The Abdus Salam
International Centre for Theoretical Physics

United Nations
Educational, Scientific and
Cultural Organization

IAEA
International Atomic Energy Agency

SMR/1845-8

Conference on Structure and Dynamics in Soft Matter and Biomolecules: From Single Molecules to Ensembles

4 - 8 June 2007

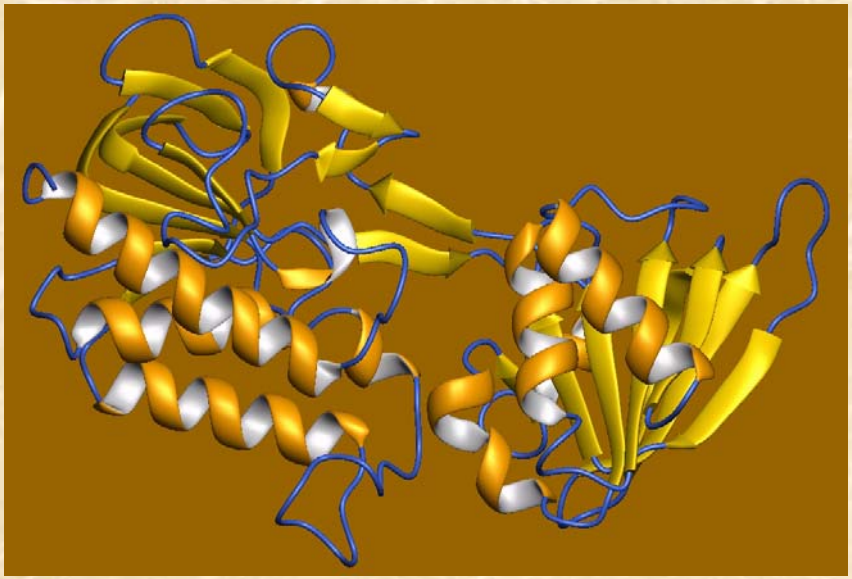
Native states of proteins and computational modeling of their structure and structure-function relationships

Maria KURNIKOVA

Carnegie Mellon University
Department of Chemistry
4400 Fifth Avenue
503 Mellon Institute
Pittsburgh, PA 15213-3890

Native State Proteins
Introductory Lecture

Maria Kurnikova
Chemistry Department
Carnegie Mellon University
Pittsburgh, USA



Proteins – Multiple Functions in Living Cell

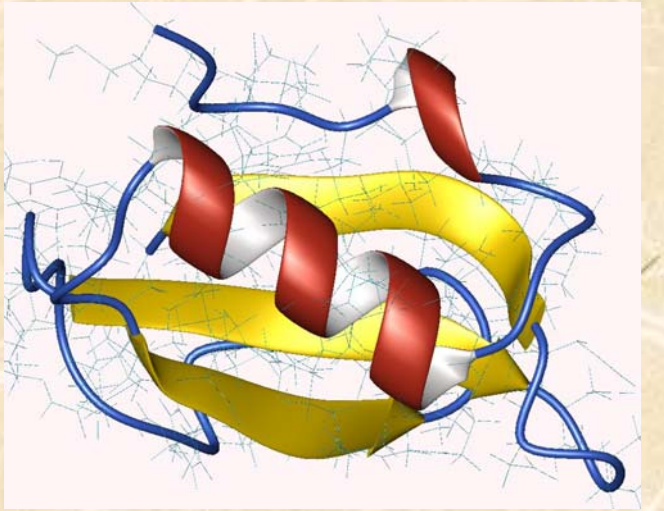
- Molecular machines
- Building blocks
- weapons

- Biochemical catalysis
- Regulation of gene expression
- Reception and Signaling
- Bioenergetics: photosynthesis, breathing, mechano-chemical work

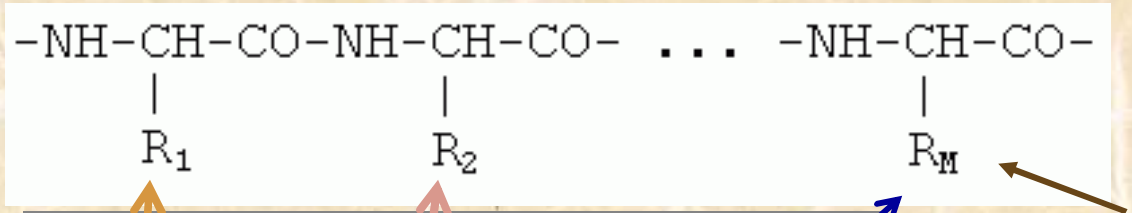
Key Property of Proteins:
Highly Specific to their Function

Highly Specific Molecular Interactions

What is a Protein?

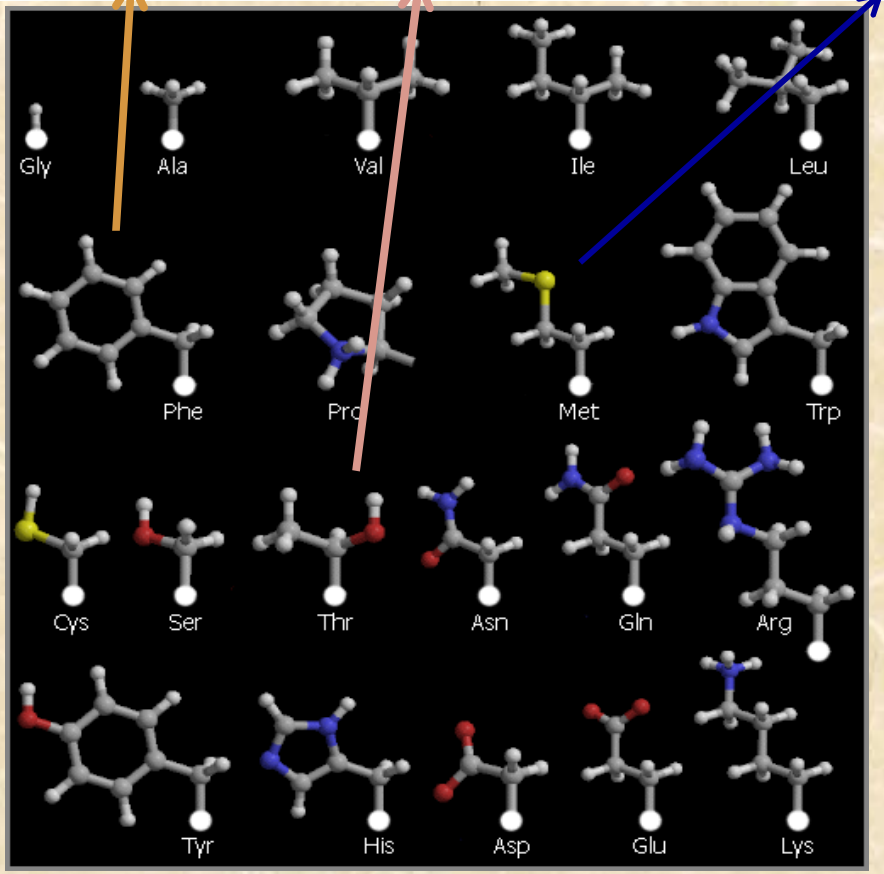


Protein is a Linear Polymer
Synthesized by a cell using gene code



←main chain

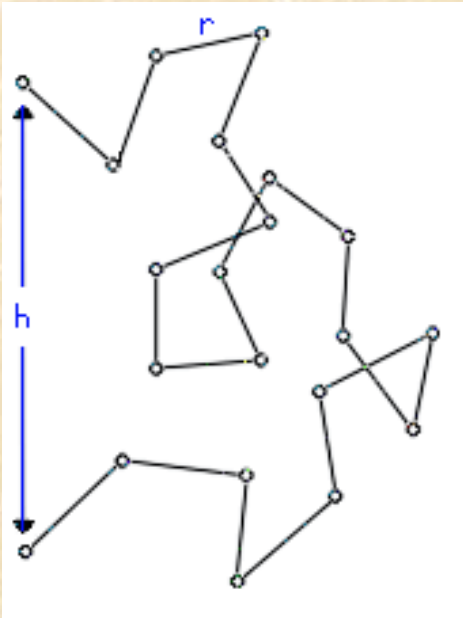
side chains



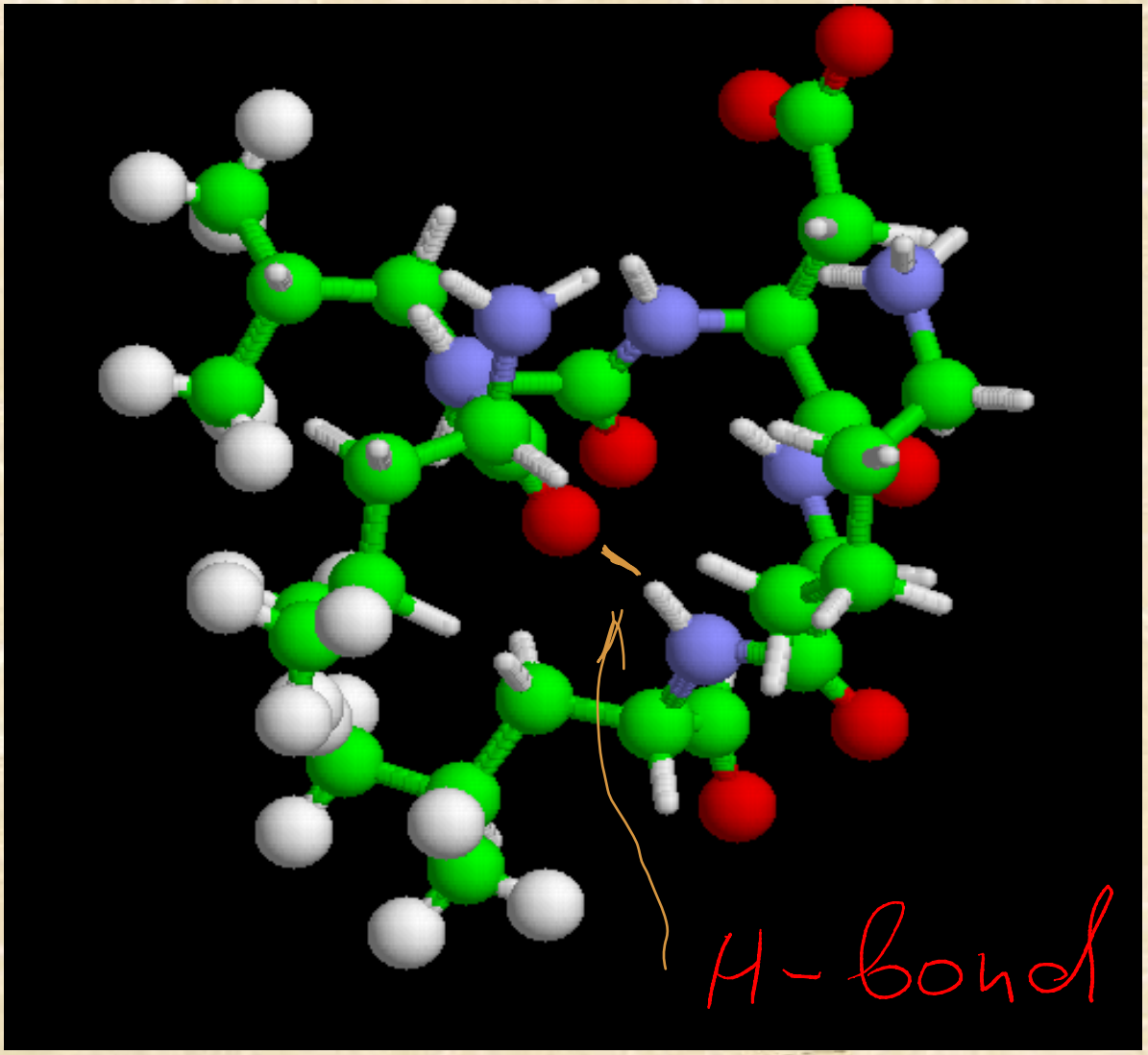
Residues of Amino Acids of native proteins (20)

Random Coil

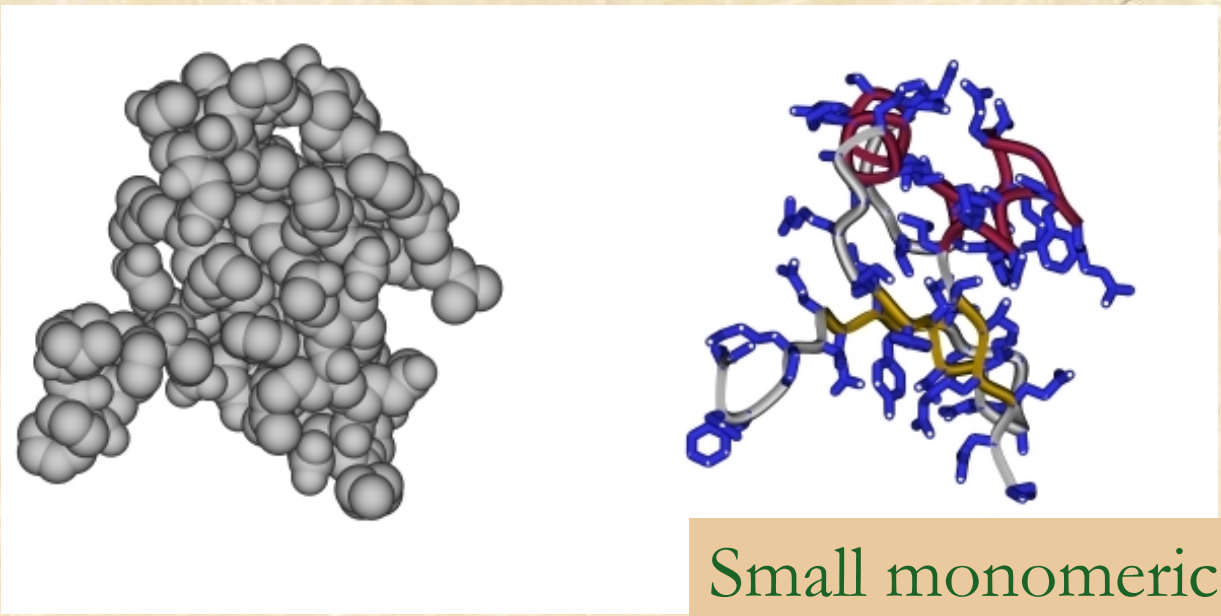
- Not Tightly packed – large in size
- Has water inside



Tight network of bonds in a folded protein



All atom Representation of a Protein

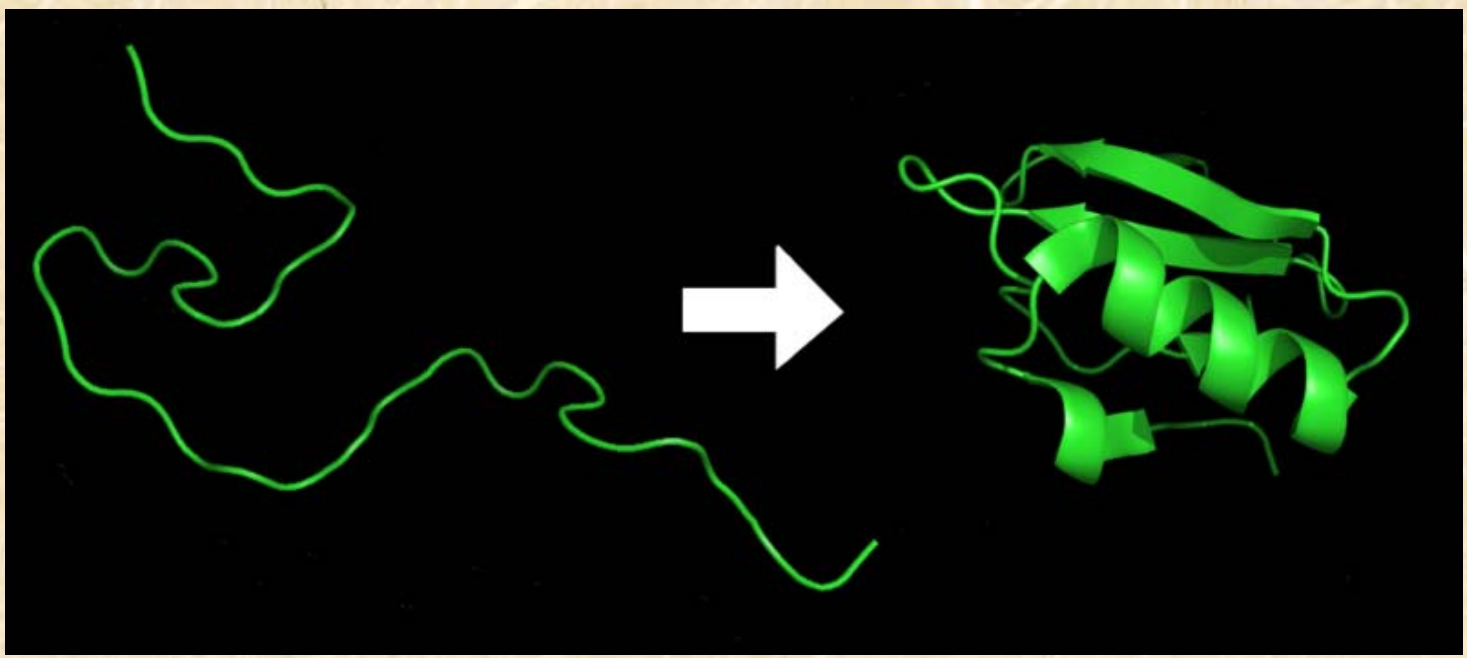


Small monomeric proteins

- Protein chain is tightly packed
- No space inside
- Fairly Rigid – important for Function

