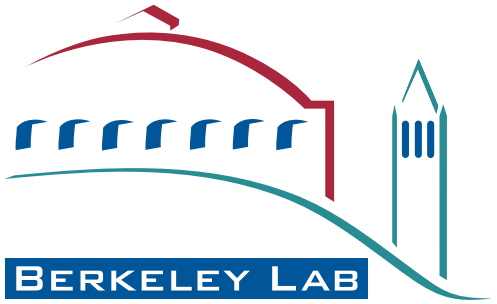
HIGH-THROUGHPUT IMAGING OF MULTI-PROTEIN COMPLEXES

INNOVATIONS IN NANOSCALE MODELING AND IMAGING OF BIOLOGICAL SYSTEMS

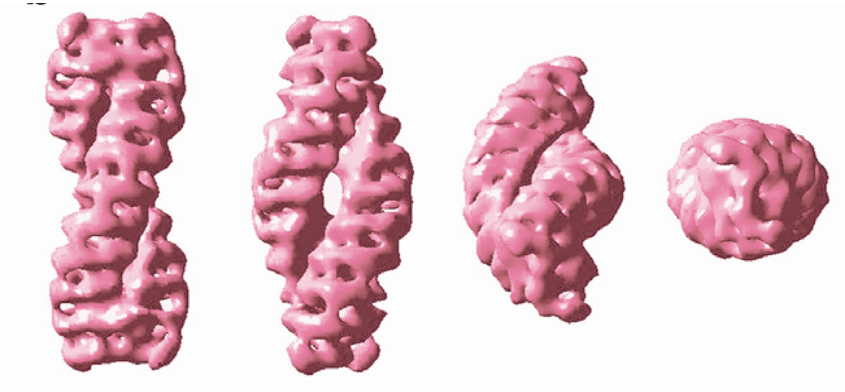
HOUSTON, APRIL 17-22, 2006



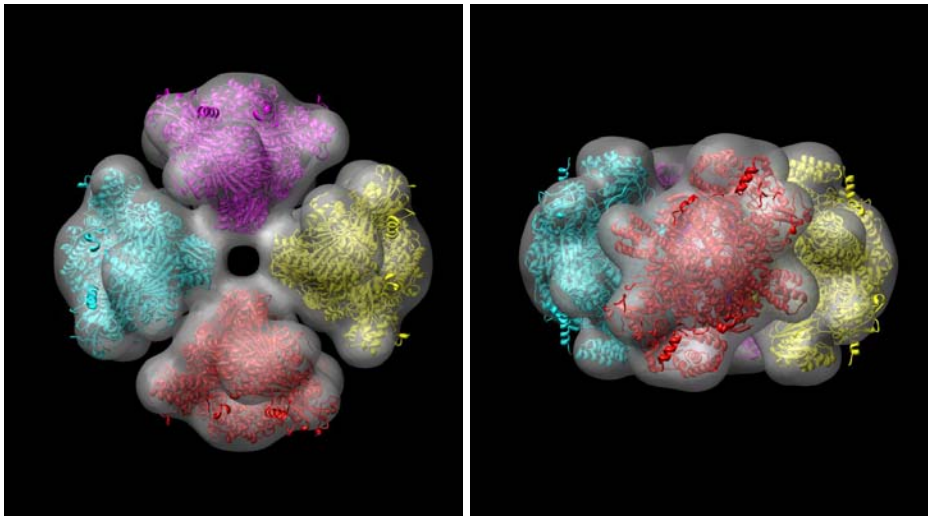
ROBERT M. GLAESER

SINGLE-PARTICLE EM IS AN EXCELLENT WAY TO CHARACTERIZE THE STRUCTURE OF A MULTIPROTEIN COMPLEX

- **Tripeptidylpeptidase II is a "giant protease with a twist"**
  - **Rockel et al. (2005) PNAS 102:10135-10140**



