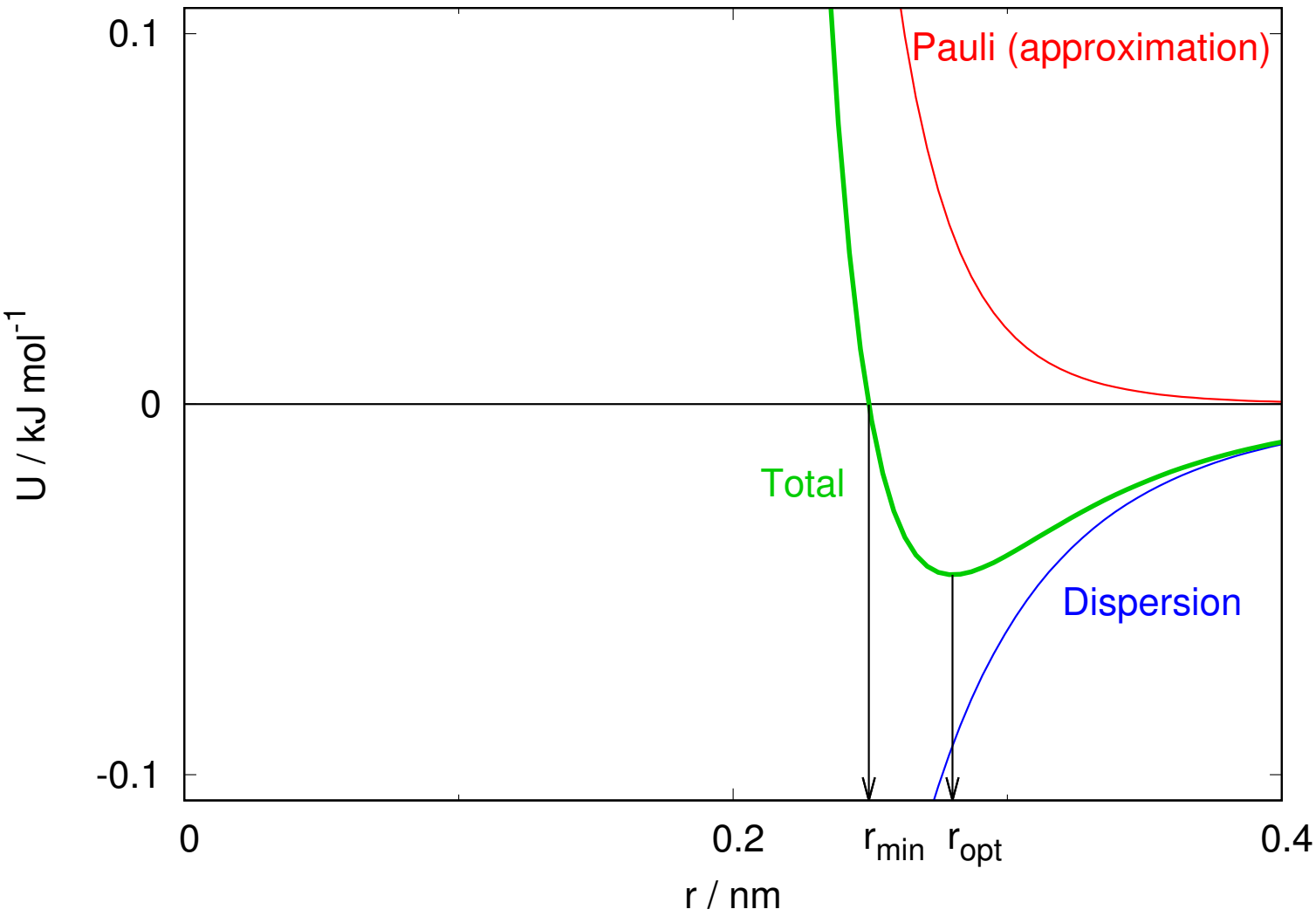


IMPOSSIBLE !

Interactions of nonpolar molecules

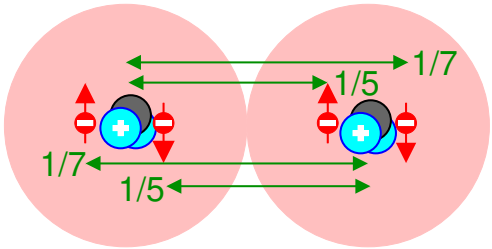


Dispersion forces



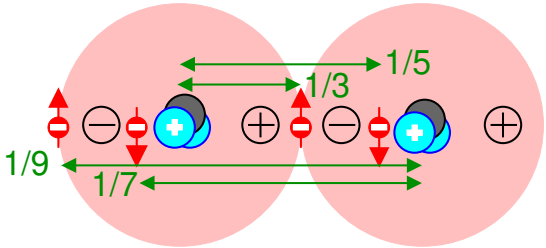
Dispersion forces

intermolecular energy: relative repulsive energy is identical
relative attractive energy $= -1/r$



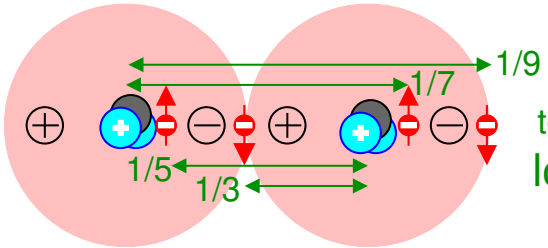
total relative intermolecular attractive energy $= -(1/7+1/5+1/7+1/5) = -4.114$

reference energy



total relative intermolecular attractive energy $= -(1/5+1/3+1/9+1/7) = -4.724$

lower energy

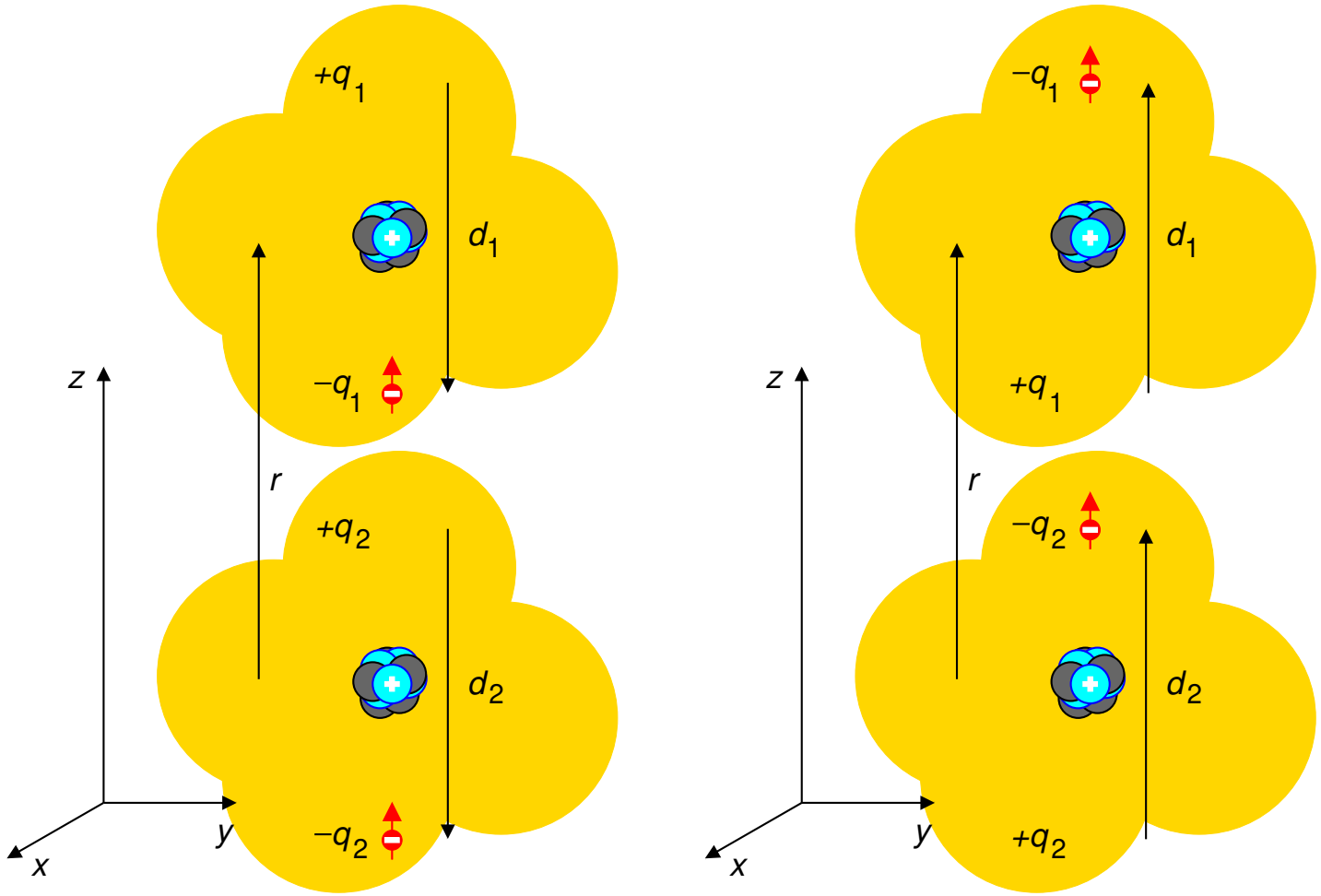


total relative intermolecular attractive energy $= -(1/5+1/3+1/9+1/7) = -4.724$

lower energy

Dispersion forces

Depends on vibrations: $\langle U \rangle = \frac{3h\nu}{4} \frac{\alpha^2}{(4\pi)^2} \frac{1}{r^6}$ (identical molecules)



Dispersion forces and Pauli repulsion in proteins

van der Waals interactions

- **Dispersion force:**
  universal (polar and nonpolar molecules/groups)
  backbone and sidechains
- **Pauli repulsion:**
  steric hindrance – limits possible torsion angles
  backbone: $\phi, \psi, (\omega)$ Ramachandran diagram
  sidechains: $\chi^1, \chi^2, \ldots$

Lennard-Jones potential:

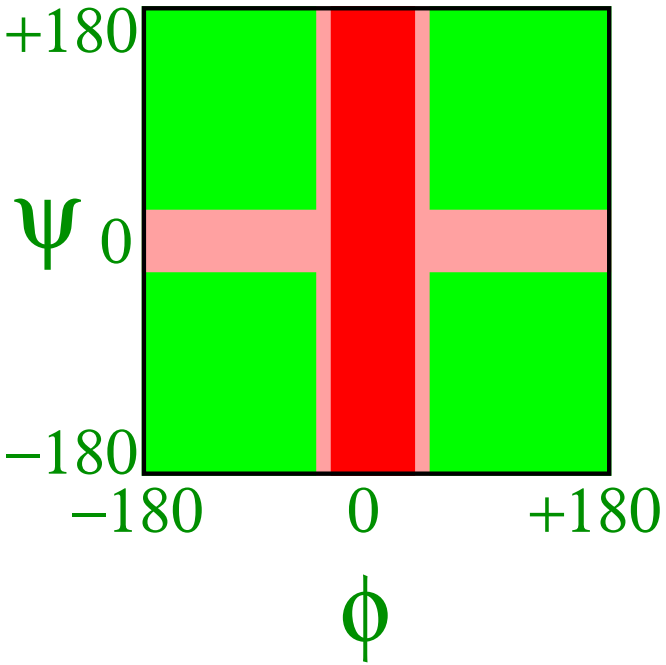
$$U = U_{\text{opt}} \left(\left(\frac{r_{\text{opt}}}{r} \right)^{12} - 2 \left(\frac{r_{\text{opt}}}{r} \right)^{6} \right)$$

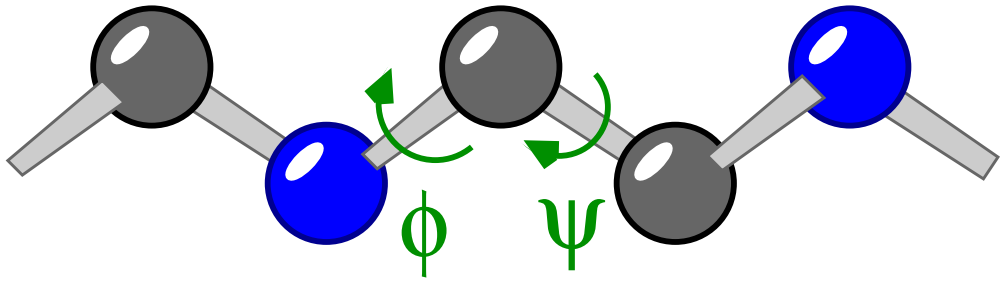
Van der Waals parameters

Atom···atom	U_{opt} / kJ mol^{-1}	r_{opt} / nm	r_{min} / nm
He···He	0.05	0.28	0.25
–H···H–	0.50	0.24	0.20
–C···C–	0.50	0.34	0.30
–N···N–	0.85	0.31	0.27
–O···O–	0.95	0.30	0.27

Ramachandran diagram

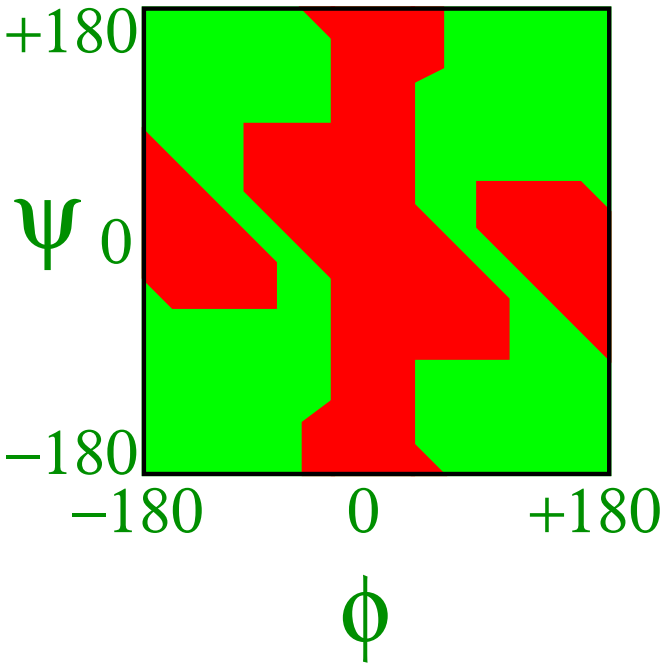
Repulsion of backbone C and N only

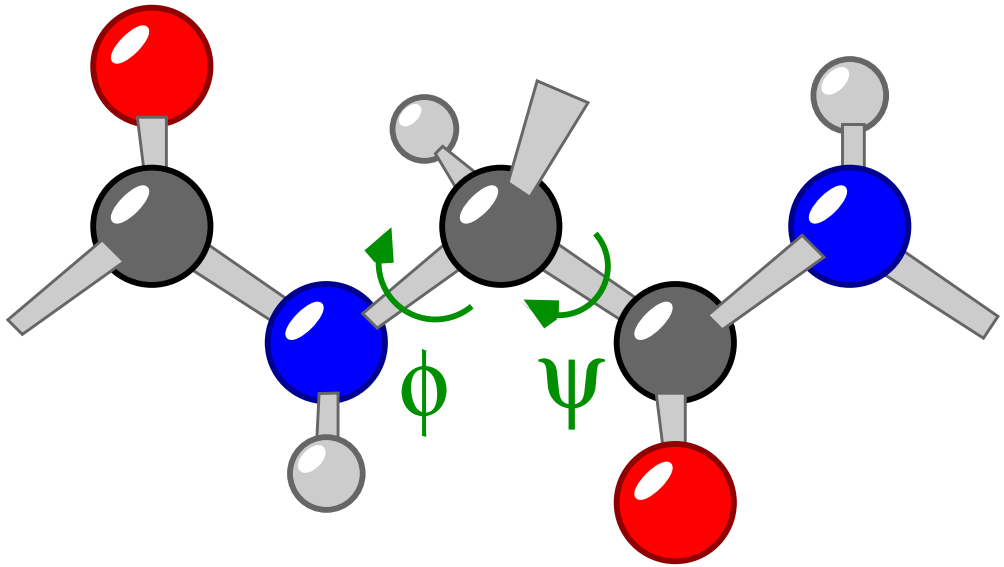




Ramachandran diagram

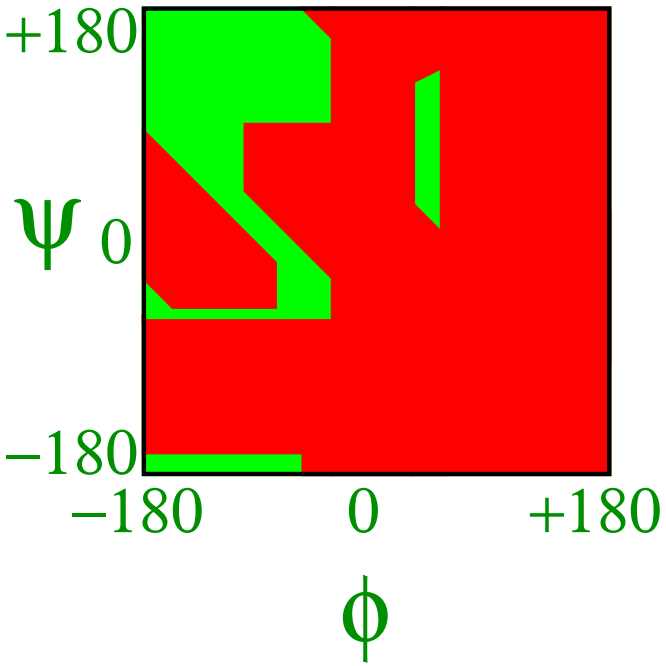
Repulsion including backbone amide H and O