The Surface and its Shape is most important for function – the rest – forms "rigid" supportive skeleton

"Functional" Protein has Definite Spatial Organization – 3D Structure – Native State

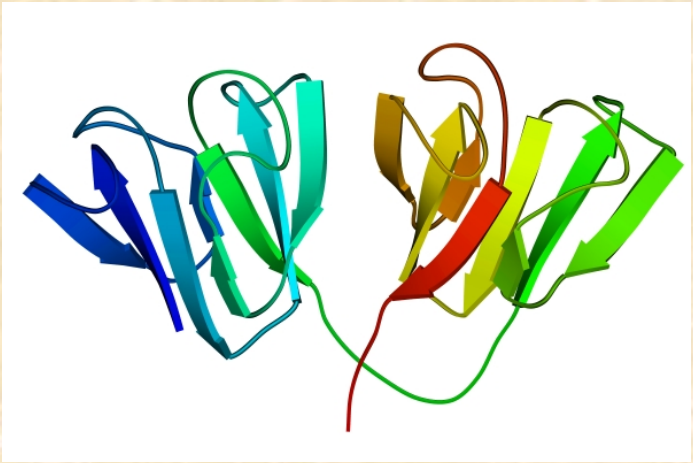


- First Protein Crystal – Hemoglobin (ca. 1860)
- First X_ray structures (late 1950)

Physical Properties of Protein
Important for Function – Aperiodic crystal

✓ Small – 50 -150 aa – very dense crystal-like globules 25-40 Å

✓ Large – several separate globular domains



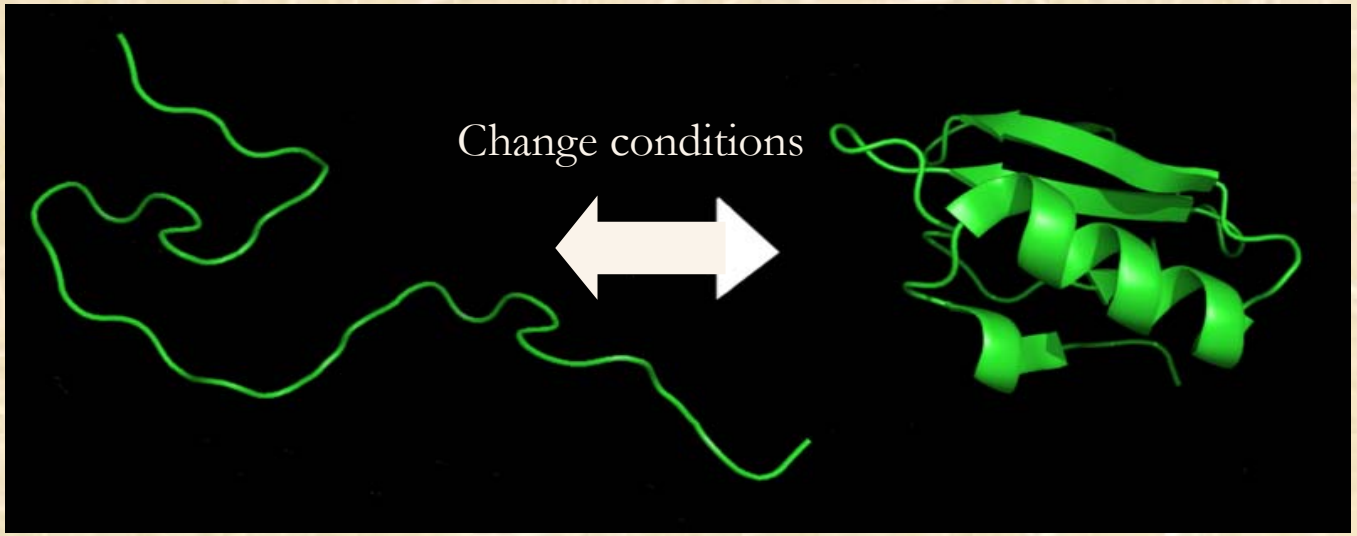
- Well-defined 3D Structure – not random
- **Melts like a crystal – not like glass – all at once…**

why?

??Functional robustness??

Proteins are capable of self organization

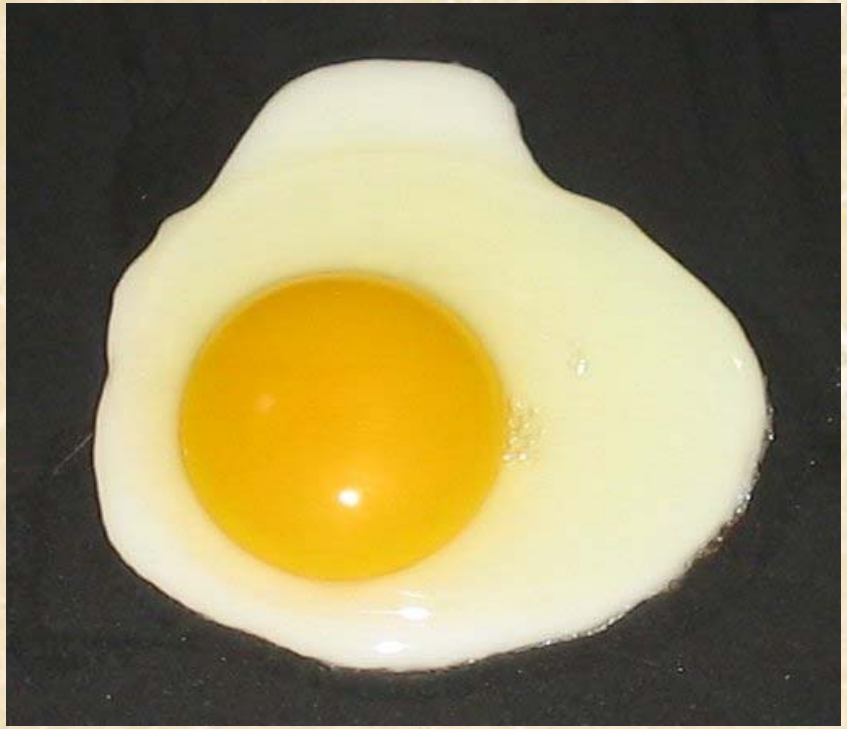
- Asfinsen (ca. 1960): *in vitro* spontaneous refolding of a mildly denatured protein in absence of any other macromolecules



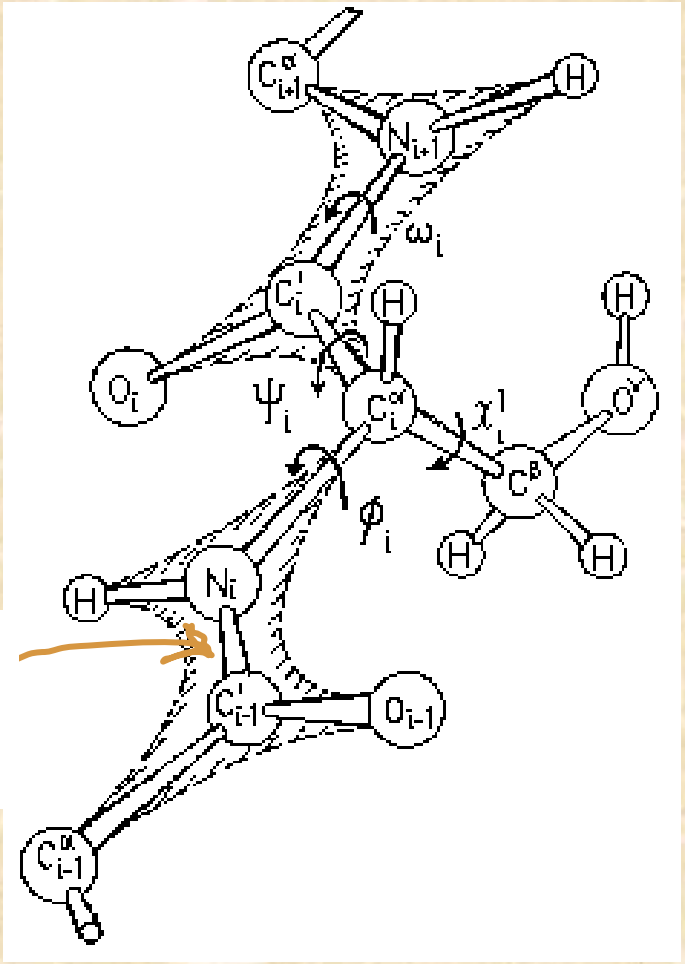
term: *in vitro* – in a test-tube
in vivo – in a biological organism

Strongly Denatured Protein

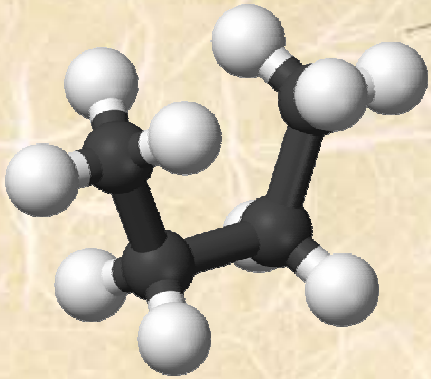
- IRREVERSIBLE

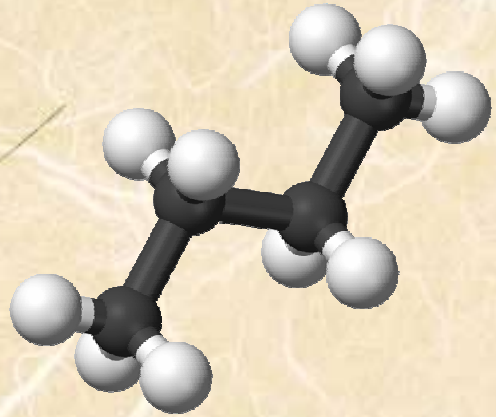


Flexibility of a Polypeptide Chain



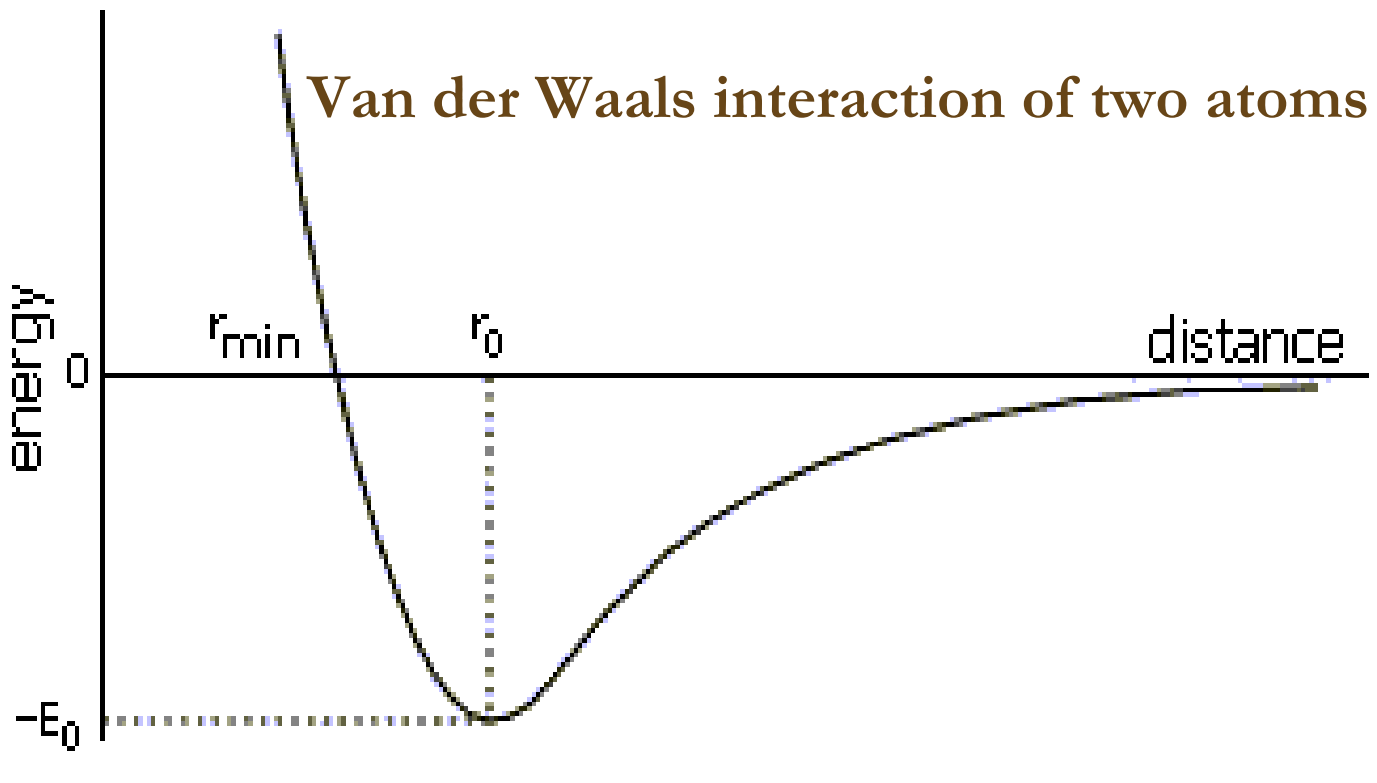
Peptide bond





Not all conformations of a polypeptide main chain are OK

- Steric overlap of atoms = increase of interaction energy between two atoms if they are too close



Typical parameters of van der Waals interaction potentials.

