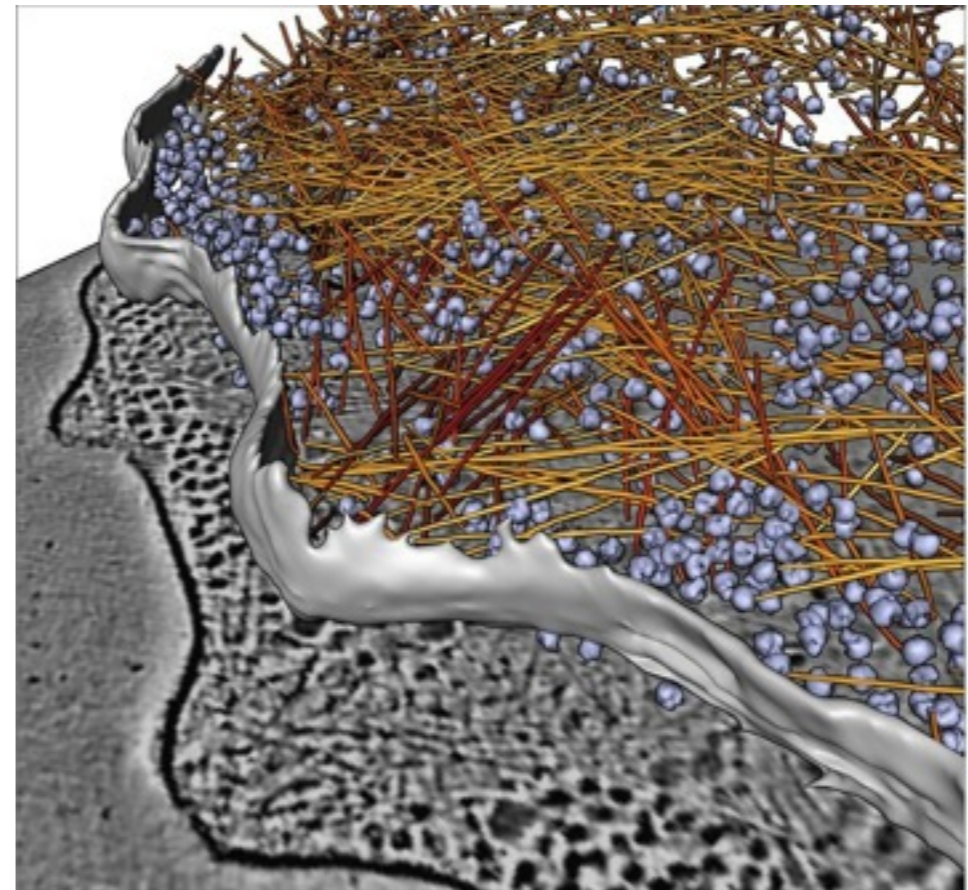


# Interactions between proteins, DNA and RNA

---

The energy, length and time coordinate system to find your way in the cell



# Coordinates for my lecture Tuesday, 14:15-15:45

## BioQuant, seminar room 044

Karsten Rippe

BioQuant

Room 645, 6th floor

Telefon: 54-51376

e-mail: [Karsten.Rippe@dkfz.de](mailto:Karsten.Rippe@dkfz.de)

Overview on learning Biophysics in Heidelberg

<http://malone.bioquant.uni-heidelberg.de/teaching>

Material for the lecture: Biophysical concepts and theoretical descriptions

[http://malone.bioquant.uni-heidelberg.de/BPC\\_2014/BPC\\_1+2.html](http://malone.bioquant.uni-heidelberg.de/BPC_2014/BPC_1+2.html)

**Username: teaching**

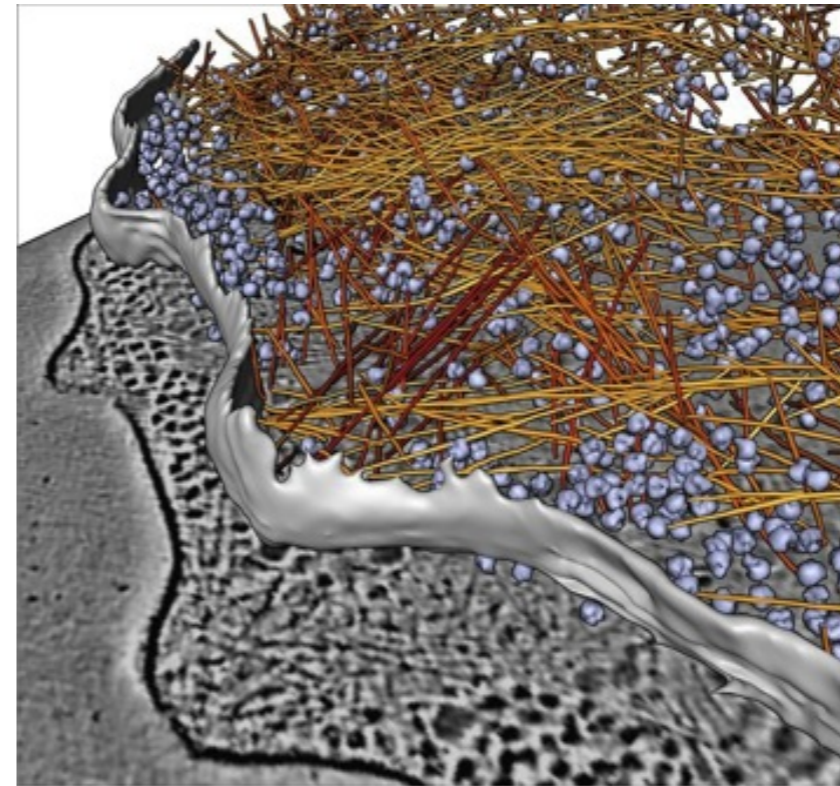
**Password: nonukes**

**(currently not active)**

# Problem sets

- Starting today there will be problem sets every 1-2 weeks (8-10 total)
- These have to be returned until the following Tuesday at the beginning of lecture
- Then answers to the problem set (and any problems) are discussed
- Corrected problem sets are returned the next Tuesday (ideally...)
- “Benotete Scheine” will be given according to the results from the problem set
- You can select one bonus problem set of your choice, i. e. one can get 100% with one missing or  $> 100\%$  if one does all problem sets correctly

# The energy, length and time coordinate system to find your way in the cell



- Length: *E. coli* - Yeast cell - Human cell
- Speed: *E. coli* - how fast does it stop and go?
- Energy: Hydrolysis of ATP?

