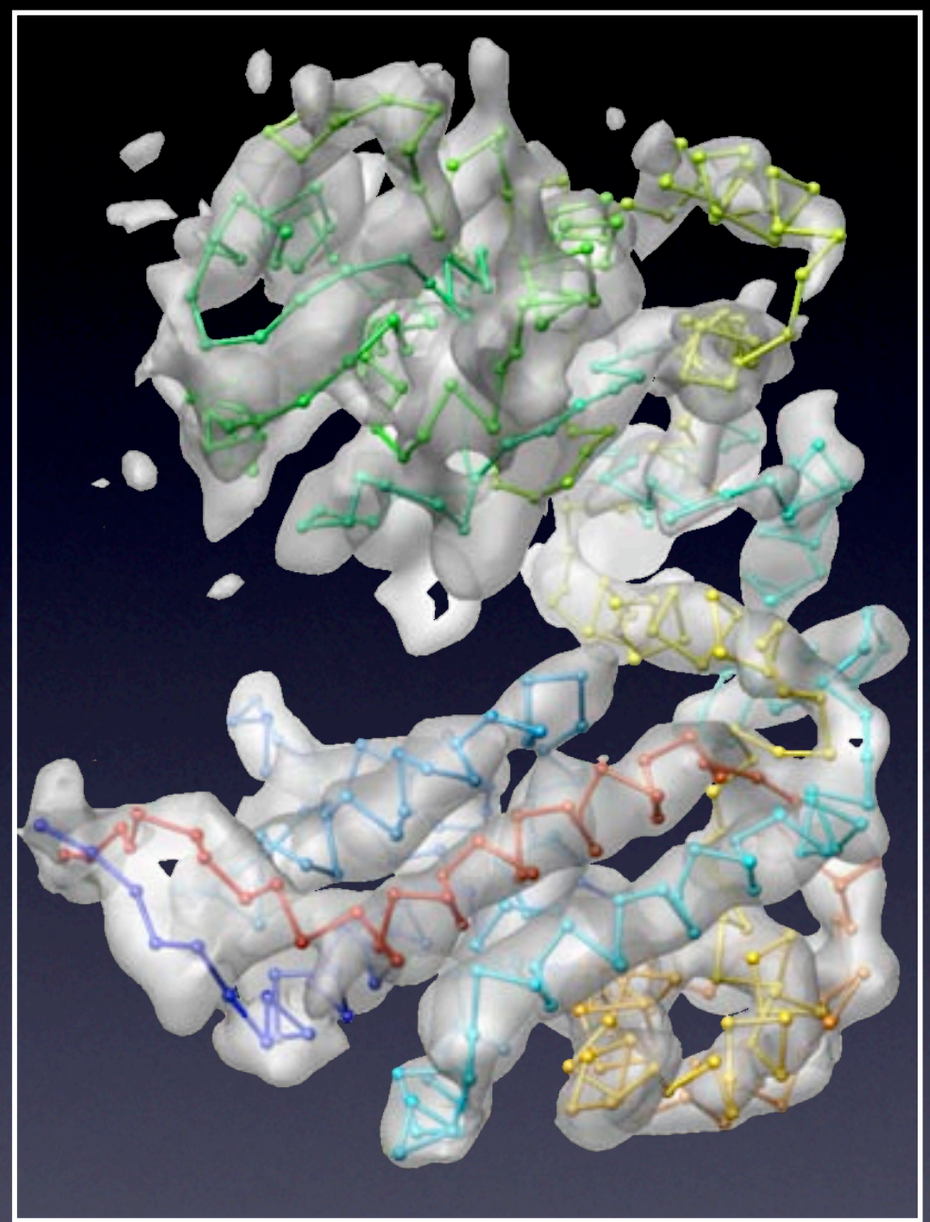
Single Particle Reconstruction: Resolution, Heterogeneity and Dynamics



Steve Ludtke

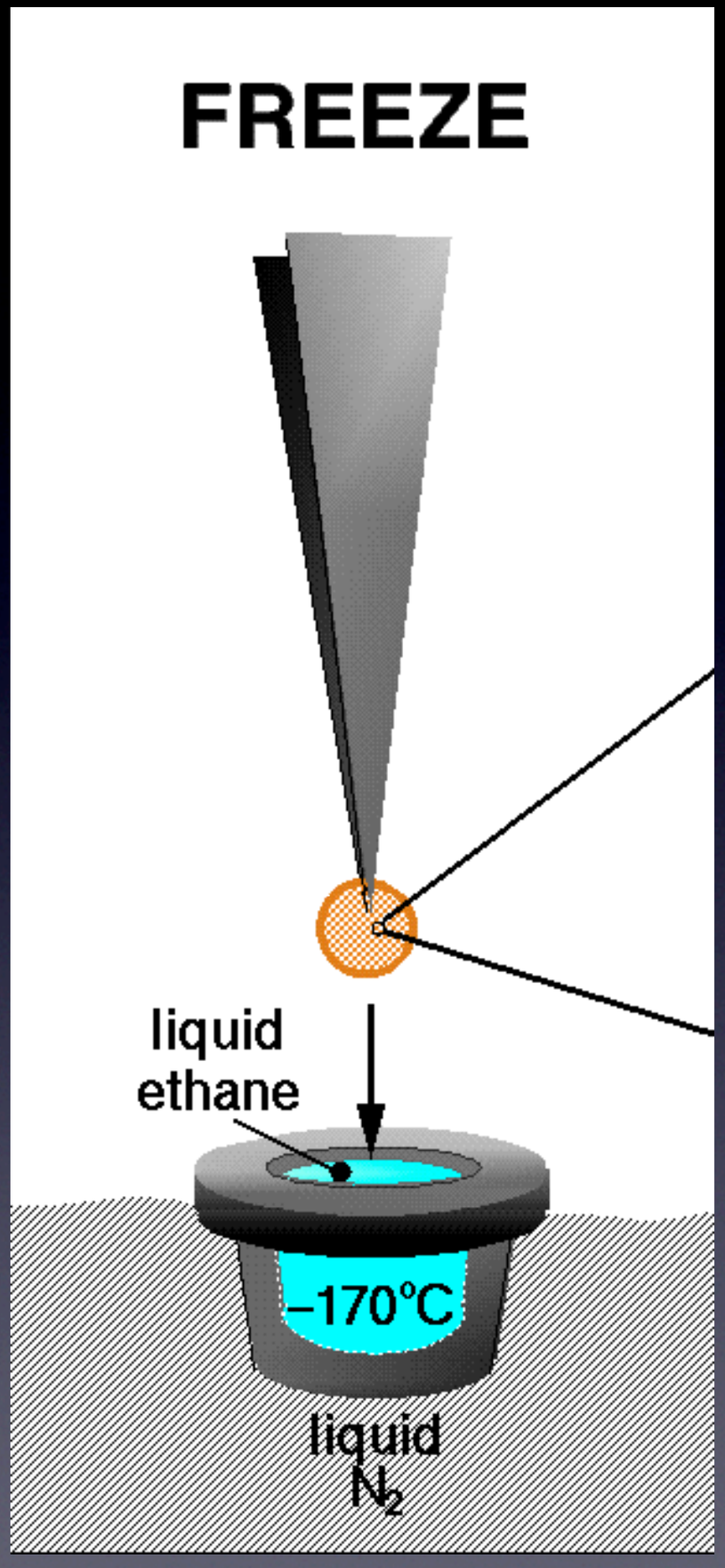
National Center for Macromolecular Imaging
Department of Biochemistry &
Molecular Biology
Baylor College of Medicine