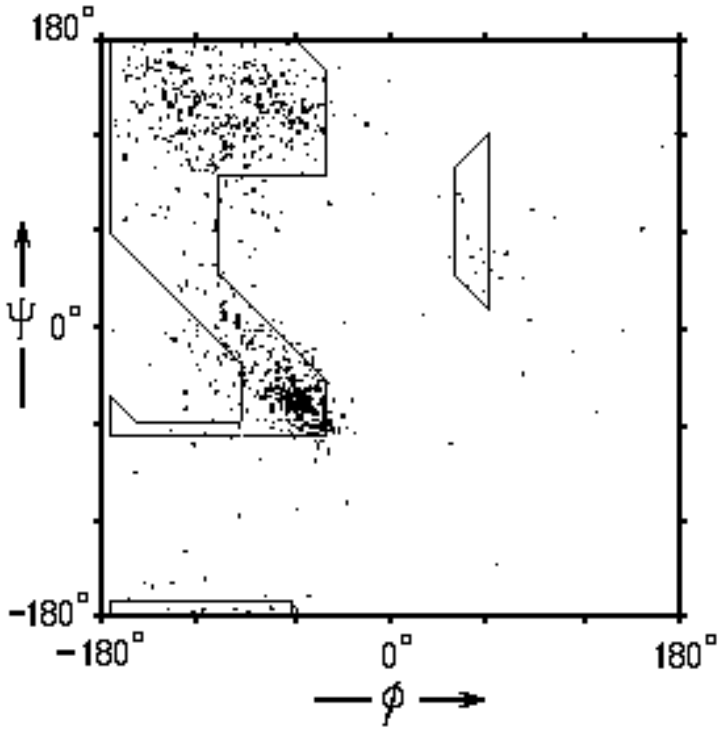
Interaction	E_0, $\frac{kcal}{mol}$	r_0, Å	r_{min}, Å	Min. v. d. Waals radii of atoms, Å	
H....H	0.12	2.4	2.0	H:	1.0
H....C	0.11	2.9	2.4		
C....C	0.12	3.4	3.0	C:	1.5
O....O	0.23	3.0	2.7	O:	1.35
N....N	0.20	3.1	2.7	N:	1.35
CH_2...CH_2	≈0.5	≈4.0	≈3.0	CH_2:	≈1.5

R.A.Scott, H.A.Scheraga, *J. Chem. Phys.* (1965) **45**:2091

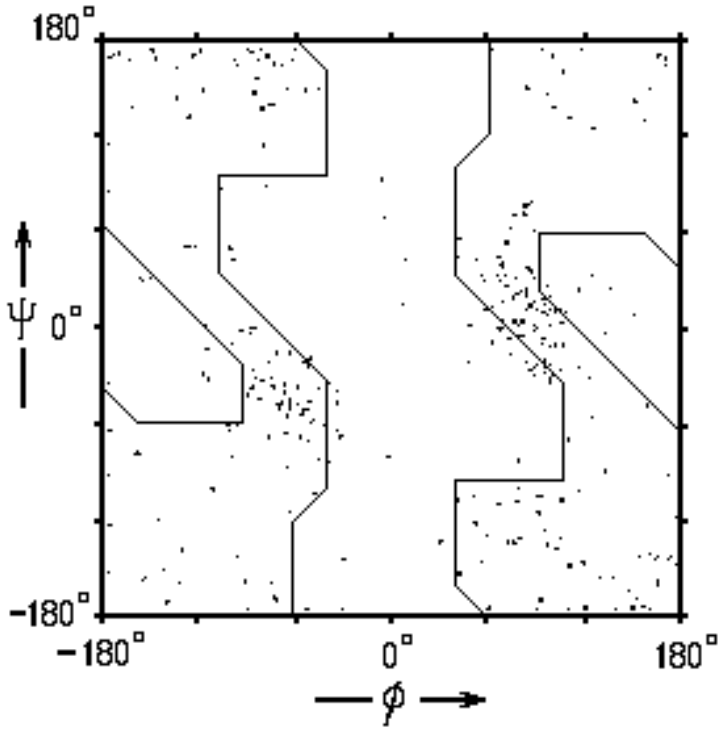
G.N.Ramachandran, V.Sasisekharan, *Adv. Prot. Chem.* (1968) **23**:283

Ramachandran Plots

All Residues



Glycine



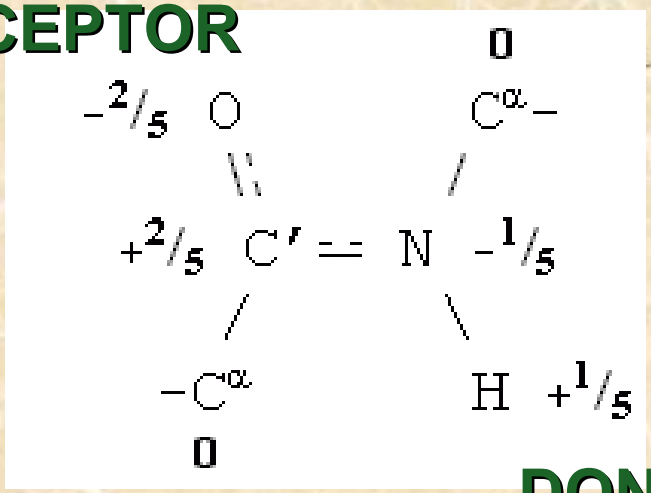
Contour lines show allowed and disallowed vdw regions

Dots – values from real protein structures

Hydrogen Bonds in Polypeptides

- O—H : : : O
- N—H : : : O
- N—H : : : N

ACCEPTOR



DONOR

H : : : O or H : : : N distance

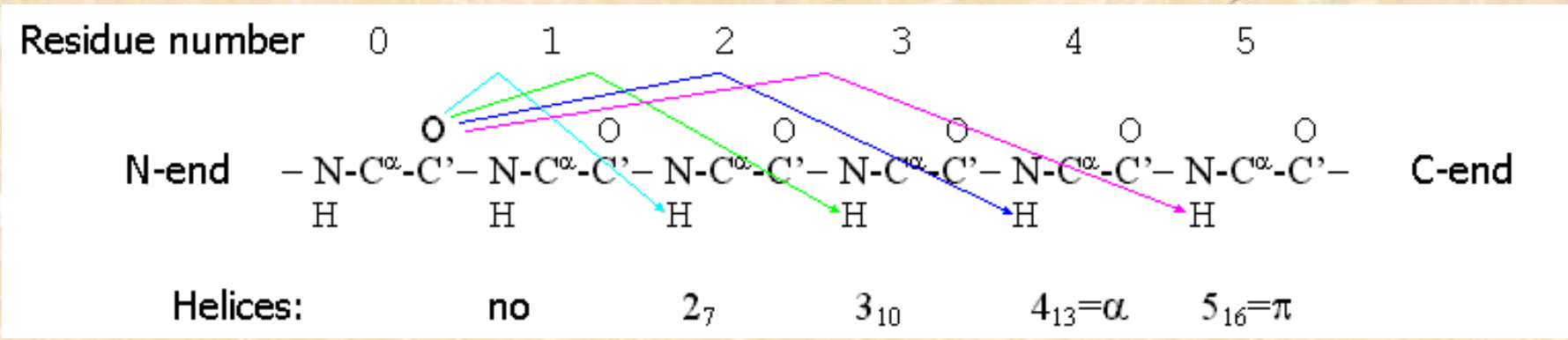
2.35 — 2.75 – normal van der Waals interaction

