Advanced biochemistry and its methods
Lectures 4 and 5

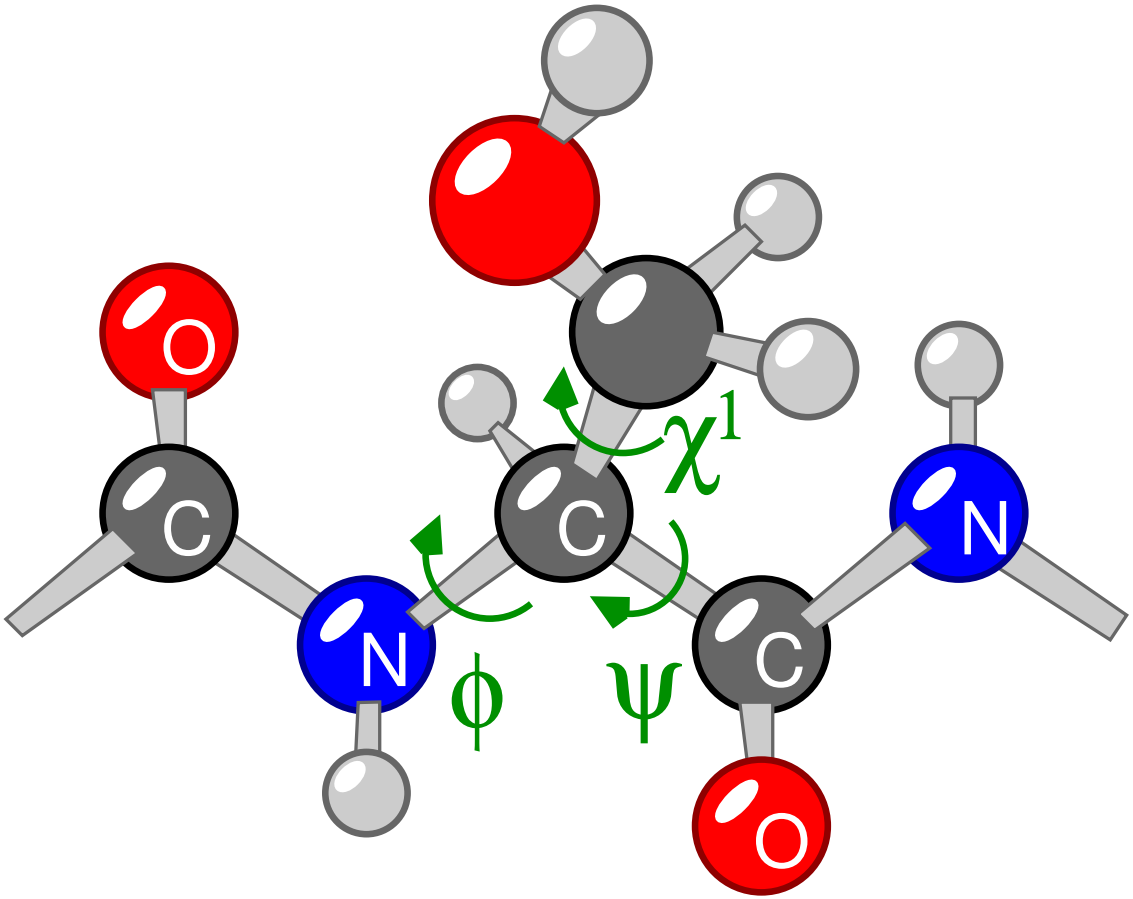
Lukáš Žídek

lzidek@chemi.muni.cz

Finkelstein and Ptitsyn: Protein Physics, Academic Press 2002
Daune: Molecular Biophysics, Molecular Biophysics, Oxford University Press 1999

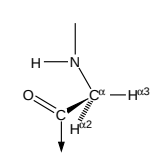
Proteins

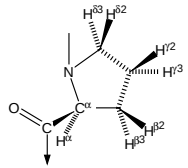
Amino acids connected by peptide bonds

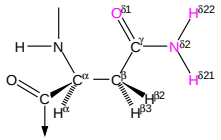


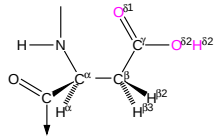
Protein structure = conformation
defined by torsion angles $(\phi, \psi, \chi^1, \ldots)$

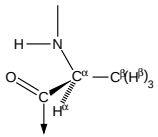
Amino acids

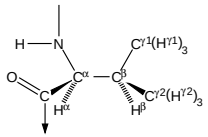

Gly (G)

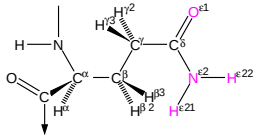

Pro (P)

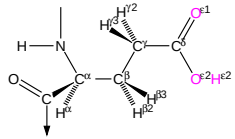

Asn (N)

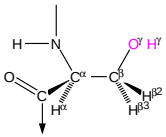

Asp (D)

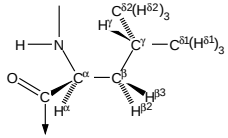

Ala (A)

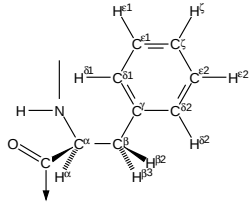

Val (V)

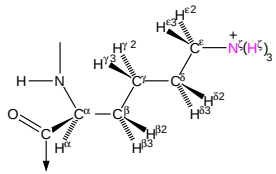

Gln (Q)

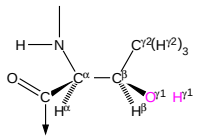

Glu (E)

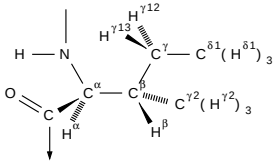

Ser (S)

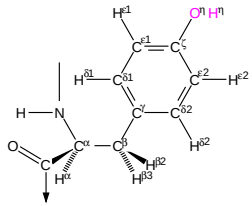

Leu (L)

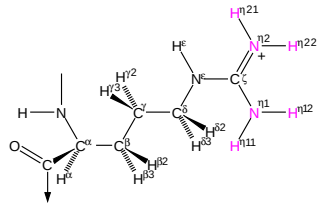

Phe (F)

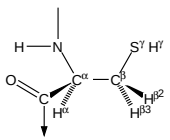

Lys (K)

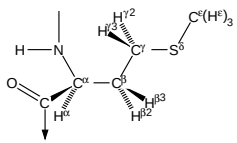

Thr (T)

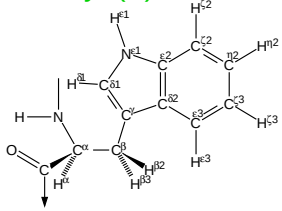

Ile (I)

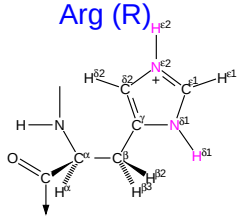

Tyr (Y)


Arg (R)


Cys (C)


Met (M)


Trp (W)


His (H)

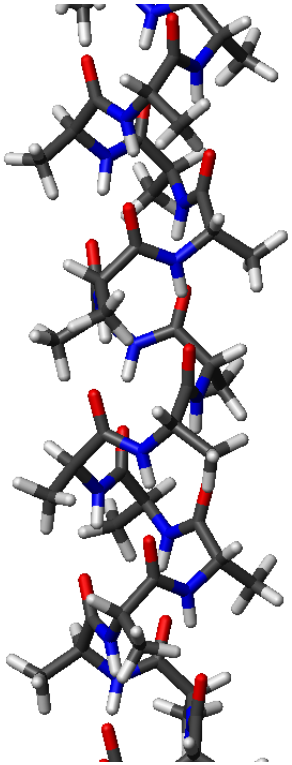
Primary structure

Amino acid sequence

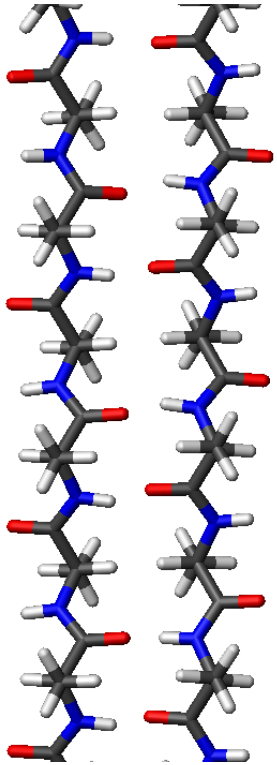
SAKIIHLTDDSFDTDVLKAILVDFW
AEWCGPCKMIAPILDEIADEYQGKL
TAPKYGIRGIPTLLLFKNGEVAATK
VGALSKGQLKEFLDANLA

Secondary structure

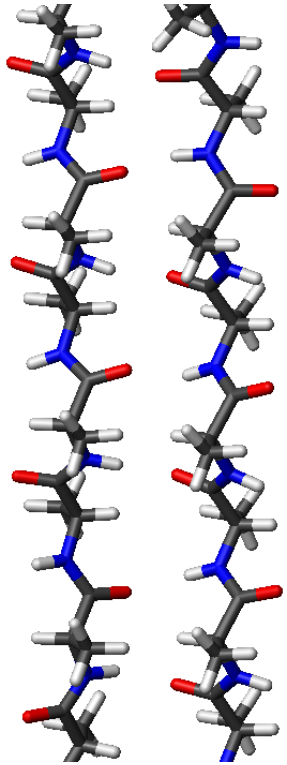
Conformation of protein backbone
regular universal repetitive motifs



α-helix

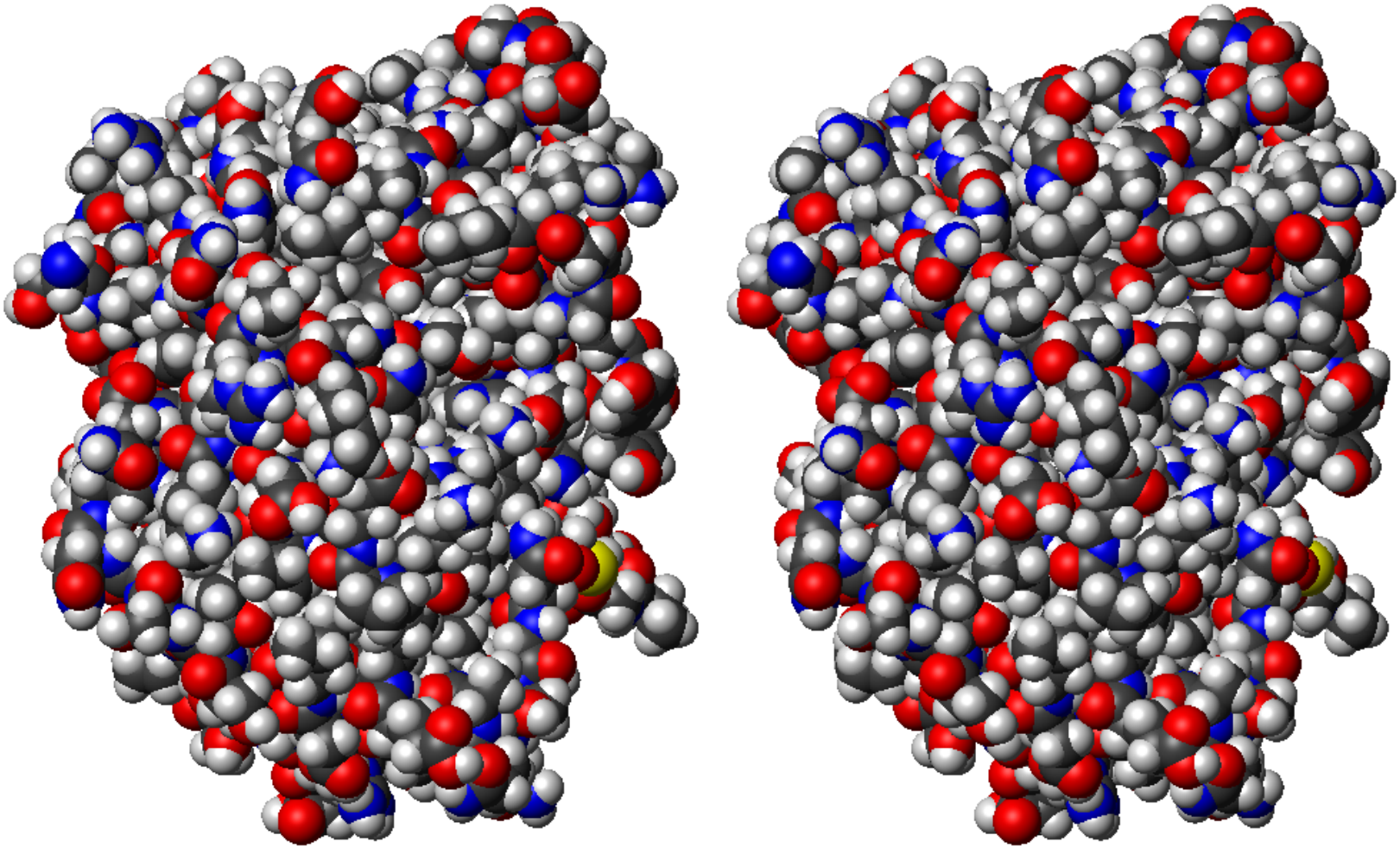


antiparallel β-sheet

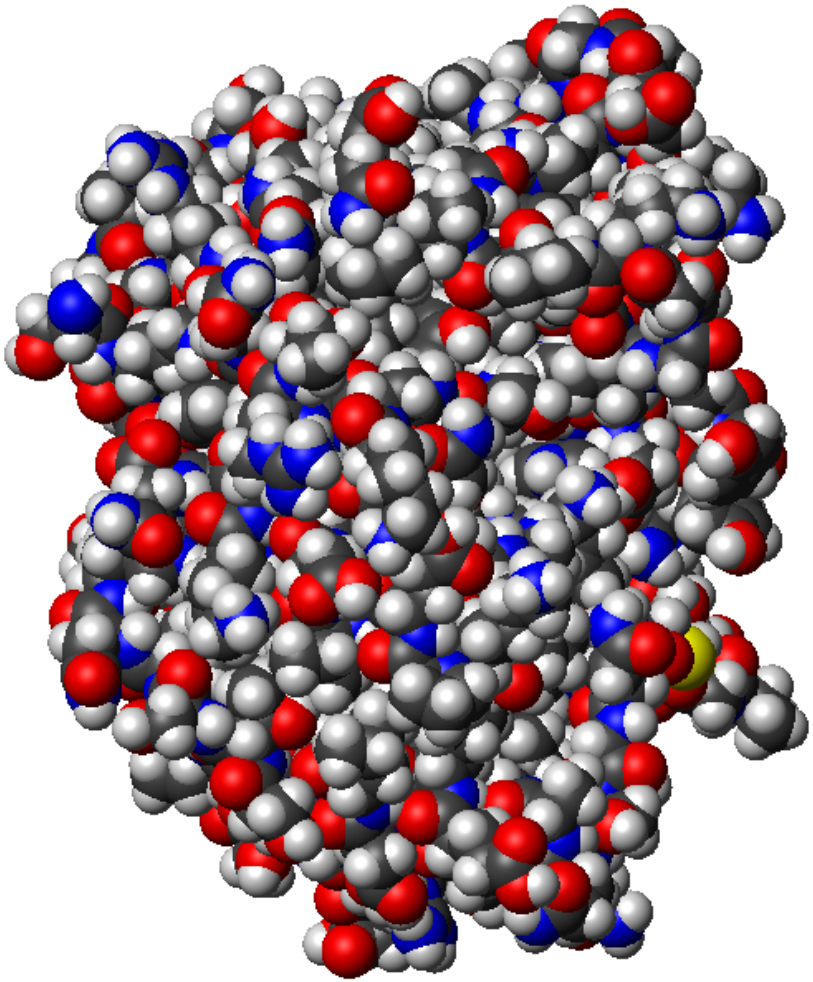


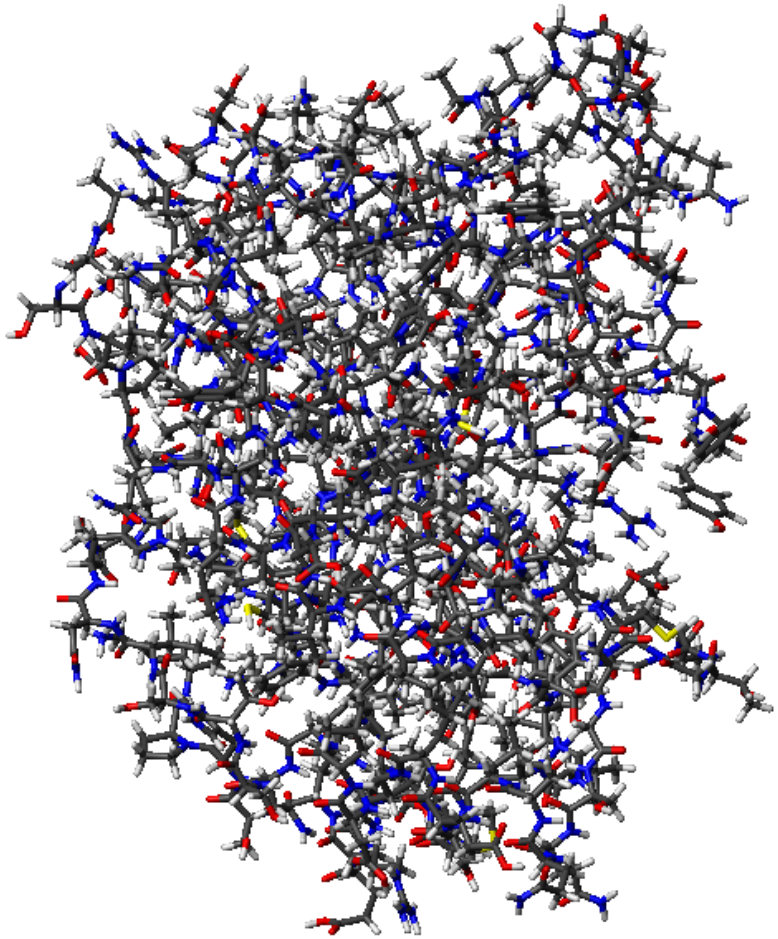
parallel β-sheet

Tertiary structure

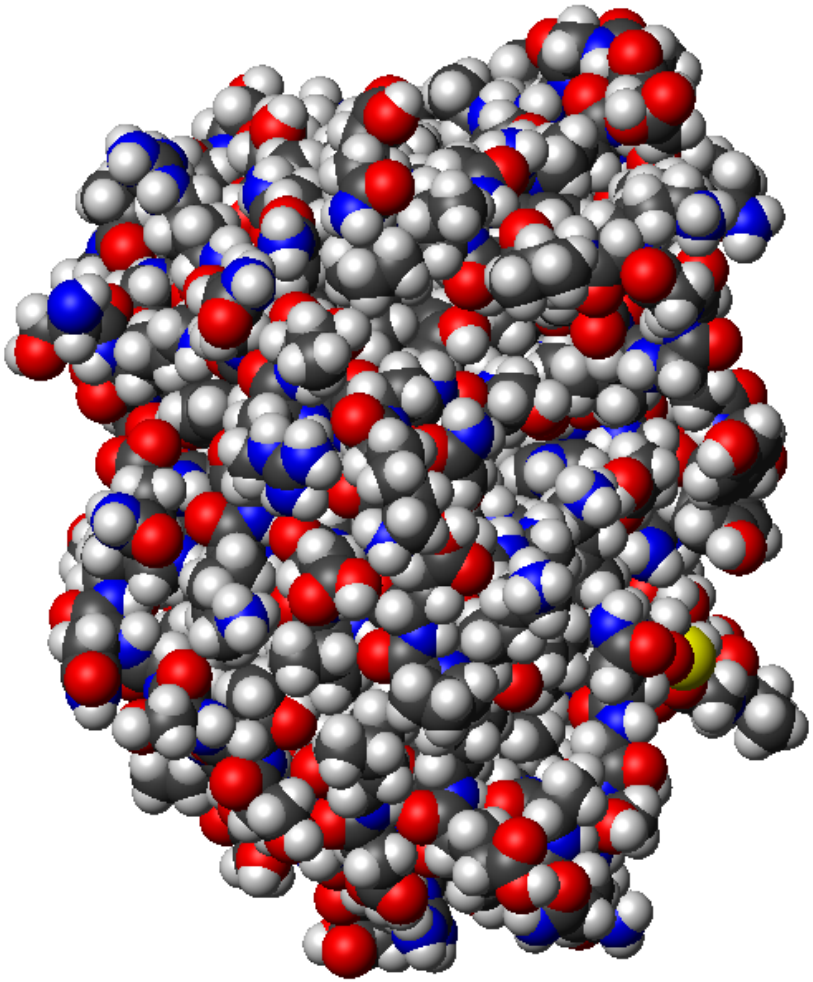


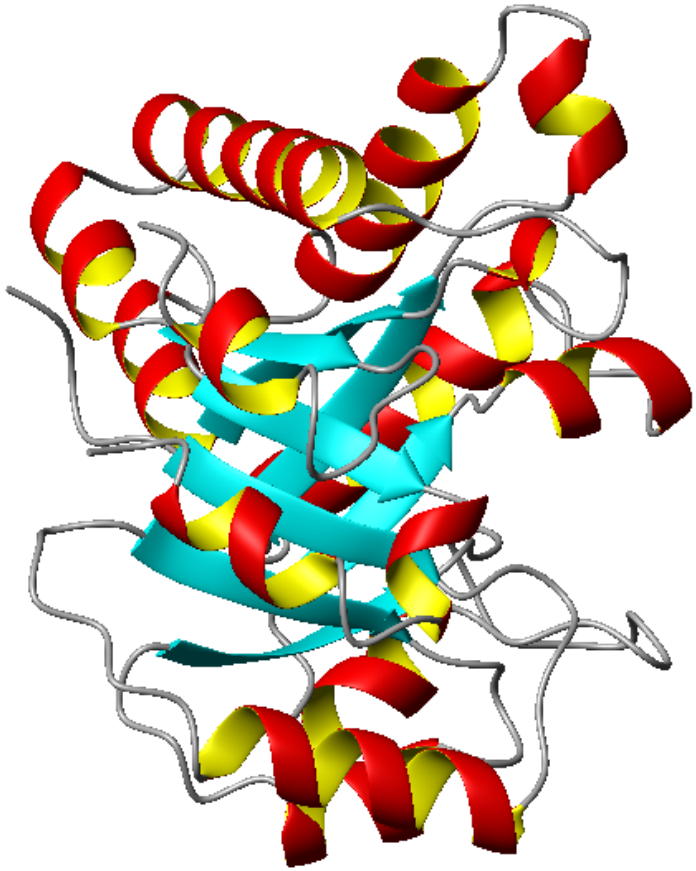
Tertiary structure





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 $S = R \ln \Omega$ (Ω = 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 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

Chemistry: electromagnetic force only

Coulomb's law:

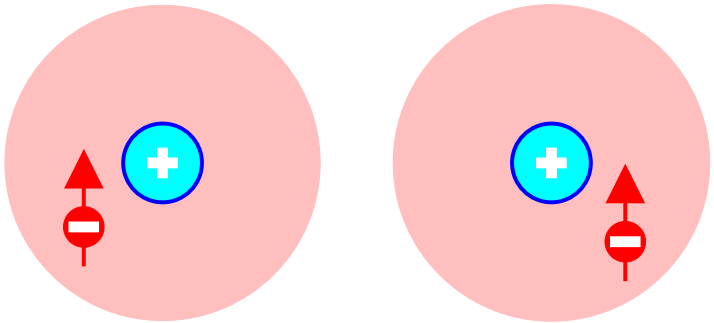
$$F = \frac{1}{4\pi\epsilon_0}\frac{Q_1 Q_2}{r^2}$$

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

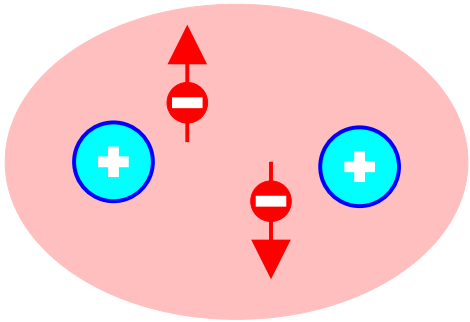
- Force is a vector: $\vec{F} = \frac{1}{4\pi\epsilon_0}\frac{Q_1 Q_2}{r^2}\cdot\underbrace{\frac{\vec{r}}{r}}_{\text{unit vector}}$
- Electric intensity: $\vec{E} = \frac{1}{4\pi\epsilon_0}\frac{Q}{r^2}\cdot\frac{\vec{r}}{r}$
- $U = \frac{N_A}{4\pi\epsilon_0}\frac{Q_1 Q_2}{r}$ if expressed in kJ/mol

Covalent bonds

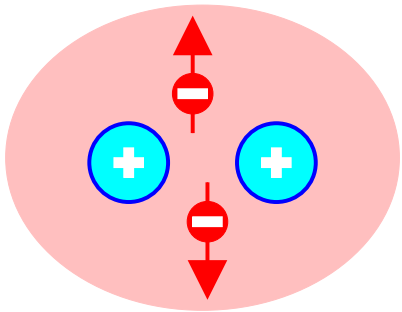
Quantum mechanics



reference energy

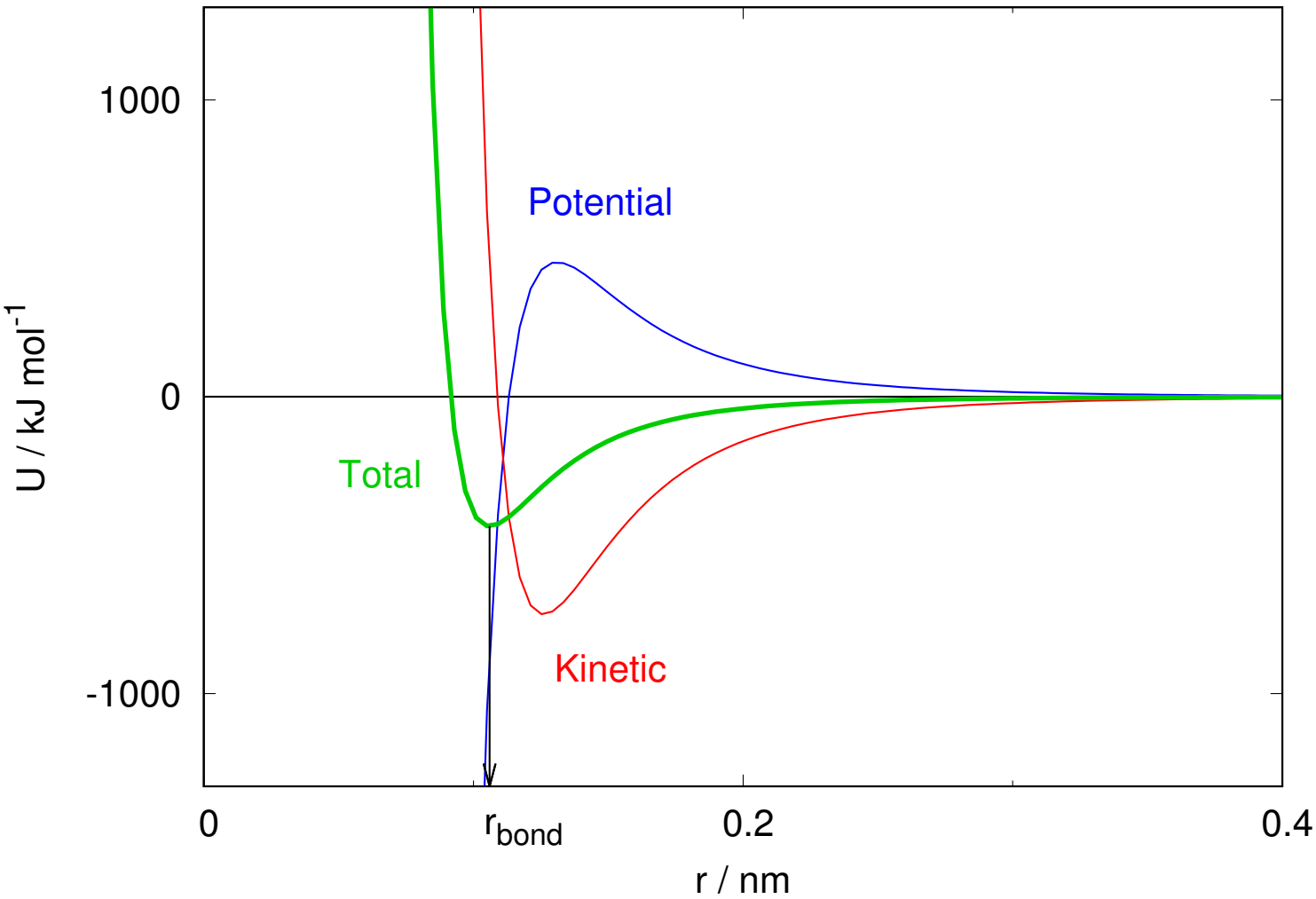


lower energy



higher energy

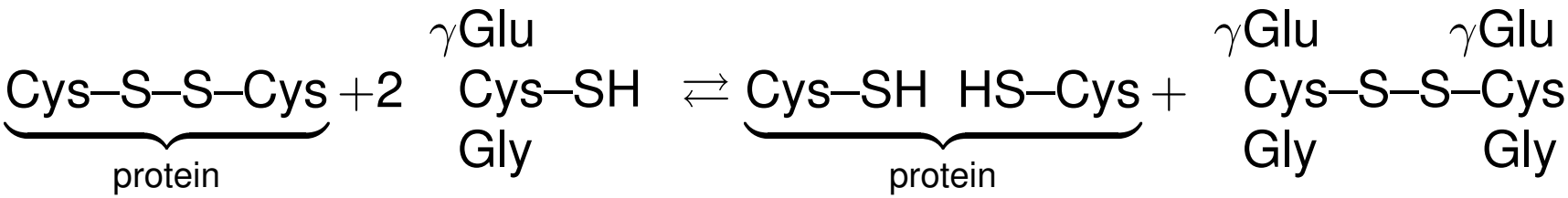
Covalent bonds



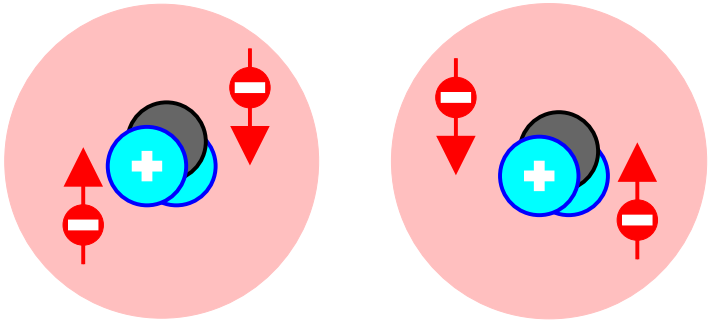
Covalent bonds in proteins

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

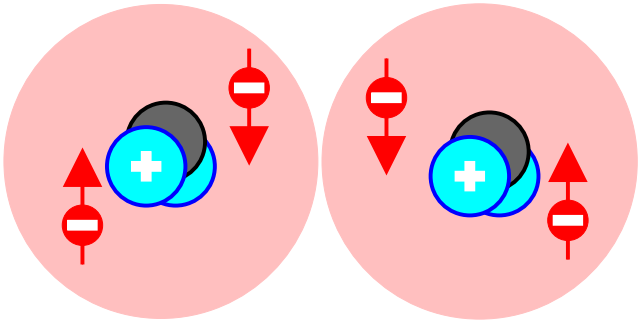
S–S bridges important (and frequent) in extracellular proteins but play marginal structural role in intracellular proteins: Exchange with glutathione ($\Delta G \approx 0$)



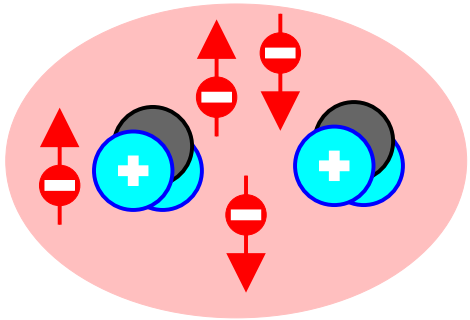
Interactions of nonpolar molecules



reference energy

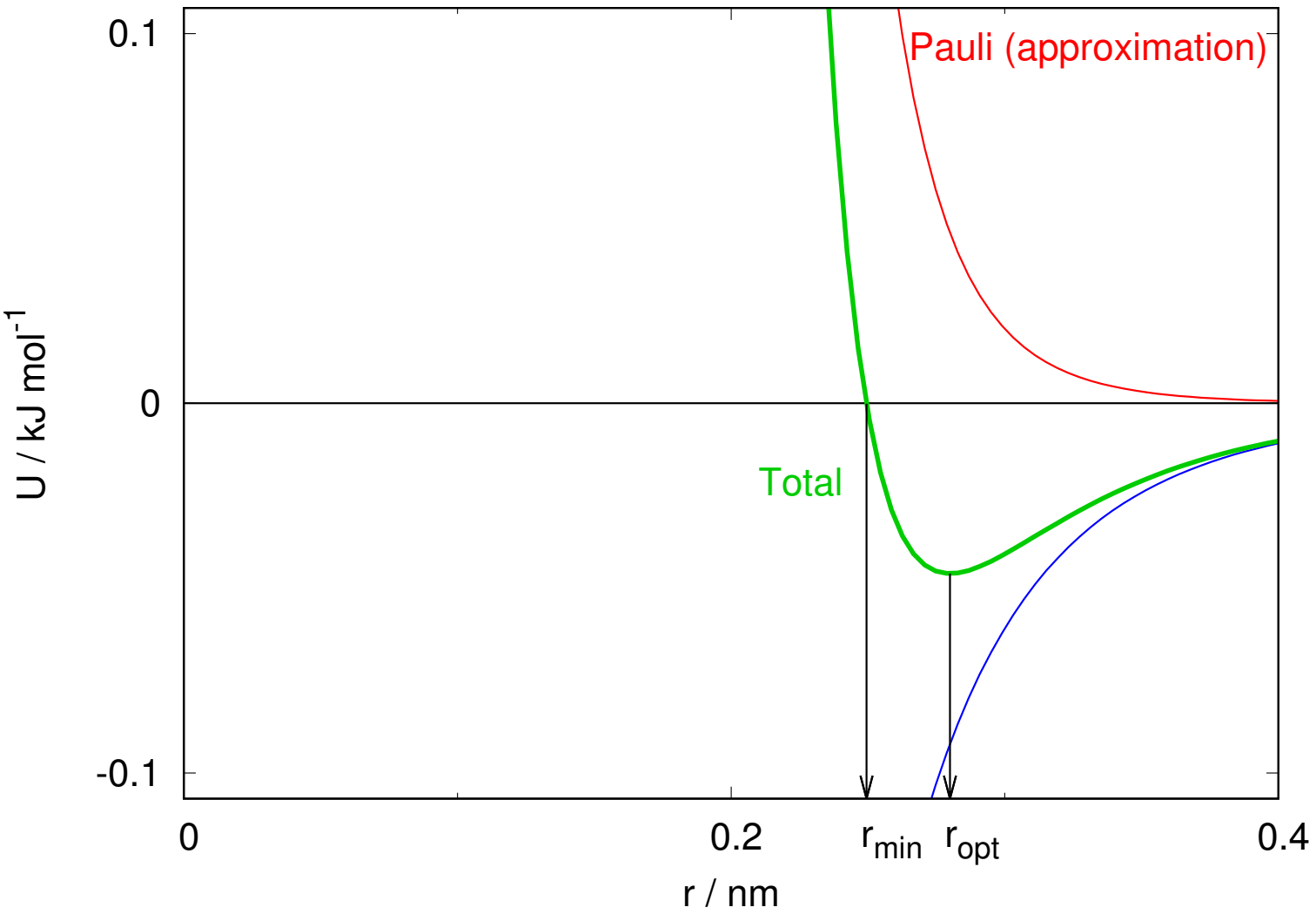


lower energy

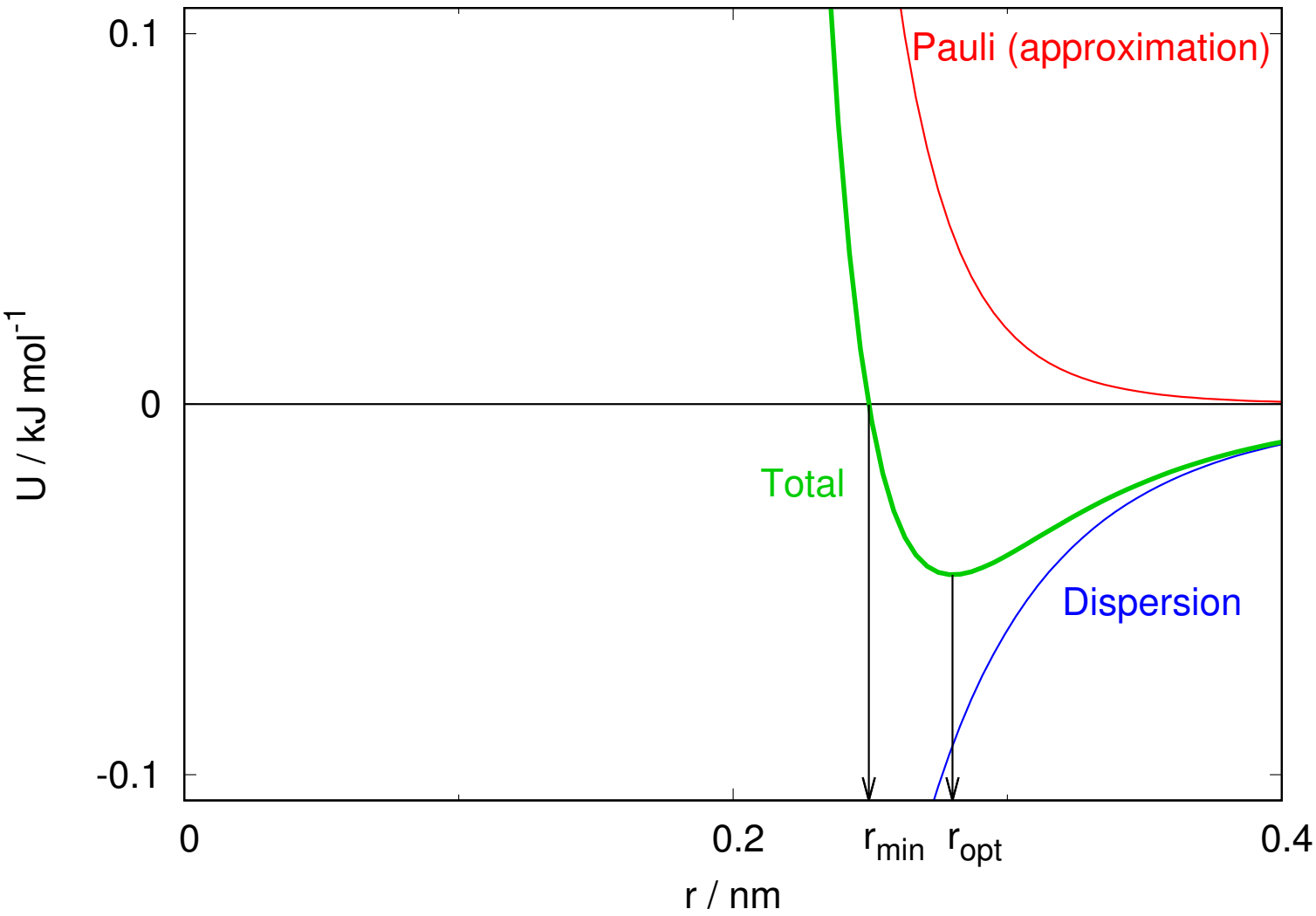


IMPOSSIBLE !