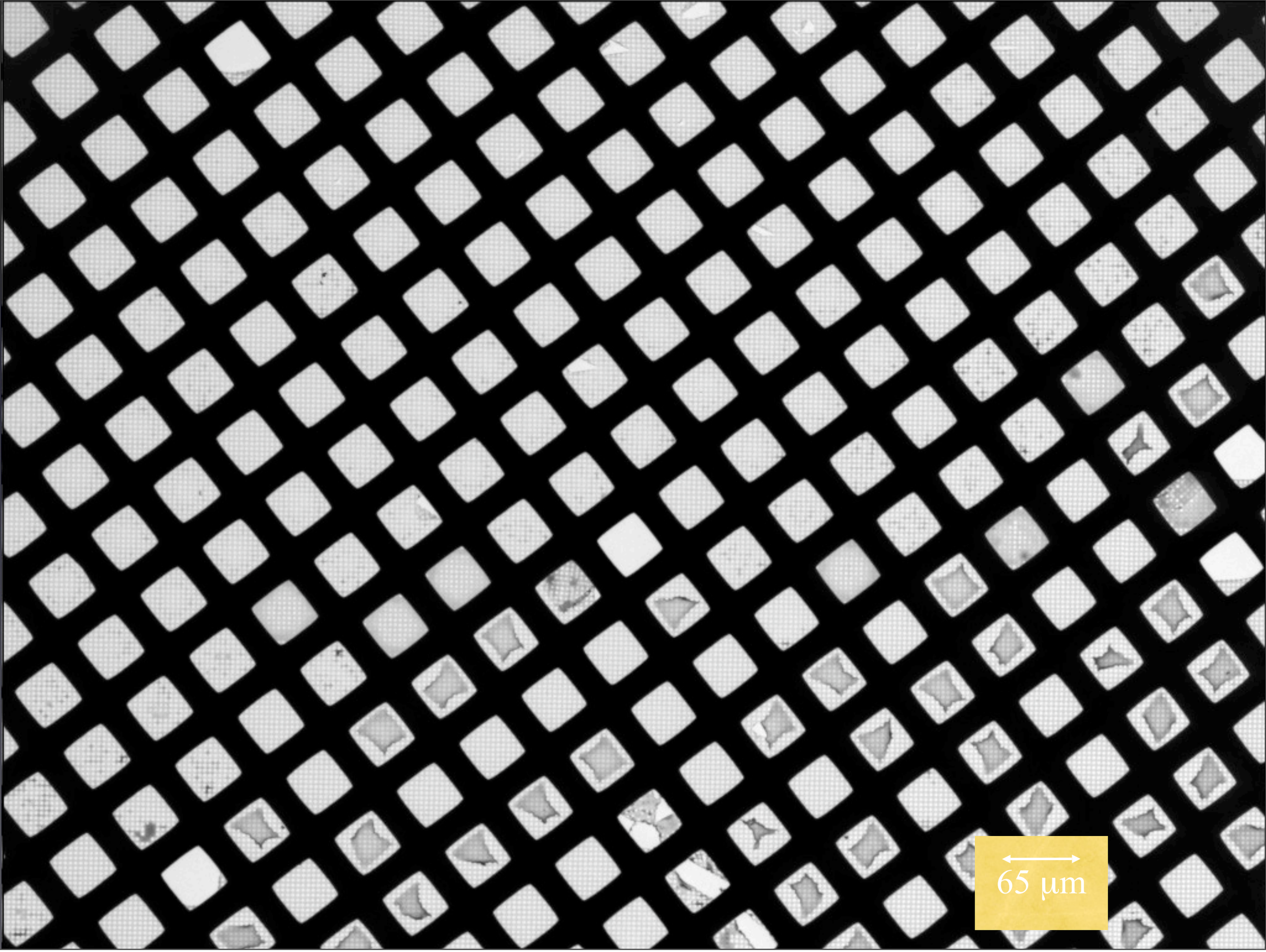
BCM Baylor College of Medicine

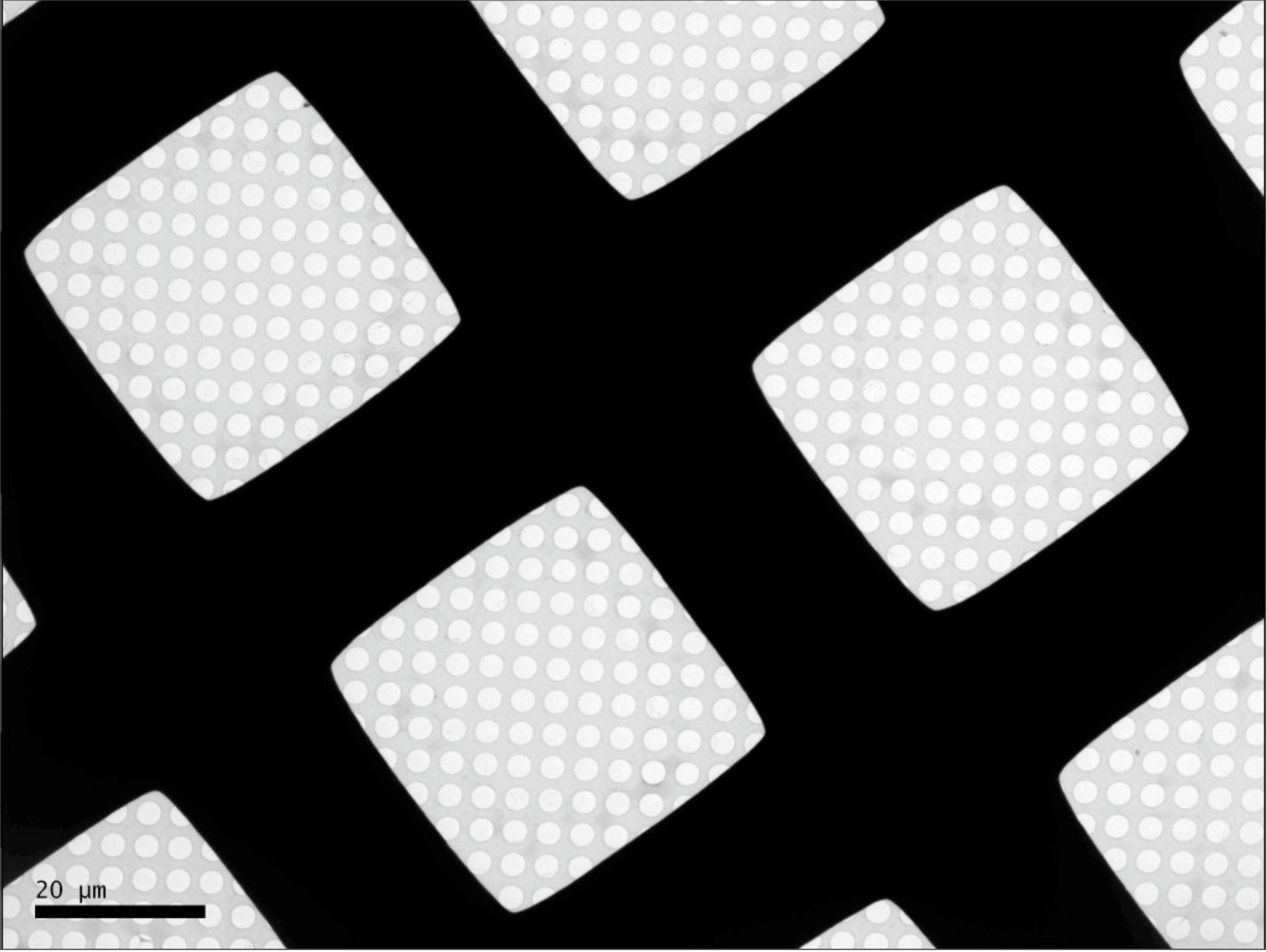
NCRR

NCMI National Center for Macromolecular Imaging

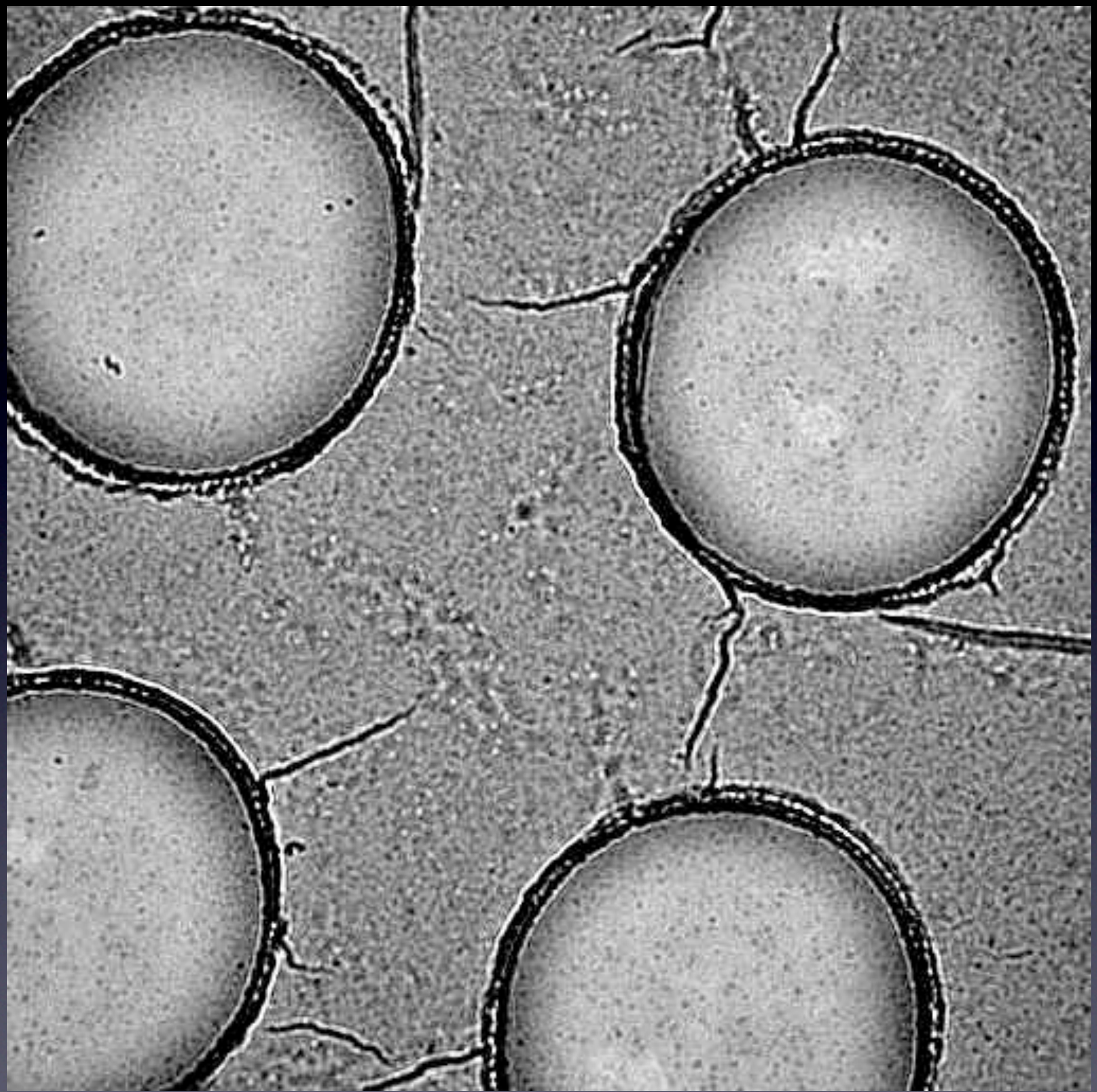
NIH Roadmap

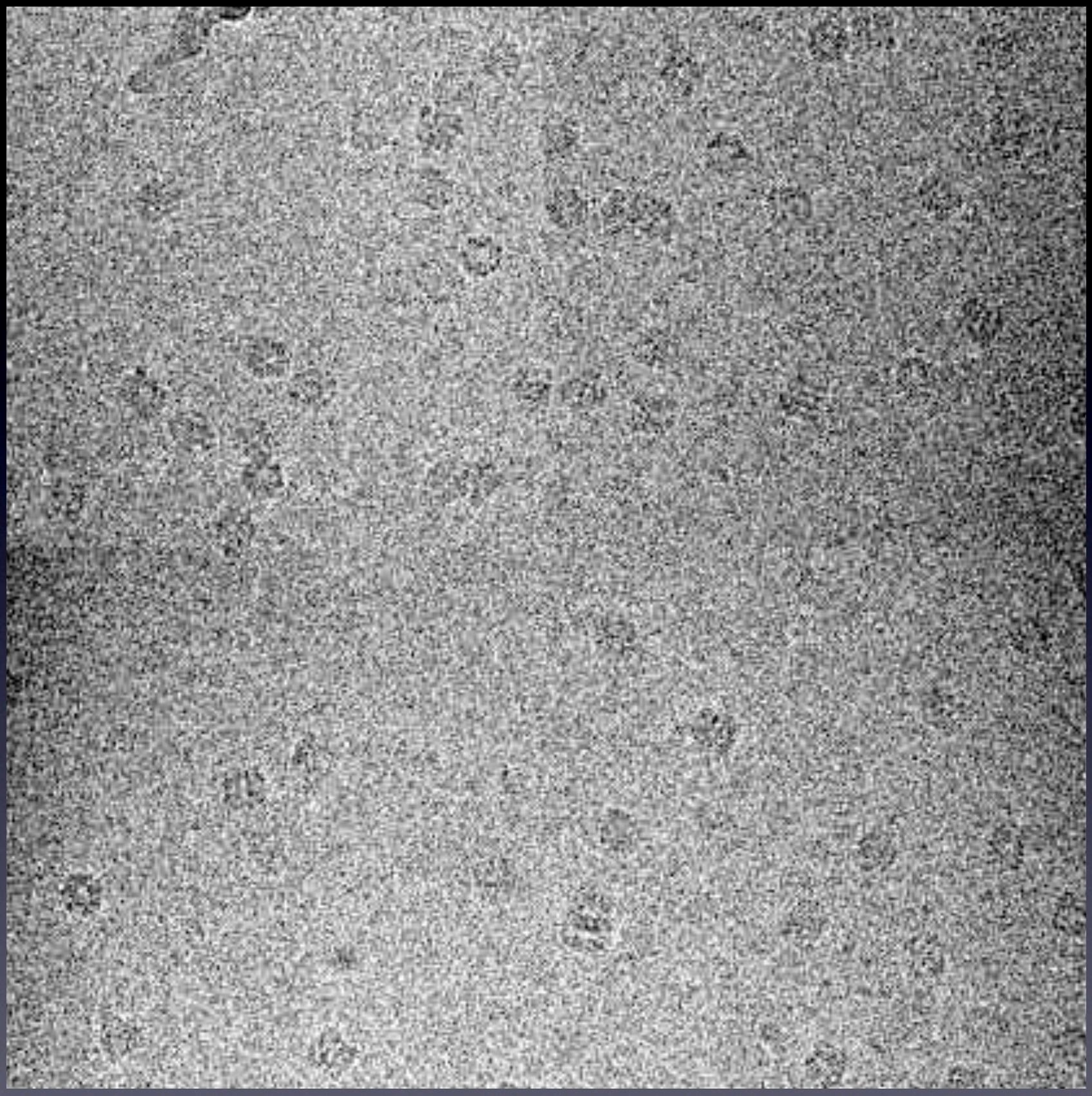

FREEZE
liquid
ethane
–170°C
liquid
N_2

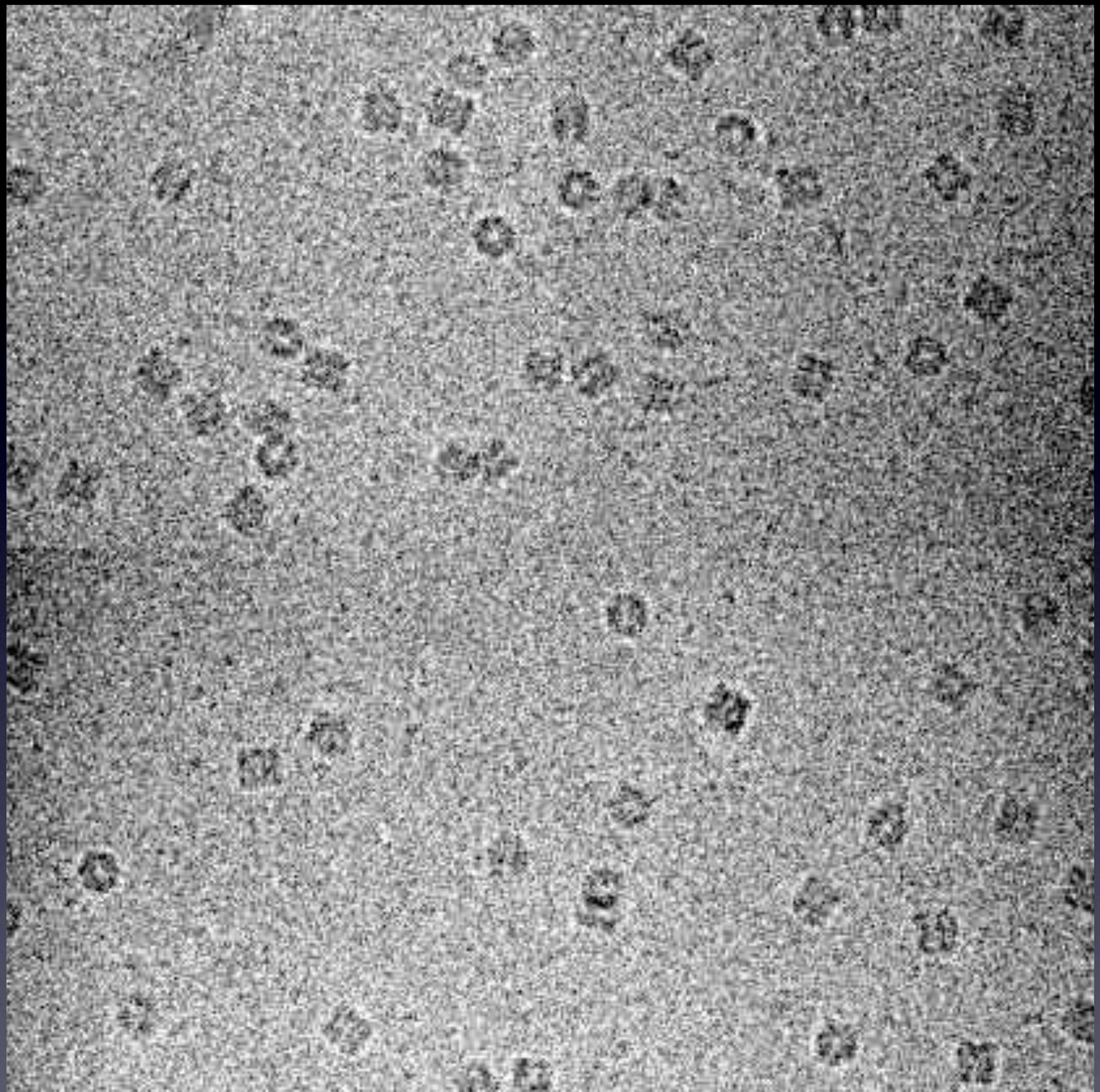

65 µm

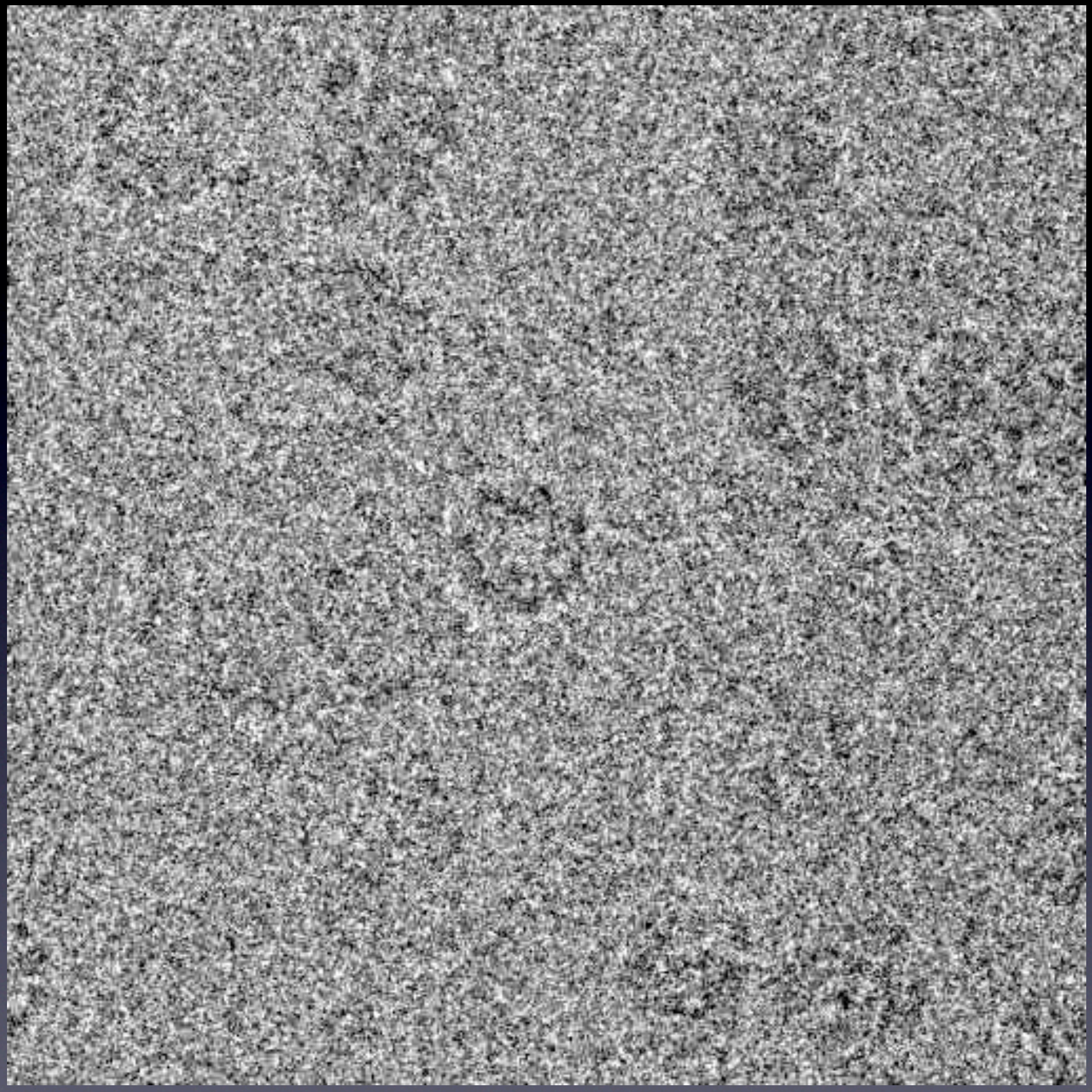


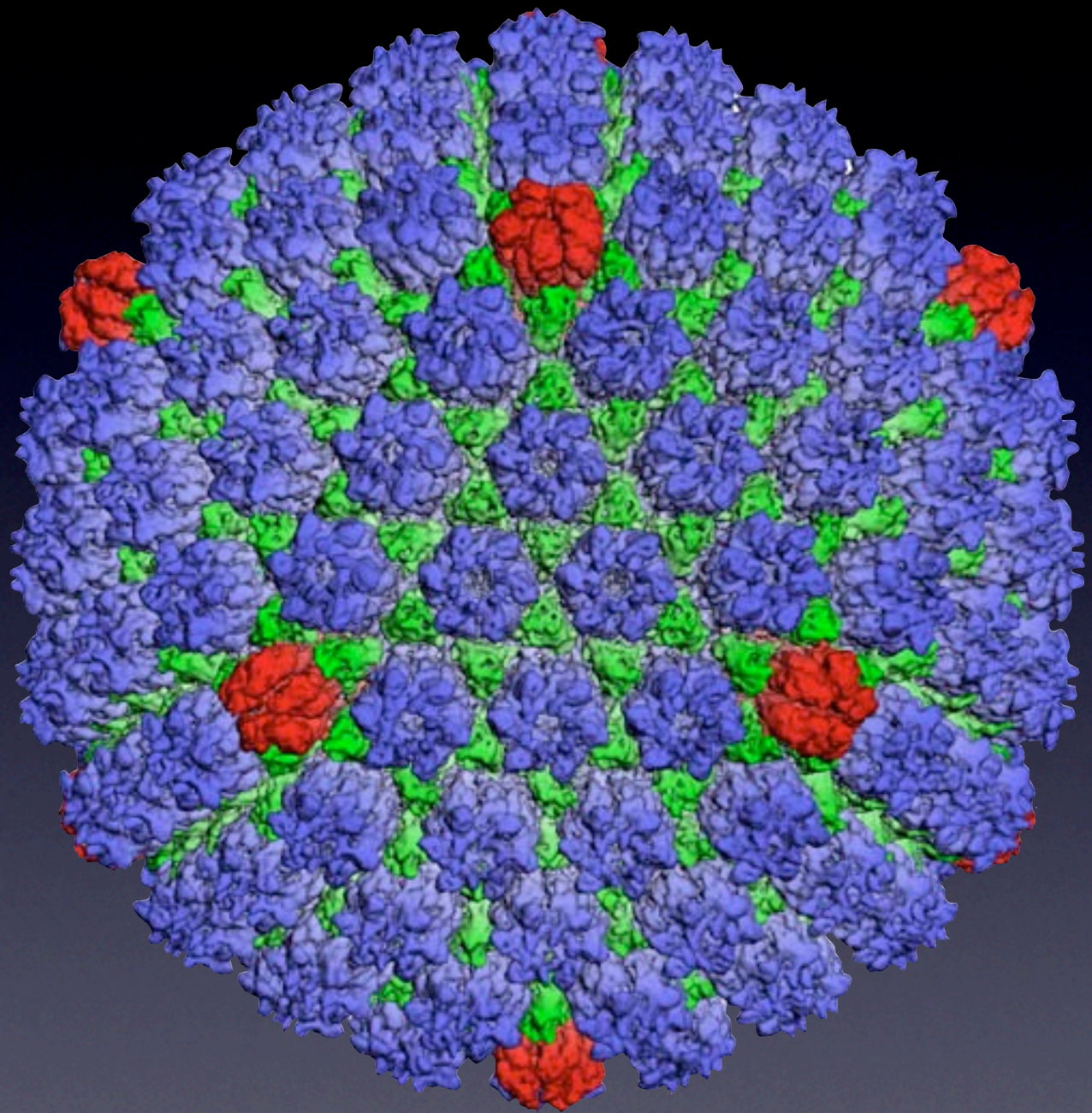
Tuesday, January 29, 2008


~1 μm

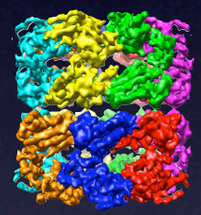

1000 Å


1000 Å

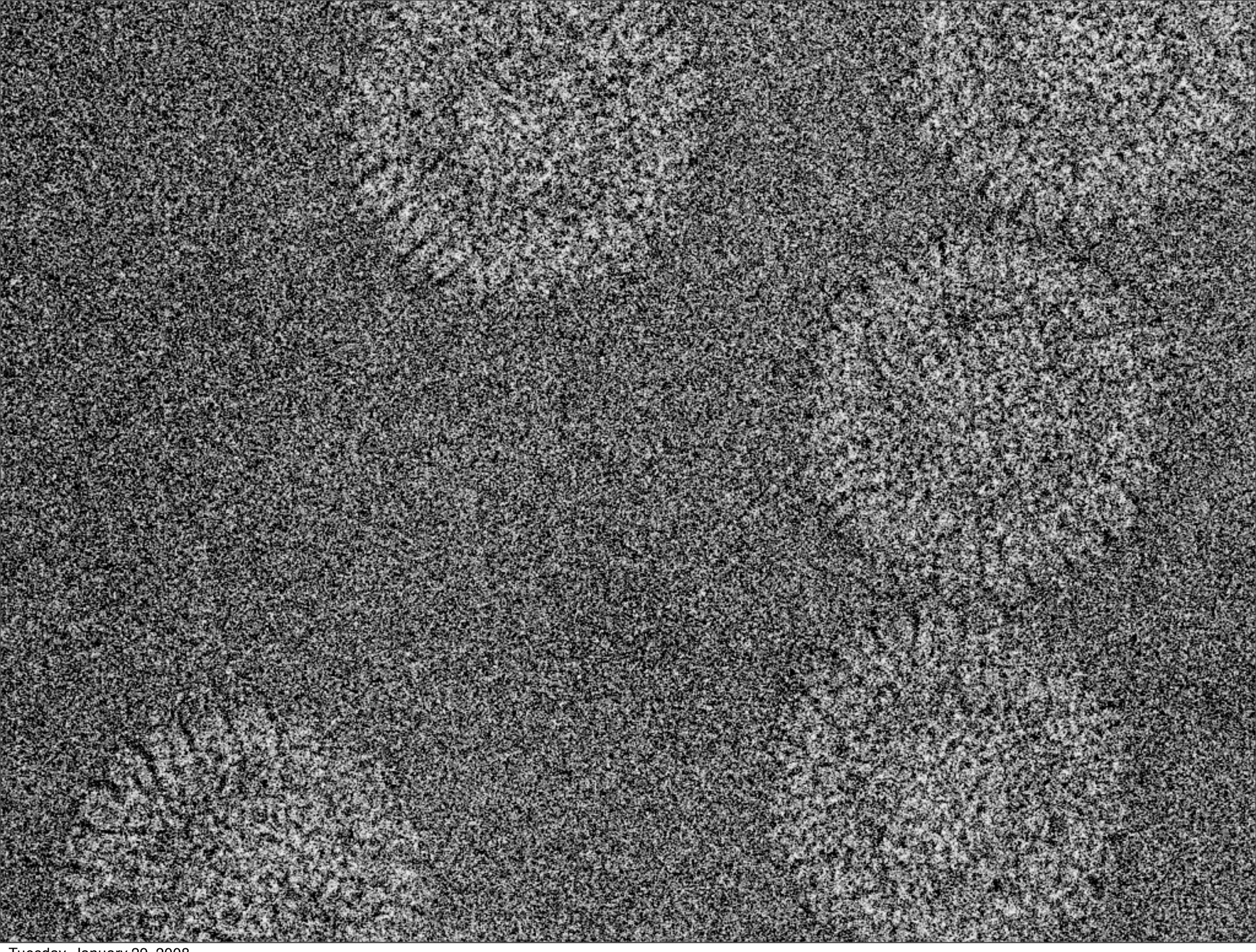

200 Å



200,000 kDa

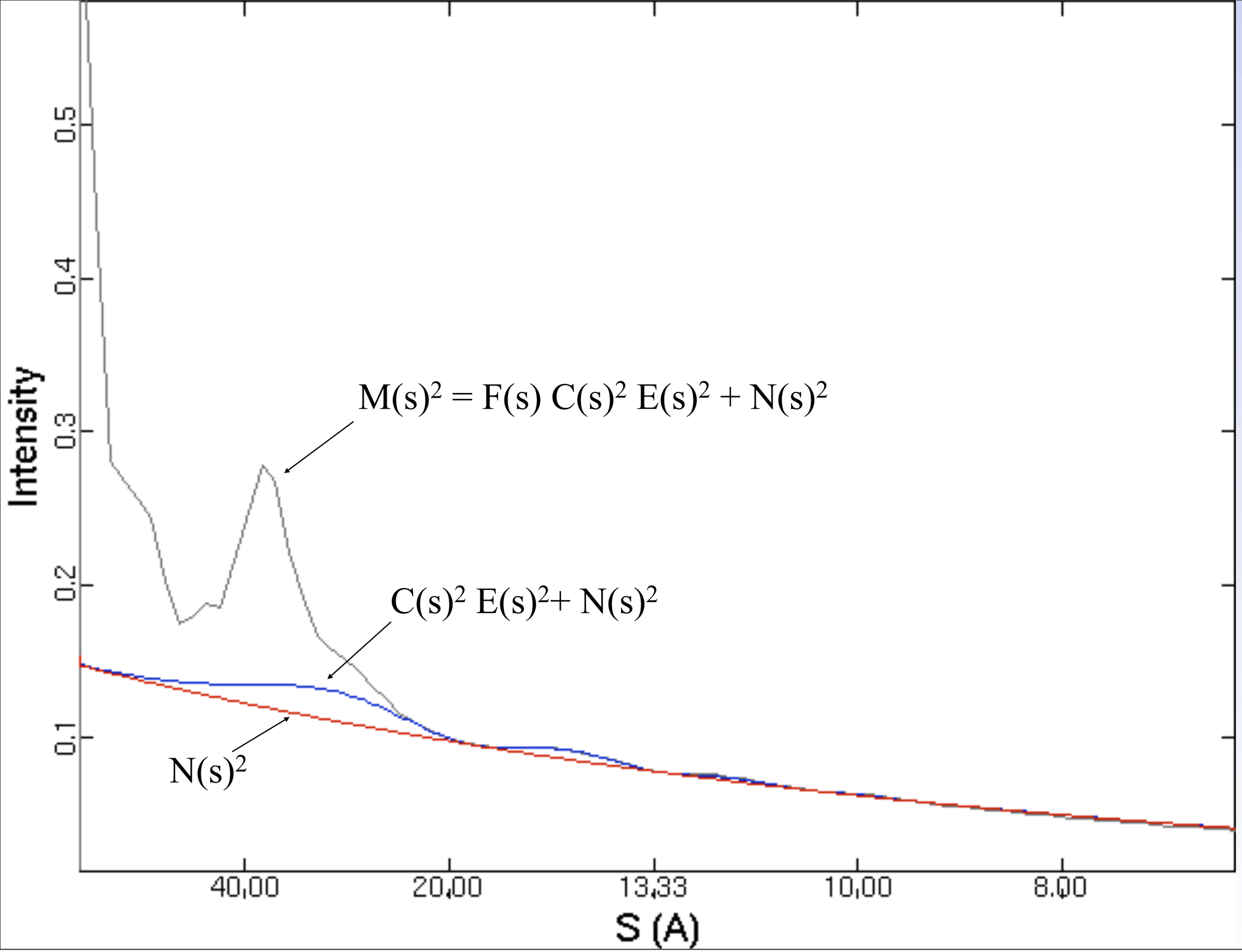


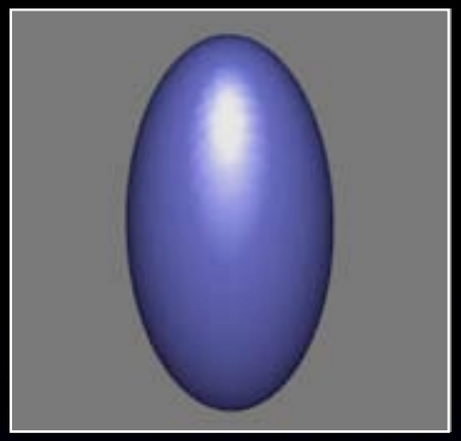
800 kDa



Tuesday, January 29, 2008

Tuesday, January 29, 2008


Intensity
0,5
0,4
0,3
0,2
0,1
M(s)² = F(s) C(s)² E(s)² + N(s)²
C(s)² E(s)²+ N(s)²
N(s)²
40,00
20,00
13,33
10,00
8,00
S (A)



