Protein Physics 2016 Lecture 7, February 9

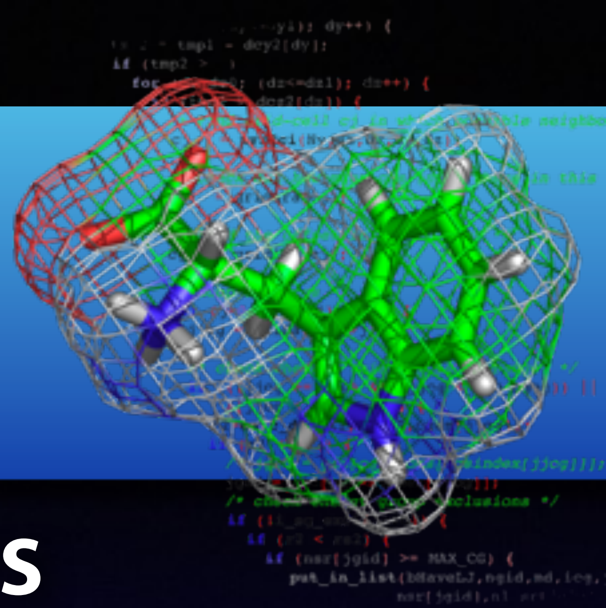
Protein structure: Fibrous proteins & Globular proteins

Magnus Andersson

magnus.andersson@scilifelab.se

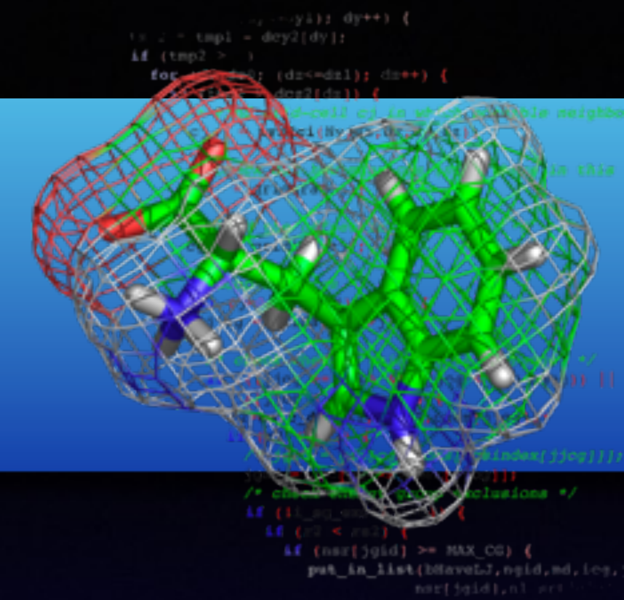
Theoretical & Computational Biophysics

Recap



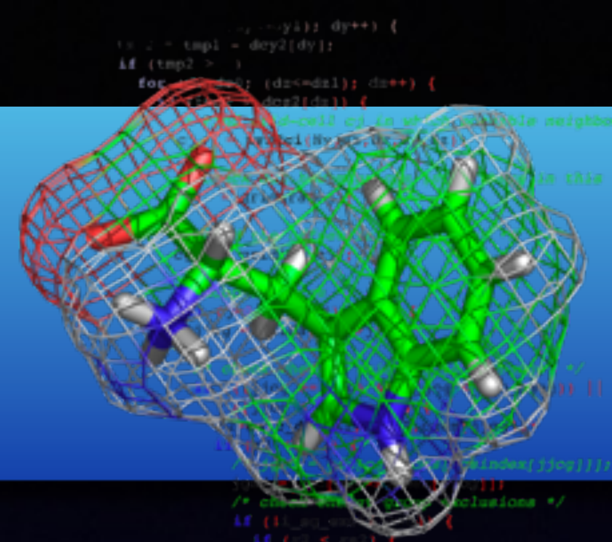
- Secondary structure thermodynamics
- Alpha helices
  - Local interactions, fast folding
  - Energy/entropy balance
  - Helix-coil co-existence
- Beta sheets
  - Non-local interactions, slow folding
  - Natural sizes (length/width) for stability!
  - All-or-none phase transition

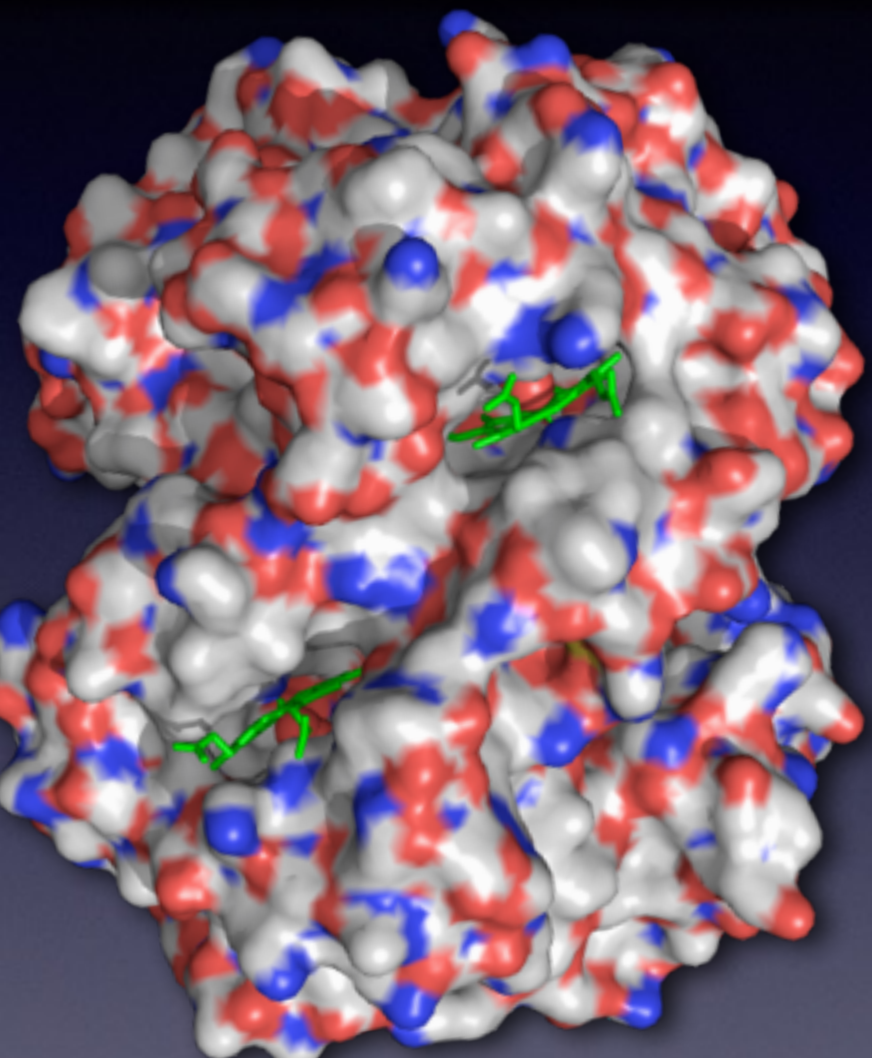
Outline today



- Looking at real proteins
- Complex assemblies from simple α-helix/β-sheet building blocks
- Supersecondary structure & "motifs"
- Fibrous proteins
- Globular proteins
  - α-helix, β-sheet proteins
  - Mixed structures

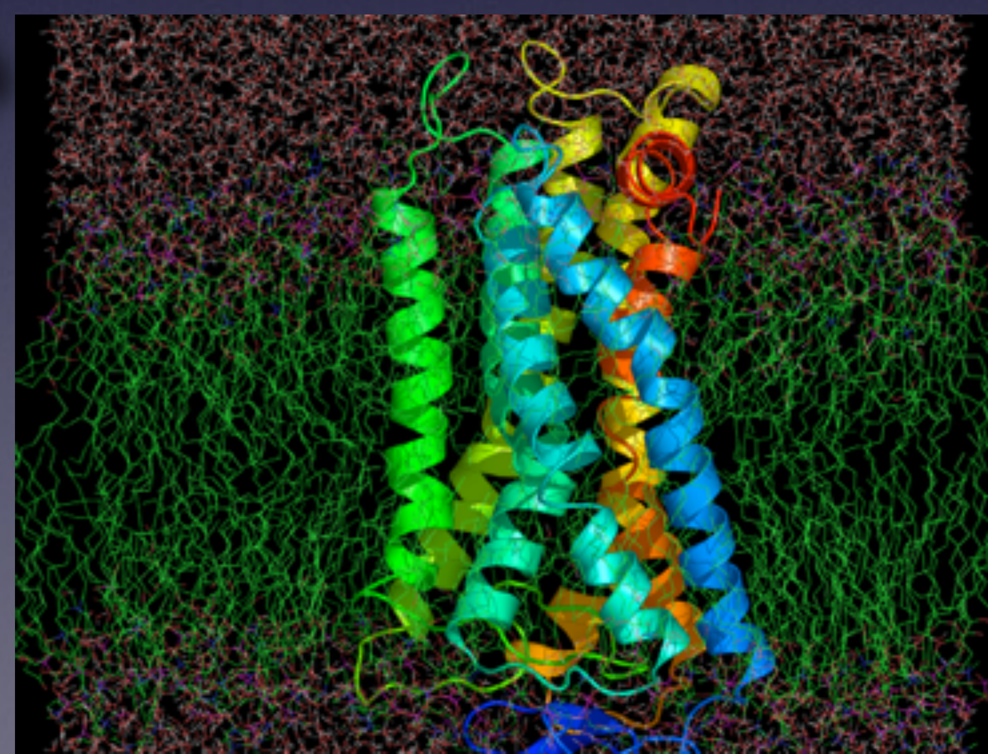
Types of proteins

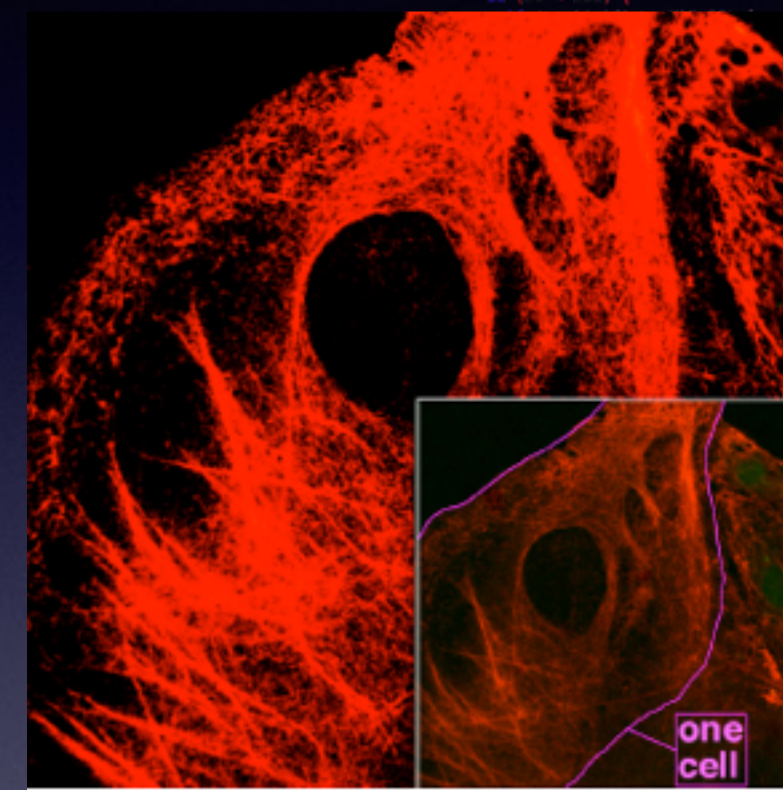




Hemoglobin
Water-soluble
("globular")

Rhodopsin
Membrane protein

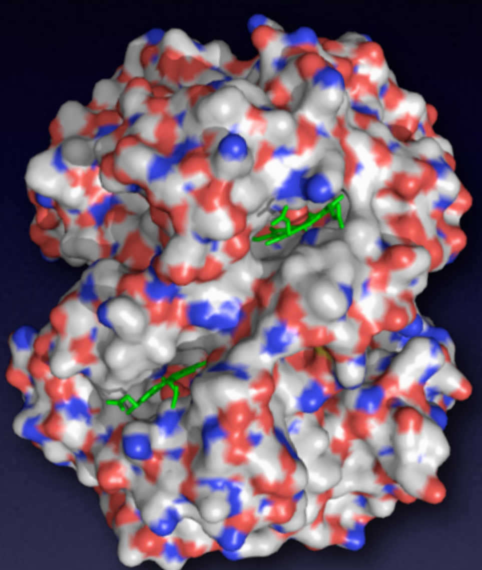


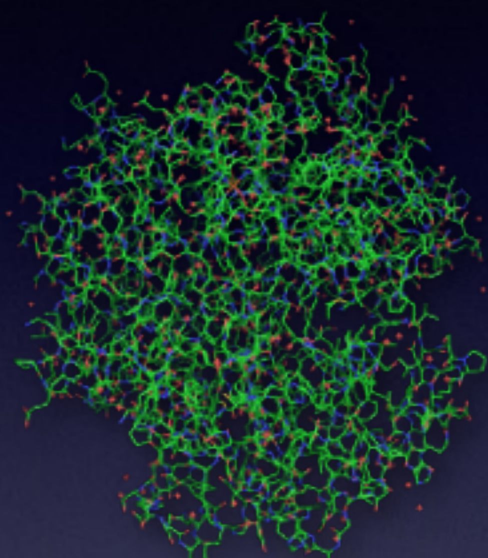


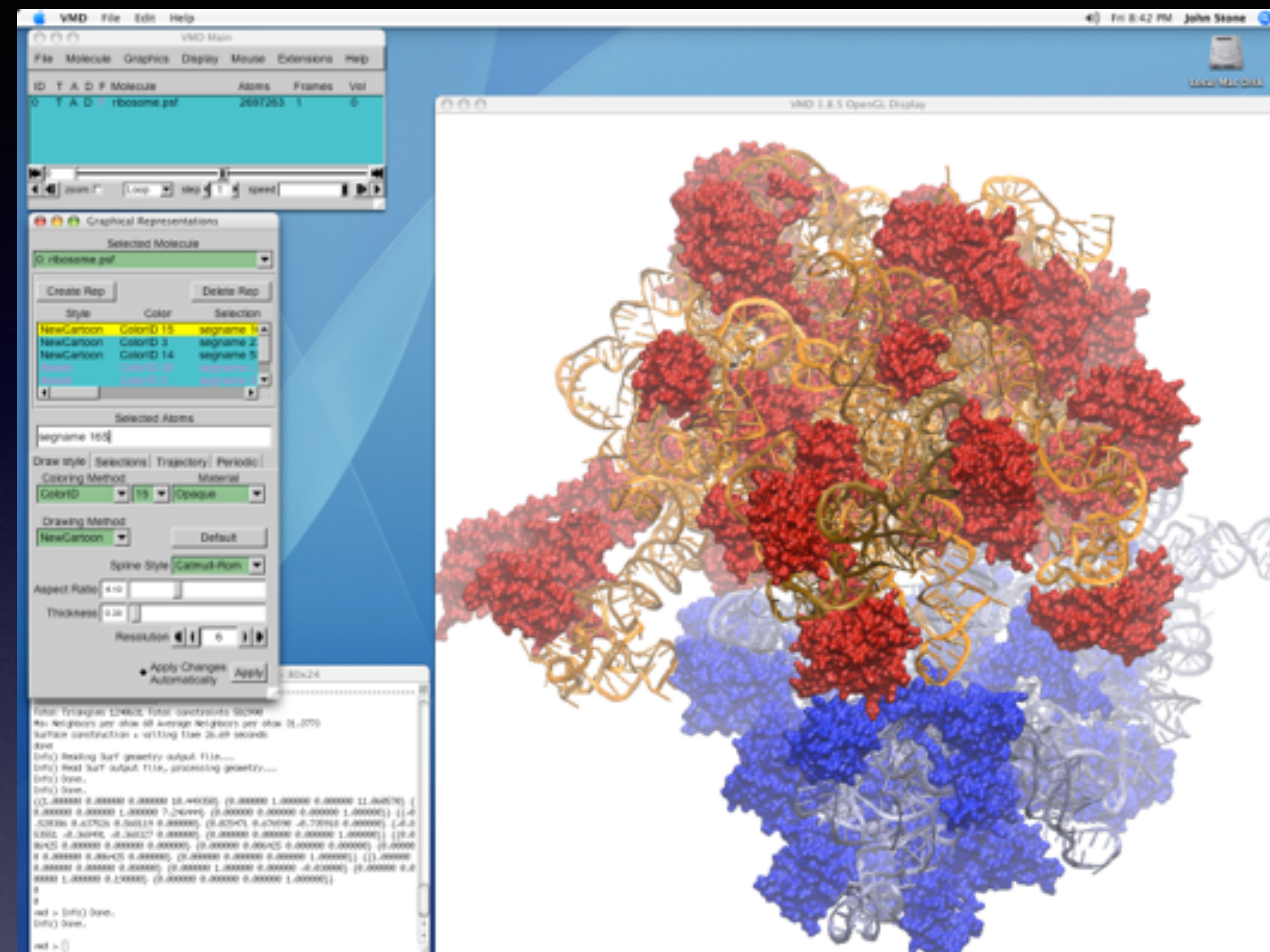
Keratin
Fibrous protein
"Building material"

Tools for studying protein structure: VMD

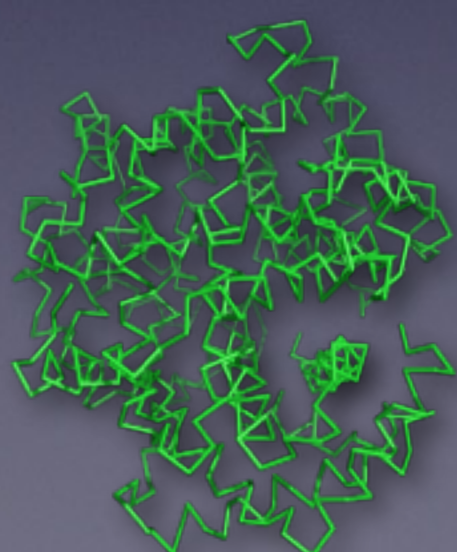
http://www.ks.uiuc.edu/Research/vmd/

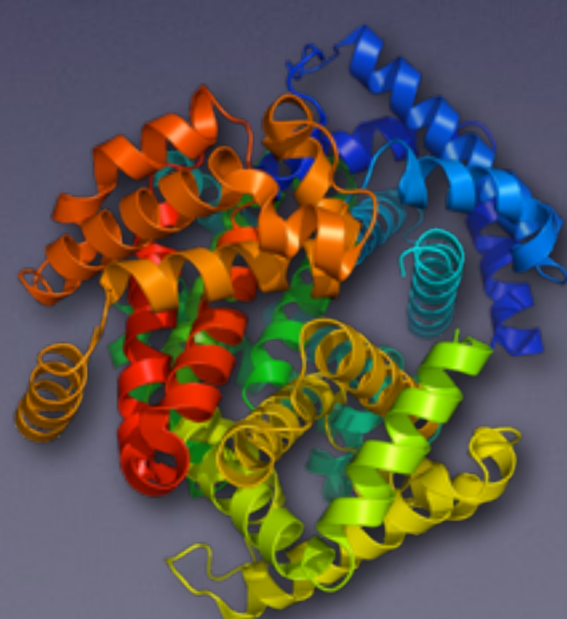


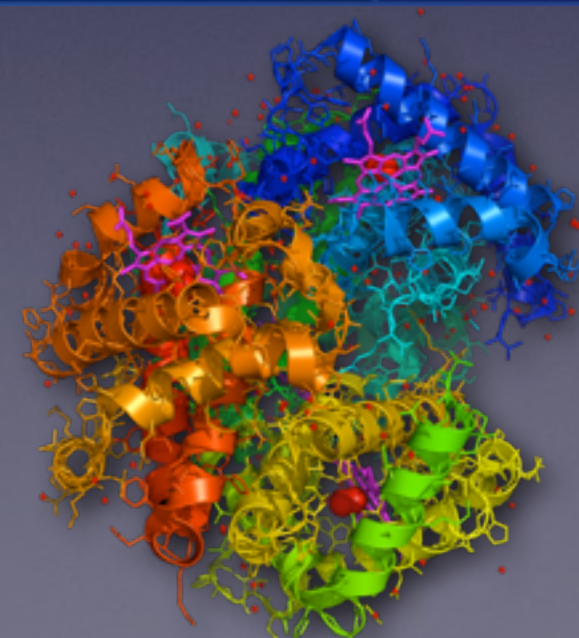


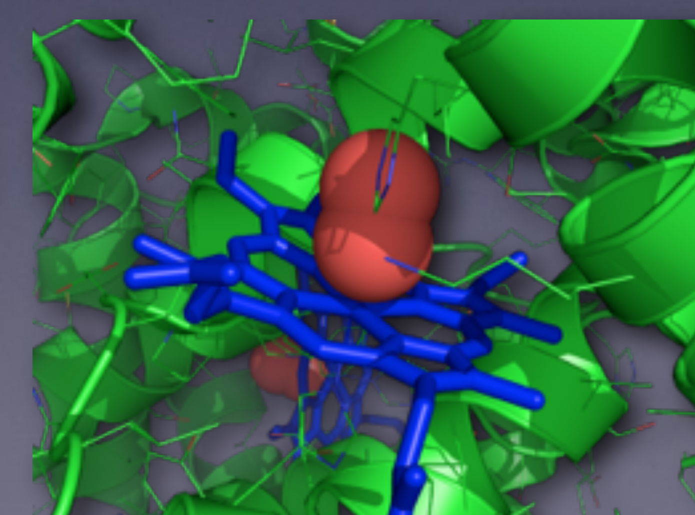


Several representations of hemoglobin

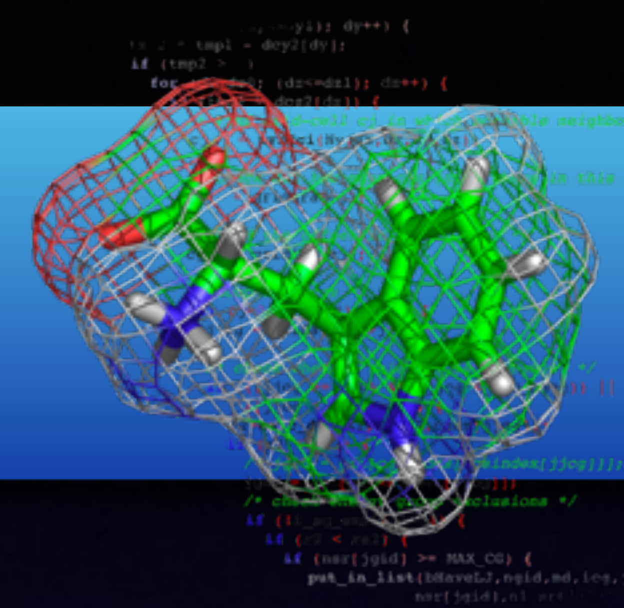




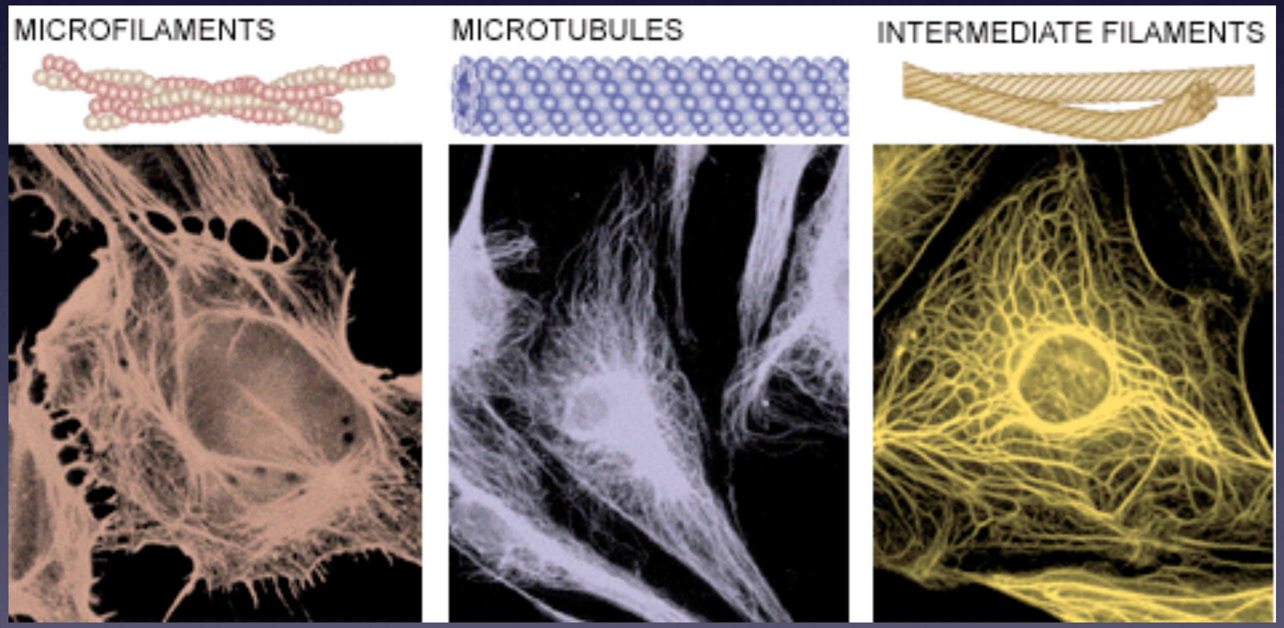


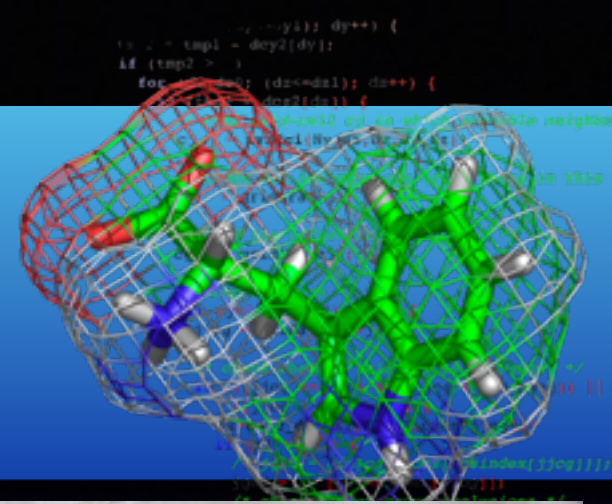


Fibrous proteins



- Structural building blocks
- Less specific biological function
- Microfilaments, tubules
- Fibrils, Hair, Nails, Shells, Claws, etc.
- Often large proteins
- But *aggregates* of them are even larger
- Regular, simple interactions (like H-bonds)



