

Combining Experimental Data with Rosetta Computation Models

Elleansar Okwei
(e.okwei@vanderbilt.edu)

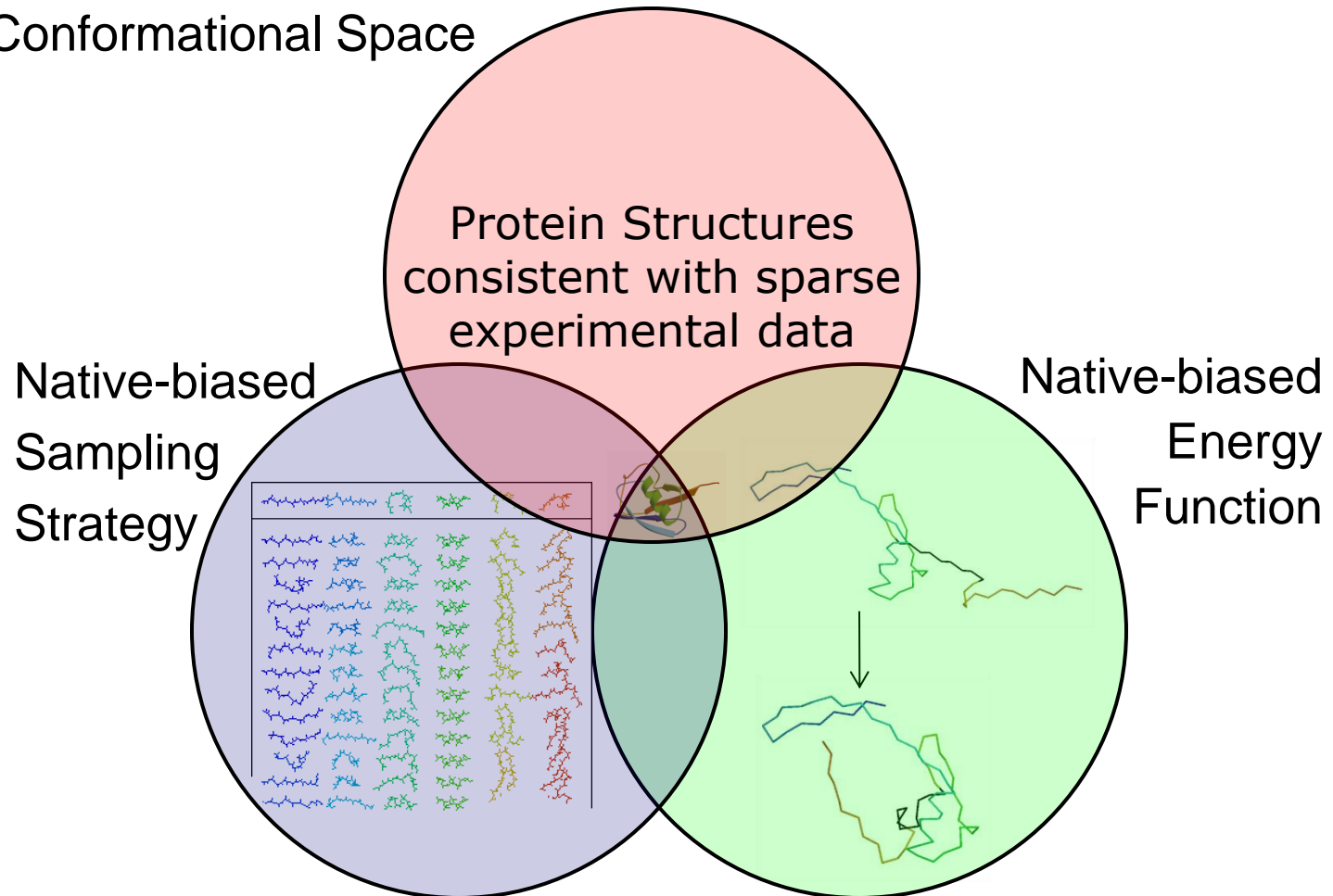


Center for Structural Biology
and the Department of Chemistry

Combining Strengths: Building Accurate Models from Limited Data

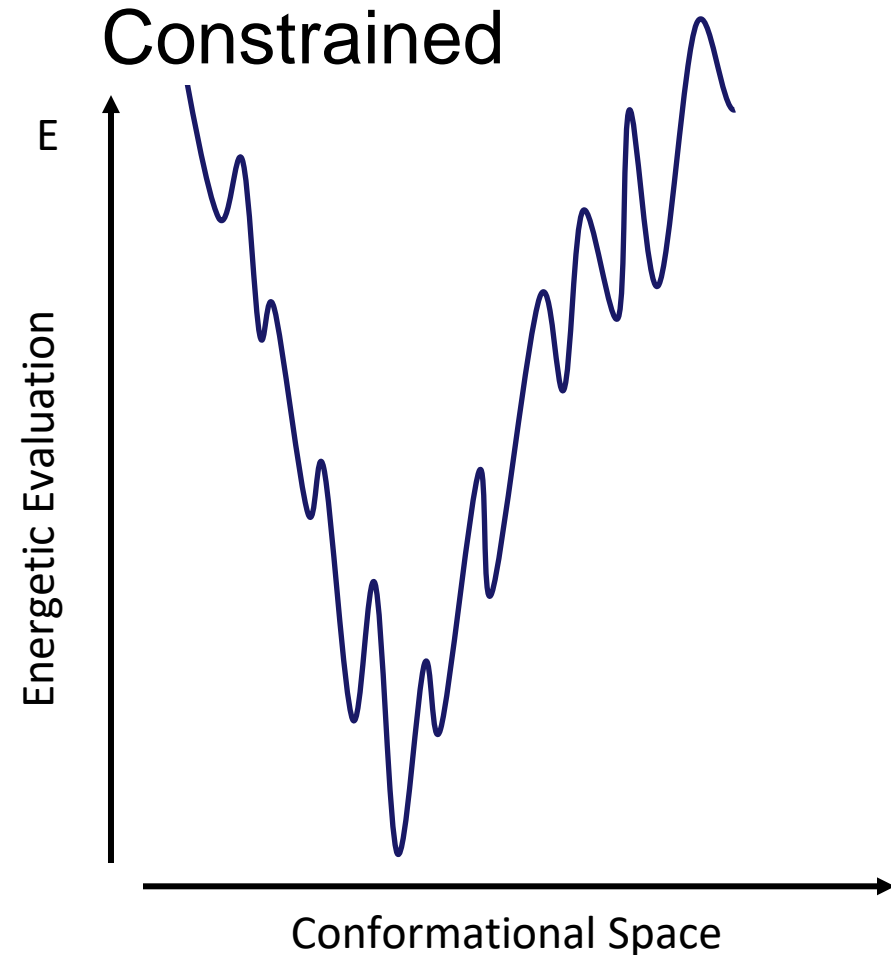
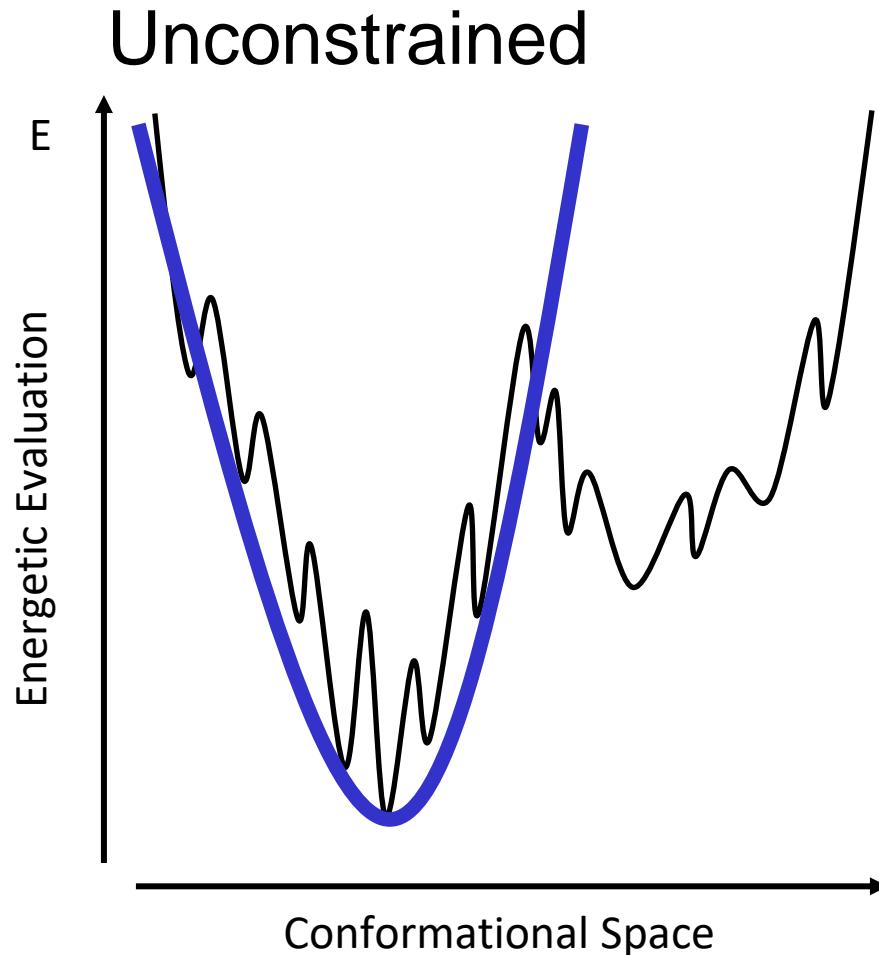