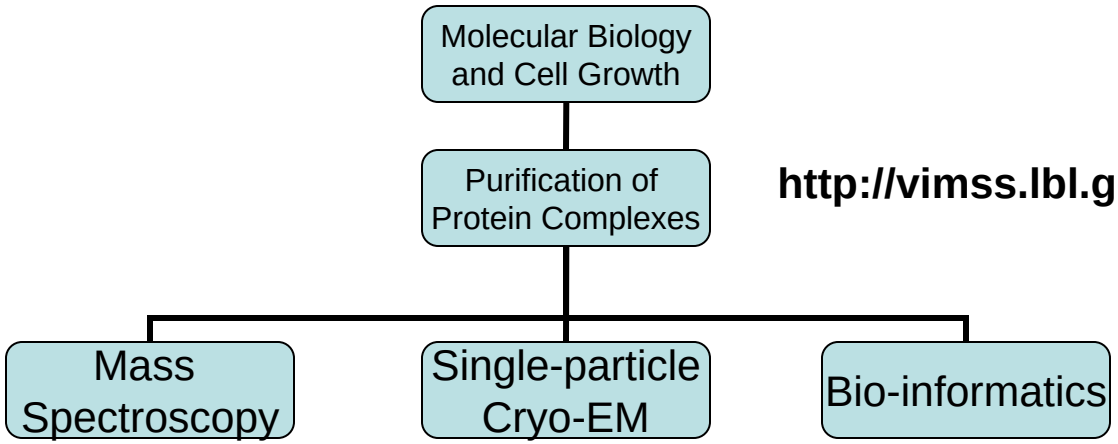
- **Pyruvate-ferredoxin oxidoreductase is an octamer in *Desulfovibrio vulgaris* but a dimer in *Desulfovibrio africanus***
  - **X-ray structure of the dimer, "docked" into the 2nm reconstruction of the octamer**



Florian Garczarek et al. (2006) Unpublished

AN APPLICATION THAT REQUIRES HIGH THROUGHPUT:

Dept. of Energy's PROTEIN COMPLEX ANALYSIS PROJECT (PCAP) AT LBNL (Mark Biggin, PI)

