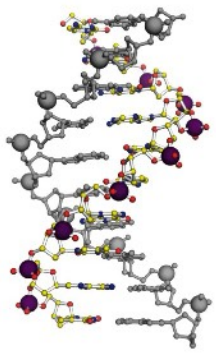
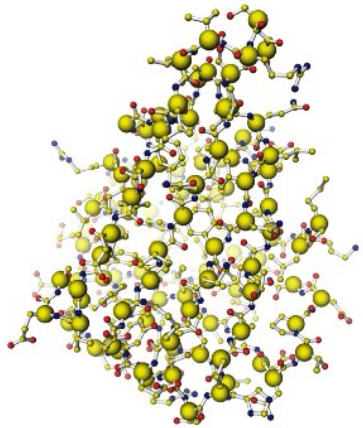
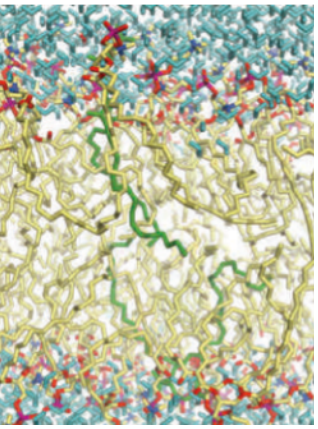


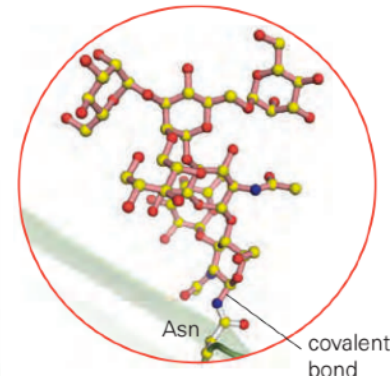
Chimie Biologique I
Biological Chemistry I
BIO-212



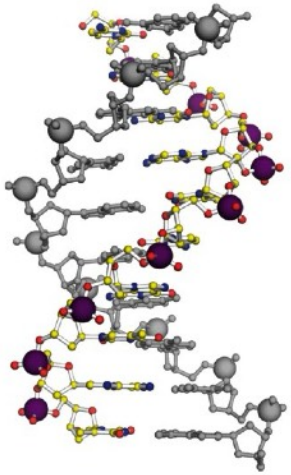
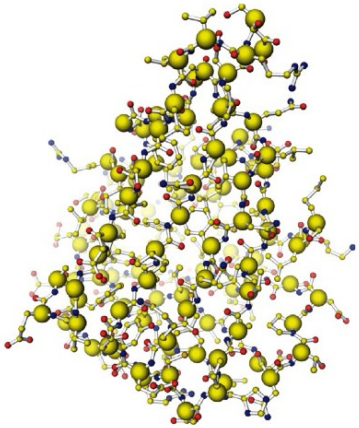
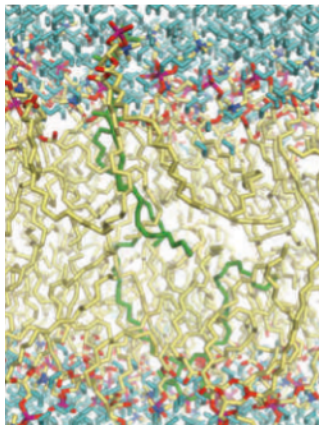
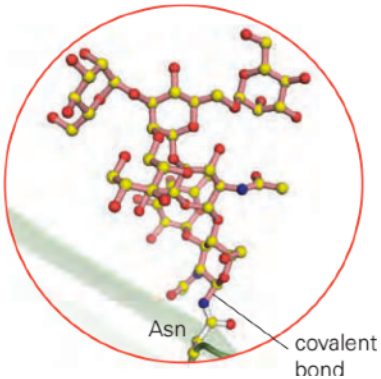
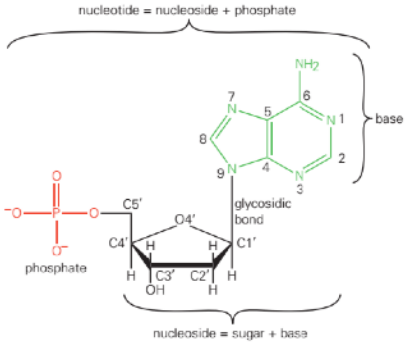
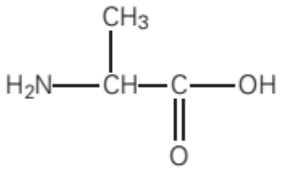
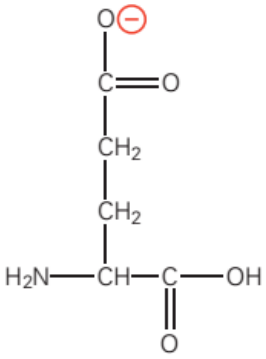
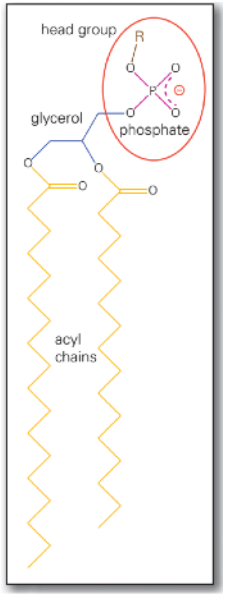
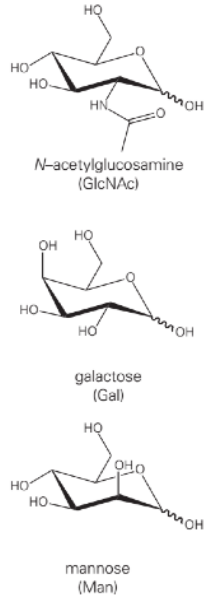
Lecture 4

Matteo Dal Peraro





The Molecules of Life

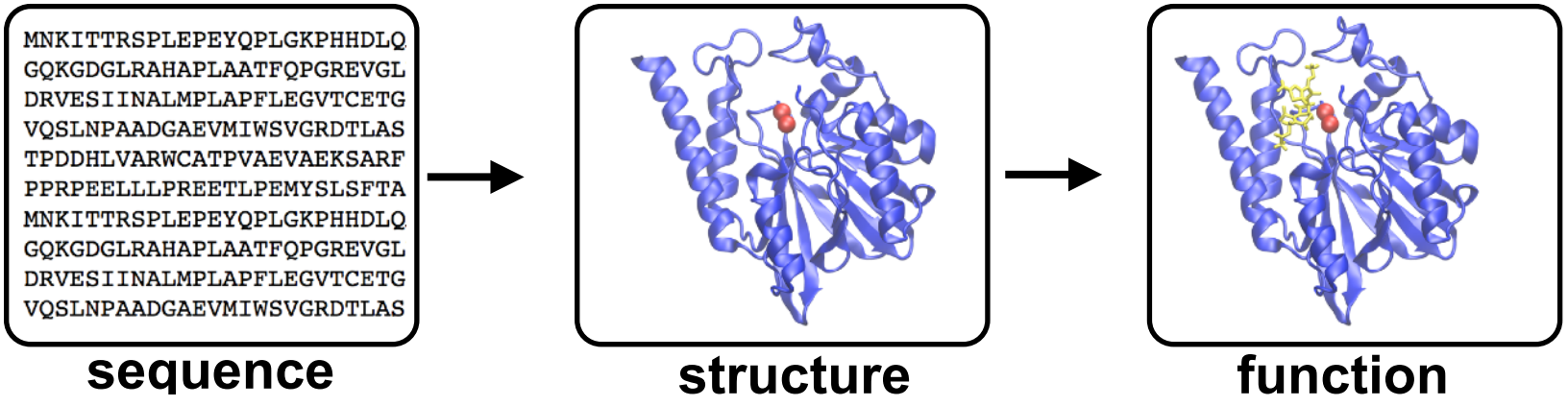
	Nucleic Acids	Proteins	Lipids	Glycans
Macromolecular Structure				
Building Block		 		

Rhetorical Question:

Why do we care about structure ?

Structure largely determines function and the activities that different macromolecules have in cells.

Paradigm in biochemistry



Proteins: Sequence Determines Structure

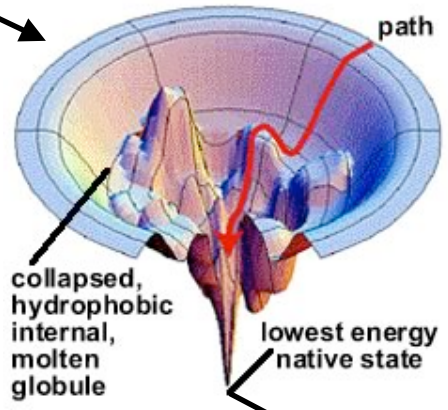
```
MNKITTRSPLEPEYQPLGKPHHDLQ
GQKGDGLRAHAPLAATFQPGREVGL
DRVESIINALMPLAPFLEGVTCETG
VQSLNPAADGAEVMIWSVGRDTLAS
TPDDHLVARWCATPVAEVAEKSARF
PPRPEELLLPREETLPEMYSLSFTA
MNKITTRSPLEPEYQPLGKPHHDLQ
GQKGDGLRAHAPLAATFQPGREVGL
DRVESIINALMPLAPFLEGVTCETG
VQSLNPAADGAEVMIWSVGRDTLAS
```

primary sequence

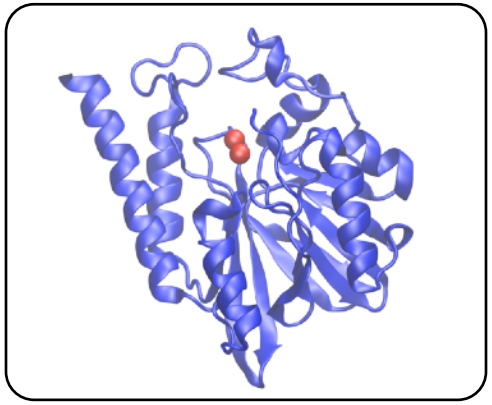
Anfinsen dogma (1954):

protein structure is determined by its sequence

folding pathway



Dill@UCSF



folded native structure

Levinthal paradox (1969)

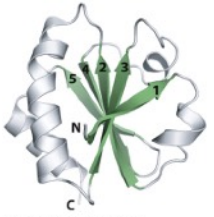
100 residue-long peptide

$3^{198} \sim 10^{94}$ torsional degrees of freedom

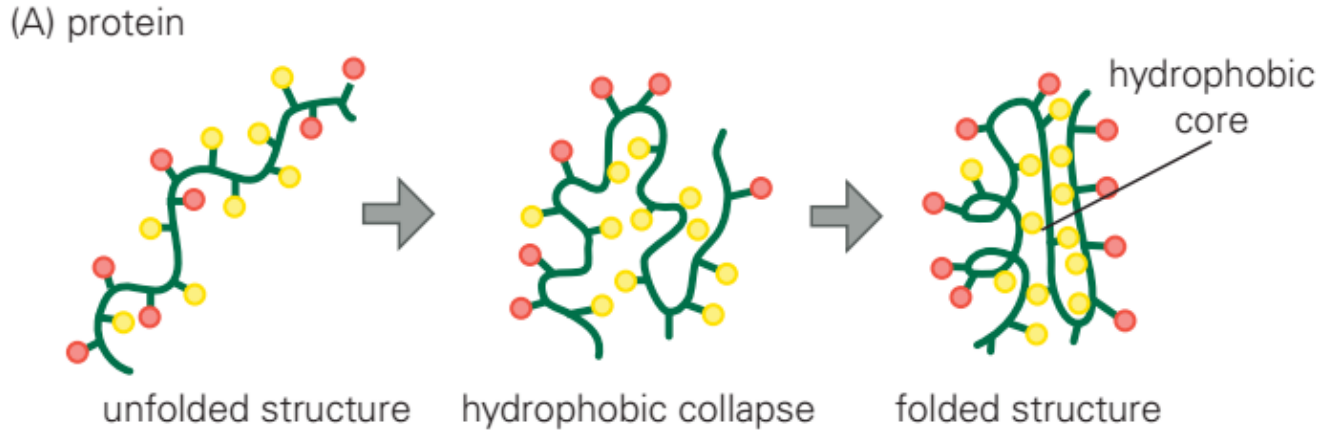
? why is this a paradox ?

because all these conformations are not samples, and the native states is quickly found within milliseconds thanks to thermodynamics

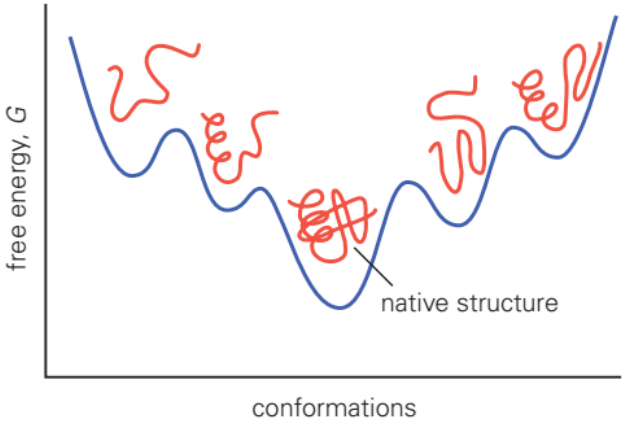
Sequence Determines Structure



- Proteins fold into defined 3D structures

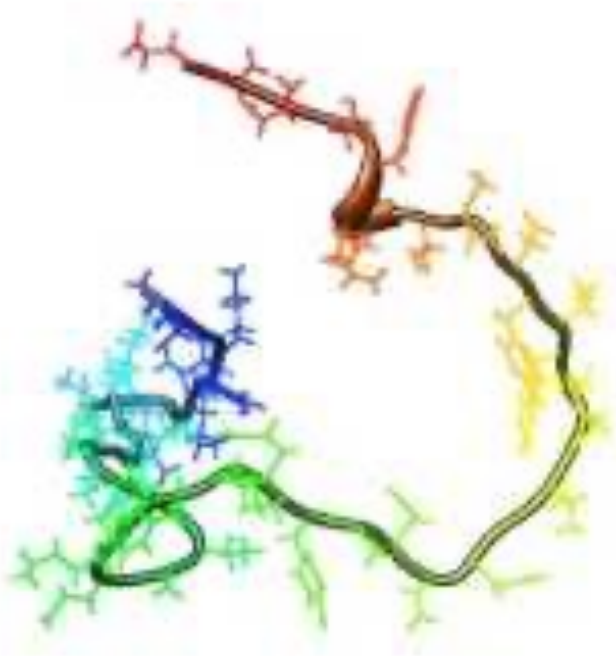


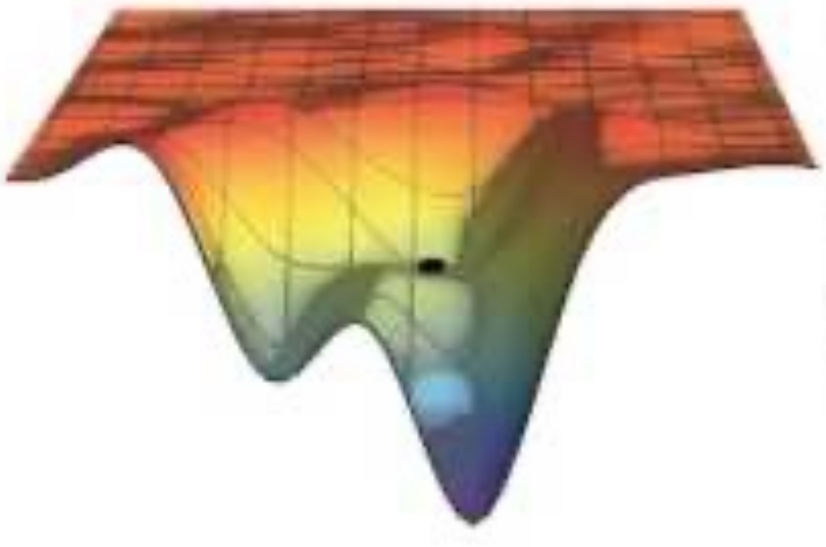
- the folding process is energetically driven and proteins tend to fold to what we call an energy global minima (spontaneously or aided by molecular chaperones)



