



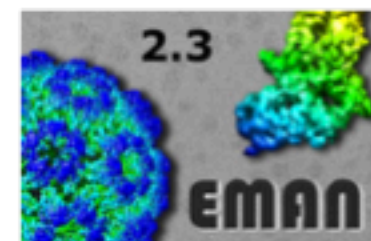
# Introduction to Single Particle Reconstruction

**Steve Ludtke**

Charles C. Bell Professor  
Biochemistry and Molecular Biology  
Director, CryoEM/CryoET Core  
Co-director CIBR Center  
Baylor College of Medicine

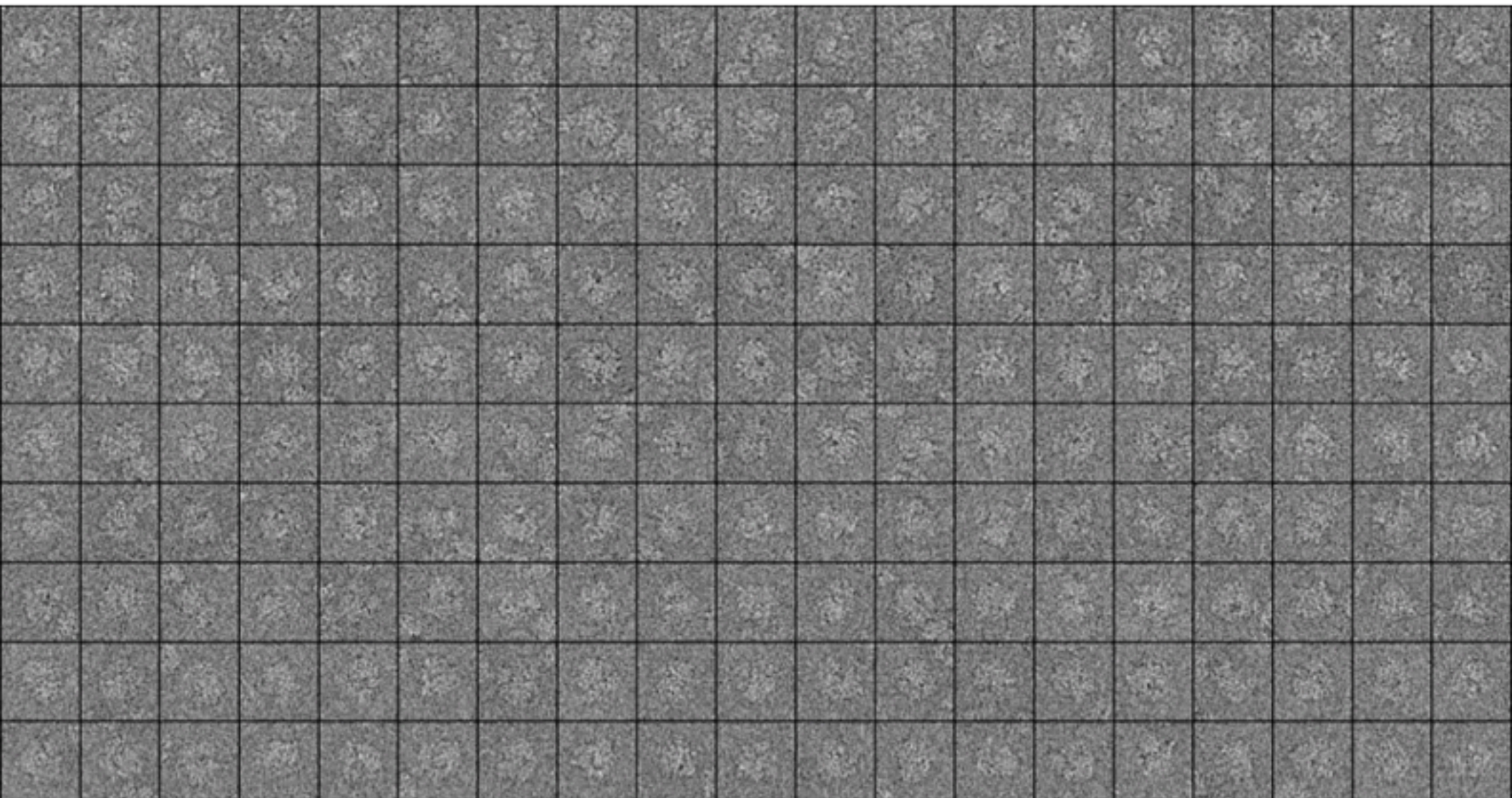
Baylor  
College of  
Medicine

VERNA & MARRS MCLEAN  
DEPARTMENT OF  
BIOCHEMISTRY AND  
MOLECULAR BIOLOGY

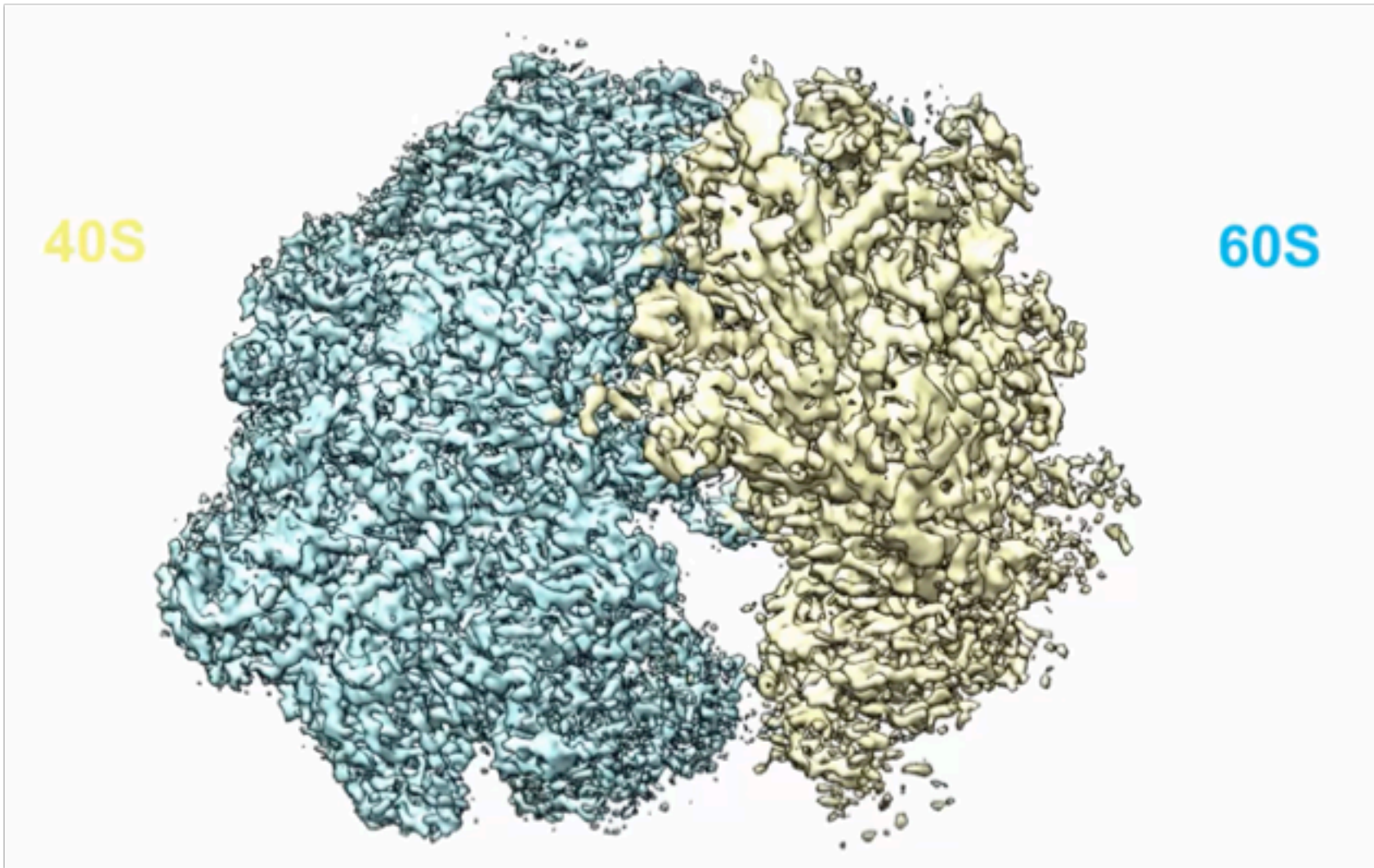


5/2019, Ludtke, UTMB

190/~30,000 Particle Images

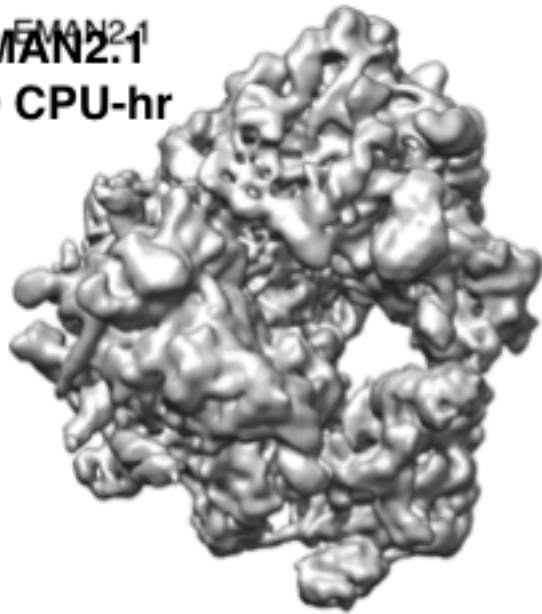




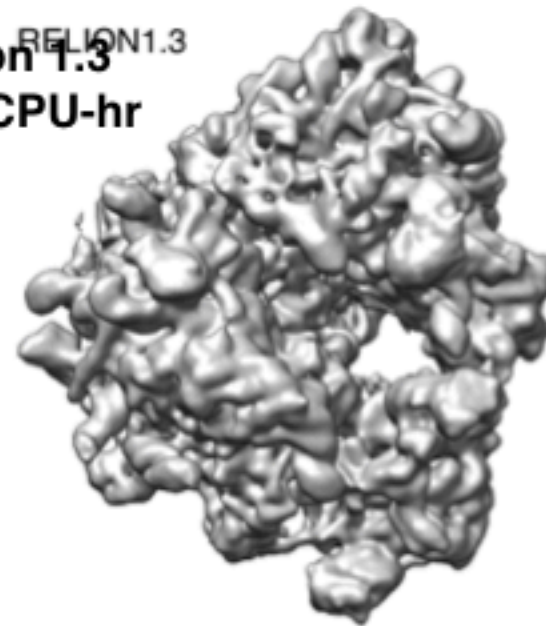


Bai, X. C., Fernandez, I. S., McMullan, G. & Scheres, S. H. Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. *Elife* 2, e00461 (2013). PMC3576727.

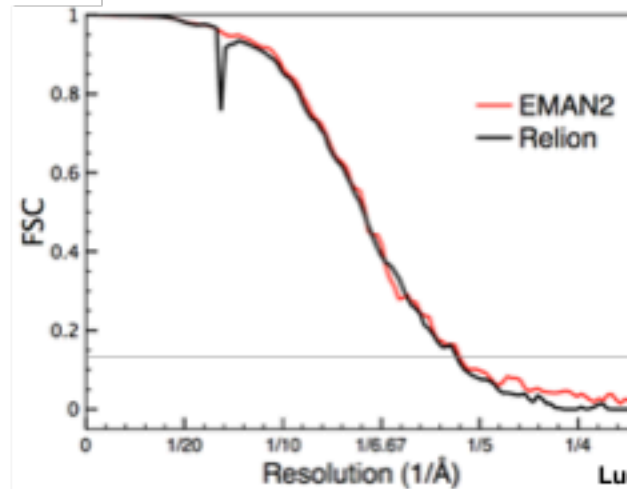
<sup>EMAN2.1</sup>  
**EMAN2.1**  
320 CPU-hr



<sup>RELION1.3</sup>  
**Relion 1.3**  
2200 CPU-hr

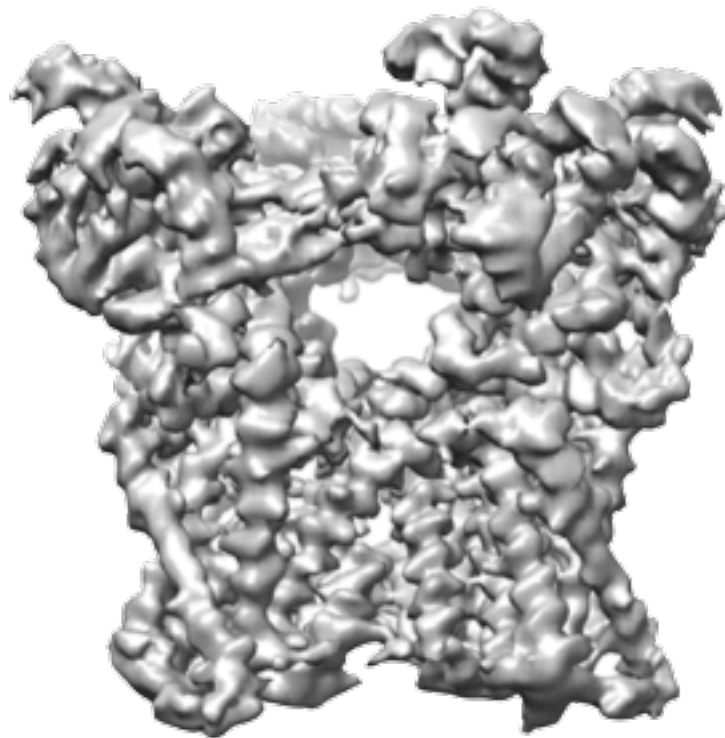


**S. Cerevisiae 80S Ribosome (EMD-2275)**  
Data taken from PDBe 3DEM test data  
no movie alignment performed  
Dataset 10002 (Bai XC, Fernandez IS,  
McMullen G, Scheres SH)



