Cryo-Electron Microscopy
(Single-Particle Cryogenic Electron Microscopy)
(Cryo-EM)

CS/CME/BioE/Biophys/BMI 279

Nov. 5 and 10, 2020

Ron Dror

THE REVOLUTION WILL NOT BE CRYSTALLIZED

MOVE OVER X-RAY CRYSTALLOGRAPHY. CRYO-ELECTRON MICROSCOPY IS KICKING UP A STORM IN STRUCTURAL BIOLOGY BY REVEALING THE HIDDEN MACHINERY OF THE CELL.

BY EWEN CALLAWAY

In a basement room, deep in the bowels of a steel-clad building in Cambridge, a major insurgency is under way.

A hulking metal box, some three metres tall, is quietly beaming terabytes' worth of data through thick orange cables that disappear off through the ceiling. It is one of the world's most advanced cryo-electron microscopes: a device that uses electron beams to photograph frozen biological molecules and lay bare their molecular shapes. The microscope is so sensitive that a shout can ruin an experiment, says Sjors Scheres, a structural biologist at the UK Medical Research Council Laboratory of Molecular Biology (LMB), as he stands dwarfed beside the £5-million (US$7.7-million) piece of equipment. "The UK needs many more of these, because there's going to be a boom," he predicts.

In labs around the world, cryo-electron microscopes such as this one are sending tremors through the field of structural biology. In the past three years, they have revealed exquisite details of protein-making ribosomes, quivering membrane proteins and other key cell molecules,

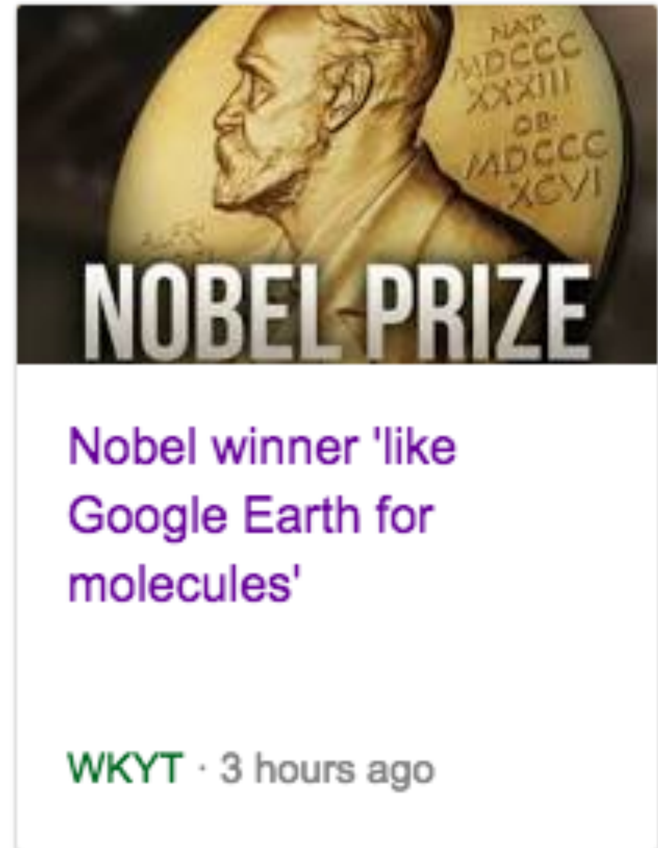
ILLUSTRATION BY VIKTOR KOEN

© 2015 Macmillan Publishers Limited. All rights reserved

Nature,
Sept. 10, 2015

2017 Nobel Prize in Chemistry

Awarded to Jacques Dubochet, Joachim Frank and Richard Henderson and "For developing cryo-electron microscopy for the high-resolution structure determination of biomolecules in solution"



Outline

- Overview of single-particle electron microscopy (EM)
- Cryo-EM images are *projections*
- Sample preparation
- Computational reconstruction methods
  - 2D image analysis
    - Image preprocessing
    - Particle picking
    - Image clustering and class averaging
  - 3D reconstruction
    - Reconstruction with known view angles
    - Structure refinement with unknown view angles
    - Calculating an initial structure
    - Fitting atomic-resolution models to lower-resolution EM structures

Overview of single-particle electron microscopy (EM)

The basic idea

- We want the structure of a "particle": a molecule (e.g., protein) or a well-defined complex composed of many molecules (e.g., a ribosome)
- We spread identical particles out on a film, and image them using an electron microscope
- The images are two-dimensional (2D), and each particle is positioned with a different, unknown orientation.
- Given enough 2D images of particles, we can computationally reconstruct the 3D shape of the particle

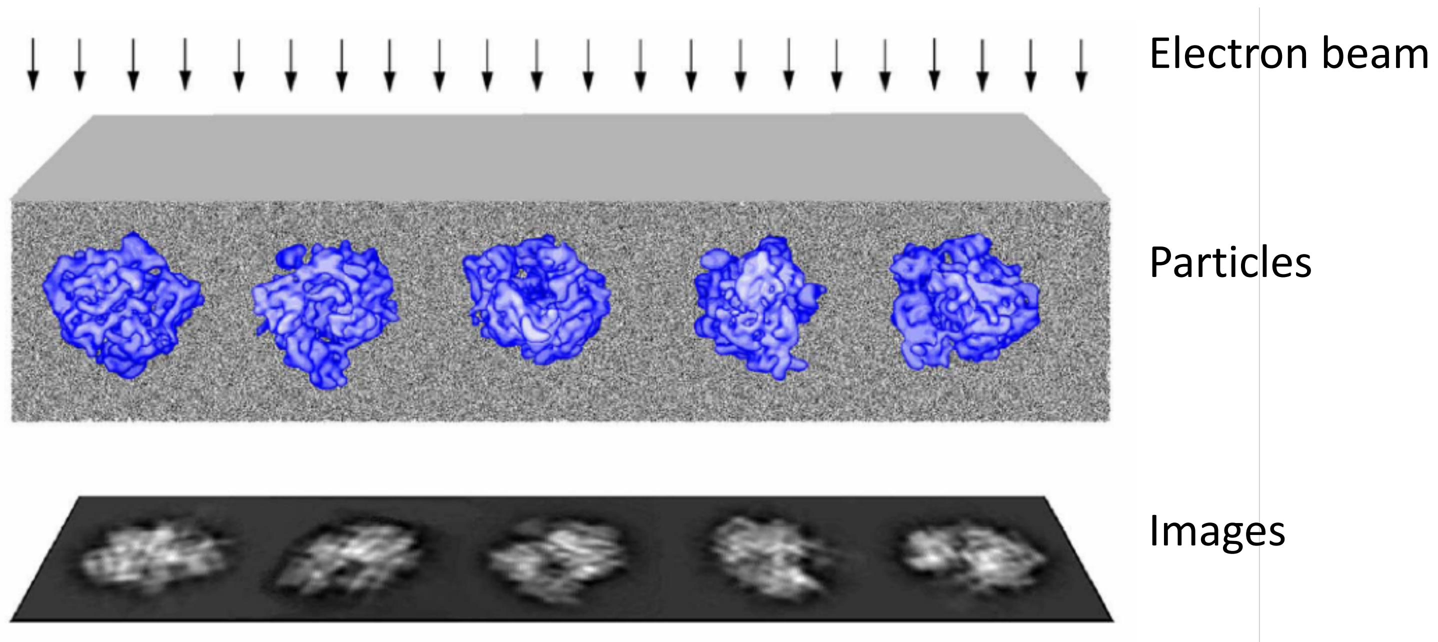


Image from Joachim Frank
http://biomachina.org/courses/structures/091.pdf

A high-end cryo-electron microscope

Dramatic recent improvements

- Single-particle EM has been around for decades, but it has improved *dramatically* in the last several years due to:
  - Invention of better cameras
    - Until recently, electrons were detected either by photographic film, or by scintillator-based digital cameras which converted electrons to photons for detection
    - New “direct-electron detectors” can detect electrons directly, substantially improving image resolution and quality
  - Better computational reconstruction techniques
- Single-particle cryo-EM is thus coming into much wider use, and may challenge crystallography as the dominant way to determine experimental structures

Comparison to x-ray crystallography

- Cryo-EM's major advantage over crystallography is that it does not require formation of a crystal
  - Particularly advantageous for large complexes, which are usually difficult to crystallize
  - Also avoids structural artifacts due to packing in a crystal lattice. In EM, particles are in a more natural environment.
- On the other hand:
  - Cryo-EM's resolution is (typically) lower than that of crystallography
  - Reconstructing structures of small proteins from EM images is difficult, because images from different orientations look similar (i.e., "a blob")

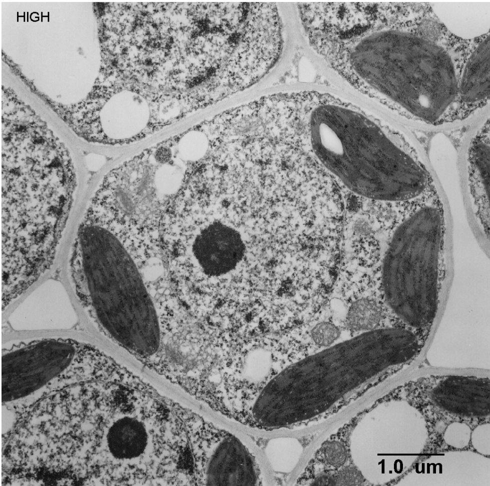
- Bottom line: Cryo-EM is particularly advantageous for large complexes, because:
  - Large complexes tend to be harder to crystallize
  - The computational reconstruction problem in cryo-EM is usually easier to solve for large particles than for small ones

Cryo-EM images are *projections*

Cryo-EM uses *transmission* electron microscopy

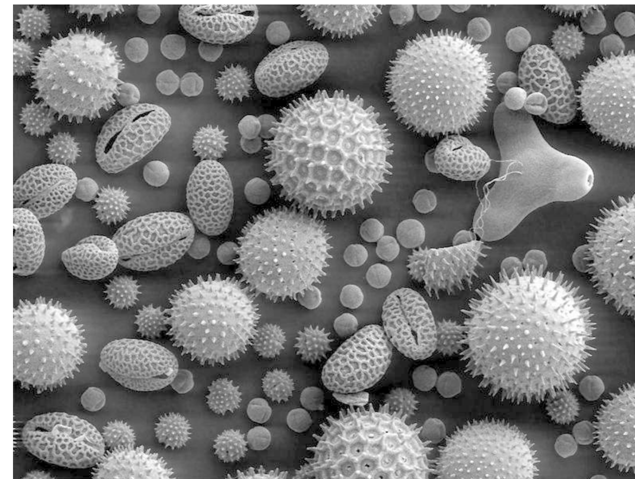
- In transmission electron microscopy, a beam of electrons pass through a thin sample before forming an image

