Small Molecules Rotamers: Generating and Testing

Kristian Kaufmann
Vanderbilt University

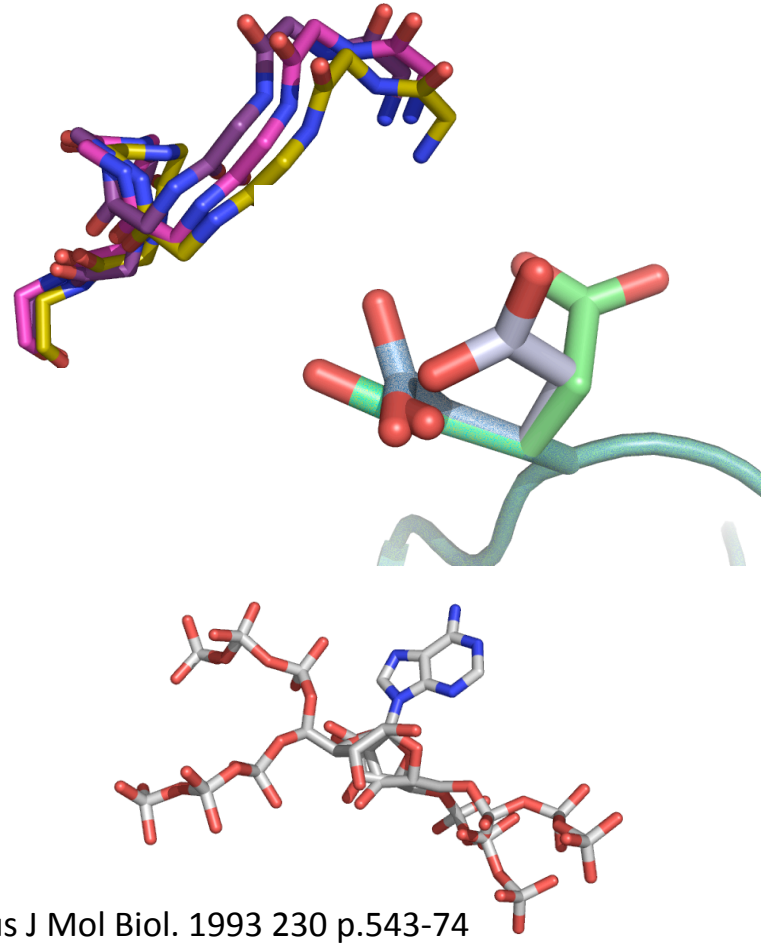
RosettaCon 2008
July 22nd -25th 2008

Generating and Using Rotamers for Small Molecules

- Rotamers construction strategy
- Predictions of conformations on a set of 628 small molecules
- Modifications to ROSETTALIGAND to incorporate small molecule rotamers
- Evaluation of Docking on a set of 7 proteins in complex with 10 small molecules

Building molecular models from fragments molecules captured in crystal structures has proven fruitful

- Fragments for *de novo* protein modelling
- Rotamers for protein side chain modelling
- Corina generates 3D structures from statistics garnered from the Cambridge Structural Database (CSD) of small molecules.
- FleXX an incremental construction algorithm for small molecule docking and design also makes use of the CSD in it small molecule construction