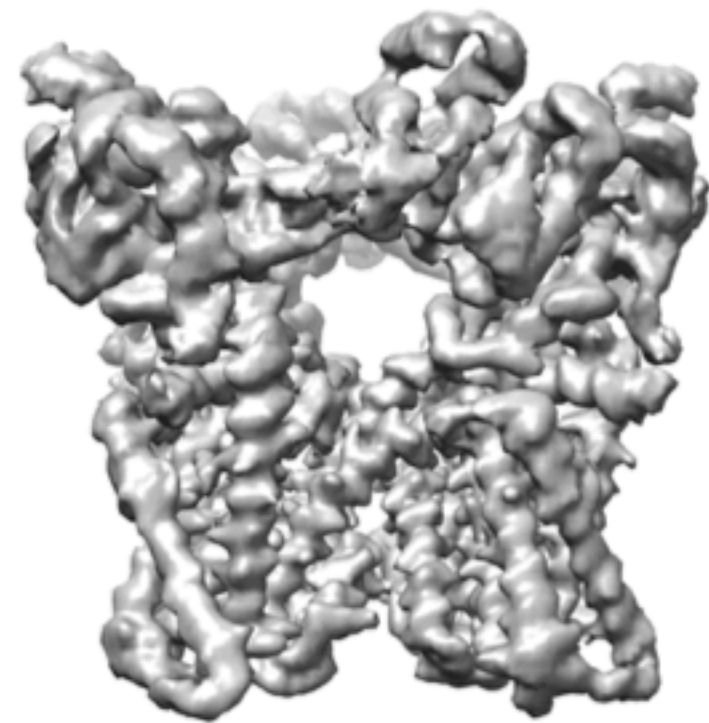
Ludtke, 1/28/16

# TRPV1



EMAN2

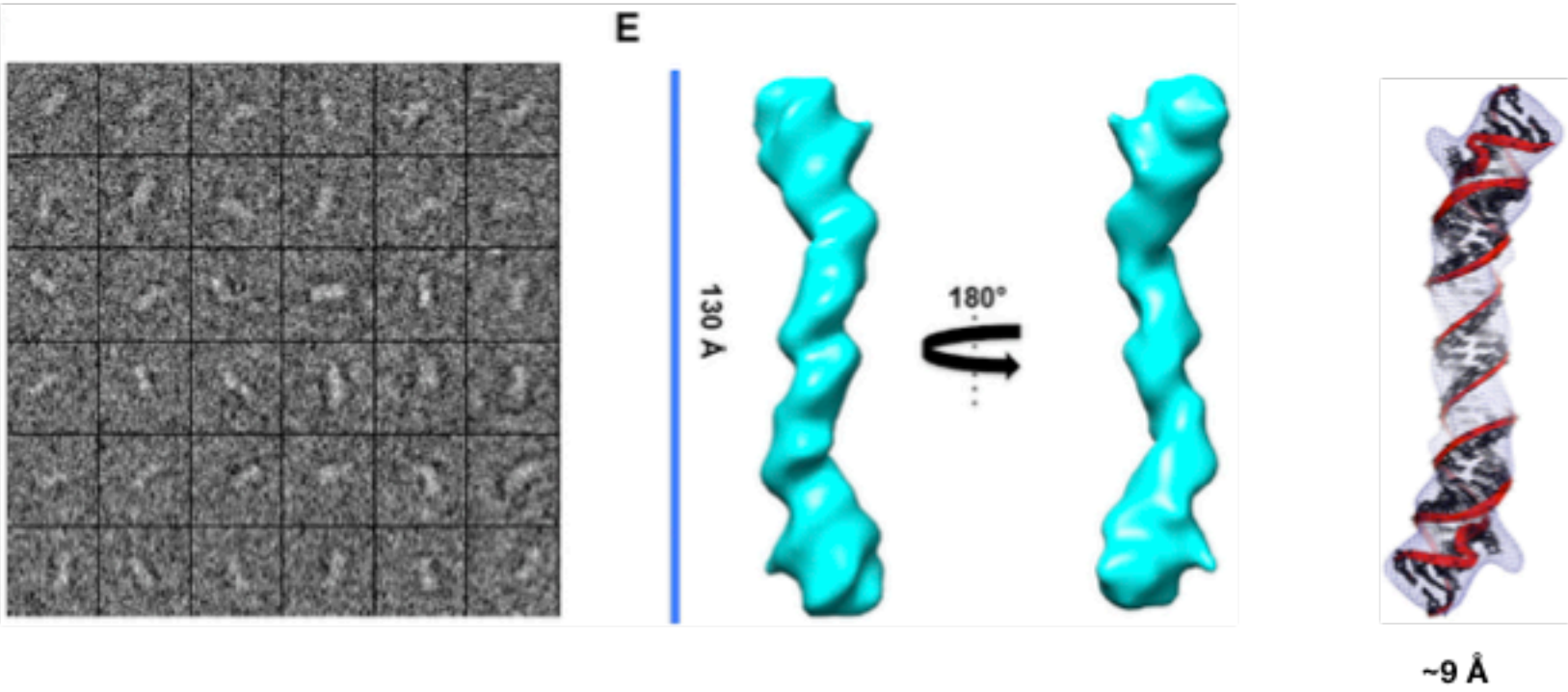
3.7 Å



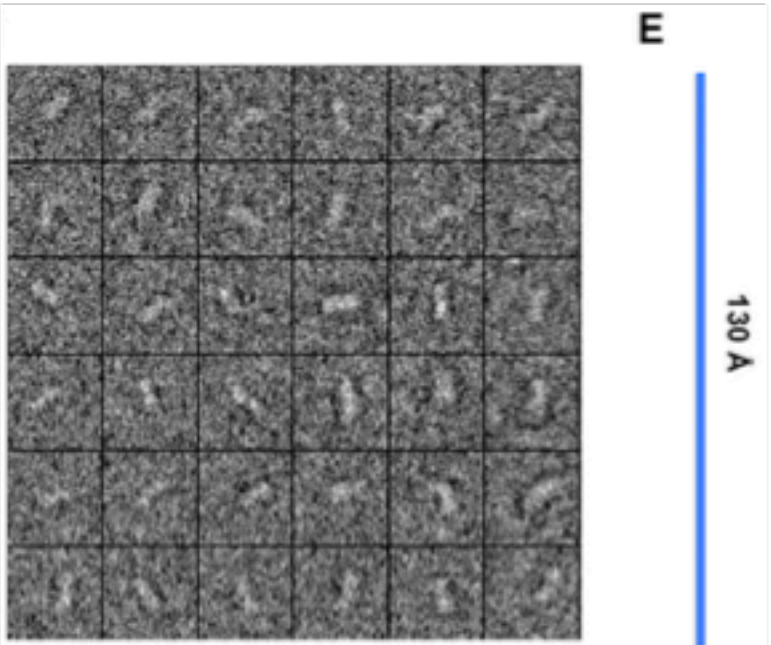
Relion

Liao, M., Cao, E., Julius, D., and Cheng, Y. (2013). Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature*. 504:107-112.

