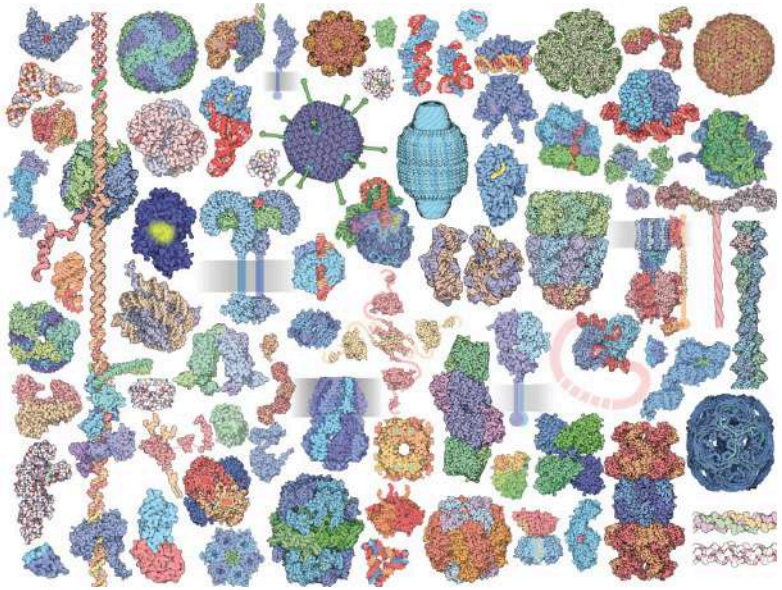
Single-particle Cryogenic Electron Microscopy ("cryo-EM")

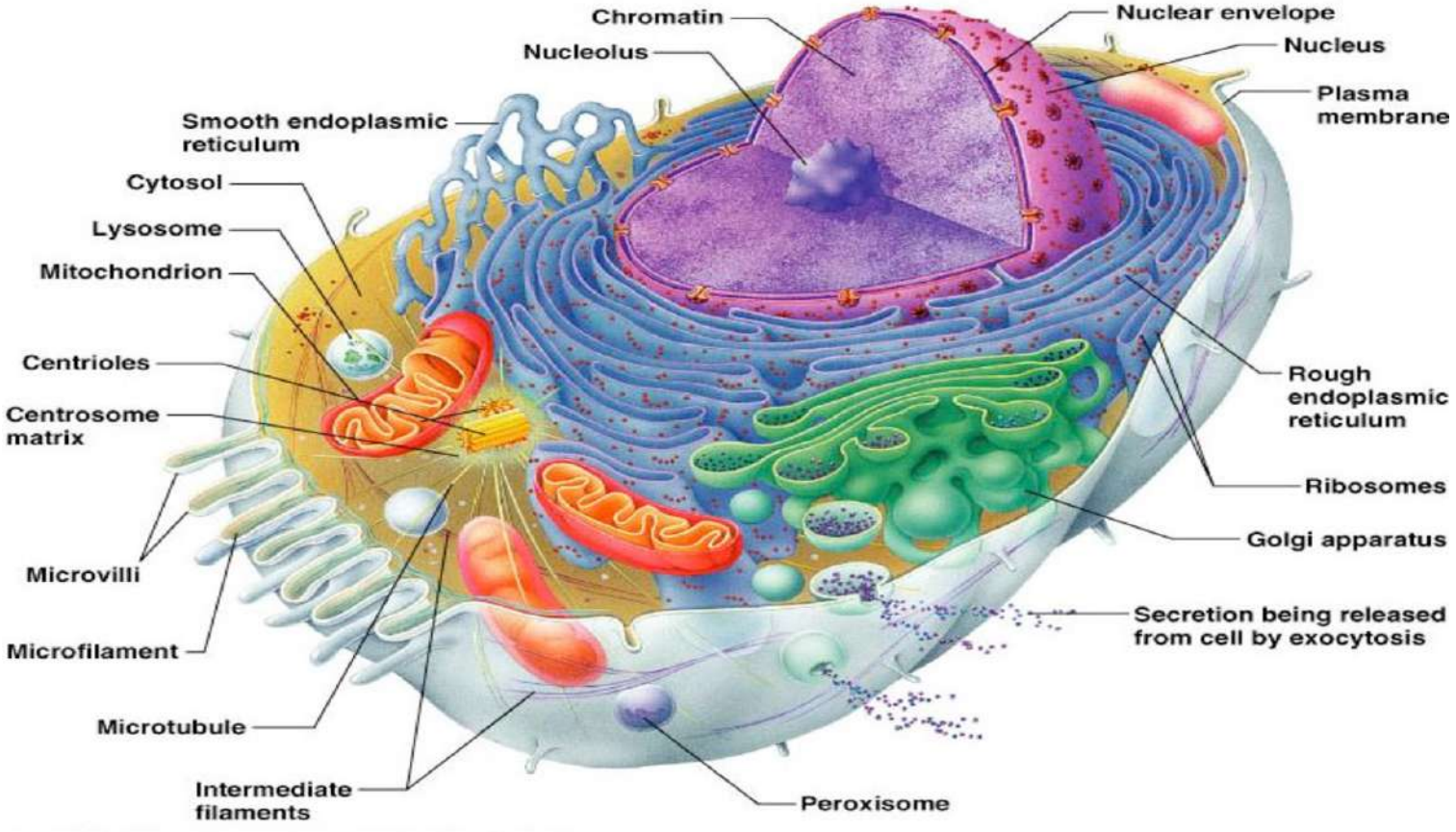


Joachim Frank,
Biochemistry and Molecular Biophysics/Biological Sciences
Columbia University

NYSBC spring course 2022

72nd Street: Eleanore Supports Ukraine

Structure of a Generalized Cell



The Nobel Prize in Chemistry 2017
Jacques Dubochet, Joachim Frank, Richard Henderson

Share this: 991

The Nobel Prize in Chemistry 2017

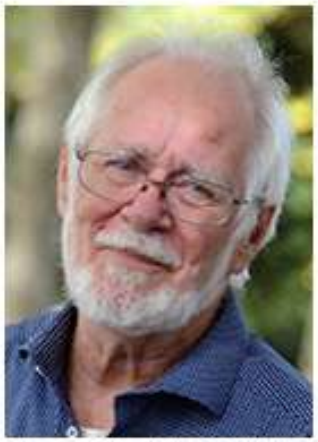


Photo: Félix Imhof © UNIL [CC BY-SA 4.0]
Jacques Dubochet
Prize share: 1/3

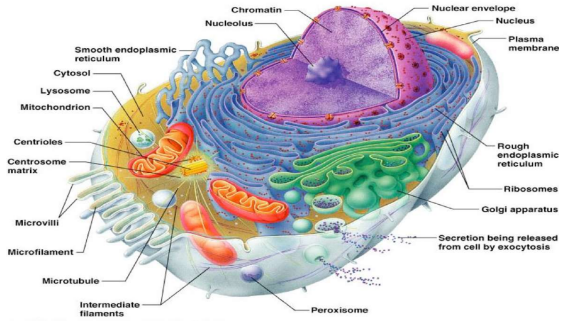
Photo: B. Winkowski © Columbia University Medical Center
Joachim Frank
Prize share: 1/3

Photo: MRC Laboratory of Molecular Biology
Richard Henderson
Prize share: 1/3

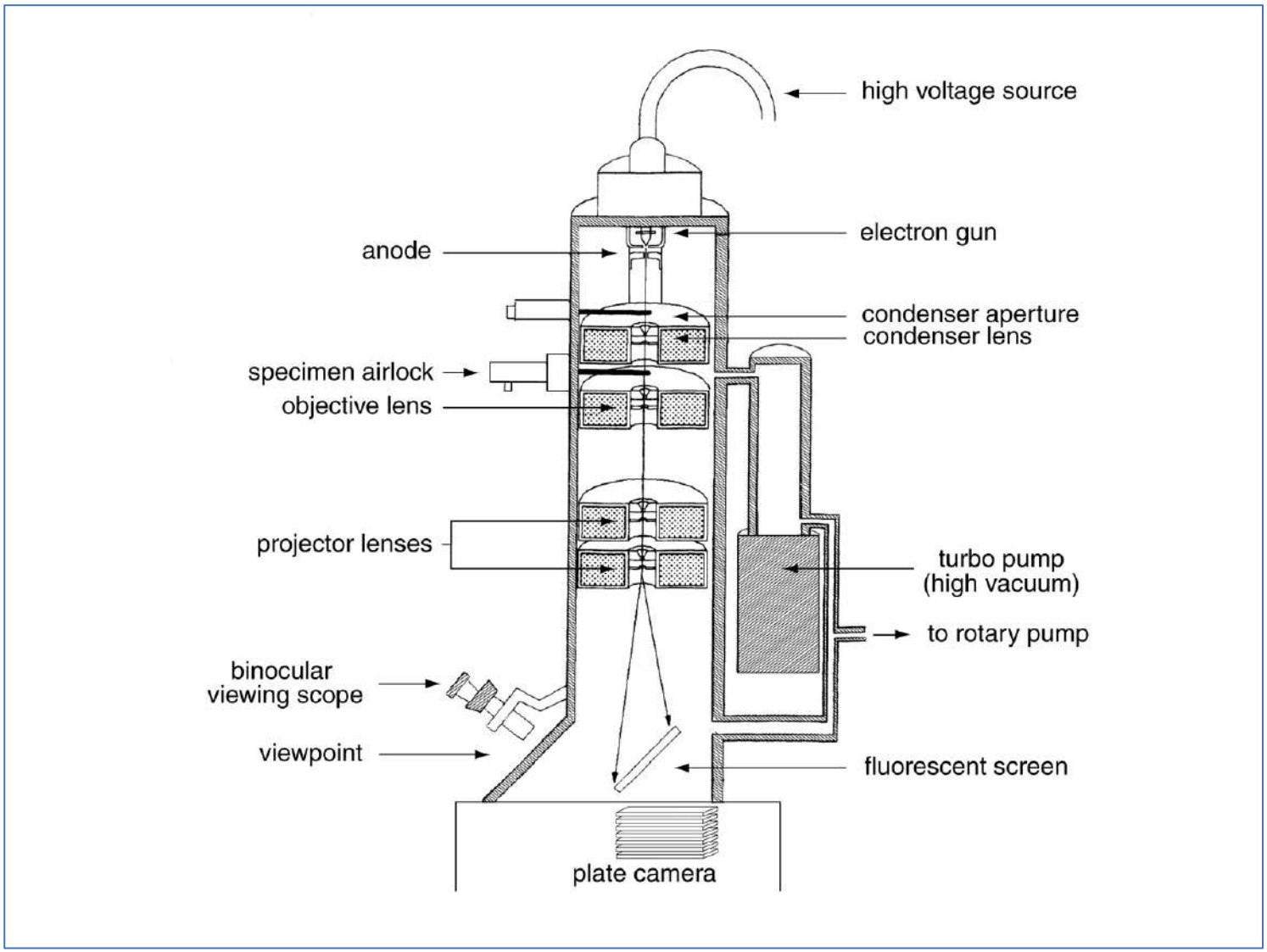
The Nobel Prize in Chemistry 2017 was awarded to Jacques Dubochet, Joachim Frank and Richard Henderson *"for developing cryo-electron microscopy for the high-resolution structure determination of biomolecules in solution".*

". . . for developing cryo-electron microscopy for the high-resolution structure determination of biomolecules in solution."

"REDUCTIONISM"

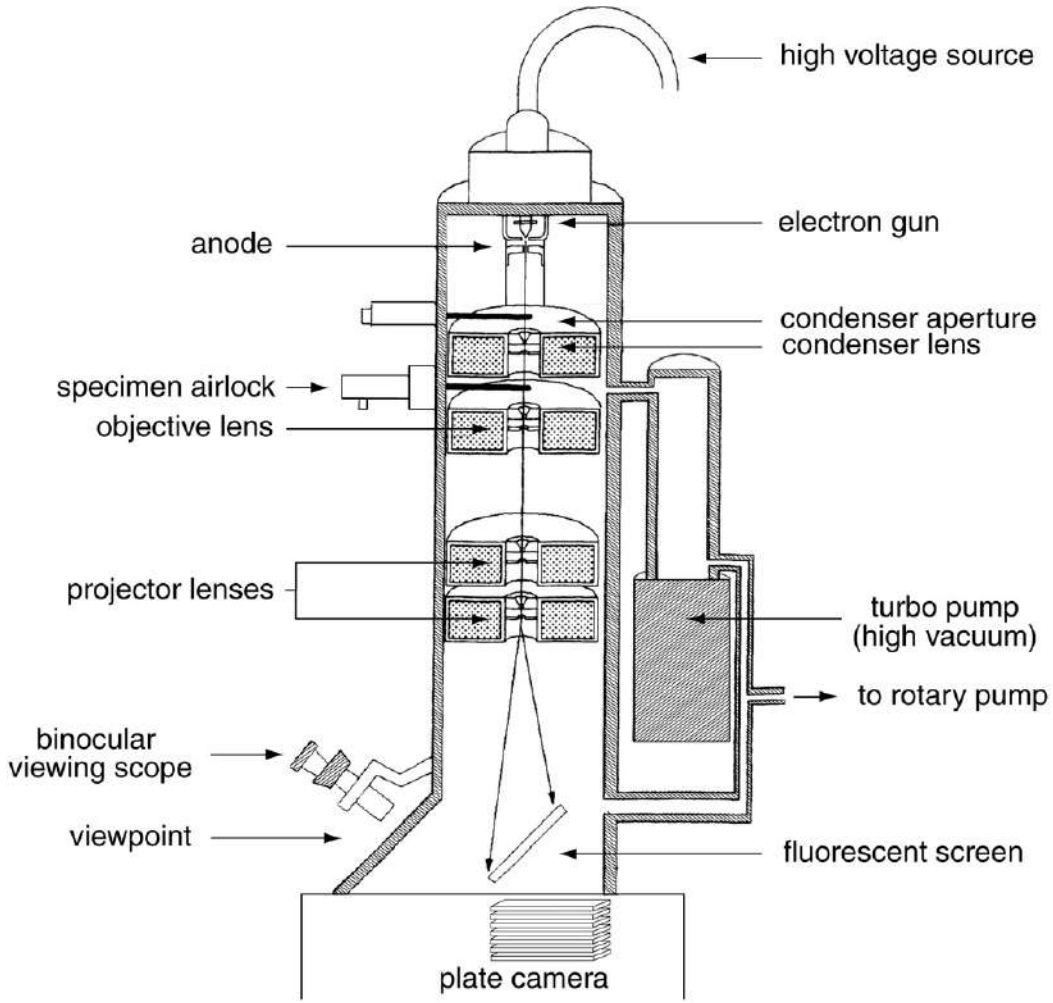


→



The Transmission Electron Microscope (TEM)

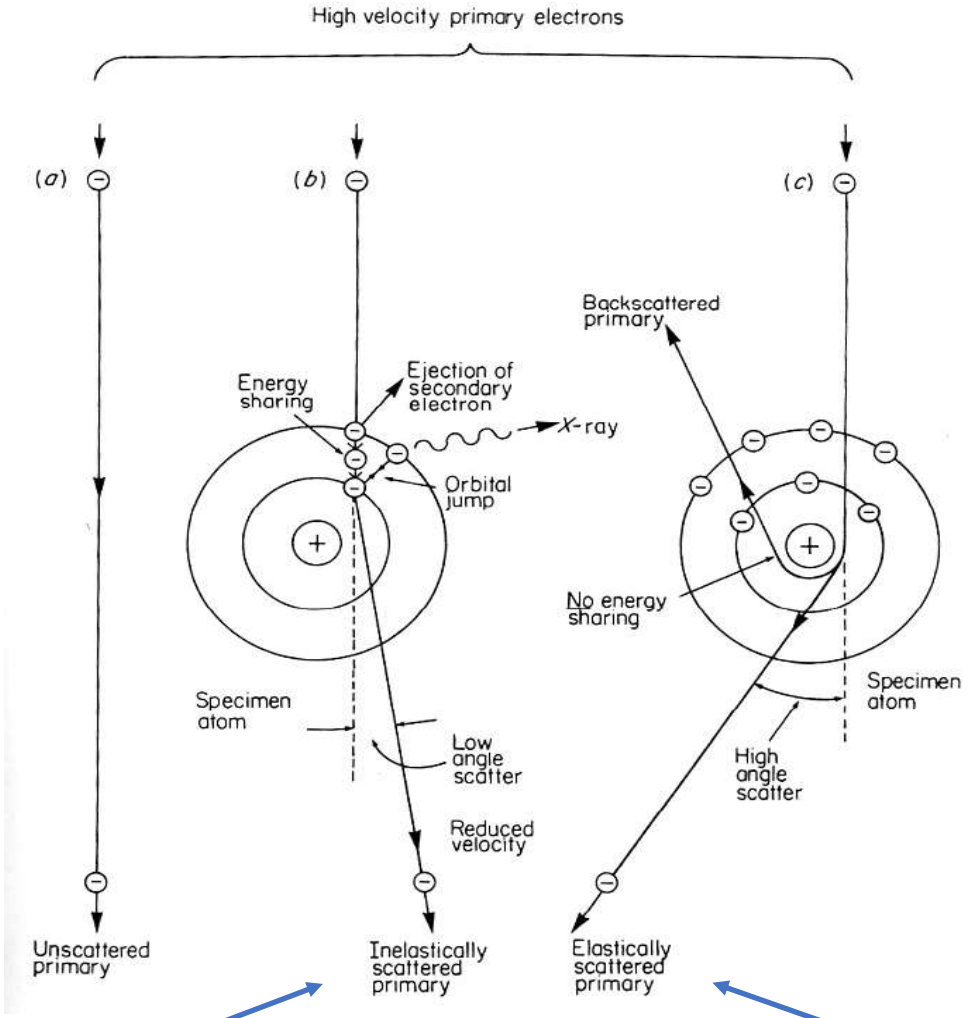
transmission → image is formed from electrons that have penetrated the sample



The Transmission Electron Microscope (TEM)

transmission → image is formed from electrons that have penetrated the sample

Three classes of scattering outcomes



Forward scattered beam "sees" the Coulomb potential

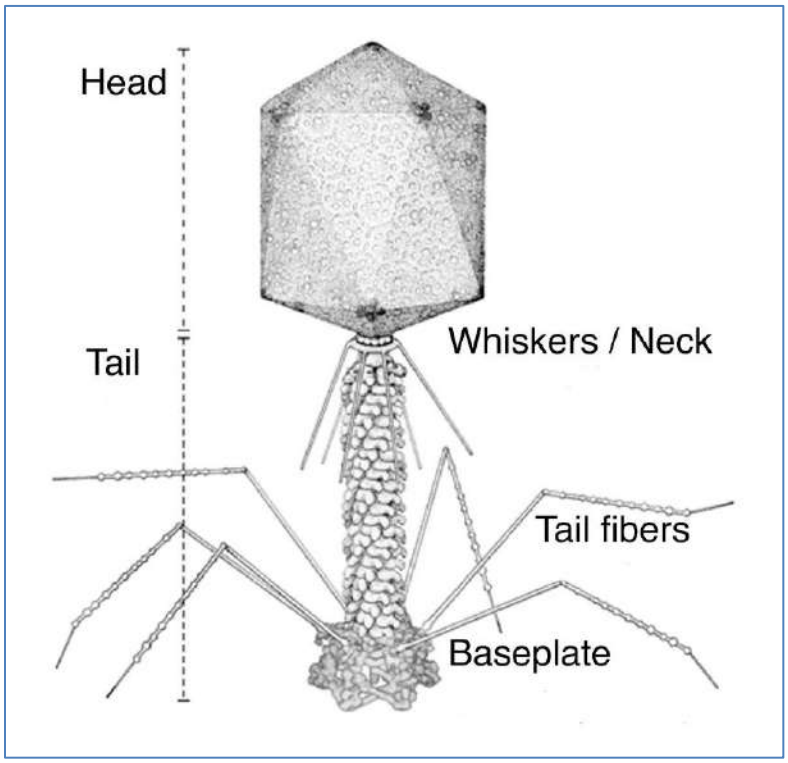
Low-resolution image
energy loss, incoherent

High-resolution image
no energy loss, coherent


Object projection =
integral over Coulomb
potential distribution

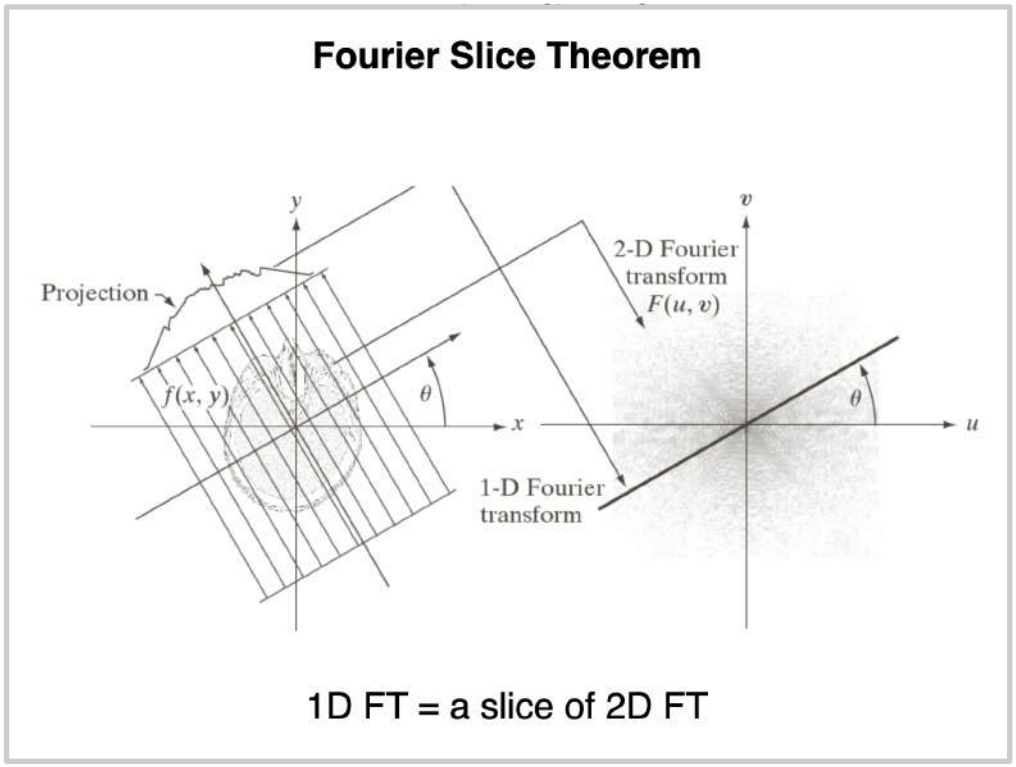
3D reconstruction of biological molecules – how did it get started?