Transmission electron microscopy



http://www.cas.miamioh.edu/~meicenrd/ANATOMY/Ch2_Ultrastructure/Tempcell.htm

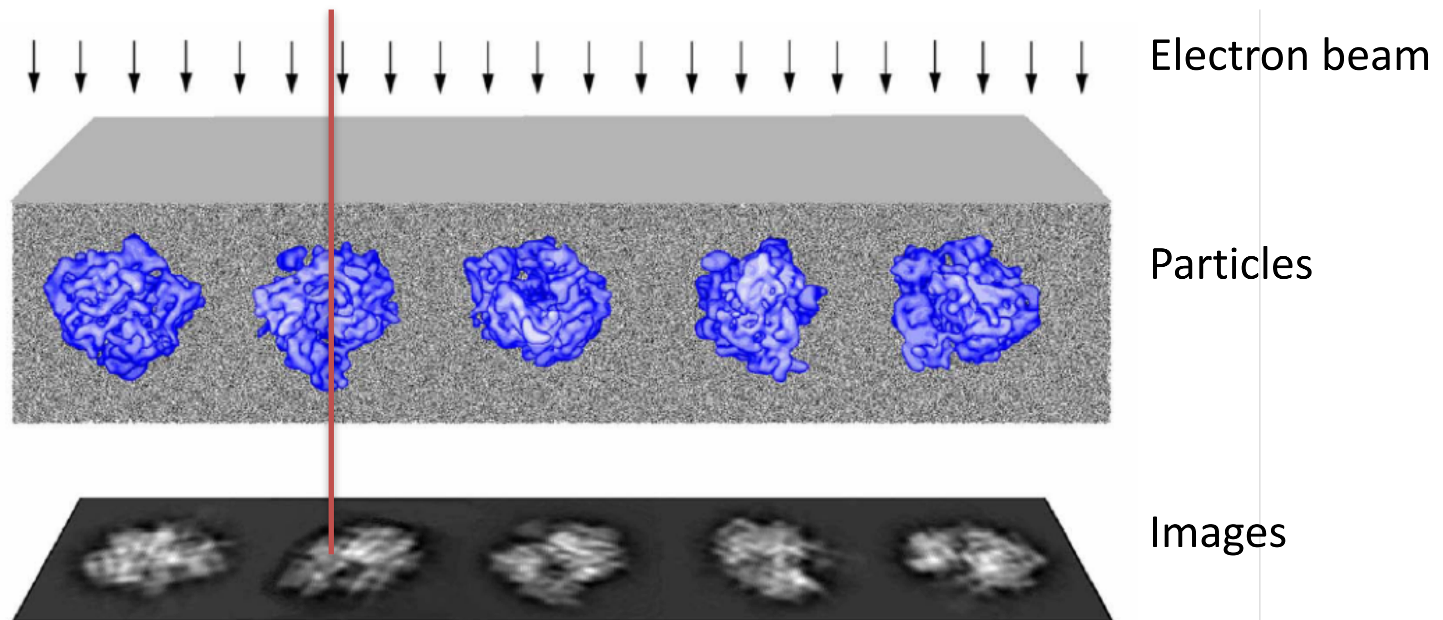
Scanning electron microsopy

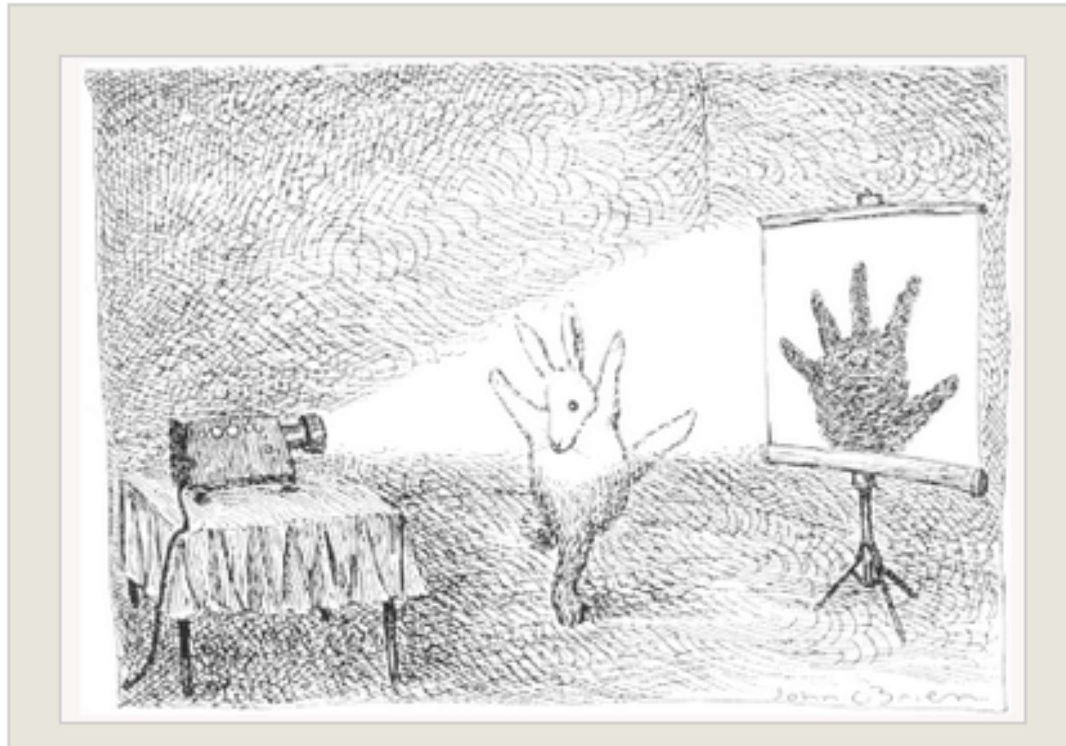


http://www.newscientist.com/data/images/ns/cms/dn14136/dn14136-1_788.jpg

Cryo-EM images are projections

- Each recorded 2D image is thus a projection of the 3D shape (density) we want to reconstruct
  – That is, we can think of each pixel value in the 2D image as a sum of the values along a line through the 3D sample (in the direction of the electron beam)





From Joachim Frank, *Three-dimensional electron microscopy of macromolecular assemblies: Visualization of biological molecules in their native state*, 2006

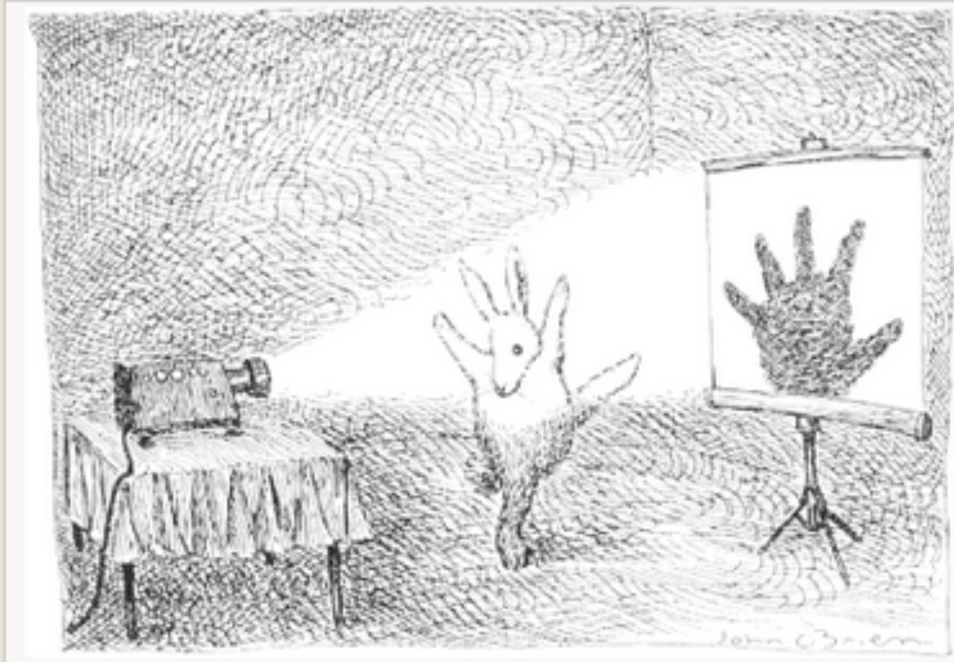


Figure 5.1 A single projection image is plainly insufficient to infer the structure of an object. (Note, though, that TEM projections do not merely give the outline, as in this drawing, but internal features, too—the bones and internal organs of the rabbit, which we would see if the projector were to emit X-rays.) (Drawing by John O'Brien; © 1991 *The New Yorker*.)

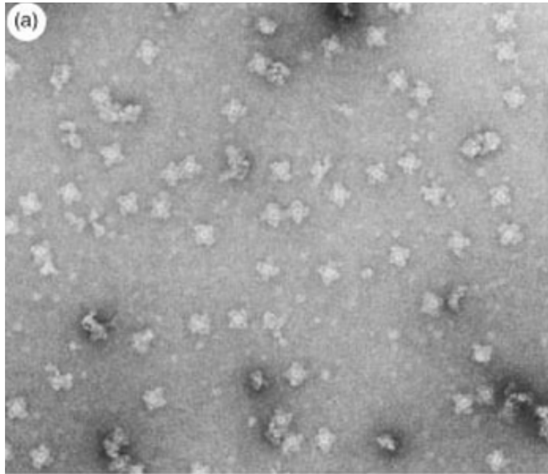
In transmission EM, the image would look more like an "x-ray" of the bunny than a shadow of the bunny

From Joachim Frank, *Three-dimensional electron microscopy of macromolecular assemblies: Visualization of biological molecules in their native state*, 2006

Sample preparation

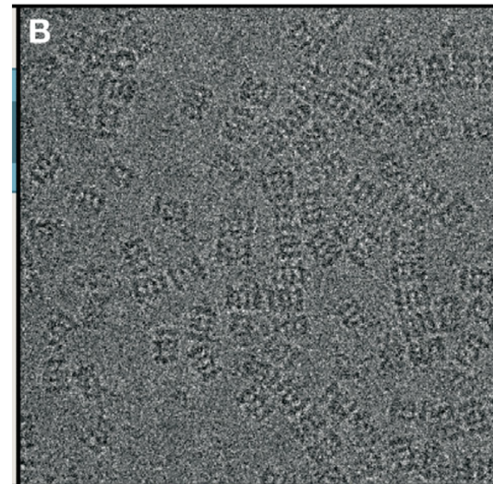
Sample preparation

- To survive in the electron microscope (in a vacuum, under electron bombardment), particles are usually prepared in one of two ways:
  - Negative staining
    - Coat particles with heavy metal salt crystals
    - This increases contrast (particles are easy to pick out from background)
    - It limits resolution to ~20 Å and can introduce artifacts
  - Vitrification
    - Particles are embedded in ice (vitreous ice: flash frozen, not crystalline)
    - This gives less contrast, but enables much higher resolution (below 4 Å)
    - High-resolution single-particle EM relies on vitrification and is thus referred to as cryo-electron microscopy (cryo-EM)



Negative stain.
Particles are easy to pick out.

Frank, 2006



Vitreous ice.
Particles are harder to pick out, even though this is a very good ("easy") case.

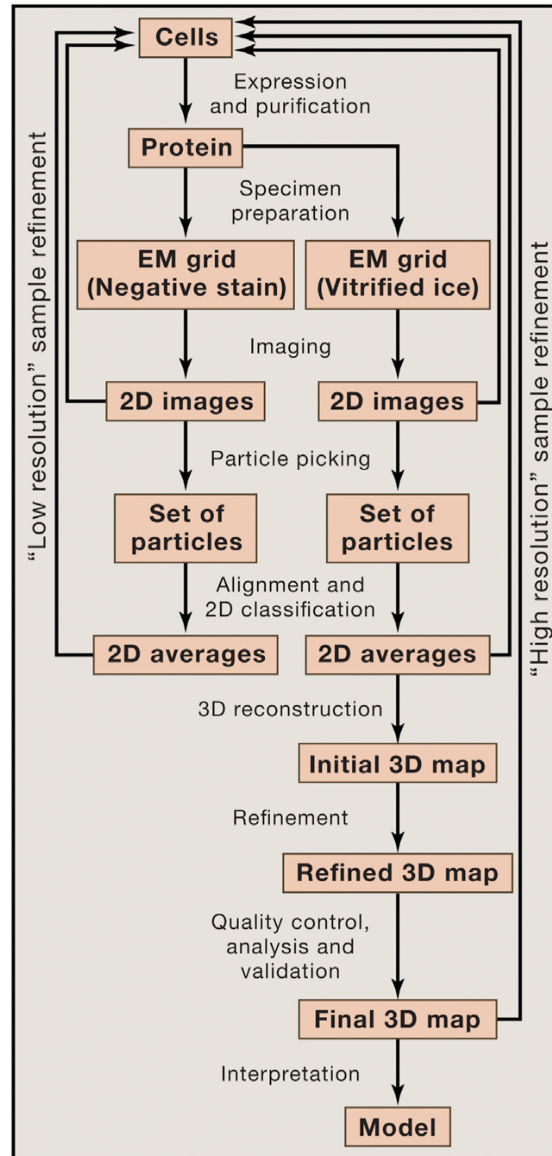
Cheng, *Cell* 161:450 (2015)