# 30 kDa HIV-1 RNA Dimerization Signal



# 30 kDa HIV-1 RNA Dimerization Signal

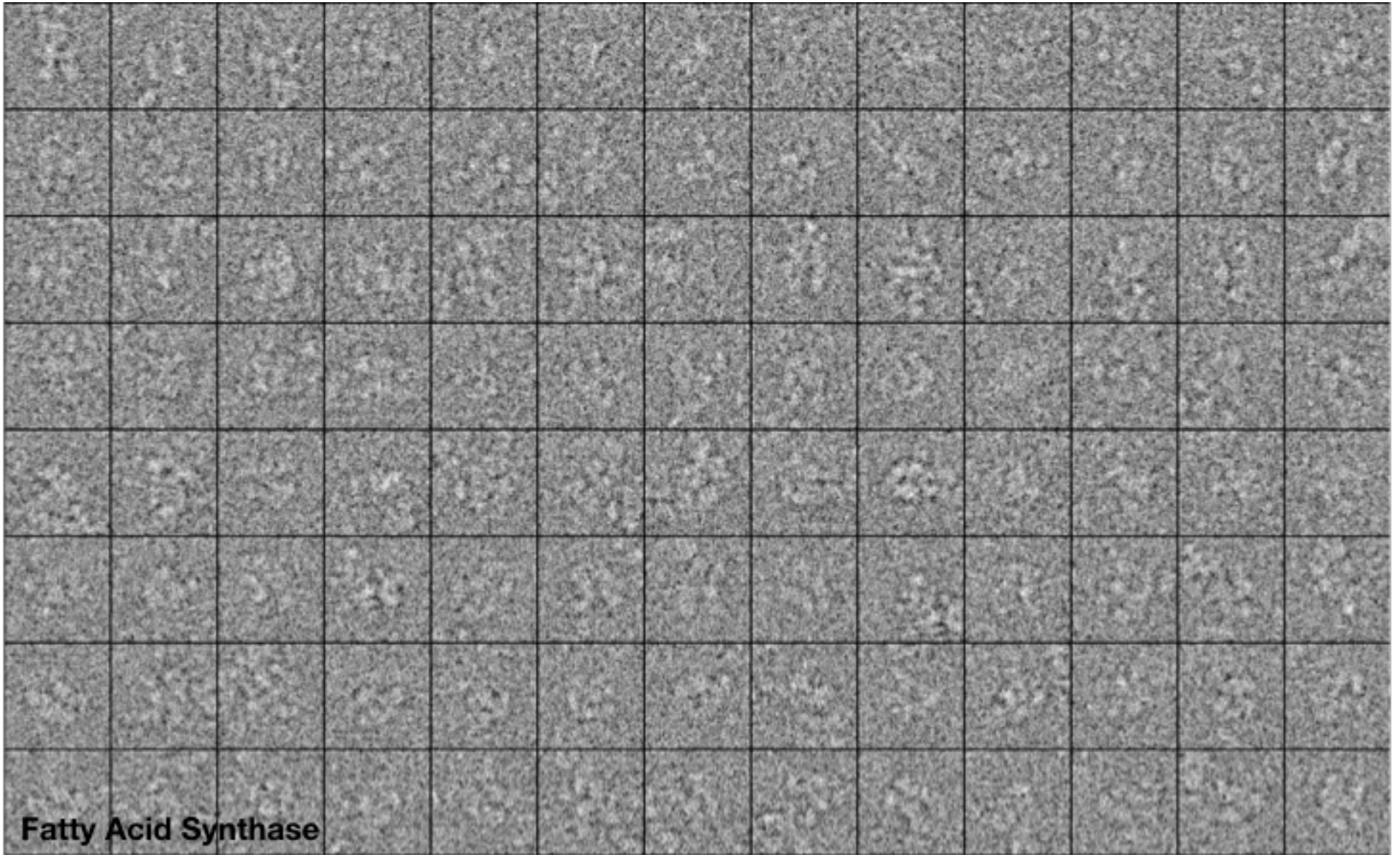


~15 Å  
Relion

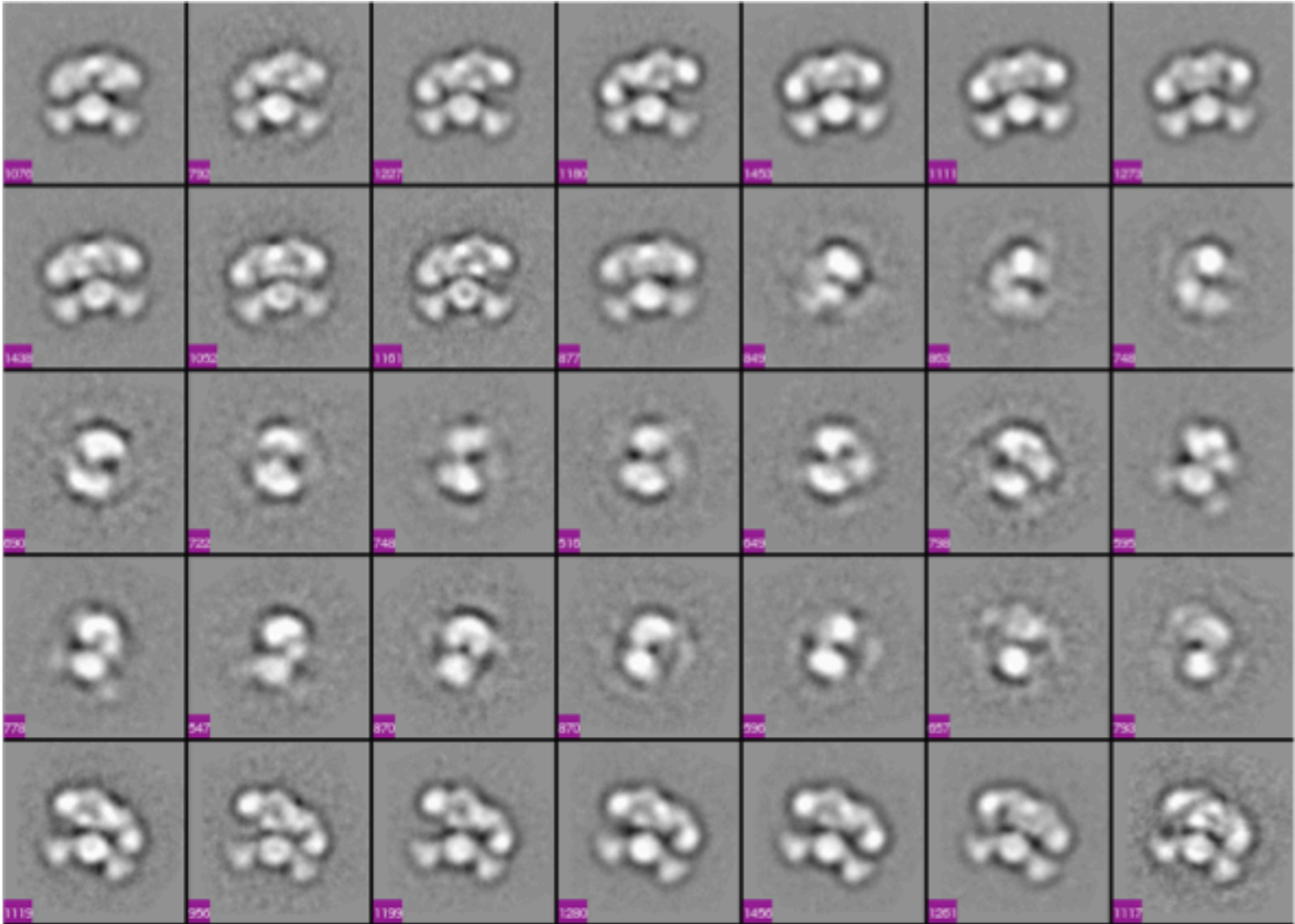


~9 Å  
EMAN2

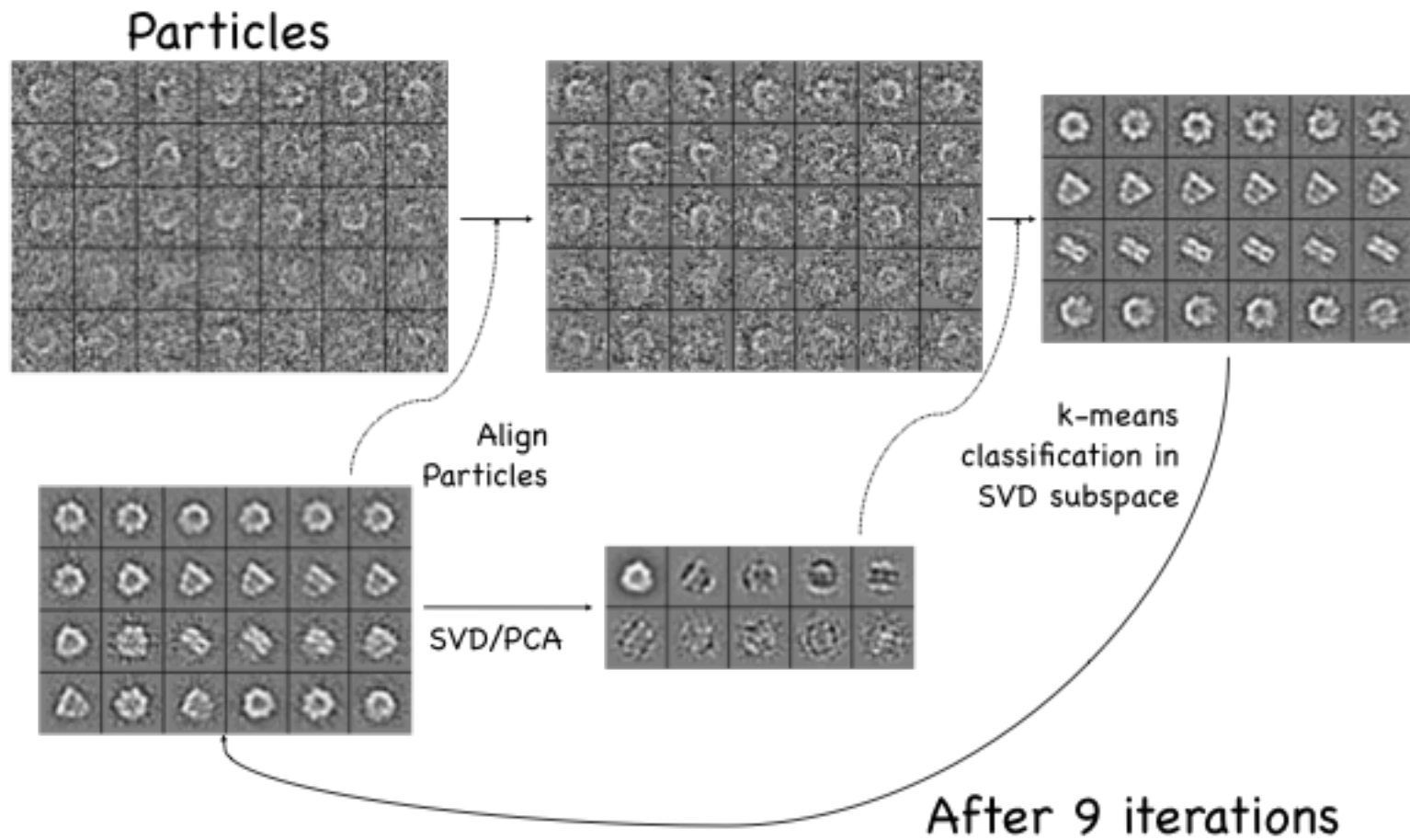






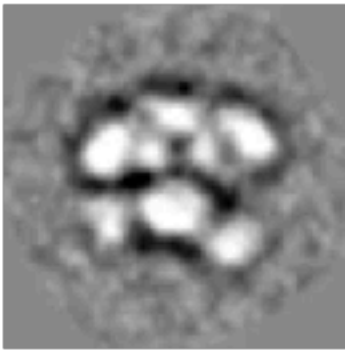


# Unsupervised Classification

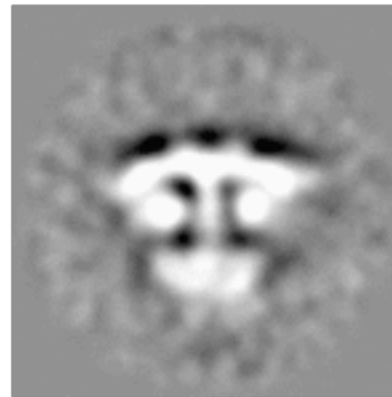


GroEL single ring mutant + GroES

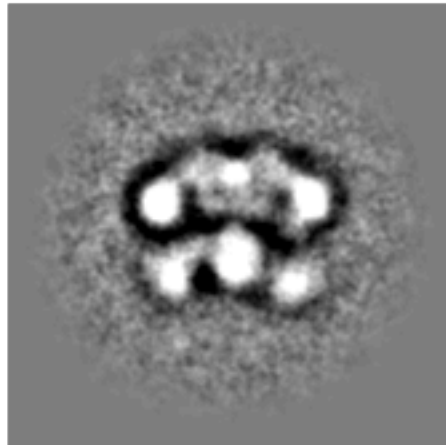
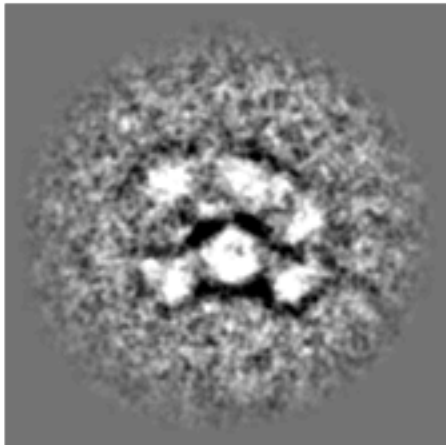
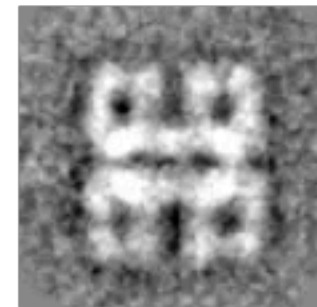
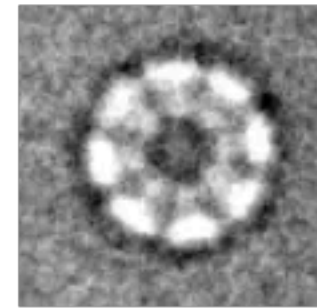
Fatty Acid  
Synthase  
~30 Å motion



IP3 Receptor  
~18 Å motion



TriC (open)  
~12 Å motion



E

E

B

E

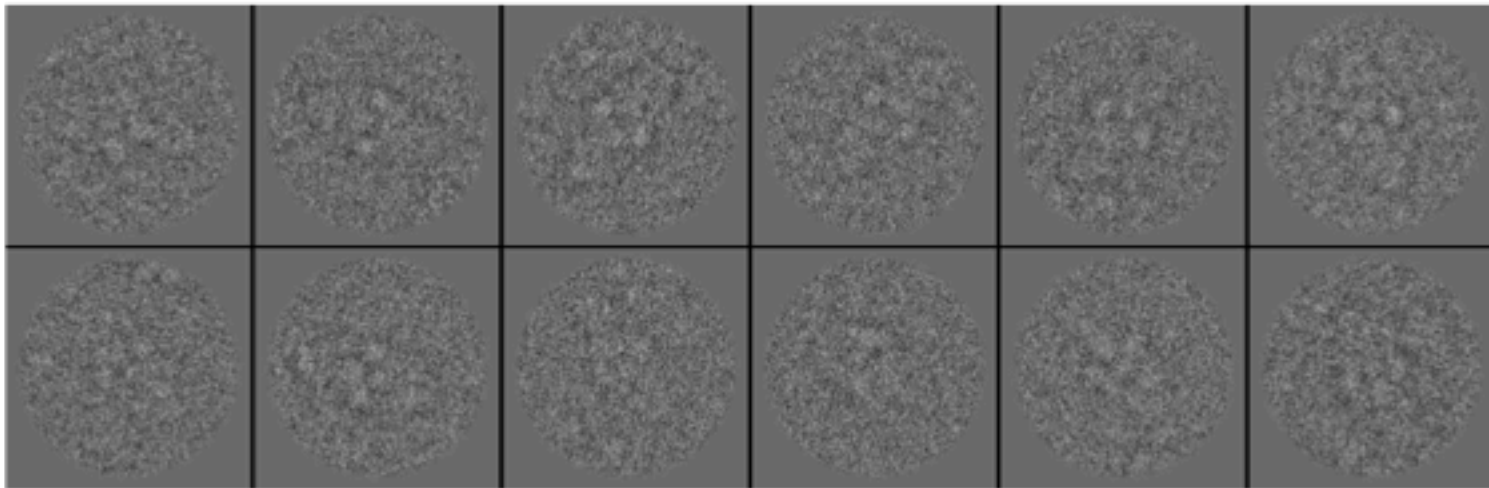
E

B

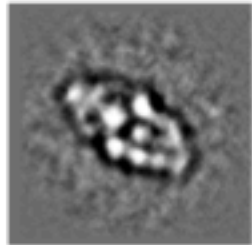
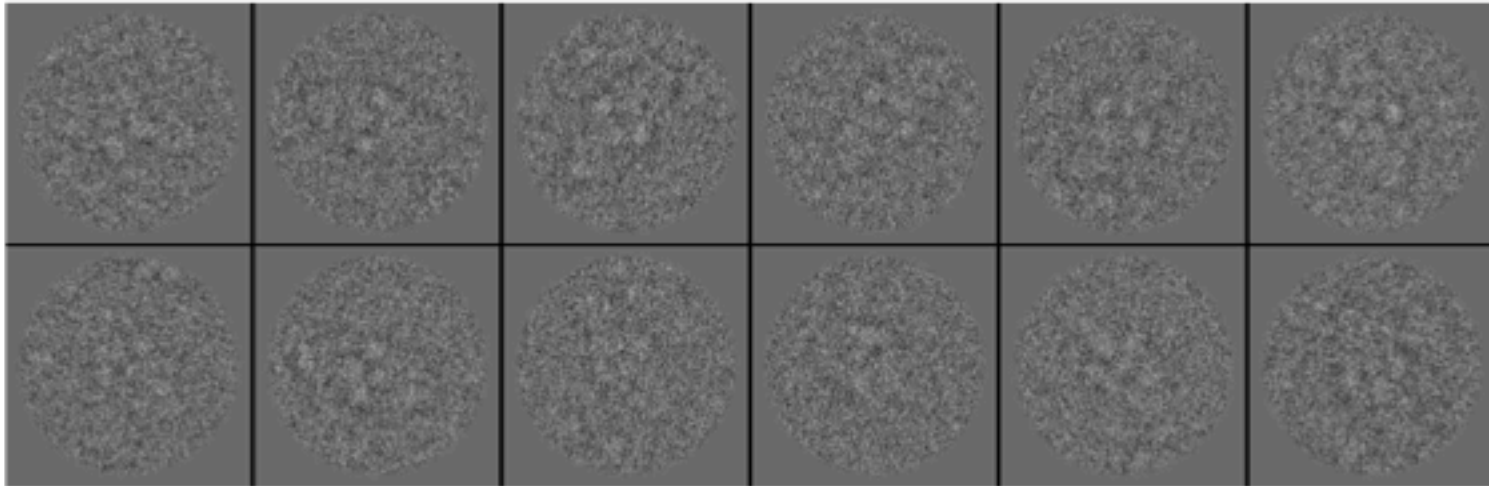
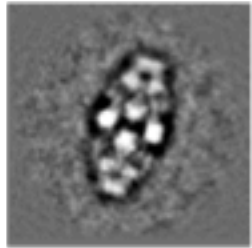
B