Initial 3D Model

Uniform Projections

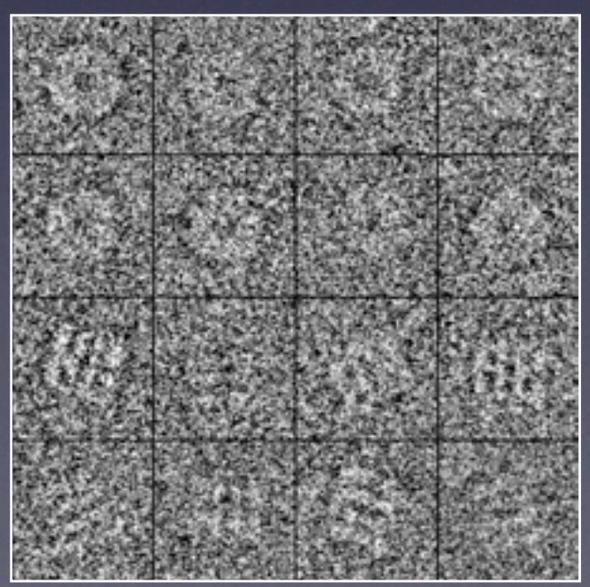
Build New 3D Model

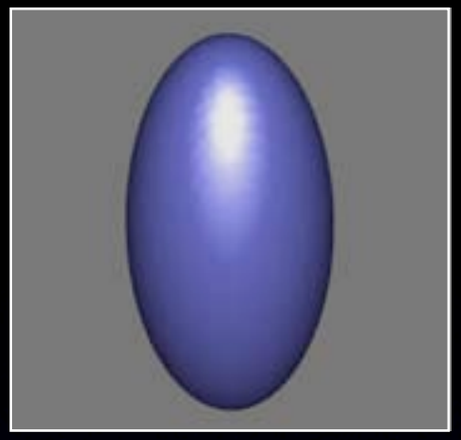
Final 3D Model

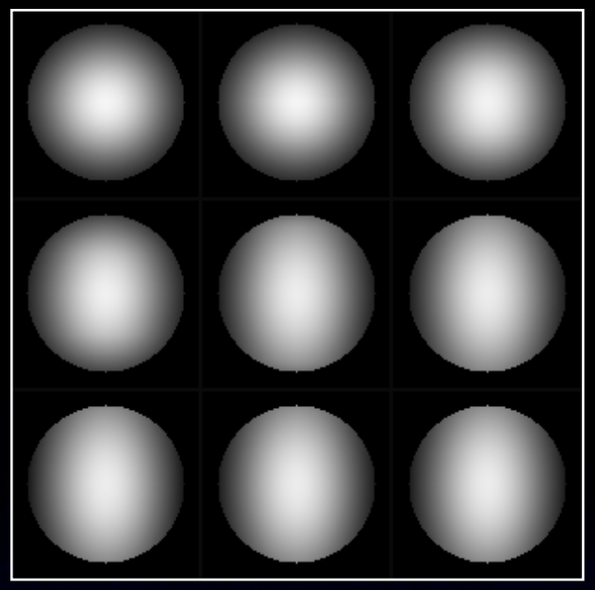
Particle Images

Classify Particles

Align and Average Classes







Initial 3D Model

Uniform Projections

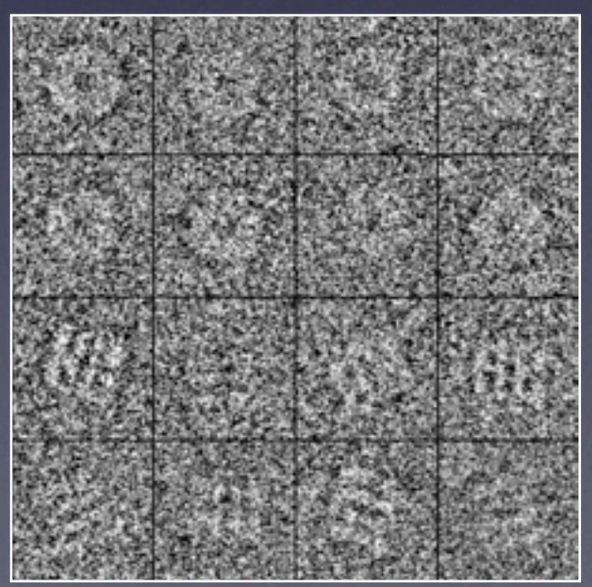
Build New 3D Model

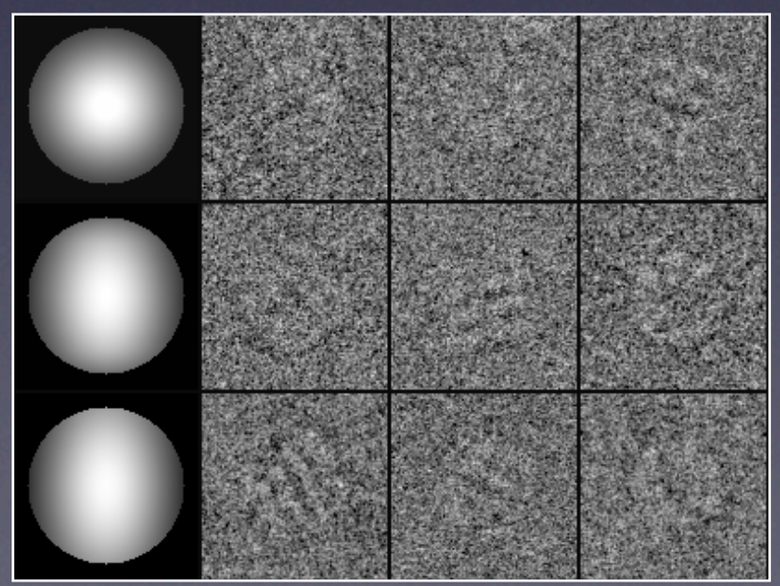
Final 3D Model

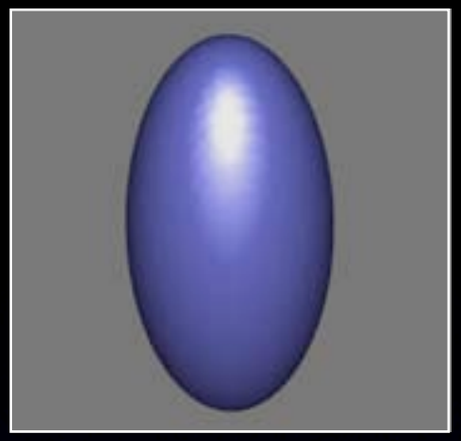
Particle Images

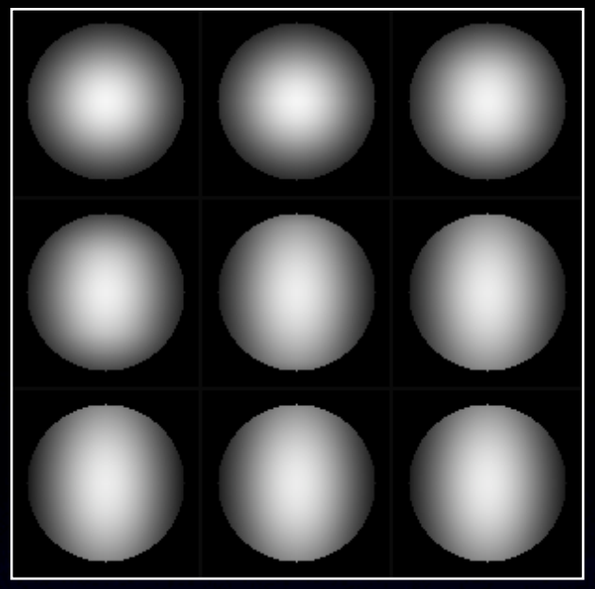
Classify Particles

Align and Average Classes









Initial 3D Model

Uniform Projections

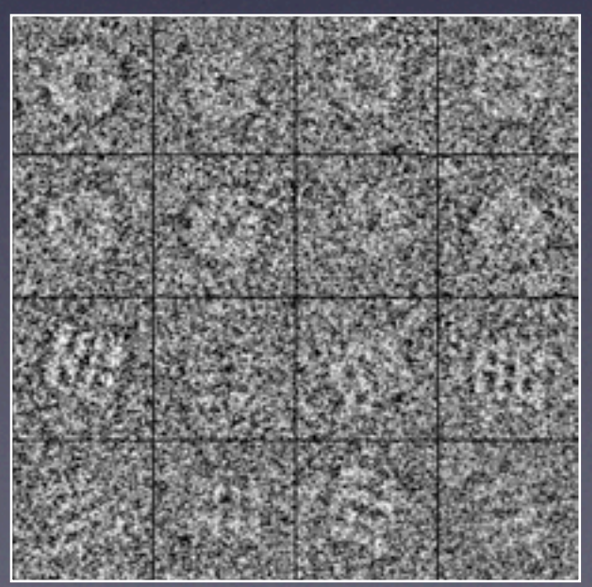
Build New 3D Model

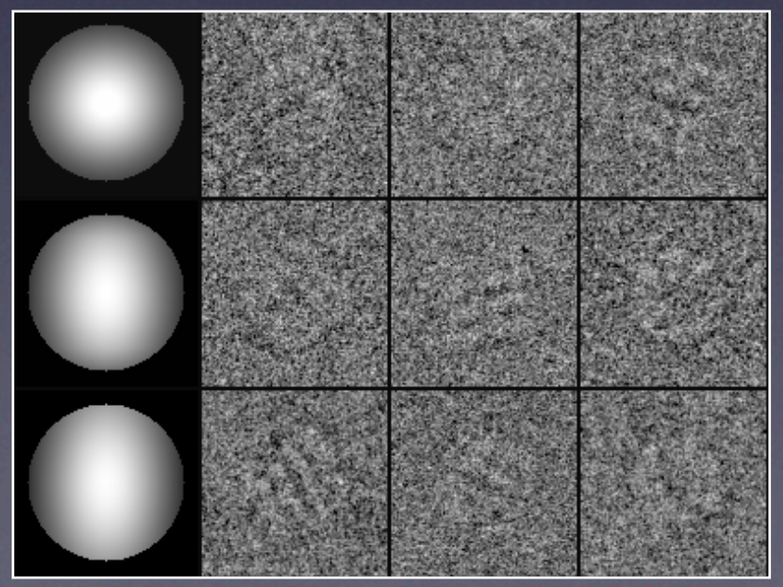
Final 3D Model

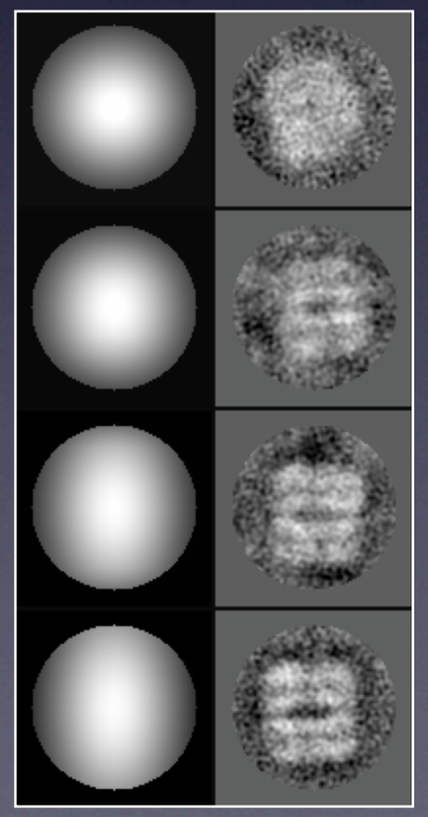
Particle Images

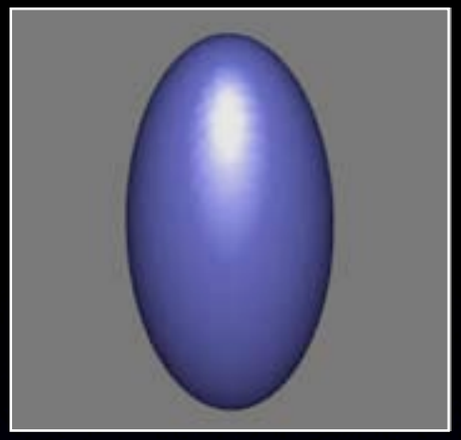
Classify Particles

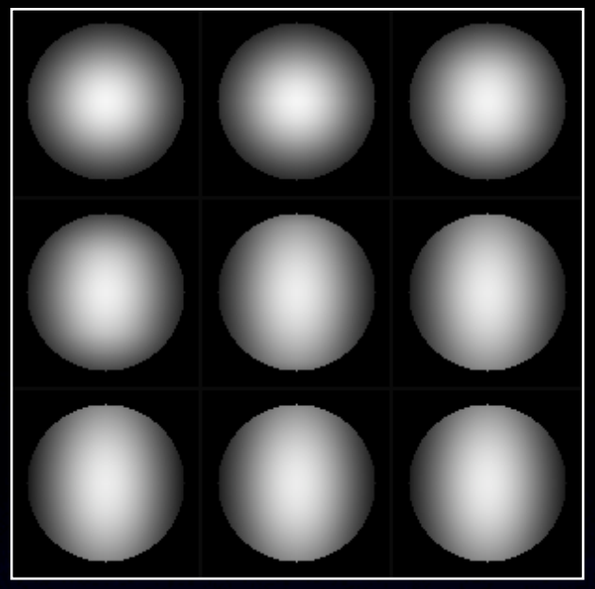
Align and Average Classes

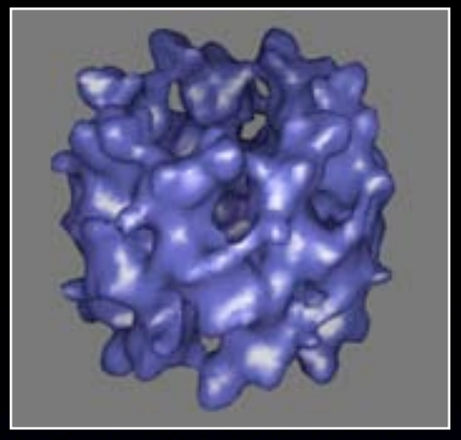












Initial 3D Model

Uniform Projections

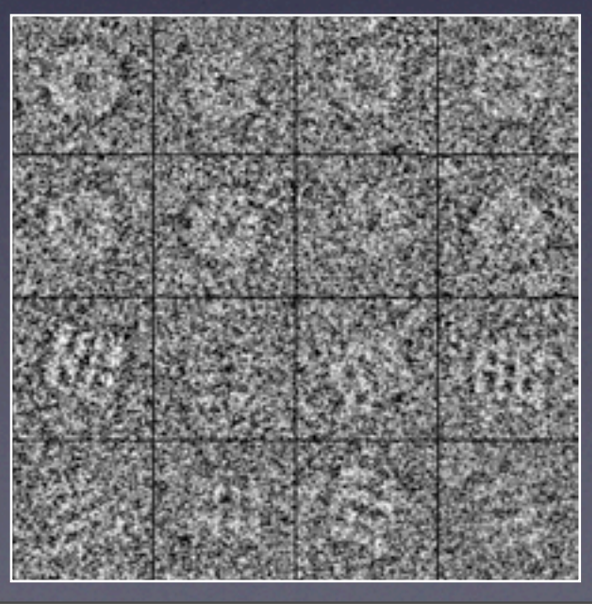
Build New 3D Model

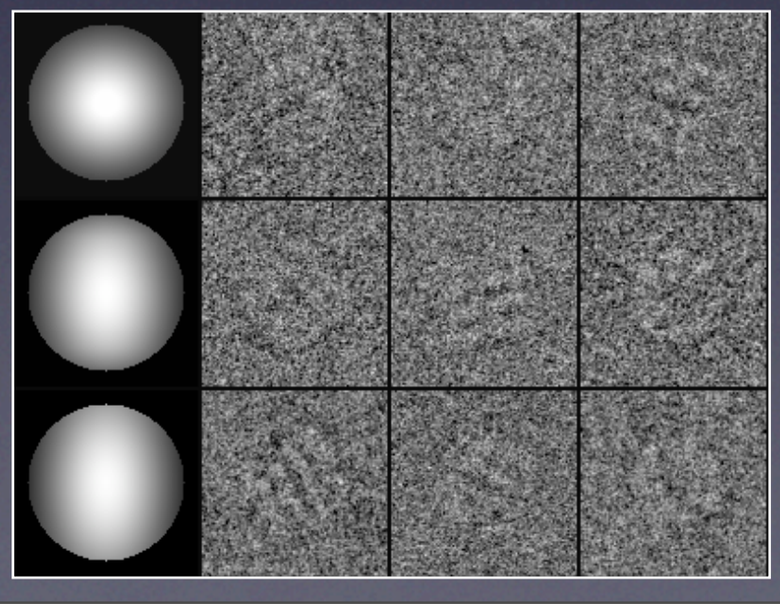
Final 3D Model

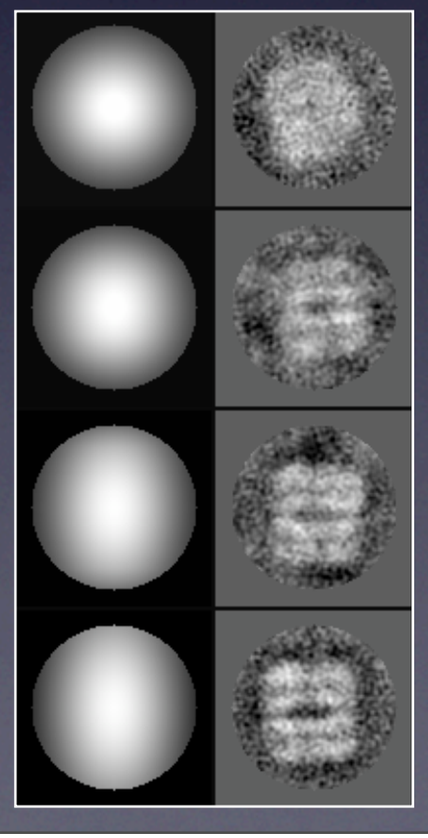
Particle Images

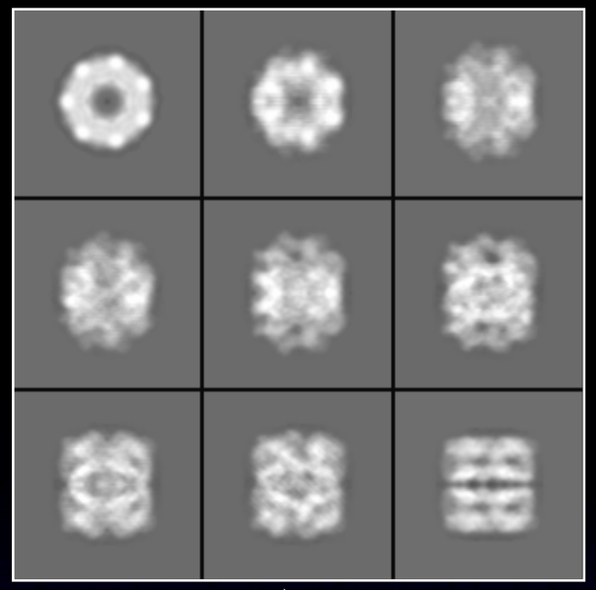
Classify Particles

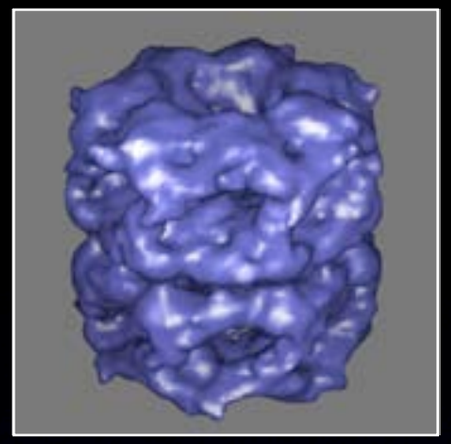
Align and Average Classes

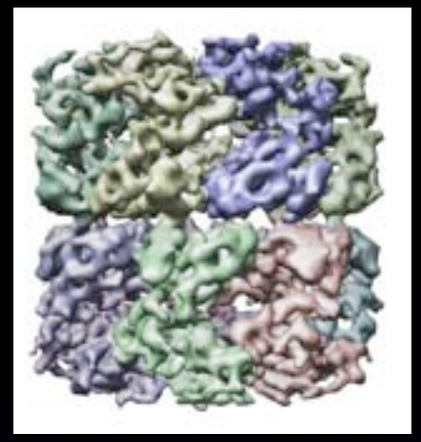












Initial 3D Model

Uniform Projections

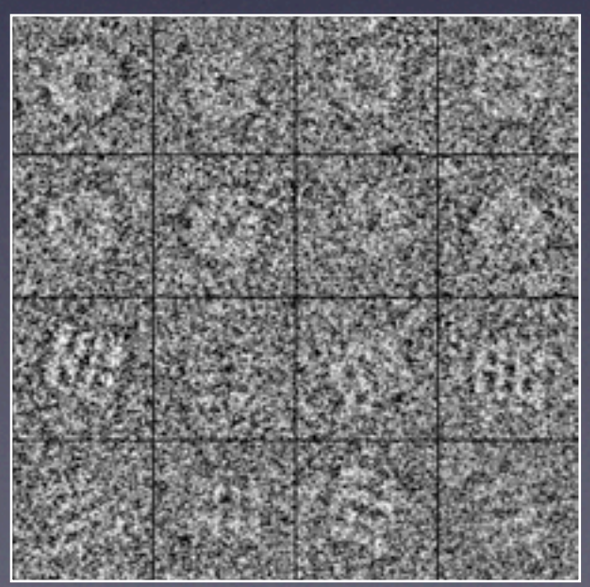
Build New 3D Model

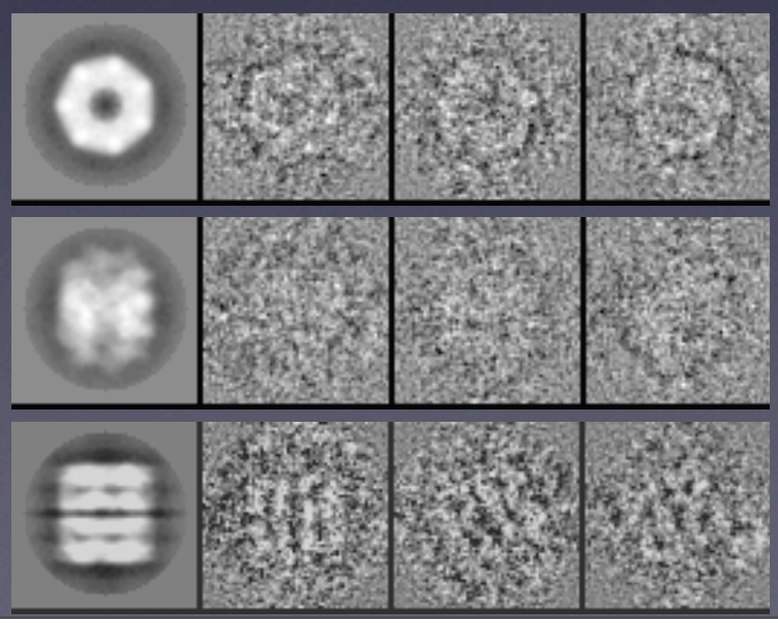
Final 3D Model

Particle Images

Classify Particles

Align and Average Classes



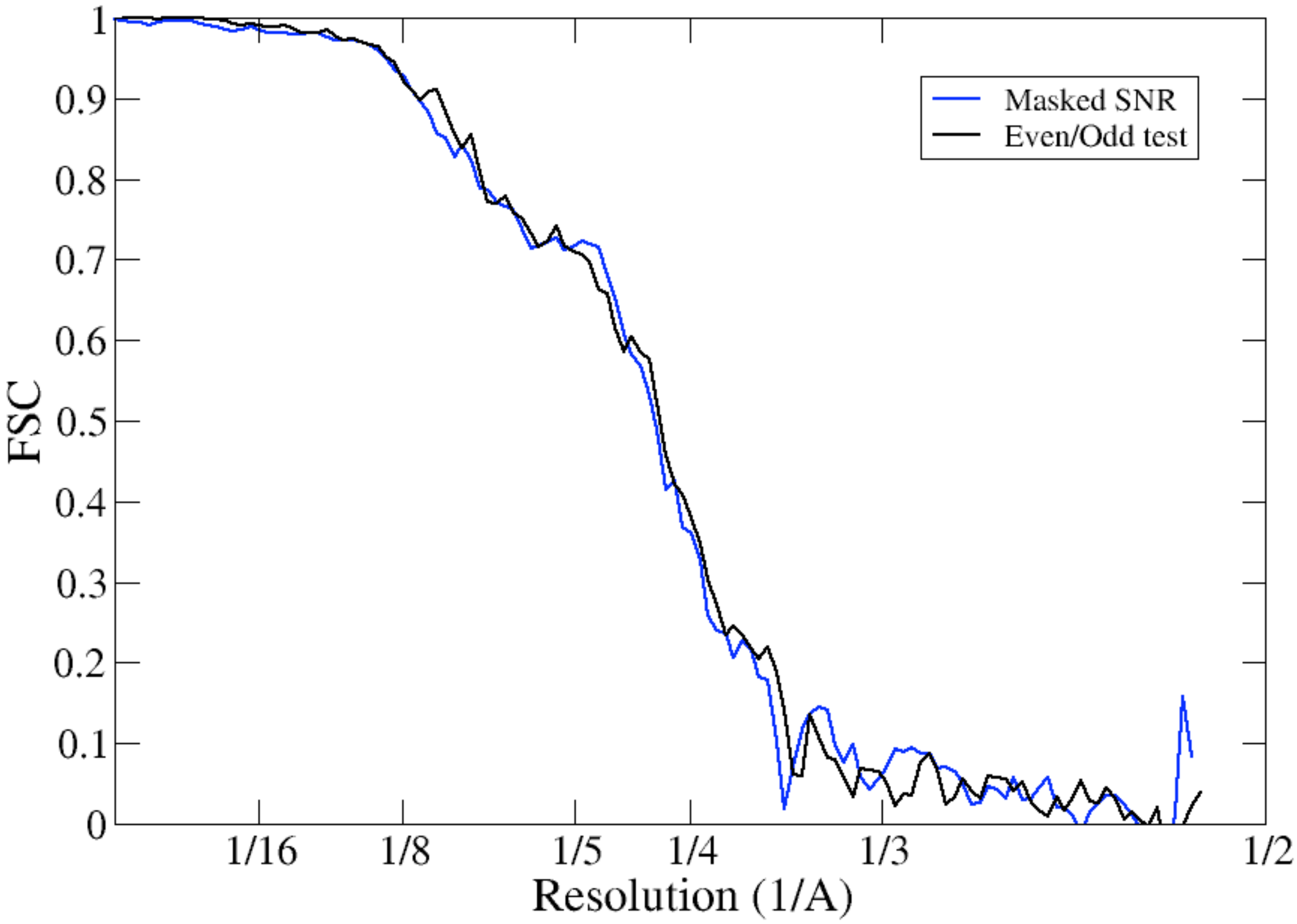


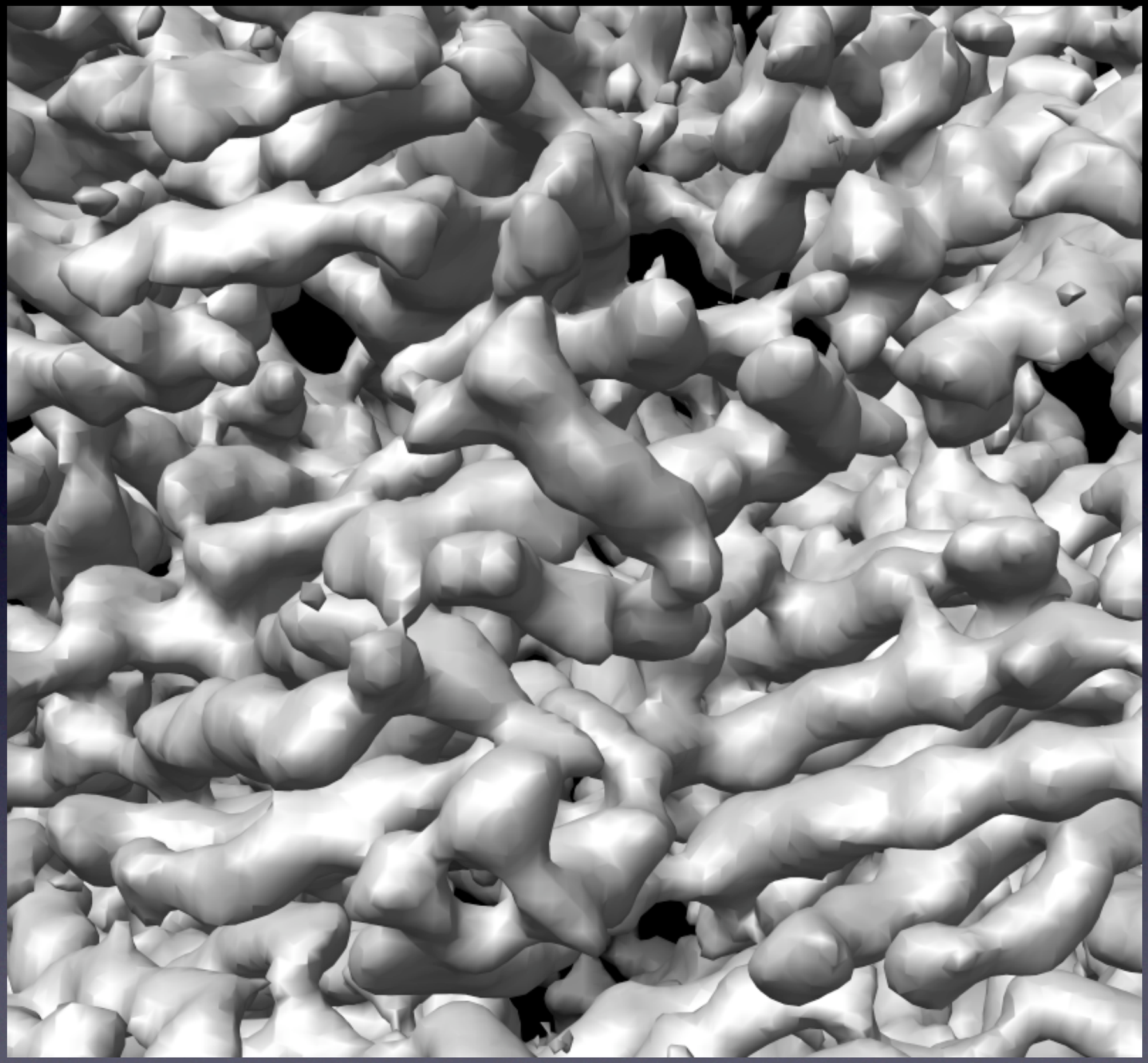
GroEL

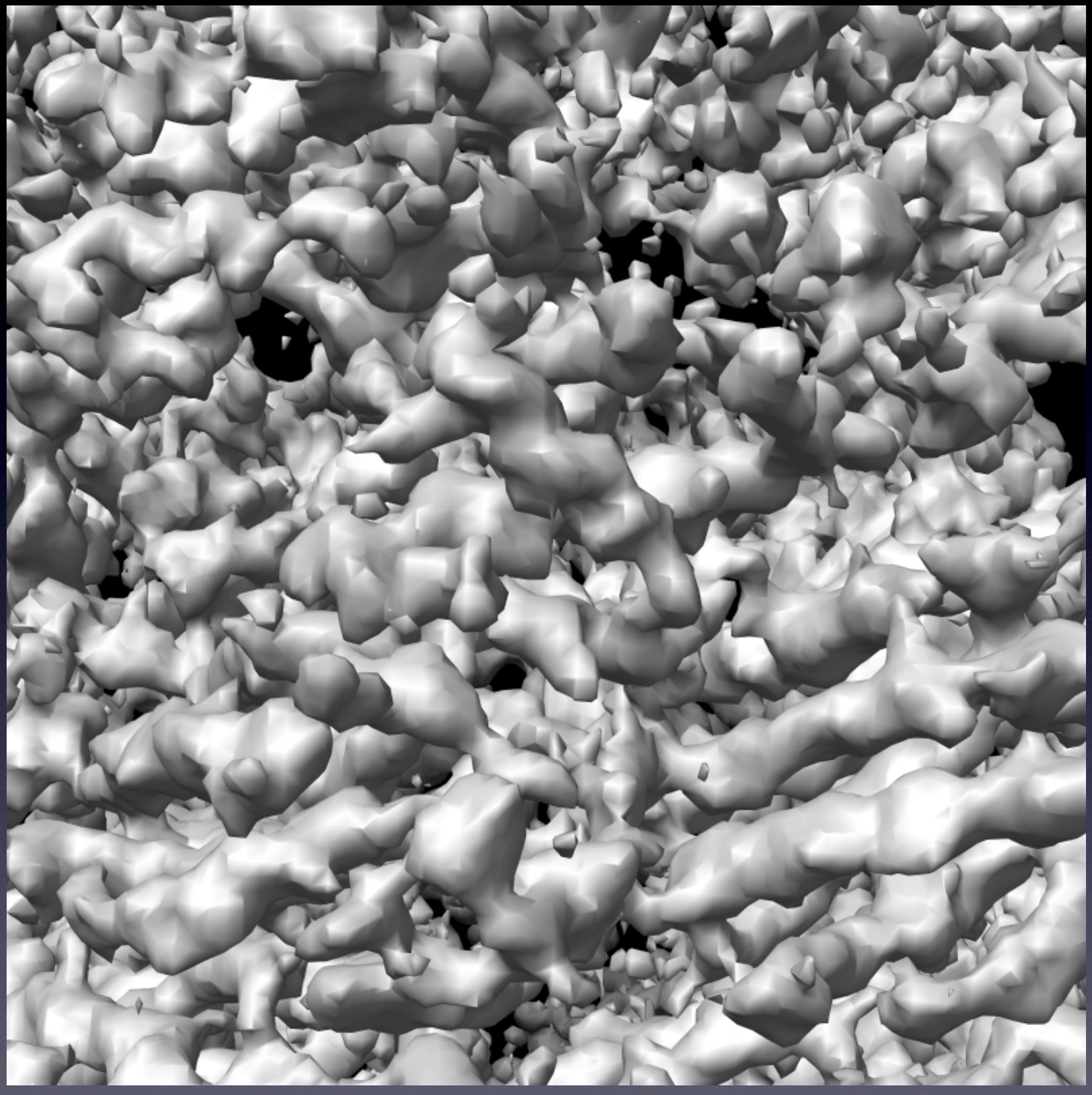
- Type I chaperonin → GroES co-chaperonin
- homo 14-mer
- 2 back to back 7 membered rings
- ~800 kDa, ~200 Å

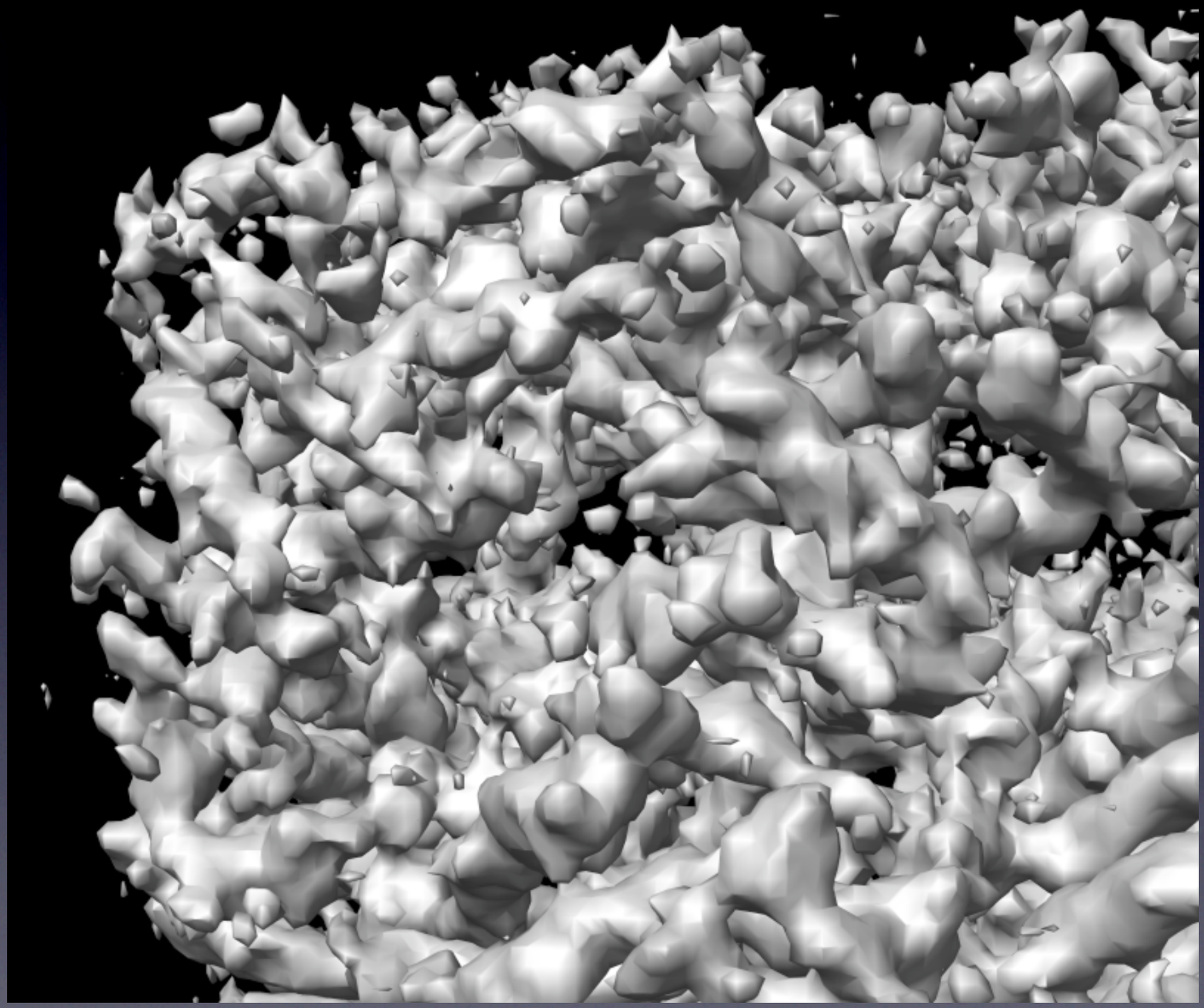
Data Collected 2005

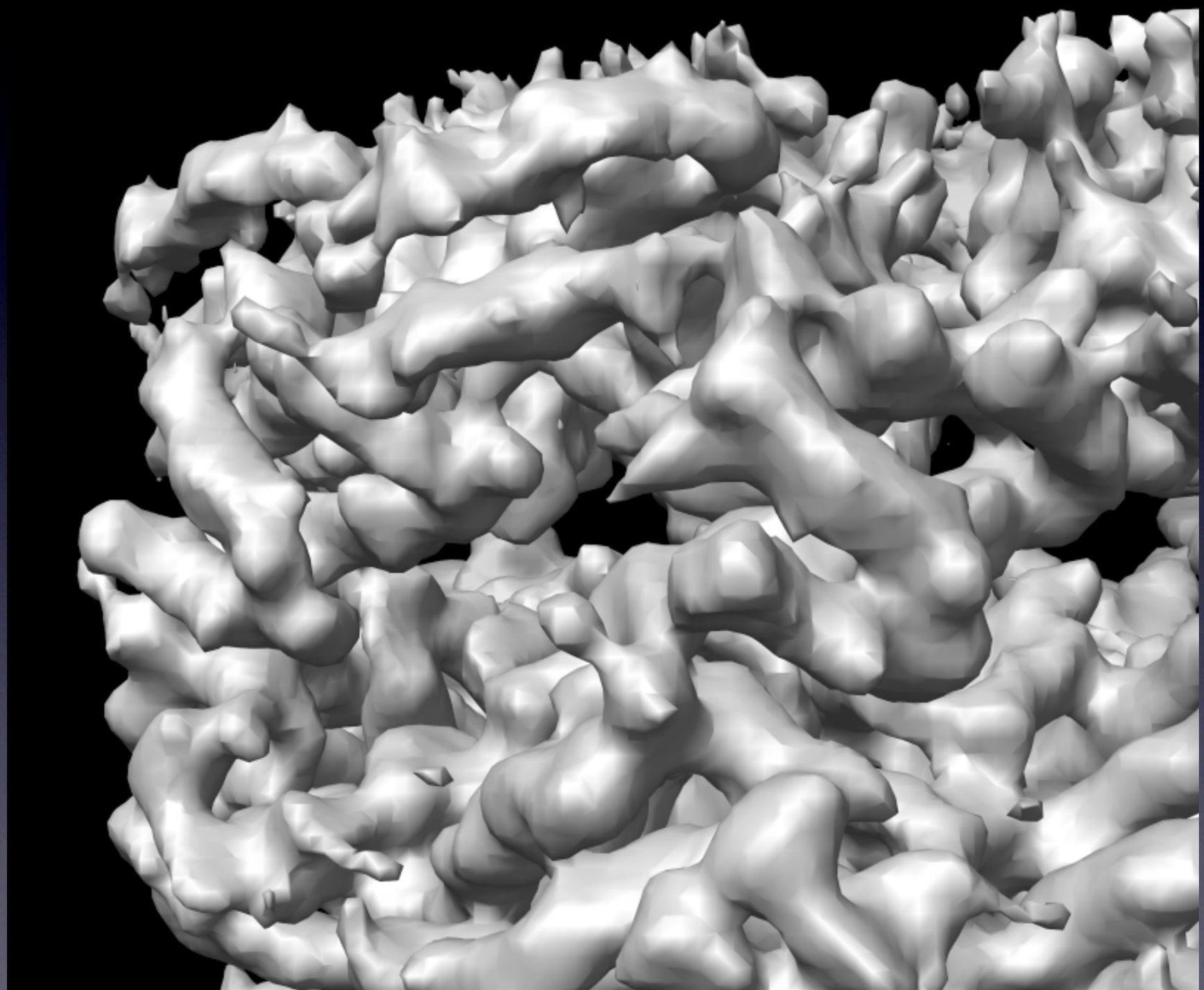
- Native, unliganded GroEL, no ATP/ADP (?)
- JEOL 3000SFF (Yoshi-style) at LHe temp
- 6 microscopy sessions, Film
- 825 micrographs, Nikon 9000 @ 6.35 µm scan step
- 60k mag → 1.06 Å/pix
- 135 micrographs used → 20,401 particles
- Defocus 1.2 – 2.2 µm

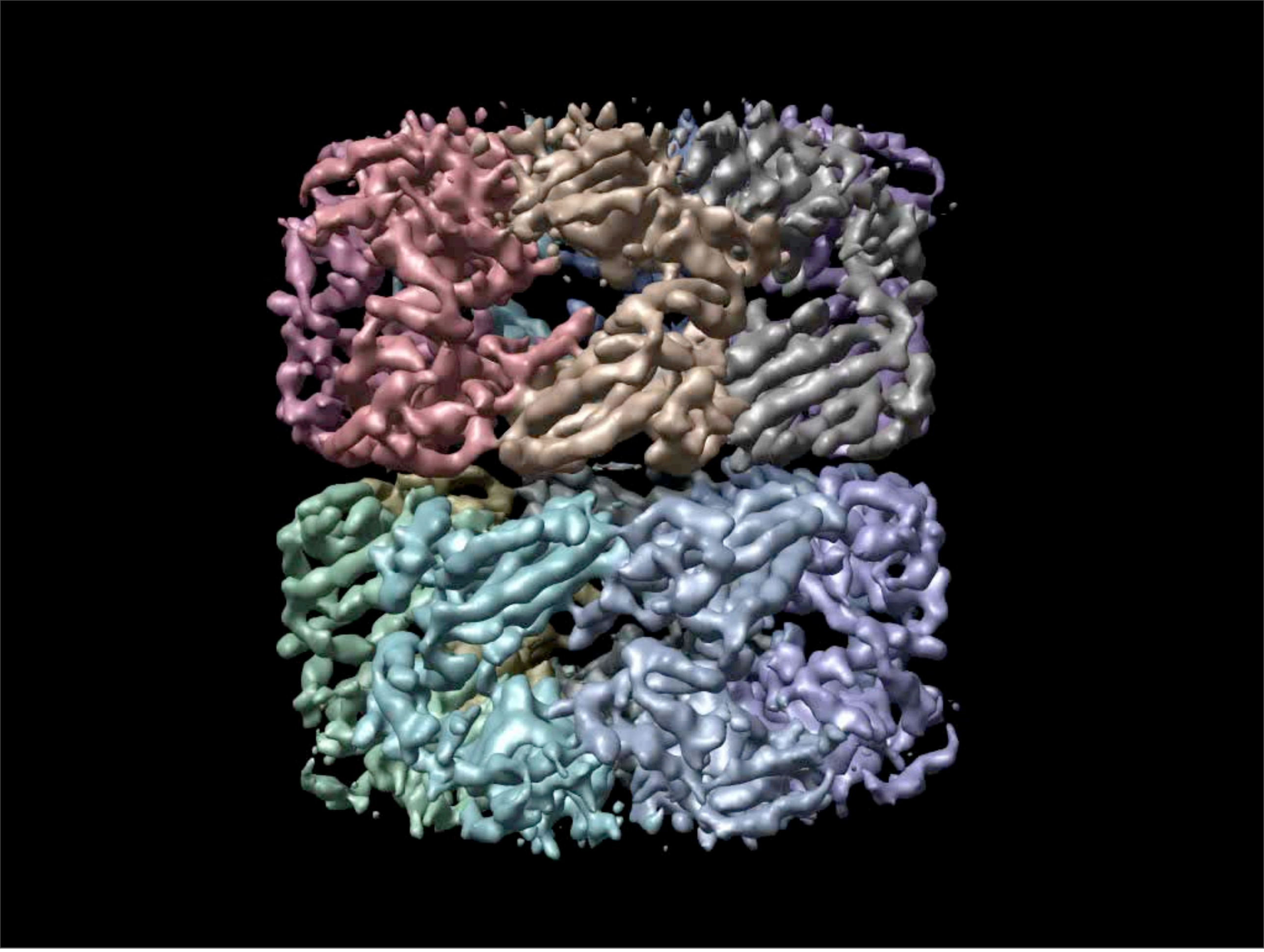

1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
FSC
Masked SNR
Even/Odd test
1/16
1/8
1/5
1/4
1/3
1/2
Resolution (1/A)

