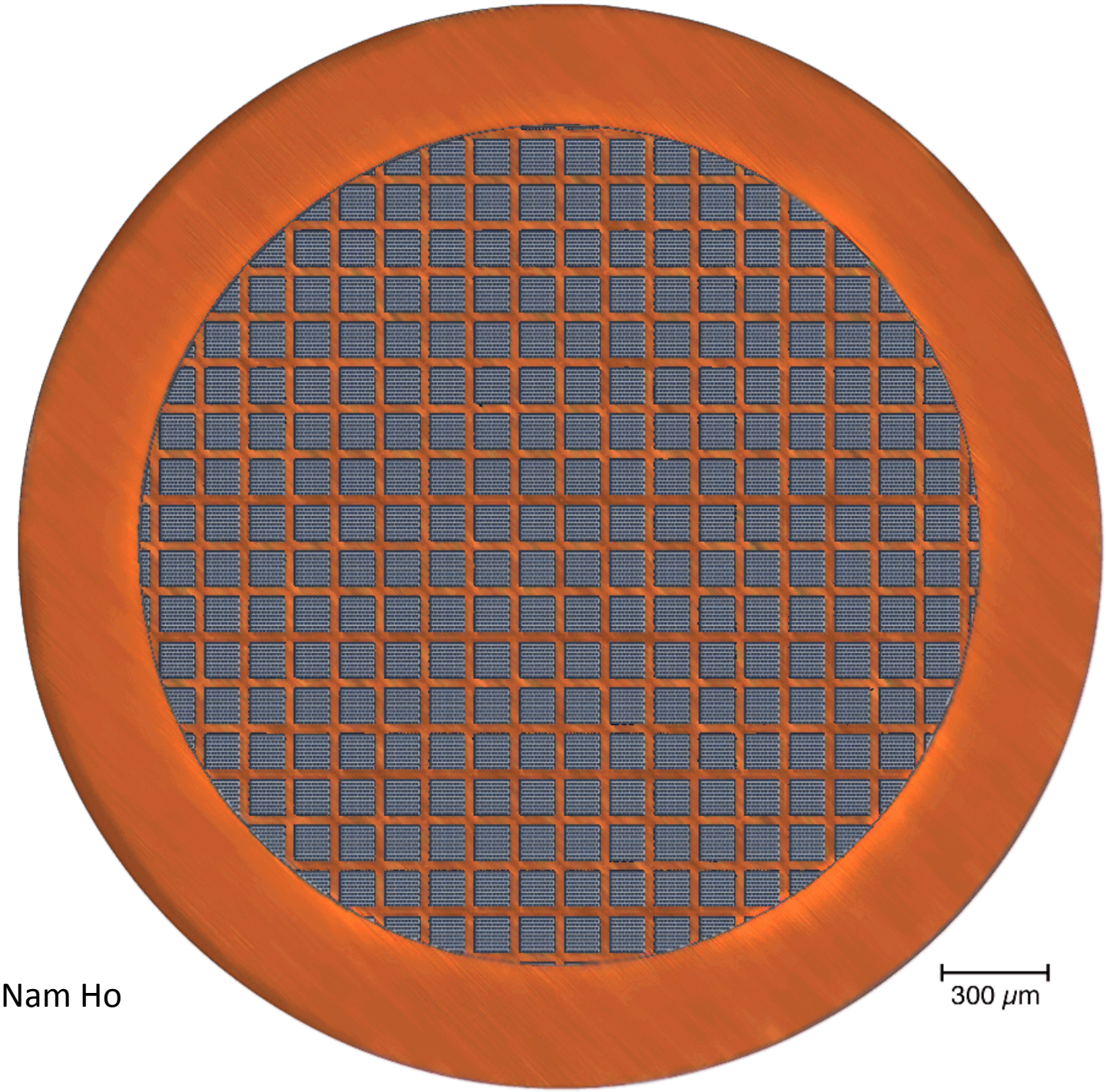
Single-particle Cryo-EM – the sky is the limit.

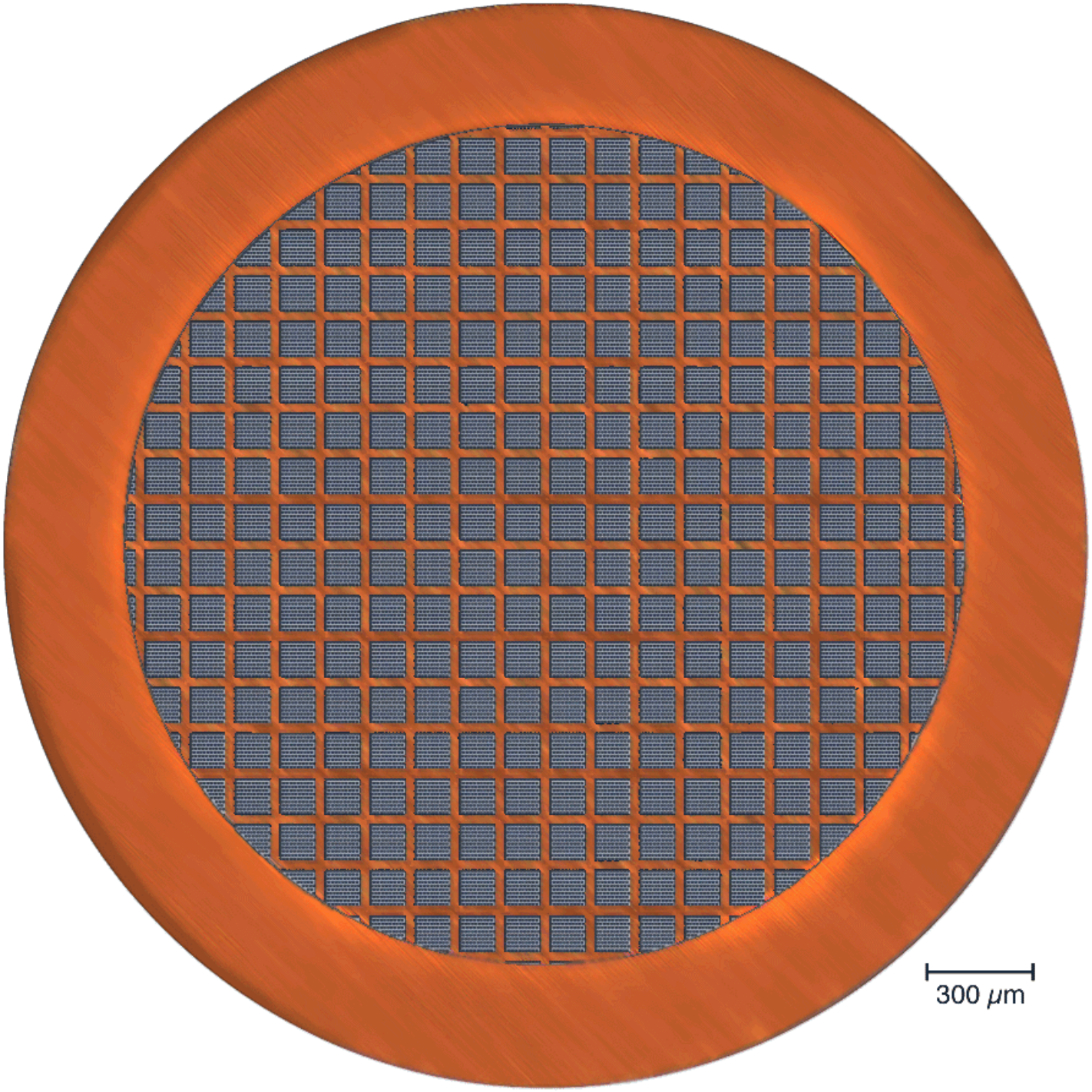
Joachim Frank

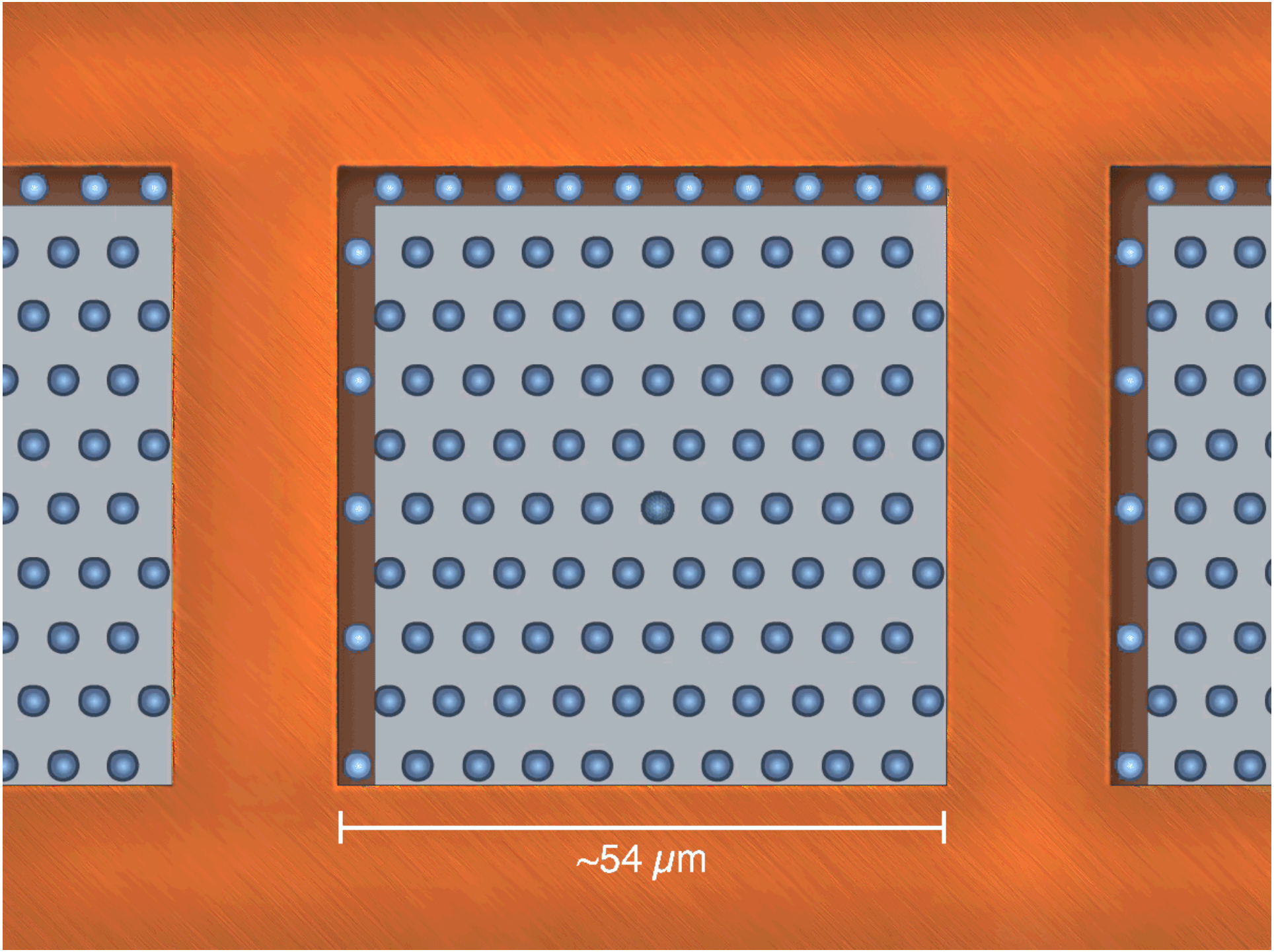
Howard Hughes Medical Institute
Department of Biochemistry and Molecular Biophysics
and Department of Biological Sciences
Columbia University