# Quantitative biology needs numbers...

**Cell Biology by the Numbers** (upcoming book by Rob Philipps & Ron Milo, Lecture by Ron Milo).

Book draft: <http://book.bionumbers.org>

Bionumbers database

<http://bionumbers.hms.harvard.edu>

- Develop a coordinate system and intuition about the molecular life of a cell - “the microscopic world”
- “Sanity” checks (lecture, discussion, design of experiment...)
- Try to rationalize **why** the cell operates in a certain manner

How many mRNA molecules  
are in an *E. coli* cell?

a)  $10^3 - 10^4$

b)  $10^5 - 10^6$

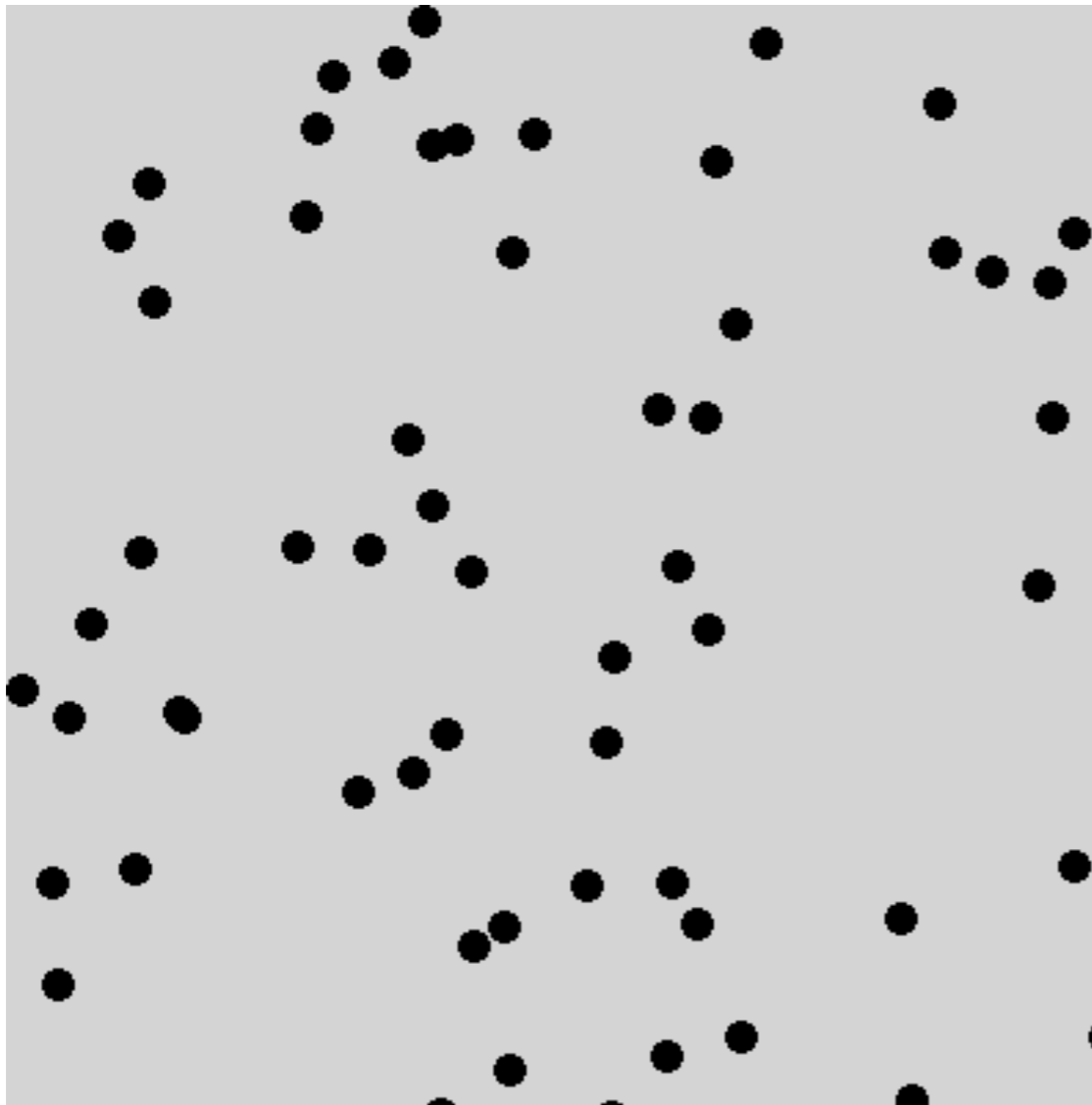
c)  $10^7 - 10^8$

d)  $10^9 - 10^{10}$

Before Ludwig Boltzmann got depressed and killed himself, he did some really great things, like inventing the Boltzmann constant



# Molecules in an ideal gas



the average kinetic energy of a particle with mass  $m$  along the x-axis is related to  $k_B T/2$

