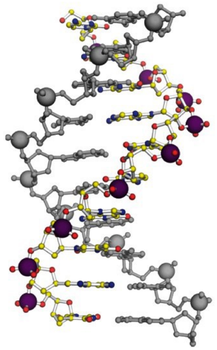
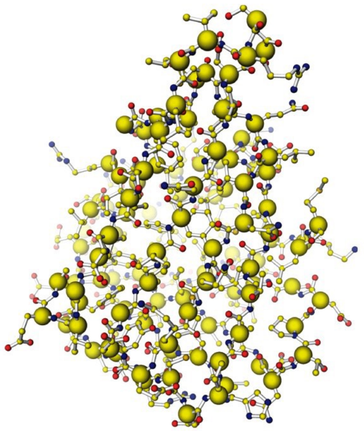
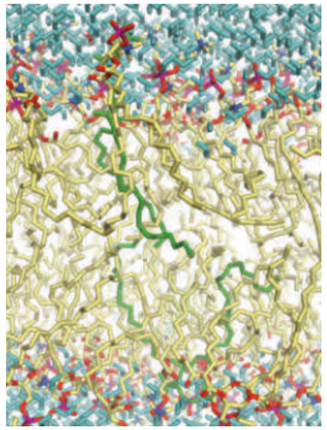


Chimie Biologique I
Biological Chemistry I
BIO-212

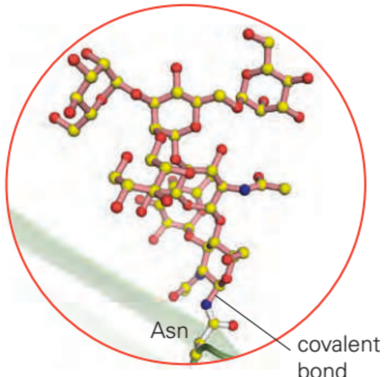


Lecture 11

Matteo Dal Peraro



Chapters 9 and 10 TMOL



Gibbs Free Energy

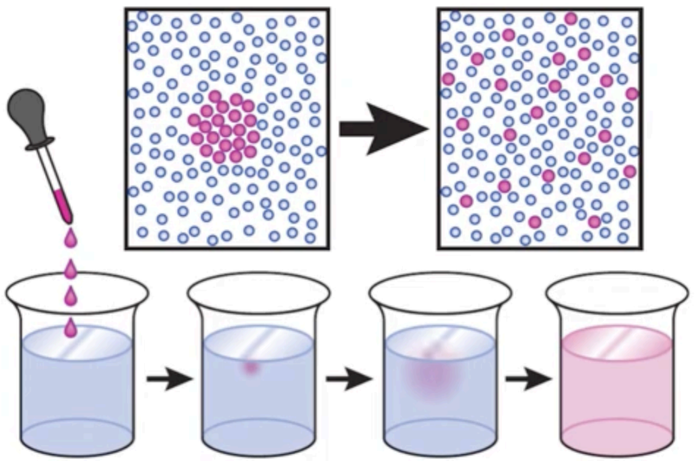
How do we know when a biological process is at equilibrium or at least tend to the equilibrium conditions? We introduce a new function that will include the 1st and 2nd law, ie. conservation of energy and maximisation of entropy. This is the free energy (**Gibbs free energy, G** at constant pressure)

$$G = H - TS \qquad (H = U + pV)$$

$$S = k_B ln W$$

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

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



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

Gibbs Free Energy

The Gibbs free energy **G = H - TS** measures the maximum amount of energy in a system that can be converted into useful (non-volume) work while maintaining constant temperature and pressure.

$$dG = dH - TdS = \delta W_{non-volume}$$

Under constant T and P:

1. Volume work (*PdV*) is automatically handled by the H term in G = H - TS.

2. *G* focuses on the remaining energy available for “useful” work, like electrical work, chemical reactions, or mechanical work. *dG* quantifies this available energy.

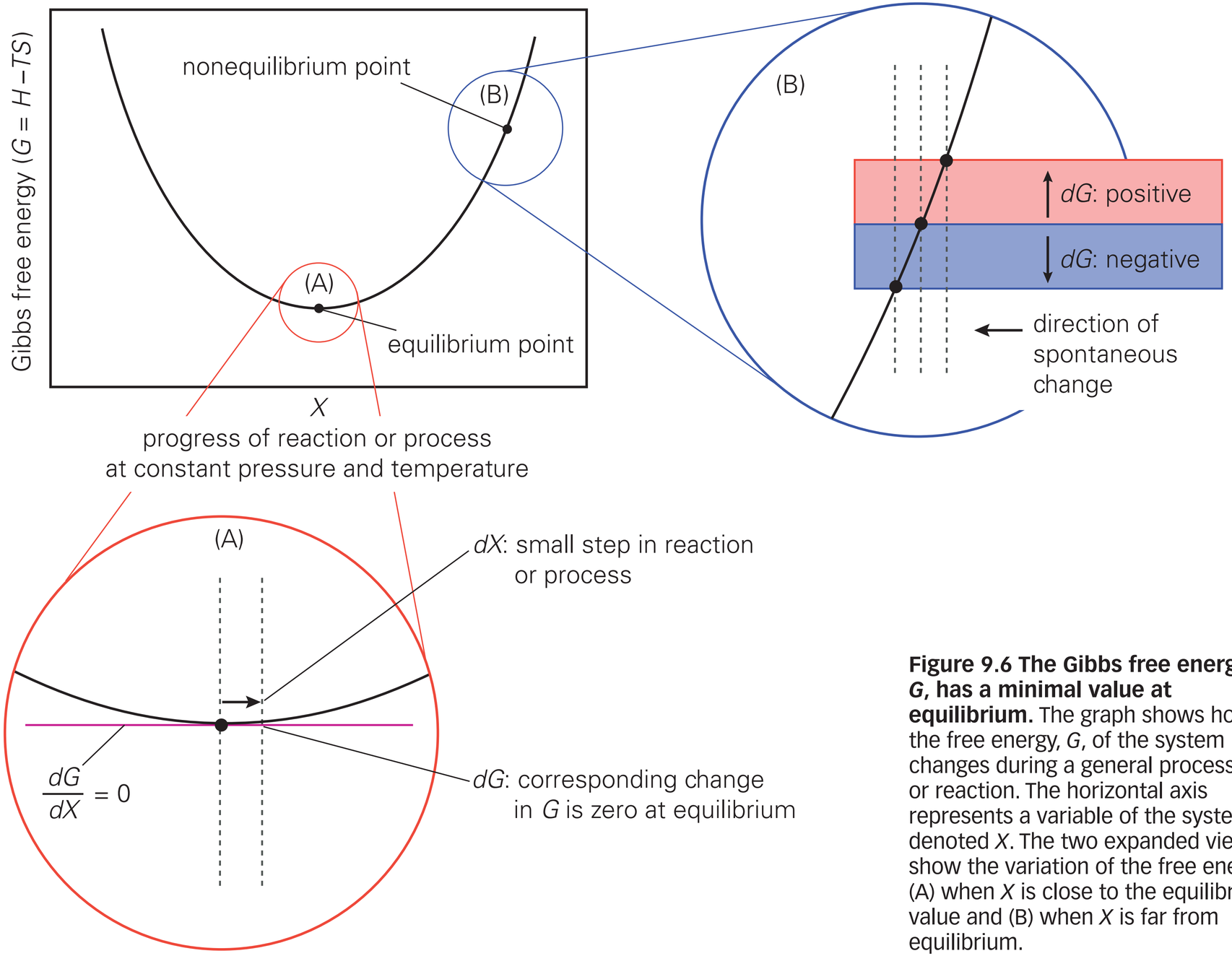
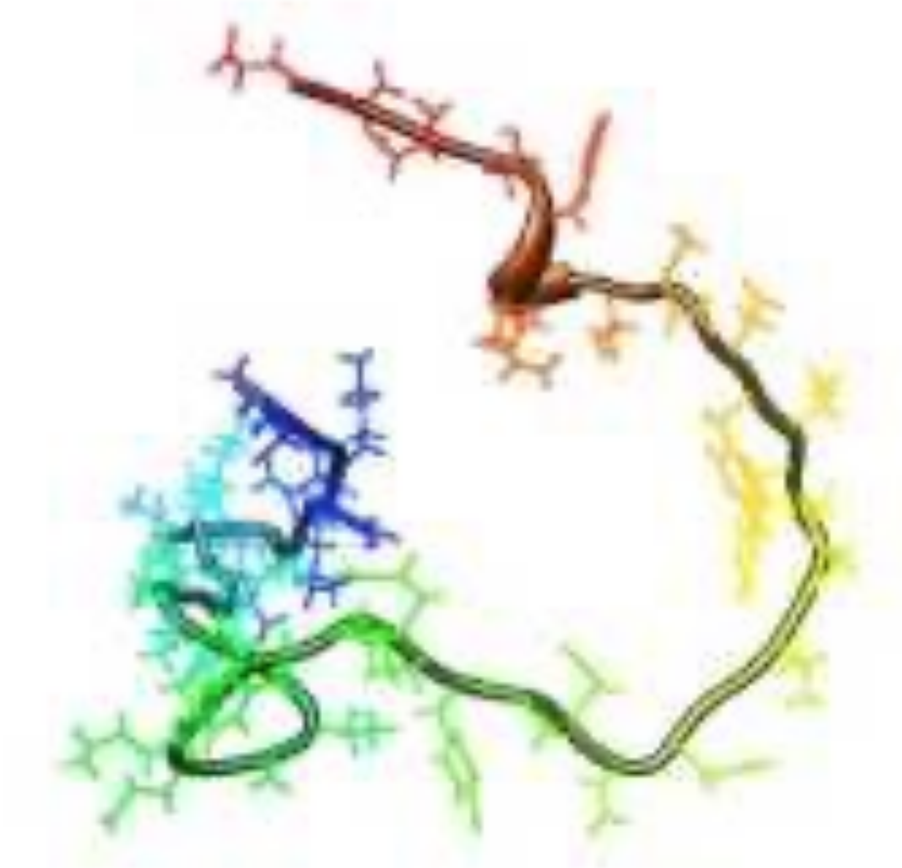
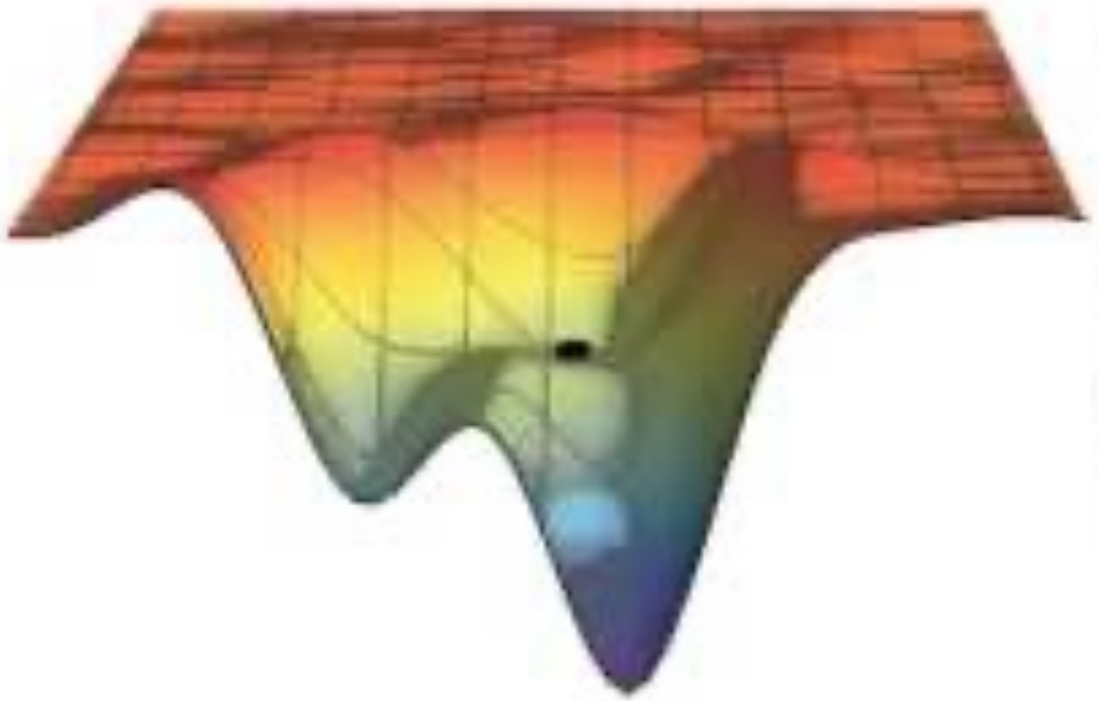


Figure 9.6 The Gibbs free energy, *G*, has a minimal value at equilibrium. The graph shows how the free energy, *G*, of the system changes during a general process or reaction. The horizontal axis represents a variable of the system, denoted *X*. The two expanded views show the variation of the free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

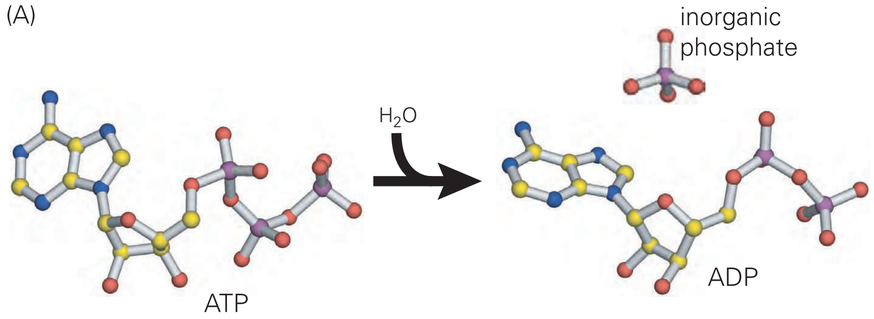




Free Energy of chemical reactions

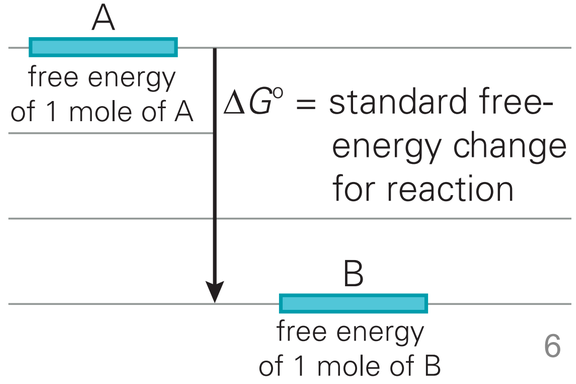
Consider ATP hydrolysis:

$$ATP + H_2O \rightarrow ADP + P_i$$

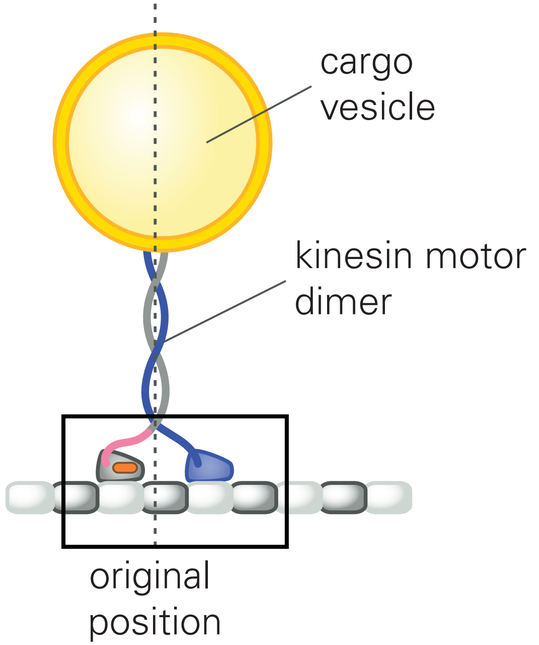


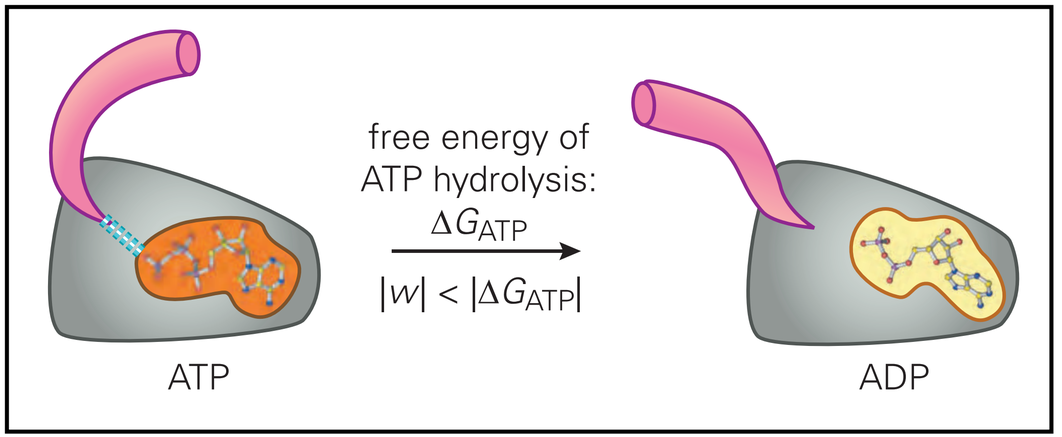
$$\Delta G = \int_{\text{reactants}}^{\text{products}} dG = G(\text{products}) - G(\text{reactants}) = G(\text{ADP} + P_i) - G(\text{ATP} + H_2O)$$

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 $\Delta G^0 = -28$ kJ•mol^{-1}

