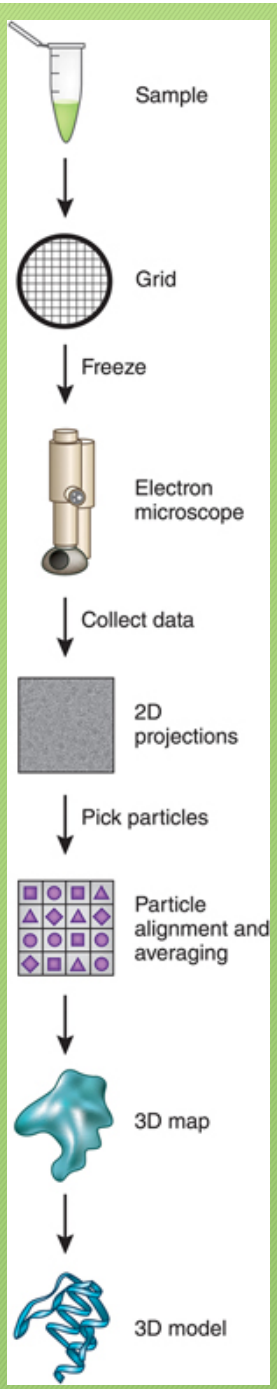
Single-Particle Analysis
Part III – Case Studies

Yong Zi Tan
Carragher, Potter and Mancia Lab
NYSBC/Columbia University

SPA Workflow



Doerr, 2016

- Many different ways to proceed in each step of the workflow
  - How to pick and choose the packages?
- Looking at examples of things that worked
  1. TRPV1 Channel to 3.4 Å (Liao *et al.*, 2014)
  2. Proteasome to 2.8 Å (Campbell & Veesler *et al.*, 2015)
  3. 60S Ribosome to 2.9 Å (Passos & Lyumkis, 2015)

Case Study 1: TRPV1 Channel

nature International weekly journal of science

Search Go

▸ Advanced search

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video | For Authors

Archive > Volume 504 > Issue 7478 > Articles > Article

NATURE | ARTICLE

日本語要約

Structure of the TRPV1 ion channel determined by electron cryo-microscopy

Maofu Liao, Erhu Cao, David Julius & Yifan Cheng

Affiliations | Contributions | Corresponding authors

Nature **504**, 107–112 (05 December 2013) | doi:10.1038/nature12822

Received 24 July 2013 | Accepted 30 October 2013 | Published online 04 December 2013

PDF | Citation | Reprints | Rights & permissions | Article metrics

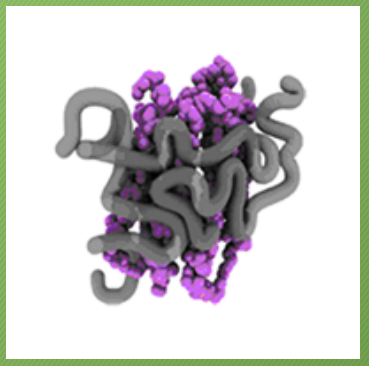
Information obtained from Liao *et al.*, 2013 and Cheng, 2014

1) Sample Preparation

- TRPV1 Characteristics
  - Membrane channel
  - 292 kDa tetramer (C4 symmetry)
- MBP affinity purification
- Solubilized in DDM and also amphipols

1) Sample Preparation

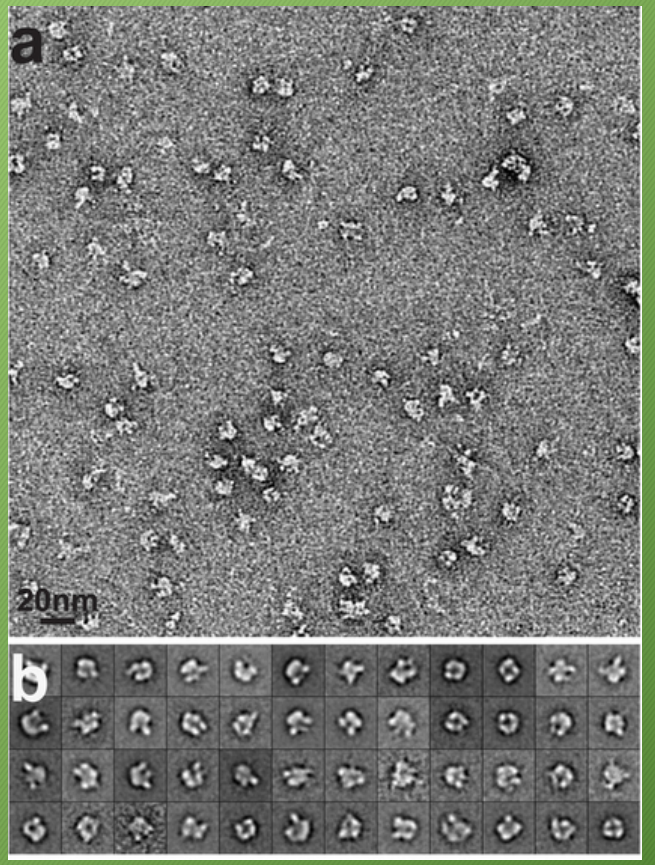
- TRPV1 Characteristics
  - Membrane channel
  - 292 kDa tetramer (C4 symmetry)
- MBP affinity purification
- Solubilized in DDM and also amphipols
  - Short amphipathic polymers that are able to keep individual membrane protein water-soluble in their native state under the form of small hydrophilic complexes
  - Pros: More stability, no detergent required
  - Other possibilities: Nanodiscs, liposomes, SMALPs (Postis *et al.*, 2015) and saposin-lipoprotein (Frauenfeld *et al.*, 2016)



Popot (2010), Popot *et al.*, (2003, 2011), Breyton *et al.*, (2010), Gohon & Popot (2003), Sanders *et al.*, (2004)

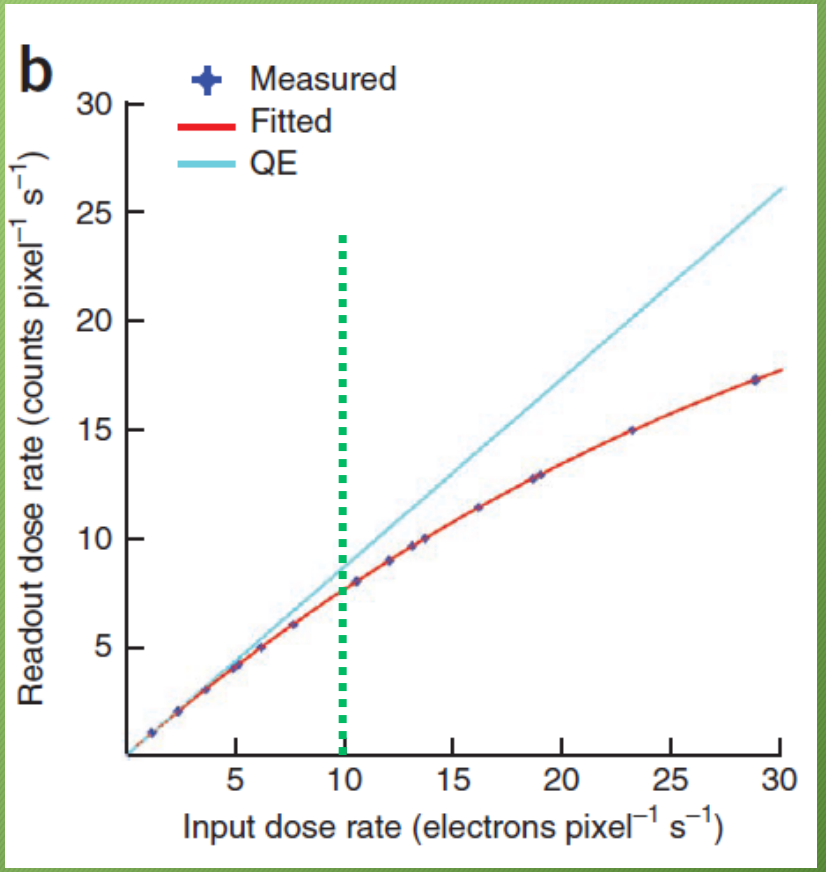
2) Screening and Characterization

- Using negative stain
- Random conical tilt to obtain initial model
  - Pros: *Ab initio*, useful when common lines does not work (preferred orientation), useful for heterogeneity
  - Cons: Possible flattening of protein, resolution restricted, missing cone