Complete Conformational Space



“Constraints” (Restrains) in Rosetta

“Constraints” alter the energy function



Separation of Measurement and Scoring



AtomPair	NE2	13	V3	32	HARMONIC	0.0	0.2				
Angle	CD2	13	NE2	13	ZN	32	HARMONIC	2.09	0.35		
Dihedral	CG	13	CD2	13	NE2	13	ZN	32	CIRCULARHARMONIC	3.14	0.35

Separation of Measurement and Scoring



```
AtomPair NE2 13 V3 32 HARMONIC 0.0 0.2
```

```
Angle CD2 13 NE2 13 ZN 32 HARMONIC 2.09 0.35
```

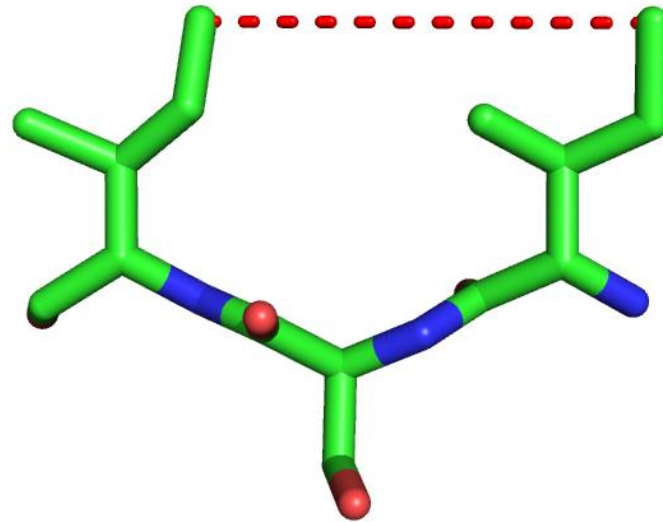
```
Dihedral CG 13 CD2 13 NE2 13 ZN 32 CIRCULARHARMONIC 3.14 0.35
```

- Distance (AtomPair)
- Angle
- Dihedral
- Coordinate
- AmbiguousConstraint
- KofNConstraint
- Harmonic
- CircularHarmonic
- Flat-bottomed Harmonic
- Sigmoid
- Bounded

Measurement Types

```
AtomPair CD1 52 CD1 54 HARMONIC 6.0 0.2
```

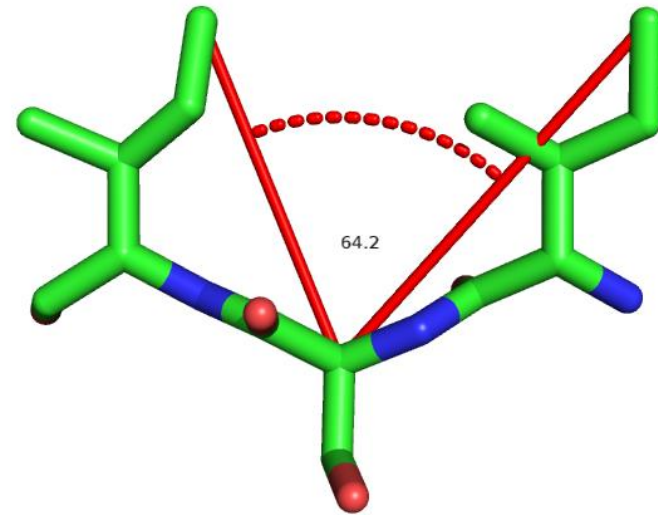
- Distance (AtomPair)
- Angle
- Dihedral
- Coordinate
- AmbiguousConstraint
- KofNConstraint



Measurement Types

Angle	CD1	52	CA	53	CD1	54	HARMONIC	60.0	0.2
-------	-----	----	----	----	-----	----	----------	------	-----

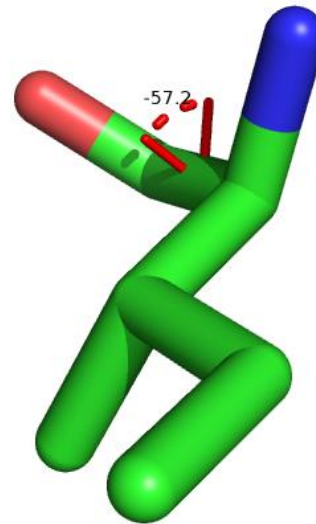
- Distance (AtomPair)
- Angle
- Dihedral
- Coordinate
- AmbiguousConstraint
- KofNConstraint



Measurement Types

```
Dihedral N 52 CA 52 C 52 O 52 HARMONIC 0.0 0.2
```

- Distance (AtomPair)
- Angle
- **Dihedral**
- Coordinate
- AmbiguousConstraint
- KofNConstraint

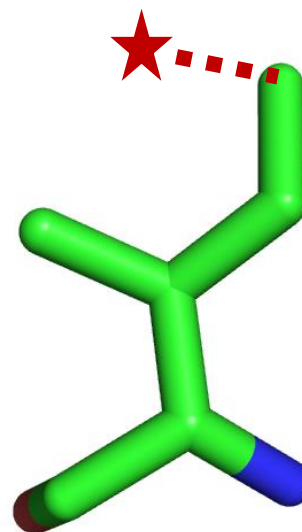


Measurement Types



```
CoordinateConstraint CD1 54 CA 54 X Y Z HARMONIC 0.0 0.2
```