David DeRosier and Aaron Klug, Nature 1968



Three-dimensional reconstruction

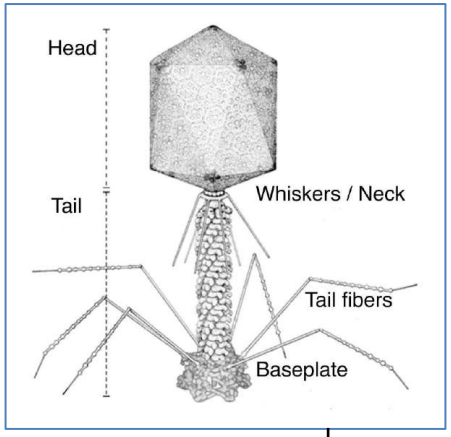
projection-slice theorem,
central slice theorem, or



Johannes Radon, 1917

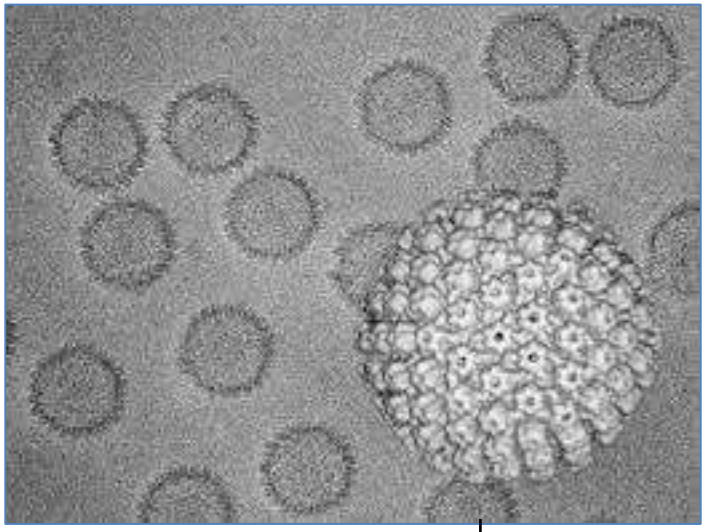
EM of crystalline samples (3 different types of crystal order)

HELICAL



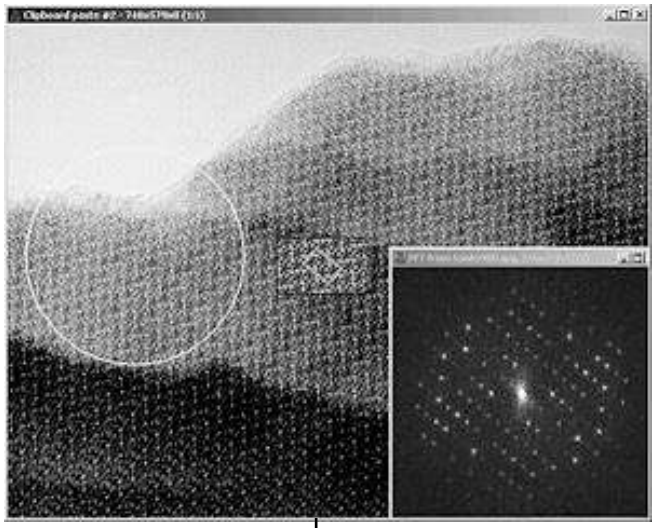
DE ROSIER AND KLUG 1968

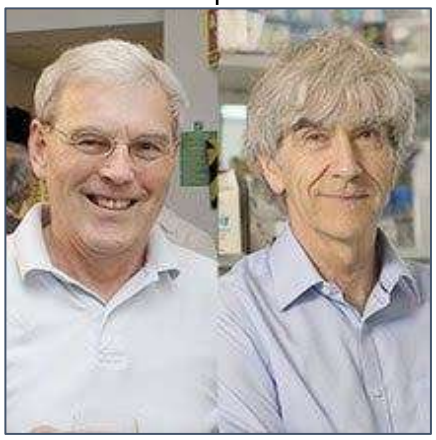
ICOSAHEDRAL (VIRUSES)



CROWTHER 1970

2-DIMENSIONAL

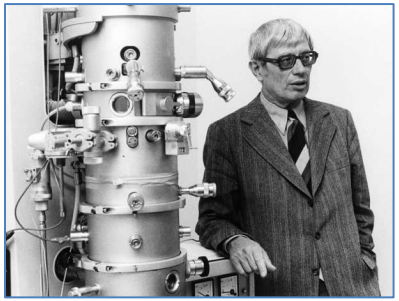




HENDERSON AND UNWIN 1975

X-ray and EM roots

My mentor Walter Hoppe, originally an X-ray crystallographer, turned his attention to electron microscopy in the early 1960s.



Walter Hoppe

Important for the rapid dissemination of ideas and techniques in the early days, Hoppe along with Max Perutz co-organized several meetings, starting with one in Hirschegg (1968), that brought X-ray crystallographers together with the early pioneers of electron crystallography.

Max Perutz/John Kendrew

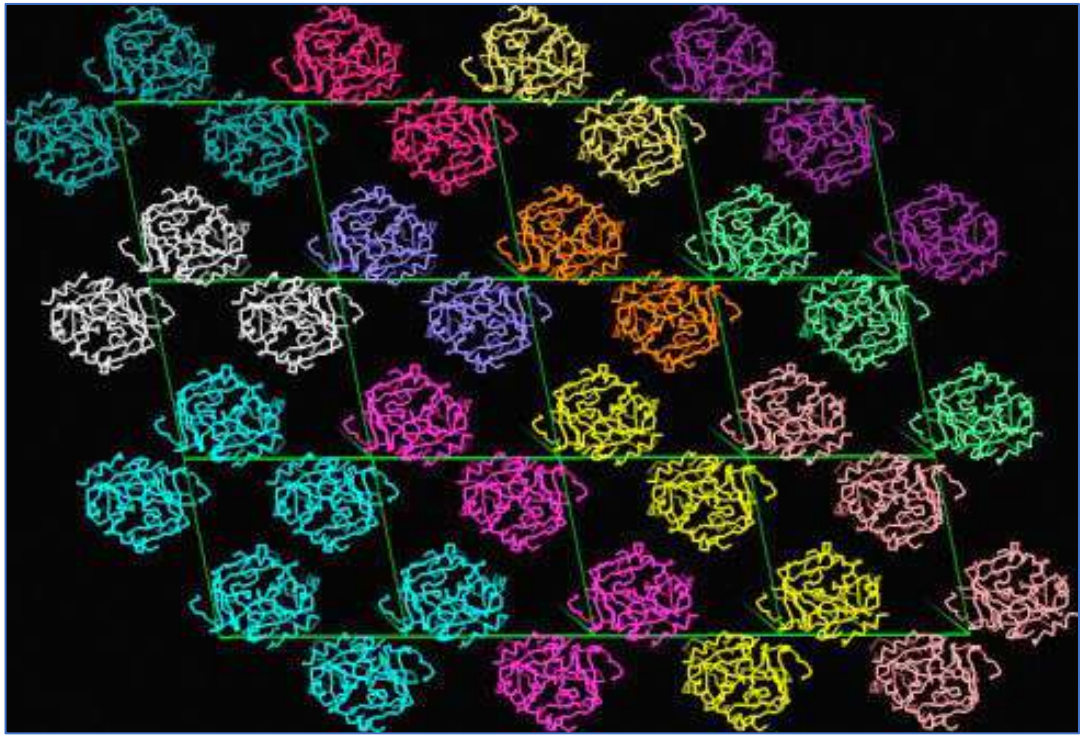
One of the first Hybrid meetings!

Hirschegg, site of 1968 workshop on X-ray crystallography and EM of proteins organized by Walter Hoppe and Max Perutz.

Harold Erickson, Richard Henderson, Ken Holmes, Hugh Huxley, and Nigel Unwin

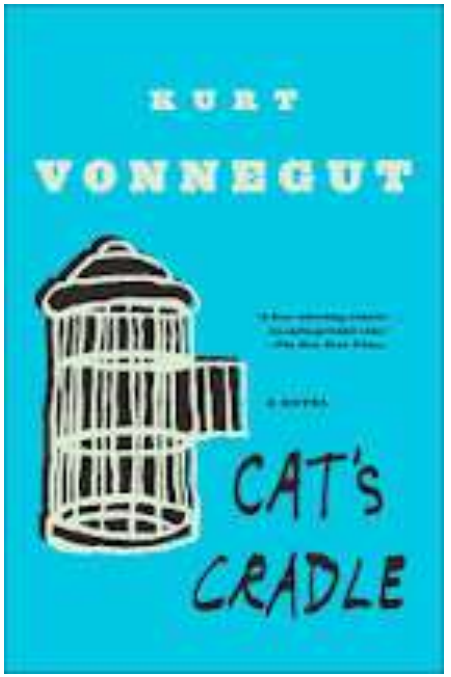
. . . and many others

Crystals are not good examples of life-like environments for molecules. They act as energy traps and make me think of ice nine.

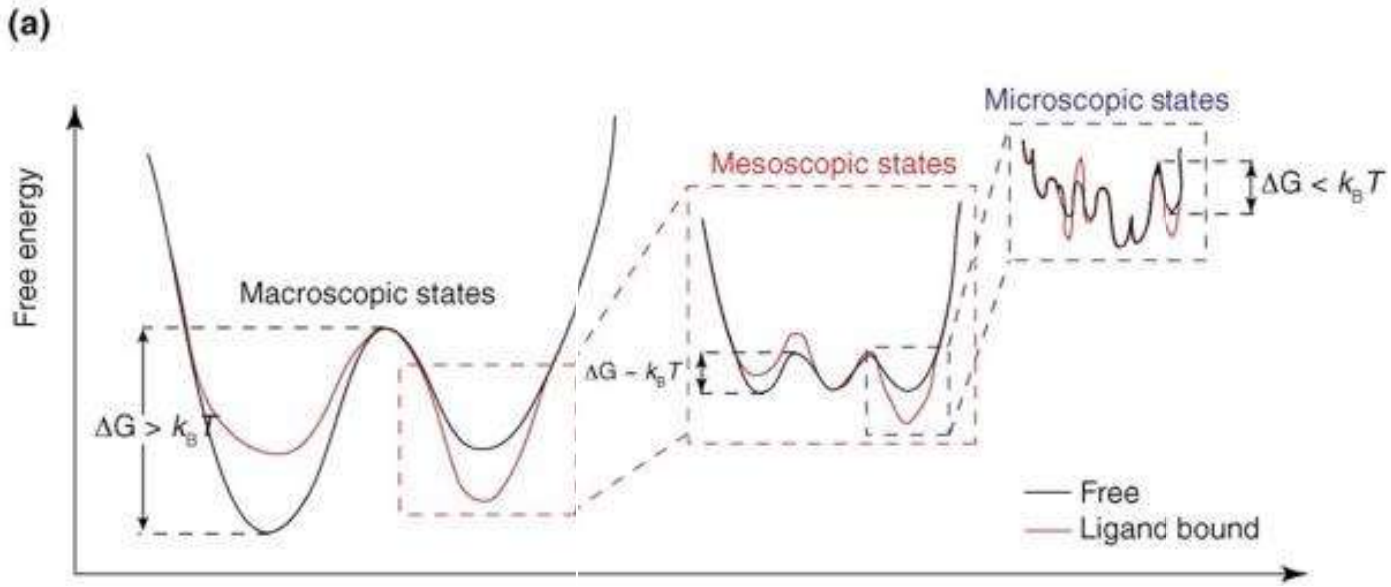


Ice-nine vs. conditions for Life

Ice-nine is a fictitious alternative structure of water that is solid at room temperature. When a crystal of ice-nine contacts liquid water, it acts as a seed crystal that makes the molecules of liquid water arrange themselves into the solid form of lowest energy, namely ice-nine. **Life ceases under these conditions.**

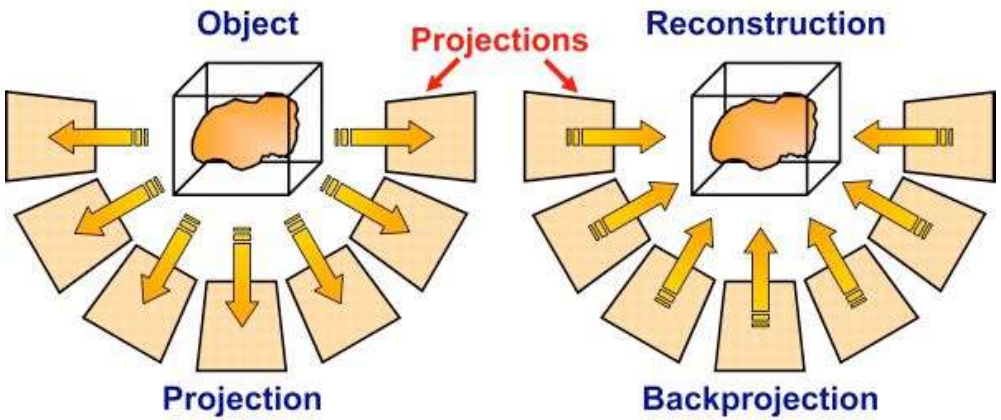


Life requires a multiplicity of states that are accessible at normal temperatures.



Munro et al., Trends Biochem Sci. 2009 Aug; 34(8): 390–400.

Do we need crystals?
3D reconstruction of asymmetrical molecules by electron tomography

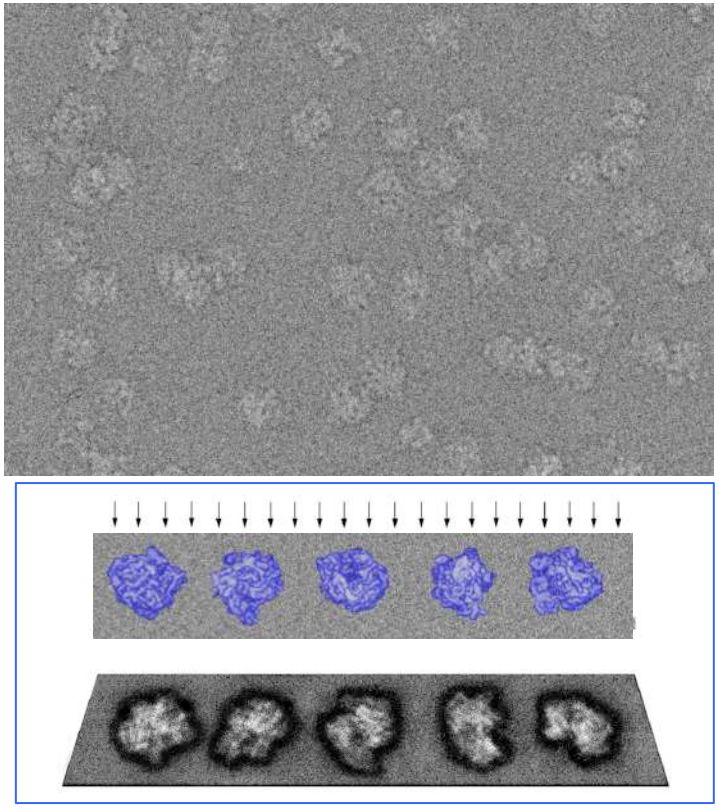


Walter Hoppe (MPG Archive)

- Electron Tomography of single molecules
- Examples: fatty acid synthetase and ribosome
- BUT: Accumulated electron exposure exceeded 1000 e^-/A^2

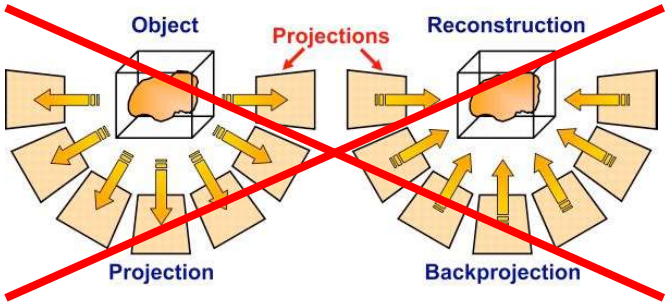
Do we need crystals?
3D reconstruction of asymmetrical molecules by single-particle techniques – the concept

- Single-particle techniques: structural information from images of single (i.e., unattached) molecules in many copies.
- *The molecules are randomly oriented.*
- The molecules are free to assume all naturally occurring conformations.
- *A single snapshot may already give us hundreds of particle views.*
- As we collect more snapshots, more orientations will be covered, until we have enough for reconstructing the molecule in three dimensions.
-



Where did the idea come from?

- Distrust of electron tomographic approach (Walter Hoppe's idea) because of accumulation of radiation damage [this was before the dose fractionation concept]



- Discovery (in my thesis project) that successive micrographs of the same area of a carbon film can be aligned by cross-correlation with high accuracy (better than 3Å) despite shot noise and buildup of contamination between exposures
-

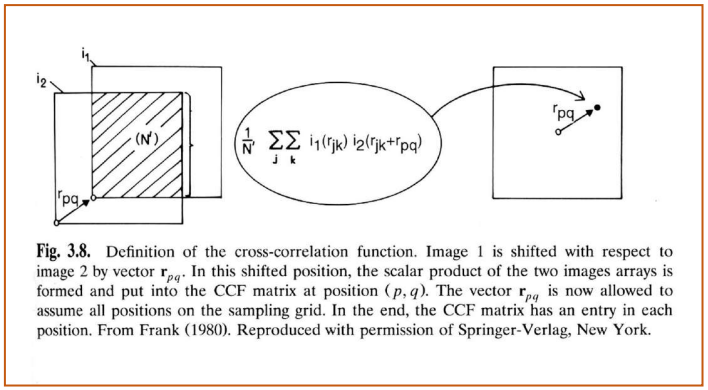
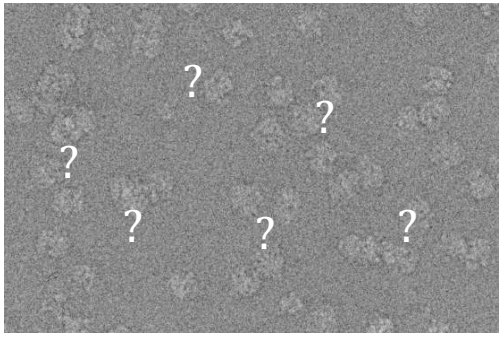
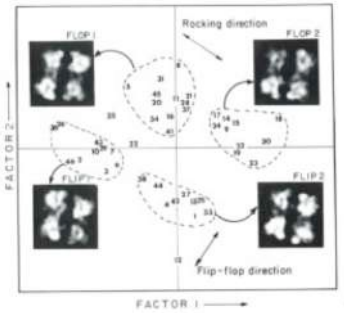


Fig. 3.8. Definition of the cross-correlation function. Image 1 is shifted with respect to image 2 by vector $\mathbf{r}_{pq}$. In this shifted position, the scalar product of the two images arrays is formed and put into the CCF matrix at position (p, q). The vector $\mathbf{r}_{pq}$ is now allowed to assume all positions on the sampling grid. In the end, the CCF matrix has an entry in each position. From Frank (1980). Reproduced with permission of Springer-Verlag, New York.

Why was the idea considered heretic?

- Is a structure even *defined* if it is not part of a crystal?
- Computational problem I: is it possible to determine the orientations of a molecule with unknown structure from its projections?
- Computational problem II: Is it possible to sort molecule images objectively?
- Computational problem III: is it possible to accurately align molecule images taken with low exposure?
- Specimen problem: Are single molecules (i.e., without a surrounding matrix of a crystal) stable enough for high-resolution averages to be formed?
- (when applied to negatively stained samples) What is the relationship between structure inferred from negatively stained molecule and the real structure?
- (when applied to frozen samples) What is the relationship between the structure of a molecule before and after freezing?