E





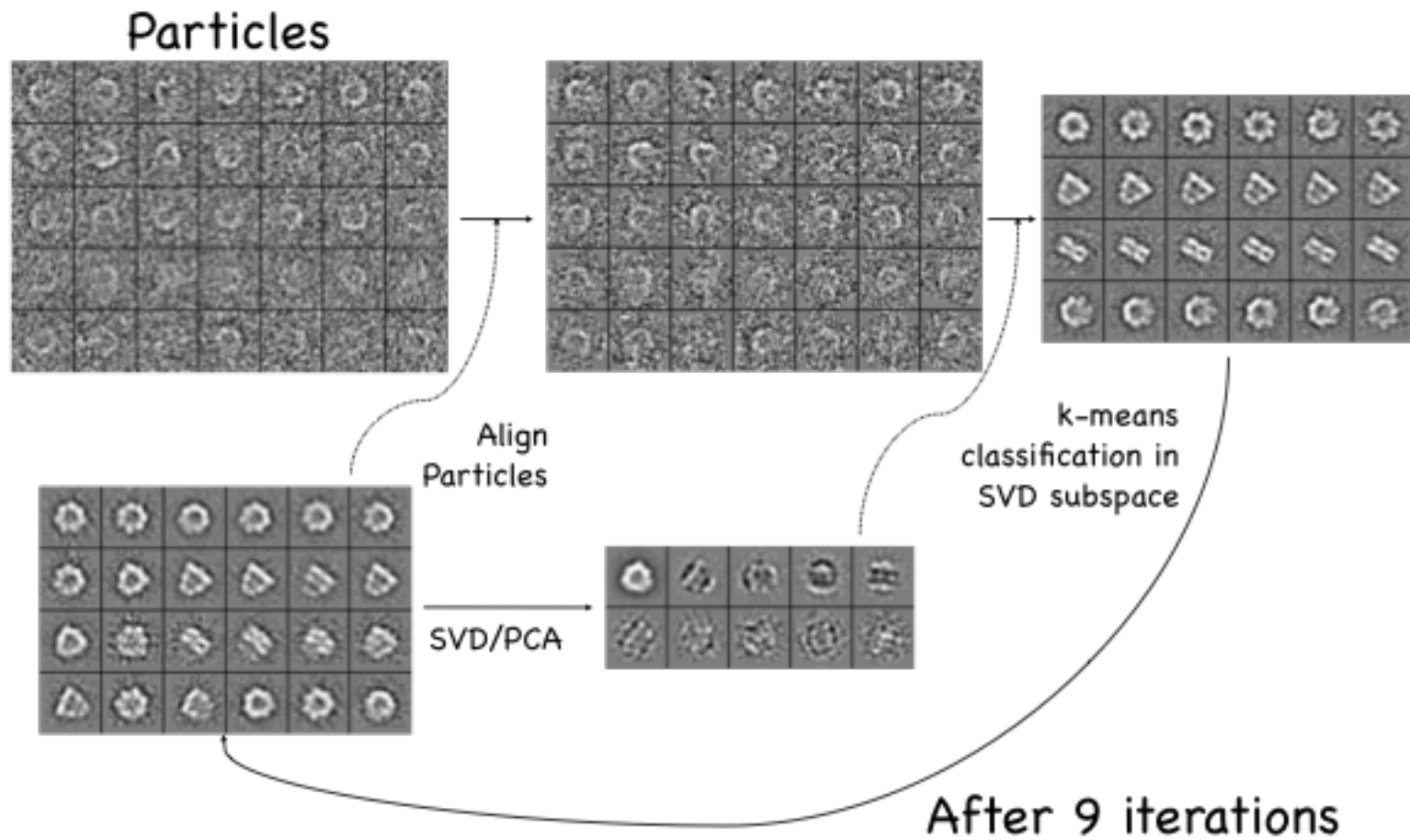
**$\beta$  Galactosidase**



$\beta$  Galactosidase

(this is real  $\sim 2.2$  Å B-gal data)

# Unsupervised Classification



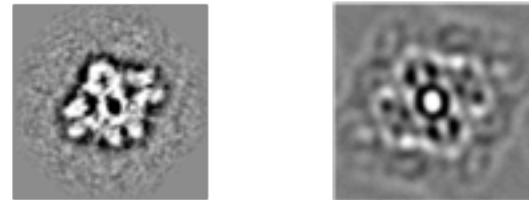
GroEL single ring mutant + GroES

# Invariants ?

- An image derived from another image which doesn't change under some operation
- A translational invariant is an image that doesn't change when its source is moved around in the box
- A rotational invariant is an image that doesn't change when its source is rotated
  
- If we had a good one we could skip costly alignments!

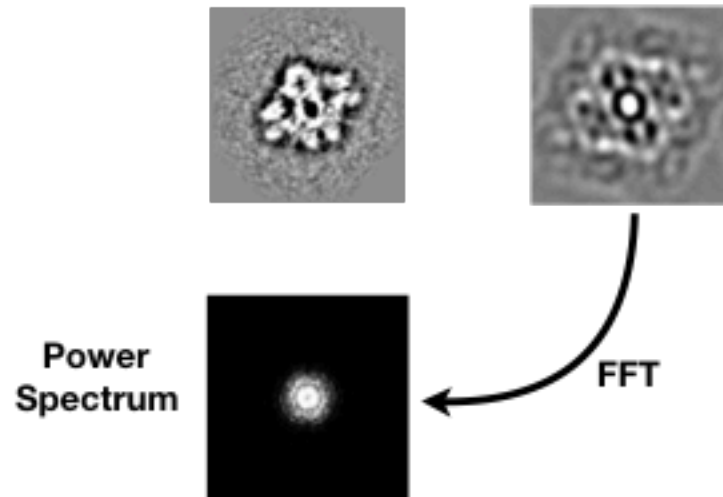


# Autocorrelation



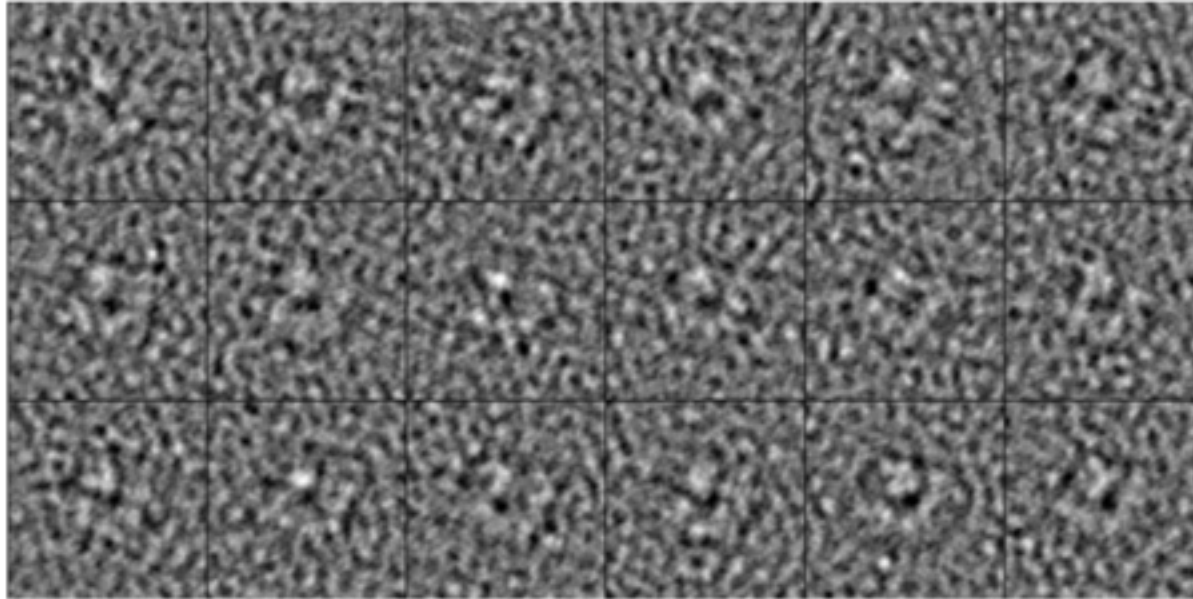
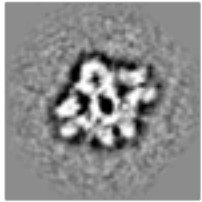
$\beta$  Galactosidase

# Autocorrelation

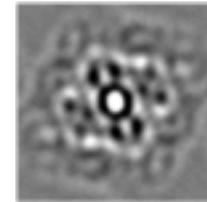


$\beta$  Galactosidase

# Autocorrelation



Any One of These

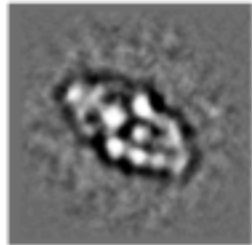
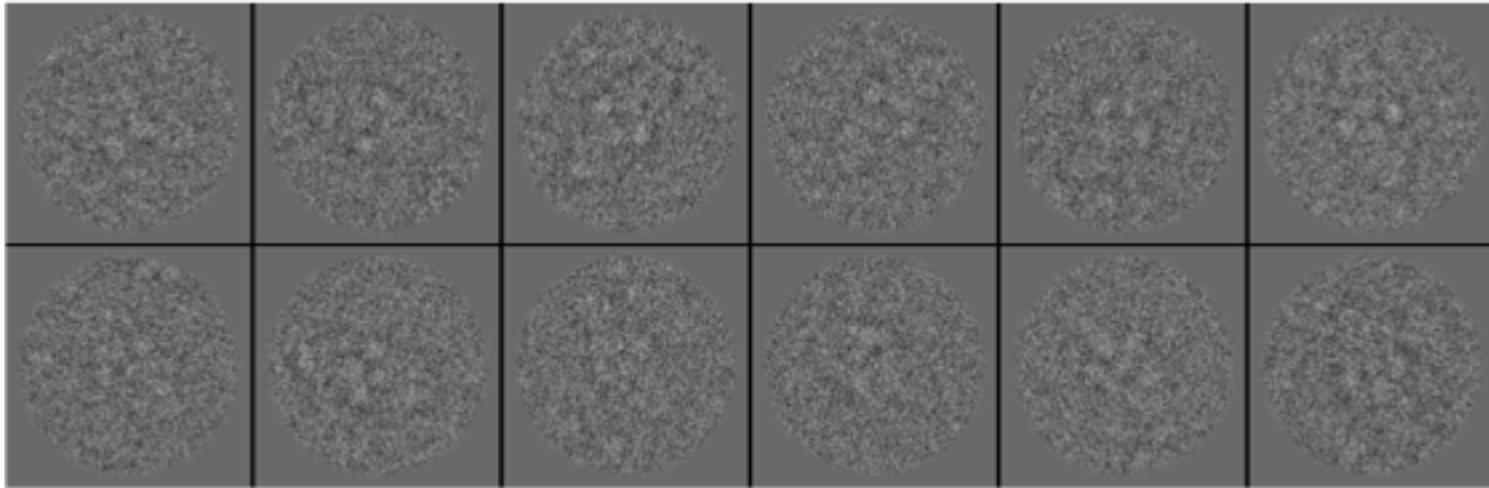
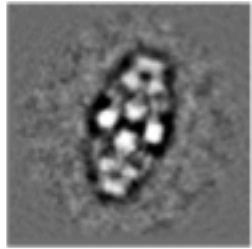


And we still haven't done rotation....

# Problems with Autocorrelation

- We lose phases, twice. Significant information loss
- Handedness lost
- Noise additivity
  - $P(A+B+C+D) \neq P(A)+P(B)+P(C)+P(D)$





(this is real  $\sim 2.2 \text{ \AA}$  B-gal data)

# Bispectrum

- Power spectrum:

$$P(\bar{s}) = |F(\bar{s})|^2$$

- Bispectrum

$$B(\bar{s}_1, \bar{s}_2) = F(\bar{s}_1)F(\bar{s}_2)F^*(\bar{s}_1 + \bar{s}_2)$$



**4-D !?!**

**3 point correlation in real space  
translationally invariant**

# Bispectrum

- Power spectrum:

$$P(\bar{s}) = |F(\bar{s})|^2$$

- Bispectrum

$$B(\bar{s}_1, \bar{s}_2) = F(\bar{s}_1)F(\bar{s}_2)F^*(\bar{s}_1 + \bar{s}_2)$$

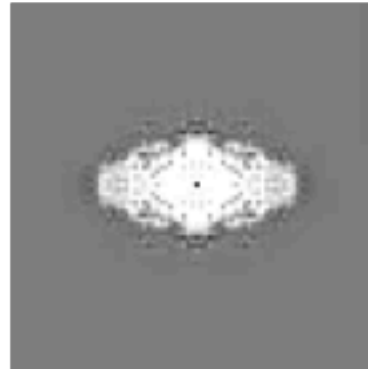


**4-D !?!**

**3 point correlation in real space**

**... So we compute a carefully selected 3-D subspace from the bispectrum integrated rotationally.**

# Bispectrum

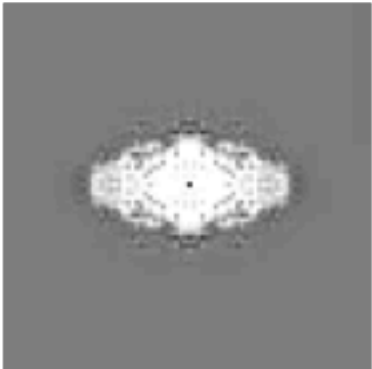


Significant changes with 3-D orientation

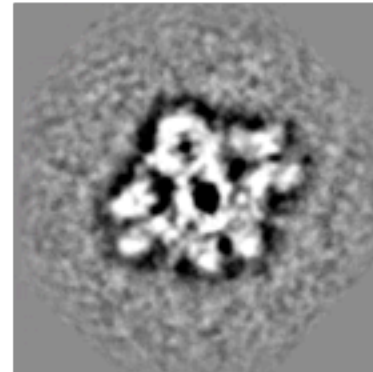
$\beta$  Galactosidase



# Bispectrum

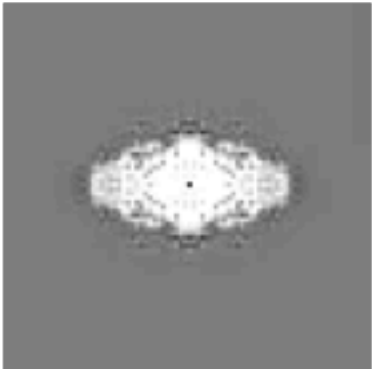


$\beta$  Galactosidase

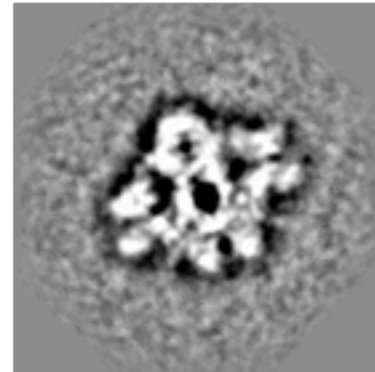


Good translational invariance

# Bispectrum

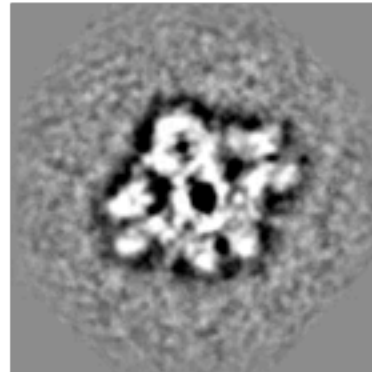


$\beta$  Galactosidase



Good rotational invariance

# Bispectrum



Good noise behavior



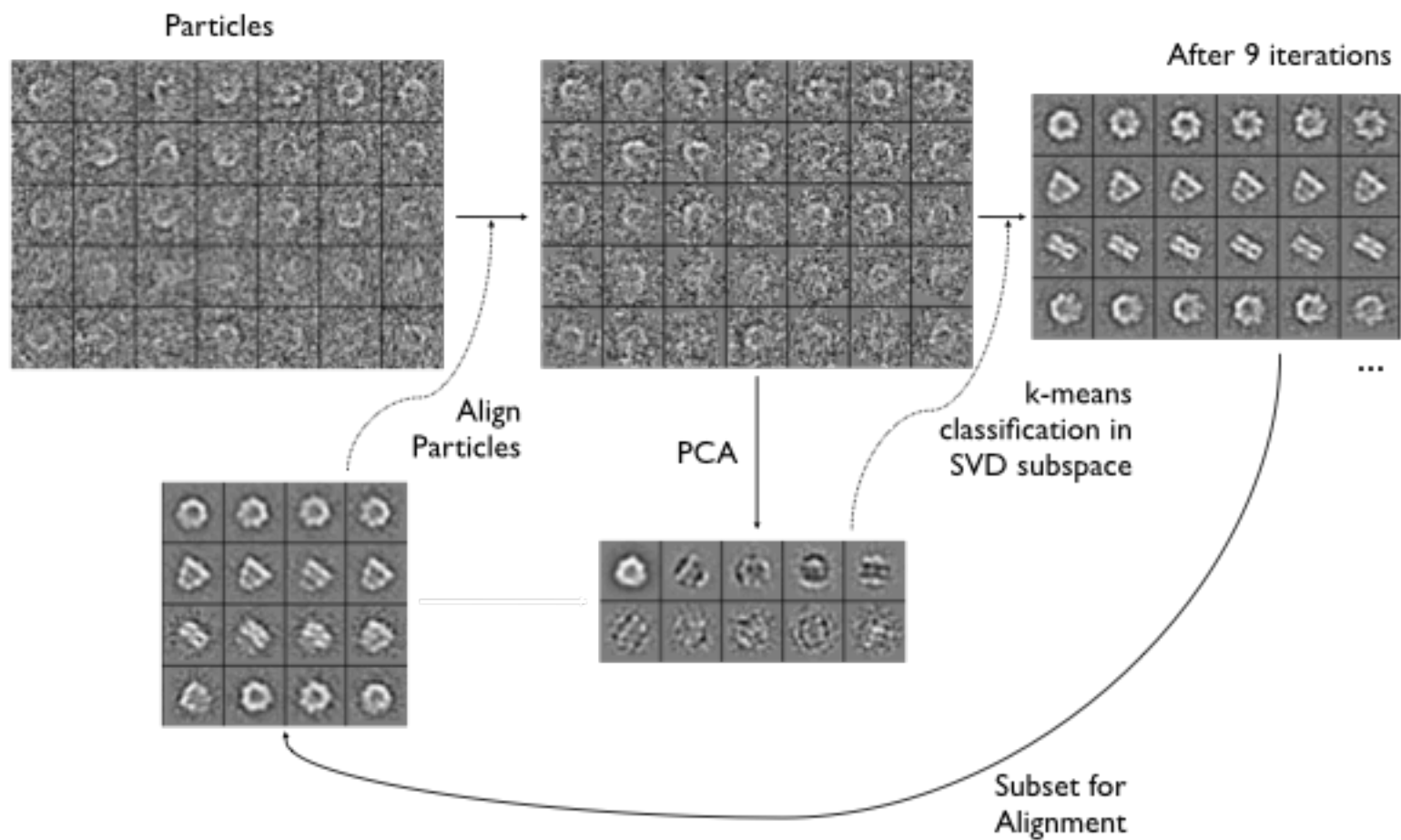
$\beta$  Galactosidase



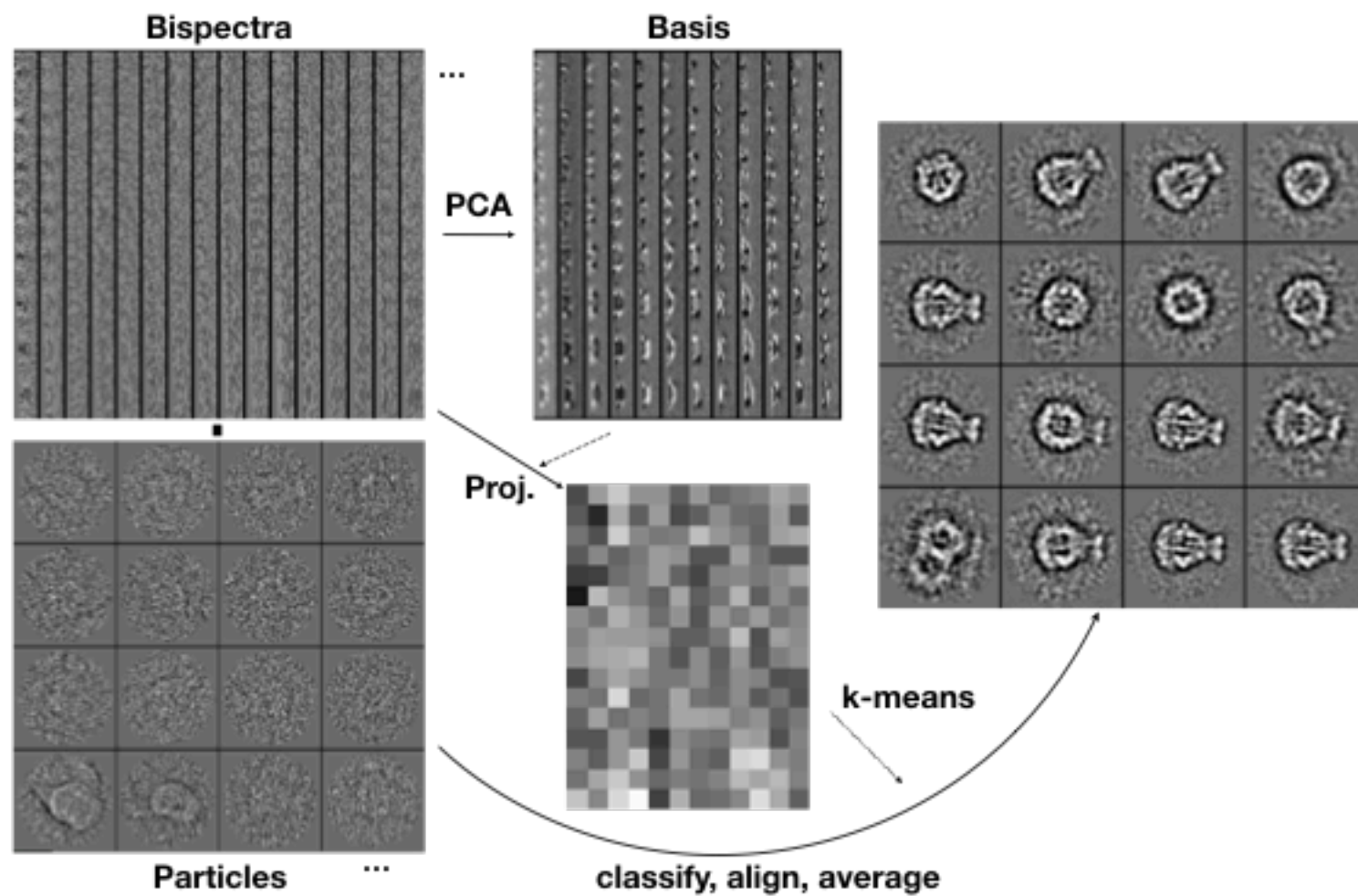
# Bispectral Invariants

- Preserve more information
- Good noise resistance
- Noise is still noise and can cancel out with averaging
- Relatively inexpensive (generate ~50 each sec per CPU)

