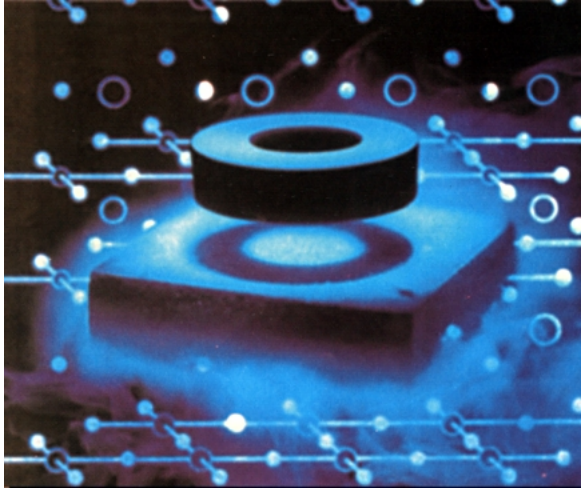
Chemical Physics II

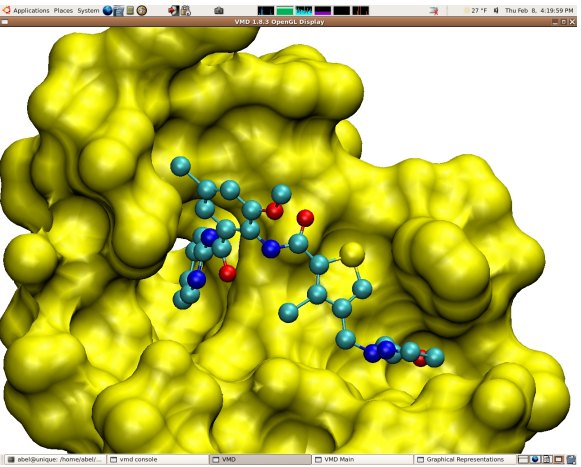


More Stat. Thermo

Kinetics

Protein Folding

...



http://www.nmc.ctc.com/images/project/proj15Thumb.jpg
http://nuclearweaponarchive.org/Usa/Tests/Ukgrable2.jpg
http://www.photolib.noaa.gov/corps/images/big/corp1417.jpg

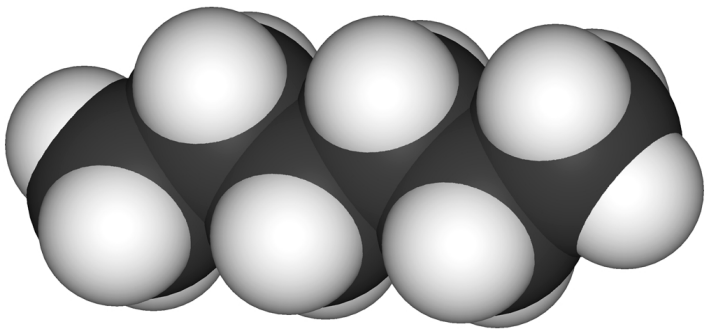
Continuing Theme

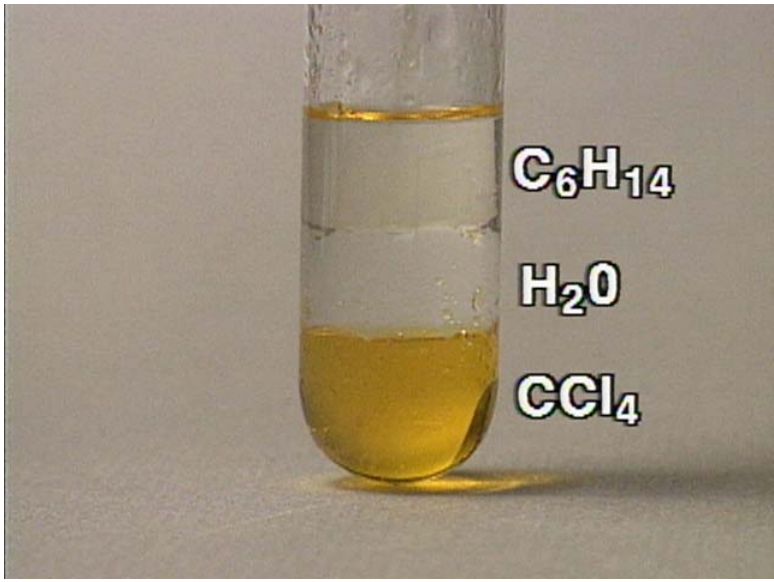
How can we understand the bulk properties of a substance from the properties of individual molecules

$H_3C\text{-}CH_3\text{-}CH_3\text{-}CH_3\text{-}CH_3\text{-}CH_3$

?