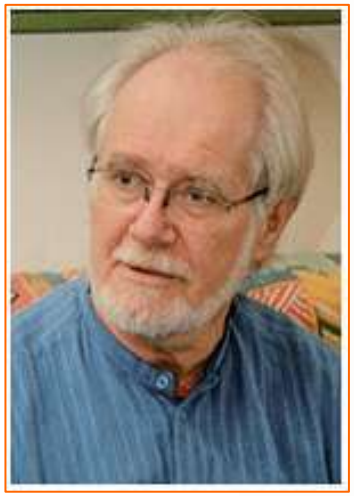




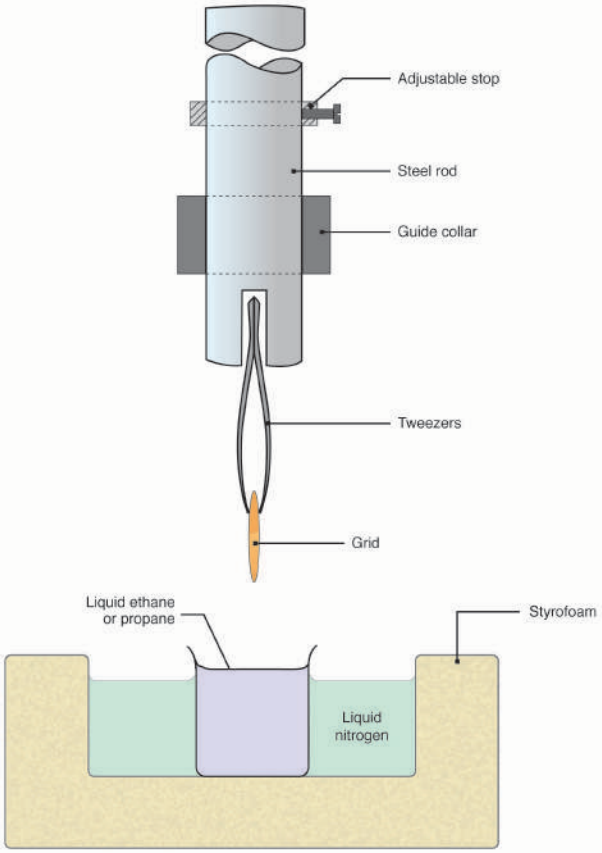
$$D \geq \frac{3}{c^2 dp_{crit}}$$

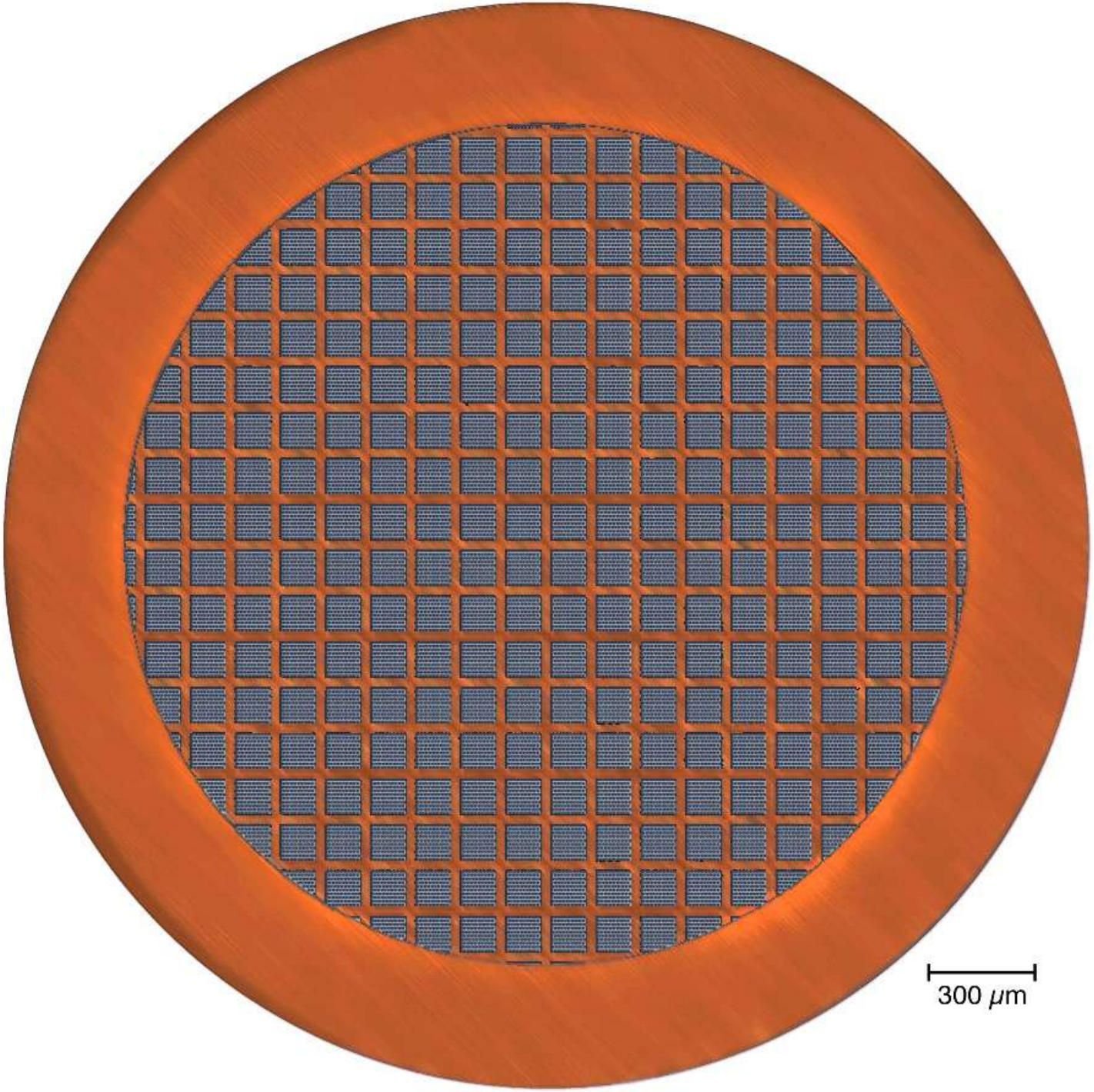
Vitreous ice embedding

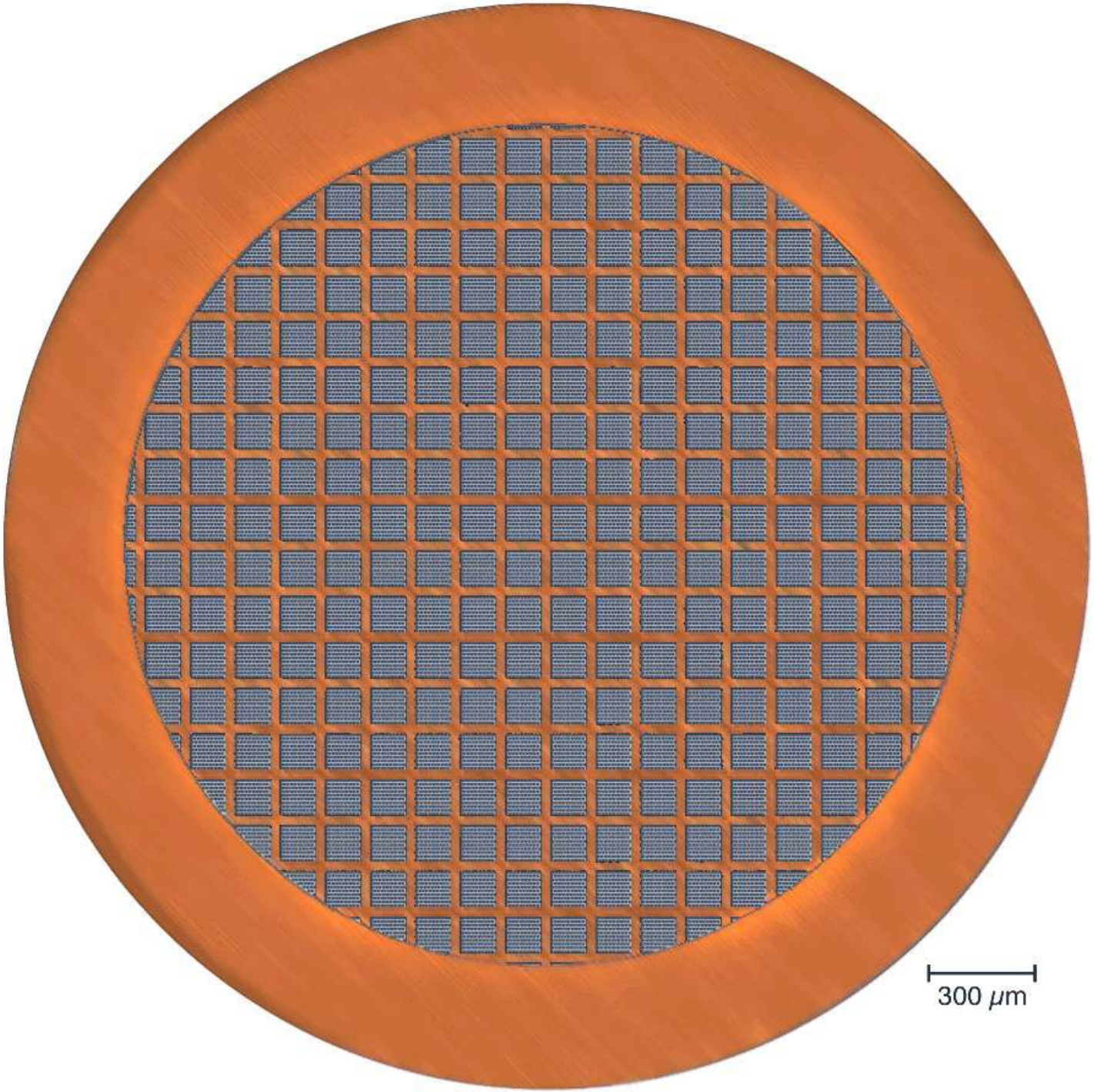
Robert M. Glaeser

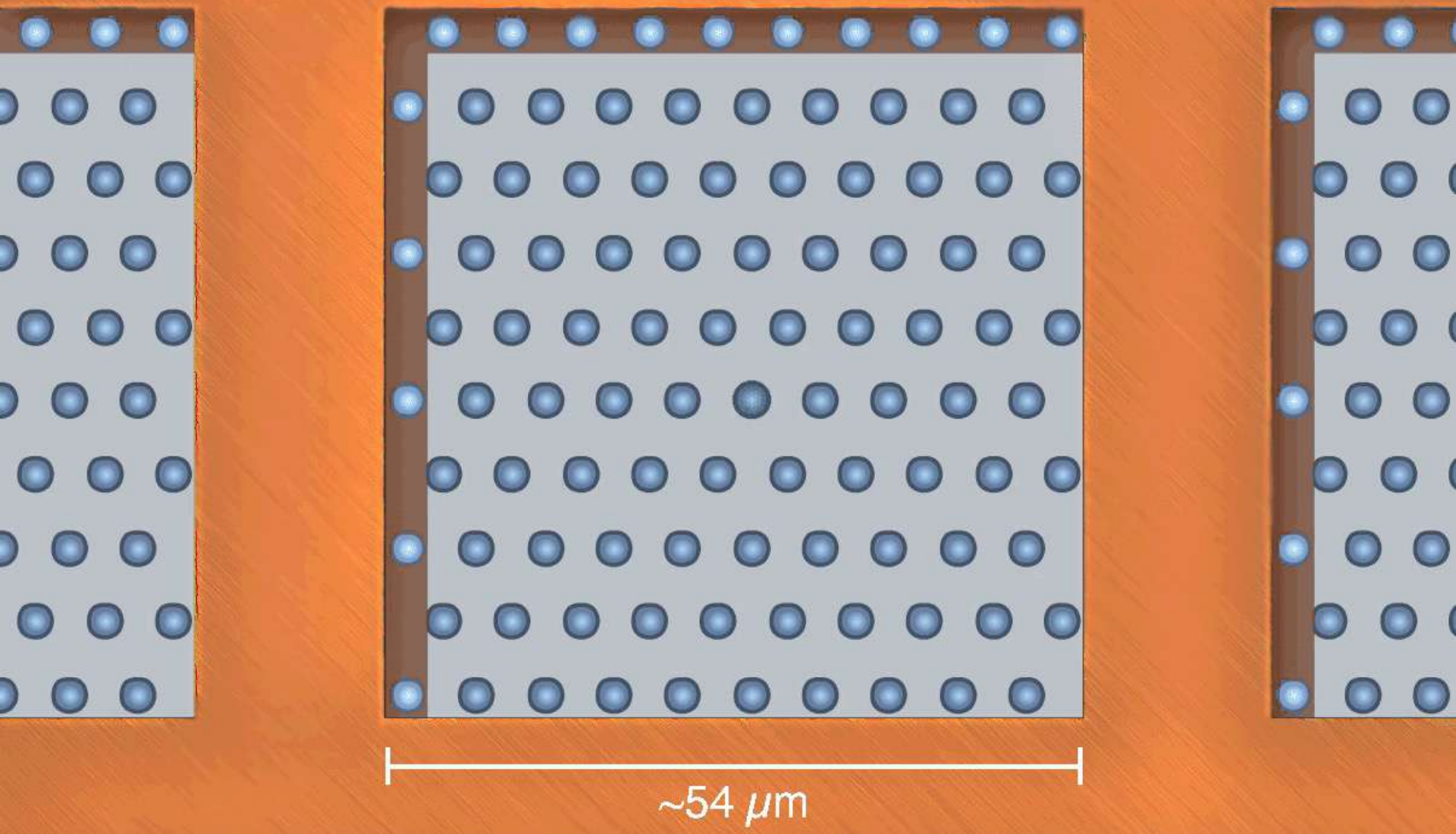


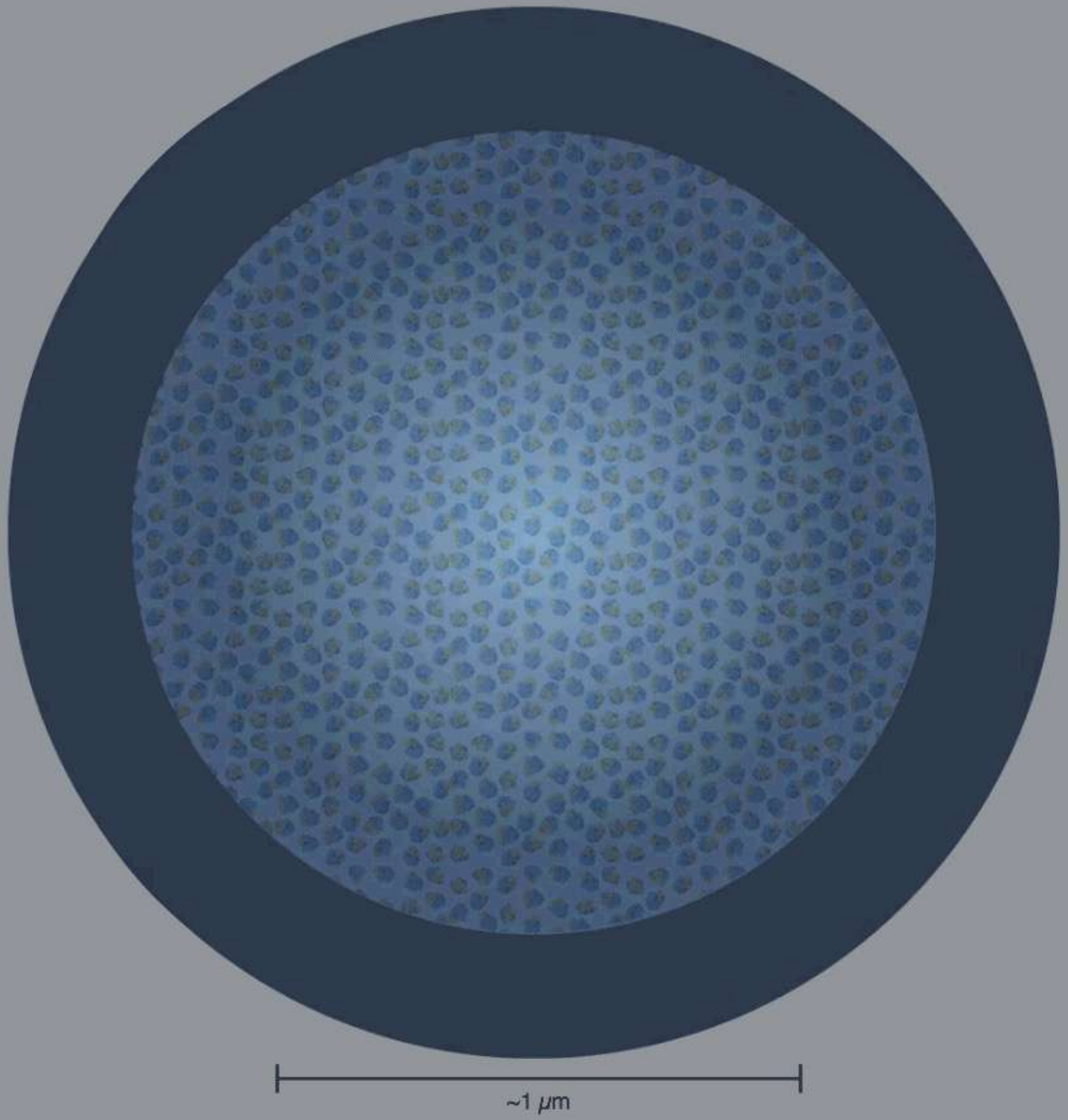
Jacques Dubochet
vitreous ice

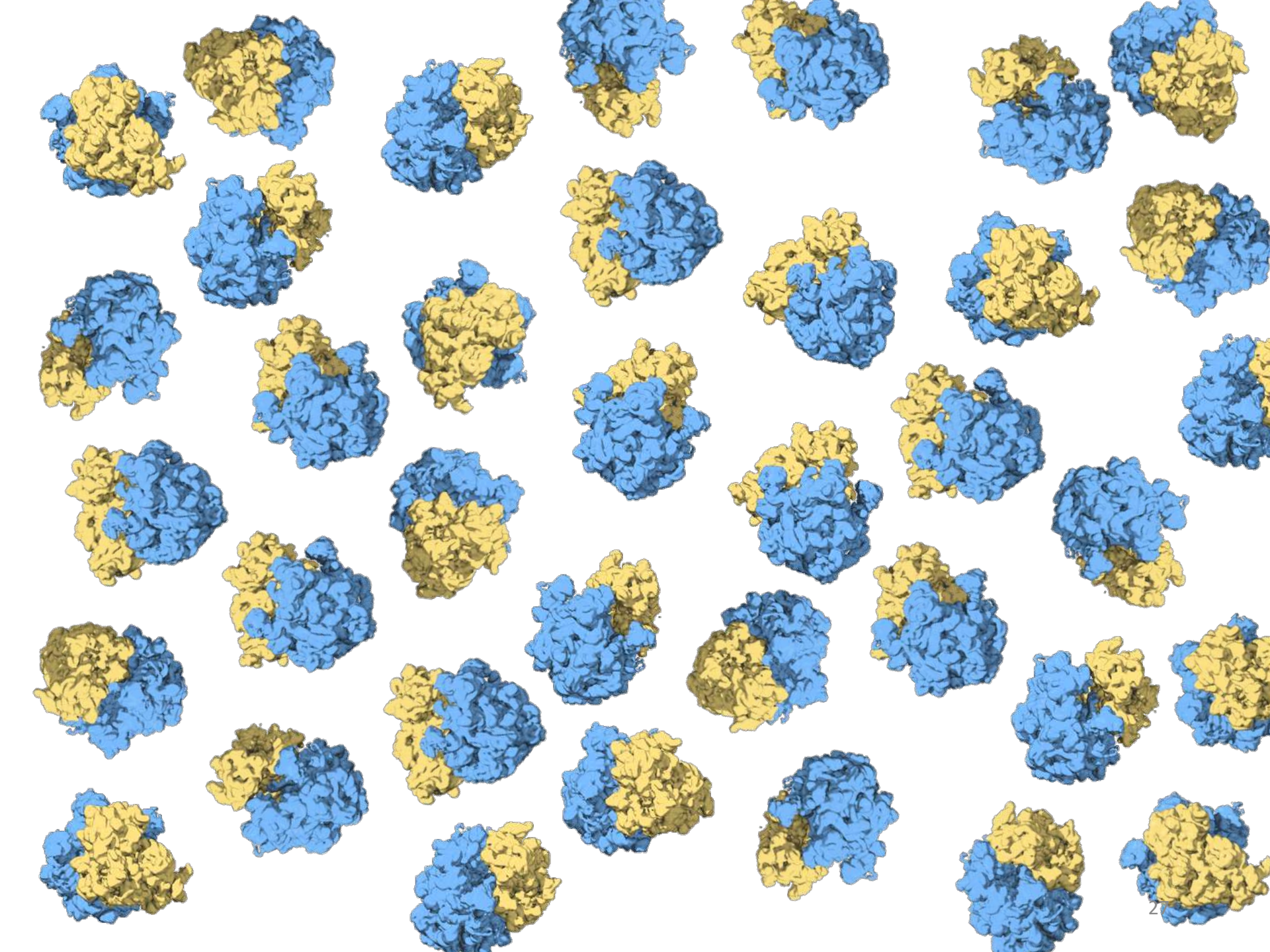


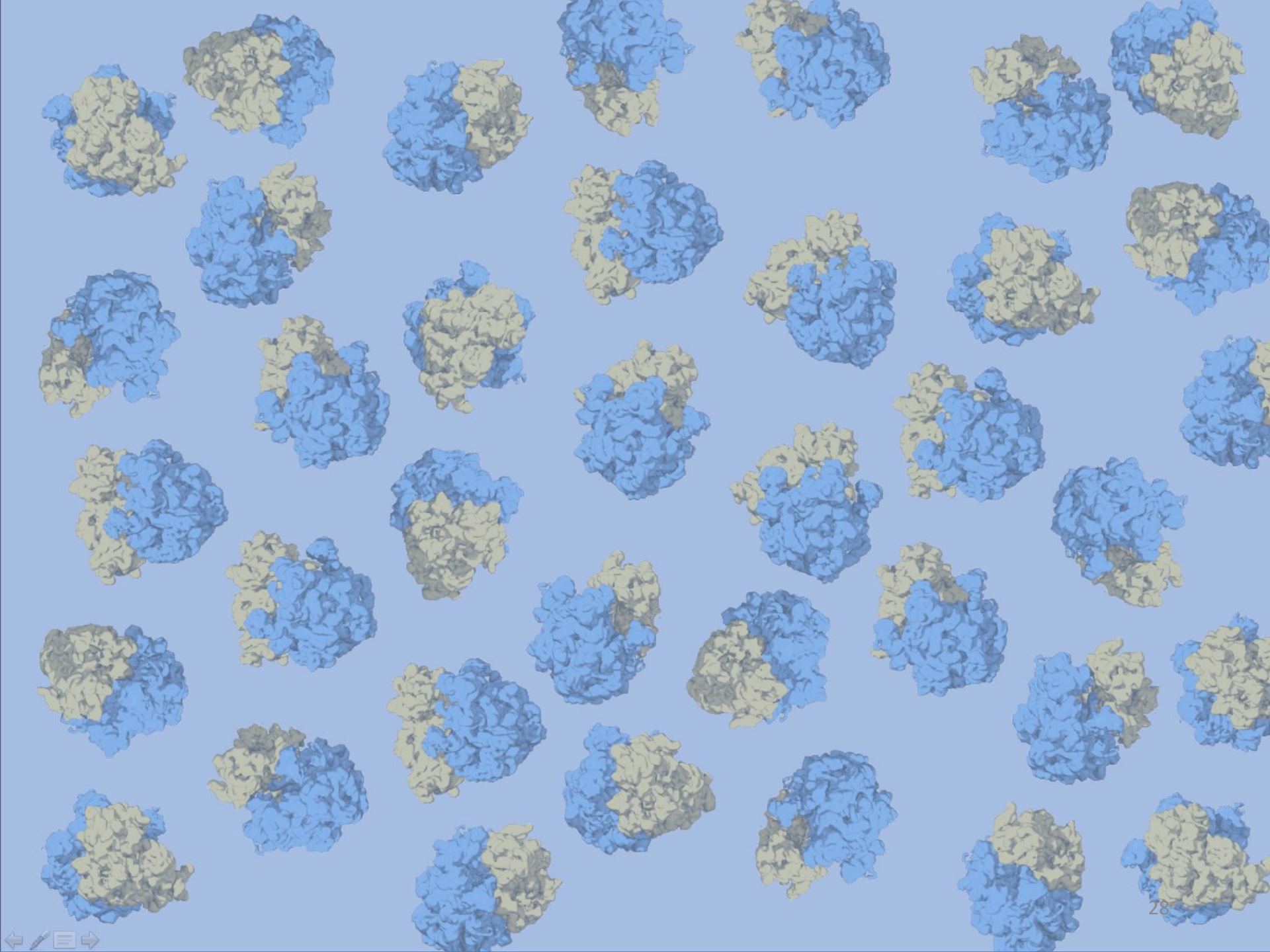

300 μm

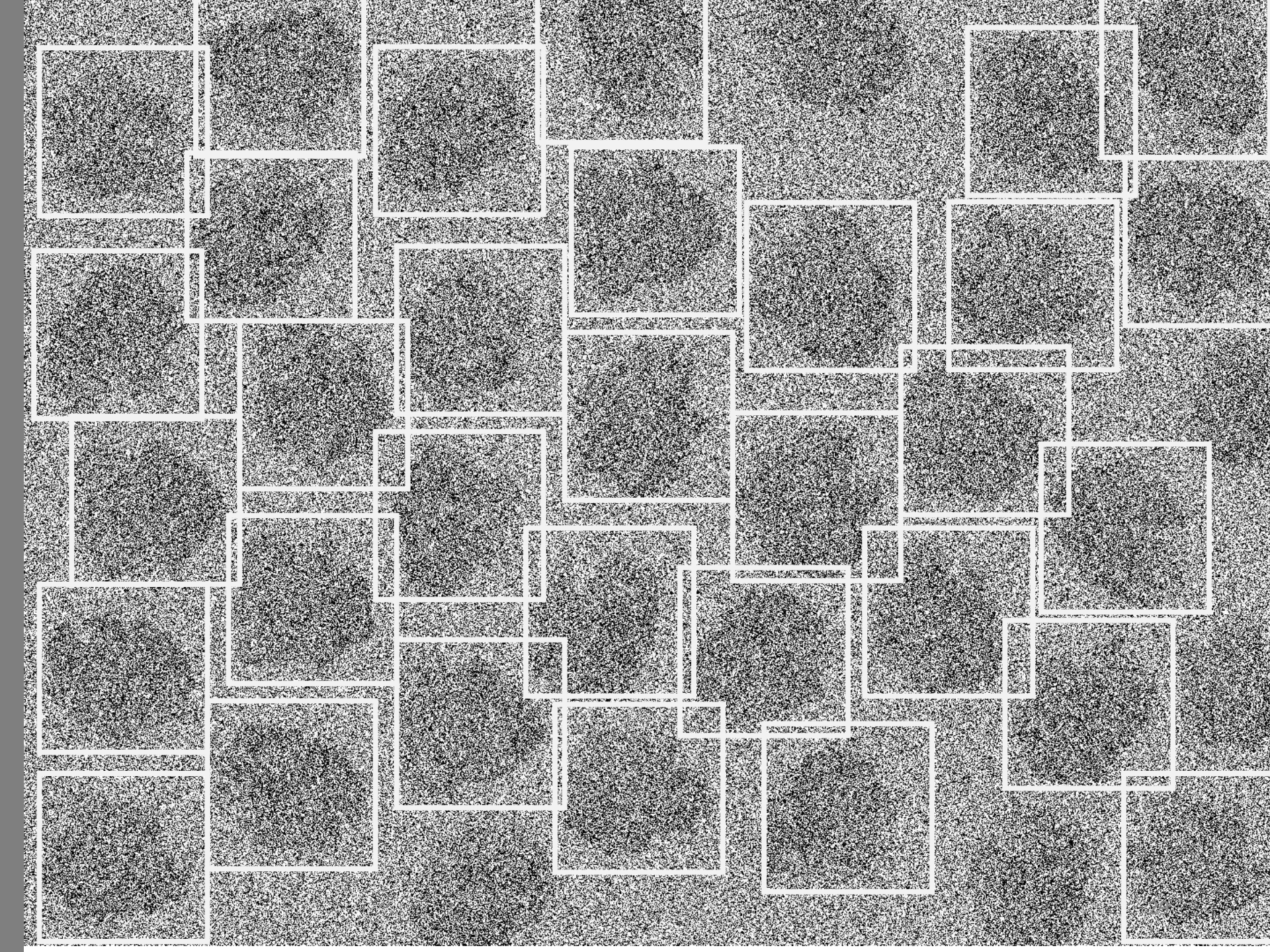

300 μm

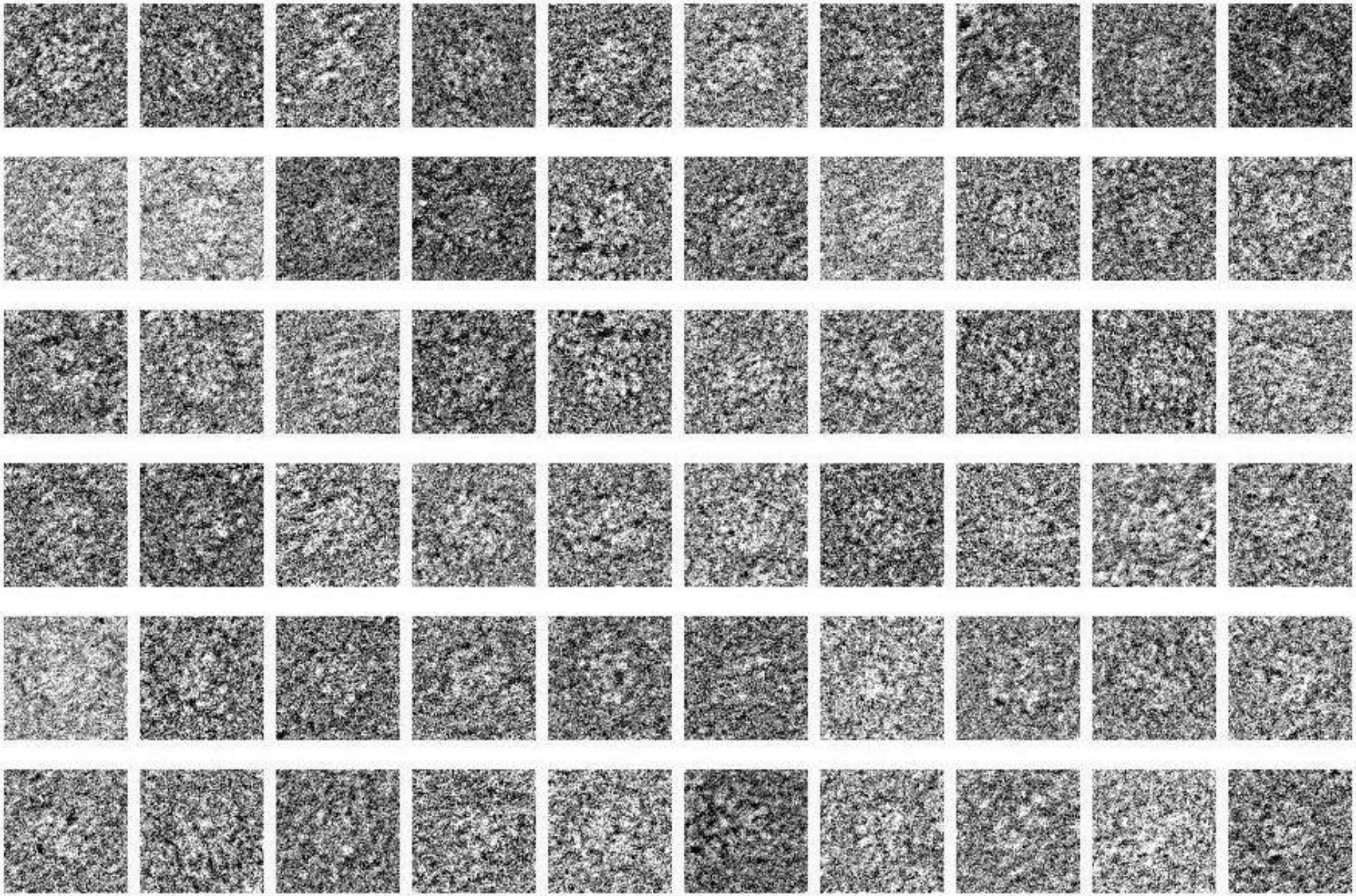

~54 μm


~1 μm

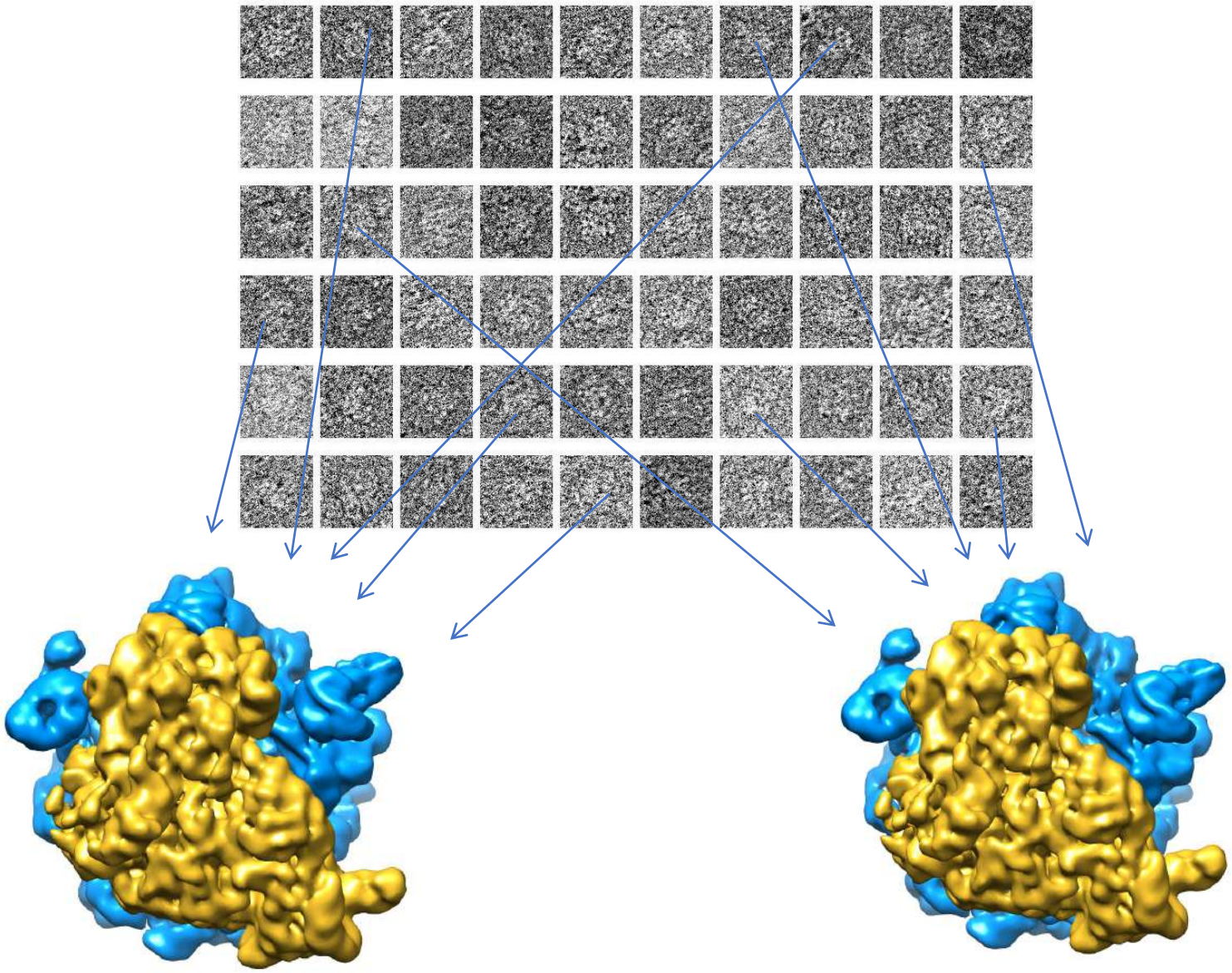








How do we get from the 2D projections to the 3D structures they arise from?



1. Align images

2. Inventory of views: 2D classification

3. Determine viewing direction for each image

4. Classify by 3D structure the image comes from

5. Three-dimensional reconstruction of each class

1. Alignment of single-particle projections ("particles") is achieved by cross-correlation

- Translational cross-correlation function
- Rotational cross-correlation function

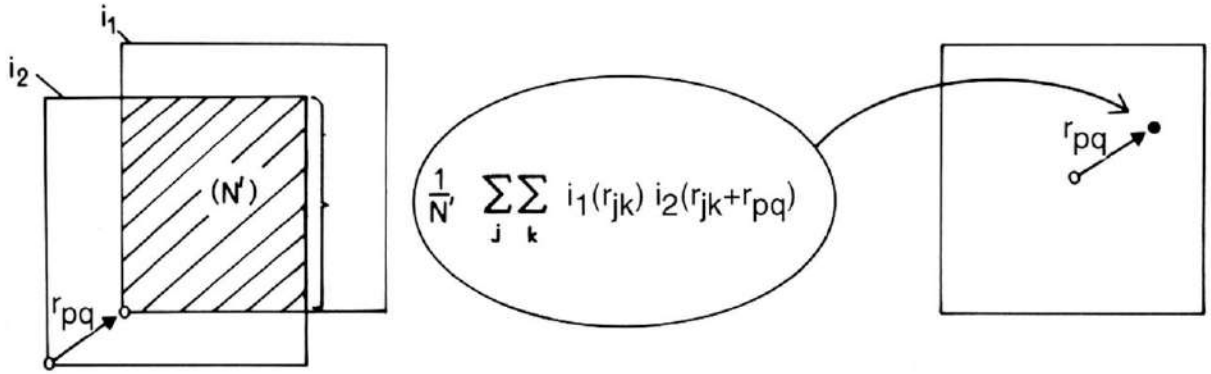
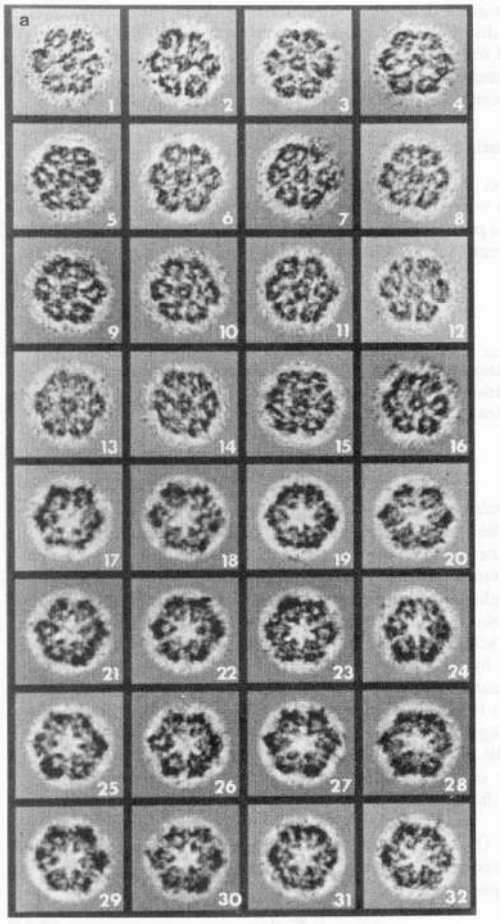
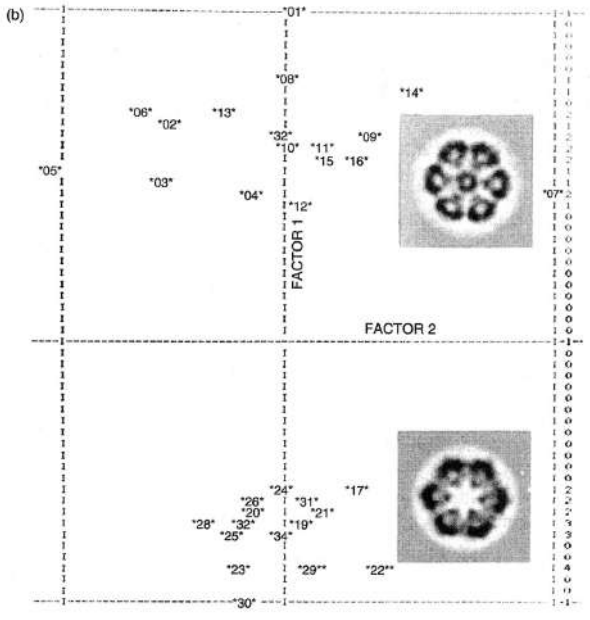


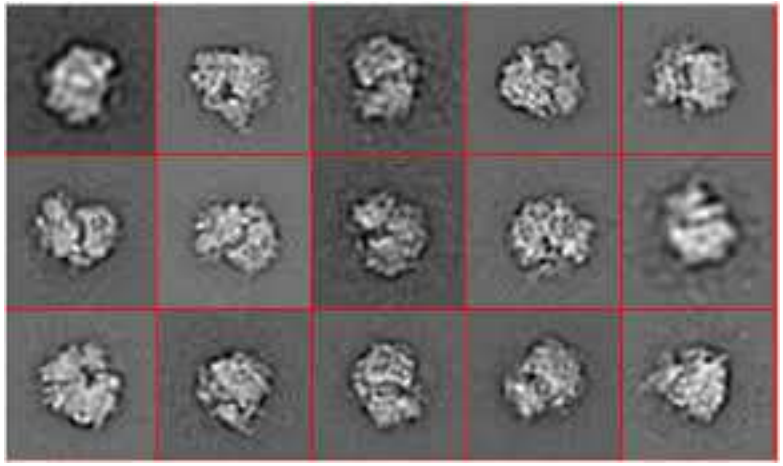
