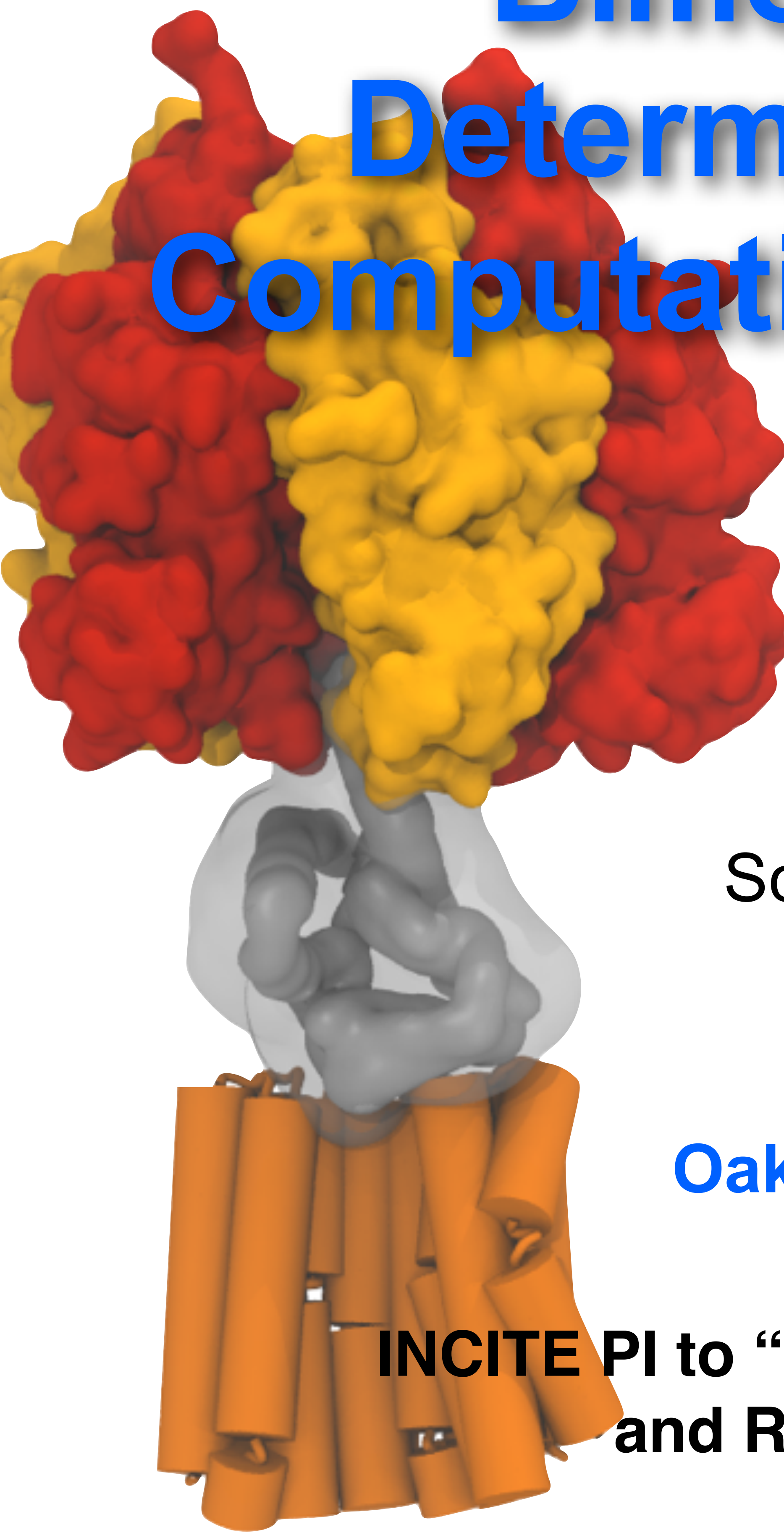
Bimolecular Structure Determination with NAMD: Computational Cryo-EM on Titan



May 16th, 2018

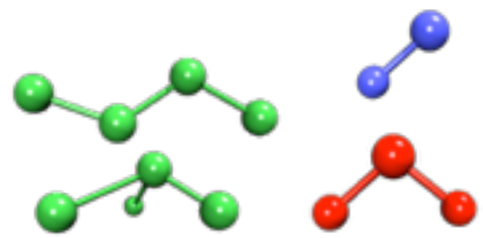
Abhishek Singharoy
School of Molecular Sciences
Arizona State University

OLCF Users Meeting
Oak Ridge National Laboratory

INCITE PI to "All-atom Simulations of Photosynthetic and Respiratory Energy Conversion"

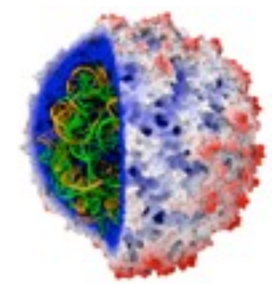
Biomolecules in Action: Using Titan as a "Computational Microscope" with NAMD

Chemistry

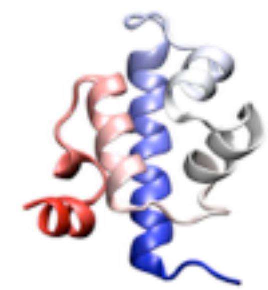


$$U(\vec{R}) = \underbrace{\sum_{bonds} k_i^{bond}(r_i - r_0)^2}_{U_{bond}} + \underbrace{\sum_{angles} k_i^{angle}(\theta_i - \theta_0)^2}_{U_{angle}} + \underbrace{\sum_{dihedrals} k_i^{dihe}[1 + \cos(n_i\phi_i + \delta_i)]}_{U_{dihedral}} + \underbrace{\sum_i \sum_{j \neq i} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \sum_i \sum_{j \neq i} \frac{q_i q_j}{\epsilon r_{ij}}}_{U_{nonbond}}$$

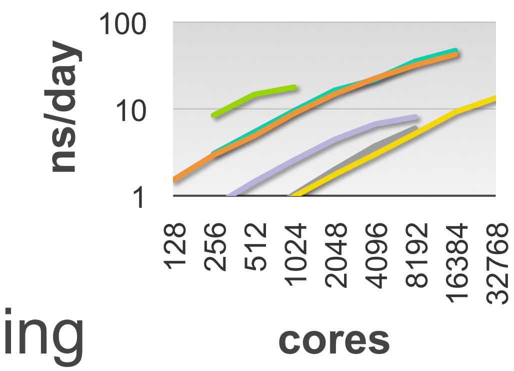
NAMD Software



Virus



Protein Folding



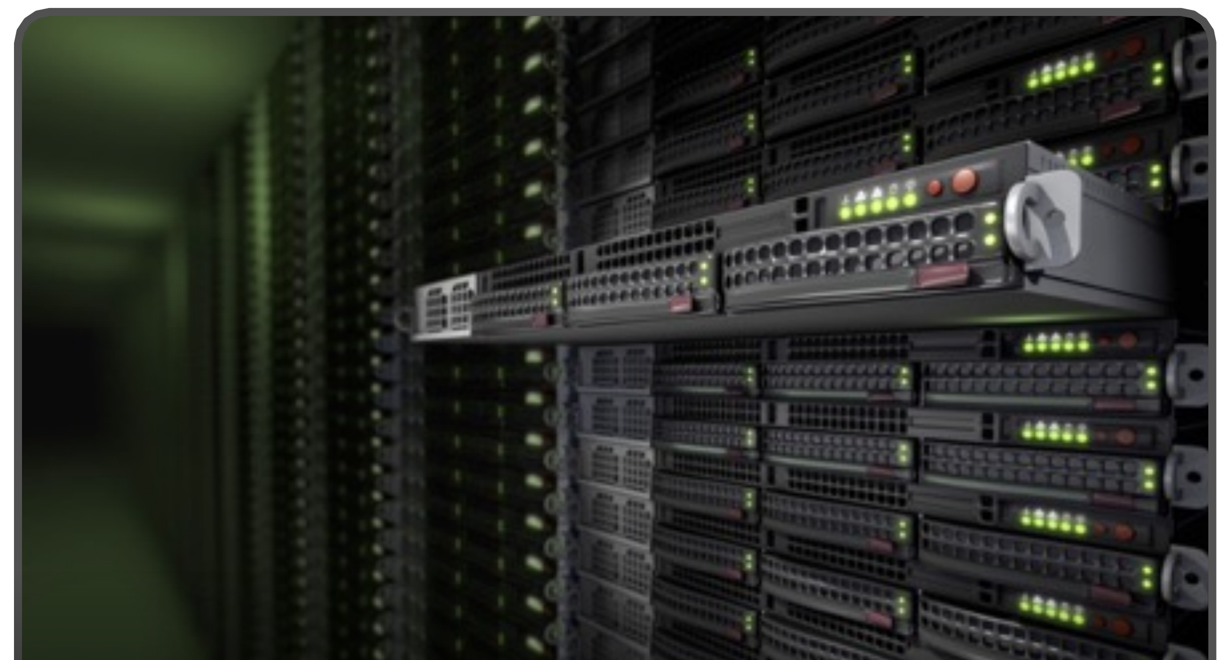
Physics

$$m_i \frac{d^2\vec{r}_i}{dt^2} = \vec{F}_i = -\vec{\nabla} U(\vec{R})$$

Math

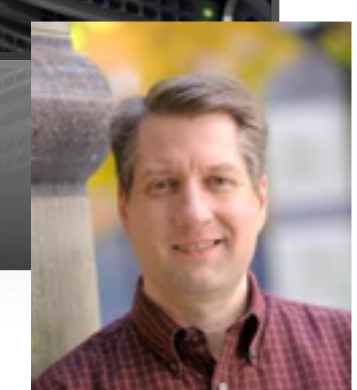
$$\vec{r}_i(t + \Delta t) = 2\vec{r}_i(t) - \vec{r}_i(t - \Delta t) + \frac{\Delta t^2}{m_i}\vec{F}_i(t)$$

(repeat ***one billion times*** = microsecond)



..and Supercomputers

David Hardy

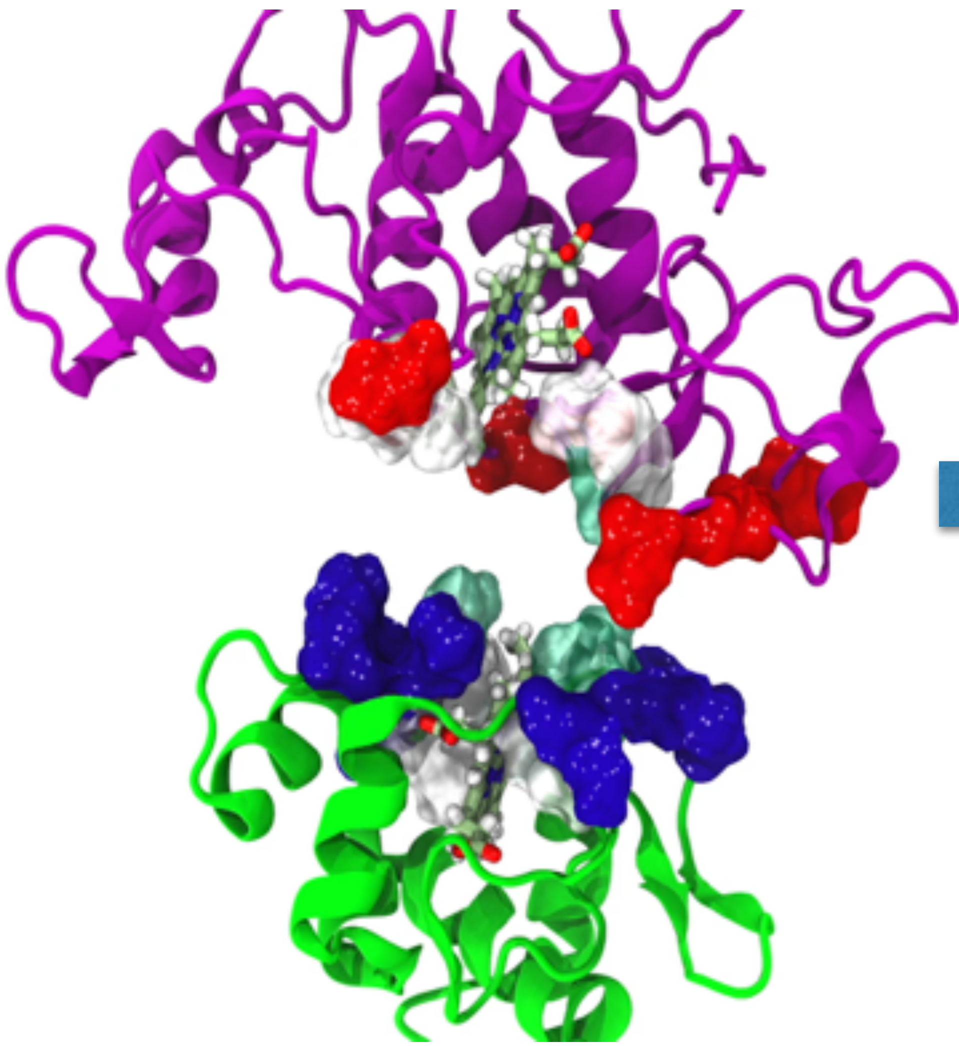


Jim Phillips

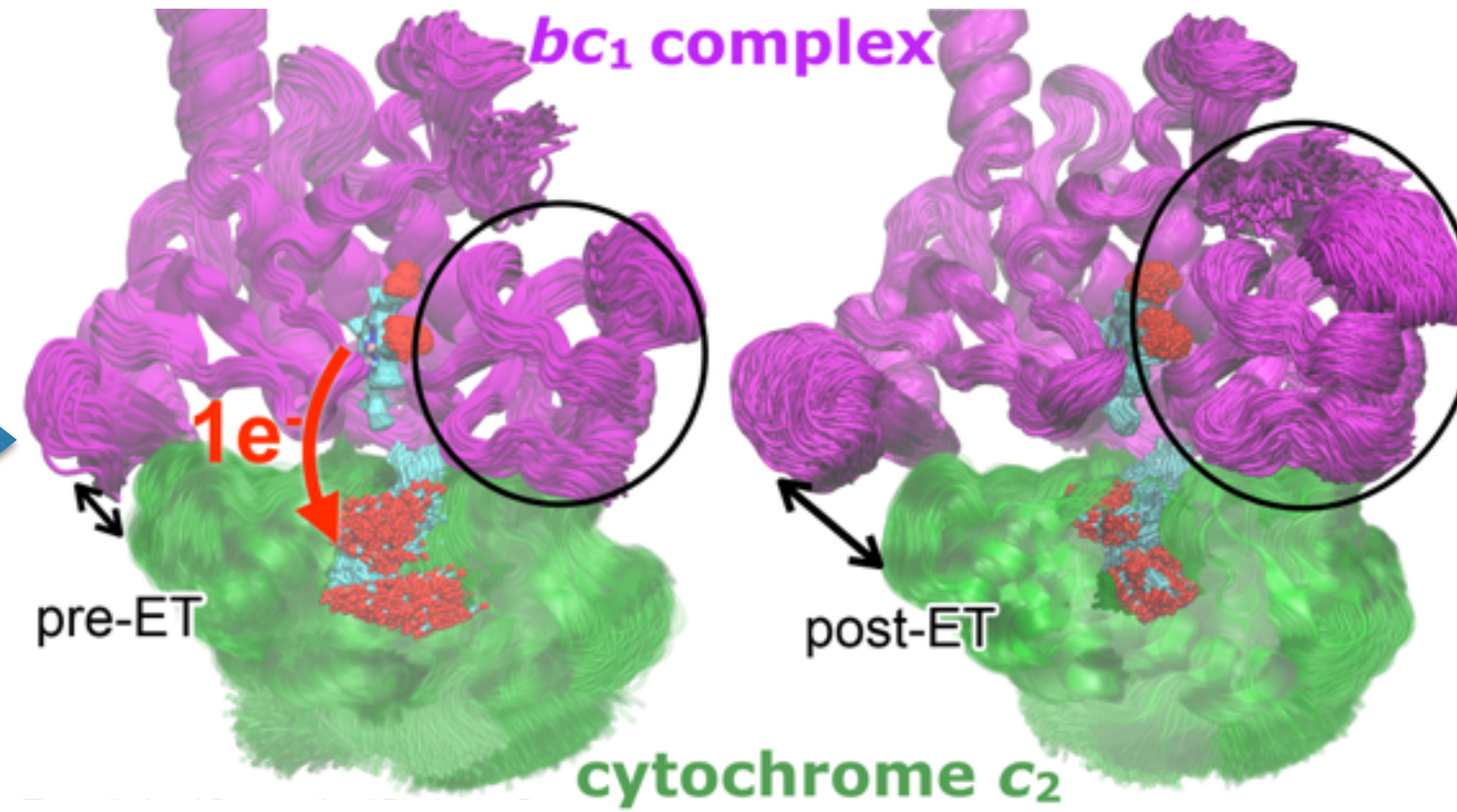
Molecular Dynamics (MD) simulations

Why Does One Need a Supercomputer ?

