Cryo-Electron Microscopy: Data Analysis and Reconstruction Workflow

David Jeruzalmi
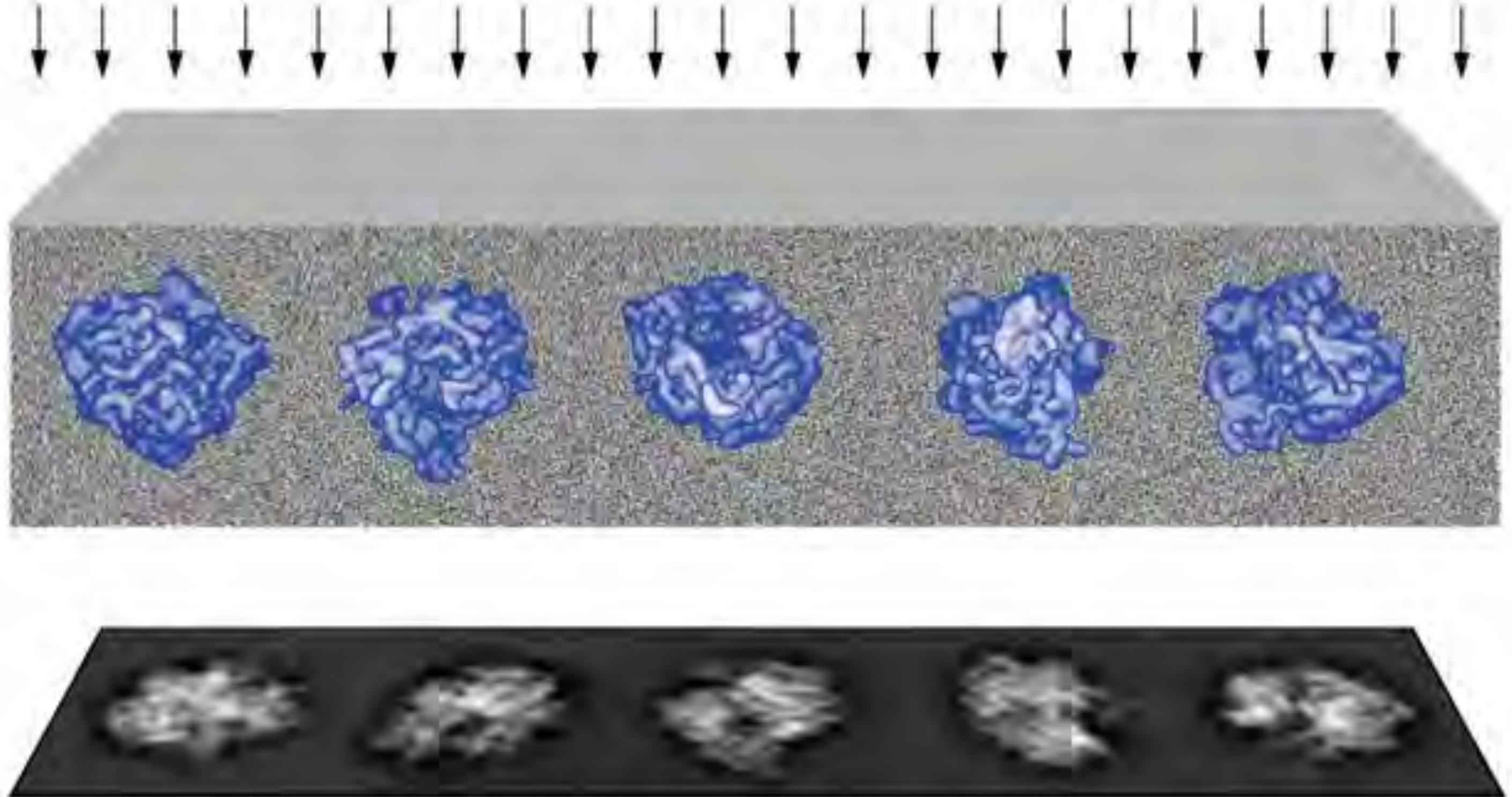
Chemistry & Biochemistry

City College of New York
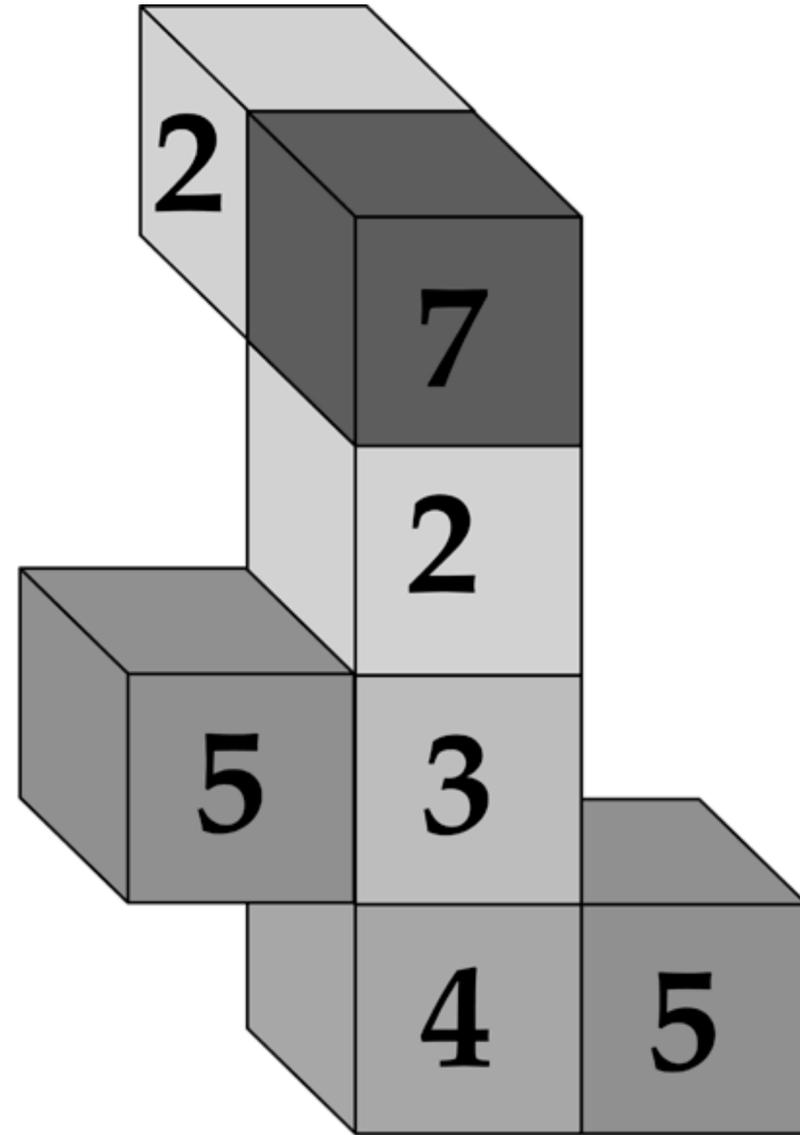
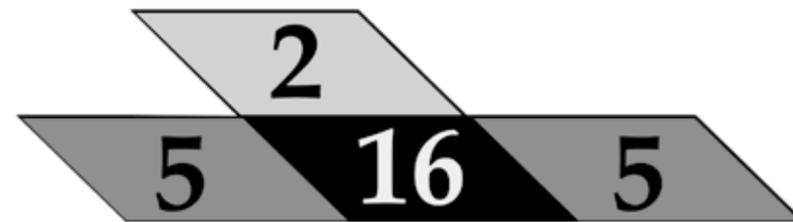Graduate Center of the City University of New York
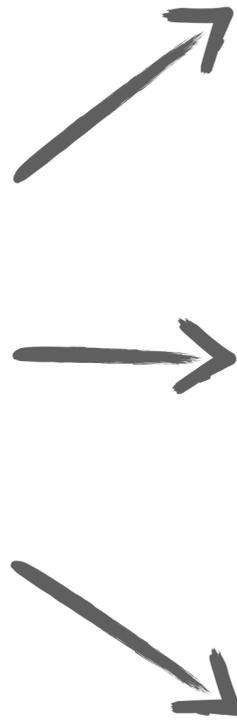
dj@ccny.cuny.edu

1

electrons

2

2026

'3D-Image'

2D-Projection

Three-dimensional structure

Thuman–Commike, FEBS Lett. 2001

Three-dimensional structure

A   Micrograph

Thuman-Commike, FEBS Lett. 2001

C Aligned particle images

A Micrograph

E Oriented class averages

Side   Top   Front

   2026
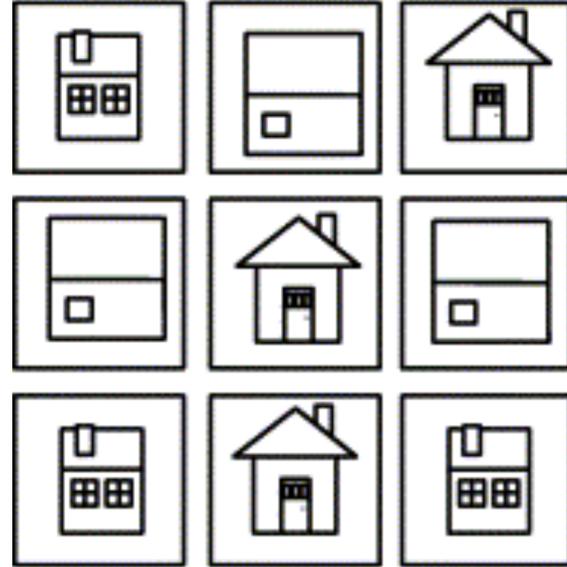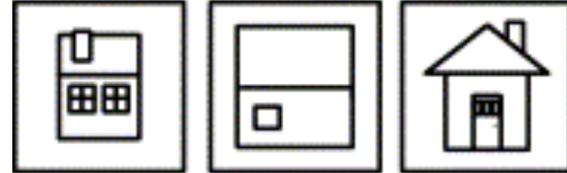
C   Aligned particle images

D   Classified particle images

Class 1   Class 2   Class 3

E   Oriented class averages

Side   Top   Front

F   Three-dimensional structure

2026

Particle Selection

Alignment

Visualization and Structure Interpretation

roof

door

windows

Three-Dimensional Reconstruction

Classification

Class 1   Class 2   Class 3

Determine Orientations

Side   Top   Front

Thuman-Commike, FEBS Lett. 2001

Three-dimensional structure

FSC=0.143

1.70Å

half-map FSC

resolution (1/Å)

Nakane, T. et al. Single-particle cryo-EM at atomic resolution. Nature **587**, 152–156 (2020).

Thuman–Commike, FEBS Lett. 2001

light source

electron gun

condenser lens

specimen

objective lens

eyepiece lens

projector lens

image viewed directly

viewing screen or photographic film
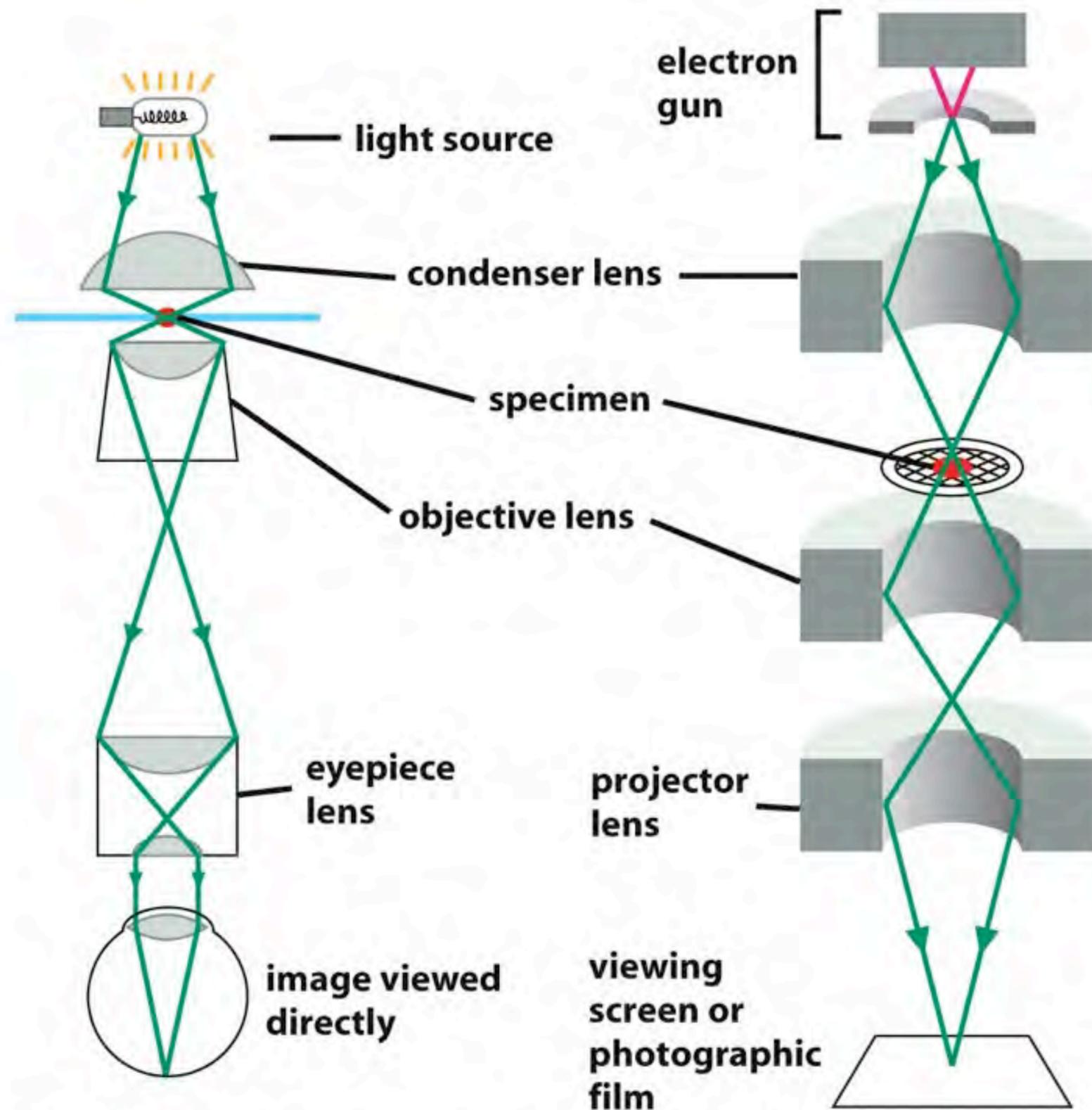
Figure 9-42 part 1 of 2 Molecular Biology of the Cell 5/e (© Garland Science 2008)

"That much was indeed at once obvious to the marvelously quick brain of Leo Szilard, who in 1928 suggested to me that I should make an electron microscope.

To this suggestion I gave the answer, which, I believe, would have been given by almost all physicists: "What is the use of it? Everything under the electron beam would burn to a cinder!"

"Let it burn, but let us look at the cinder."

Dennis Gabor, Preface to:
Marton, "Early History of the Electron Microscope", 1968, 1994

Shape/
Solvent

Fold/
Atom

$0.28 \times N_{atoms}$

larger
structure

lnF

$\sqrt{N_{atoms}}$

smaller
structure

Number of
Images

N=1

10

100

1000

10,000

100,000

1,000,000

Noise
Amplitude
$(F_{NOISE(N)})$

B_{image}

$B_{overall}=$
$B_{image} +$
$B_{computation}$

10 Å

resolution
limit

$1/d^2 \ \text{Å}^{-2}$

Dose-weighting

Per-particle
CTF correction

Focused classification
AI-based sharpening
AI-based model-building

Lomakin, J. Investig. Dermatol, 2025

13

2026

stack of movie frames

apply dose weighting scheme to each movie frame

early frame   intermediate frame   late frame

increasing radiation damage

stronger down-weighting of high frequencies

sum movie frames into single micrograph

https://cryoem101.org/chapter-5/

14

The Algorithm:
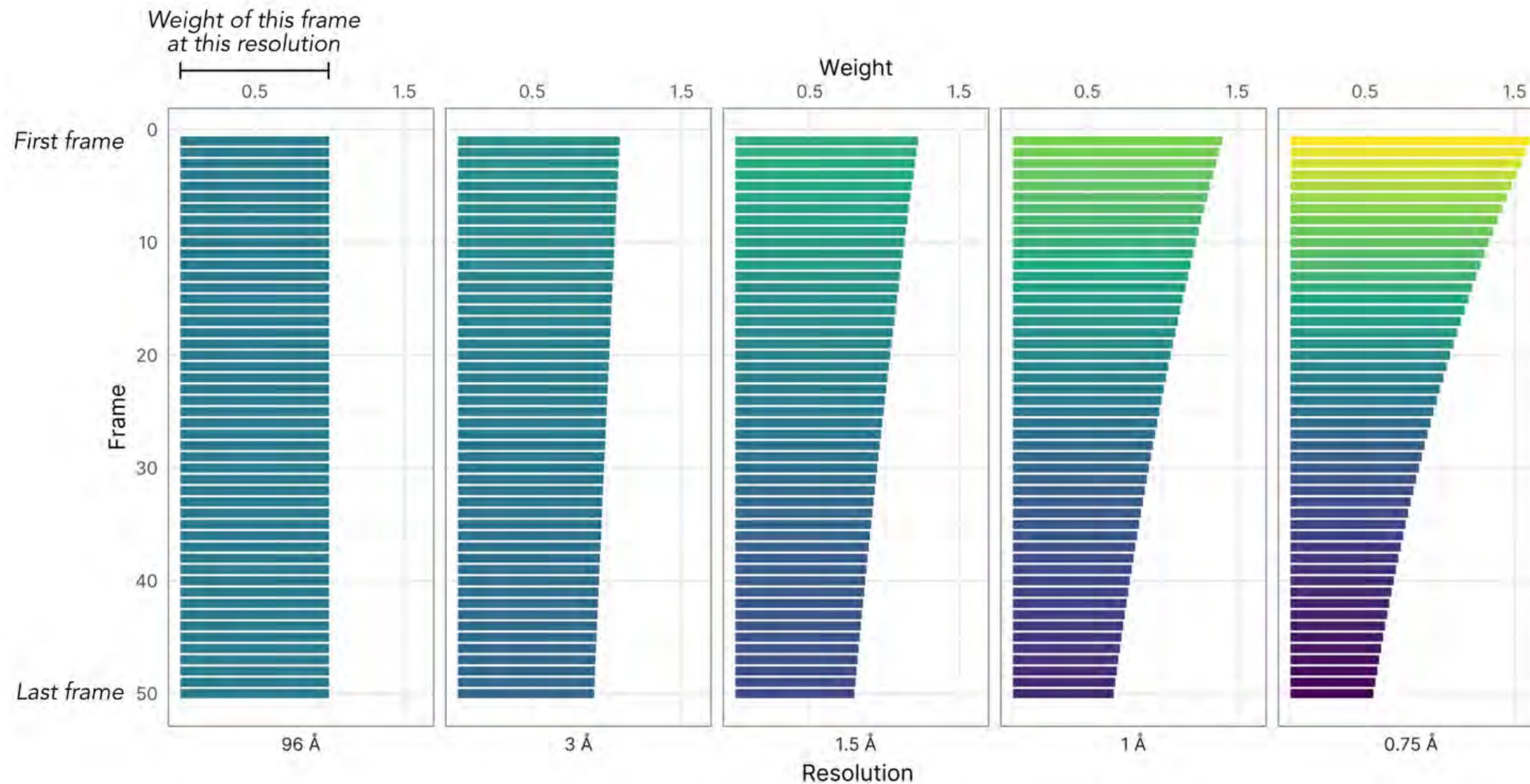Dose weighting applies a frequency-dependent filter, statistically prioritizing early frames for high-resolution details and later frames for low-frequency contrast.

