Proteins are not rigid structures:

Protein dynamics, conformational variability, and thermodynamic stability

Dr. Andrew Lee

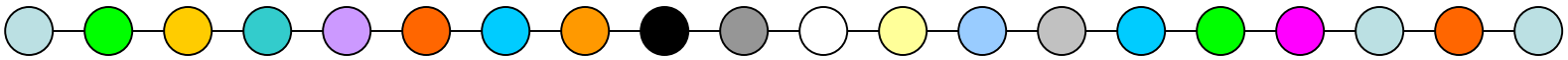
UNC School of Pharmacy (Div. Chemical Biology and Medicinal Chemistry)

UNC Med School (Biochemistry and Biophysics)

drewlee@unc.edu

4109 Marsico Hall

Proteins are long polypeptide chains of ~50 or more "residues".

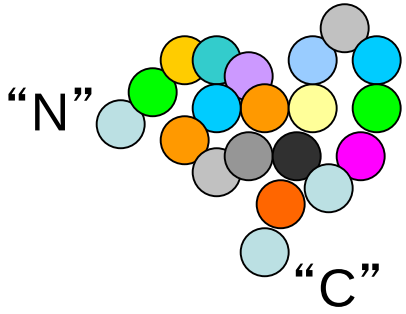


amino terminus

"N"

carboxy terminus

"C"



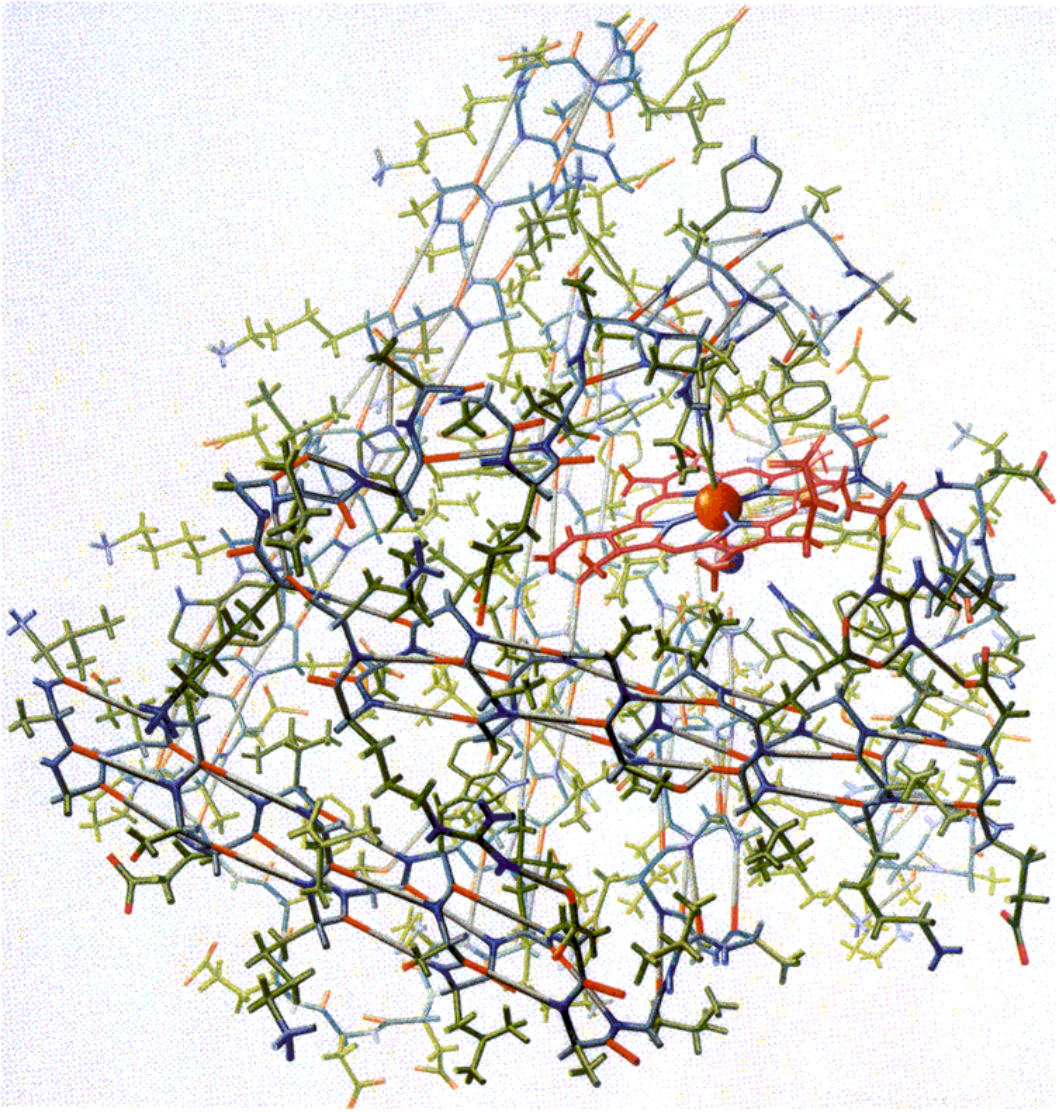
globular fold

- **Protein sequences are given as linear sequences of their one-letter amino acids:**

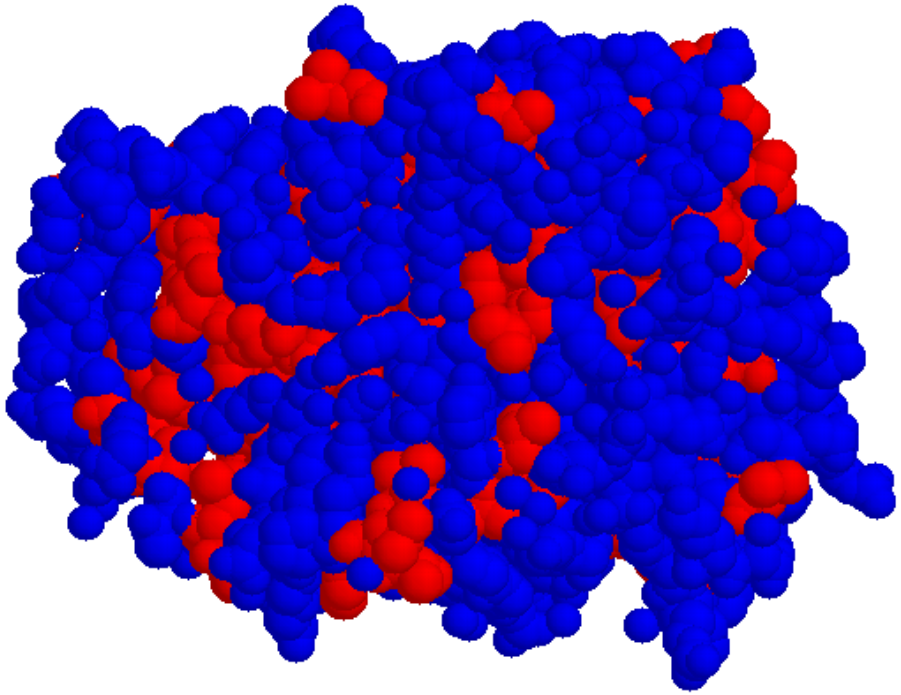
MAKRNIVTATTSKGEFTMLGVHDNVAILPTHASPGESIVIDGKE-
VEILDAKALEDQAGTNLEITIITLKRNEKFRDIRPHIPTQITETNDG-
VLIVNTSKYPNMYVPVGAVTEQGYLNLGGRQTARTLMYNFPTRAGQ….

Proteins are typically “globular” in shape

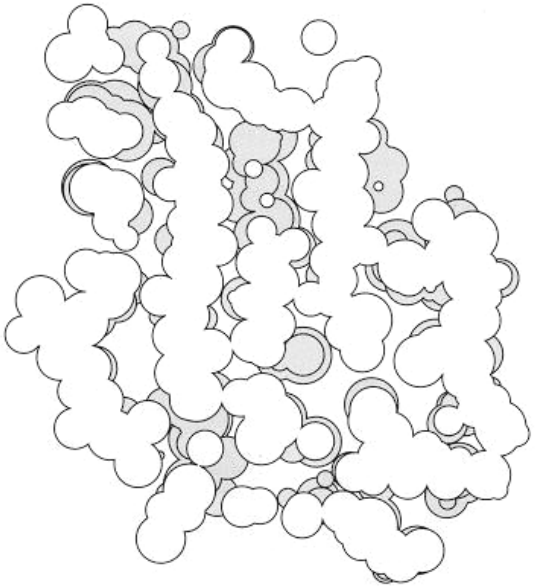
myoglobin
(17 kDa)



Protein structure is “tightly packed”, like a solid



“cutaway view”



taken from: “Protein Structure and Function” (Petsko & Ringe)

- a single protein chain “folds” into a compact structure
- hydrophobic residues on inside (red), hydrophilic on outside (blue)