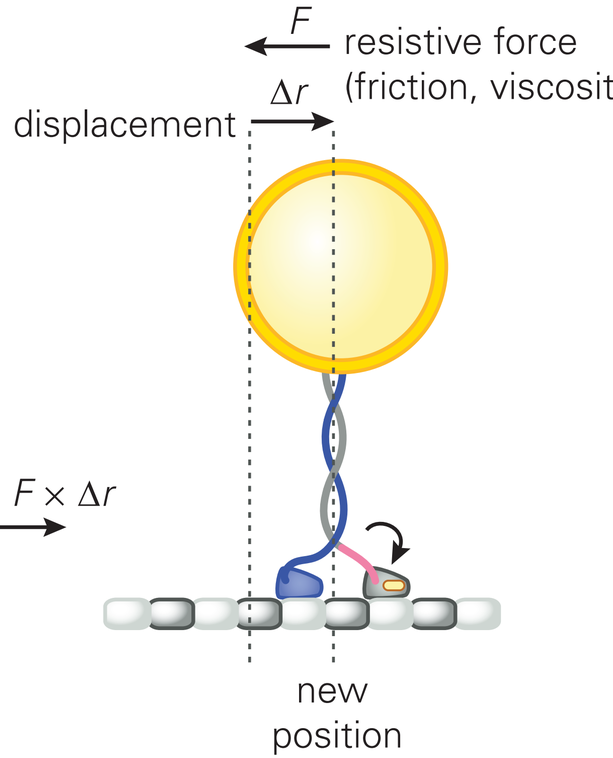


Coupling of ATP hydrolysis to work

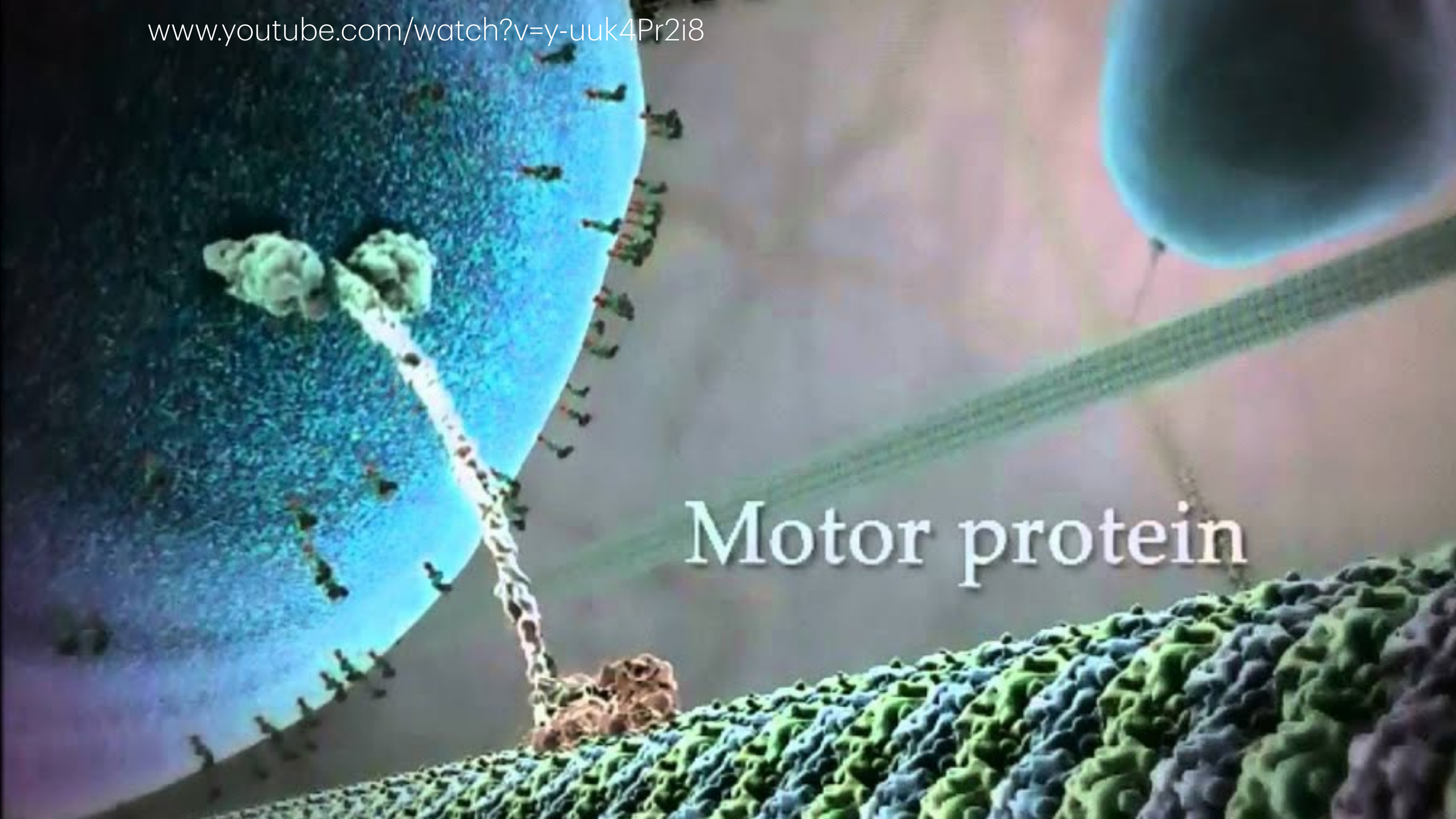




work, $w = F \times \Delta r$



www.youtube.com/watch?v=y-uuk4Pr2i8

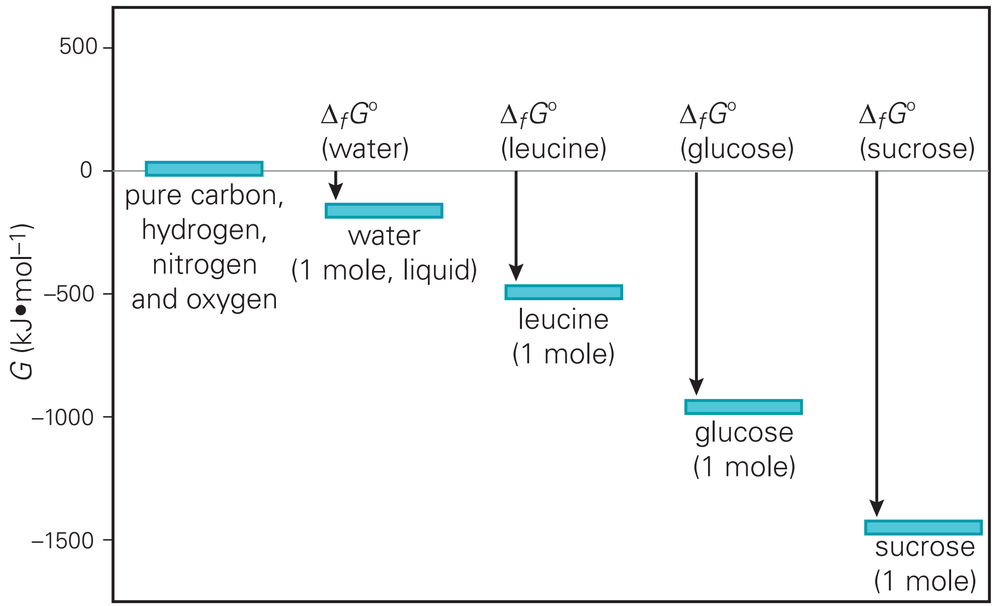


Free Energy of chemical reactions

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

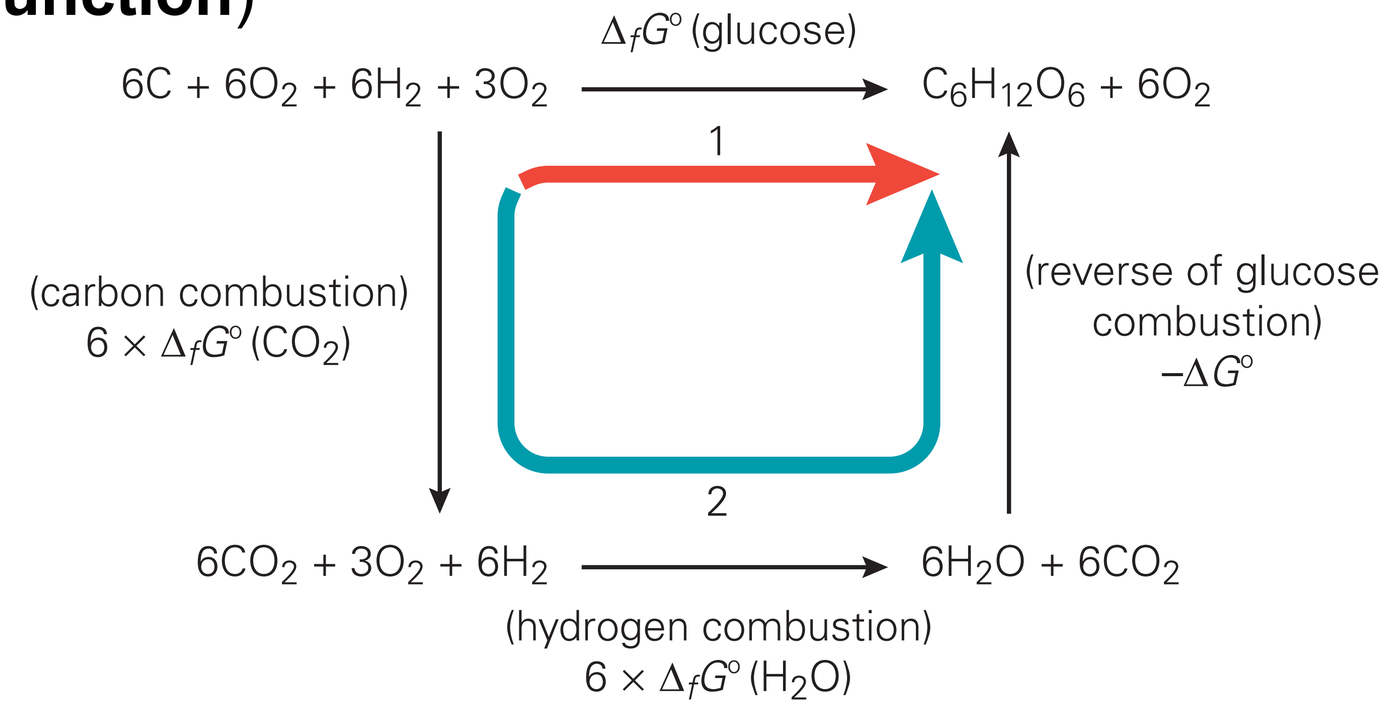
How to calculate these free energies? Using the free energy of formation of the molecules involved in the reactions, starting from the composing elements

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

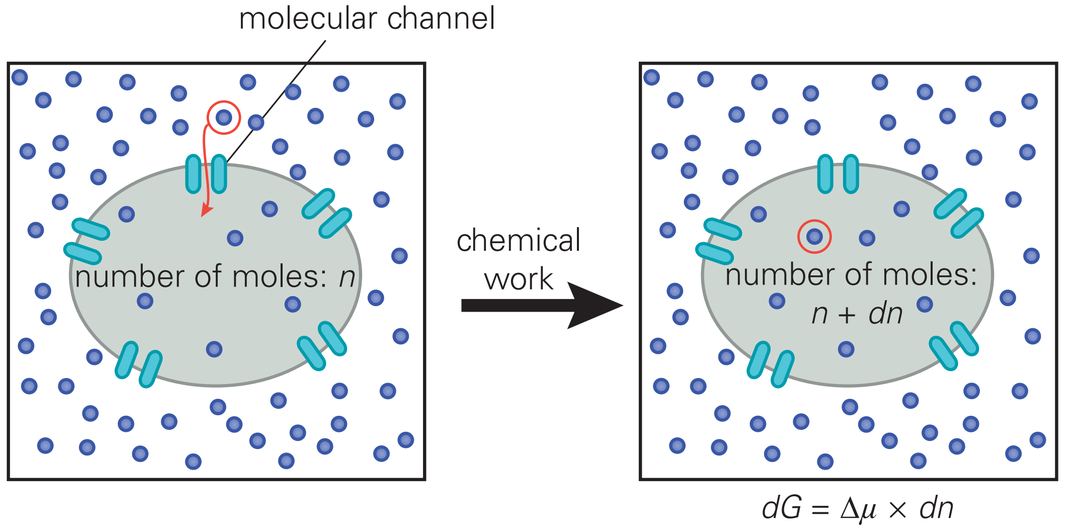


Free Energy of chemical reactions

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**)



Chemical work



- the most important source of free energy for the cell comes from chemical work - e.g. transformation or transfer of matter - changes in the free energy of the system that result from changes in the number of molecules.

$$\Delta G = \int \Delta\mu \; dn$$

- the **chemical potential** takes into account when a system loses or gains molecules, or molecules change states.

- chemical potential of a type of molecule is simply the free energy of one mole of these molecules under the specified conditions (units are in fact J/mole)

$$\mu_i = \left(\frac{\partial G}{\partial N_i} \right)_{T,P,N_{i \neq j}} \approx \frac{\Delta G}{\Delta N_i} = G \text{ (for one molecule)}$$

Chemical potential

- it is the thermodynamic quantity that determines the equilibrium conditions, the tendency of a molecule to escape from a region to another and/or to react from a given state to another

- like T is the entropic force driving energy transfer, the chemical potential drives the net transfer or transformation of particles

- from the maximization of entropy, you obtain the **equilibrium condition**:

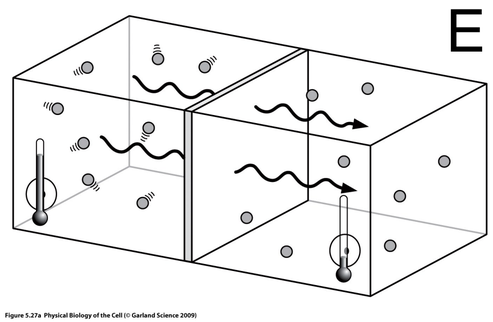


Figure 5.27a Physical Biology of the Cell (© Garland Science 2009)

$$T_A = T_B$$

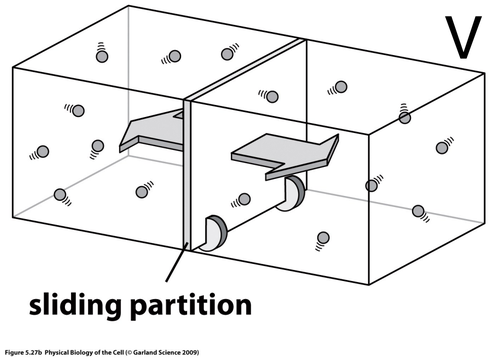


Figure 5.27b Physical Biology of the Cell (© Garland Science 2009)

$$p_A = p_B$$

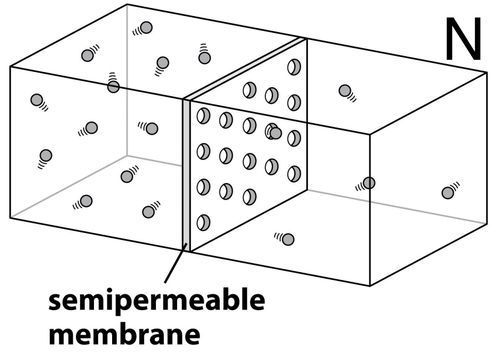


Figure 5.27c Physical Biology of the Cell (© Garland Science 2009)

$$\mu_A = \mu_B$$

Chemical potential driving forces

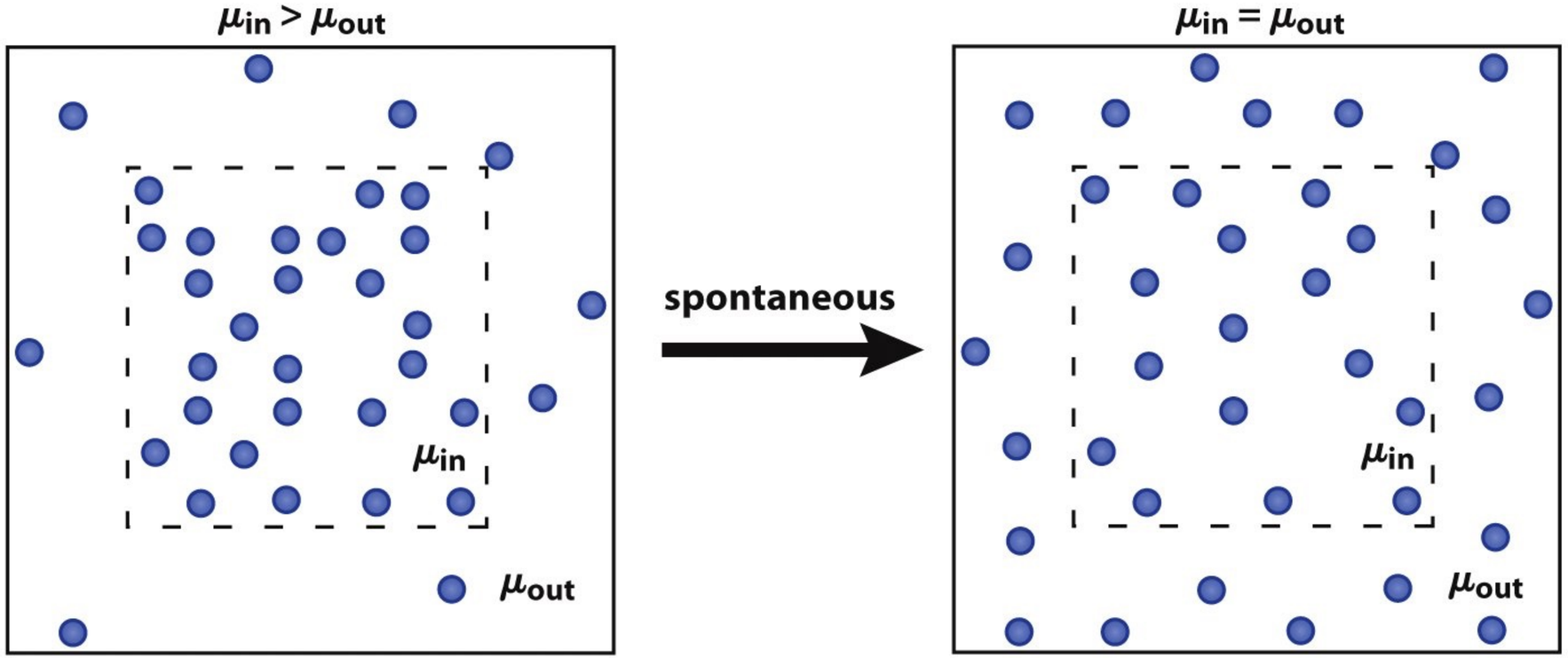


Figure 10.1 The Molecules of Life (© Garland Science 2013)

$$dG = (\mu_{in} - \mu_{out})\, dN_{in} < 0$$

therefore molecules move spontaneously from regions of high chemical potential to regions of low chemical potential

Chemical potential driving forces

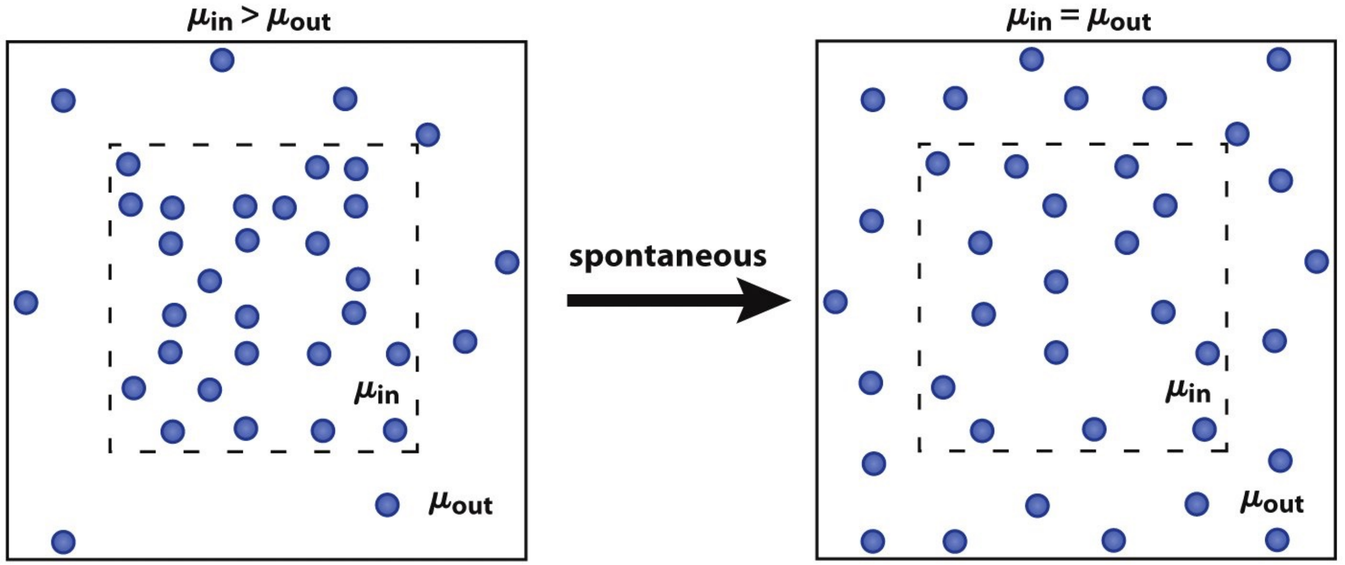
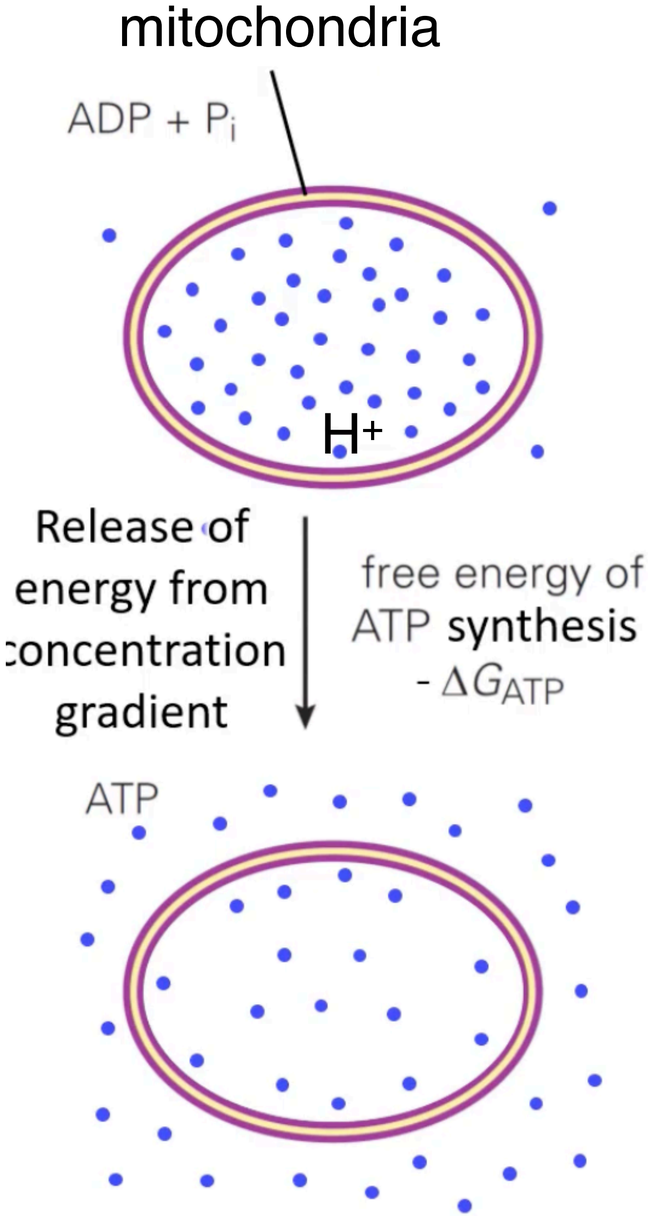


Figure 10.1 The Molecules of Life (© Garland Science 2013)

$$dG = (\mu_{in} - \mu_{out})\, dN_{in} < 0$$

therefore molecules move spontaneously from regions of high chemical potential to regions of low chemical potential



www.youtube.com/watch?v=CSrtewCJbpg

Chemical potential and concentration

- chemical potential of a solute is related to the **logarithm of the concentration** of the solute as a consequence of the positional entropy of the solute
- if you consider a ideal solute solution ($N_{waters} >> N_{molecules}$)