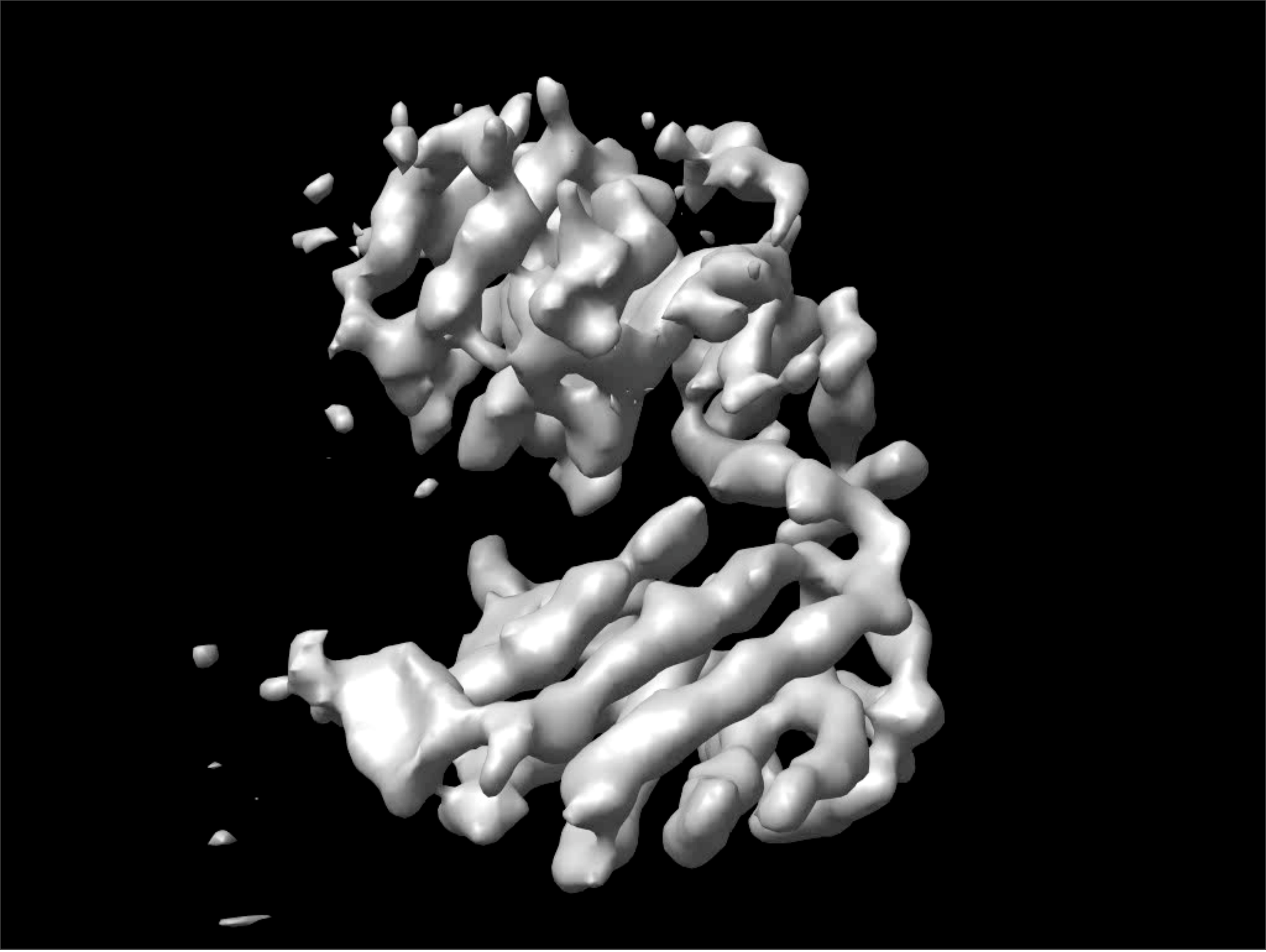


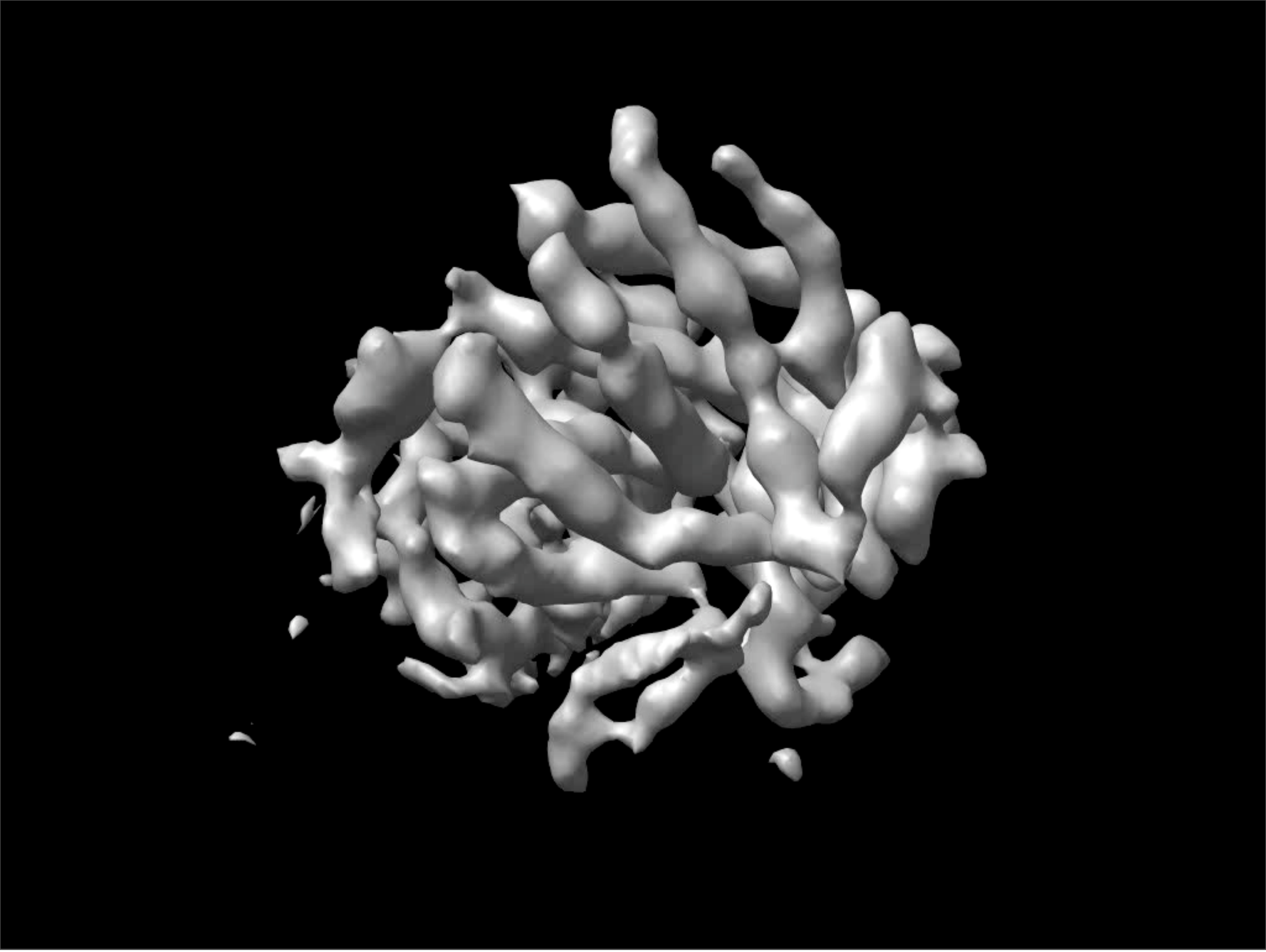






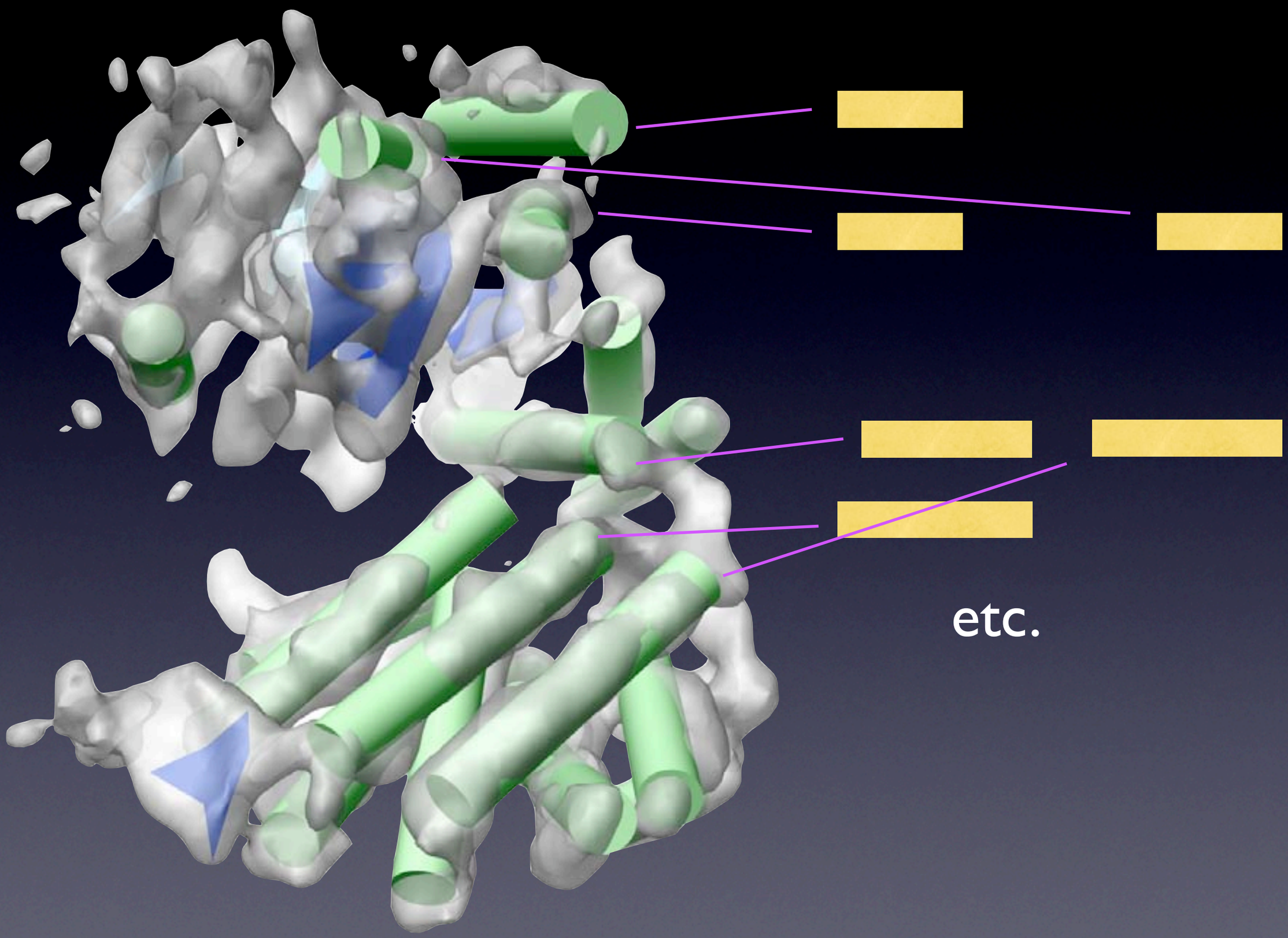




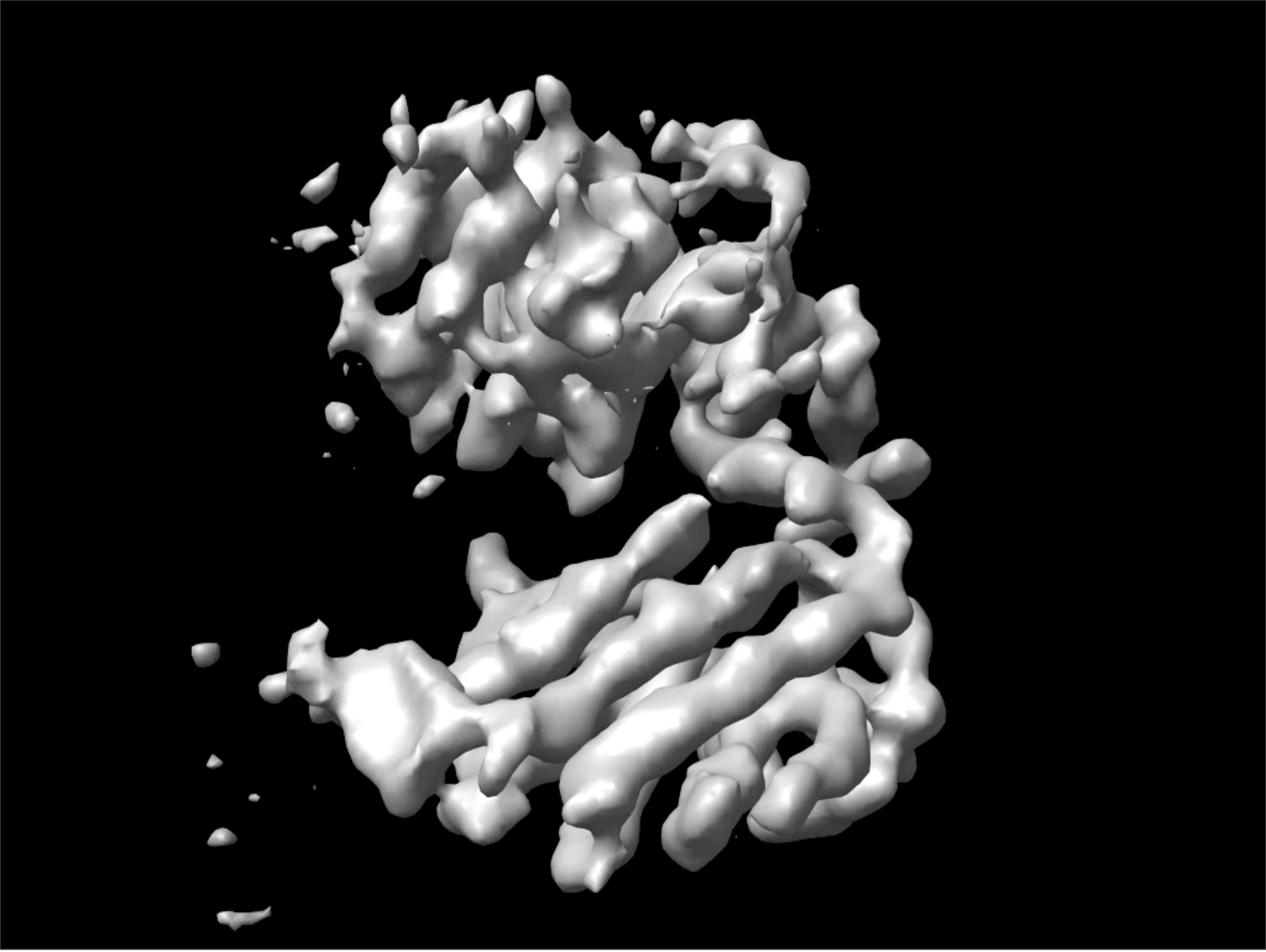


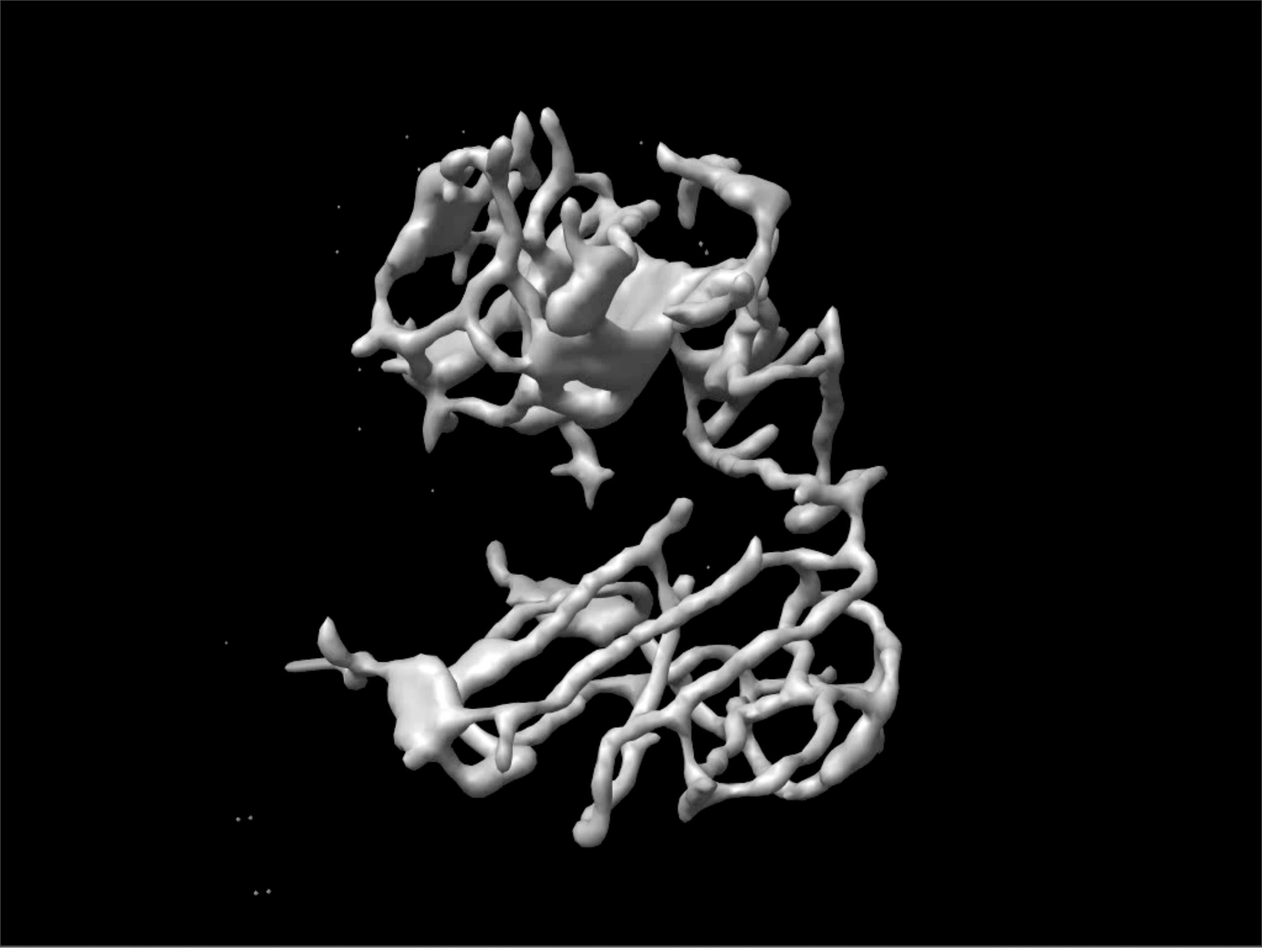
How much can we get at 4 Å ?

SSEHunter



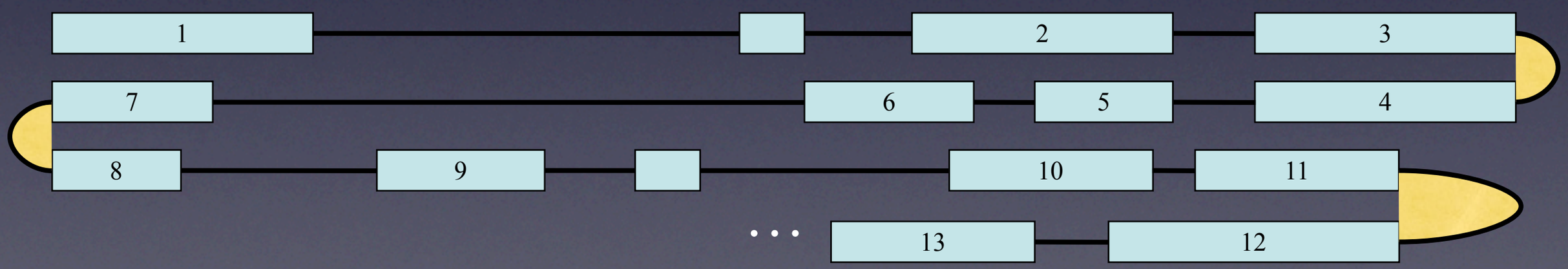
etc.





Sequence Based Secondary Structure Prediction

```
1    AAKDVKFGNDAGVKMLRGVNVLADAVKVTLGPKGRNVVLDKSFGAPTITKDGVSVAREIE

61   LEDKFENMGAQMVKEVASKANDAAGDGTTTATVLAQAIITEGLKAVAAGMNPMDLKRGID

121  KAVTVAVEELKALSVPCSDSKAIAQVGTISANSDETVGKLIAEAMDKVGKEGVITVEDGT

181  GLQDELDVVEGMQFDRGYLSPYFINKPETGAVELESPFILLADKKISNIREMLPVLEAVA

241  KAGKPLLIIAEDVEGEALATAVVNTIRGIVKVAAVKAPGFGDRRKAMLQDIATLTGGTVI

301  SEEIGMELEKATLEDLGQAKRVVINKDTTTIIDGVGEEAAIQGRVAQIRQQIEEATSDYD

361  REKLQERVAKLAGGVAVIKVGAATEVEMKEKKARVEDALHATRAAVEEGVVAGGGVALIR

421  VASKLADLRGQNEDQNVGIKVALRAMEAPLRQIVLNCGEEPSVVANTVKGGDGNYGYNAA

481  TEEYGNMIDMGILDPTKVTRSALQYAASVAGLMITTECMVTDLPKNDAADLGAAGGMGGM

541  GGMGGMM
```

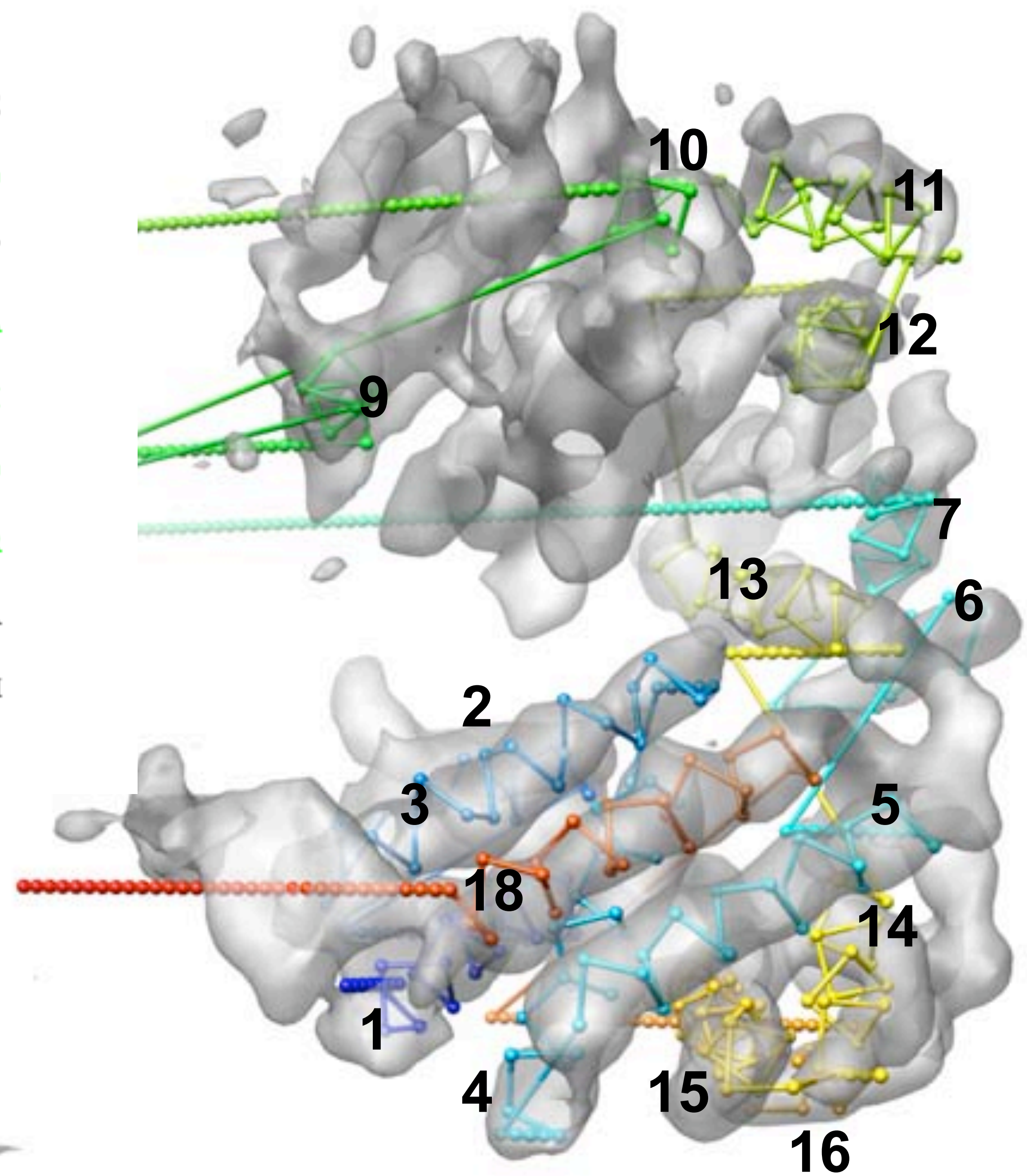


Put it all together

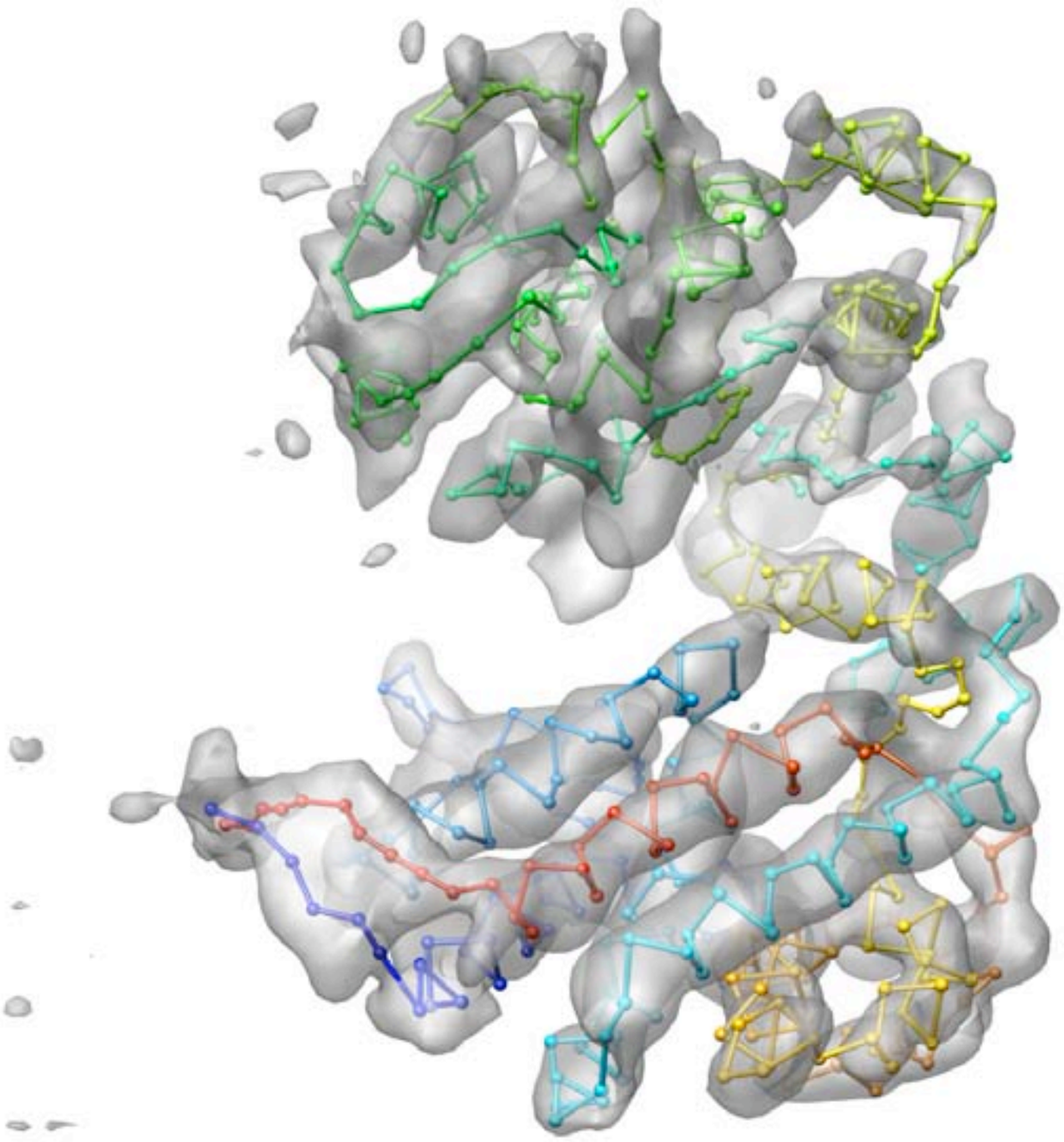
- SSEHunter
  - lengths of helices
- Skeleton
  - Pathlengths between helix endpoints, possible connectivity
- Sequence based Sec. Structure Pred.
  - Lengths of and maximum distances between helices. Relation to primary sequence. Errors ?
- MANY permutations,
  - eg. > 10^{16} possible helix assignments

Build the Model

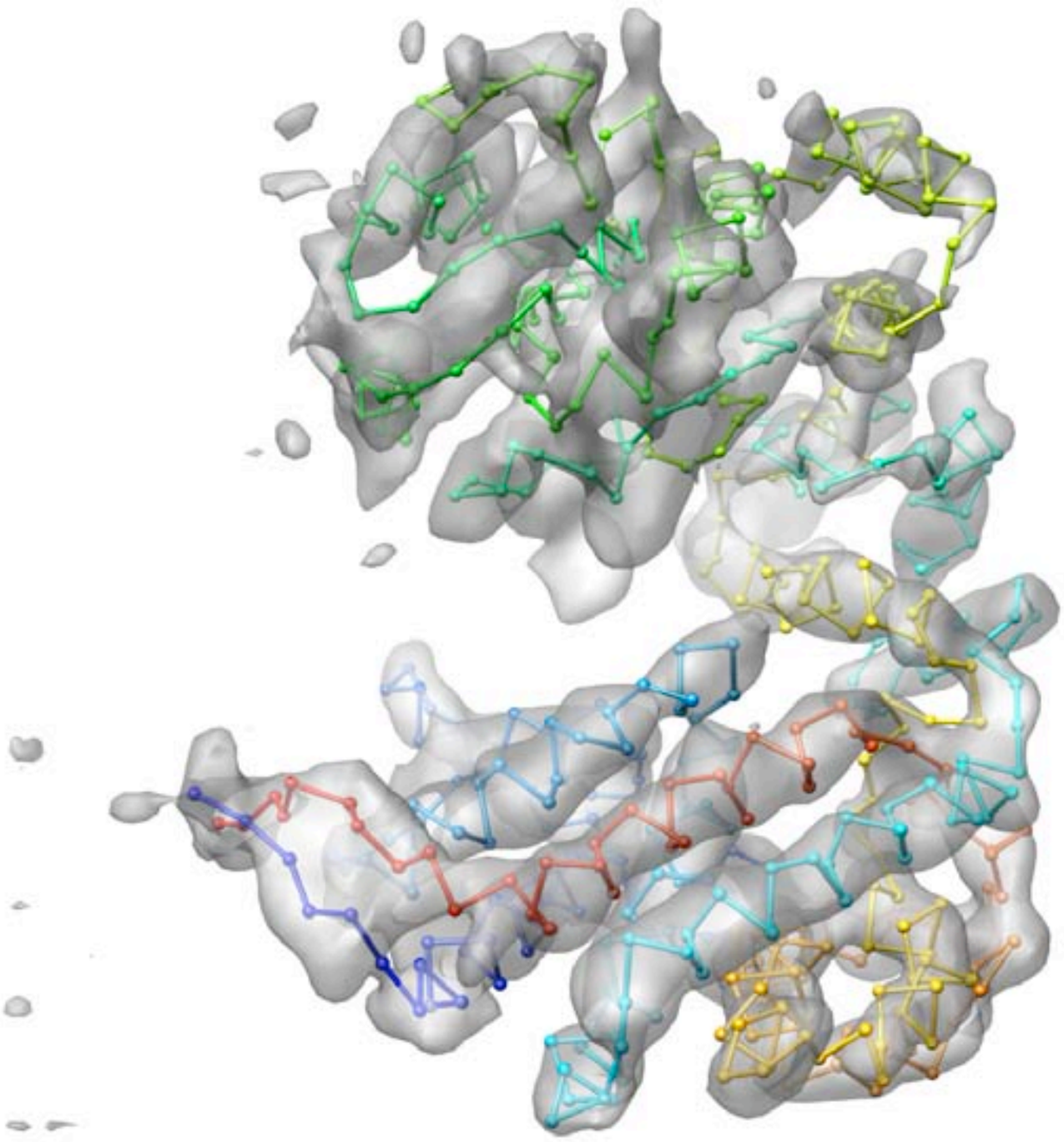
```
                1                                                      2
1      AAKDVKFGNDAGVKMLRGVNVLADAVKVTLGPKGRNVVLDKSFGAPTITKDGVSVAREIE
            3                     4                            5
61     LEDKFENMGAQMVKEVASKANDAAGDGTTTATVLAQAIITEGLKAVAAGMNPMDLKRGID
                           6             7
121    KAVTVAVEELKALSVPCSDSKAIAQVGTISANSDETVGKLIAEAMDKVGKEGVITVEDGT
                                                                8
181    GLQDELDVVEGMQFDRGYLSPYFINKPETGAVELESPFILLADKKISNIREMLPVLEAVA
                   9                                       10
241    KAGKPLLIIAEDVEGEALATAVVNTIRGIVKVAAVKAPGFGDRRKAMLQDIATLTGGTVI
                                                           11
301    SEEIGMELEKATLEDLGQAKRVVINKDTTTIIDGVGEEAAIQGRVAQIRQQIEEATSDYD
          12                         13                        14
361    REKLQERVAKLAGGVAVIKVGAATEVEMKEKKARVEDALHATRAAVEEGVVAGGGVALIR
                             15            16
421    VASKLADLRGQNEDQNVGIKVALRAMEAPLRQIVLNCGEEPSVVANTVKGGDGNYGYNAA
          17                    18
481    TEEYGNMIDMGILDPTKVTRSALQYAASVAGLMITTECMVTDLPKNDAADLGAAGGMGGM

541    GGMGGMM
```