Outline

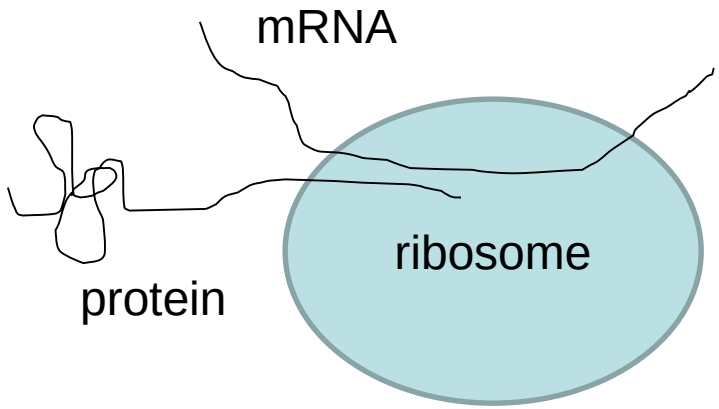
- Thermodynamic stability of proteins

  - contributions from enthalpy and entropy

- Protein dynamics

  - how dynamics has been measured/detected
  - relation to function

When proteins come off the ribosome, they typically fold



Why do they fold? Thermodynamics

How do they fold? Protein folding kinetics

Four Forces Driving Protein Structure Formation

1. hydrophobic effect

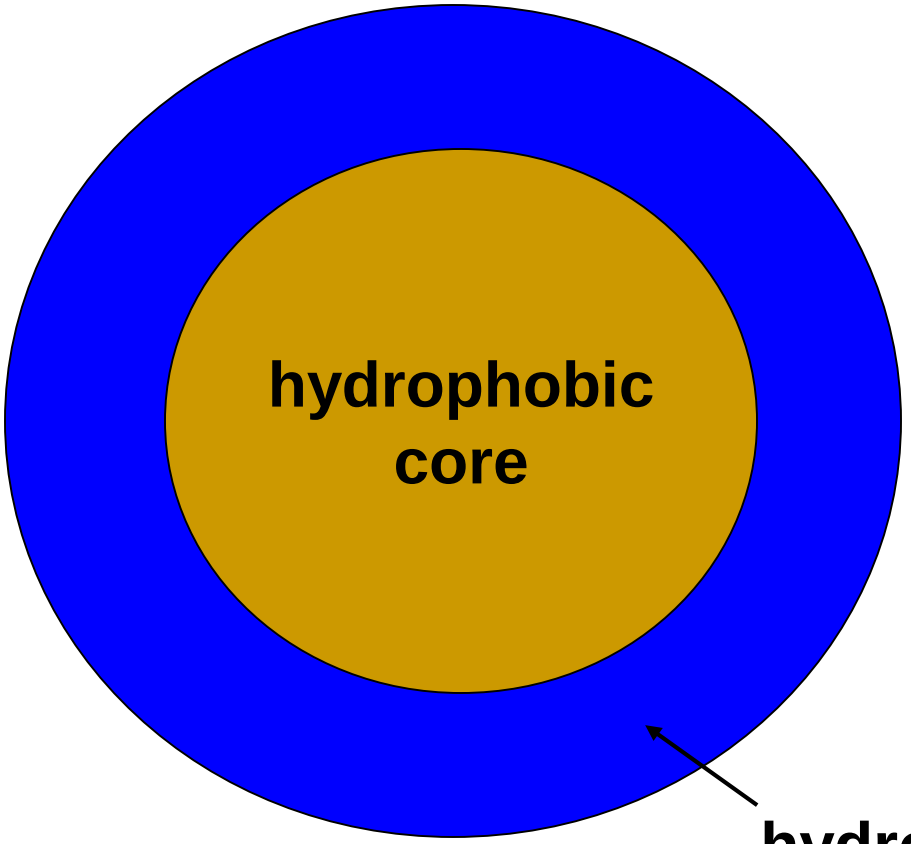
2. van der Waals forces

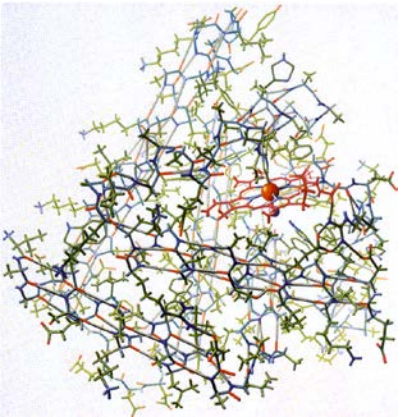
3. H-Bonds

4. electrostatic interactions

Force 1: The "hydrophobic effect"

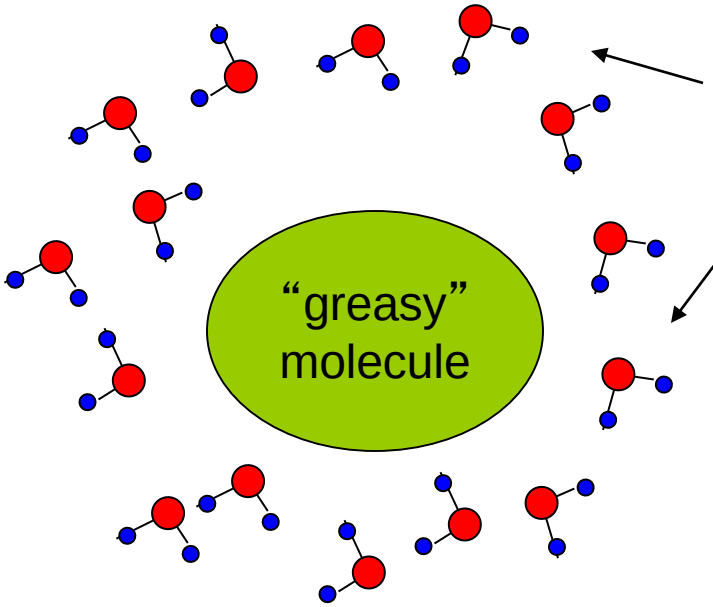
protein





“hydrophobic effect” (continued)

“clathrate” cage structures



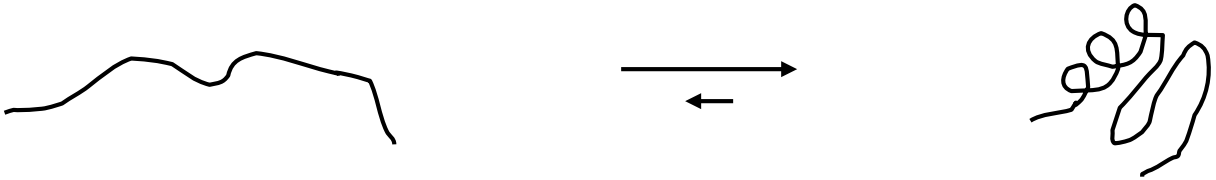
protein unfolded: hydrophobic side chains exposed => clathrate structures
protein folded: hydrophobic side chains are buried

“ordered” H_2O in clathrates are entropically unfavorable

Protein Folding and the 2-State Approximation

unfolded "state": U

folded state: F

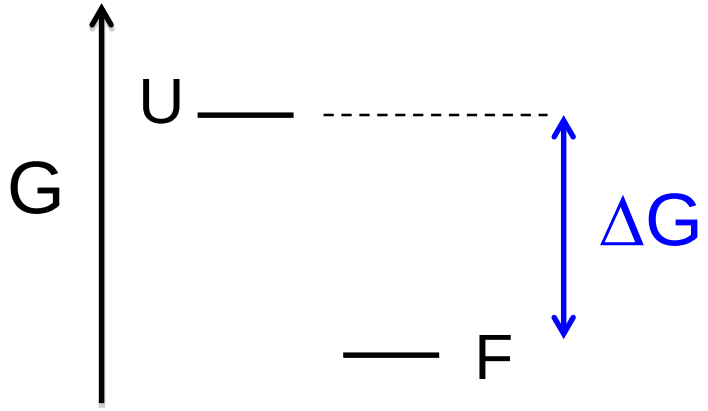


"random coil": large # of conformations

uniquely folded structure

Equilibrium Constants:

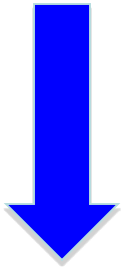
$$K_{fold} = \frac{[\text{ folded }]}{[\text{ unfolded }]} \qquad K_u = 1/K_f$$



$$\Delta G_f = - RT \ln K_f$$

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

enthalpy entropy



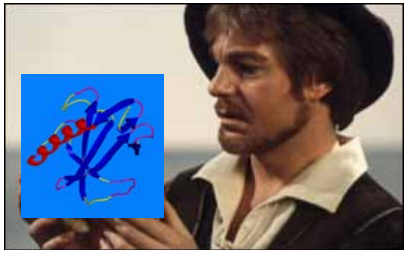
non-covalent bonding

- van der Waals interactions
- H-bonding
- electrostatic interactions

- hydrophobic effect (desolvation)

- conformational entropy

Individual classes of interactions can be strongly energetically favorable or strongly energetically unfavorable.

