

# Molecular Dynamics Flexible Fitting

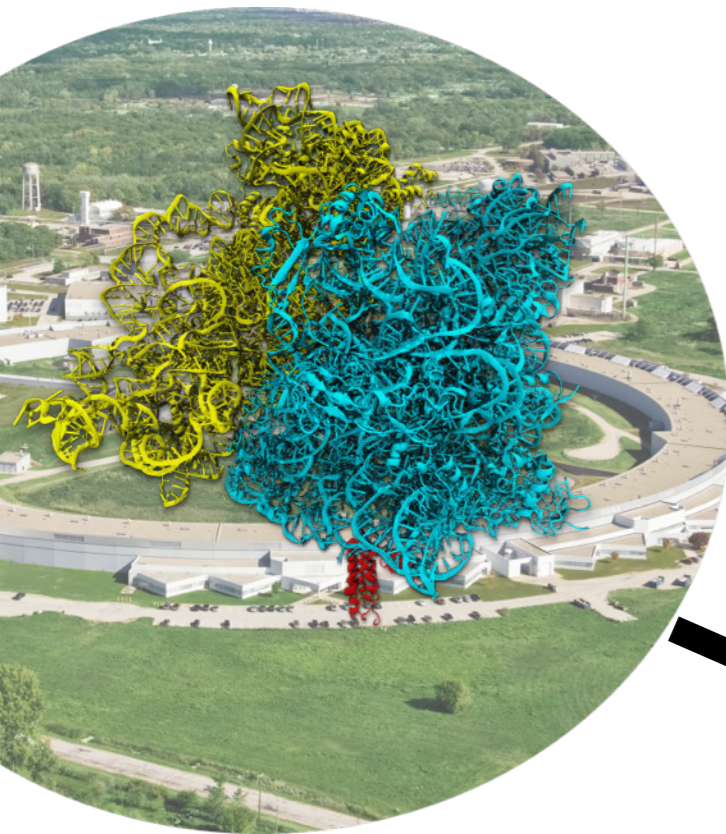
Ryan McGreevy  
Research Programmer

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

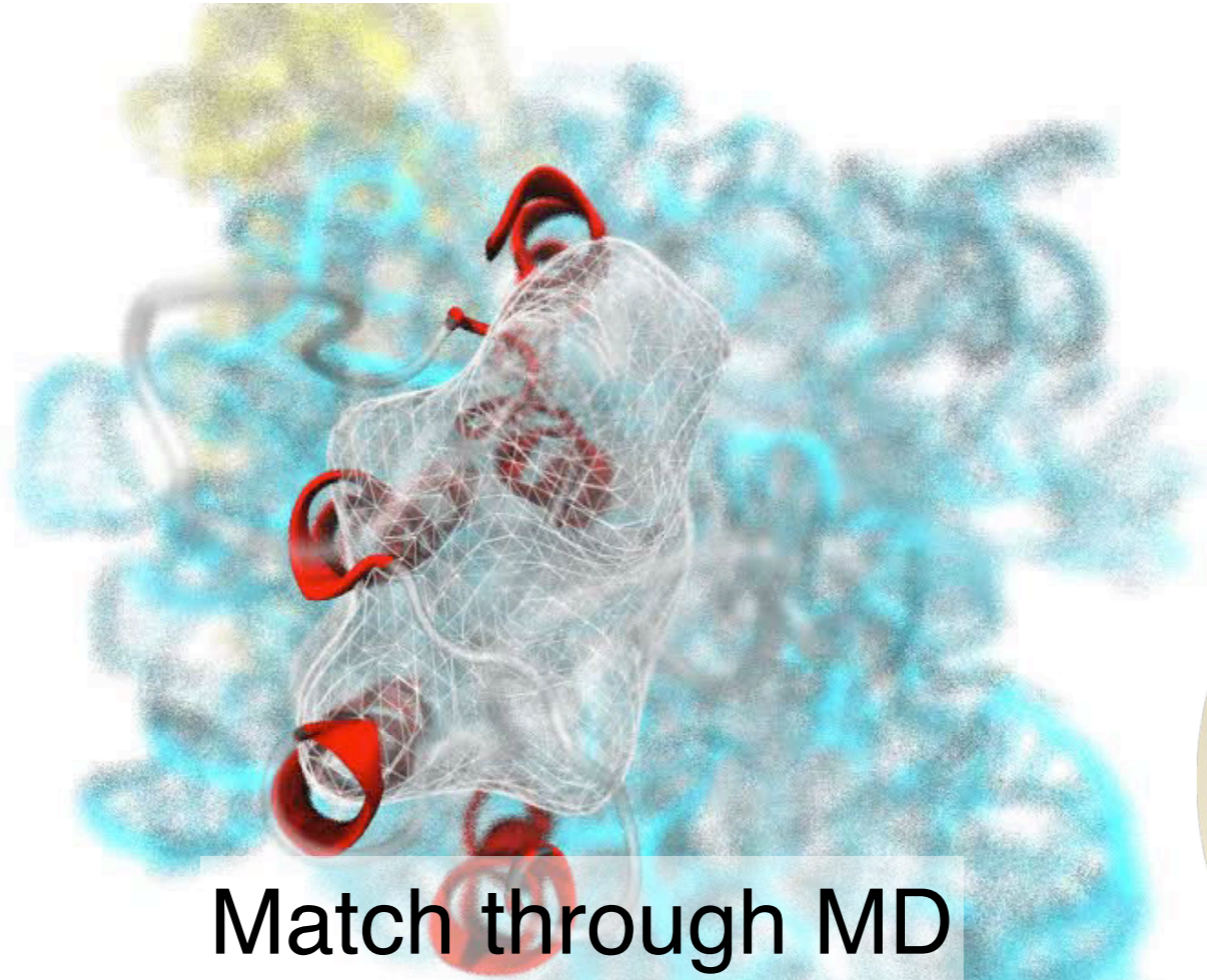
# Molecular Dynamics Flexible Fitting

(Ribosome-bound YidC)