Fig. 3.8. Definition of the cross-correlation function. Image 1 is shifted with respect to image 2 by vector $\mathbf{r}_{pq}$. In this shifted position, the scalar product of the two images arrays is formed and put into the CCF matrix at position (p, q). The vector $\mathbf{r}_{pq}$ is now allowed to assume all positions on the sampling grid. In the end, the CCF matrix has an entry in each position. From Frank (1980). Reproduced with permission of Springer-Verlag, New York.

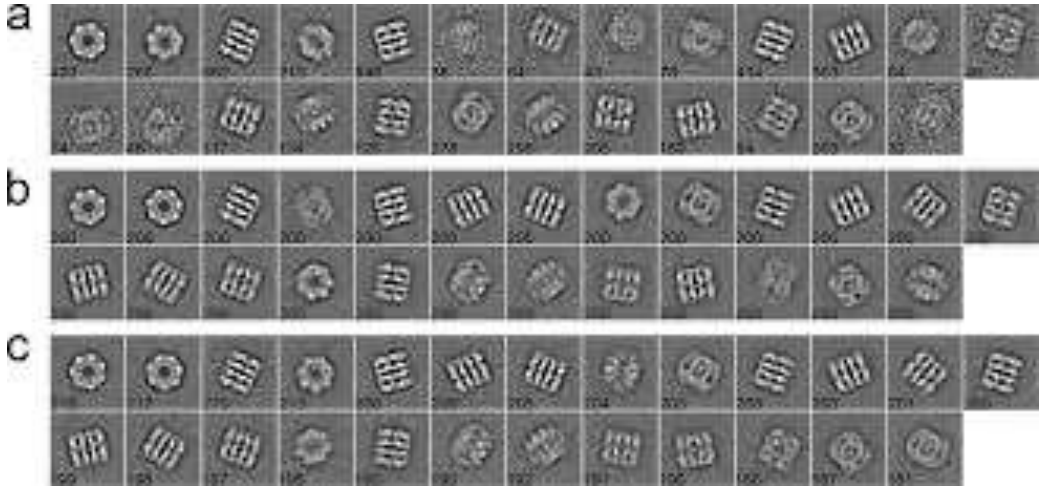
2. Inventory of views

Multivariate statistical analysis (van Heel and Frank, 1981)








a
b
c

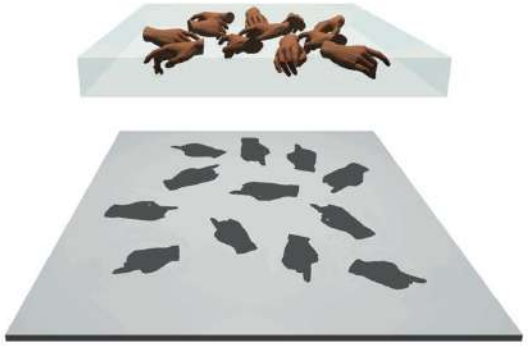
3. Determine viewing directions

bootstrap techniques:

-- random-conical reconstruction

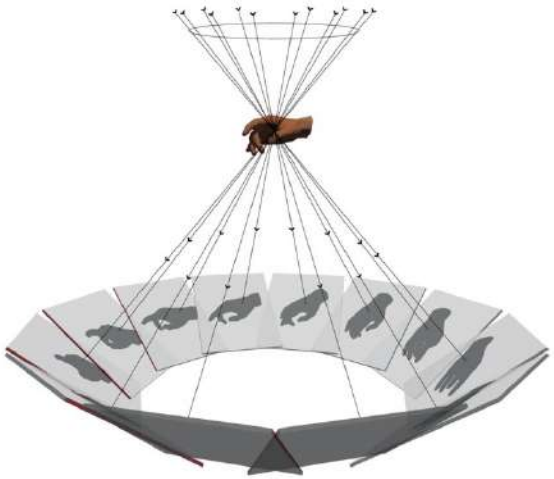
-- “angular reconstitution”

HOW TO DETERMINE ORIENTATIONS OF PARTICLE PROJECTIONS



Random-conical reconstruction

Frank et al., Ultramicroscopy 1978
Radermacher et al., 1986



4. Classify by 3D structures the image comes from

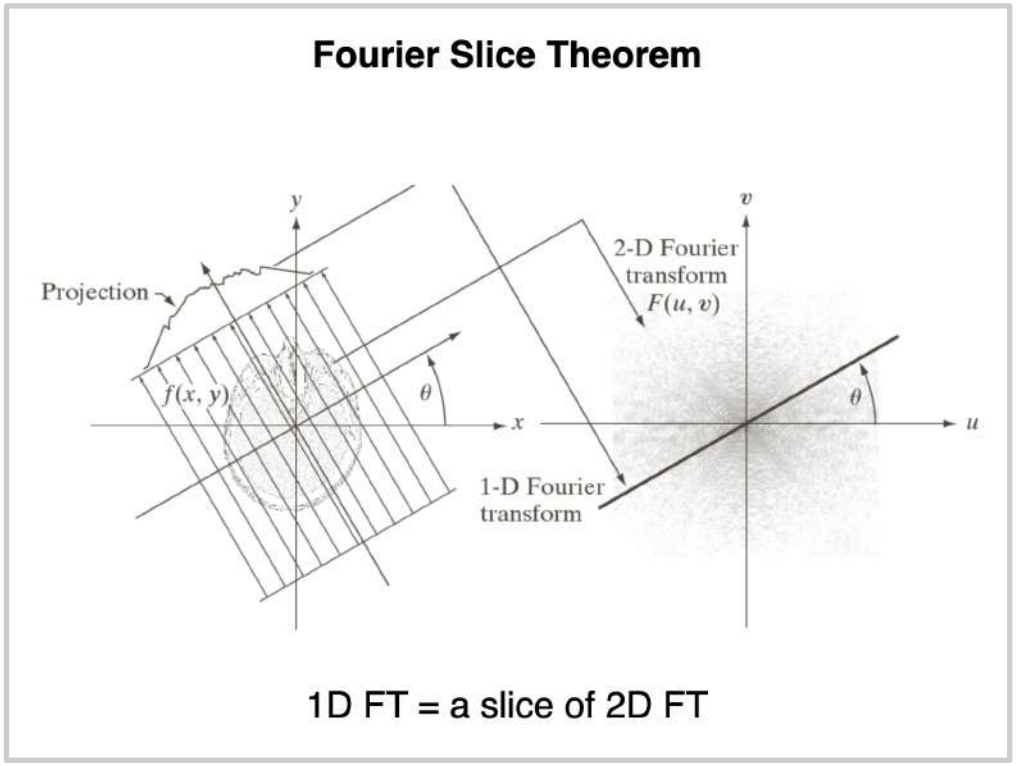
Maximum likelihood classification
S. Scheres et al. Nature Methods 2007
S. Scheres JMB 2012 (Relion)

Sort out subsets, then reconstruct them separately

“Story in a Sample”

5. Three-dimensional reconstruction

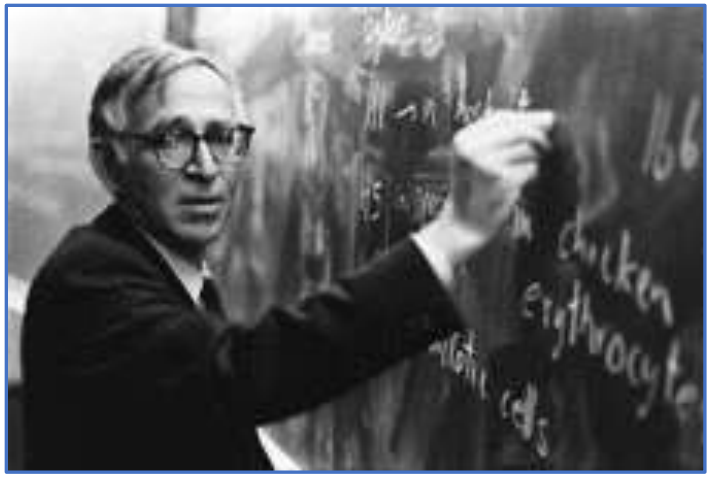
projection-slice theorem,
central slice theorem, or