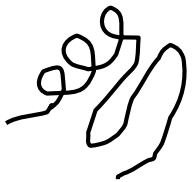


to fold

+

Favorable interactions

- enthalpy from van der Waals packing interactions
- hydrophobic effect (H_2O entropy)
- gain of protein-protein H-bonds
- electrostatic effects



or not to fold

-

Unfavorable Interactions

- protein conformational entropy
- loss of protein-water H-bonding

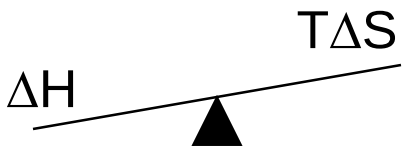
Proteins are typically "stable" by only 5-10 kcal/mole

Compare to other bond energies:

Bond Type	ΔG (kcal/mole)
hydrogen bond	1-3
ATP hydrolysis	~7
C-H covalent bond	~100

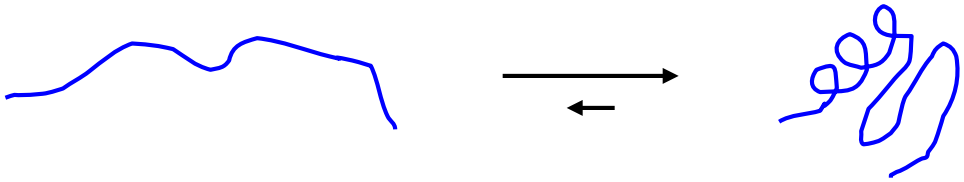
Protein folding stability is precariously balanced $\Delta G = \Delta H - T\Delta S$

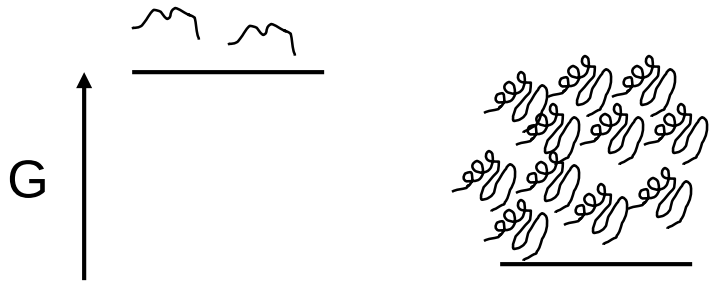
enthalpically favored
entropically unfavored (or is it?)



Proteins are in equilibrium with the denatured state.

Because the ΔG is ~5-10 kcal/mole, there is a small (but not insignificant) population of unfolded proteins.





Protein Dynamics

Fine, proteins have shapes and stable structure.

So what's the big deal?

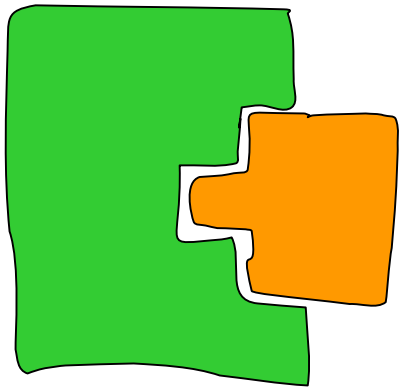
Proteins actually DO THINGS!!

- bind other molecules (proteins, DNA, metabolites)
- catalyze reactions (enzymes)
- movement, such as muscle contractions, rotary motors, etc.

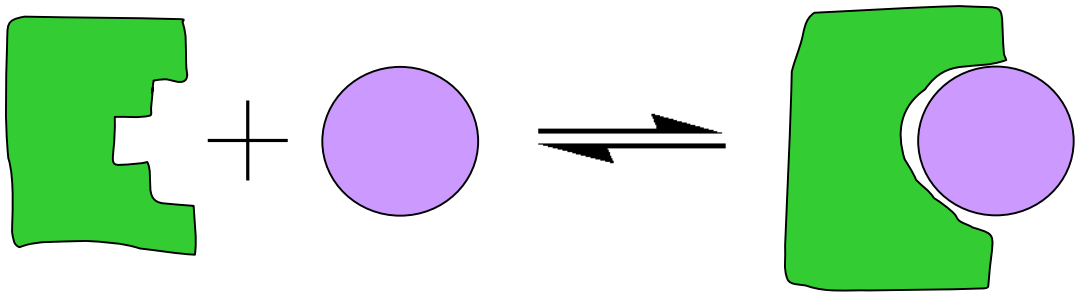
Ligand binding: Proteins can change their shape

- Protein structures fluctuate on many different timescales
- They can unfold (and refold)
- They can switch to another conformation

Lock and Key



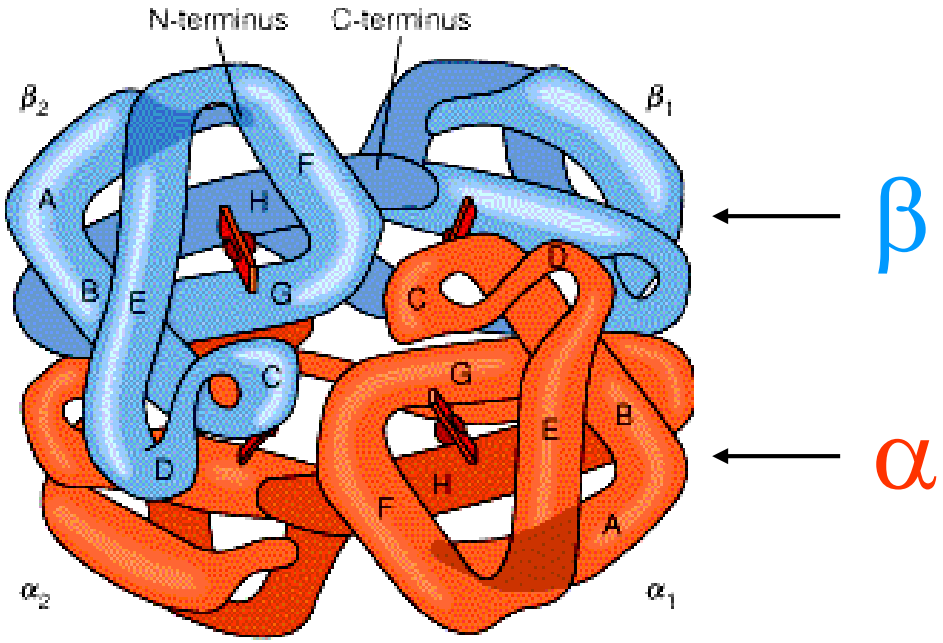
"Induced Fit"



Hemoglobin: an allosteric protein

4 chains: 2 "α" chains
2 "β" chains

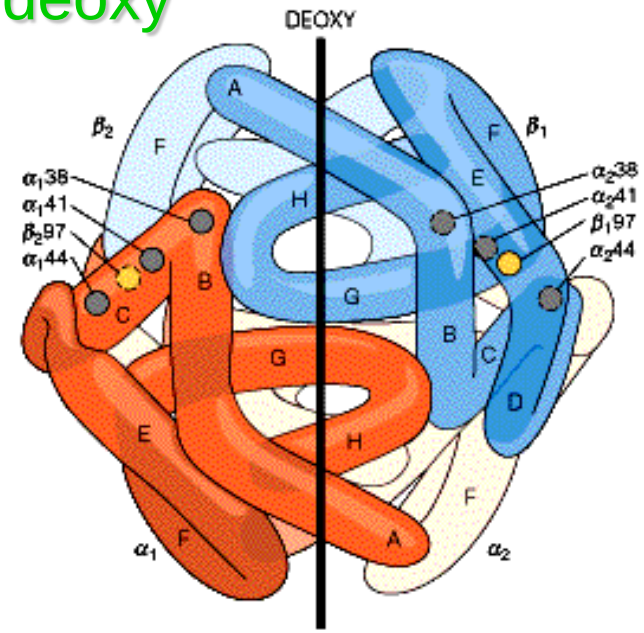
4 hemes
4 O_2 binding sites



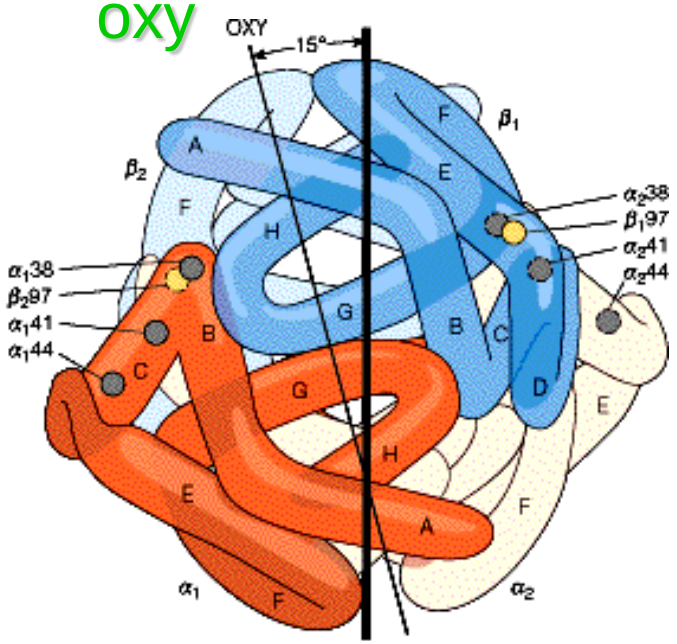
O_2 binding at the 4 sites are NOT independent!