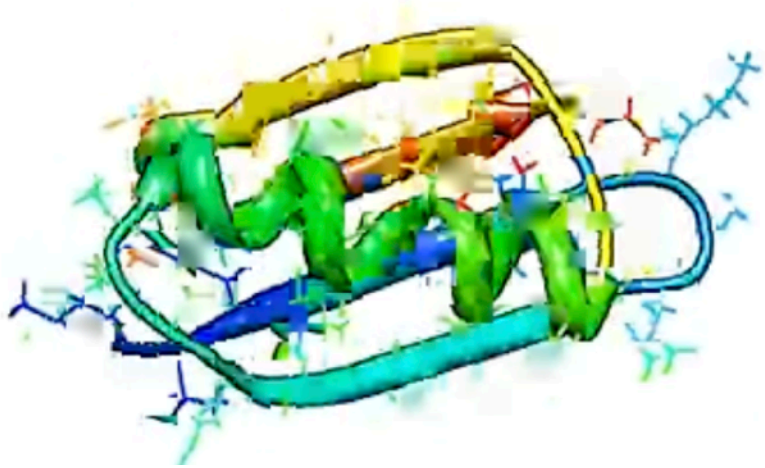
thermodynamic hypothesis

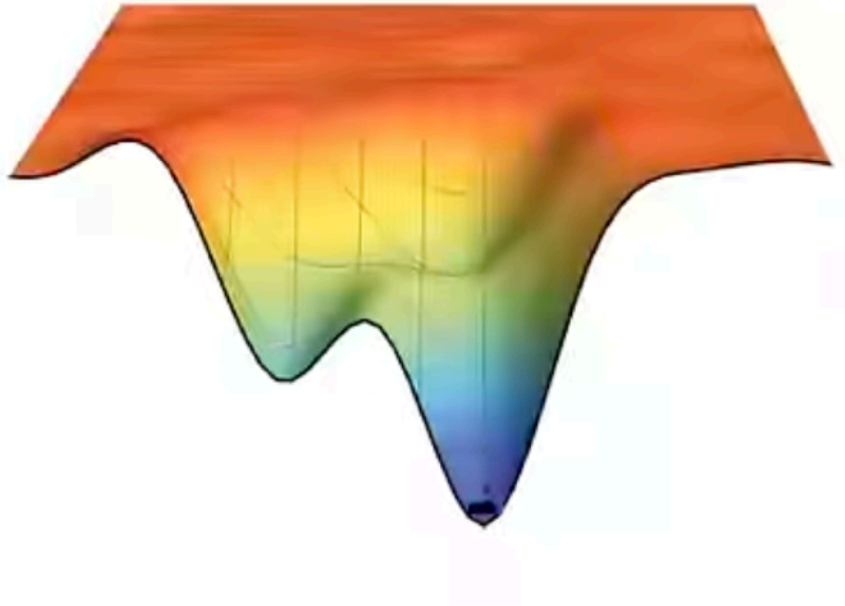
Protein folding





Protein folding





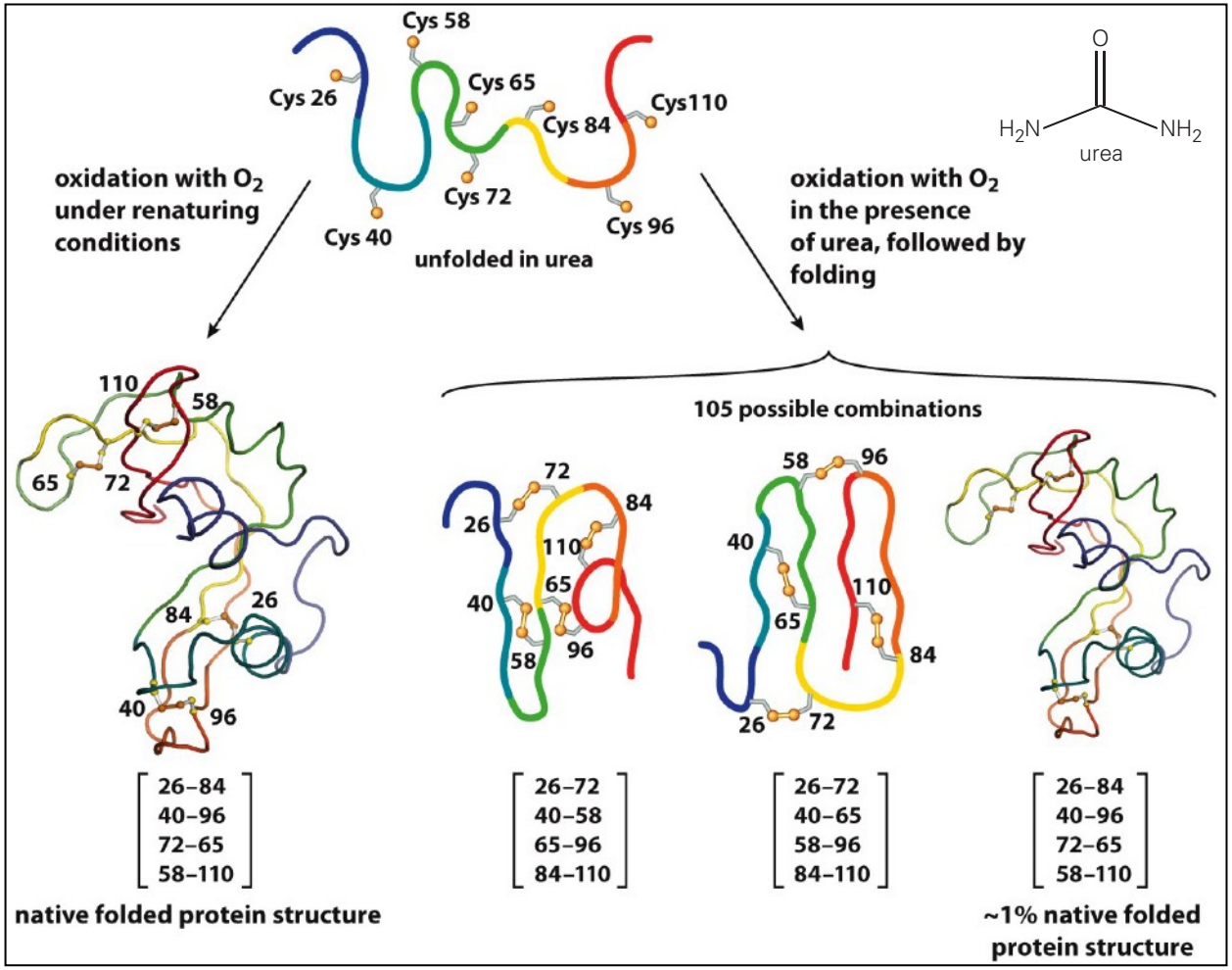
Sequence Determines Structure

The Anfinsen experiment (1954)

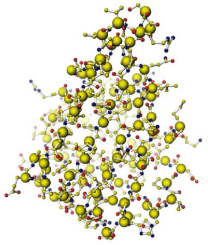
- A Nobel prize (1972) experiment that by measuring enzymatic activity figured out the principles of protein folding

-Ribonuclease refolded and oxidized recovers 100% activity

-Ribonuclease oxidized and refolded in presence of urea recovers 1% activity

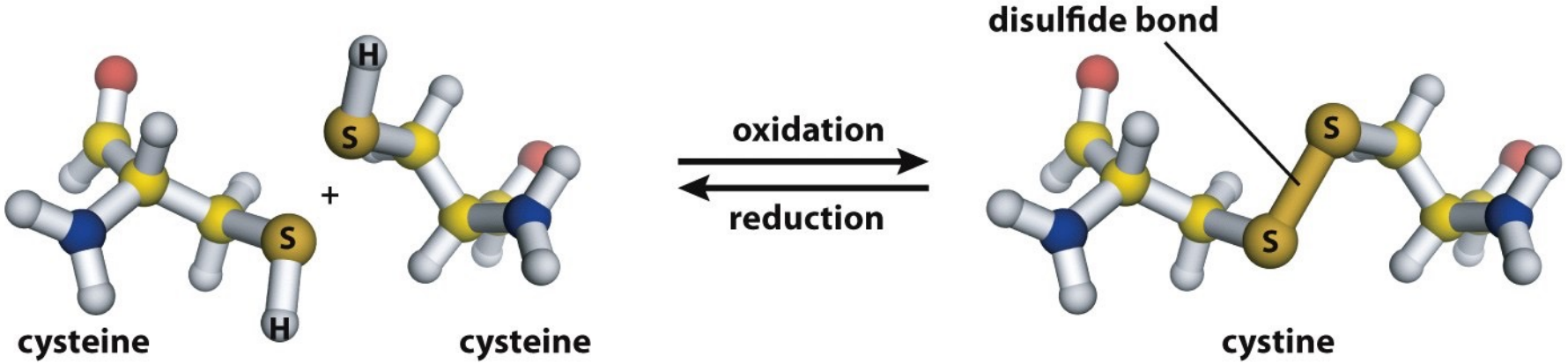


Remember that:



- Cysteine – a special amino acid

- Cysteines can from disulfide bonds



-Disulfide bonds are covalent (reversible) bonds which can play major roles in protein stabilization
- in vivo: mostly present in secreted proteins - rarely found inside the cell.
- in vitro: reducing agent to break them tris (2-carboxyethyl) phosphine hydrochloride (TCEP), beta-mercaptoethanol (BME), and dithiothreitol (DTT).

Protein folding is driven by the formation of a hydrophobic core

-Packing of secondary structural elements bring together the hydrophobic side chains that form the hydrophobic core.

lots of ordered waters (<S)

few residues interactions (>H)

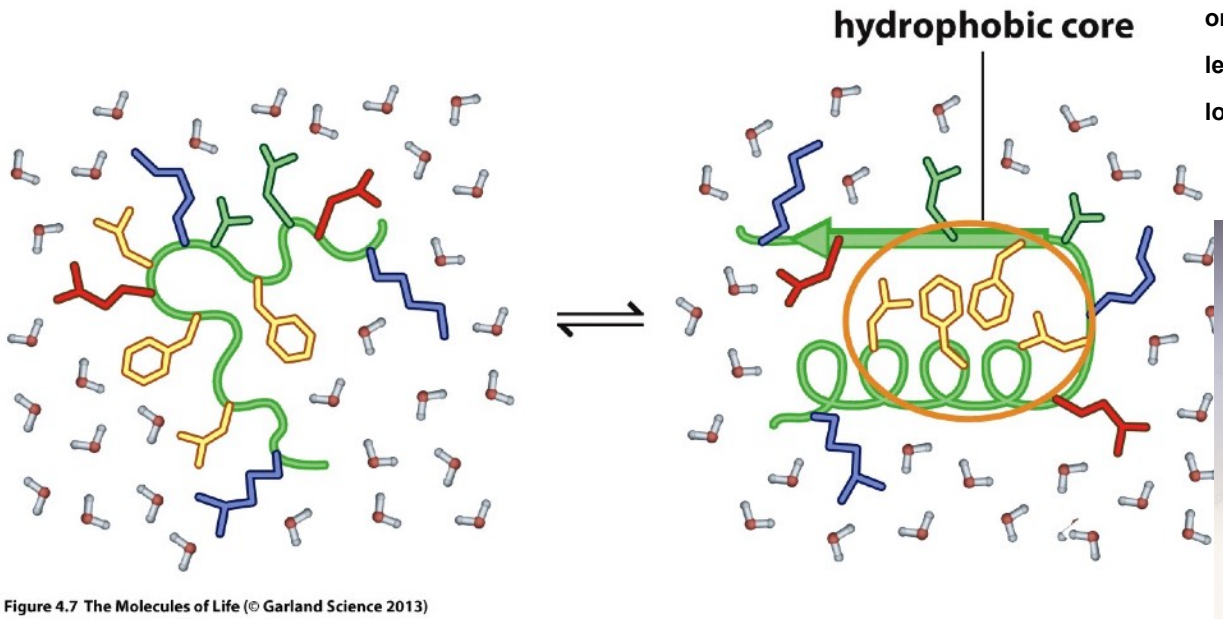


Figure 4.7 The Molecules of Life (© Garland Science 2013)

ordered core (<S)

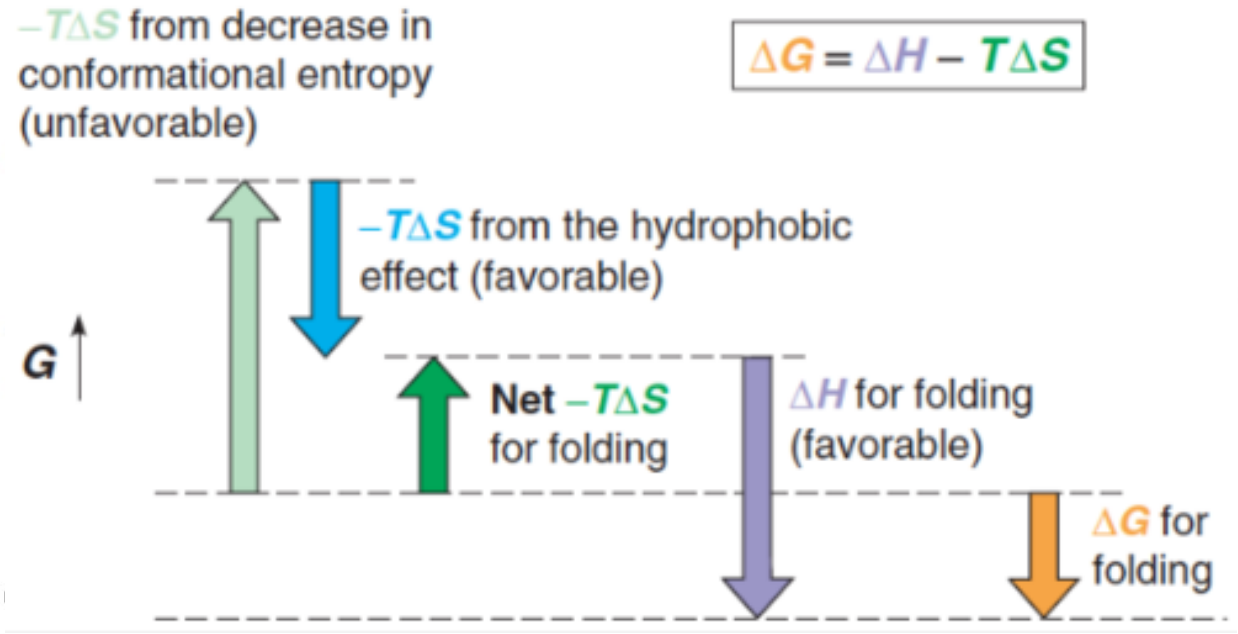
less ordered waters (>S)

lots residues interactions (<H)

-The stability of the folded structure results primarily from the hydrophobic side chains clustering together away from the water (**the hydrophobic effect**)
- folding is a subtle free energy optimization exercise, ***G = H -TS***

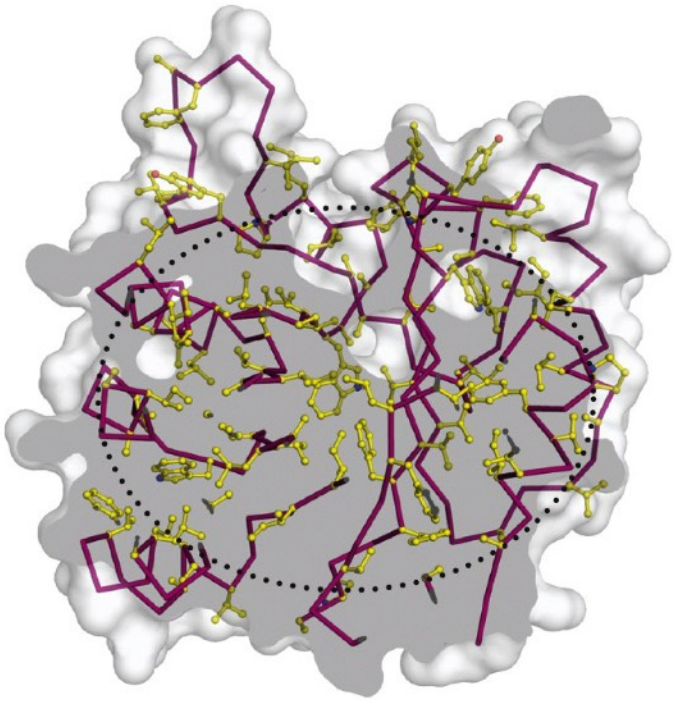
Protein folding is driven by the formation of a hydrophobic core

-Packing of secondary structural elements bring together the hydrophobic side chains that form the hydrophobic core.

