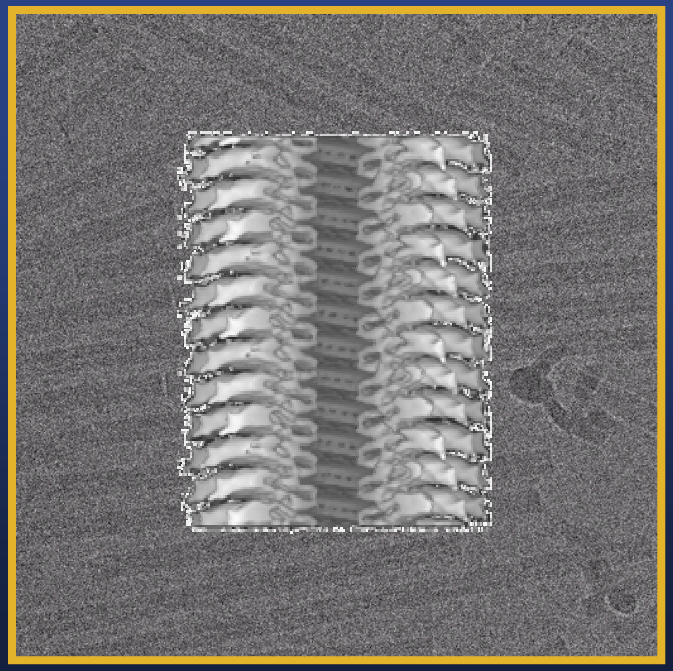
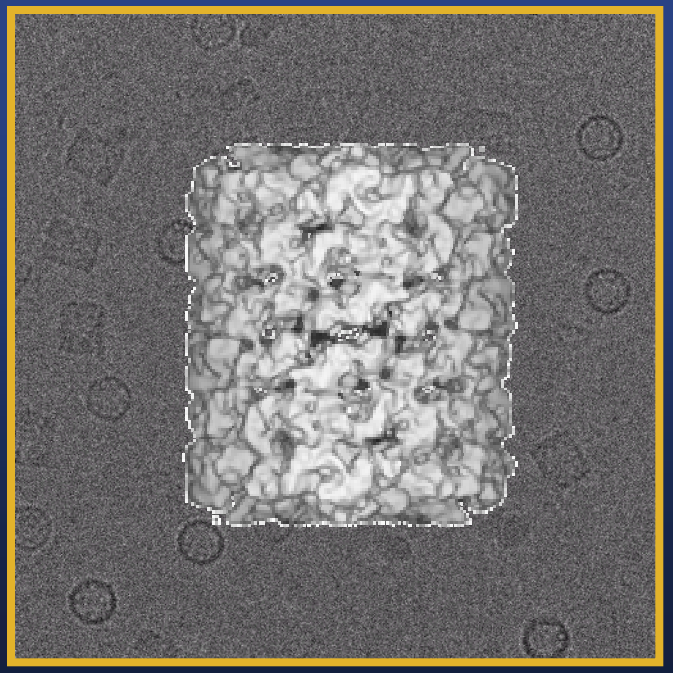
This material is provided for educational use only. The information in these slides including all data, images and related materials are the property of :

Bridget Carragher
Department of Cell Biology, CB129
The Scripps Research Institute,
10550 North Torrey Pines Road
La Jolla, CA 92037, USA
Tel: (858) 784-9050
Fax: (858) 784-9090

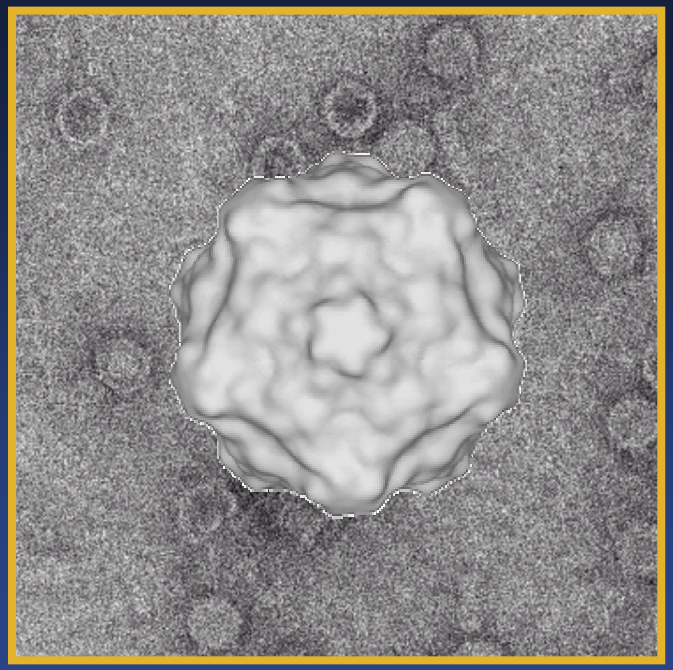
No part of this material may be reproduced without explicit written permission.

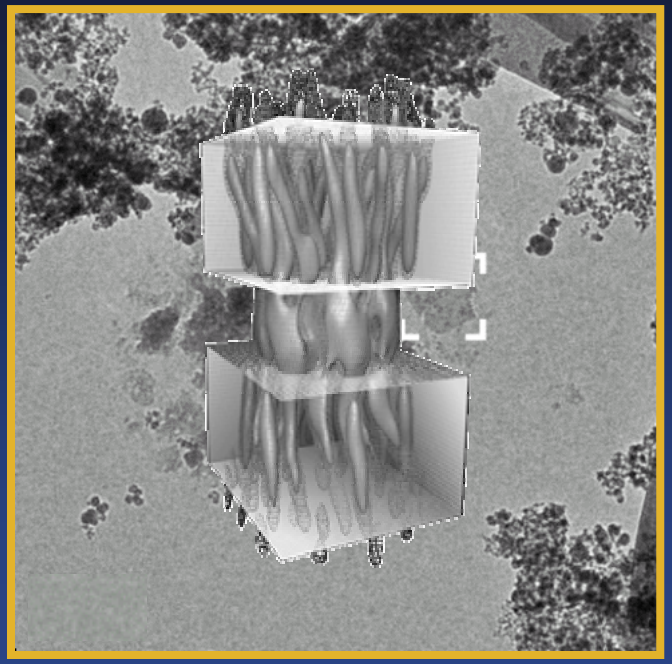




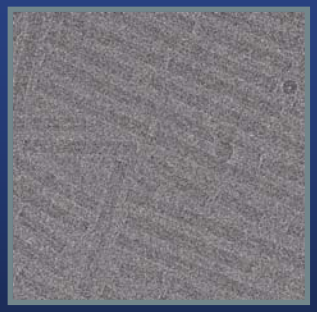
Automated Processing and Reconstruction

Bridget Carragher
Clint Potter

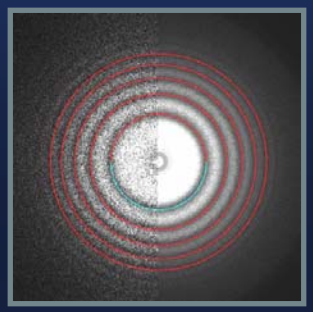




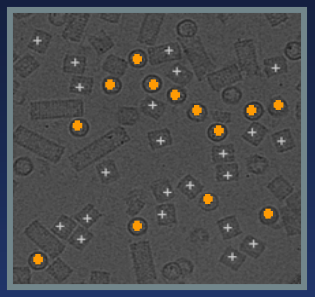
Automated image processing and reconstruction



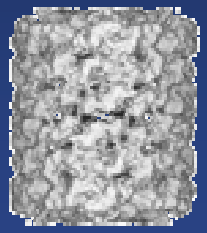
Assess the images



Determine the CTF

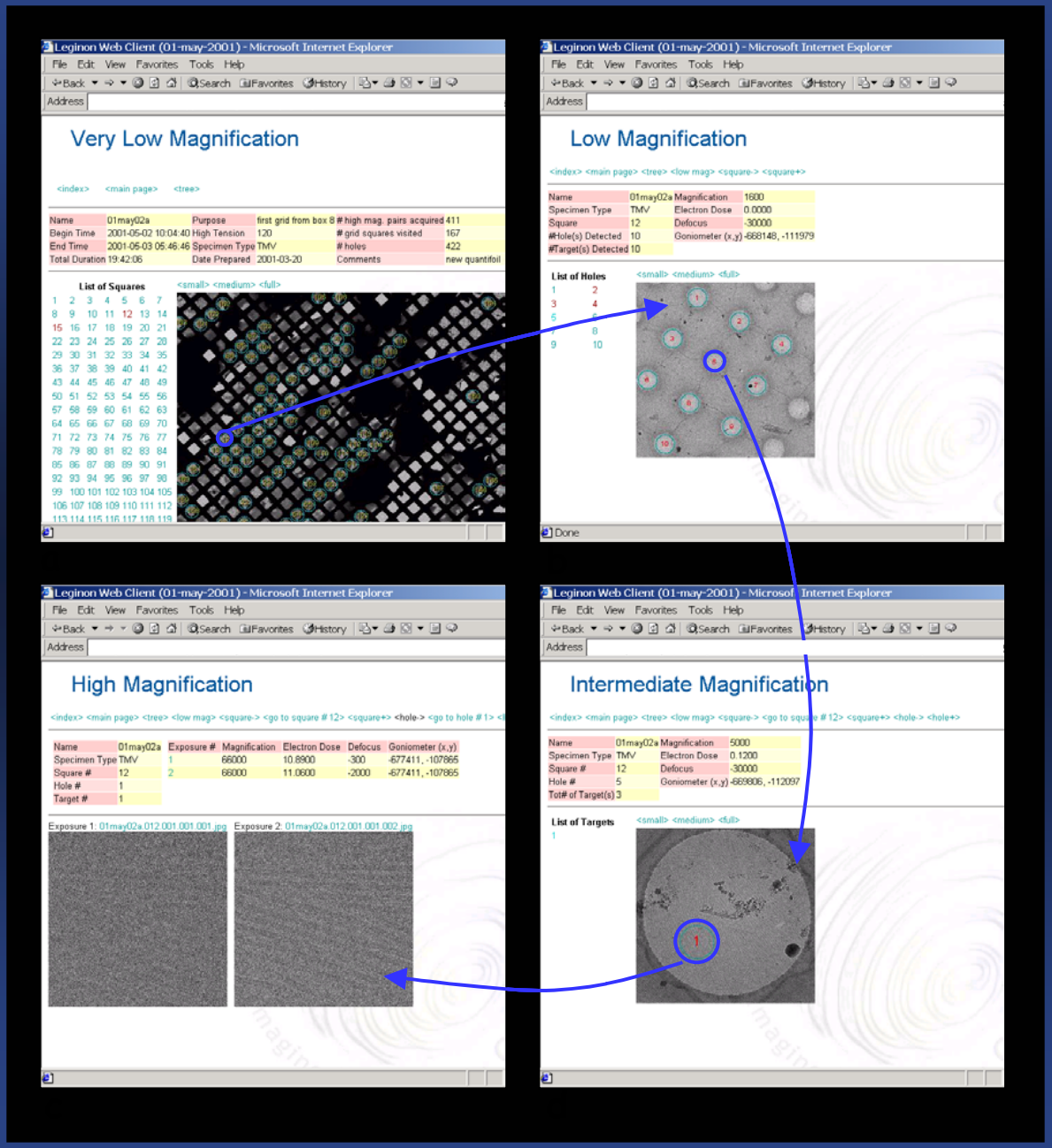


Select and segment particles



Reconstruct 3D map

Leginon Database



Database (at ~2 years):

# of experiments:	~300
# of images :	~235,000
# of records:	~45,000,000
Online data:	~1 Terabyte
Archived data:	~360 Gbytes
Size of database:	~270 Mbytes

"Typical experiment"

5000 images collected
20 Gbyte data

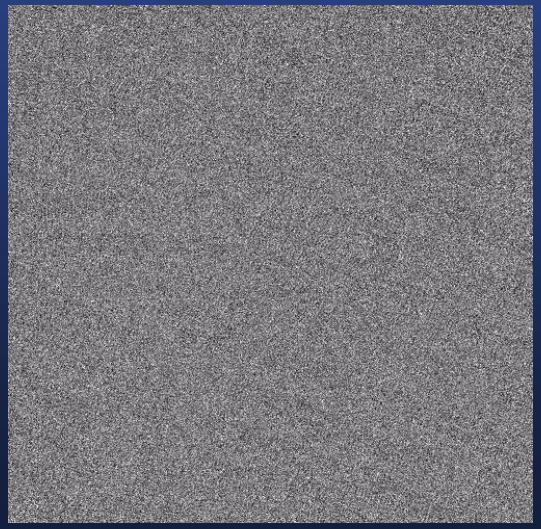
- Provides a means of systematically assessing the quality of the data.
- Provides a system for delivering "standard" datasets to other interested communities.

http://nramm.scripps.edu/

Fellman, et al, J. Struct. Biol, 137, p. 273-282 (2002)