“T” “R”

deoxy

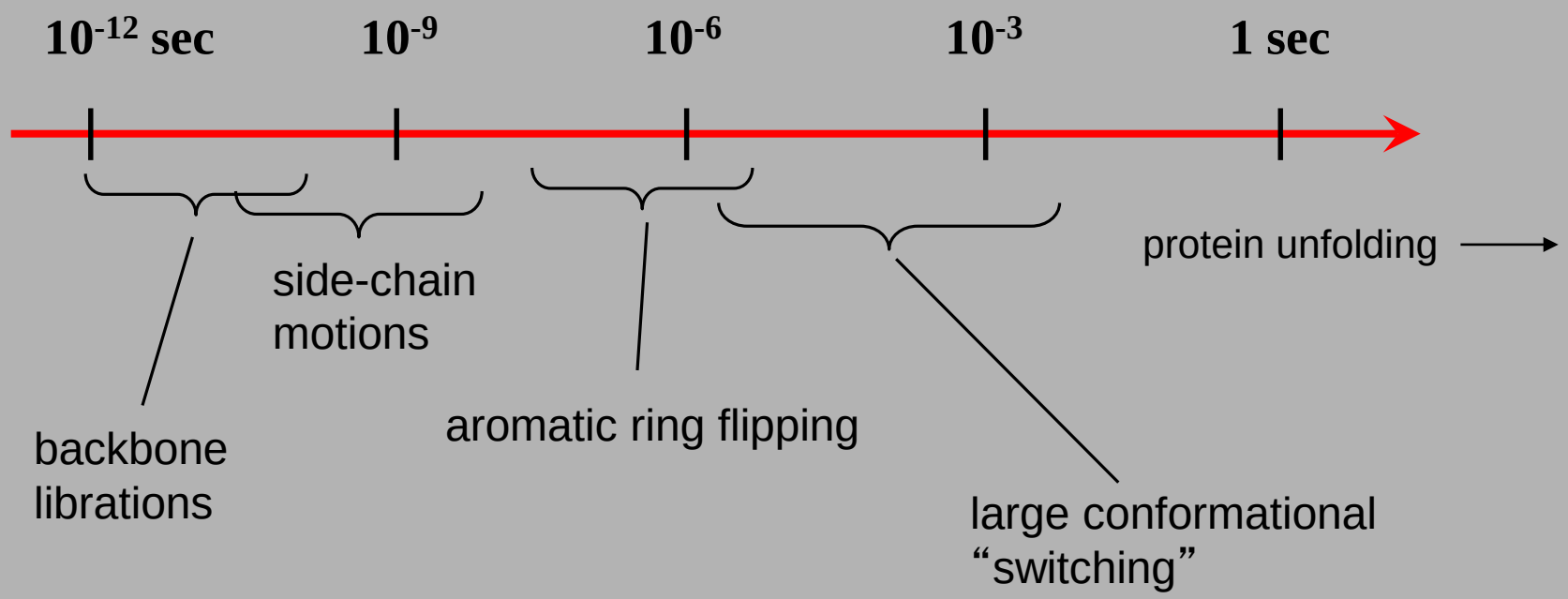


oxy

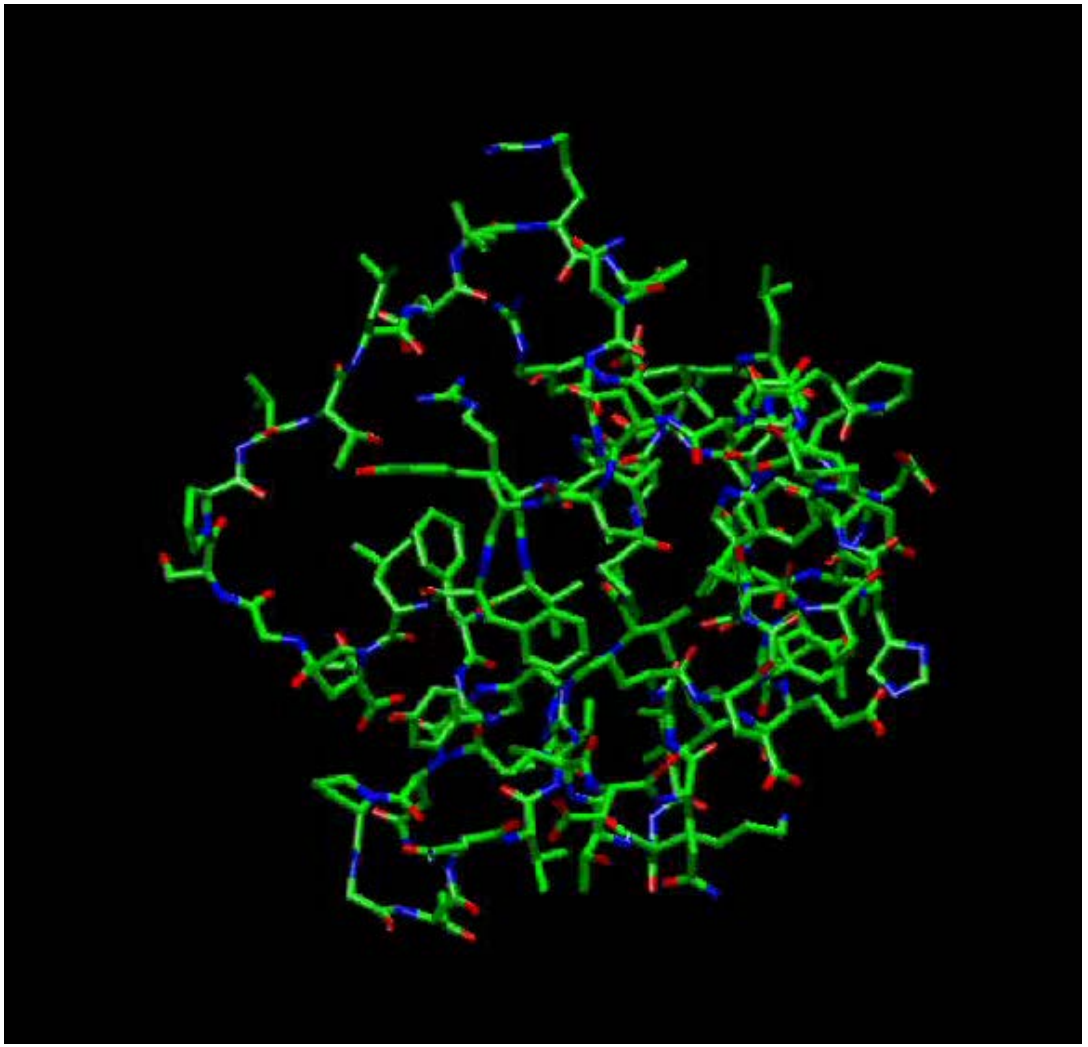


- Adjustment (small) of tertiary structure in monomers
- Adjustment (large) of quarternary structure at the chain <u>interfaces</u> (loss of deoxy interactions “paid” for by binding of O_2)
- **allosteric “communication”** between O_2 binding sites is possible because of hemoglobin’s tetrameric structure

