Molecular Modeling 2018--lecture 6

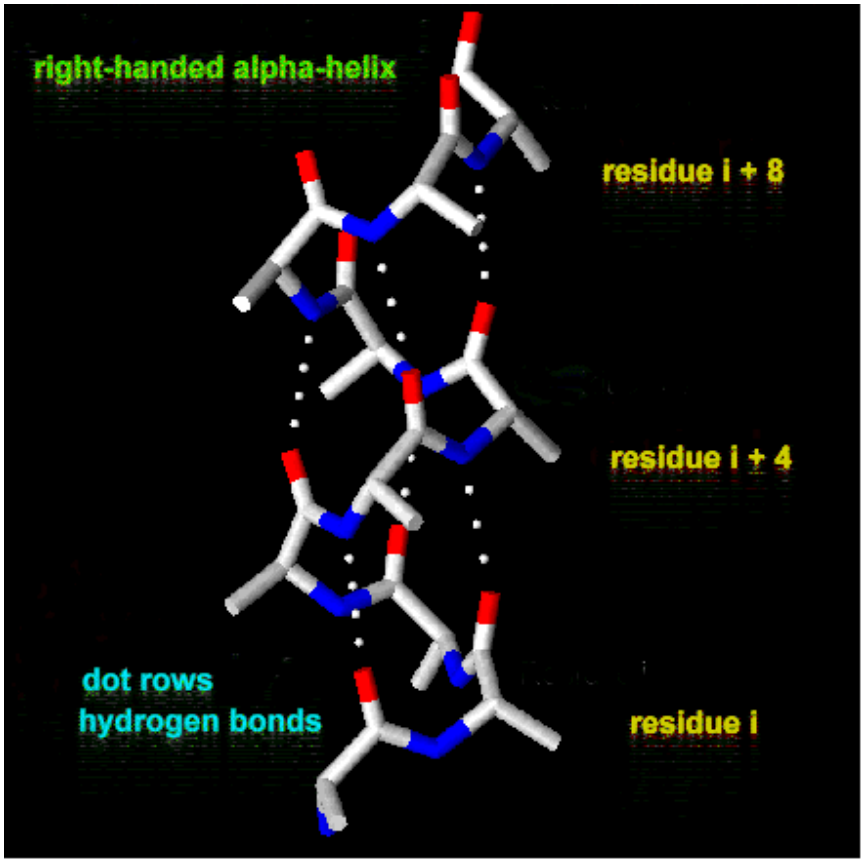
Secondary structure

Secondary structure prediction

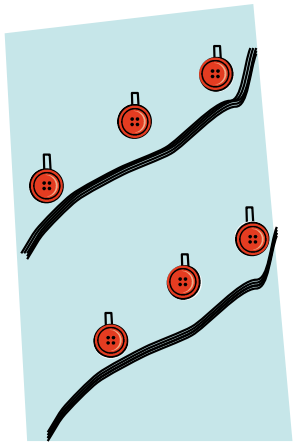
Alpha helix

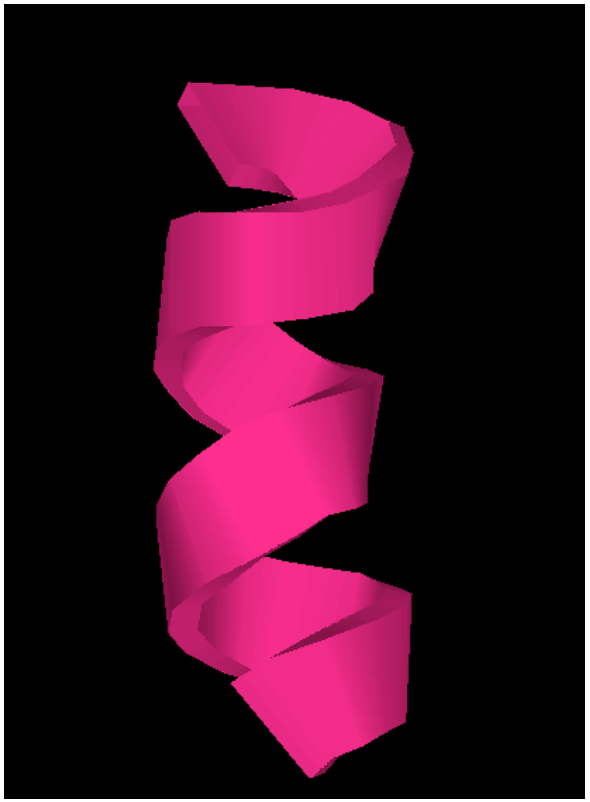
Right-handed helix. H-bond is from the oxygen at *i* to the nitrogen at *i+4*. α-helices have an overall dipole because the H-bonds are all in the same direction. Must be > 3 residues.

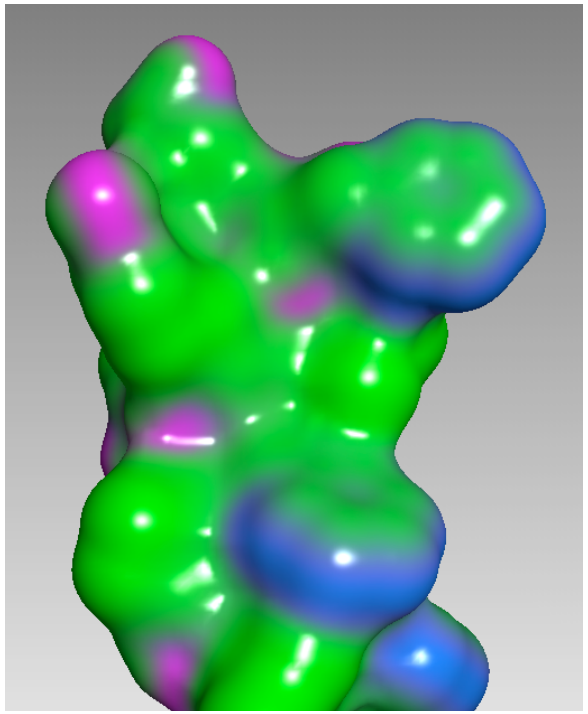
H-bond rule (NH->O): i->i+4



dipole

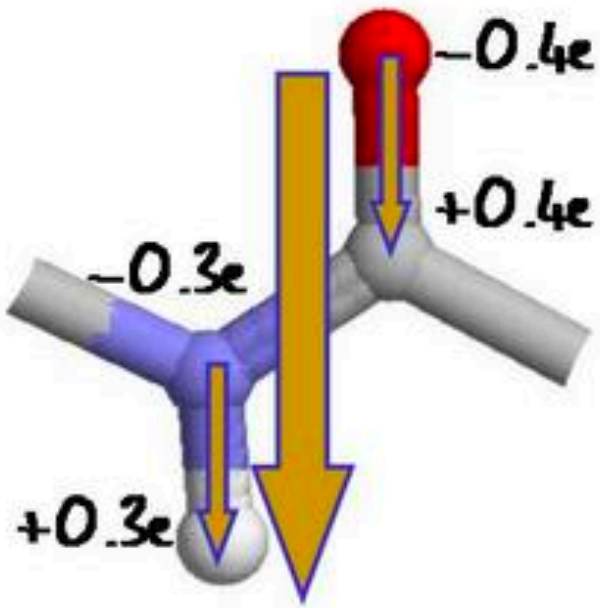






Helices are not as "cylindrical" as the cartoon suggests.

ALPHA-HELIX DIPOLE 1



- The peptide group has a strong dipole moment due to partial charges on NH and CO groups.

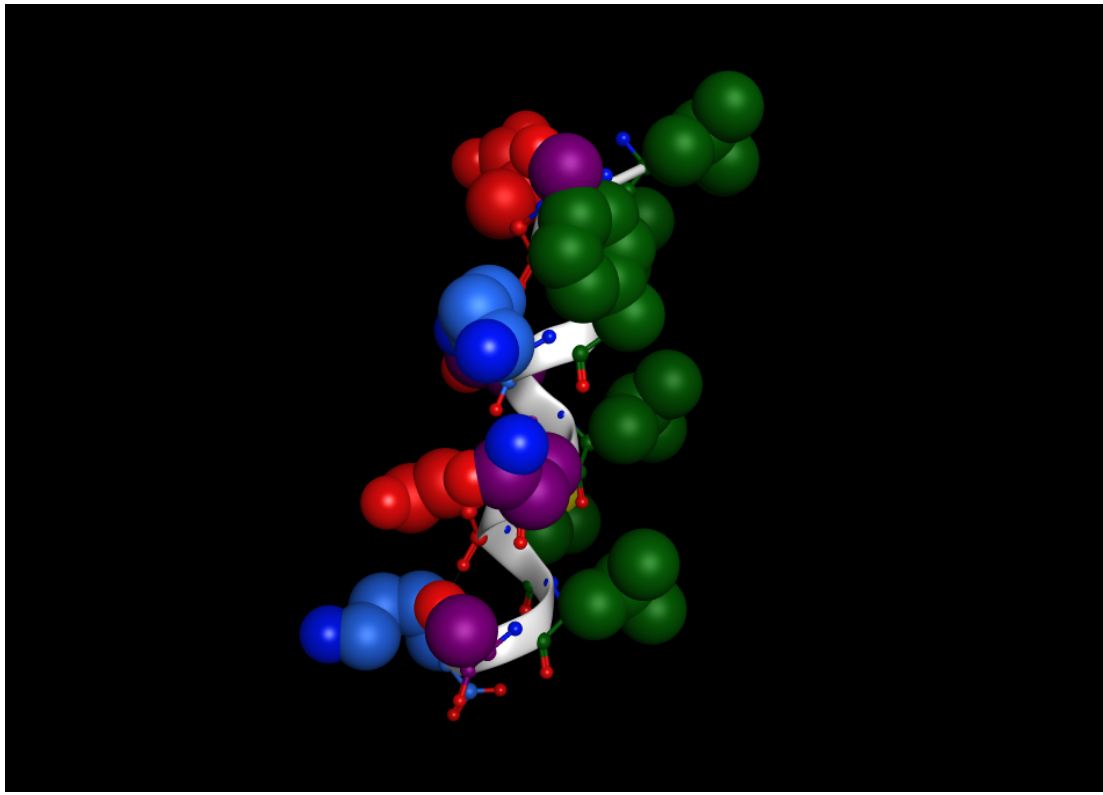
https://web.stanford.edu/class/sbio228/public/lectures/lec2/Lecture2/The_Alpha_Helix/images/Alpha-Helix_Dipole_1.jpg