3) Cryo Data Collection

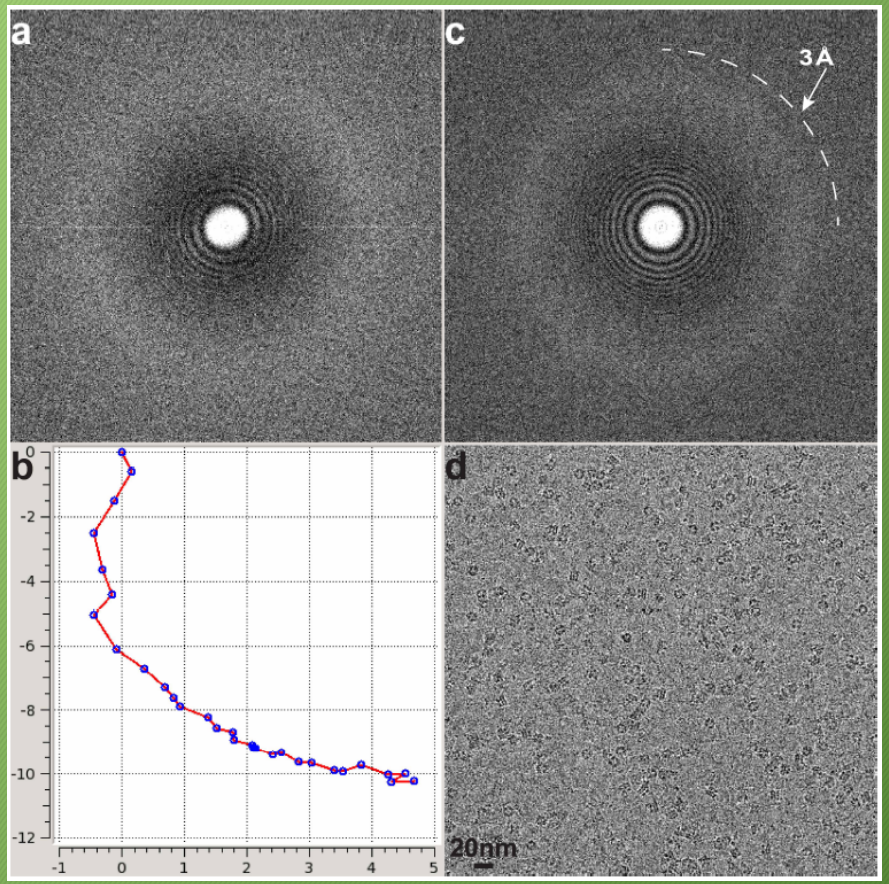
- Carbon grids
- TF20
  - 200 kV, CCD Camera
- Polara
  - 300 kV
  - K2 camera
  - Dose rate: 9.9 e-/pixel/s
    - Minimize coincidence loss
  - Defocus range: 1.5 to 3.0 mm
- Semi-automated collection using UCSFImage4



Li *et al.*, 2012

3) Cryo Data Collection

- Inspect your micrographs
  - Thon rings going out far after drift correction means good ice thickness

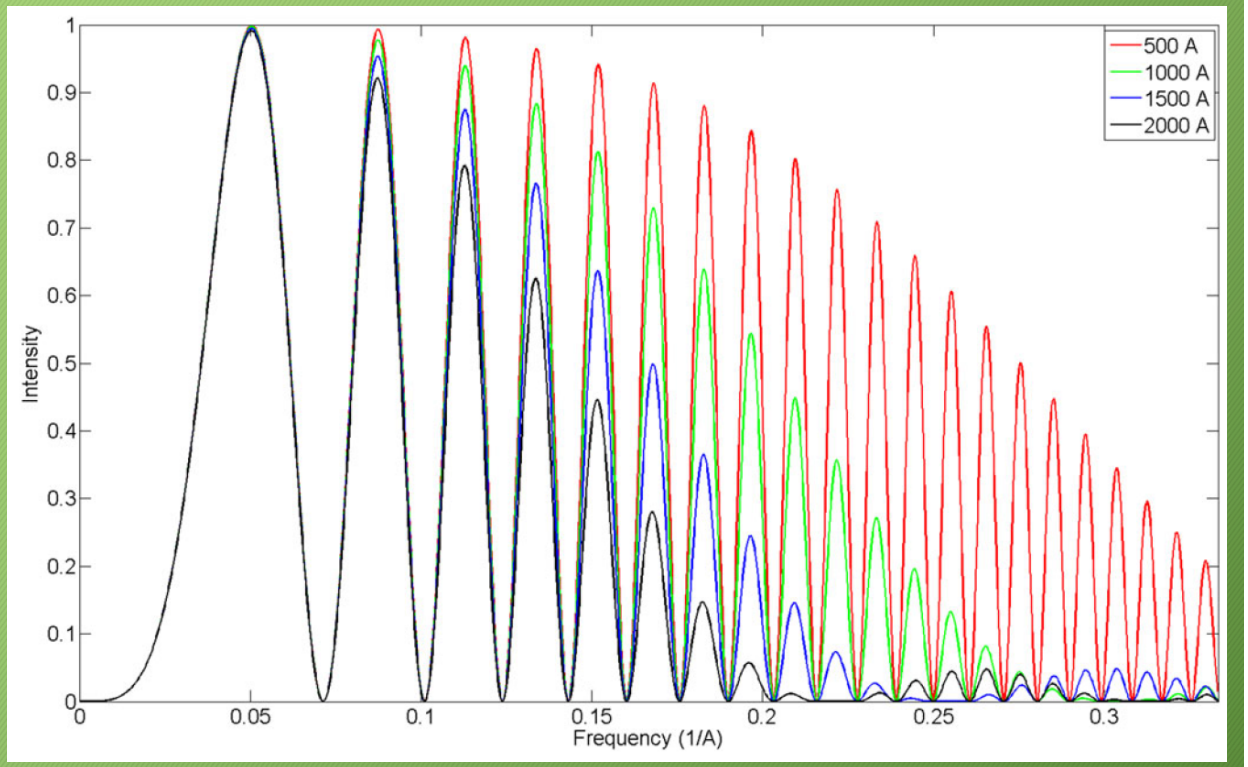


3) Cryo Data Collection

Theoretical Look at Ice Thickness

Loss of resolution due to

1. More noise from the ice
2. Averaging particles with defocii distributed over a range due to the ice thickness
3. Inelastic scattering

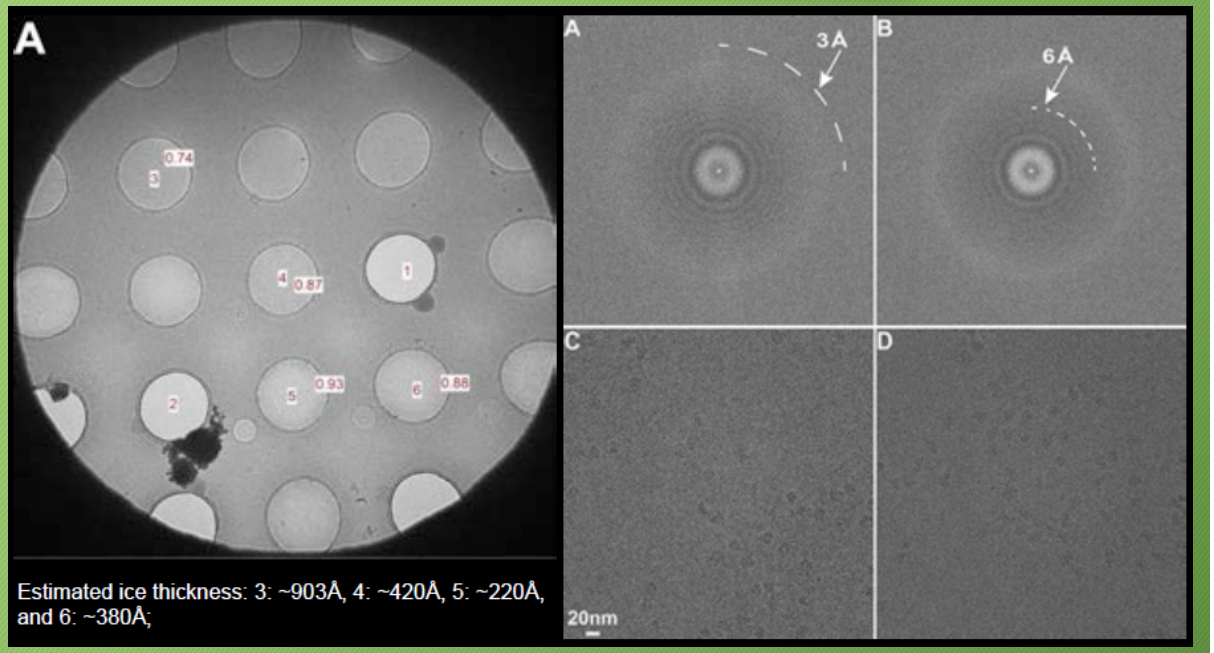


Wu *et al.*, 2015

3) Cryo Data Collection

Empirical Look at Ice Thickness

- Thicker ice, less Thon rings



Cheng, 2014

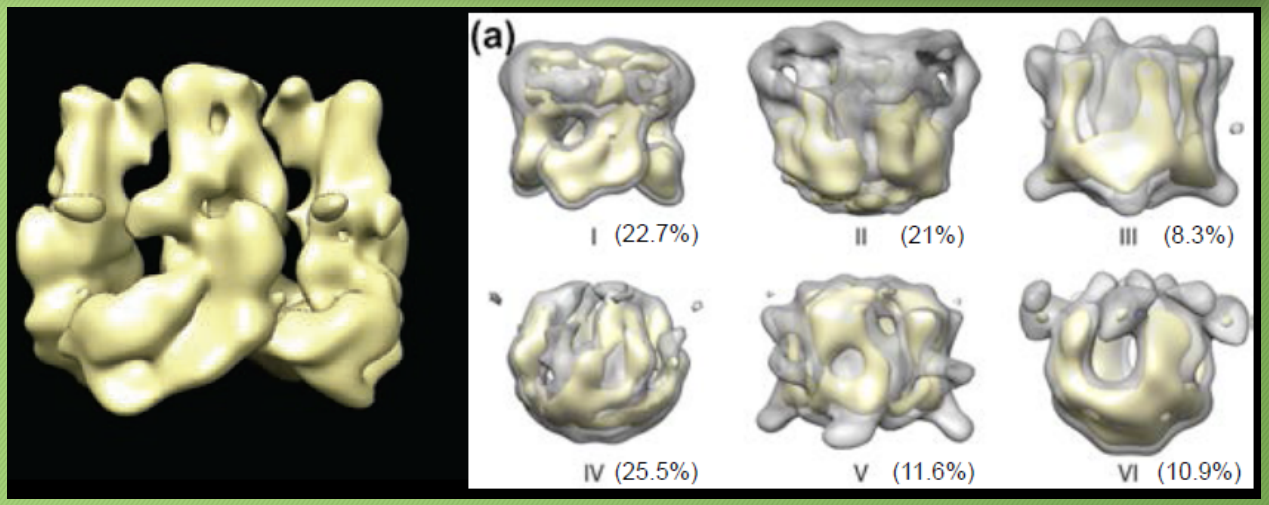
4) Data Processing

- Motioncorr: Whole frame alignment
- SamViewer: Particle picking
- CTFFind3 and CTFTilt: CTF estimation
- Spider: 2D Classification

4) Data Processing

- Frealign: 3D Reconstruction for RCT data
- Relion: 3D Classification and Reconstruction for Cryo data

4) Data Processing



Classification versus Refinement

What are the differences?