Density=0.6548 g/ml
Boiling point=69 °C
Melting point=–95°C
etc.





http://en.wikipedia.org/wiki/Image:Hexane-3D-space-filling.png

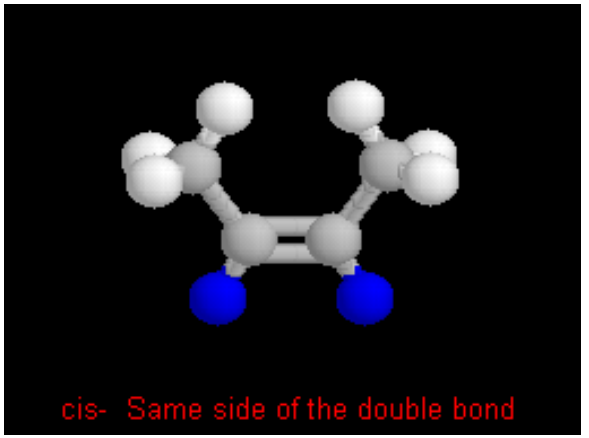
http://jchemed.chem.wisc.edu/JCESoft/CCA/CCA2/ST/LIKEDIS1/64A.JPG/19.JPG

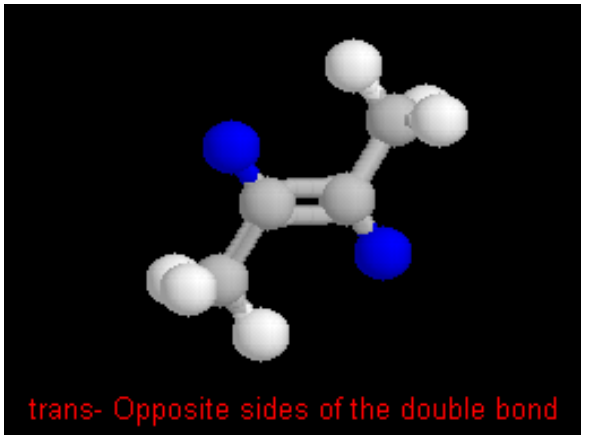
The Game Plan

- Review the two-state system
  - butene cis/trans example
  - activation energy and kinetics
- Brief intermezzo on the free energies
- Generalize the two-state problem to an N-state problem
  - We will see how this relates to the folding of proteins
  - This will give a new perspective on entropy

Butene Gas

- Recall, some molecules of the gas will be in the cis conformation, others will be in the trans conformation:

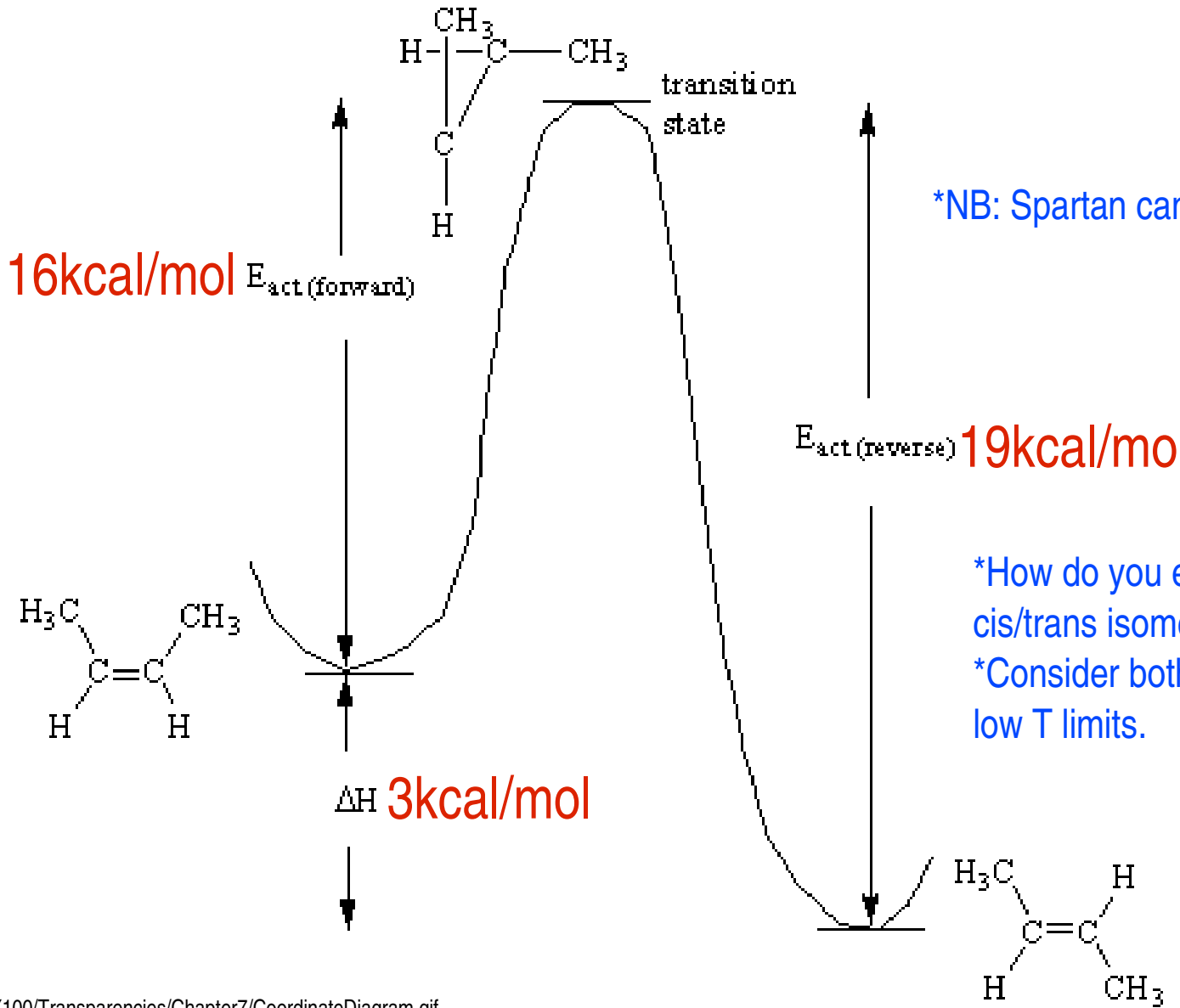




http://www.carlton.srsd119.ca/chemical/equilibrium/images/

Butene Gas

- The butene **potential energy surface**



*NB: Spartan can compute these

*How do you expect the ratio of cis/trans isomers to behave?
*Consider both the high T and low T limits.

http://usm.maine.edu/~newton/CHY100/Transparencies/Chapter7/CoordinateDiagram.gif

Butene Gas cis/trans Probability

Three key properties of our probability function :

1. It should be a function of energy and temperature

$$P_c = P_c(E_c, T) \quad \text{and} \quad P_t = P_t(E_t, T)$$

2. Independent events should multiply , ie for any two butene molecules

$$P_{c1,c2} = P_c \cdot P_c \;\; ; \;\; P_{c1,t2} = P_{t1,c2} = P_c \cdot P_t \;\; ; \;\; P_{t1,t2} = P_t \cdot P_t$$

3. Probabilities should be normalized to 1

$$P_c + P_t = 1$$

Three key properties of our probability function:

1. It should be a function of energy and temperature

$$P_c \propto e^{-E_c/RT} \quad \text{and} \quad P_t \propto e^{-E_t/RT}$$

Boltzmann Factor

2. Independent events should multiply

$$P_{c1,t2} = P_c \cdot P_t \propto e^{-E_c/RT} \cdot e^{-E_t/RT} = e^{-(E_c+E_t)/RT}$$

Probabilities should be normalized to 1

$$P_c + P_t = 1 = \frac{e^{-E_c/RT}}{Q} + \frac{e^{-E_t/RT}}{Q} = \frac{e^{-E_c/RT} + e^{-E_t/RT}}{Q}$$

Partition Function

$$\text{where} \quad Q = e^{-E_c/RT} + e^{-E_t/RT}$$

Butene Gas cis/trans Probability

So our final result was,

$$P_c = \frac{e^{-E_c/RT}}{Q} = \frac{e^{-E_c/RT}}{e^{-E_c/RT} + e^{-E_t/RT}}$$

$$P_t = \frac{e^{-E_t/RT}}{Q} = \frac{e^{-E_t/RT}}{e^{-E_c/RT} + e^{-E_t/RT}}$$

Butene Gas Isomerization Kinetics