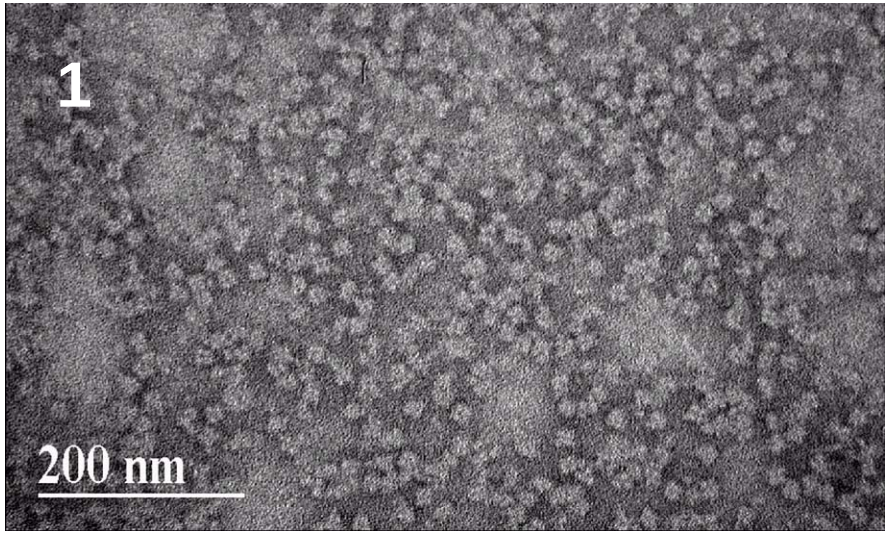


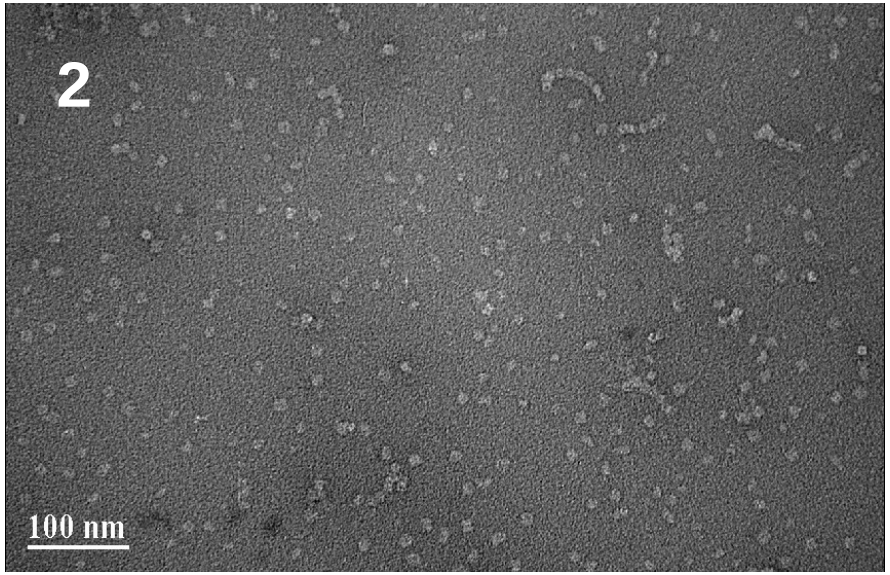
http://vimss.lbl.gov/projects/pcap.html

Systems-biology characterization and imaging of microbial multi-protein complexes for organisms capable of reducing metals and radionuclides

Understand the role of all multi-protein complexes in the microorganism's response to a variety of stresses in their environment

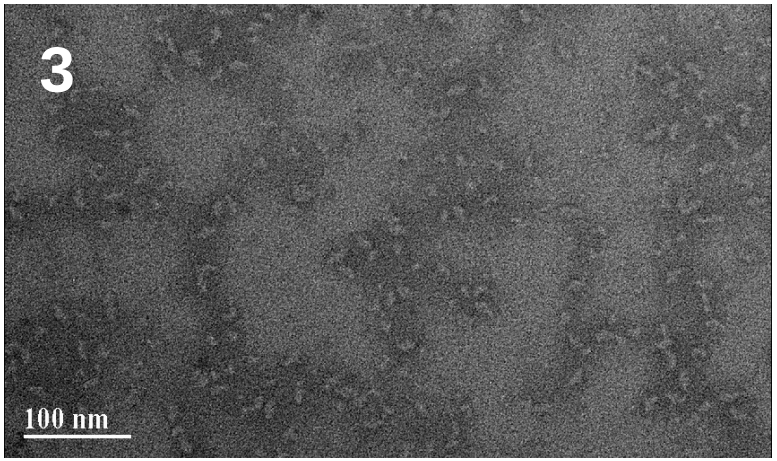
DECIDING WHETHER SAMPLES MIGHT GIVE GOOD MAPS IS ALREADY AS EASY AS RUNNING A GEL





- **Prepare samples with 3 negative stains on glow-discharged grids (<45 min)**
- **Evaluate whether the sample is well dispersed and homogeneous in at least one of the negative stains (~45 min each)**

1. **Pyruvate:ferredoxin oxidoreductase (PFOR);**
   $M_r > 10^6$ Da – good prospect
2. **Putative protein;**
   M_r ~ 400 kDa – maybe it would work
3. **Phosphoenol pyruvate synthetase (peps);**
   M_r ~ 400 kDa – poor prospect, but try it