Interactions of nonpolar molecules

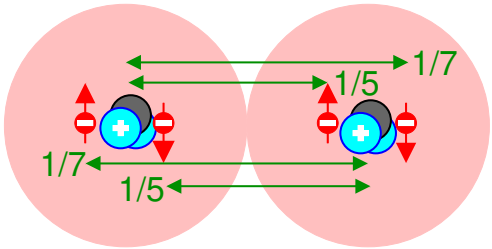


Dispersion forces



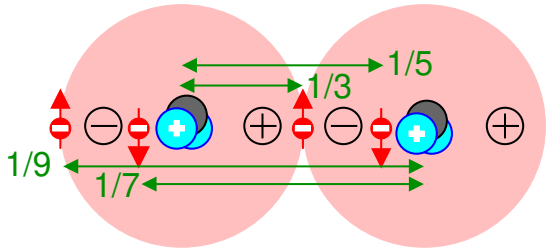
Dispersion forces

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



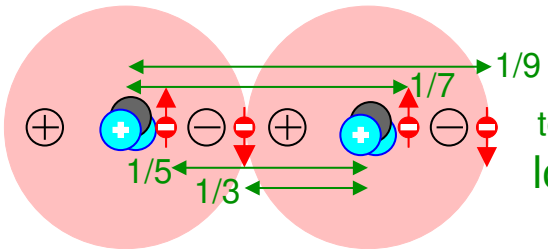
total relative intermolecular attractive energy $=-(1/7+1/5+1/7+1/5)=-4.114$

reference energy



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

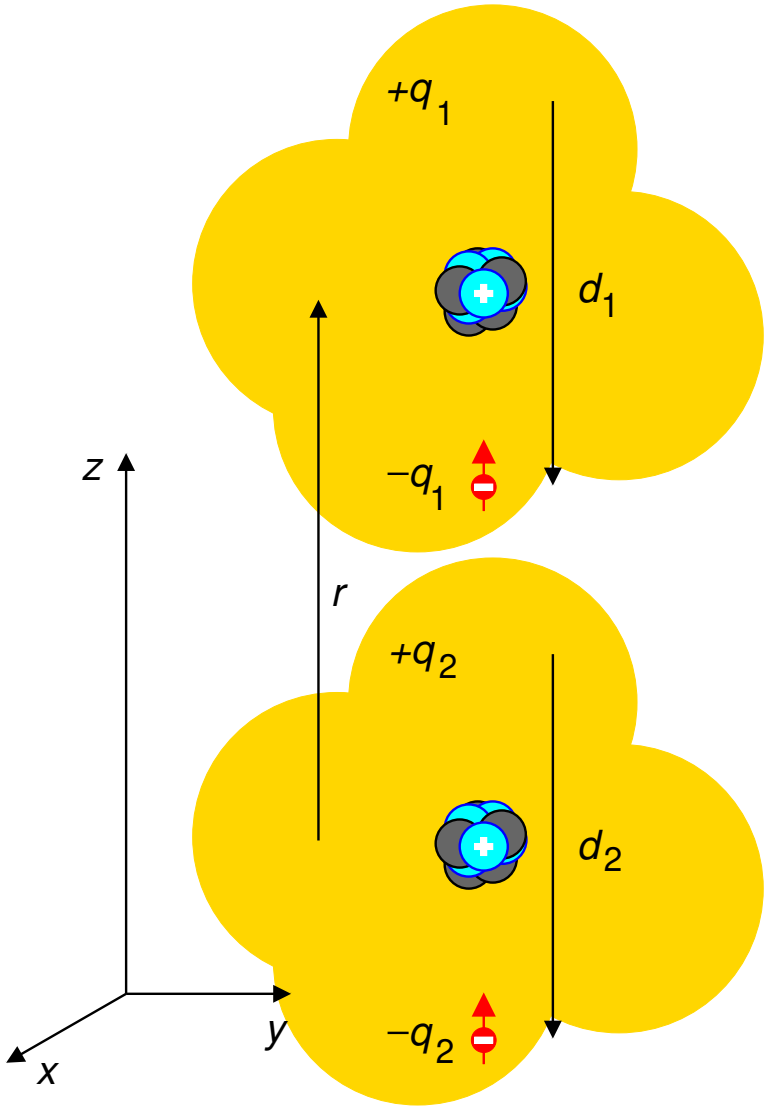
lower energy

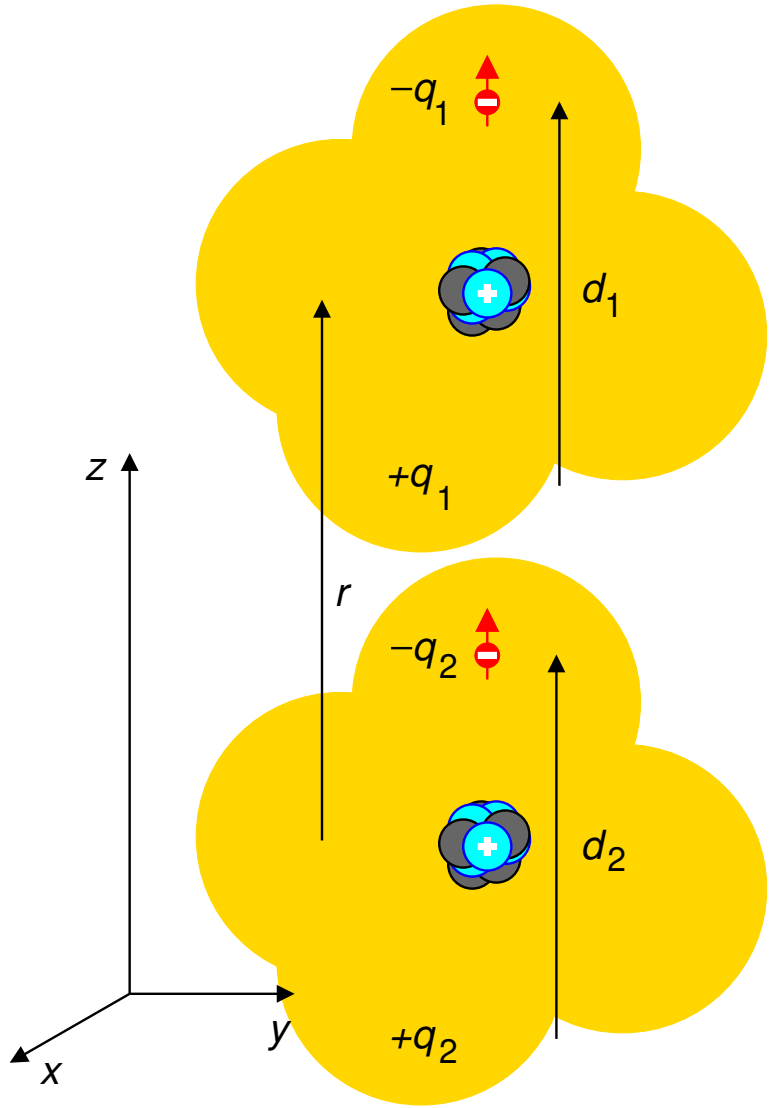


total relative intermolecular attractive energy $=-(1/5+1/3+1/9+1/7)=-4.724$

lower energy

Dispersion forces





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:

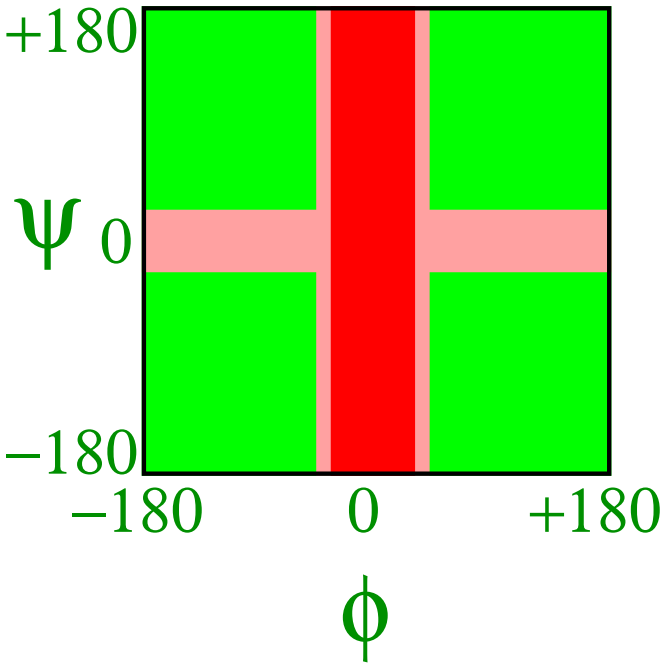
$$U = U_{\text{opt}} \left(\left(\frac{r_{\text{opt}}}{r} \right)^{12} - 2 \left(\frac{r_{\text{opt}}}{r} \right)^{6} \right)$$

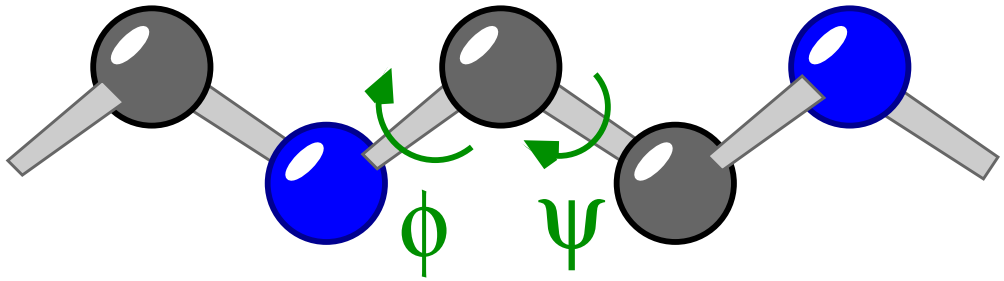
Van der Waals parameters

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

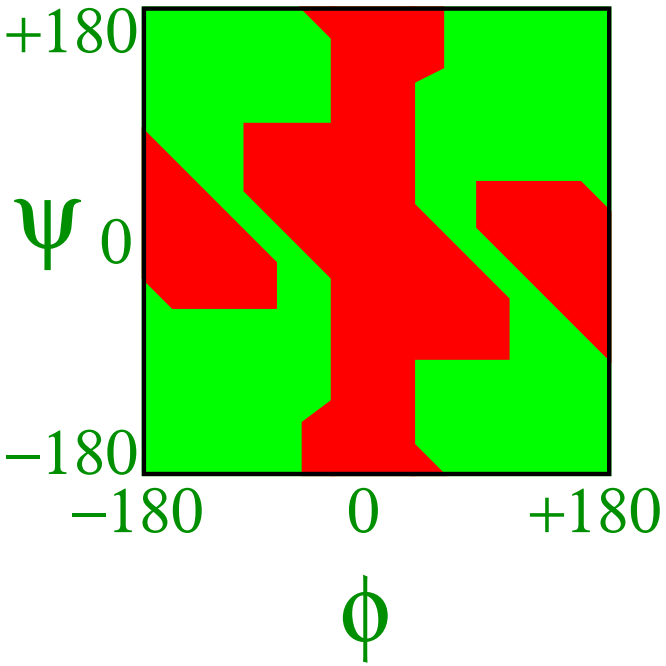
Repulsion of backbone C and N only

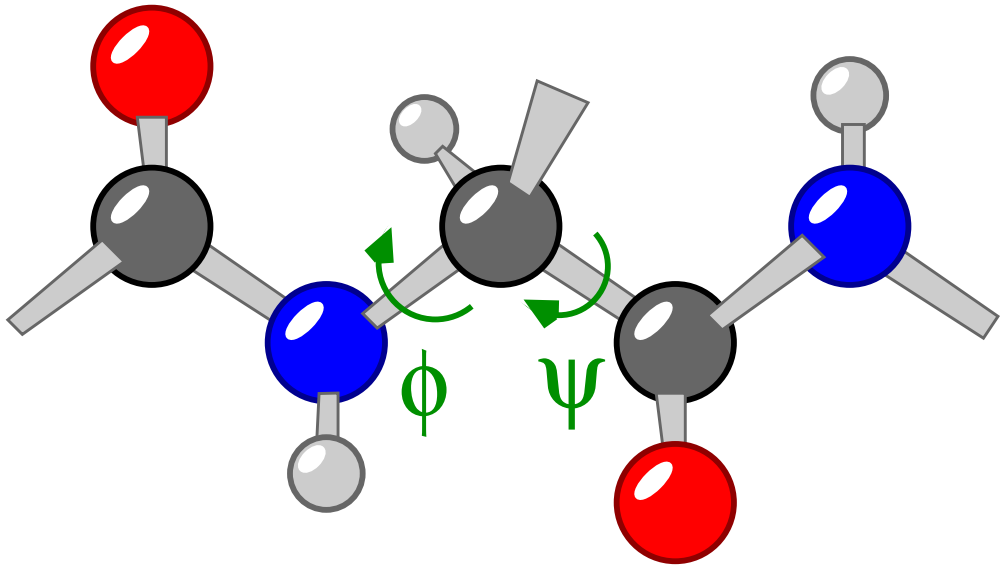




Ramachandran diagram

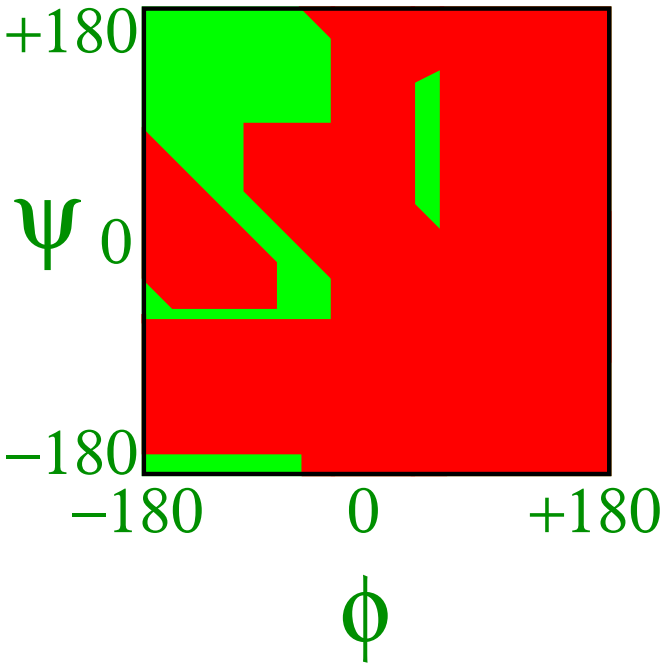
Repulsion including backbone amide H and O

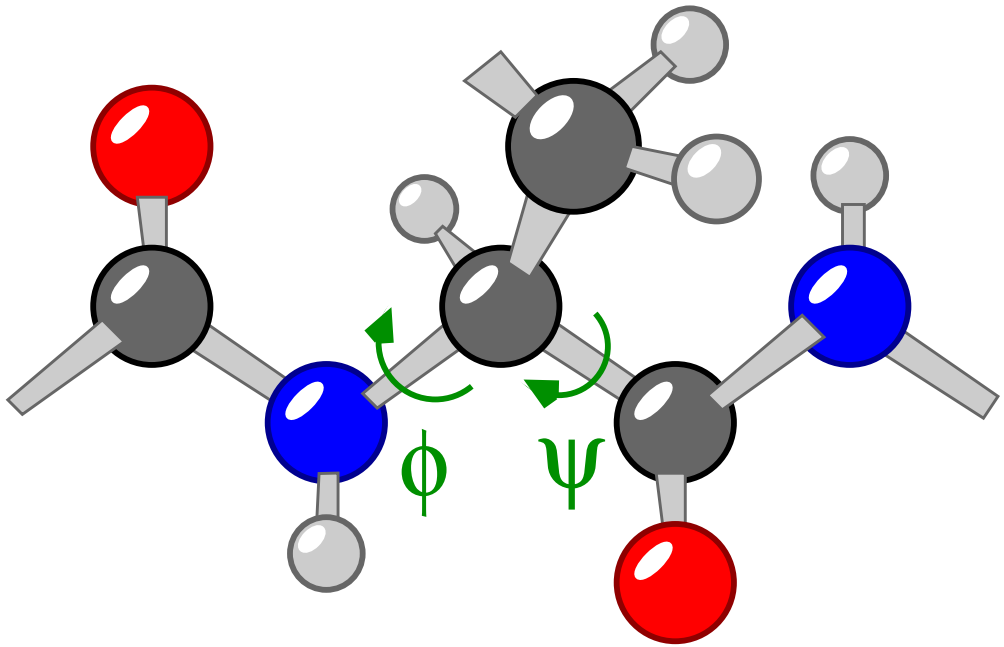




Ramachandran diagram

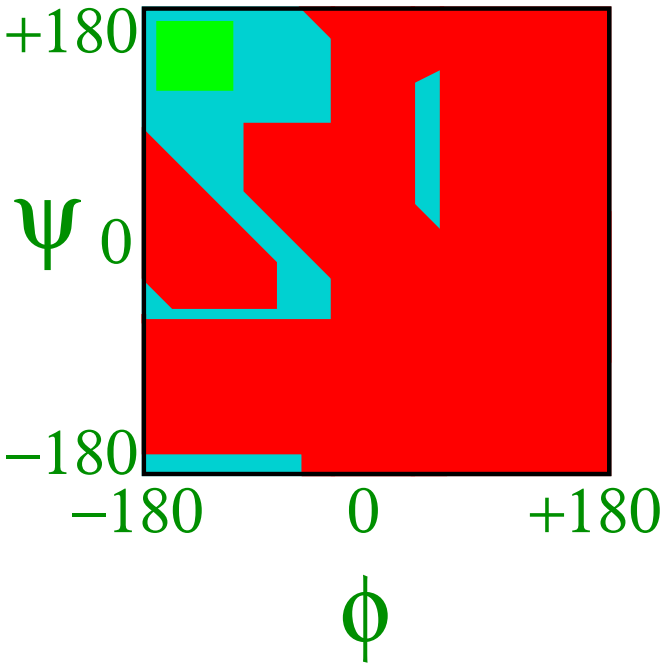
Repulsion including C^{β}

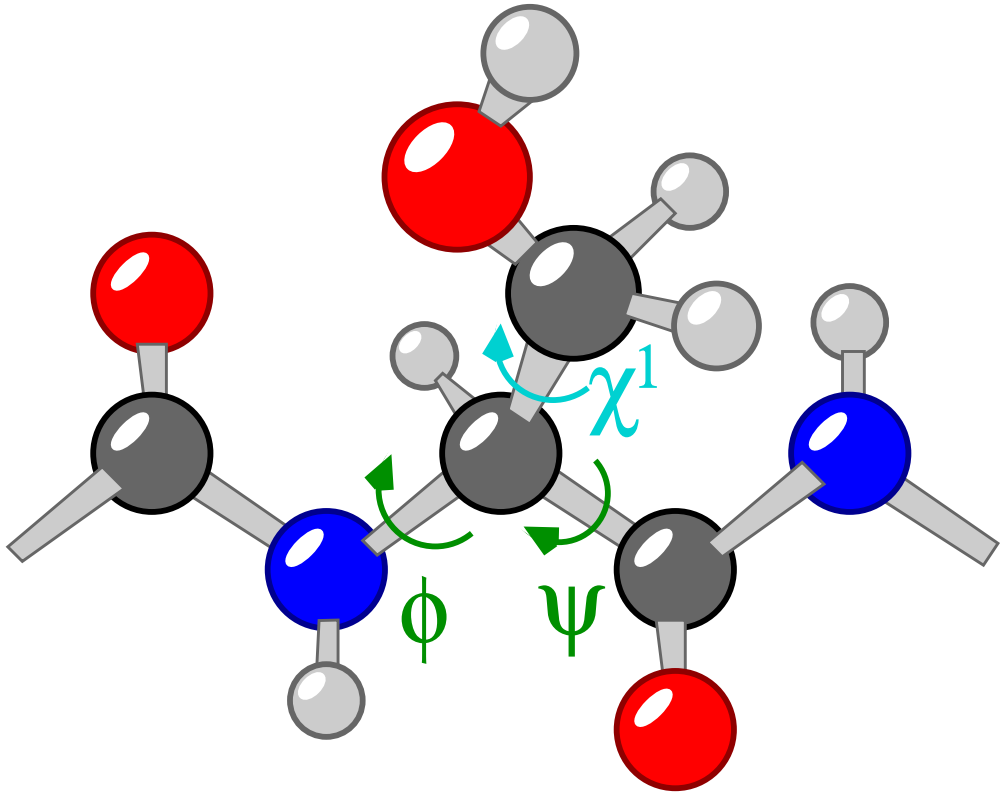




Ramachandran diagram

Repulsion including side chains (all, side chain dependent)





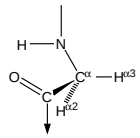
Electrostatic interactions

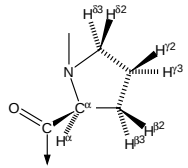
Charged groups (ions):

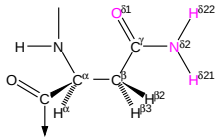
$$F = \frac{1}{4\pi\epsilon_0}\frac{Q_1 Q_2}{r^2} \qquad U = \frac{1}{4\pi\epsilon_0}\frac{Q_1 Q_2}{r}$$

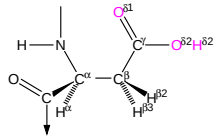
$\Delta G = 460\,\text{kJ/mol}$ for charges $0.3\,\text{nm}$ appart

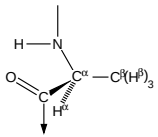
Charged amino acids

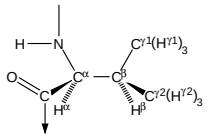

Gly (G)

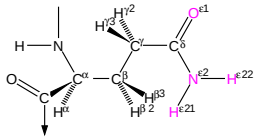

Pro (P)

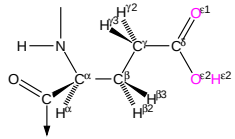

Asn (N)

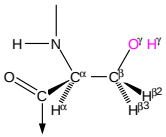

Asp (D)

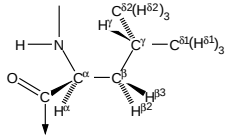

Ala (A)

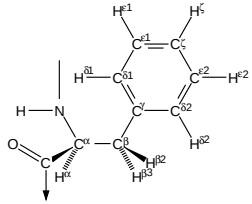

Val (V)

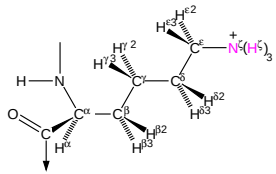

Gln (Q)

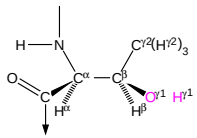

Glu (E)

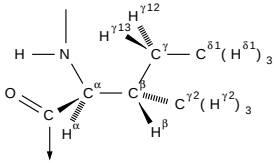

Ser (S)

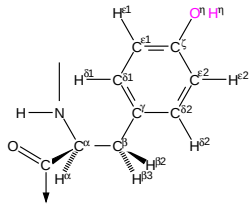

Leu (L)

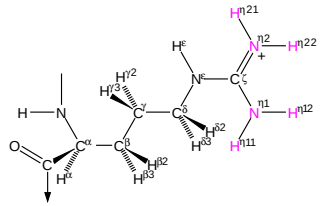

Phe (F)

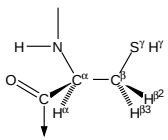

Lys (K)

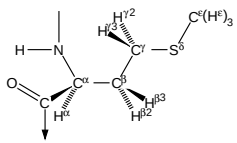

Thr (T)

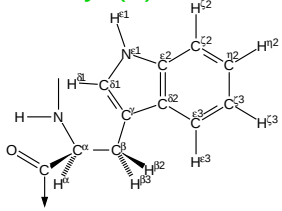

Ile (I)

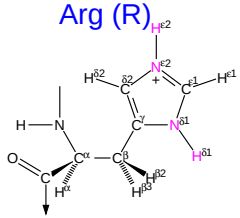

Tyr (Y)


Arg (R)

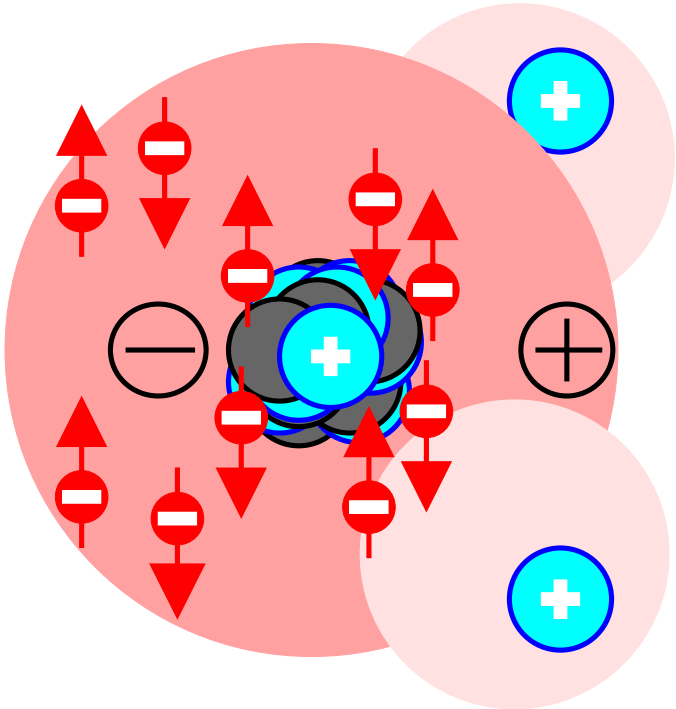

Cys (C)

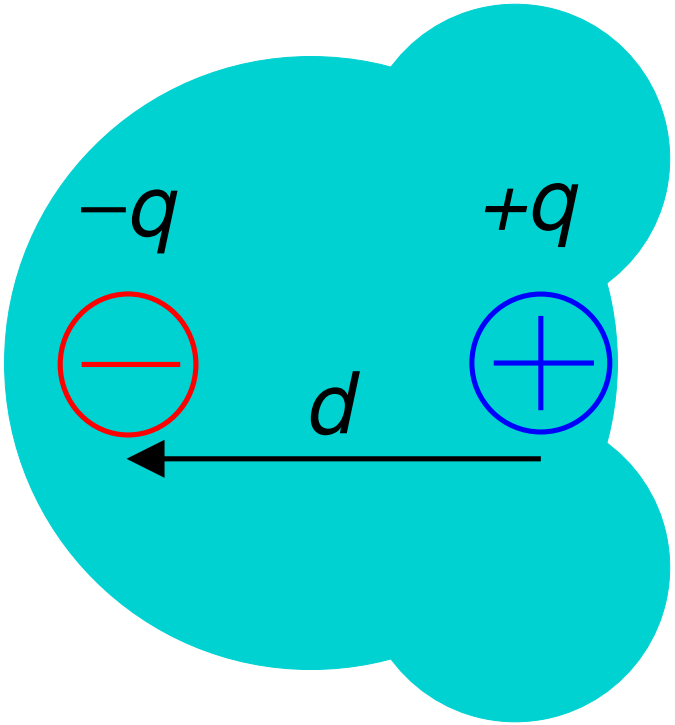

Met (M)


Trp (W)


His (H)

Polar molecules



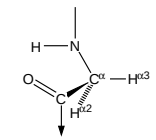


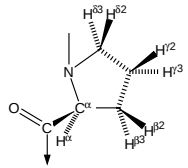
Permanent electric dipoles:

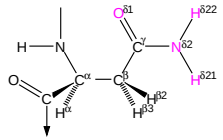
zero net charge but partial charges $\pm q$ separated by distance d

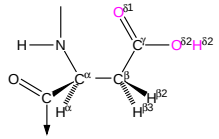
polar groups in molecules

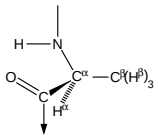
Polar amino acids

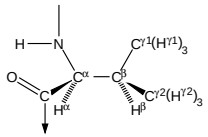

Gly (G)

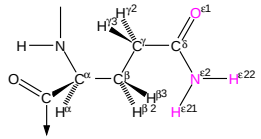

Pro (P)

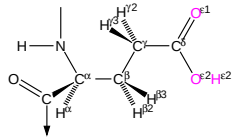

Asn (N)

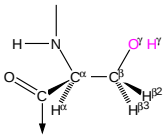

Asp (D)

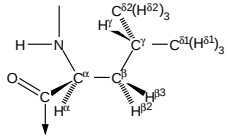

Ala (A)

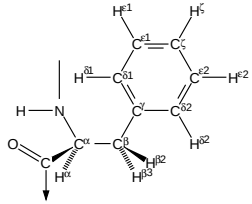

Val (V)

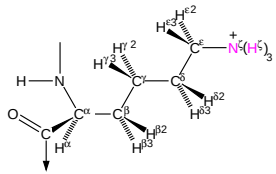

Gln (Q)

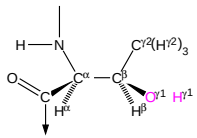

Glu (E)

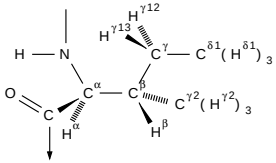

Ser (S)

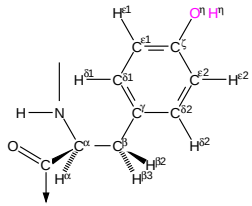

Leu (L)

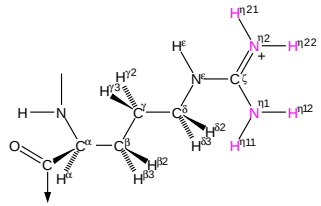

Phe (F)

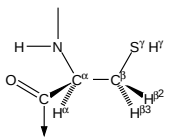

Lys (K)

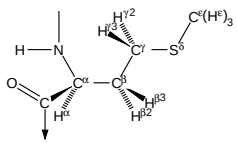

Thr (T)

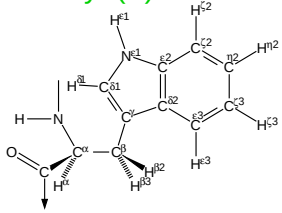

Ile (I)

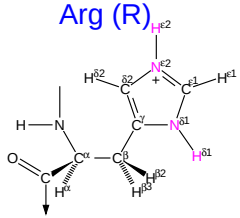

Tyr (Y)


Arg (R)


Cys (C)

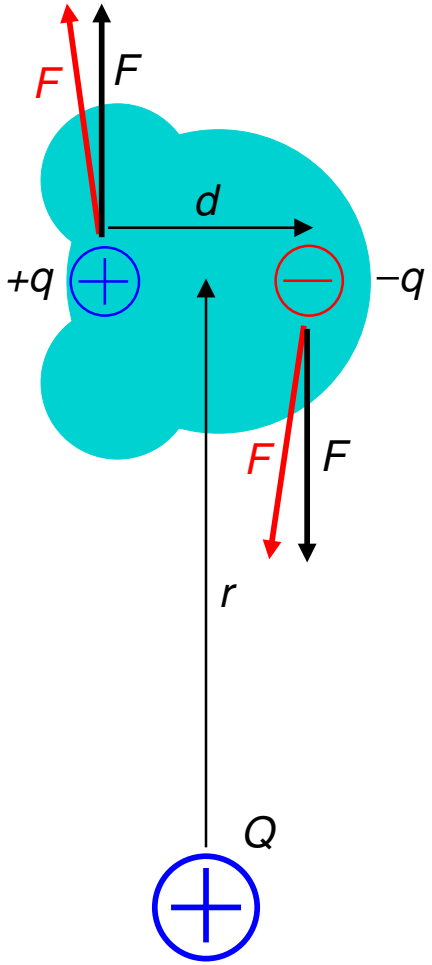

Met (M)

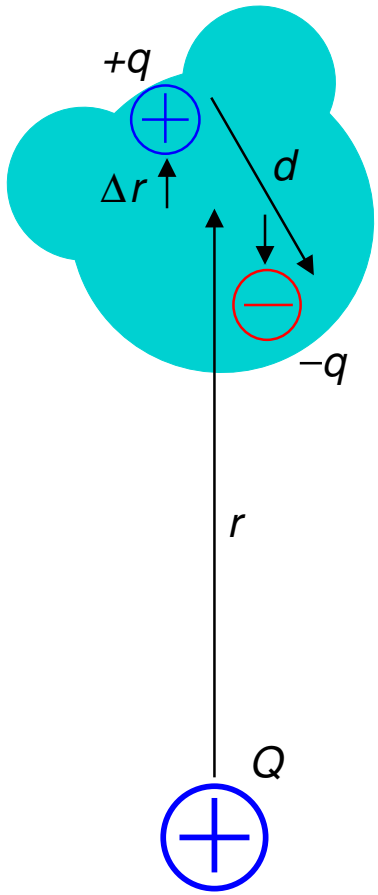

Trp (W)


His (H)

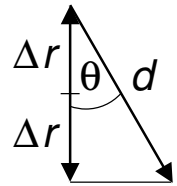
Permanent electric dipoles

Moment of forces (torque) rotates the dipole





if $d \ll r$



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

$F = F$

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

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

Why d/r ?

Potential energy of rotating dipole from perpendicular orientation to orientation tilted by θ from electric force lines:

$$U = \int_r^{r+\Delta r} F_1 \mathrm{d}r' + \int_r^{r-\Delta r} F_2 \mathrm{d}r' = \frac{-qQ}{4\pi\epsilon_0} \int_r^{r+\Delta r} \frac{1}{r^2}\mathrm{d}r' + \frac{qQ}{4\pi\epsilon_0} \int_r^{r-\Delta r} \frac{1}{r^2}\mathrm{d}r' =$$

$$= -\frac{qQ}{4\pi\epsilon_0}\frac{2\Delta r}{r^2} = -\frac{1}{4\pi\epsilon_0}\frac{dqQ}{r^2}\cos\theta$$

Permanent electric dipoles

Why $(\cdots)^2/3RT$?

orientation (θ) is averaged in solution

Averaging in general: $\langle X \rangle = \sum_i P_i X_i \longrightarrow \int P(u)X(u)\mathrm{d}u$

P_i probability that $X = X_i$; $P(u)$ probability that $X = X(u)$

Averaging of electric dipoles:

$u = \cos\theta$ orientation; $\frac{\mathrm{d}u}{\mathrm{d}\theta} = \frac{\mathrm{d}\cos\theta}{\mathrm{d}\theta} = -\sin\theta \Rightarrow -\sin\theta\mathrm{d}\theta = \mathrm{d}u$

$U = -\frac{1}{4\pi\epsilon_0}\frac{dqQ}{r^2}\cos\theta = -\frac{1}{4\pi\epsilon_0}\frac{dqQ}{r^2}u$ energy of a dipole tilted by θ

Boltzmann: $P(u) = Z^{-1}\mathrm{e}^{-U(u)/RT} = Z^{-1}\mathrm{e}^{uw}$; $w = \frac{1}{4\pi\epsilon_0}\frac{dqQ}{r^2RT}$

Z = sum of all possible e^{uw} ("partition function"):

$$Z = \int_0^\pi P(\cos\theta)\sin\theta\mathrm{d}\theta = \int_1^{-1} -\mathrm{e}^{uw}\mathrm{d}u = \frac{\mathrm{e}^w - \mathrm{e}^{-w}}{w}$$

Permanent electric dipoles

$$\langle U \rangle = \int_1^{-1} P(u)U(u)\mathrm{d}u = \frac{RTw}{\mathrm{e}^w - \mathrm{e}^{-w}} \int_1^{-1} \mathrm{e}^{uw} u \mathrm{d}u = RT\left(w\frac{\mathrm{e}^w + \mathrm{e}^{-w}}{\mathrm{e}^w - \mathrm{e}^{-w}} - 1\right)$$

where $\int_1^{-1} u\mathrm{e}^{-uw}\mathrm{d}u$ was solved using the chain rule.

If w is small (relatively small dipole relatively far from Q),

$e^{\pm w} \approx 1 \pm w + \frac{1}{2}w^2 \pm \frac{1}{6}w^3 + \cdots$

$RT\left(w\frac{\mathrm{e}^w+\mathrm{e}^{-w}}{\mathrm{e}^w-\mathrm{e}^{-w}} - 1\right) = RT\frac{(w-1)\mathrm{e}^w-(w+1)\mathrm{e}^{-w}}{\mathrm{e}^w-\mathrm{e}^{-w}} = RT\frac{w\left(\mathrm{e}^w+\mathrm{e}^{-w}\right)-\left(\mathrm{e}^w-\mathrm{e}^{-w}\right)}{\mathrm{e}^w-\mathrm{e}^{-w}}$

$\mathrm{e}^w - \mathrm{e}^{-w} \approx 1 + w + \cdots - 1 + w - \cdots \approx 2w$

but $w\left(\mathrm{e}^w + \mathrm{e}^{-w}\right) - \left(\mathrm{e}^w - \mathrm{e}^{-w}\right)$ more tricky:

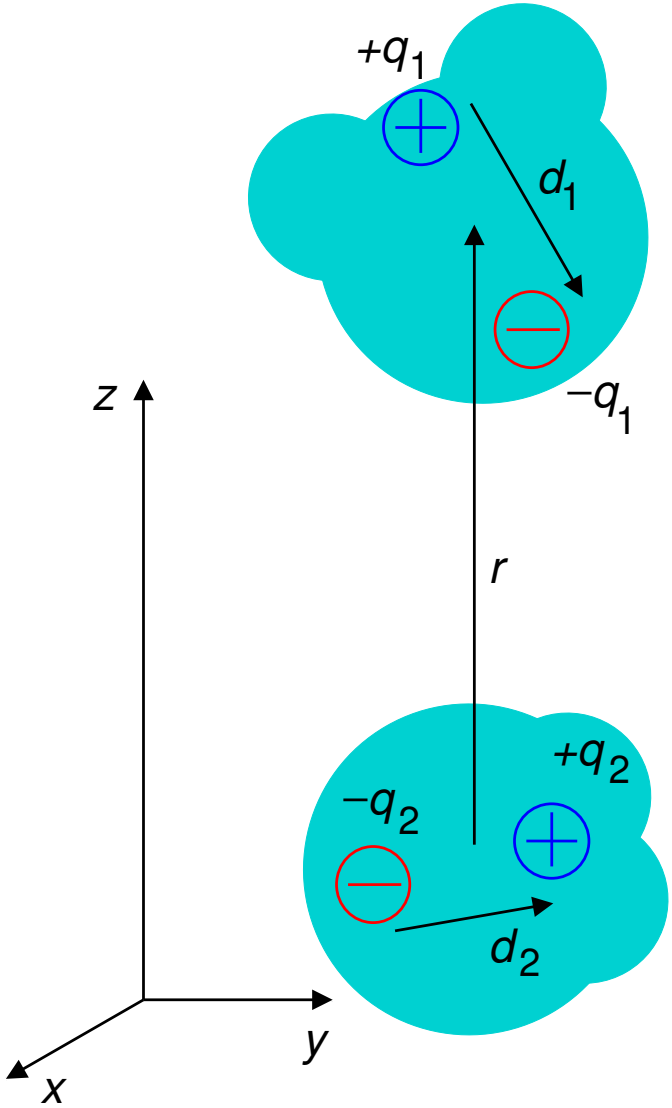
$w\left(\mathrm{e}^w + \mathrm{e}^{-w}\right)\left(\mathrm{e}^w - \mathrm{e}^{-w}\right) \approx w\left(2 + w^2 + \cdots\right) - \left(2w + \frac{1}{3}w^3 + \cdots\right)$

$\left(2w + w^3 + \cdots\right) - \left(2w + \frac{1}{3}w^3 + \cdots\right) = \frac{2}{3}w^3$

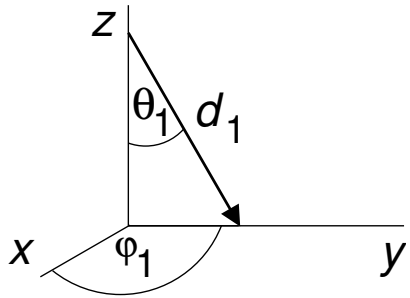
$$RT\frac{(w-1)\mathrm{e}^w - (w+1)\mathrm{e}^{-w}}{\mathrm{e}^w - \mathrm{e}^{-w}} \approx RT\frac{\frac{2}{3}w^3}{2w} = \frac{RTw^2}{3} = -\frac{1}{3RT}\left(\frac{1}{4\pi\epsilon_0}\frac{qdQ}{r^2}\right)^2$$

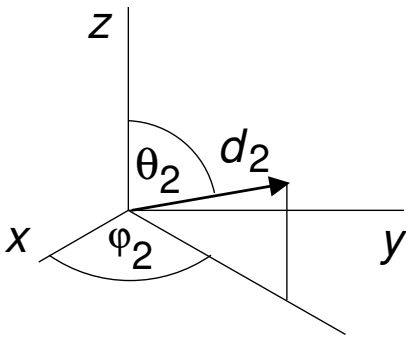
Permanent electric dipoles

4 charges in space $\Rightarrow$ must be analyzed in 3D



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





Permanent electric dipoles

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

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

Permanent dipoles in different molecules

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

Permanent electric dipoles

Calculation is even more tedious, analysis of all forces gives

$$U = \frac{1}{4\pi\epsilon_0}\frac{q_1 q_2}{r}\left(\frac{\vec{d}_1}{r}\cdot\frac{\vec{d}_2}{r} - 3\left(\frac{\vec{d}_1}{r}\cdot\frac{\vec{r}}{r}\right)\left(\frac{\vec{d}_2}{r}\cdot\frac{\vec{r}}{r}\right)\right)$$

which can be expressed in terms of angles $\theta_1, \theta_2, \phi_1, \phi_2$:

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

Probability that the dipoles have a particular orientation is again given by the Bolzmann's law $P(\theta_1, \theta_2, \phi_1, \phi_2) = Z^{-1}\mathrm{e}^{-U(\theta_1,\theta_2,\phi_1,\phi_2)/RT}$, but U depends on 4 angles $\theta_1, \theta_2, \phi_1, \phi_2$. Therefore, averaging (integration) must be performed over all 4 angles.

Induced dipoles

Induced electric dipoles:
polar and nonpolar groups in molecules
the induced dipole is proportional to the inducing force:
$q\vec{d} = \alpha\epsilon_0\vec{E}$
$\vec{E}$ is electric intensity ($\vec{F}/q$, force per unit charge)
α is *polarizability* (how easy is to move electrons)

Charge Q – induced dipole $q \cdot d$

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

Induced dipoles

We already derived that

$$U = -\frac{1}{4\pi\epsilon_0}\frac{qdQ}{r^2}\cos\theta = -qdE\cos\theta$$

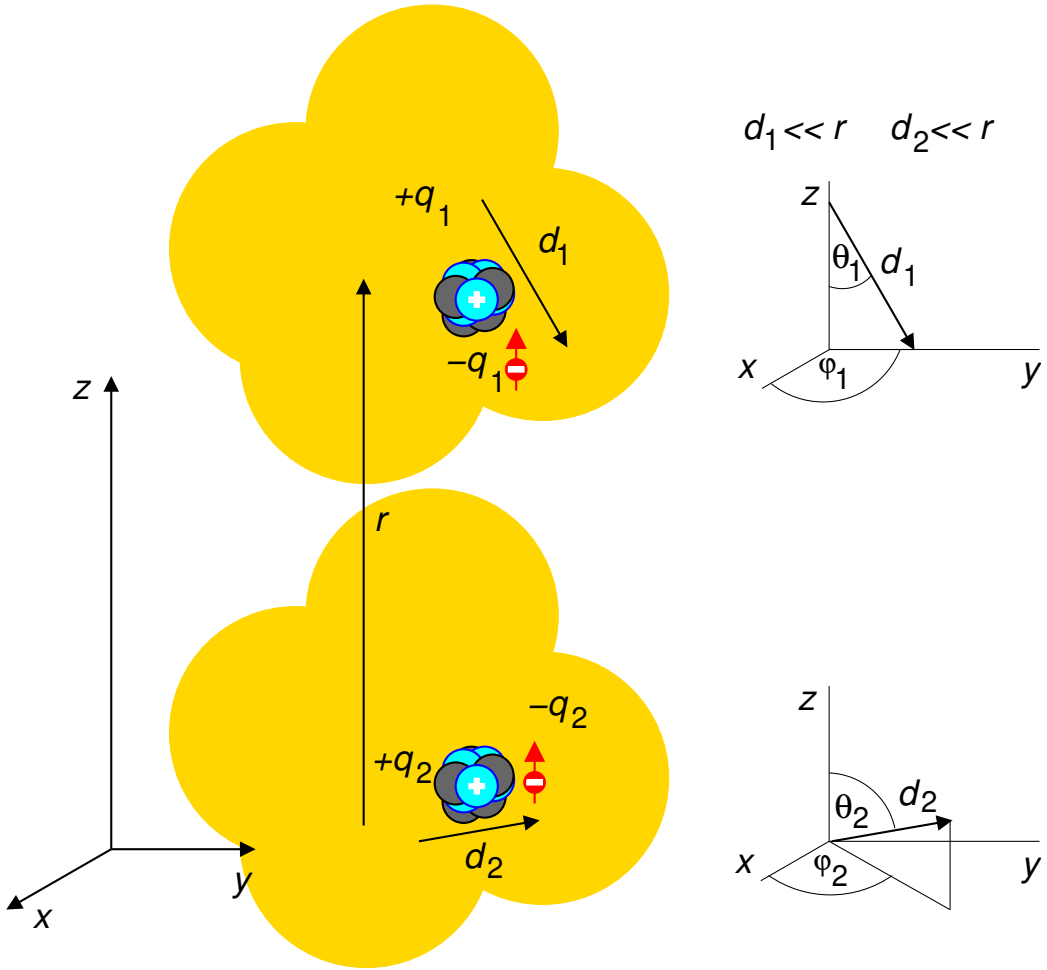
If we assume that the dipole is induced in the direction of $\vec{E}$, at each point of the molecule,

$$\mathrm{d}U' = -q\vec{d}\cdot\mathrm{d}\vec{E}' = -\alpha\epsilon_0\vec{E}'\cdot\mathrm{d}\vec{E}' = -\frac{1}{2}\alpha\epsilon_0\mathrm{d}(E')^2$$

$$U = -\frac{1}{2}\int_0^E \alpha\epsilon_0\mathrm{d}(E')^2 = -\frac{1}{2}\alpha\epsilon_0 E^2 = -\frac{1}{2}\alpha\epsilon_0\left(\frac{1}{4\pi\epsilon_0}\frac{Q}{r^2}\right)^2$$

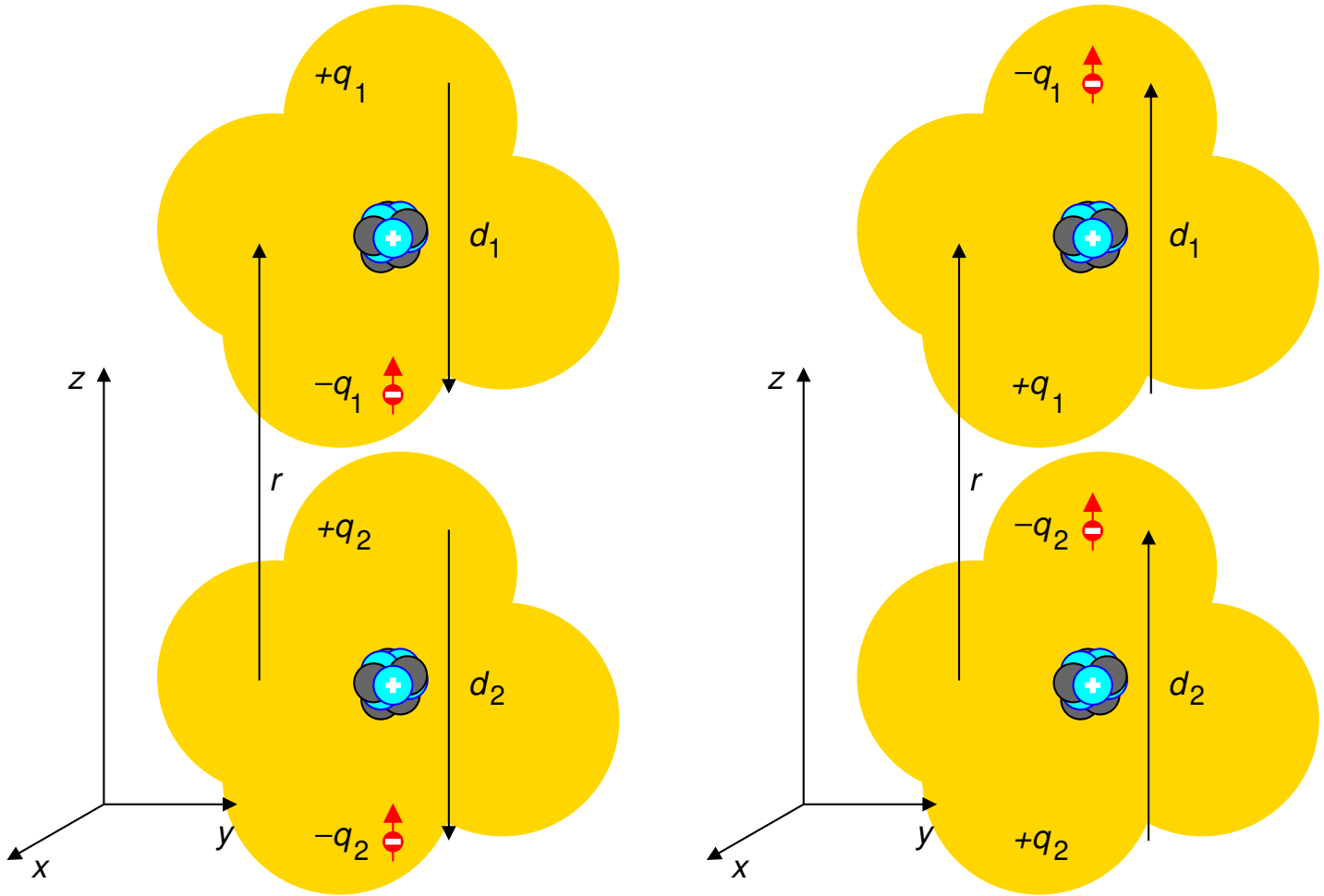
Induced dipole – induced dipole

In principle, the same relation like permanent dipoles



Induced dipole – induced dipole

In reality vibrations: $\langle U \rangle = -3h\alpha\nu/4r^6$ (identical molecules)



