LECTURE 5

BB 101

MODULE: *PHYSICAL BIOLOGY*

Ambarish Kunwar

Lab No. 204
Department of Biosciences and Bioengineering
IIT Bombay

akunwar@iitb.ac.in

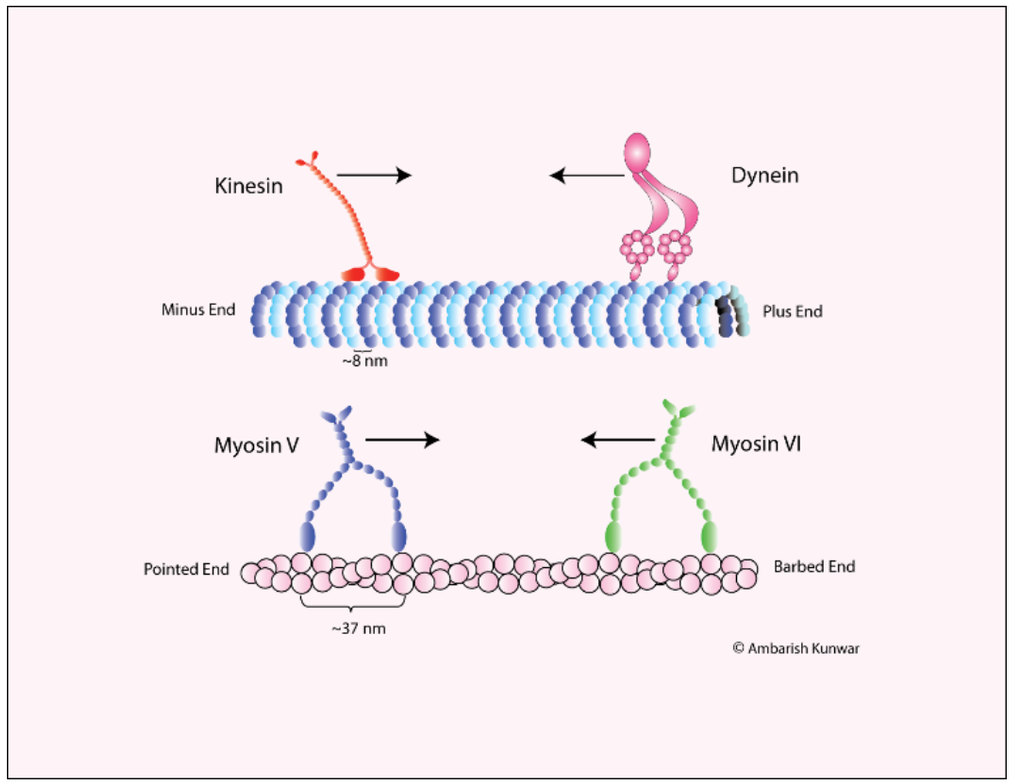
http://www.bio.iitb.ac.in/~akunwar/

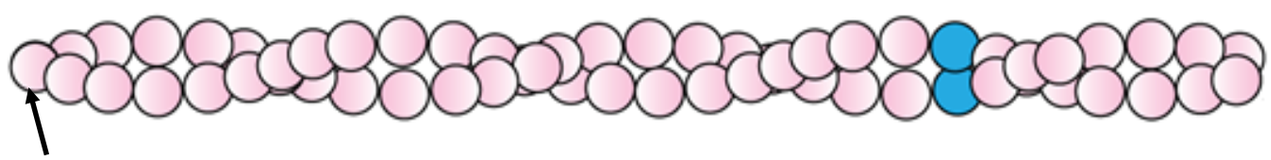
Review of Lecture 4

- Fick's law, Continuity Equation and Diffusion Equation
- Free diffusion from a point source
- Need for active transport
- Einstein Relation and its significance
- Reaction-Diffusion System