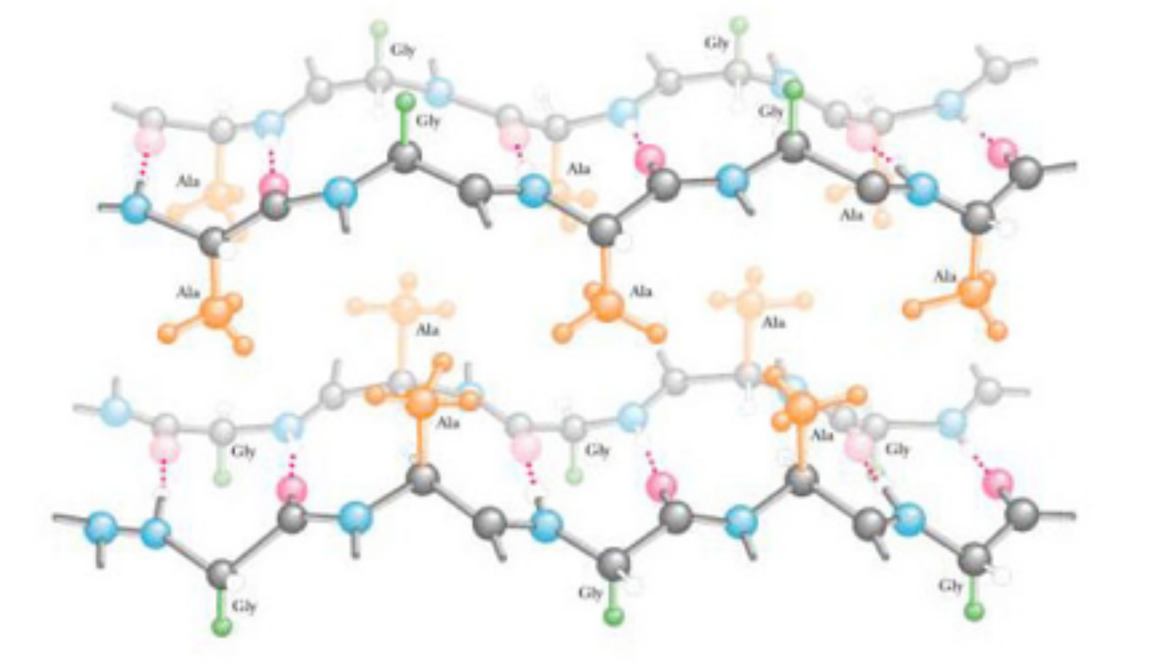


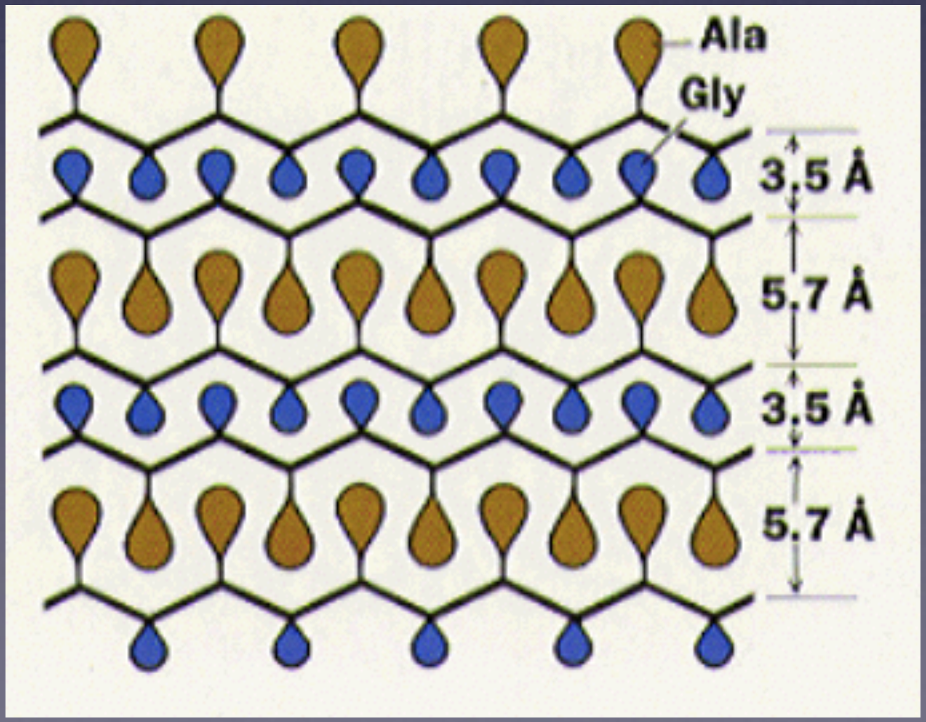
Silk fibroin

- 80% antiparallel β-sheets

—(Ser \ Gly / Ala \ Gly / Ala \ Gly)—

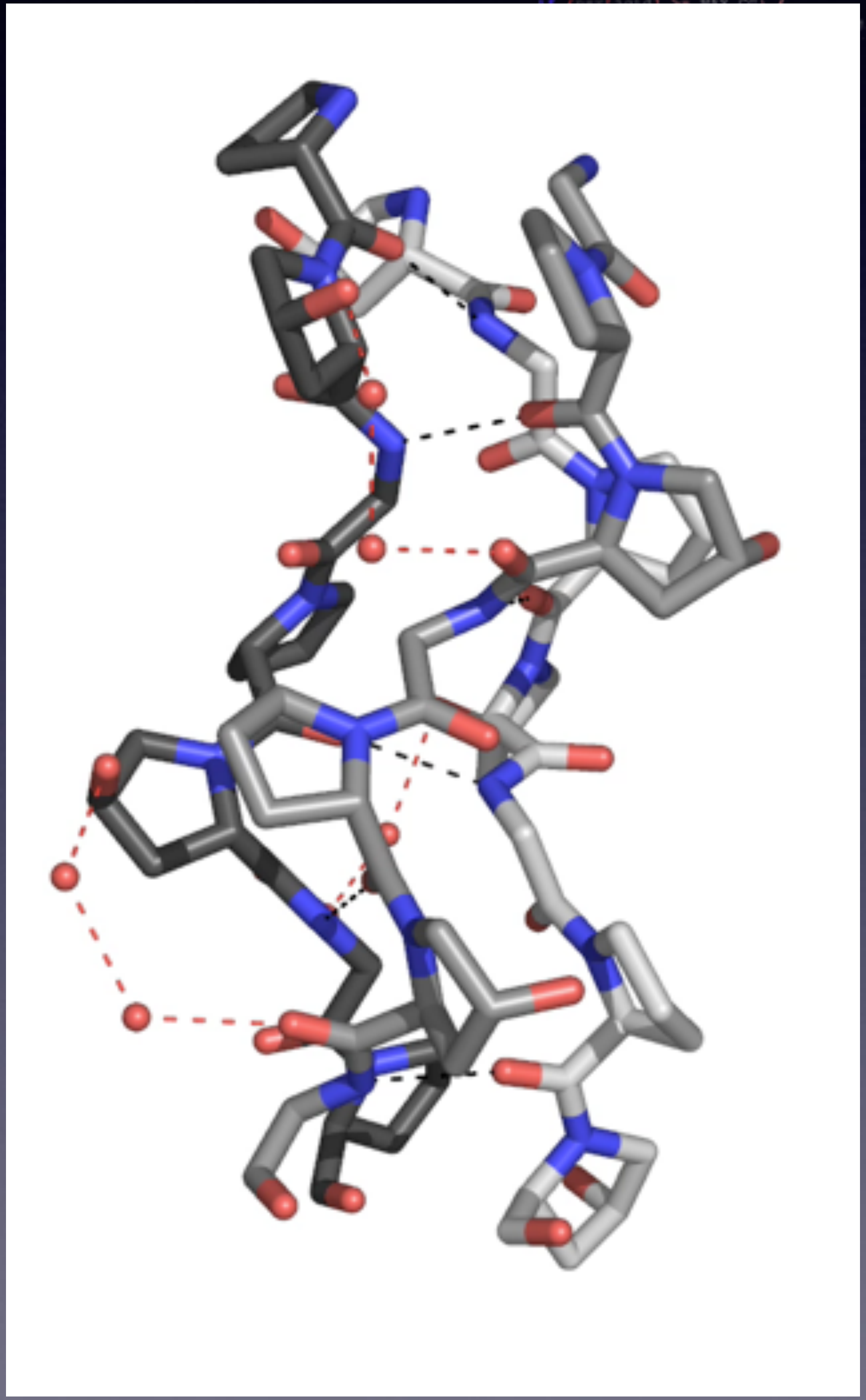
- hydro-phobic/philic surfaces
- Quasi-crystals





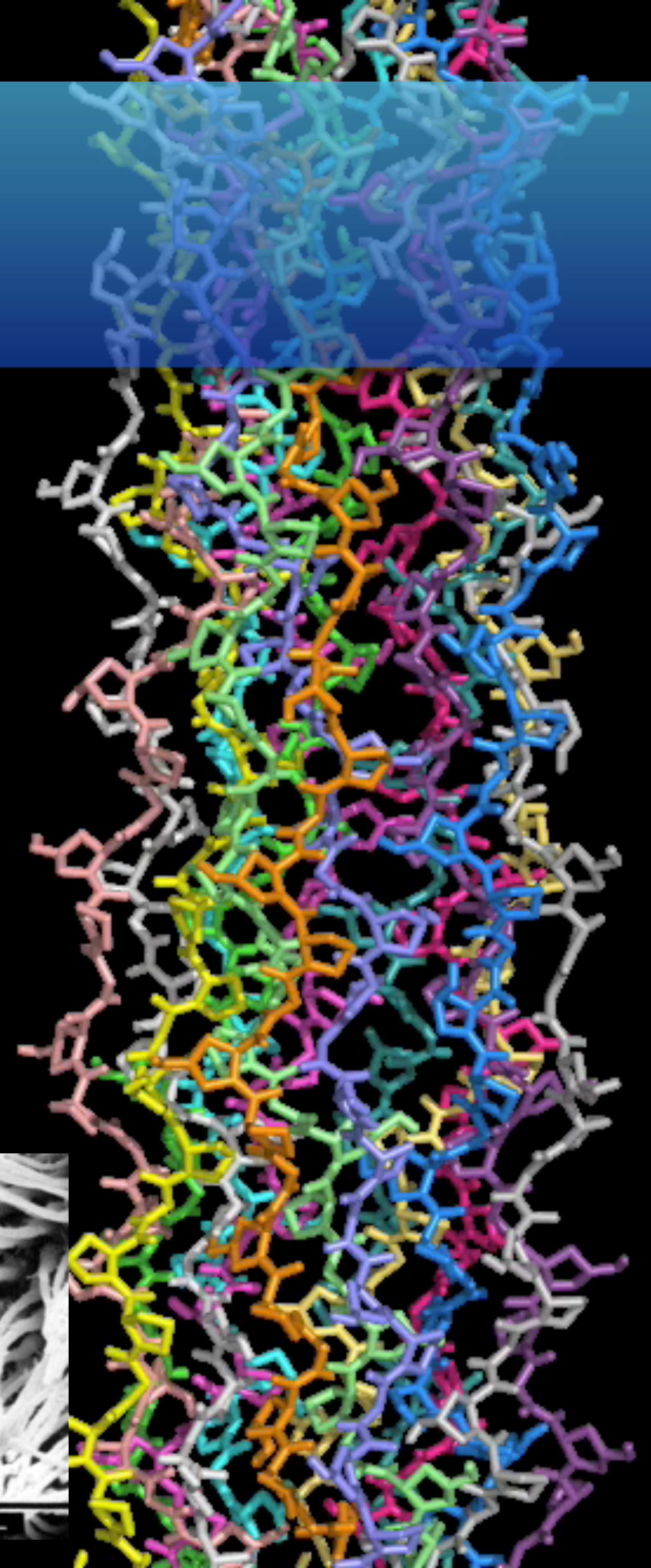
Collagen

- Triple-chain helix
- (Glycine-Proline-Proline)$_n$
- No hydrogen bonds within chains
- 25% of the protein in your body!
- Bone, teeth, skin, etc.
- 15Å wide, ~3000Å long
- Aggregates into larger *quaternary* structures

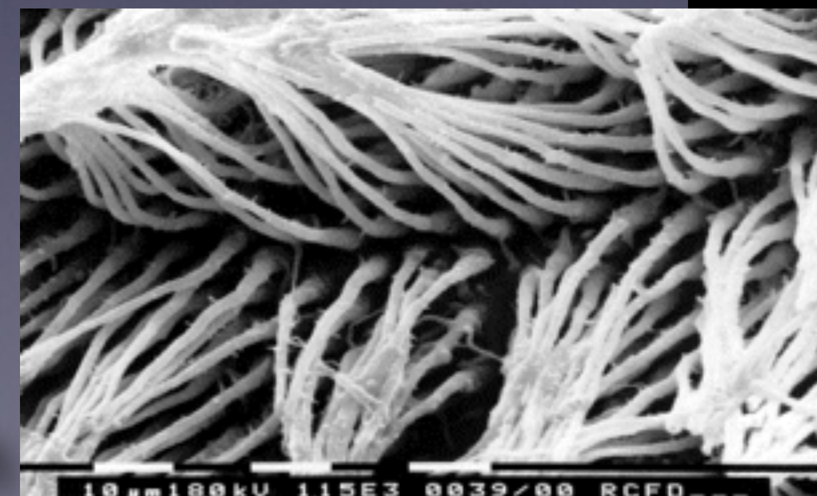


Collagen

- Chains are *not* identical (2+1)
- 3 chains form a superhelix
- Superhelices aggregate to fibrils

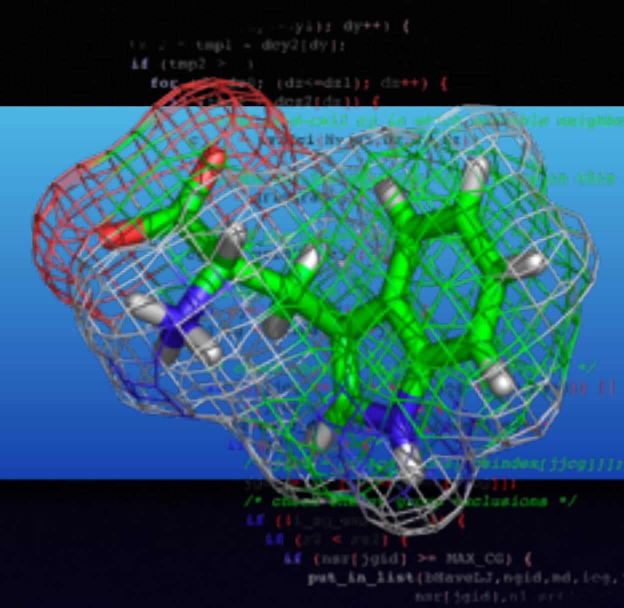


Mutations G->X cause brittle-bone disease

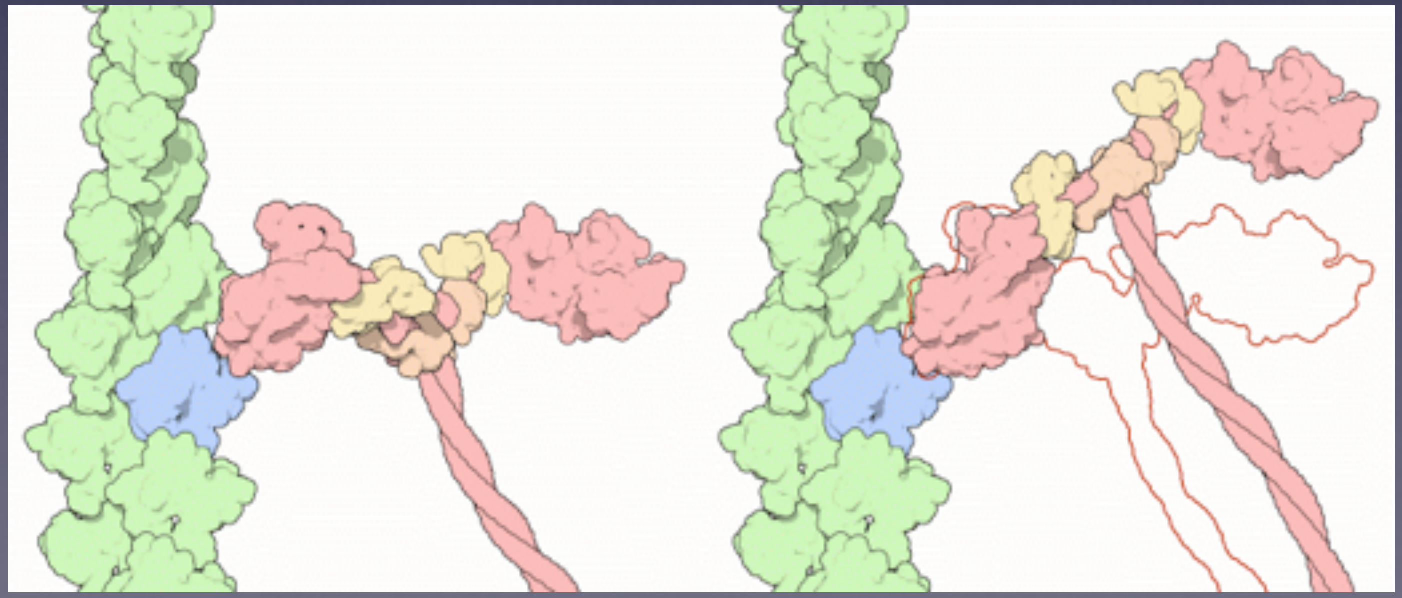


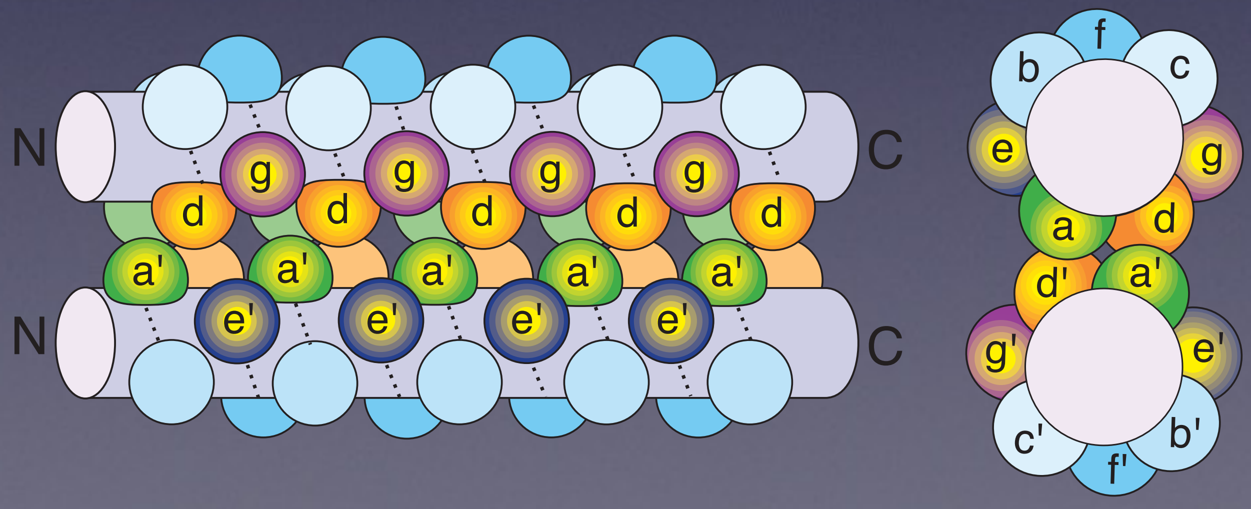
Dentin fibrils in tooth

Coiled coil helices



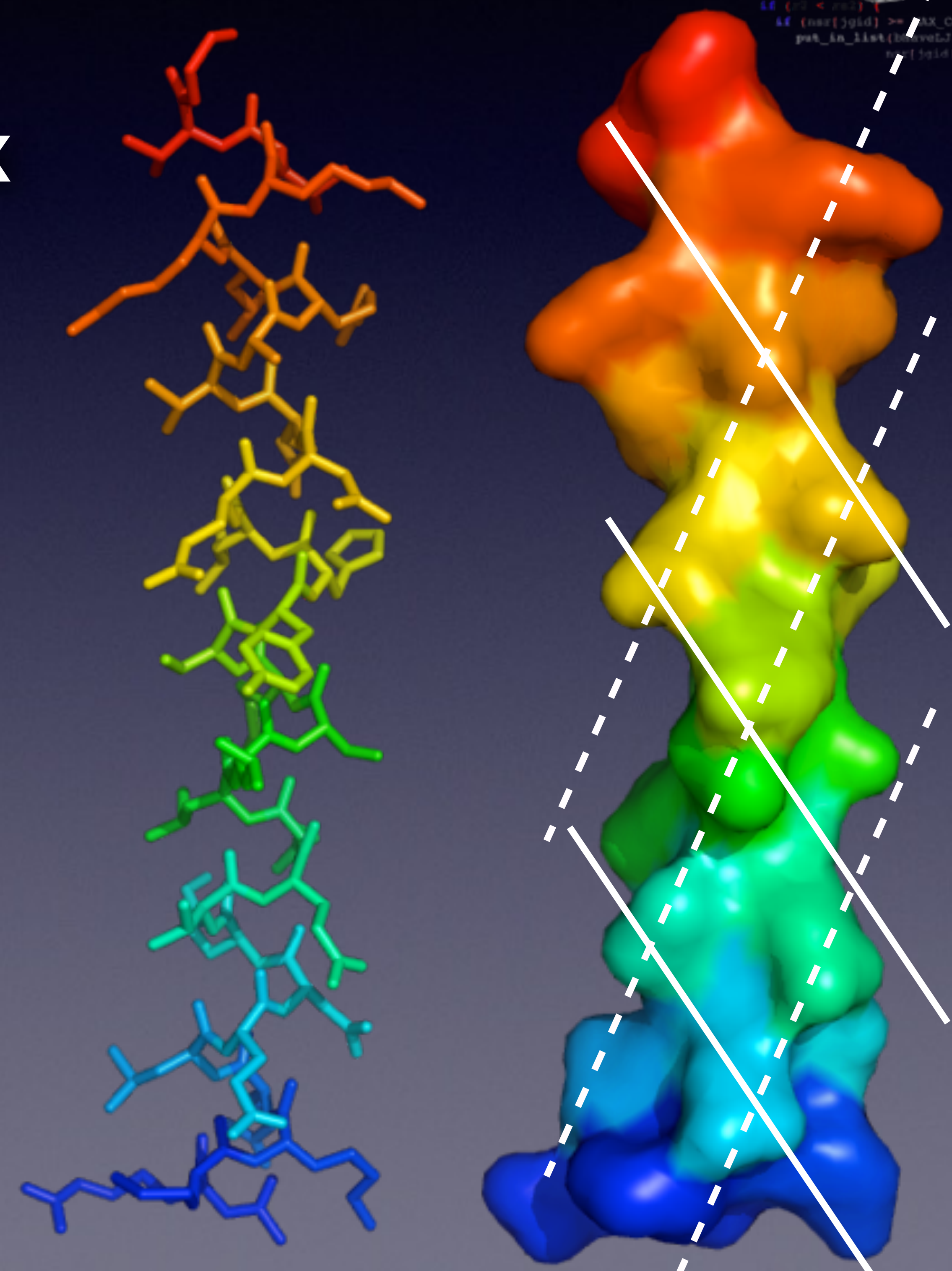
- Associated α-helices
- Normally 2, but sometimes 3 or more
- Myosin proteins in muscles
- Often 3.5 residues per turn instead of 3.6 Why?





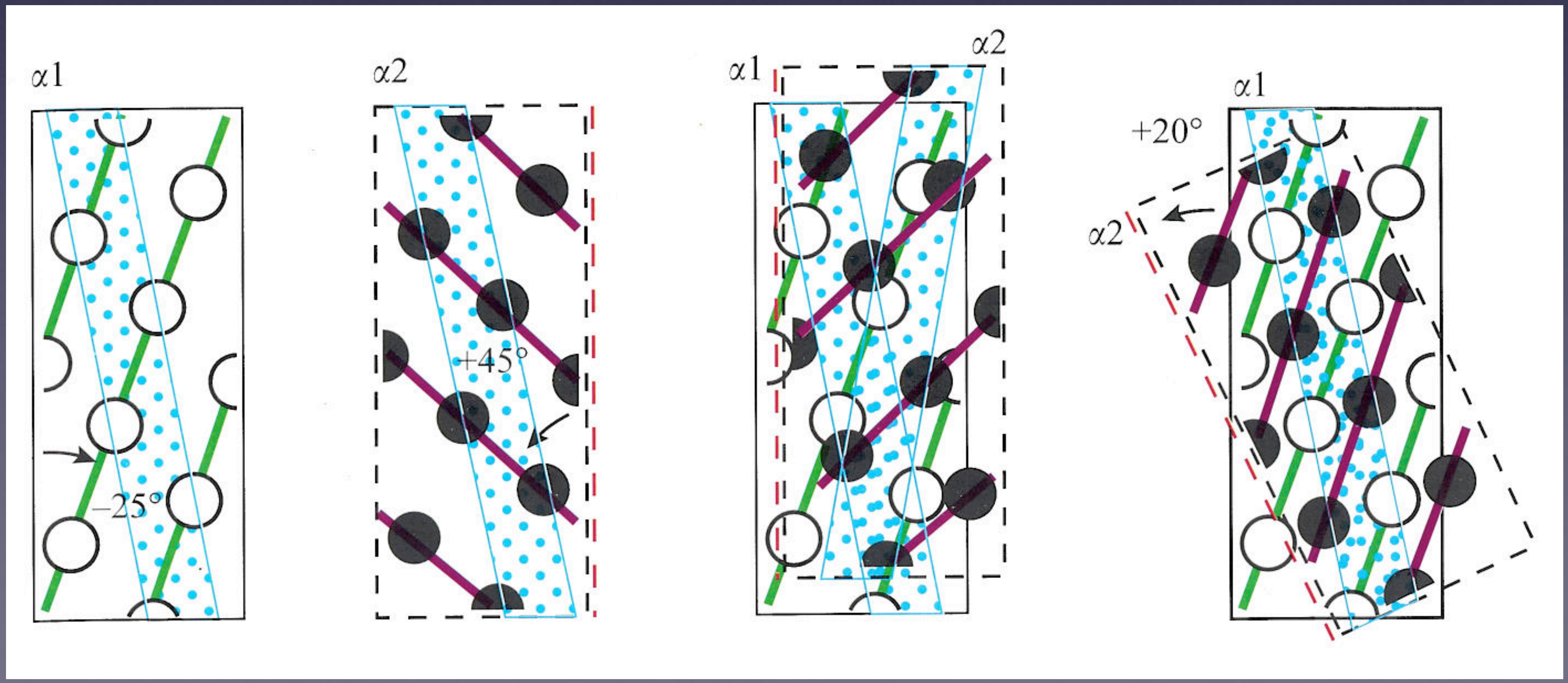
Helix geometry

- Ridges on surface of helix
- i, i+3, i+6 (45 deg): ——
- i, i+4, i+8 (-25 deg): - - - -
- What happens when two helices interact?