Structural transitions



Ensemble of trajectories

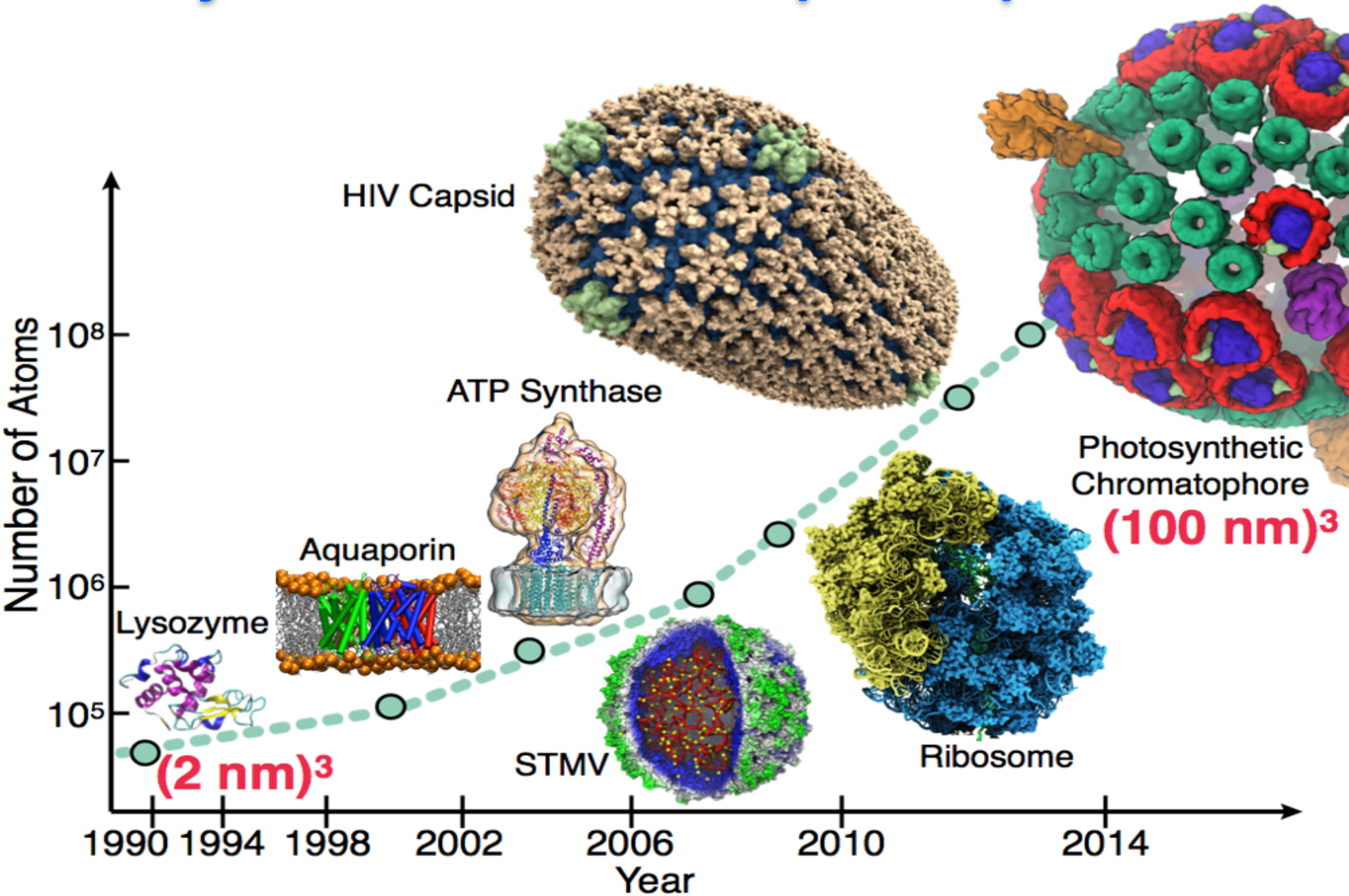


MD_1 MD_2 MD_3 .... MD_{n-1} MD_n

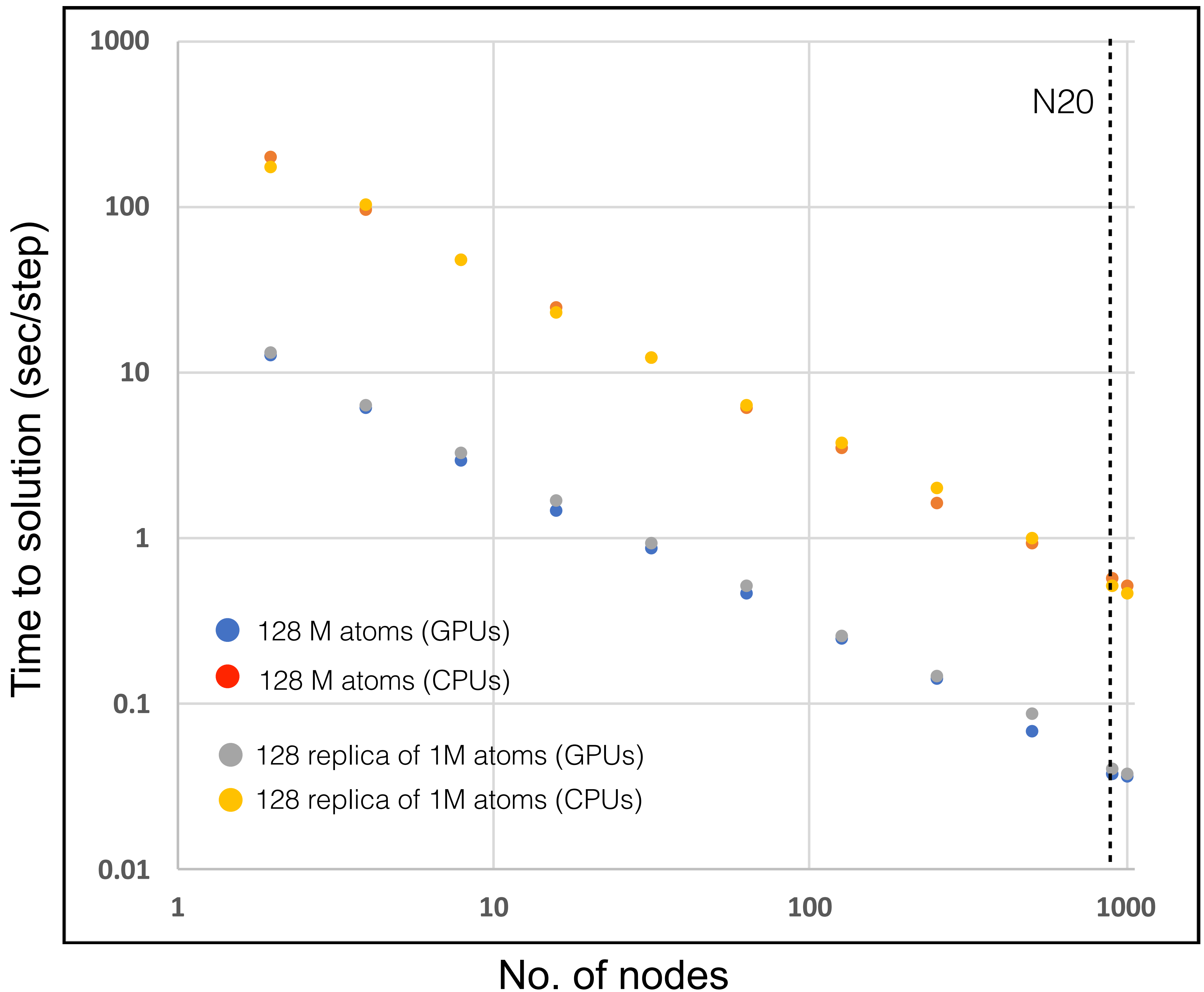
Multiple replica required !!

J.Am.Chem.Soc 138, 12077 (2016)

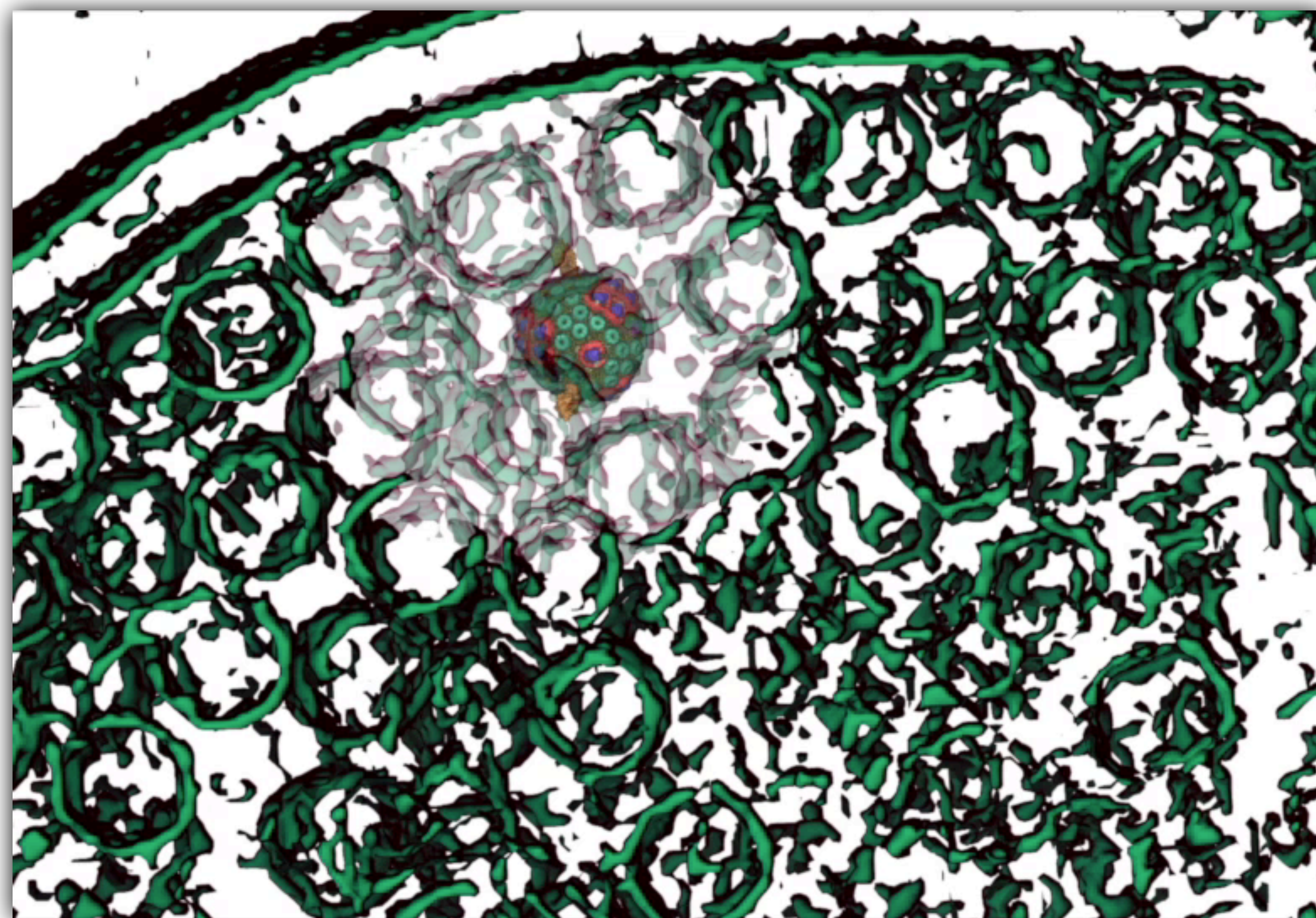
Why Does One Need a Supercomputer ?

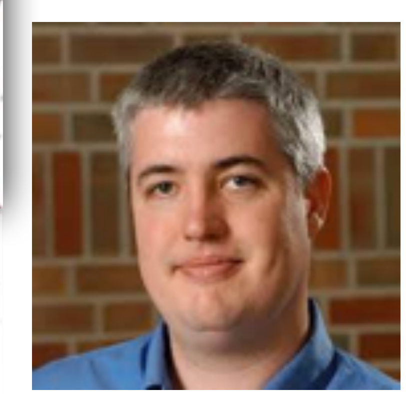


Parallel Performance of NAMD on Summit



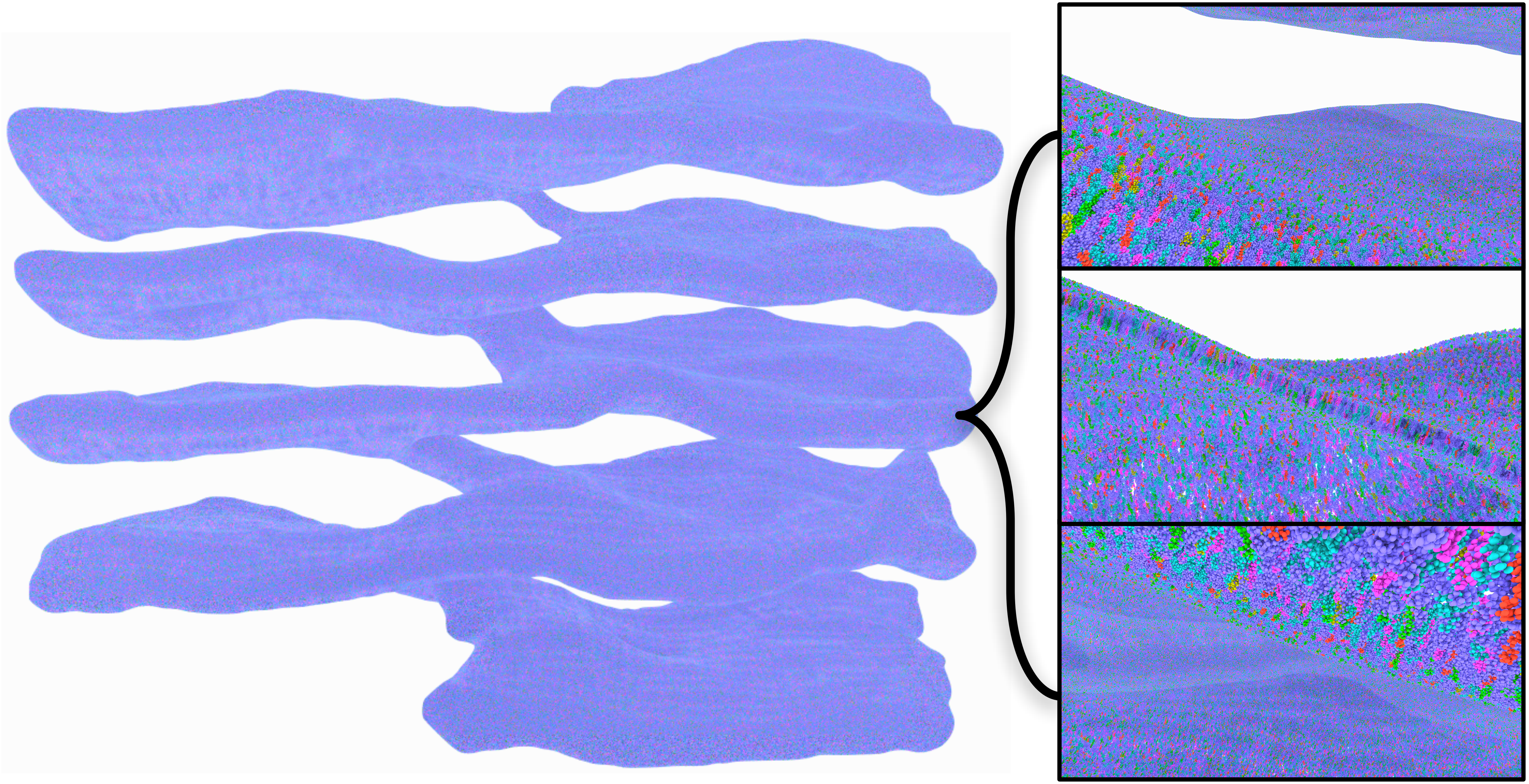
Visualization and Analysis: VMD





John Stone

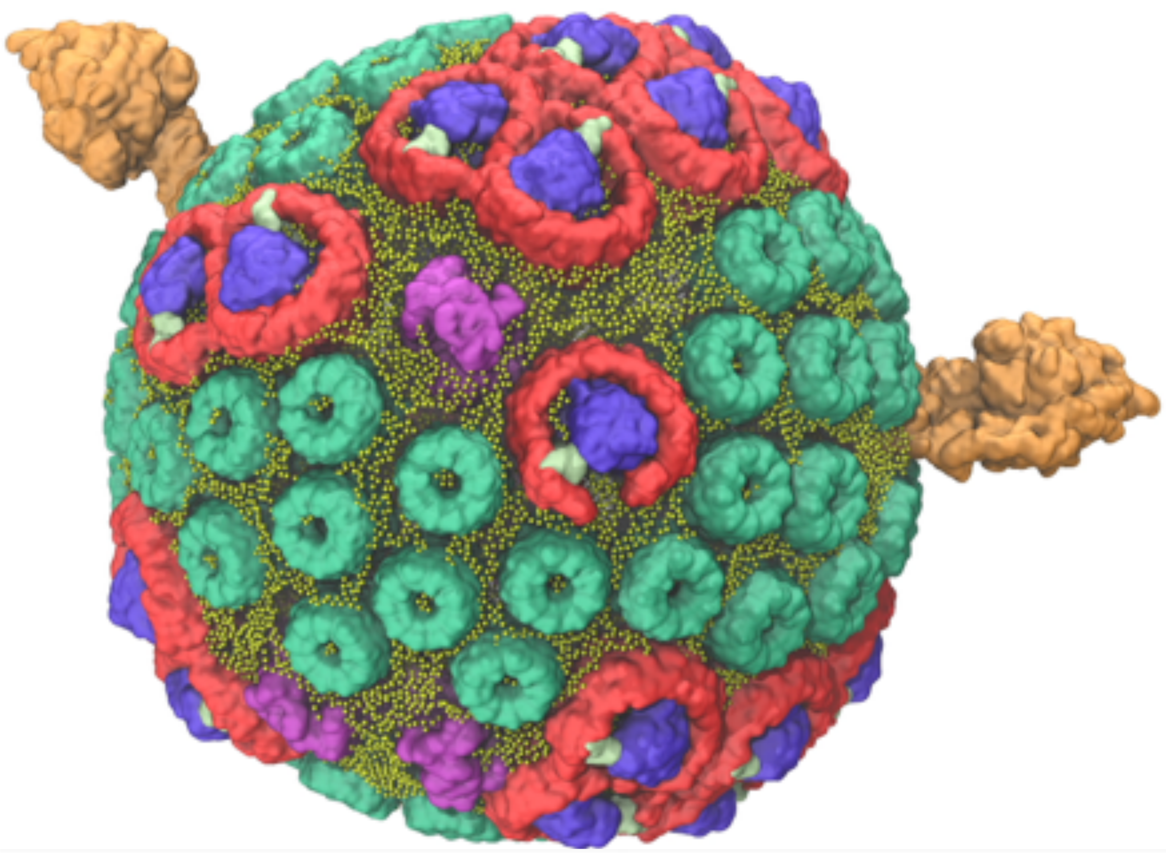
In-situ Visualization of Billion Atoms : SIGHT