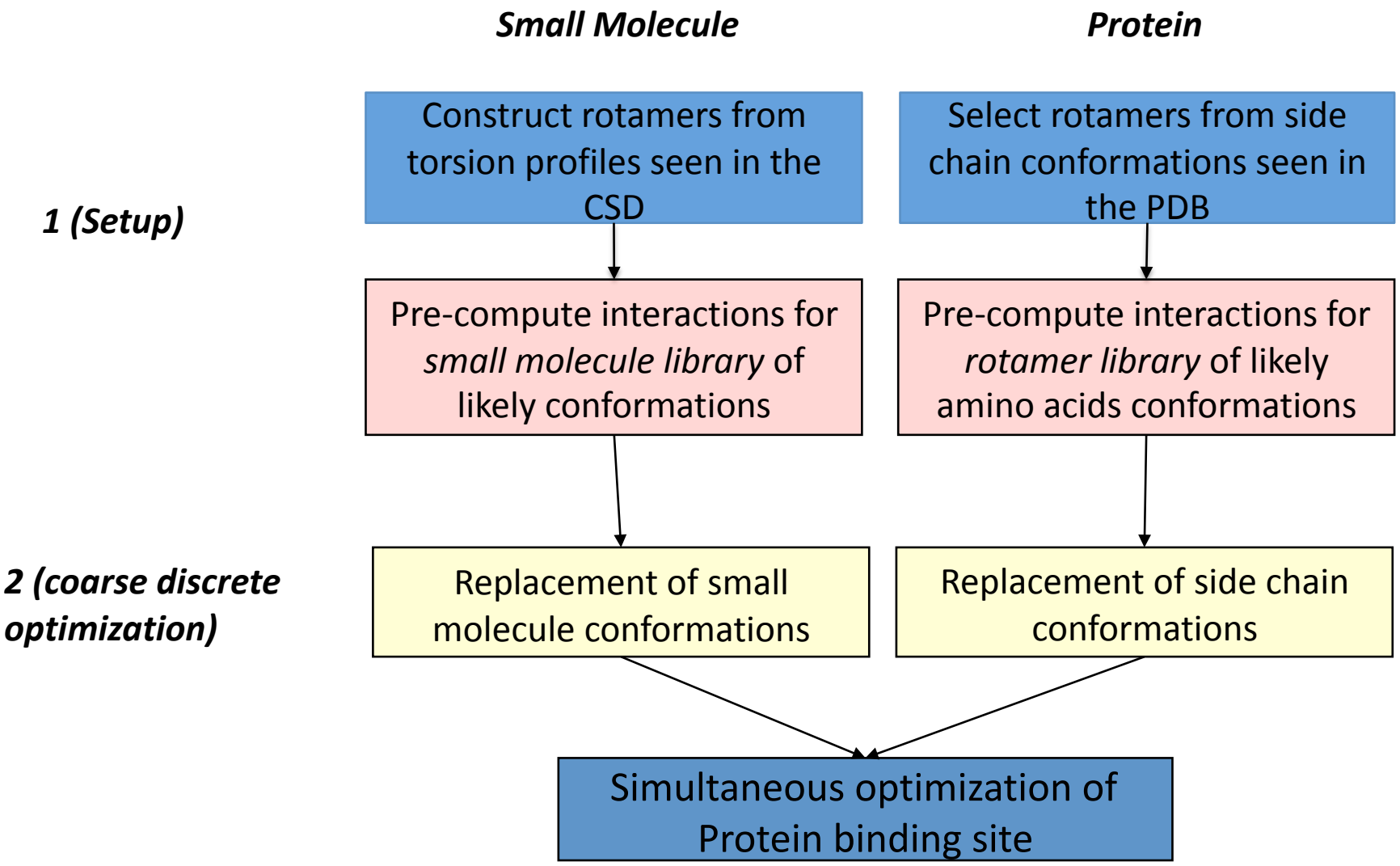
Rohl et al. Methods Enzymol. 2004 383 p.66-93 Dunbrack and Karplus J Mol Biol. 1993 230 p.543-74
Gasteiger et al. Tetrahedron Comp. Method. 1990, 3, 537-547 Rarey et al. J Mol Biol. 1996 261 470-89

Powerful Energy functions can be derived from crystal structures

- Hydrogen bonding function derived through analysis of high resolution protein crystal structures.
- Dunbrack rotamer energy derived from the probability of observing the rotamer in the PDB
- Van der Waals forces derived from atomic distances in crystal structures

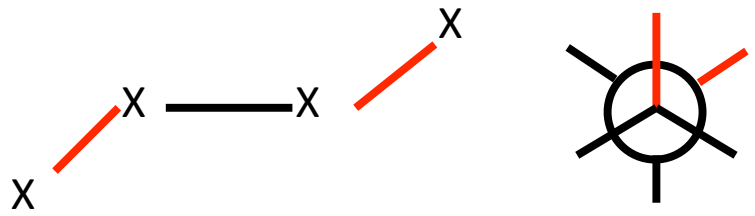
Simons et al. Proteins. 1999 34 p.82-95
Kuhlman et al. Science. 2003 302 p.1364-8

Implementing Small Molecule Flexibility using Rotamers while Retaining Rosetta's Functionality

