Rotamers
Manual protein design
Packing
Adding waters
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.

Newman projections

"m" -60° gauche

"t" 180° anti/trans

"p" +60° gauche
Fig. 1. Illustration of the relationships and nomenclature for the side-chain dihedral angle $\chi_1$. Each of the three staggered conformations is labeled with its $\chi_1$ angle (measured from the backbone N), its officially correct $g^+$ or $g^-$ designation, and its $p$, $t$, or $m$ nomenclature as used in this work. Note that earlier studies have used opposite $g^+$, $g^-$ designations, as discussed in the text (Nomenclature).
General generic rotamer preference order for $\chi_1$

Energy of...

$m < t < p$

...but, actual rotamer preference depends on
1) the amino acid
2) the backbone conformation
3) packing.
$\chi_1$ backbone determines the preference for $\chi_2$ which determines the preference for $\chi_3$, an energetic decision tree of rotamers.
The energy of a rotamer can be calculated two ways.

– Using a force field. (not very accurate)
– using statistics from the protein data bank. (empirical and accurate)

\[ E_{\text{rot}} = -RT \log \left( \frac{P(r)}{1-P(r)} \right) \]

where \( P(r) \) is the probability of rotamer \( r \).
Rotamer Libraries

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
Richardson rotamer library

G 3 0
-0.000000        G        -       p:0000    900.    900.    900.    900.    0.000
0.0000  0.0000  0.0000  N
1.4700  0.0000  0.0000  CA
2.0200 -0.7140 -1.2400  HB

A 3 0
-0.000000        A        -       p:0000    900.    900.    900.    900.    13.255
0.0000  0.0000  0.0000  N
1.4700  0.0000  0.0000  CA
2.0200 -0.7140 -1.2400  CB

C 4 1
0.377250        C        -       p:1000     52.    900.    900.    900.    33.851
0.0000  0.0000  0.0000  N
1.4700  0.0000  0.0000  CA
2.0200 -0.7140 -1.2400  CB
1.2558 -0.0213 -2.7466  SG
1.6794 -0.5784 -3.7820  HG
0.377250        C        -       p:2000     62.    900.    900.    900.    34.855
0.0000  0.0000  0.0000  N
1.4700  0.0000  0.0000  CA
2.0200 -0.7140 -1.2400  CB
1.4900  0.1650 -2.7500  SG
2.1829 -0.2329 -3.7575  HG
0.377250        C        -       p:3000     72.    900.    900.    900.    33.924
0.0000  0.0000  0.0000  N
1.4700  0.0000  0.0000  CA
2.0200 -0.7140 -1.2400  CB
1.7470  0.3158 -2.7228  SG
2.1829 -0.2329 -3.7575  HG
0.232074        C        -       t:1000    173.    900.    900.    900.    33.457
0.0000  0.0000  0.0000  N

AA 1-letter code
rotamer name, dihedral angle(s).  900 means "n/a"

Stats compiled from a set of 240 high resolution PDB structures.

Rotamer stability depends on $\phi\psi$

W sidechain is shown here lying over Thr backbone

Rotamers of W*: $\chi_1, \chi_2$

| Rotamers of W* | $\chi_1$ | $\chi_2$ | $P|\phi=-140,\psi=160$ | $P|\phi=-60,\psi=-40$ |
|----------------|---------|---------|----------------|----------------|
| p-90           | +60     | -90     | 0.372          | 0.079          |
| p90            | +60     | +90     | 0.238          | 0.005          |
| t-105          | 180     | -105    | 0.033          | 0.251          |
| t90            | 180     | 90      | 0.021          | 0.268          |
| m0             | -65     | 5       | 0.038          | 0.124          |
| m95            | -65     | 95      | 0.183          | 0.203          |

Roland Dunbrack's is a backbone-dependent rotamer library
\( \phi \psi \) determines preference for \( \chi_1 \), determines the preference for \( \chi_2 \), determines the preference for \( \chi_3 \).
Exercise 21.2: Protein design using Protein | Protein builder

• Open "messedup.moe"

• Using **Protein | Protein builder**, find a better rotamer.

• Select several sidechains that are in mutual contact. Click **REPACK**. wait. What happened?