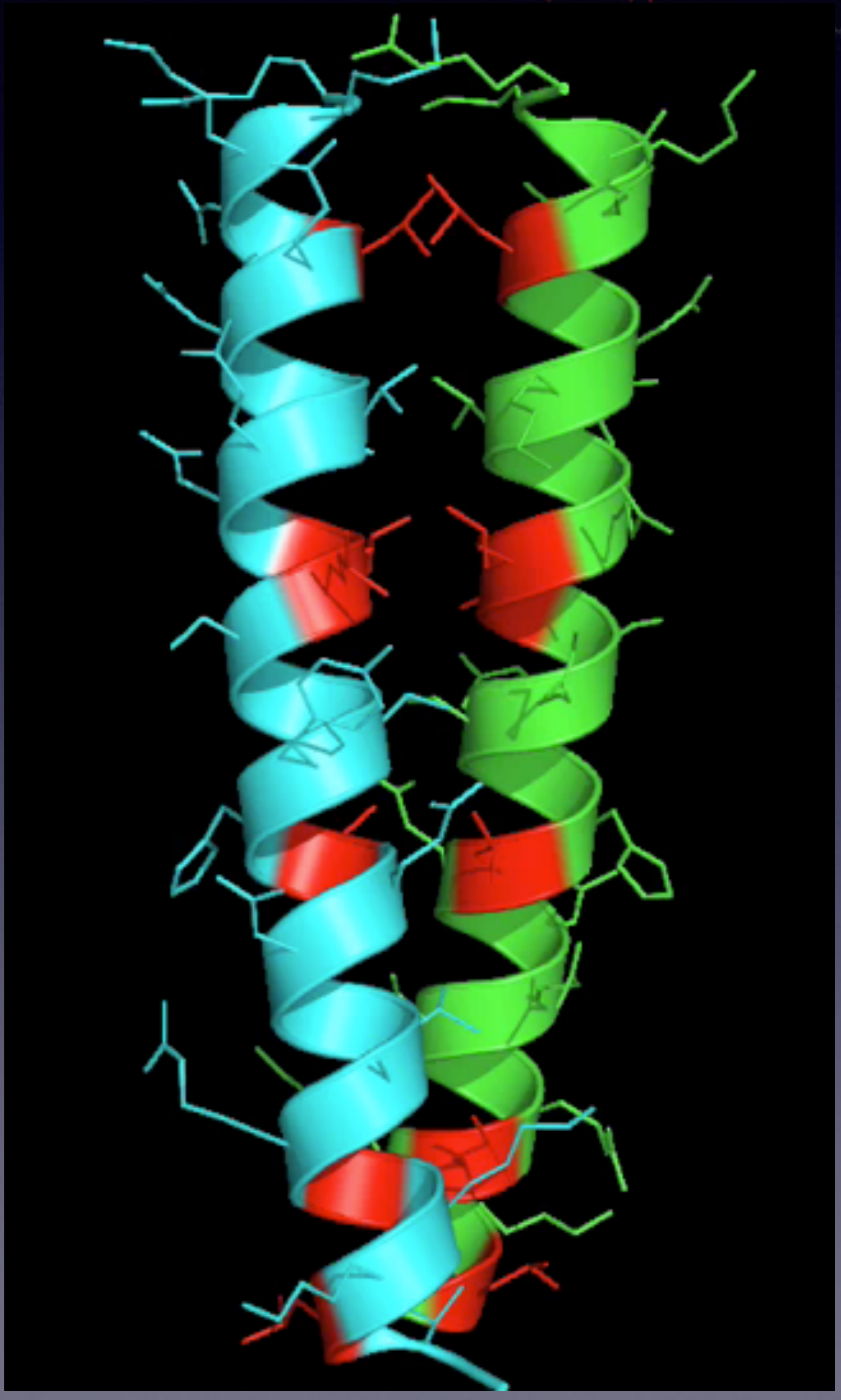
Helix packing

- Try packing helices by rotating so that the i,i+3 ridge of one is parallel to i,i+4 of the other
- Leads to very tight contact for long helices
- Predicted by Francis Crick, 1953

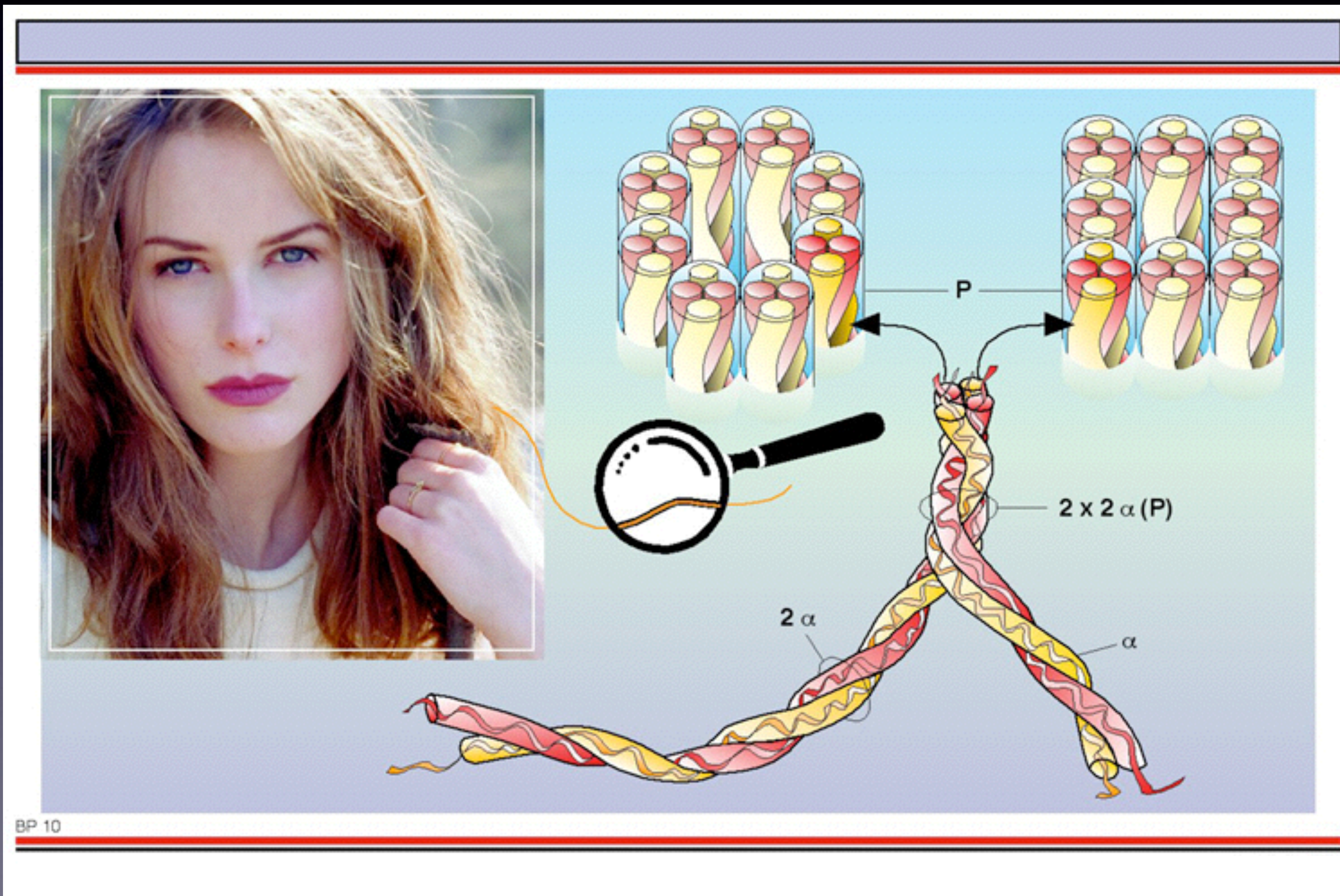


α-Keratin

- Entire protein is a coiled coil
- Every 7th residue is Leucine (red)
- Packed, 20 degree angles
- “Leucine zippers”
  - Hydrophobic sidechains pack
- 11% of the residues are Cystein
  - Stabilization by disulphide bonds of thiol groups

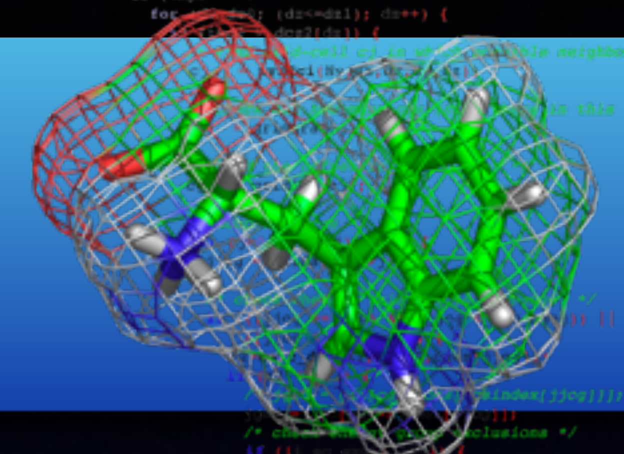


α-Keratin

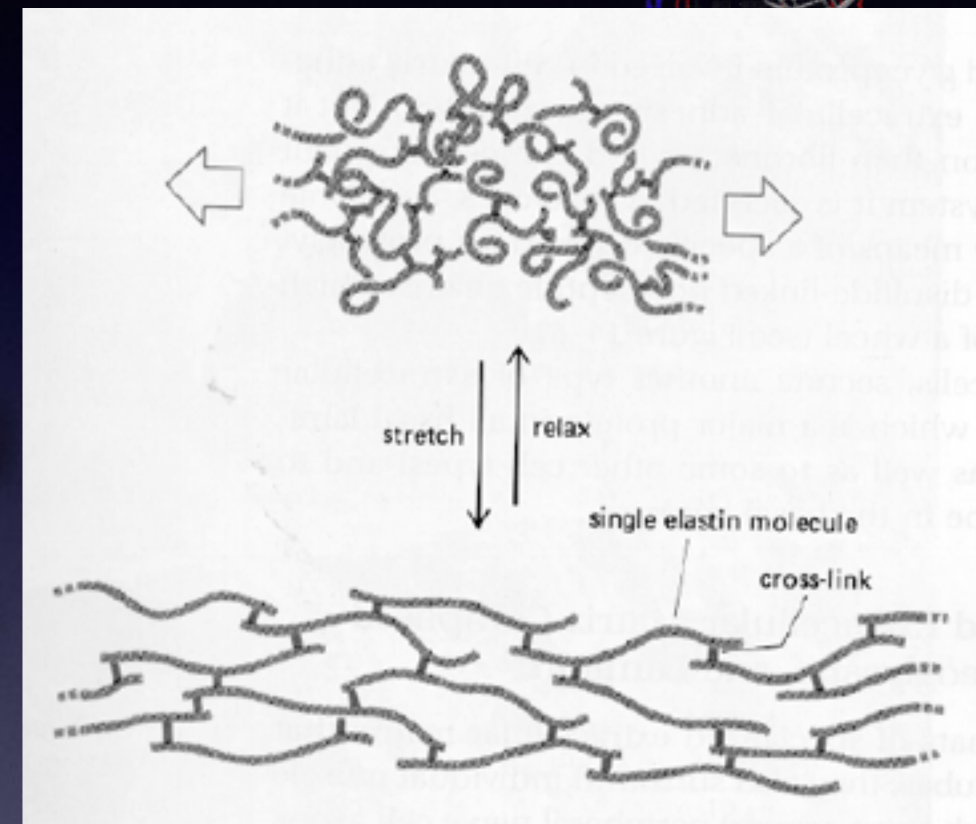


~10 turns of alpha helix produced per second!

Elastin

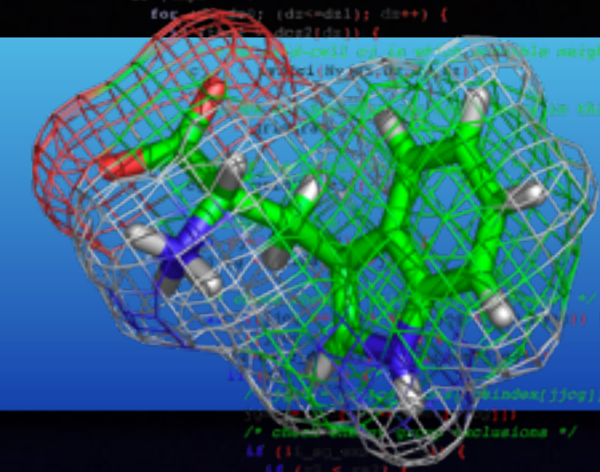


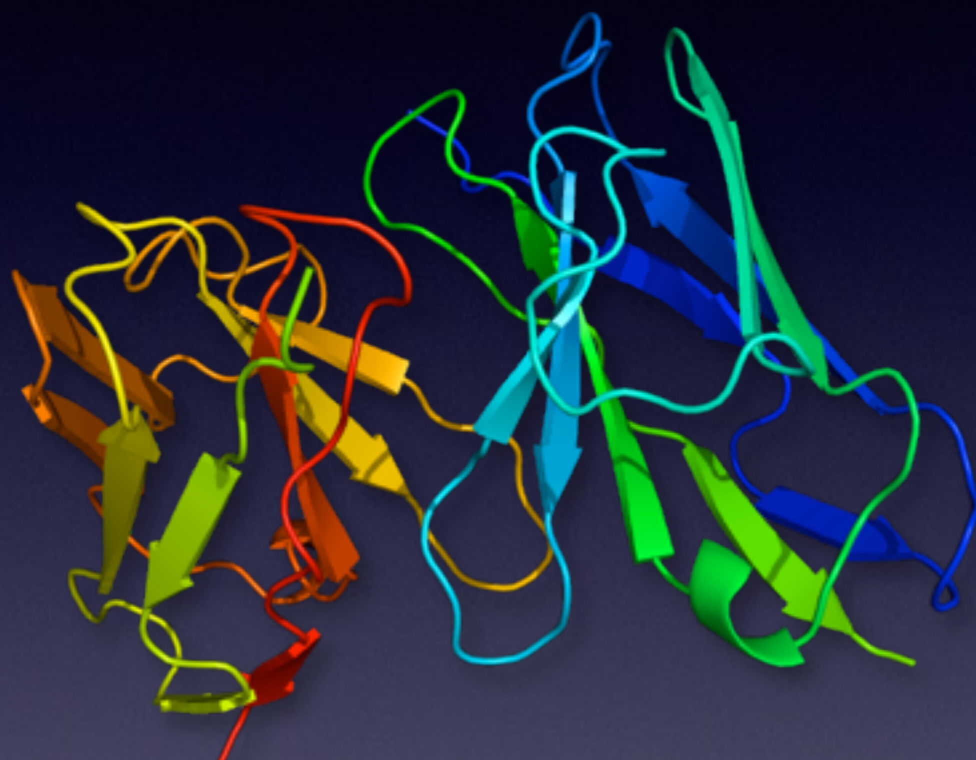
- Highly elastic fibrous protein
- Similar to collagen, but ~10% helix, 45% sheet, 45% coil
- Cross-linked by modified Lysine
- Deficiency in Lysine-modifying enzymes can lead to loss of elasticity of vessels, and in worst case aorta rupture!



Real aorta vs. elastin biomaterial

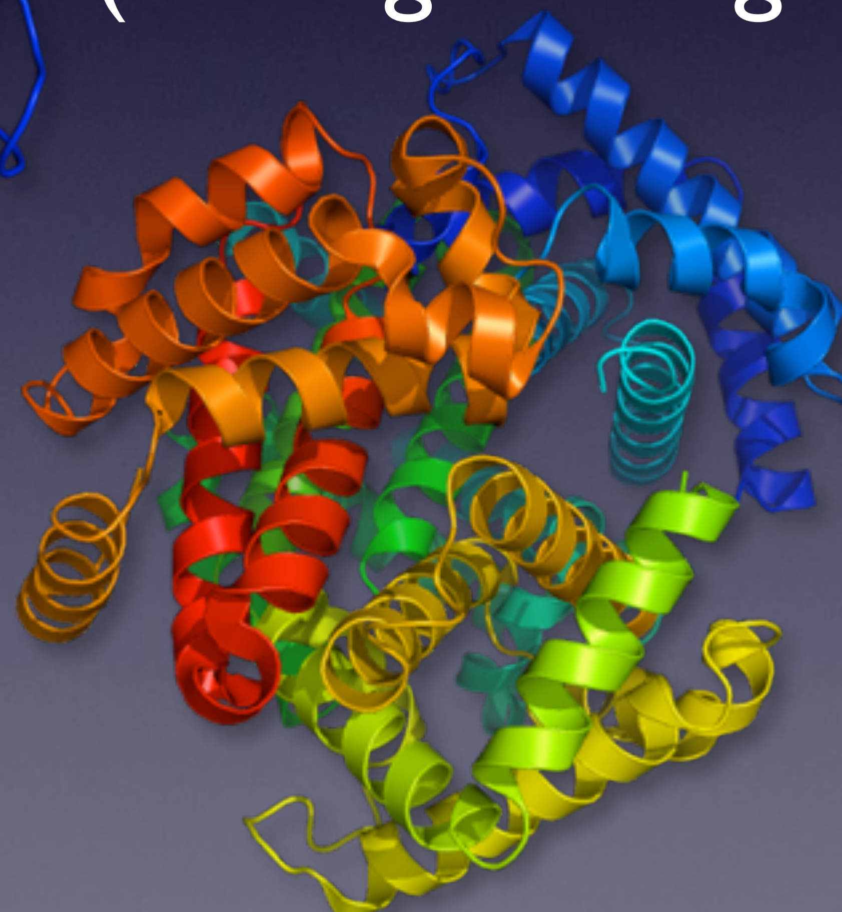
Globular proteins

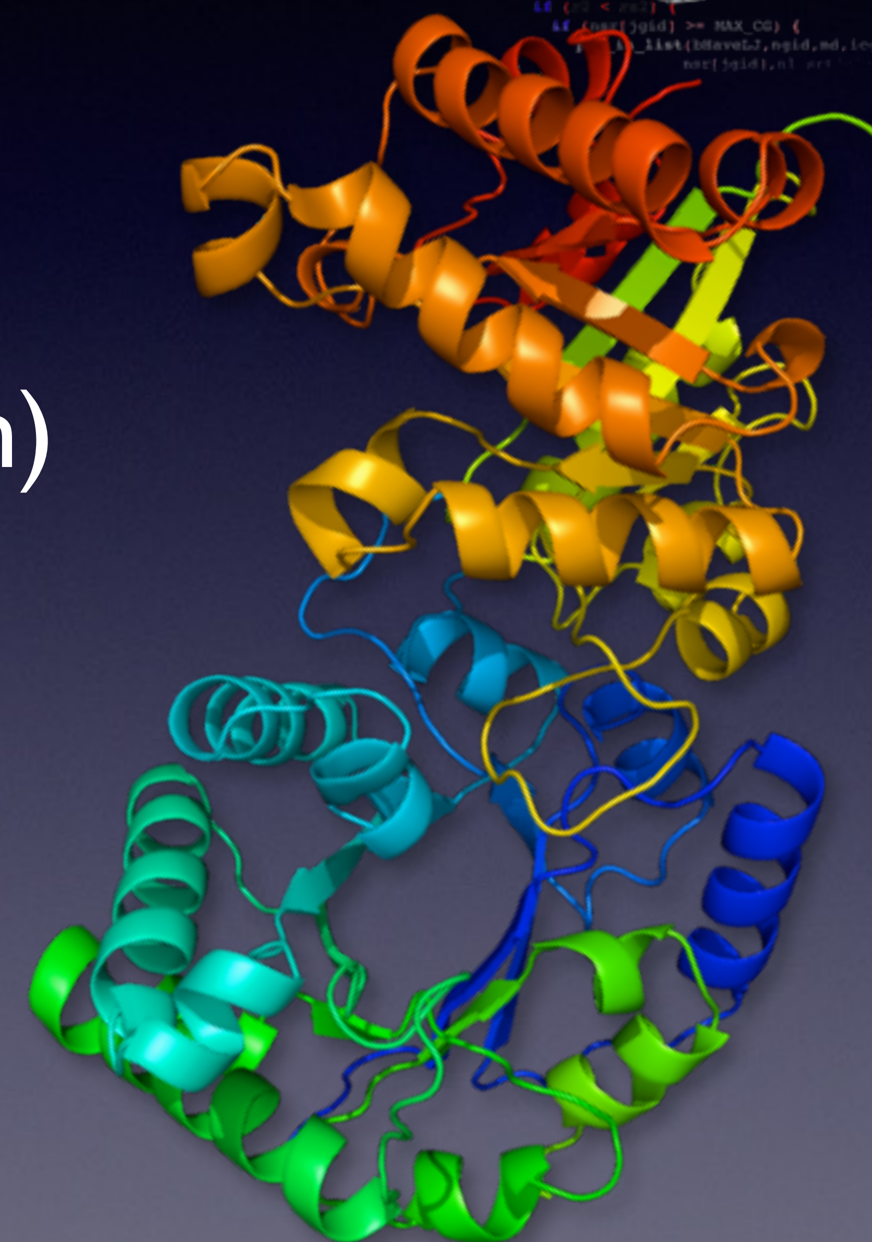




All β-sheet

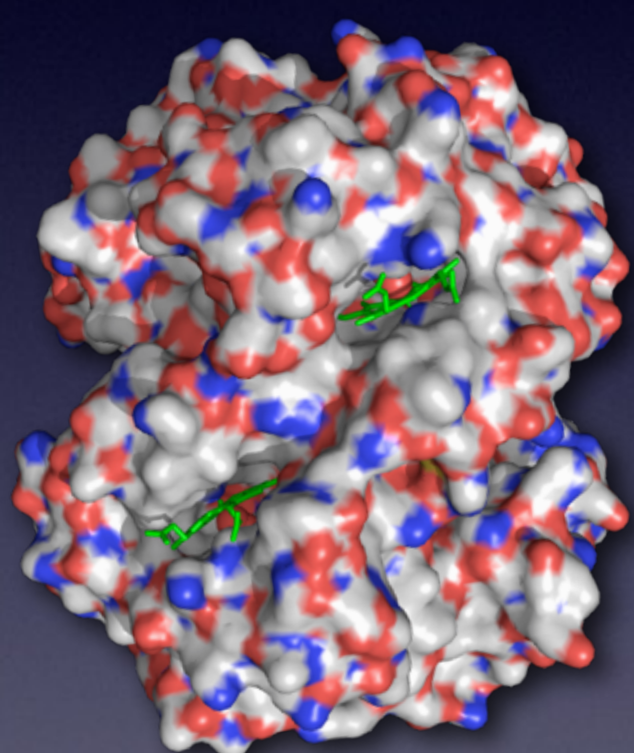
All α-helix
(hemoglobin again)



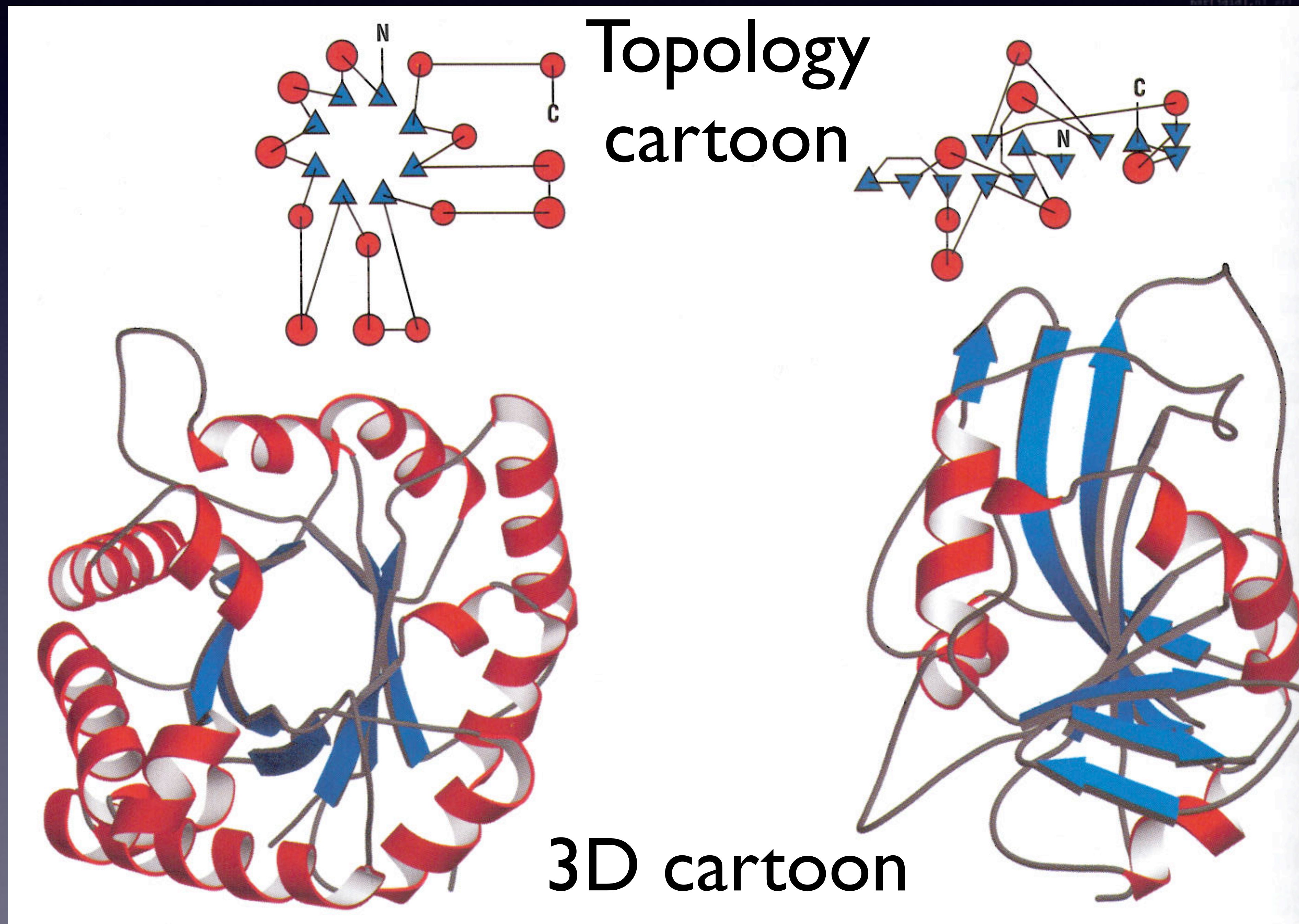


Mixed helix/sheet

Protein topologies

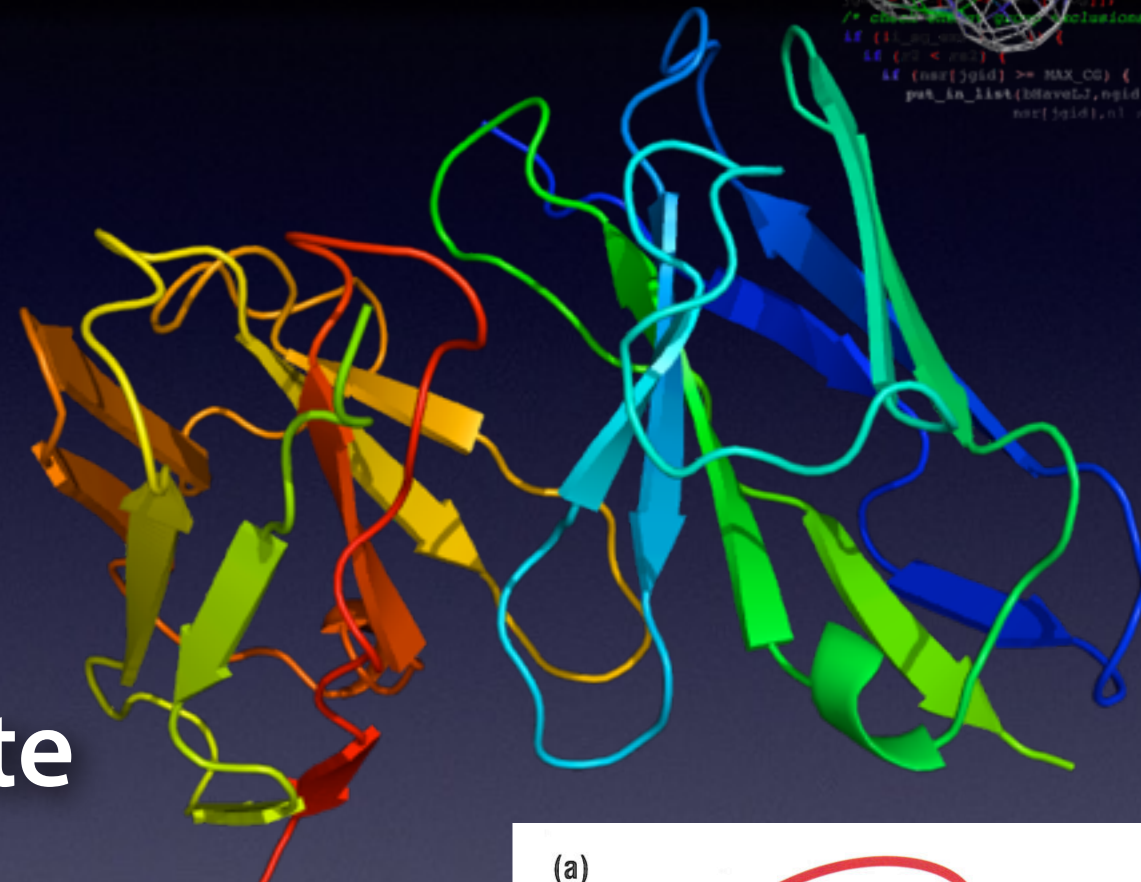


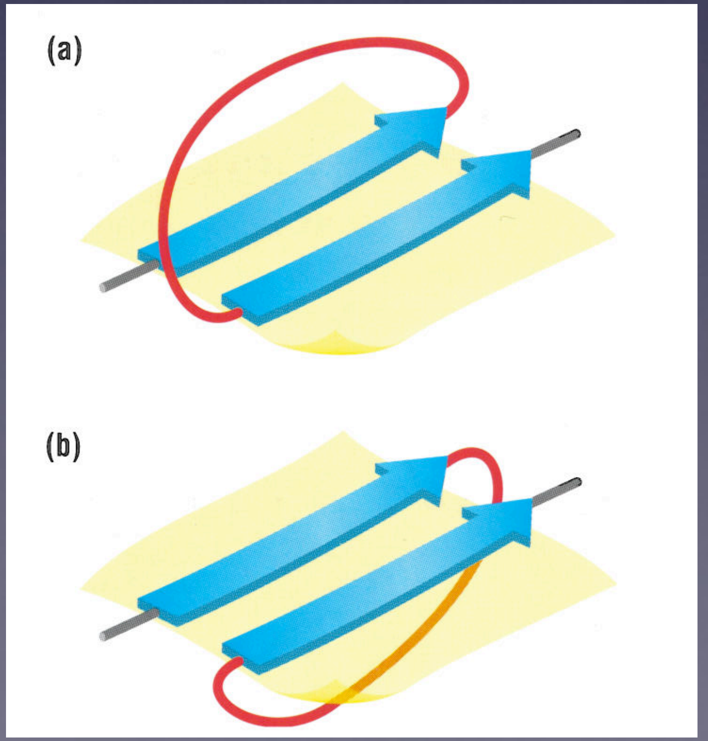
Structure



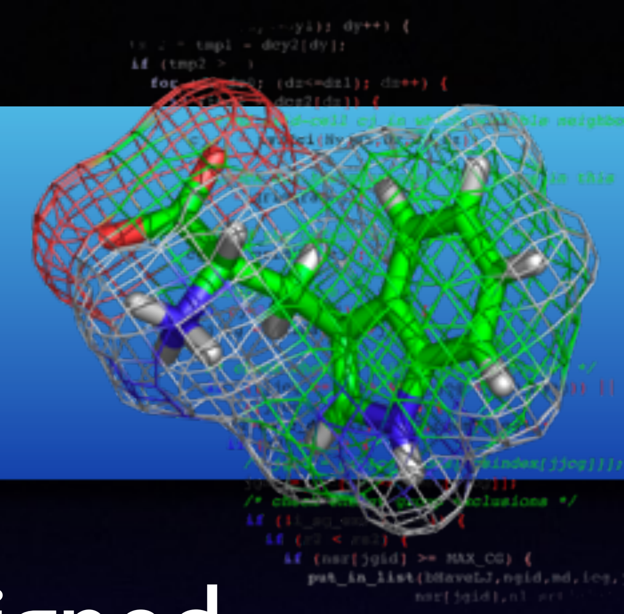
Globular β structure

- β-structures are simple
  - Only continuous sheets
  - 1-2 stacked sheets
- Antiparallel sheets dominate
- Sheets are slightly twisted
- Parallel sheets typically have right-handed geometry of the cross-over coil (left very rare)