GETTING 3-D MAPS SUITABLE FOR "DOCKING" SHOULD ALSO BE AS EASY AND FAST AS RUNNING A GEL

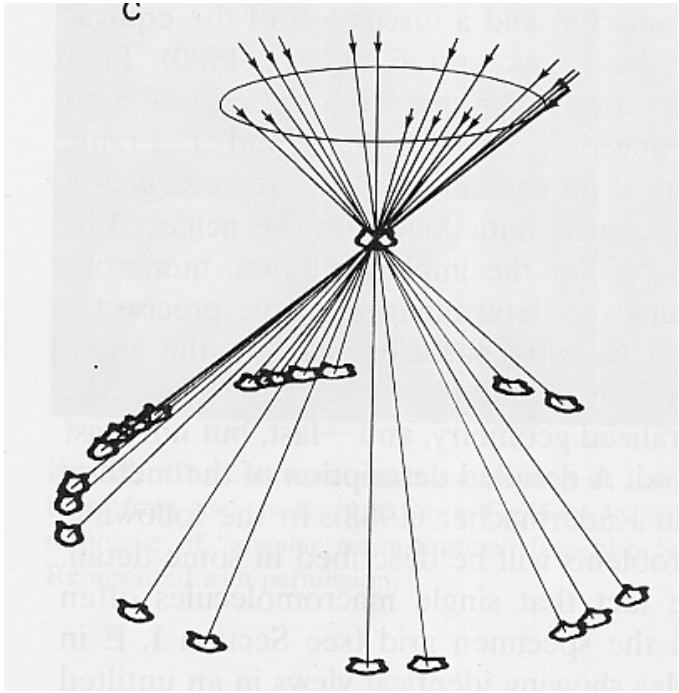
For this to be possible, the following must be in place:

- **Determining whether a preparation of purified protein is suitable for getting a 3-D reconstruction must be fast and efficient – already in place**

1. **Obtaining an *initial model* suitable for refinement must be fast and fool-proof**
2. ***Data collection* must be fast and easy – automated microscope operation can help a lot**
3. ***Automated particle-boxing* can also speed up the process (James Chen)**
4. ***Computation* may need to be done with large-scale, highly parallelized clusters**

1. Obtaining an initial model suitable for refinement must be fast and fool-proof

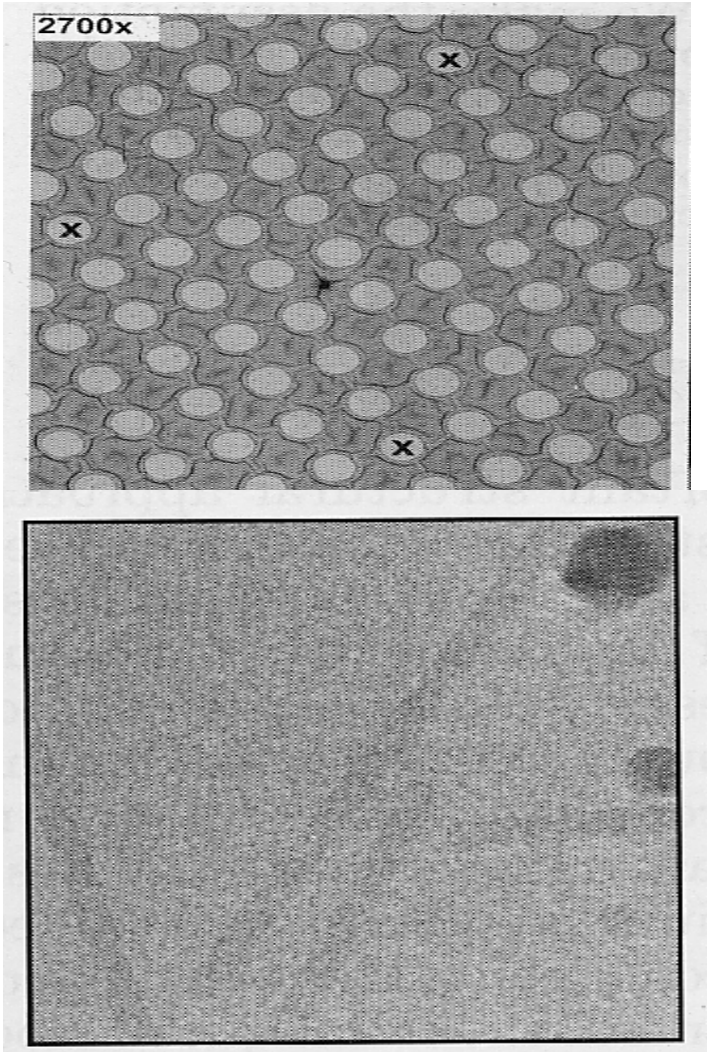
- **Currently used methods are rather slow and *may* give distorted and uncertain results**
  - **Reconstructions may suffer from a large missing cone**
  - **Unreliable in the face of structural heterogeneity**
  - **Therefore need careful "user intervention" to succeed**
- **New methods based on symmetrical tilt pairs**
  - **Merge data from image pairs, not just "the tilted image" in a pair**
  - **Obtain full angular coverage – i.e. no missing cone of information (for example, see Leschziner et al., JSB)**