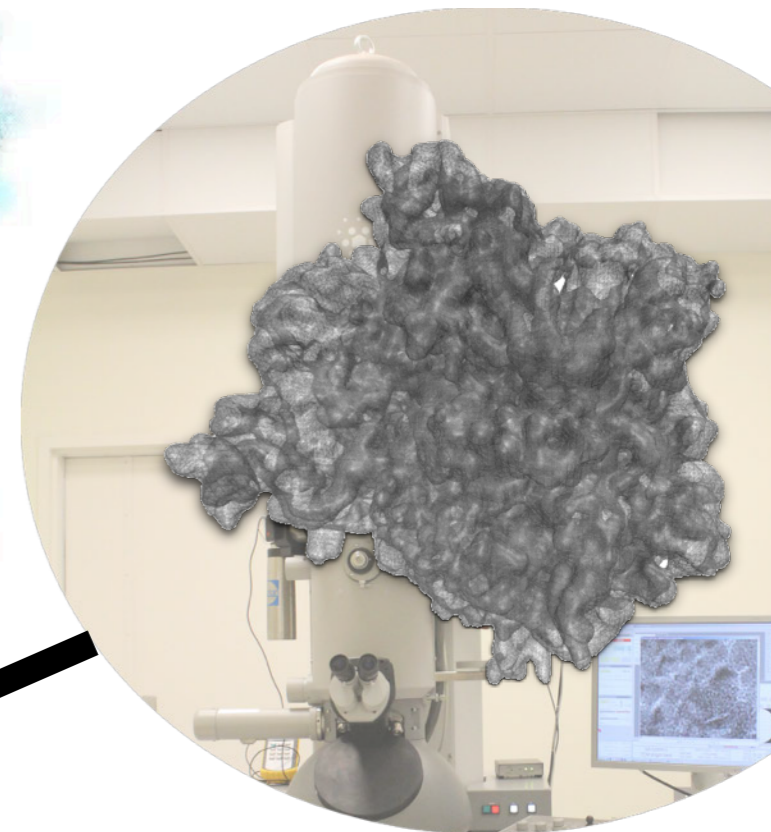
crystallographic  
structure



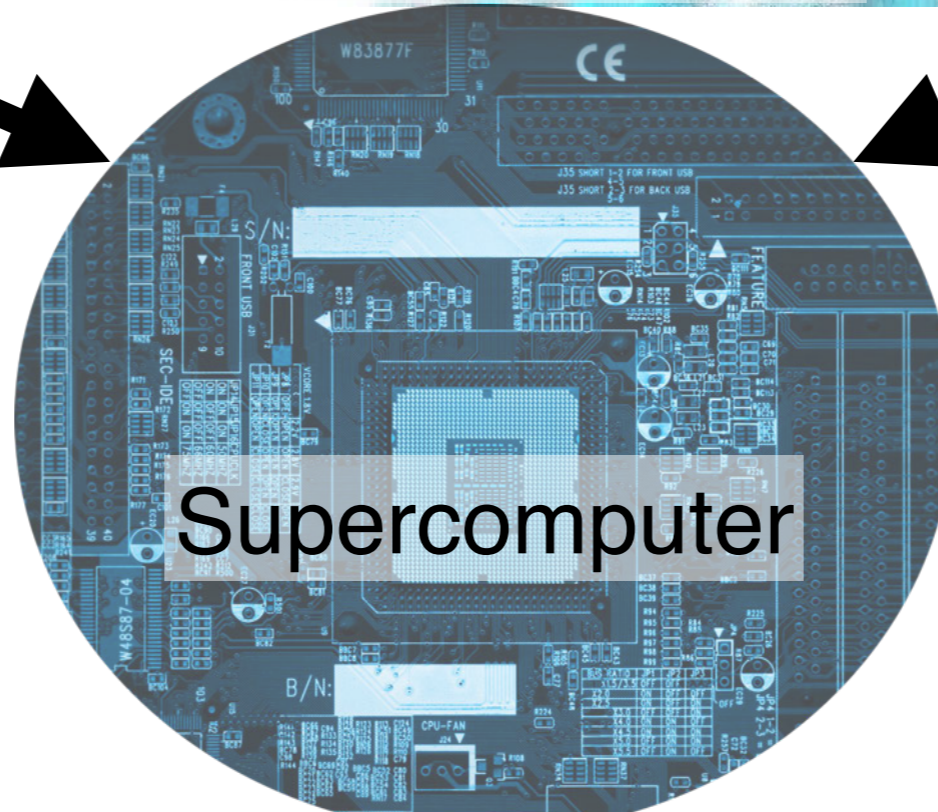
APS  
Synchrotron



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

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

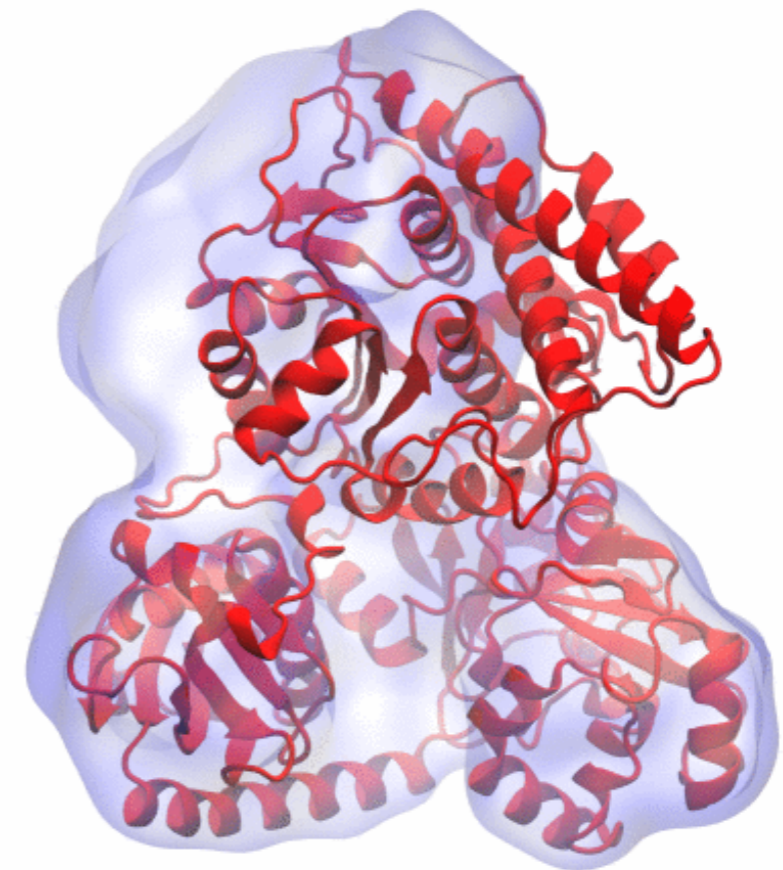
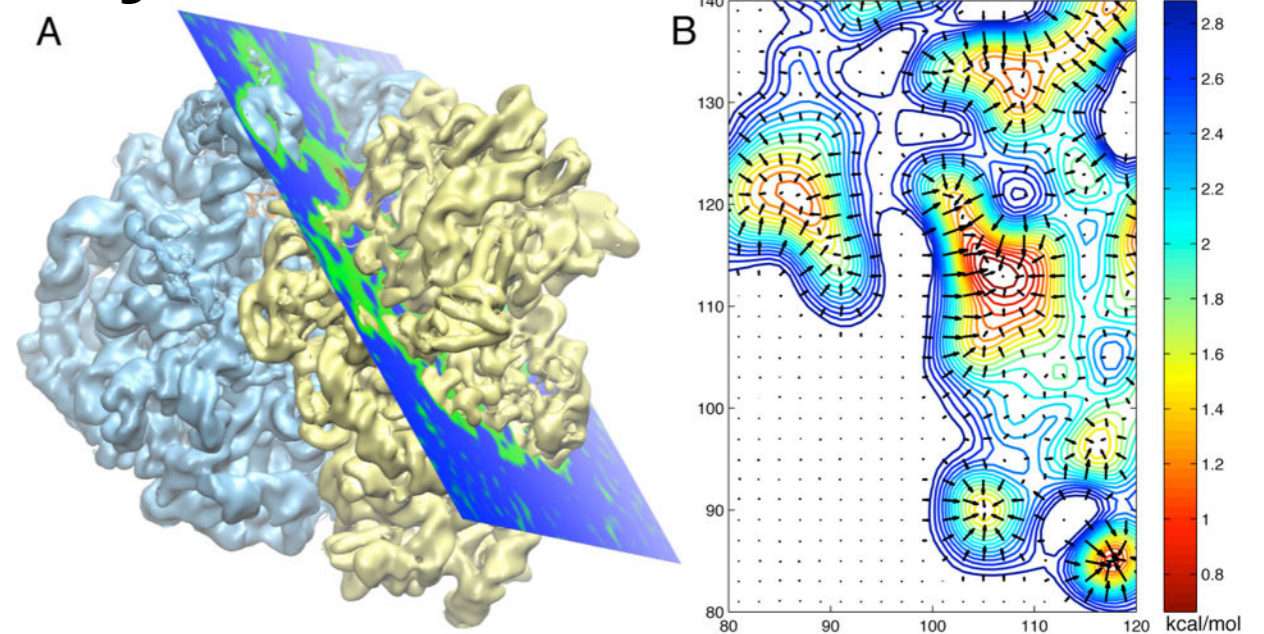
$$V_{EM}(\mathbf{r}) = \begin{cases} \xi \left( 1 - \frac{\Phi(\mathbf{r}) - \Phi_{thr}}{\Phi_{max} - \Phi_{thr}} \right) & \text{if } \Phi(\mathbf{r}) \geq \Phi_{thr}, \\ \xi & \text{if } \Phi(\mathbf{r}) < \Phi_{thr}. \end{cases}$$

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.