4) Data Processing

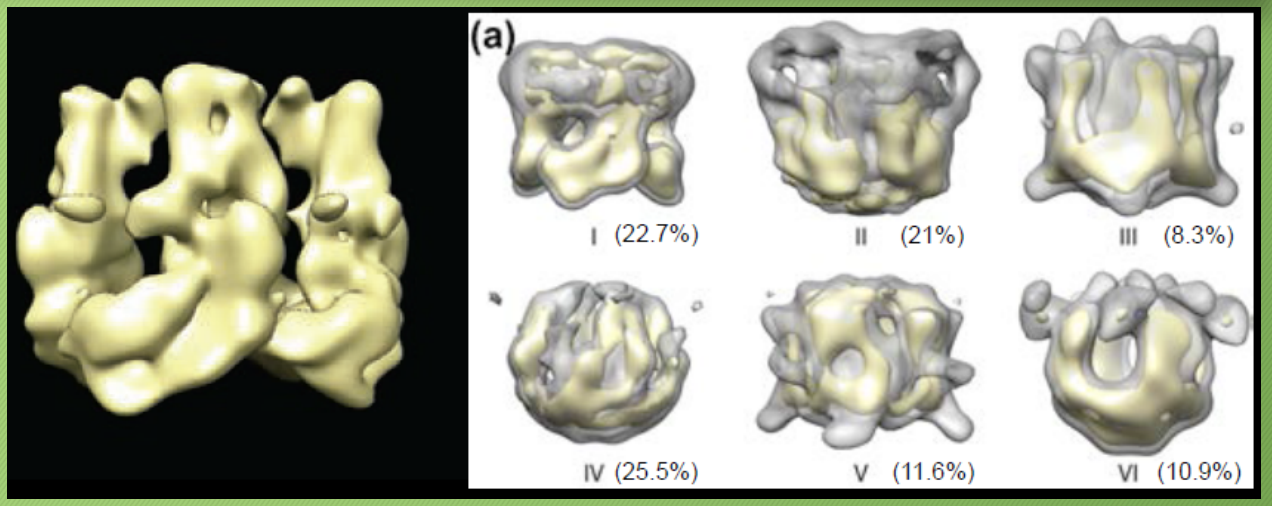
3D Classification

- Randomly split data up
- Coarse angular sampling usually used
- Euler angles and shifts can be refined together with classification, or done out of sync, or decoupled fully
- Use to clean up dataset, and tease out heterogeneity

3D Refinement

- All data refined against 1 model (data split into 2 half maps)
- Progressively finer angular sampling used
- Use to push resolution

4) Data Processing



Initial Model from RCT → 3D Classify TF20 Data → Pick Best Classes, Combine → 3D Autorefine → 8.8 Å structure

Low Pass Filter TF20 Structure → 3D Classify Polara Data → Pick Best Classes, Combine → 3D Autorefine → **3.4 Å Final Structure**

4) Data Processing

- Dose fractionation
  - Removal of first 2 frames for each movie stack
    - Affected most by beam induced movement
  - Removal of last 14 frames
    - Affected most by radiation damage
  - Now done more finely by
    - Relion
    - Tim Grant's exposure weighting software

5) Model Building

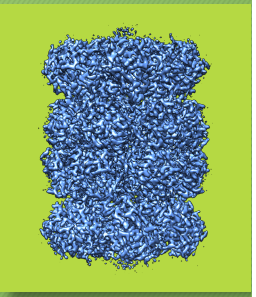
- Coot: Model building
  - Not all C_α or side chains built
- Validation of EM Map: Gold standard refinement and resolution
- Validation of atomic model: Ramachandran plot

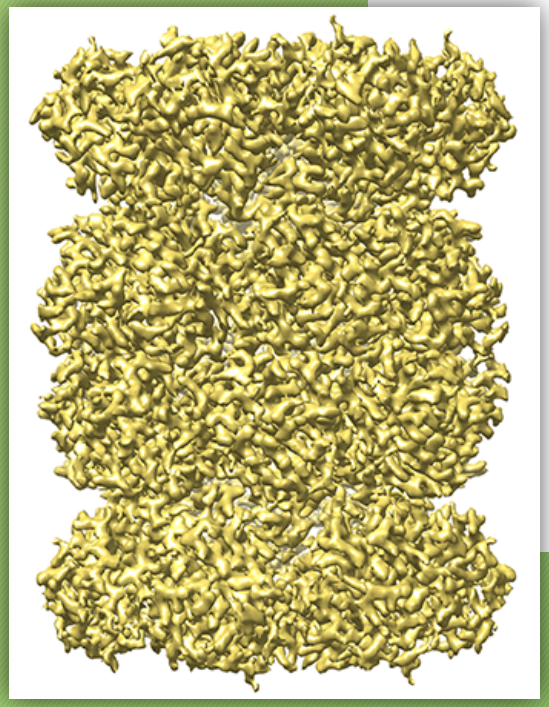
Quote from Yifan Cheng – What was required to get to 3.4 Å

1. Production of high quality and biochemically stable proteins
2. Available and well characterized pharmacological reagents
3. Camera related new technologies: high-DQE and dose fractionation
4. Classification of heterogeneous particles

Cheng, 2014

Case Study 2: Proteasome





eLIFE
elifesciences.org

SHORT REPORT

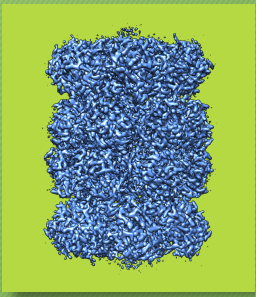
2.8 Å resolution reconstruction of the *Thermoplasma acidophilum* 20S proteasome using cryo-electron microscopy

Melody G Campbell[1,2†], David Veesler[1,2,3†], Anchi Cheng[1,2,4], Clinton S Potter[1,2,4], Bridget Carragher[1,2,4*]

[1]National Resource for Automated Molecular Microscopy, The Scripps Research Institute, La Jolla, United States; [2]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, United States; [3]Department of Biochemistry, University of Washington, Seattle, United States; [4]New York Structural Biology Center, New York, United States

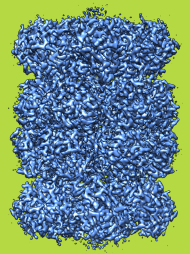
Information obtained from Campbell & Veesler 2014, Campbell & Veesler *et al.*, 2015

1) Sample Preparation



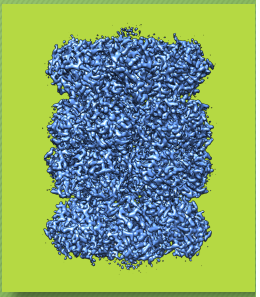
- T20S Proteasome Characteristics
  - Soluble protein
  - 700 kDa, D7 symmetry
- Common test sample
  - 3.3 Å cryo-EM structure (Li *et al.*, 2012)
  - Minimal structural heterogeneity

1) Sample Preparation

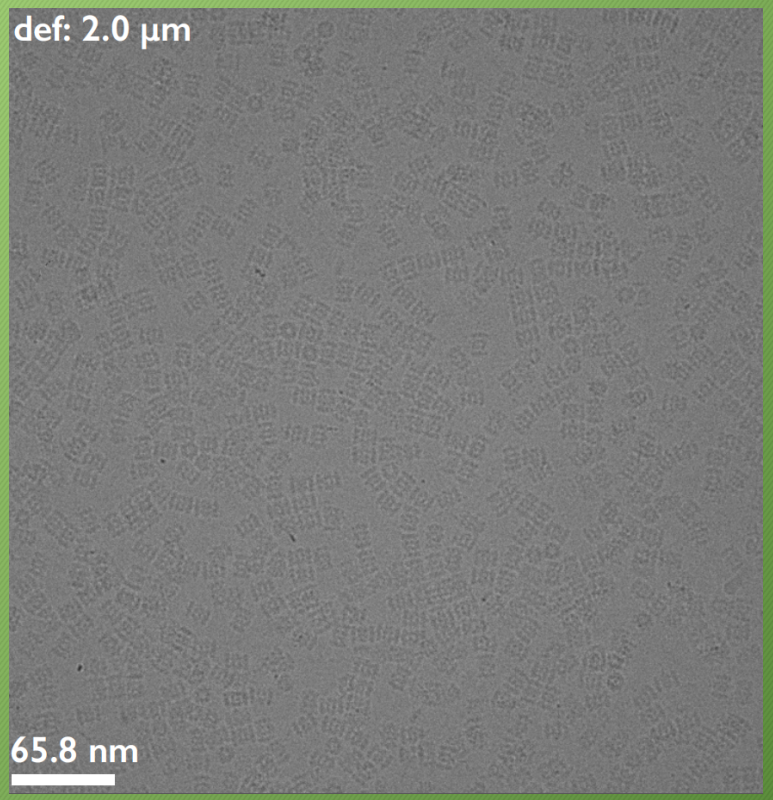


- T20S Proteasome Characteristics
  - Soluble protein
  - 700 kDa, D7 symmetry
- Common test sample
  - 3.3 Å cryo-EM structure (Li *et al.*, 2012)
  - Minimal structural heterogeneity