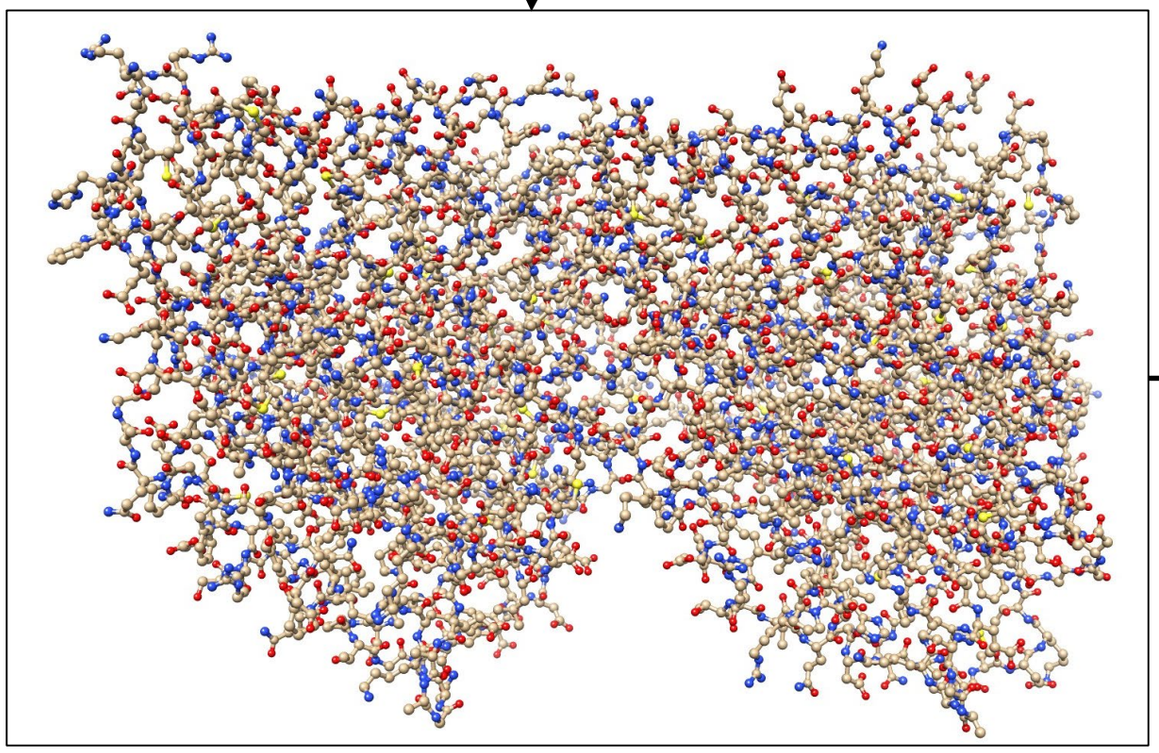
Proteins and their structures

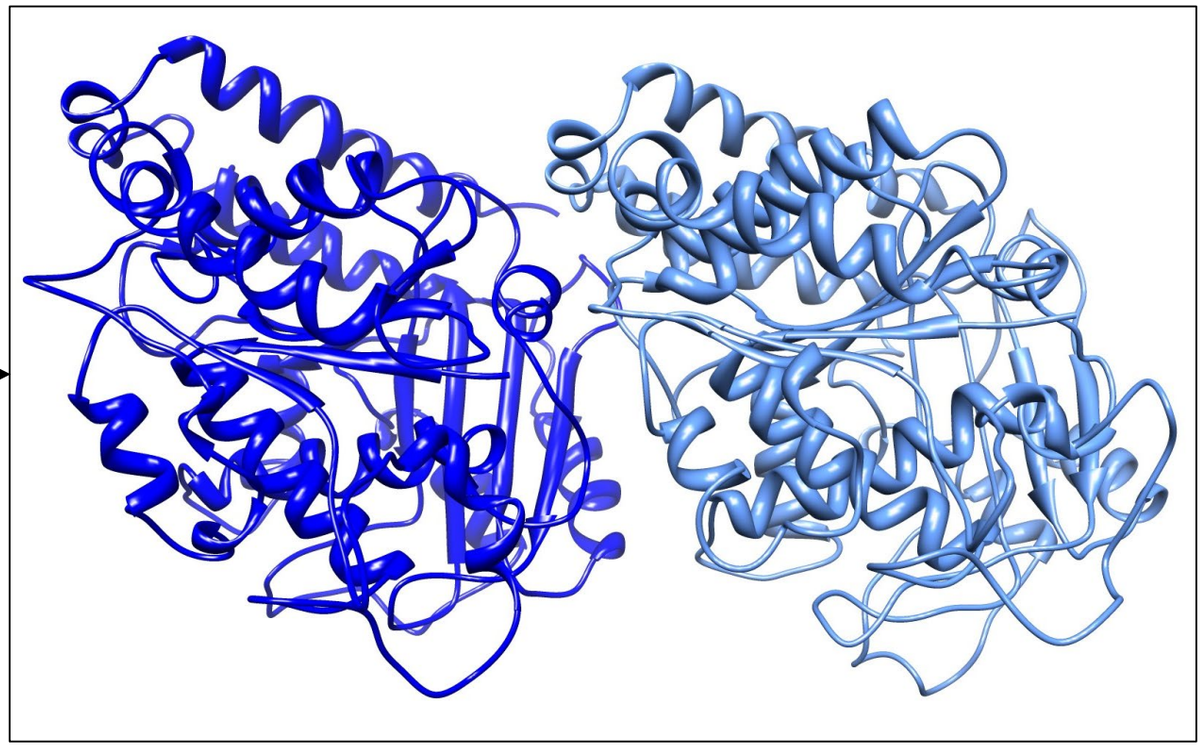
Tubulin and actin monomers are protein molecules. Microtubule is made by tubulin dimer and actin filaments are made from action monomers





Visualizing Protein Structures





α-tubulin Protein Sequence

RECISIHVGQAGVQIGNACWELYCLEHGIQPDGQMPSDKTIGGGDDSFNTFFSETGAGKH
VPRAVFVDLEPTVIDEVRTGTYRQLFHPEQLITGKEDAANNYARGHYTIGKEIIDLVLDR
IRKLADQCTGLQGFSVFHSFGGGTGSGFTSLLMERLSVDYGKKSKLEFSIYPAPQVSTAV
VEPYNSILTTHTTLEHSDCAFMVDNEAIYDICRRNLDIERPTYTNLNRLIGQIVSSITAS
LRFDGALNVDLTEFQTNLVPYPRIHFPLATYAPVISAEKAYHEQLSVAEITNACFEPANQ
MVKCDPRHGKYMACCLLYRGDVVPKDVNAAIATIKTKRTIQFVDWCPTGFKVGINYEPPT
VVPGGDLAKVQRAVCMLSNTTAIAEAWARLDHKFDLMYAKRAFVHWYVGEGMEEGEFSEA
REDMAALEKDYEEVGV

β-tubulin Protein Sequence

REIVHIQAGQCGNQIGAKFWEVISDEHGIDPTGSYHGDSDLQLERINVYYNEAAGNKYVP
RAILVDLEPGTMDSVRSGPFGQIFRPDNFVFGQSGAGNNWAKGHYTEGAELVDSVLDVVR
KESESCDCLQGFQLTHSLGGGTGSGMGTLLISKIREEYPDRIMNTFSVVPSPKVSDTVVE
PYNATLSVHQLVENTDETYCIDNEALYDICFRTLKLTTPTYGDLNHLVSATMSGVTTCLR
FPGQLNADLRKLAVNMVPFPRLHFFMPGFAPLTSRGSQQYRALTVPELTQQMFDAKNMMA
ACDPRHGRYLTVAAVFRGRMSMKEVDEQMLNVQNKNSSYFVEWIPNNVKTAVCDIPPRGL
KMSATFIGNSTAIQELFKRISEQFTAMFRRKAFLHWYTGEGMDEMEFTEAESNMNDLVSE
YQQYQDA

Number of amino acids in α-tubulin = 436

Number of amino acids in β-tubulin = 427

Total number of amino acids in tubulin heterodimer = 863

Total number of atoms in tubulin heterodimer = 8319 atoms

Proteins and their structures

- Proteins perform their function by folding into different shapes

- Proteins are sequence of amino acids

- Given a sequence of amino acids can be predict the structure of the protein?