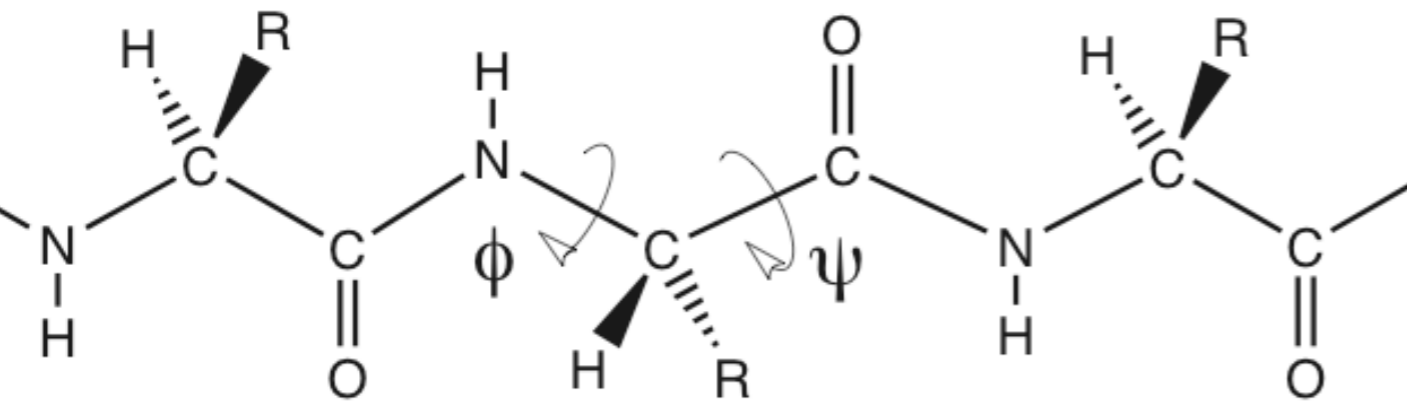
Acetyl – CoA Synthase

# Secondary structure restraints

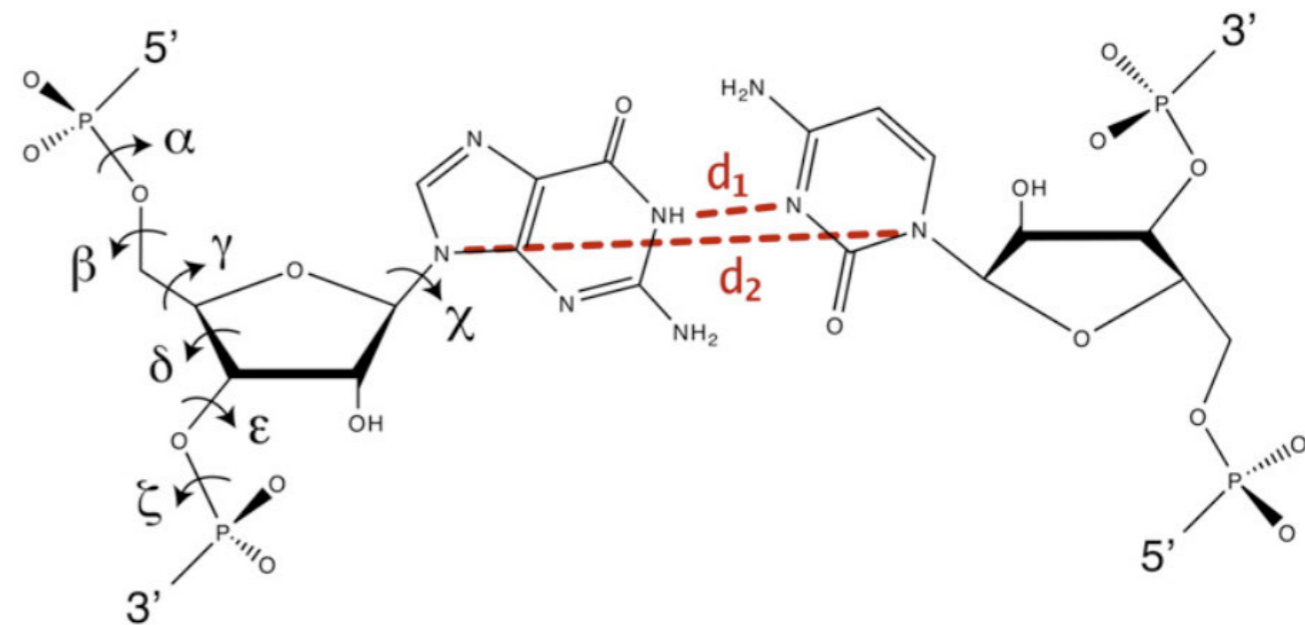
Harmonic restraints are applied to preserve secondary structure of proteins and nucleic acids, avoiding “overfitting.”

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

For proteins,  $\phi$  and  $\psi$  dihedral angles of residues within helices or beta strands are restrained.



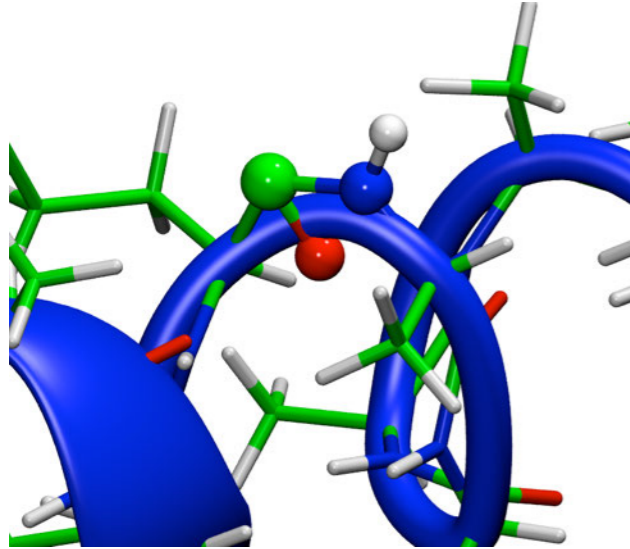
For nucleic acids, distance and dihedral restraints are applied to a selected set of base pairs.