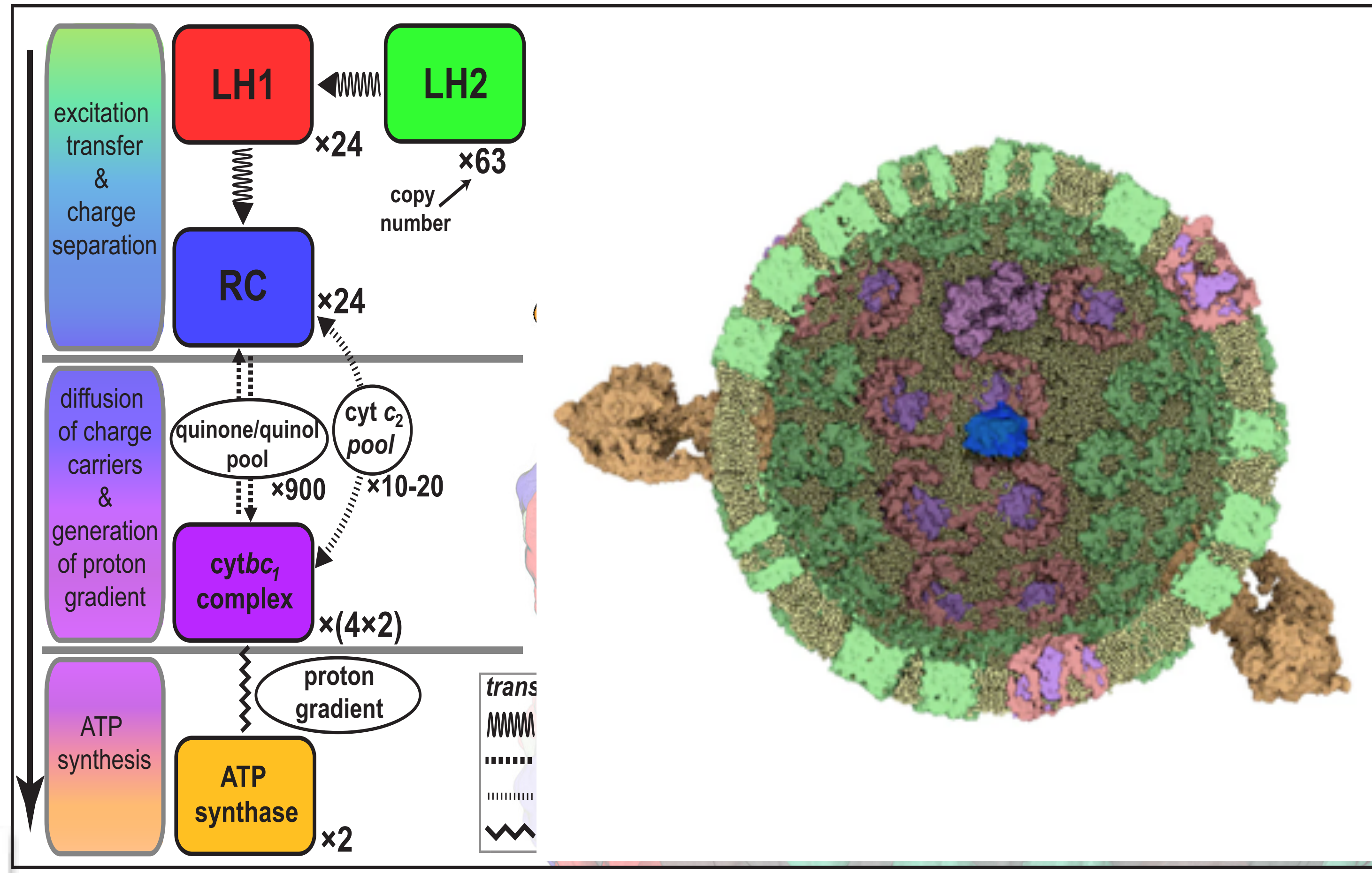


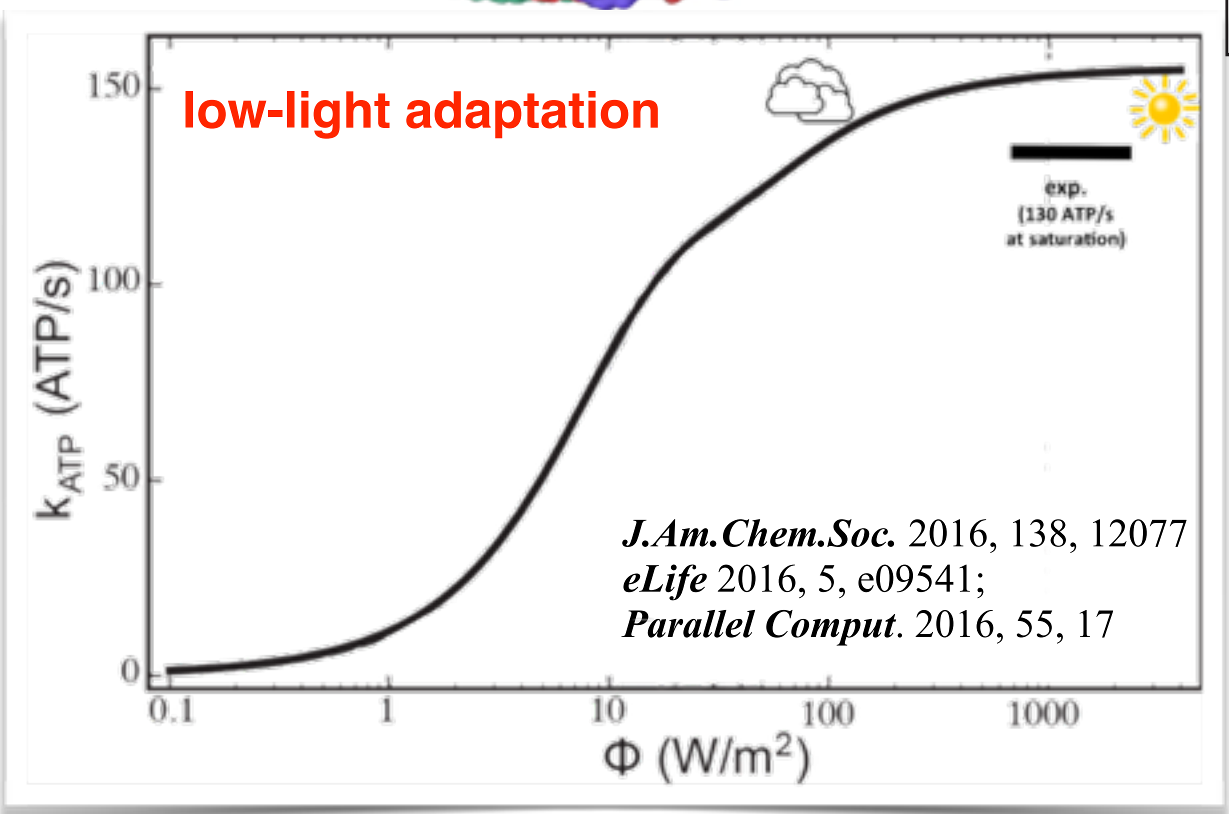
Noah Trebesch

Ben Hernandez (OLCF)

Scientific Accomplishment # 1: Energy Conversion in Bacterial Photosynthesis





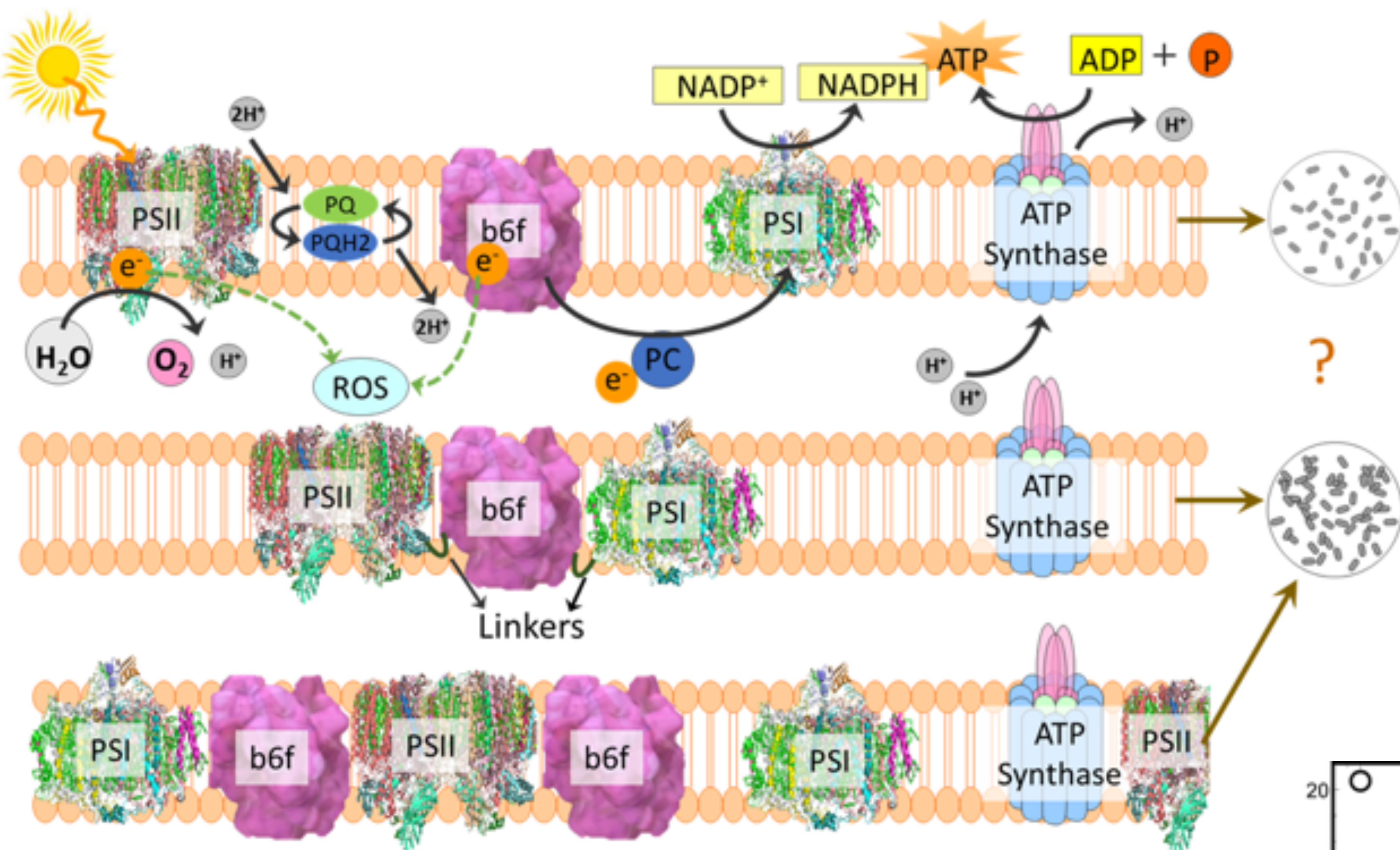


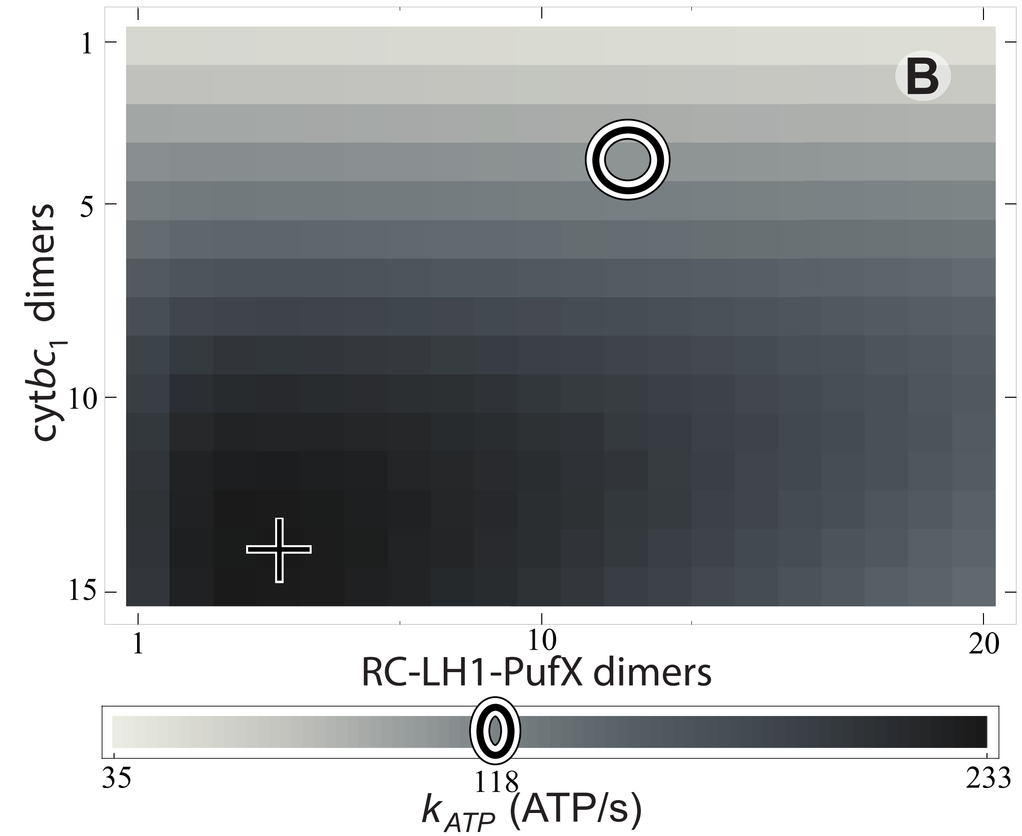
$$k_{ATP}(I) = \frac{1}{2} Iq \left(1 + \frac{1}{2} Iq \tau_{RC}(I) \frac{1}{n_{RC}}\right)$$

$$\tau_{RC}(I) = 1 + (\tau_H - 1)\left(1 - e^{\frac{Iq}{2B}}\right)$$

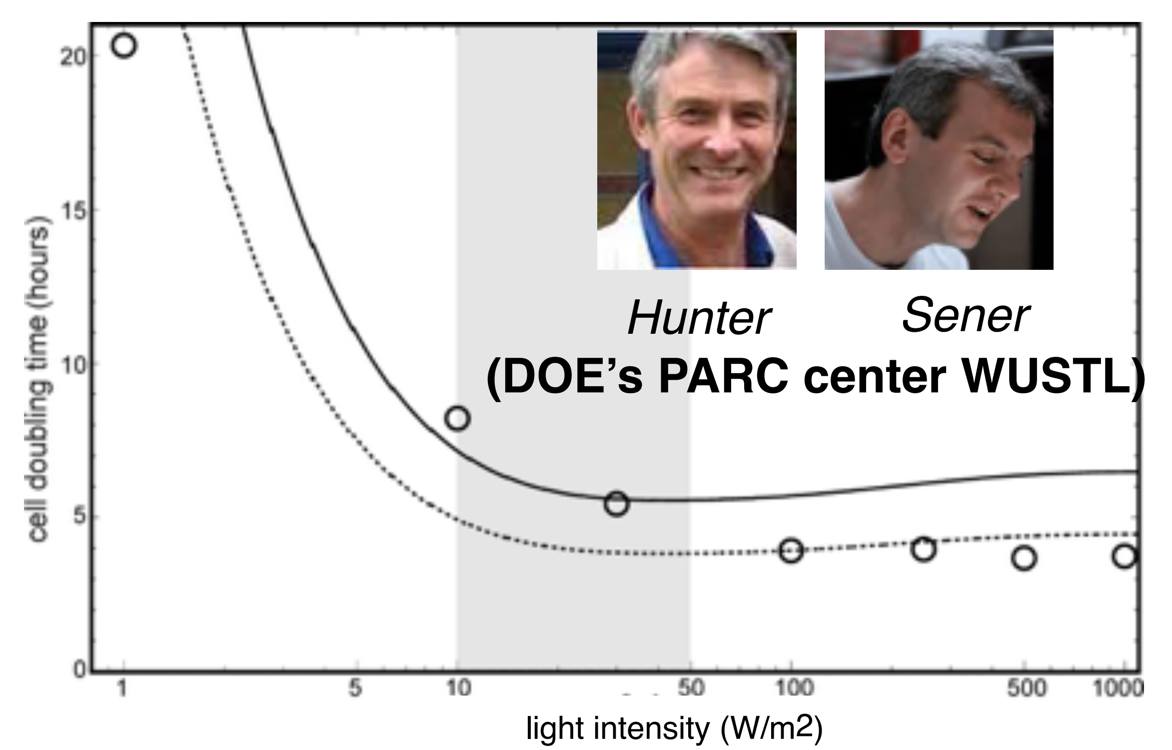
$$\tau_H = \frac{n_{RC}}{n_B}\tau_B;\ B = \frac{2n_B}{\tau_B}$$