Data Quality (y-axis)
Frame Number (Time) (x-axis)
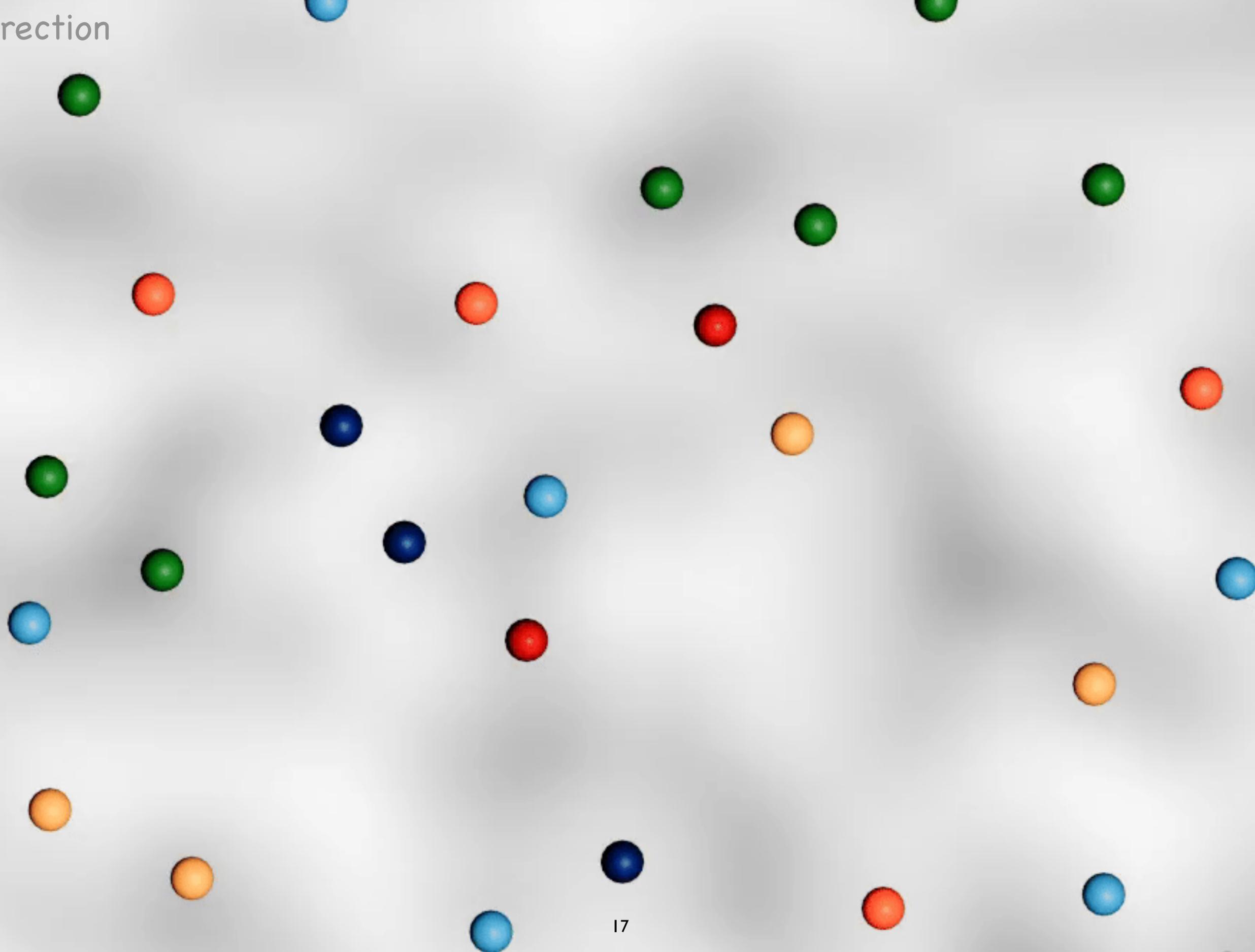Resolution
Motion

15
2026

Each frame of the movie receives a **weight** for each resolution. If all weights are **1.0** for a given resolution, then all frames contribute equally to the micrograph at that resolution.
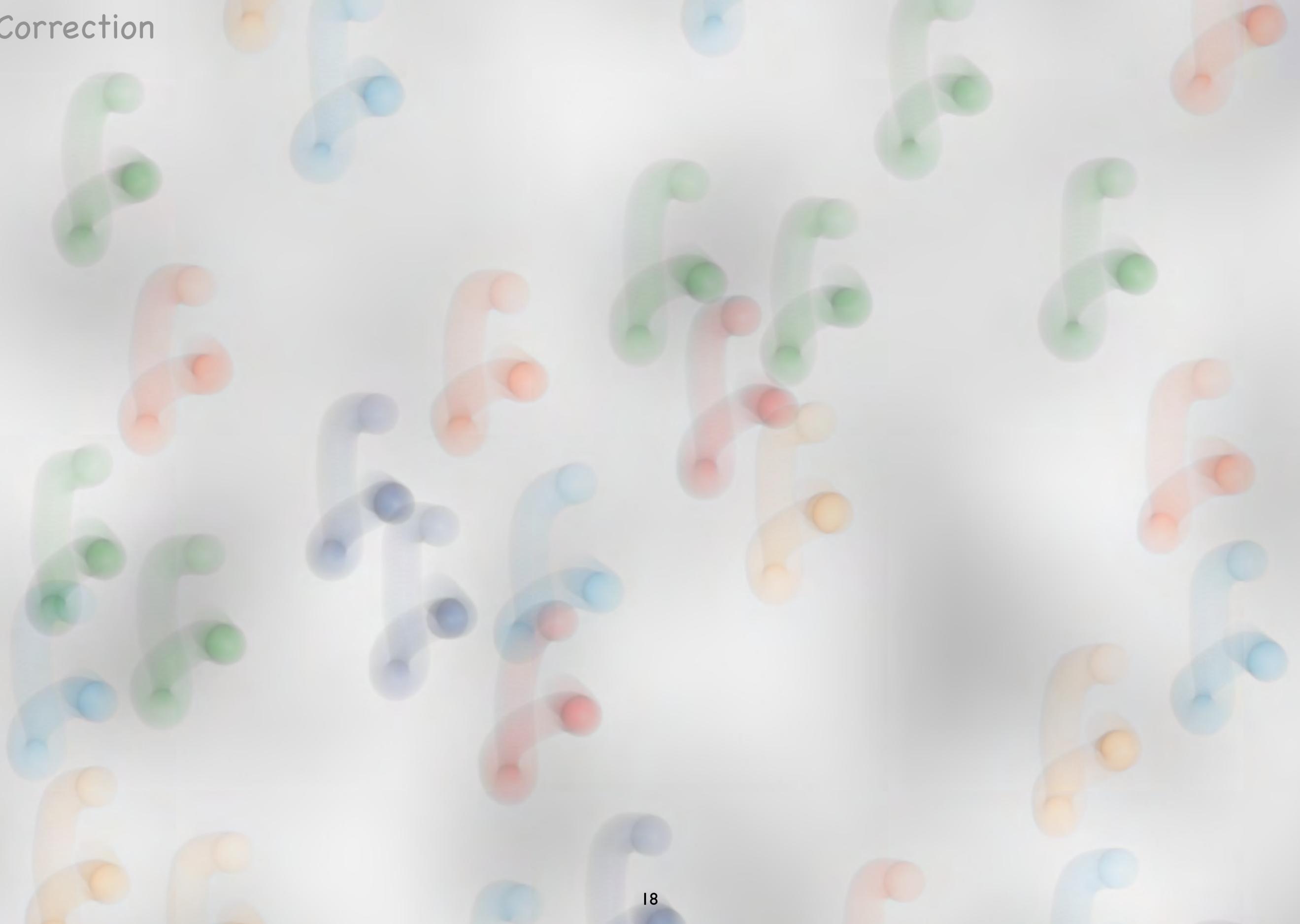
At high resolution, radiation damage becomes highly significant. Earlier frames therefore have **higher weights**, since they have been less damaged. Later frames have **lower weights** due to the significant radiation damage in those frames.

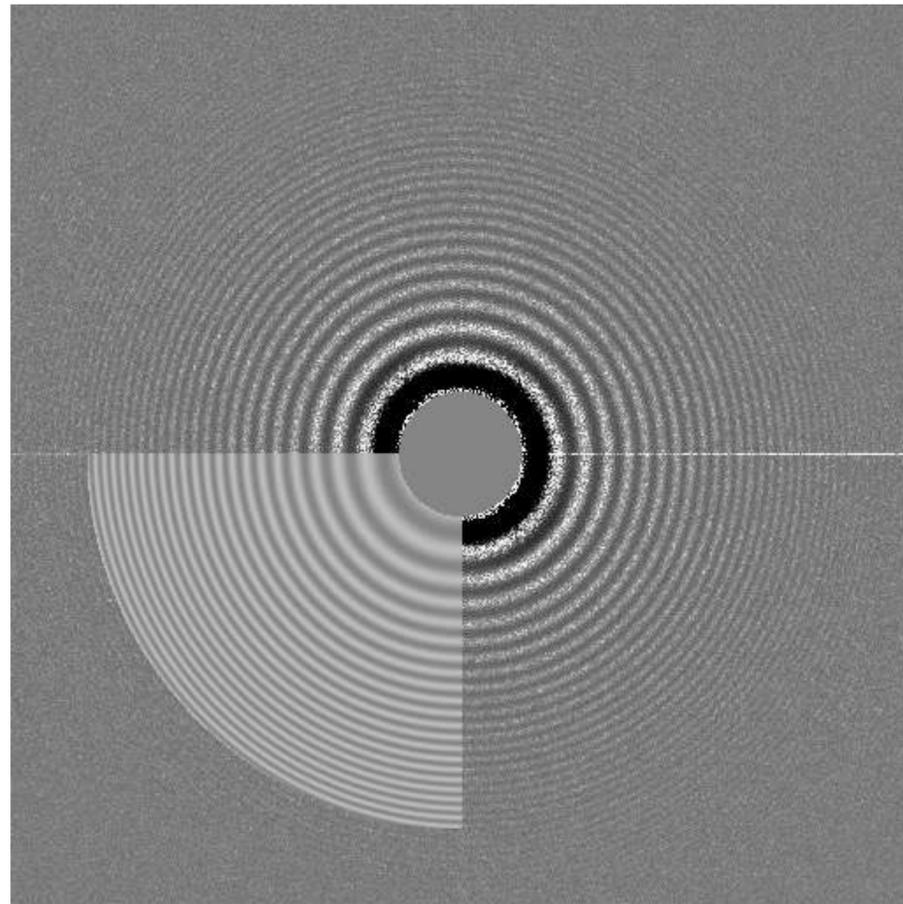https://guide.cryosparc.com/processing-data/all-job-types-in-cryosparc/motion-correction

- CTFFIND4
- Sparx/EMAN
- GCTF
- Warp
- CryoSPARC (patch method)

CTFFIND4

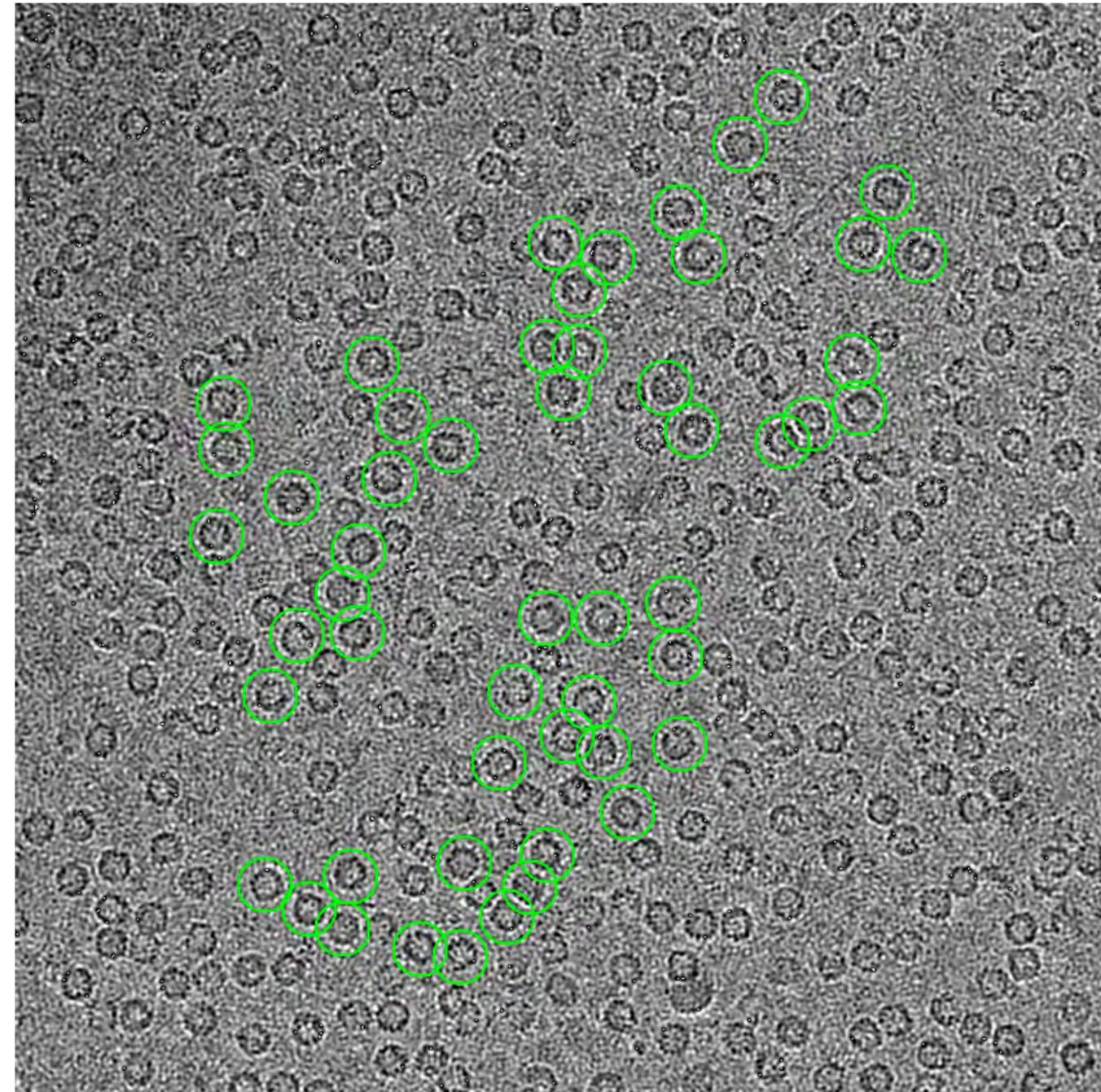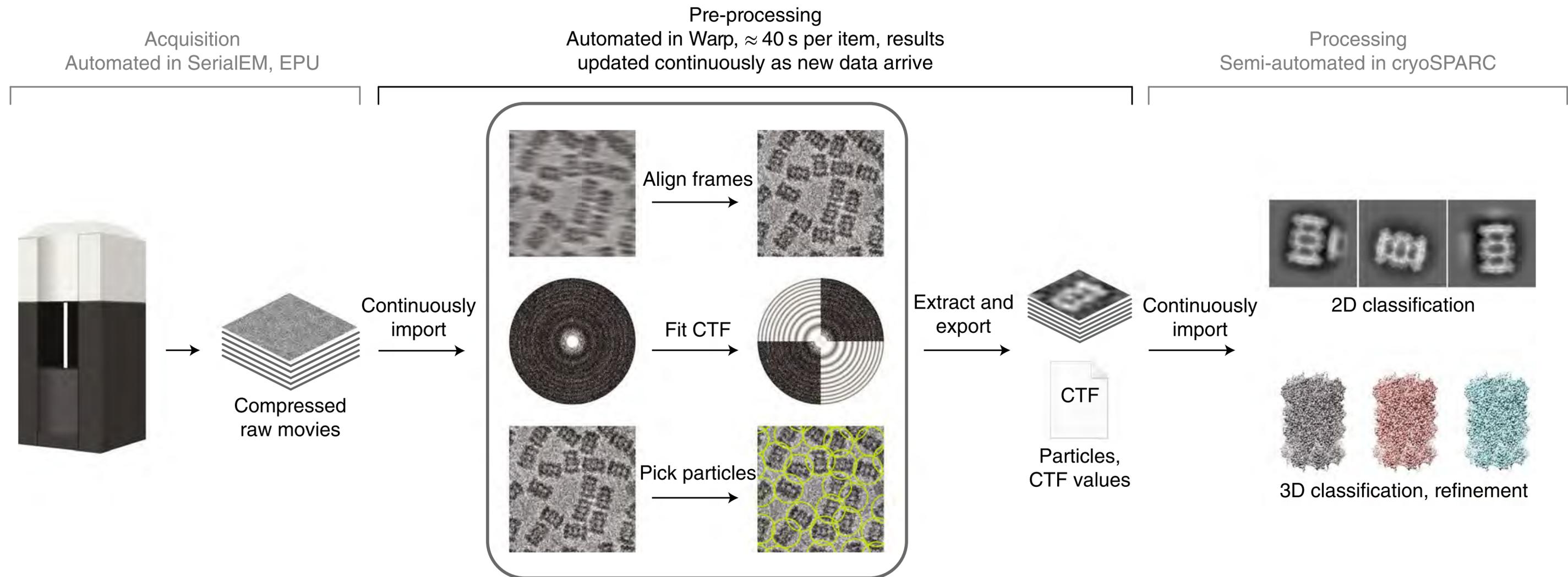DF1 15064.3 DF2 12356.3 ANGAST 0.037 PHASE 0.000 FIT 4.148 A
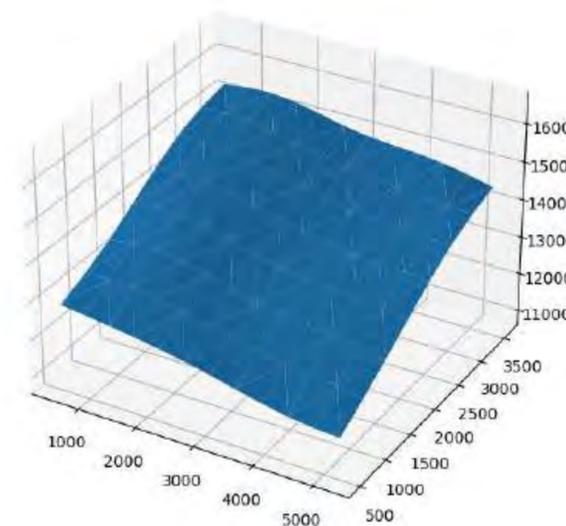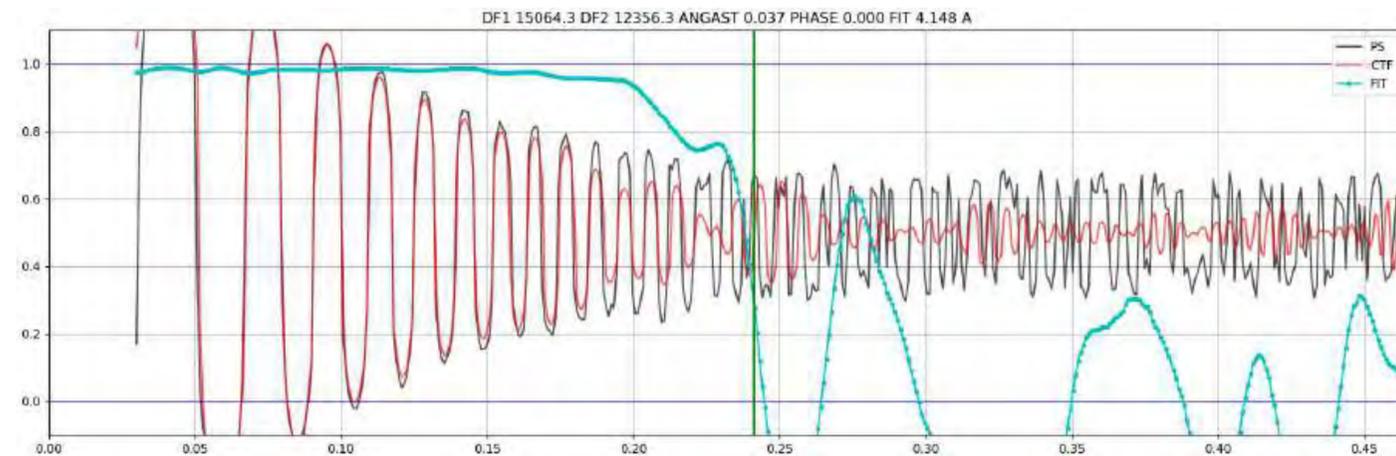
My stuff (CryoSPARC)

Manual – Blob – Template – Neural network

- Topaz
- Gautomatch
- Warp
- crYOLO
- FindEM
- Blob Picker
- DeepPicker
- DeepEM
- PIXER
- DRPnet
- DeepCryoPicker
- AutoCryoPicker
- *Start with provided model, get 2D classes, and retrain*

21

2026

Acquisition
Automated in SerialEM, EPU

Pre-processing
Automated in Warp, ≈ 40 s per item, results
updated continuously as new data arrive

Processing
Semi-automated in cryoSPARC

Compressed
raw movies

Continuously
import

Align frames

Fit CTF

Pick particles

Extract and
export

CTF

Particles,
CTF values

Continuously
import

2D classification

3D classification, refinement

Tegunov, D. & Cramer, P. Real-time cryo-electron microscopy data preprocessing with Warp. *Nat Methods* **16**, 1146–1152 (2019).