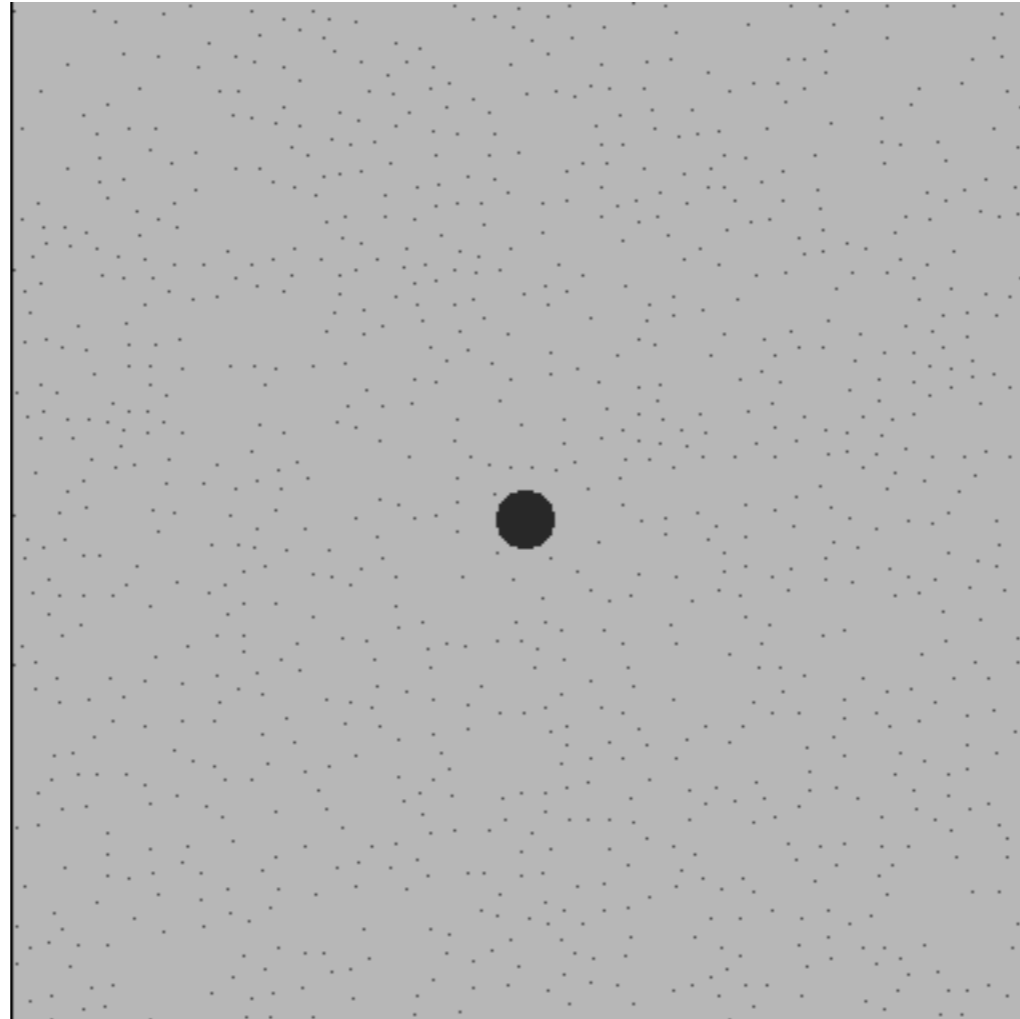
$$\left\langle \frac{m v_x^2}{2} \right\rangle = \frac{kT}{2} \quad \left\langle v_x^2 \right\rangle = \frac{kT}{m}$$

$$\sqrt{\left\langle v_x^2 \right\rangle} = \sqrt{\frac{kT}{m}}$$

A protein like lysozyme ( $M = 14$  kDa) at room temperature would have a speed of  $\langle v_x^2 \rangle^{1/2} = 13$  m/sec **in vacuum**.



# Movements of a macromolecule in solution



- The macromolecule collides with water molecules and moves in a “random walk”.
- The length of the free path is much smaller than the diameter of the particle.
- The average distance from the starting point is proportional to the square root of time.

$k_B T$  is the energy available for spontaneous reactions

$$P_i \propto g_i \cdot \exp\left(\frac{-E_i}{k_B T}\right)$$

The Boltzmann equation yields the probability  $P_i$  to find a molecule with energy  $E_i$

- $g_i$ : number of different states with energy  $E_i$
- $k_B$ : Boltzmann constant
- $T$ : Temperature

probability to find a particle with an energy

- of  $k_B T$  or larger: 0.37  $\Rightarrow$  processes that requires an energy of  $k_B T$  occur spontaneously
- of  $10 k_B T$  or larger: 0.00005  $\Rightarrow$  these processes will not occur spontaneously

at 25 °C  $k_B T = 4.1 \cdot 10^{-21}$  J or  **$k_B T = 4.1$  pN·nm**

$k_B T$  refers to a single molecule

for of 1 mol of particles one has to use  $k_B T \times 6.022 \cdot 10^{23} = RT$

at 25 °C with  $= 8.3 \text{ J} \cdot \text{mol}^{-1} \cdot \text{K}^{-1} \Rightarrow$   **$RT = 2.5$  kJ/mol or 0.6 kcal/mol**

Hydrolysis of ATP:  $\sim 20 k_B T/\text{ATP}$ , 10 - 15 kcal/mol or 40 - 60 kJ/mol

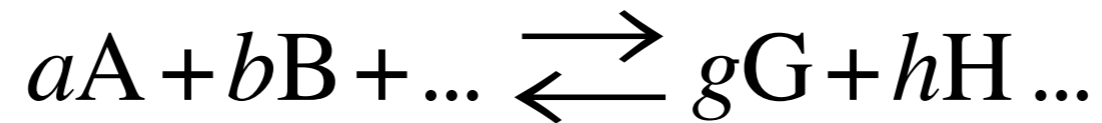
# The free energy $\Delta G$

- At constant pressure  $P$  and constant temperature  $T$  the system is described by the Gibbs free energy:

$$G \equiv H - TS \qquad \Delta G = \Delta H - T \Delta S$$

- $H$  is the enthalpy or heat content of the system,  $S$  is the entropy of the system
- a reaction occurs spontaneously only if  $\Delta G < 0$
- at equilibrium  $\Delta G = 0$
- for  $\Delta G > 0$  the input of energy is required to drive the reaction

# $\Delta G$ of an reaction in equilibrium



$$0 = \Delta G^0 + RT \ln \left( \frac{[G]^g [H]^h \dots}{[A]^a [B]^b \dots} \right)_{\text{Eq}}$$

$$\Delta G^0 = -RT \ln \left( \frac{[G]^g [H]^h \dots}{[A]^a [B]^b \dots} \right)_{\text{Eq}} = -RT \ln K$$

$$K = \left( \frac{[G]^g [H]^h \dots}{[A]^a [B]^b \dots} \right)_{\text{Eq}} = \exp \left( \frac{-\Delta G^0}{RT} \right)$$