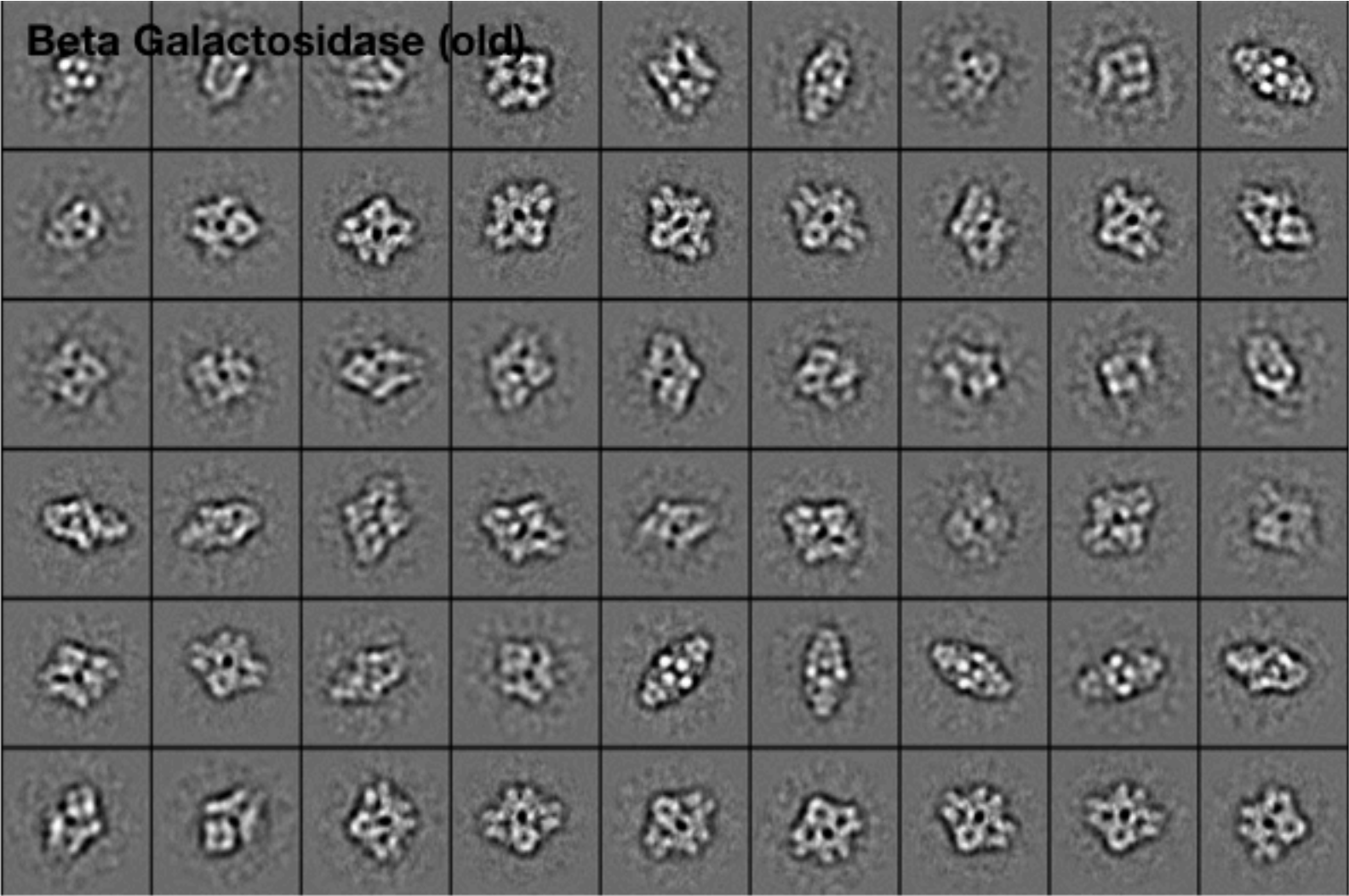
# (old) 2D Refinement



# (new) 2D Refinement

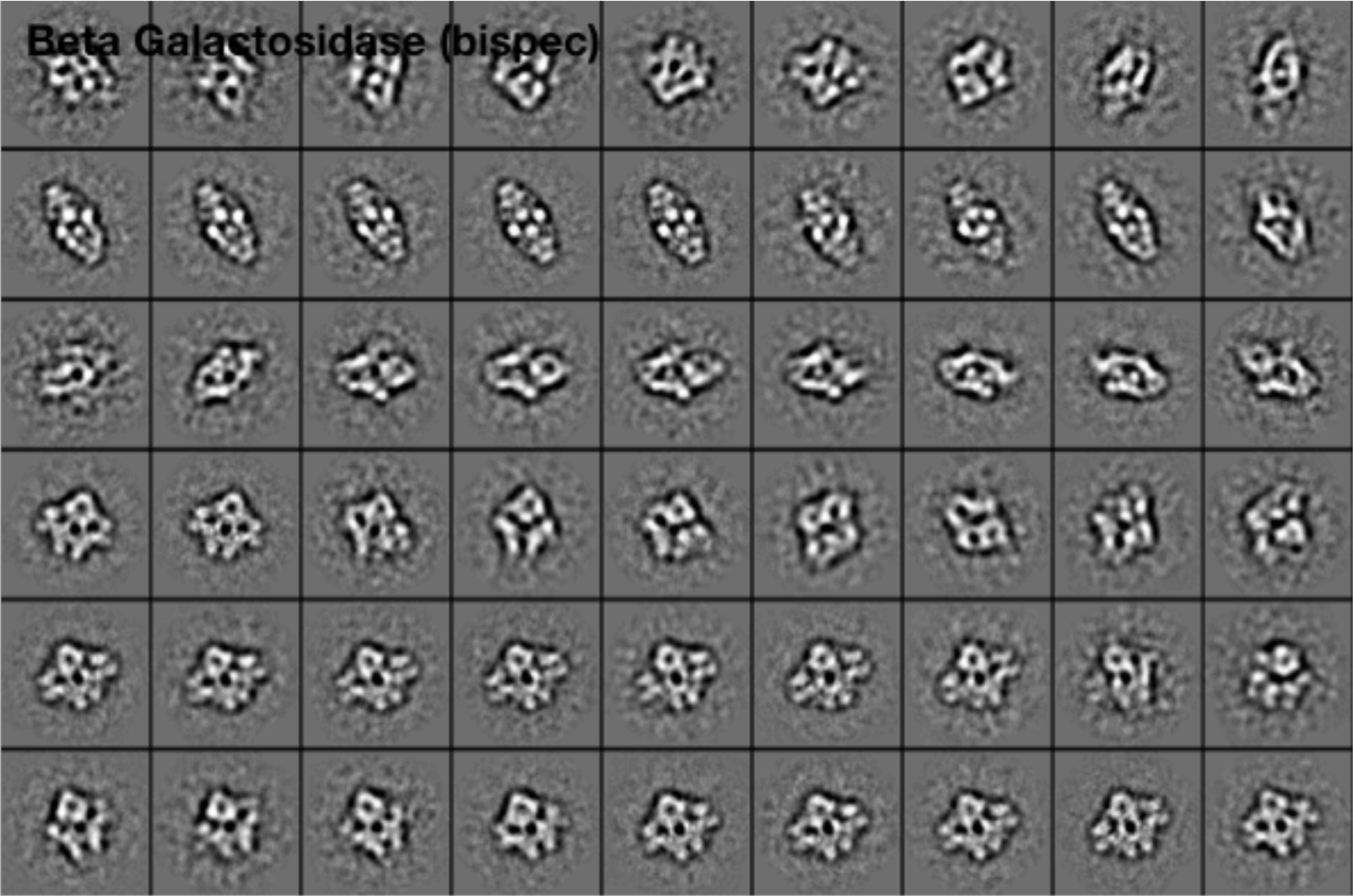


TcdA1



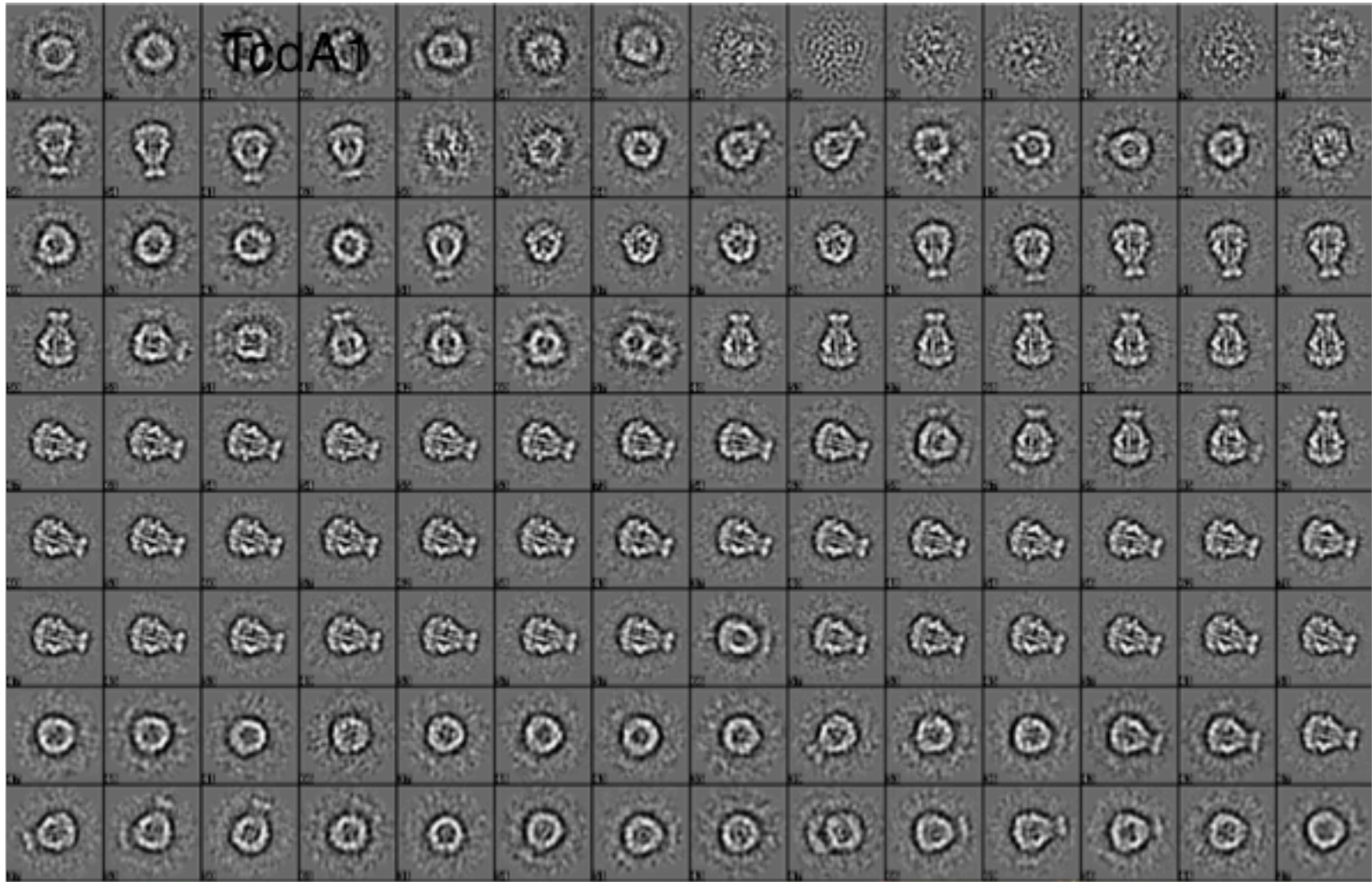
5513 particles  
240x240 box size  
64 Classes

2 h 27 m  
1 Laptop (4 core)



5513 particles  
240x240 box size  
64 Classes

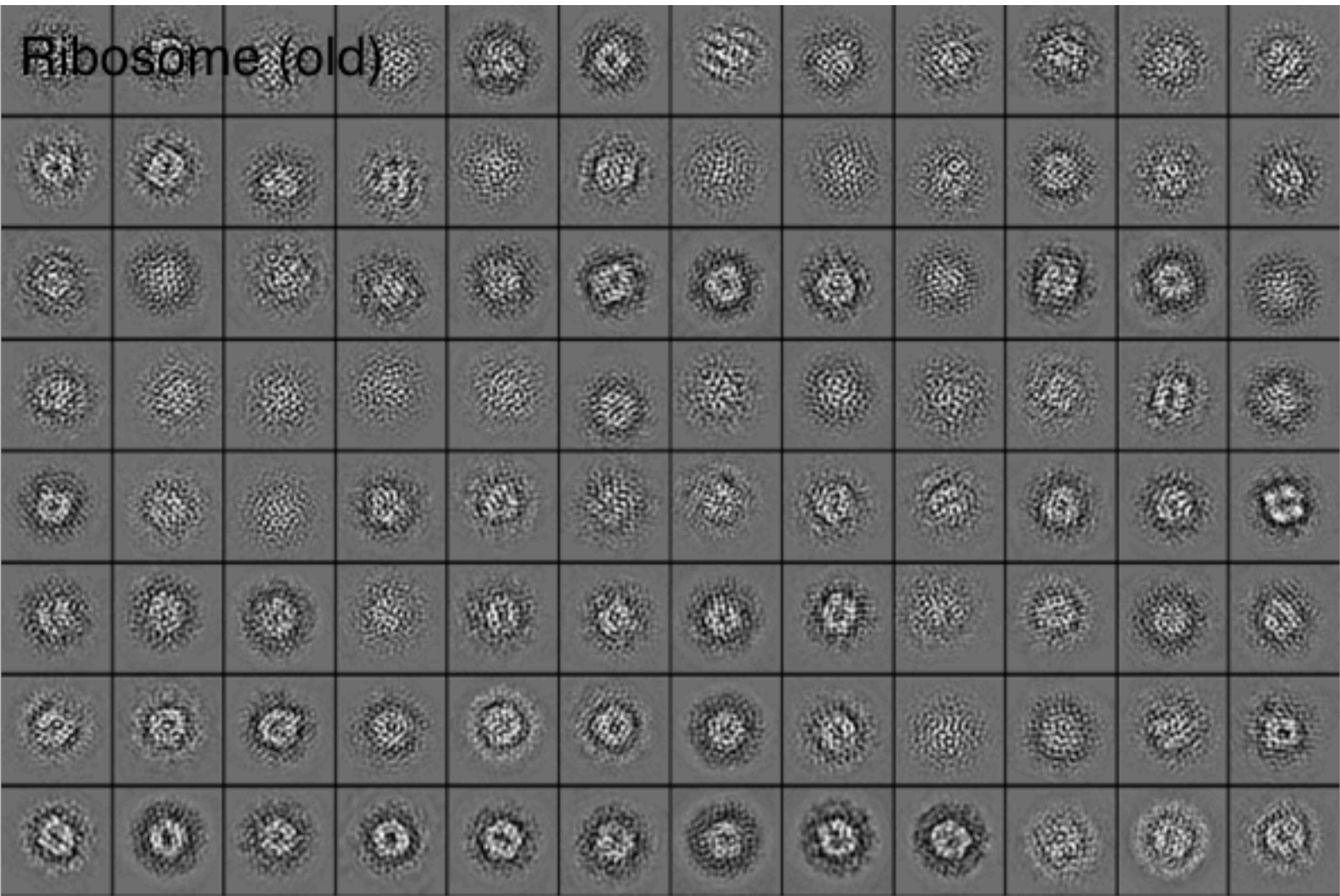
7 min  
1 Laptop (4 core)



