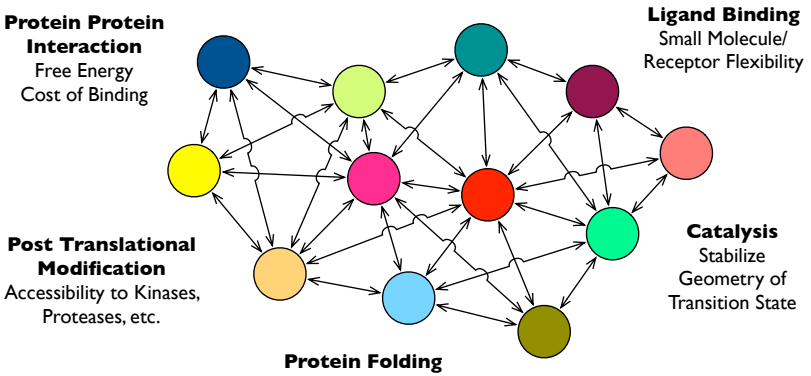
Metropolis Hastings Monte Carlo simulations with RosettaScripts

Colin A. Smith - Kortemme Lab
RosettaCON - Saturday, August 6, 2011

Predicting Flexibility With Monte Carlo Simulations



Rigorous Monte Carlo Must Preserve Detailed Balance

Analogous to the Law of Mass Action:

$A \Leftrightarrow B$

$$K_{eq} = \frac{[B]}{[A]} = \frac{k_{forward}}{k_{reverse}}$$

Solution #2:

$$a_{AB} = \frac{1}{q_{AB}} e^{(E_A - E_B)/kT}$$

$$p_{AB} = q_{AB} a_{AB}$$

$$\frac{\pi_B}{\pi_A} = \frac{p_{AB}}{p_{BA}} = \frac{q_{AB} a_{AB}}{q_{BA} a_{BA}}$$