Summary of electrostatic interactions

Interaction	Energy
ion – ion	$C\frac{Q_1 Q_2}{r}$
ion – permanent dipole	$C\frac{qdQ}{r^2}\cos\theta$
- in different molecules	$\frac{C^2}{3RT}\frac{q^2 d^2 Q^2}{r^4}$
permanent dipole – permanent dipole	$C\frac{q_1 d_1 q_2 d_2}{r^3}K$
- in different molecules	$-\frac{2C^2}{3RT}\frac{q_1^2 d_1^2 q_2^2 d_2^2}{r^6}$
charge – induced dipole	$-\frac{C^2\epsilon_0}{2}\frac{\alpha Q^2}{r^4}$
permanent dipole – induced dipole	$-C^2\epsilon_0\frac{\alpha q^2 d^2}{2r^6}(3\cos^2\theta+1)$
- in different molecules	$-C^2\epsilon_0\frac{\alpha q^2 d^2}{r^6}$
induced dipole – induced dipole	$-\frac{h}{2}\frac{\nu_1\nu_2}{\nu_1+\nu_2}\frac{\alpha_1\alpha_2}{r^6}$

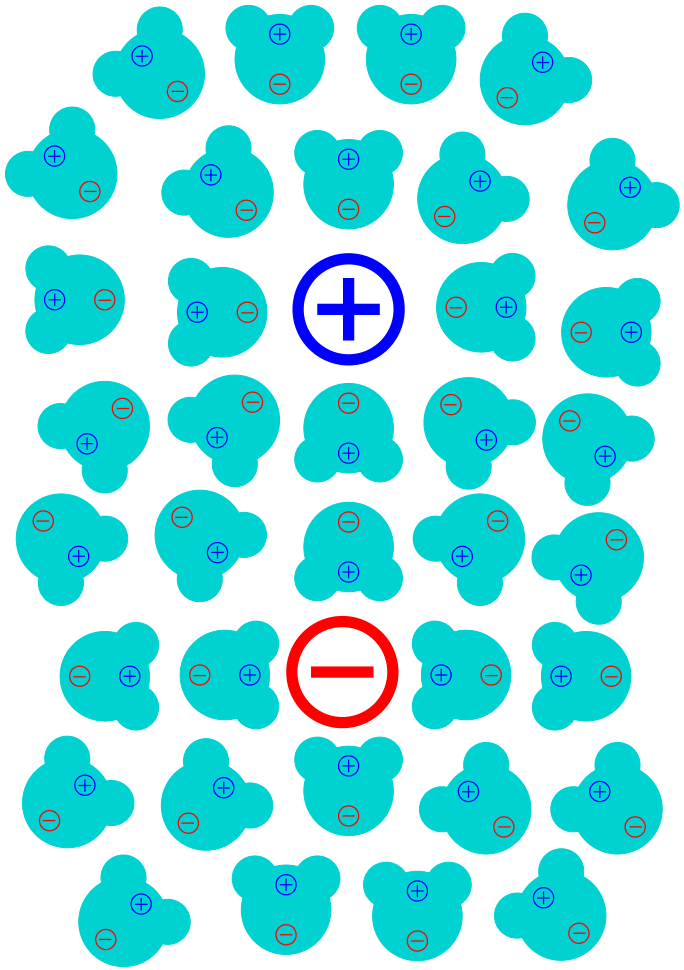
$C = 1/4\pi\epsilon_0, \quad K = \sin\theta_1 \sin\theta_2 \cos(\phi_1 - \phi_2) - 2\cos\theta_1 \cos\theta_2$

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

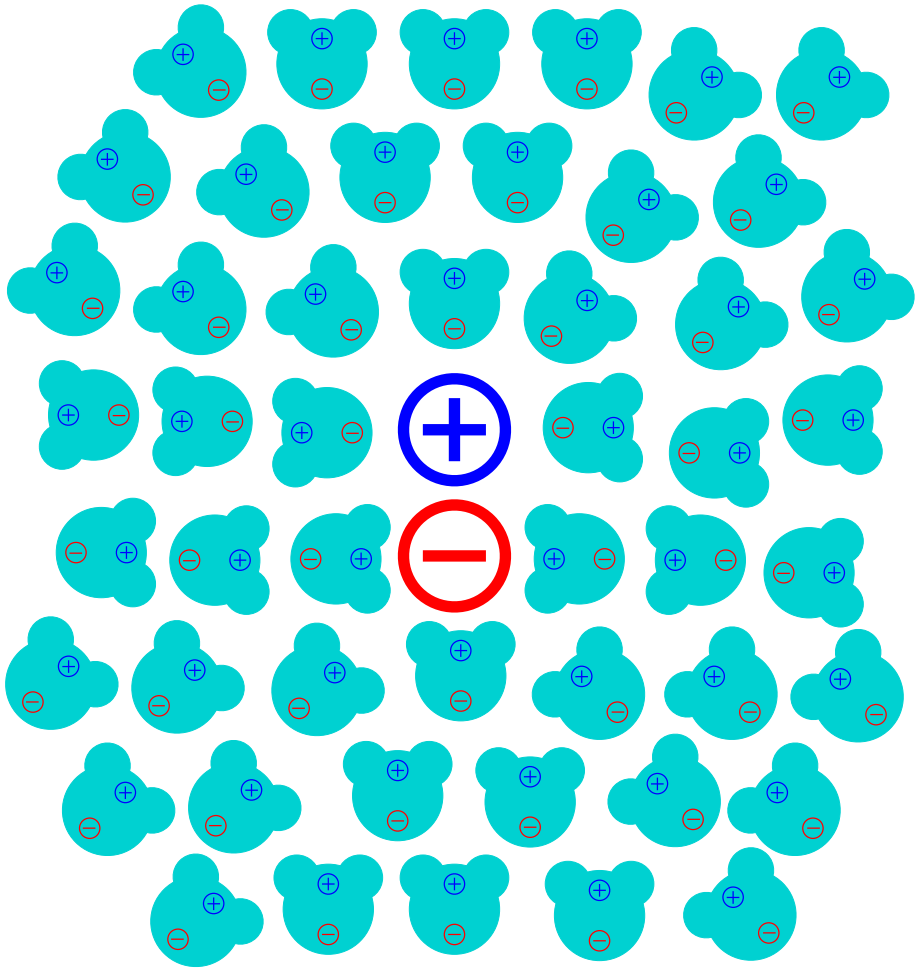
- polarization/orientation of atoms/groups in the molecule
- orientation of solvent molecules
- to maximize energy (enthalpy) of their electrostatic interactions at the cost of lowering entropy
- water does not work as an electrostatic "barrier"
- formally decreases constant in Coulomb's law
  $\Rightarrow$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

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

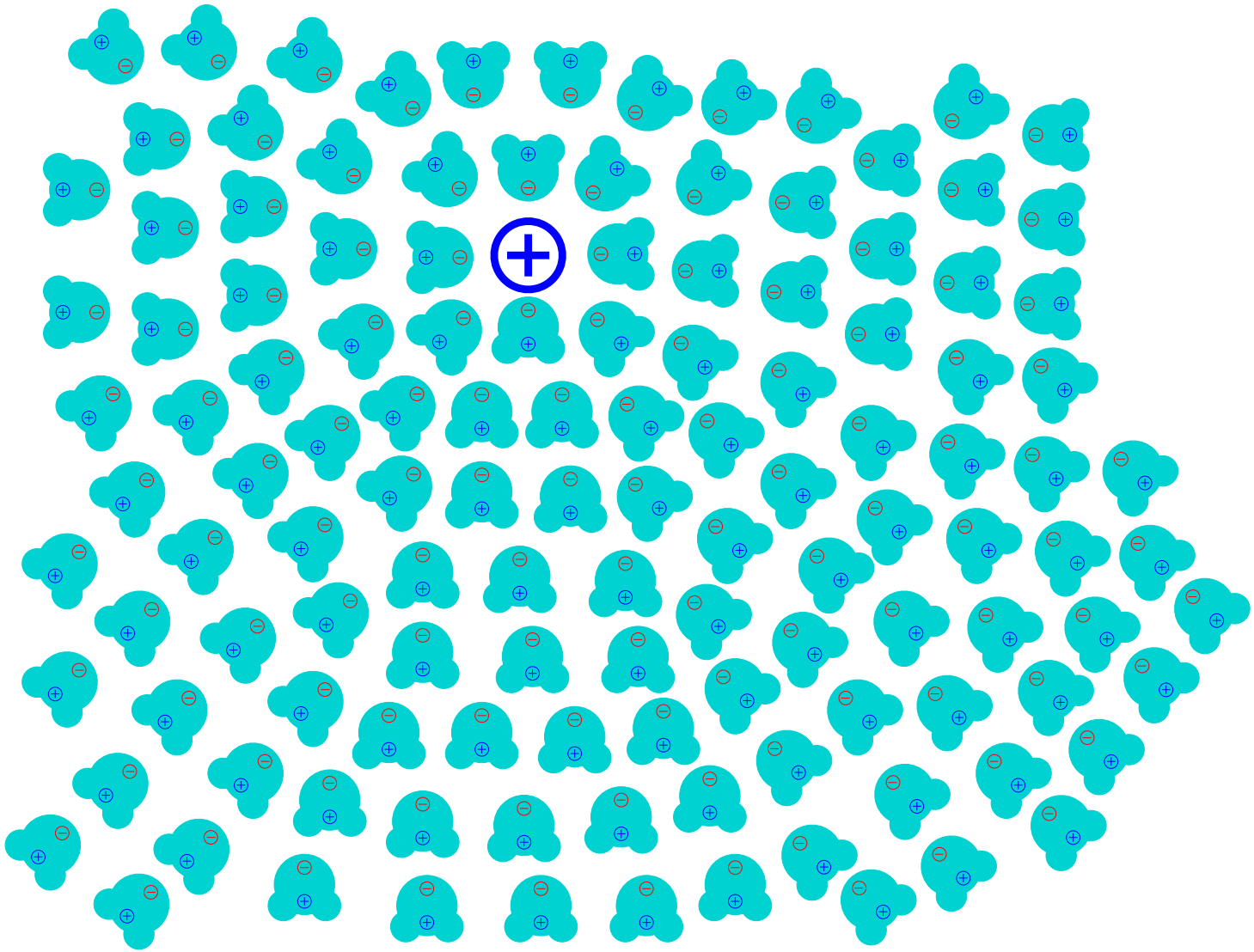
$\Delta G = 6\,\mathrm{kJ/mol}$ for charges 0.3 nm appart

Solvation of charges

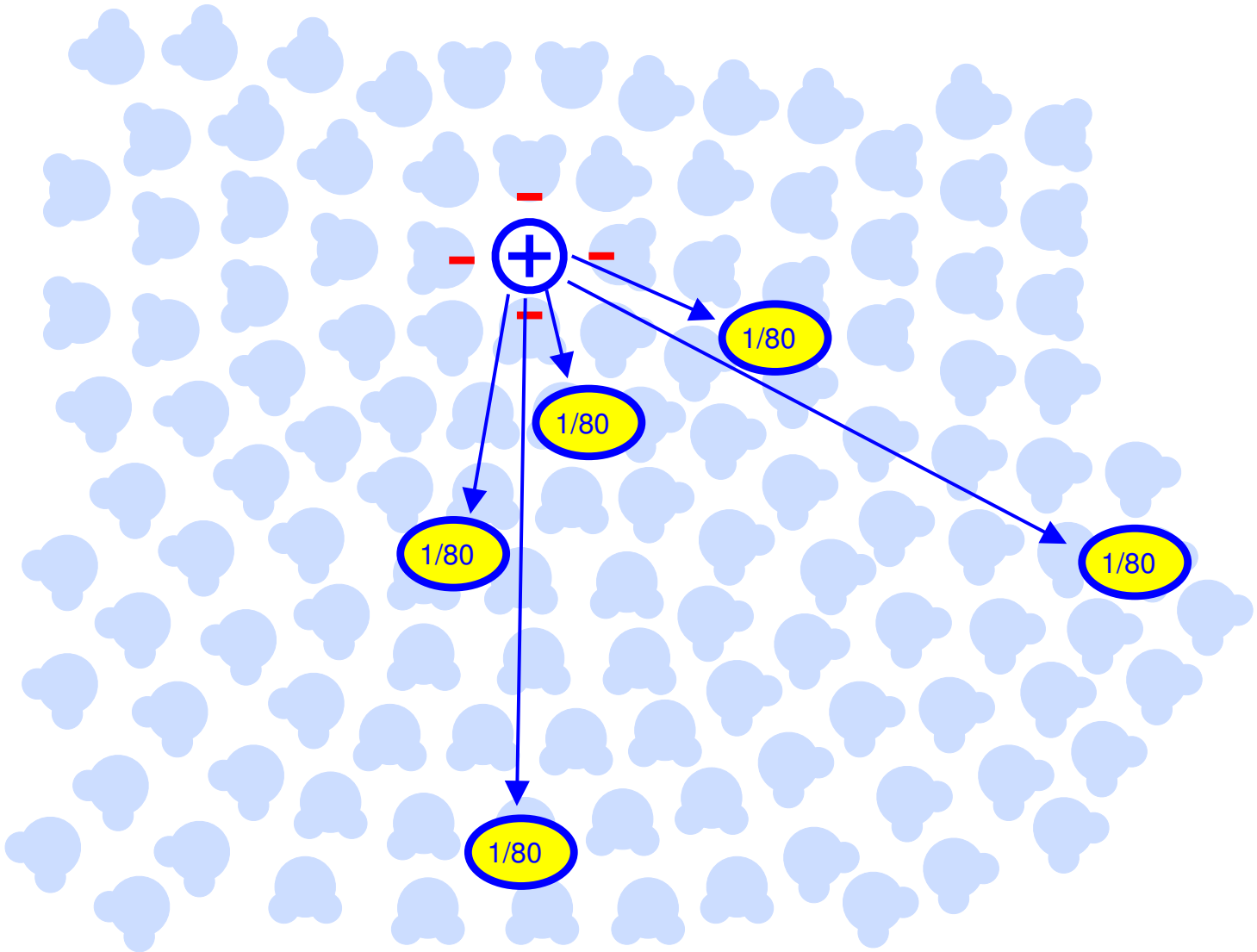
Effect of orientation of water molecules,
water does not need to be between charges



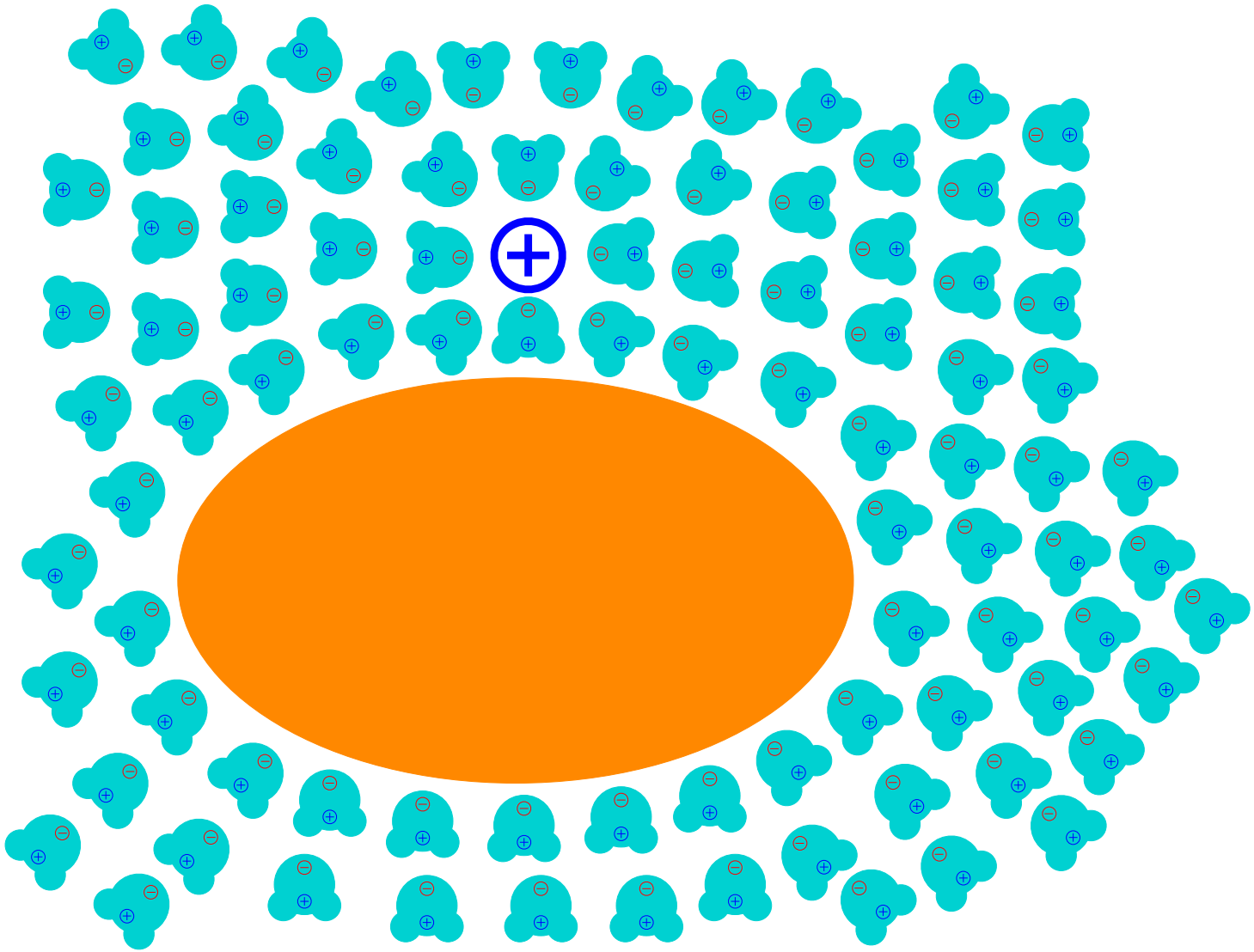
Interactions with charge in bulk water



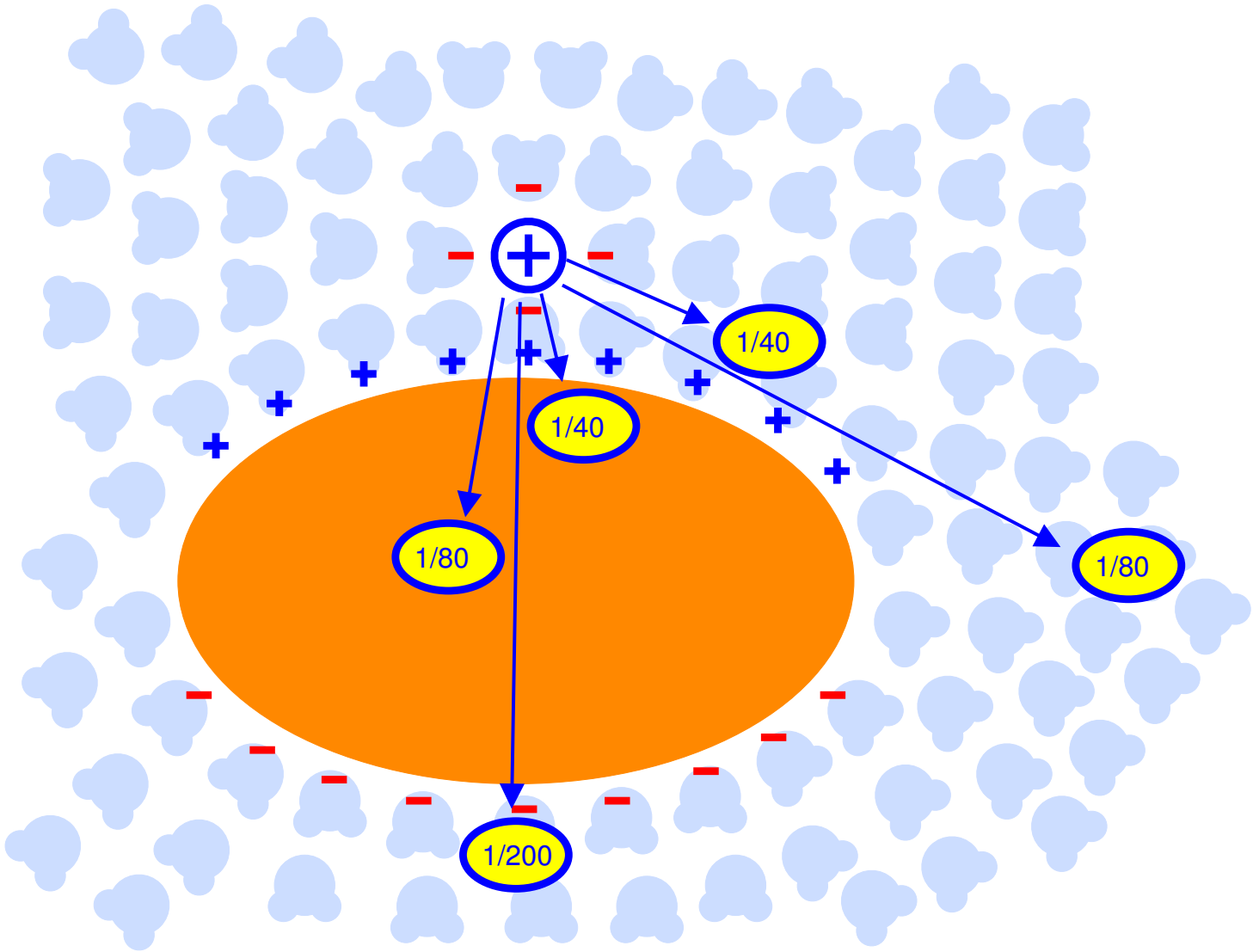
Interactions with charge in bulk water



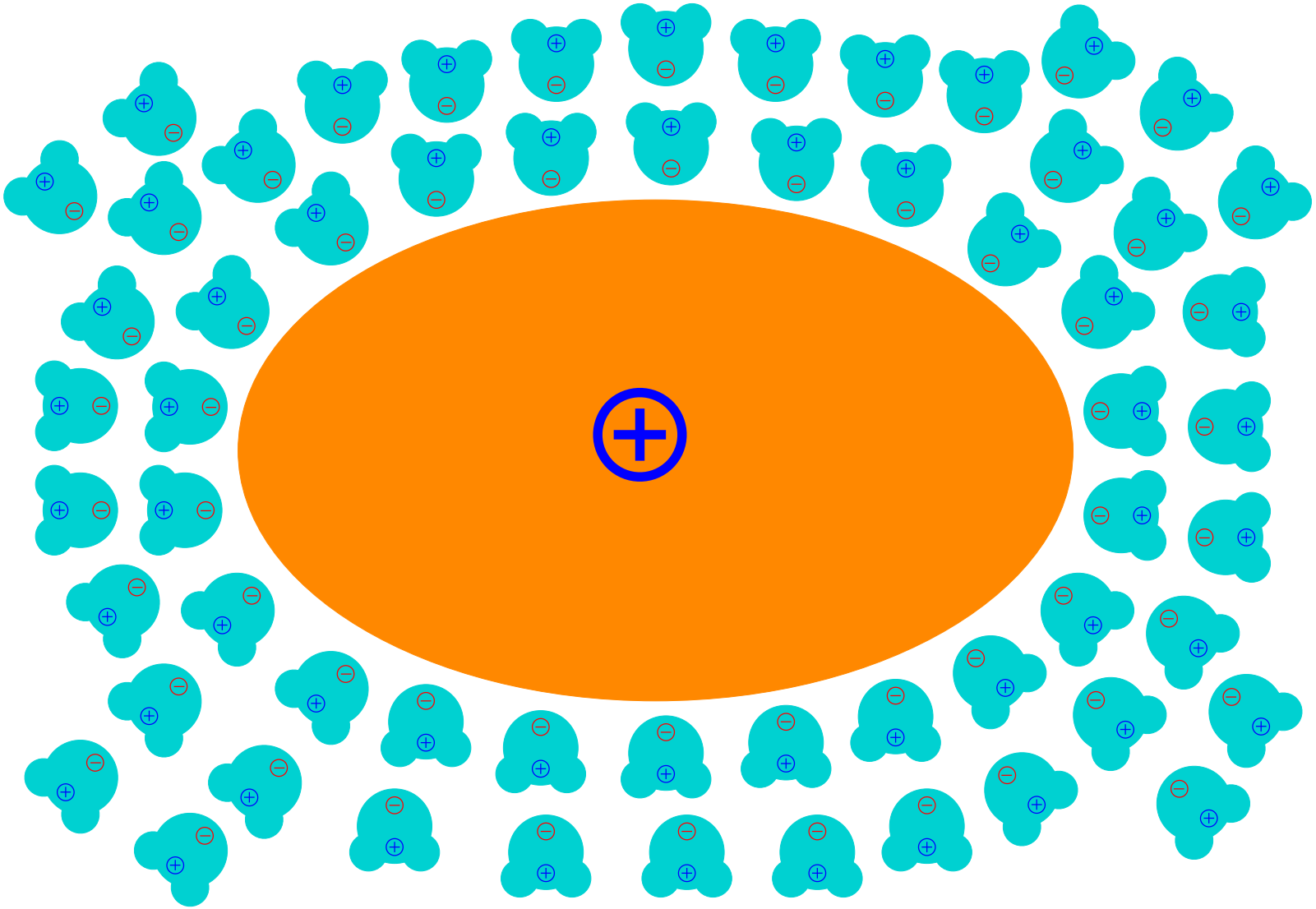
Interactions with charge at protein sufrace



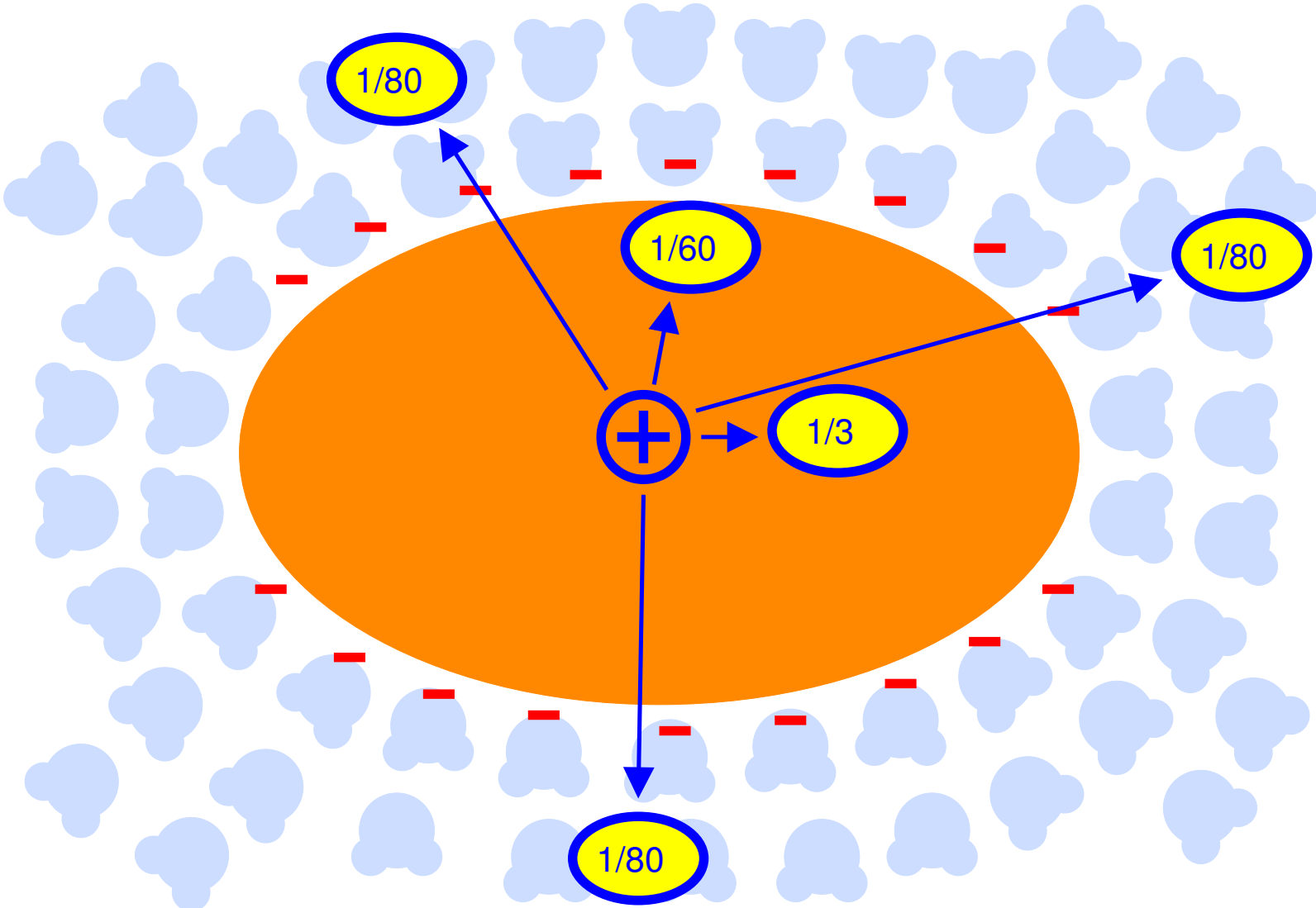
Interactions with charge at protein sufrace