# Weak forces maintain macromolecular structure

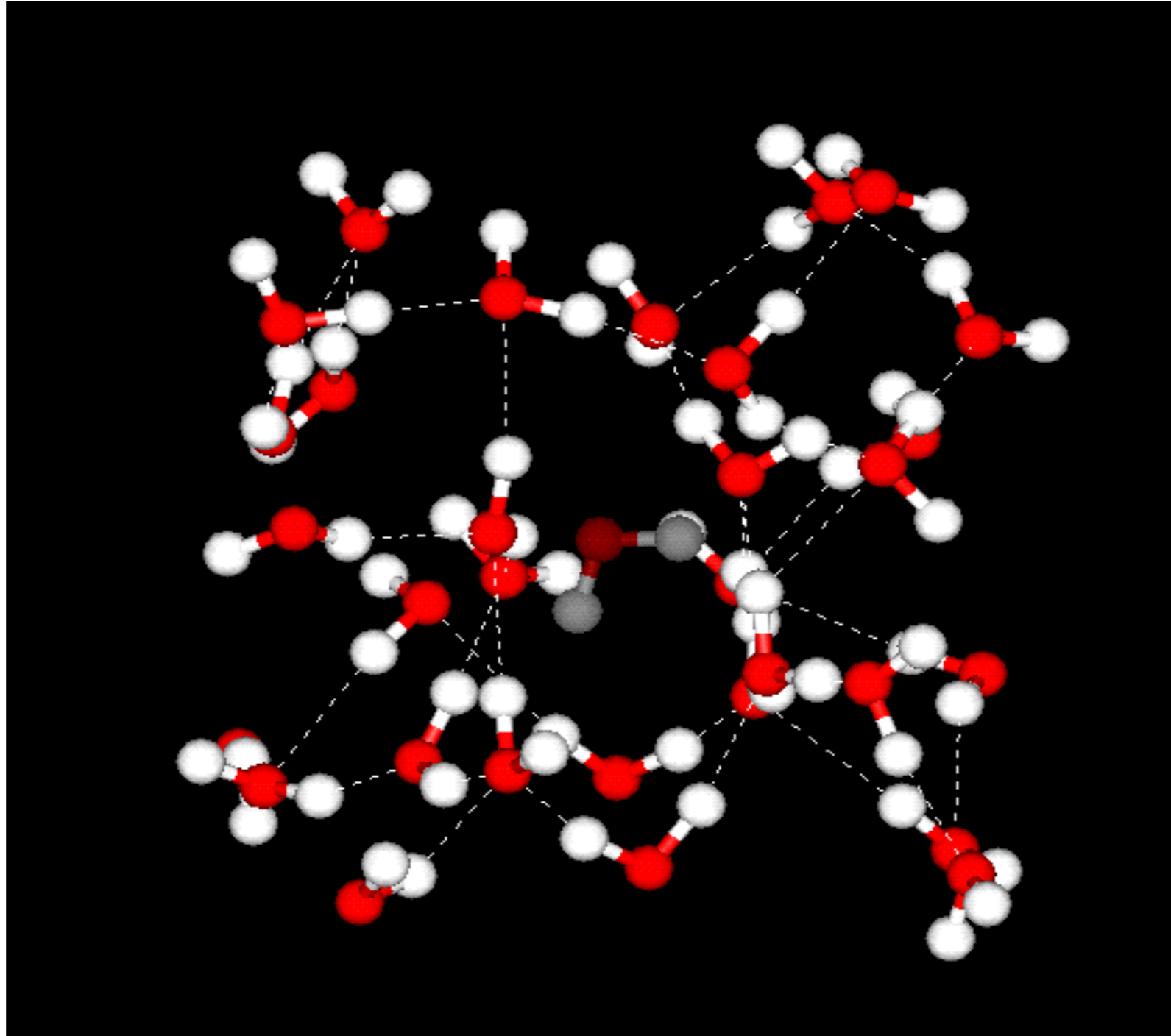
At 25 °C, 0.6 kcal/mol or 2.5 kJ/mol thermal energy: enough to break non-covalent bonds