Sequence patterns for alpha helix

- nppnnpp,
  where n = non-polar, p = polar
- Example:

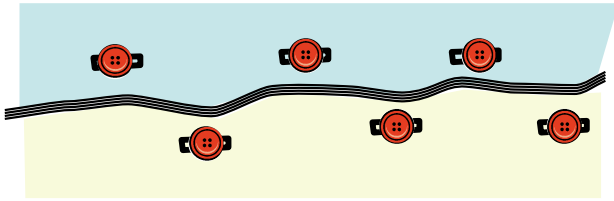
  LSELFKNLQDMLSK

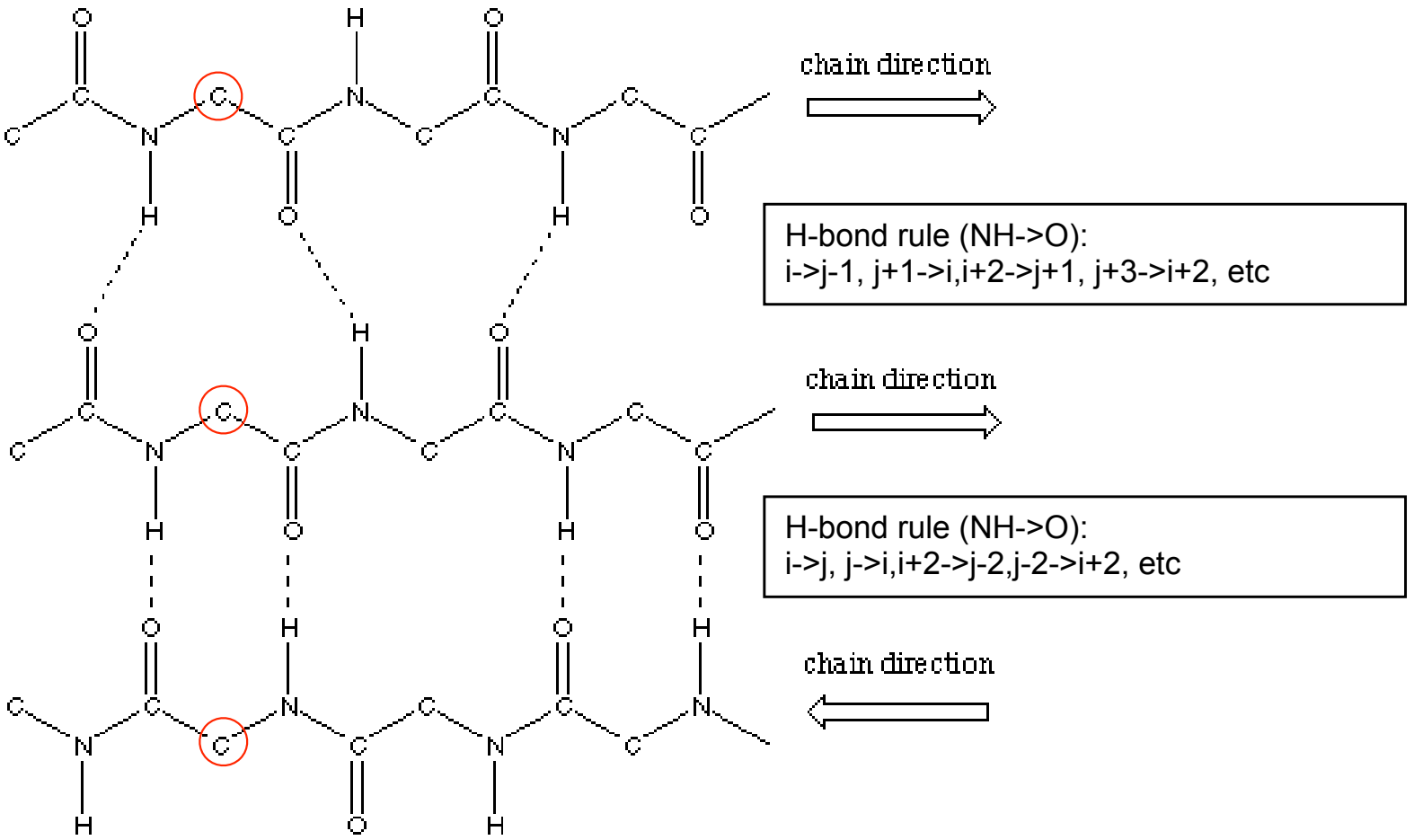
Hydrophobic all on one side



The helix is held together by the hydrophobic effect.

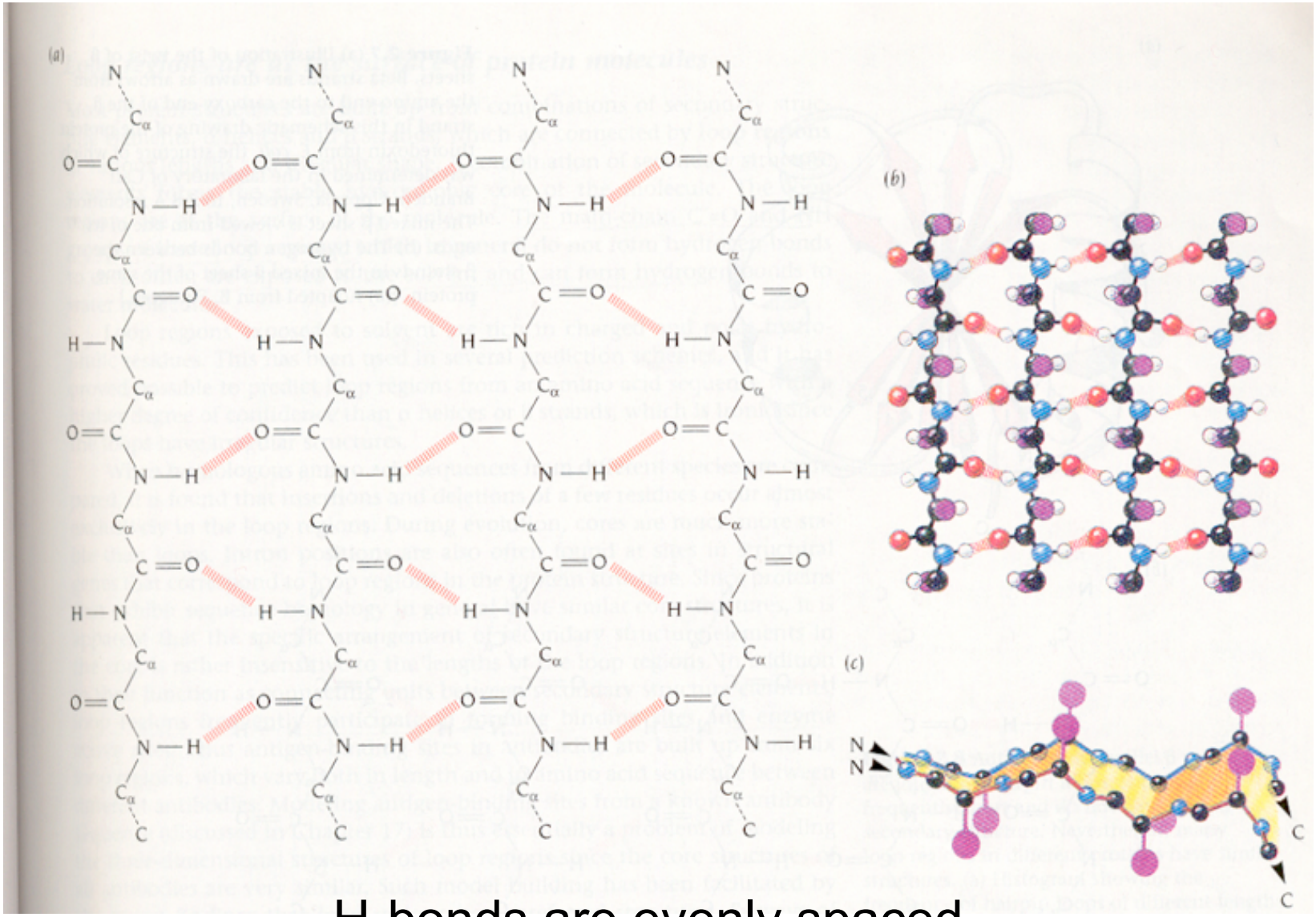
beta sheets





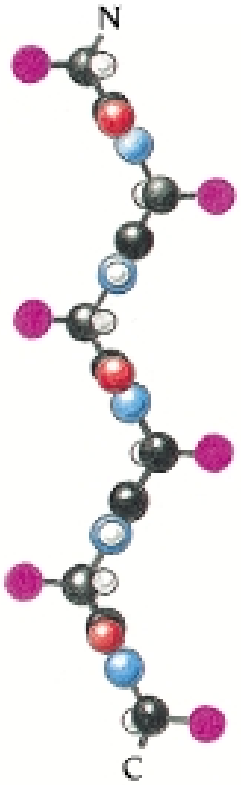
In both parallel and anti-parallel,
sidechains alternate above and below the plane of the sheet.

