Using MDFF
Examples from Modern Research

Till Rudack
Klaus Schulten Group - Theoretical and Computational Biophysics Group
NIH Center for Macromolecular Modeling and Bioinformatics
University of Illinois at Urbana-Champaign
Density → Structure → Function

Computational Modeling → Molecular Dynamics
A Sampling of TCBG‘s MDFF Projects

- Chemosensory Array
- Chromatophore
- Rous Sarcoma Virus
- Rabbit Hemorrhagic Disease
- HIV
- 26S Proteasome
Integrating experimental methods into computational modeling to obtain complete structural models
The Recycling System of the Cell
The ubiquitin proteasome proteolytic pathway

1. Substrate tagging by Ubq4
2. Ubq4-substrate recognition
3. Substrate degradation

Kisselev Cancer Cell 2013
Near-atomic model of the 26S proteasome

Cryo-EM density

Subunits from X-ray crystallography, NMR, and homology modeling

Wolfgang Baumeister
Friedrich Foerster

PDB-ID 4CR2
EMDB-ID 2594
Resolution 7.7 Å
Unverdorben et al. PNAS 2014

Molecular Dynamics Flexible Fitting (MDFF):
Trabuco et al. Structure 2008
Functional subunits of the 26S proteasome

Ubiquitin Recognition (Rpn10, Rpn13, Rpn1)

Deubiquitylation (Rpn11)

Substrate Unfolding (ATPase-ring)

Substrate Degradation (α-ring, β-ring)
Deubiquitylation subunit: Rpn11

Complete models are a basic prerequisite to perform MD simulations

Active site of Rpn11: substrate is cleaved from ubiquitin tag

Deubiquitylation (Rpn11)

Missing segments
- highly flexible
- ambiguous density

Chain V of PDB-ID 4CR2

www.ks.uiuc.edu/~trudack

Unverdorben et al. PNAS 2014
Combining Rosetta and MDFF

incomplete structural model deposited in the PDB

de novo structure prediction → energy ranking → model filtering → interactive MDFF of cryo-EM data

complete structural model that fits cryo-EM data

**Rosetta**
Leaver-Fay *et al.* Methods Enzymol. 2011

**VMD/NAMD**
Humphrey *et al.* J. Mol. Graph. 1996

Integrating user expertise into *de novo* structure prediction
Model filtering by secondary structure

Secondary structure histogram of predicted ensembles of Rpn11's C-terminal tail

- α-Helix
- 3-10 Helix
- Turn
- Coil

Averaged pattern

Near atomic model (4cr2)
**Predicted model**

Representative model of the predicted averaged secondary structure pattern for Rpn11’s C-terminal tail (purple).

Rosetta tends to build compact structures.

Secondary structure pattern of amino acids 217-306 (purple)
Visual inspection of cryo-EM density
Predicted model to initiate MDFF

Representative model of predicted ensemble for Rpn11′s C-terminal tail

Secondary structure pattern of amino acids 217-306 (purple)
Interactive Molecular Dynamics Flexible Fitting

MDFF can be run on Cloud computing for low cost!

www.ks.uiuc.edu/~trudack

MDFF Tutorial on YouTube and at http://www.ks.uiuc.edu/Research/mdff/
Complete model of Rpn11 fitted to density
Quality check by cross-correlations

Rpn11 colored by local cross correlations
Incomplete vs. complete model

Incomplete model

Cross correlation 0.61

Complete model

Cross correlation 0.63

Rosetta/MDFF
Low vs. high resolution density model

**Red:** 3.5 Å cryo-EM model of Rpn11 within the isolated proteasomal lid

**Purple:** completed Rpn11 model within the 7.7 Å proteasomal cryo-EM density

- **Isolated lid cryo-EM model**
  - Gabriel Lander / Andreas Martin
  - PDB-ID 3JCK    EMDB-ID 6479
  - Resolution 3.5 Å
  - Dambacher et al. eLife 2016

- **26S proteasome cryo-EM density**
  - Wolfgang Baumeister
  - EMDB-ID 2594
  - Resolution 7.7 Å
  - Unverdorben et al. PNAS 2014
Low vs. high resolution density model

Structure predicted for low resolution matches structure of high resolution.