Bepler, T., Kelley, K., Noble, A. J. & Berger, B. Topaz-Denoise: general deep denoising models for cryoEM and cryoET. *Nat. Commun.* **11**, 5208 (2020).
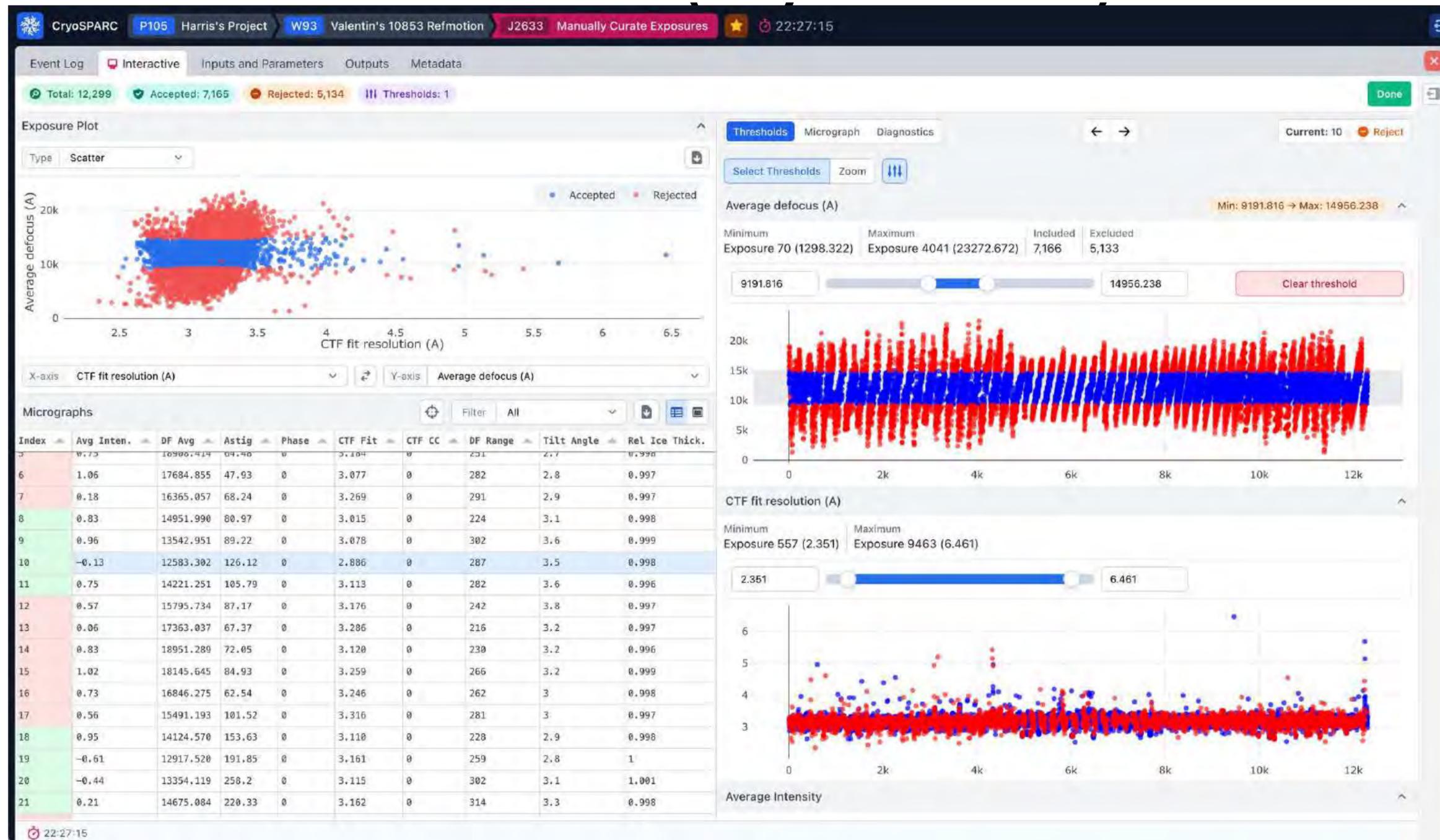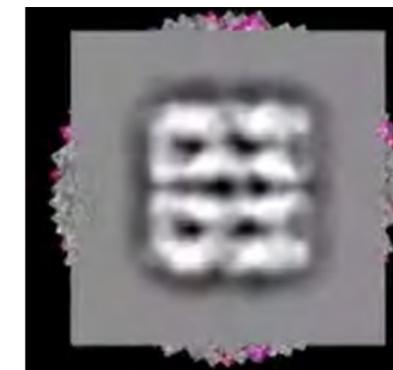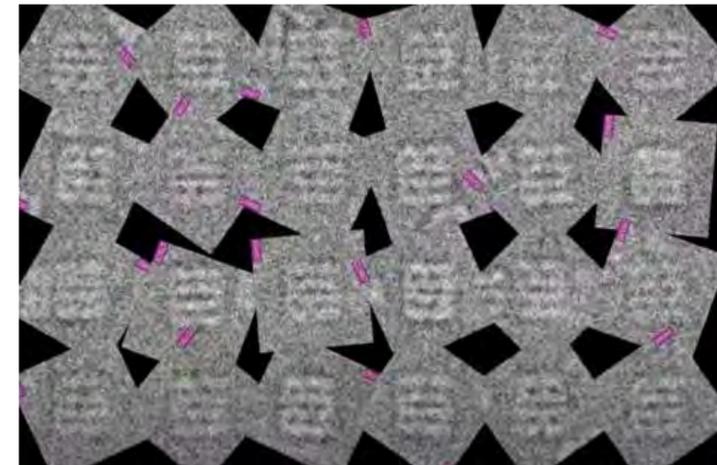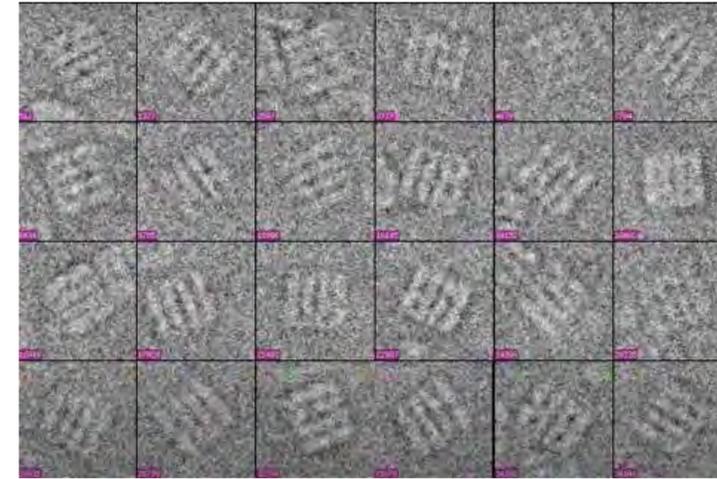
- Defocus range
- CTF fit resolution
- Number of particles

- Tilt angle
- Astigmatism
- Ice thickness

23

- cryoSPARC
- Relion
- Sparx/EMAN2
- ISAC
- Spider
- Simple
- *Remove "bad" particles*

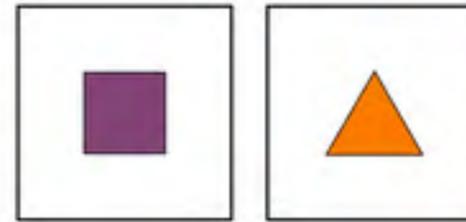*Lander 2009*
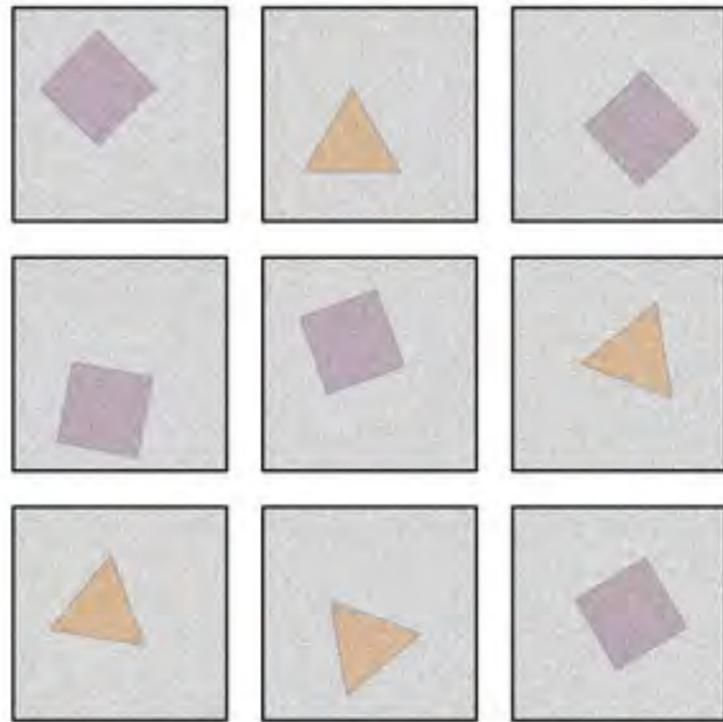
Align, classify, and average

'Initial cleanup' of particles

Particles are only rotated and translated in-plane
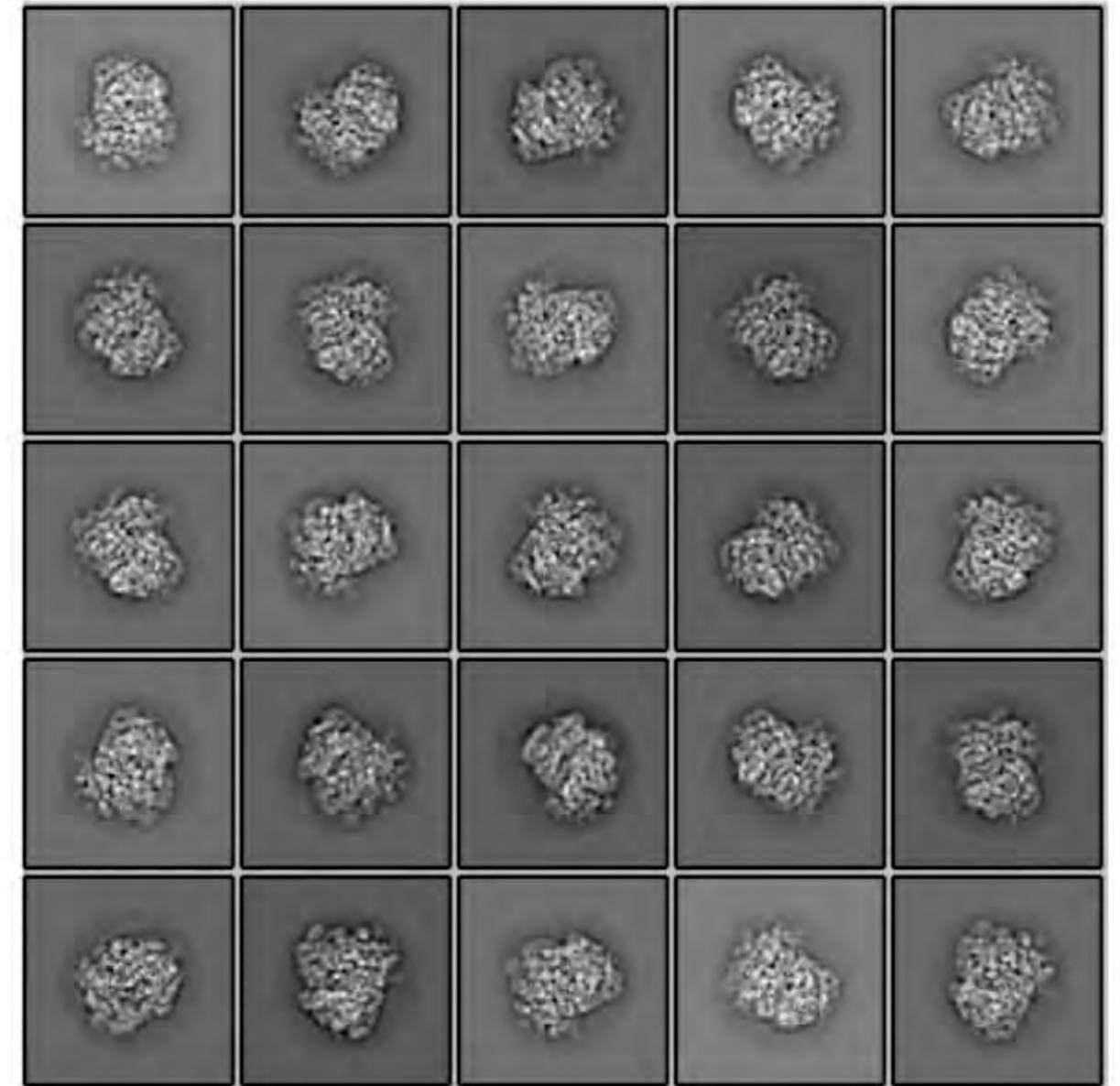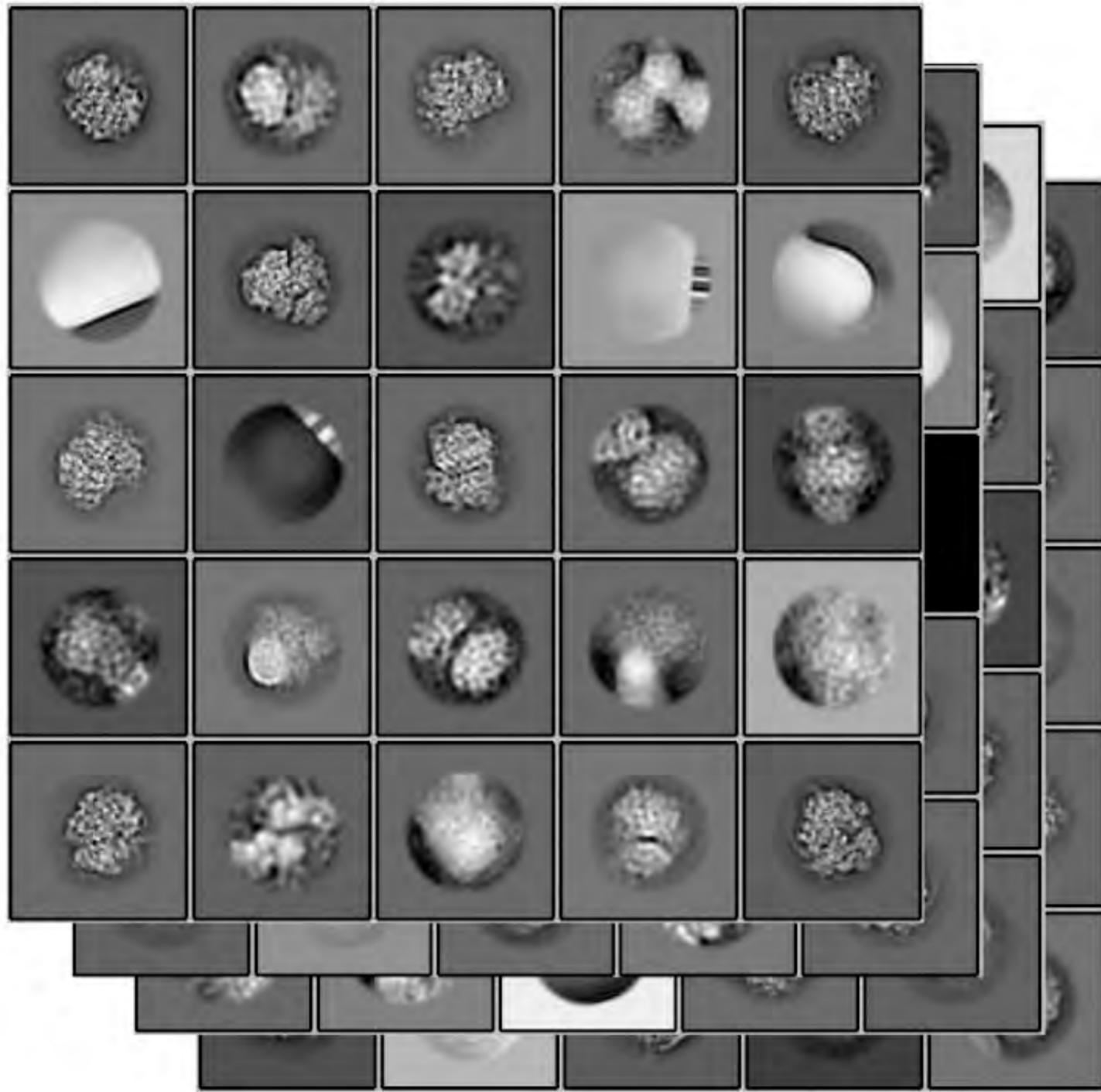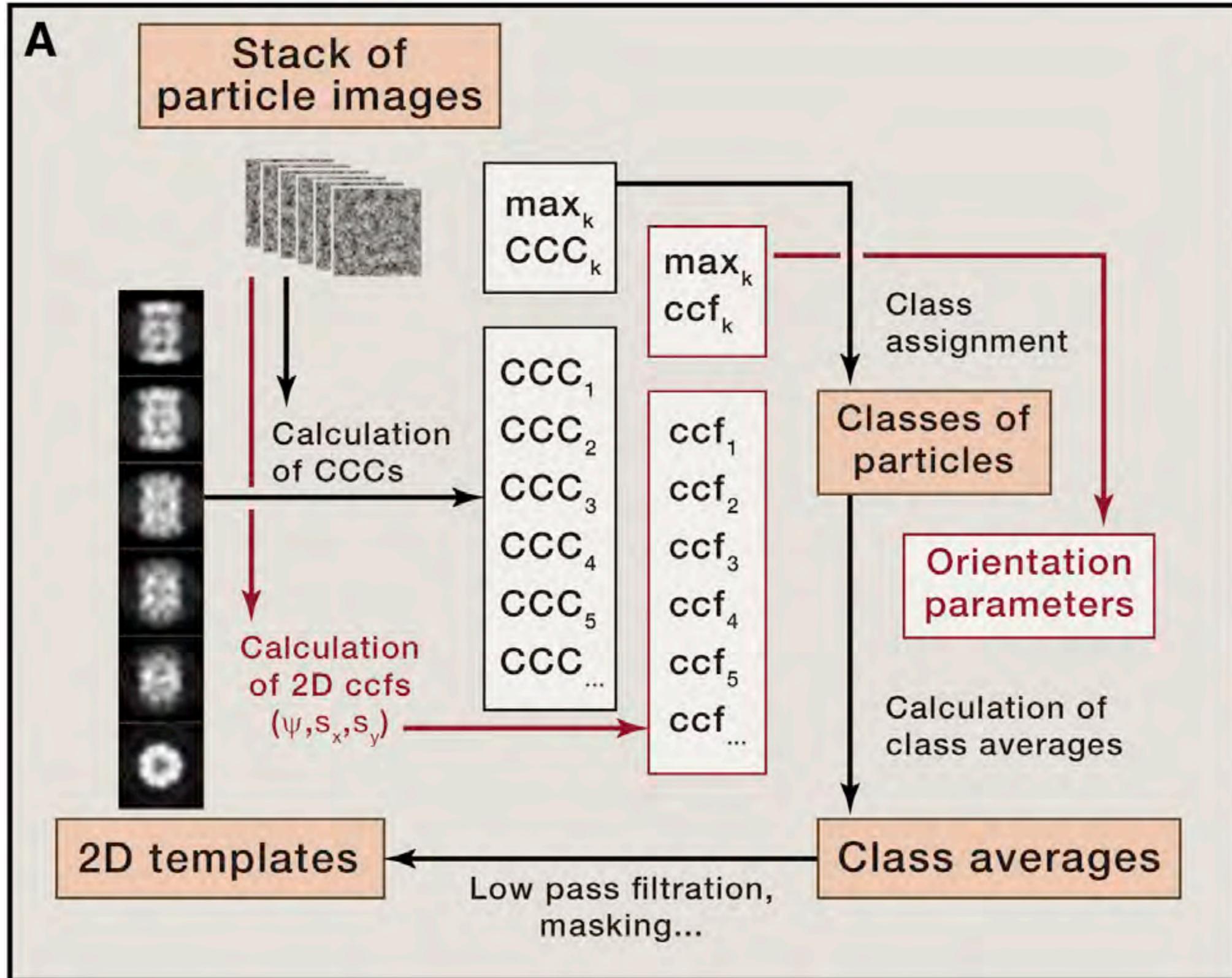
Class average recovered

Inputs: (CryoSPARC)

Number of classes

Maximum resolution
Maximum alignment resolution
Minimum alignment resolution
Plotting sort method
Initial classification uncertainty factor
Circular mask diameter (A)

- Anticipate 100 to 400 particles per class
- Don't ask for too many classes
- I split my particle stack into stacks of 100K particles and process each separately to get clean-vs-dirty particles
  - Screen through various values for radius
  - Relion
    - Tau fudge
    - CTF
  - cryoSPARC
    - Turn off Force Max over poses/shifts
    - Initial classification uncertainty factory (2 and above)
    - Number of iteration to anneal sigma as high as 25
    - Set online-EM iterations to 40
    - Set Batchsize per class to 400
    - Change Re-center mask threshold (possibly as high as 0.75) for centering particles and smearing neighbors
    - set White noise model to off

https://guide.cryosparc.com/processing-data/all-job-types-in-cryosparc/particle-curation/job-2d-classification
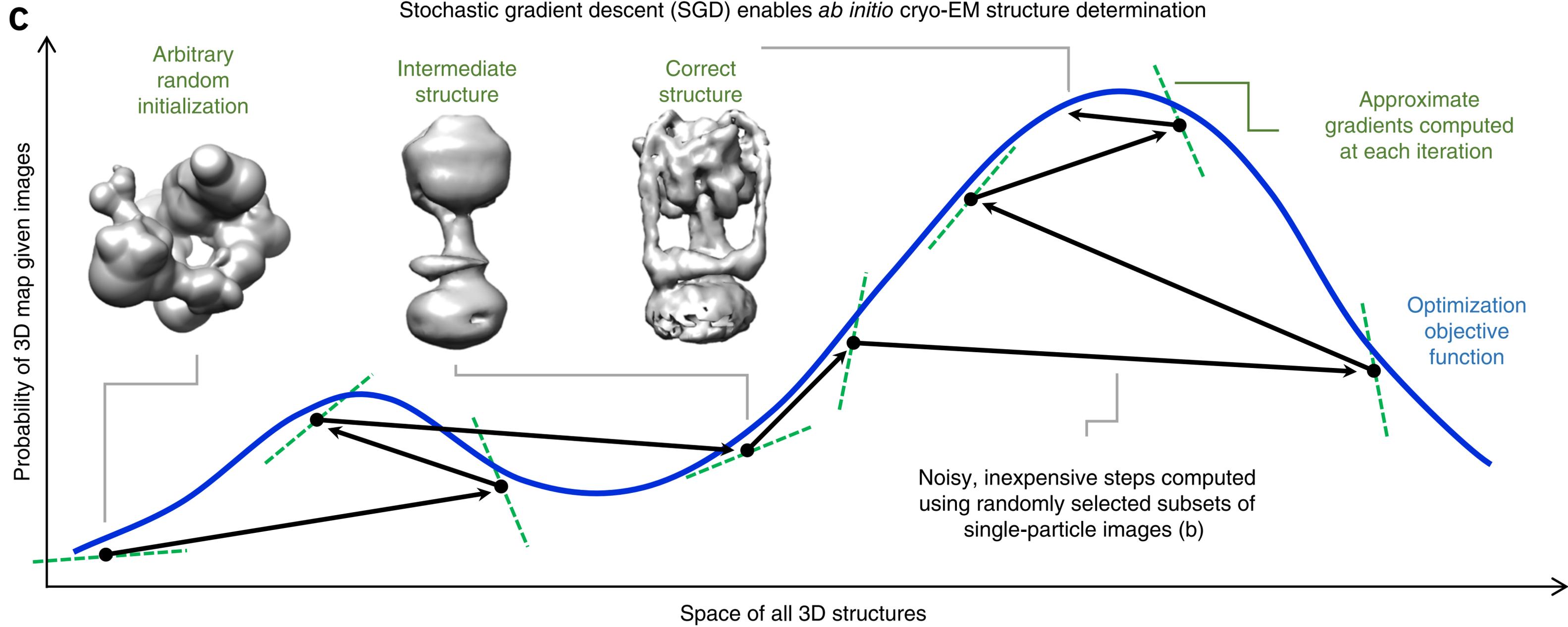
2026

Cheng, Y., Grigorieff, N., Penczek, P. A. & Walz, T. A Primer to Single-Particle Cryo-Electron Microscopy. Cell 161, 438–449 (2015).

