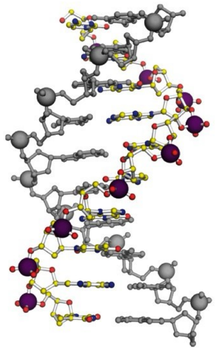
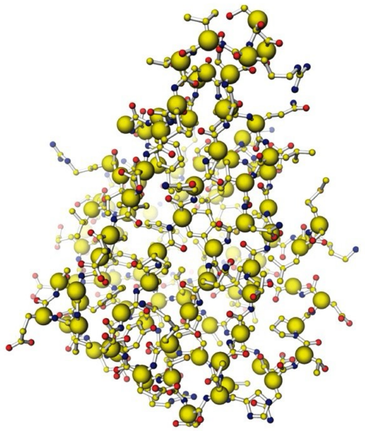
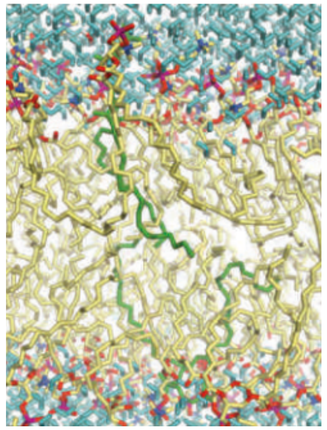


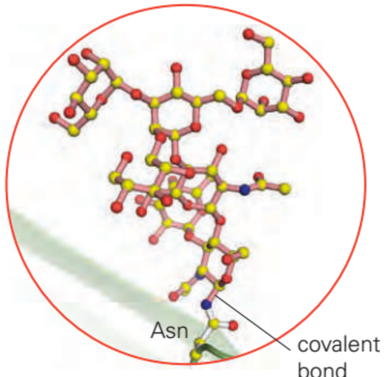
Chimie Biologique I
Biological Chemistry I
BIO-212



Lecture 10

Matteo Dal Peraro





where are we so far …

molecular interactions
nucleic acids

proteins

structural biology

Bioenergetics

Binding and kinetics

lipids, glycans

Protein visualization

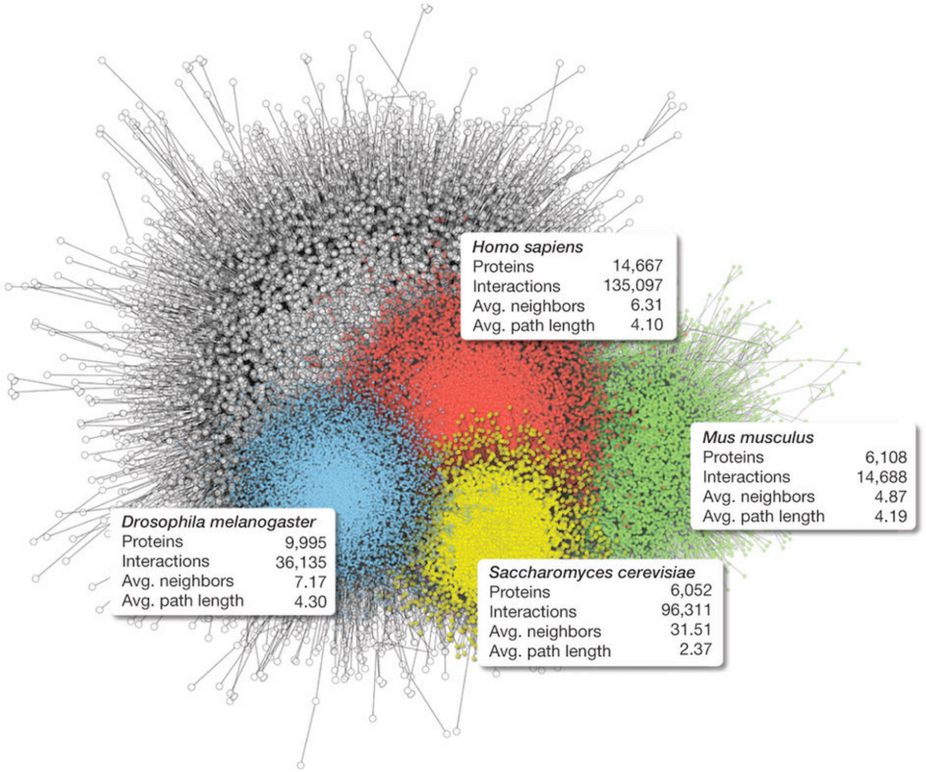
Protein production and analysis

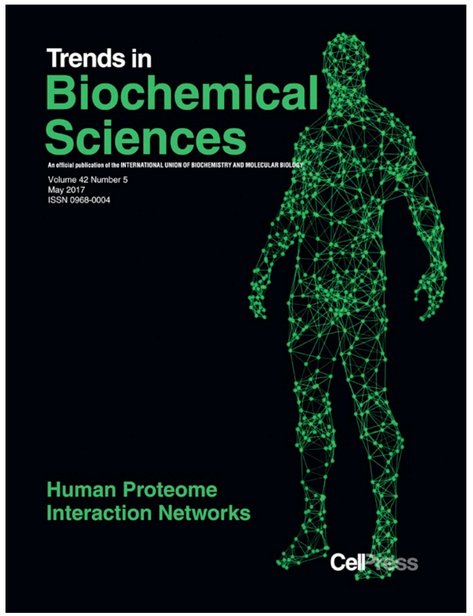
- how biomolecules interact
- how biomolecules are structured
- how we reveal their structure
- how we can visualise them in 3D
- how we produce them for analysis in vitro
- **how we account for energy in biochemistry**
- how to characterise molecular binding

Why binding is that important ?

Because all cellular processes are controlled by the way that different molecules interact with each other, for example recognition of proper substrates by enzymes, the transmission of cellular signals, the recognition of one cell by another, the control of transcription and translation, and the fidelity of DNA replication.

If one maps protein-protein interactions across different organisms, this is what you get:





-Dots are proteins, edges represent interactions

Molecular Interactions in Biomolecules

Atomic Interactions

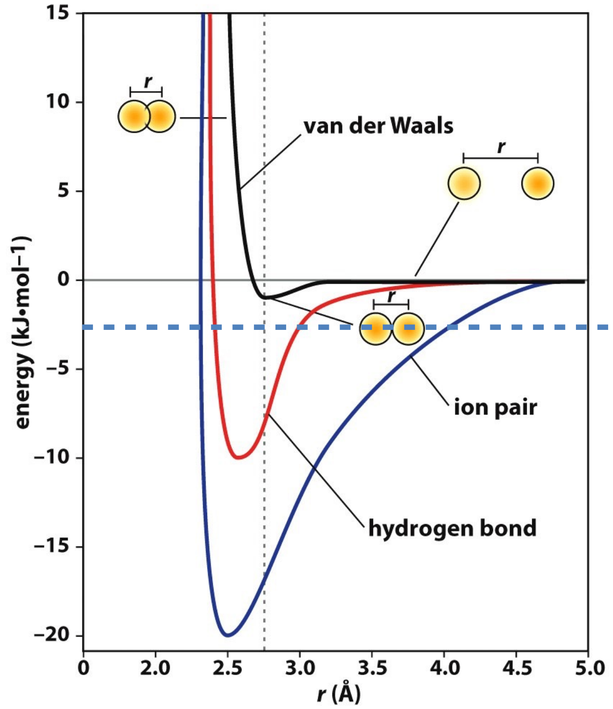
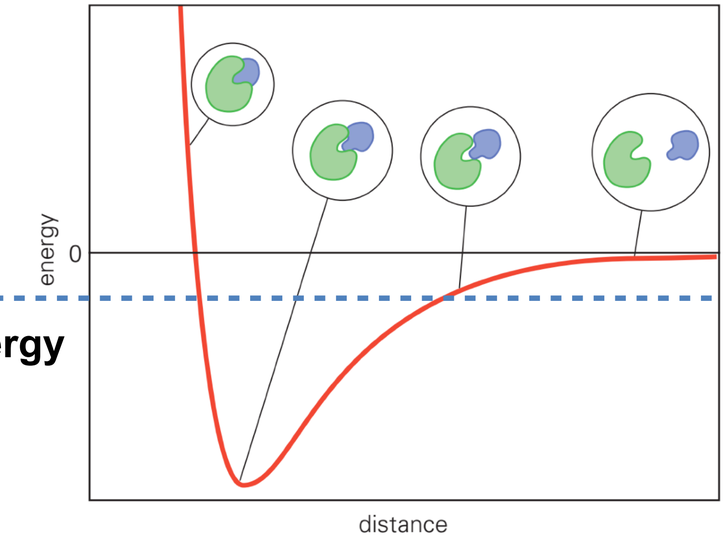


Figure 1.12 The Molecules of Life (© Garland Science 2013)

Macromolecular interactions



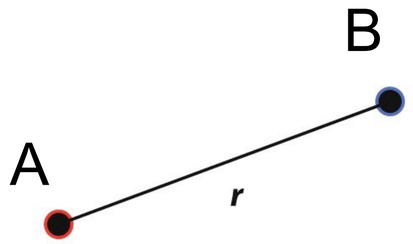
thermal energy — **2.5 kJ/mol**

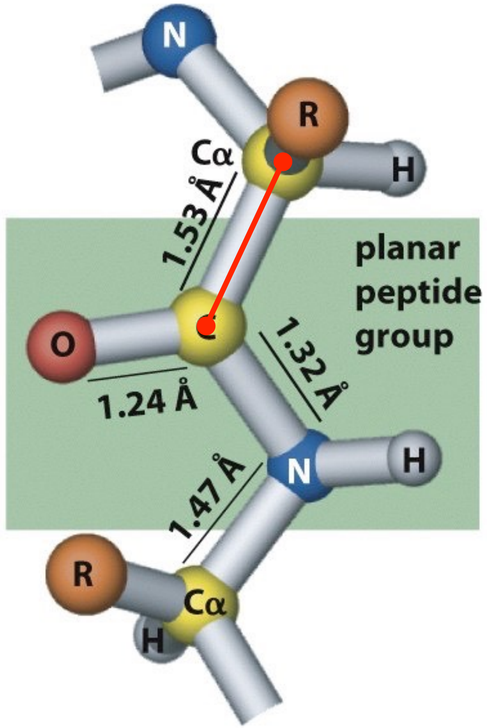
$$U(r) \sim \sum_{i<j}^{N} U(r_{ij}) \qquad \text{(addictive)}$$

$$U(r) = \frac{1}{4\pi\epsilon_0}\frac{1}{D}\frac{q_1 q_2}{r} \qquad \text{(for ex Coulomb)}$$

We can now understand some of the energetic principles that govern the interactions between macromolecules and determine how complex biological processes occur

Potential energy for a chemical bond





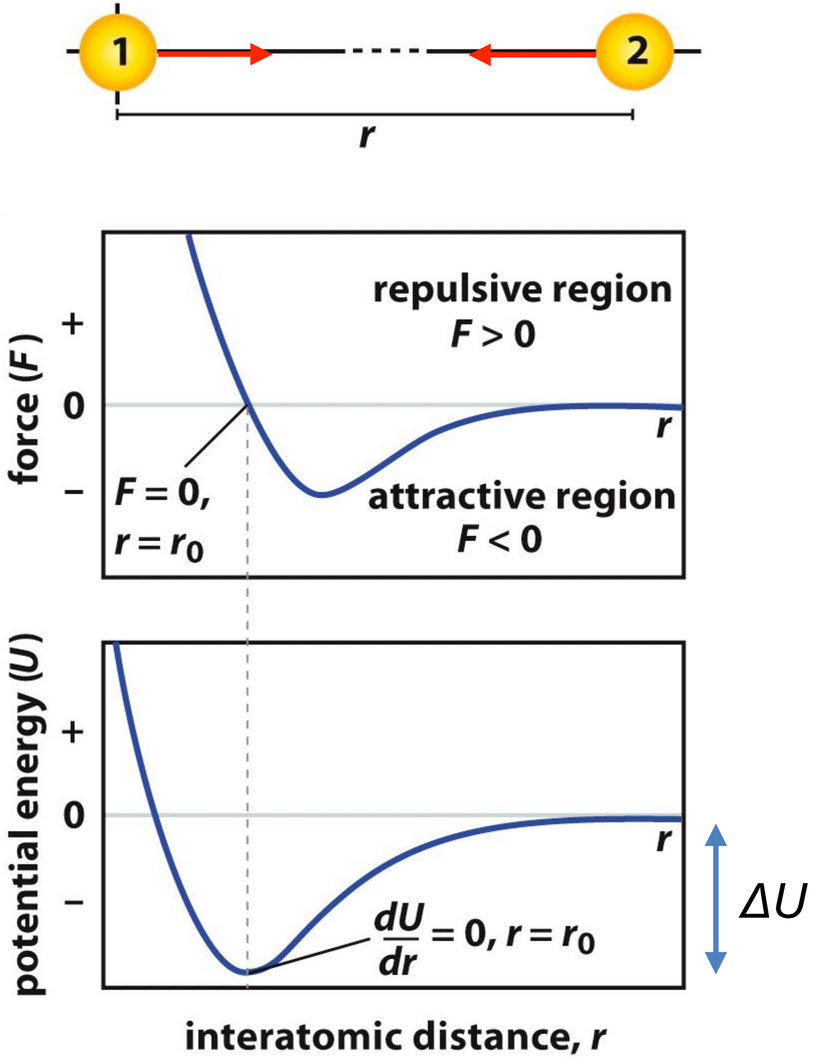
how such potential, *U(r)* would look like?

➡ at very large distance *r* ?

➡ at short distance?

➡ at equilibrium distance, r_0 ?

Potential energy and forces



- potential energy is the energy stored in the system
- a simple case is when U depends on distance ***r*** only
- *U(r)* can have complex forms, from which forces on each atom are derived:

$$F(r) = -\frac{dU(r)}{dr}$$

$$U(r) = -\int_{\infty}^{r} F(\tilde{r})d\tilde{r}$$

Covalent bond potential energy function

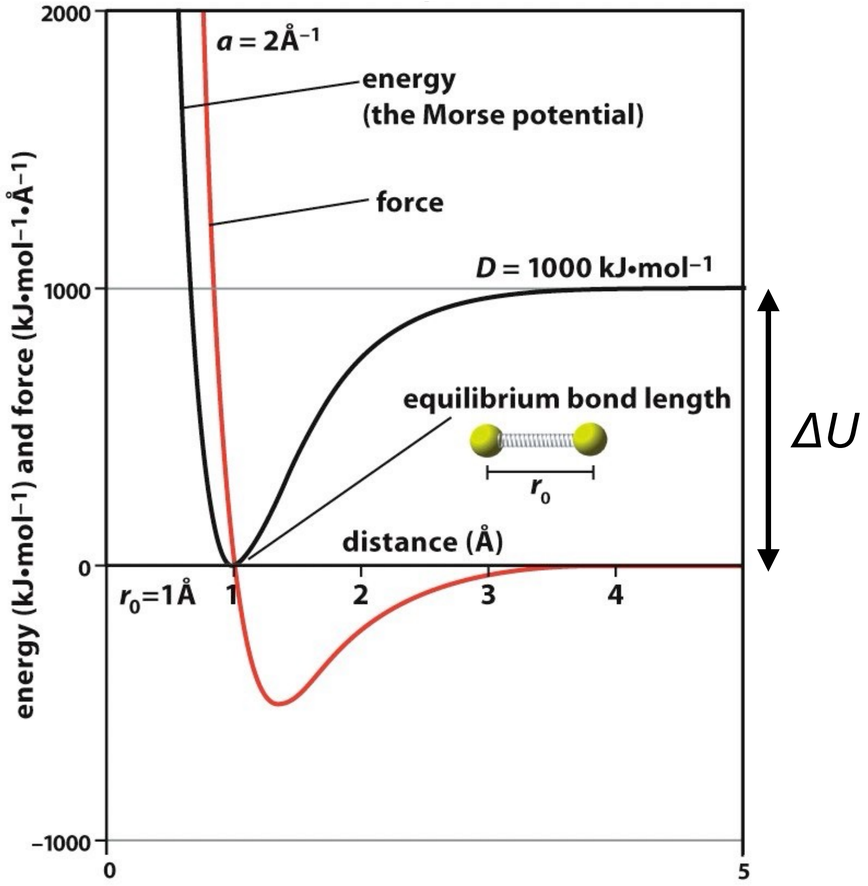
$$U_{bond}(r) = D[1 - e^{-a(r-r_0)}]^2$$

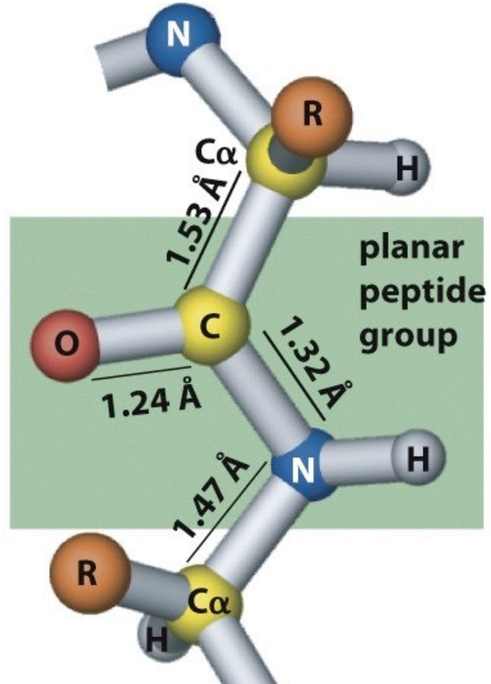
Morse potential

D: dissociation energy (related to H of bond formation)

r₀: equilibrium bond distance

a: force constant





Energetics of molecular interactions

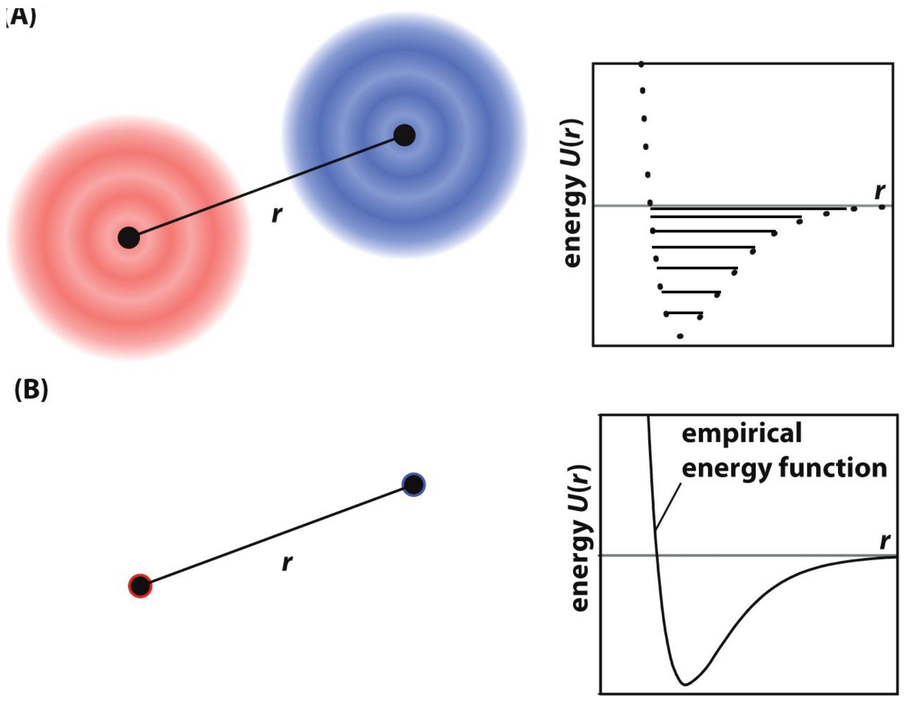


Figure 6.19 The Molecules of Life (© Garland Science 2013)