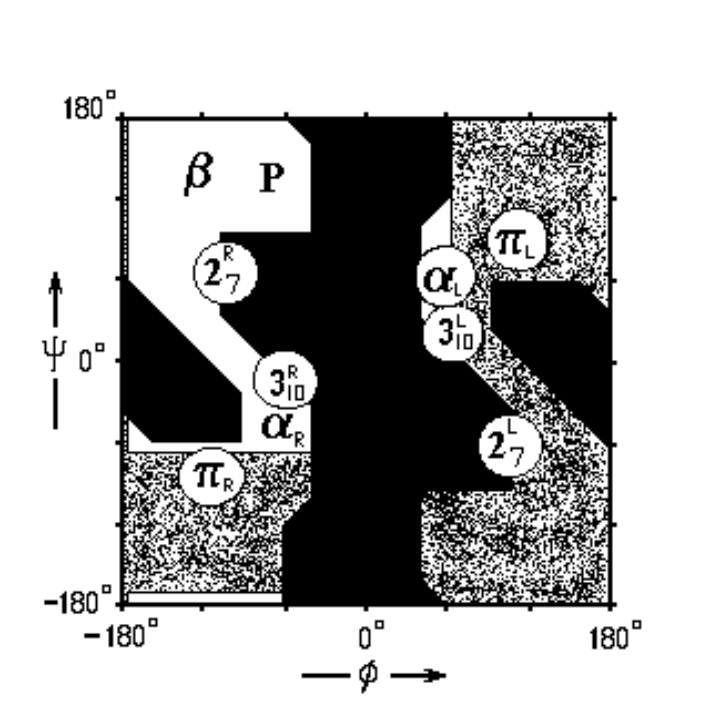
1.8 — 2.1 - hydrogen bond

H-bond energy: 5 kcal/mol

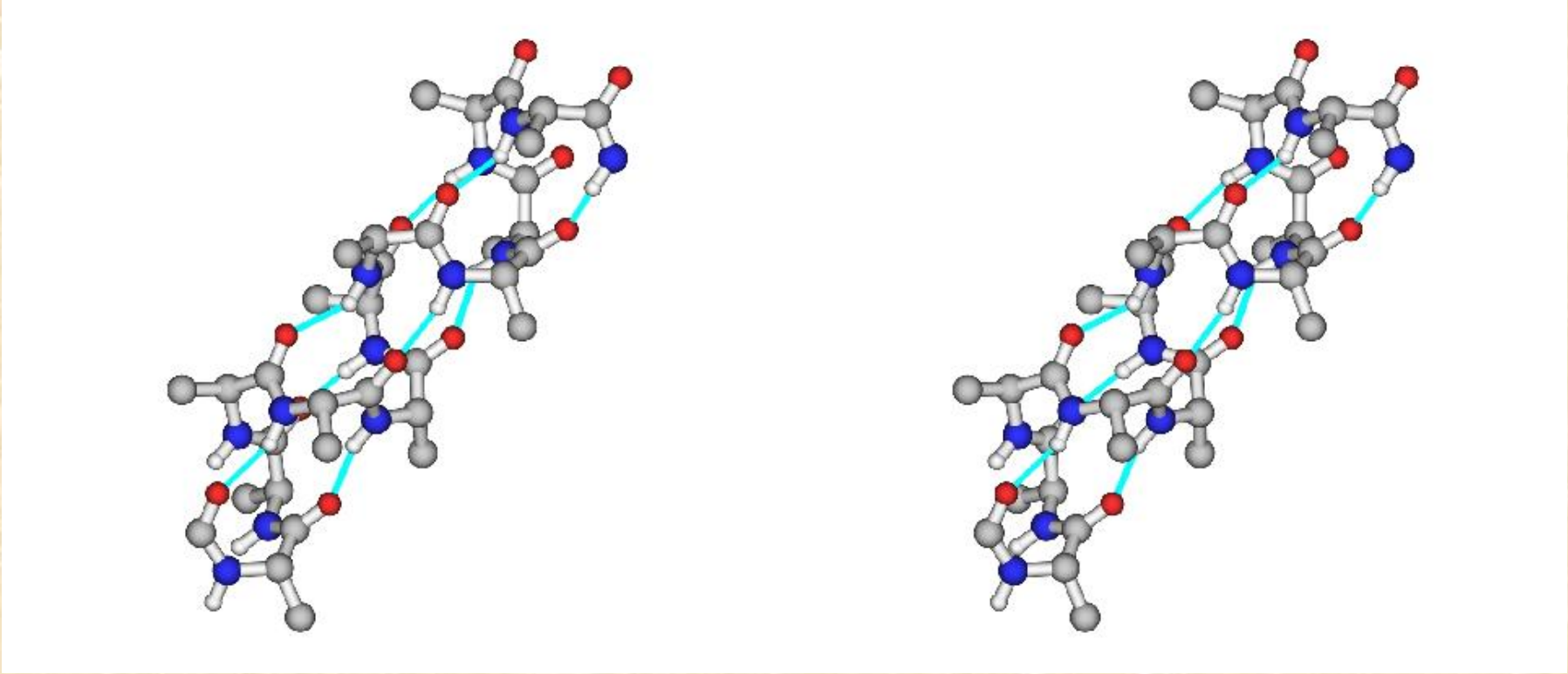
kT at room temperature is 0.6 kcal/mol

Secondary Structure: alpha-helix

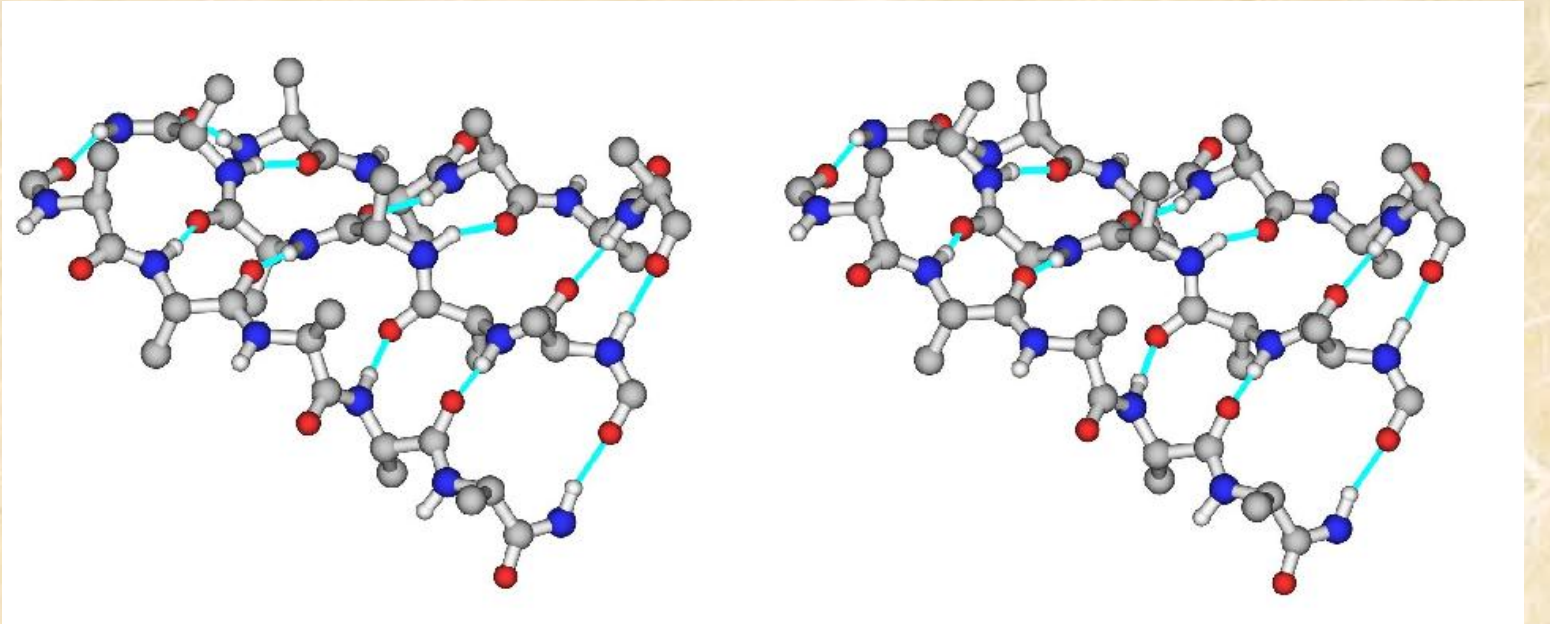


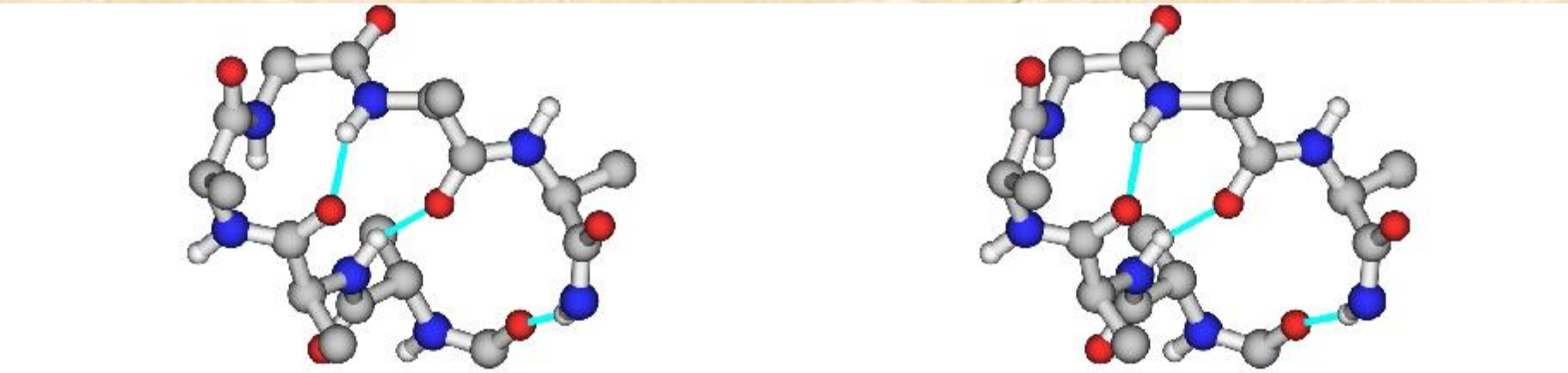


Right alpha-Helix



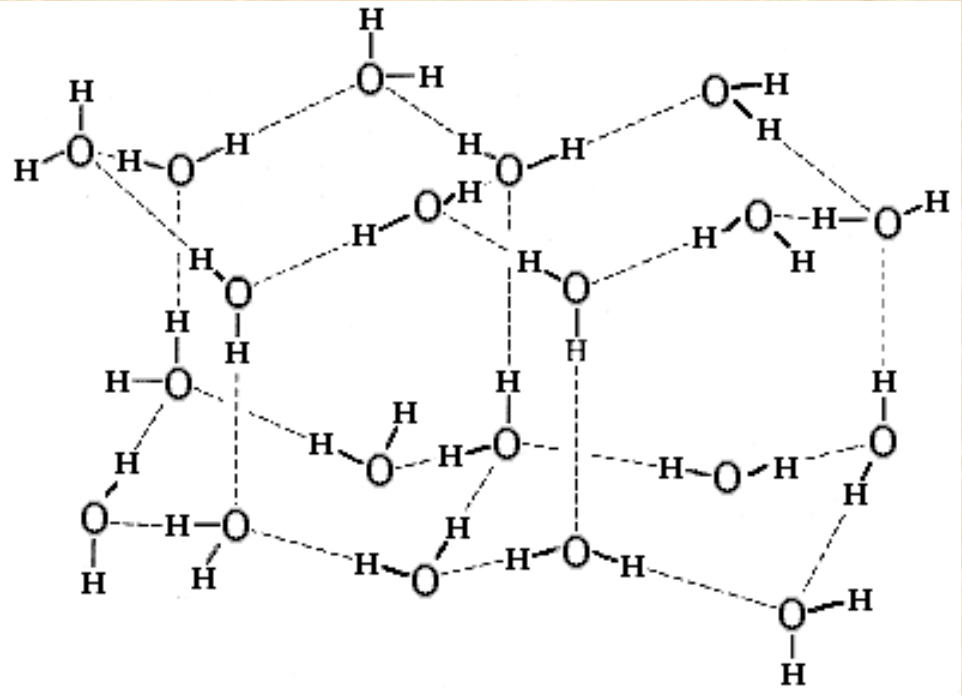
Beta - Structures





Hydrogen Bonds in Water

Structure of ice

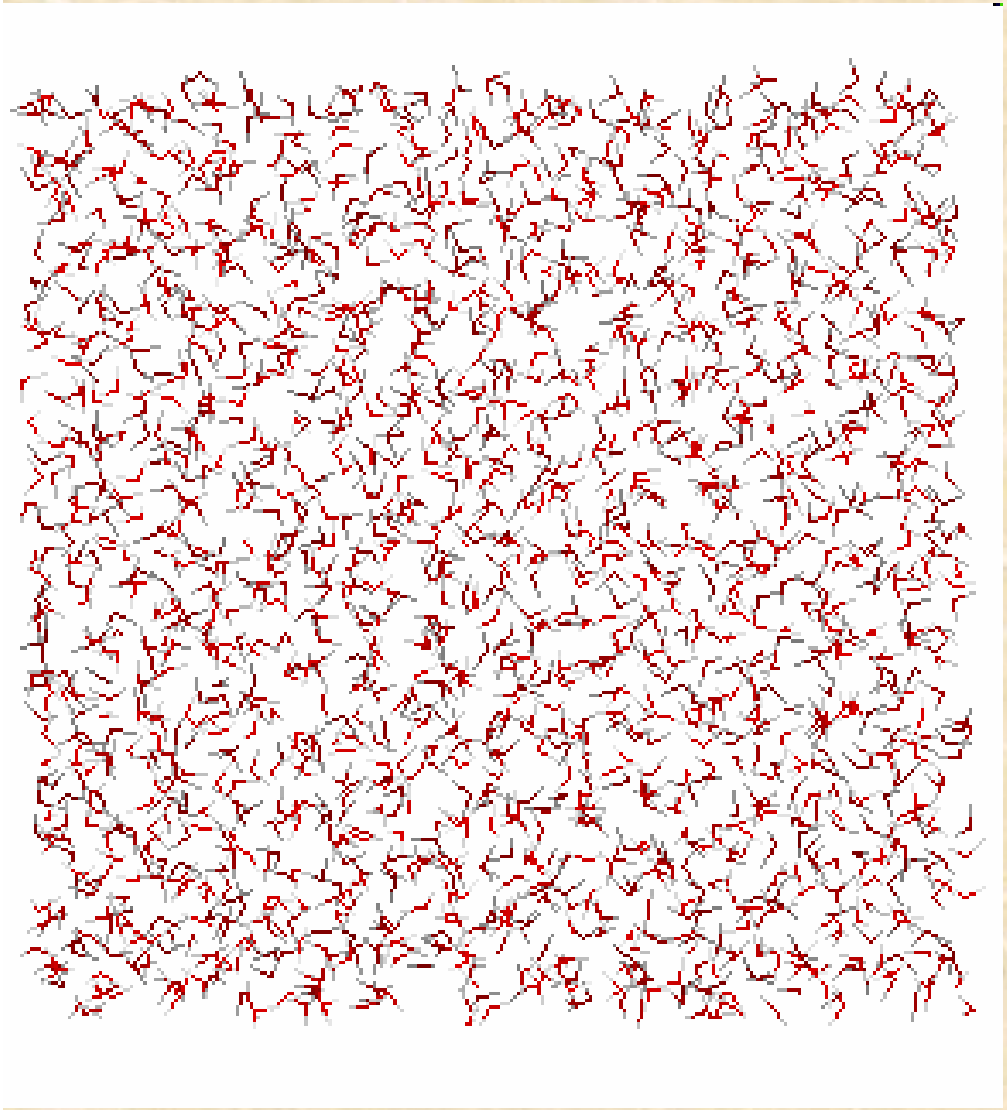


Good for h-bonds
Bad for van der Waals contacts

→

Ice is less dense then water
Easily melts with pressure

Liquid water



Almost all molecules form h-bonds most of the time

but each one is short lived ~ 1-2ps

Molecules are very mobile!

Free Energy

$$F = E - TS$$

Probability ~ $\exp(-F/k_B T) \rightarrow$ max

$F \rightarrow$ min

Entropy $S = k_B \ln$ (#STATES)

Peptide H-Bond is weaker in water then in vacuum due to entropy gain

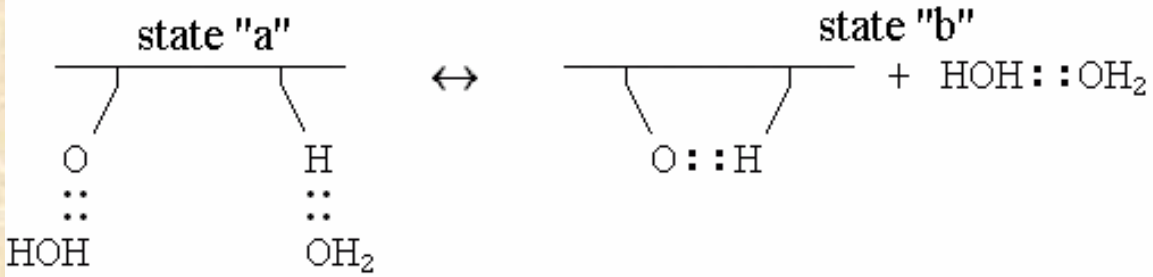
In vacuum:

$E_a=0,\ S_a=0$ $\qquad$ $E_6=E_H,\ S_6=0$

$E_6-E_a=E_H,\quad S_6-S_a=0$

$F_6-F_a=(E_6-E_a)-T(S_6-S_a)=E_H$

In water:



$E_a=2E_H,\ S_a=0$ $\qquad$ $E_6=2E_H,\ S_6=S_H$

$E_6-E_a=0,\quad S_6-S_a=S_H$

$F_6-F_a=(E_6-E_a)-T(S_6-S_a)=-TS_H$

h-bond is stable if $E_H < -TS$ $\qquad$ *water is liquid not gas*

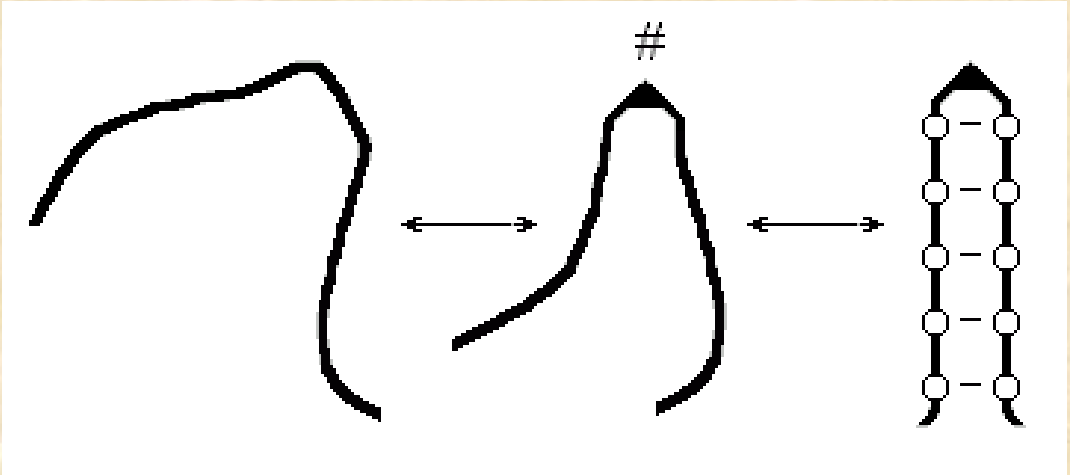
Energetics and Kinetics of Secondary structure formation: Helix-Coil

- Free Energy of formation of an h-bond in alpha-helix is -2 kcal/mol

  (Loss of Entropy is ~2 kcal/mol)
- Free energy of helicity of Ala ~ -0.4 kcal/mol
- Free energy of helicity of Gly ~+1 kcal.mol
- Rate: nanoseconds/residue
- **coil <-> helix**

not a phase transition but a cooperative transition

Beta Structures in Polypeptides – form slowly



2D structure – phase transition

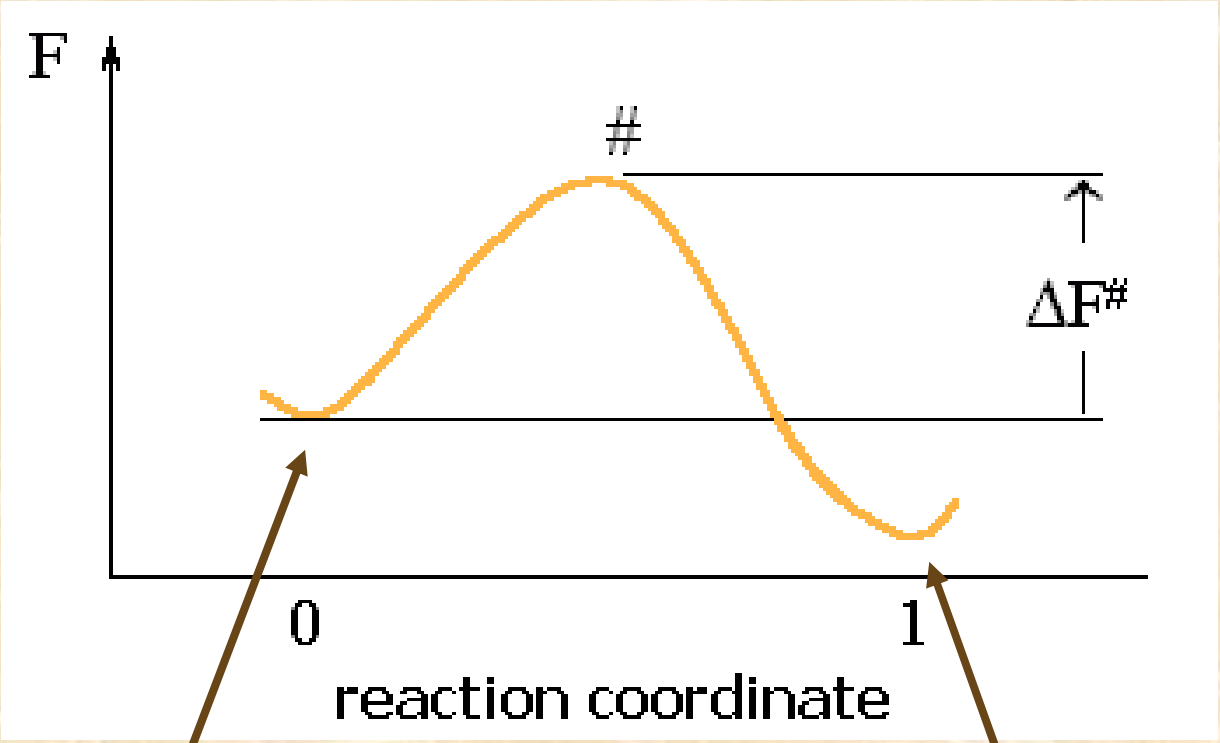
$t_{\beta\text{-HAIRPIN}} \sim 3000$ ns

In globular Proteins - fast

Stable – fast

Non-stable – slow – *high activation free energy*

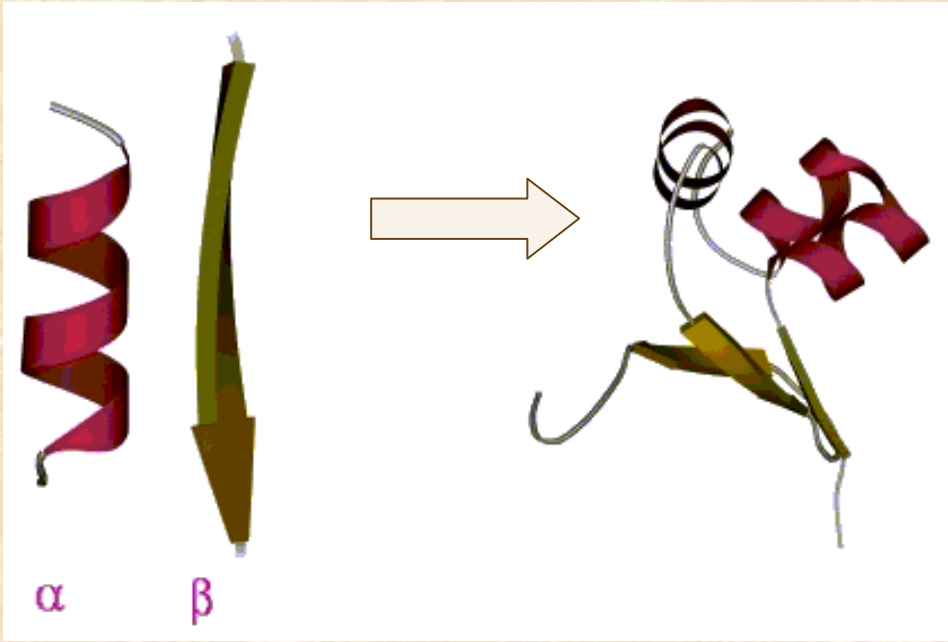
Activation Free Energy



coil

β-pin

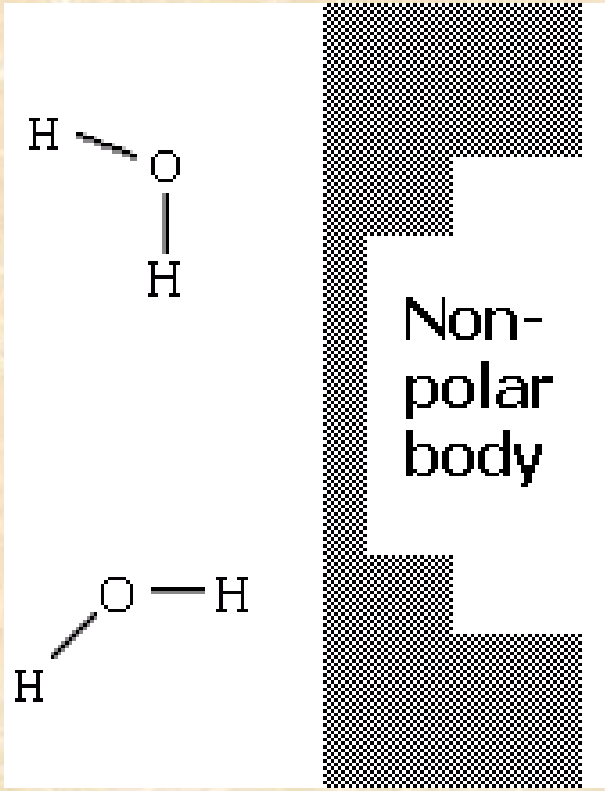
Formation of Tertiary structure



Interaction of Protein and Water
Hydrophobic Effect

- ¾ of work for protein folding is due to the hydrophobic effect – water is bad solvent

