Protein folding
Local structure
Handedness
SCOP -- a hierarchy

- [http://scop.berkeley.edu](http://scop.berkeley.edu)

1. class
2. fold
3. superfamily
4. family
5. protein
6. species

SCOP’s hierarchy is sequence centered
Folding -- also a hierarchy

unfolded

intermediate

folded

$\text{t} = 0$

$\text{t} = 1 \text{ ns}$

$\text{t} = 1 \text{ ms}$

$t = \text{time after leaving the ribosome}$

More about protein folding in later lectures
Efimov showed that almost all proteins can be classified as being one of just 7 "structural trees". In each tree, structures are "grown" by adding secondary structure elements one at a time.
How to grow a Efimov structural tree.

Evolutionary significance?
Folding pathway?
Maybe it's both!

commom core
A classification / folding hierarchy

**Classification**

Early

- Secondary structure
- Local structure
- Super-secondary structure
- Tertiary structure
- Quaternary structure

Late
Folding step 1?

Secondary structure
About the Alpha helix

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

H-bond rule for donor to acceptor (NH->O): \(i\) to \(i+4\)

Helices are right-handed

Helices are bumpy, not “cylindrical”.
The peptide dipole makes Hydrogen bonding stronger.
Sequence pattern for the amphipathic alpha helix

- npnnpp, where n = non-polar, p = polar

- Example: `LSELFKNLQDMLSK`

The helix is held together by the hydrophobic effect.

Amphipathic helices stick to other amphipathic helices.
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, or nnnn, where n=non-polar, p=polar

Non-polar residues (green, purple) mostly on the face.

Charged residues (blue, red) mostly on the ends.
Mathematical expressions for H-bonding patterns

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

<table>
<thead>
<tr>
<th>next H-bond donor</th>
<th>multiply by donor</th>
<th>multiply by acceptor</th>
<th>add to donor</th>
</tr>
</thead>
<tbody>
<tr>
<td>next H-bond acceptor</td>
<td>multiply by donor</td>
<td>multiply by acceptor</td>
<td>add to acceptor</td>
</tr>
</tbody>
</table>

For example, for an alpha helix....

\[
\begin{align*}
150 & \quad = \quad 101 \\
146 & \quad = \quad 011
\end{align*}
\]

or

\[
\begin{align*}
150 & \quad = \quad 100 \\
146 & \quad = \quad 100
\end{align*}
\]
Secondary structure using matrices: antiparallel sheet

Use the augmented matrices to find the next H-bond before/after (donor, acceptor) = (102, 3) in a antiparallel sheet
Secondary structure using matrices: parallel sheet

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>C</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>N</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>-2</td>
<td></td>
</tr>
</tbody>
</table>

Use the augmented matrix to find the next H-bond before/after (donor, acceptor) = (103, 2) in a parallel sheet
Are beta sheets really secondary structure? Or tertiary structure?

Early folding is local. Late folding is non-local.

Alpha helix
• H-bonds are local
• Side chain contacts are local
• Folding is fast (ns)

Beta sheet
• H-bonds are non-local
• Side chain contacts are non-local
• Folding is slow (μs)

Is early folding just a decision of helix versus not-helix?

J. Richardson’s Ramachandran plot

Φ = -139° Ψ = 135°

Φ = -58° Ψ = -47°
Folding step 2?

Local structure
Turns and caps

- Short pieces of protein chain sample conformational space randomly, driven by energy.

- Sequence determines structure. Non-polar sidechains and glycines are especially important.

- Usually reverse the chain direction.
beta turns
4-residues

Residue 1 hydrogen bonds to residue 4

Type I (most common). Oxygen points away when viewed clockwise.

Type II (less common). Oxygen points toward when viewed clockwise.
Backbone angles and sequence of beta turns

Backbone angles ±30°

<table>
<thead>
<tr>
<th>Type</th>
<th>$\phi_{i+1}$</th>
<th>$\psi_{i+1}$</th>
<th>$\phi_{i+2}$</th>
<th>$\psi_{i+2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>-60</td>
<td>-30</td>
<td>-90</td>
<td>0</td>
</tr>
<tr>
<td>II</td>
<td>-60</td>
<td>120</td>
<td>80</td>
<td>0</td>
</tr>
<tr>
<td>VIII</td>
<td>-60</td>
<td>-30</td>
<td>-120</td>
<td>120</td>
</tr>
<tr>
<td>I'</td>
<td>60</td>
<td>30</td>
<td>90</td>
<td>0</td>
</tr>
<tr>
<td>II'</td>
<td>60</td>
<td>-120</td>
<td>-80</td>
<td>0</td>
</tr>
<tr>
<td>Vla1</td>
<td>-60</td>
<td>120</td>
<td>-90</td>
<td>0*</td>
</tr>
<tr>
<td>Vla2</td>
<td>-120</td>
<td>120</td>
<td>-60</td>
<td>0*</td>
</tr>
<tr>
<td>Vlb</td>
<td>-135</td>
<td>135</td>
<td>-75</td>
<td>160*</td>
</tr>
<tr>
<td>IV</td>
<td>turns excluded from all the above categories</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*have cis-peptide bond at $i+2$