Computational reconstruction methods

Overview of computational methods

- **2D image analysis**: First, go from raw image data to higher-resolution 2D projections
  - Image preprocessing
  - Particle picking
  - Image clustering and class averaging
- **3D reconstruction**: Then use these higher-resolution projections to build a 3D model
  - Background: Reconstruction with known view angles
  - Structure refinement with unknown view angles
  - Calculating an initial structure
  - Fitting atomic-resolution models to lower-resolution EM structures

Overview of computational methods



Cheng et al., *Cell* 161:438 (2015)

Computational reconstruction methods

2D image analysis

The raw images don’t look so good

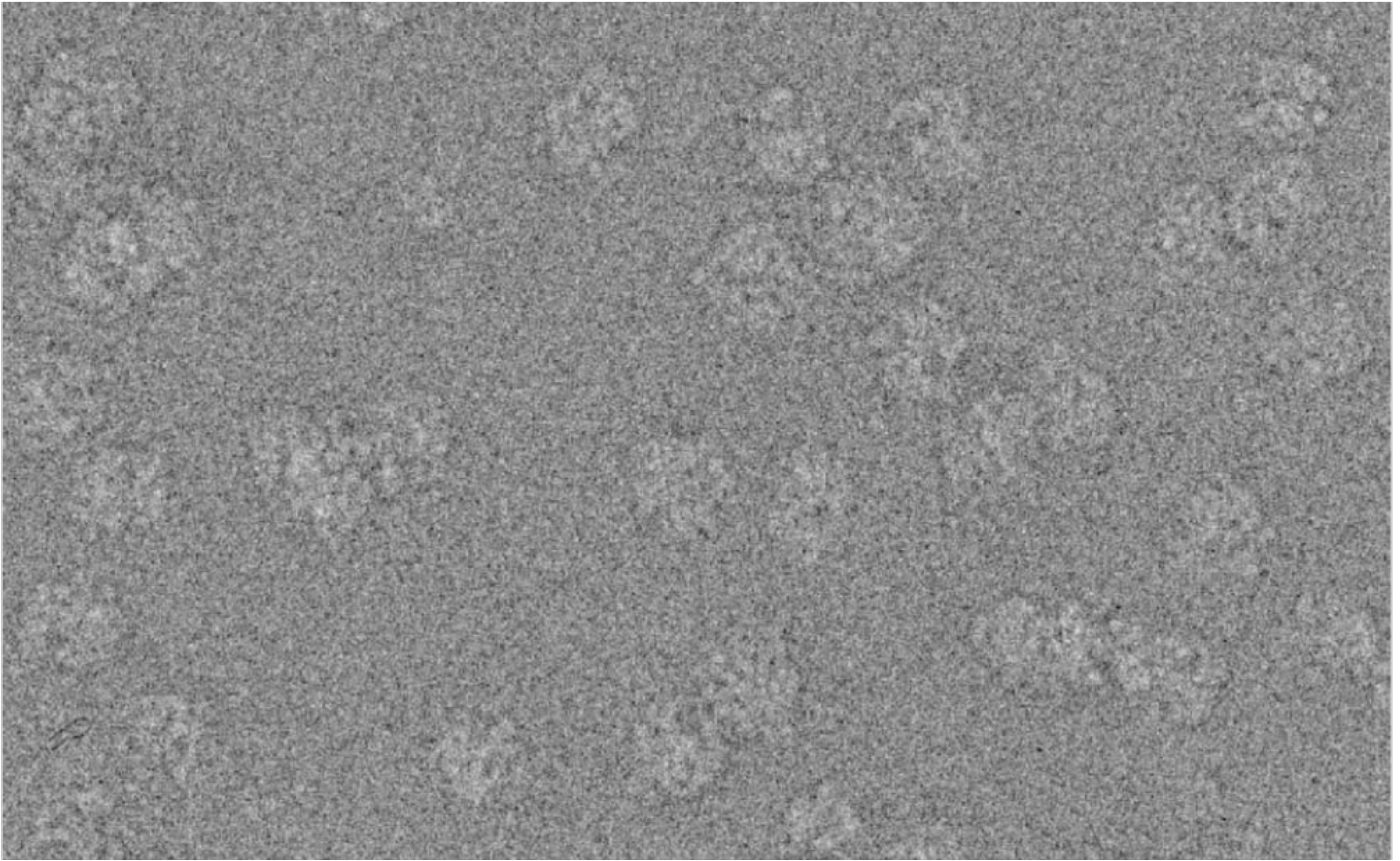


Image from Joachim Frank
http://biomachina.org/courses/structures/091.pdf

Before attempting any 3D reconstruction, we do several types of processing on the images

Computational reconstruction methods

2D image analysis

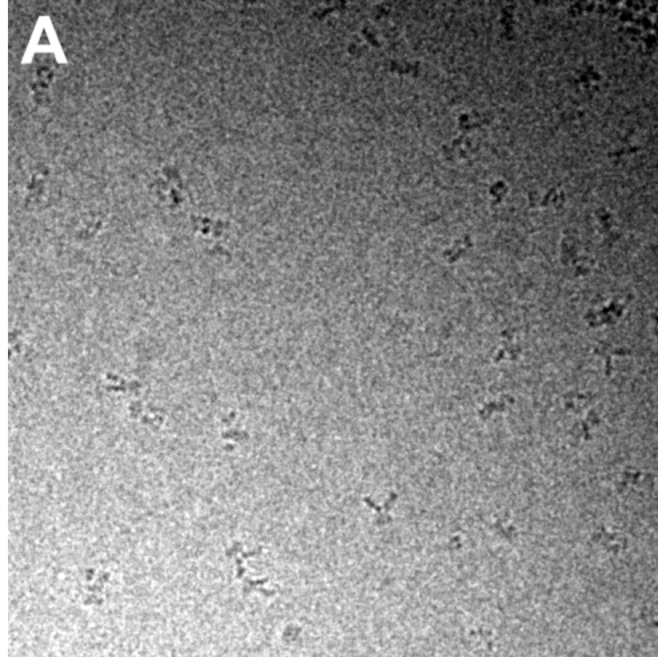
Image preprocessing

Image preprocessing

- **Problem 1**: The sample tends to move slightly during imaging, blurring the image
- **Solution**
  - Direct electron detectors are fast enough to record a movie instead of a single image
  - Align the movie frames computationally, then average them together

Image preprocessing

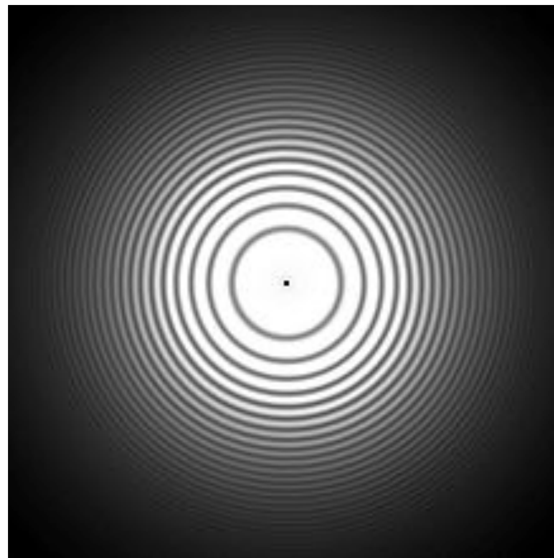
- **Problem 2**: Overall brightness is often non-uniform (due to uneven illumination or sample thickness)
- **Solution**: high-pass filter the image



Cheng et al., *Cell* 161:438 (2015)

Image preprocessing

- **Problem 3**: The optics cause the recorded image to be a blurred version of the ideal image
  - This blurring is a convolution, and can thus be expressed as a multiplication in the frequency domain, where the ideal image is multiplied by the "contrast transfer function"
- **Solution**: Estimate parameters of the contrast transfer function, then correct for it
  - Some of the parameters are known (from the optics), while others are estimated from the images
  - Correction is generally done in the frequency domain



https://en.wikipedia.org/wiki/Contrast_transfer_function

A typical contrast transfer function, in the frequency domain (zero frequency at the center)