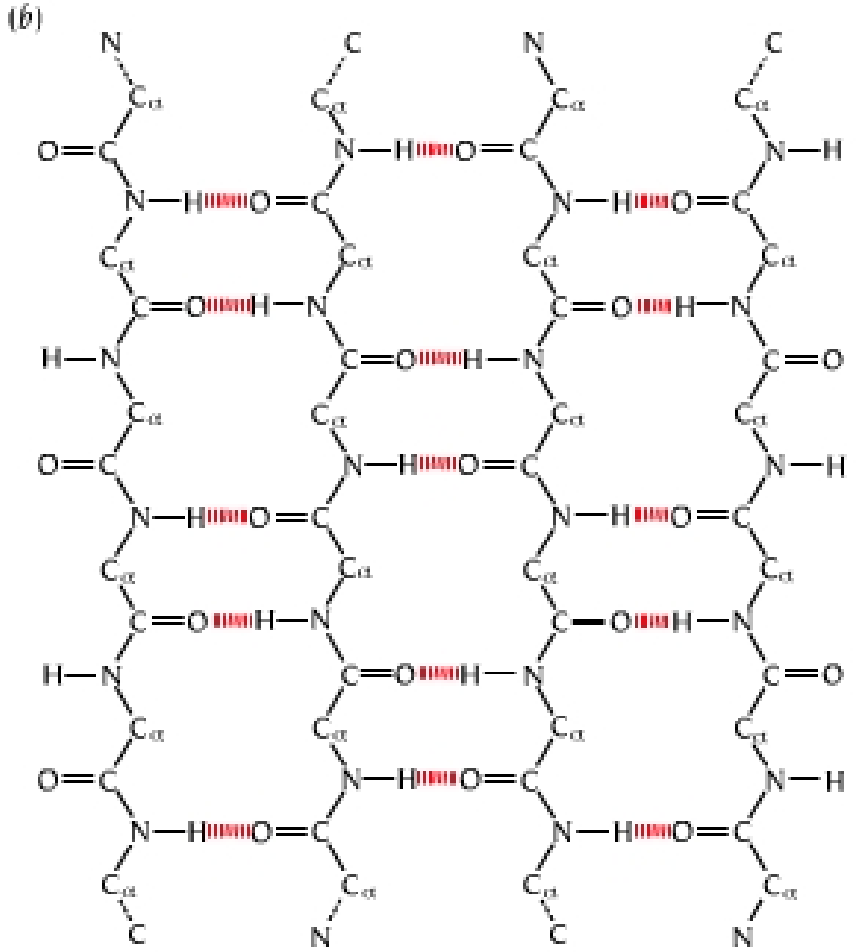
Parallel beta sheet



H-bonds are evenly spaced.
H-bonds are not 90° to the chain.

Anti-parallel beta sheet

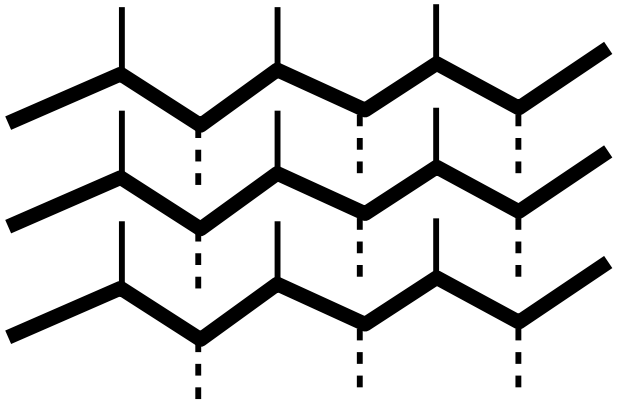


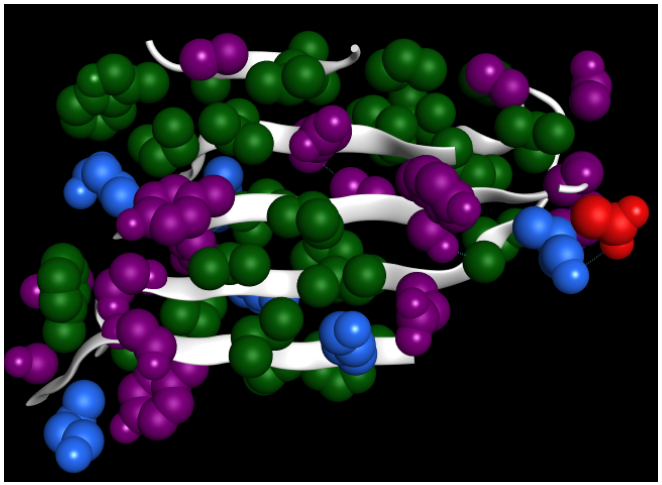


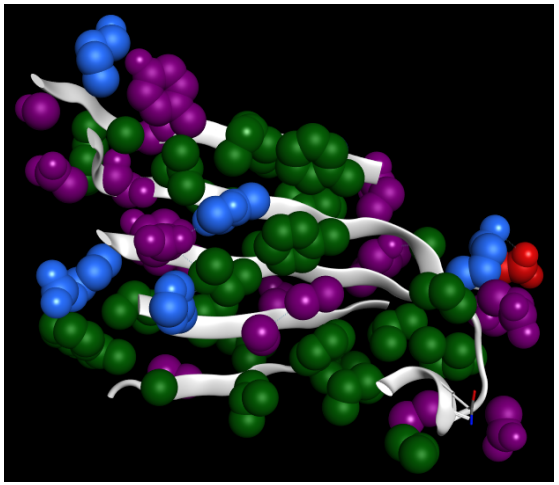
H-bonds are unevenly spaced.
H-bonds are 90° to the chain.

Sequence patterns for beta sheet

- npnp, where n=non-polar, p=polar
- nnnn







Charged residues mostly on the ends.

Secondary structure using matrices

An H-bonding pattern can be expressed using "augmented" matrix notation.

next H-bond donor
next H-bond acceptor

=

multiply by donor	multiply by acceptor	add to donor
multiply by donor	multiply by acceptor	add to acceptor

X

current H-bond donor
current H-bond acceptor

For example, for an alpha helix....

150
146

=

1	0	1
1	0	-3

X

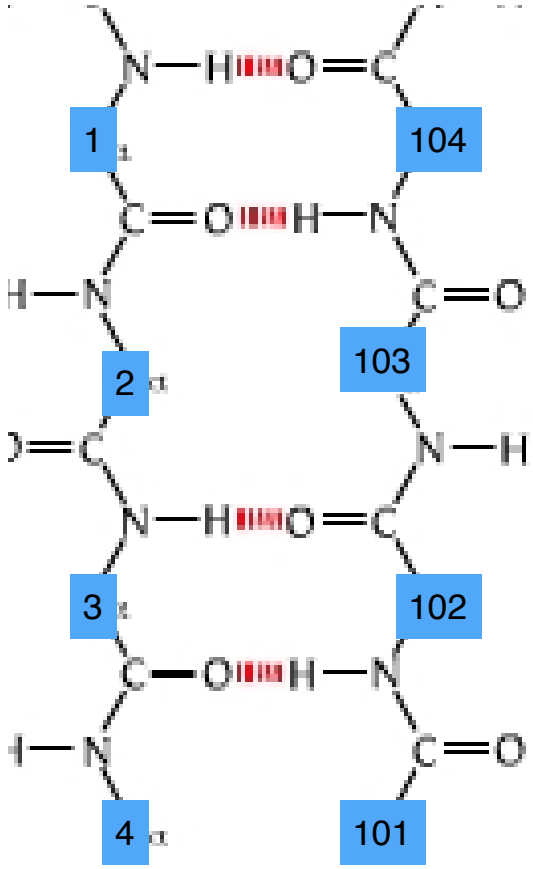
149
145

In a helix, donor NH is always +4 to acceptor O.