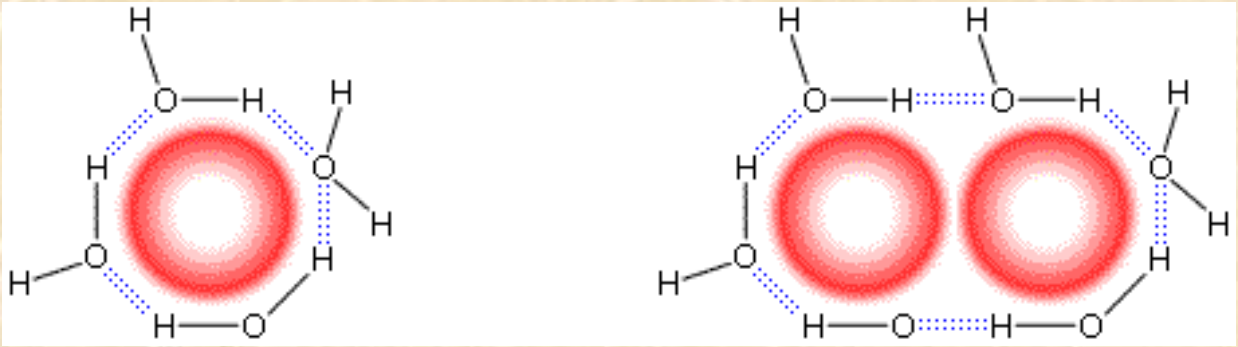
overall collapse ~100 kcal/mol



Hydrophobic Collapse in Globular Proteins

- Non-specific
- Hydrophobic penalty in large molecules ~

~ *surface area*



Forms non-crystal, non-rigid, unspecific globule

Final Crystal Structure

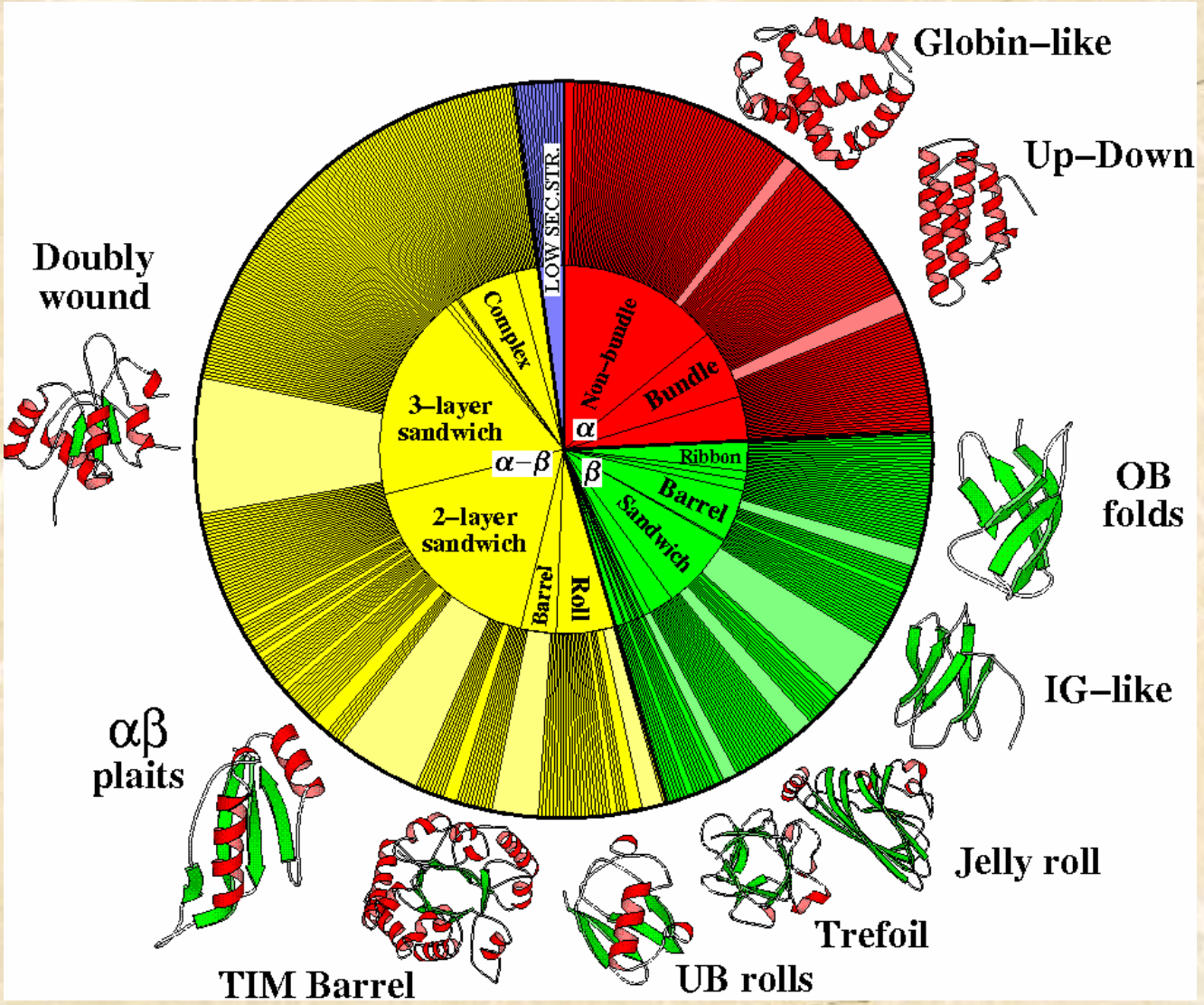
- Fine tuning by van der Waals interactions – packing
- Hydrogen bonds formed inside globule (expensive to loose 5 kcal/mol)
- Exclusion of water

•Stability difference between

Native state and denatured protein

~10 kcal/mol

Typical Globular Architectures



Main Principles of Globular protein Structure

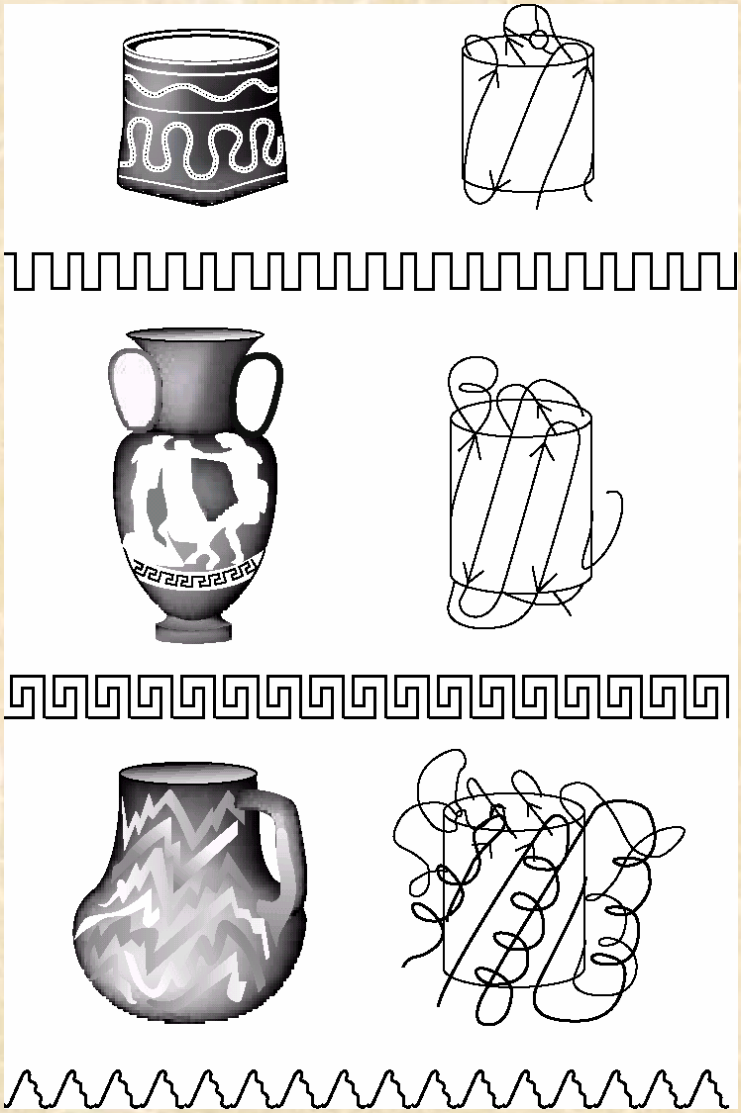
- Quasi-random sequence

1:1 distribution of *polar : hydrophobic* residues

- Secondary structure elements form hydrophobic nucleus – all hydrogen bonds are satisfied internally

- Loops are mostly on the surface –
  1) flexible extended – minimum entropy penalty
  2) hydrogen bonds with water

Typical Folding Patterns



Non-selfcrossing chain threading –

The tightest packing with no energetic penalty?

J. Richardson, 1977

How Unique are Protein Sequences?

1 0F ~10^9 RANDOM SEQUENCES MAKES A "PROTEIN-LIKE" STRUCTURE

(SOLID, WITH A SPECIFIC BINDING).

Where to look for Proteins?

- The UniProtKB/Swiss-Prot
  http://www.ebi.ac.uk/swissprot/

an annotated protein sequence database

now: 269293 **sequence entries**

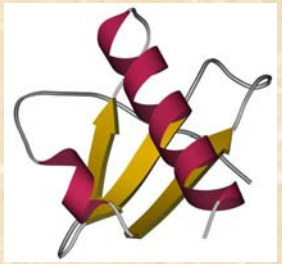
- Protein Data Bank (PDB)

  http://www.rcsb.org - a protein structure database obtained by x-ray and NMR

now: 43755 **3D structures**

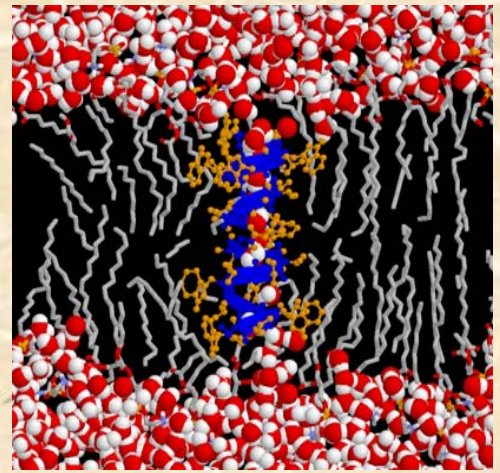
Sequence Dictates Structure in a given environment

- **Globular Proteins** – water soluble



- **Membrane Proteins** –

associated with lipid membrane

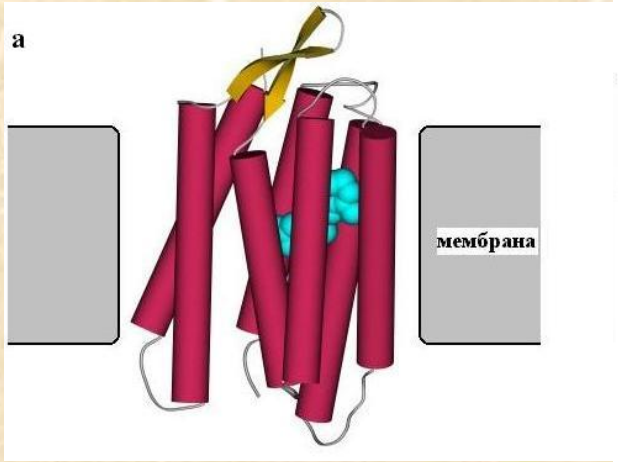


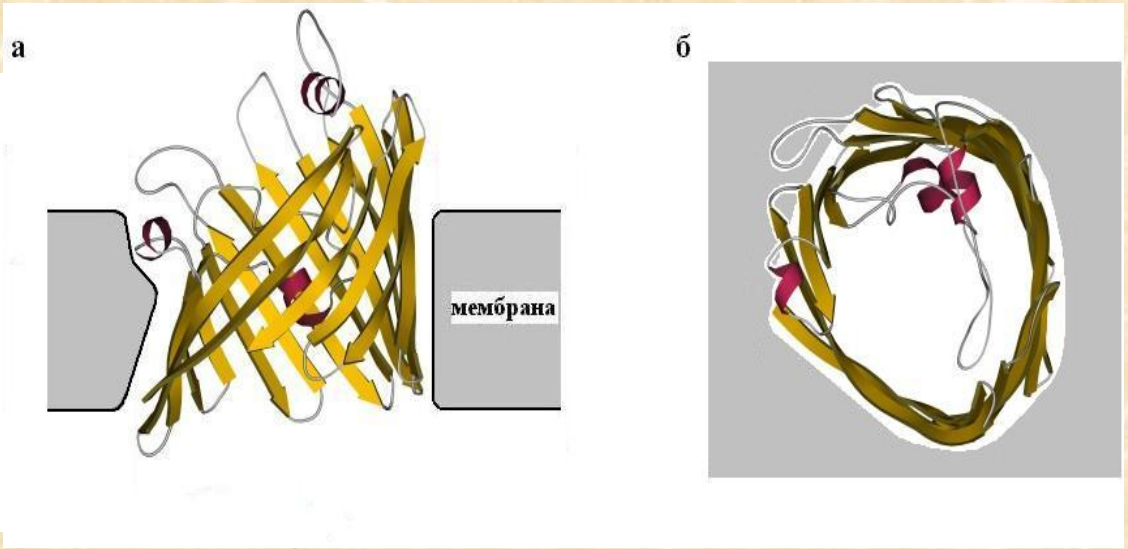
- **Fibrous proteins** –

large aggregates of regular structures

Architechture of Membrane Proteins

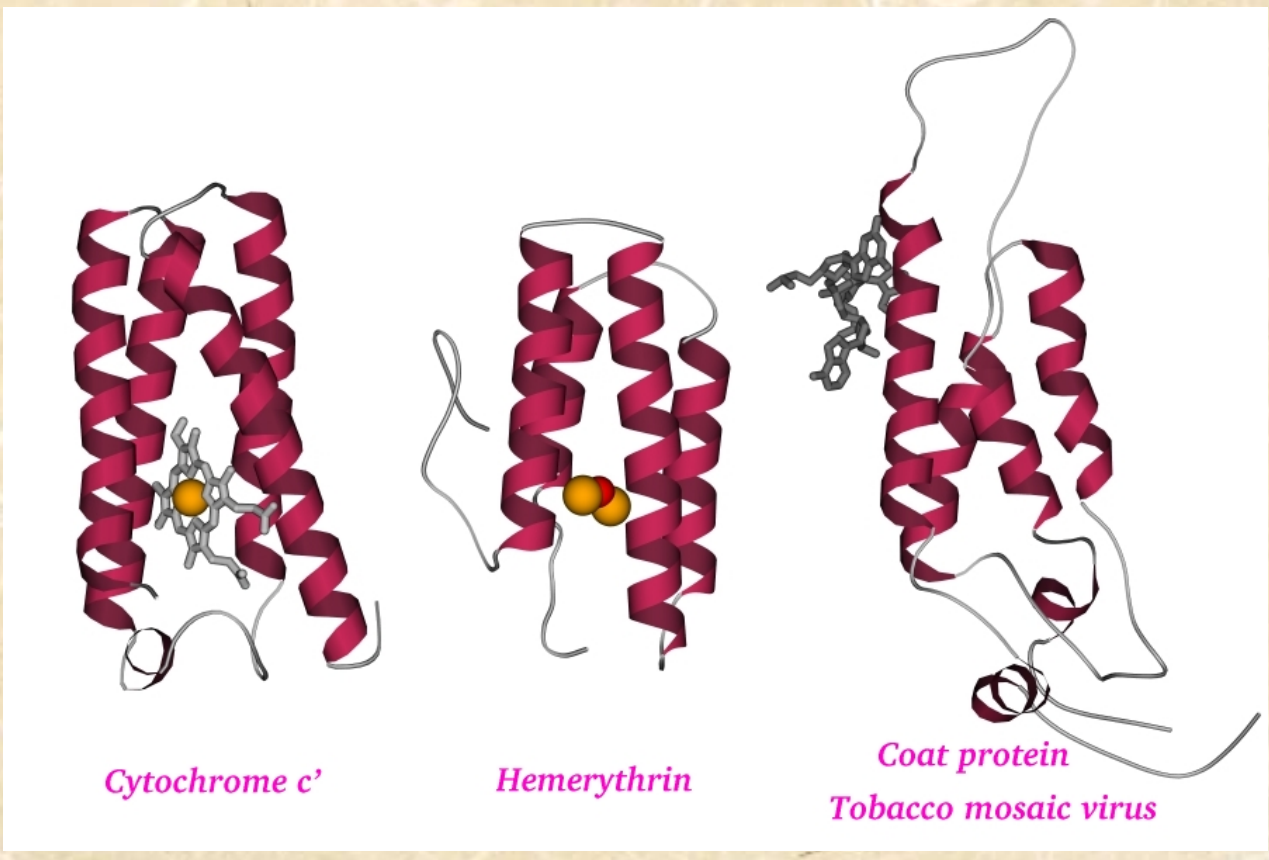
- Not quasi-random – blocks of hydrophobic and quasi-random sequences
- hydrophobic blocks form transmembrane structures





