Overview of lecture
* General comment on peptide bond
* Discussion of backbone dihedral angles
* Discussion of Ramachandran plots
* Description of helix types.
* Description of $\beta$ structures
* NMR patterns associated with $2^\circ$ structures
* Chemical Shift Index
* TALOS
Useful background reading


Web Tutorial: http://broccoli.mfn.ki.se/pps_course_96/ss-960531_1.html
Details of the peptide bond

* Usually in the trans configuration.
* It has a partial (40%) double bond character which results in the six atoms being in one plane
* N is partial positive and the O is partial negative
Six atoms of a peptide bond lie in a plane
Consequences of the amide plane

* Angle about $C_\alpha$–N bond is denoted $\phi$.
* Angle about the $C_\alpha$-C bond is denoted $\varphi$.
* The entire path of the peptide backbone is known if $\phi$ and $\varphi$ is specified.
* Some values of $\phi$ and $\varphi$ are more energetically favorable.
Pictorial representation of peptide plane with $\phi$ and $\varphi$ represented.
Ramachandran Plots

* Not all combinations of $\phi$ and $\varphi$ angles are energetically favorable.

* Ramachandran was the first to demonstrate the acceptable values by plotting $\phi, \varphi$ combinations from known structures.

* The sterically favored combinations are the basis for preferred secondary structures.
Energetically preferred combinations.
Combinations found in proteins.
General properties of helices

* 32-38% or residues in helices.
* Side chains protrude from helix every 100°.
* Axis of the helix usually has a bend that may be as much as 20°.
* They ideally have $\phi = -57.8^\circ$ and $\varphi = -47.0^\circ$ which is in a favorable and steep minimum.
* There is a tendency to change from apolar to polar residues every 3-4 residues.
Types of Helices

3.10- $\phi = -74.0^\circ$ and $\varphi = -4.0^\circ$

pi- $\phi = -57.1^\circ$ and $\varphi = -69.7^\circ$
$\alpha$-Helix

* Originally described in 1951 to explain x-ray seen with $\alpha$-keratin.


* Pauling realized that the $\alpha$-helix had 3.6 residues per turn.

* Concluded there was a hydrogen bond between $>C=0$ of residue $n$ and $>N-H$ or residue $n+4$.

\[
>\text{C}=0 \quad --------\text{H-N}<$

\[
\text{n} \quad \text{n+4}
\]
* Predicted that L-amino acids give a right-handed helix since a left-handed one would have too close an approach between $>\text{C}=\text{O}$ and $\text{C}_\beta$.

* The final evidence came from the X-ray structure of myoglobin.

The $3_{10}$ helix

* The hydrogen bond is between the $>\text{C}=\text{O}$ of residue $n$ and the $>\text{N}-\text{H}$ or residue $n+3$.

* This results in backbone dihedral angles of $\phi = -74.0^\circ$ and $\varphi = -4.0^\circ$ which is also in an energy minimum.

* This is very common for short helical stretches.

* Often seen in the last turn of a helix.
* Since H-bond is separated by 4 amino acid this means first 3 >N-H groups and the four >C=O groups do not participate in H-bond.

* Thus over 2/3rds of amino acids in helices are not fully H-bonded.
Helix Capping

* Due to limited solvent access at first turn it necessitates the need for alternate H-bonds.

* >N-H at n-terminal are satisfied by side-chain H-bond acceptors.

* >C=O at the c-terminal are satisfied predominantly by backbone >N-H groups from the turn following the helix.

* There are currently 7 common capping motifs.

β-Sheets

* 20-28% of residues found in sheets.
* Pauling and Corey described the correct H-bonding pattern in 1951.
* The backbone in fully extended
* They ideally have $\phi = -120^\circ$ and $\varphi = +120.0^\circ$ which is in a favorable and steep minimum.
* There extended structure results in no intra-segment H-bonds or van der Waals interactions.

Anti-parallel $\beta$-sheet
Parallel β-sheet
Parallel vs. anti-parallel $\beta$-sheets

* Parallel sheets tend to be more regular if one examines the $\phi$ and $\varphi$ angles.

* Anti-parallel sheets have H-bonds perpendicular to the strands and alternate narrow and widely spaced pairs.

* Parallel sheets have equally spaced H-bonds that angle across the strands.
Hydrogen bonds is β-sheet
NOE patterns is β-sheet
Overview of tight(β) turns

* They have a succession of different $\phi$ and $\varphi$ values for each residue.
* They have a H-bond between the $>\text{C}=0$ of residue i and the $>\text{N}-\text{H}$ of residue i+3.
* Backbone at either end of Type I or II turn is in the right position to continue an anti-parallel $\beta$-ribbon.