- Distance (AtomPair)
- Angle
- Dihedral
- **Coordinate**
- AmbiguousConstraint
- KofNConstraint



Measurement Types

KofNConstraint 2

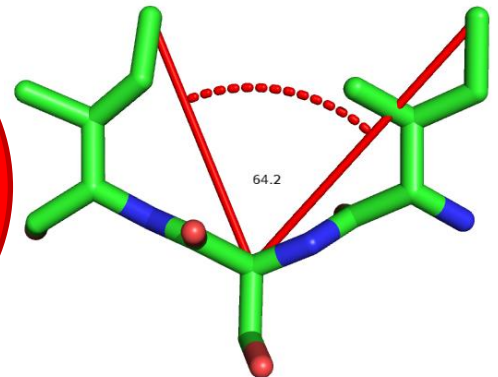
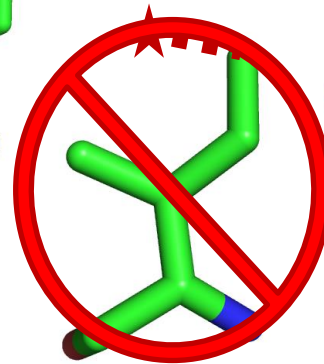
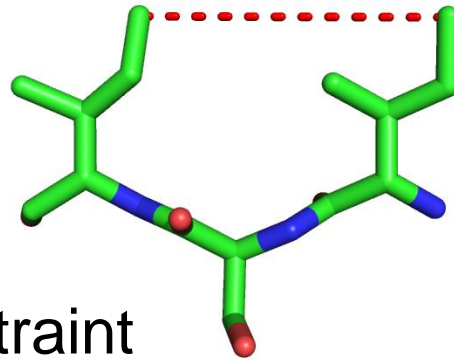
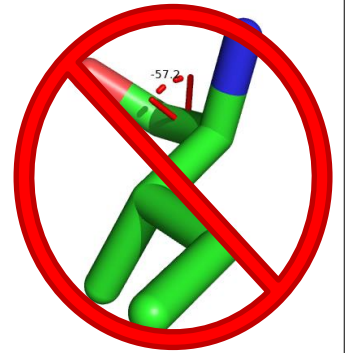
AtomPair CD1 52 CD1 54 HARMONIC 6.0 0.2

