

Molecular Dynamics Flexible Fitting

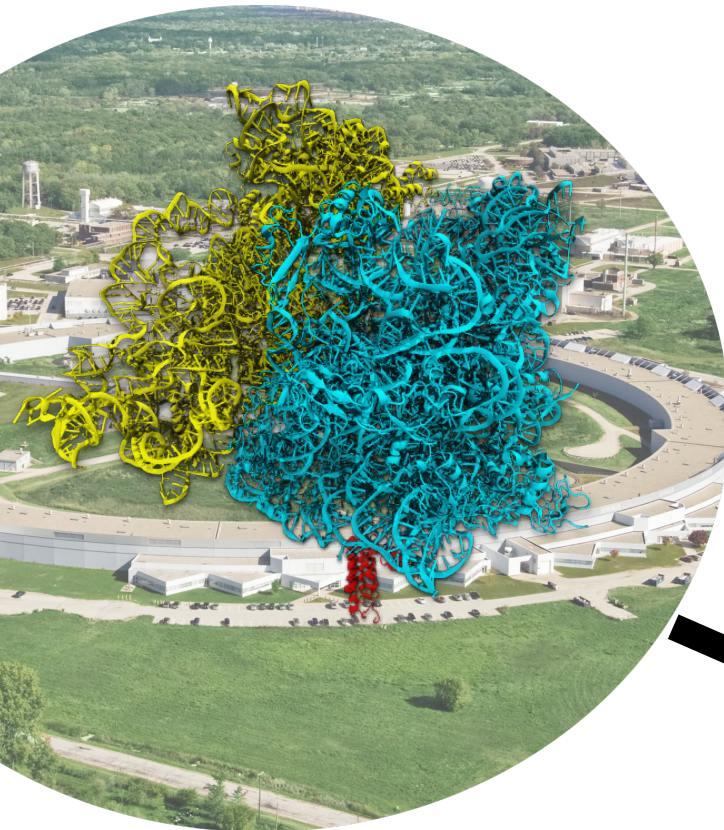
Ryan McGreevy
Research Programmer

University of Illinois at Urbana-Champaign
NIH Resource for Macromolecular Modeling and
Bioinformatics

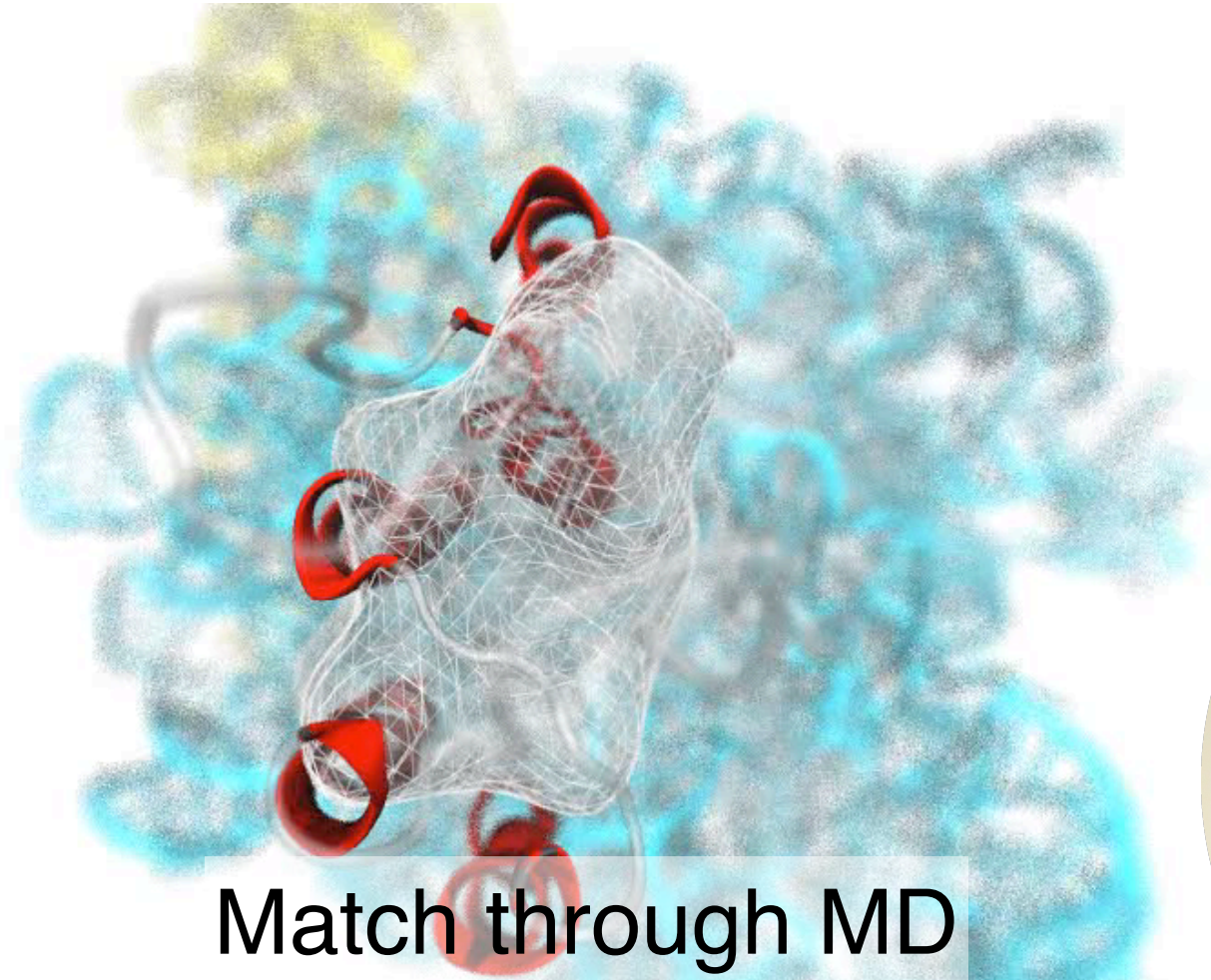
Molecular Dynamics Flexible Fitting

(Ribosome-bound YidC)

crystallographic
structure

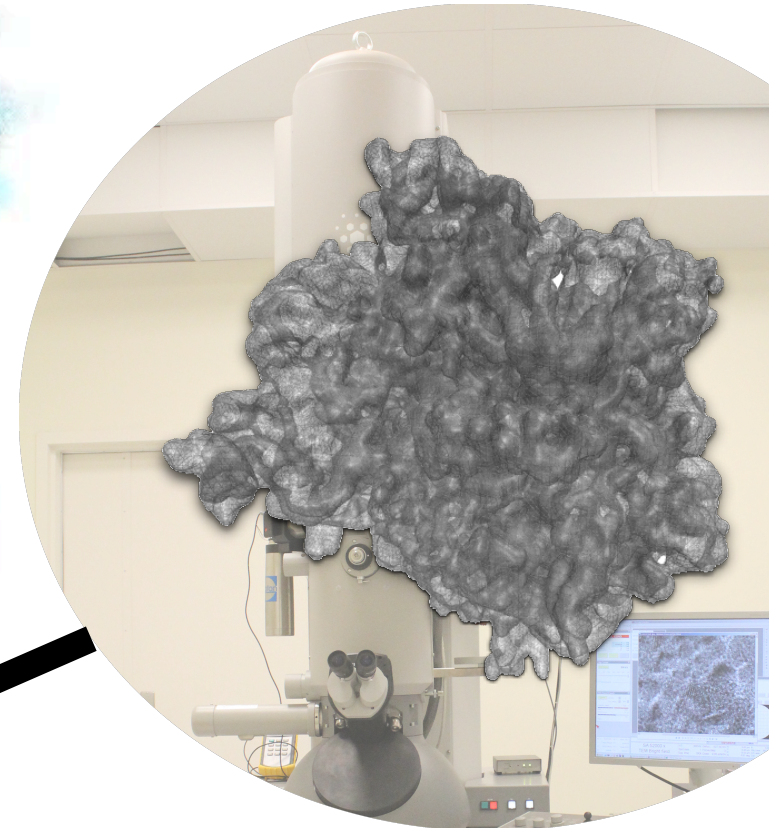


APS
Synchrotron

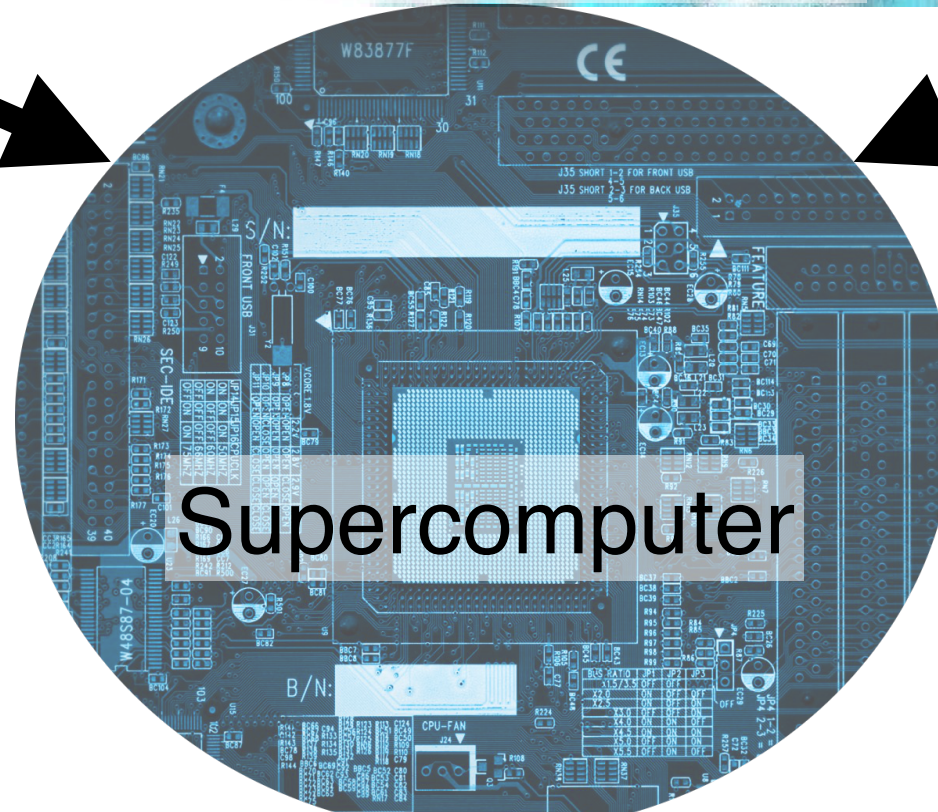


Match through MD

cryo-EM density
map



Electron
Microscope



Supercomputer

Molecular Dynamics Flexible Fitting - Theory

Two terms are added to the MD potential

$$U_{total} = U_{MD} + U_{EM} + U_{SS}$$

An external potential derived from the EM map is defined on a grid as

$$U_{EM}(\mathbf{R}) = \sum_j w_j V_{EM}(\mathbf{r}_j)$$

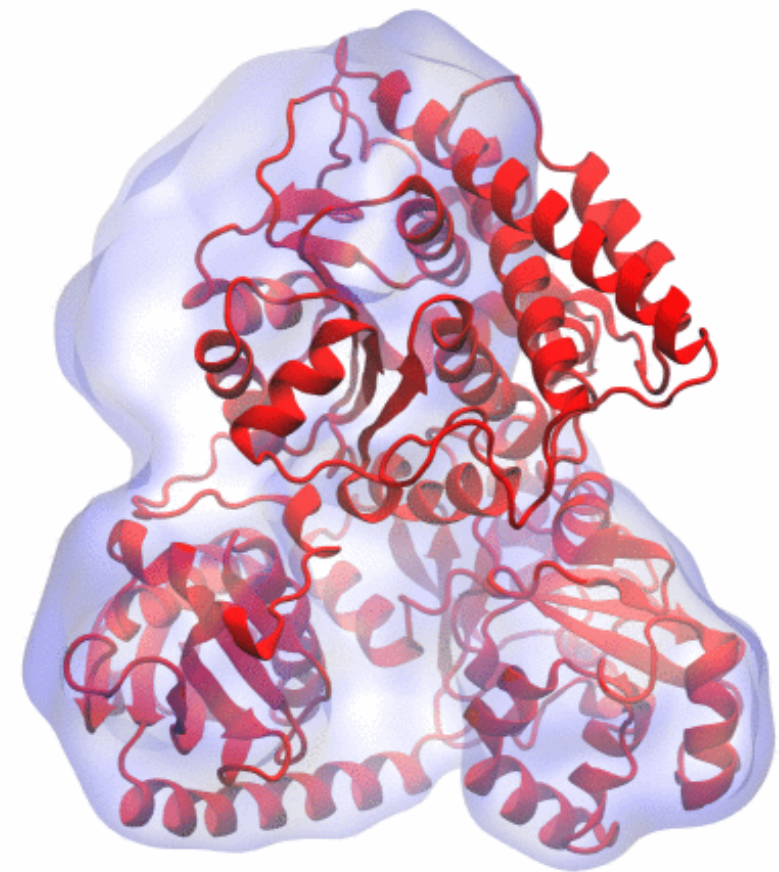
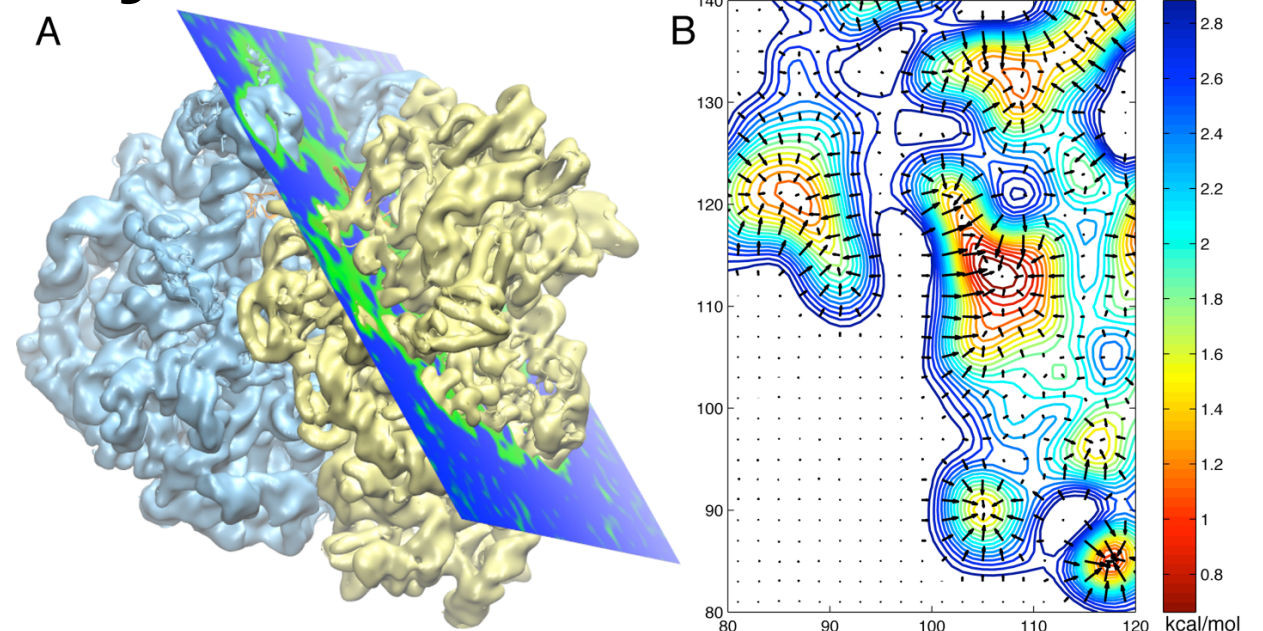
$$V_{EM}(\mathbf{r}) = \begin{cases} \xi \left(1 - \frac{\Phi(\mathbf{r}) - \Phi_{thr}}{\Phi_{max} - \Phi_{thr}} \right) & \text{if } \Phi(\mathbf{r}) \geq \Phi_{thr}, \\ \xi & \text{if } \Phi(\mathbf{r}) < \Phi_{thr}. \end{cases}$$

A mass-weighted force is then applied to each atom

$$\mathbf{f}_i^{EM} = -\nabla U_{EM}(\mathbf{R}) = -w_i \partial V_{EM}(\mathbf{r}_i) / \partial r_i$$

[1] Trabuco et al. *Structure* (2008) 16:673-683.

[2] Trabuco et al. *Methods* (2009) 49:174-180.



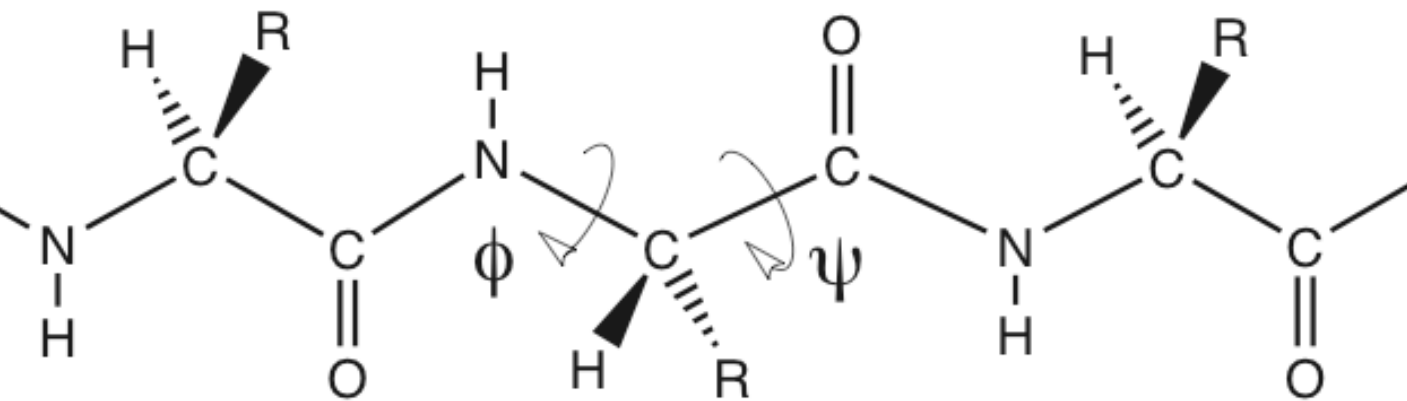
Acetyl – CoA Synthase

Secondary structure restraints

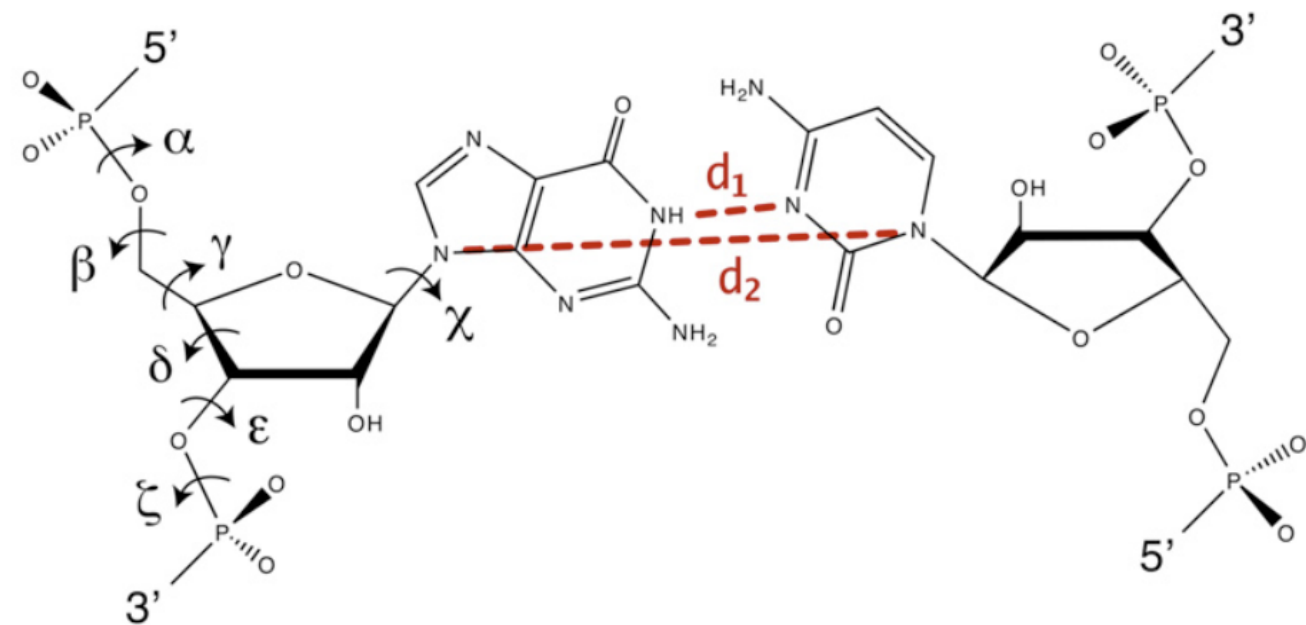
Harmonic restraints are applied to preserve secondary structure of proteins and nucleic acids, avoiding “overfitting.”

$$U_{SS} = \sum_{restraints} k_{\mu} (\mu - \mu_0)^2$$

For proteins, ϕ and ψ dihedral angles of residues within helices or beta strands are restrained.



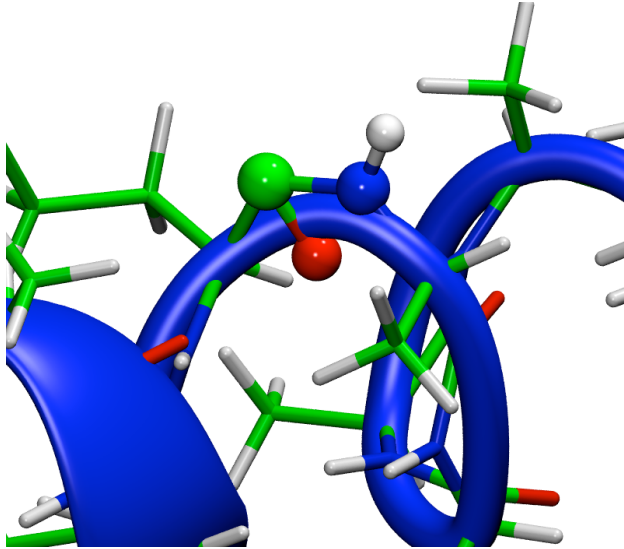
For nucleic acids, distance and dihedral restraints are applied to a selected set of base pairs.