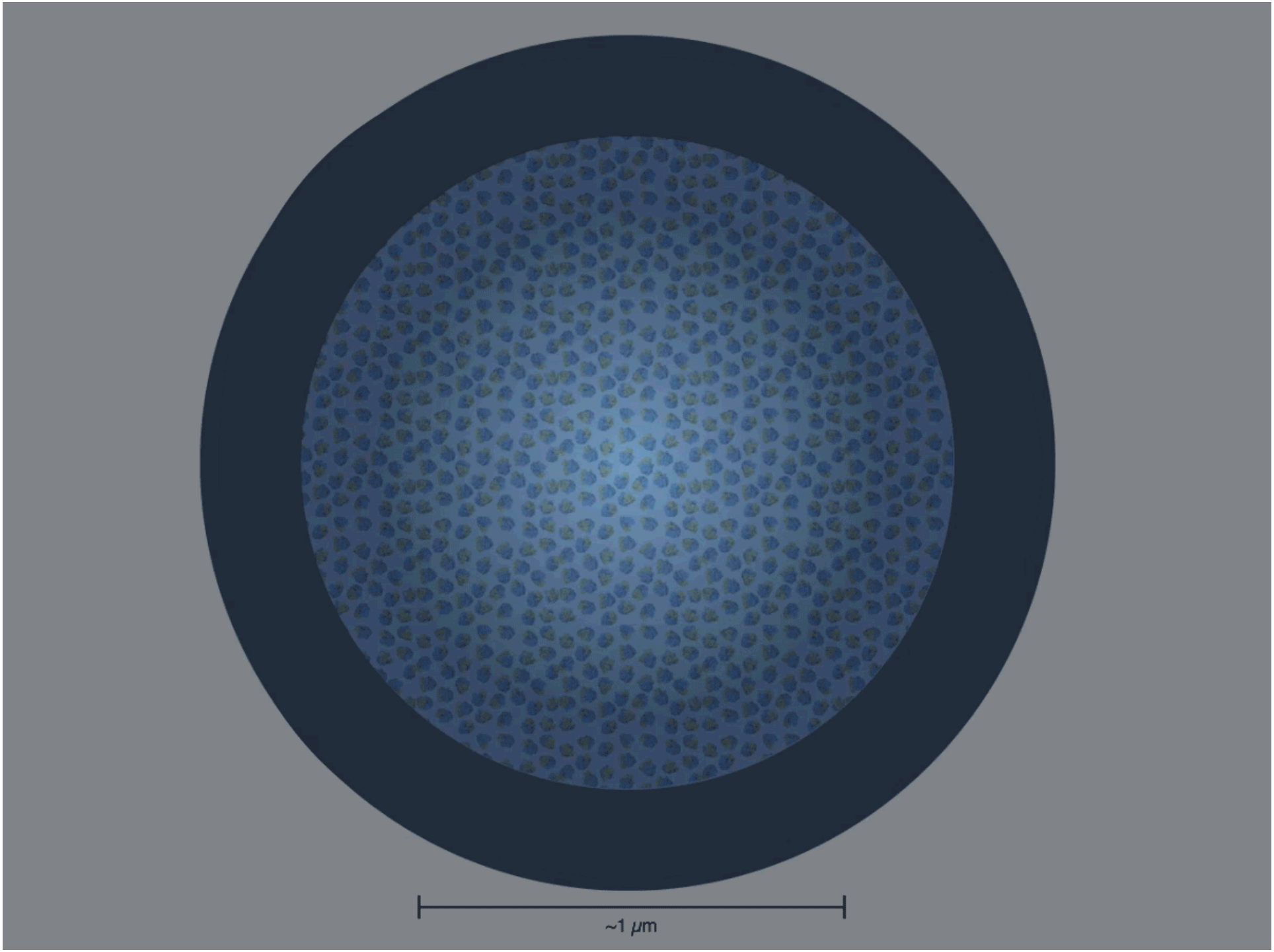
Funding: HHMI and NIH

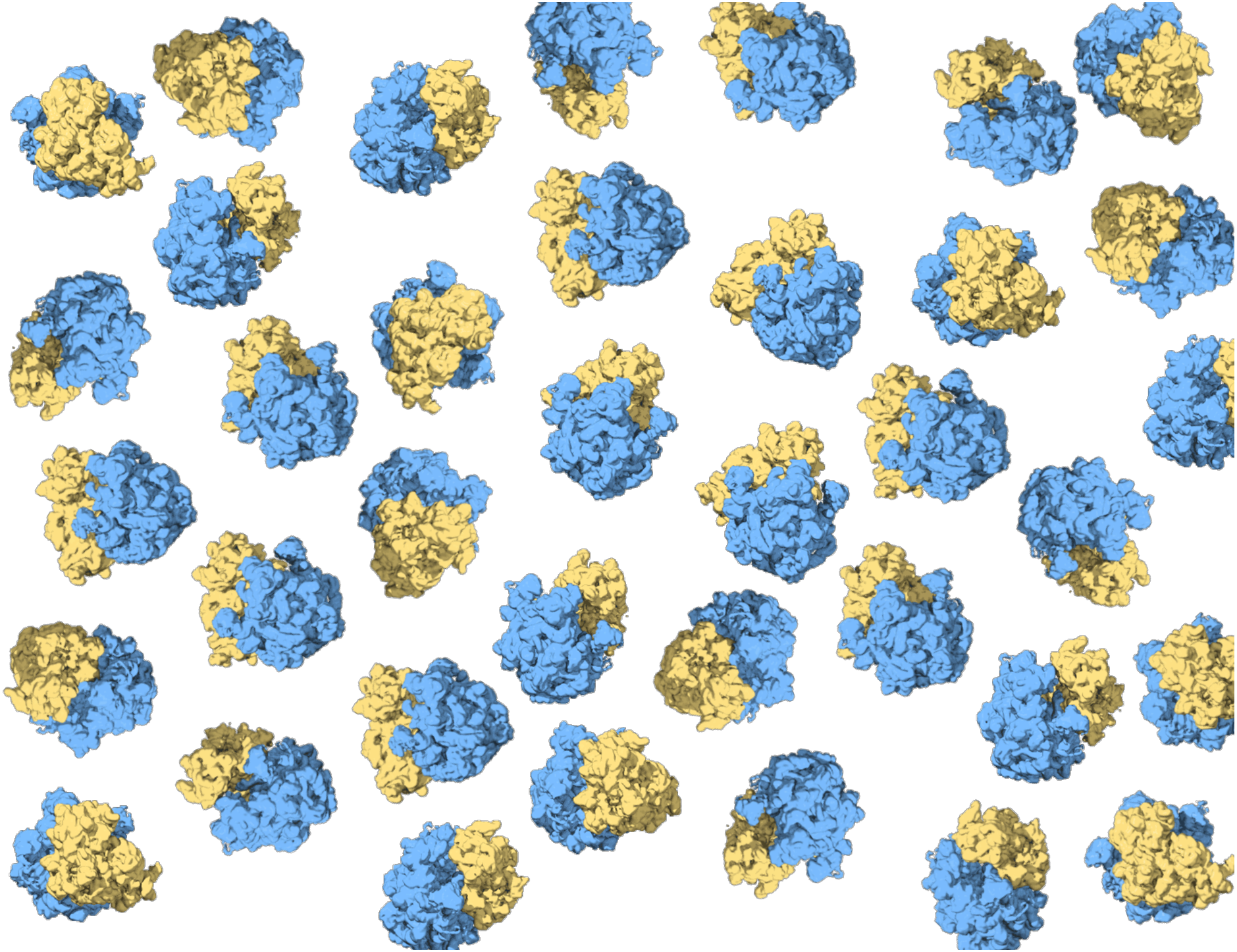


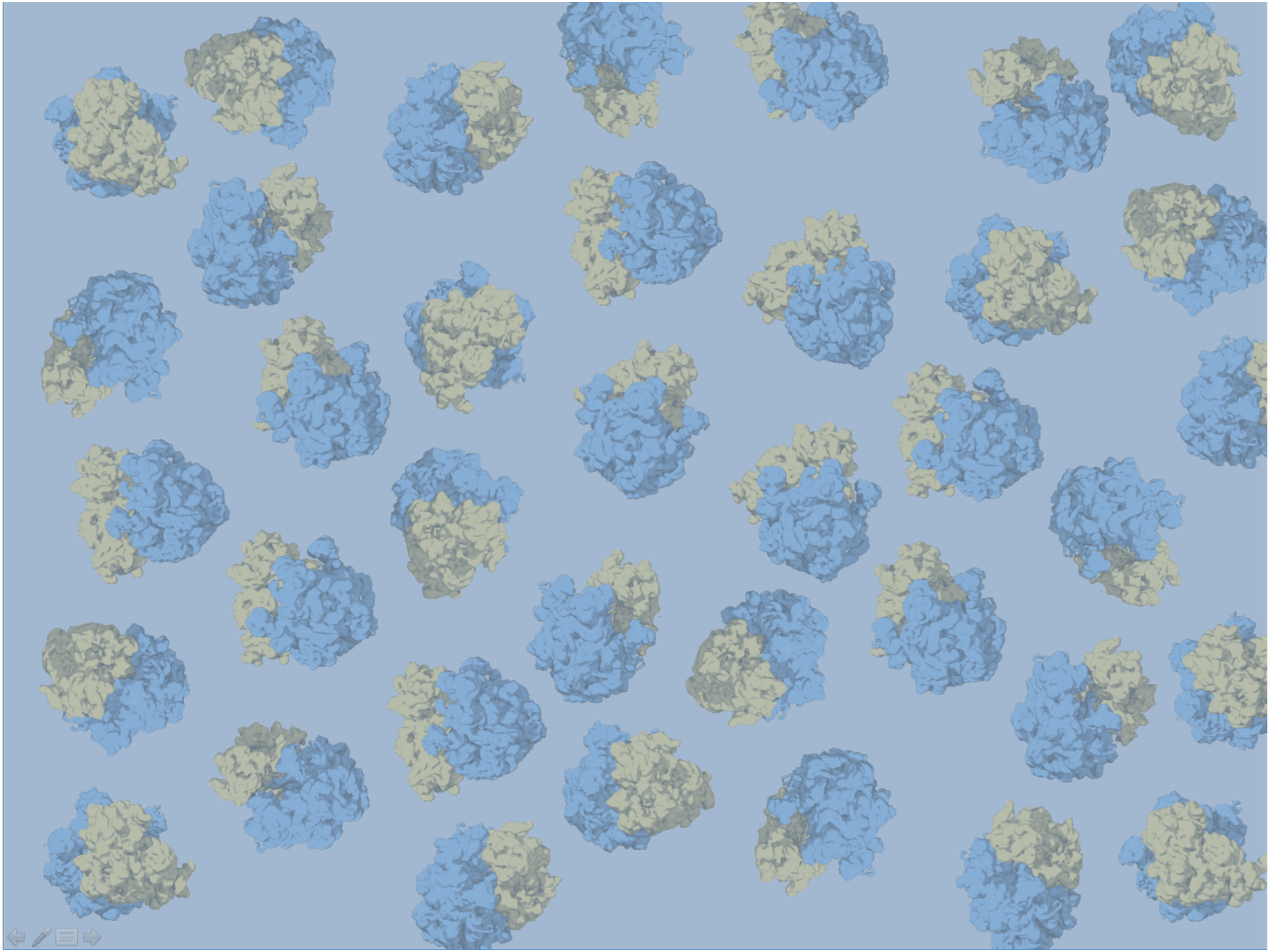
Danny Nam Ho

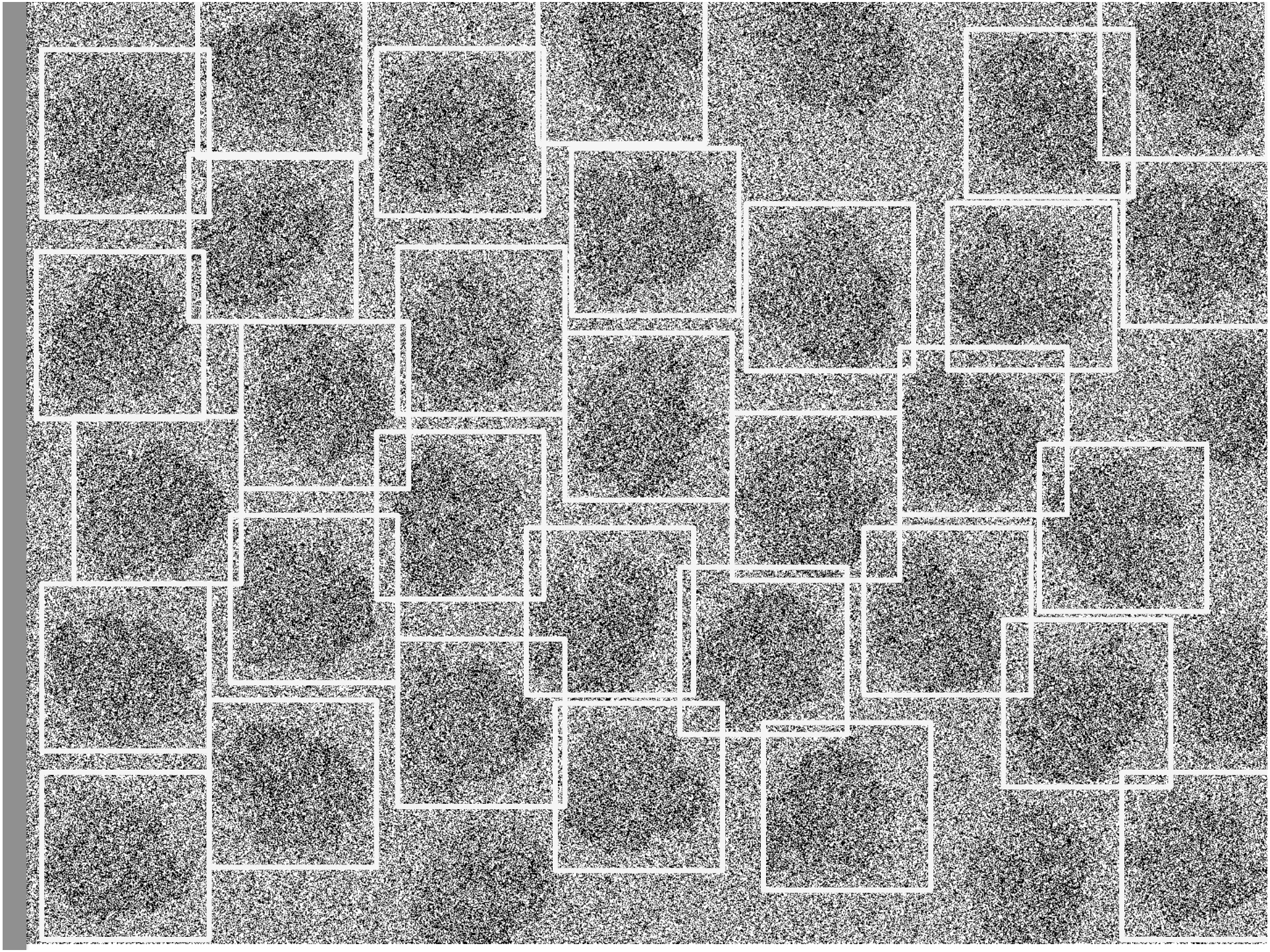

300 μm

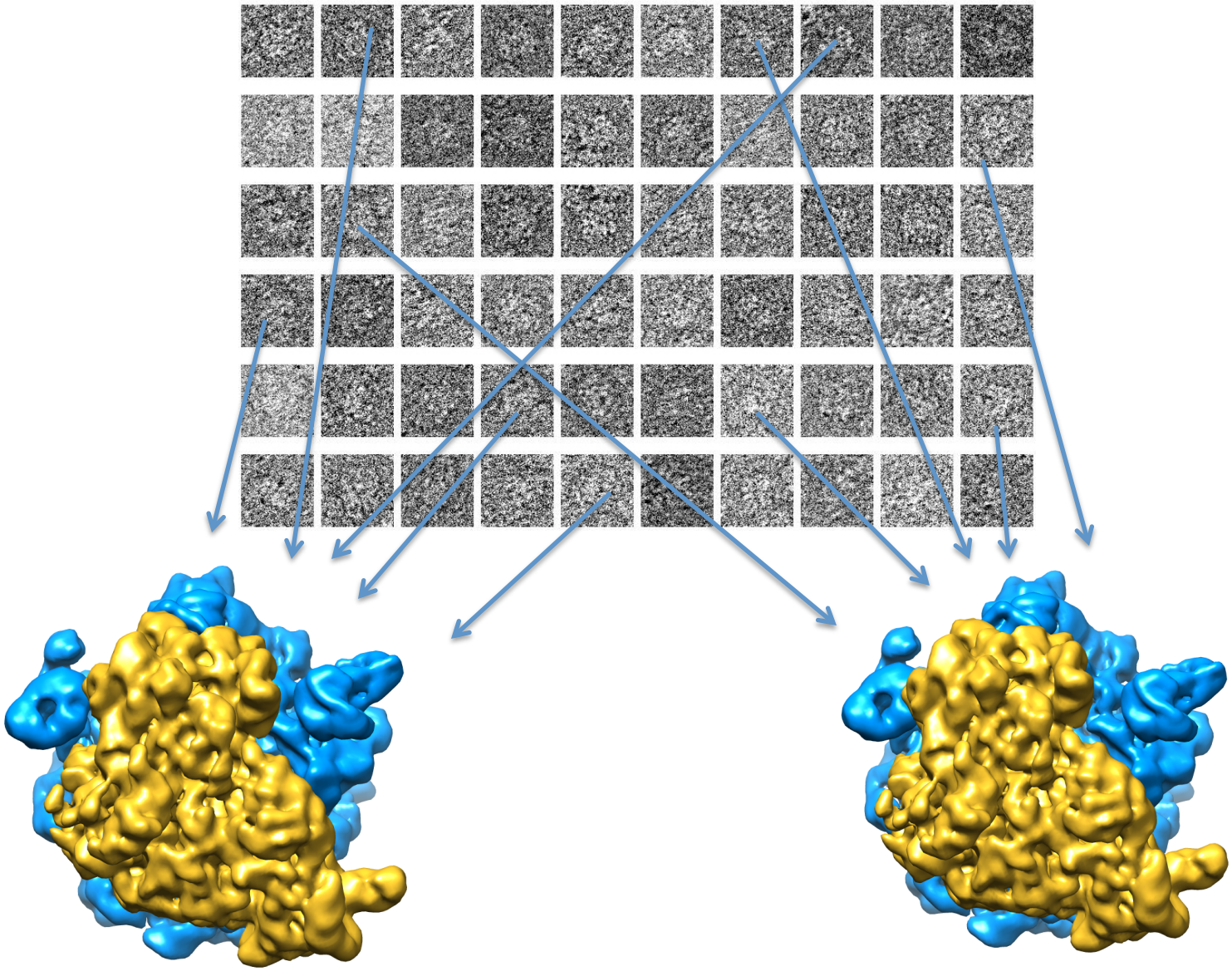

~54 μm


~1 μm









Electron Microscopy: Imaging Molecules in Three Dimensions

Up to now, x-ray diffraction and transmission electron microscopy have been largely complementary techniques for imaging biological structures. Crystallographers use x-ray diffraction for high-resolution mapping of the structure of biological molecules, ranging from relatively simple amino acids to complex proteins, provided that they can be prepared in crystalline form. Electron microscopy has been useful for directly imaging structural features on a much wider variety of specimens but also at a much lower resolution.

This state of affairs may soon change, however, because in the last year Nigel Unwin and Richard Henderson of the Medical Research Council (MRC) Laboratory of Molecular Biology, Cambridge, England, have successfully combined the mathematical technique of image reconstruction from projections with new methods for preparing specimens and minimizing radiation damage. Thus they were able to reconstruct images of a protein molecule to a resolution of 7 angstroms, nearly three times the best resolution previously attainable. This combination of techniques promises to bring to electron microscopy both high resolution and the ability to image periodic assemblies of molecules that are too small to yield information by examination with x-rays.

The protein molecules were components of the purple membrane, which itself is a specialized part of the cell membrane of the bacterium *Halobacterium halobium*. This organism has been much in the news recently as the only known example of a cell that does not contain chlorophyll but that can convert sunlight into energy for its own metabolism. Scientists hope that understanding this unique process will clarify such cell processes as the production of adenosine triphosphate, the energy-storing molecule in cells, as well as lead to new solar energy technologies.

Unwin and Henderson's first innovation was to preserve the purple membrane (an oval sheet with a diameter of about 1 micrometer and a thickness of 45 angstroms) in a dilute glucose solution. This procedure is necessary because in the moderately high vacuum [10^{-3} pascal or better (1 pascal equals 7.53×10^{-3} torr)] of an electron microscope, dehydration of an untreated biological specimen normally causes cracking, fragmentation, or other trauma that distorts its structure.

Moreover, since the atoms in biological molecules are light, they do not strongly scatter electrons. Because of the resulting lack of contrast, microscopists have had to invent staining techniques to produce an artificial contrast. One widely used method, which was a breakthrough in its own right when introduced in the mid-1950's, is to coat the specimen with a dilute solution of a heavy metal salt, such as uranyl acetate. The solution dries and leaves the metal salt to cover the surface and fill crevices on the specimen. This technique is called negative staining because contrast, provided by the metal, is associated with an absence of molecular material.

Negative staining not only provides contrast for imaging but also stabilizes a specimen against the effects of dehydration. Nonetheless, the stabilization is at best incomplete, and only the features of a specimen that have dimensions greater than 20 angstroms are preserved. Moreover, because the stain only coats its surface, features inside the specimen cannot be seen at all. Thus, the 2- or 3-angstrom resolution possible with good electron microscopes cannot be utilized in practice.

According to Unwin and Henderson, the use of glucose alleviates the dehydration problem because, when dried, the surface of the solid glucose looks much like water, but is not volatile. And, since it contains no heavy atoms, glucose does not prevent the internal features of the membrane from being imaged. Electron diffraction patterns taken from purple membrane lattices and also from beef liver catalase crystals indicated that details with dimensions as small as 3.5 angstroms were preserved in the glucose-embedded specimens. This is within shooting distance of what researchers can achieve with x-ray diffraction of materials (including catalase) that crystallize in three dimensions, and indicates that the ultimate resolution attainable with electron microscopy may yet come close to that of x-ray crystallography.

Unfortunately, glucose does not solve the problem of the intrinsically low contrast of biological molecules. The MRC investigators overcame this difficulty by taking advantage of a special property of the purple membrane that also enabled them to solve another perennial problem of imaging biological materials with electron microscopes: Biological molecules are extremely susceptible to radiation

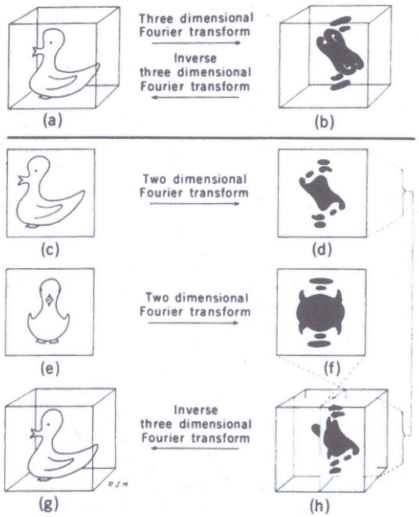


Fig. 1. The mathematical principles of three-dimensional reconstruction. A three dimensional duck (a) and its Fourier transform (b) are approximated as follows. Required are (c) a two-dimensional projection of (a); (d) the two-dimensional Fourier transform of (c); (e) another projection of the duck; and (f) its two-dimensional Fourier transform. The three dimensional duck (g) is calculated from an approximate three-dimensional Fourier transform (h) which was reconstructed from the two-dimensional transforms (d) and (f). [Source: James Lake, New York University Medical School]

Arthur L. Robinson
Science 192 (1976) 360-363

Arthur L. Robinson, at the very end of a 1976 commentary on the first 3D reconstruction of the purple membrane protein by Unwin and Henderson (Nature, 1975):

"Joachim Frank of the State of New York Department of Health, Albany, has been exploring theoretical methods for averaging data from arrays of identical objects that are not periodic. If such methods were to be perfected, then, in the words of one scientist, the sky would be the limit."