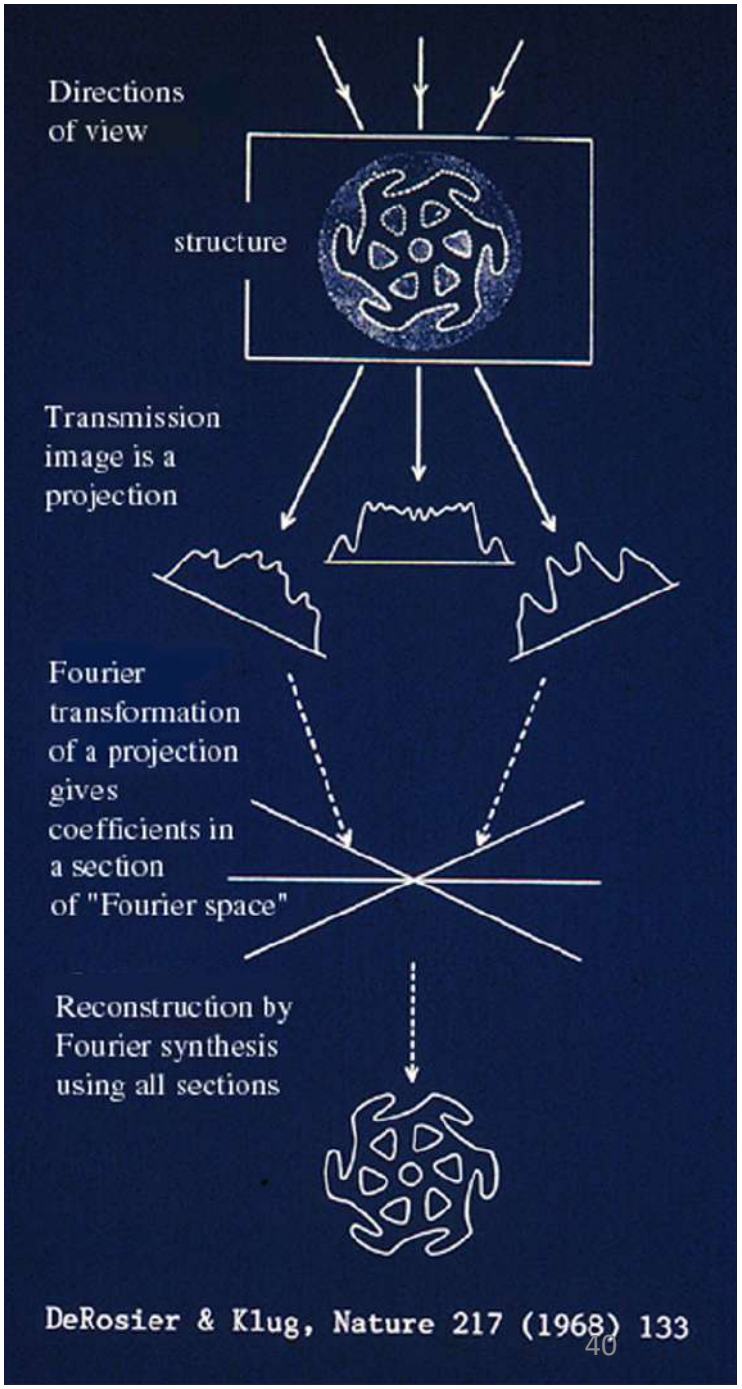
Johannes Radon, 1917

3D RECONSTRUCTION
Multiple TEM views are needed

Pioneering work: 3D reconstruction of bacterio-phage tail using the Fourier-Bessel approach, 1968



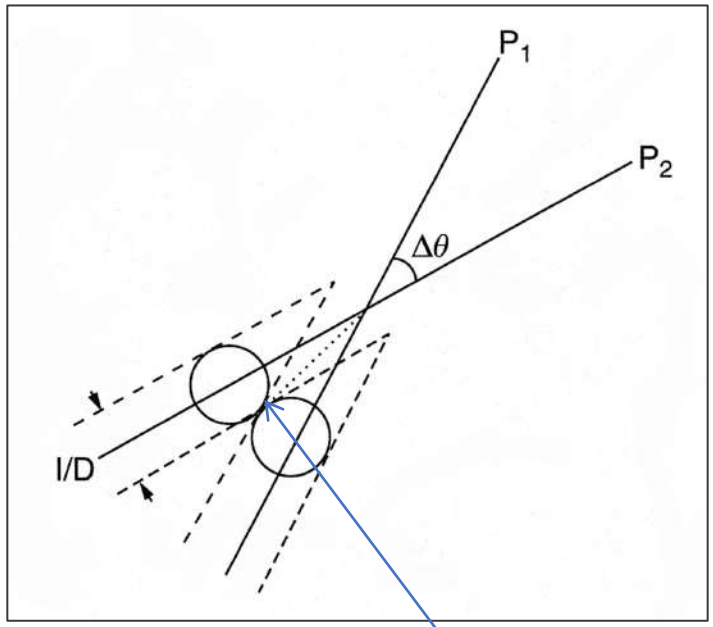
Aaron Klug and David DeRosier, LMB/MRC Cambidge



THREE-DIMENSIONAL RECONSTRUCTION

Crowther's formula:

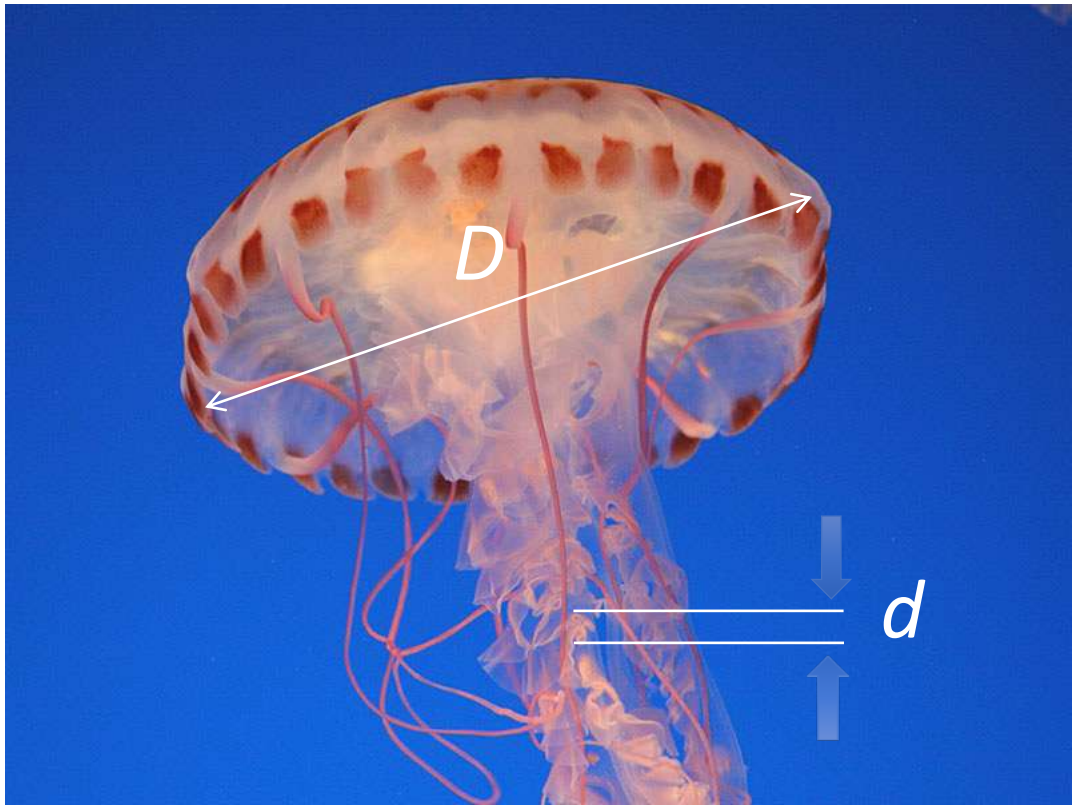
$$N = \frac{\pi D}{d}$$

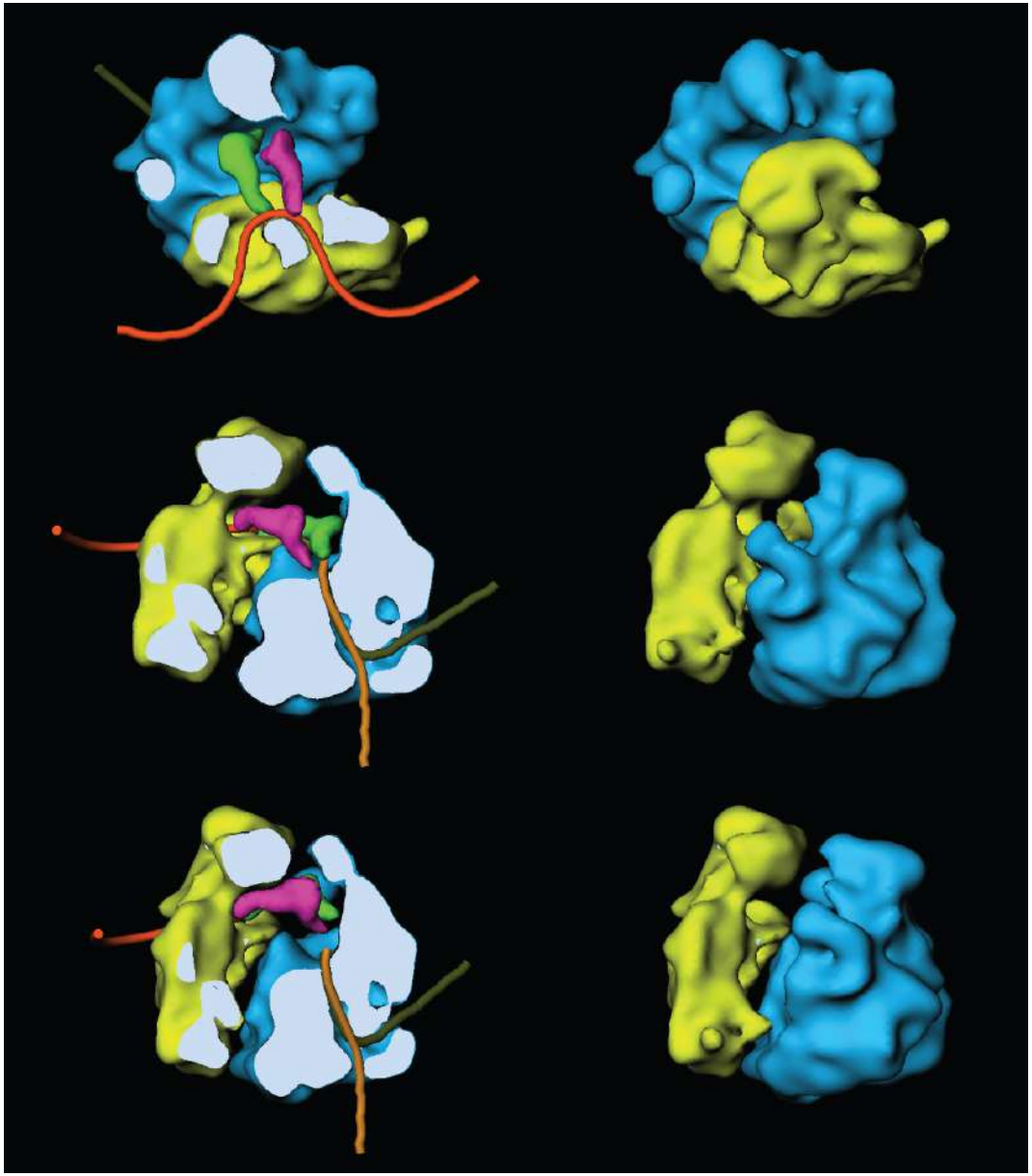


N number of projections
D particle size
d = 1/R = 1/resolution

<u>Example: ribosome at 1/10 Å^{-1} resolution.</u>
N ~ 3 x 250 /10 = 75 equispaced noise-free projections are sufficient to reconstruct the 70S ribosome at 1/10 Å^{-1} resolution!

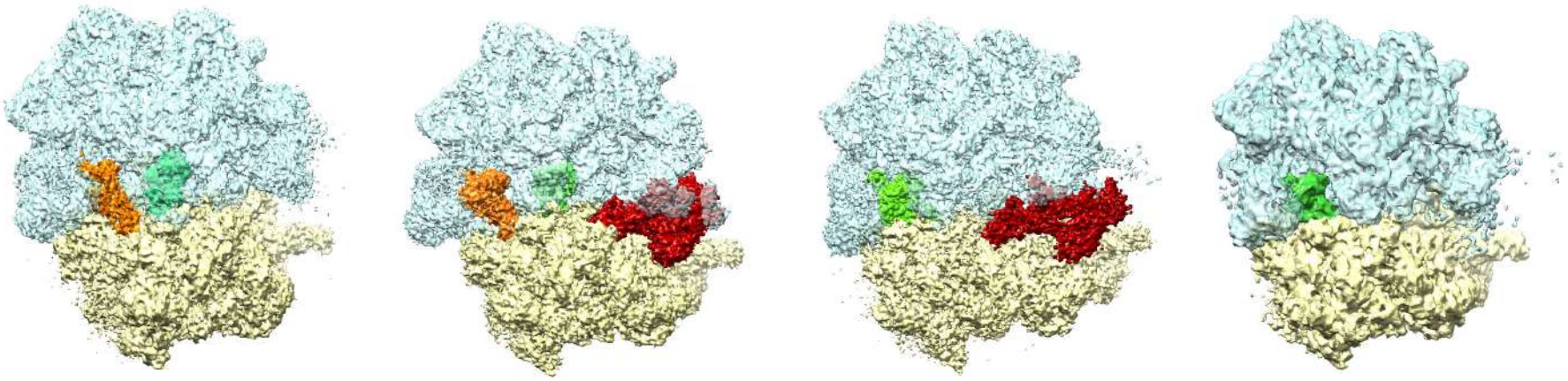
“The larger an object is (*D*), and the smaller the detail (*d*) is that you are interested in, the more views it takes to reconstruct the object.”





Frank et al., Nature 1995

nr 70S--P-E nr 70S--EF-G—P-E r 70S—EF-G—P/E r 70S—P/E

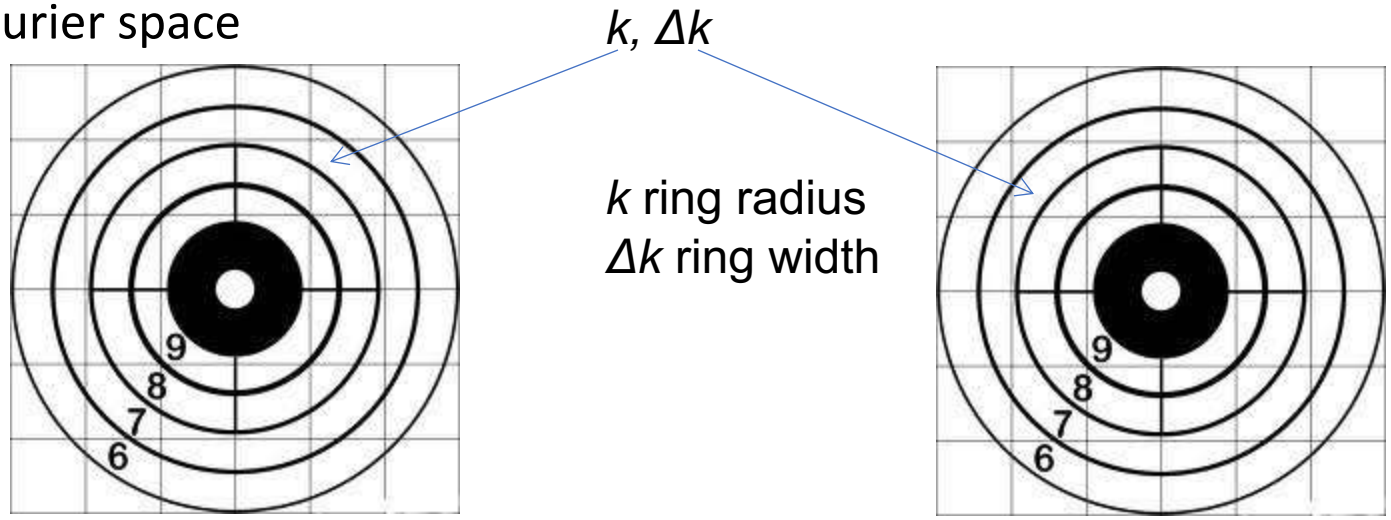


Maps	Annotation of States: (r70S: rotated) (nr70: non-rotated)	Resolution(Å)	Number of particles
1	nr70S-PtRNA-EtRNA	4.00	~50K
2	nr70S-PtRNA-EtRNA-EFG	3.61	~90K
3	r70S-P/EtRNA-EFG	4.22	~35K
4	r70S-P/EtRNA	5.70	~15K

“Story in a Sample”

Resolution definition -- determination in Fourier space

- Resolution is a reciprocal quantity, measured in Fourier space
- Defined as the spatial frequency [1/Å] up to which information is reproducible, by some measure of reproducibility
- Decomposition of information, by Fourier rings
- Randomly picked halfsets (e.g., odd- vs. even-numbered images)
- Compare averages [reconstructions] from halfsets over rings (shells) in Fourier space



Fourier ring/shell correlation

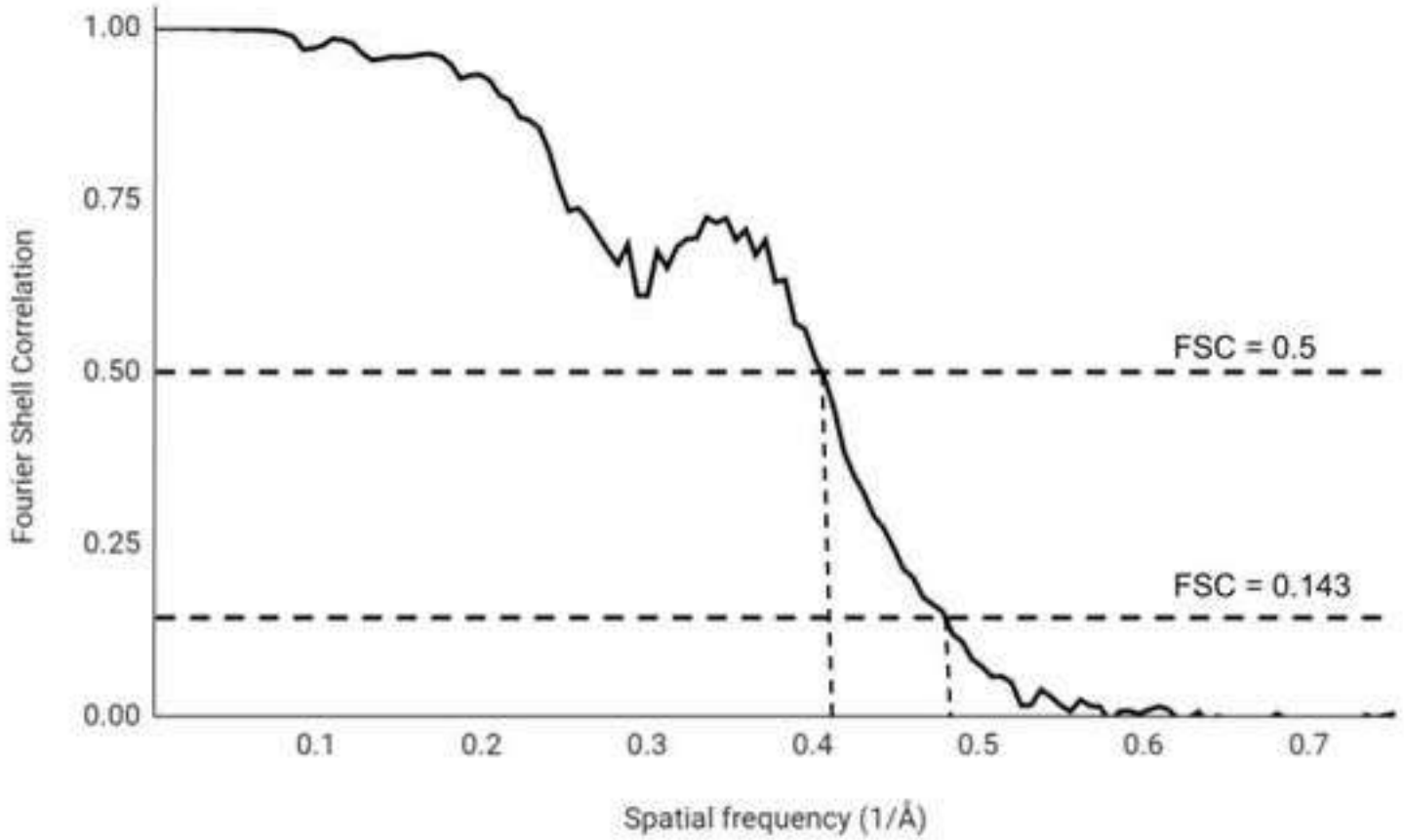
$$FSC_{1,2}(r) = \frac{\sum_{r_i \in r} F_1(r_i) \cdot F_2(r_i)^*}{\sqrt{\sum_{r_i \in r} |F_1(r_i)|^2 \cdot \sum_{r_i \in r} |F_2(r_i)|^2}}$$

r *corresponding shell in Fourier space*

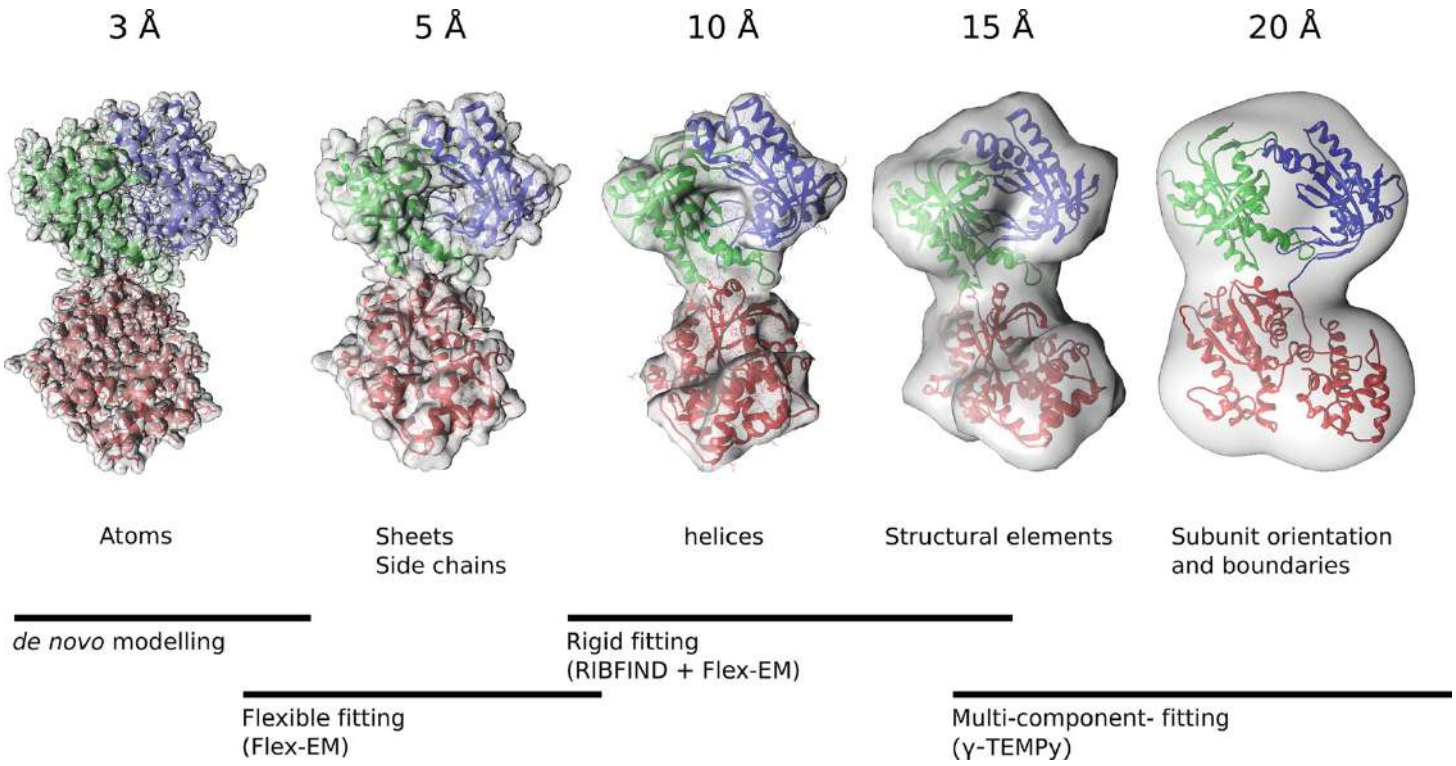
$F(r_i)$ *complex "structure factor" at* r_i *in Fourier space*

