Automated homology modeling.
Manual re-alignment
Deletion within a helix. What happens?

True indels within a contiguous alpha helix are rare. Alignment errors are pretty common.
Deletion within a strand. What happens?

Before deletion. Residue in green will be deleted.


Therefore...

Don't make a deletion with a secondary structure element.

Move it to a turn region.
Move beta strand indel to turn

- Moving a deletion out of a helix or strand avoids disruption of the SSE caused by the energy minimization (which is incapable of saving a bad alignment).
- Moving a deletion move the disruption to where it can be tolerated.

**alignment**

```
XXXXXXXXXXXXXXXX
YYY~YYYYYYYYYY
```

**before**

```
XXXXXXXXXXXXXXXX
YYY~YYYYYYYYYY
```

**after deletion**

```
XXXXXXXXXXXXXXXX
YYY~YYYYYYYYYY
```

**after energy min.**

```
XXXXXXXXXXXXXXXX
YYY~YYYYYYYYYY
```

**alignment**

```
XXXXXXXXXXXXXXXX
YYY~YYYYYYYYYY
```

**bad. distorsion. H-bonds lost.**

```
XXXXXXXXXXXXXXXX
YYY~YYYYYYYYYY
```

**good . no distorsion. H-bonding retained.**

```
XXXXXXXXXXXXXXXX
YYY~YYYYYYYYYY
```
Overrule automatic alignment when it...

- puts an indel in the **middle of a helix**
- puts an indel in the **middle of a strand**
- puts the anchor residues of a deletion **too far apart**
- puts an insertion at a buried position
- puts an insertion in the middle of a helix
- forces a database search of a loop of length zero.
- forces a database where the anchors are adjacent in the template.

How to intervene:

- Find a new spot for your indel
- Push the gap to the new spot before running (or re-running) Homology Model.
move the anchor points by "unaligning"

Alignment forces 2 residues into the space of ... zero!

Alignment forces 4 residues into the space of 2. Doable!
move insertion point to a turn

Alignment forces insertion into a helix!

Target:
- - - - - - - - - -

Template:
- - - - - - - - - -

Alignment forces 2 residues into the space of 1. Doable!

Target:
- - - - - - - - - -

Template:
- - - - - - - - - -

anchors
An exception to the *no-insertions-in-strand* rule

**Beta-bulge:**

Two sidechains (usually polar) point to one side of the sheet, instead of just one. This causes a kink in an otherwise continuous pairing.

A beta-bulge is almost always a 1-residue insertion, here, usually a polar amino acid, most frequently D.

Occasionally, a many residue insertion occurs here. In that case it is called a “*beta-blowout.*”
An exception to the *no-insertions-in-helix* rule

Nature has selected an insertion of strongly anti-helix residues.
Not an exception to the *no-deletions-in-helix* rule

Deletion in the sequence leads to shorter helix, not a helix with a "deletion" in it.

This shows two actual structures. The one with the deletion is more extended, to span the distance.
Demonstration of the power of re-alignment

1. Load "sequence 1" again (or go back to your MOE file from the exercise)
2. Re-run the PDB search if necessary. Load 2IGD.
3. Align (use the defaults)
4. Your alignment should look like this:
5. Save as "before.moe"

6. Protein > Homology Model
   Name Current System: automatic.moe
   Name Output Database: automatic.mdb
   Sequence: Chain #1
   Template: Chain #2
   Uncheck "disable C-terminal ..."
   Set Models: 3
   Forcefield: Amber12
   Leave everything else as is.
   OK
Demonstration of the power of re-alignment

7. Close
8. Open "before.moe"
9. Push the alignment to make it look like this

10. **Protein > Homology Model**
    - Name **Current System**: realigned.moe
    - Name **Output Database**: realigned.mdb
    - Sequence: Chain #1
    - Template: Chain #2
    - Uncheck "disable C-terminal ..."
    - Set **Models**: 3
    - **Forcefield**: Amber12
    - Leave everything else as is.

**OK**
Demonstration of the power of re-alignment

11. Close
12. Open "automatic.mdb"
13. Mouse-over right-mouse > send to MOE each entry.
   When it asks whether to clear system, say No.
14. Protein > Geometry > Phi-Psi plot (Ramachandran plot)
15. How many outliers? Where are they?

16. Look at the results

17. Close
18. Open "realigned.mdb"
19. Mouse-over right-mouse > send to MOE each entry.
   When it asks whether to clear system, say No.
20. Protein > Geometry > Phi-Psi plot (Ramachandran plot)
21. How many outliers? Where are they?