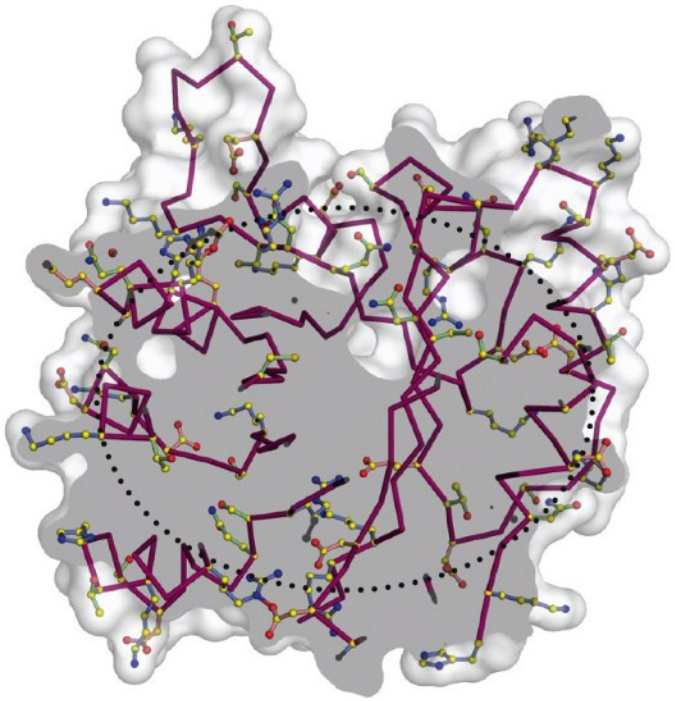


-The stability of the folded structure results primarily from the hydrophobic side chains clustering together away from the water (**the hydrophobic effect**)
- folding is a subtle free energy optimization exercise, ***G = H -TS***

Proteins have hydrophobic cores and hydrophilic surfaces



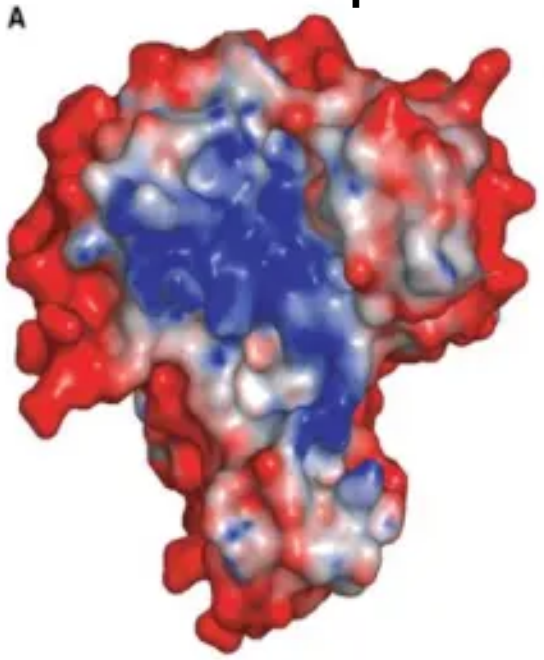
Hydrophobic core



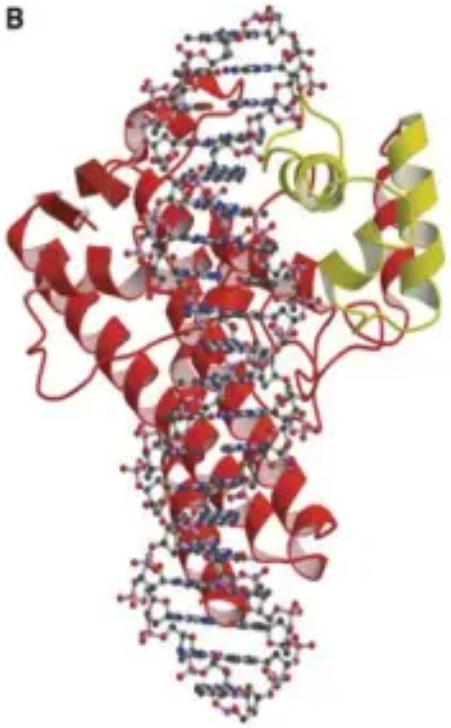
Hydrophilic surface

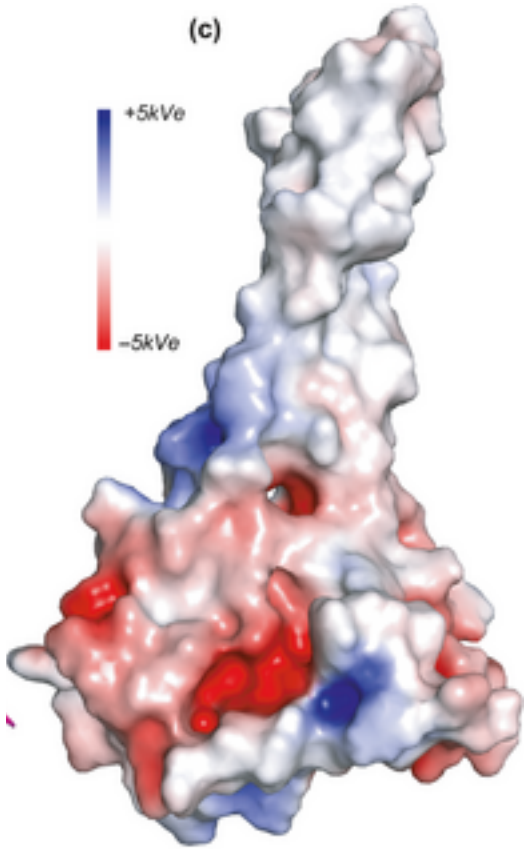
Proteins have hydrophobic cores and hydrophilic surfaces

electrostatic potential



Topoisomerase





blue is positive
red is negative
white is hydrophobic

membrane protein domain

A brief note on membrane proteins

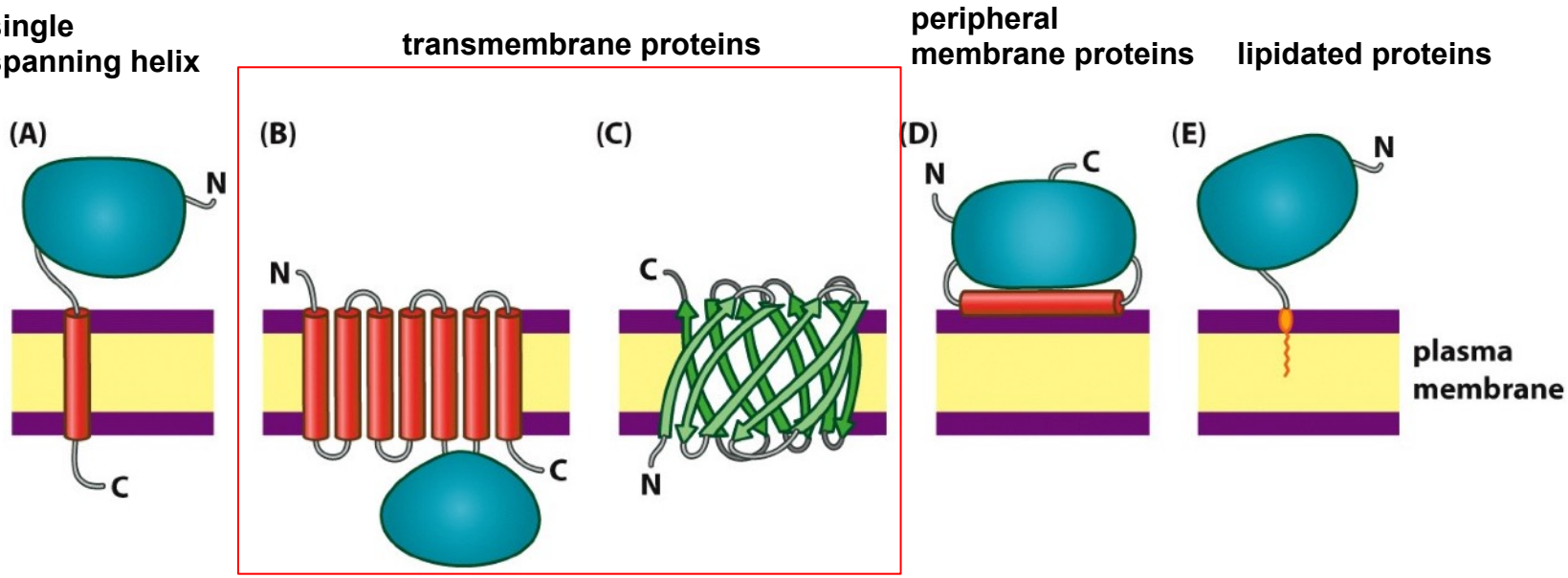
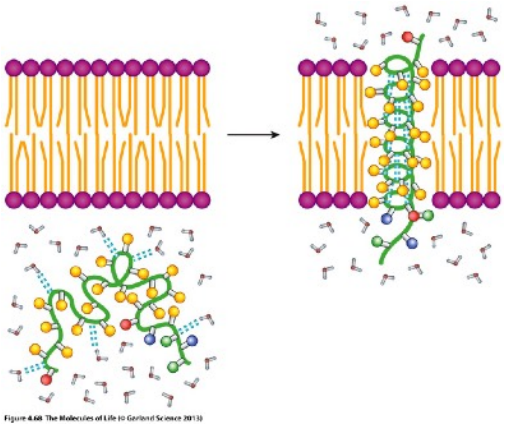


Figure 4.67 The Molecules of Life (© Garland Science 2013)

Given that these proteins are embedded in the membrane their surfaces are hydrophobic.



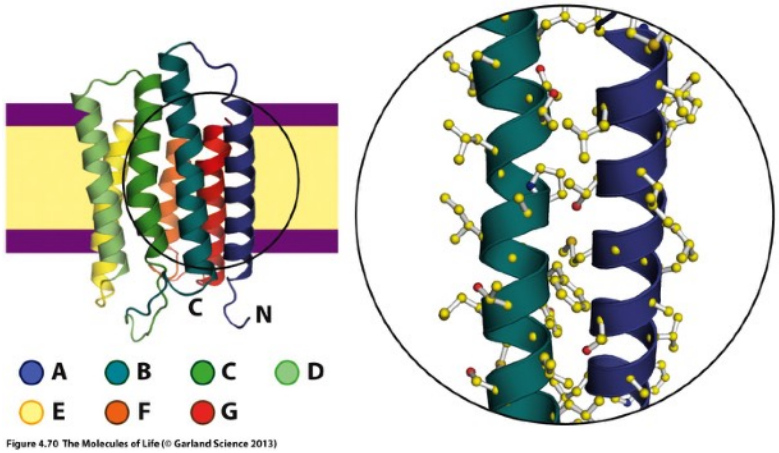
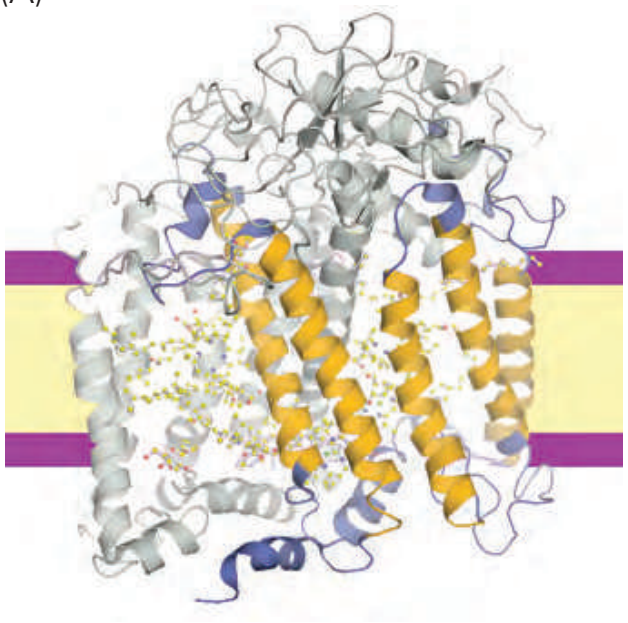
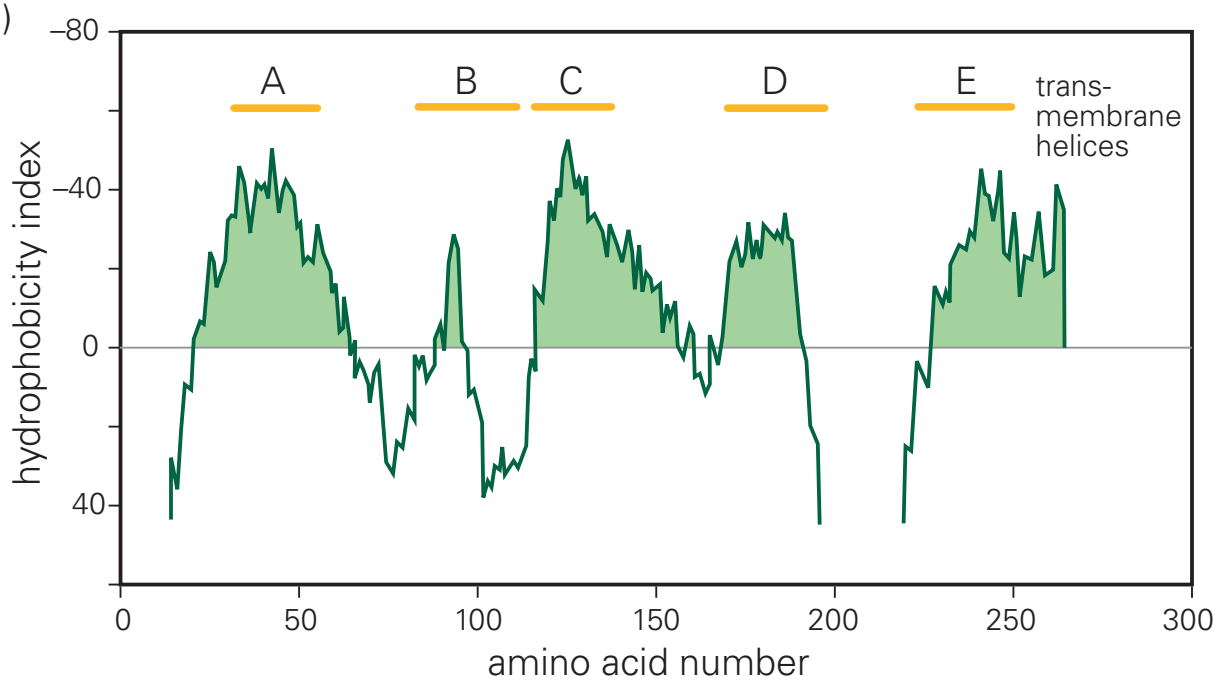


Figure 4.70 The Molecules of Life (© Garland Science 2013)

(A)



(B)



Trp	Phe	Leu	Ile	Met	Tyr	Val	Cys	Pro	His^0	Thr	Ser	Ala	Gln	Asn	Gly	Arg	His$^+$	Lys	Glu	Asp
–8.8	–7.1	–5.0	–4.5	–2.9	–2.9	–2.1	0.0	+0.4	+0.4	+0.8	+2.1	+2.1	+3.3	+3.8	+4.5	+7.5	+9.5	+11.7	+15.0	+15.0

Hydrophobic **Hydrophilic**

The **hydrophobicity index** at any position along the sequence is the aggregate value of the water/octanol transfer free energies for 19 contiguous residues in the sequence, centered on the residue in question.

https://dtu.biolib.com/app/DeepTMHMM/run

UniProt BLAST Align Peptide search ID mapping SPARQL Release 2022_03 | Statistics Help

Find your protein

UniProtKB ▾ Advanced | List Search

Examples: **Insulin**, **APP**, **Human**, **P05067**, **organism_id:9606**

UniProt is the world's leading high-quality, comprehensive and freely accessible resource of protein sequence and functional information. Cite UniProt ”

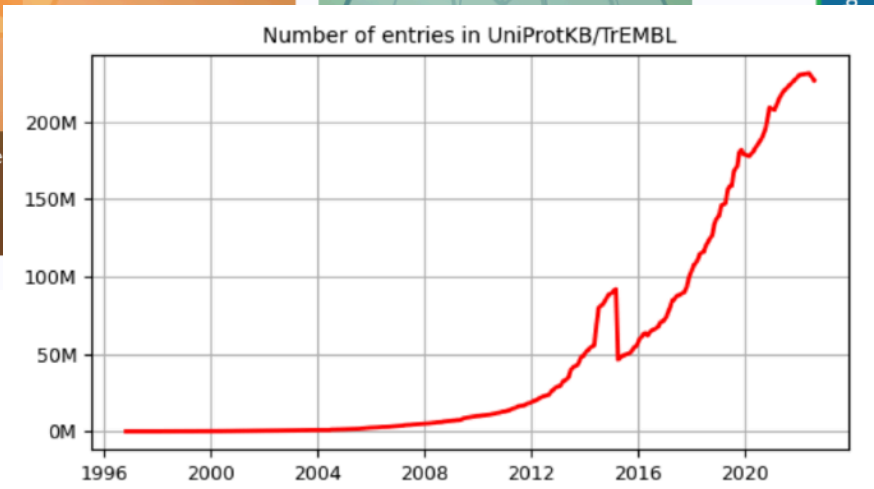
Accessing UniProt programmatically? Have a look at the **new API documentation**.
If you still need it, the **legacy version of the website** is available until the 2022_04 release.

Proteins
UniProt Knowledgebase
Reviewed (Swiss-Prot) 568,002
Unreviewed (TrEMBL) 226,771,948

Species
Proteomes
Protein sets for species with sequenced genomes from across the tree of life

Protein Clusters
UniRef
Cluste

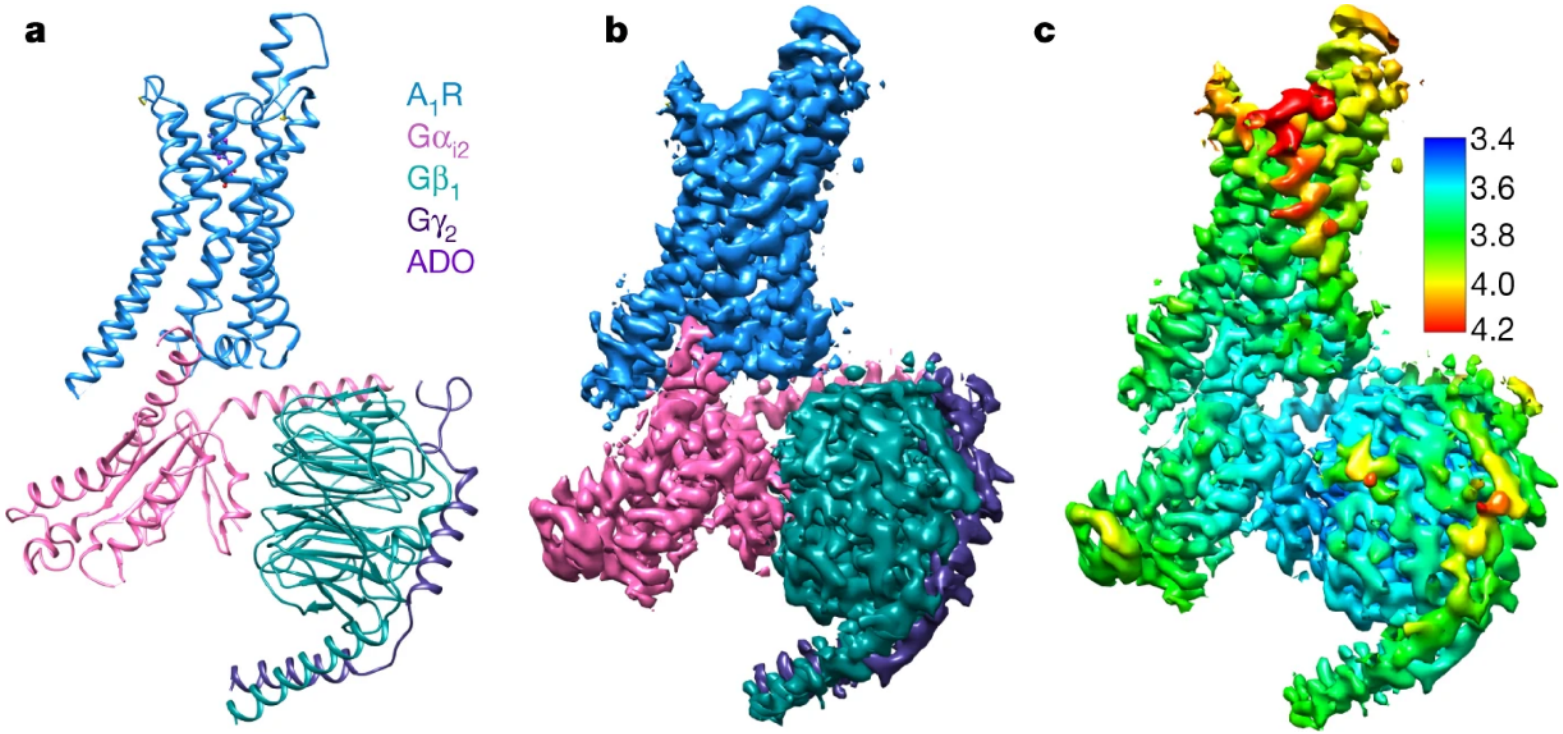
Sequence Archive
UniParc



https://www.uniprot.org

GPCRs are involved in sight, taste, smell, behavior, mood, and immune system regulation. Even though the signaling molecules, types of GPCR, and mechanisms of action may differ for all these roles, they all involve certain extracellular signals that are converted into a cellular response.