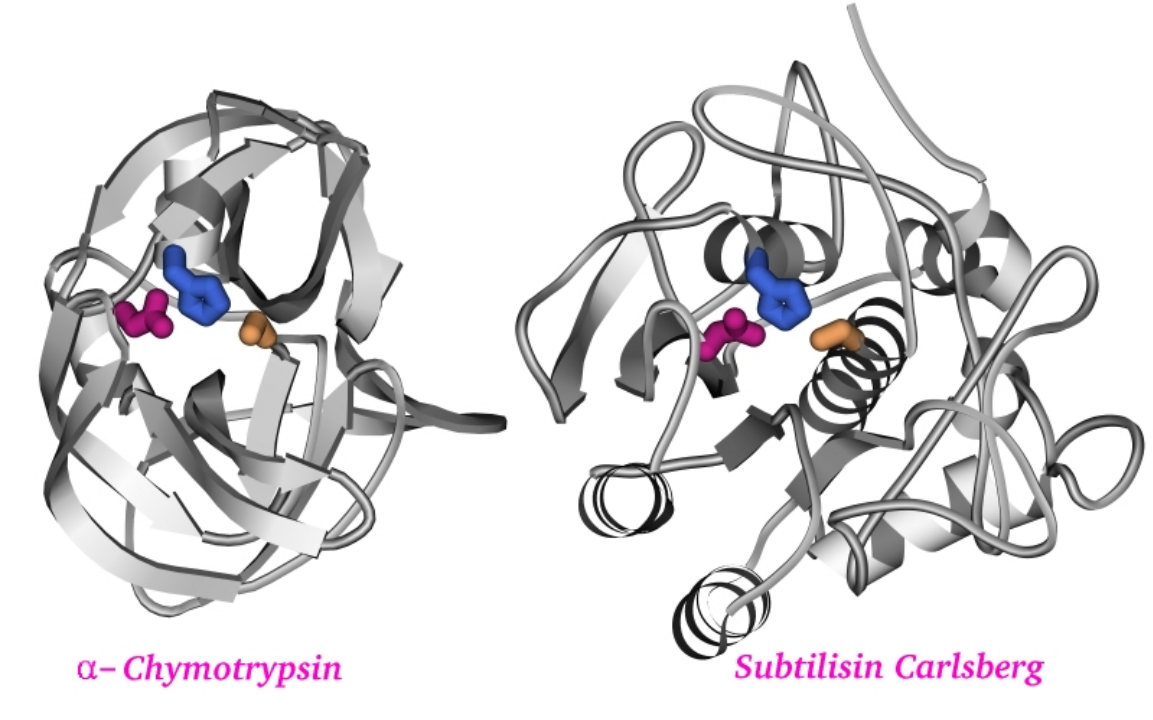
Structure – Function Relations

- Typically shape – defines structure
- Sometimes similar structures - different function



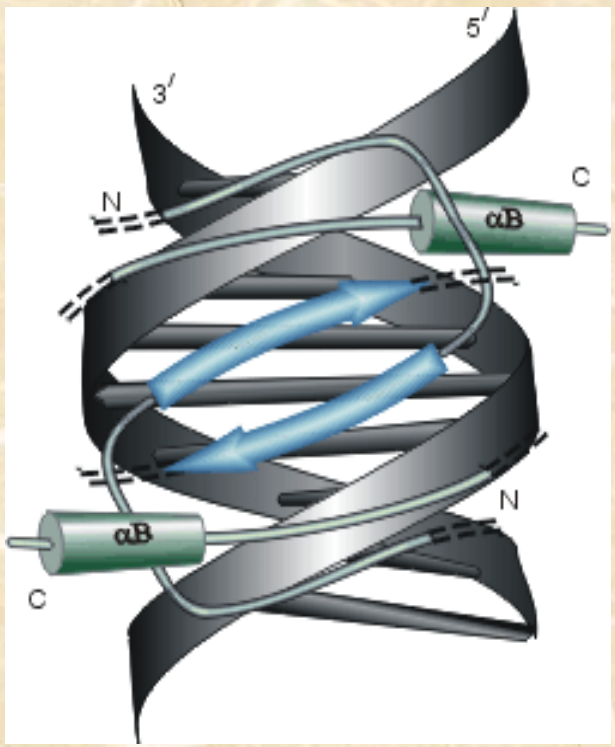
Structure – Function Relations

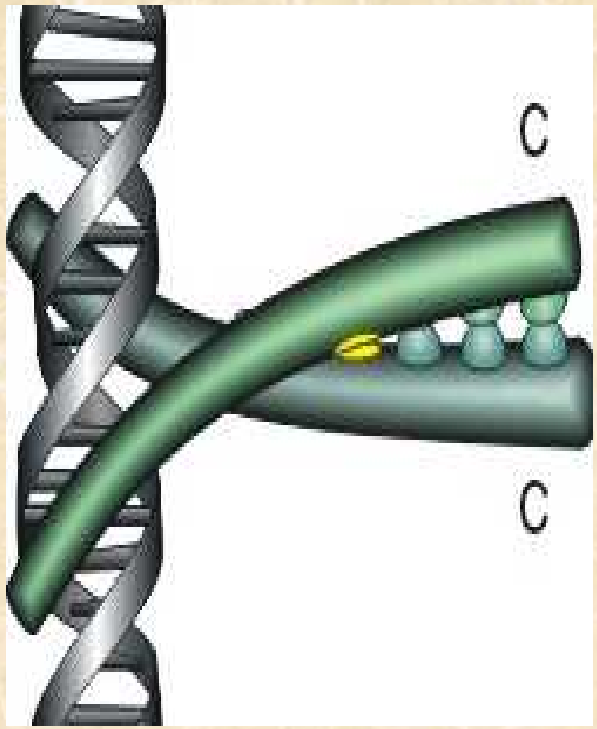
- Typically shape – defines structure
- Sometimes similar function – different structures



PROTEINS AT ACTION:
BIND → TRANSFORM → RELEASE

- BINDING





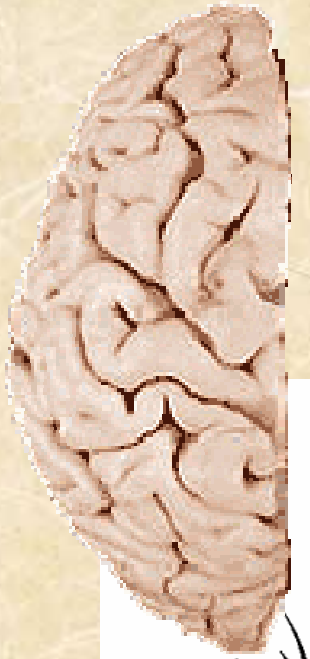
PROTEINS AT ACTION: BIND → TRANSFORM → RELEASE

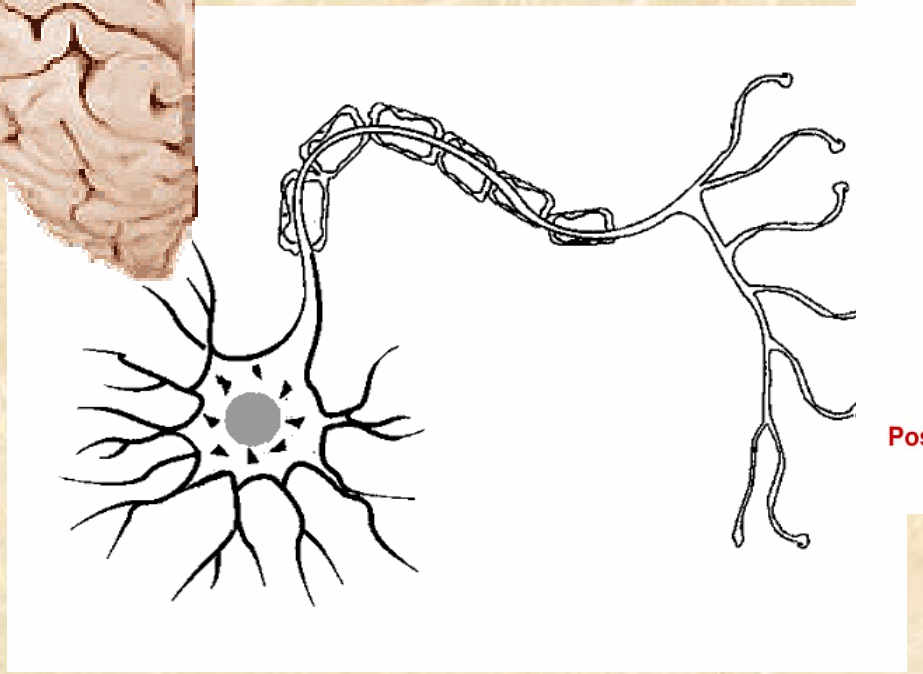
A Ligand-gated ion channel:

Multi-subunit, multi-domain membrane-associated protein

Function: facilitates excitatory transmission between neuronal cells

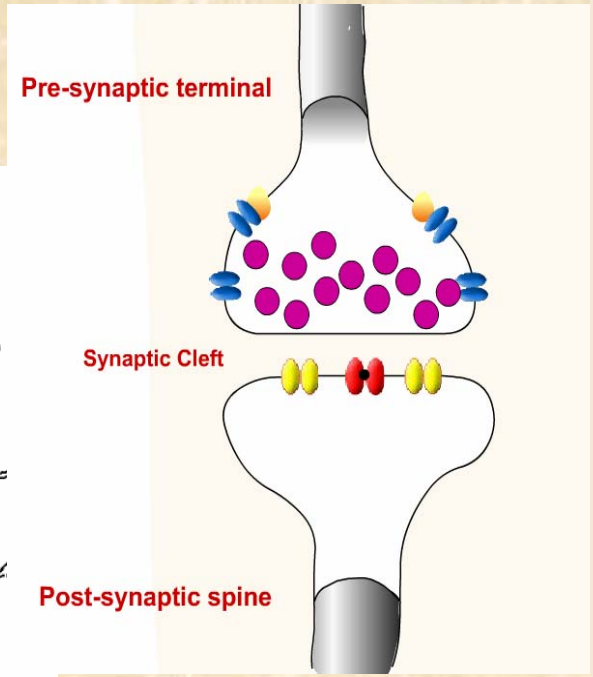
Glutamate receptors are found in brain



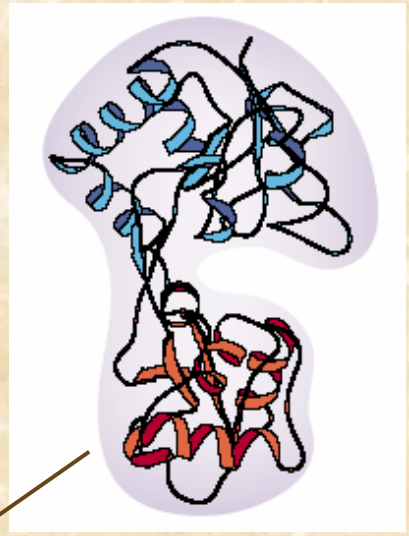


neuron

synapse

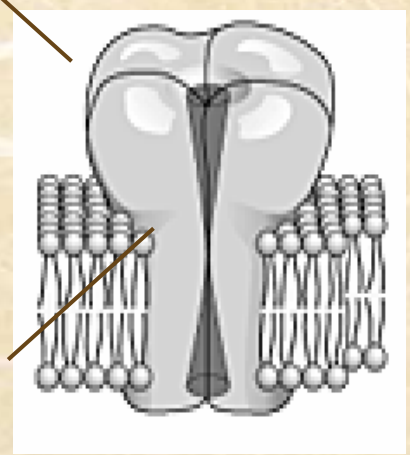


Channel Opening

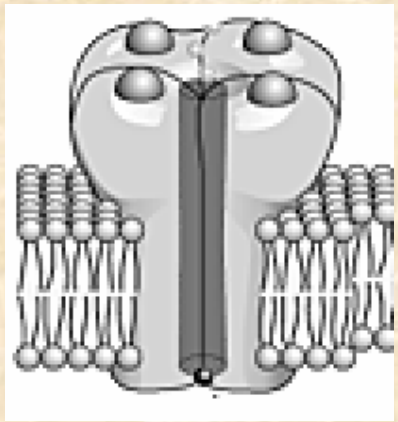


Binding core

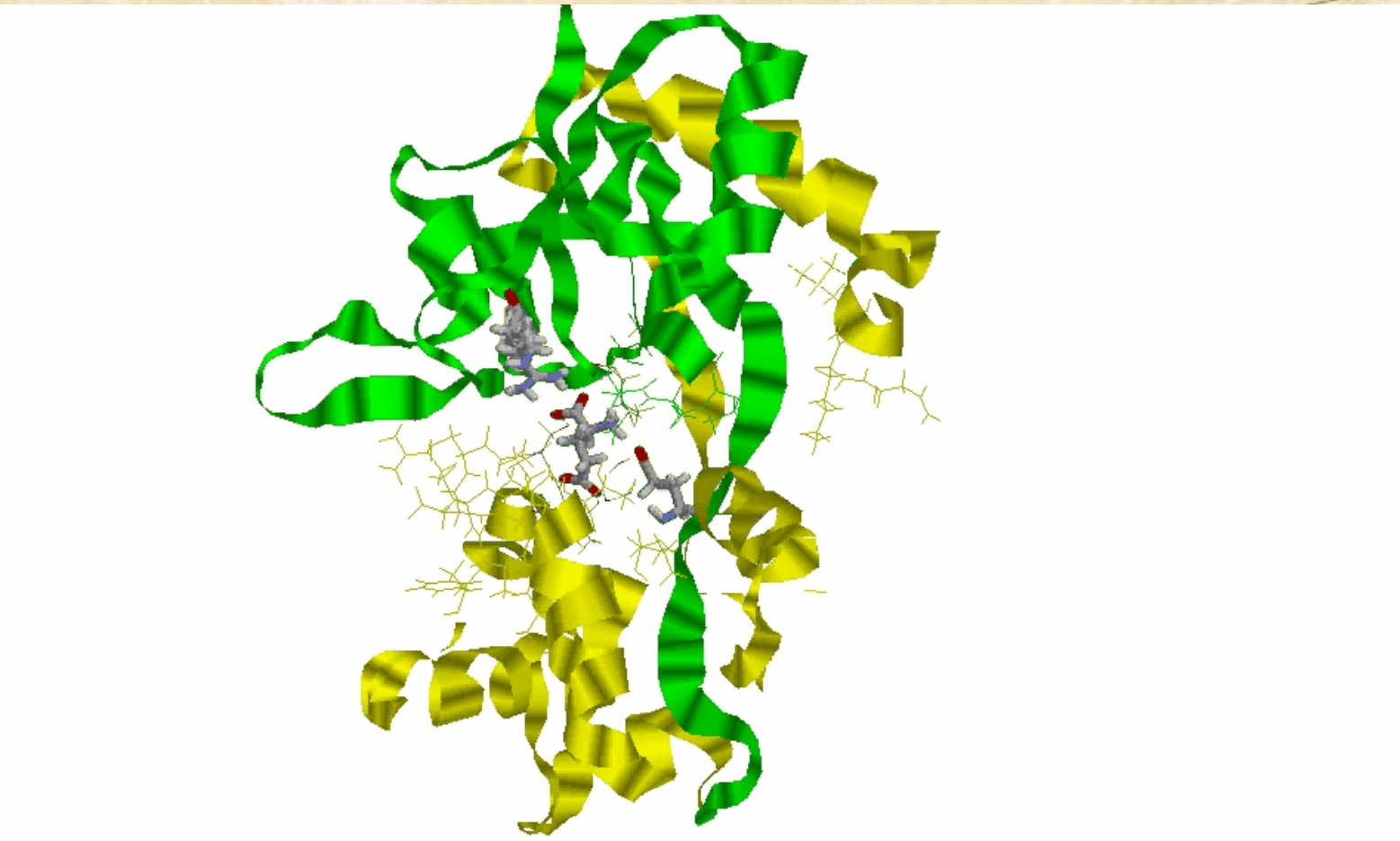
Glutamate



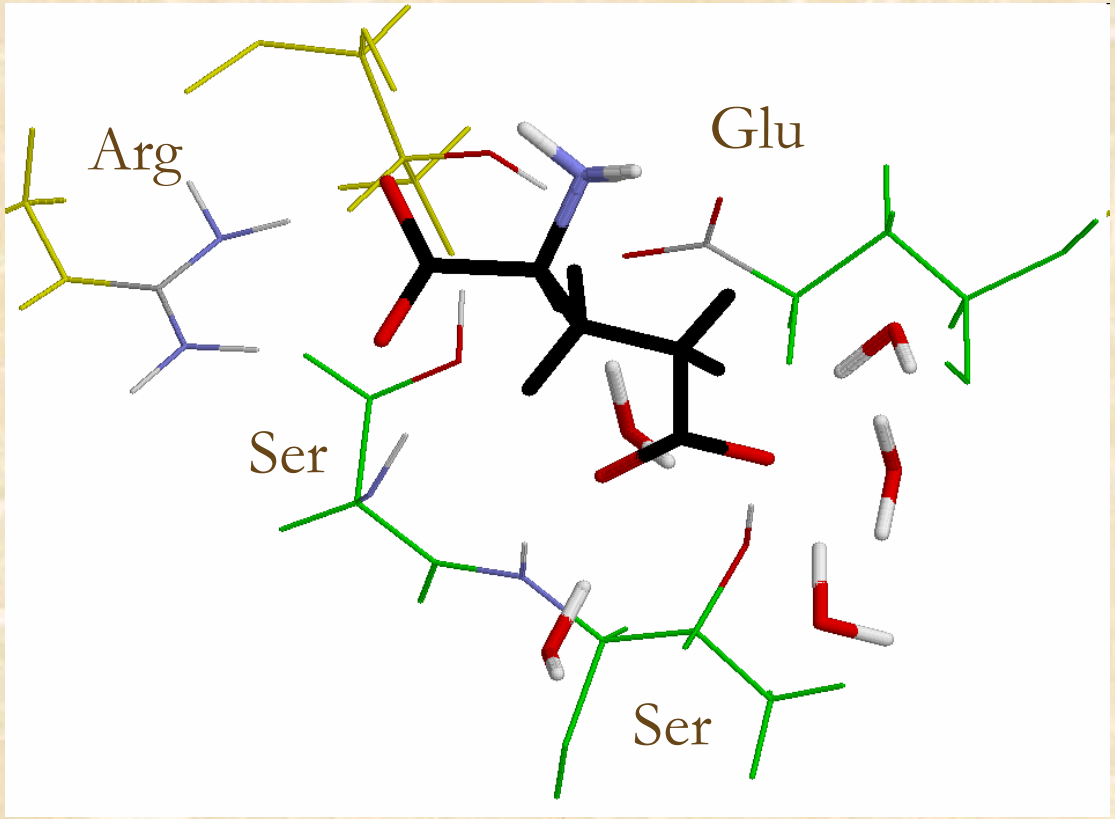
Transmembrane channel



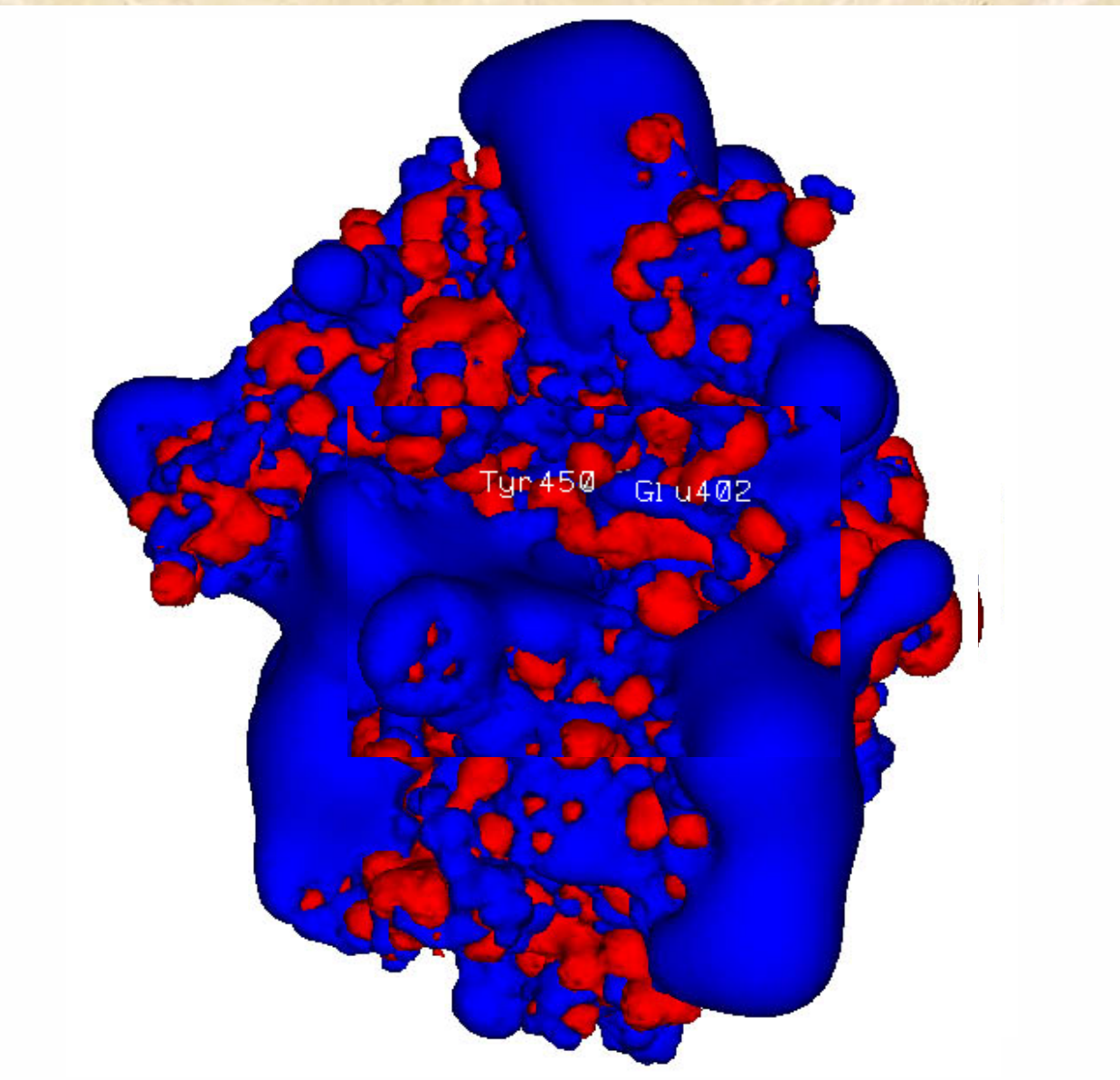
Structure of the Glutamate Binding Site



Binding Site structure

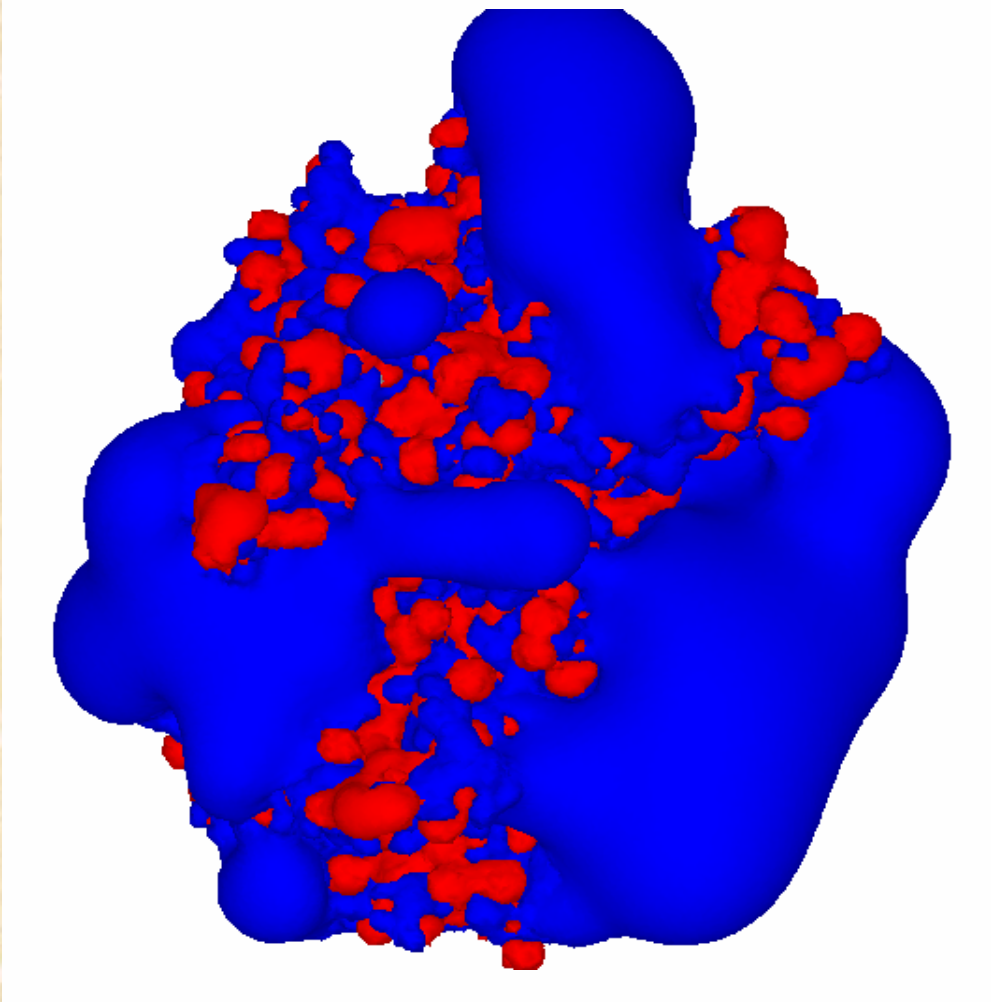


Electrostatic Potential Apo-Glur2



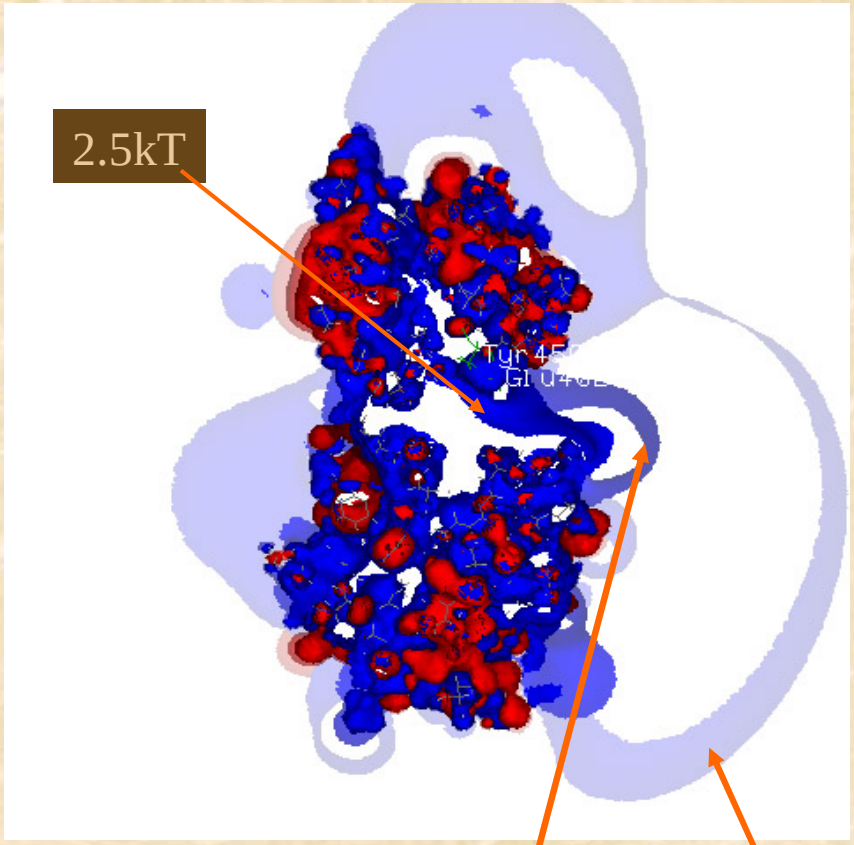
Electrostatic Potential.