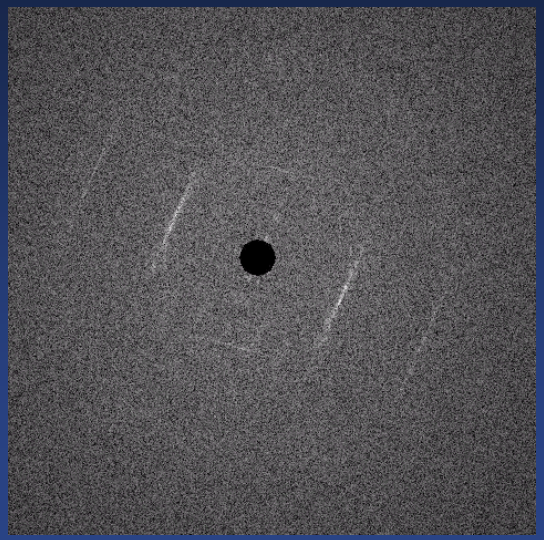
Assessing the images: TMV

Exposure 1

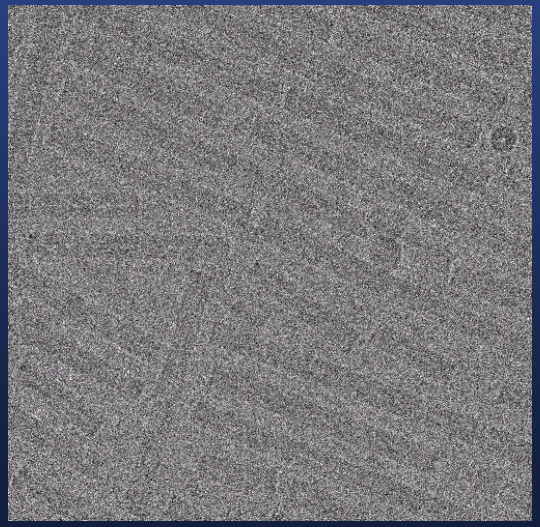


dose ~12e^{-}/Å^2

Δf = -30nm

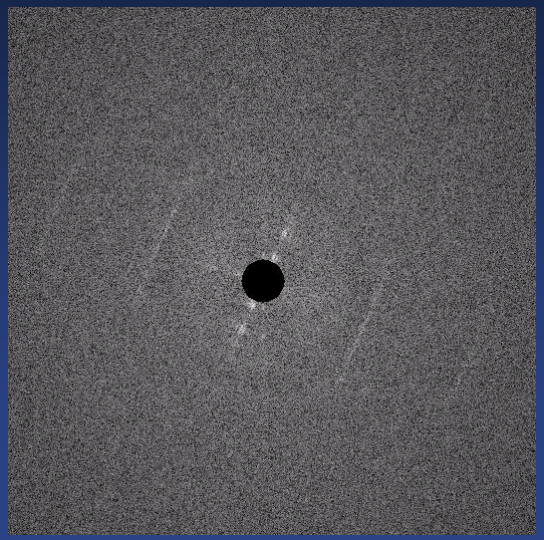


Exposure 2

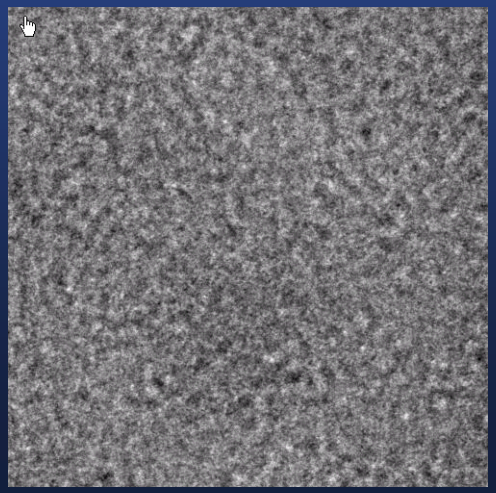


dose ~24e^{-}/Å^2

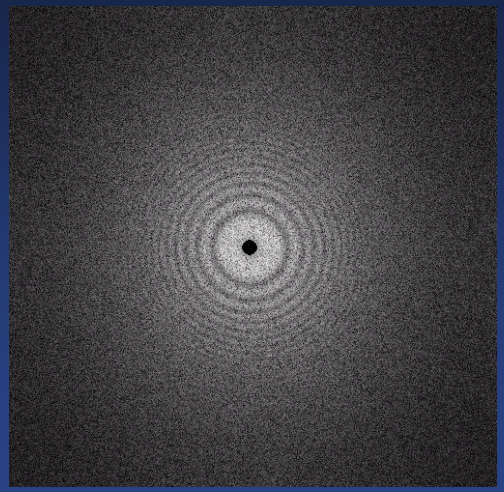
Δf = -200nm



Focus image

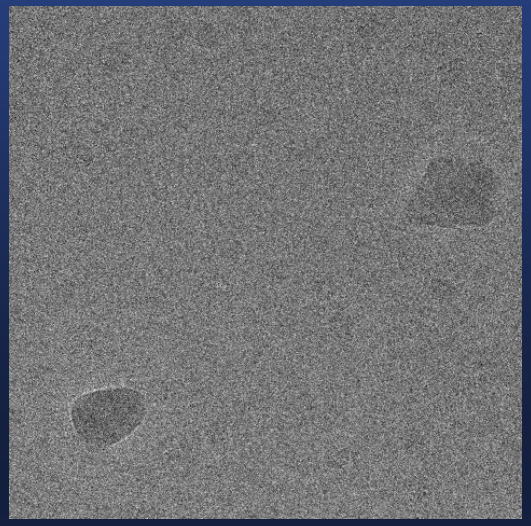


Δf = -200nm



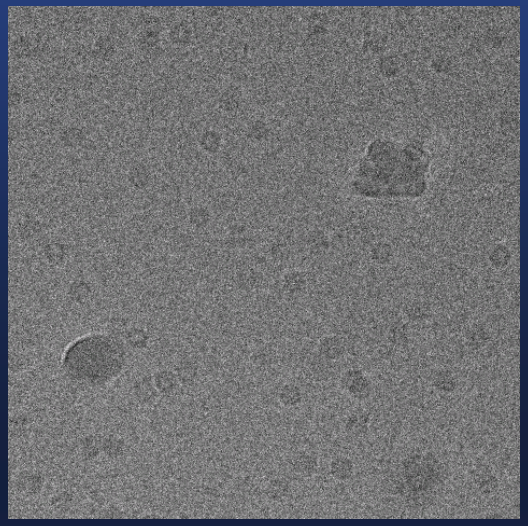
Assessing the images: GroEL

Exposure 1

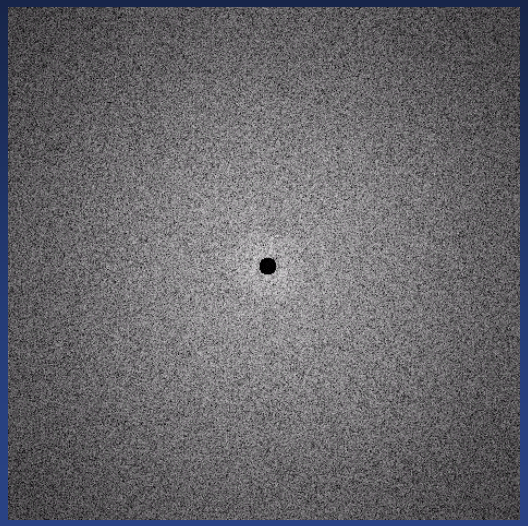


Δf = -60nm

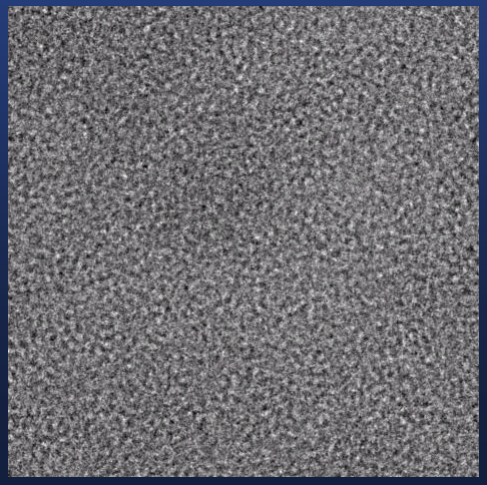
Exposure 2



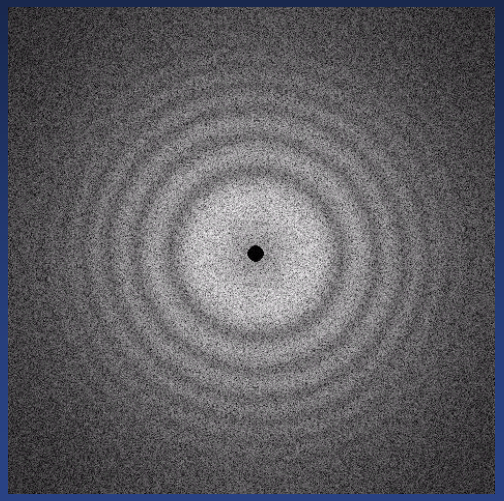
Δf = -200nm



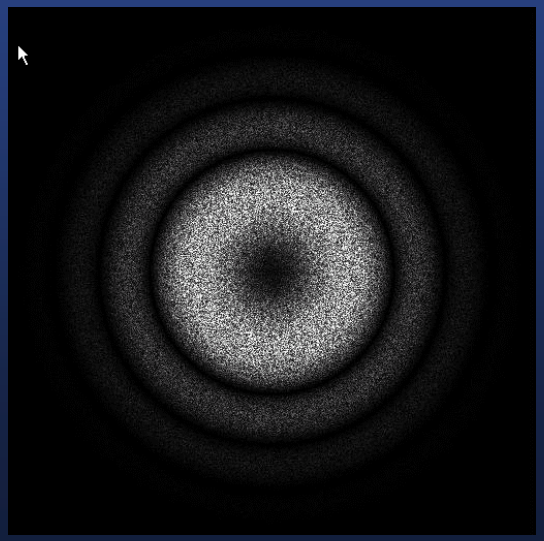
Focus image

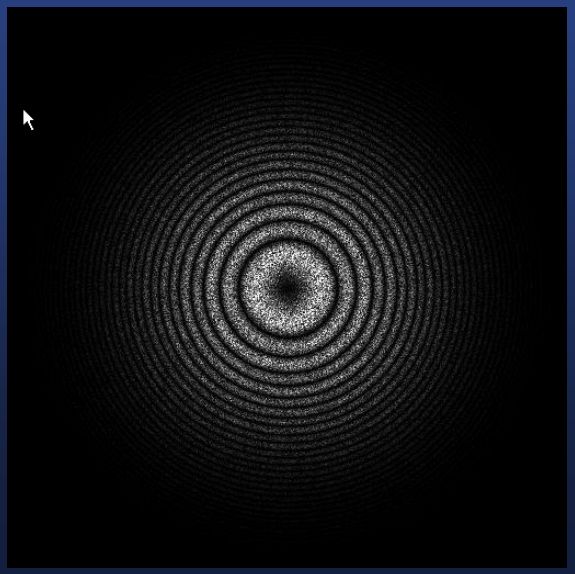


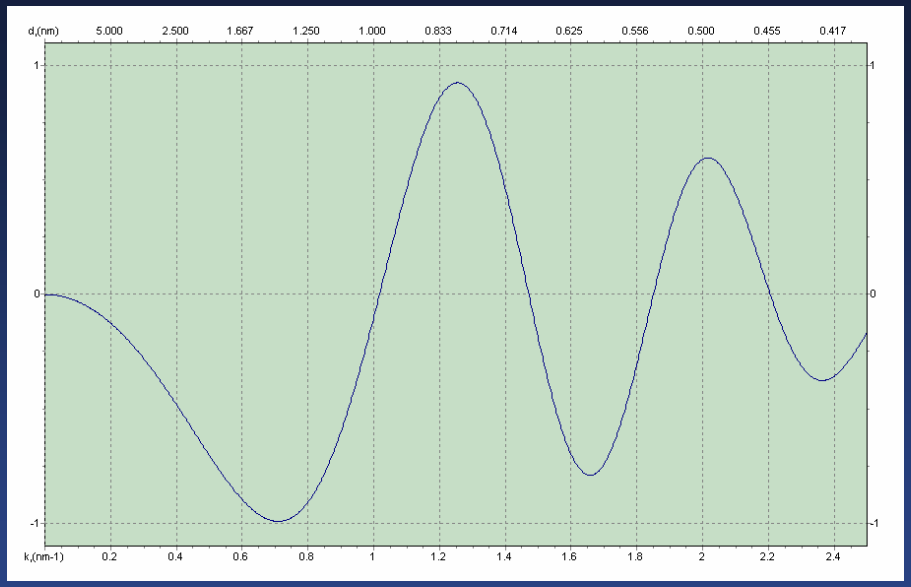
Δf = -60nm

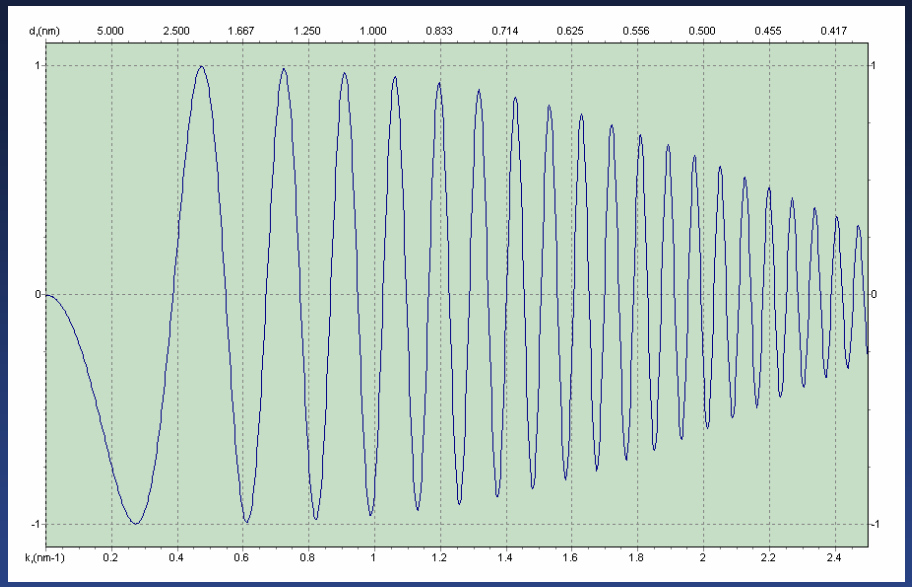


Defocus determination



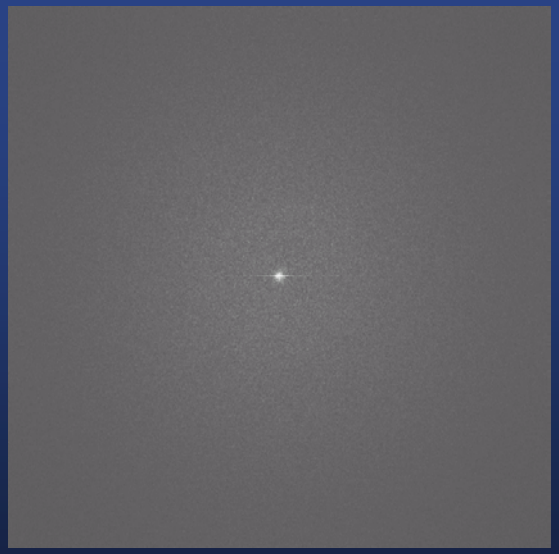




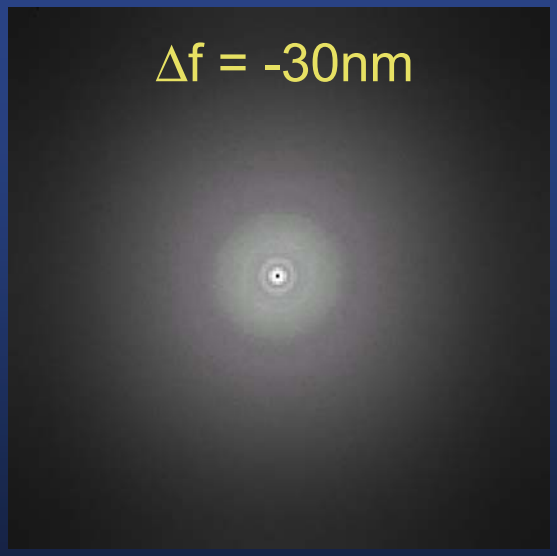


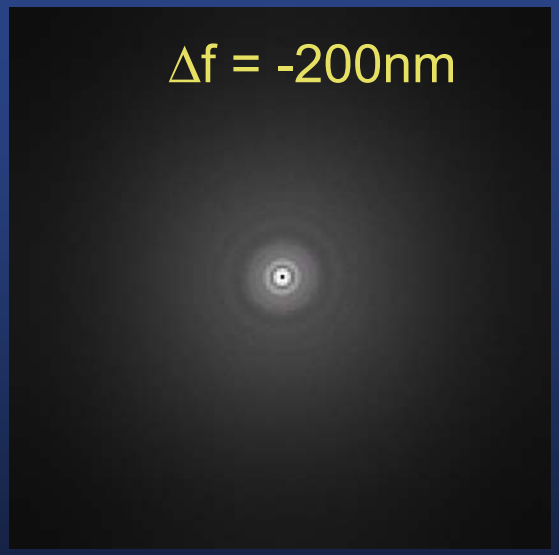
Defocus –300nm

Defocus –2000 nm

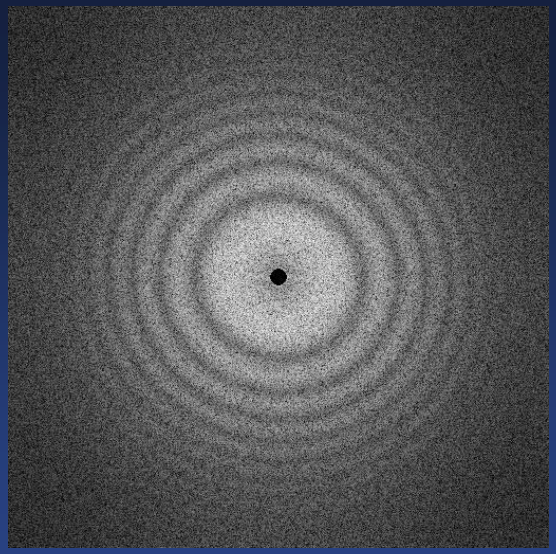


1 image : 17 particles

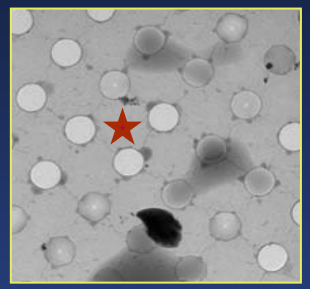




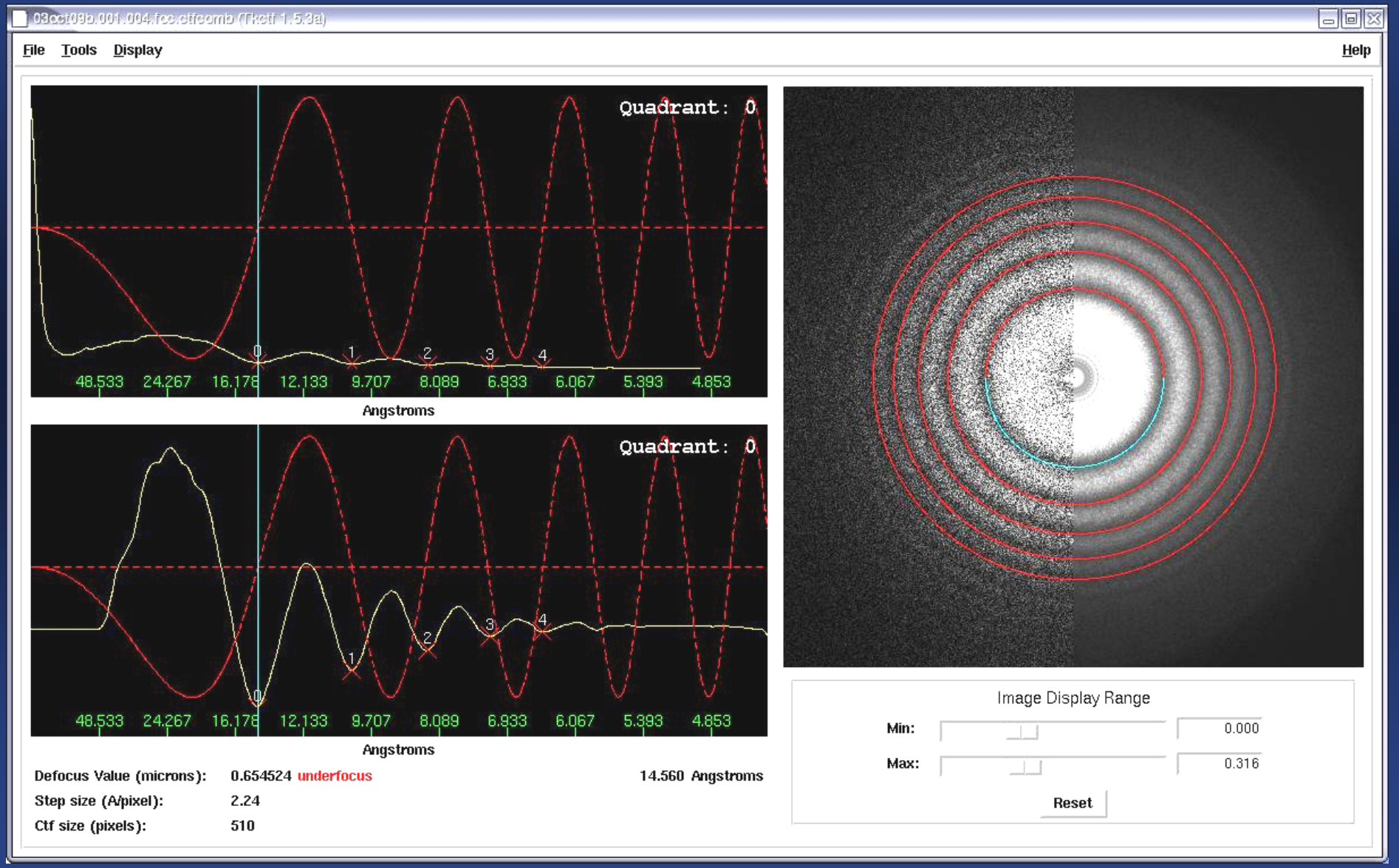
Set of images : ~ 2000 particles



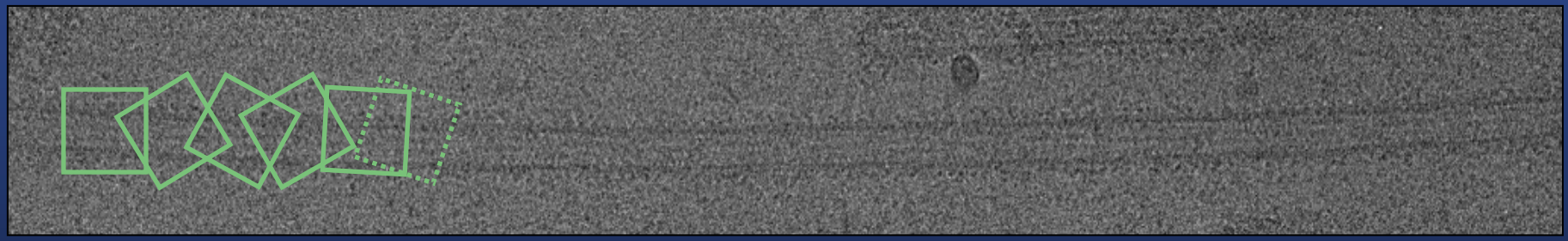
focus power spectrum



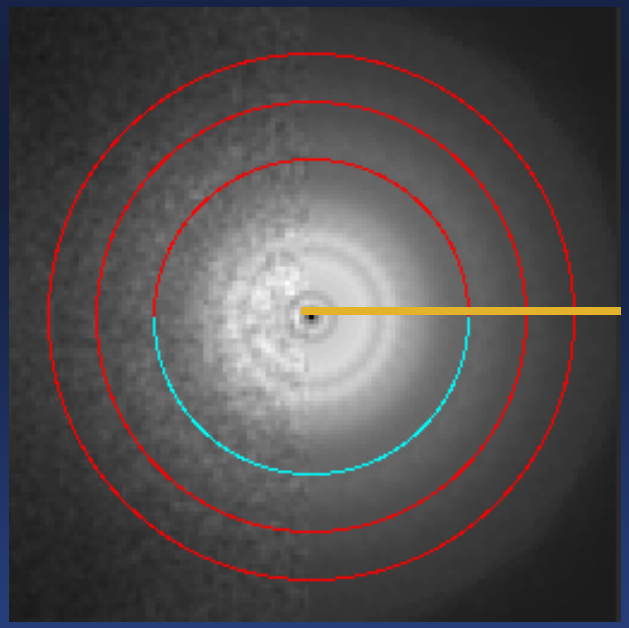
“Automated” defocus determination

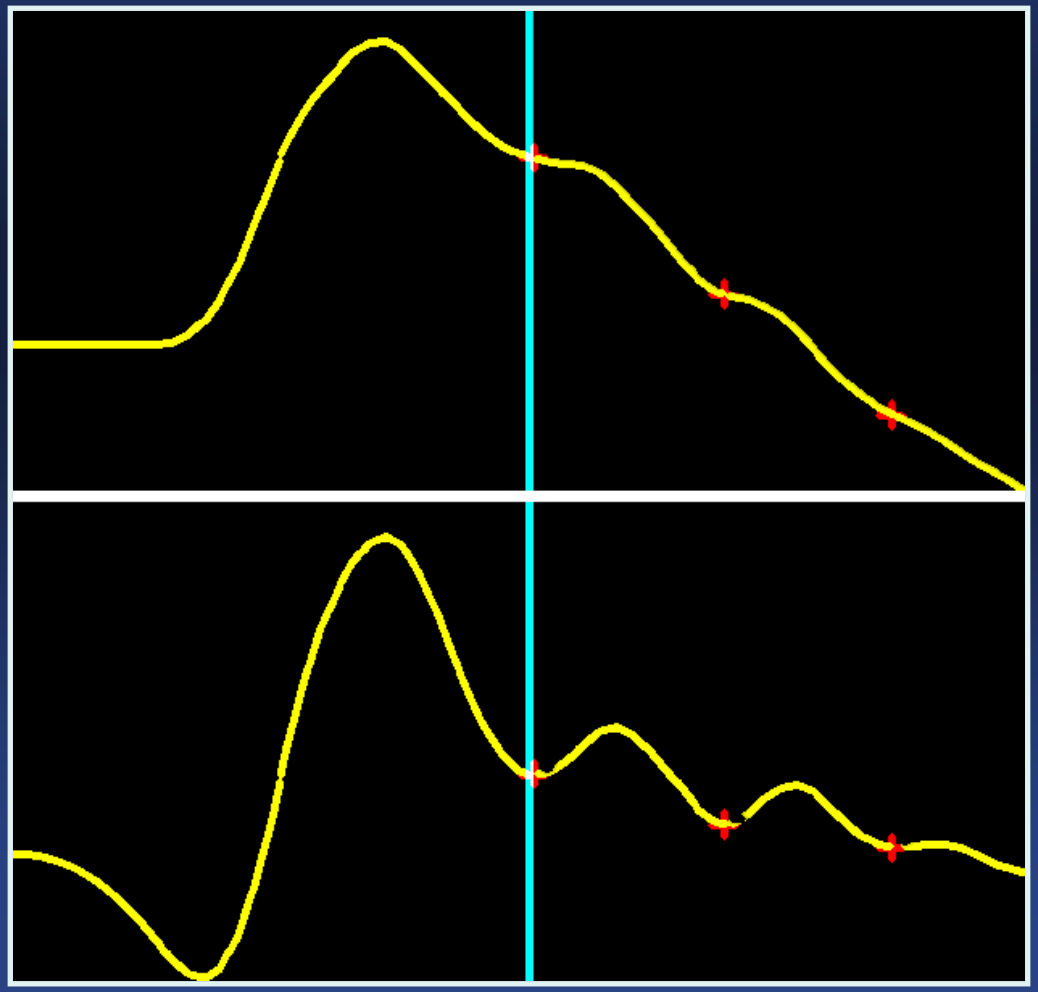


Defocus determination using incoherent sum of structure



FT(incoherent sum)