$\sum_{r_i \in r}$ *summation over all Fourier space voxels* r_i *in shell* r

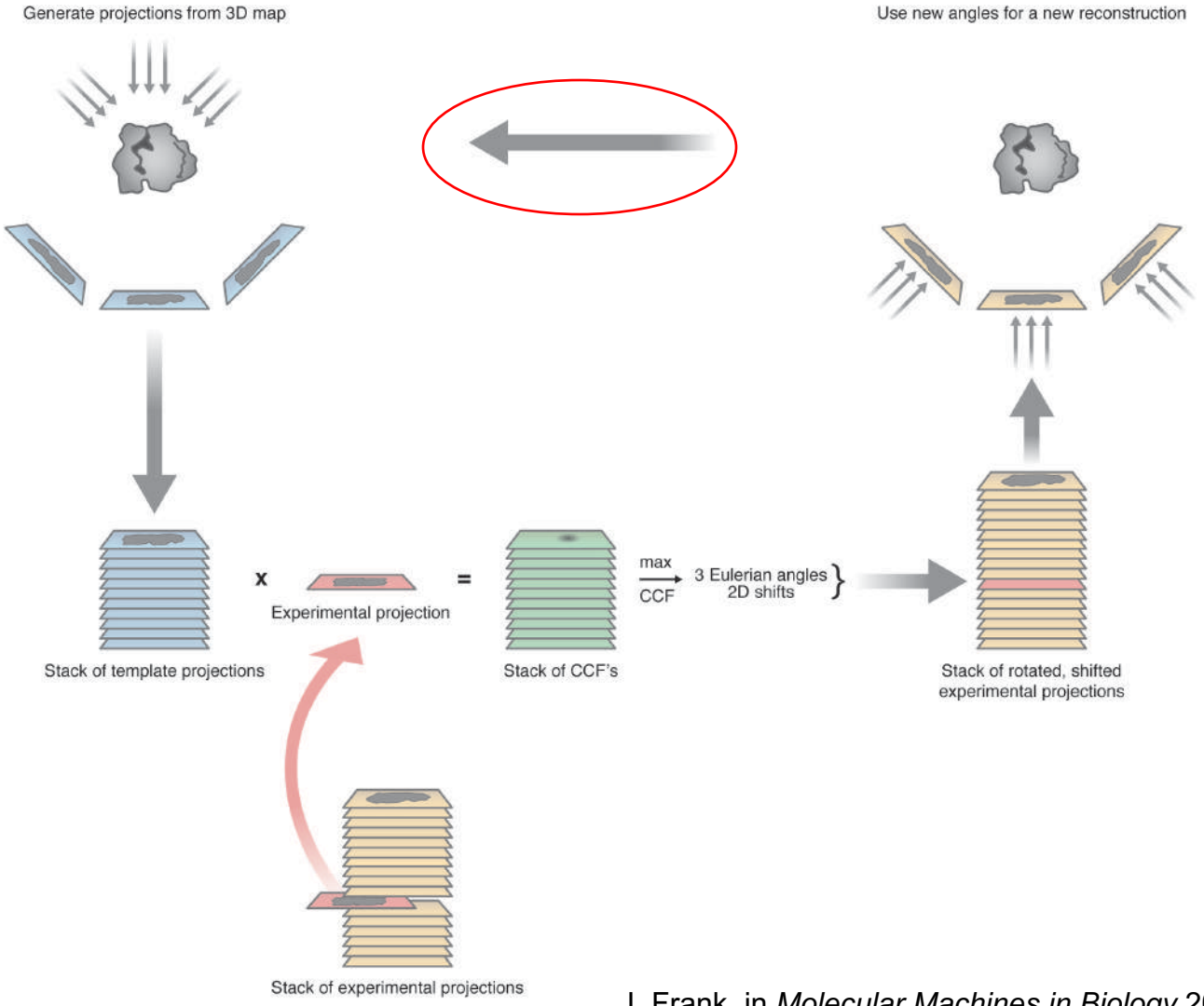
Resolution via Fourier shell correlation



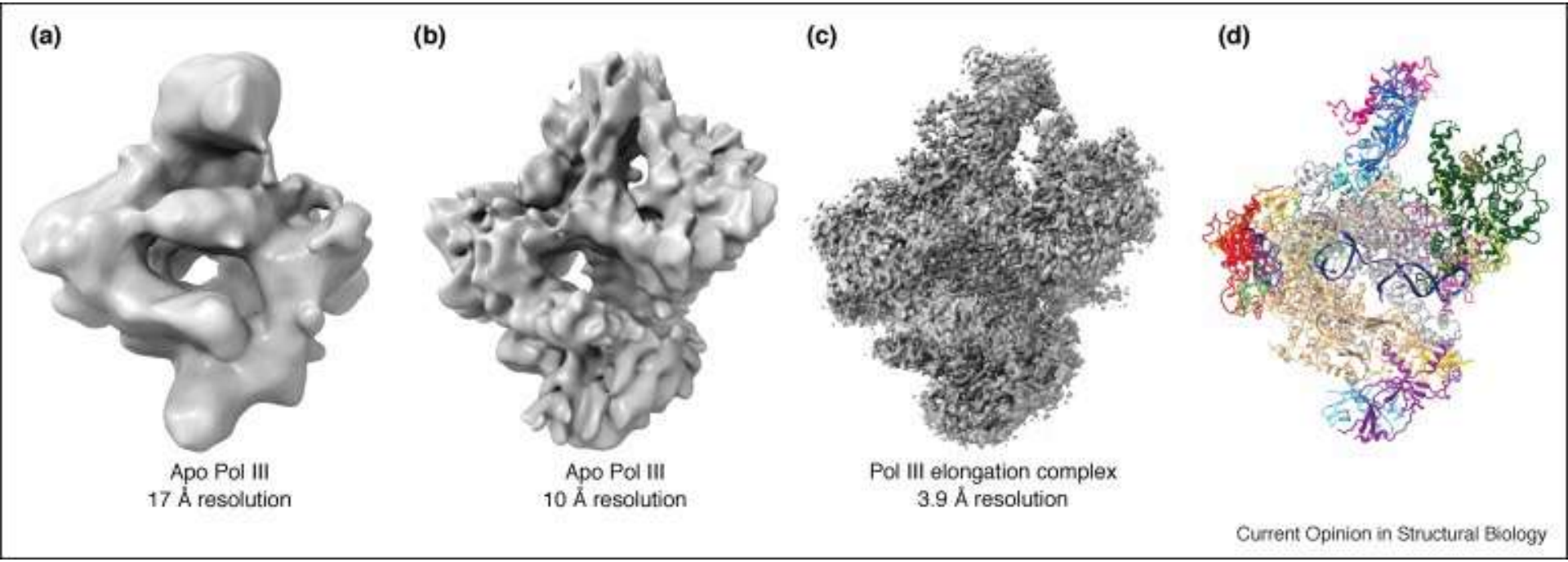
Resolution pragmatically defined through visibility of structural elements

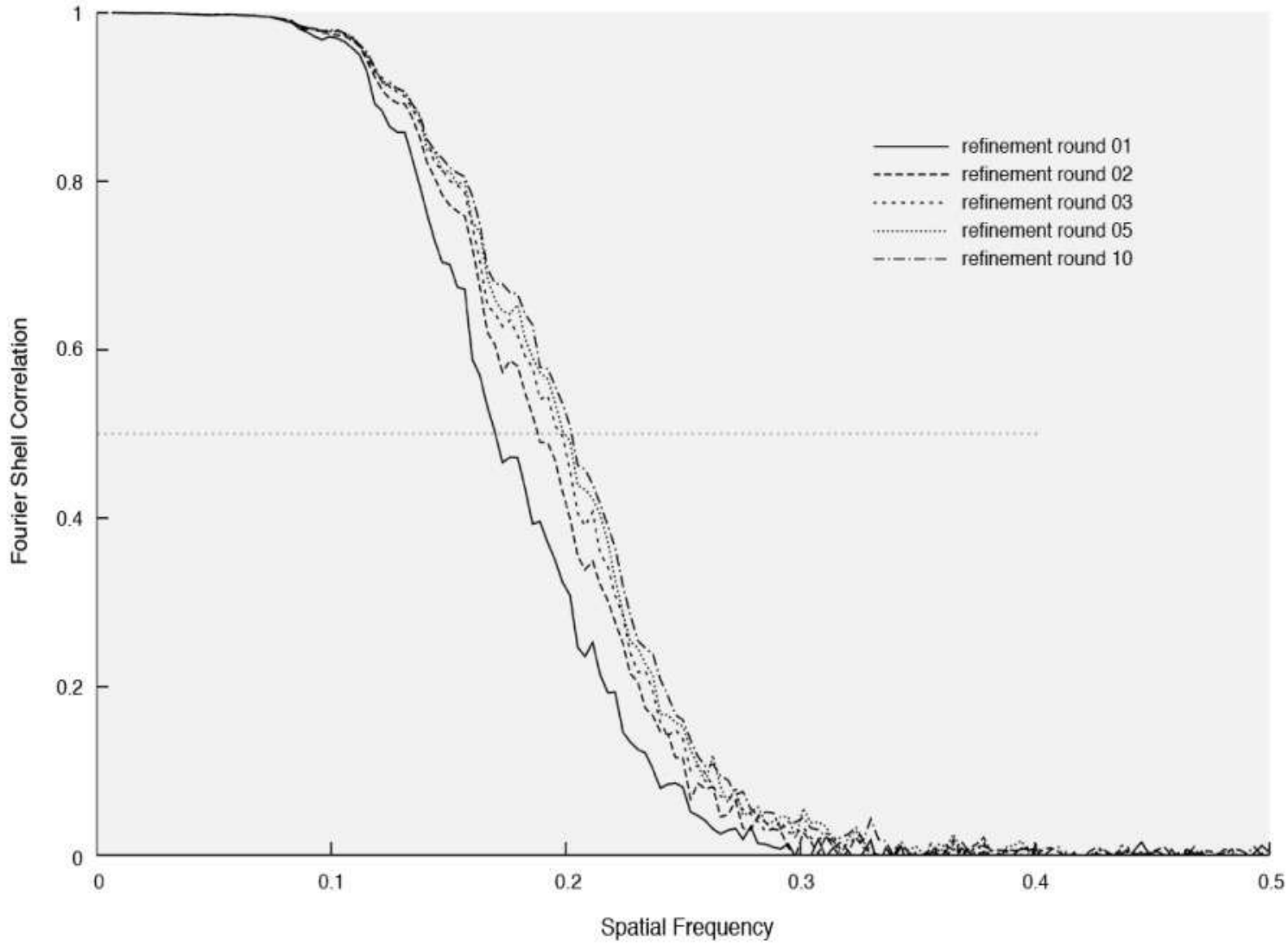


Iterative Angular Refinement

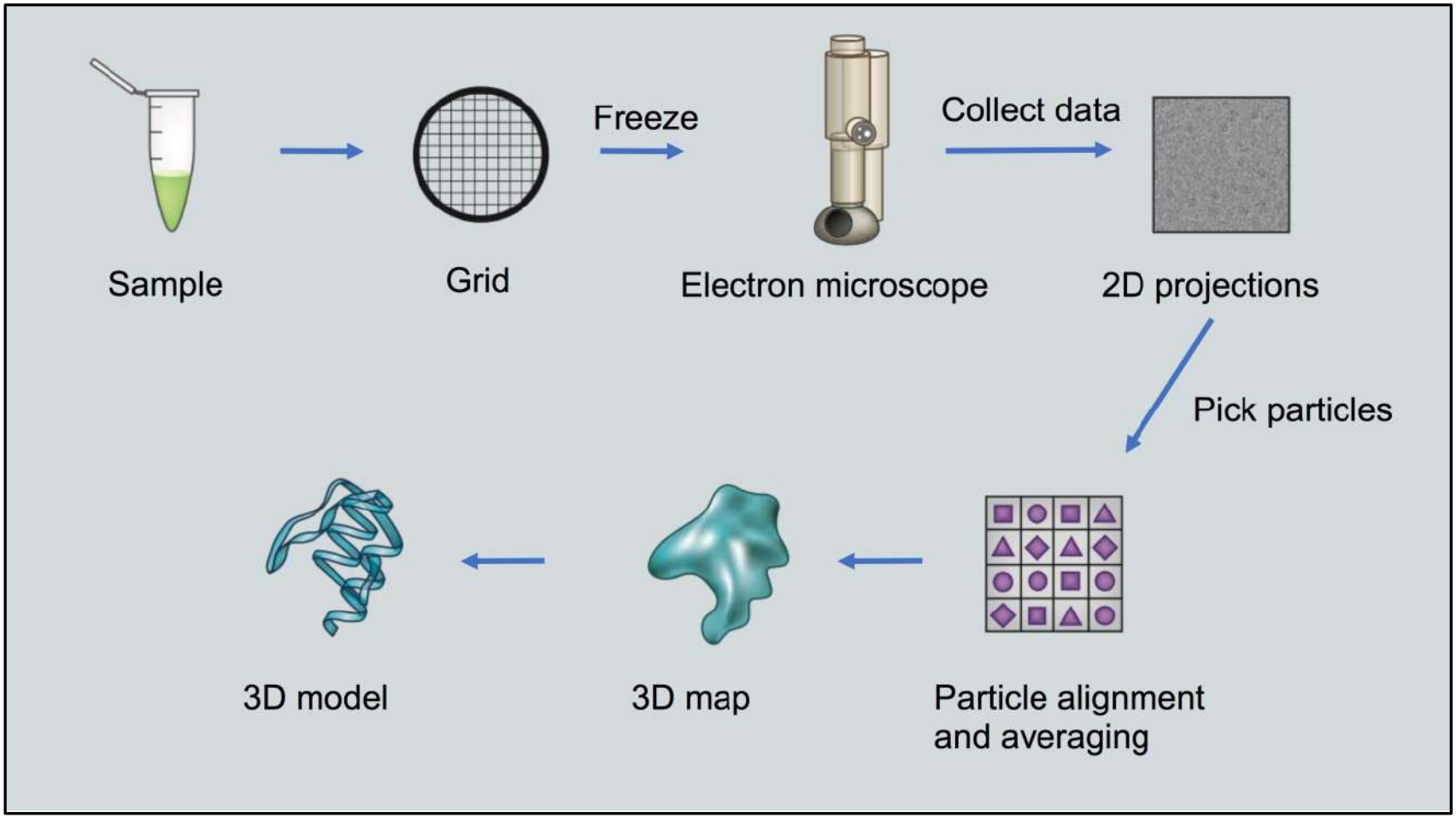


J. Frank, in *Molecular Machines in Biology* 2011


(a)
(b)
(c)
(d)
Apo Pol III
17 Å resolution
Apo Pol III
10 Å resolution
Pol III elongation complex
3.9 Å resolution
Current Opinion in Structural Biology

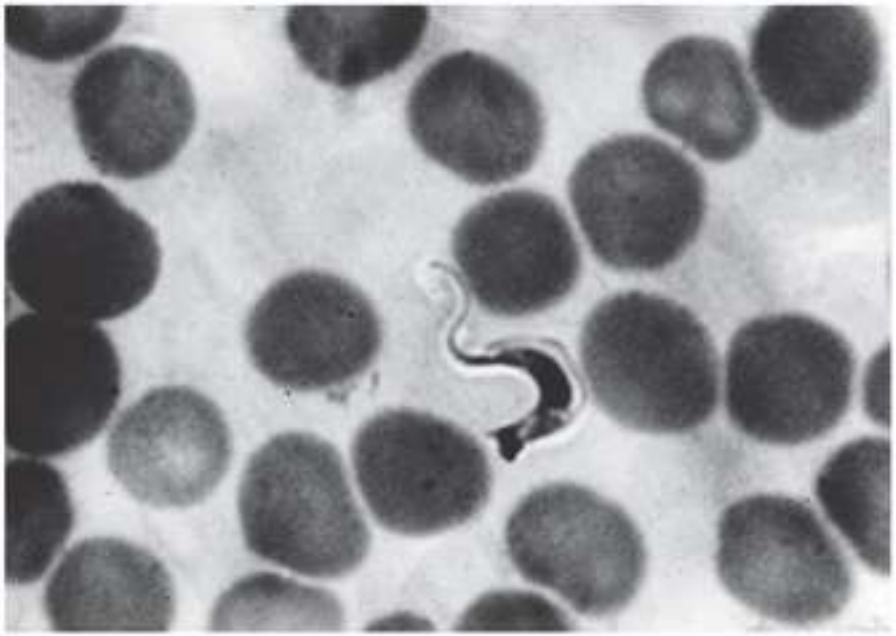


FSC following progress of refinement



Creative Biostructure
Jan 24, 2018

Case study: *Trypasonoma cruzi* ribosome



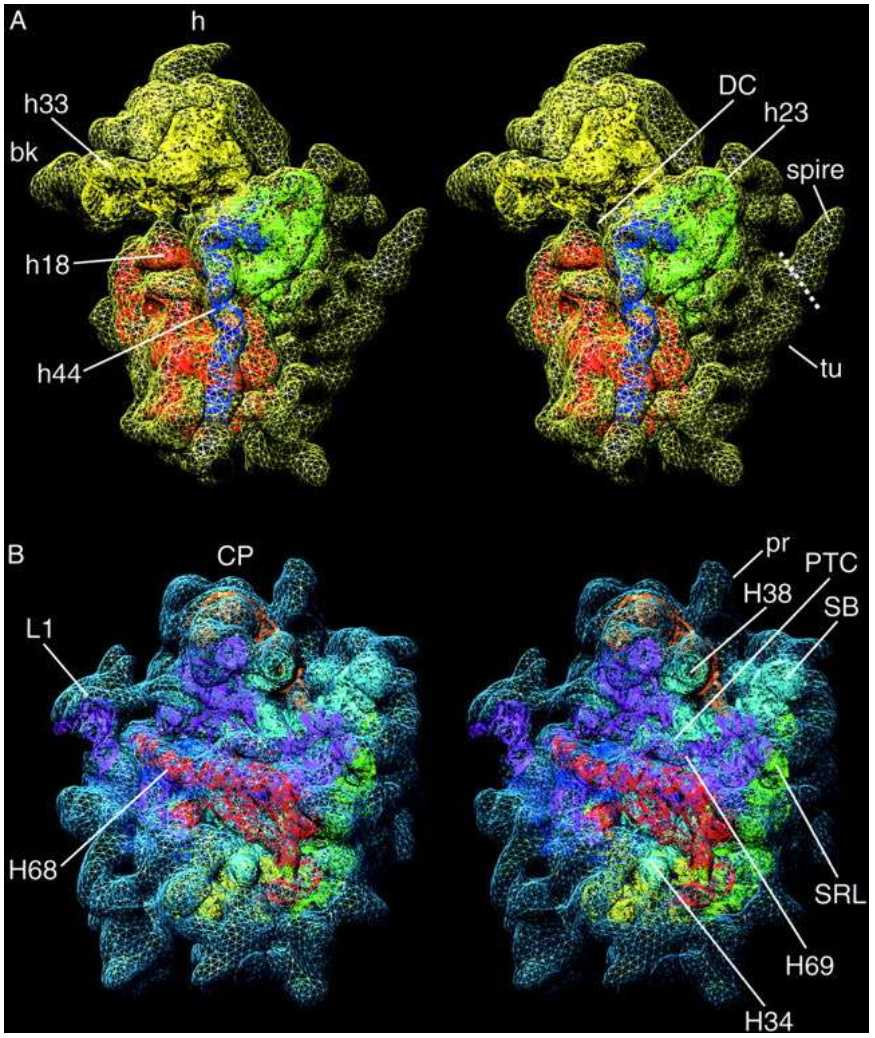
Mariano Levin

FIGURE 44.3 *Trypanosoma cruzi* parasite in human blood smear. Giemsa stain, magnification ×625. (Courtesy Dr. Maria Shikanai Yasuda, São Paulo, Brazil.)