=> Noncovalent interactions constantly forming and breaking

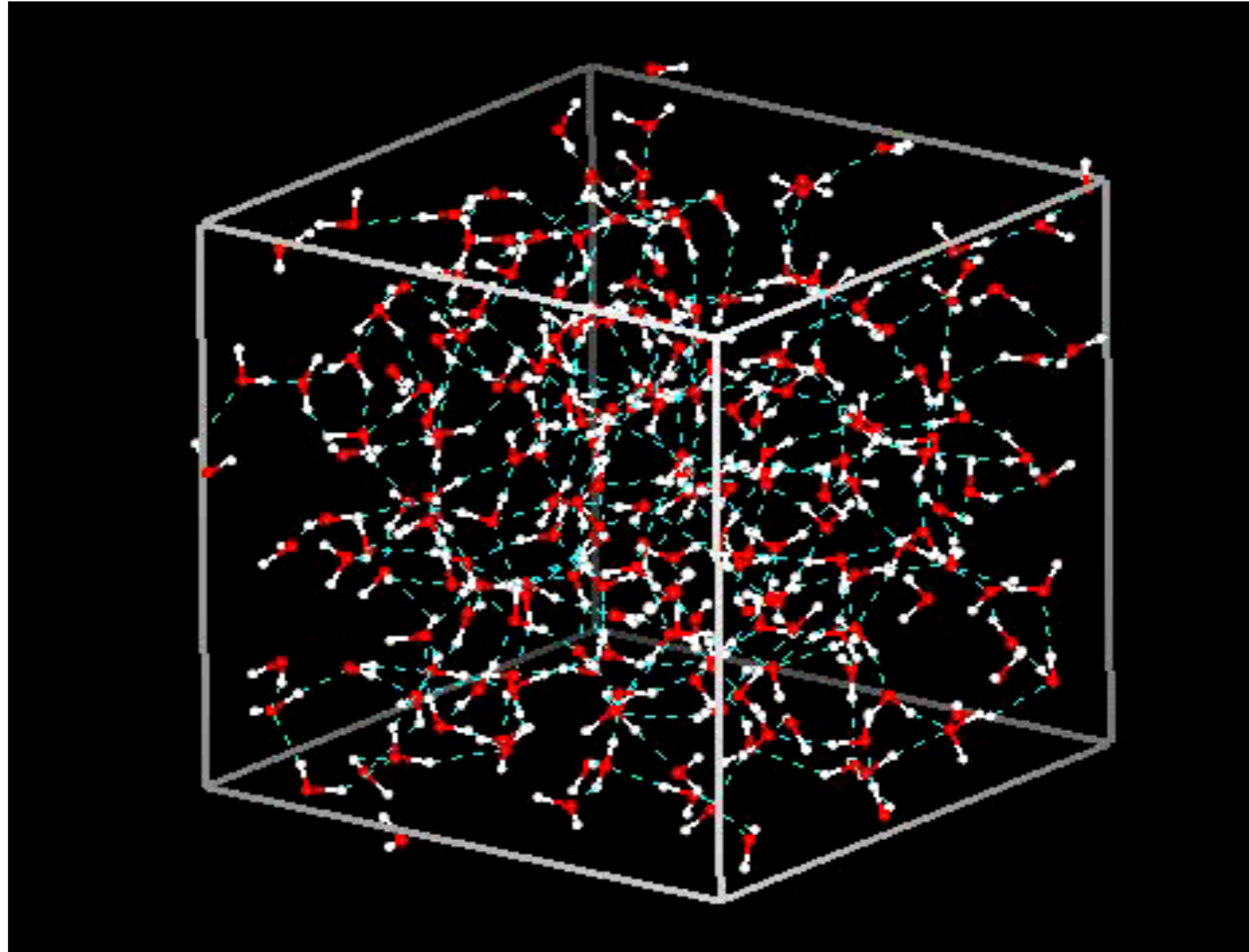
- Van der Waals Attractive Forces: 0.1 - 1.0 kcal/mol,  $1/r^6$
- Hydrogen Bonds: 1 - 2 kcal/mol,  $1/r^3$
- Ionic Interactions: 5 kcal/mol physiological,  $1/r$
- Covalent bonds: 50 - 200 kcal/mol,  $1/r^{\text{very high}}$
- Hydrolysis of ATP: 10 - 15 kcal/mol, 40 - 60 kJ/mol,  $\sim 20 k_B T/\text{ATP}$