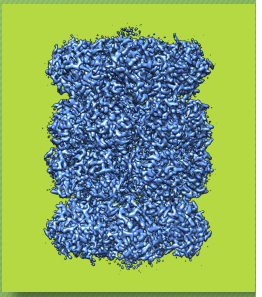
2) Cryo Data Collection



- Carbon grids
- Krios
  - 300 kV
  - K2 camera in super-resolution mode
    - Dose rate: 12 e-/pixel/s
    - Defocus range: 0.9 to 2.4 mm
  - Falcon II was also used and benchmarked
- Leginon automated data collection

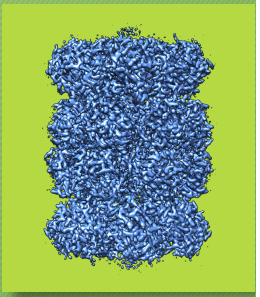


3) Data Processing

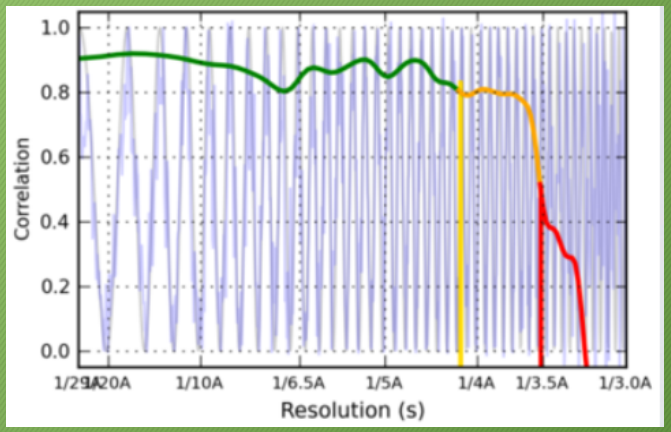


- Appion Package
  - Motioncorr: Whole frame alignment
  - Done simultaneously
    - FindEM: Template particle picking
    - CTFFind3: CTF estimation
  - Xmipp CL2D: 2D Classification
    - Downsampled to change pixel size from 0.6575 Å to 5.26 Å
    - Rationale: Computing speed up
  - PDB2MRC: Initial Model
- Relion

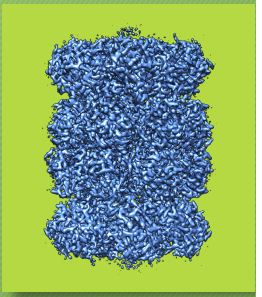
3) Data Processing



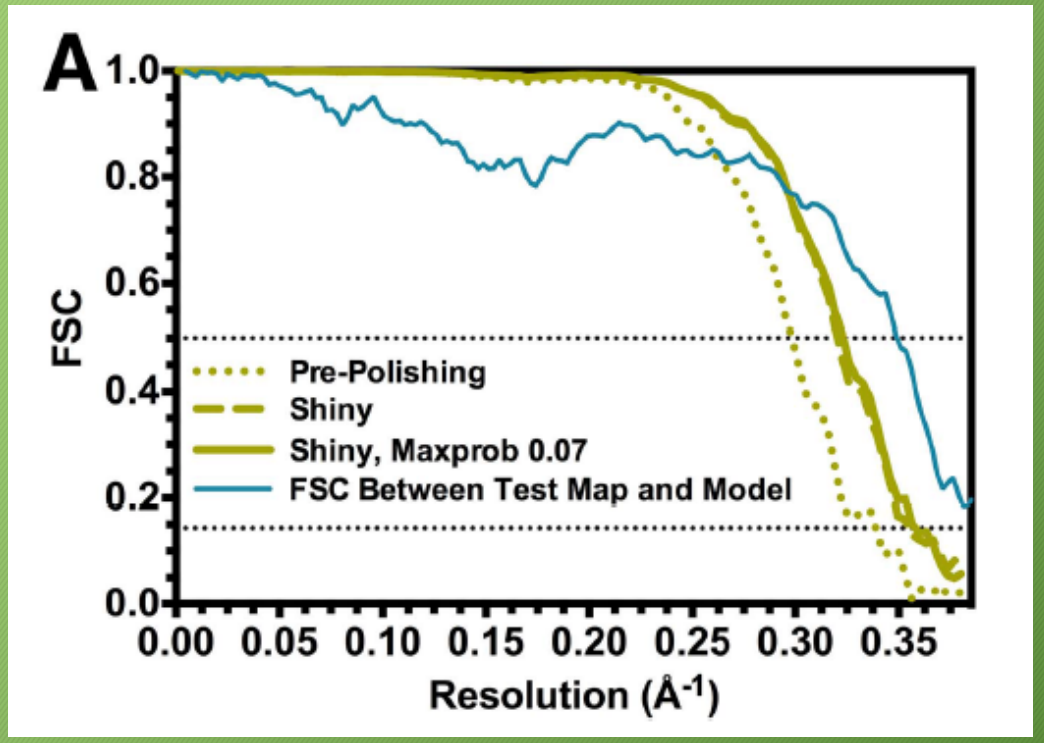
- Selection of Micrographs
  - By eye
    - Select squares with thin ice to collect
  - Computationally
    - Criteria: Cross-correlation coefficients (CC) between the 1-D radially averaged power spectrum of each micrograph and the calculated Contrast Transfer Function (CTF).
    - CC ≥ 80% at a resolution of 4 Å or better kept
    - 196 out of 985 selected
    - Part of Appion



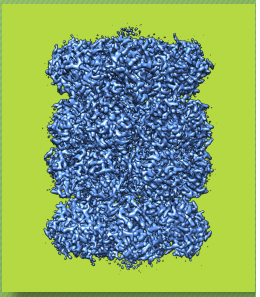
3) Data Processing



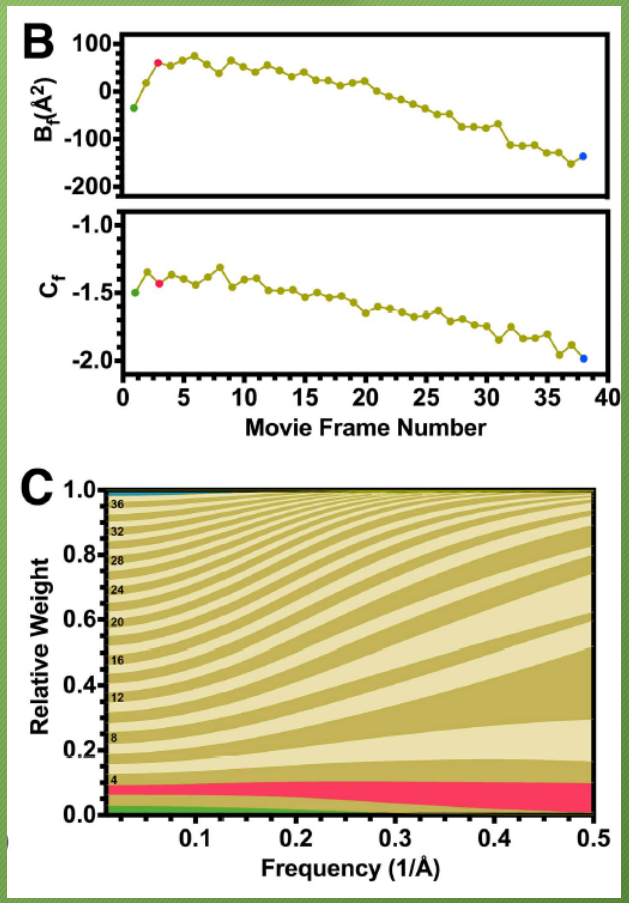
- Selection of Particles
  - 2D Classification
  - MaxProb from Relion
    - Value indicates uncertainty in both class and orientation assignments
    - 1/6 of particles discarded – Less is more



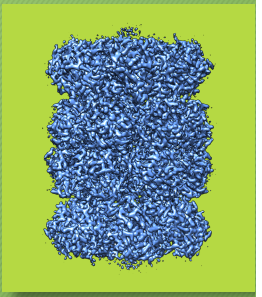
3) Data Processing



- Relion Particle Polishing
  - Correct individual beam-induced particle translations
  - Calculate and apply a frequency dependent weight for the contribution of individual movie frames to the reconstruction
    - More refined approach of dealing with different information content of frames
- Can done using experimental measured (Grant & Grigorieff 2015)
  - Which one is better?