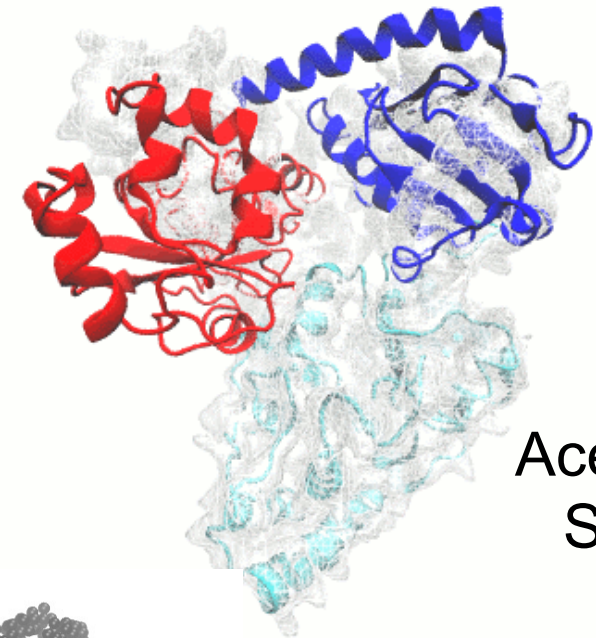
Additional Restraints

Cis-peptide and Chirality

Eduard Schreiner, et al. BMC Bioinformatics, 12, 190, 2011

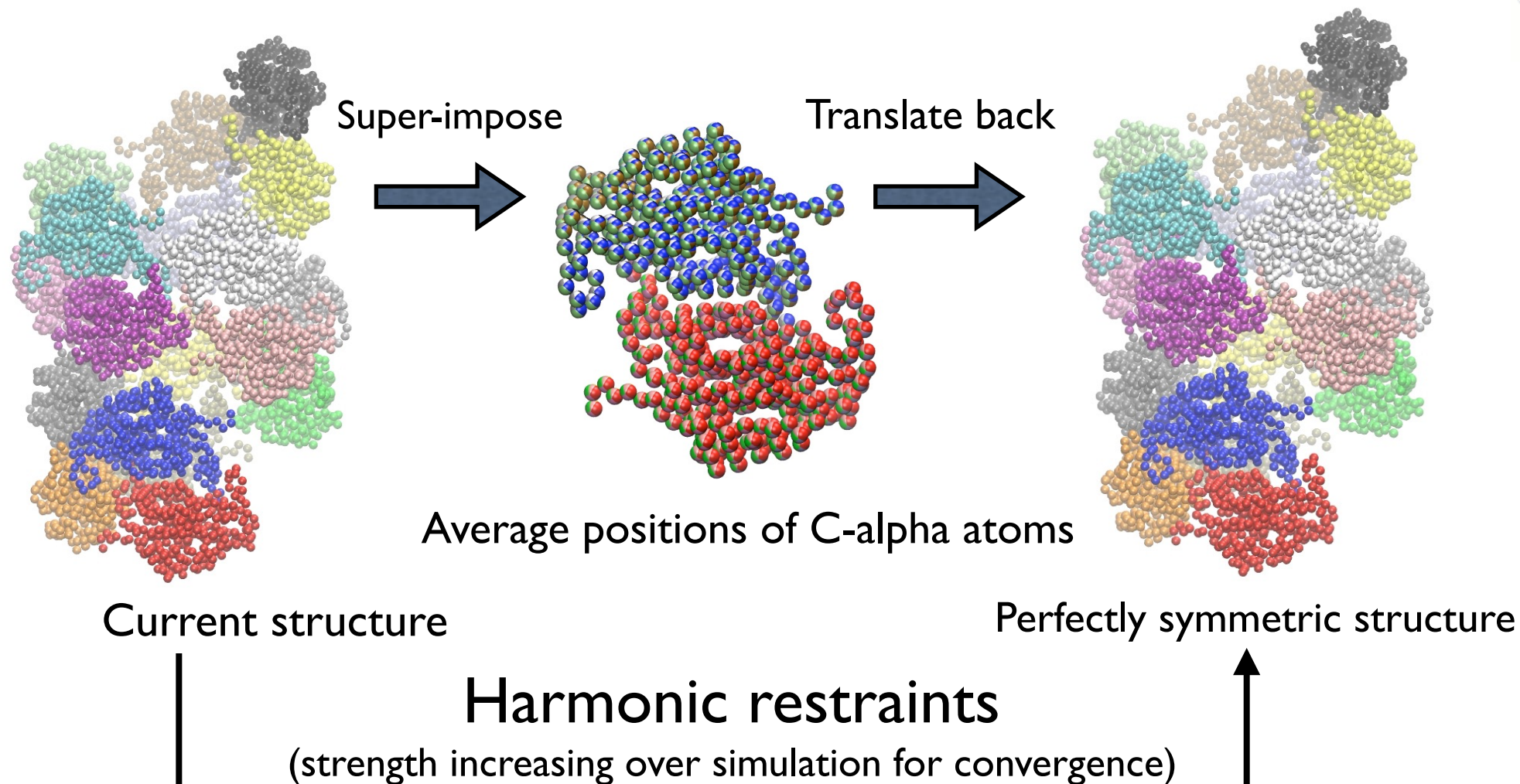


Domain-wise



Acetyl – CoA
Synthase

Symmetry



Kwok-Yan Chan, et al. Structure, 19, 1211-1218, 2011



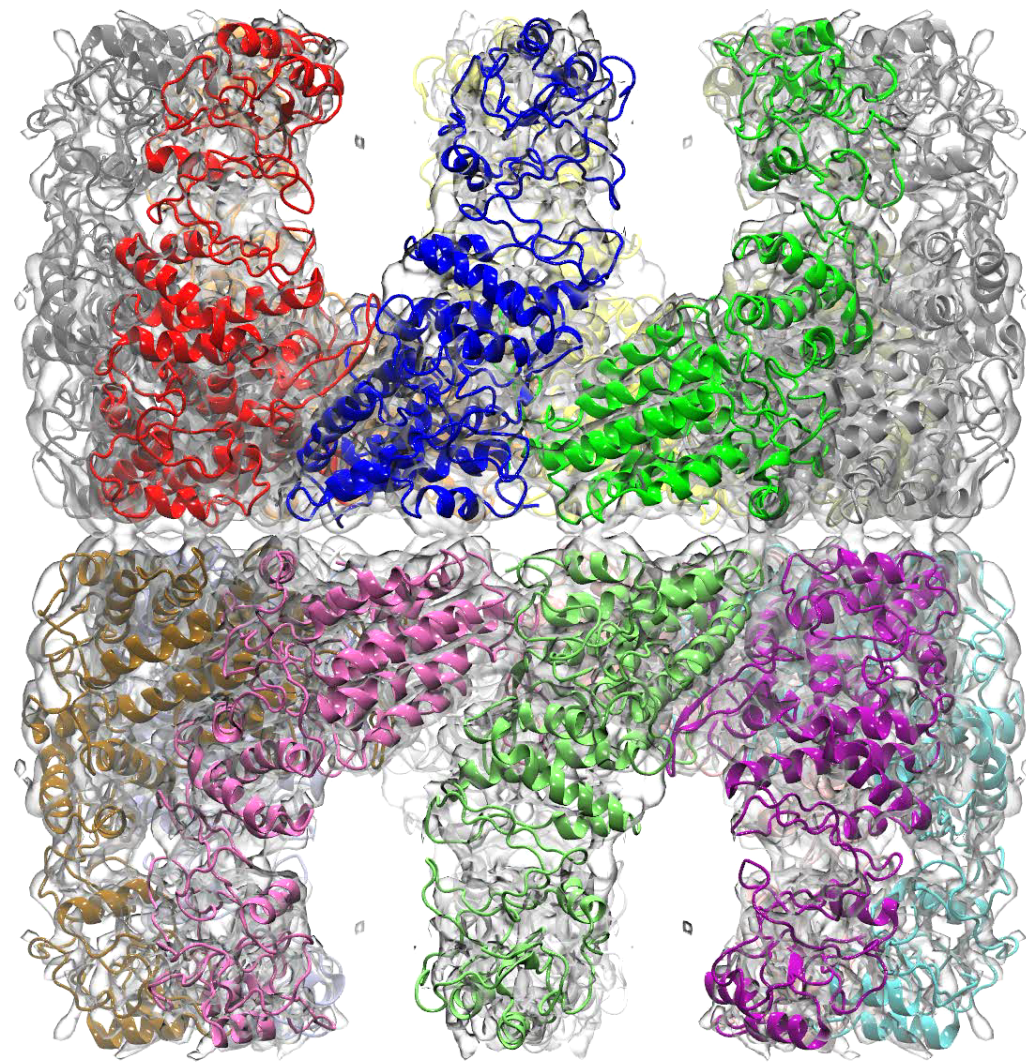
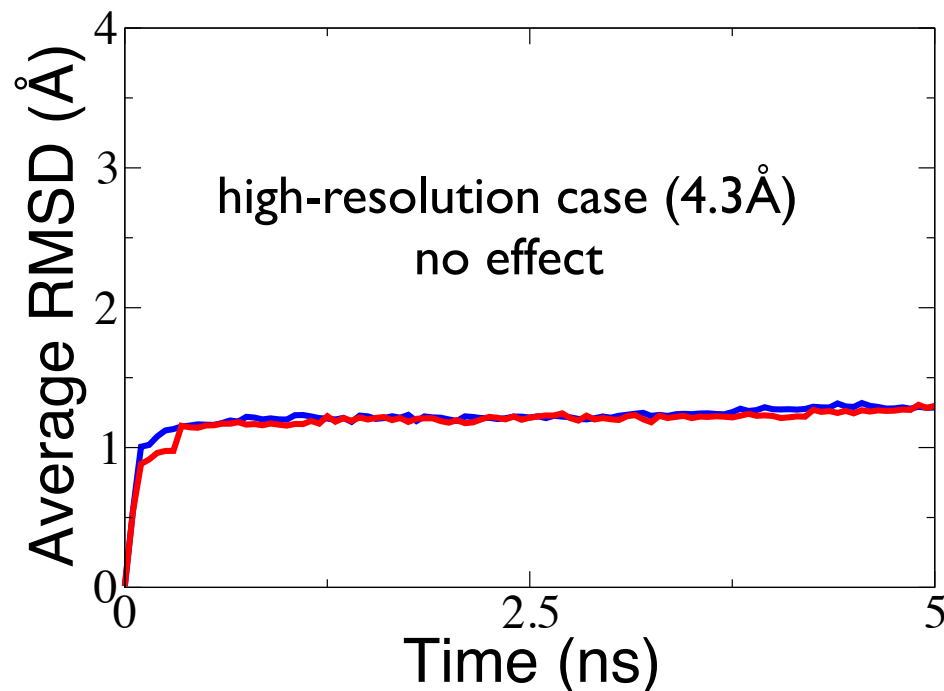
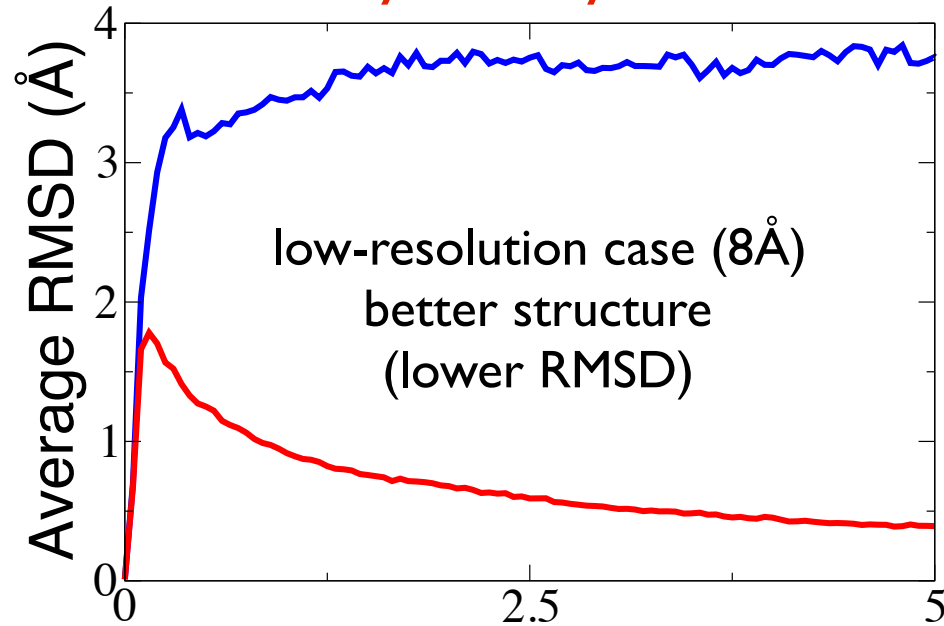
B. pumilus cyanide hydratase

Symmetry restrained MDFF - Test Case 1

Improve quality of fit for low-resolution data

Blue: without symmetry restraints

Red: with symmetry restraints



Archaeal group II chaperonin from *M. maripaludis* (Mm-cpn)
8-fold rotational + 2 fold reflection symmetry
homology model (based on PDB 3LOS) fitted into EM map (EMDB 5140)

Chan et al. *Structure* (2011) 19:1211-1218.

Symmetry restrained MDFF - Test Case 2

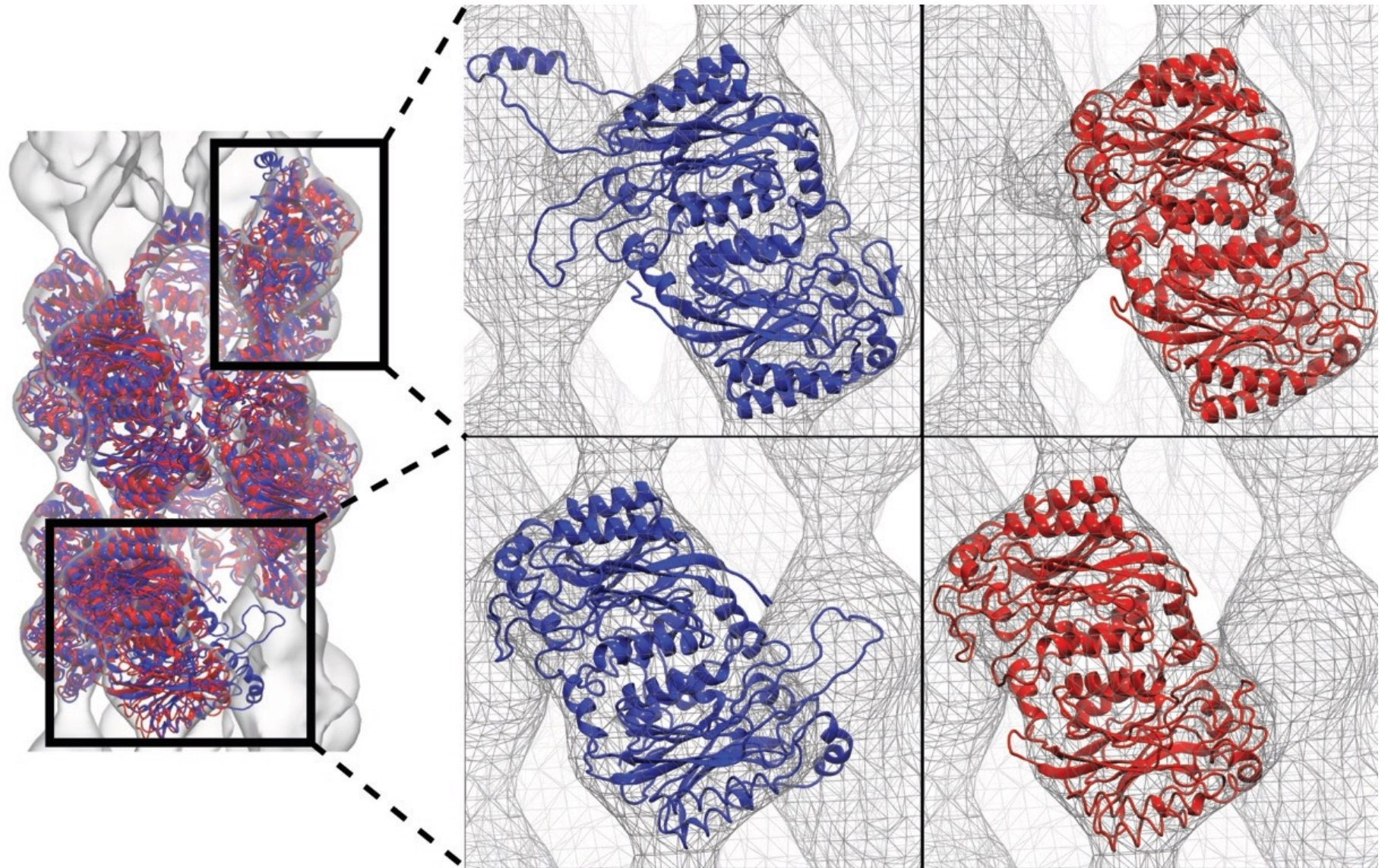
Prevent “edge distortion effect”

Finite-size Simulation
(9 dimers)

helical symmetry

Fitted models of J1 nitrilase
from *R. rhodochrous*

homology model and EM map (EMD 1313)
from collaborator T. Sewell, U. of Cape
Town



Without Symmetry
Restraints

With Symmetry
Restraints

Chan et al. *Structure* (2011) 19:1211-1218.

Simulation Environment

MDFF can be run in different environments:

1. Vacuum

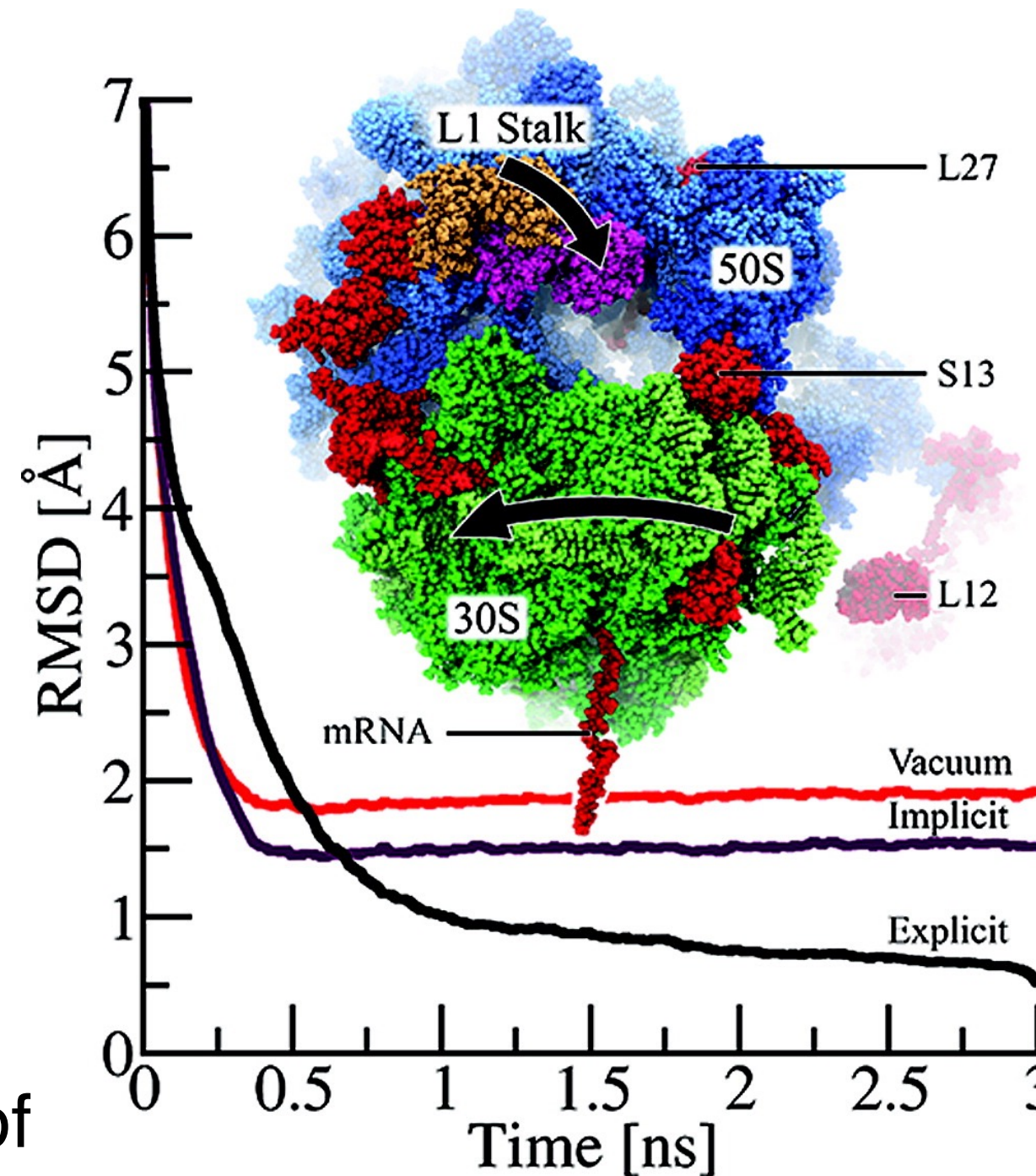
- No water molecules
- Fastest but potentially inaccurate

2. Explicit Solvent

- Explicit atomic detail water molecules
- Computationally slow and introduces effects of viscous drag