As such they are key targets for drug developments - it is estimated that ~700 approved drugs target GPCRs, implying that approximately 35% of approved drugs target GPCRs.



https://doi.org/10.1038/s41586-018-0236-6

Where to find protein structures?

http://www.rcsb.org

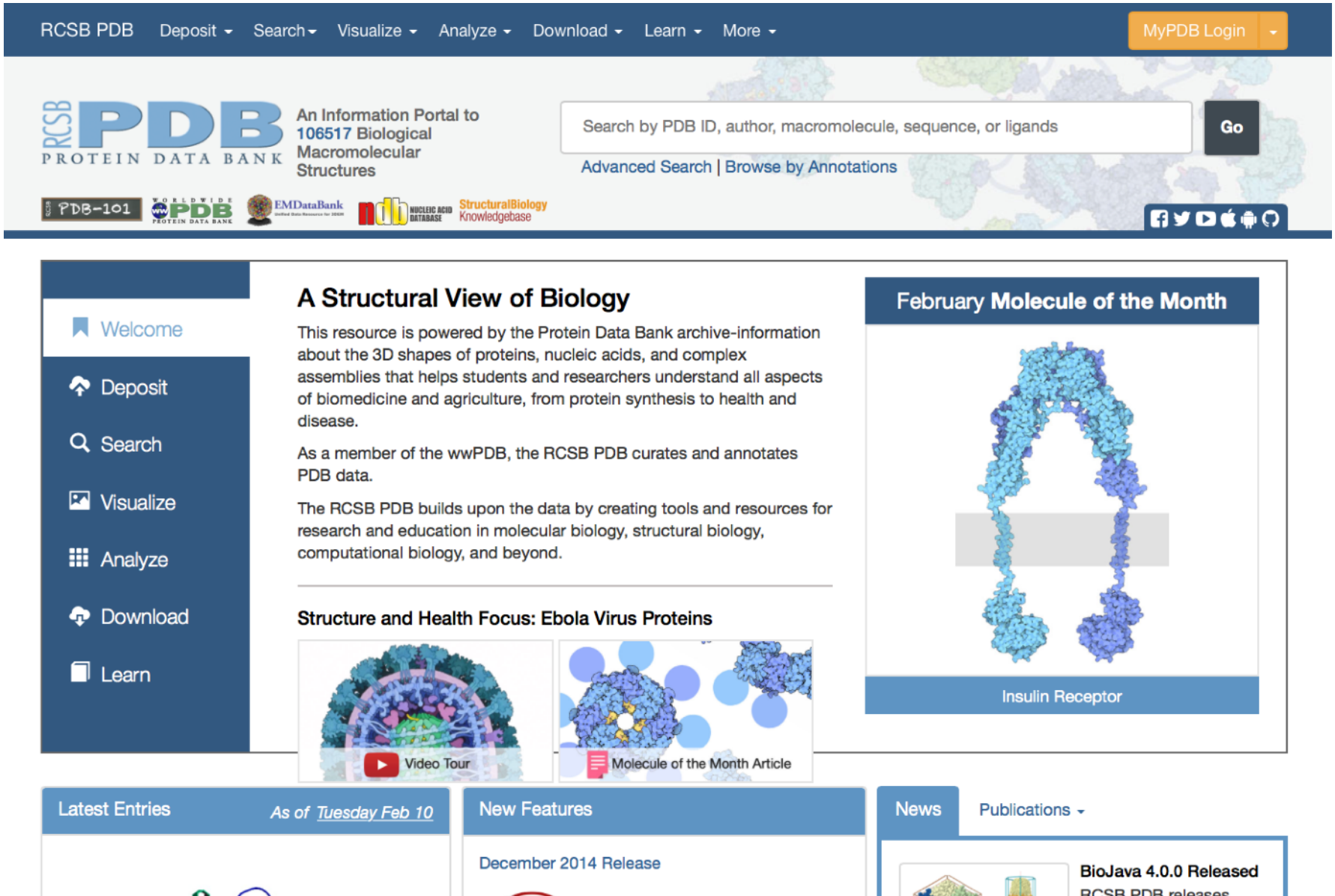


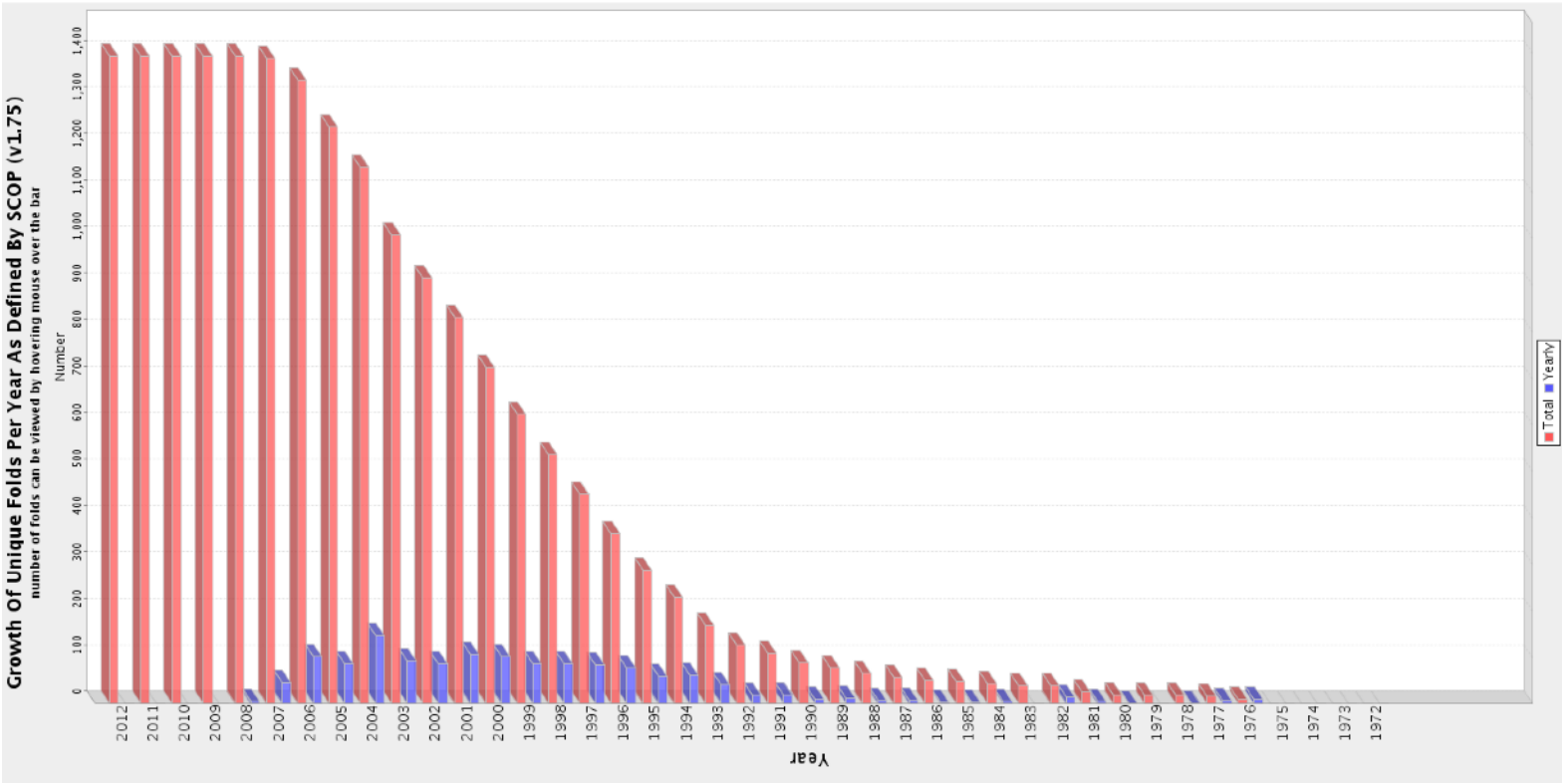
Table 5.1 Number of genes, families, and folds in different microorganisms.

Species	Number of proteins in the genome	Total number of families	Number of structurally characterized families	Predicted number of folds
Aggregate		4000–7000	1000	900–1300
M. genitalium	480	400–600	70	250–350
R. prowazekii	834	750–950	122	350–500
A. aeolicus	1522	950–1100	154	400–550
M. jannaschii	1715	850–950	74	300–400
A. pernix	1760	950–1000	62	300–450
Synechocystis sp.	3169	1700–2200	220	450–650
M. tuberculosis	3900	1500–2000	200	450–700
B. subtilis	4100	1800–2100	260	450–700
E. coli	4289	2000–2600	353	550–800
S. cerevisiae	6530	2400–4500	234	500–720

(Adapted from Y.I. Wolf, N.V. Grishin, and E.V. Koonin, *J. Mol. Biol.* 299: 897–905, 2000. With permission from Elsevier.)

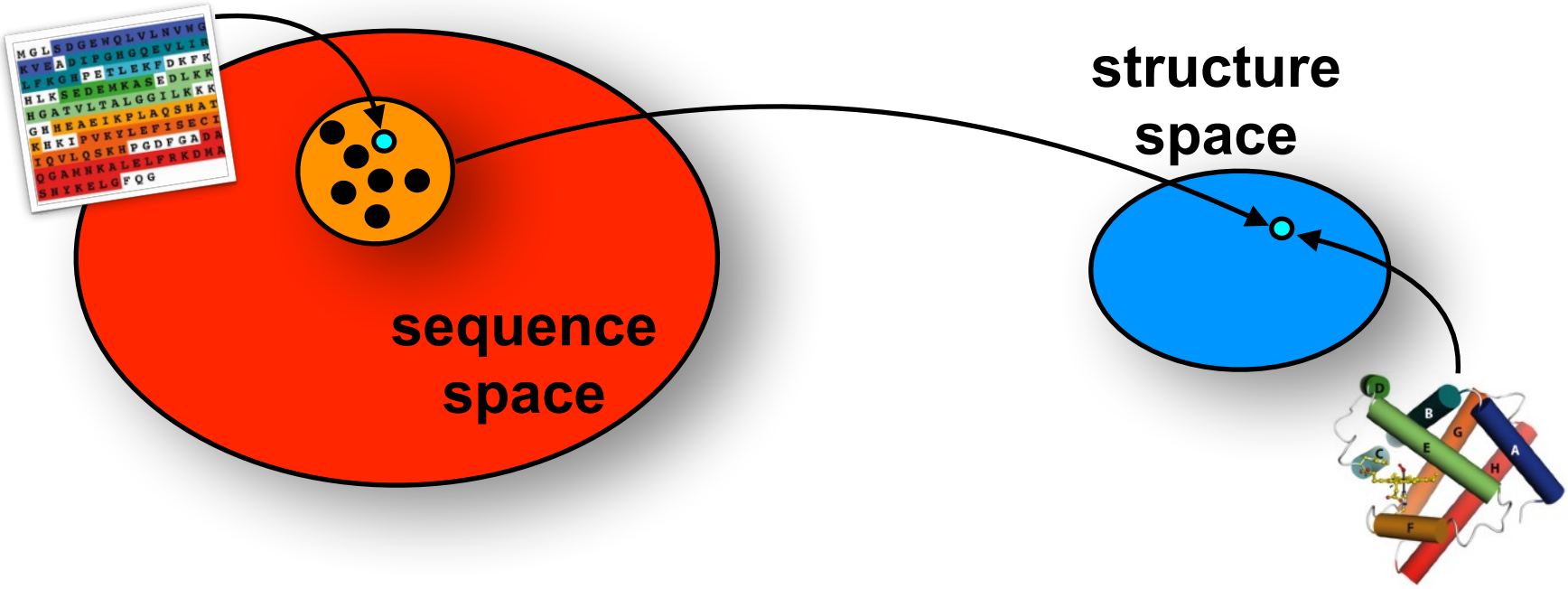
the number of folds seems to be limited across species

Protein folding motifs are limited



- It seems that all possible folds of proteins have been discovered already
- sequence space is much larger than folding space

The sequence space is enormous



- potential sequence space for protein of 150 a.a. ~20^{150} = ~10^{195}
- atoms in the observed universe ~10^{80}
- the sequences explored by evolution are much less (~$10^{10\text{-}20}$), structures lesser
- structure space is much more limited

Foldseek Search

Foldseek enables fast and sensitive comparisons of large structure sets

G LAB STEINEGGER LAB

Results: job.pdb_A METALLO-BETA-LACTAMASE

ALL DATABASES | AFDB-PROTEOME (373) | AFDB-SWISSPROT (732) | AFDB50 (1000) | CATH50 (330) | GMGCL_ID (472) | MGNIFY_ESM30 (1000) | PDB1C

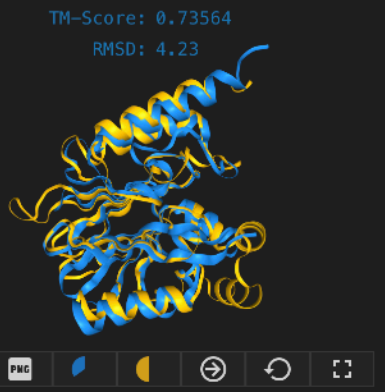
AFDB-PROTEOME 373 hits

GRAPHICAL | NUMERIC

Target	Description	Scientific Name	Prob.	Seq. Id.	E-Value	Position in query	Alignment
AF-Q86346-F1-model_v4	Possible bifunctional enzyme…	Mycobacterium tuberculosis …	1.00	21.3	3.46e-16	11–223	
AF-X8F1W6-F1-model_v4	Metallo-beta-lactamase supe…	Mycobacterium ulcerans str. …	1.00	20.5	2.42e-15	11–223	
AF-O07720-F1-model_v4	Lactamase_B domain-contai…	Mycobacterium tuberculosis …	1.00	19.6	1.88e-15	11–228	

```
Q  11 ITQLSDKVYTYVSLAEIEGMVPSNGMIVINNHQAALLDTPINDAQTETLVNWVADSLHAKVTTFIPNHWHGDC-IGGLGY
        +L+D V+    L   +        +V       L+DT    ++   +   V      +VT  +  H H D  +G   +
T   5 WERLTDSVHRC-RLPFCD----VTVGLVRGRTGILLVDTGTTLGEATAIAADVKQIAGCQVTHVVLTHKHFDHVLGSSVF
Q  90 LQKKGVQSYANQMTIDL----------AKEKGL-P----------VPEHGFTDSLTVSLDGMPLQCYYLGGGHATDNI
            + +     ++              A   G               P+HG      V L    +   + G GH  T ++
T  79 ---DQAEVFCAPEVVEYLRSATDRLREDALSYGADTAEVDRAIAALKPPQHGI-YDAAVDLGDRTVTITHPGSGHTTADL
Q 147 VVWLP----TEN--ILFGGCMLKDNQAT-SIGNISDADVTAWPKTLDKVKAKFPSARYVVPGHGDYGGTELIEHTKQIVN
      VV  P     +   ++F G ++++  A   I    +D+D+ AWP TLD+V A        VPGHG     +++   +  +
T 156 VVVAPATGHADGPTVVFTGDLVEE-SADPDID--ADSDLAAWPATLDRVLAIGGPDASYVPGHGKVVDAQFVRRQRAWLR
Q 220 QYIESTSKP
      ++  +P
T 233 T--RASRQP
```

Select target residues to highlight their structure



AF-U7Q482-F1-model_v4	Lactamase_B domain-contai…	Sporothrix schenckii ATCC 5…	1.00	19.8	5.15e-12	3–225
AF-O69728-F1-model_v4	Possible hydrolase	Mycobacterium tuberculosis …	1.00	13.9	1.50e-11	11–219
AF-Q68D91-F1-model_v4	Metallo-beta-lactamase dom…	Homo sapiens	1.00	14.8	4.26e-12	3–227
AF-Q9I5I9-F1-model_v4	SDS hydrolase SdsA1	Pseudomonas aeruginosa P…	1.00	15	1.16e-11	11–225
AF-D4A249-F1-model_v4	Metallo-beta-lactamase dom…	Rattus norvegicus	1.00	14.7	1.24e-11	3–223
AF-X8FG82-F1-model_v4	Metallo-beta-lactamase supe…	Mycobacterium ulcerans str. …	1.00	17.5	5.25e-11	9–220
AF-C0NH59-F1-model_v4	Lactamase-like protein nscB	Histoplasma capsulatum G1…	1.00	17.7	1.09e-11	3–223
AF-Q57544-F1-model_v4	Hydroxyacylglutathione hydr…	Haemophilus influenzae Rd …	1.00	19.2	5.48e-12	22–219
AF-P32717-F1-model_v4	Putative alkyl/aryl-sulfatase …	Escherichia coli K-12	1.00	11.9	4.63e-11	11–227
AF-A0A044SS06-F1-model_v4	Lactamase_B domain-contai…	Onchocerca volvulus	1.00	20	1.03e-11	3–222
AF-O06154-F1-model_v4	Conserved protein	Mycobacterium tuberculosis …	1.00	19.2	4.35e-11	

Nature Biotechnology doi: 10.1038/s41587-023-01773-0

Proteins - Discussion

Statement: can two proteins with different sequence have the same fold?