Angle CD1 52 CA 53 CD1 54 HARMONIC 60.0 0.2

Dihedral N 52 CA 52 C 52 O 52 HARMONIC 0.0 0.2

CoordinateConstraint CA 1 CD1 54 HARMONIC 0.0 0.2

END



- AmbiguousConstraint
- KofNConstraint

Scoring Types



AtomPair CD1 52 CD1 54 HARMONIC 6.0 1.0

HARMONIC 6.0 3.0

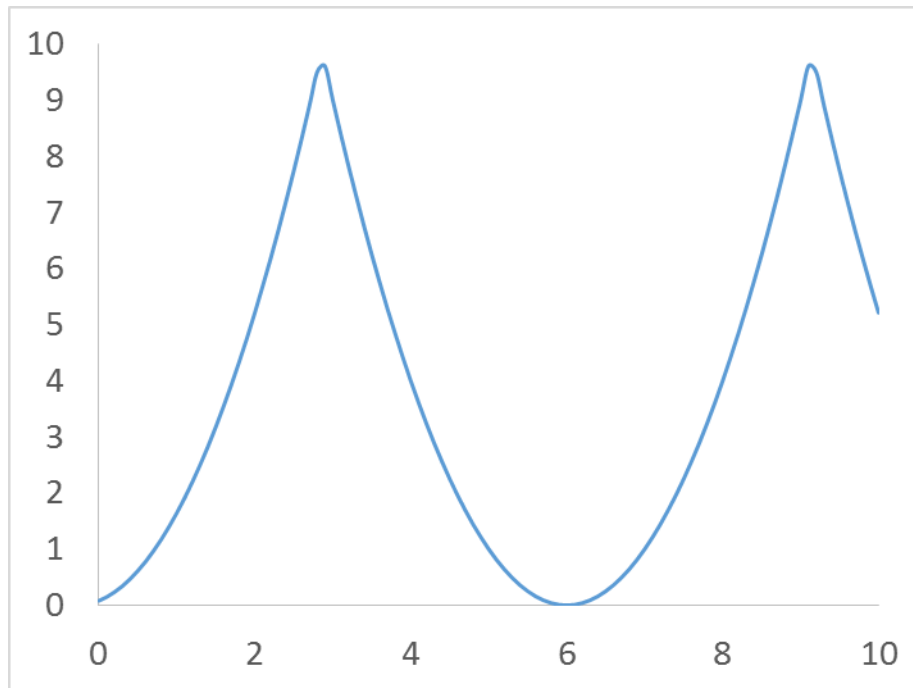


- Harmonic
- CircularHarmonic
- Flat-bottomed Harmonic
- Sigmoid
- Bounded

Scoring Types



```
Dihedral N 52 CA 52 C 52 O 52 CIRCULARHARMONIC 6.0 1.0
```

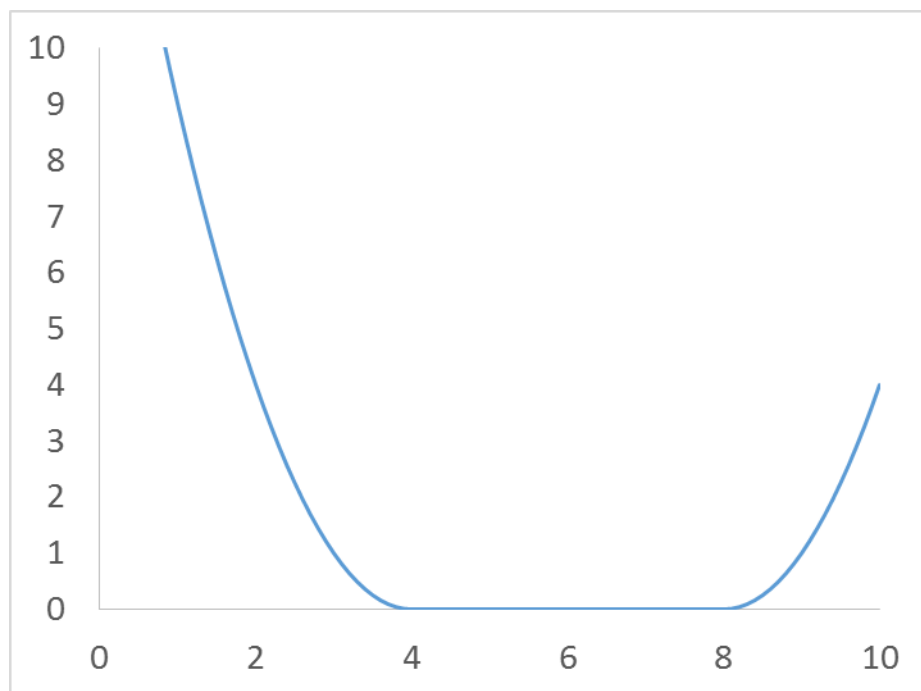


- Harmonic
- **CircularHarmonic**
- Flat-bottomed Harmonic
- Sigmoid
- Bounded

Scoring Types



AtomPair CD1 52 CD1 54 FLAT_HARMONIC 6.0 1.0 2.0

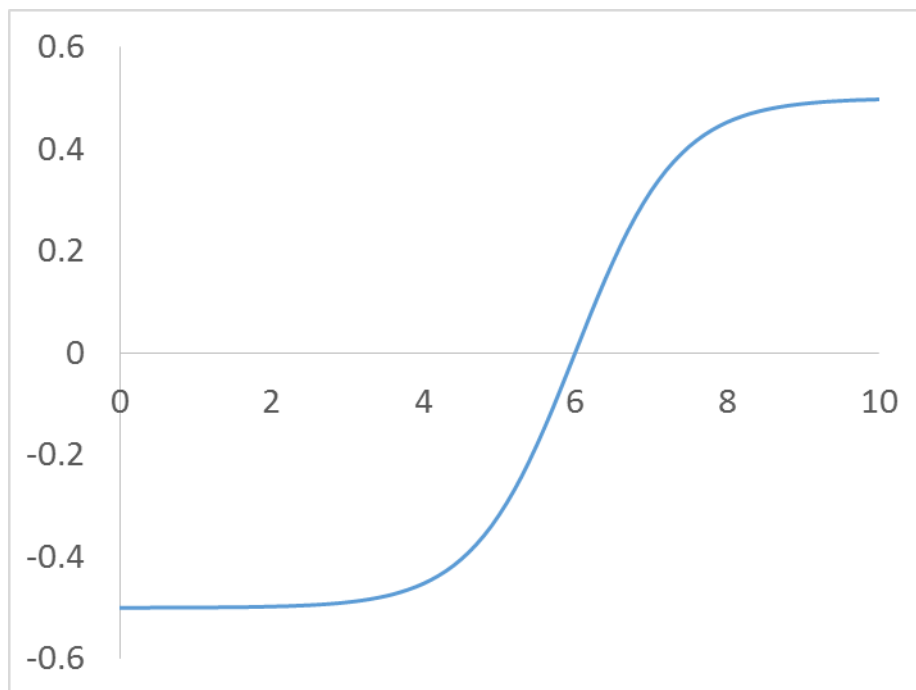


- Harmonic
- CircularHarmonic
- Flat-bottomed Harmonic
- Sigmoid
- Bounded

Scoring Types



AtomPair CD1 52 CD1 54 SIGMOID 6.0 5

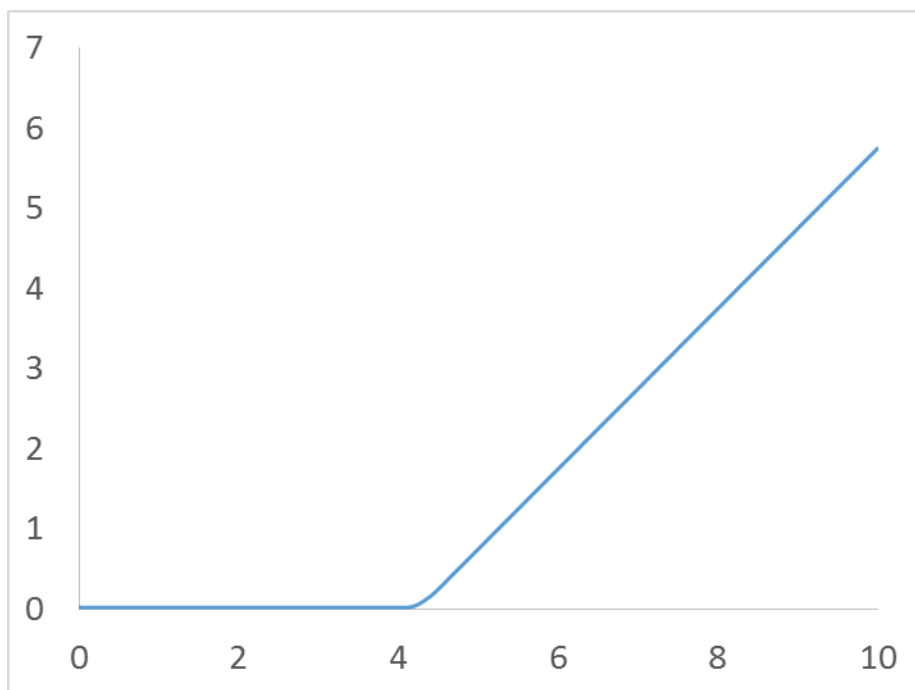


- Harmonic
- CircularHarmonic
- Flat-bottomed Harmonic
- Sigmoid
- Bounded

Scoring Types



AtomPair CD1 52 CD1 54 BOUNDED 0.0 4.0 1.0 0.5 TAG



- Harmonic
- CircularHarmonic
- Flat-bottomed Harmonic
- Sigmoid
- Bounded

Using Constraints: Two Parts



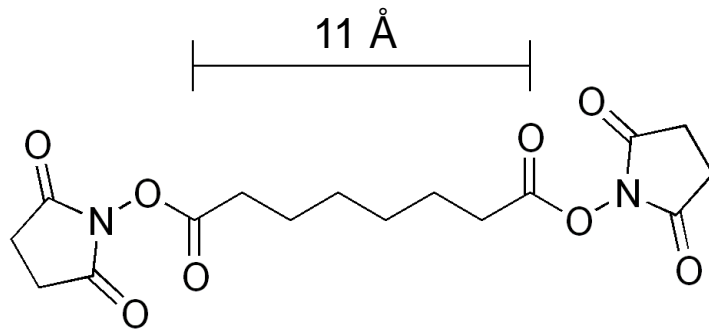
Constraints in the Pose

- Constraint file

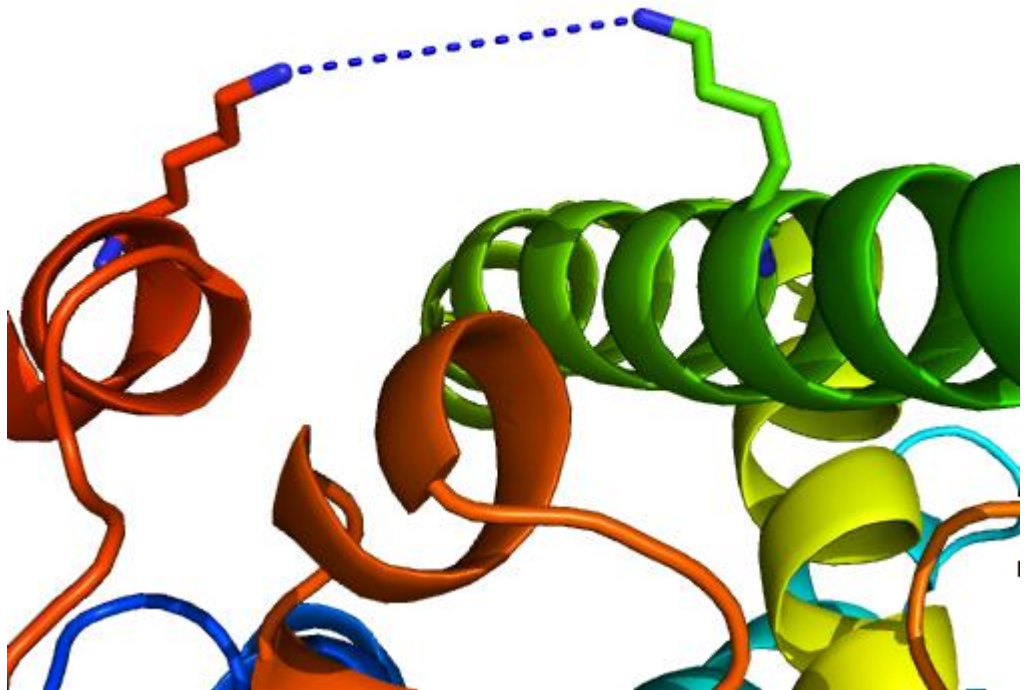
Constraints in the Scorefunction

- Weights File

Crosslinks – direct distance measurement



Disuccinimidyl suberate (DSS)

