- Copy demo data from DVD
- Handouts
- Registration desk for talk signup
- Friday Dinner & Taxi
- Overflow in 216
- Teaching classroom

Single Particle Refinement Strategies & Introduction to EMAN2

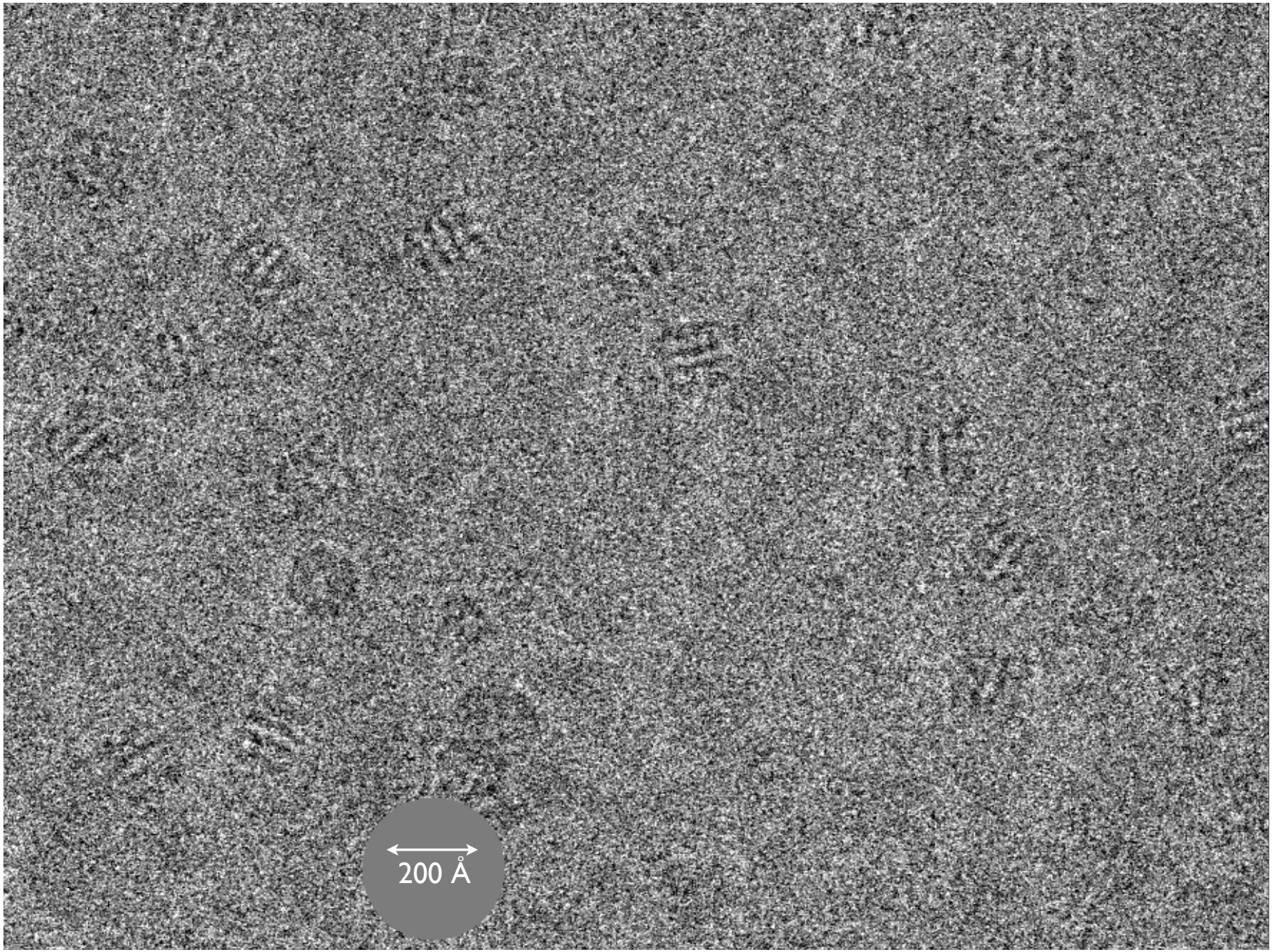
Steve Ludtke
National Center for Macromolecular Imaging
Baylor College of Medicine