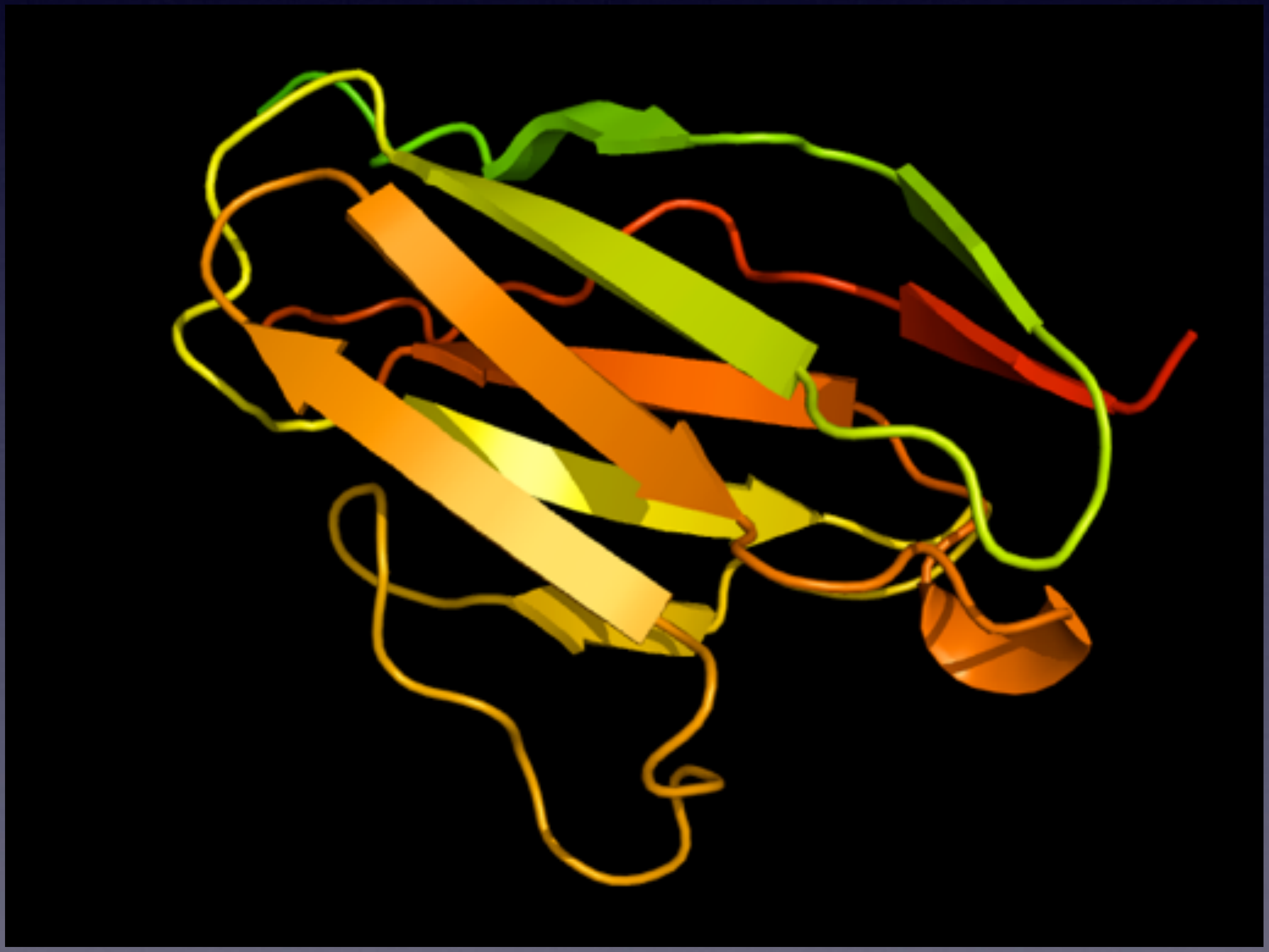
β-sheet packing



Orthogonal

"β cylinder/barrel": FABP

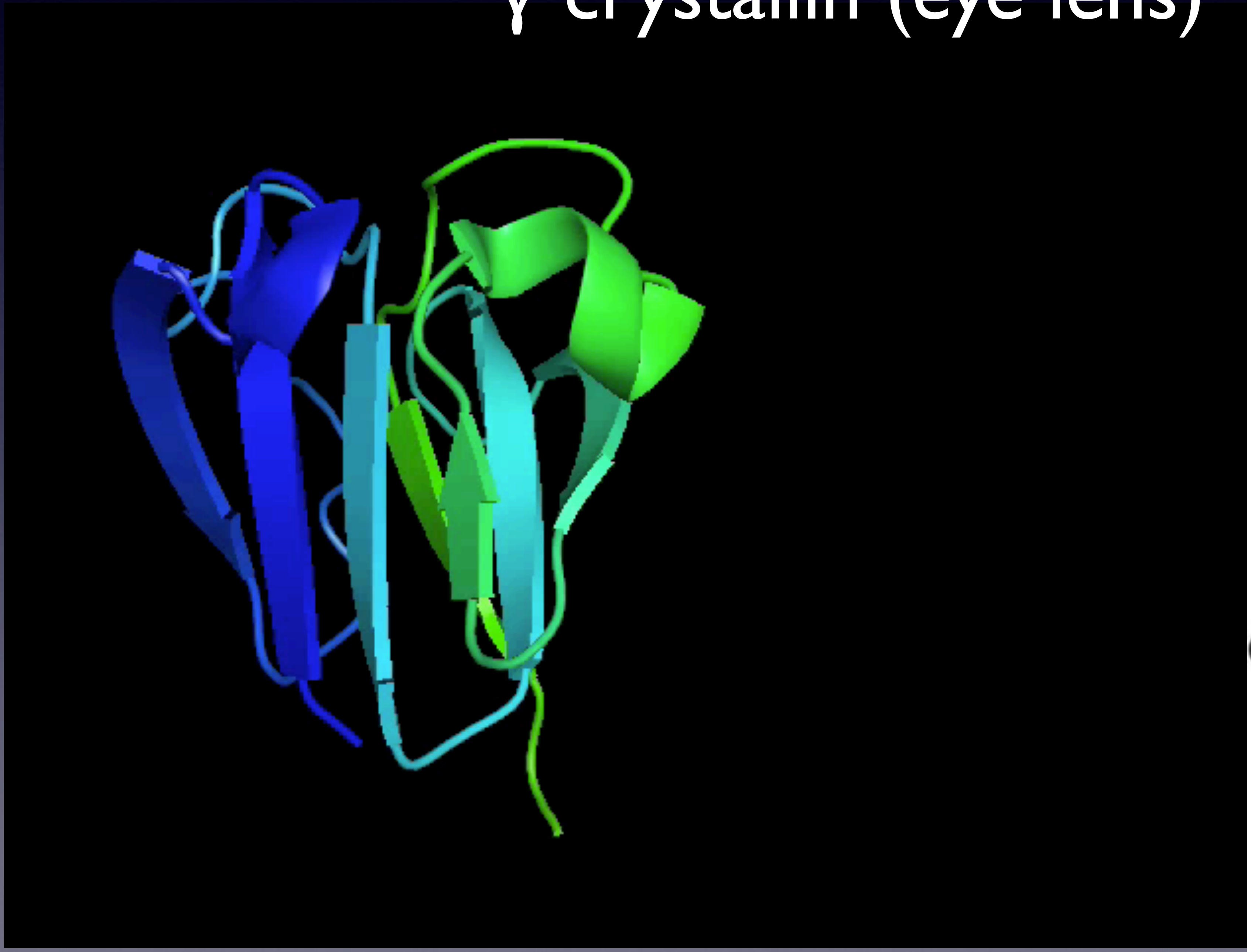
Aligned

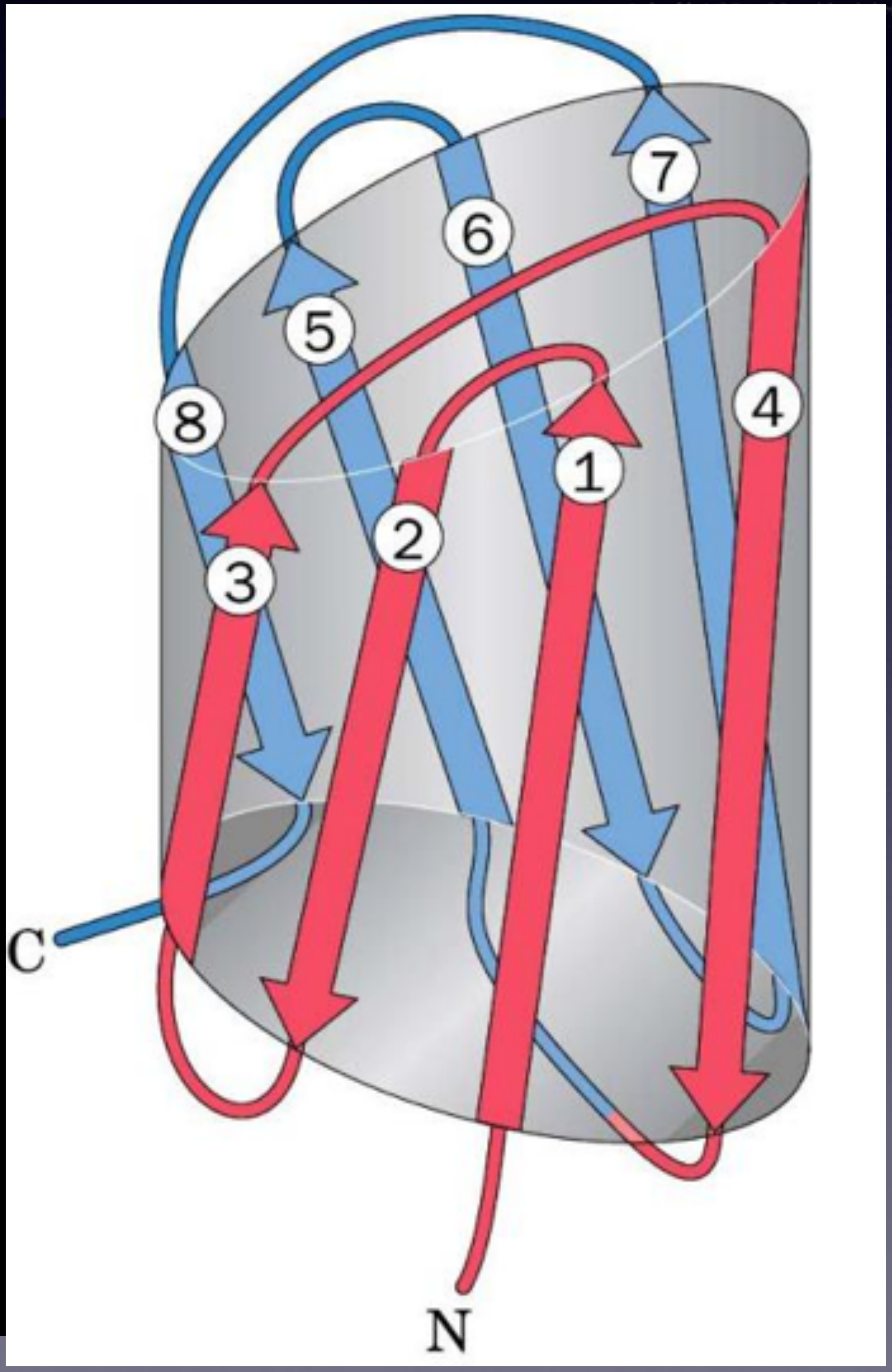


"β sandwich": Immunoglobulin

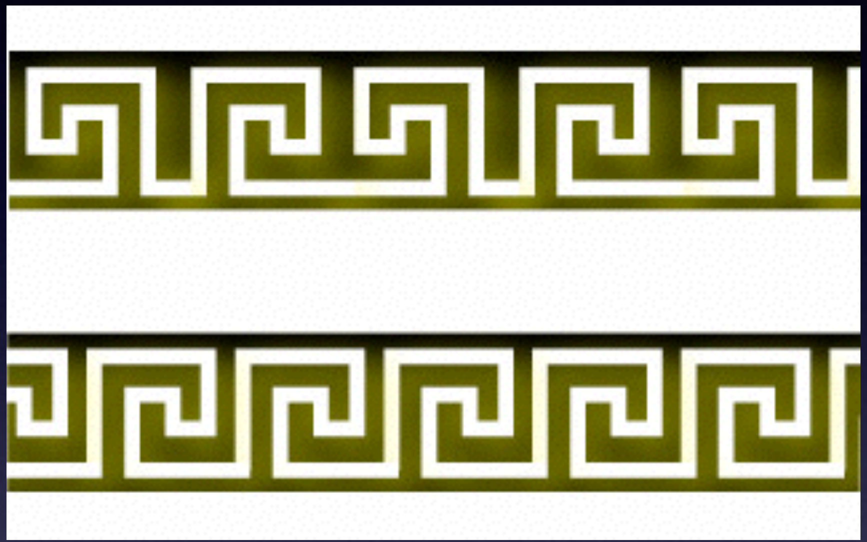
β-sheet topologies

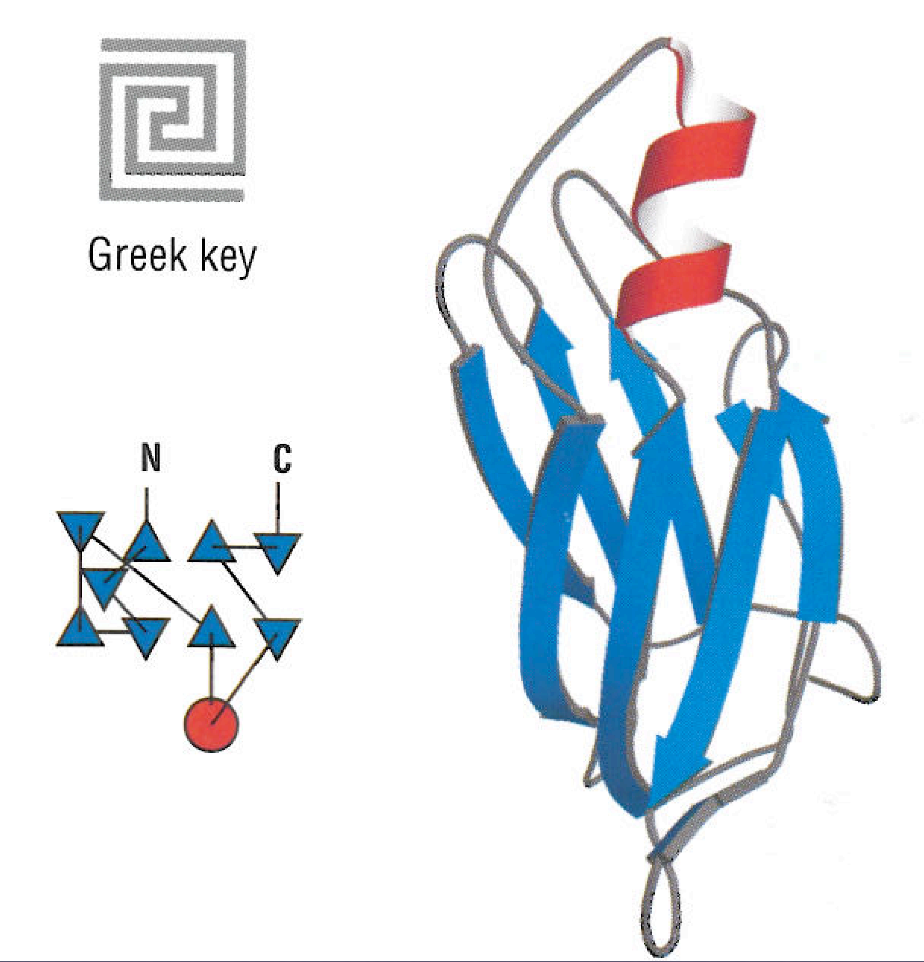
γ crystallin (eye lens)



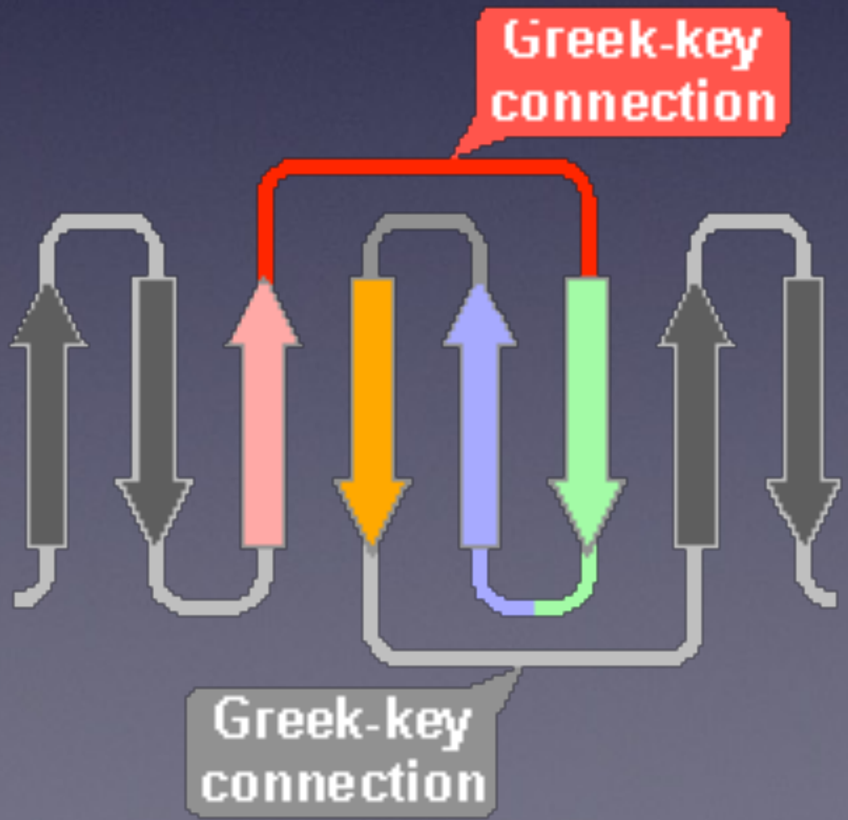


Greek keys

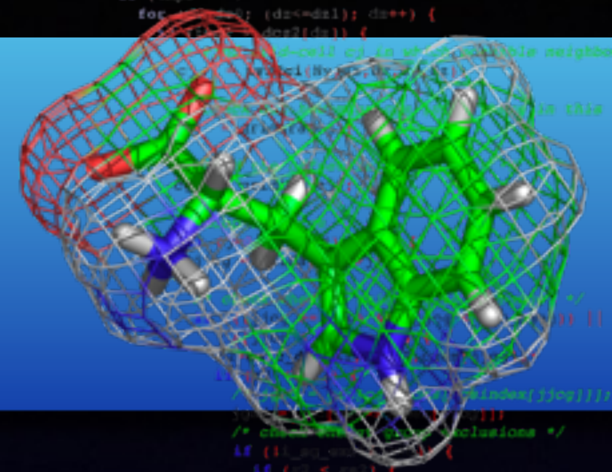


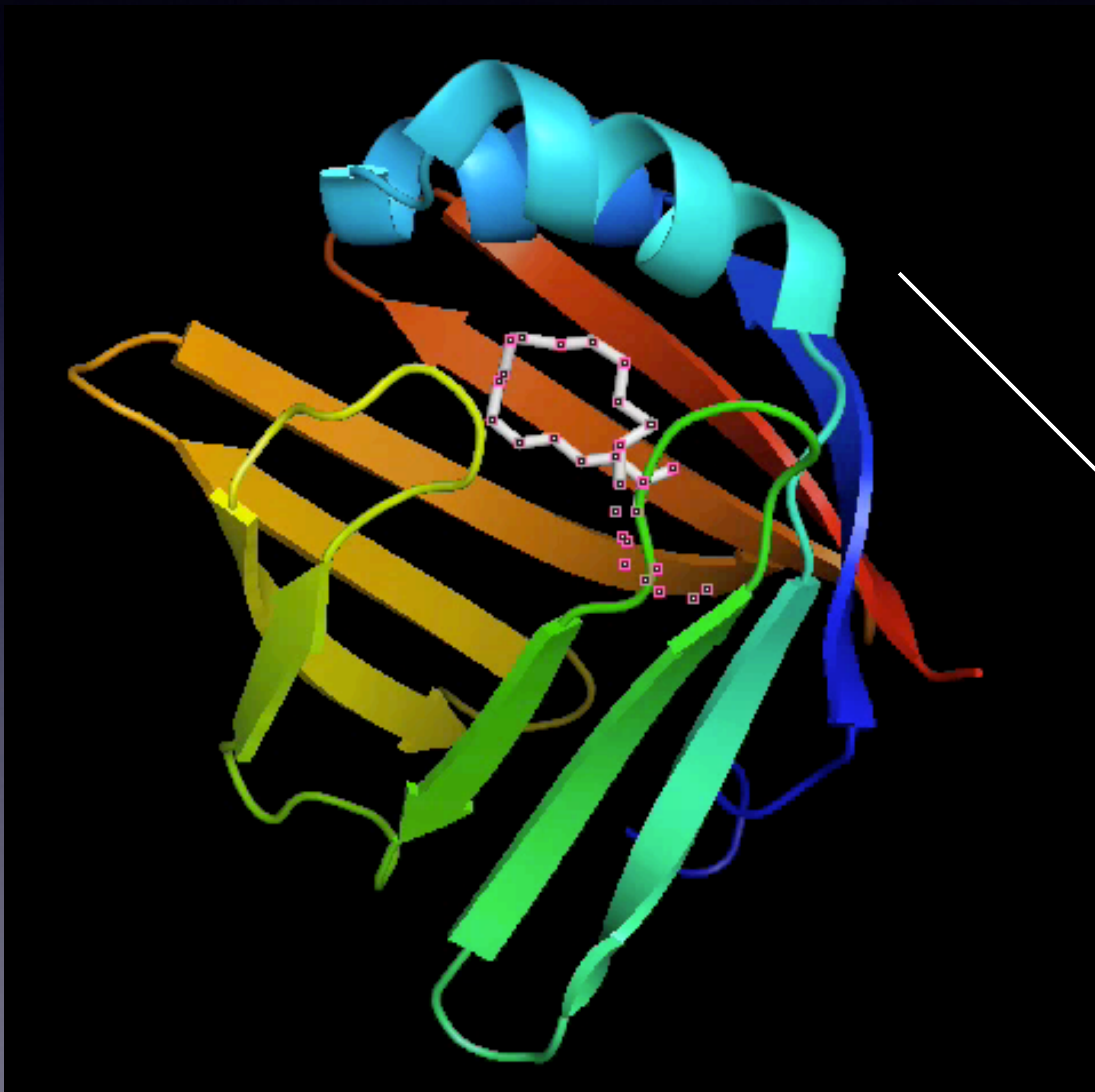


pre-Albumin



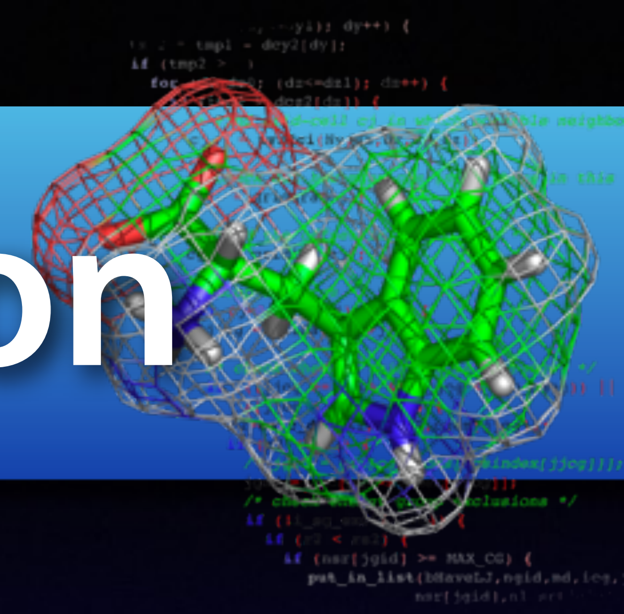
FABP





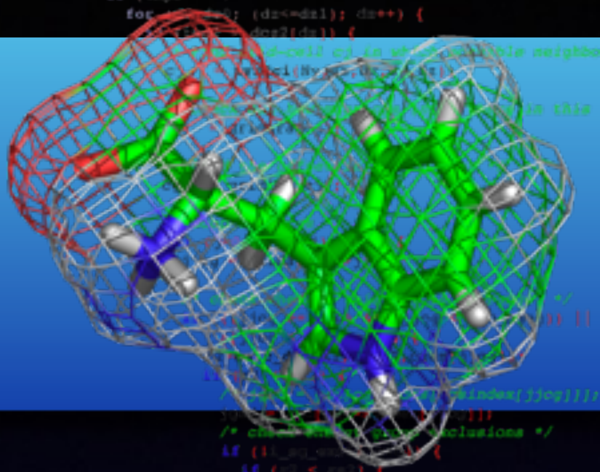
- Fatty acid binding protein
- h-phobic shielding
- Hydrophilic outside, Hydrophobic inside
- Binds oleic acid here
- β-meander motif