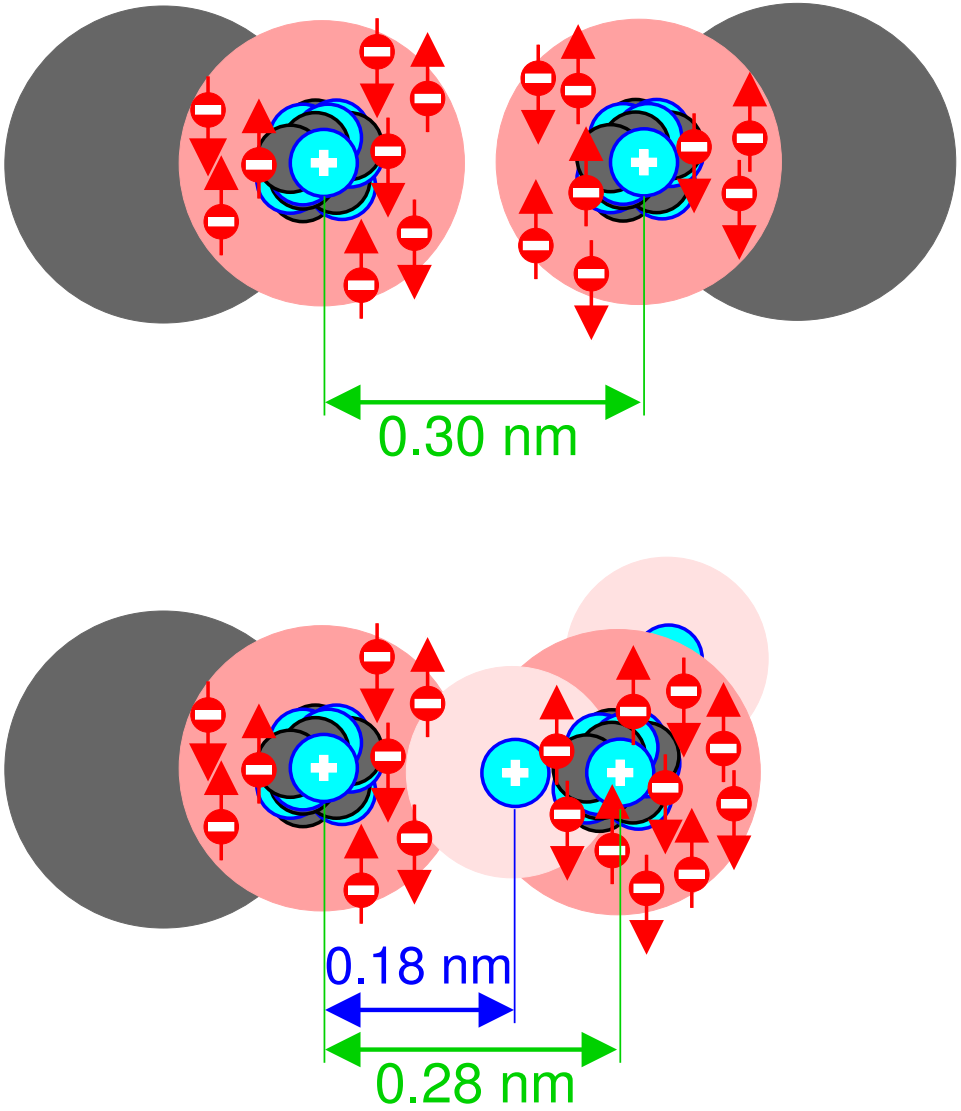
Interactions with charge inside protein



Interactions with charge inside protein



Hydrogen bonds



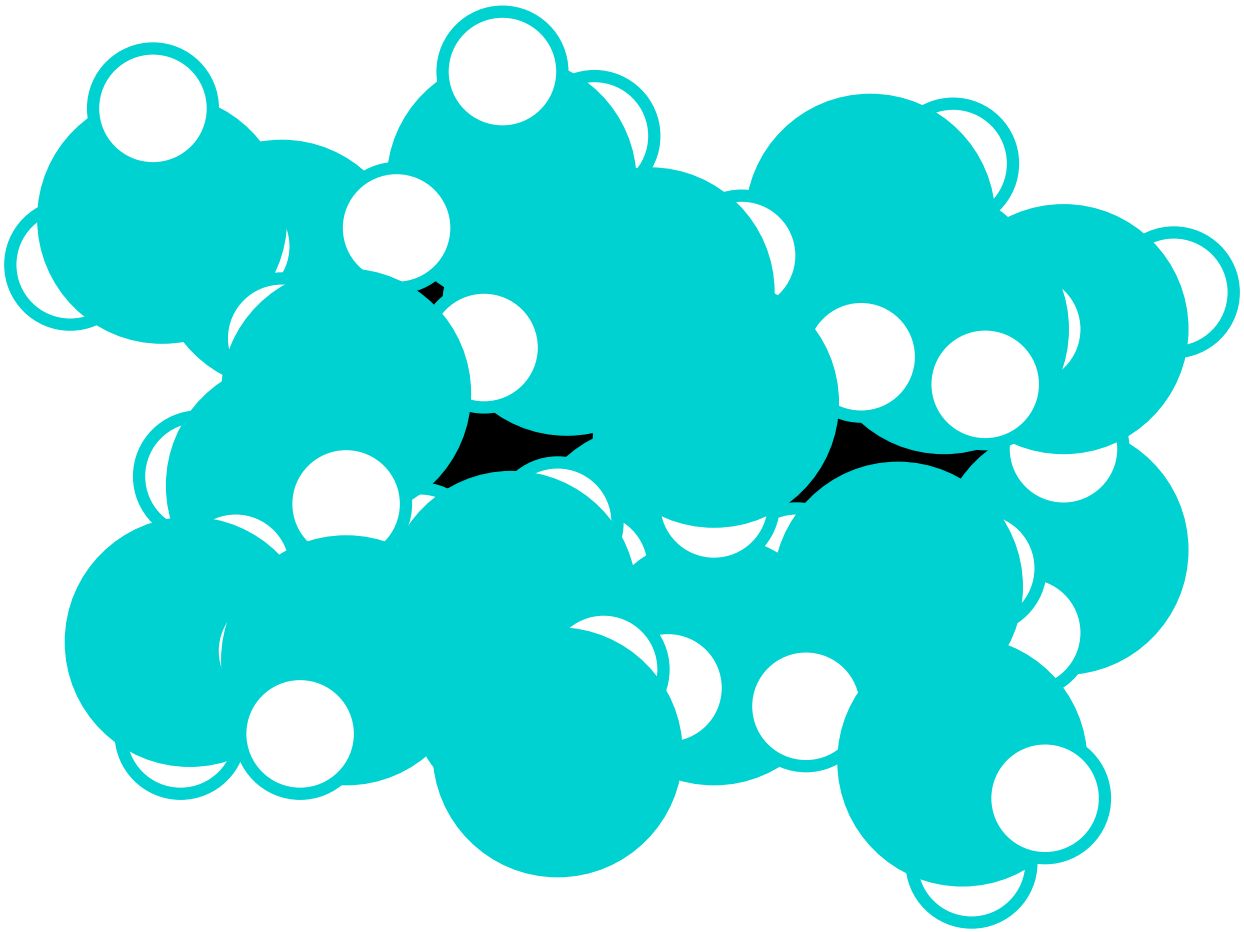
Hydrogen bonds

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	

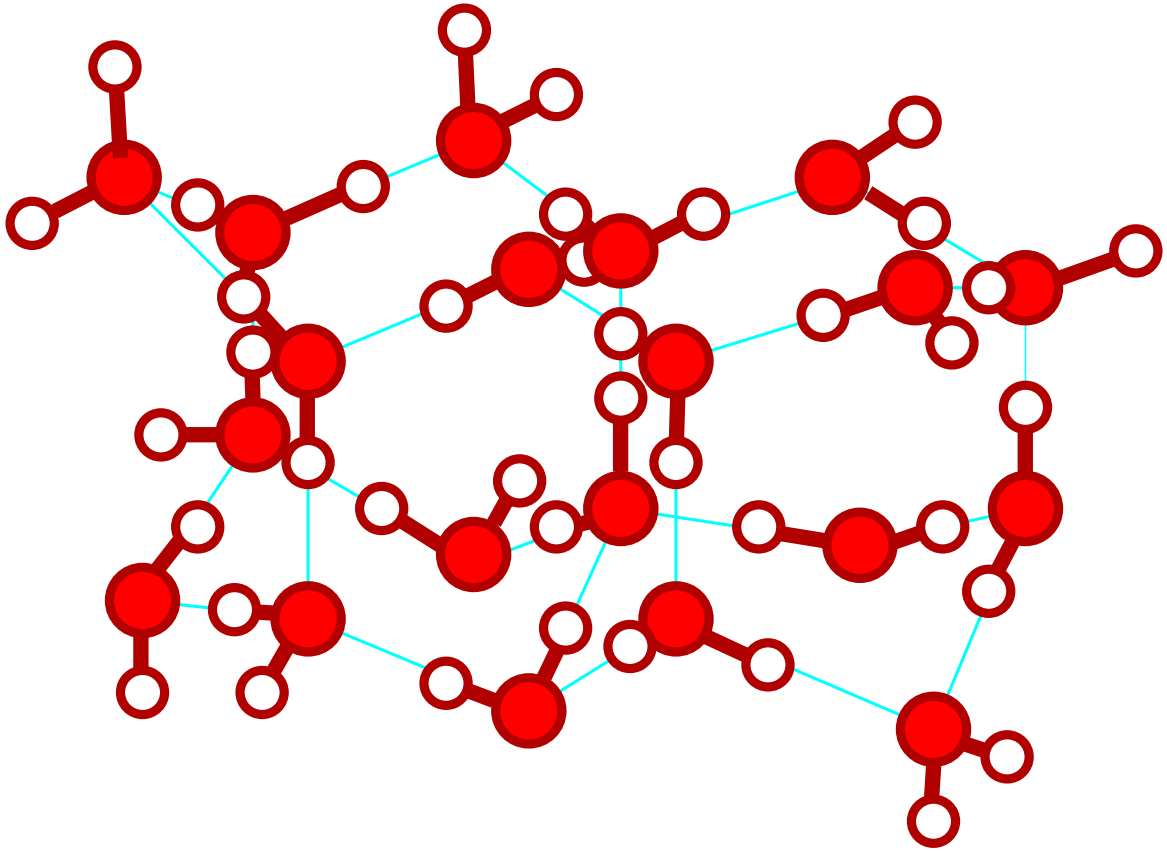
U(H-bond)$=$ 20 kJ/mol

Ice



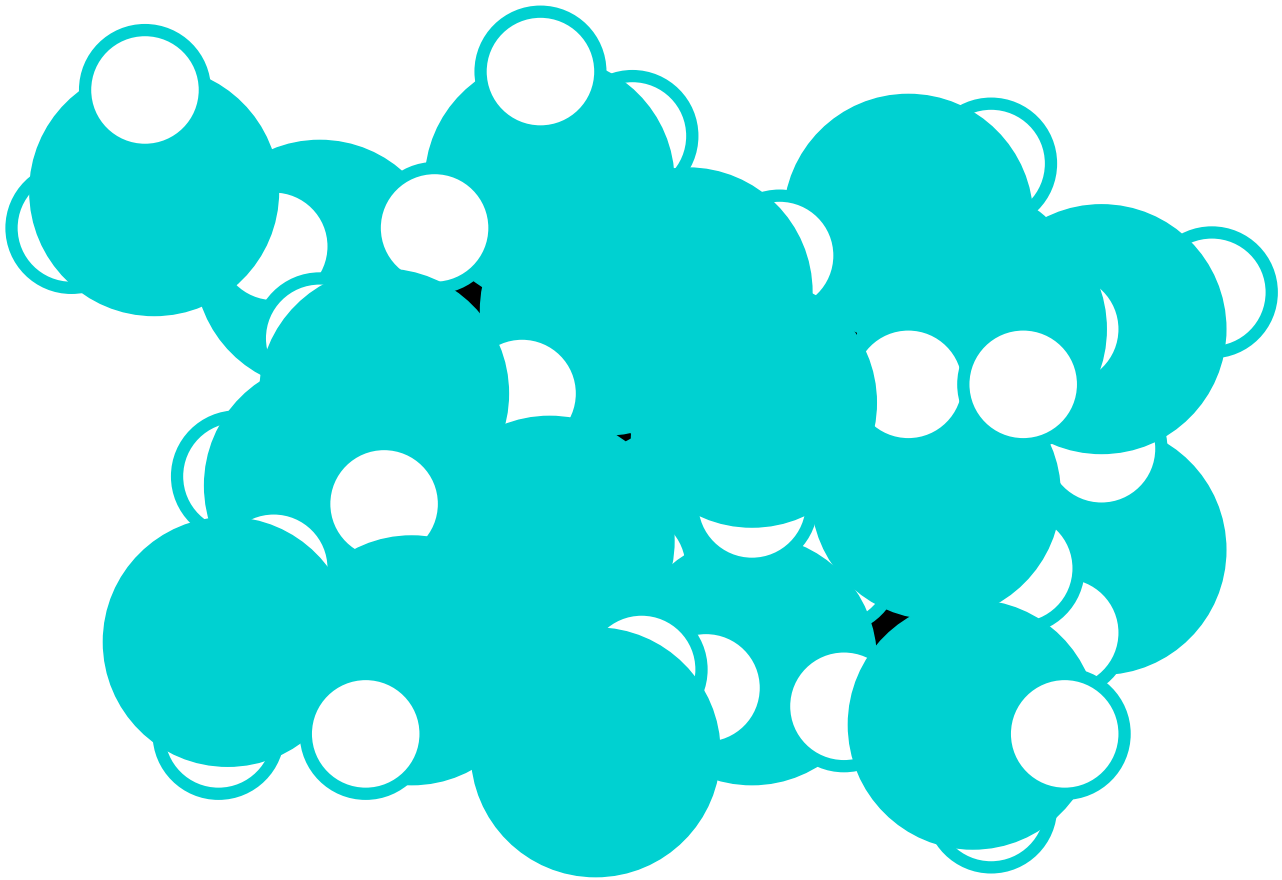
void space

Ice



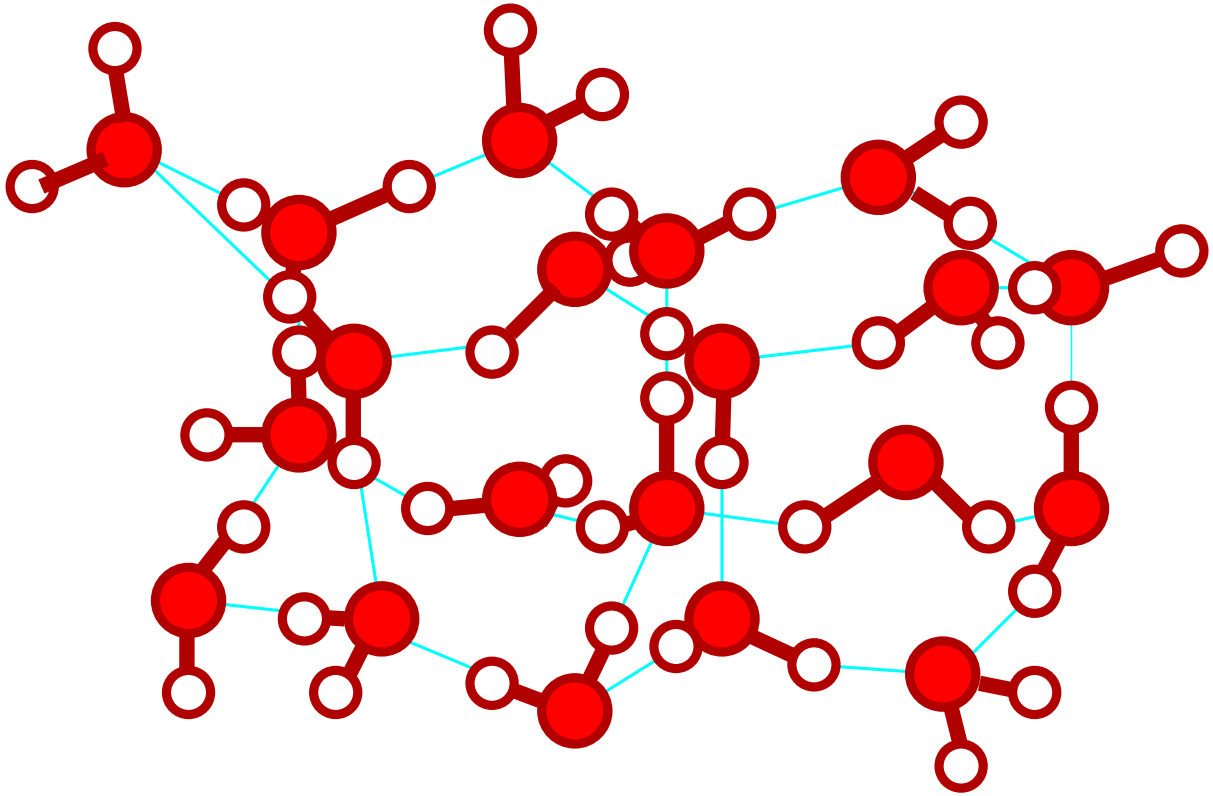
Hydrogen bonds : 50 kJ/mol

Liquid water



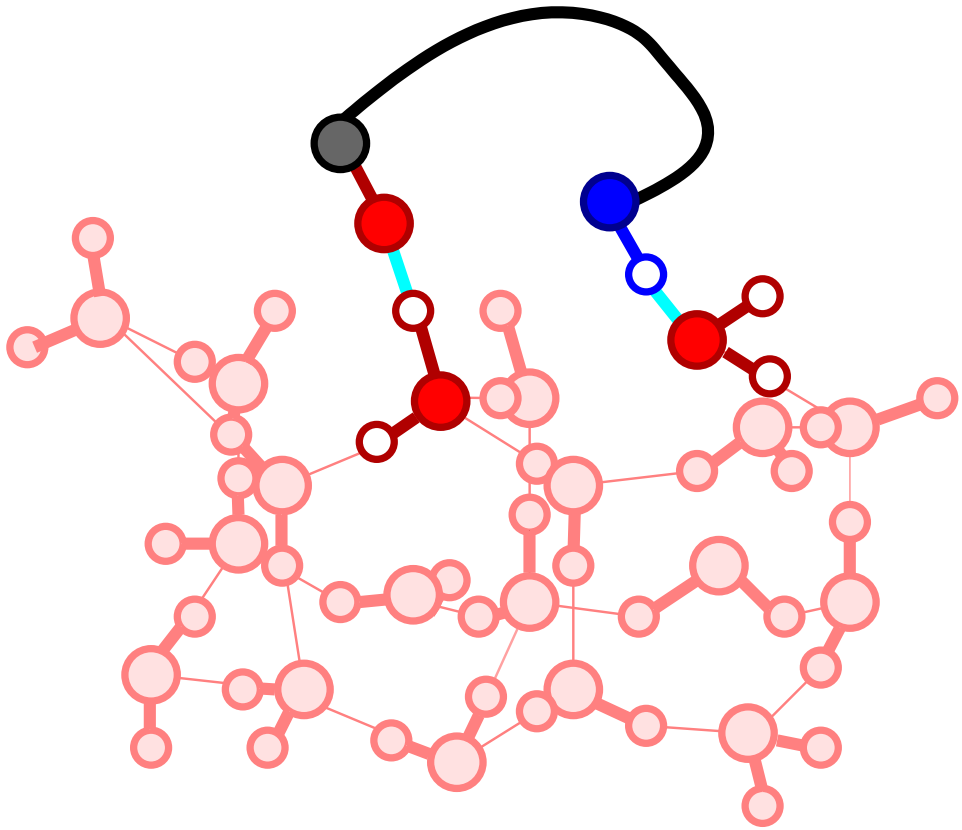
void space less than in ice

Liquid water



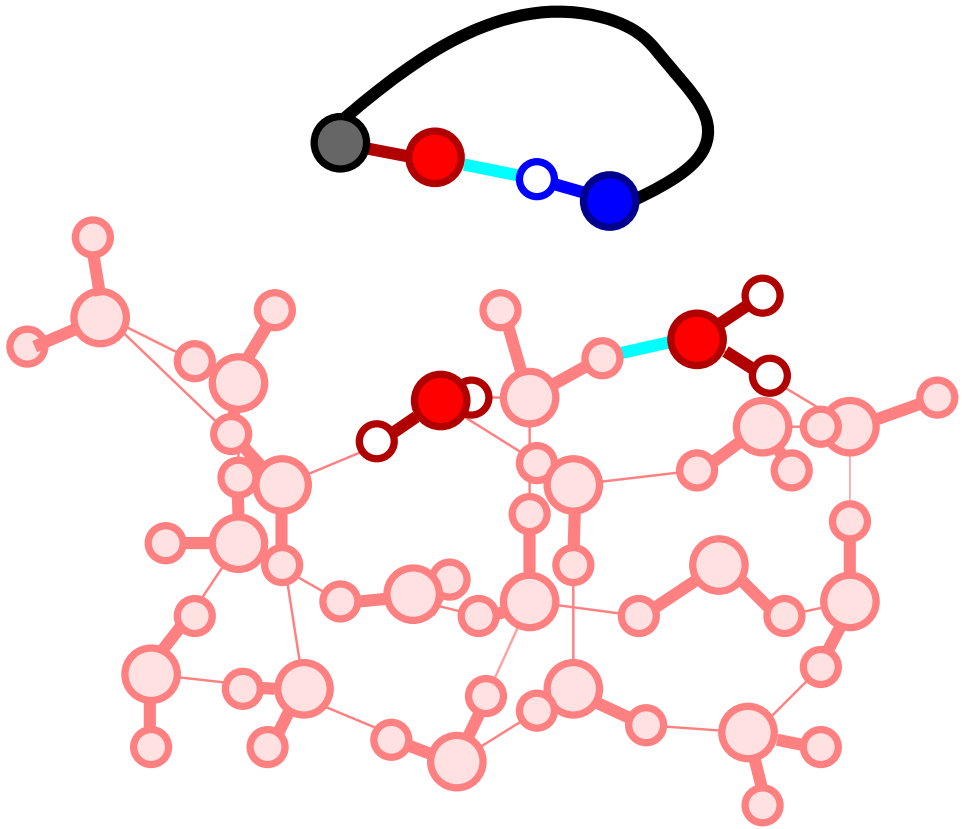
Hydrogen bonds : 40 kJ/mol

Hydrogen bonds in proteins



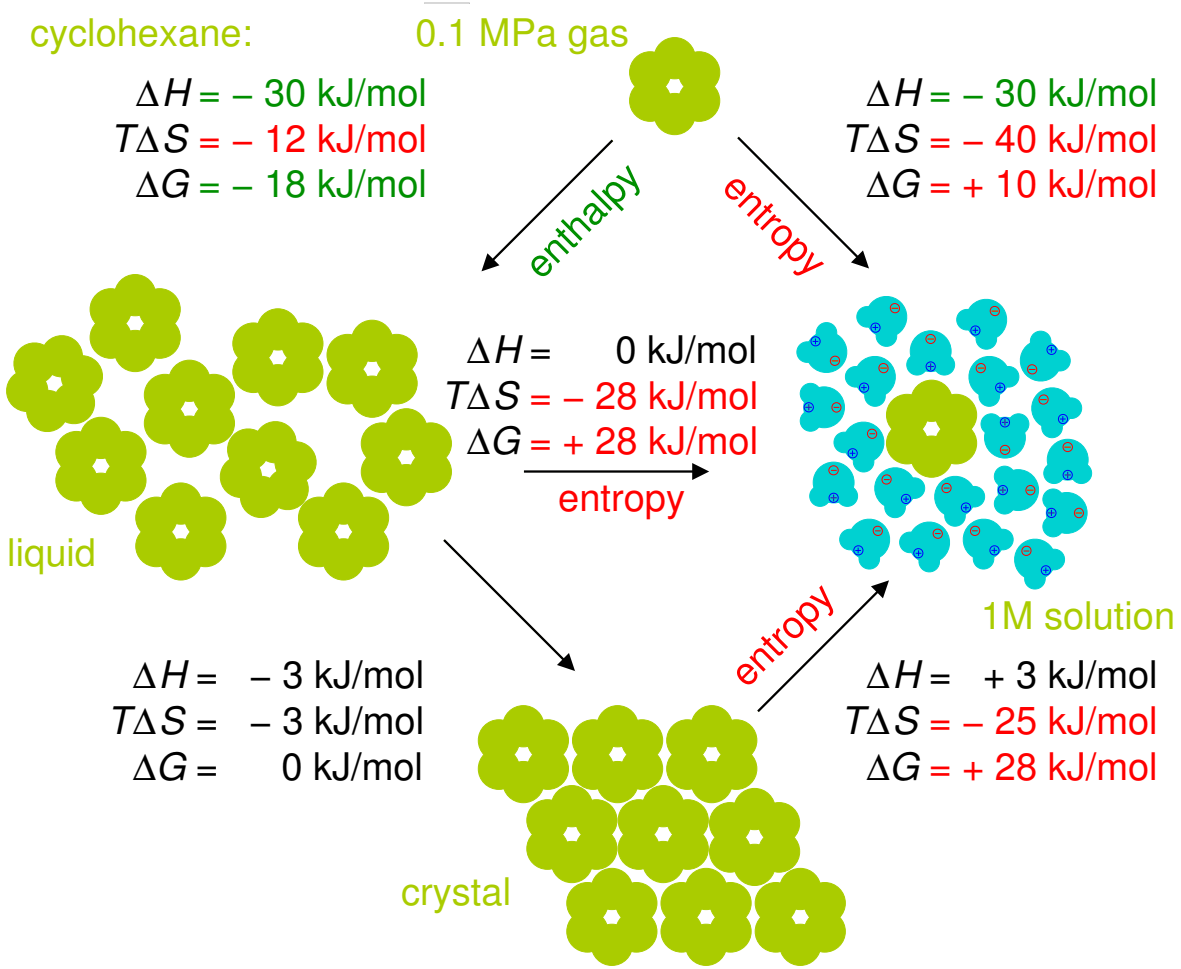
$$\Delta G = 0$$

Hydrogen bonds in proteins

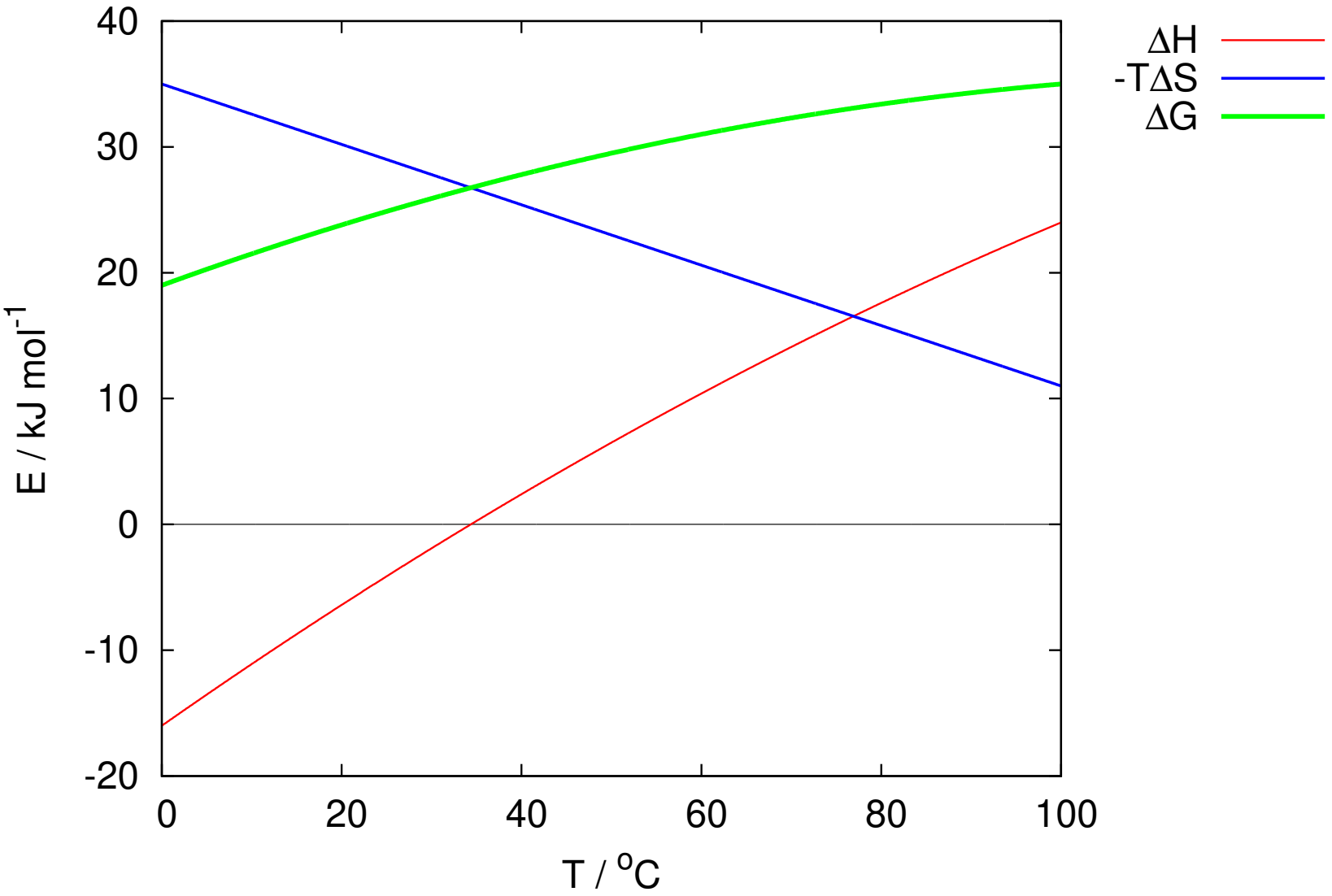


$\Delta G = -12\,\mathrm{kJ/mol}$

Hydrophobic effect

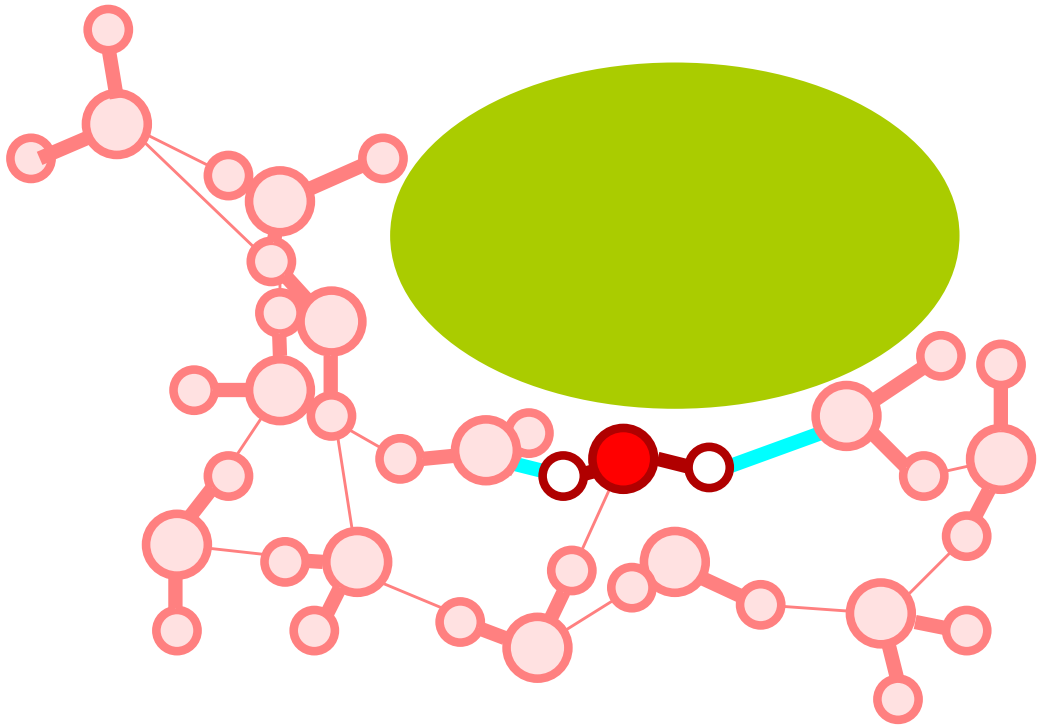


Hydrophobic effect



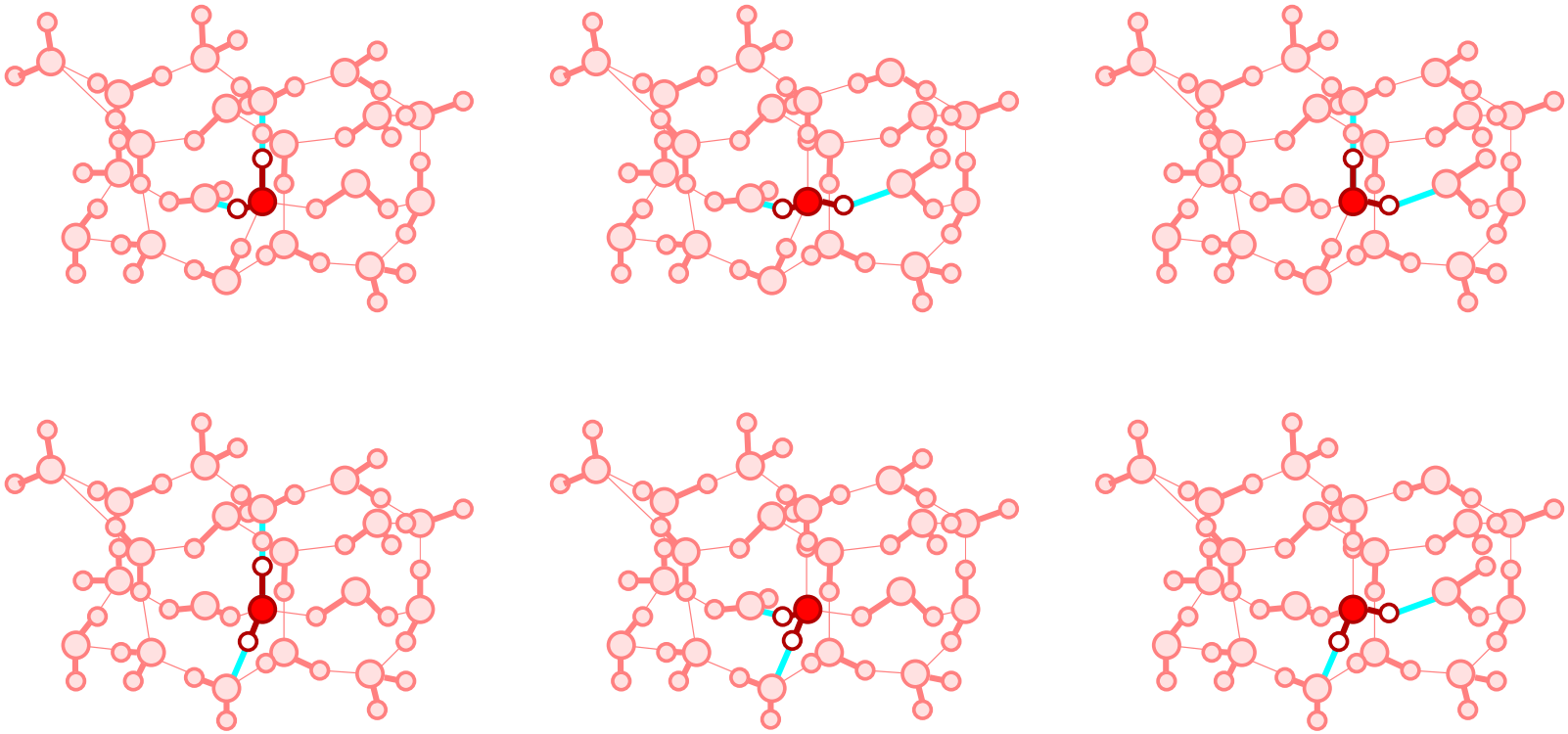
Hydrophobic effect

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



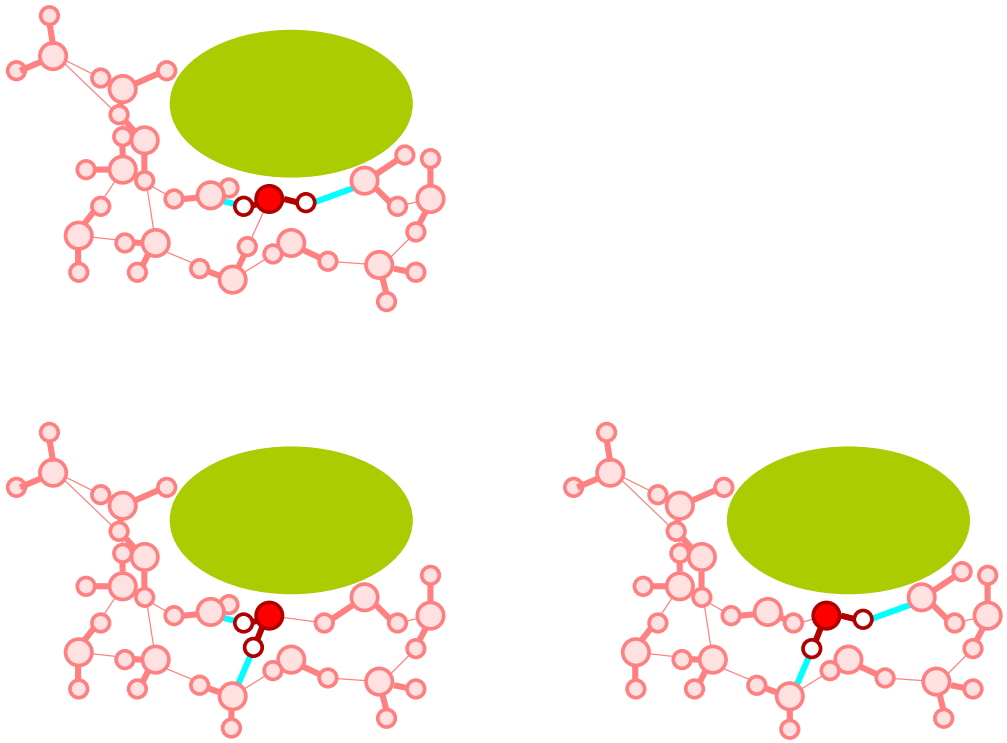
Hydrophobic effect

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

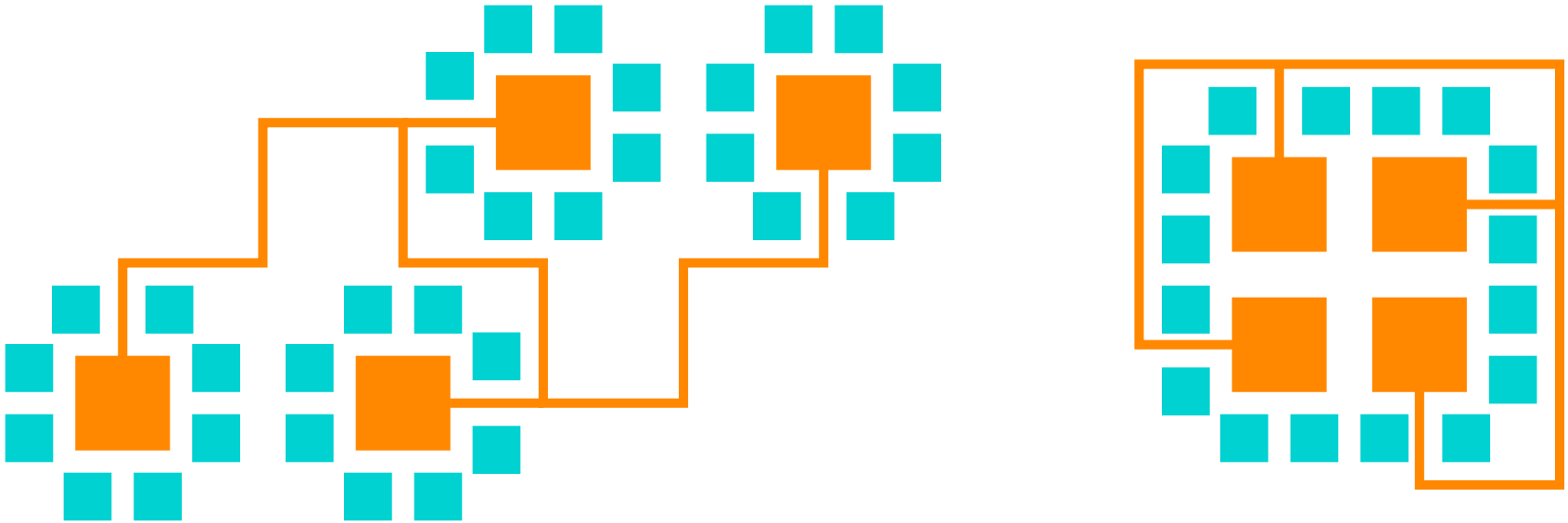


Hydrophobic effect

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



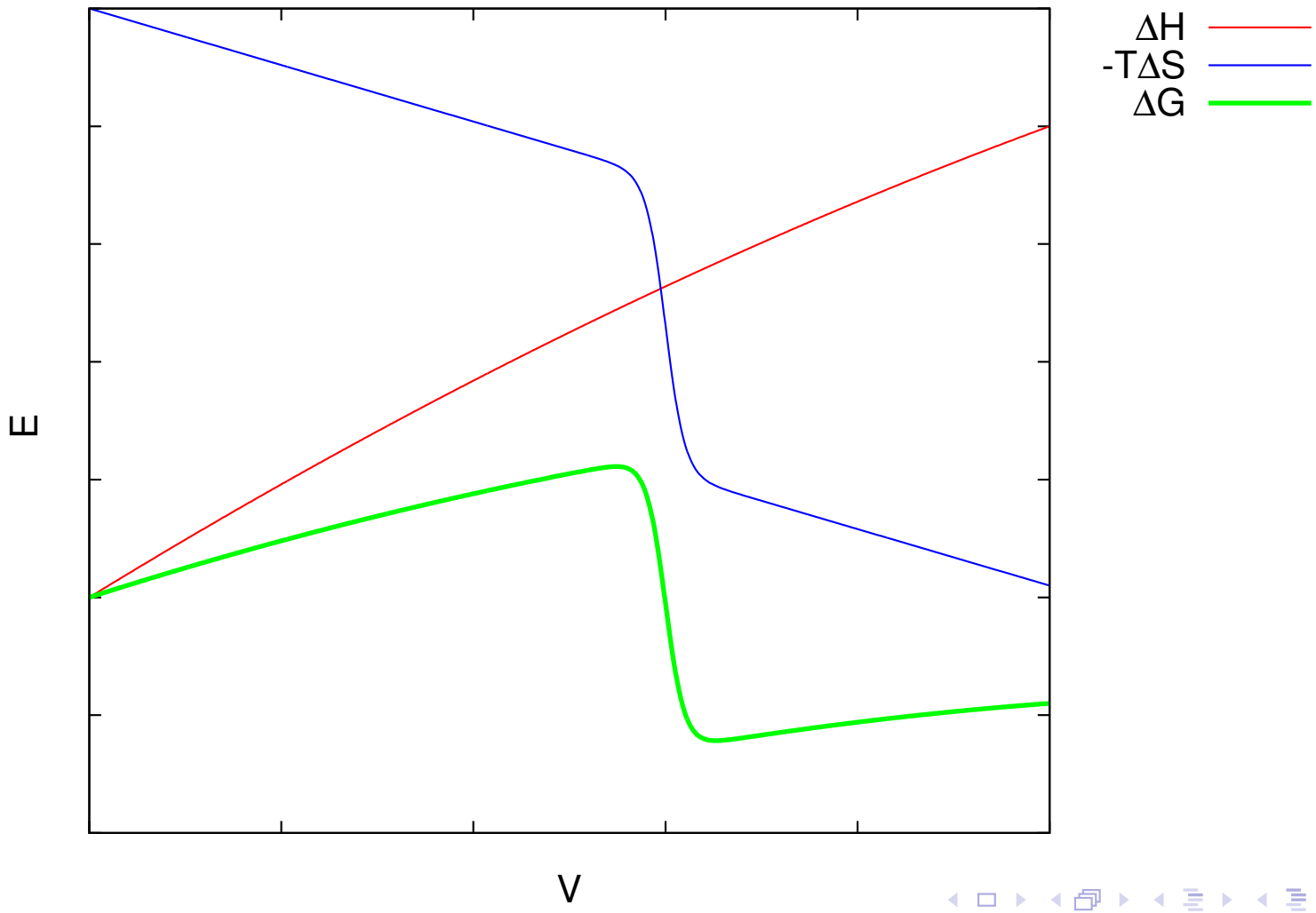
Hydrophobic effect in proteins



- 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

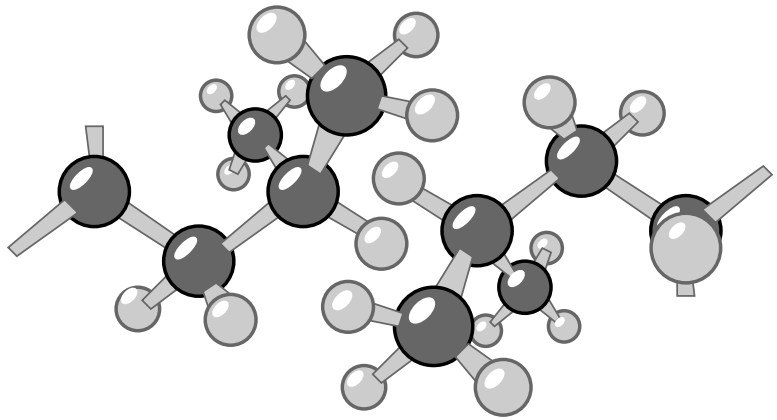
Protein stability

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



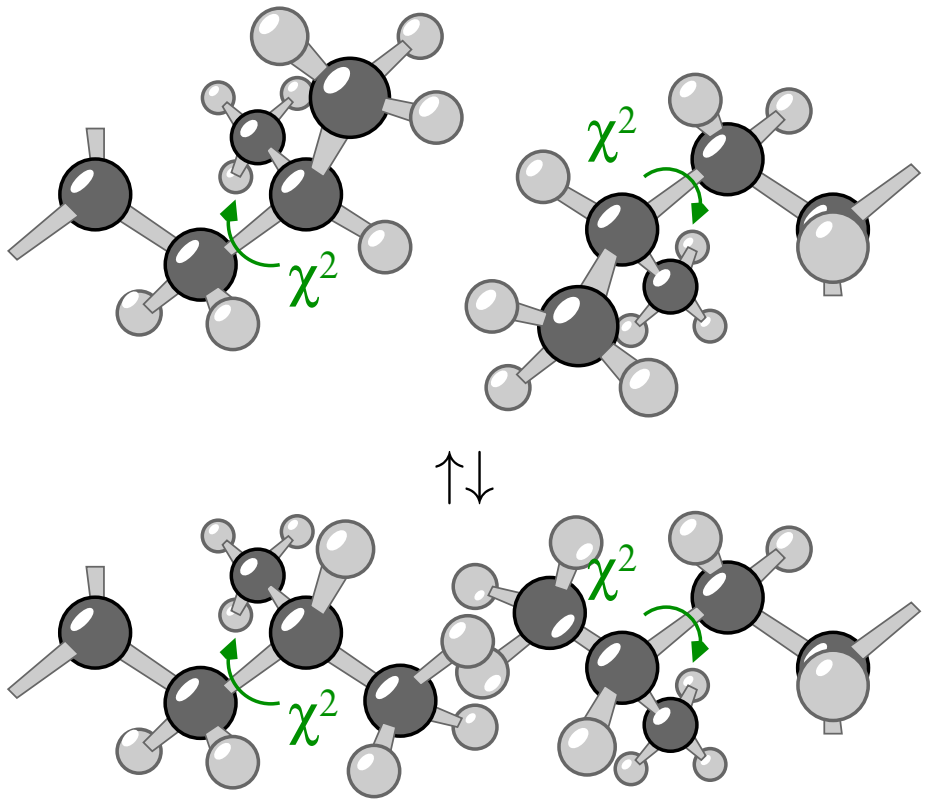
Protein stability

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

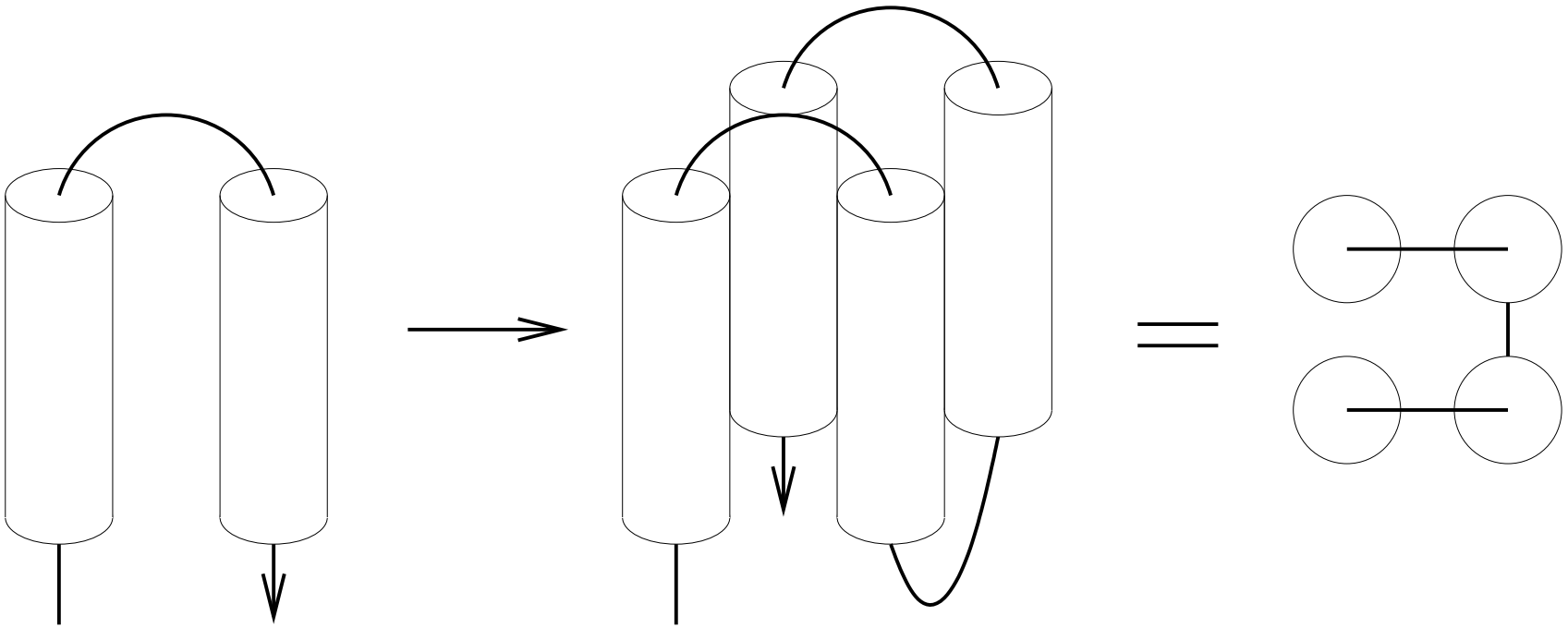
ion with charge $+1/-1$, dipole of peptide bond (1.2×10^{-29} C m)

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

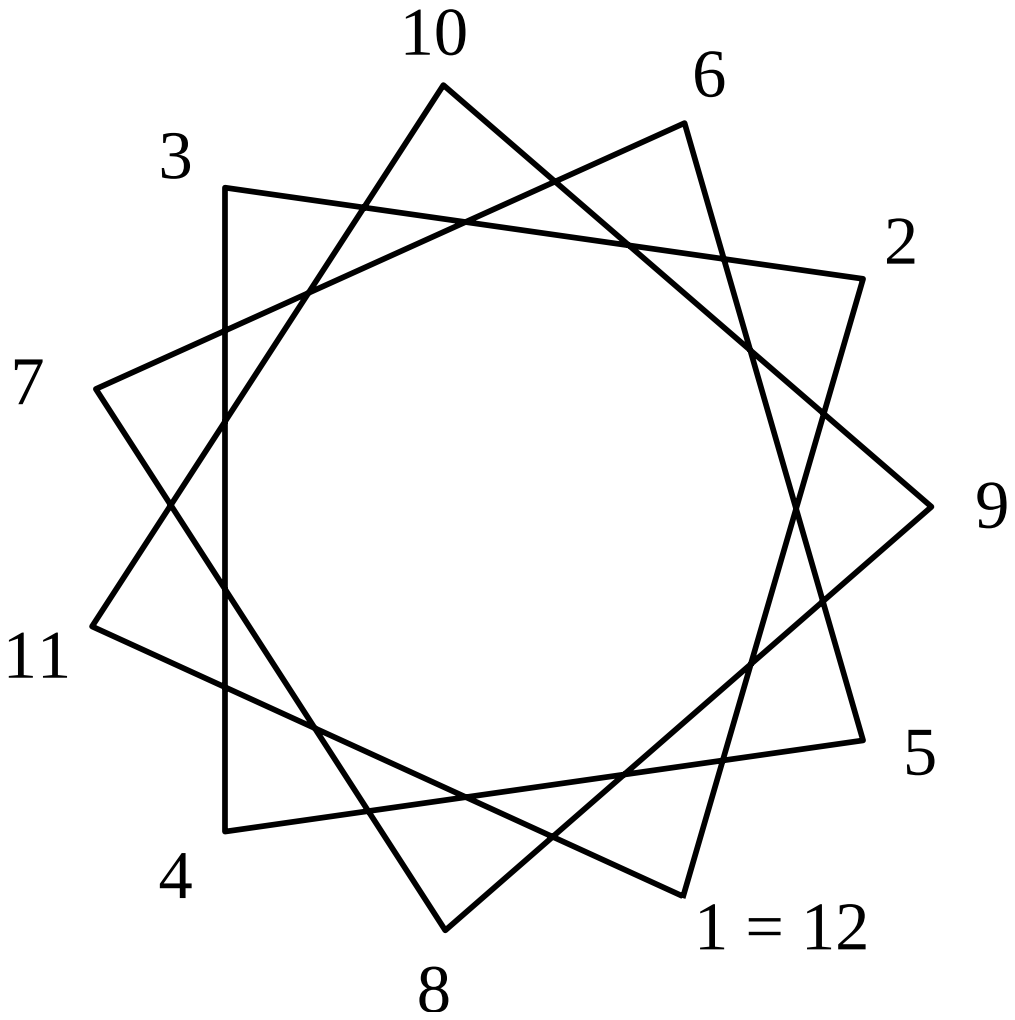
Packing of α-helices

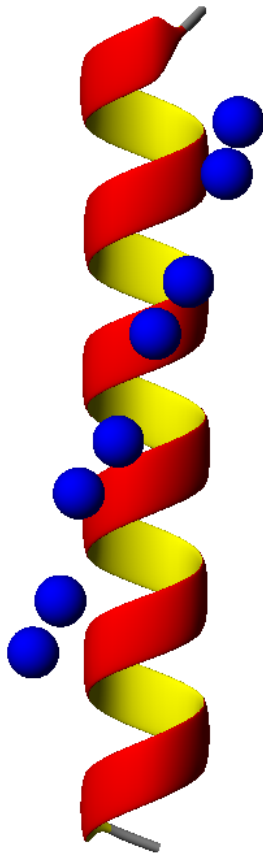
Amphiphilic helices: nonpolar sidechains inside



Packing of α-helices

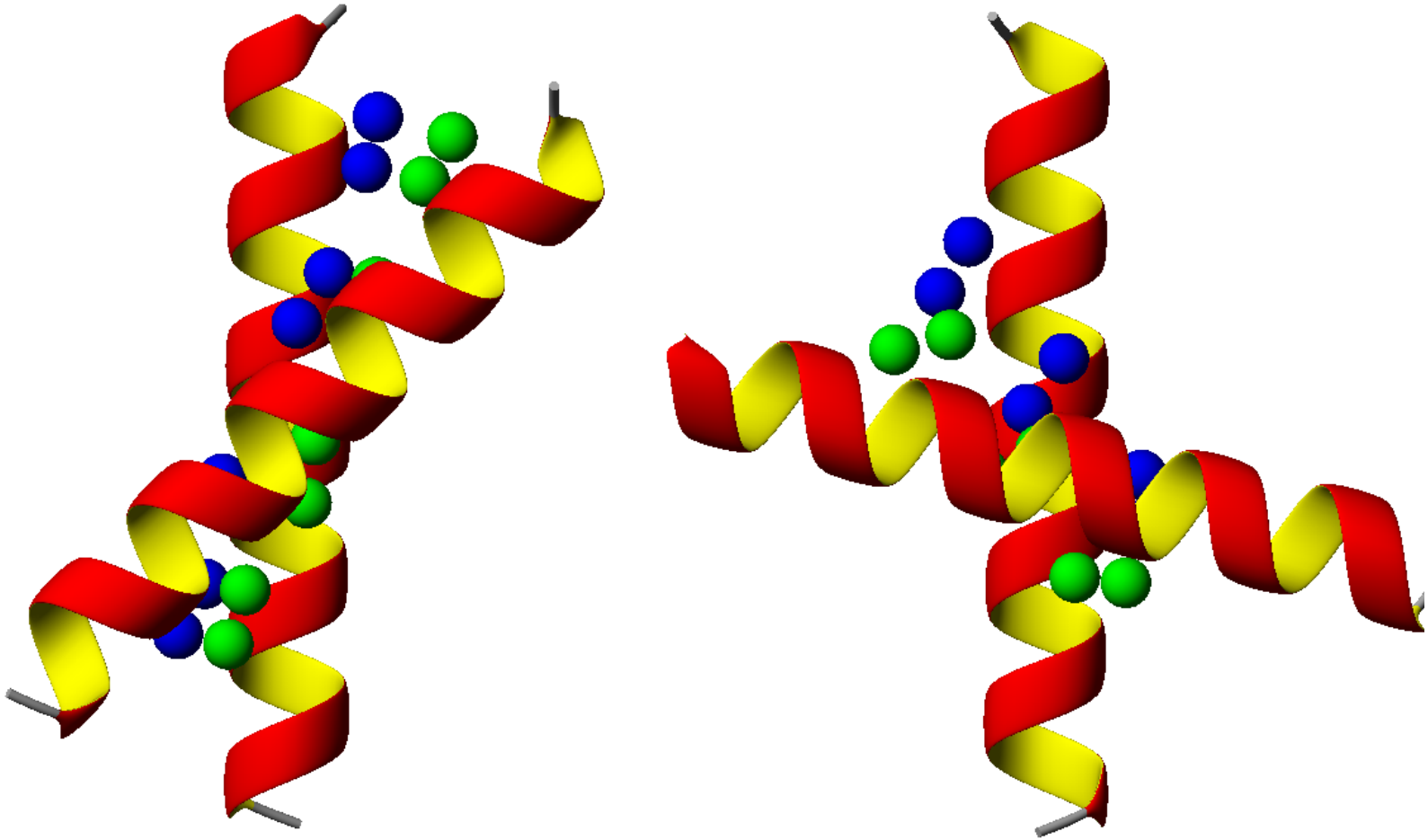
to compact nonpolar sidechains in amphiphilic helices





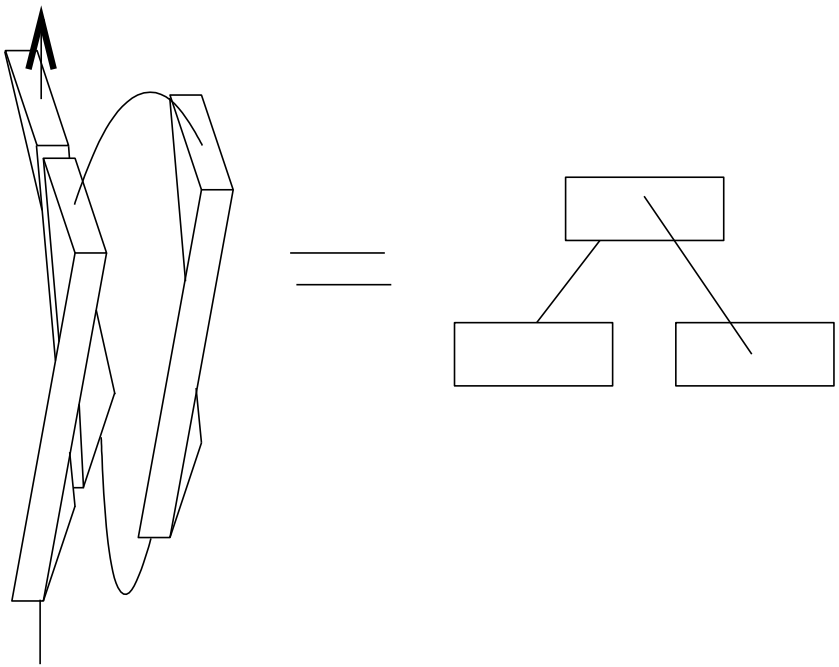
Packing of α-helices

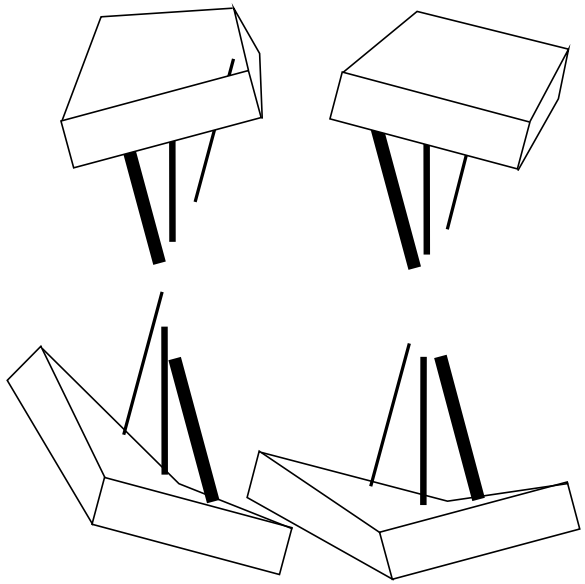
Hydrophobic side chains (blue/green) spheres packed positions 1, 5, 9, 13 (left) and 1, 4, 7 (right)