-Yes or No ?

- Why ?

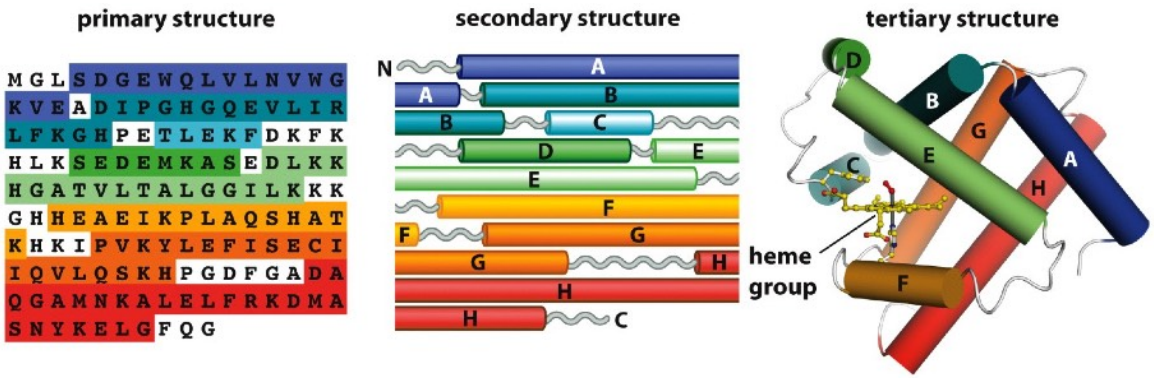
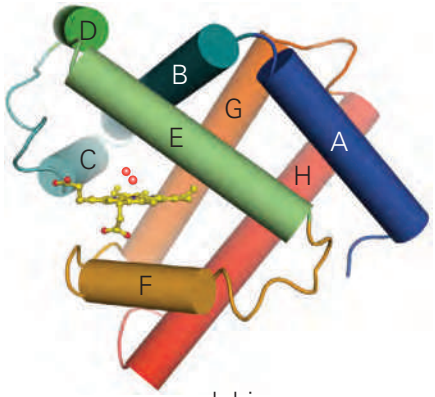
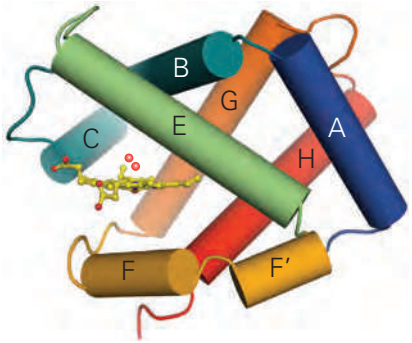


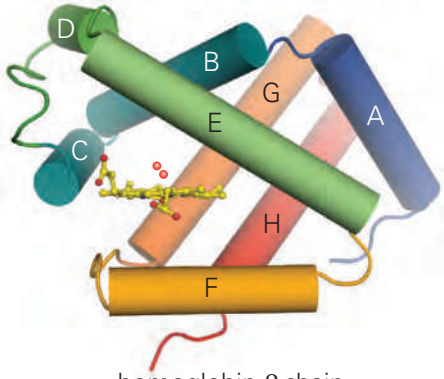
Figure 4.1 The Molecules of Life (© Garland Science 2013)



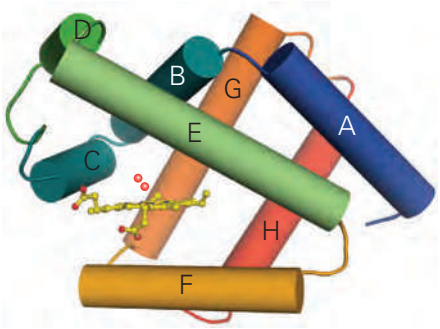
myoglobin



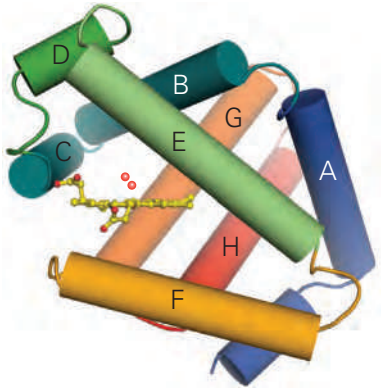
hemoglobin α chain



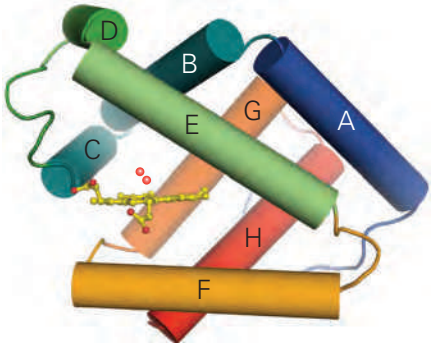
hemoglobin β chain



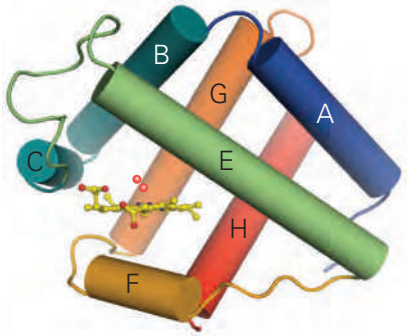
erythrocruorin



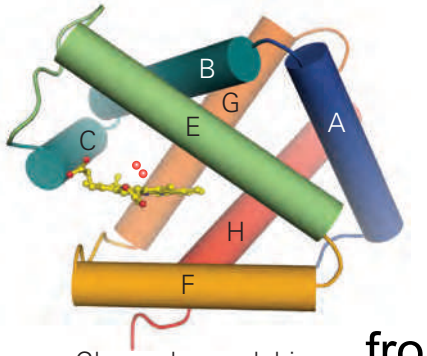
clam hemoglobin



worm hemoglobin



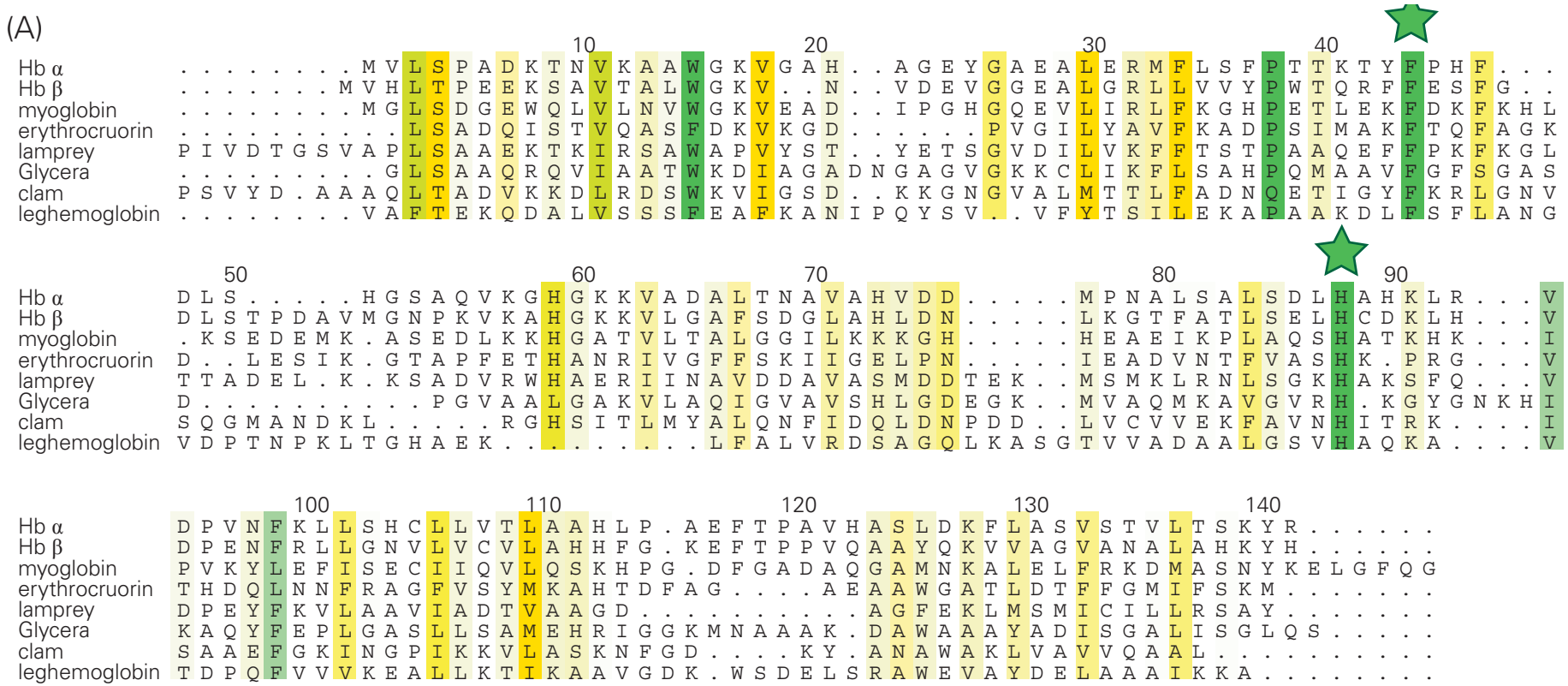
leghemoglobin



Glycera hemoglobin

globin fold

from different organisms

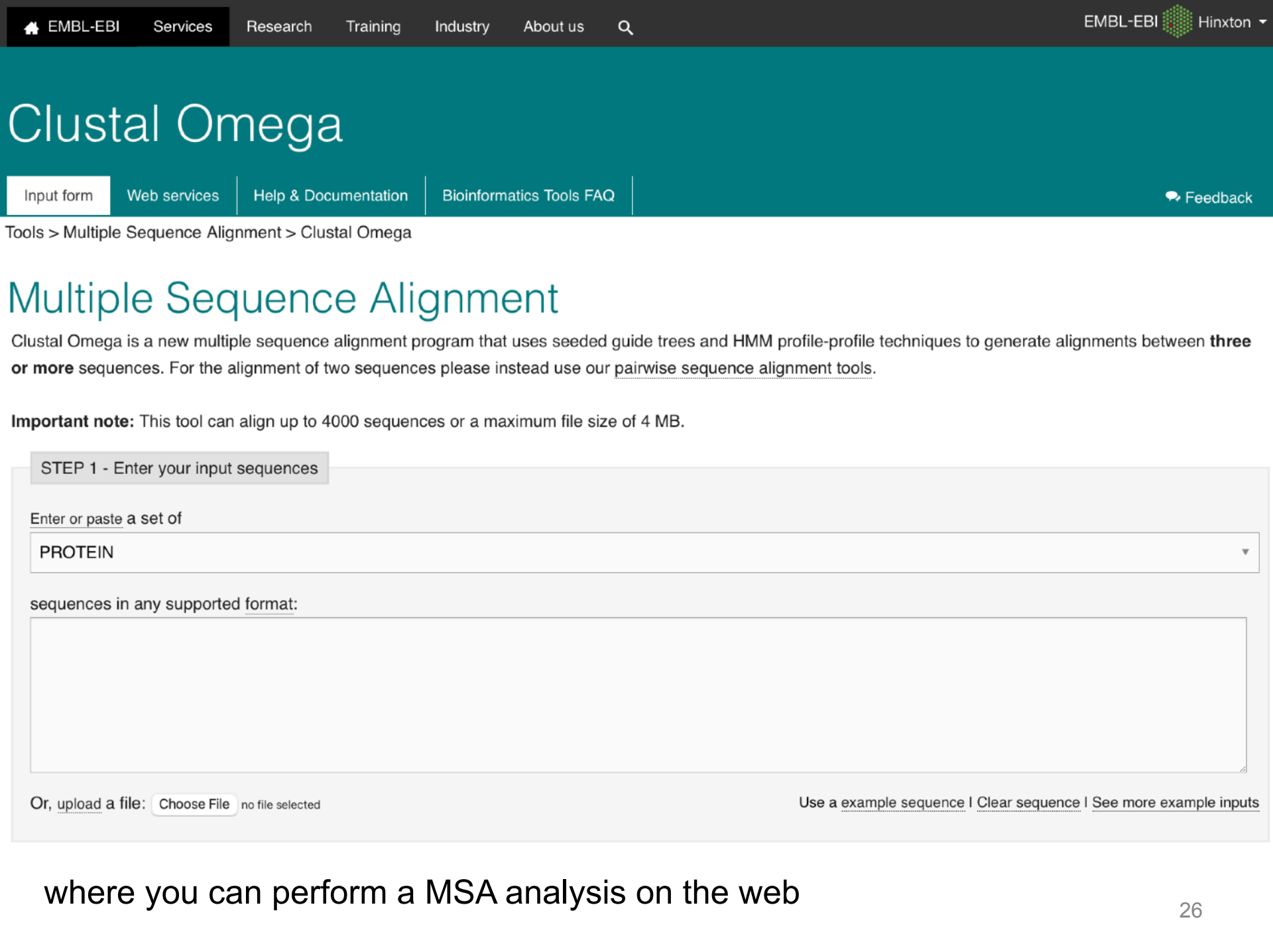


if you do a MSA of global sequences you can see that their identity can be quite low

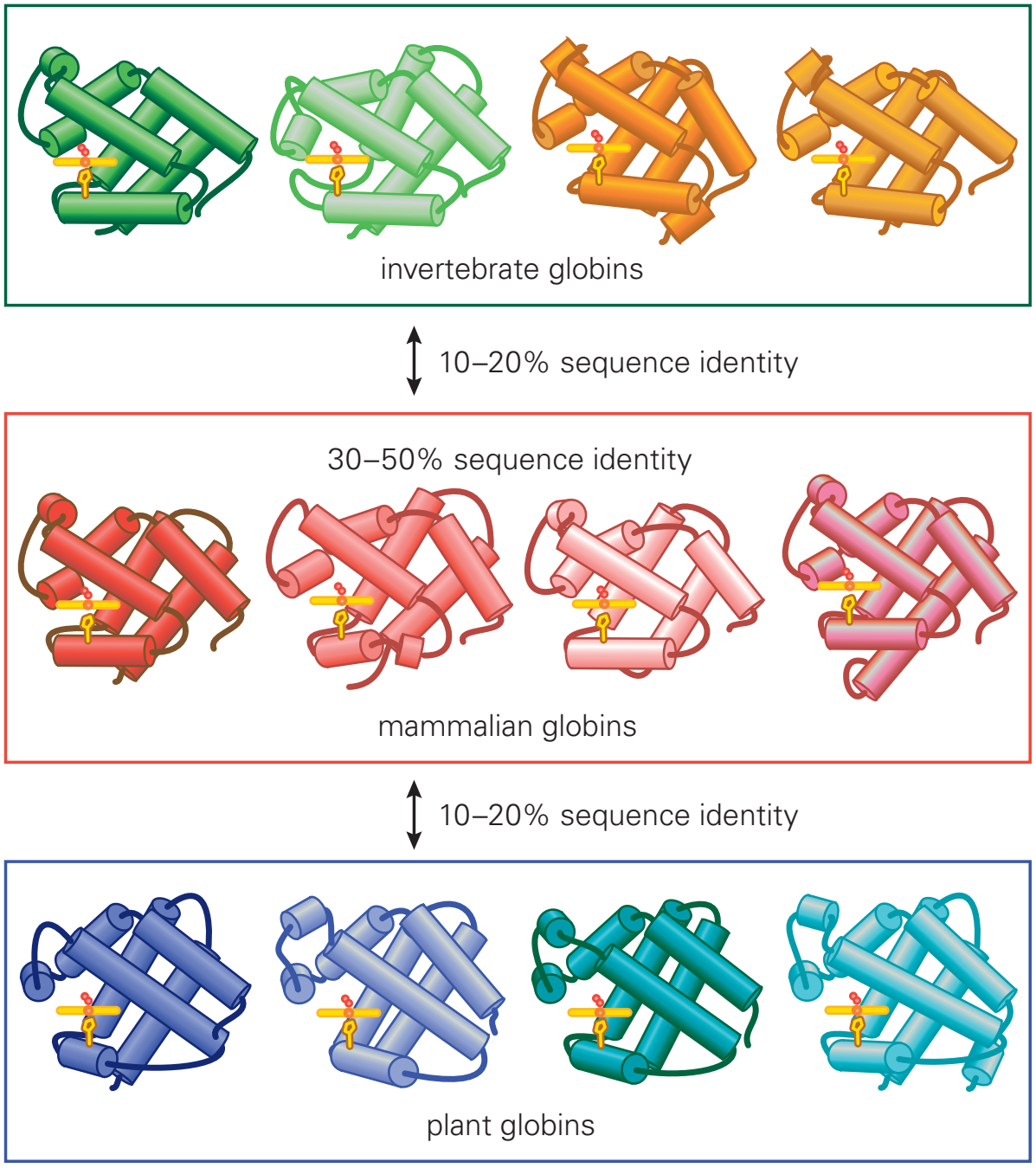
notice the green sites that are highly conserved - these are the ones involved in holding the heme group, they are conserved to preserve the function

multiple sequence alignment (MSA)