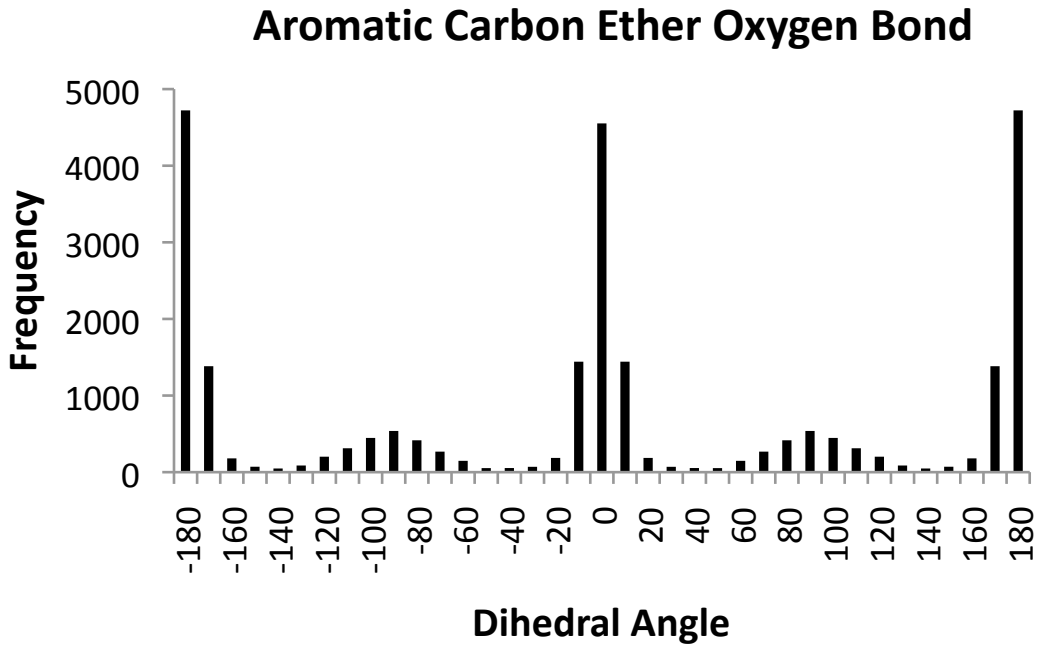


Generating Torsion Profile from the Cambridge Structural Database

Define Atom Types to capture chemical characteristics of dihedral angles

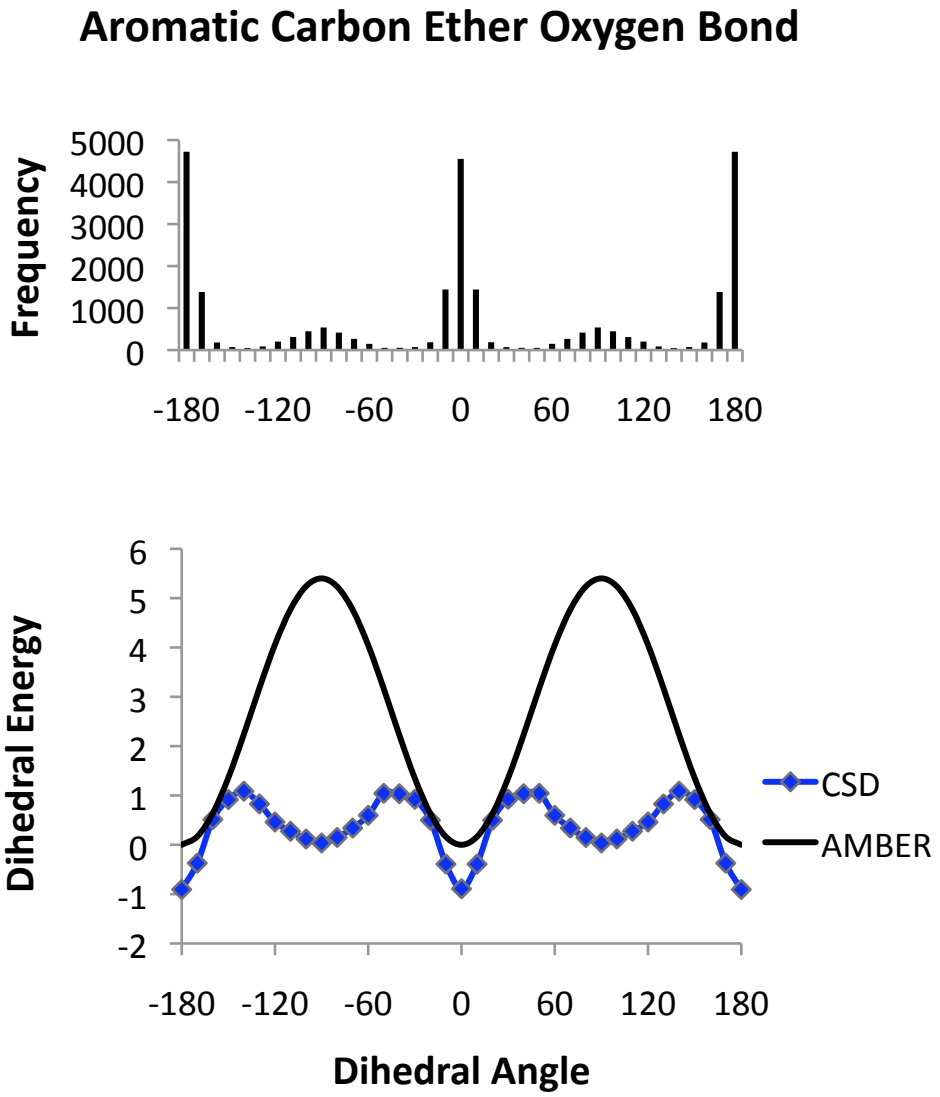


Search Cambridge Structural Database for all atom pairs excluding those in ring systems and measure dihedral. Bin every ten degrees.