Packing of β-sandwiches

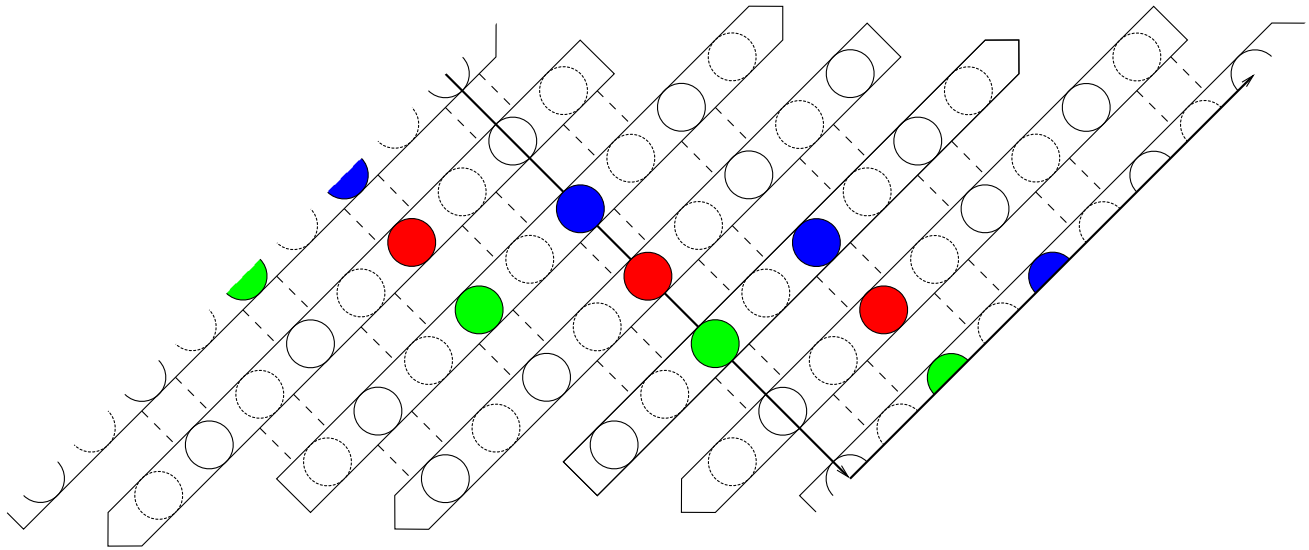
Hydrophobic residues inside

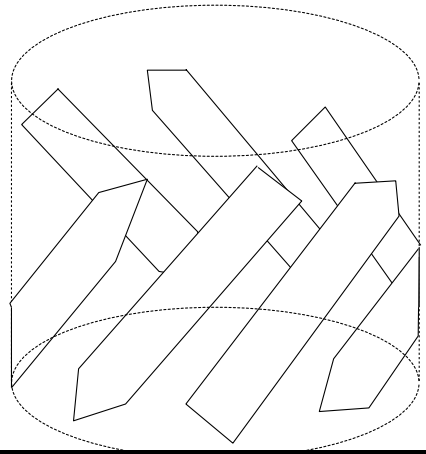


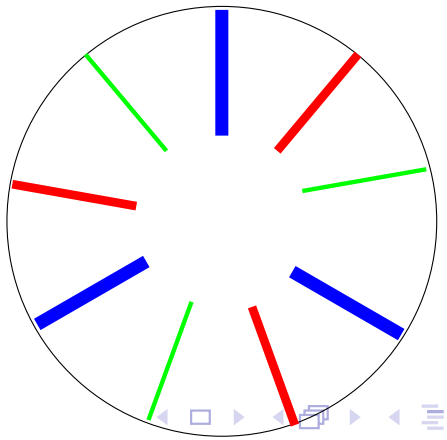


Packing of β-barrels

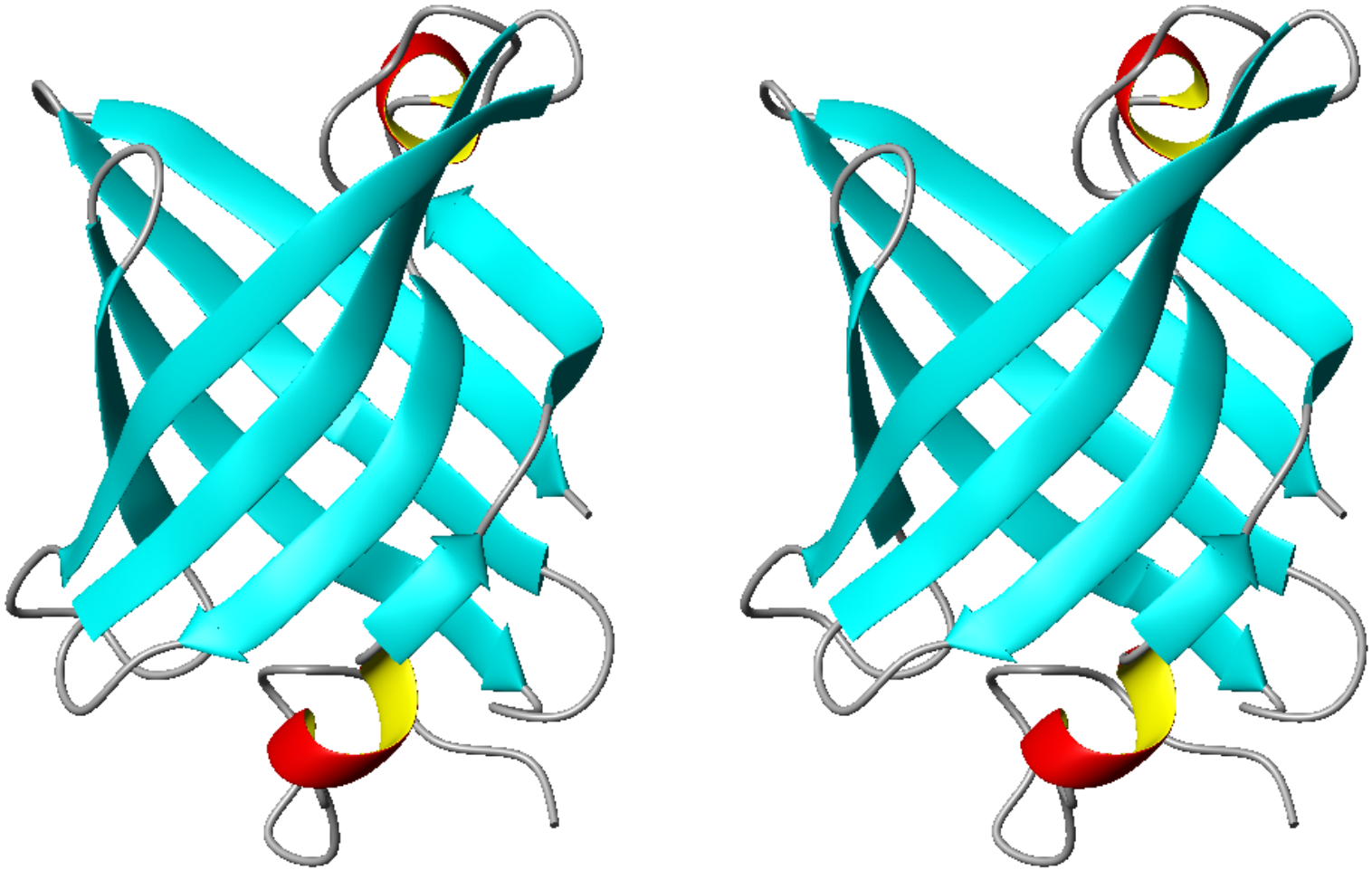
Hydrophobic residues inside



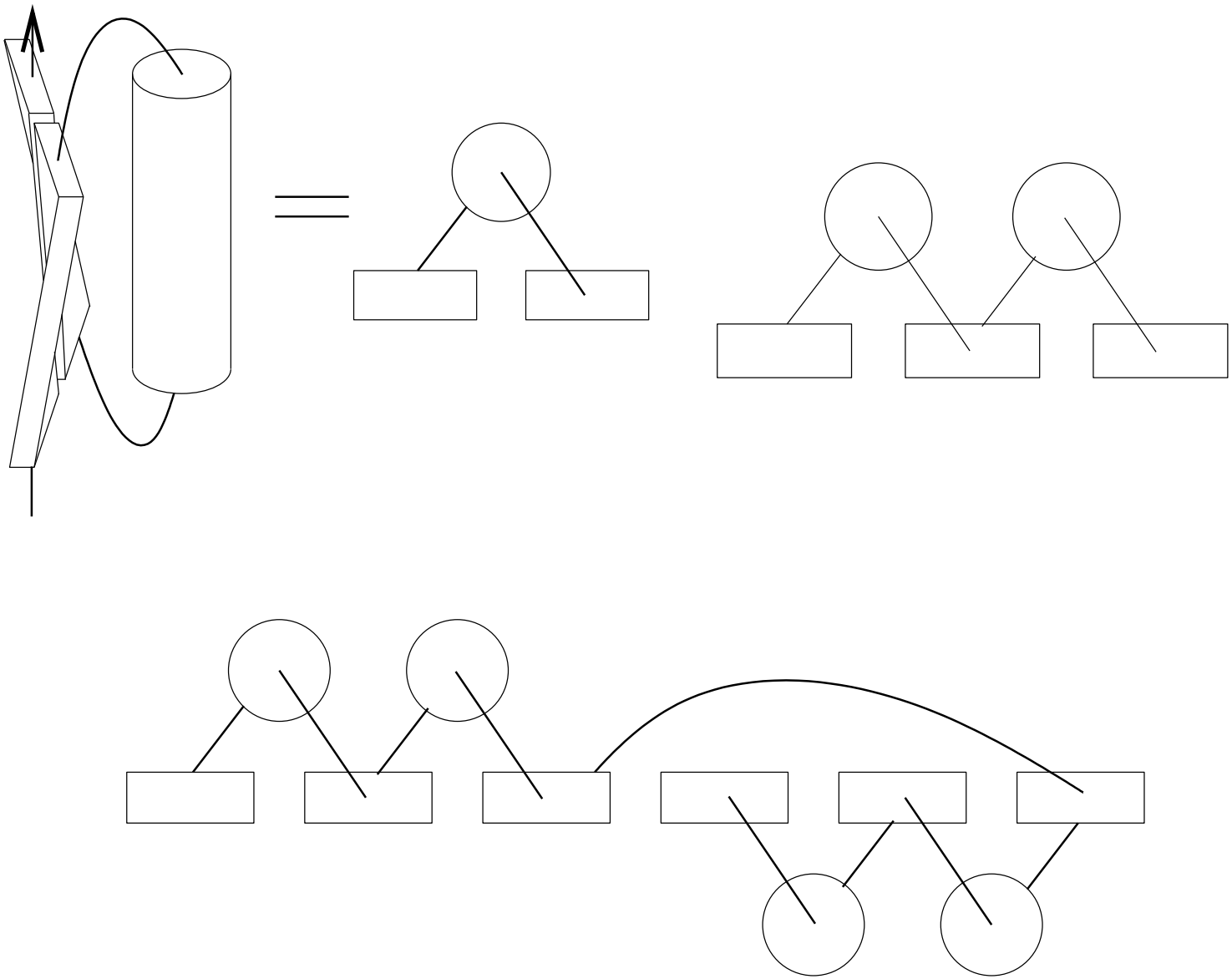




Packing of β-barrels

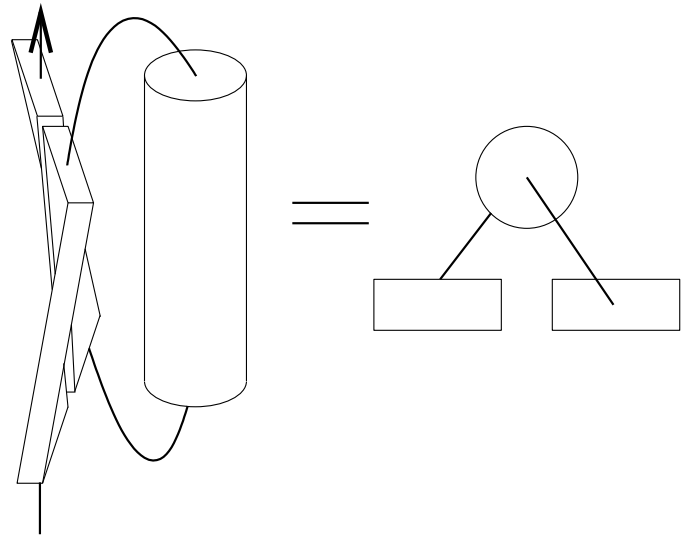


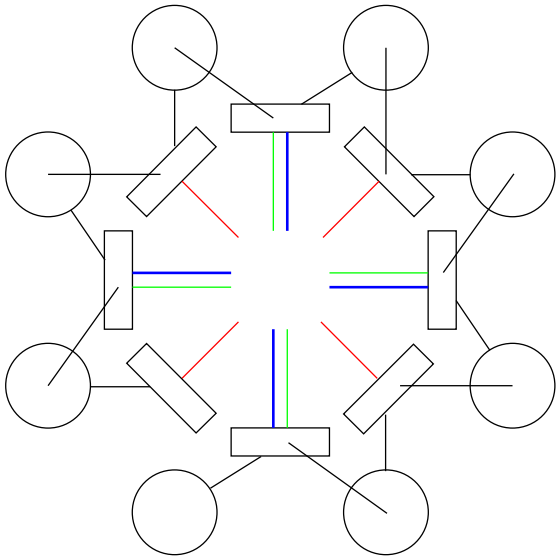
α/β-proteins

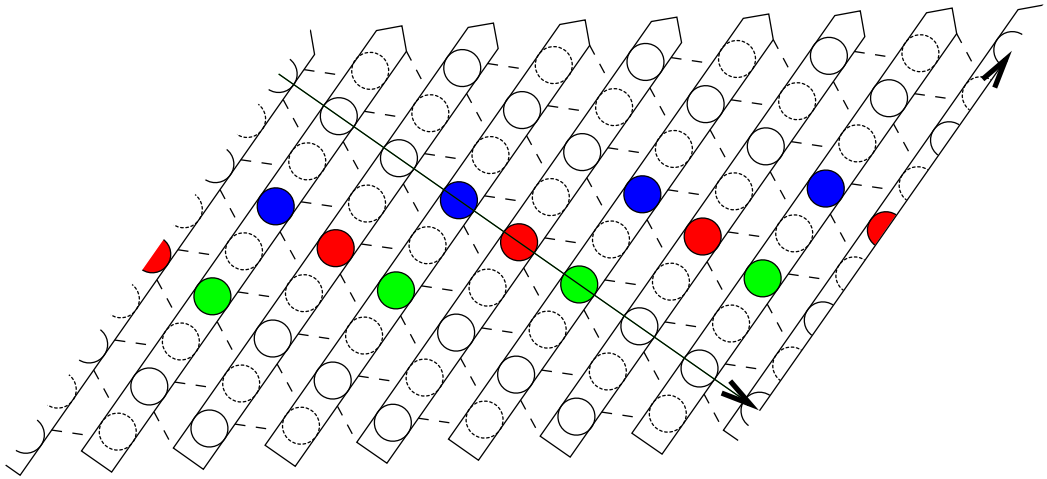


Packing in α/β-proteins (TIM-barrel)

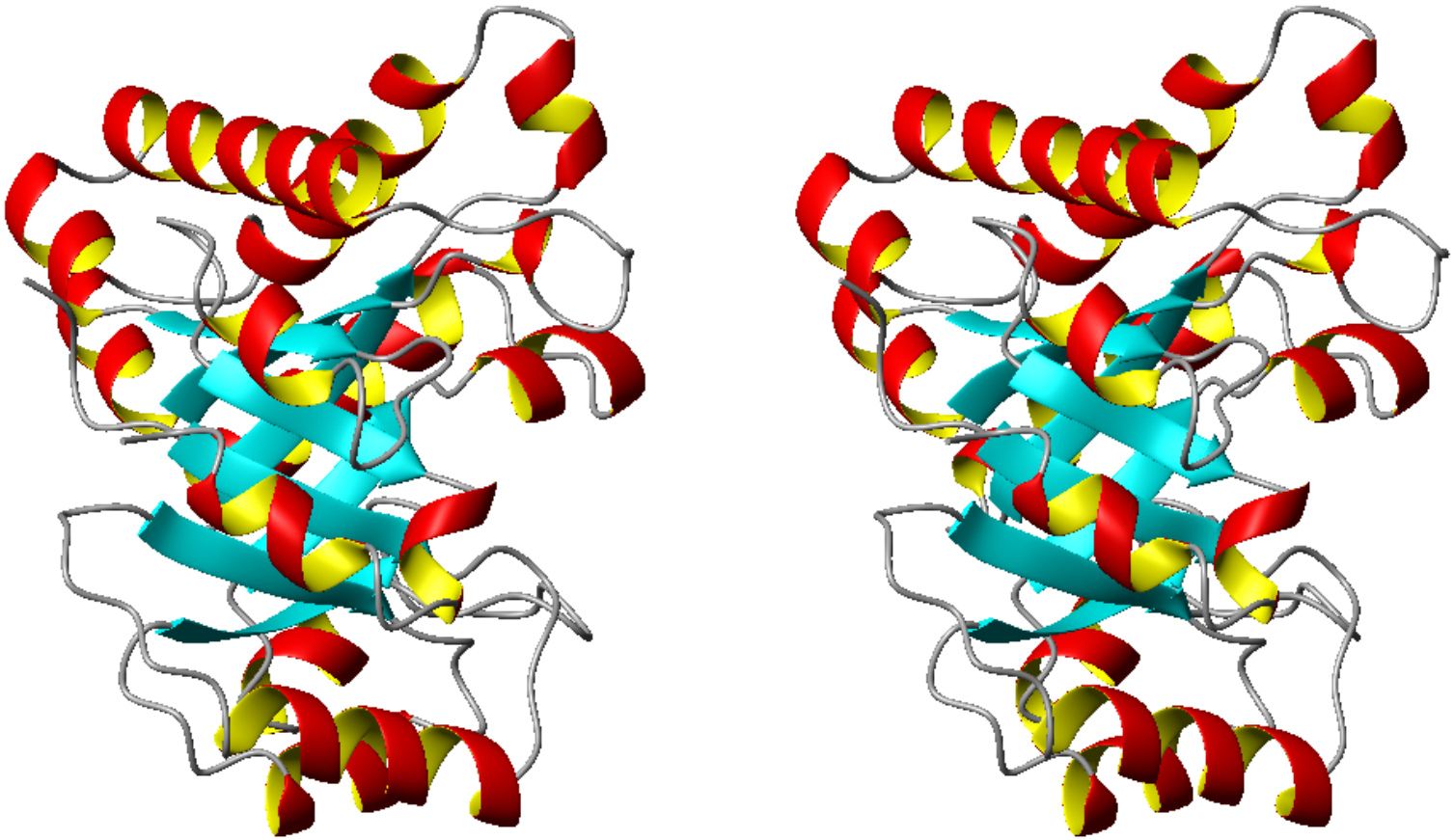
Hydrophobic residues inside







Packing of TIM-barrel



Methods of protein structure determination

COMPUTATIONAL:
De novo structure calculation not reliable

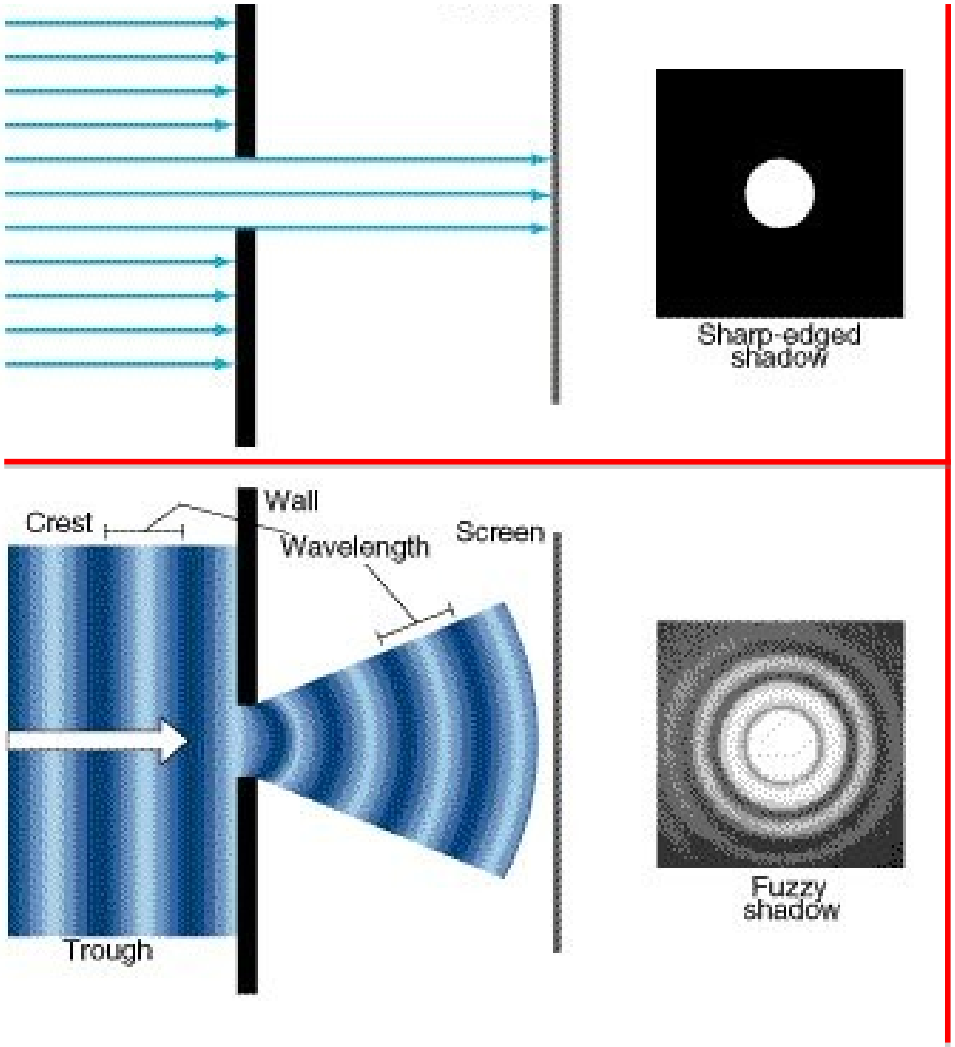
- **Homology modeling**
  using a similar known structure as a starting model

EXPERIMENTAL:
Interactions of electromagnetic (other) waves with molecules

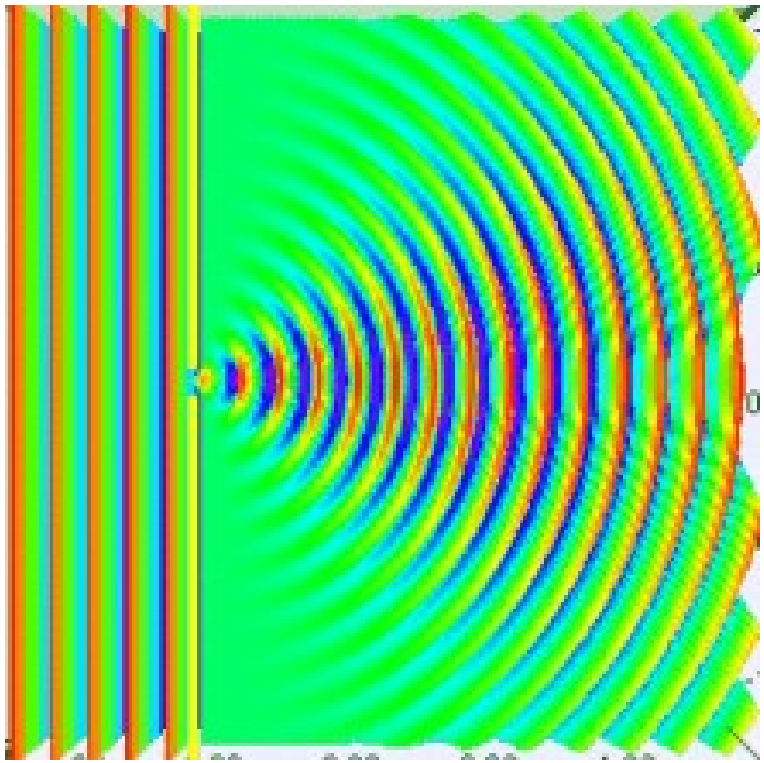
- **Spectroscopy**
  how molecules change characteristics of the wave (intensity, phase, polarization, frequency)
- **Microscopy**
  how molecules change direction of wave in space

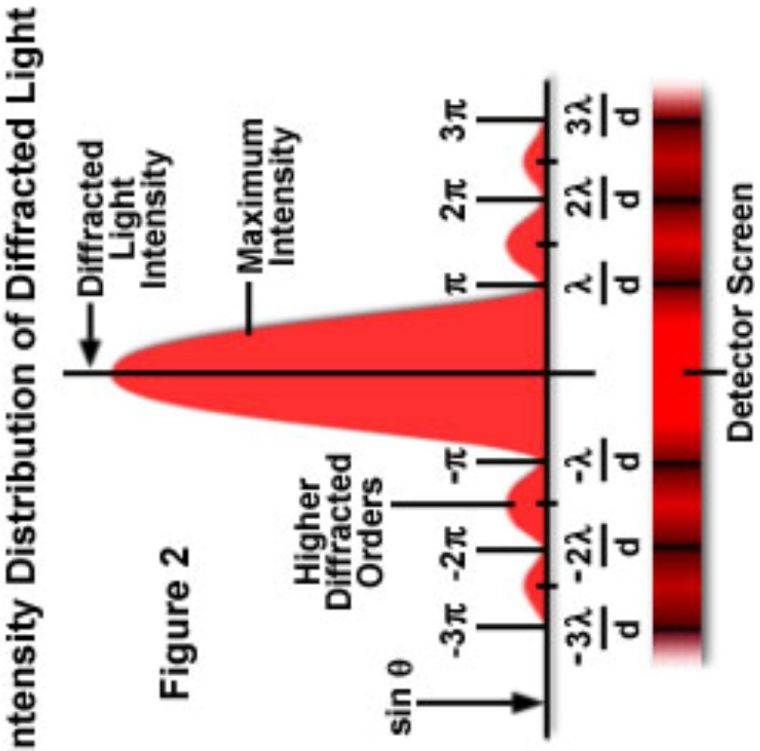
Scattering by molecules

Particles vs waves

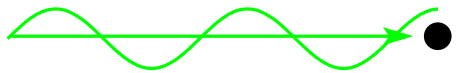


Diffraction of light

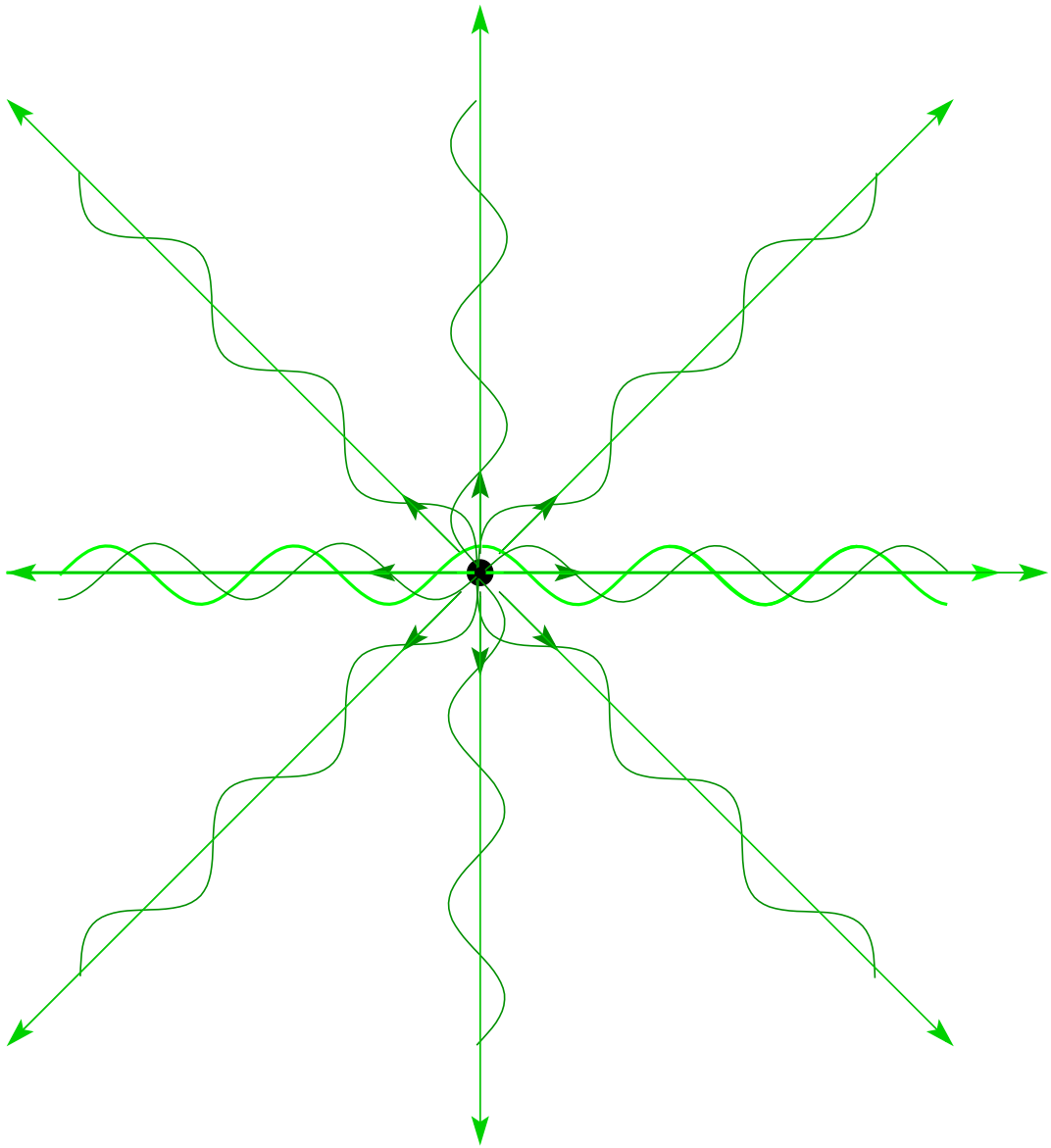




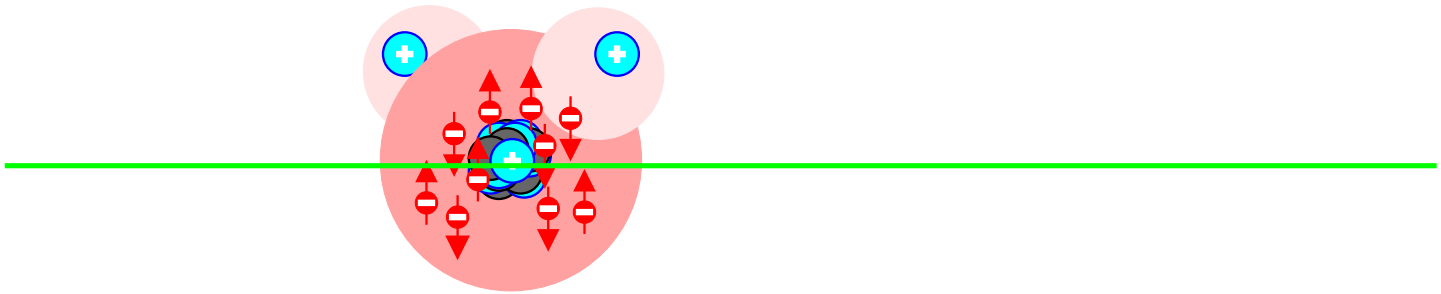
Scattering by molecules



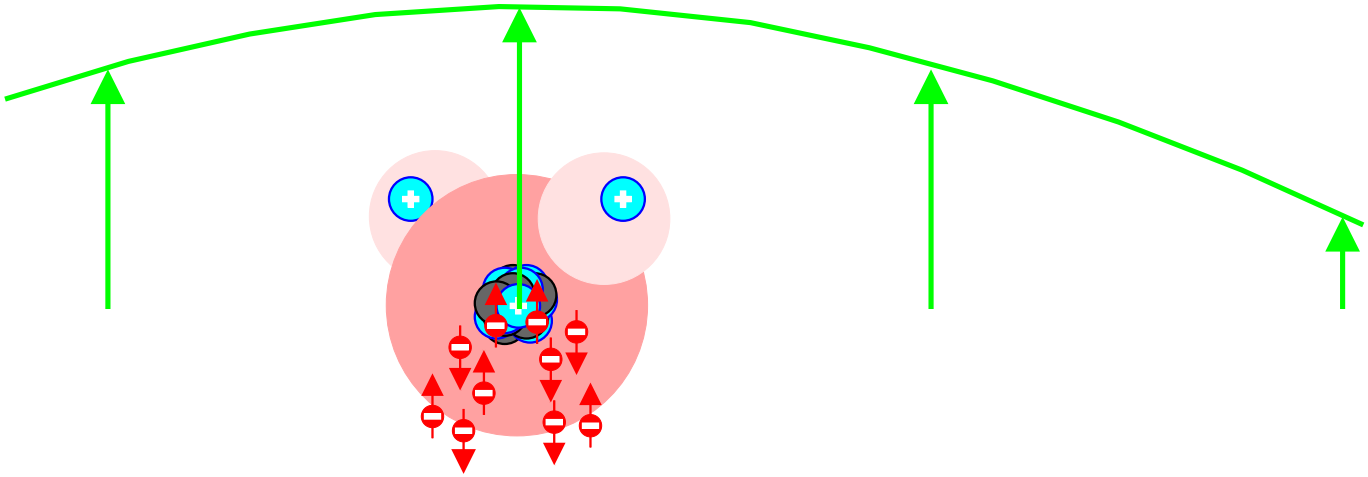
Scattering by molecules



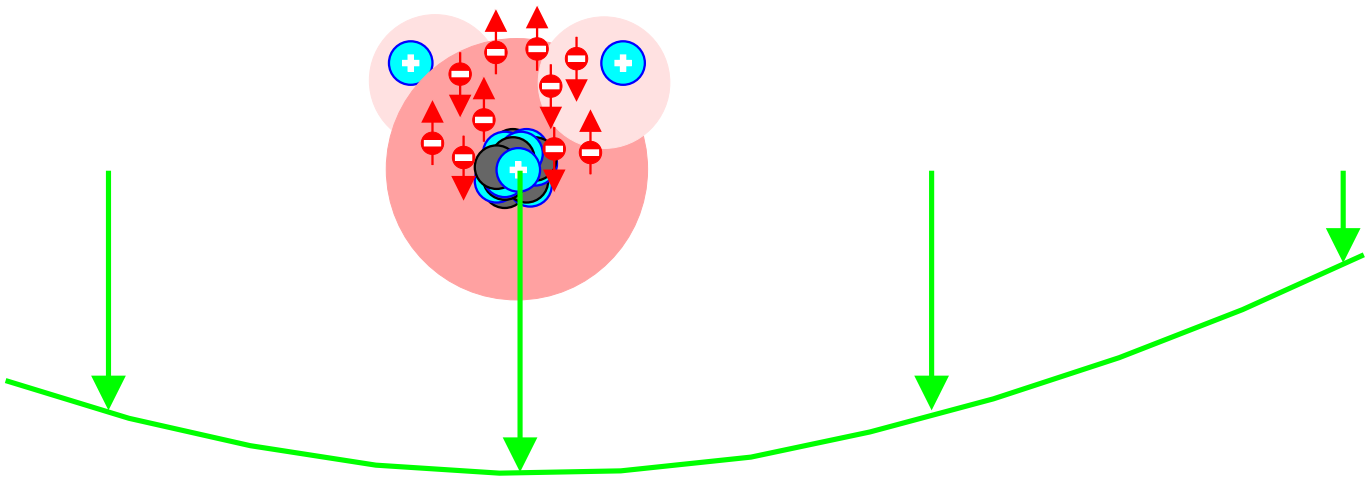
Scattering by molecules



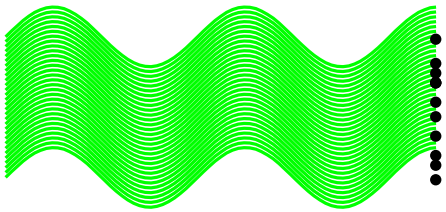
Scattering by molecules



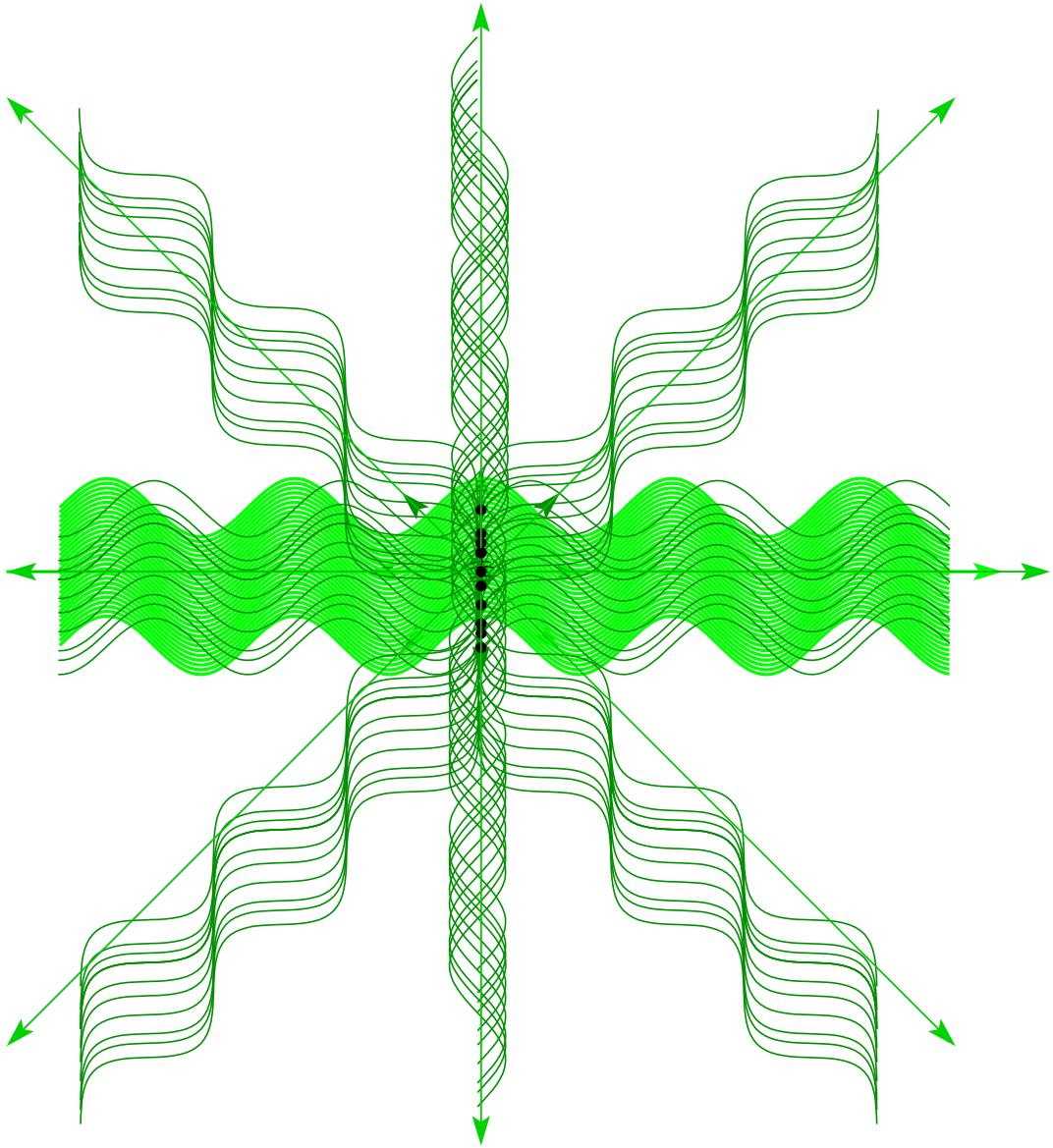
Scattering by molecules



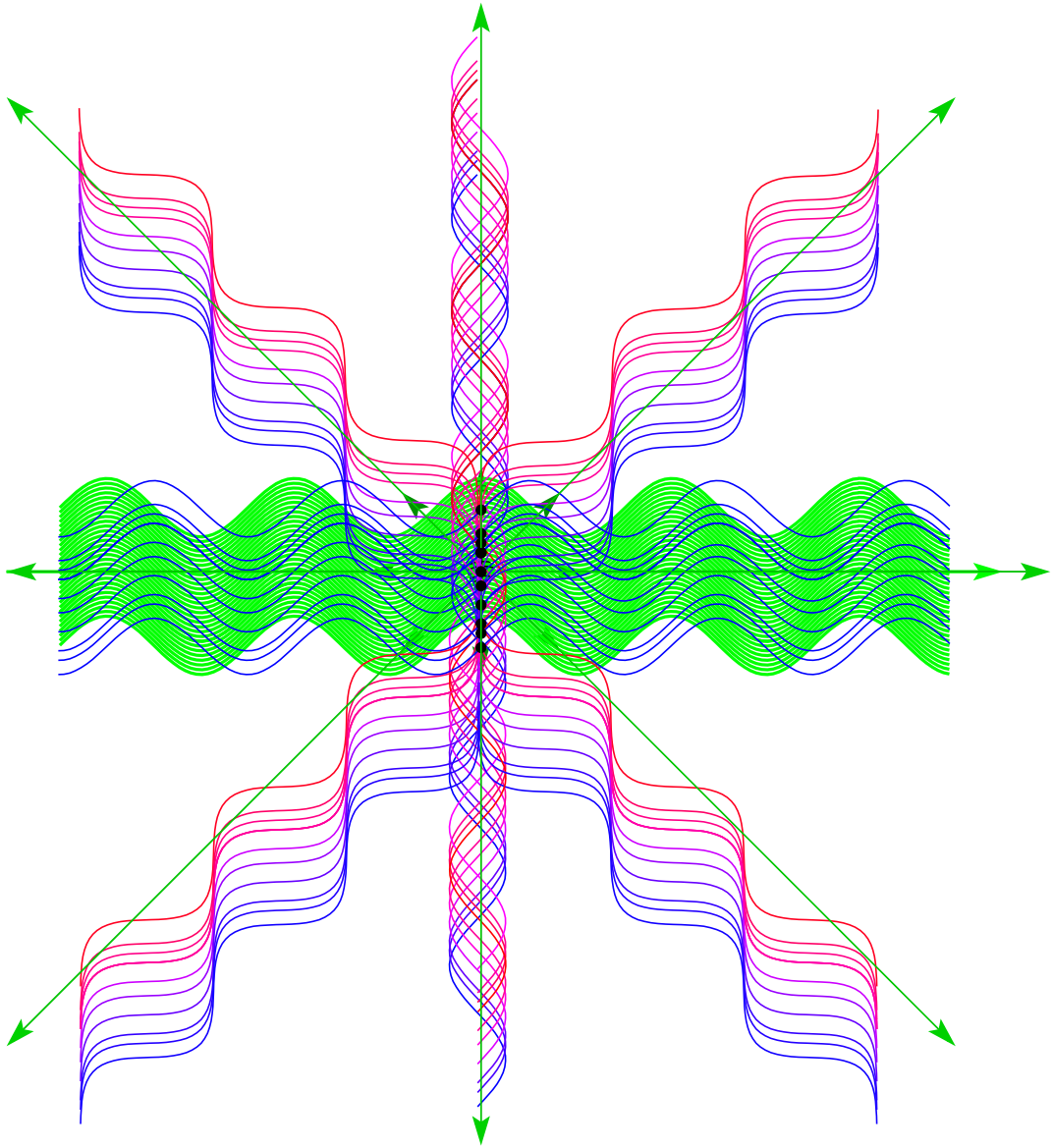
Scattering by molecules



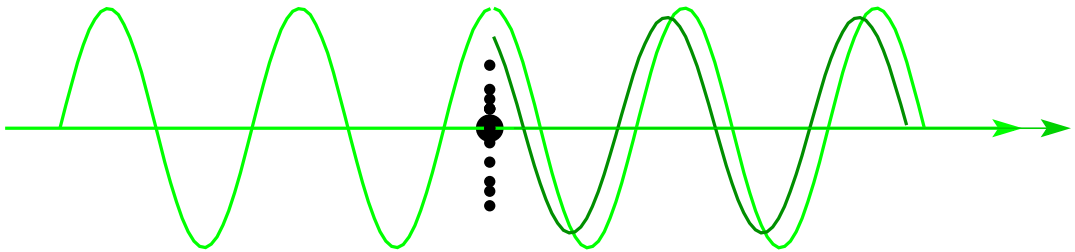
Scattering by molecules



Scattering by molecules



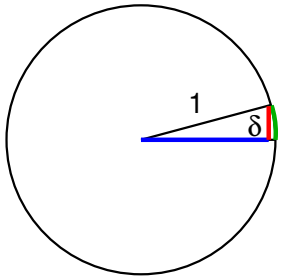
Scattering by molecules



Why 90 ° phase shift?