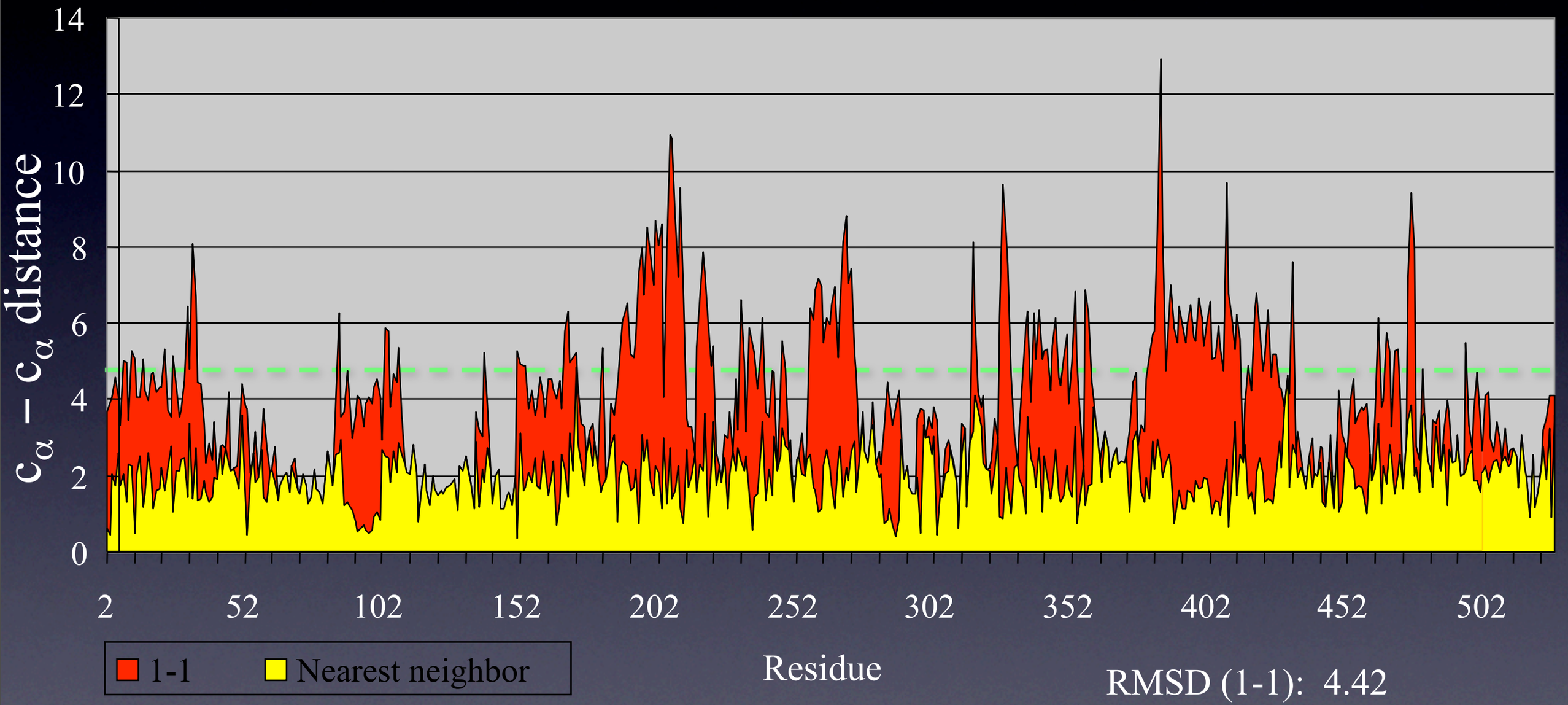
GroEL Chain Trace

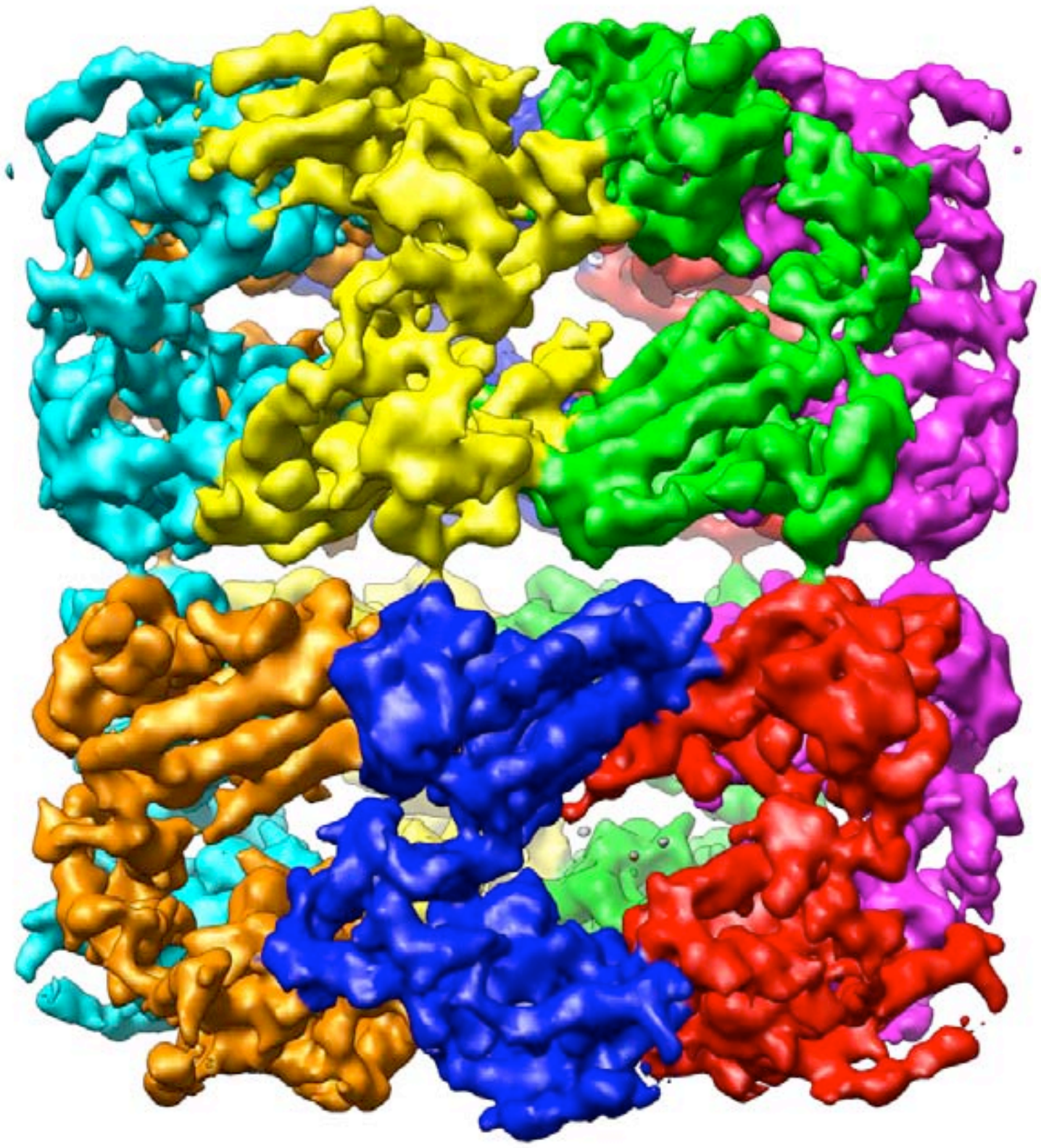


GroEL Chain Trace

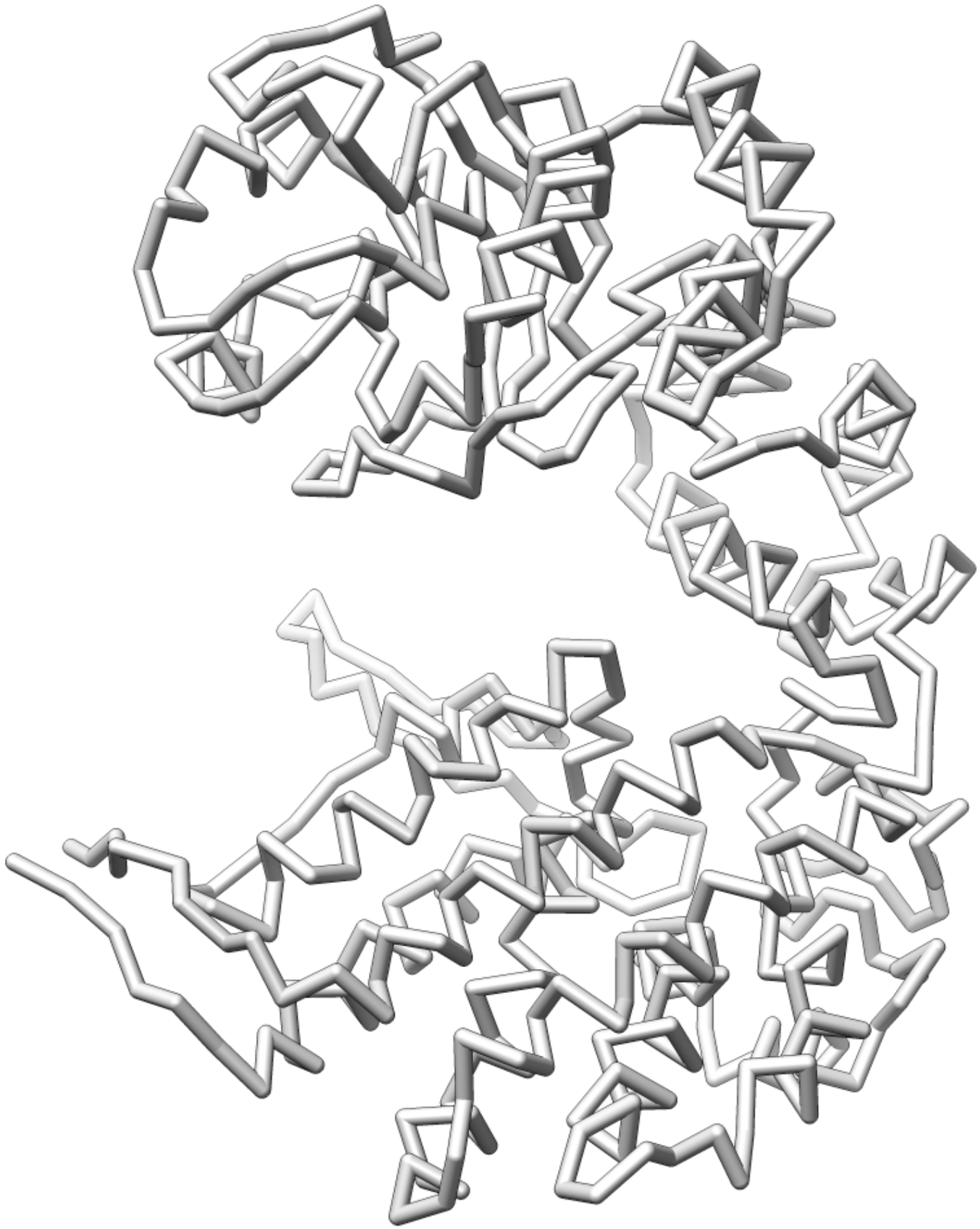


RMSD vs. PDB Structure



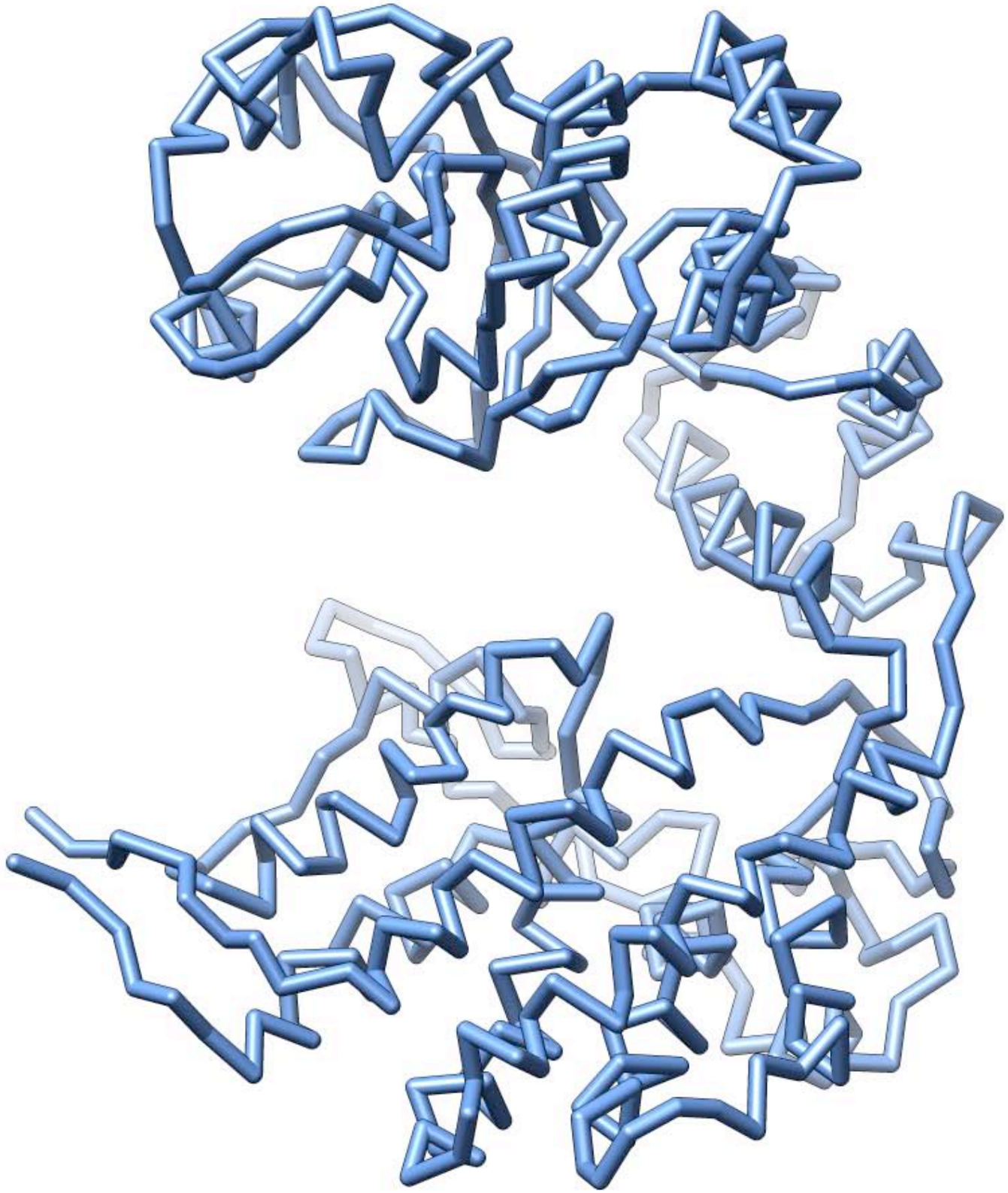


D7



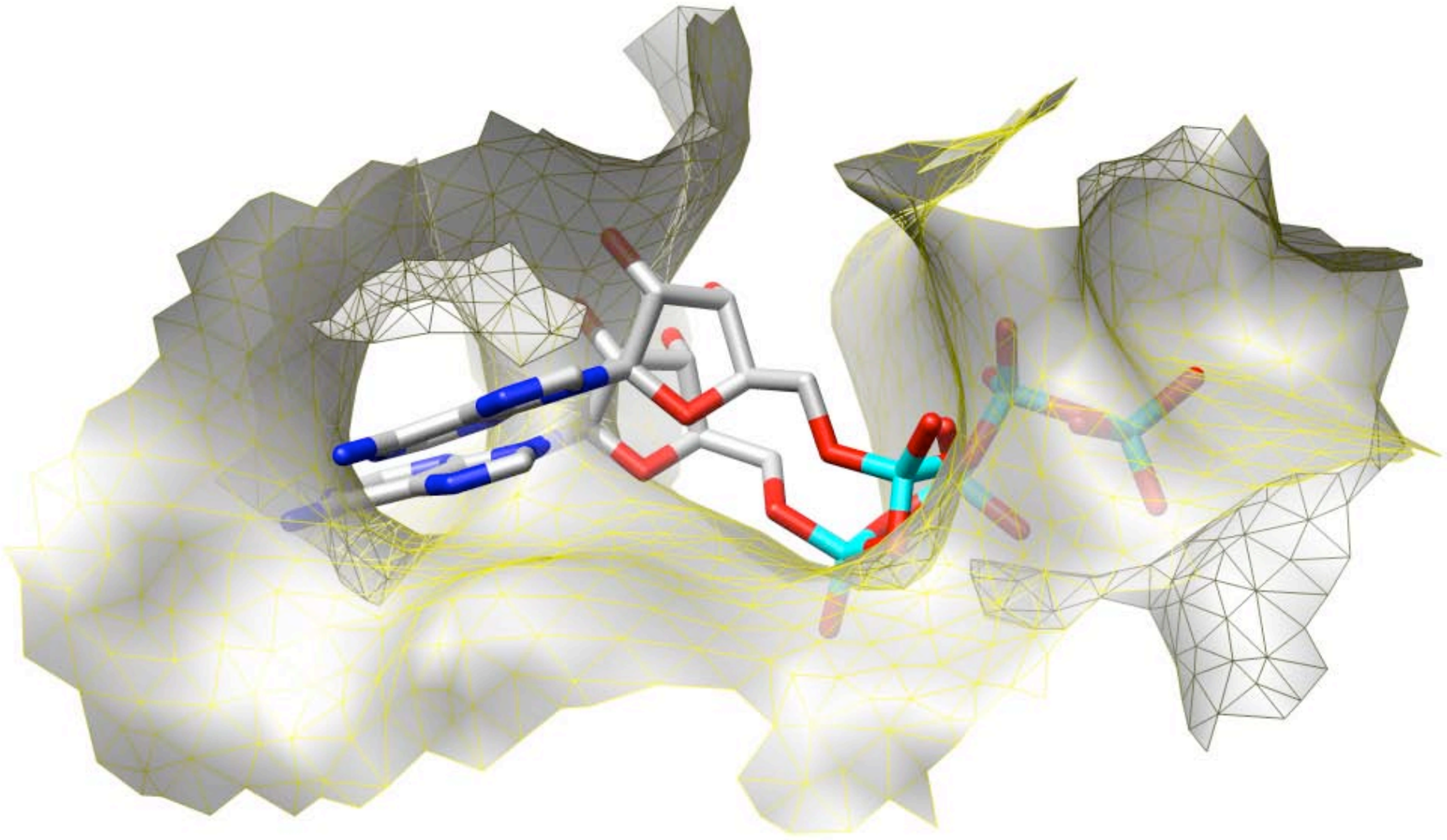
Bottom Ring C7

Top Ring C7



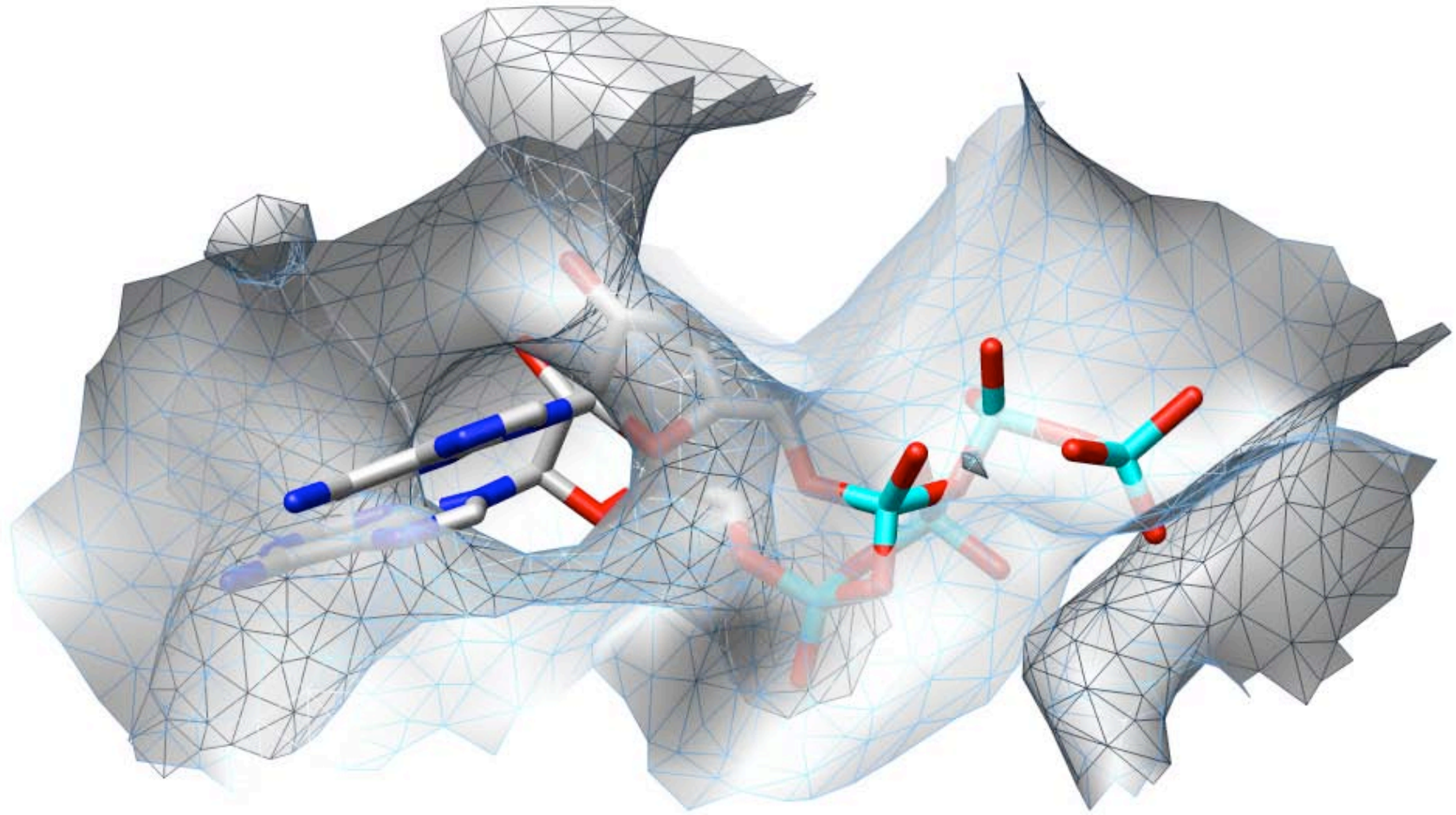
Bottom Ring

(ATP Binding Site)



Top Ring

(ATP Binding Site)