Summit Goals 1: From First-Principles to Phenotypic Behaviors



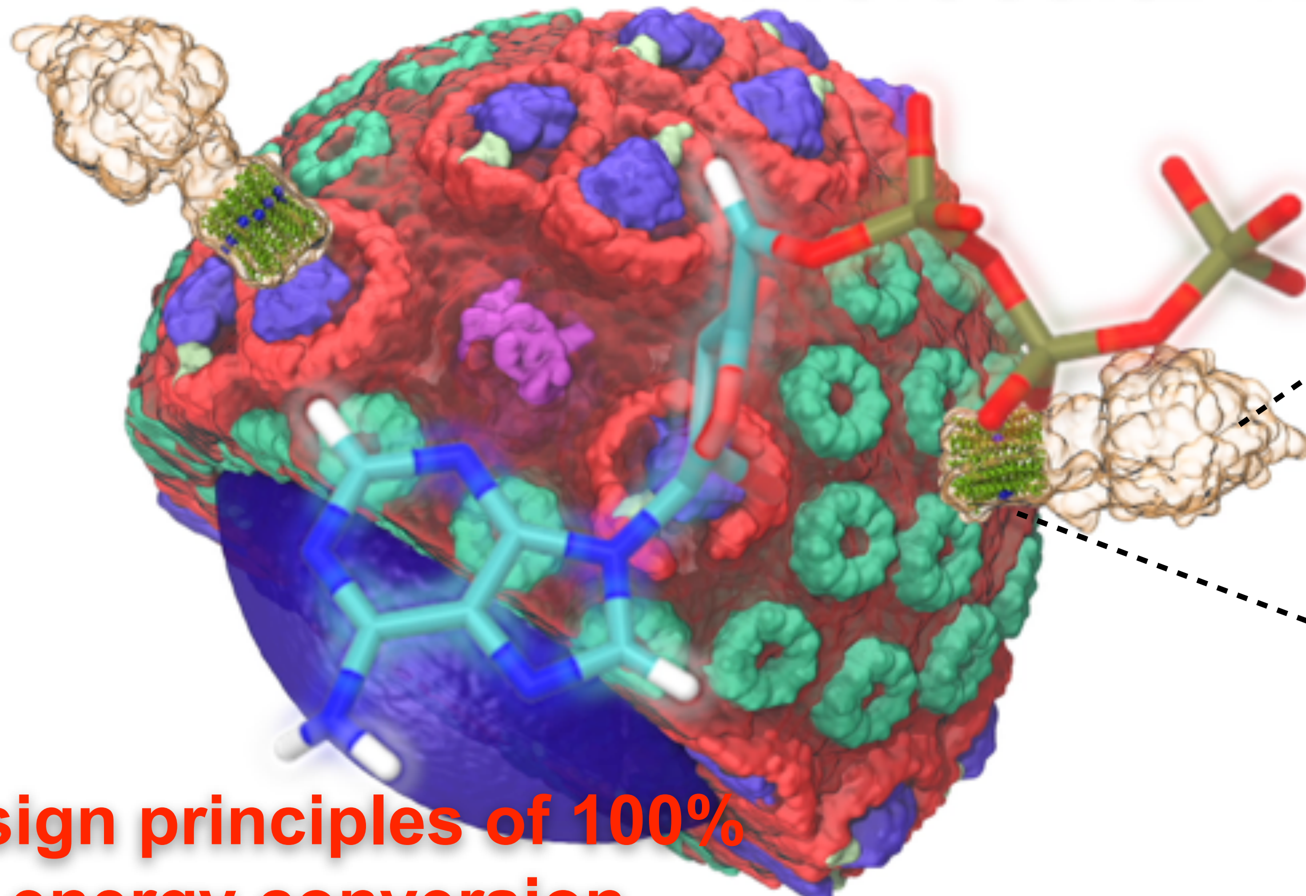


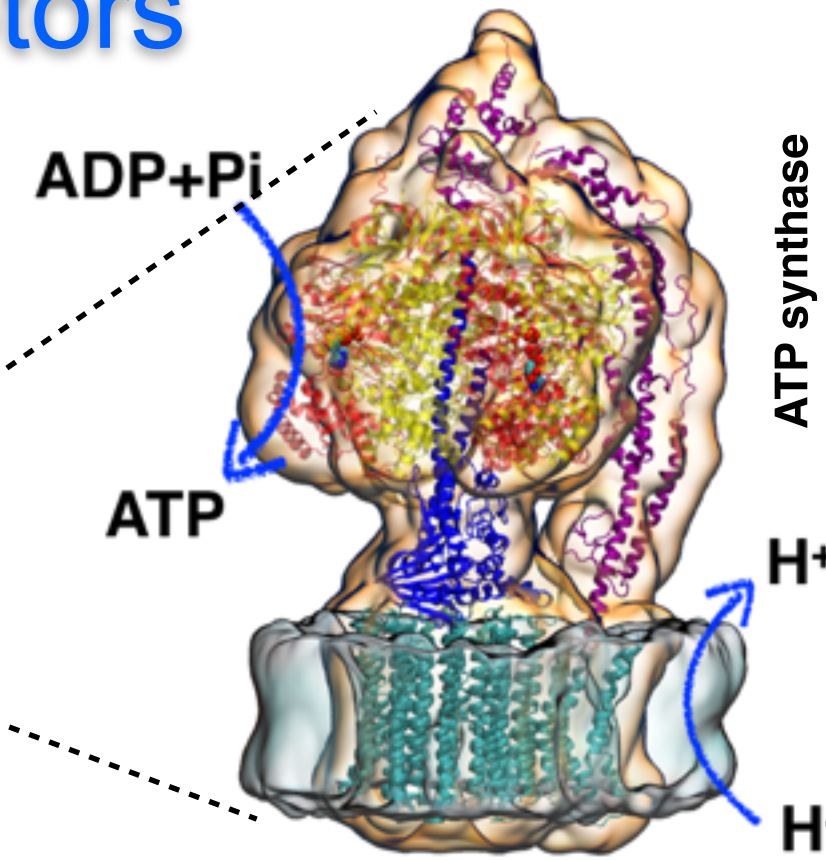
J.Phs.Chem.B **2017**, *121* , 3787–3797



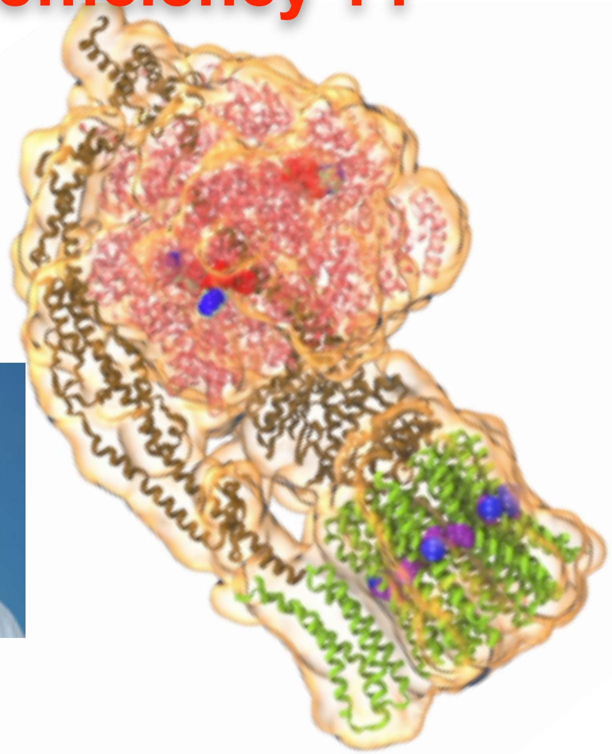
Can we make a **model for growth** from first principles ??

Summit Goals 2: Conformational Transition in Molecular Motors



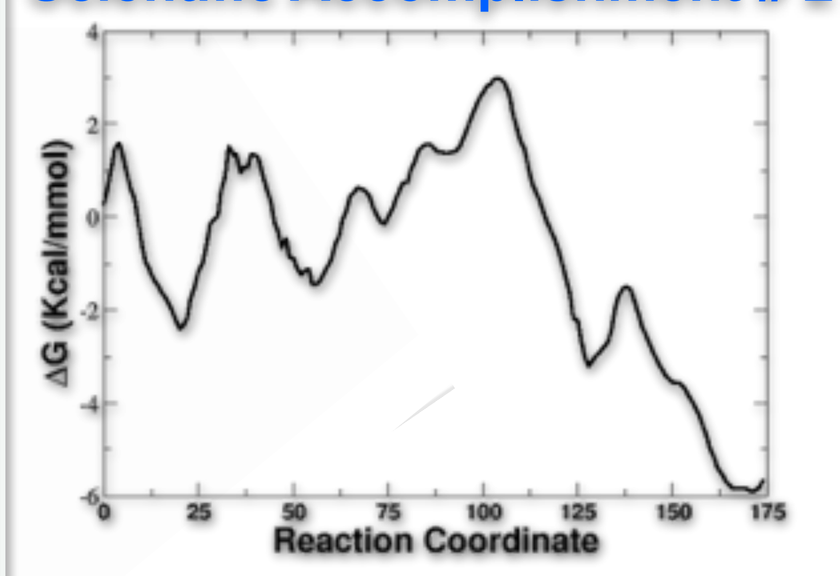


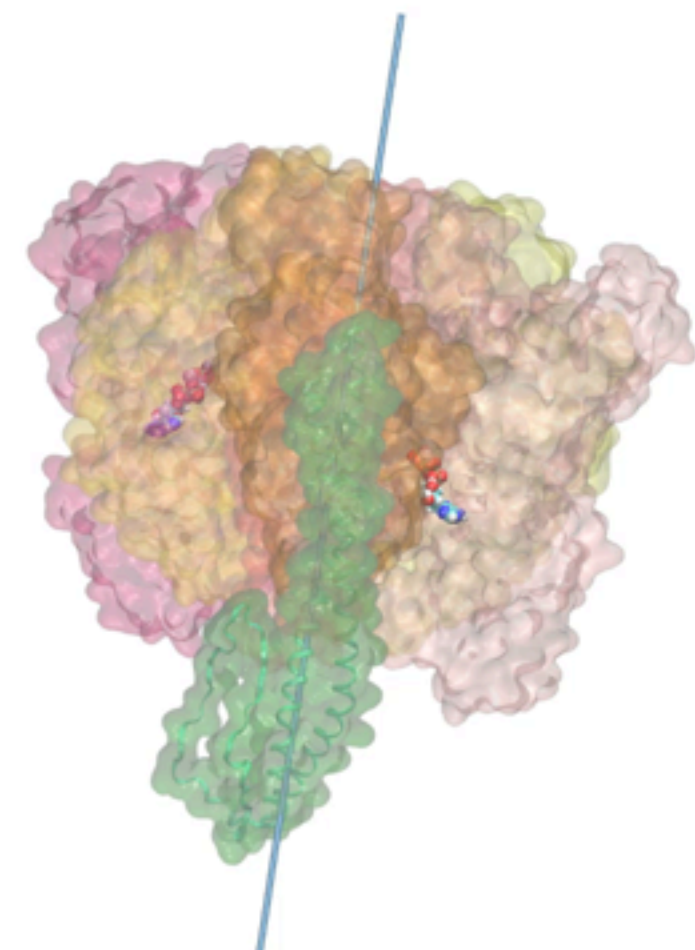
Design principles of 100% energy conversion efficiency ??



Wah Chiu
(Stanford)

Scientific Accomplishment # 2

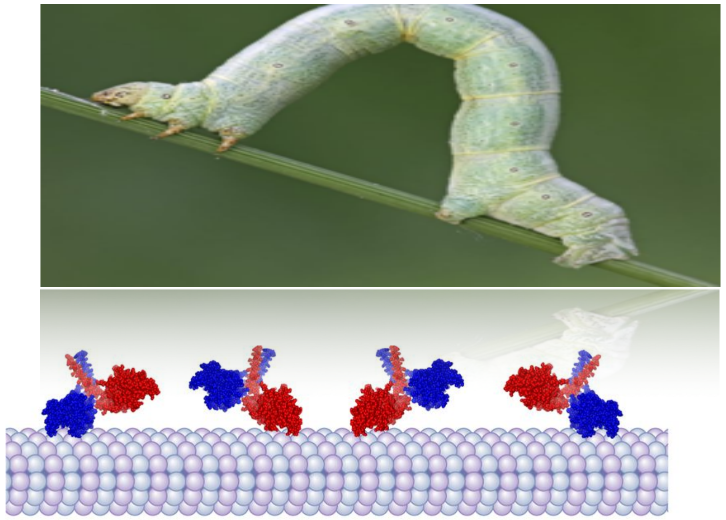




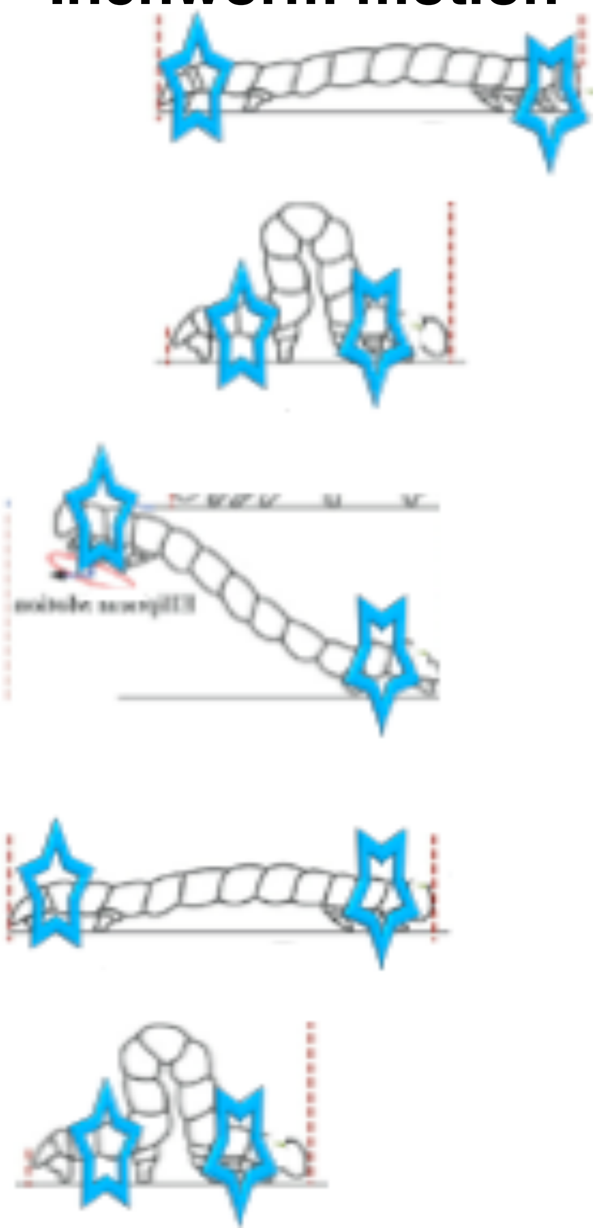
J.Am.Chem.Soc. 2017, 139, 293
J.Phys.Chem.B 2016, 121, 350

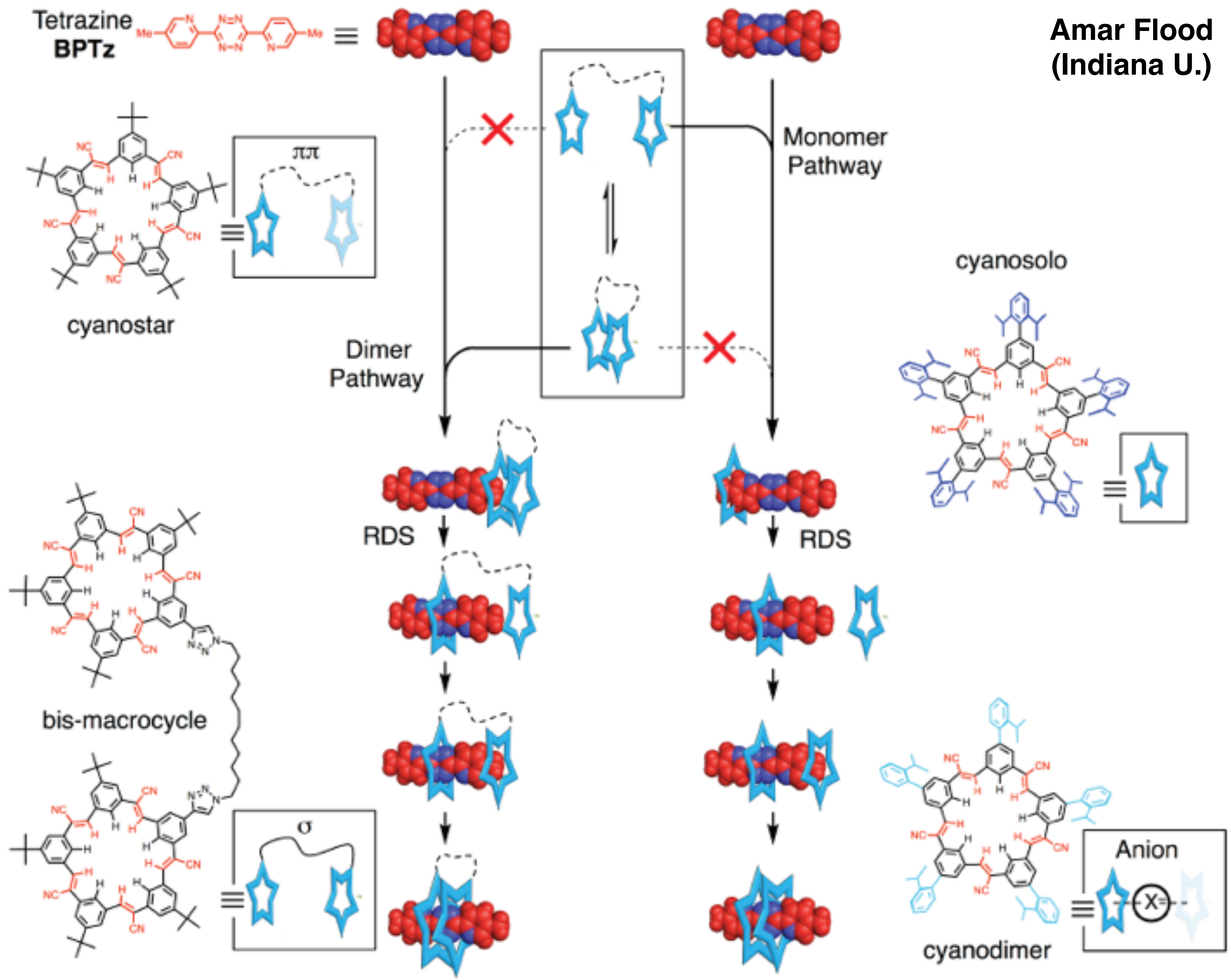
Scientific Accomplishment # 3: Synthesis of Artificial Motors

Amar Flood
(Indiana U.)