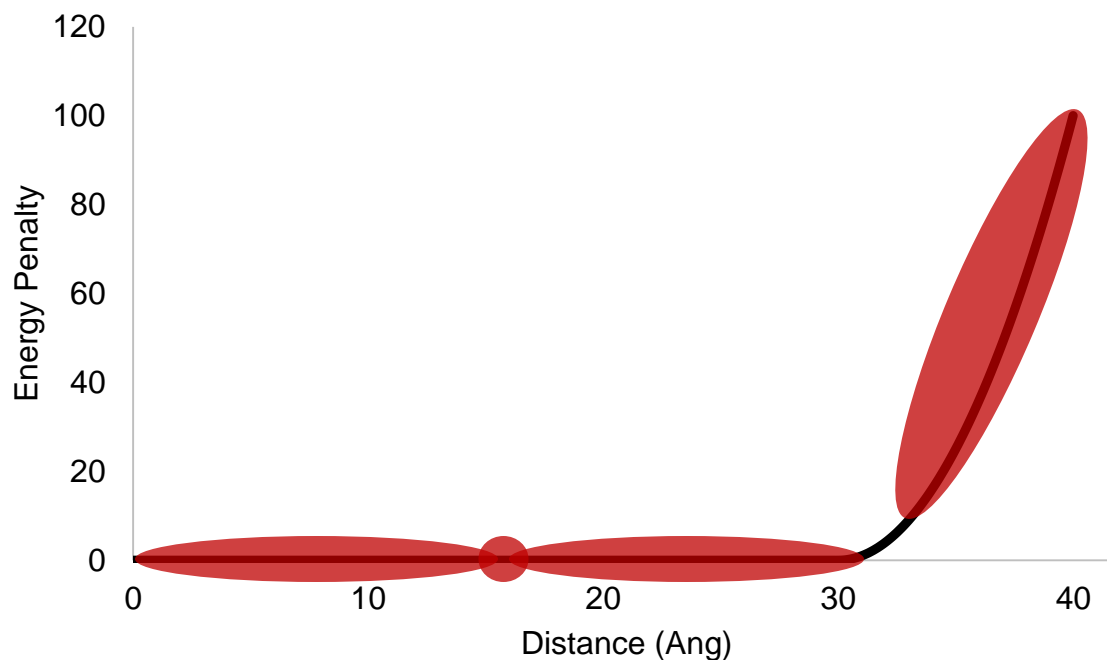


Kahraman et al. PLOS One, 2013, 8(9) e73411
rosetta.demos/protocol_capture/xl_driven_protein_docking/

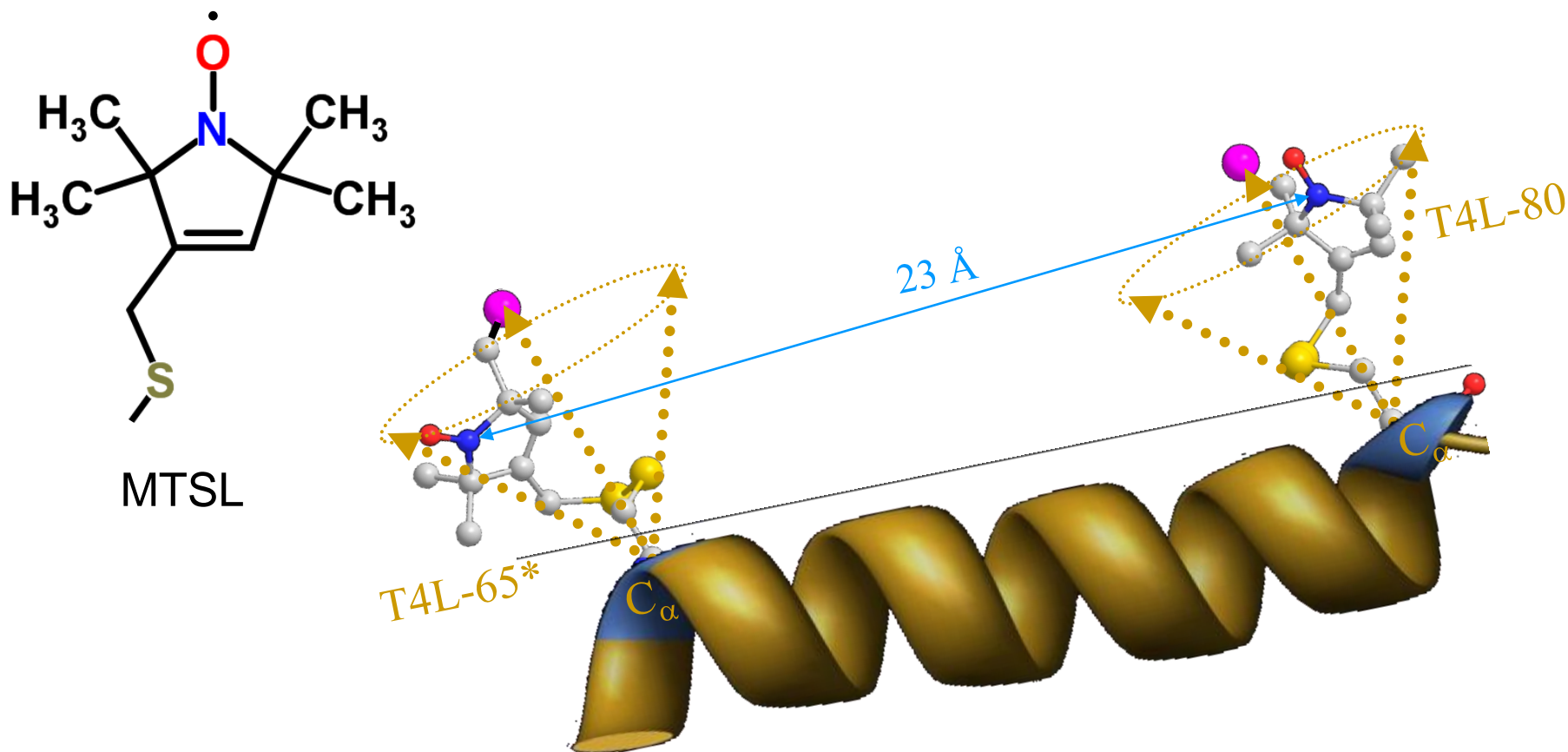
Chemical crosslinking gives distance restraint information



AtomPair CB 67 CB 124	FLATHARMONIC 15.0 1.0 15.0
-----------------------	----------------------------