(read chapter 5.6-5.11 BLOSUM-62 matrix in preparation of practical of week 6-7)



where you can perform a MSA analysis on the web

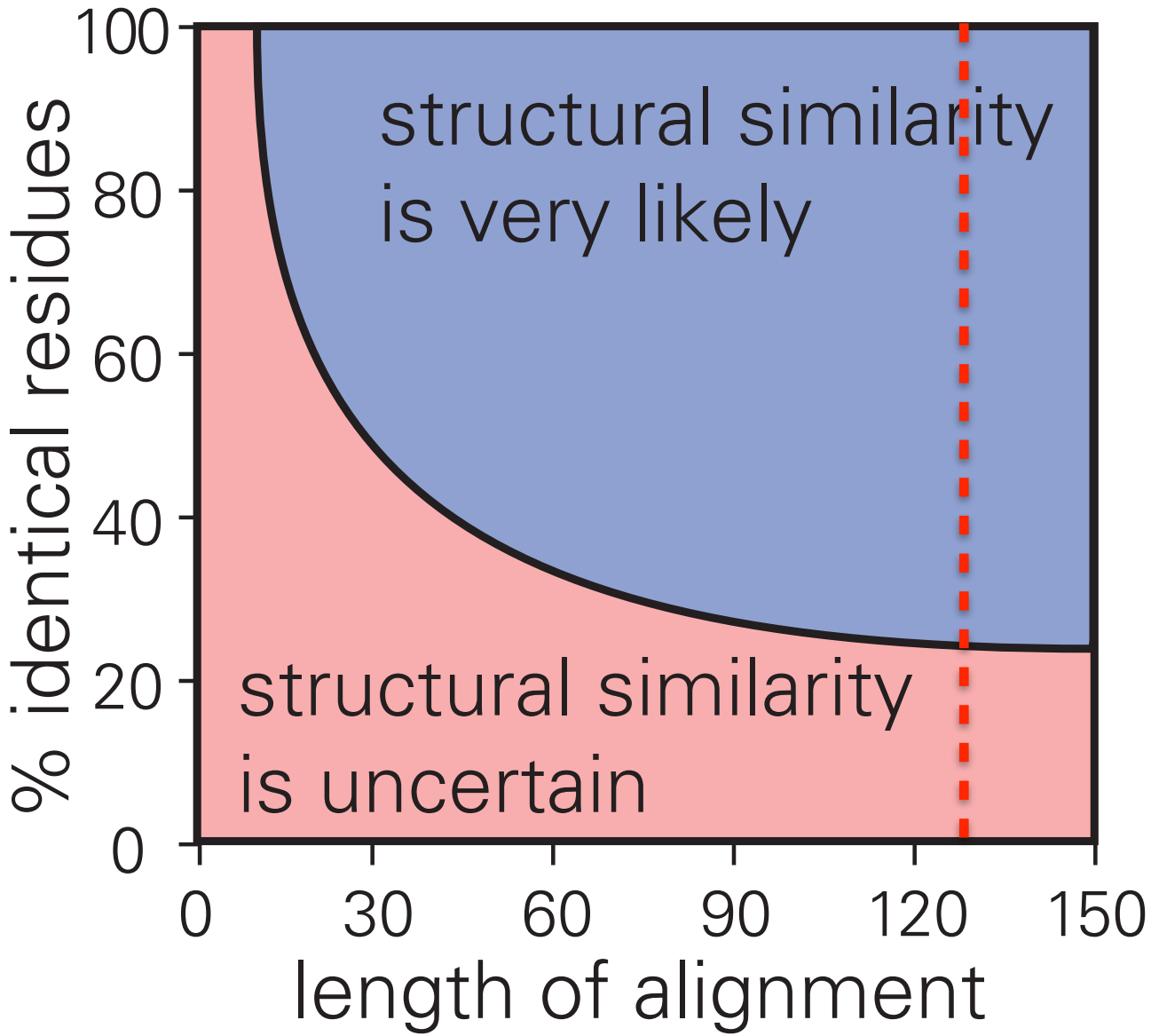


different organisms have very sequence identity. This can be high within more similar organisms but can be as low as 10% if you compare plants and mammals

with only 10% sequence identity you can still preserve the same overall fold

This is telling you that the structural space is more confined that the sequence space and will give you a practical way to model structure from sequence

practical guidelines to assess if you can structurally model a sequence based on the structure of a known protein that has a certain % of homology (~identity)



Proteins – Take Home Messages

- Proteins fold spontaneously minimising their free energy

- Anfinsen experiment demonstrated the direct relationship between sequence, structure and function

- Proteins fold thanks to the hydrophobic effect creating an hydrophobic core and hydrophilic surface

- The hydrophobic effect is the principal driving force underlying protein folding

- Membrane proteins do not follow the same rule as they have to partition to the hydrophobic membrane environment

Thermodynamics - the system

Revisitation of some thermodynamic concepts with particular relevance to biochemistry and biological systems (aka **bioenergetics**)
- In order to be able to study a system, one has to define **what that system is:**

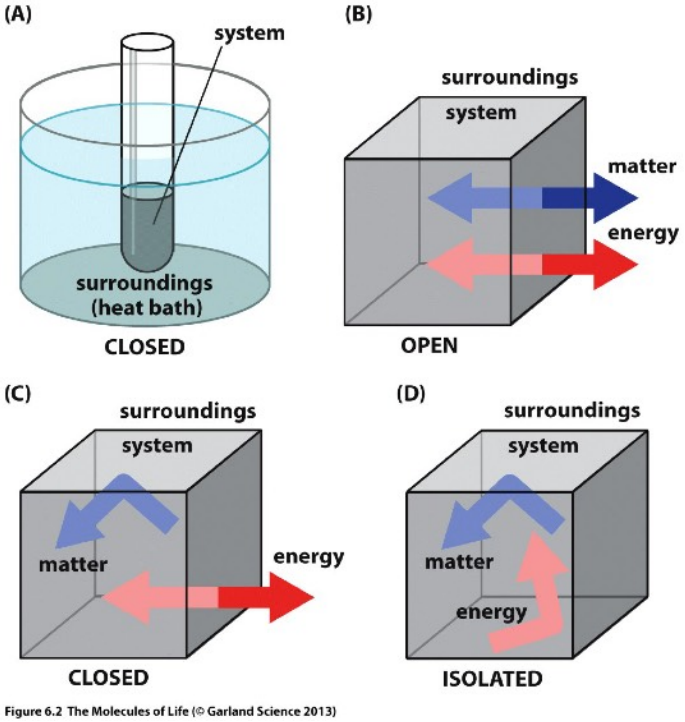
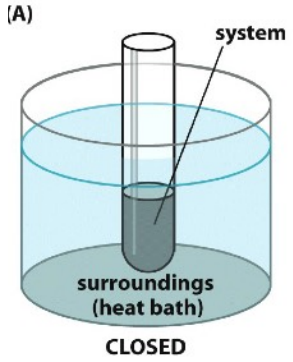


Figure 6.2 The Molecules of Life (© Garland Science 2013)

- In (A) we have an example of a **closed system** in a test tube where we can keep track of the flow of energy in a meaningful way. This is the most common situation when we work in the lab (*in vitro*), but cells can be also sometime considered open systems.
- In such system one can measure the **heat** released or taken up as a process proceeds. This can be measured by **calorimeters**.

Work and heat

- Energy can be exchanged between a closed system and the surroundings by doing work (*w*) or by heating (*q*).
- **Work** and **heat** are two **modes of transfer of energy** - not energy forms per se
- Eg an exothermic reaction occurring in the system – energy is released



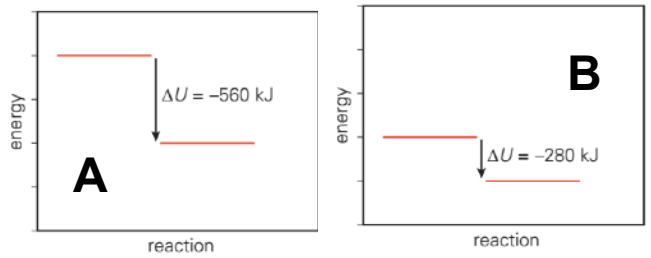
Combustion of glucose

$$C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O \qquad \Delta U = -2801\ \text{kJ}\bullet\text{mol}^{-1}$$

$$\boxed{\Delta U = U\,(\text{final}) - U\,(\text{initial})}$$

where *U* is the internal energy of the system (U = K+P, kinetic + potential energy)

This reaction is called **exothermic** because the energy of the product state is lower. On the opposite when the energy of the products is higher, the reaction is called **endothermic.**



which reaction is the most favorable?

Work and heat

But what are heat and work? how do we interpret this energy transfer at the molecular level?

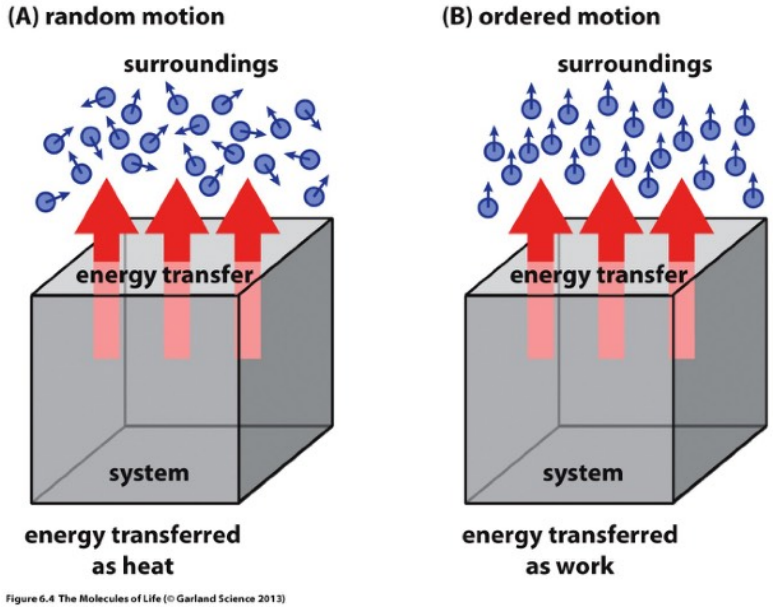


Figure 6.4 The Molecules of Life (© Garland Science 2013)

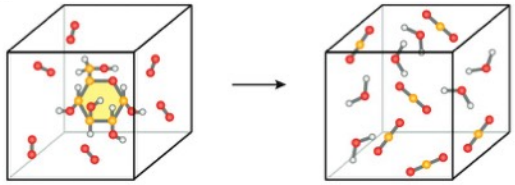
- Energy transferred to the surroundings as **heat** stimulates the random motion of molecules in the surroundings (i.e. increases their velocity and temperature)

- When the system does mechanical **work** on the surroundings, it causes the order movement of some part of the surroundings - this energy can be better stored

Work and heat

But what are heat and work? how do we interpret this E transfer at the molecular level?

Combustion of glucose



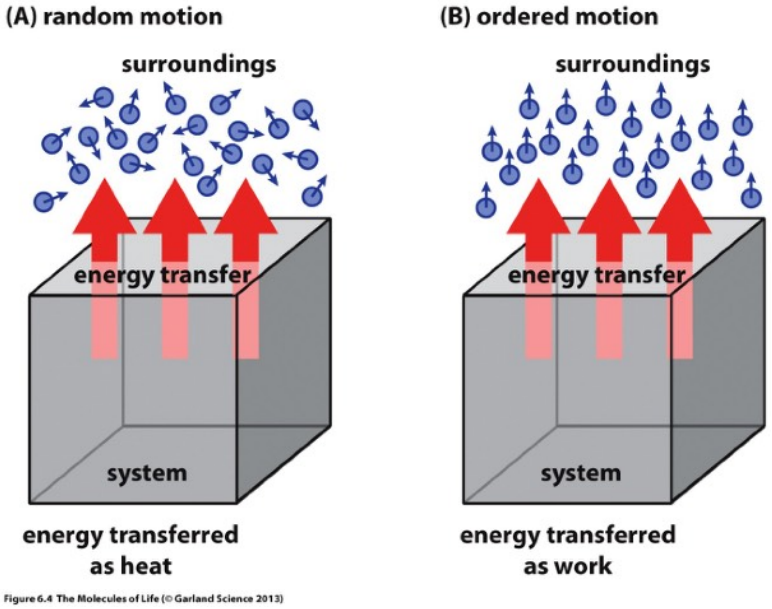
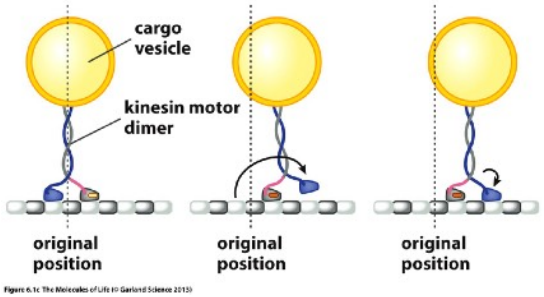


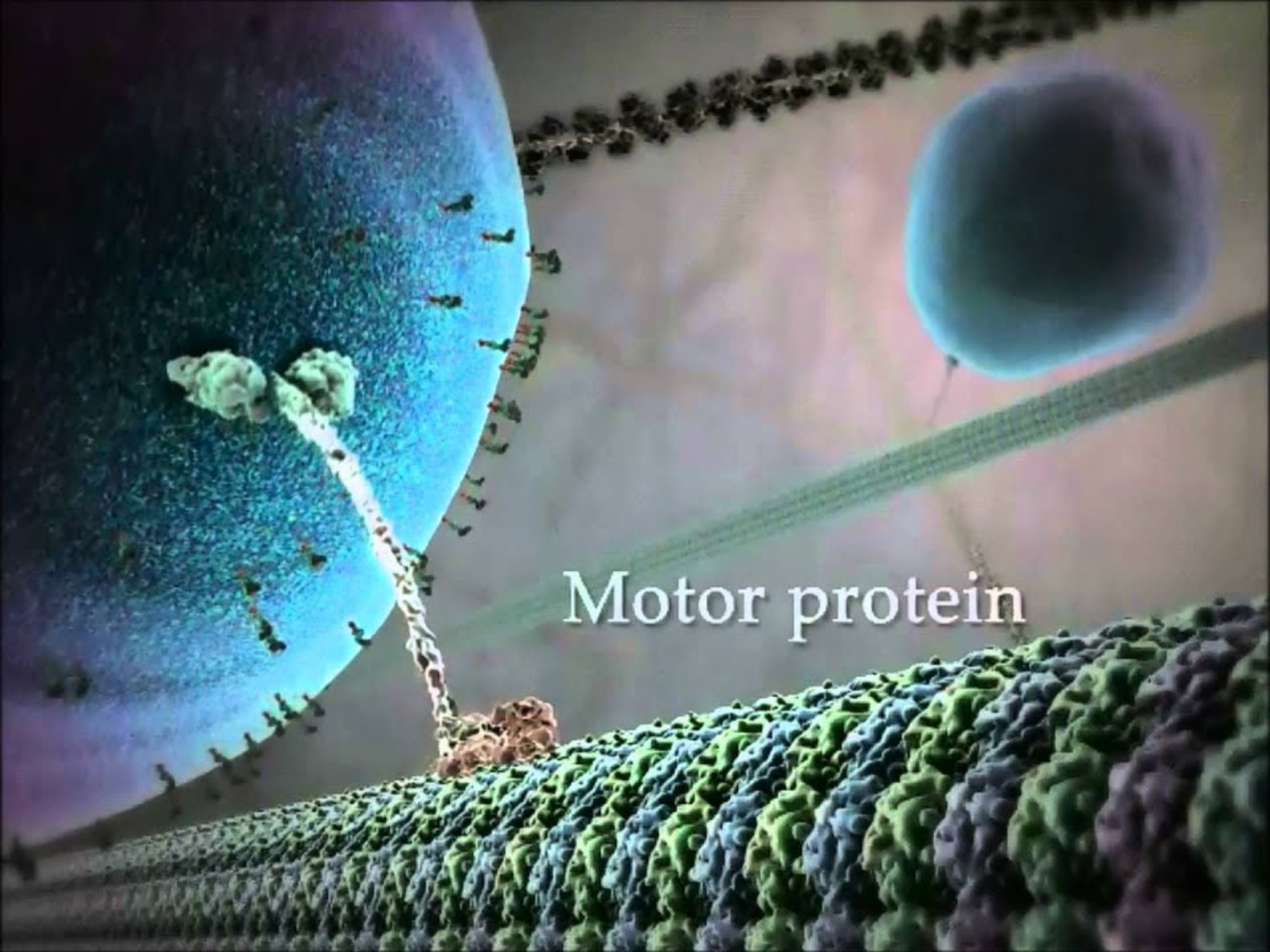
Figure 6.4 The Molecules of Life (© Garland Science 2013)

Kinesin walks



- Energy transferred to the surroundings as **heat** stimulates the random motion of molecules in the surroundings (i.e. increases their velocity and temperature)

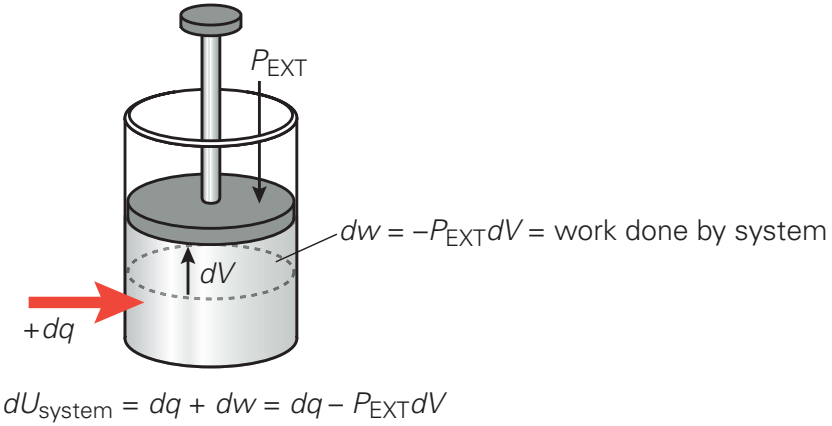
- When the system does mechanical **work** on the surroundings, it causes the order movement of some part of the surroundings - this energy can be better stored


Motor protein

1st Law of Thermodynamics or Conservation Law

- In all physical and chemical processes the total energy of the system and the surroundings stays constant (or of the system only if it is isolated)

$$dU_{\text{total}} = dU_{\text{system}} + dU_{\text{surroundings}} = 0$$

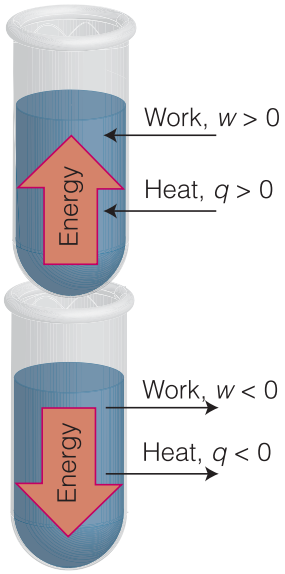


- consider an example system where you can control/measure both temperature and pressure (work)

- In principle, to keep track of the changes of internal energy in the system we need to account for both changes in heat and work that are transferred

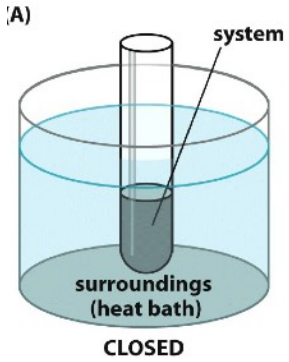
$$dU = dq + dw$$

How do we measure work and heat?



sign convention in thermodynamics:

- *w* and *q* are positive if *E* enters the system
- *w* and *q* are negative when *E* leaves the system (for example glucose combustion, or any exothermic reaction)



How do we measure work?