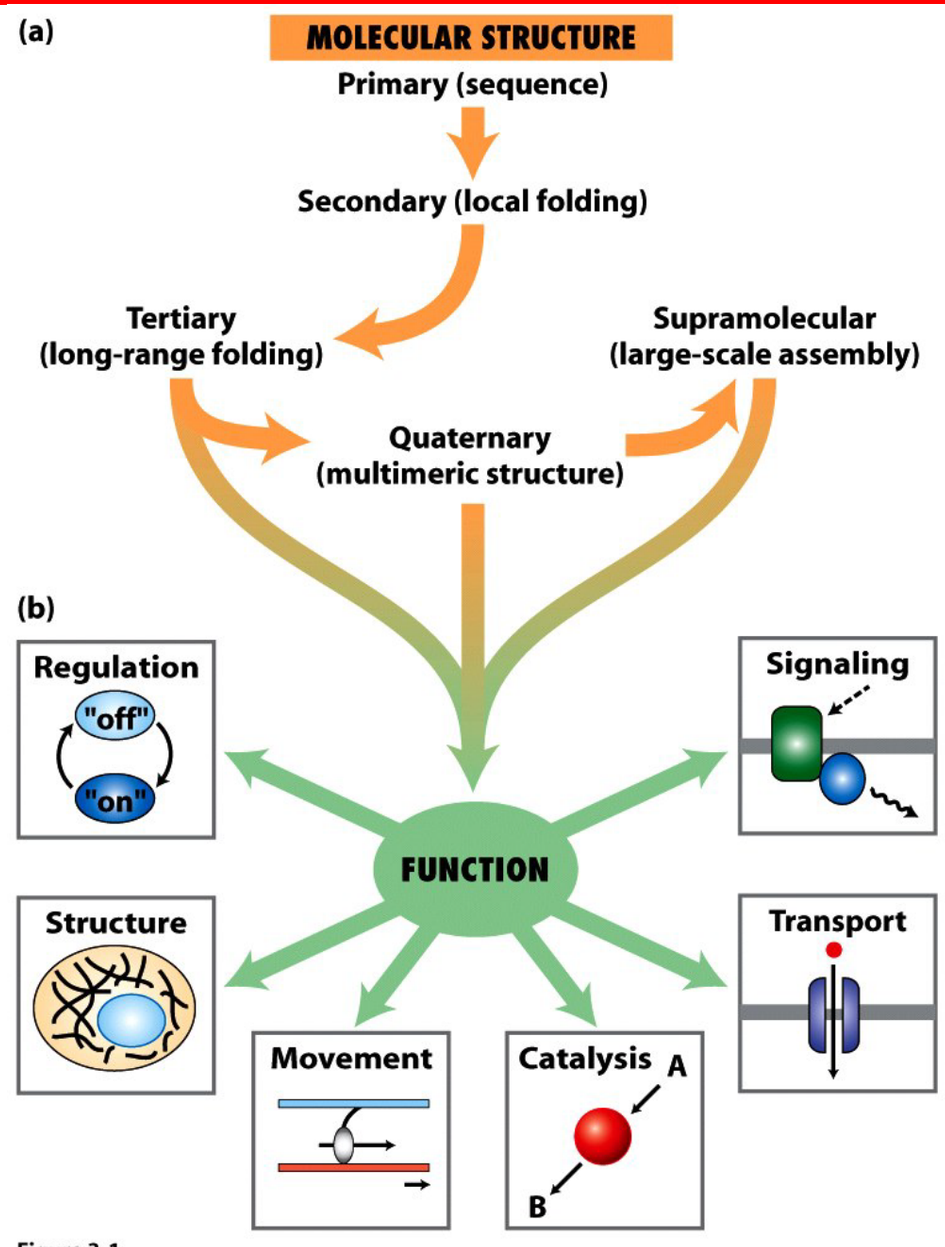


Figure 3-1
Molecular Cell Biology, Sixth Edition

Figure Source: http://www.mastcell-basophil.net/wiki/wiki-start/microtubules-and-mast-cell-signaling/

Figure Source: Otterbein et al., Vol. 293, 708-711 (2001)