Topology composition

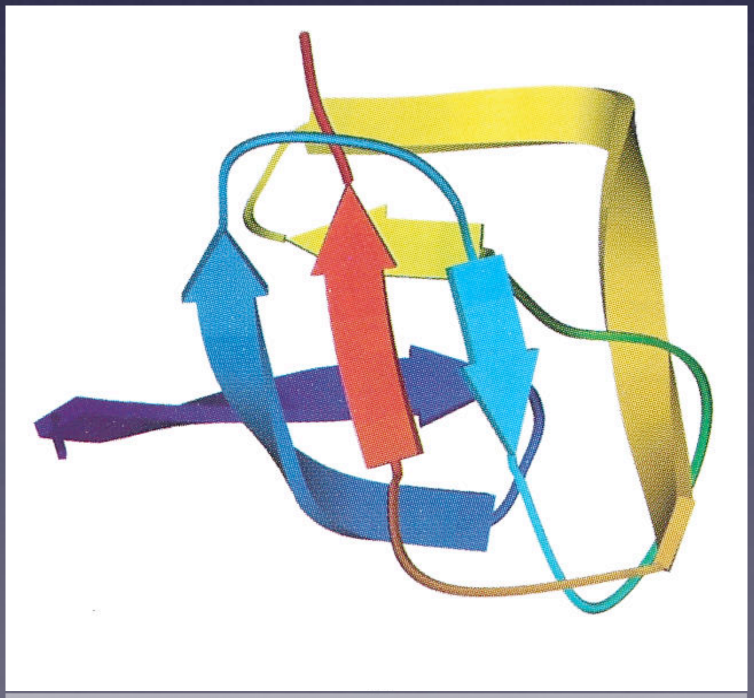


- Only limited amount of topologies observed
- Examples:
  - Almost never see mixed parallel & anti-parallel β-sheets
  - Left-handed crossover form sheets rare
- Properties of amino acids limits conformations
- Stable proteins require stable building blocks

Topology features

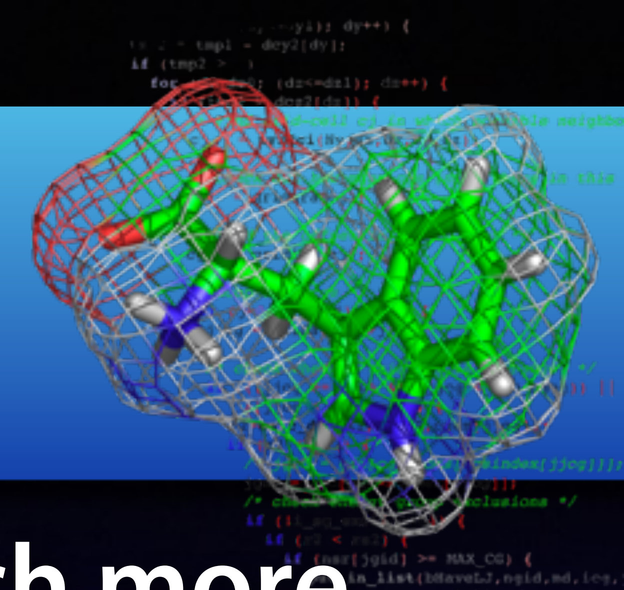


- Anti-parallel β-sheets most abundant
  - Based on hairpins
  - Likely much easier (faster) to form
- Loops do not overlap
- Knots are not allowed

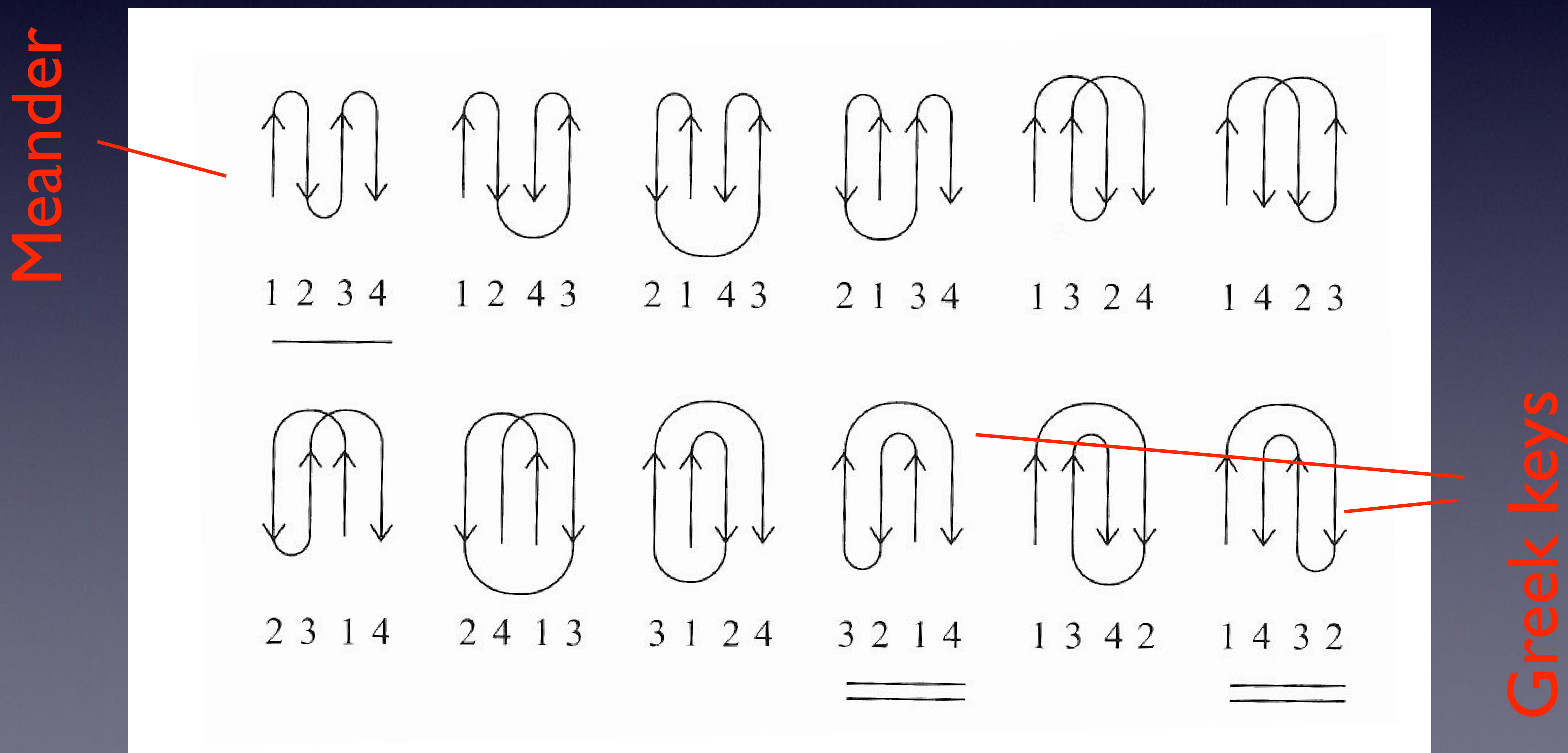


Pepsin is an exception...

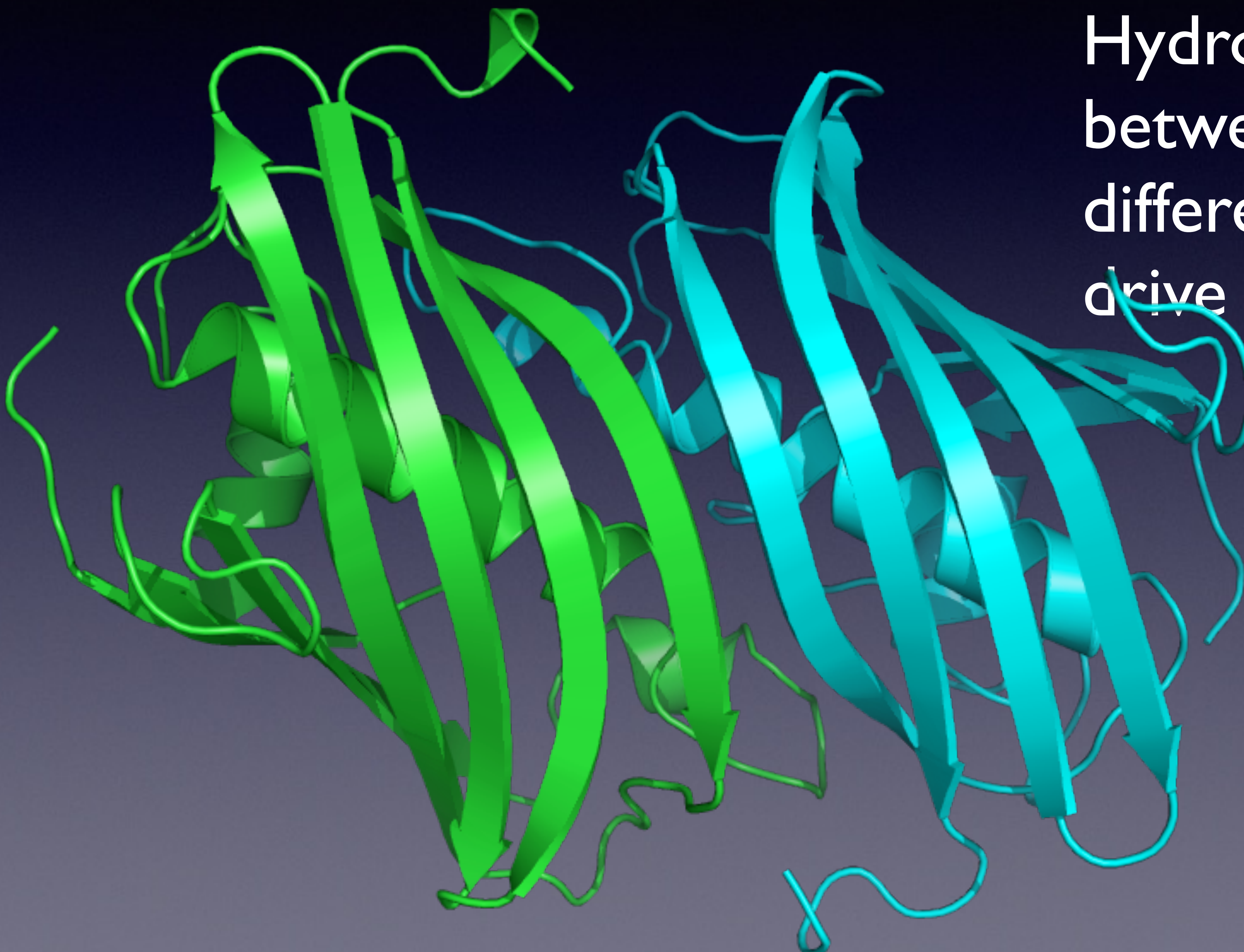
Topology features



- Some supersecondary structures much more common than others!

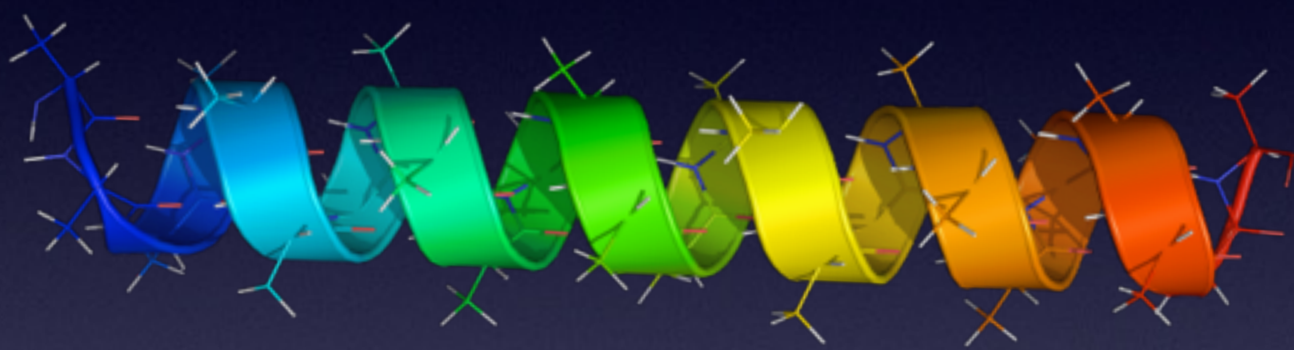


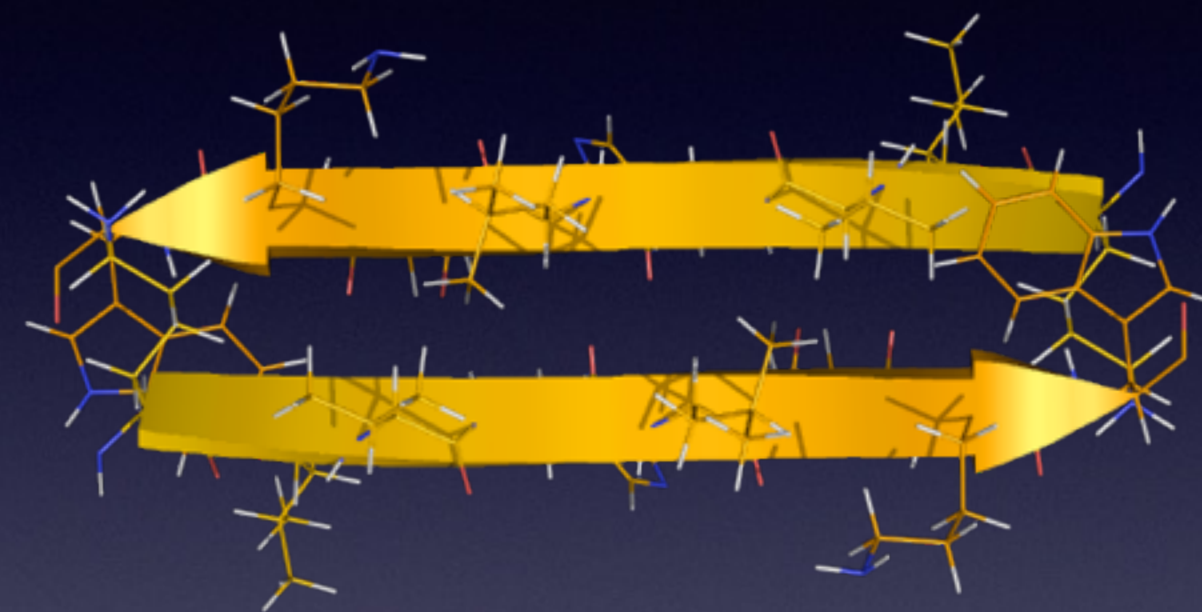
Dimerization



Hydrogen bonds between sheets of different monomers drive association

α-helices vs. β-sheets





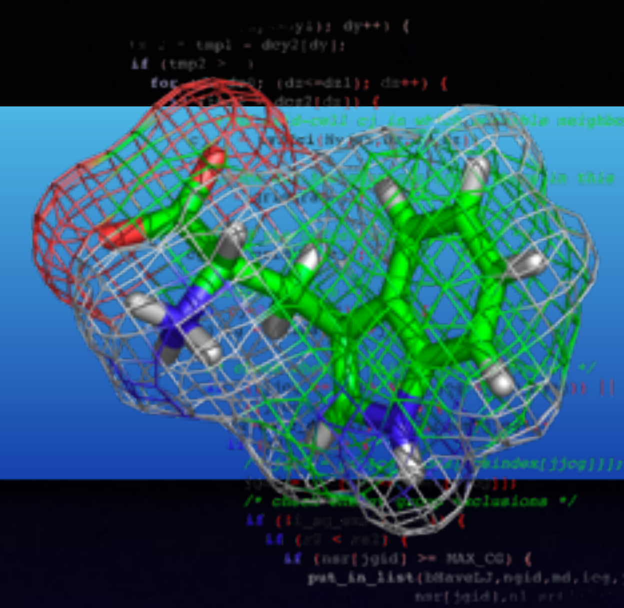
Local h-bond interactions
Rigid & isolated helix cylinders
Few h-bonds between helices
=>Few constraints

Non-local h-bonds
Flexible, interacting strands
All h-bonds between strands
=> Lots of constraints

3D-organization

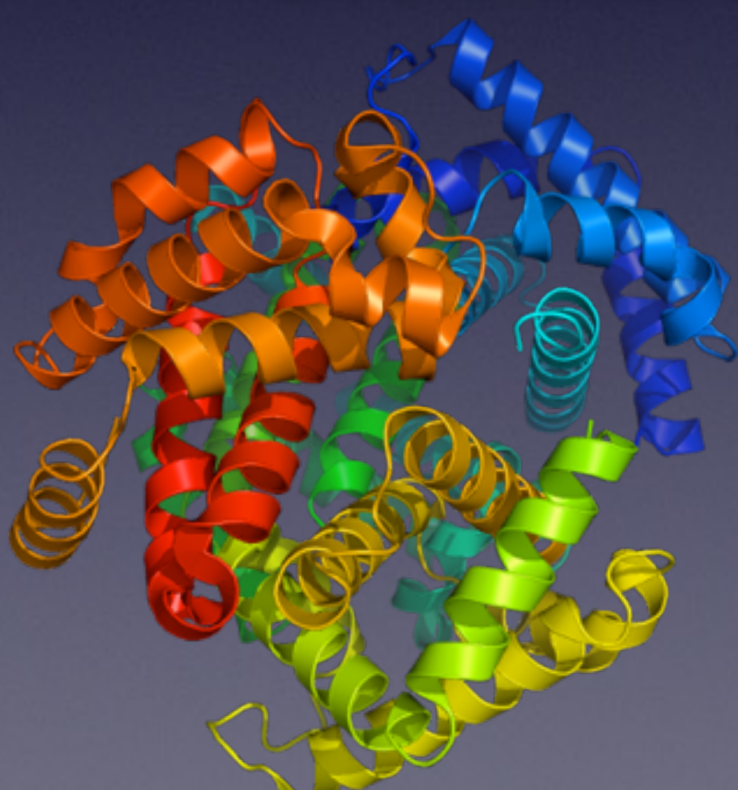
Parallel or antiparallel layers of β-sheets

Only requirement is packing of helices

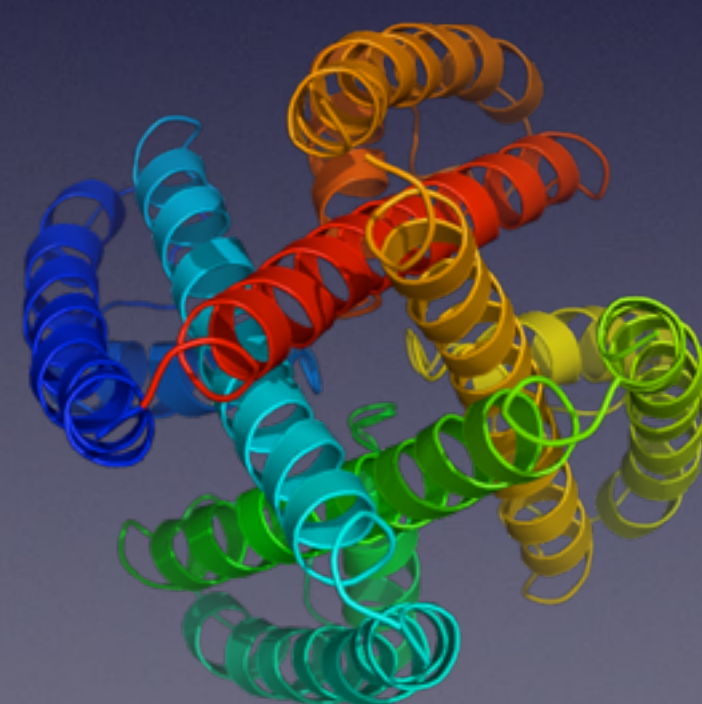


Helical diversity

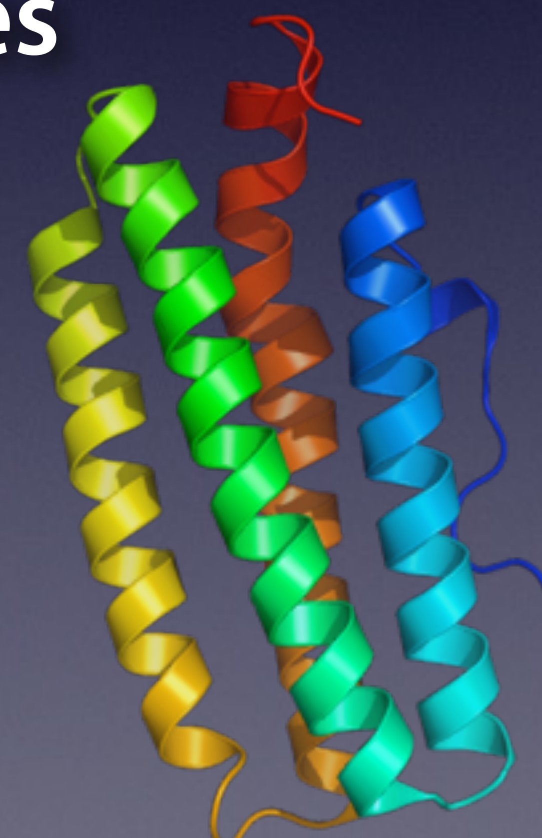
- Helical structures exhibit more diversity, and are harder to classify than all-β
- Even more true for mixed structures



Hemoglobin

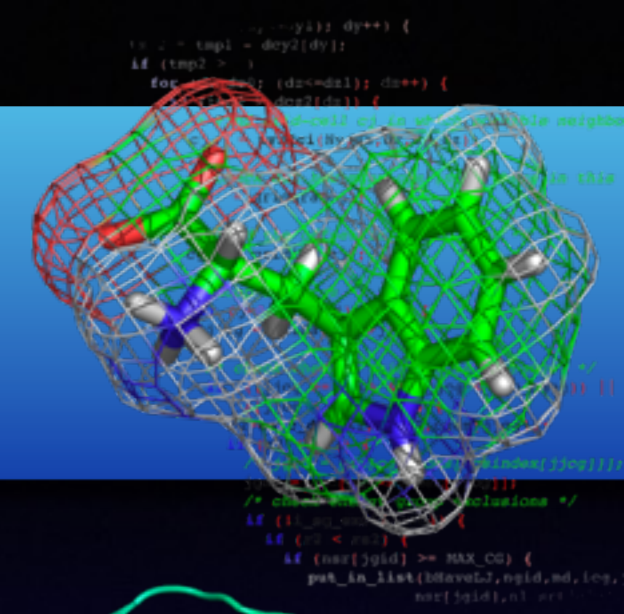


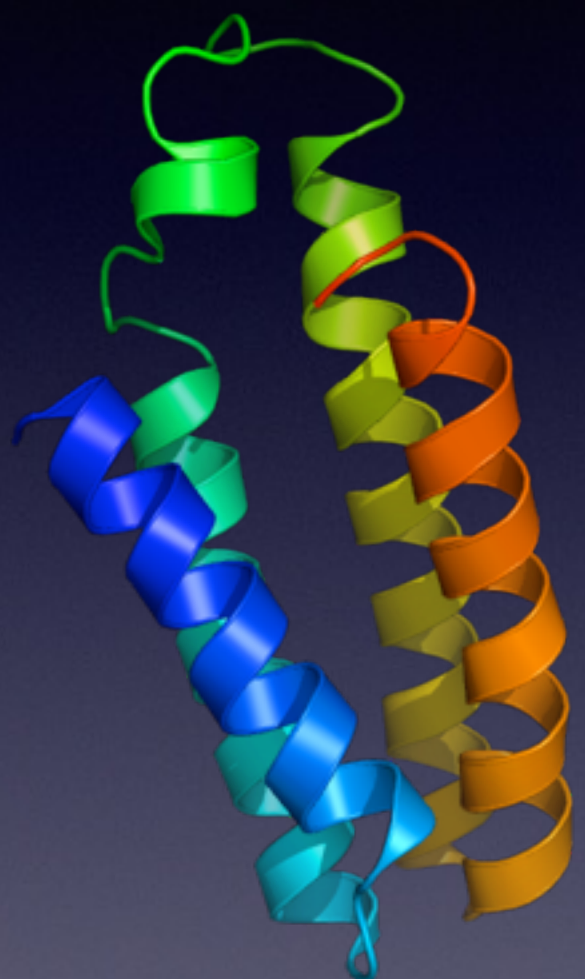
KcsA K-channel



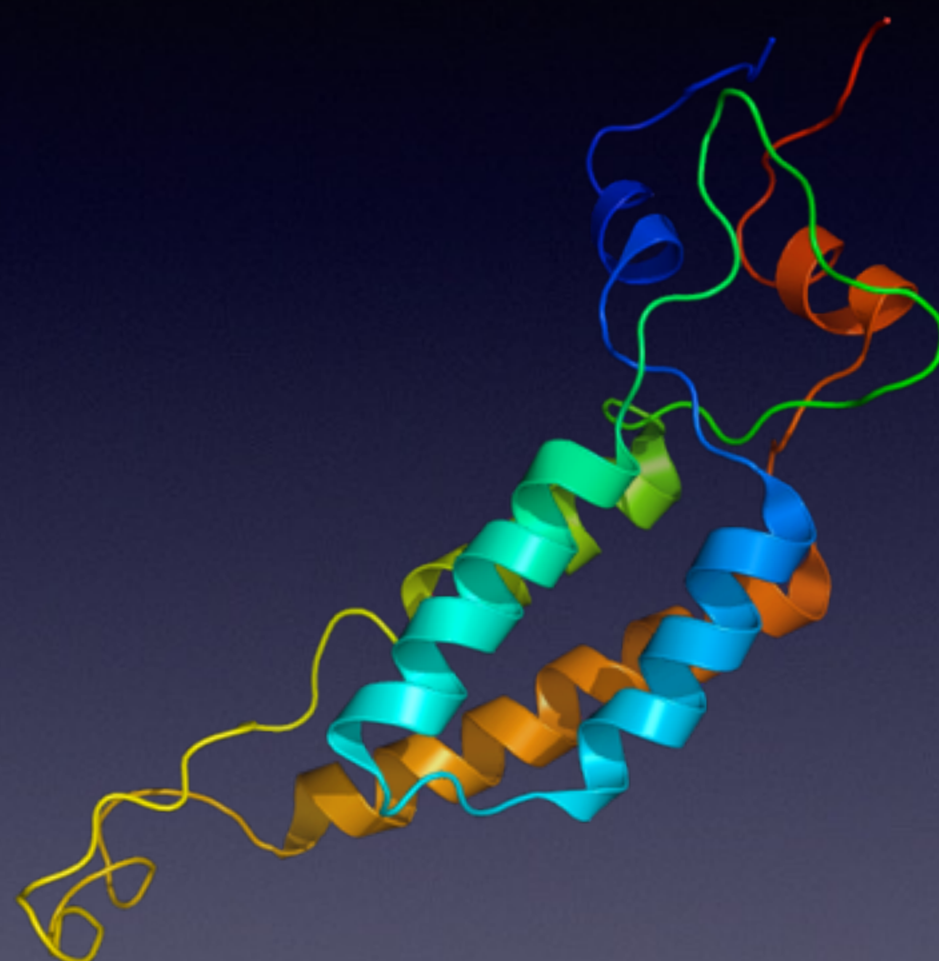
Hemerythrin

4-helix bundles

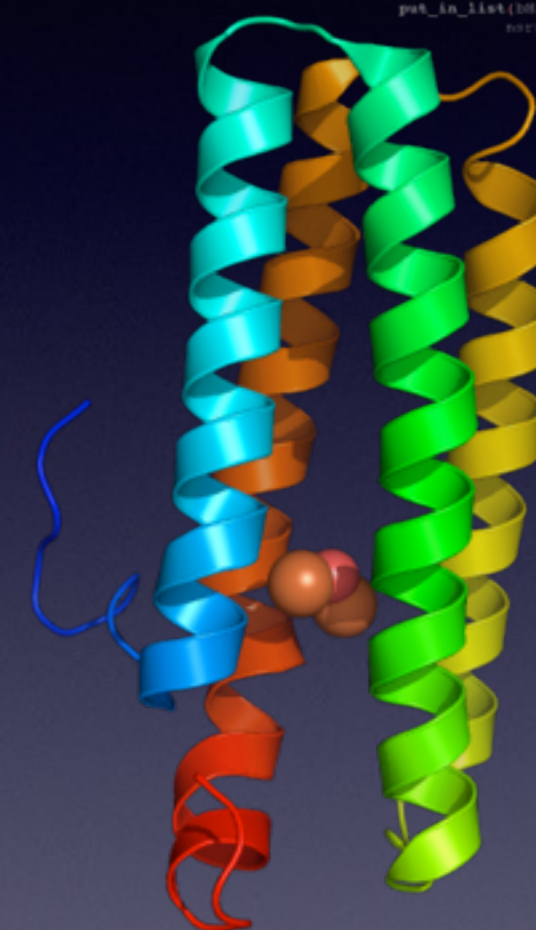




Cytochrome C



TMV Coat protein
(Tobacco Mosaic Virus)