Secondary structure using matrices: antiparallel sheet

0	1	0
1	0	0

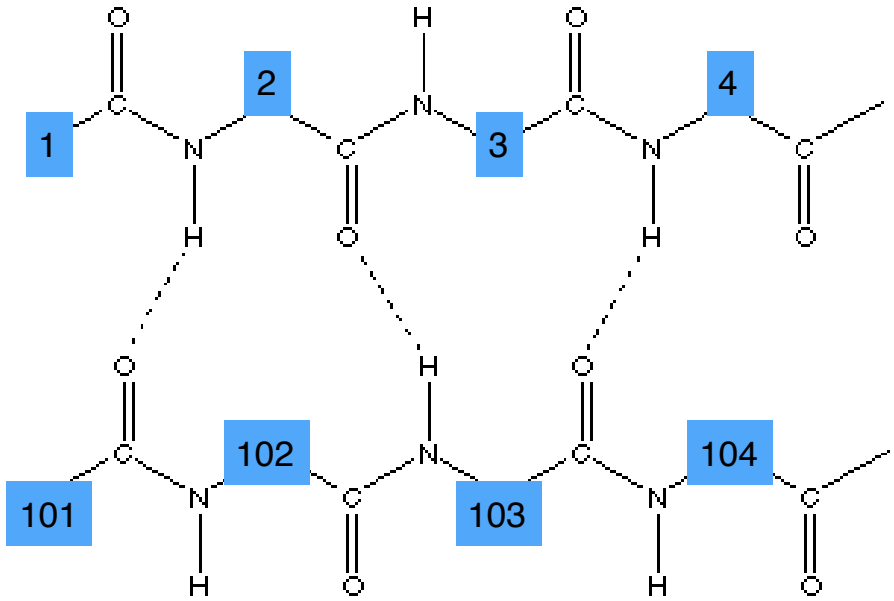
0	1	-2
1	0	+2



In a helix, donor NH is always +4 to acceptor O.

Secondary structure using matrices: parallel sheet

0	1	0
1	0	0



0	1	-2
1	0	+2

In a helix, donor NH is always +4 to acceptor O.

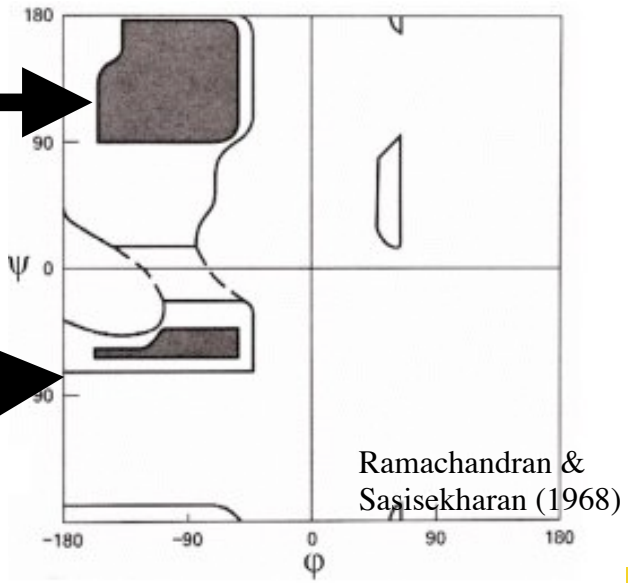
Secondary structures in Ramachandran Plot

beta sheet

-180° < Φ < 0°
90° < Ψ < 180°

alpha helix

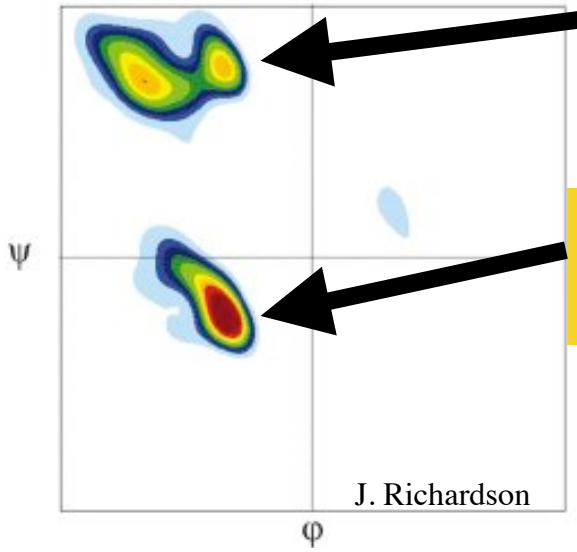
-100° < Φ < -40°
-80° < Ψ < -30°



Ramachandran & Sasisekharan (1968)

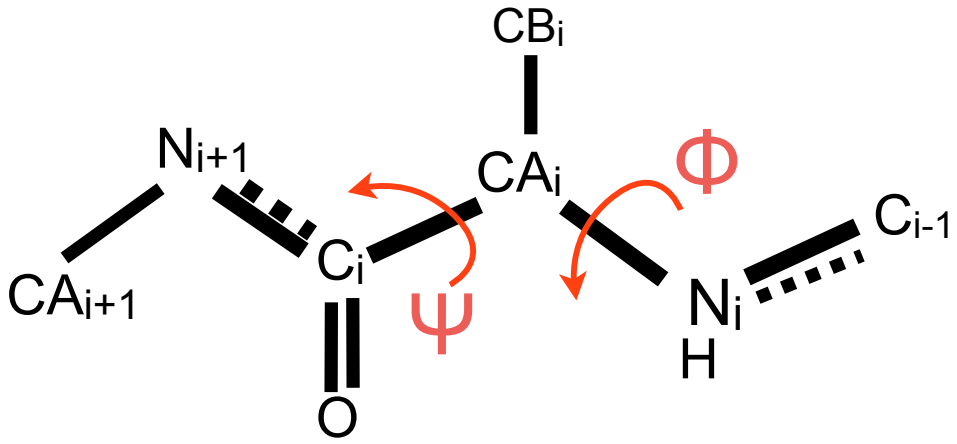
(a)