- We estimated the rate at which the a jug of all cis butene will isomerize to trans butene by finding
  - what fraction of the molecules at a given time are at the transition state
  - this is the same calculation we just did, but now ΔE=16kcal/mol instead of 3kcal/mol

$$k_{cis\rightarrow trans} \propto P(tst)\, s^{-1} = \frac{e^{\frac{-\Delta E_{c\,tst}}{RT}}}{1+e^{\frac{-\Delta E_{c\,tst}}{RT}}} s^{-1} = \frac{e^{\frac{-16\text{kcal/mol}}{(.002\,\text{kcal/mol}\cdot\text{K})300\text{K}}}}{1+e^{\frac{-16\text{kcal/mol}}{(.002\,\text{kcal/mol}\cdot\text{K})300\text{K}}}} s^{-1}$$

$$\Rightarrow k_{cis\rightarrow trans} \propto 10^{-12}/\text{s} \quad \text{ie} \quad 30000\text{yr halflife}$$

Free Energy Intermezzo

From last time,

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

$$\Delta H = \Delta E + P\Delta V, \quad \text{but} \quad P\Delta V \approx 0 \quad \text{for bio.}$$

$$\Rightarrow \quad \Delta H \approx \Delta E$$

$$\Rightarrow \Delta G \approx \Delta E - T\Delta S \equiv \Delta A$$

Gibbs Free Energy

Helmholtz Free Energy
(More important too us than it might look at first glance.)

Free Energy Intermezzo

- We have now defined,
  - The partition function: Q
  - The Helmholtz free energy: ΔA
- If we can relate ΔA and Q, then
  - can express the free energy in terms of Boltzmann factors
  - can derive other thermodynamic parameters such as pressure, entropy, heat capacity, isothermal compressibility, etc

Free Energy Intermezzo

It can be shown ,

$$\Delta A = -RT \ln(Q)$$

$$Pr = RT \frac{\partial \ln Q}{\partial V}$$

$$\langle E \rangle = RT^2 \frac{\partial \ln Q}{\partial T}$$

$$S = \frac{\partial}{\partial T}(RT \ln Q)$$

etc...