Population

Transition Probability

Mover Proposal Density

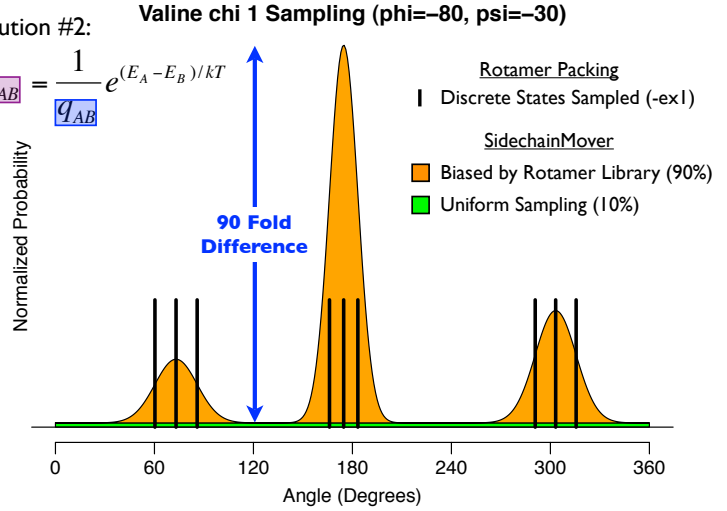
Metropolis Criterion

Problem: $q_{AB} \neq q_{BA}$ Solution #1: $q_{AB} = q_{BA}$

Rotamer Packing vs. SidechainMover

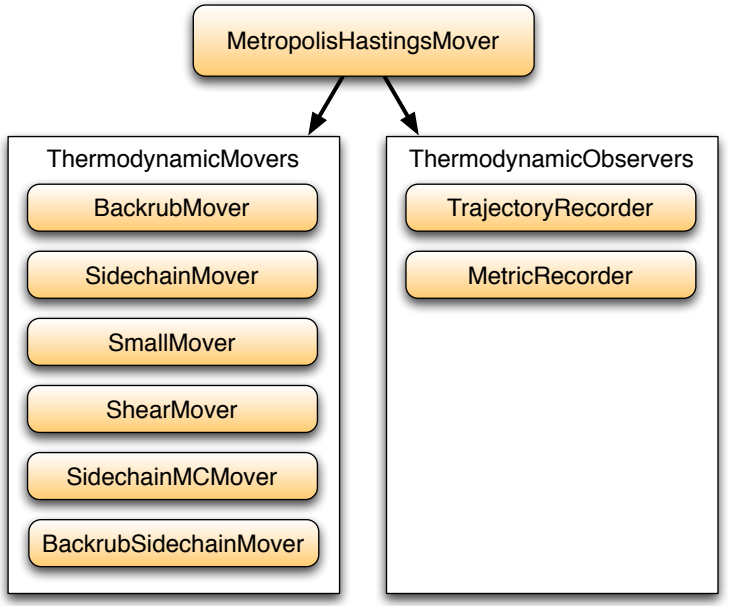
Solution #2:

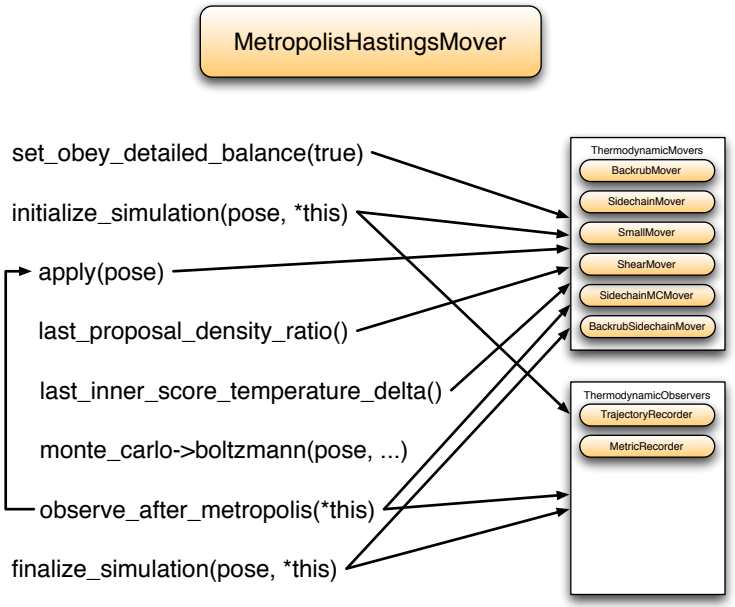
$$a_{AB} = \frac{1}{q_{AB}} e^{(E_A - E_B)/kT}$$



Problems with the backrub app

- Can't upweight sampling of certain regions, only on/off
- No way to automatically use movers others implement
- Many movers have overlapping command line flags
- Difficult to combine with other movers in a larger protocol
- Side chains use PDB numbering, backbone uses absolute Rosetta numbering
- Doesn't support JD2





MetropolisHastingsMover

```
<SCOREFXNS>
   <score12_mh weights=score12_full>
      <Reweight scoretype=mm_bend weight=1/>
   </score12_mh>
</SCOREFXNS>
```

```
<MetropolisHastings name=metropolis_hastings
                    scorefxn=score12_mh
                    temperature=4.8
                    trials=10000000>
```

Including Movers

Traditional

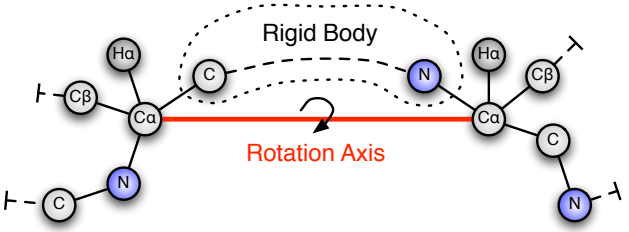
```
<Backrub name=backrub .../>
<MetropolisHastings name=metropolis_hastings ...>
  <Add mover_name=backrub sampling_weight=(1 &Real)/>
</MetropolisHastings>
```