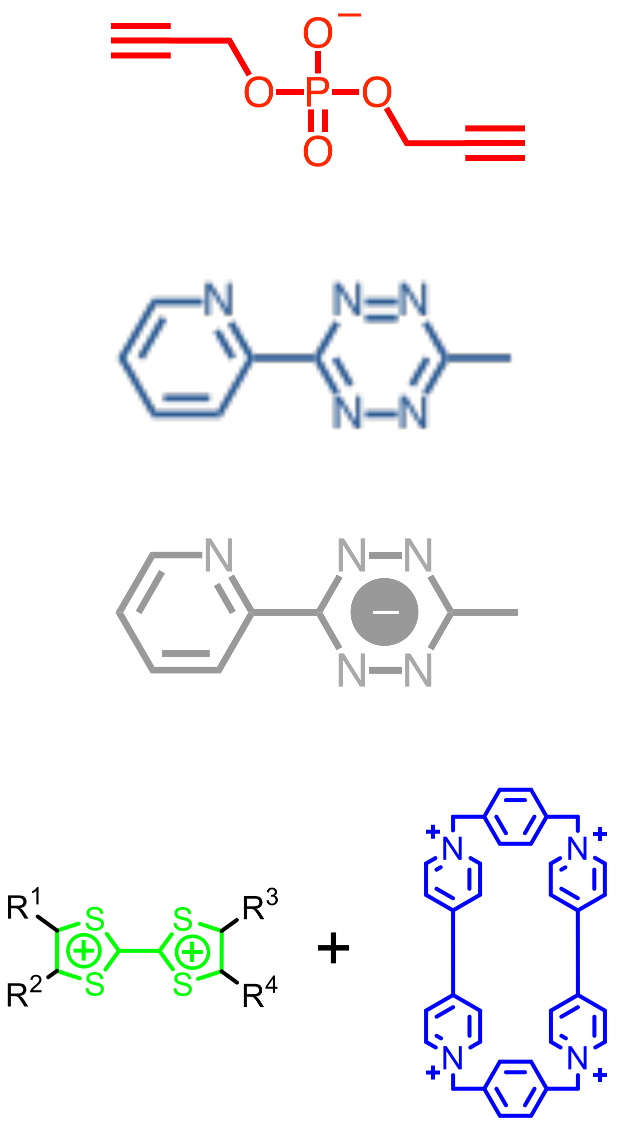
Inchworm motion





***Proc. Nat. Acad. Sci.* 2018:** https://doi.org/10.1073/pnas.1719539115
(Accepted for special issue commemorating the **2016 Nobel Prize in Chemistry**)

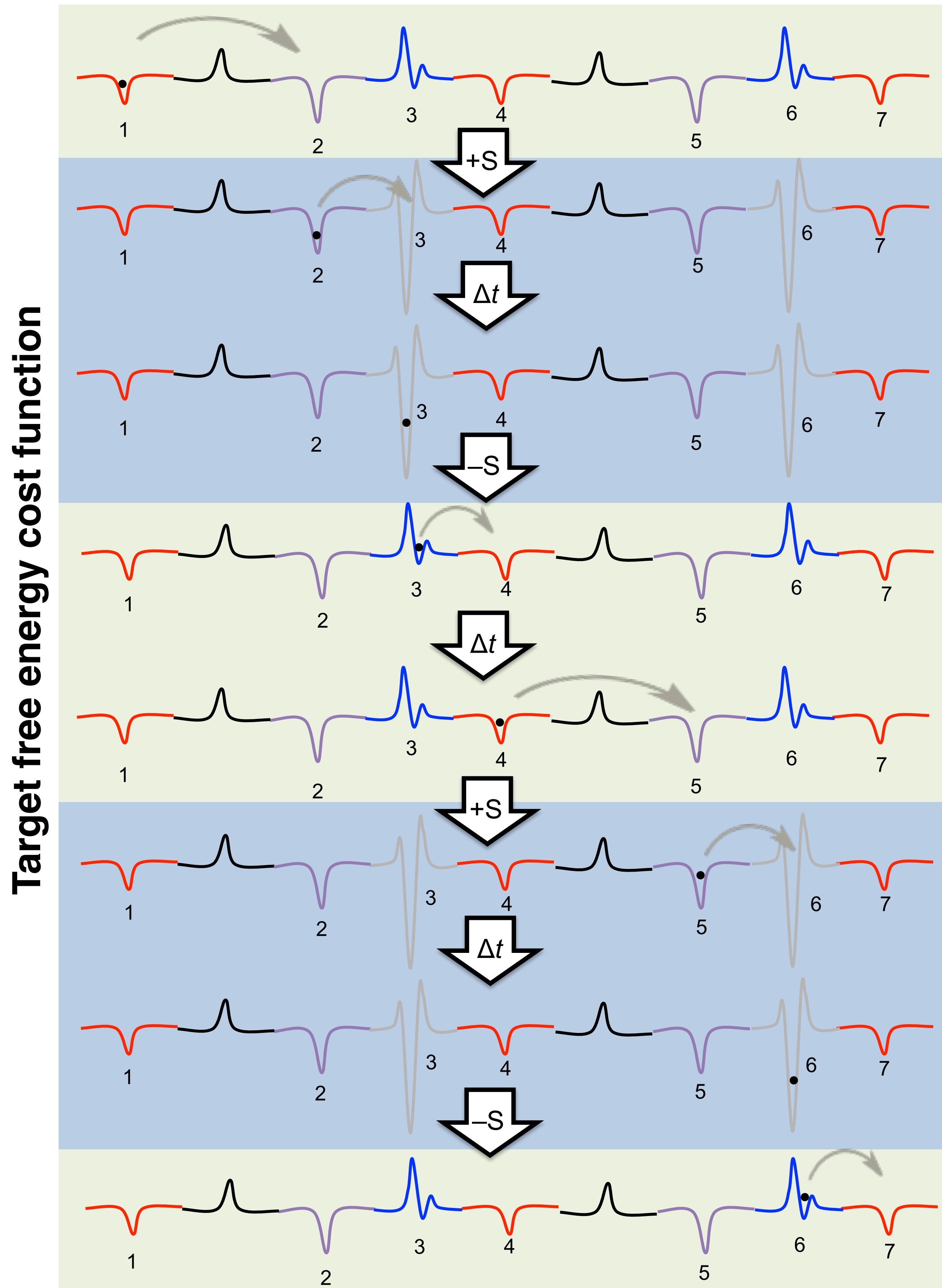
Scientific Challenge # 3: High-throughput design of Artificial Motors



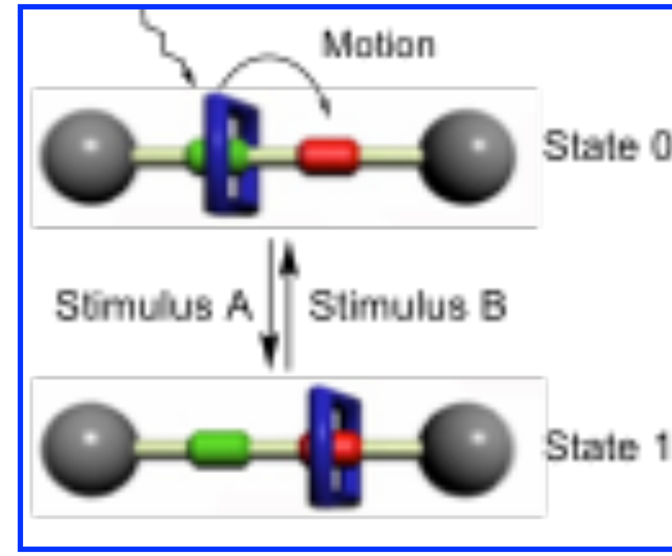
TTF^{2+} CBPQT^{4+}

Space of components

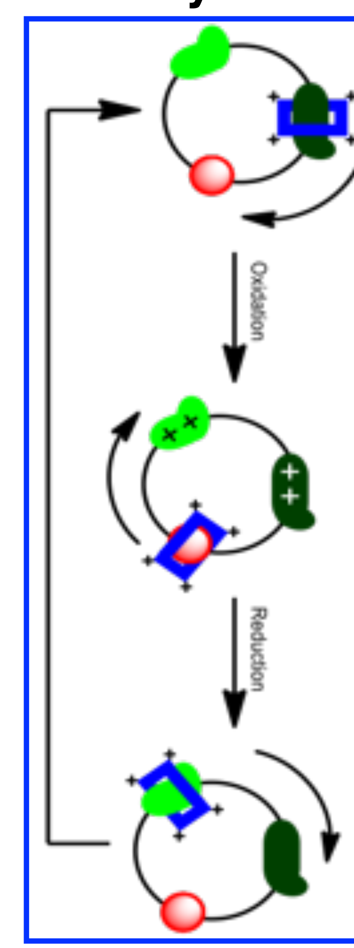
Machine Learning application



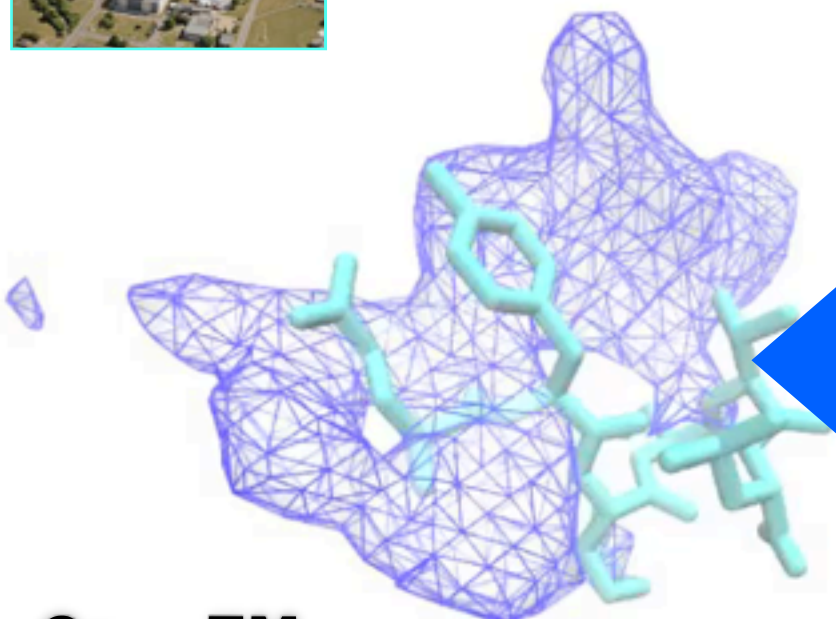
Linear motion

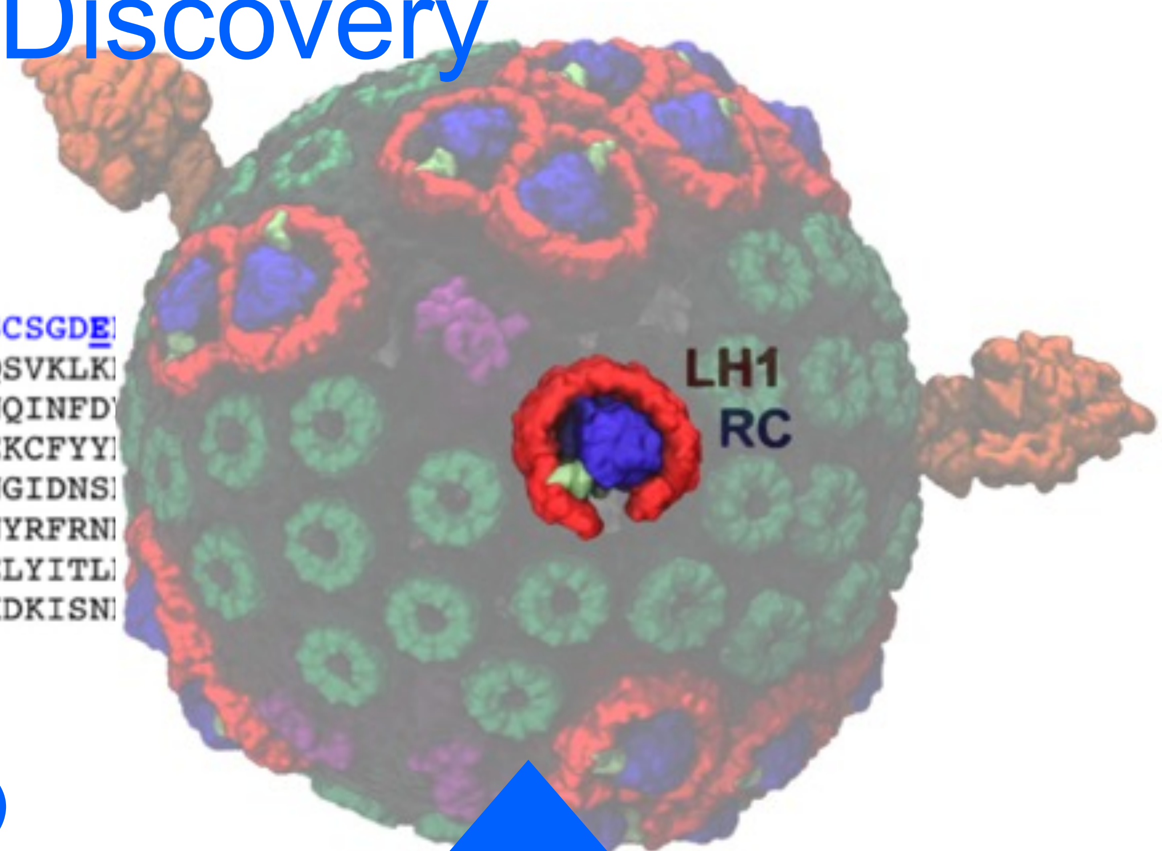


Rotatory motion

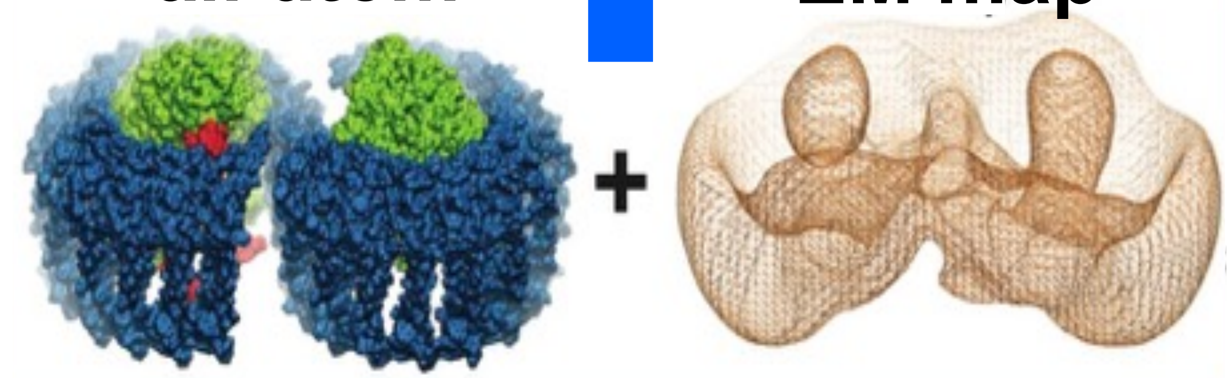


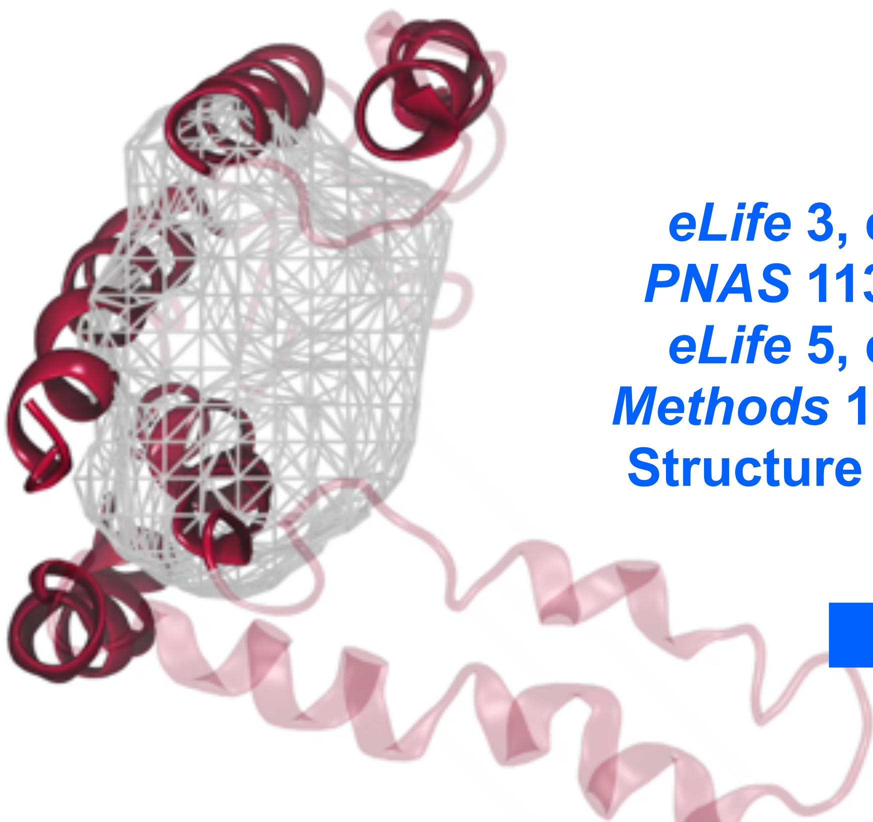
Scientific Accomplishment # 4: Data-guided Structure Discovery