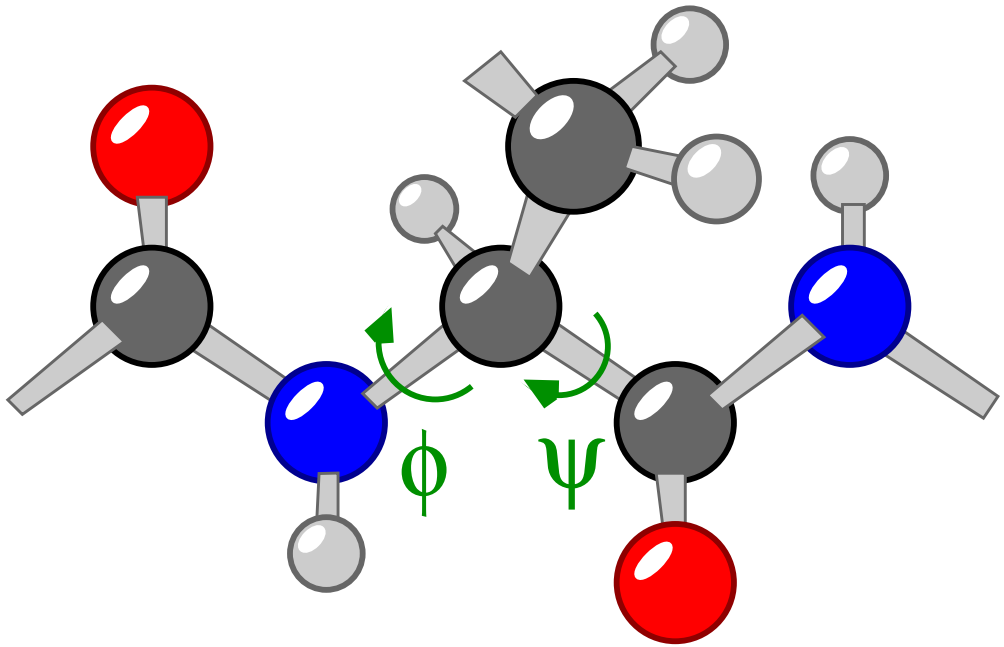




Ramachandran diagram

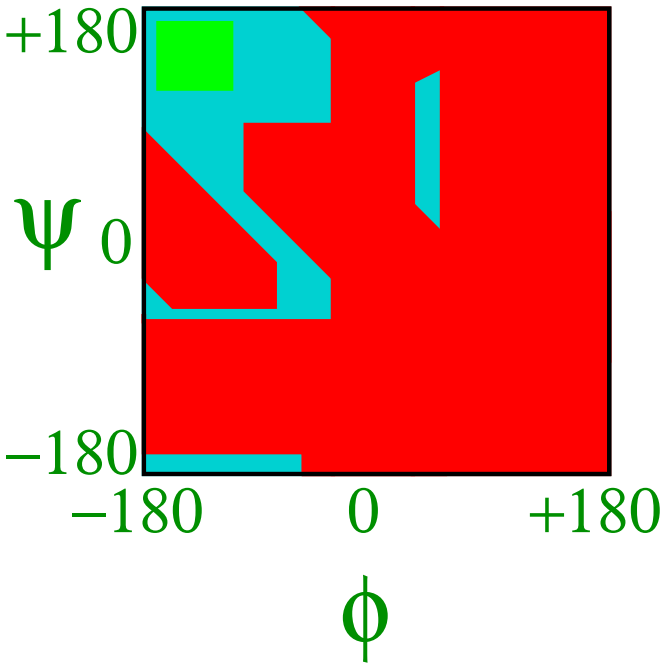
Repulsion including C^{β}

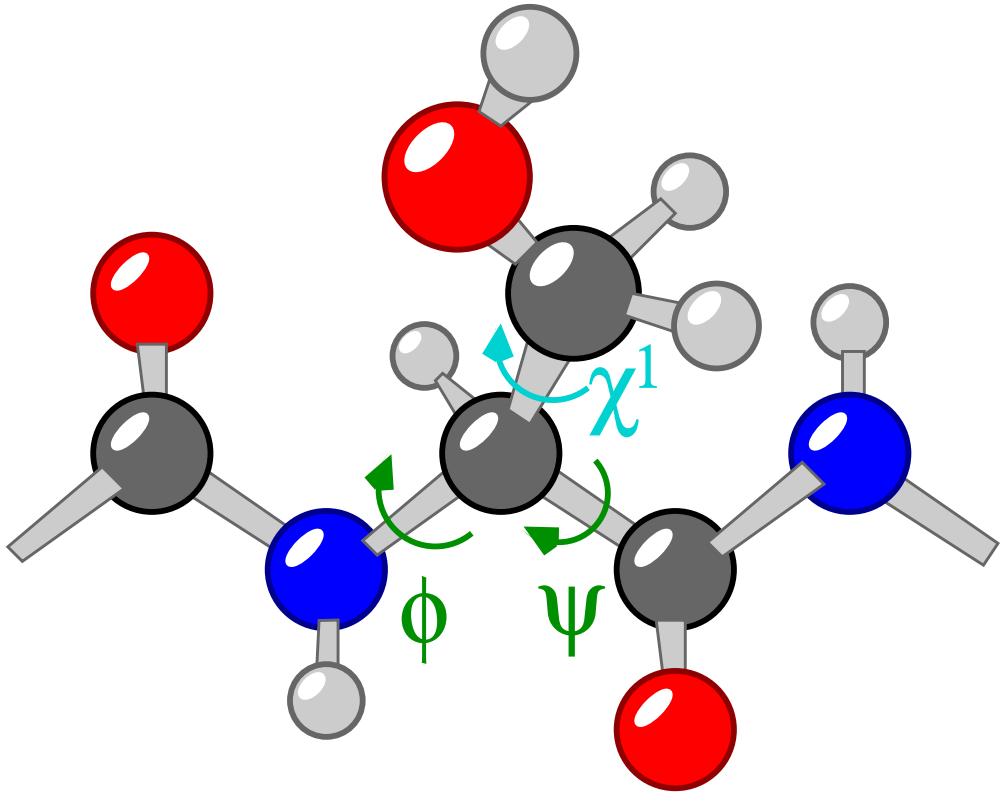




Ramachandran diagram

Repulsion including side chains (all, side chain dependent)





Interactions of/in charged and polar molecules

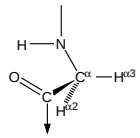
- Charged molecules (ions)
- Neutral polar molecules

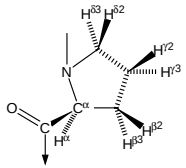
Charged groups (ions):

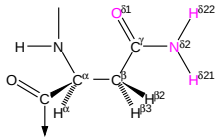
$$F = \frac{1}{4\pi\epsilon_0}\frac{Q_1 Q_2}{r^2} \qquad U = \frac{1}{4\pi\epsilon_0}\frac{Q_1 Q_2}{r}$$

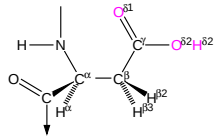
$\Delta G = 460\,\text{kJ/mol}$ for charges $0.3\,\text{nm}$ appart

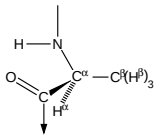
Charged amino acids

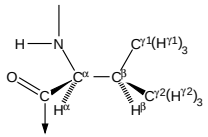

Gly (G)

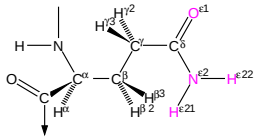

Pro (P)

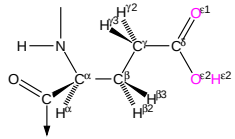

Asn (N)

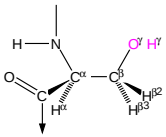

Asp (D)

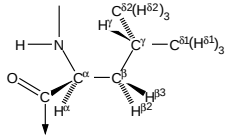

Ala (A)

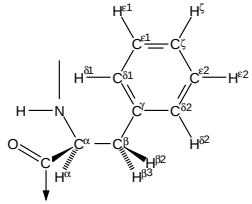

Val (V)

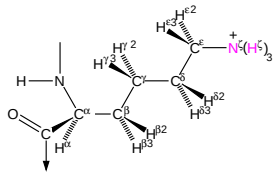

Gln (Q)

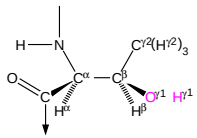

Glu (E)

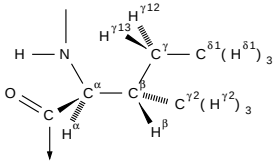

Ser (S)

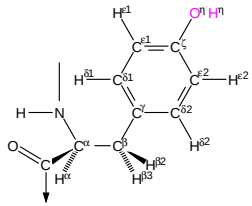

Leu (L)

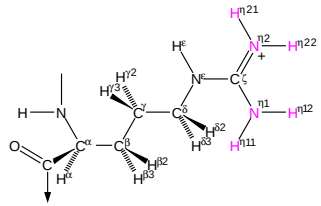

Phe (F)

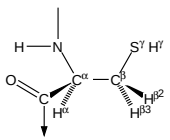

Lys (K)

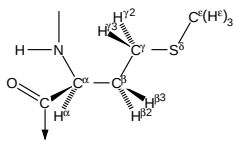

Thr (T)

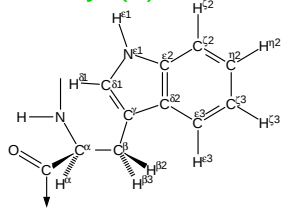

Ile (I)

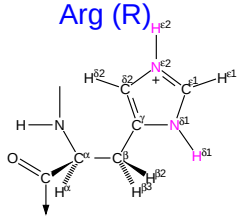

Tyr (Y)


Arg (R)

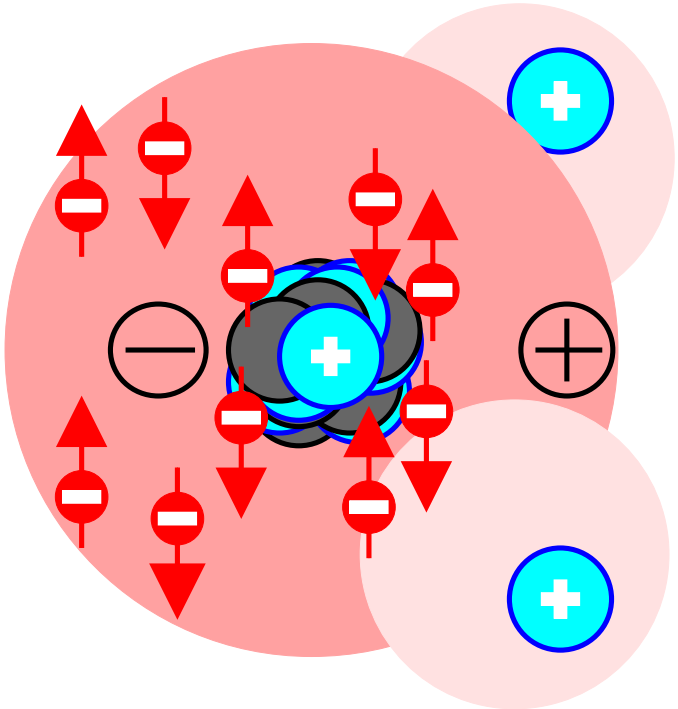

Cys (C)

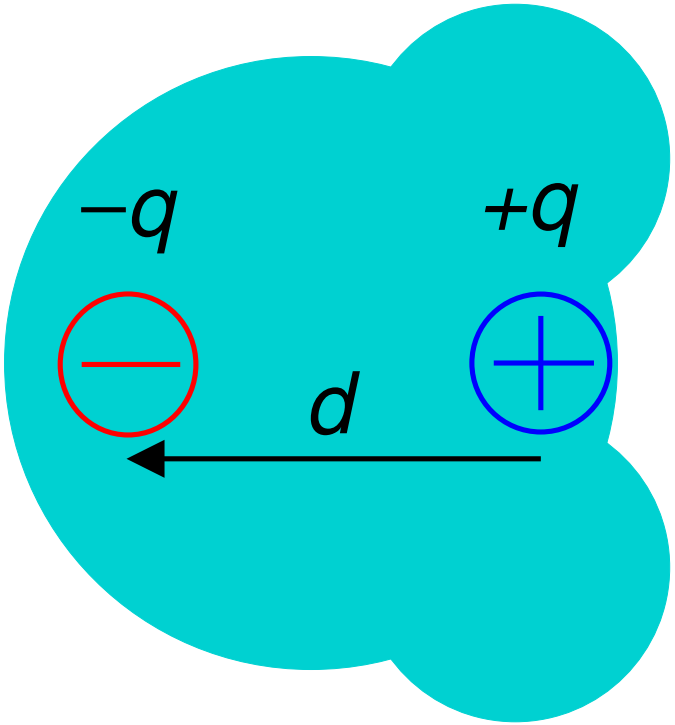

Met (M)


Trp (W)


His (H)

Polar molecules



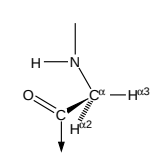


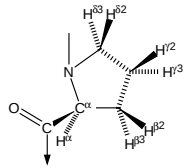
Permanent electric dipoles:

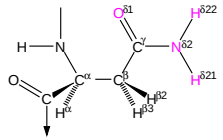
zero net charge but partial charges $\pm q$ separated by distance d

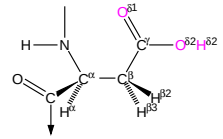
polar groups in molecules

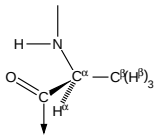
Polar amino acids

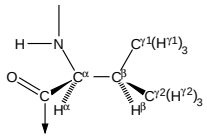

Gly (G)

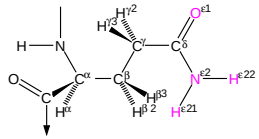

Pro (P)

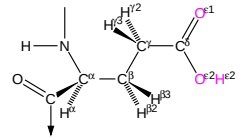

Asn (N)

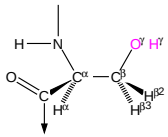

Asp (D)

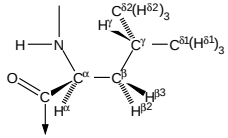

Ala (A)

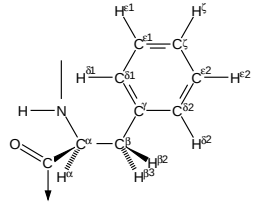

Val (V)

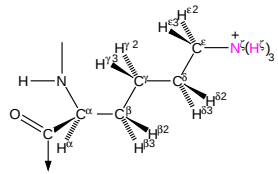

Gln (Q)

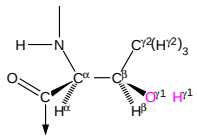

Glu (E)

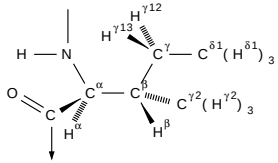

Ser (S)

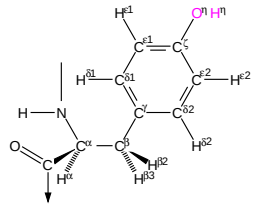

Leu (L)

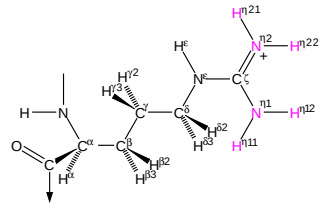

Phe (F)

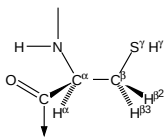

Lys (K)

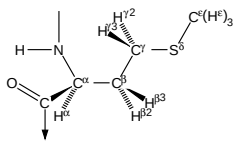

Thr (T)

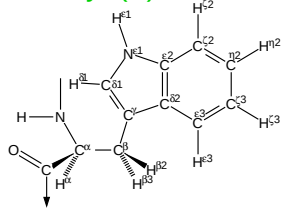

Ile (I)

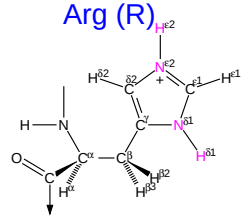

Tyr (Y)


Arg (R)

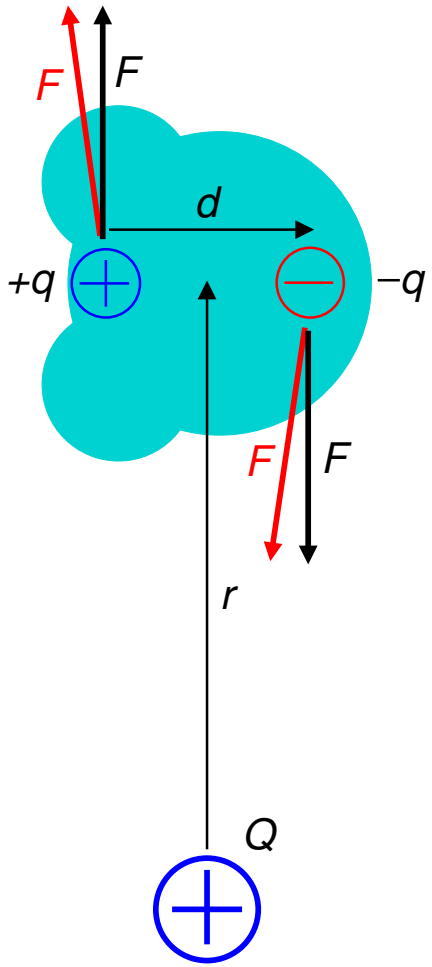

Cys (C)

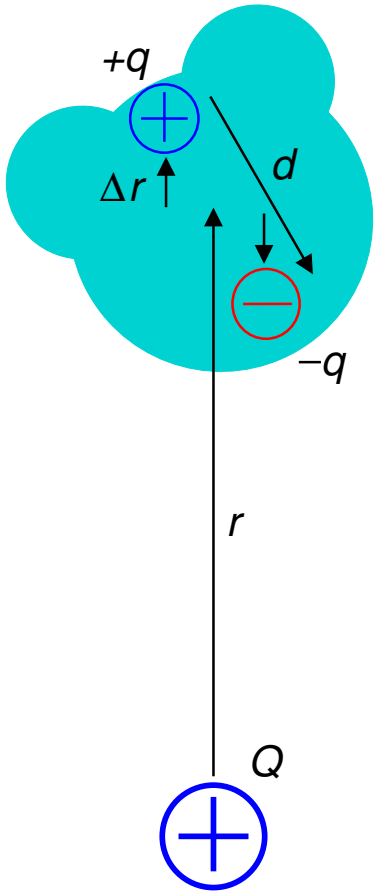

Met (M)

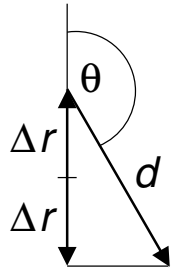

Trp (W)


His (H)

Permanent electric dipoles







$2\Delta r = -d\cos(\pi - \theta)$

$2\Delta r = d\cos\theta$

$F = F$

$d << r$

Permanent electric dipoles

Charge Q – permanent dipole $q \cdot d$
Charge and permanent dipole in the same molecule

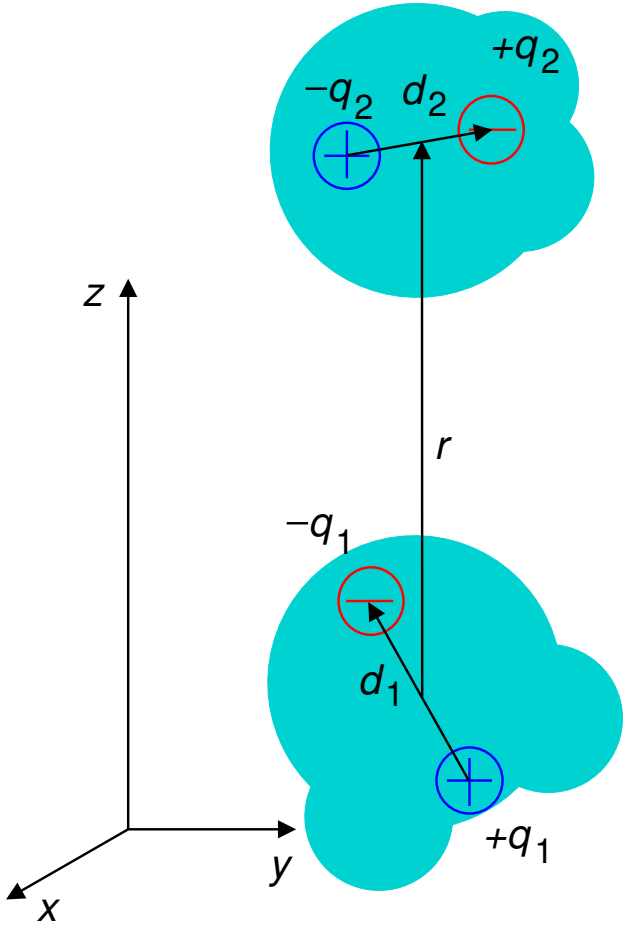
$$U = \frac{1}{4\pi\epsilon_0} \frac{qQ}{r} \frac{d}{r} \cos\theta$$