- nuclei and electrons are described by **quantum mechanics** by the *Schrodinger equation*

$$H\Psi(\mathbf{r},\mathbf{R}) = E\Psi(\mathbf{r},\mathbf{R})$$

- QM calculations more exact but computationally demanding

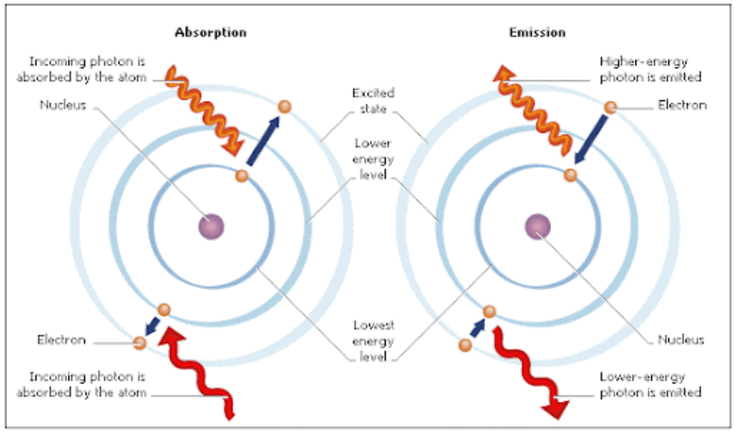
- simplified *"ball-and-spring"* models following the laws of classical physics are good approximations

- Empirical energy potential *U(R)* are applicable to all classes of macromolecules

Energy Levels

- We need to think about molecules as entities that despite their homogenous composition populate multiple energetic levels, this is due to the quantum mechanical nature of matter - biological matter included

- An example for some simple molecules:



high energy

low energy

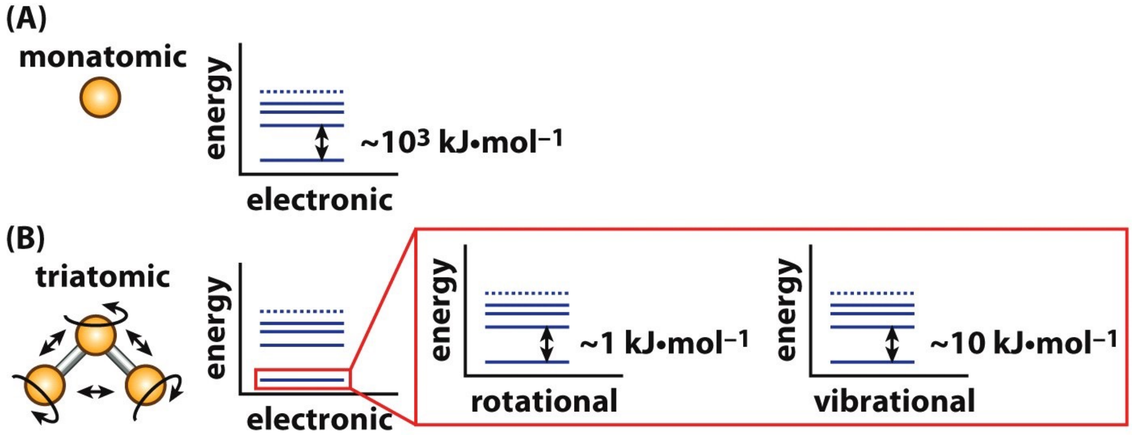
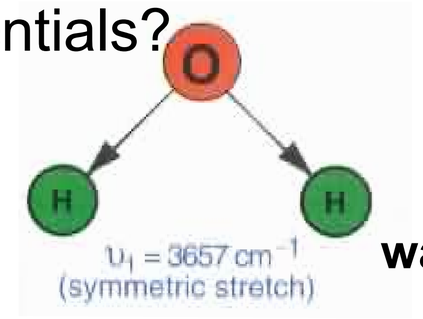


Figure 6.15 The Molecules of Life (© Garland Science 2013)

- Imagine how the energy levels would look for molecules with thousands of atoms

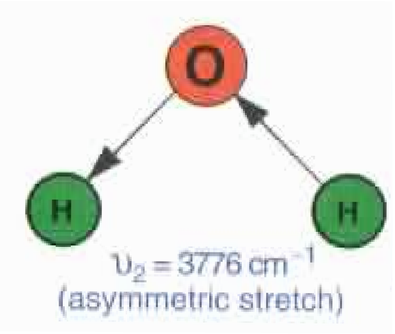
Vibrational modes

how do we derive the parameters for these potentials?

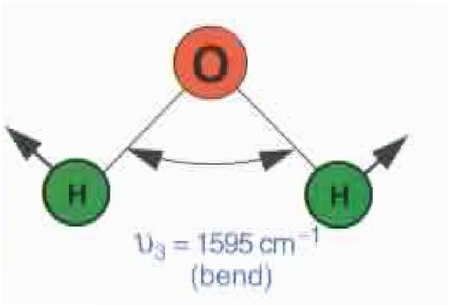


wavenumber [cm^{-1}]

- vibrational spectra give information about **bonded** interactions (3N-6 modes, with N number of atoms)



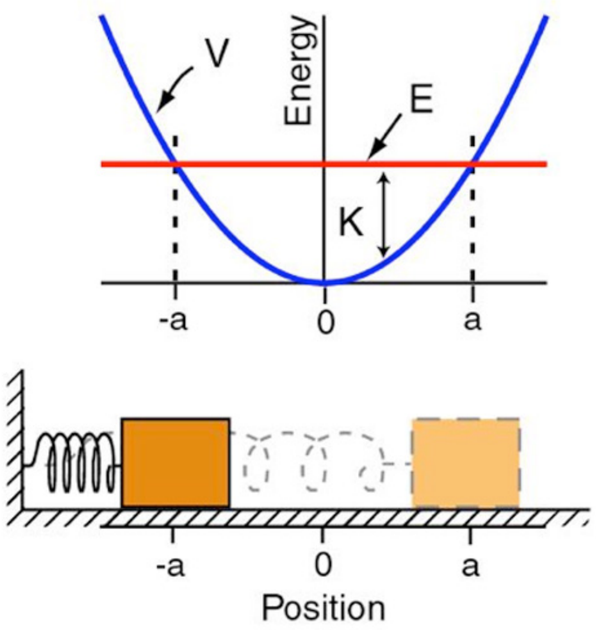
- derived from infrared spectroscopy (IR: 300-3000 cm^{-1}), Raman spectroscopy or QM calculations

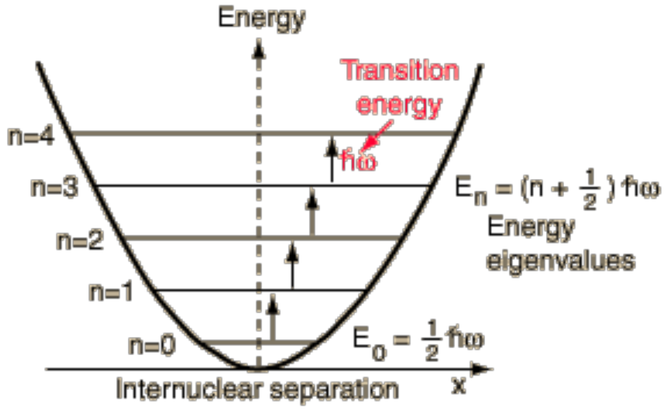


- the higher the wavenumber the more difficult the deformation, but also depends on the environment and mass atoms

water vibrational modes

Vibrational modes as harmonic oscillators





- vibrational degrees of freedom can be modeled as harmonic oscillators, following **Hooke's law**

$$\boldsymbol{F} = -k\,\boldsymbol{x}$$

- the potential energy associated is:

$$U(\boldsymbol{x}) = k/2\;\boldsymbol{x}^2$$

- if one solves the Schrodinger equation with this potential, the solutions are:

$E_\upsilon = (\mathrm{n}+1/2)\,h\nu\;;\; n = 0, 1, 2\;\ldots$ quantum number

$\nu = 1/2\pi\,(k/m)^{1/2}$ **wavenumber** [cm^{-1}]

$\omega = \sqrt{\dfrac{k}{m}} = 2\pi\,\nu$ **angular frequency**

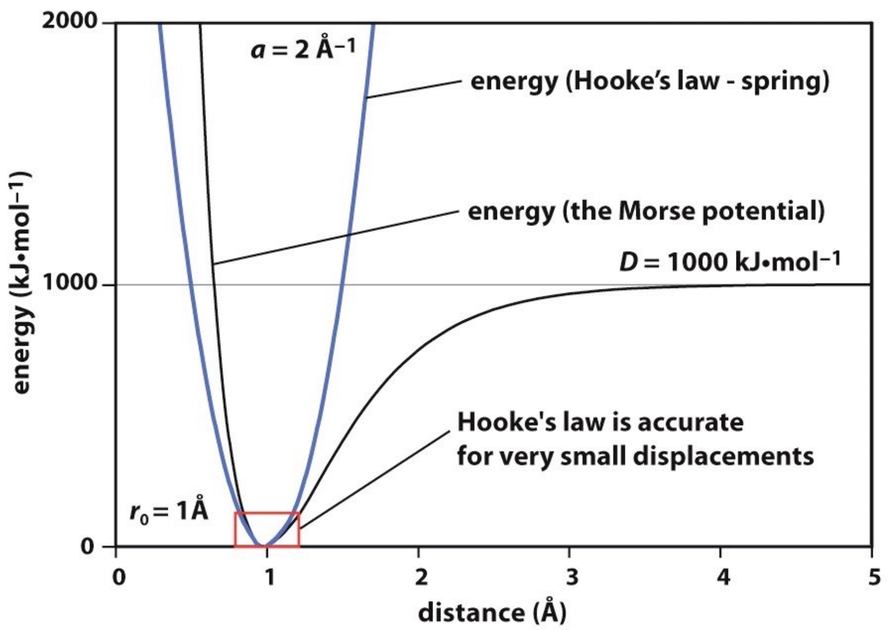
Harmonic approx for covalent bonds

$$U_{bond}(r) = \frac{k_b}{2}(r - r_0)^2$$

- equilibrium bond lengths are obtained from X-ray or QM calculations (r_0)
- deviation from reference value (r_0) is modeled as an **harmonic** potential
- k_b is the **force constant**

$$k_b = (2\pi\nu)^2\mu = \omega^2\mu$$
$$\mu = m_1 m_2/(m_1 + m_2)$$

- reasonable approximation for small deviation from r_0 (~0.1 Å), dissociation for larger values

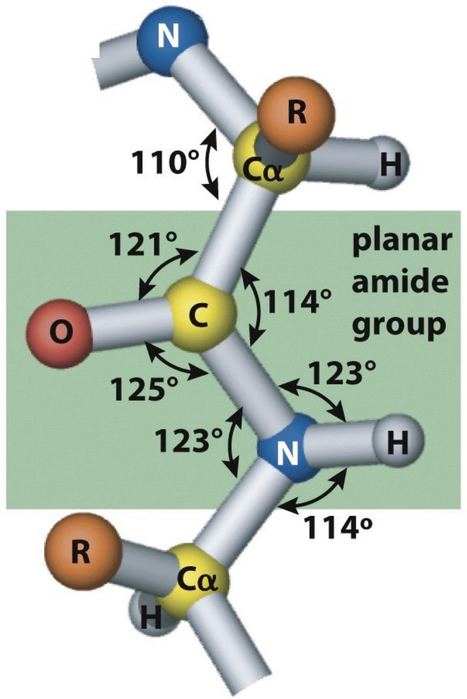


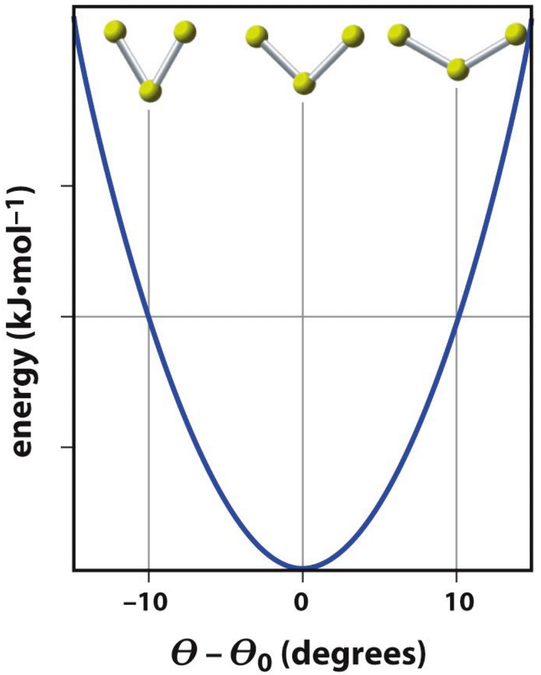
Bond	l_0 (Å)	k (kcal mol^{-1} Å^{-2})
Csp^3-Csp^3	1.523	317
Csp^3-Csp^2	1.497	317
$Csp^2=Csp^2$	1.337	690
$Csp^2=O$	1.208	777
Csp^3-Nsp^3	1.438	367
C–N (amide)	1.345	719

Bending potential energy functions

- angles depend on hybridization of electronic **orbitals**
- rule of thumb: sp:~180°; sp^2:~120°; sp^3:~109.5°