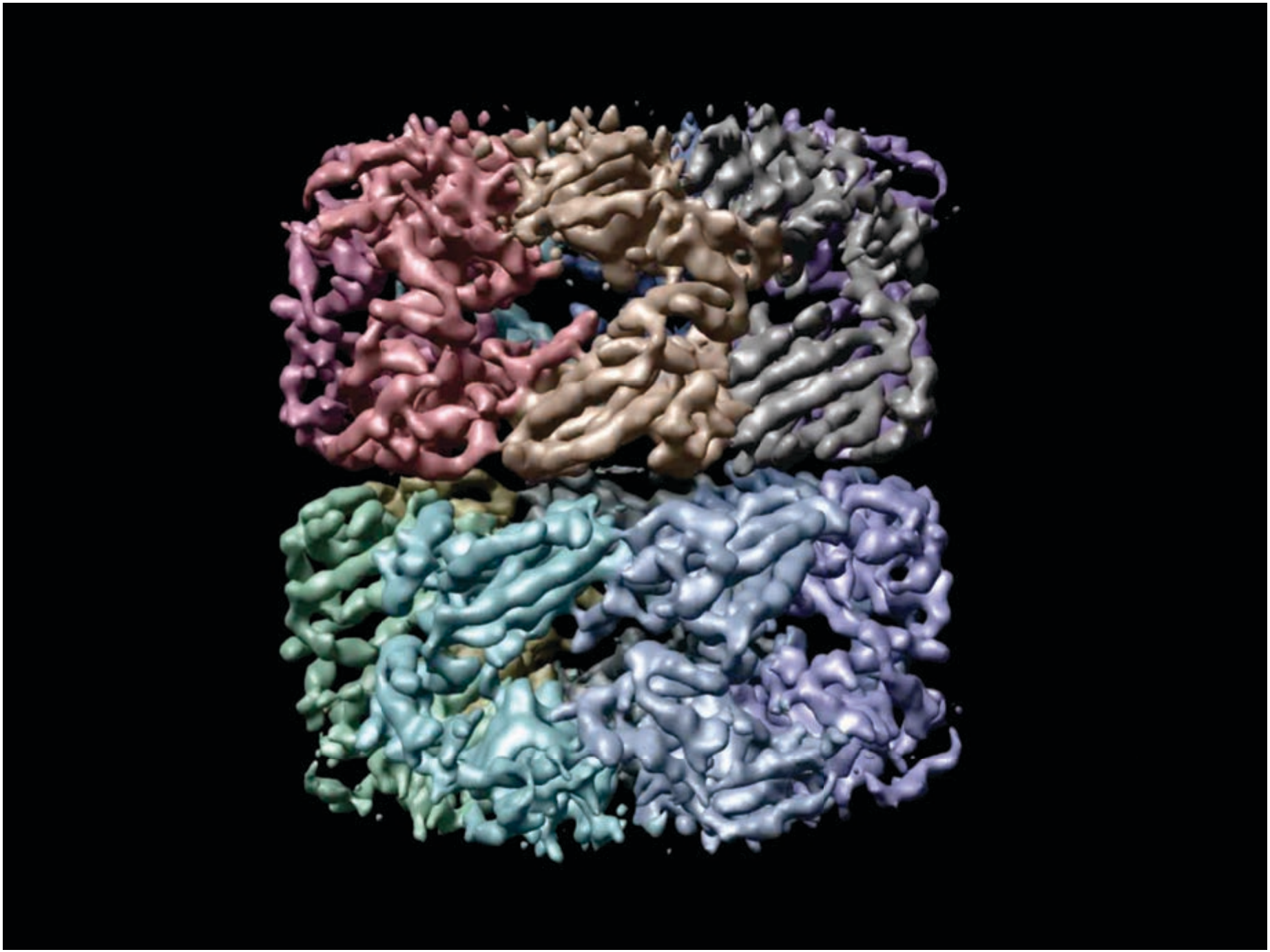
BCM Baylor College of Medicine

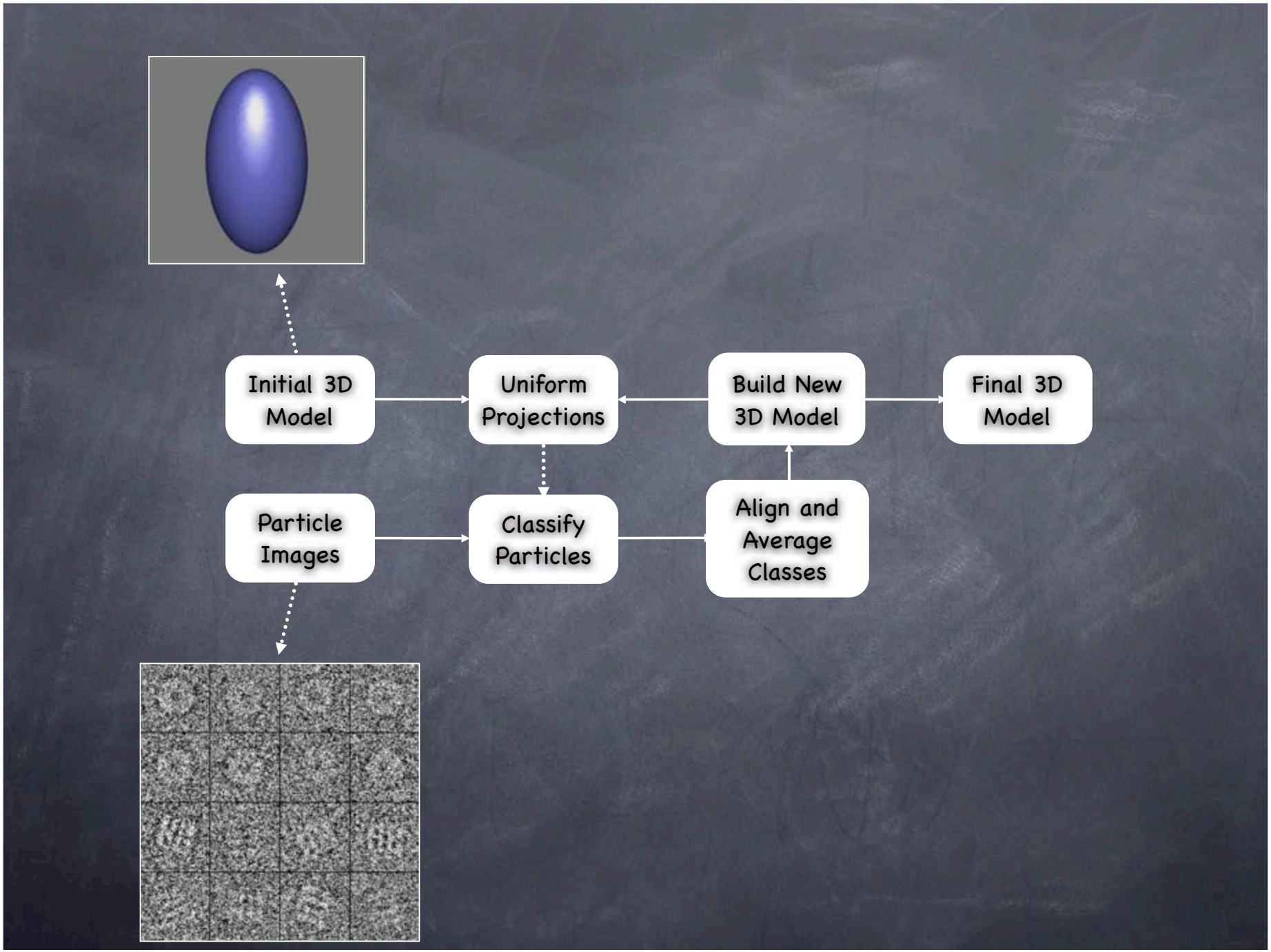
NCRR

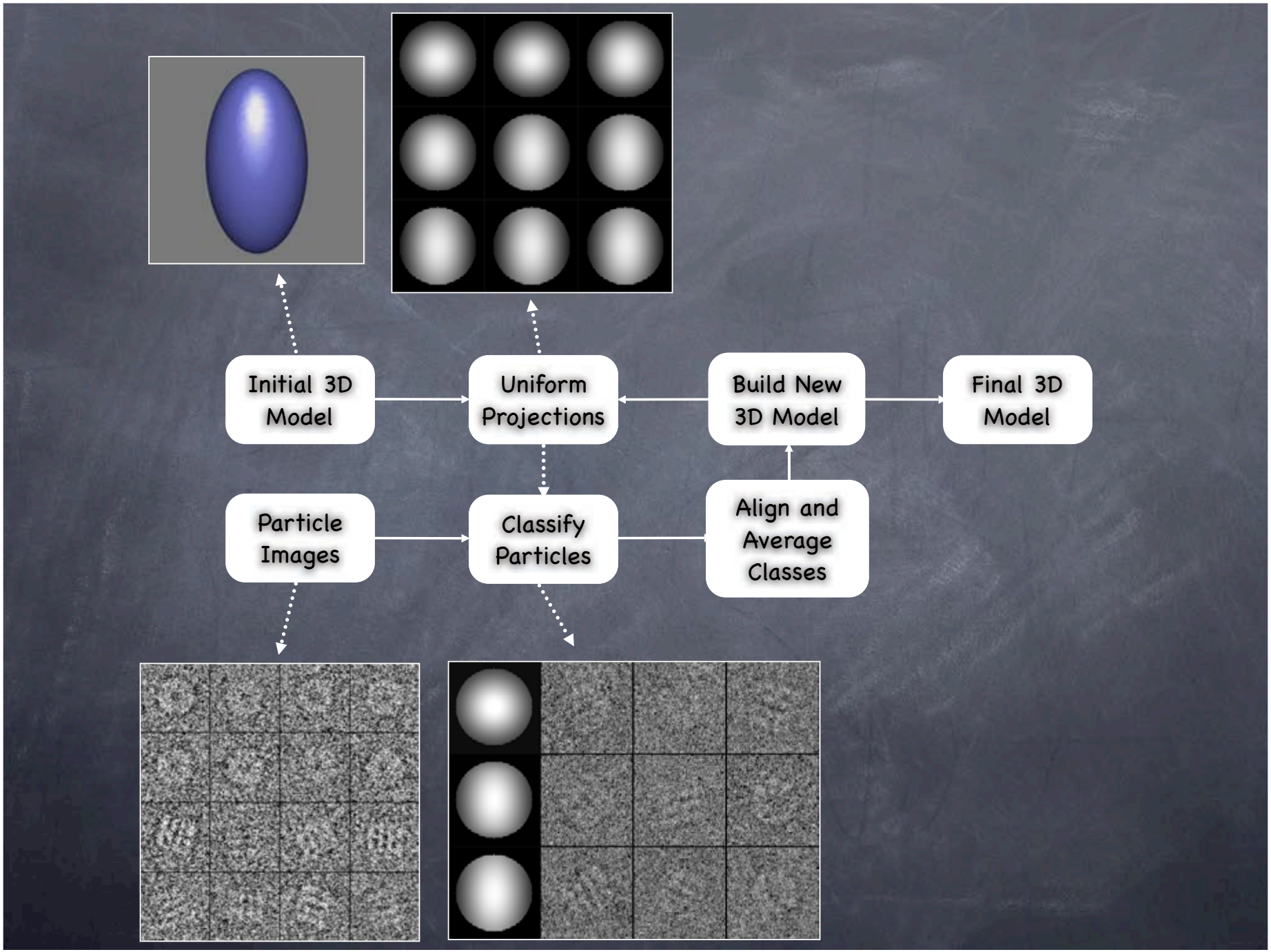
NCMI

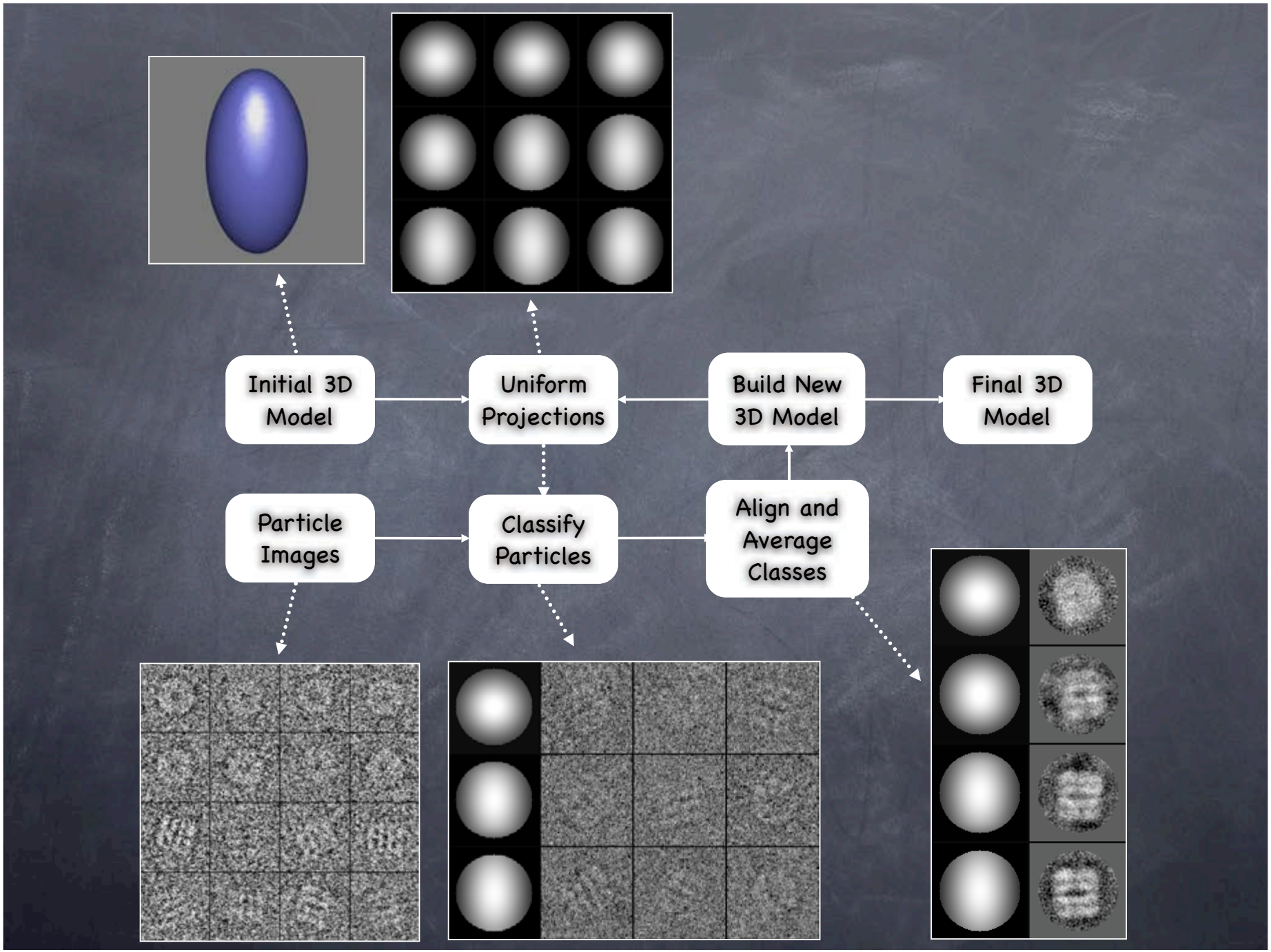
NIH Roadmap

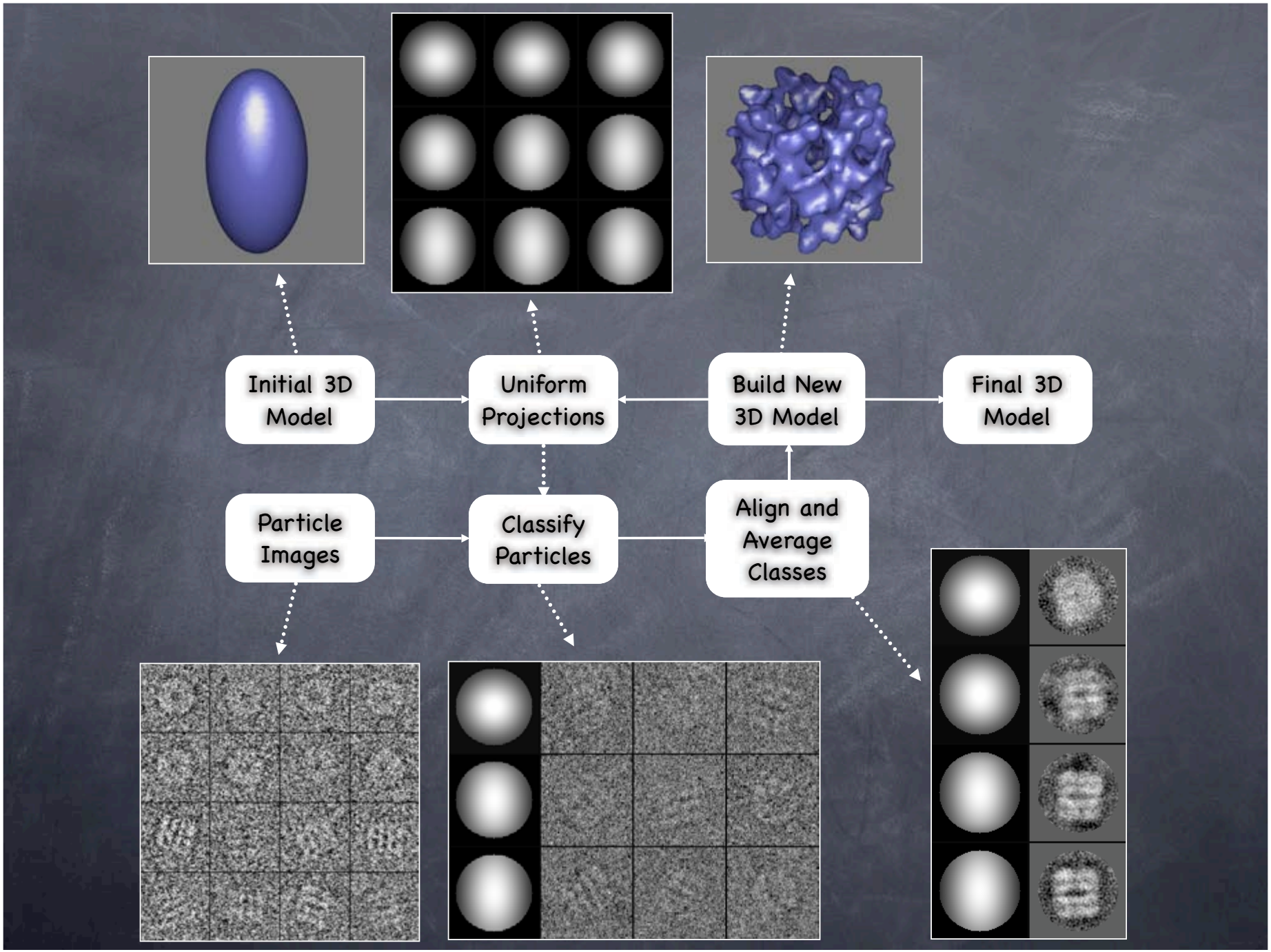

200 Å

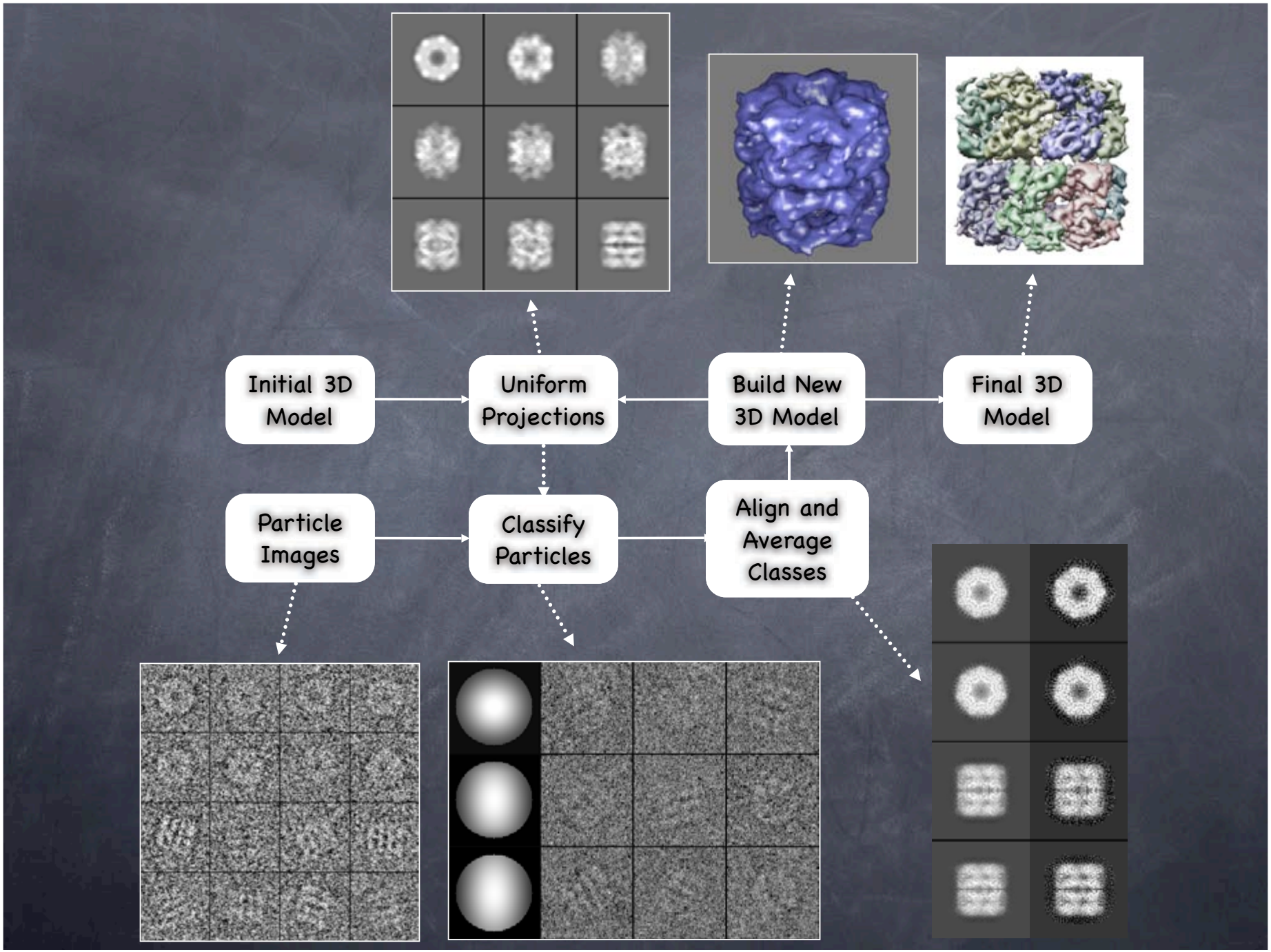


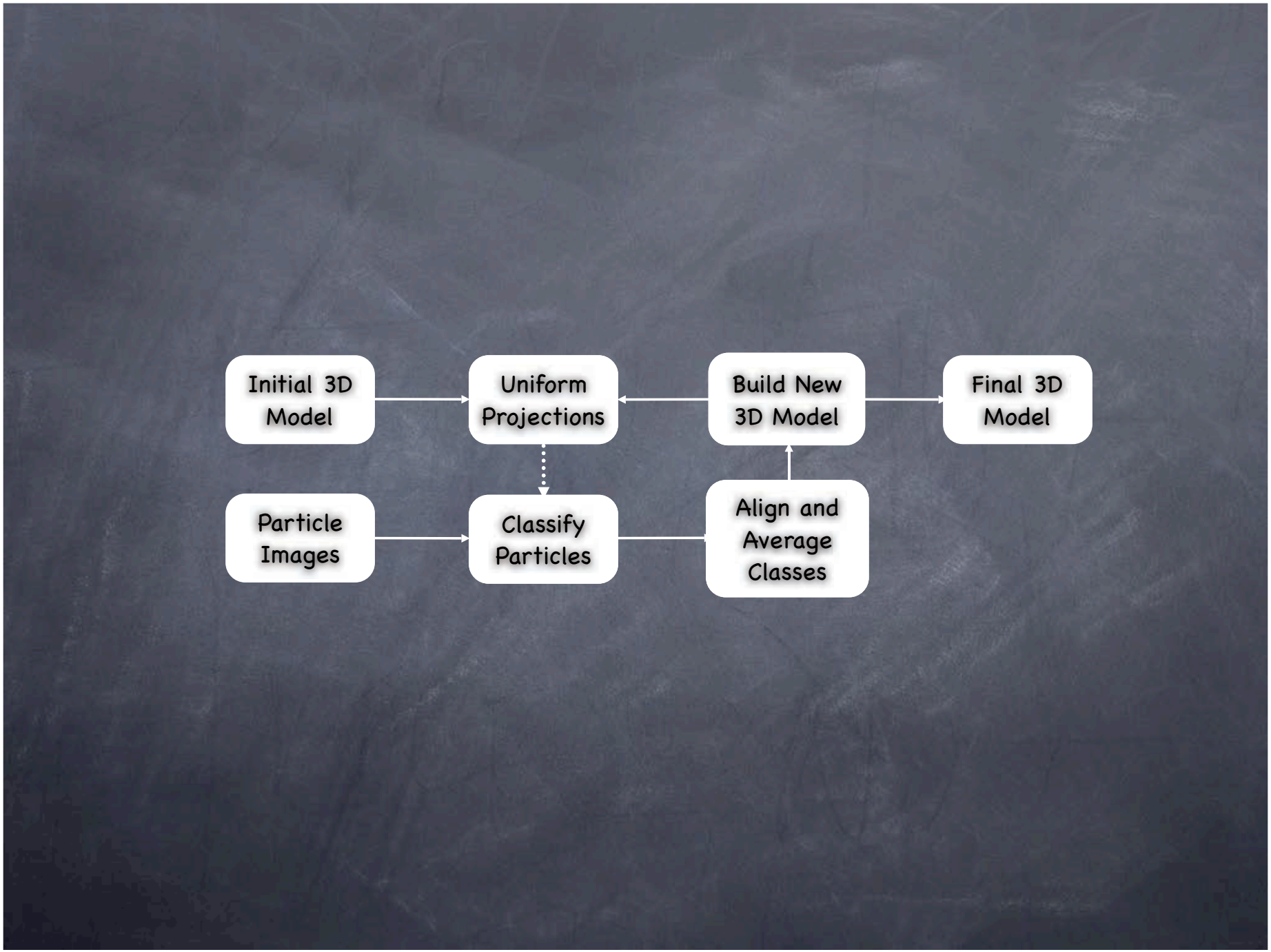

Initial 3D Model
Uniform Projections
Build New 3D Model
Final 3D Model
Particle Images
Classify Particles
Align and Average Classes

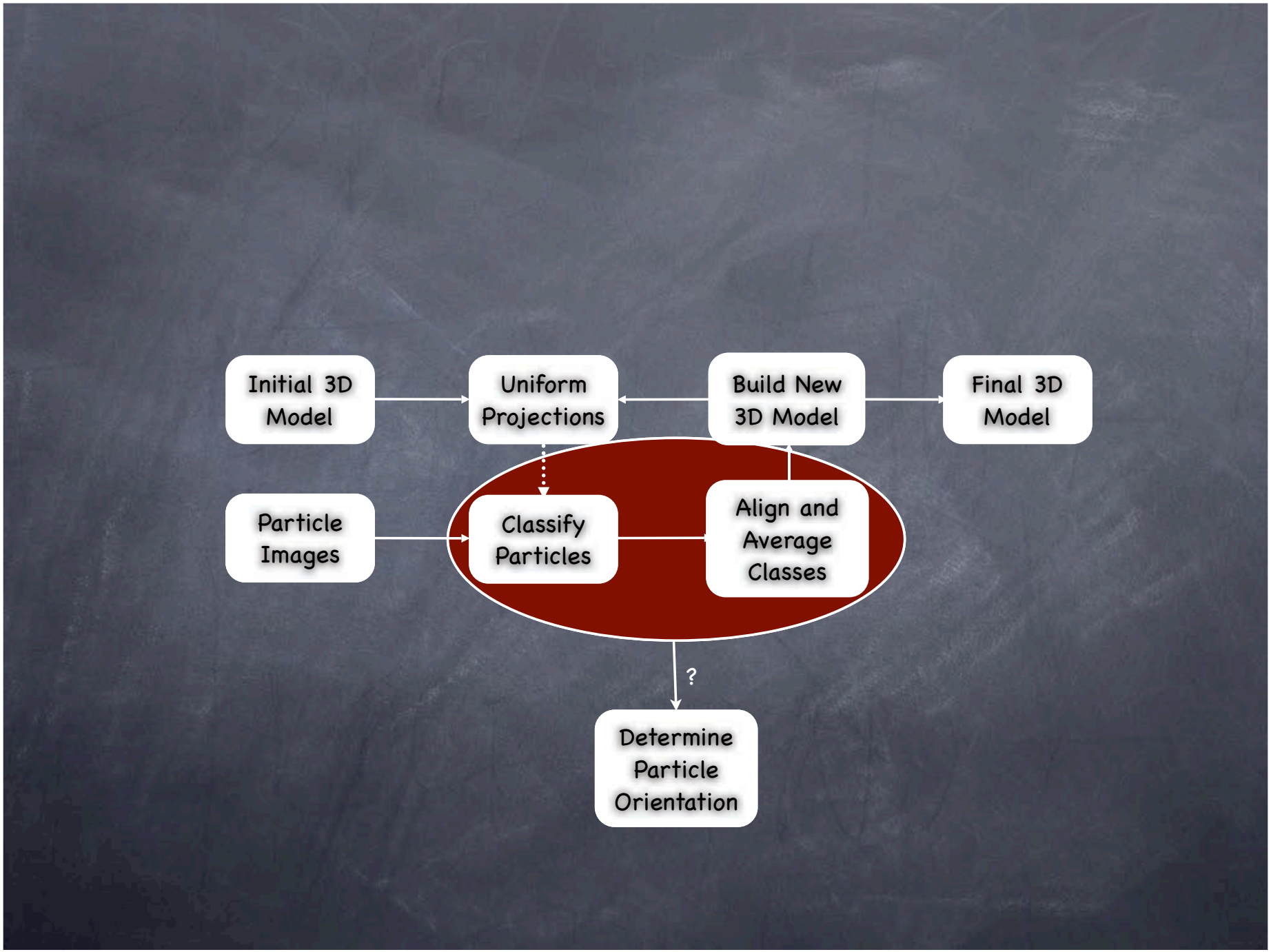

Initial 3D Model
Uniform Projections
Build New 3D Model
Final 3D Model
Particle Images
Classify Particles
Align and Average Classes

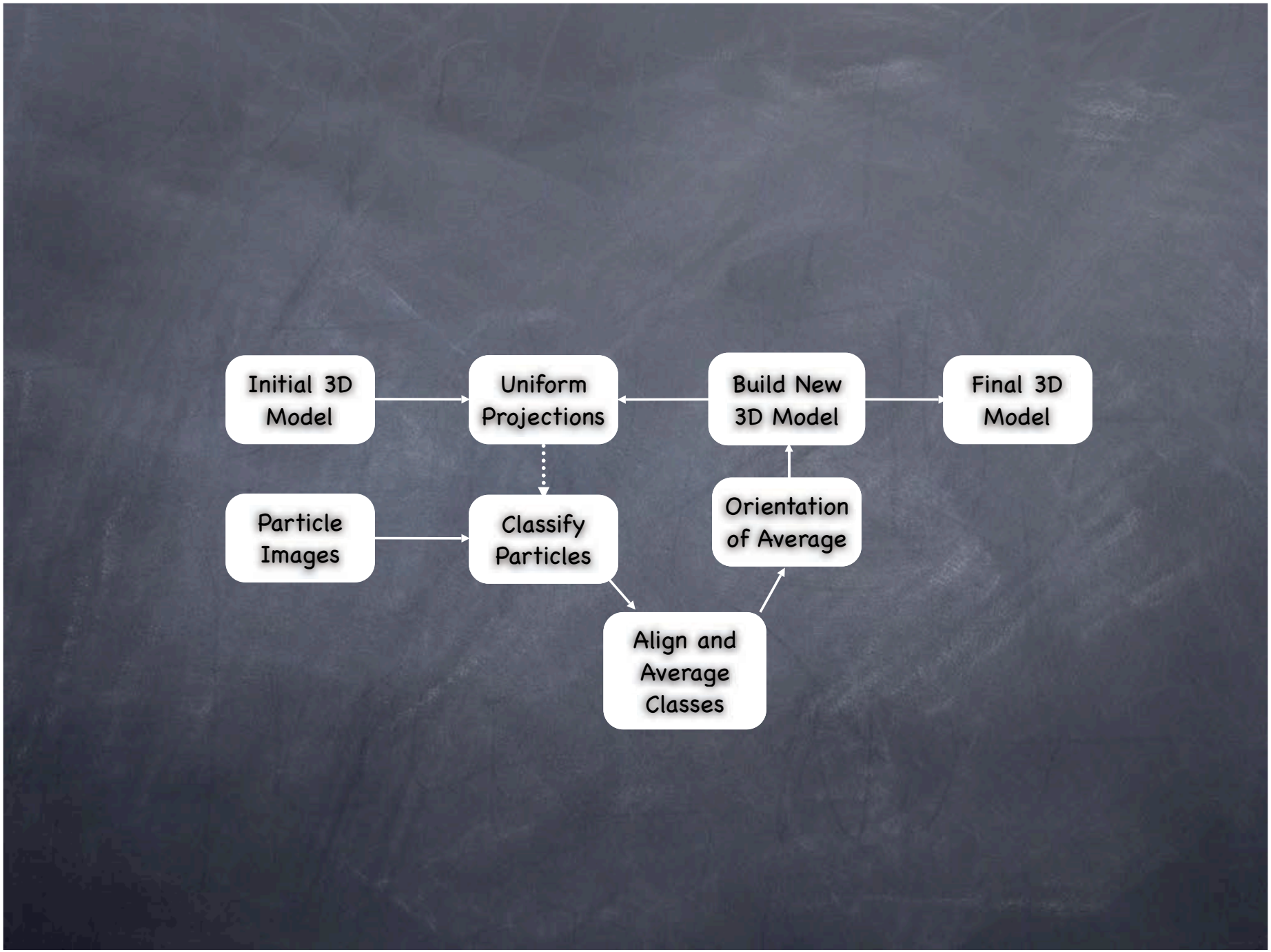

Initial 3D Model
Uniform Projections
Build New 3D Model
Final 3D Model
Particle Images
Classify Particles
Align and Average Classes