Generalizing the 2-State Solution

- We have now seen that we can think of our butene molecule rattling between three different states
  - The trans state (lowest in energy, most stable)
  - The cis state (higher in energy, but still long lived)
  - The **transition state** (very high in energy, short lived)
- We should be able to derive the the probability of a butene molecule occupying any of these three states

Generalizing the 2-State Solution

For 3 states we find ,

$$P_t = \frac{e^{-E_t/RT}}{Q} = \frac{e^{-E_t/RT}}{e^{-E_c/RT} + e^{-E_t/RT} + e^{-E_{tst}/RT}} = .99_{298K} \quad ; \quad .82_{1000K}$$

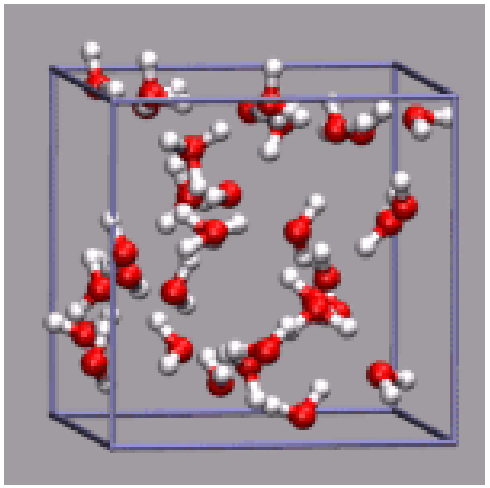
$$P_c = \frac{e^{-E_c/RT}}{Q} = \frac{e^{-E_c/RT}}{e^{-E_c/RT} + e^{-E_t/RT} + e^{-E_{tst}/RT}} = .01_{298K} \quad ; \quad .18_{1000K}$$

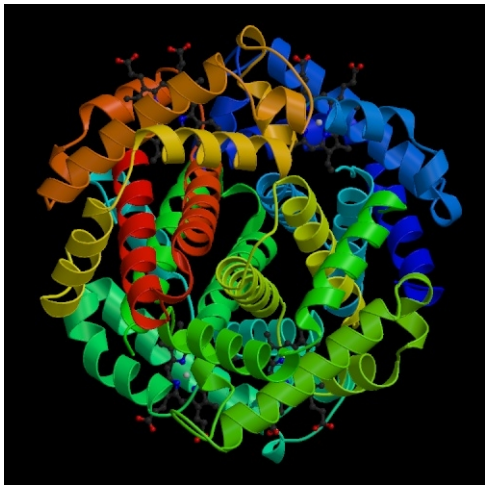
$$P_{tst} = \frac{e^{-E_{tst}/RT}}{Q} = \frac{e^{-E_{tst}/RT}}{e^{-E_c/RT} + e^{-E_t/RT} + e^{-E_{tst}/RT}} = 10^{-12}{}_{298K} \quad ; \quad 10^{-4}{}_{1000K}$$

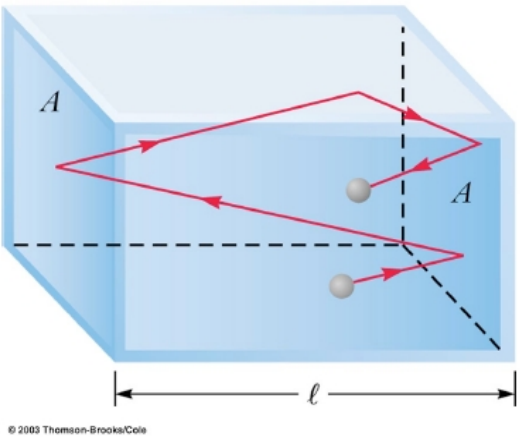
where $E_t = 0 kcal/mol$; $E_c = 3 kcal/mol$; $E_{tst} = 16 kcal/mol$

Many State Problems

- Most “real world” chemical problems have very many states:
  - The number of ways liquid water molecules can arrange themselves in a cup
  - The number different shapes a protein can trace
  - The number of different positions and velocities gas particles can assume in a box