$$U_{angle}(\theta) = \frac{k_\theta}{2}(\theta - \theta_0)^2$$

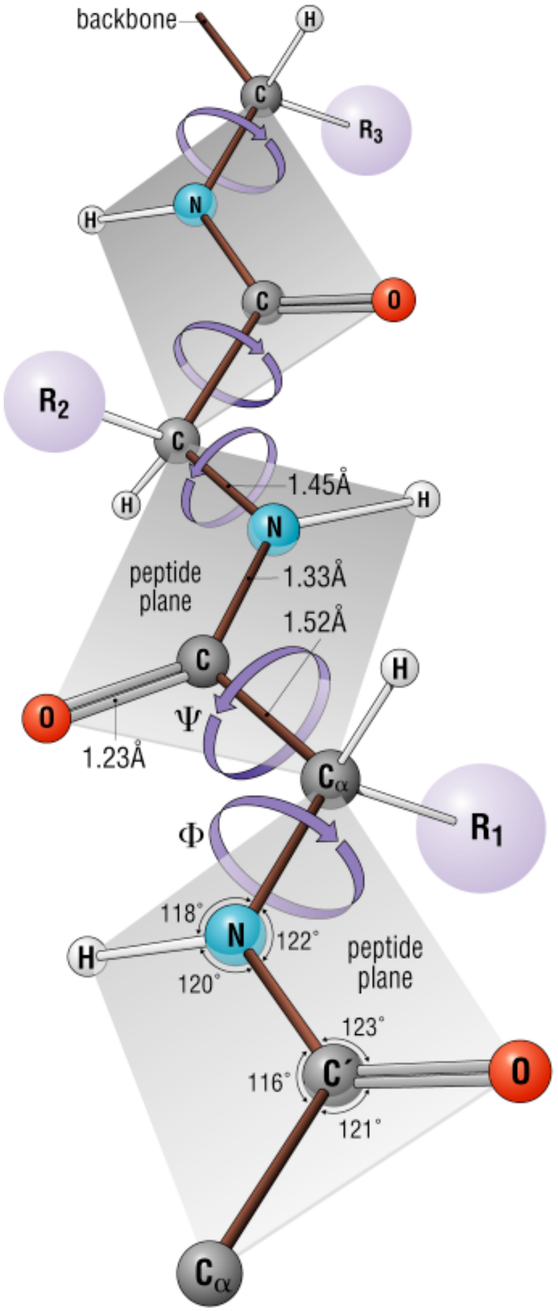


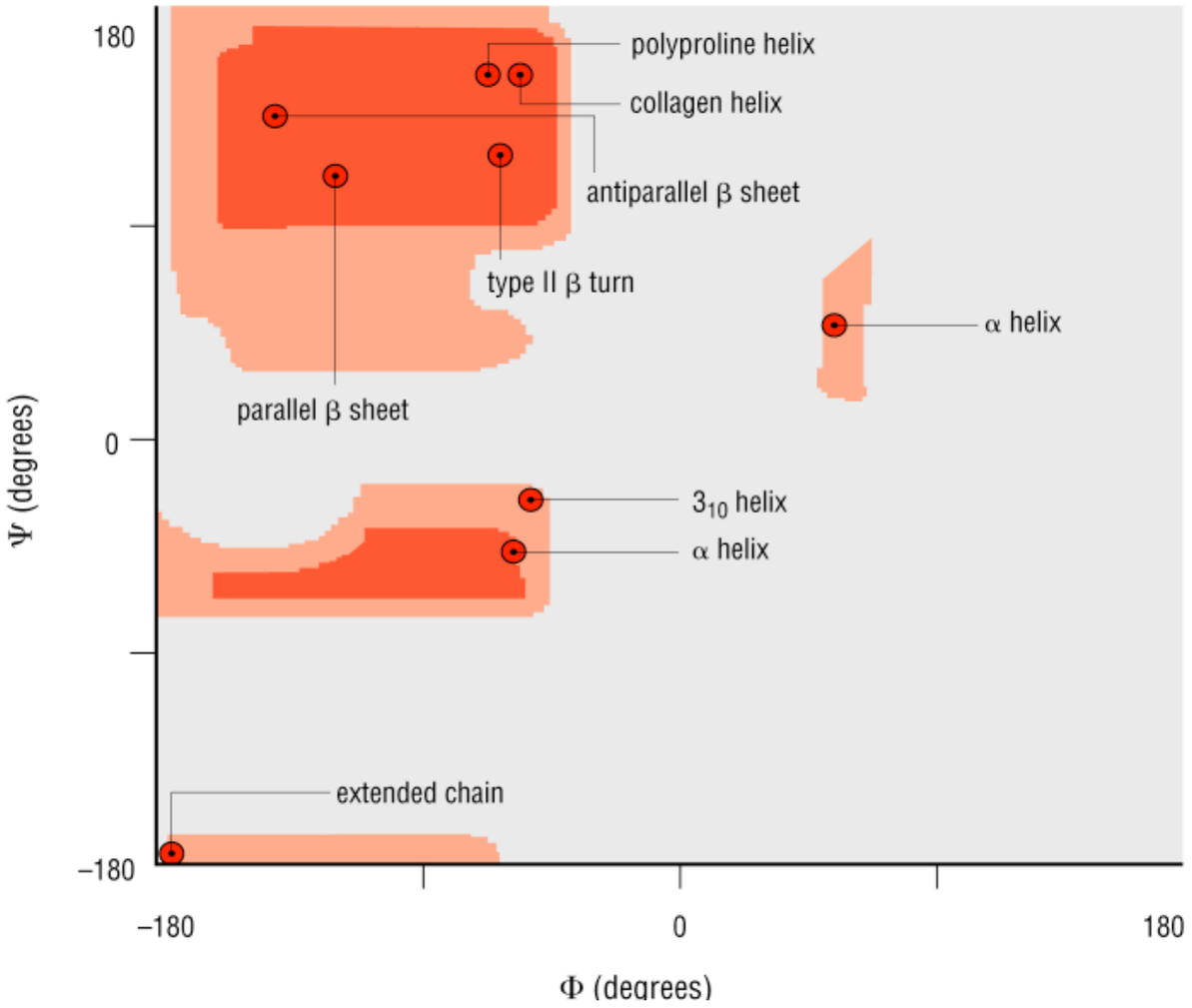


H–O–H, H–N–H bend	1600
H–C–H bend	1500
H–C–H scissor	1400
H–C–H rock	1250
H–C–H wag	1200
H–S–H bend	1200
O–C=O bend	600
C–C=O bend	500
S–S–C bend	300
C=C torsion	1000
C–O torsion	300–600
C–C torsion	300
C–S torsion	200

Angle	θ_0	k (kcal mol^{-1}deg^{-1})
Csp^3–Csp^3–Csp^3	109.47	0.0099
Csp^3–Csp^3–H	109.47	0.0079
H–Csp^3–H	109.47	0.0070
Csp^3–Csp^2–Csp^3	117.2	0.0099
Csp^3–Csp^2=Csp^2	121.4	0.0121
Csp^3–Csp^2=O	122.5	0.0101

Secondary structure geometry





Ramachandran plot

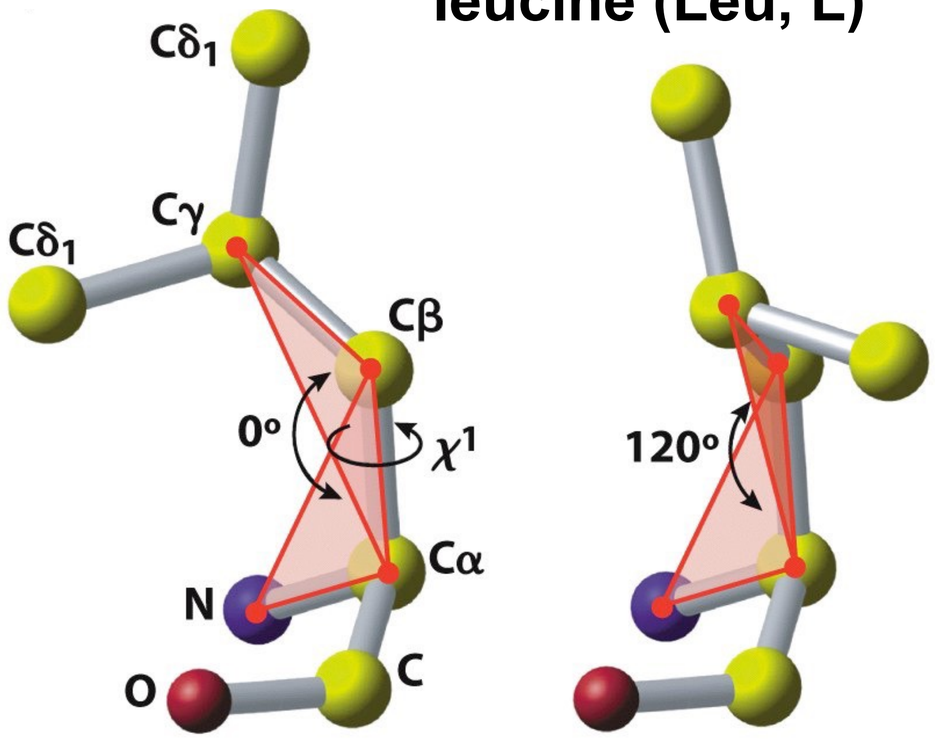
Procheck: http://www.ebi.ac.uk/pdbsum/

Torsional potential energy functions

- barriers for torsional rotations need a periodic potential term
- n is the **periodicity**, k_φ is the barrier height, δ is called **phase** $(0, \pi)$
- Fourier series of torsional potentials (n=1,2,3,4)
- parameters are again derived from **spectroscopy** or **quantum mechanics** calculations

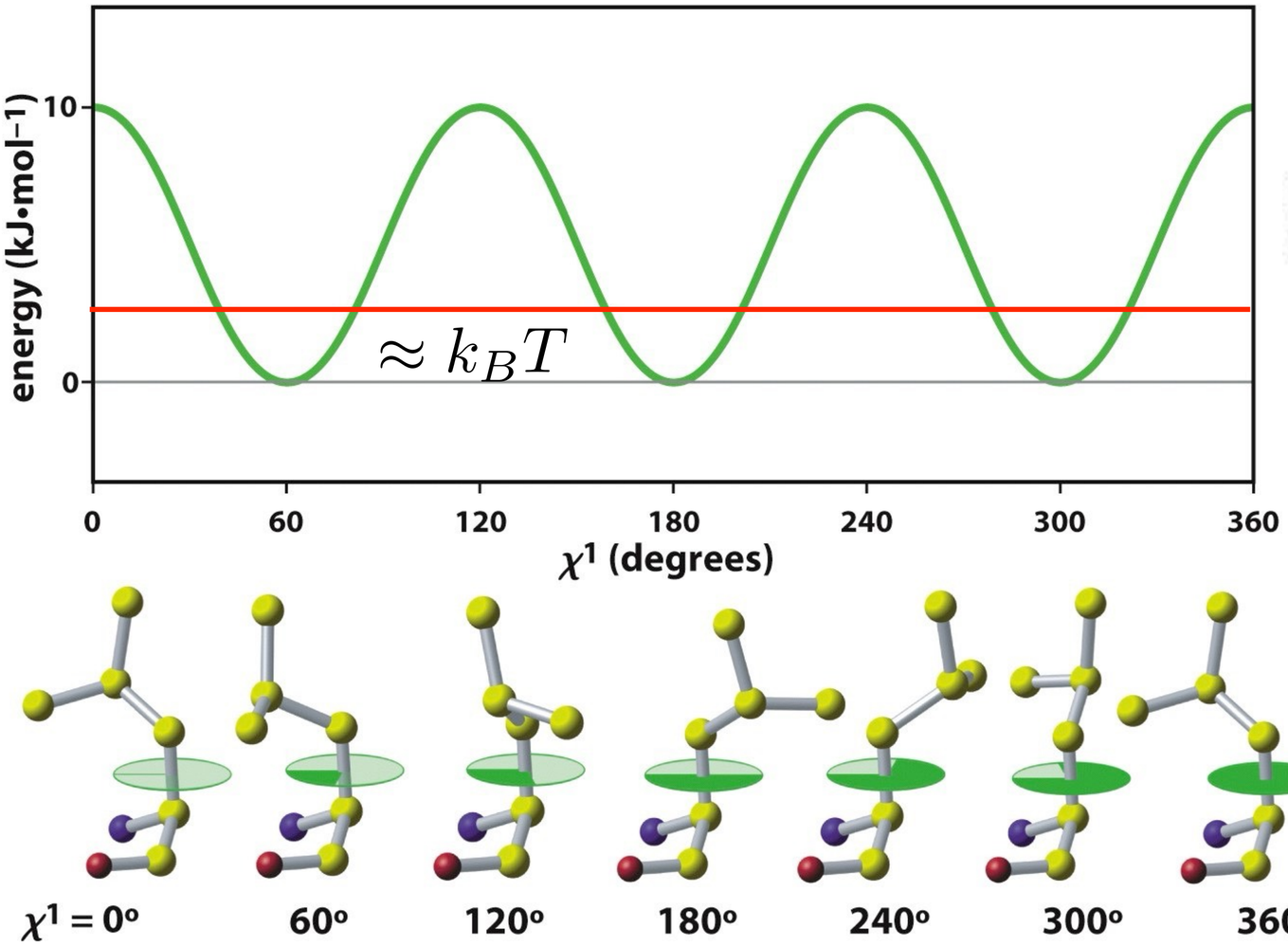
$$U_{torsion}(\phi) = \sum_n \frac{k_{\phi,n}}{2}[1 + cos(n\phi - \delta)]$$

leucine (Leu, L)

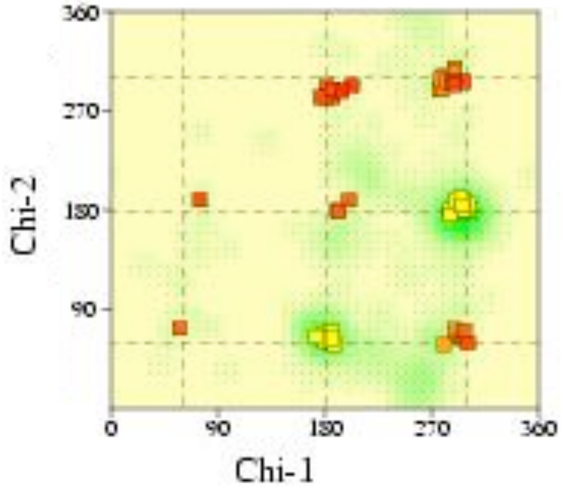


$$\chi^1 = \phi$$

$$U(\phi) = \frac{k_\phi}{2}(1 + cos3\phi)$$



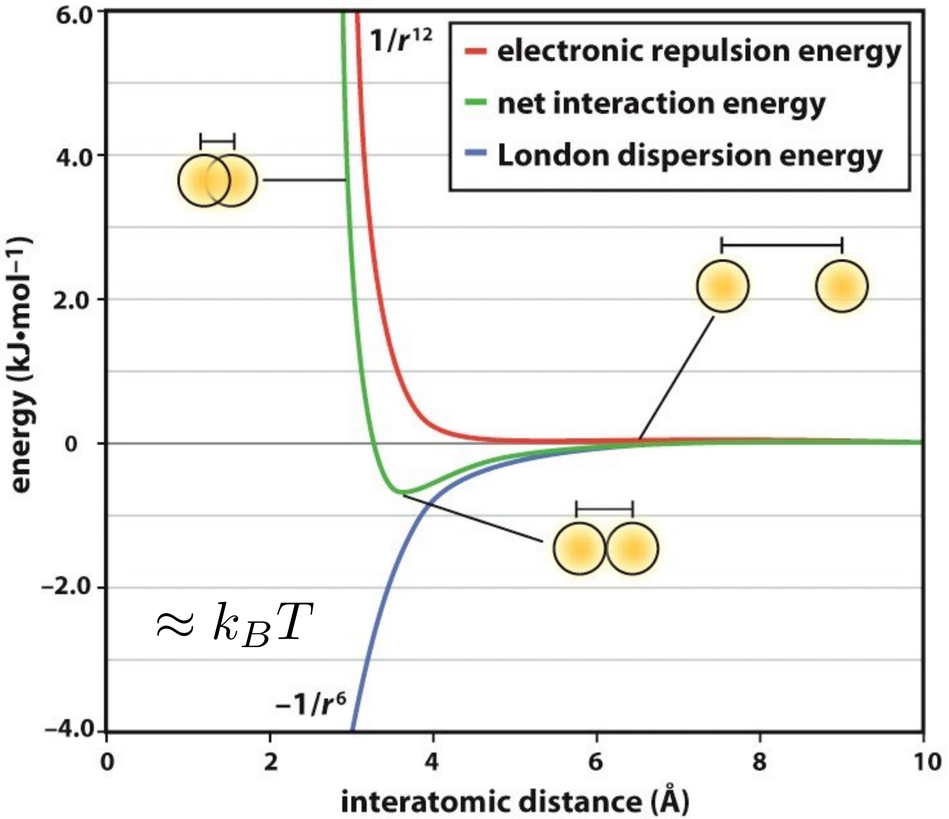
leucine (Leu, L)



Procheck: http://www.ebi.ac.uk/pdbsum/

van der Waals interactions

London forces are at the basis of the attractive interactions and coupled with repulsive interactions given by the Pauli principle define **vdW interactions** which are commonly modeled by **Lennard-Jones** (12-6) energy potentials:

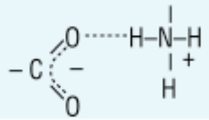
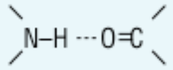
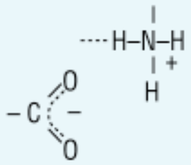
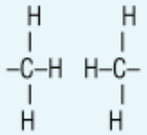


$$U_{VdW} = \sum_{i>j}^{N} \epsilon \left[\left(\frac{r_0}{r_{ij}} \right)^{12} - 2 \left(\frac{r_0}{r_{ij}} \right)^{6} \right]$$

$$U_{VdW} = \sum_{i>j}^{N} \frac{A}{r_{ij}^{12}} - \frac{C}{r_{ij}^{6}}$$

r_0 = eq. interatomic separation,
$r_0/2$ = vdW radius if *i* and *j* are the same atom type

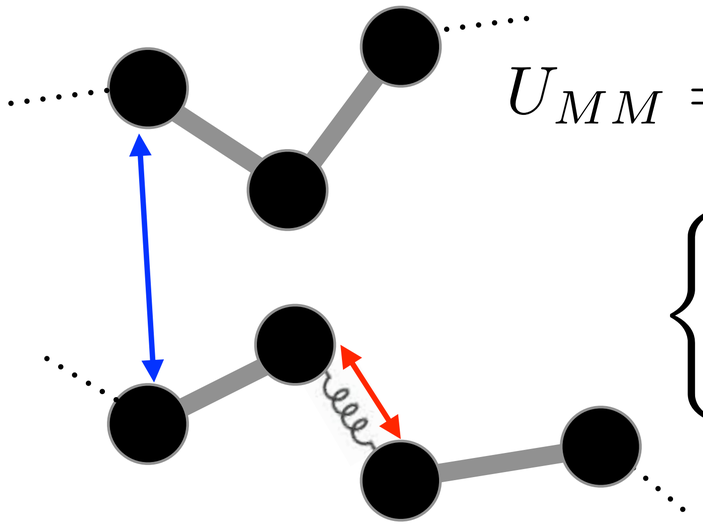
Molecular interactions in biomolecules (bonded and non-bonded)

Interaction	Example	Distance dependence	Typical distance	Free energy (bond dissociation enthalpies for the covalent bonds)
Covalent bond	–C_α–C–	-	1.5 Å	356 kJ/mole (610 kJ/mole for a C=C bond)
Disulfide bond	–Cys–S–S–Cys–	-	2.2 Å	167 kJ/mole
Salt bridge		Donor (here N), and acceptor (here O) atoms <3.5 Å	2.8 Å	12.5–17 kJ/mole; may be as high as 30 kJ/mole for fully or partially buried salt bridges (see text), less if the salt bridge is external
Hydrogen bond		Donor (here N), and acceptor (here O) atoms <3.5 Å	3.0 Å	2–6 kJ/mole in water; 12.5–21 kJ/mole if either donor or acceptor is charged
Long-range electrostatic interaction		Depends on dielectric constant of medium. Screened by water. 1/r dependence	Variable	Depends on distance and environment. Can be very strong in nonpolar region but very weak in water
Van der Waals interaction		Short range. Falls off rapidly beyond 4 Å separation. $1/r^6$ dependence	3.5 Å	4 kJ/mole (4–17 in protein interior) depending on the size of the group (for comparison, the average thermal energy of molecules at room temperature is 2.5 kJ/mole)

k_BT(300 K) = 0.6 kcal/mole - 2.5 kJ/mole - 4.1 pN nm - 4.1 * 10^{-21} J
(1 kcal/mol = 4.184 kJ/mole)