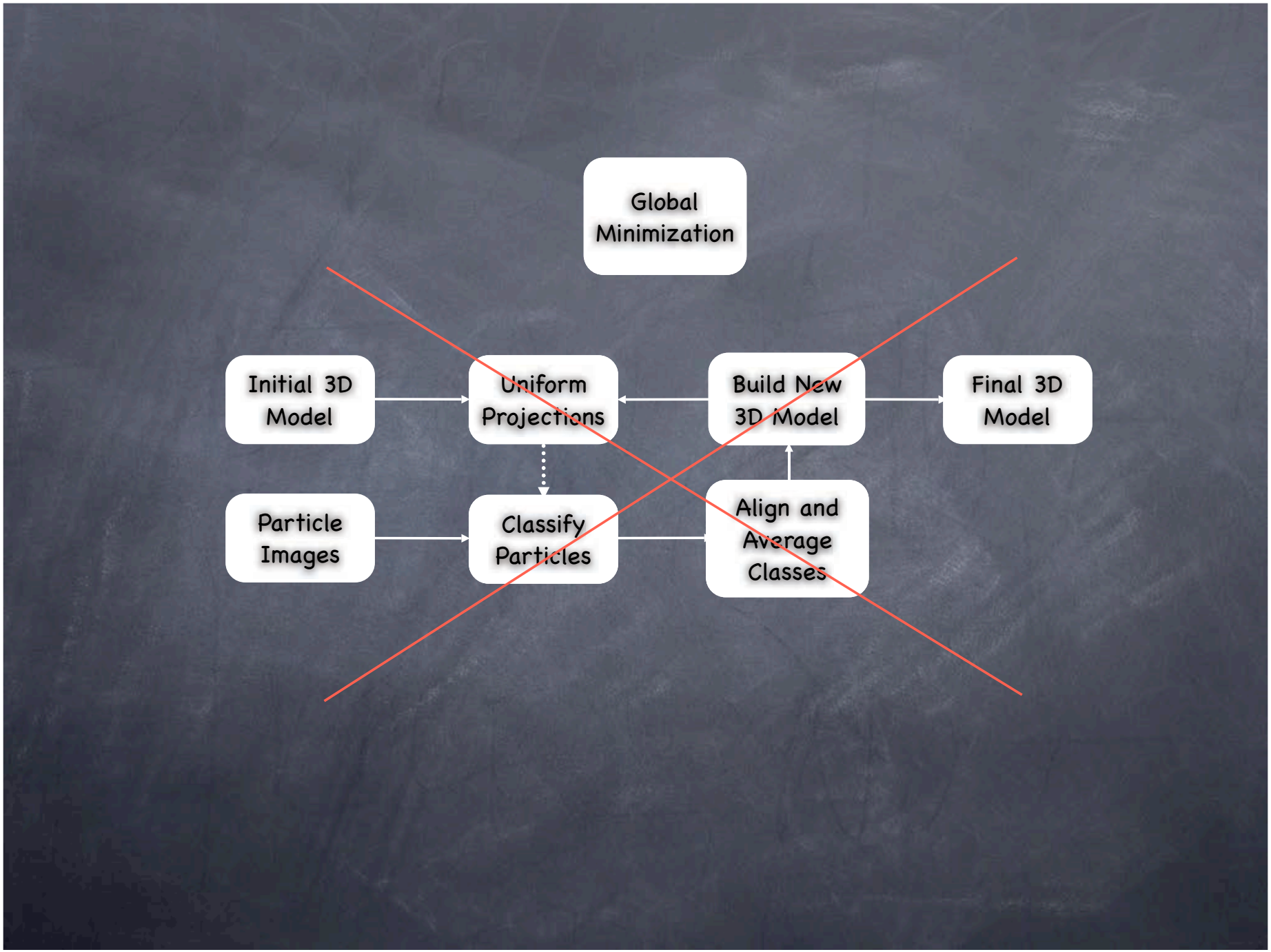

Initial 3D Model
Uniform Projections
Build New 3D Model
Final 3D Model
Particle Images
Classify Particles
Align and Average Classes


Initial 3D Model
Uniform Projections
Build New 3D Model
Final 3D Model
Particle Images
Classify Particles
Align and Average Classes


Initial 3D Model
Uniform Projections
Build New 3D Model
Final 3D Model
Particle Images
Classify Particles
Align and Average Classes

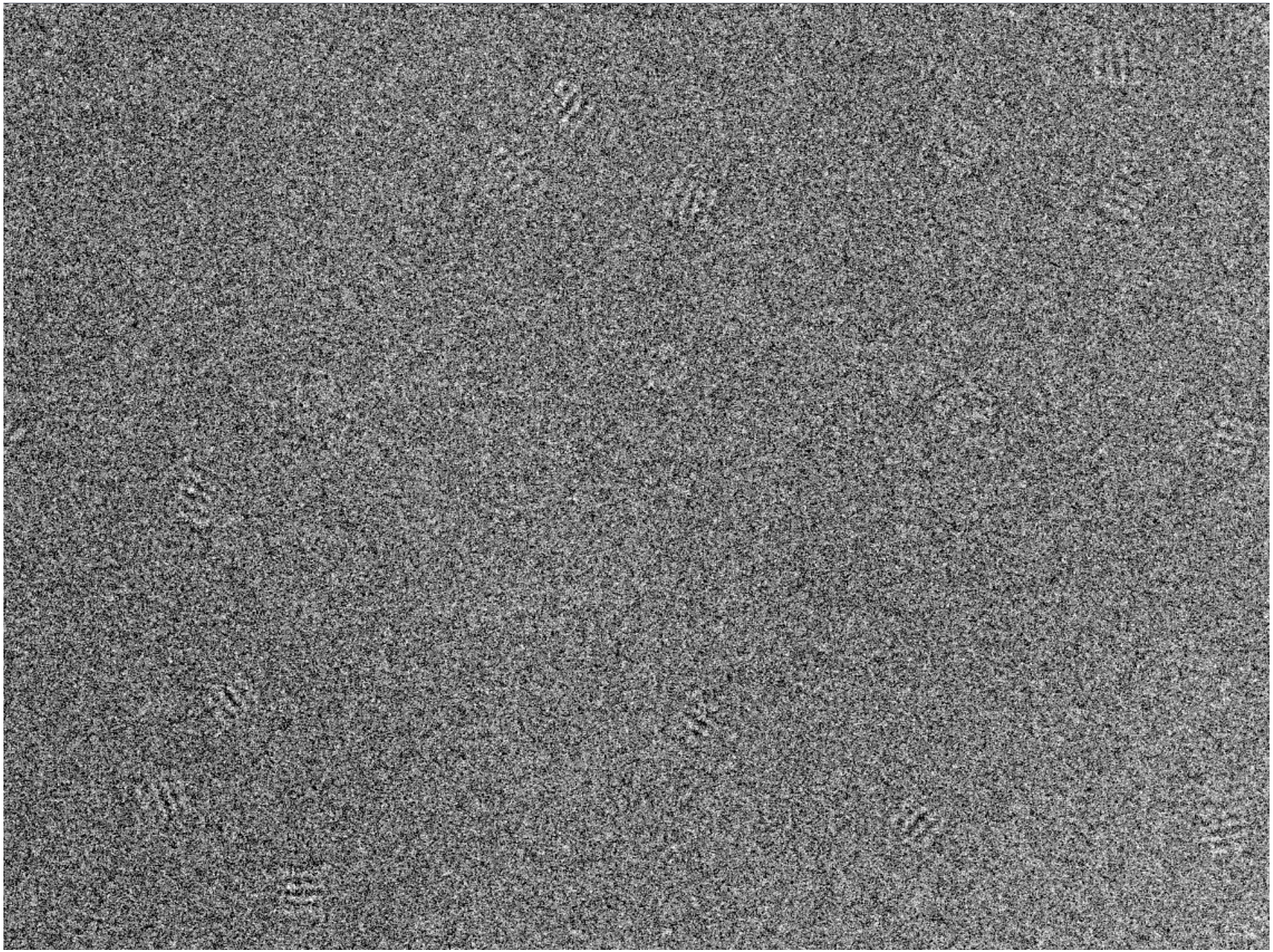

Initial 3D Model
Uniform Projections
Build New 3D Model
Final 3D Model
Particle Images
Classify Particles
Align and Average Classes
?
Determine Particle Orientation

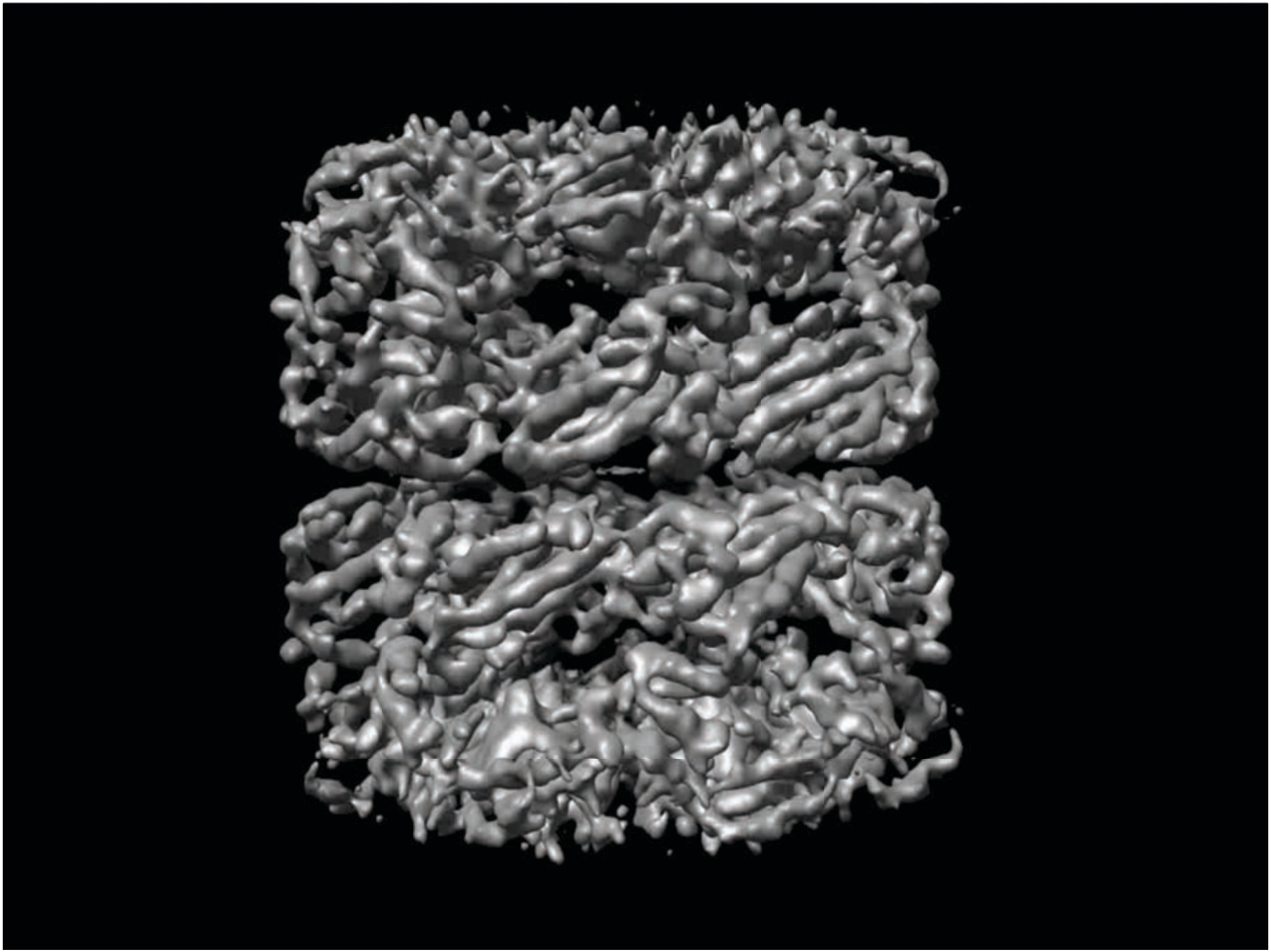

Initial 3D Model
Uniform Projections
Build New 3D Model
Final 3D Model
Particle Images
Classify Particles
Orientation of Average
Align and Average Classes

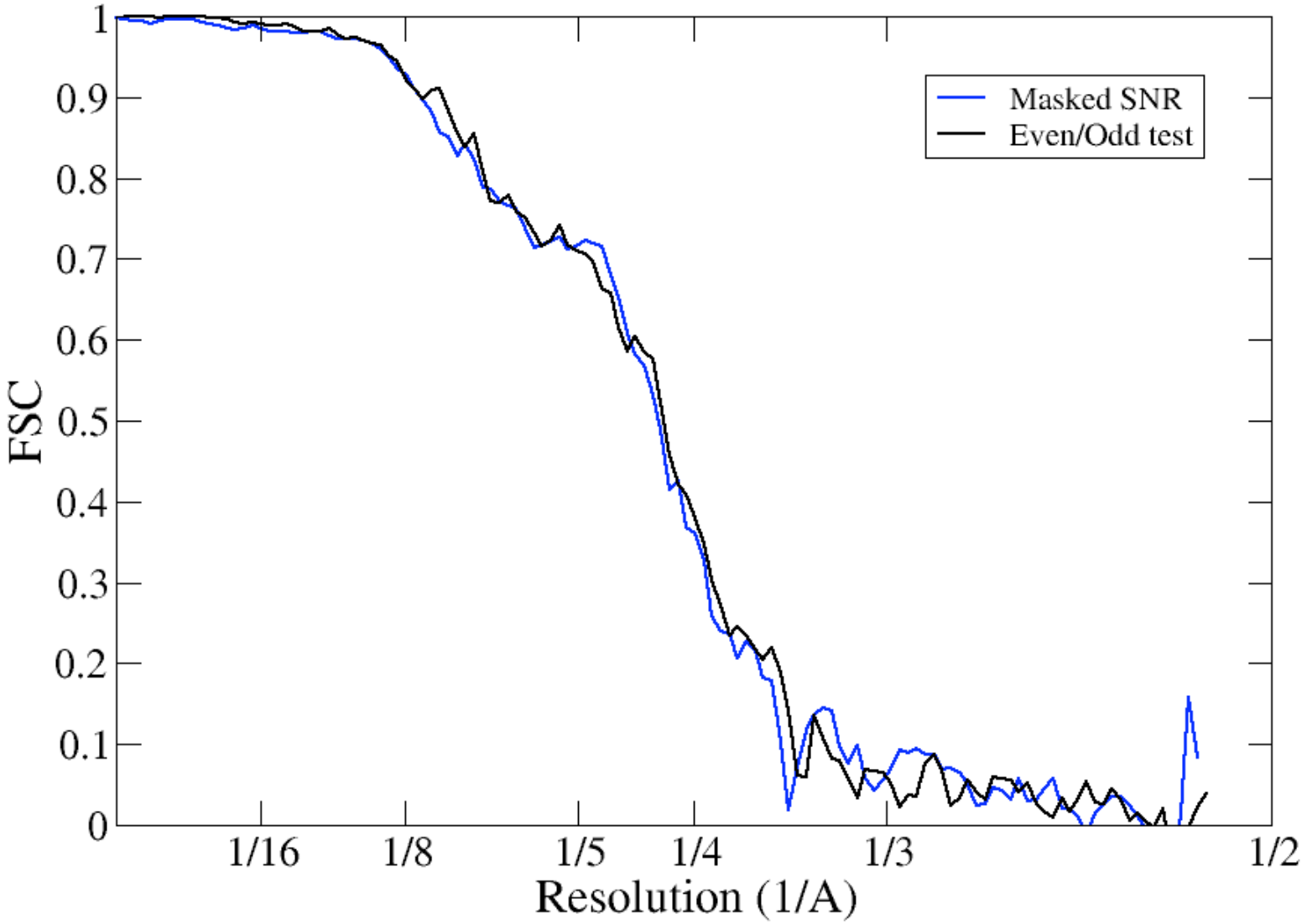

Global Minimization
Initial 3D Model
Uniform Projections
Build New 3D Model
Final 3D Model
Particle Images
Classify Particles
Align and Average Classes

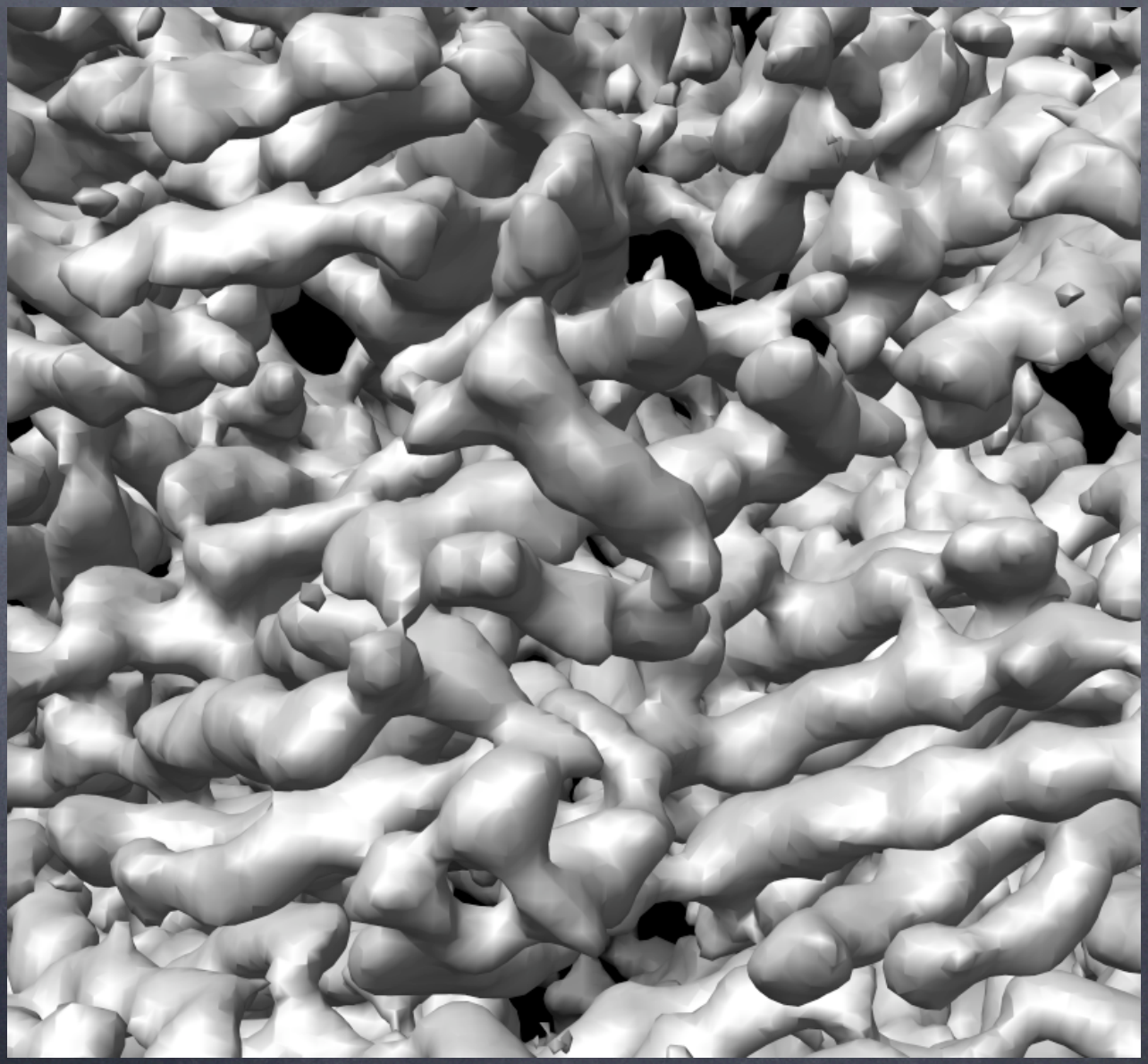
GroEL 2005

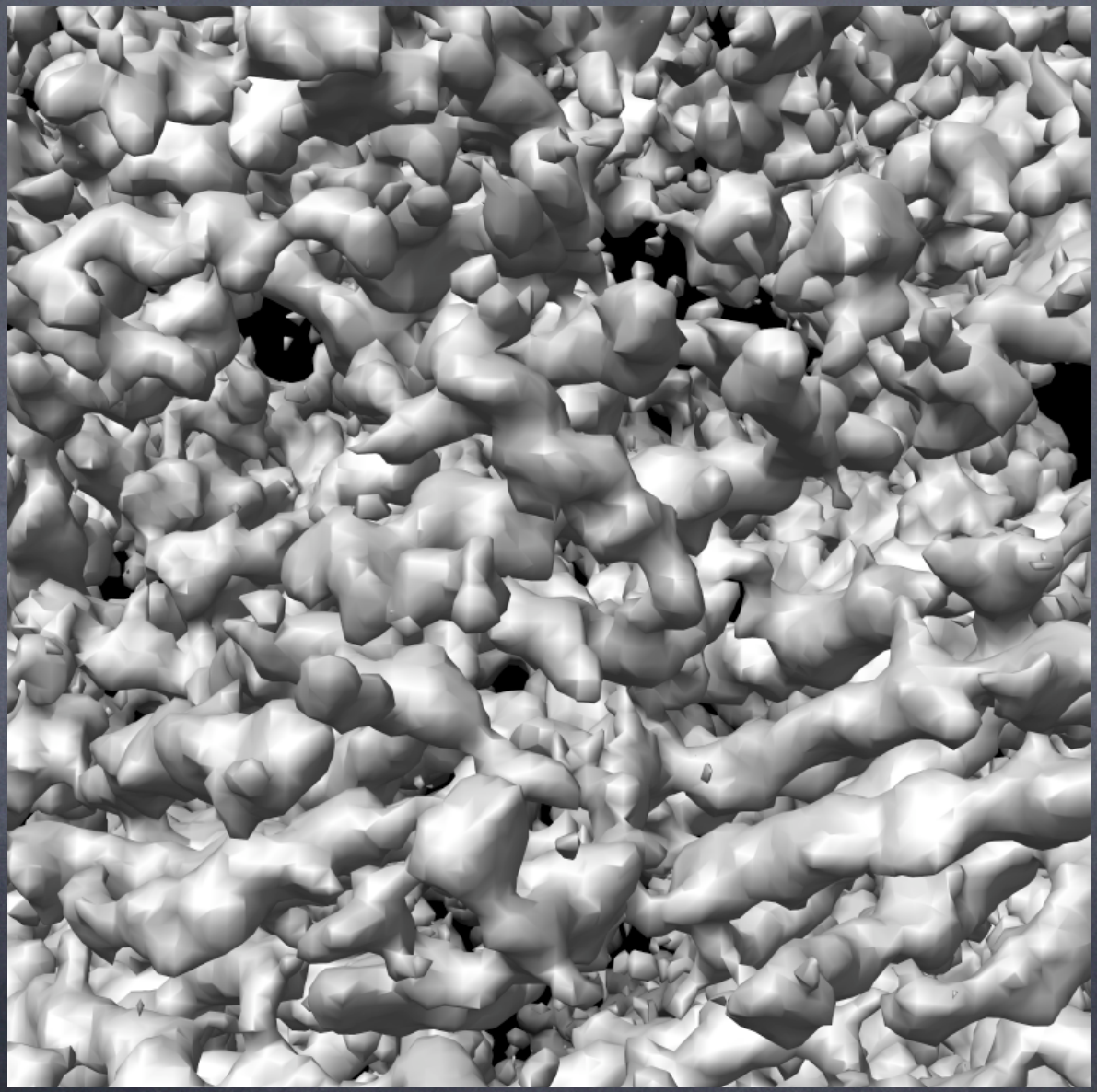
- Native, unliganded GroEL, no ATP/ADP (?)
- JEOL 3000SFF (Yoshi-style) at LHe temp
- 6 microscopy sessions, Film
- 825 micrographs, Nikon 9000 @ 6.35 µm scan step
- 60k mag → 1.06 Å/pix
- 135 micrographs used → 20,401 particles