**but they can occur between macromolecule and water/ions just as well...**

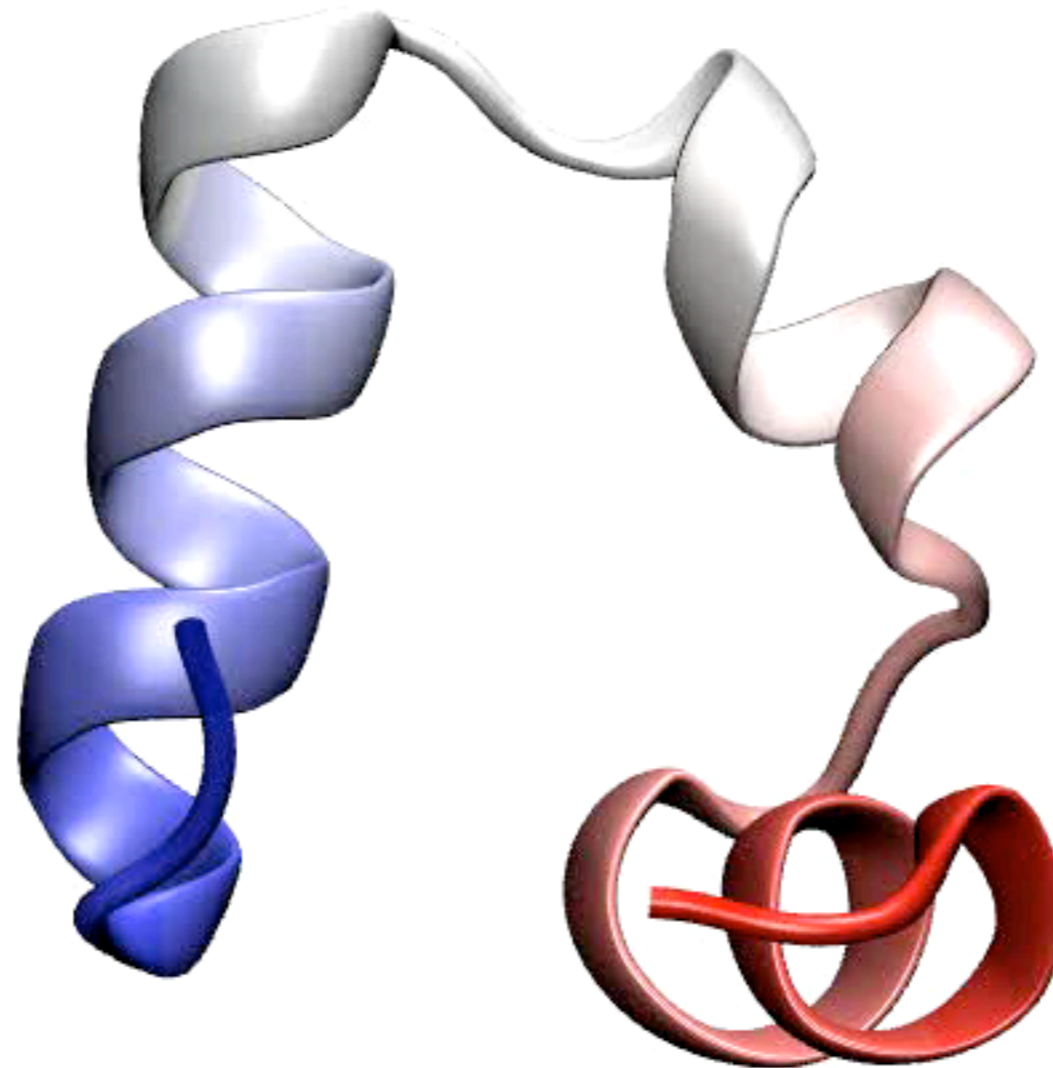
# Hydrogen bonding of liquid water



# Hydrogen bonding of liquid water

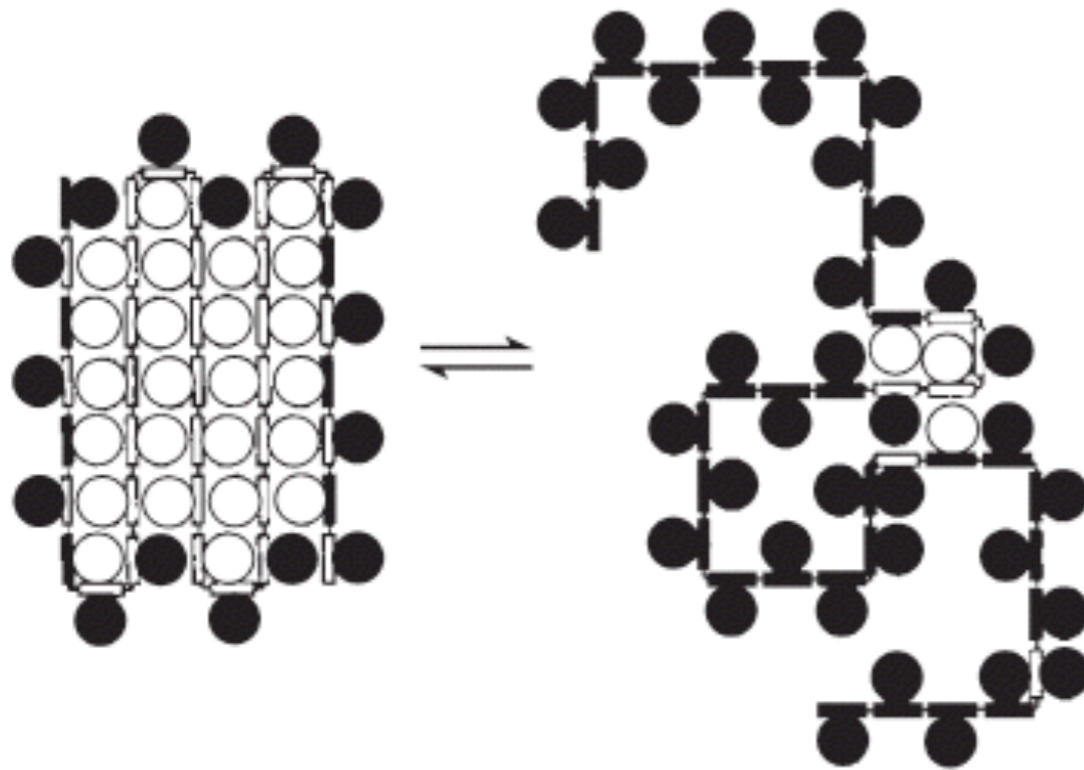


# Protein folding as seen in molecular dynamics simulations of the villin protein headpiece ( $\sim 10 \mu\text{s}$ time scale)





# Scheme of protein folding with solvent accessible amino acids in black



folded state

random coil

Table 1. Folded globular proteins.

	buried <sup>a</sup>
non-polar side chains (Ala, Val, Ile, Leu, Met, Phe, Trp, Cys)	83%
peptide groups (-CO-NH-)	82%
<i>ca.</i> 1.1 intramolecular hydrogen bonds formed per residue <sup>b</sup>	

<sup>a</sup> Lesser & Rose (1990).

<sup>b</sup> Stickle *et al.* (1992).

# Calculating entropy: how probable or disordered is the final state?

Entropy provides that measure  
(Boltzmann)...

$$S \equiv k_B \ln W$$

Molecular Entropy

Boltzmann Constant

Number of microscopic ways in which a particular outcome (macroscopic state) can be attained

**Criterion for Spontaneity:**

For Avogadro number's  
of molecules...

$$S = \underbrace{(N_{\text{Avogadro}} k_B)}_{R \text{ (gas constant)}} \ln W$$

Therefore: the most probable  
outcome maximizes entropy  
of isolated systems

$$\Delta S > 0 \text{ (spontaneous)}$$

$$\Delta S < 0 \text{ (non-spontaneous)}$$

# Unfavorable conformation entropy for protein folding

for the folded state:  $\sim 1$  conformation

$$S_{\text{folded}} = R \ln(1) = 0$$

for the unfolded state:  $x$  is the number of flexible points per residue and  $z$  is the number of possible orientations of equal energy at each point.

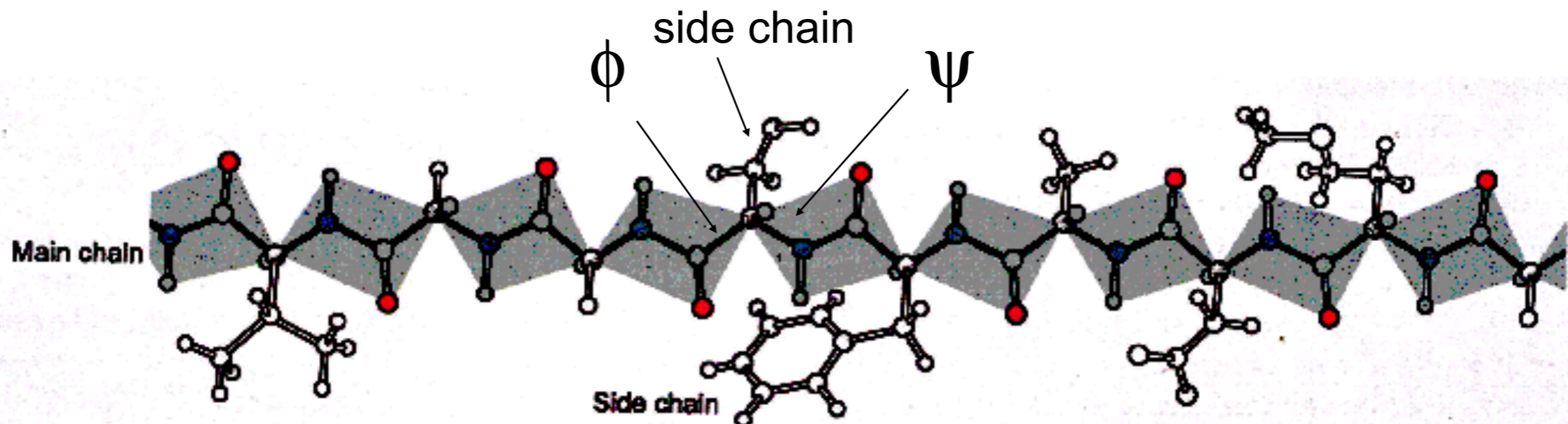
$$S_{\text{unfolded}} = R \ln(z^x)$$

# Estimating the unfavorable conformational entropy for protein folding

$$S_{\text{conf}} = R \ln(z^x) \quad \Delta S_{\text{fold}} = R \ln\left(\frac{W_{\text{unfolded}}}{W_{\text{folded}}}\right)$$