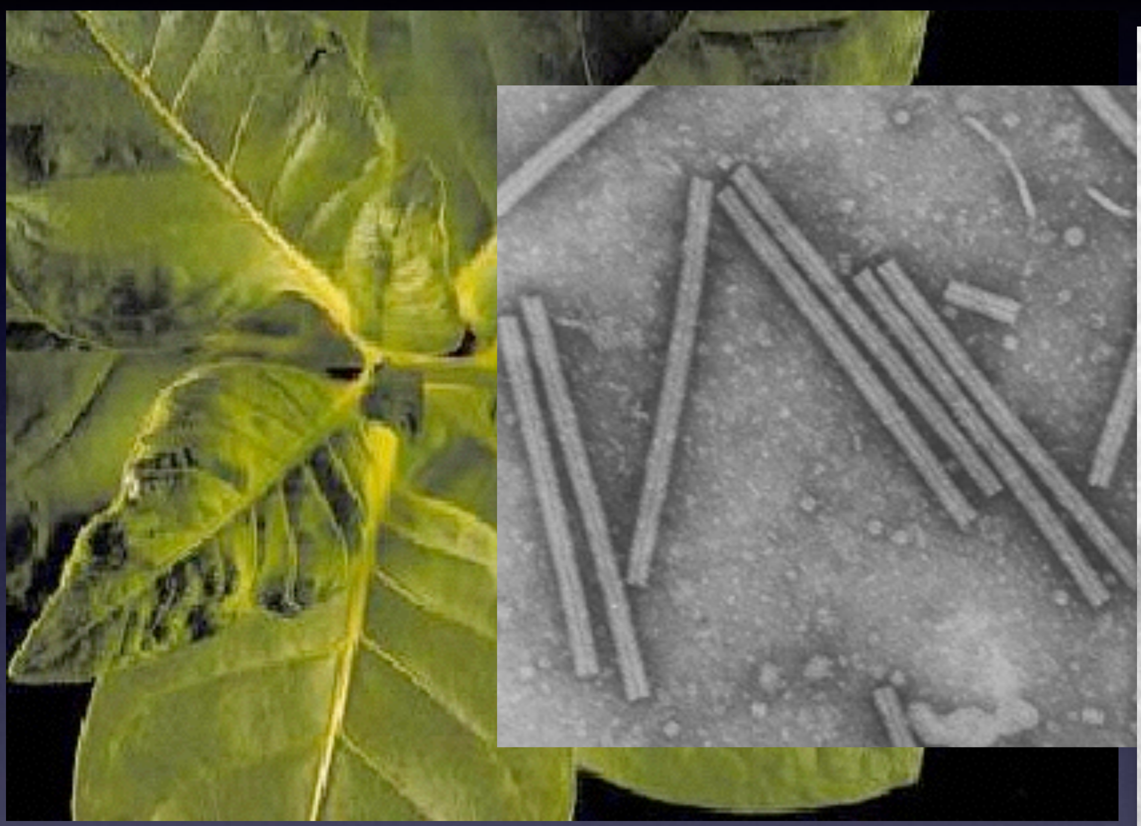
Hemerythrin
binding oxygen

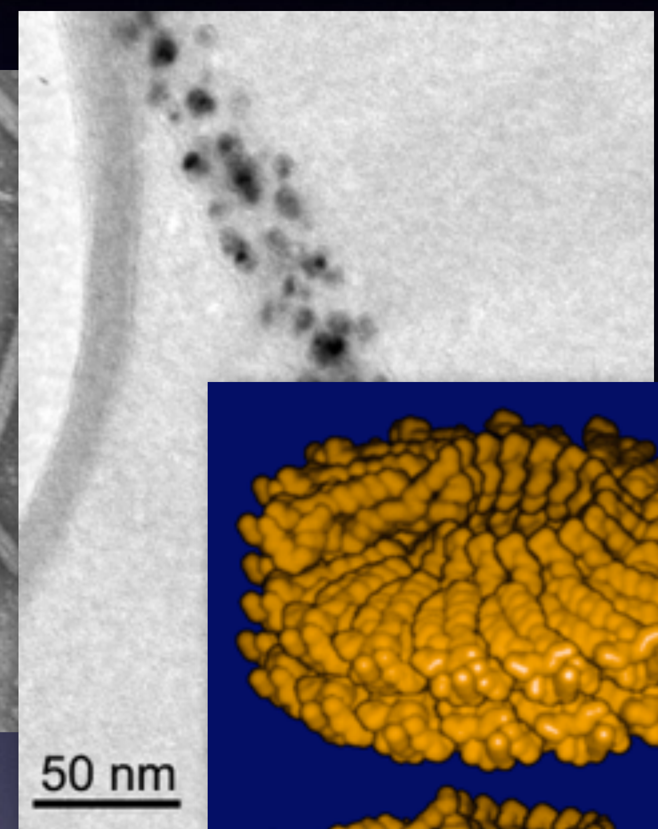
Note: Neighbor helices are anti-parallel!

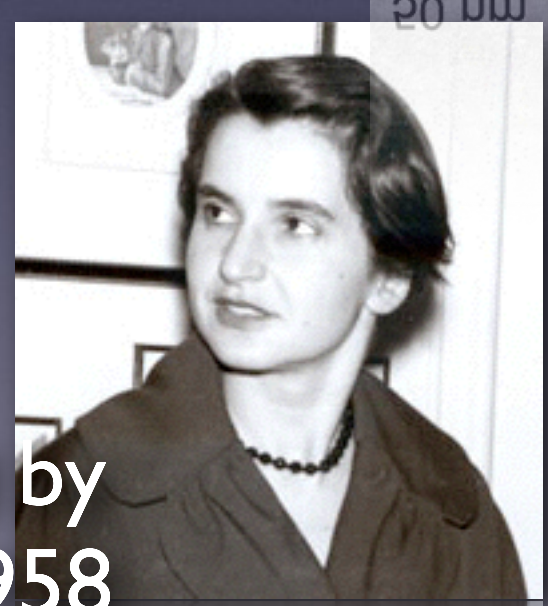
Cytochrome folds

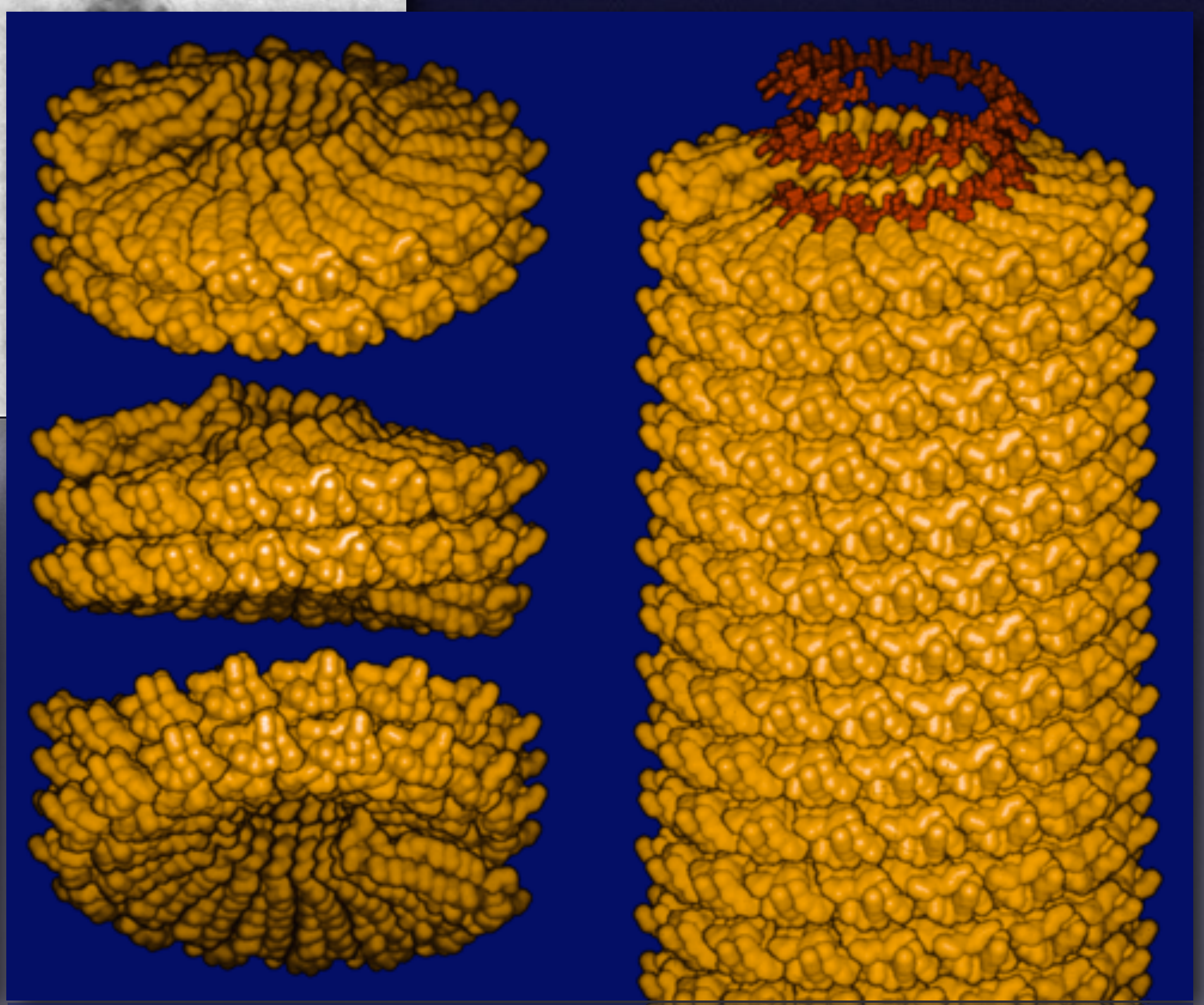
- Diverse domain, often carries out electron transport or binds metals
- Cytochrome domains are extremely abundant in *Shewanella Oneidensis MR-1*, a bacterium that can bind heavy metals

Tobacco mosaic virus



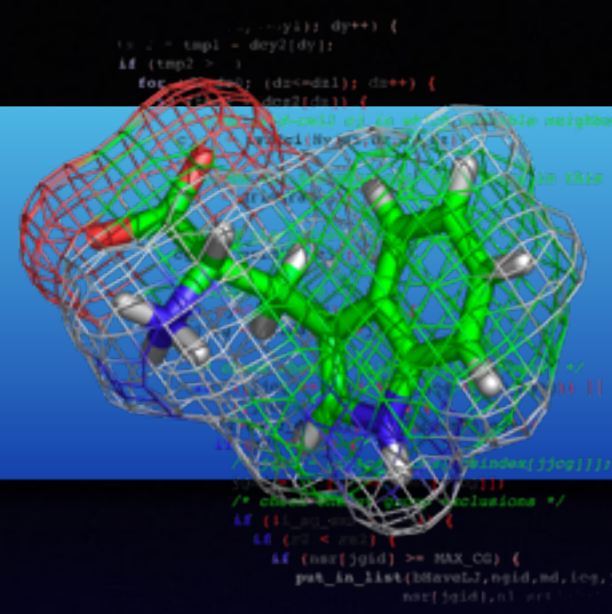


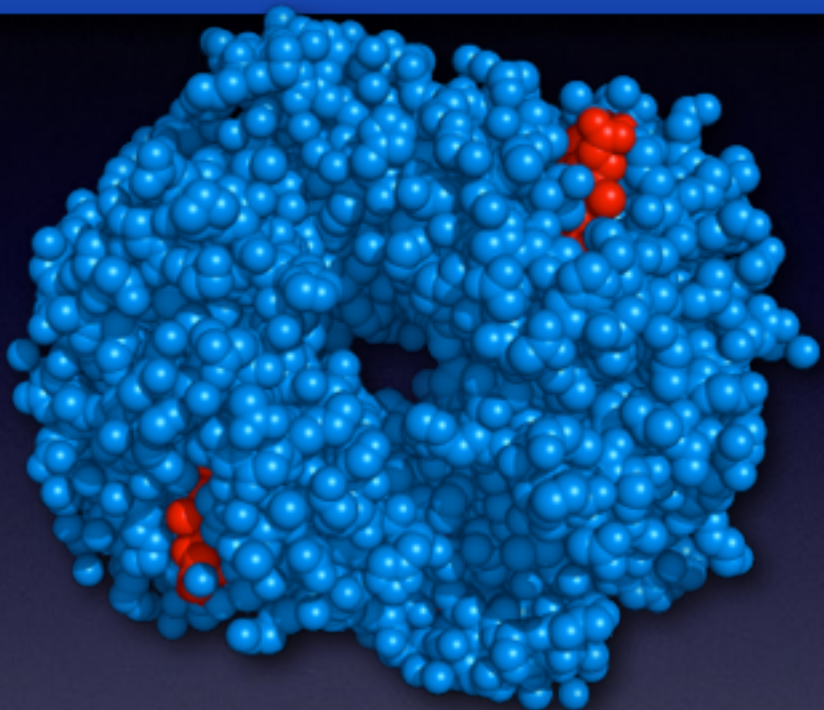


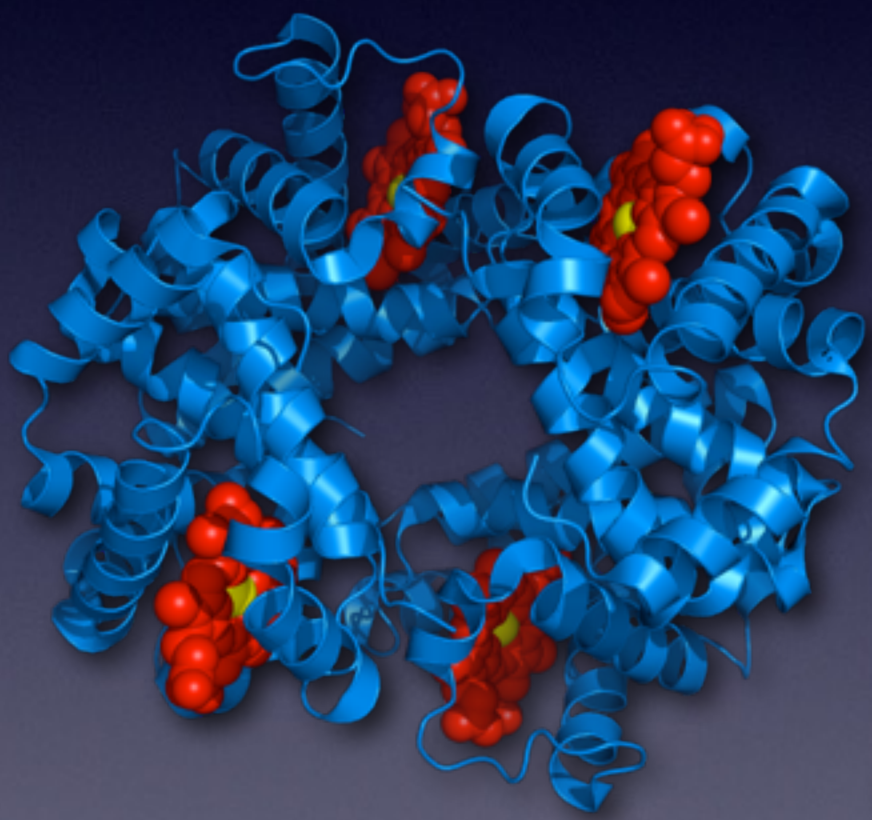


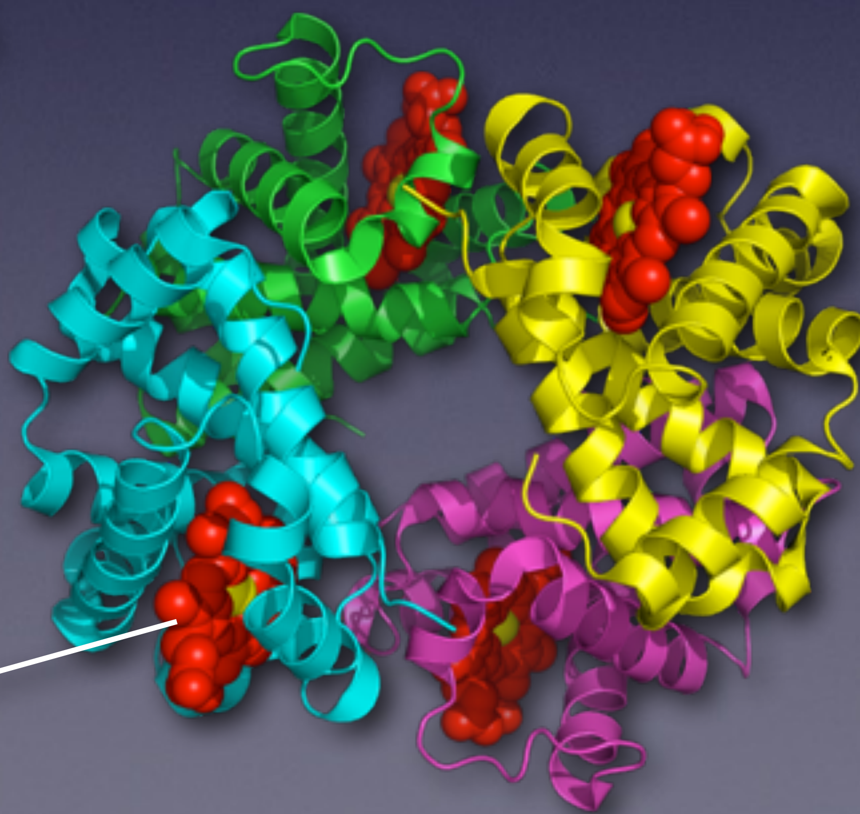
Structure proposed by
Rosalind Franklin, 1958

Hemoglobin

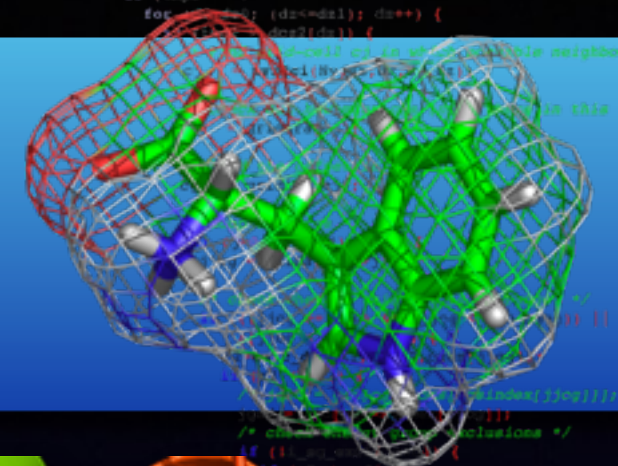


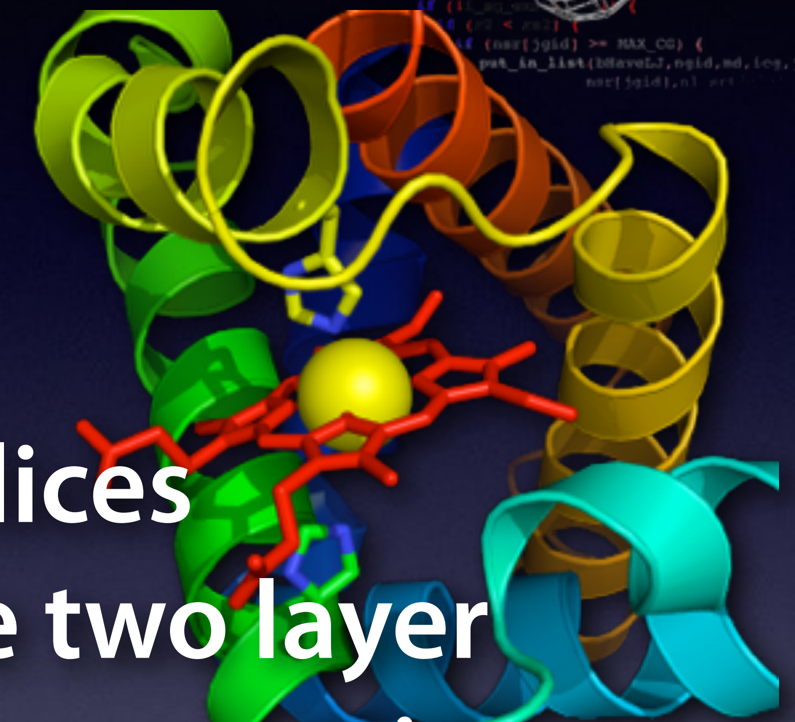






Hemoglobin

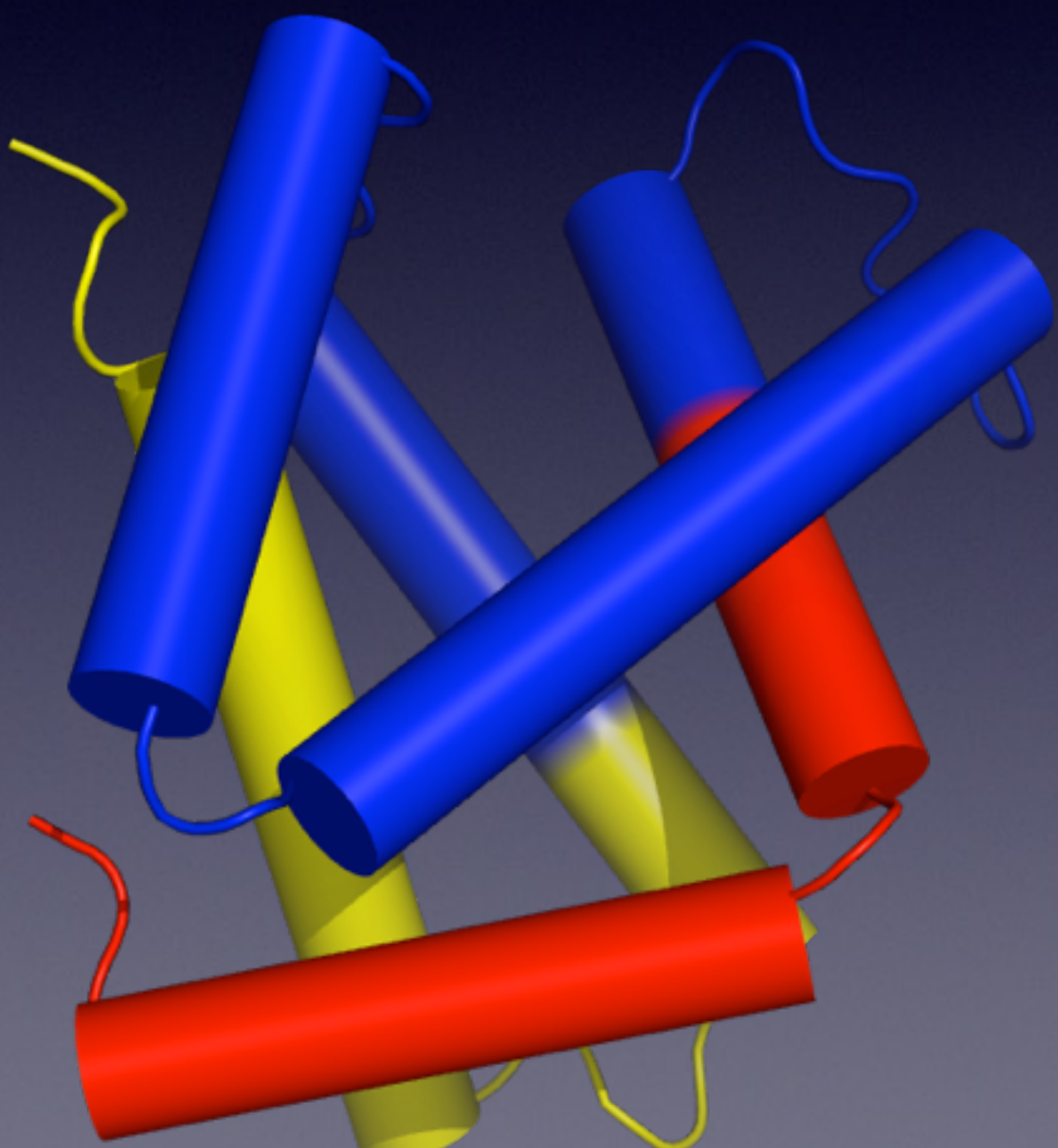




Single unit - globin fold

- 6 alpha helices
- Almost like two layer
- Binds heme group in pocket
- Myoglobin: similar monomer
- Why? Difference between Hemoglobin & Myoglobin?

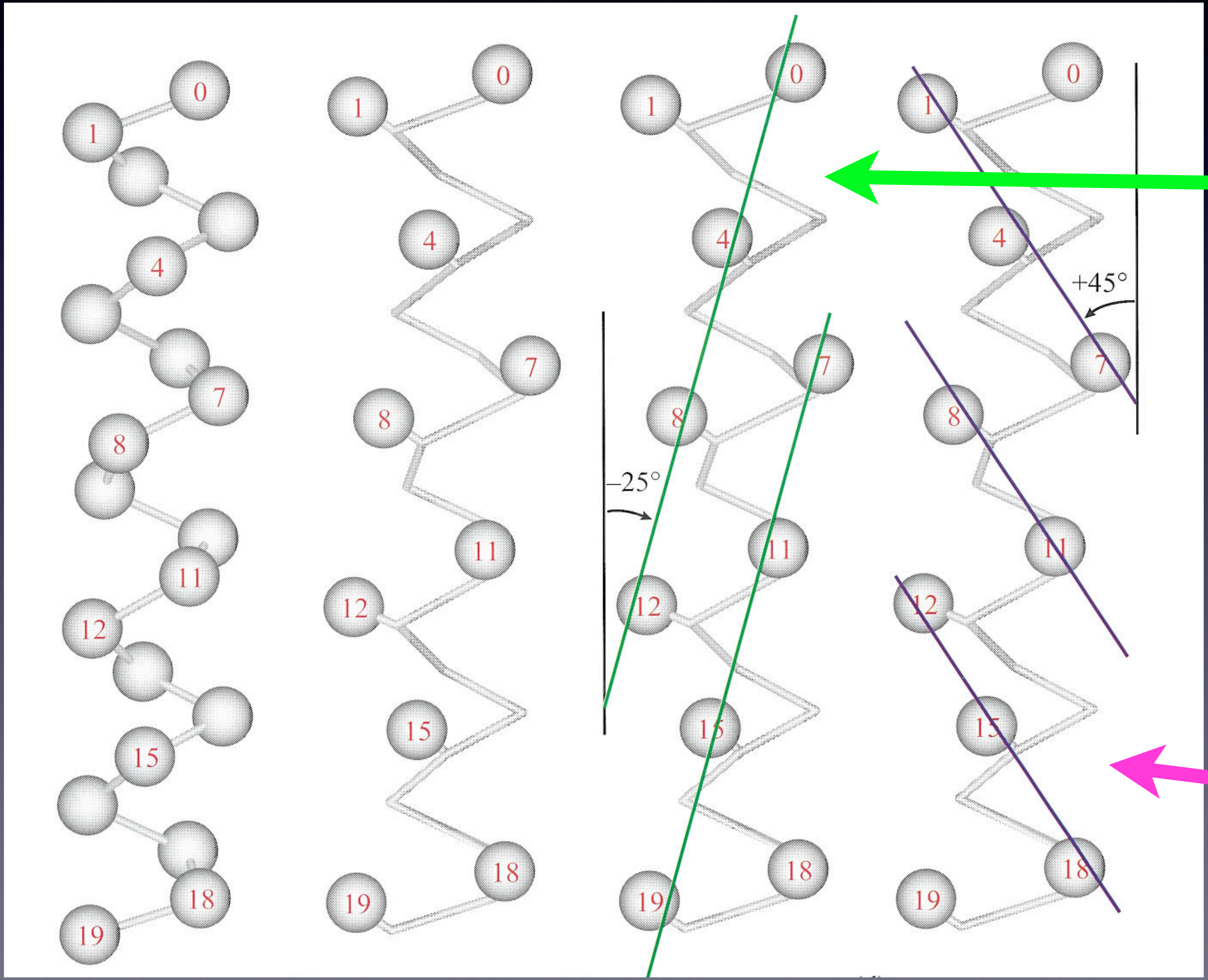
Hemoglobin intron/exons



Note that the exons do *not* correspond to domains or even secondary structure!