3. Implicit Solvent

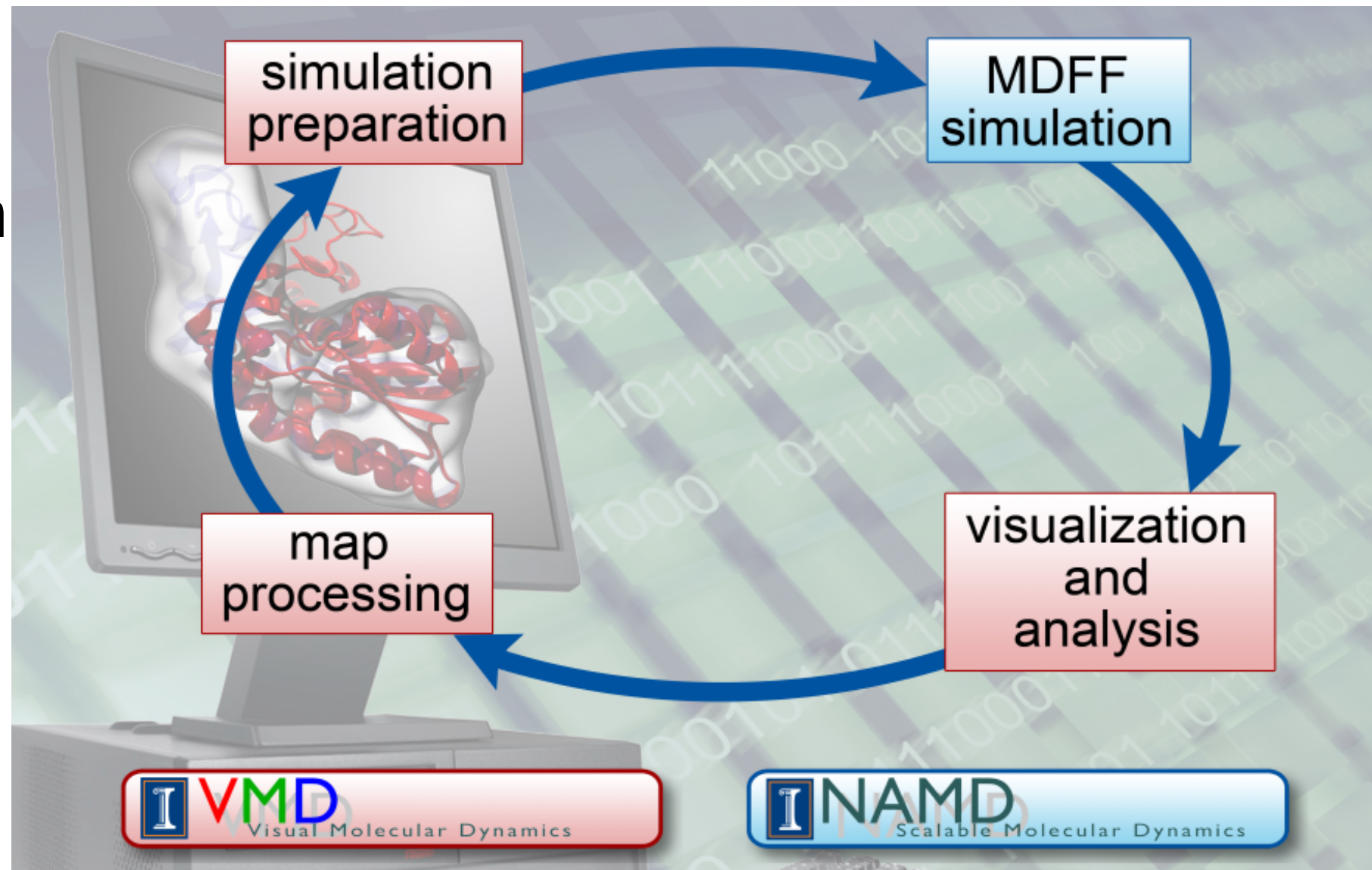
- generalized Born approximation of electrostatics
- Compromise between speed and accuracy



Tanner, et al. *Journal of Chemical Theory and Computation* 7(11) 3635–3642, 2011.

MDFF Software Suite

- NAMD and VMD used together to run MDFF
- Every NAMD and VMD feature is available in MDFF



Fitting time is dependent on:

- system size
- map and structure quality
- Generally need ~ 1 ns or less (much shorter than MD)

Input: MDFF only requires a PDB, PSF, and density map

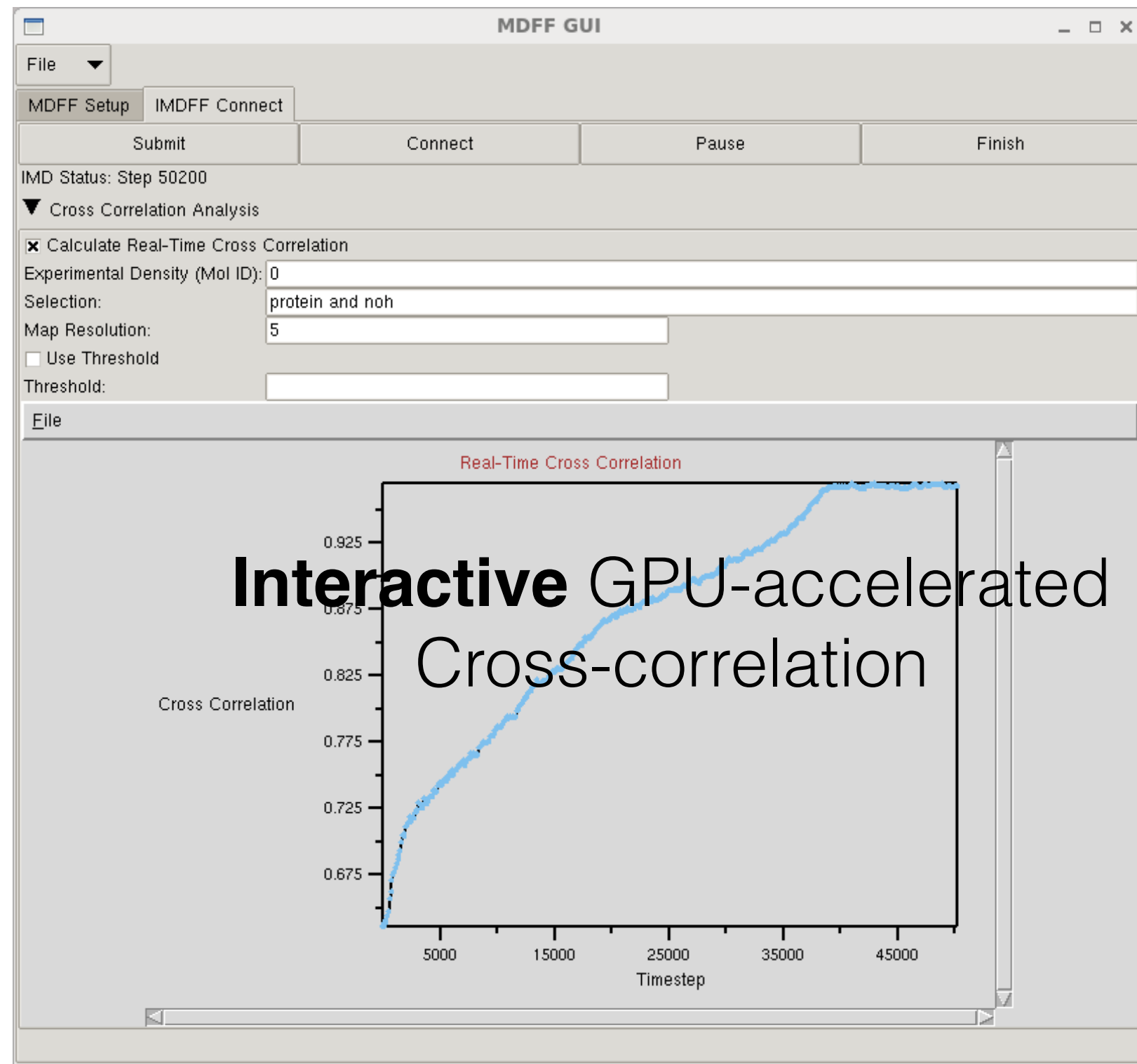
Output: produces simulation trajectory from which an ensemble of structures can

<http://www.ks.uiuc.edu/Research/mdff/> be extracted

MDFF Software Suite

- system sizes up to 100 million atoms (viruses, chromatophore)
- maps from 3 to 15 Å
- runs on laptops to petascale computing resources (Blue Waters, Titan)

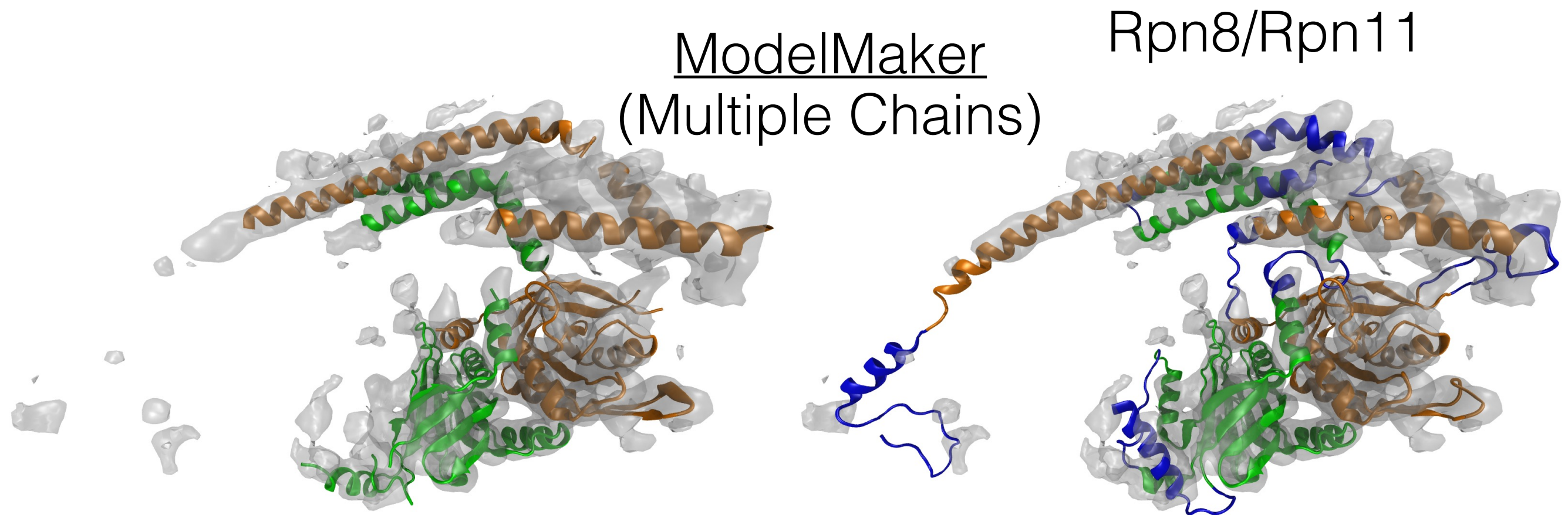
New MDFF GUI (VMD 1.9.2) makes setting up, running, and analyzing fitting simulations even easier



<http://www.ks.uiuc.edu/Research/mdff/>

Obtaining Initial Structures

1. X-ray crystallography or NMR structures
2. Refine structures from low-res X-ray data with xMDFF
Ryan McGreevy*, Abhishek Singharoy*, et al. Acta Crystallographica D70, 2344-2355, 2014
3. Homology or ab initio modeling with Modeller, Rosetta, MUFOLD
(Ci-VSP, YidC, Holotranslocon)

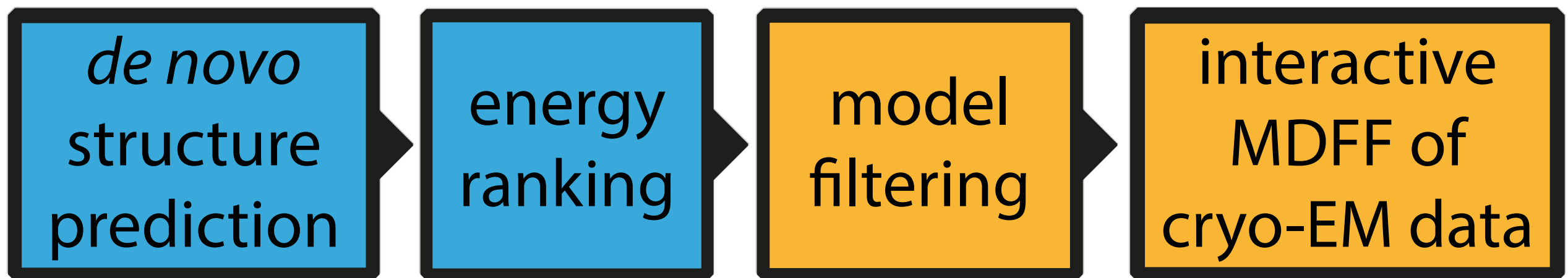


Rosetta structure prediction to fill **missing pieces** and MDFF to filter, refine and validate candidate structures

ModelMaker Interactive Modeling

Combining structure prediction with the user's expertise to interpret densities

incomplete structural model deposited in the PDB



complete structural model that fits cryo-EM data

Rosetta

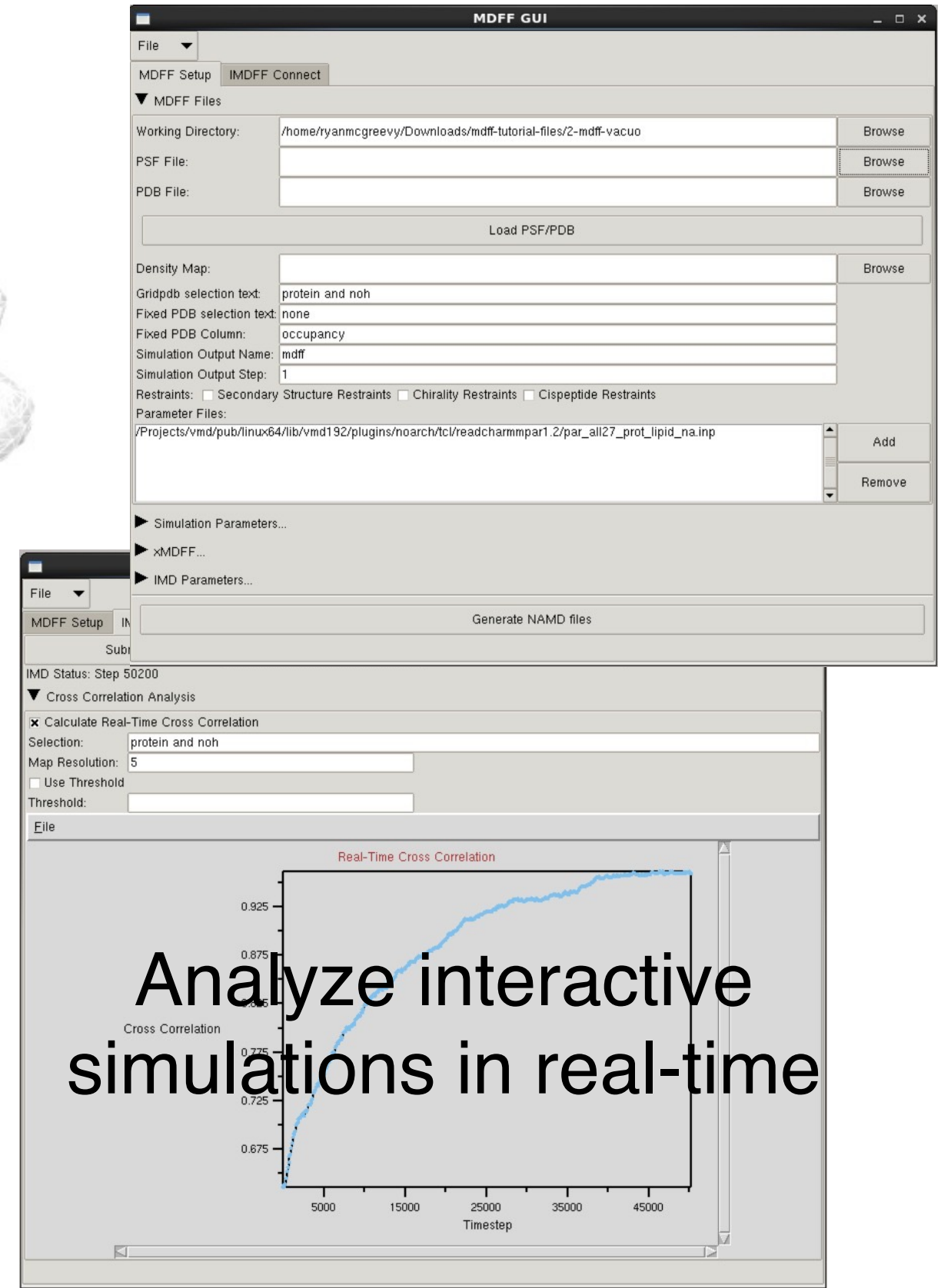
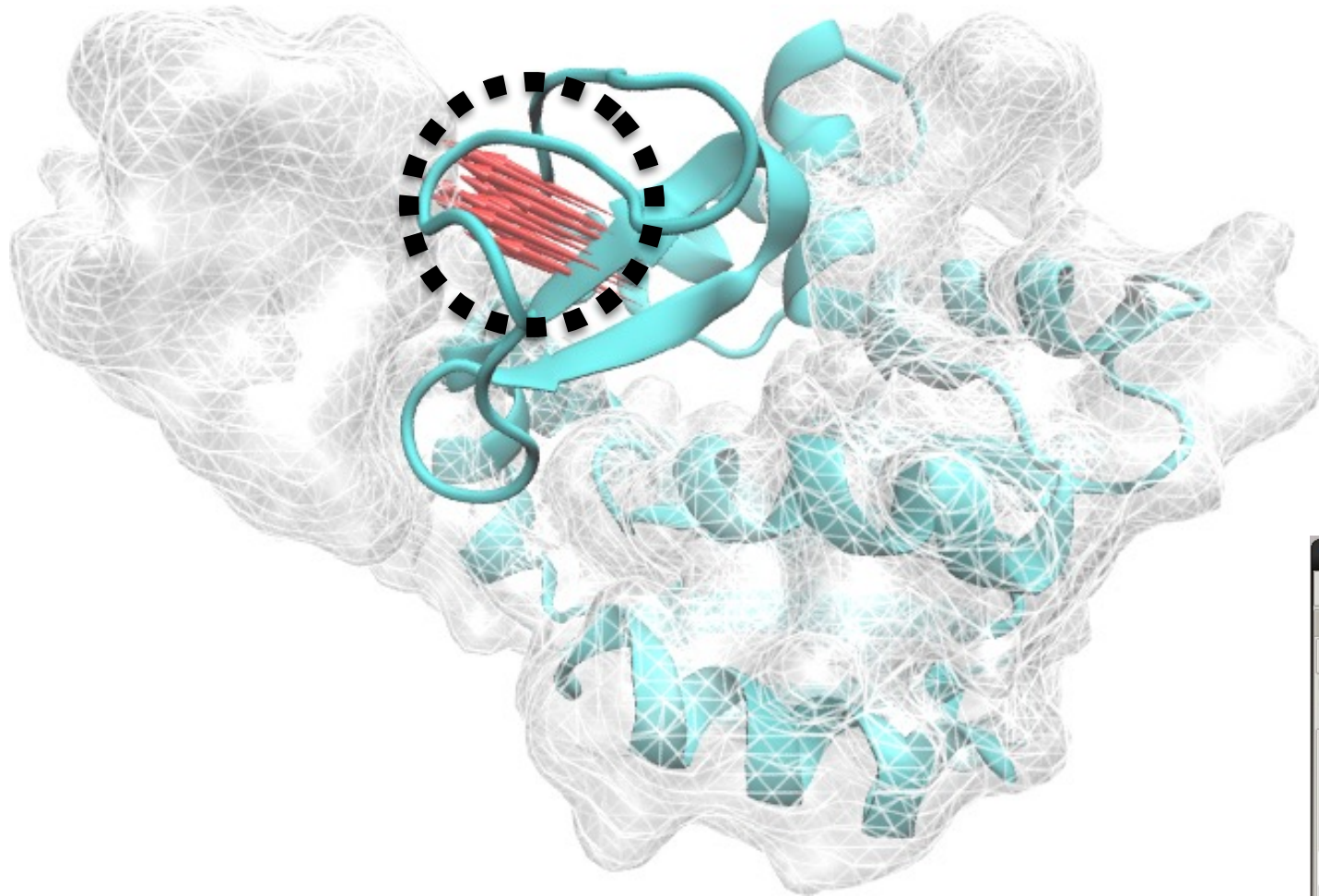
Leaver-Fay *et al.* Methods Enzymol. 2011
Porter *et al.* PLoS One 2015

VMD/NAMD

Humphrey *et al.* J. Mol. Graph. 1996
Philips *et al.* J. Comput. Chem. 2005

Interactive Modeling with MDFF GUI

- Apply forces to manually manipulate structure into the density
- Useful for difficult to fit structures with large conformational changes