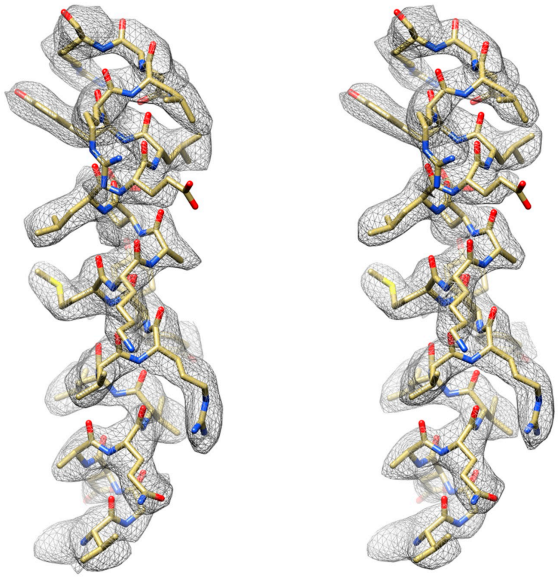


3) Data Processing

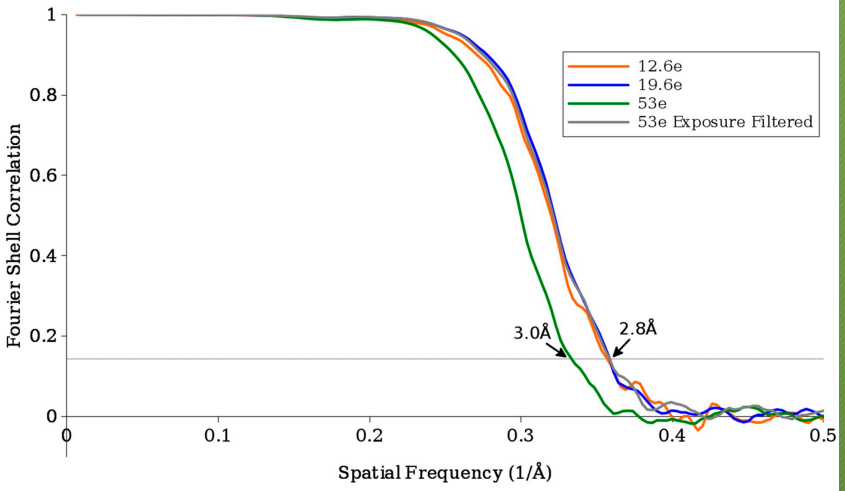


- Answer: Both works!

A



B

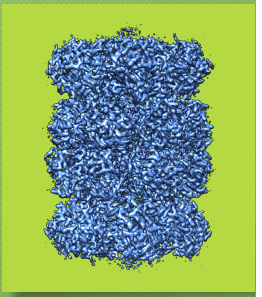


Relion + B-factor

Frealign + Exposure Weighting

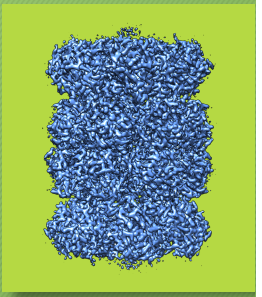
Grant & Grigorieff 2015

5) Model Building

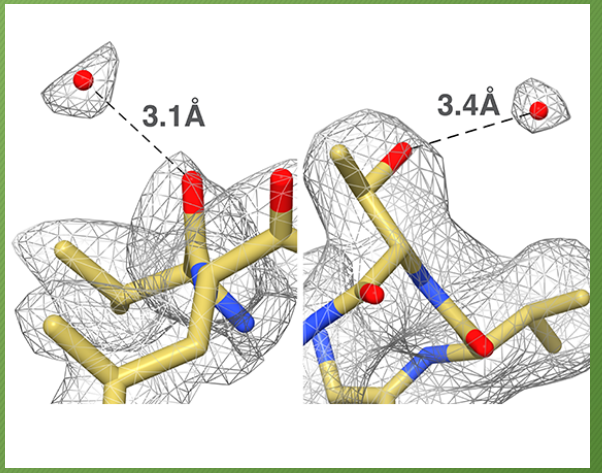


- UCSF Chimera: Known crystal structure docked in
- Rosetta: Atomic model refinement
  - Coot: Check model, add water models, iterative refinement
  - Refinement done against 1 half map, FSCs calculated against the other

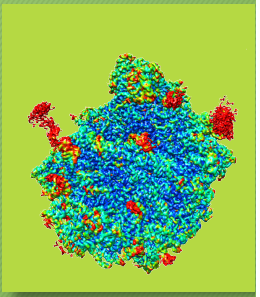
What was required to get to 2.8 Å and see water molecules



- Proteasome is a rigid and homogenous sample
- Relatively high electron dose
  - More signal for better particle alignment
- Mechanical stage movement used instead of beam-tilt at high magnification exposure
  - Avoid introducing phase shift
- Picking out thin ice by a trained eye
- Utilization of algorithmic advances
  - Projection matching, particle polishing



Case Study 3: 60S Ribosome



Contents lists available at ScienceDirect

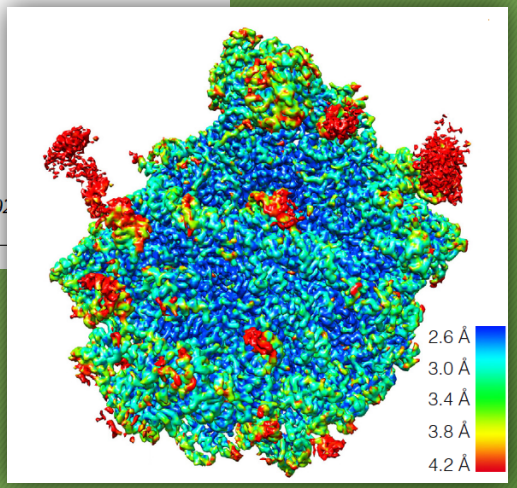
Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Single-particle cryoEM analysis at near-atomic resolution from several thousand asymmetric subunits

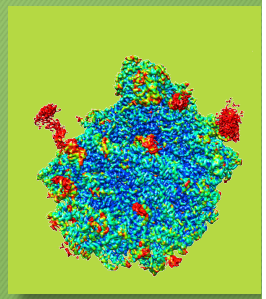
Dario Oliveira Passos, Dmitry Lyumkis *

Laboratory of Genetics and Helmsley Center for Genomic Medicine, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92



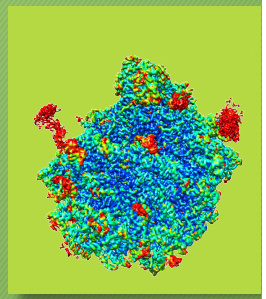
Information obtained from Passos & Lyumkis, 2015

1) Sample Preparation

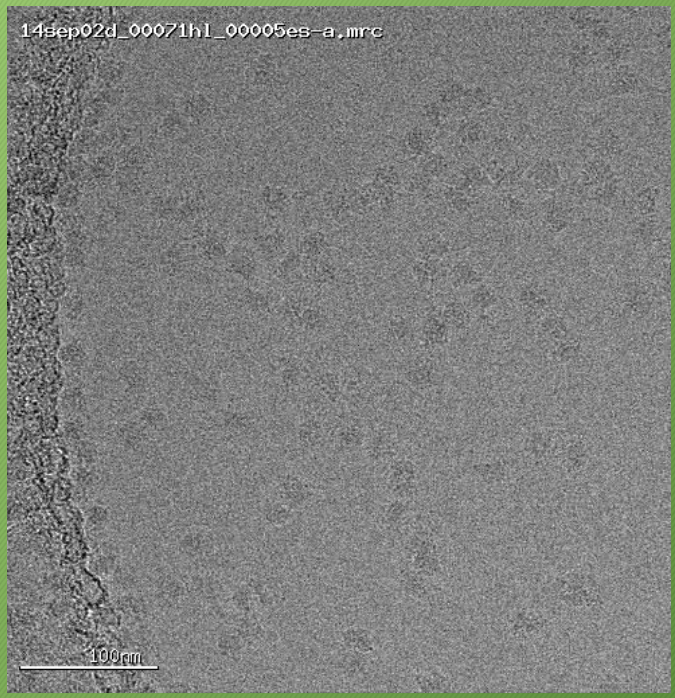


- 60S Ribosome Characteristics
  - Complex made of RNA and proteins
  - 2.5 MDa in size
  - No symmetry

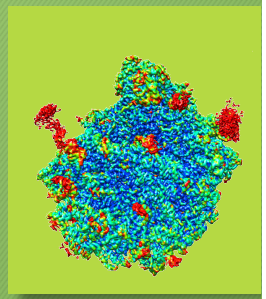
2) Cryo Data Collection



- Holey carbon grids – over holes
- Krios
  - 300 kV
  - K2 camera in super-resolution mode
    - Dose rate: 8.5 e-/pixel/s
    - Defocus range: 0.5 to 2.5 mm
- Leginon automated data collection

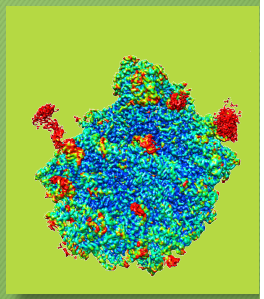


3) Data Processing



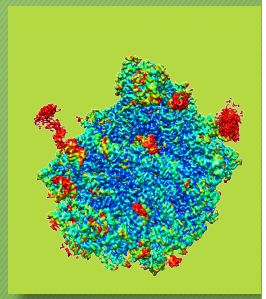
- Appion Package
  - Motioncorr: Whole frame alignment
    - Binned by 2 for more cost-effective processing
  - Done simultaneously
    - Manual masking
    - Manual particle picking (of ~300 particles)
    - CTFFind3: CTF estimation
  - Xmipp CL2D: 2D Classification for good class averages to use as templates
  - FindEM: Template particle picking