Timescales of Motion in Proteins



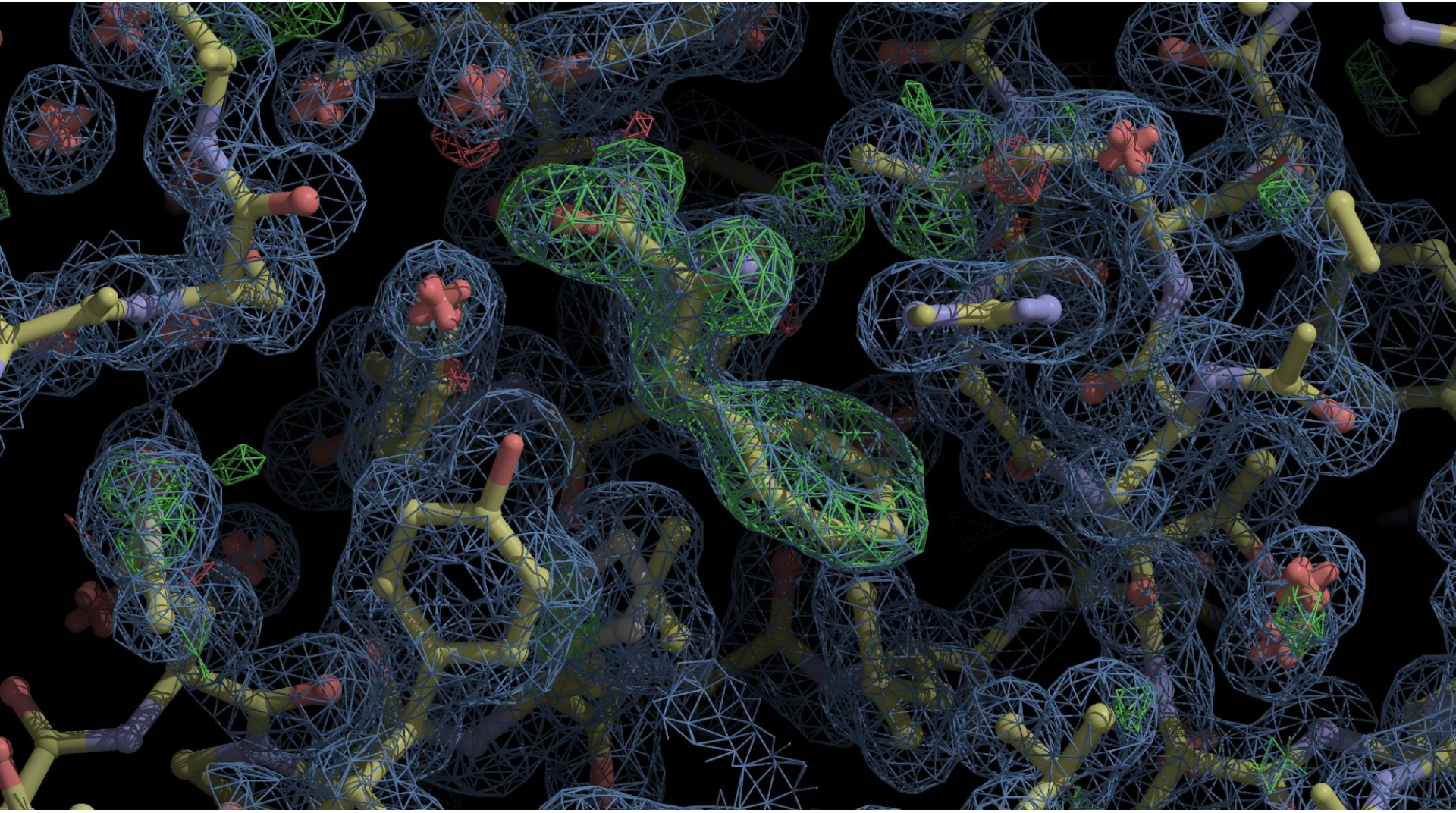
Proteins have dynamic flexibility



length of movie = 5 nanoseconds (5×10^{-9} sec)

Yet, the “static model” of protein structure is firmly embedded in our psyche:

“Seeing is believing”



2 Key Questions

- How do we know that proteins fluctuate? Can we see it experimentally?
- Are fluctuations in the structure important for function?

Timeline for key experiments/observations in protein dynamics

Evidence for:

Spectroscopy: G. Weber described proteins as "kicking and screaming" 1975

Computer simulations (first time to "see" dynamics) 1980's

☞ Cold crystal experiment (x-ray) 1992

☞ The protein energy landscape (applies to dynamics too) 1991

Actually "seeing" it experimentally in proteins (NMR):

Clear demonstration of aromatic side chains "flipping" by NMR 1980's

☞ NMR relaxation, (and later, Vendruscolo visualization) 1990's (2005)

☞ Wand nature paper (entropy, the "line") 2007

DHFR dynamics (function) 2006

Experimental evidence for importance of dynamics for function

Rassmussen et al., *Nature* (1992)

The "glass transition": protein motions "freeze out" at 220 Kelvin

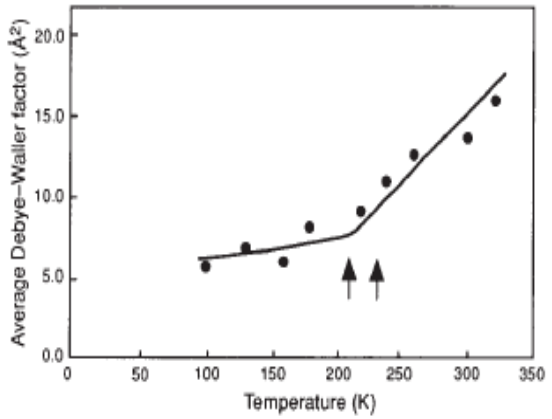


FIG. 1 Plot of the average Debye–Waller factor for all of the non-hydrogen protein atoms in crystalline bovine pancreatic ribonuclease A as a function of temperature (redrawn from data in ref. 1). The plot is biphasic, but the exact transition temperature is hard to estimate. This difficulty in part arises because the transition is broad, but there is also evidence that it depends on solvent composition and on the rate of cooling (see for example, ref. 16). Our estimate of the midpoint at 220 K is approximate; it is possible that the transition is still occuring at 200 K under the conditions of our experiment. We chose the two temperatures indicated by arrows for binding experiments so that they were as close as possible while still representing the two regions of behaviour.

RNase A experiment

- crystal at room temp: cuts RNA in two
- crystal at 230 K: slowly cuts RNA in two
- crystal at 210 K: no cutting

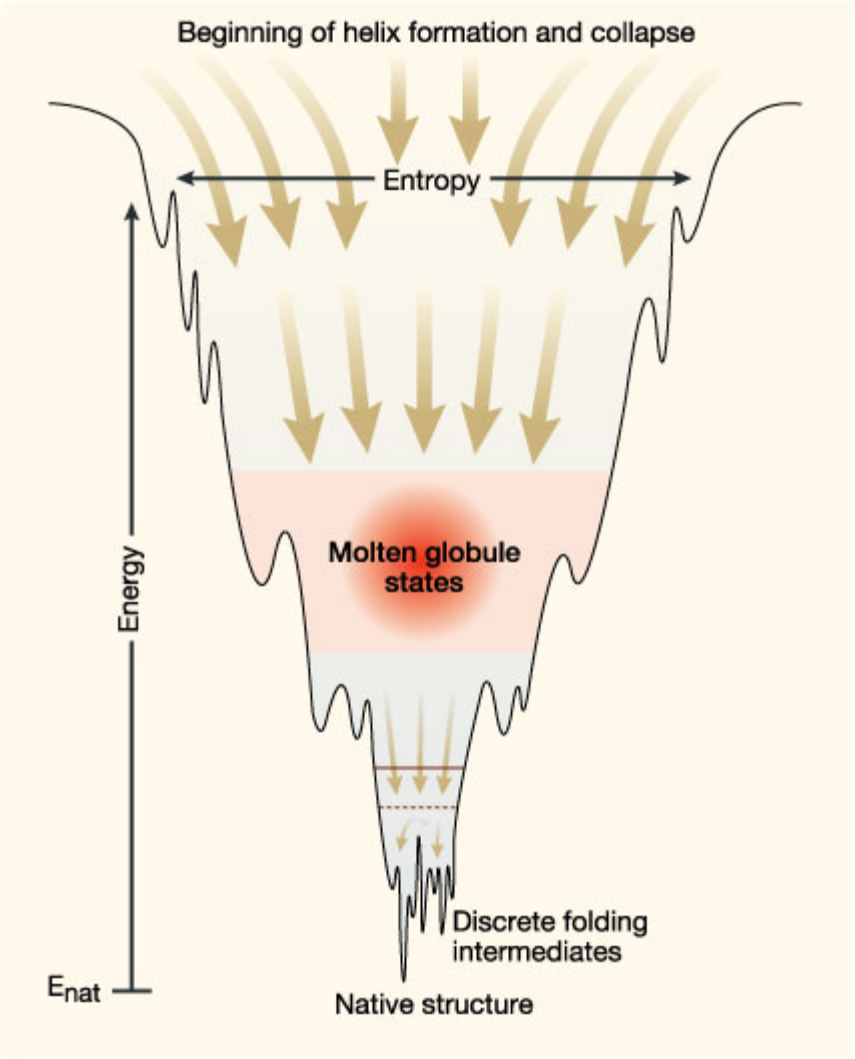
from x-ray diffraction, no difference in structure betweetn 210 and 230 K!!!