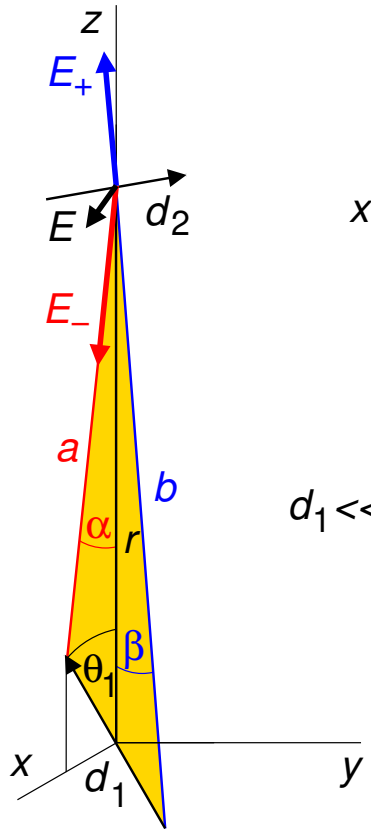
Charge and permanent dipole in different molecules

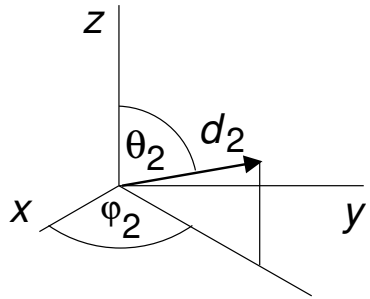
$$\langle U \rangle = -\frac{1}{3RT} \left(\frac{1}{4\pi\epsilon_0} \frac{qQ}{r} \frac{d}{r} \right)^2$$

Derived in Žídek: Strukturní biochemie, dodatek A

Permanent electric dipoles







$d_1 << r \quad d_2 << r$

Permanent electric dipoles

Permanent dipole $q_1 \cdot d_1$ – permanent dipole $q_2 \cdot d_2$
Permanent dipoles in the same molecule

$$U = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r} \frac{\vec{d}_1}{r} \frac{\vec{d}_2}{r} (\sin\theta_1 \sin\theta_2 \cos(\phi_1 - \phi_2) - 2\cos\theta_1 \cos\theta_2)$$

Permanent dipoles in different molecules

$$\langle U \rangle = -\frac{2}{3RT} \left(\frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r} \frac{d_1}{r} \frac{d_2}{r} \right)^2$$

Derived in Žídek: Strukturní biochemie, dodatek A

Induced dipoles

induced dipole is proportional to inducing force: $q\vec{d} = \alpha\epsilon_0\vec{E}$

Charge Q – induced dipole $\alpha\epsilon_0\vec{E}$

$$\langle U \rangle = -\frac{\alpha\epsilon_0}{2} N_A \left(\frac{1}{4\pi\epsilon_0}\right)^2 \frac{Q^2}{r^2}$$

Permanent dipole $q \cdot d$ – induced dipole $\alpha\epsilon_0\vec{E}$

$$\langle U \rangle = -\frac{\alpha\epsilon_0}{2} N_A \left(\frac{1}{4\pi\epsilon_0}\right)^2 \frac{q^2}{r^2}\frac{d^2}{r^2}(1 + 3\cos^2\theta)$$

induced dipole $\alpha\epsilon_0\vec{E}$ – induced dipole $\alpha\epsilon_0\vec{E}$ (dispersion forces)

$$\langle U \rangle = -\frac{3h\nu N_A}{4} \left(\frac{\alpha\epsilon_0}{4\pi\epsilon_0}\right)^2 \frac{1}{r^6}$$

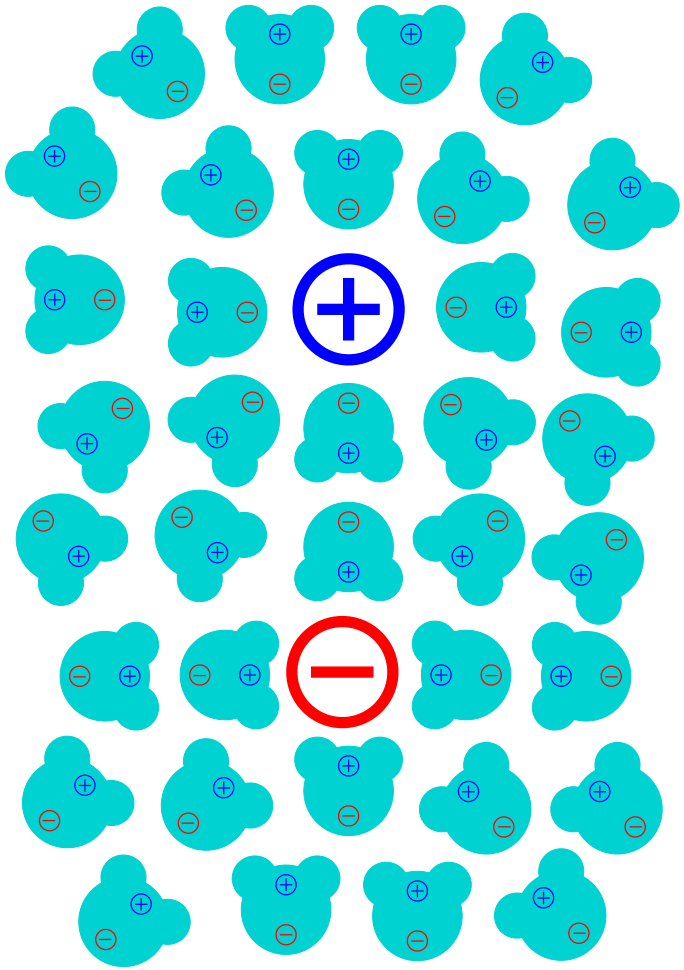
Derived in Žídek: Strukturní biochemie, dodatek A

Electrostatic interactions in proteins

- backbone (C=O, N–H $\Rightarrow$ dipole of α-helices)
- sidechains (nonpolar/polar/charged)
- **WATER**

Solvation of charges

Interaction of charges with water dipoles greatly reduces interaction between charges



Environment reduces electrostatic interactions

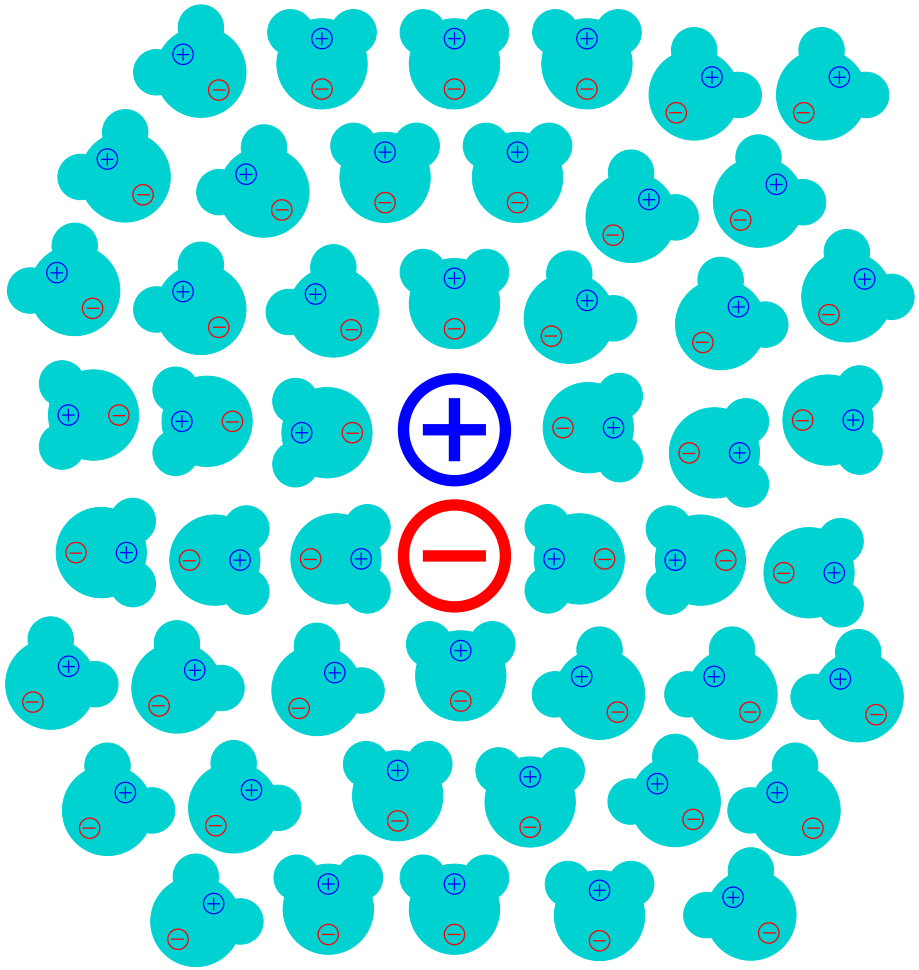
- polarization/orientation of atoms/groups in the molecule
- orientation of solvent molecules
- to maximize energy (enthalpy) of their electrostatic interactions at the cost of lowering entropy
- water does not work as an electrostatic "barrier"
- formally decreases constant in Coulomb's law
  $\Rightarrow$ increases $\epsilon_0 \rightarrow \epsilon_r \epsilon_0$

$$F = \frac{1}{4\pi\epsilon_r\epsilon_0}\frac{Q_1 Q_2}{r^2}$$

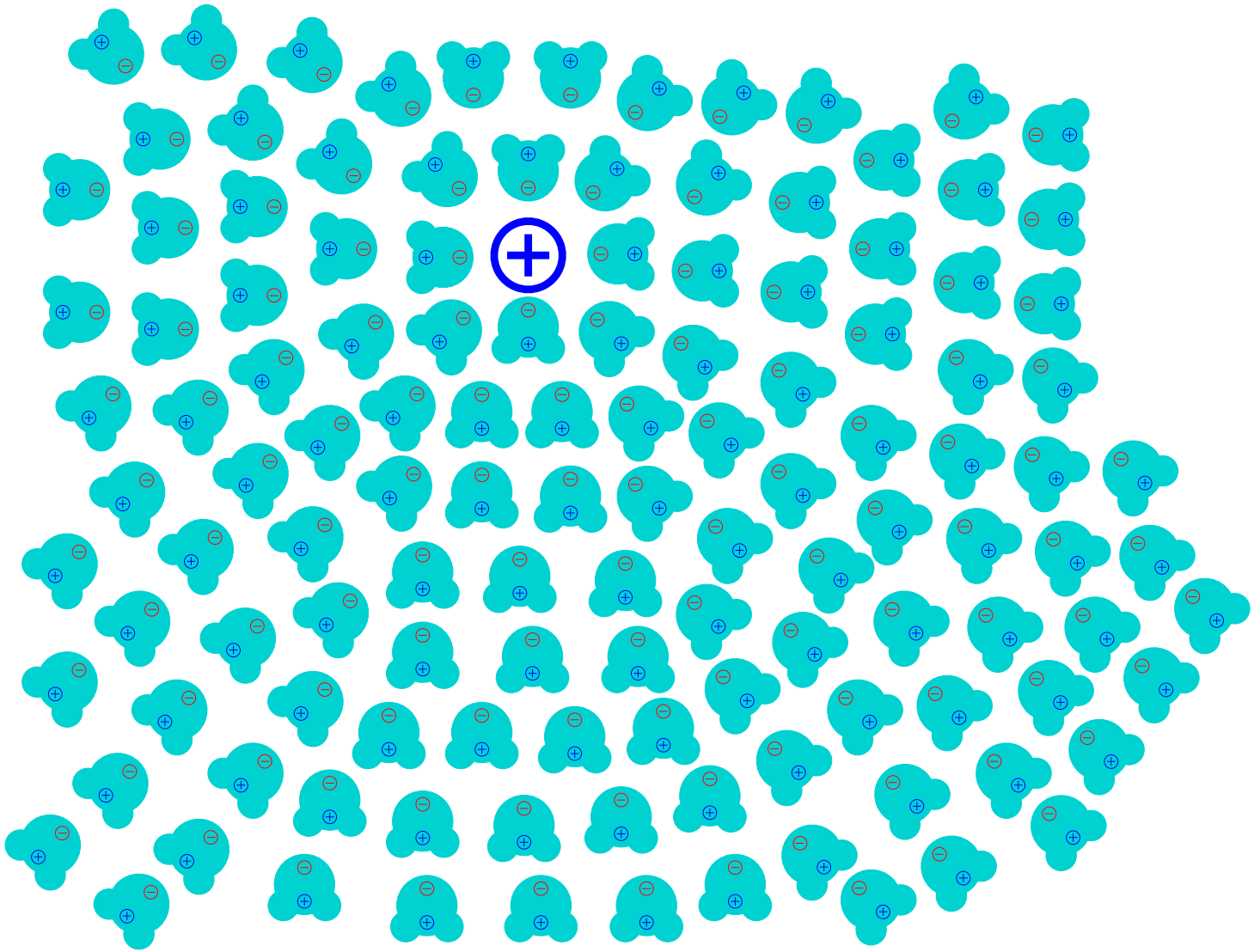
$\Delta G = 460\,\text{kJ/mol} \rightarrow 6\,\text{kJ/mol}$ for charges 0.3 nm appart

Solvation of charges

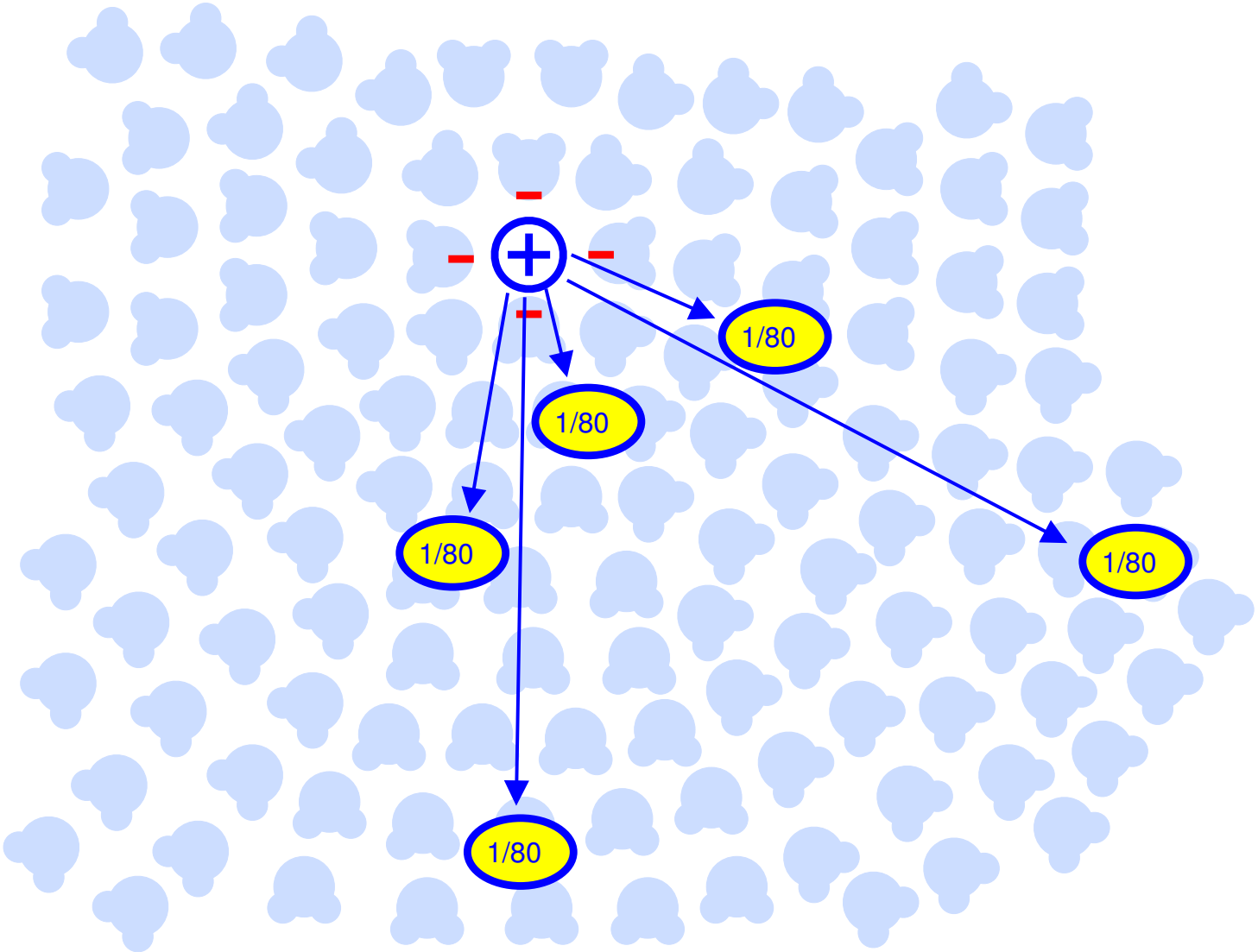
Effect of orientation of water molecules,
water does not need to be between charges