Streamlined

```
<MetropolisHastings name=metropolis_hastings ...>
  <Backrub sampling_weight=(1 &Real) .../>
</MetropolisHastings>
```

BackrubMover

```
<Backrub  sampling_weight=0.6
          pivot_residues=54A,...,130A
          pivot_atoms=CA (default)
          require_mm_bend=1 (default)/>
```

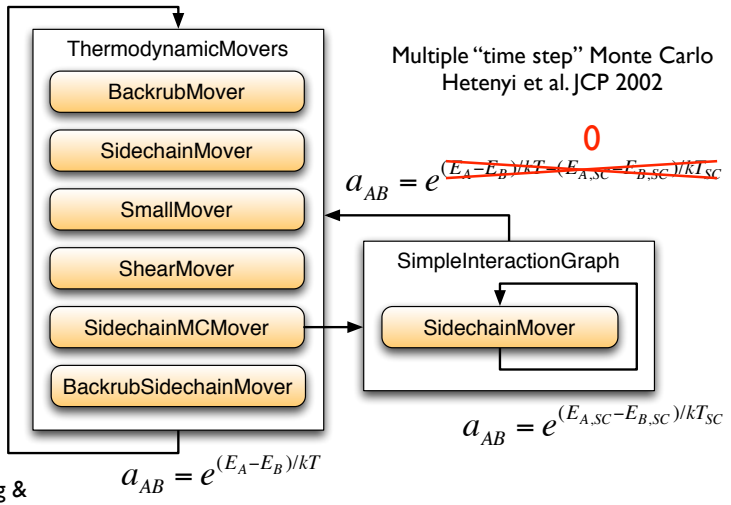


SidechainMover

```
<TASKOPERATIONS>
   <DesignAround name=da repack_on=0 design_shell=0
                 resnums=8A,...,156A/>
   <RestrictToRepacking name=rtrp/>
   <PreserveCBeta name=pcb/> (important for backrub)
</TASKOPERATIONS>
```

```
<Sidechain sampling_weight=0
           task_operations=da,rtrp,pcb
           prob_uniform=0.1
           prob_withinrot=0.2/>
```

SidechainMCMover



Liz Kellogg &
Andrew Leaver-Fay

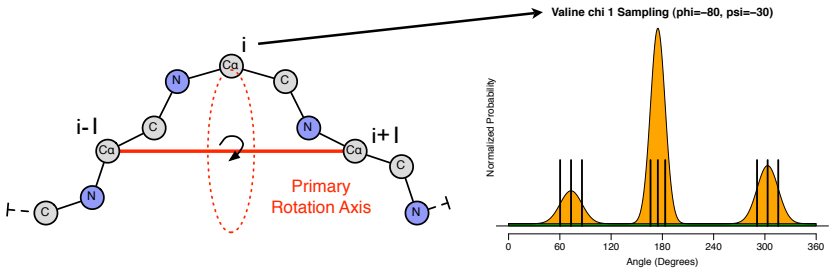
SidechainMCMover

```
<SidechainMC sampling_weight=0.1
             task_operations=da,rtrp,pcb
             prob_uniform=0.1
             prob_withinrot=0.2
             ntrials=1000
OR           scorefxn=score12_mh
             temperature=4.8
             inherit_scorefxn_temperature=1/>
```

Liz Kellogg &
Andrew Leaver-Fay

BackrubSidechainMover

```
<BackrubSidechain sampling_weight=0.3
                  pivot_residues=112A,113A,114A
                  task_operations=da,rtrp,pcb/>
```



ThermodynamicObservers

TrajectoryRecorder

```
<TrajectoryRecorder stride=5000
                    filename=traj.pdb.gz/>
```