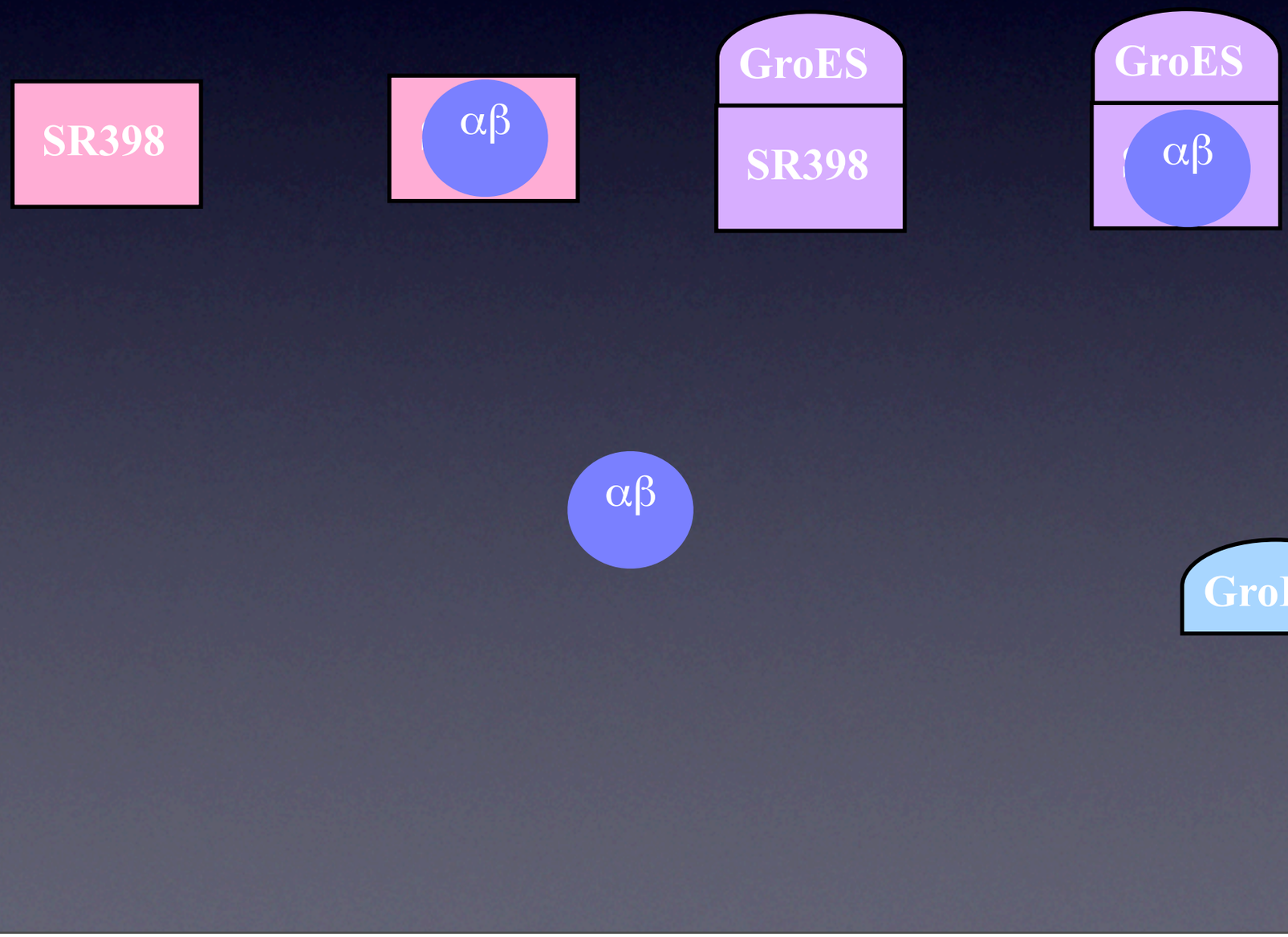
GroEL results

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

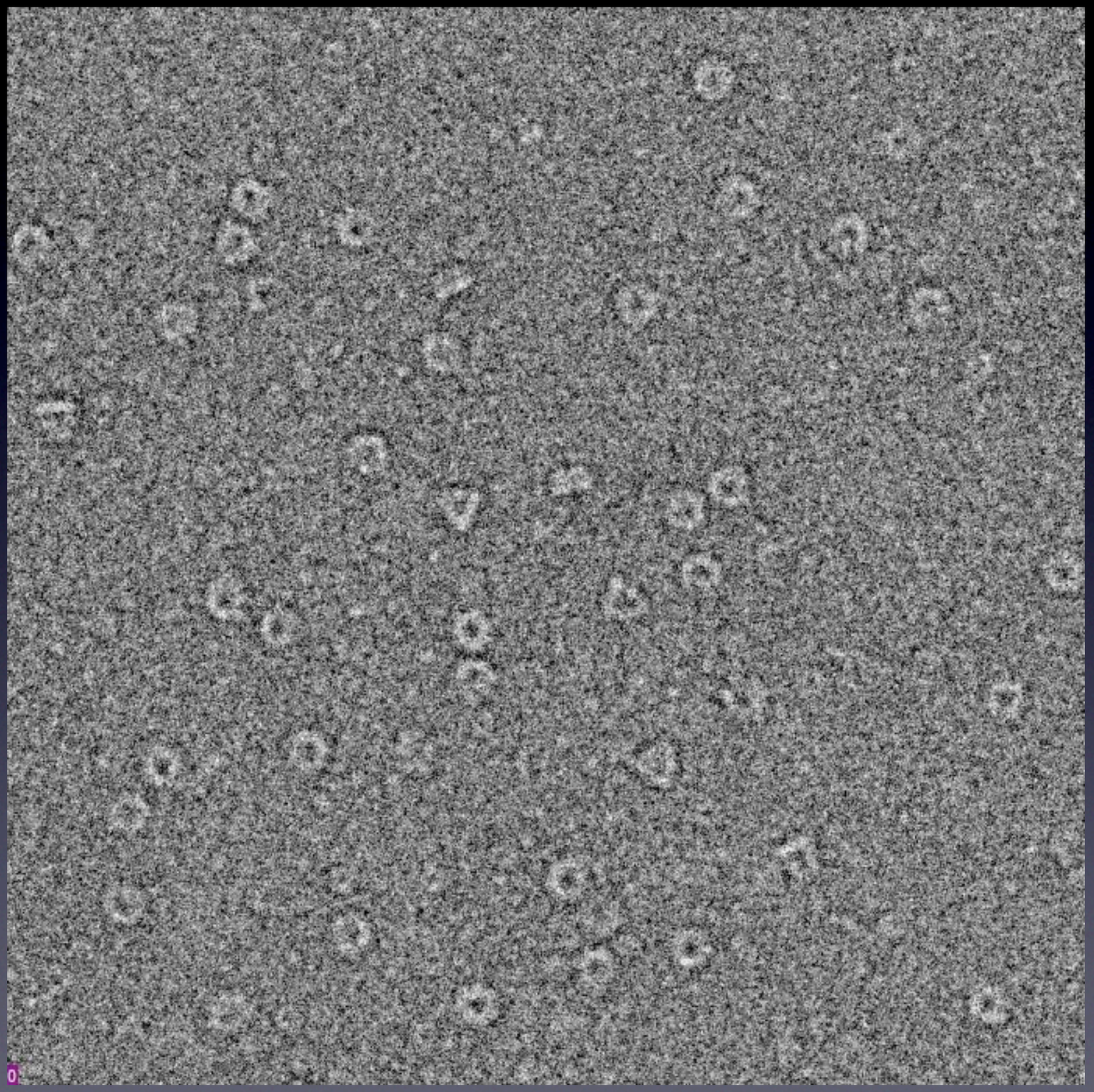
Multi-component Systems

GroEL, GroES & Substrate

Produces a Mixed Population

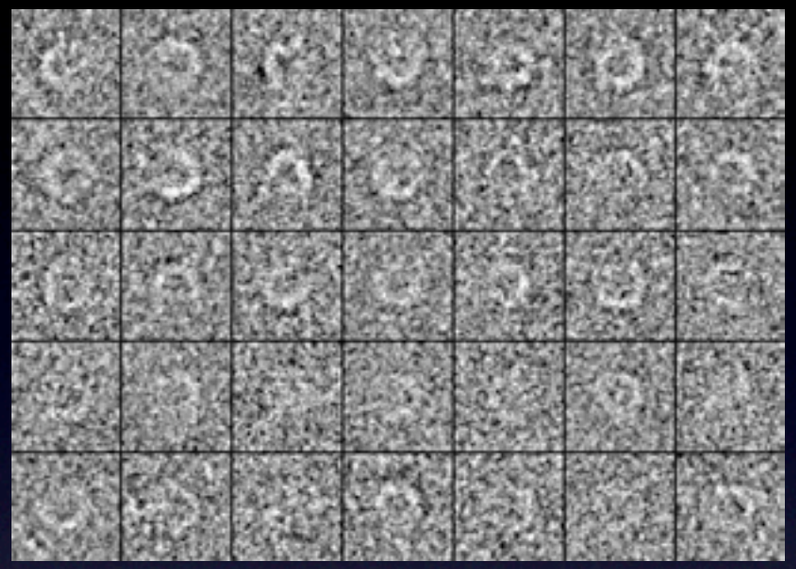


SR398+GroES



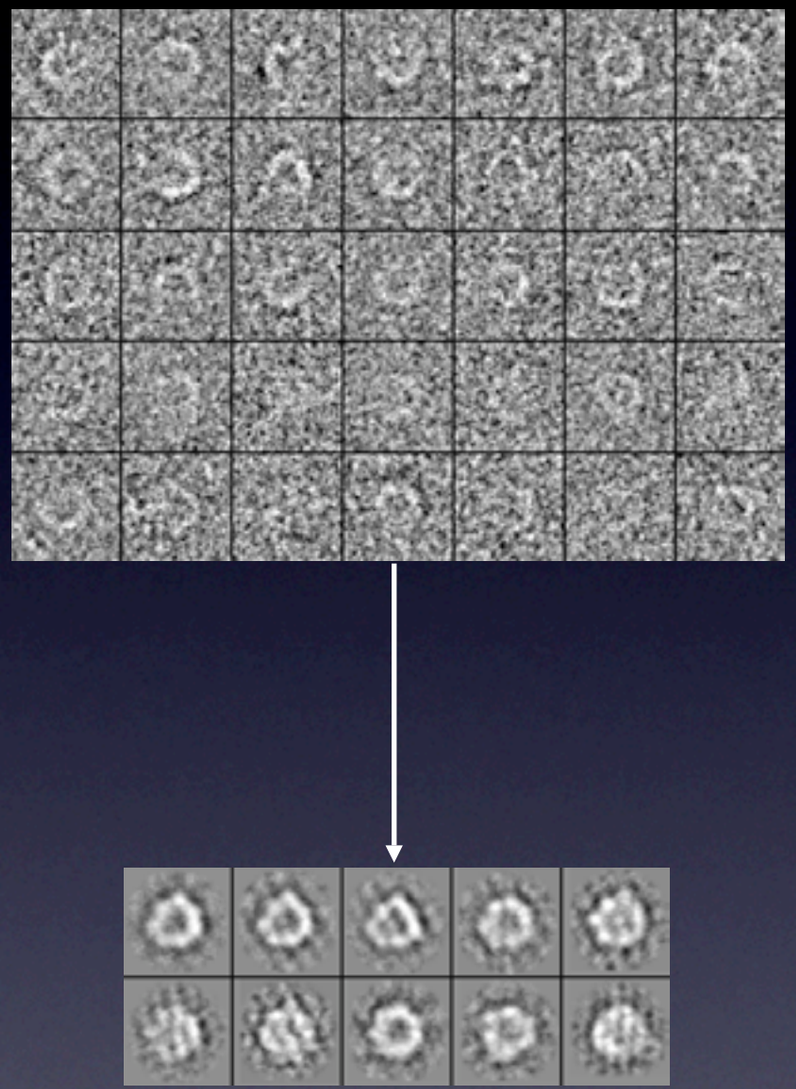
2D Refinement

Particles



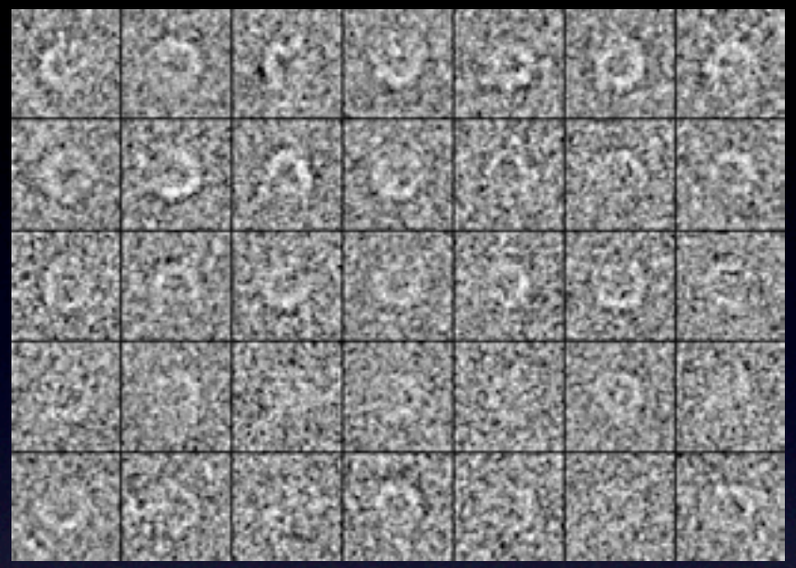
2D Refinement

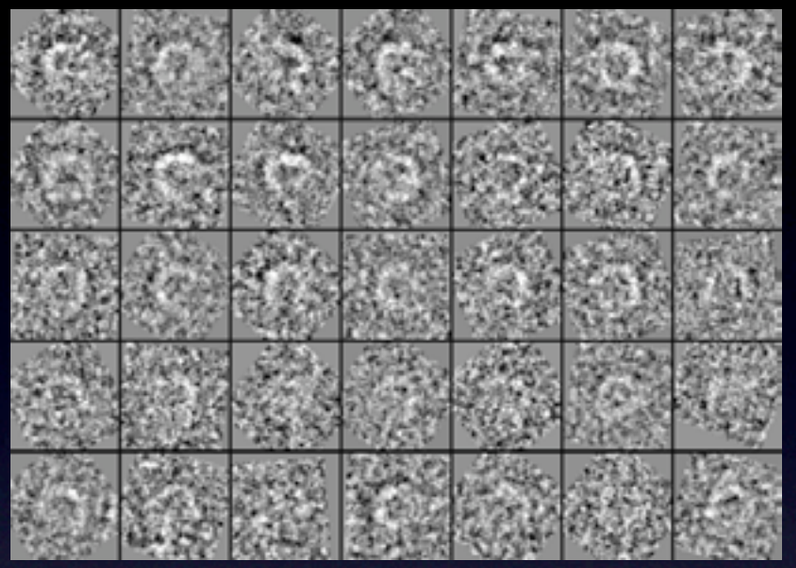
Particles



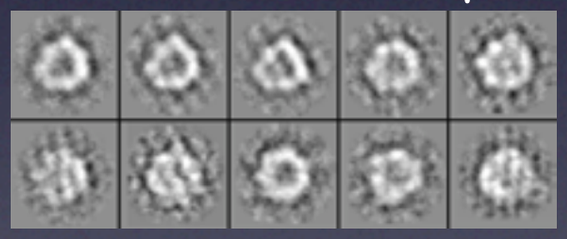
2D Refinement

Particles



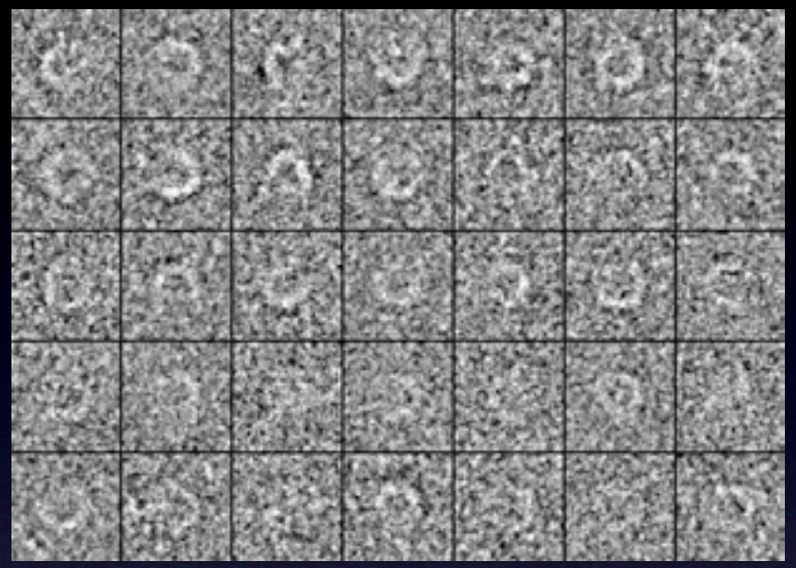


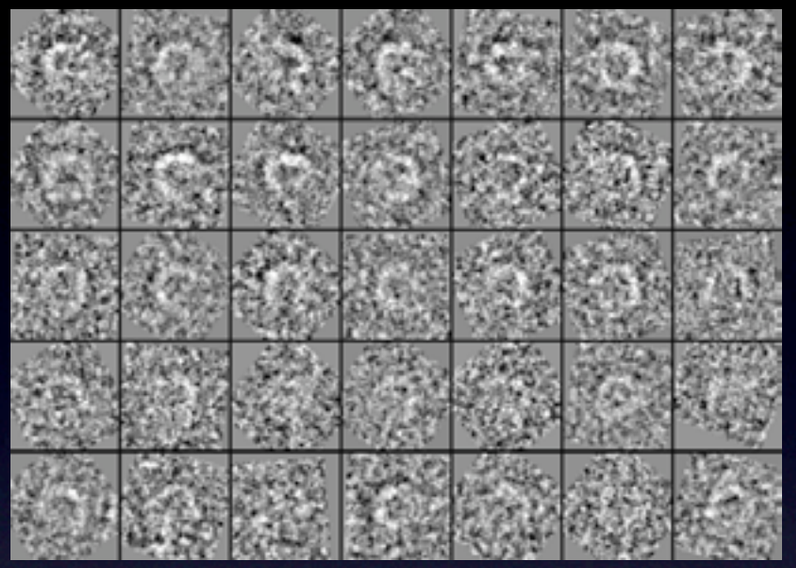
Align Particles



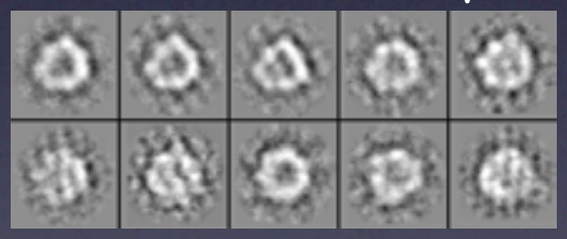
2D Refinement

Particles

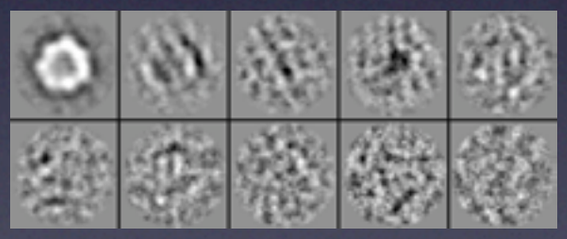




Align Particles

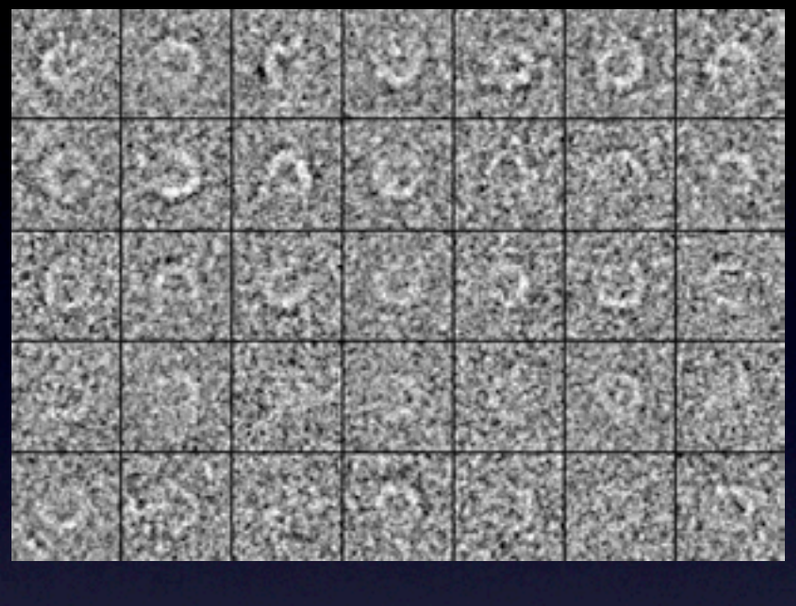


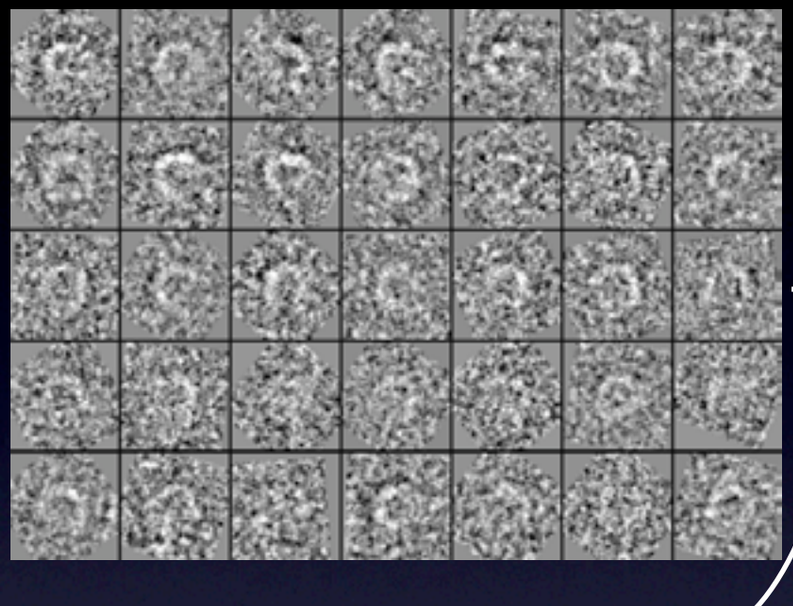
SVD/PCA

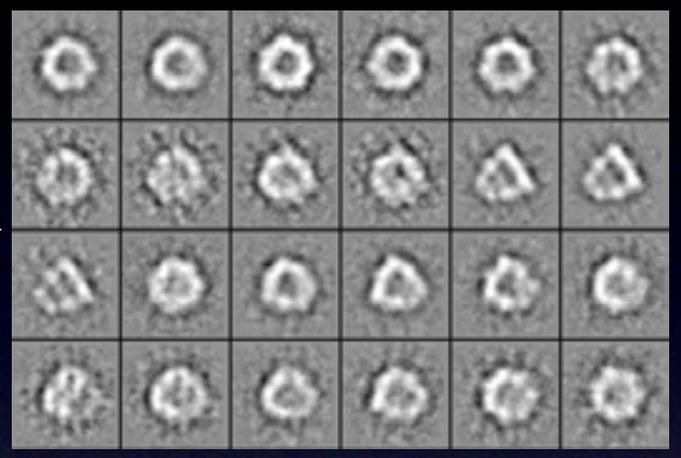


2D Refinement

Particles

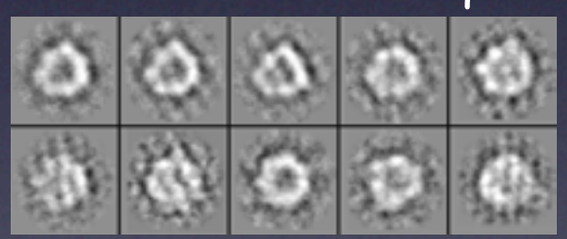




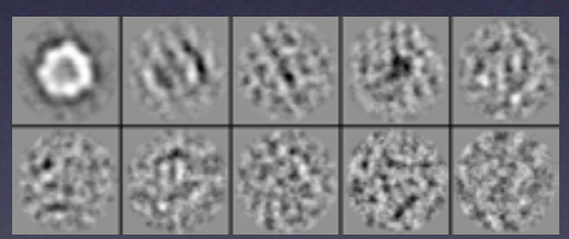


Align Particles

k-means classification in SVD subspace

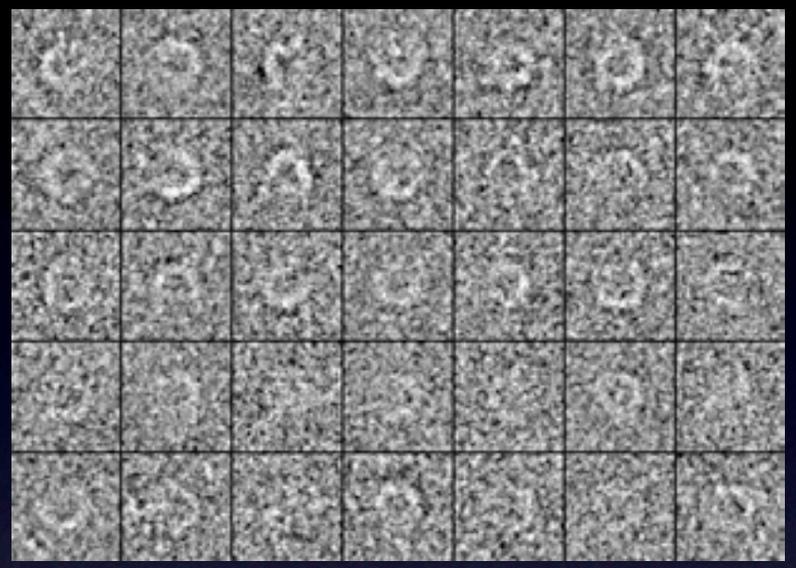


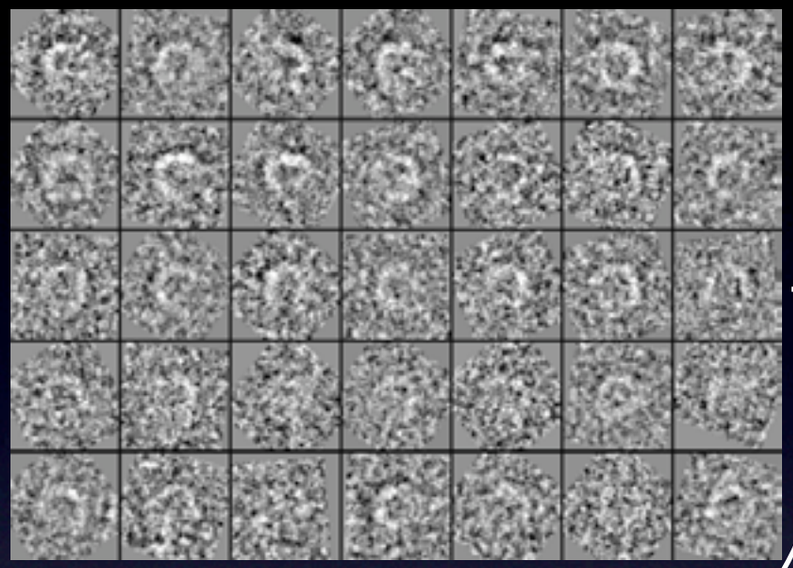
SVD/PCA

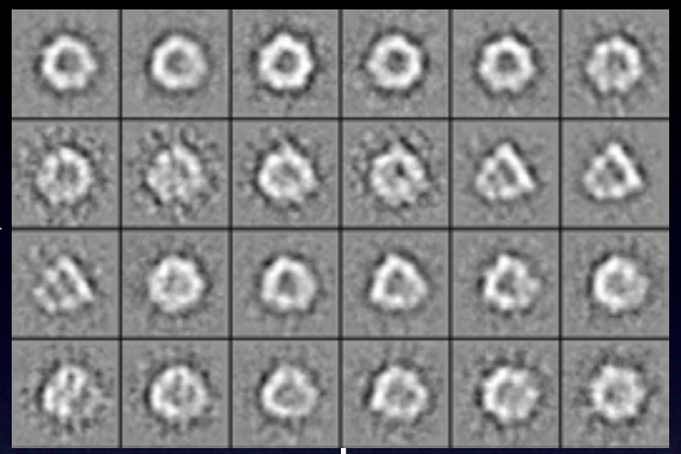


2D Refinement

Particles

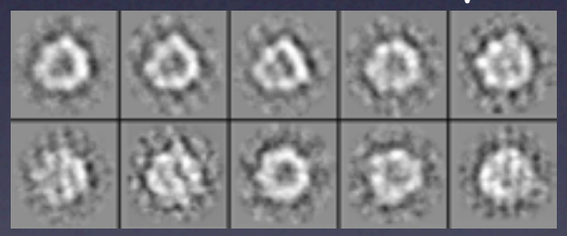




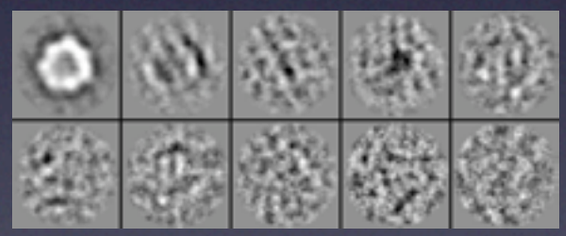


k-means classification in SVD subspace

Align Particles

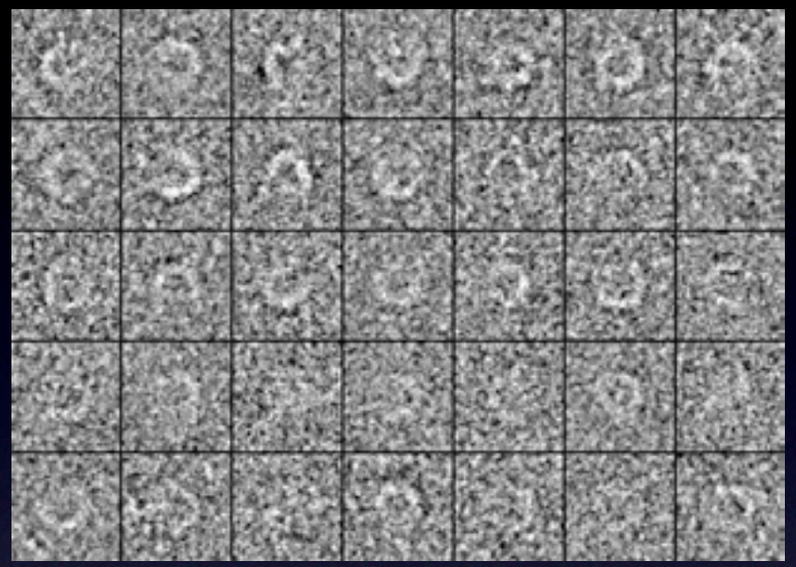


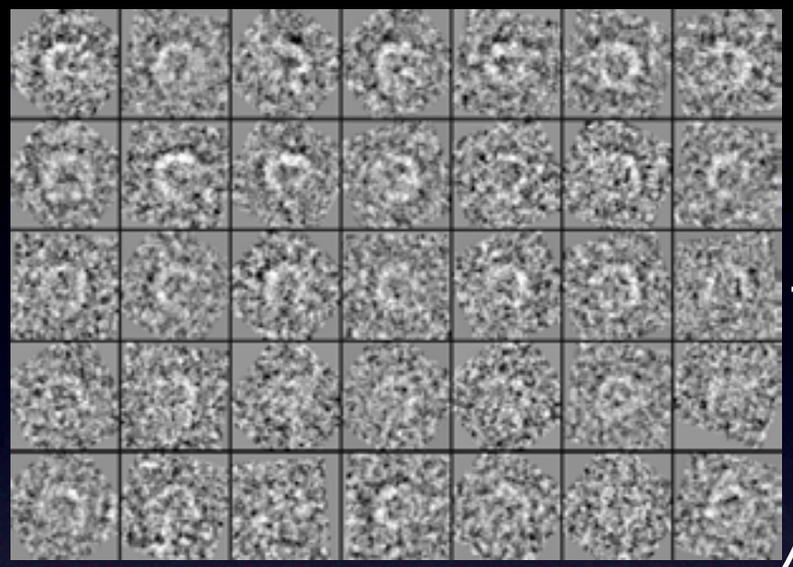
SVD/PCA

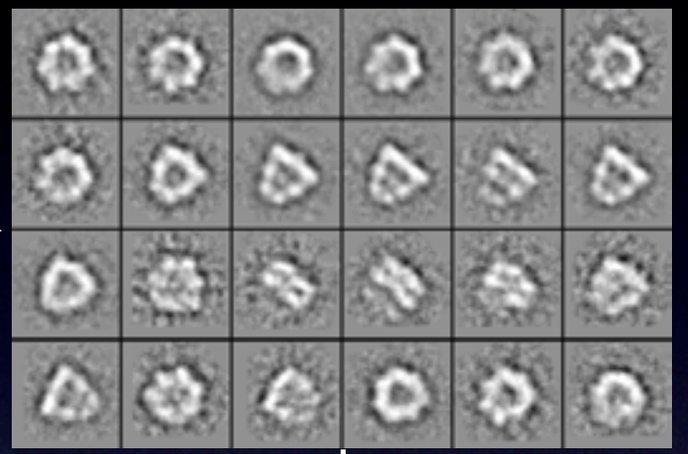


2D Refinement

Particles

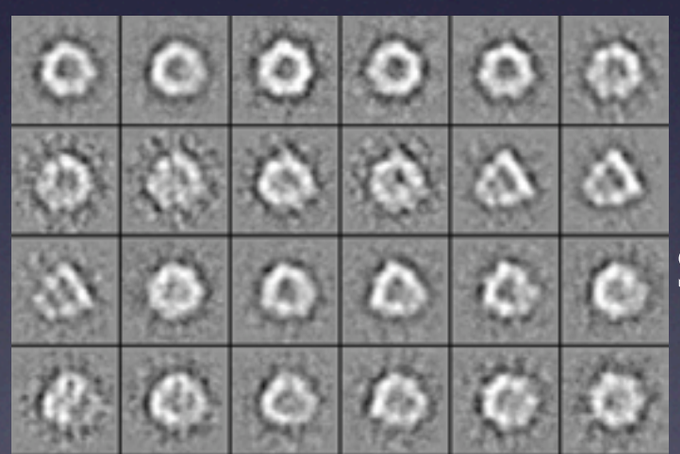