Proteins and their structures

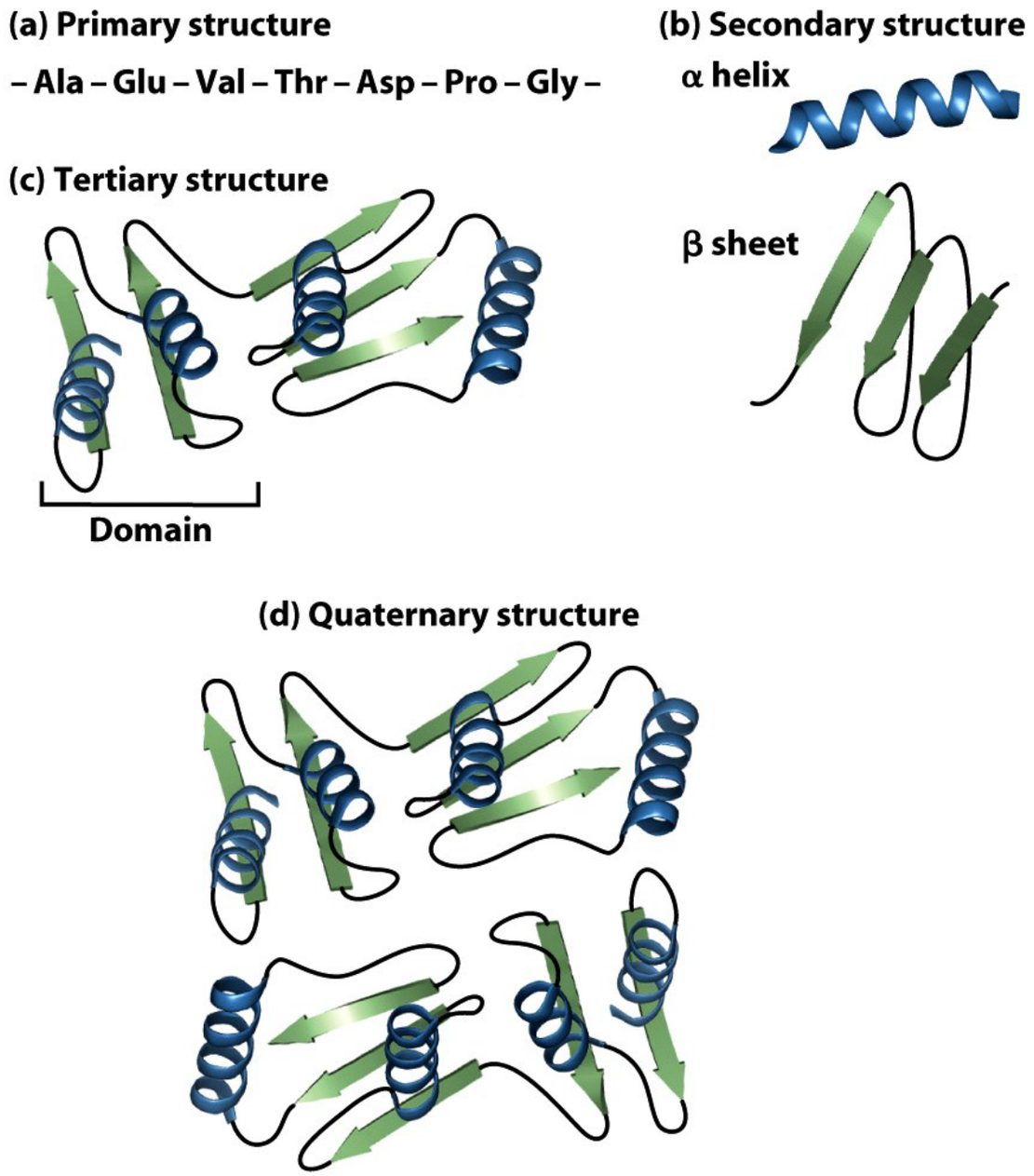


Figure 3-2
Molecular Cell Biology, Sixth Edition

Proteins and their structures

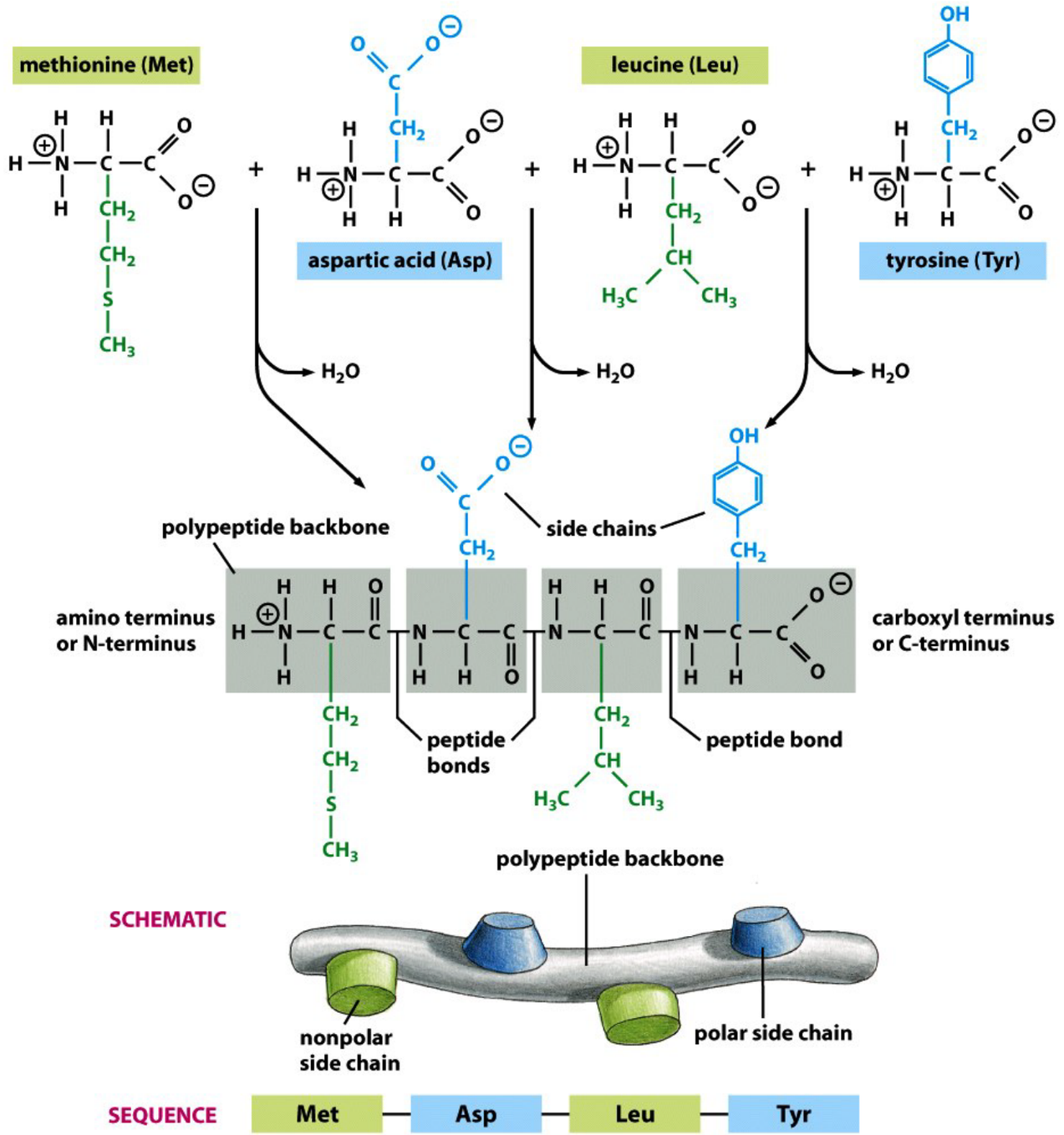


Figure 3-1 *Molecular Biology of the Cell* (© Garland Science 2008)