MetricRecorder

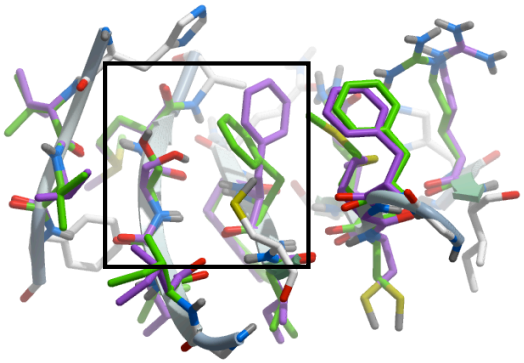
```
<MetricRecorder stride=100 filename=metrics.txt>
    <Torsion rsd=99A type=CHI torsion=1 name="Ser 99 Chi 1"/>
    <Torsion rsd=113A type=CHI torsion=1 name="Phe 113 Chi 1"/>
    <Torsion rsd=113A type=CHI torsion=2 name="Phe 113 Chi 2"/>
</MetricRecorder>
```

Crystallography shows alternate conformations at room temperature

Most of the alternate conformations occur in the 2 beta sheets of the beta sandwich

Fraser et al. *Nature* 2009

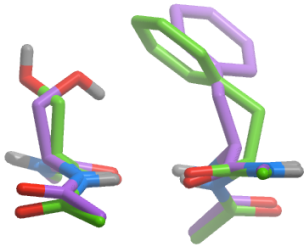
Crystallography shows alternate conformations at room temperature



Major Conformation (A)
Minor Conformation (B)

Fraser et al. *Nature* 2009

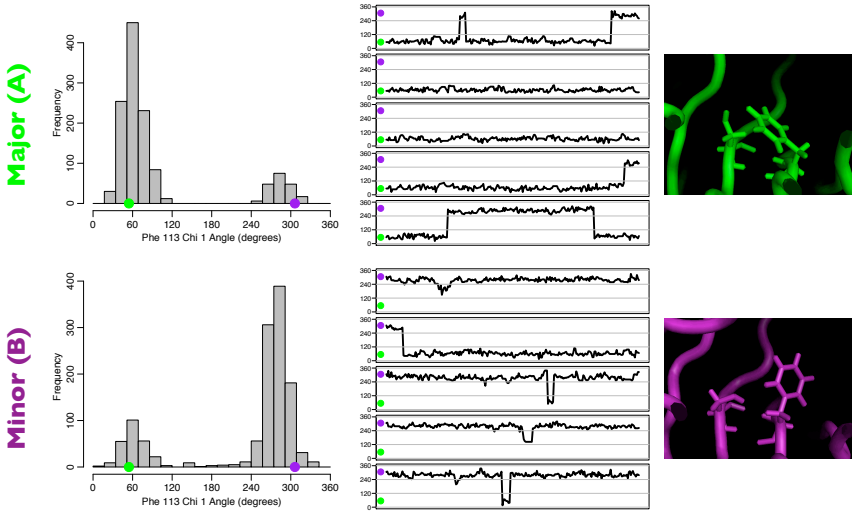
Crystallography shows alternate conformations at room temperature



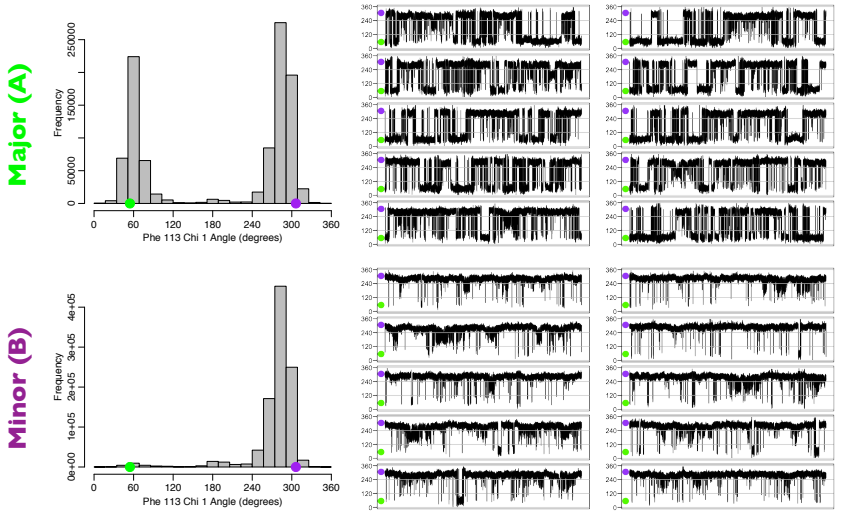
Major Conformation (A)
Minor Conformation (B)