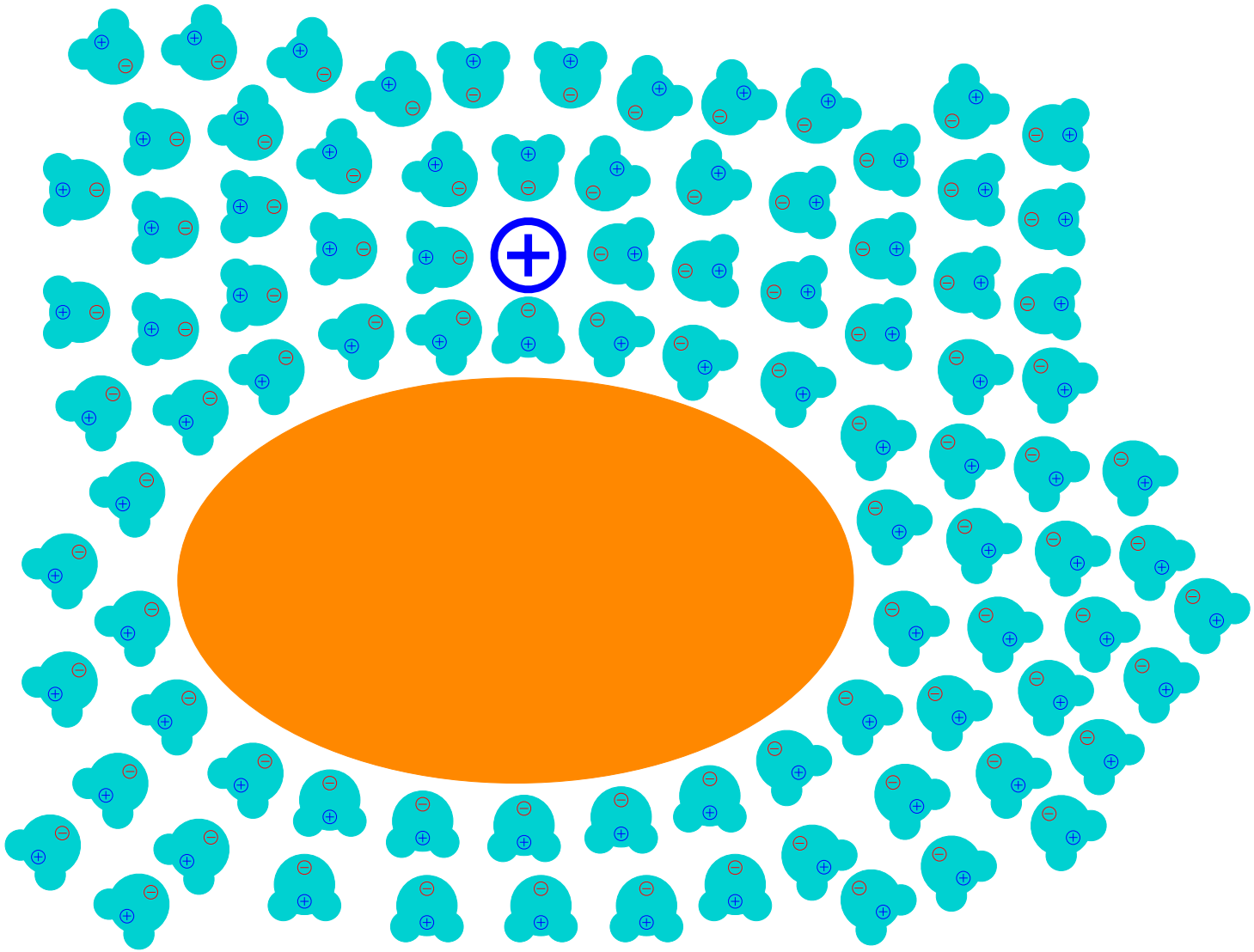
Interactions with charge in bulk water



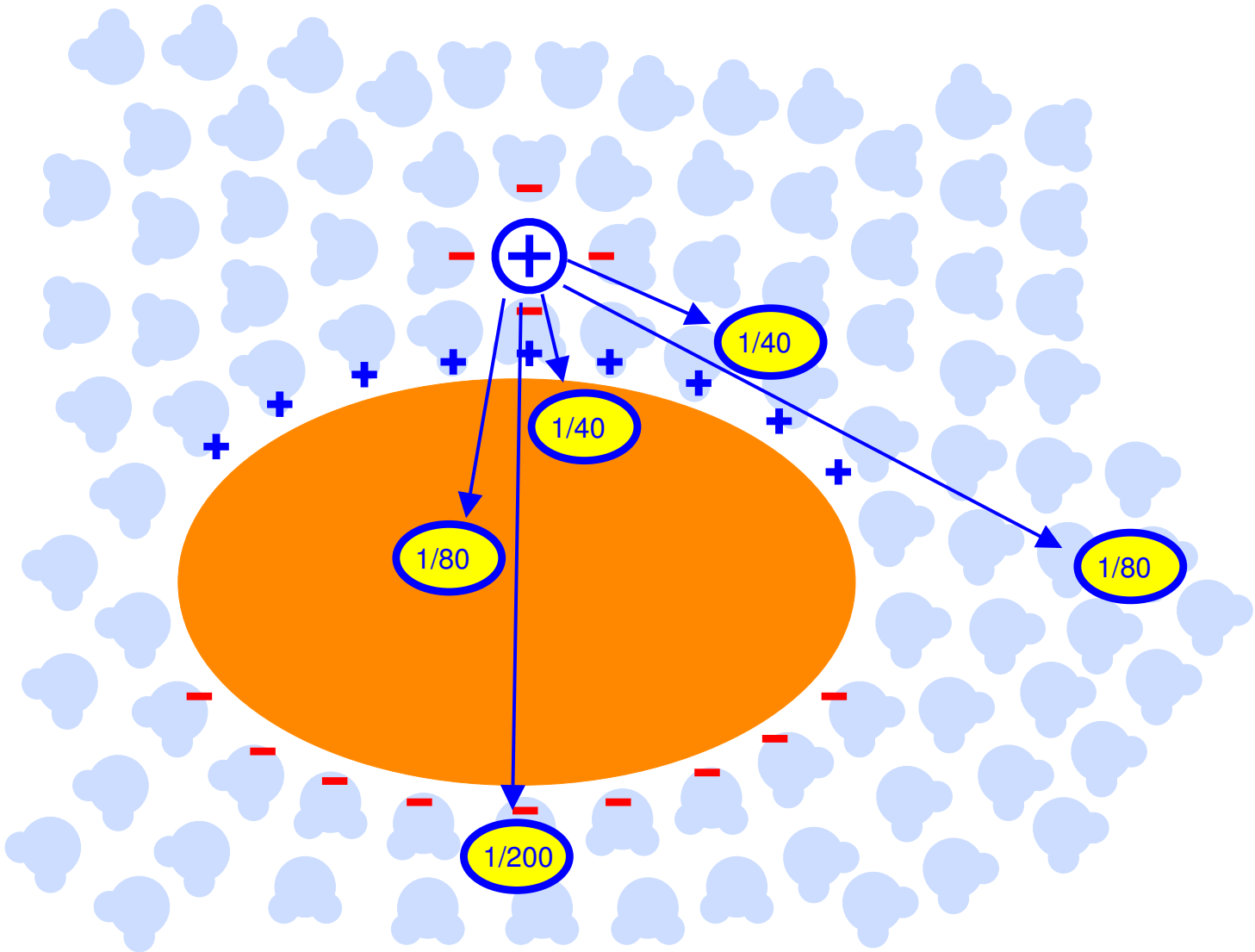
Interactions with charge in bulk water



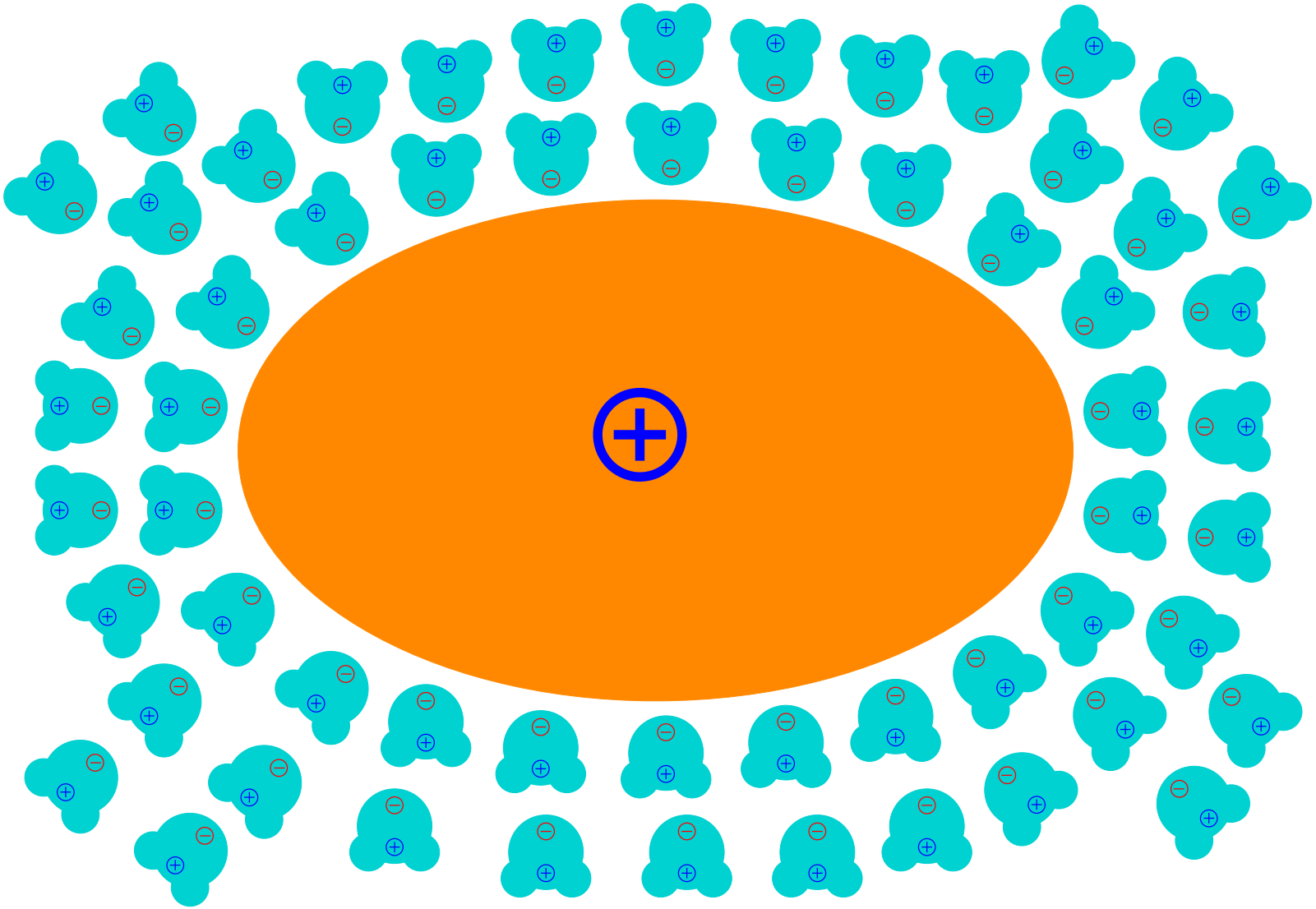
Interactions with charge at protein sufrace



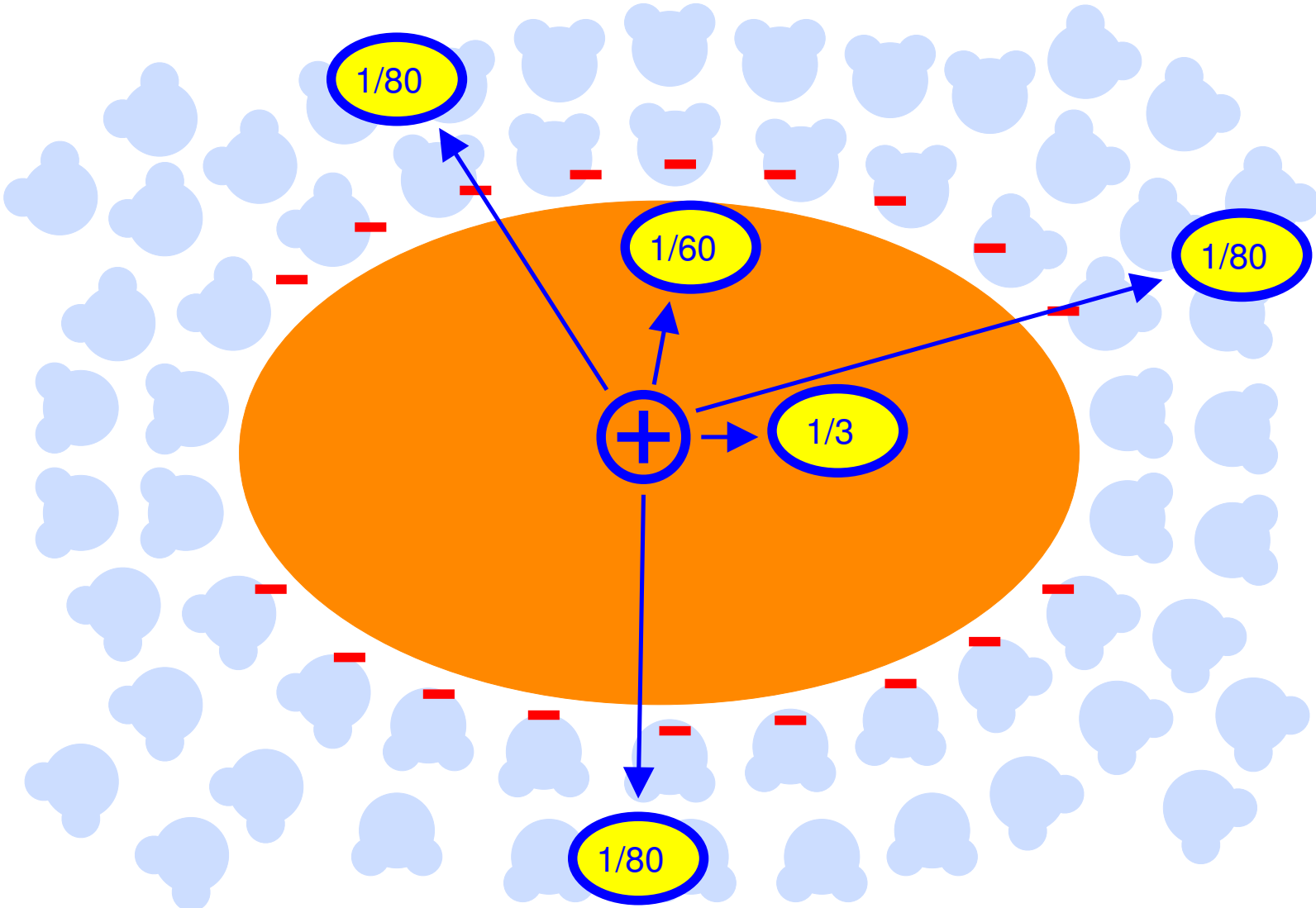
Interactions with charge at protein sufrace



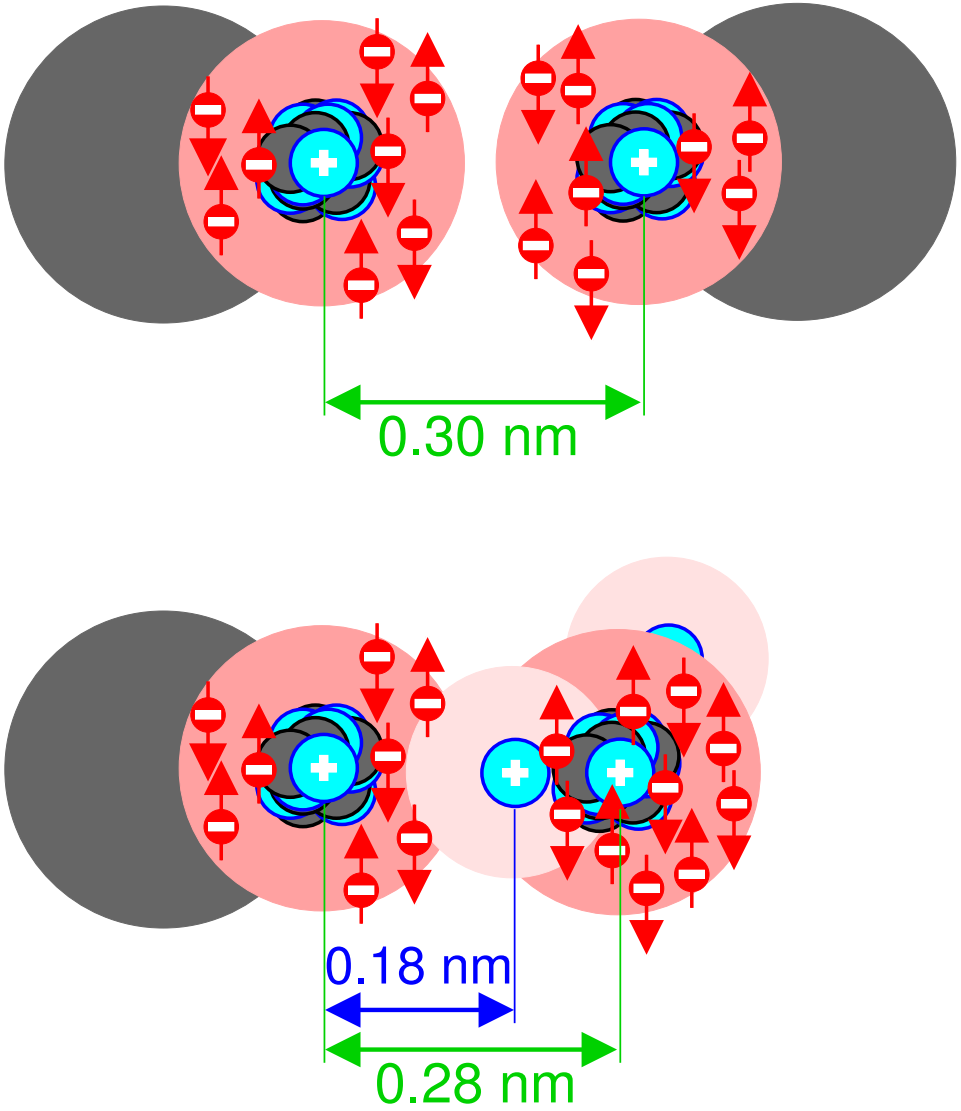
Interactions with charge inside protein



Interactions with charge inside protein



Hydrogen bonds



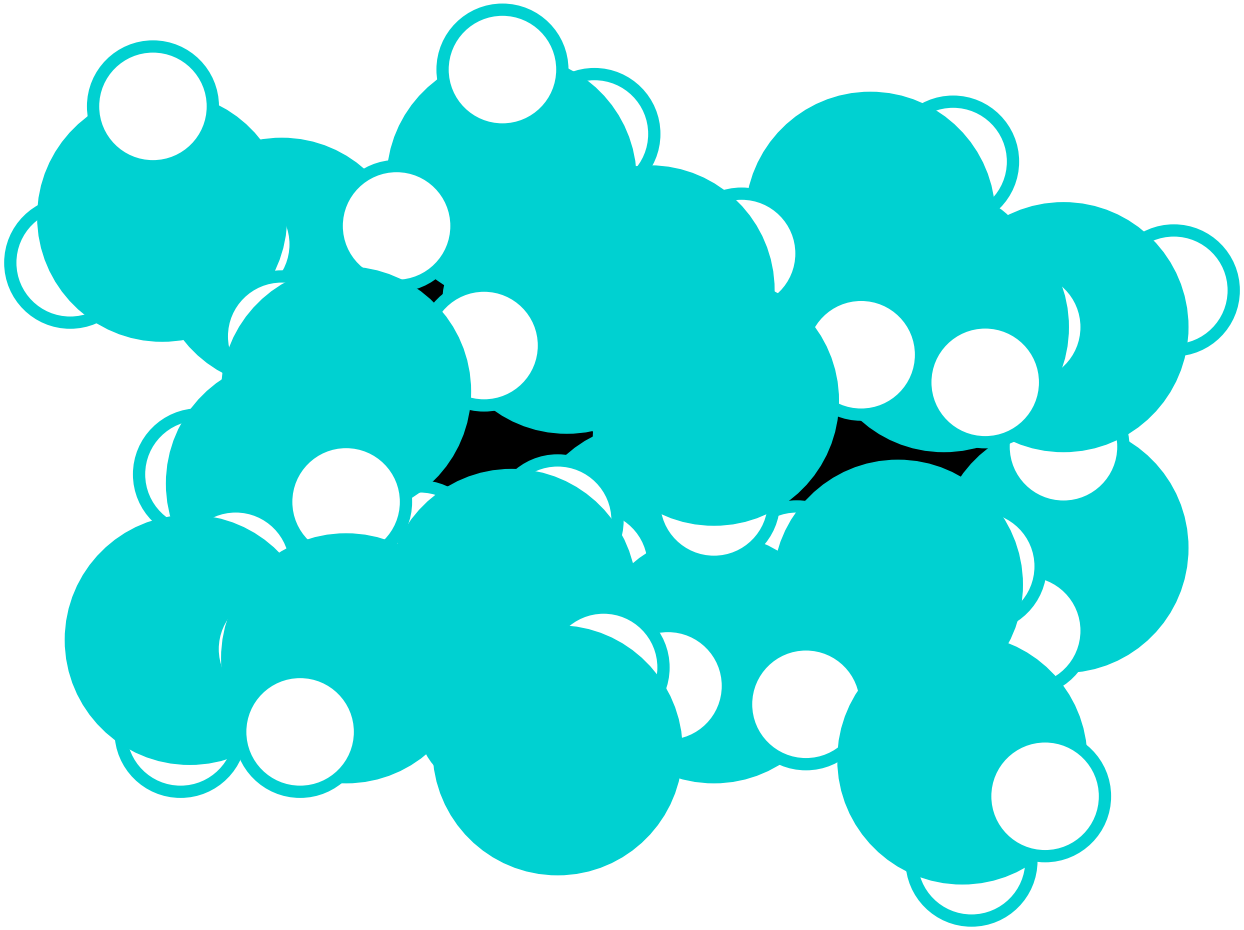
Hydrogen bonds

Hydrogen between atoms shortens their optimum distance

Atom···atom	U_{opt} / kJ mol^{-1}	r_{opt} / nm	r_{min} / nm
He···He	0.05	0.28	0.25
–H···H–	0.50	0.24	0.20
–C···C–	0.50	0.34	0.30
–N···N–	0.85	0.31	0.27
–NH···N–		0.31	
–O···O–	0.95	0.30	0.27
–OH···O–		0.28	