Type I: $\phi_2 = -60^\circ$, $\varphi_2 = -30^\circ$, $\phi_3 = -90^\circ$, $\varphi_3 = 0^\circ$
Type II: $\phi_2 = -60^\circ$, $\varphi_2 = 120^\circ$, $\phi_3 = 90^\circ$, $\varphi_3 = 0^\circ$
Type I turn
Type II turn
Correlation between $^3J_{HN\alpha}$ and 2° Structure

NOE patterns associated with 2° structure
3D- HNHA experiment for $^3J_{hn\alpha}$
NOE pattern in an $\alpha$-helix

Strips taken from a 3D $^{15}$N-HSQC-NOESY Experiment
Chemical shifts and 2°structure

* One can often quickly determine the $^1\text{H}$, $^{15}\text{N}$ and $^{13}\text{C}$ chemical shifts with triple resonance experiments.

* It soon become obvious that there was a correlation between chemical shift deviations and secondary structure elements.

* These deviations are from random coil values

* Most important ones are CO, HA and CA.

Pattern between $2^\circ$ Structure and $^{13}$C Shifts

Deviations from random coil values

Chemical shift index (CSI)

* Fast, simple and reliable method to assign secondary structure.
* It is based on a statistical analysis of chemical shifts in proteins of known structure.
* CSI=0: $\delta$ in range of random coil values
* CSI=1 or -1: $\delta$ is greater or less than random coil chemical shift values.
* Must consider at least four residues to define an element of secondary structure.

## CSI RANDOM COIL VALUES

<table>
<thead>
<tr>
<th>RES</th>
<th>HA</th>
<th>C</th>
<th>CA</th>
<th>CB</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>4.32</td>
<td>177.8</td>
<td>52.5</td>
<td>19.1</td>
<td>123.8</td>
</tr>
<tr>
<td>C</td>
<td>4.55</td>
<td>174.6</td>
<td>58.2</td>
<td>28.0</td>
<td>118.8</td>
</tr>
<tr>
<td>C</td>
<td>4.71</td>
<td>174.6</td>
<td>55.4</td>
<td>41.1</td>
<td>118.6</td>
</tr>
<tr>
<td>D</td>
<td>4.64</td>
<td>176.3</td>
<td>54.2</td>
<td>41.1</td>
<td>120.4</td>
</tr>
<tr>
<td>E</td>
<td>4.35</td>
<td>176.6</td>
<td>56.6</td>
<td>29.9</td>
<td>120.2</td>
</tr>
<tr>
<td>F</td>
<td>4.62</td>
<td>175.8</td>
<td>57.7</td>
<td>39.6</td>
<td>120.3</td>
</tr>
<tr>
<td>G</td>
<td>3.96</td>
<td>174.9</td>
<td>45.1</td>
<td>9999.9</td>
<td>108.8</td>
</tr>
<tr>
<td>H</td>
<td>4.73</td>
<td>174.1</td>
<td>55.0</td>
<td>29.0</td>
<td>118.2</td>
</tr>
<tr>
<td>I</td>
<td>4.17</td>
<td>176.4</td>
<td>61.1</td>
<td>38.8</td>
<td>119.9</td>
</tr>
<tr>
<td>K</td>
<td>4.32</td>
<td>176.6</td>
<td>56.3</td>
<td>33.1</td>
<td>120.4</td>
</tr>
<tr>
<td>L</td>
<td>4.34</td>
<td>177.6</td>
<td>55.1</td>
<td>42.4</td>
<td>121.8</td>
</tr>
<tr>
<td>M</td>
<td>4.48</td>
<td>176.3</td>
<td>55.4</td>
<td>32.9</td>
<td>119.6</td>
</tr>
<tr>
<td>N</td>
<td>4.74</td>
<td>175.2</td>
<td>53.1</td>
<td>38.9</td>
<td>118.7</td>
</tr>
<tr>
<td>P</td>
<td>4.42</td>
<td>177.3</td>
<td>63.3</td>
<td>32.1</td>
<td>135.8</td>
</tr>
<tr>
<td>Q</td>
<td>4.34</td>
<td>176.0</td>
<td>55.7</td>
<td>29.4</td>
<td>119.8</td>
</tr>
<tr>
<td>R</td>
<td>4.34</td>
<td>176.3</td>
<td>56.0</td>
<td>30.9</td>
<td>120.5</td>
</tr>
<tr>
<td>S</td>
<td>4.47</td>
<td>174.6</td>
<td>58.3</td>
<td>63.8</td>
<td>115.7</td>
</tr>
<tr>
<td>T</td>
<td>4.35</td>
<td>174.7</td>
<td>61.8</td>
<td>69.8</td>
<td>113.6</td>
</tr>
<tr>
<td>V</td>
<td>4.12</td>
<td>176.3</td>
<td>62.2</td>
<td>32.9</td>
<td>119.2</td>
</tr>
<tr>
<td>W</td>
<td>4.66</td>
<td>176.1</td>
<td>57.5</td>
<td>29.6</td>
<td>121.3</td>
</tr>
<tr>
<td>Y</td>
<td>4.55</td>
<td>175.9</td>
<td>57.9</td>
<td>37.8</td>
<td>120.3</td>
</tr>
<tr>
<td></td>
<td>α-helix</td>
<td>β-sheet</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>---------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSI (CO)</td>
<td>+1</td>
<td>-1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSI (Cα)</td>
<td>+1</td>
<td>-1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSI (Hα)</td>
<td>-1</td>
<td>+1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSI (Cβ)</td>
<td>n.a.</td>
<td>+1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example of CSI analysis
TALOS

* Database system for empirical prediction of phi and psi backbone angles.
* Uses a combinations of five kinds (HA, CA, CB, CO, N) of chemical shifts.
* Like CSI relies on fact that many secondary chemical shifts are highly correlated with aspects to protein secondary structure.
* It uses triplets of amino acids to make predictions.