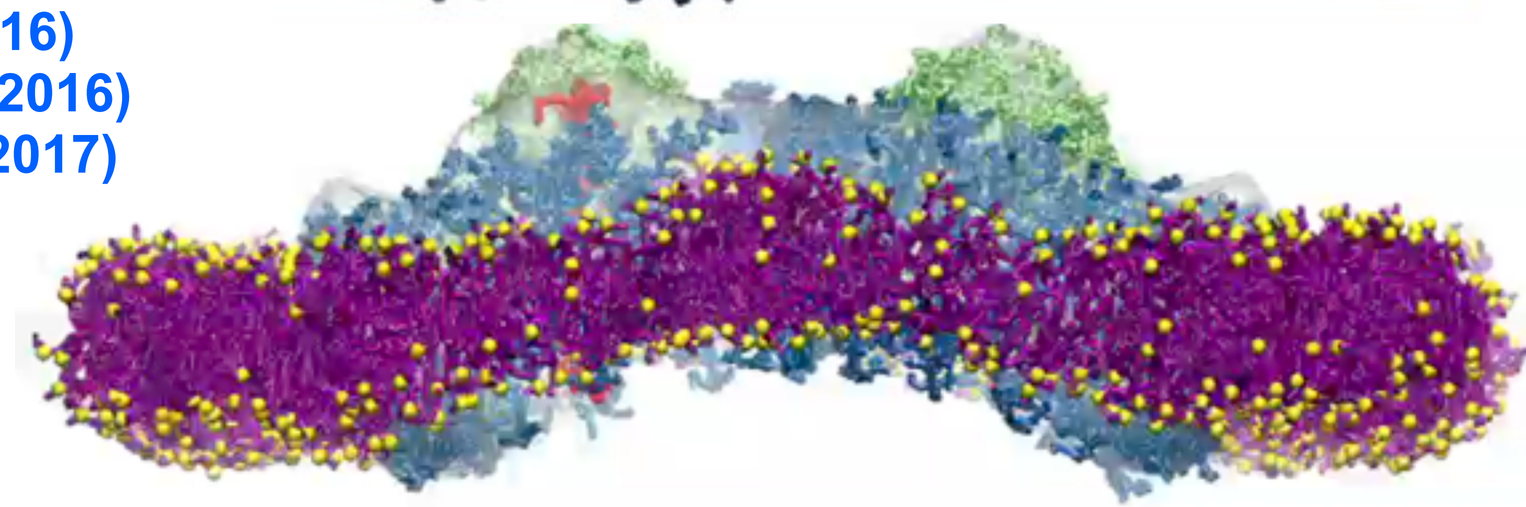


Acta. Cryst. D 70, 2344 (2014)
Nat. Struct. Mol. Biol. 21, 244 (2014)
J. Am. Chem. Soc. 137, 8810 (2015)
J. Am. Chem. Soc. 138, 4843 (2016)





eLife 3, e03035 (2014)
PNAS 113, 10310 (2016)
eLife 5, e16105 (2016)
Methods 100, 50-60 (2016)
Structure 24, 2102 (2017)



Molecular Dynamics Flexible Fitting - creation of data-driven force fields

Two terms are added to the MD potential

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

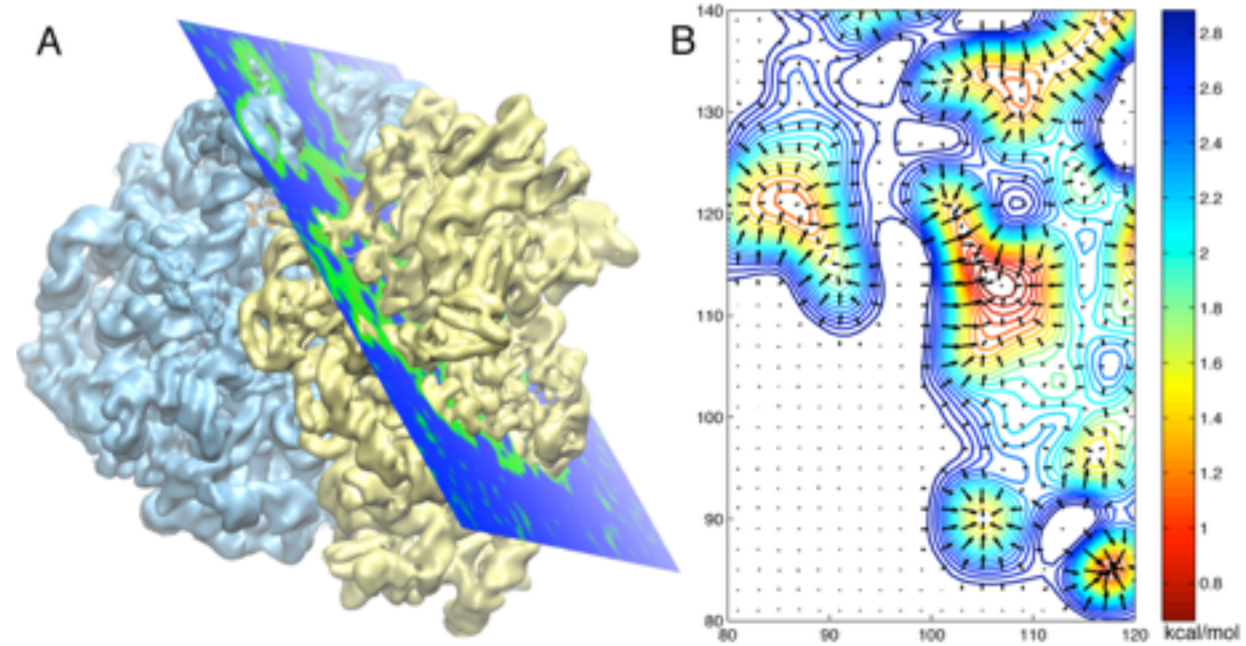
An external potential derived from the EM map is defined on a grid as

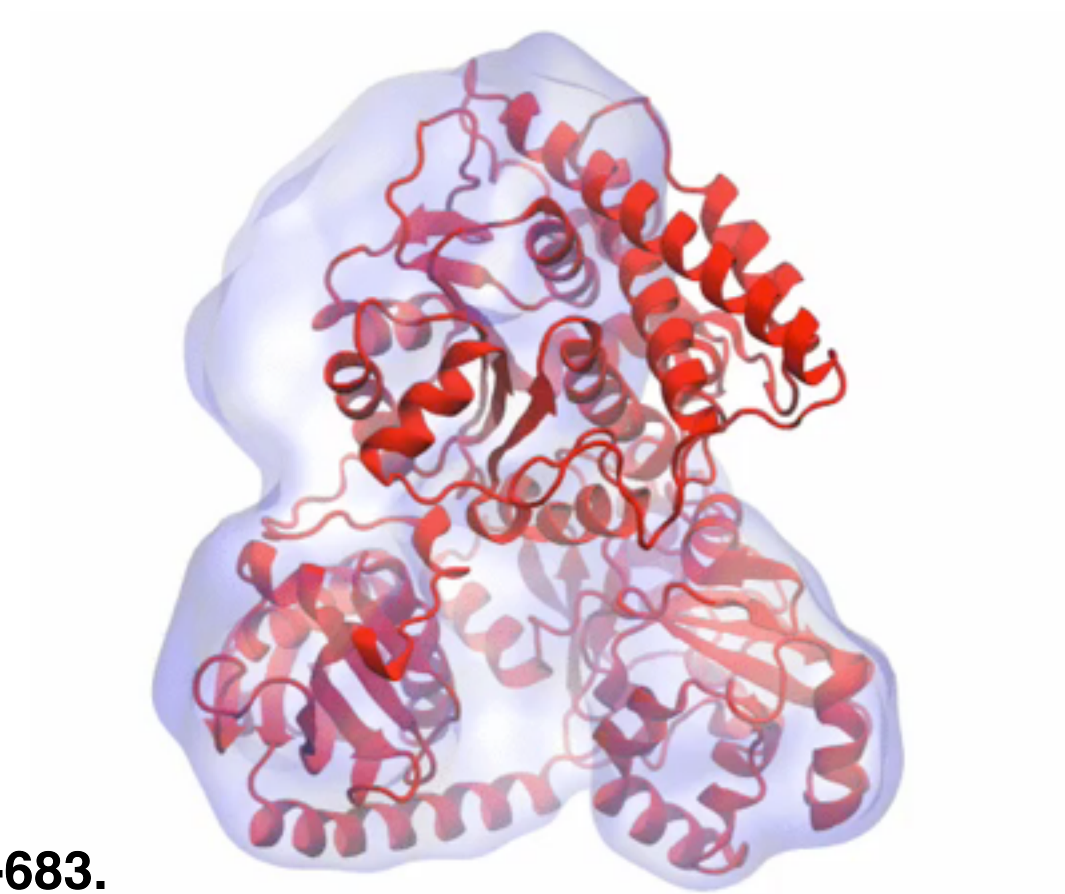
$$U_{EM}(\mathbf{R}) = \sum_j w_j V_{EM}(\mathbf{r}_j)$$

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

A mass-weighted force is then applied to each atom

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



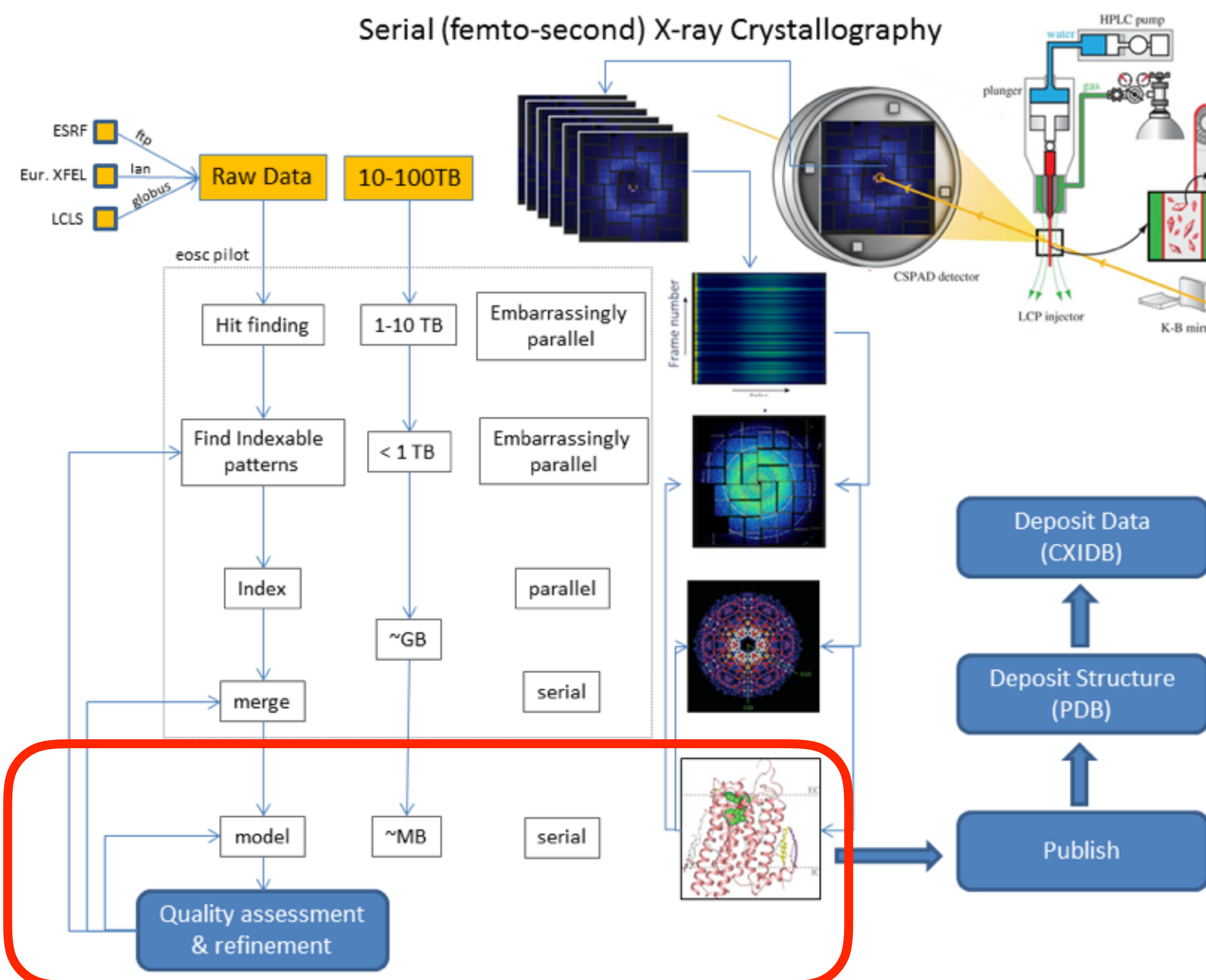


Acetyl – CoA Synthase

[1] Trabuco,Villa..Frank*, Schulten *Structure* (2008) 16:673-683.
[2] Trabuco,Villa..Frank*, Schulten *Methods* (2009) 49:174-180.

