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**

\[
U(\vec{R}) = \sum_{\text{bonds}} k_{\text{bond}} (r_{ij} - r_{0})^2 + \sum_{\text{angles}} k_{\text{angle}} (\theta_{ij} - \theta_{0})^2 + \sum_{\text{dihedrals}} k_{\text{dihedral}} [1 + \cos(n_{ij} \phi_{ij} + \delta_{ij})] + \sum_{\text{strands}} k_{\text{strands}} [\frac{(\alpha_{ij} - \beta_{ij})^2}{\gamma_{ij}}] + \sum_{\text{crosses}} \frac{q_{ij} q_{kl}}{r_{ij}}
\]

**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)

**NAMD Software**

- Virus
- Protein Folding

**Supercomputers**

- 128
- 256
- 512
- 1024
- 2048
- 4096
- 8192
- 16384
- 32768

- Jim Phillips
- David Hardy

Molecular Dynamics (MD) simulations
Why Does One Need a Supercomputer?

Structural transitions

Ensemble of trajectories

Multiple replica required!!

Why Does One Need a Supercomputer?
Parallel Performance of NAMD on Summit

![Graph showing performance of NAMD on Summit](image)
Visualization and Analysis: VMD
In-situ Visualization of Billion Atoms : SIGHT

Noah Trebesch

Ben Hernandez (OLCF)
Scientific Accomplishment # 1: Energy Conversion in Bacterial Photosynthesis

**Low-light adaptation**

![Diagram of bacterial photosynthesis components](image)

**Equation: k_{ATP}(I) = \frac{1}{2} I q \left( 1 + \frac{1}{2} I q \tau_{RC}(I) \frac{1}{n_{RC}} \right) \)**

**Equation: \tau_{RC}(I) = 1 + (\tau_H - 1) \left( 1 - e^{\frac{I q}{2B}} \right) \)**

**Equation: \tau_H = \frac{n_{RC}}{n_B} \tau_B; B = \frac{2n_B}{\tau_B} \)**

*eLife* 2016, 5, e09541;
*Parallel Comput.* 2016, 55, 17

**Processes involved in energy conversion in the photosynthetic chromatophore.**

**Figure 4.** Processes involved in energy conversion in the photosynthetic chromatophore. (ATP synthase, cyt bc, cyt b, complex, quinone/quinol complexes, ATP synthase)
Summit Goals 1: From First-Principles to Phenotypic Behaviors

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


Figure 3. Effect of vesicle composition on steady-state ATP production at different light intensities. Vesicle composition is given in terms of the number of cytbc dimers (nB) and of RC-LH1-PufX dimers (nL) for vesicles featuring identical surface area; LH2 composition of the vesicle is determined by k_ATP (ATP/s).

Sener et al. eLife 2016;5:e09541. DOI: 10.7554/eLife.09541
Summit Goals 2: Conformational Transition in Molecular Motors

Scientific Accomplishment #2

Design principles of 100% energy conversion efficiency ??

J. Am. Chem. Soc. 2017, 139, 293
Scientific Accomplishment # 3: Synthesis of Artificial Motors

Inchworm motion

Tetrazine
BPTz

Monomer Pathway

Dimer Pathway

RDS

Bis-macrocycle

cyanosolo

cyanodimer

(Accepted for special issue commemorating the 2016 Nobel Prize in Chemistry)

Amar Flood
(Indiana U.)
Scientific Accomplishment # 4: Data-guided Structure Discovery

Nat. Struct. Mol. Biol. 21, 244 (2014)

eLife 3, e03035 (2014)
PNAS 113, 10310 (2016)
eLife 5, e16105 (2016)
Methods 100, 50-60 (2016)
Structure 24, 2102 (2017)
An external potential derived from the EM map is defined on a grid as

$$U_{EM} = \sum_j w_j V_{EM}(r_j)$$

Two terms are added to the MD potential

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

A mass-weighted force is then applied to each atom

$$f_i^{EM} = -\nabla U_{EM}(R) = -w_i \partial V_{EM}(r_i) / \partial r_i$$

Data-acquisition Pipelines at NSF BioXFEL Center

Serial (femto-second) X-ray Crystallography

Raw Data → 10-100TB

- Hit finding → 1-10 TB
- Find Indexable patterns → < 1 TB

- Index
- merge

→ model → ~MB

Quality assessment & refinement

- Deposit Data (CXIDB)
- Deposit Structure (PDB)

Publish

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
Results of the Cryo-EM Structure Challenge (EM193)

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

CCF

LAP

ENV

Sequence Match

TRPV1

β-galactosidase
Vision: Structure Discovery with Summit

Currently on BlueWaters