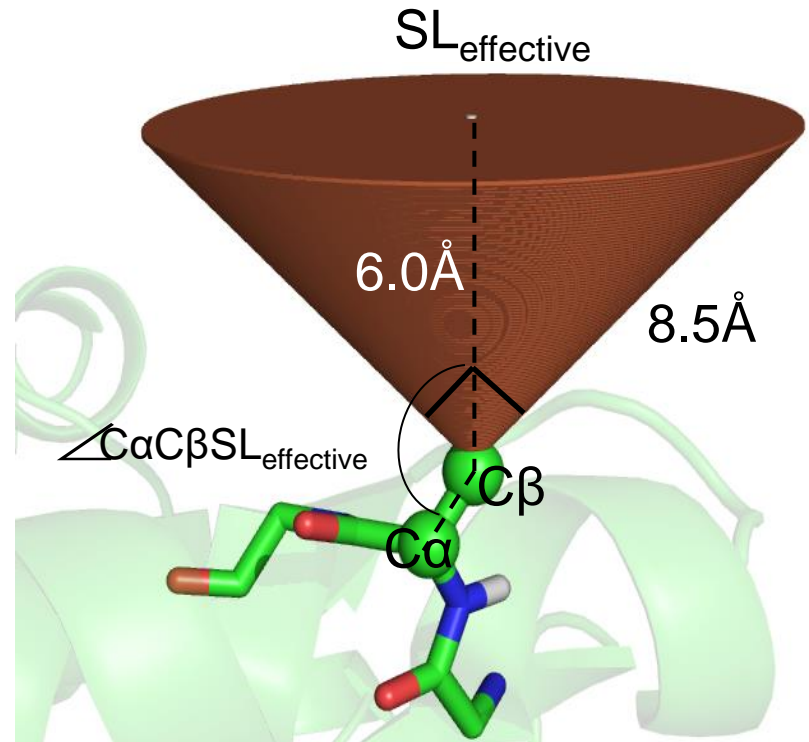
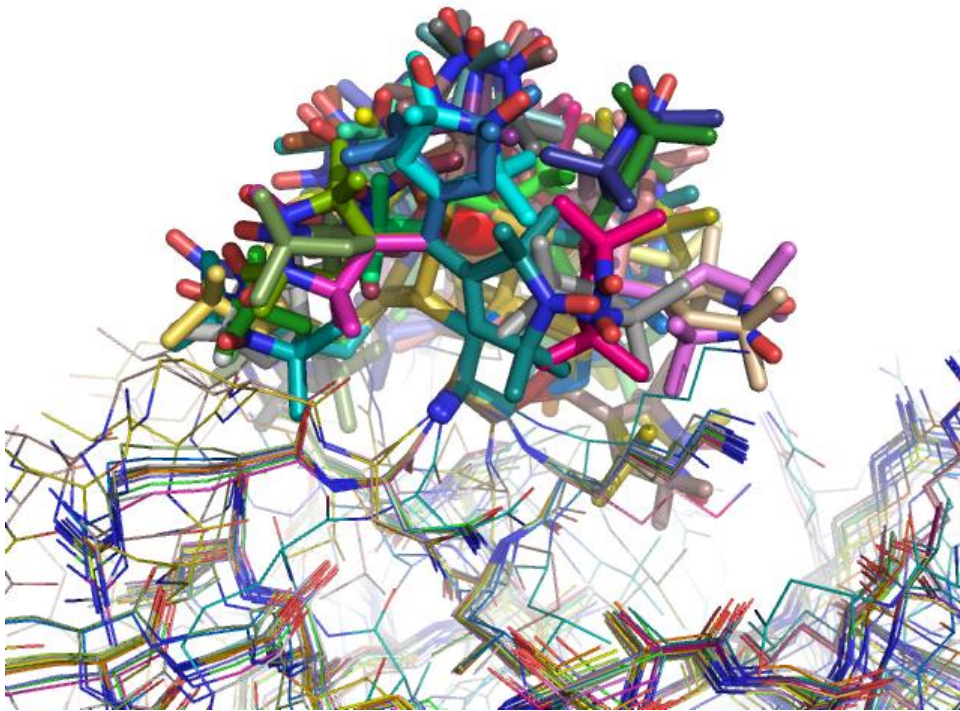


Double Electron-Electron Resonance gives distance information



Borbat, P. P.; McHaourab, H. S.; Freed, J. H.,
J Am Chem Soc **2002**, 124, (19), 5304-14.

Conformations of the spin label are modeled as a conical distribution



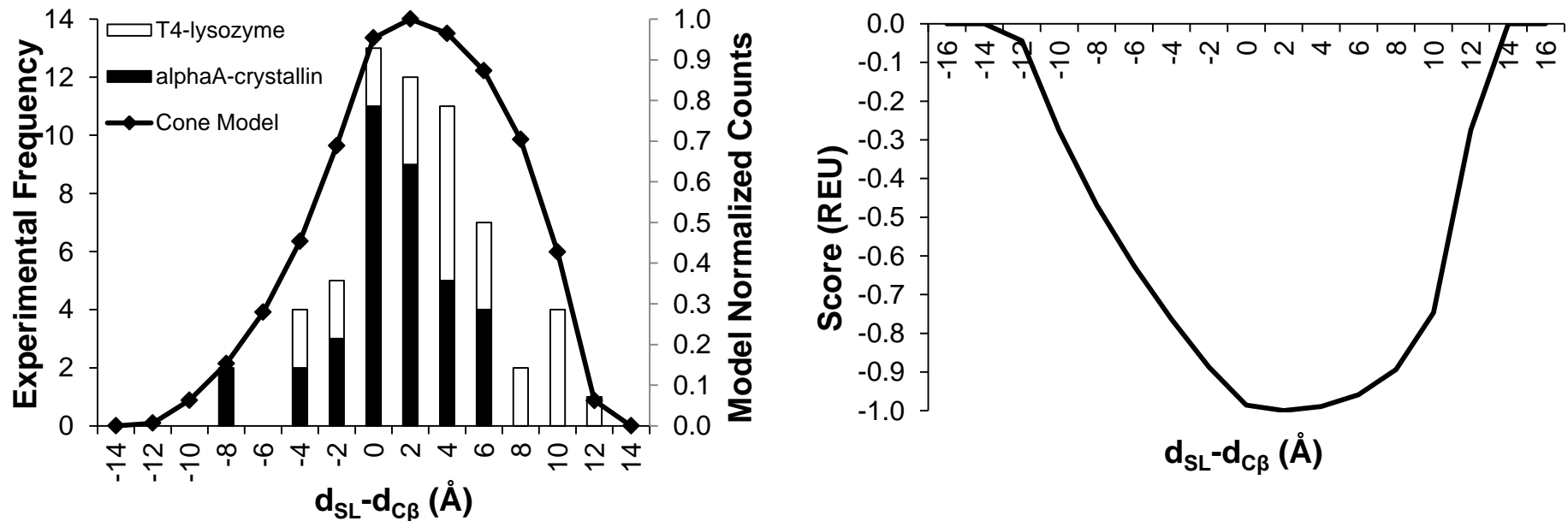
Alexander, N.; Al-Mestarihi, A.; Bortolus, M.; McHaourab, H.; Meiler, J. "De Novo High-Resolution Protein Structure Determination from Sparse Spin-Labeling EPR Data" *Structure* **2008**, *16*, 181-95.

S. J. Hirst, N. Alexander, H. S. McHaourab and J. Meiler; "RosettaEPR: an integrated tool for protein structure determination from sparse EPR data"; *J Struct Biol*; **2011**; Vol. 173 (3): p. 506-14.