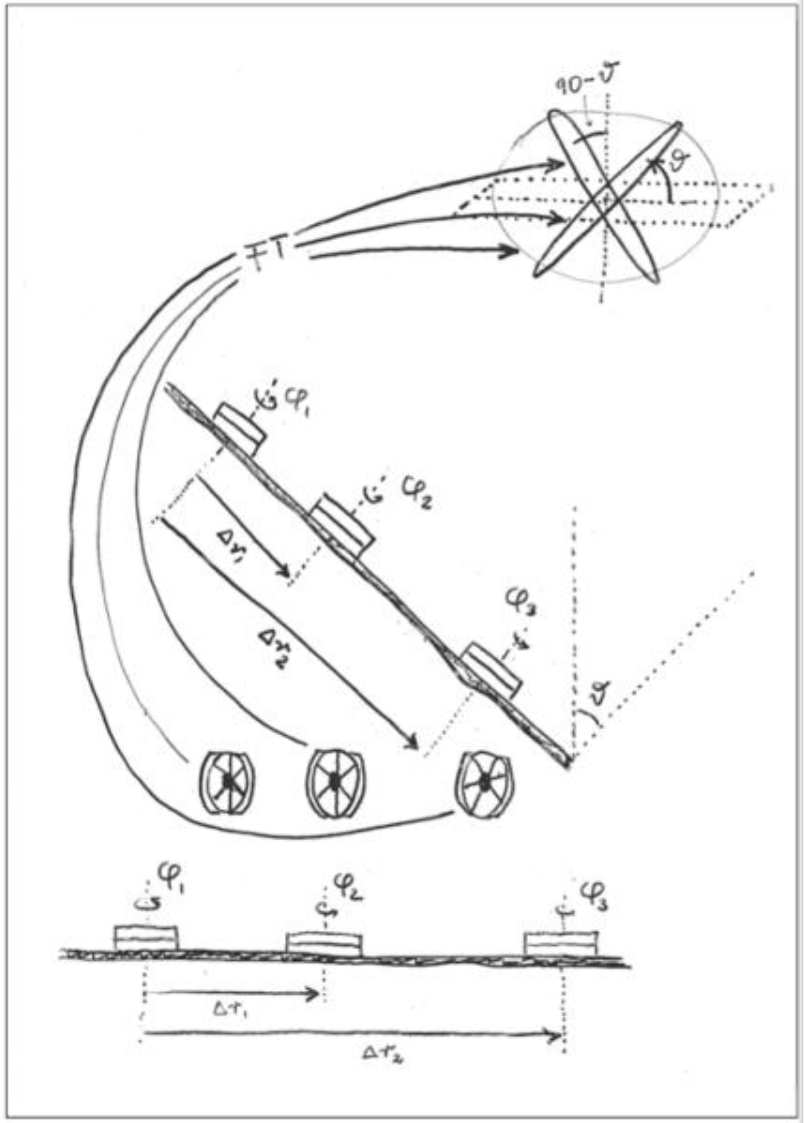
Science 192 (1976) 360-363

1976

an idea


Single-particle reconstruction: 40 Years

RANDOM-CONICAL RECONSTRUCTION

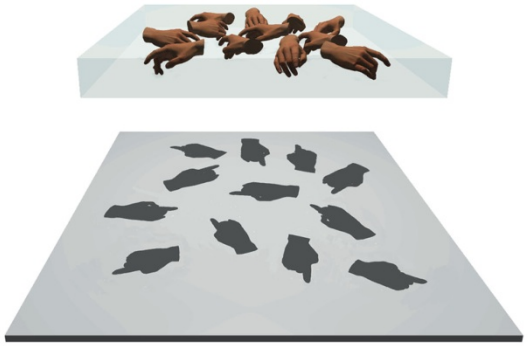


“Overhead”

1979

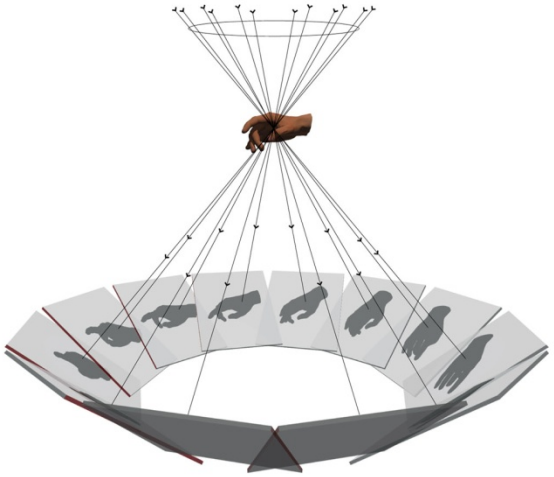
an idea

HOW TO DETERMINE ORIENTATIONS OF PARTICLE PROJECTIONS



Random-conical data collection

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



1 General packages WIKIPEDIA 2016

- 1.1 Appion
- 1.2 Bsoft
- 1.3 Cyclops
- 1.4 EMAN2
- 1.5 EMAN
- 1.6 Eos
- FREALIGN
- 1.7 IMAGIC
- 1.8 IPLT
- 1.9 MDPP
- 1.10 MRC IMAGE PROCESSING PACKAGE
- 1.11 RELION
- Scipion
- 1.12 SIMPLE
- 1.13 SPARX
- **1.14 SPIDER ---------- 1981**
- 1.15 Suprim
- 1.16 Xmipp

-- early fundamentals (80s) --

- Alignment by cross-correlation
- Feasibility of alignment at low dose
- Multivariate statistical analysis and 2D classification
- Resolution measured by extent of reproducibility in Fourier space
- Bootstrap reconstructions: random conical and via common lines
- Reconstruction with arbitrary irregular geometry
- Multireference methods in 2D classifiaction

Cross-corrrelation alignment of particles

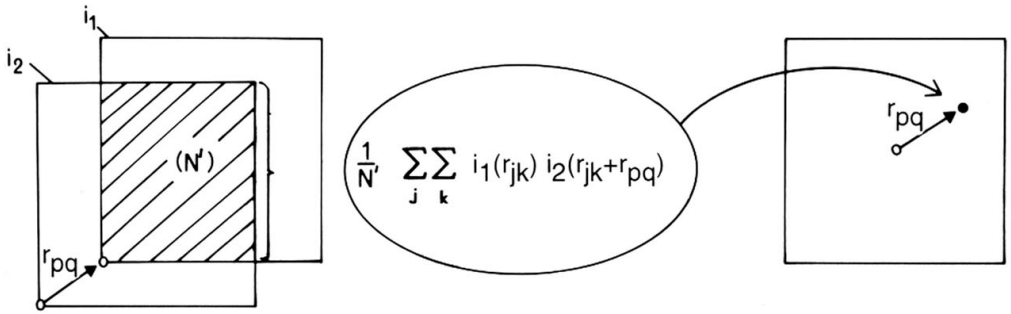
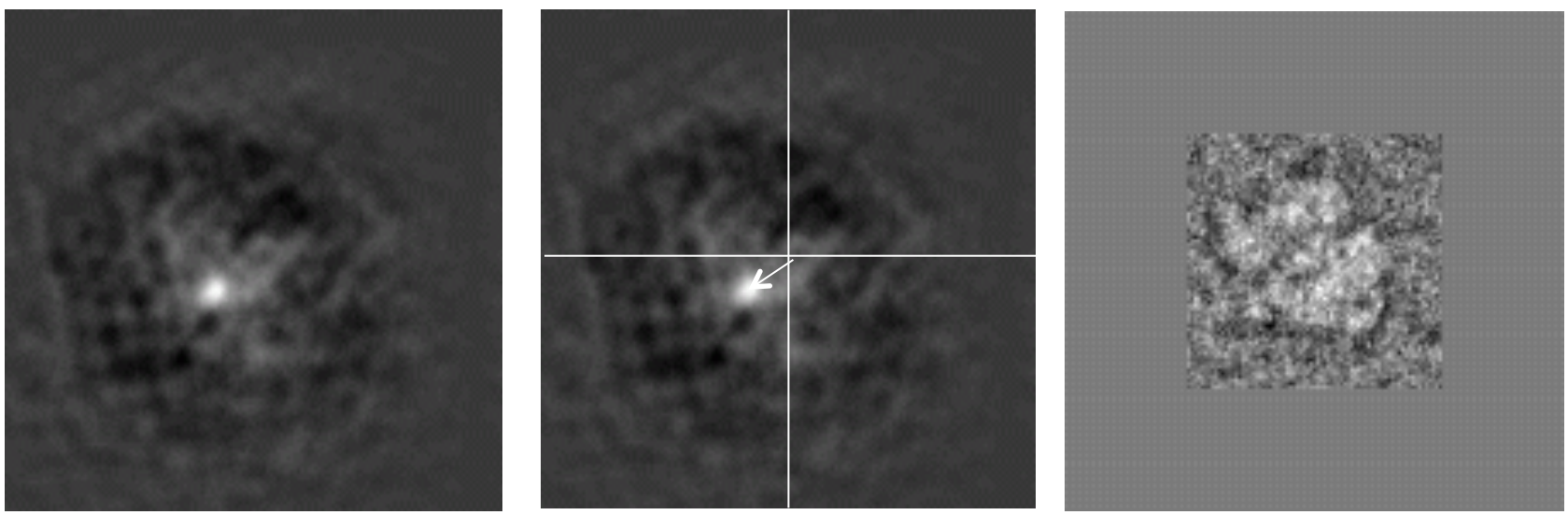
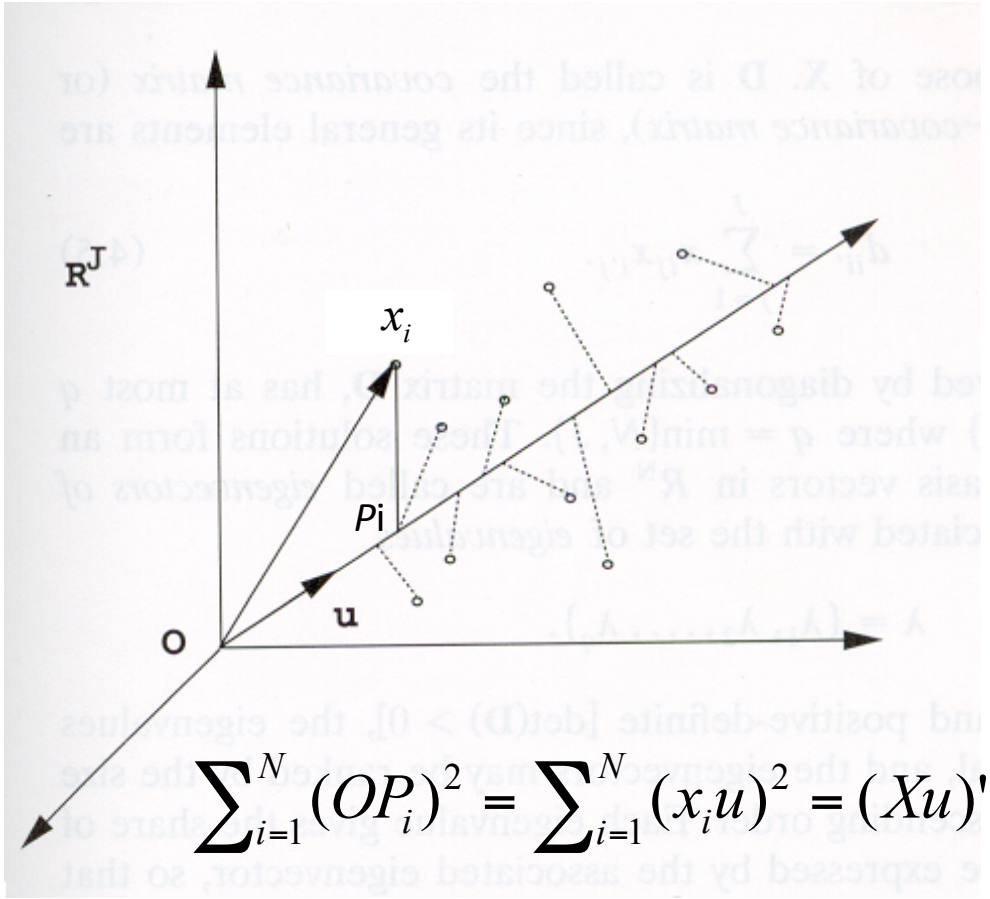
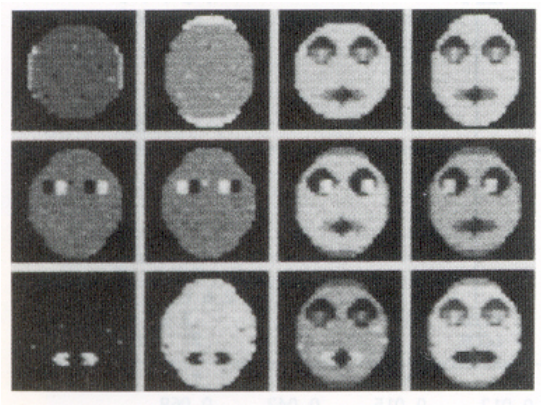


Fig. 3.8. Definition of the cross-correlation function. Image 1 is shifted with respect to image 2 by vector $\mathbf{r}_{pq}$. In this shifted position, the scalar product of the two images arrays is formed and put into the CCF matrix at position (p, q). The vector $\mathbf{r}_{pq}$ is now allowed to 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.



J. Frank, 3D Electron Microscopy of Macromolecular Assemblies, Oxford 2006





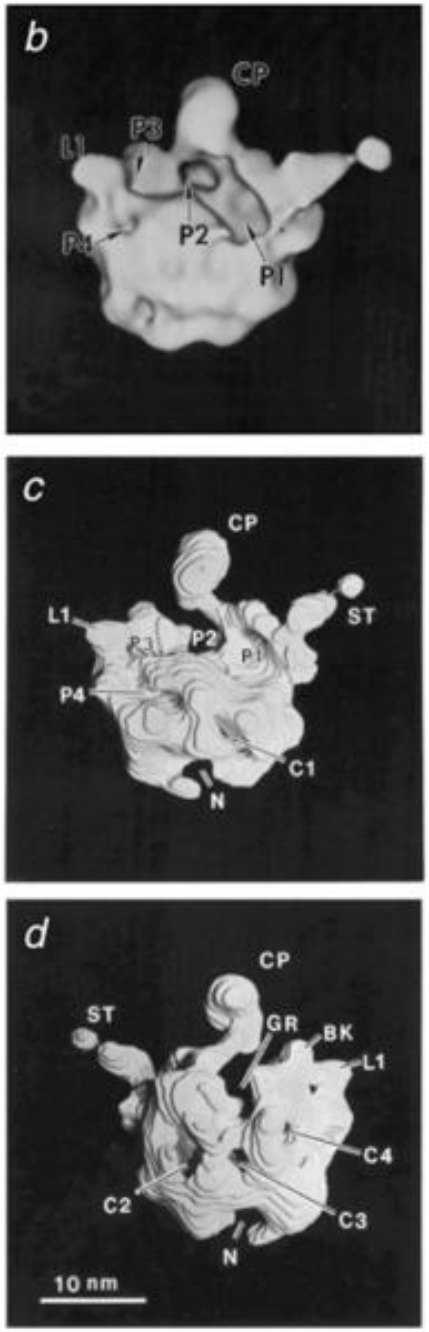
X= matrix containing N image vectors (each with J elements) as rows

$$\sum_{i=1}^{N}(OP_i)^2 = \sum_{i=1}^{N}(x_i u)^2 = (Xu)'X'u = u'X'Xu \xrightarrow{!} \max$$

J. Frank, 3D Electron Microscopy of Macromolecular Assemblies, Oxford 2006

- A fast view back: 30 years of progress in ribosome reconstructions
- What's new and promising in developments of sample preparation, instrumentation, and data analysis
- Our own contributions to the future: (1) time-resolved imaging (2) analysis of continuous conformational changes by manifold embedding