- warm crystal to 230K: slow cutting again

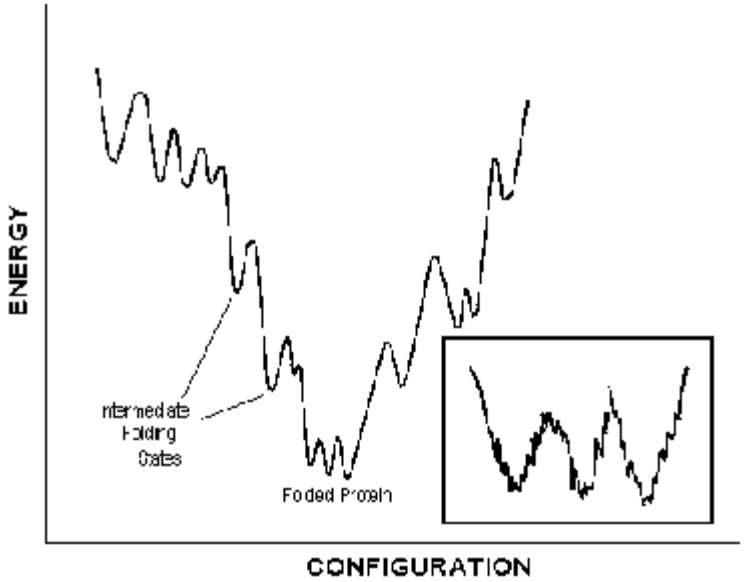
conclusion: thermal motions required for RNase A activity.

Protein energy landscape

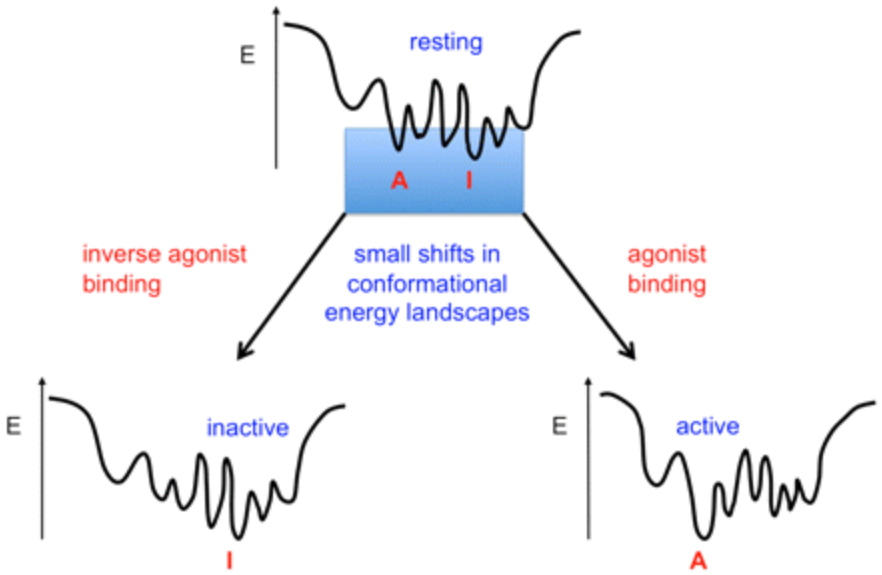
The “accepted” view



The “new” view



for G-protein coupled receptors (GPCRs)



Chen et al., *PNAS* (2012), 33, 13284-13289.

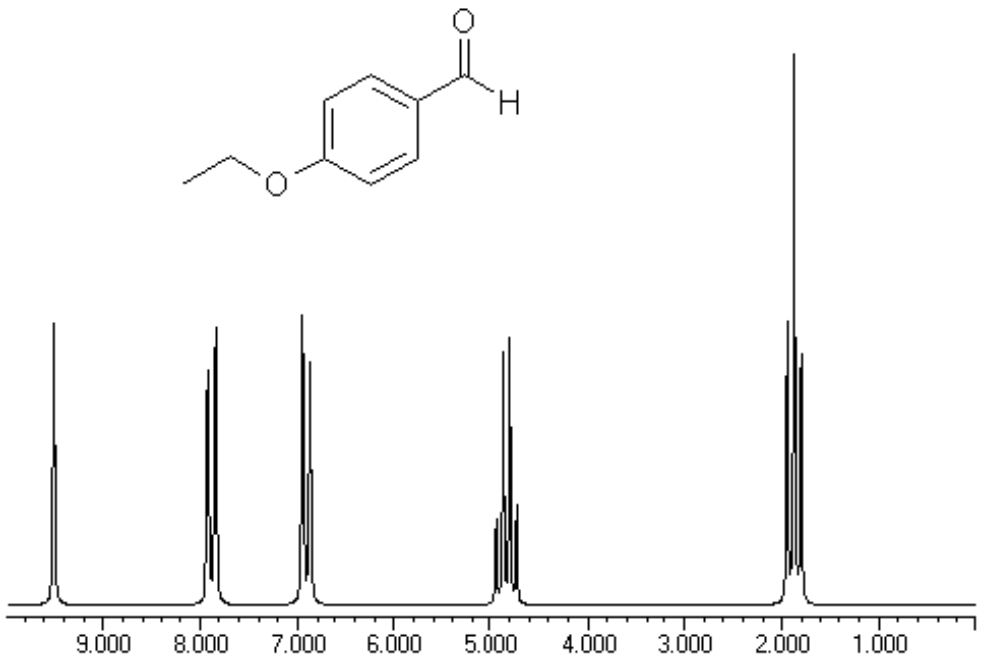
NMR Spectroscopy

- Proteins are studied in solution
- Structural information obtained by nuclear spin spectroscopy


OXFORD
VNMR Command
and Parameter
Reference

Chemical Environment:

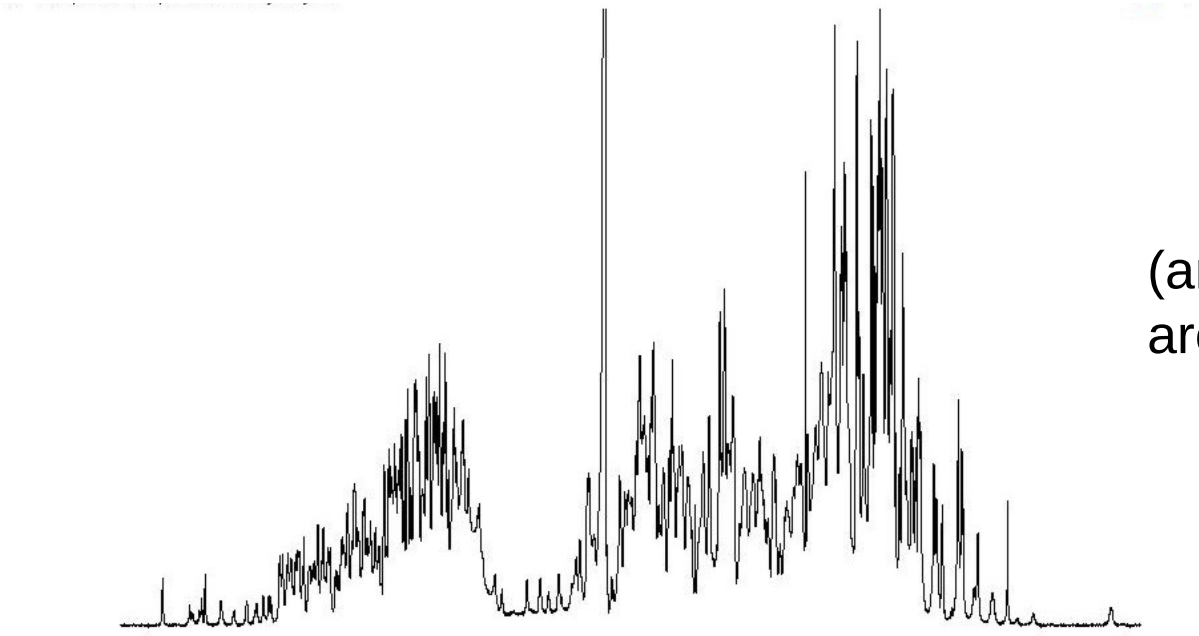
- chemical shift
- J-coupling



^{1}H (ppm)

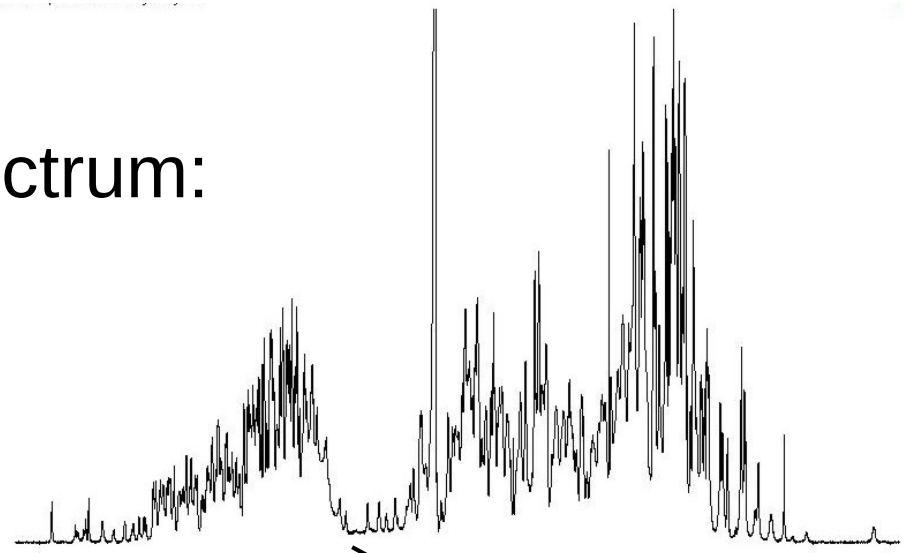
- For ^{1}H and ^{13}C spins, 1D FT-NMR spectra can be obtained.
- However, spectral complexity increases with increasing molecular weight.