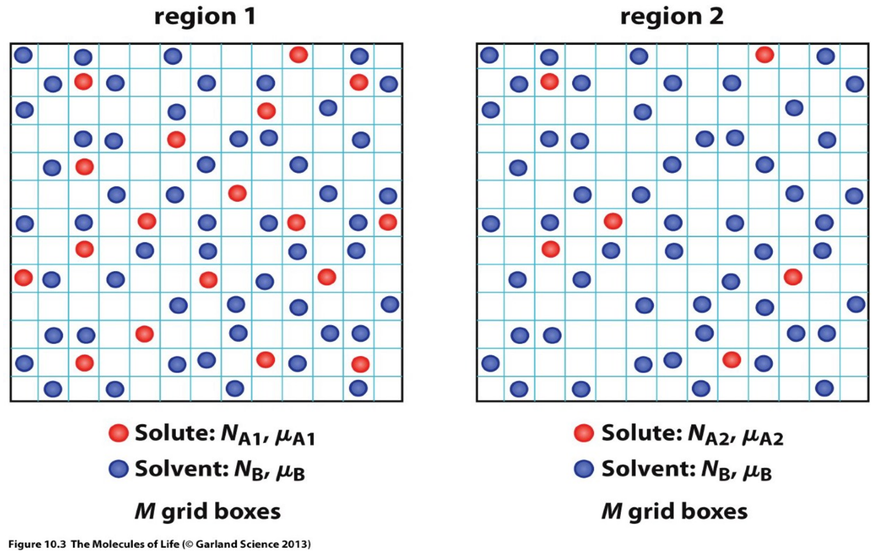


Figure 10.3 The Molecules of Life (© Garland Science 2013)

$$\Delta\mu = \mu_2 - \mu_1 = RT\ln\left(\frac{C_2}{C_1}\right)$$

we expect molecules to move spontaneously from region 1 (high concentration) to region 2 (low concentration), which is in the direction of of decreasing chemical potential.

in reality the chemical potential is also always defined with respect to standard conditions

$$\mu = \mu^\circ + RT\ln\left(\frac{C}{C^\circ}\right) = \mu^\circ + RT\ln C$$

Equilibrium Constants

for a generic reaction $\upsilon_A A + \upsilon_B B \rightleftharpoons \upsilon_C C + \upsilon_D D$ (eg $ATP + H_2O \rightarrow ADP + P_i$)

where $\boldsymbol{\nu}_i$ are the stoichiometric coefficients, imposing the free energy is at the minimum ie dG=0, equilibrium conditions, you obtain that

$$dG = \mu_A\, dn_A + \mu_B\, dn_B + \mu_C\, dn_C + \mu_D\, dn_D = 0 \quad \text{thus} \quad \boxed{\upsilon_A\,\mu_A + \upsilon_B\,\mu_B = \upsilon_C\,\mu_C + \upsilon_D\,\mu_D}$$

using $dN_i = \nu_i d\xi$ ξ :: reaction coordinate

Therefore, isolating the chemical potential at standard conditions and using the relationship between chemical potential and concentration:

$$\upsilon_C\,\mu_C^o + \upsilon_C\,\mu_D^o - \upsilon_A\,\mu_A^o - \upsilon_B\,\mu_B^o = -RT\ln\frac{[C]_{eq}^{\upsilon_C}\,[D]_{eq}^{\upsilon_D}}{[A]_{eq}^{\upsilon_A}\,[B]_{eq}^{\upsilon_B}} = \Delta G^o$$

Defining the **equilibrium constant *K_{eq}***

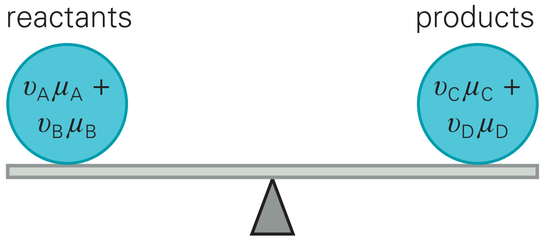
$$K_{eq} = \frac{[C]_{eq}^{\upsilon_C}\,[D]_{eq}^{\upsilon_D}}{[A]_{eq}^{\upsilon_A}\,[B]_{eq}^{\upsilon_B}} \quad \text{we obtain that} \quad \boxed{\Delta G^o = -RT\ln K_{eq} \quad \text{and} \quad K_{eq} = e^{-\left(\frac{\Delta G^o}{RT}\right)}}$$

Equilibrium Constants

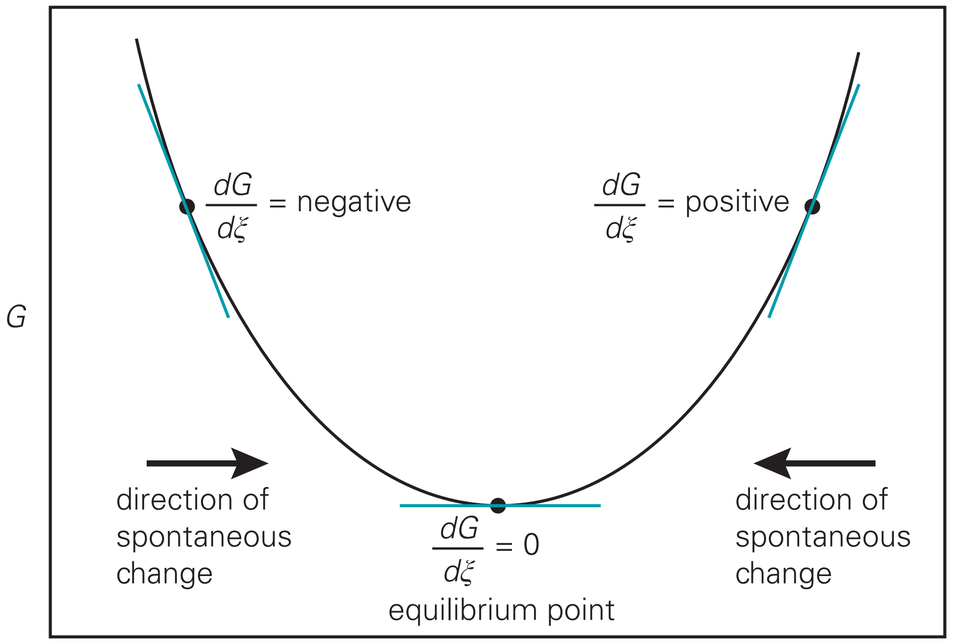
Remember that the equilibrium constant K_{eq} is adimensional because also in the case the stoichiometric coefficients do not cancel out - all concentrations are refereed with respect to the standard conditions

$$K_{eq} = \frac{\left(\frac{[C]_{eq}}{[C]^o}\right)^{\upsilon_C}\left(\frac{[D]_{eq}}{[D]^o}\right)^{\upsilon_D}}{\left(\frac{[A]_{eq}}{[A]^o}\right)^{\upsilon_A}\left(\frac{[B]_{eq}}{[B]^o}\right)^{\upsilon_B}}$$

$$\Delta G^o = -RT \ln K_{eq}$$



pictorial representation
on how to reach equilibrium ->



$$dN_i = \nu_i d\xi$$

ξ:: reaction coordinate

Example of ATP hydrolysis

$$ATP + H_2O \rightarrow ADP + P_i \qquad \Delta G^o = -28\ kJ \bullet mol^{-1}$$

Equilibrium constant (water is neglected as concentration does not change):

$$K = \frac{[ADP][P_i]}{[ATP]}$$

We can compute K using $K_{eq} = e^{-\left(\frac{\Delta G^o}{RT}\right)}$

$$K = e^{+28/2.478} = e^{+11.3} \approx \left(10^{0.43}\right)^{11.3} \approx 10^5$$

Assuming that $[P_i]$ is constant 10^{-2} M (10 mM within physiological levels)

$$\frac{[ADP]}{[ATP]} = 10^7$$

What do you think its the extent of this reaction ?

Example of ATP hydrolysis

$$ATP + H_2O \rightarrow ADP + P_i \qquad \Delta G^o = -28\ kJ \bullet mol^{-1}$$

Equilibrium constant (water is neglected as concentration does not change):

$$K = \frac{[ADP][P_i]}{[ATP]}$$

We can compute K using $K_{eq} = e^{-\left(\frac{\Delta G^o}{RT}\right)}$

$$K = e^{+28/2.478} = e^{+11.3} \approx \left(10^{0.43}\right)^{11.3} \approx 10^5$$

Assuming that $[P_i]$ is constant 10^{-2} M (10 mM within physiological levels)

$$\frac{[ADP]}{[ATP]} = 10^7$$

Thus at equilibrium almost all the ATP will indeed have converted to ADP, except for roughly 1 part per 10 million. The substantial free energy difference between ATP and its hydrolysis products does indeed **drive the reaction nearly to completion.**

Mass action ratio

If we are far from equilibrium, dealing with observed concentrations not at equilibrium conditions, how do we calculate ΔG?

$$\Delta G = \left[\upsilon_C \mu_C^o + \upsilon_D \mu_D^o - \upsilon_A \mu_A^o - \upsilon_B \mu_B^o \right] + RT \ln \frac{[C]_{obs}^{\upsilon_C} [D]_{obs}^{\upsilon_D}}{[A]_{obs}^{\upsilon_A} [B]_{obs}^{\upsilon_B}}$$

$$\Rightarrow \boxed{\Delta G = \Delta G^o + RT \ln Q}$$

and we can define the **reaction quotient** as Q:

$$Q = \frac{[C]_{obs}^{\upsilon_C} [D]_{obs}^{\upsilon_D}}{[A]_{obs}^{\upsilon_A} [B]_{obs}^{\upsilon_B}}$$

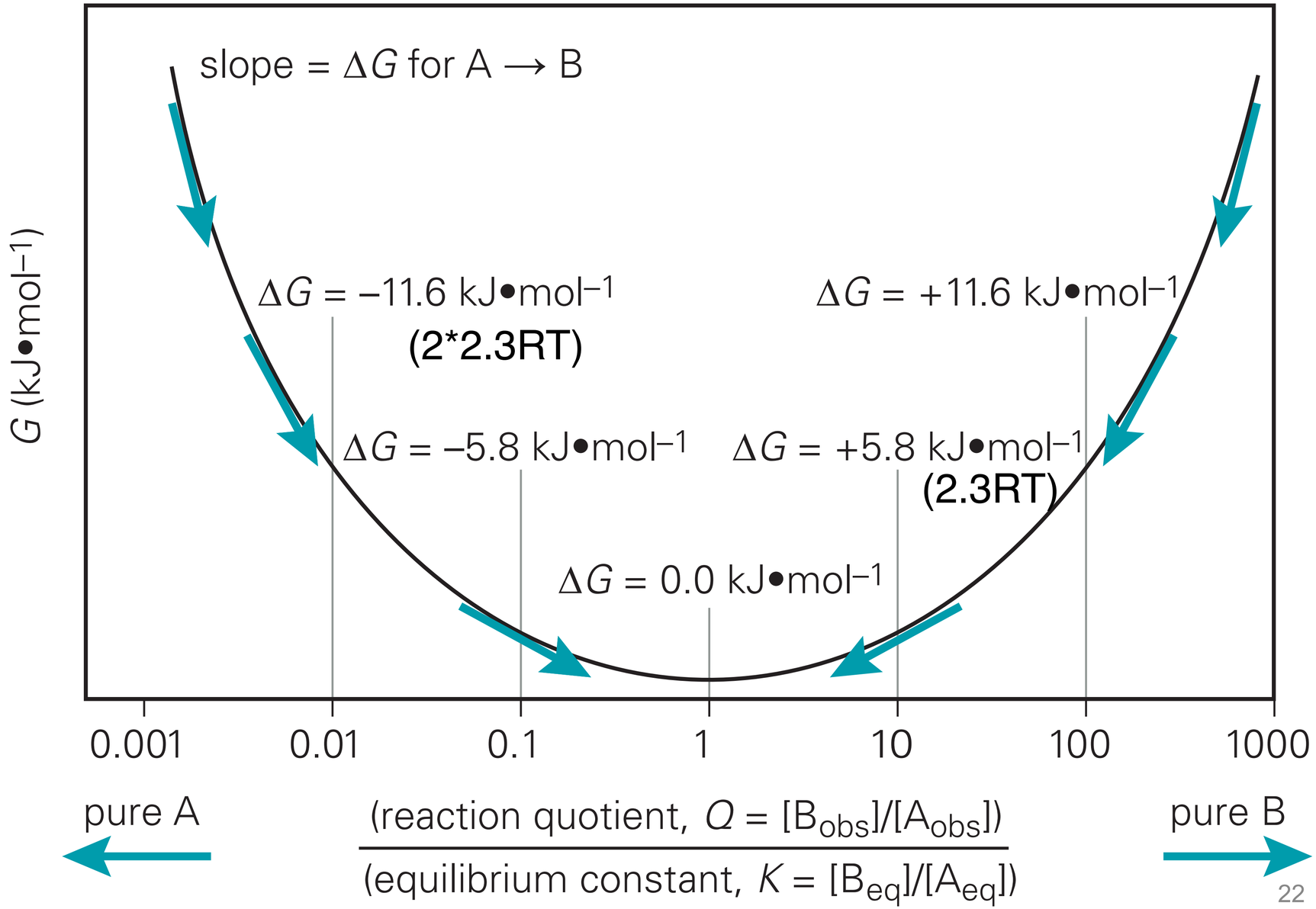
introduction the equilibrium constant for better comparison:

$$\Delta G = \Delta G^o + RT \ln Q = -RT \ln K + RT \ln Q$$

$$\Rightarrow \boxed{\Delta G = +RT \ln \frac{Q}{K} = +2.3RT \log_{10} \frac{Q}{K}}$$

the **ratio Q/K, the mass action ratio** gives us a way to understand where a reaction is going. If Q/K < 1 implies that ΔG < 0 and the reaction is going forward. When Q=K the reaction is at equilibrium.

$$\Delta G = +RT\ln\frac{Q}{K} = +2.3RT\log_{10}\frac{Q}{K}$$



However, in living cells, the ratio of ATP to ADP is held nearly constant, the reaction does not go to completion because the action of metabolic reactions that consume sugars and other sources of energy used to synthesize ATP.

$$\mathrm{ATP} + \mathrm{H_2O} \rightarrow \mathrm{ADP} + \mathrm{P_i} \qquad \Delta G^{o} = -28\ \mathrm{kJ \bullet mol^{-1}}$$

$$K = \frac{[\mathrm{ADP}][\mathrm{P_i}]}{[\mathrm{ATP}]} \qquad \frac{[\mathrm{ADP}]}{[\mathrm{ATP}]} = 10^{7}$$

Q	*Q*/*K*	ΔG (kJ•mol^{-1})	[ATP]/[ADP]	Relevant condition
10^{5}	1	0	10^{-7}	equilibrium
10^{3}	10^{-2}	–11	10^{-5}	
1	10^{-5}	–28	10^{-2}	standard condition
10^{-3}	10^{-8}	–46	10^{1}	mitochondrial matrix
10^{-5}	10^{-10}	–57	10^{3}	cytoplasm

For instance ATP concentrations in the cytoplasm are held so that [ATP]/[ADP] is ~1000, storing more free energy than in standard conditions. Thus, the ATP hydrolysis reaction can be coupled to other reactions that are uphill in terms of free energy, making such transformations possible.

Acid-base equilibrium

Acids dissociate in water to release protons

$$HA \rightleftharpoons H^+ + A^-$$

Acid dissociation constant:

$$K_a = \frac{[H^+][A^-]}{[HA]}$$

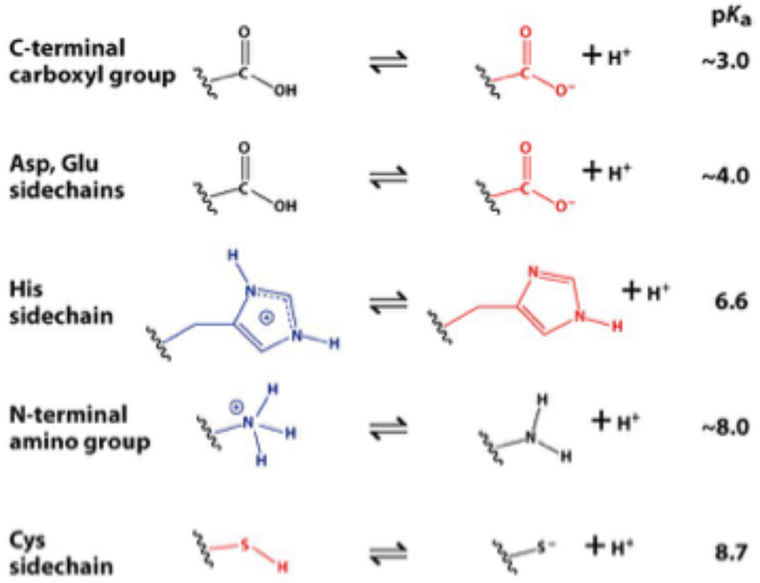
Applying the $\log_{10}$ and recalling that $pH = -\log_{10}([H^+])$ and $pK_a = -\log_{10}(K_a)$

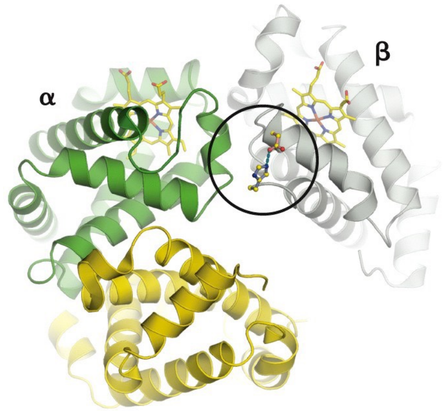
We obtain the Henderson-Hasselbalch equation:

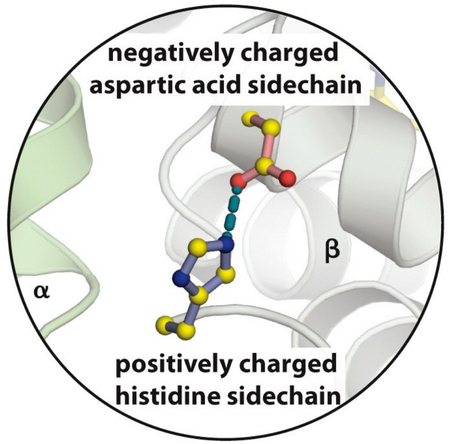
$$pH = pK_a + \log_{10}\left(\frac{[A^-]}{[HA]}\right)$$

We will revive it in the exercise session

Acid-base equilibrium and protonation states







With the **Henderson-Hasselbalch** equation and a few concepts from the **free energy and equilibria** you will be able to compute pKas for residues within the protein environment (see exercises)

His for example has a pKa of 6 and at pH 7 we have that $\frac{[\mathrm{His}]}{[\mathrm{His}^+]} = 10$
and thus $\Delta G^{o\prime} = -RT\ln K = -8.314 \times 300 \times \ln 10 = -5.7\ \mathrm{kJ \bullet mol^{-1}}$
for the deprotonation reaction of His ($His^+ \longrightarrow His + H^+$).
If His is close to an Asp in a protein environment its pKa can raise up !

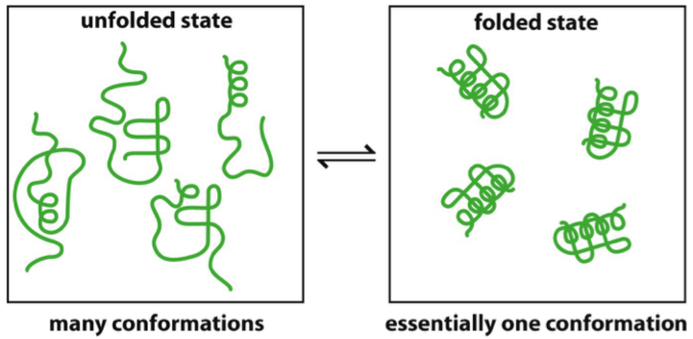
ex. page 437 (TMOL)

Free energy change in protein folding

Two concepts we are going to use to describe protein folding

$$K = e^{-\Delta G^{o}/RT} \qquad \Delta G^{o} = \Delta H^{o} - T\Delta S^{o}$$

Protein folding reaction can be simplified as transition from the unfolded state (U) to the folded state (F).