Exon 2 binds a heme group even when expressed isolated from exons 1 & 3, but it does not bind O_2. (Craik, Nature 1981)

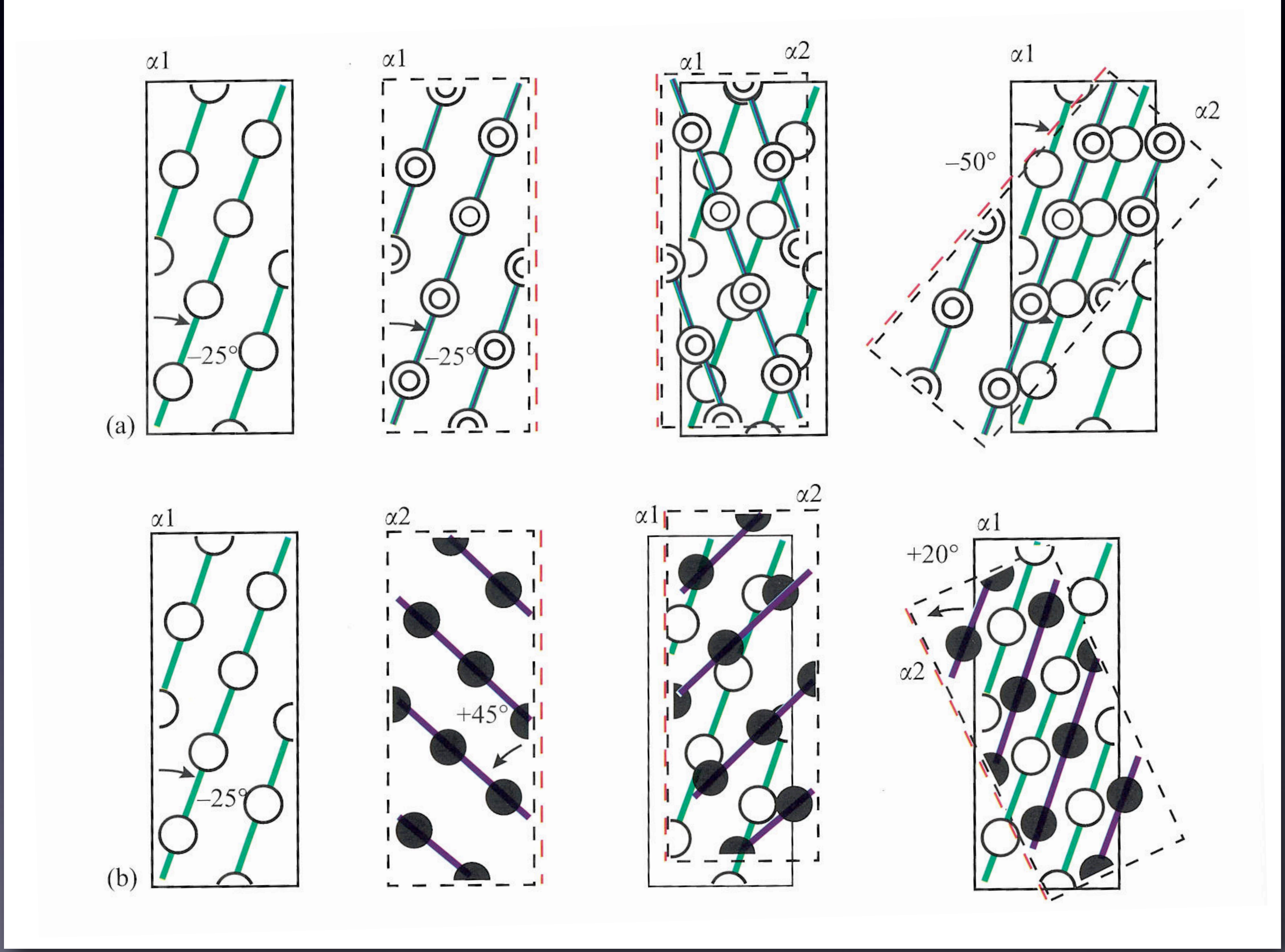
Helix ridges/grooves



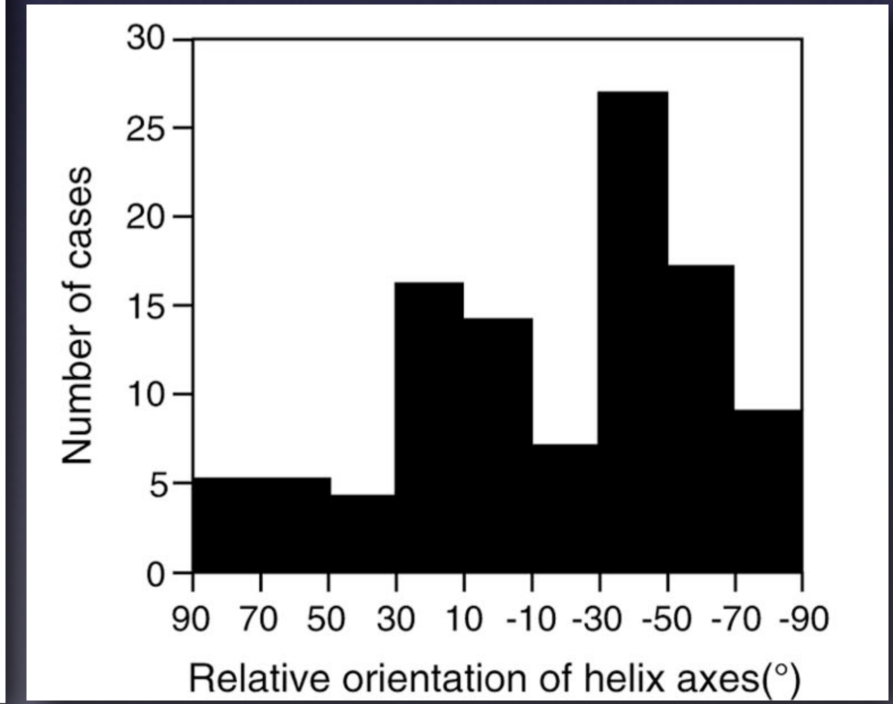
i,i+4 ridges
(-25° tilt)

i,i+3 ridges
(+45° tilt)

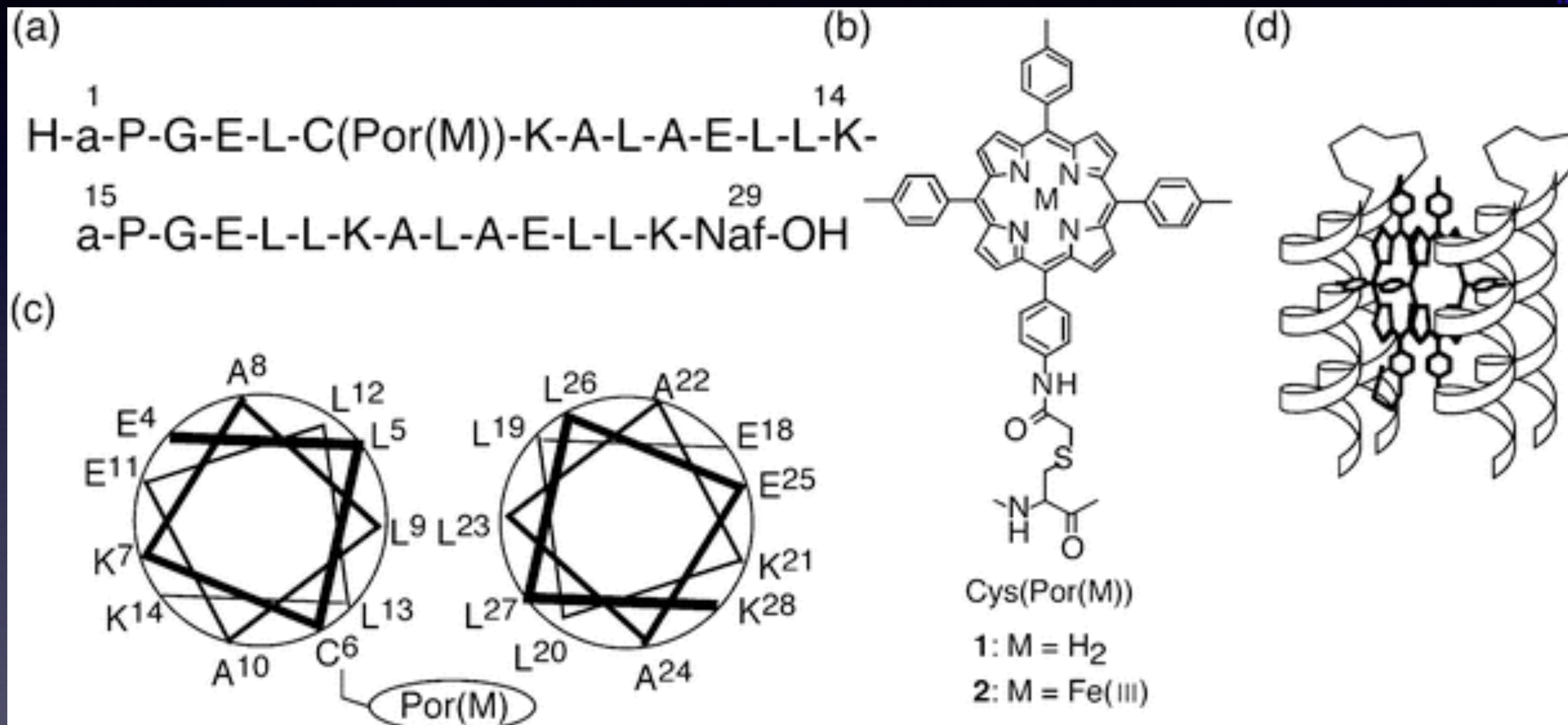
Helix pair packing



-50° or +20°



Hydrophobic moment



Many helices are amphiphilic

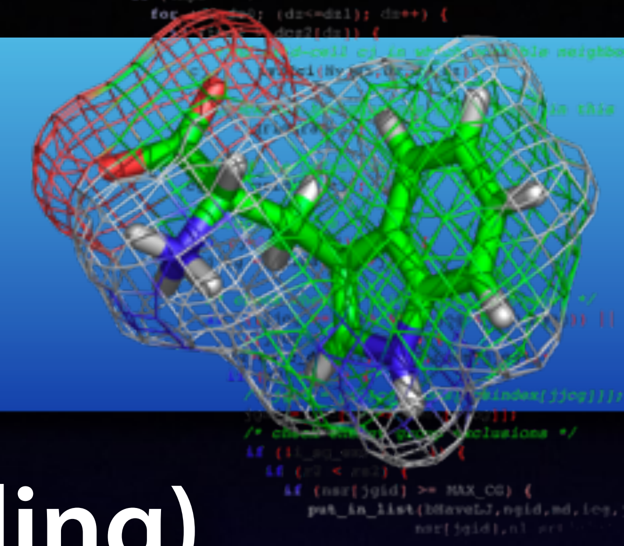
Used for protein design, e.g. heme binding (artificial blood!)

Amphiphilic helices are used as emulsifiers in many low-calorie products, but there is one big drawback - can you guess what?

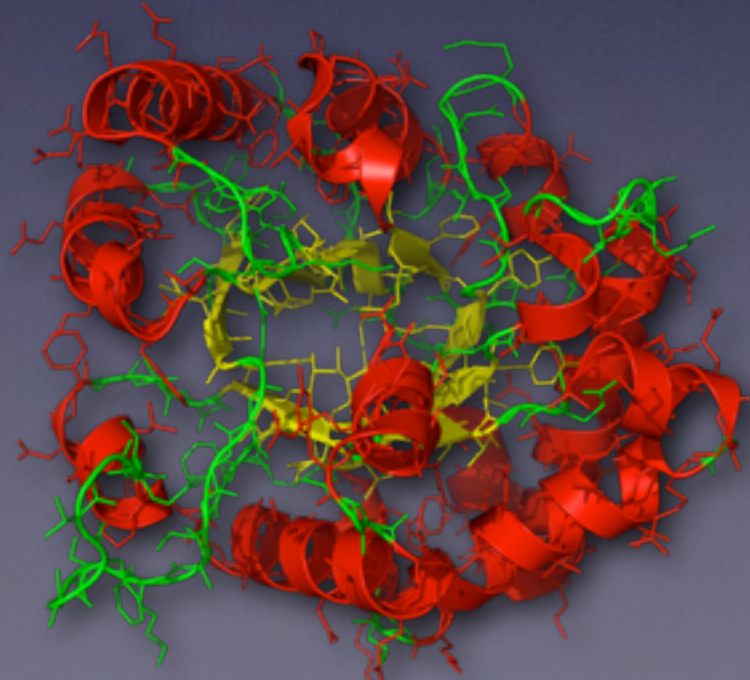
Mixed α and β domains

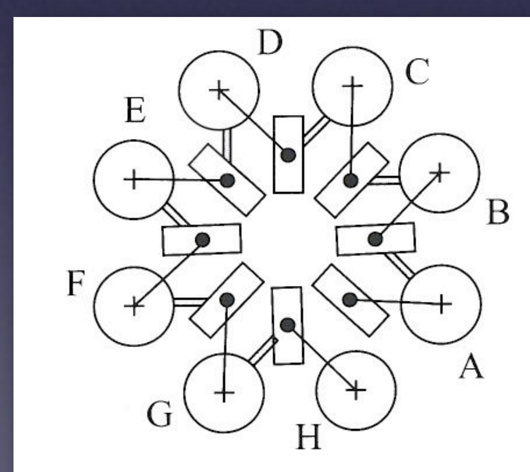
- Cannot mix single β-strand directly with α-helices
- Two types of mixed structure:
  - α/β: Alternating helix & sheet (how?)
  - α+β: Separate helix & sheet parts

α/β structure

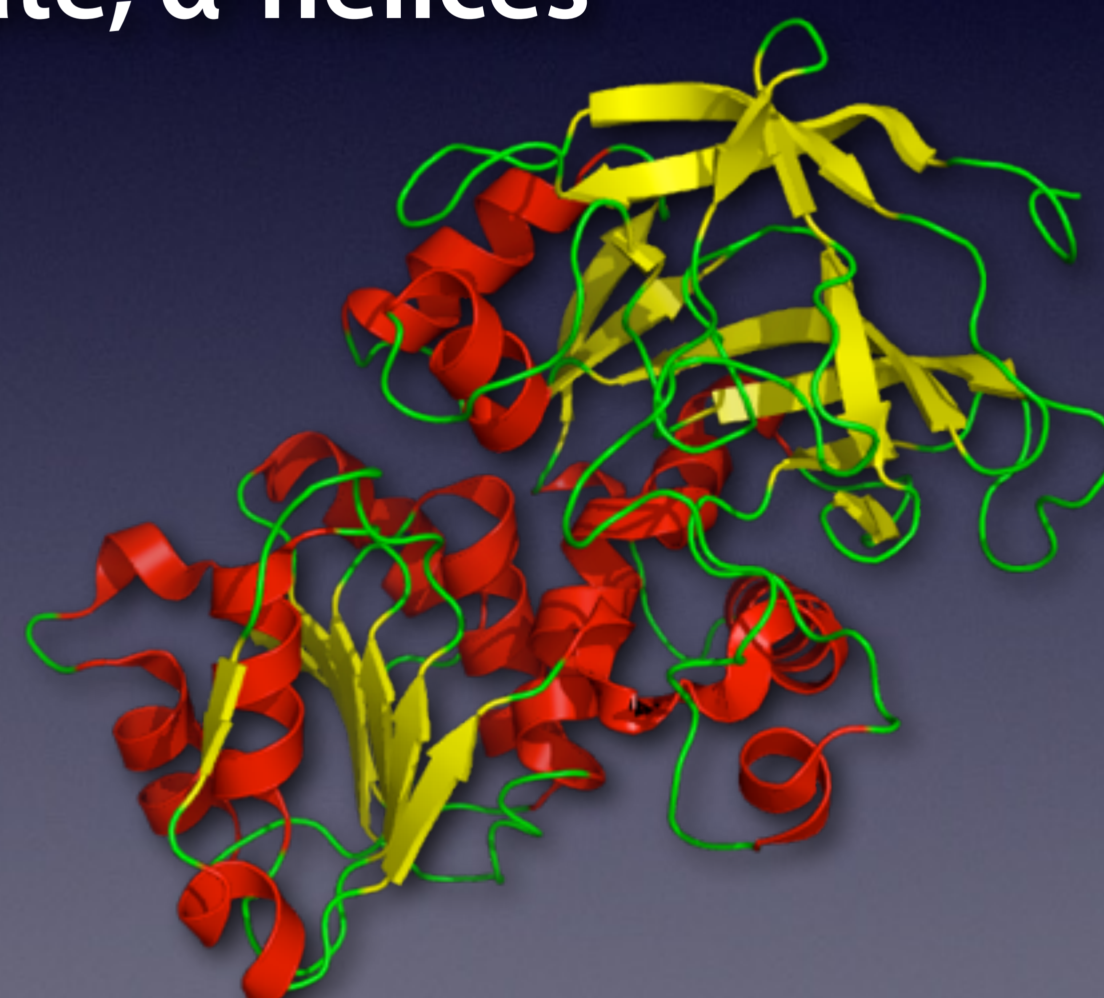


- Mainly parallel β-strands (h-bonding)
- Adjacent, but separate, α-helices
- βαβαβαβα...





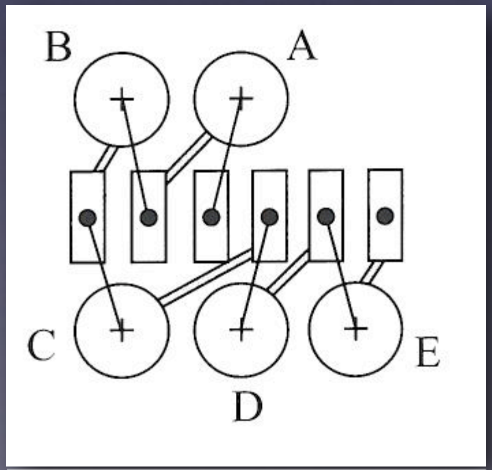
TIM barrel

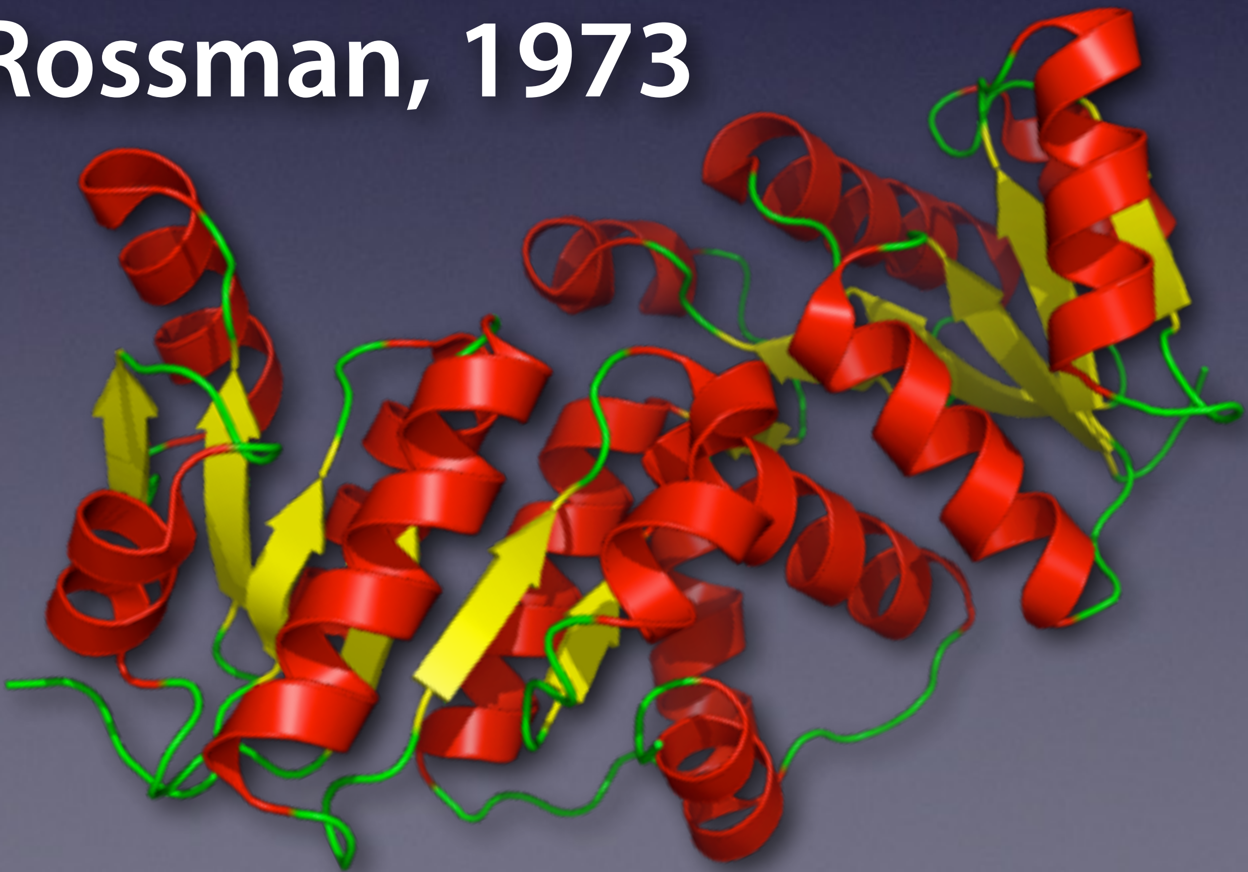


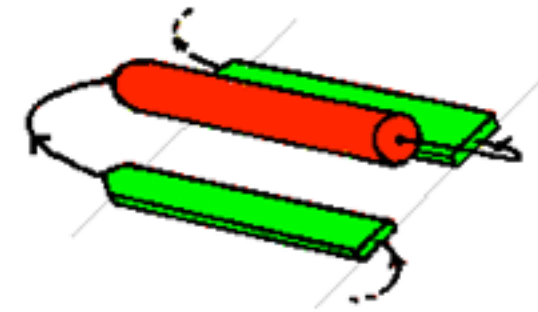
Alcoholdehydrogenase

Rossman folds

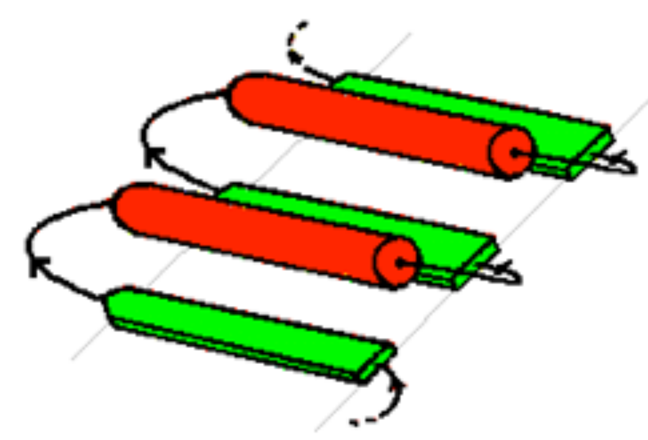
- βαβαβ motif
- Common for binding nucleotides, NAD cofactors
- Michael Rossman, 1973





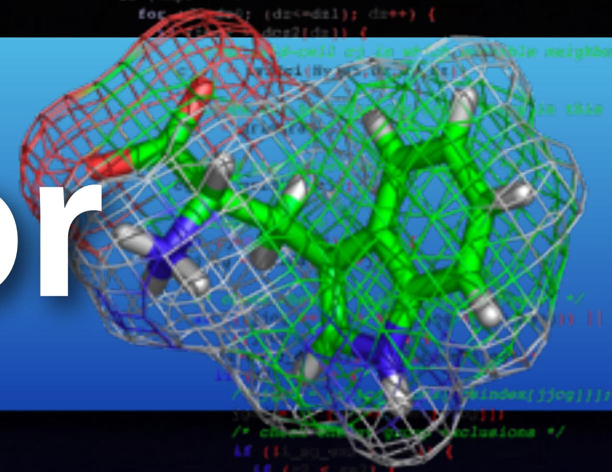


The right-handed beta-alpha-beta unit. The helix lies above the plane of the strands.