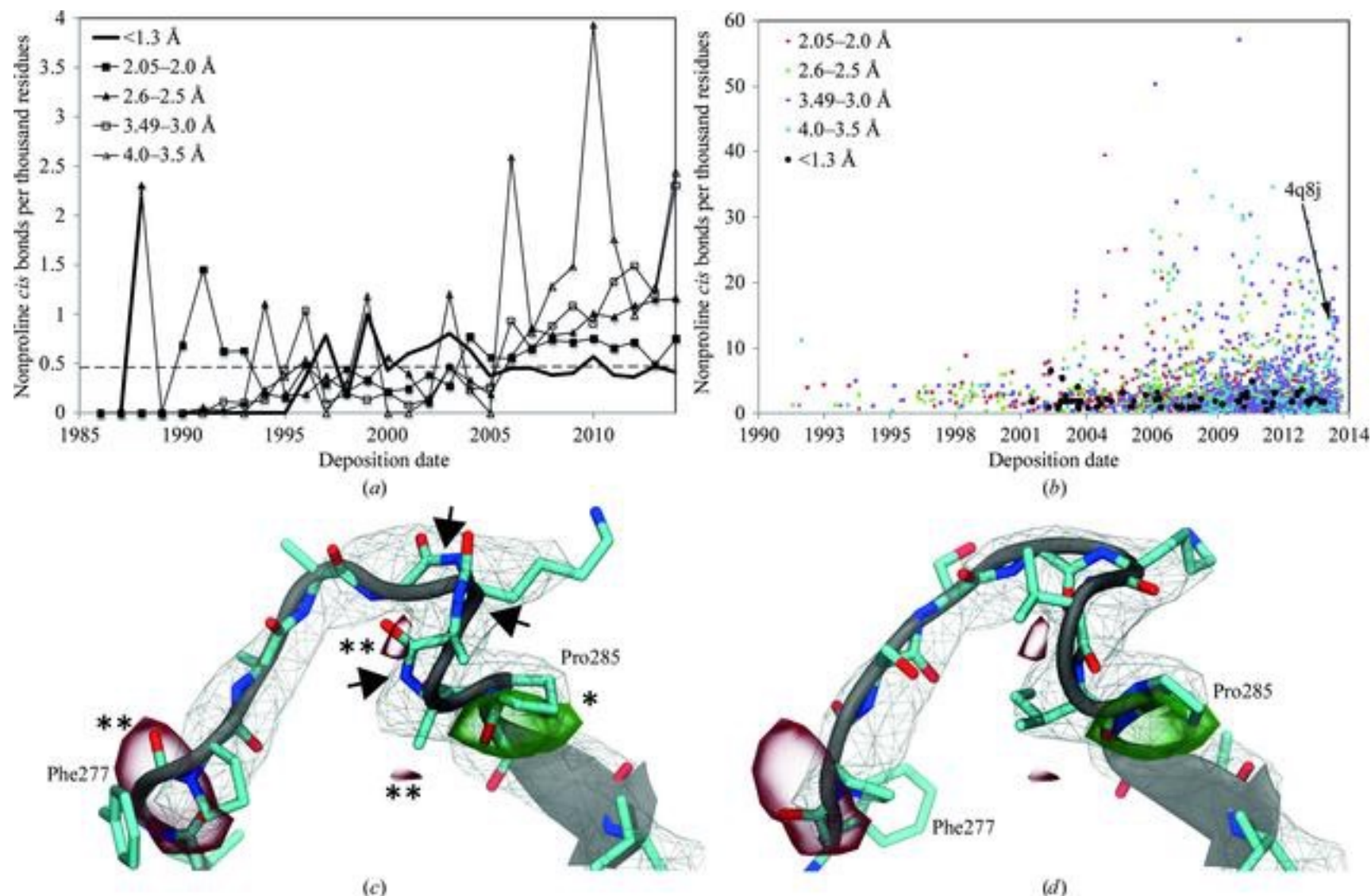
New MDFF GUI in VMD 1.9.2
Set up and run interactive (or
traditional) MDFF/xMDFF
simulations

Analyze interactive
simulations in real-time

Importance of Checking Initial Structure

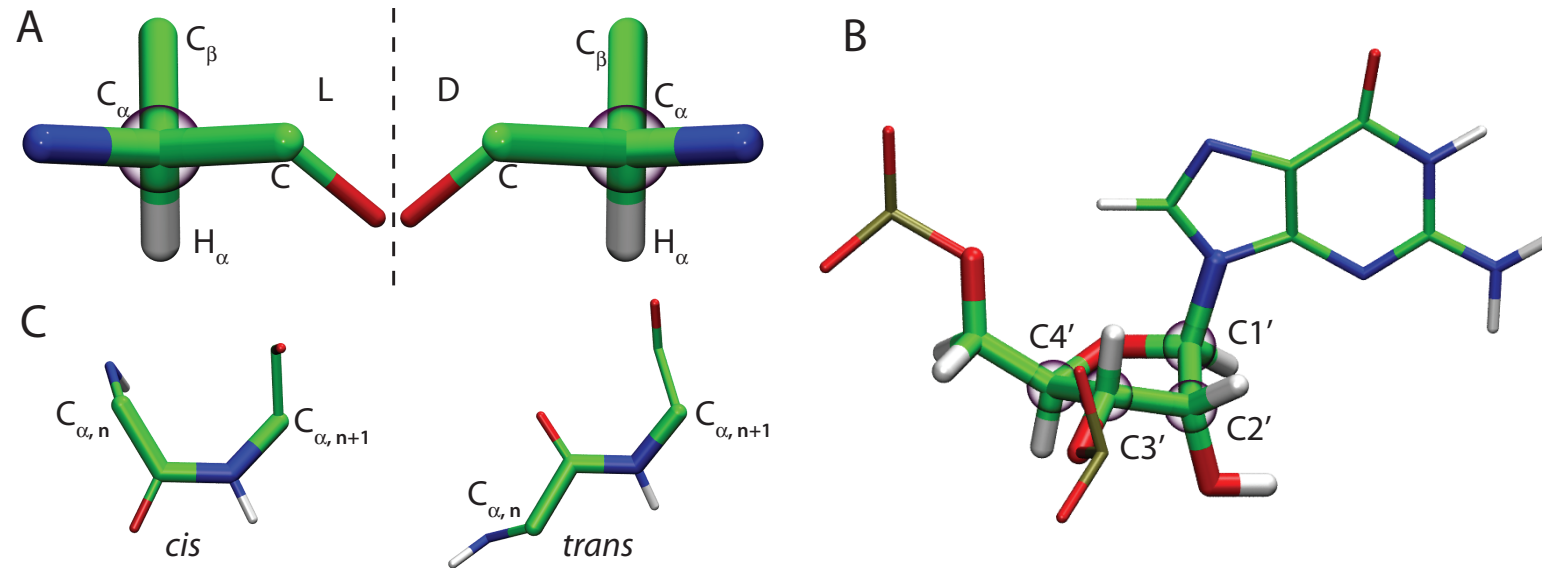
<0.05% non-proline bonds found in the cis conformation natively, however:

- The frequency of non-proline cis-peptide bond errors has been increasing for low-resolution
- These errors can hide issues in other parts of the structure



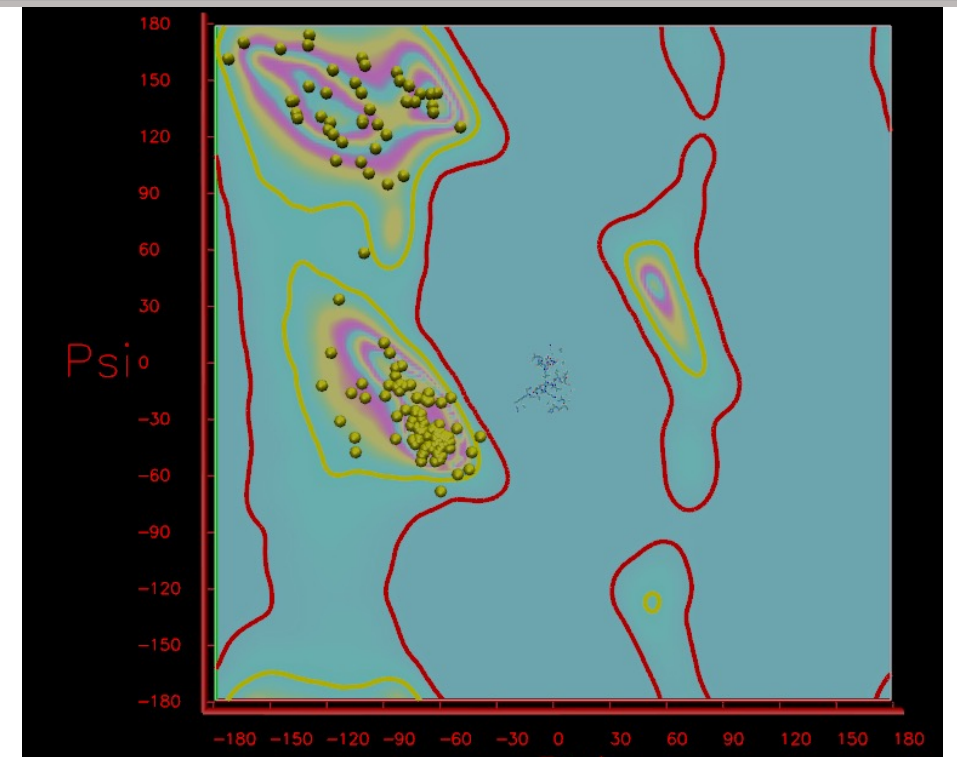
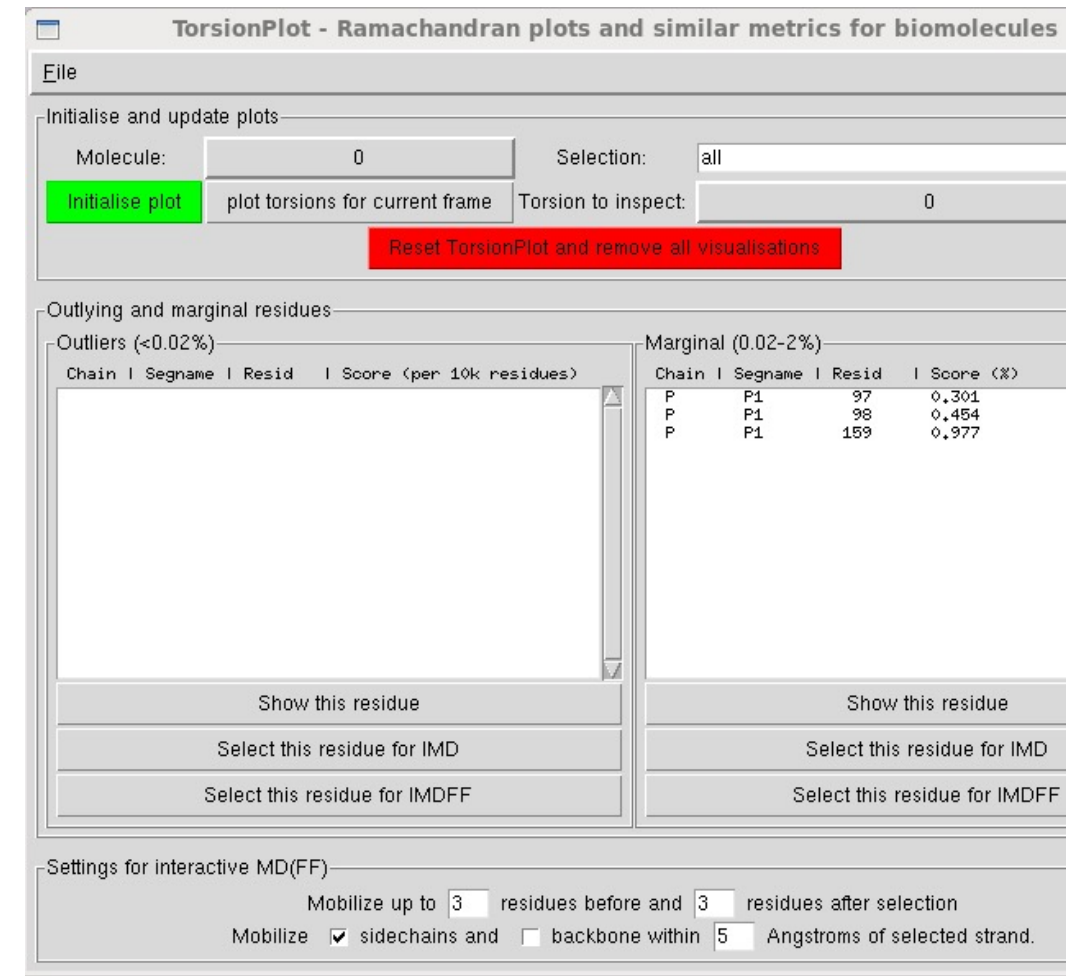
Tristan Croll. Acta Crystallographica D71, 706-709, 2015.

Structure Checking Plugins in VMD



Eduard Schreiner, et al. BMC Bioinformatics, 12, 190, 2011

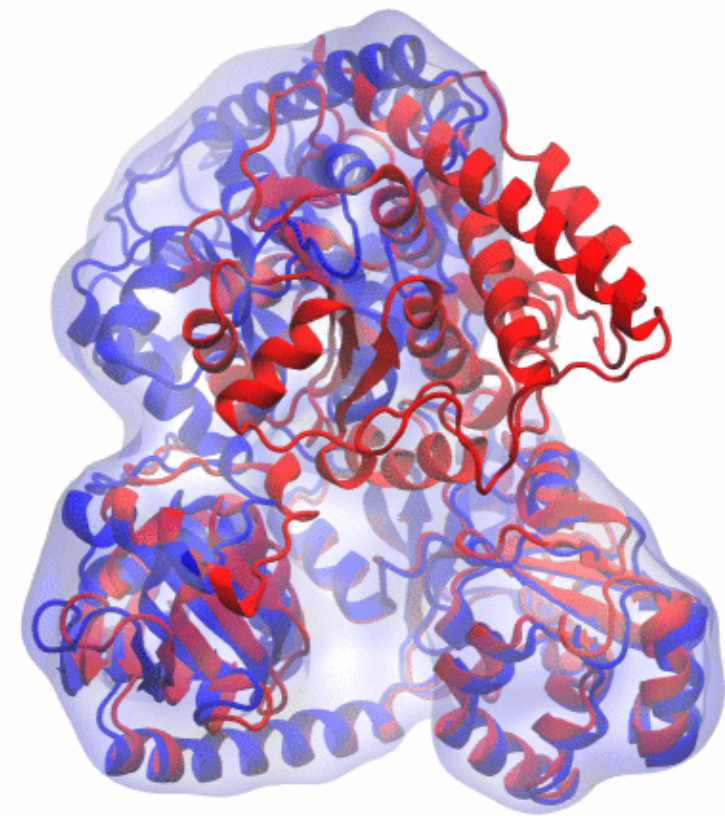
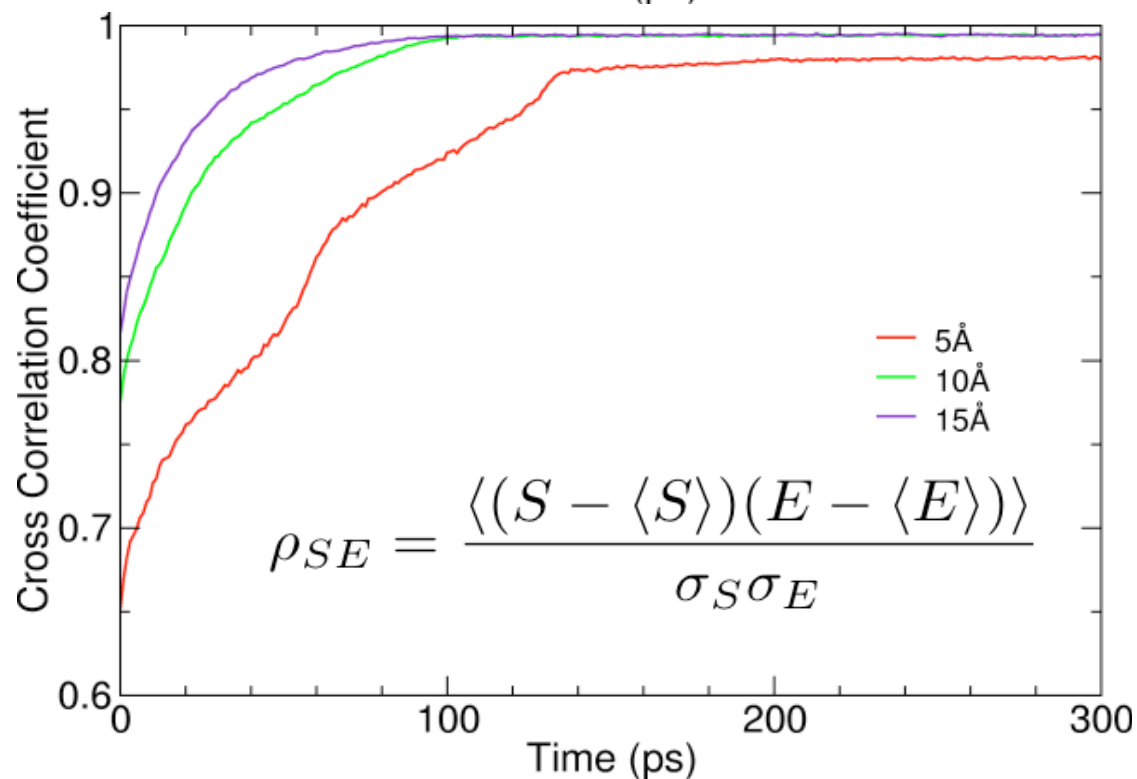
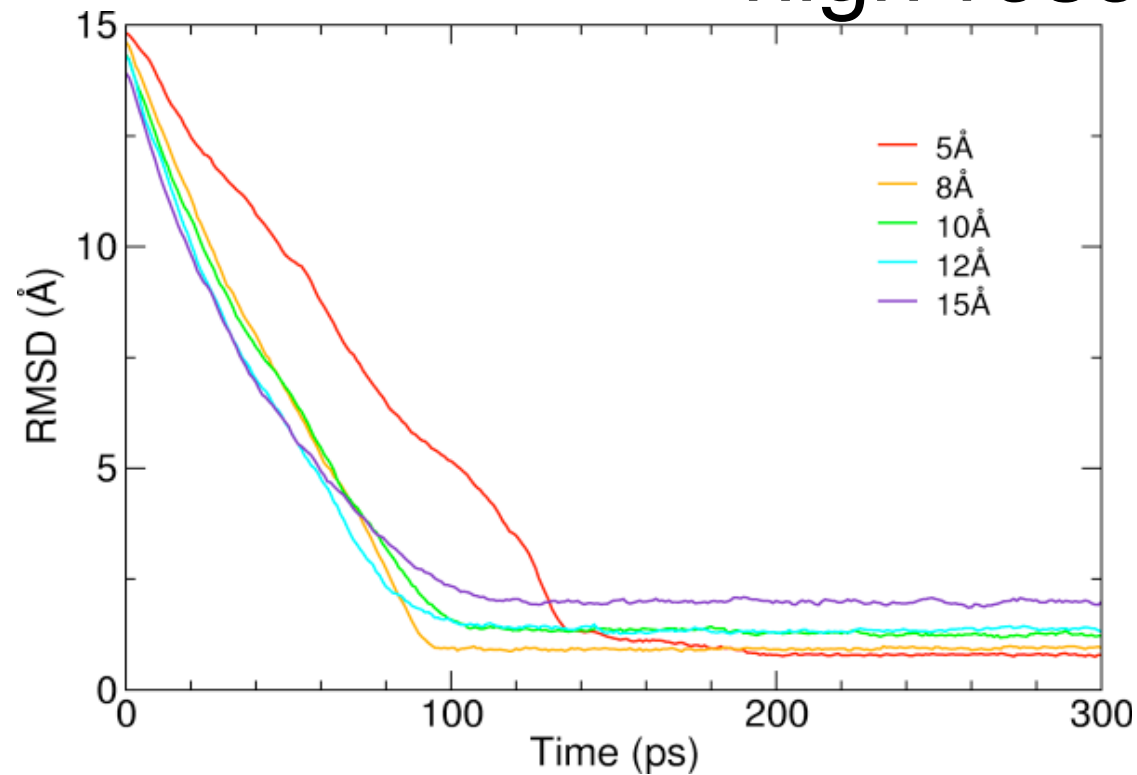
- Wrong chirality, cis-peptide bonds, and torsion angle outliers may arise during modeling
- VMD provides tools to check, visualize, and correct these errors
- These tools, together with MD force fields, produce models with good structural geometry



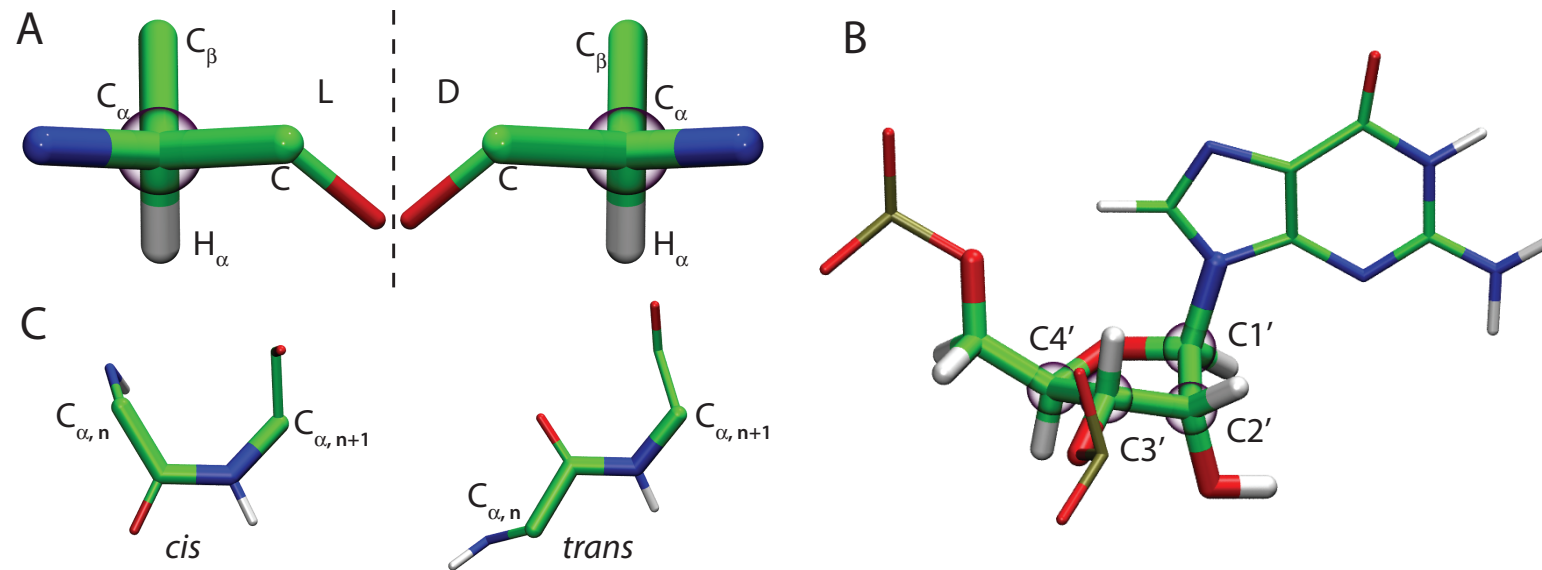
TorsionPlot Plugin new in VMD 1.9.3

Analyzing MDFF Model Quality 0: **Known Structures**

MDFF has been validated against a wide-ranging set of known high-resolution structures



Analyzing MDFF Model Quality 1: Structure Checking



xMDFF refined structures

PDB ID	Molprobit	
	initial (published)	final
1AV1	3.72	1.94
1YE1	2.68	1.89
1JL4	3.24	1.47
1AOS	3.40	2.45
1XDV	2.87	2.01
1YI5	3.08	1.73