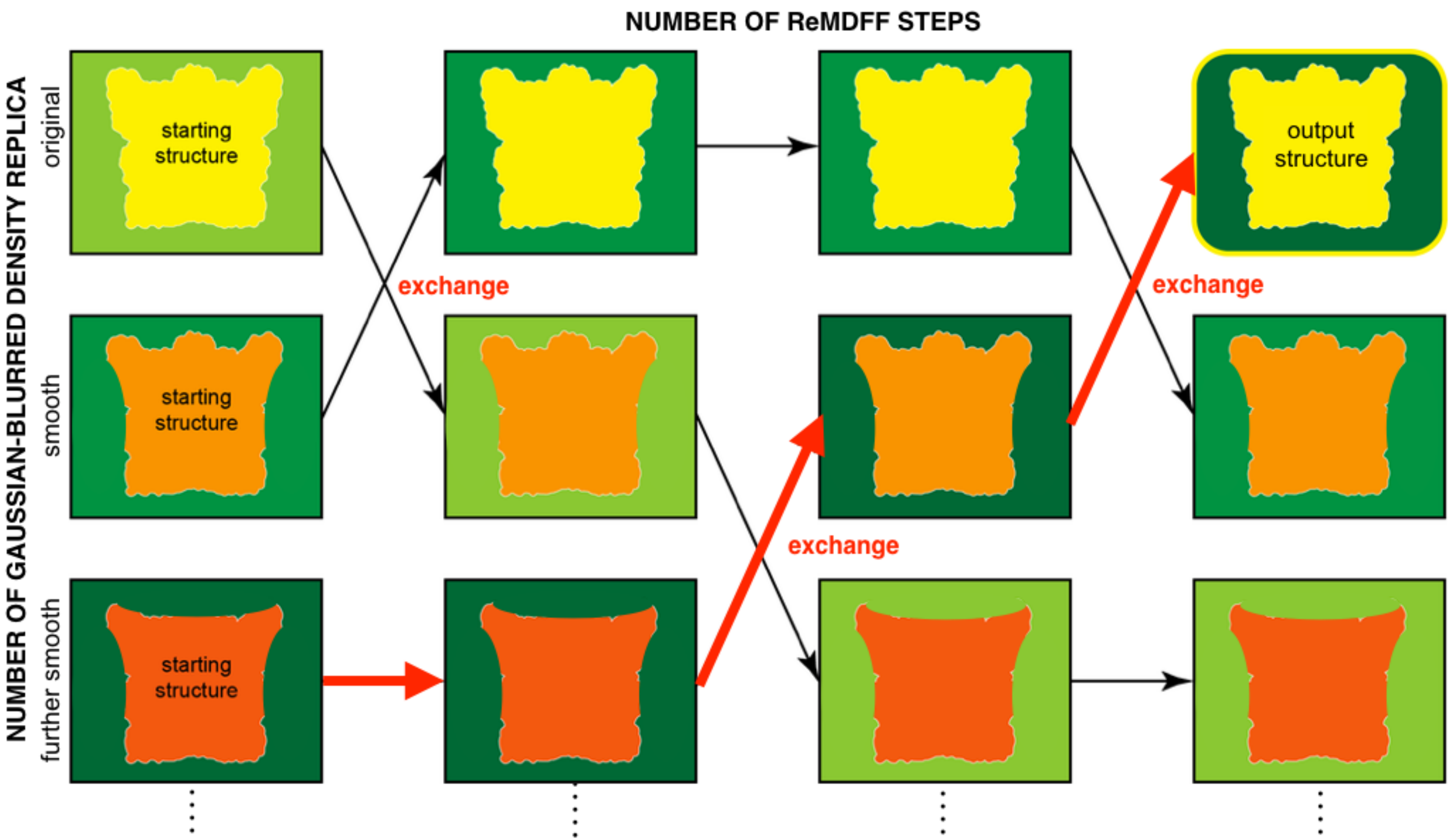
***2017 Nobel Prize in Chemistry**

Data-acquisition Pipelines at NSF BioXFEL Center

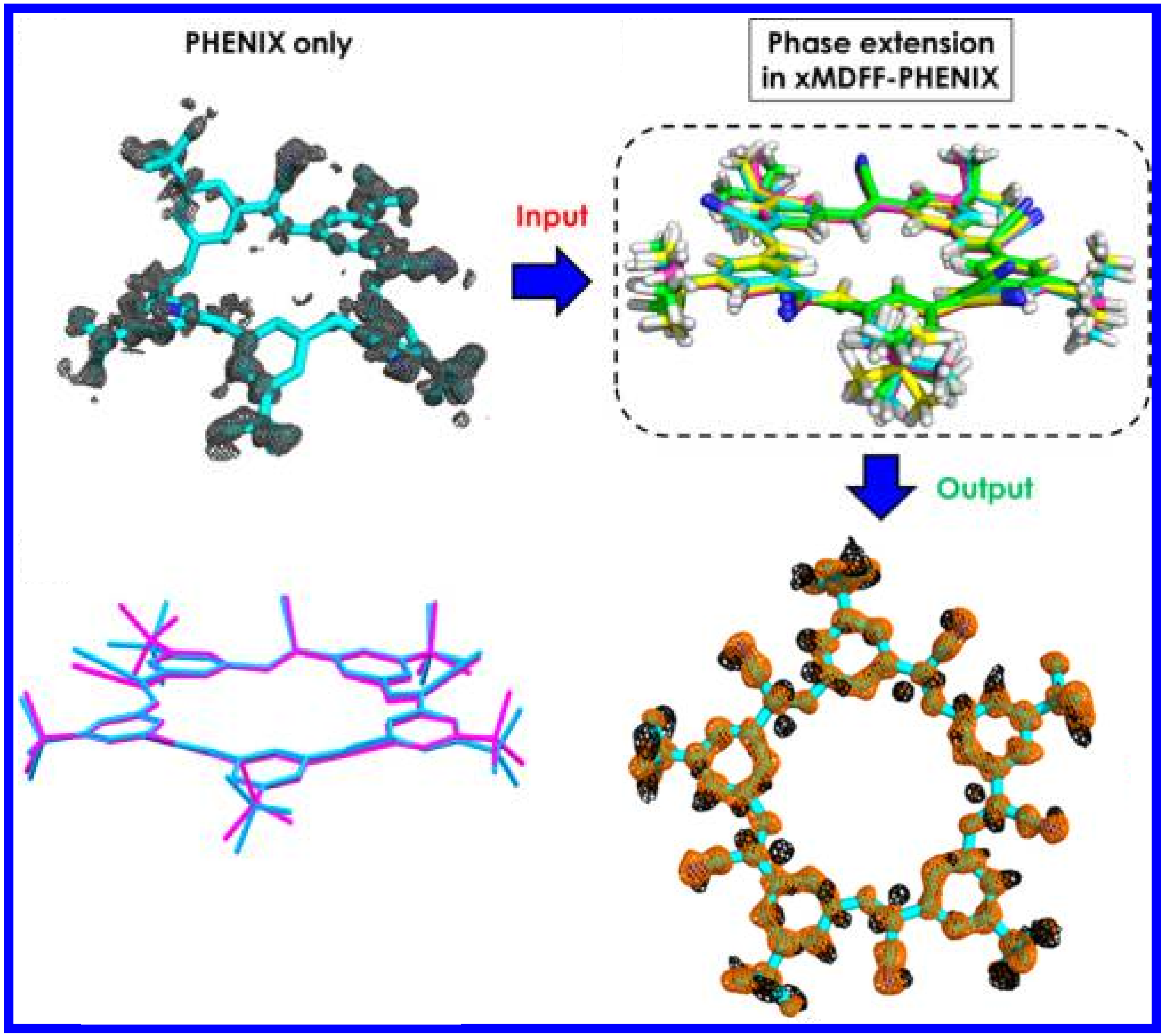


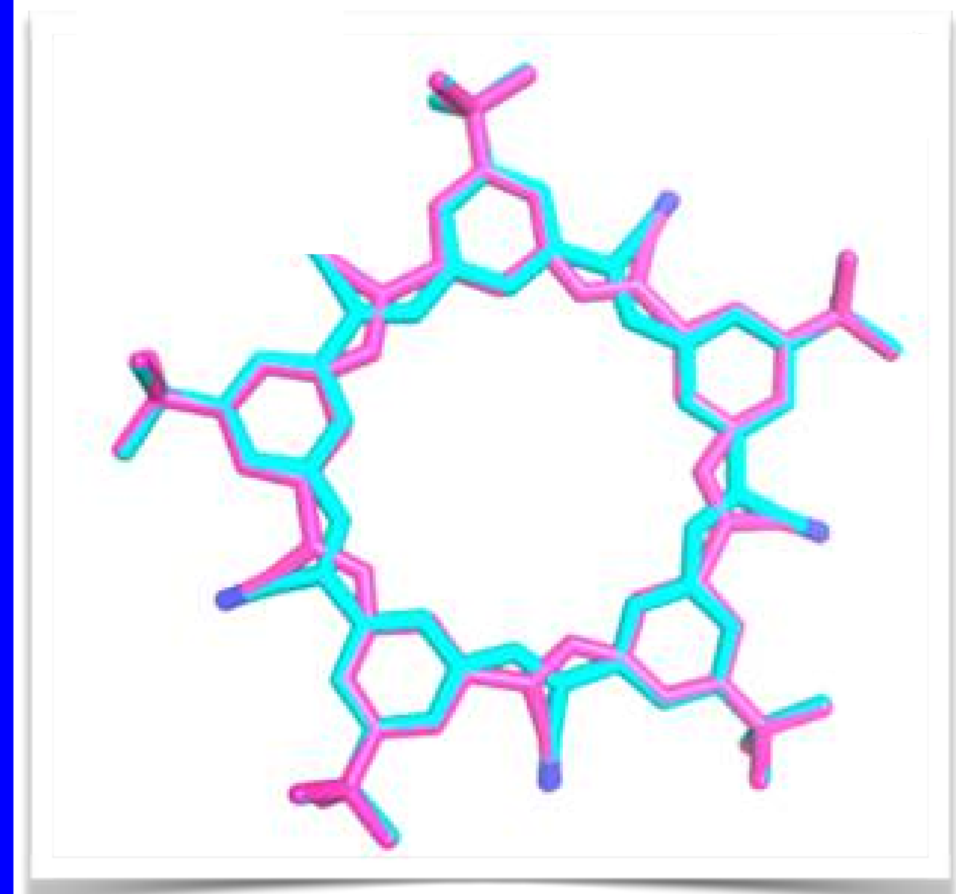
Fromme (ASU)

Scientific Challenge # 4: Structure Discovery with Supercomputers



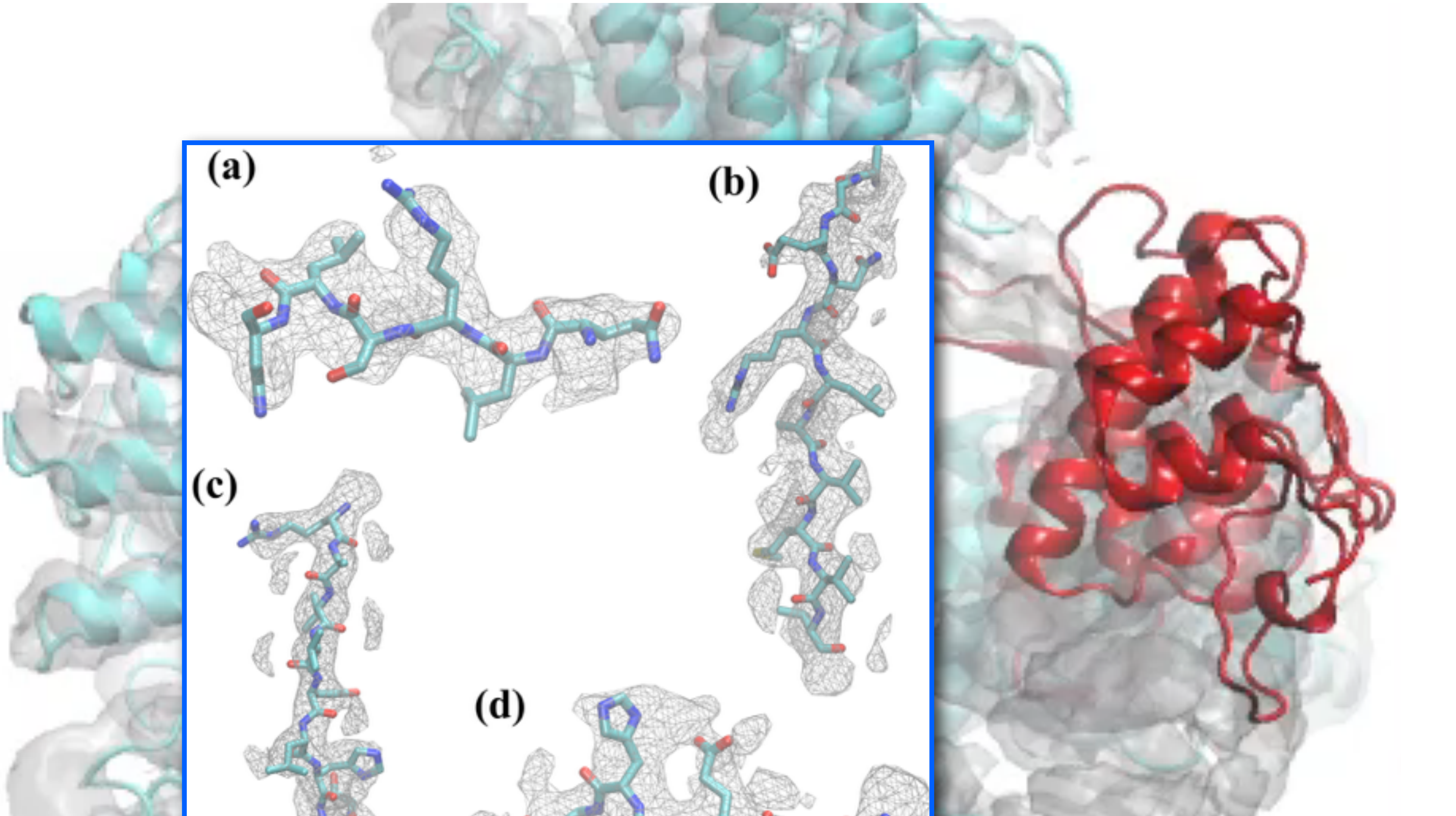
Combination of “Enhanced-sampling” with Flexible Fitting

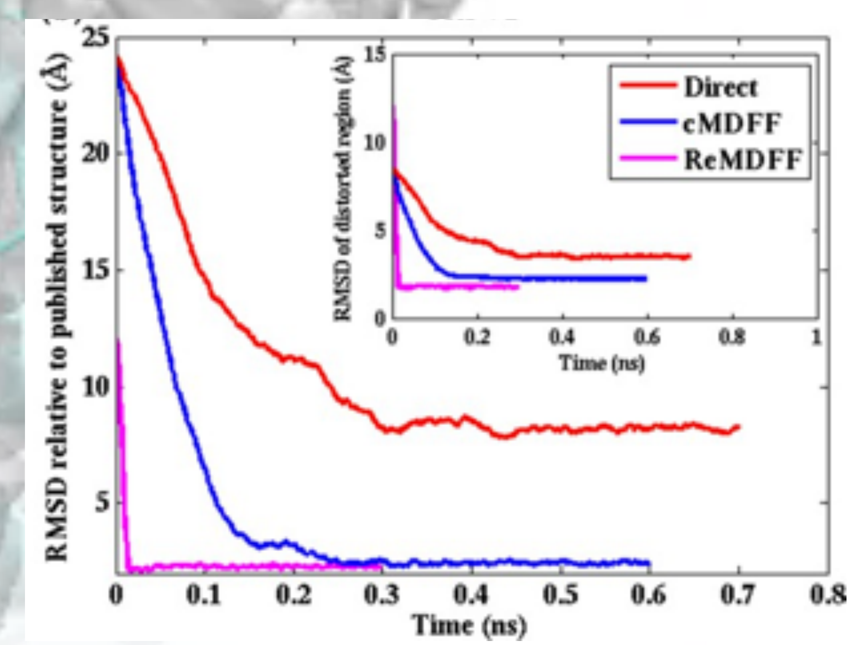




Multiple stereoisomers
in the same 1.8 Angstrom data

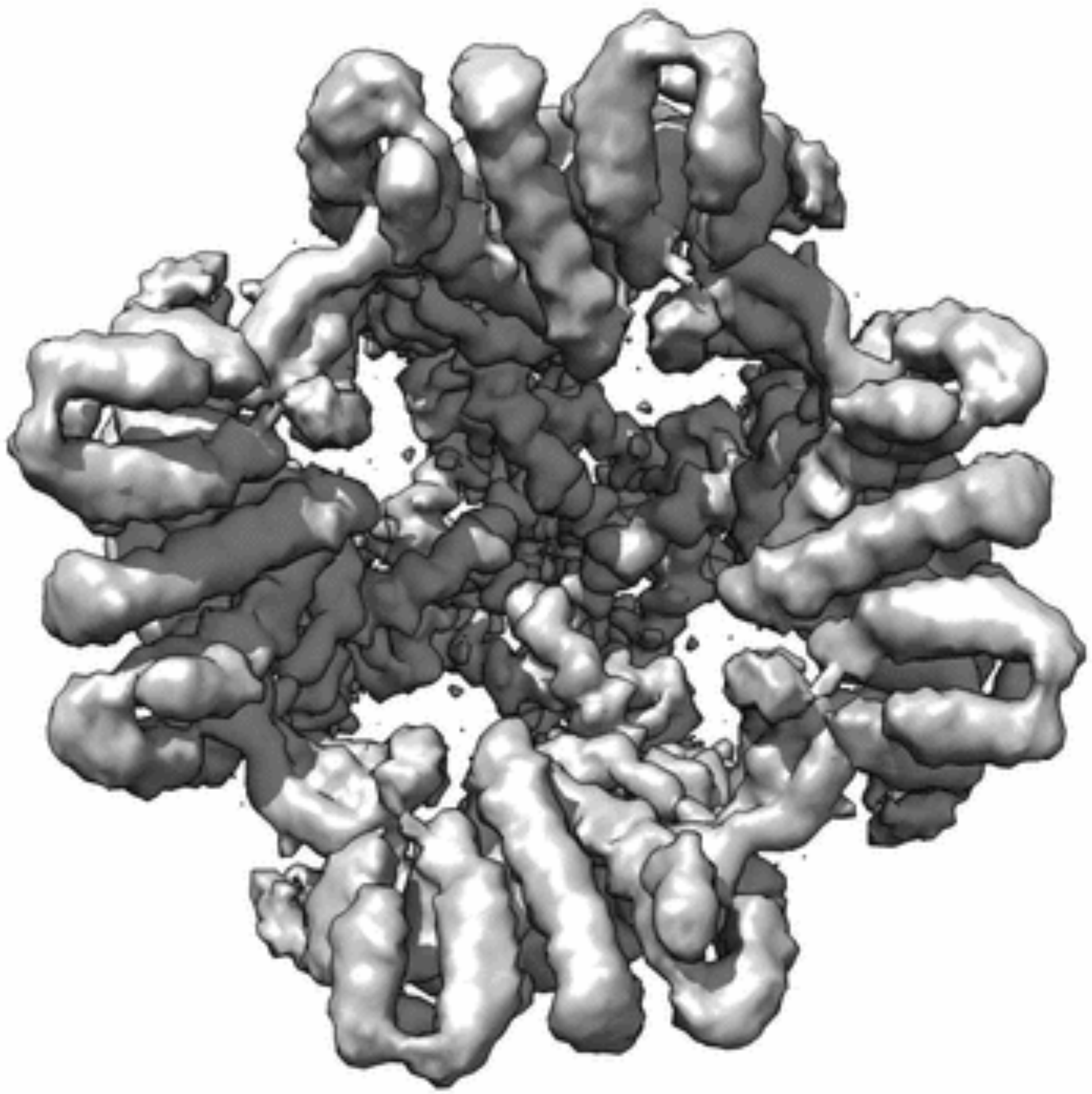
J Am. Chem. Soc. 2015, 137, 8810
J Am. Chem. Soc. 2016, 138, 4843