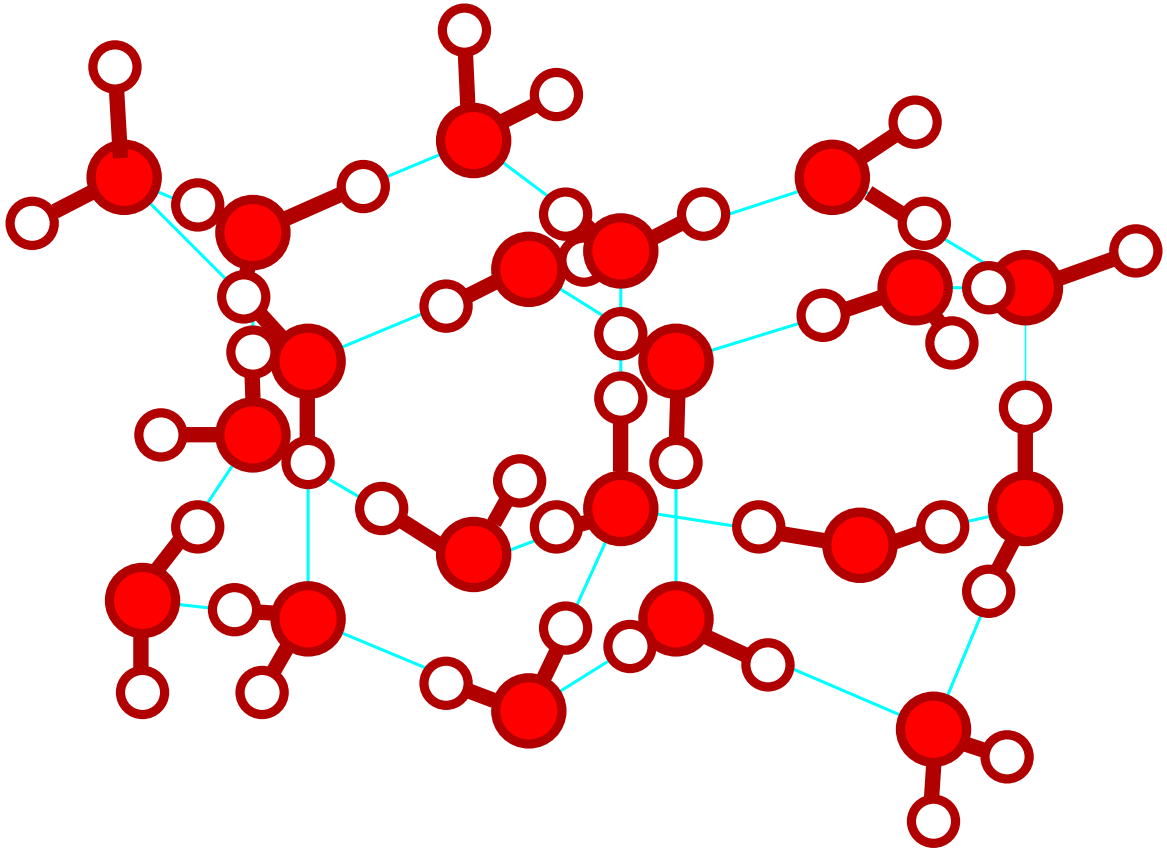
U(H-bond)$=$ 20 kJ/mol

Ice



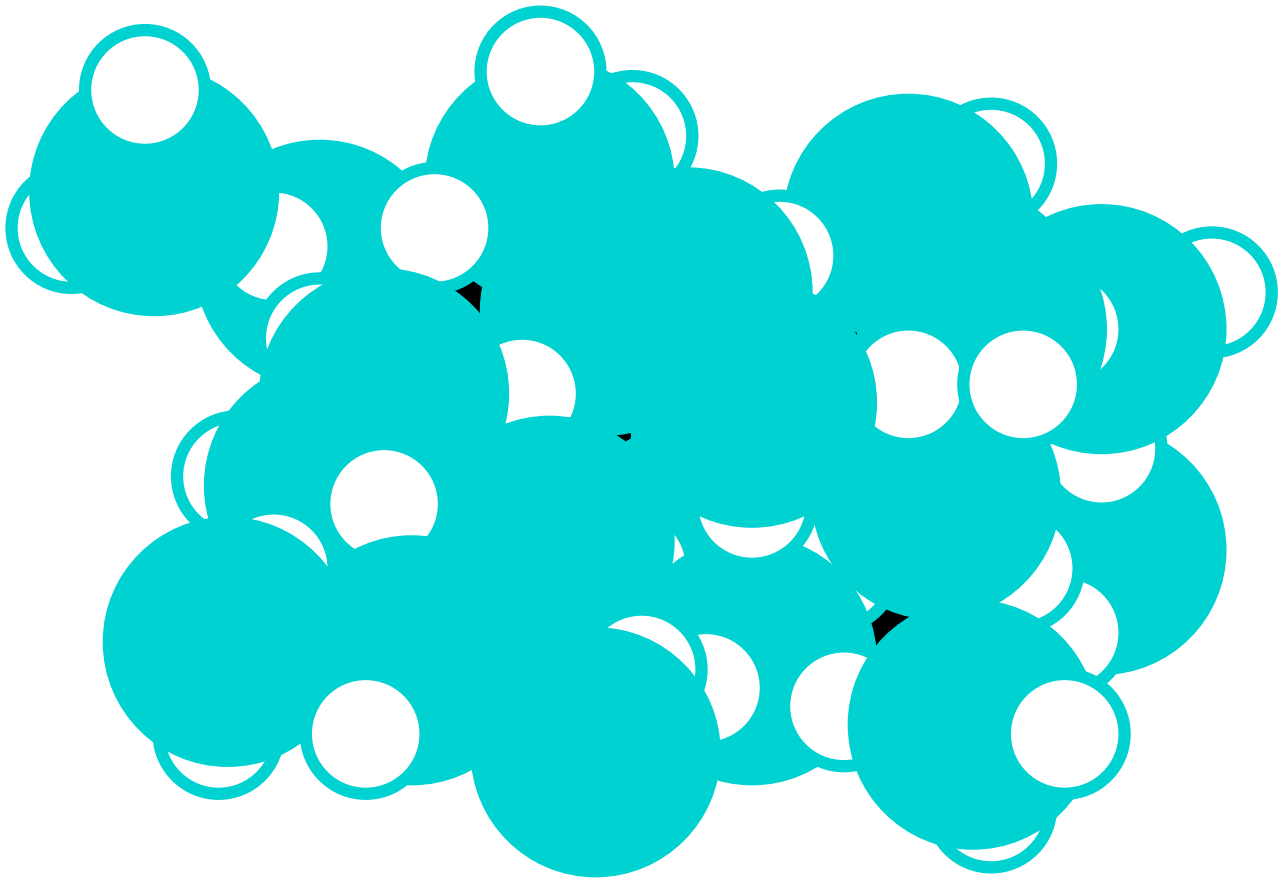
void space

Ice



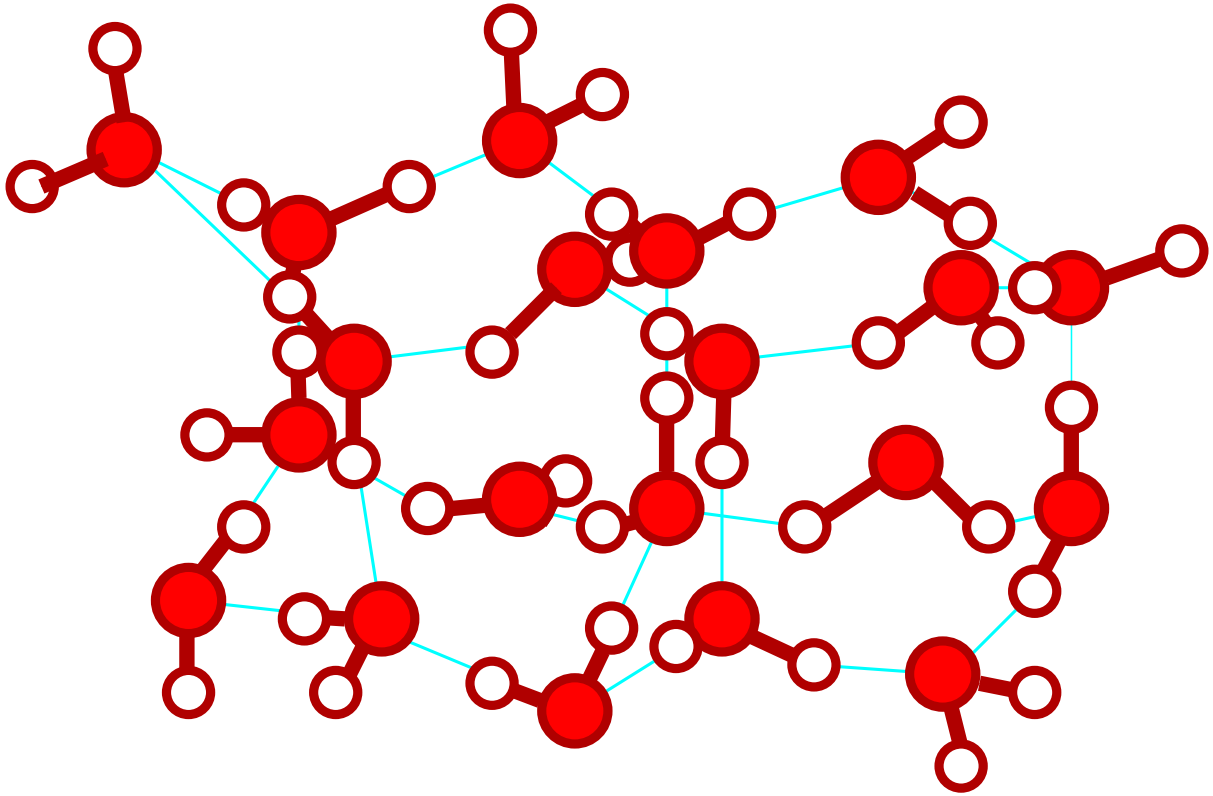
Hydrogen bonds : 50 kJ/mol

Liquid water



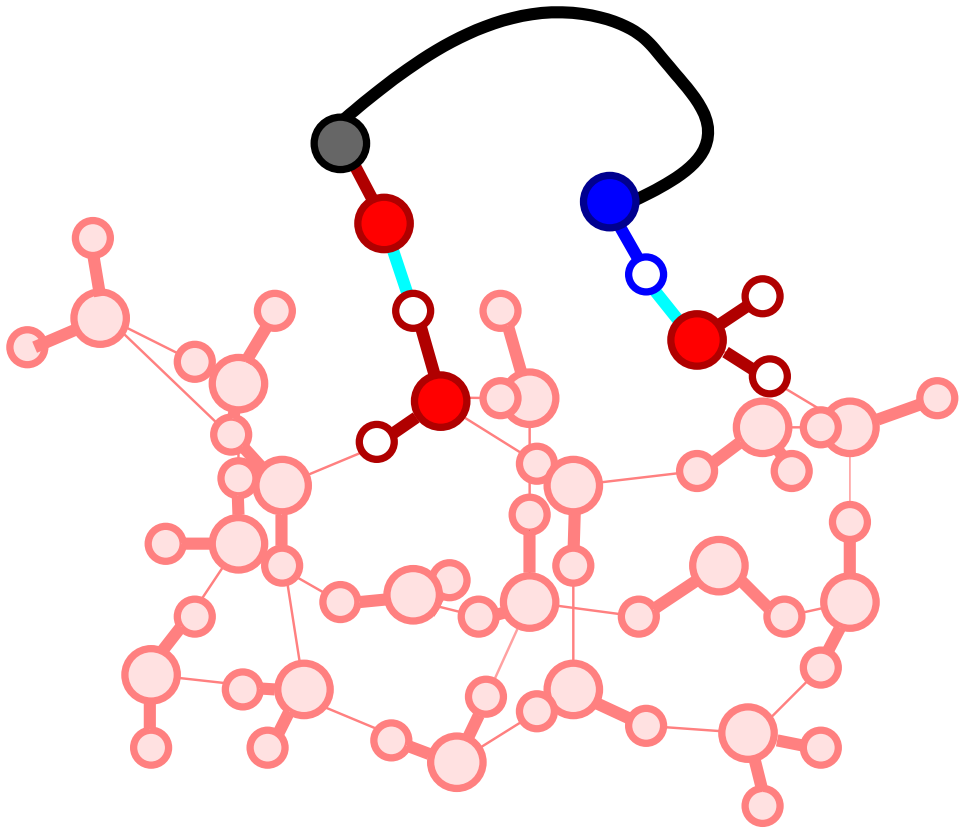
void space less than in ice

Liquid water



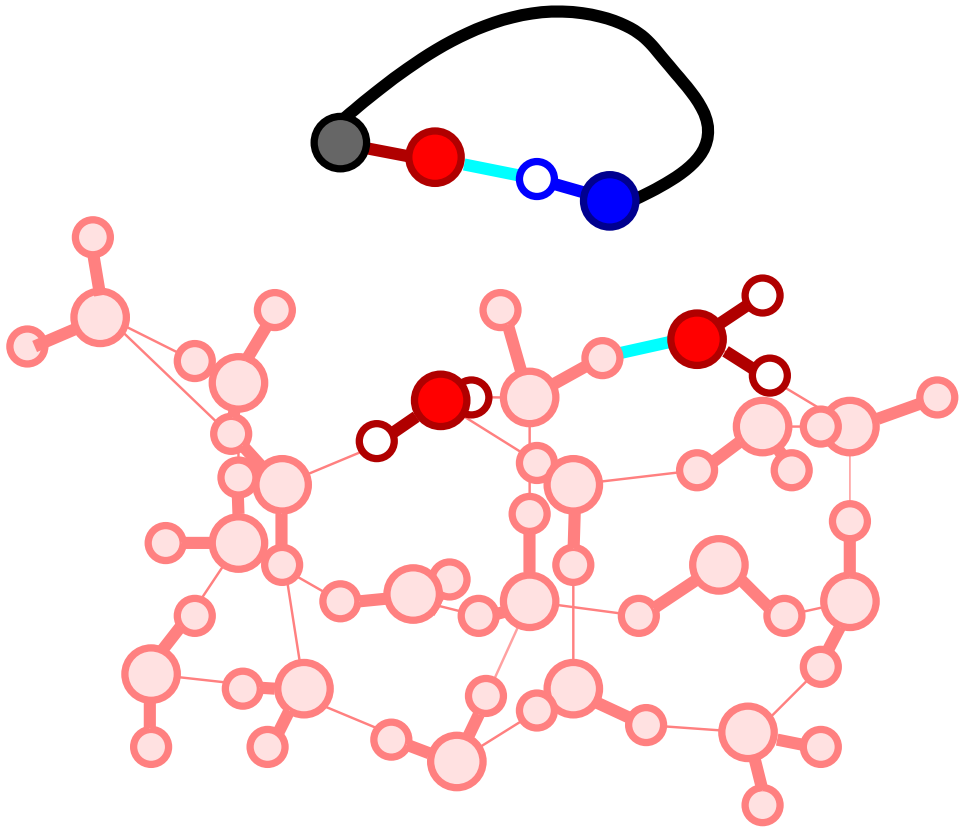
Hydrogen bonds : 40 kJ/mol

Hydrogen bonds in proteins



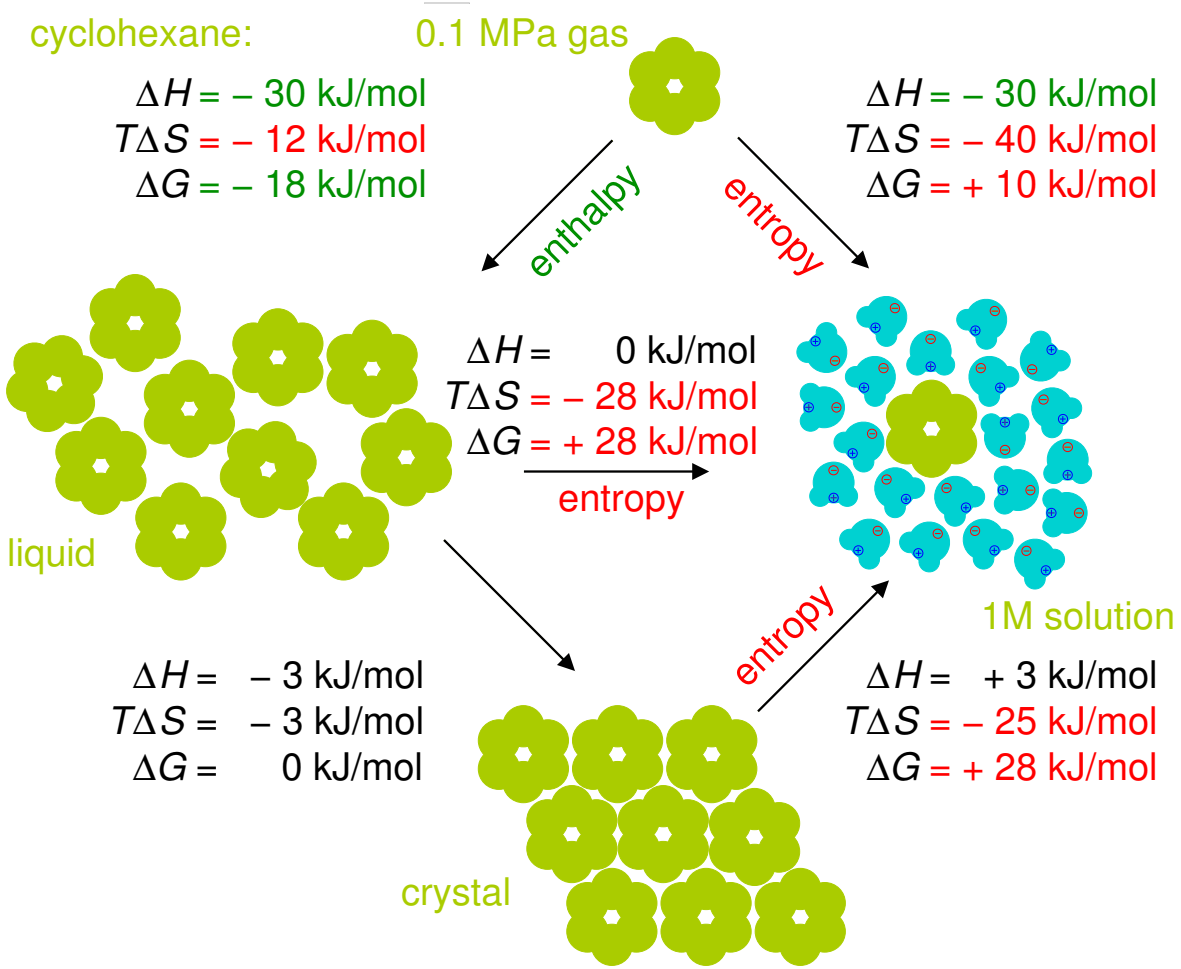