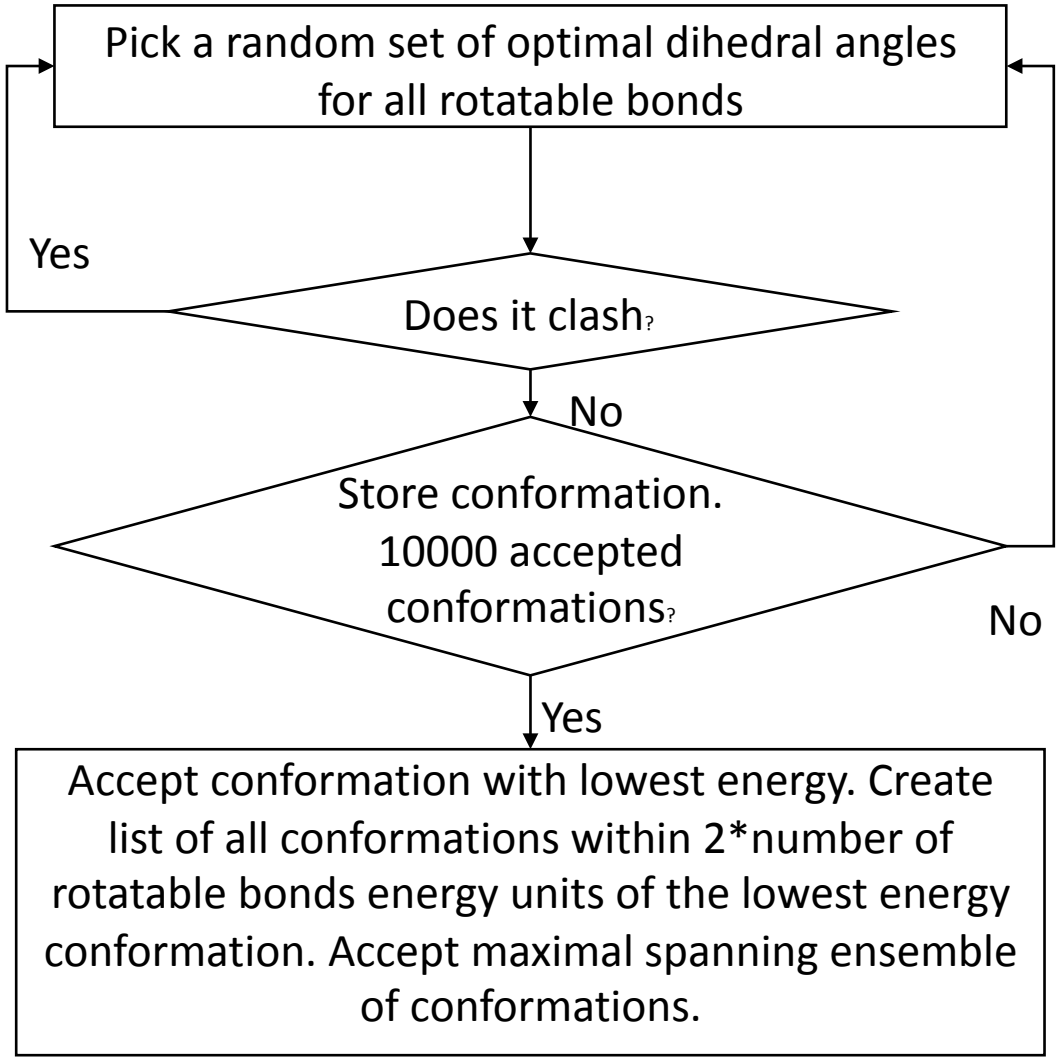


Generating Energies from a Torsion Profile

- Energies can be generated by computing the negative log of the propensity of a state
- The propensity is the probability normalized by the random probability of selecting a state.
- Any angle with a energy less than zero has a greater than random probability.
- Minima in the energy profile can then be used in constructing small molecule conformations