Thus that the folding reaction is

$$\mathrm{U} \rightleftharpoons \mathrm{F}$$

$$K_{\text{folding}} = \frac{[\mathrm{F}]}{[\mathrm{U}]} \quad (\text{at equilibrium})$$

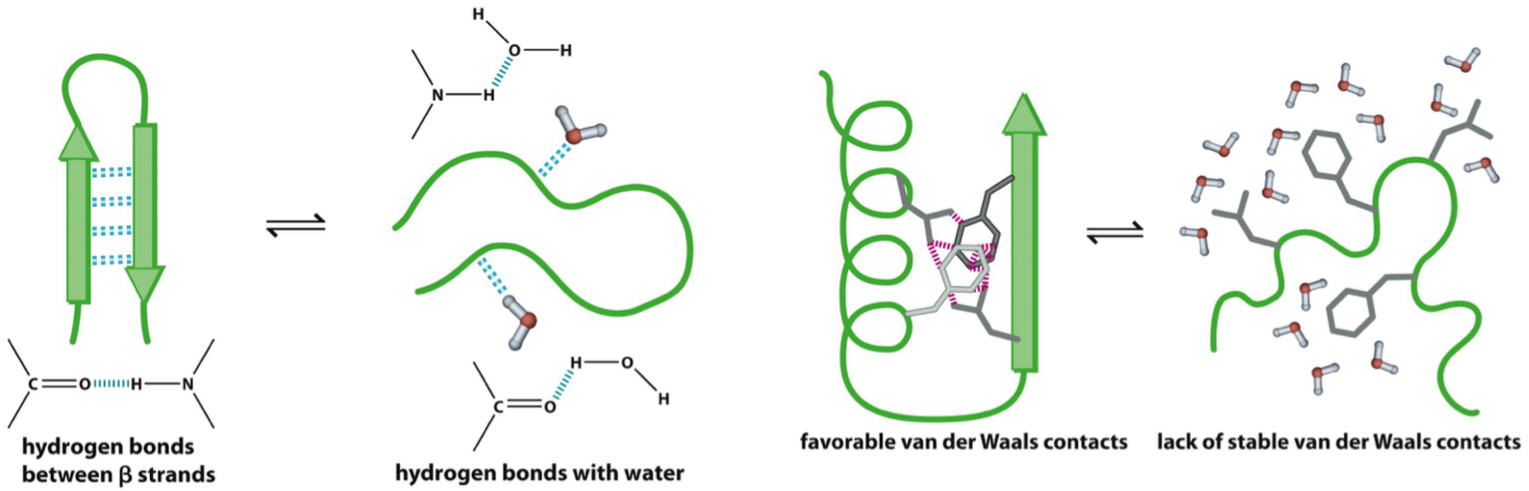
for the unfolding reaction we have $K_{\text{unfolding}} = \dfrac{[\mathrm{U}]}{[\mathrm{F}]} = K_{\text{folding}}^{-1}$

Energetic components in protein folding

Protein folding folding result rom a balance between enthalpy and entropy

$$\Delta G^o = \Delta H^o - T\Delta S^o$$

As you know there are several types of molecular interactions in a folded protein



Which component gives a higher contribution and why ?

a) Enthalpy

b) Entropy

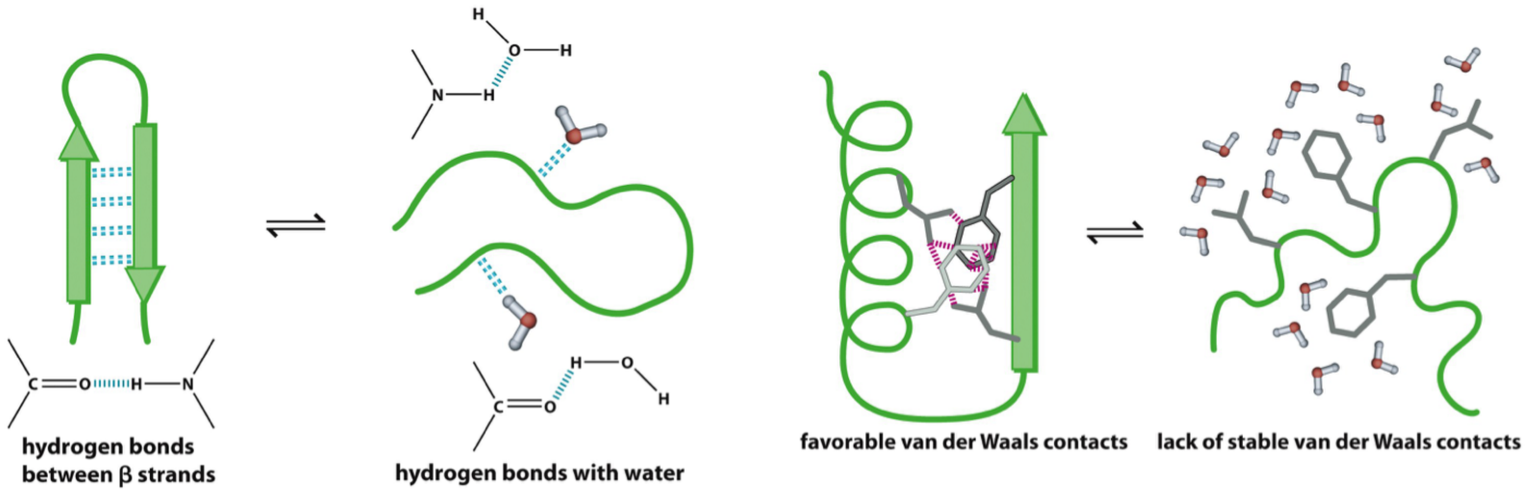
c) Both the same

Energetic components in protein folding

Protein folding folding result rom a balance between enthalpy and entropy

$$\Delta G^o = \Delta H^o - T\Delta S^o \quad < 0$$

As you know there are several types of molecular interactions in a folded protein



Enthalpy: H-bonds do not have a major effect, in U many with water are replaced in F within the protein. vdW interactions are more favourable in F but are not so strong - overall you get a small favourable H balance ($\Delta H<0$)

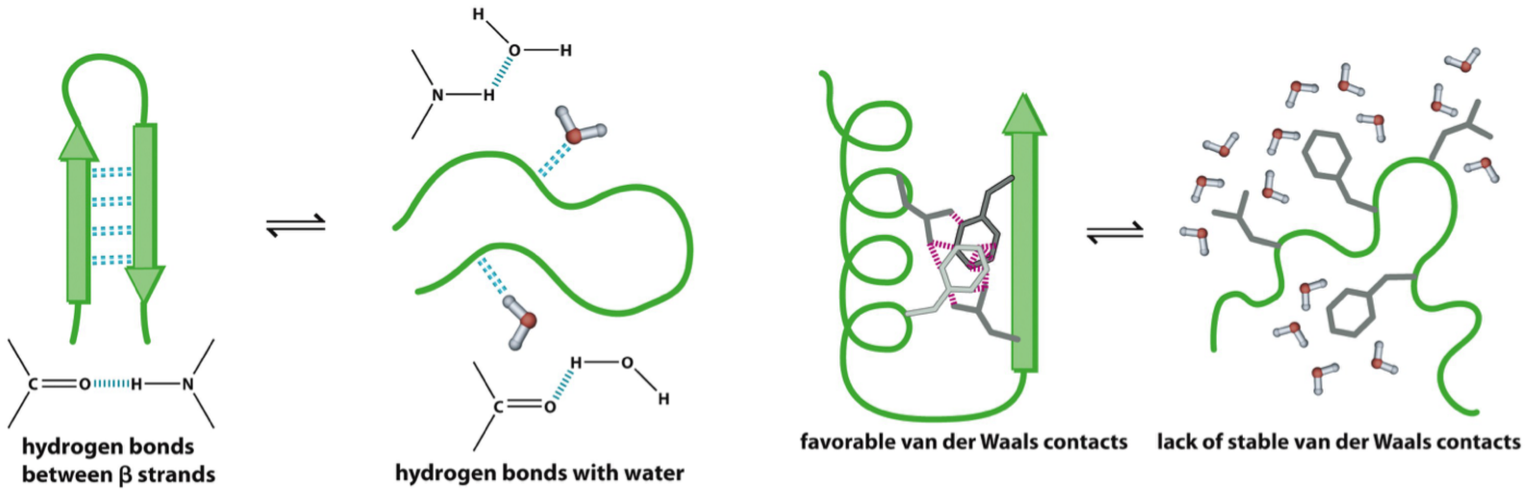
This can be around ΔH = -150 kJ/mol for a small size protein (~50 aa)

Energetic components in protein folding

Protein folding folding result rom a balance between enthalpy and entropy

$$\Delta G^{o} = \Delta H^{o} - T\Delta S^{o} \quad < 0$$

As you know there are several types of molecular interactions in a folded protein



Entropy: U state has high entropy (many different unfolded configurations).
The F state has low entropy (only one configuration, S=0)

$$\Delta S = R \ln N_C \qquad \Delta S^{o} = 0 - S^{o}_{\text{unfolded}} = -S^{o}_{\text{unfolded}}$$

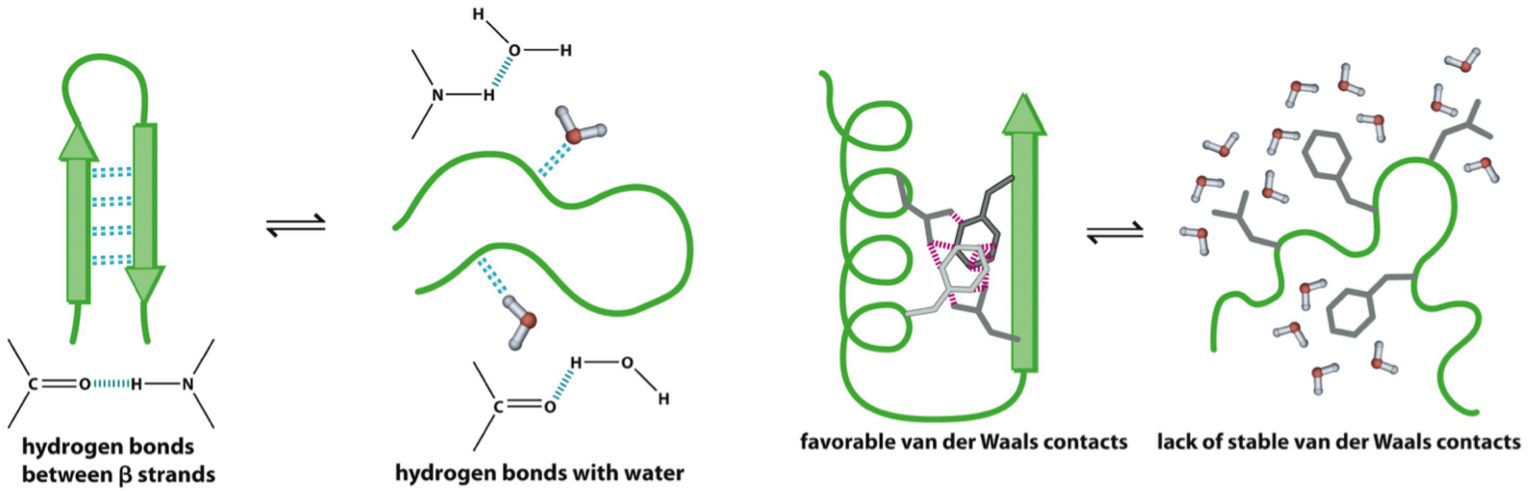
Thus $-T\Delta S^{0}$ can be around +200 kJ/mol for a small size protein

Energetic components in protein folding

Protein folding folding result rom a balance between enthalpy and entropy

$$\Delta G^o = \Delta H^o - T\Delta S^o$$

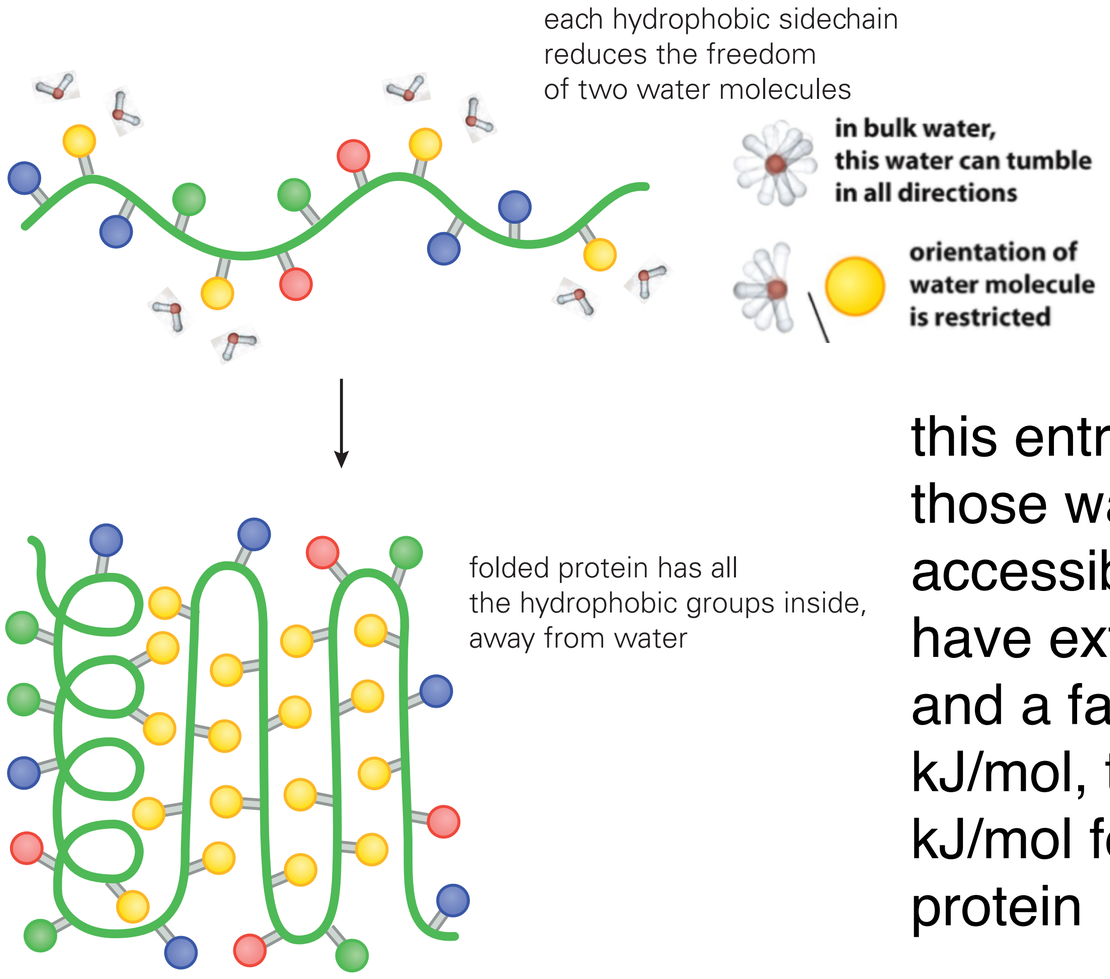
As you know there are several types of molecular interactions in a folded protein



Thus the overall ΔG is positive, ΔH - TΔS = -150 - (-200) ~50 kJ/mol, which is not what we expect as the protein actually folds, ie ΔG < 0.
Something is missing - namely the hydrophobic effect

Energetic components in protein folding

Increase of entropy on the solvent (e.g. water) upon folding due to the so-called **hydrophobic effect,** provides another favorable term to the free energy of folding

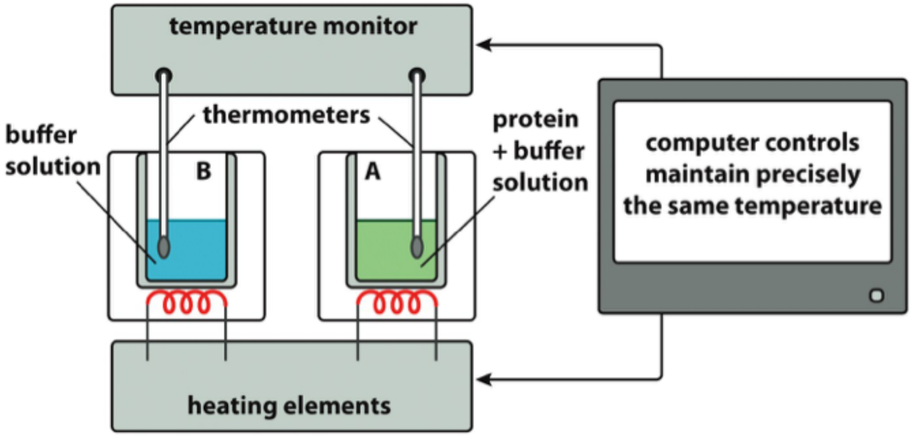


this entropy contributions is low in U state

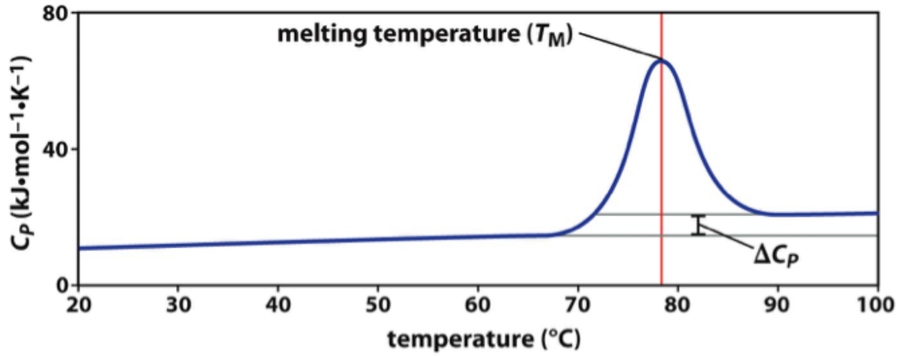
this entropy contribution is hight as those waters can have more accessible conformations, thus we have extra $\Delta S > 0$ term from water and a favourable term $-T\Delta S = -100$ kJ/mol, thus overall the $\Delta G \sim -50$ kJ/mol for our exemplary small protein

Thermodynamic properties of proteins

Enthalpic and entropic energy contributions for protein folding can be in fact measured by using calorimeters, which measure the heat capacity as we have seen



Instrument :
Differential Scanning Calorimeter



Data: Melting Curve

Temperature scan of the sample

- Measure heat capacity associated to protein folding
- Determines the melting temperature

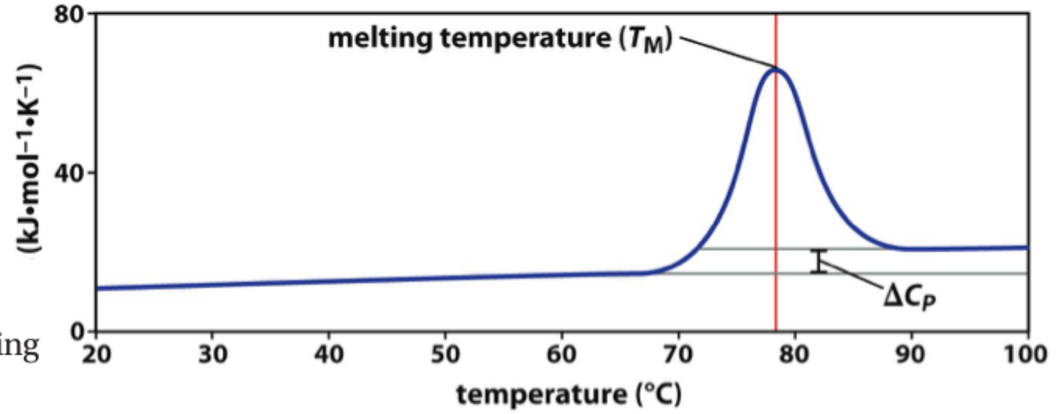
Thermodynamic properties of proteins

with calorimetry we follow the unfolding reaction

$$F \rightleftharpoons U$$

$$K_{\text{unfolding}} = \frac{[U]}{[F]}$$

$$\Delta G^{o}_{\text{unfolding}} = \Delta H^{o}_{\text{unfolding}} - T\Delta S^{o}_{\text{unfolding}}$$

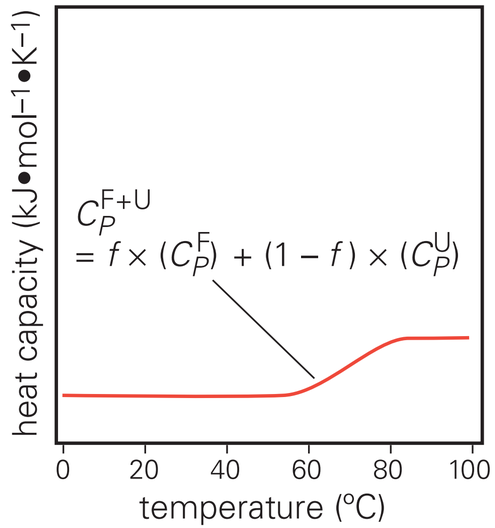


$\Delta C_p > 0$

As the temperature increases more of the protein unfolds, until at a certain temperature (T_M - melting temperature), the ΔG^o becomes zero (remember that $\Delta G^o = -RT \ln K_{eq}$)

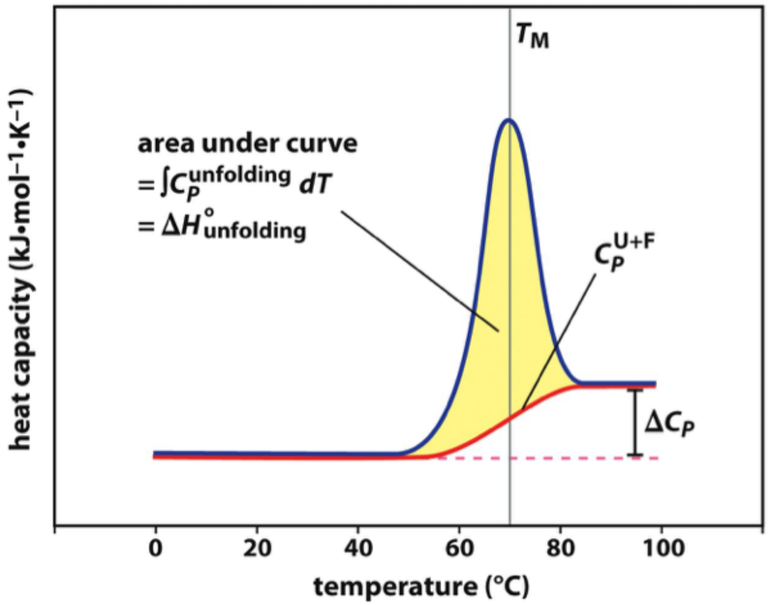
The measured heat capacity includes contributions of both the folded population(f) and the unfolded population(1-f)

$$C_P^{F+U} = f\, C_P^{F} + (1-f)\, C_P^{U}$$



Thermodynamic properties of proteins

But looking at the data something is missing which is the peak reporting the heat capacity due to the interactions that are broken passing from the folded to unfold state