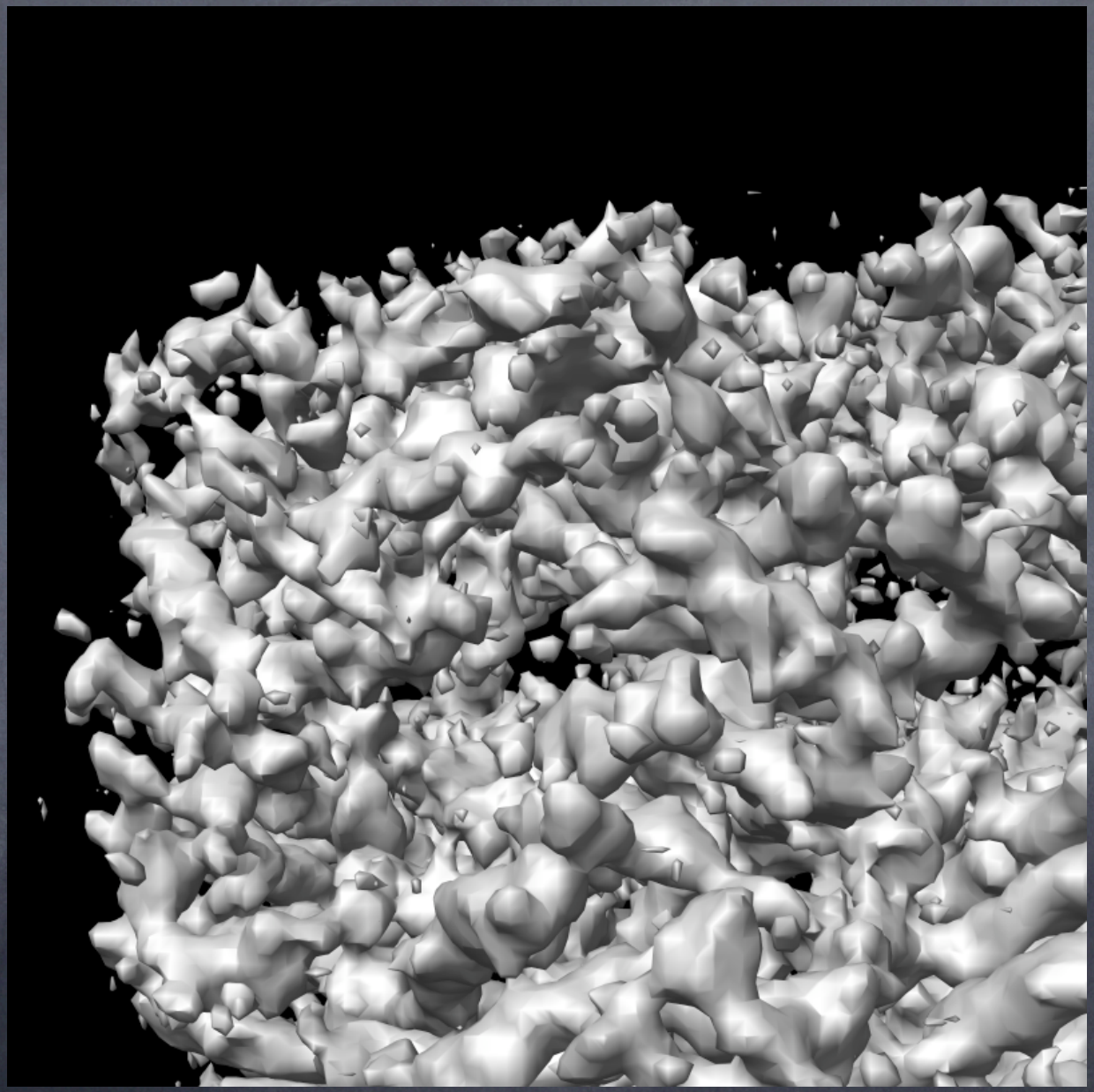


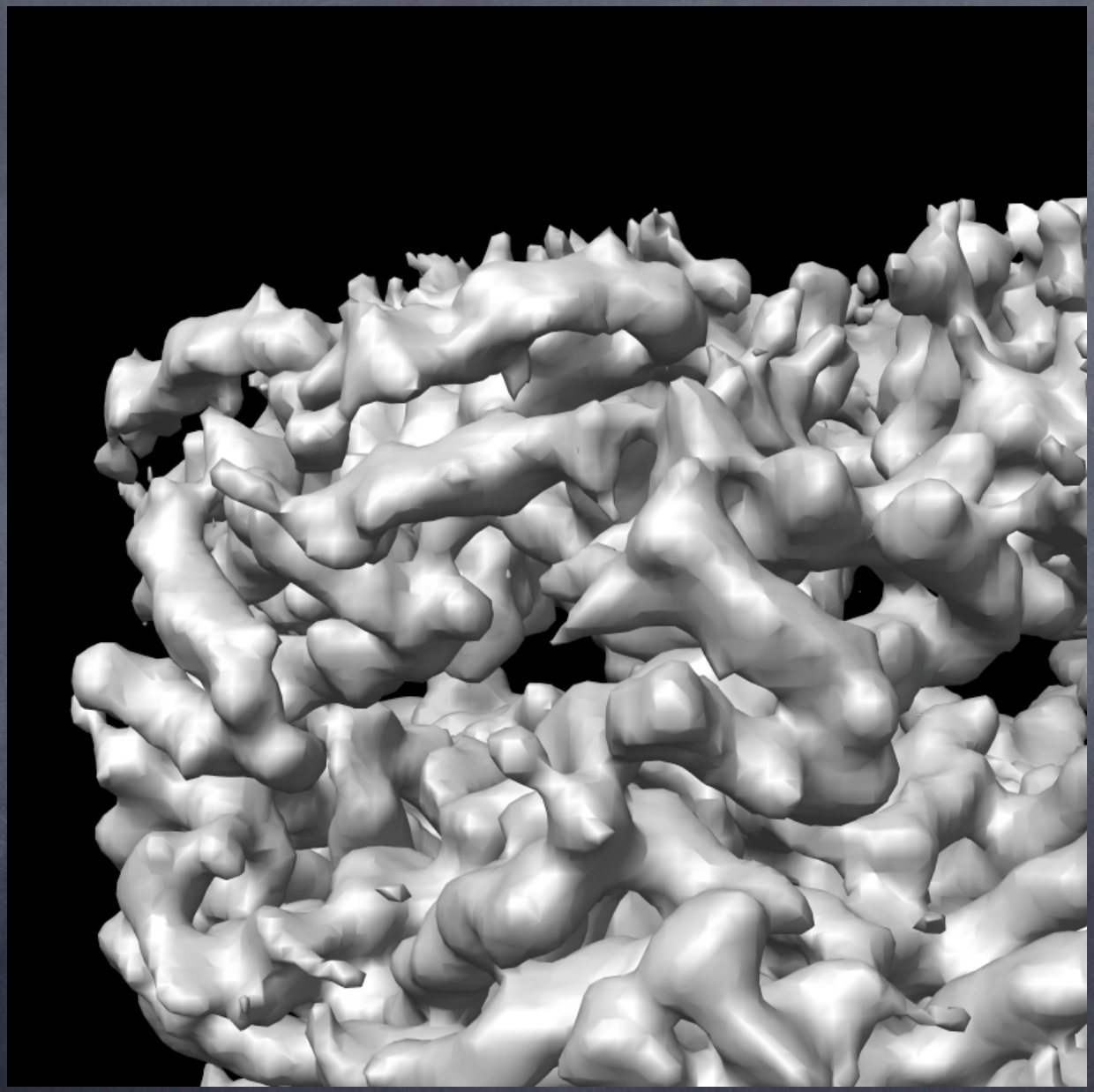


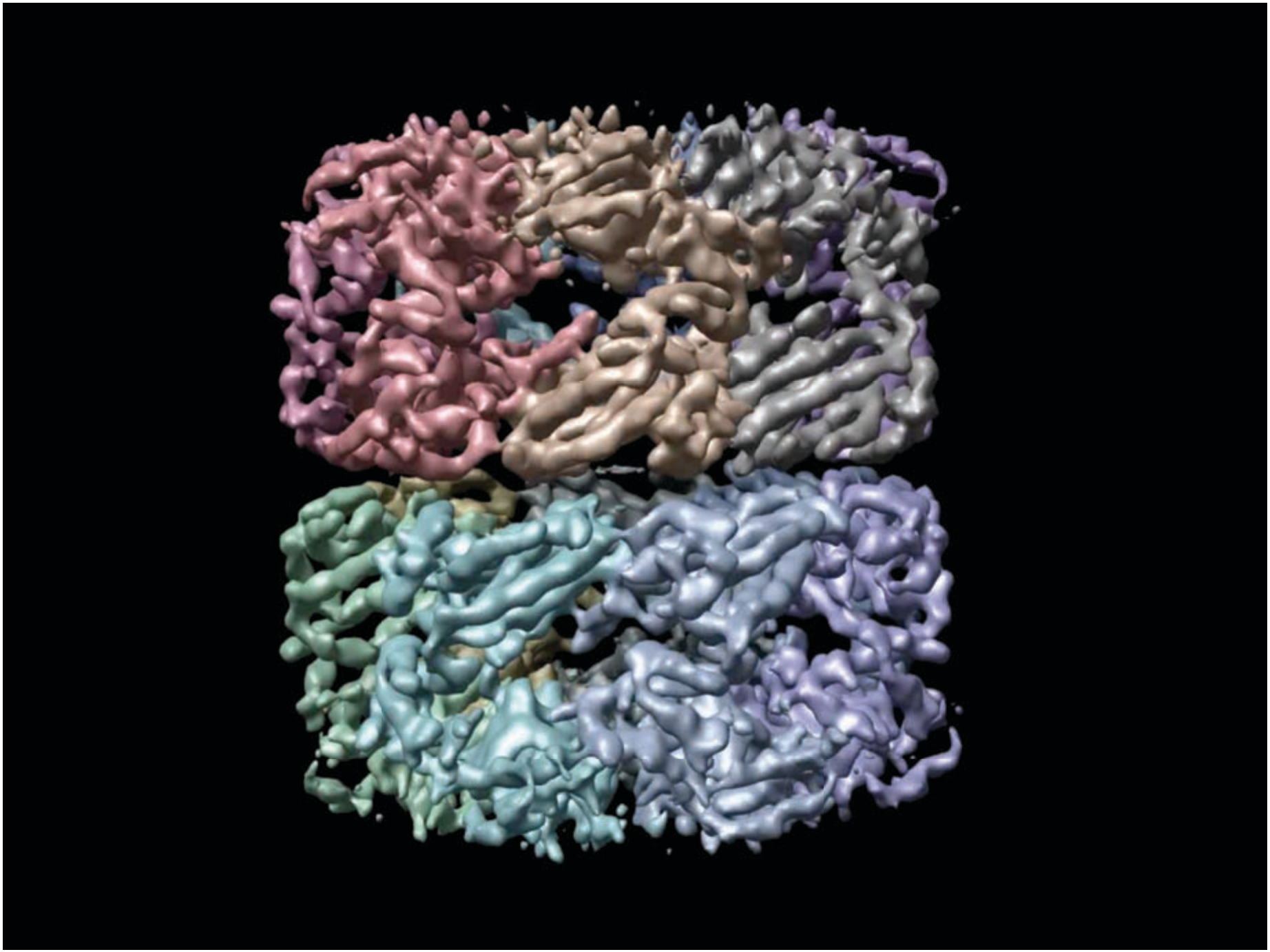

Masked SNR
Even/Odd test
FSC
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1/16
1/8
1/5
1/4
1/3
1/2
Resolution (1/A)