$\Delta G = 0$

Hydrogen bonds in proteins

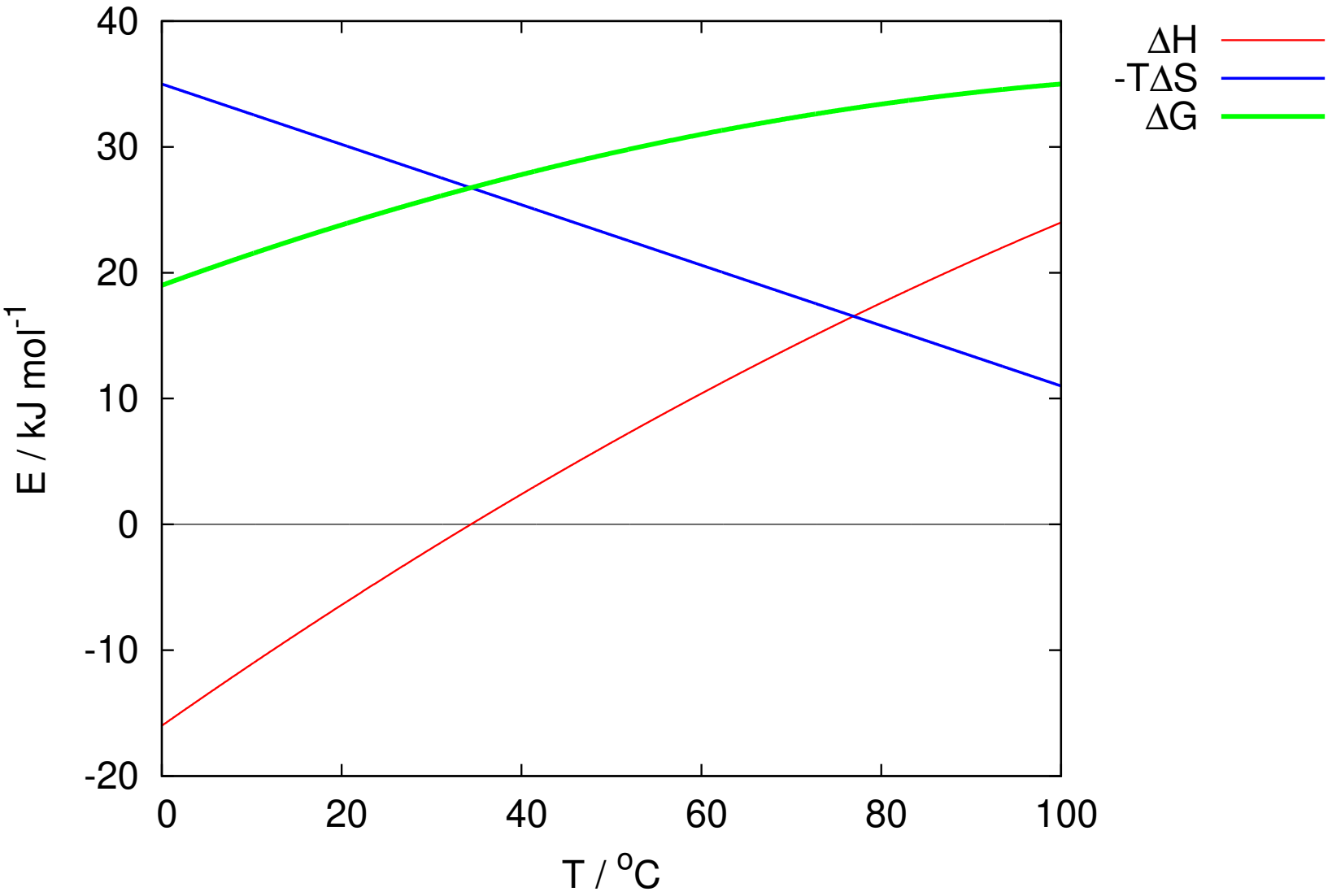


$\Delta G = -12\,\text{kJ/mol}$ (entropy of water)

Hydrophobic effect

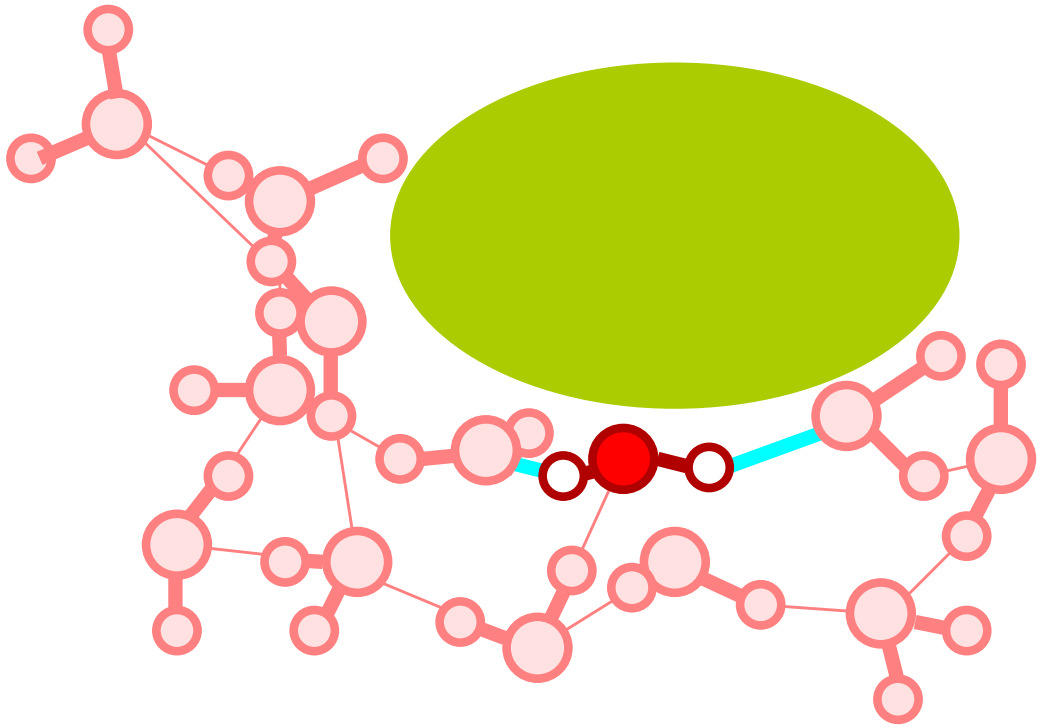


Hydrophobic effect



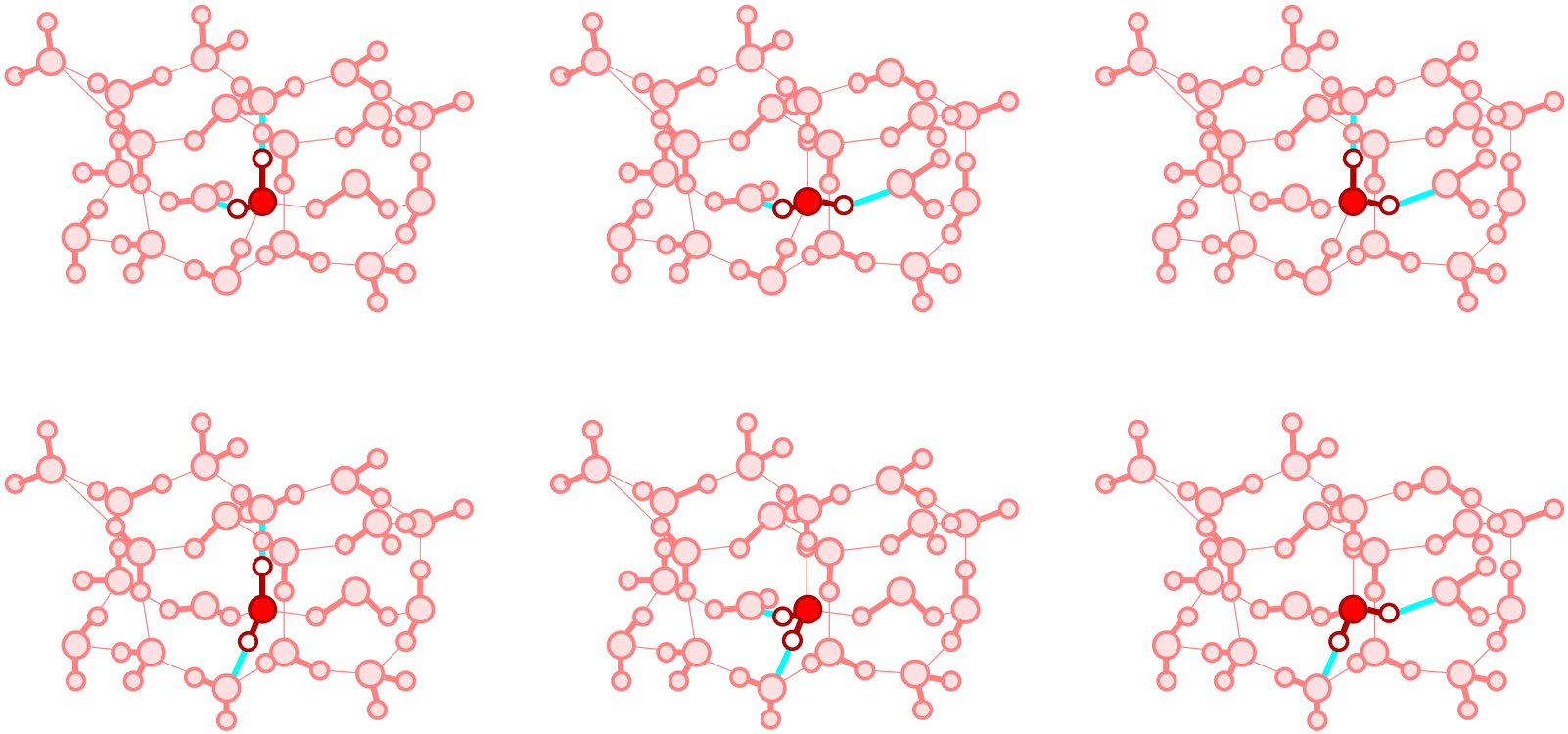
Hydrophobic effect

- orientation of solvent molecules
- to maximize energy (enthalpy) of their hydrogen bonds at the cost of lowering entropy