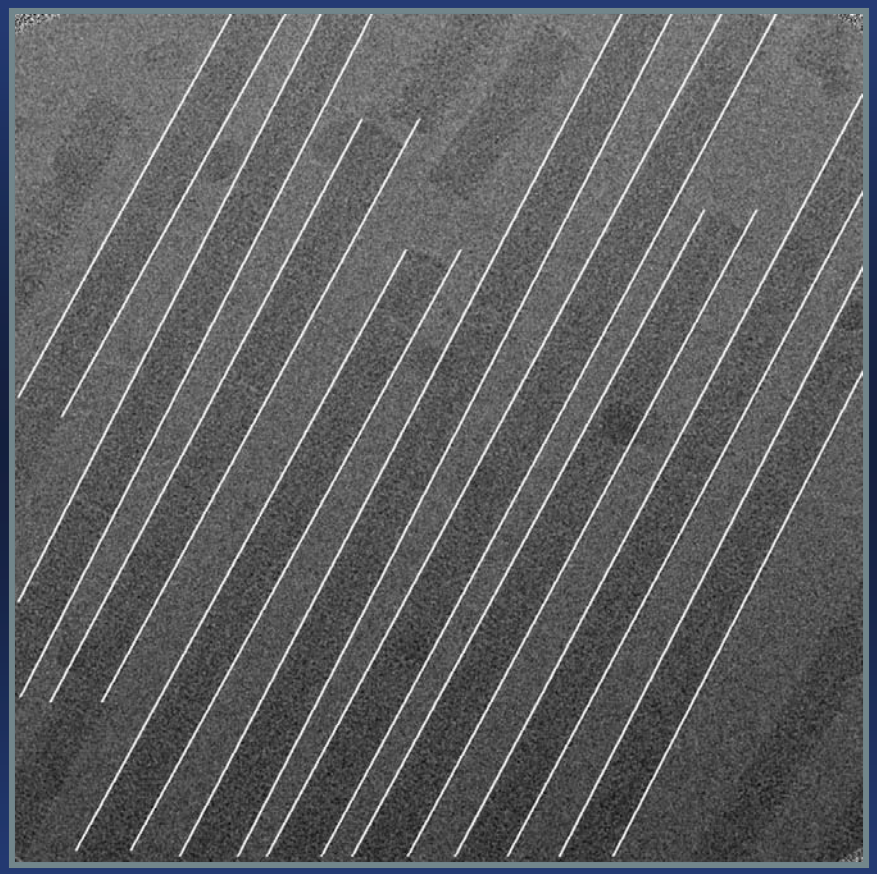
A few references on automated CTF determination:

Automated Determination of Parameters Describing Power Spectra of Micrograph Images in Electron Microscopy
Zhong Huang, Philip R. Baldwin, Srinivas Mullapudi, and Pawel A. Penczek
J. Struct. Biol., In press.

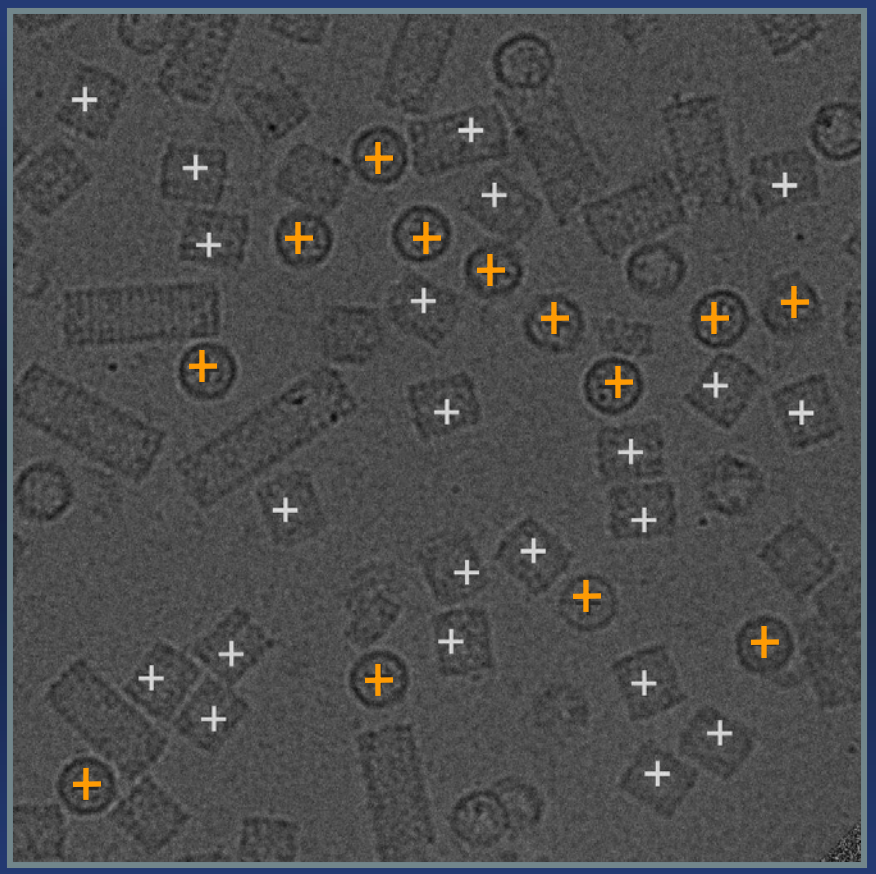
Automatic CTF correction for single particles based upon multivariate statistical analysis of individual power spectra.
Sander, B., Golas, M. M., Stark, H.
J. Struct. Biol. 142, 392-401 (2003)

Accurate determination of local defocus and specimen tilt in electron microscopy.
Mindell, J. A., Grigorieff, N.
J. Struct. Biol. 142, 334-347 (2003)

Automated specimen selection and segmentation



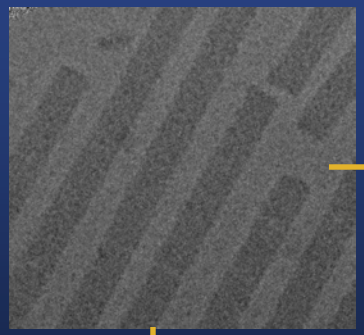
Filaments

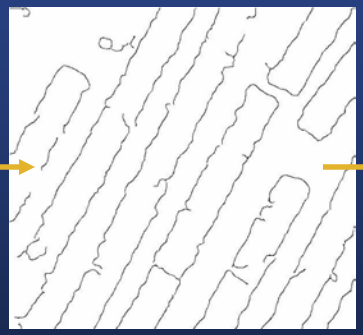


Single particles

Filament selection and segmentation

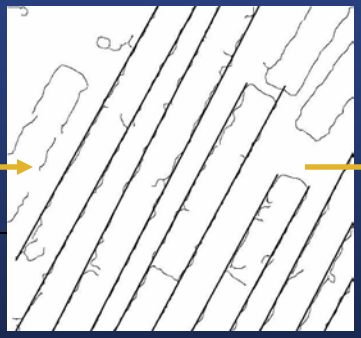
far from focus



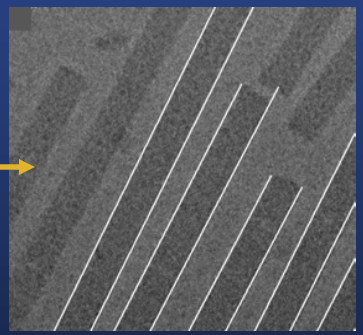


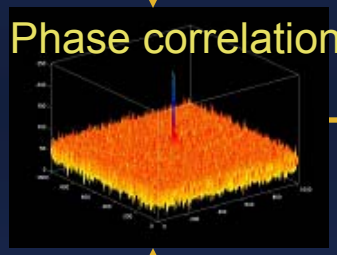
edges

grouping



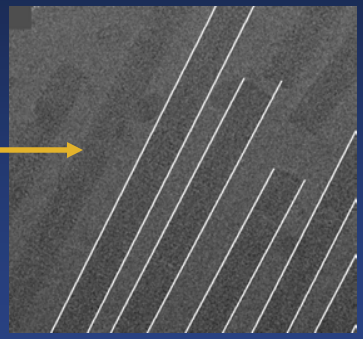
organization





Shift

near to focus



Zhu et al (2001) Automated identification of filaments in cryoelectron microscopy images, JSB, 135, 302-312.

Single particle selection and segmentation

Algorithm:

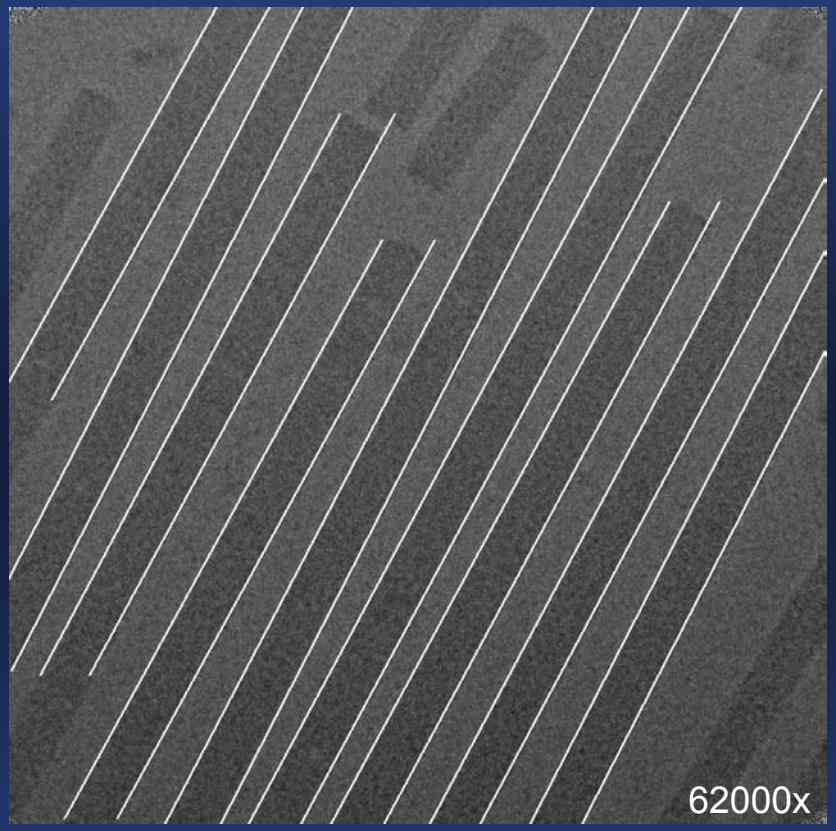
(i) Find the edges w/ Canny

(ii) Find the shapes w/ Hough

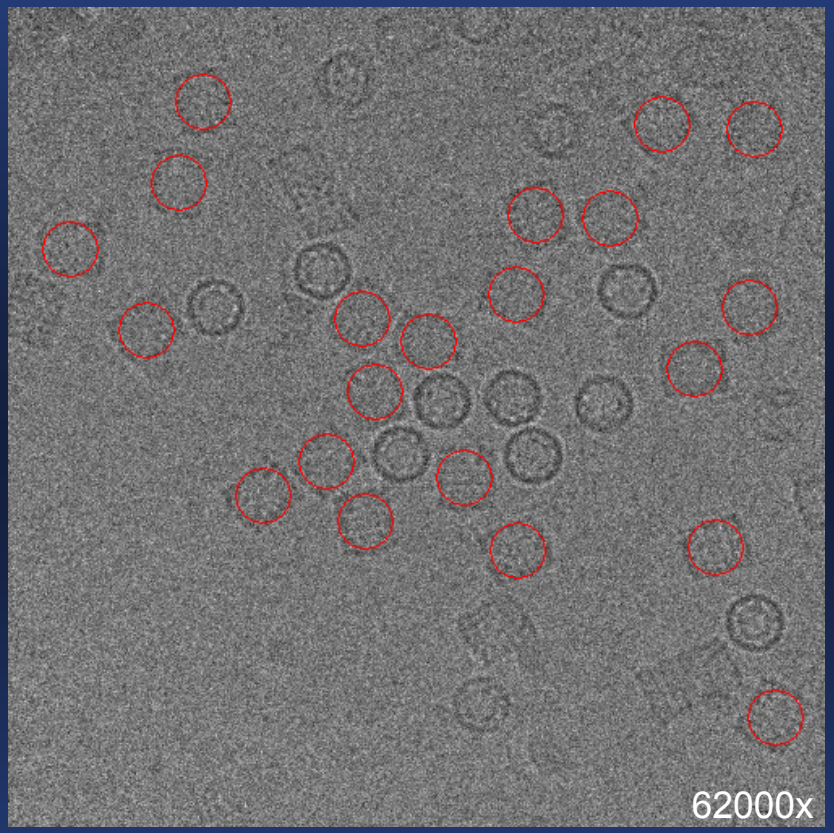
(iii) Pruning

Zhu et al. (2003) "Automatic Particle Detection through Efficient Hough Transforms," IEEE Transactions on Medical Imaging, 22, 1053-1062

Some examples



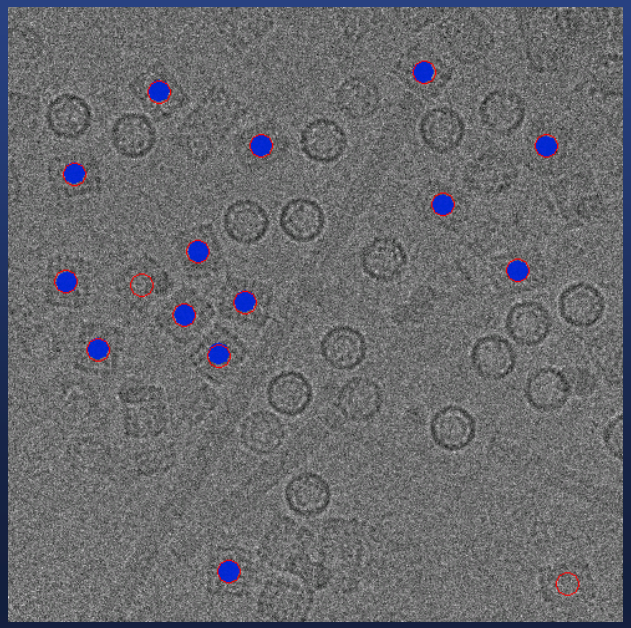
411 image pairs
686 filaments automatically selected

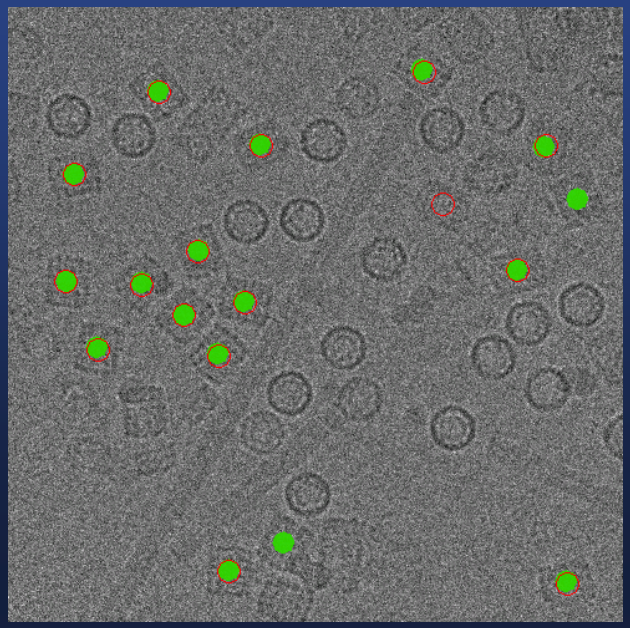


816 image pairs
23,000 particles automatically selected

Automated particle selection “bakeoff”