Experimentally we can measure the work done by expansion of the system against an external pressure p_{ext} (for instance reactions producing gases at const pressure):

$$w = -p_{ext}\,\Delta V$$

where the minus sign is because the system does work, thus loses energy when it expands. This is not a common case in bioenergetics, as volume remains usually constant.

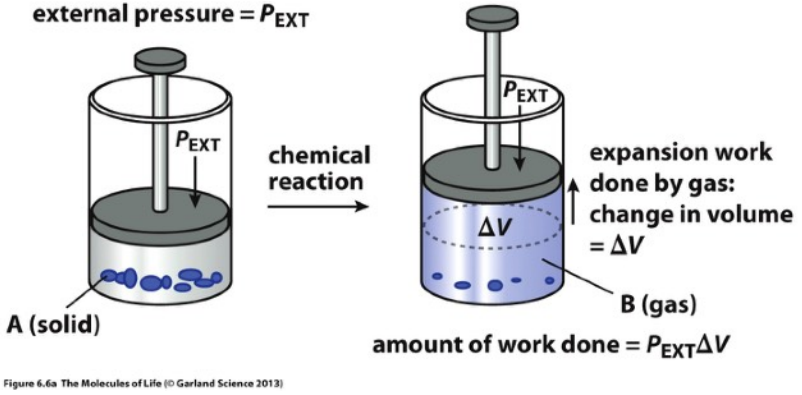


Figure 6.6a The Molecules of Life (© Garland Science 2013)

How do we measure work and heat?

How do we measure heat?

Remember the isothermal expansion of an ideal gas (isolated system), in this case $w = -q$, and this will give you a way to measure q from the work done.
When a system is heated, its temperature T rises, and the ΔT depends on the internal properties of the substance - this is called ***heat capacity C***, defined as:

$$\boldsymbol{C = q / \Delta T} \text{ , or } \boldsymbol{q = C \Delta T}$$

Heat capacity is defined as the amount of heat required to increase the temperature of the system by 1 degree kelvin.

C is an *extensive property* (depends on the size of substance, units kJ/K), thus you can also work with the molar heat capacity that C per mole (kJ/K*mol).

You might recall from Physics II that *C* changes if considered at constant volume *V* (C_V) or constant pressure *p* (C_p), **in bioenergetics the most relevant one is C_p** as in the lab or in the cell processes happen at constant pressure.

Note: 1 cal = 4.184 J is defined by the heat capacity of water, and 1 cal is the amount of heat required to raise the T of 1 g of water by 1 K at 287.5 K and 1 atm.

Internal energy (U)

The **internal energy *U*** is the sum of the kinetic and potential energy of your system, and while it might be difficult to calculate this in absolute terms, it is possible to calculate change in *U* given by transfer of work and heat:

$$dU = dw + dq$$

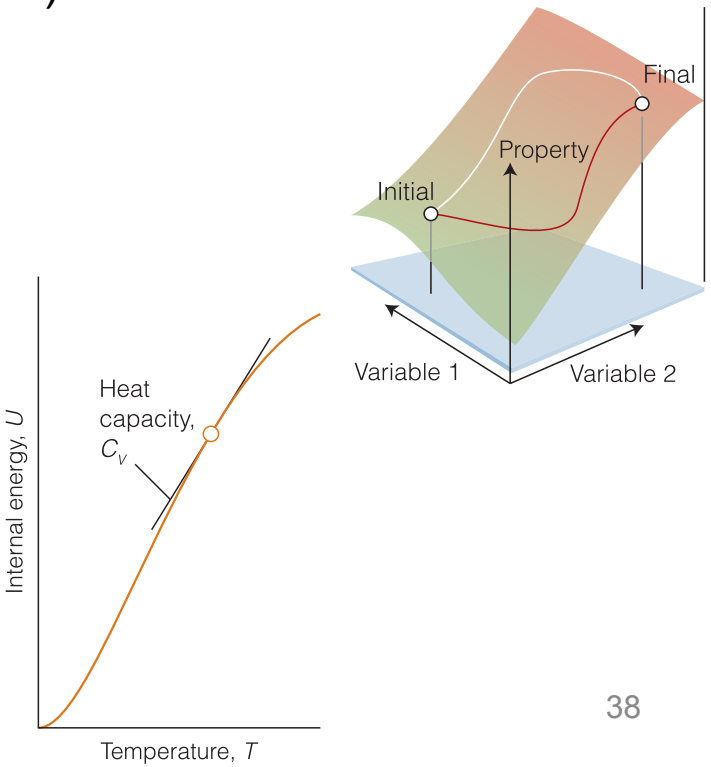
for instance in the case of ideal gas isothermal expansion $\Delta U = 0$. Also remember that you can work with U per mole (measured in kJ/mol). **U is a state function** !

This definition of ΔU provides a way to measure the internal energy of a system when a process/reaction takes place. In fact if we work at constant volume the work = 0 and thus $\Delta U = q_V$ (this for instance could be the case of a biological cell).

if $\Delta U = q_V$ then $C_V = \Delta U / \Delta T$ or

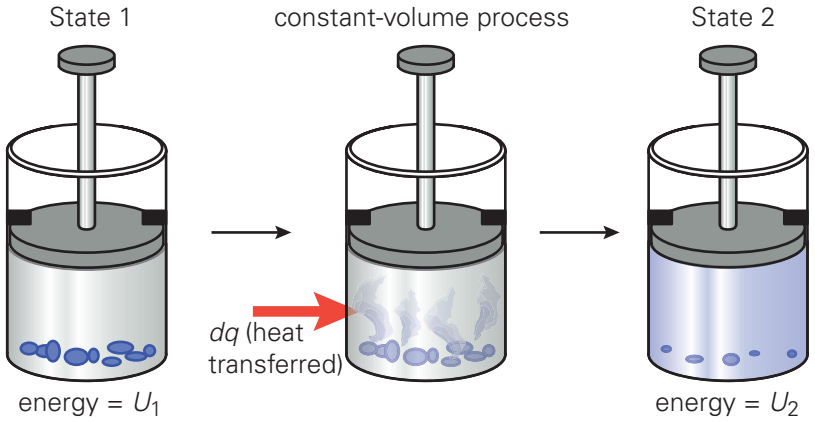
$C_V = dU / dT$ in differential form, or

$C_V = [\partial U / \partial T]_V$

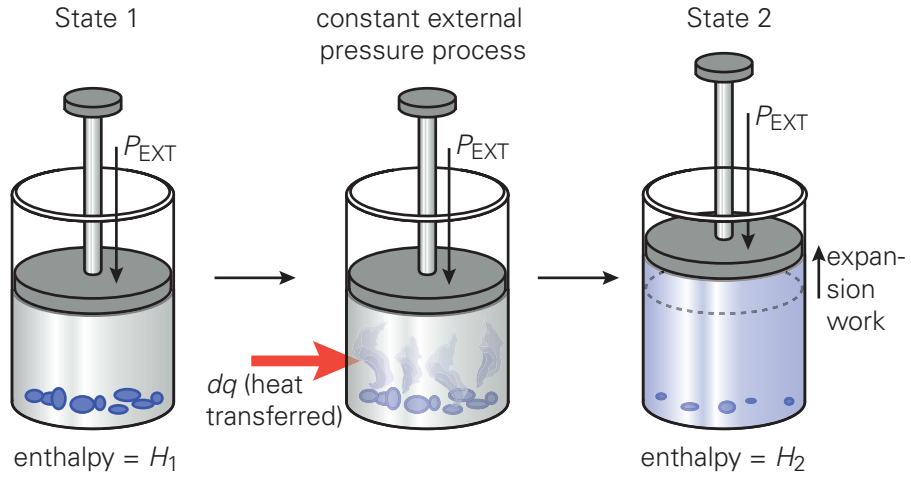


Enthalpy

If we work at constant pressure, then you can have the case in which there is no change in volume (w=0) where the heat transferred is ***q = ΔU***



or the system is free to change volume and the heat is ***q = ΔU + p ΔV***



Enthalpy

Therefore to characterize these types of systems we introduce another variable of the system – **Enthalpy (H)** as the sum of the internal energy *U* and pressure times volume (*pV*) that takes into account the the potential work done by the system

$$H = U + pV$$

thus the change of *H* in case the pressure remains constant is :

$$\Delta H = \Delta U + p\Delta V$$

however, the change in *H* boils down to the heat transferred at constant pressure

$\Delta H = q_p$ or in differential notation $dH = dq$

this is because $\Delta U + p\Delta V = (w + q) + p\Delta V = (-p\Delta V + q) + p\Delta V = q$

Thus now we have identified **a state function *H*, enthalpy** that allows us to characterise processes at constant pressure. Thus exothermic reaction are now characterised by $\Delta H < 0$ and vice versa endothermic reactions have $\Delta H > 0$

Heat Capacity at constant pressure

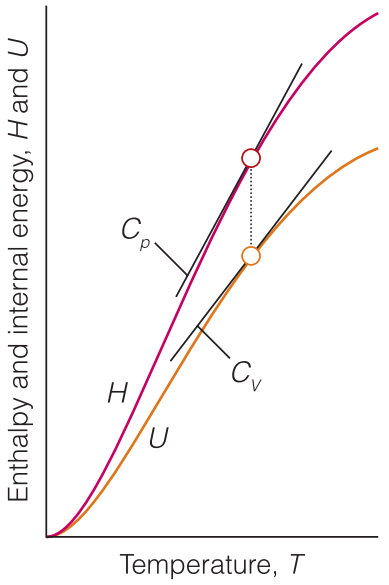
Even more important is that enthalpy can be measured experimentally (calorimetry) through the heat-capacity (C_p). In fact if $\Delta H = q_p$ and $C=q/\Delta T$, you obtain that $C = \Delta H/\Delta T$ at constant pressure or

$\boldsymbol{C_p = \Delta H / \Delta T}$ or $\boldsymbol{\Delta H = C_p \Delta T}$ or in differential notation $\boldsymbol{dH = C_p\, dT}$ *or* $\boldsymbol{C_p = [\partial H/\partial T]_p}$

Heat capacity is defined as the amount of heat required to increase the temperature of the system by 1 kelvin.

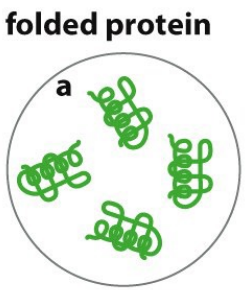
Moreover, C at constant pressure is always higher than at constant volume, as remember from ideal gases that $C_p - C_V = R \sim 8.314\ \mathrm{J \cdot K^{-1} \cdot mol^{-1}}$

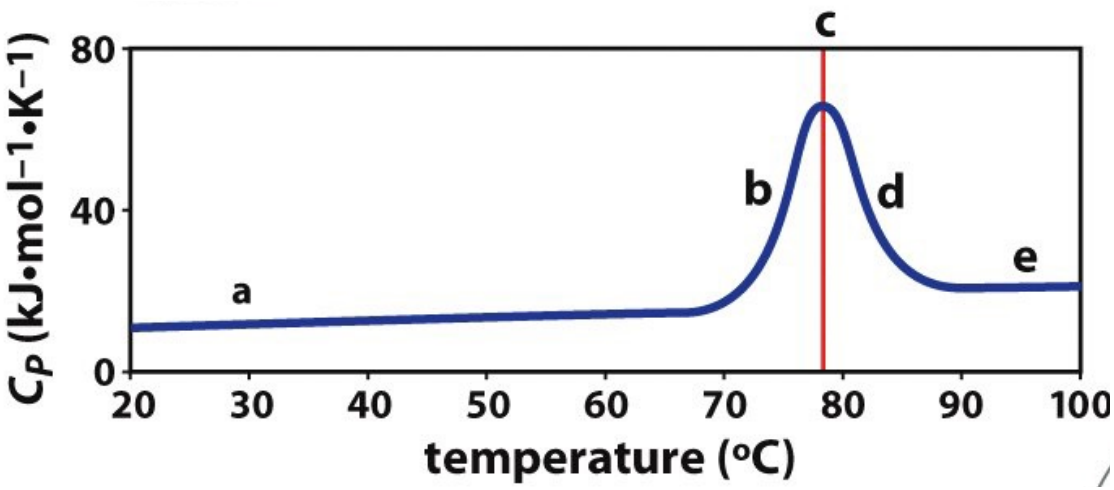
This is because if a system is free to expand at constant pressure, some energy supplied as heat can be transferred to the surrounding as work. This also implies that H is always greater than U.

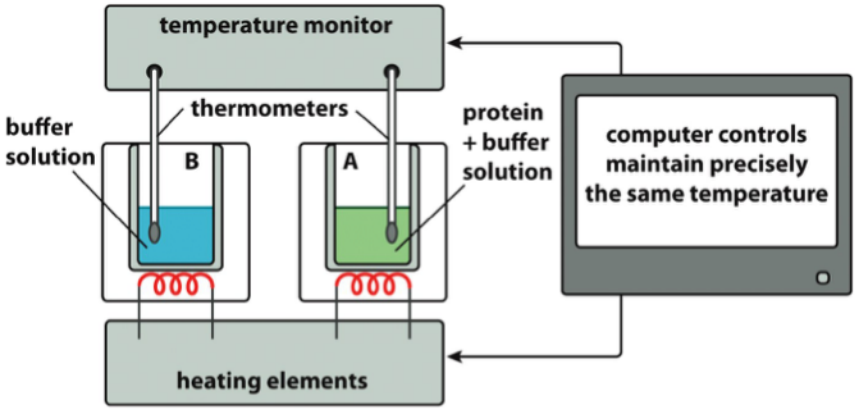


Heat Capacity at constant pressure

Practical example in biochemistry: unfolding of a protein by temperature.
This is a usual results from a calorimetric experiment



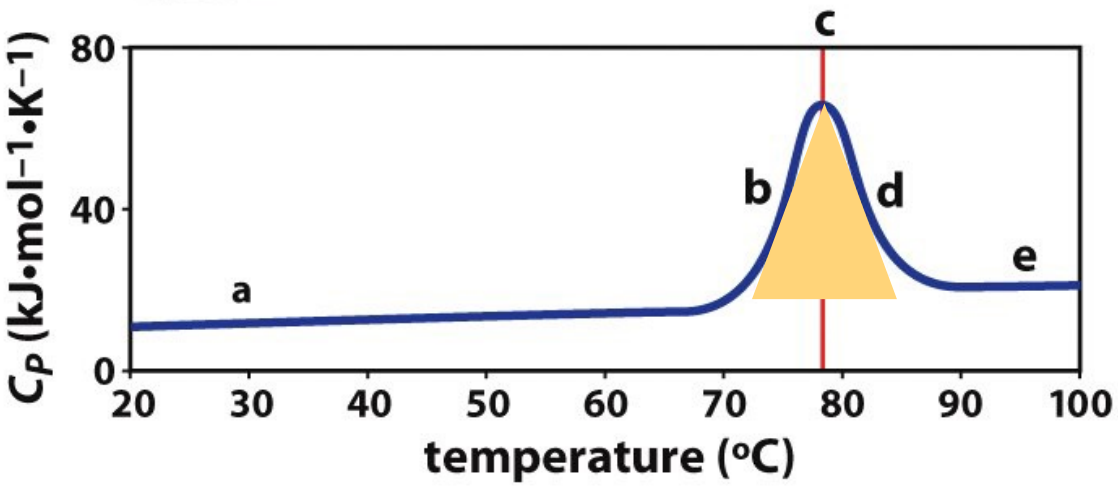




Differential Scanning Calorimeter

Let's Discuss !!!!!