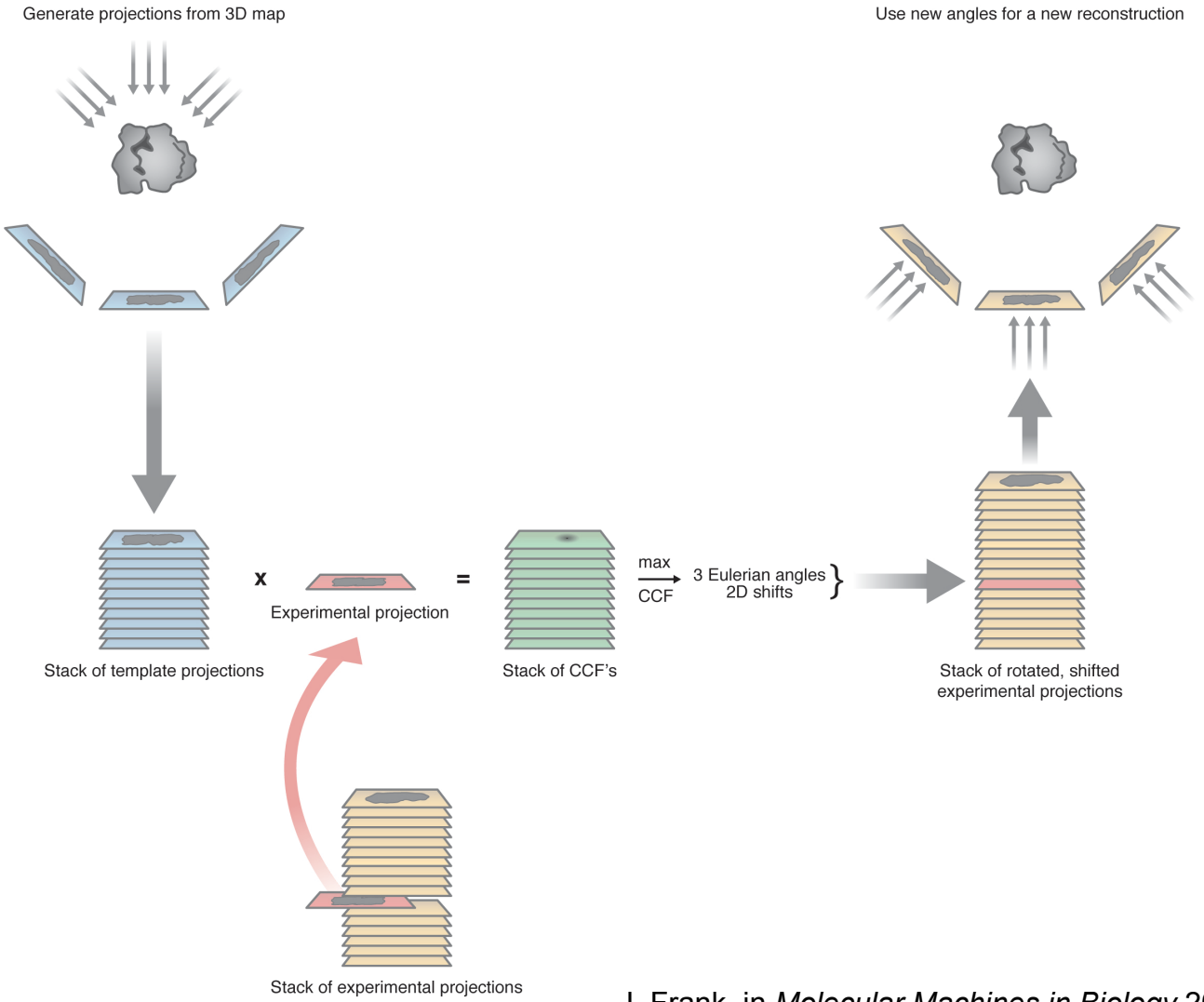
First random-conical reconstruction
Negative stain



1987
25 Å

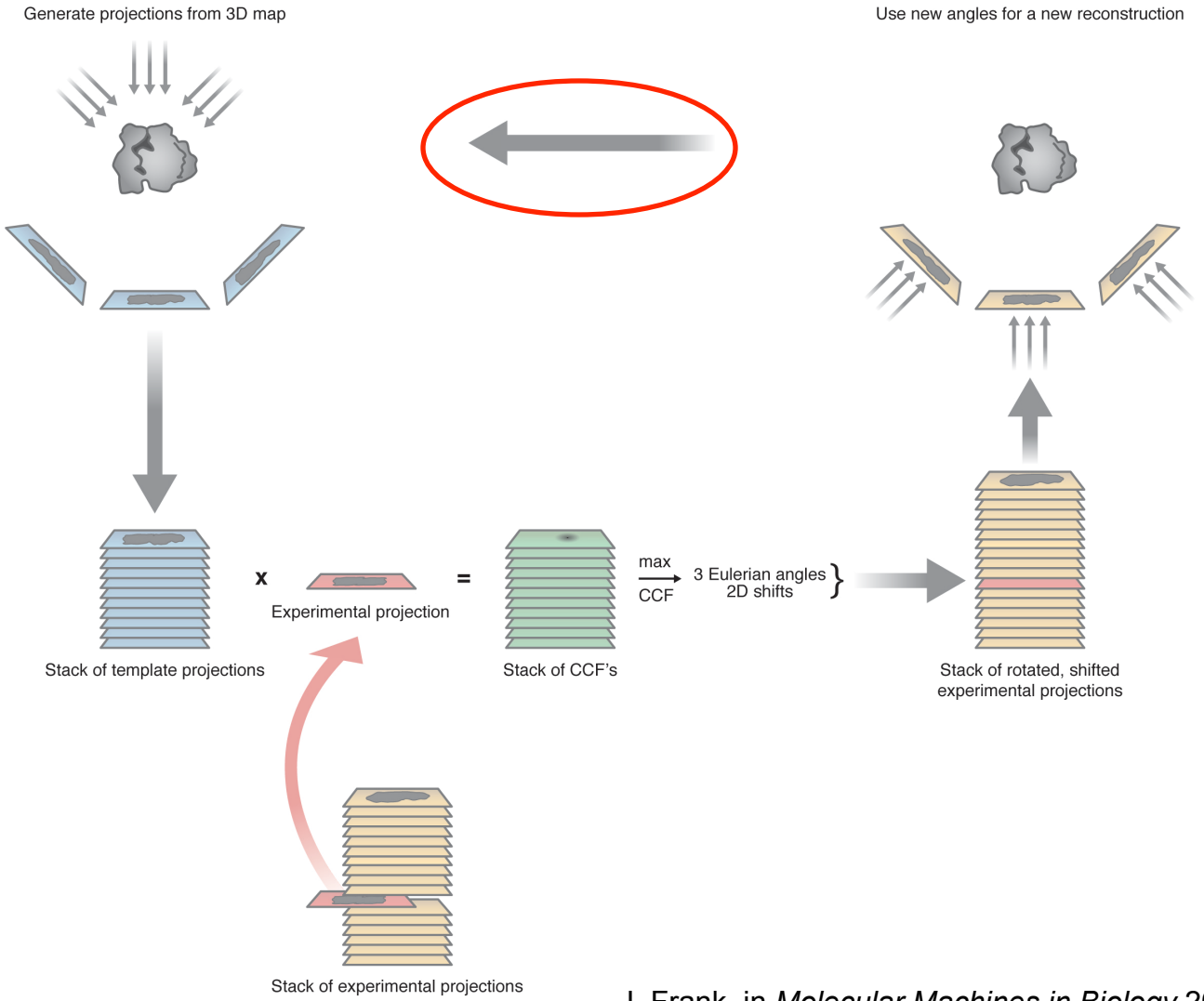
Radermacher et al.,
J. Micrscopy

Orientation determination by reference to an existing reconstruction (supervised classification)

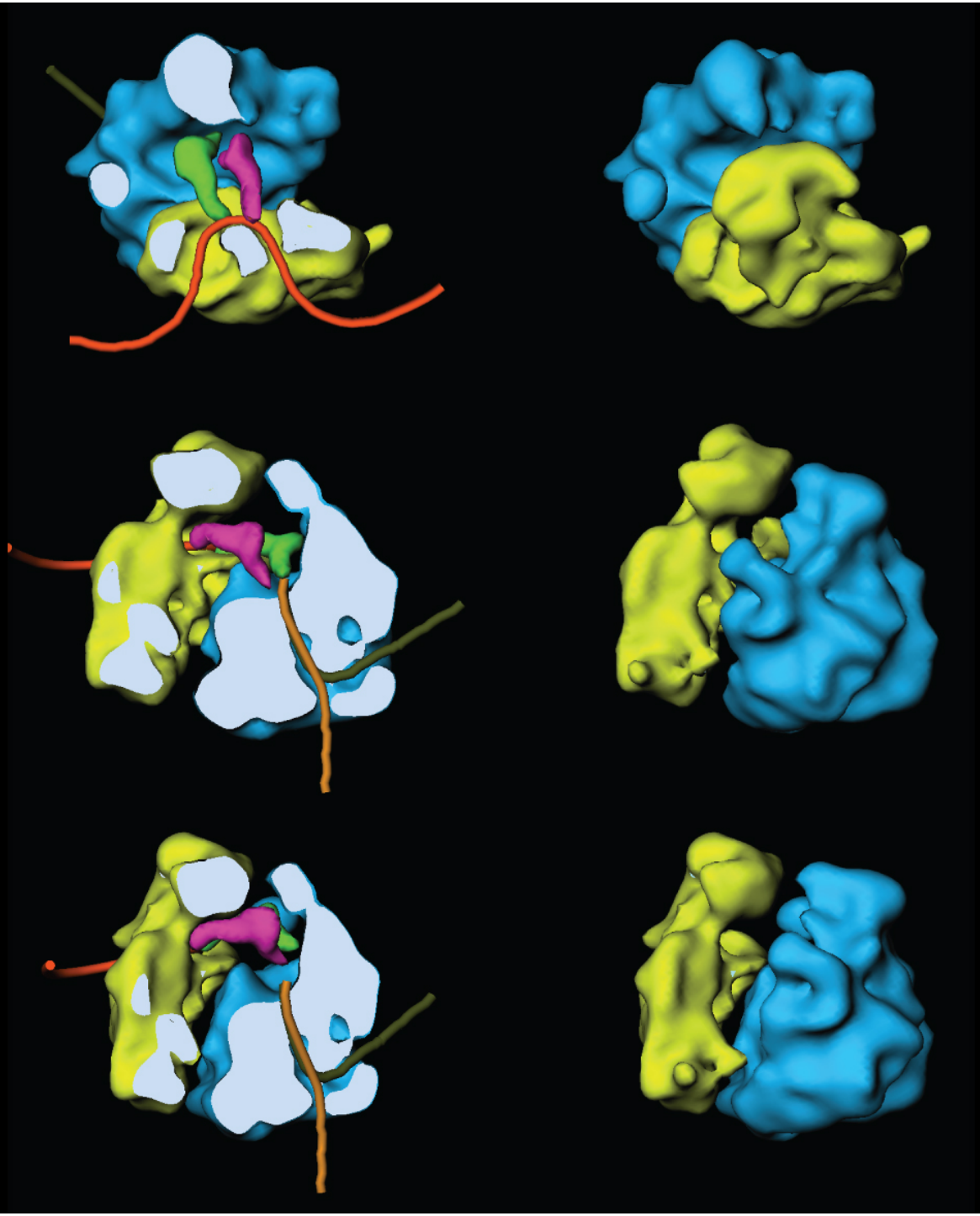


J. Frank, in *Molecular Machines in Biology* 2011

Iterative Angular Refinement



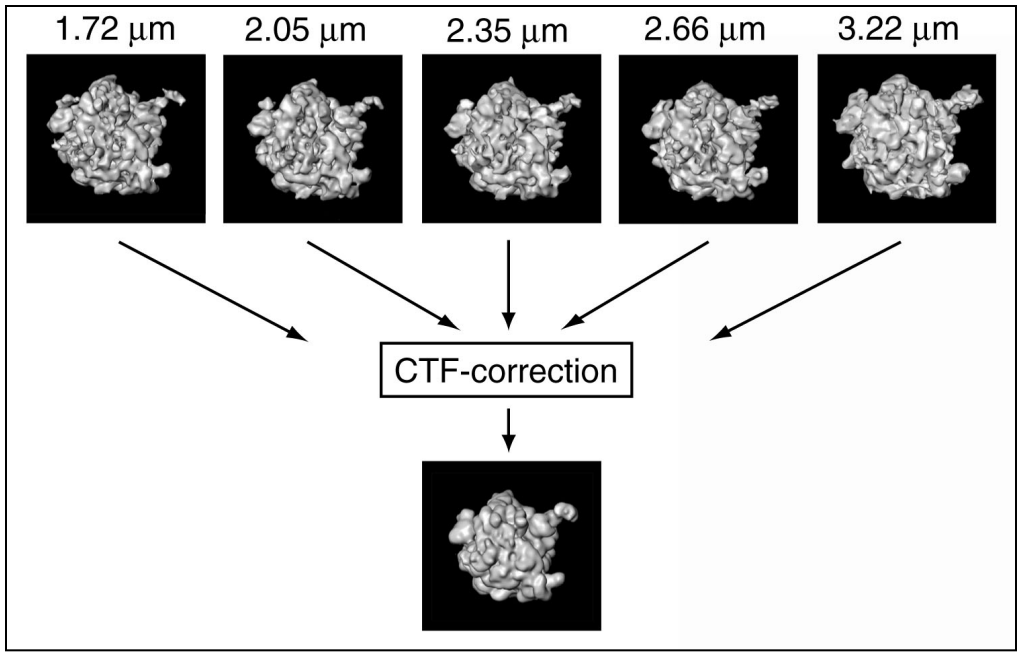
J. Frank, in *Molecular Machines in Biology* 2011



1995
25 Å

Frank et al.,
Nature

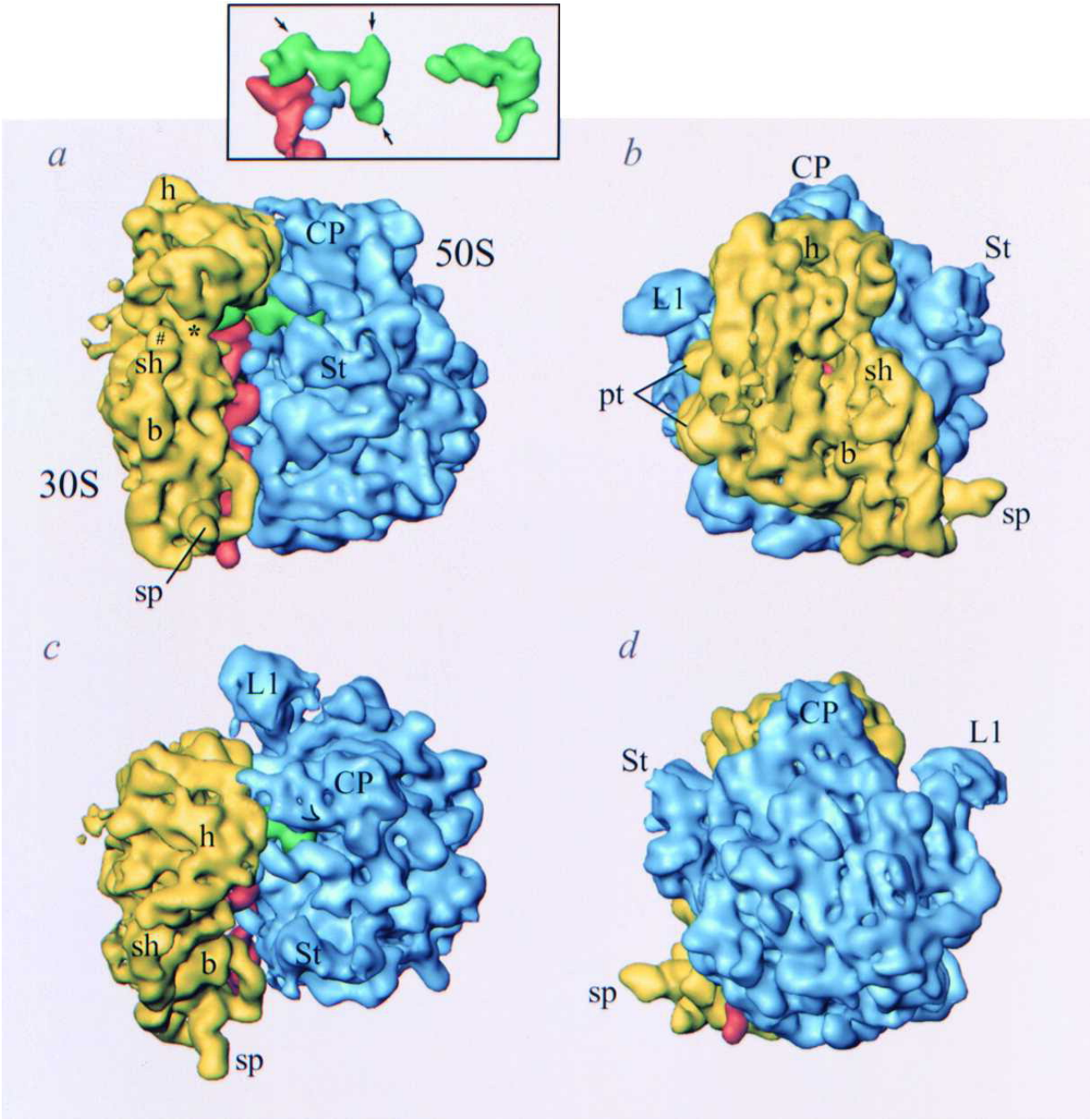
CTF correction and merging of defocus group reconstructions by Wiener filtering



$$F(\mathbf{k}) = \sum_{n=1}^{N} W_n(k) F_n(\mathbf{k})$$

$$W_n(k) = \frac{SNR_n(k) H_n^*(k)}{\sum_{n=1}^{N} SNR_n(k) \, | H_n(k) |^2 + 1}$$