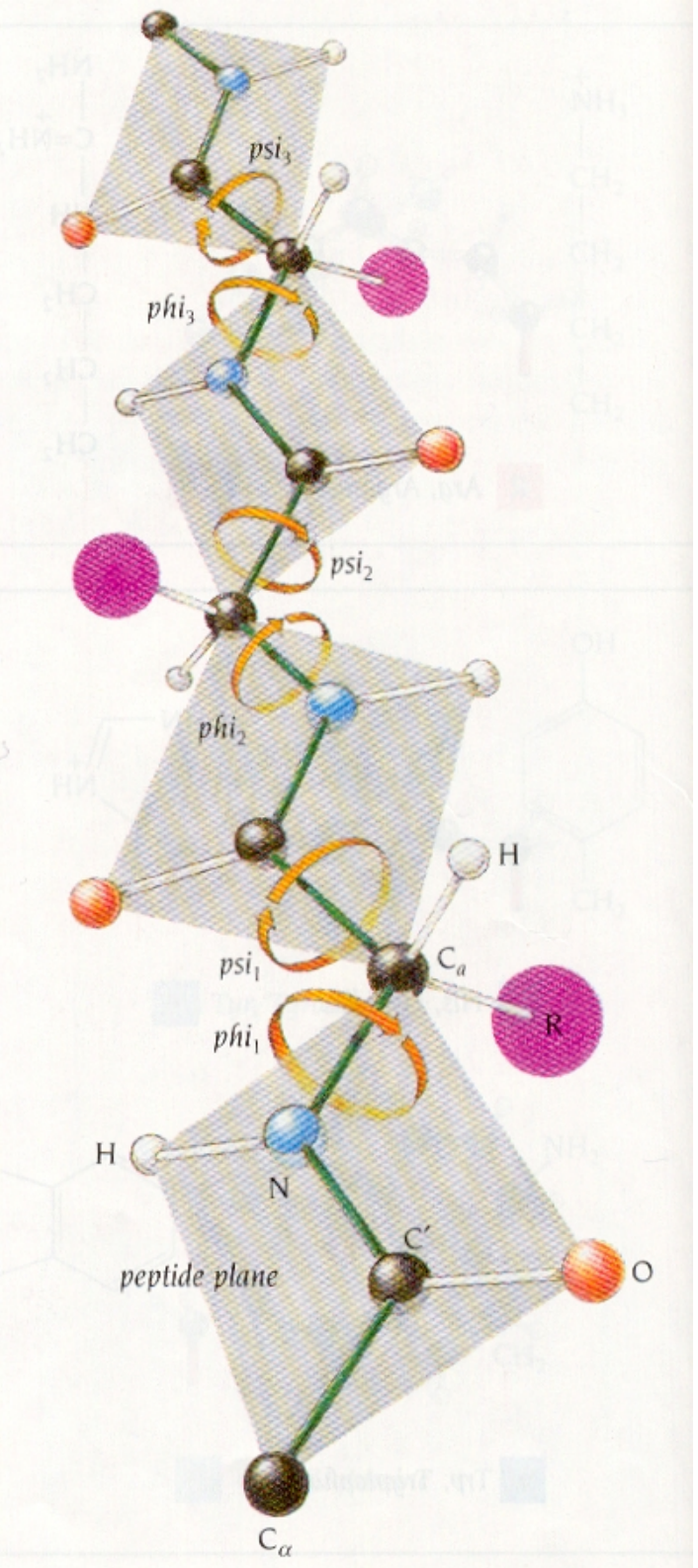


http://www.theochem.ruhr-uni-bochum.de/~axel.kohlmeyer/cpmd-vmd/unitcell-movie.gif

http://iwrwww1.fzk.de/biostruct/Assets/1a00x500.jpg

Protein Folding

- Proteins are polymers
  - think of this as a string of beads
  - each bead is an amino acid
  - each amino acid has two rotatable bonds (phi, psi) connecting it to each other amino acid
  - the string changes shape by rotating around these bonds
  - When in water, each protein will **spontaneously fold into a specific, well defined, shape**



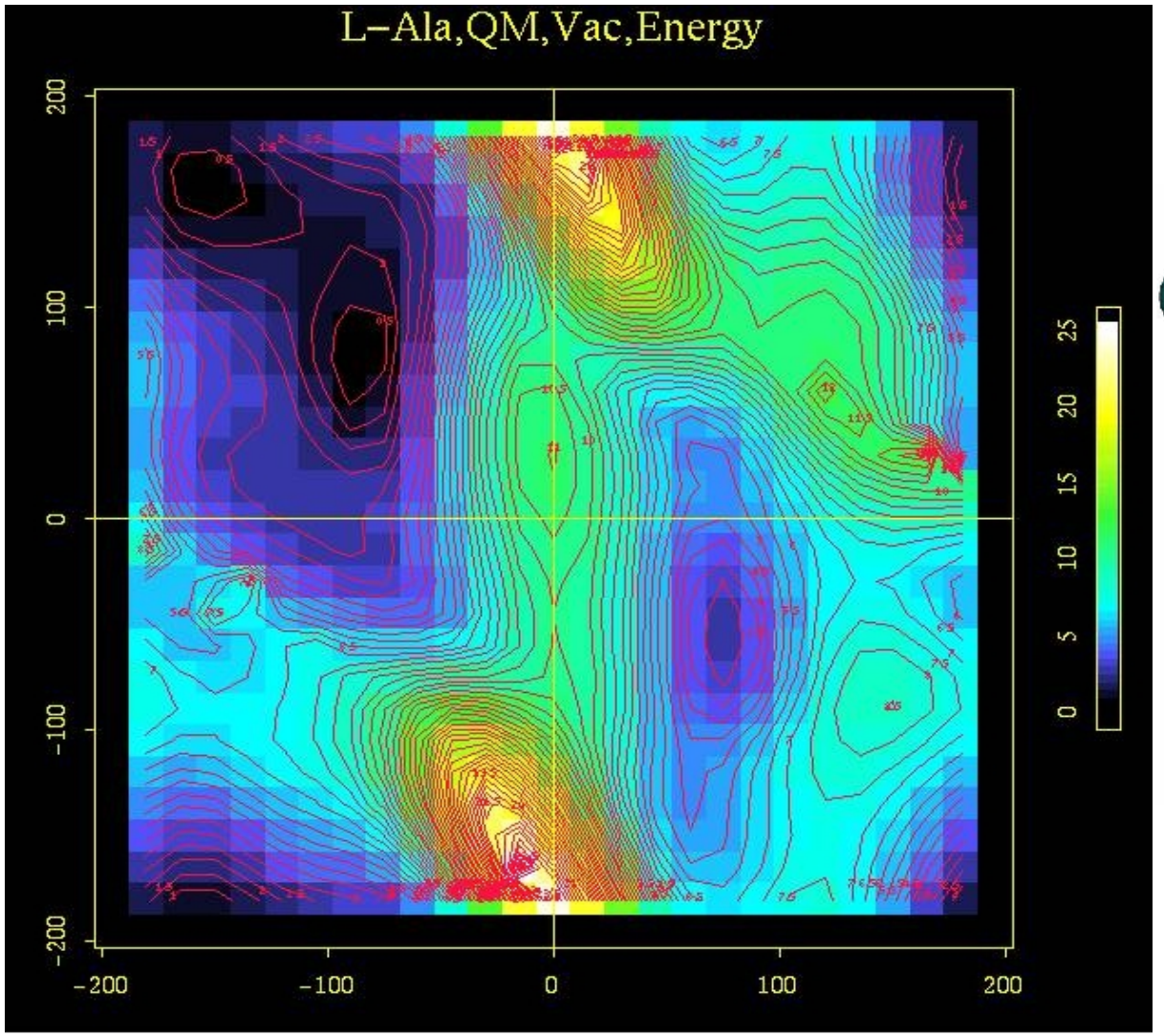
http://tigger.uic.edu/classes/phys/phys461/phys450/ANJUM05/psi_phi_angles.jpg