Proteins and their structures

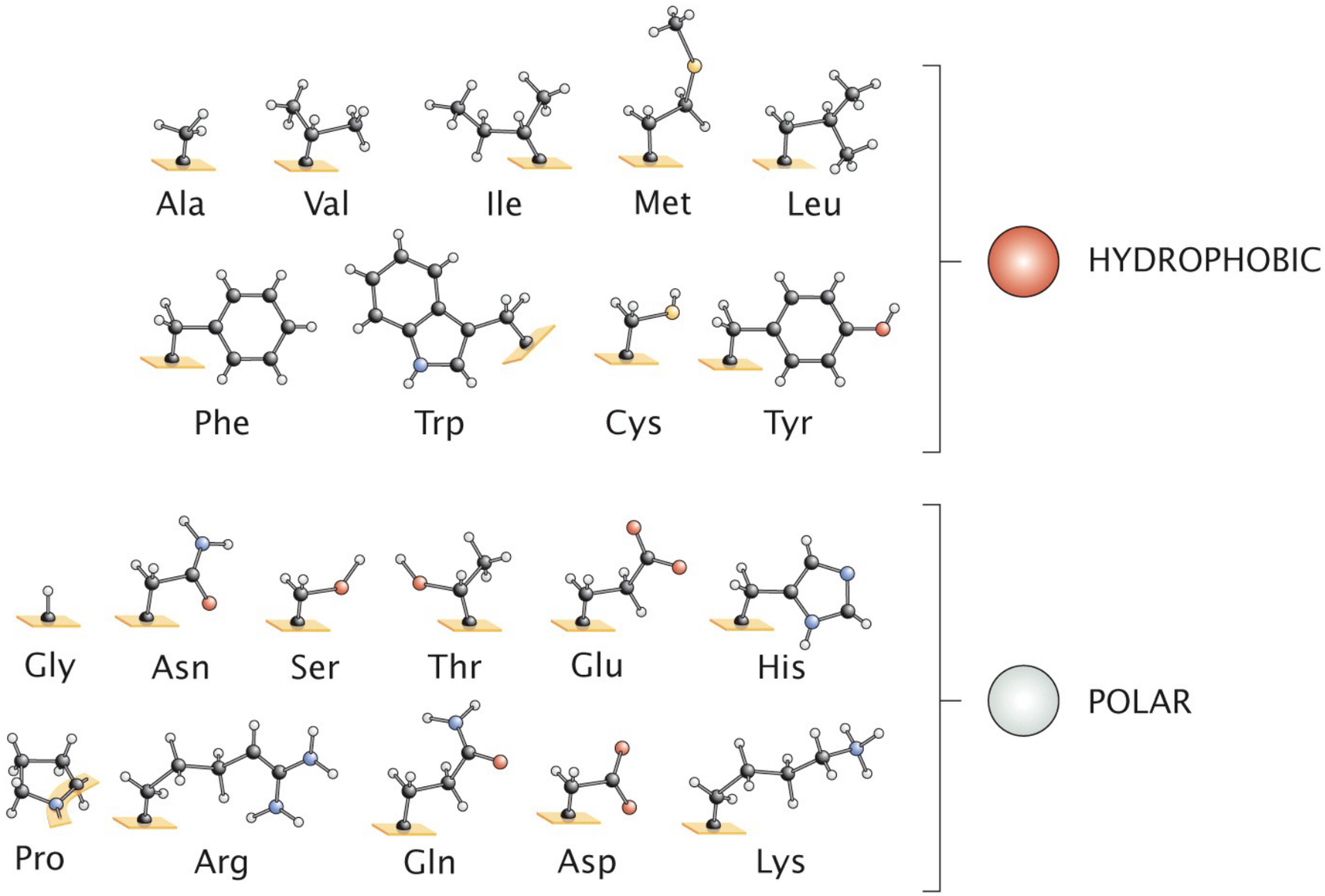


Figure 8.28 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Figure 3-2 *Molecular Biology of the Cell* (© Garland Science 2008)

Protein Structures and Folding

- Proteins are sequence of amino acids

- Given a sequence of amino acids can be predict the structure of the protein?

- Classical mechanics tell us that protein would prefer the structure/conformation which minimizes energy

- However, this is not true when thermal fluctuations are present

Protein Structures are free energy minimizers

- In presence of thermal fluctuations, a protein folds into the structure which minimizes free energy, out of all the possible ways that a particular chain of amino acids can fold up

- What is free energy?

- Helmholtz free energy

$$A = U - TS$$

- Gibbs free energy

$$G = H - TS$$

$$H = U + pV$$

Free Energy

- Free energy=Energy- Temperature x Entropy

- Free energy can be minimized by either decreasing energy or increasing entropy

- In order to calculate free energy we have to calculate entropy

- What is entropy? How to calculate entropy?

Entropy

- Entropy is a measure of the microscopic degeneracy of a macroscopic state ("macro-state"

- In other words, entropy can be computed by counting the number of possible microscopic arrangements/states ("micro-states") for a given macroscopic state ("macro-state")