# Additional Restraints

## Cis-peptide and Chirality

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

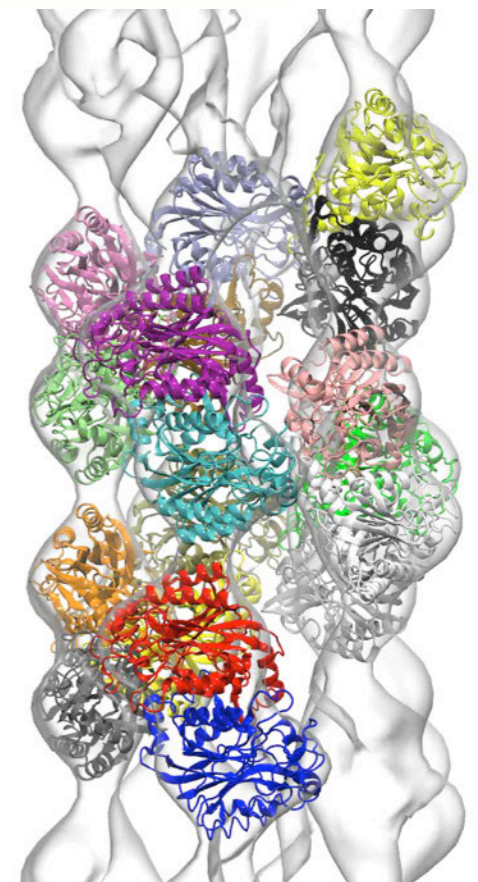
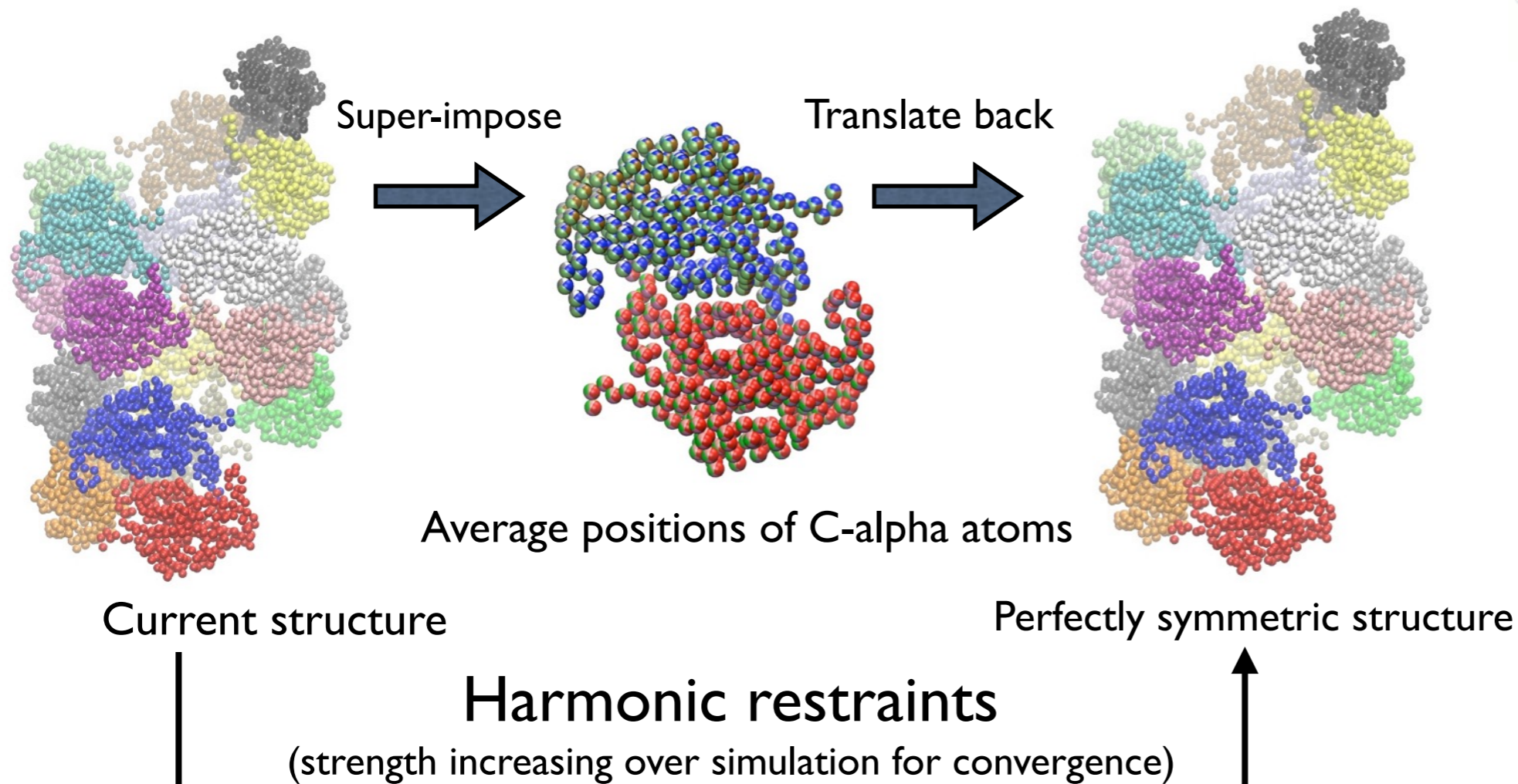


## Domain-wise



Acetyl – CoA Synthase

## Symmetry



B. pumilus cyanide dihydratase

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

# Simulation Environment

MDFE can be run in different environments:

## 1. Vacuum

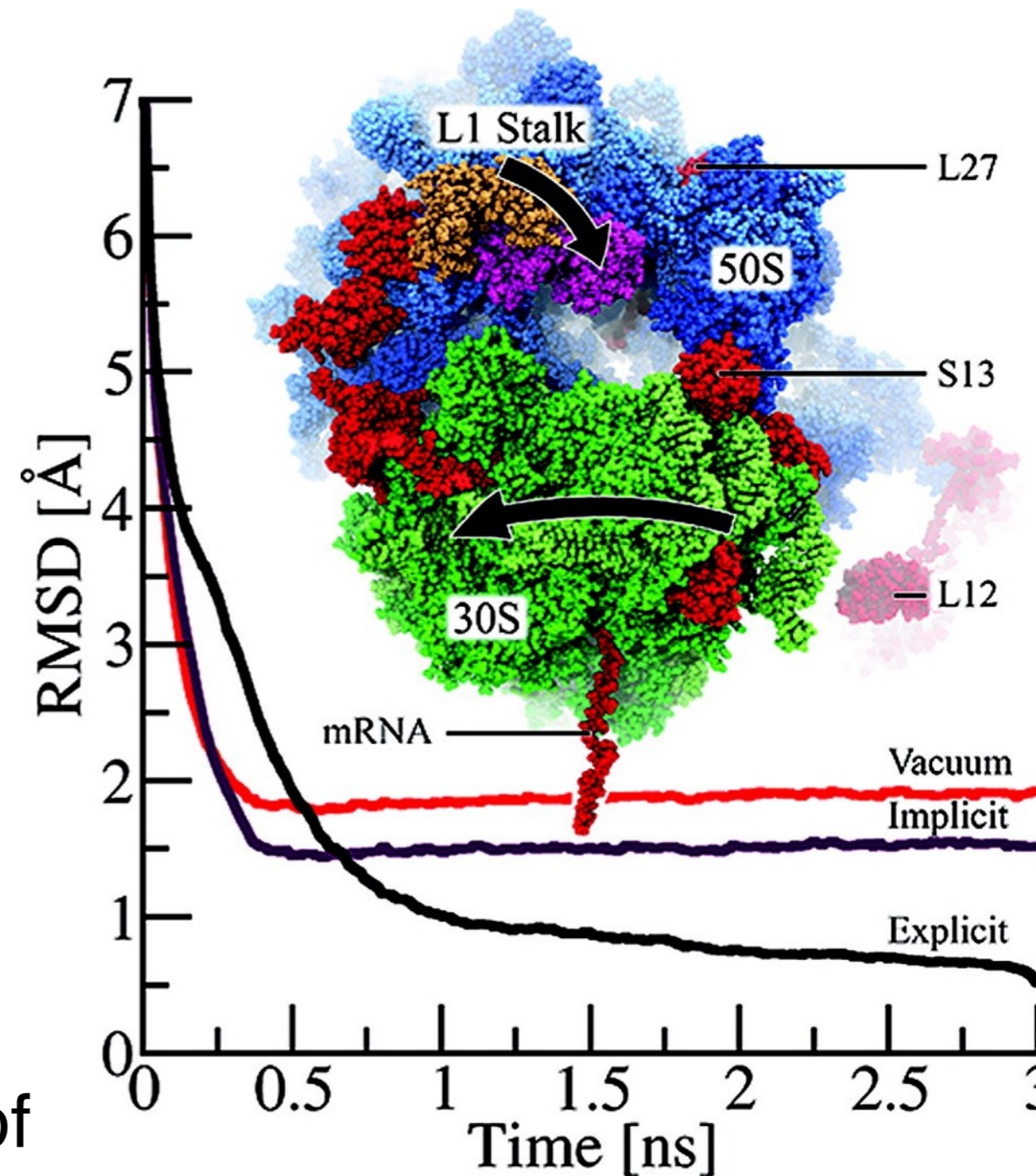
- No water molecules
- Fastest but potentially inaccurate