Chagas disease

Trypanosoma cruzi

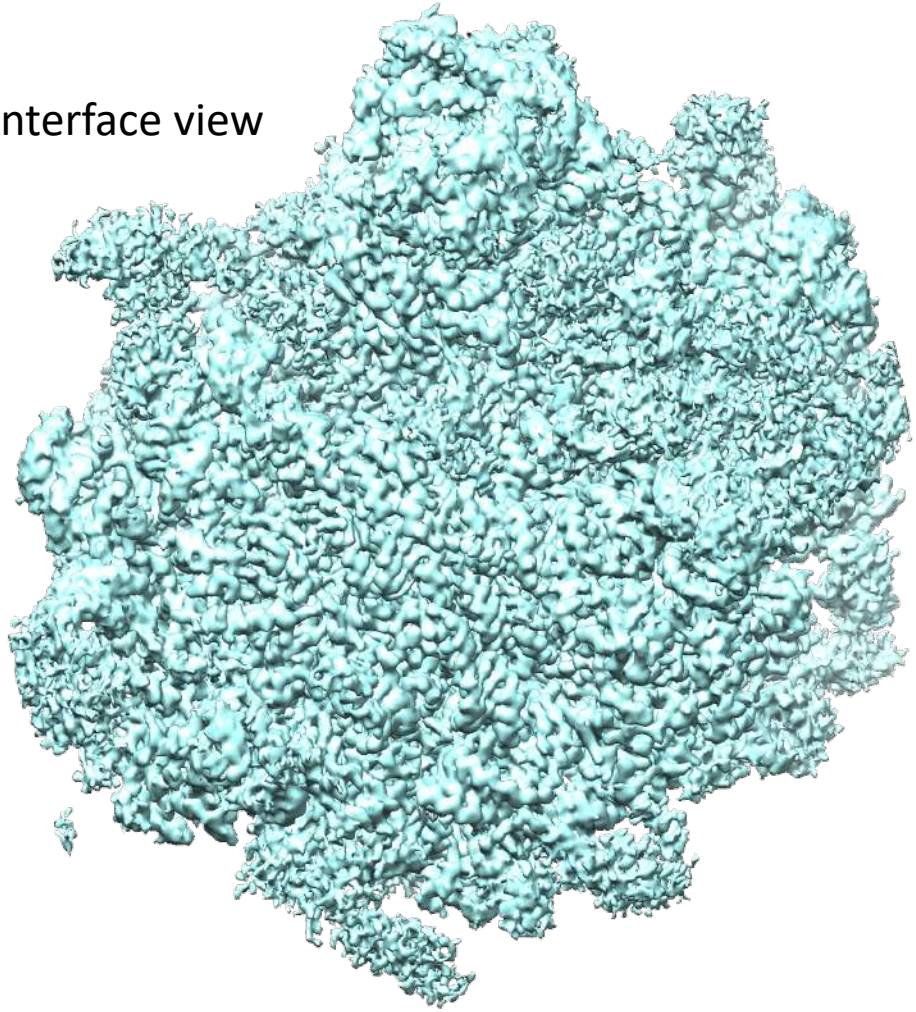
- Parasite causing Chagas disease
- Widespread in South America
- Transmitted by a bug
- Cryo-EM reconstruction in 2005 to 12Å resolution (with Mariano Levin) recorded on film
- Unusually large rRNA expansion segments

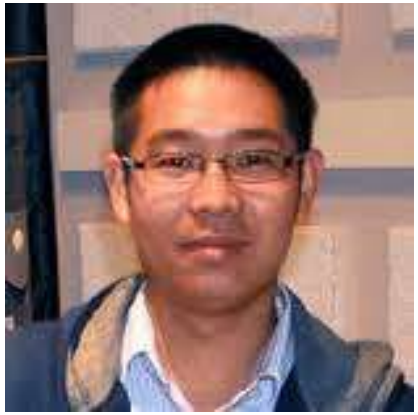


Haixiao Gao et al. PNAS 2005;102:10206-10211

Trypanosome cruzi ribosome (large subunit) at 2.5 Å average resolution from ~160,000 particles

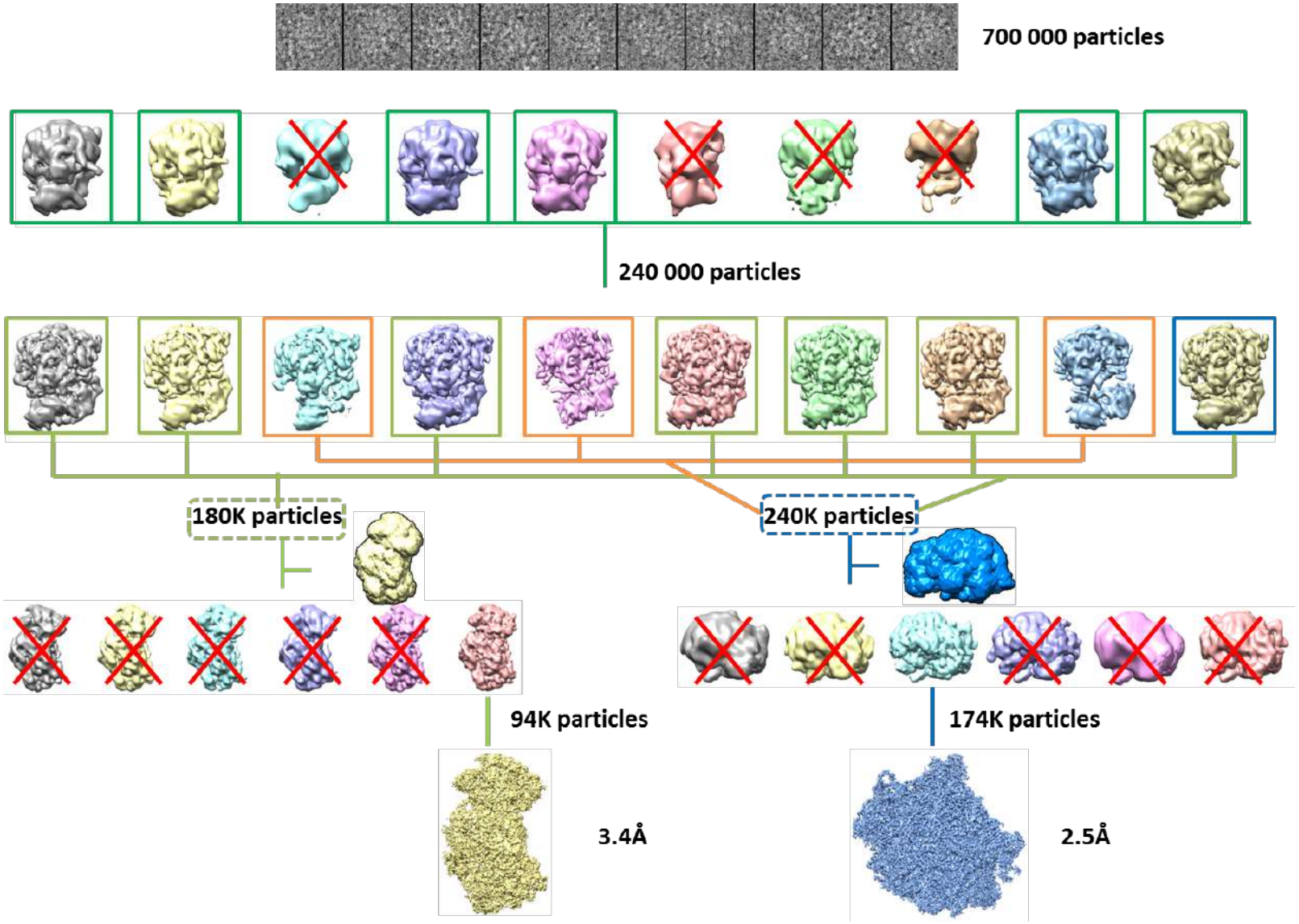
60S interface view

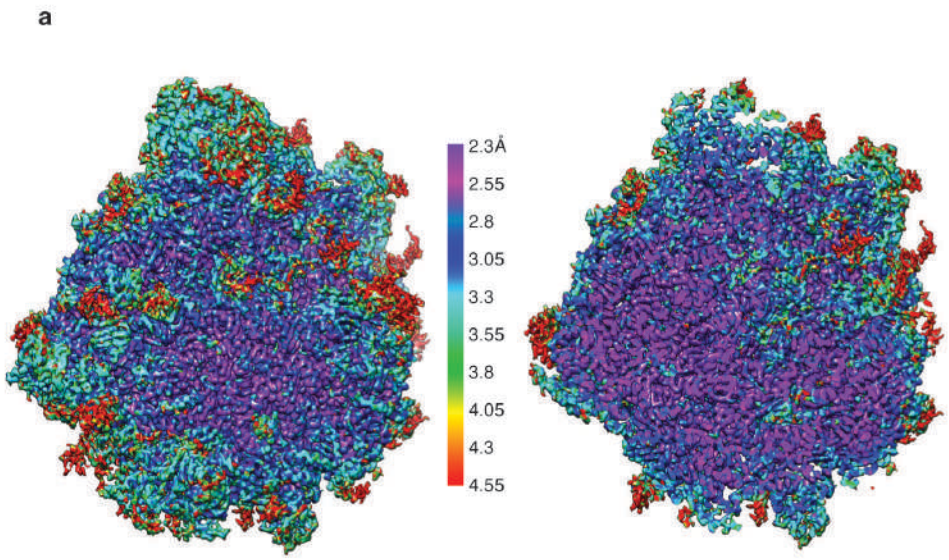


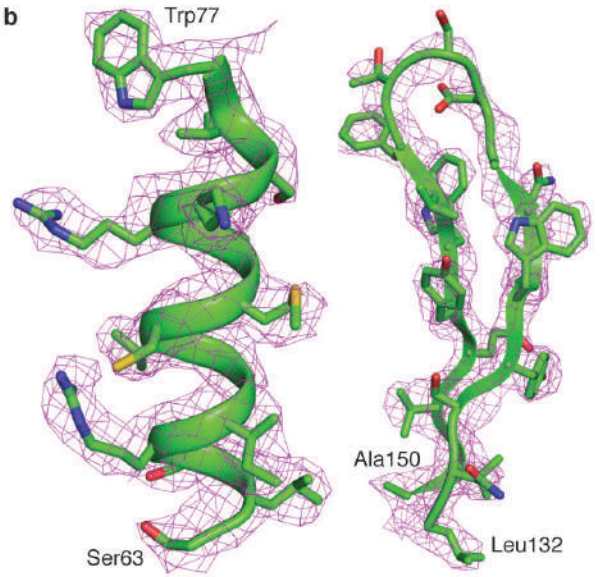


Zheng Liu

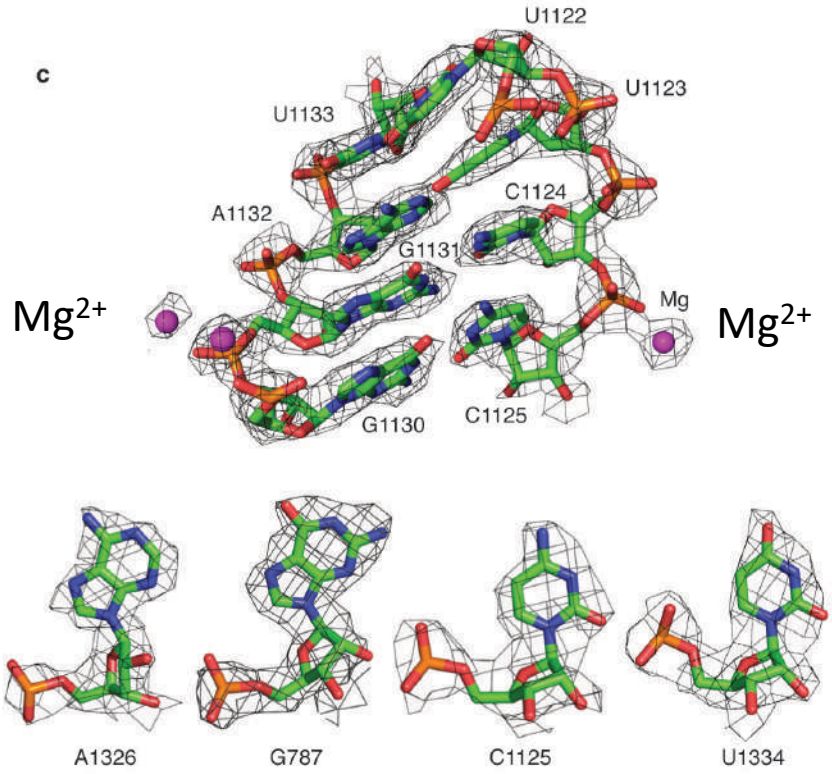
Liu et al., PNAS 2016


700 000 particles
240 000 particles
180K particles
240K particles
94K particles
174K particles
3.4Å
2.5Å

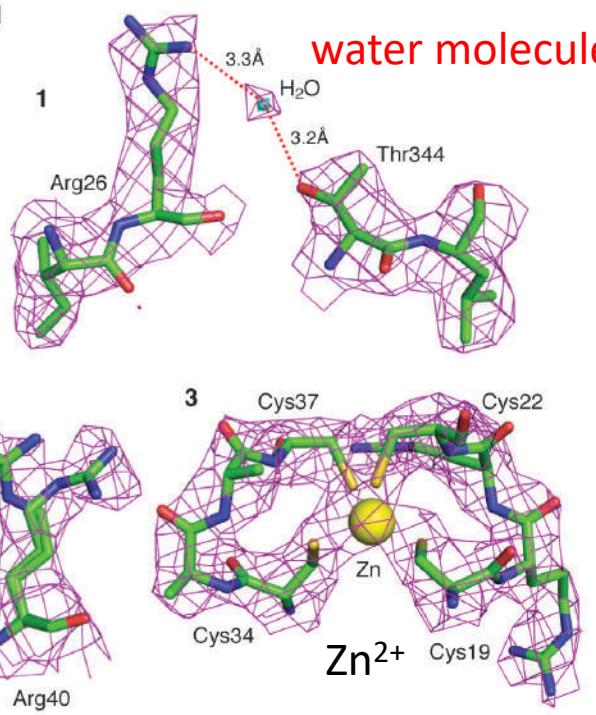




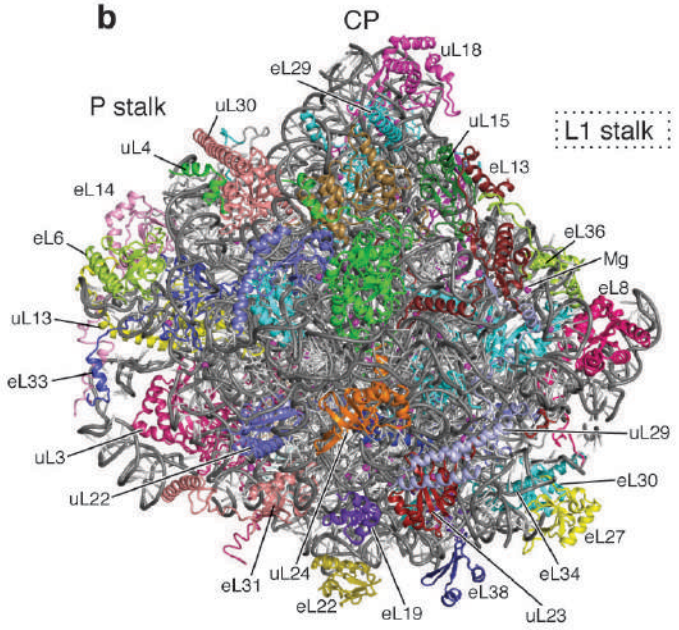
α-helix 2 strands of β-sheet



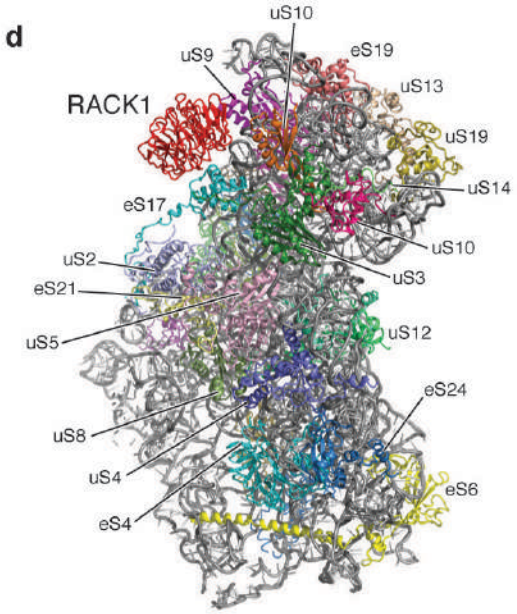
Mg^{2+} Mg^{2+}



Zn^{2+}



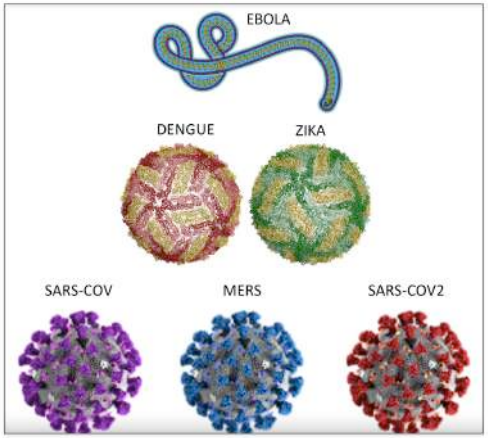
60S

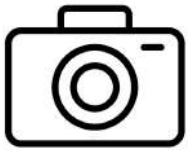


40S

Liu et al., PNAS 2016

Just in Time: high-resolution single-particle cryo-EM and the new pandemics





2012

2012 – 2015 — MERS

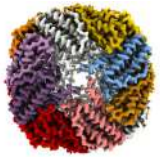
2015 – 2016 — Zika

2014 – 2016 — Ebola

2019 – 2020 — Dengue

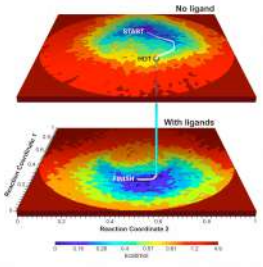
2020 - ???? — SARS-Cov2

Future directions

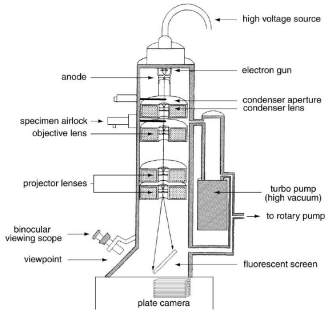


Higher spatial resolution

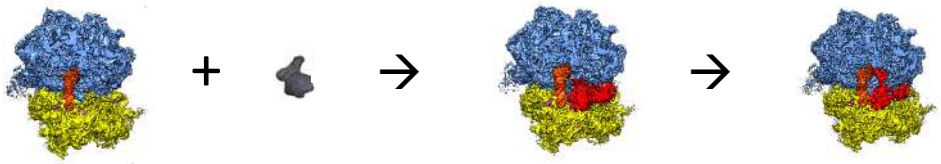
Time resolution
microfluidics



State resolution
machine learning
energy landscape



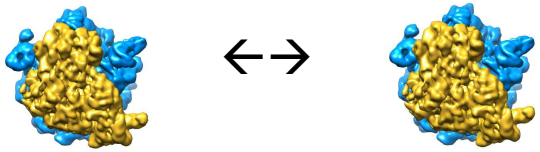
Capturing Motion by Single-Particle Cryo-EM -- two ways:



TIME RESOLUTION

Capture short-lived states in a reaction of two or more components of a machine on the way toward thermal equilibrium

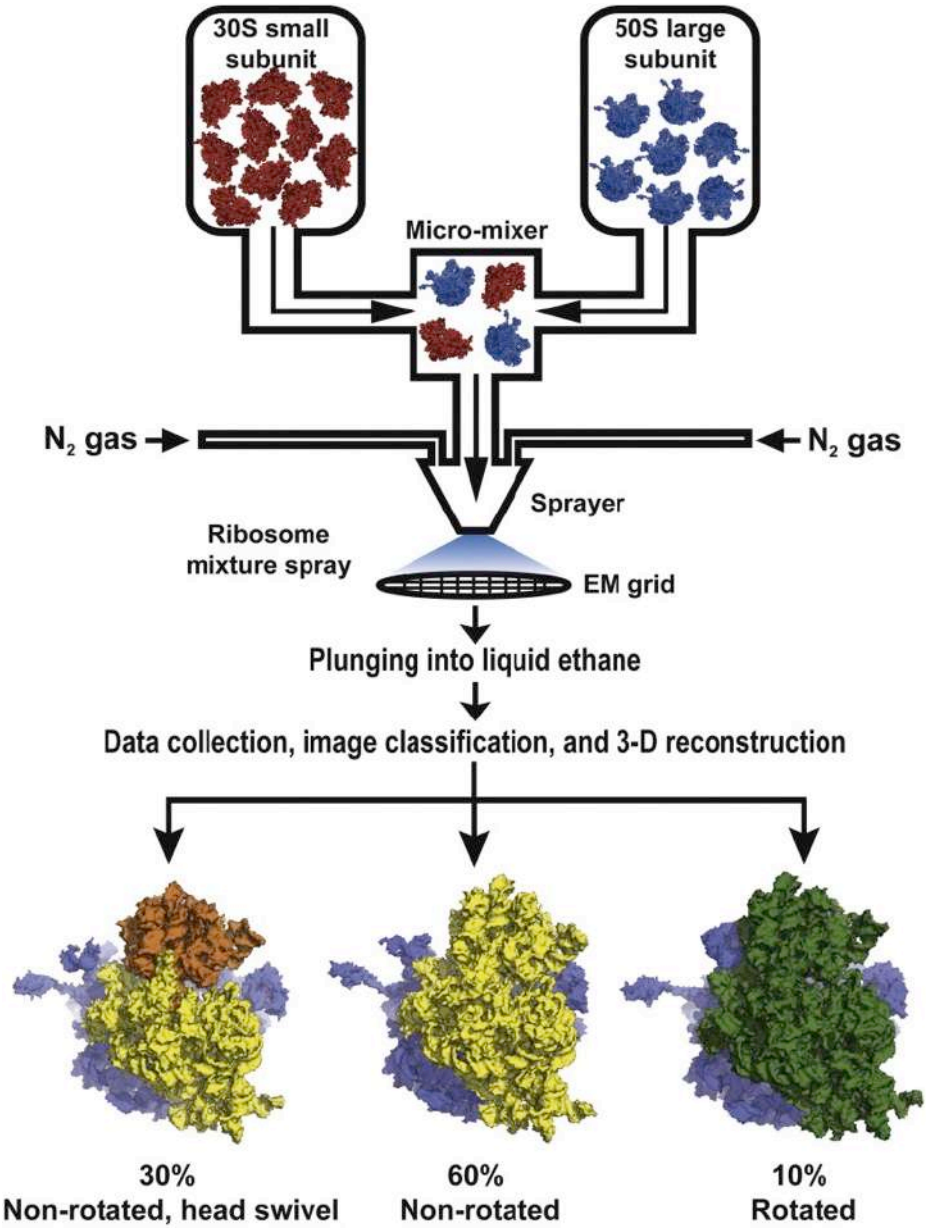
Time-resolved cryo-EM



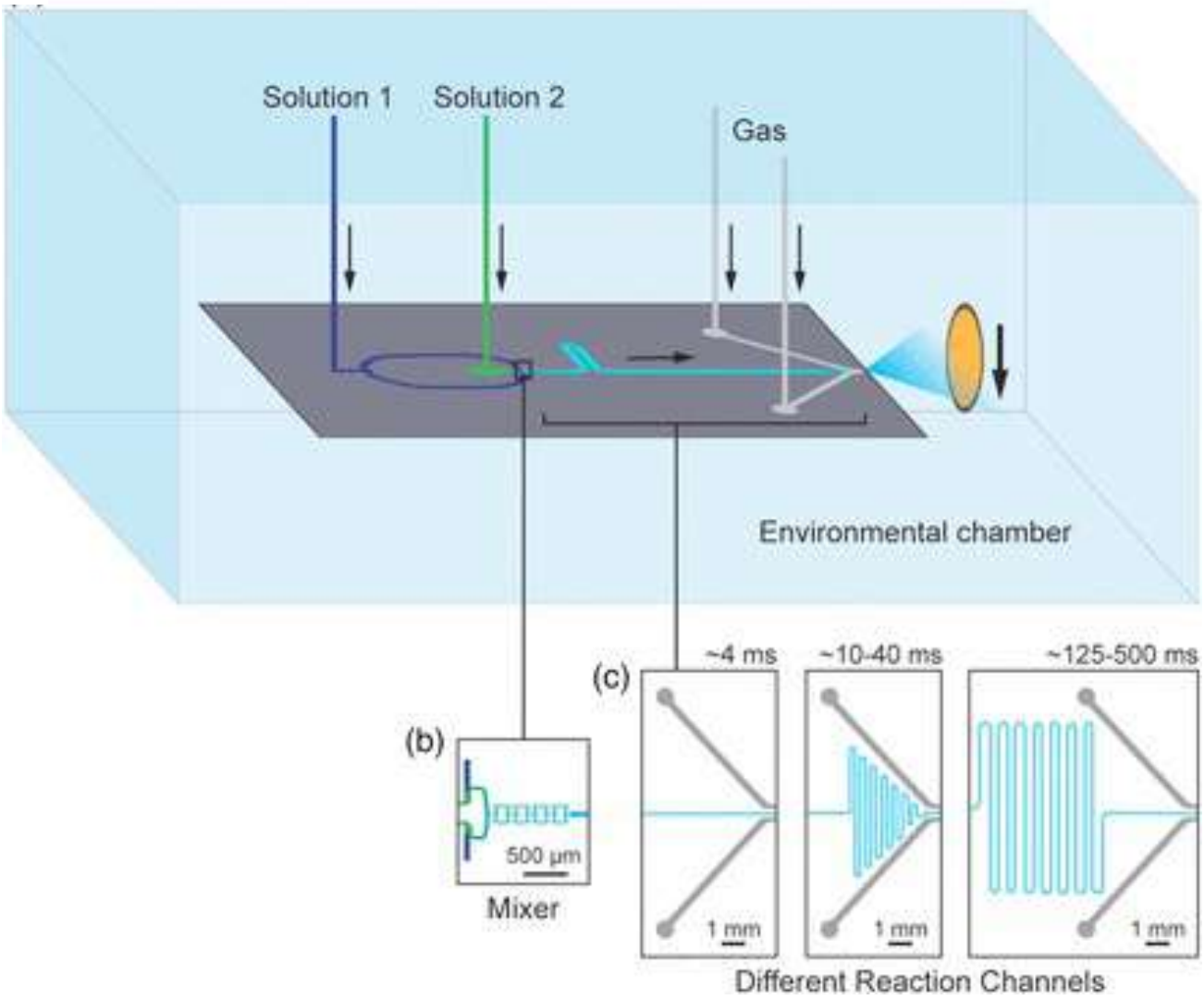
STATE RESOLUTION

Collect many (in the 100,000s) snapshots of molecules or molecular machines exhibiting conformational changes in thermal equilibrium