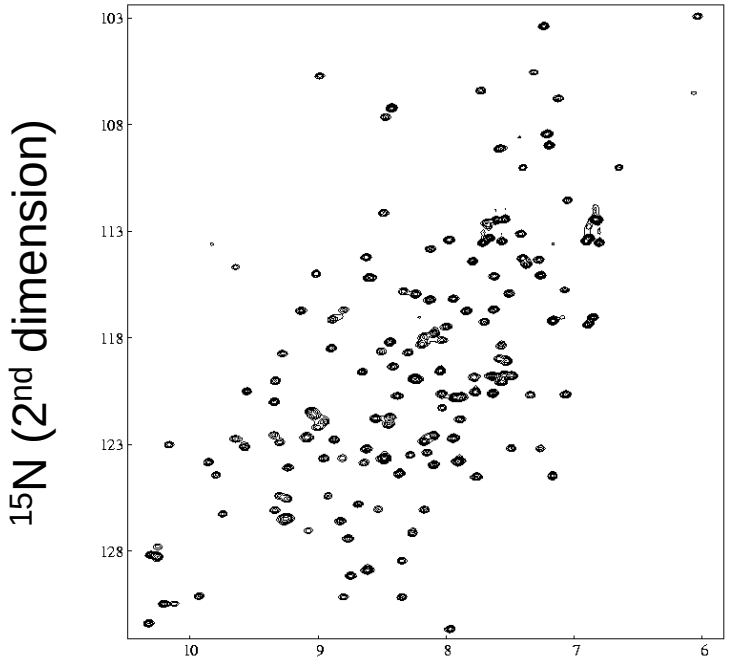
^{1}H spectrum of lysozyme at 800 MHz



(and the J-couplings aren't even seen here!!!)

^{1}H NMR spectrum:

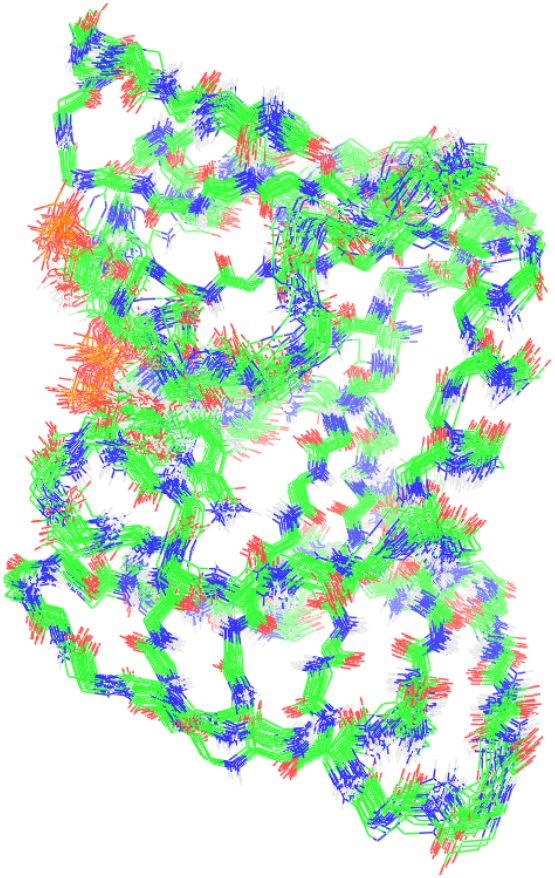




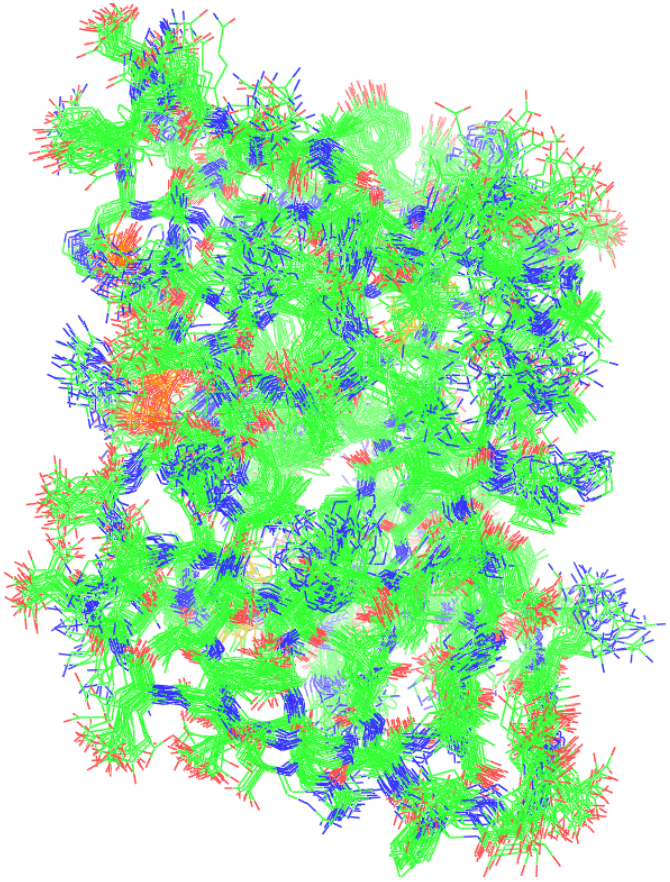
Can resolve more “peaks” by adding a 2nd dimension.

Structure solved in solution by NMR:

Proteins exist as conformational ensembles

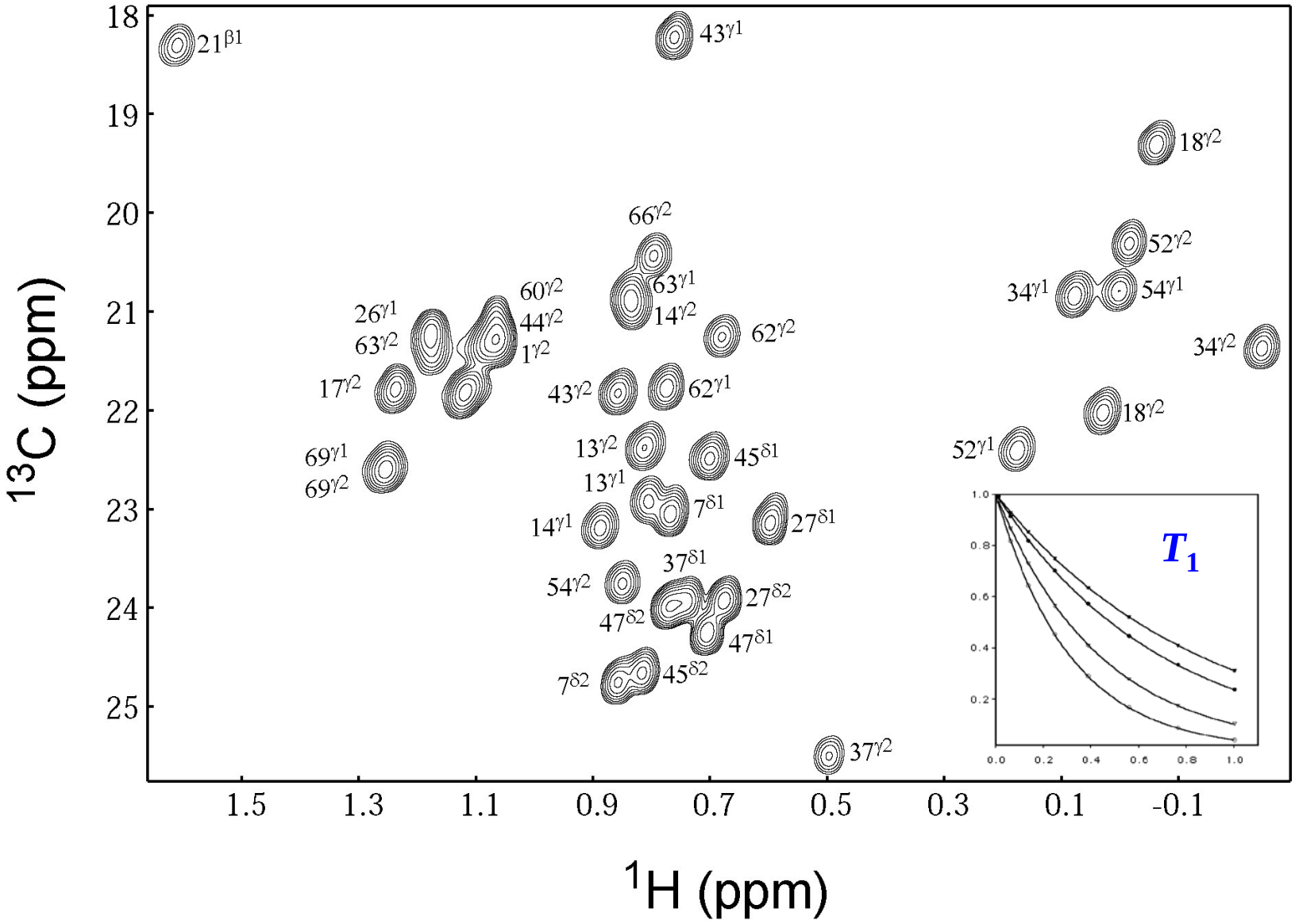


backbone atoms only



all atoms (i.e. with side chains)