Statistics of $D_{SL} - D_{C\beta}$ calculated from cone model match experiment

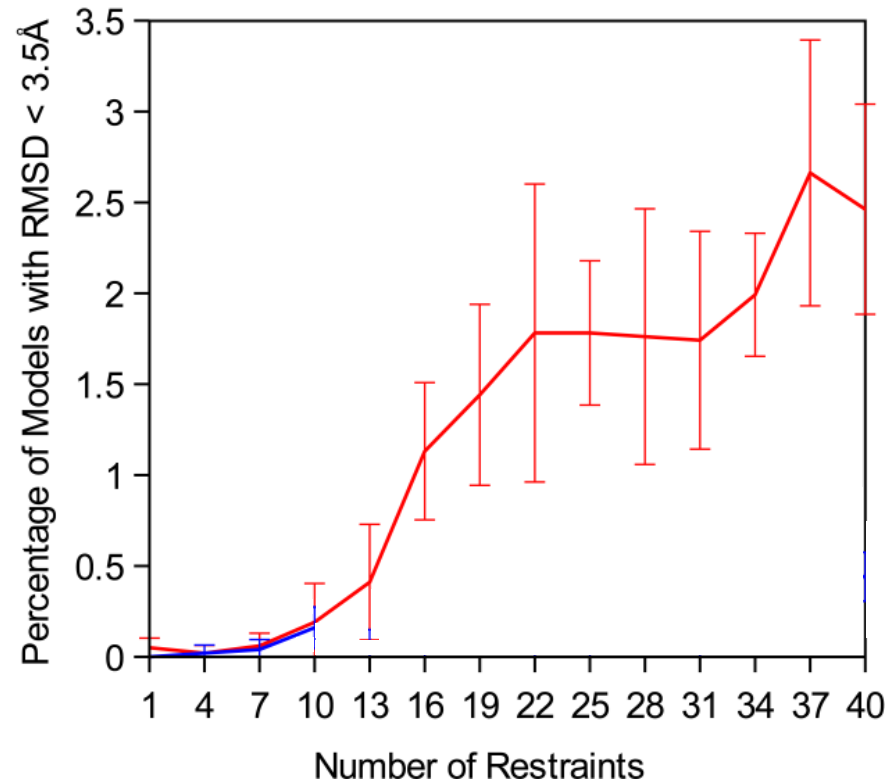
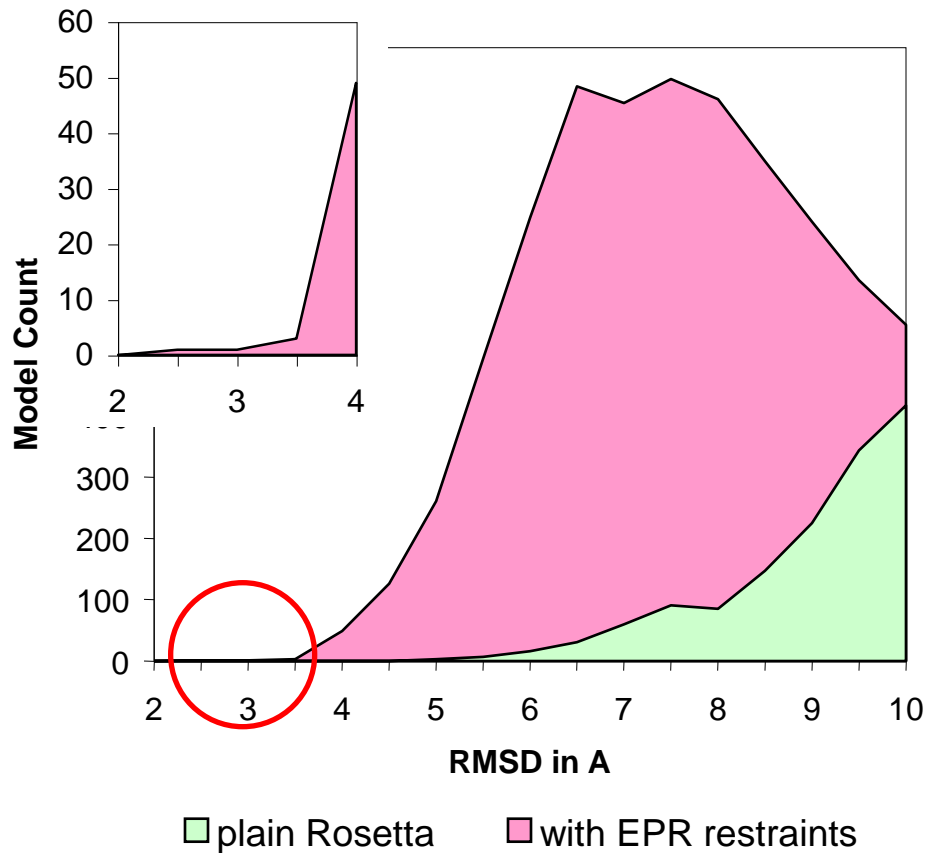


Allows the creation of a scoring function indicating how well a protein model agrees with EPR distance data



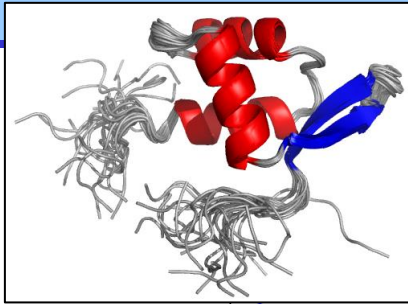
```
AtomPair CB 65 CB 80 SPLINE EPR_DISTANCE 16.0 4.0 0.5
```

Influence of Experimental Data on Sampling and Model Quality

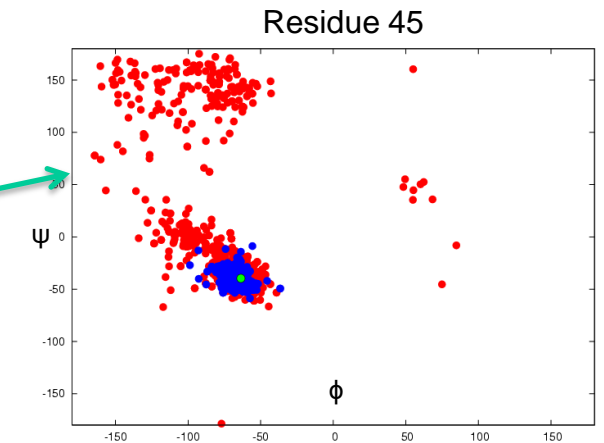
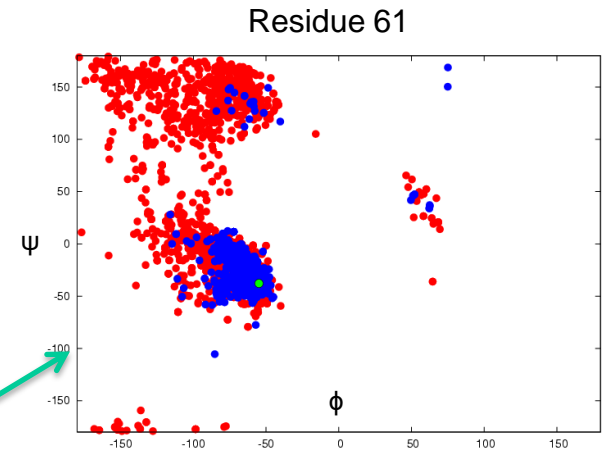
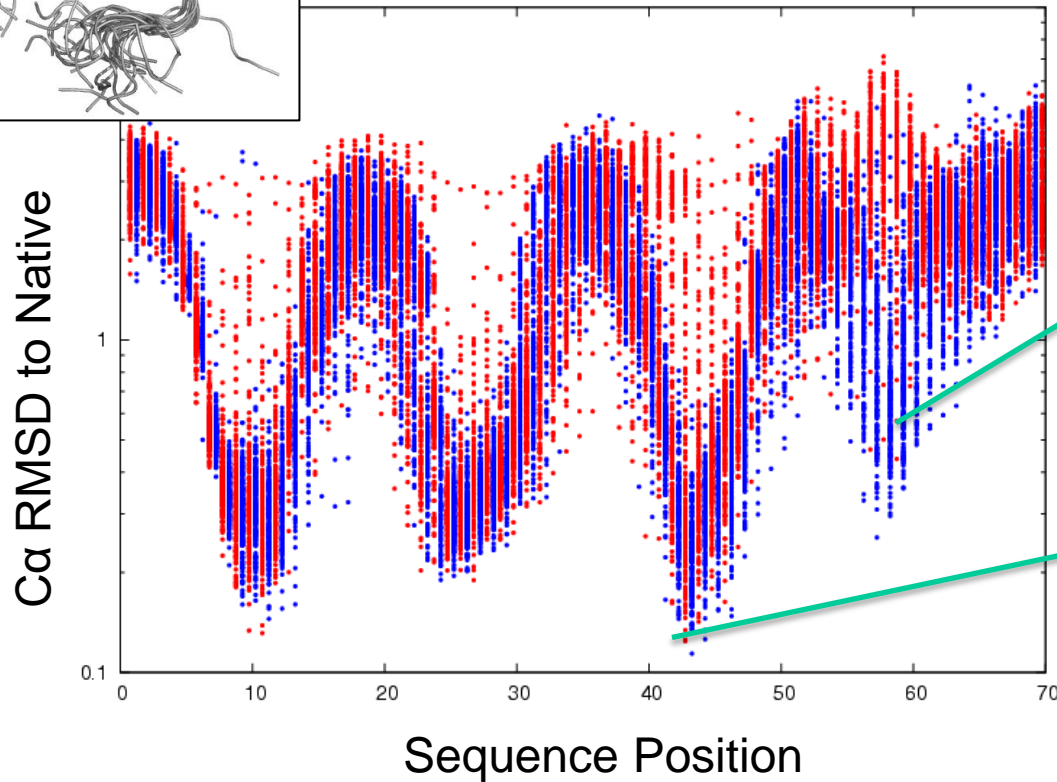