$$S = k_B \ln W$$

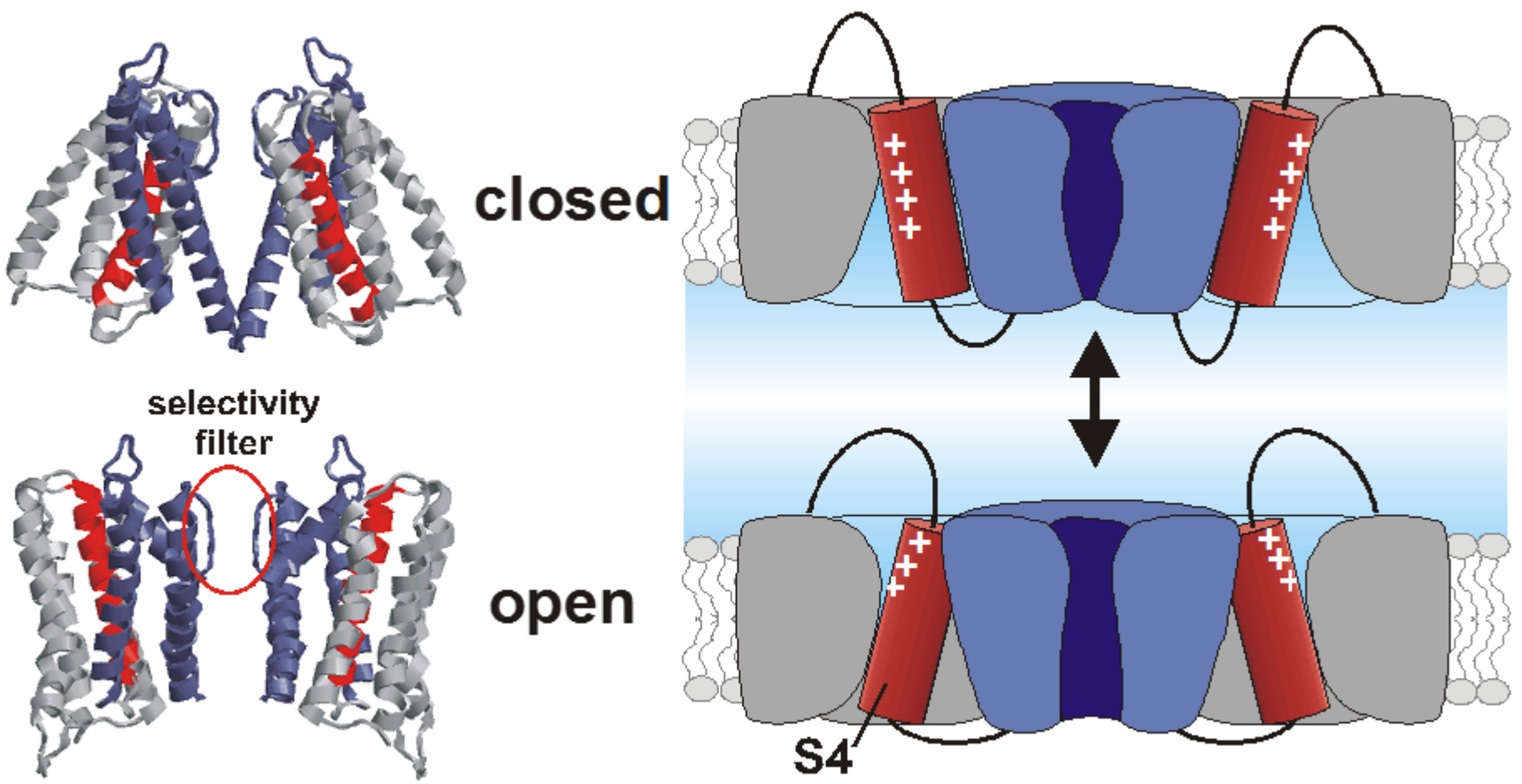
W is the number of possible microscopic arrangements (or "micro-states")

Entropy

Figure Sources: http://en.wikipedia.org/wiki/Boltzmann%27s_entropy_formula#/media/File:Zentralfriedhof_Vienna_-_Boltzmann.JPG

Micro-state and Macro-state

- Consider a protein that can exist in two **structural states.** For example open (1) and closed (2)



http://www.biophys.umontreal.ca/Bluncklab/research.html

Micro-state and Macro-state

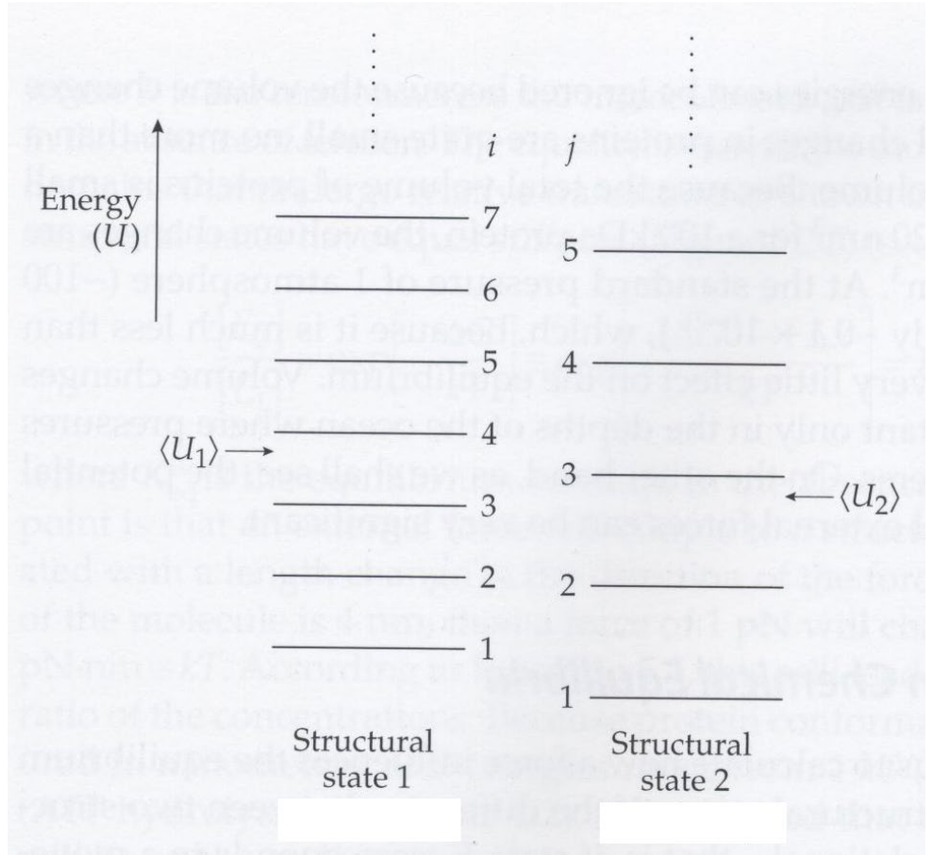
- However, due to thermal fluctuations, these structural states consists of enormous number of **conformations states**, where a conformational state is a set of positions of all the atoms

- Structural states $\Rightarrow$ “Macro-states”

- Conformational states $\Rightarrow$ “Micro-states”

Micro-state and Macro-state

- We know that probability of finding protein any microstate i or j is given by Boltzmann law



$$p_i = \frac{1}{Z} e^{-\frac{U_i}{k_B T}} \qquad p_j = \frac{1}{Z} e^{-\frac{U_j}{k_B T}}$$

- What is the probability of finding protein in a given macrostate?