## Talos Random Coil Table

<table>
<thead>
<tr>
<th>RES</th>
<th>HA</th>
<th>C</th>
<th>CA</th>
<th>CB</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>4.32</td>
<td>177.8</td>
<td>52.3</td>
<td>19.0</td>
<td>123.8</td>
</tr>
<tr>
<td>C</td>
<td>4.55</td>
<td>174.6</td>
<td>56.9</td>
<td>28.9</td>
<td>118.8</td>
</tr>
<tr>
<td>C</td>
<td>4.71</td>
<td>174.6</td>
<td>55.4</td>
<td>43.7</td>
<td>118.6</td>
</tr>
<tr>
<td>D</td>
<td>4.64</td>
<td>176.3</td>
<td>54.0</td>
<td>40.8</td>
<td>120.4</td>
</tr>
<tr>
<td>E</td>
<td>4.35</td>
<td>176.6</td>
<td>56.4</td>
<td>29.7</td>
<td>120.2</td>
</tr>
<tr>
<td>F</td>
<td>4.62</td>
<td>175.8</td>
<td>58.0</td>
<td>39.0</td>
<td>120.3</td>
</tr>
<tr>
<td>G</td>
<td>3.96</td>
<td>174.9</td>
<td>45.1</td>
<td>9999.9999</td>
<td>108.8</td>
</tr>
<tr>
<td>H</td>
<td>4.73</td>
<td>173.3</td>
<td>54.5</td>
<td>27.9</td>
<td>118.2</td>
</tr>
<tr>
<td>I</td>
<td>4.17</td>
<td>176.4</td>
<td>61.3</td>
<td>38.0</td>
<td>119.9</td>
</tr>
<tr>
<td>K</td>
<td>4.32</td>
<td>176.6</td>
<td>56.5</td>
<td>32.5</td>
<td>120.4</td>
</tr>
<tr>
<td>L</td>
<td>4.34</td>
<td>177.6</td>
<td>55.1</td>
<td>42.3</td>
<td>121.8</td>
</tr>
<tr>
<td>M</td>
<td>4.48</td>
<td>176.3</td>
<td>55.3</td>
<td>32.6</td>
<td>119.6</td>
</tr>
<tr>
<td>N</td>
<td>4.74</td>
<td>175.2</td>
<td>52.8</td>
<td>37.9</td>
<td>118.7</td>
</tr>
<tr>
<td>P</td>
<td>4.42</td>
<td>177.3</td>
<td>63.1</td>
<td>31.7</td>
<td>135.8</td>
</tr>
<tr>
<td>Q</td>
<td>4.34</td>
<td>176.0</td>
<td>56.1</td>
<td>28.4</td>
<td>119.8</td>
</tr>
<tr>
<td>R</td>
<td>4.34</td>
<td>176.3</td>
<td>56.1</td>
<td>30.3</td>
<td>120.5</td>
</tr>
<tr>
<td>S</td>
<td>4.47</td>
<td>174.6</td>
<td>58.2</td>
<td>63.2</td>
<td>115.7</td>
</tr>
<tr>
<td>T</td>
<td>4.35</td>
<td>174.7</td>
<td>62.1</td>
<td>69.2</td>
<td>113.6</td>
</tr>
<tr>
<td>V</td>
<td>4.12</td>
<td>176.3</td>
<td>62.3</td>
<td>32.1</td>
<td>119.2</td>
</tr>
<tr>
<td>W</td>
<td>4.66</td>
<td>176.1</td>
<td>57.7</td>
<td>30.3</td>
<td>121.3</td>
</tr>
<tr>
<td>Y</td>
<td>4.55</td>
<td>175.9</td>
<td>58.1</td>
<td>38.8</td>
<td>120.3</td>
</tr>
</tbody>
</table>
The TALOS triplet system

TALOS searches database for the ten best matches to a given triplet in the target protein (Data base uses 20 proteins representing 3,000 triplets).
Reliability of TALOS

* makes no prediction on 20-45% of the residues in proteins.
* makes predictions for about 65% of the residues on average.
* in 5 of 20 proteins studied the results included no bad predictions.
* about 3% of the predictions were bad!!!
* he uncertainty for good predictions was 14 degrees for phi and 13 degrees for psi.

Example of TALOS output