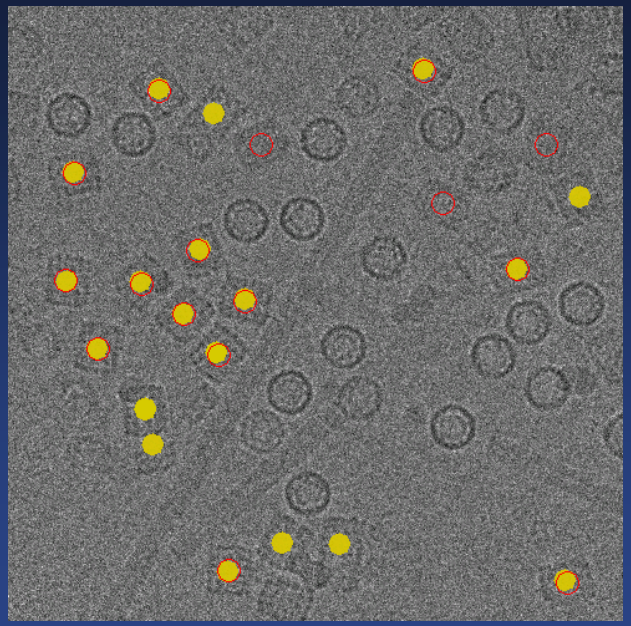
human vs. human

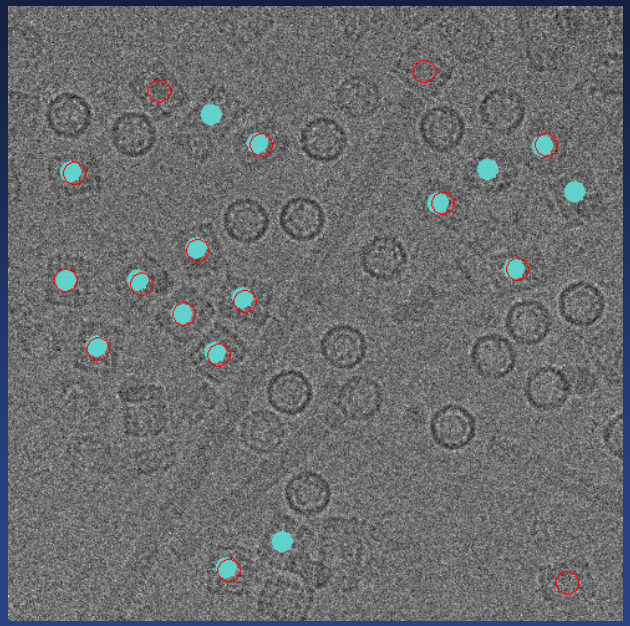




human vs. template matching

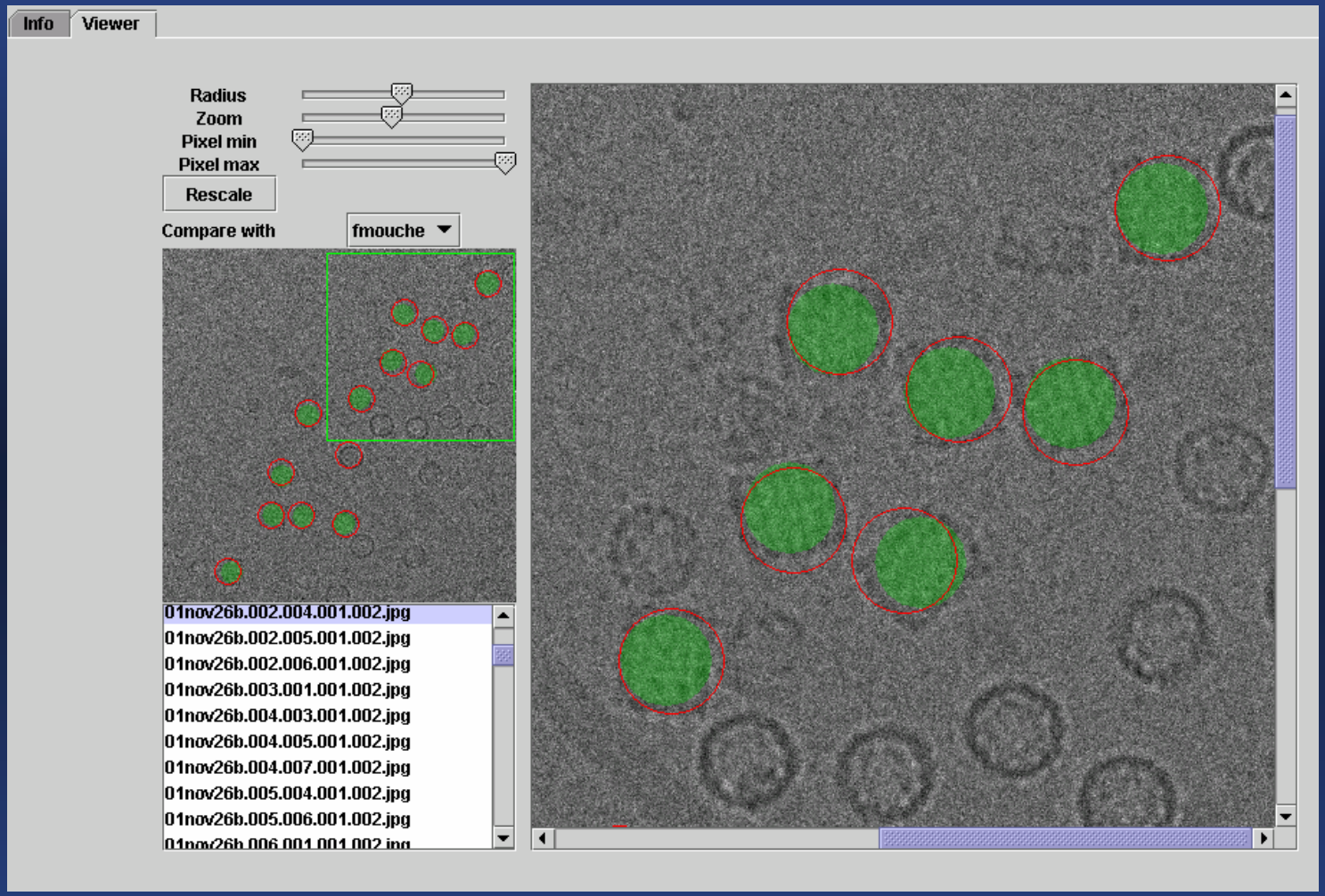
human vs. feature recognition





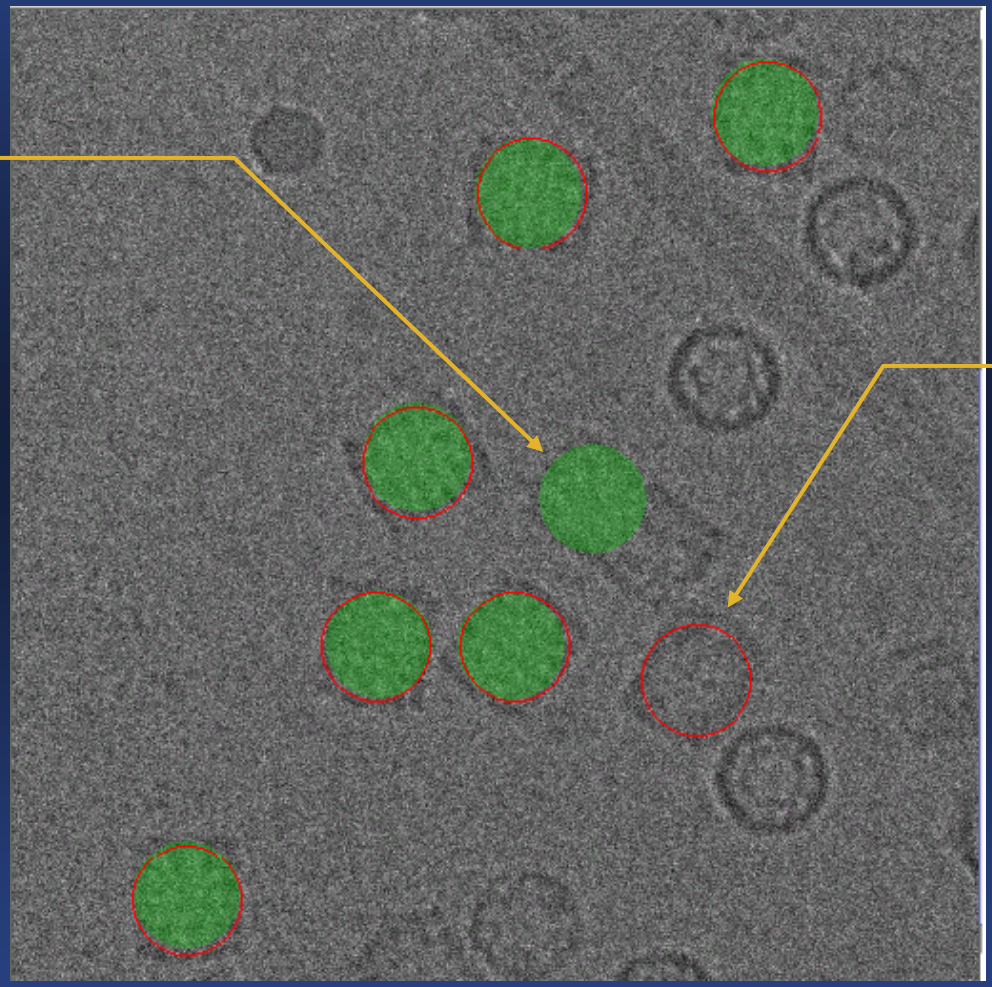
human vs. machine learning

Automated particle selection Bakeoff



http://ami.scripps.edu/leginon/particle_viewer/

Bakeoff Results

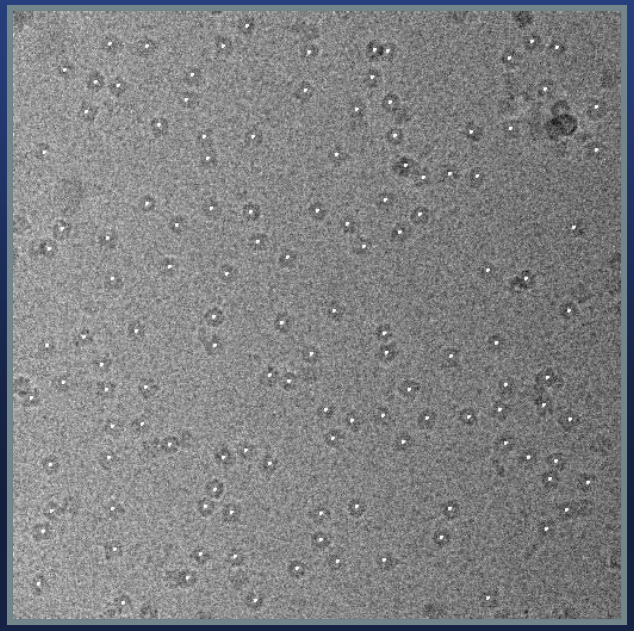


Bakeoff Results

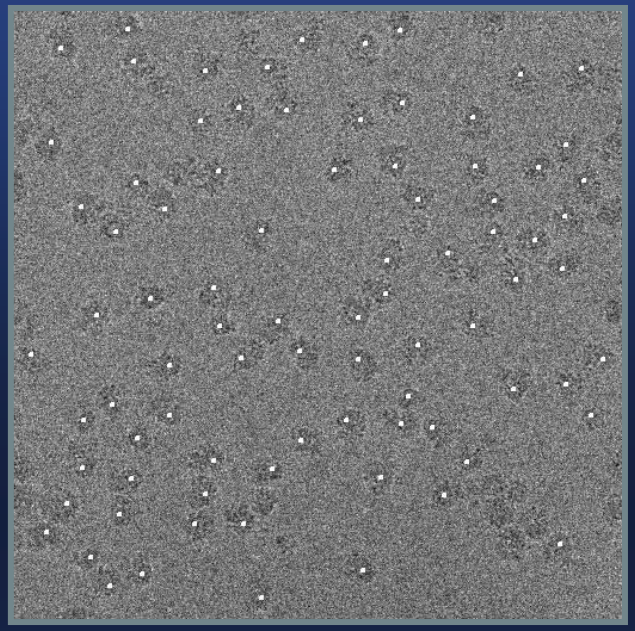
Test / Truth		Bern	Mouche (Manual)	Haas (Manual)	Hall	Ludtke (EMAN)	Melanson[1] (Manual)	Roseman	Sigworth	Bajaj	Zhu
Bern (948) [2]			16.2 23.8	21.5 21.0	36.3 37.7	43.1 30.3	91.0 24.1	10.3 30.3	64.5 6.6	11.5 33.9	17.1 28.0
Mouche (1042)		23.8 16.2		11.7 2.3	27.4 22.0	43.4 23.7	91.1 17.0	2.4 16.6	65.9 1.7	8.3 24.7	9.7 13.7
Haas (944)		21.0 21.5	2.3 11.7		26.2 28.2	41.1 28.4	91.0 24.1	1.5 23.9	64.2 6.4	7.0 31.0	8.8 21.3
Hall (969)		37.7 36.3	22.0 27.4	28.2 26.2		52.0 39.9	92.3 33.0	19.3 35.8	72.1 25.2	24.3 42.2	25.7 33.7
Ludtke (775)		30.3 43.4	23.7 43.4	28.4 41.1	39.9 52.0		92.1 45.5	20.3 49.4	63.6 22.2	21.0 51.9	23.5 45.4
Melanson (112)		24.1 91.0	17.0 91.1	24.1 91.0	33.0 92.2	45.5 92.1		13.4 92.0	63.4 88.6	17.0 92.7	23.2 92.2
Roseman (1219)		30.3 10.3	16.6 2.4	23.9 1.5	35.8 19.3	49.4 20.3	92.0 13.4		70.8 1.4	14.0 17.4	17.5 7.8
Sigworth (361)		6.6 64.5	1.7 65.9	6.4 64.2	25.2 72.1	22.2 63.6	88.6 63.4	1.4 70.8		5.0 73.0	6.1 68.7
Bajaj (1269)		33.9 11.5	24.7 8.3	31.0 7.0	42.2 24.3	51.9 21.0	92.7 17.0	17.4 14.0	73.0 5.0		24.0 11.4
Zhu (1109)		28.0 17.1	13.7 9.7	21.3 8.8	33.7 25.7	45.4 23.5	92.2 23.2	7.8 17.5	68.7 6.1	11.4 24.0	
Median /Mean	FNR	28.0/26.2	16.6/15.3	23.9/21.8	33.7/33.3	45.4/43.8	92.0/91.4	10.3/10.4	65.9/67.4	11.5/13.3	17.1/15.9
	FPR	21.5/34.6	23.8/31.5	21.0/29.2	28.2/41.5	28.4/38.1	24.0/29.0	30.3/38.9	6.4/18.1	33.9/43.4	28.0/35.8
Std.	FNR	9.0	8.4	8.0	6.0	9.0	1.2	7.6	3.8	6.5	7.2
	FPR	27.6	30.0	31.0	25.6	24.4	16.1	27.0	27.8	25.1	28.5

Zhu, Y., B. Carragher, R.M. Glaeser, D. Fellmann, C. Bajaj, M. Bern, F. Mouche, F. de Haas, R.J. Hall, S.C. Ludtke, P.A. Penczek, A.M. Roseman, F.J. Sigworth, N. Volkmann, and C.S. Potter (2003) "Automatic Particle Selection: Results of A "Bakeoff"", JSB, In press.

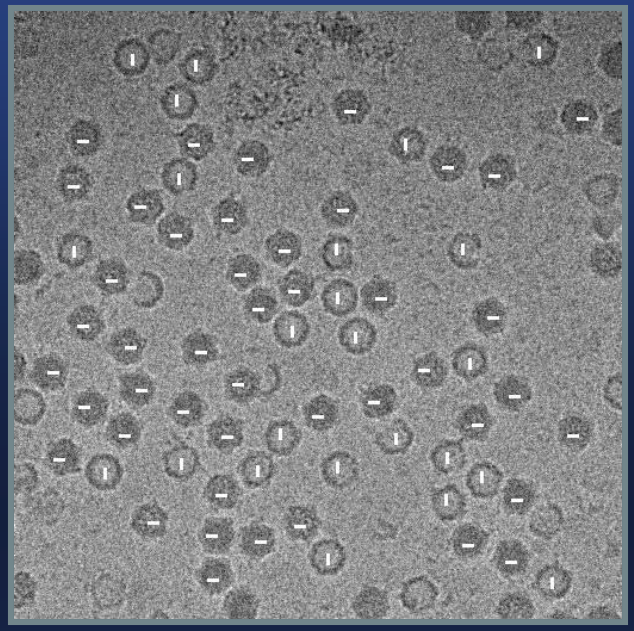
Some examples using the “Roseman” algorithm (+)



GroEL
images: 364
particles: 23,000
Time: 20 hours



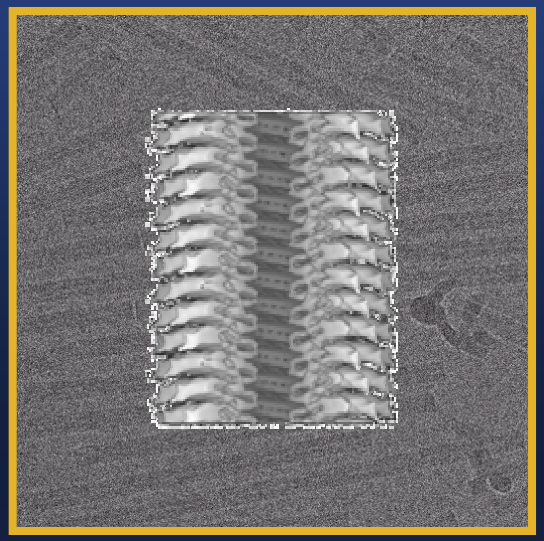
Ribosomes
images: 551
particles: 26,000
Time: 36 hours

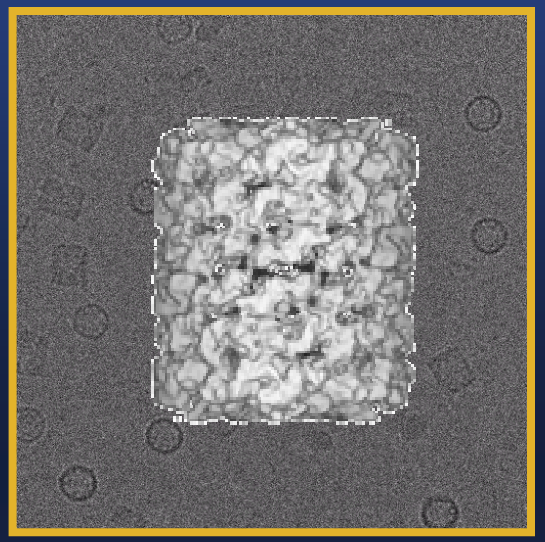


CPMV
images: 300
particles: 14,000
Time: 6 hours

Reconstruction:

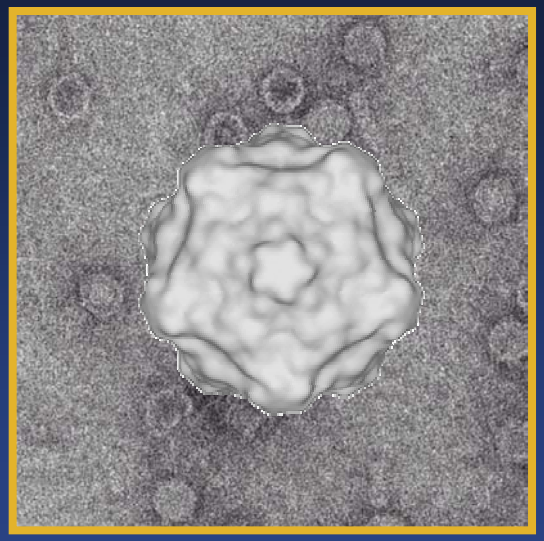
helices

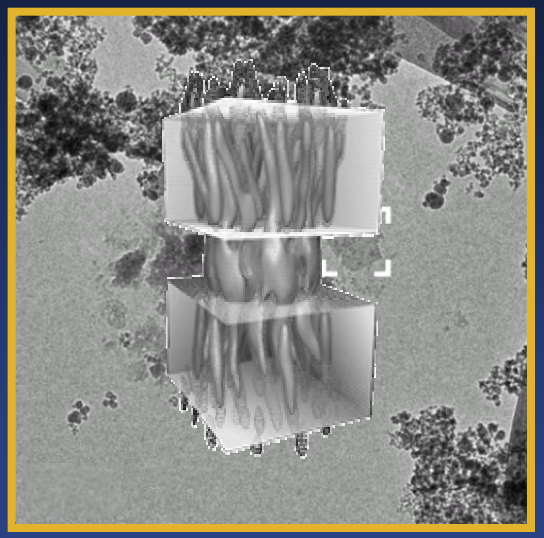




single particles

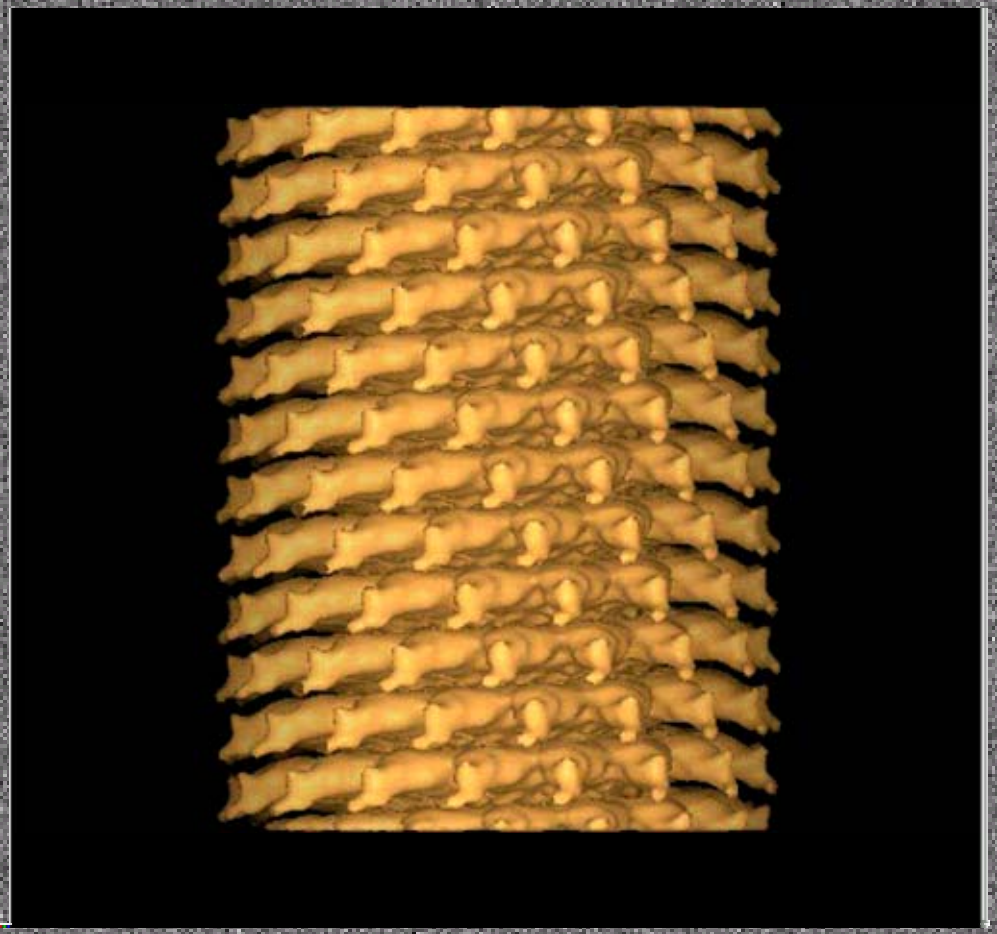
viruses





2D crystals

Reconstruction of TMV
Grid to map: 10Å within 24 hours.