- The probability of finding protein in a given macrostate is given sum of all p_i or p_j

Micro-state and Macro-state

- Thus, Probability of finding in a macrostate X is given by

$$p(X) = \sum_{iX} p_i = \sum_{iX} \frac{1}{Z} e^{-\frac{U_i}{k_B T}}$$

- Using relation $G = -k_B T ln Z$ between partition function and free energy it can be shown that

$$\boxed{p(X) = \frac{1}{Z} e^{-\frac{G(X)}{k_B T}}}$$

- Similar to Boltzmann formula for microstate, energy replaced by free energy of macrostate

Proof of $G = -k_B T \ln Z$

$$G \equiv < U > -TS$$

$$< U >= \sum_{i=1}^{N} U_i P_i$$

$$S = -k \sum_{i=1}^{N} p_i \ln p_i$$

Shannon's Entropy

$$\Rightarrow G = \sum_{i=1}^{N} U_i P_i + kT \sum_{i=1}^{N} p_i \ln p_i = \sum_{i=1}^{N} P_i \left\{ U_i + kT \ln p_i \right\}$$

$$P_i = \frac{1}{Z} \exp\left[-\frac{U_i}{kT} \right]$$

$$\Rightarrow G = \sum_{i=1}^{N} \frac{1}{Z} \exp\left[-\frac{U_i}{kT} \right] \left\{ U_i + kT \ln\left(\frac{1}{Z} \exp\left[-\frac{U_i}{kT} \right] \right) \right\} = \frac{1}{Z} \sum_{i=1}^{N} \exp\left[-\frac{U_i}{kT} \right] \left\{ U_i - kT \ln Z - U_i \right\}$$

$$Z = \sum_{i=1}^{N} \exp\left[-\frac{U_i}{kT} \right]$$

$$\Rightarrow G = -kT \ln Z \Rightarrow Z = \exp\left(-\frac{G}{kT} \right)$$

So far…

- Proteins can exist in many structural states “macrostates”

- Each “macrostate” consists of many “microstates”

- “macrostate” with minimum free energy is preferred

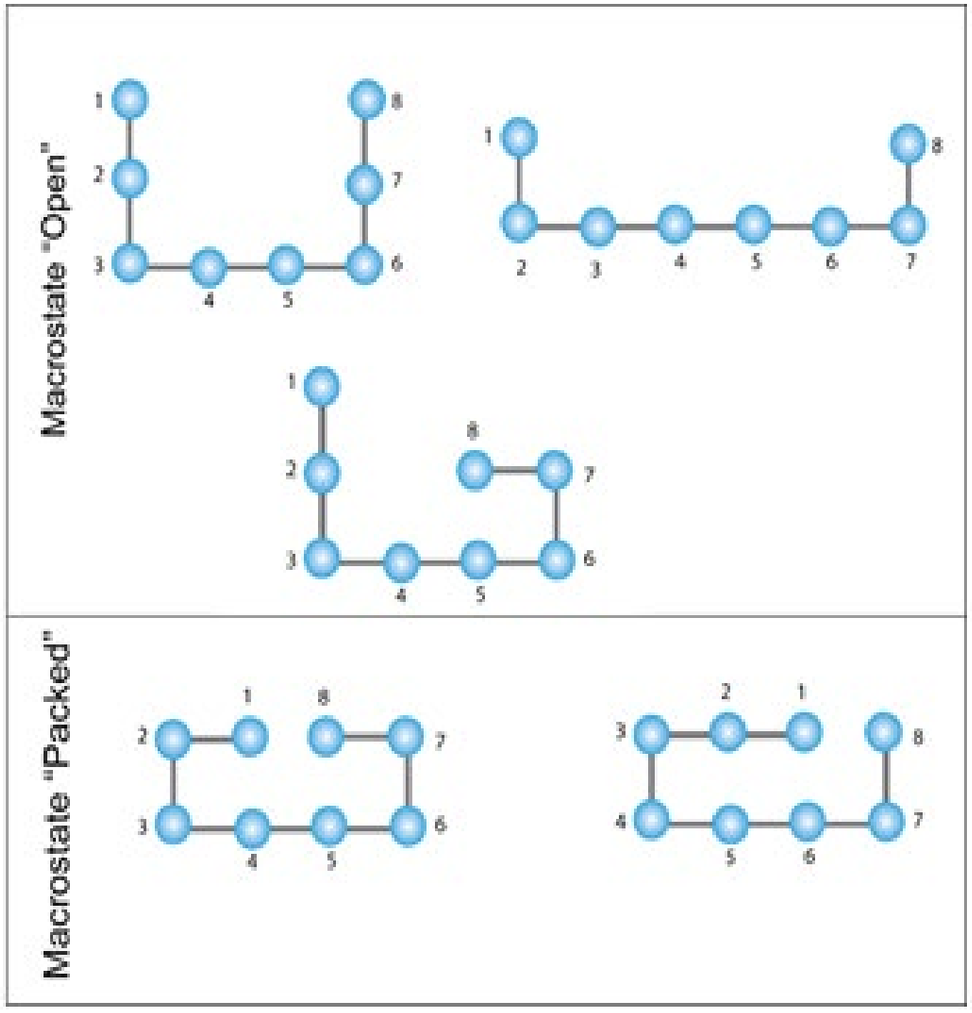
- $G = H - TS$, $G = -k_B T ln Z$ and $\mathrm{S} = k_B \ln W$