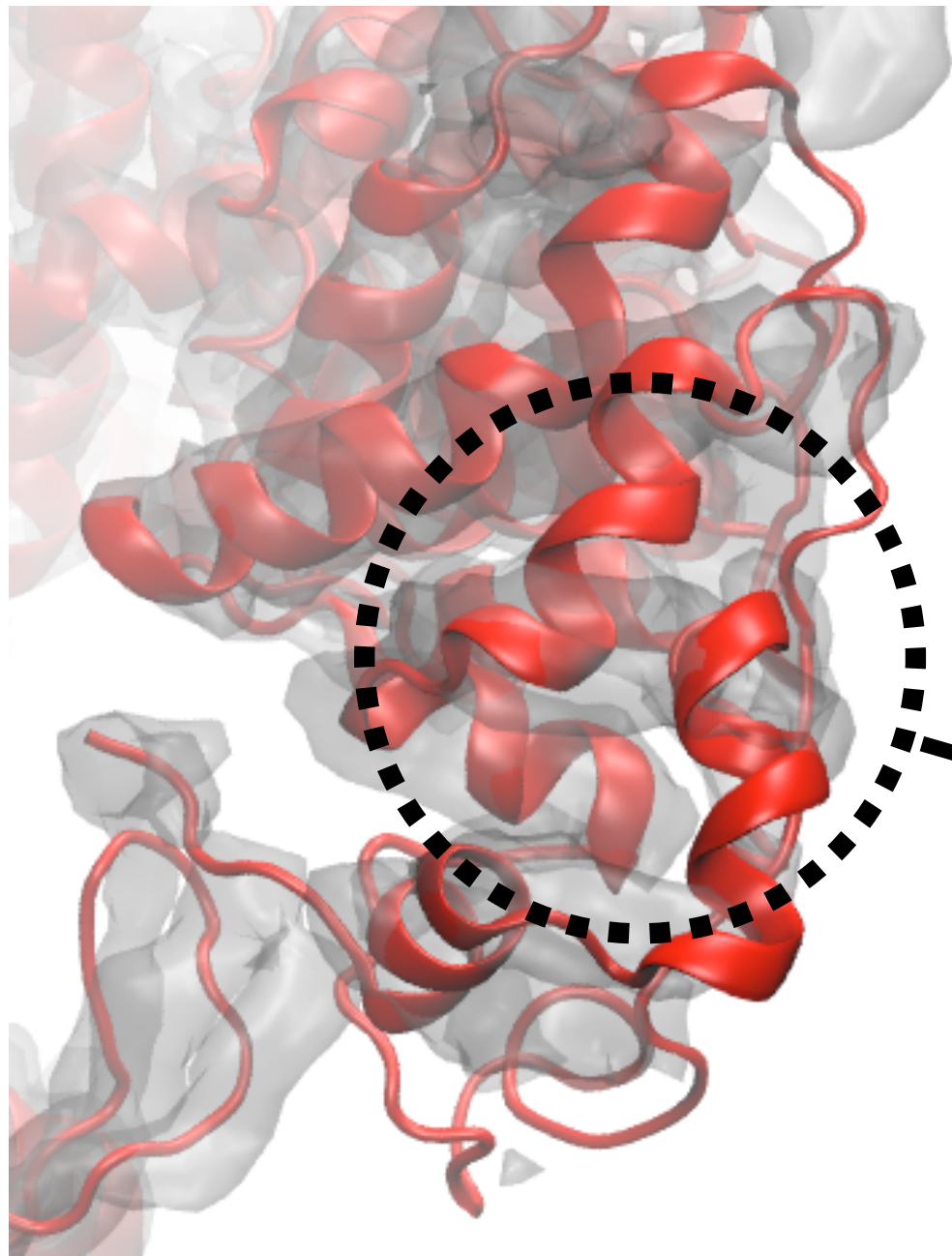
Eduard Schreiner, et al. BMC Bioinformatics, 12, 190, 2011

- Wrong chirality, cis-peptide bonds, and torsion angle outliers may arise during modeling
- VMD provides tools to check, visualize, and correct these errors
- These tools, together with MD force fields, produce models with good structural geometry

Ryan McGreevy*, Abhishek Singharoy*, et al. Acta Crystallographica D70, 2344-2355, 2014

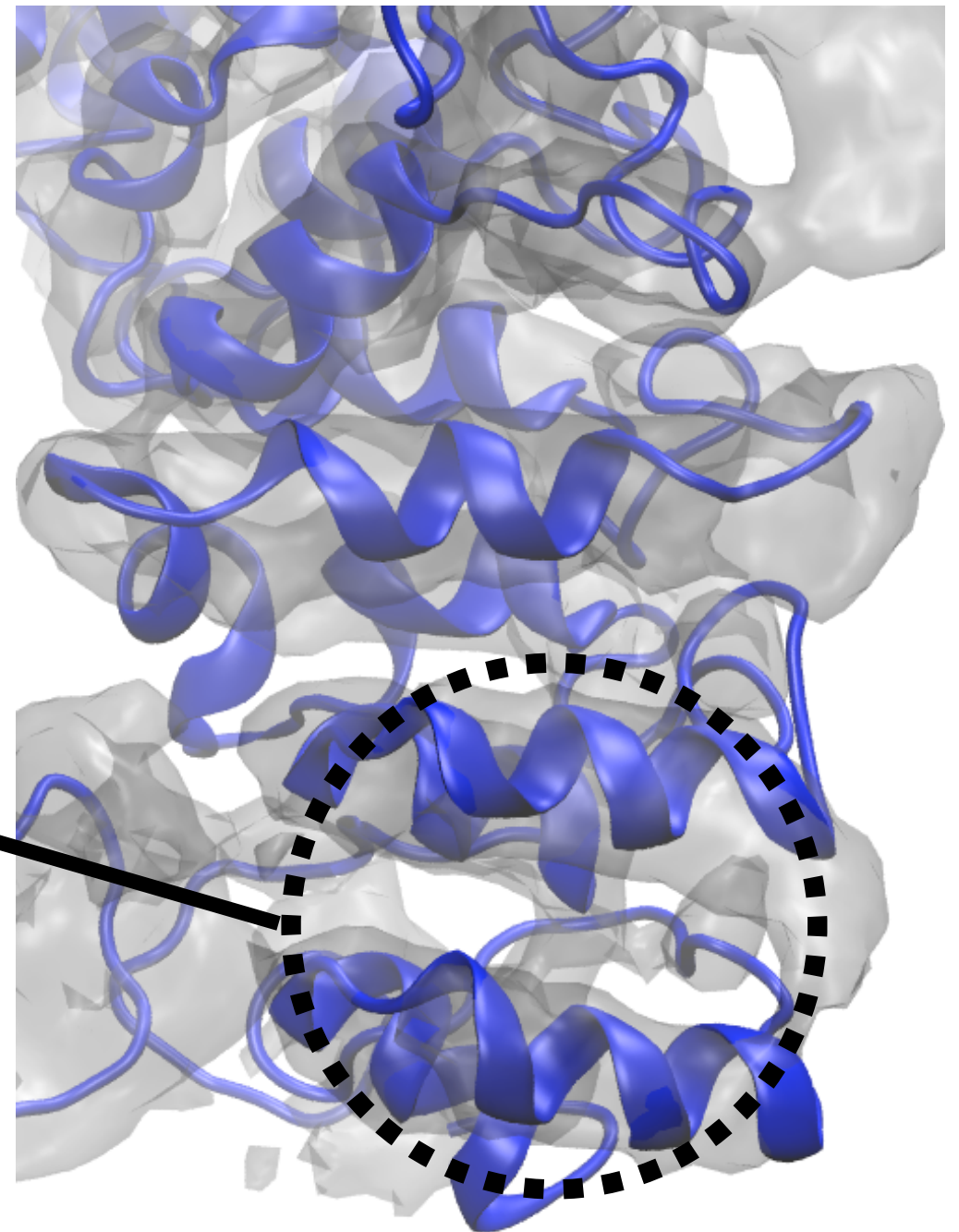
Analyzing MDFF Model Quality 2: **Global Cross Correlation**

Global CC is not always a good indicator of fit



CC = 0.728

RMSE(reference) = 6.23 Å



CC = 0.723

RMSE(reference) = 2.30 Å

Analyzing MDFF Model Quality 2: **Local Cross Correlation**

- Local cross correlation indicates quality of fit of specific regions across the entire structure
- New parallel CPU and GPU algorithms provide significant speed up (**25-50x speedup over Chimera**), allowing for fast computation along fitting trajectories

Structure is colored by cross correlation, along with **Timeline** analysis of the trajectory

(a)

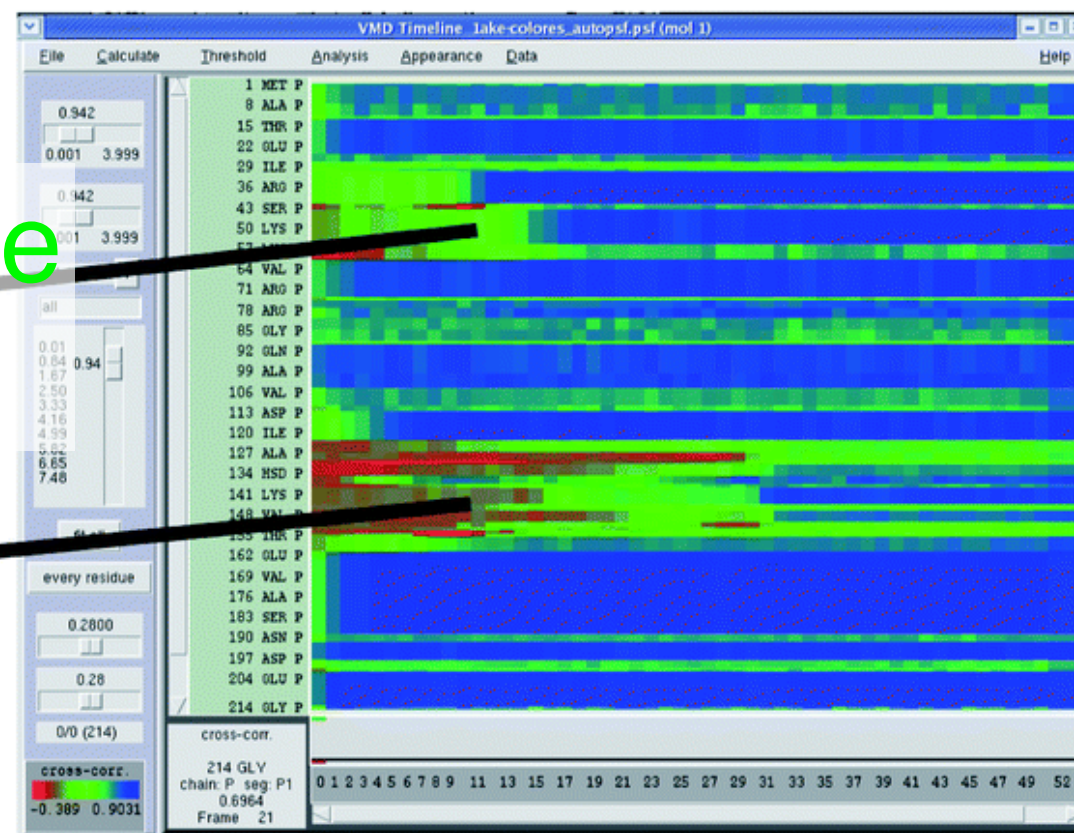
Good Fit

Intermediate Fit

(b)

Bad Fit

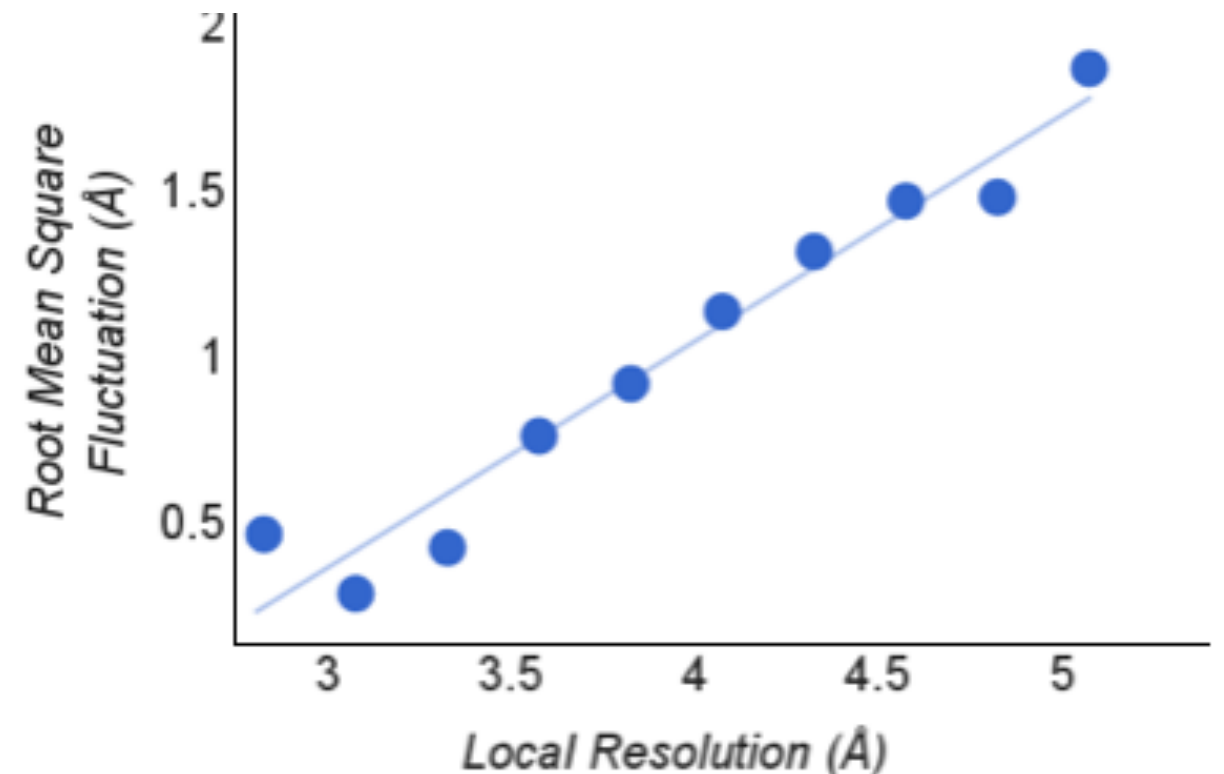
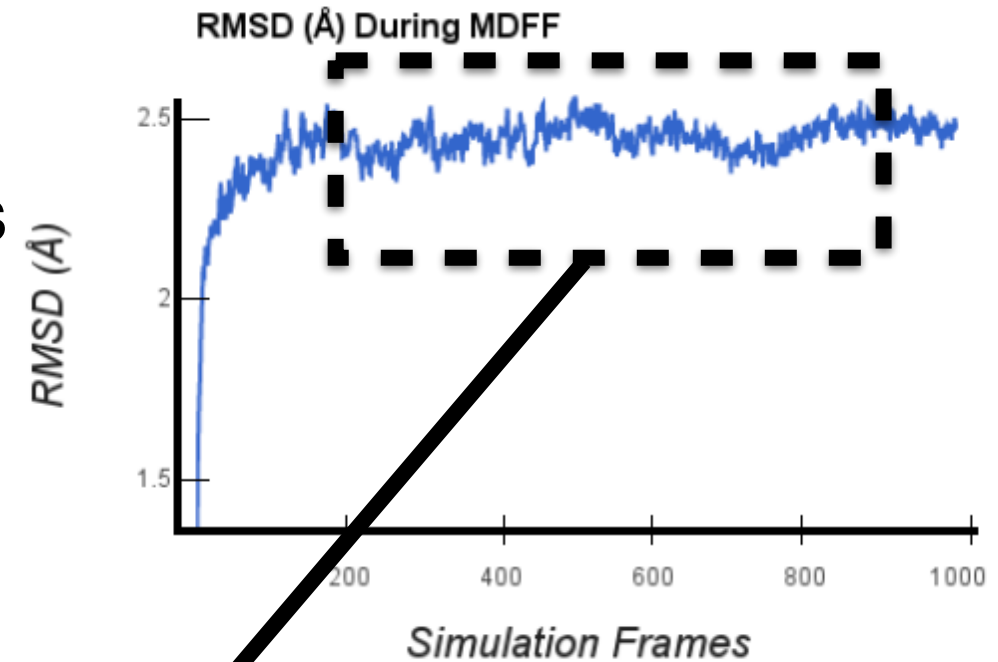
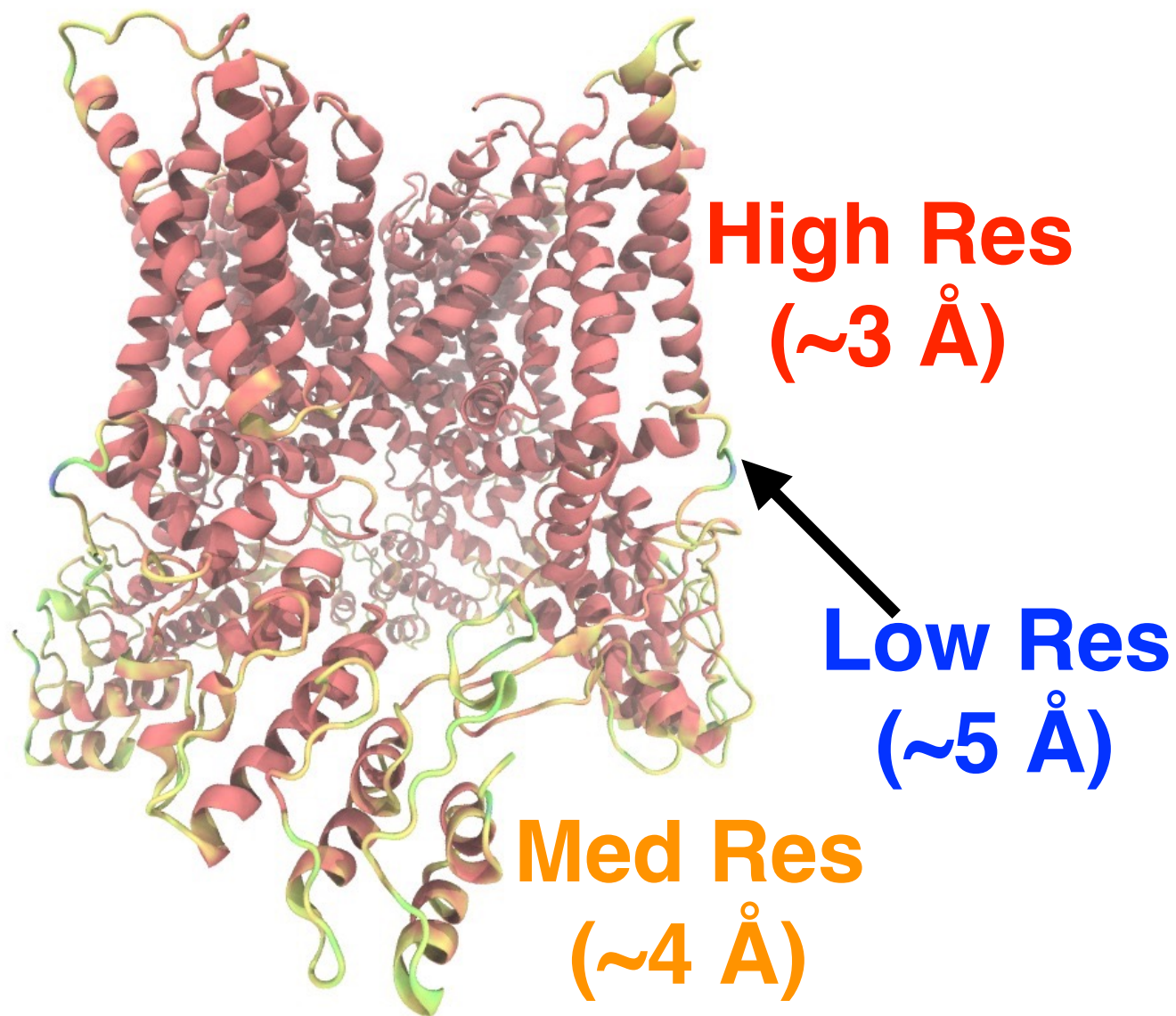
(c)



Analyzing MDFF Model Quality 3: **Local Resolution Analysis**

Local resolution of the experimental density from ResMap for error analysis and simulation parameterization

Root Mean Square Fluctuation (RMSF) correlates highly with local resolution

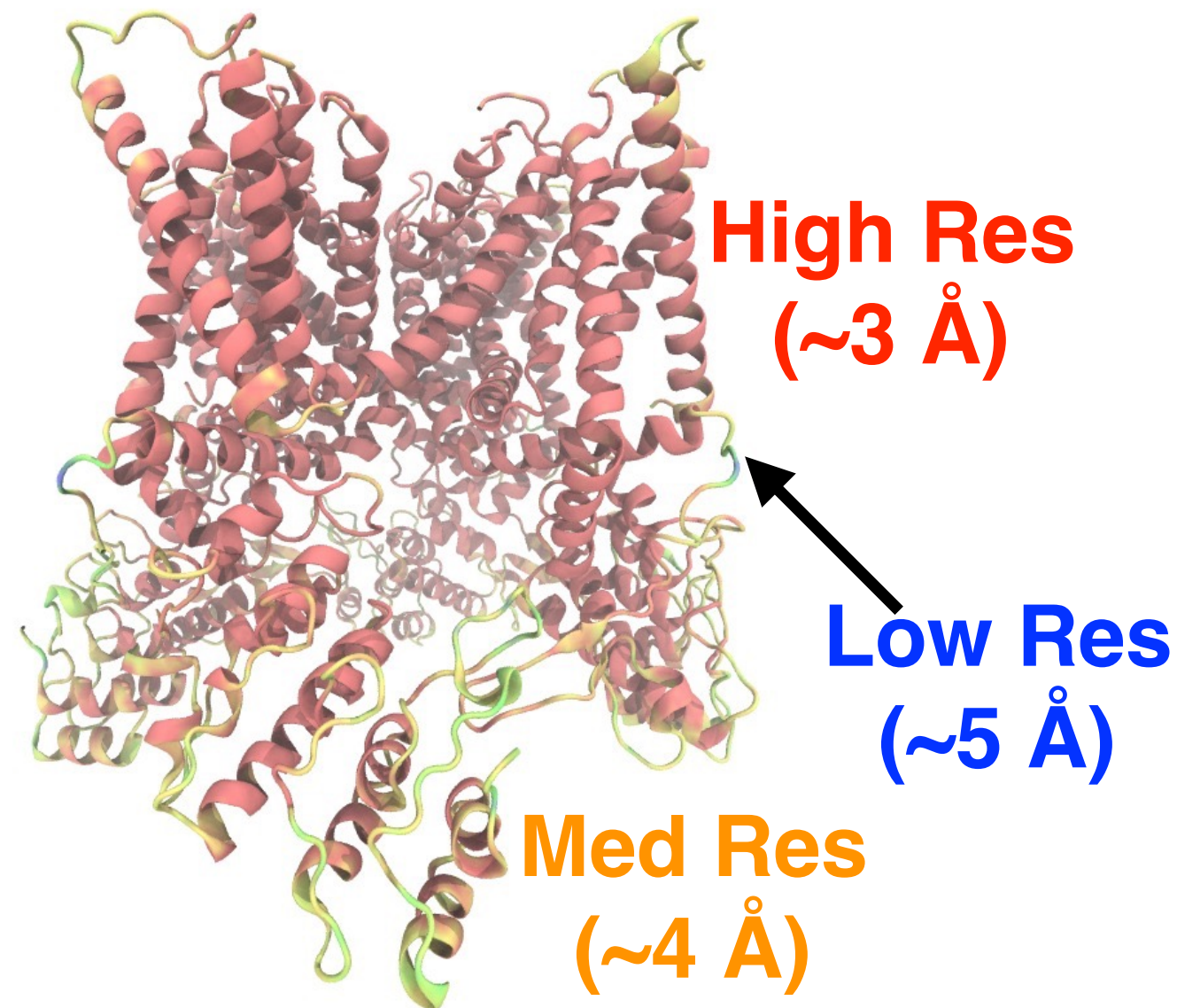


A. Kucukelbir, F.J. Sigworth, H.D. Tagare, Quantifying the Local Resolution of Cryo-EM Density Maps, Nature Methods, Volume 11, Issue 1, Pages 63-65, 2014.