3) Data Processing



- Appion Package
  - Xmipp CL2D: Round 2 2D Classification for template picked particles
  - Optimod: Generate initial model using common lines method
  - Xmipp Reconstruction: Obtain initial angles for good particles
- Frealign
  - Single model refinement
  - 1 round of refinement → ~ 5 Å
  - 3 rounds of refinement → ~ 3 Å
  - 10 rounds of refinement → 2.9 Å

4) Model Building



- Looked at eL6, a protein missing from crystal structures
  - Rosetta: Atomic model refinement
  - Molprobity: Validation of atomic model
  - Map-to-model FSC: Validation

Resolution limitations

particle homogeneity

image quality

CTF

E_{det}

E_{tot}

alignment precision

E_{trans}
E_{rot}
E_{mag}
E_{det}
$\propto e^{-ck^2}$

Δt

particles averaged

$N_{views} = \pi D k$

$N_{images} = \pi D k \left(\frac{SNR_{desired}}{SNR_{preset}} \right)^2$

$N_{images} = \pi D k \left(\frac{SNR_{desired}}{SNR_{ideal\ image} \cdot E_{image\ quality} \cdot E_{alignment\ precision} \cdot E_{particle\ heterogeneity}} \right)^2$

10^3
10^4
10^7
$10^8 - 10^9$

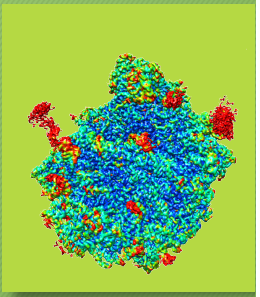
beam-induced specimen movement
modulation transfer function (MTF)
distortions, aberrations

$E_{image\ quality} = E_{sc} E_{tc}$

$E_{alignment\ precision} = E_{trans} E_{rot} E_{mag} E_{det}$

$\sigma_{trans} < 1\ \text{Å}$
$\sigma_{rot} < 1°$
$\sigma_{det} < 200\ \text{Å}$

Have I collected enough particles?

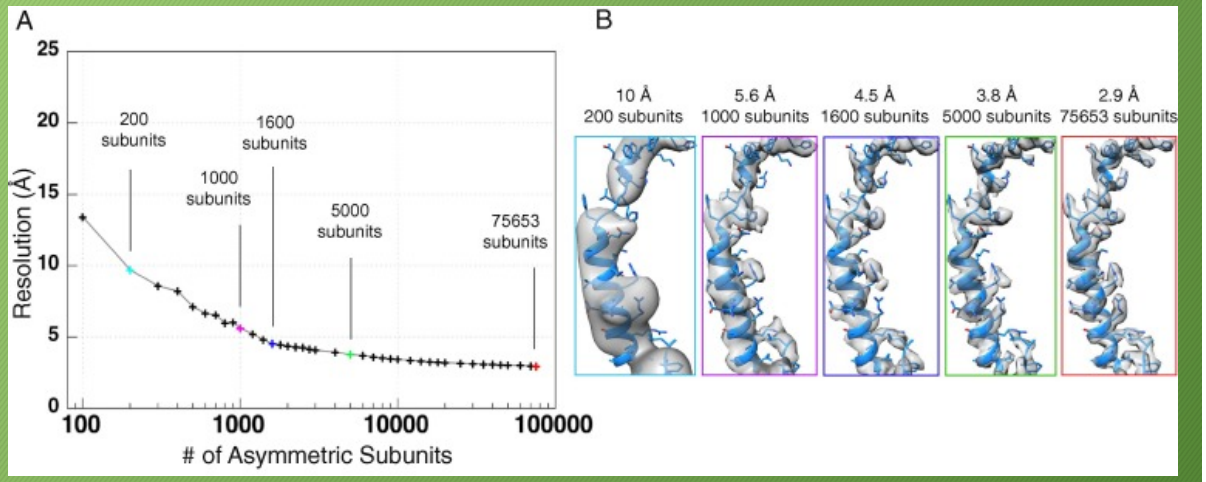


- Law of diminishing returns
  - Collecting more particles not always the answer
  - Extrapolate: 10x more particles to gain 0.1 Å

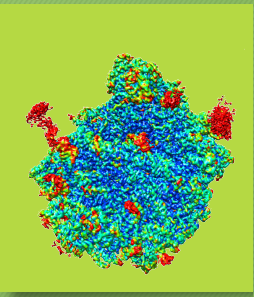
2010: 5.5 Å for 1.4 million 80S (Armache *et al*)

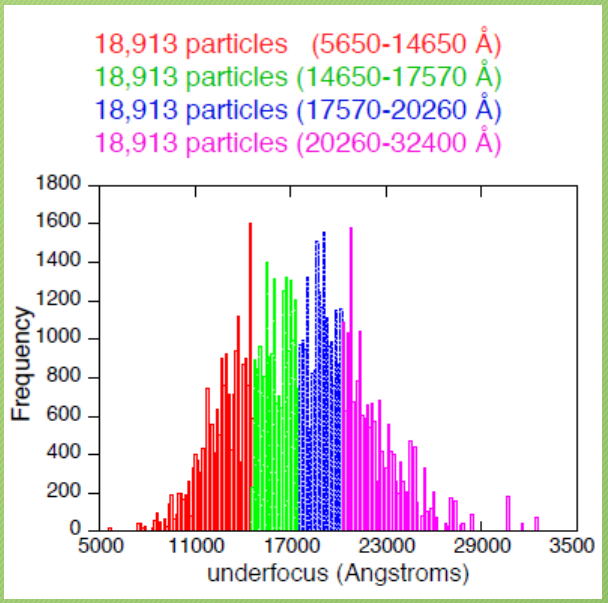
2013: 4.5 Å for 30,000 80S (Bai *et al*.)