22. Look at the results
Demonstration of the power of re-alignment

16. Look at the results: automatic alignment

<table>
<thead>
<tr>
<th>Tag</th>
<th>Chain</th>
</tr>
</thead>
<tbody>
<tr>
<td>⭐ 1: strepto 3</td>
<td>TYKLIVKGSNTFSGETTTKAVDVETAESFKQYANENGKEKQYGEWSFDDTKLTFVTVAE</td>
</tr>
<tr>
<td>2IGD 2: 2IGD.A</td>
<td>MTGAVTYKLVINGKTLKGETTTKAVDADTAEKFQKYGANDNG--VDGVWYTDDATK--TFTVTE</td>
</tr>
</tbody>
</table>

Lots of outliers. Not coincidentally, in the indels.
Demonstration of the power of re-alignment

**22. Look at the results**

<table>
<thead>
<tr>
<th>Tag</th>
<th>Chain</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: strepto 3</td>
<td></td>
<td>TYKLIVKGNFSTGETTTTAKAVDETAEKSFKQYNENG-EKVYGEWSFD-TKLTFTVTAE</td>
</tr>
<tr>
<td>2IGD 2: 2IGD.A</td>
<td></td>
<td>MTPAVTYKLIVINGKTLKGETTTTAKAETAEKAFKQYNANGVD--GVWTYDDAT--KTFTVTE</td>
</tr>
</tbody>
</table>

Few outliers. Still in the indels
23. Find outliers in the **phi-psi plot**.
24. Select the sequence around one of the outliers.
   - **Unfix** it. **Invert** selection. **Fix**.
25. *run* 'gizcolorf.svl'
   - Blue means little force. Red means high force, on an atom.
26. *run* 'gizmin.svl'
27. Pull and push to make regular H-bonds, common secondary and local structures. Make sure it doesn't get too red!
28. Keep **Protein > Geometry > phi-psi** plot on. You are done when all of the outliers have moved to the allowed regions.
29. Unfix additional residues if necessary.
30. Save and upload the MOE file as "Exercise 5". Due Feb 17.

Continue working with your model
Rotamers

Sidechain conformations fall into three classes called rotational isomers, or **rotamers**.

A random sampling of Phenylalanine sidechains, w/ backbone superimposed
Sidechain rotamers

1-4 interactions differ greatly in energy depending on the moieties involved.

-60° gauche
180° anti/trans
+60° gauche
Rotamer stability is dependent on the backbone $\phi\psi$ angles.

W sidechain is shown here lying over Thr backbone.

<table>
<thead>
<tr>
<th>Rotamers of W*: $\phi$ $\psi$</th>
<th>$P_{\phi=-140,\psi=160}$</th>
<th>$P_{\phi=-60,\psi=-40}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>p-90 +60 -90</td>
<td>0.372</td>
<td>0.079</td>
</tr>
<tr>
<td>p90  +60 +90</td>
<td>0.238</td>
<td>0.005</td>
</tr>
<tr>
<td>t-105 180 -105</td>
<td>0.033</td>
<td>0.251</td>
</tr>
<tr>
<td>t90  180 90</td>
<td>0.021</td>
<td>0.268</td>
</tr>
<tr>
<td>m0  -65 5</td>
<td>0.038</td>
<td>0.124</td>
</tr>
<tr>
<td>m95  -65 95</td>
<td>0.183</td>
<td>0.203</td>
</tr>
</tbody>
</table>
Rotamer libraries have been compiled by clustering the sidechains of each amino acid over the whole database. Each cluster is a representative conformation (or rotamer), and is represented in the library by the best sidechain angles (chi angles), the "centroid" angles, for that cluster.

Two commonly used rotamer libraries:

* Jane & David Richardson: http://kinemage.biochem.duke.edu/databases/rotamer.php

Roland Dunbrack: http://dunbrack.fccc.edu/bbdep/index.php

*rotamers of W on the previous page are from the Richardson library.
Rotamer explorer

- Pulls side chain rotamers from a rotamer library. Superposes on the backbone.
- Note difference in chi angles (side chain torsion angles)
- Only displays rotamers that have a good energy.
- Energy is calculated relative to the best rotamer.
- *Not shown in Rotamer Explorer*: Rotamers have intrinsic energies. Some are intrinsically better than others.
Try Rotamer explorer

- Study your structure. Find a side chain that is not "happy". Select it.
- **Protein | Rotamer explorer**
- Get from MOE
- **Explore**
  - click through results
  - inspect.
- **Mutate.**
- Repeat.
Review questions

• After Homology Model, are we done making a homology model?
• If you have a glycine and it is strictly conserved, where might $\phi$ lie in the $\phi$-$\psi$ (Ramachandran) plot?
• How do you as a protein modeler feel about a buried, unsatisfied, backbone hydrogen bond donor or acceptor?
• How can we guess which donor H-bonds to which acceptor?
• Are side chain rotamers conserved in regions where sequence is conserved? (usually, never, always)
• What variables define side chain rotamers?
• What makes some rotamers intrinsically lower energy than others?
• What locations tend to have many possible rotameric states?
• Which amino acids have only 1 rotamer?
• Which amino acids have the most rotamers?