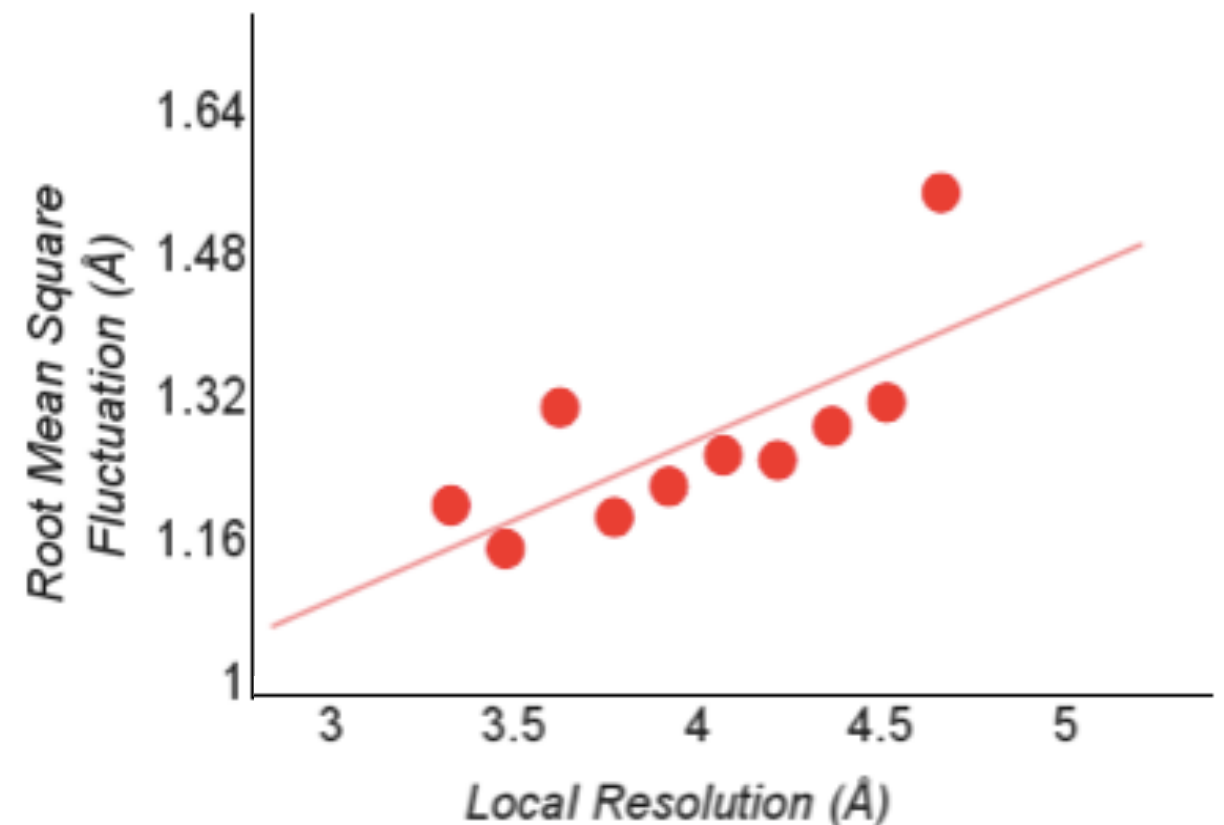
Analyzing MDFF Model Quality 3: **Local Resolution Analysis**

Local resolution of the experimental density from ResMap for error analysis and simulation parameterization

Root Mean Square Fluctuation (RMSF) correlates highly with local resolution



RMSF During Equilibration



A. Kucukelbir, F.J. Sigworth, H.D. Tagare, Quantifying the Local Resolution of Cryo-EM Density Maps, Nature Methods, Volume 11, Issue 1, Pages 63-65, 2014.

Analyzing MDFF Model Quality 4: **Cross-validation correlation**

Cascade and **direct** fitting structure to one half map and calculating the cross correlation to the other

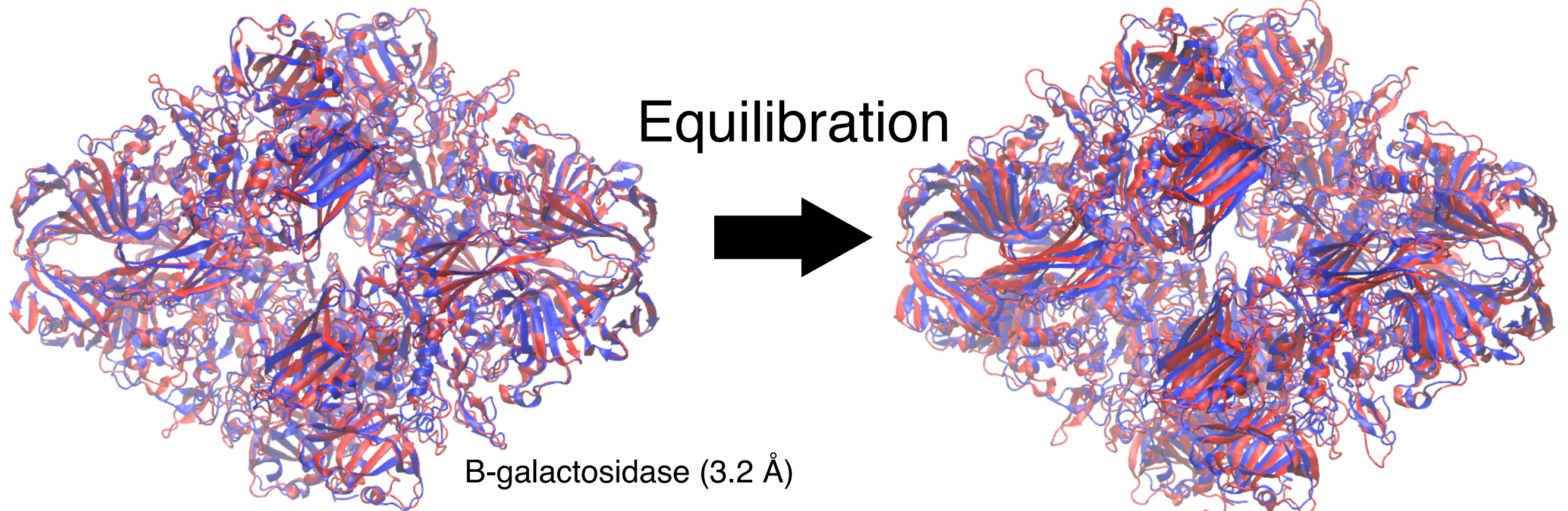
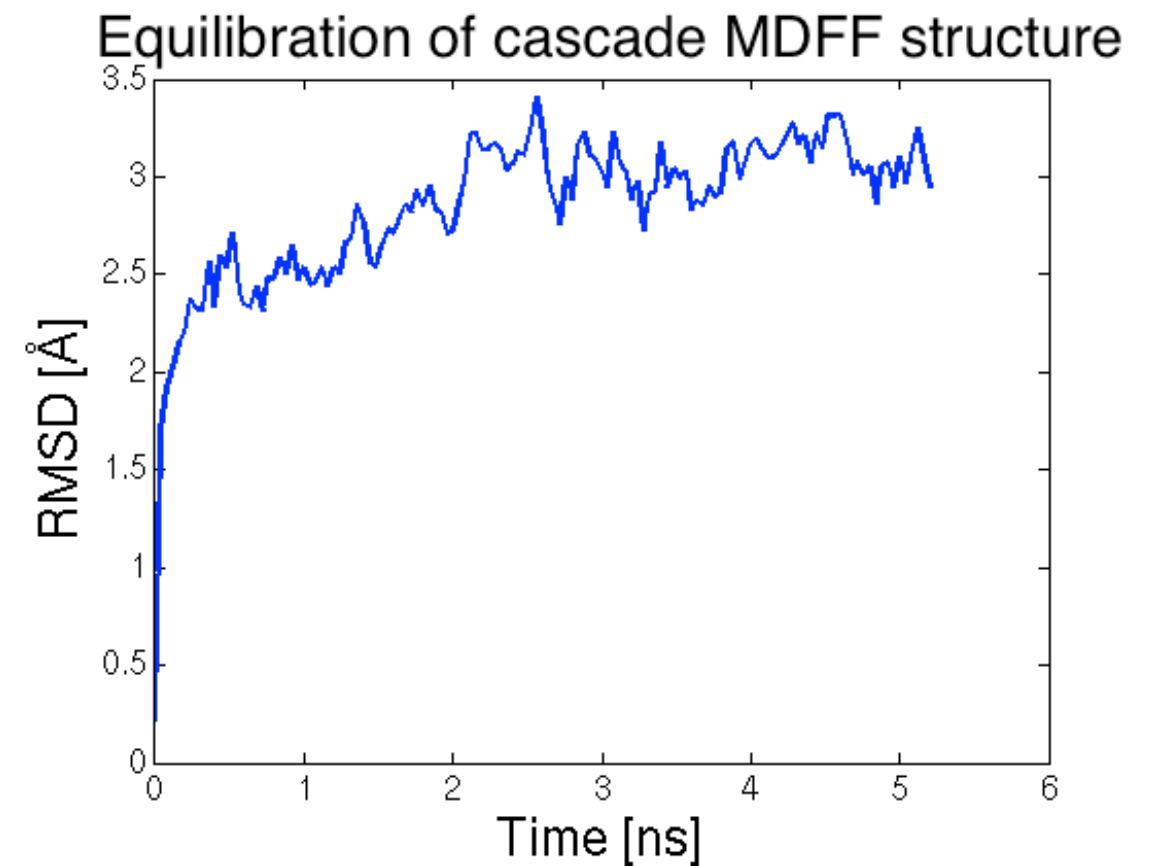
Fit to \ CC w.r.t.	Halfmap I	Halfmap II
	Halfmap I	Halfmap II
Halfmap I	0.715 (0.686)	0.714 (0.685)
Halfmap II	0.716 (0.688)	0.716 (0.688)

CC of reference structure w.r.t. each half map was 0.719

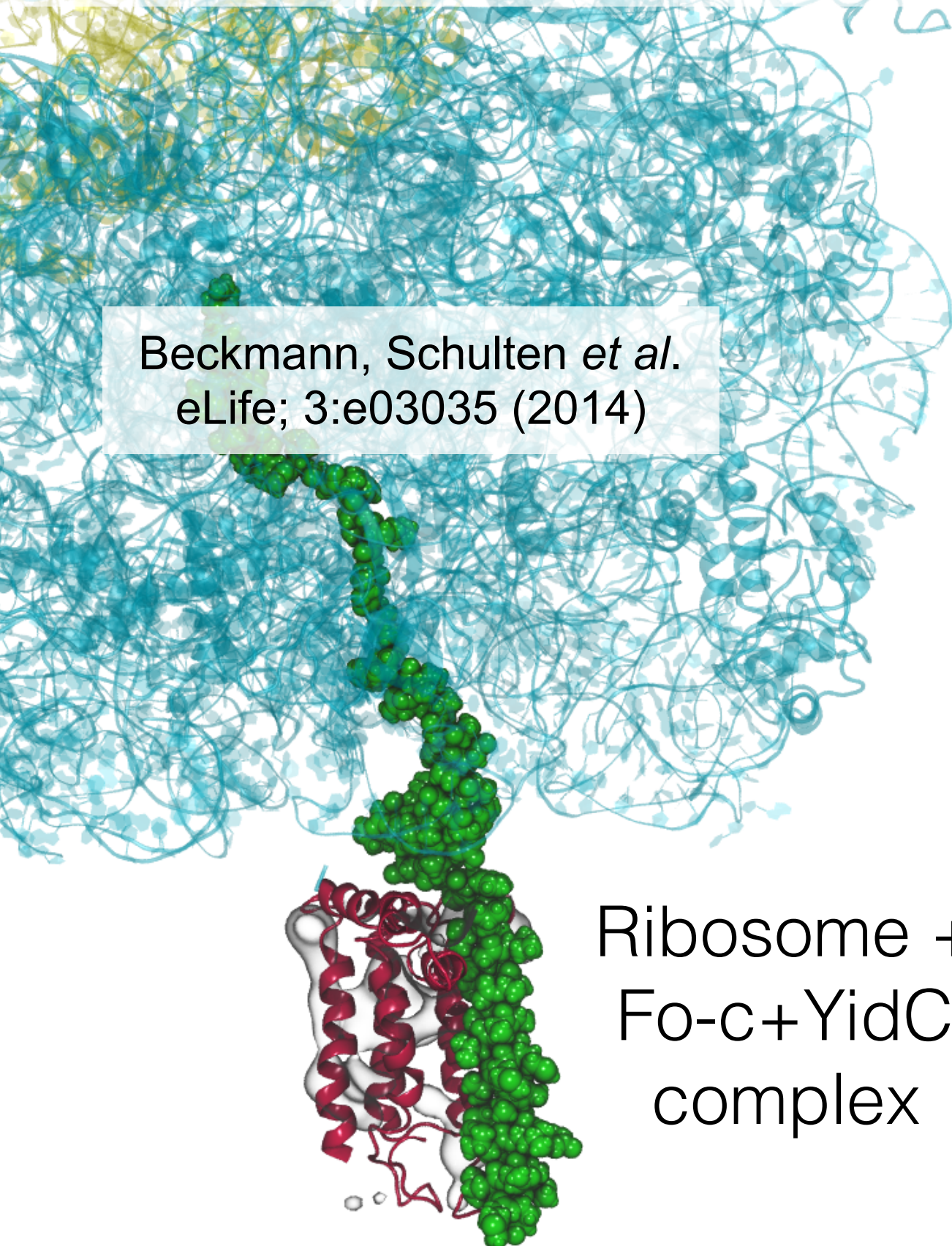
Analyzing MDFF Model Quality 5: MD post-processing

Stability of structure during equilibration

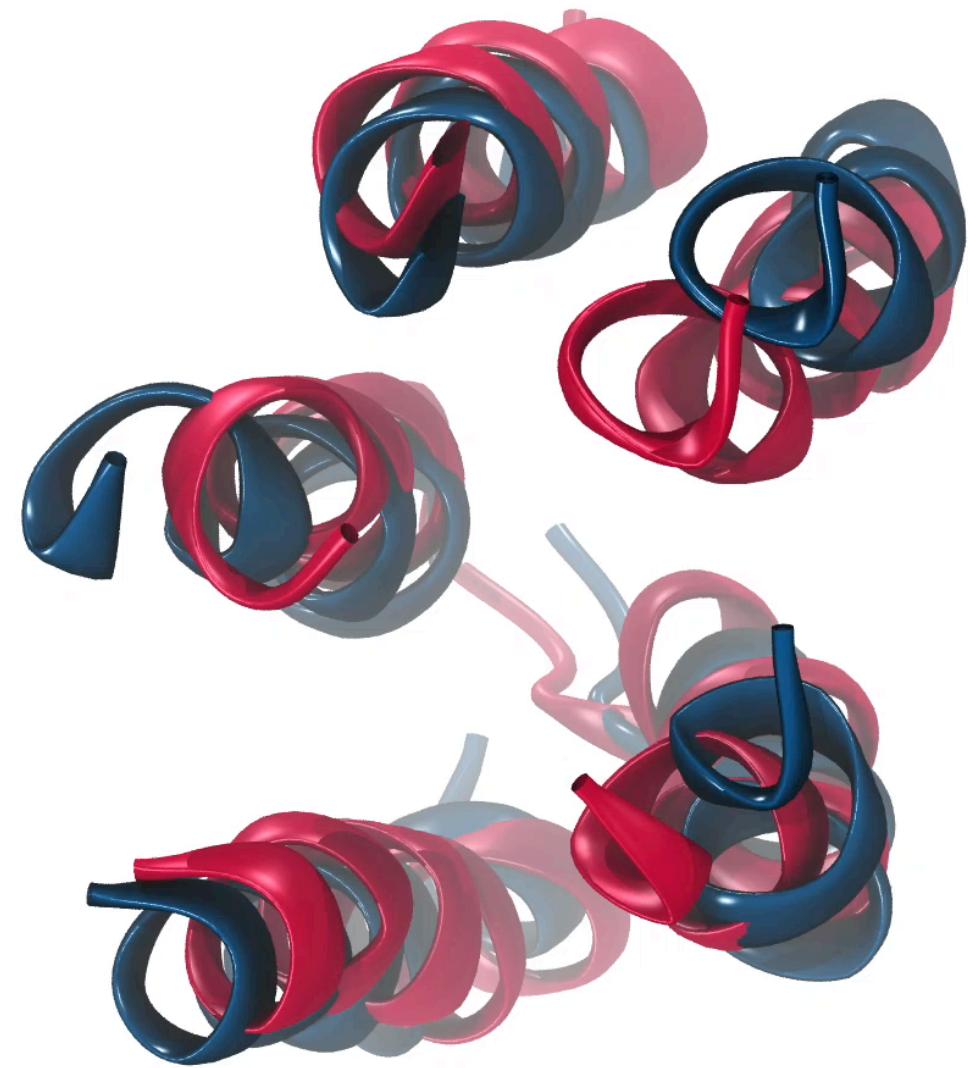
Deviation from fitted structure after equilibration is within map resolution ($\sim 3\text{\AA}$)



Analyzing MDFF Model Quality 6: Agreement with Experiment



**Ribosome-bound structure predicted
by MDFF from cryo-EM map ~ 7.5 Å**



Crystal Structure (3WVF) 3.2 Å

Kumazaki *et al.* Nature (2014)

Nascent chain confirmed also by chemical cross-linking, gel filtration chromatography and mass spectroscopy.