8,290 particles  
384x384 box size  
128 Classes

15 sec to classify  
4 min total  
1 workstation

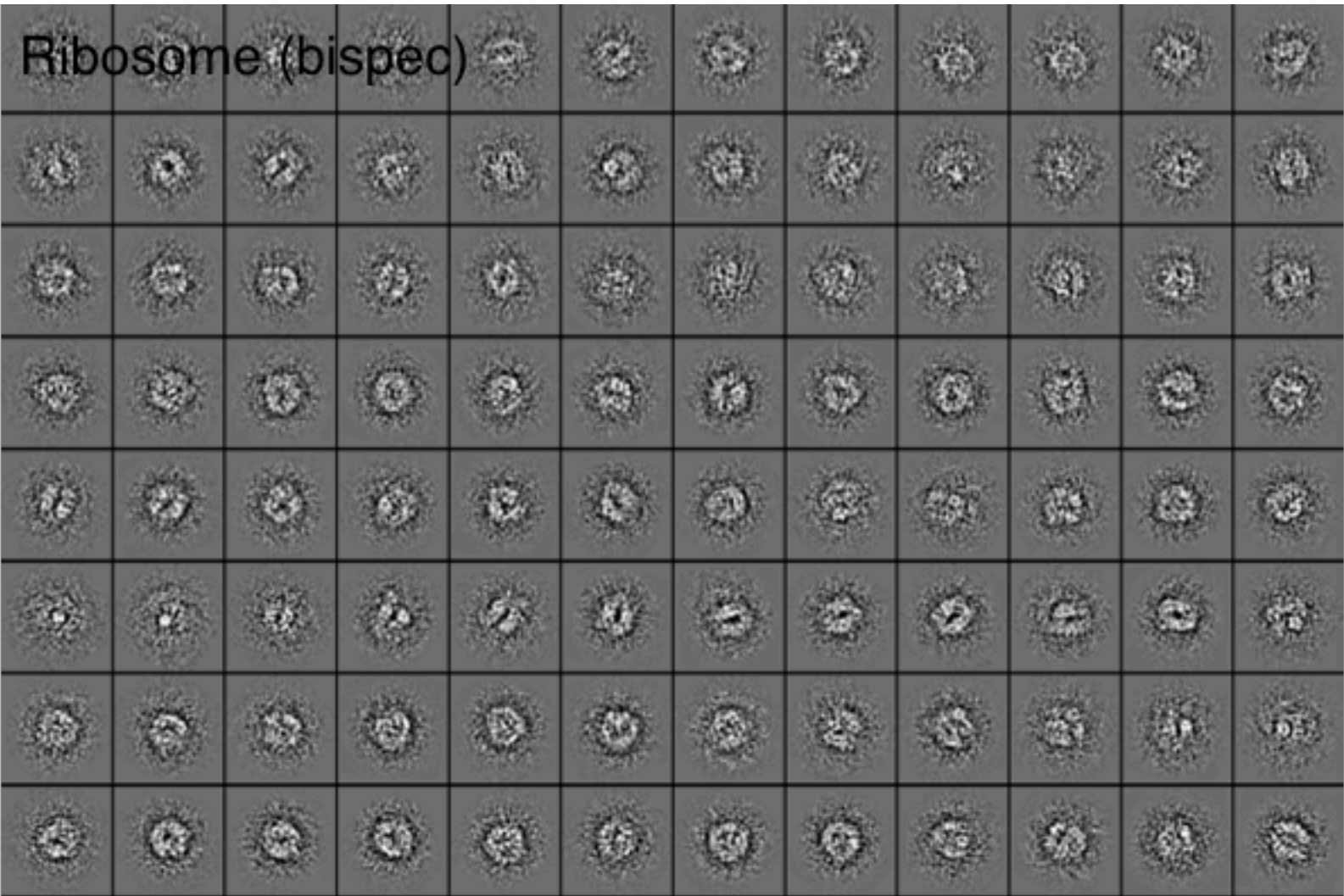




10,533 particles  
208x208 box size  
128 Classes

2h 39m  
1 laptop

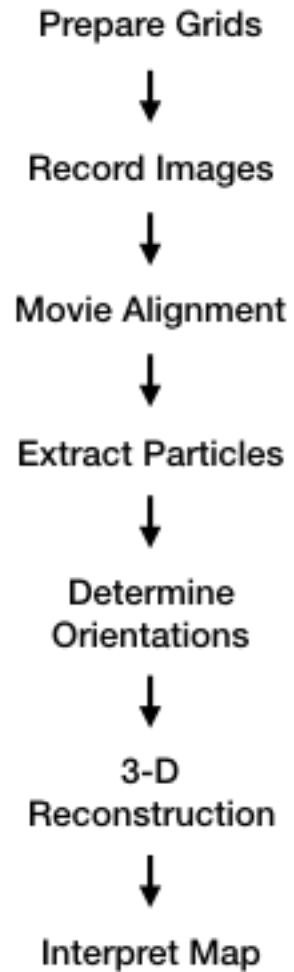




10,533 particles  
208x208 box size  
128 Classes

12 min  
1 laptop

# Single Particle Analysis



# Single Particle Analysis

Extract Particles



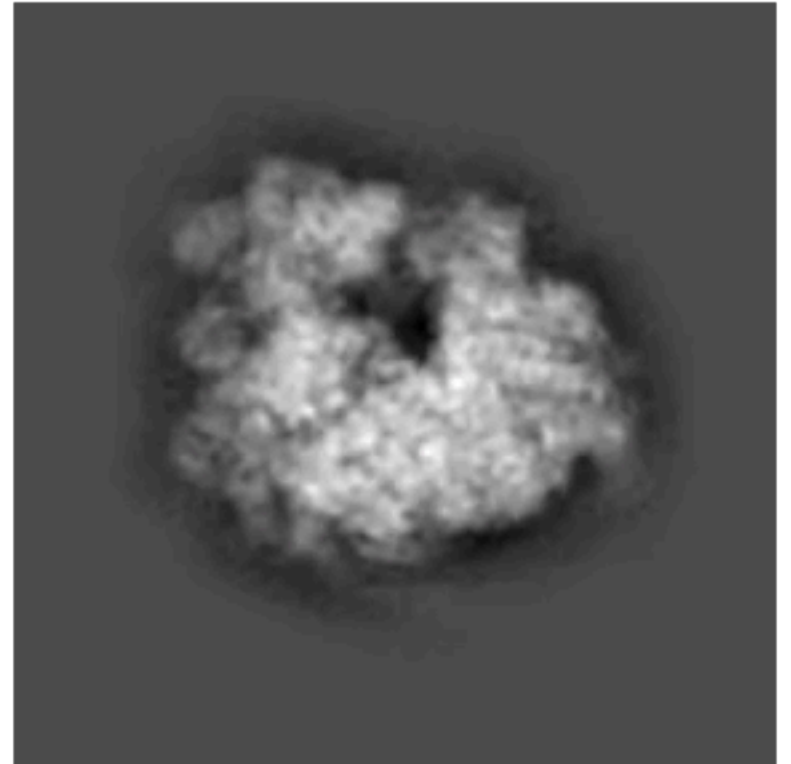
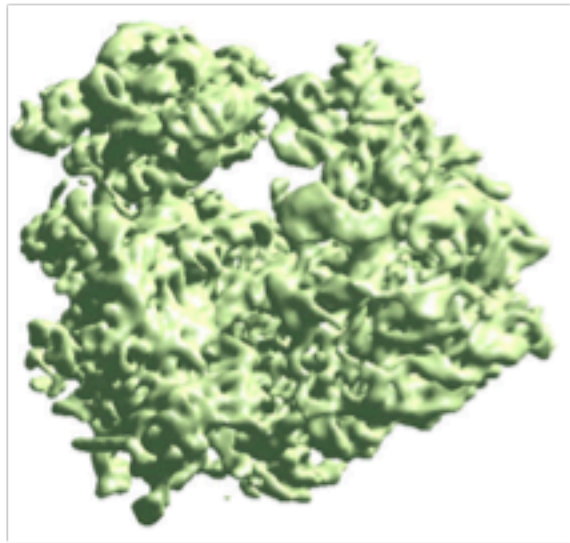
Determine  
Orientations



3-D  
Reconstruction

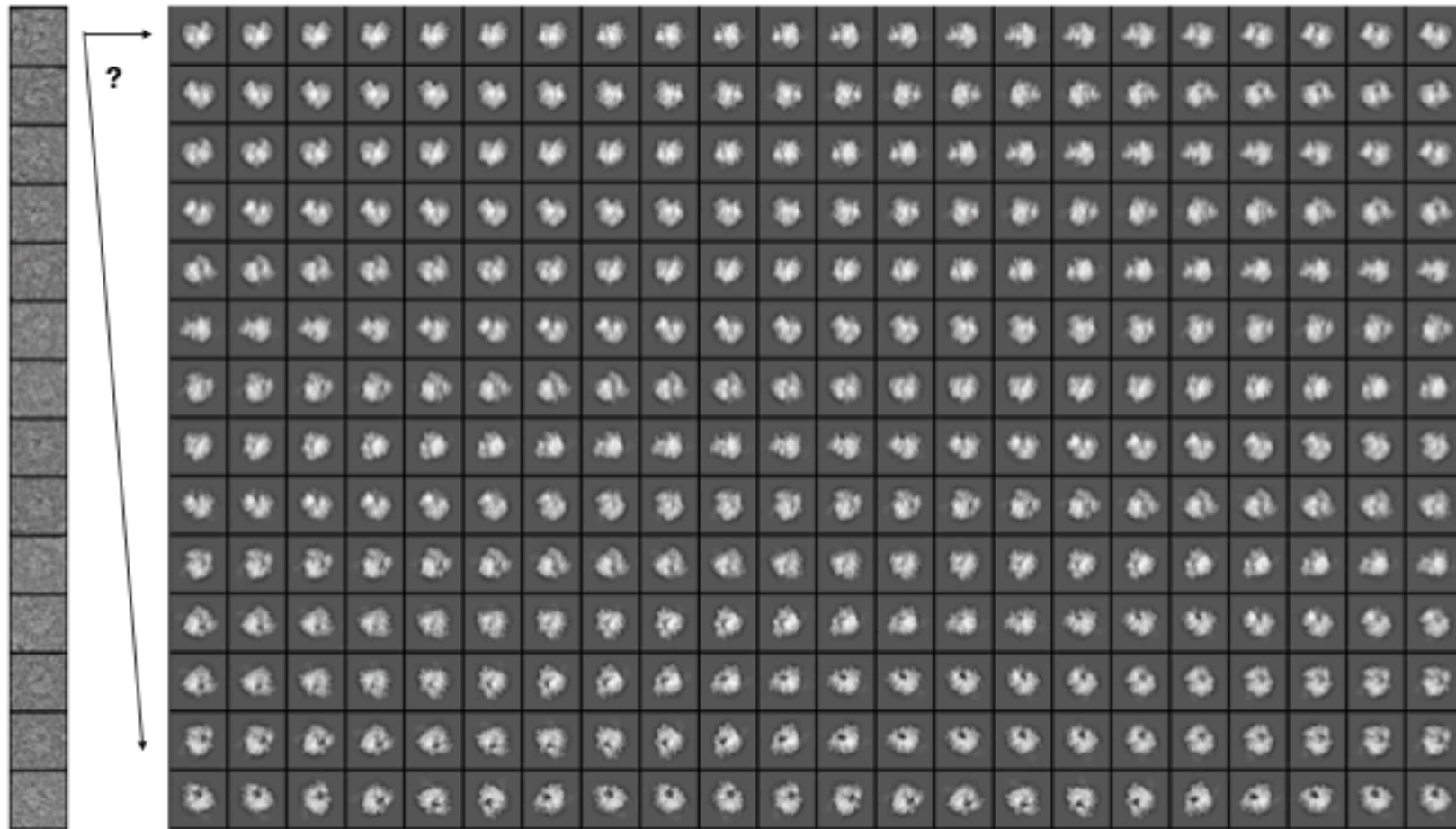


Interpret Map



**7.5° Angular Step**

# Determine Particle Orientations



Every particle vs every possible projection?

# Determine Particle Orientations

(Ribosome at  $\sim 4 \text{ \AA}$ )

- \* 30,000 particles
  - \* 3000 projections ( $\sim 2.6$  degree sampling)
  - \* 180 rotations
  - \* 10 x 10 translations (if centering is decent)
  - \* 65,536 pixels (256x256 image)
  - \* 20 FLOPS/pixel
- =  $2.1 \times 10^{18}$  FLOPS (2 exaFLOPS)
- @ 100 gigaFLOPS/s = 6,000 hours ( $\sim 240$  days on one typical workstation)
  - Clearly we need to be smarter...
  - (Actual EMAN2 time  $< 3$  hours on ... even more cleverness possible though)

# Determine Particle Orientations

- \* 30,000 particles
  - \* 3000 projections (~2.6 degree sampling)
  - \* 180 rotations
  - \* 10 x 10 translations (if centering is decent)
  - \* 65,536 pixels (256x256 image)
  - \* 20 FLOPS/pixel
- =  $2.1 \times 10^{18}$  FLOPS (2 exaFLOPS)
- @ 100 gigaFLOPS/s = 6,000 hours (~ 240 days on one typical workstation)
  - Clearly we need to be smarter...
  - (Actual EMAN2 time <3 hours ... even more cleverness possible though)



# Single Particle Analysis

Extract Particles



Determine  
Orientations



3-D  
Reconstruction

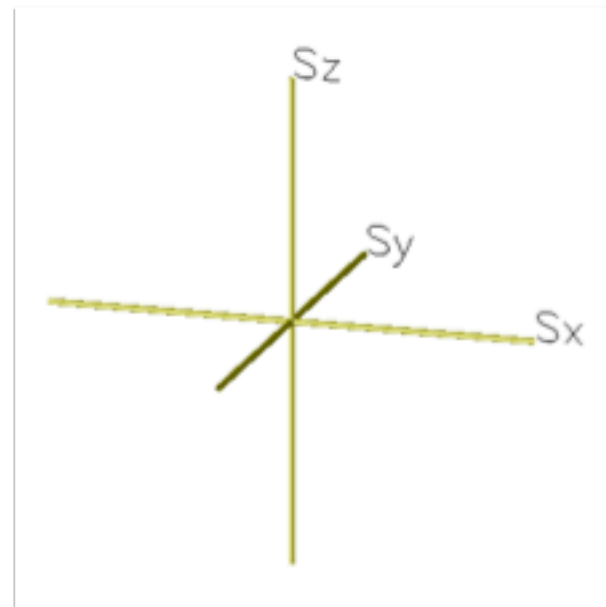


Interpret Map

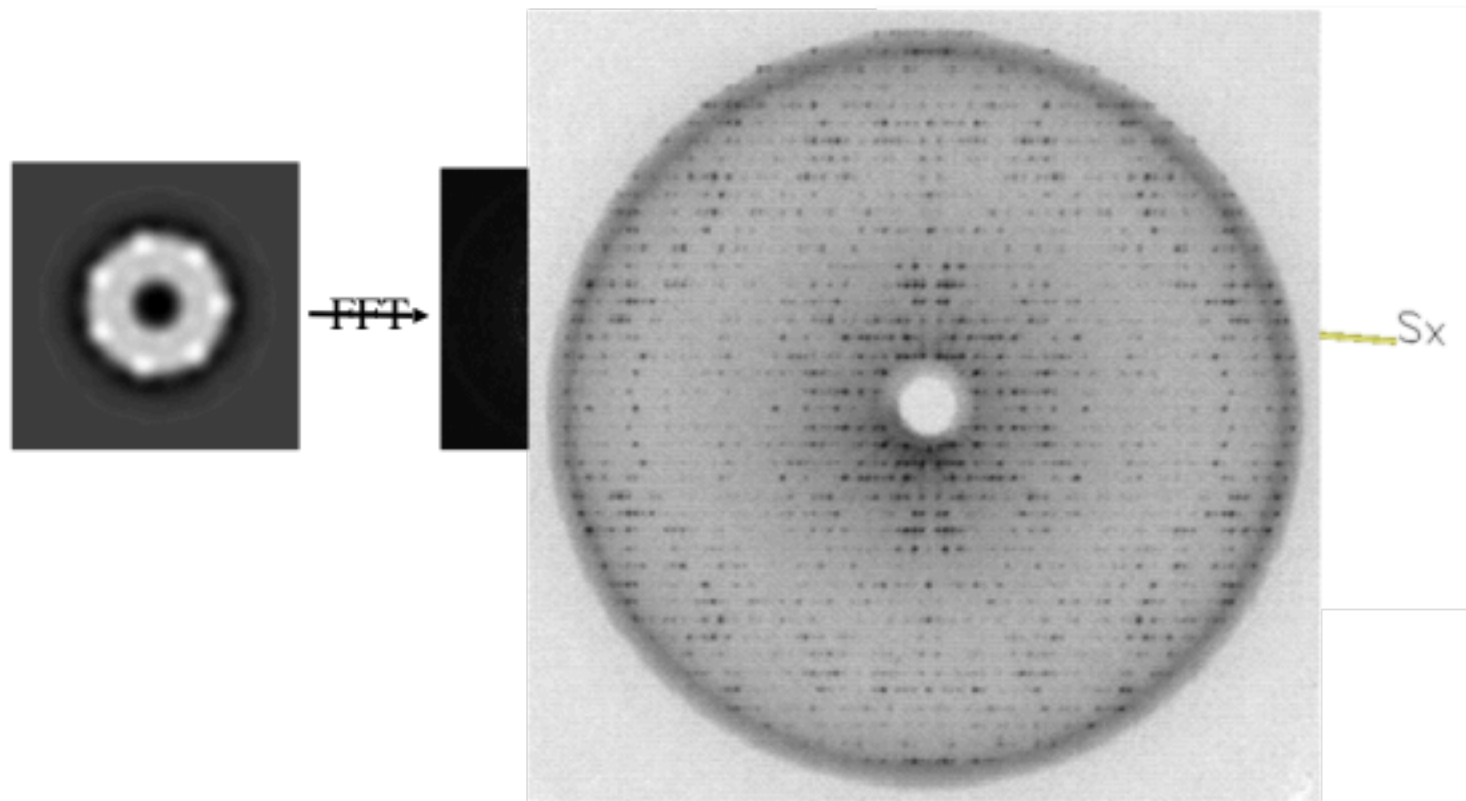
# Reconstruction Algorithms

- Back Projection
- Filtered Back Projection
- Direct Fourier Inversion
- SIRT
- SART
- ...