Φ = -139°
Ψ = 135° β



J. Richardson

Φ = -58°
Ψ = -47° α

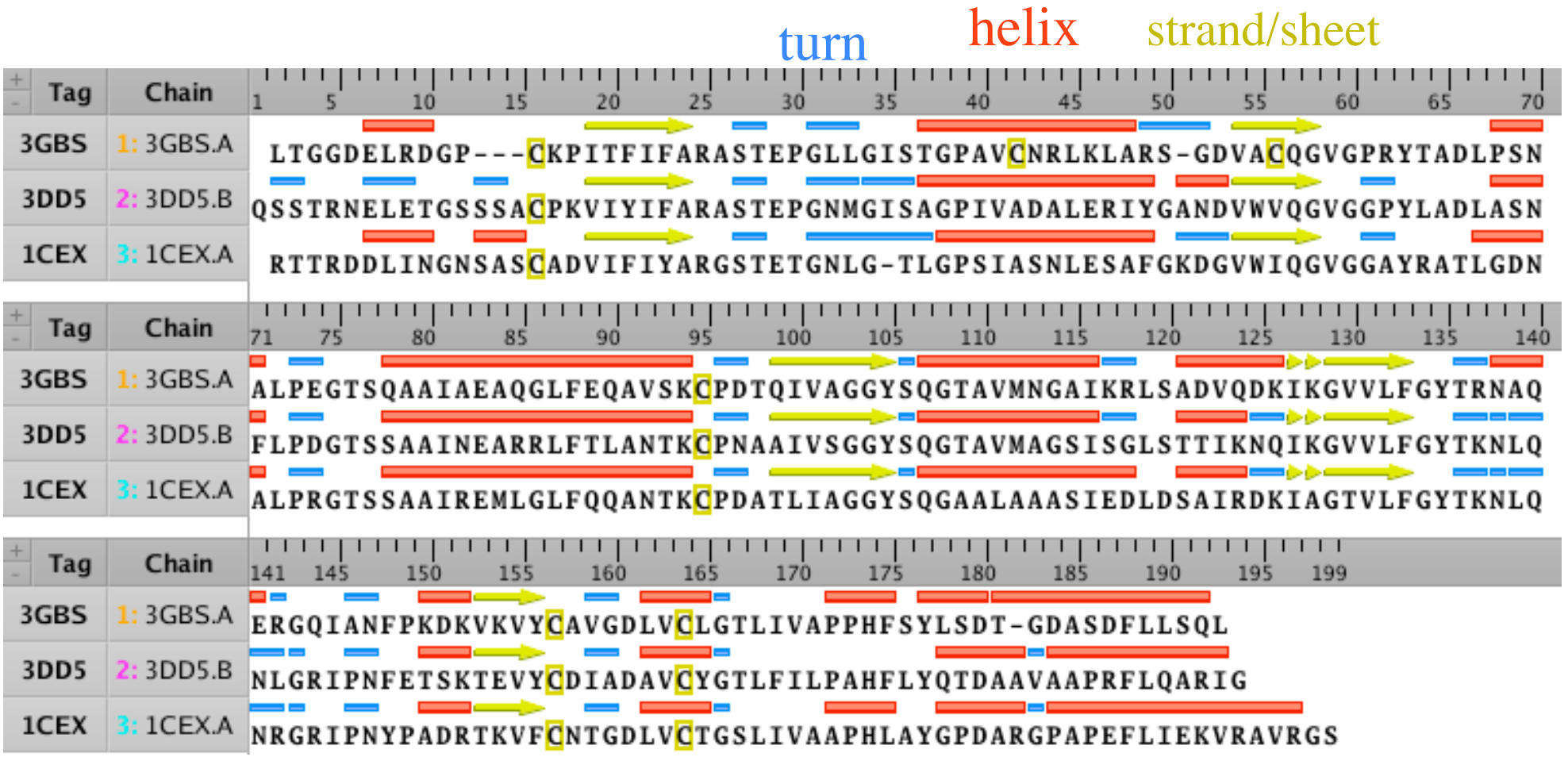


Predicting secondary structure from primary structure

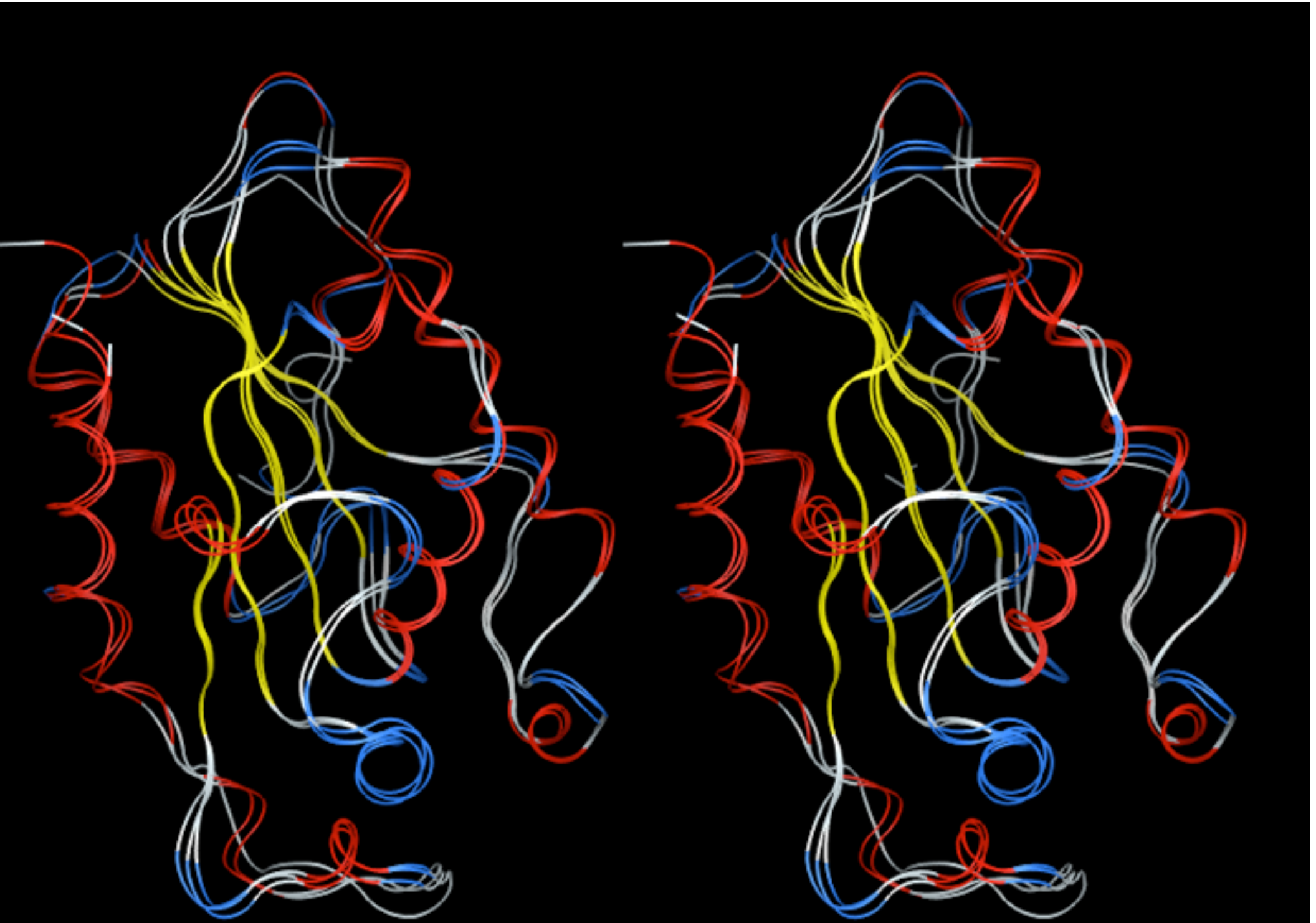
assumes

1. Secondary structures have sequence patterns
2. Those patterns are conserved across homolog proteins.

Secondary structure is strongly conserved among even remote homologs.



cutinases, 48 - 53% sequence identity.



Global positioning of SSEs is conserved.

The rule: similar sequence means similar structure

sequences that…

…have a common ancestor

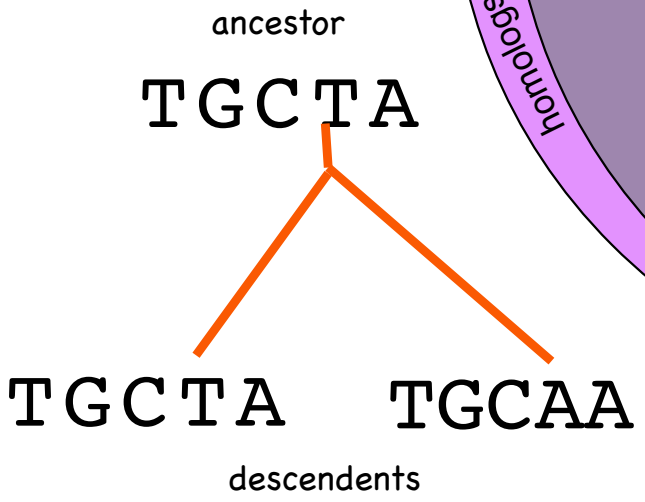
…superimpose in space

..Venn diagram..

homologs that don't *superpose*

most homologs are superposable

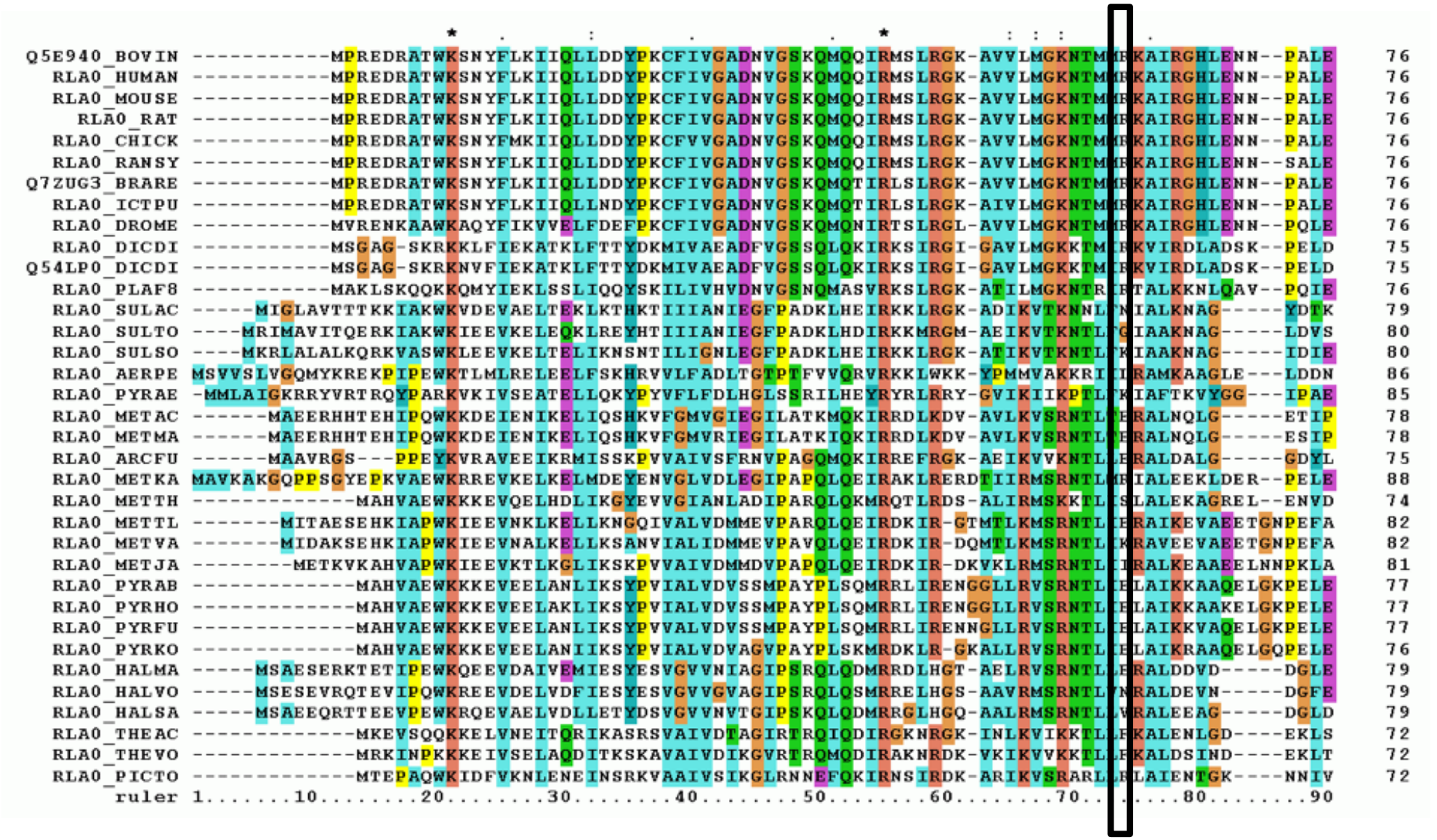
results of convergent evolution



Amino acid sequence profiles have patterns in them

- Positions in homologs conserve location, side chain conformation, packing environment.
- Evolution has sampled the low energy ways to fill each position.
- Multiple sequence alignments inform us about the nature of the position.
  - buried vs exposed.
  - alpha vs beta vs loop