Molecular mechanics potentials

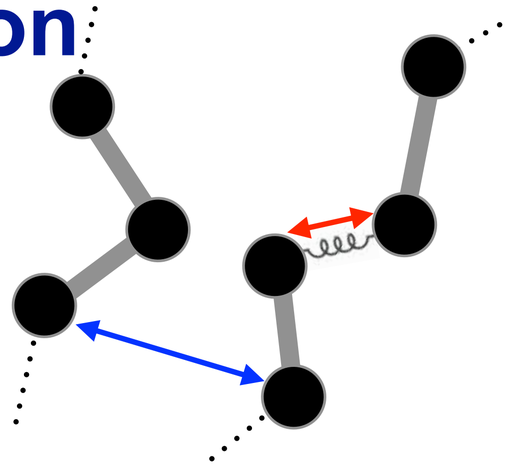
- **molecular mechanics** (MM) potential energy gives minimum-energy conformation of a molecule
- based on **physics**, but uses simplified “ball-and-spring” models (**classical** physics), which mask the quantum nature (*Schrodinger equation*)
- are **empirical**, i.e. calibrated to describe the quantum nature of chemical bonds and short-range interactions



$$U_{MM} = U_{bonded} + U_{non-bonded}$$

$$\begin{cases} U_{bonded} = U_{bond} + U_{angle} + U_{torsion} \\ U_{non-bonded} = U_{electrostatics} + U_{VdW} \end{cases}$$

Empirical potential energy function



The force field:

- large number of parameters fitted to represent experimental data or QM calculated quantities
- “trial and error” or least-squares fitting methods to converge to a consistent set of parameters
- assumption that parameters can be transferable to different contexts (specialized vs. generalized FF)

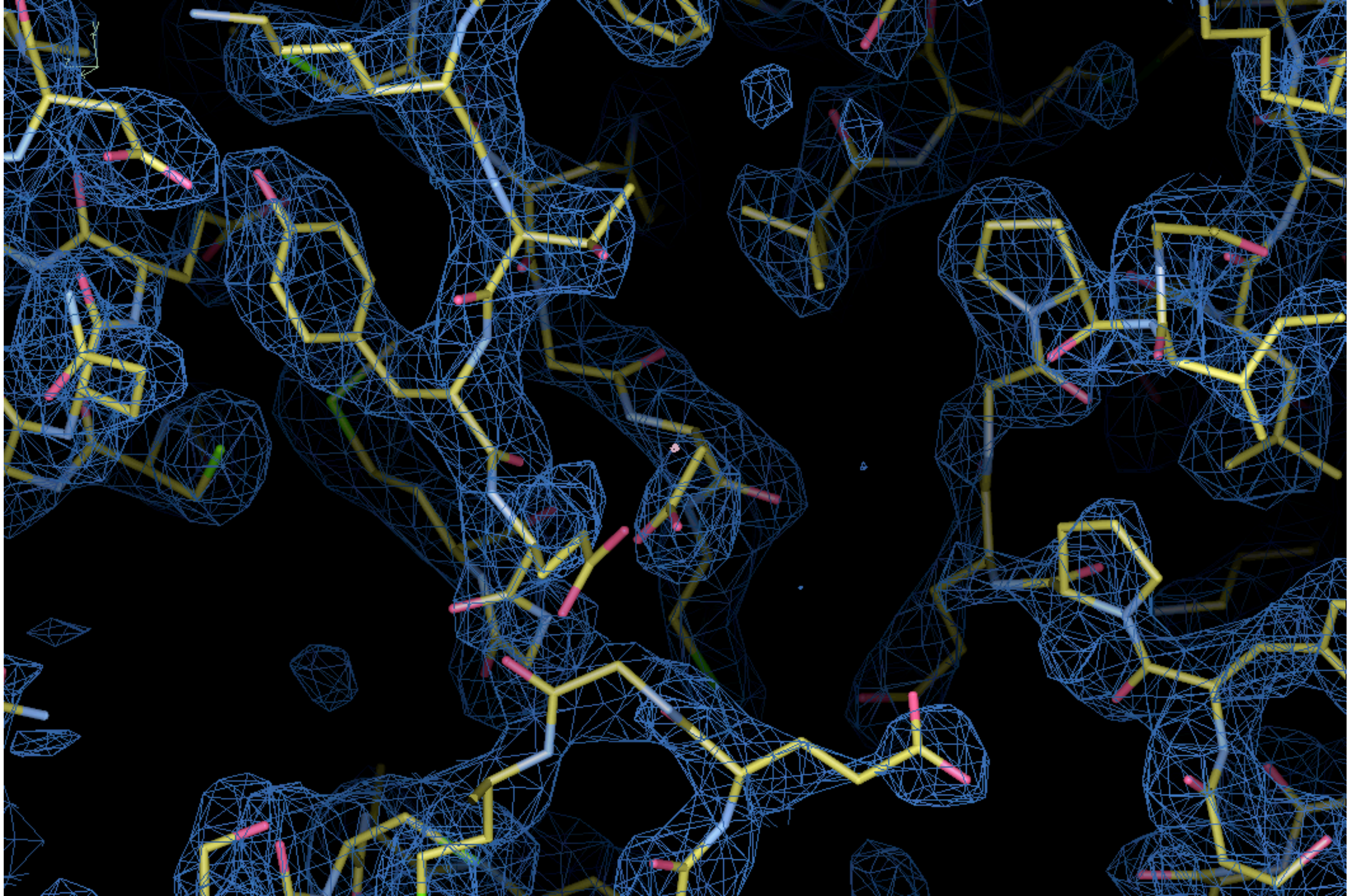
$$U_{MM} = U_{bonded} + U_{non-bonded}$$

$$\sum_{\text{All Bonds}} \tfrac{1}{2}K_b(b-b_o)^2 + \sum_{\text{All Angles}} \tfrac{1}{2}K_\theta(\theta-\theta_o)^2$$

$$+ \sum_{\text{All Torsion Angles}} K_\phi[1-\cos(n\phi+\delta)]$$

$$+ \sum_{\text{All nonbonded pairs}} \varepsilon\left[\left(r_o/r\right)^{12} - 2\left(r_o/r\right)^{6}\right]$$

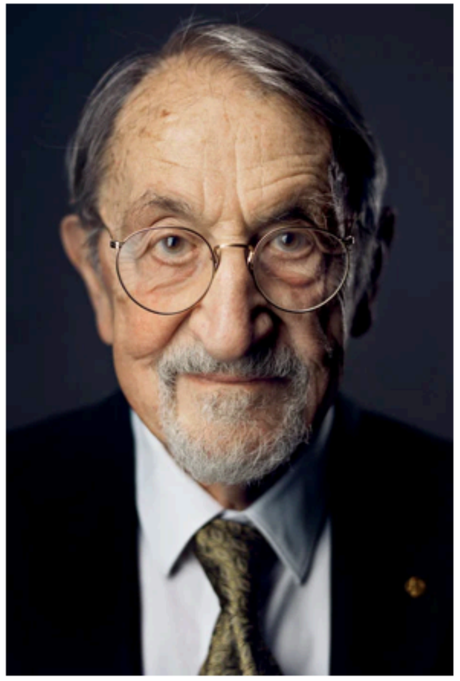
$$+ \sum_{\text{All partial charges}} 332 q_i q_j / r$$



2013 NOBEL PRIZE IN CHEMISTRY

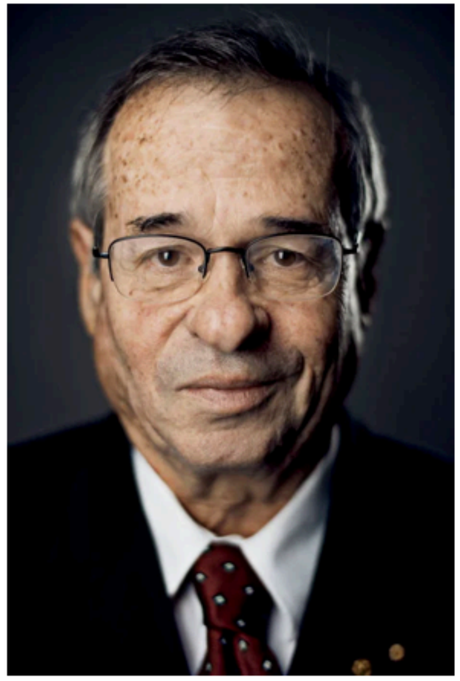
Martin Karplus
Michael Levitt
Arieh Warshel

"for the development of multiscale models for complex chemical systems"



Martin Karplus

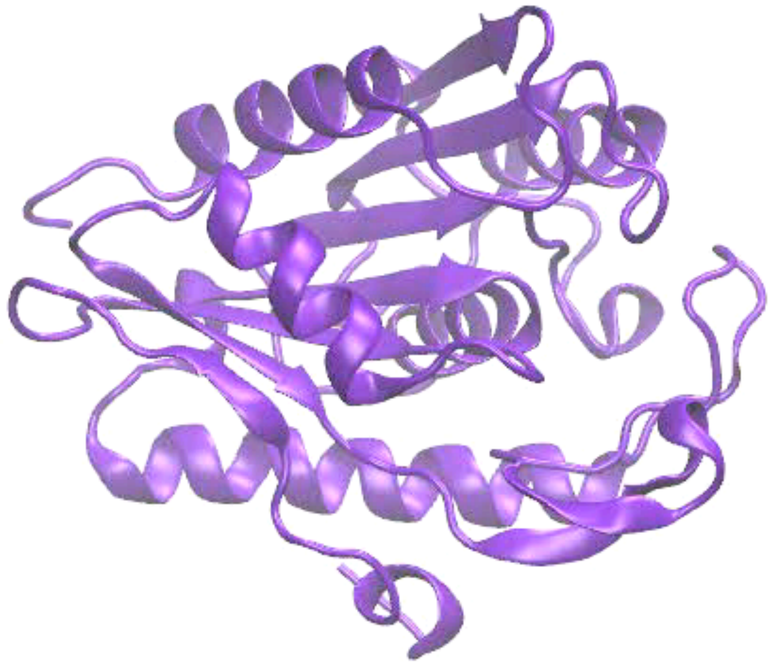
Michael Levitt



Arieh Warshel

X-ray crystallography

$\{x_i, y_i, z_i\}_{i=1,\ldots,N}$

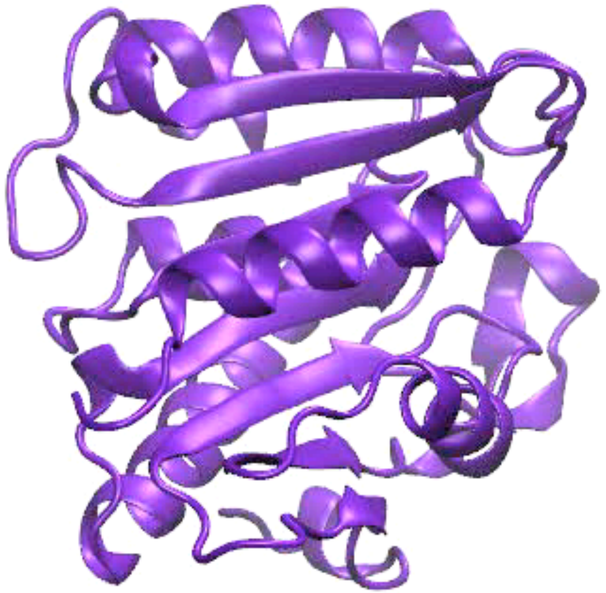


(human acyl-protein thioesterase)

molecular modeling and simulations

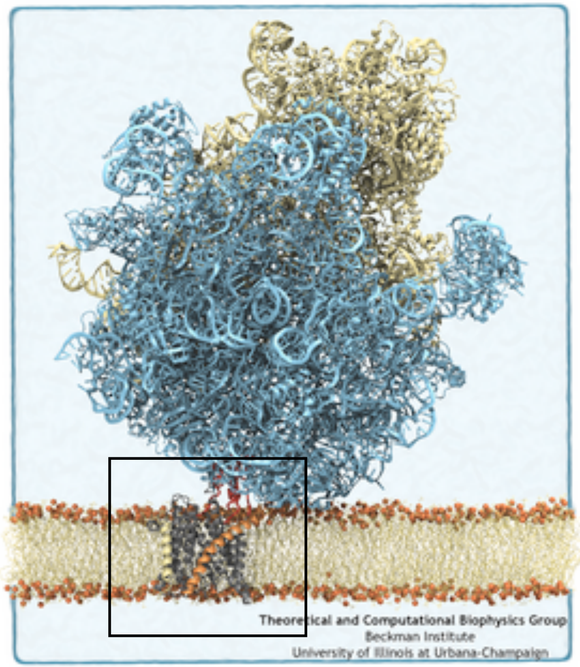
$$\{x_i(t), y_i(t), z_i(t)\}_{i=1,\ldots,N}$$

solvation
pH
post-translational modifications
interactions network
temperature effects (k_BT)
.....



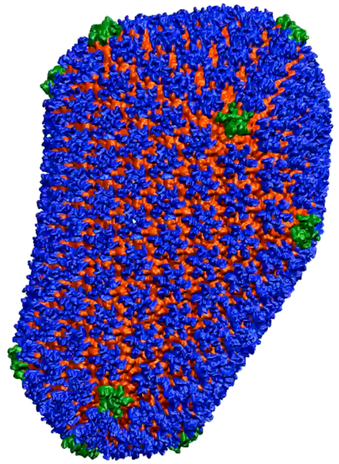
State-of-the-art of molecular simulations

- up to **10^2 millions** of atoms (e.g. viruses, ribosome)

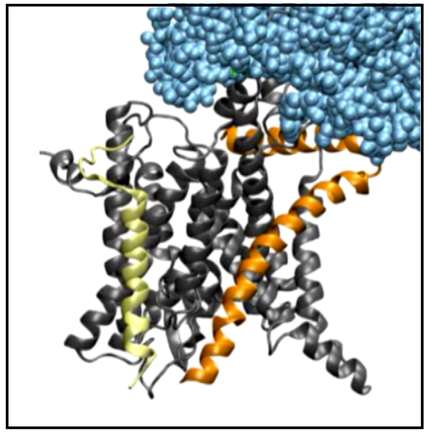


James Gumbart, et al.
Structure, 17:1453-1464, 2009.