GluR2 with the Glutamate bound

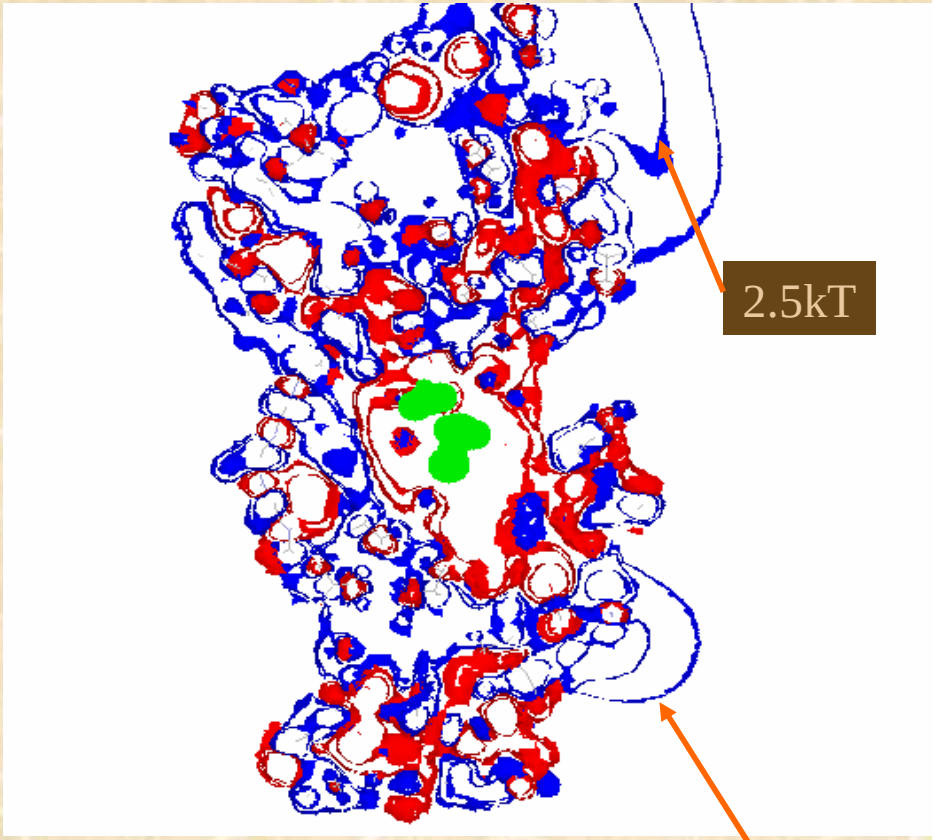


Slice through the center of the protein

Apo



Glutamate-bound

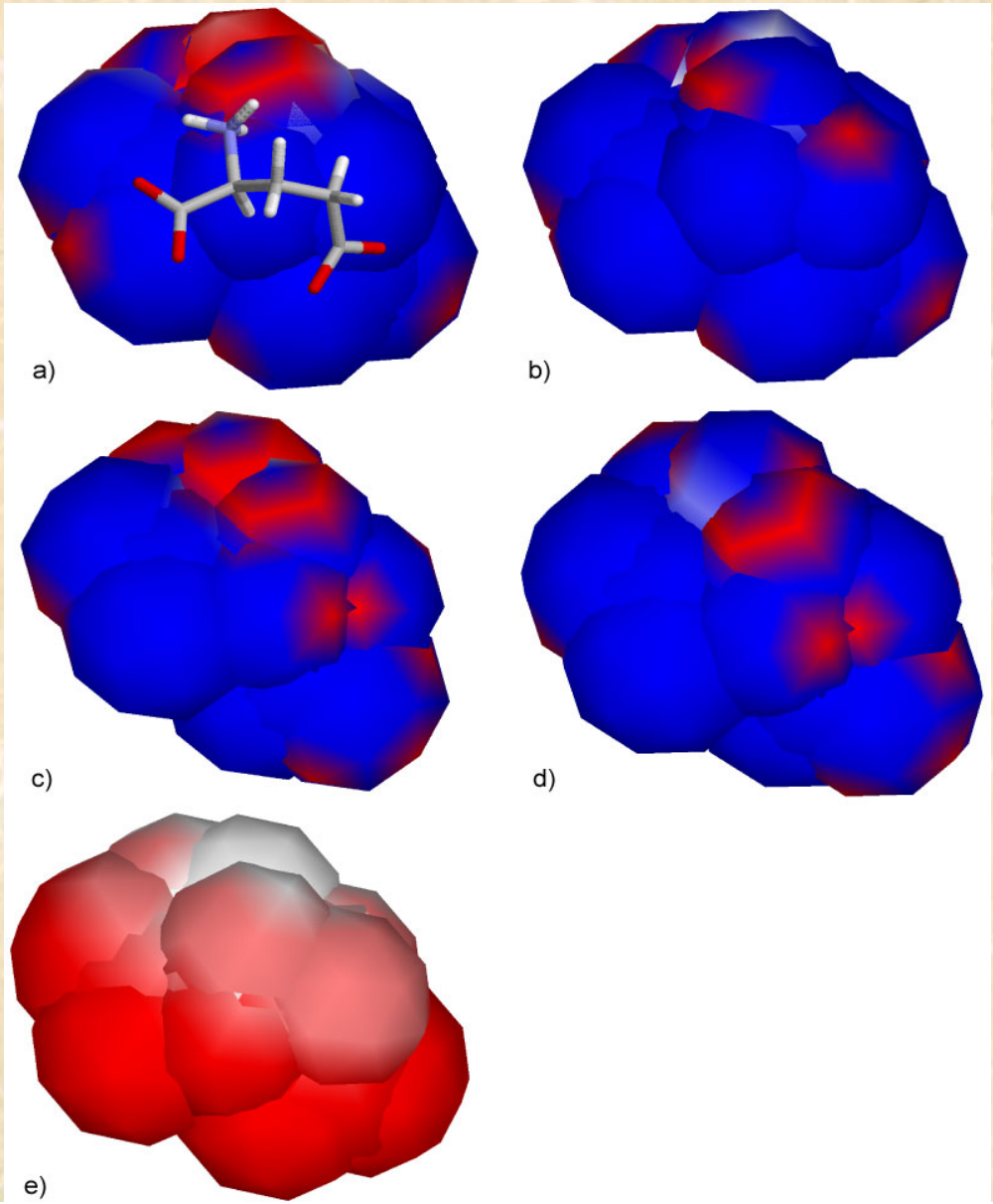


Ligand-Protein Electrostatic Complementarity

Wild Type

Mutant

Glutamate in water



No Glu705

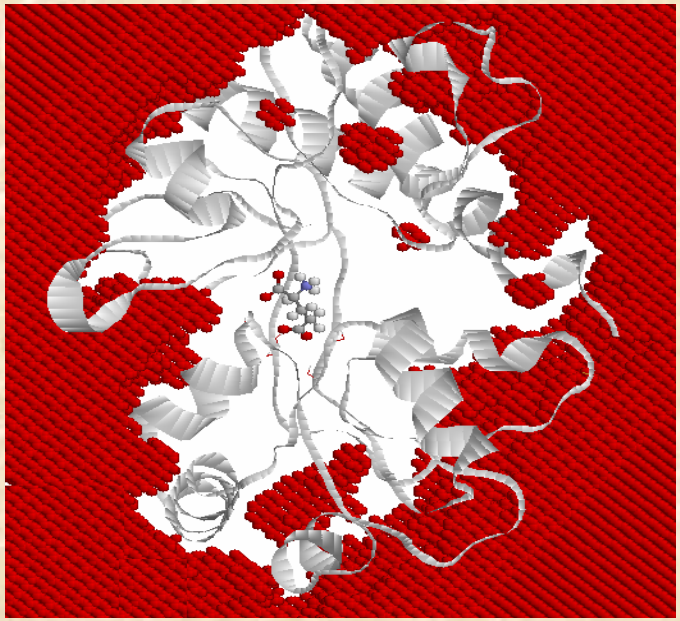
No Asp705

Free Energy of Ligand Binding with Protein in Closed Conformation Using MD/PBSA

Binding energy is calculated as:

$$\left\langle \Delta G_{\text{int}} + \Delta G_{PB} + \Delta G_{SASA} - T\Delta S \right\rangle_0$$

The dielectric map of GluR2

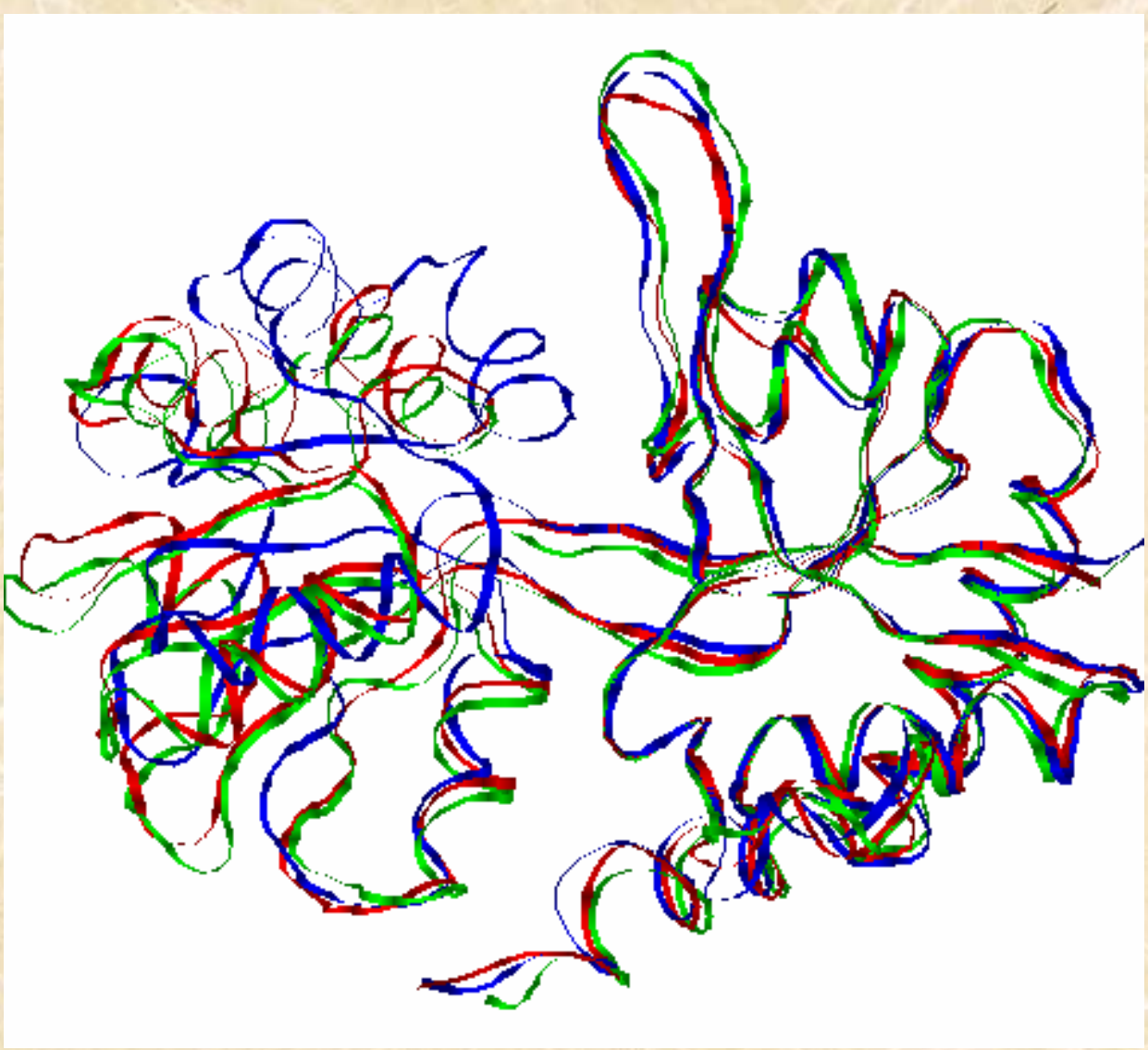


Glu- GluR2 binding energy

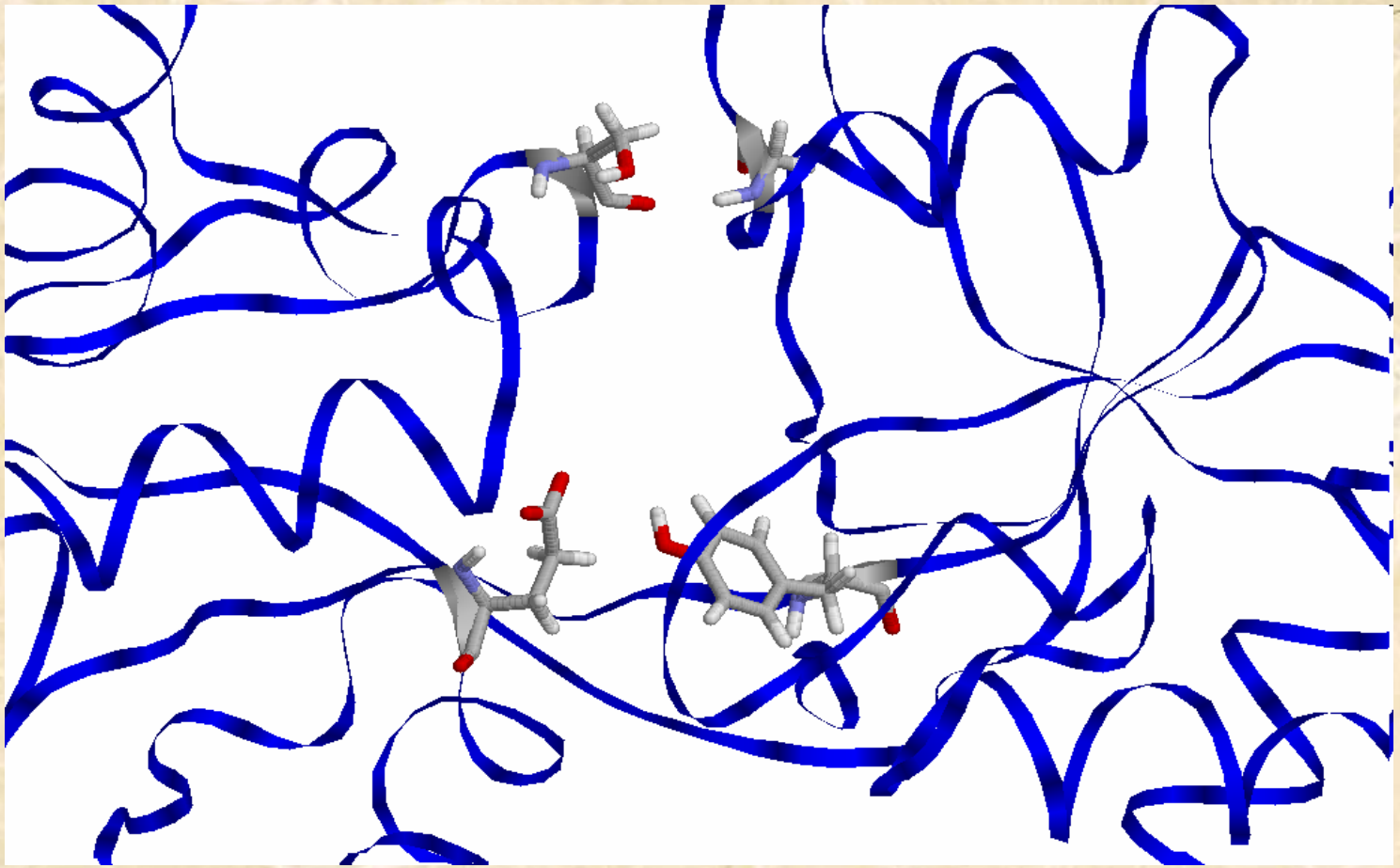
	WT	E705D
Exp. GluR4 (*Abele et al. 2000, JBC 275, 21355)*	-8.53±0.06	-6.14 ±0.17
Calculated GluR2	-13±4	-8±4

K. Speranskiy and M. Kurnikova, Biochemistry 2005

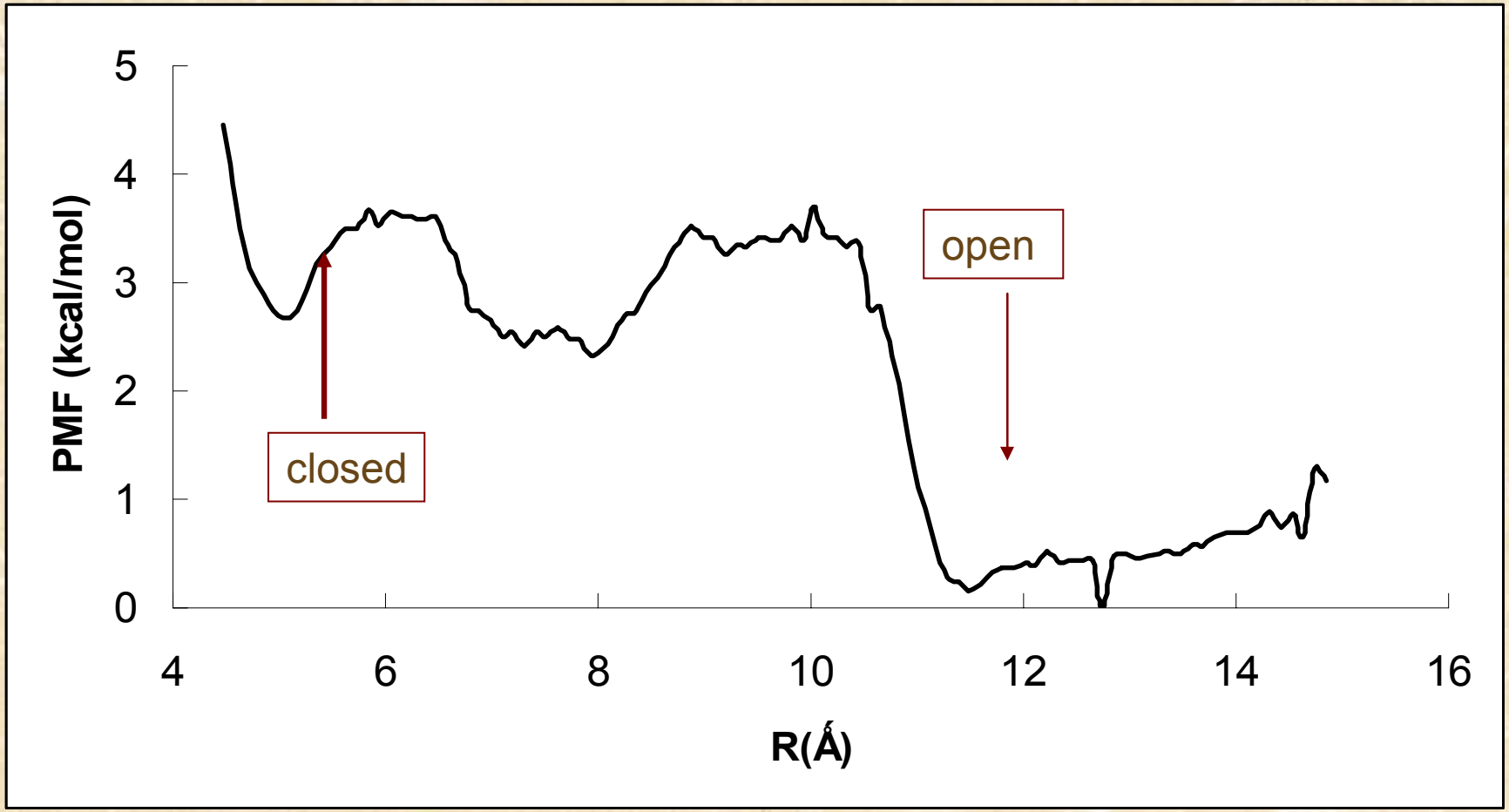
Open-Closed conformation



Lock Mechanism



Free Energy of Conformational Rearrangement of $E705^{-0.75}$ GluR2



Free Energy of Transition Calculated Thermodynamic Integration/ Umbrella Sampling Combination

Transition	Conformation	ΔG, kcal/mol
E705 → $E705^{-0.75}$	closed	33.01
$E705^{-0.75}$ → $E705^{-0.75}$	closed → open	-3.53
$E705^{-0.75}$ → E705	open	- 33.9
TOTAL: WT protein	closed → open	-4.44

References

- ✓ The main source of this presentation have been an excellent book

  by Alexei V. Finkelstein and Oleg Ptitsyn from the famous Protein Institute of Russian Academy of Science (http://phys.protres.ru/)

  Protein Physics : A Course of Lectures *(Academic Press, 2002)*

  based on lecture notes for a course taught in late eighties at my alma-mater Moscow PhysTech

(Or if you can read Russian you can view an electronic version of this course at http://phys.protres.ru/lectures/protein_physics/)

- ✓ A slide set of lectures given by A. Finkelstein at HHMI

  http://phys.protres.ru/lectures/slides/slides.html

- ✓ Wikipedia – is a great source of knowledge about everything!

Acknowledgments

Carnegie Mellon
- Dr. Tatyana Mamonova
- Dr. Kirill Speranskiy

Kirill and Tatyana have done all the work on the Glutamate receptor

kurnikova group at http://crete.chem.cmu.edu

Funding:
Research Corporation
NIH
CMU/Pittsburgh LifeSciences Greenhouse foundation
Faculty Development Berkman fund (CMU)
Pittsburgh Supercomputing Center (NSF-PACI grant)

Inquiries are welcome

- My email: kurnikova@cmu.edu – feel free to contact me with questions related or not to theoretical biophysics and our PhD programs in chemistry, biophysics and computational biology:

http://www.biophysics.pitt.edu

http://www.compbio.cmu.edu/

http://www.chem.cmu.edu/grad/index.html

- or drop by my lab for a coffee if you happened to be in Pittsburgh