Non-Constraint-Based Experimental Data Incorporation

Chemical Shifts help fragment picking



Fragment Comparison



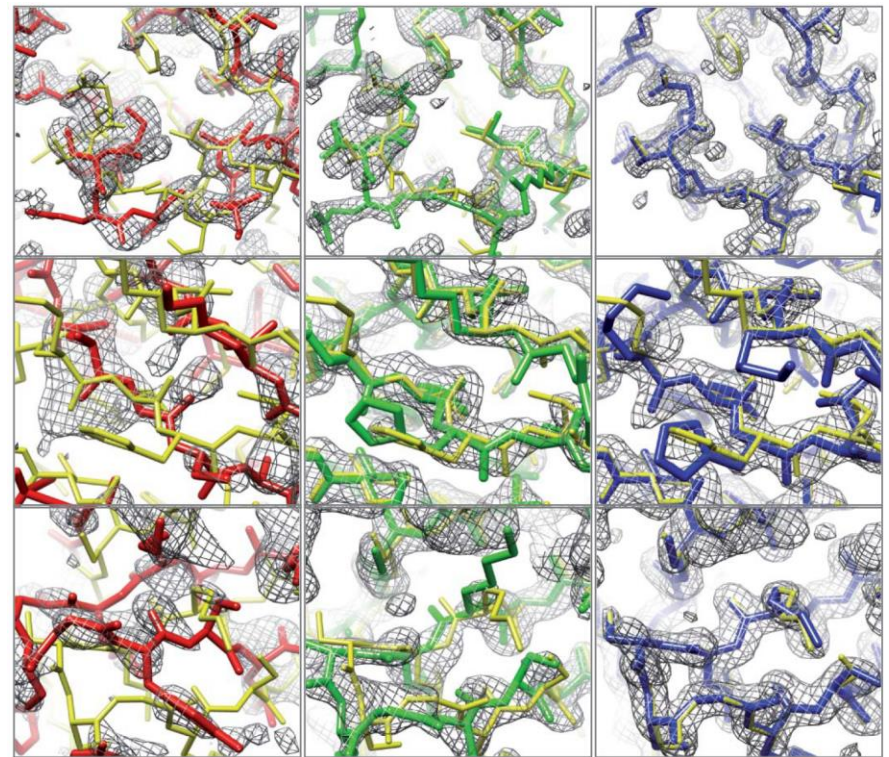
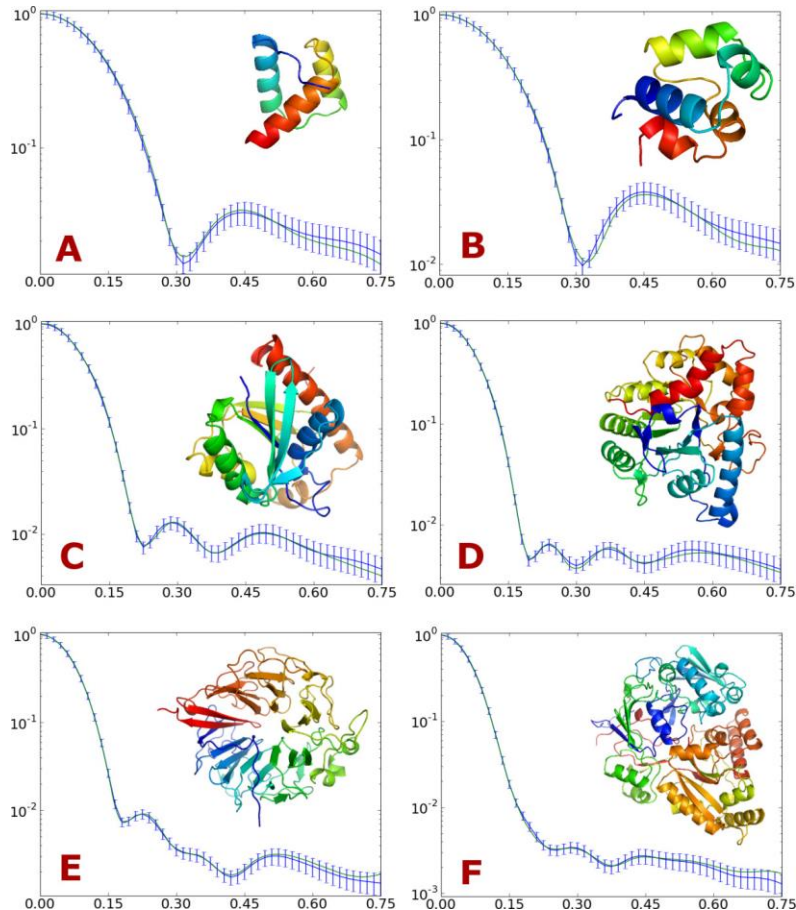
Red: Rosetta Fragments, picked by Sequence profile & Secondary Structure Prediction
Blue: CSRosetta Fragments, picked by CS Comparison & Sequence Matching

SAXS and EM restraints in Rosetta



(Stovgaard et al., BMC Bioinformatics, 2010)

DiMaio F. (2011) Nature



Initial
molecular
replacement

Improve Phases

Model rebuilding
and energy
optimization



References and additional info

- Rosetta Documentation

https://www.rosettacommons.org/docs/latest/rosetta_basics/file_types/constraint-file

- *Kahraman et al. PLOS One, 2013, 8(9) e73411*
- [rosetta/demos/protocol_capture/xl_driven_protein_docking](#)
- Stovgaard et al., BMC Bioinformatics, 2010 – SAXS
- DiMaio F. (2011) Nature - EM
- https://faculty.washington.edu/dimaio/files/rosetta_density_tutorial_aug18.pdf