Penczek et al., Scanning Microscopy 1997



2000

11.5 Å

Gabashvili et al., Cell

- Reference-based 3D classification
- MDFF and other flexible fitting approaches

SUPERVISED CLASSIFICATION

(Valle et al., EMBO J. 2002)

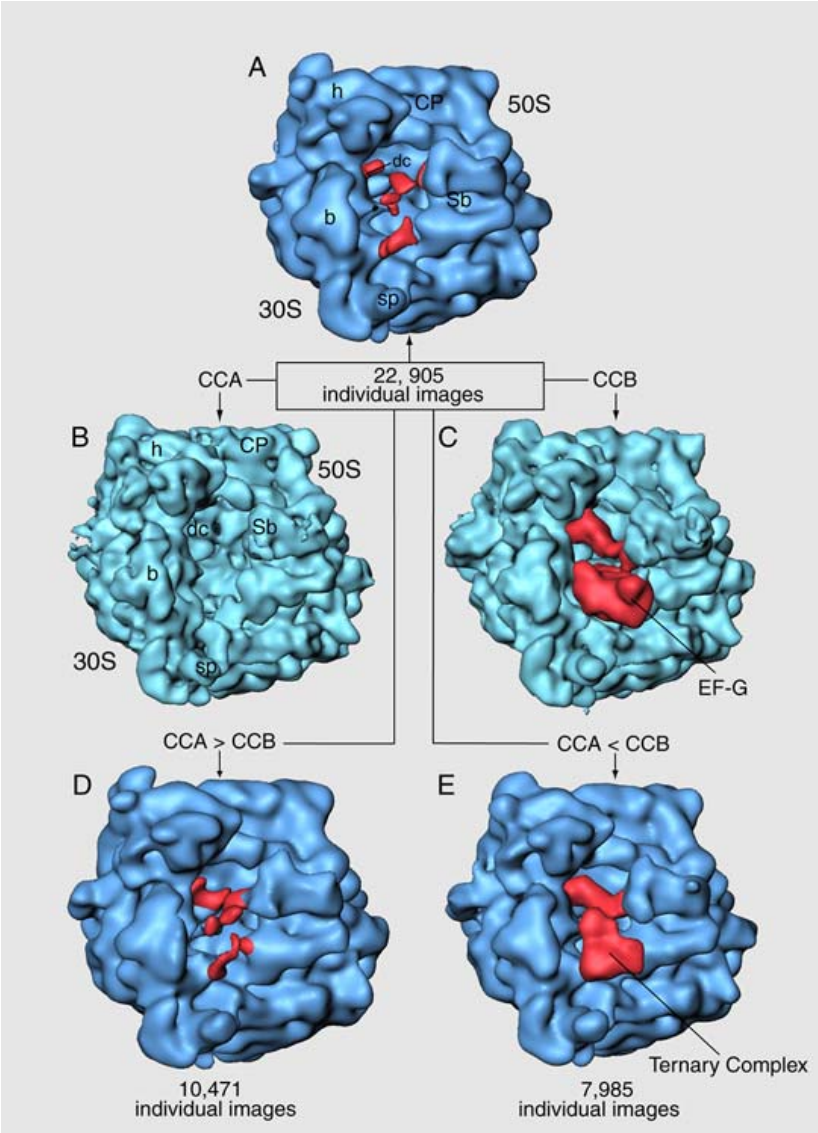
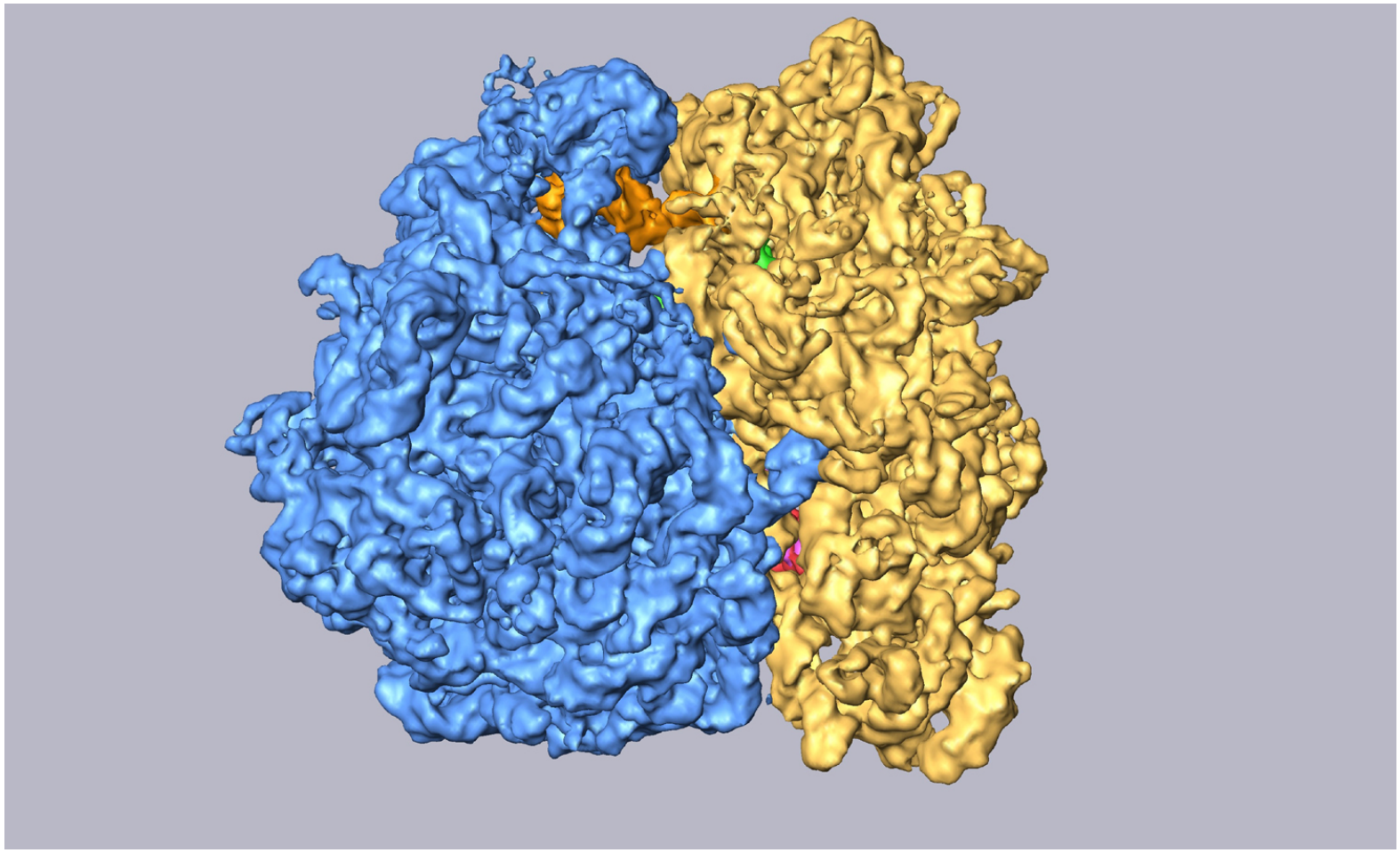


Fig. 1. General scheme of the classification strategy based on the comparison of the cross-correlation coefficients (CC) of the individual particles with two different references (CCA = with reference A, *versus* CCB = with reference B). (A) Cryo-EM map of the 70S ribosome-ternary complex using 22,905 particles. (B) Reference A. Cryo-EM map of 70S ribosome·fMet-tRNA$_f^{Met}$ complex (Gabashvili et al., 2000). (C) Reference B. Cryo-EM map of 70S ribosome·EF-G·GDP·fusidic acid complex (M. Valle, unpublished results). (D) Cryo-EM map #1, computed from 10,471 individual particles, with a resolution of 16.8 Å. (E) Cryo-EM map #2, computed from 7,985 individual particles, with a resolution of 16.8 Å.
Landmarks are as follows: CP, central protuberance; Sb, L7/L12 stalk base; sp, spur; b, body; h, head; dc, decoding center.



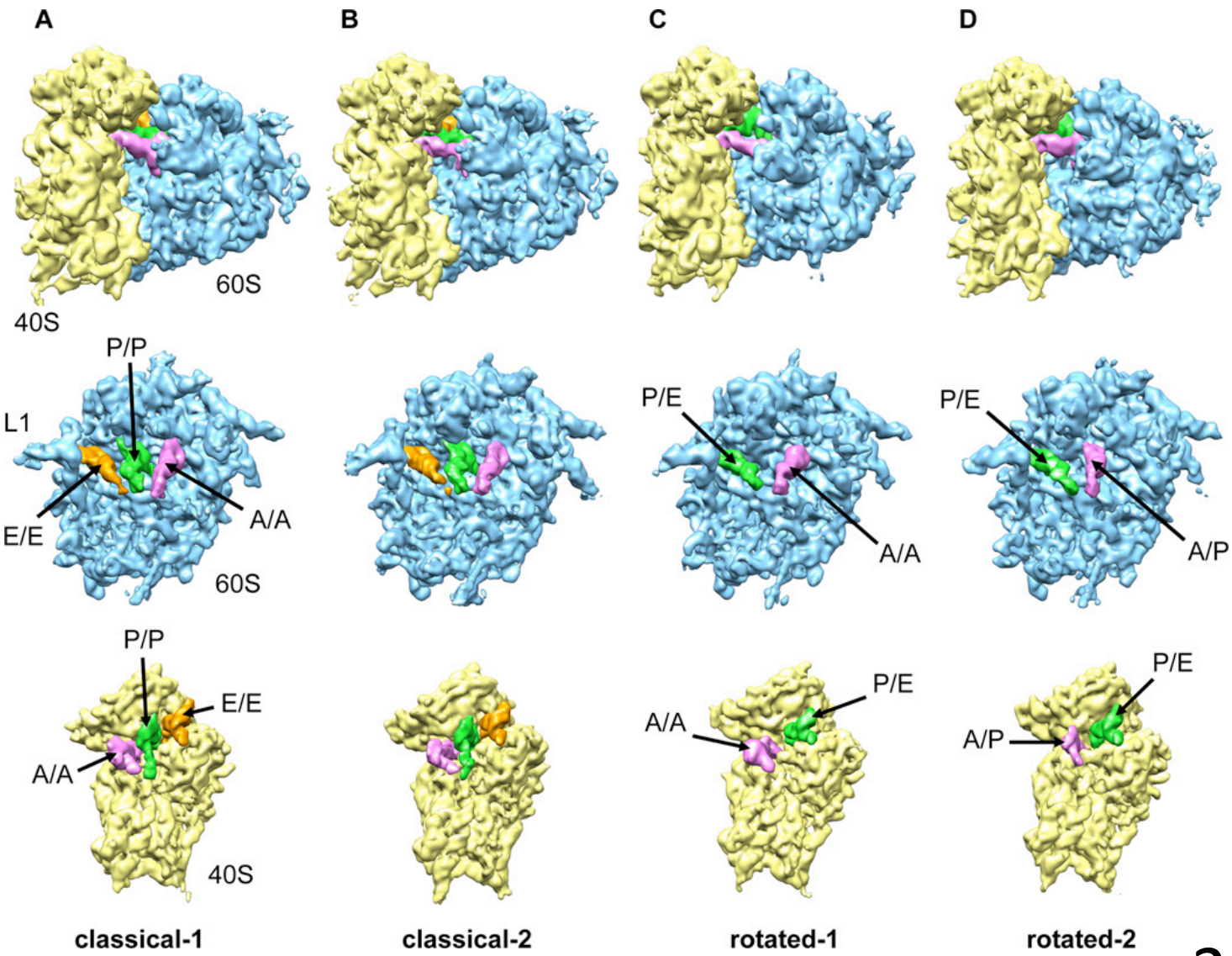
E. coli 70S with ternary complex bound

2008

6.7 Å

LeBarron et al.,
J. Struct. Biol.

- Multiparticle refinement
- “Story in a sample”



Multiple structures from the same sample (mammalian 80S PRE compex)

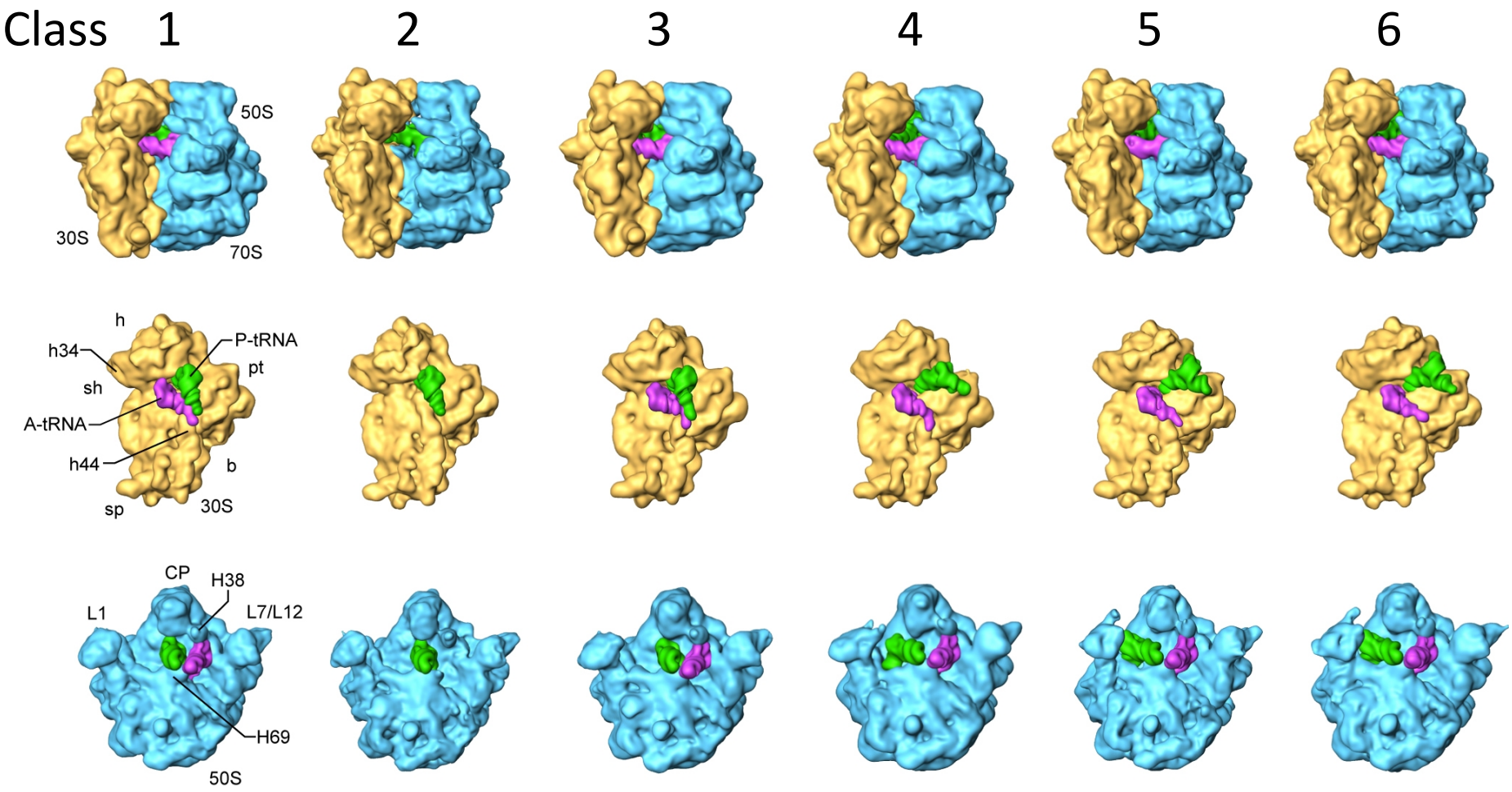
2011

9 to 11 Å

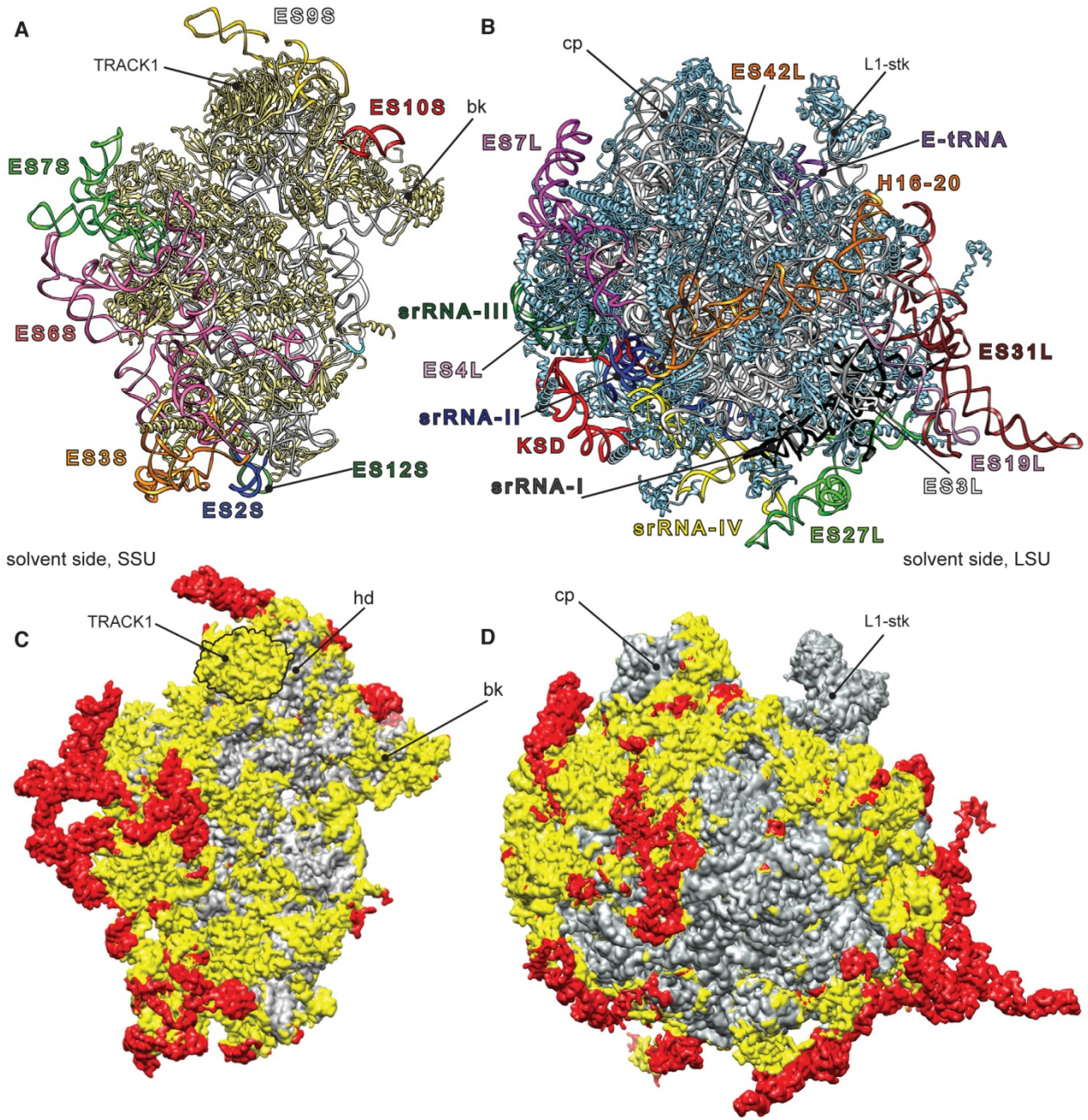
Budkevich et al.
Mol. Cell

- Maximum likelihood classification
- 3D variance estimation

“**STORY IN A SAMPLE**” -- intermediate states in the ratchet-like motion and hybrid tRNA positions in the absence of EF-G