TMV

Leginon for TMV

Calibration and alignment

Very low magnification image acquisition and analysis (60x)

Low magnification image acquisition and analysis (660x)

Intermediate magnification image acquisition and analysis (6600x)

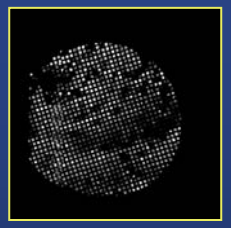
Drift check, focus and astigmatism correction

High magnification image acquisition (66,000x)

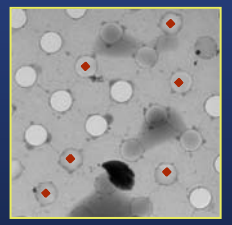
Identification and segmentation of filaments

Helical analysis

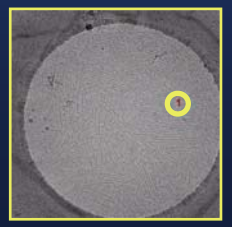
Interpretation of map



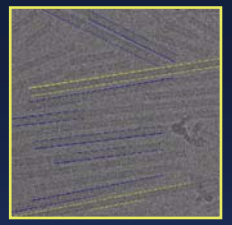
-Survey grid
-Identify potential grid squares
Dose = 0



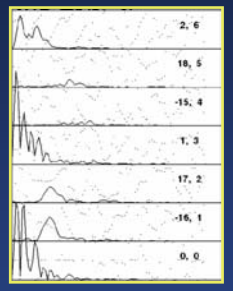
-Identify location of holes
-Estimate ice thickness
Dose = 0.001 e/A^2



-Estimate probability of specimen
-Identify location of best specimen
Dose = 0.1 e/A^2



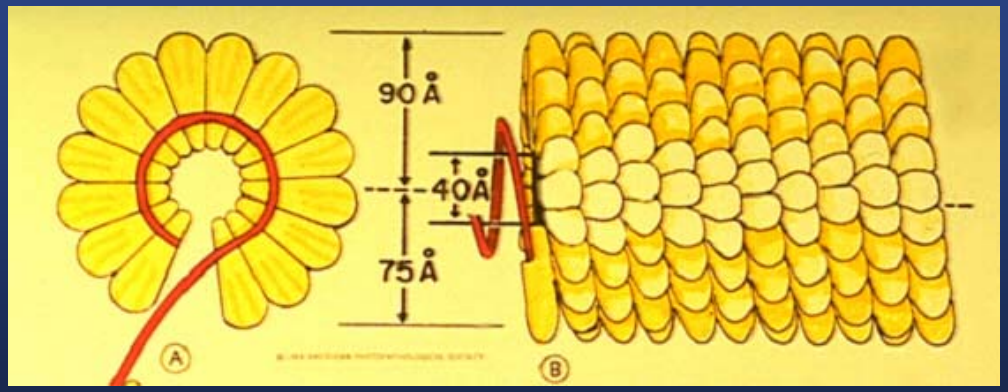
-Identify and segment filaments
Dose = 10 e/A^2

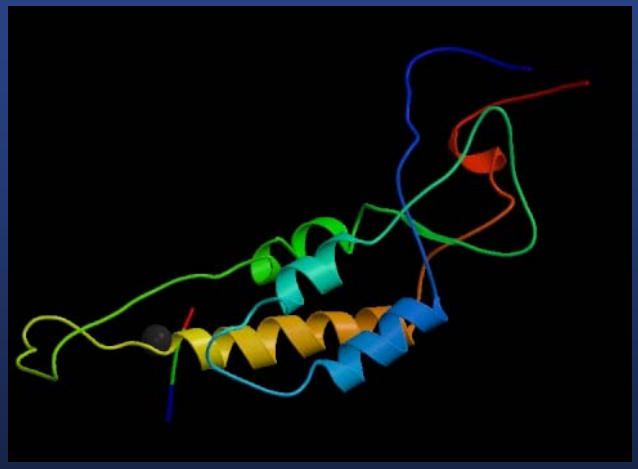


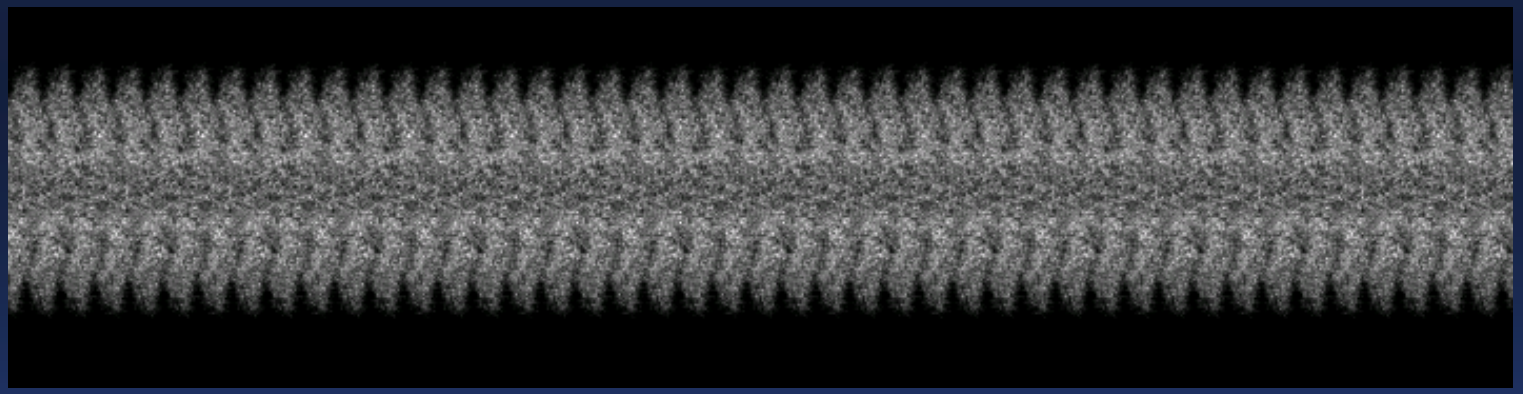
-Find layer lines
-CTF correction
-Tilt and shift correction
-Fitting and averaging

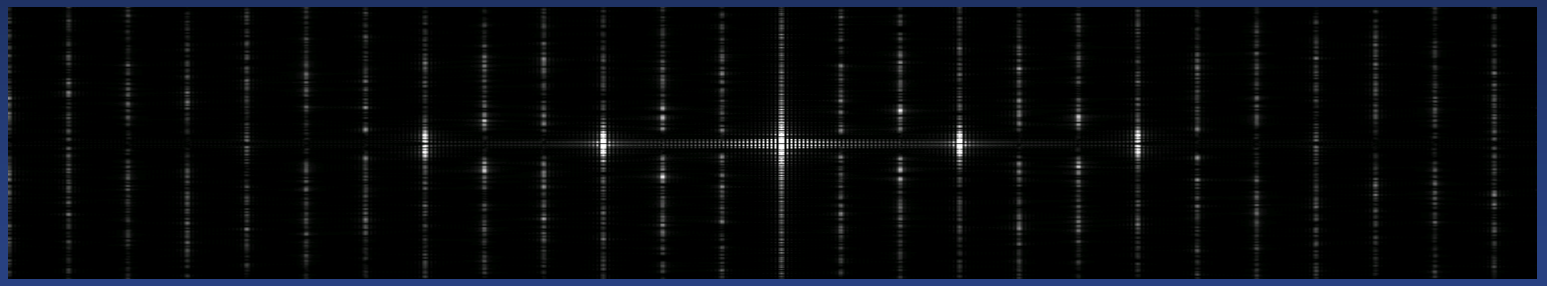
Potter, et al. Ultramicroscopy, 77 (1999) 153-161
Carragher, et al., J. Struct. Biology 132 (2000) 33-45

Why TMV?



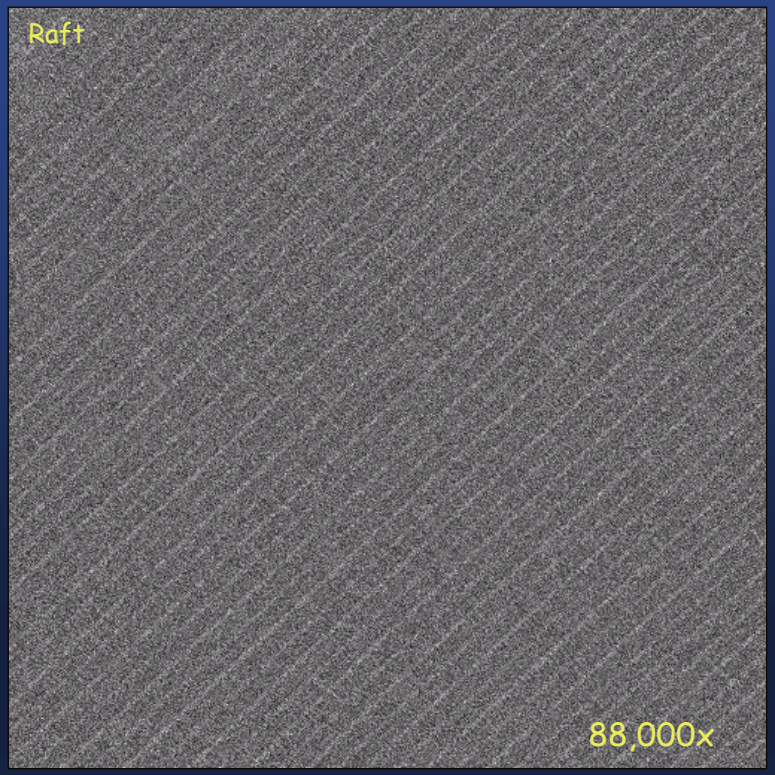




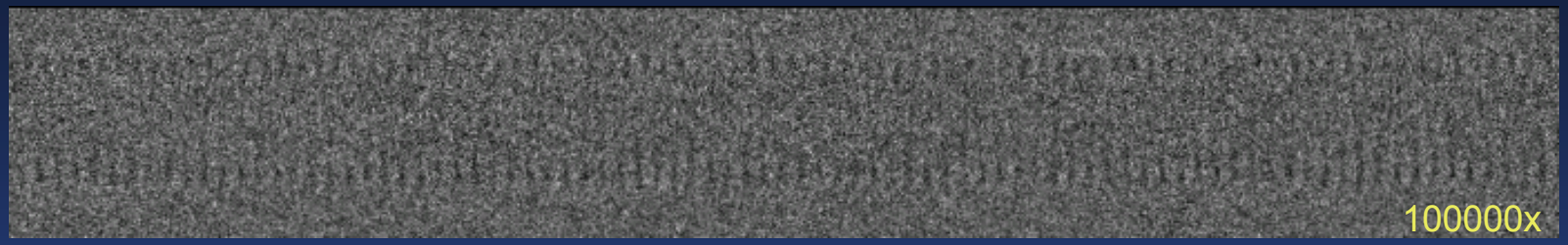


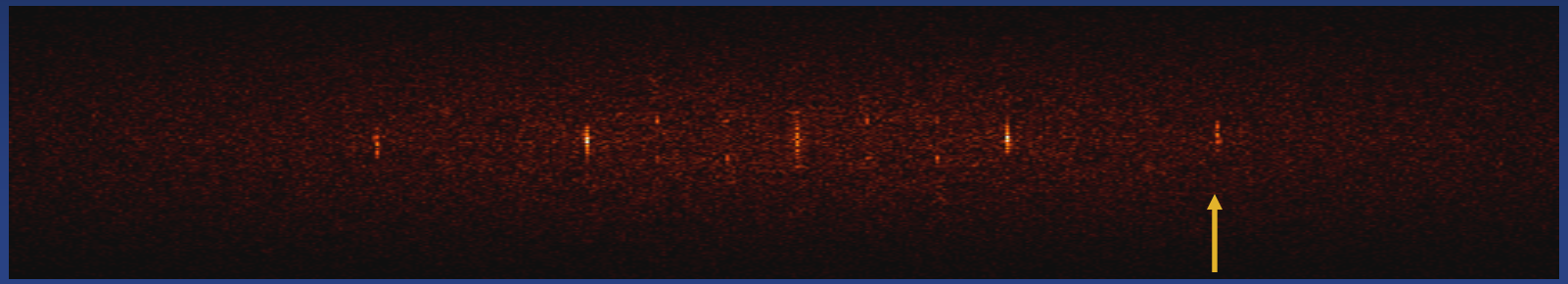
1: 69 Å
2: 34.5 Å
3: 23 Å
4: 17.2 Å
5: 13.8 Å
6: 11.5 Å
9: 7.6 Å
12: 5.8 Å