You're not responsible for the particular form of the contrast transfer function

Computational reconstruction methods

2D image analysis

Particle picking

Pick out the particles in the 2D images

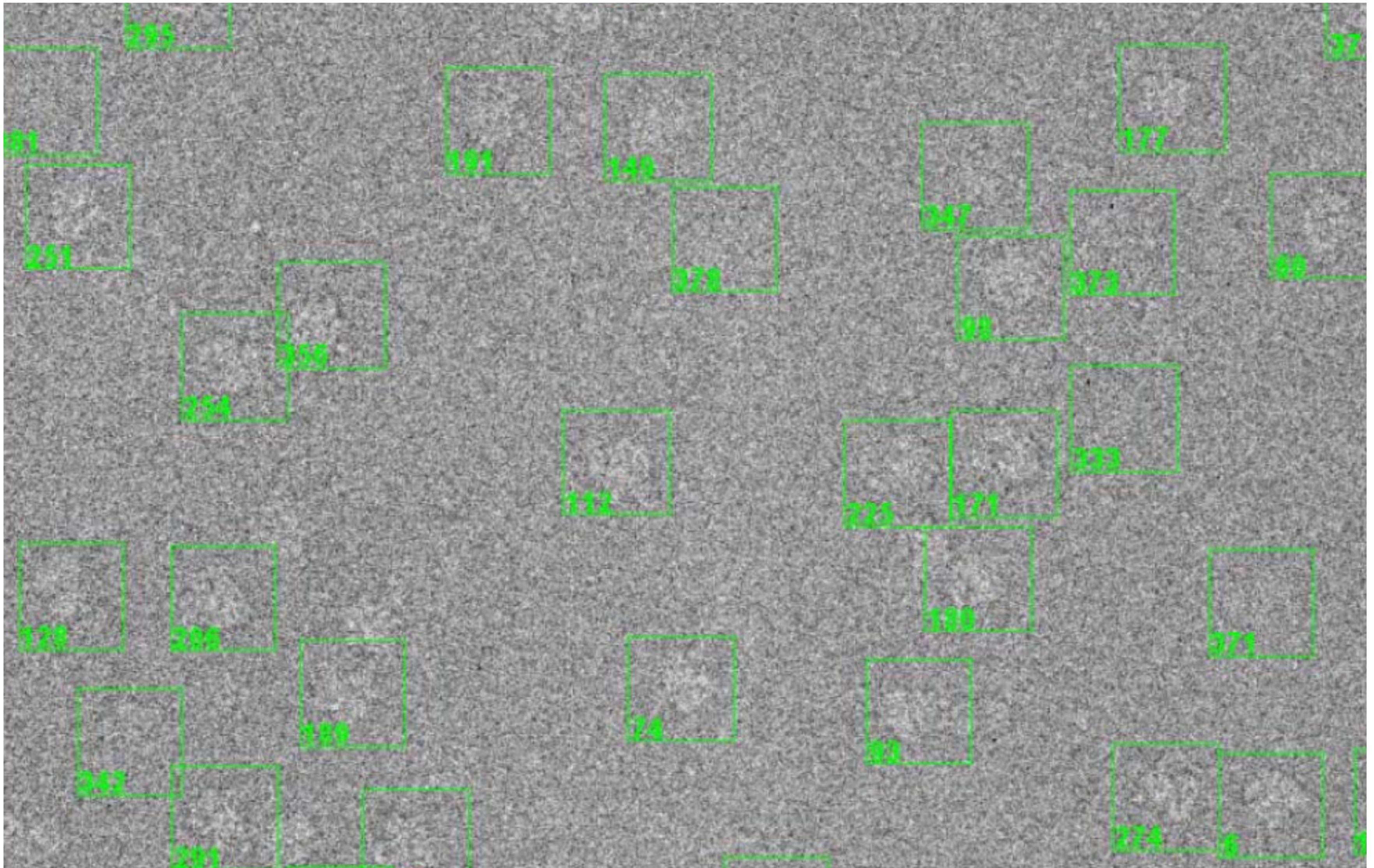


Image from Joachim Frank
http://biomachina.org/courses/structures/091.pdf

Particle picking results

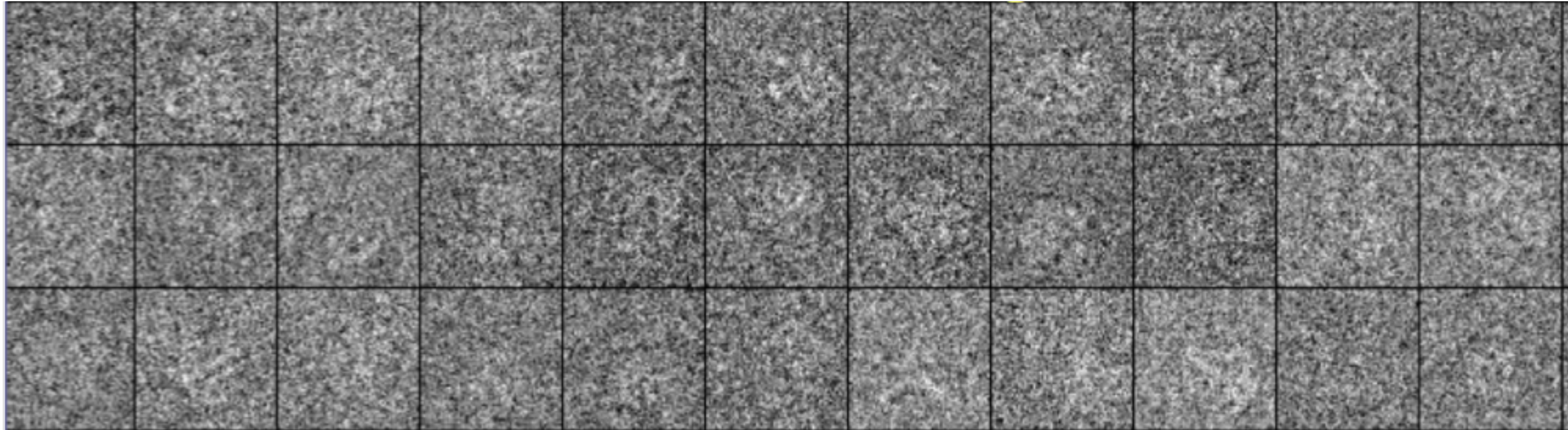
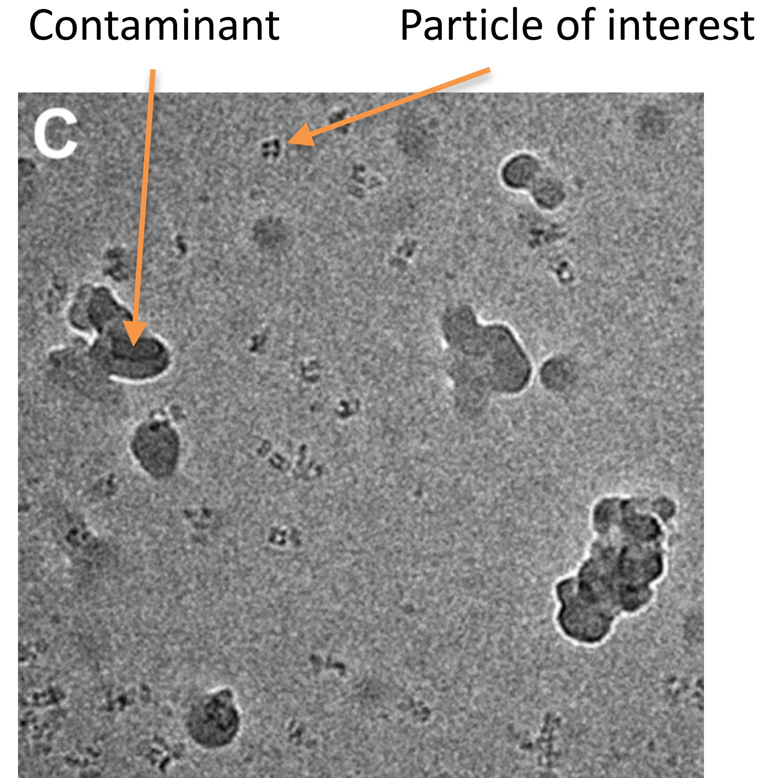


Image from Joachim Frank
http://biomachina.org/courses/structures/091.pdf

Particle picking methods

- Particle picking can be difficult, because the images are low-contrast and noisy
  - Images may also have contaminants that should be ignored
- A variety of automated and semi-automated methods have been developed
  - For example, matching to templates, or picking out high-contrast regions
- Often this is still done manually, at least to seed automated methods with suitable templates



Cheng et al., *Cell* 2015

Computational reconstruction methods

2D image analysis

Image clustering and class averaging

Averaging similar images reduces noise

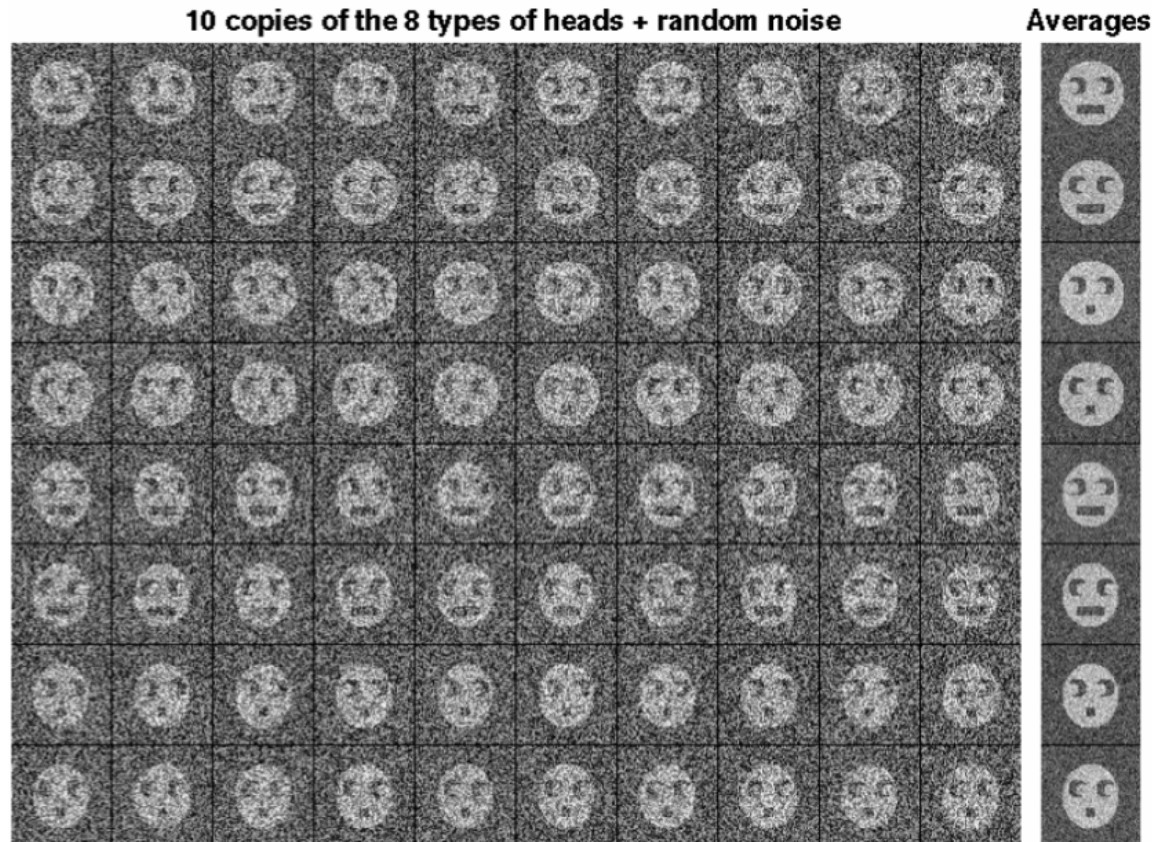


Image from Joachim Frank
http://biomachina.org/courses/structures/091.pdf

- The images in each row above represent the same ideal image but with different corrupting noise
- If we average the images in each row (that is, average corresponding pixels), we end up with a less noisy image, because the noise in the different images tends to cancel out

Goal: cluster the particle images into classes of similar images

- Group together images with similar view angles
  - Then align them to one another and average them together to reduce noise
- To do this, divide images into several classes (with each class representing a set of similar view angles)
- We need to determine both what the classes are and which images should be assigned to each class
- This is a *clustering* problem
  - Group images such that the images within a group are similar, but images in different groups are different
  - In machine learning terminology, this is “unsupervised learning”

Standard approach: *k*-means clustering

- Pick *k* random images as class exemplars
- Then iterate the following:
  - Assign each image to the closest exemplar
  - Average all the images in each class to determine a new class exemplar

- Notes:
  - In the assignment step, we need to align each particle image against the exemplar images
  - We need to specify the number of classes (*k*) in advance, or experiment with different values of *k*
  - *k*-means clustering is guaranteed to converge, but not guaranteed to find a globally optimal solution
  - Indeed, the solution may depend heavily on the initialization conditions, and may be heavily suboptimal

Caveat: Potential model bias in clustering/alignment

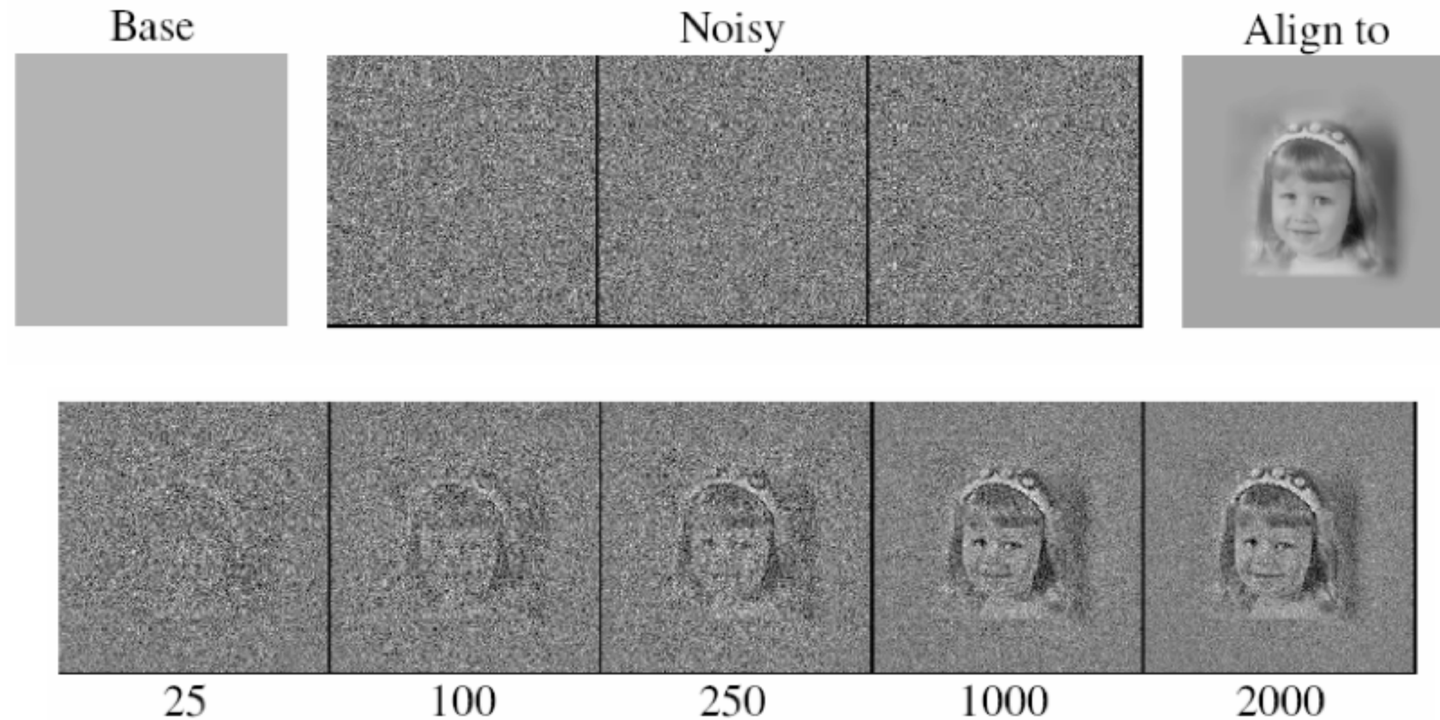


Image from Steve Ludtke
http://biomachina.org/courses/structures/091.pdf

In this case, the images are just noise, but by selecting images and alignments that best match a given template, we get a class average that looks like the template.