HIV-1 capsid

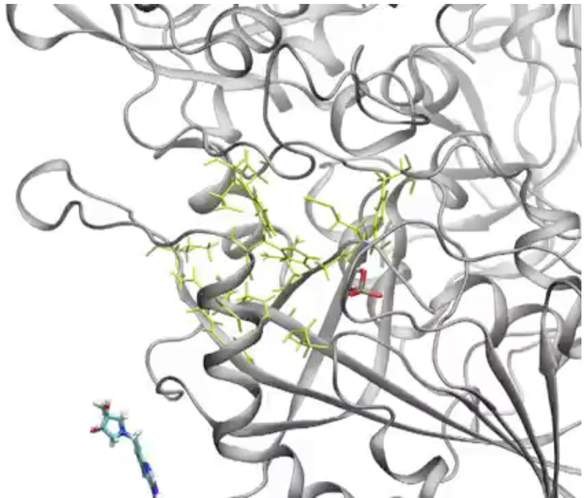


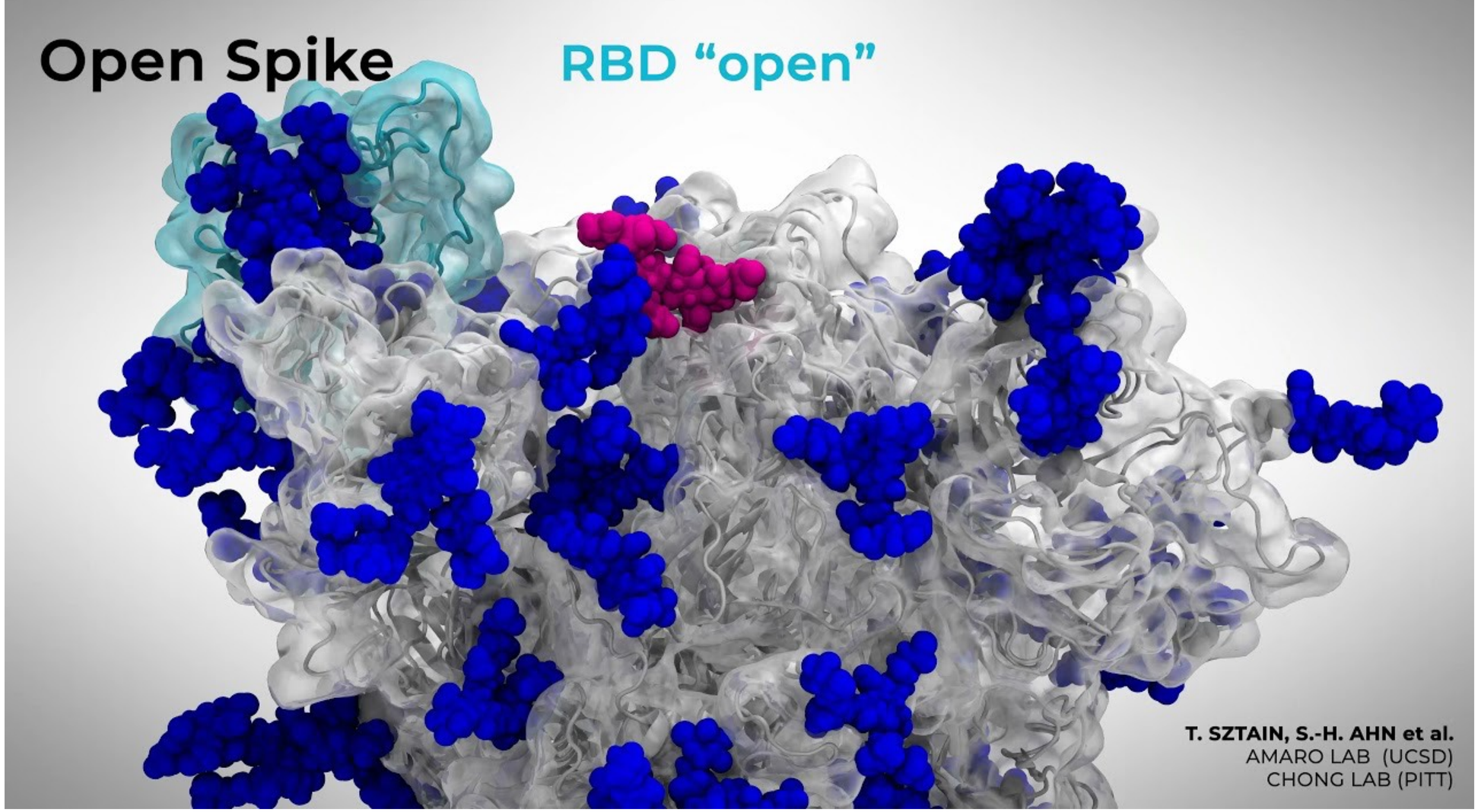
Zhao et al. *Nature*, 497:643-646, 2013



protein translocation

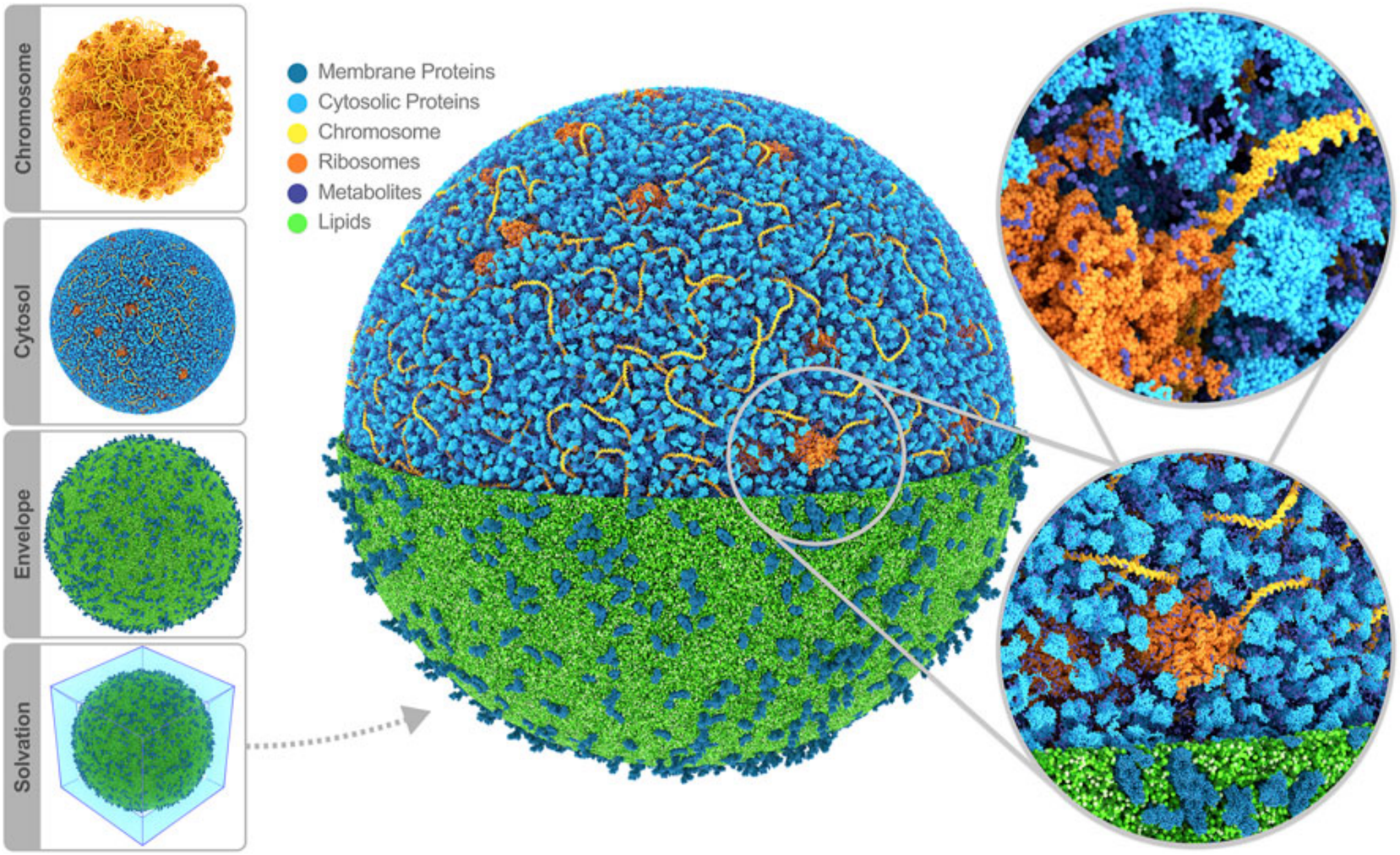
drug binding on a kinase





Molecular mechanism of SARS-CoV2 spike opening

State-of-the-art of molecular simulations

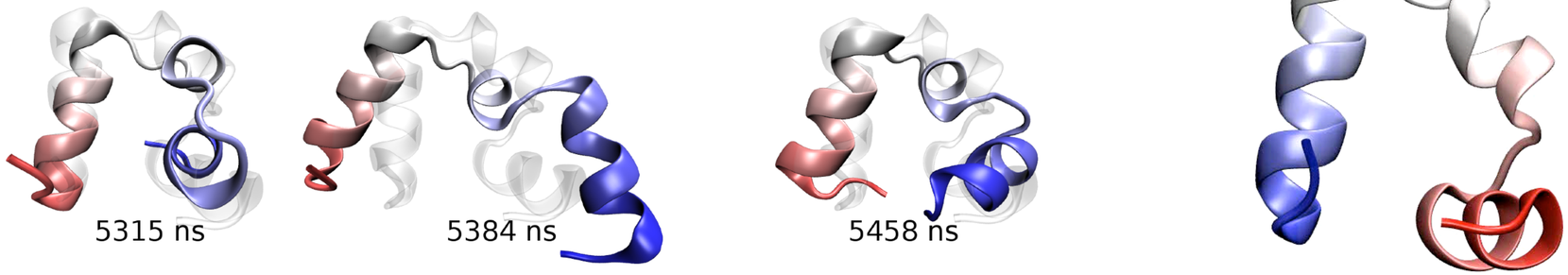


Whole-cell Martini model of **JCVI-syn3A**. The four stages of cell building are shown on the side. The final system contains 60,887 soluble proteins (light blue), 2,200 membrane proteins (blue), 503 ribosomes (orange), a single 500 kbp circular dsDNA (yellow), 1.3 million lipids (green), 1.7 million metabolites (dark blue), 14 million ions (not shown) and 447 million water beads (not shown) for a total of 561 million beads representing more than **six billion atoms**.

https://doi.org/10.3389/fchem.2023.1106495

State-of-the-art of molecular simulations

- up to the **millisecond** timescale

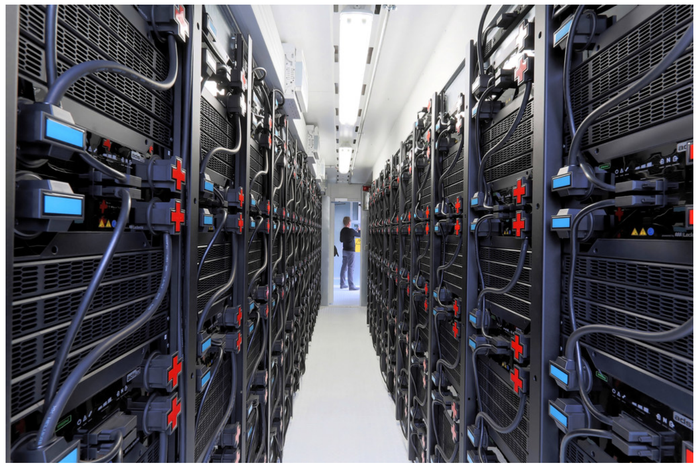


villin folding Freddolino, et al.. *Biophysical Journal*, 94:L75-L77, 2008.

Voelz *et al. J. Am. Chem. Soc.*, **2010**, *132*, 1526

High-Performance Computing (HPC) resources

Frontier - Oak Ridge 1.2 exaFlops



HPC@EPFL Kuma (GPU H100)

CSCS Alps - 270 PetaFlops

Anton D.E. Shaw Research

What to know...

- Simplified energy functions are used to calculate molecular potential energies rapidly using molecular simulation techniques

- The energies of covalent bonds are approximated by functions such as the Morse, which account for covalent bond breakage, or by harmonic potentials more easily

- The Lennard-Jones potentials describe weak attractions and strong repulsions between atoms.

- The interactions between charged atoms is governed by Coulomb’s law.

- Interactions with water weaken the effective strengths of hydrogen bonds and ionic interactions in proteins.

Spontaneous Reactions

Favorable energy changes in energy are not sufficient to indicate the direction of the spontaneous change

$$A + B \rightarrow C + D \quad \Delta U = \text{negative} \quad (\text{energy released})$$

Examples for spontaneous macroscopic reactions:

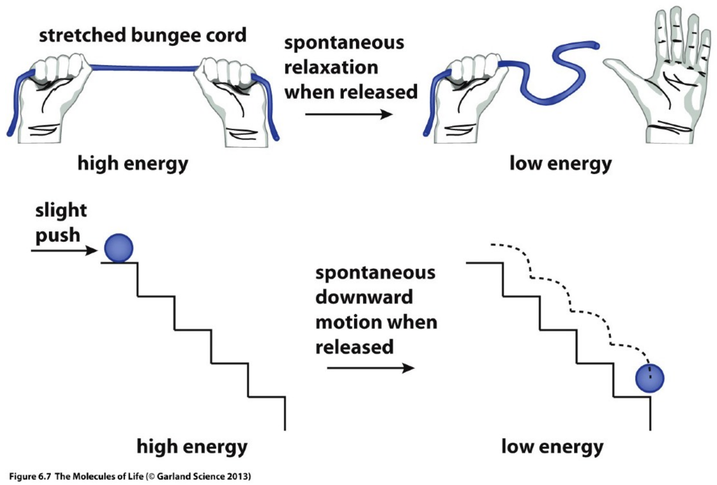
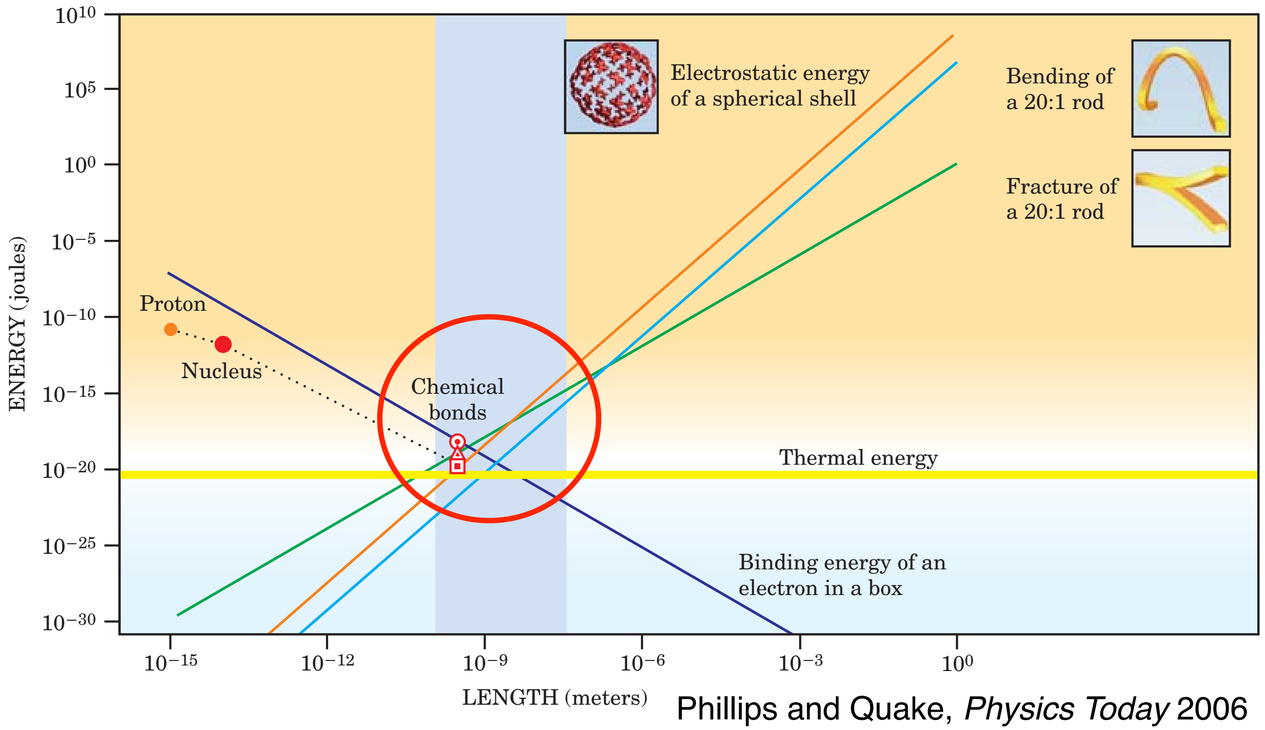


Figure 6.7 The Molecules of Life (© Garland Science 2013)

This intuition for microscopic systems is not always correct, because:
- The collective behavior also depends on **entropy** (see protein folding case)
- or there are **kinetic barriers** for the reaction to be overcome

The intriguing nature of biological interactions

- biological systems are subjected to deterministic forces (enthalpy) and thermal forces (entropy)
- at the dimension scale of biological systems these are however on the same order of magnitude
- all transformations in cells are thus determined by this subtle interplay, defined by the free energy of the system (**G=H-TS**) (accuracy in a noisy world, and use of thermal fluctuations to deploy biological function)



Phillips and Quake, *Physics Today* 2006

E_{det}/k_BT

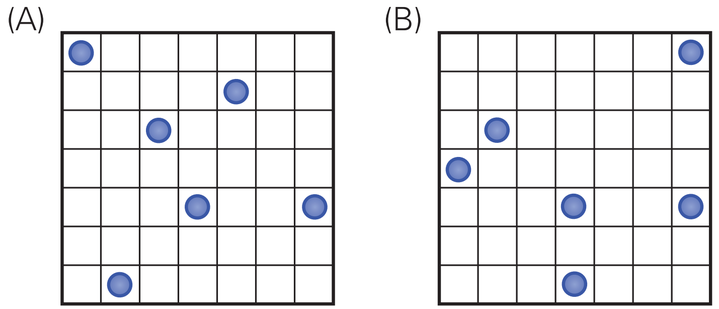
k_BT = 4.1 pN*nm
= 0.6 kcal/mol
= 2.5 kJ/mol
= 0.025 eV

at room temperature (300K)

Macrostates and Microstates

Energy alone does not indicate the direction of spontaneous change (eg isothermal expansion of ideal gas). For microscopic systems we use a **statistical approach** to describe them and use the concept **entropy** which describe their **randomness**.

This is related to the **multiplicity** of states of the system. The system is characterized by a **macrostate** with some global properties like temperature, pressure, and number of molecules. A **microstate** is a specific configuration of molecules that is consistent with that **macrostate**. Each of them corresponds to many different microstates.



microstate A microstate B

M potential configurations
N molecules

$$W(M,N) = \frac{M!}{N!(M-N)!}$$

the simple example of tossing coins

4 coins, 16 possible outcomes (4^2), each coin can give head (H) or tail (T), 2 states

Macrostate	Individual Microstates	Number of Microstates
4 H 0 T	HHHH	1
3 H 1 T	HHHT, HHTH, HTHH, THHH	4
2 H 2 T	HHTT, HTHT, THHT, THTH, TTHH, HTTH	6
1 H 3 T	HTTT, THTT, TTHT, TTTH	4
0 H 4 T	TTTT	1

N= number of coins with H

M=4, N=4
M=4, N=3
M=4, N=2
M=4, N=1
M=4, N=0

number of microstates for each macrostate :

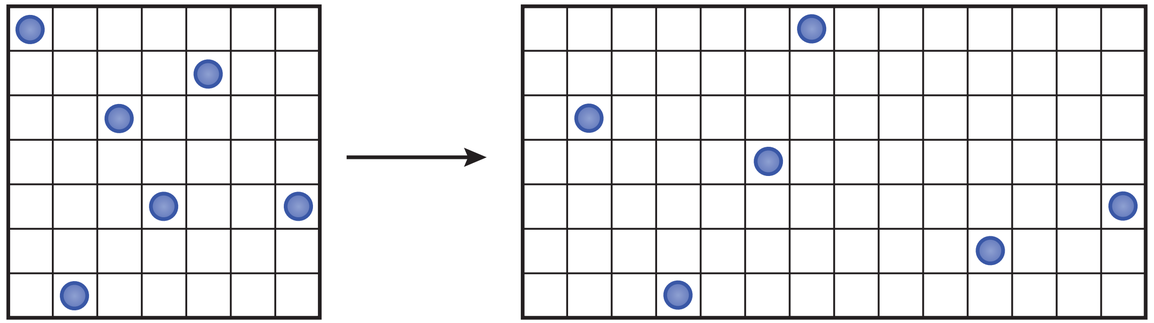
$$W(M,N)=\frac{M!}{N!(M-N)!}$$

Ex: what is the # of microstates for having 5H and 5T if you toss 10 coins?

If number of possible states is more than 2,
let's say ***t***, the formula becomes in general equal to

$$W=\frac{N!}{N_1!N_2!\ldots N_t!}$$

Microstates



$M = 49$ $M = 98$

$$W(49,6) = \frac{49!}{6!(49-6)!} = \frac{6.08\times10^{62}}{720\times6.04\times10^{52}} \approx 1.4\times10^{7}$$

using Stirling's approx $\ln n! \approx n\ln n - n$

$$\ln W(98,6) = 98\times\ln(98) - 98 - \ln(720) - (92\times\ln 92 - 92)$$

$$\Rightarrow W(98,6) = e^{20.72} = (10^{0.434})^{20.72} \approx 10^{9}$$

a system will tend to states of increased multiplicity or with maximal number of microstates.