First make a multiple sequence alignment

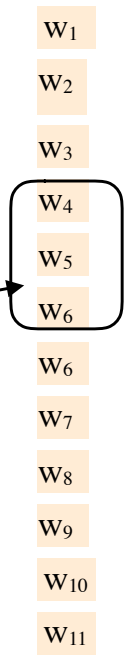


Each position in a MSA is a column of AA's representing the evolutionary history of one position.

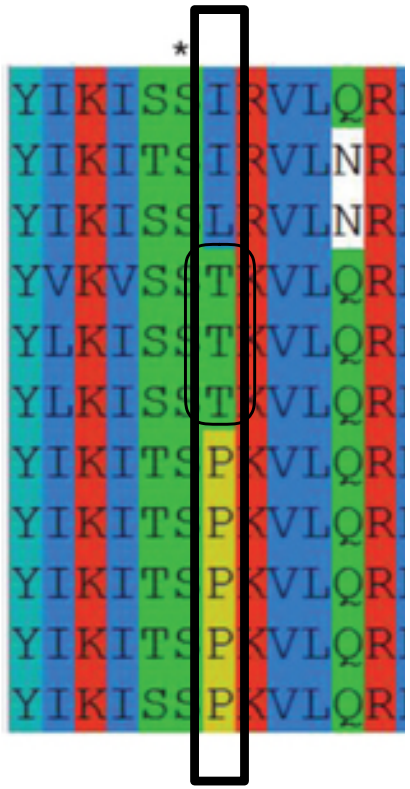
MSA is condensed to a matrix of AA probabilities.

$$P(T|7) = \frac{\sum_{i \,\forall\, S_{(7)}=T} w_i}{\sum_{all\ i} w_i}$$

Prob of Thr@ position 7 is the sum of the weights.



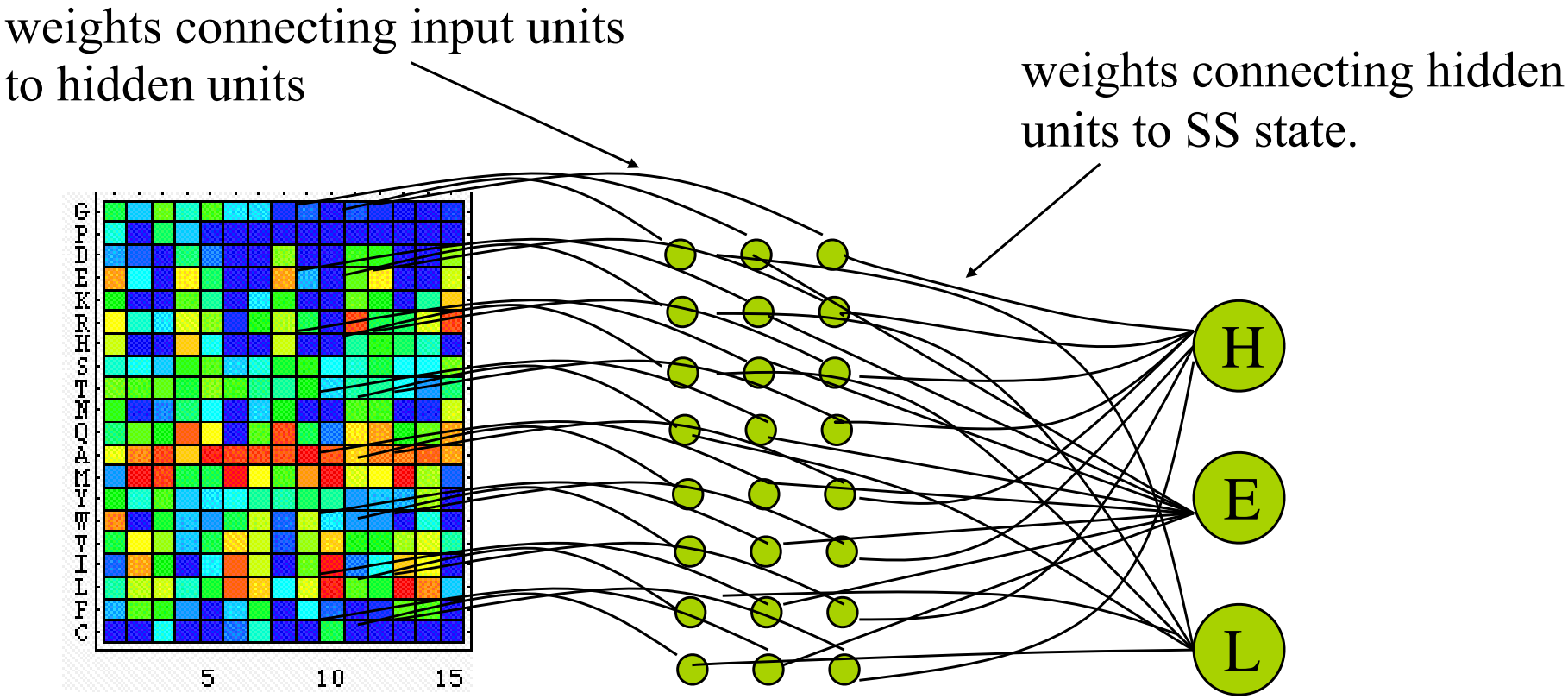
Sequences in the MSA are "weighted".



[The probability of amino acid *T* at position *7* is the sum of the sequence weights w_i over all sequences *i* such that the amino acid at position *7* of that sequence is *T*, divided by the sum over the sequence weights w_i.]

Psi-Pred: training a neural network to find patterns

First pass: NN encodes AA-dependence of SS.