- Random conical tilt

- Orthogonal conical tilt

- Common-lines

- Tomography with STA

- Random initial parameters, optimize with stochastic gradient descent (SIMPLE, cryoSPARC, and Relion).

- SAXS/SANS

- Structure prediction (calculate map of PDB)

- Generate multiple initial models if uncertain in model
  - – Look for continuity in density
  - – Look for sausages to indicate α-helices
  - – Are projections comparable to class averages?
- Ask for multiple models to be generated
- CryoSPARC's starting frequency should have more information than particle_size / 5 (e.g. 300 / 5 = 60Å)
- Use C_1 symmetry

c

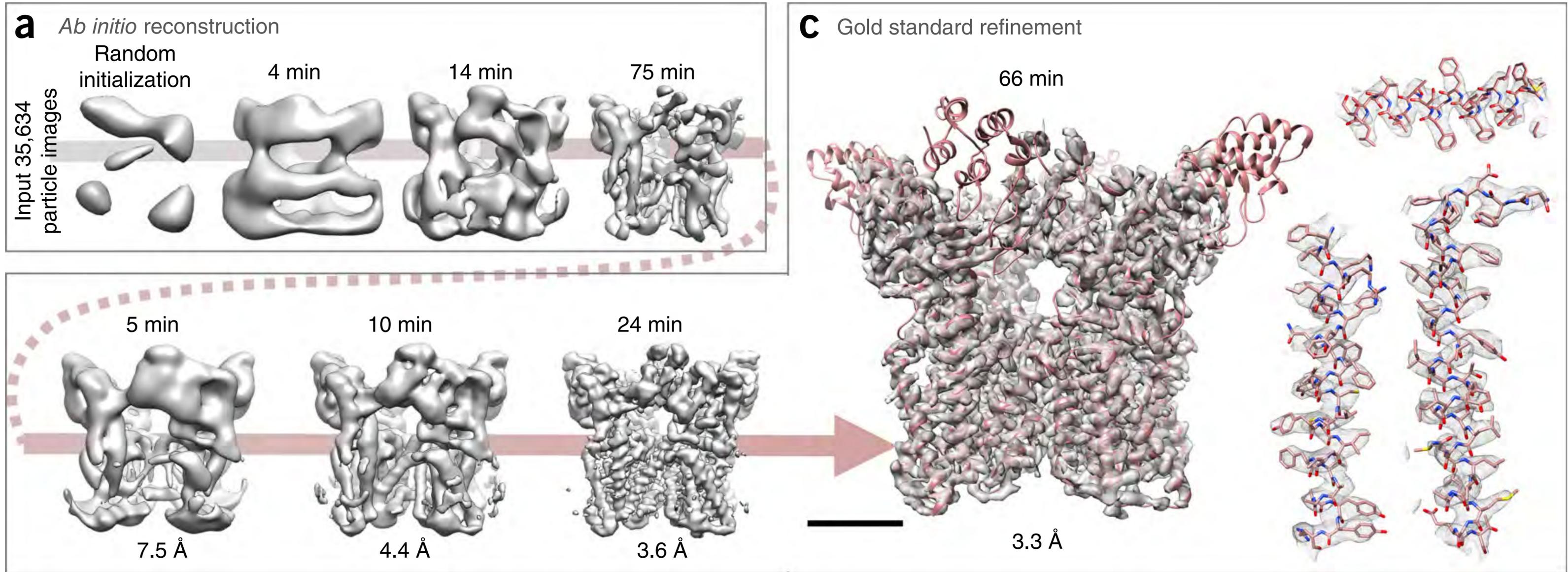Stochastic gradient descent (SGD) enables *ab initio* cryo-EM structure determination

Probability of 3D map given images

Arbitrary
random
initialization

Intermediate
structure

Correct
structure

Approximate
gradients computed
at each iteration

Optimization
objective
function

Noisy, inexpensive steps computed
using randomly selected subsets of
single-particle images (b)
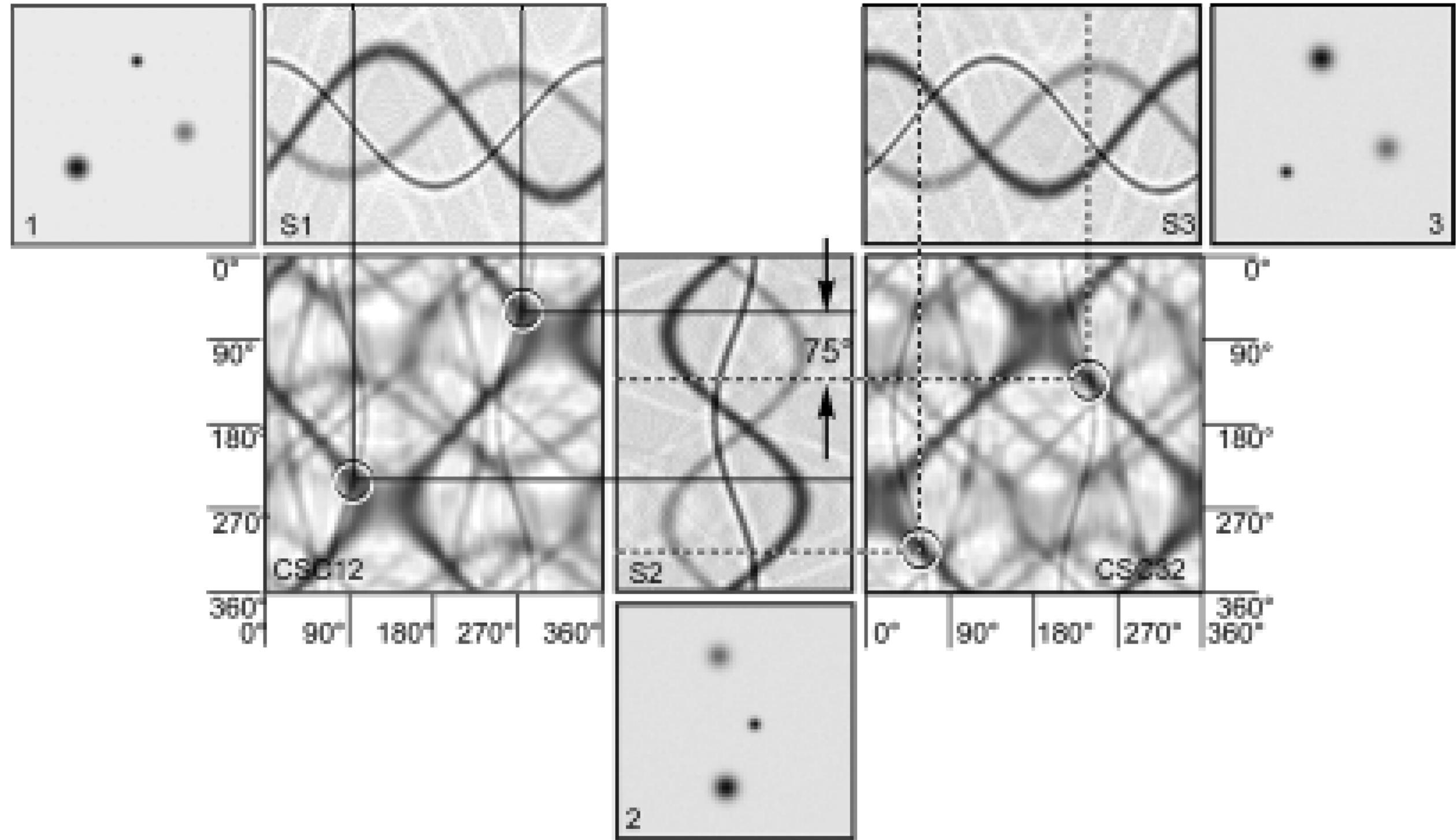
Space of all 3D structures

Punjani, A., Rubinstein, J. L., Fleet, D.
J. & Brubaker, M. A. cryoSPARC:
algorithms for rapid unsupervised
cryo-EM structure determination.
Nat. Methods 14, 290–296 (2017).

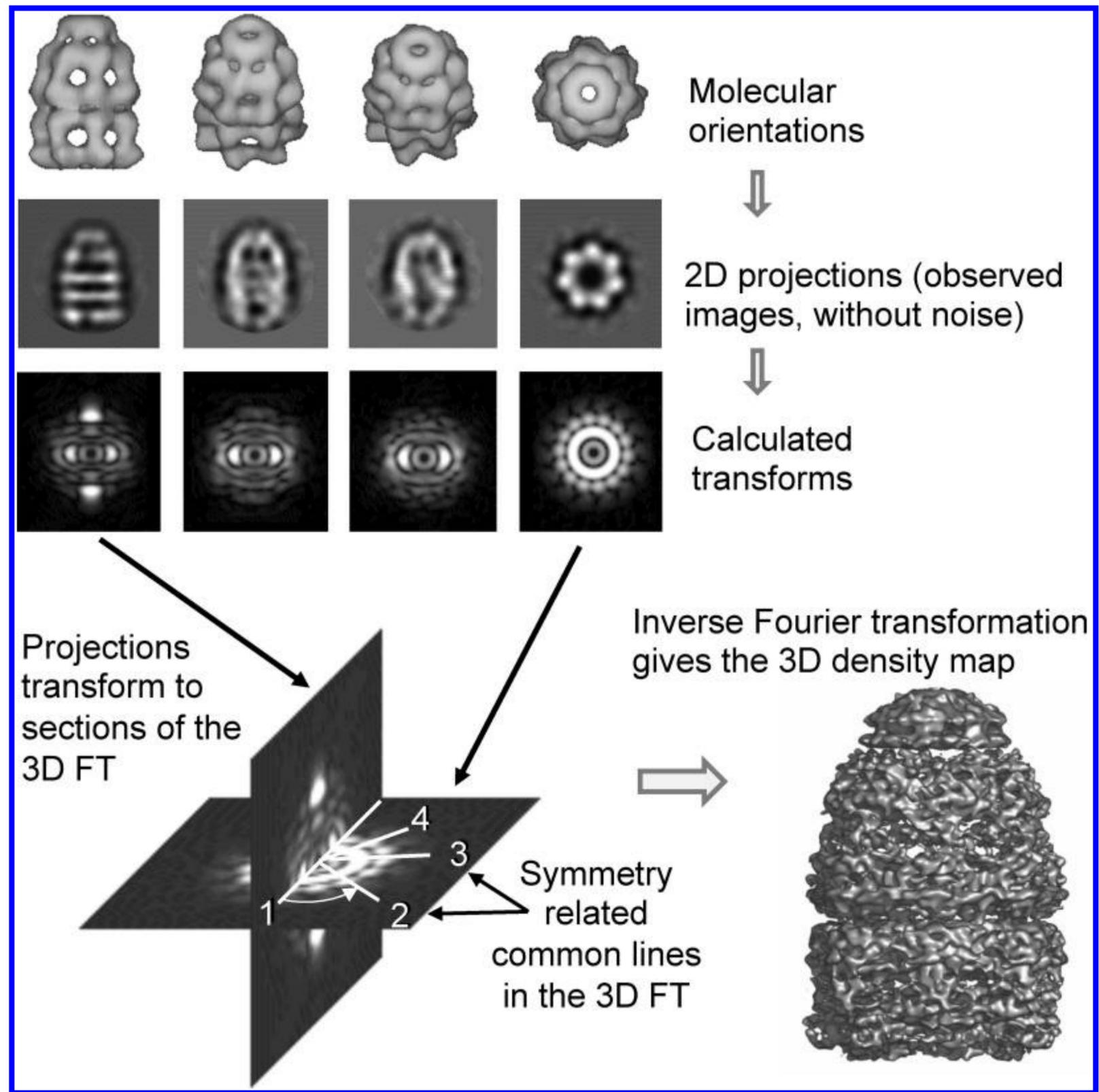CryoSPARC: Stochastic Gradient Descent

a *Ab initio* reconstruction

Input 35,634 particle images

Random initialization

4 min

14 min

75 min

5 min

10 min

24 min

7.5 Å

4.4 Å

3.6 Å

c Gold standard refinement

66 min

3.3 Å

NCCAT

Orlova, E. V. & Saibil, H. R. Structural Analysis of Macromolecular Assemblies by Electron Microscopy. Chem. Rev. 111, 7710–7748 (2011).

32

2026

Molecular orientations

2D projections (observed images, without noise)

Calculated transforms

Projections transform to sections of the 3D FT

Inverse Fourier transformation gives the 3D density map

Symmetry related common lines in the 3D FT

Orlova, E. V. & Saibil, H. R. Structural Analysis of Macromolecular Assemblies by Electron Microscopy. Chem. Rev. 111, 7710–7748 (2011).

33

3D Auto-Refinement (Homogeneous/Uniform Refinement)

Projection Matching

Angular Reconstitution Refinement

3D Classification

Masked (Focused) 3D Auto-Refinement

Partial Signal Subtraction

Multi-Body Refinement

Non-Uniform Refinement
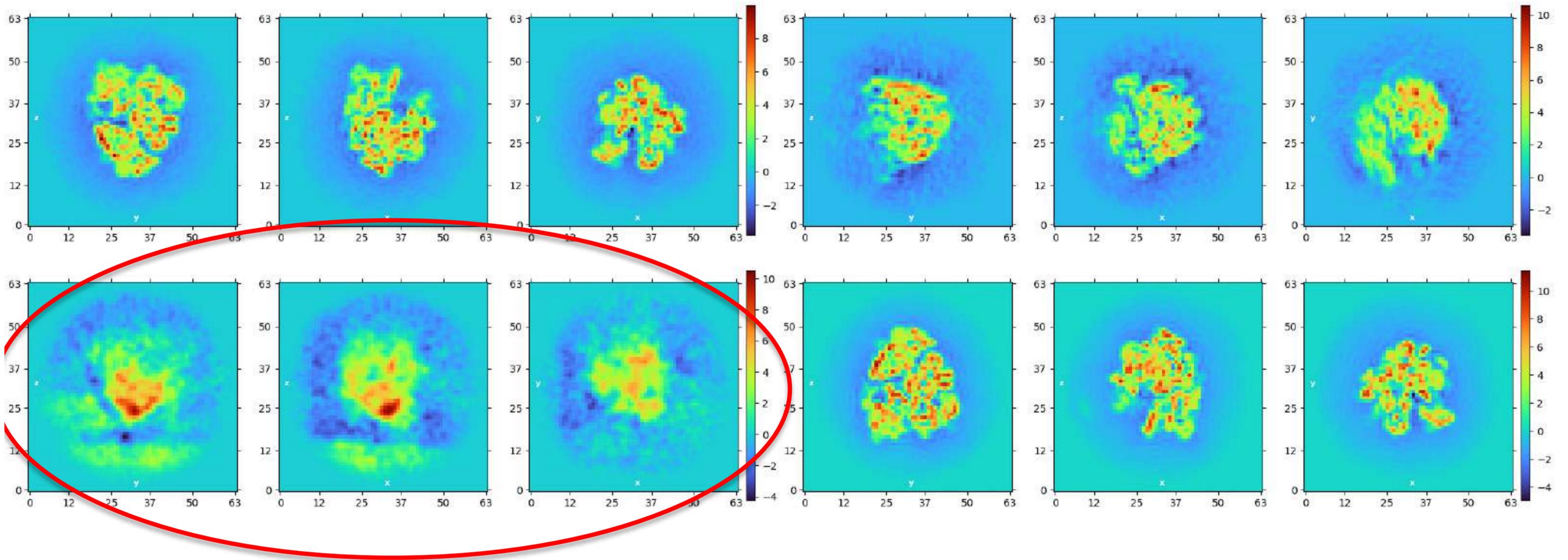
3D Flexible Refinement (3DFlex)

CTF Refinement

Bayesian Polishing (Particle Polishing)
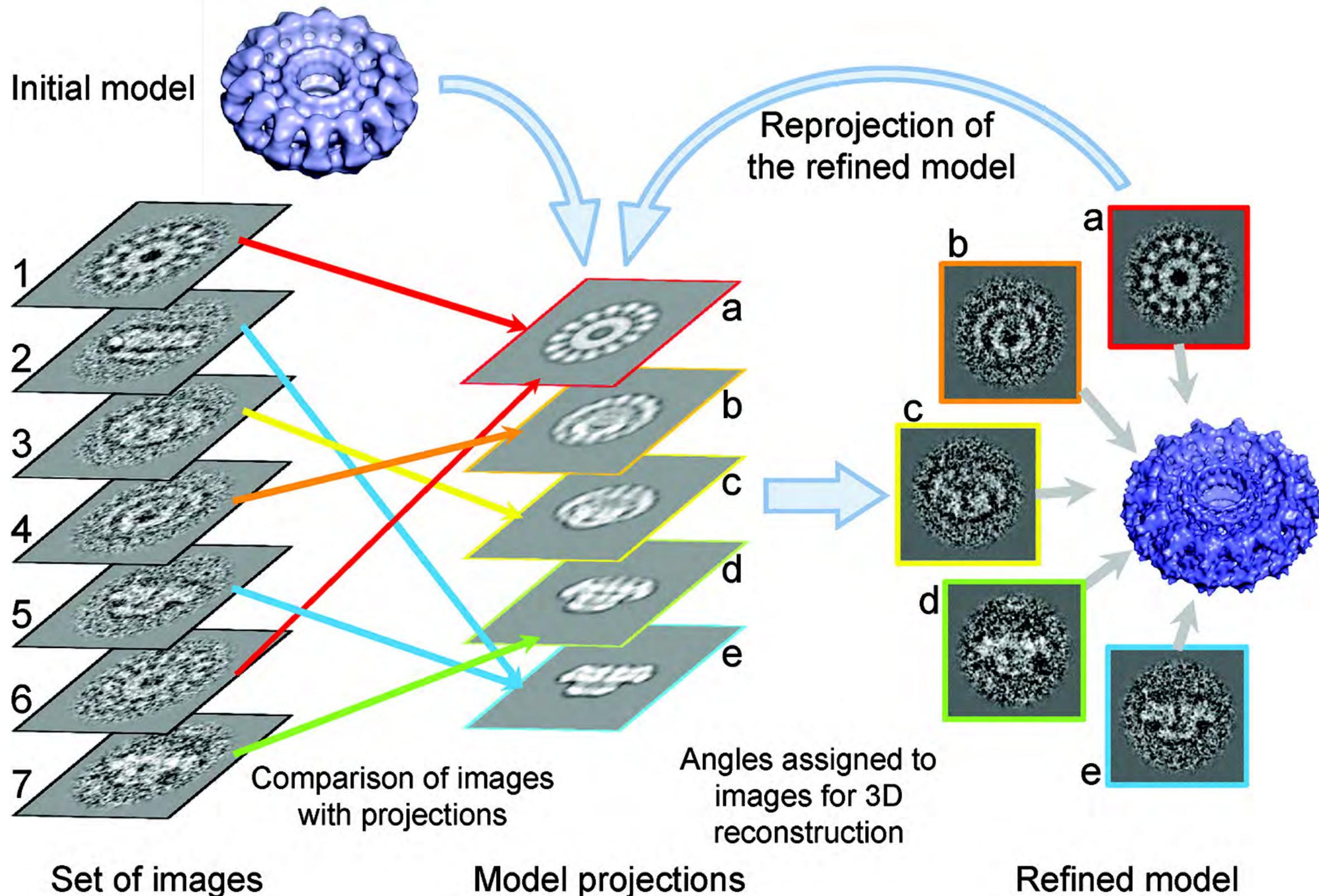
Initial model