# Reconstruction & the Asymmetric Triangle

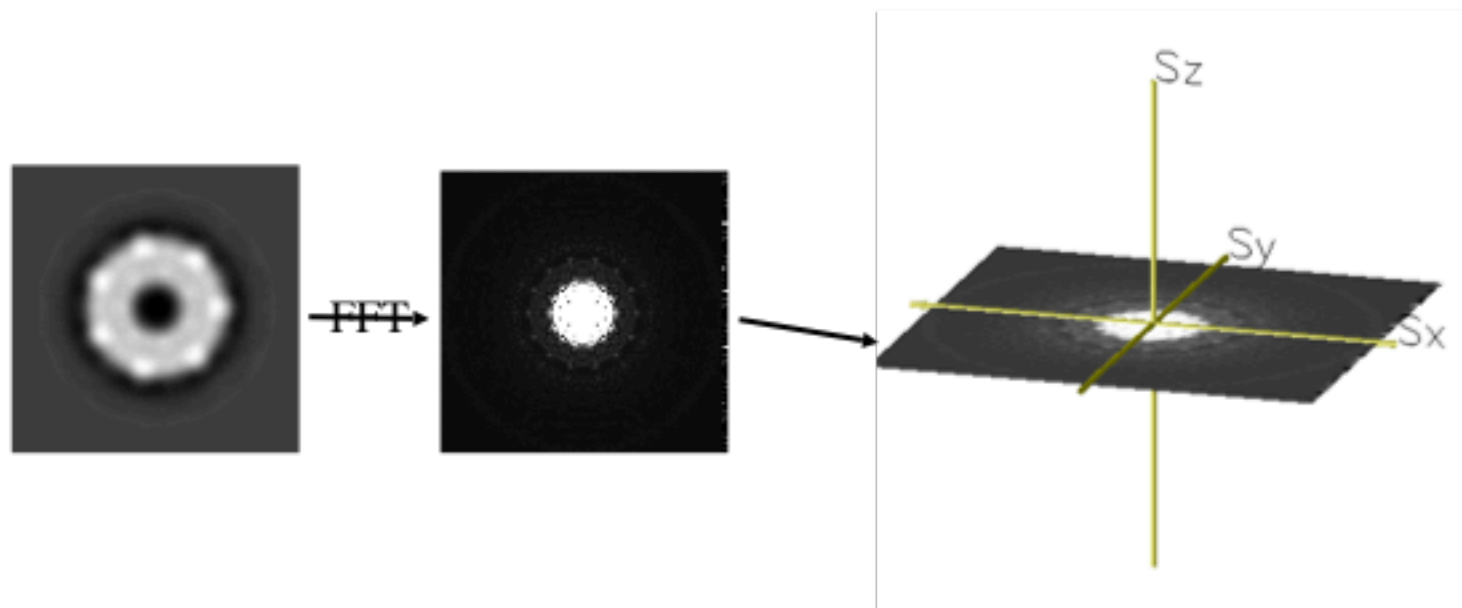


# Reconstruction & the Asymmetric Triangle

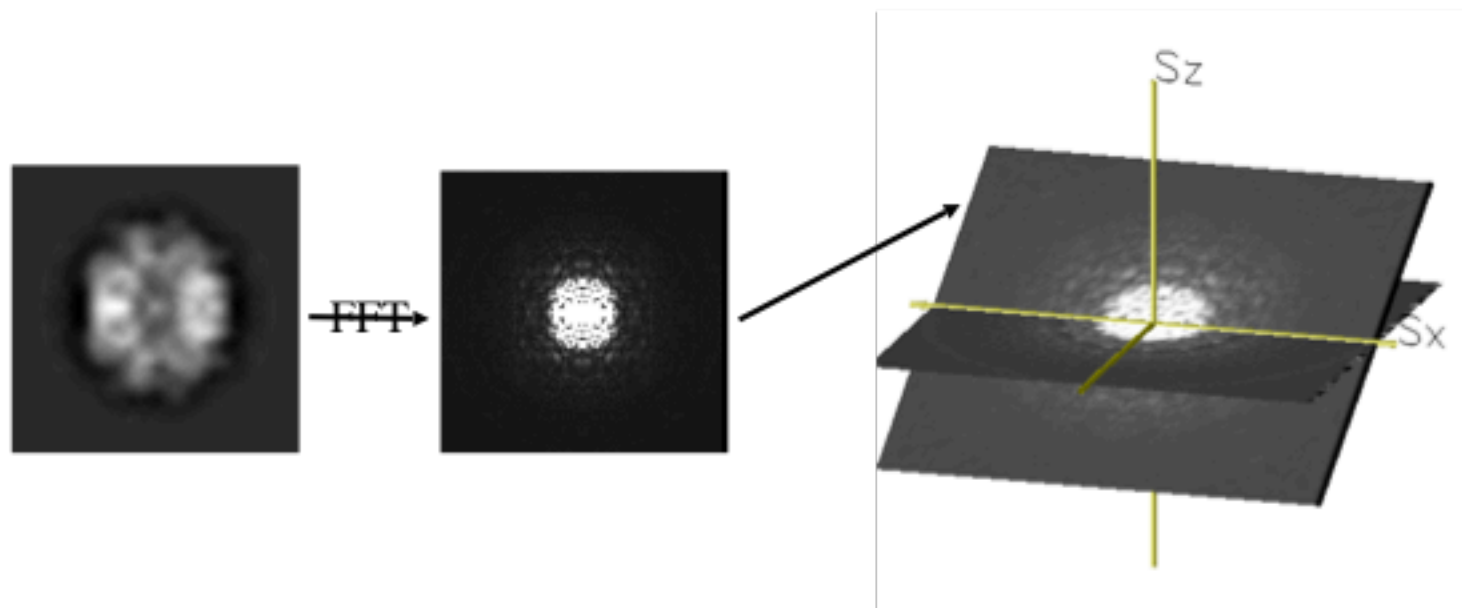


**Crystals have spots**  
**Particles are ~continuous and have phases!**

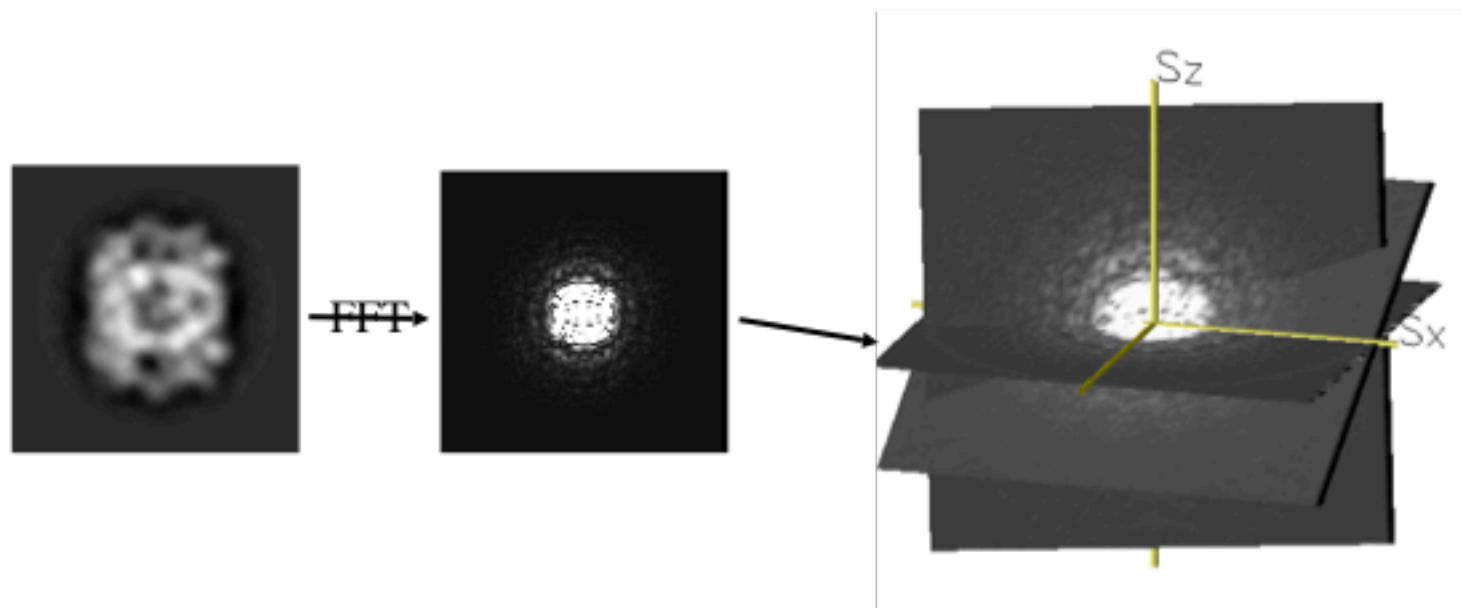
# Reconstruction & the Asymmetric Triangle



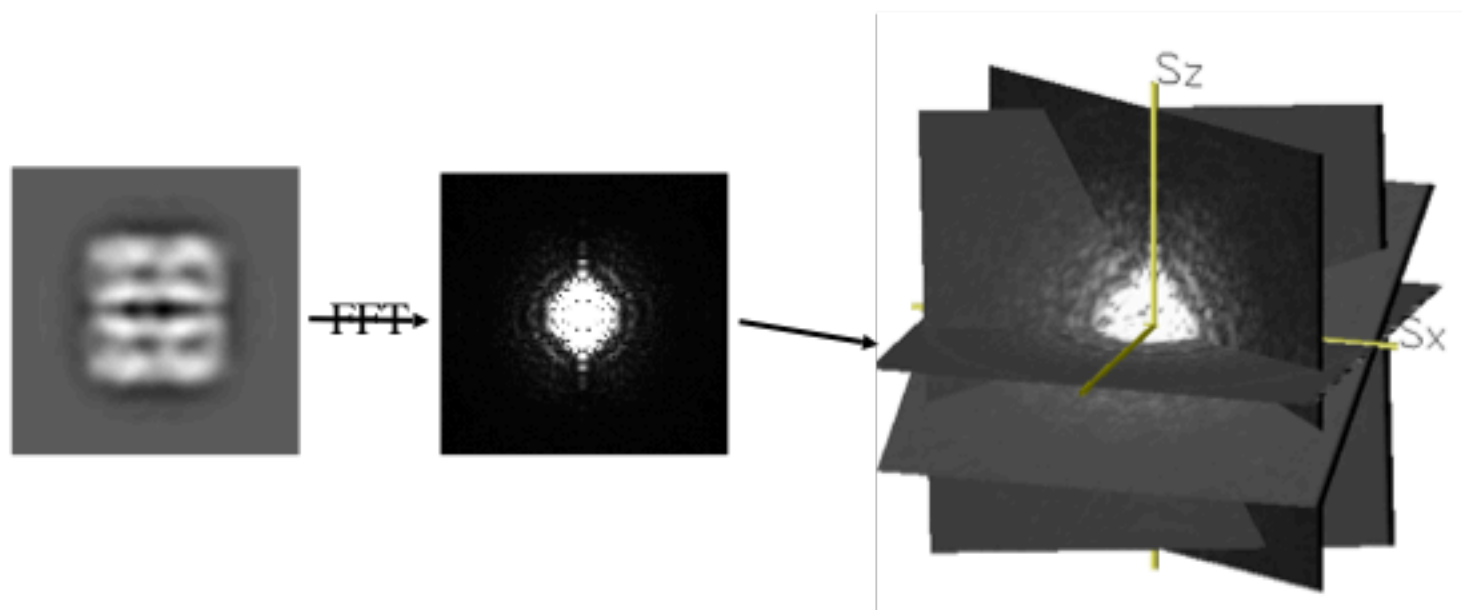
# Reconstruction & the Asymmetric Triangle



# Reconstruction & the Asymmetric Triangle

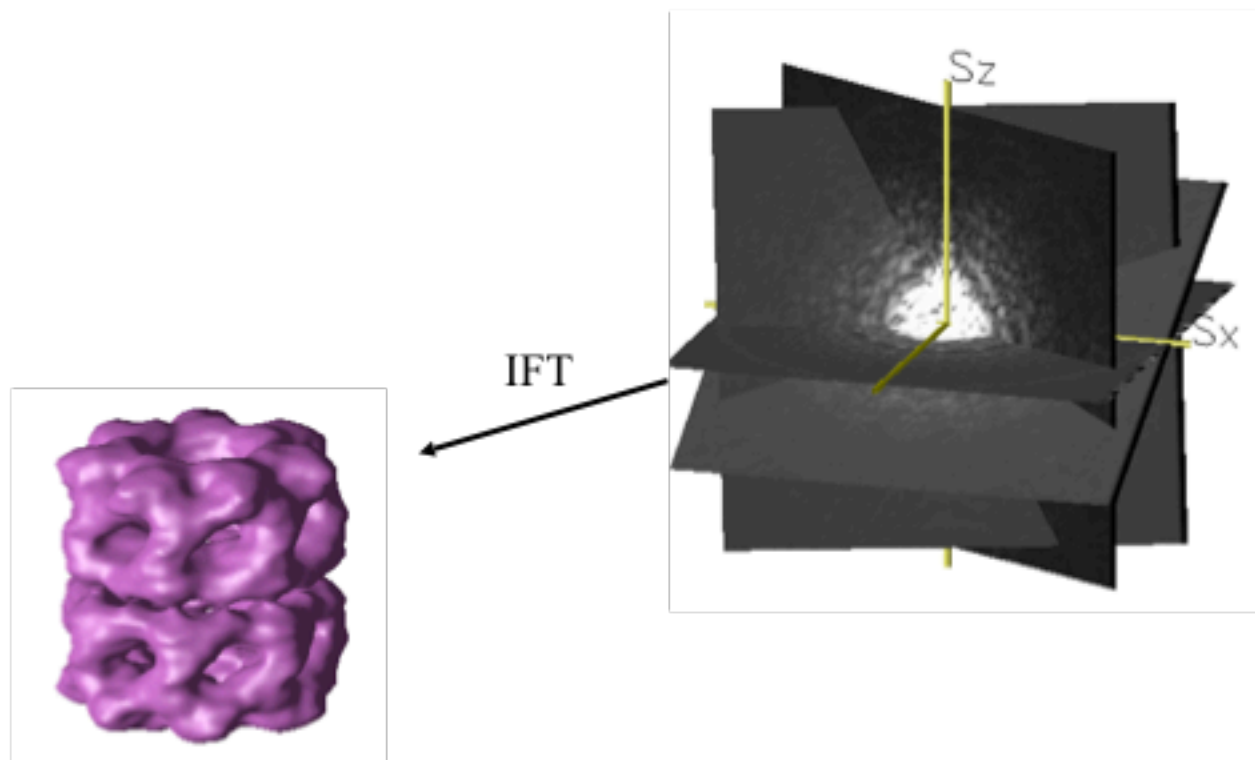


# Reconstruction & the Asymmetric Triangle





# Reconstruction & the Asymmetric Triangle



# One problem...

- To determine the particle orientations, we needed a 3-D reference
- catch 22
- All current CryoEM Single Particle solutions are iterative!