Caveat: Potential model bias in clustering/alignment

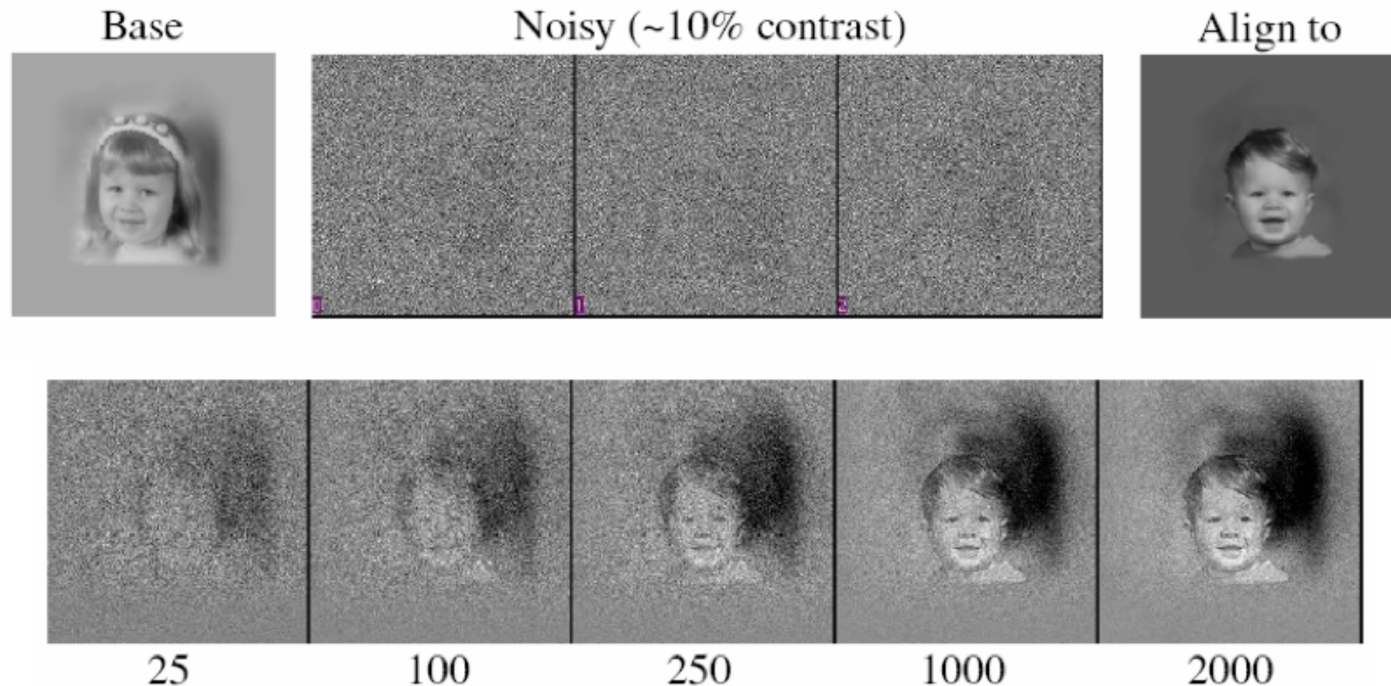


Image from Steve Ludtke
http://biomachina.org/courses/structures/091.pdf

In this case, the images are noisy versions of one face, but by selecting images and alignments that best match a second face, we get a class average that looks like the second face.

Avoiding these problems

- A variety of more sophisticated clustering methods ameliorate these problems
  - Some involve modifications to k-means (including the recently developed Iterative Stable Alignment and Clustering method)
  - Some involve principal components analysis or other dimensionality reduction techniques

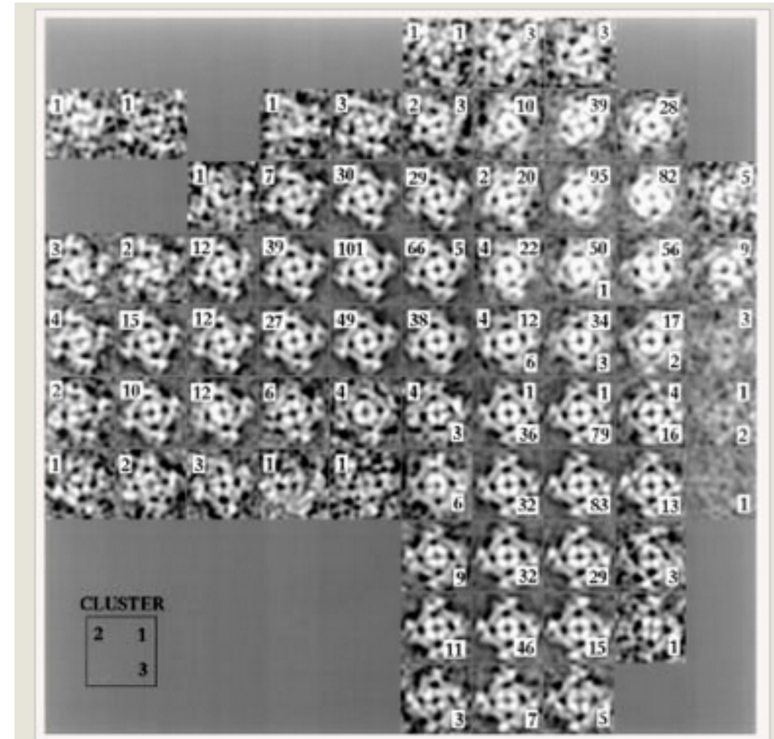
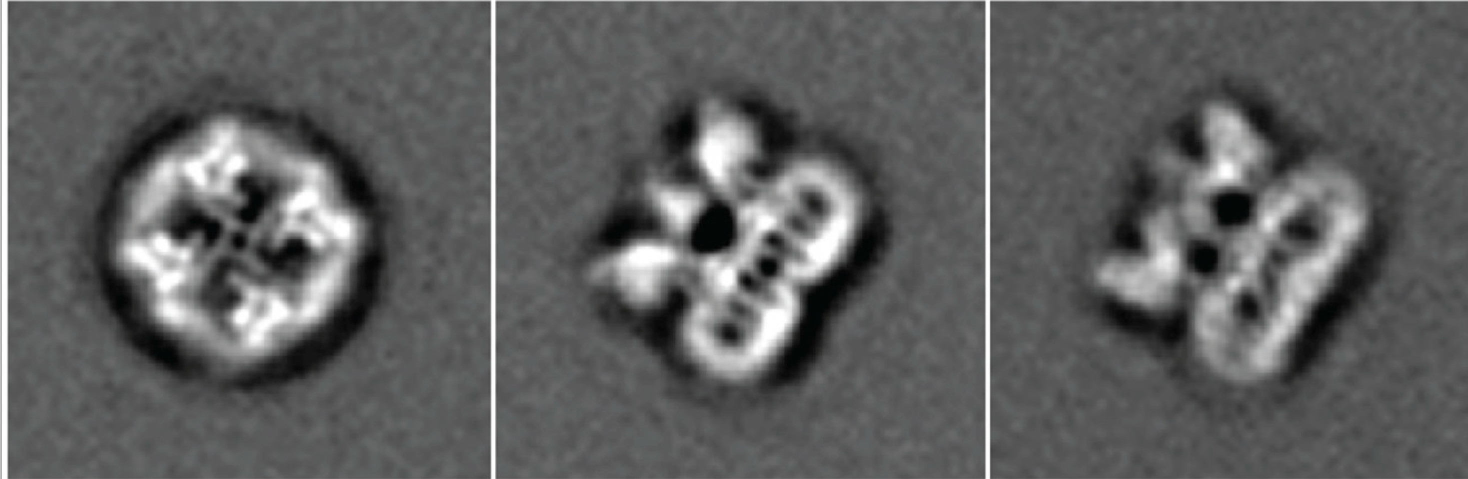


Figure 4.14 Checkerboard display of local averages (calcium release channel), each computed from images falling on a grid in factor space (factors 1 versus 2). The number on top of each average indicates the number of images falling into that grid space. Empty regions are bare of images. The distinction of main interest is between molecules lying in different orientations, related by flipping. It is seen from the peripheral pinwheel features pointing either clockwise (on the left) or counterclockwise (on the right). From Frank et al. (1996), reproduced with permission of Elsevier.

You're not responsible for these methods

Class averaging results



Cheng, *Cell* 161:450 (2015)

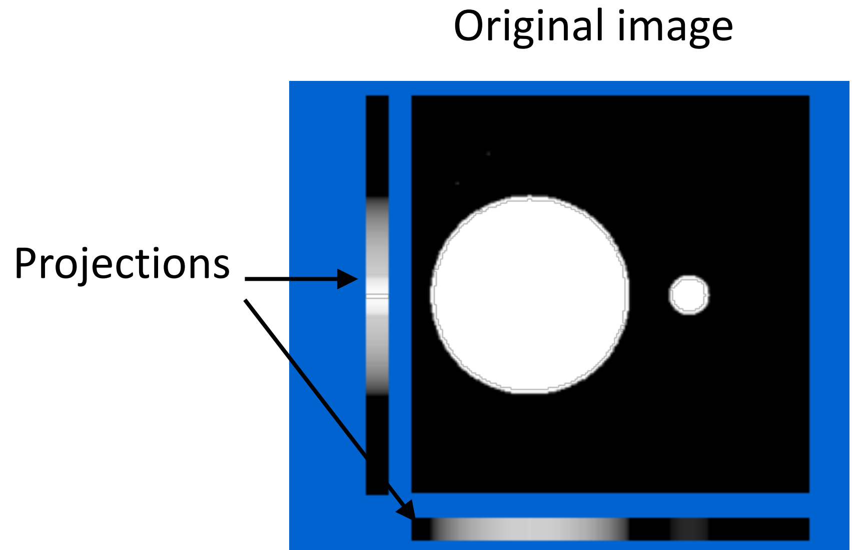
These are considered good class averages
(from a high-resolution single-particle EM study)

Computational reconstruction methods

3D reconstruction

Problem

- Suppose you're given many projections of a 2D image, and you want to reconstruct the original image. How would you do it?



- How would you do it if you know the view angle for each projection?
- How would you do it if you *don't* know the view angle for each projection (as is the case for cryo-EM)?