Standard cryo-EM
+ deep data analysis



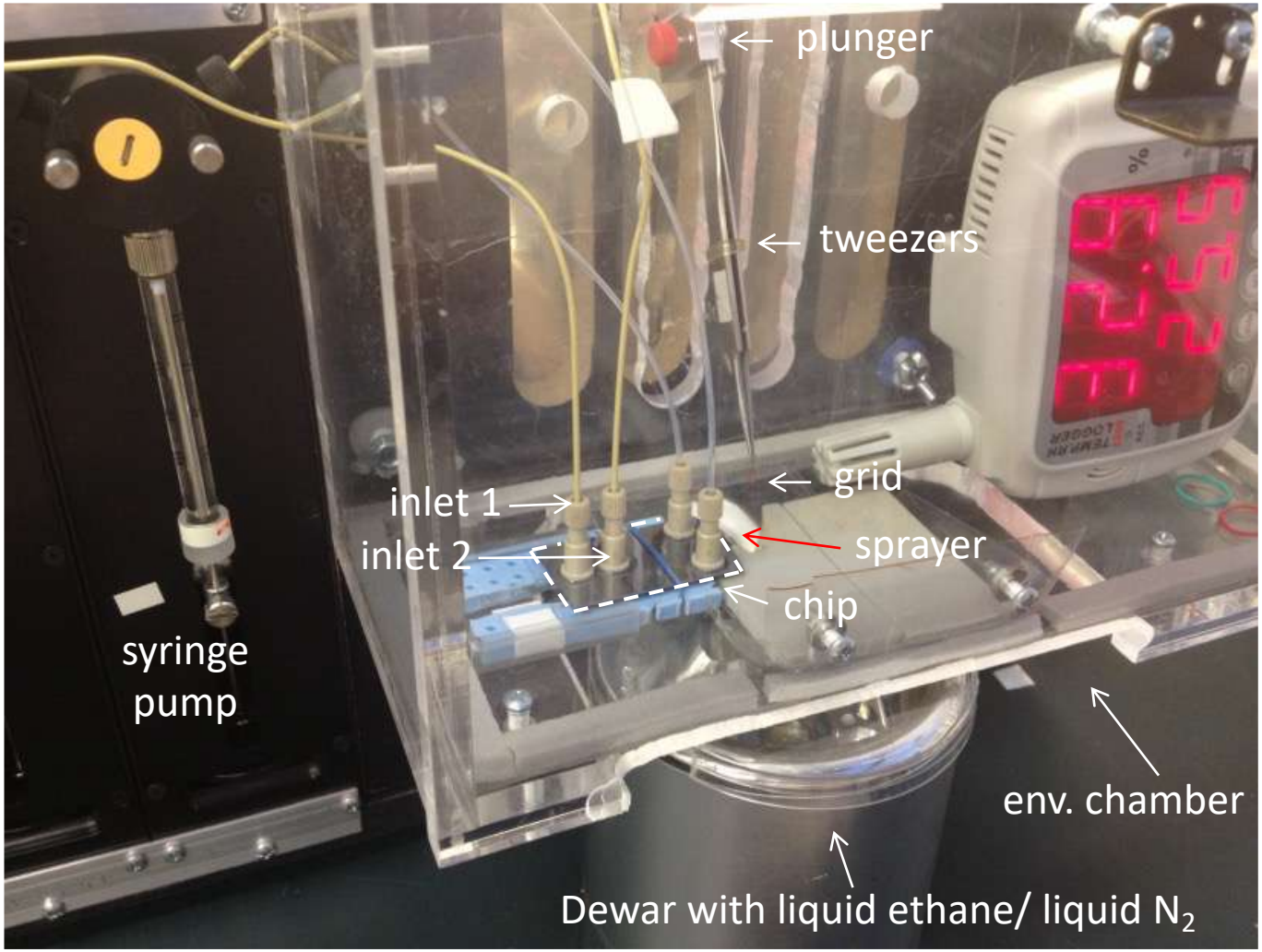
Commentary to Chen et al. , Structure 2015
by
Lin and Ermolenko, Structure 2015



Lu et al., JSB 2009

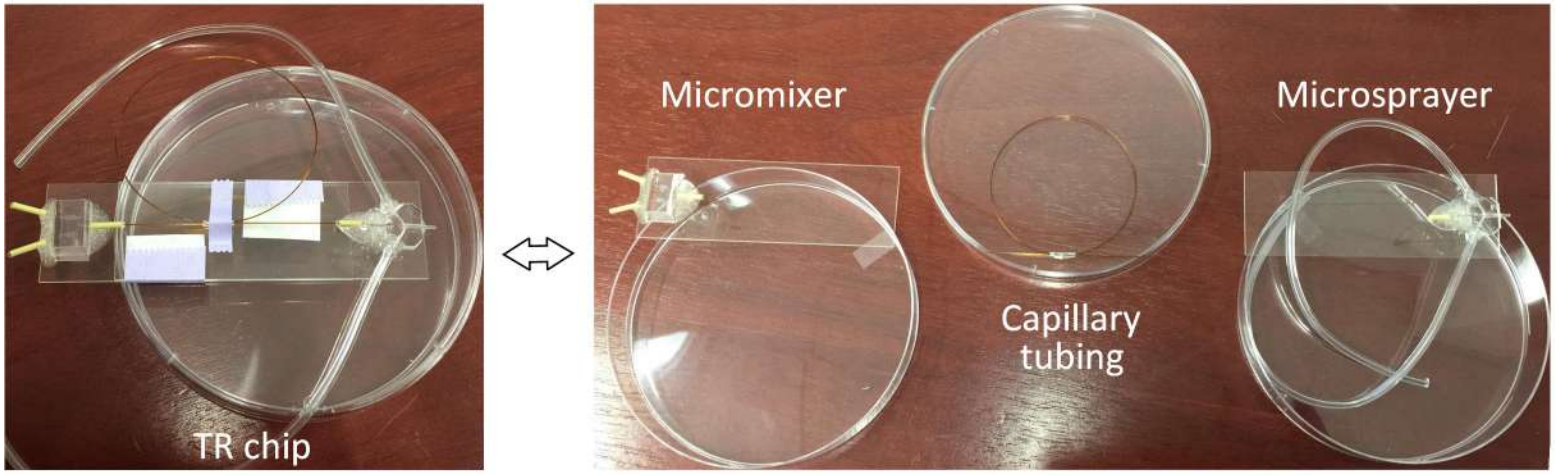
Image: Chen and Frank, Microscopy 2015

Apparatus for TR cryo-EM at Columbia



Modular PDMS-based microfluidic chip

(1) Easy and cheap to fabricate; (2) more efficient 3D mixer; (3) modular design – reuse of components; (4) length of glass capillary defines reaction time.



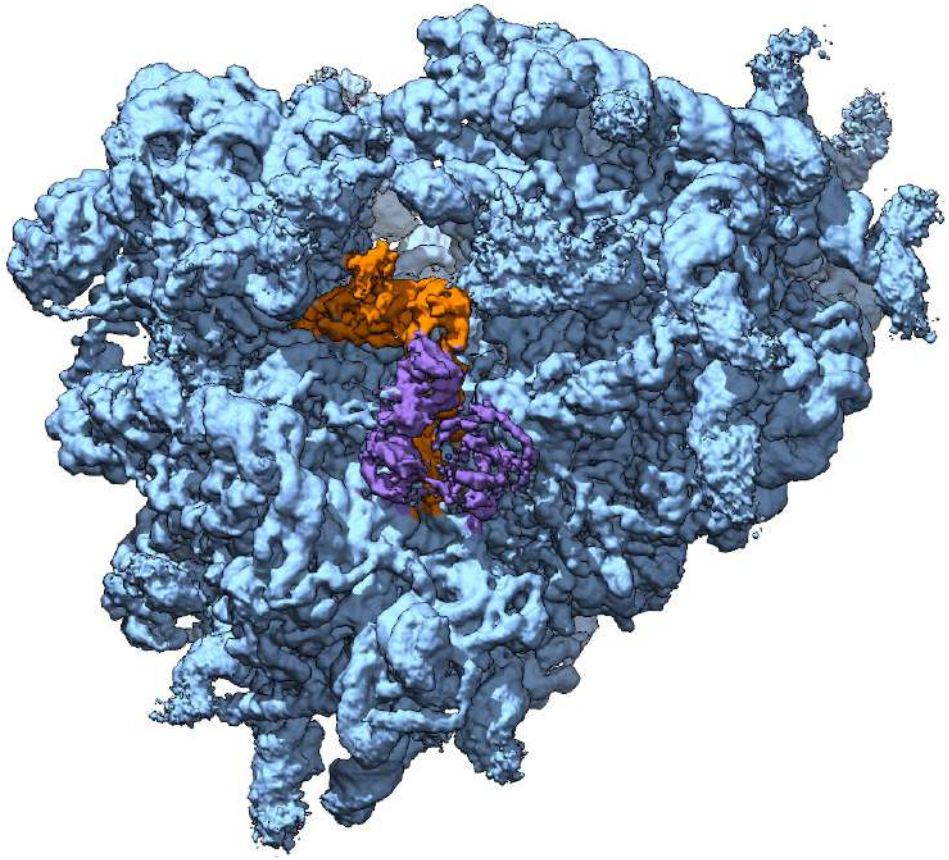
assembled

disassembled

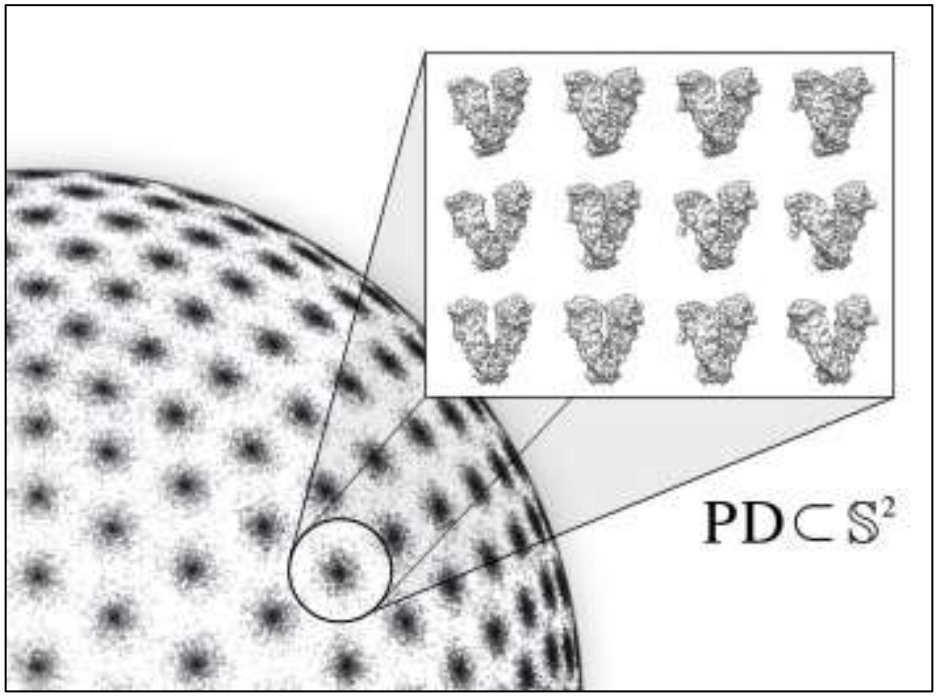
Xiangson Feng
Frank Lab

Qiao Lin
Dept Mechanical Engineering

Translation termination in mammalian ribosome: eRF1-eRF3 on the 80S ribosome captured prior to GTP hydrolysis

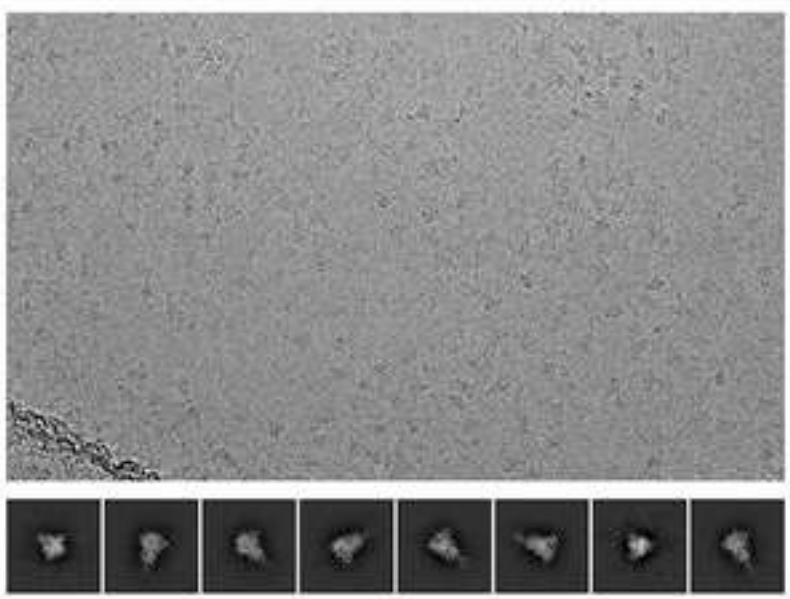


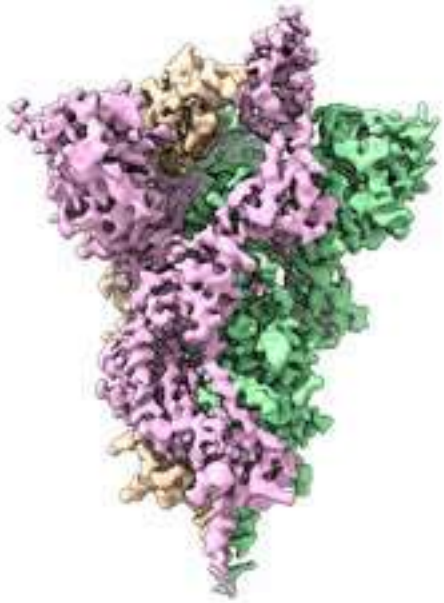
ManifoldEM -- What is the principle?



Angular sphere S^2 and its projection directions (PDs).
Initial analysis is for one PD at a time.

SARS CoV-2 virus spike protein

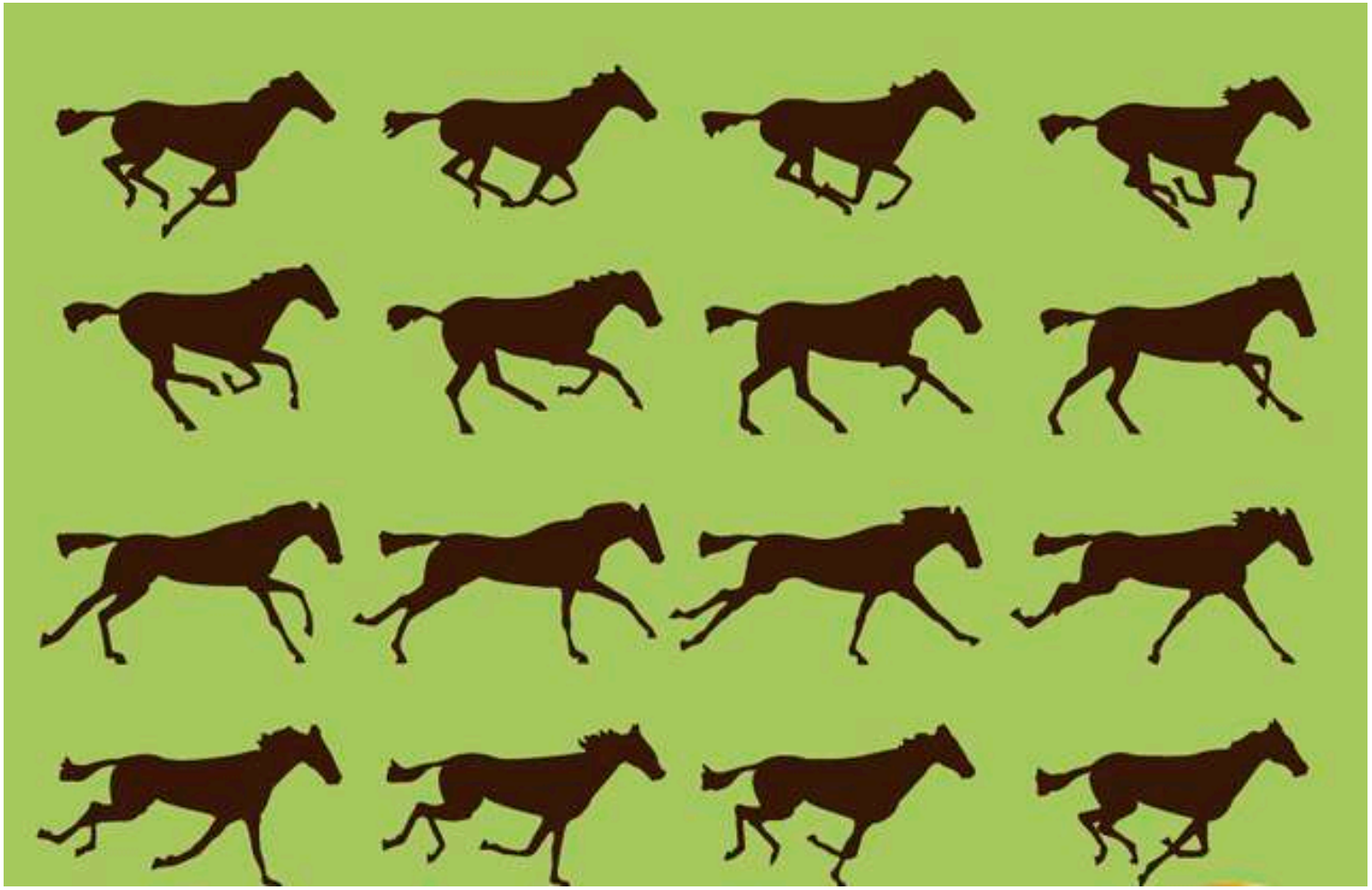




Wrapp et al., Science 2020


10,000 horses galloping

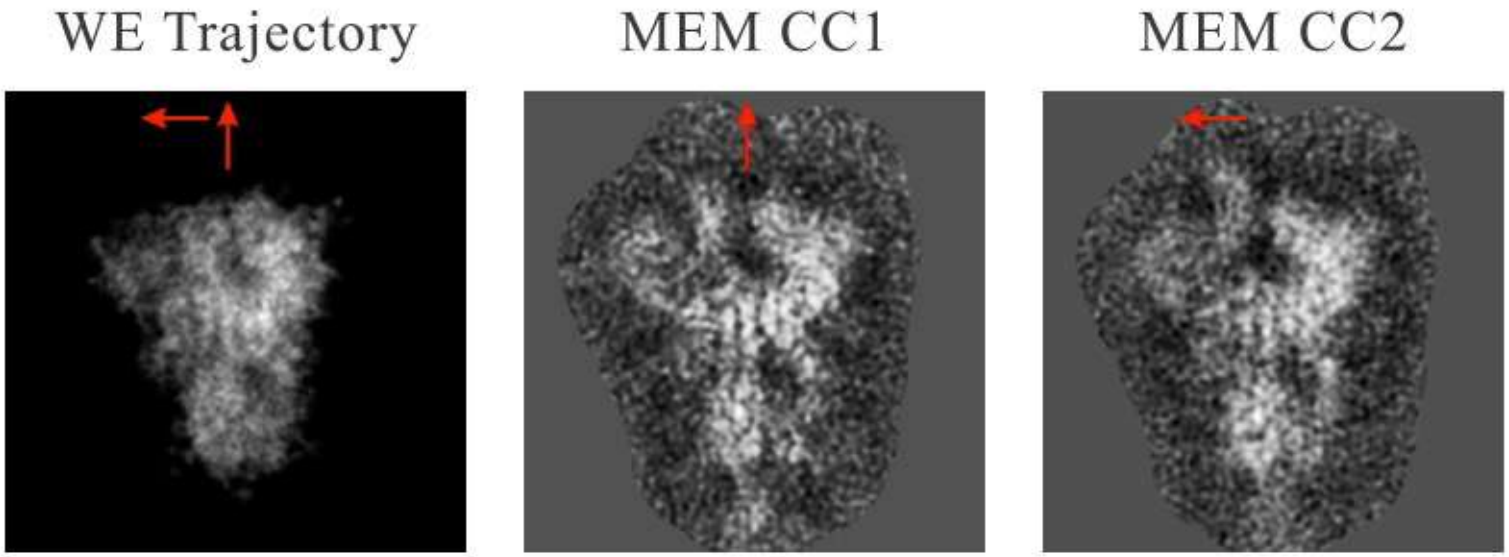
DO for each viewing direction (projection direction on the angular sphere):
order the images sequentially by similarity



(This is a one-dimensional conformational motion component: one complete cycle of gallop)

By splicing all images from different horses in the sequence of their similarity in forward and backward direction, we get a movie of the average horse galloping

Side view (PD 1386)

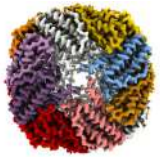


Playback: 0

WE = Weighted ensemble MD

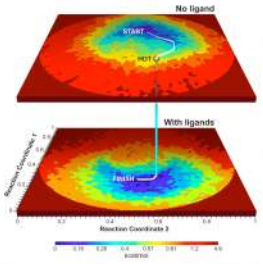
MEM = ManifoldEM

Future developments: there is more than better spatial resolution!



Higher spatial resolution

Time resolution
microfluidics



State resolution
machine learning
energy landscape