Since *W* is a very high number for molecular systems (N_A), it is handier to use *lnW*, which has the nice property to be **extensive** (as most of the thermodynamic functions)

$$W_{A+B} = W_A \times W_B$$

$$\Rightarrow \ln W_{A+B} = \ln W_A + \ln W_B$$

Entropy (statistical definition)

$$S = k_B \ln W$$

which is therefore extensive and a state function, and has the unit of energy/temperature, J/K, like the Boltzmann constant. This is equivalent to the **thermodynamic definition** of S:

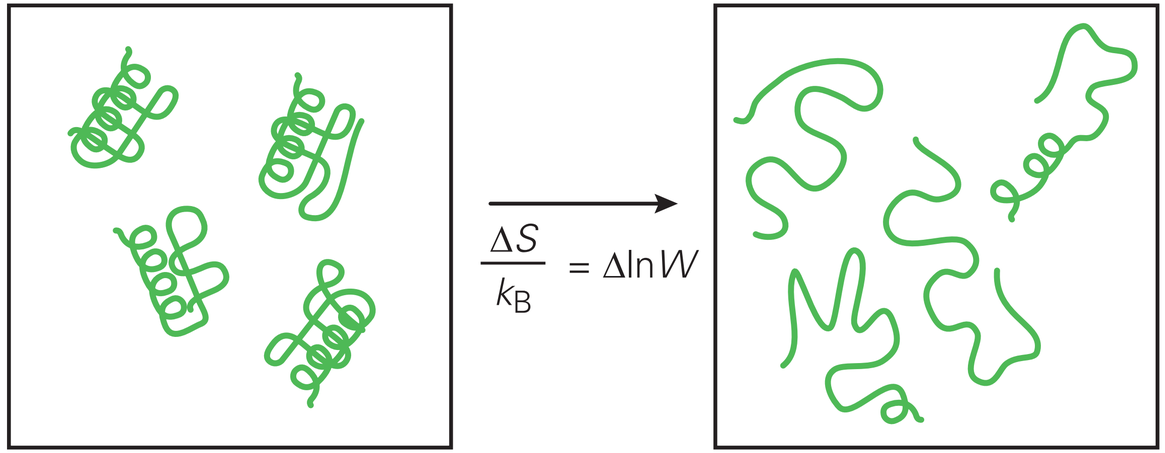
$$\Delta S = \frac{q_{rev}}{T}$$

that you have seen derived from the study of heat engines (see page 330 Chapter 7 for a demonstration for ideal gas). Thus spontaneous processes will increase entropy and at equilibrium S will be maximal, ie:

$dS_{sys} + dS_{surr} > 0$ and $dS_{sys} + dS_{surr} = 0$ at equilibrium

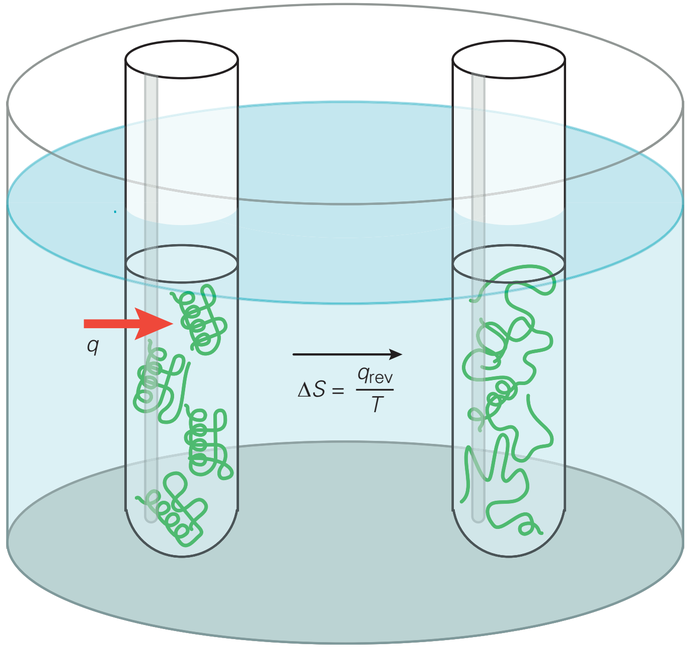
2nd law of thermodynamics - maximal entropy principle

statistical definition of S



W = number of conformations or configurations

thermodynamic definition of S



W is difficult to be measured, but heat can be measured and this equivalence provides a way to reconnect *S* to the molecular features of the system - the link is T, temperature

Energy distribution

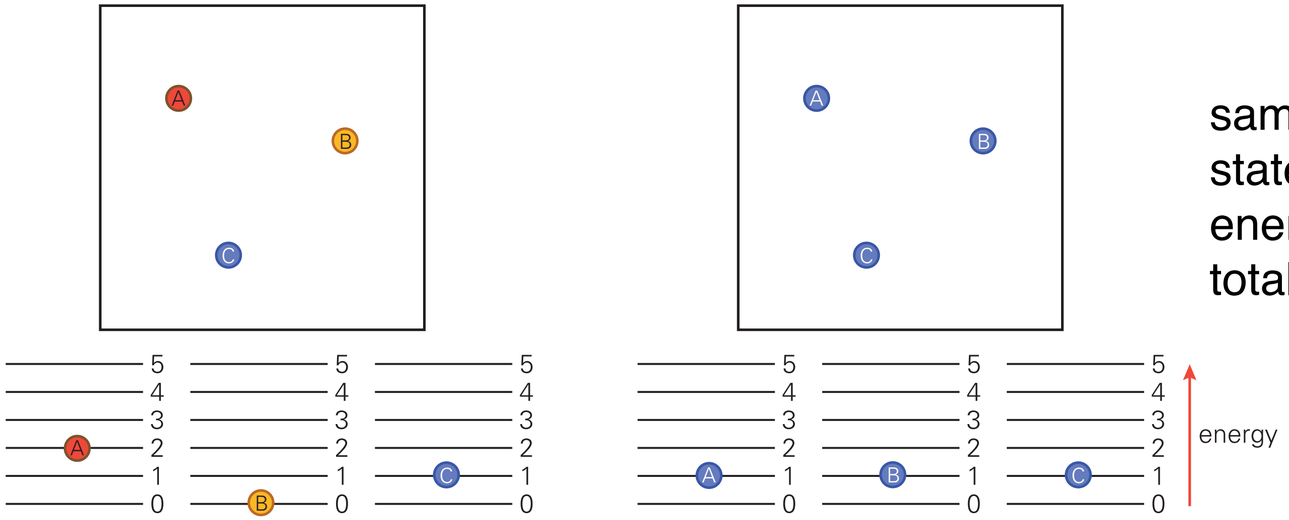
if we define the state of the system by the volume (V), the number of particles (N), and the total energy (U), this has a given *W*(positional), multiplicity in the configurational space and also a energy multiplicity, thus that in total for the system:

$$W_{\text{total}} = W_{\text{positional}} \times W_{\text{energy}}$$

thus:

$$S_{\text{total}} = k_{\text{B}} \ln W_{\text{total}}$$

$$= S_{\text{positional}} + S_{\text{energy}}$$

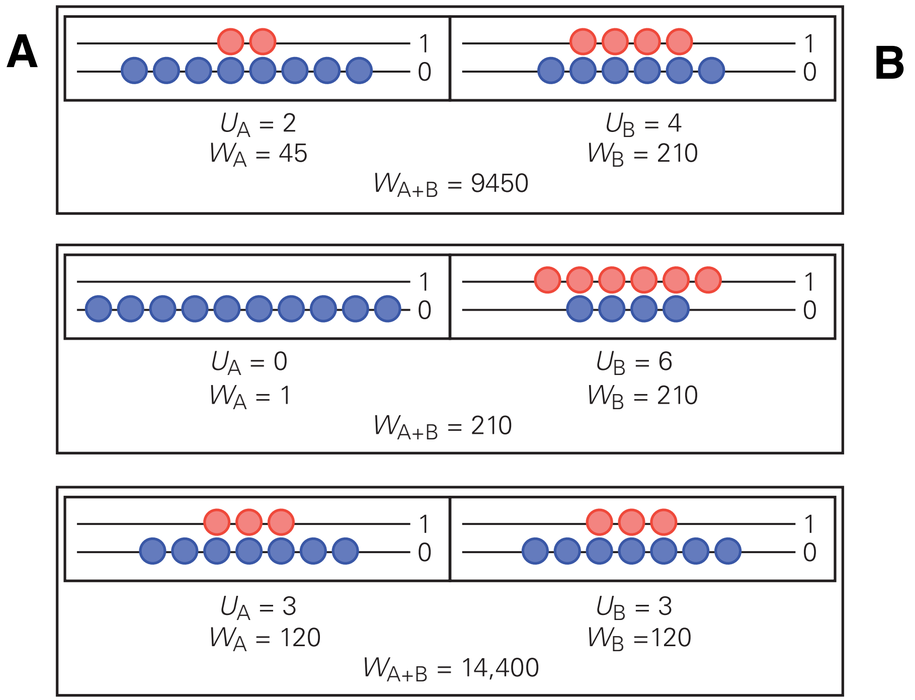


same configuration (micro state) with two possible energy distributions when total energy is U = 3 au

Energy distribution

Using similar counting strategies as before, this is obtained by: $$W = \frac{N!}{N_1!\,N_2!\ldots N_t!}$$

for *N* molecules distributed among a total of *t* different energy levels, where N_1, N_2, ... , N_i are the numbers of molecules in the first, second, etc., energy level (up to the highest level, *t*)



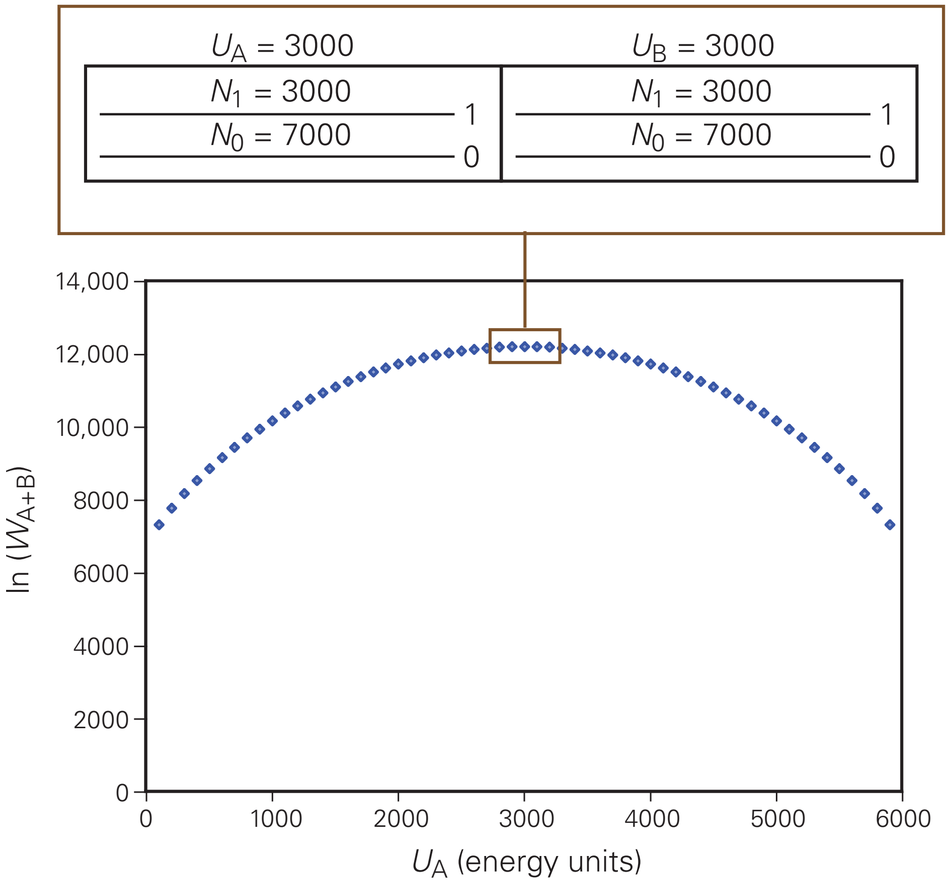
initial state $W_A = \frac{10!}{2!\,8!} = 45$ $W_B = \frac{10!}{4!\,6!} = 210$

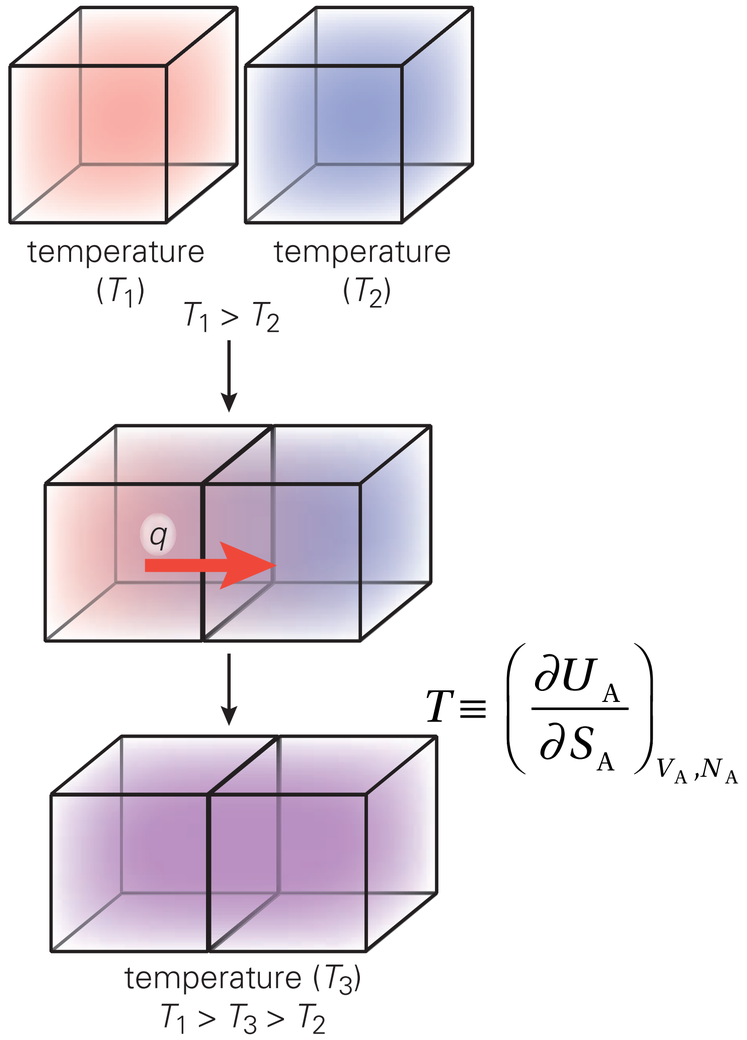
Energy is transferred from A to B

Energy is equally distributed in A a B

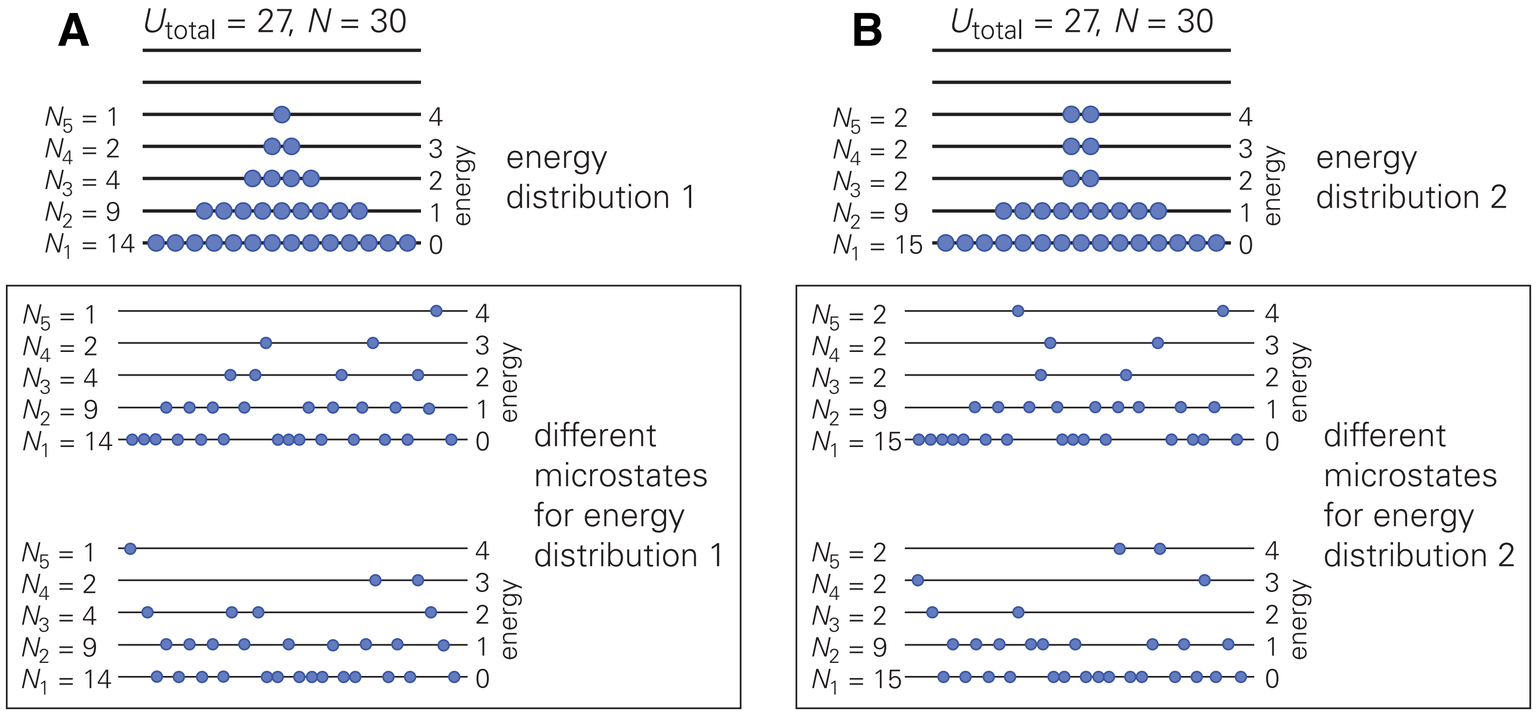
see how the combined energy multiplicity for the system A+B is larger when both has higher multiplicity and in fact the value is maximal when the energy is equally distributed between A and B