Layer Line Numbers: 1 2 3 6 9 12

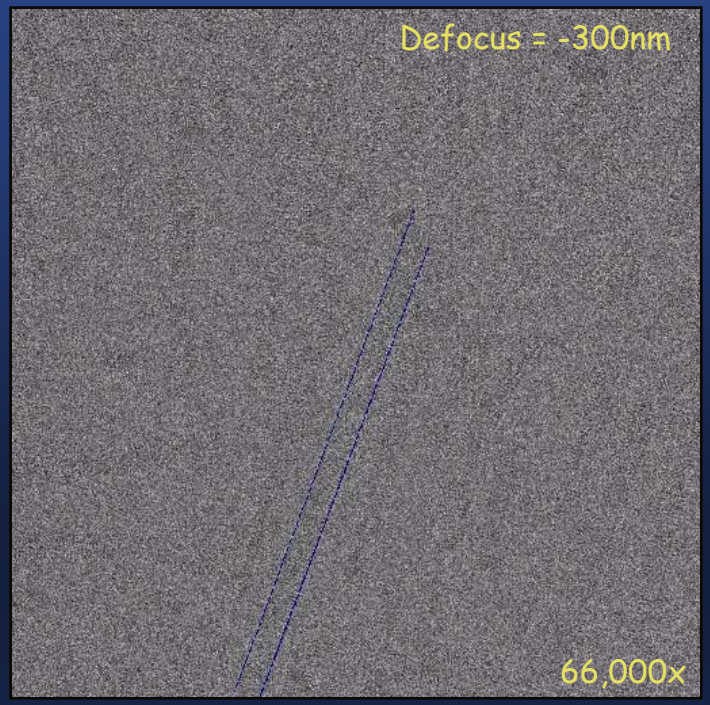

Raft
88,000x

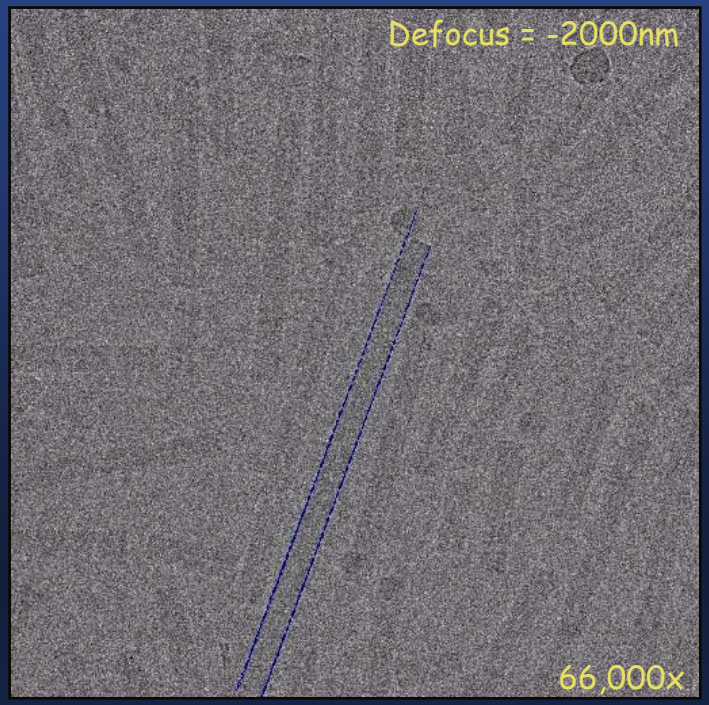
7.6 A

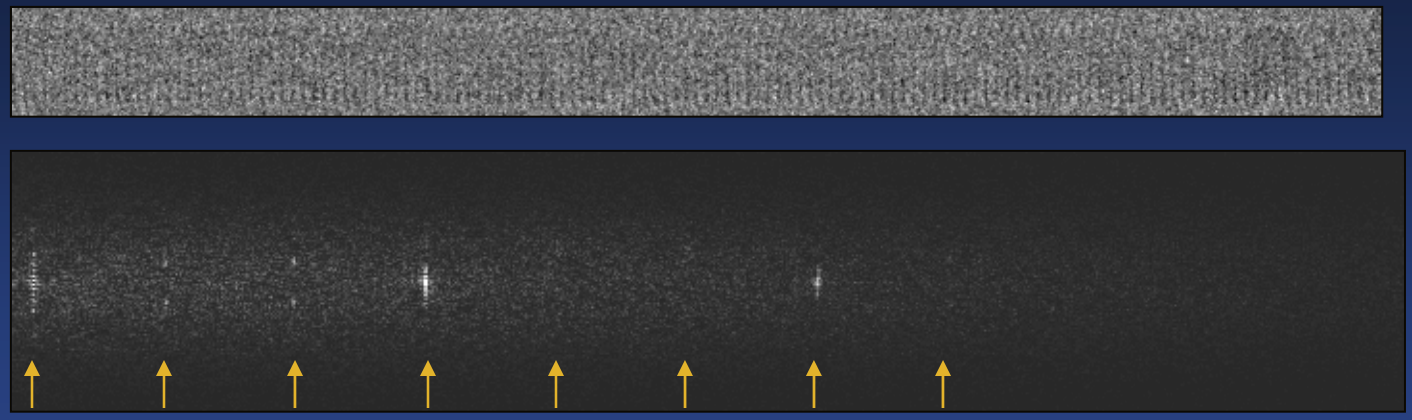

100000x


11.5 A

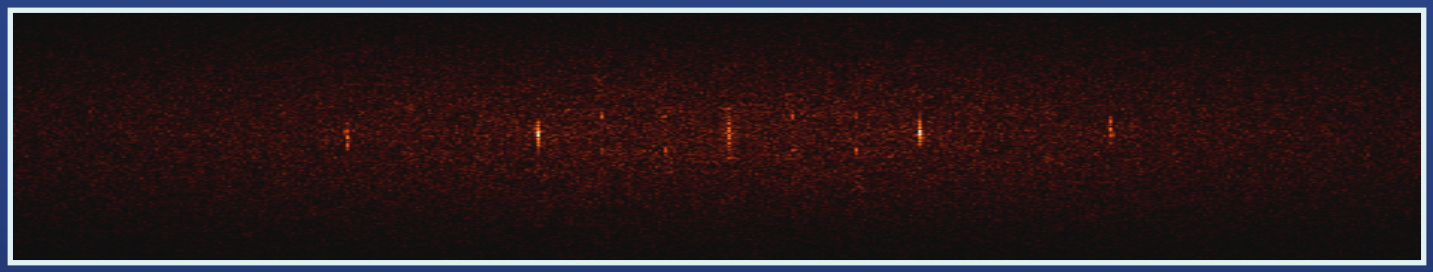
Helical analysis using Phoelix



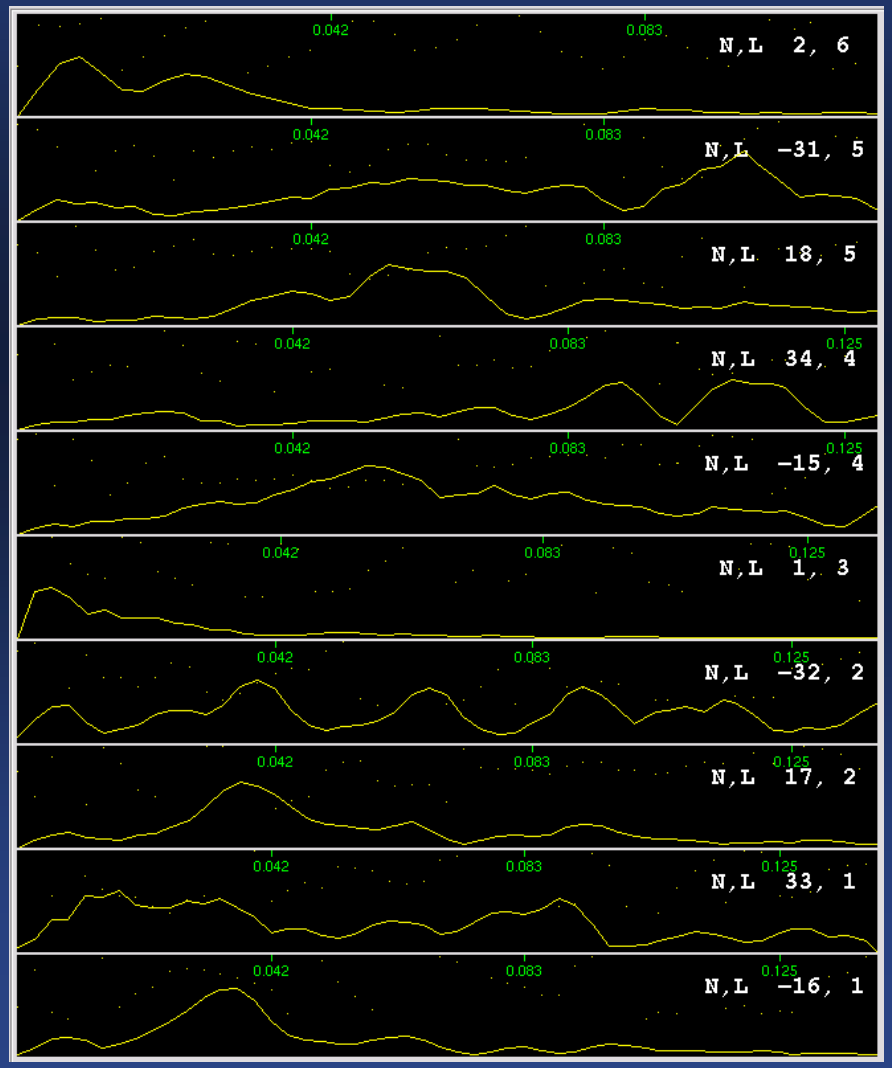


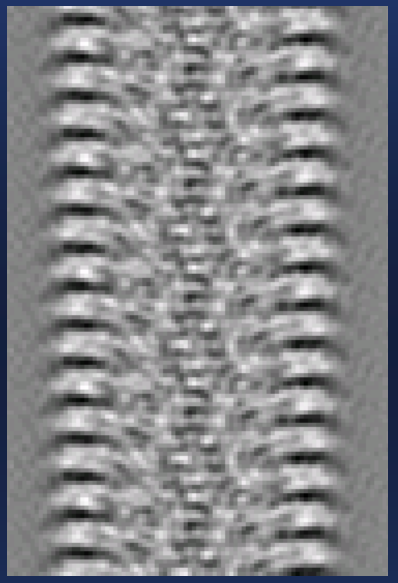


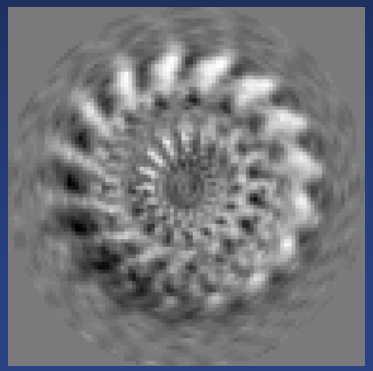
Σ

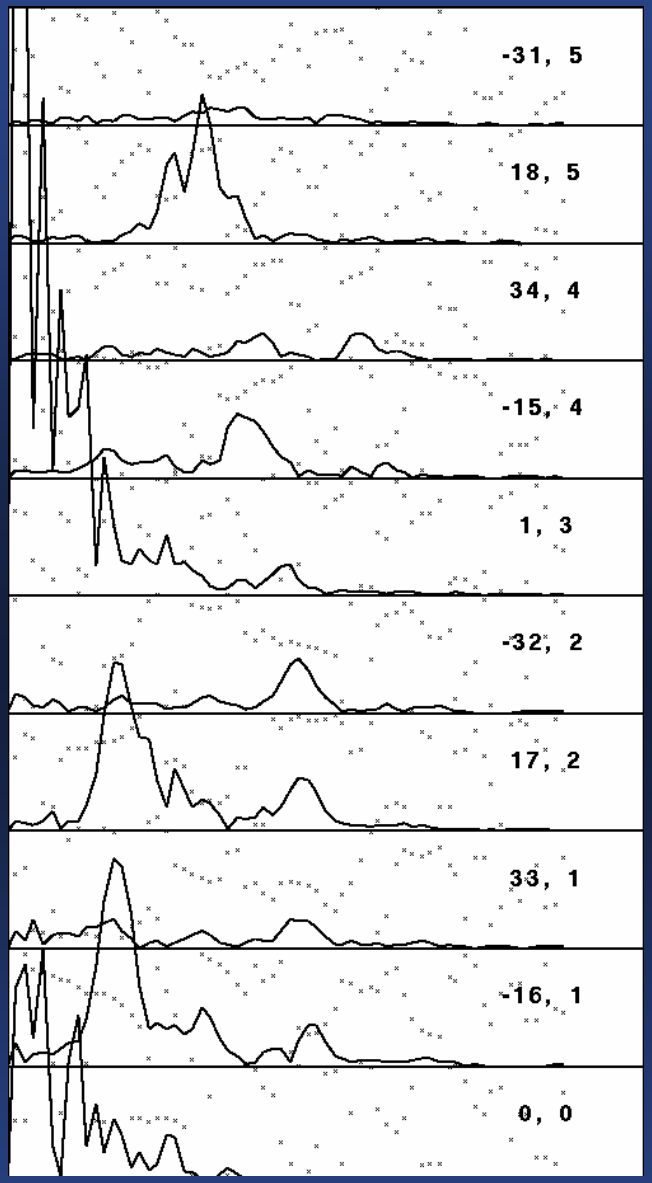


||


0.042
0.083
N,L 2, 6
0.042
0.083
N,L -31, 5
0.042
0.083
N,L 18, 5
0.042
0.083
0.125
N,L 34, 4
0.042
0.083
0.125
N,L -15, 4
0.042
0.083
0.125
N,L 1, 3
0.042
0.083
0.125
N,L -32, 2
0.042
0.083
0.125
N,L 17, 2
0.042
0.083
0.125
N,L 33, 1
0.042
0.083
0.125
N,L -16, 1

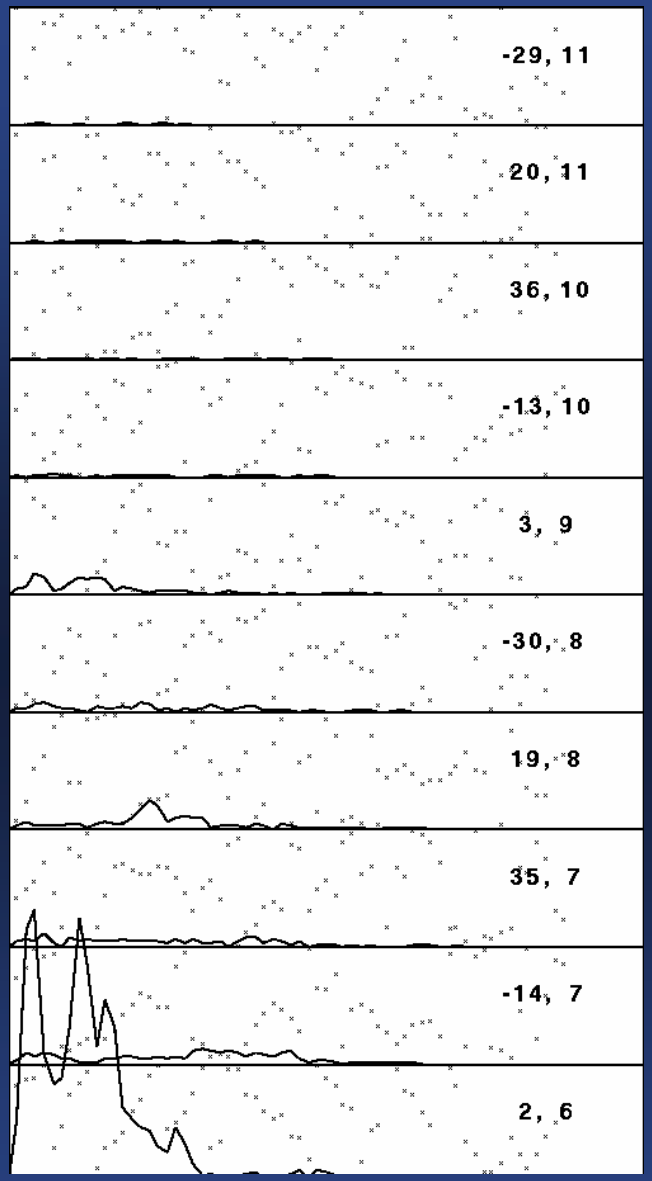






23 Å

69 Å

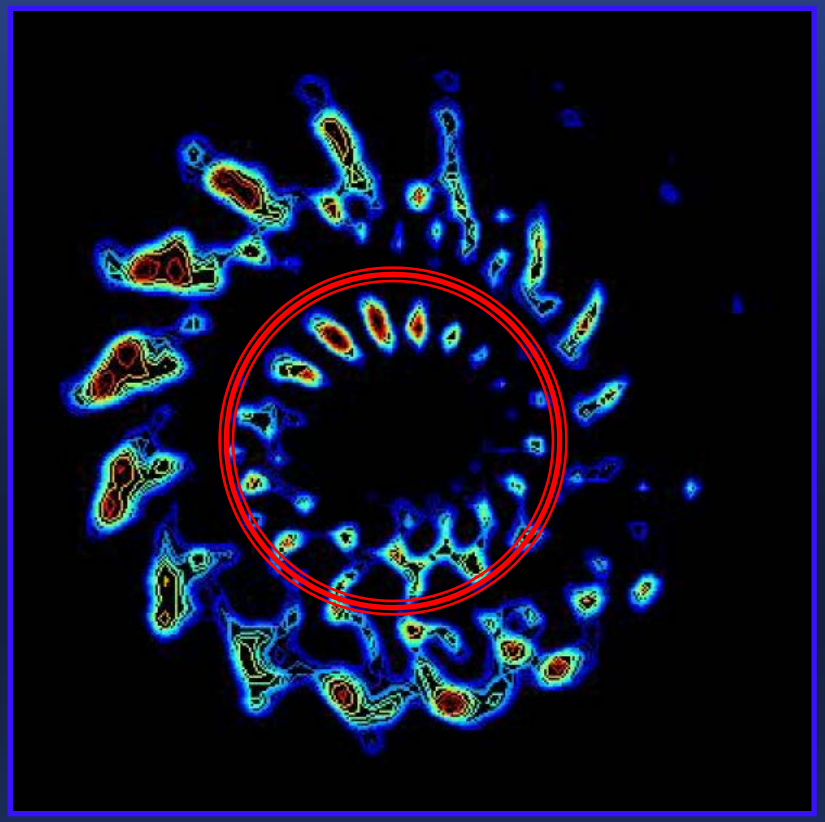


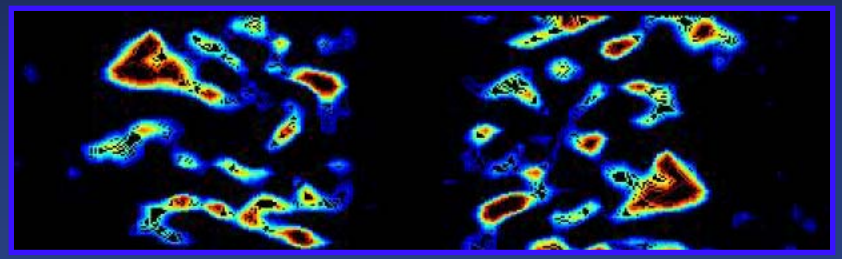
7.6 Å

11.5 Å

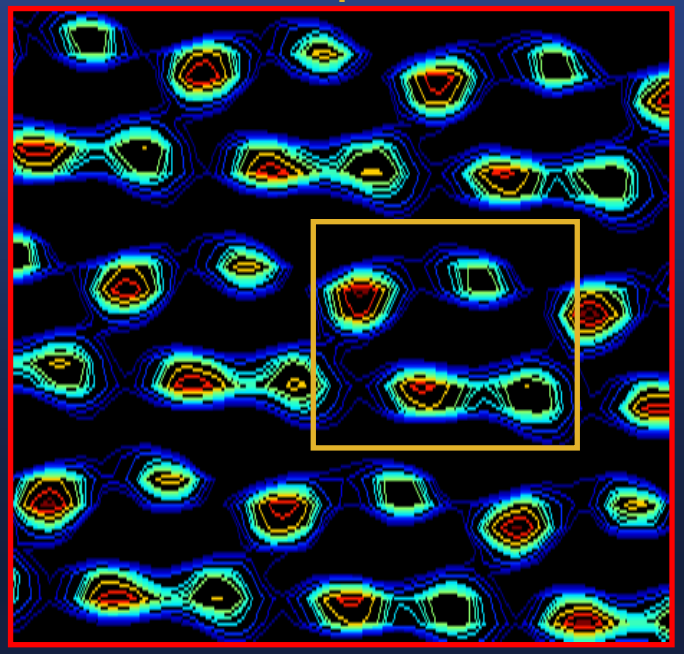
Fourier Shell Coefficient (FSC) for this experiment: 5.8A

Interpretation of the map

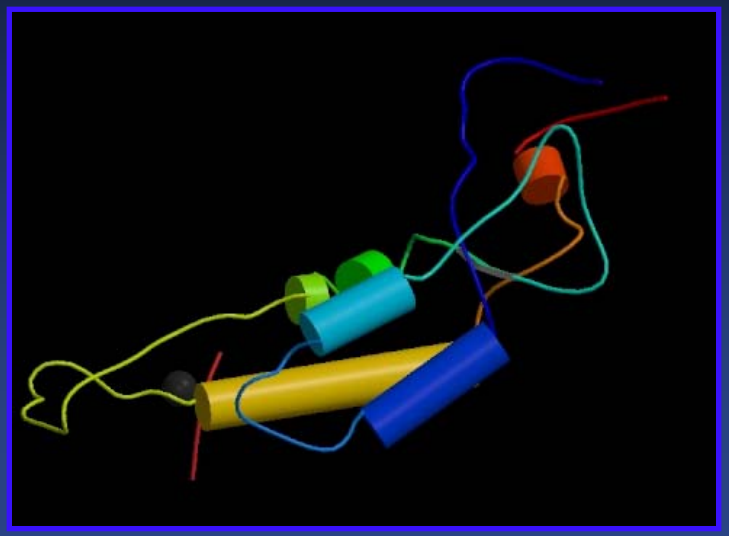




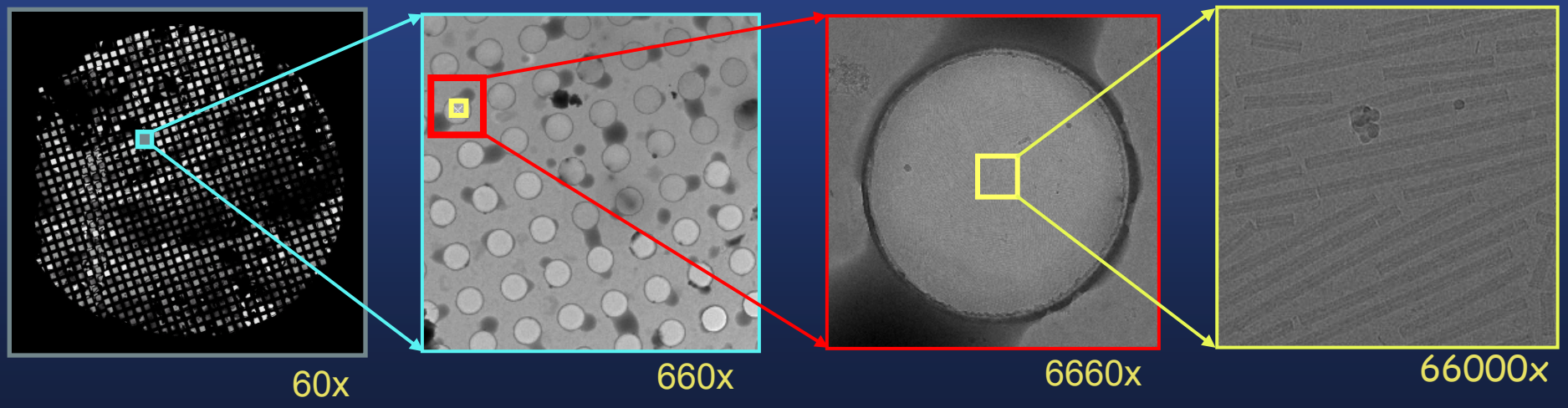
Helical axis



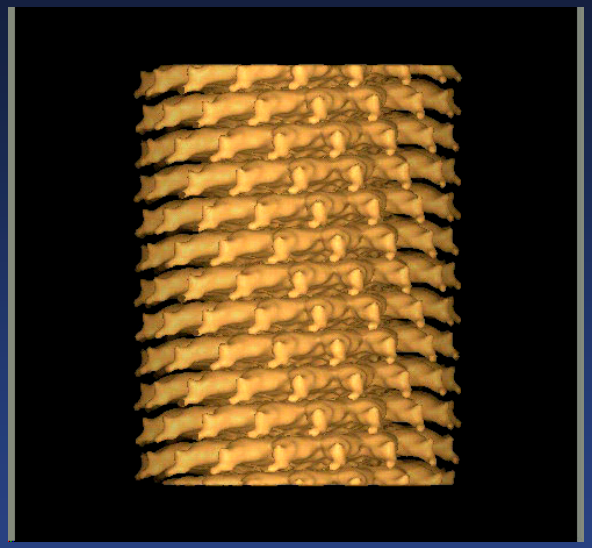
Cylindrical section at radius 66Å



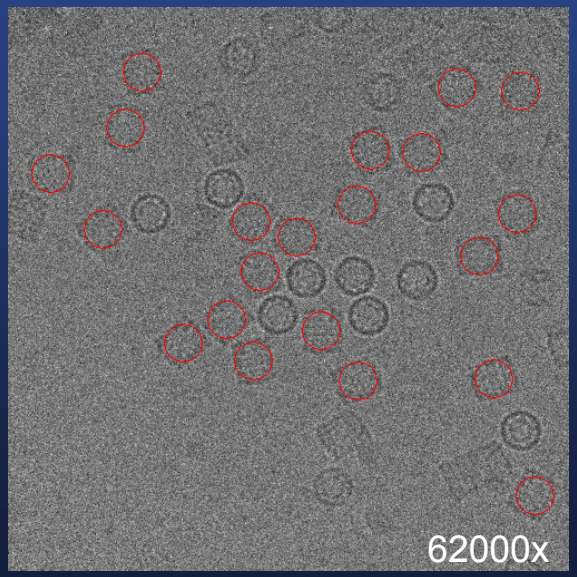
What is the throughput?