Generating Small Molecule Rotamers



Predicting Bioactive Conformations

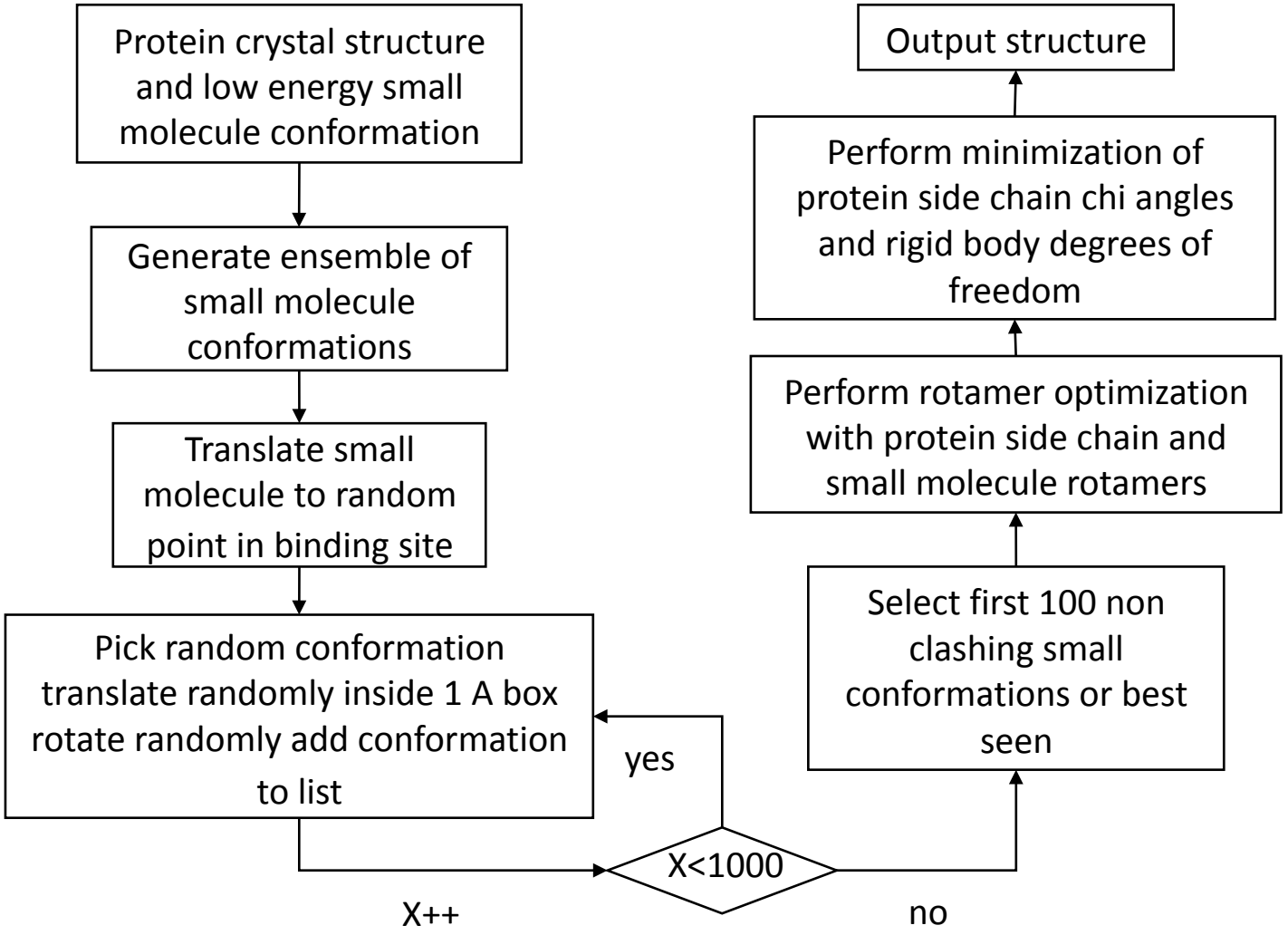
- Test set taken from PDBBind a collection of crystal structures for small molecules from protein complexes with known binding constants
- PDBBind was culled to contain only molecules with ≤ 6 rotatable bonds
- Up to 500 rotamers were generated for each small molecule

Atomic RMSD to crystallized conformation

#of Rotatable Bonds	#of Molecule	Average RMSD of closest conformation	Average RMSD of furthest conformation
1	92	0.14±0.16	1.12±0.47
2	118	0.33±0.26	1.74±0.69
3	118	0.41±0.22	2.13±0.62
4	135	0.47±0.21	2.45±0.69
5	97	0.61±0.30	2.83±0.81
6	118	0.79±0.32	3.07±0.87
Overall Total	628	0.46±0.31	2.23±0.94

Wang et al. J Med Chem. 2005 48 p.4111-9.