Protein Folding

- Most proteins contain 100-200 amino acids
- Each amino acid can assume roughly ten different pairs of phi and psi angles
- So the total **conformational space** (ie all the shapes it can fold into) is 10^{100} states
- **Our goal is to understand the thermodynamics of protein folding**

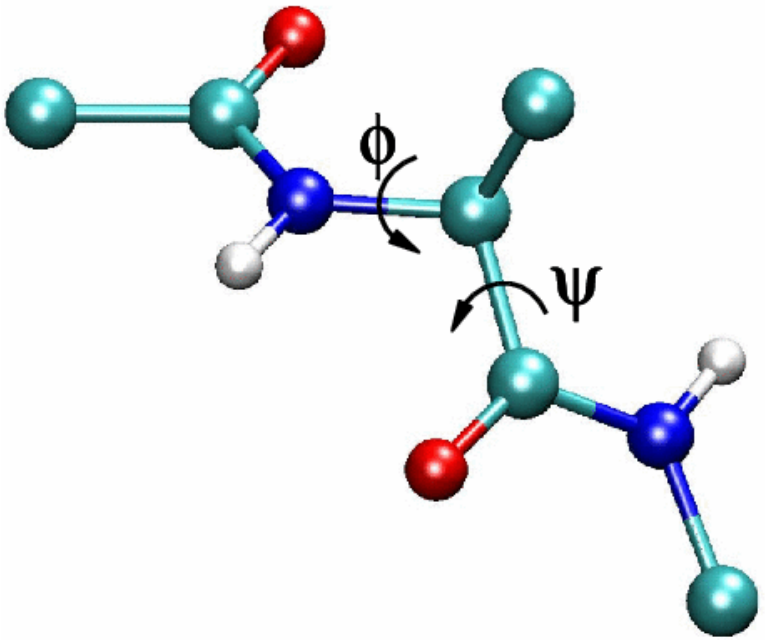
Protein Folding

- Visualization of the ~10 stable states for each amino acid



ψ

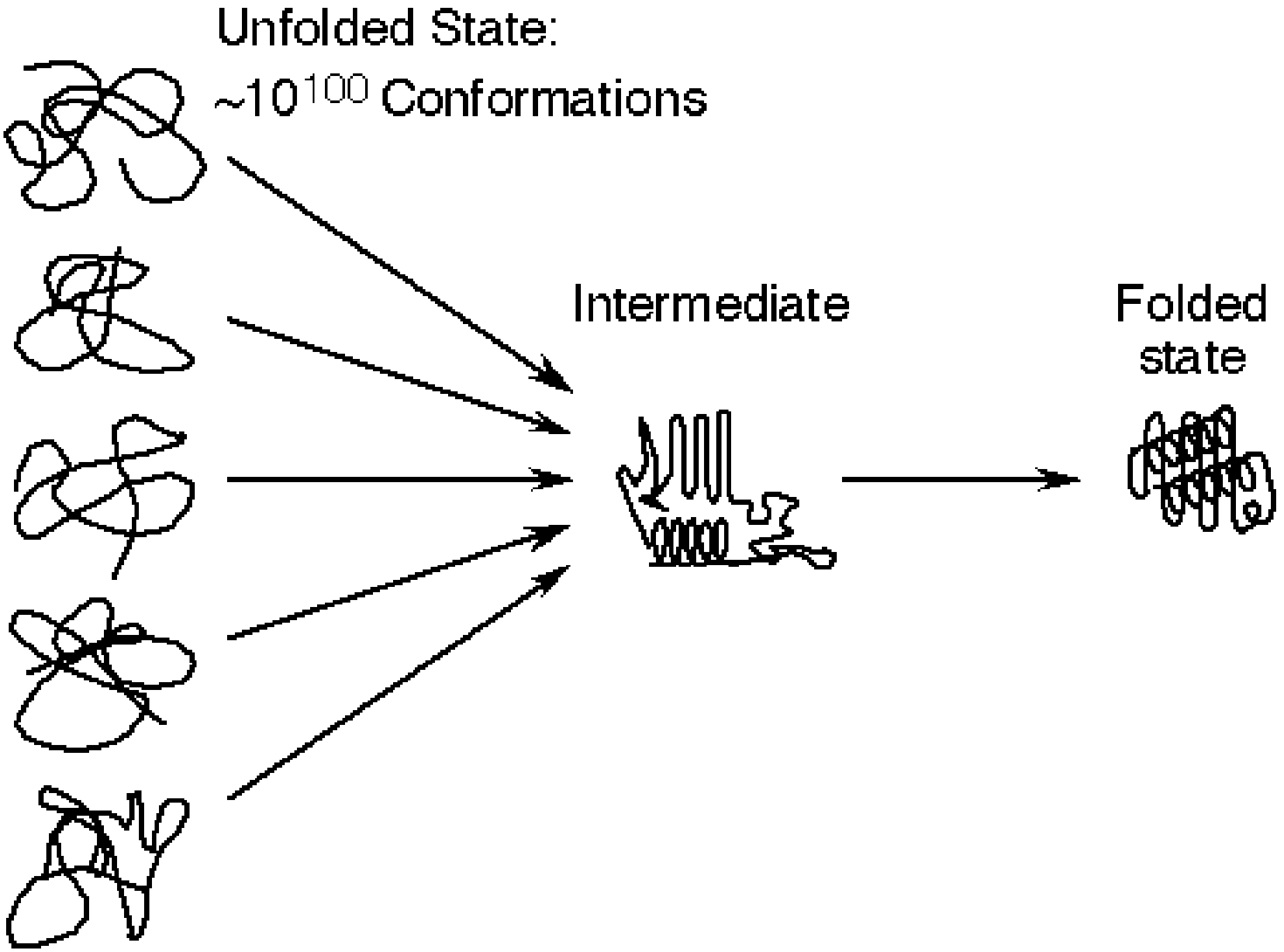
ϕ



http://helix.nih.gov/research/janakim_jan2005.jpg

http://www.math.princeton.edu/~weiqing/alanine/ala2.gif

Protein Folding



http://www.pitb.de/nolting/prot00/6.gif

Protein Folding

- The single folded state of a given protein is typically quite complicated
- It must be in this fold to function
- Prion diseases are caused by unfolded proteins

Protein Folding

- This is a many state problem
  - There exists 1 single folded state with energy E_F
  - There exist ~10^{100} unfolded states with energy E_u
- We will treat each of the unfolded states as though they have the same energy
- We will see
  - the large number of unfolded states opposes folding **(entropy)**
  - the favorable interaction energy of the folded state promotes folding **(enthalpy)**

Protein Folding

We find for protein being folded or unfolded,

$$Q = e^{-E_F/RT} + 10^{100}\, e^{-E_U/RT} = e^{-E_F/RT} + \Omega\, e^{-E_U/RT}$$

$$P_F = \frac{e^{-E_F/RT}}{Q} = \frac{e^{-E_F/RT}}{e^{-E_F/RT} + 10^{100}\, e^{-E_U/RT}}$$

"Degeneracy"=10^{100}

$$P_U = 10^{100}\, \frac{e^{-E_U/RT}}{Q} = \frac{e^{-E_U/RT}}{e^{-E_F/RT} + 10^{100}\, e^{-E_U/RT}}$$

What can we learn about protein folding from this solution?

Protein Folding

- Experimentally,
  - Most proteins are nearly entirely folded at 25°C
  - Most proteins are half unfolded at 75°C
  - Most proteins are nearly entirely unfolded at 90°C
- Notice that this process is much more temperature dependent than the butene isomerization
- For the rest of the discussion we will choose our energy scale such that $E_F=0$ and $E_U=\Delta E_U$