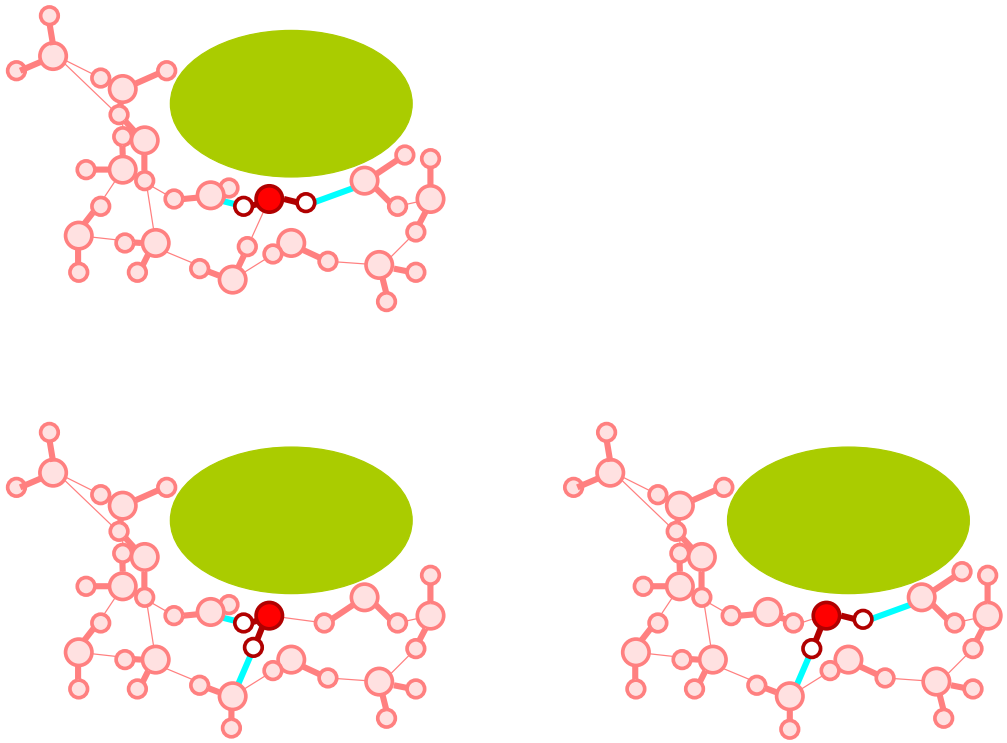
Hydrophobic effect

6 possible orientations:
entropic contribution $-RT \ln 6 = -15\,\text{kJ/mol}$

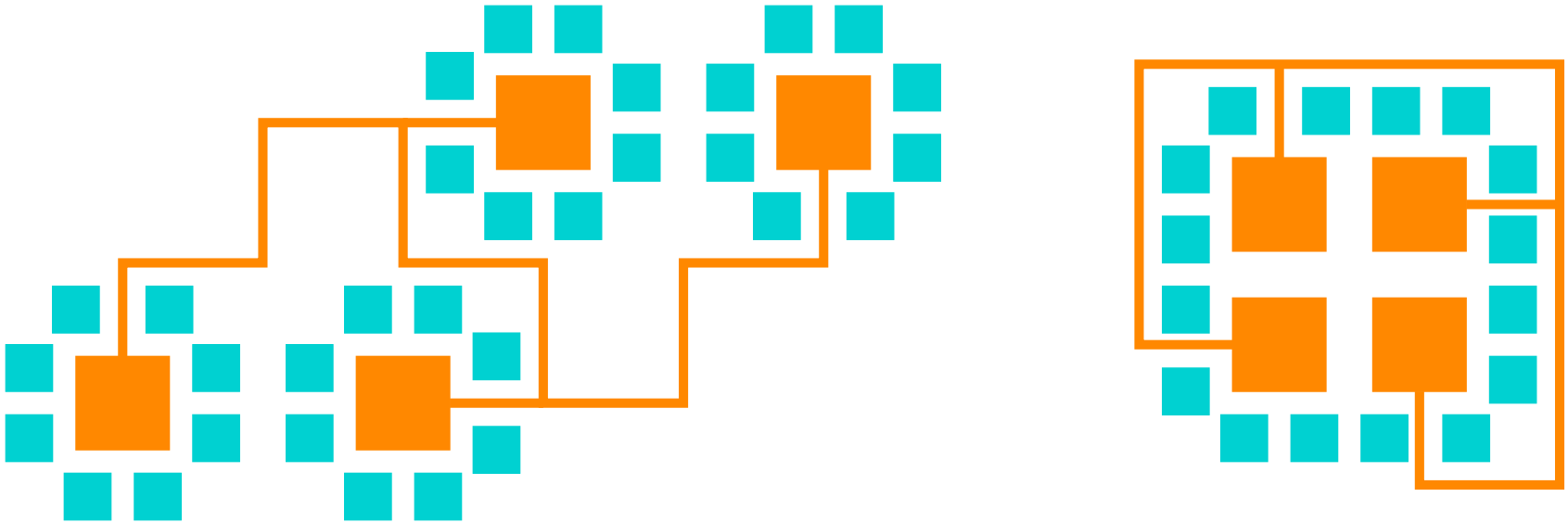


Hydrophobic effect

3 possible orientations:
entropic contribution $-RT \ln 3 = -7.5\,\text{kJ/mol}$



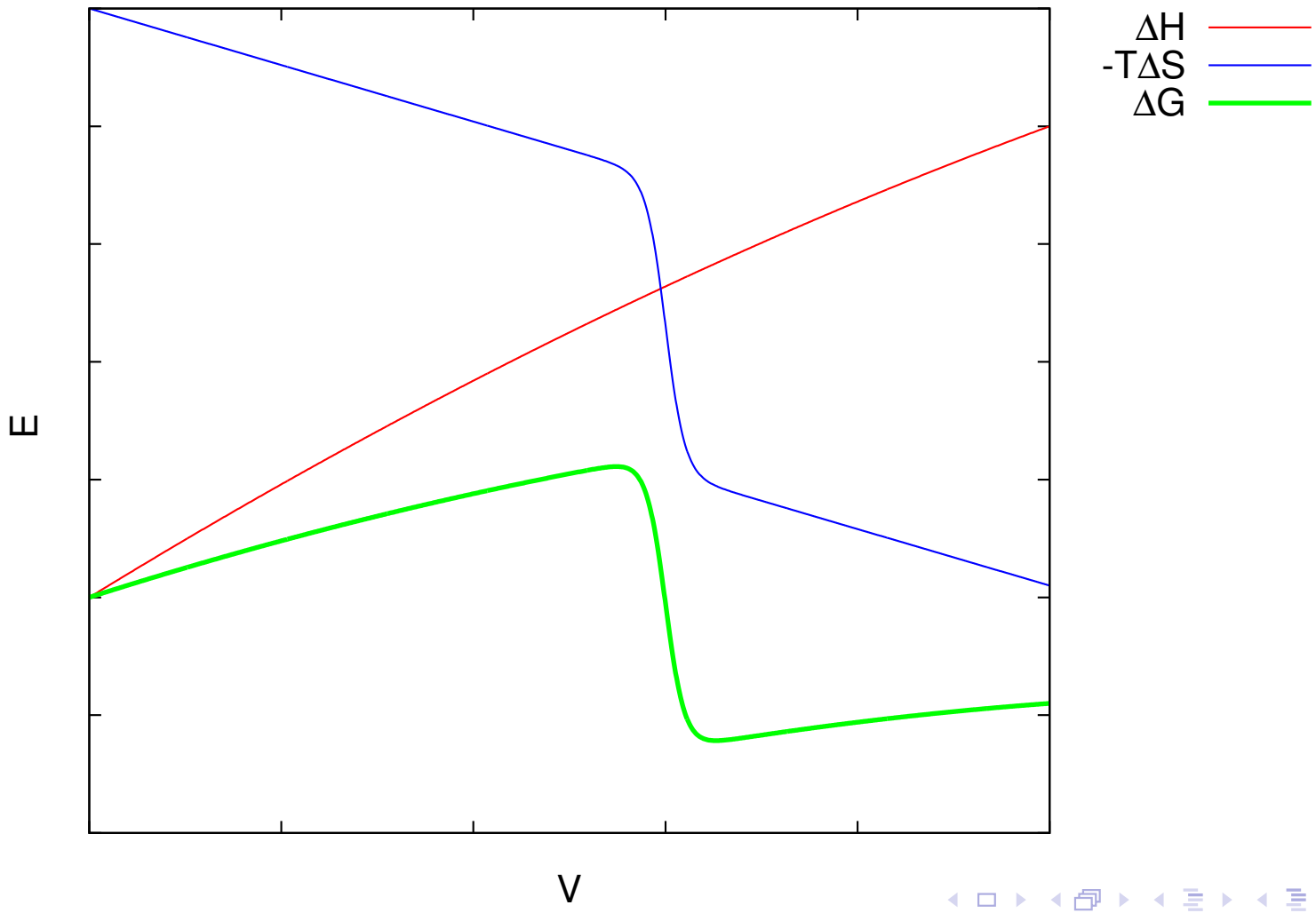
Hydrophobic effect in proteins



- packing nonpolar sidechains reduces entropy cost (less water molecules with restricted orientation)
- **the most important contribution to $-\Delta G$**
  Ala: 2.5 kJ/mol, Leu: 8 kJ/mol, Phe: 12 kJ/mol
- no specificity

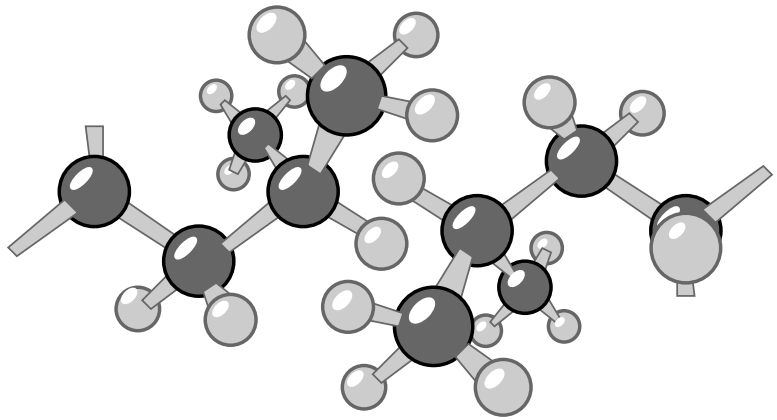
Protein stability

Loss of compactness = ↗ volume V during denaturation
High cooperativity (sharp drop of ΔG)



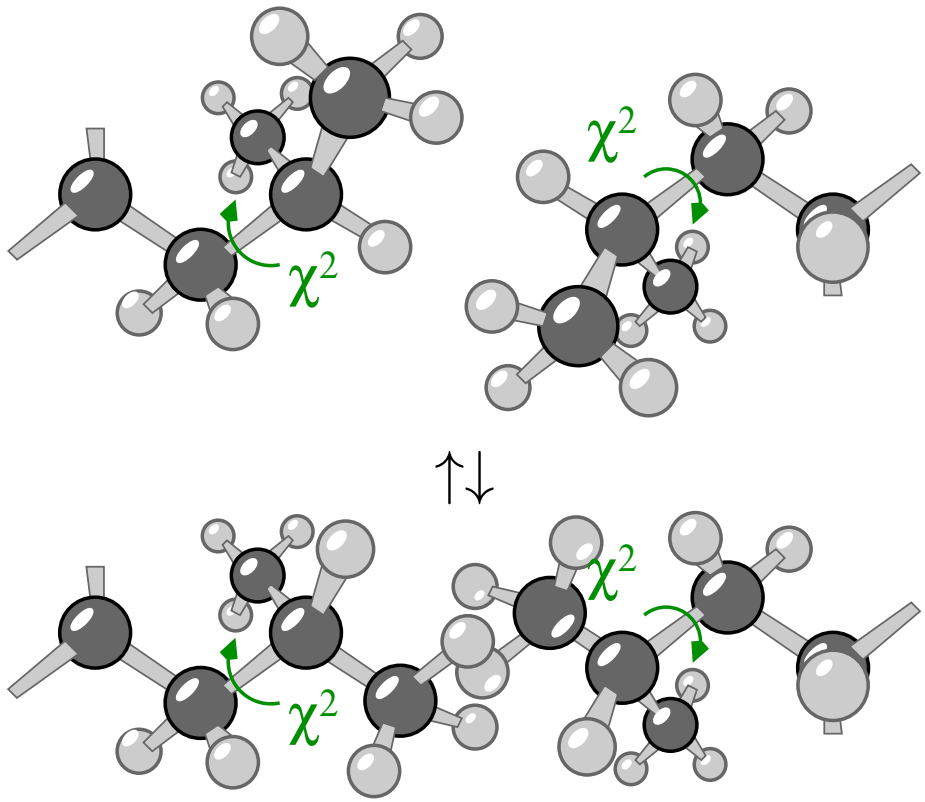
Protein stability

Packed side chains in compact folded proteins
No side chain rotation possible
1 side chain orientation: entropic contribution $-RT \ln 1 = 0$



Protein stability

Less compact protein (”molten globule”)
Reduced dispersion energy (less $-H \Rightarrow \Delta H > 0$)
but side chain rotation possible ($S \nearrow \Rightarrow -T\Delta S \ll 0$)



Comparison of energy (ΔG) values

Type	kJ/mol	condition
thermal RT	2.5	at 300 K (27 °C)
covalent bond	350	C–C
ion-ion	460	0.3 nm apart in vacuum
ion-ion	150	0.3 nm apart inside protein
ion-ion	12	0.3 nm apart at protein surface
dipole-dipole	30	0.3 nm apart in vacuum
dipole-dipole	10	0.3 nm apart inside protein
ion-dipole	41	0.5 nm apart in vacuum
ion-dipole	14	0.5 nm apart in protein
hydrogen bond	20	in vacuum ($\Delta G \approx \Delta H$)
hydrogen bond	6	in water ($\Delta G \approx -T\Delta S$)
hydrophobic effect	8	per Leu side chain
hydrophobic effect	12	per Phe side chain

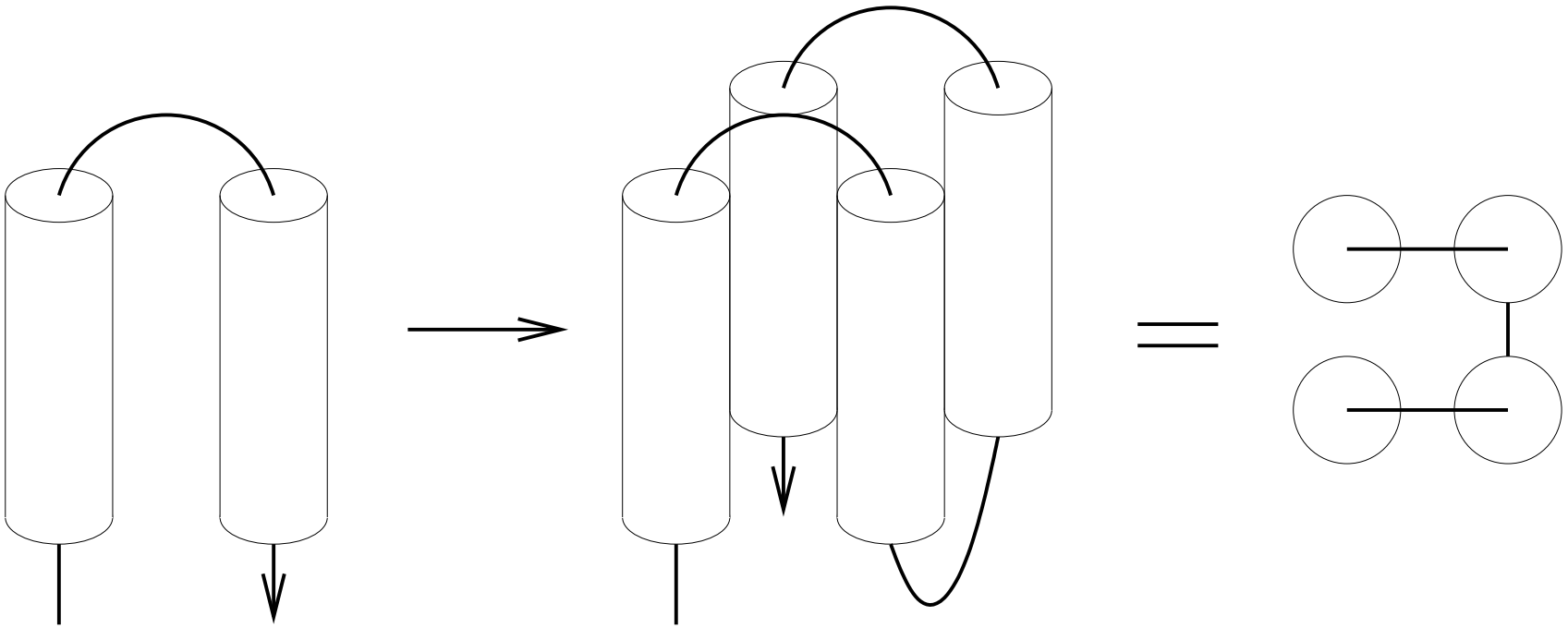
ion with charge $+1/-1$, dipole of peptide bond (1.2×10^{-29} C m)

Summary of interactions stabilizing proteins

- Covalent bonds define primary structure
- Disulfide bridges important outside cell
- Structures limited by steric requirements
- Dominant role of solvent (hydrophobic effect)
- Compaction due to hydrophobic effect
- Exact architecture due to electrostatics, hydrogen bonds

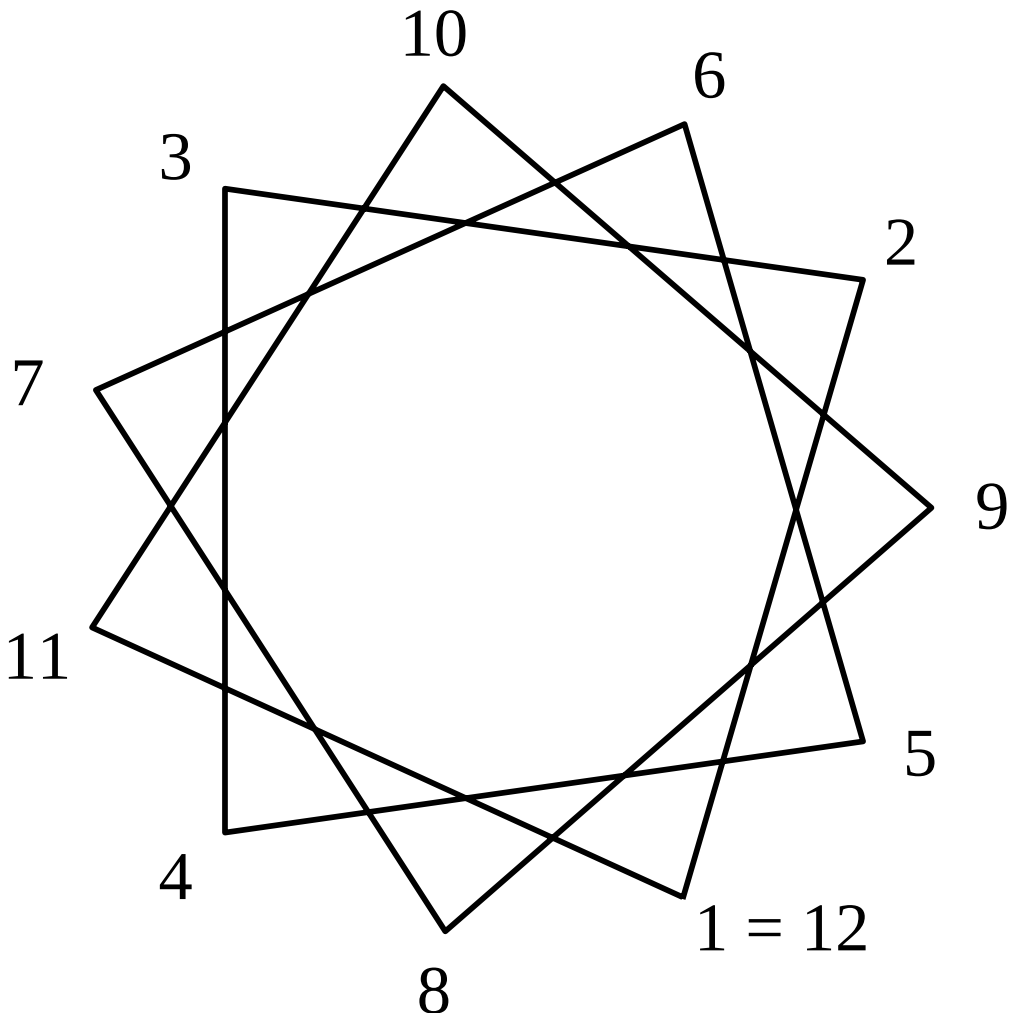
Packing of α-helices

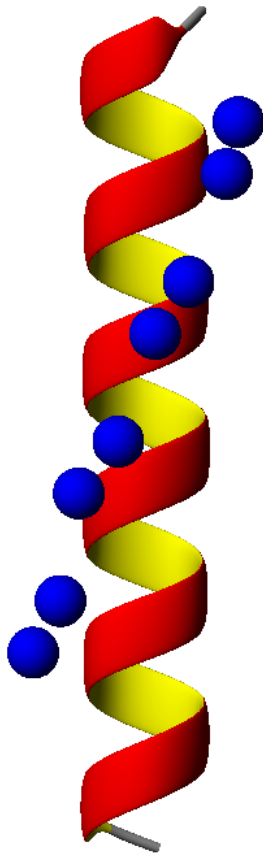
Amphiphilic helices: nonpolar sidechains inside



Packing of α-helices

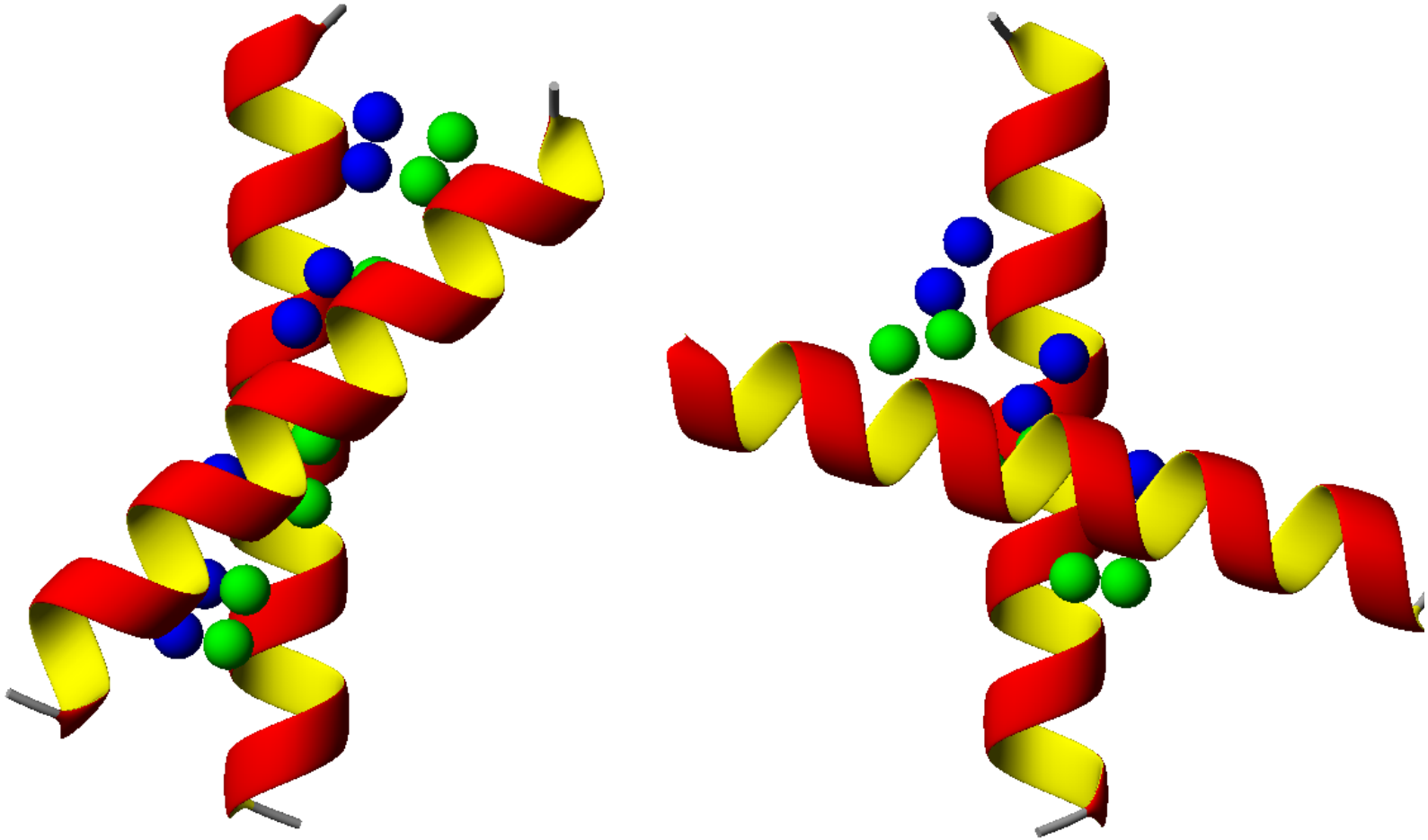
to compact nonpolar sidechains in amphiphilic helices