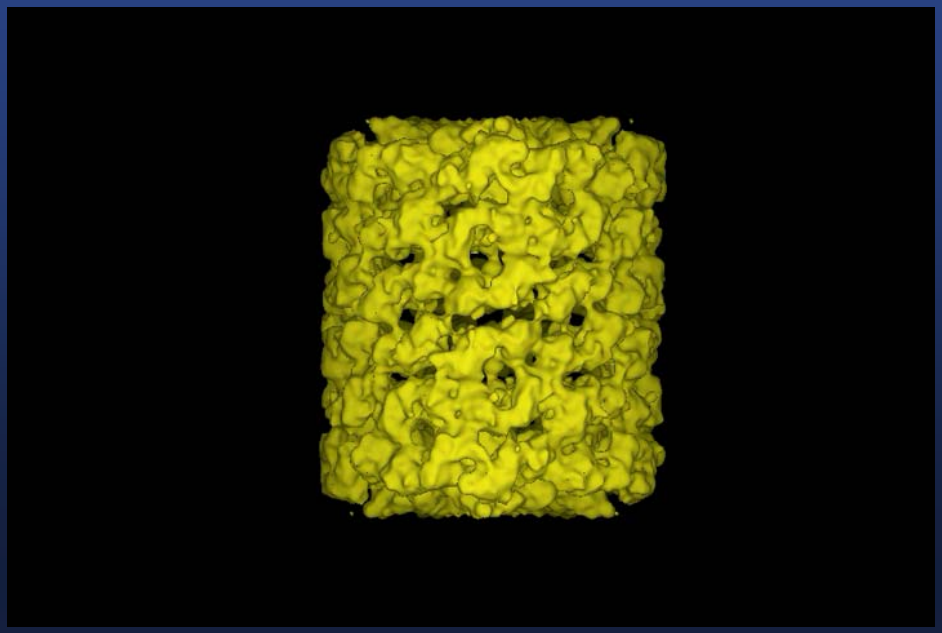


# grids:	1
# squares:	19
# holes:	91
# defocus pairs:	131
# filaments found:	243
Duration:	12 hrs
Phoelix reconstruction	
# filaments used in map:	45
# molecules in map:	~70,000
Resolution of map:	~7.6Å
Yield:	~20%

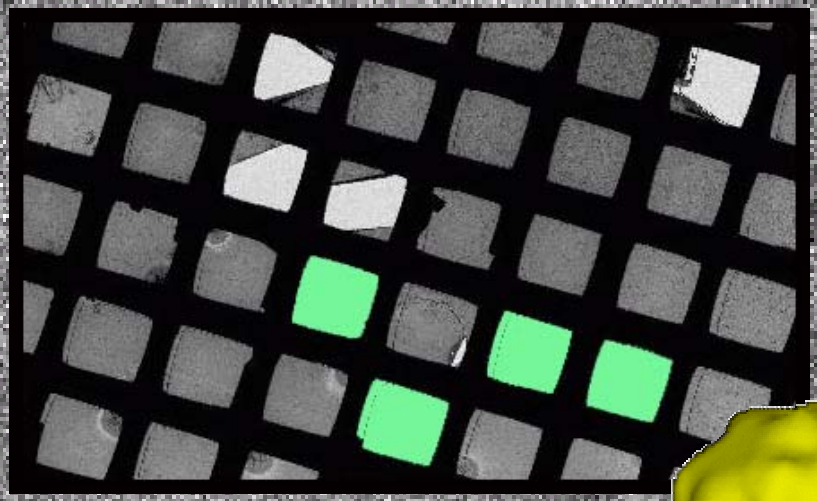


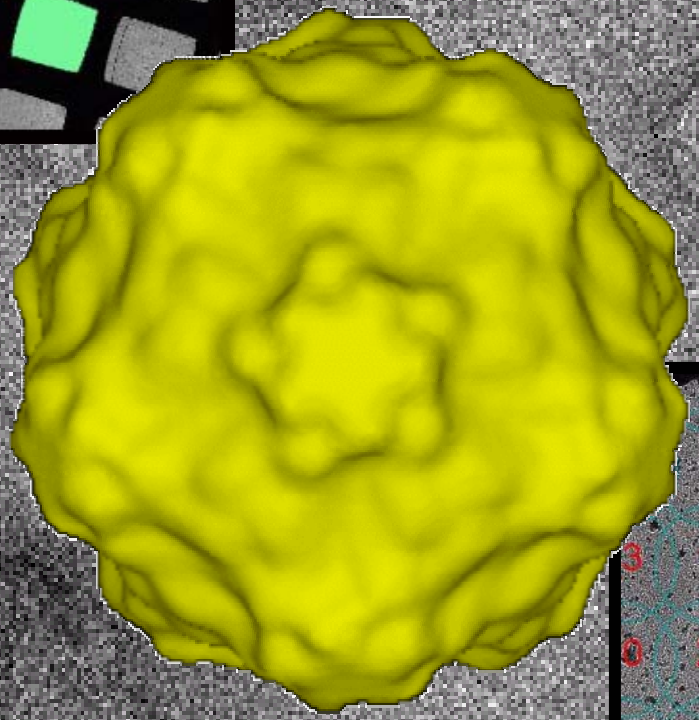
Single particle example: 3D reconstruction of Hemocyanin

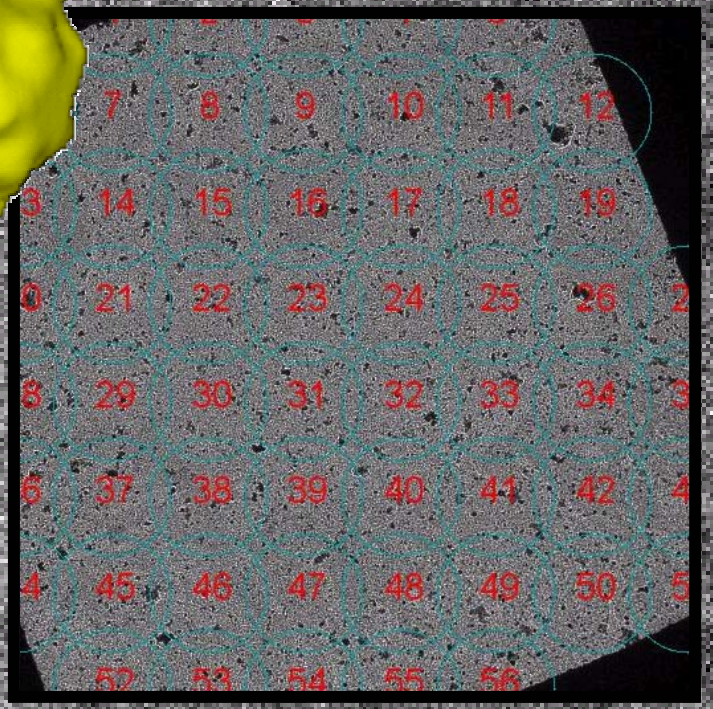




# grids:	1
# squares:	10
# holes:	209
# defocus pairs:	816
# particles found:	21,000
Duration:	28 hrs
Spider reconstruction	
# particles used in map:	~11,000
# molecules in map:	~110,000
Resolution of map:	~12 Å (FSC 3σ), ~15 Å (FSC 0.5)



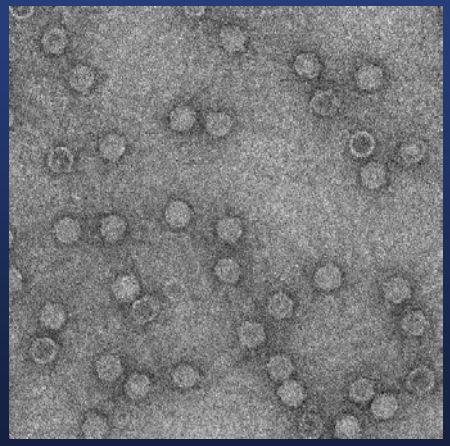




3D reconstruction of CPMV in stain

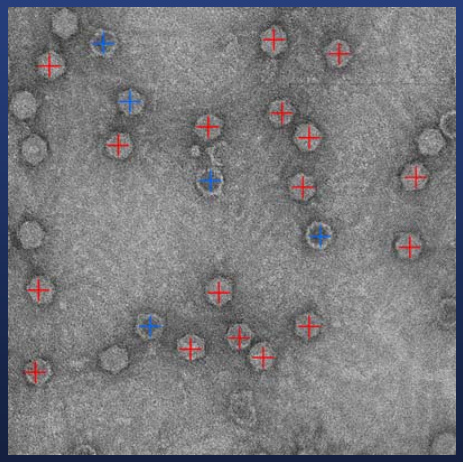
Icosahedral virus example: 3D reconstruction of CPMV in stain

Acquisition



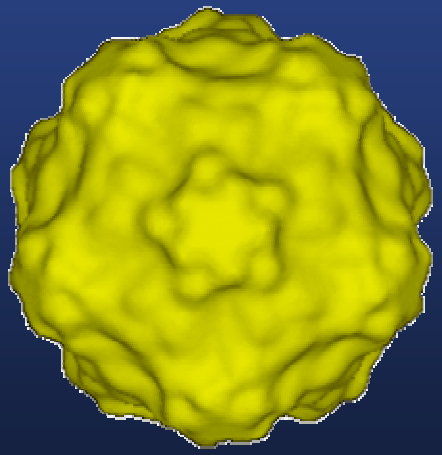
~2,000 particles/hr
(100 images/hr)

Selection



~3,800 particles/hr

Reconstruction



<1 hr/map
(200 particles)

# grids:	1
# squares:	10
# holes:	-
# defocus pairs:	816
# particles found (full):	24,000
Duration:	17 hrs
EMAN reconstruction	
# particles used in map:	~200
# molecules in map:	12,000
Resolution of map:	~40A

What next?

Improve throughput!

Rate

Data acquisition and targeting algorithms

Cryostage stability

CCD cameras

Yield

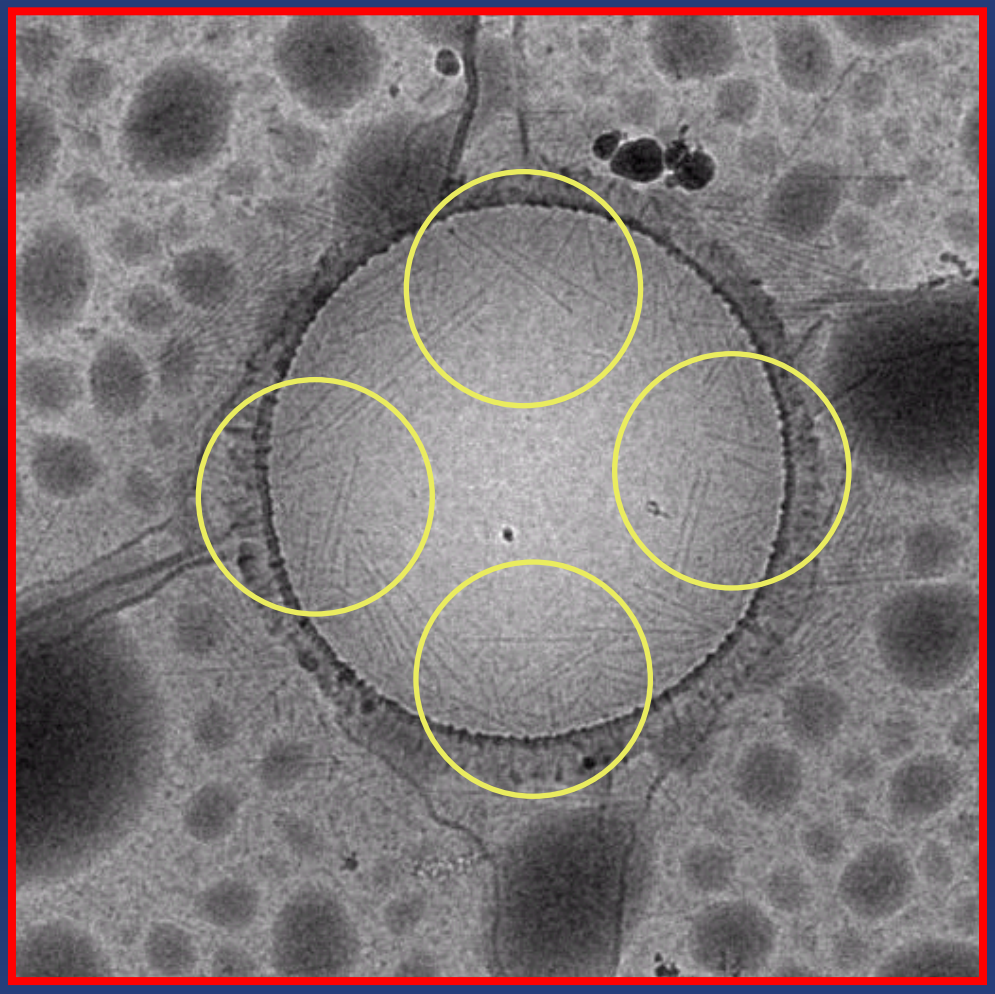
Learning and intelligence (databases)

Sustainability

Automated cryostage filling

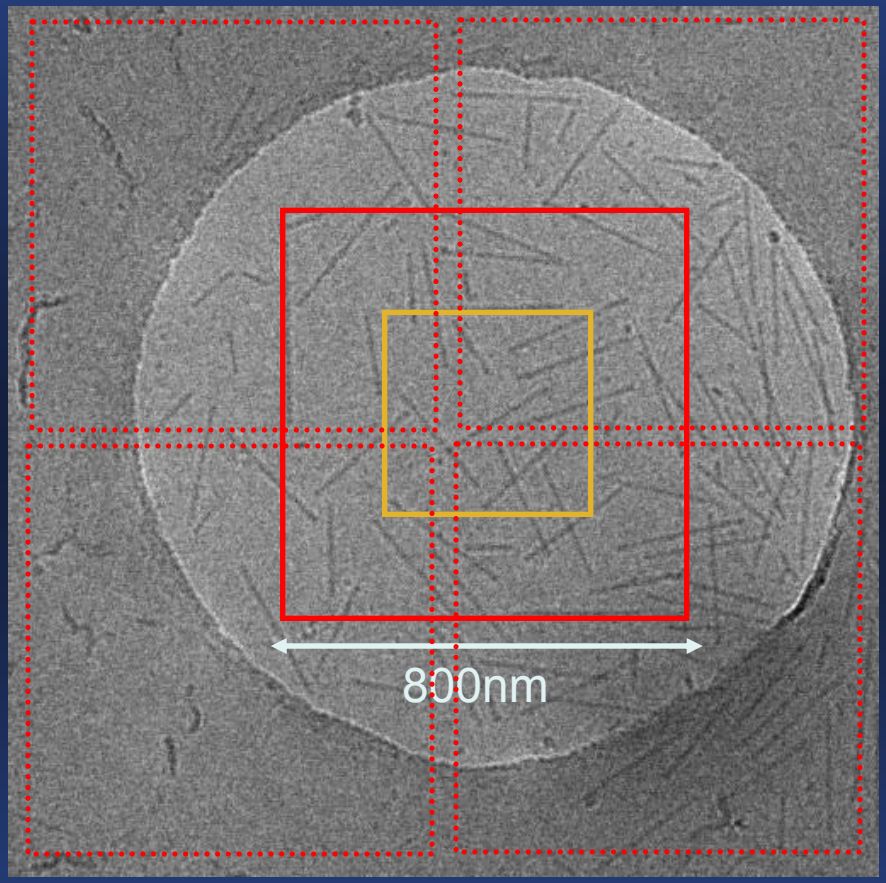
Robotic specimen handling

Throughput: 2K camera



1700x

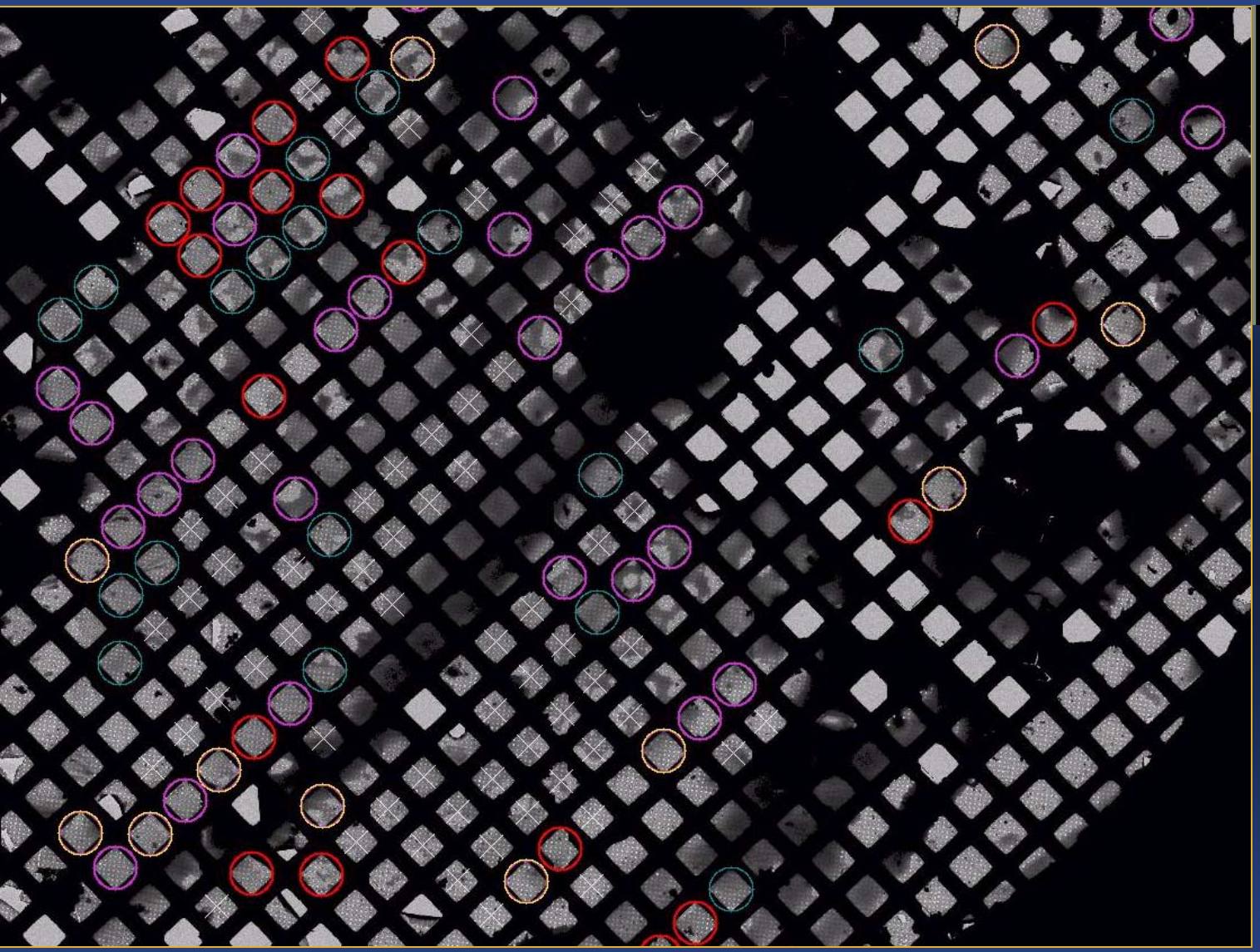
Improving throughput: 4K CCD Camera

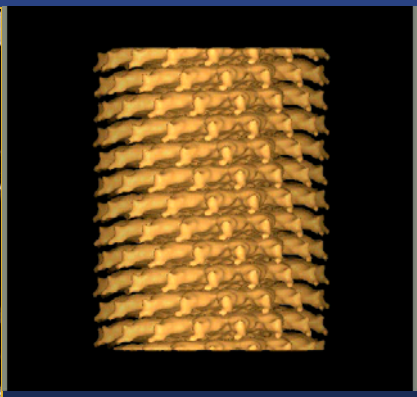


Acquisition area at 50,000x for 4K camera

Acquisition area at 62,000x for 2K camera

Improving yield:





X Selected squares (167)
Holes found (90)
Filaments found (66)
LL identified (33)
Fitted into average (20)

filaments used in final map: ~100
squares used in final map: 20
holes used in final map: 32

Improving sustainability:

Automated cryostage replenishment prototype

Automated cryostage replenishment prototype

Improving sustainability: Robotic specimen handling

Why use a robot to handle specimens?