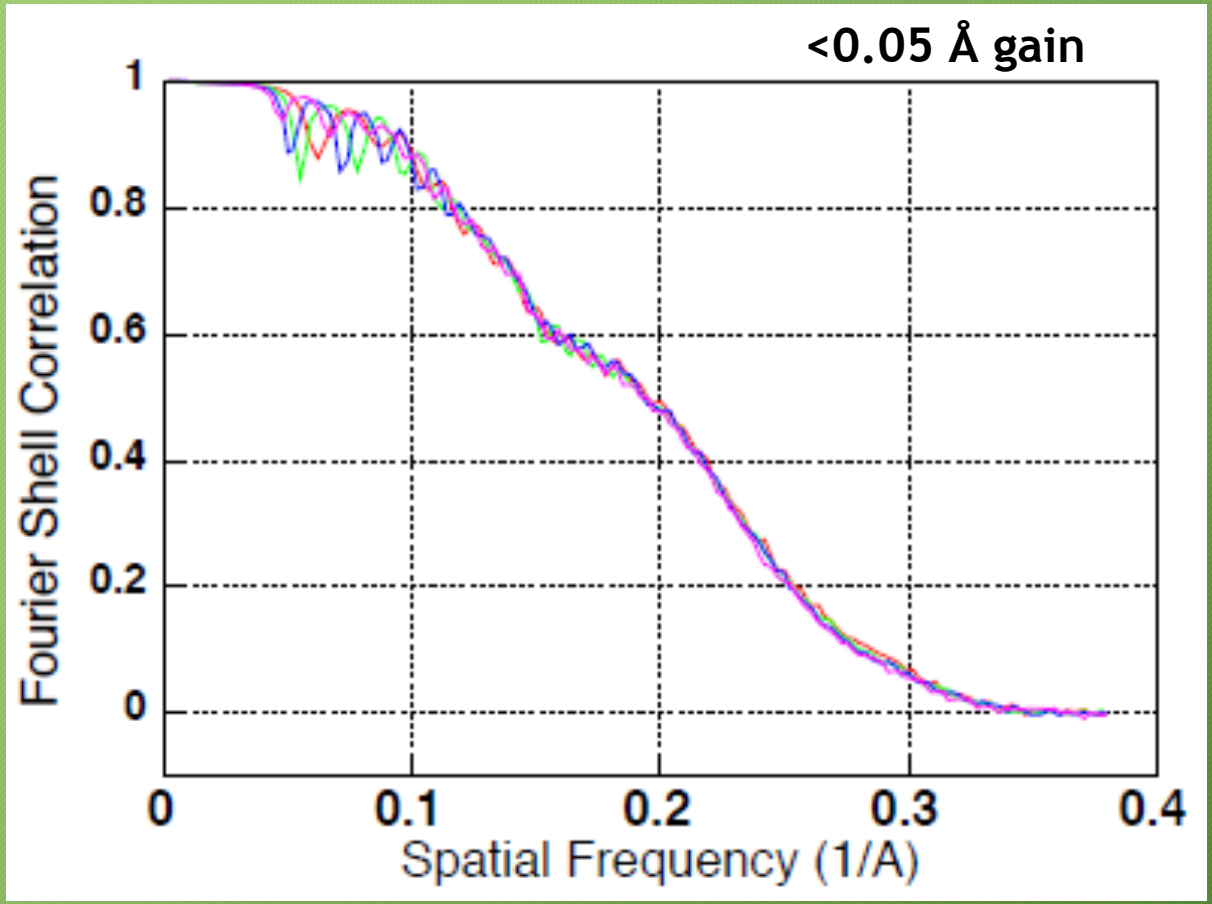
2015: 2.9 Å for 75,653 60S



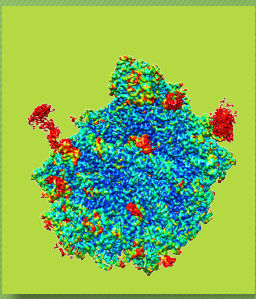
Defocus Matters… to an extent







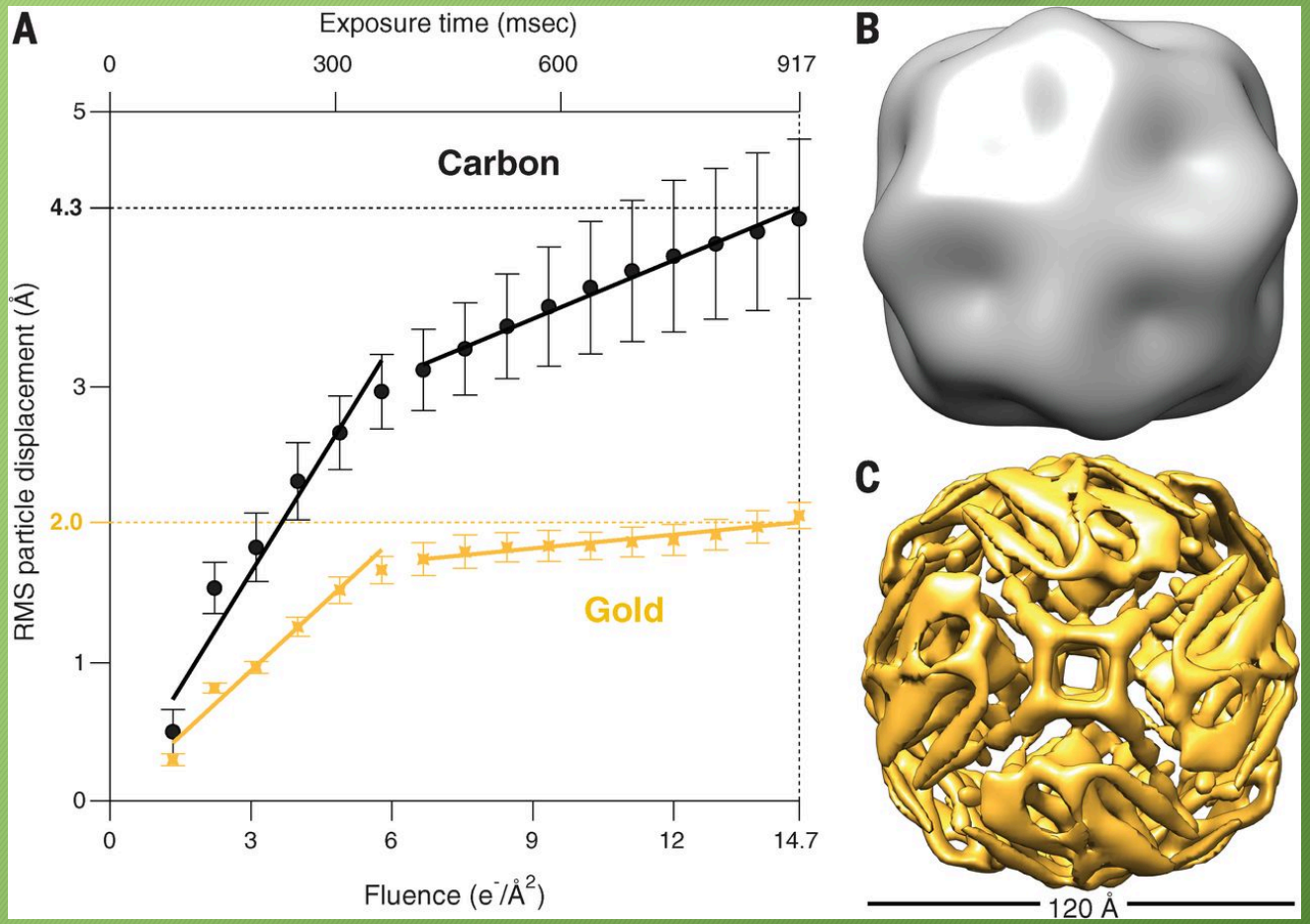
Structures under 3Å so far…



#	Resolution (Å)	Sample	Defocus Range (μm)
1	2.9	PCV2 (1.67MDa)	0.2 – 2.5
2	2.3	p97 with Inhibitor (0.54MDa)	0.7 – 2.5
3	2.4	p97 with ADP (0.54MDa)	0.7 – 2.5
4	2.9	60S (2.5MDa)	0.5 – 2.5
5	2.8	AAV-DJ (3.75MDa)	0.75 – 3.0
6	2.8	Proteasome (0.7MDa)	0.9 – 2.4
7	2.9	Cytoplasmic polyhedrosis virus with GTP (Not given)	Not given
8	2.2	Beta-galactosidase (0.465MDa)	0.6 – 2.0
9	2.9	Ribosome-EF-Tu complex (2.8 MDa)	0.7 – 2.5
10	2.9	Anthrax toxin pore (0.44 MDa)	1.8 – 5.1
11	2.6	VP6 Rotavirus (0.041 MDa)	0.4 – 2.0

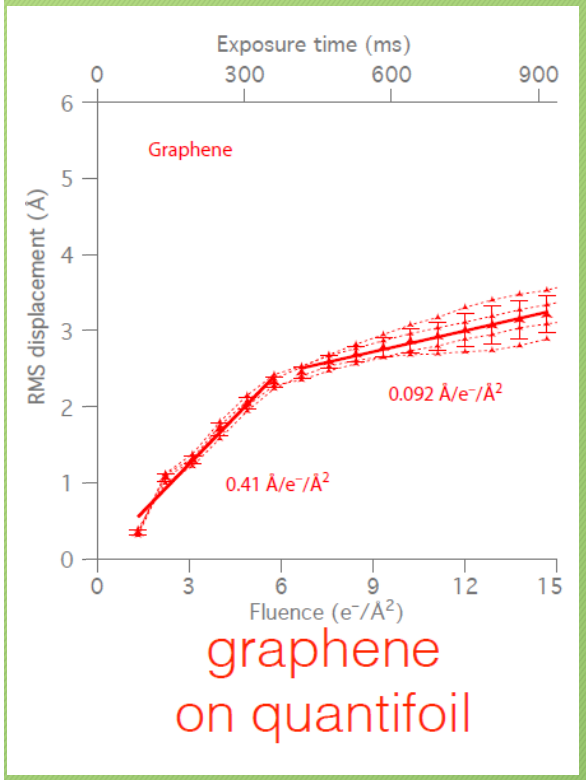
Further Advancement – Different Substrates Available

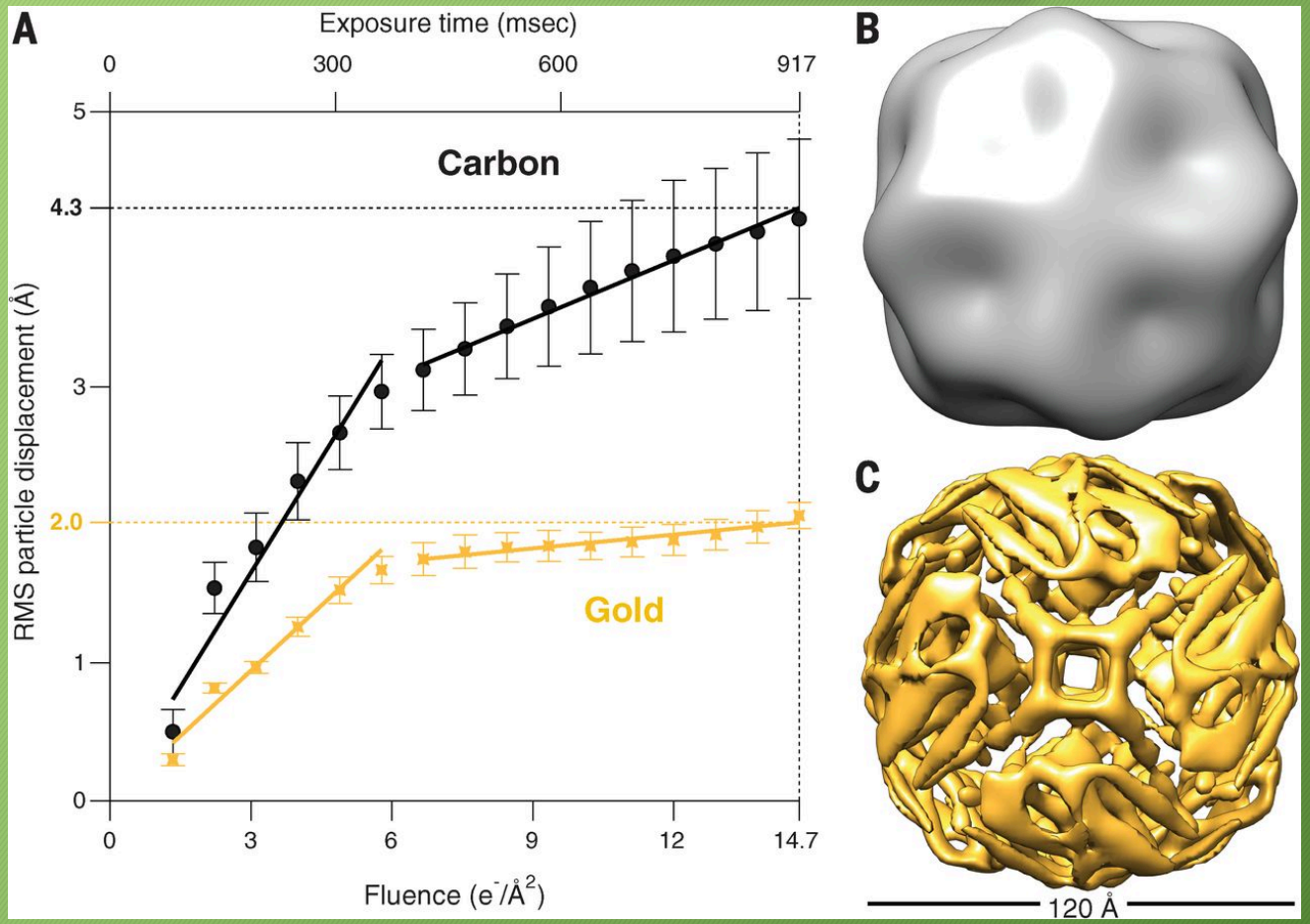
- Gold grids provide more stability compared to carbon grids