- Probability of finding a macrostate X is given by

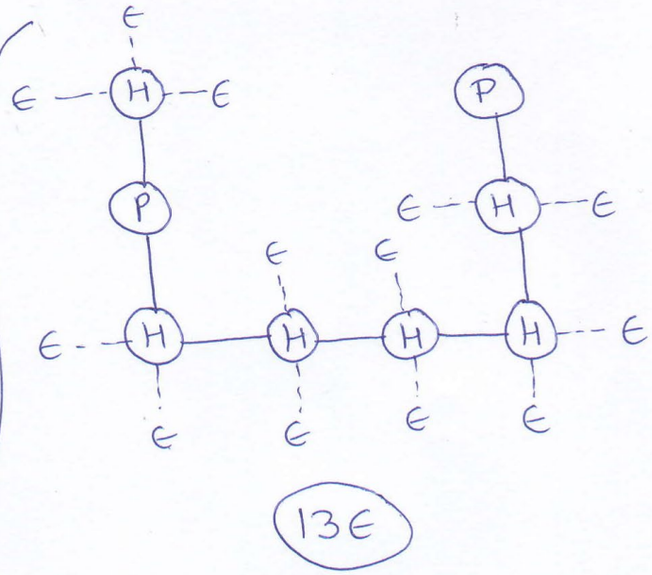
$$p(X) = \frac{1}{Z} e^{-\frac{G(X)}{k_B T}}$$

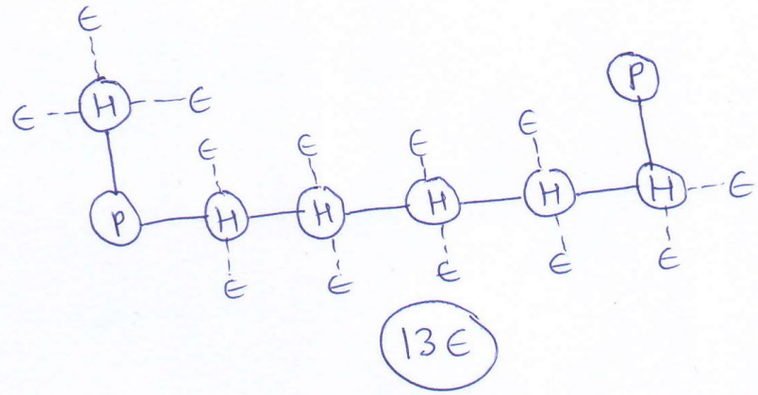
HP models of protein folding

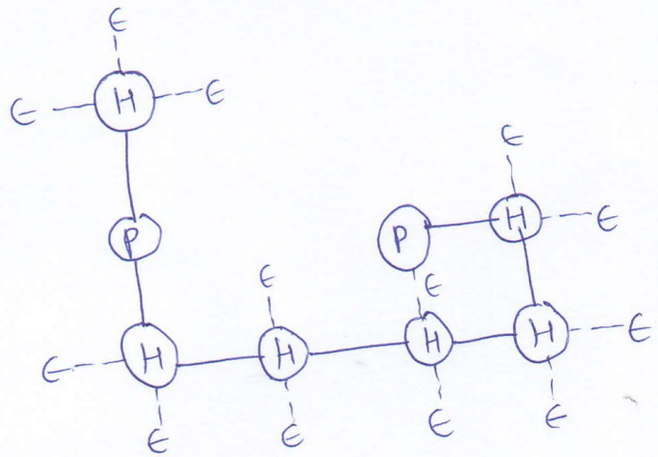
- Consider a protein consisting of eight amino acids consisting of H and P
- The sequence of the amino acids is HPHHHHHP
- Suppose that this protein can fold into five possible structures such that energy increases by ε for every contact of H with either P or solvent molecule
- Suppose these structures can be broadly classified into two macroscopic states “Open” and “Packed”
- What will be the structural state of the protein at temperature T i.e. Open or Packed?



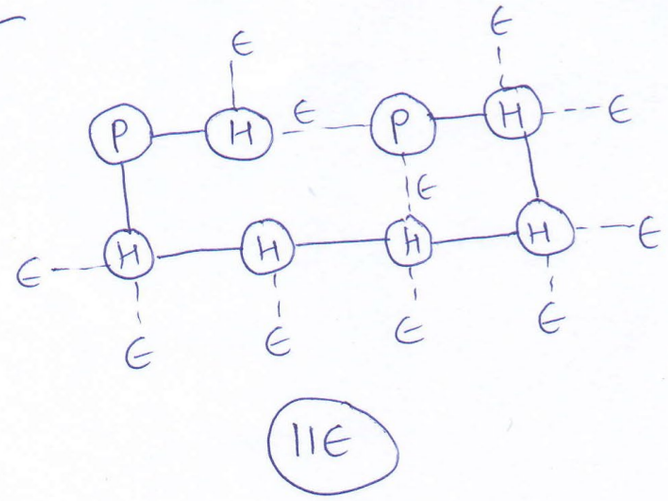
Open

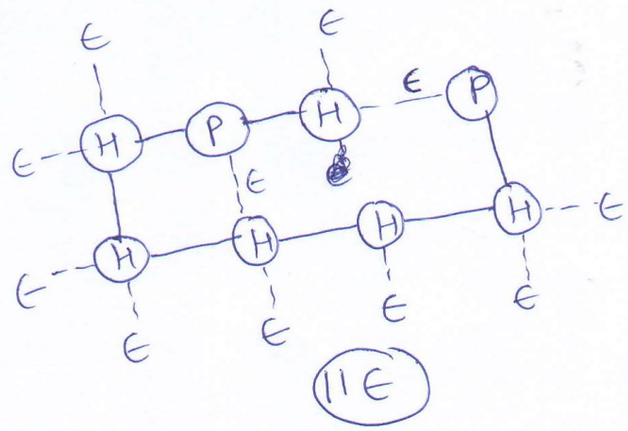






Paired





HP Models of Protein Folding

$$G_{Open} = 13\varepsilon - k_B T \, ln \, 3$$

$$G_{Packed} = 11\varepsilon - k_B T \, ln \, 2$$

Packed will be preferred as long as $G_{Packed} < G_{open}$

Or, $11\varepsilon - k_B T \, ln \, 2 < 13\varepsilon - k_B T \, ln \, 3$

Or, $T < \left(\frac{\varepsilon}{k_B}\right)\left(\frac{2}{\ln(\frac{3}{2})}\right)$

Summary

- Proteins and their structures
- Proteins are free energy minimizers
- Microstate and Macrostate
- Relations $G = H - TS$ and $G = -k_B T ln Z$
- $\mathrm{S} = k_B \ln W$, $S = -k \sum_i p_i ln p_i$ (Shannon's entropy)
- HP model of protein folding