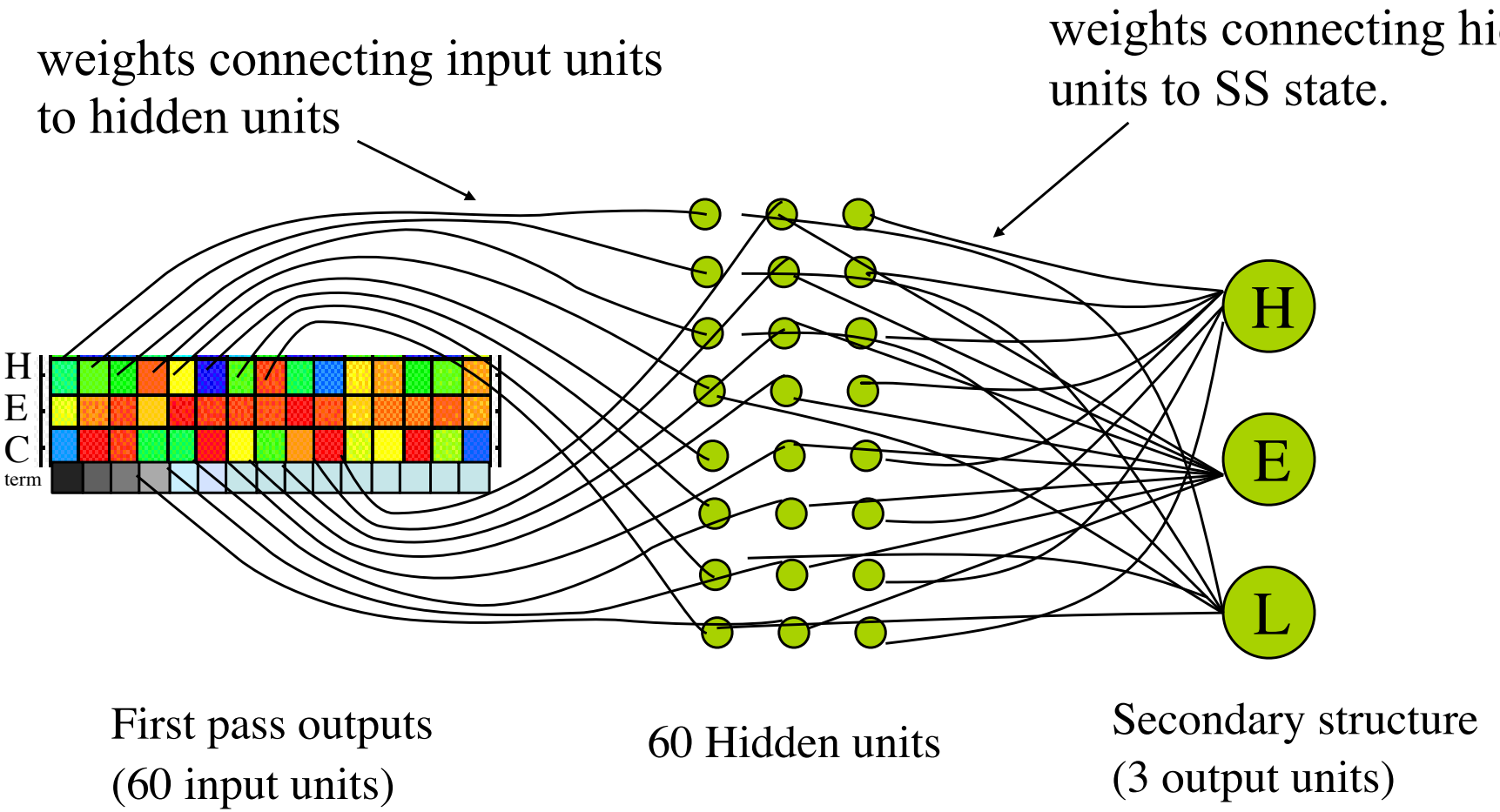


Training by back-propagation: weights are found that *minimize errors*

Burkhard Rost (1996) Methods in Enzymology

Psi-Pred: training a neural network to find patterns

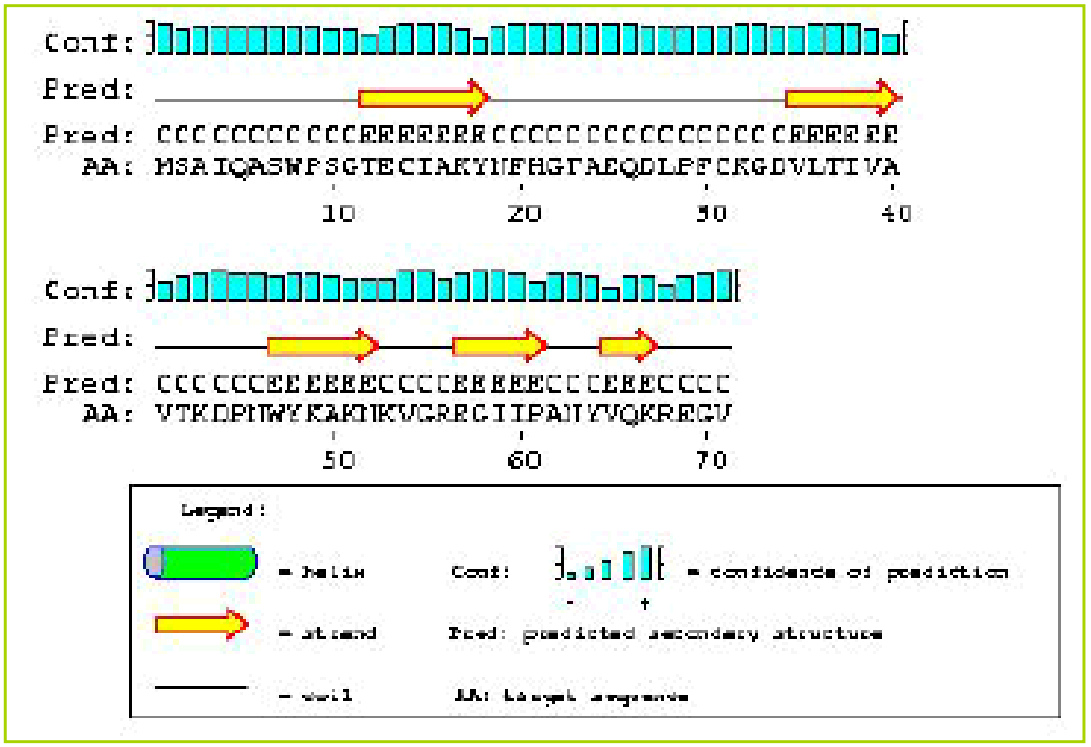
Second pass: capturing run-lengths



http://bioinf.cs.ucl.ac.uk/psipred/

PSI-pred-- a secondary structure predictor that uses profiles

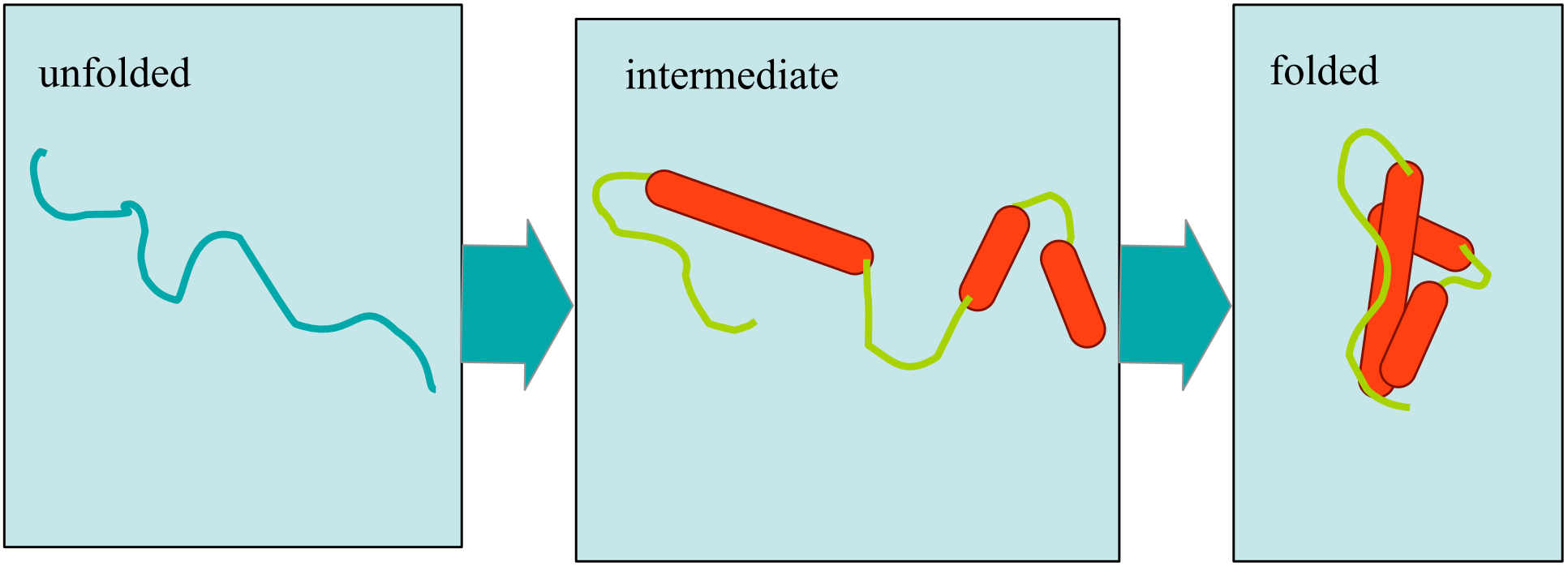
• PSI-PRED (Jones et al.) is currently the best server for secondary structure prediction, according to CASP results. • H, E or C is predicted based on an artificial neural network connecting a profile (**Psi-Blast** output) with known protein structures (**DSSP** assignments). • Predictions are assigned *confidences*. A window of 15 is used to predict the central residue.• Accuracy claimed to be 76-78% Q3.



The PSIPRED protein structure prediction server
L. J. McGuffin, K. Bryson, D. T. Jones (2000)
Bioinformatics 16 (4) p. 404-405

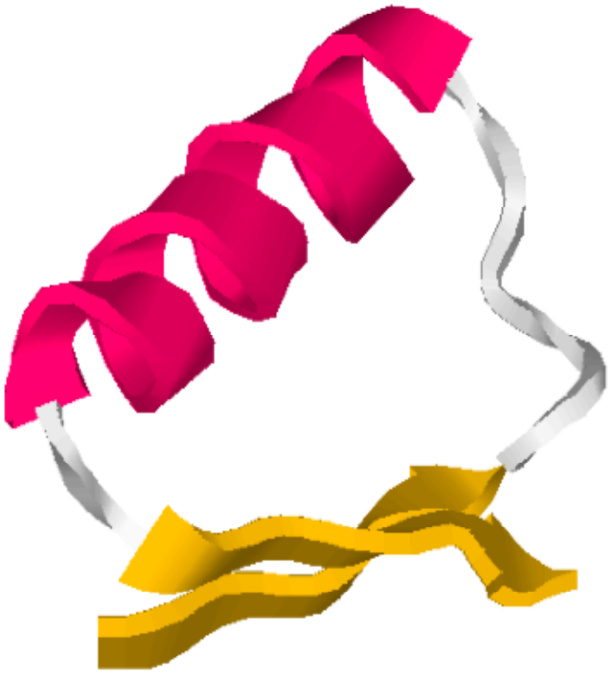
Q: Why does sliding window SS prediction work?