Agirrezabala et al., PNAS 2012



T. brucei 80S ribosome

2013
5.5 Å
Hashem et al.
Nature

2012: Direct electron detection camera

2014
3.2 Å
Wong et al.
Nature

2014
3.4 Å
Vorhees et al.
Cell

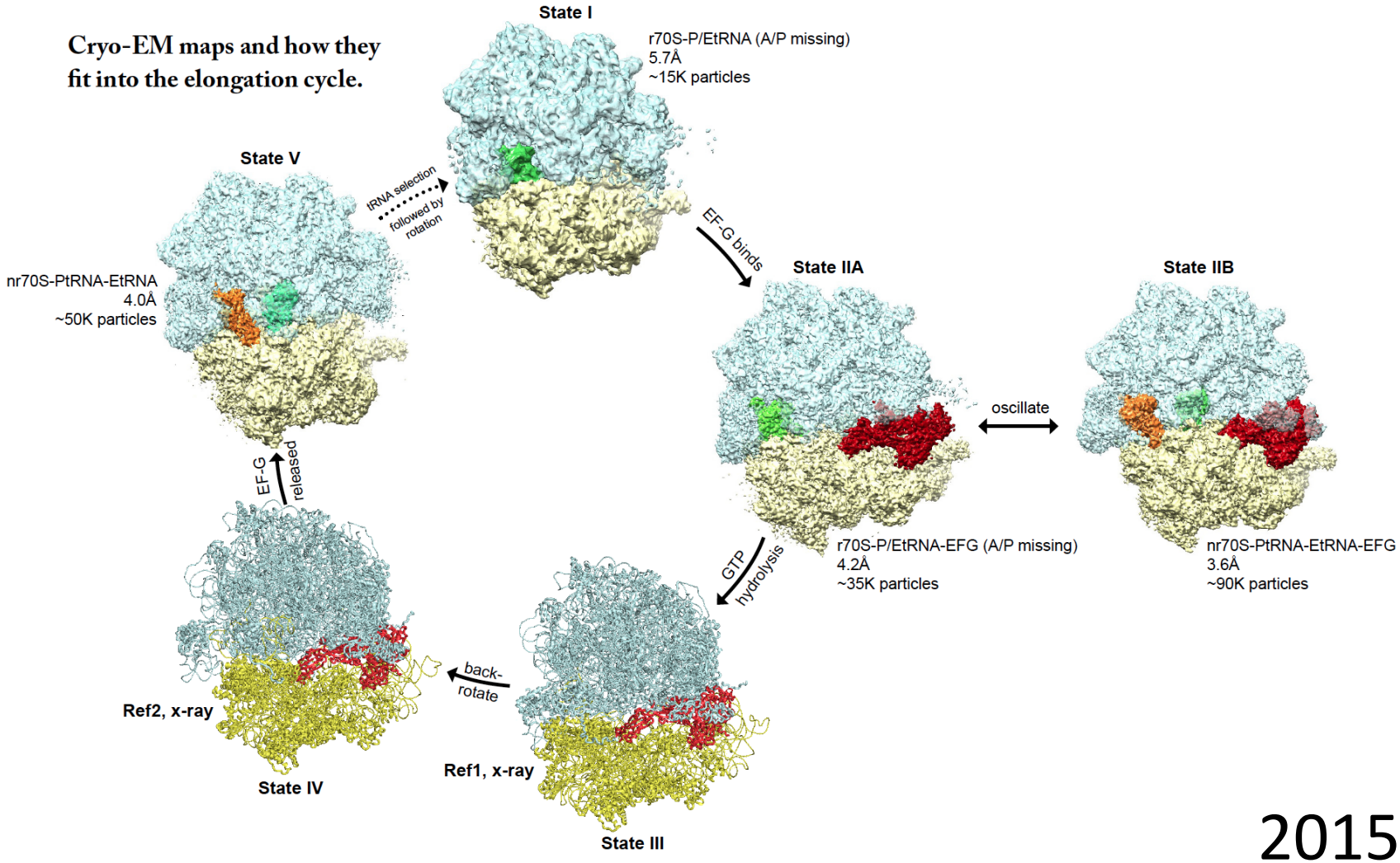
2013
4.1 Å
Bai et al.
eLife

2014
3.7 Å
Fernandez et al.
Cell

2014
3.2 Å
Amunts et al.
Science

2015
2.9 Å
Fischer et al.
Nature

Ribosomes everywhere!



Multiple structures from the same sample (70S-EF-G--H91A)

2015

3.6 Å

Li et al. Science Advances

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

Slides related to this unpublished work are not shared.

Future developments: conceivable/ likely/desirable/mandatory

SAMPLE PREPARATION:

- Self-blotting nanowire grids for better reproducibility and faster operation
- Graphene grids
- SPOTITON grids with multiple sample screening. Overcome waste of grid real estate, waste of sample, waste of time
- Time-resolved cryo-EM using mixing/spraying

INSTRUMENTATION:

- Rumors: even better Cameras are on their way!
- Inexpensive EMs that are affordable for a lab and fit into a standard room
- More competition for FEI would be VERY healthy both for FEI and the customers
- Phase plates: no more CTF correction!

COMPUTATION:

- Standardization of formats, procedures
- Interoperability
- Focused classification with background subtraction
- Increasing role of GPUs for routine tasks
- Continuous conformational changes

Graphene-Base Grids

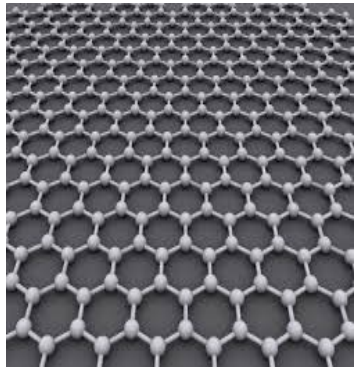
Development of graphene-base substrates for use in electron microscopy

Graphene: mono-layer of carbon atoms with ideal mechanical and electrical characteristics:

-transparent for electrons

-high conductive capacity (sample charge release)

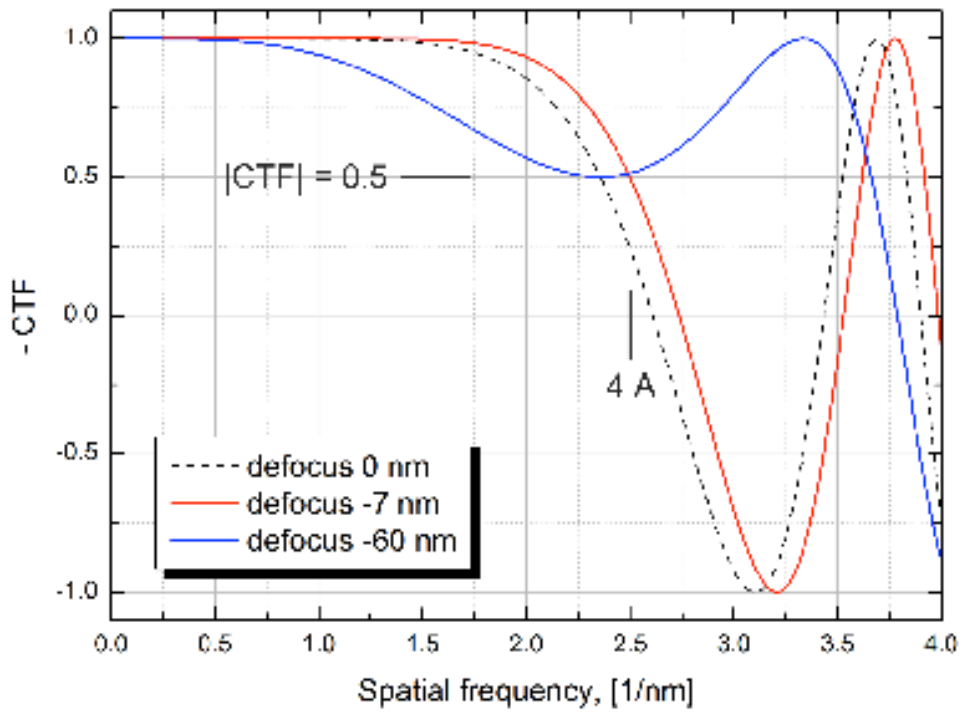
-easily modified: precise control of the number of particle adsorbed/Å^2



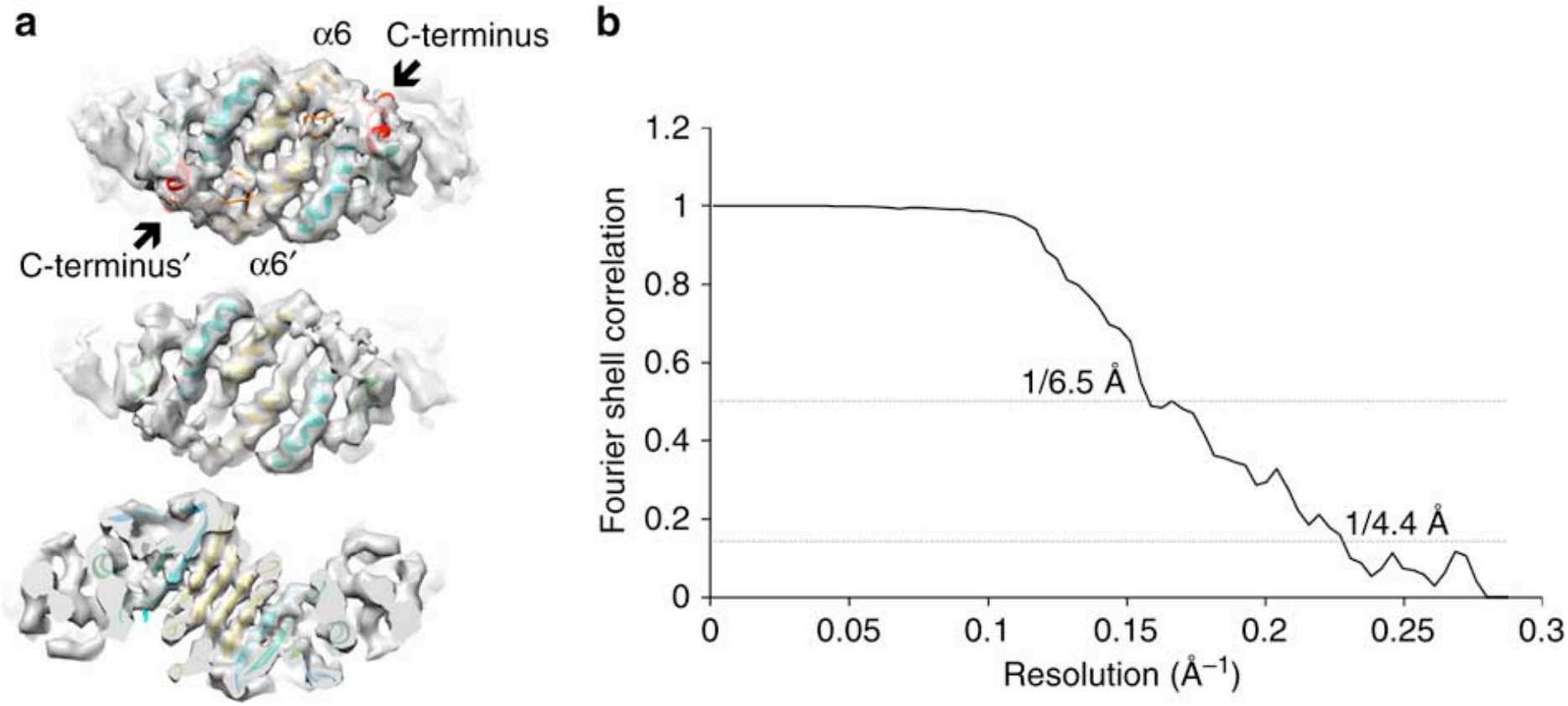
Chris Russo

Russo and Passmore.
Nat Methods. 2014

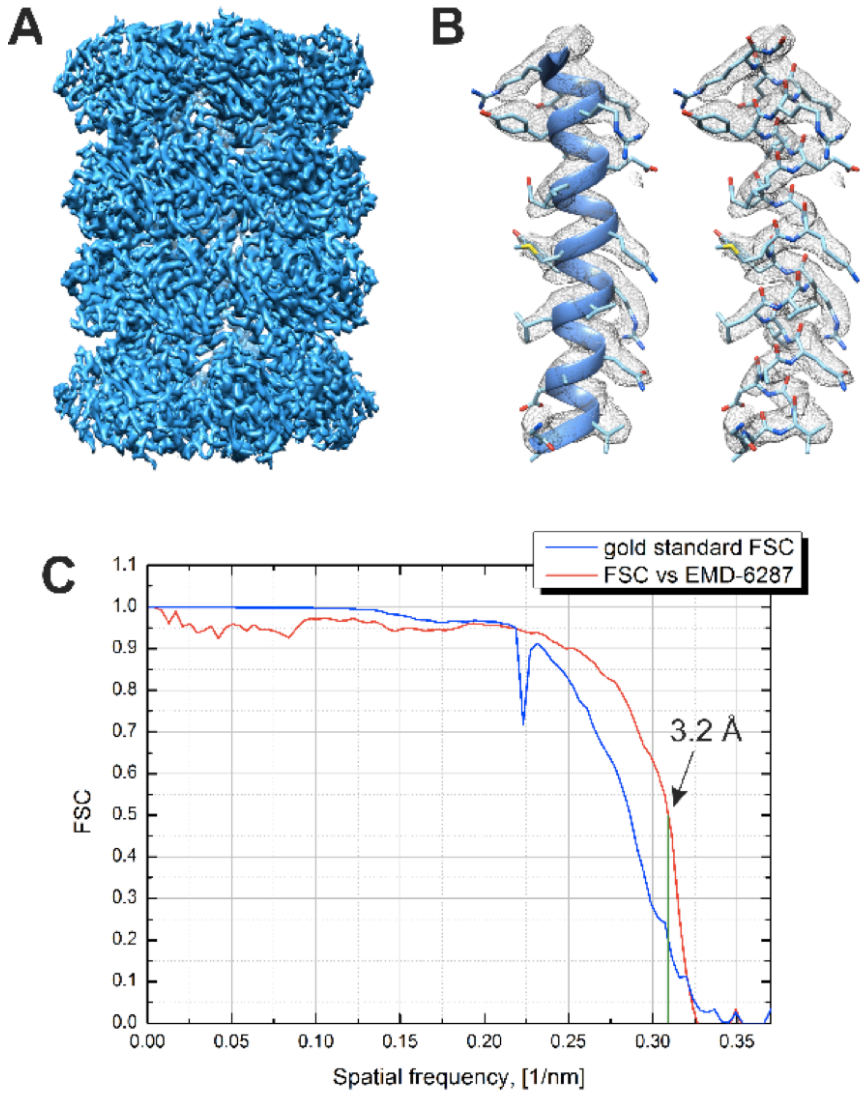
Instrumentation: PHASE PLATES



Radostin Danev & Wolfgang Baumeister, eLIFE 2016



Maryam Khoshouei, Mazdak Radjainia, Amy J. Phillips, Juliet A. Gerrard, Alok K. Mitra, Jürgen M. Plitzko, Wolfgang Baumeister & Radostin Danev, Nature Communications 2016