## 2. Explicit Solvent

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

## 3. Implicit Solvent

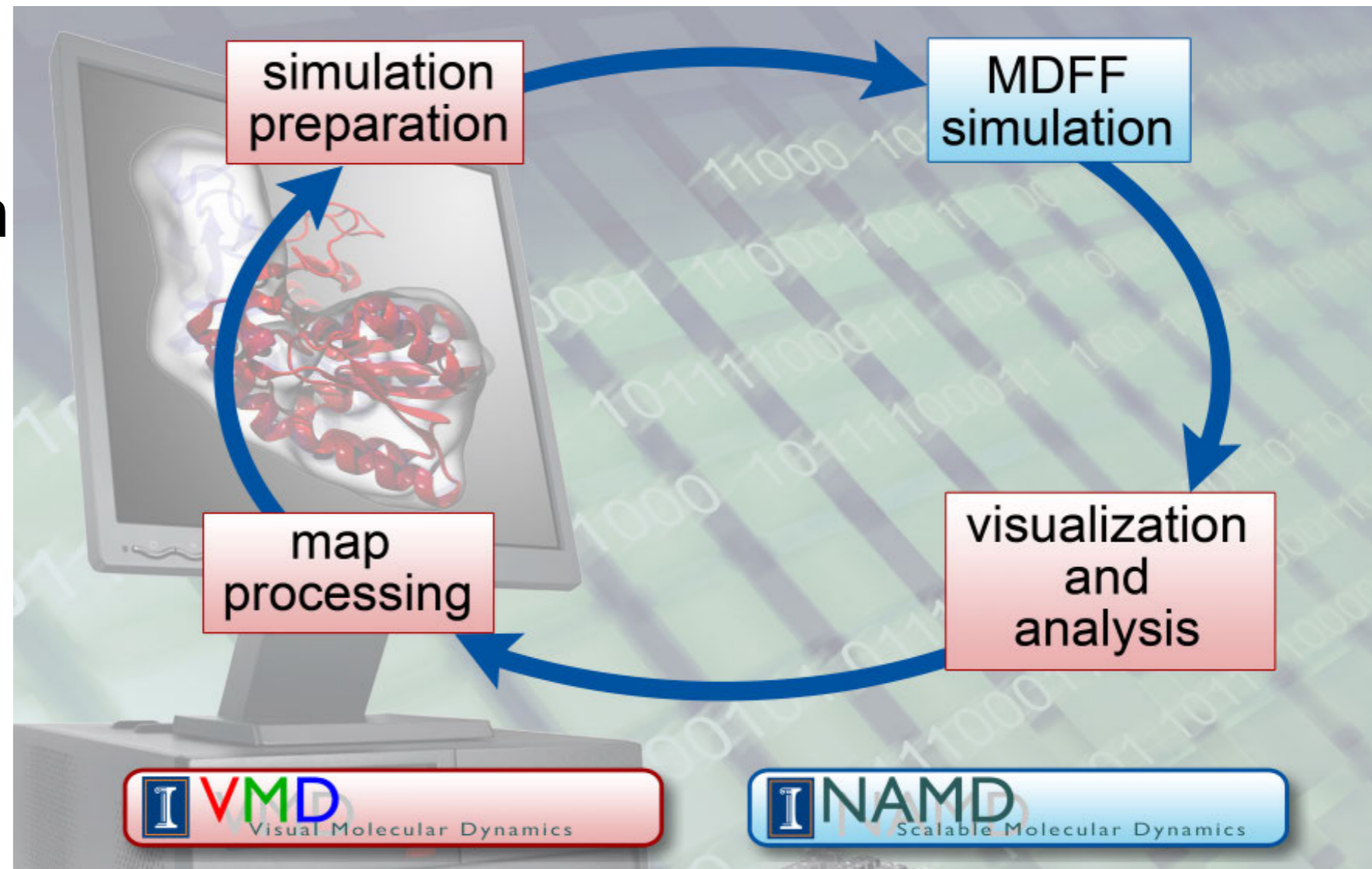
- generalized Born approximation of electrostatics
- Compromise between speed and accuracy



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  $\sim 1$  ns or less (much shorter than MD)

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

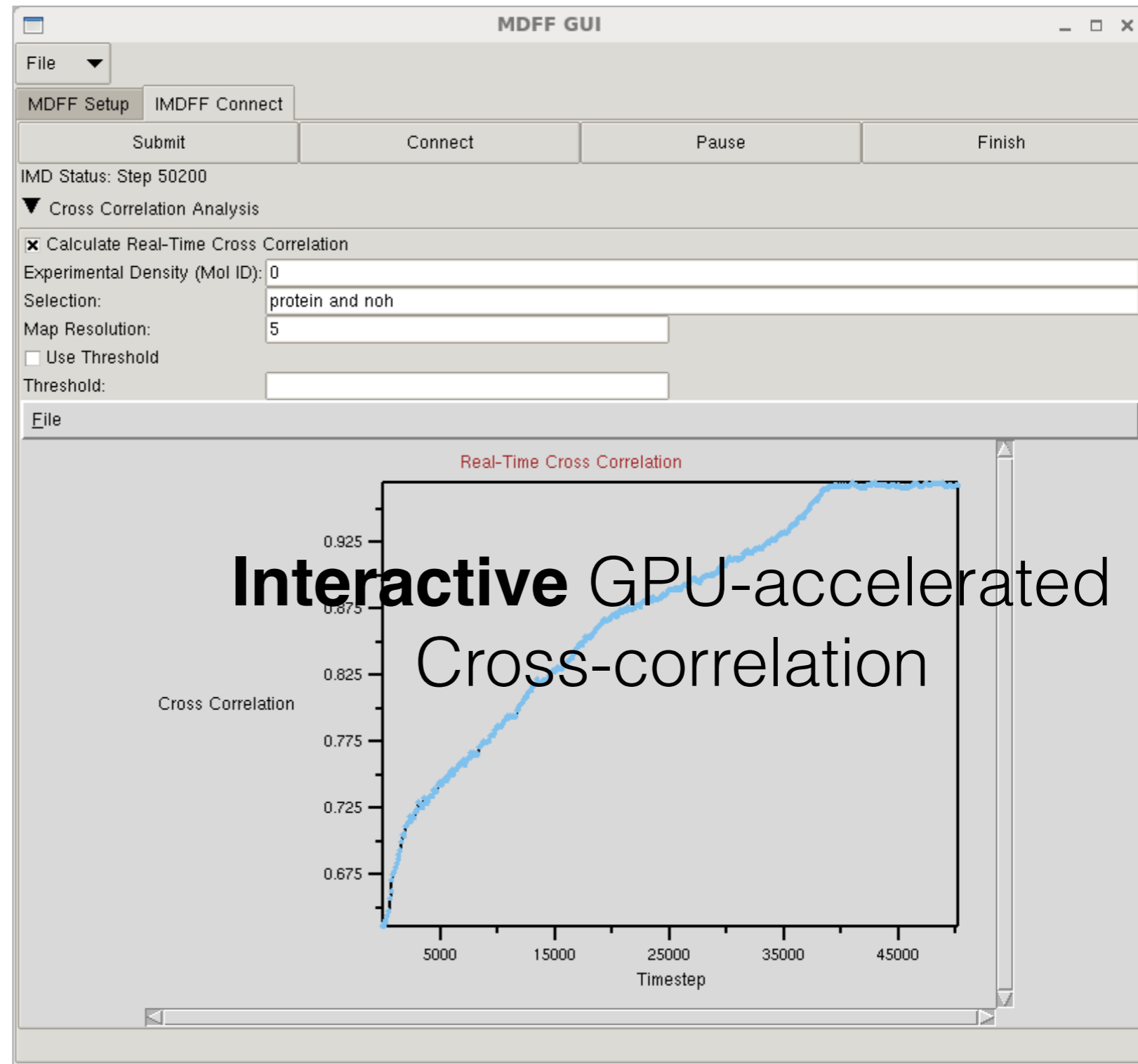
**Input:** MDFF only requires a PDB, PSF, and density map