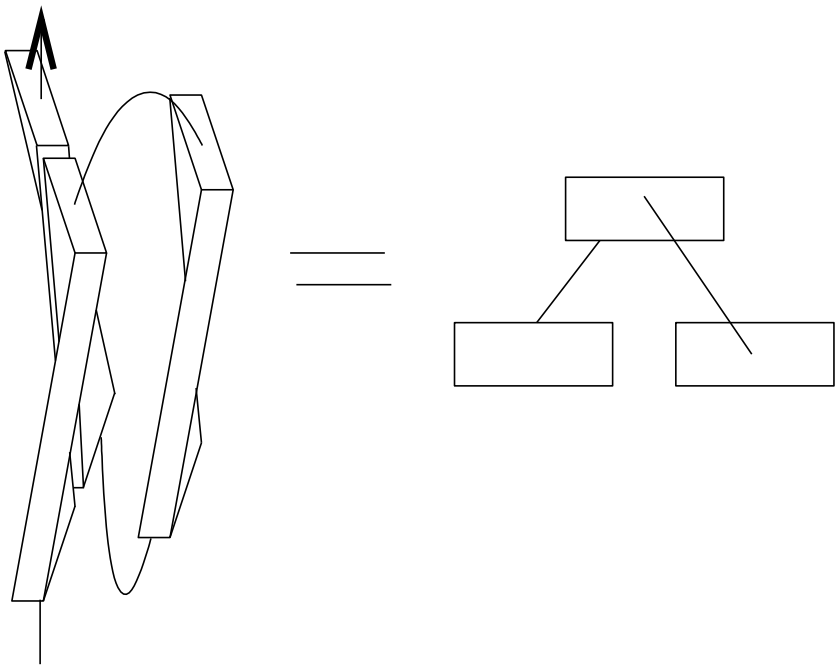
Packing of α-helices

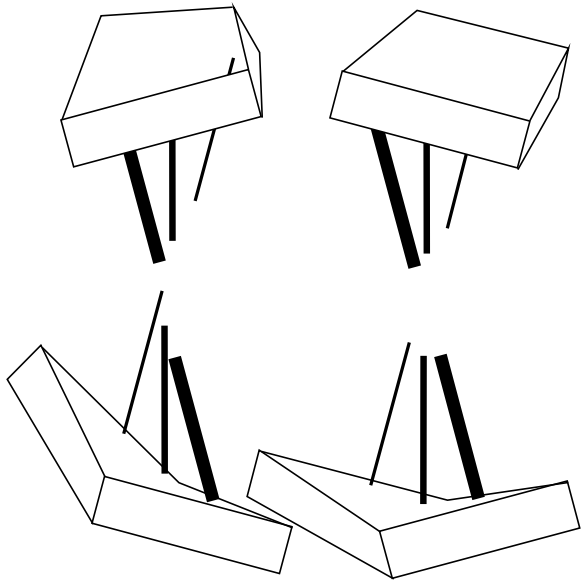
Hydrophobic side chains (blue/green) spheres packed positions 1, 5, 9, 13 (left) and 1, 4, 7 (right)



Packing of β-sandwiches

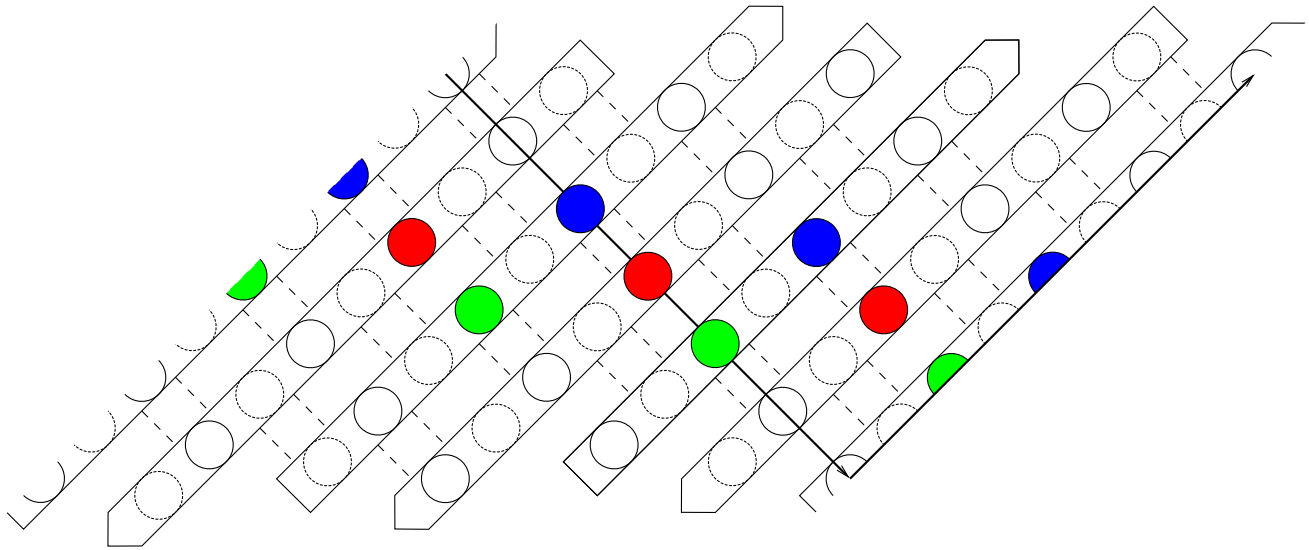
Hydrophobic residues inside

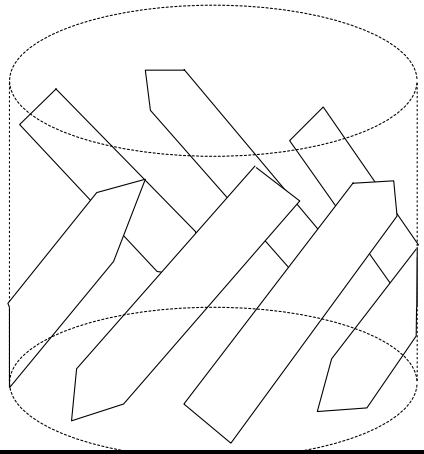


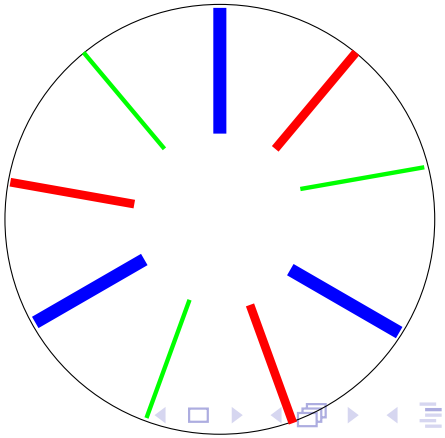


Packing of β-barrels

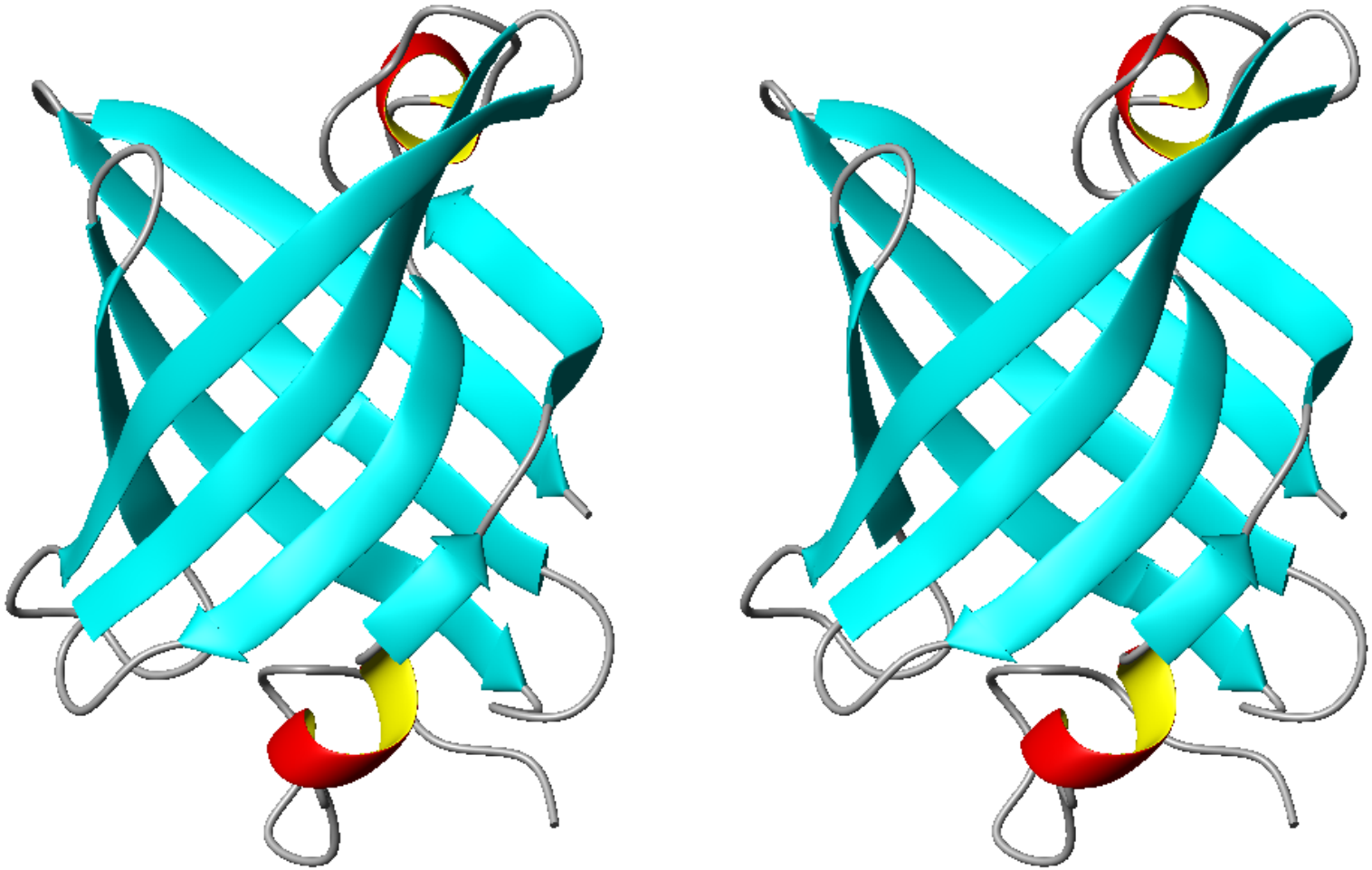
Hydrophobic residues inside



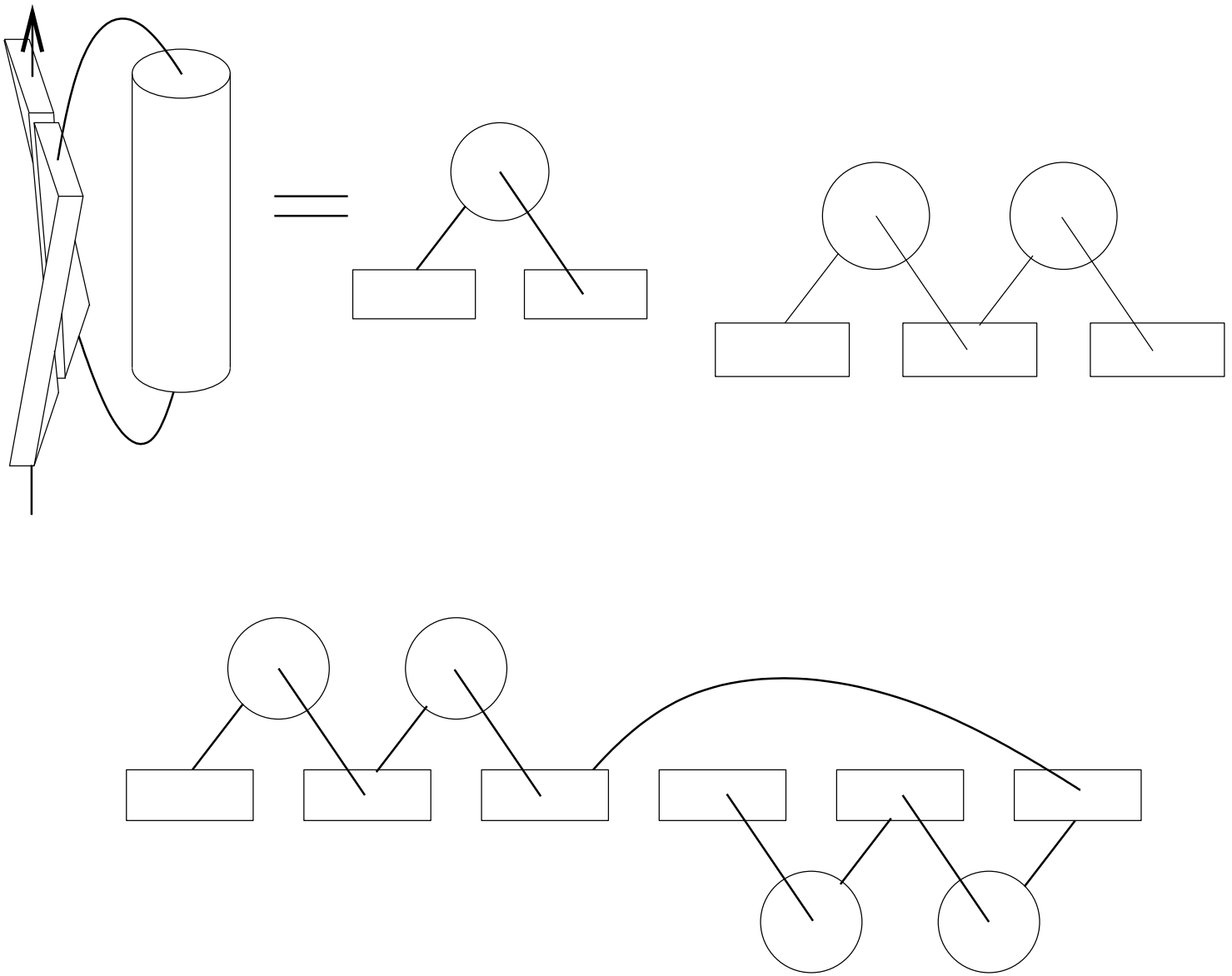




Packing of β-barrels

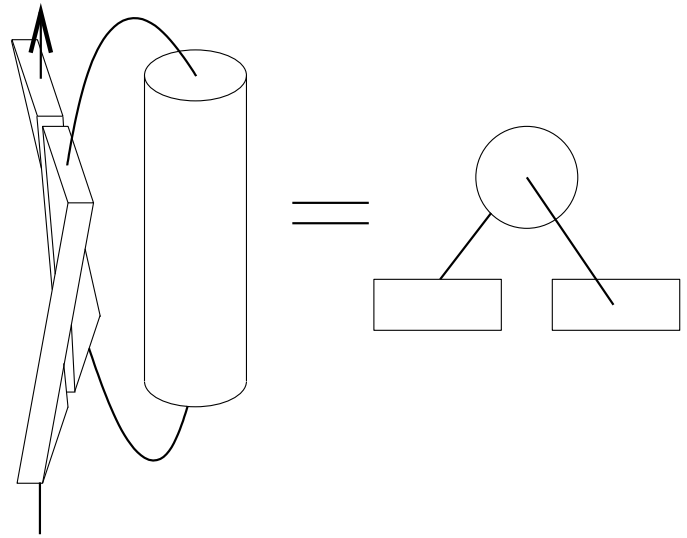


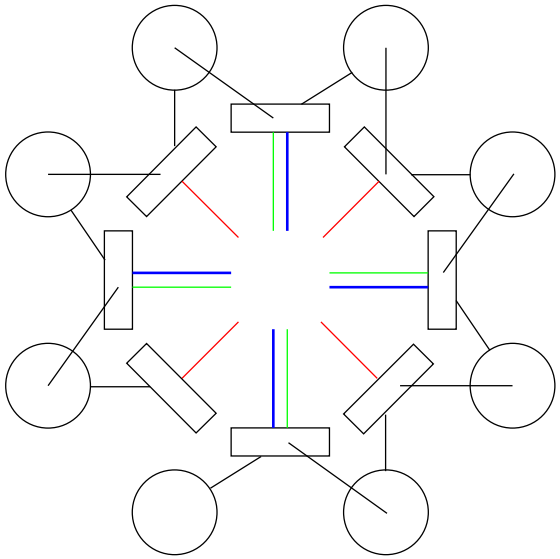
α/β-proteins

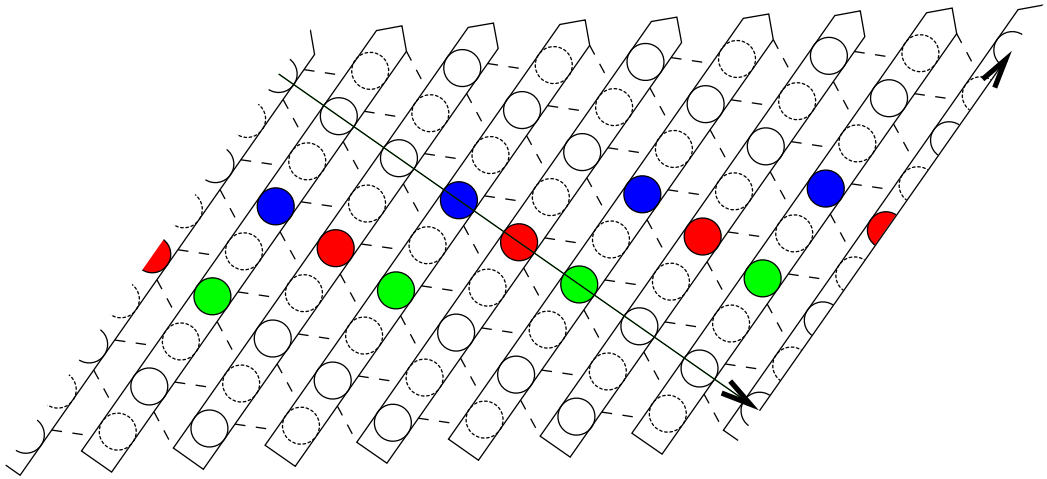


Packing in α/β-proteins (TIM-barrel)

Hydrophobic residues inside







Packing of TIM-barrel