Protein Folding

We can calculate ΔE_U by forcing $P_F = .5 \quad @\ 350K$,

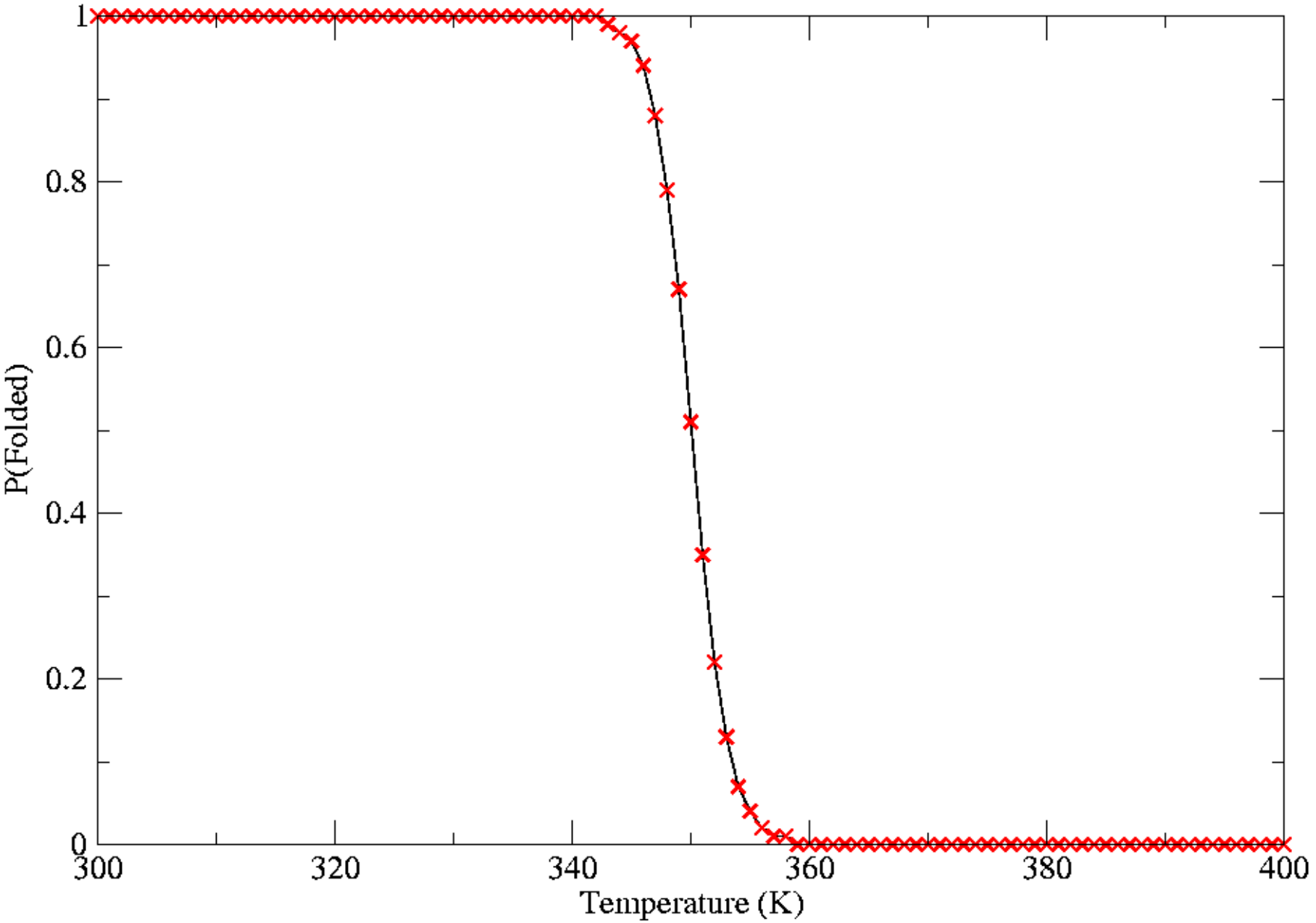
$$P_F = \frac{e^{-E_F/RT}}{e^{-E_F/RT} + 10^{100}\, e^{-E_U/RT}} = \frac{1}{1 + 10^{100}\, e^{-\Delta E_U/R\cdot 350K}} \equiv .5$$

$$\Rightarrow \Delta E_U = 161.2 \text{kcal/mol}$$

What does this say about the occupancy of the folded state at different temperatures ?

$$P_{F\ @\ 300K} = \frac{1}{1 + 10^{100}\, e^{\frac{-161.2\text{kcal/mol}}{R\cdot 300K}}} \approx 1 \quad \text{and} \quad P_{F\ @\ 390K} \approx 10^{-11}$$

Probability a generic protein is folded as a function of temperature



Protein Folding

We can calculate Δ A of unfolding from this derived data ,

$$\Delta A_{F \to U} = -RT \ln \frac{Q_U}{Q_F} = -RT \ln \left[\frac{\Omega \cdot e^{-\Delta E_U / RT}}{1} \right]$$

$$\Rightarrow \Delta A_{F \to U} = -RT \ln e^{-\Delta E_U / RT} - RT \ln \Omega$$

$$\Rightarrow \Delta A_{F \to U} = \Delta E_U - RT \ln \Omega$$

but we have already showed ,

$$\Delta A_{F \to U} = \Delta E_U - T \cdot \Delta S \quad \Rightarrow \quad \Delta S_U = R \ln \Omega$$

Protein Folding

The value of Δ A at 300K is also interesting ,

$$\Delta A_{F \to U} = -RT \ln \left[\Omega \cdot e^{-\Delta E_U / RT} \right]$$

$$\Rightarrow \Delta A_{F \to U} = \Delta E_U - RT \ln \Omega$$

$$\Rightarrow \Delta A_{F \to U} = 161.2 \text{kcal/mol} - 138.2 \text{kcal/mol} = 23 \text{kcal/mol}$$

For perspective , a water – water hydrogen bond is 5kcal / mol

Conclusions

- Proteins have a relatively small and abrupt window of thermal stability
  - this is because both the entropic and enthalpic terms are very large and opposed
- These methods of calculating ΔE, ΔA,and ΔS are at the core of computational chemistry
- We typically get ΔE from a model
- Once we have ΔE for each state we can calculate bulk thermodynamic properties