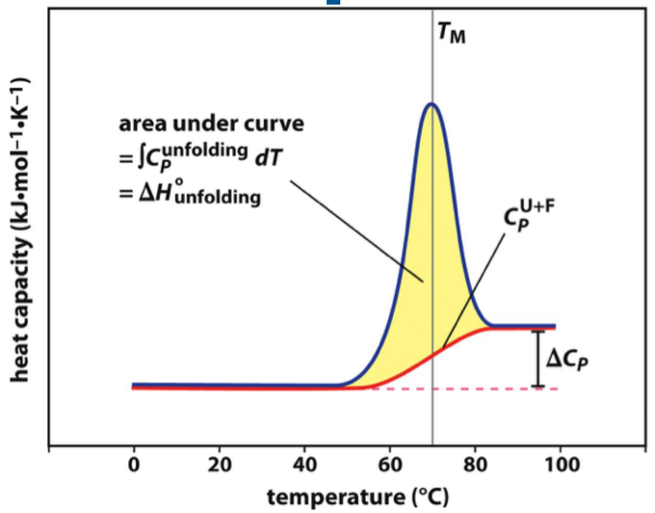
$$C_P^{\text{observed}} = C_P^{\text{F+U}} + C_P^{\text{unfolding}} = f\ C_P^{\text{F}} + (1-f)\ C_P^{\text{U}} + C_P^{\text{unfolding}}$$

Thus to estimate only the energy that accounts for the unfolding process we have to integrate $C_p^{\text{unfolding}}$ obtaining ΔH

$$\int C_P^{\text{unfolding}}\ dT = H_{\text{unfolded}} - H_{\text{folded}} = \Delta H_{\text{unfolding}}$$

Thermodynamic properties of proteins

$$\int C_P^{\text{unfolding}}\, dT = H_{\text{unfolded}} - H_{\text{folded}} = \Delta H_{\text{unfolding}}$$



At melting temperature $K(\text{at } T_M) = \dfrac{[U]}{[F]} = 1$ thus $\Delta G^o(T_M) = -RT \ln K(T_M) = 0$

if the free energy is zero we can not only estimate the enthalpy contribution to unfolding but also the entropic one - thus **at the melting T we can completely characterise the thermodynamics of folding**, in fact:

$$\Delta G^o_{\text{unfolding}}\ (\text{at } T_M) = \Delta H^o_{\text{unfolding}} - T_M \Delta S^o_{\text{unfolding}} = 0$$

$$\Rightarrow \Delta S^o_{\text{unfolding}} = \frac{\Delta H^o_{\text{unfolding}}}{T_M}$$

Thermodynamic properties of proteins

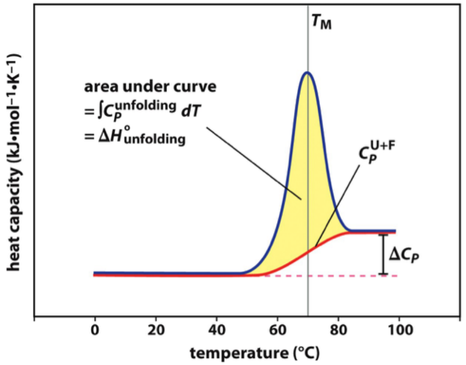
But how to characterise the process
at any given T different from T_M?
For this we can benefit from the fact that

$$\Delta C_P = C_P^{\mathrm{U}} - C_P^{\mathrm{F}}$$

and thus:

$$\Delta H^{\mathrm{o}}_{\mathrm{unfolding}}(T) = \Delta H^{\mathrm{o}}_{\mathrm{unfolding}}(T_{\mathrm{M}}) + \Delta C_P (T - T_{\mathrm{M}})$$

$$\Delta S^{\mathrm{o}}_{\mathrm{unfolding}}(T) = \Delta S^{\mathrm{o}}_{\mathrm{unfolding}}(T_{\mathrm{M}}) + \Delta C_P \ln\left(\frac{T}{T_{\mathrm{M}}}\right)$$



in case ΔC_p was zero, then the equations above would have been independent from the temperature.
$\Delta C_p > 0$ is correlated with the extent of interaction of hydrophobic groups with water

Protein	Molecular weight	ΔH^0 (25°C) [kJ•mol^{-1} per residue]	ΔS^0 (110°C) [J•mol^{-1} per residue]	ΔC_P [J•K^{-1}•mol^{-1} per residue]
Protein G-B1	7200	1.4	16.1	53
Parvalbumin	11,500	1.4	16.8	46
Cytochrome *c*	12,400	0.64	17.8	67
Ribonuclease A	13,600	2.4	17.8	44
Hen lysozyme	14,300	2.0	17.6	52
Staph. nuclease	16,800	0.85	17.5	61
Myoglobin	17,900	0.04	17.9	75
Papain	23,400	0.93	17.0	60
β-Papain	23,800	1.3	17.9	58
α-Chymotrypsin	25,200	1.1	18.0	58
Average		1.2 ± 0.7	17.4 ± 0.6	57 ± 9

What to know...

- Concentrations of reactants and products at the equilibrium constant (K) is related to the free energy change.

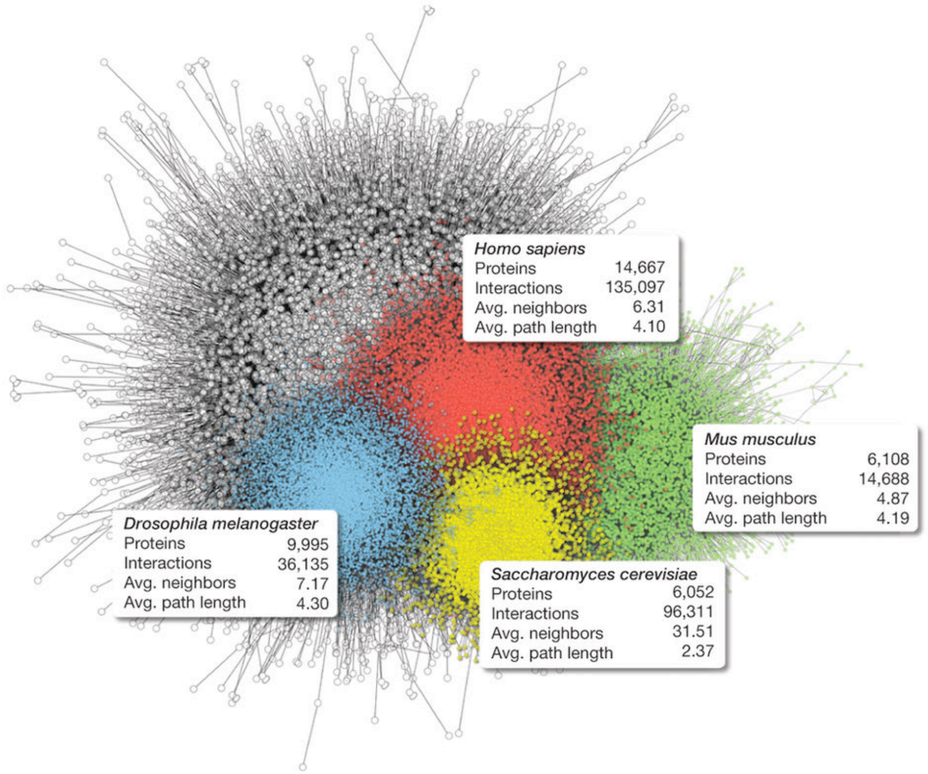
- Don’t forget the Henderson-Hasselbach equation and application to amino acid protonation states !

- Protein folding results from a balance between energy and entropy and the hydrophobic effect is a key feature

- Measurements of the difference in the heat capacity makes it possible to determine the value of ΔH and ΔS for unfolding at any temperature

Why binding is that important ?

protein-protein interaction networks

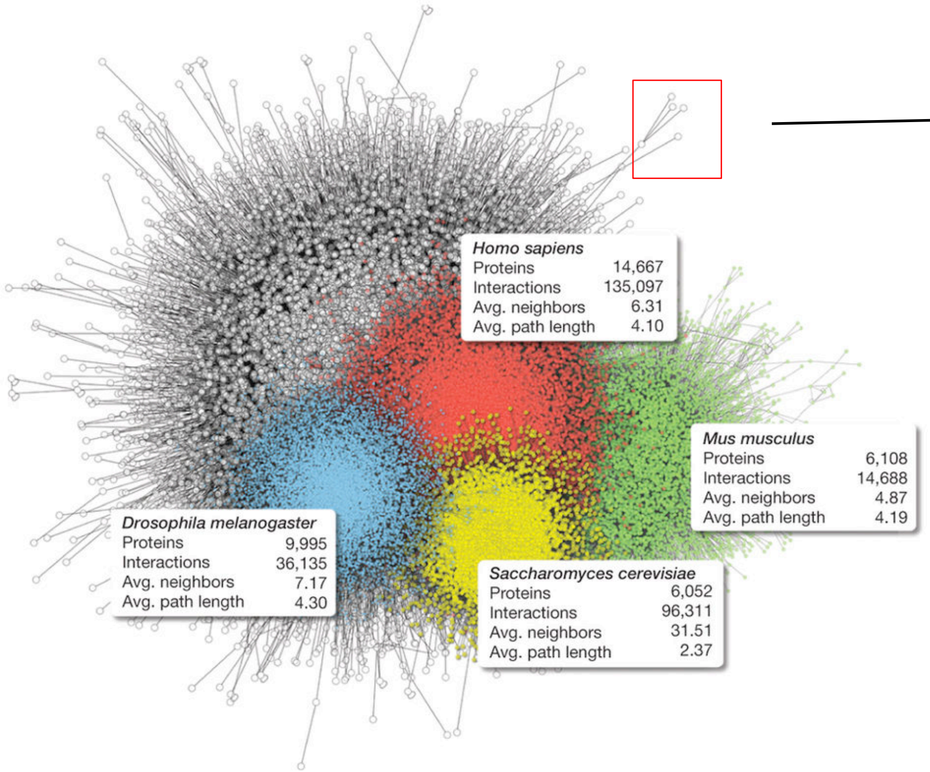


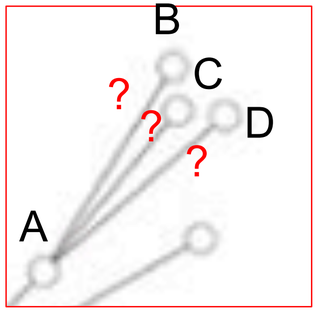
-Dots are proteins, edges represent interactions

You can imagine that such extensive interaction networks are at least as complex for:

- protein-dna interactions
- protein-small molecule interactions
- protein-lipid interactions as well have started to be recognised as very important

Why binding is that important ?



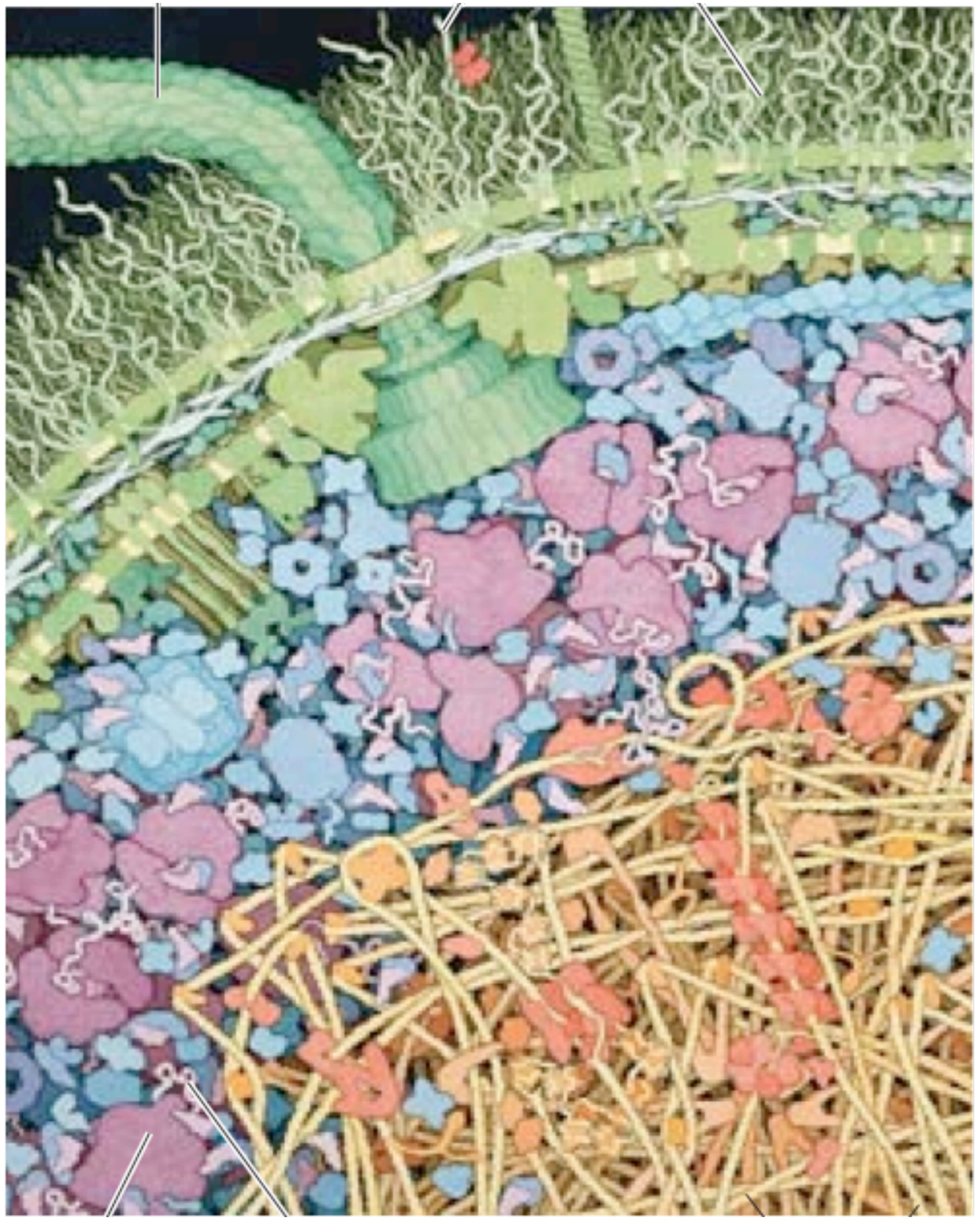


For a **qualitative** description of the this submap its enough to know that A binds B,C and D somehow

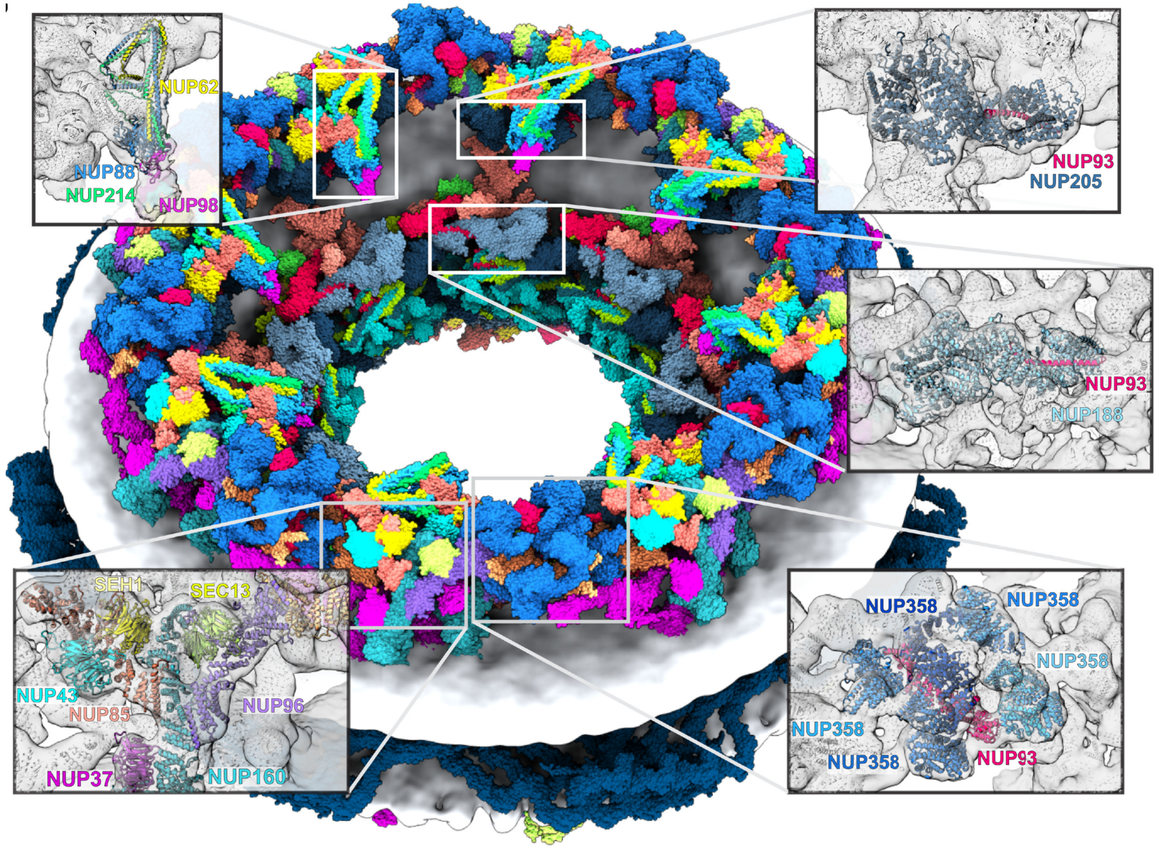
However, for a **quantitative** description it is essential to determine measurable quantities to these edges – i.e. binding affinities

Measuring the concentration of the free and associated species at equilibrium, we can calculate the strength of the molecular interactions.

-Dots are proteins, edges represent interactions

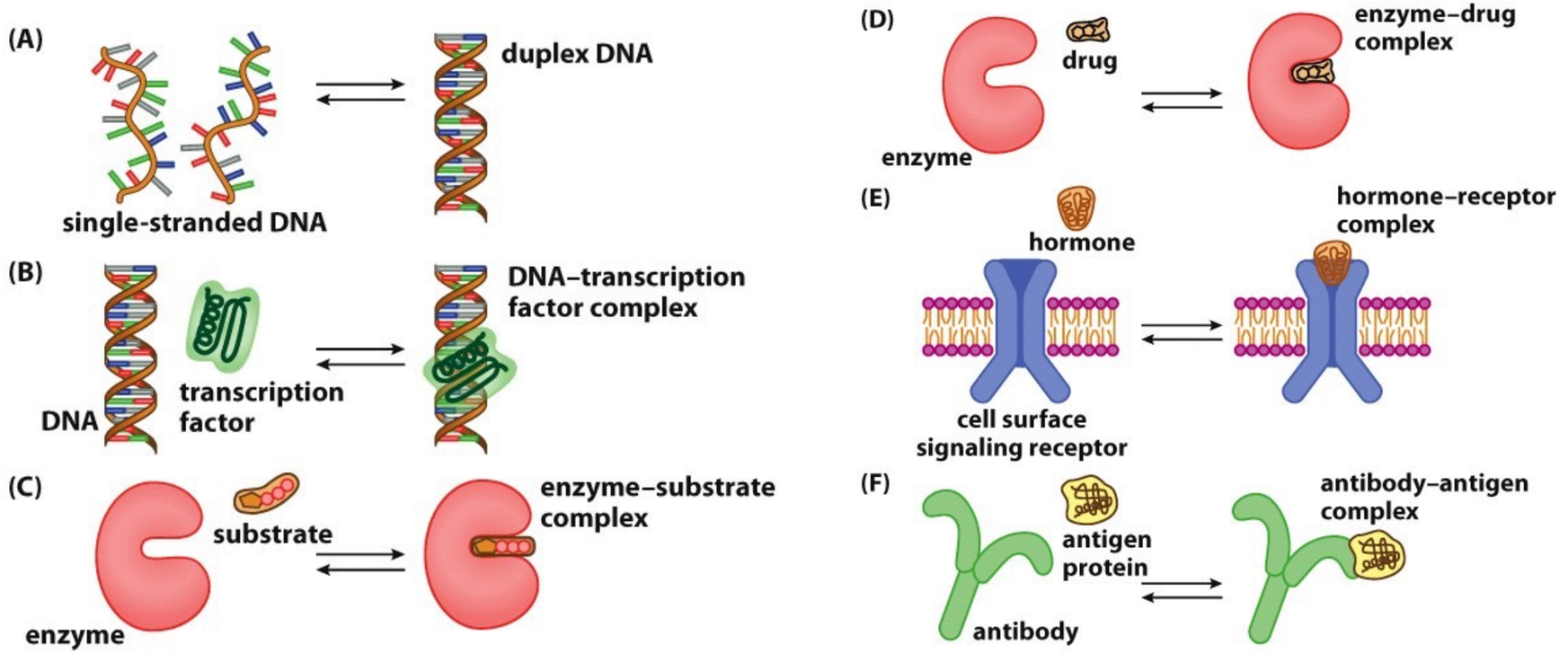


e.g., the Nuclear Pore Complex



Viewing Binding as a Chemical Reaction

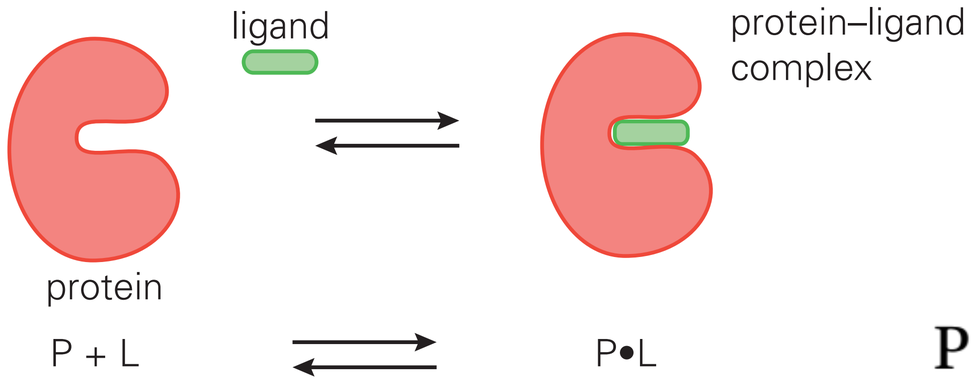
Some examples of important molecular recognition events:



we talked about equilibrium constants for chemical reactions, if in general we consider a protein as target **P** and a ligand **L** as the interacting molecule via non covalent interactions the general complex P•L can be considered the product of the following reaction:

$$P + L \rightleftharpoons P \bullet L$$

Looking at the reaction mechanism



$$P + L \rightleftharpoons P \bullet L$$

Equilibrium constant:

$$K = \frac{[P \bullet L]}{[P][L]}$$

Because it is a binding reaction we call it **association constant:**

$$K_A = \frac{[P \bullet L]}{[P][L]}$$

Thus we can define the **binding fee energy** for the association as:

$$\Delta G^{o}_{bind} = -RT \ln K_A$$

which is a measure of the **affinity** of the interaction, that is, how strongly the molecules bind to each other.