A: Local sequence has enough information to determine the secondary structure. This suggests that -- in general -- secondary structure forms early in the protein folding process, since it depends little on non-local (tertiary structure) interactions.



Super-secondary structures.

β hairpin



βαβ unit

Beta hairpin exercise

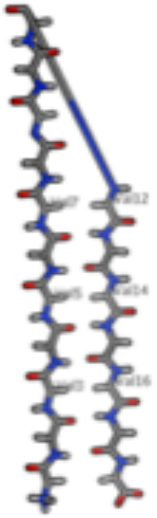
Exercise 4.1. Make a beta hairpin

Before you start, learn these mouse functions:

middlemouse	rotates all
shift-middlemouse	translates all
meta-middlemouse	rotates selected
shift-meta-middlemouse	translates selected
leftmouse	selects
shift-leftmouse	adds to selection
control-leftmouse	selects residues
control-shift-leftmouse	adds residues to selection
leftmouse double-click	selects whole chain
leftmouse click on empty space	clears selection

Create a peptide:

Edit | Build | Protein, Geometry: **anti-strand**. Residue: **ADVDVKVSPNGVEVKVRA**
(Can you convert from 1-letter to 3-letter codes?)
Center. Zoom out.



Make sure you have set **Select | synchronize** to the checked state.

In the SEQ window, select C-terminal 7 residues, **VEVKVRA**. Move the selected atoms so that the chain is antiparallel to the first 7 residues, **ADVDVKV**, with the valine sidechains lined up on one side.

Select each valine and label it (**ctrl-L** "V" return. **rightmouse drop-down menu | atoms | residue**)

Hide | side chains

Edit | Potential| Restrain.
Set Target 1.8, 1.8, Weight 50. Select H and O atoms from paired Valines.

Compute | Prepare | Structure preparation. Correct if necessary.

Start live energy minimization: **SVL: run 'gizmin.svl'.**

Make sure the hydrogen bonds are of the anti-parallel beta-sheet type. All Valines should be on the same side of the sheet.

If there are errors in the restraints, first **Cancel | GizMOE_minimizer,** then open **Potential Setup** (extreme lower left of the MOE window) | **Restraints.** Click on the restraints you want to delete or modify.

Restart gizmin: **SVL: run 'gizmin.svl'.**

Turn on **contacts | H-bond.** (Select BB. Unselect all others)

• Push and pull atoms untill you get residues **SPNG** to form a Type II beta turn (middle oxygen pointed UP when turn is viewed left-to-right , clockwise)

• Check for accidental D-amino acids. Switch them to L by pulling the alpha hydrogen straight through the alpha carbon.

• Check for any *cis*-peptides, make them *trans*.

• Make Gly 11 have a positive φ angle.

• Make salt bridges between oppositely charged side chains.

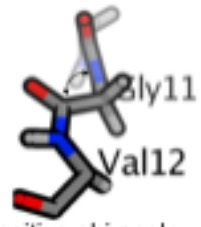
• Is the hairpin twisted? Right-handed or left-handed?

Cancel | GizMOE_minimize .

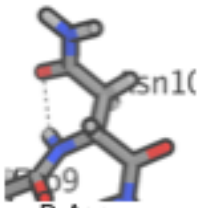
Remove the restraints. Restart **SVL: run 'gizmin.svl'.**

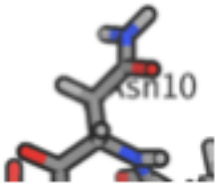
Does the structure hold together or fall apart with restraints?

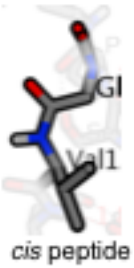
Save the MOE file. Upload to the homework server.
Check exercise4.1

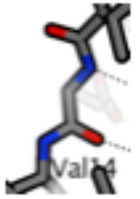


positive phi angle


D-Asn


L-Asn

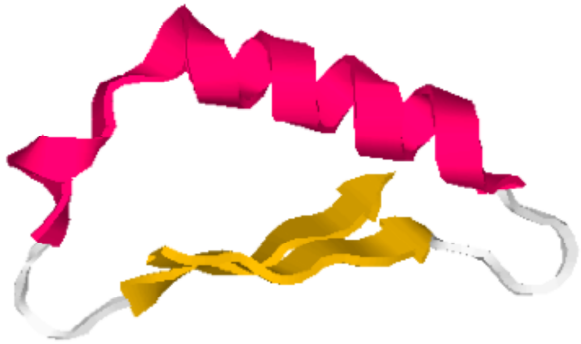

cis peptide

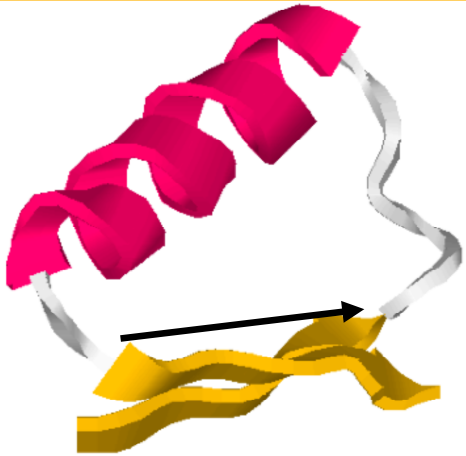

trans peptide

Final model with ribbon rendering. Beta sheet has a right-handed twist.

βαβ

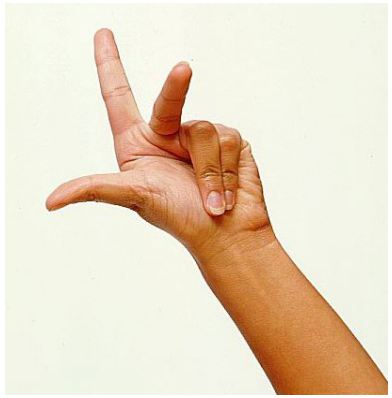
βαβ supersecondary structure units are mostly right-handed





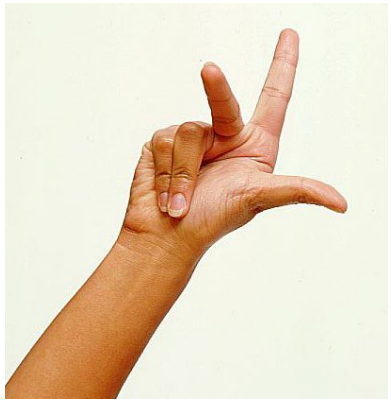
L-handed βαβ

1.5%



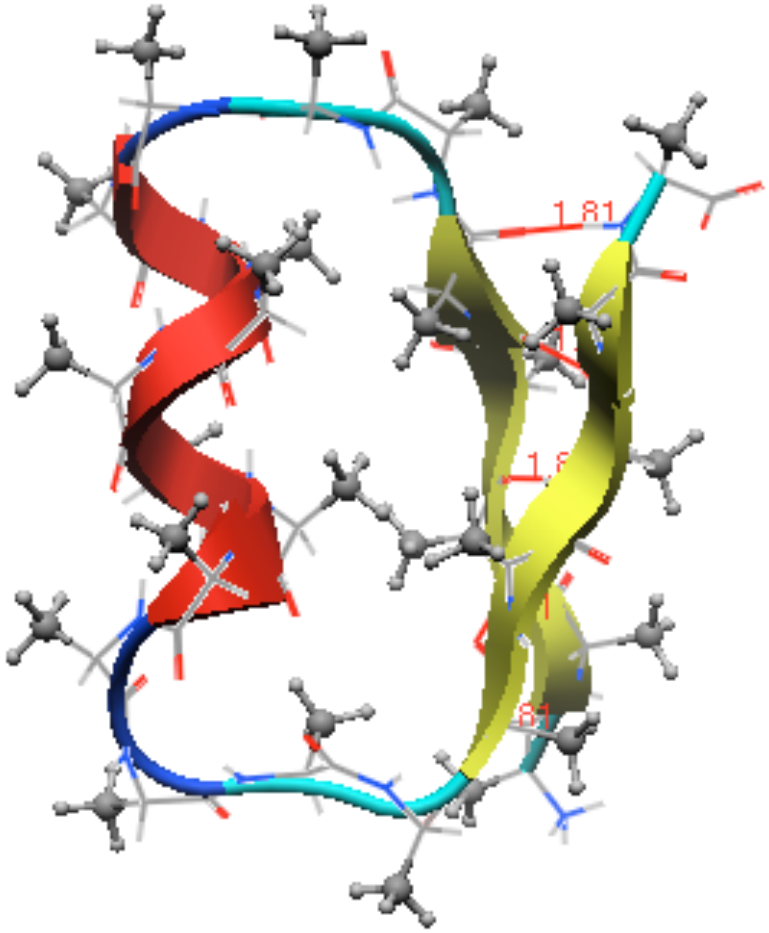
R-handed βαβ

98.5%



Homework 2

Make a right-handed βαβ unit by hand.



due Tue Feb 6

http://www.bioinfo.rpi.edu/bystrc/courses/biol4550/hw2.pdf