Reprojection of the refined model

a
b
c
d
e

Comparison of images with projections

Angles assigned to images for 3D reconstruction

Set of images

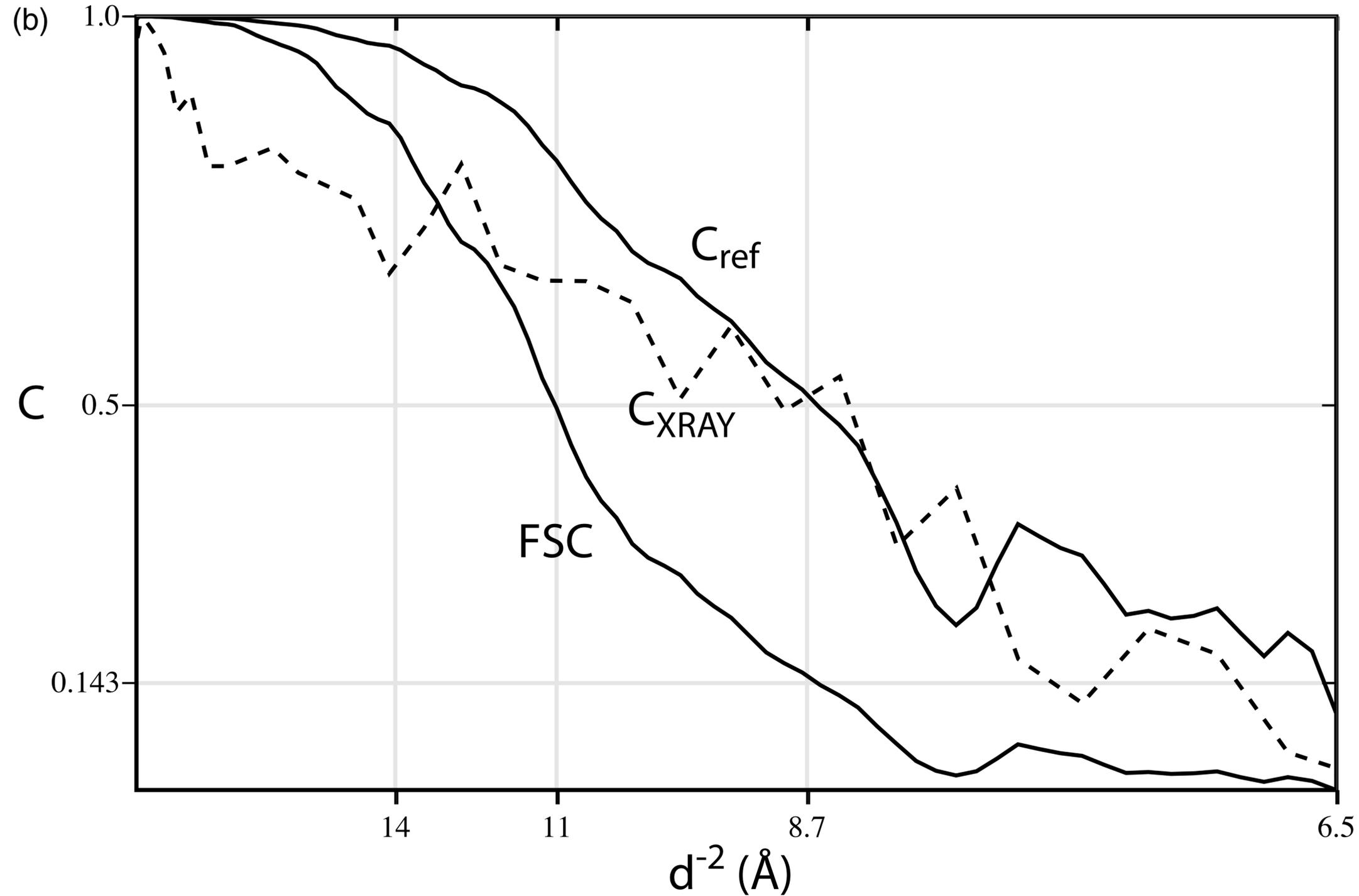Model projections

Refined model

35

- Can be used to clean data further
  - Discard "bad" particles

36

2026

- Can be used to clean data further
  - Discard "bad" particles
 - Discard some preferred orientations

37

(b)
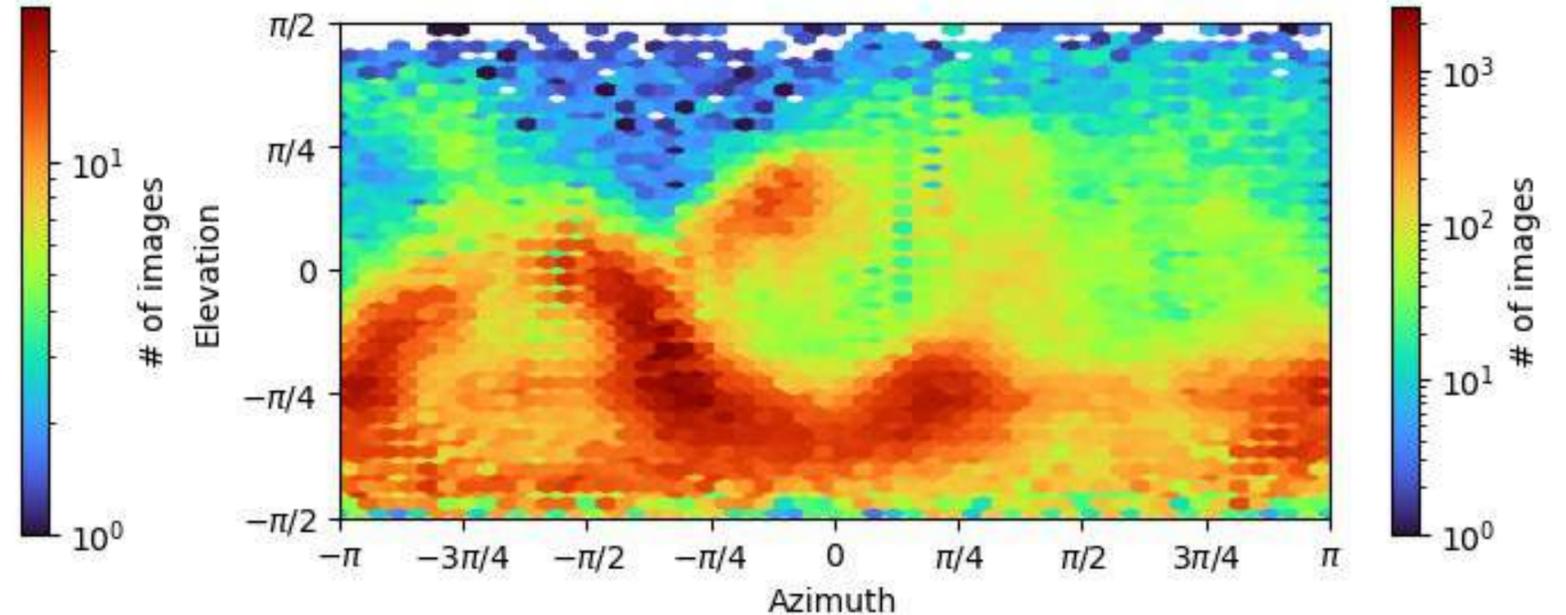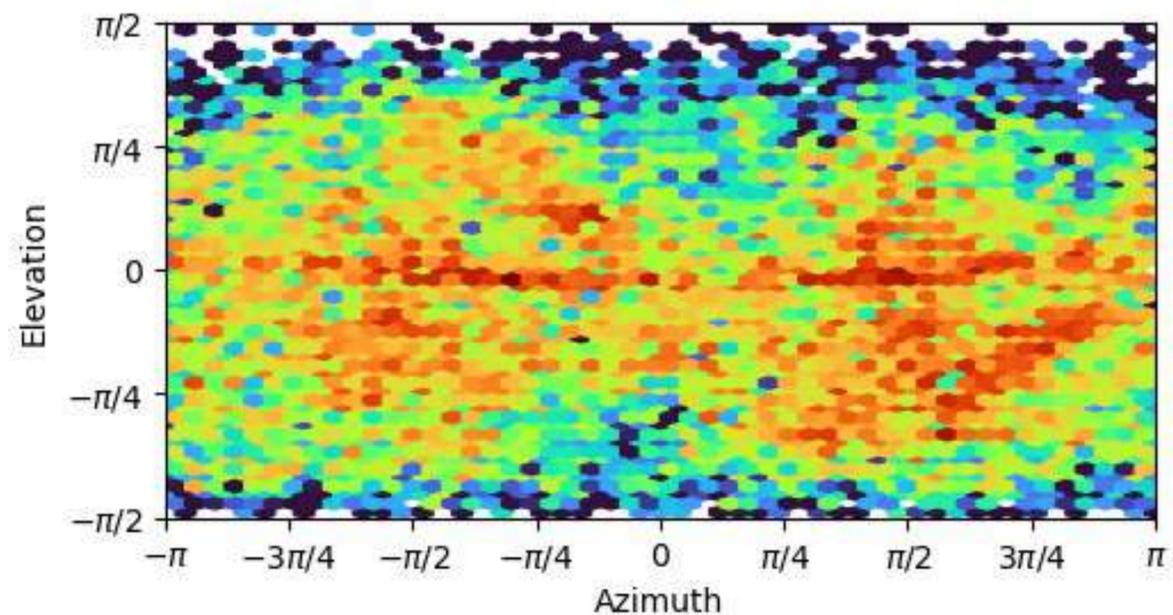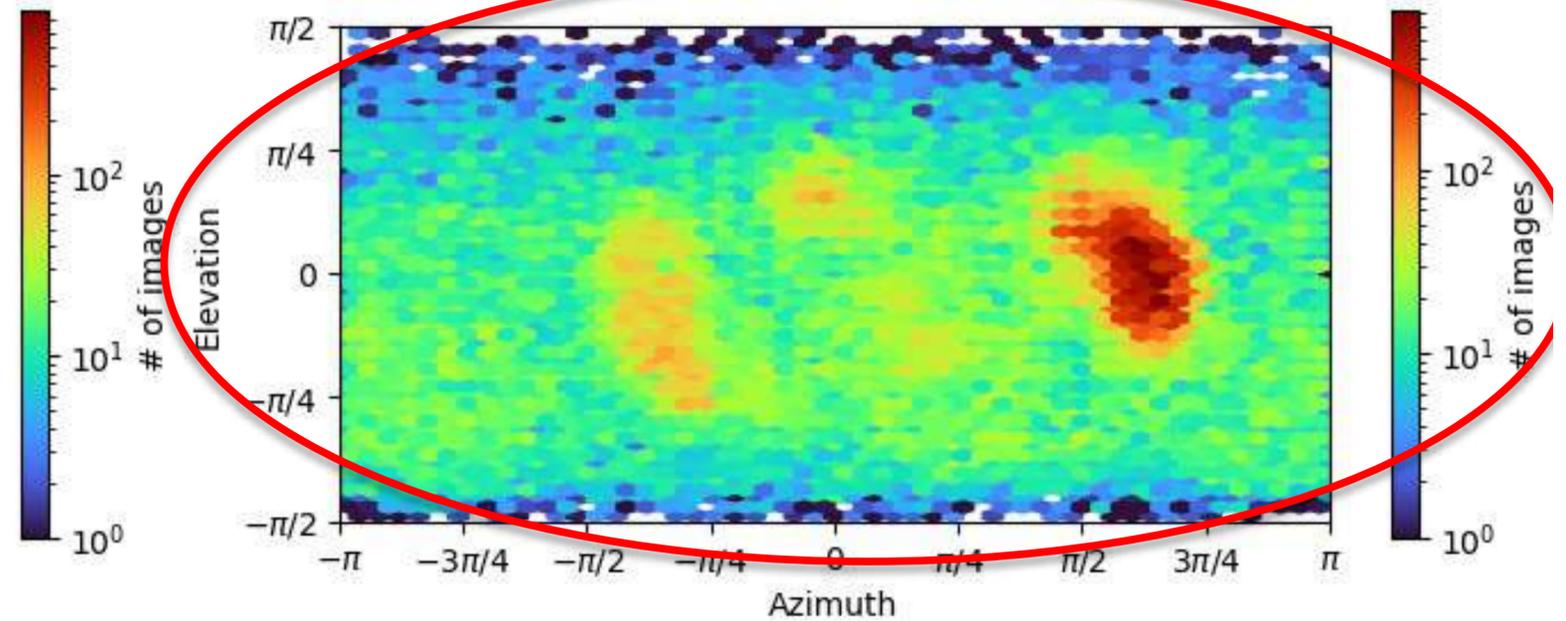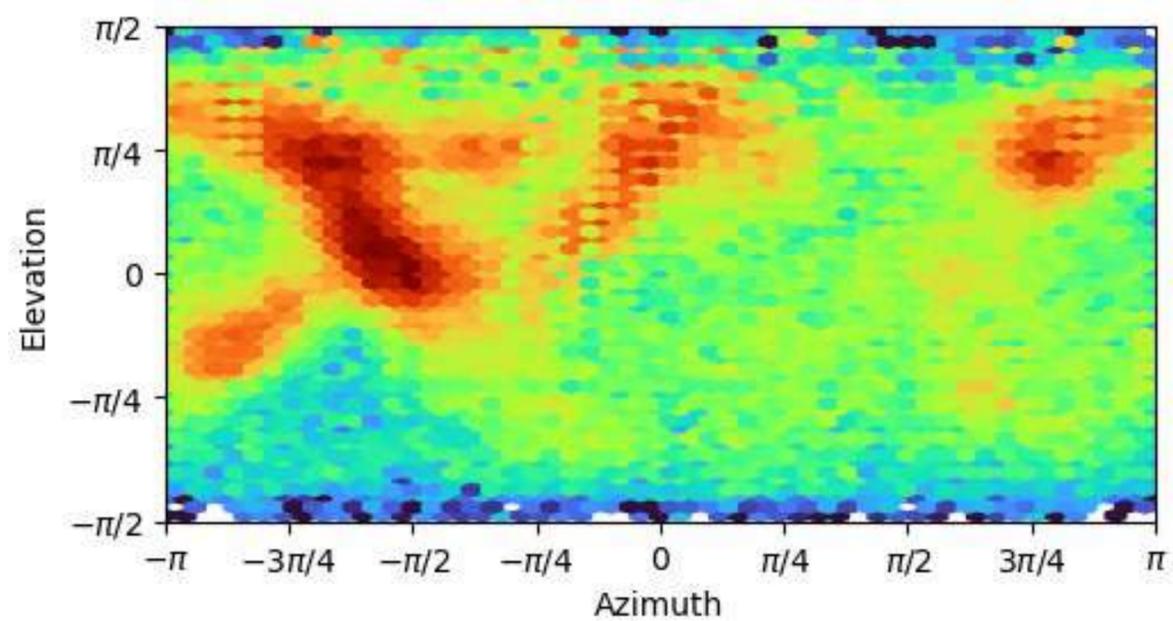
"The FSC is the correlation between two independent maps, where each map is calculated from half the images. In the Appendix, we argue that the resolution of the map should be assigned at the point where the FSC crosses a threshold of 0.143. This corresponds to the resolution at which the estimated correlation between a density map calculated from all the data and a perfect reference map ...... also plotted is 0.5."

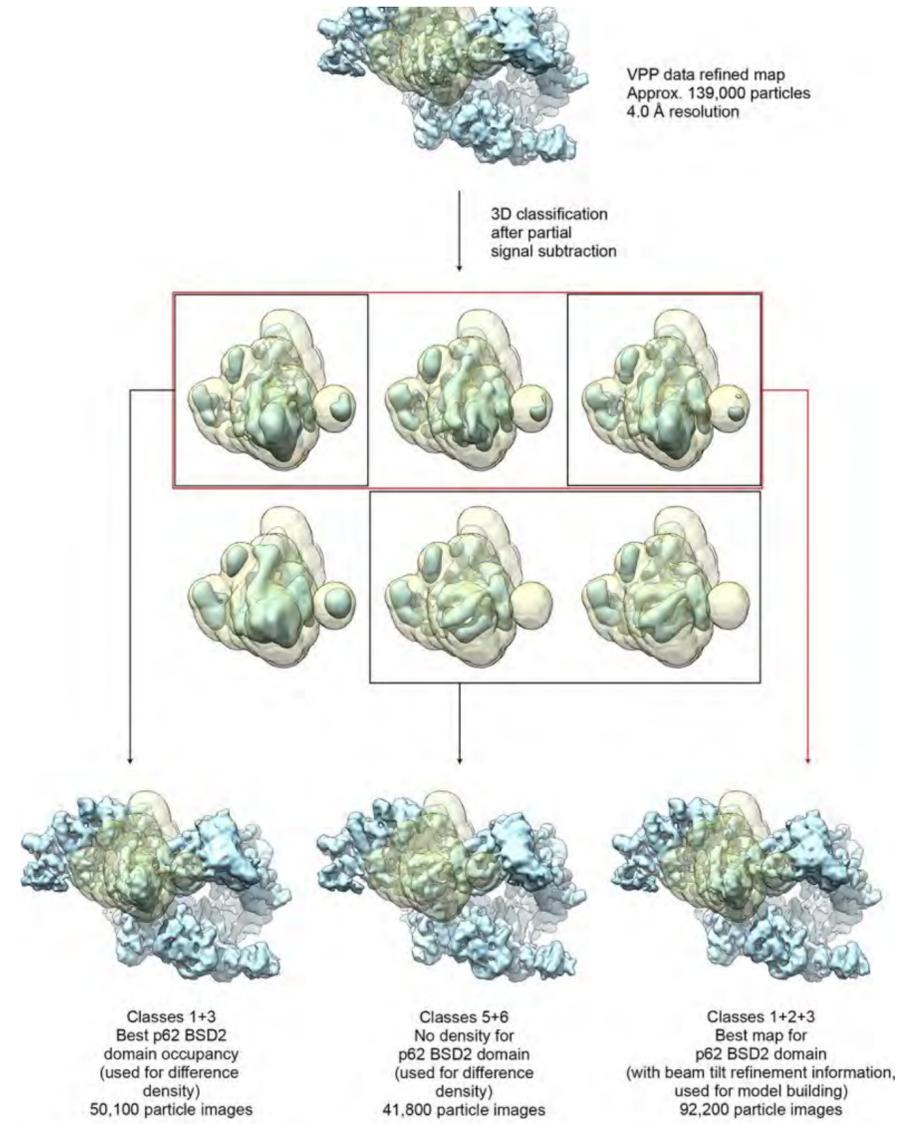Rosenthal, P. B. & Henderson, R. J. Mol. Biol. **333**, 2003.
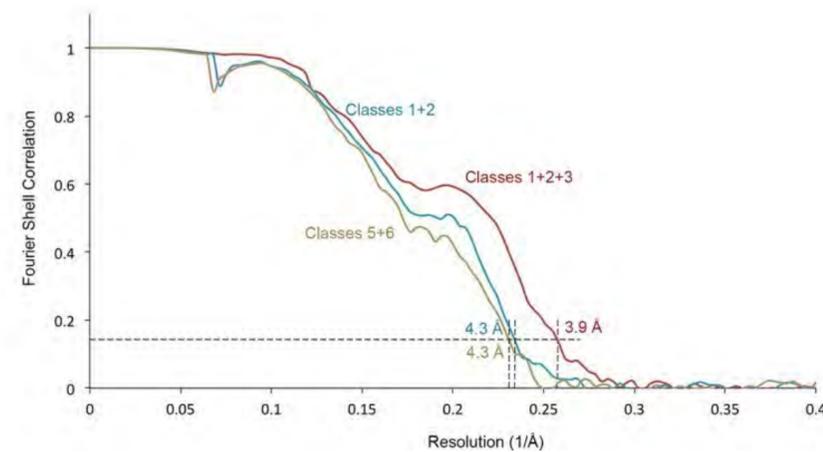
Heel, M. van & Stöffler–Meilicke, M. EMBO J. 1985
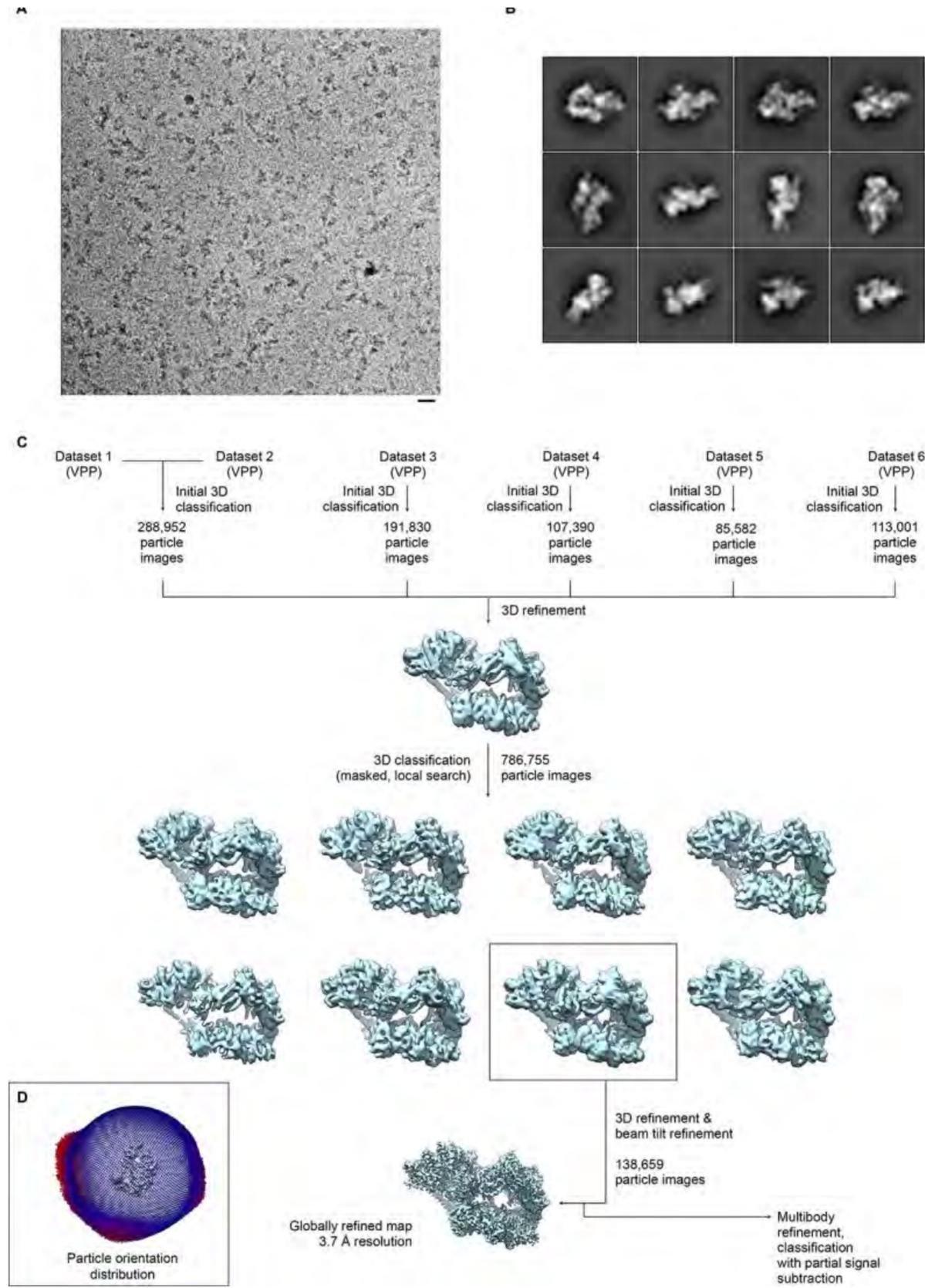
Greber, B. J., Toso, D. B., Fang, J. & Nogales, E. The complete structure of the human TFIIH core complex. eLife **8**, e44771 (2019).