Looking at the reaction mechanism

Energetic Landscape of Binding

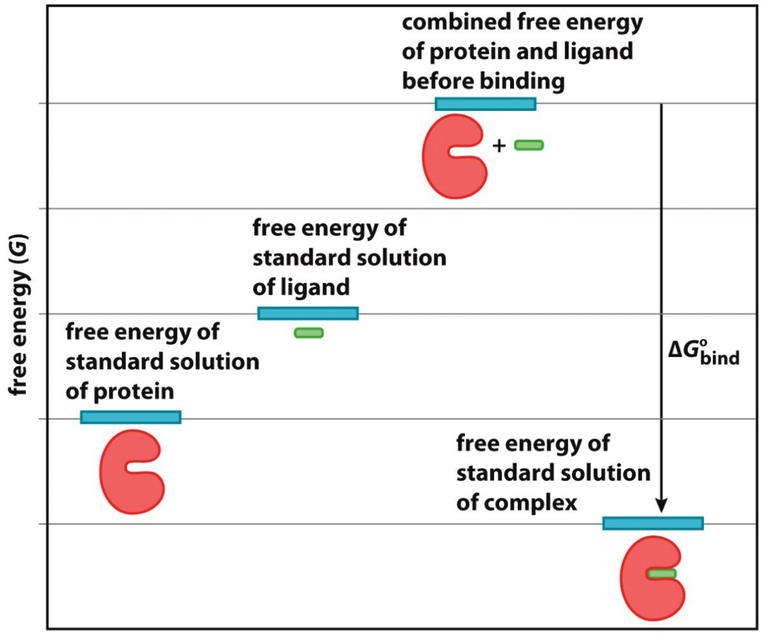


Figure 12.3 The Molecules of Life (© Garland Science 2013)

$$P + L \rightleftharpoons P \bullet L$$

$$K_A = \frac{[P \bullet L]}{[P][L]}$$

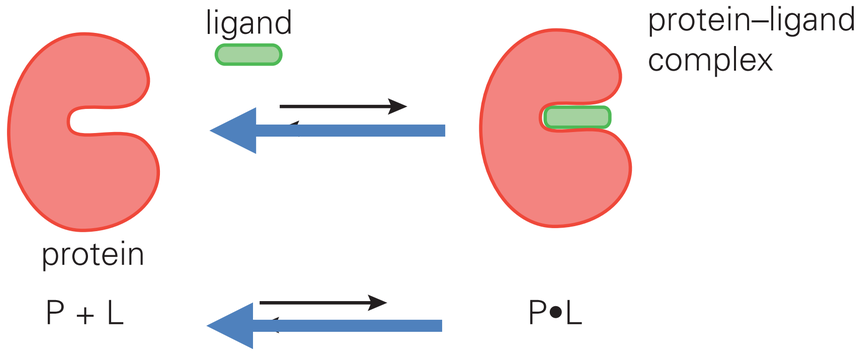
$$\Delta G^{o}_{bind} = -RT \ln K_A$$

Definition of Binding Affinity:

- it refers to the strength of a molecular interaction
- the greater the decrease in free energy upon binding the greater the affinity

Looking at the reaction mechanism

It is however common practice to characterize the affinity of a binding interaction in terms of the **dissociation reaction,** which is associated to the **dissociation constant K_D**



$$P \bullet L \rightleftharpoons P + L$$

$$K_D = \frac{[P][L]}{[P \bullet L]} = \frac{1}{K_A}$$

$$\Delta G^{o}_{bind} = +RT \ln K_D$$

K_D is a dimensionless quantity, but usually it is discussed as if it had molar units of concentration. K_Ds that range from picomolar to nanomolar (10^{-12} -10^{-9} M, ~ -50 kJ/mol) are the tightest interactions, if in the order of millimolar (10^{-3} M, ~ -15 kJ/mol) they are the weakest.

Looking at the reaction mechanism

common practice to express K_D in molar terms because concentration at standard conditions are kind of neglected

$$K_D = \left(\frac{[P \bullet L]^o}{[P]^o [L]^o} \right) \frac{[P][L]}{[P \bullet L]} = \left(\frac{[P \bullet L]^o}{[P]^o [L]^o} \right) K_D^*$$

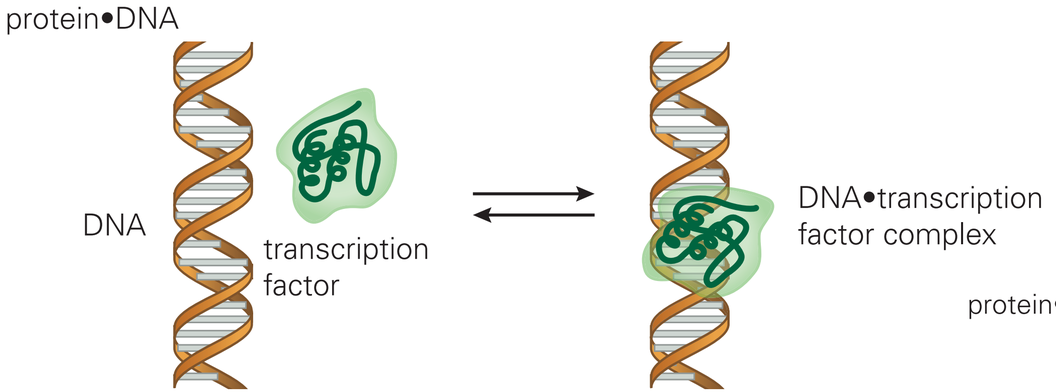
Some examples and affinity ranges :

Type of interaction	K_D (molar)	ΔG^o_{bind} (at 300 K) (kJ•mol^{-1})
Enzyme–ATP	~1 × 10^{-3} to ~1 × 10^{-6} (millimolar to micromolar)	–17 to –35
Signaling protein binding to a target	~1 × 10^{-6} (micromolar)	–35
Sequence-specific recognition of DNA by a transcription factor	~1 × 10^{-9} (nanomolar)	–52
Small molecule inhibitors of proteins (drugs)	~1 × 10^{-9} to ~1 × 10^{-12} (nanomolar to picomolar)	–52 to –69
Biotin binding to avidin protein (one of the strongest known noncovalent interactions)	~1 × 10^{-15} (femtomolar)	–86

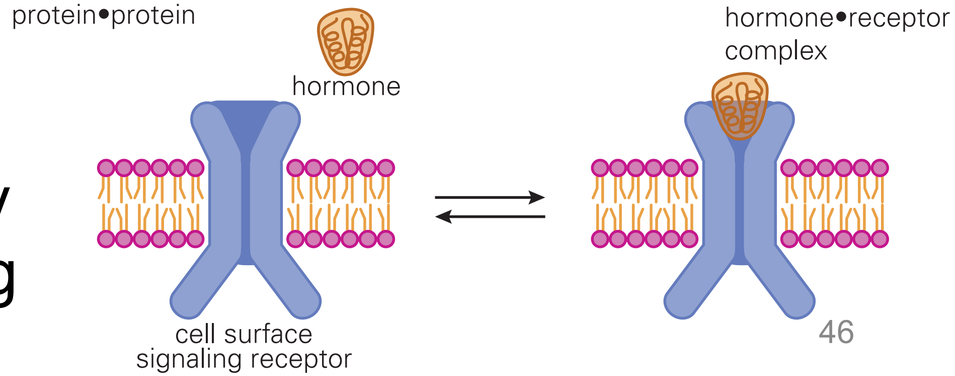
The trade-off between **affinity** and **specificity** is crucial in many biological processes

Affinity vs. specificity

Affinity of an interaction defines its strength (ie K_D).
But affinity alone is most of the times not sufficient as if you bind with high affinity to multiple targets then you will engage with targets that are not specific for a given biological function (ie off-target binding)
Specificity thus is the affinity of the ligand for one and only one specific target of interest

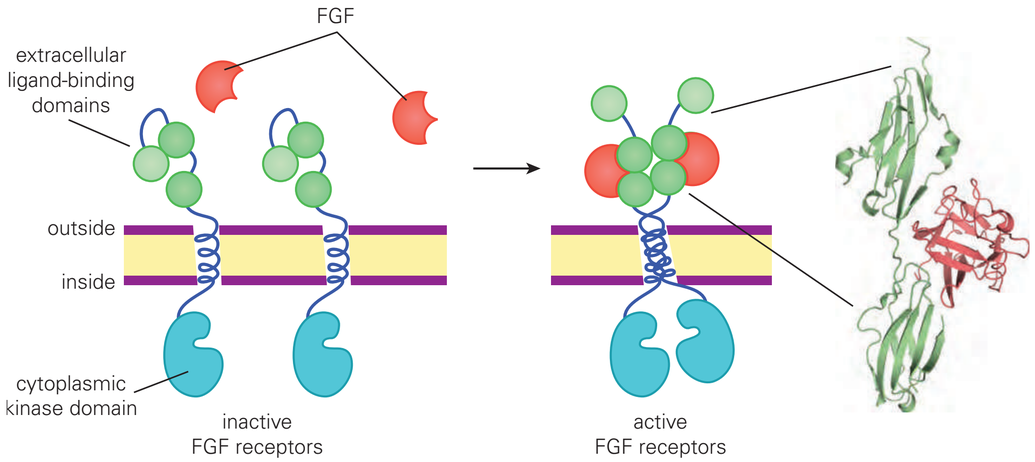


Specificity is mainly achieved by polar, directional interactions (eg H-bonds).

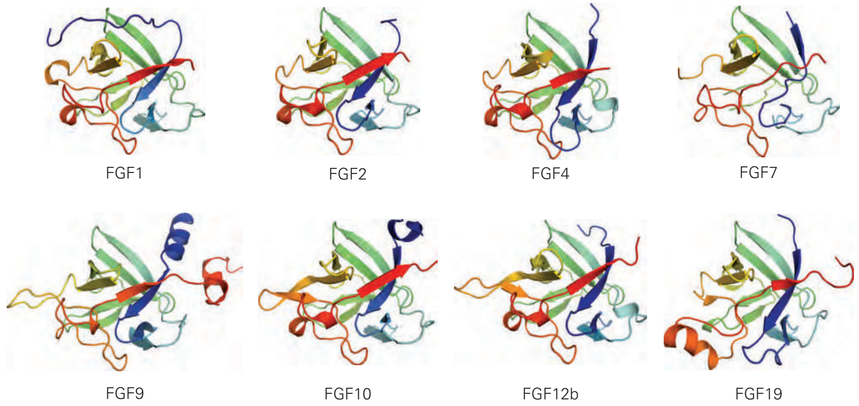


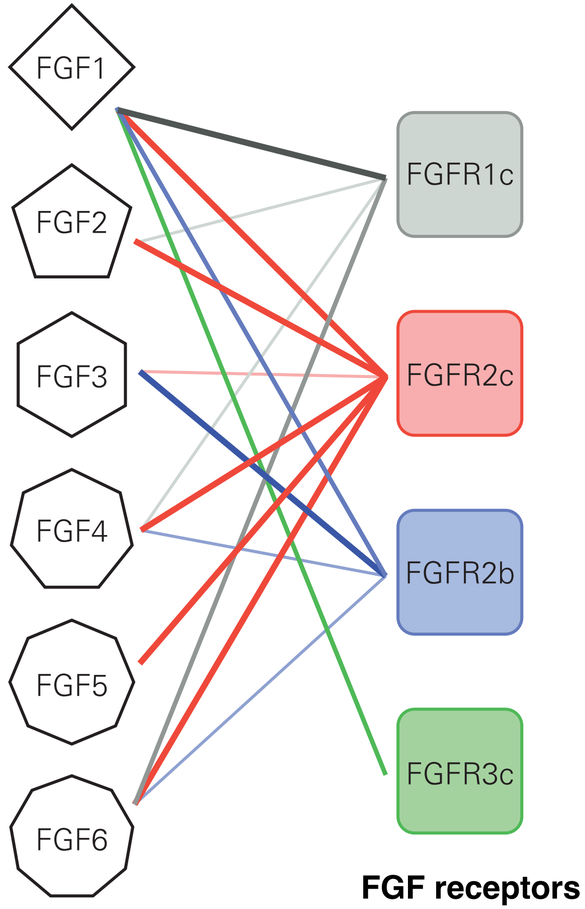
Affinity vs. specificity

Affinity of an interaction defines its strength (ie K_D).
Specificity is the affinity of the ligand for one and only one specific target of interest



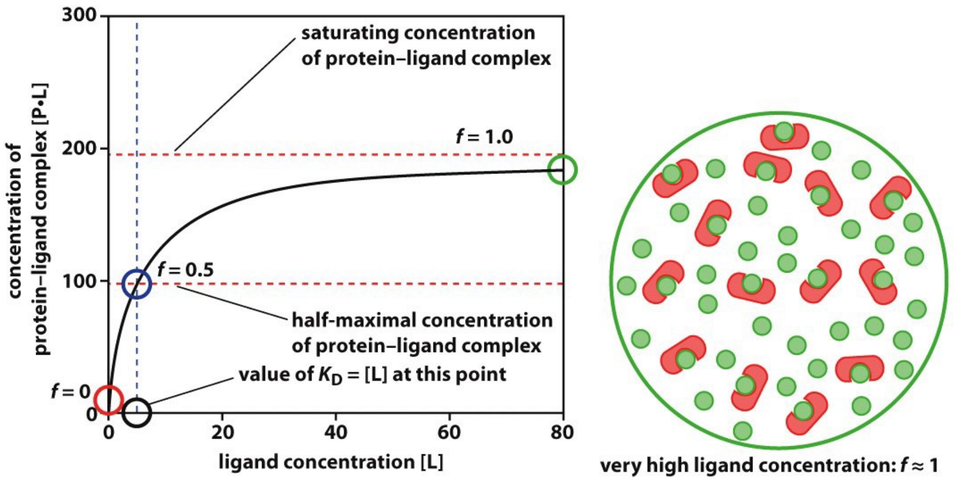
Interaction between **fibroblast growth factor** (**FGF**) and its receptor



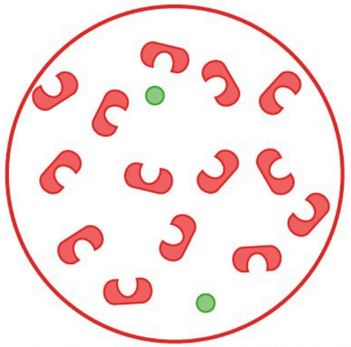


How to determine K_D

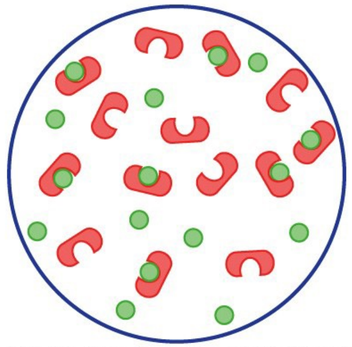
The value of K_D corresponds to the concentration of free ligand at which protein is half saturated



very high ligand concentration: $f \approx 1$



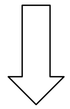
very low ligand concentration: $f \approx 0$



intermediate ligand concentration: $f = 0.5$
K_D = [L] when $f = 0.5$

Figure 12.4 The Molecules of Life (© Garland Science 2013)

Let's see why this is true:
we call ***f*** the **fractional saturation** or fractional occupancy of the ligand binding sites



which is the extent to which the binding sites on a protein are filled with ligand

How to determine K_D

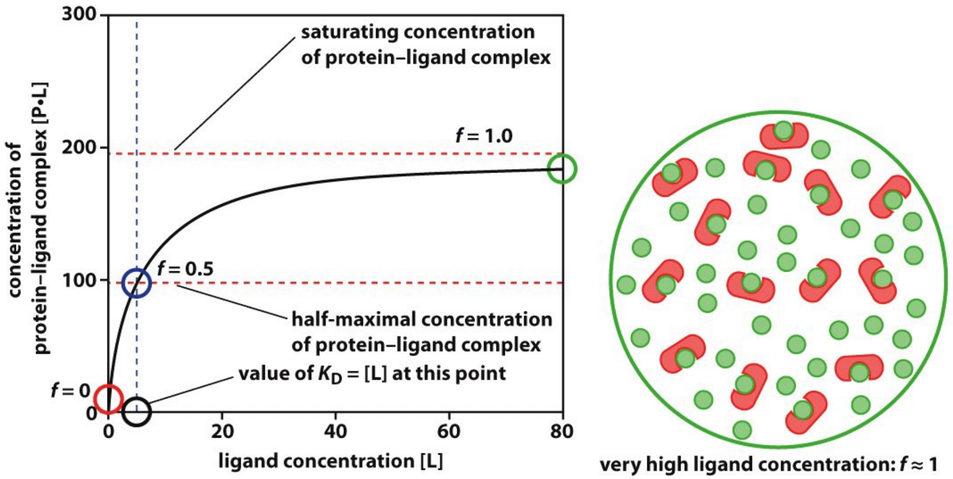


Figure 12.4 The Molecules of Life (© Garland Science 2013)

Let's see how this is true:
we call ***f*** the **fractional saturation** or fractional occupancy of the ligand binding sites

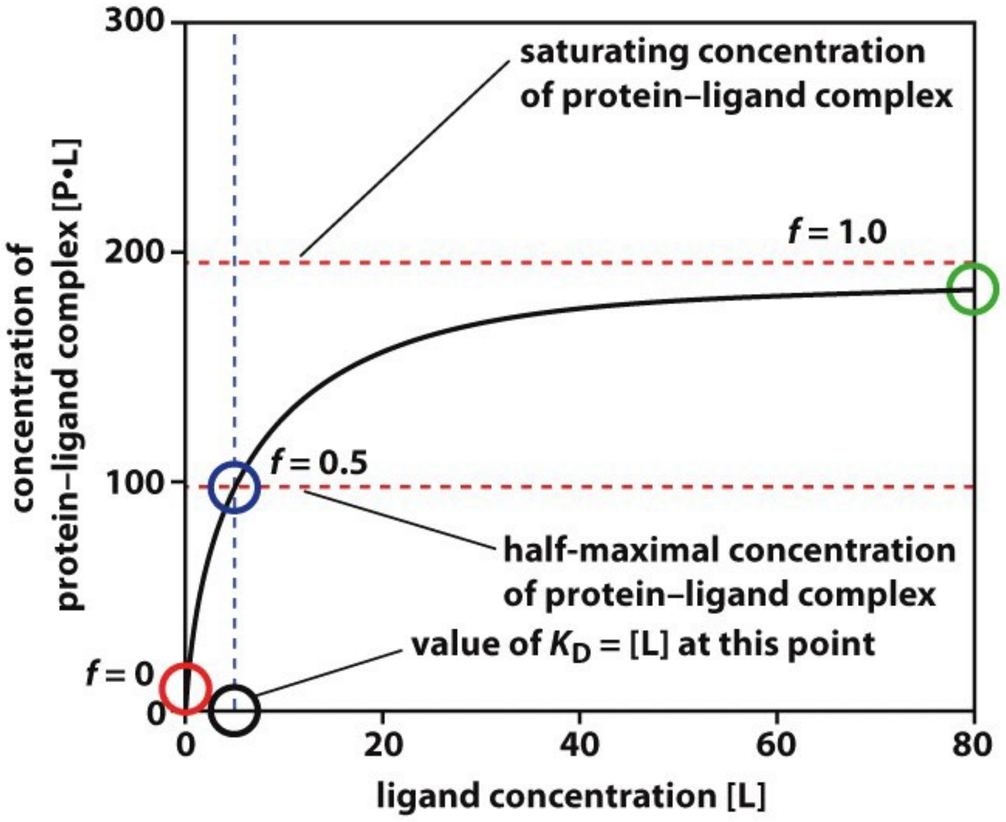
$$f = \frac{\text{concentration of protein with ligand bound}}{\text{total protein concentration}} = \frac{[\mathrm{P{\bullet}L}]}{[\mathrm{P}]+[\mathrm{P{\bullet}L}]}$$