Fraser et al. *Nature* 2009

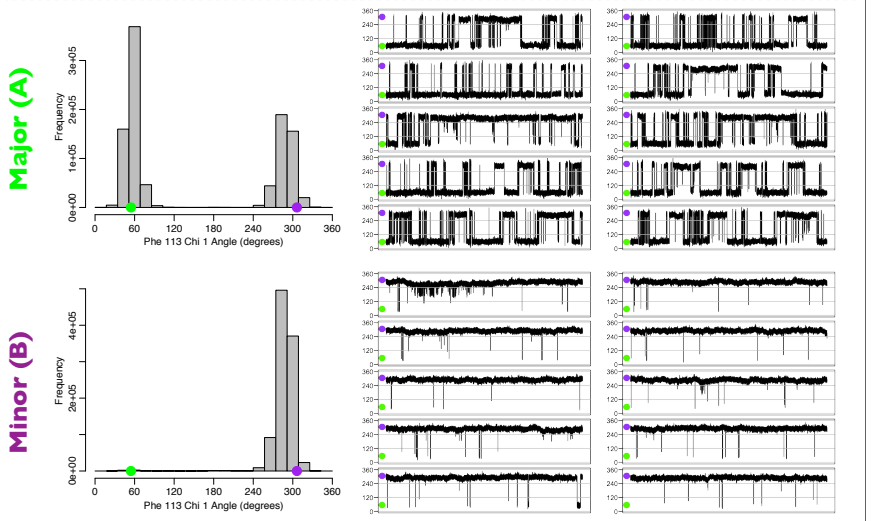
Phenylalanine Intraconverts at 4.8 kT



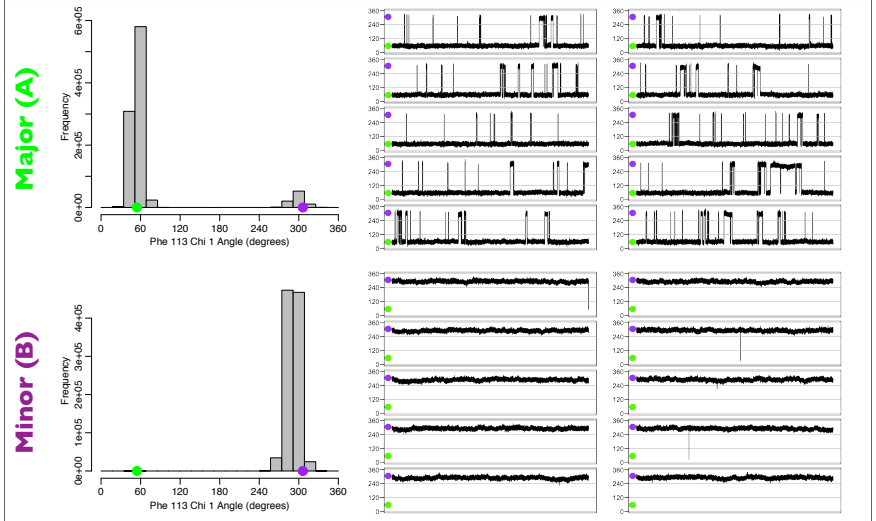

```
<MetropolisHastings temperature=4.8 trials=10000000 ...>
 <Backrub sampling_weight=0.6 pivot_residues=54A,...,130A pivot_atoms=CA/>
 <BackrubSidechain sampling_weight=0.3 pivot_residues=112A,113A,114A .../>
 <SidechainMC sampling_weight=0.1 ntrials=1000 .../>
```