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

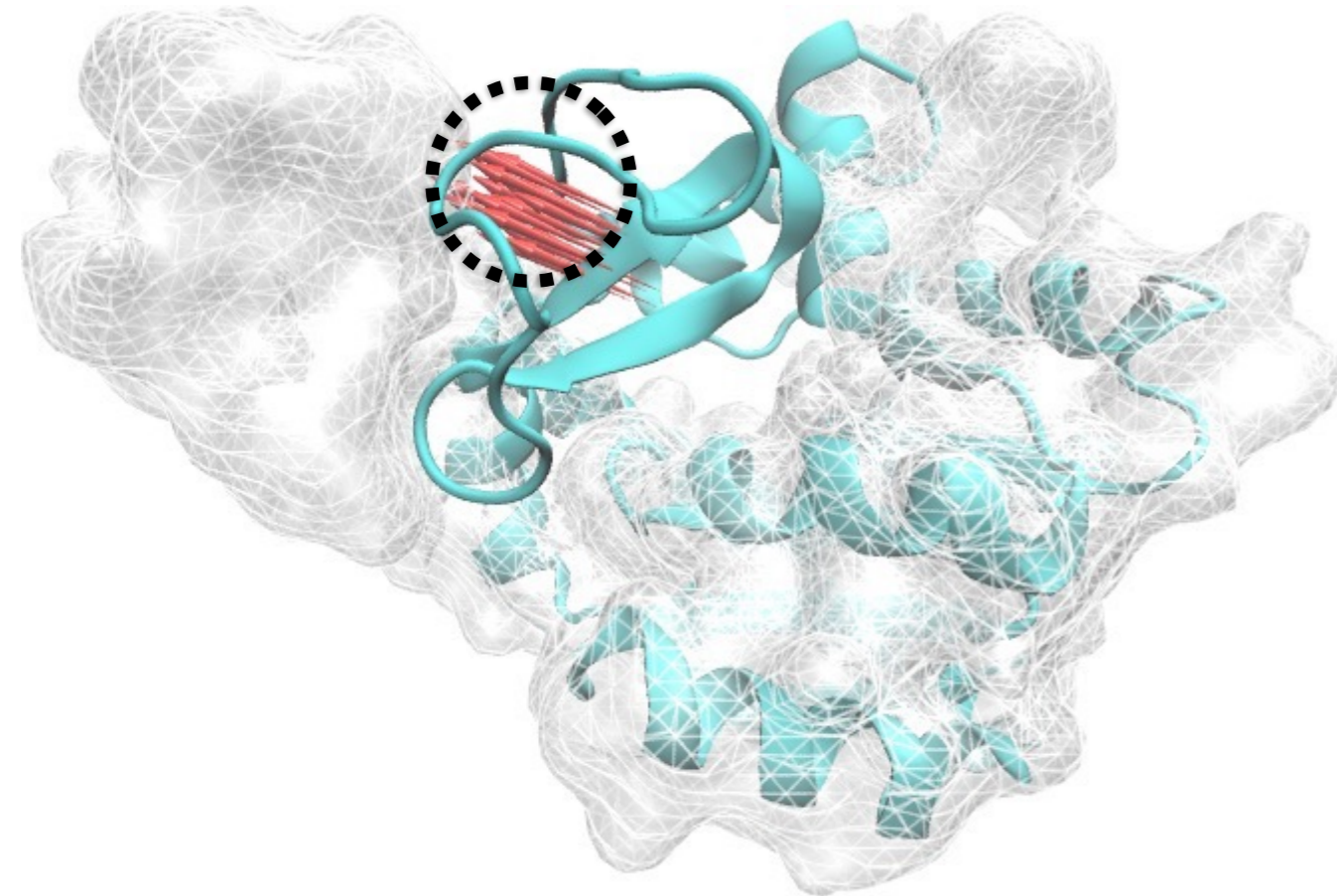
New MDFF GUI (VMD 1.9.3) makes setting up, running, and analyzing fitting simulations even easier