Recalling that : $[\mathrm{P{\bullet}L}] = \dfrac{[\mathrm{P}][\mathrm{L}]}{K_D}$

$$f = \frac{[\mathrm{P}][\mathrm{L}]}{K_D\left([\mathrm{P}] + \dfrac{[\mathrm{P}][\mathrm{L}]}{K_D}\right)}$$

How to determine K_D

$$f = \frac{\text{concentration of protein with ligand bound}}{\text{total protein concentration}} = \frac{[P \bullet L]}{[P]+[P \bullet L]}$$



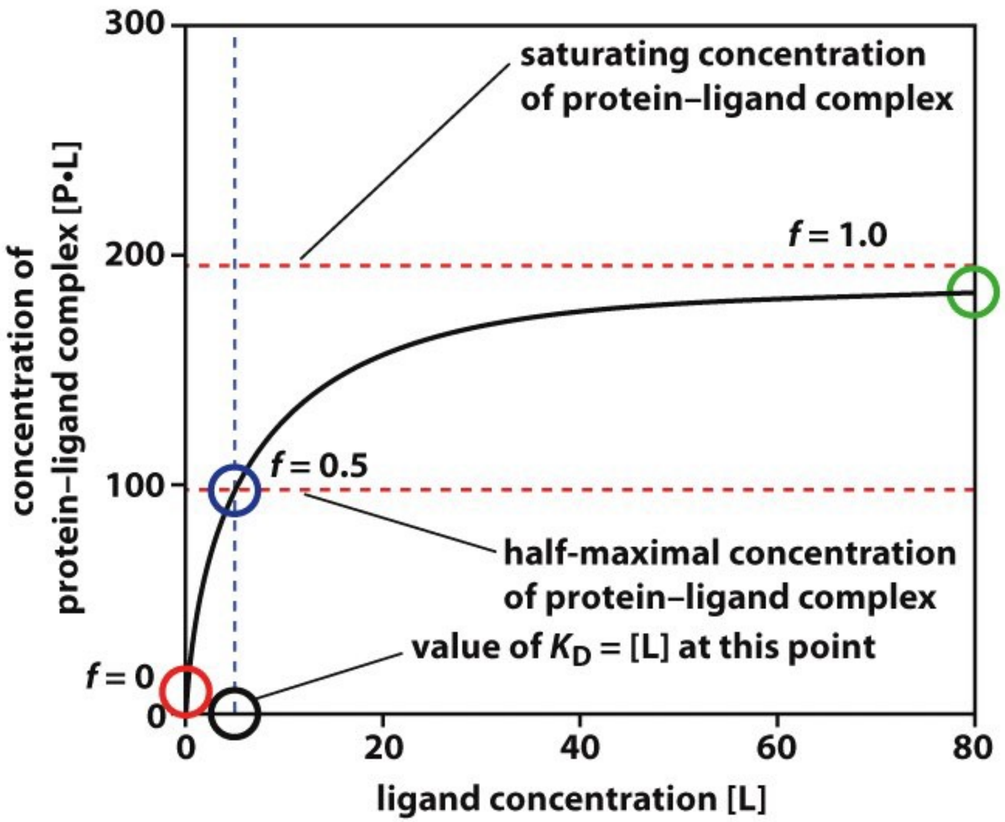
$$f = \frac{[P][L]}{K_D\left([P]+\frac{[P][L]}{K_D}\right)}$$

$$\Rightarrow f = \frac{[L]}{K_D\left(1+\frac{[L]}{K_D}\right)} = \frac{[L]}{K_D+[L]} = \frac{\frac{[L]}{K_D}}{1+\frac{[L]}{K_D}}$$

$$f = \frac{[L]}{[L]+K_D}$$

(rectangular hyperbolic function)

How to determine K_D



$$f = \frac{[\mathrm{L}]}{[\mathrm{L}] + K_\mathrm{D}}$$

When $[\mathrm{L}] = K_\mathrm{D}$

$$f = \frac{K_\mathrm{D}}{K_\mathrm{D} + K_\mathrm{D}} = \frac{1}{2}$$

Therefore, a plot of fractional saturation as function of ligand concentration is known as a **binding isotherm** or **binding curve**. The K_D value depends on the temperature.

Binding assays

In all cases to estimate dissociation constants we need to come up with some **experimental** procedure able to measure the amount of ligands bound to a protein. These are called **binding assays** - there are many ways to develop a binding assay based on many different techniques and properties of the ligand or proteins. Here is an example based on **radioactivity** applied to steroid hormone receptors

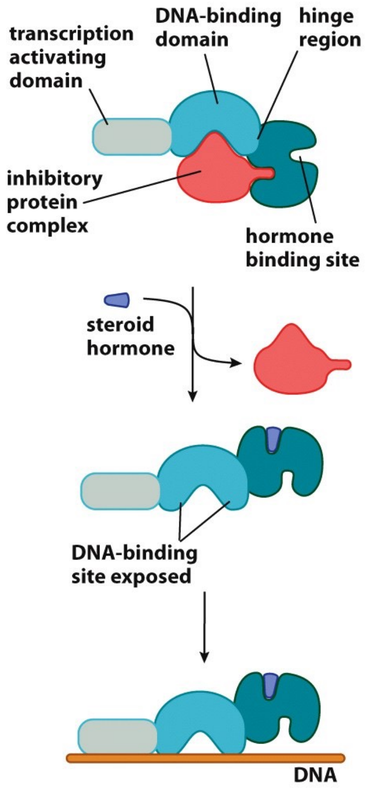
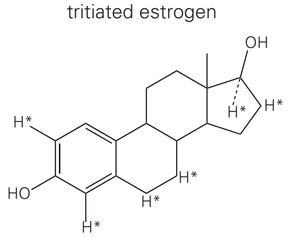
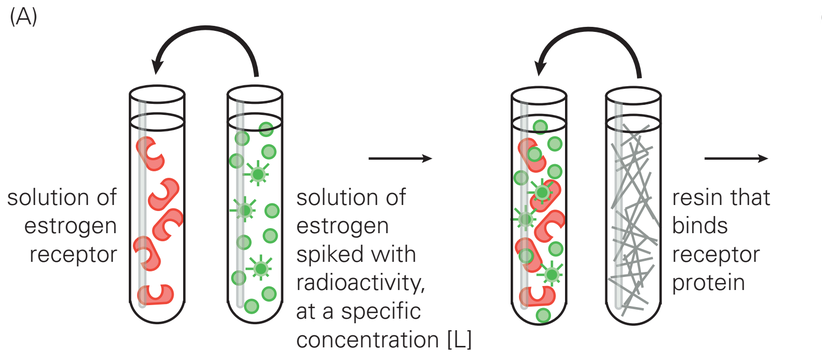
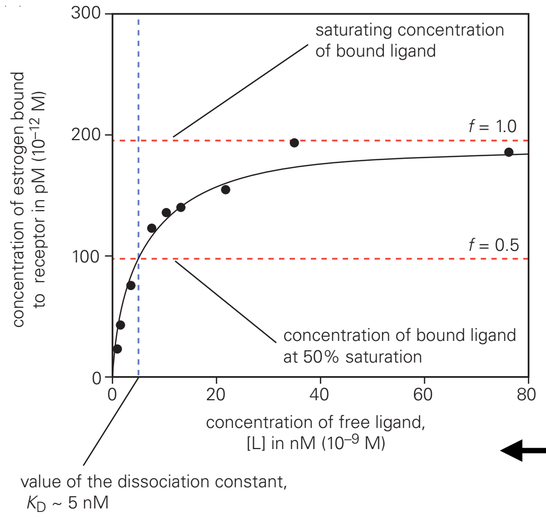


Figure 12.5b The Molecules of Life (© Garland Science 2013)

we use a radioactively labelled ligand

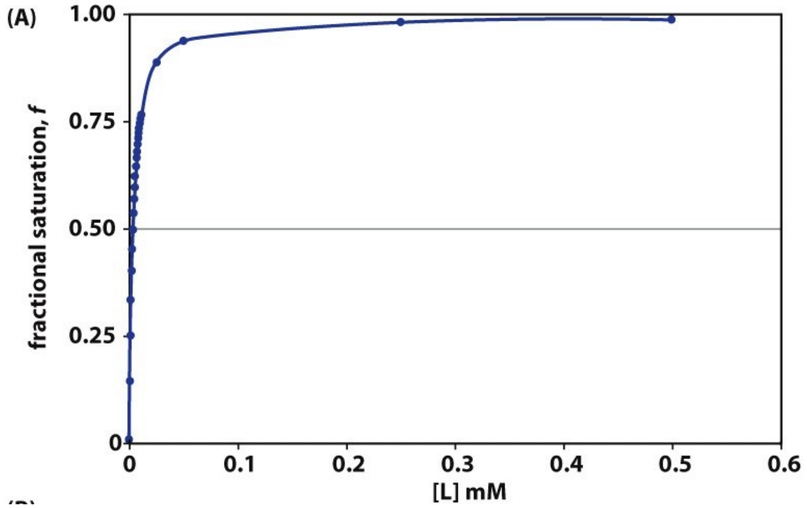


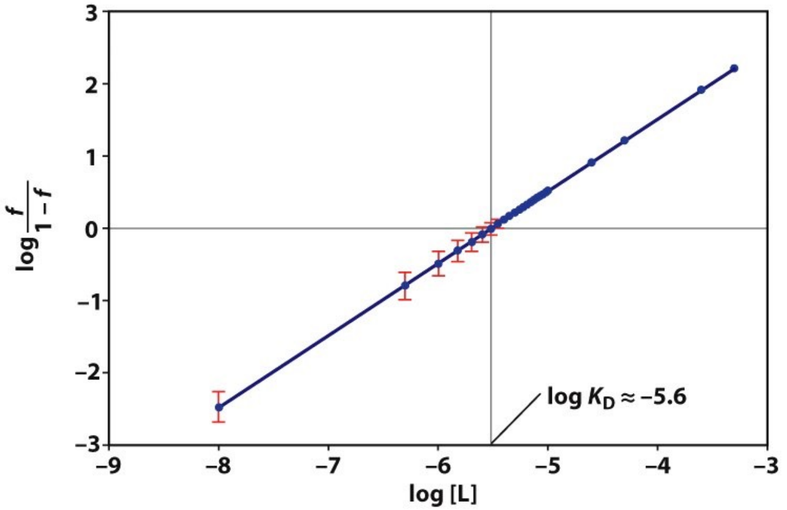




[P●L] is then plotted for various concentration of L and curve fitted to calculate K_D

Binding curves shown in different ways





$$\frac{\text{fraction bound}}{\text{fraction unbound}} = \frac{f}{1-f} = \frac{[\mathrm{L}]}{[\mathrm{L}]+K_\mathrm{D}} \frac{[\mathrm{L}]+K_\mathrm{D}}{K_\mathrm{D}} = \frac{[\mathrm{L}]}{K_\mathrm{D}}$$

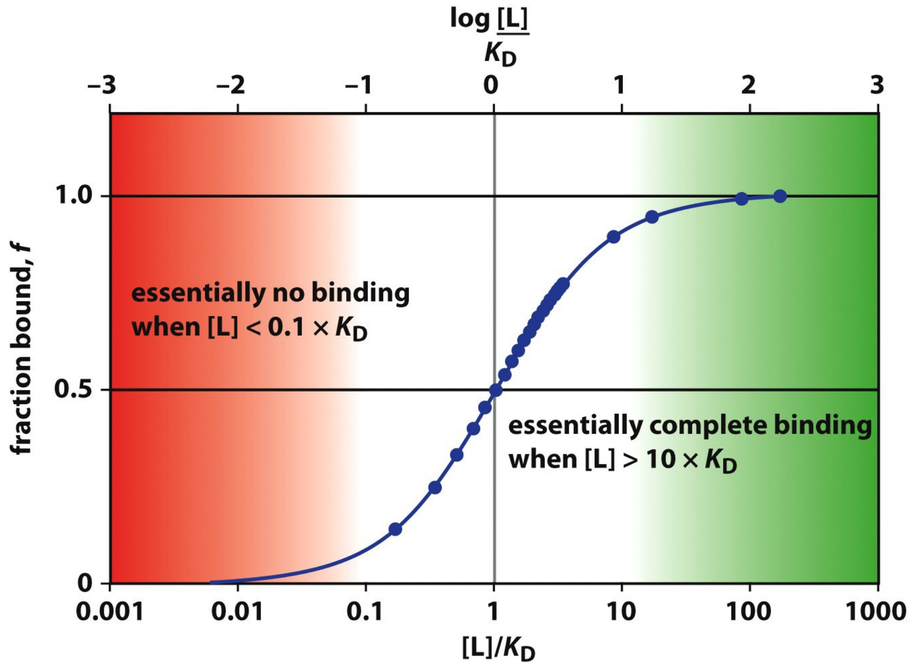
$$\log\left(\frac{f}{1-f}\right) = \log\left(\frac{[\mathrm{L}]}{K_\mathrm{D}}\right) = \log[\mathrm{L}] - \log K_\mathrm{D}$$

When $\left(\frac{f}{1-f}\right) = 1$ then the log is 0 and the intercept of the line on the horizontal axis is equal to the log of K_D

Note: we usually assume that the amount of bound ligand is very small compared to the total amount of ligand available, thus we usually use the free ligand concentration and the total ligand concentration interchangeably

if $[\mathrm{L}]_{\text{bound}} << [\mathrm{L}]_{\text{total}}$ then $[\mathrm{L}]_{\text{total}} = [\mathrm{L}] + [\mathrm{L}]_{\text{bound}} \approx [\mathrm{L}]$

Universal Binding Isotherm



Saturable binding is the hallmark of specific binding interactions, ie a simple binding mode with one protein associated with one ligand, one binding pocket at the time

The K_D defines the ligand concentration range over which the protein switches from unbound to bound - using the universal binding curve is a handy way to characterise binding because the **ligand concentration is expressed in terms of dissociation constant** (ie it is a universal curve).

Universal Binding Isotherm

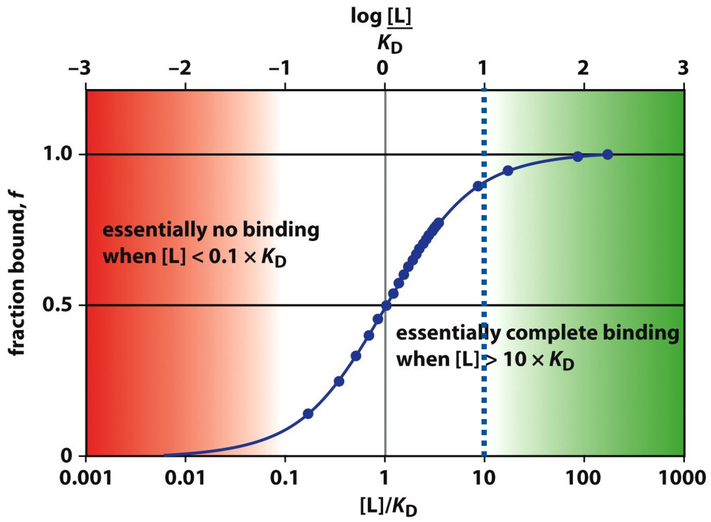


Figure 12.11 The Molecules of Life (© Garland Science 2013)

A common question in biochemistry and pharmacology is how much of a protein is bound to a ligand.

For a concentration where the free ligand is 10 times the value of K_D , the target is almost completely occupied (at 91%), in fact:

$$f = \frac{[\mathrm{L}]}{[\mathrm{L}] + K_\mathrm{D}} = \frac{10K_\mathrm{D}}{10K_\mathrm{D} + K_\mathrm{D}} = \frac{10}{11} = 0.91$$

Universal Binding Isotherm

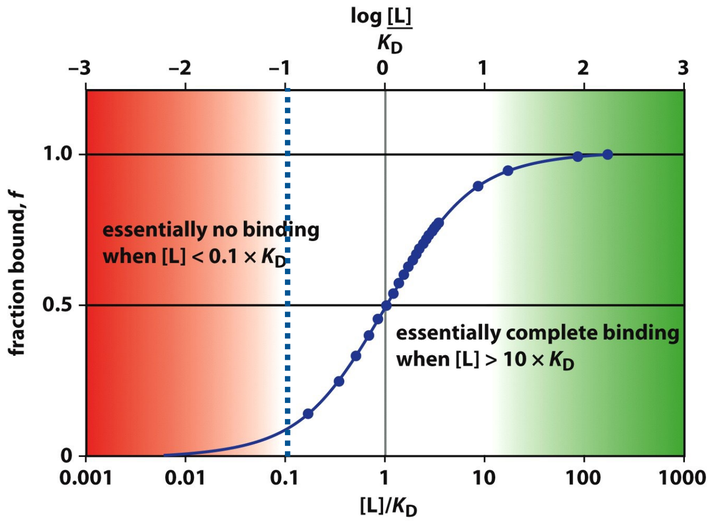


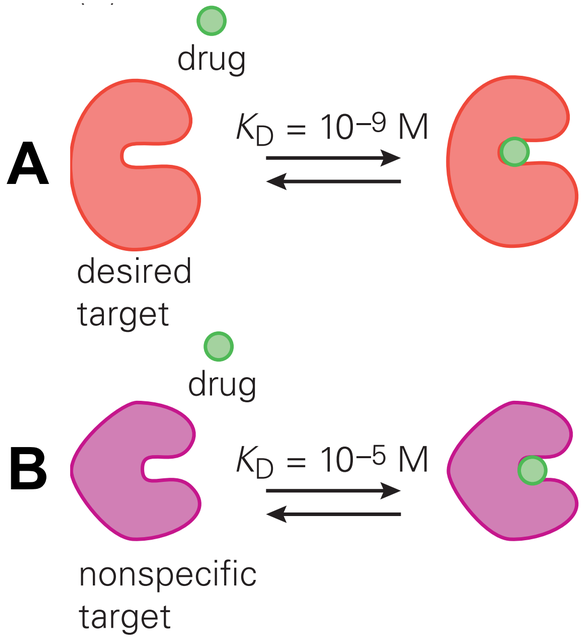
Figure 12.11 The Molecules of Life (© Garland Science 2013)

A common question in biochemistry and pharmacology is how much of a protein is bound to a ligand

For a concentration where the free ligand is 0.1 times the value of K_D, the target is occupied only at 9%

$$f = \frac{0.1K_D}{K_D + 0.1K_D} = \frac{0.1}{1.1} = 9\%$$

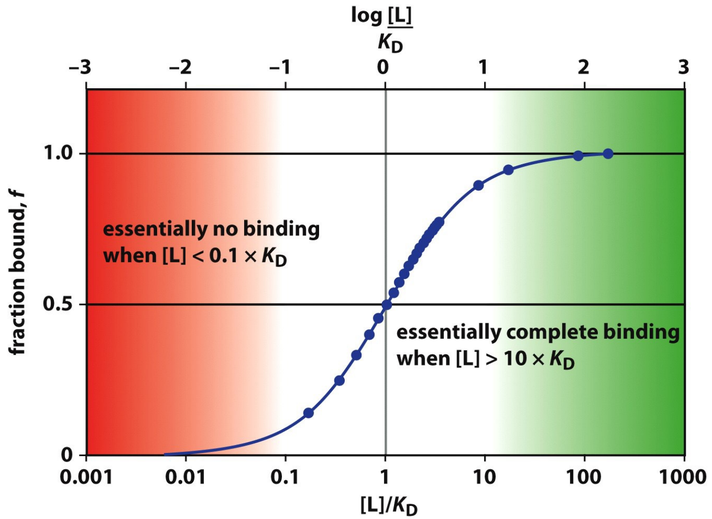
Thinking again about specificity !



Let's make use of the universal binding curve

If you have this situation in which you have to develop a drug for a given target A, which should also be selective, ie it does not bind a second target B.

What would the optimal K_Ds be in relative terms to achieve this goal?



If the drug is delivered at a 100 nM concentration, can you achieve a proper selectivity?