Glycine rules turn propensity

<table>
<thead>
<tr>
<th>position type \ type</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td></td>
<td>P</td>
<td></td>
<td>G</td>
</tr>
<tr>
<td>II</td>
<td></td>
<td>P</td>
<td>P</td>
<td>G</td>
</tr>
<tr>
<td>VIII</td>
<td></td>
<td>G/P</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>I'</td>
<td></td>
<td>G</td>
<td>G</td>
<td></td>
</tr>
<tr>
<td>II'</td>
<td></td>
<td>G</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

http://www.ebi.ac.uk
## Sequence patterns for turns, caps and secondary structure

<table>
<thead>
<tr>
<th>Motif</th>
<th>Average boundaries</th>
<th></th>
<th>Average rmsd (len)</th>
<th>Pattern of conserved non-polar residues</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mda (°)</td>
<td>dme (Å)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1  Amphipathic α-helix</td>
<td>56</td>
<td>0.71</td>
<td>0.78 (15)</td>
<td>1-4-8, 1-5-8</td>
</tr>
<tr>
<td>2  Non-polar α-helix</td>
<td>54</td>
<td>0.58</td>
<td>0.40 (11)</td>
<td>1-4-8, 1-5-8</td>
</tr>
<tr>
<td>3  Schellman cap type 1</td>
<td>81</td>
<td>1.01</td>
<td>1.02 (15)</td>
<td>1-6-9-11</td>
</tr>
<tr>
<td>4  Schellman cap type 2</td>
<td>76</td>
<td>0.94</td>
<td>0.94 (15)</td>
<td>1-6-8-9</td>
</tr>
<tr>
<td>5  Proline α-helix C cap</td>
<td>92</td>
<td>1.07</td>
<td>0.89 (13)</td>
<td>1-2-5-8</td>
</tr>
<tr>
<td>6  Frayed α-helix</td>
<td>75</td>
<td>0.96</td>
<td>0.69 (15)</td>
<td>1-5-9-13</td>
</tr>
<tr>
<td>7  Helix N capping box</td>
<td>99</td>
<td>0.95</td>
<td>0.65 (15)</td>
<td>1-6-9-13</td>
</tr>
<tr>
<td>8  Amphipathic β-strand</td>
<td>89</td>
<td>0.87</td>
<td>0.87 (6)</td>
<td>1-3, 1-3-5</td>
</tr>
<tr>
<td>9  Hydrophobic β-strand</td>
<td>101</td>
<td>0.91</td>
<td>0.91 (7)</td>
<td>1-2-3</td>
</tr>
<tr>
<td>10 β-Bulge</td>
<td>100</td>
<td>0.97</td>
<td>0.78 (7)</td>
<td>1-4-6</td>
</tr>
<tr>
<td>11 Serine β-hairpin</td>
<td>94</td>
<td>0.76</td>
<td>0.81 (9)</td>
<td>1-8</td>
</tr>
<tr>
<td>12 Type-I hairpin</td>
<td>80</td>
<td>0.94</td>
<td>1.23 (13)</td>
<td>1-7-8</td>
</tr>
<tr>
<td>13 Diverging type-II turn</td>
<td>87</td>
<td>1.04</td>
<td>1.00 (9)</td>
<td>1-7-9</td>
</tr>
</tbody>
</table>

Folding step 3?

Super-secondary structure
**β Hairpins**

Serine hairpin

Type-I hairpin

Diverging type-2 turn

Sequence patterns are expressed as log-likelihood, converted to color scale.
β sheet super-secondary structure.

meander

"greek key"

β helix (2-sided, 3-sided, left, right)

wikipedia.org/wiki/Beta_sheet
α Helical Super-Secondary Structures

- SSS contains more than one SSE.
- Beta turns and helix caps are usually involved.
- Canonical ones have names.

Coiled-coil

Helix hairpin

alpha-alpha corner

EF hand

www.cryst.bbk.ac.uk
Handedness

Right-handed helix.
Put the thumb of the right hand along the axis of rotation.