Spin-Relaxation: a more direct measure of dynamics

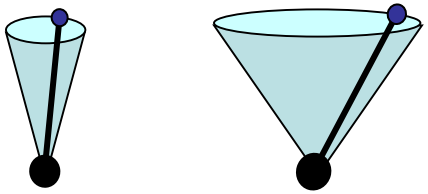


Characterization of dynamics directly from relaxation of NMR signals

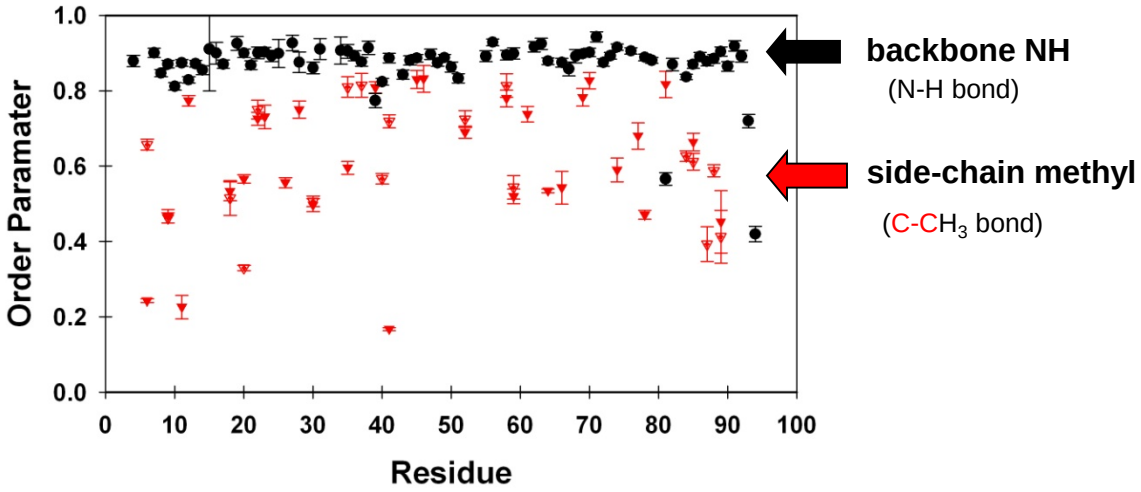
Order parameter: S^2

1 = fixed orientation
0 = no preferred orientation

captures motion on ps-ns timescale



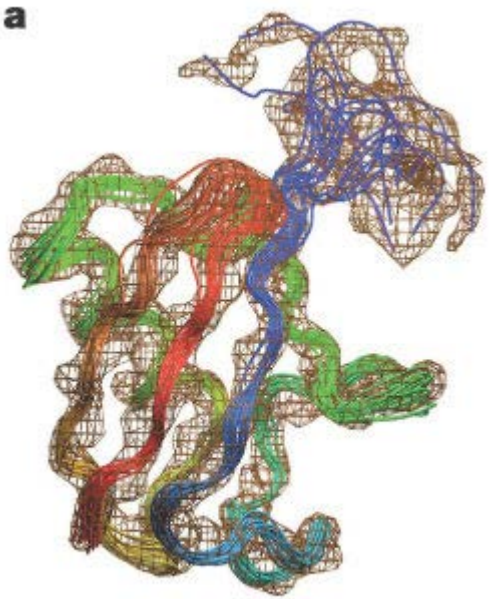
high S^2 **low S^2**



backbone NH
(N-H bond)

side-chain methyl
(C-CH_3 bond)

use S^2 for ensemble restraint.
ps-ns timescale ensemble



core is “liquid-like”

work by Michele Vendruscolo
and coworkers

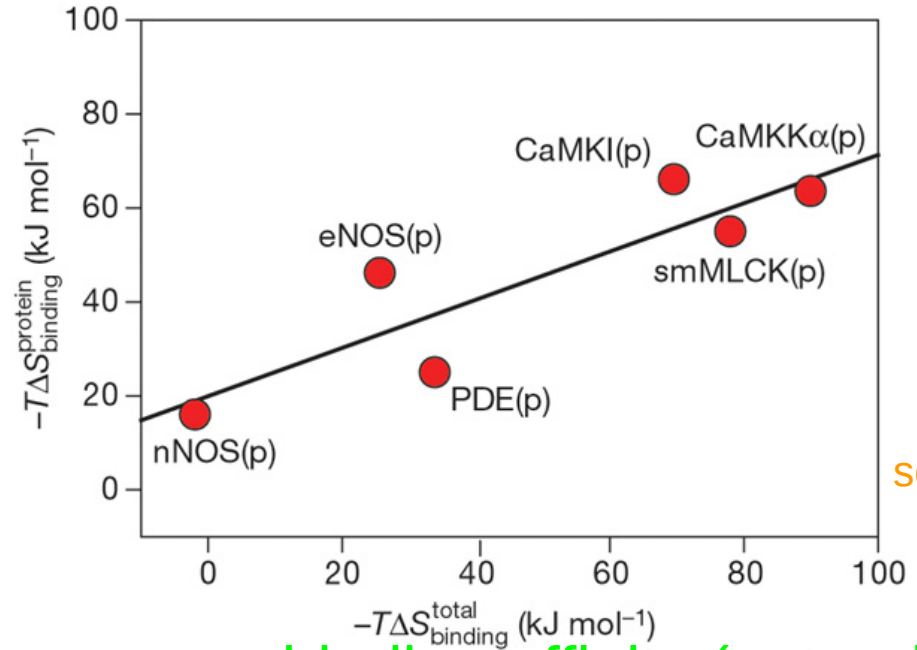
Protein flexibility ("dynamics") can affect protein-ligand affinity

ΔG (overall free energy) = $\Delta H - T\Delta S$

S = entropy (measure of "disorder")

Calculate entropy for individual residues from the order parameter:

dynamics



A. Joshua Wand and coworkers.

Nature (2007), 448, 325.

see also work by Kalodimos

Nature (2012), 488, 236.

binding affinity (entropic component)

Actual change in entropy upon binding, measured thermodynamically (actual energy!)

Where are we at now with our knowledge of protein dynamics?

- “Protein dynamics” is still a relatively new field of research
- Complementary to static structure determination (i.e. x-ray diffraction)
- Motions on order of 10 ns – 1 μs very hard to characterize in detail
- Computer simulated proteins provides the “clearest” picture
  (but computer power is an issue – hard to go longer than tens of μs)

- Dynamics important for function, but we are just beginning to understand why.

Study Questions

1. Would you expect a globular, folded protein to spontaneously unfold? If “no”, why? If “yes”, what would you expect to happen after that?

2. Based on basic principles of physical chemistry, how can dynamics affect free energy changes via entropy?

see optional question #3 on next slide……

Optional question:

4. For a protein that is stable at 3 kcal/mole (that is, ΔG_{fold} = -3 kcal/mole), calculate the percentage of proteins that will be unfolded at 25 ºC at equilibrium?

3) You encounter a shrink ray that reduces you to a 10nm sized human being. Taking advantage of your small size, you decide to scuba dive into a beaker and observe protein folding and dynamics. This protein is 20kDa in size, has both hydrophobic and hydrophillic residues, and is known to be allosteric.

a) Given your knowledge of the driving forces for protein folding and unfolding, visually represent these forces in play for a protein in water. Label your diagram and include detailed explanations for what you see, considering all the complications discussed this week for protein folding and dynamics. (E.g. Drawing representative amino acid residues within part of this protein that are known to be involved in van der Waals interactions and whether these (non-covalent “bonds”) are stronger or weaker than other types of bonds, what forces are at play with these bonds, how these forces do or do not contribute to folding, ΔH, or ΔS.)

b) The scientist whose bench you have invaded adds adequate **ligand** to bind to both binding sites on the allosteric protein. Again, visually represent what happens to the protein and the forces at play. Include in your discussion the dynamics of the protein and how it would affect ligand binding, conformational selection vs. induced fit models, and which model you think is more likely to represent your allosteric protein.