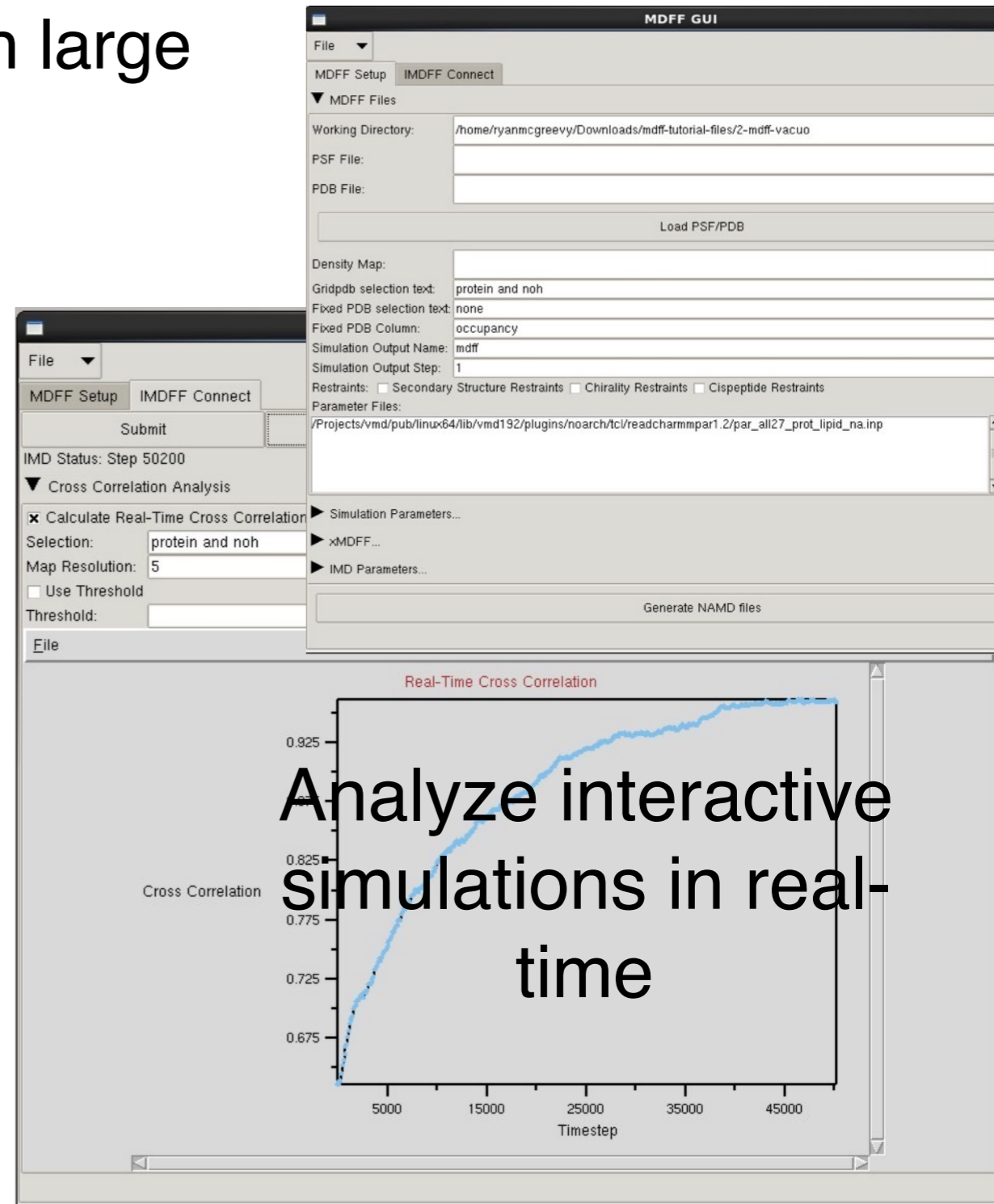
# Interactive Modeling with MDFF GUI

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

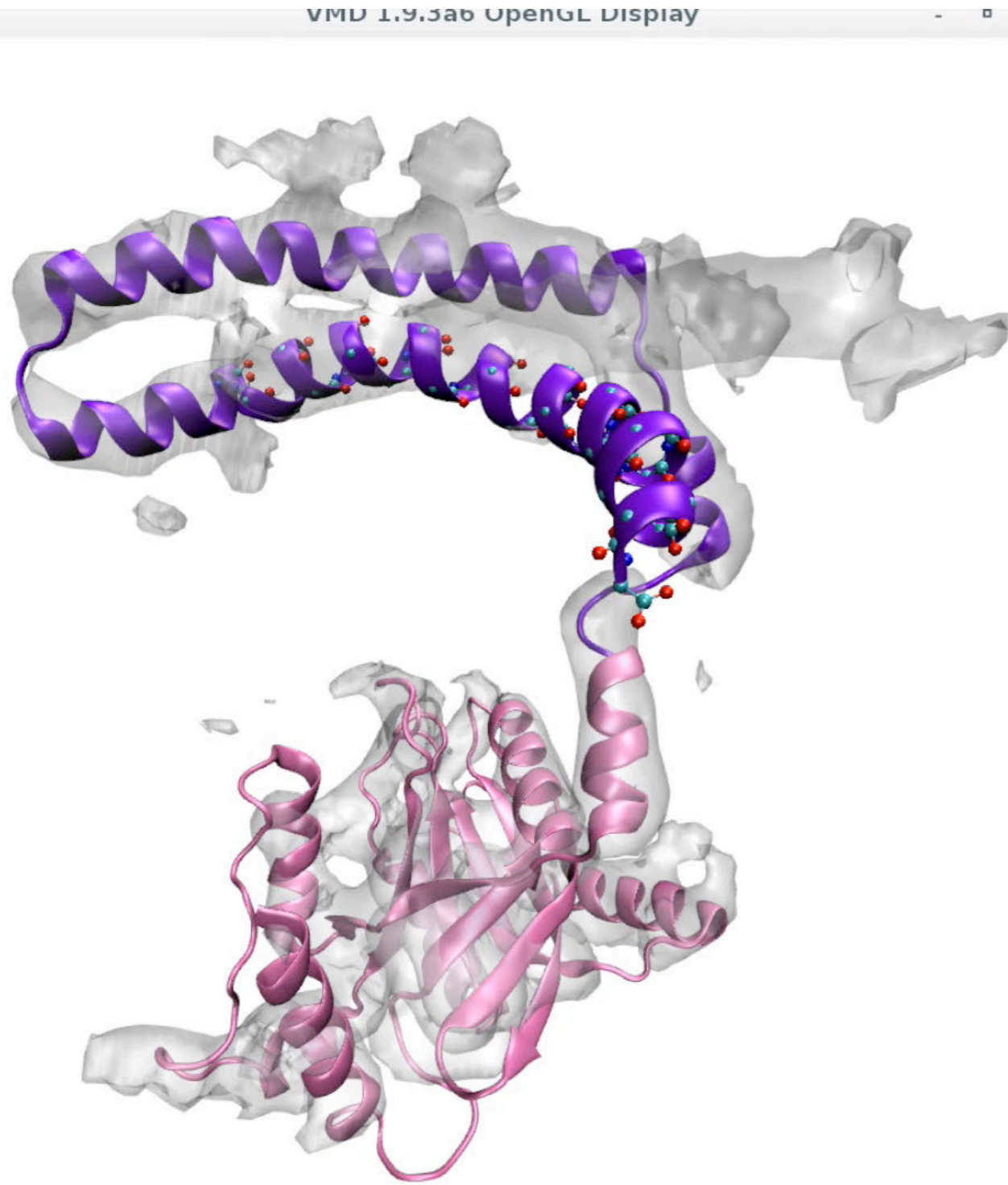


New MDFF GUI in VMD 1.9.3

Set up and run interactive (or traditional) MDFF/xMDFF simulations



# Interactive Modeling Integrates User Expertise



VMD Main

ID	T	A	D	F	Molecule	Atoms	Frames	Vol
0	T	A	D	F	start_1.psf	4822	5	0
1	A	D	F		Rpn11_2_2594_density.0	0	0	1

MDFF GUI

File

MDFF Setup | IMDF Connect

Submit | Connect | Pause | Finish

IMD Status: Step 400

▼ Cross Correlation Analysis

Calculate Real-Time Cross Correlation

Experimental Density (Mol ID): 1

Selection: protein and noh

Map Resolution: 7.7

Use Threshold

Threshold:

File

Real-Time Cross Correlation

Cross Correlation

Timestep

The MDFF GUI window displays a graph of Cross Correlation versus Timestep. The y-axis ranges from 0.696175 to 0.696575, and the x-axis ranges from 150 to 400. The graph shows a blue line with several peaks and troughs, indicating fluctuations in cross-correlation over time. The line starts at approximately 0.696175 at timestep 150 and ends at approximately 0.696575 at timestep 400.