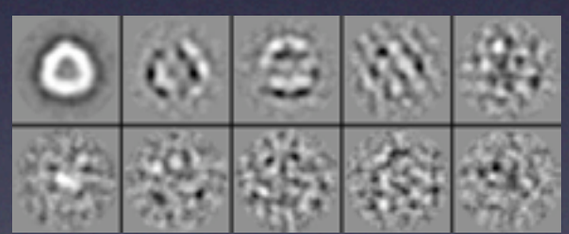


k-means classification in SVD subspace

Align Particles

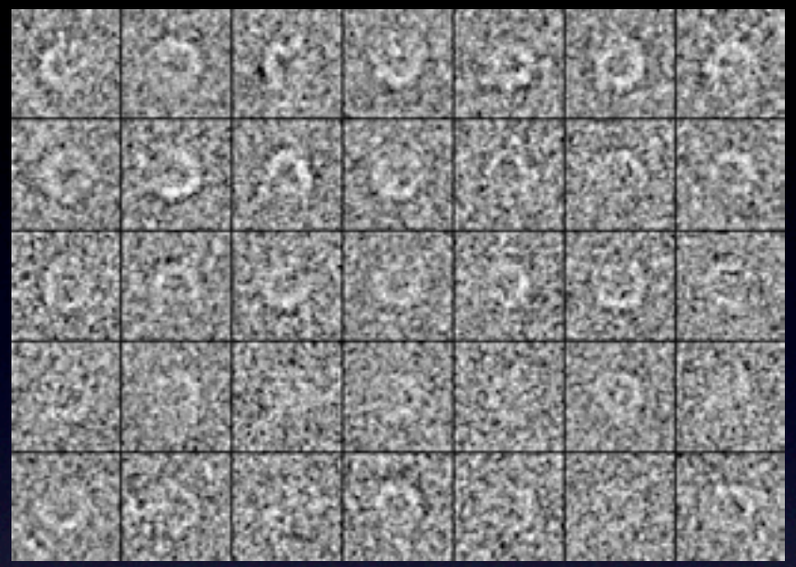


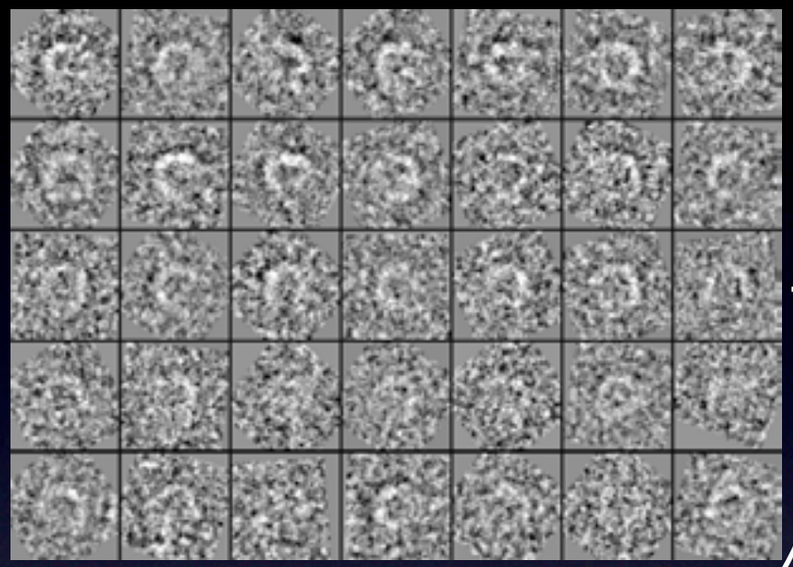
SVD/PCA

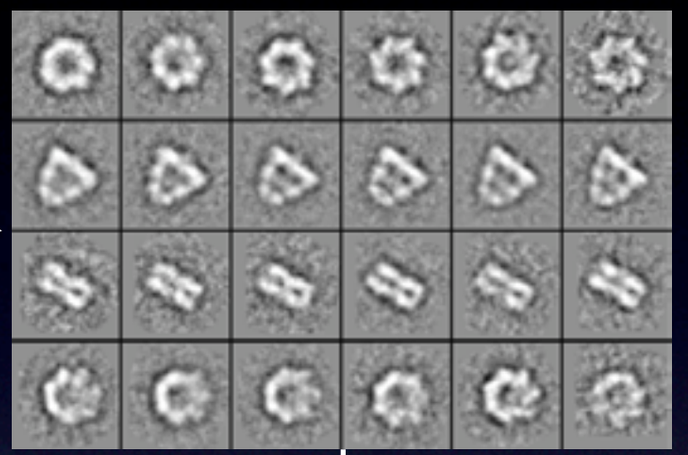


2D Refinement

Particles

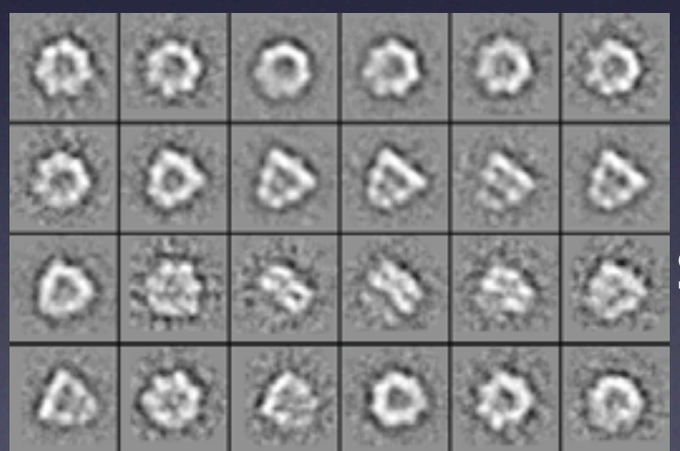




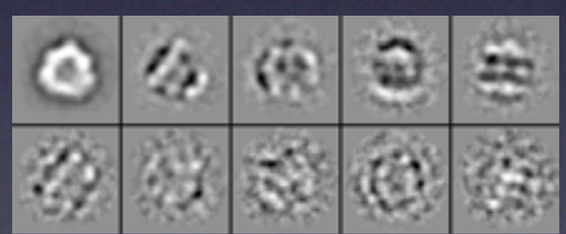


Align Particles

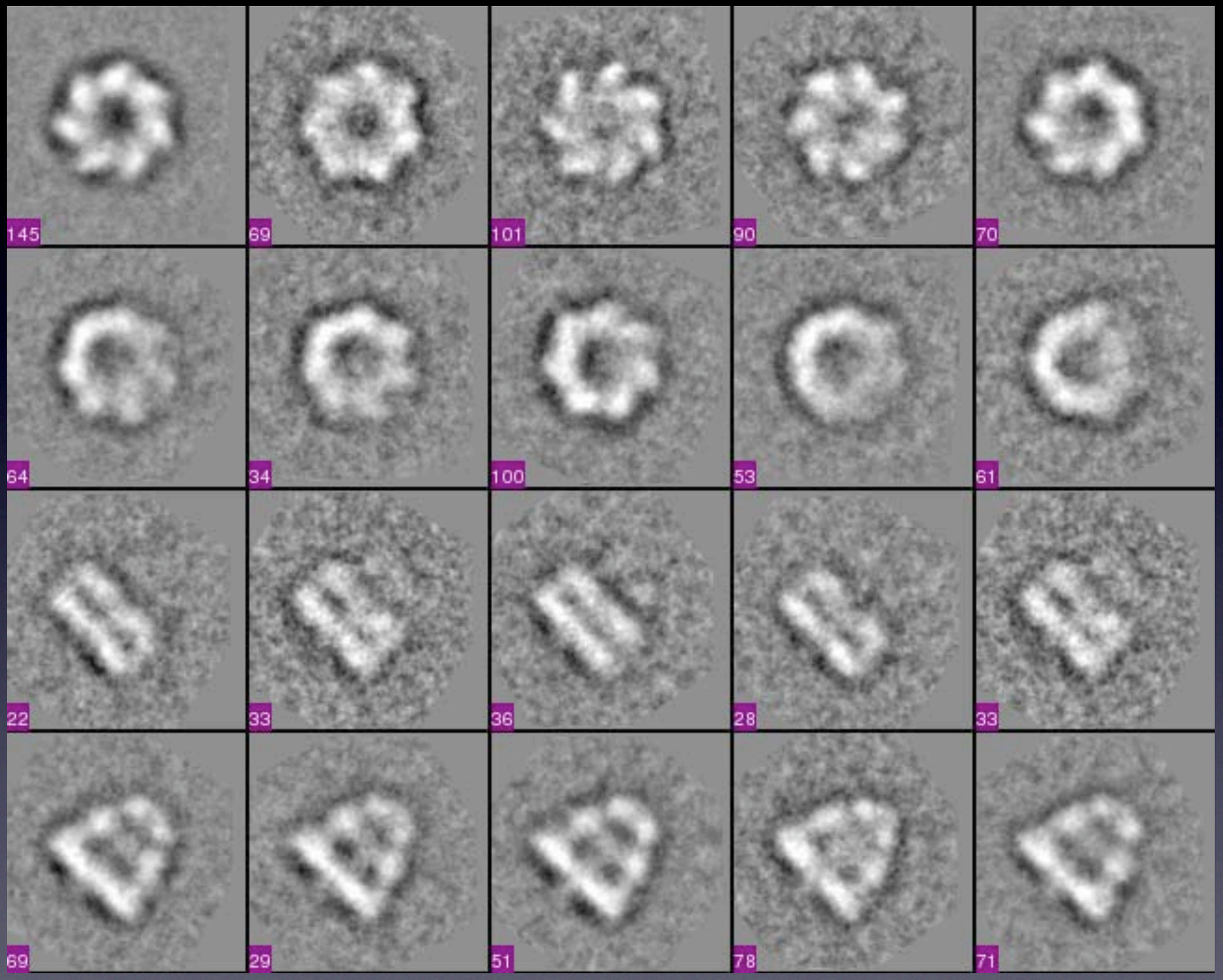
k-means classification in SVD subspace

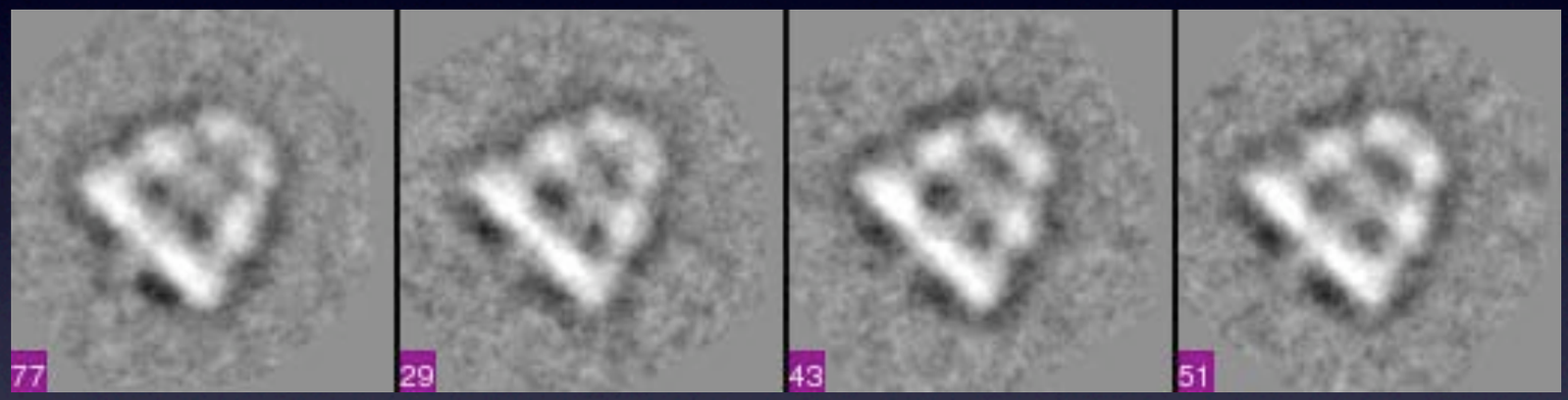


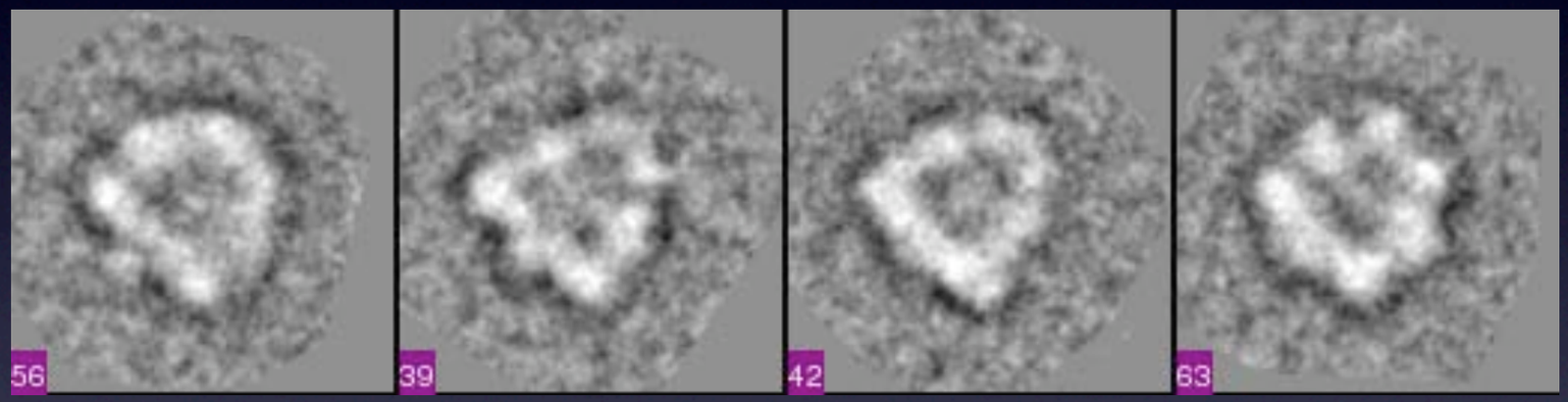
SVD/PCA



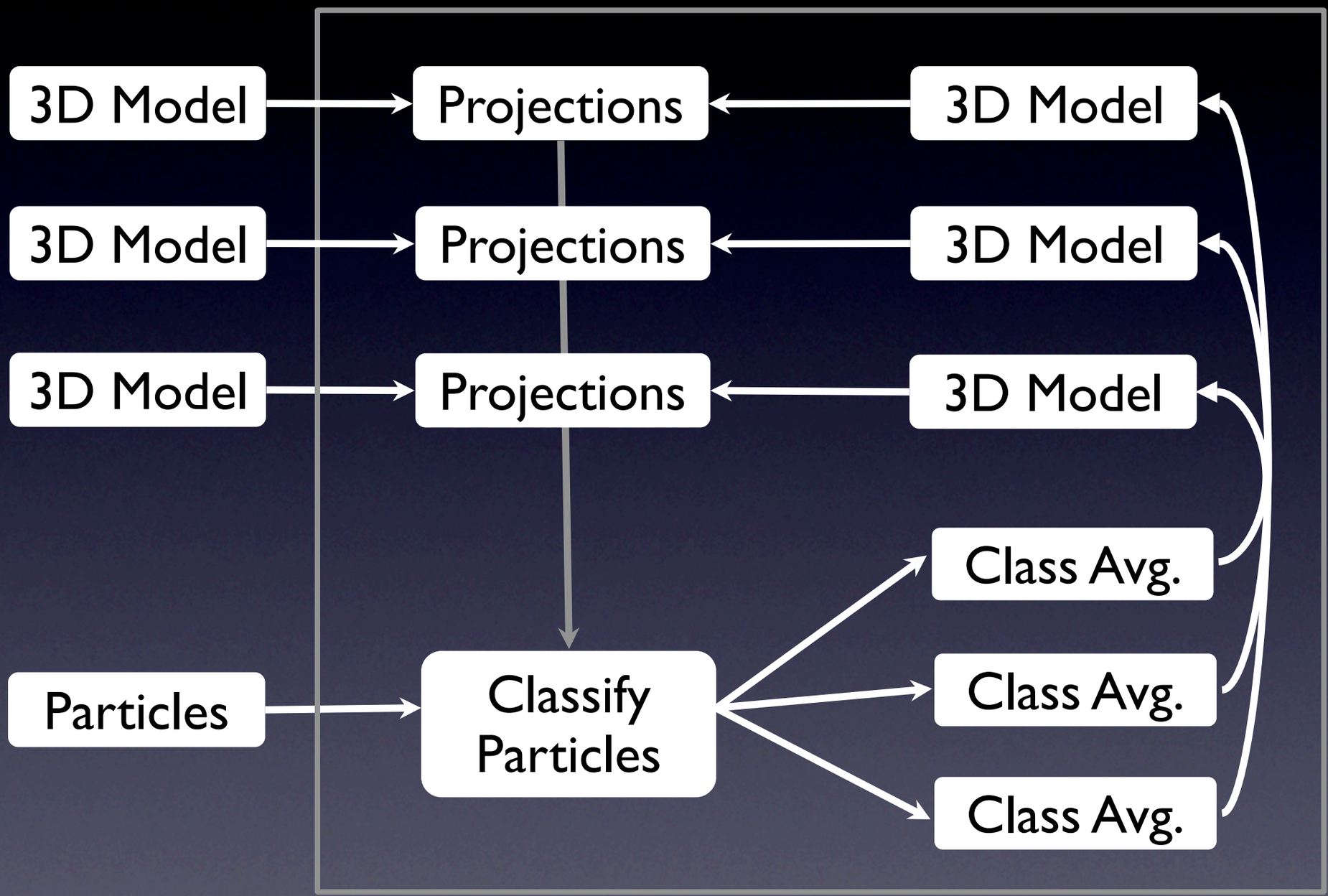
After 9 iterations


145
69
101
90
70
64
34
100
53
61
22
33
36
28
33
69
29
51
78
71


77
29
43
51


56
39
42
63

Multireference Refinement



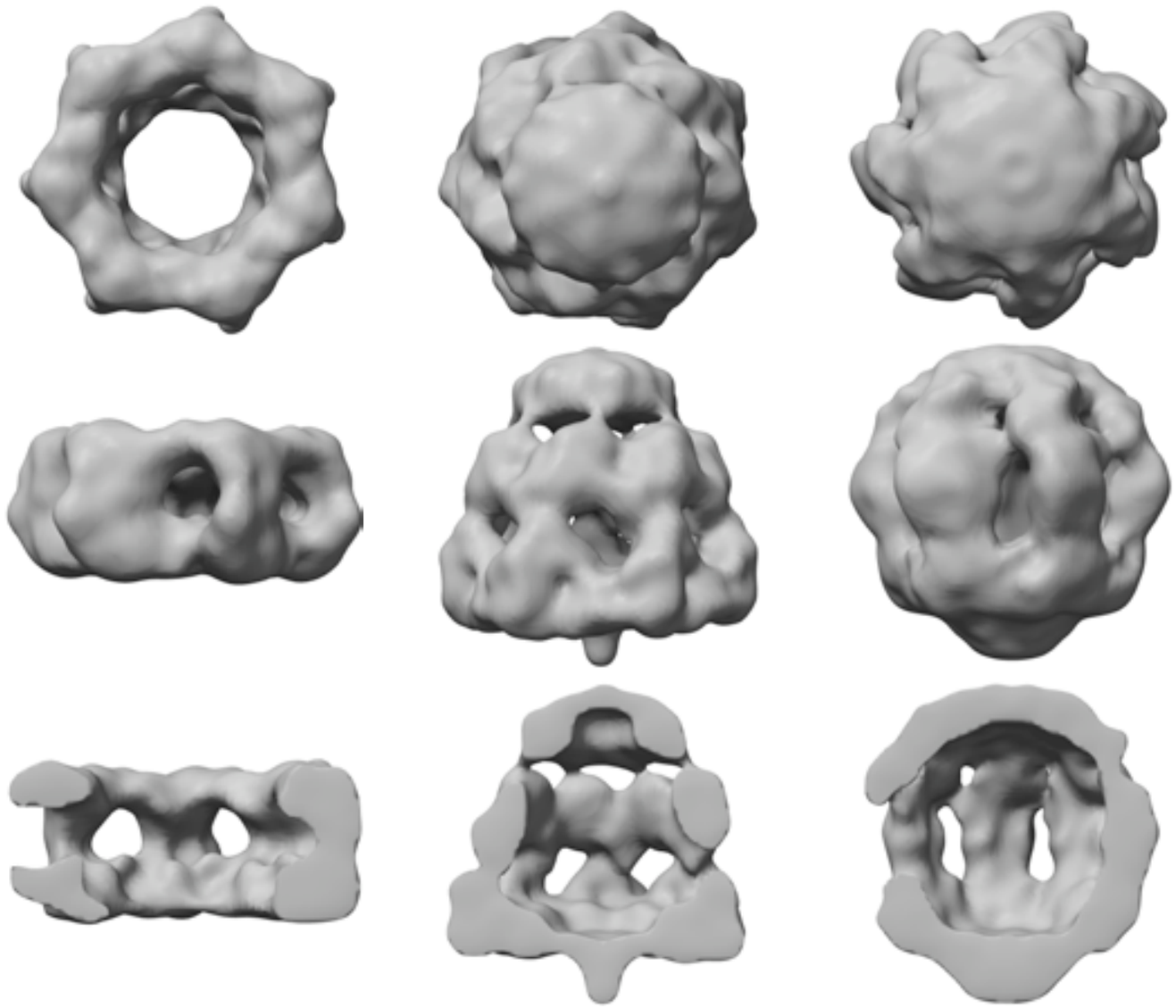
SR398+GroES+Mg-ATP

Top View

Side View

Side View of Half Map

Central Slice of Side View



SR398 heptamer with no GroES

Standard Conformation

Expanded Conformation

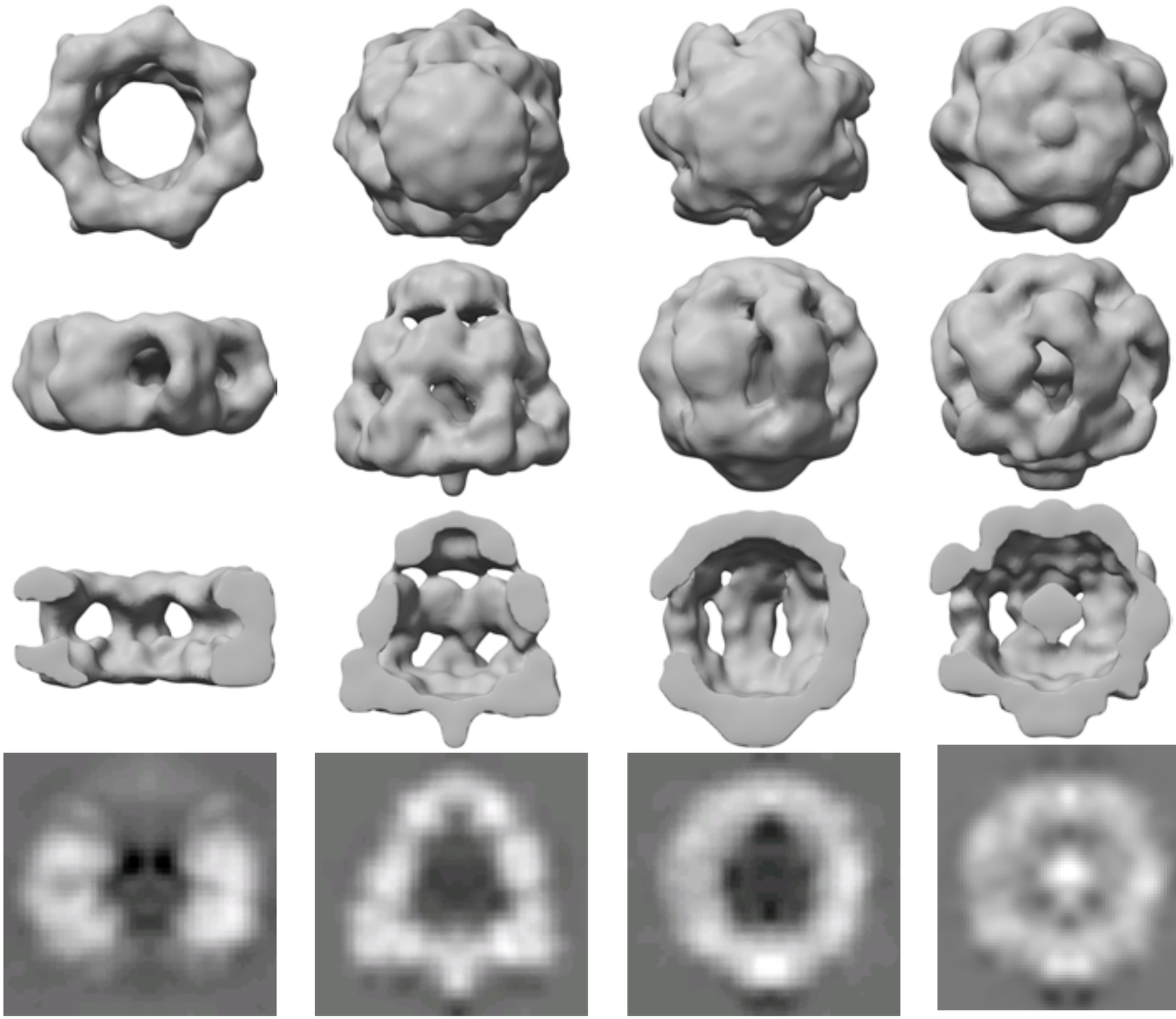
SR398+GroES+Mg-ATP

Top View

Side View

Side View of Half Map

Central Slice of Side View



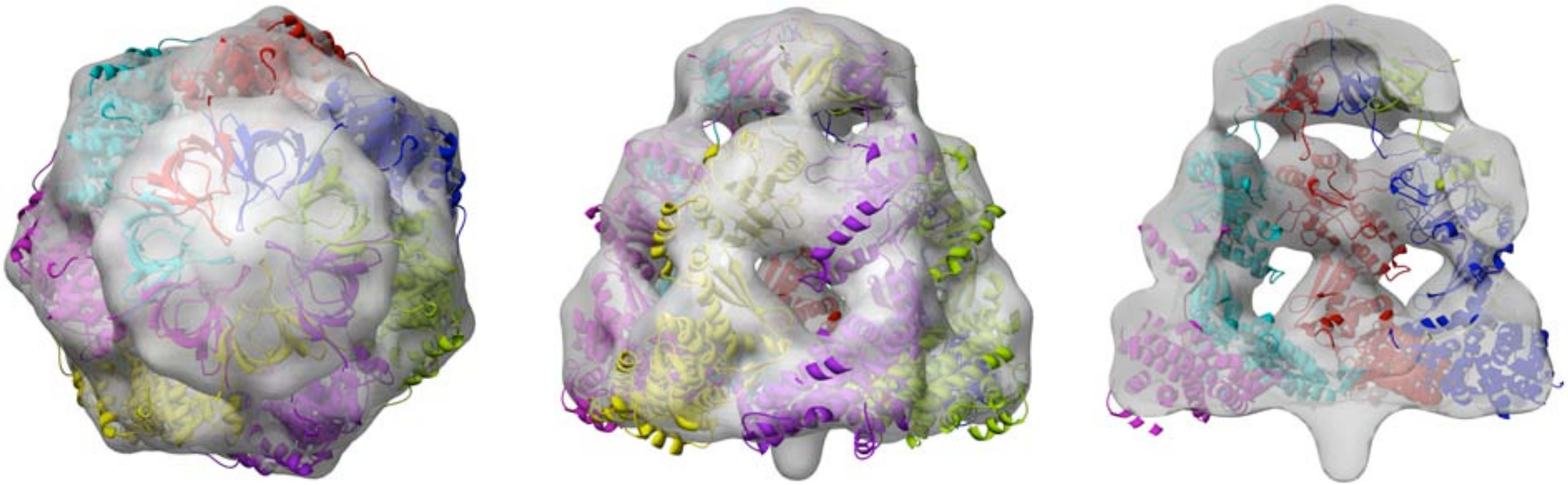
SR398 heptamer with no GroES

Standard Conformation

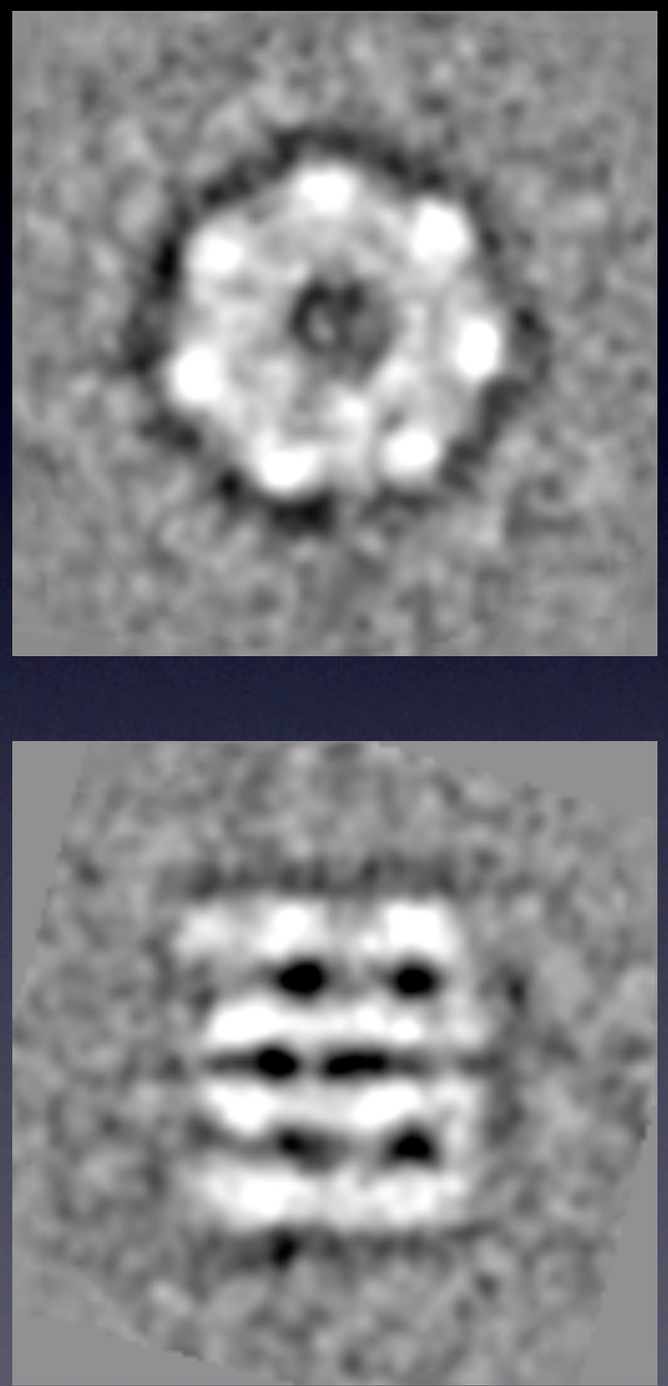
Expanded Conformation

With Substrate

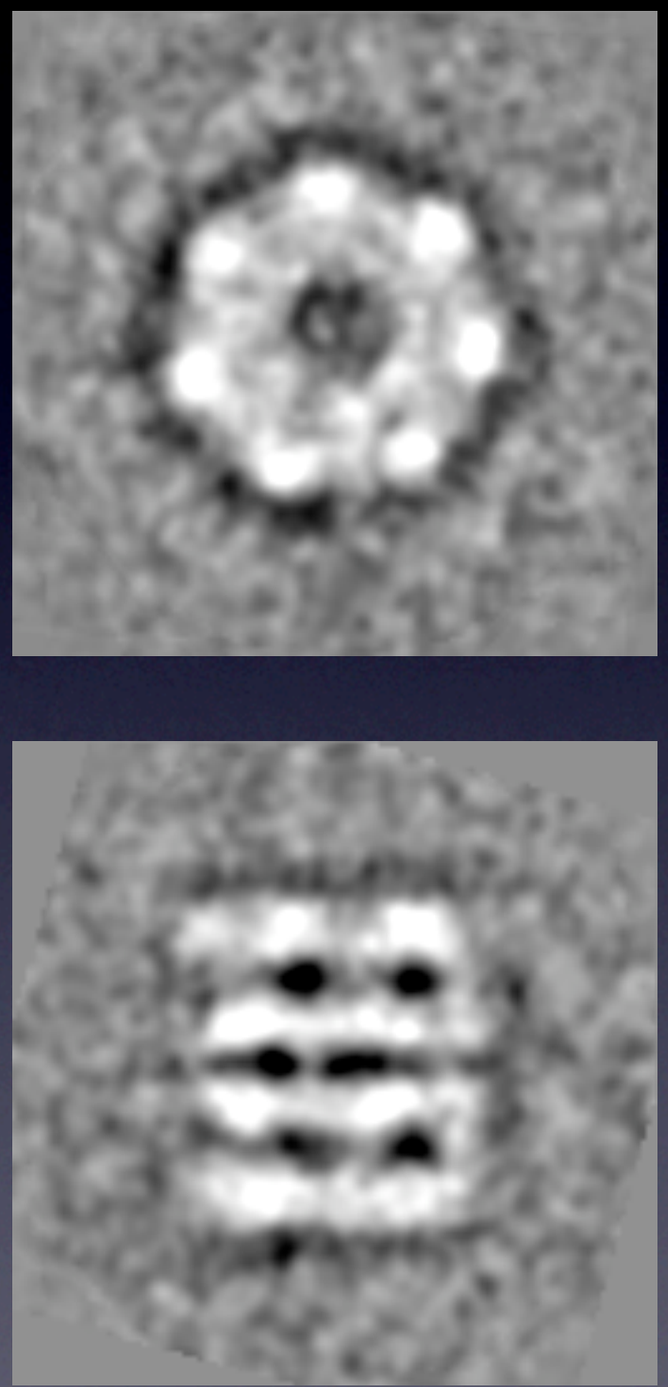
X-ray structure docked into cryo-EM density map



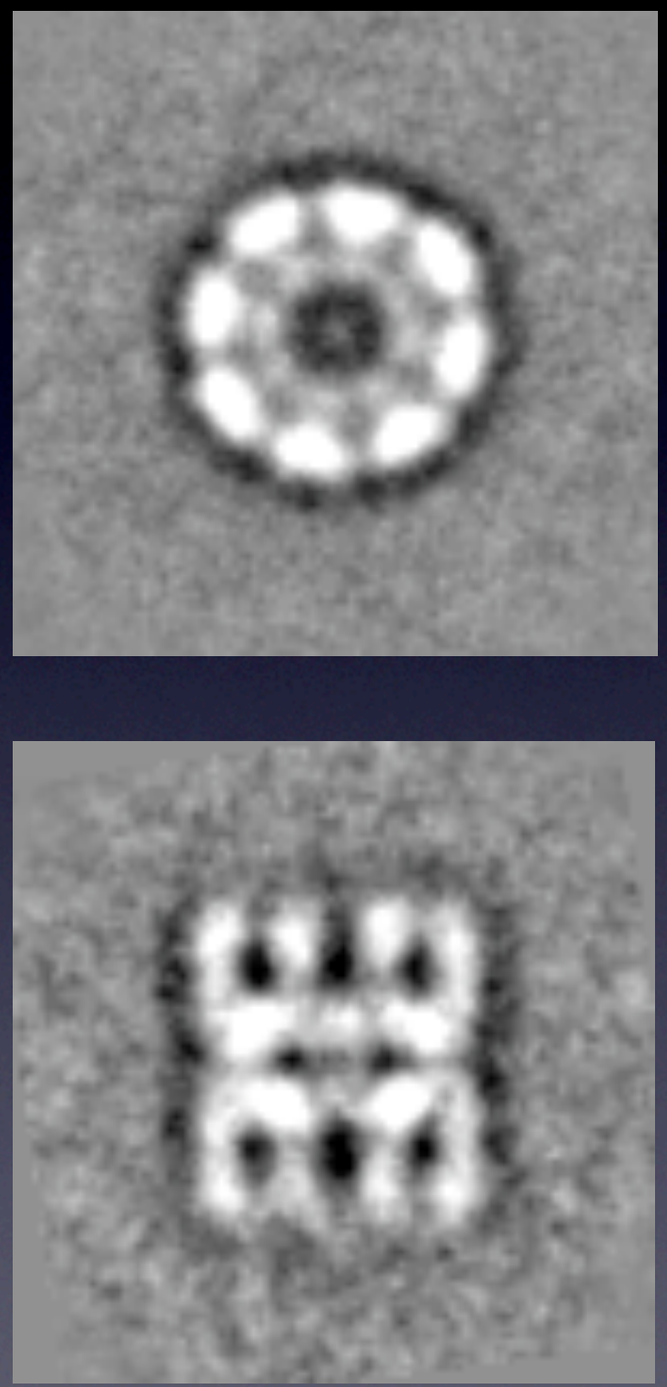
GroEL



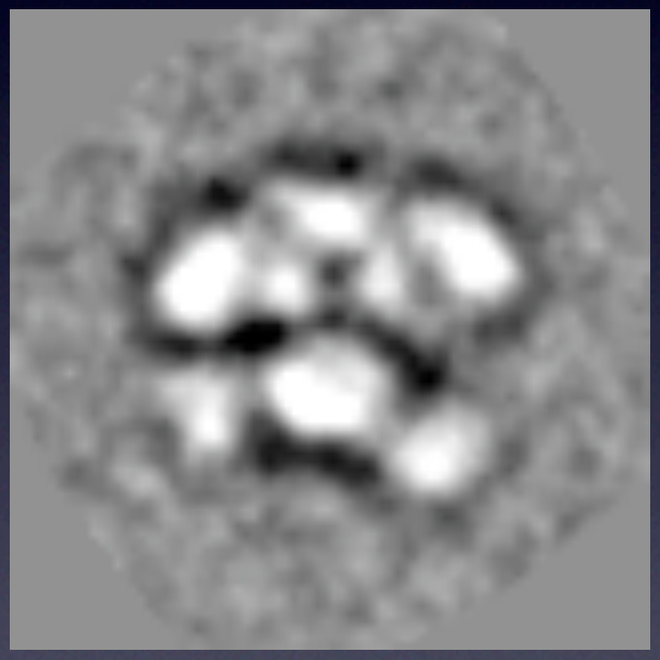
GroEL



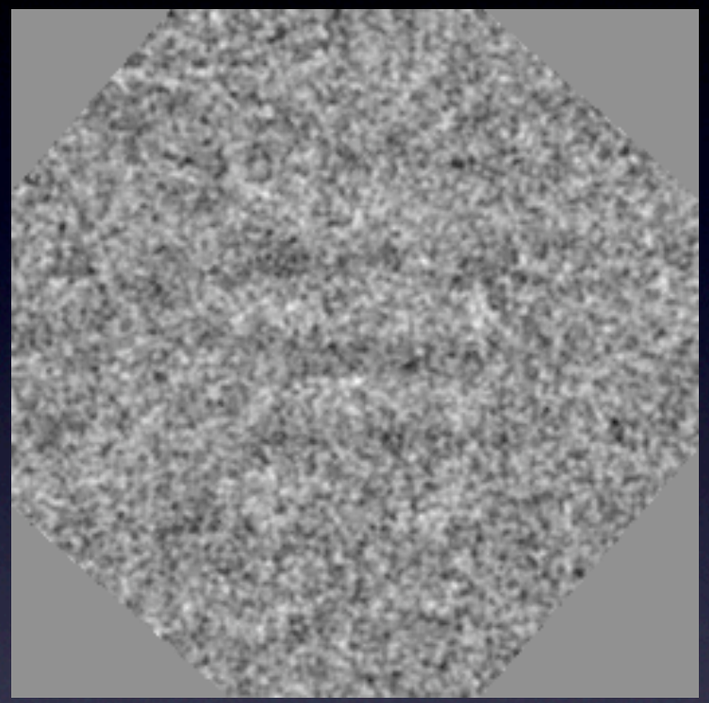
mmcpn



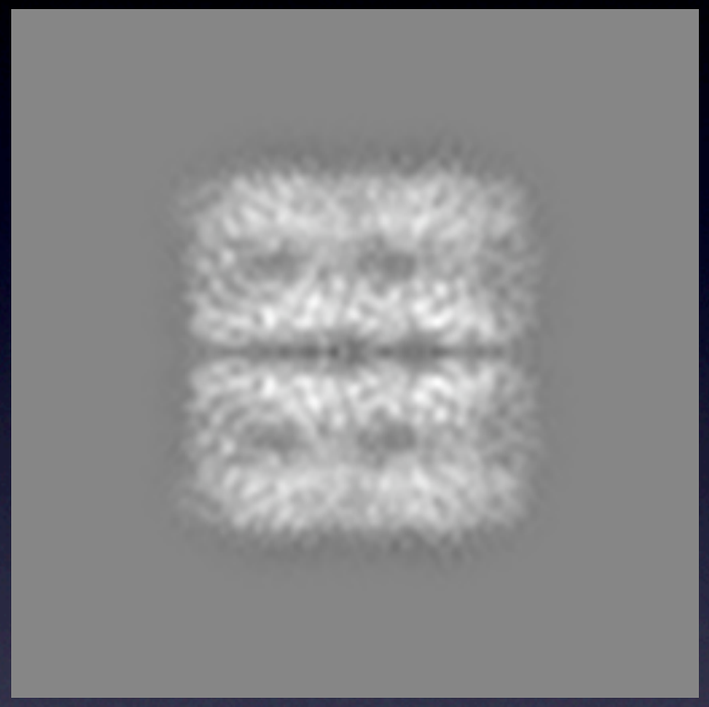
Fatty Acid Synthase



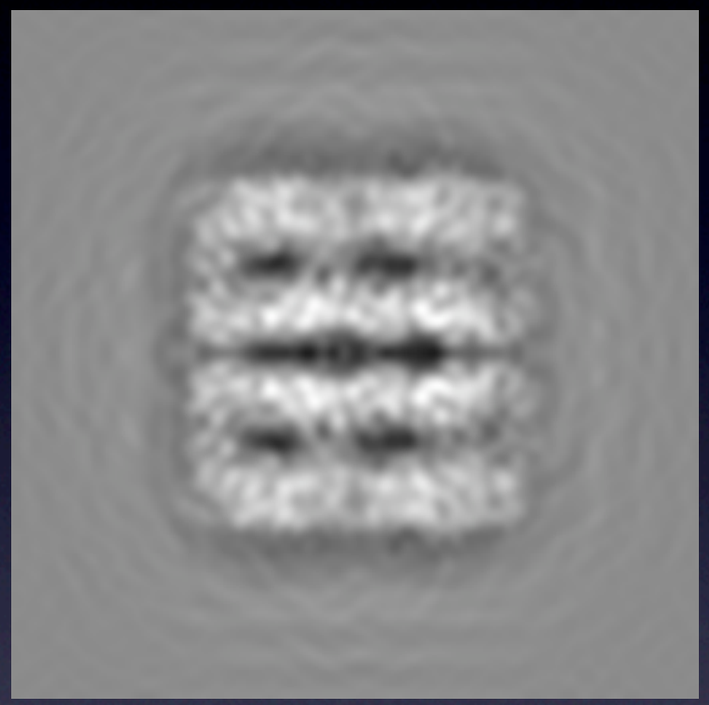
Can we subtract what we don't want ?