Energy distribution





see how the combined energy multiplicity for the system A+B is larger when both has higher multiplicity and in fact the value is maximal when the energy is equally distributed between A and B - **ie manifestation of the 2nd law of thermodynamics and maximisation of entropy S**



$$W_1 = \frac{30!}{1!2!4!9!14!} = 1.74 \times 10^{14}$$

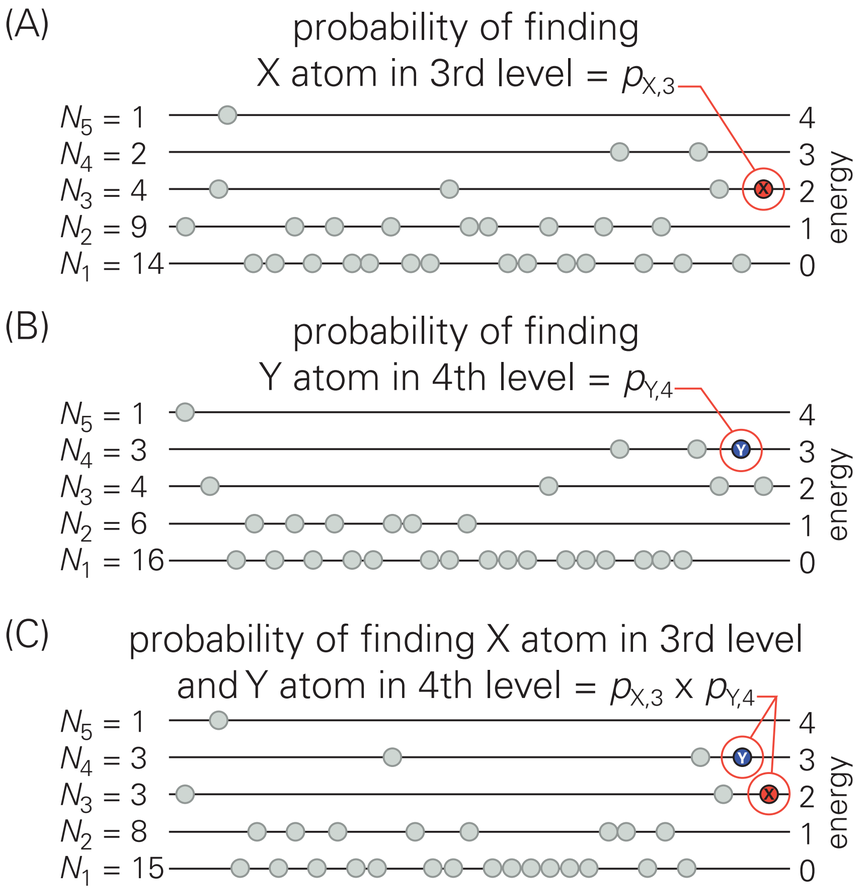
$$W_2 = \frac{30!}{2!2!2!9!15!} = 0.69 \times 10^{14}$$

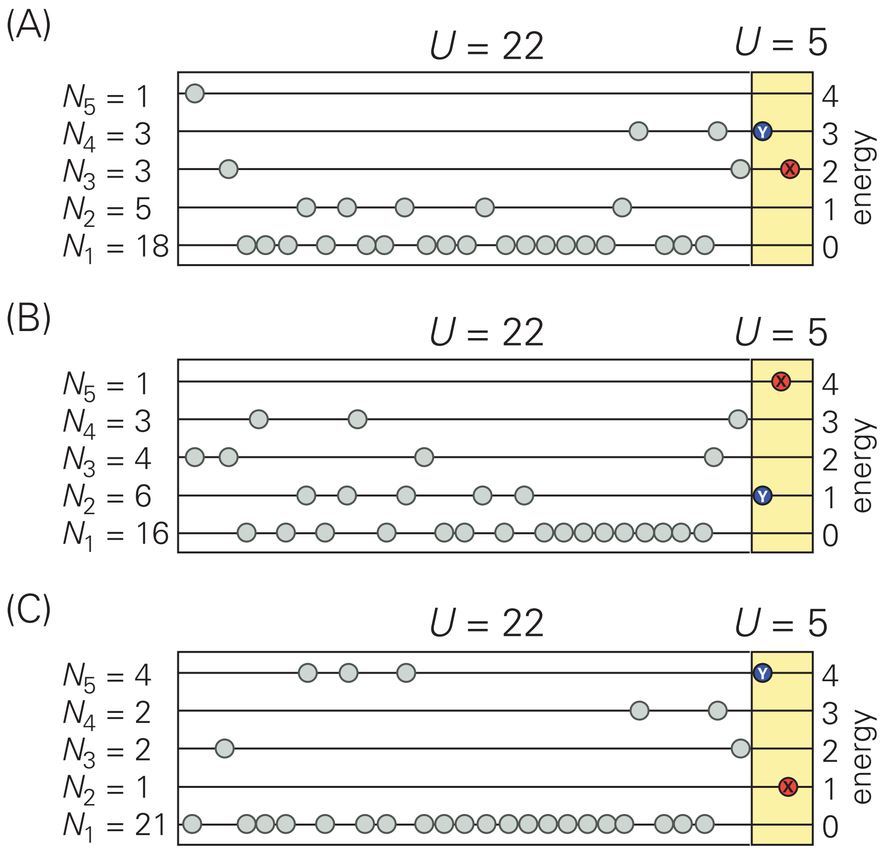
as the number of N increases one specific distribution dominates over all the other and this is the **Boltzmann distribution**, that is an **exponential** distribution, because this is the only mathematical way to reconcile that fact that

$$p_{XY} = p_{X,j} \times p_{Y,k} = Af(u_j) \times Af(u_k)$$ but also

$$p_{XY} = Bg(u_j + u_k) \qquad \Rightarrow f(u_j) \times f(u_k) \propto g(u_j + u_k)$$ thus if $$p_j = \frac{e^{-u_j / k_B T}}{Q}$$

with $$p_i = N_i / N$$





$$p_{XY} = p_{X,j} \times p_{Y,k} = Af(u_j) \times Af(u_k)$$

$$\Rightarrow f(u_j) \times f(u_k) \propto g(u_j + u_k)$$

$$f(u_j) \times f(u_k) = e^{-\beta u_j} \times e^{-\beta u_k} = e^{-\beta (u_k + u_j)} = g(u_j + u_k)$$

$$\Rightarrow \quad p_j \propto Ae^{-\beta u_j}$$

$$p_{XY} = Bg(u_j + u_k)$$

The total energy of the system remains constant

Each particle's energy distribution is independent of the others.

All microstates with the same total energy are equally likely.

Boltzmann Distribution

If you have eg an isolate system with N - number of atoms or molecules, and total energy U, when **N is large (~N_A)**, it is difficult to know how the energy is distributed through N atoms. Thus we can **only describe in statistical terms the population of a state**, i.e. the N_i – number of molecules that will be found in an energy level with energy U_i.

If you have M energy levels you can have different state distributions of this kind {N_0, N_1, … , N_M}, eg if N = 100 , {98,0,2, …} or {96,1,1,1,1, …}. The most probable state is the one with more potential configurations and it is described by the **Boltzmann distribution**:

$$p_i = {N_i}/{N} \quad , \quad N_i = \frac{Ne^{-U_i/k_B T}}{Q}$$

where Q is the **partition function** $Q = \sum_i e^{-U_i/k_B T}$

and k_B is the **Boltzmann constant** (k_B = 1.381 x 10^{-23} J/K)

From this, temperature T is a parameter that characterises the distribution

Boltzmann Distribution and Energetic Levels

- definition of the Boltzmann distribution:

$$N_i = \frac{Ne^{-U_i/k_B T}}{Q}$$

- the partition function Q is constant at a given temperature (we assume that the energies of U_i don't change with T), therefore we can say that

$$N_i \propto e^{-U_i/k_B T}$$

and thus you can estimate the ratios of between different populations at different energy levels using the following relation:

$$\frac{N_2}{N_1} = e^{-\Delta U/k_B T} \qquad \Delta U = U_2 - U_1$$

Remember that the gas constant R is the "molar" form of k_B, in fact:
$R = N_A k_B = 8.3145$ J/K*mol, thus if you work with KJ/mol you have to use RT in the Boltzmann distribution

Boltzmann Distribution and Energetic Levels

-With the formula below we can access the ratios between populations of different energy levels using the following relation:

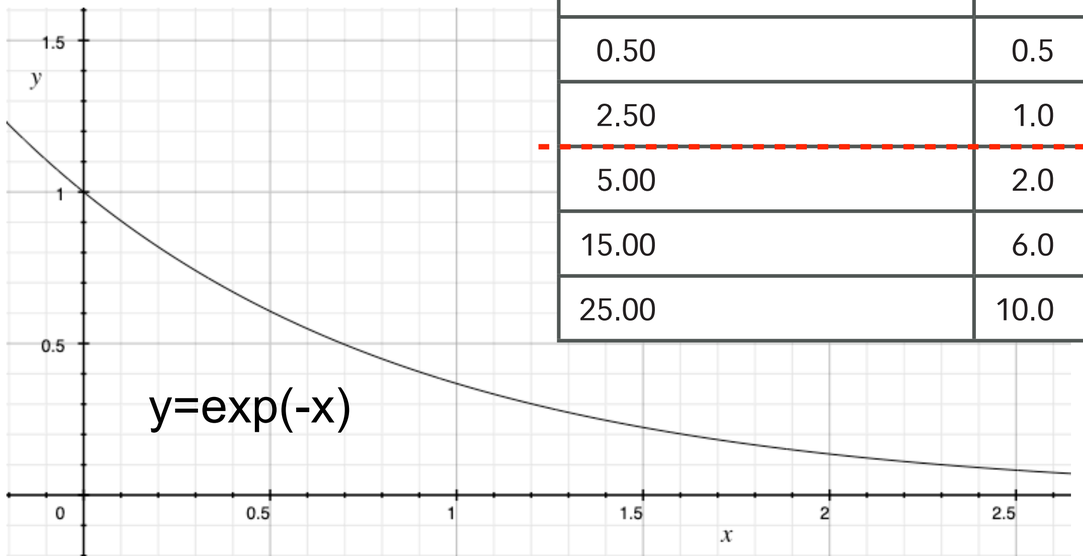
$$\frac{N_2}{N_1} = e^{-\Delta U / k_B T} \qquad \Delta U = U_2 - U_1$$

$$\frac{N_2}{N_1} = e^{-\Delta U / 2.529}$$

using as unit kJ/mol

ΔU (kJ•mol^{-1})	$\frac{\Delta U}{k_B T}$ (T = 300 K, $k_B T \approx$ 2.5 kJ•mol^{-1})	$e^{\frac{-\Delta U}{k_B T}}$
0.25	0.1	0.90
0.50	0.5	0.61
2.50	1.0	0.37
5.00	2.0	0.13
15.00	6.0	0.00067
25.00	10.0	0.0000045

$k_B T$



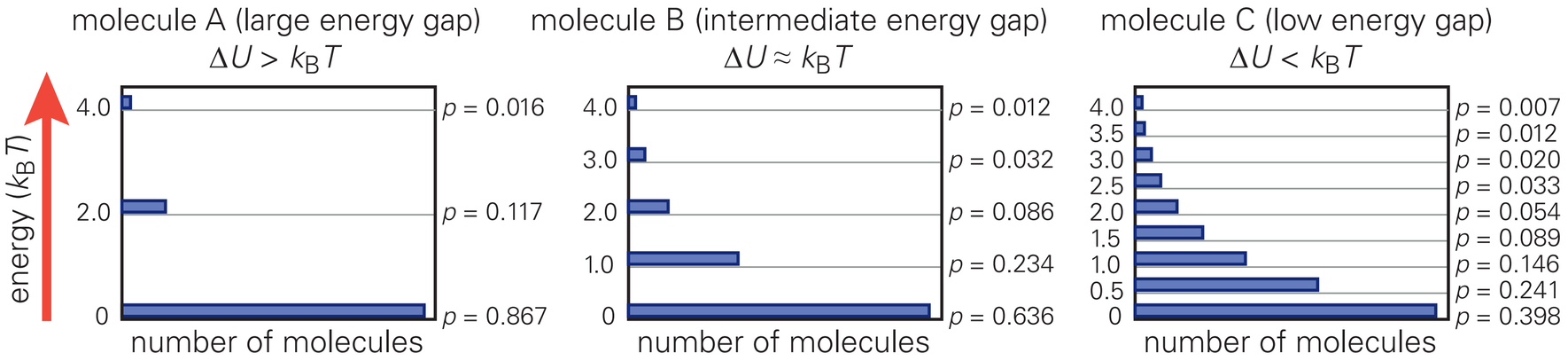
Boltzmann Distribution and Energetic Levels

-With the formula below we can access the ratios between populations of different energy levels using the following relation:

$$\frac{N_2}{N_1} = e^{-\Delta U / k_B T} \qquad \Delta U = U_2 - U_1$$

$$\frac{N_2}{N_1} = e^{-\Delta U / 2.529}$$

using as unit kJ/mol



3 molecules with different accessible energy levels - levels that are less spaced (< kT) are more accessible - at constant T

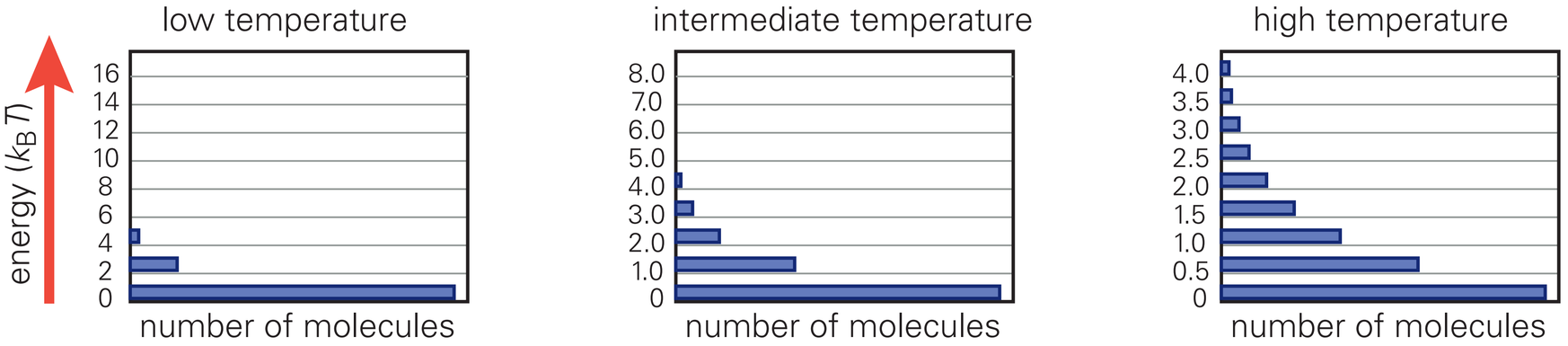
Boltzmann Distribution and Energetic Levels

-With the formula below we can access the ratios between populations of different energy levels using the following relation:

$$\frac{N_2}{N_1} = e^{-\Delta U / k_B T} \qquad \Delta U = U_2 - U_1$$

$$\frac{N_2}{N_1} = e^{-\Delta U / 2.529}$$

using as unit kJ/mol



same molecule at different T, the occupancy of energy levels increases with T

Boltzmann Distribution in macromolecules

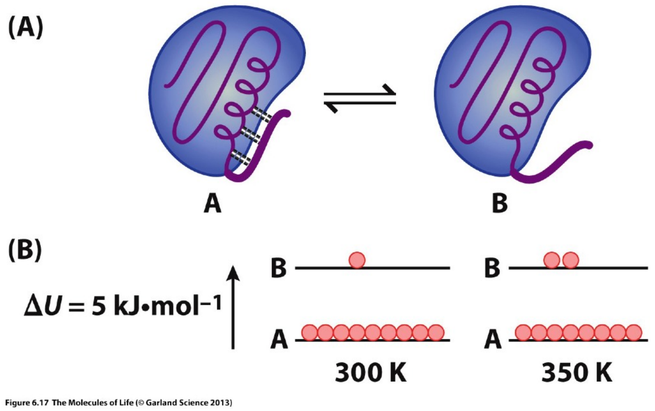


Figure 6.17 The Molecules of Life (© Garland Science 2013)

Protein molecules take up energy as they unfold

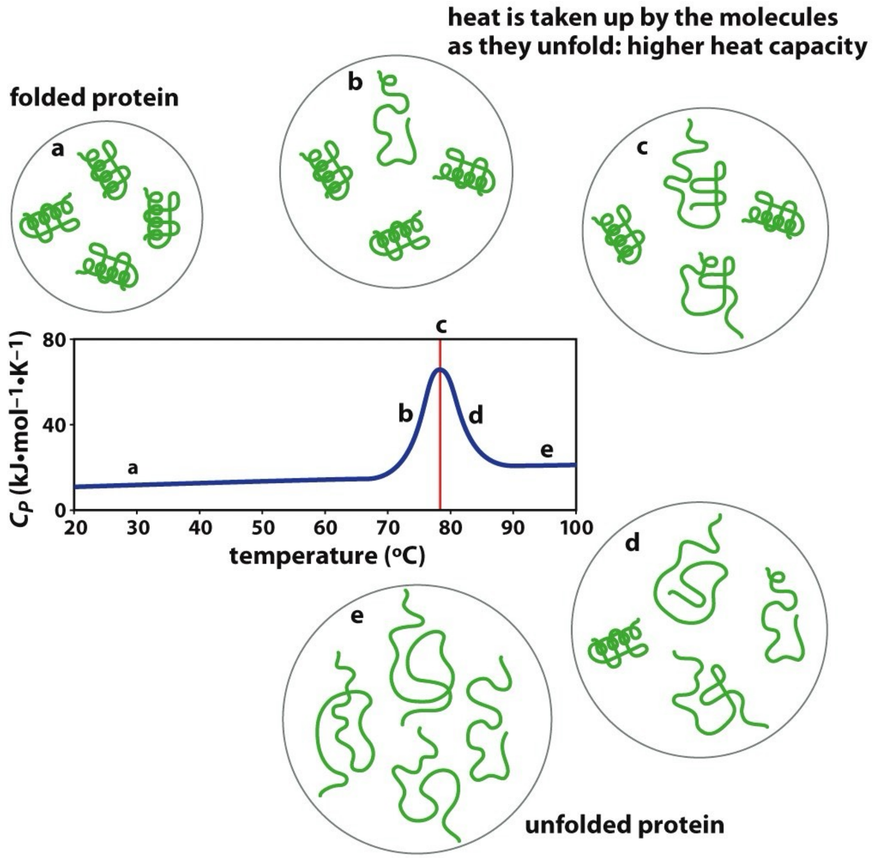


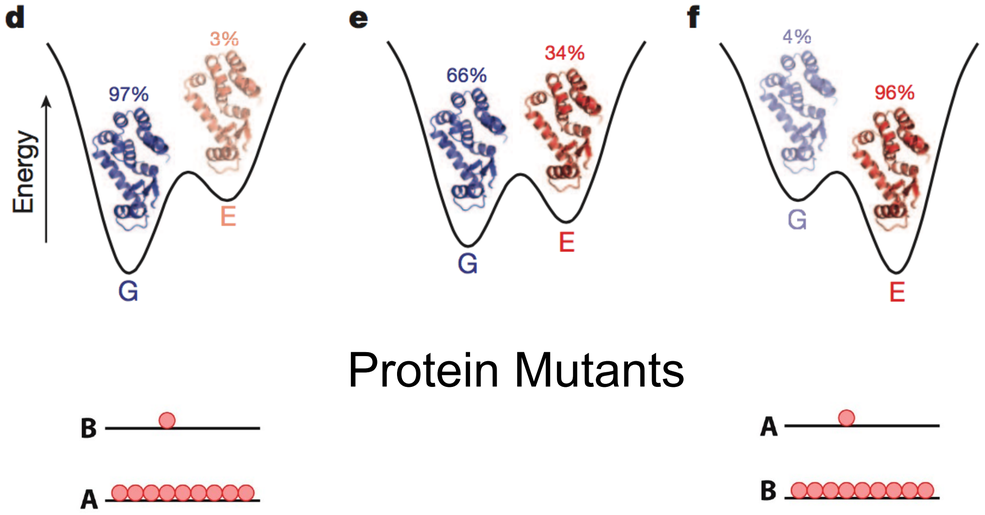
Figure 6.18 The Molecules of Life (© Garland Science 2013)