Computational reconstruction methods

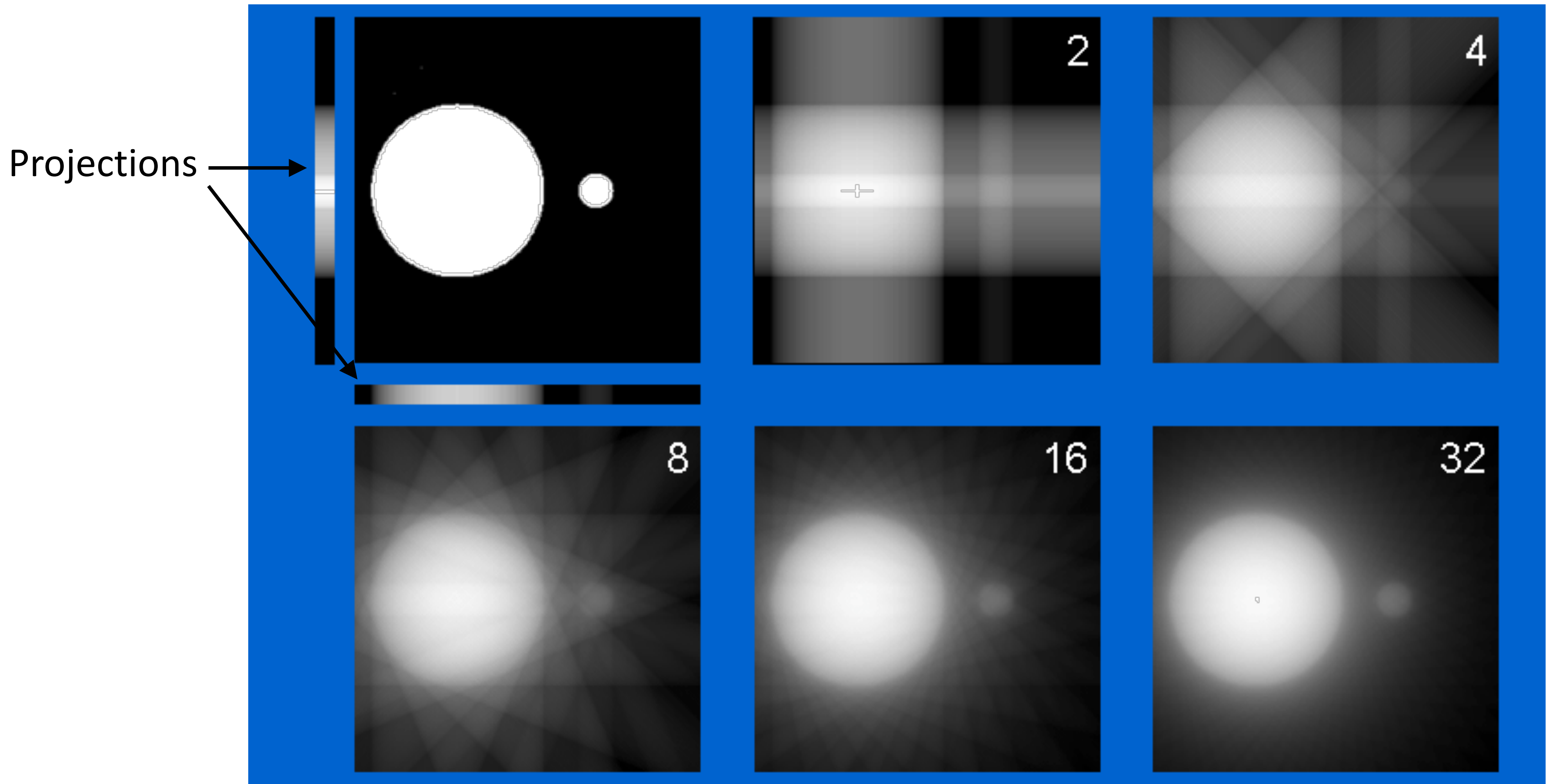
3D reconstruction

Background: Reconstruction with known view angles

Suppose you knew the view angle for each particle image

- How would you reconstruct the 3D density map from 2D projections?
  - Same problem is encountered in medical imaging (e.g. in CT scans, which are basically 3D x-rays)
- The simplest approach would be *back-projection*: reverse the projection process by “smearing” each projection back across the reconstructed image

Back-projection

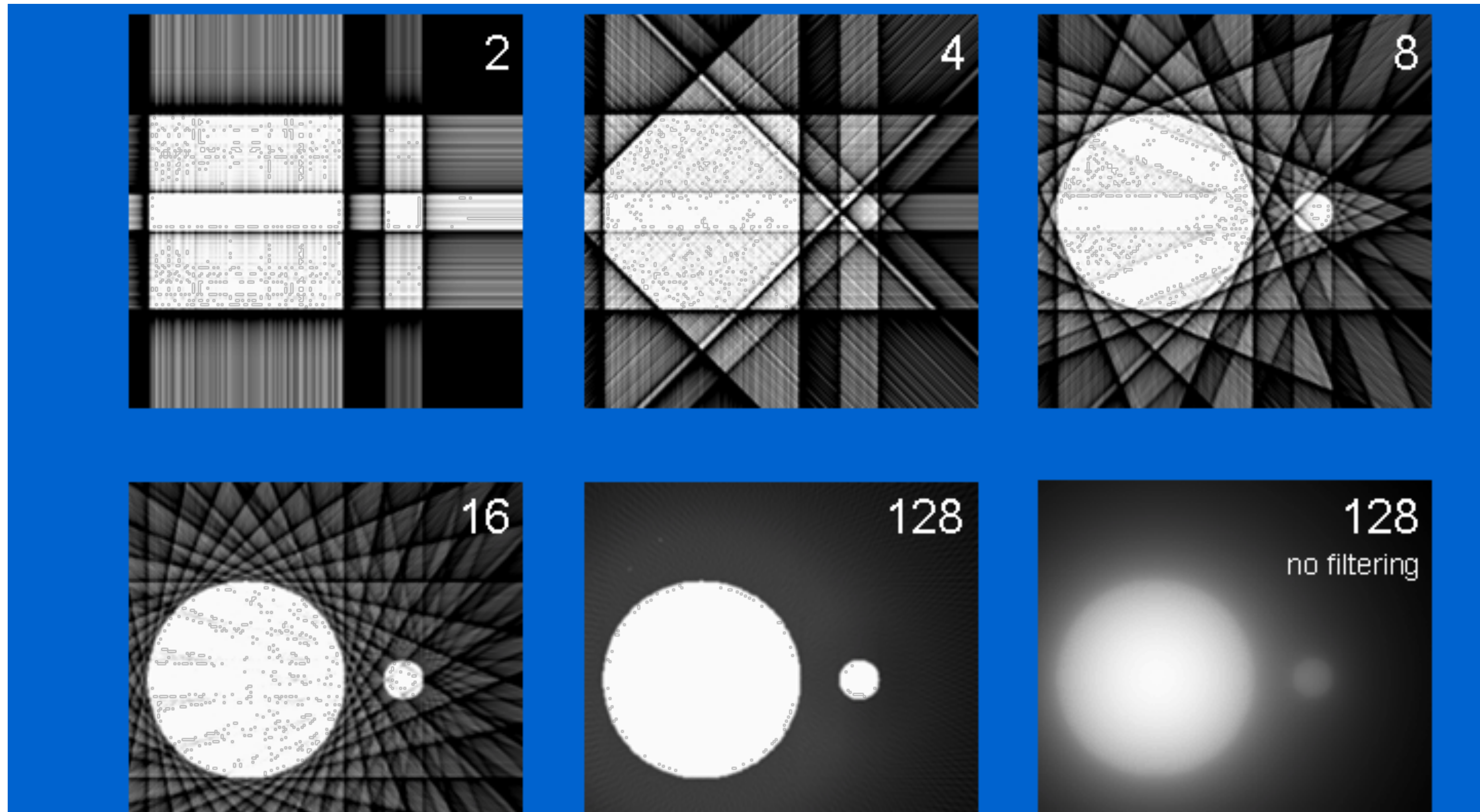


http://www.impactscan.org/slides/impactcourse/basic_principles_of_ct/img12.html

The result of back-projection is a *blurred* version of the original image.
How can we fix this?

Filtered-back projection

- It turns out we can fix this problem by applying a specific high-pass filter to each image before back-projection. This is *filtered back-projection.*

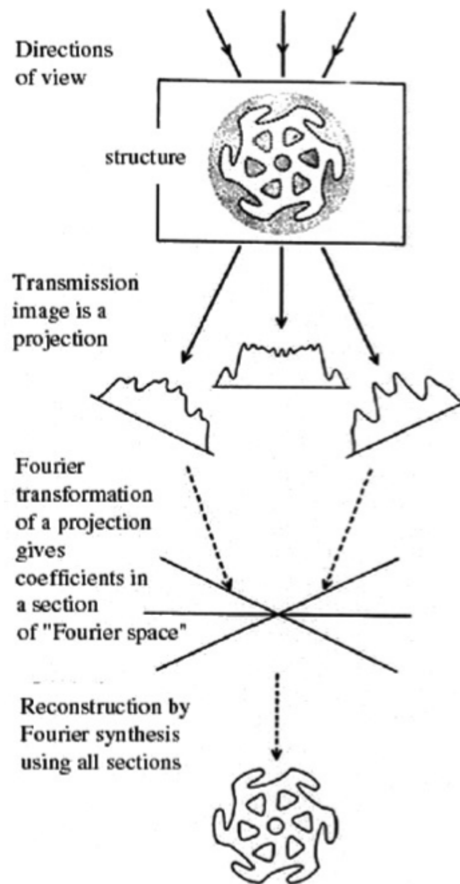


http://www.impactscan.org/slides/impactcourse/basic_principles_of_ct/img15.gif

Why does filtered back-projection work?

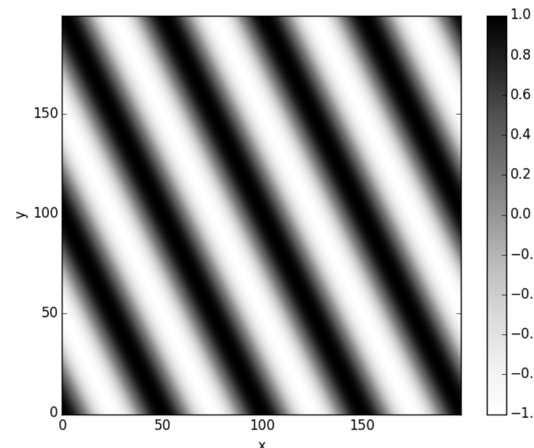
You're not responsible for this

- To answer this, use the *projection slice theorem*



Projection slice theorem (2D version):
The 1D Fourier transform of the 1D projection of a 2D density is equal to the central section—perpendicular to the direction of projection—of the 2D Fourier transform of the density

This theorem holds because each of the 2D sinusoids used in the 2D Fourier transform is constant in one direction



Why does filtered back-projection work?

You're not responsible for this

- Back-projection is equivalent to filling in central sections in the Fourier domain
- The problem is that when reconstructing by back-projection, we overweight the low-frequency values (in the figure, the density of dots is greatest near the center)
- To fix this, reduce the weights on low-frequency components.

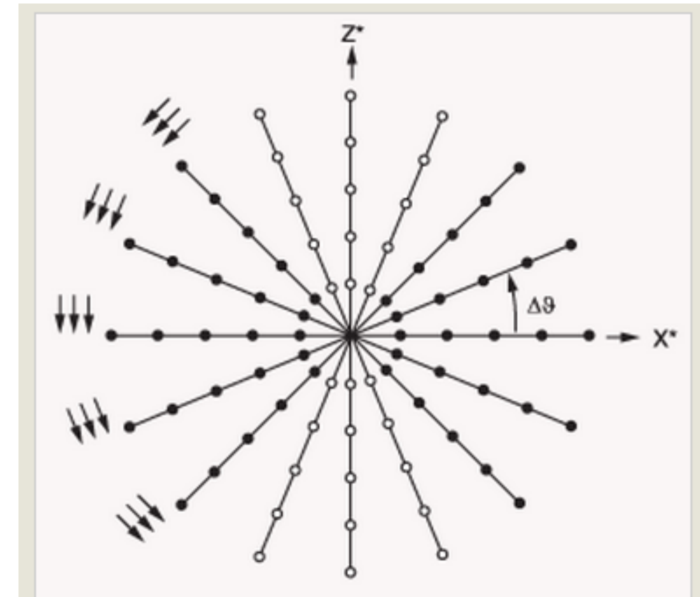
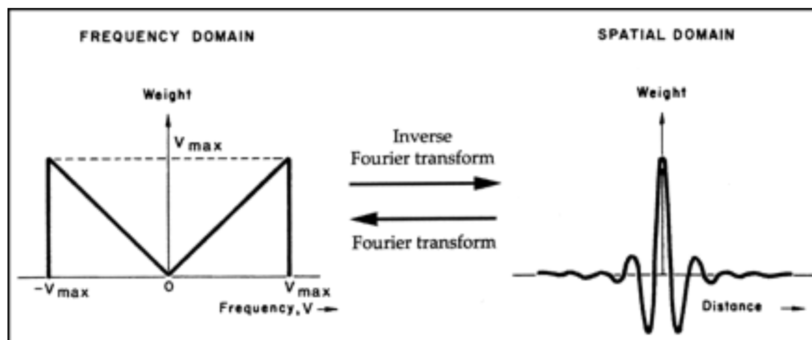


Figure 5.15 The density of sampling points in Fourier space obtained by projections decreases with increasing spatial frequency. Although this is shown here for single-axis tilting, the same is obviously true for all other data collection geometries. From Frank and Radermacher (1986), reproduced with permission of Springer-Verlag.