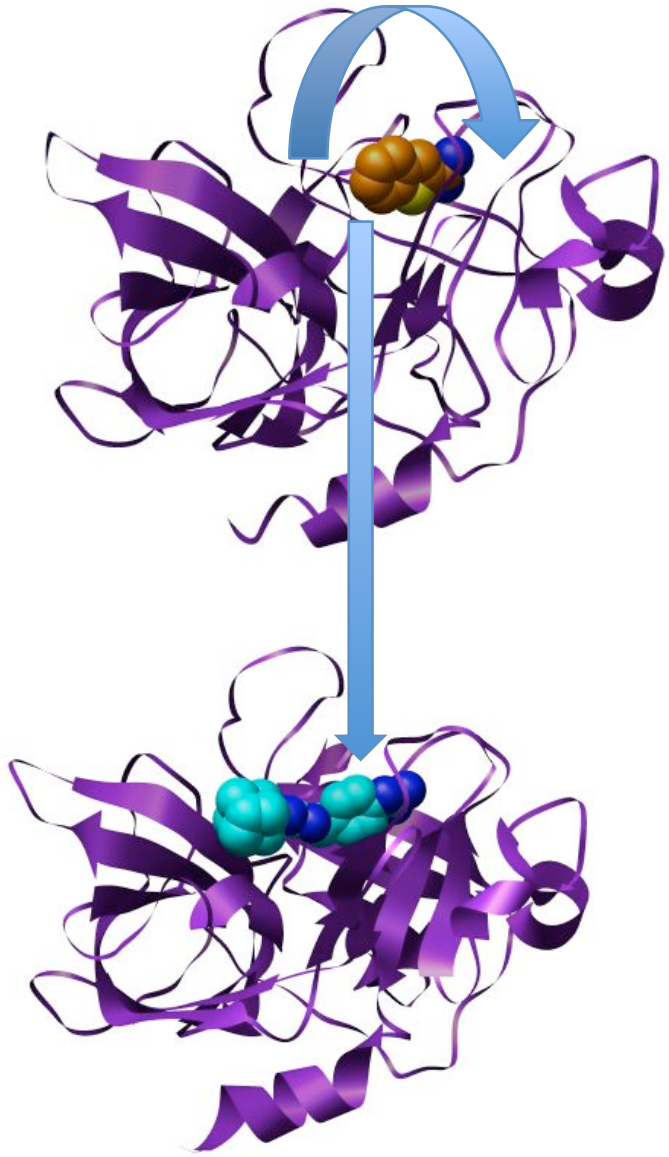
Docking with Small Molecule Rotamers in Rosetta



For original protocol see Meiler and Baker Proteins. 2006 65 p.538-48.

Docking Benchmarks

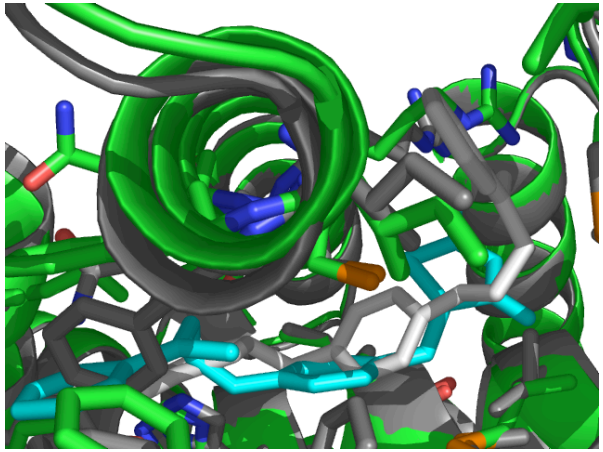
- 15 crystal structures with small molecules
  - 7 proteins, 10 ligands
- 10 self docking examples
  - Verifies protocols ability to sample and identify native state.
- 11 cross docking examples
  - Tests sensitivity to changes in backbone as would be expected from a homology model.