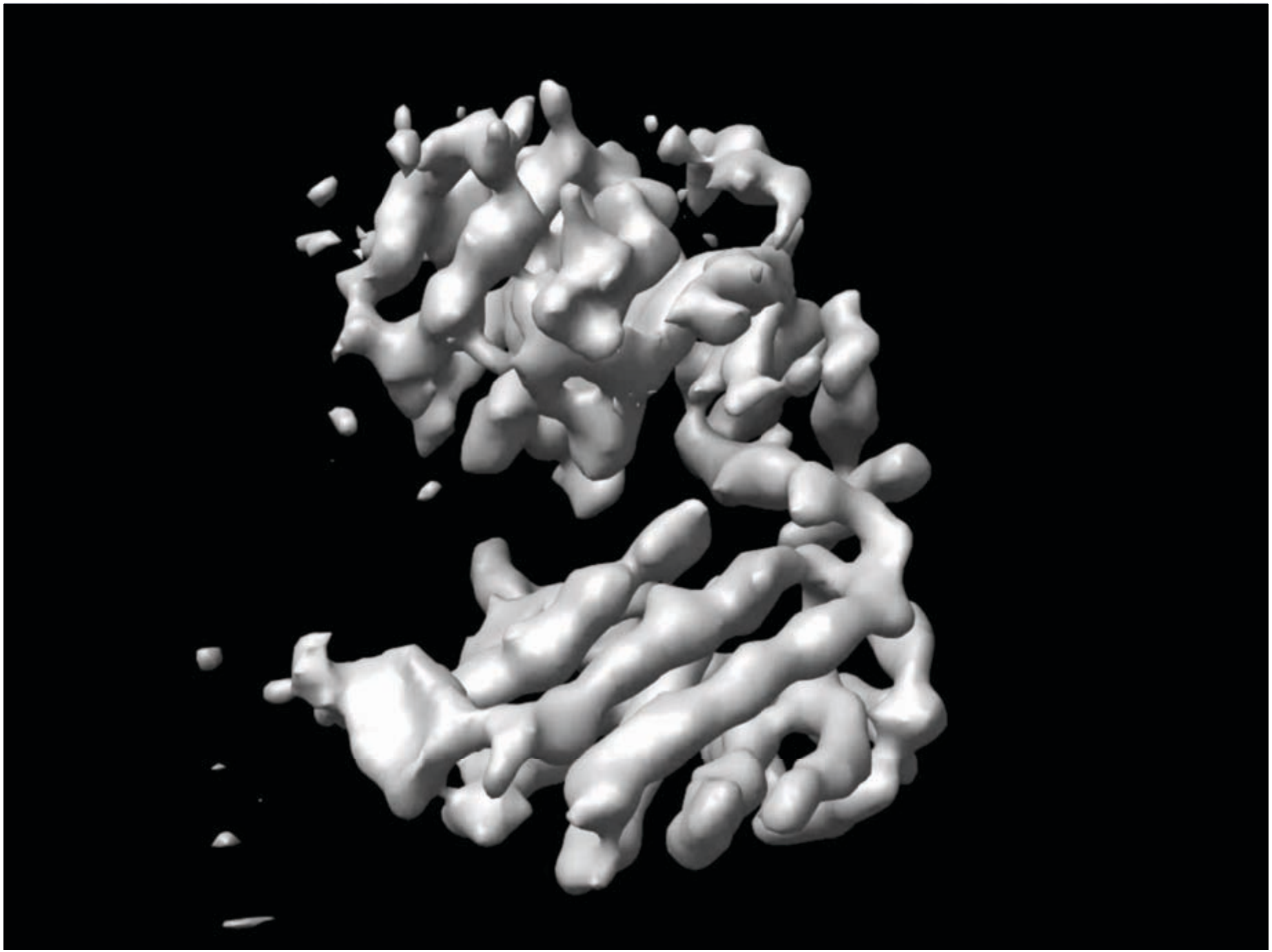


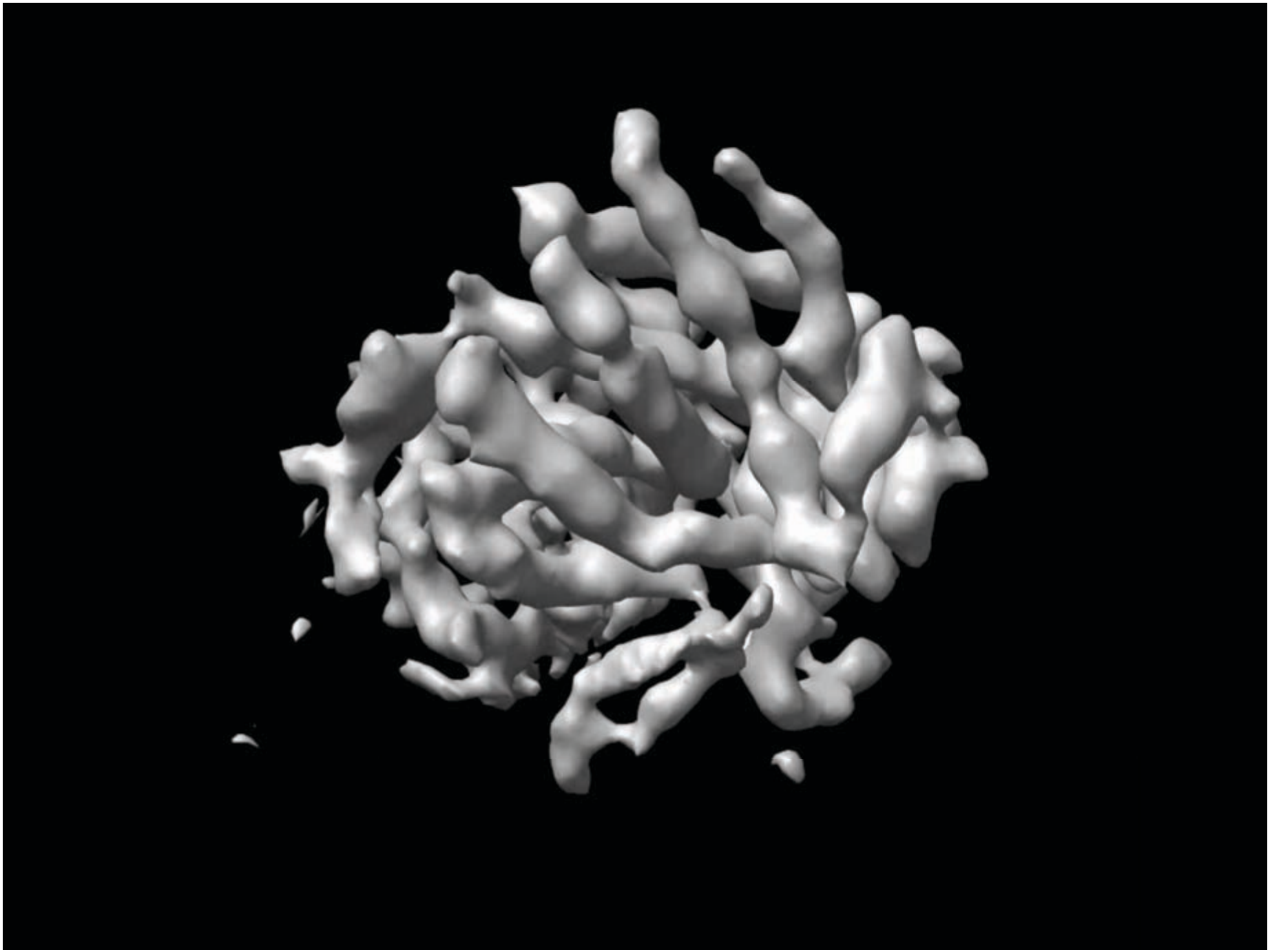




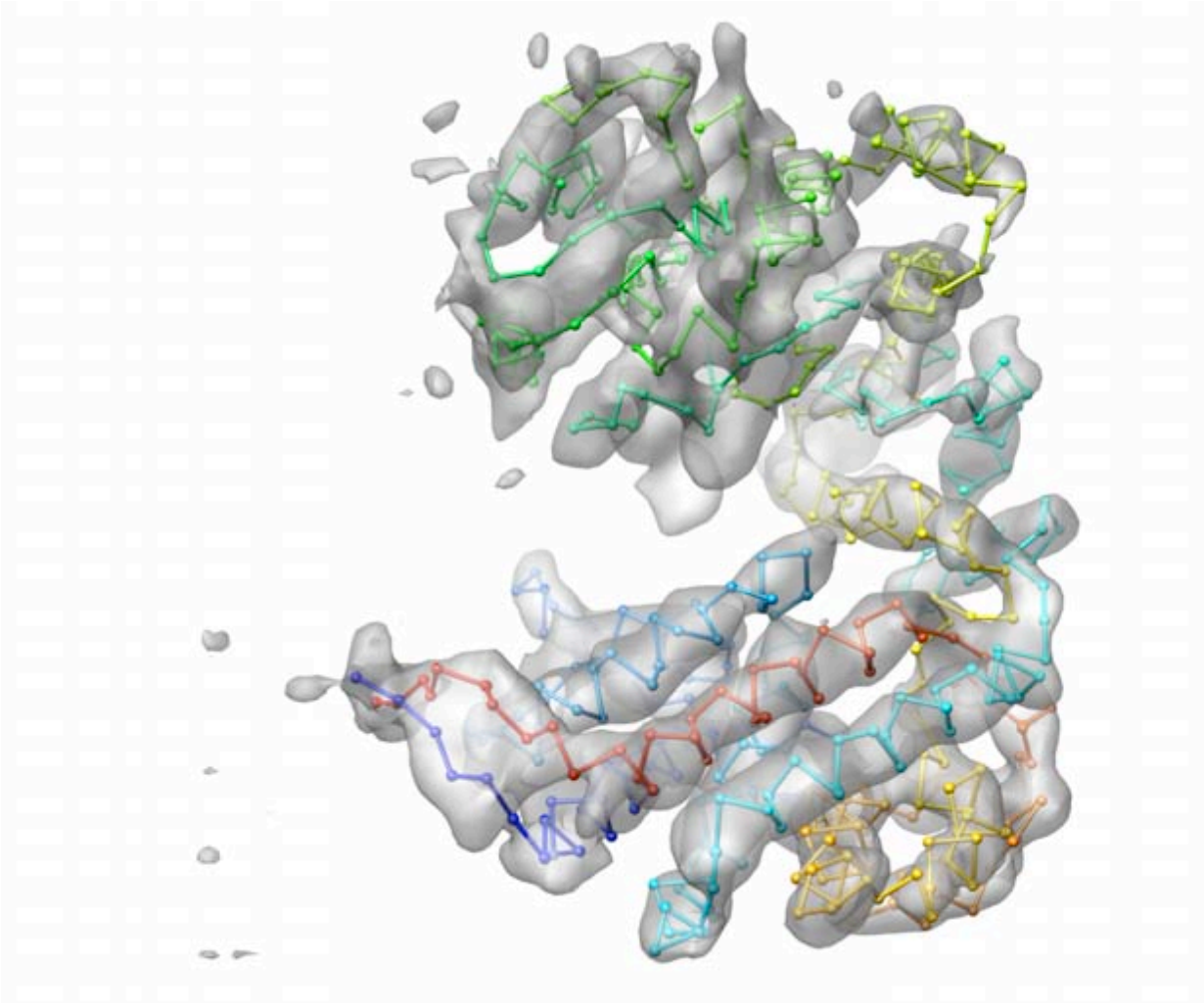




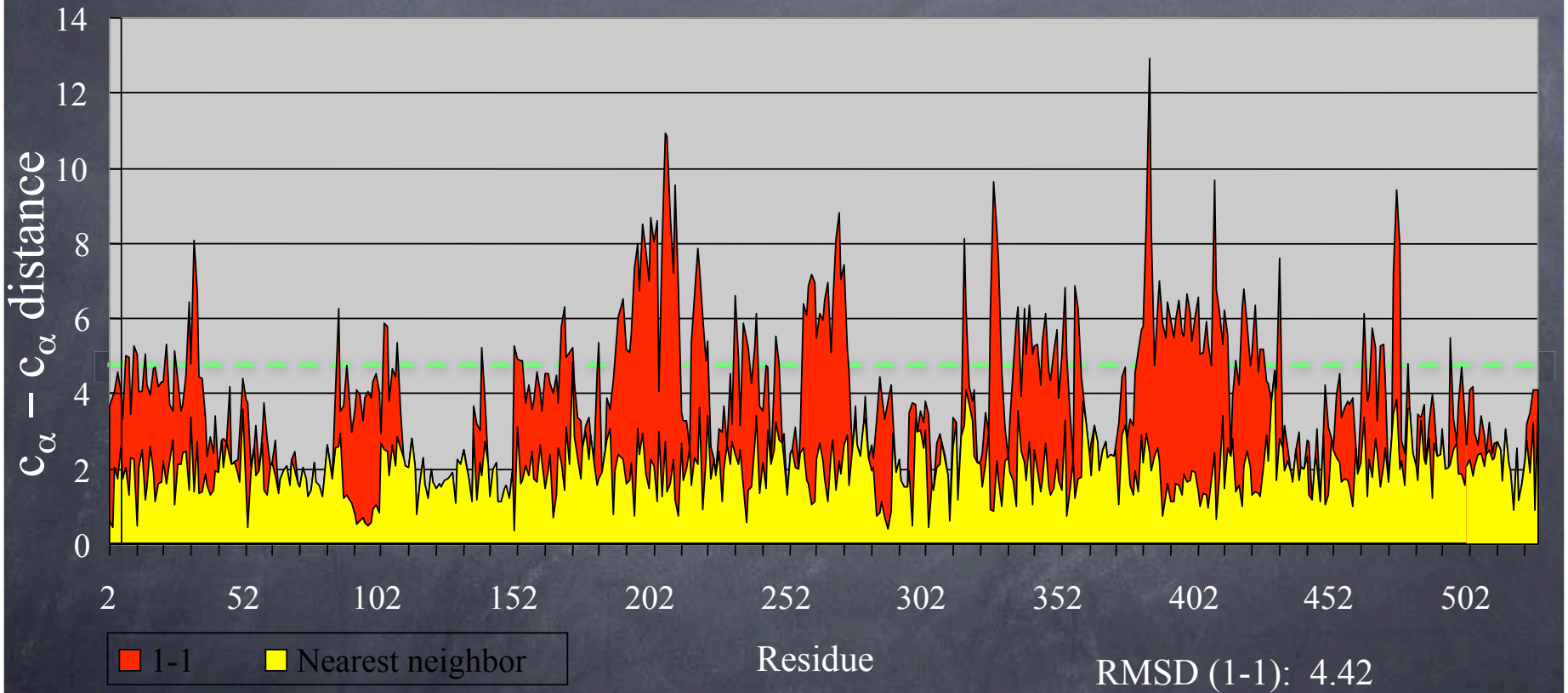




GroEL Chain Trace



RMSD vs. PDB Structure

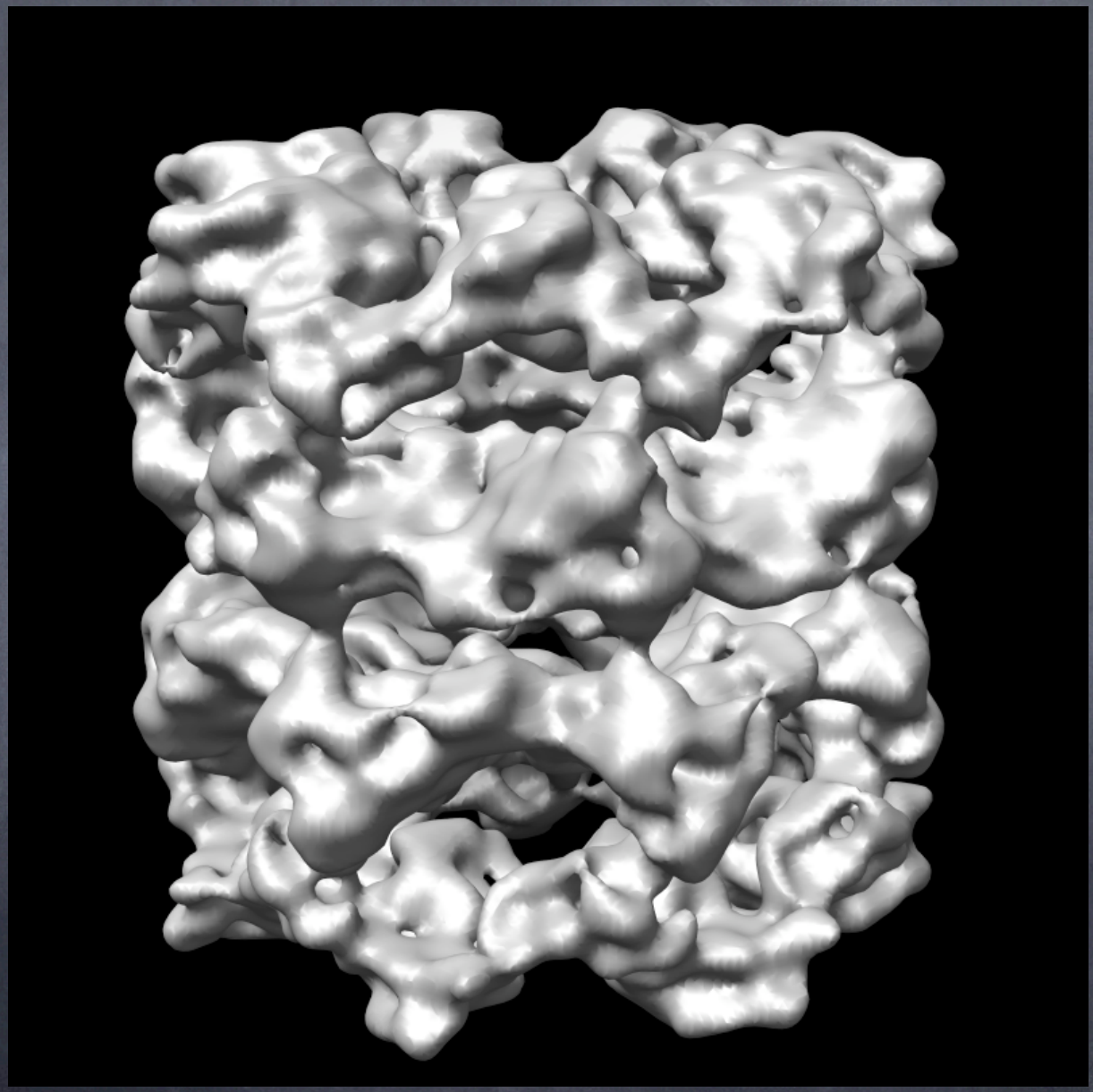


RMSD (1-1): 4.42
RMSD (NN): 2.19

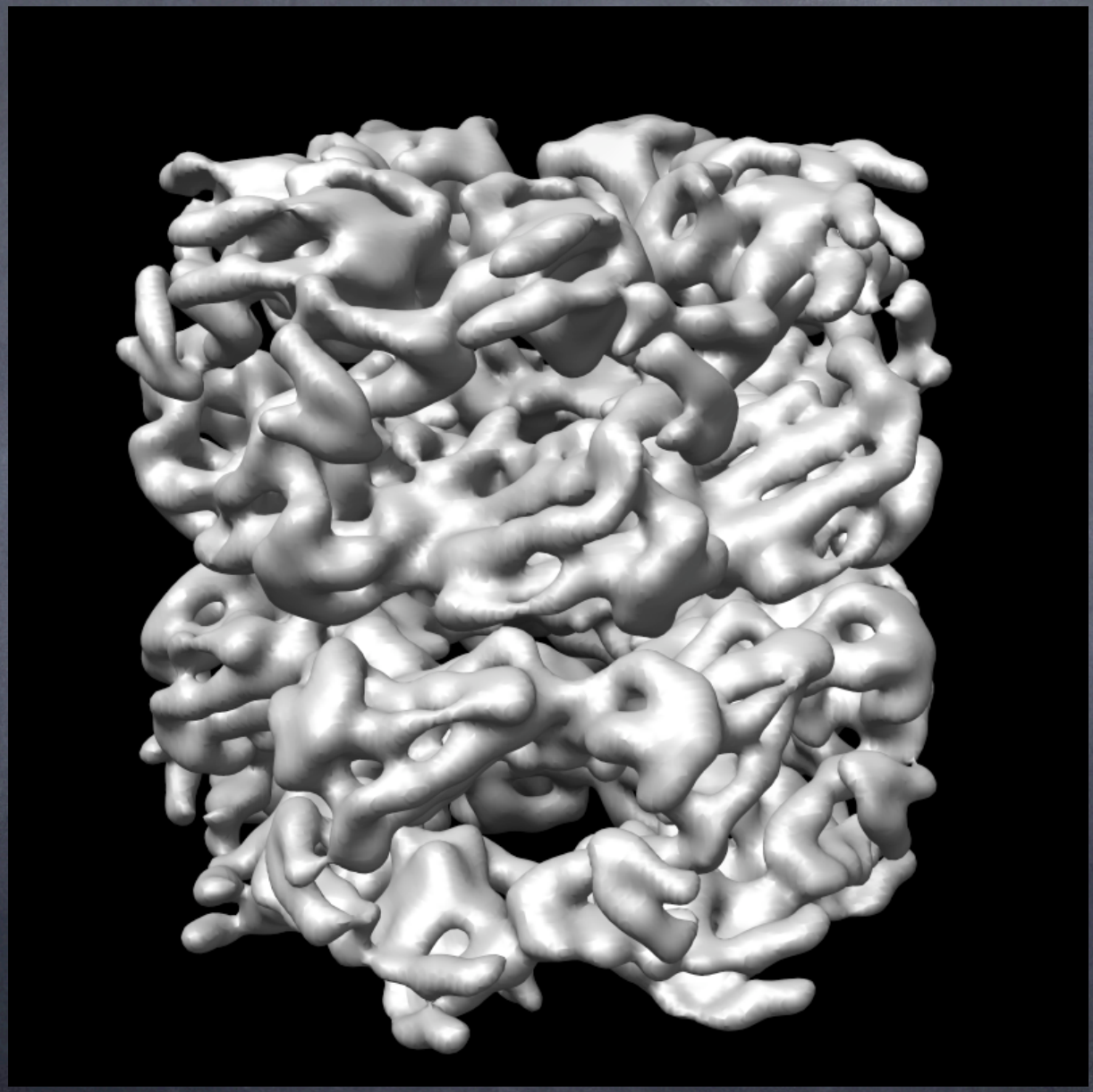
GroEL results

- ~4 Å resolution
- Single particle based Cα trace
- Asymmetric in solution
  - 1 Ring like apo crystal structure
  - Other ring similar to nucleotide bound state

11.5 Å GroEL

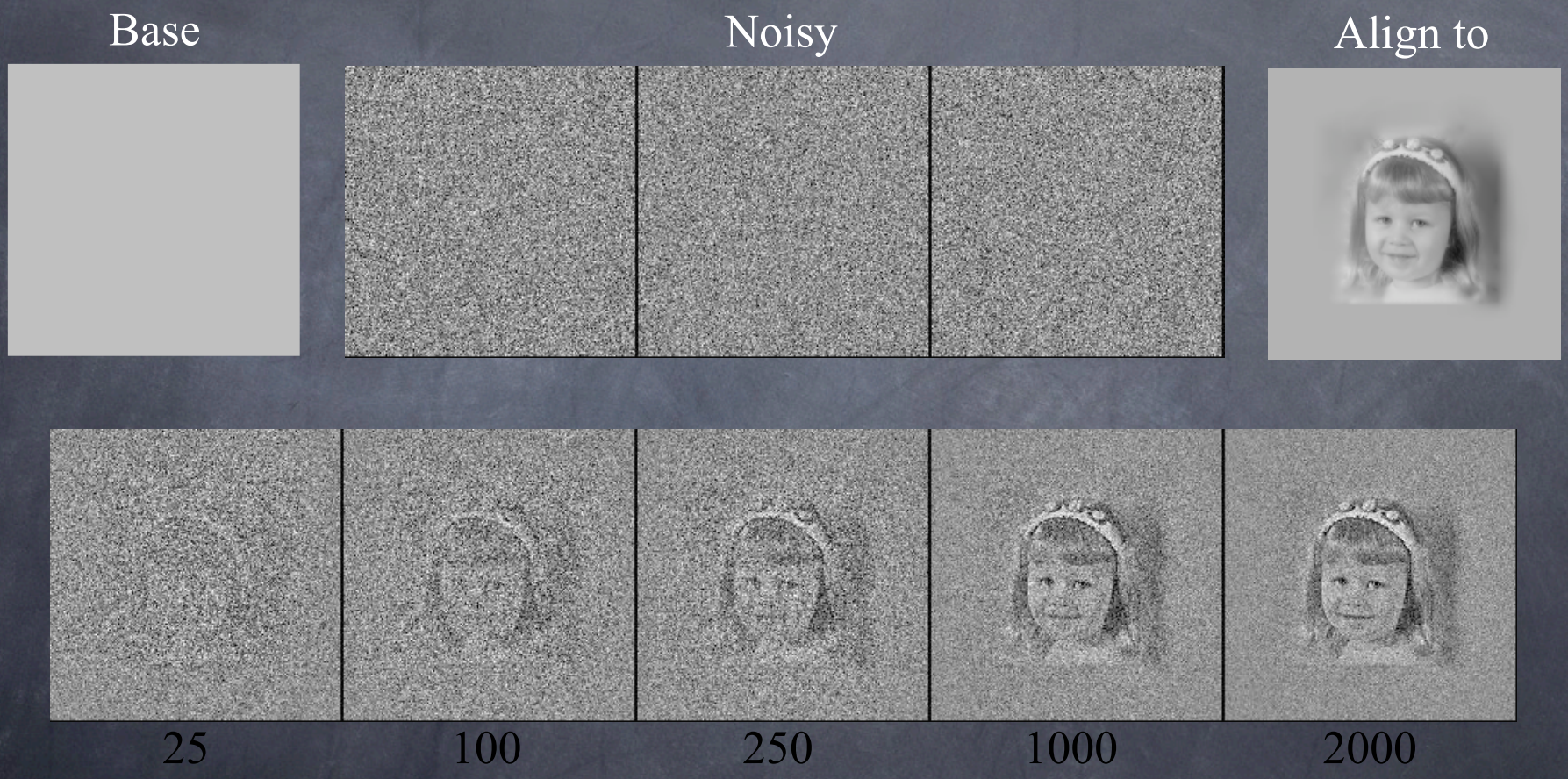


4 Å →11.5 Å GroEL



Model Bias ?