Tm= melting T

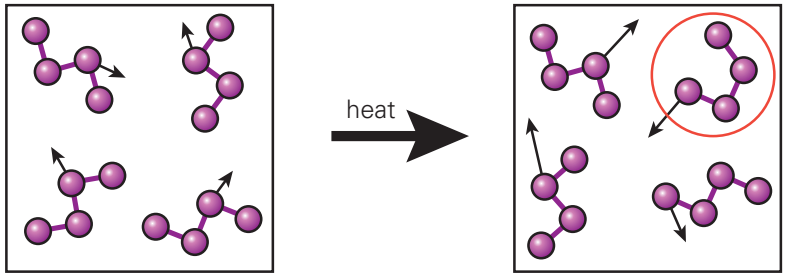


A. Pick one of the letters located in a region of the curve where:
   1) Protein is fully folded
   2) Protein is fully unfolded
   3) Half of the protein molecules are unfolded

B. If I want to determine the enthalpy of the unfolding process – how can I obtain such a value:
   1) Highest Cp value
   2) Area under the curve
   3) Delta between starting and finishing Cp

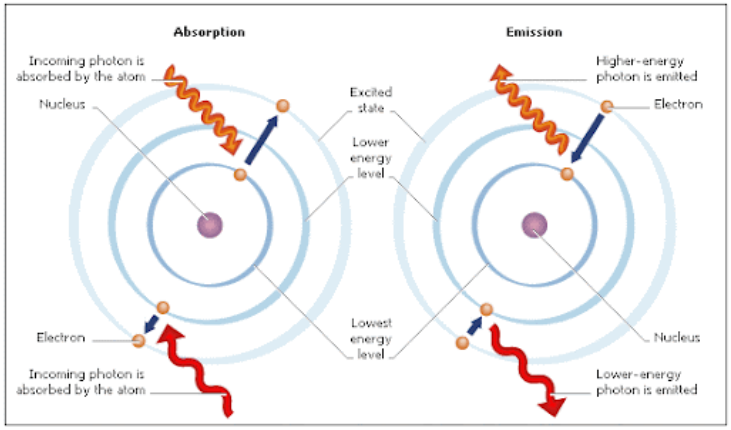
What to know ...

- Energy released by chemical reactions is converted into heat and work.
- The total energy of the system and the surroundings is conserved - this is the first law of thermodynamics.
- The heat transferred to a system under conditions of constant pressure is equal to the change in enthalpy of the system.
- The heat capacity of a macromolecular solution increases and then decreases with temperature as the macromolecule unfolds because it can take up energy by passing to conformations at higher energy



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:

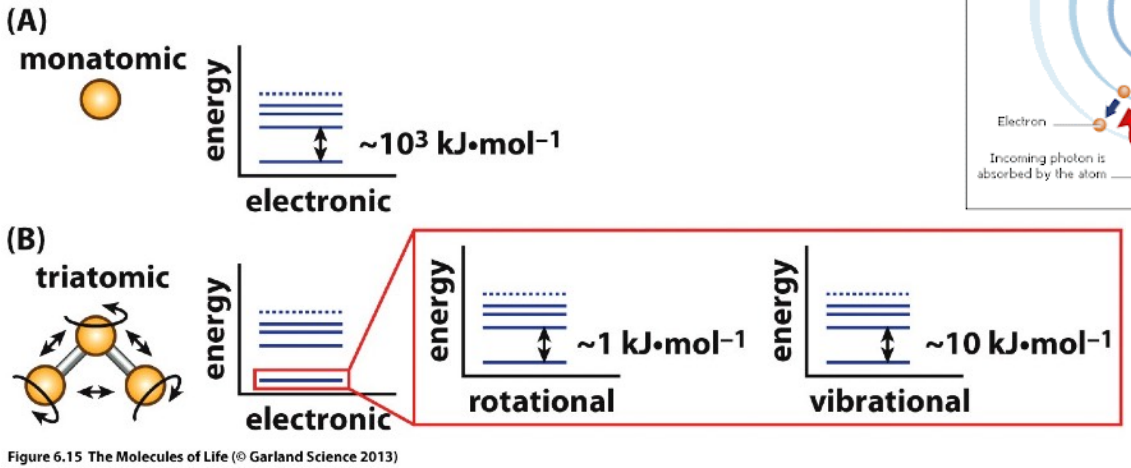


Figure 6.15 The Molecules of Life (© Garland Science 2013)

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

- The **Boltzmann distribution** provides the theoretical framework to quantify the populations at different energetic levels

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 E_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 (this is called the multiplicity W=M!/(N!(M-N)!), check the book for more details) and it is described by the **Boltzmann distribution**:

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

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:

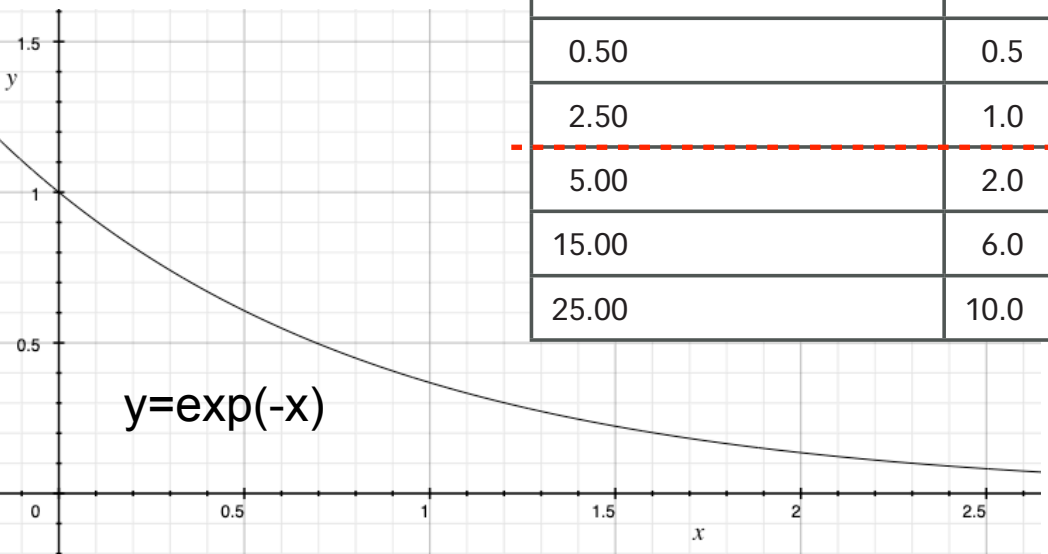
$$\frac{N_2}{N_1} = e^{-\Delta U / k_B T} \qquad \Delta U = U_2 - U_1$$

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

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

$k_B T$



y=exp(-x)

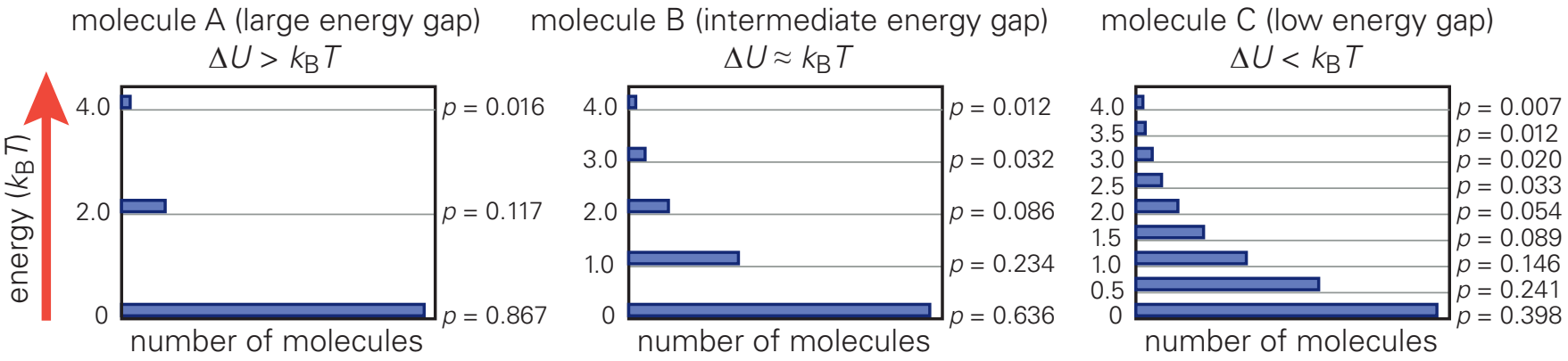
Boltzmann Distribution and Energetic Levels

-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

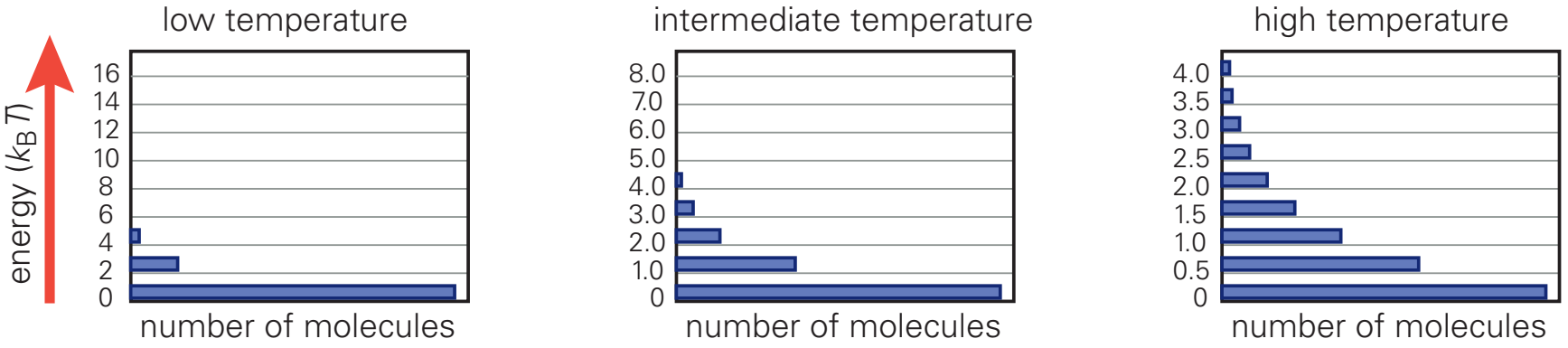
Boltzmann Distribution and Energetic Levels

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

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

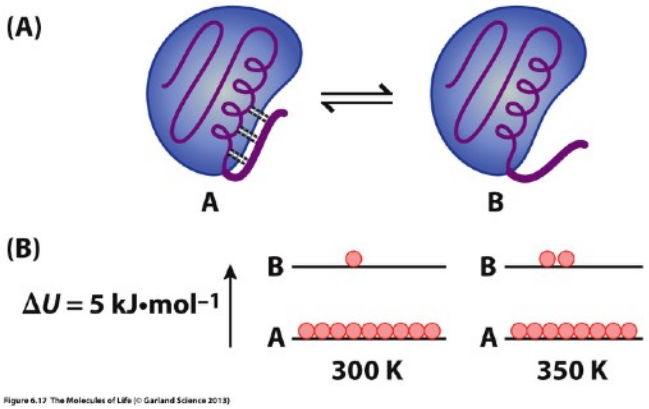
$$\frac{N_2}{N_1} = e^{-\Delta U / 2.529}$$

using as unit kJ/mol



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

Boltzmann Distribution in macromolecules



Protein molecules take up energy as they unfold

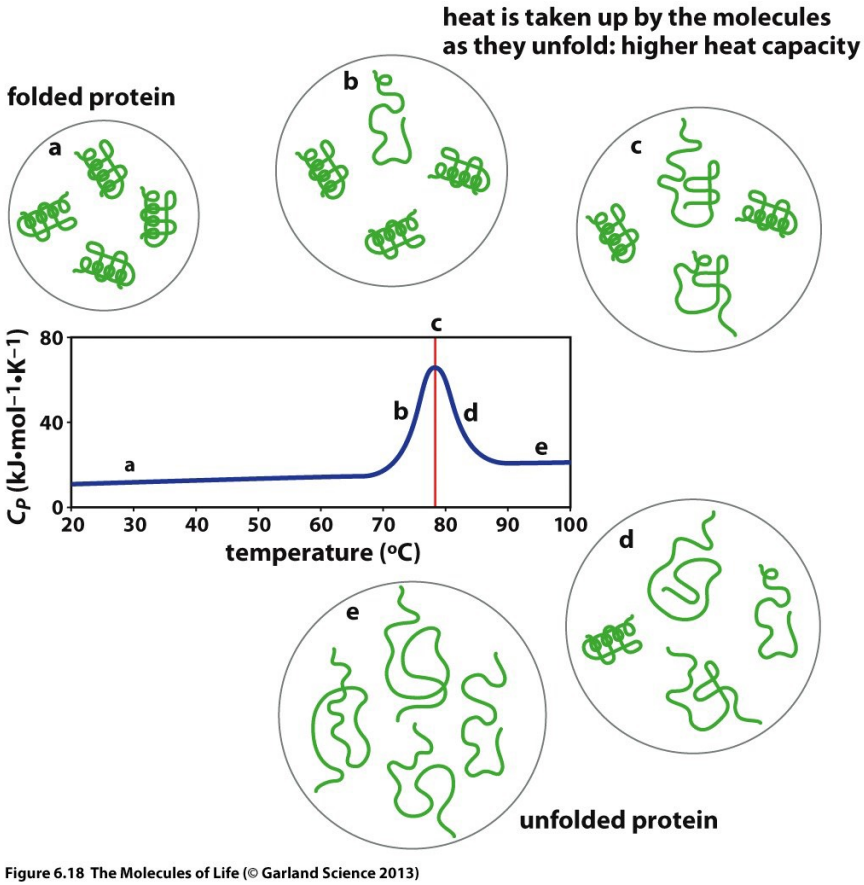


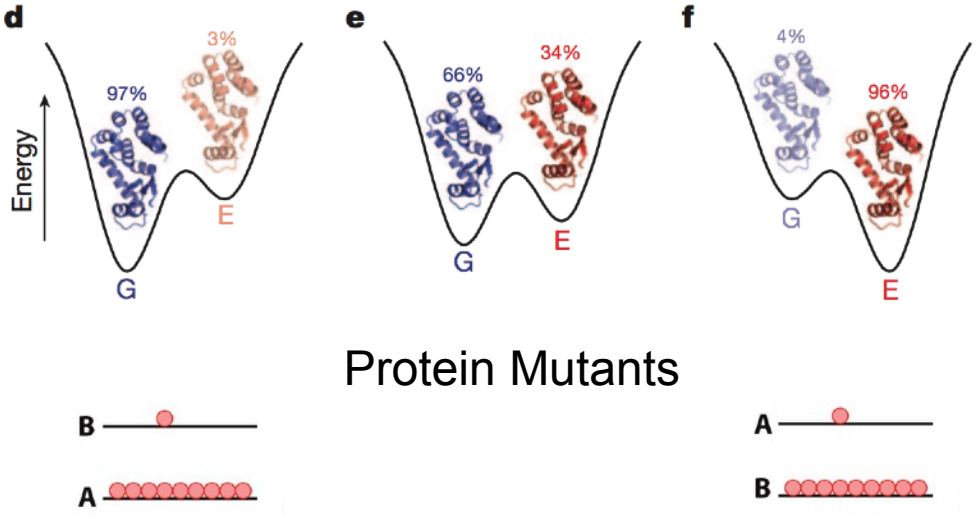
Figure 6.18 The Molecules of Life (© Garland Science 2013)

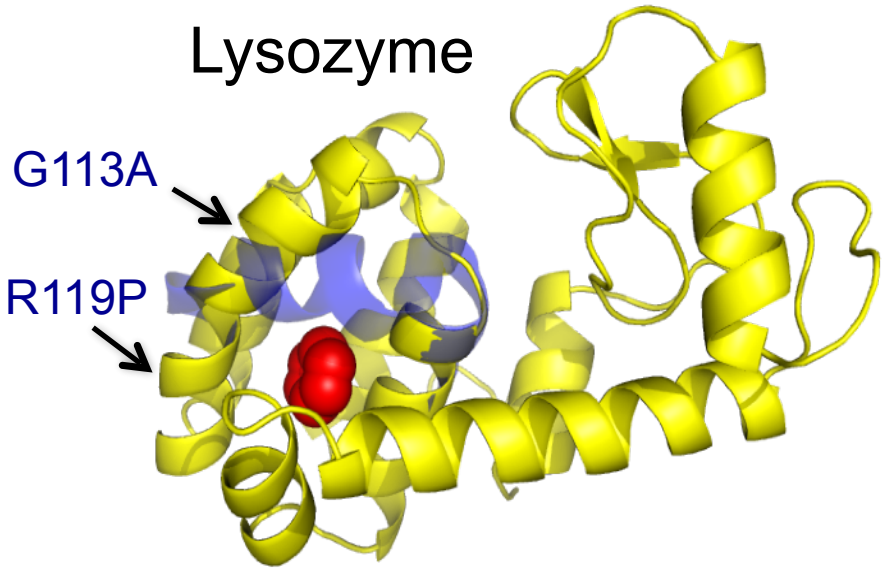
Shifting the distribution of populations with temperature

One can see how the formalism of the Boltzmann distribution helps us to describe 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.