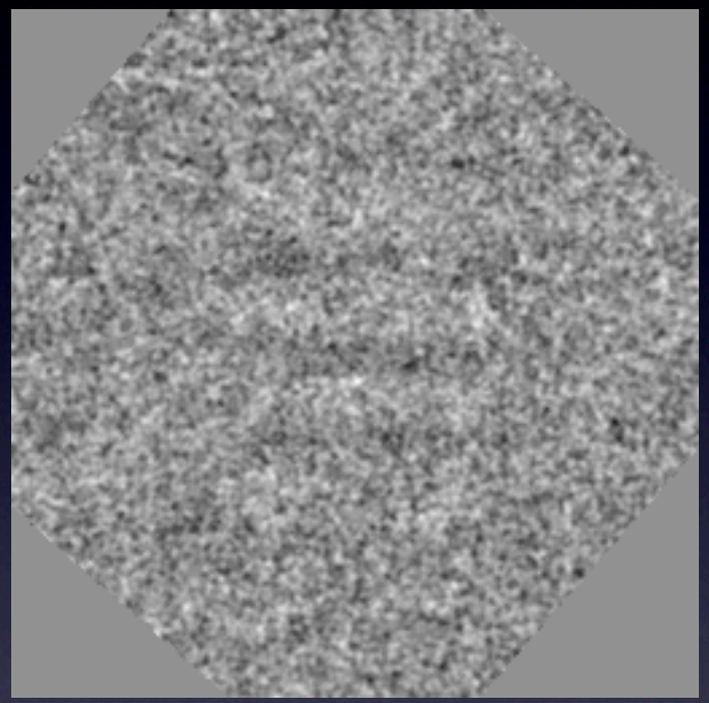
Can we subtract what we don't want ?



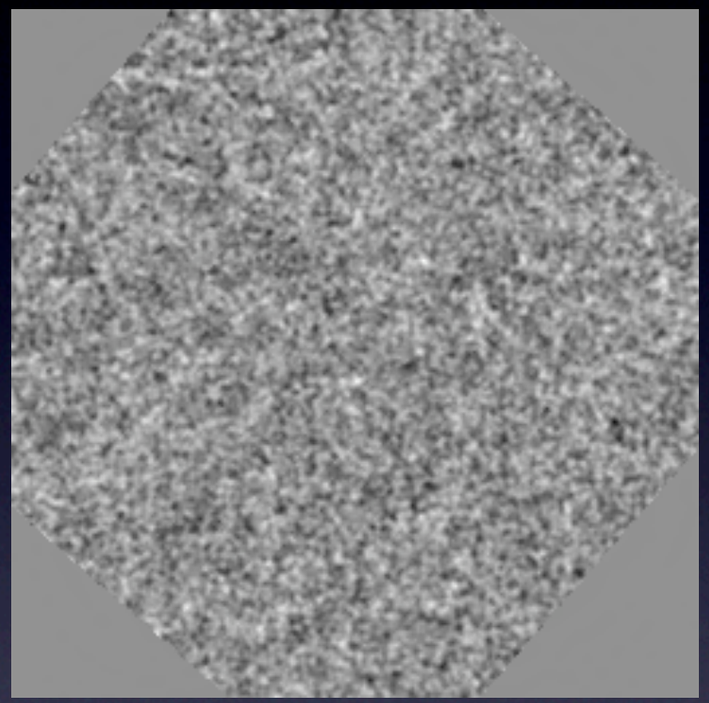
A lot of math →



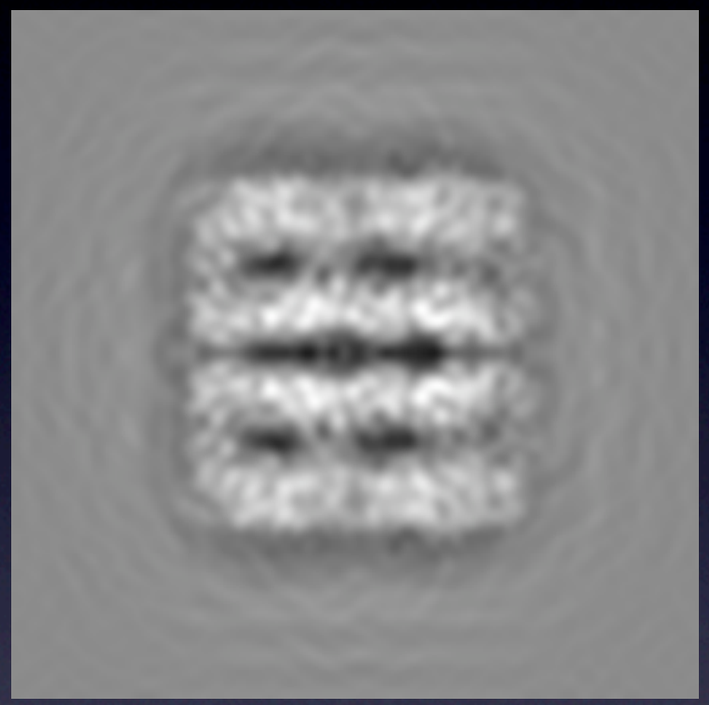
Can we subtract what we don't want ?



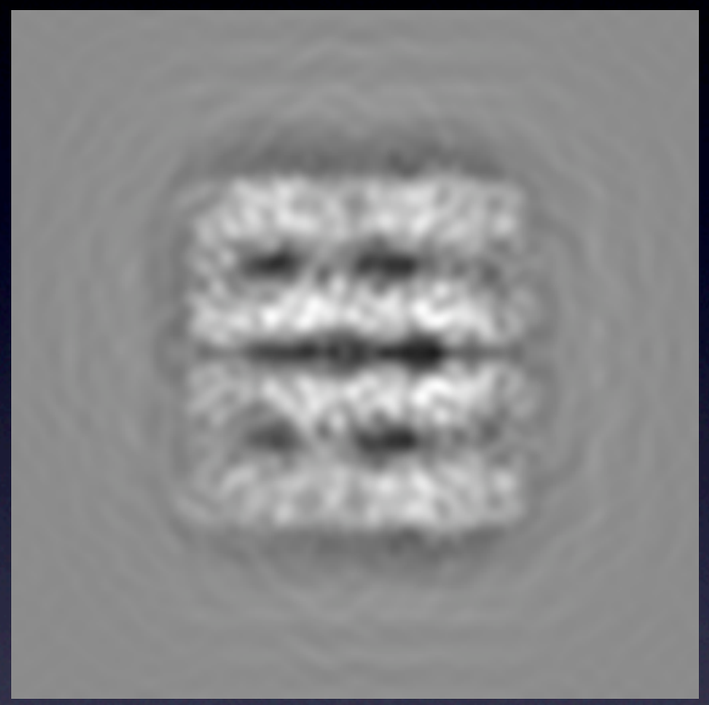
Can we subtract what we don't want ?



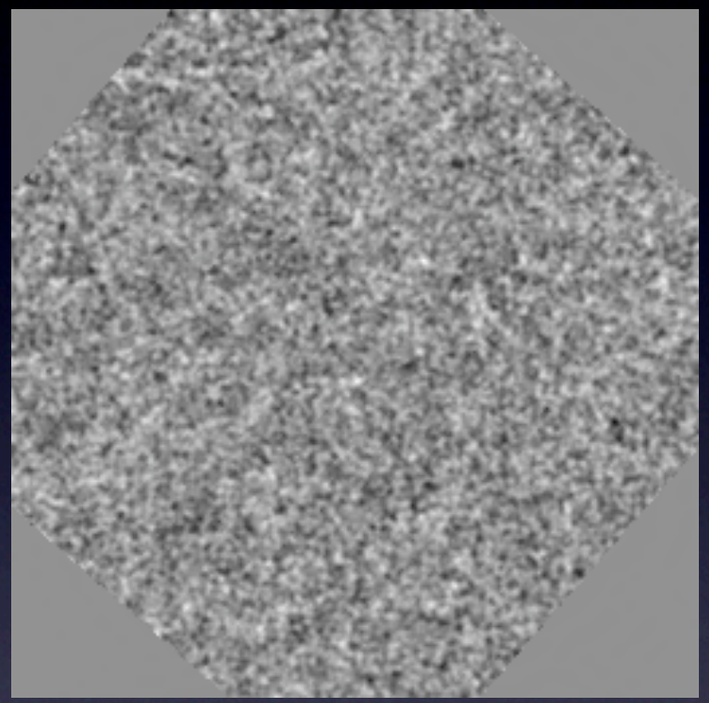
Remove One Monomer



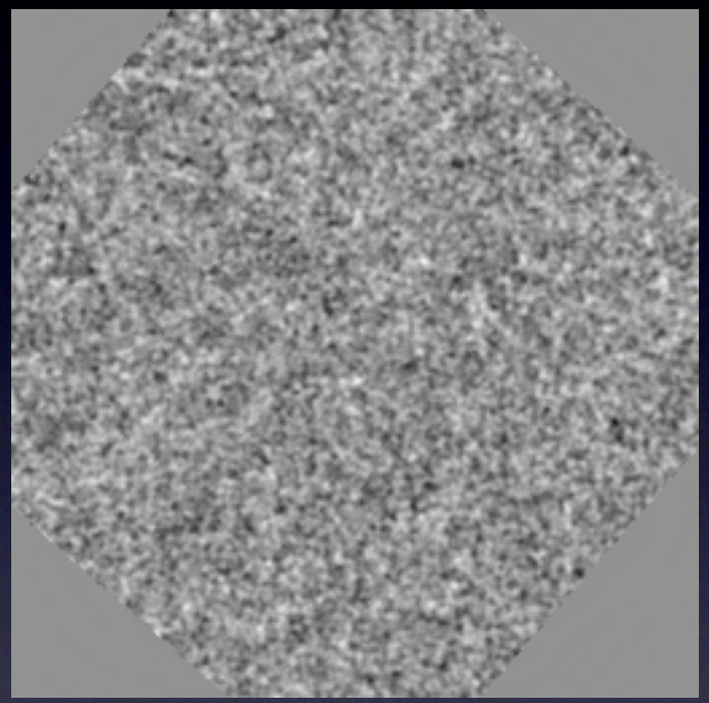
Remove One Monomer



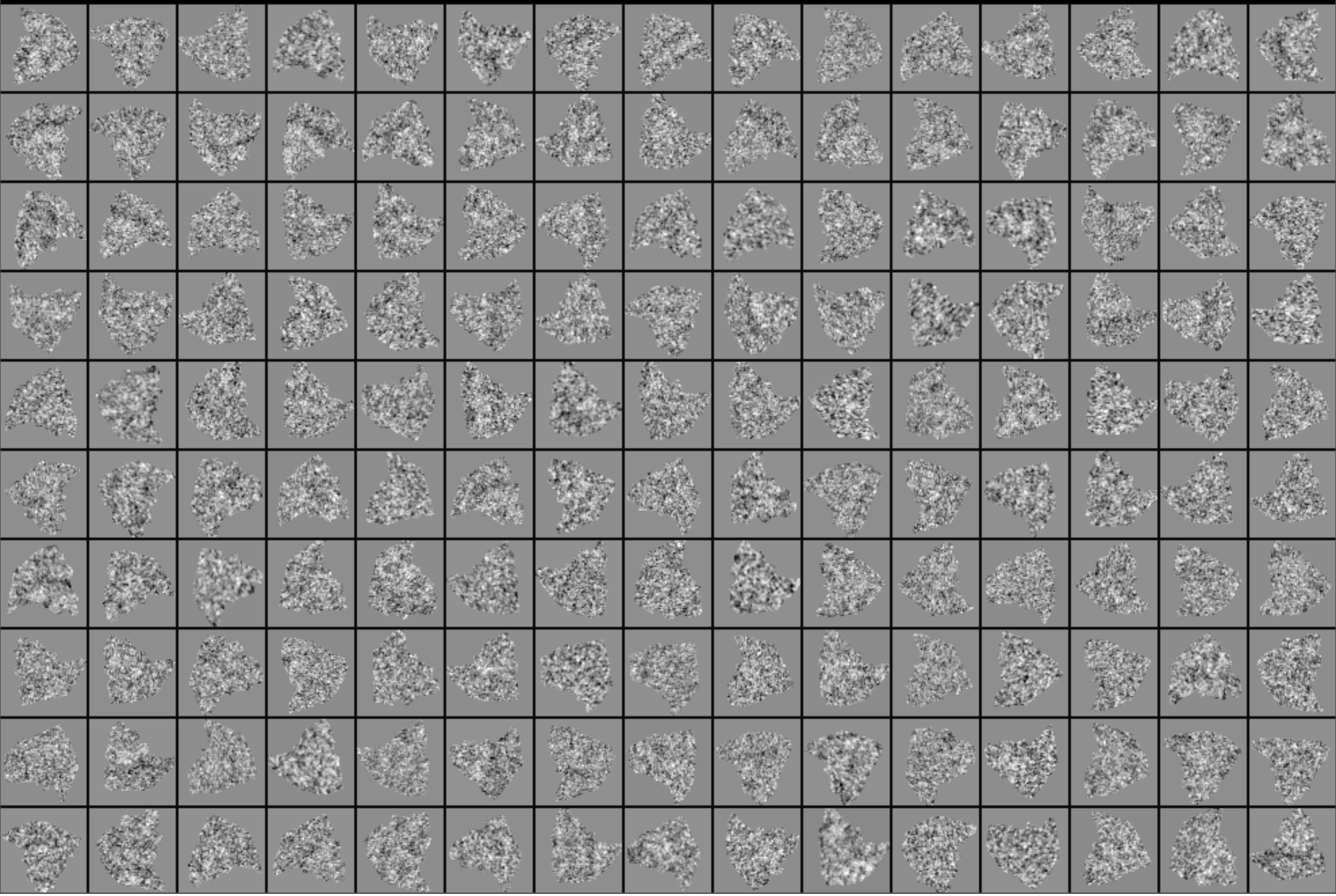
Remove One Monomer



Remove One Monomer



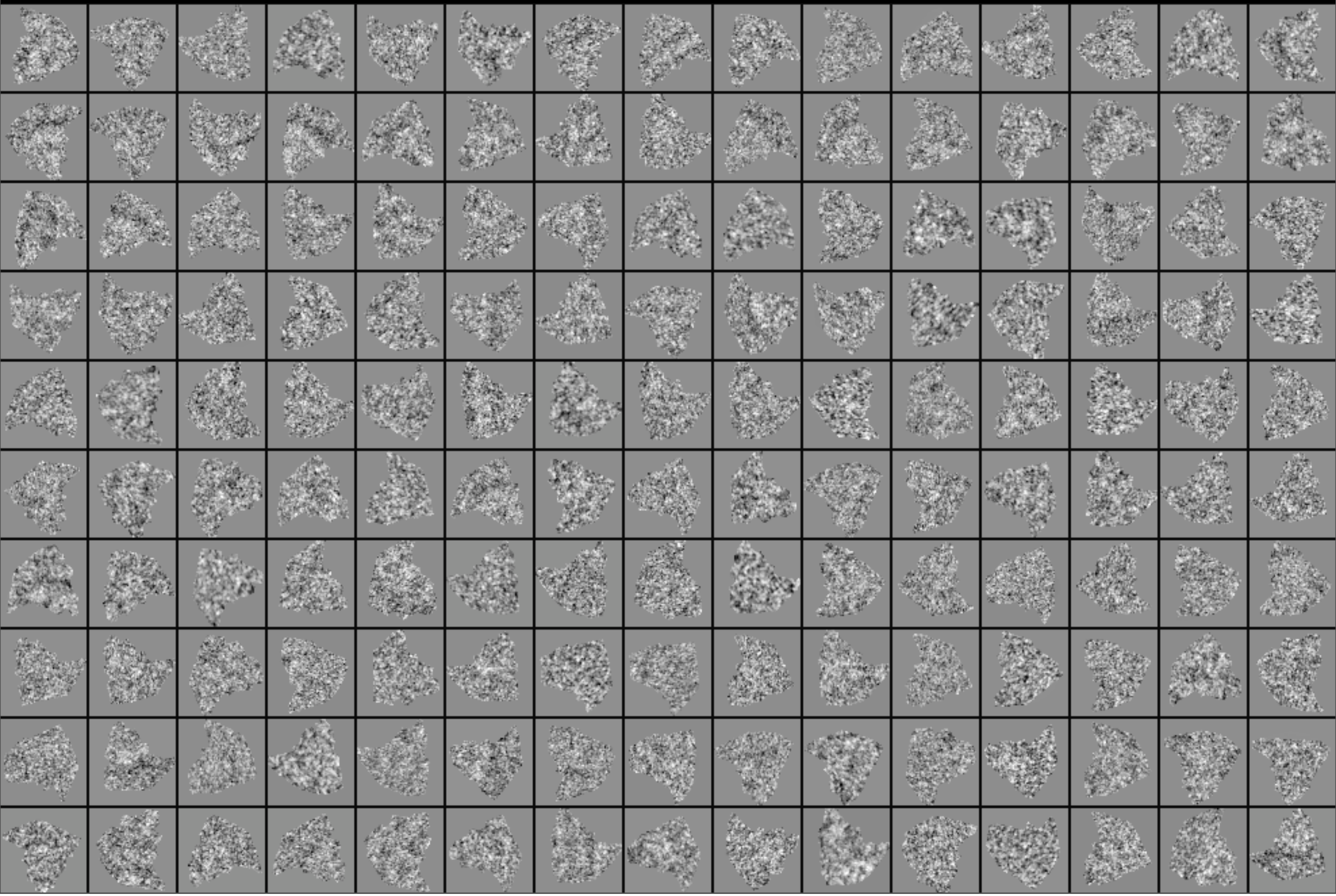
Many particles

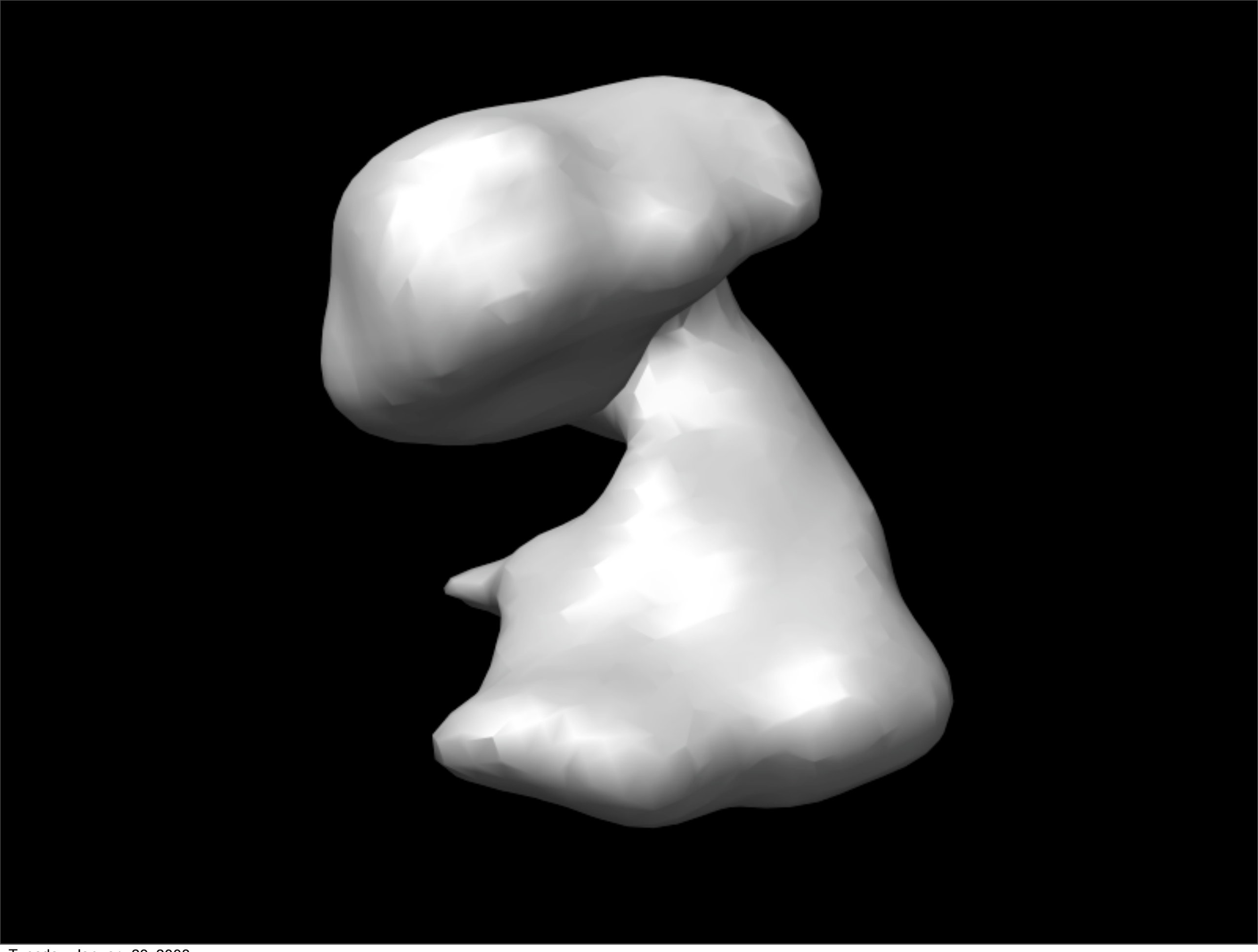


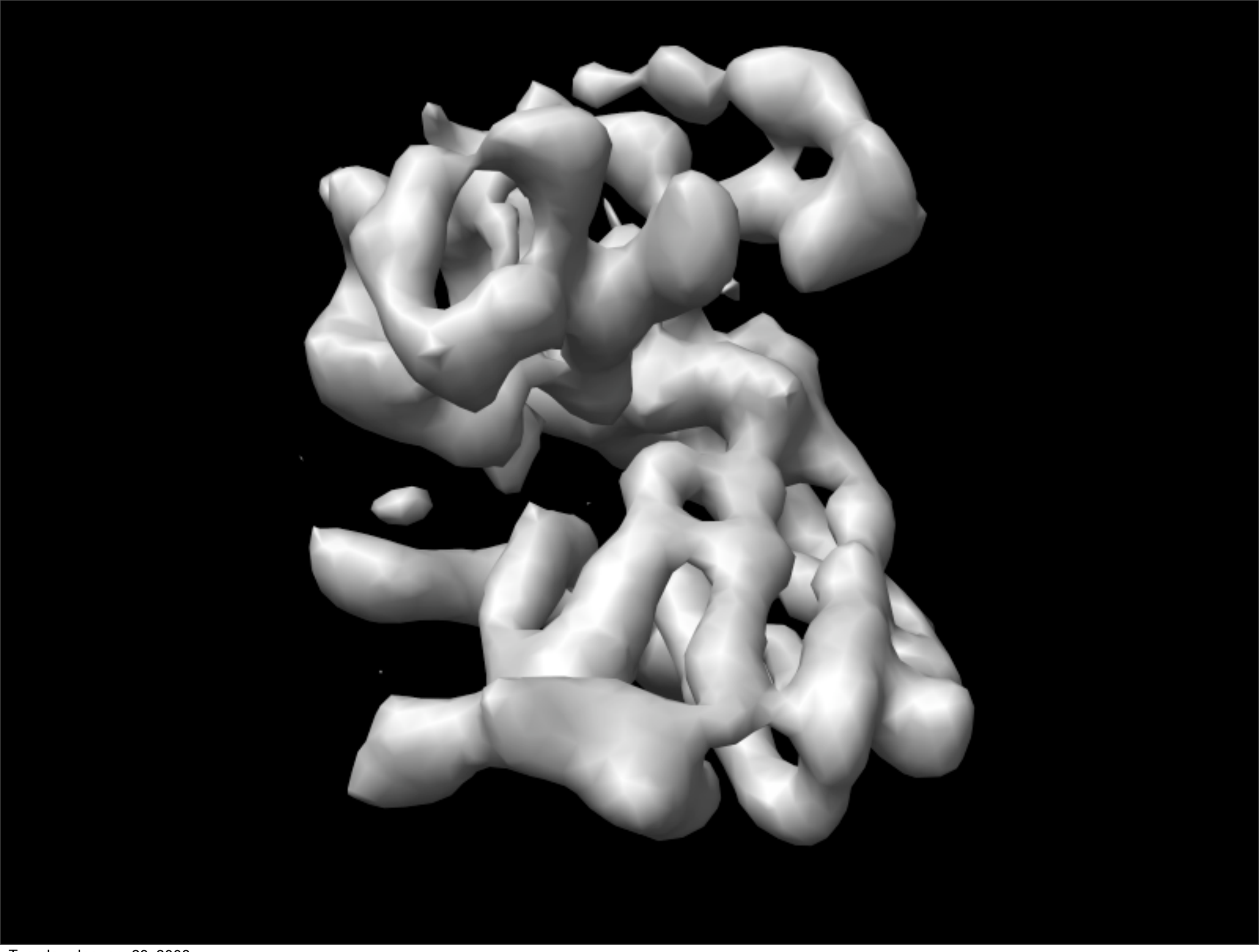
Many Particles

- 40,000 GroEL -> 560,000 monomers
- 128x128 with mask @ 1 Å/pixel
- ~250,000 CPU-hr for a full refinement
- Multi-model → millions of CPU-hr
- So,
- 64x64, 100,000 particles

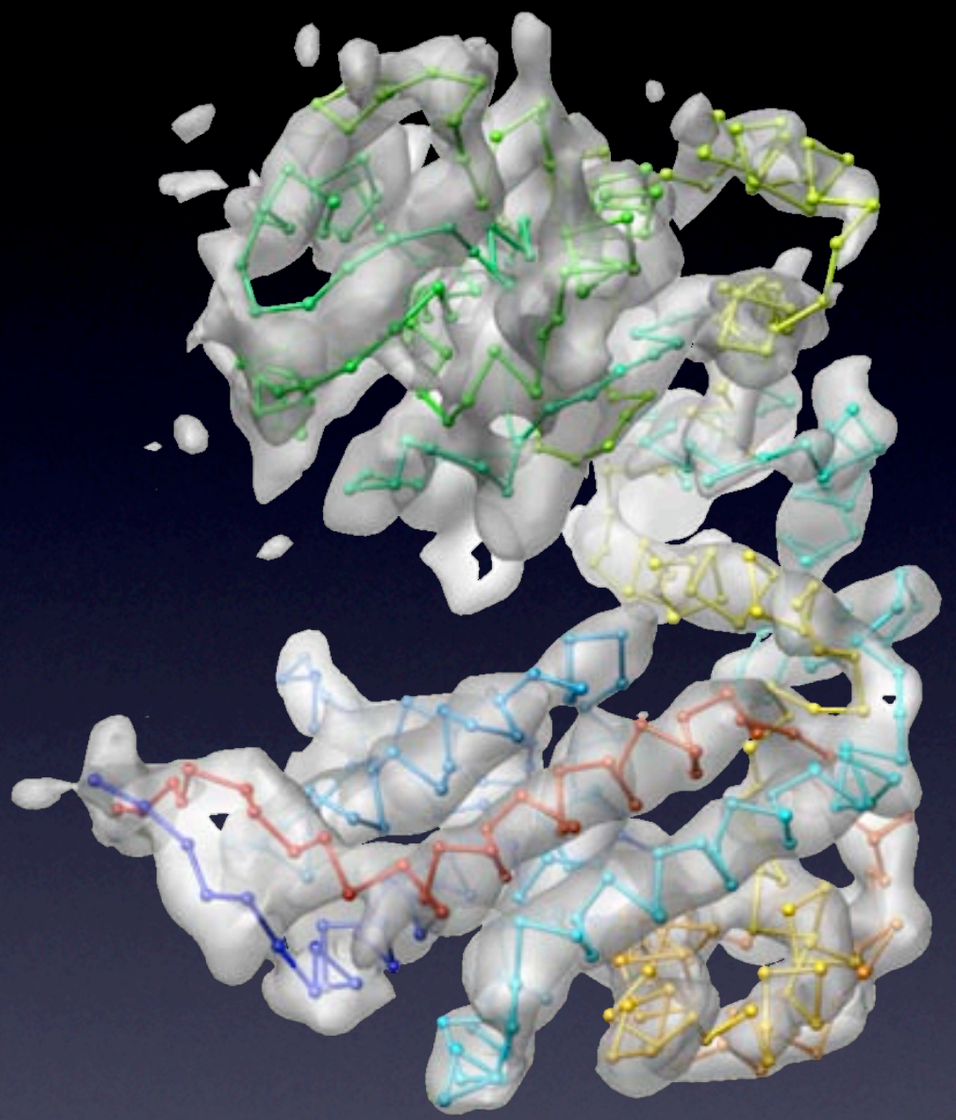
Many particles







Acknowledgements



GroEL

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

FAS

- Wei Zhang
- Sub Chirala
- Ziwei Gu
- Wah Chiu
- Salih Wakil
- Flo Quiocho

Mmcpn

- Wah Chiu
- Junjie Zhang
- Donghua Chen
- Judith Frydman (Stanford)
- Stefanie Reissmann (Stanford)

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

Graphics produced using UCSF Chimera and EMANimator.