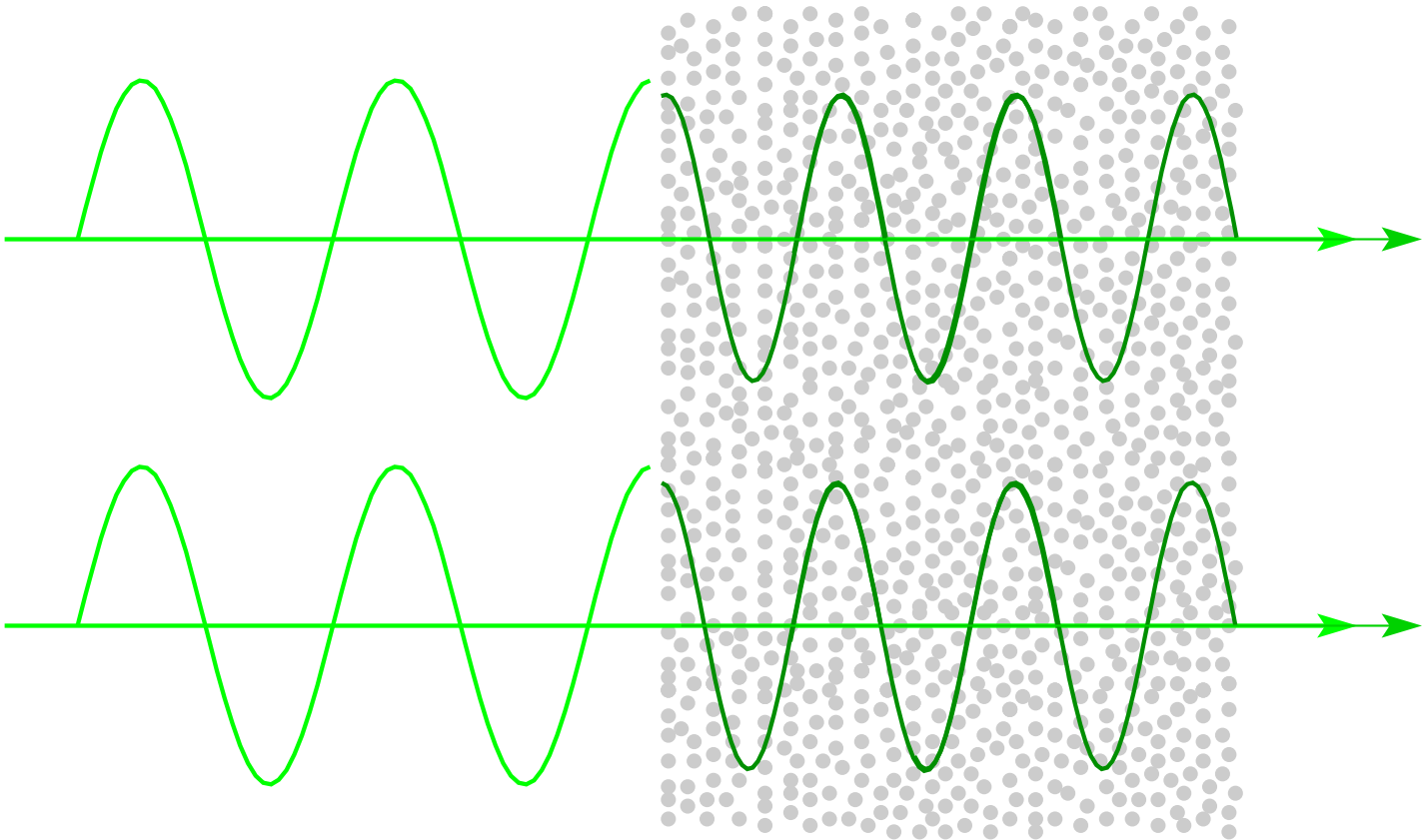
(1) Scattered wave is added to the original wave:
$A\cos\left(\frac{2\pi}{T}\left(t-\frac{z}{c}\right)\right)+Af\cos\left(\frac{2\pi}{T}\left(t-\frac{z}{c}\right)-\phi\right)=A\cos\alpha+Af\cos(\alpha-\phi)$
Scattered amplitude is a small fraction (f) of the original one

(2) We know that wave interacting weakly with molecules is apparently slightly retarded ($c'=c/n$, $n=$ refractive index)
Original wave: $A\cos\left(\frac{2\pi}{T}\left(t-\frac{z}{c}\right)\right)=A\cos\alpha$
Retarded wave: $A'\cos\left(\frac{2\pi}{T}\left(t-\frac{z}{c'}\right)\right)=A'\cos\left(\frac{2\pi}{T}\left(t-\frac{z}{c}-\frac{n-1}{c}z\right)\right)=$
$A'\cos(\alpha-\delta)=A'\cos\delta\cos\alpha+A'\sin\delta\sin\alpha\approx A'\cos\alpha+A'\delta\sin\alpha$
because n is only slightly $>1\Rightarrow n-1$ small $\Rightarrow\delta$ is small (we use arc lenght to measure angle and for a small angle δ, $\cos\delta\approx 1$, $\sin\delta\approx\delta$)

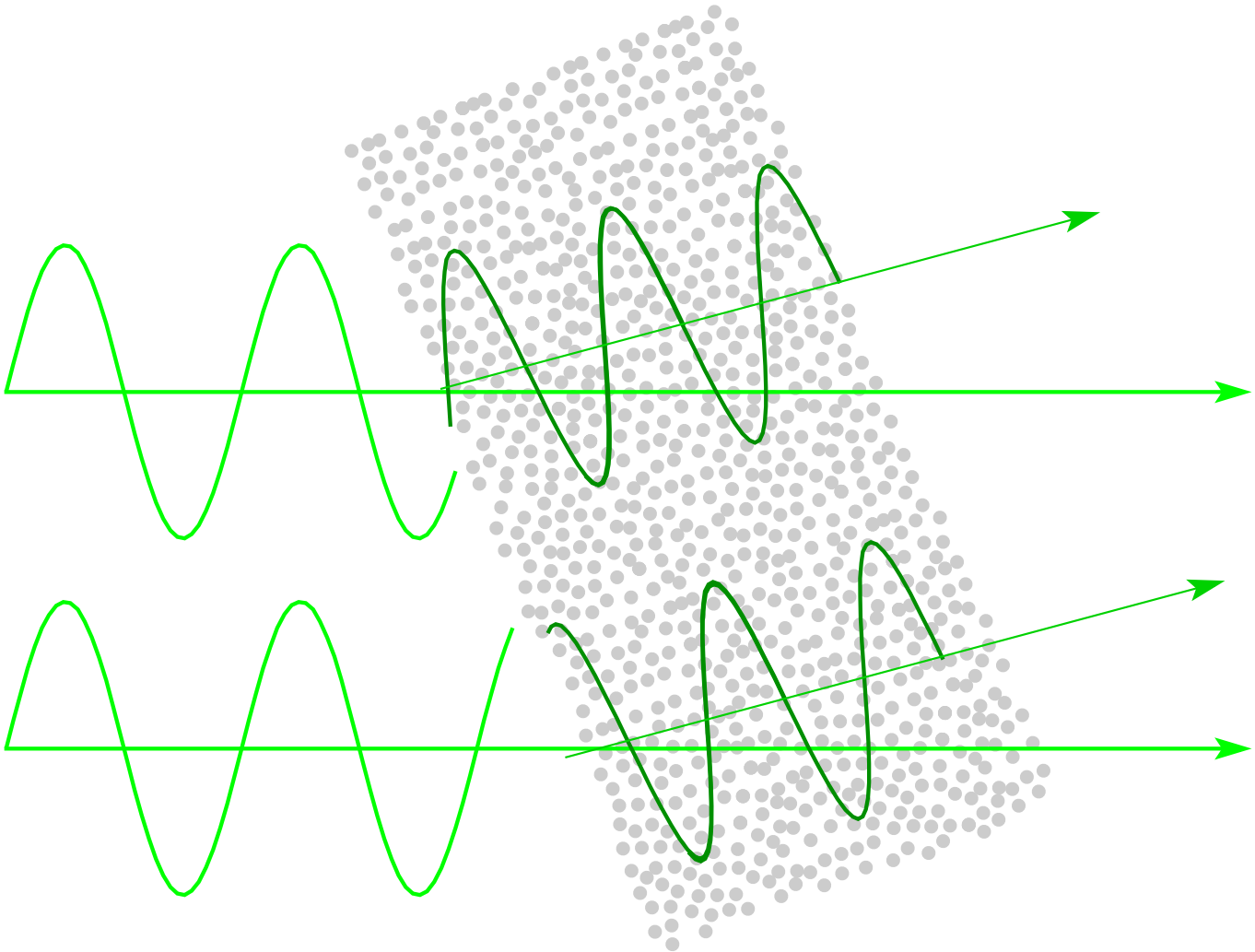


(3) $A\cos\alpha+Af\cos(\alpha-\phi)=A'\cos\alpha+A'\delta\sin\alpha$
if $A=A'$, $f=\delta$, $\phi=\pi/2$ (because $\sin\alpha=\cos(\alpha-\pi/2)$)

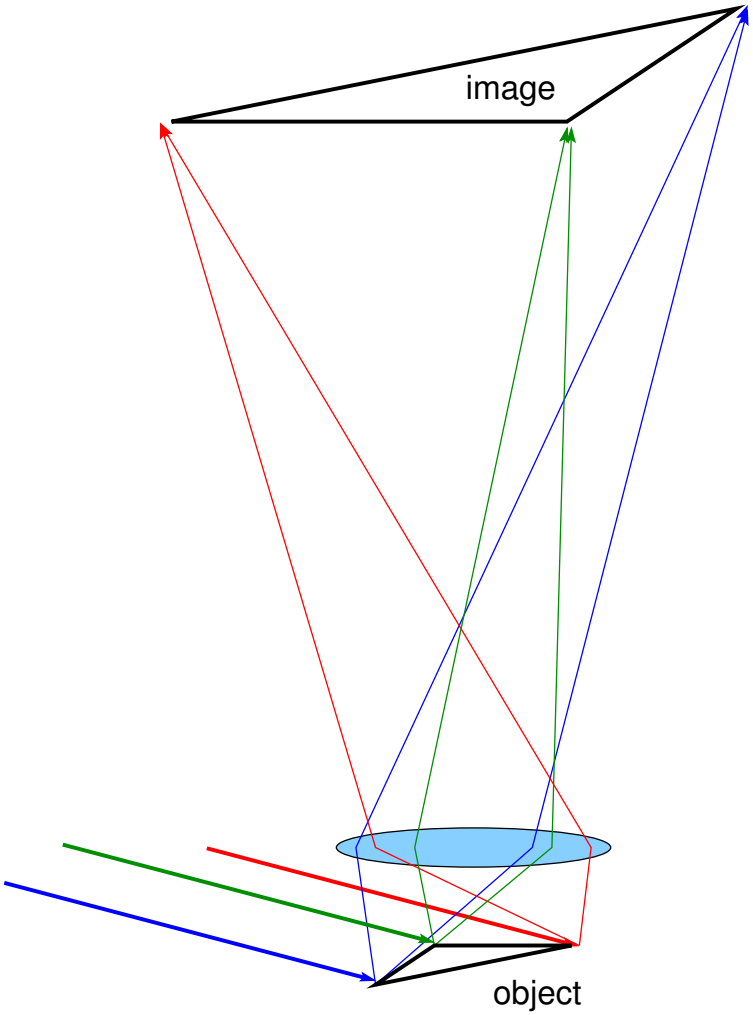
Refractive index

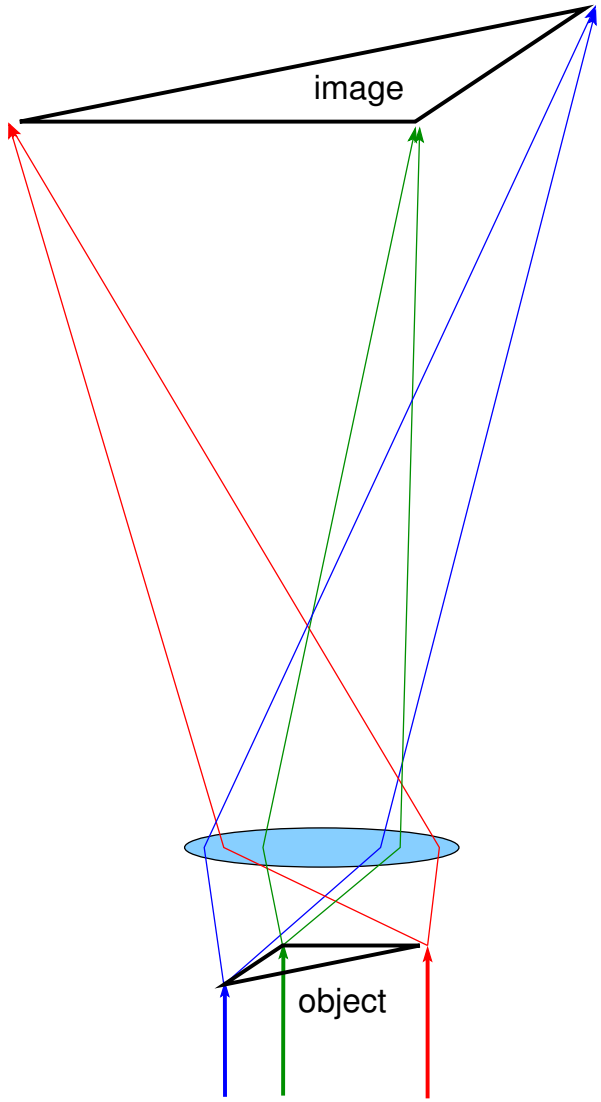


Refractive index

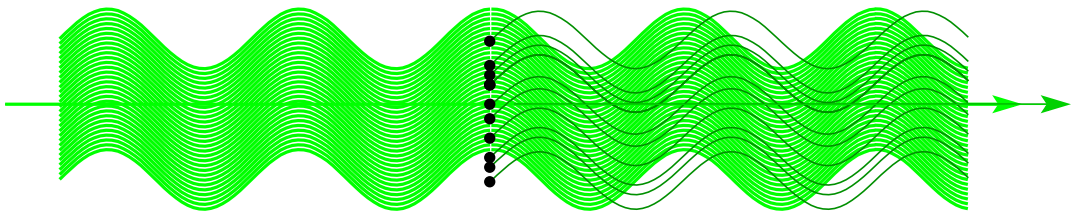


Microscopy

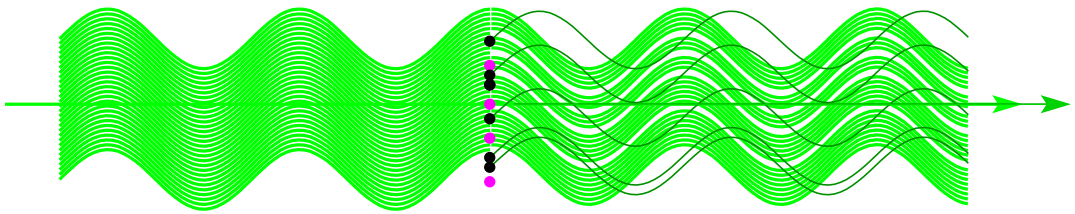




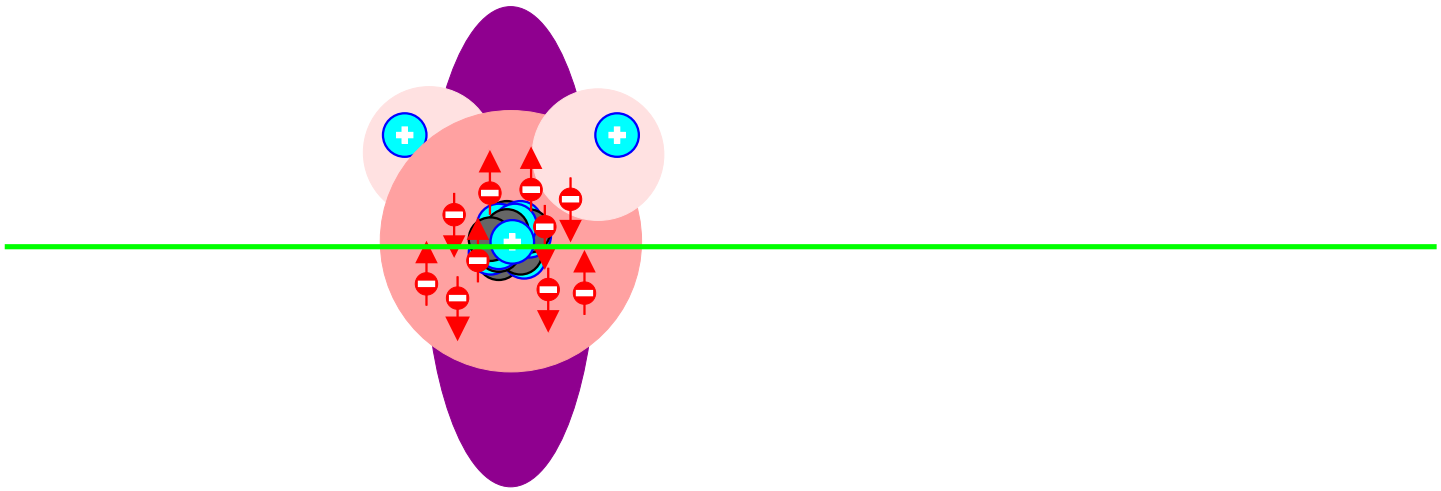
Absorption by molecules



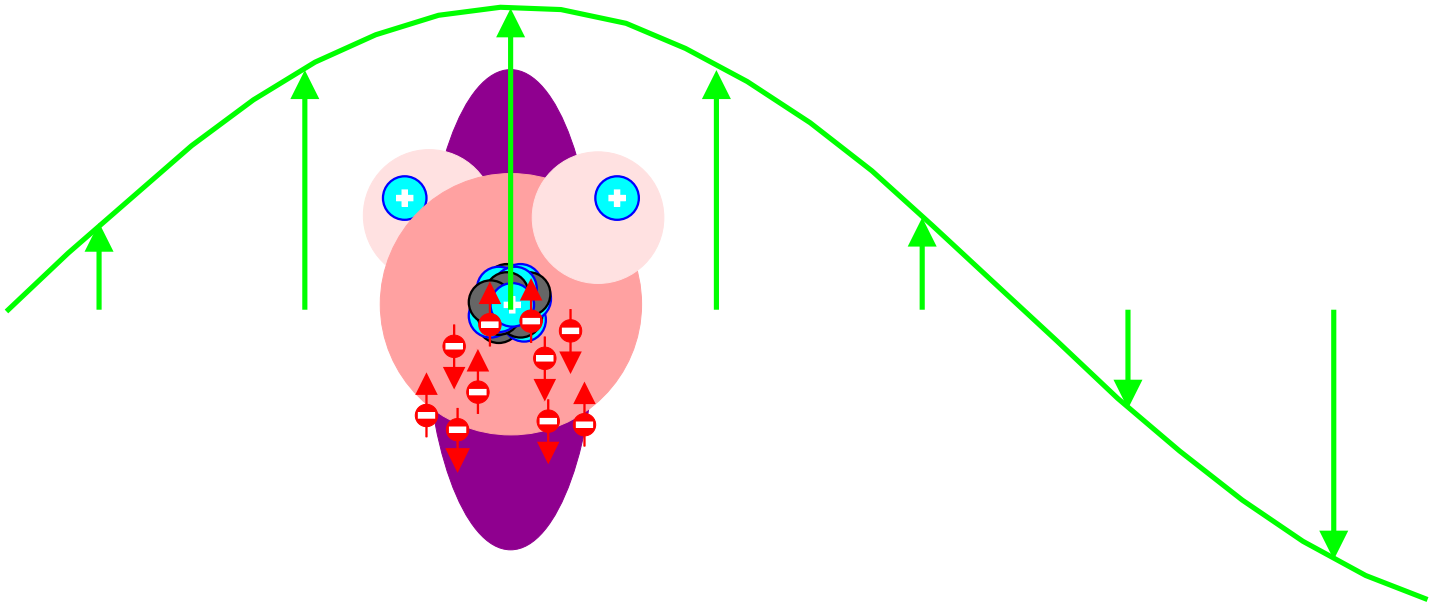
Absorption by molecules



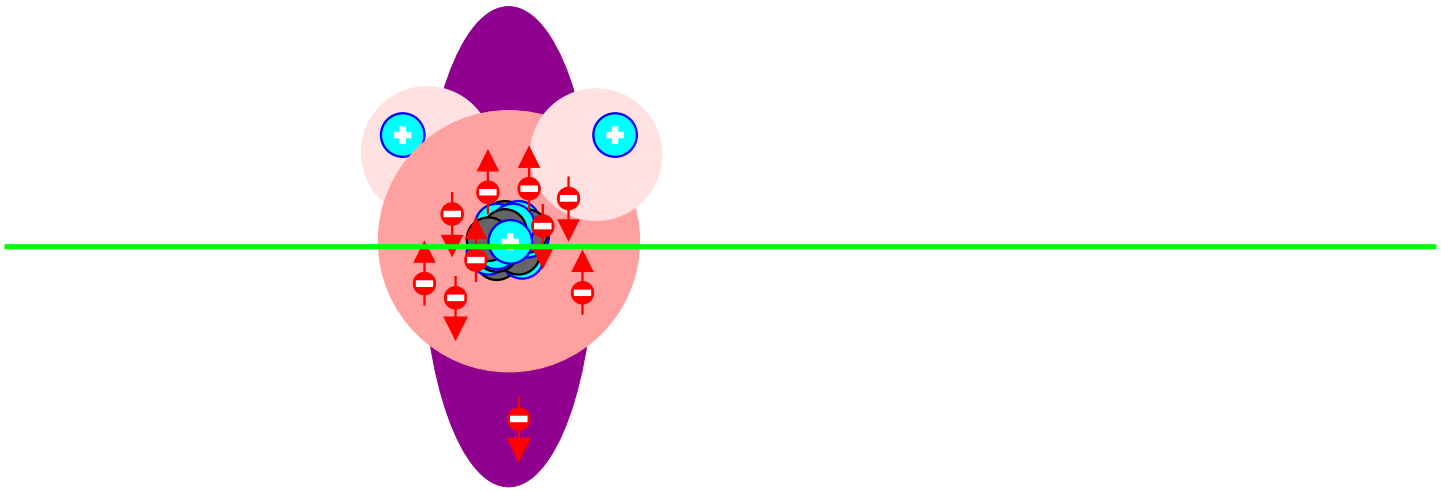
Absorption by molecules



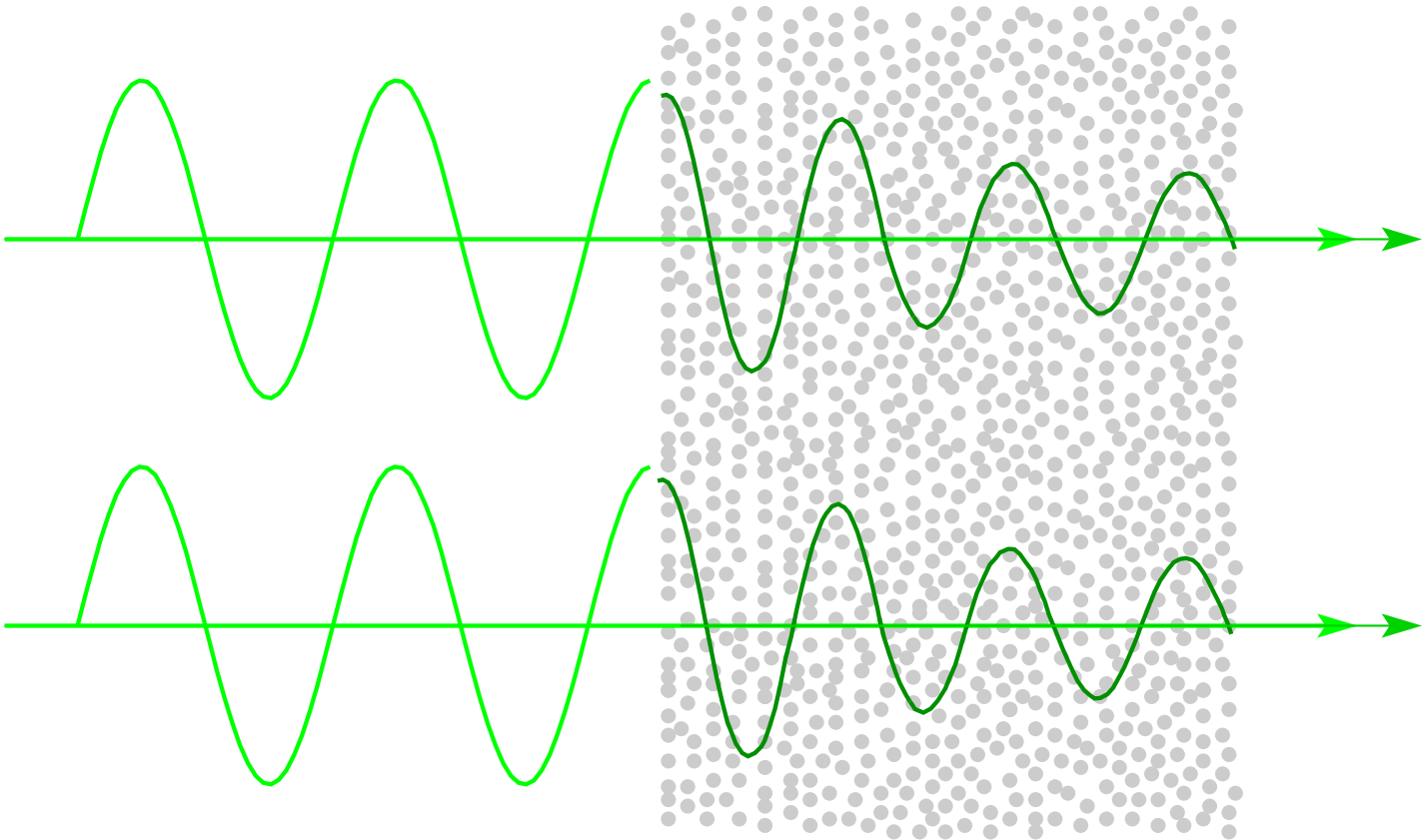
Absorption by molecules



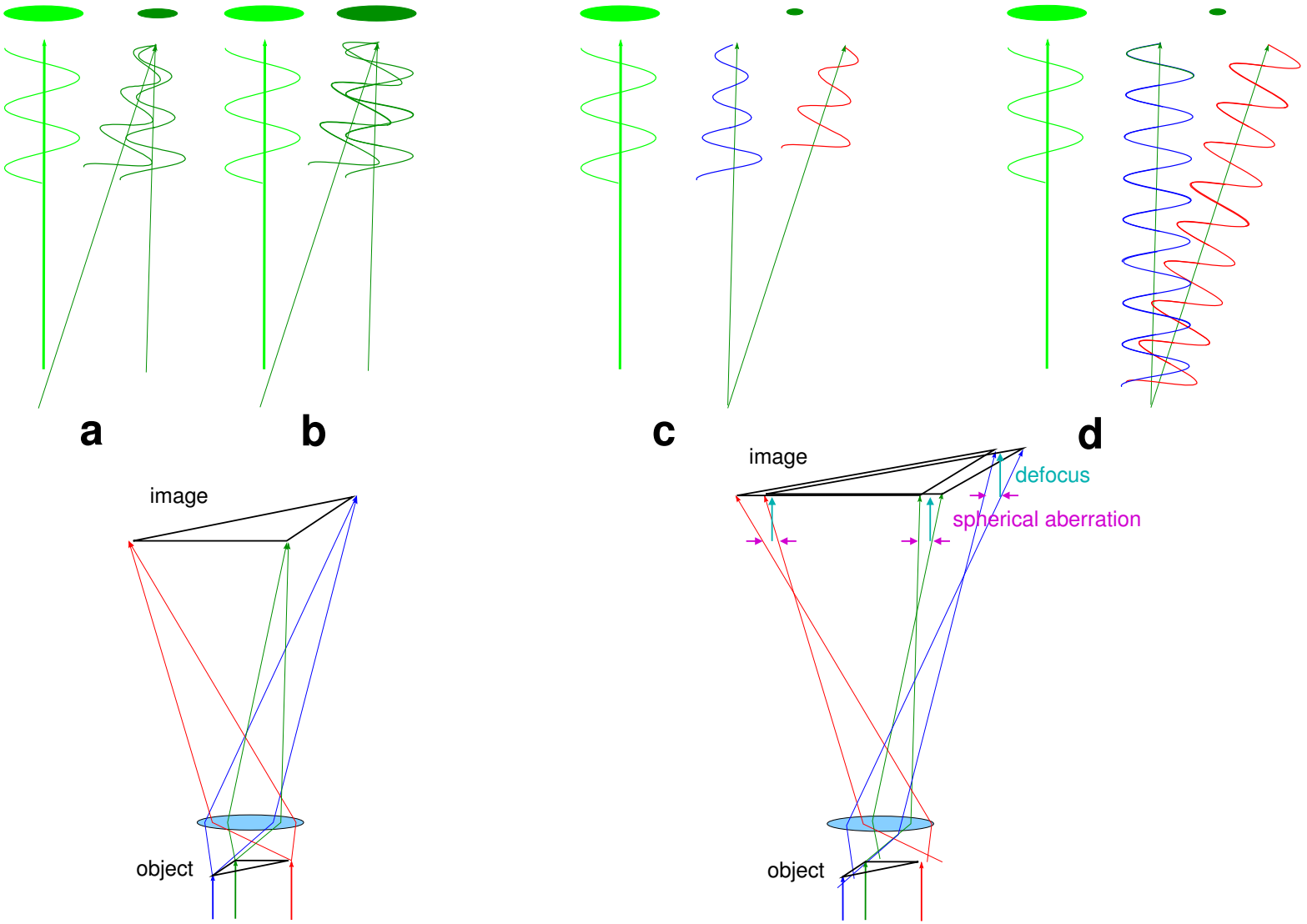
Absorption by molecules



Absorption by molecules



Phase vs amplitude objects



Phase vs amplitude objects

- **a** Perfect lens in focus provide contrast in intensity of scattered vs. original wave if the scattered wave is absorbed by the sample
- **b** Perfect lens in focus provide no contrast in intensity of scattered vs. original wave if the scattered wave is not absorbed by the sample, but only phase-shifted by 90 °
- **c** Imperfect and defocused lens provide contrast in intensity of scattered vs. original wave if the scattered wave is absorbed by the sample
- **d** Defocus and spherical aberration of lens provide contrast in intensity of scattered vs. original wave if the scattered wave is not absorbed by the sample, but only phase-shifted by 90 °. Defocus and spherical aberration introduce another phase shift $\Rightarrow$ scattered have oposite phase and cancel each other.

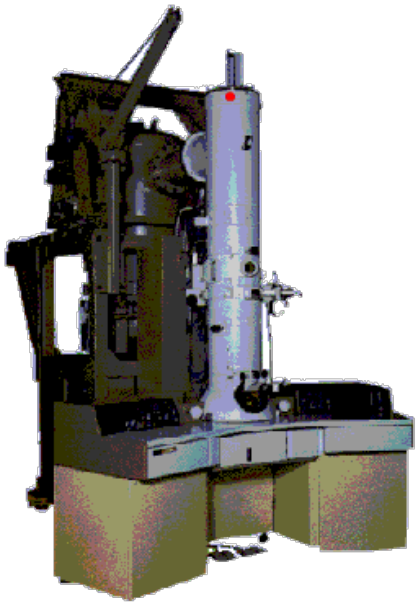
Limitation of optical microscopy

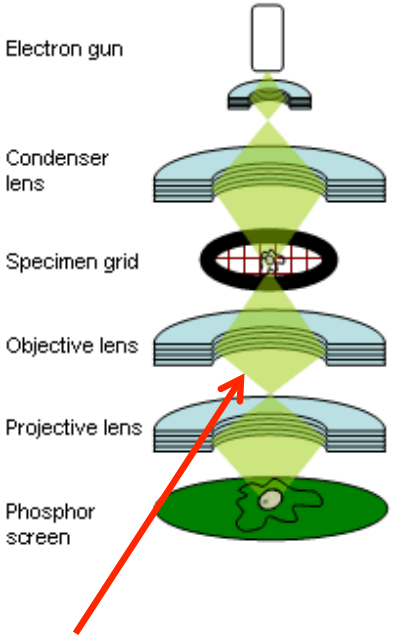
Photons scatter as waves $\Rightarrow$ limited resolution
resolution $< 0.61\lambda/n$ ($\lambda =$ wavelegth, $n =$ refractive index
$\Rightarrow \lambda \approx 0.1\,nm$ (distances of atoms in molecules), X-rays
Lens are not available for X-rays ($\lambda \approx 0.1$ nm)
no material has sufficient refractive index

SOLUTIONS:

- **Microscopy with electrically charged waves**
  electrons scatter as waves
  electron beams are bent in electromagnetic field
  $\Rightarrow$ **Electron microscopy**
- **Analysis of diffraction patterns**
  intensity enhanced if molecules are aligned in crystals
  $\Rightarrow$ **X-ray crystallography**
  applicable also to electron and neutron waves

Transmission Electron Microscope





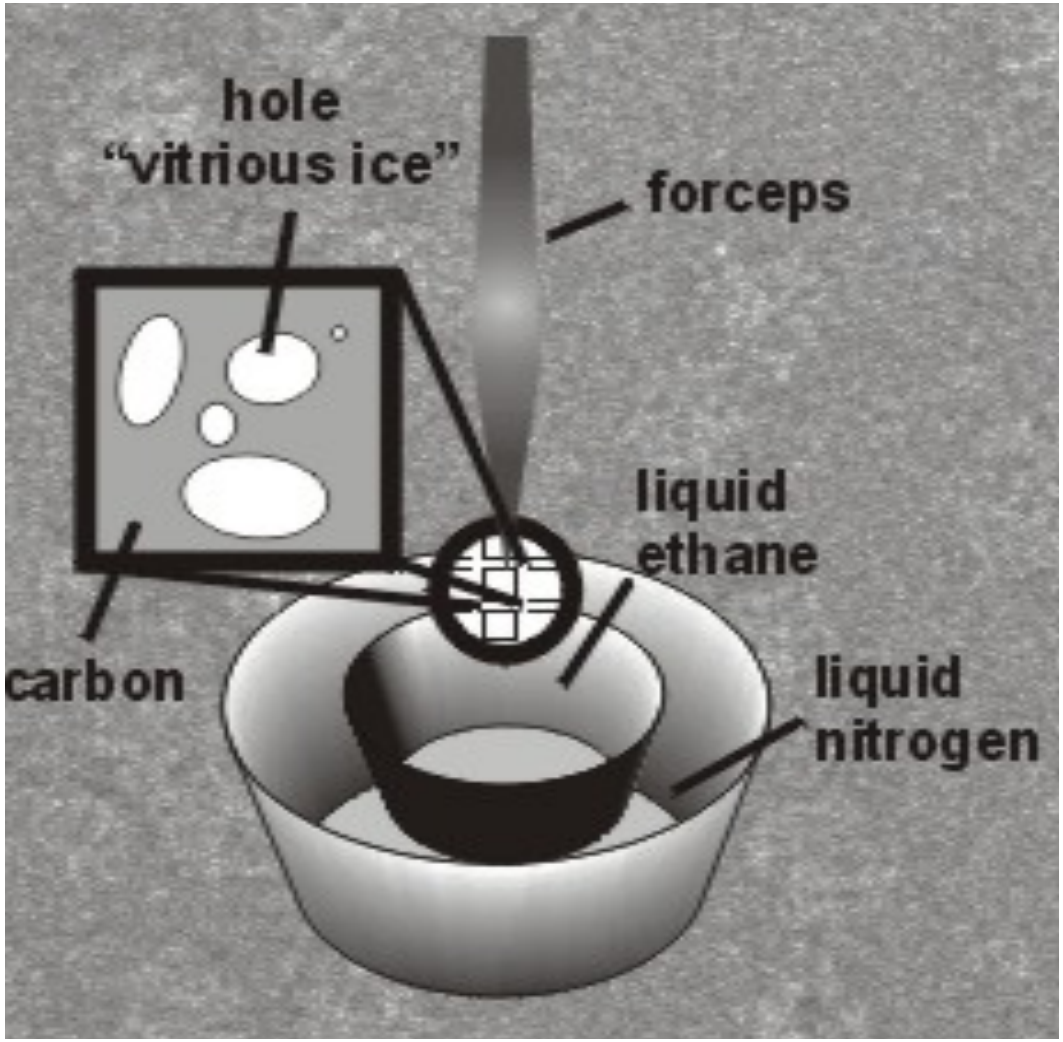
Electron source:
Thermal emission from heated cathode

Focussing:
Electro-magnetic Lenses

Detection:
Phosphor screen or CCD camera (former times: negative)