Screening trials

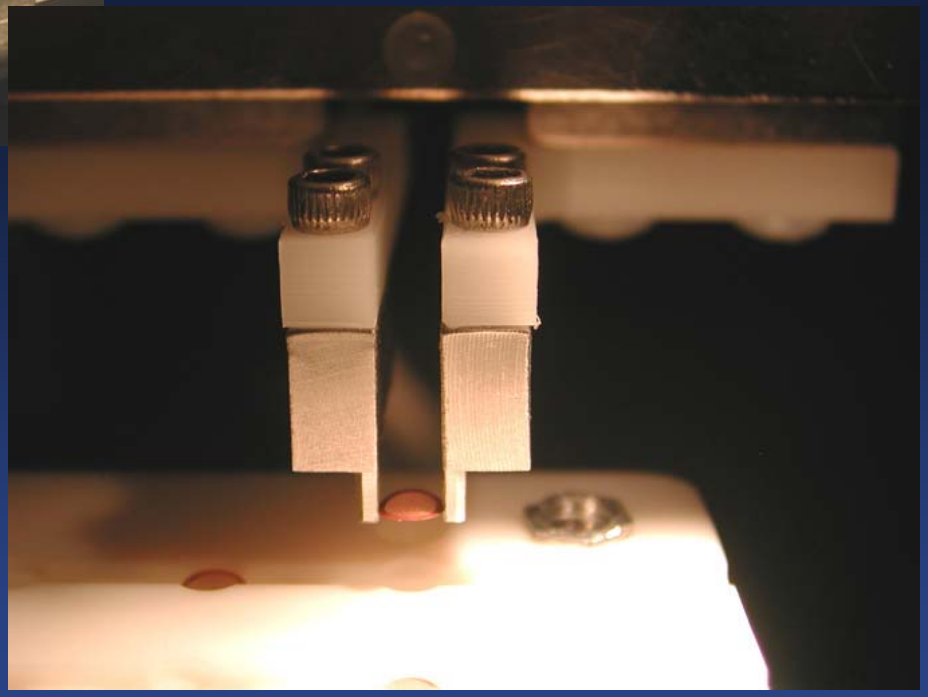
Virus characterization

…

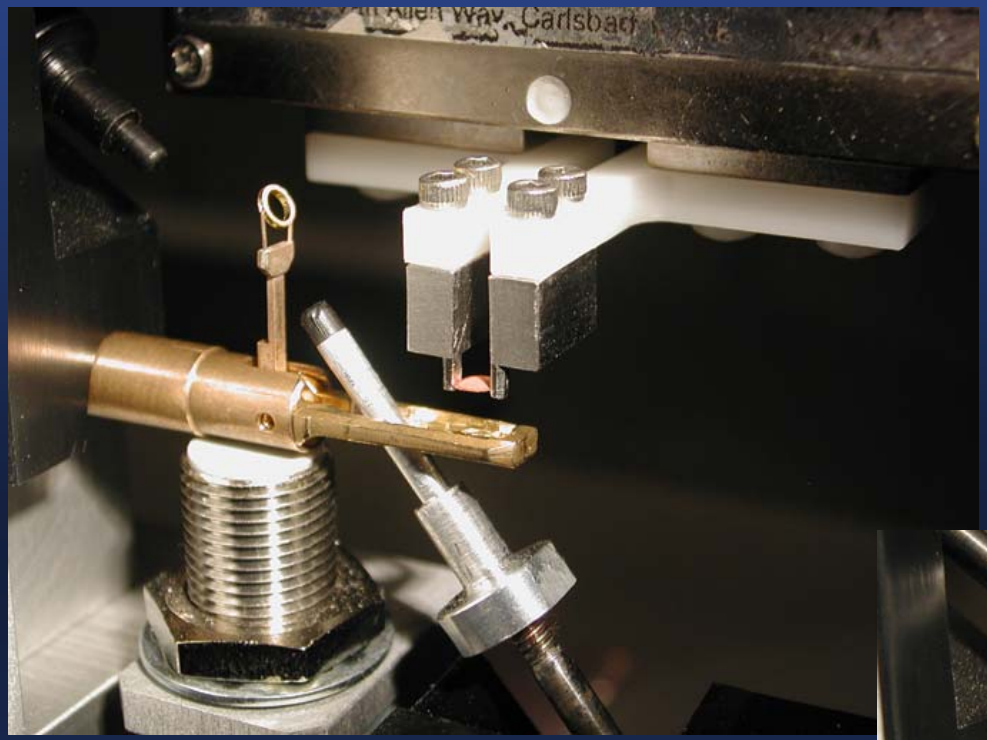
Throughput: 10’s of grids per day

Stain not ice!

Robotic specimen handling

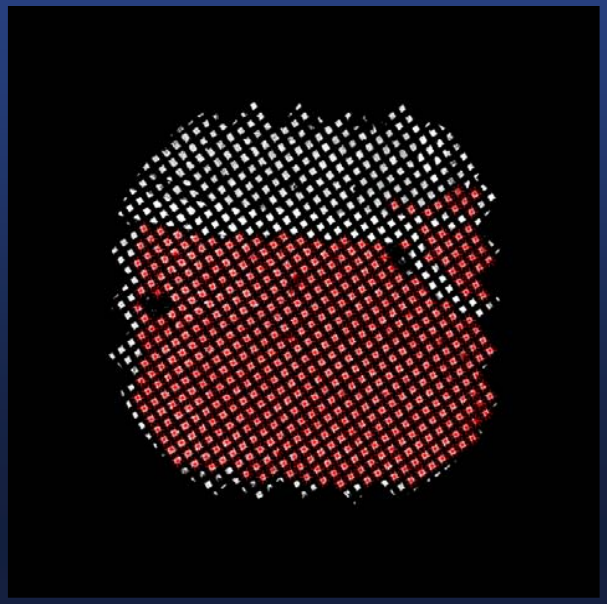


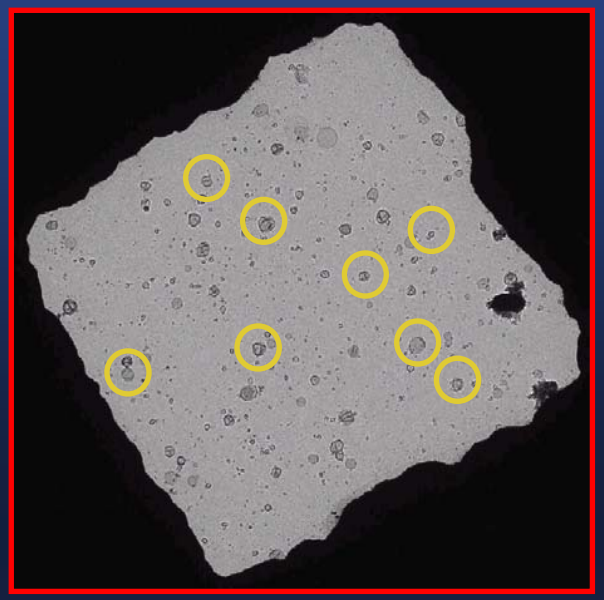
Robotic specimen handling

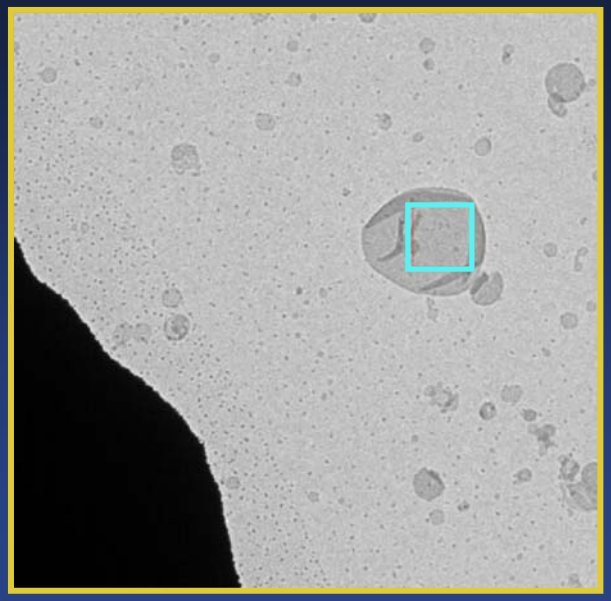


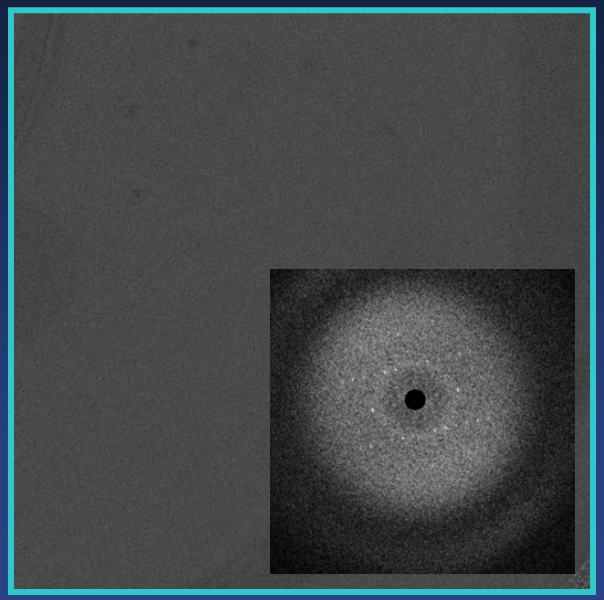
Robotic specimen handling

Automation goal: Screening trials
Example: Rhodopsin crystals



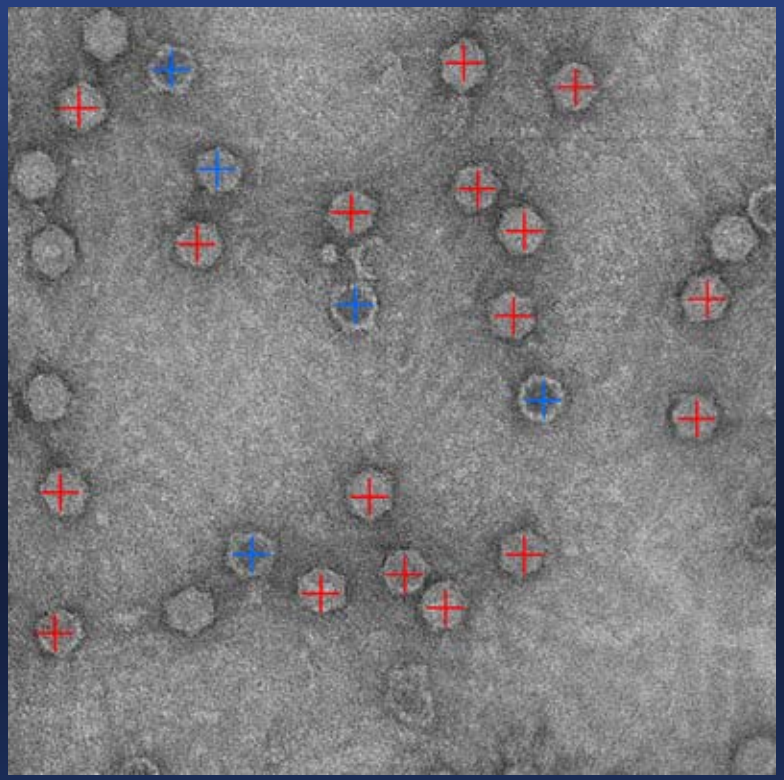




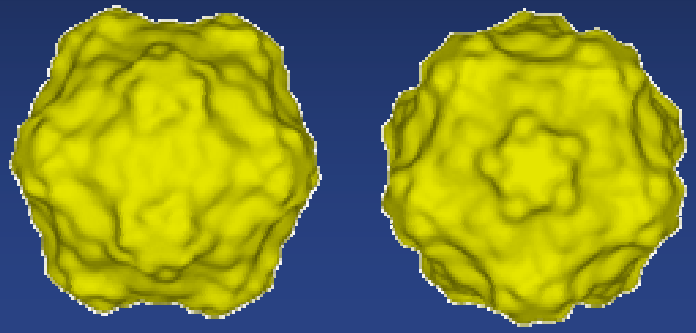


Courtesy of
Yoshi Fujiyoshi

Automation goal: Virus characterization
Example: CPMV



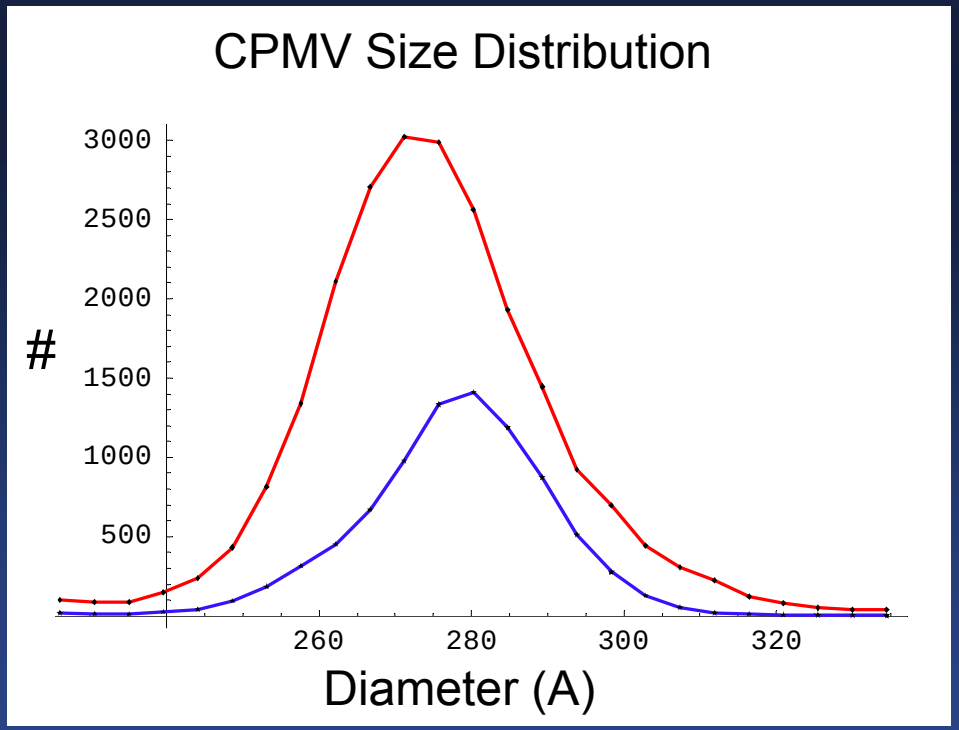
~1000's particles/hour



200 particles per map

CPMV Classifier

Full	23,277	(63%)
Empty	8,678	(24%)
Junk	4,664	(13%)



NRAMM

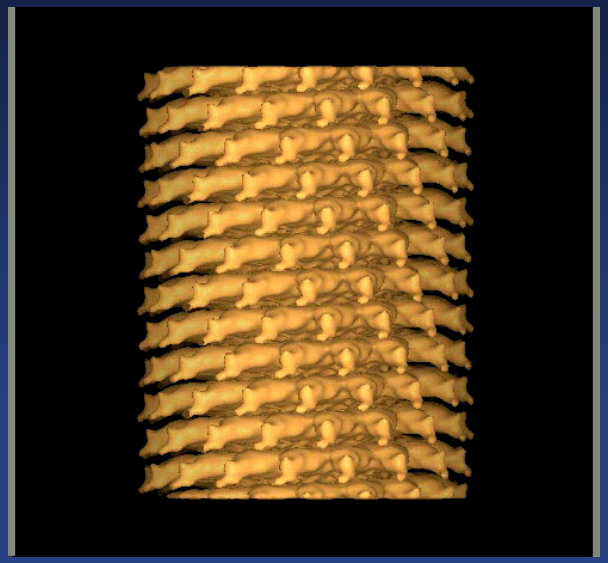
National Resource for Automated Molecular Microscopy

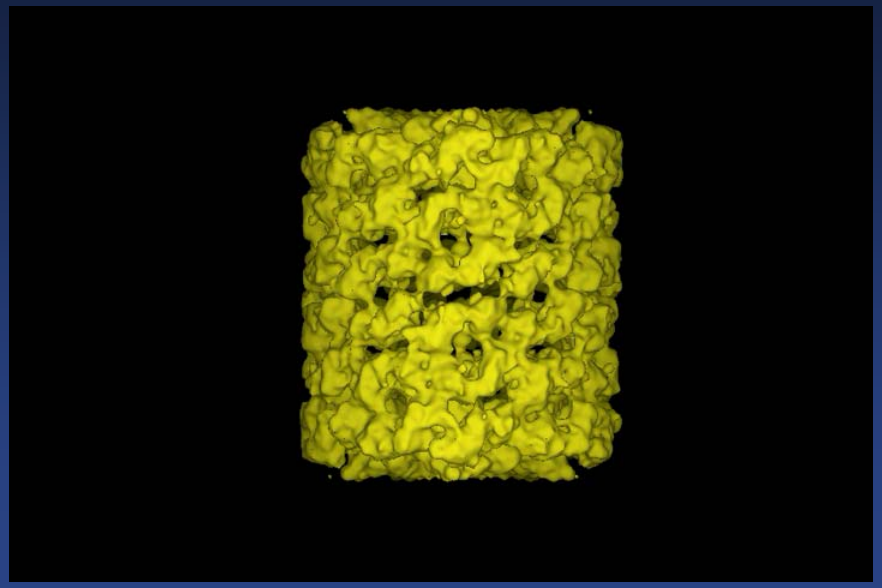
Core Technology Projects:

- Specimen handling
- Automated image acquisition
- Automated processing
- Information handling

Automation goals:

- 7-15 Å resolution
- high throughput
- mainstream





NRAMM

National Resource for Automated Molecular Microscopy

National Center for
Research Resources

A Practical Course in Molecular Microscopy

November 12-20, 2003

Center for Integrative Molecular Biosciences (CIMBio)
The Scripps Research Institute (TSRI), La Jolla, California

http://nramm.scripps.edu

AMI / NRAMM Team:

Jim Pulokas

Denis Fellmann

Christian Suloway

Yuanxin Zhu

Francisco Guerra

Joel Quispe

Anchi Cheng

Scott Stagg

Brian Sheehan

Cynthia Ona

Bridget Carragher

Clint Potter

Support:

NCRR (RR 17573)

National Institutes of Health (GM61939)

National Science Foundation (DBI-9730056, DBI-0296063)