Yes - If you administer this drug at 100 nM (10^{-7} M), you will obtain that 99% of the target A is bound with the drug, while only 1% of target B will be occupied

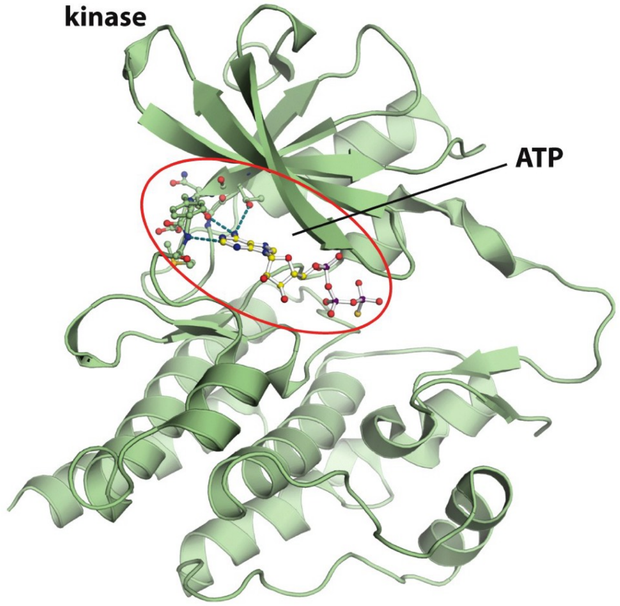
Drug Binding by Proteins

One of the main criteria that drives **drug development (DD)** processes is the binding affinity of the drug to a protein

Generally it starts that have low affinity (ie **lead** compounds), which will then improved by the **lead optimisation** process

Let’s look at the example of kinases – key enzymes in signalling and involved in a number of diseases such as cancer

Kinases natively bind ATP and transfer one phosphate group to Ser, Thr or Tyr of protein substrates



Drug Binding by Proteins

Structural information does help DD

If we look at the active site interactions they can help us to rationally design drug compounds

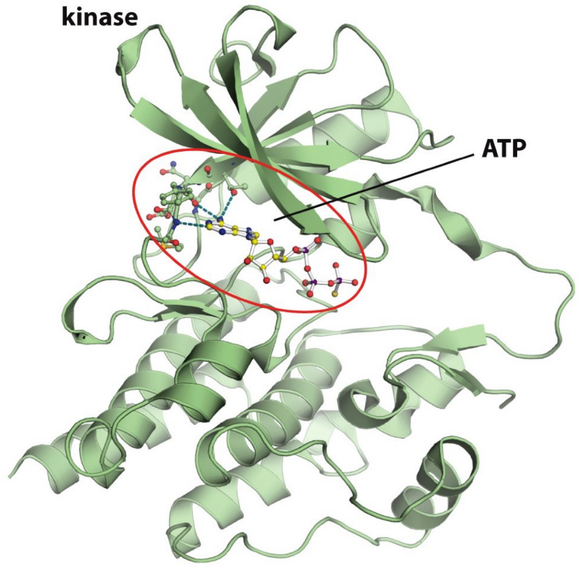


Figure 12.13a The Molecules of Life (© Garland Science 2013)

For example the malfunctioning of the tyrosine-protein kinase Abl causes chronic myelogenous leukemia. Inhibitor of Abl known as imatinib (marketed as Gleevec) blocks the action of Abl and is an effective treatment for the leukemia.

How and where to start with the DD process?

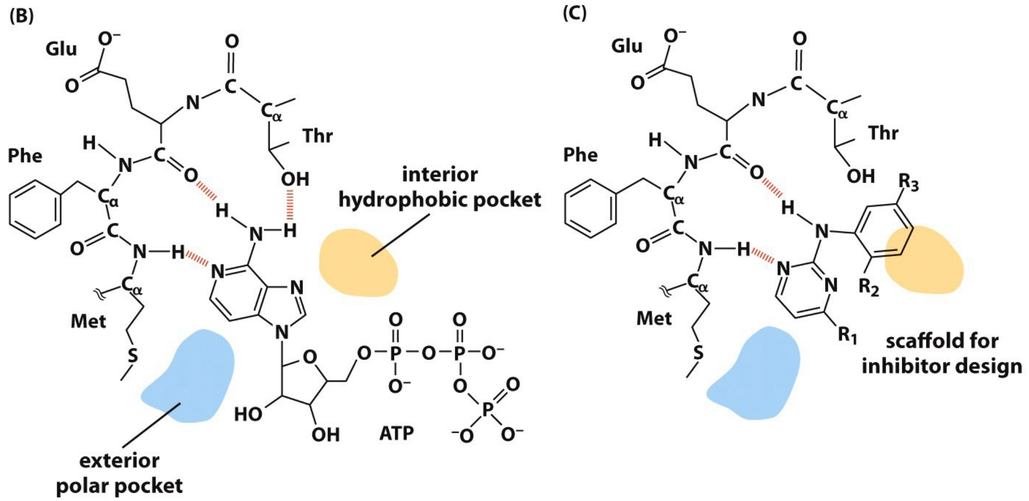


Figure 12.13bc The Molecules of Life (© Garland Science 2013)

The Drug Development Process

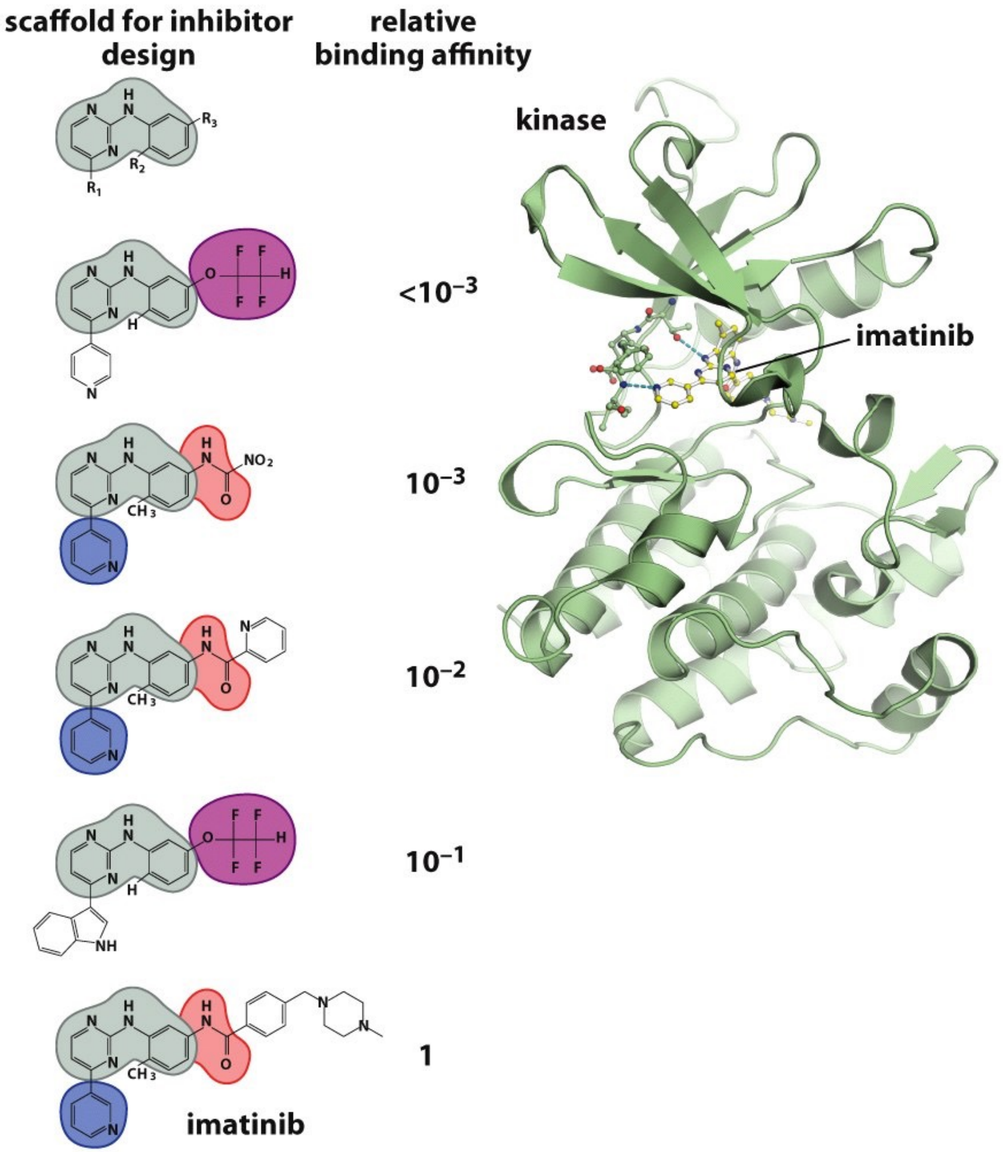
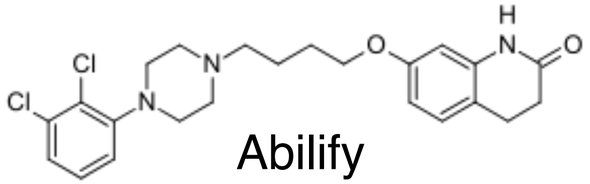


Figure 12.14 The Molecules of Life (© Garland Science 2013)

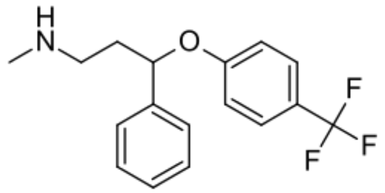
Lead optimization

We can see in this example how different changes in the structure of the small molecule change binding affinities, to reach up to ~10 nM

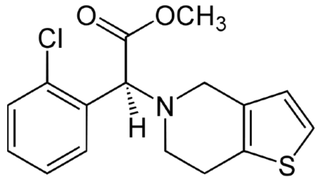
Small molecules drugs



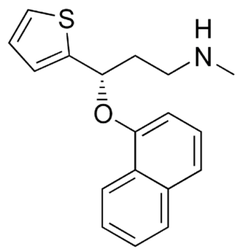
Abilify
(antipsychotic)



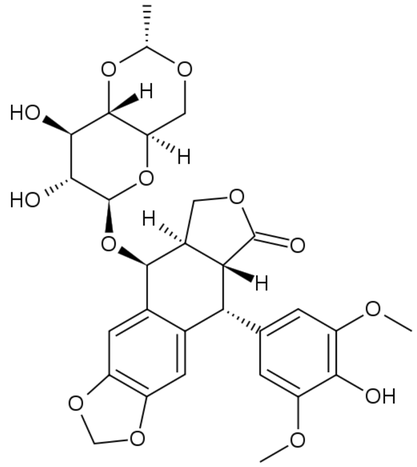
Prozac
(antidepressant)



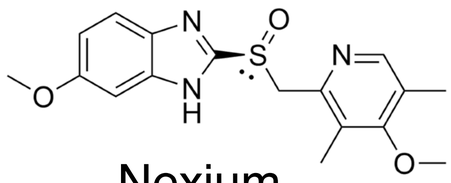
Plavix
(antiplatelet agent)



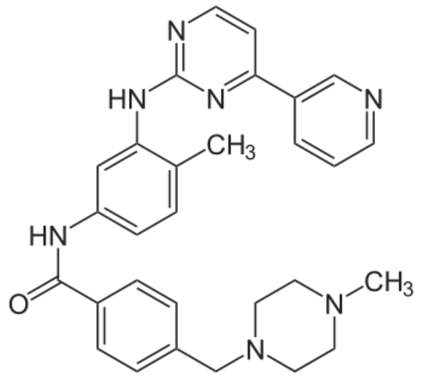
Cymbalta
(pain and anxiety)



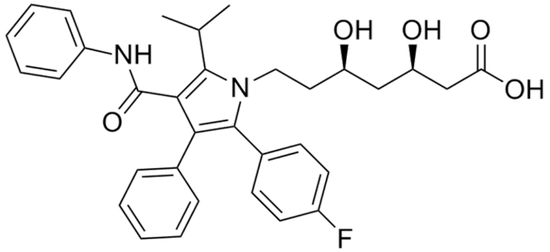
Etoposide
(cancer)



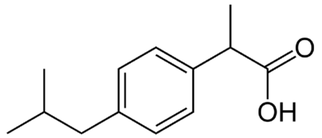
Nexium
(gastric acid)



Gleevec
(cancer)



Lipitor
(Cholesterol)



Ibuprofen
(inflammation)

Drug Discovery & Development

FDA approvals in 2020

53 New Drugs Approved

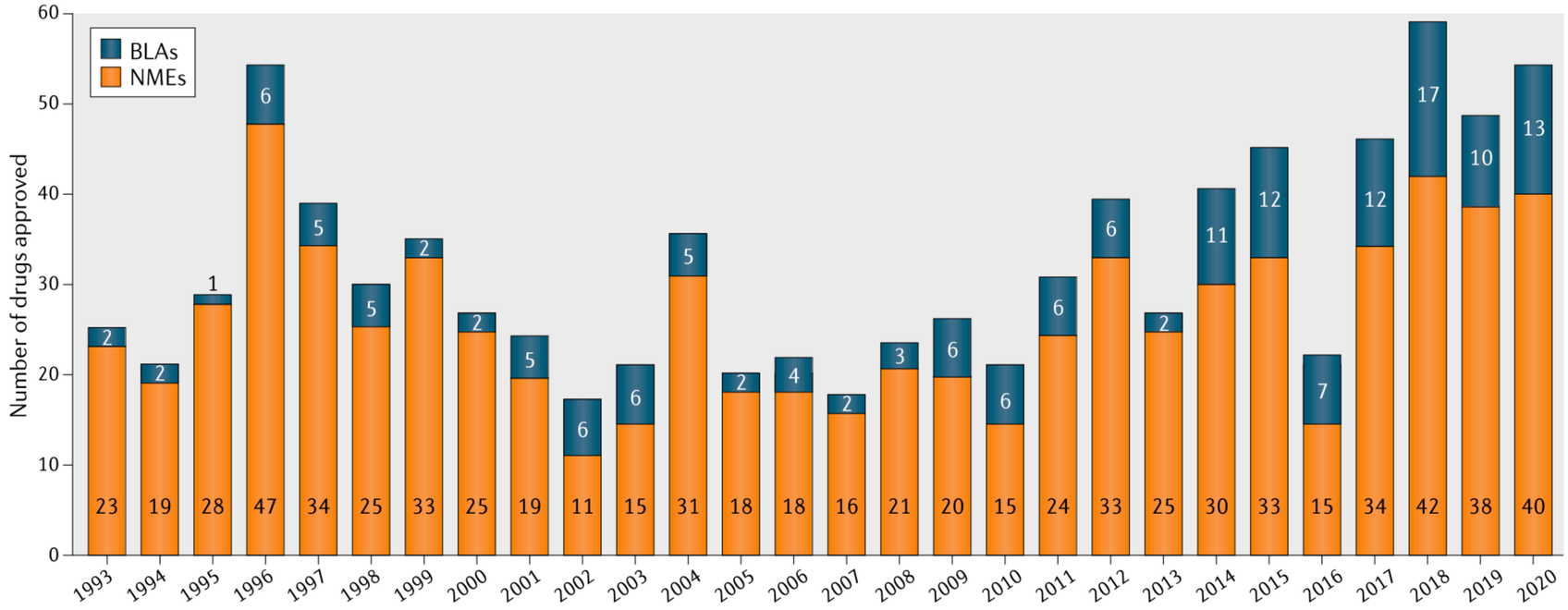


Fig. 1 | **Novel FDA approvals since 1993.** Annual numbers of new molecular entities (NMEs) and biologics license applications (BLAs) approved by the FDA's Center for Drug Evaluation and Research (CDER). See TABLE 1 for new approvals in 2020. Approvals by the Center for Biologics Evaluation and Research (CBER), for products such as vaccines and gene therapies, are not included in this drug count (see TABLE 2). Source: FDA.

Source:
Asher Mullard,
Nature Reviews Drug Discovery, February (**2021**)

Energetics of Binding

it is always true that entropy is lost upon binding of the small molecule, but

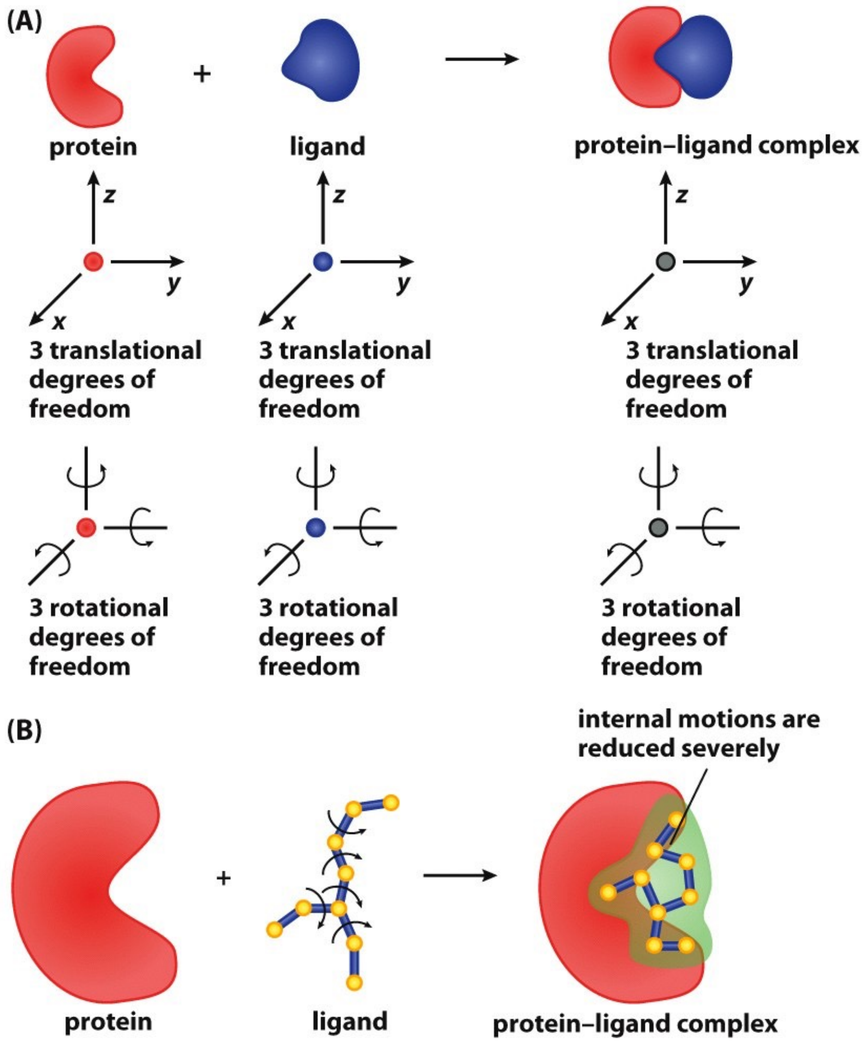


Figure 12.33 The Molecules of Life (© Garland Science 2013)

$$\Delta G = \Delta H - T\Delta S$$

- There are entropic losses on binding which are not energetically favorable

- Enthalpic gains on the interactions of the protein with the ligand

- Entropic gains on the solvent (ie hydrophobic effect)

thus overall $\Delta G < 0$

Energetics of Binding

Very similar principles to the folding reaction

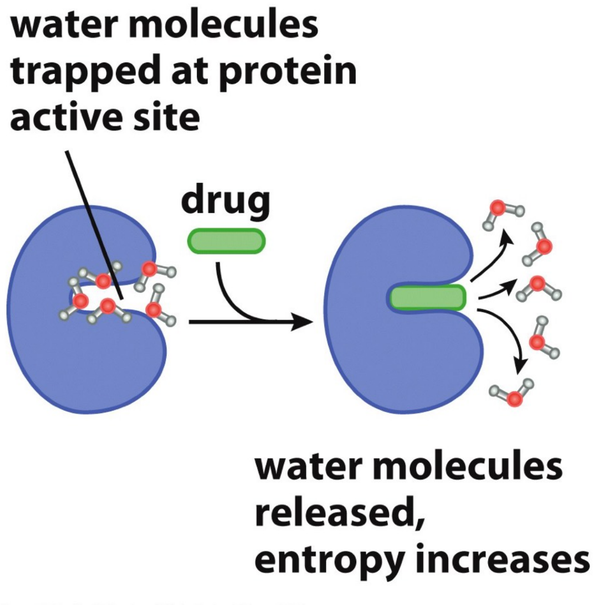


Figure 12.35a The Molecules of Life (© Garland Science 2013)

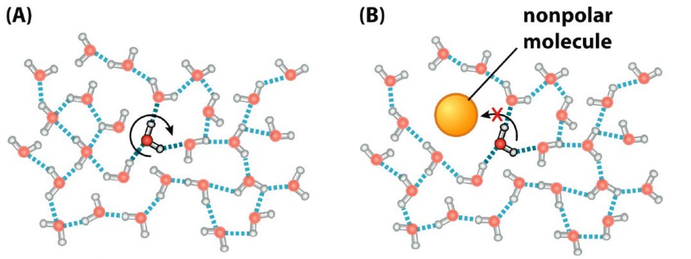


Figure 12.34 The Molecules of Life (© Garland Science 2013)

- Enthalpic gains on the interactions of the protein with the ligand come **mostly from H-bonds** - this contribution is not very large
- Entropic gain in free energy due to the solvent (ie **hydrophobic effect**) is instead quite significant
- As a consequence an optimal drug should **maximise hydrophobic features**
- In practice, drugs cannot be only hydrophobic as they will be **insoluble**, complicating adsorption and administration
- Also hydrophobic interactions tend to be not specific, while polar are more specific
- Thus in the development of a drug one needs to find the **best compromise** between polar and hydrophobic interactions

What to know ...

- The affinity of a protein for a ligand is characterized by the dissociation constant K_D

- The value of the K_D for a binding interaction is the ligand concentration at which half the receptors are bound to ligand

- The value of the K_D determines the concentration range of the ligand over which the receptor switches from unbound to bound

- The K_D for a physiological ligand is usually close to the natural concentration of the ligand, this is the result of evolution

- Key formulas :

$$K_D = \frac{[P][L]}{[P \bullet L]} = \frac{1}{K_A} \qquad \Delta G^o_{bind} = +RT \ln K_D$$

$$f = \frac{[L]}{[L] + K_D} \qquad \log\left(\frac{f}{1-f}\right) = \log\left(\frac{[L]}{K_D}\right) = \log[L] - \log K_D$$