As you travel up the helix (going in the direction of your right thumb) the line curve in the direction of your fingers.

Yes, that means you are turning left when you walk up a right-handed spiral staircase, and right when you are walking up a left-handed spiral staircase.
Super-secondary structure.

βαβ supersecondary structure units are mostly right-handed

L-handed βαβ
1.5%

R-handed βαβ
98.5%
Sternberg & Thornton: Twist of beta sheet makes right-handed crossover more of a straight line.
Theories for why $\beta\alpha\beta$ units are right-handed.
Theories for why $\beta\alpha\beta$ units are right-handed.

Phone Cord Effect: Northern versus Southern route to helix

Theories for why $\beta\alpha\beta$ units are right-handed.

left-handed torque turns left-handed $\beta\alpha\beta$ to right-handed $\beta\alpha\beta$

### Phone cord: Brownian Dynamics Simulations

#### Diagram
- **(a)**: B1, L1, H, L2, B2
- **(b)**: Right-handed
- **(c)**: Helical
- **(d)**: Ambiguous

#### Data Table
<table>
<thead>
<tr>
<th>Ψ</th>
<th>0% South</th>
<th>50% South</th>
<th>100% South</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trials</td>
<td>2738</td>
<td>1164</td>
<td>501</td>
</tr>
<tr>
<td>Collapsed</td>
<td>2540</td>
<td>1066</td>
<td>418</td>
</tr>
<tr>
<td>Helical</td>
<td>851</td>
<td>578</td>
<td>286</td>
</tr>
<tr>
<td>Ambiguous</td>
<td>456</td>
<td>299</td>
<td>131</td>
</tr>
<tr>
<td>Right-handed</td>
<td>124</td>
<td>130</td>
<td>107</td>
</tr>
<tr>
<td>Left-handed</td>
<td>271</td>
<td>149</td>
<td>48</td>
</tr>
</tbody>
</table>

[Phone cord video](http://www.youtube.com/watch?feature=player_embedded&v=6hQYjtmU6E0)
### 3-helix bundles are also right-handed

61% R, 39% L, n=431, p(R=0.50)<0.01
Folding step 4?

Tertiary structure
Folding pathway is an ensemble

Individual chains may form structure in different orders, but all end up in the native state (or misfolded, or aggregated).

[Image: https://commons.wikimedia.org/wiki/File:ACBP_MSM_from_Folding@home.tiff]
Folding is a funnel

...with lots of energetic barriers.

Protein folding is an ensemble on a noisy funnel

Dead ends are misfolded states.

Ensemble: Many paths to the folded state.


How to force hydrogen bonds using restraints

- **To add a restraint**
  - **Edit | Potential | Restrain**, distance, Target 1.8, 1.8, Weight 50
  - Pick amide H and carbonyl O.
  - Click **Create**.
  - Cancel | Restrain (or esc) when done

- **Energy minimize**
  - **Compute | prepare | Structure preparation**
    - Checks for missing atoms, assigns energies.
  - SVL: run ‘gizmin.svl’
    - When finished, be sure to Cancel | GizMOE_Minimizer

- **To remove or modify restraints**
  - **Potential setup** (button at far lower left)
  - Restraints tab
Exercise 4.1

Make a beta hairpin

anti-parallel sheet with valine side chains all on the same side of the sheet.

Edit | Build | Protein, Geometry: anti-strand. Residue: ADVDVKVSPNGVEVKVRA

Zoom out.

Select the second half of the chain starting with NG.

Rotate and translate it (shift-alt-middlemouse) so that the first three valines (3,5,7) are lined up with other there valines (12,14,16), and the valine backbone H-bonding groups (NH and CO) are close to the H-bonding distance (1.8Å from H to O)

Hide side chains to help see the backbone atoms better.

Edit | Potential | Restraine Set Target 1.8, 1.8, Weight 50. Select H and O atoms. Create. When done you have 2 restraints for each of the three paired valines for a total of 6 restraints.

Compute | Prepare | Structure preparation. Hit Correct if necessary. Protonate3D.

SVL: run 'gizmin.svl'.

If there are errors in the restraints, Cancel/GizMOE, open Potential Setup (extreme lower left of the MOE window). Restraints. Click on restraints to delete or modify them.

Restart SVL: run 'gizmin.svl'.

Look at out the structure.
It should have beta pleating when viewed from the edge of the sheet. Sidechains should alternate up and down in that view. Residues SPNG form a beta-turn.

Cancel/Gizmin. Remove the restraints. Restart SVL: run 'gizmin.svl'.

Does the structure hold together or fall apart?