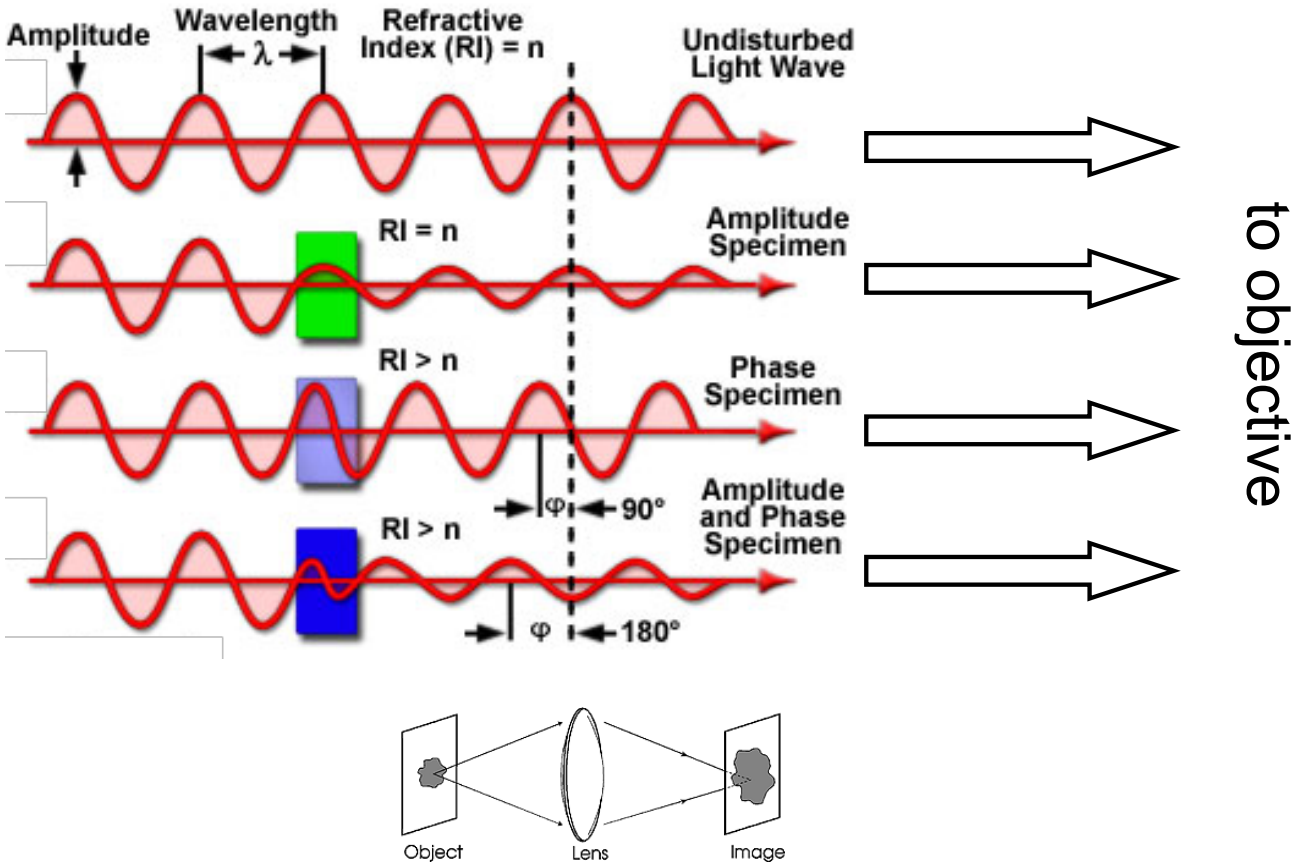
Cryo-electron microscopy

Proteins in vitrous ice, can reach atomic resolution



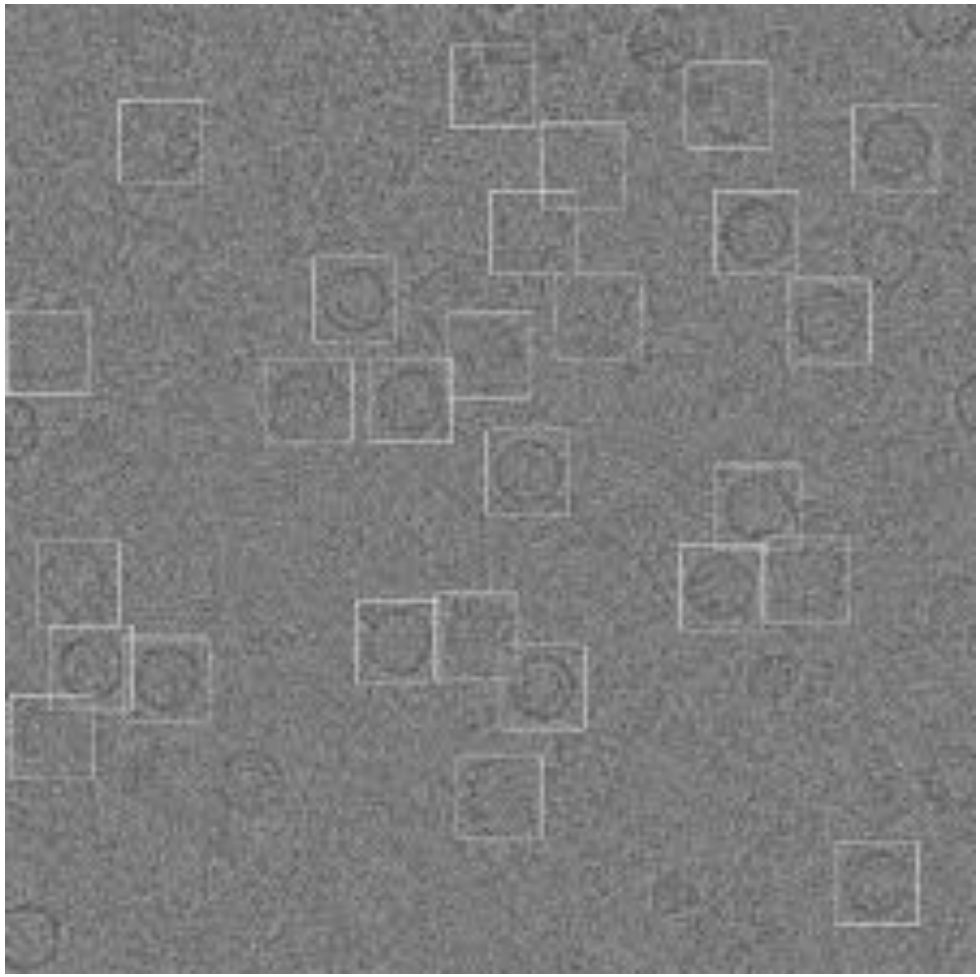
Amplitude vs. phase objects

Macromolecules in water / vitreous ice are phase objects



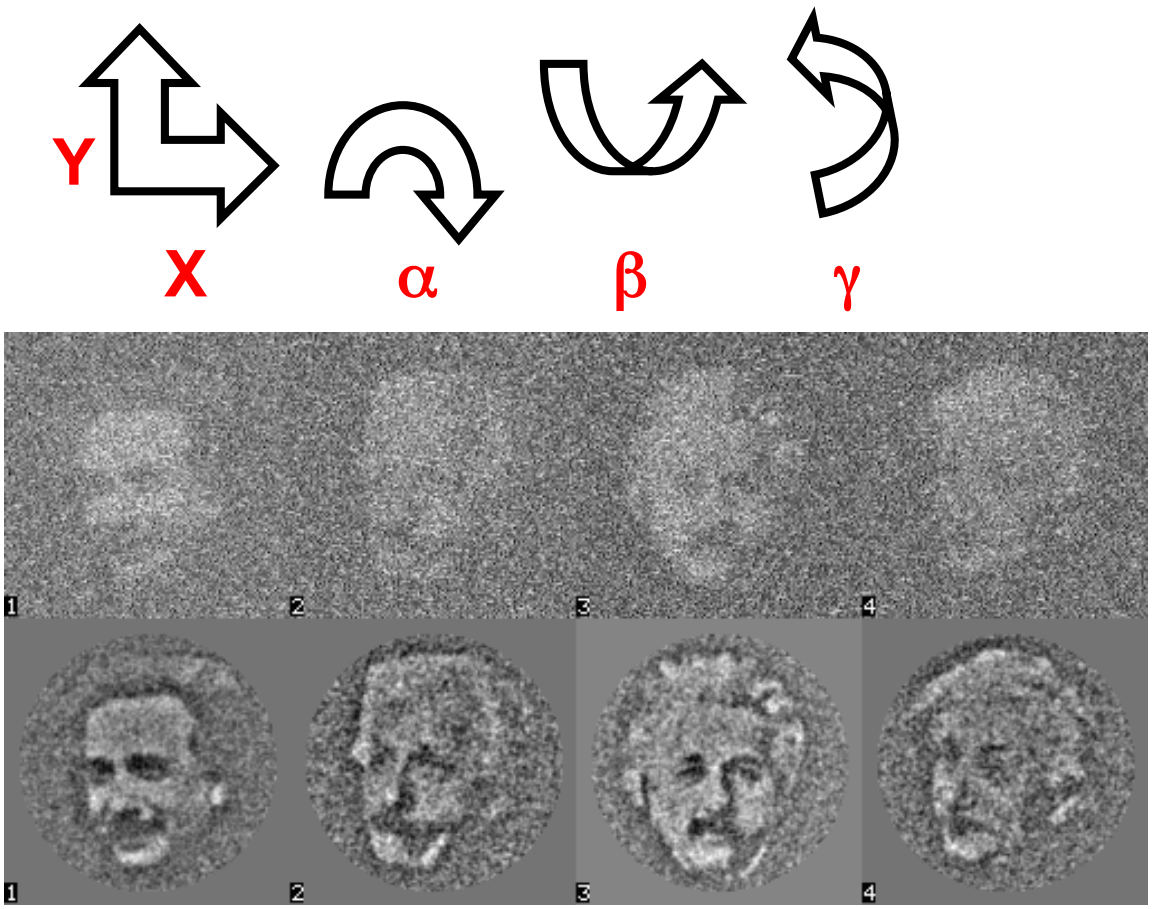
Samples damaged by electrons

Signal to noise ratio

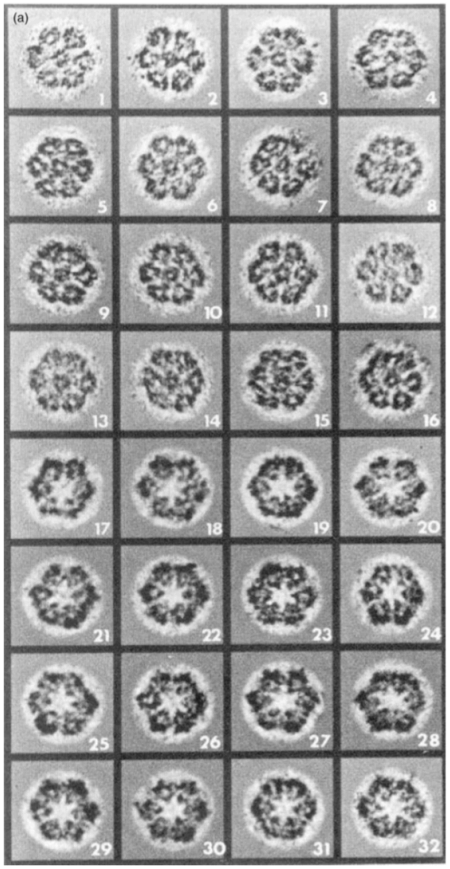


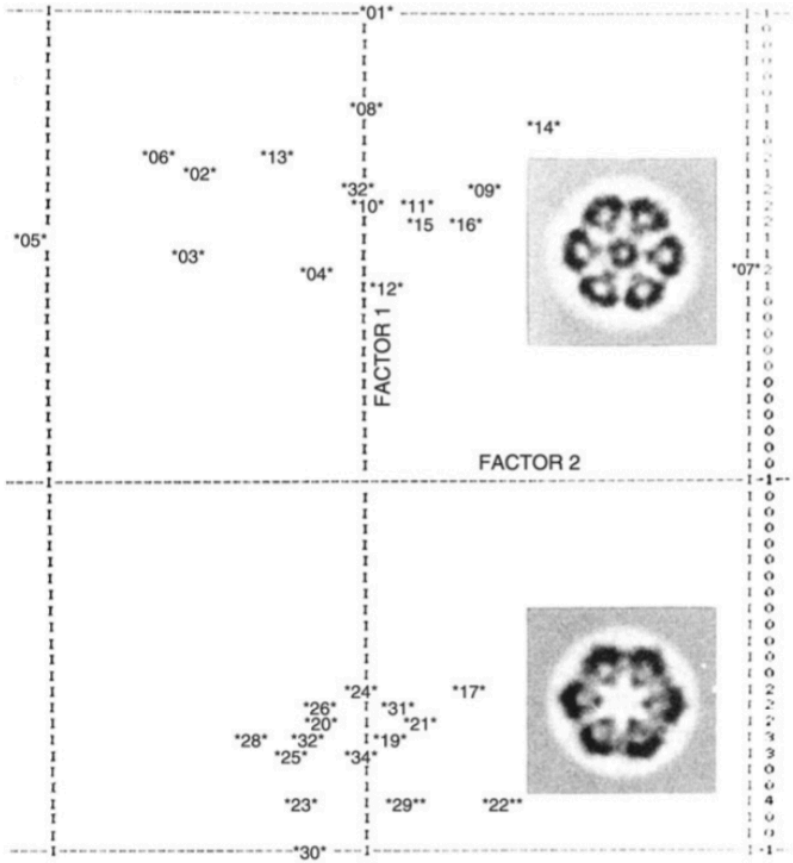
Alignment and classification

Images contain diffrent views of possibly different molecules



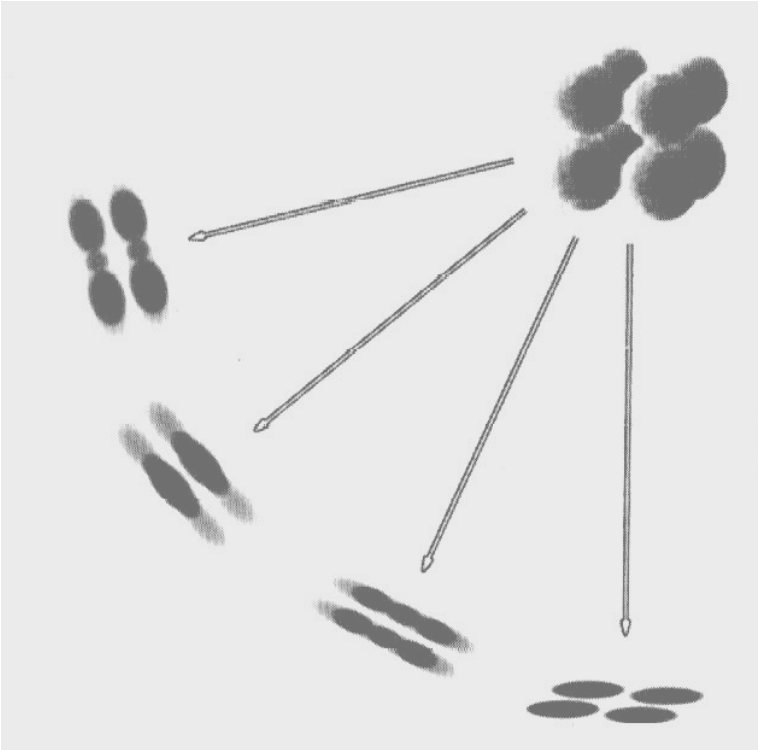
Classification and averaging (principal component analysis)

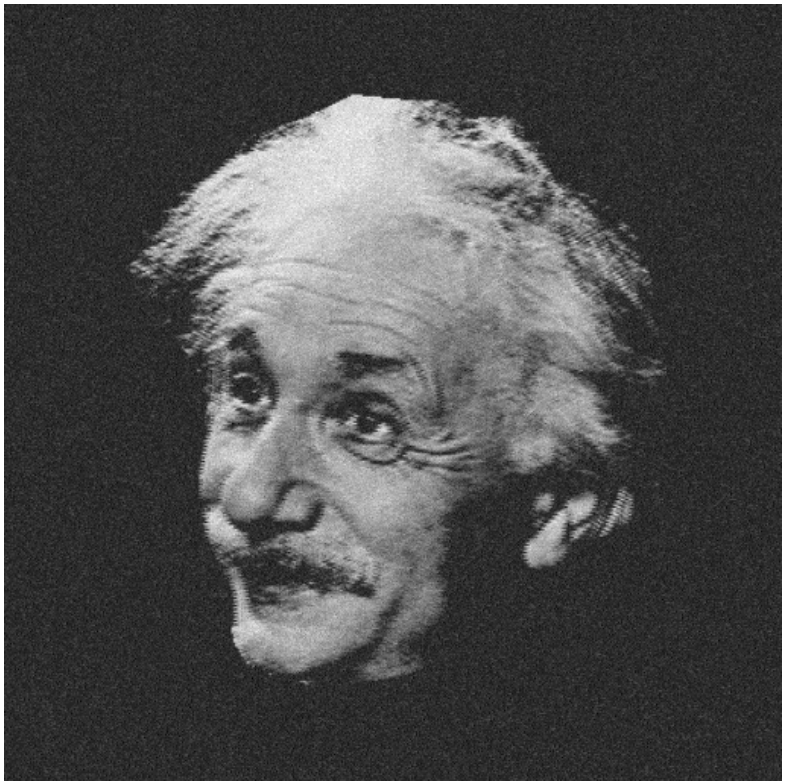




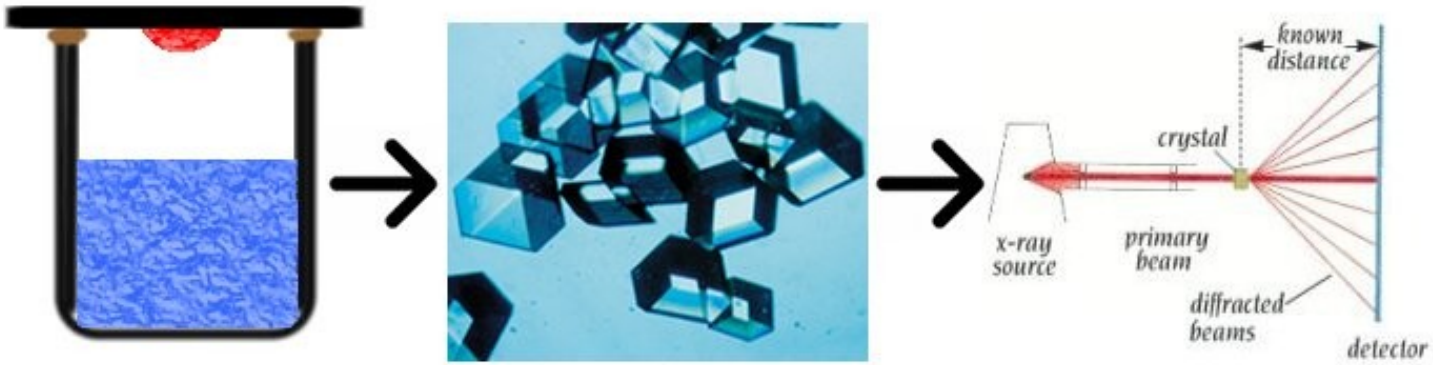
3D reconstruction

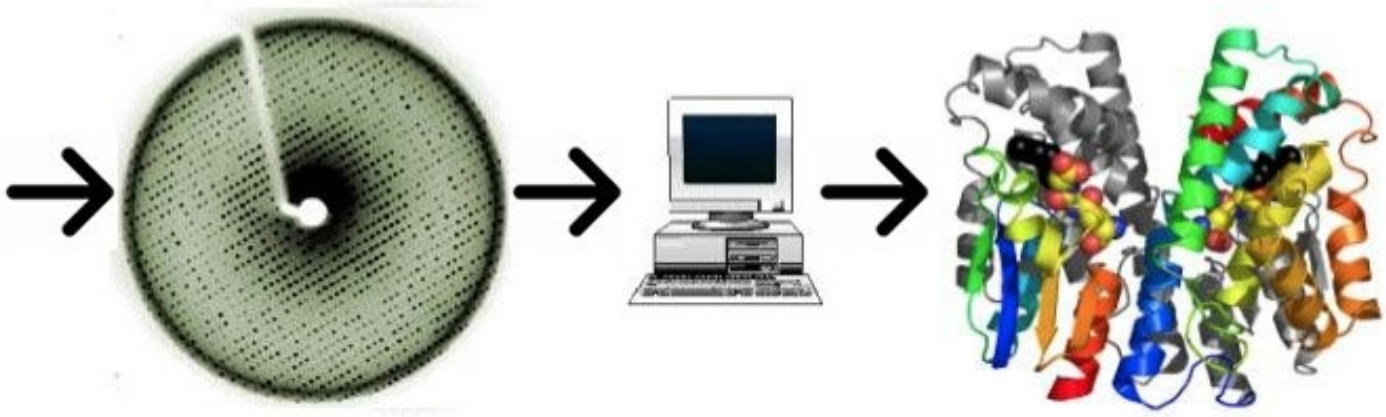
Iterative process:
2D projections are calculated from a 3D model
alignment and classification are improved iteratively



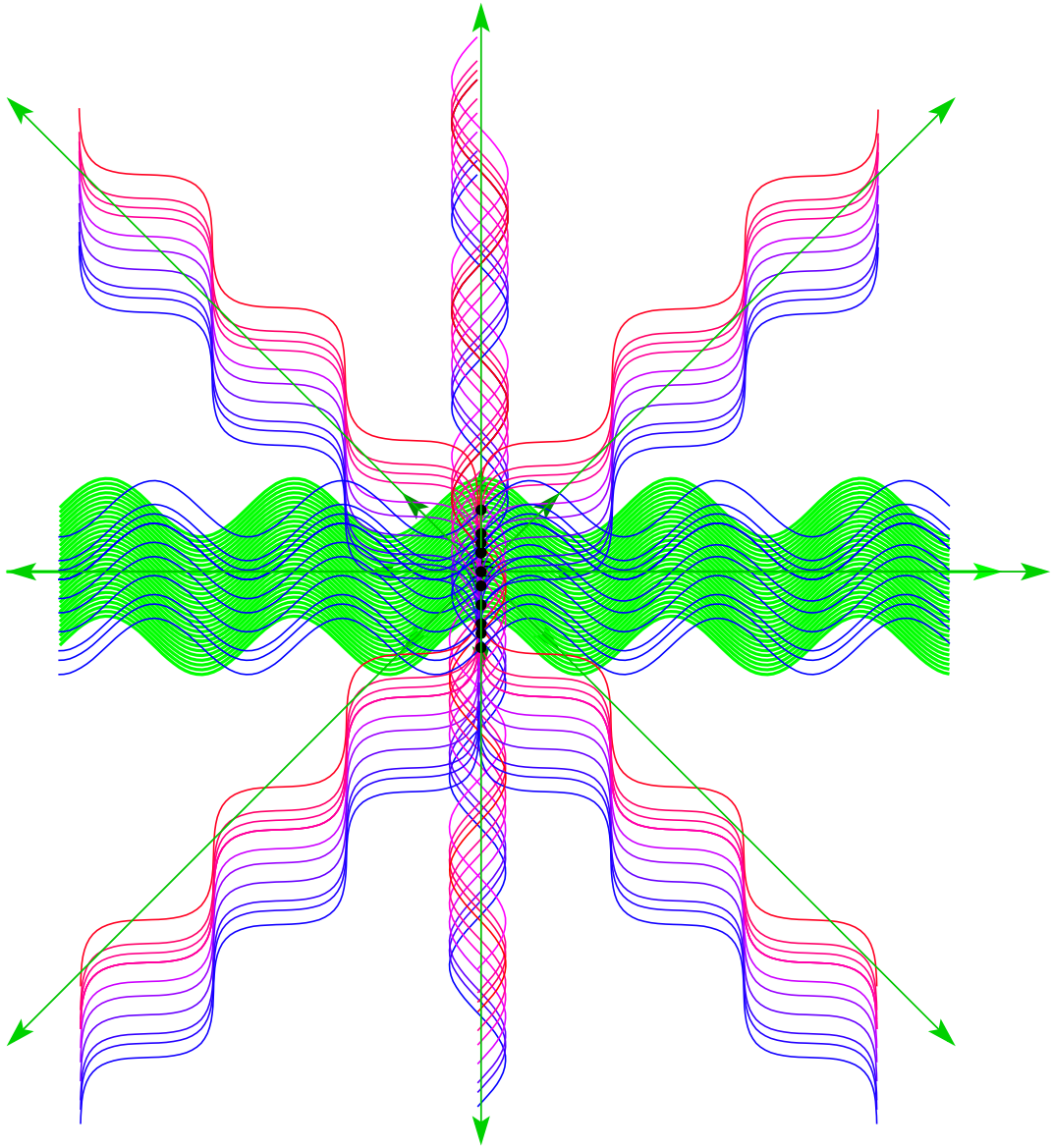


X-ray crystallography

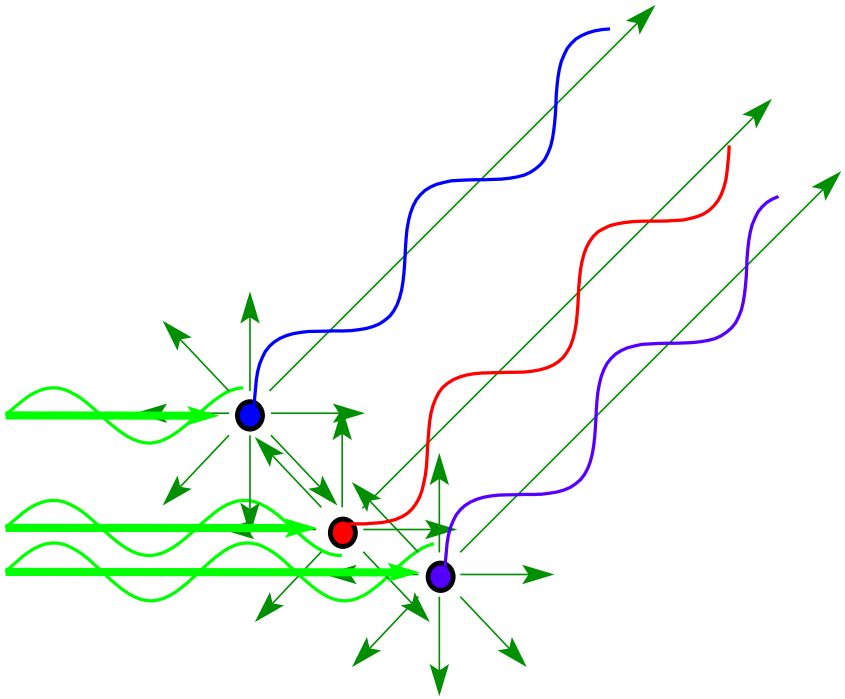




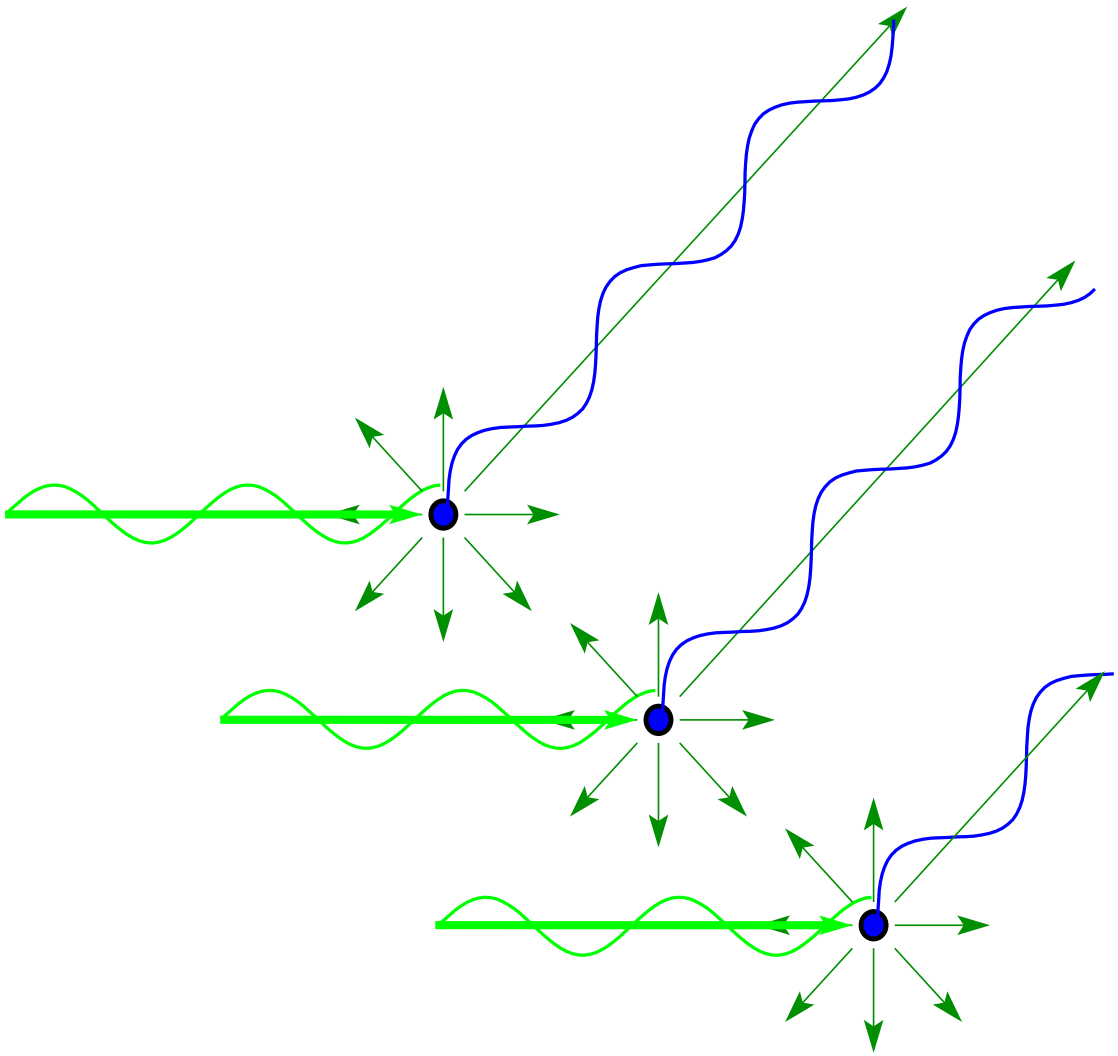
Diffraction



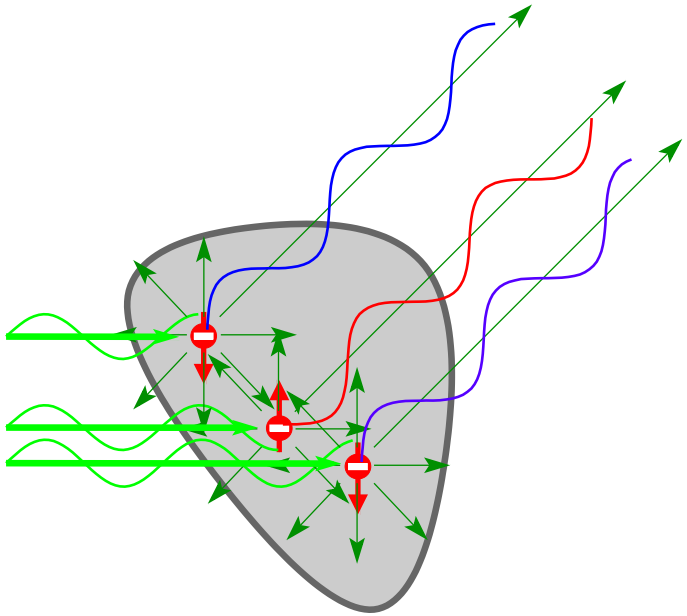
Interference



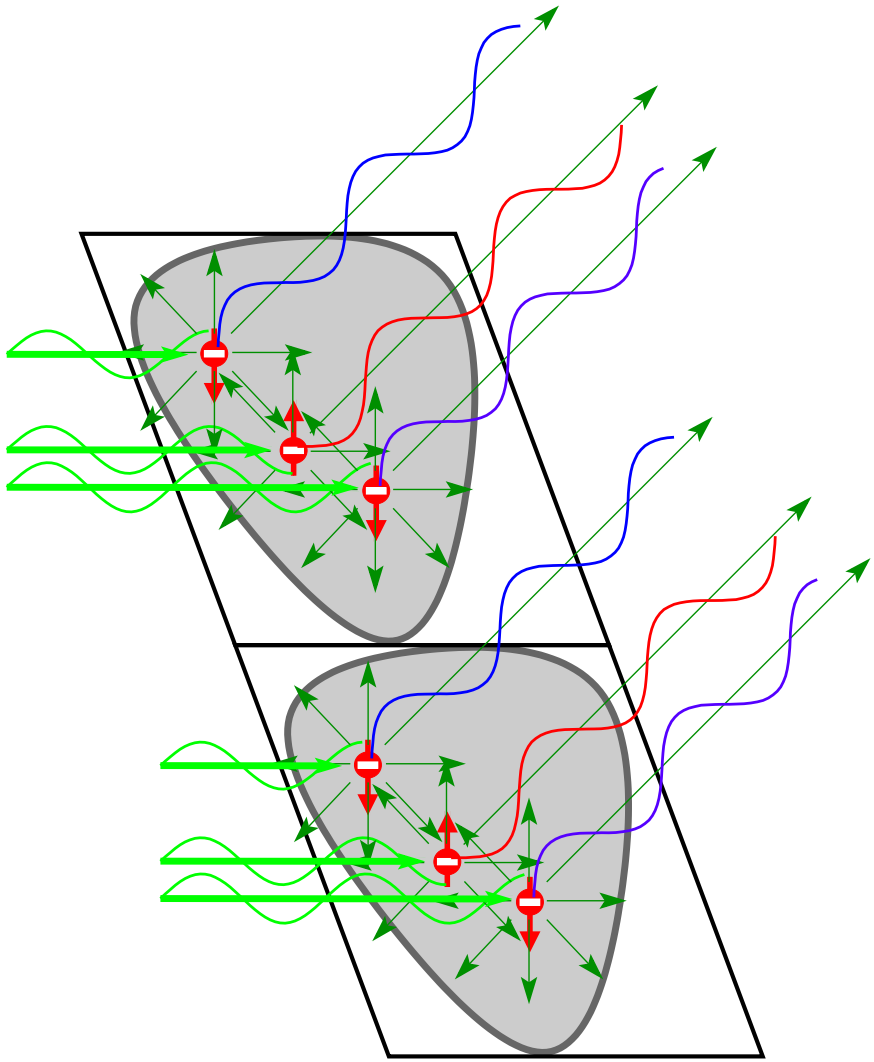
Constructive interference in crystals



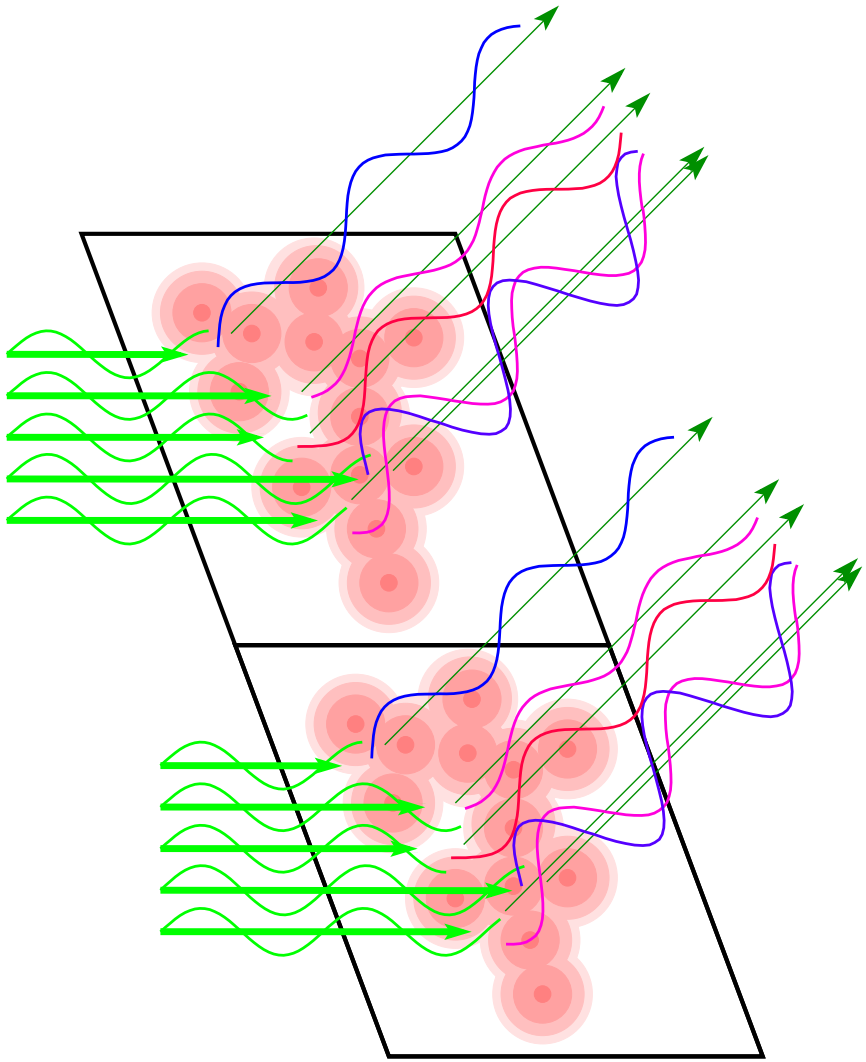
Diffraction by electrons



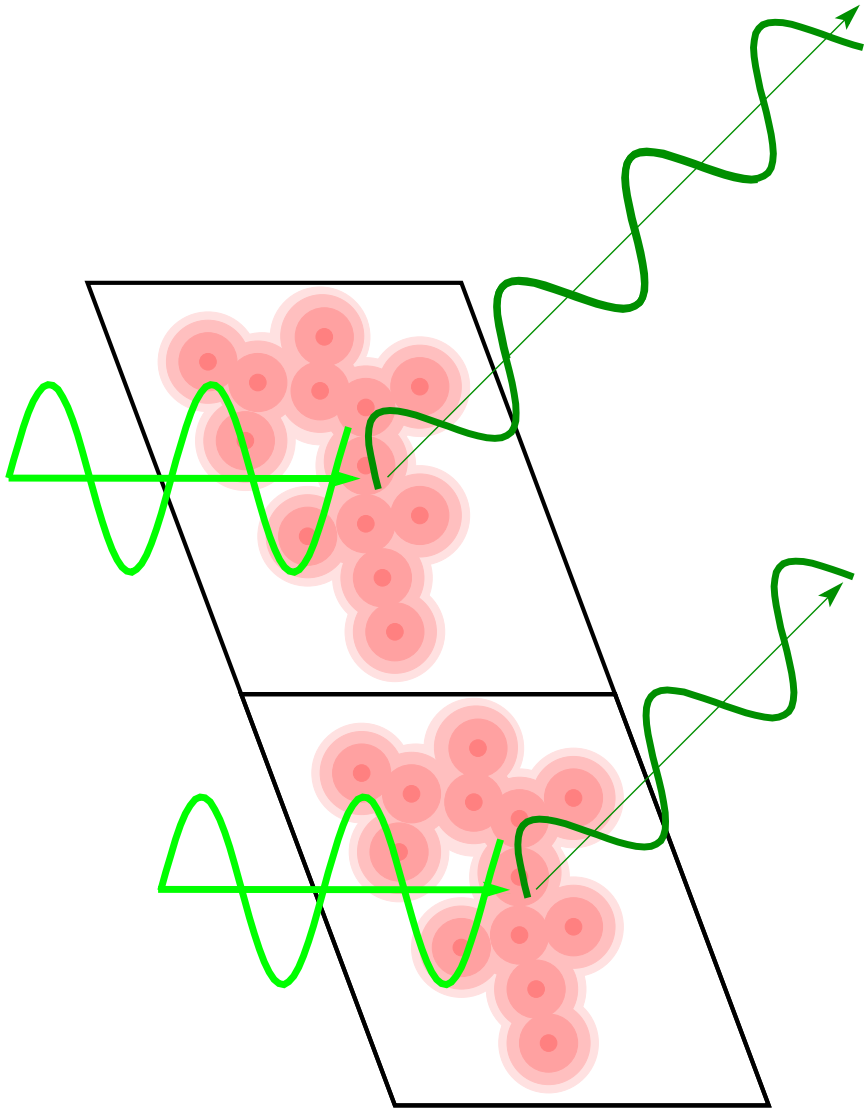
Diffraction by electrons



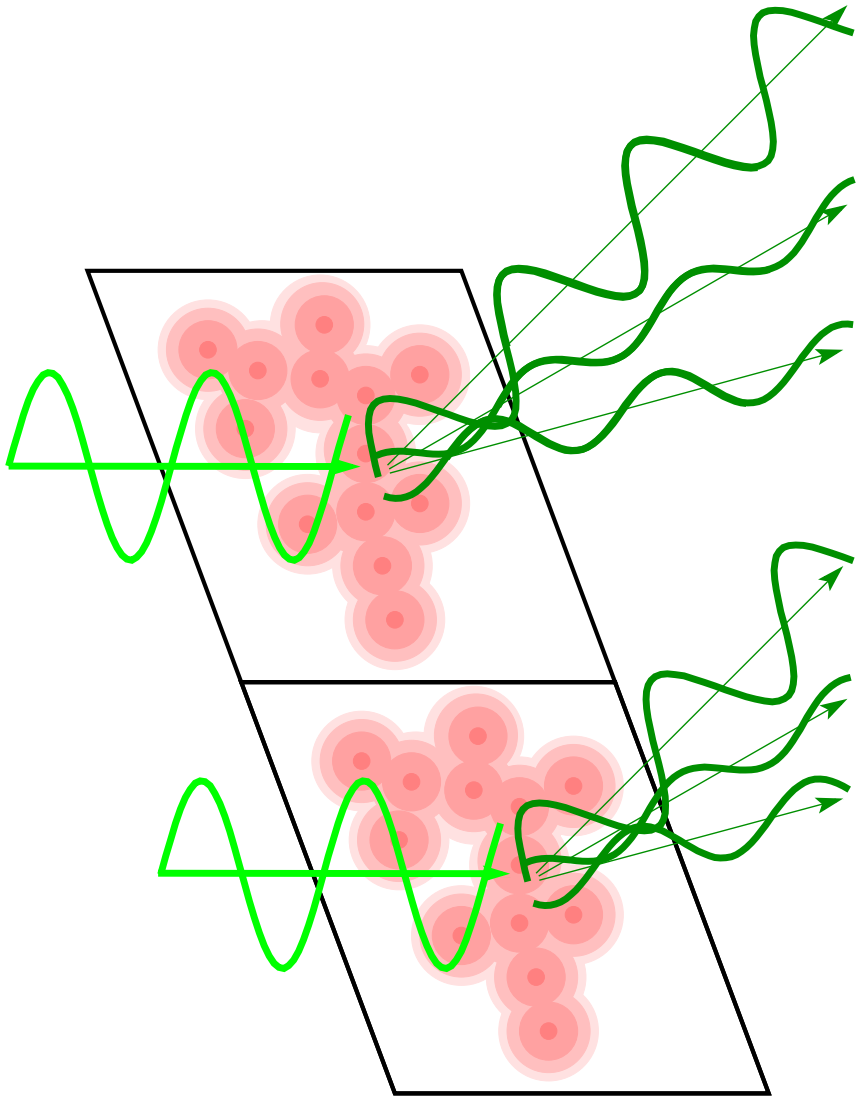
Electron density



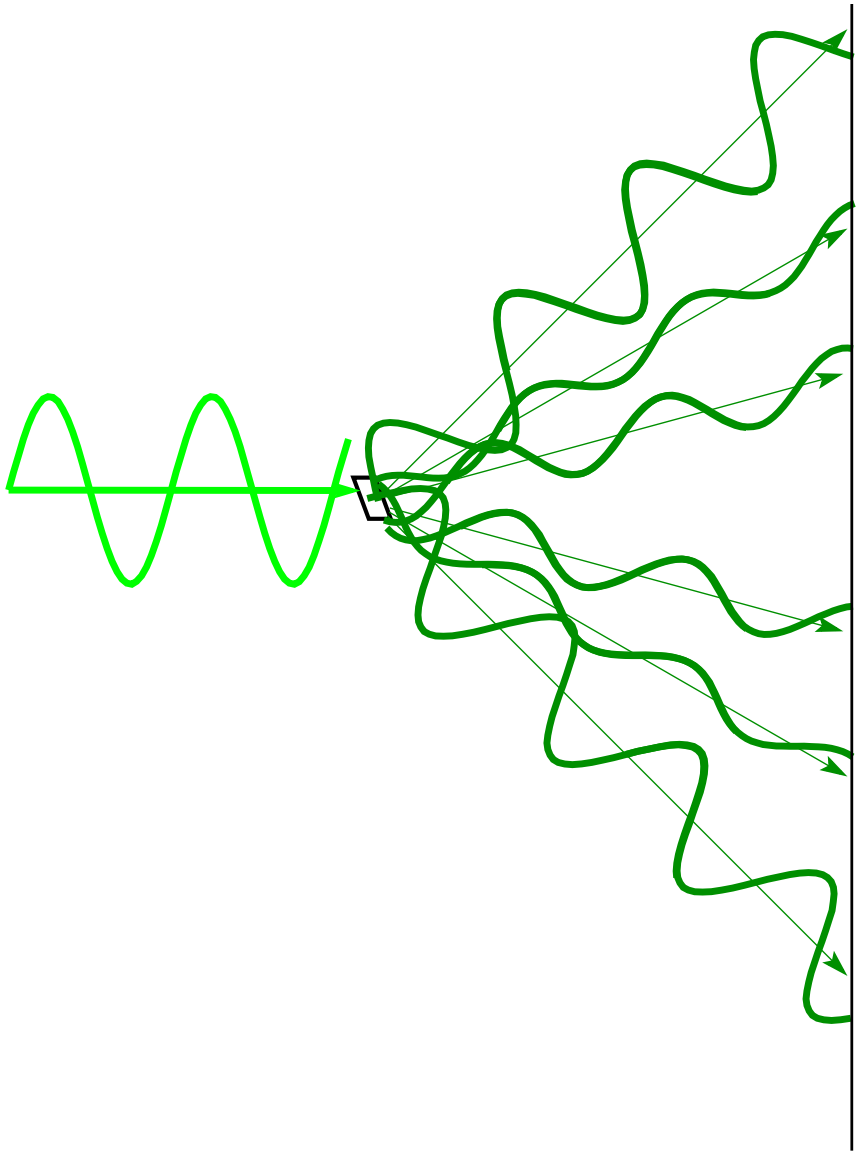
Structure factor



Structure factor



Only amplitude is detected

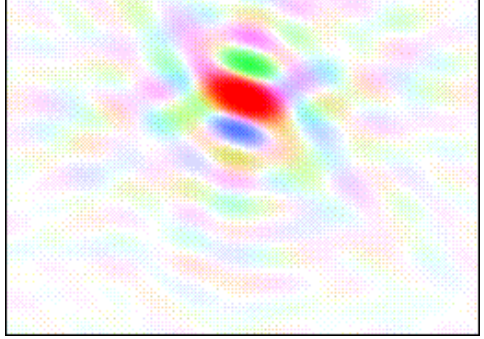


Structure factor

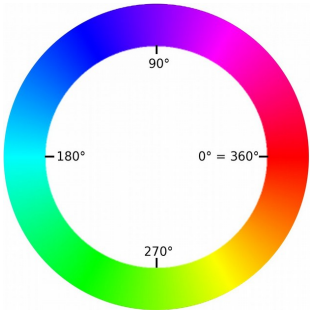
Real space cat

Fourier transform

Fourier amplitudes and phases

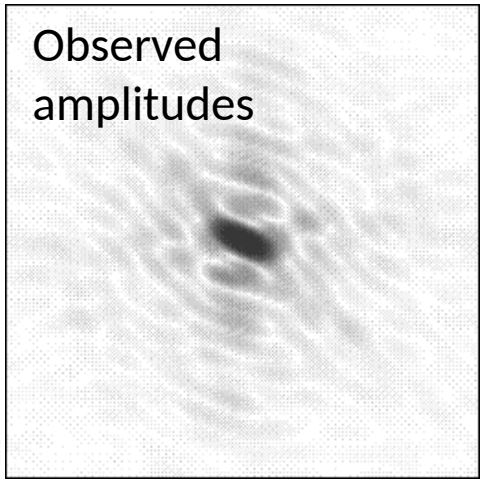


Circular rainbow scale of phases



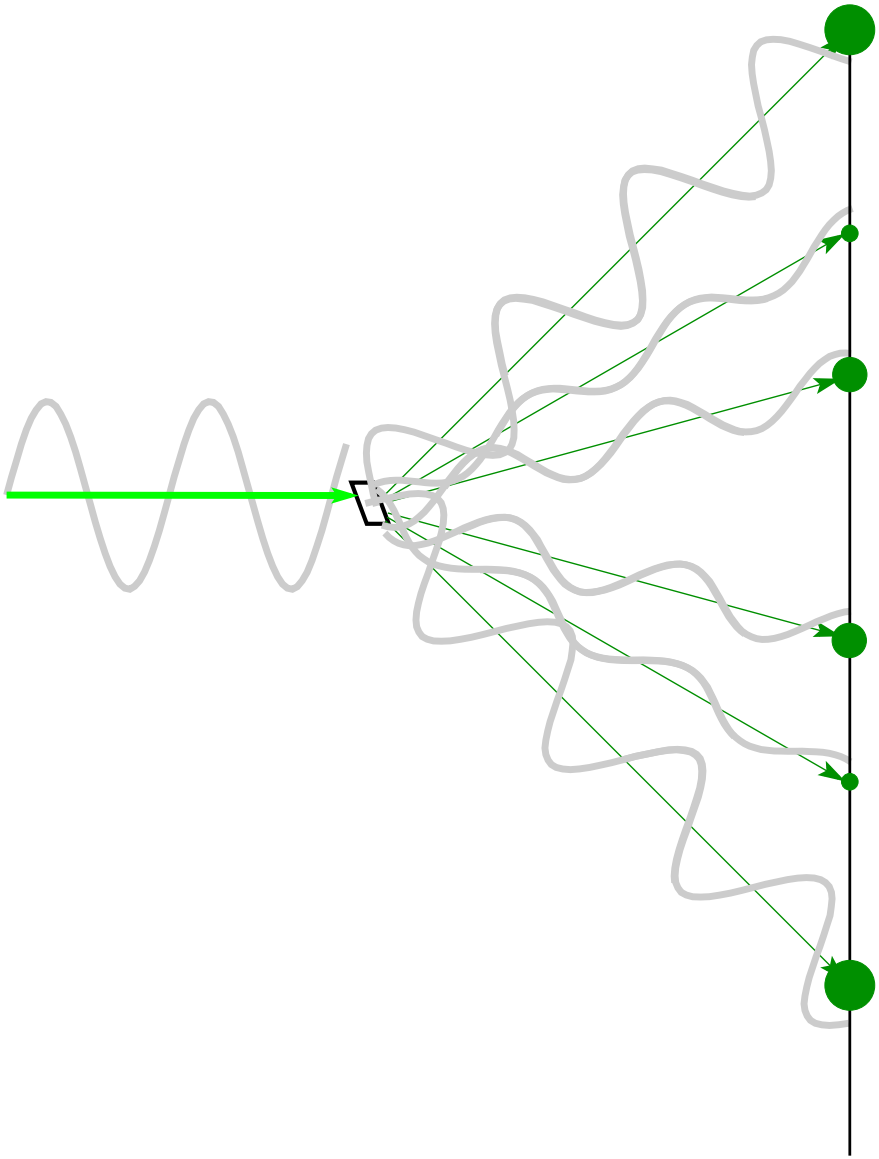
Linear intensity scale of amplitude size

Observed amplitudes



$$F(h\,k\,l) = V \int_{x=0}^{1} \int_{y=0}^{1} \int_{z=0}^{1} \rho(x\,y\,z) \exp[2\pi i(hx + ky + lz)]\,dx\,dy\,dz$$