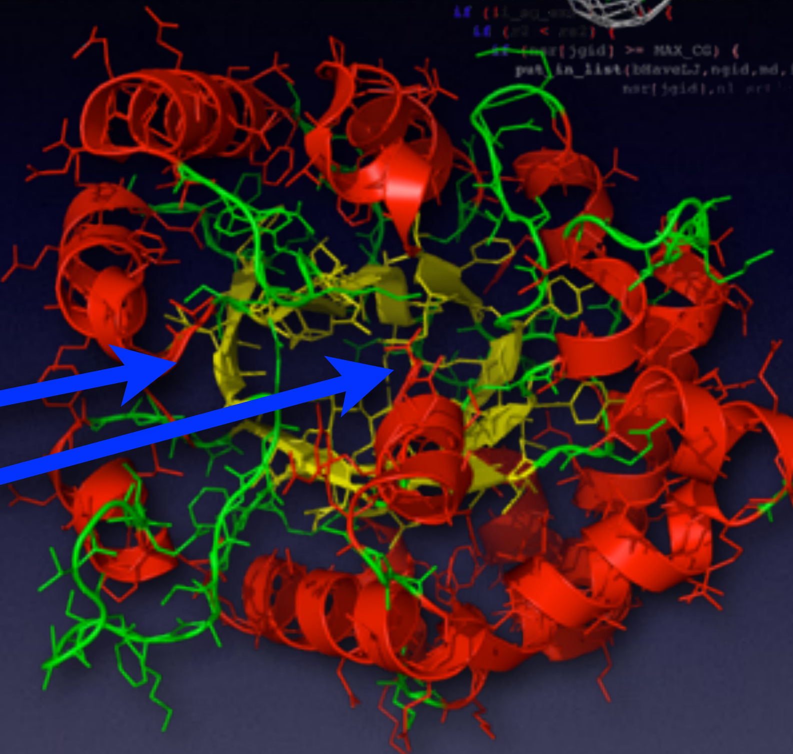


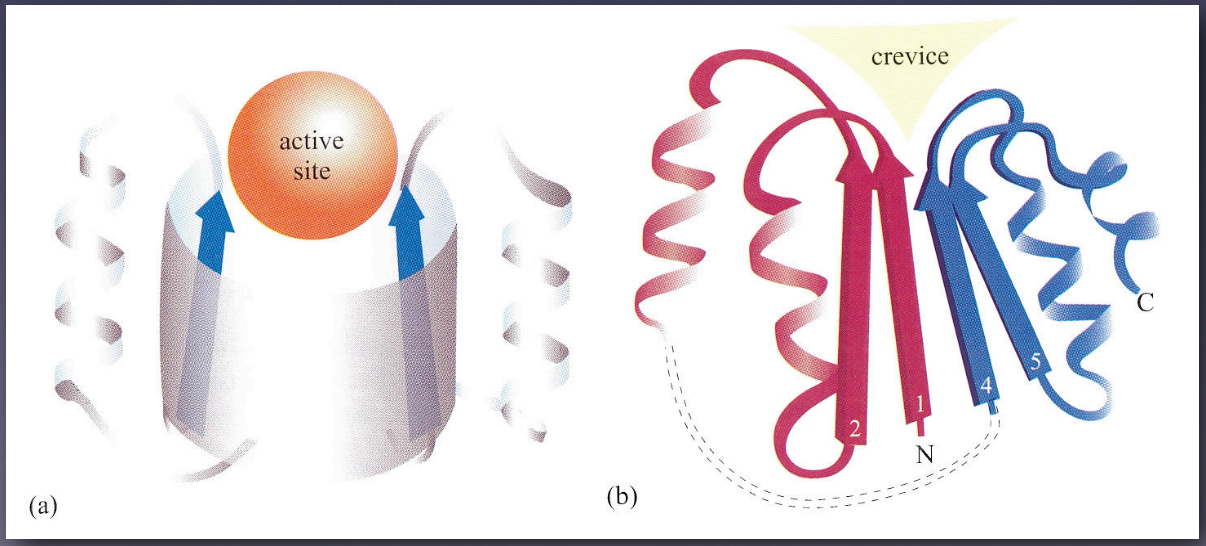
The Rossman fold

α/β structure interior



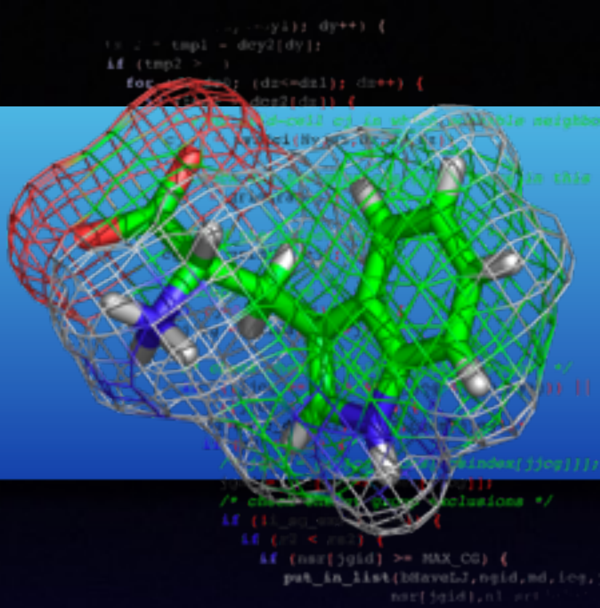
- Normally two hydrophobic cores
- TIM:
  - Between sheets and helices
  - Inside sheets
- Rossman:
  - On both sides of sheet



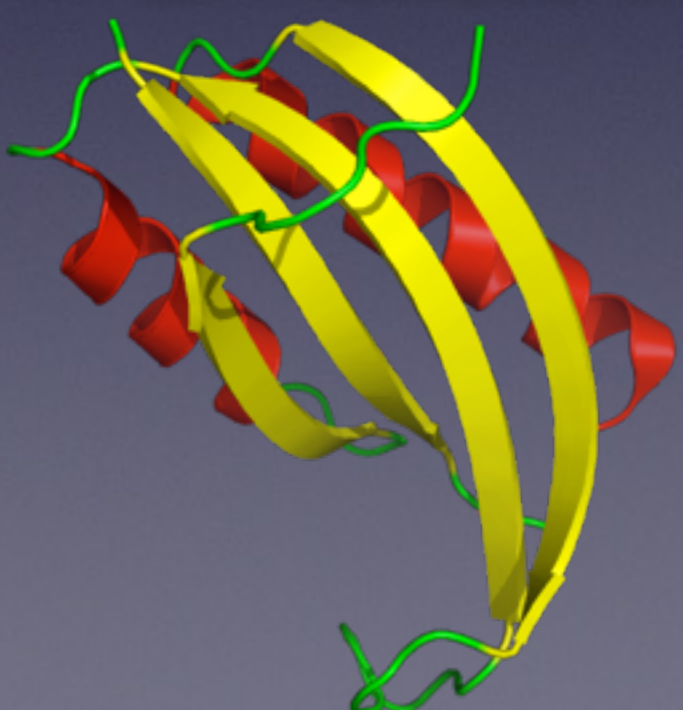


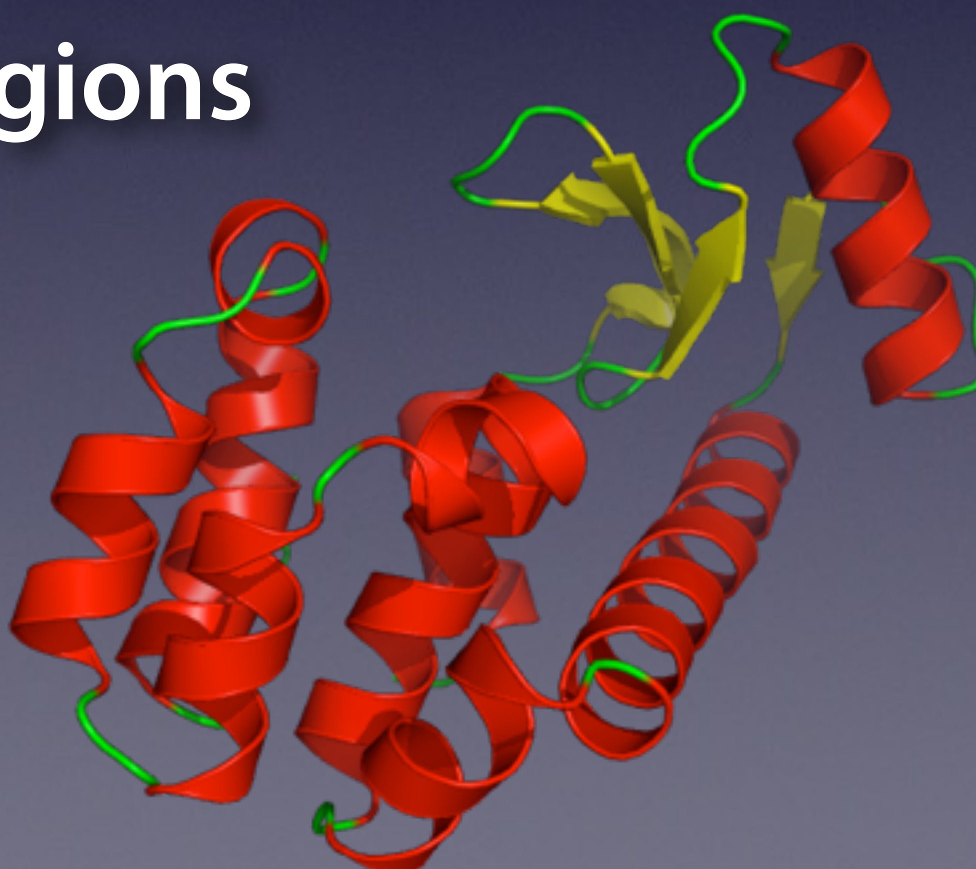
Common as binding sites

α+β structure



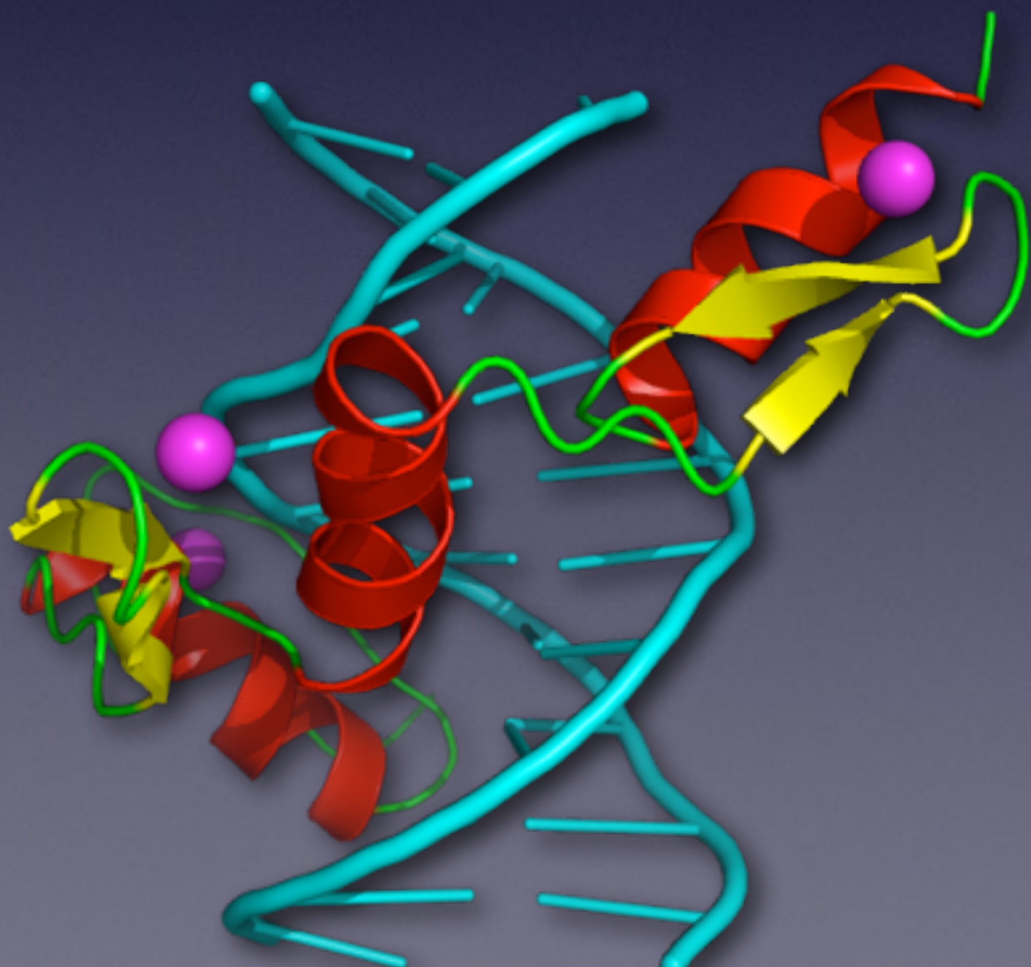
- Mainly anti-parallel β-strands
- 1: Alternating pairs of helices & strands
  - ββααββββααββββααββ...
- 2: Helices in separate regions



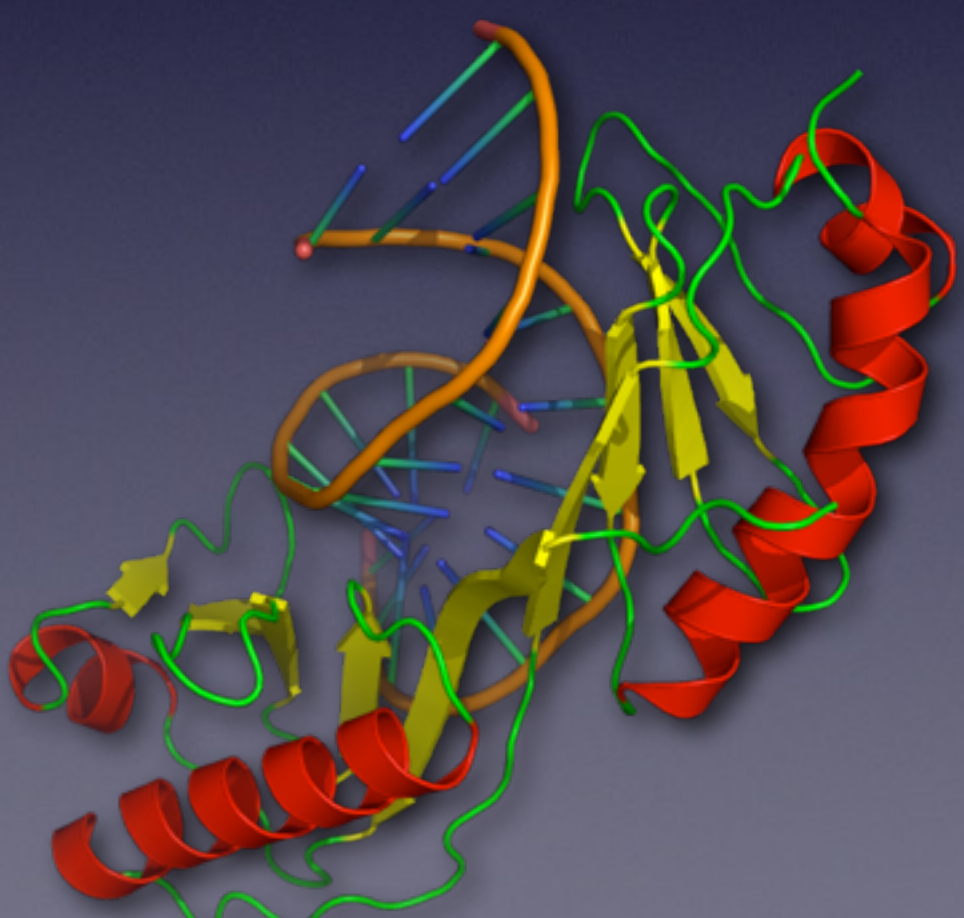


DNA-binding α+β proteins

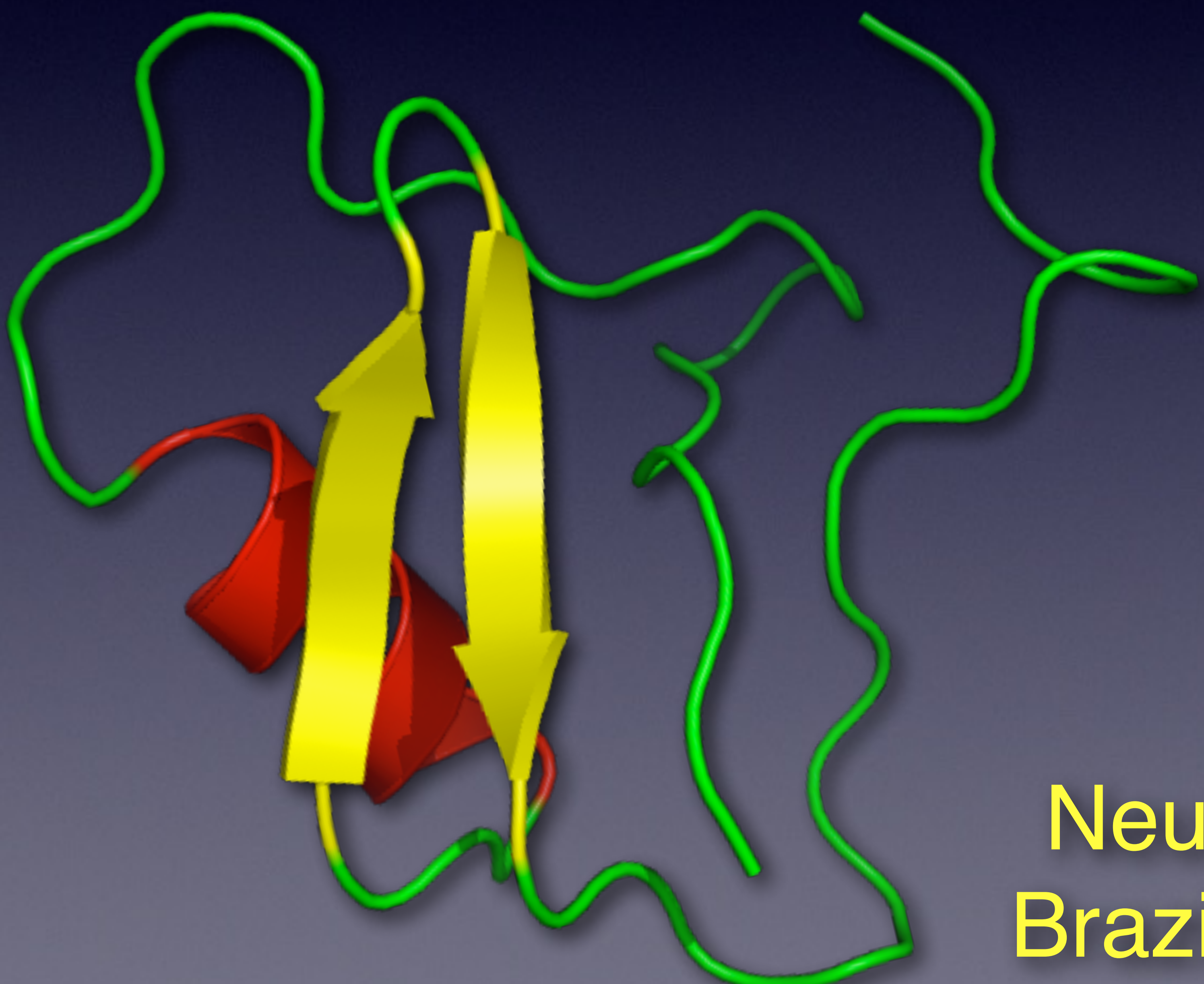
Zinc finger: stabilized by bound Zinc ion DNA-binding motif



TATA-binding protein Binds to DNA 5'-TATAA-3' to initiate transcription: separates DNA strands

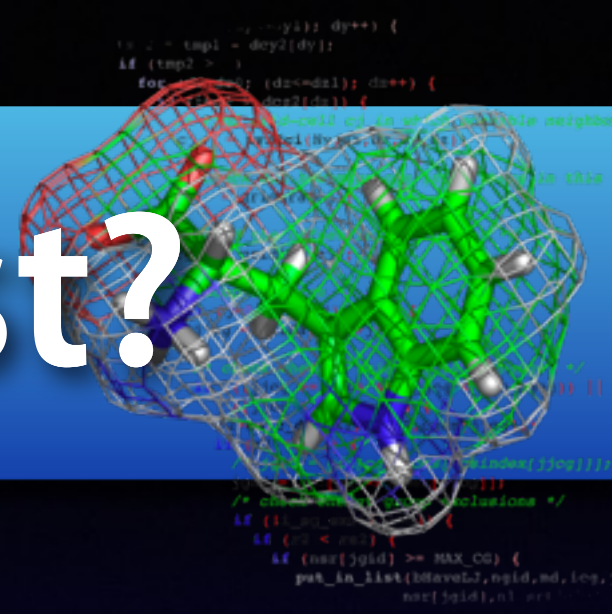


Irregular structures?

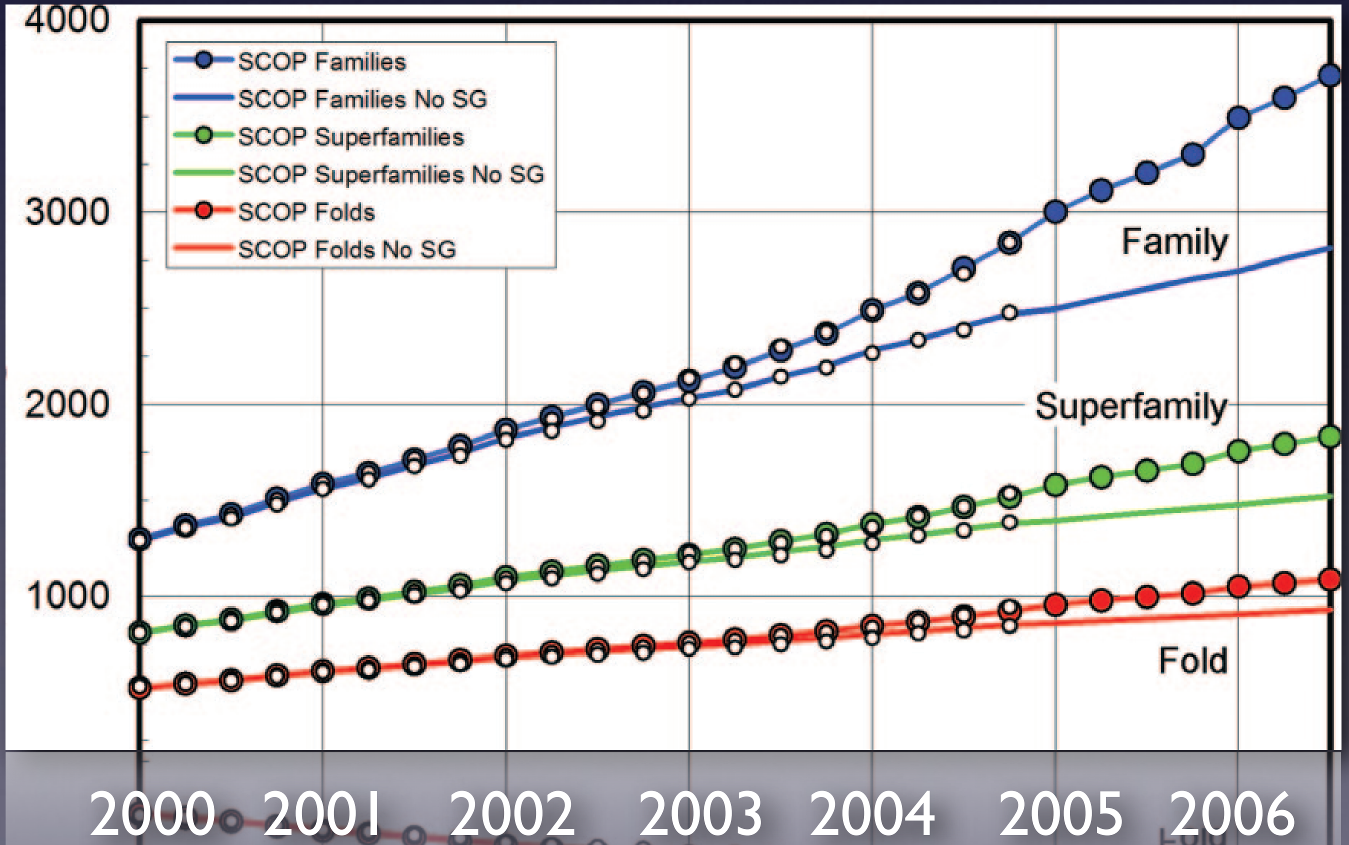


Neurotoxin from
Brazilian scorpion

How many folds exist?

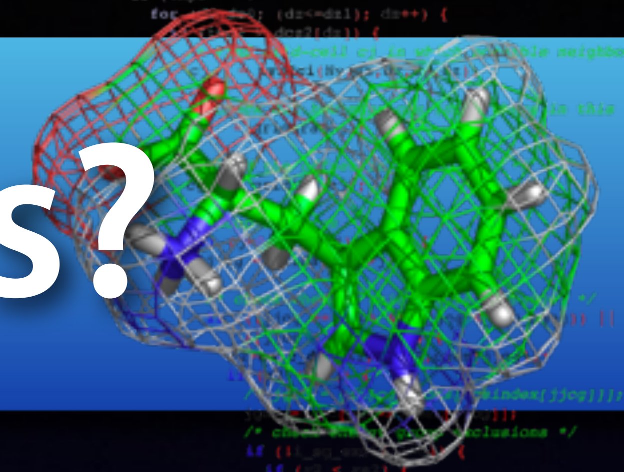


- Even proteins with unrelated sequences seem to "reuse" the basic structural building blocks
- Chotia 1992:
  1000 folds?
- Today:
  1300 & growing (slowly)
- Fold space is limited!

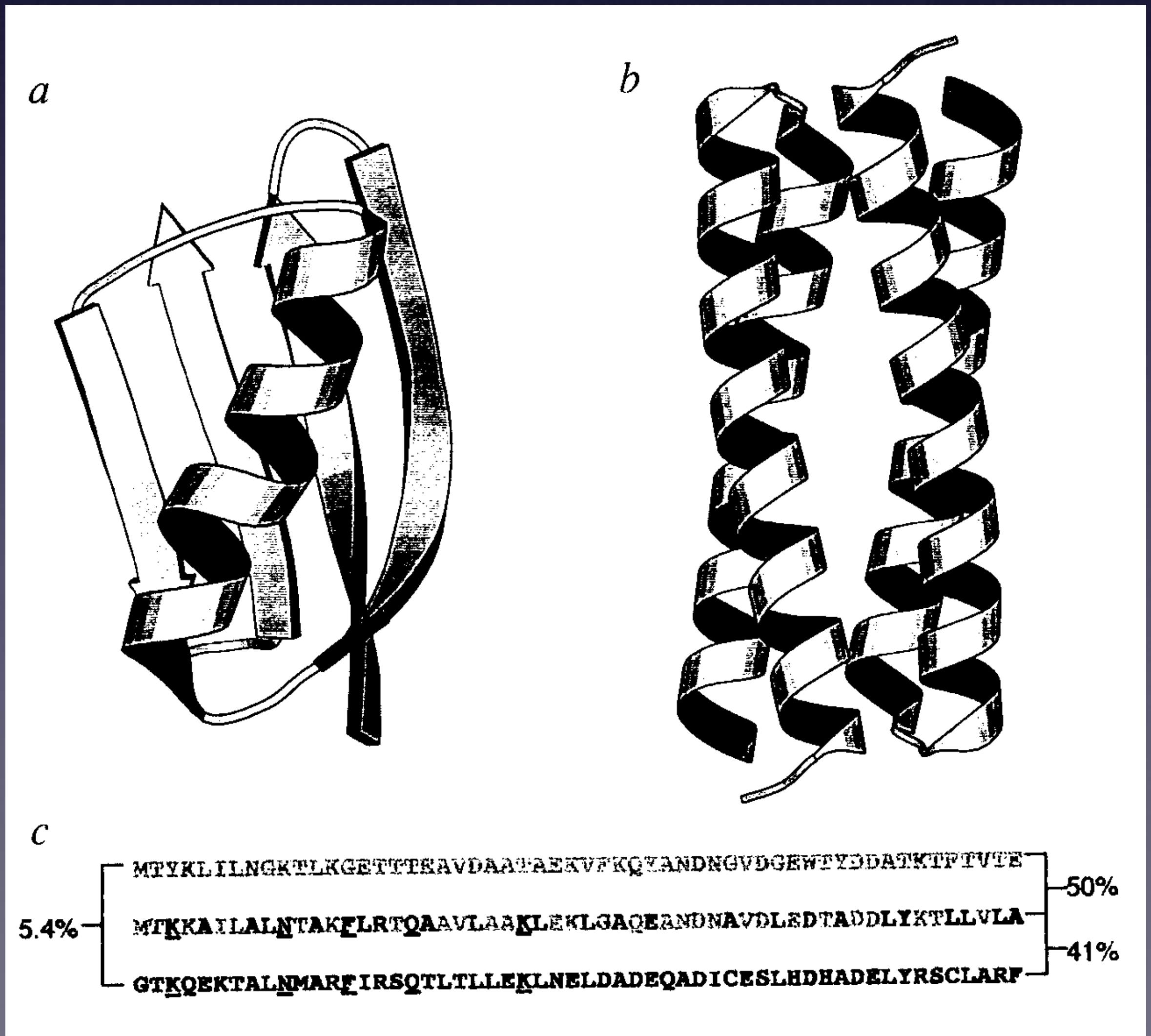


Levitt, 2007

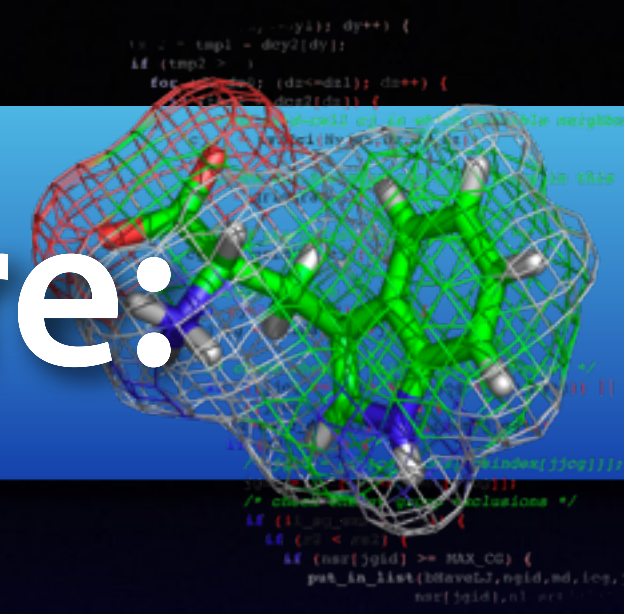
How stable are folds?



- Sequence determines tertiary structure
- Folds seem to be quite stable against mutations
- We “know” that similar sequences share structure
- Protein alchemy by Lynne Regan, 1997: By changing less than 50% of residues, it was possible to turn a β-protein into a four-helix bundle α-protein...



Summary of structure:



- Book chapters 11,13 & 14
- Fibrous: Collagen, Silk, Keratin, Elastin
- β proteins: Greek keys, dimerization
- α-helix proteins: Globin and other folds
  - Larger diversity than β-proteins
- α/β proteins: TIM barrel, Rossman fold
- α+β proteins: Nuclease, Zinc fingers
- Supersecondary structure