39

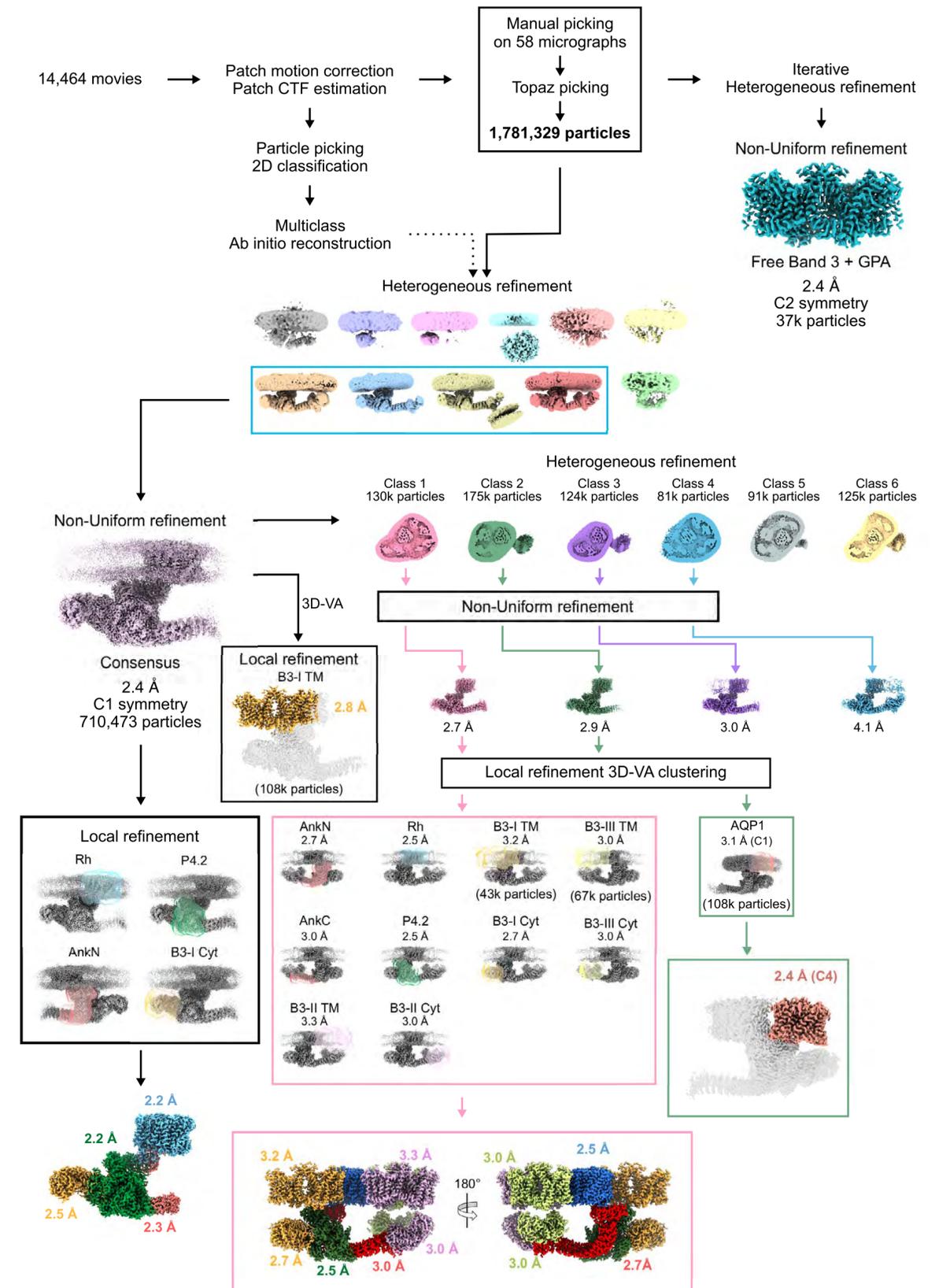Greber, B. J., Toso, D. B., Fang, J. & Nogales, E. The complete structure of the human TFIIH core complex. eLife **8**, e44771 (2019).
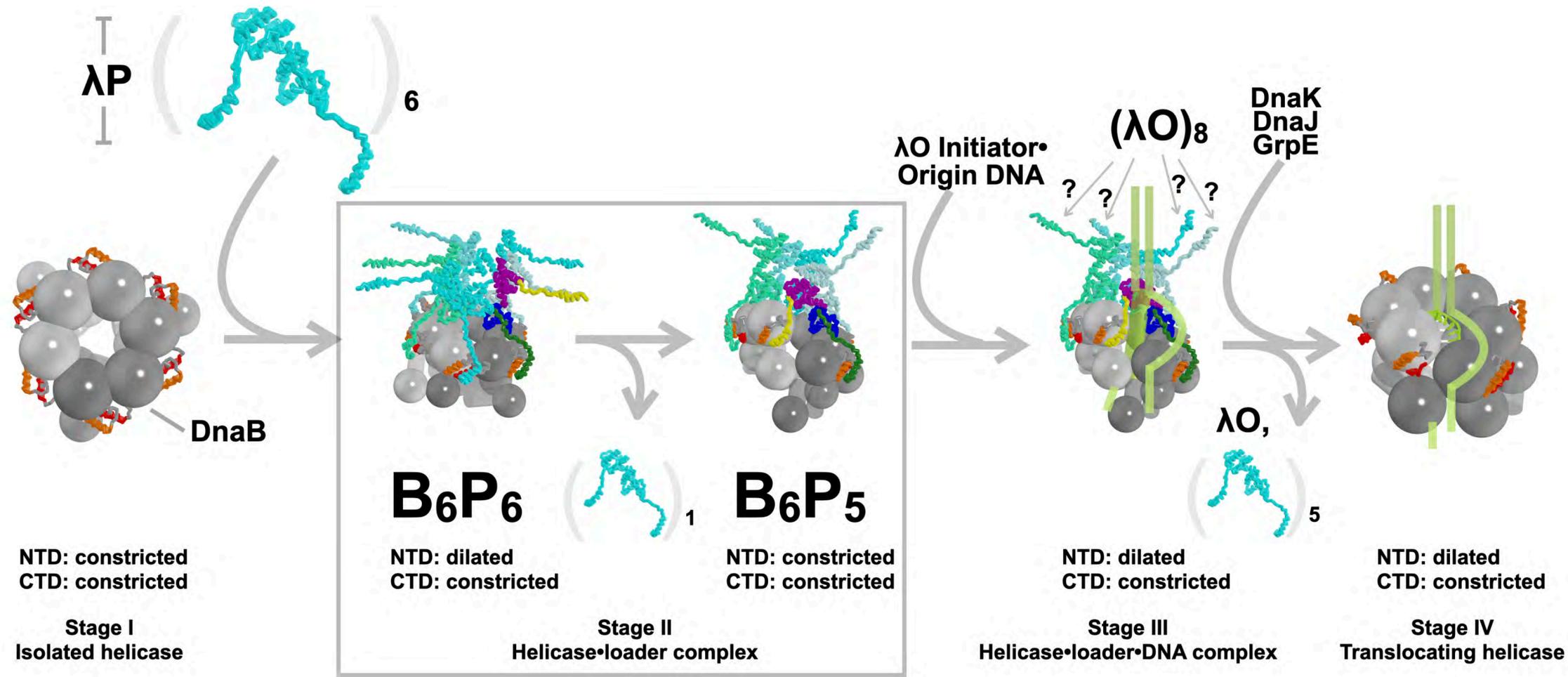
40

- Human erythrocyte ankyrin-1 complex

- Multiple ab initio models generated

- 3D classification to identify compositional differences

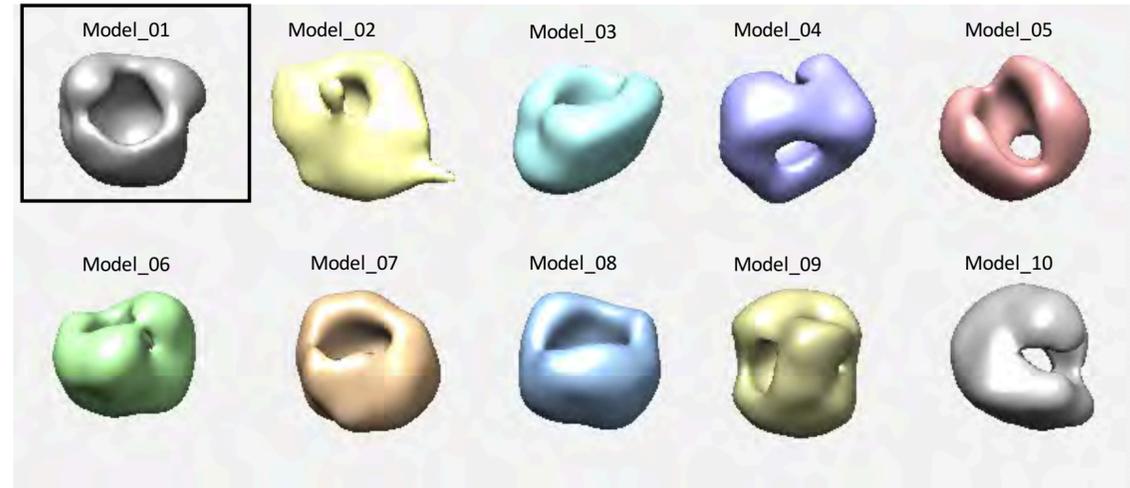- Signal subtraction and local refinement used to identify conformational differences

Vallese, F. et al. Architecture of the human erythrocyte ankyrin-1 complex. Nat. Struct. Mol. Biol. **29**, 706–718 (2022).
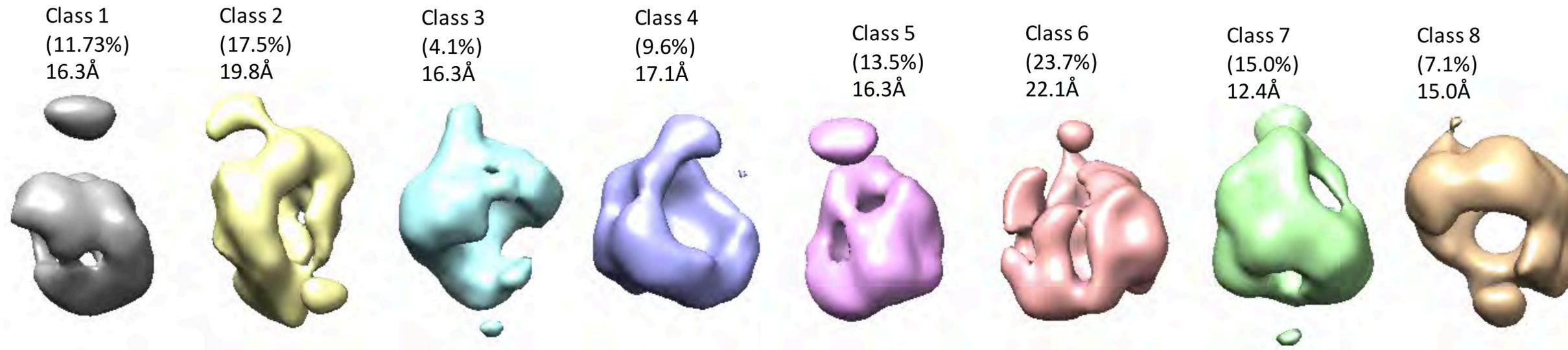
41

2026

1) ab initio

2) DnaB helicase (closed planar ring)

3) DnaB helicase • DnaC helicase loader (open spiral)

4) DnaB helicase • ssDNA (closed spiral)

5) Cryo-electron tomography (BP)

Chase, J. et al. eLife 2018.
Noble, A. J. et al. eLife 2018.
Chase, J., Berger, J. & Jeruzalmi, D. Trends Biochem. Sci. 2022.
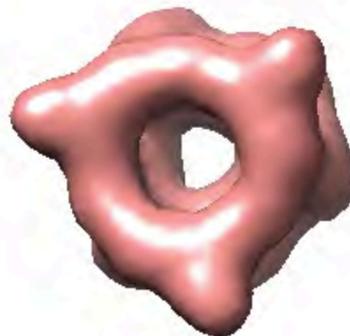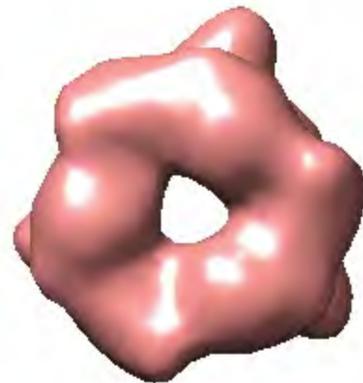Shatarupa, A. , Brown D. et al. Nucleic Acids Res. 2025.

Model_01 Model_02 Model_03 Model_04 Model_05
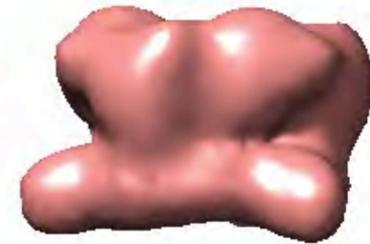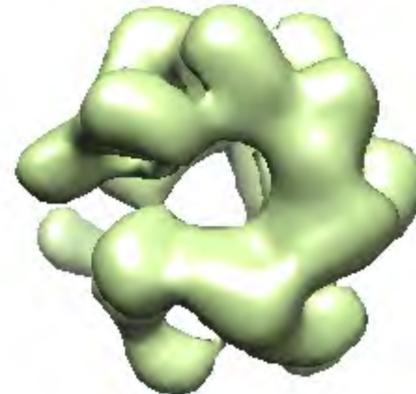Model_06 Model_07 Model_08 Model_09 Model_10

Initial model
(ab initio)
(EMAN)

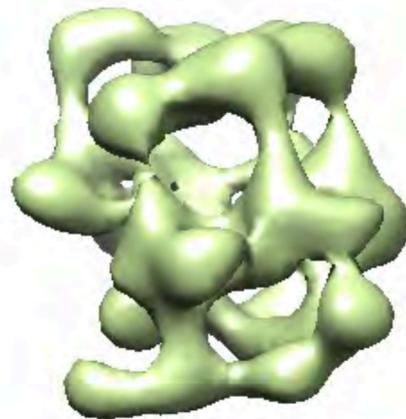3D Classification
Using Model 1
15,529 Particles

Class 1
(11.73%)
16.3Å

Class 2
(17.5%)
19.8Å

Class 3
(4.1%)
16.3Å

Class 4
(9.6%)
17.1Å

Class 5
(13.5%)
16.3Å

Class 6
(23.7%)
22.1Å

Class 7
(15.0%)
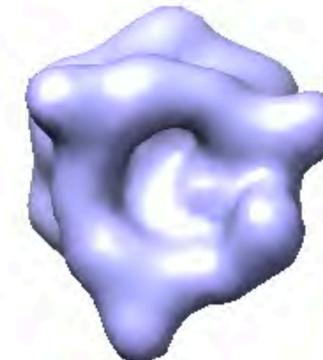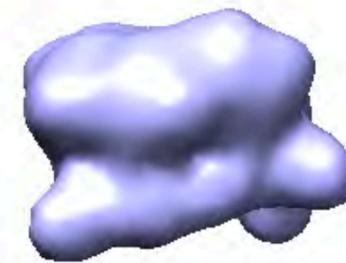12.4Å

Class 8
(7.1%)
15.0Å

43

2026

PDB 2R6A
DnaB
Closed, Planar
30Å

EMDB 2321
DnaB • DnaC
Open Spiral
25Å

PDB 4ESV
DnaB
Closed Spiral
30Å

3D Reconstruction using closed planar form of DnaB (PDB = 2R6A)

	3D_Class1 15.2% 19.8Å	3D_Class2 15.1% 16.3Å	3D_Class4 13.9% 16.3Å	3D_Class6 20.7% 22.1Å	30.0Å
Side View					
"DnaB CTD" Looking dowr					
DnaB NTD looking up					

Bailey, Science 2007

45

2026

	3D_Class3 27.2% 22.0Å	3D_Class5 13.1% 16.0Å	3D_Class7 14.9% 18.8Å	3D_Class8 13.2% 19.8Å	25.0Å
Side View					
"DnaB CTD" Looking down					
DnaB NTD looking up					