# Analyzing MDFF Model Quality: **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)

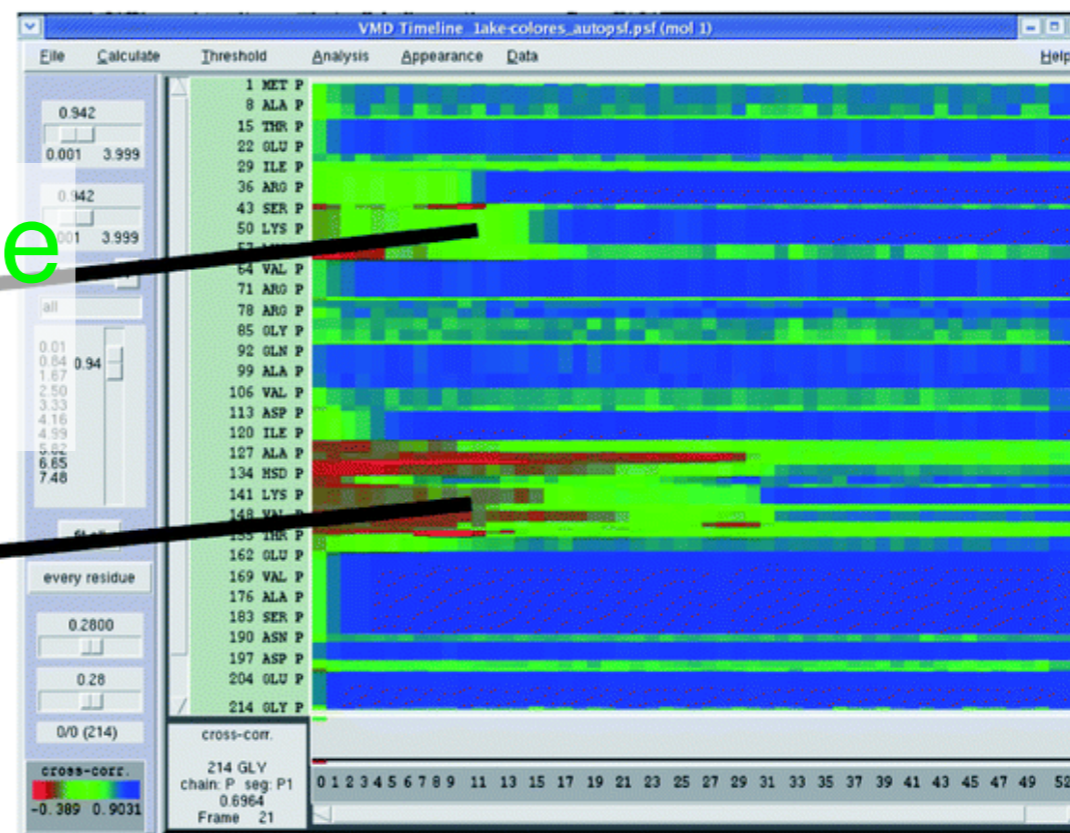
Good Fit

Intermediate Fit

(b)

Bad Fit

(c)



# MDFFF on the Cloud Costs Less than a Cup of Coffee

VMD, NAMD, MDFFF now available on Amazon Cloud

Focus on the scientific challenges of your project without having to worry about local availability and administration of suitable computer hardware and installing or compiling software

Acetyl-CoA Synthase (PDB 1OAO) 11469 atoms, 6 Parallel Replicas



Instance Type	CPU	Performance (ns/day)	Time (hours)	Simulation Cost (\$)
c3.8xlarge	30	5.88	0.41	1.68
c3.4xlarge	12	3.33	0.72	0.84
c3.2xlarge	6	1.35	1.78	0.84



Singharoy, *et al.* eLife 2016

Easy, 1-click launch for fast access to MDFFF on HPC hardware

The screenshot shows the AWS Marketplace interface for the product 'NAMDA, VMD, and MDFFF'. The header includes the AWS Marketplace logo, a search bar, and navigation links. The product title is 'NAMDA, VMD, and MDFFF' with a sub-header 'Sold by: TCBG'. The description states: 'VMD is designed for modeling, visualization, and analysis of biological systems such as proteins, nucleic acids, lipid bilayer assemblies, etc. It may be used to view more general molecules, as VMD can read standard Protein Data Bank (PDB) files and display the contained structure. VMD provides a wide variety of methods for rendering and coloring a molecule: simple points and lines, CPK spheres and cylinders, licorice bonds, backbone tubes and ribbons, cartoon drawings, and others. VMD can be used to animate and analyze the trajectory of a molecular dynamics (MD) simulation. In particular, ... Read more'. Below the description are sections for 'Customer Rating' (5 stars), 'Latest Version' (0.1.0), 'Operating System' (Linux/Unix, Ubuntu Ubuntu Server 16.04 LTS), 'Delivery Method' (64-bit Amazon Machine Image (AMI) [Read more]), 'Support' (See details below), 'AWS Services Required' (Amazon EC2, Amazon EBS), and 'Highlights' (Highly Scalable). On the right, there is a 'Continue' button and a 'Pricing Information' section with a region dropdown menu set to 'US East (N. Virginia)' and a 'Free Tier Eligible' badge.

# 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:**

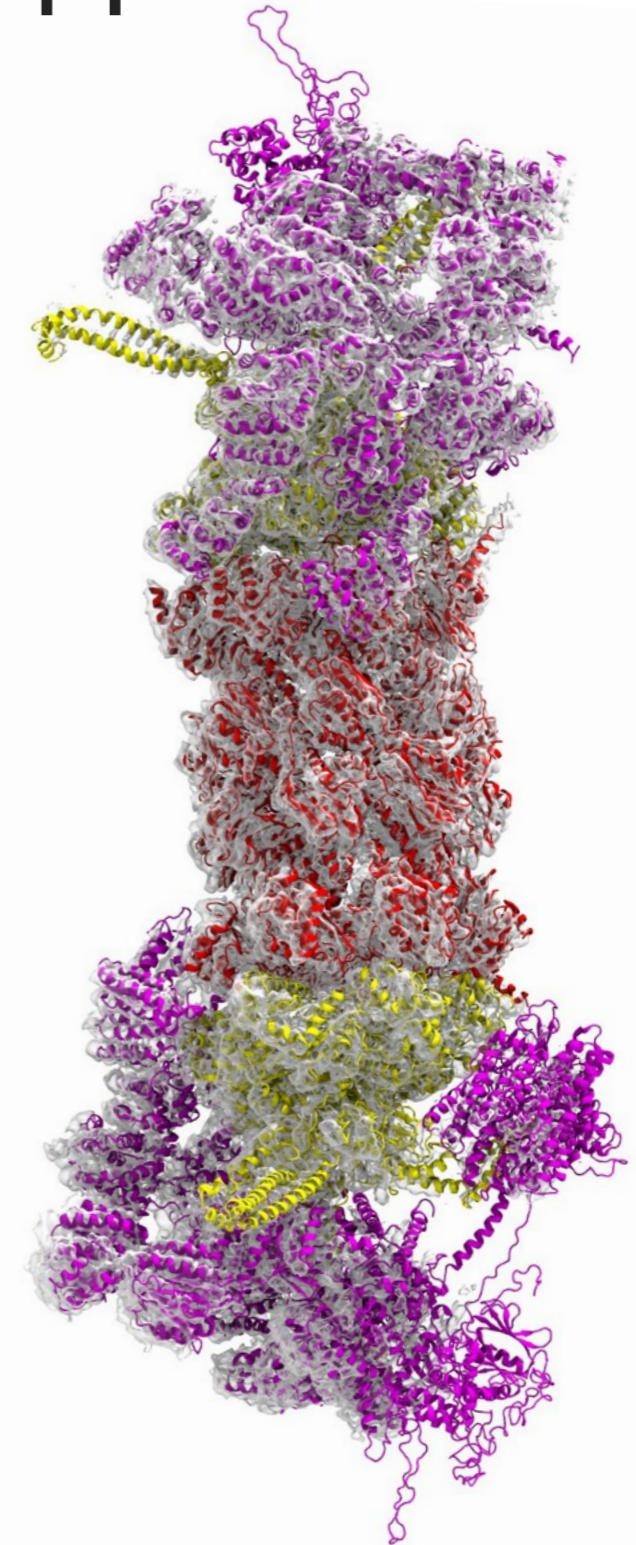
Singharoy\*, Teo\*, McGreevy\*, et al. *eLife* (2016)

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

# Further Information

Find out more about MDFF including:

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

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