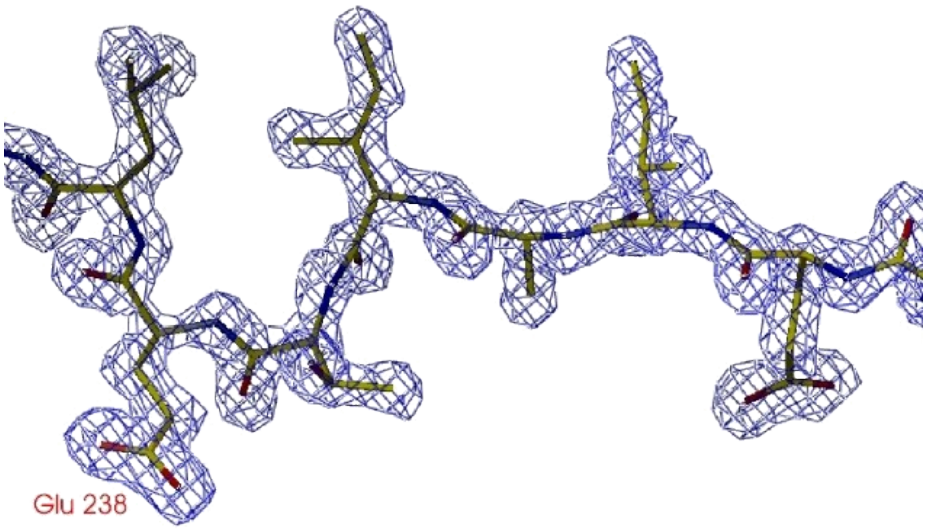
Only amplitude is detected

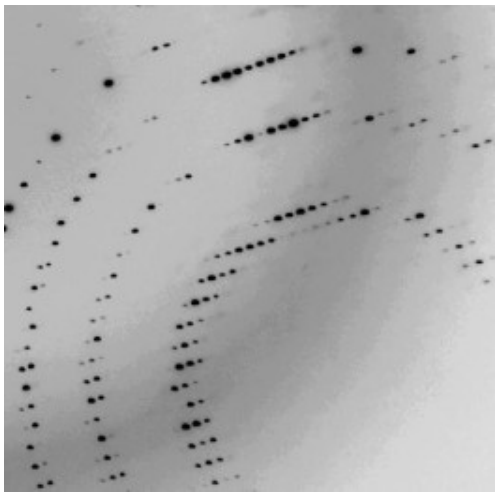


Electron density equation & PHASE PROBLEM

$$\rho(x\,y\,z) = \frac{1}{V}\sum_h\sum_k\sum_l \left|F(h\,k\,l)\right| \exp\left[-2\pi i(hx+ky+lz)+i\alpha(h\,k\,l)\right]$$

$$F(h\,k\,l) = |F(h\,k\,l)|e^{i\alpha(h\,k\,l)}$$



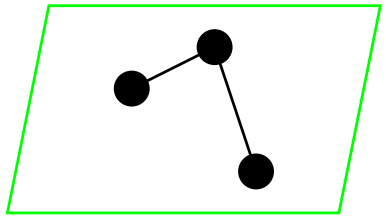


Phase problem

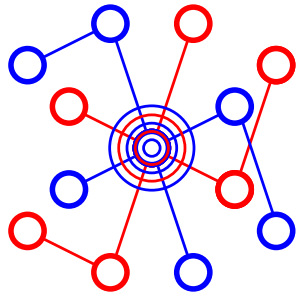
- **Direct interpretation of amplitudes**
  mutual positions of atoms calculated from amplitudes
  for simple molecules (Patterson function/map)
- **Using heavy atoms**
- **Molecular replacement**
  Diffraction back-calculated from a known structure
  similar to the studied proteins
  Orientation and position of the molecule in the crystal
  obtained by searching for the match of diffraction patterns
  (measured vs. back-calculated)
  Calculated phases used for the unknown molecule

Patterson maps

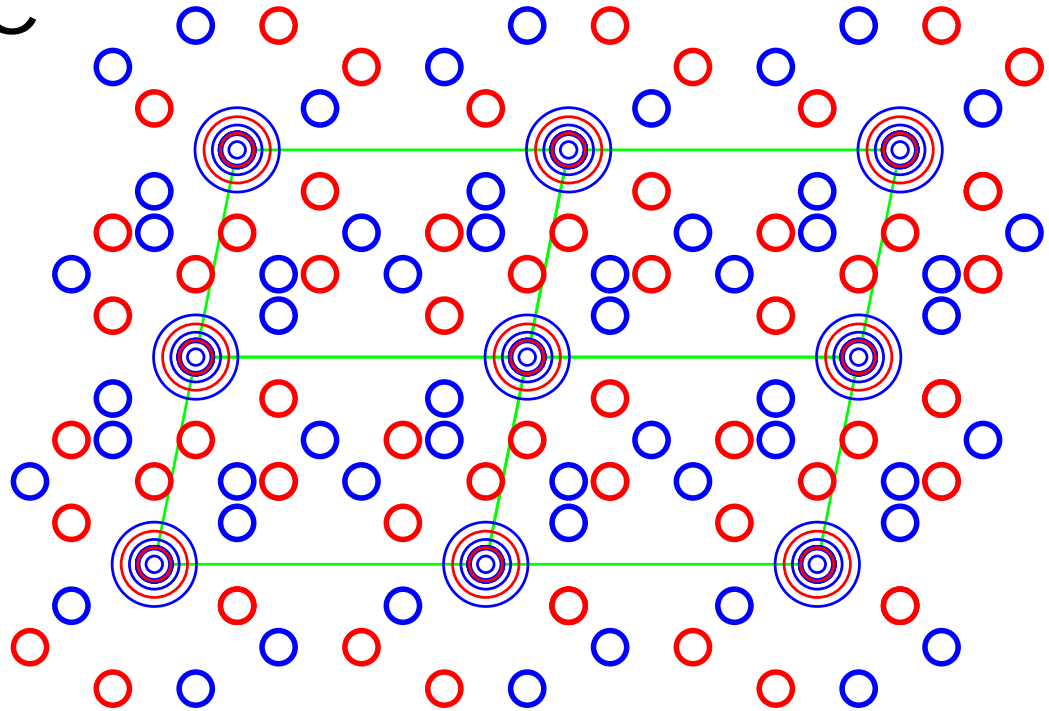
A



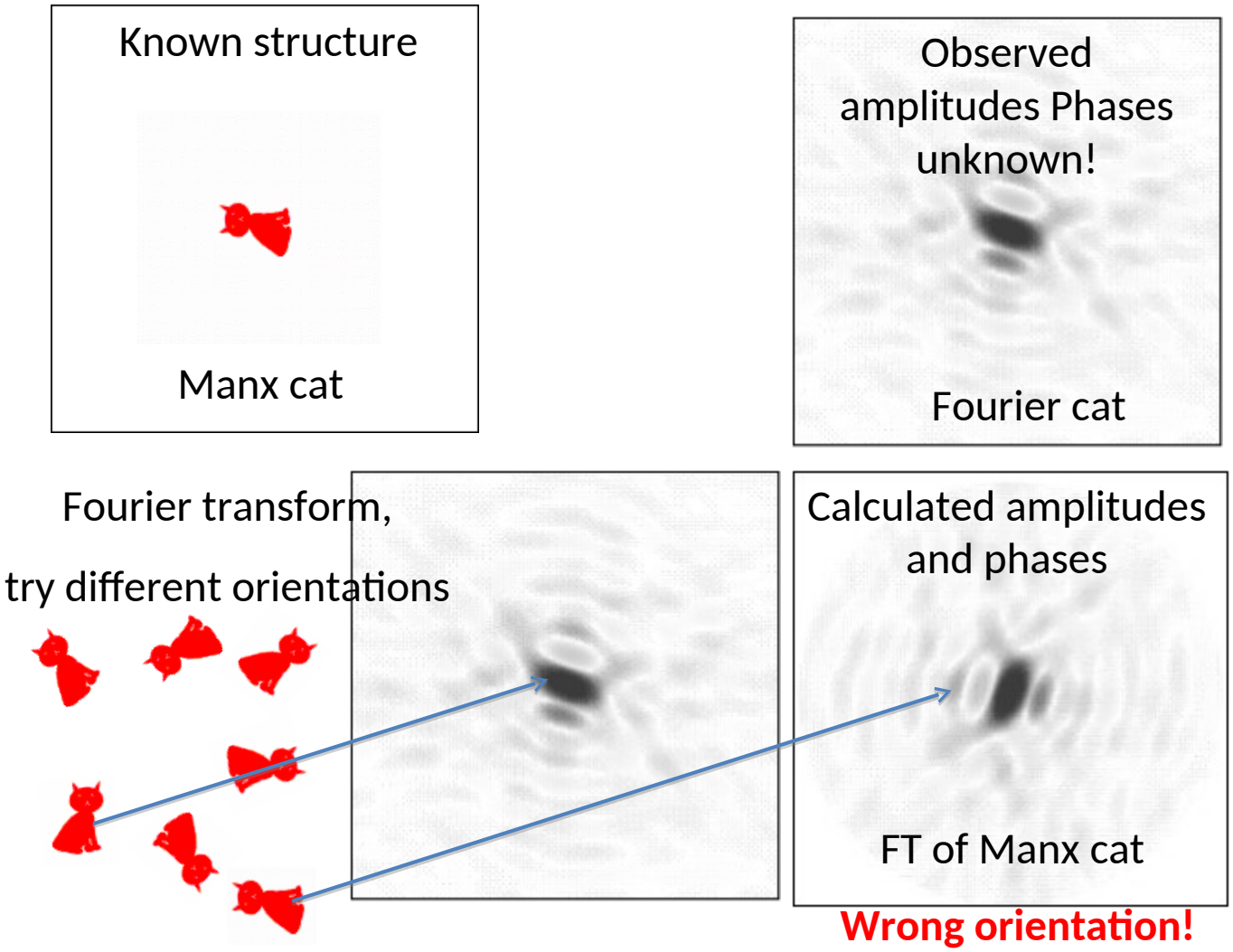
B



C

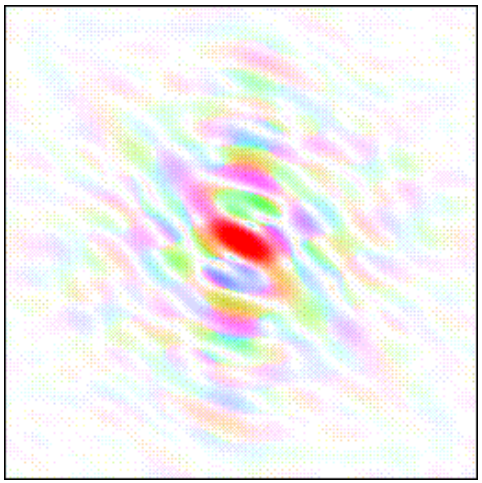


Molecular replacement

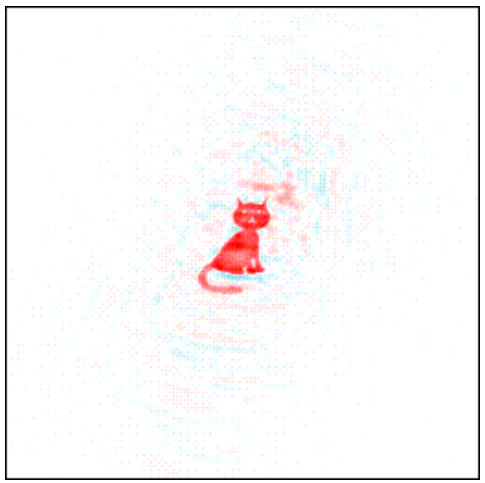


Molecular replacement

Observed **amplitudes** (tailed cat),
calculated **phases** (Manx cat)

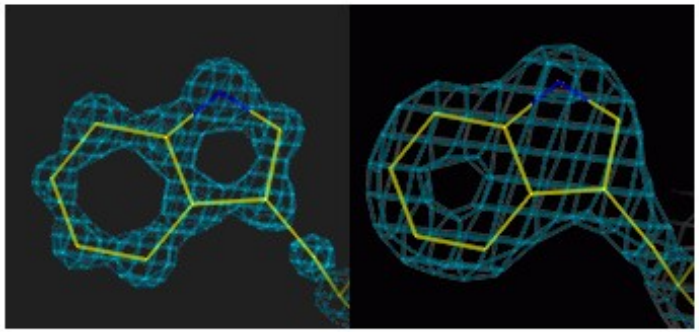


Inverted
Fourier
transform

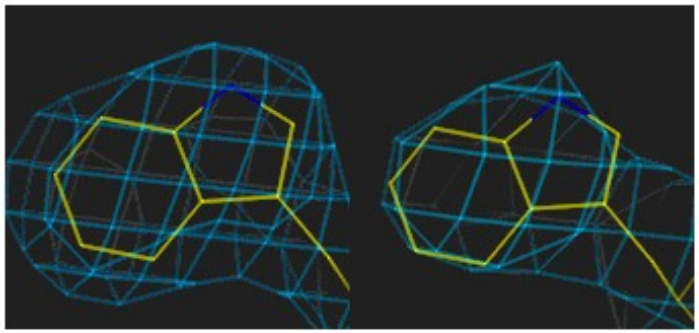


Even the tail becomes visible!

Model building & resolution

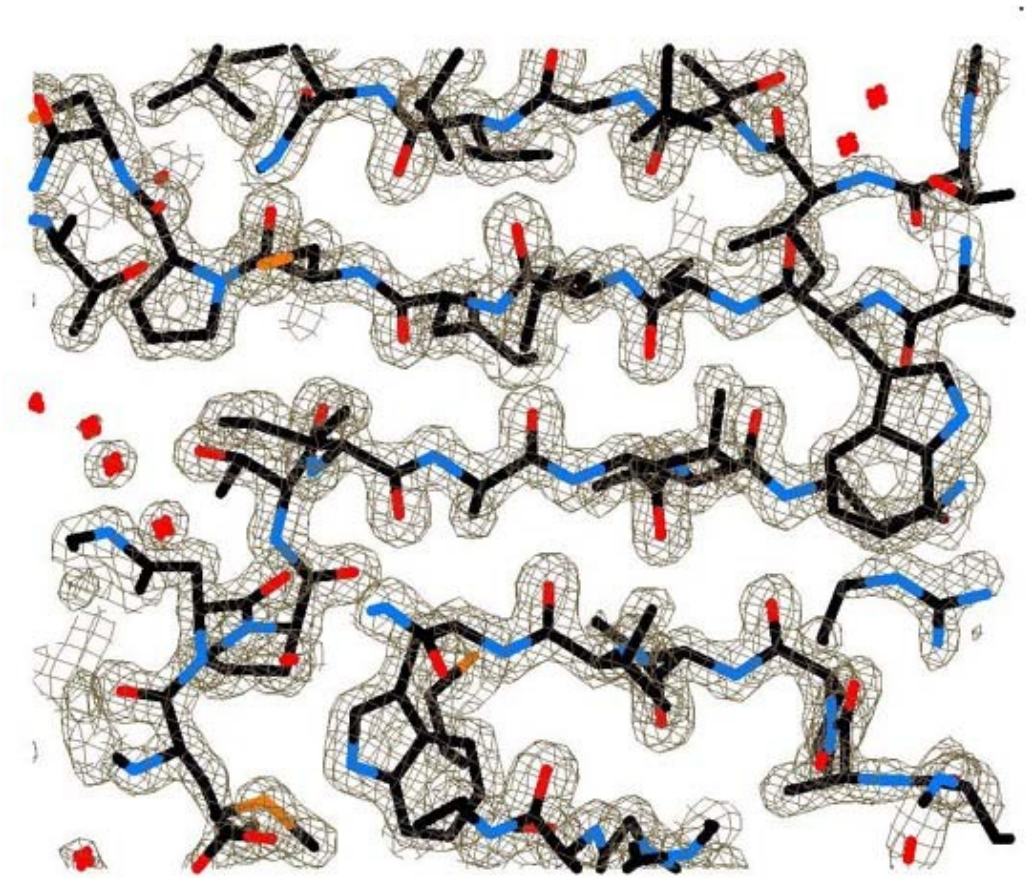


1.0Å 2.5Å

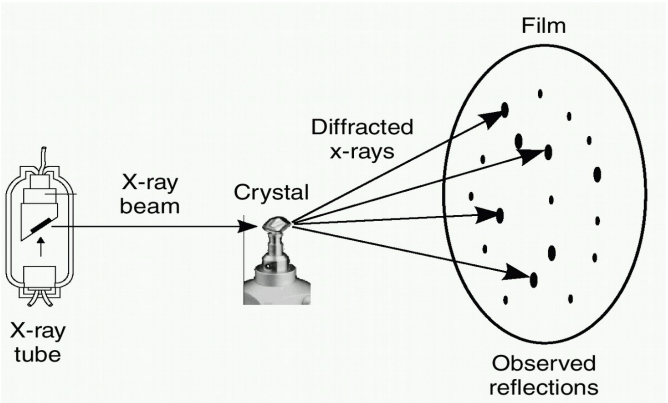


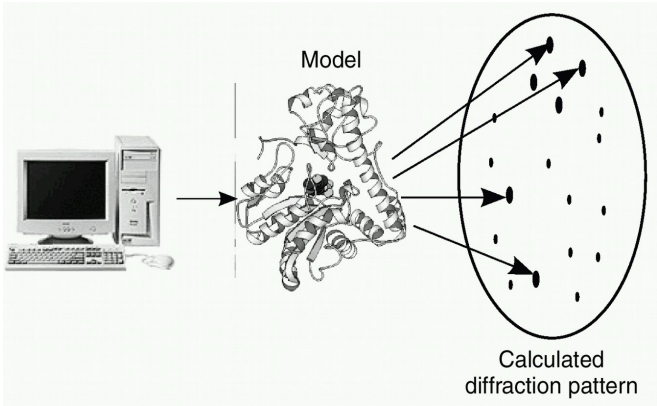
3.0Å 4.0Å

Model building & refinement



R-factor, R_{free} factor





R-factor

$$R = \frac{\sum_{hkl} ||F_{obs}| - k|F_{calc}||}{\sum_{hkl} |F_{obs}|}$$

R_{free} factor

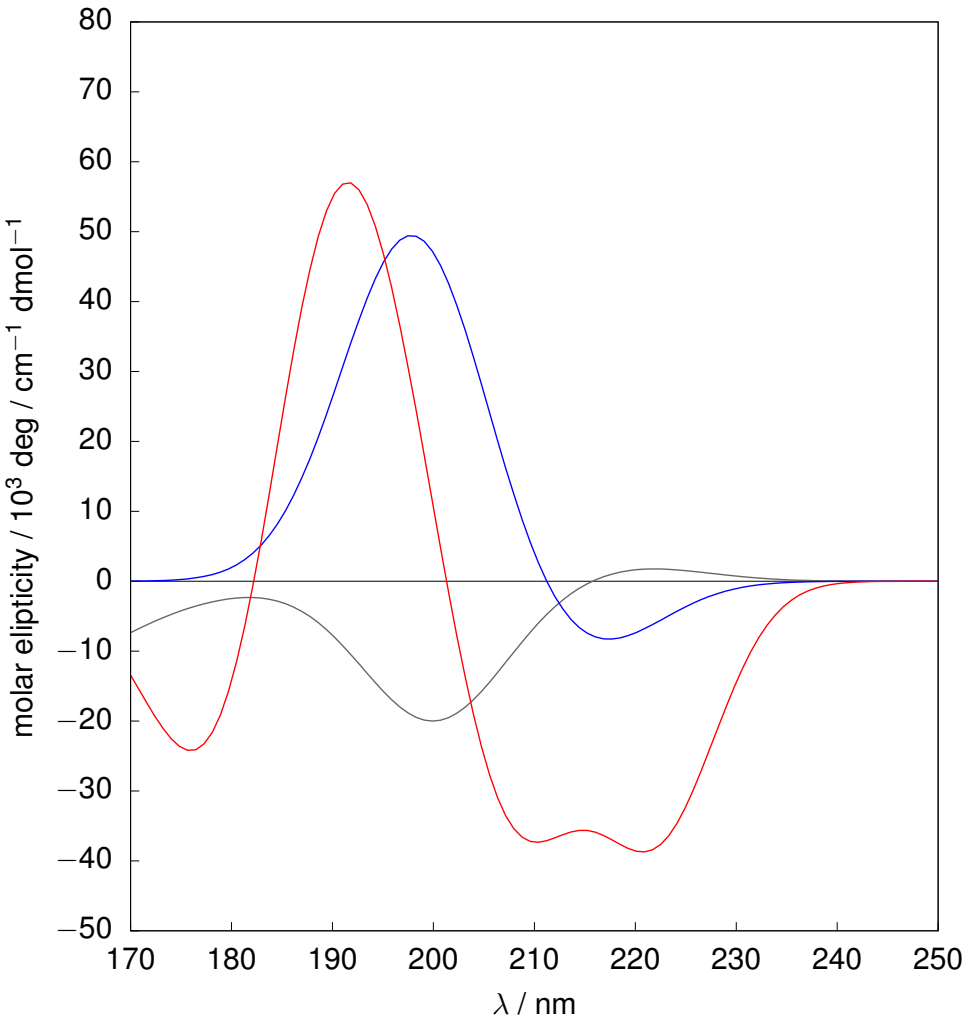
$$R_{free} = \frac{\sum_{hkl \subset T} ||F_{obs}| - k|F_{calc}||}{\sum_{hkl \subset T} |F_{obs}|}$$

Spectroscopy

- **UV/VIS spectrophotometry**
  absorption, transition of electrons to higher orbitals
  concentration, content of aromatic amino acids, heme, prosthetic groups
- **CD spectroscopy**
  absorption differences of polarized light by chiral molecules
  overall content of secondary structures
- **IR spectroscopy**
  absorption, transition of nuclei to higher vibration states

CD spectroscopy and secondary structure

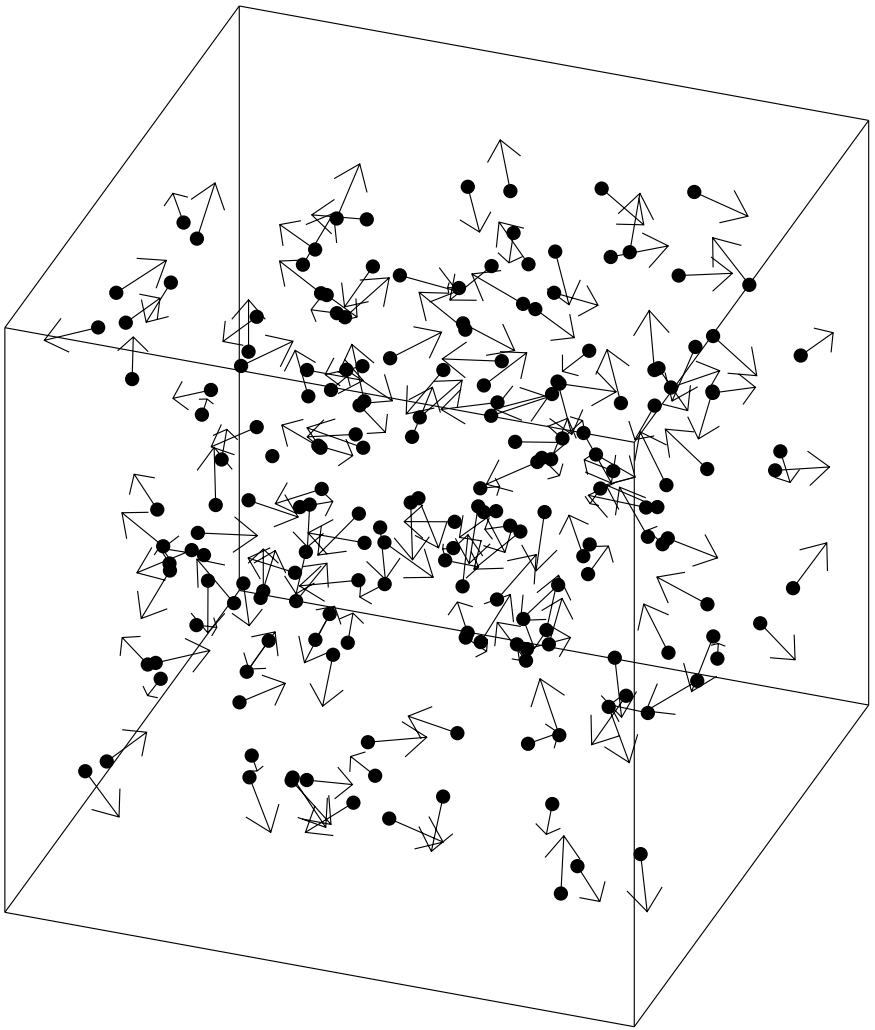
Random coil α-helix β-sheet



Spectroscopy

- **NMR spectroscopy**
  precession of magnetic moments of nuclei in magnetic field
  magnetic moments are slightly aligned in a static magnet
  axis of alignment is tilted by electromagnetic (radio) waves
  aligned magnetic moments precess about the static field
  resulting oscillating magnetic field is measured
  we do not observe the applied electromagnetic waves
  interactions of magnetic moments (mutual, with electrons)
  $\Rightarrow$ resolution, structural information
  atomic resolution structure determination,
  dynamics, interactions

Nuclear magnetic resonance



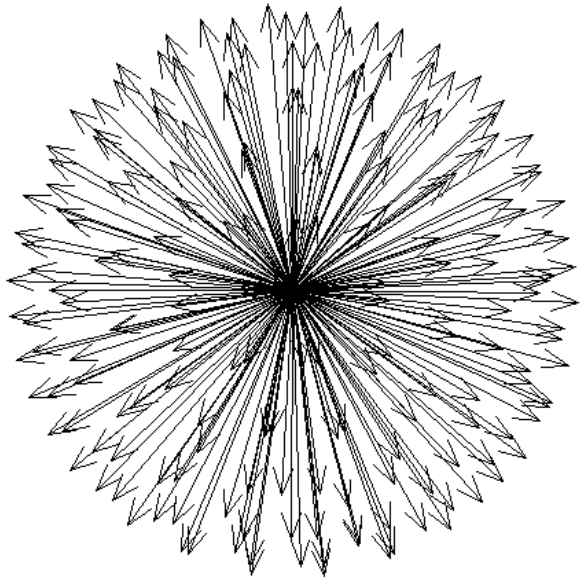
- molecule
- magnetic moment

	S	$\frac{10^{-9}\gamma}{\mathrm{rad\,s^{-1}T^{-1}}}$	%
e^-	1/2	−182.000	100
^{1}H	1/2	0.277	99.98
^{13}C	1/2	0.067	1.1
^{14}N	1	0.019	99.6
^{15}N	1/2	−0.027	0.4
^{17}O	5/2	−0.036	0.04
^{19}F	1/2	0.252	100
^{31}P	1/2	0.108	100
^{129}Xe	1/2	−0.075	24.4

quadrupolar (relax fast)

rare isotopes (enrichment)

NMR sample outside magnet

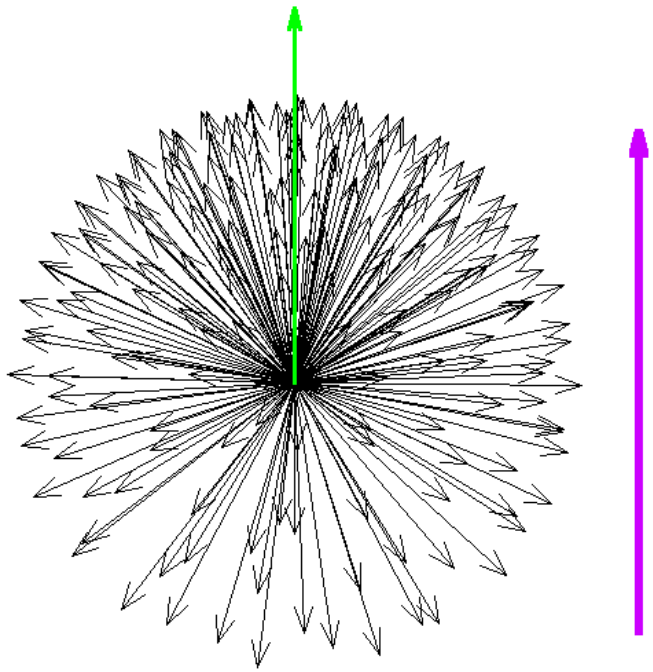


in equilibrium (spherical symmetry)

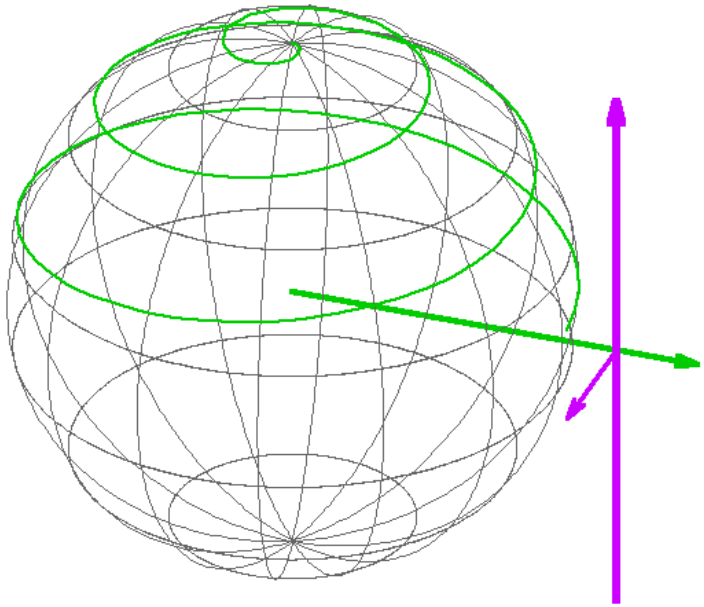
Polarization

Boltzmann distribution: $P(\theta) \propto e^{-\frac{E}{k_B T}} = e^{\frac{\vec{\mu}\cdot\vec{B}}{k_B T}} \Rightarrow M_z = \frac{N}{V}\frac{\mu^2 B}{3k_B T}$

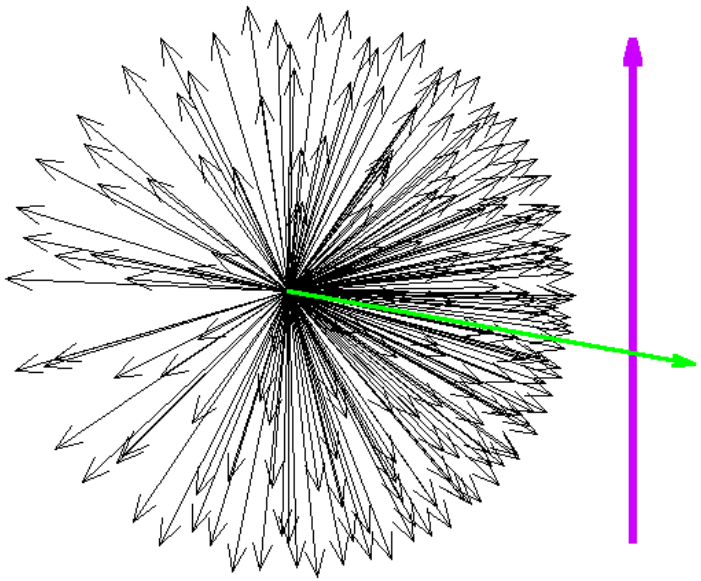
Precession (angular momentum in a field): $\vec{\omega} = -\gamma B$



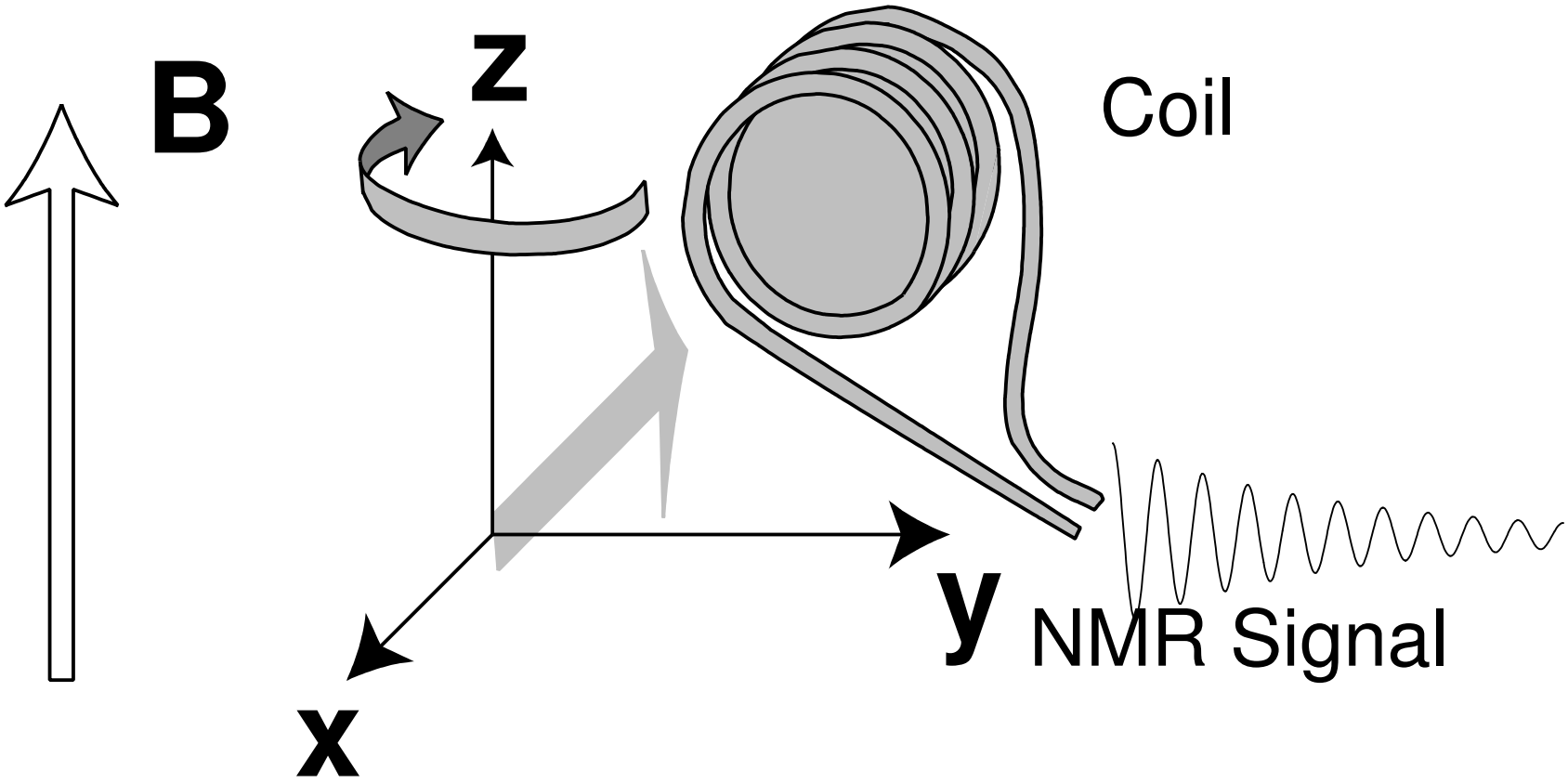
Excitation



Coherent evolution

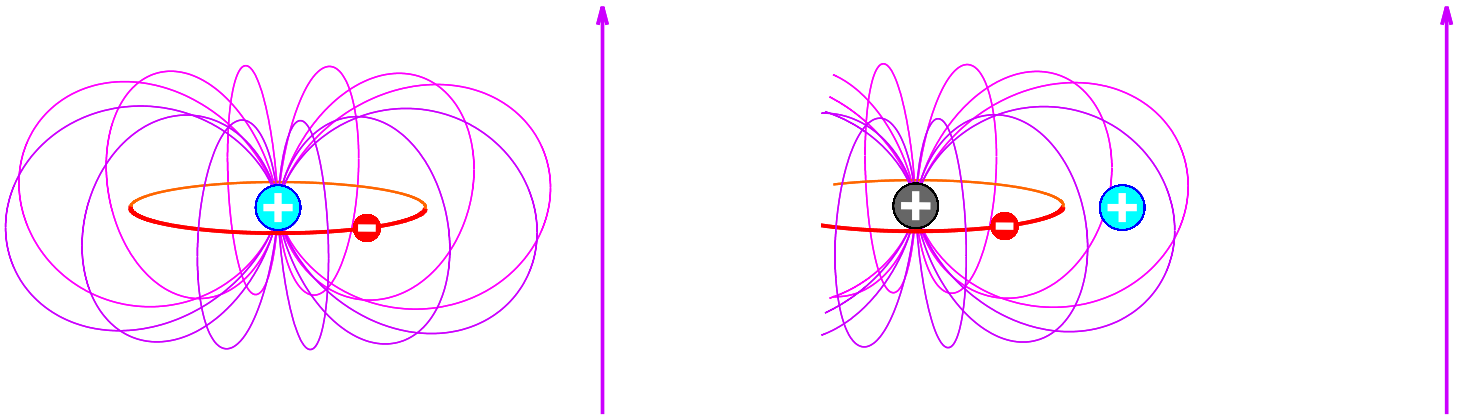


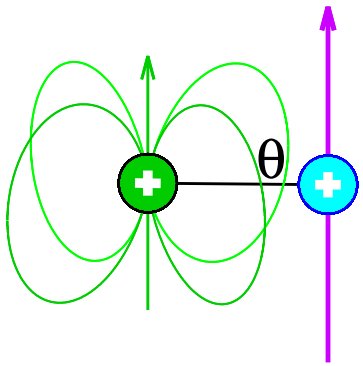
NMR signal detection



reproduced from M. H. Levitt: Spin Dynamics

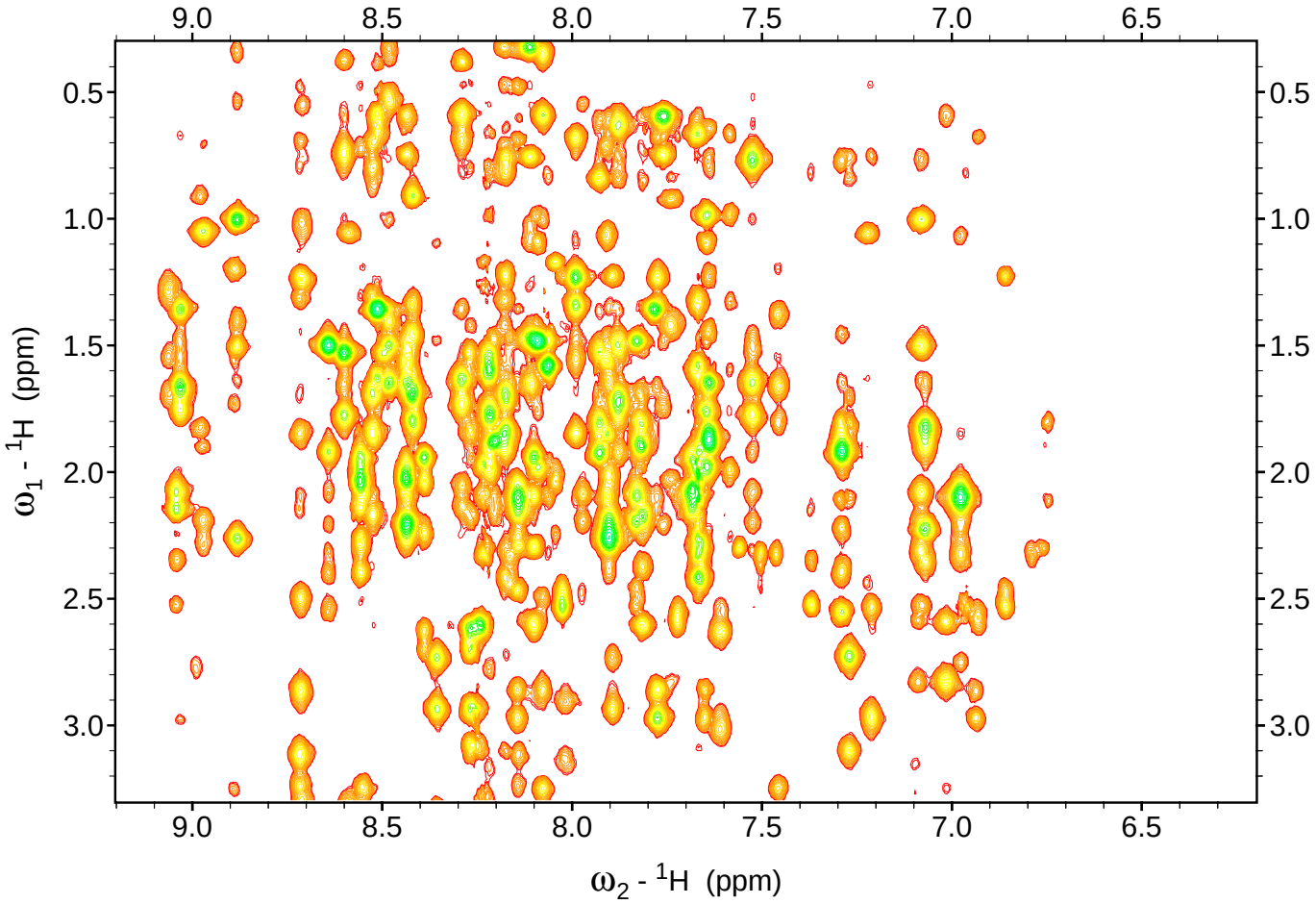
Interactions with other nuclei and electrons





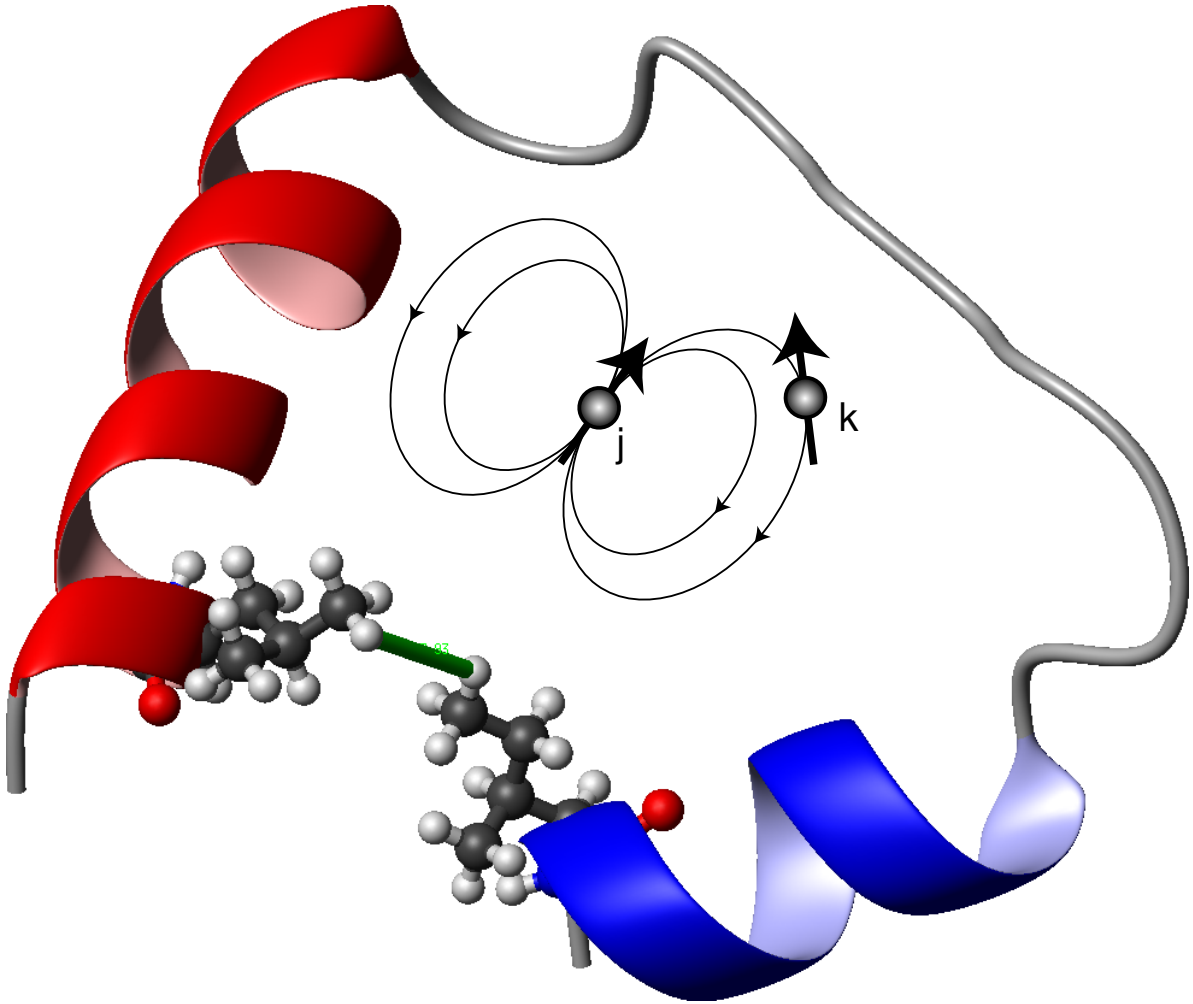
Nuclear Overhauser effect from NOESY spectra

Nuclear Overhauser effect: result of dipole-dipole interactions
Peak intensities are proportional to nuclear Overhauser effect



Distances from nuclear Overhauser effect

Nuclear Overhauser effect proportional to $1/d^6$ (d = distance)



Distance based structure calculation

Model of the protein built from known distances