Thermoplasma acidophilum 20S proteasome at 3.2Å resolution

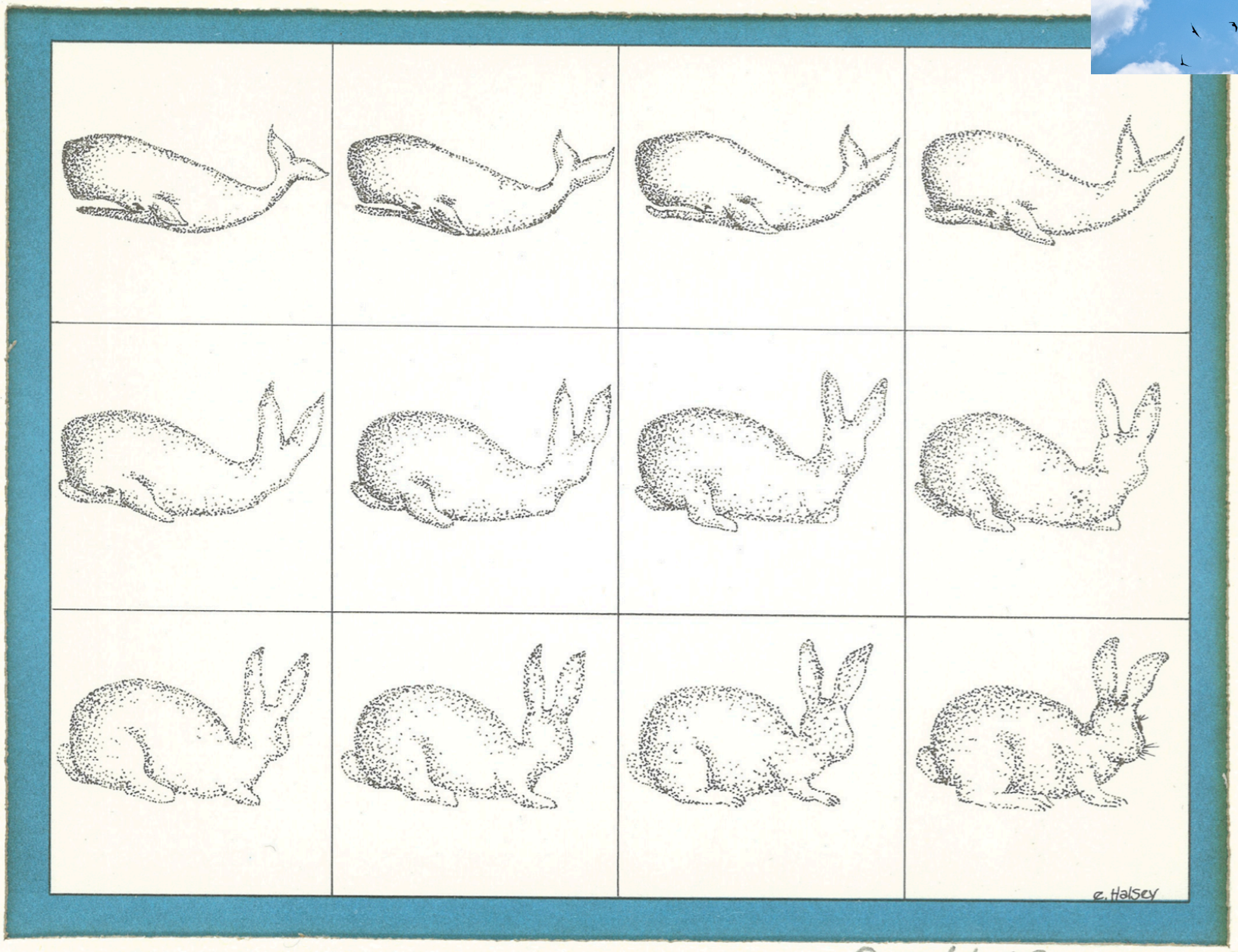
Danev & Baumeister, eLIFE

Time-resolved cryo-EM

- Applications to ribosome recycling
- Newer apparatus
- New plastic-based chip

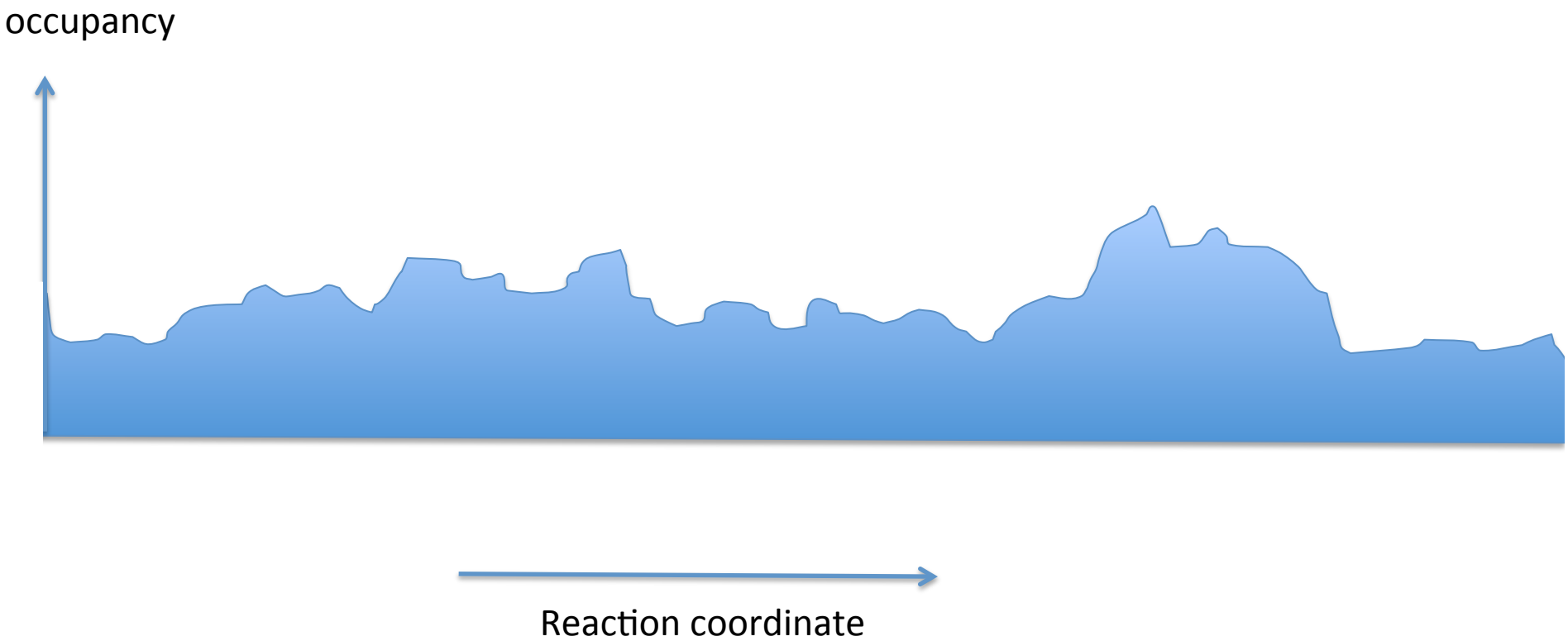
All this work is unpublished, not shared.

Continuous changes in conformation

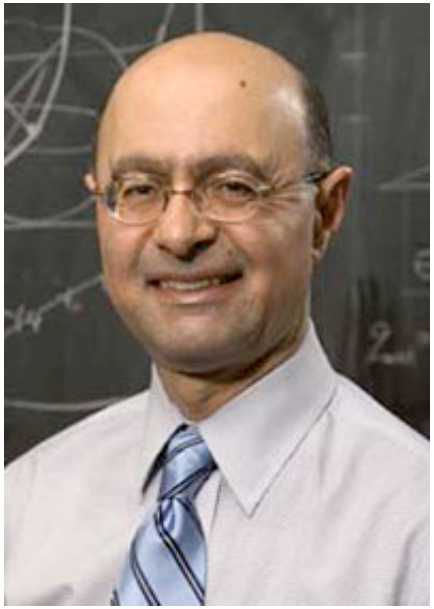


© E. Halsey 1978

Fine and coarse granularity in the occupancy of of “states”



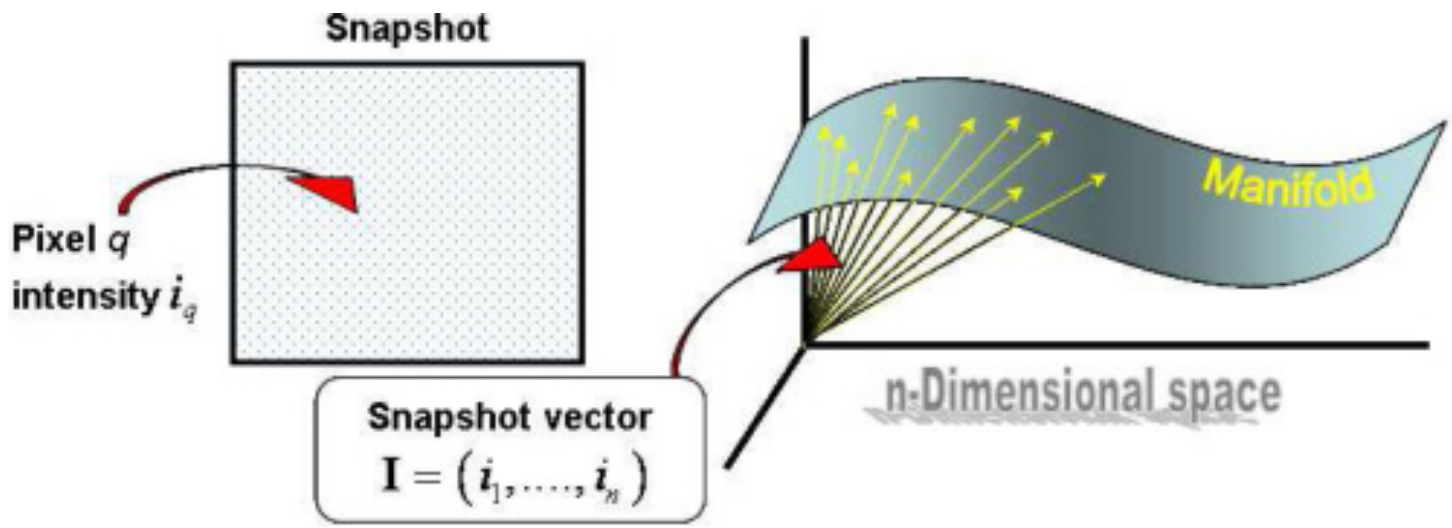
MANIFOLD EMBEDDING APPLIED TO CRYO-EM DATA



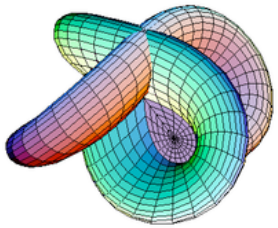
Abbas Ourmazd Ali Dashti
University of Wisconsin at Milwaukee

Dashti et al., PNAS 2014; Chen and Frank, Microscopy 2015; Frank and Ourmazd, Methods 2016

Representation of (aligned) images as vectors in an n-dimensional space

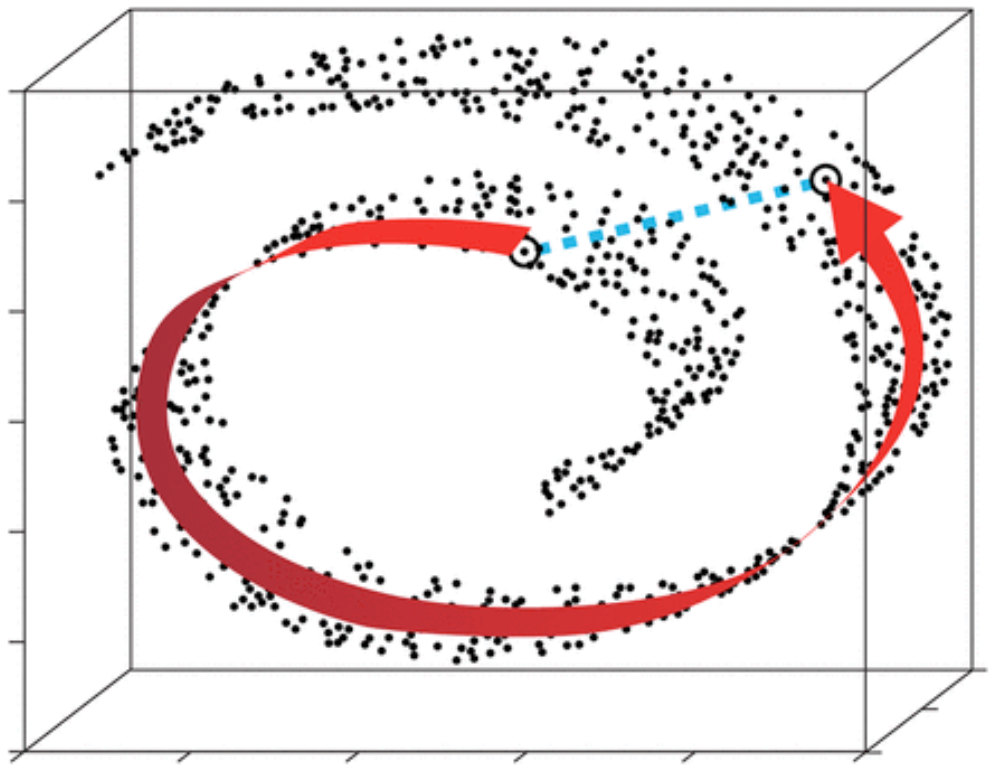


Introduced into EM by M. van Heel and J. Frank, Ultramicroscopy 1981
Figure from Schwander et al., "Conformations of Macromolecules and their Complexes from Heterogeneous Datasets." *Phil. Trans. B* 369.1647 (2014).



WIKI: “In mathematics, a manifold is a topological space that resembles Euclidean space near each point. More precisely, each point of an n-dimensional manifold has a neighborhood that is homeomorphic to the Euclidean space of dimension n.”

Riemann: “Mannigfaltigkeit” (German)



Conformationally heterogeneous molecules: <u>subset in any projection direction forms a manifold.</u>

Chacterization of manifold requires a mapping that follows the topography of the manifold.

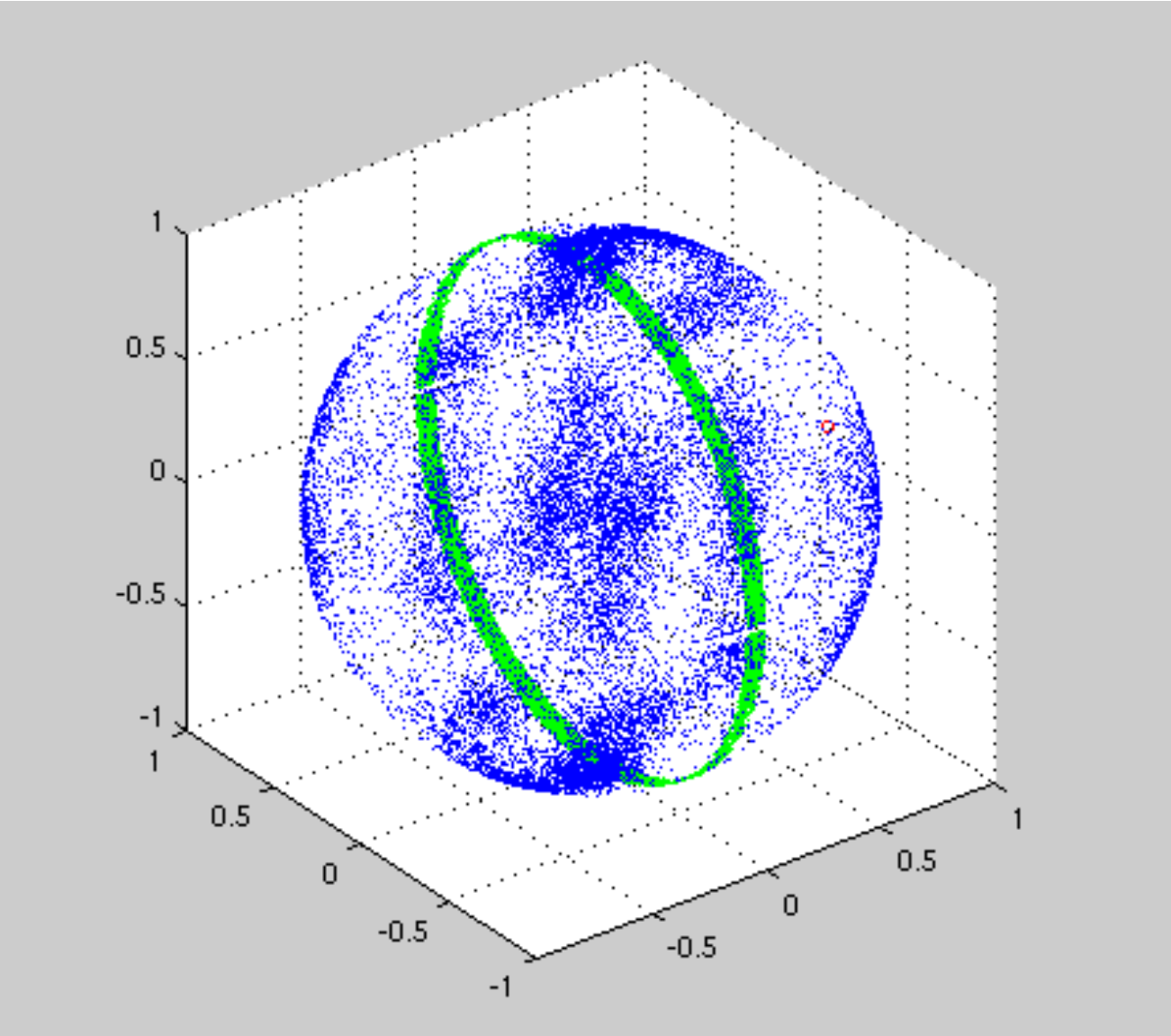
Eigen decomposition, factors with highest ranking.