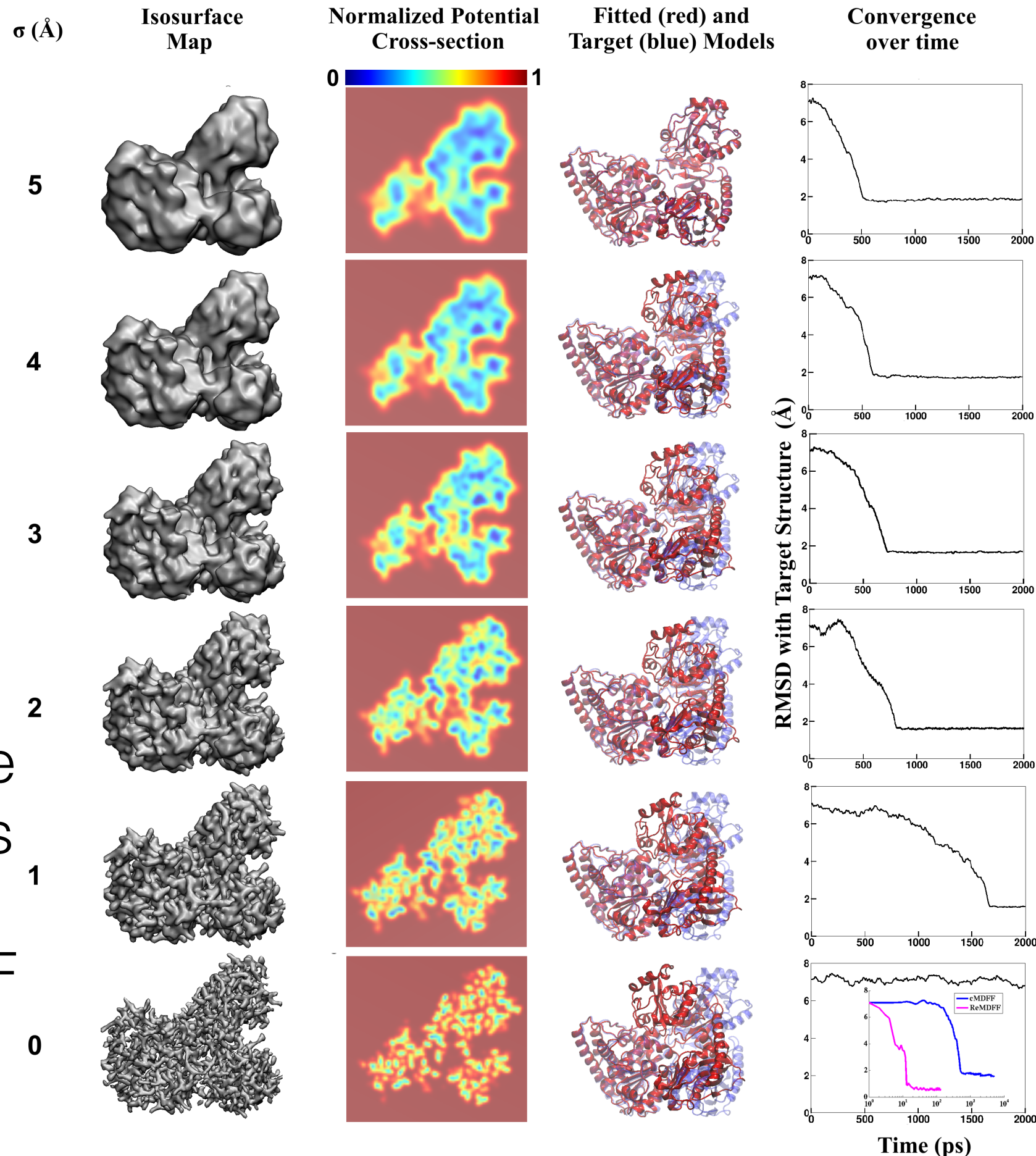
Molecular Dynamics Flexible Fitting Advanced Techniques

Ryan McGreevy
Research Programmer

University of Illinois at Urbana-Champaign
NIH Resource for Macromolecular Modeling and
Bioinformatics

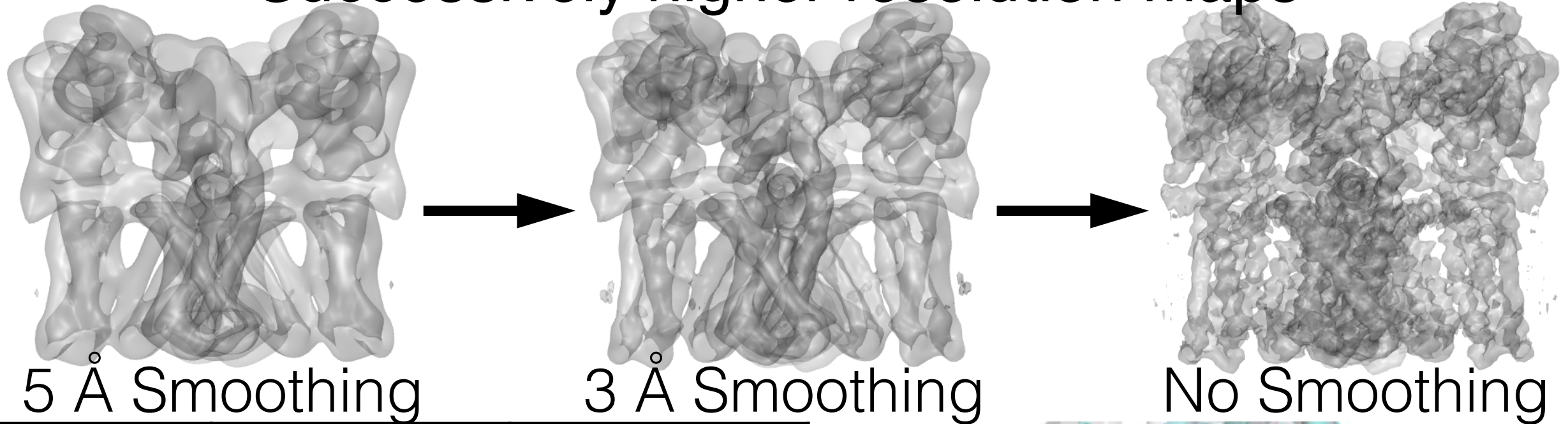
MDFF for high-resolution cryo-EM

- MDFF made in a time of lower resolution ($\sim 8\text{-}15\text{ \AA}$) EM maps
- High-resolution ($< 5\text{ \AA}$) now more easily obtainable
- Structure can become trapped in steep wells of high-resolution potential during MDFF

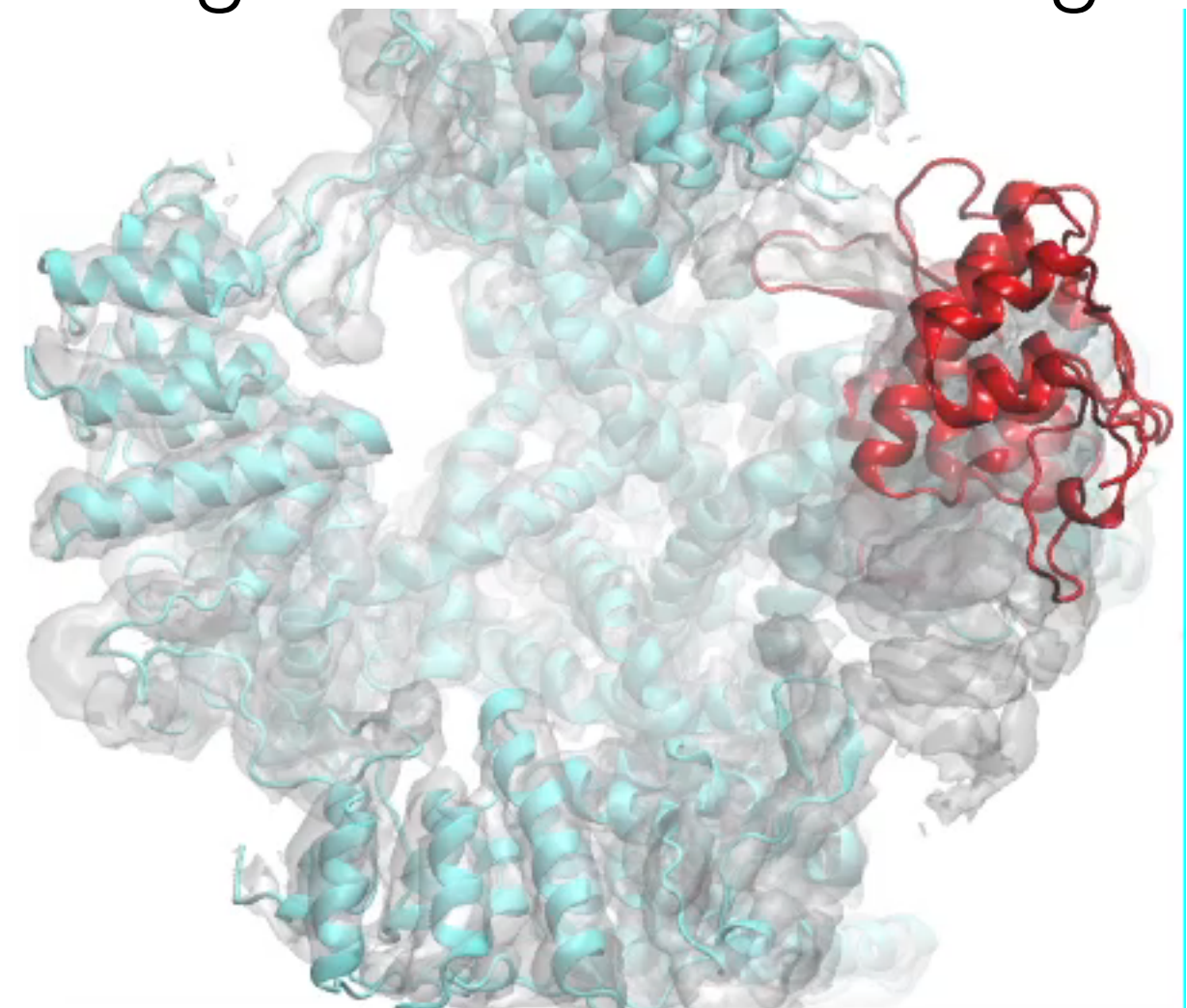


Cascade MDFF for **high-resolution** cryo-EM:

Successively higher resolution maps



Protocol	Global correlation	RMSD (Å)
Reference	0.732	-
Direct	0.699	12.41
Cascade	0.724	2.30



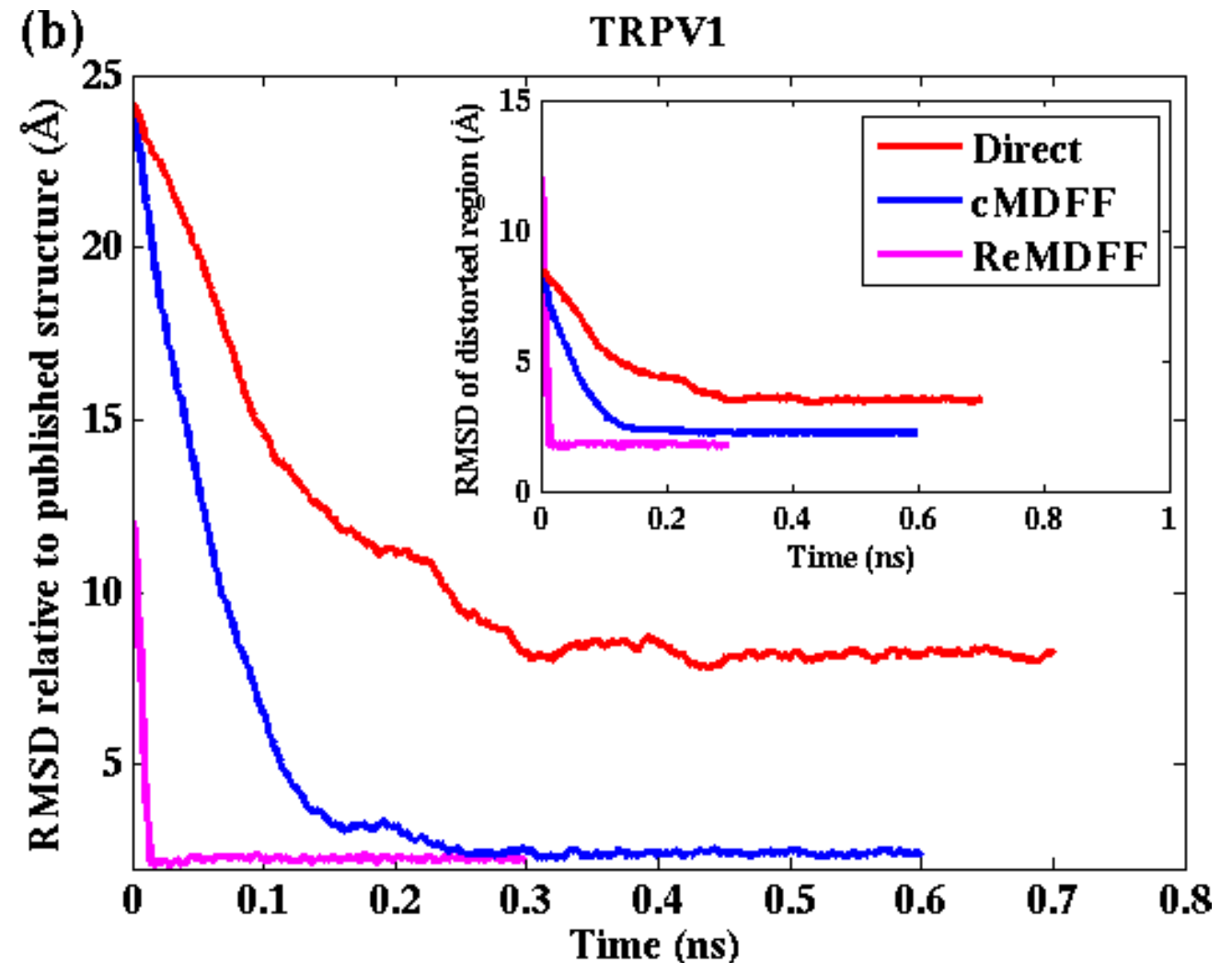
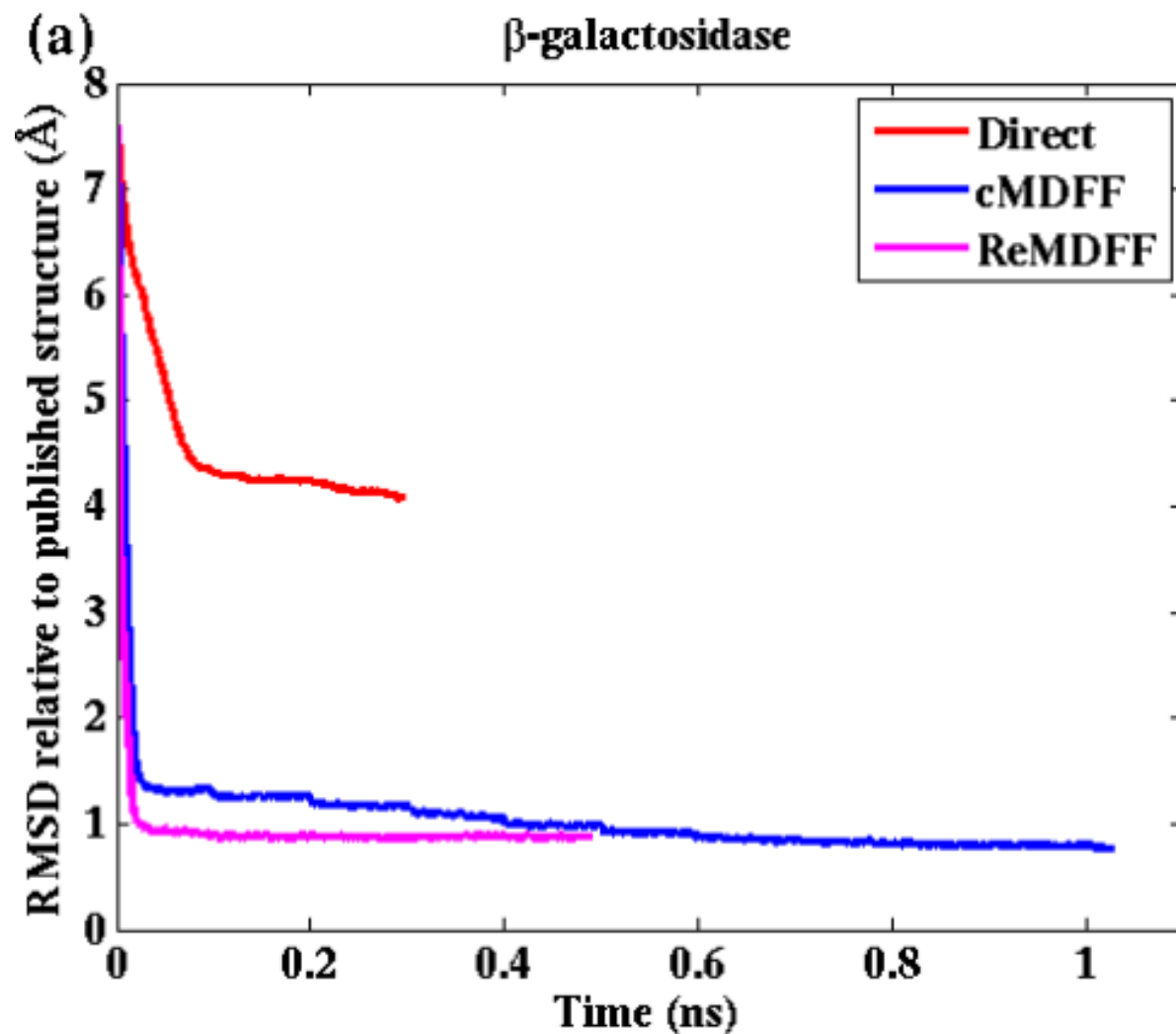
TRPV1 (3.4 Å)

Singharoy, Teo, McGreevy, et. al. eLife, 2016

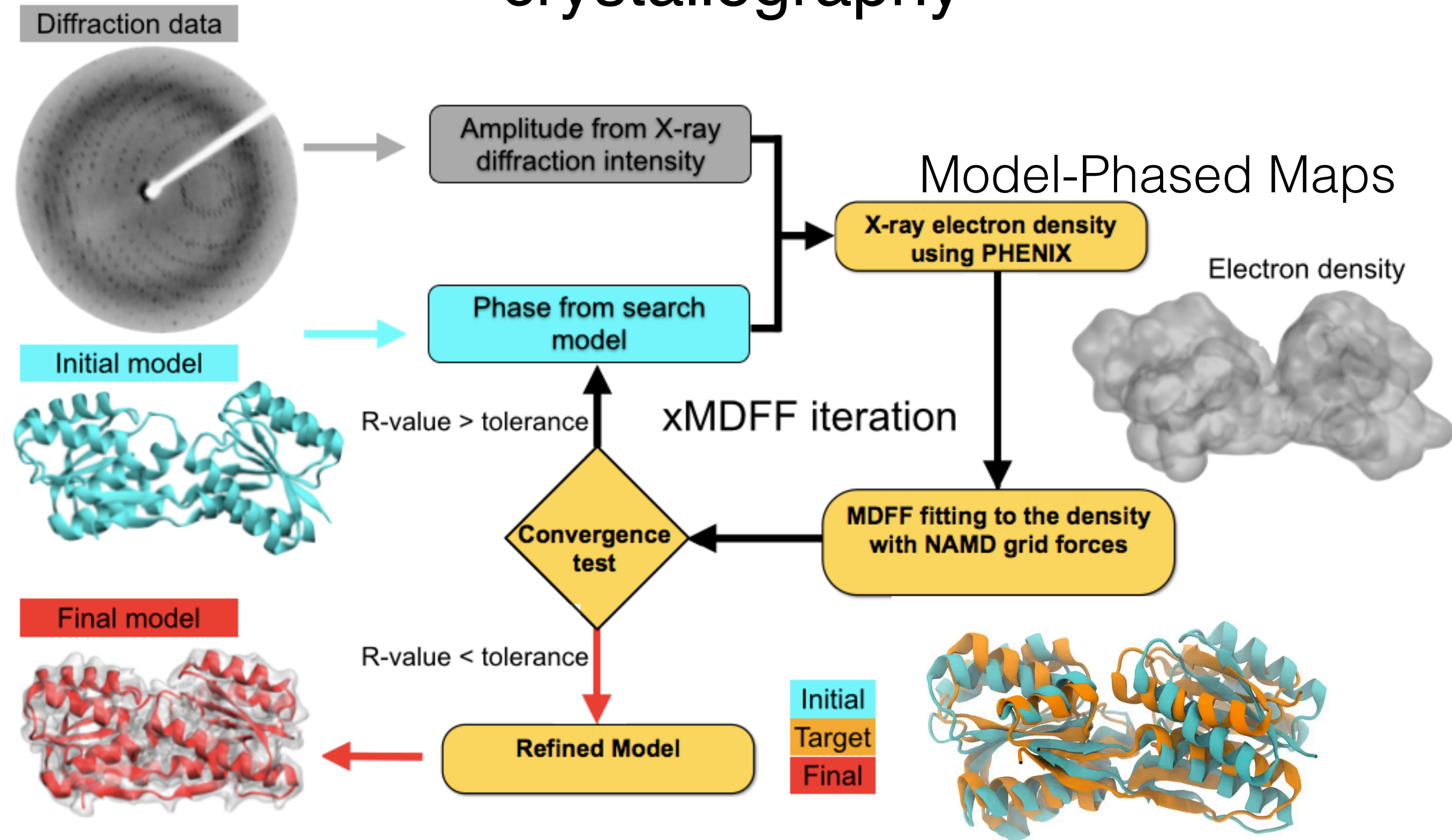
M. Liao, E. Cao, D. Julius, Y. Cheng, Nature 504, 2013

Resolution Exchange MDFF for **high-resolution** cryo-EM

- multiple maps blurred to varying resolution, like cMDFF
- independent parallel replicas (like Replica Exchange)
- each replica fits to a different map
- periodically exchange maps between replicas
- currently random exchange vs. parallel tempering which requires sufficient potential overlap of the energy distributions between neighboring replicas