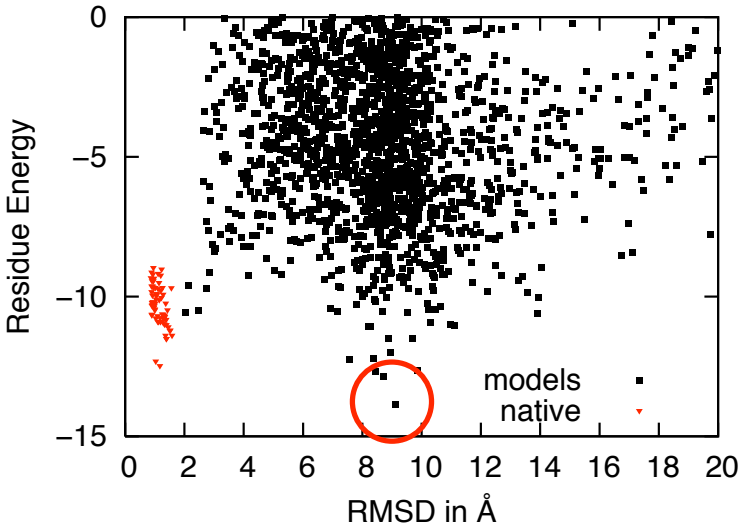


Docking Benchmark

2PRG docking to 1FM9

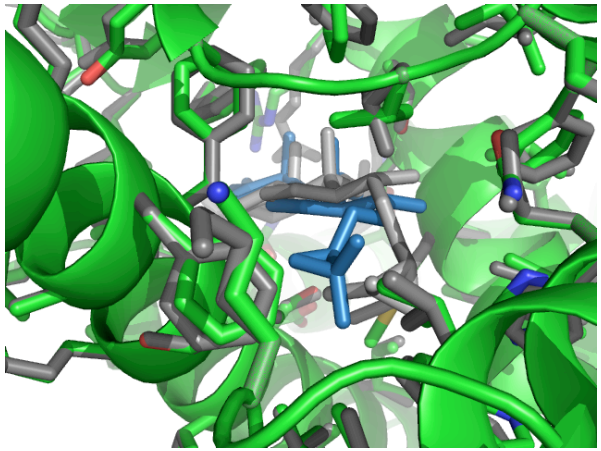
ligand binding domain of the human peroxisome proliferator activated receptor γ

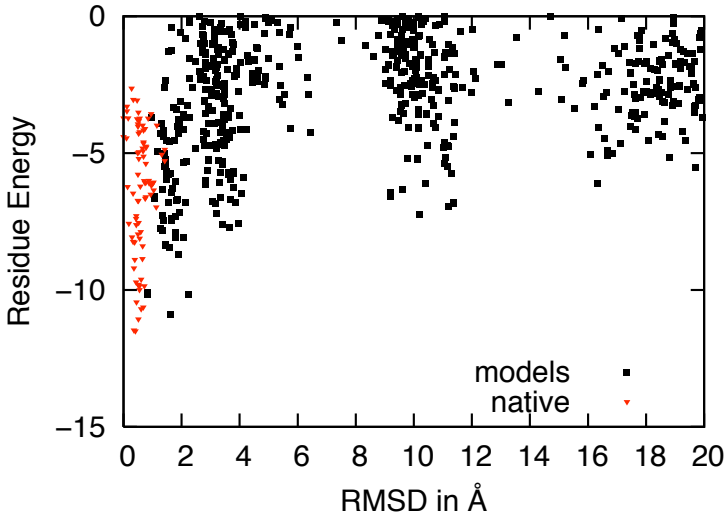




1P8D docking to 1P8D

Oxysterols receptor liver X receptor β.