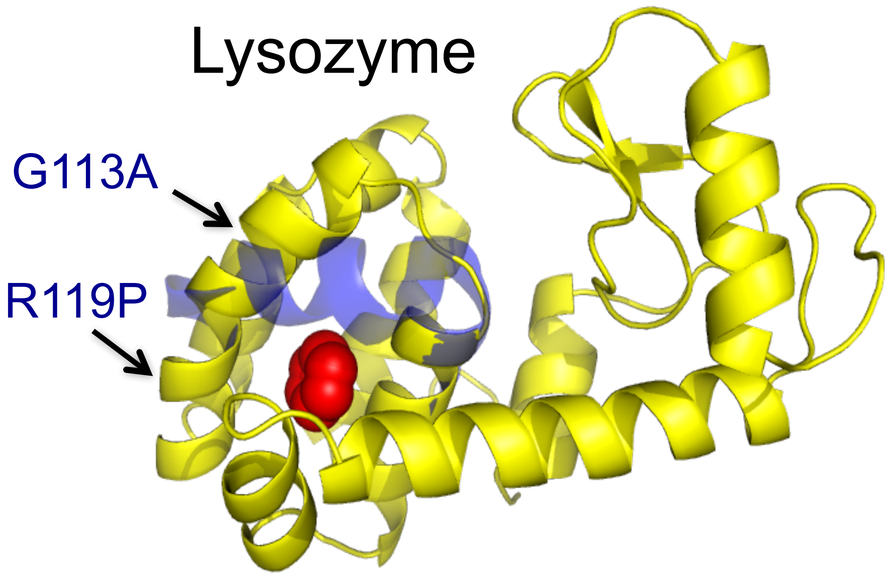
Shifting the distribution of populations with temperature

One can see how the formalism of the Boltzmann distribution helps us to understand what occurs in proteins and other biomolecules.

Engineering different protein states

Solution structure of a minor and transiently formed state of a T4 lysozyme mutant





Ground State (xtal-3dmv)
Excited State
Benzene
Excited State Stabilizing Mutations

- Proteins can co-exist in multiple conformational states (e.g. ground and excited)

What to know...

- The Boltzmann distribution describes the populations of molecules in different energy levels.

- Energy levels corresponding to energies much greater than k_BT above the lowest energy level are not highly populated.

- The energy required to break molecular interactions in folded macromolecules gives rise to the peak in heat capacity when the temperature is increased.

Gibbs Free Energy

How do we know when a biological process is at equilibrium or at least tend to the equilibrium conditions? We introduce a new function that will include the 1st and 2nd law, ie. conservation of energy and maximisation of entropy. This is the free energy (**Gibbs free energy, G** at constant pressure)

$$G = H - TS$$

which <u>always</u> decreases when a process occurs spontaneously and it is at a minimum at equilibrium at constant pressure and temperature

$$dG = dH - TdS \qquad \text{with} \qquad dG \leq 0$$

- G is only dependent on the system
- It accounts for both enthalpic and entropic contributions

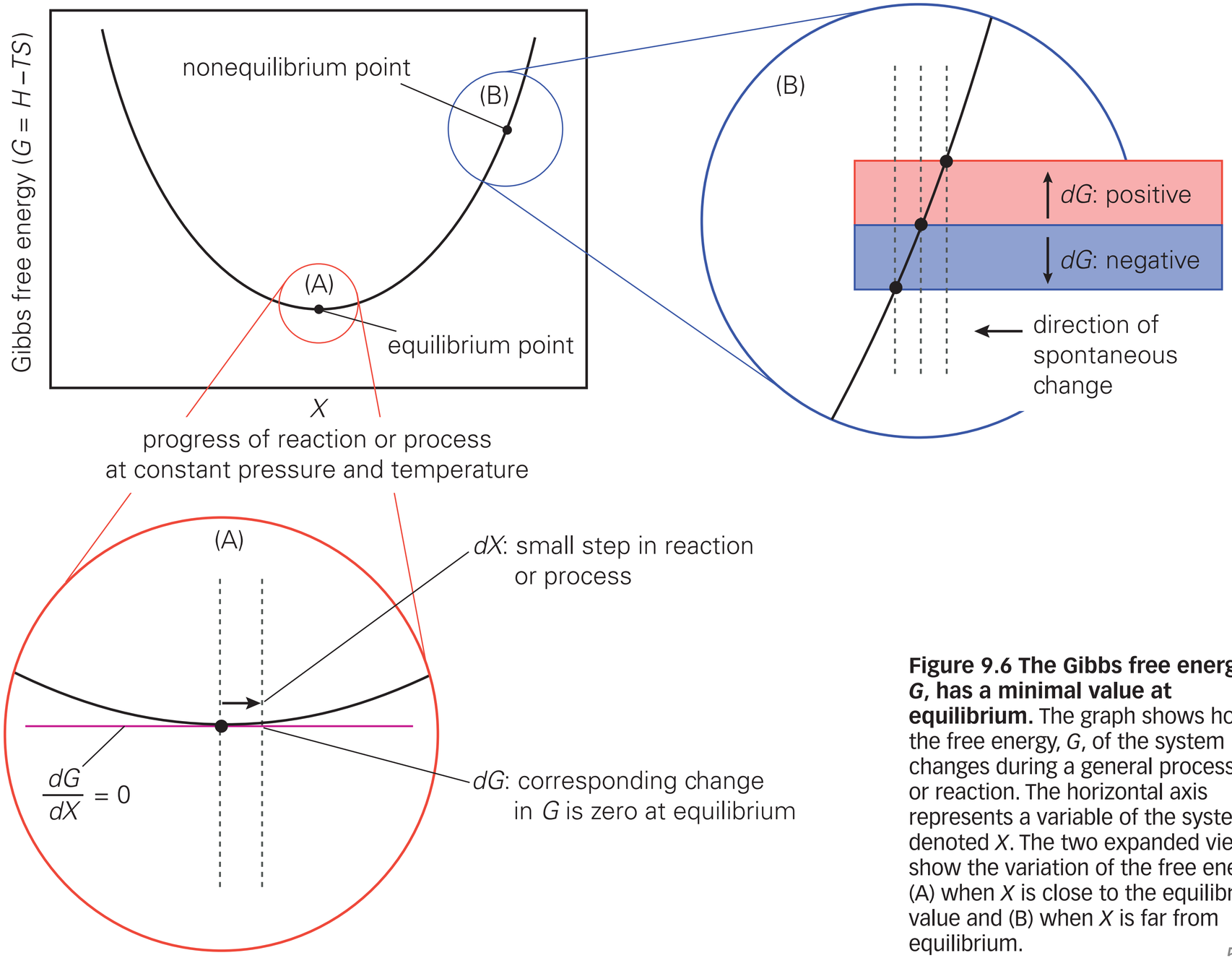
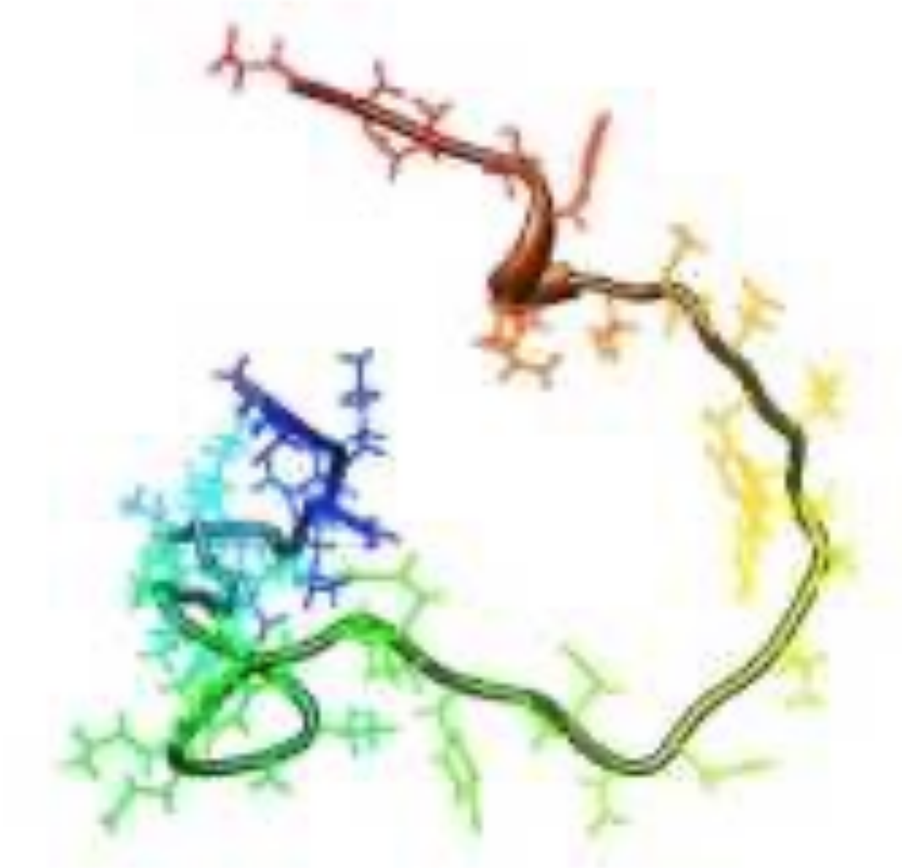
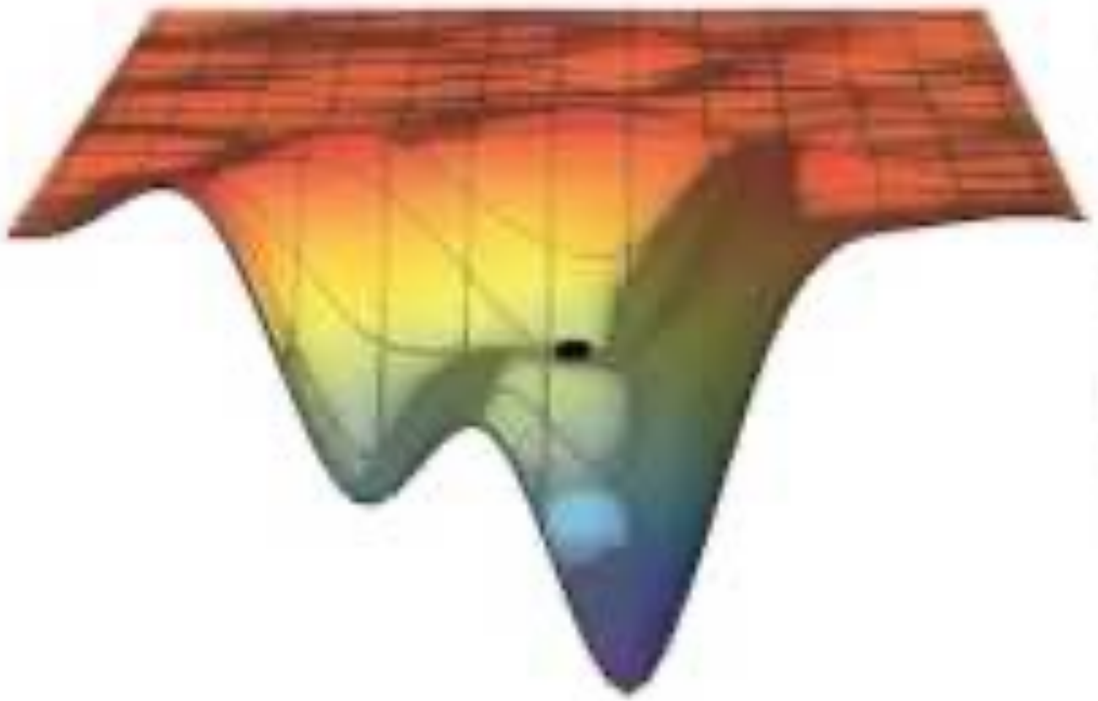


Figure 9.6 The Gibbs free energy, *G*, has a minimal value at equilibrium. The graph shows how the free energy, *G*, of the system changes during a general process or reaction. The horizontal axis represents a variable of the system, denoted *X*. The two expanded views show the variation of the free energy (A) when *X* is close to the equilibrium value and (B) when *X* is far from equilibrium.





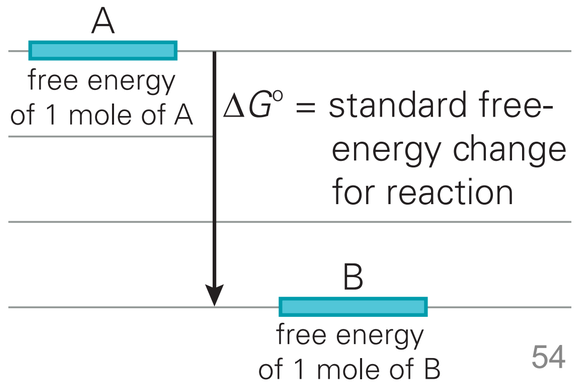
Free Energy of chemical reactions

Consider ATP hydrolysis:

$$\mathrm{ATP} + \mathrm{H_2O} \rightarrow \mathrm{ADP} + \mathrm{P_i}$$

$$\Delta G = \int_{\text{reactants}}^{\text{products}} dG = G(\text{products}) - G(\text{reactants}) = G(\mathrm{ADP} + \mathrm{P_i}) - G(\mathrm{ATP} + \mathrm{H_2O})$$

Usually we refer to the standard ΔG^{o} at standard conditions: ie pressure 1 atm, 1 M of solute, apart for water (55 M), and room temperature 298 K. Thus at these conditions for ATP hydrolysis is $\Delta G^{0} = -28\ \mathrm{kJ \cdot mol^{-1}}$

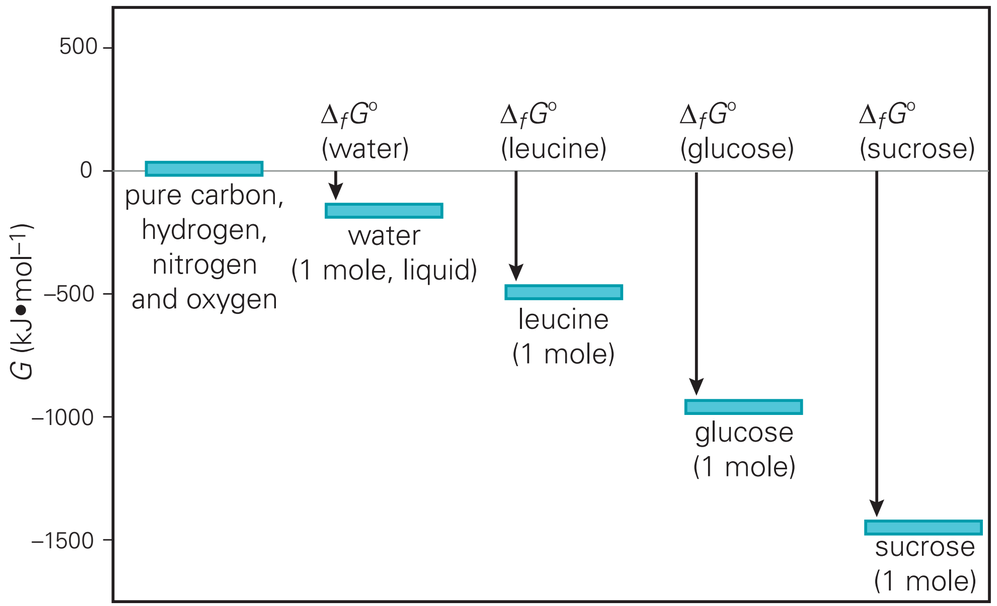


Free Energy of chemical reactions

Compound	$\Delta G^{o\prime}$ (kJ•mol^{-1})
acetate$^-$	–369.2
CO_2 (gas)	–394.4
CO_2 (aqueous solution)	–386.2
carbonate ion	–587.1
ethanol	–181.5
fructose	–915.4
fructose-6-phosphate^{2-}	–1758.3
α-D-glucose	–917.2
glucose-6-phosphate^{2-}	–1760.2
H^+ (aqueous solution)	0.0
H_2 (gas)	0.0
H_2O (liquid)	–237.2
isocitrate^{3-}	–1160.0
lactate$^-$	–516.6
OH^-	–157.3
pyruvate$^-$	–474.5
succinate^{2-}	–690.2

How to calculate these free energies? Using the free energy of formation of the molecules involved in the reactions, starting from the composing elements

$$\Delta G^o = \sum_{\text{all products}} \Delta_f G^o(\text{product}) - \sum_{\text{all reactants}} \Delta_f G^o(\text{reactant})$$



Free Energy of chemical reactions

Since it might be difficult to measure the free energy of formation for some given products in one single step, the reaction is broken down in intermediate steps involving less complex reactions and thermodynamic cycles are used to calculate the final free energy (this is working because **G is a state function**)