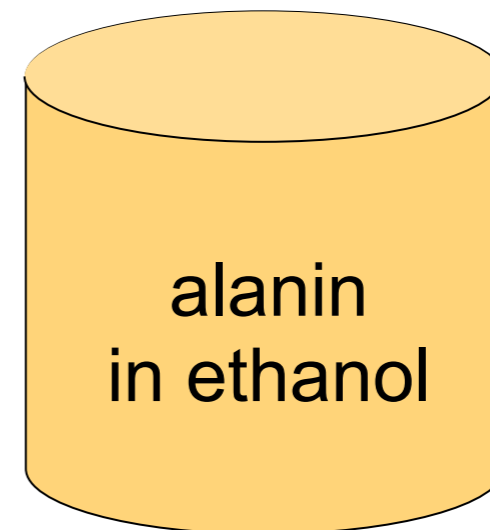
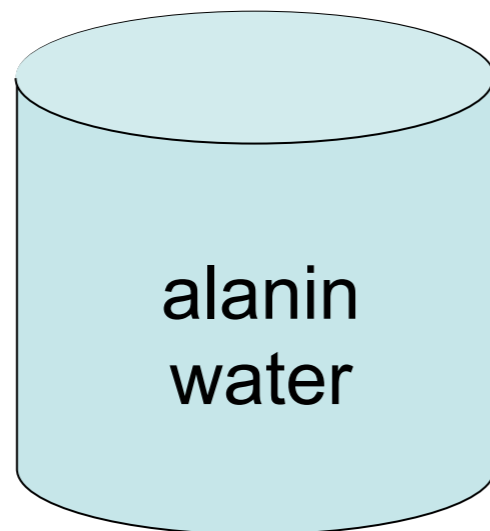
Tanford: For three flexible positions ( $\phi$ ,  $\psi$ , side chain) with two possible orientations each we have  $2^3$  conformations per residue:

- $T\Delta S = 1.2$  kcal/mol or  $5.0$  kJ/mol (estimate Tanford 1962)  
 $1.7$  kcal/mol (Spolar and Record, Science)





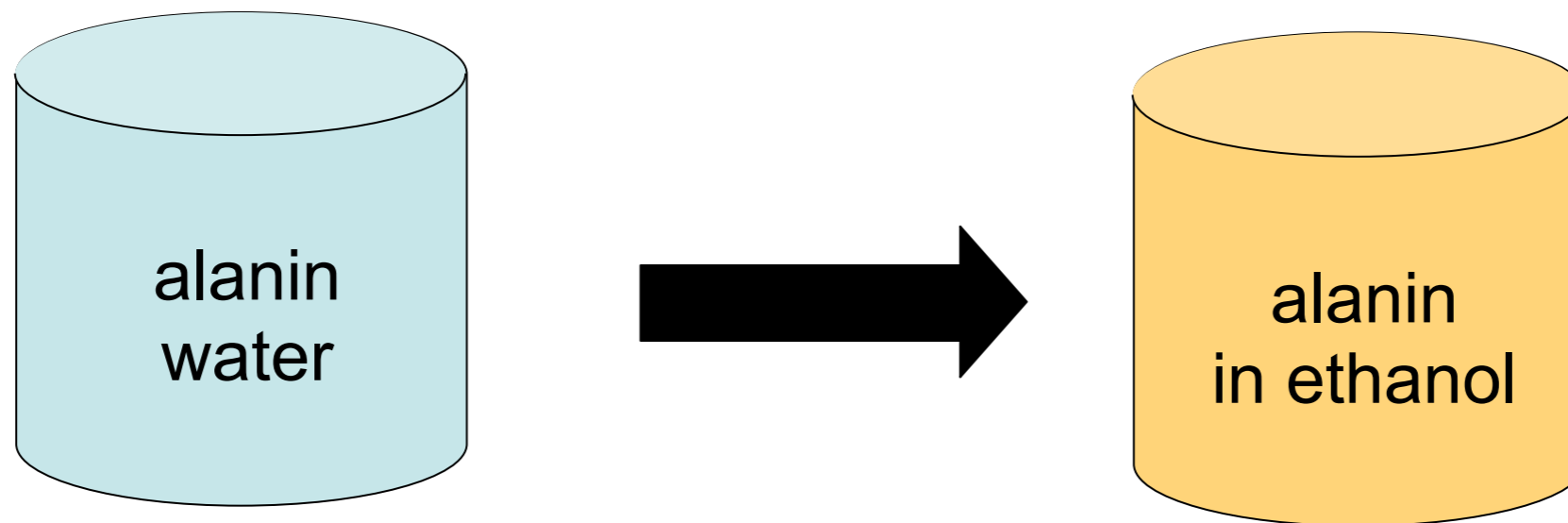
Measure solubility of amino acid in ethanol (= inside the folded protein) and in water (= unfolded state or at the protein surface)