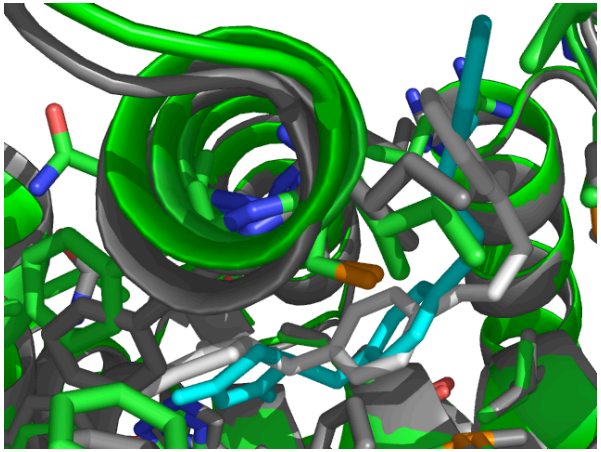


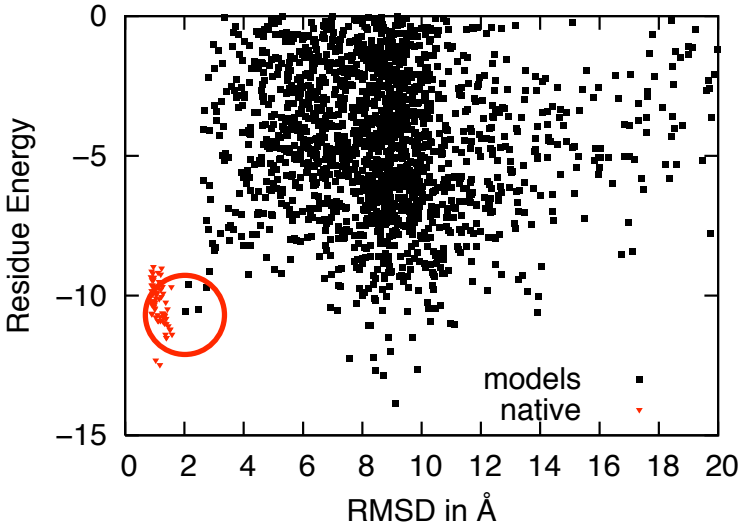


Docking Benchmark

2PRG docking to 1FM9

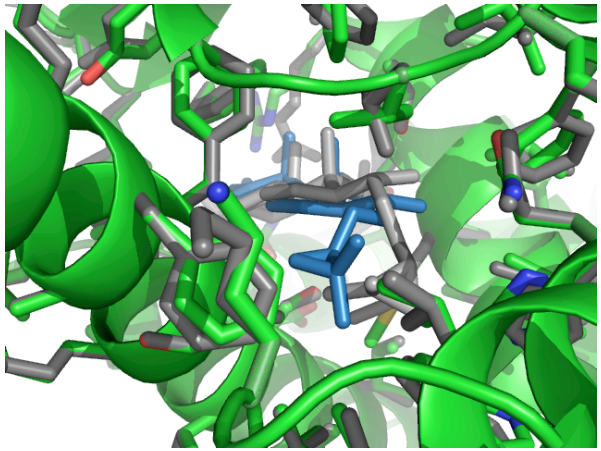
ligand binding domain of the human peroxisome proliferator activated receptor γ

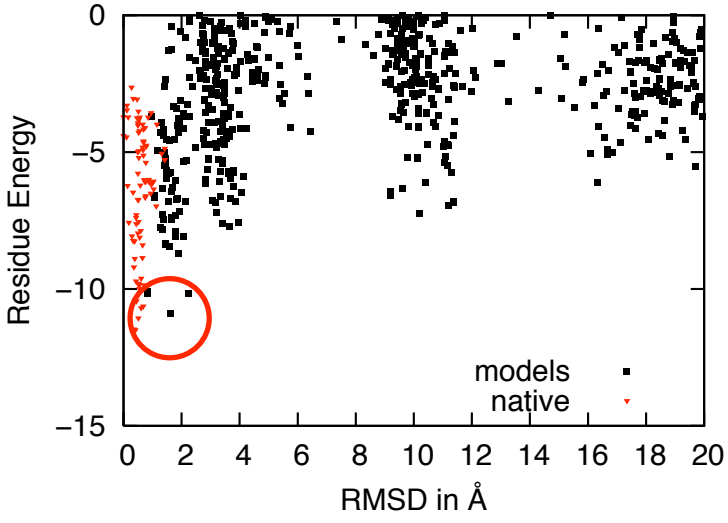




1P8D docking to 1P8D

Oxysterols receptor liver X receptor β.





Performance in a Docking Benchmark

Self Docking Results

Ligand	Protein	rank	RMSD
1aq1	**1aq1**	**1**	**0.56**
1dm2	**1dm2**	**1**	**0.31**
1dbj	**1dbj**	**1**	**1.36**
2dbl	**2dbl**	**1**	**1.45**
1p8d	**1p8d**	**1**	**1.63**
4tim	**4tim**	**1**	**1.87**
6tim	**6tim**	**1**	**1.77**
2ctc	*2ctc*	*3*	*0.82*
1pph	*1pph*	*6*	*1.49*
2prg	2prg	639	1.94

In 9 out of 10 cases self docking was successful when looking at the top 1% by energy

Cross Docking Results

Ligand	Protein	rank	RMSD
1dm2	**1aq1**	**1**	**0.56**
1dbj	**2dbl**	**1**	**1.80**
1pph	*1ppc*	*2*	*1.96*
4tim	*6tim*	*2*	*1.90*
2ctc	*7cpa*	*3*	*0.95*
6tim	*4tim*	*5*	*1.77*
1p8d	*1pqc*	*10*	*1.28*
2prg	*1fm9*	*16*	*2.02*
1p8d	1pq6	181	1.62
2dbl	1dbj	468	3.49
1aq1	1dm2	4296	1.87

In 8 of 11 cases cross docking successful when looking at the top 1% by energy

Conclusions and Future Directions

- Small molecule rotamers present a viable method of representing small molecule flexibility in Rosetta
- 90% of of self docking cases identified native-like structure in the top 1% by energy.
- 72% of cross docking cases identified a native-like structure in the top 1% by energy
- Optimization of the scoring function for discrimination of native like models should yield improved results
- Melding small molecule rotamer approach with incremental construction approach is underway for application to larger small molecules.

Acknowledgements

- Jens Meiler
- Ralf Mueller
- Kristin Glab
- DARPA and NIH
  Molecular Biophysics
  Training Grant