Frank, 2006

Ideal filter shape grows linearly with frequency.



http://jnm.snmjournals.org/cgi/content-nw/full/42/10/1499/F2

Filtered back-projection is a common technique, but there are several alternatives, including direct Fourier-domain reconstruction

This carries over to the 3D case

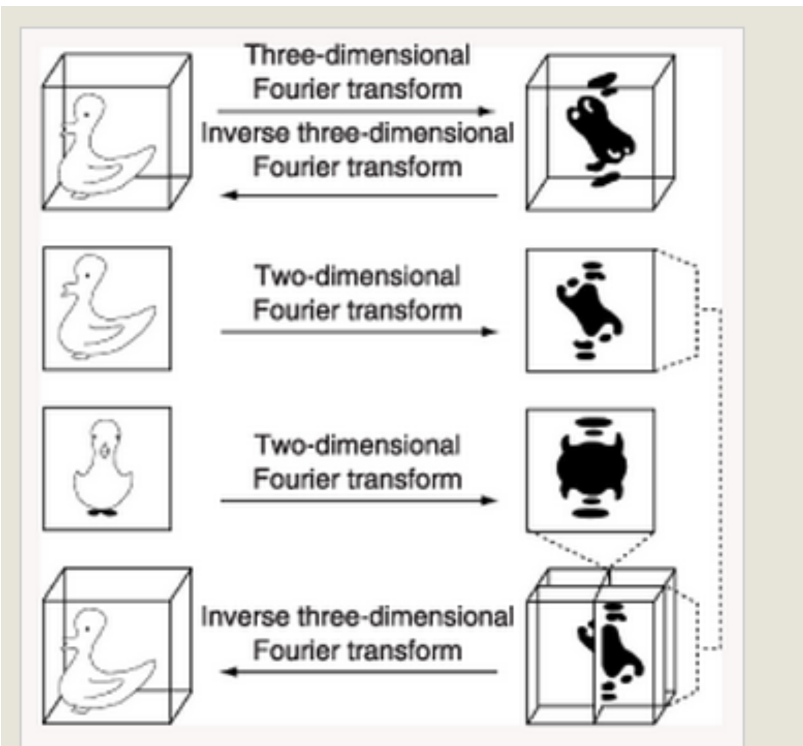


Figure 5.2 Illustration of the projection theorem and its use in 3D reconstruction. From Lake (1971), reproduced with permission of Academic Press Ltd.

Frank, 2006

Projection slice theorem (3D version): The 2D Fourier transform of the 2D projection of a 3D density is equal to the central section—perpendicular to the direction of projection—of the 3D Fourier transform of the density

You're not responsible for this

Computational reconstruction methods

3D reconstruction

Structure refinement with unknown view angles

Refining a structure

- If we’re not given the view angles for each particle, but we have a decent initial 3D model, then iterate the following steps to improve the model:
  - For each projection (i.e., each class average), find the view angle that best matches the 3D model
  - Given the newly estimated view angles, reconstruct a better 3D model (e.g., using filtered back-projection)
- This is called *3D projection matching*

An example

Class averages (starting point for reconstruction)

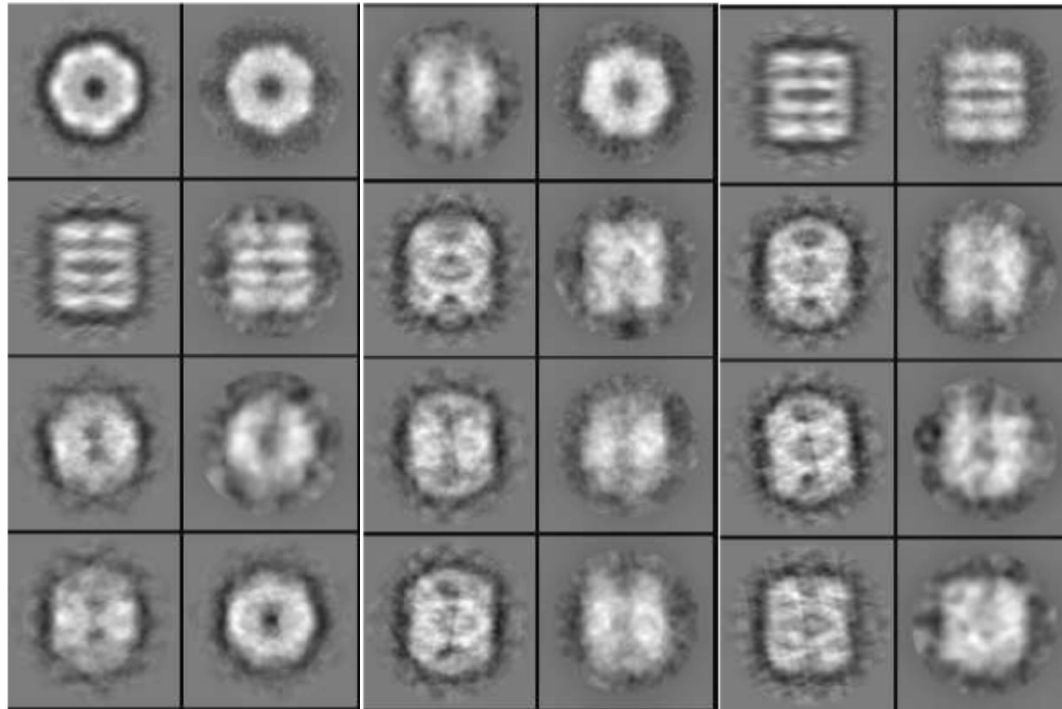
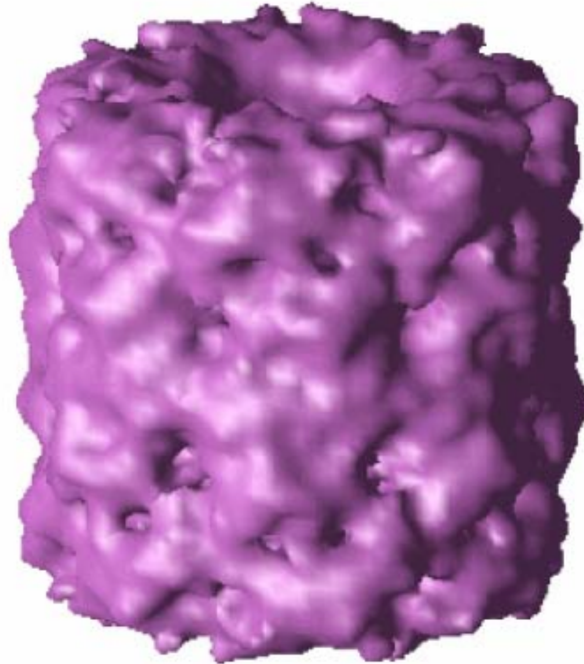


Image from Steve Ludtke
http://biomachina.org/courses/structures/091.pdf

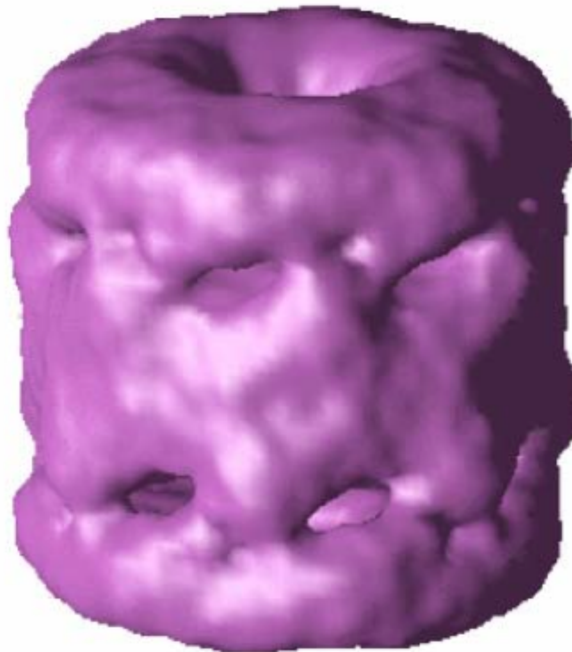
Iteration 1



This surface is a contour map. Estimated density is greater than a threshold inside the surface and less than the threshold outside it. "Density" here corresponds roughly (not precisely) to electron density.