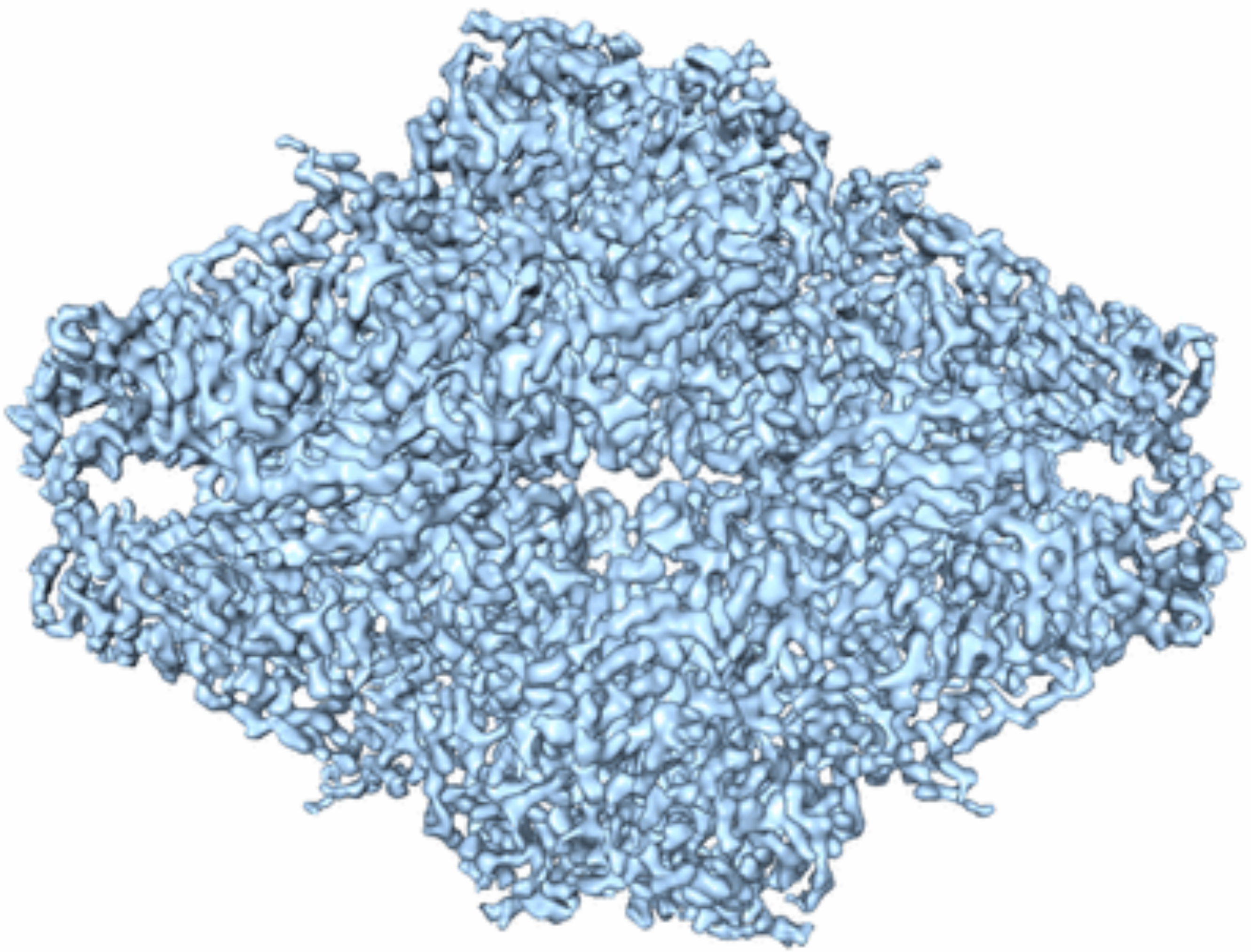


eLife 5, e16105 (2016)
Methods 100, 50-60 (2016)

Results of the Cryo-EM Structure Challenge (EM193)



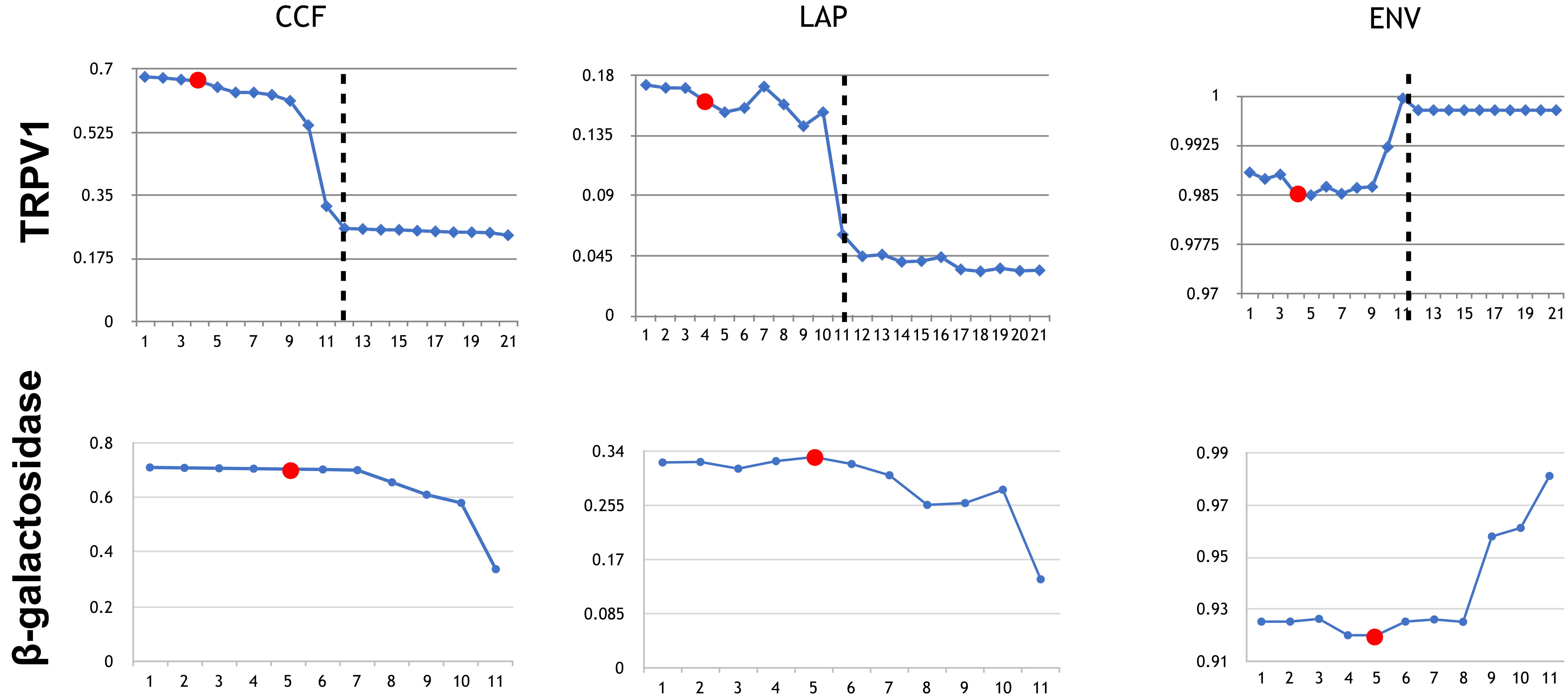
emb_5778



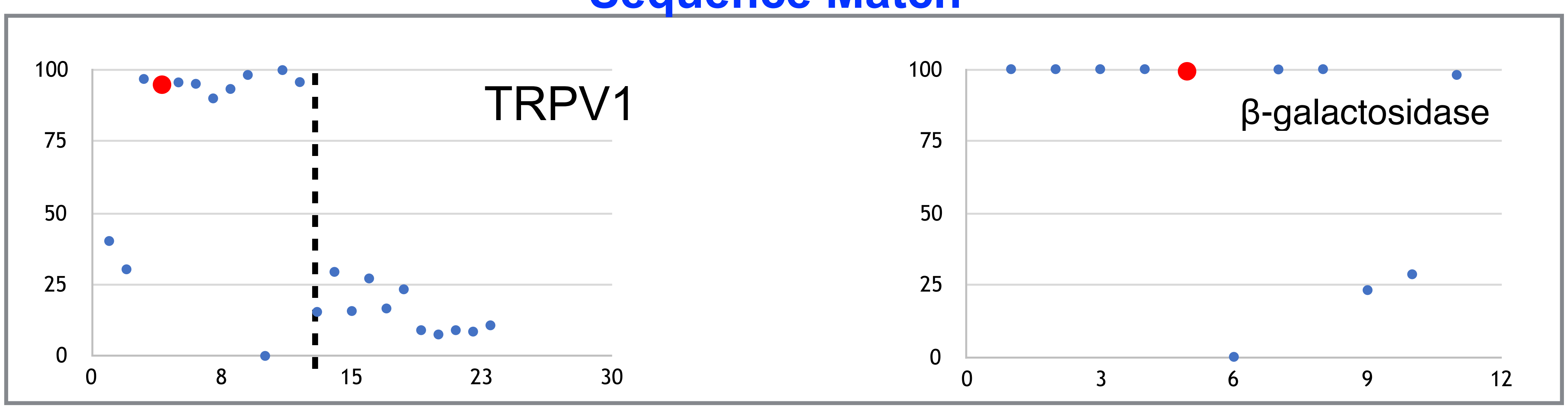
emb_5995

Journal of Structural Biology (Under review at for special section commemorating 2017 Nobel Prize)

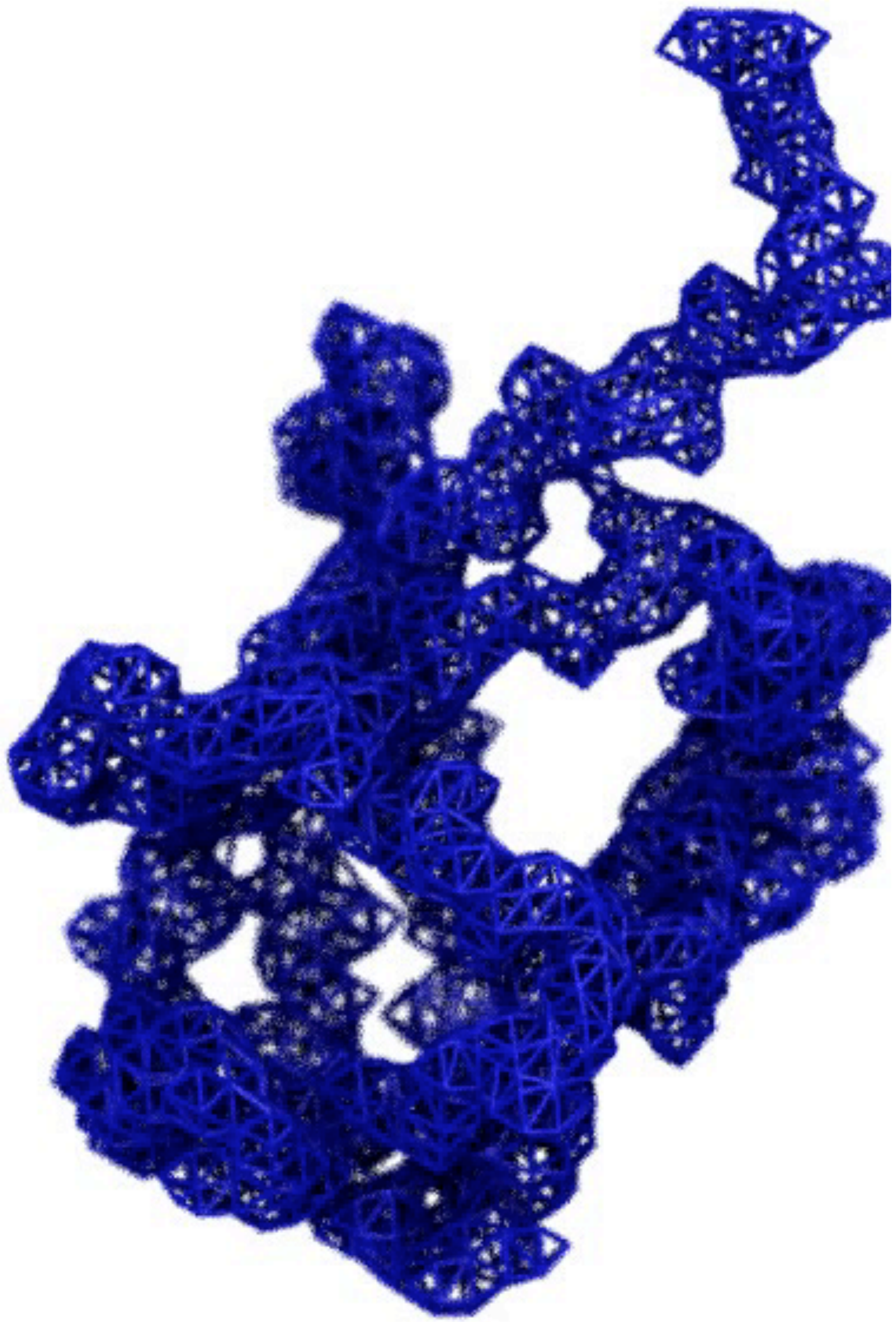
Quality of fitting

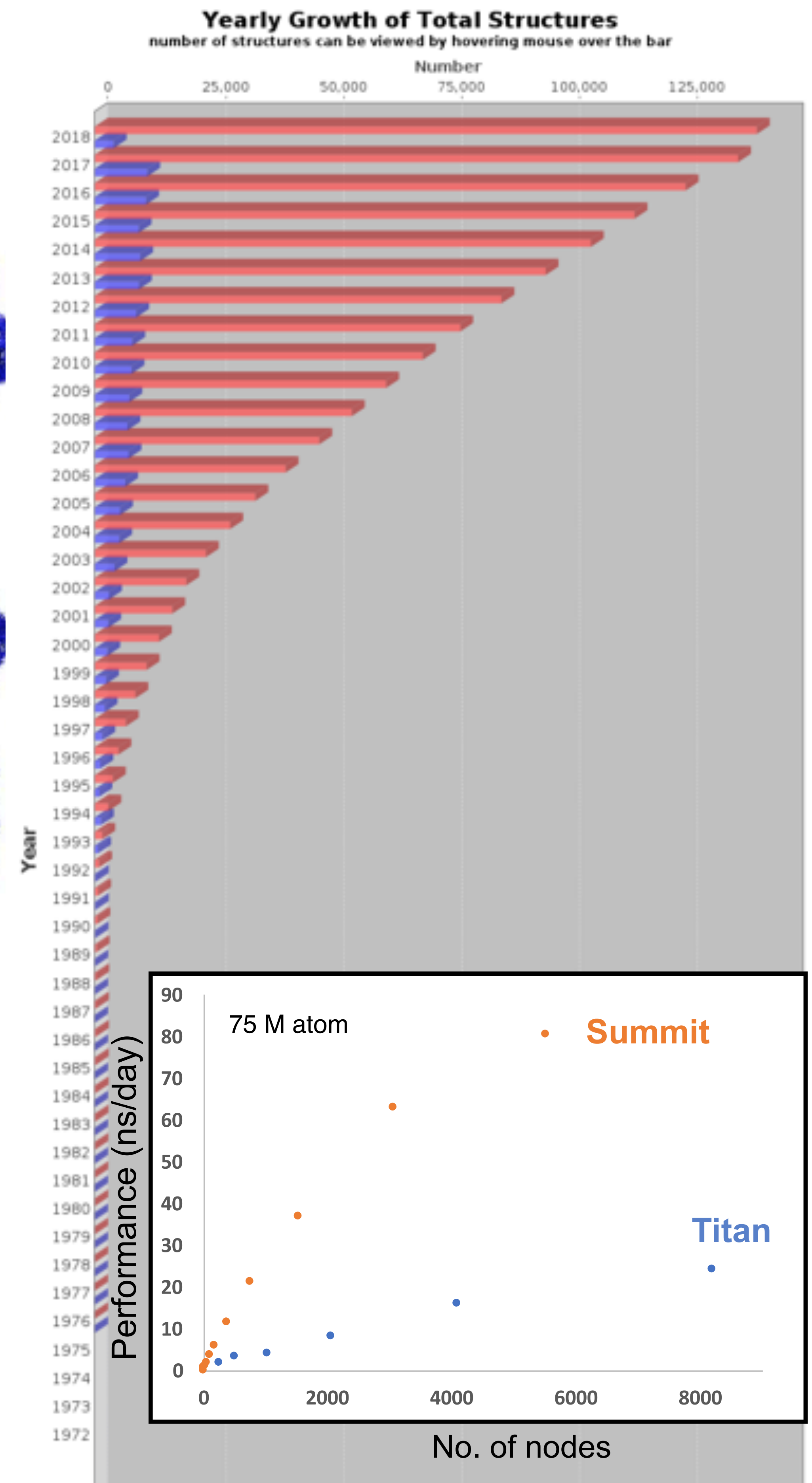


Sequence Match



Vision: Structure Discovery with Summit





Tajkhorshid

Dill (Stony-Brooks U.)

Richardson (Duke)

Currently on BlueWaters