Arias-Palomo Cell 2013

46

	3D_Class2 17.1% 26.8Å	3D_Class4 12.4% 22.1Å	3D_Class5 17.3% 26.8Å	3D_Class8 20.0% 25.0Å	25.0Å
Side View					
"DnaB CTD" Looking down					
DnaB NTD looking up					
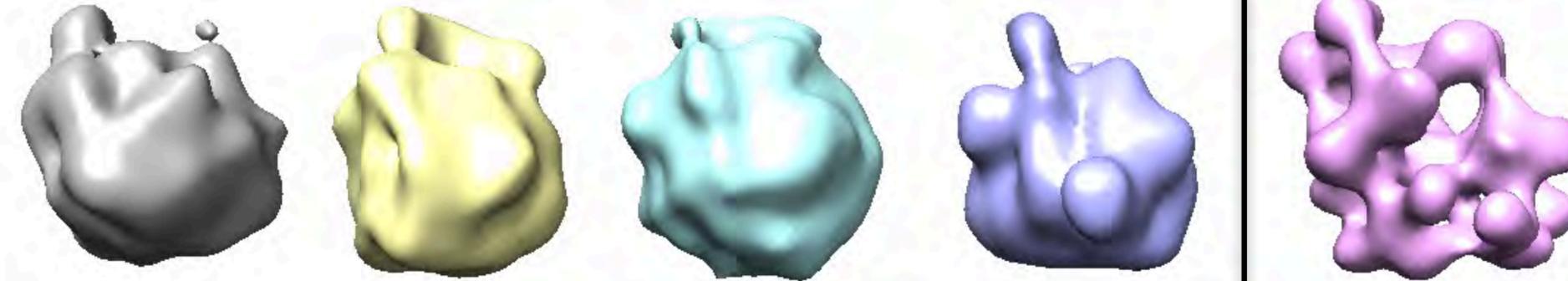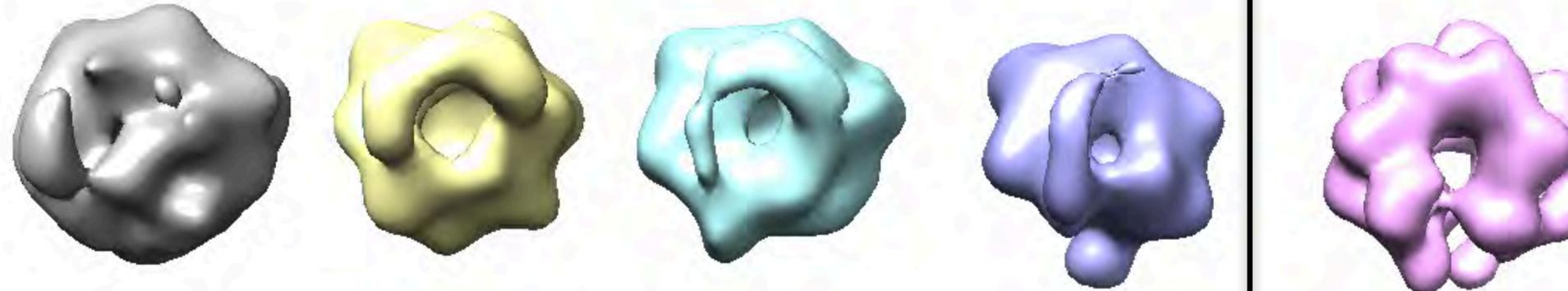
Itsathitphaisarn, Cell 2012
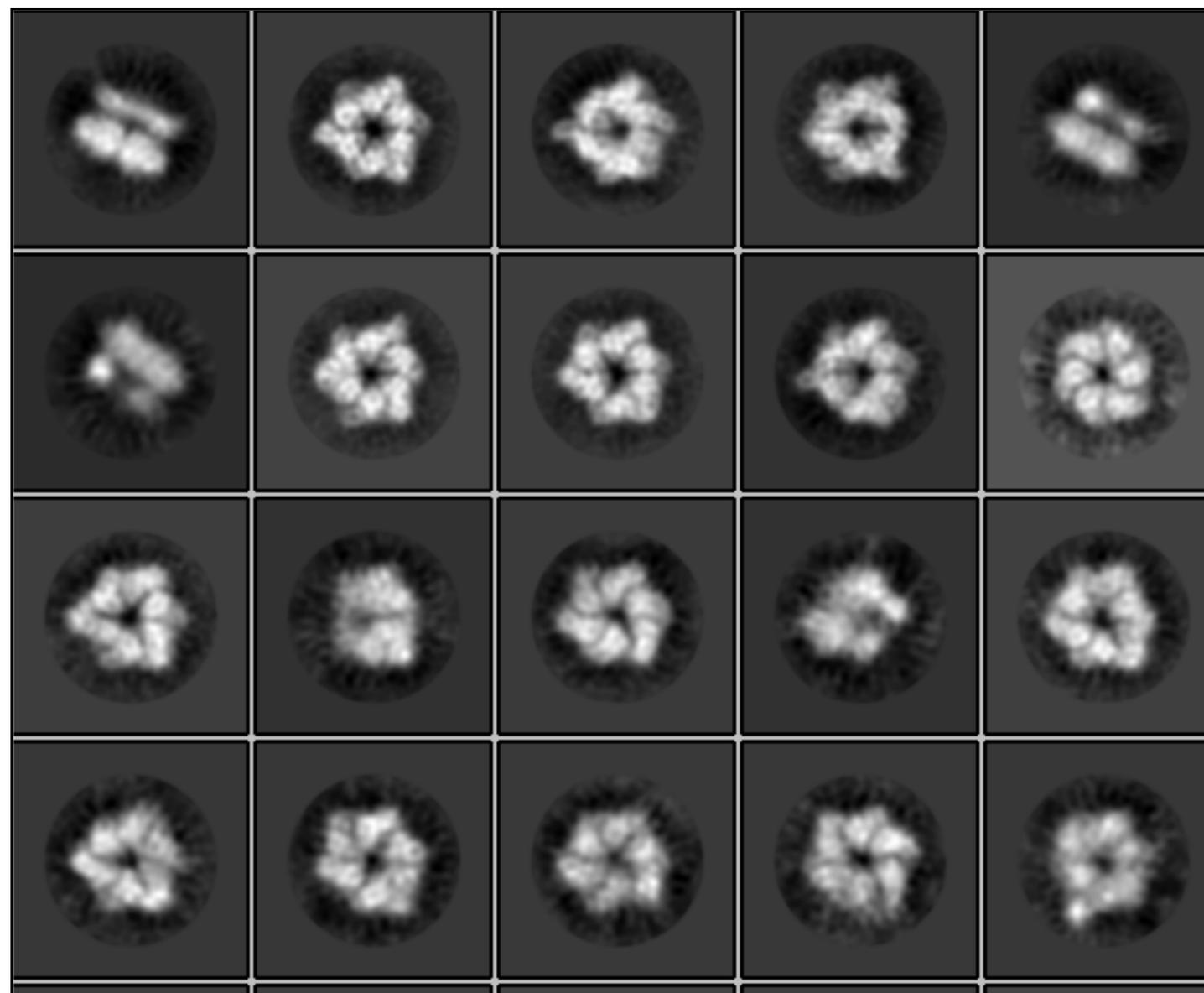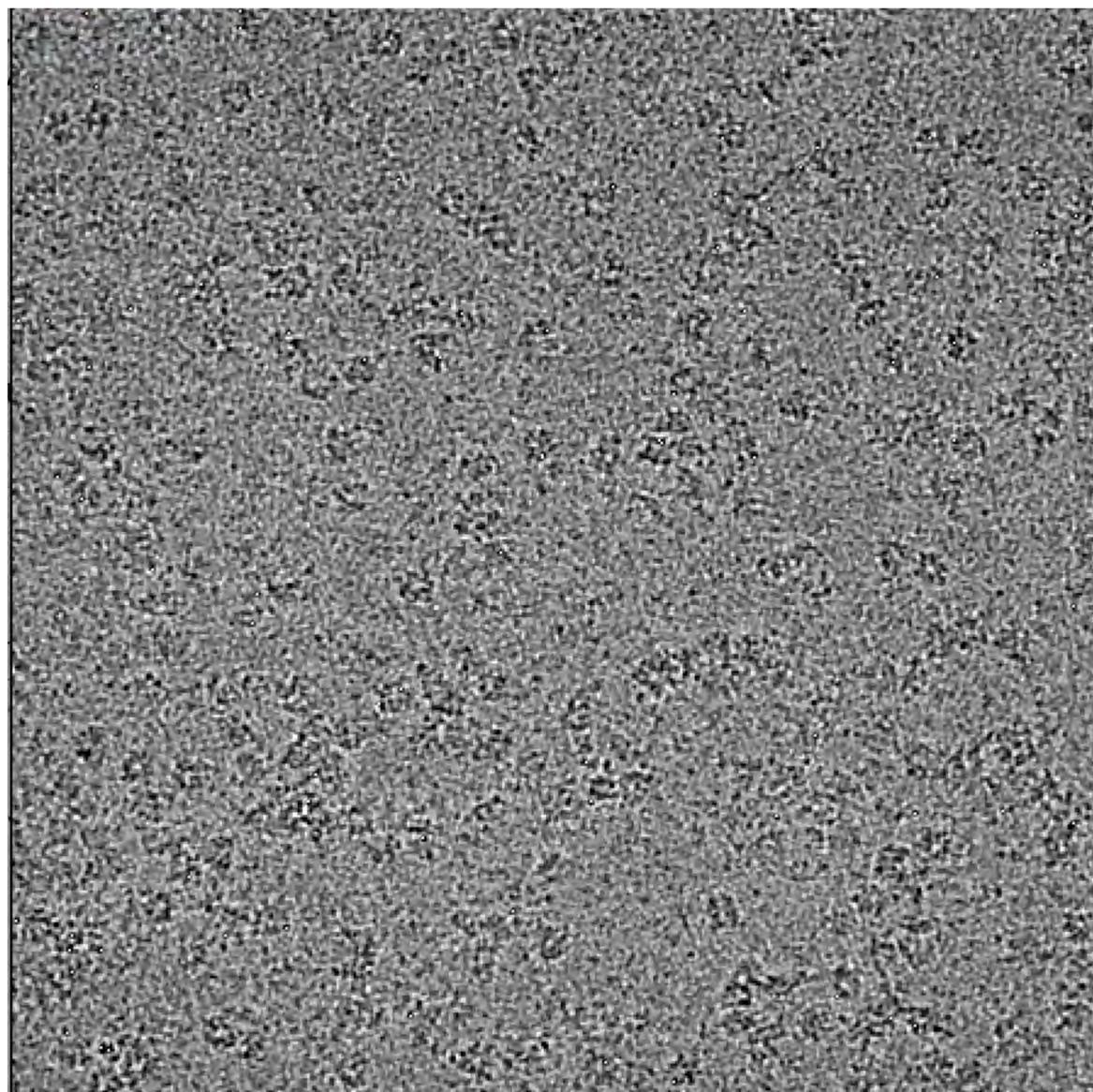
Krios/K2 Detector

**1.07 Å/px, 256x256 px
15,000 Particles
RELION**

48

Gaussian particle picking
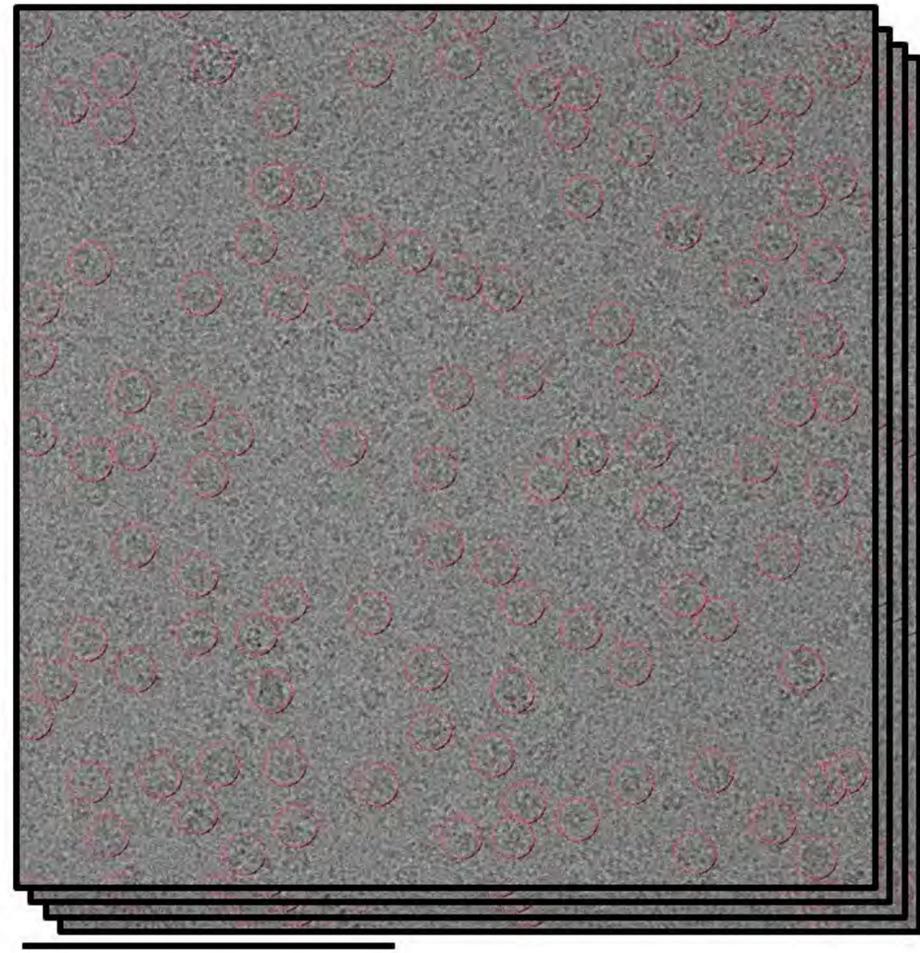
2D classification
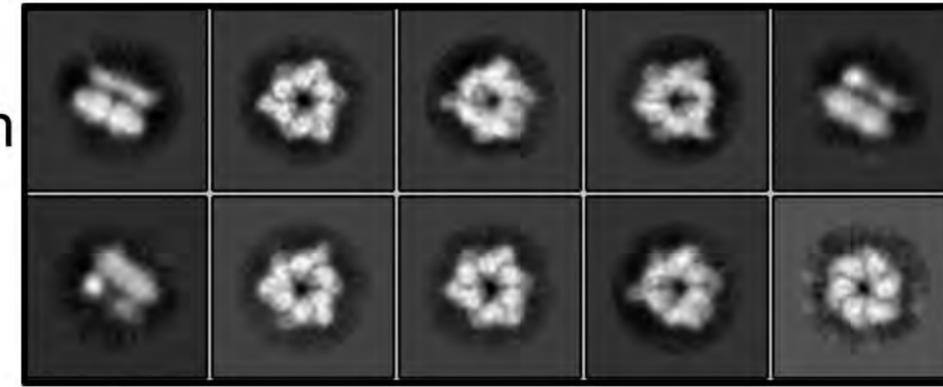
Incoherent ab initio reconstruction

(a) particles

+60°

0° ice slab

−60°

record images

(b) e⁻

tiltseries

(c) 3D tomogram from back-projected tiltseries

(d) 3D tomogram — subtomograms or "particles"

Noble, eLife 2018
Cheng, Chinese Physics 2018

Sample #33
DNAB Helicase-helicase Loader

110 nm ice

-36:21:3°, Gold Quantifoil, Krios, K2

250 nm

Noble, eLife 2018

Five tomograms;
~1,000 particles

SPT

45°

45°

52

DnaB·λP Particle Distribution

Jillian Chase

FSC of DnaB• λP
EM Map : 4.1 Å

Jillian Chase

Particle picking (Blob picker)

7,250,518 particles, 100%

Inspect picks, Extract from micrographs (150 px), 2D, Template picker, Inspect picks, Extract (200 px)

3,709,955 particles, 51.16%

2D classification

2,798,195 particles, 38.59 %

Ab-initio

Class A-00	Class A-01	Class A-02	Class A-03
780,554 P/V	473,341 P/V	996,526 P/V	547,774 P/V

2,798,195 particles, 38.59%
Heterogeneous Refinement

Class B-00, 4.40 Å
480,863 P, 6.63%

Class B-01, 4.40 Å
1,505,831 P, 20.76%

Class B-02, 7.83 Å
283,435 P, 3.90%

Class B-03, 4.40 Å
528,066 P, 7.28%

Non-uniform Refinement

Non-uniform Refinement

Class C, 2.75 Å

4.44 Å (Optimize per-particle defocus enabled)

Non-uniform Refinement
+ B-01 Volume

4.44 Å (Optimize per-group CTF params enabled)

Extraction with 400 px

1,500,880 particles, 20.70%

+ B-01 Volume

Non-uniform Refinement

X- indicates discarded particles or volumes not further pursued in this study

DeepEMhancer

Class E, 2.66 Å

Class D, 2.66 Å (Optimize per-particle defocus enabled)

Shatarupa, A. , Brown D. et al. Nucleic Acids Res. 2025.

2026

A

B GSFSC Resolution: 2.84Å

- No Mask (3.3Å)
- Spherical (3.1Å)
- Loose (3Å)
- Tight (2.8Å)
- Corrected (2.8Å)

C

0°   72°   144°   216°   288°
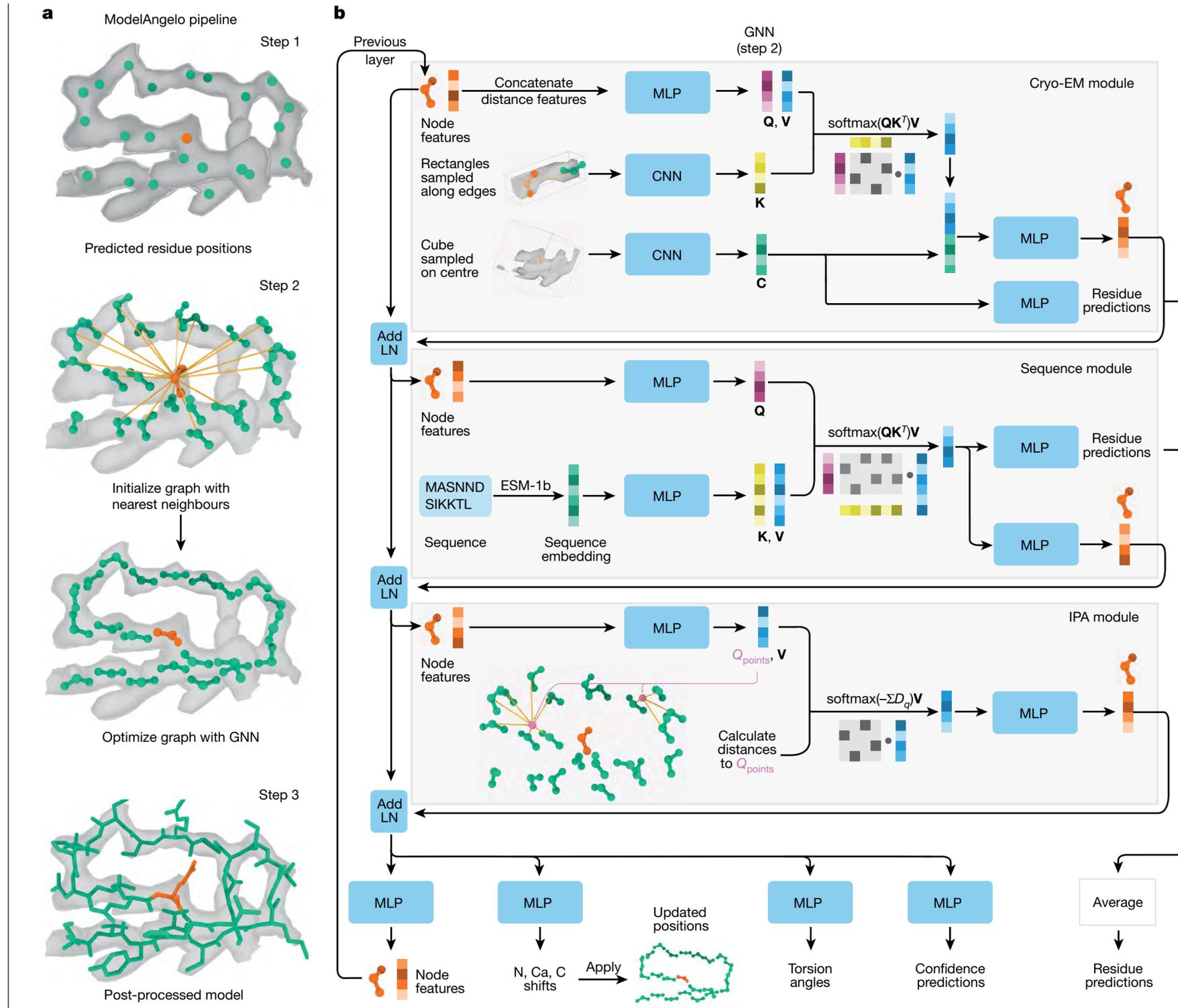
4.8   3.8   2.8
Resolution (Å)
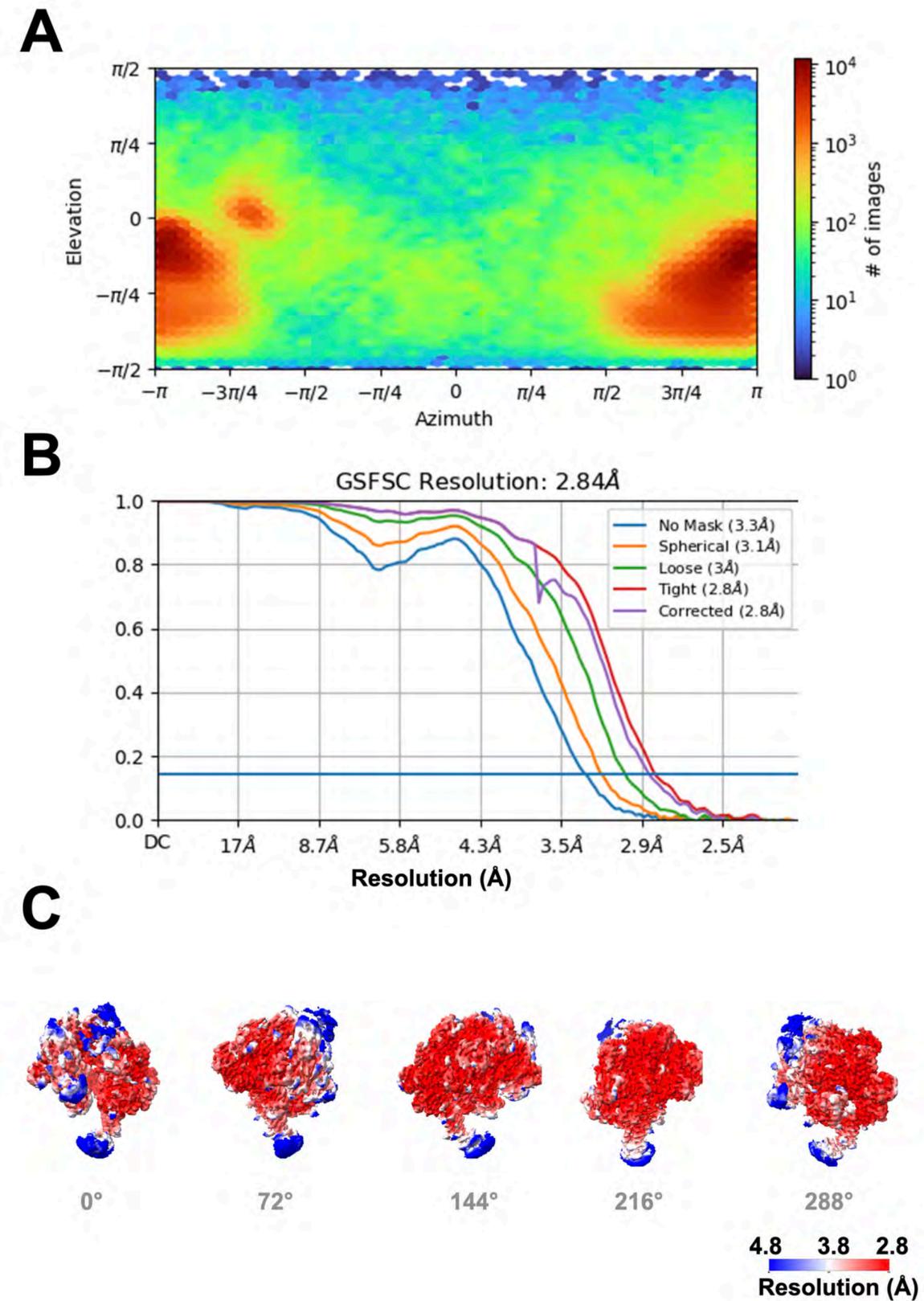
communications
biology