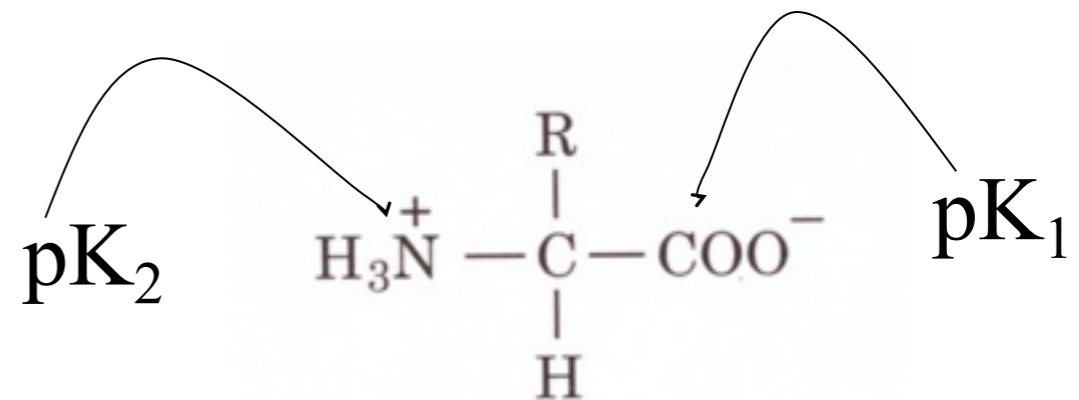
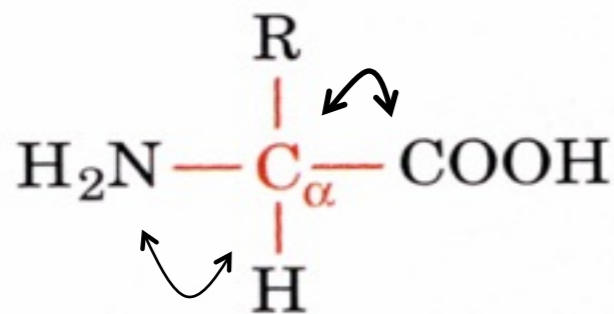
partition coefficient  $K_D = \frac{\text{solubility } (Alanin_{EtOH})}{\text{solubility } (Alanin_{H_2O})}$

Calculate the free energy from transferring an amino acid from water to ethanol

$$\Delta G_{\text{tr}} = -RT \ln \left( \frac{N_{\text{EtOH}}}{N_{\text{H}_2\text{O}}} \right) = -RT \ln(K_D)$$



Free amino acids carry a positive and a negative charge that is not present in the peptide chain



$\alpha$  amino acids because of the  $\alpha$  carboxylic and  $\alpha$  amino groups  
 $\text{pK}_1$  and  $\text{pK}_2$  respectively  $\text{pK}_R$  is for R group  $\text{pK}$ 's

$\text{pK}_1 \approx 2.2$  while  $\text{pK}_2 \approx 9.4$

In the physiological pH range, both carboxylic and amino groups are completely ionized



[CONTRIBUTION FROM THE DEPARTMENT OF BIOCHEMISTRY, DUKE UNIVERSITY, MEDICAL CENTER, DURHAM,  
NORTH CAROLINA]

## Contribution of Hydrophobic Interactions to the Stability of the Globular Conformation of Proteins

BY CHARLES TANFORD

RECEIVED APRIL 9, 1962

TABLE I<sup>a</sup>

FREE ENERGY CHANGE IN CALORIES PER MOLE FOR TRANSFER FROM ETHANOL TO WATER AT 25°

Tanford 1962

	$\Delta F_t$ , whole molecule	$\Delta f_t$ , side chain contribution	
Non-polar side chains			
Glycine	-4630	0	
Alanine	-3900	+ 730	
Valine	-2940	+1690	
Leucine	-2210	+2420	
Isoleucine	-1690 <sup>b</sup>	+2970 <sup>b</sup>	
Phenylalanine	-1980	+2650	
Proline	-2060 <sup>c</sup>	+2600 <sup>c</sup>	
Other side chains			
Methionine	-3330	+1300	
Tyrosine	- 930 <sup>d</sup>	+2870 <sup>d</sup>	
Threonine	-4190	+ 440	
Serine	-4590	+ 40	
Asparagine	-4640	- 10	
Glutamine	-4730	- 100	
Aspartic acid <sup>e</sup>	-4090	+ 540	uncharged
Glutamic acid <sup>e</sup>	-4080	+ 550	uncharged

# Burying a charged amino acid in the interior (Born expression)

$$W_B = \frac{q^2}{4\pi\epsilon_0 r} \left( \frac{1}{\epsilon_1} - \frac{1}{\epsilon_2} \right)$$

$W_B$  is the free energy of transfer in moving a charged body from a region with a relative dielectric constant  $\epsilon_2$  to a medium with a relative dielectric constant  $\epsilon_1$ . The parameter  $r$  is the radius of the charge.

$q$  (charge of an electron) =  $1.60 \times 10^{-19}$  C

dielectric constant in vacuum  $\epsilon_0 = 8.85 \times 10^{-12}$  C<sup>2</sup> J<sup>-1</sup> m<sup>-1</sup>

$r$  is ionic radius, with is typically 1-2 Å

Sharp, K.A. and Honig, B. (1990) Electrostatic interactions in macromolecules: theory and applications. *Annu Rev Biophys Biophys Chem*, 19, 301-332.

TABLE III

CONTRIBUTION OF THE MOST IMPORTANT HYDROPHOBIC INTERACTIONS TO THE FREE ENERGY OF UNFOLDING AT 25°

Side chain	$\Delta f_u$ per side chain, cal./mole	Number present in			
		myo-globin <sup>a</sup>	$\beta$ -lacto-globulin <sup>b</sup>	ribo-nuclease <sup>c</sup>	
Tryptophan	3000	2	2	0	
Isoleucine	2970	9	10	3	
Tyrosine	2870	3	4	6	
Phenylalanine	2650	6	4	3	
Proline	2600	4	8	4	
Leucine	2420	18	22	2	
Valine	1690	8	10	9	
Lysine	1500	19	15	10	
Methionine	1300	2	4	4	
Alanine	730	17	14	11	
Arginine	730	4	3	4	
Threonine	440	5	8	10	
Total number of residues		153	162	124	
$-T\Delta S_{conf}$ , kcal./mole		-184	-194	-149	conformation entropy
$\Sigma\Delta f_u$ , kcal./mole		+173	+192	+100	hydrophobic effect

# What did we learn today?

- Using numbers for cellular processes
- Energy scales
  - $k_B T$  (single molecule) or  $RT$  (per mole) reference scale
  - Equilibrium free energy  $\Delta G$  or equilibrium binding constant
  - Entropy from changes of number of states
- Macromolecular interactions in aqueous solutions
  - non-covalent interactions
  - hydrophobic effect