Secondary structure pattern obtained by Rosetta/MDFF employing a 7.7 Å density.

Secondary structure pattern of a structure modeled into a 3.5 Å density.

Deubiquitylation (Rpn11)
Functional subunits of the 26S proteasome

Ubiquitin Recognition (Rpn10, Rpn13, Rpn1)

Deubiquitylation (Rpn11)

Substrate Unfolding (ATPase-ring)

Substrate Degradation (α-ring, β-ring)
Ubiquitin recognition by Rpn10

Ubiquitin Recognition

δ-  Rpn10 (UIM)

δ+  Monoubiquitin

δ+  Tetraubiquitin
Ubiquitin recognition by Rpn10

Ubiquitin Recognition (Rpn10)
Ubiquitin recognition and deubiquitylation

Rpn10

PDB-ID 2X5N
(S. Pombe)

Ubiquitin Recognition (Rpn10)

Deubiquitylation (Rpn11)

Ubiquitin Transport

UIM

PDB-ID 2KDE
(human)
Generalized Simulated Annealing – GSAFold

GSAFold NAMD Plugin – Allows *ab initio* structure prediction

New implementation of GSA on supercomputers allows the conformational search for large flexible regions.

- Amino acid residues connecting Rpn10’s UIM with the proteasome are likely to be disordered and stochastic searching algorithms such as GSA can be used to explore their conformational space.

- GSAFold coupled to NAMD searches low-energy conformations to be used as starting points for the molecular dynamics studies.

Rafael C. Bernardi  Marcelo Melo
The conformational space of the Rpn10 linker is highly flexible.
Ubiquitin Transport to Deubiquitinase Rpn11

Ubiquitin Recognition (Rpn10)

Deubiquitylation (Rpn11)

Ubiquitin Transport
Functional subunits of the 26S proteasome

Substrate Unfolding (ATPase-ring)
The Motor of the Proteasome
Resolved nucleotides are needed
3.9 Å Resolution Density of the Human 26S Proteasome
Advantage:
Positions of bulky side chains can be observed from density

Challenge:
o no detailed side chain orientation
X-ray structure refinement tools failed in the range of 4-5 Å resolution

Solution:
combining MDFF with
monte carlo based backbone and side chain rotamer search algorithms
in an iterative manner
The ATPase Motor of the 26S Proteasome


PDB-IDs: 5L4G, 5L4K
EMDB-ID: 4002
The Motor Action of protein unfolding

Rpt3  Rpt4  Rpt5  Rpt1  Rpt2  Rpt6

Coiled-Coil  
OB fold  
Large AAA+  
Small AAA+  
HbX  
20S CP α-ring

Rpt3  Rpt4  Rpt5  Rpt1  Rpt2  Rpt6
The atomic structure enable detailed investigations of the unfolding process by QM/MM simulations combined with path sampling techniques.

NAMD QM/MM interface with MOPAC and ORCA will be released in the second semester of 2016.
ModelMaker

incomplete protein model

VMD

no density

mid-resolution density

mid-resolution or high-resolution density

ModelCompleter

ModelFitter

ModelRefiner

complete protein model

complete protein model fitted to mid-resolution cryo-EM data

complete protein model refined to high-resolution cryo-EM data

de novo structure prediction

model ranking

model filtering

interactive MDFF

backbone and sidechain refinement

iterative real-space refinement

highres MDFF

Rosetta

NAMD

VMD

Rosetta

NAMD
In order to obtain complete protein structures different experimental and computational methods need to be integrated.

Automation is important but user expertise is equally important.
Acknowledgments

The Schulten Group

Theory

Klaus Schulten
Ryan McGreevy

Experiment

Wolfgang Baumeister
Friedrich Förster
Eri Sakata

ModelMaker

Ryan McGreevy

GSA

Rafael Bernardi
Marcelo Mello

Maximilian Scheurer
Marc Siggel
Justin Porter

www.ks.uiuc.edu/~trudack