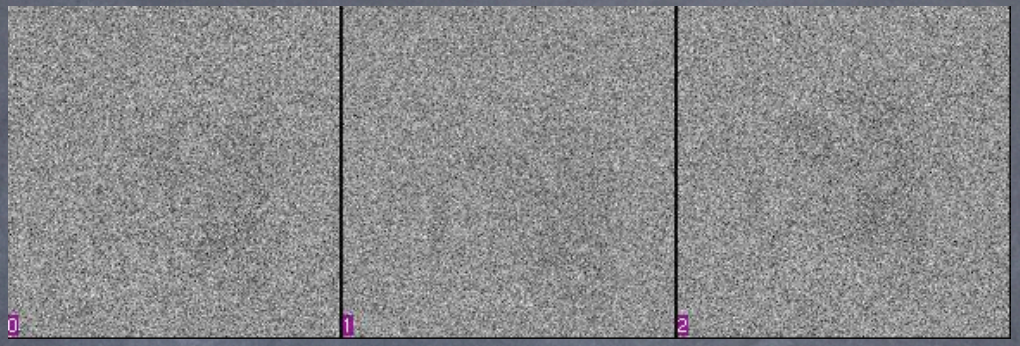
Model Bias



Model Bias

Base

Noisy (~10% contrast)



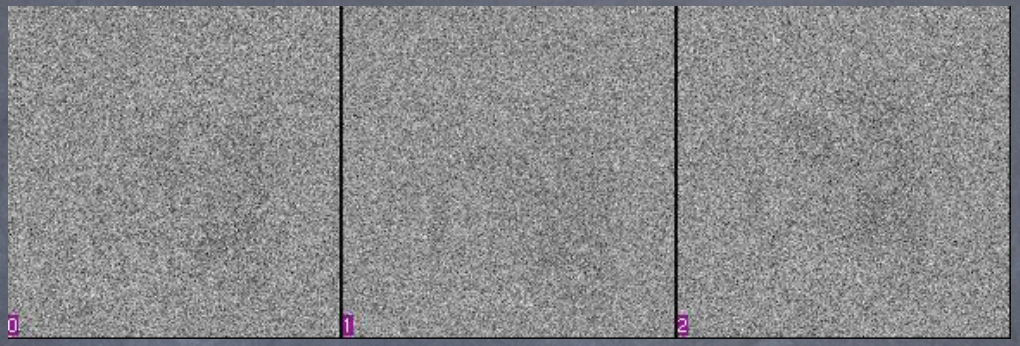
Align to

25 100 250 1000 2000

Model Bias

Base

Noisy (~10% contrast)



Align to

25 100 250 1000 2000

Model Bias

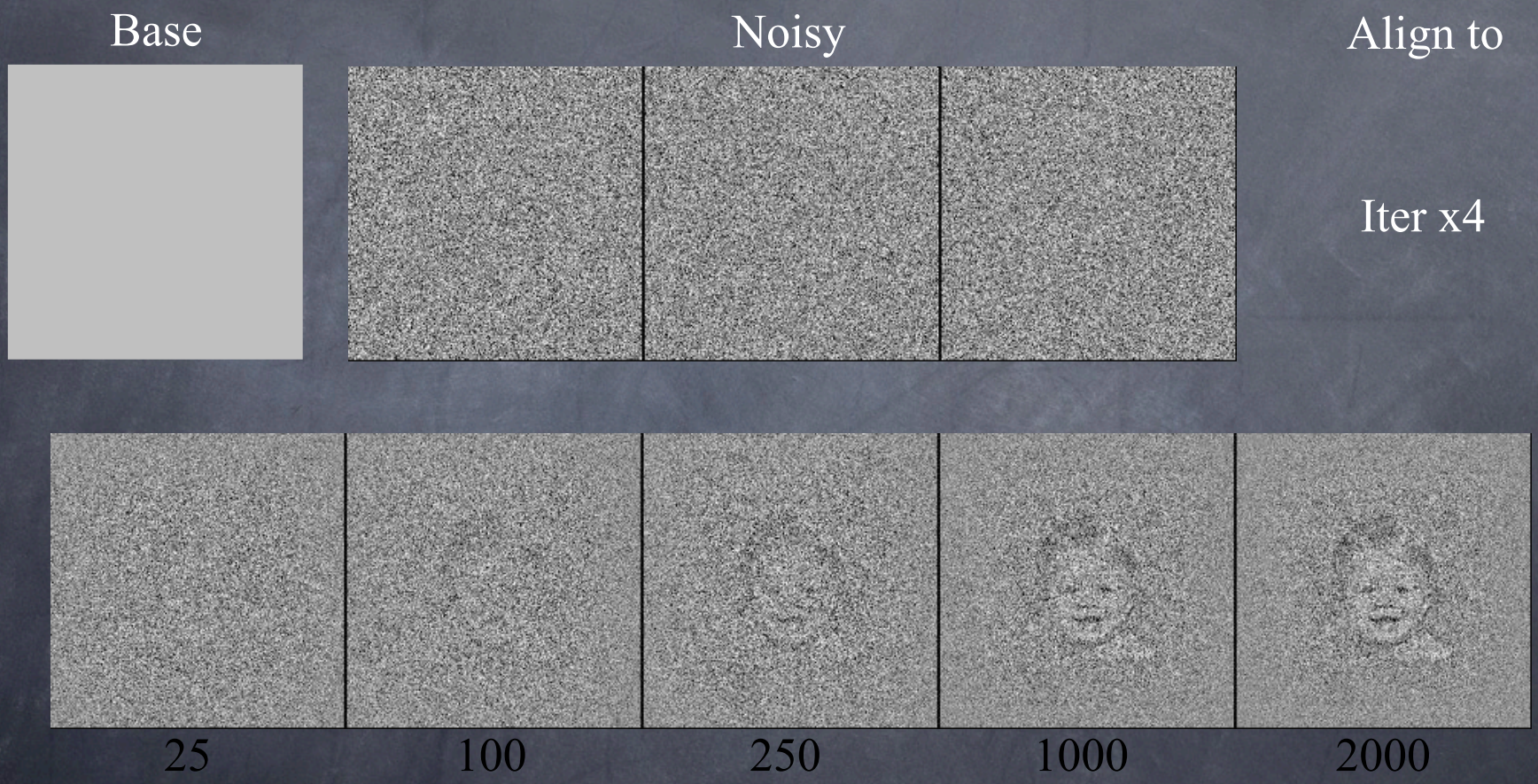
Base

Noisy

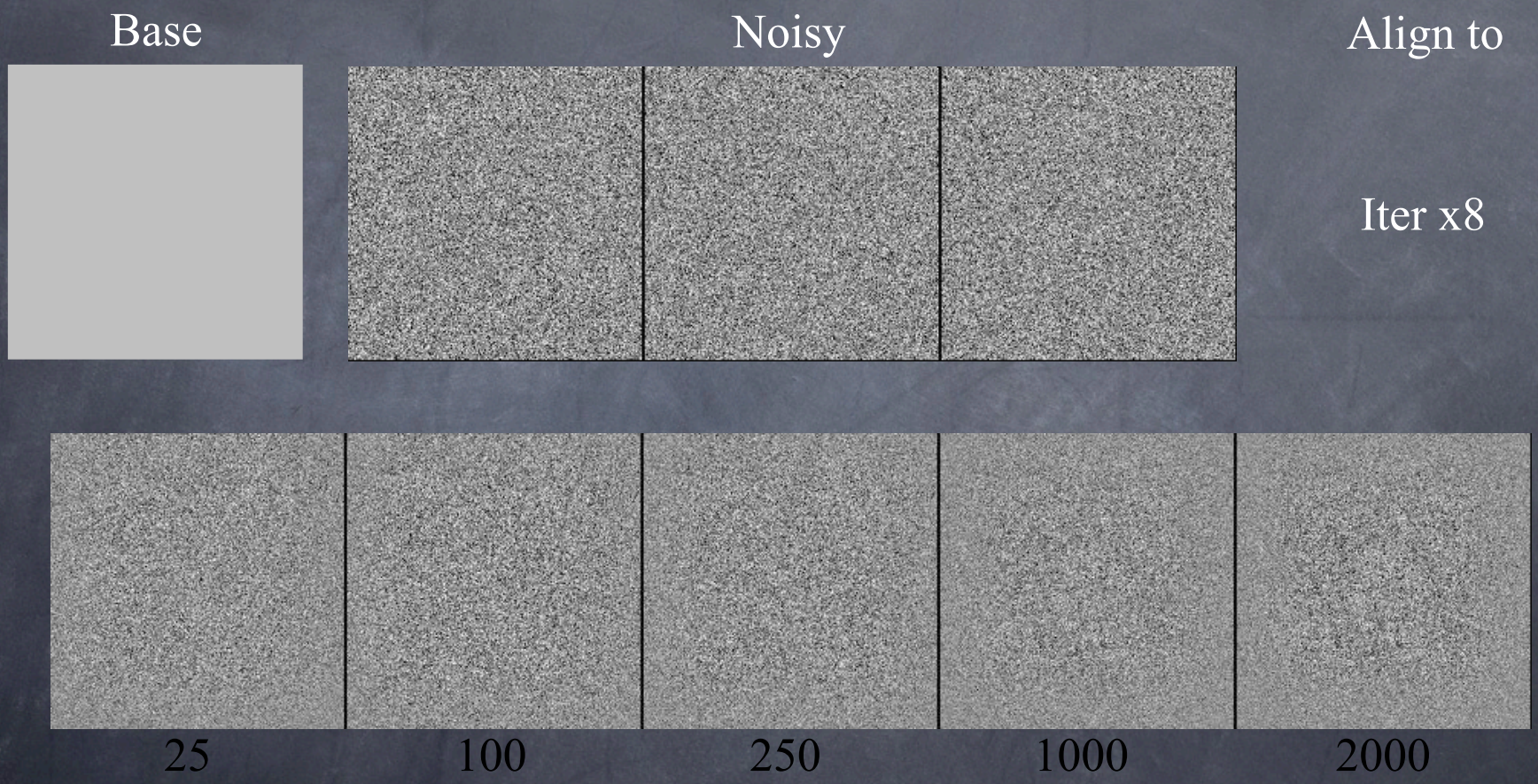
Align to

25 100 250 1000 2000

Model Bias



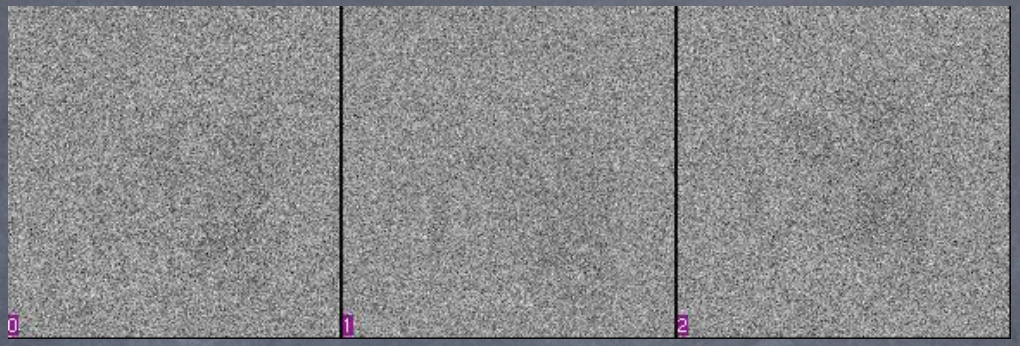
Model Bias



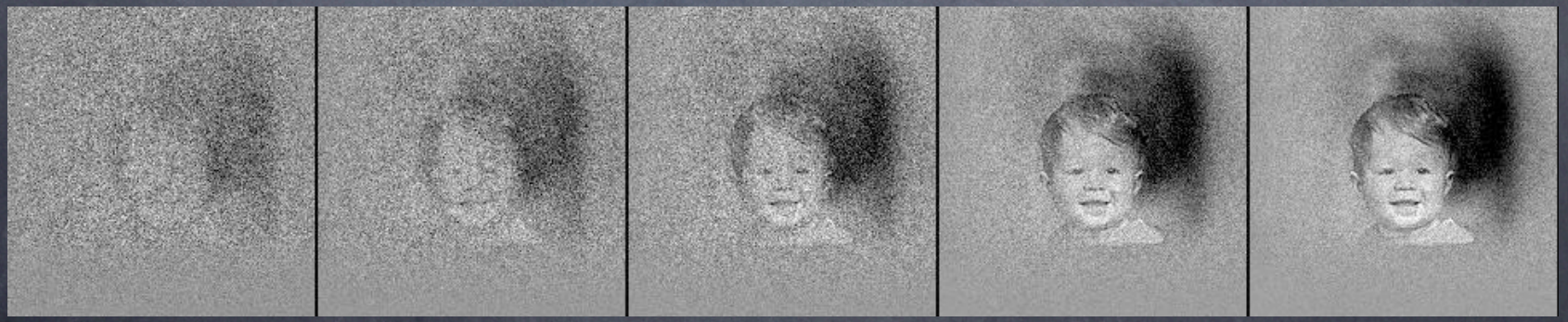
Model Bias

Base

Noisy (~10% contrast)