• **Protein design**: Select a buried sidechain that is too small, In **targets** add large sidechains Trp, Phe. Hit **Rotamers**. Inspect. Select a Trp rotamer. **Keep**.

• Select sidechains near the new Trp side chain. Click **REPACK**. Is the new Trp "happy" where you put it? (Happy means no clashes, no buried H-bond donor/acceptors, no holes, good shape complimentarity.)

• **Design** more residues this way.
The Protein Core:
Nature abhors a vacuum
Complementary surfaces leave relatively little unfilled (void) space. Protein cores are well-packed with little empty space.
close packing

is not the way of proteins.$^{16}$
Protein side chain packing is loose.
Empty space inside a protein is wide enough for wiggle room, but not wide enough to drive a water molecule through.

About 1/2 an atom wide.

If you add that to the radius of two carbon atoms, you get a typical carbon–carbon non-bonded distance = $1.5 + 1.5 + 0.75 = 3.75\text{Å}$
8Å slab stereo!
waters fills polar pockets and cavities
Adding waters one at a time

- Locate an *unsatisfied hydrogen bond donor or acceptor* with enough space to fit a water, and which is in a buried cavity or pocket.

- **Edit | Build | Molecule** (or **Builder**)
- Clear selections.
- Click **O** (oxygen). OK.
- Select the new water molecule. **EPUSIEPF. Minimize.**
Exercise 21.3  
Adding freeze dried waters

Part 1 -- Get started
*Make sure Select | Synchronize is checked.*
1. Open 1rx5 from PDB within MOE.
2. Potential Setup lower corner menu | Load | Amber14EHT
   1. Select maximum threads.
   2. Fix hydrogens.
   3. Fix charges.
   4. OK.
   You are ready to add waters.

Part 2 -- Hydrate your protein
3. Compute | Simulations | Dynamics
   1. Solvent Setup :
   2. Layer, Water, NaCl 0M, 4.0, Delete far, OK.
   3. Cell Setup: No periodicity (don't change it)
   4. Constrain: light bonds
   5. Rigid water
   6. Time step 0.002 ps
   7. NPA algorithm
   8. OK
4. in SEQ window: select all waters and ions.
5. In MOE window: EPUSIEPF*
6. Minimize.

Now your protein is hydrated. Go to Part 3.

*EPUSIEPF = Edit | Potential | Unfix, Select | Invert, Edit | Potential | Fix
**Freeze dried protein hydrate: remove waters that are exposed to bulk solvent or move too much**

Part 3 -- Freeze dry
1. In SEQ window, select water and ions chains.
2. Select **Selector**, Click **UI (user interface)**
   1. Check Selected Chains
   2. Operation: or
   3. Connectivity | Accessibility
   4. Probe radius 5.0. <---- criterion for bulk water exposure
   5. Exposed. (Some waters and ions are selected)
   6. Molecule. (Now complete water molecules are selected.)
   7. Note the number of atoms selected. Number of waters is that number divided by 3.
   8. In MOE window, Delete selected.
   9. Repeat 5-8, until...
3. No more exposed waters? Is the number of waters left less than 20? Stop. Go to Part 4.
4. Minimize.
   Now your protein is freeze-dried.

Part 4 -- Molecular dynamics
5. Select **Solvent**
6. EPUSIEPF
7. Compute | Simulations | Dynamics
   1. Change name to water.mdb
   2. Uncheck "rigid water"
   3. Protocol: prod {ps=250 T=500} (You may explore a higher or lower temperature if you do this a second time.)
   4. OK. If the simulation does not finish in time, **Cancel | Dynamics** when told.

Part 5 -- Find stable waters
8. Open water.c.250.mdb (opens in database viewer, DBV)
9. DBV: **File | Browse**
10. Hit the play button. Use the slider to set the speed of playback.
11. In MOE window, remove protein atoms, display ribbon, and make waters **spacefill**. Waters sitting in deep energy wells move very little. Waters in shallow energy wells move alot and escape to other energy wells.
12. Select the **five least mobile waters** and color them **light blue**.
13. Stop the animation. Go to the last frame and hit **keep**. (sends frame to MOE window) Close.
14. Display protein as ribbon only. Save the MOE file.
   Upload it to the homework server as Exercise 21.