Russo & Passmore 2015

Further Advancement – Different Substrates Available



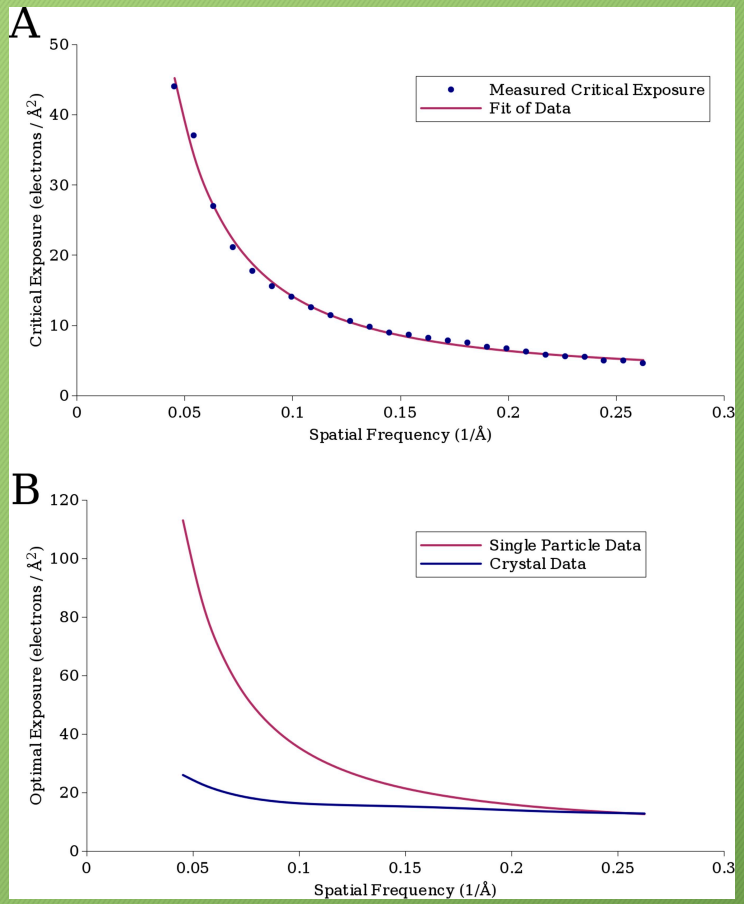


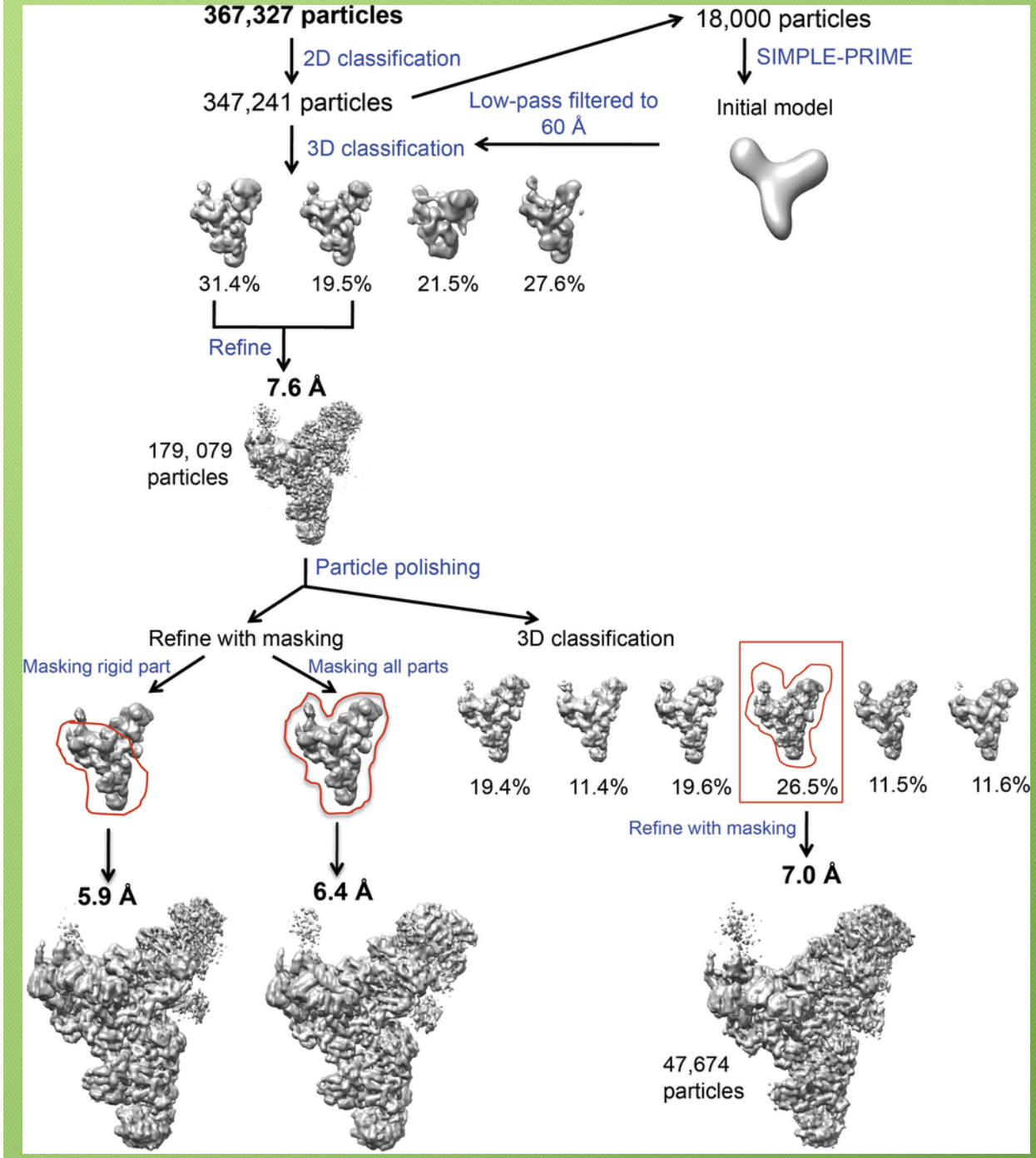
Russo & Passmore 2015

Thank you

Any questions?

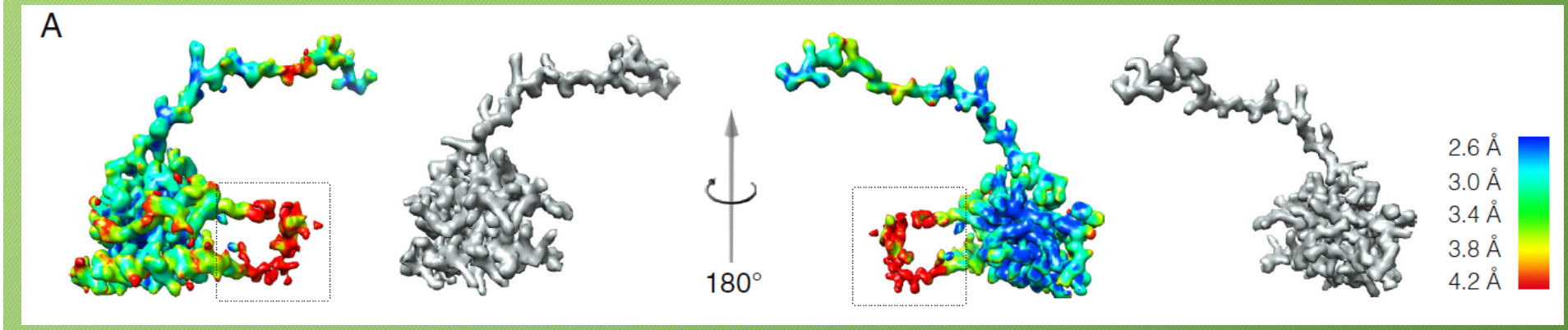
Exposure Weighting





The architecture of the spliceosomal U4/U6.U5 tri-snRNP

Nguyen *et al.*, 2015


A
180°
2.6 Å
3.0 Å
3.4 Å
3.8 Å
4.2 Å