Align to

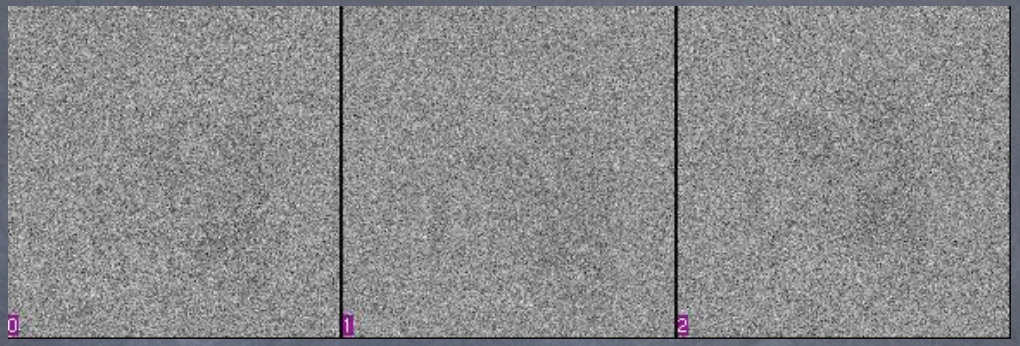


25 100 250 1000 2000

Model Bias

Base

Noisy

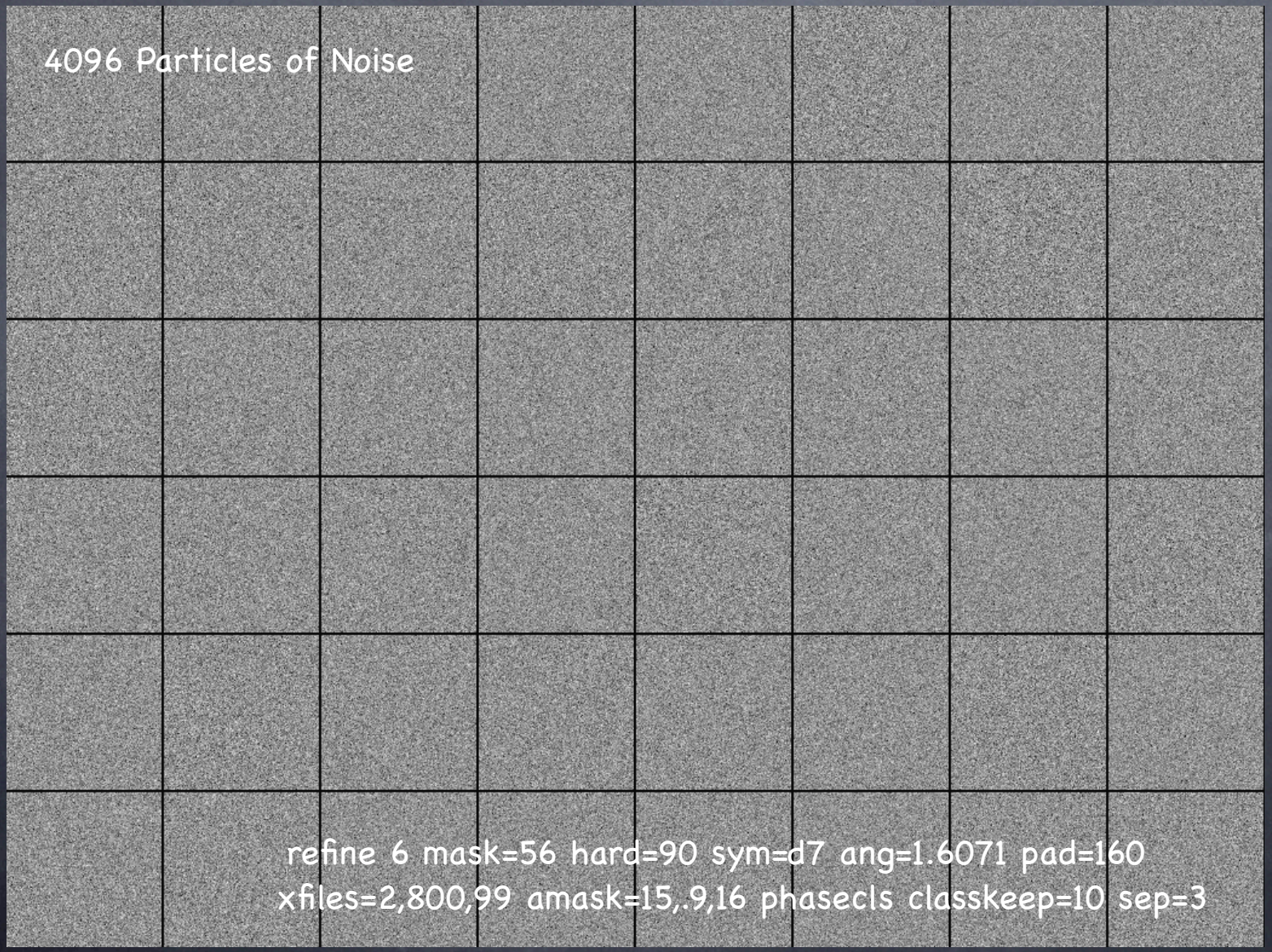


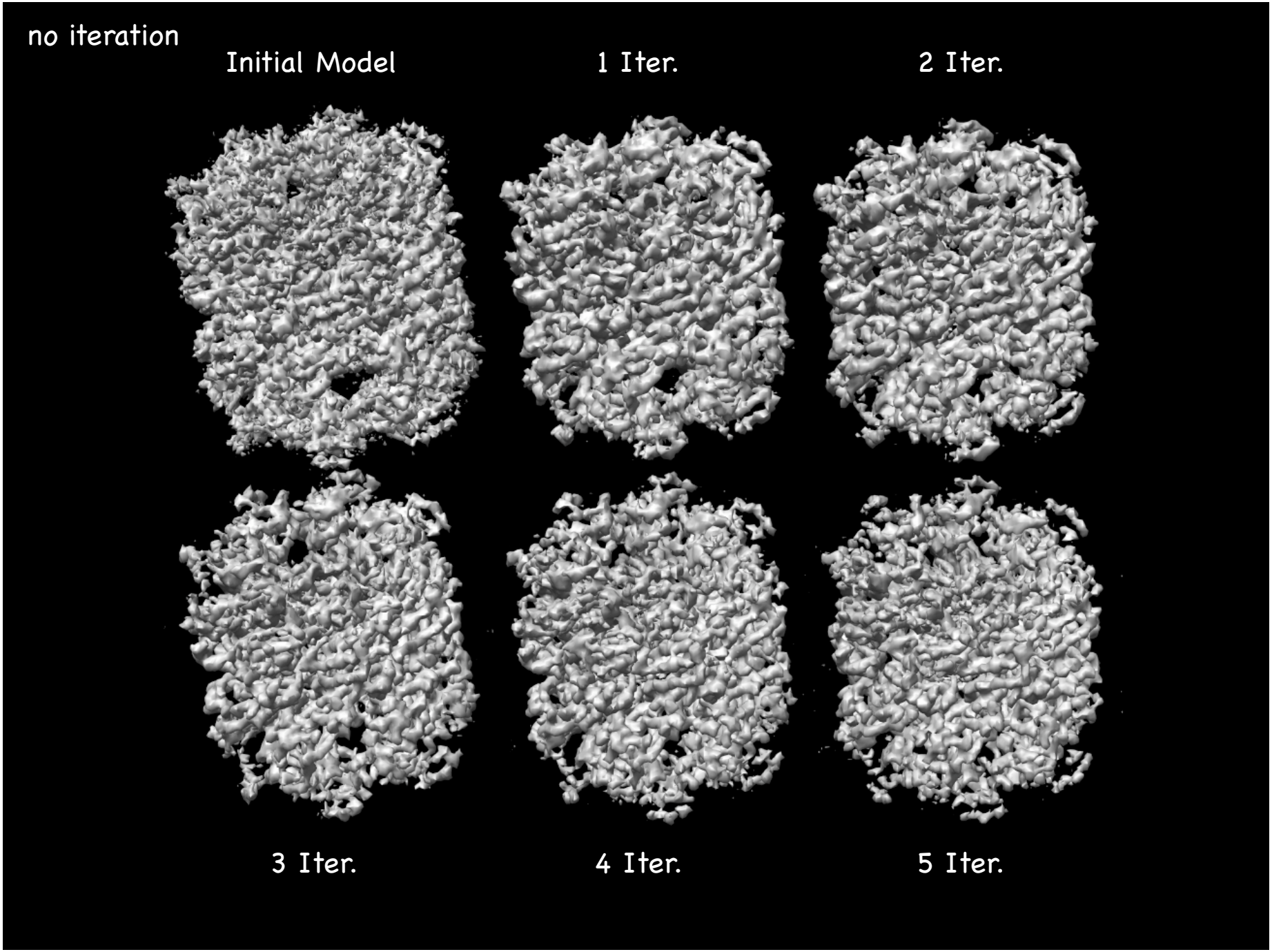
Align to

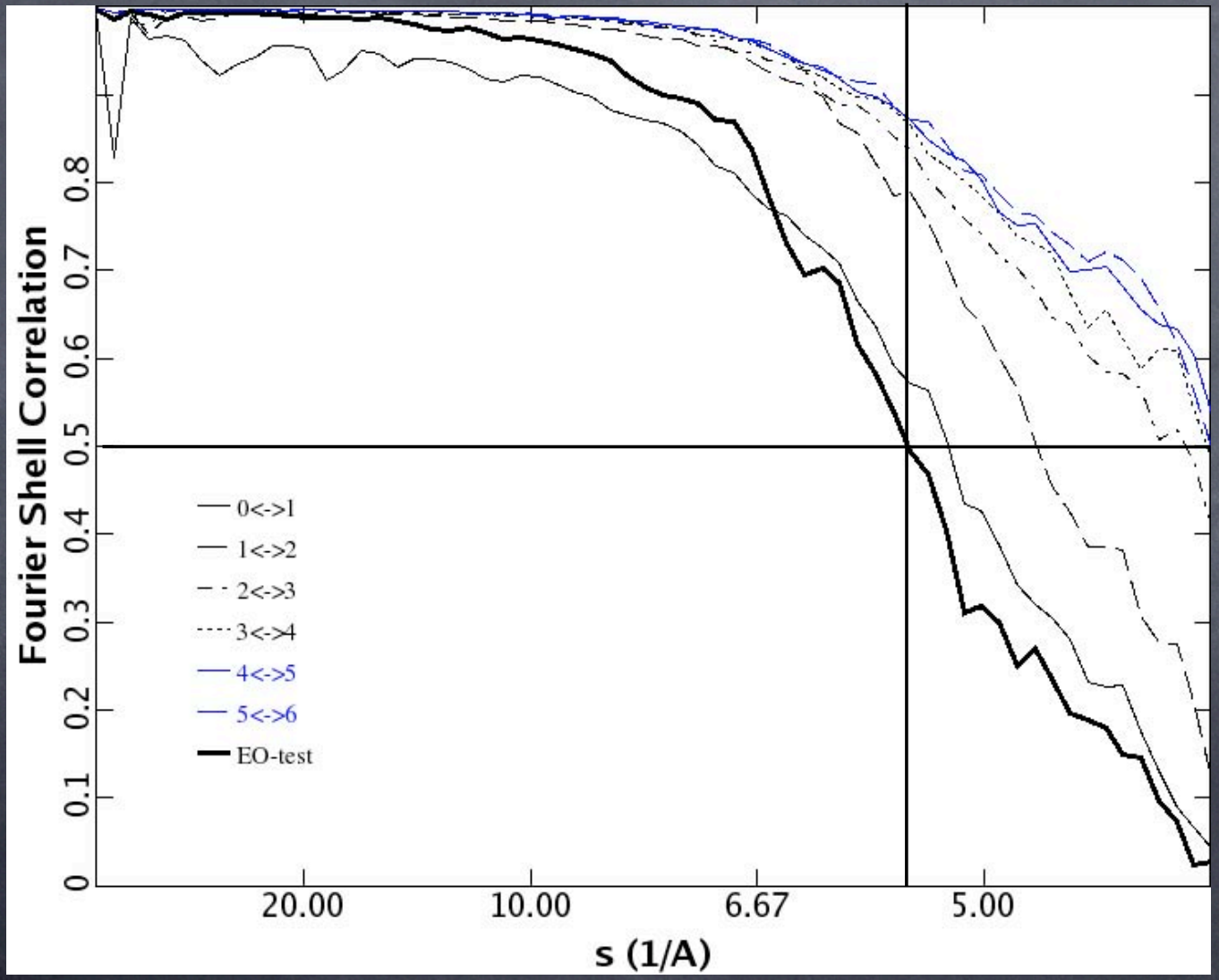
Iter x4

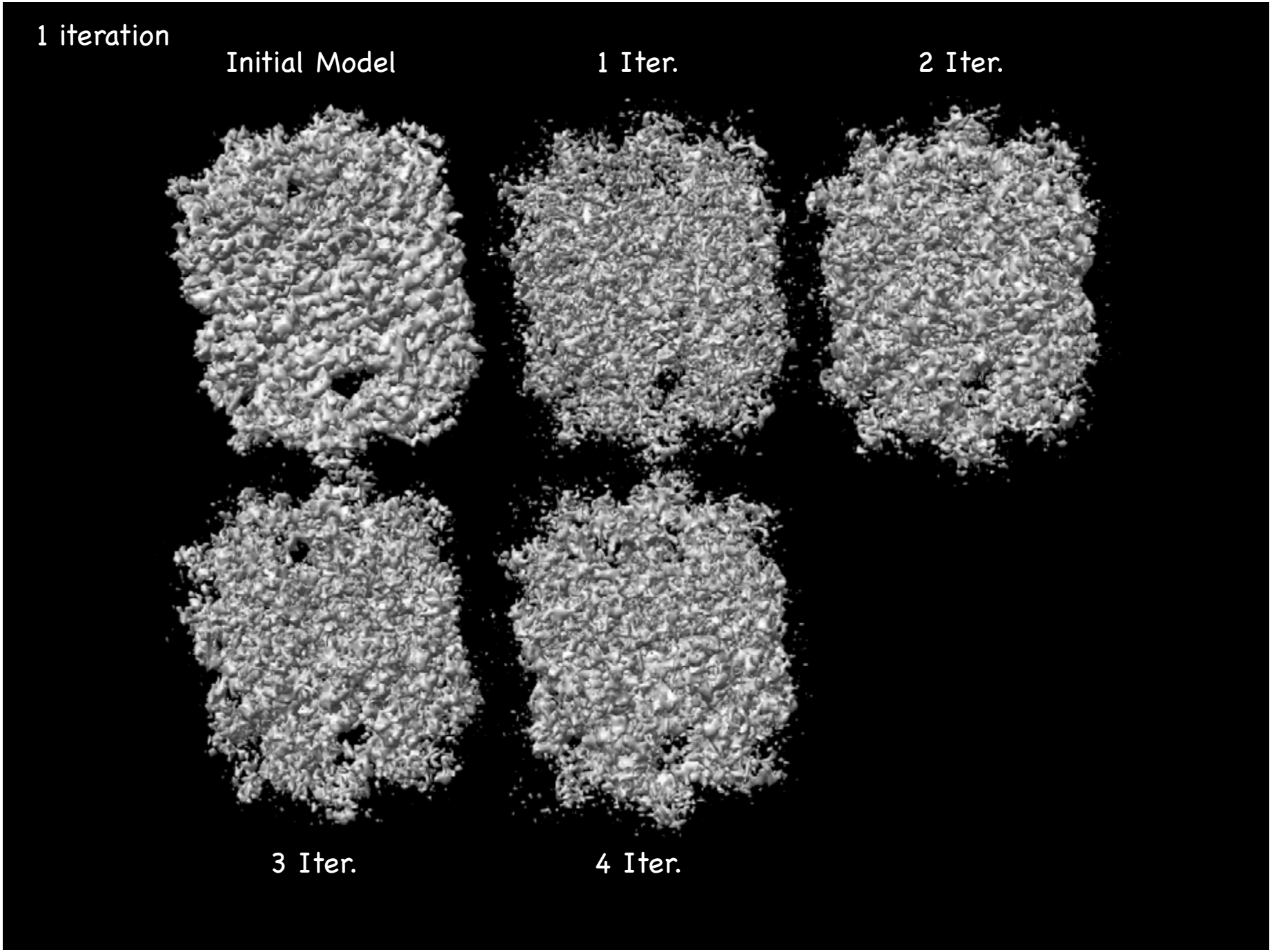
25 100 250 1000 2000

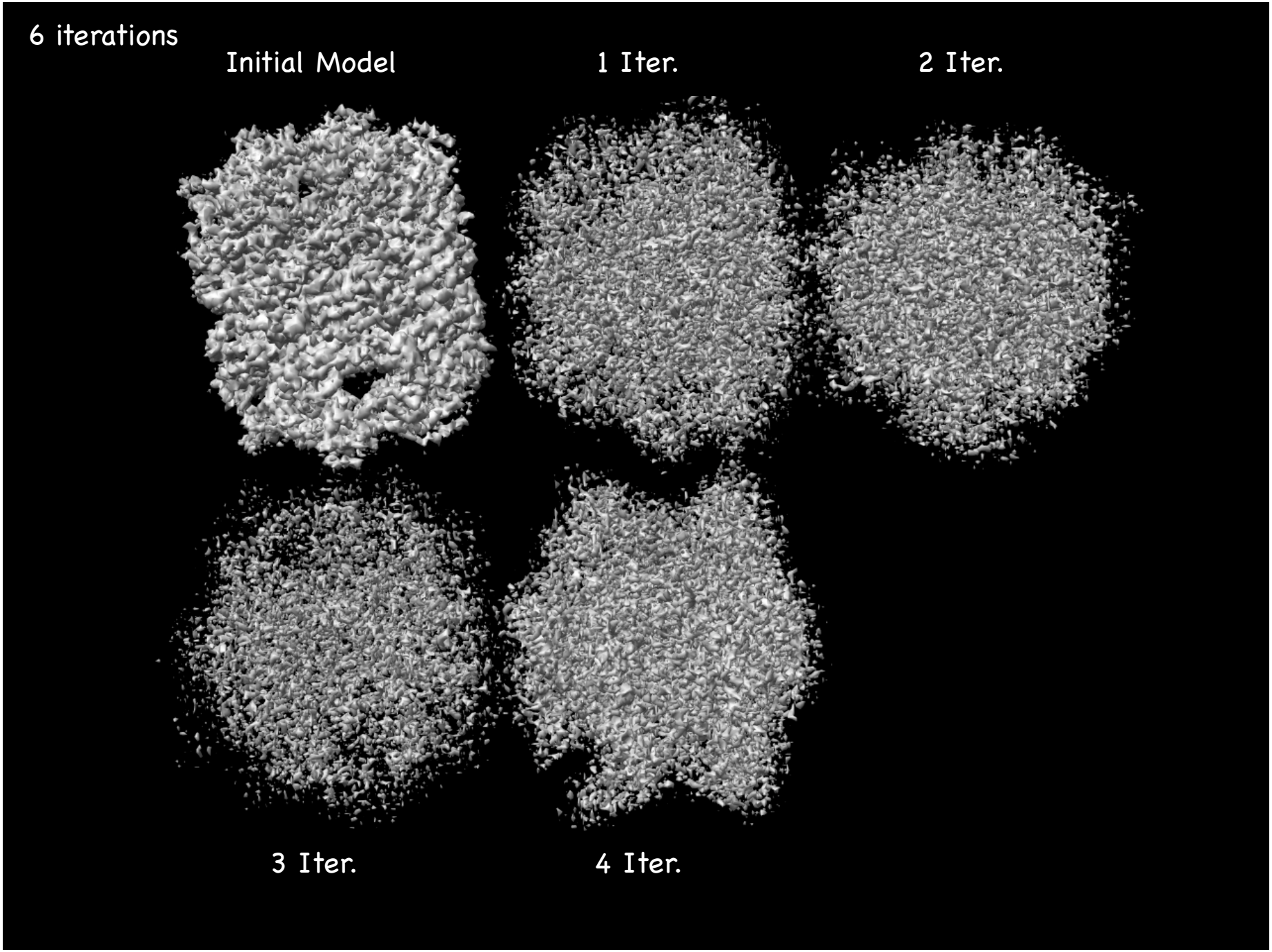
How About 3-D ?

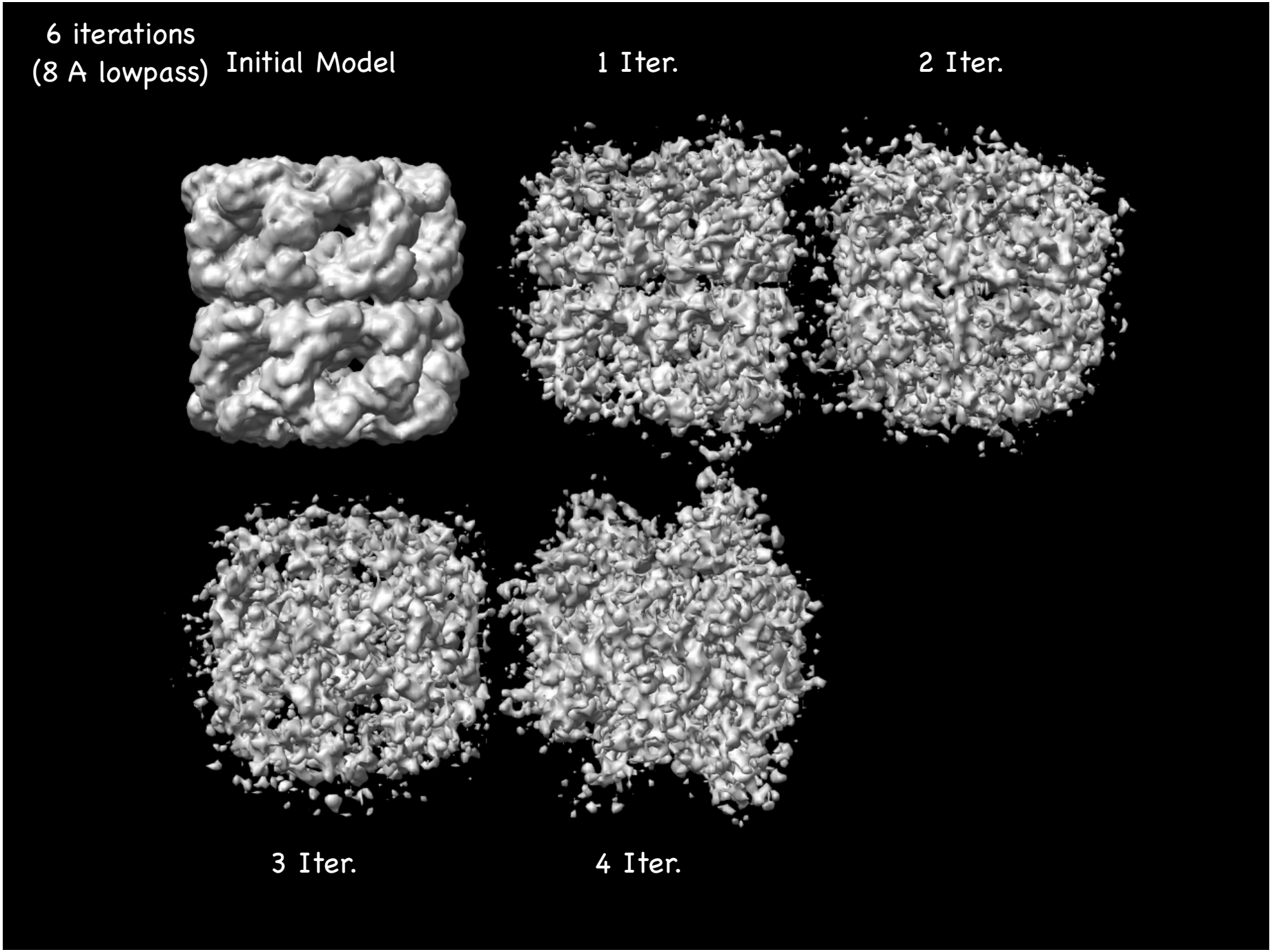

4096 Particles of Noise
refine 6 mask=56 hard=90 sym=d7 ang=1.6071 pad=160
xfiles=2,800,99 amask=15,.9,16 phasecls classkeep=10 sep=3

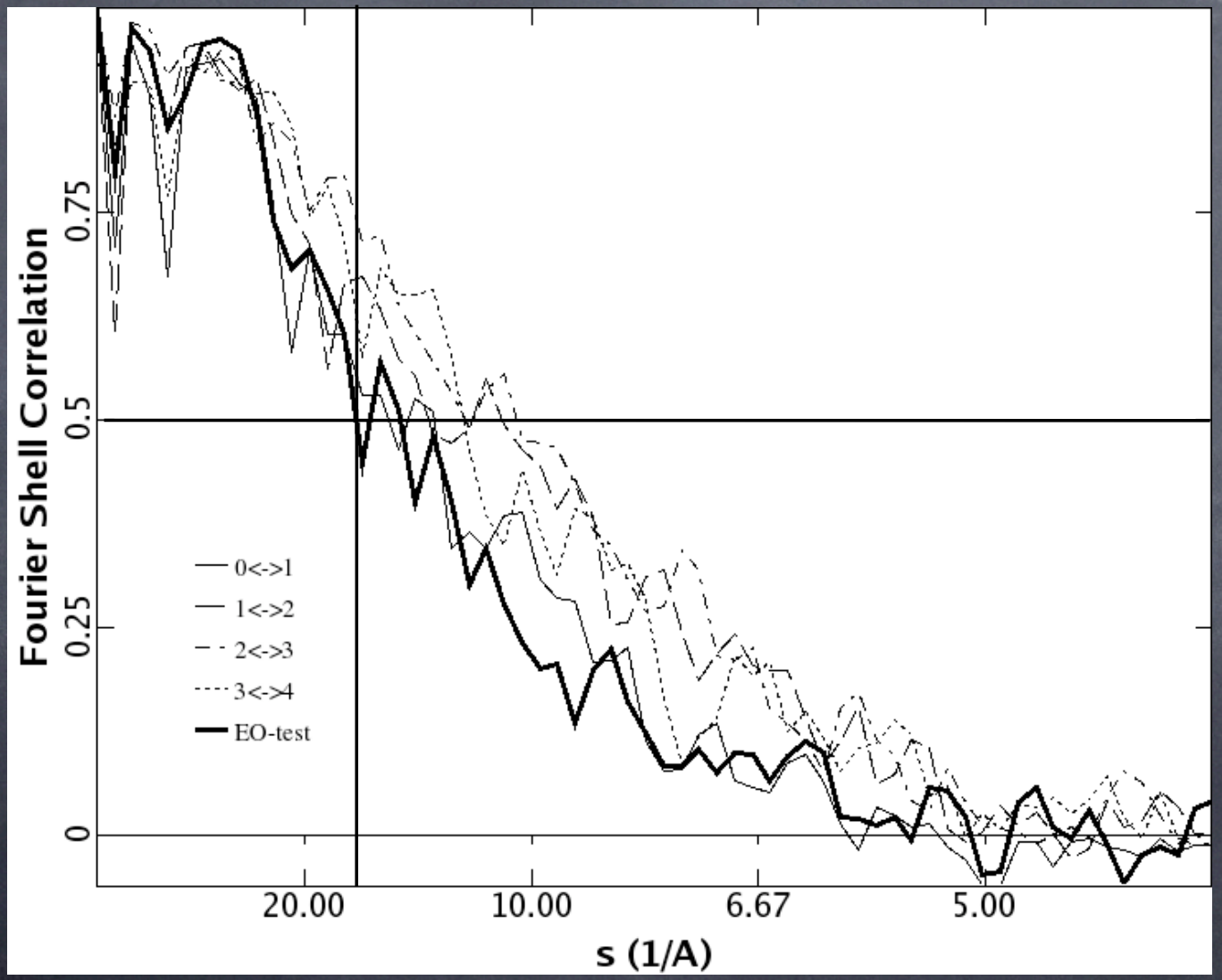

no iteration
Initial Model
1 Iter.
2 Iter.
3 Iter.
4 Iter.
5 Iter.