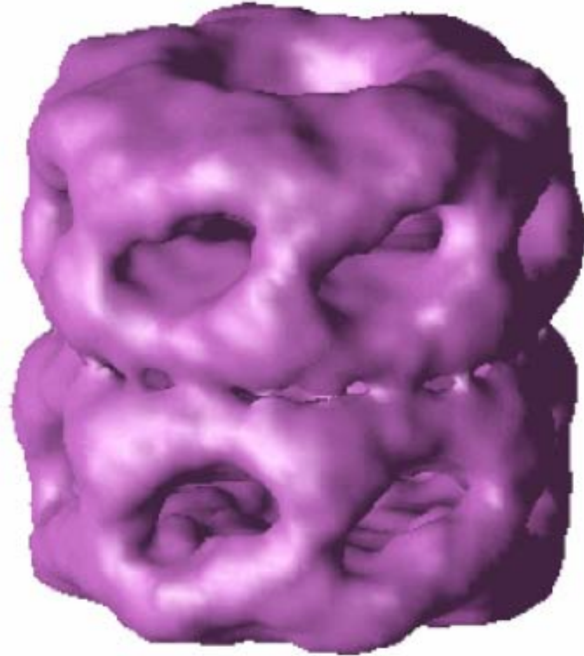
© Steve Ludtke

Iteration 2



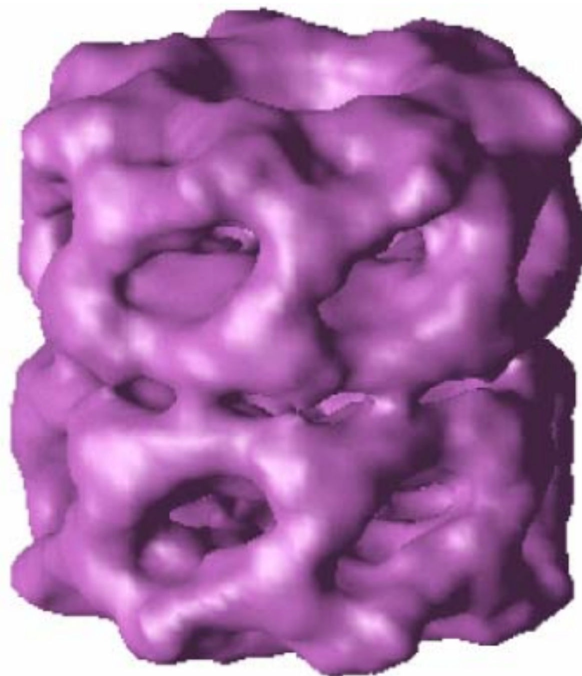
© Steve Ludtke

Iteration 3



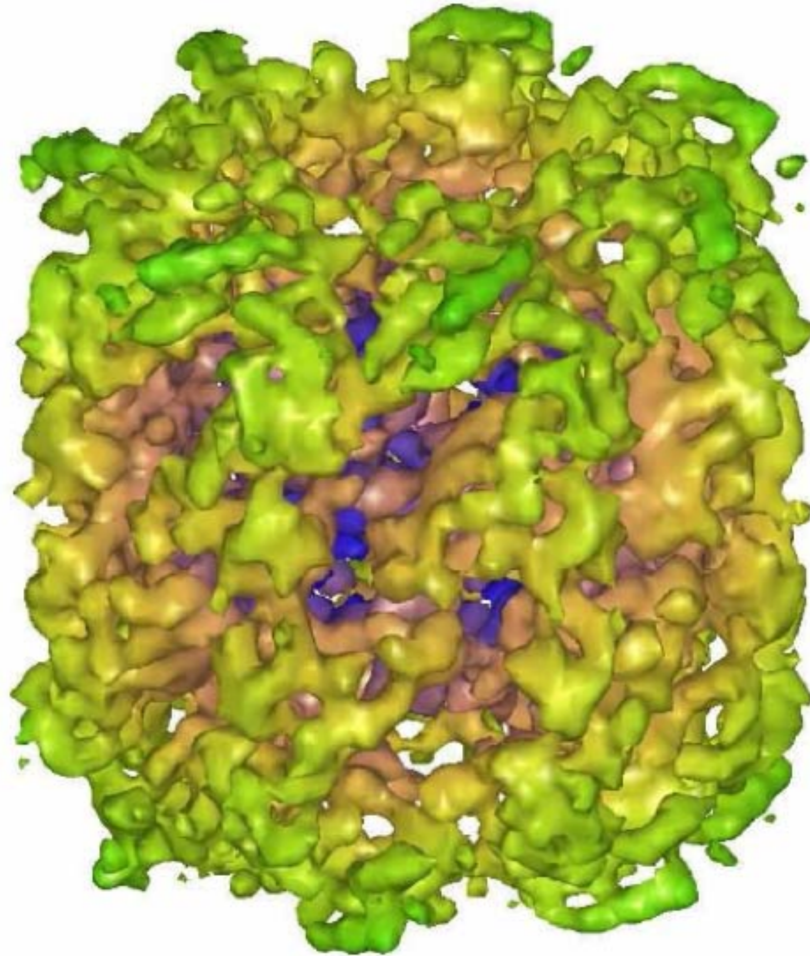
© Steve Ludtke

Iteration 4



© Steve Ludtke

Final reconstruction

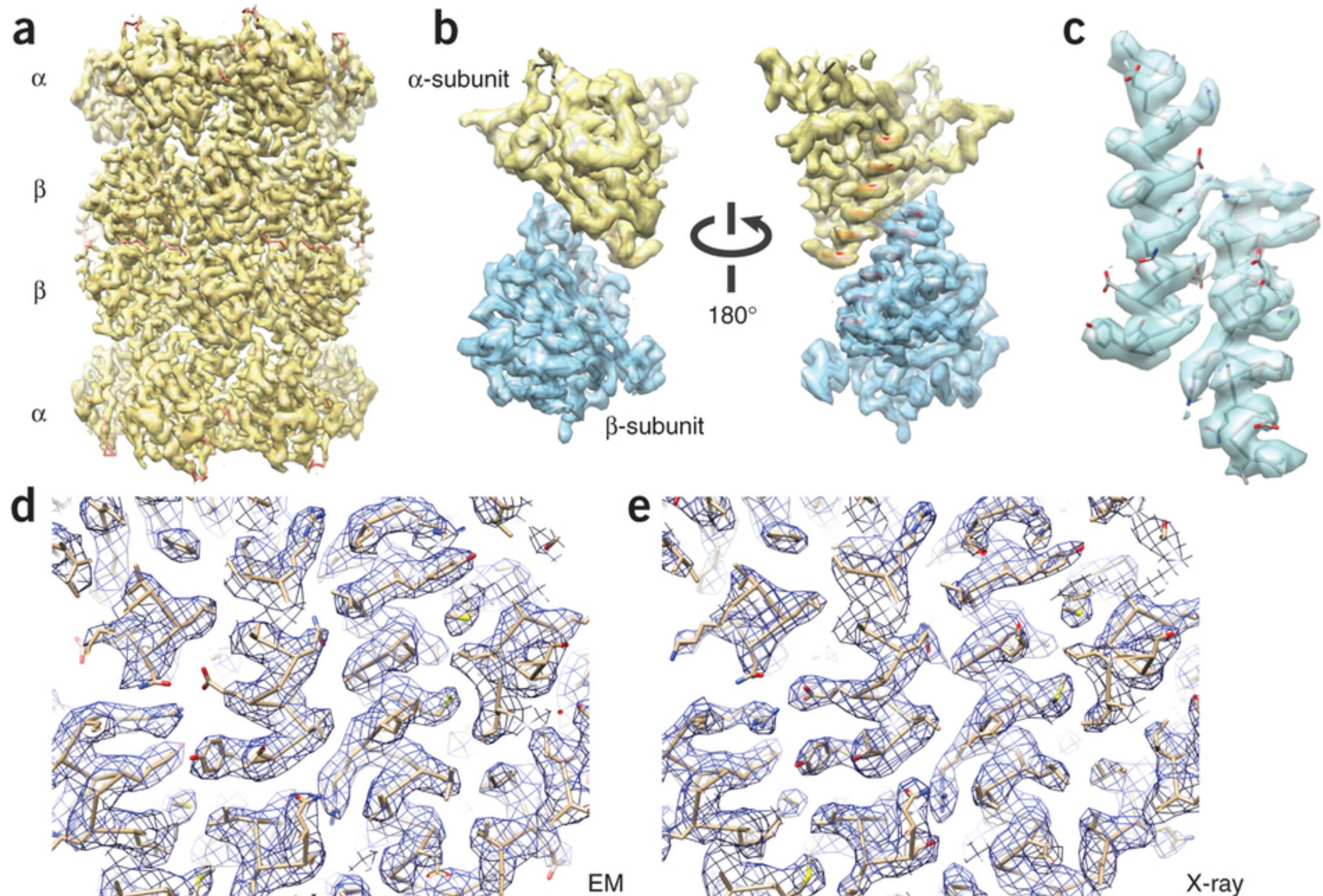


Protein: GroEL
6.5 Å resolution

Ignore the color coding

© Steve Ludtke

A high-resolution single-particle EM structure



A 3.3 Å resolution EM structure

Li et al., Nature Methods 10:584 (2013)

Caveat

- Structure refinement methods are prone to *overfitting*
  - Converged model can show features that don’t really exist and just reflect noise in the images (analogous to the issue with image clustering)
  - A variety of methods have been developed recently to deal with this issue
    - Some use Bayesian approaches (e.g., RELION software)
    - Some of the most important recent algorithmic developments in single-particle EM are in this area.

Computational reconstruction methods

3D reconstruction

Calculating an initial structure

How do we get an initial structural model?

You're not responsible for this

- Multiple options:
  - Might have one from prior experimental work (e.g., a homologous protein)
  - Conduct specialized experiments, often at lower resolution
    - Example: *random canonical tilt* approach, which requires collecting each image twice, from different angles
  - Direct computational solution
    - Common lines method: relies on the fact that Fourier transforms of different 2D projections share a common line
    - Stochastic hill climbing: a robust projection matching (refinement) approach that often allows random initialization

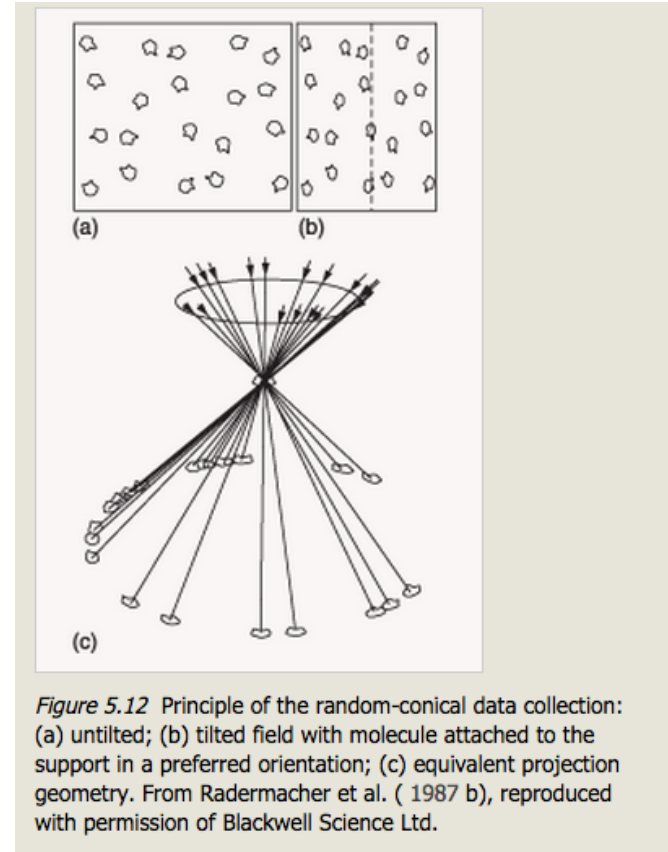


Figure 5.12 Principle of the random-conical data collection: (a) untilted; (b) tilted field with molecule attached to the support in a preferred orientation; (c) equivalent projection geometry. From Radermacher et al. (1987 b), reproduced with permission of Blackwell Science Ltd.

Frank, 2006

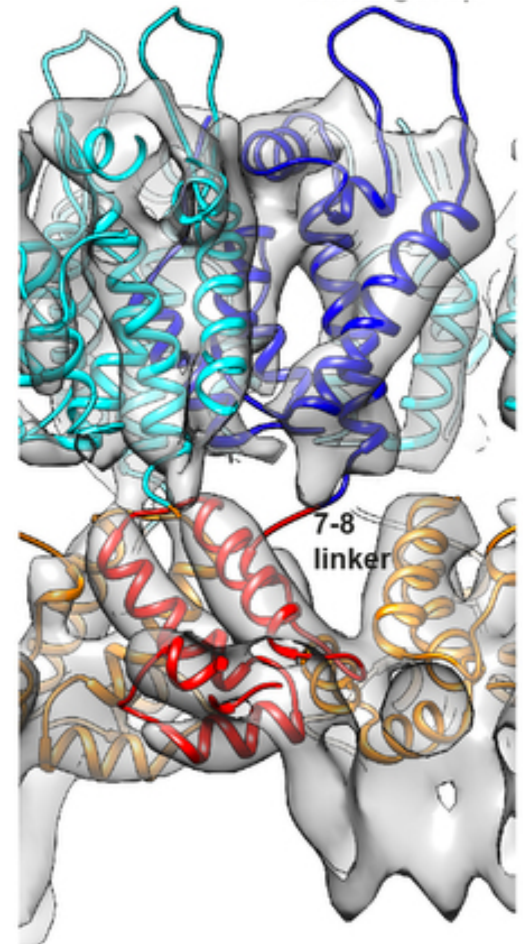
Computational reconstruction methods

3D reconstruction

Fitting atomic-resolution models to lower-resolution EM structure

Obtaining atomic-resolution models from lower-resolution EM

- Often we have high-resolution x-ray crystallography structures of each individual protein in a complex whose low-resolution structure was determined by single-particle EM.
- We can fit the high-resolution structures into the EM density computationally.



Schur et al., Nature 517:505 (2015)

Obtaining atomic-resolution models from lower-resolution EM

- Approaches based on molecular dynamics simulations can be used to allow the proteins to relax away from their crystallographic structures to better fit the EM density.

THEORETICAL AND COMPUTATIONAL
BIOPHYSICS GROUP

NIH Center for Macromolecular Modeling and Bioinformatics
www.ks.uiuc.edu

presents

Desktop MDFF

https://www.youtube.com/watch?v=6Knykqcxzfg

A very recent development: Atomic-resolution Cryo-EM

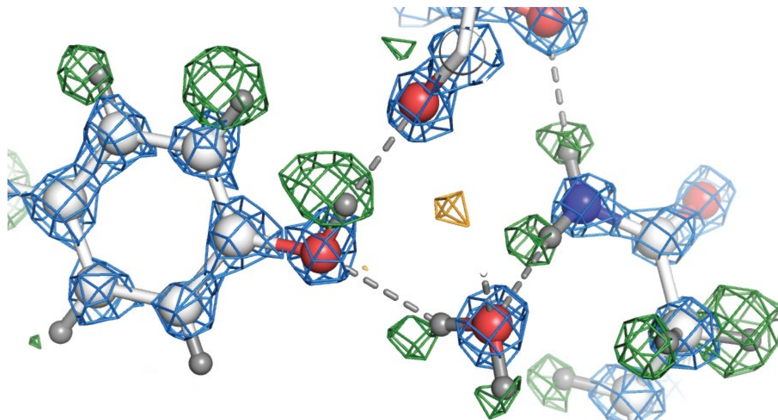
Nakane … Scheres, *Nature*, **Nov. 5, 2020**

Single-particle cryo-EM at atomic resolution

Yip … Stark, *Nature*, **Nov. 5, 2020**

Atomic-resolution protein structure determination by cryo-EM

- New technology (energy filter and new electron source, camera, software) allows resolution of 1.2 Å, at least for one protein
  - This resolution allows one to see density for individual atoms, even hydrogen



Nakane et al, *Nature*, 2020