DeepEMhancer: a deep learning solution for cryo-EM volume post-processing

Ruben Sanchez-Garcia [1,4], Josue Gomez-Blanco[2,3], Ana Cuervo[1], Jose Maria Carazo [1], Carlos Oscar S. Sorzano [1✉] & Javier Vargas[2,3✉]
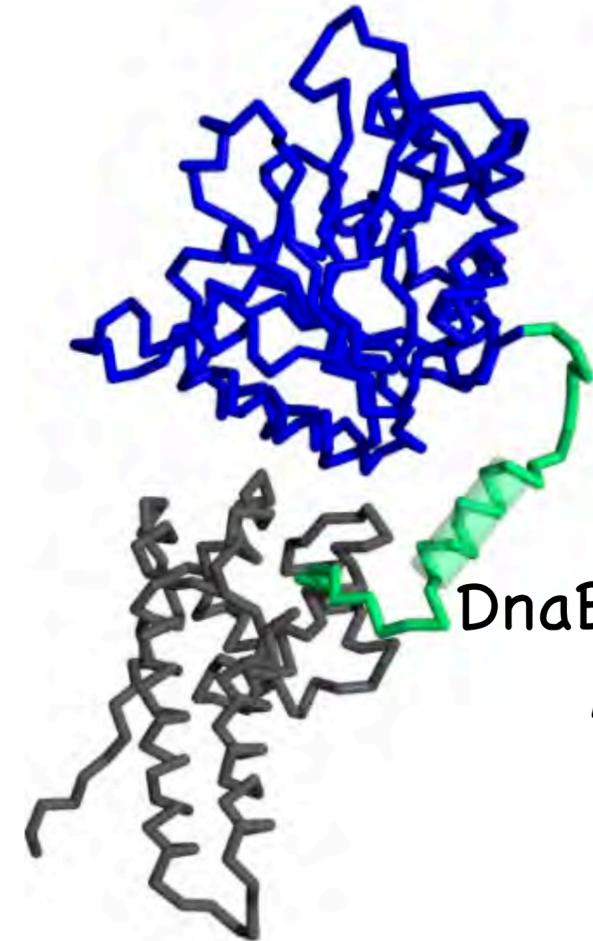
60

Lamba P:

1 - 119: Alphafold

120-192: 1.86 Å crystal structure

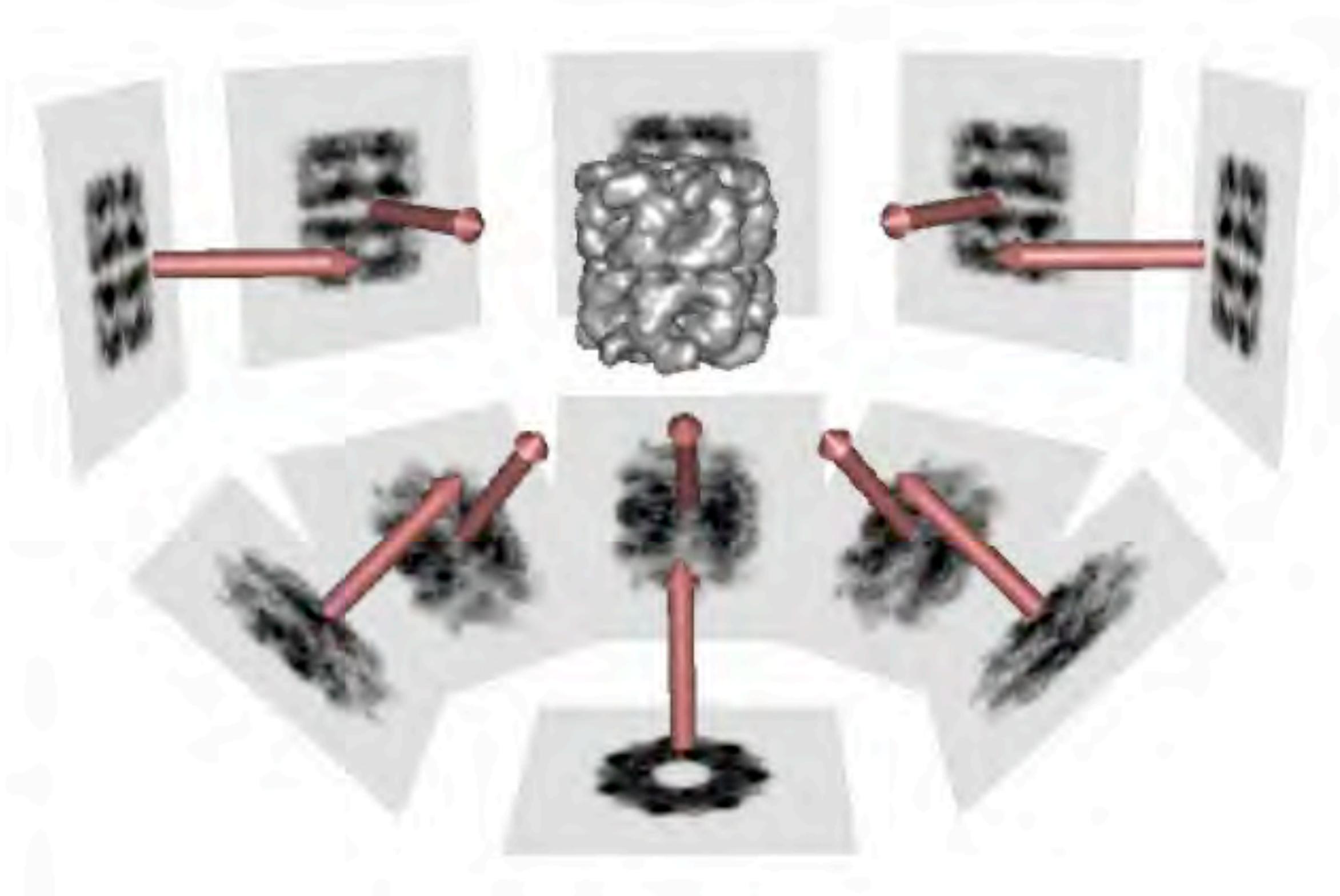193-299: Alphafold

DnaB C-Term:
Alphafold

DnaB linker-helix:
Alphafold

DnaB N-Term:
1B79 (2.30 Å)

Orlova, E. V., & Saibil, H. R. (2011). Structural analysis of macromolecular assemblies by electron microscopy. Chemical Reviews, 111(12), 7710–7748. http://doi.org/10.1021/cr100353t

63

2026

Our study addressed the so-called 'initial volume problem' in the single-particle reconstruction,

2026

Molecular Electron Microscopy

65

66

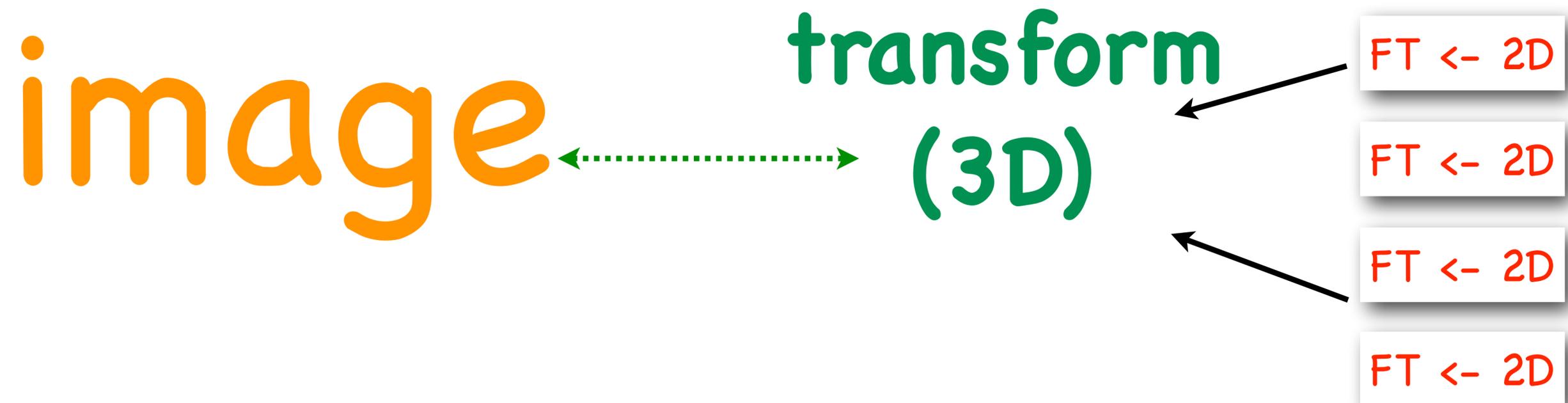FT (2D)   FT (2D)   FT (2D)   FT (2D)   FT (2D)

FT (3D)
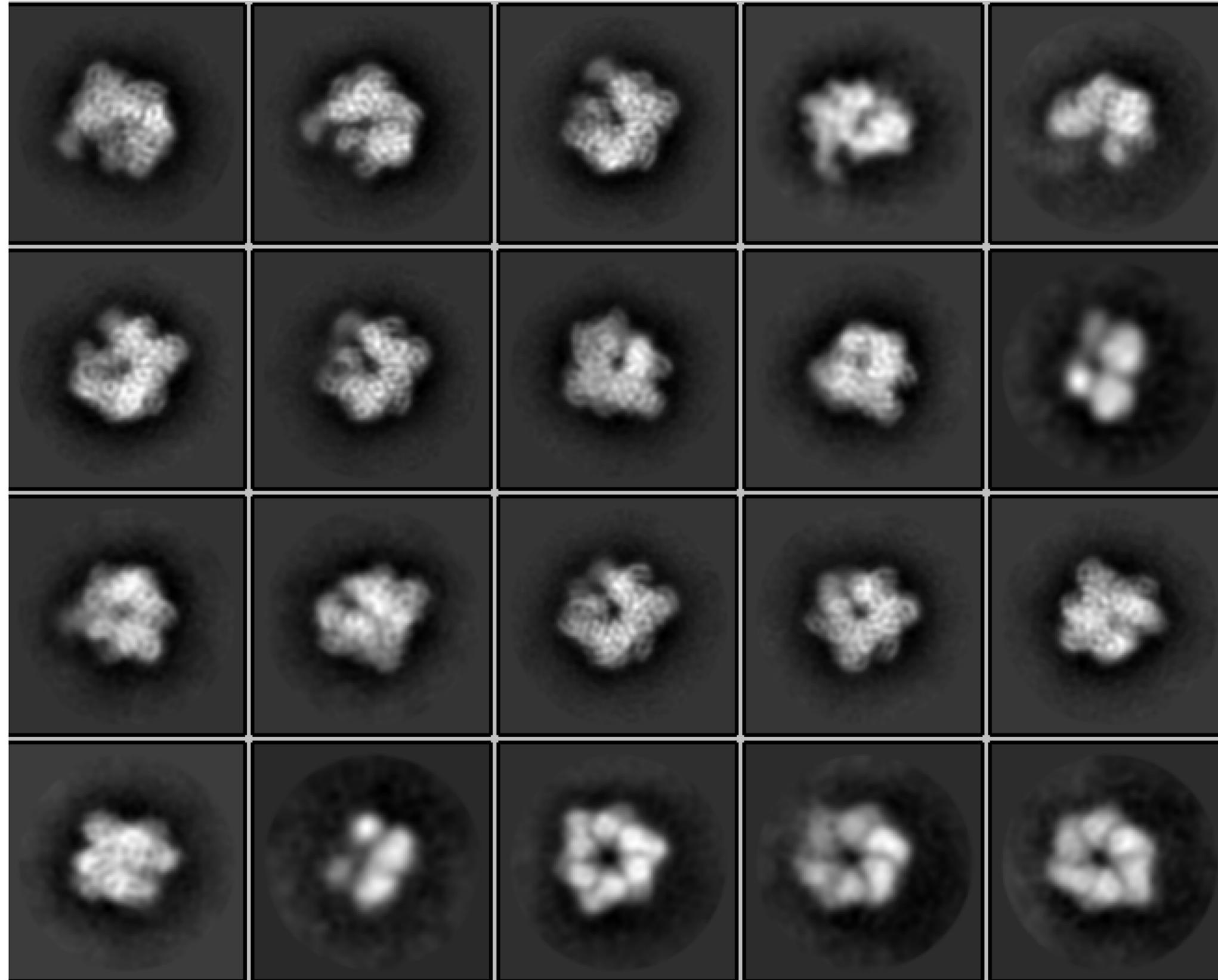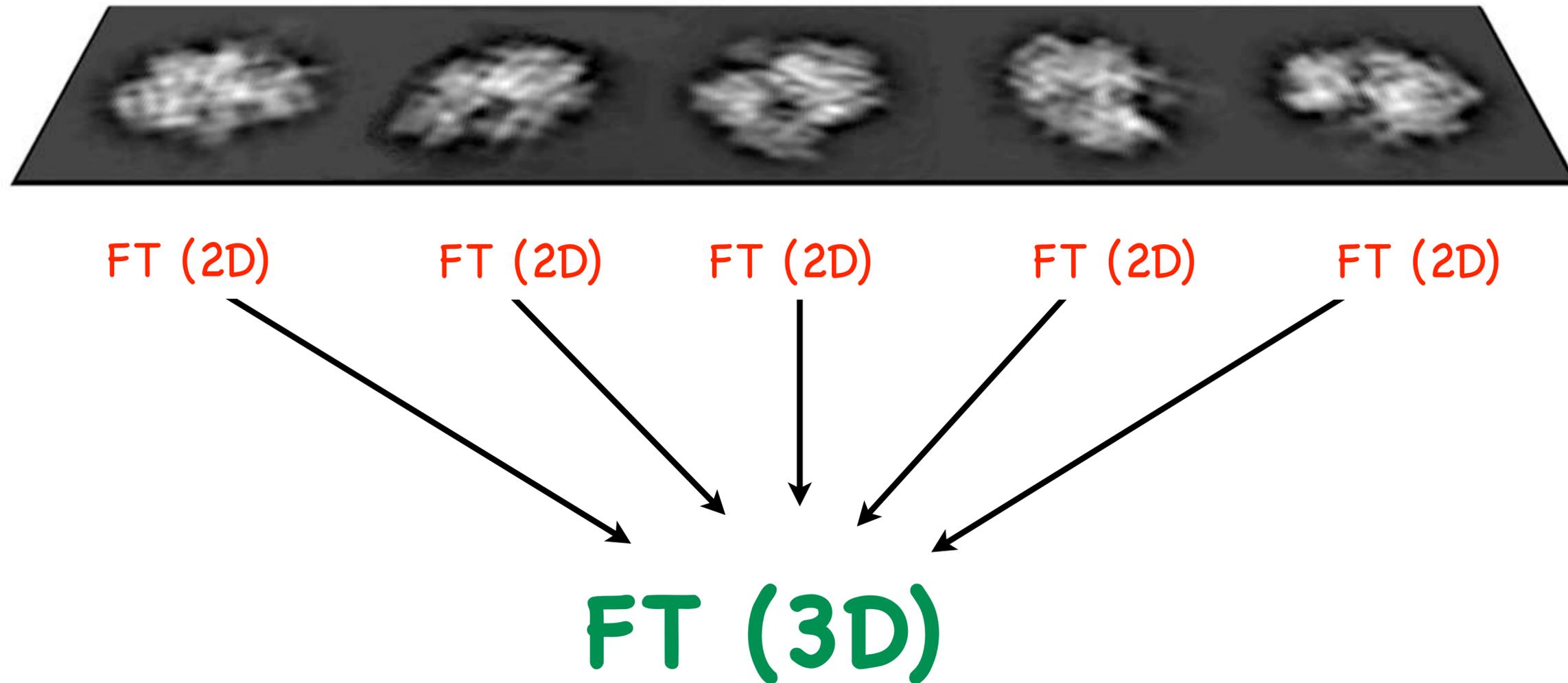
The Fourier Transform of a projection (2–D) of an object corresponds to a <u>section</u> through the center of 3D–Fourier transform of object. Johann Radon, 1917
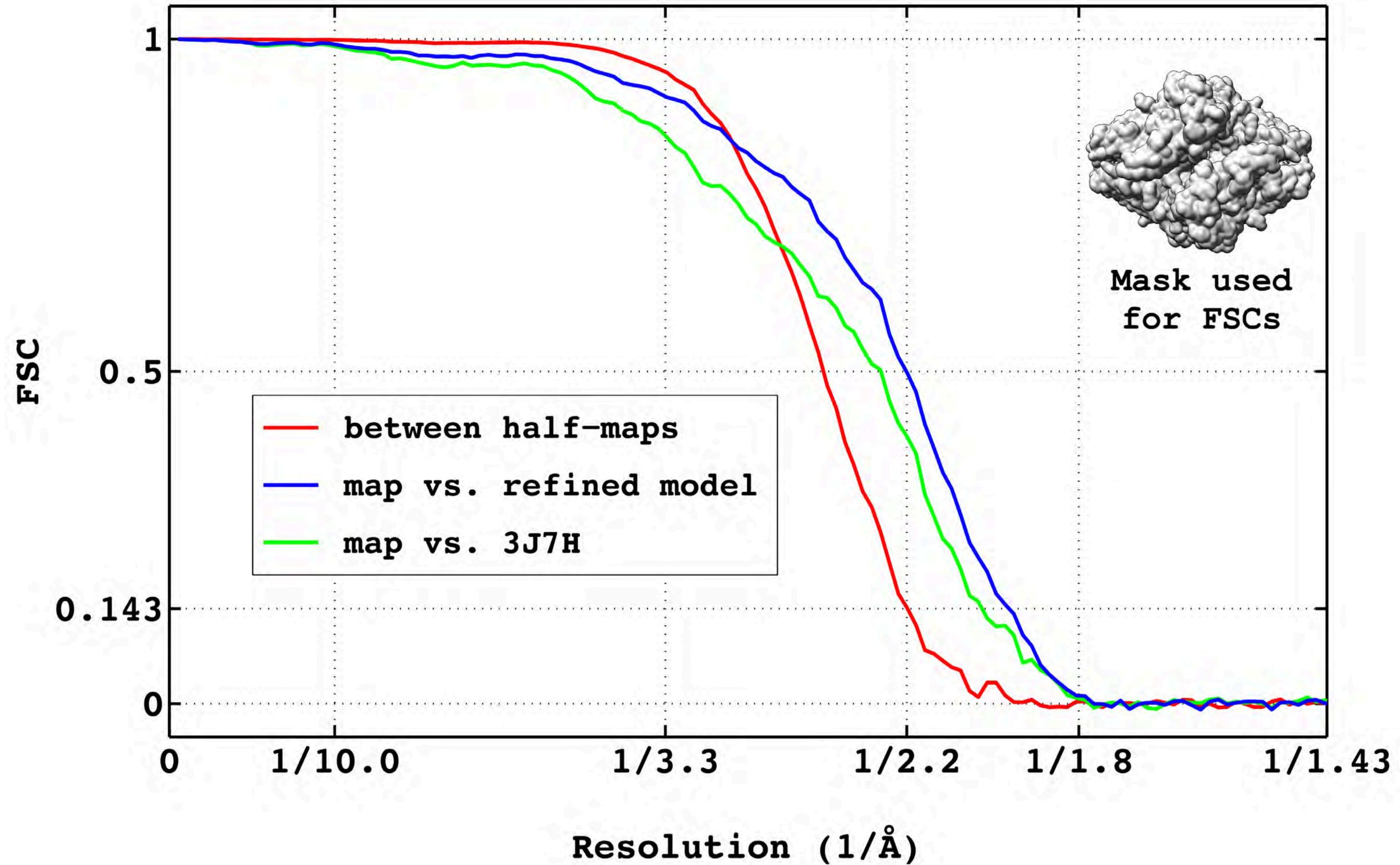Klug, Desrosier, Crowther, 1968, 1970

Fourier Transform

Real space ←----------------→ Reciprocal space

image ←----------------→ **transform (3D)**

FT <- 2D

FT <- 2D

FT <- 2D

FT <- 2D

Mask used for FSCs

Legend:
- **between half-maps** (red)
- **map vs. refined model** (blue)
- **map vs. 3J7H** (green)

Y-axis: FSC (1, 0.5, 0.143, 0)

X-axis: Resolution (1/Å) — 0, 1/10.0, 1/3.3, 1/2.2, 1/1.8, 1/1.43

# of particles
1:200
2:1000
3:3000
4:6384
5:6384
+ symmetry applied

Böttcher–Nature, 1997

70

GSFSC Resolution: 4.01$\mathring{A}$

No Mask (6.7$\mathring{A}$)
Spherical (4.7$\mathring{A}$)
Loose (4.4$\mathring{A}$)
Tight (4$\mathring{A}$)
Corrected (4$\mathring{A}$)

71

2026