Fourier Shell Correlation
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
20.00
10.00
6.67
5.00
s (1/A)
0<->1
1<->2
2<->3
3<->4
4<->5
5<->6
EO-test


1 iteration
Initial Model
1 Iter.
2 Iter.
3 Iter.
4 Iter.


6 iterations
Initial Model
1 Iter.
2 Iter.
3 Iter.
4 Iter.


6 iterations
(8 A lowpass)
Initial Model
1 Iter.
2 Iter.
3 Iter.
4 Iter.


Fourier Shell Correlation
0.75
0.5
0.25
0
0<->1
1<->2
2<->3
3<->4
EO-test
20.00
10.00
6.67
5.00
s (1/A)

How Do we Stop This ?

- (In EMAN) use classiter>3 for a few rounds
- Use several different (random) starting models and insure that you get a good answer
- Compare 3D models with results of 2D analysis

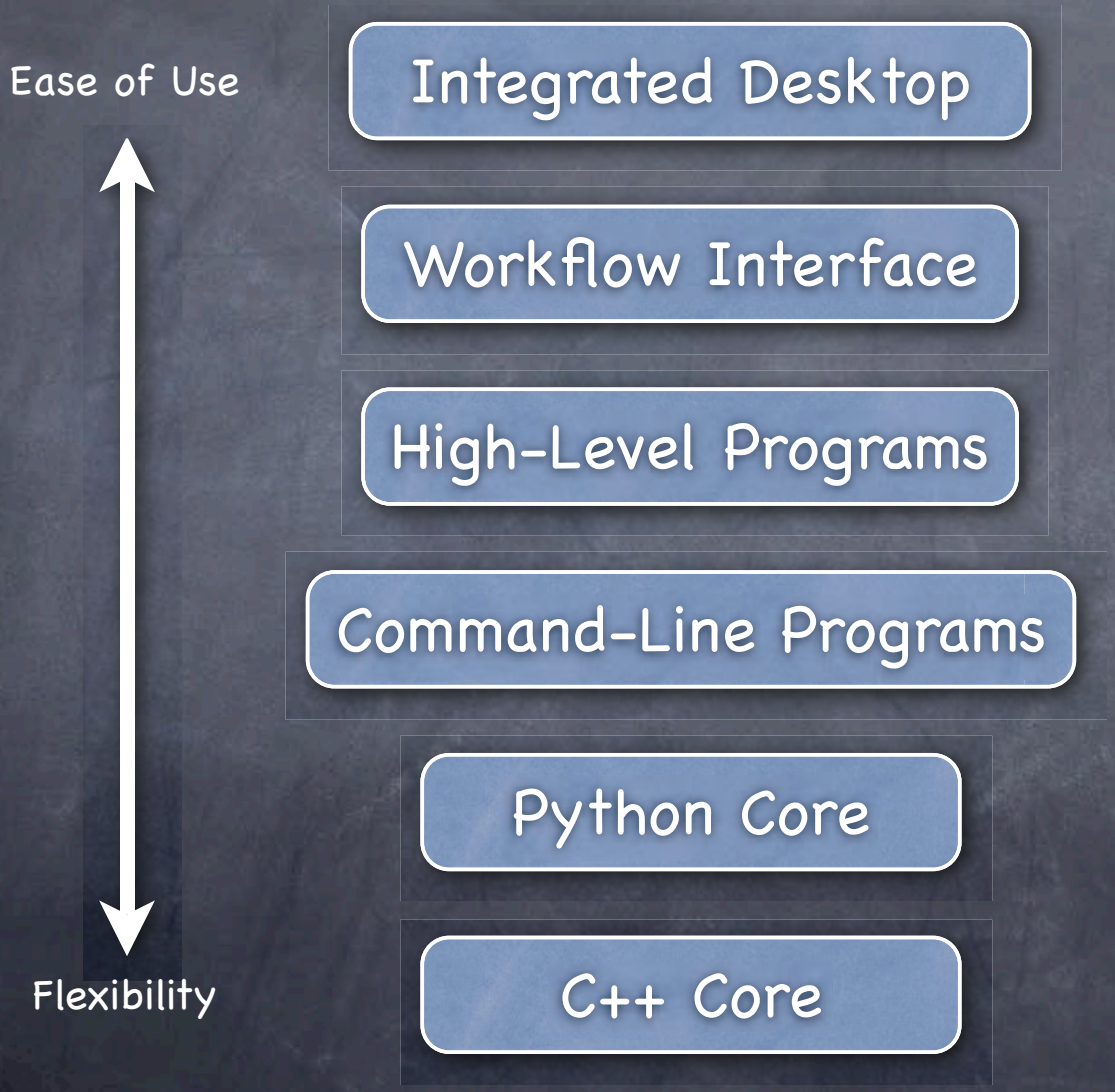
EMAN2

EMAN2 vs. EMAN1

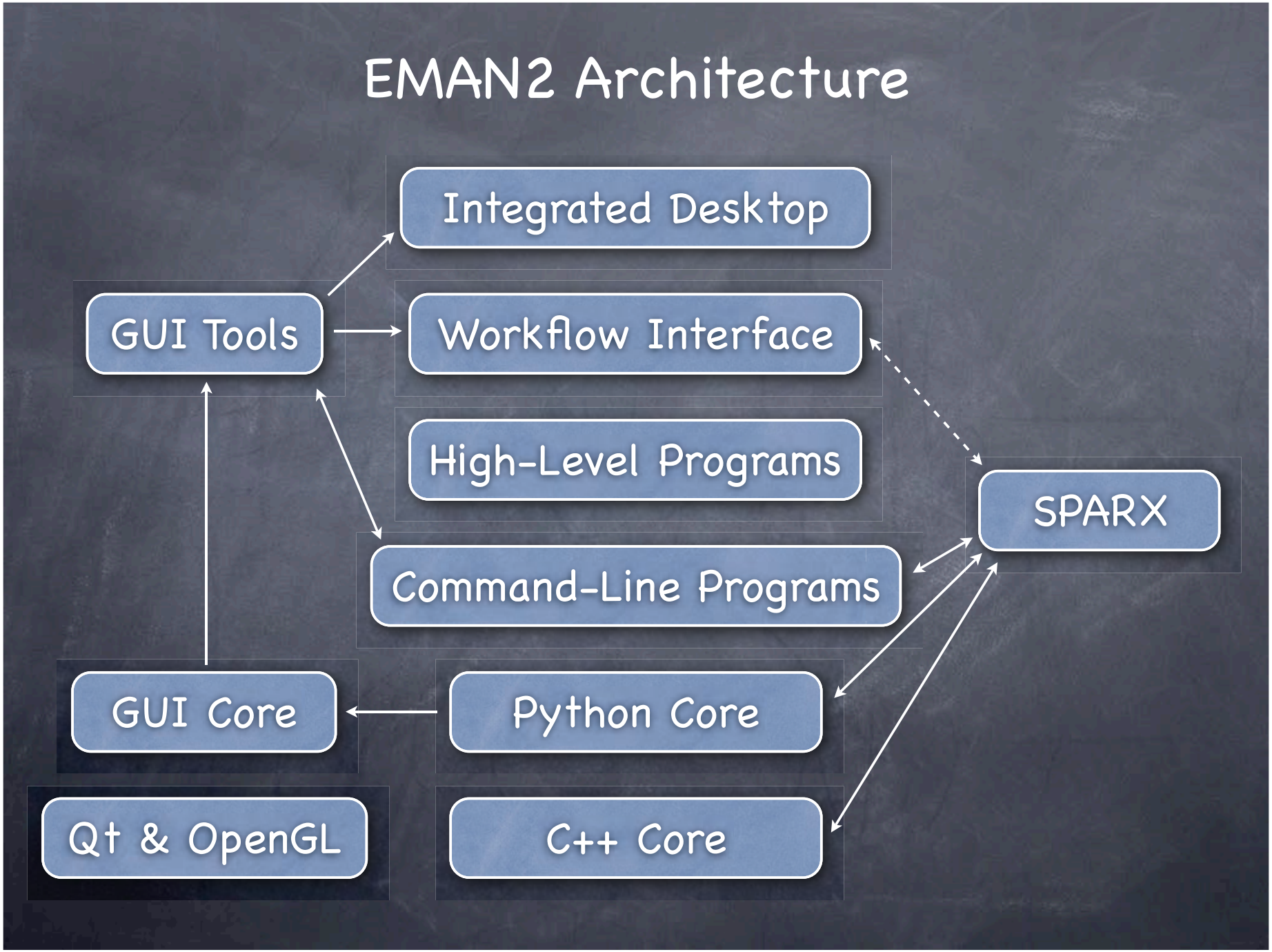
- Improved CTF model
  - Automatic fitting, Astigmatism*, Energy filtered data
- New openGL based GUI
- Workflow infrastructure
- Embedded database for data storage and metadata archival
- EMEN2 Integration *
- Easily extensible image processing infrastructure
- New parallelism strategy *
- CUDA support *

* - not yet ready for use

EMAN2 Architecture



EMAN2 Architecture



Extensible Core

Type	Description	#
Processor	Generic image processing algorithms, filters, masks, thresholds, etc.	157
Aligner	Algorithms used to align 2 images or volumes to each other	11
Projector	Routines to generate 2-D projections of 3-D objects	7
Reconstructor	Routines to reconstruct 3-D objects from 2-D projections	11
Cmp	Similarity metrics used to compare two images or volumes	9
Averager	Average together stacks of images in various ways	9
Analyzer	Perform various operations on sets of images, such as classification or PCA	6
Orientgen	Routines describing how projections cover the asymmetric triangle	6

File Formats

BDB +

MRC	R/W	IMAGIC	R/W
SPIDER	R/W	HDF5	R/W
PIF	R/W	ICOS	R/W
VTK	R/W	PGM	R/W
Amira	R/W	Xplor	W
Gatan DM2	R	Gatan DM3	R
TIFF	R/W	Scans-a-lot	R
LST	R/W	PNG	R/W
Video-4-Linux	R	JPEG	W

Processors
(categories & examples)

- filter
  - filter.lowpass.gauss
  - filter.homomorphic.tophat
- mask
  - mask.sharp
  - mask.gaussian
- math
  - math.sqrt
  - math.laplacian
- misc
  - misc.localnorm
- normalize
  - normalize
  - normalize.edgemean
- testimage
  - testimage.scurve
- threshold
  - threshold.binary
  - threshold.clampminmax
- xform
  - xform.centerofmass
  - xform.fourierorigin.tocenter

Similarity Metrics
(cmp)

With Default options, SMALLER -> more similar

- dot - dot product (negative by default)
- frc - Fourier ring correlation (weighted)
- optvariance - 'optimized variance' (EMAN1)
- phase - mean phase error
- quadmindot - Worst of quadrant dot products
- sqeuclidean - sum $(a-b)^2/n$

Programs

- 49 Command-Line Programs (EMAN2)

  syntax:

  e2<name>.py --help

  e2<name>.py <file> [--option=value] [--option] [-O]

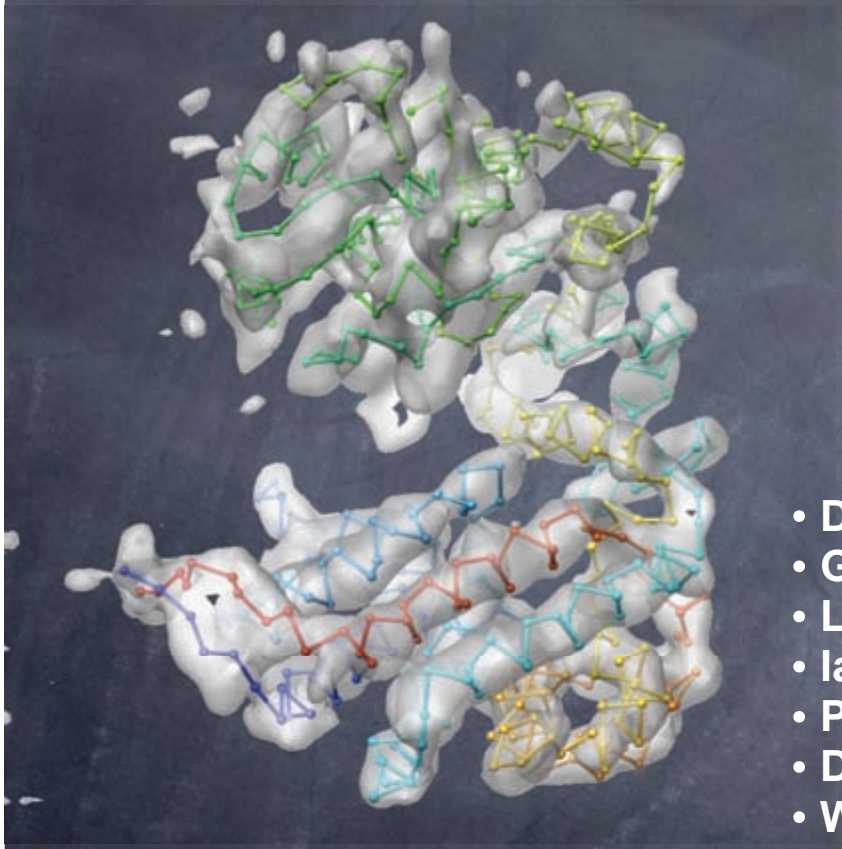
  <> - required parameter

  [] - optional parameter

GUI

- e2desktop.py (may not be stable yet)
- e2workflow.py
- e2display.py
- and other programs with the --gui option

Acknowledgements



GroEL

- Matt Baker
- Donghua Chen
- Jiu-Li Song (UTSW)
- David Chuang (UTSW)
- Wah Chiu

EMAN2

BCM

- David Woolford
- Guang (Grant) Tang
- Liwei Peng
- Ian Rees
- Phil Baldwin
- Deepy Mann
- Wen Jiang (Purdue)

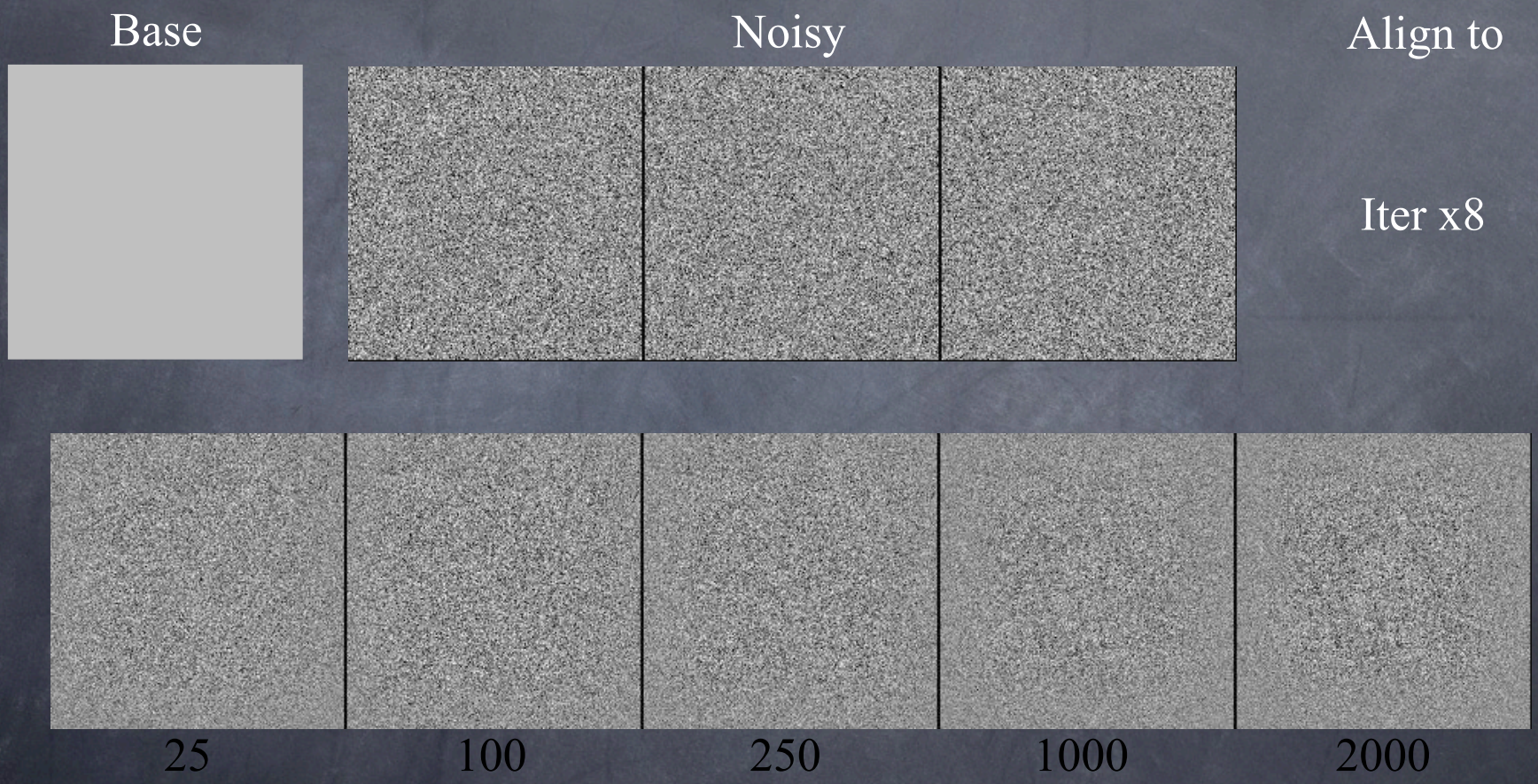
Via SPARX

- Pawel Penczek (UTH)
- Wei Zhang (UTH)
- Zhengfan Yang (UTH)
- Julien Bert (UTH)
- Stefan Raunser (MPI)
- Christian Spahn (Charité)
- Justus Loerke (Charité)
- Chao Yang (LBNL)

Supported by NIH Roadmap Initiative, NCRR, NIGMS, and the Welch Foundation.

Graphics produced using UCSF Chimera and EMANimator.

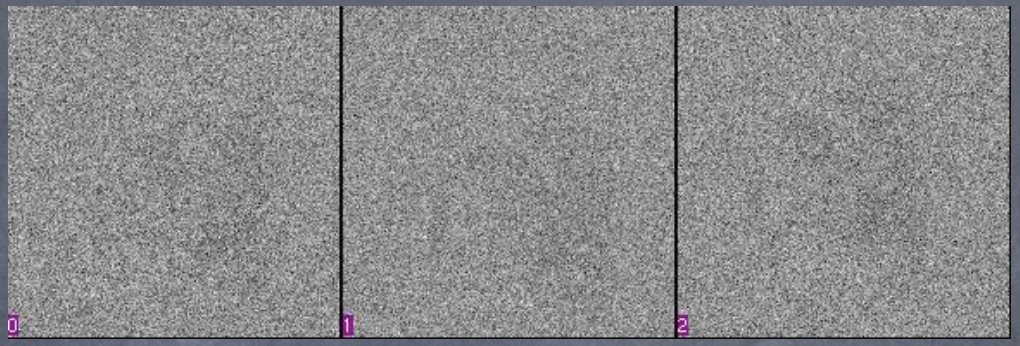
Model Bias



Model Bias

Base

Noisy



Align to

Iter x4

25 100 250 1000 2000