Combine factorial representations (“equalization”) from all projection directions.

Yeast ribosomes purified from cell extract

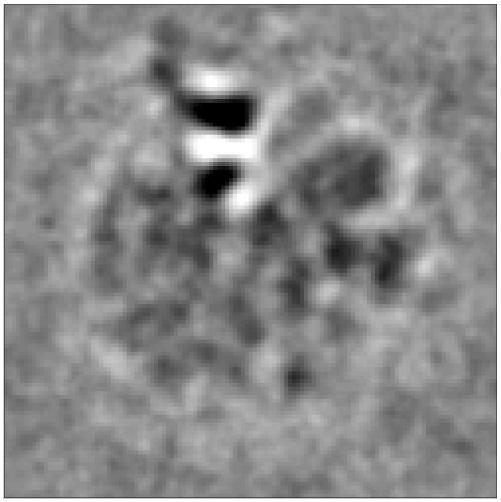
850,000 particles.

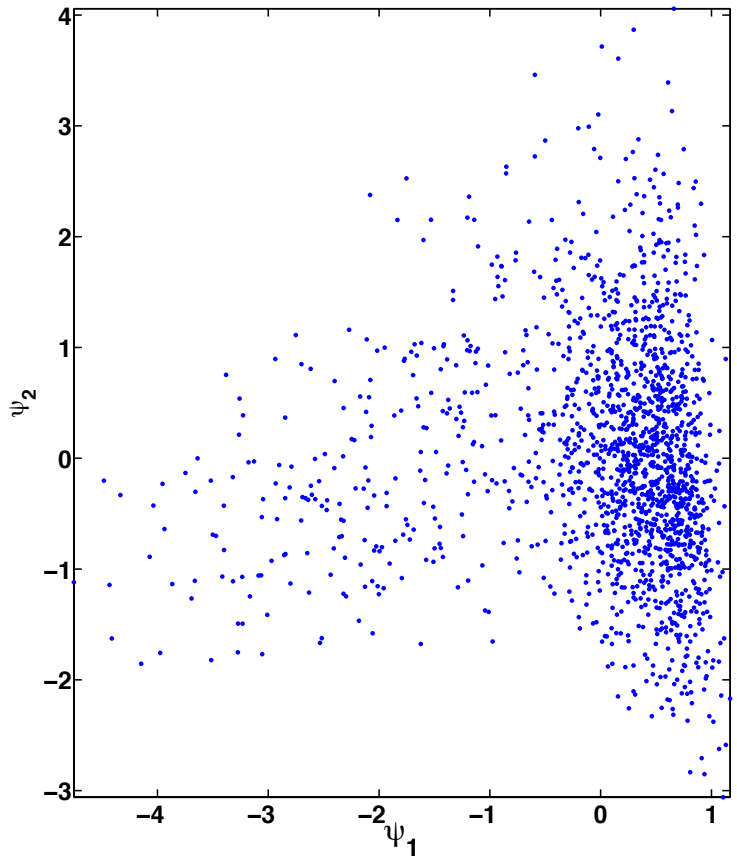
Selection of data along viewing directions along a great circle



Dashti et al., PNAS 2014

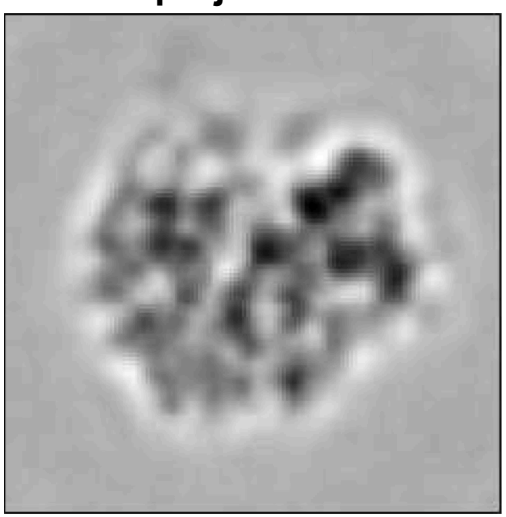
Topo2 along ψ_2

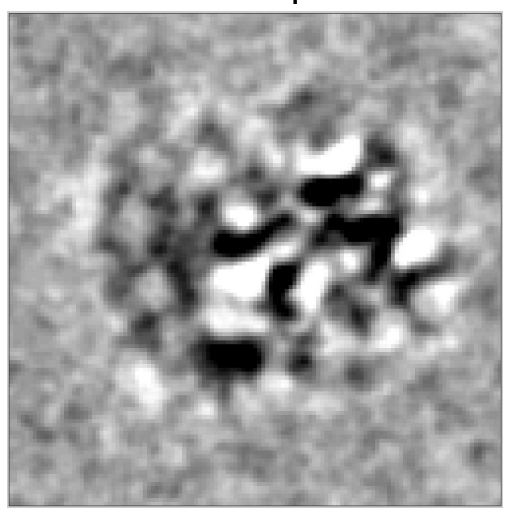




Projection's Euler angles:
Φ: 59.41
θ: 77.66
Ψ: 300.58

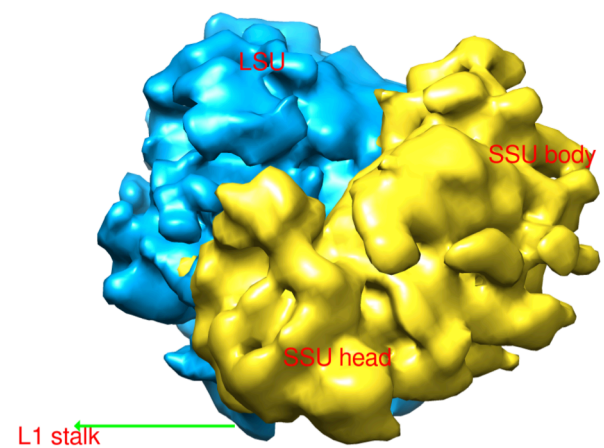
2D projection view





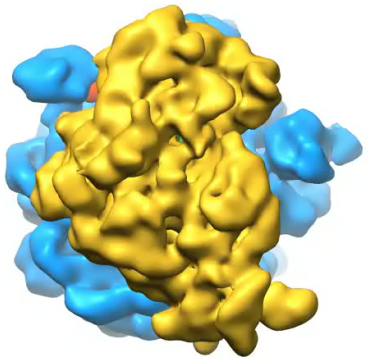
Topo2 along ψ_1

3D projection view



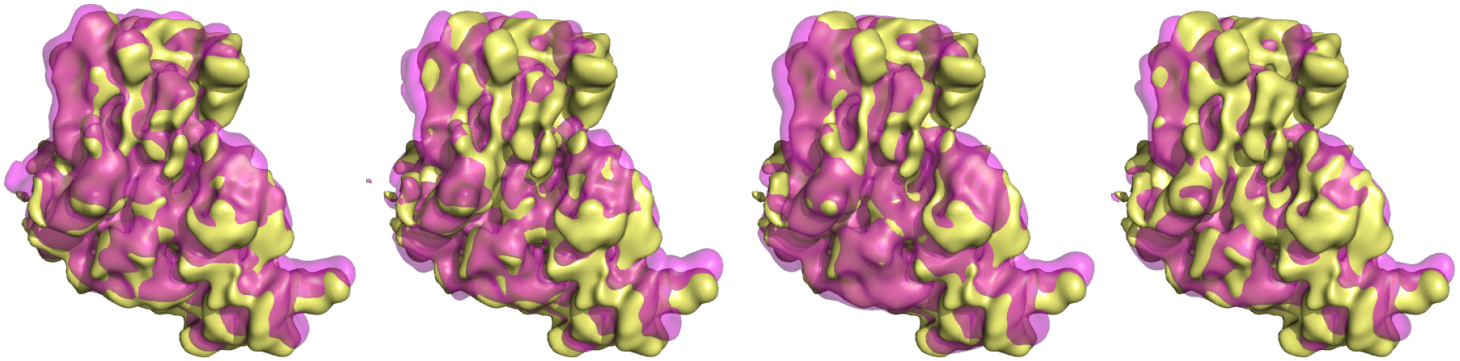
Dashti et al. PNAS 2014

Most prominent variatiation: intersubunit motion



Valle, M. *et al. Cell* 114, 123–134 (2003)

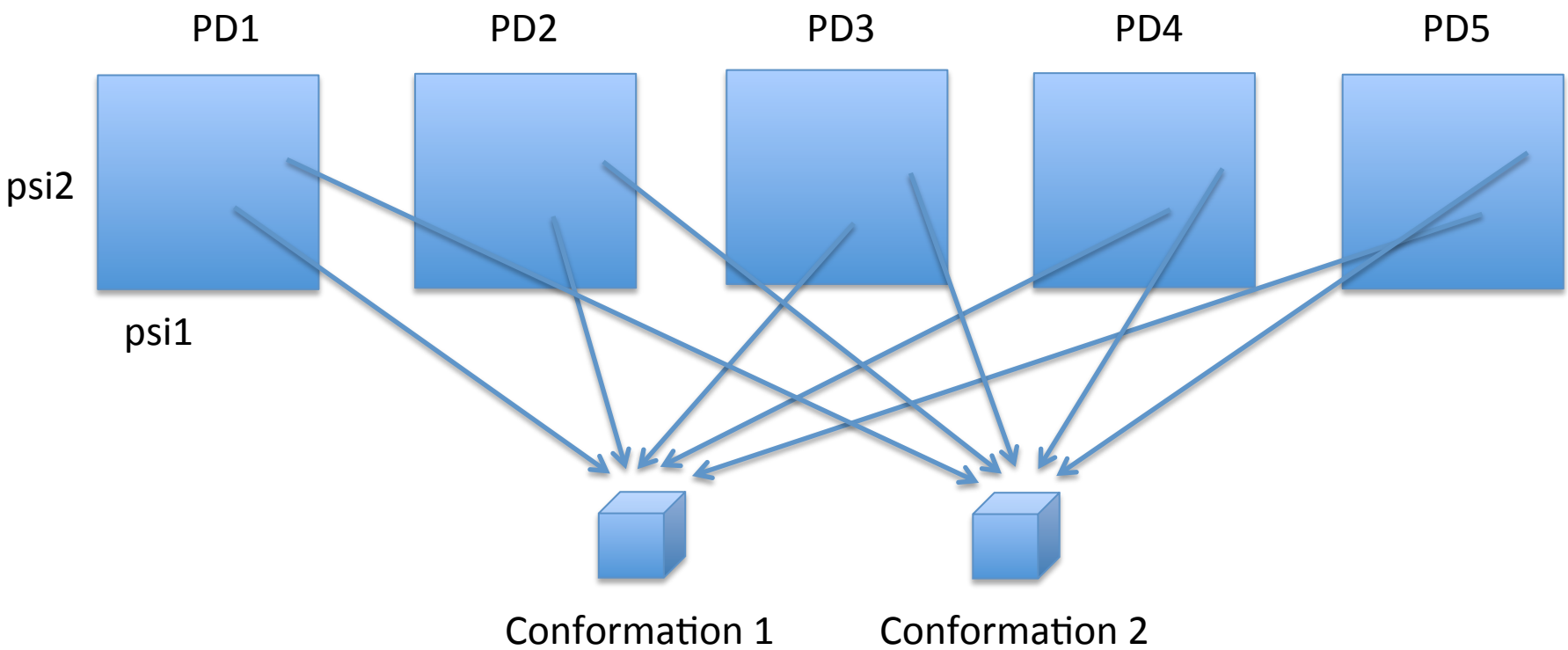
Ratchet-like motion of the small subunit relative to the large subunit, encountered in . . .



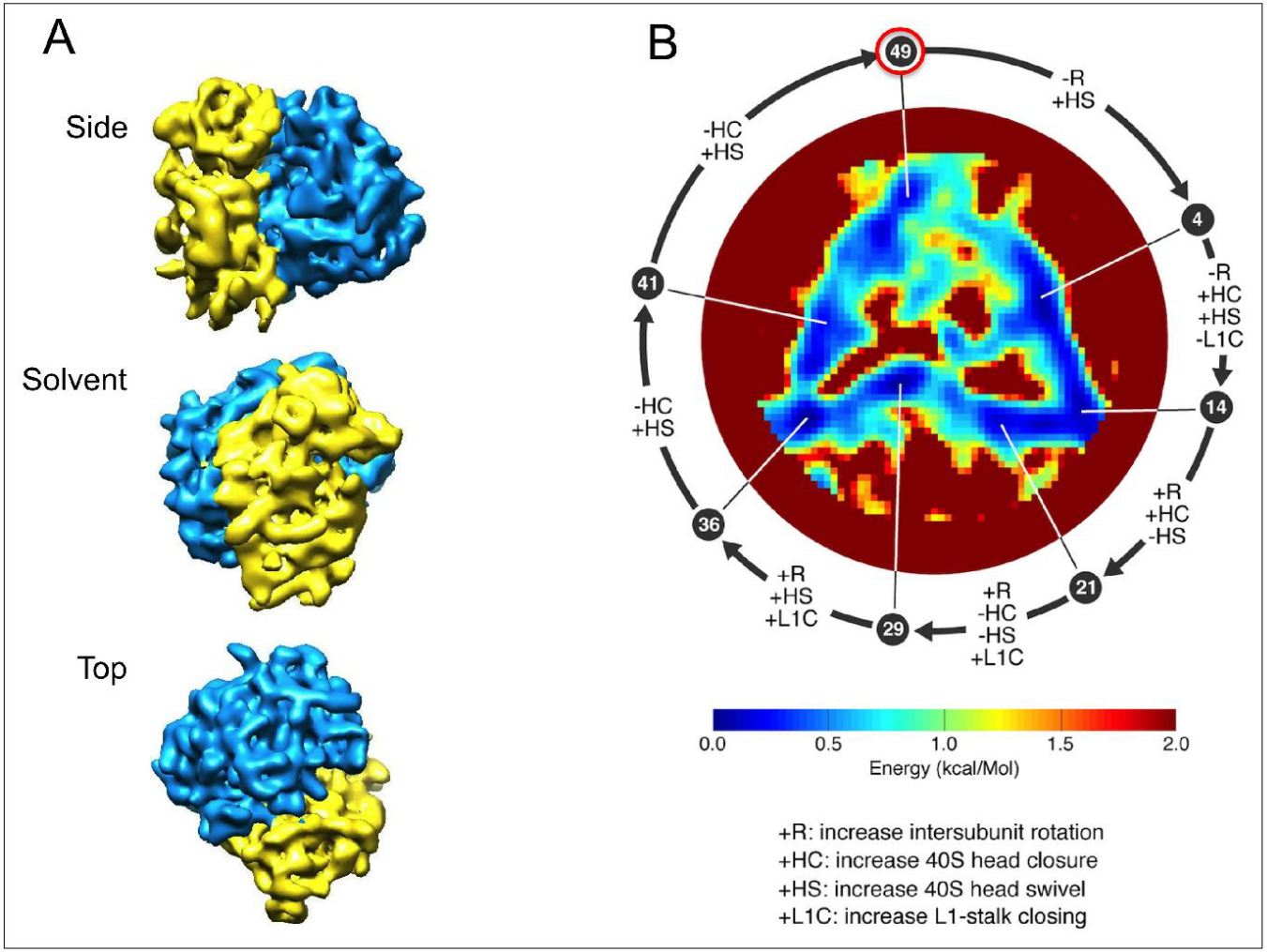
initiation, translocation, termination, and recycling

Retrieval of 3D structures

after coordinate systems {psi1, psi2, ...} have been equalized



One pixel in the psi1, psi2 coordinate system across all PDs represents the projections belonging to a single conformation. From these, the molecule presenting this conformation can be reconstructed.
(Here is an example where the molecule is reconstructed in 2 conformations)



Dashti et al. PNAS 2014

- Manifold embedding as applied to the calcium release channel – unpublished work, not shared

OUTLOOK

- Think about it! Water molecules! Who would have thought!
- Far from settled and complete, single-particle cryo-EM is a dynamic field which will have many surprises
- The ability to retrieve multiple conformations is a particular forte of single-particle reconstruction as it relates to the functional dynamics of a molecule
- Free-energy landscapes of molecular machines
- Titanic Centers *a la* Synchrotrons versus “table top” instruments affordable to individual labs?
- Automation of structure determination?

The Frank Lab

Bo **Chen**
Amedee **Des Georges**
Marcus **Fislage**
Jack **Fu**
Bob **Grassucci**
Dominique **Guiterrez**
Cristina **Gutierrez-Vargas**
Sandip **Kaledhonkar**
Harry **Kao**
Amy **Lobe**
Wen **Li**
Hstau **Liao**
Zheng **Liu**
Kelsey **Lynch**
Suvrajit **Maji**
Andrey **Malyutin**
Masgan **Saidi**
Nadia **Severina**
Bingxin **Shen**
Ming **Sun**
Melissa **Thomas**
Ed **Twomey**