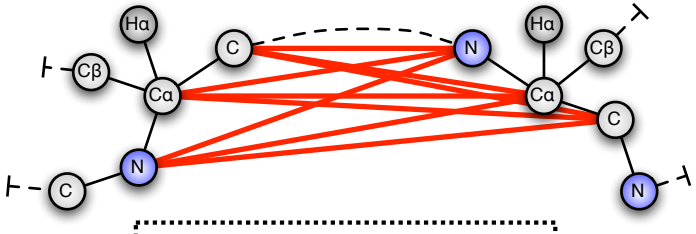
```
<MetropolisHastings temperature=2.4 trials=10000000 .../>
 <Backrub sampling_weight=0.6 pivot_residues=54A,...,130A pivot_atoms=CA/>
 <BackrubSidechain sampling_weight=0.3 pivot_residues=112A,113A,114A .../>
 <SidechainMC sampling_weight=0.1 ntrials=1000 .../>
```




```
<MetropolisHastings temperature=1.2 trials=10000000 .../>
 <Backrub sampling_weight=0.6 pivot_residues=54A,...,130A pivot_atoms=CA/>
 <BackrubSidechain sampling_weight=0.3 pivot_residues=112A,113A,114A .../>
 <SidechainMC sampling_weight=0.1 ntrials=1000 .../>
```

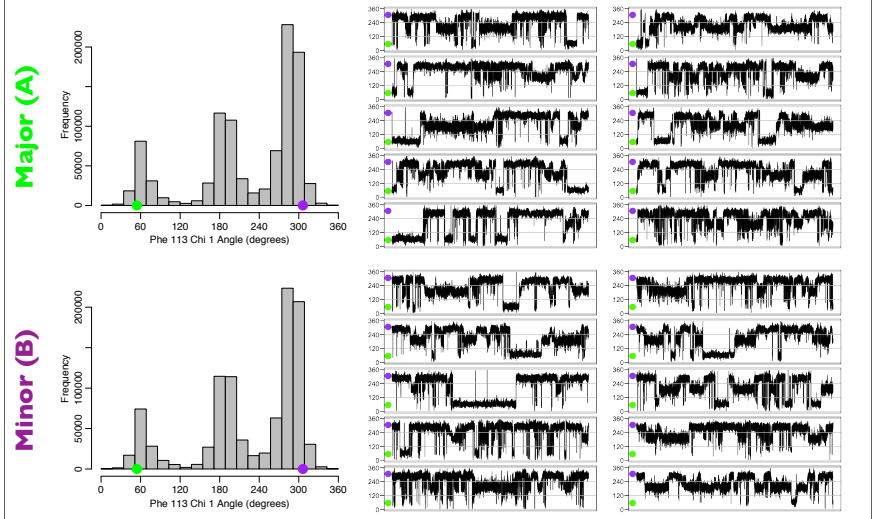


Are unsampled backbone degrees of freedom preventing convergence?




```
<Backrub  ...
          pivot_atoms=N,CA,C/>
```

```
<MetropolisHastings temperature=4.8 trials=10000000 .../>
 <Backrub sampling_weight=0.6 pivot_residues=54A,...,130A pivot_atoms=N,CA,C/>
 <BackrubSidechain sampling_weight=0.3 pivot_residues=112A,113A,114A .../>
 <SidechainMC sampling_weight=0.1 ntrials=1000 .../>
```



Conclusions

- RosettaScripts is awesome. Don't ask why you should use it, ask why you shouldn't.
- Sampling between states at near "room temperature" is still an unsolved issue for long time scale intraconversions like CypA Phe113
- Non-sampled degrees of freedom can bias sampling in important ways

Thanks

David Mavor (UCSF)
James Fraser (UCSF)
Thomas Bliska (Williams College)
Tanja Kortemme (UCSF)
Jerome Nilmeier (UCB)
Yuan Liu (UW)
Liz Kellogg (UW)
Andrew Leaver-Fay (UNC)
David Baker (UW)

NSF, DOD, Genentech