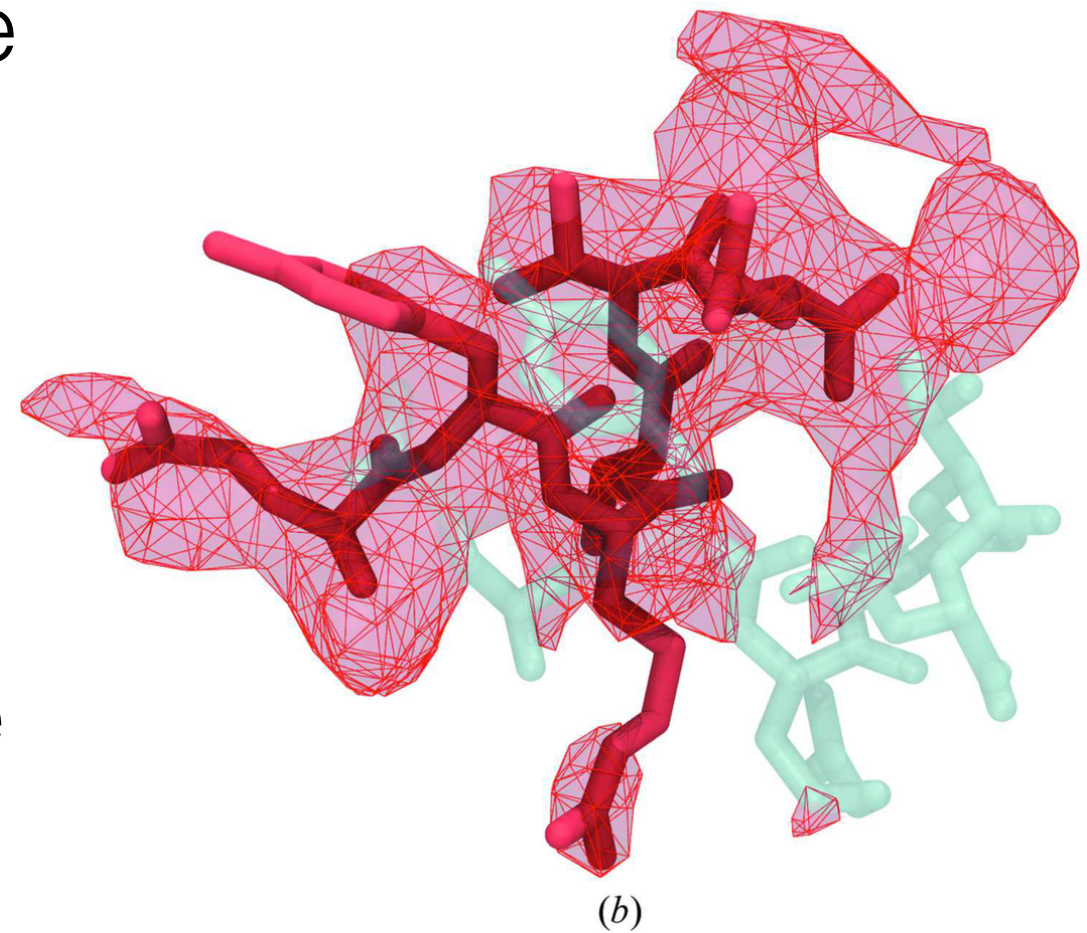
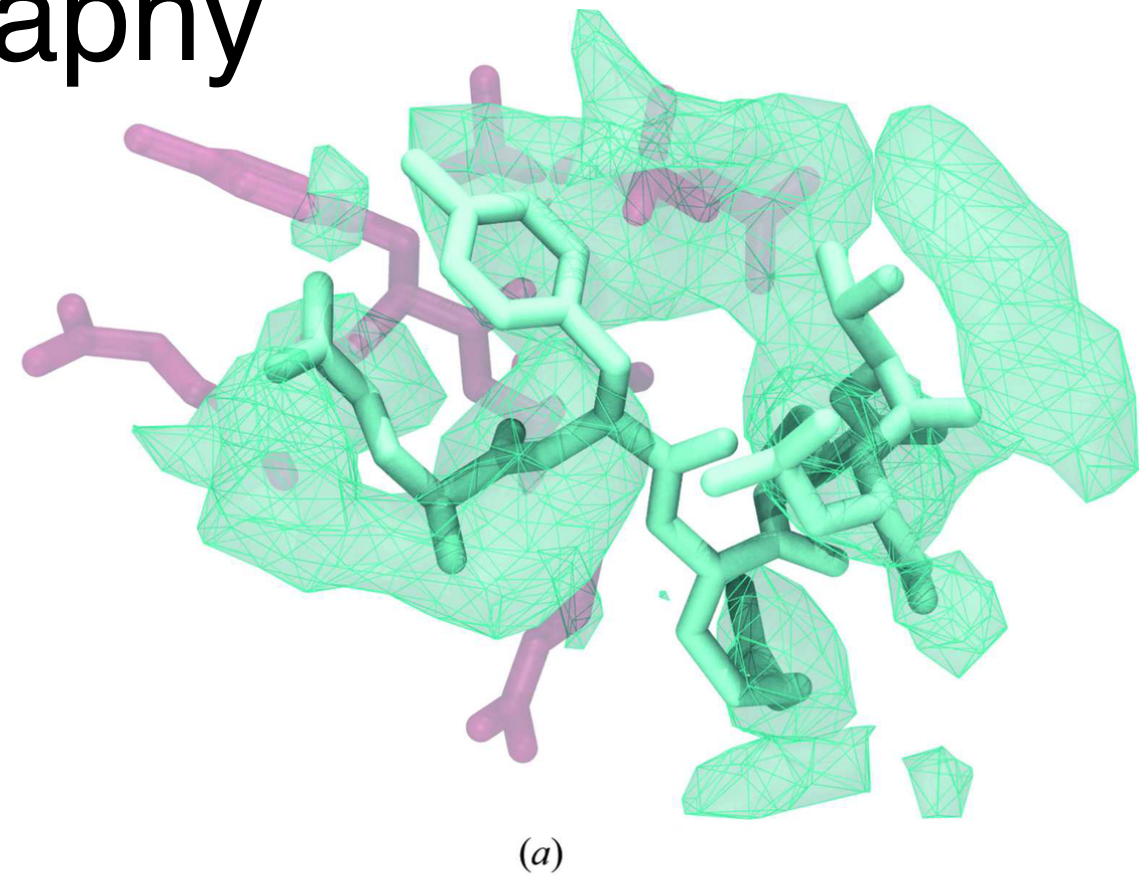


xMDFF: MDFF for low-resolution x-ray crystallography



xMDFF: MDFF for low-resolution x-ray crystallography

- Periodically generate new 2mFo-DFc maps using phenix.maps
- “Difference” maps amplify the regions of the map in which portions of the true model are missing
- Can use any phenix.maps parameters, e.g., “Feature-enhanced maps” which reduce model bias and noise

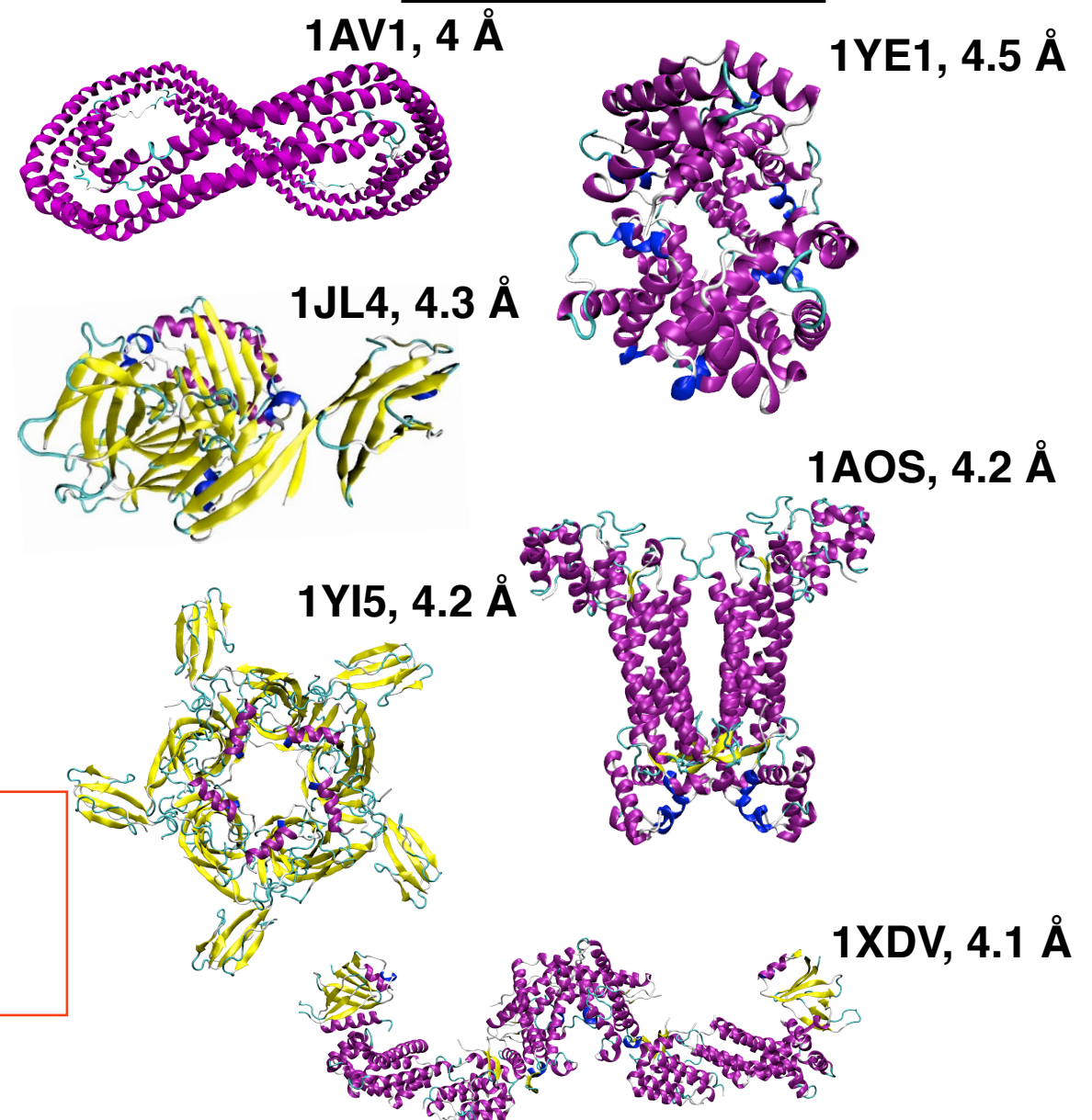


xMDFF Improves Structures Posted at the Protein Data Bank

Refinement statistics

PDB ID	Molprobity		R-work		R-free	
	initial	final	initial	final	initial	final
1AV1	3.72	1.94	0.38	0.33	0.42	0.34
1YE1	2.68	1.89	0.25	0.23	0.27	0.24
1JL4	3.24	1.47	0.36	0.33	0.42	0.38
1AOS	3.40	2.45	0.21	0.20	0.24	0.23
1XDV	2.87	2.01	0.39	0.29	0.41	0.33
1YI5	3.08	1.73	0.27	0.26	0.31	0.29

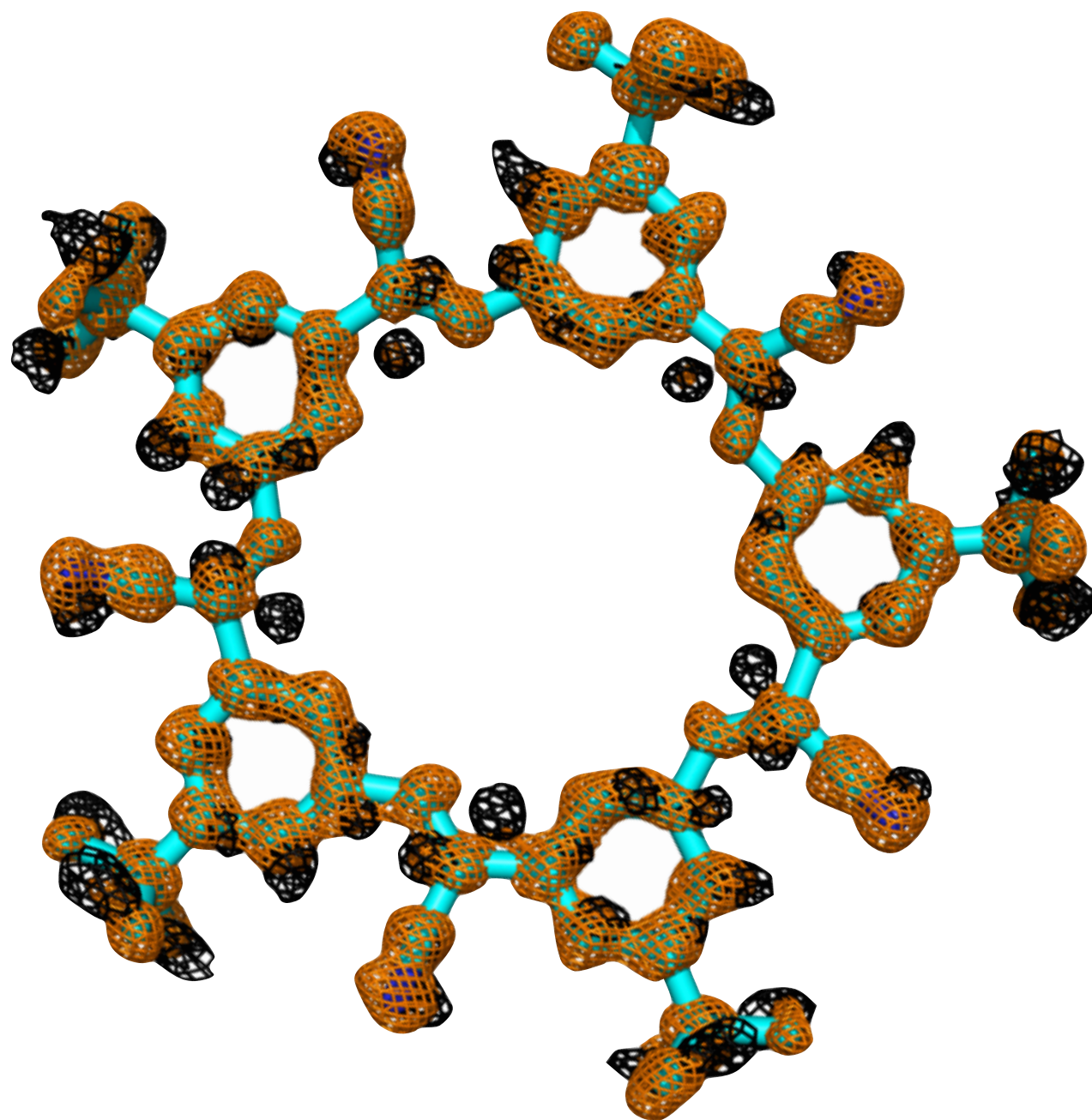
Structures



- Better R-work and R-free values than published before.
- Close R-work and R-free implies less over-fitting.
- Improved geometry implied by low Molprobity score.

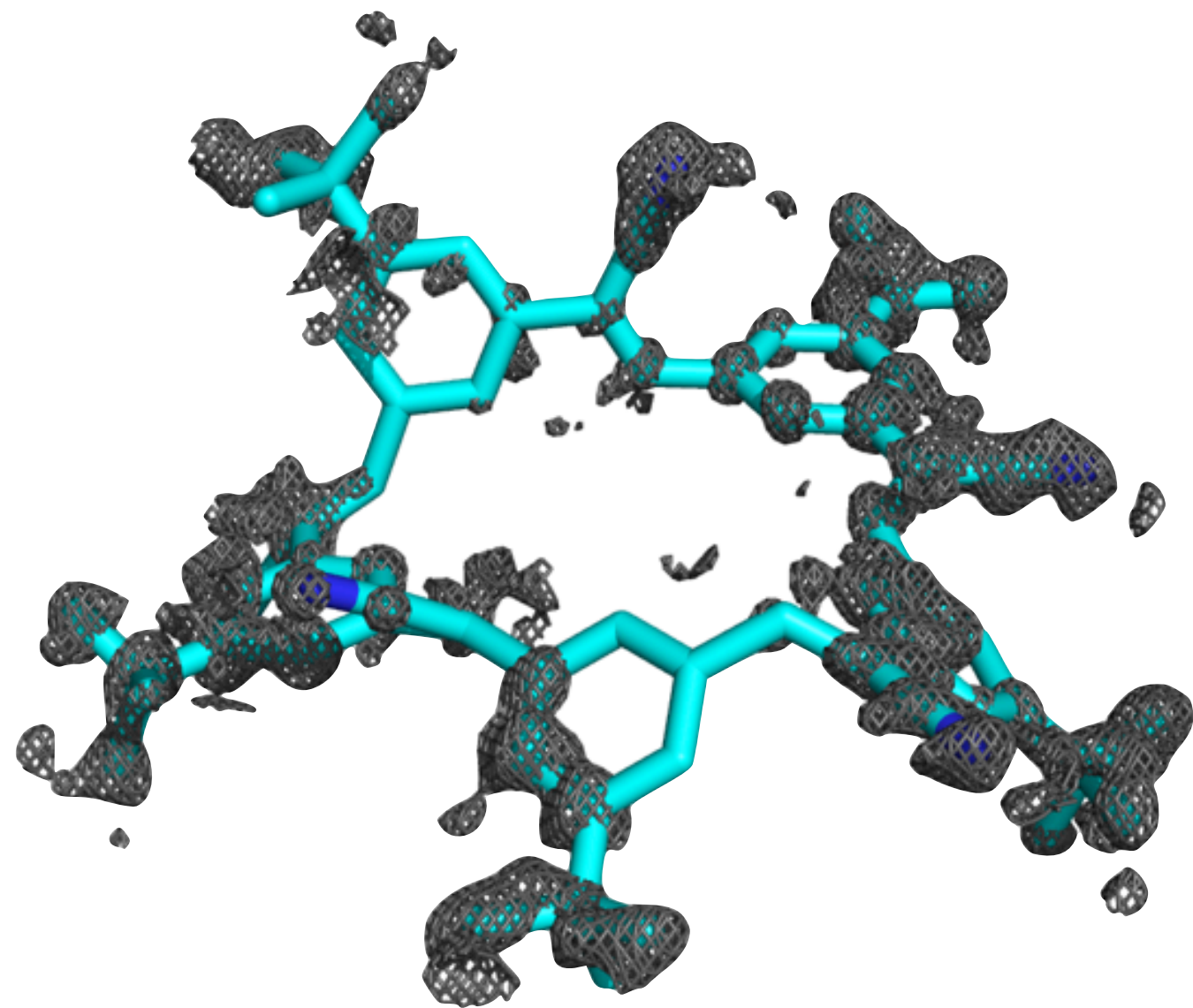
xMDFF for Abiological Materials

Cyanostar (2Å)



xMDFF-Phenix

(dual occupancy of CS shown in black and orange)



Phenix-only

Singharoy, et al. J. Am. Chem. Soc. 137 (27), pp 8810–8818, 2015.

MDFF Has a Wide Range of Applications

Over 100 reported MDFF applications:

- **By intramural Researchers:**

Schweitzer et al. ***PNAS*** (2016): Human 26S Proteasome

Cassidy et al. ***eLife*** (2016): Chemosensory array

Qufei Li et al. ***Nat. Struct. Mol. Biol.*** (2014): Structural mechanism of voltage-sensing protein

Zhao et al. ***Nature*** (2013): All-atom structure of HIV-1 capsid

Agirrezabala et al. ***PNAS*** (2012): Ribosome translocation intermediates

- **By extramural Researchers:**

He et al. ***Nature*** (2016): human pre-initiation complex

Li et al. ***Nature*** (2016): 20S proteasome

Barrio-Garcia et al. ***Nat. Struct. Mol. Biol.*** (2016): pre-60S-ribosome remodeling

Gogala et al. ***Nature*** (2014): Ribosome Sec61 complex

Unverdorben et al. ***PNAS*** (2014): 26S proteasome

Bharat et al. ***PNAS*** (2014): Tubular arrays of HIV-1 Gag

Park et al. ***Nature*** (2014): SecY channel during initiation of protein translocation

Hashem et al. ***Nature*** (2013): *Trypanosoma brucei* ribosome

Becker et al. ***Nature*** (2012): Ribosome recycling complex

Lasker et al. ***PNAS*** (2012): Proteasome

Strunk et al. ***Science*** (2011): Ribosome assembly factors

MDFF/xMDFF Methodological Articles:

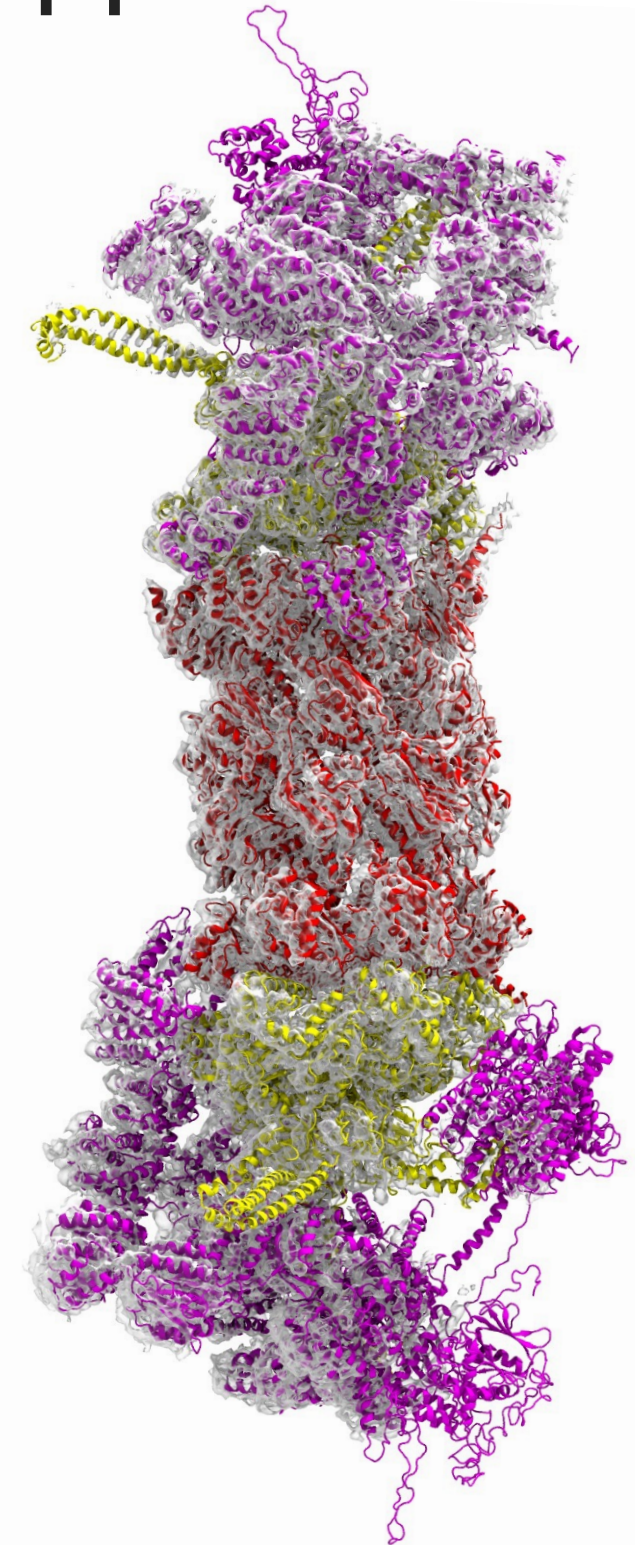
Singharoy*, Teo*, McGreevy*, et al. ***eLife*** (2016)

McGreevy et al. ***Methods*** (2016) 100:50-60

McGreevy*, Singharoy*, et al. ***Acta Crystallographica*** (2014) D70, 2344-2355

Trabuco et al. ***Structure*** (2008) 16:673-683.

Trabuco et al. ***Methods*** (2009) 49:174-180.



Full structure of the human 26S
Proteasome
Schweitzer et al., ***PNAS***. 2016

Acknowledgements and Further Information

Find out more about MDFF including:

- software downloads
- publications
- documentation
- **tutorials**

<http://www.ks.uiuc.edu/Research/mdff/>



Abhi Singharoy



Ivan Teo



Till Rudack