Random conical tilt method uses images from untilted specimens to identify (classify) particles with the same orientation, and then uses images from tilted grids to get many views, as above (J. Frank, 1996)

2. AUTOMATED IMAGING CAN MAKE DATA COLLECTION EASY, ONCE IT IS KNOWN THAT THE BIOCHEMICAL PREPARATION IS GOOD

- **Software packages have been described**
  - **Carragher & Potter**
  - **Zhang et al. (Subramaniam)**
- **A 3-D reconstruction of Tobacco Mosaic Virus has been produced at a resolution of <1 nm in less than 24 hours**
  - **Zhu et al (2001) JSB 135:302-312**
- **Images of 100,000 GroEL particles have been collected in a single day**
  - **Suloway et al. (2005) JSB 151:41-60**

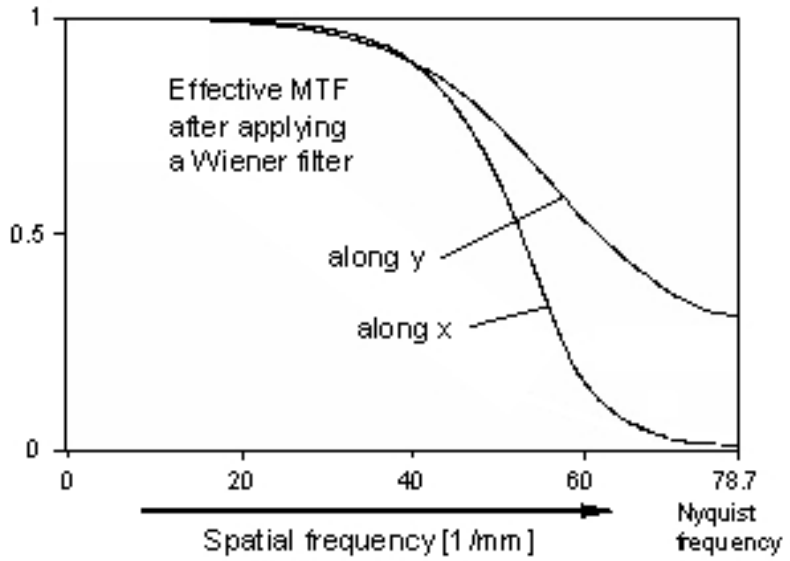


FILM GIVES 20X MORE PARTICLES PER IMAGE THAN ARE AVAILABLE FROM CCD IMAGES

- **The proper comparison of alternative detectors should consider**
  - **Detective Quantum Efficiency**
    - **Hard to exceed the value of ~0.7 for film**
  - **Define the area of the detector as the number of pixels, not the number of square centimeters**
    - **When sampling the image intensity at a frequency where the detector MTF > 0.5 (for example)**
- **The comparison depends considerably on the accelerating voltage of the EM**
  - **CCD cameras perform best at lower voltages**
- **Even at 100 kV**
  - **Best CCDs give < 2kx2k pixels (i.e. even when the chip is 4kx4k)**
  - **Film is ~7kx10k (10kx14k raw pixels) when scanned with the Nikon densitometer**
  - **Film is even more advantageous for the high-end, 200 kV and 300 kV microscopes**

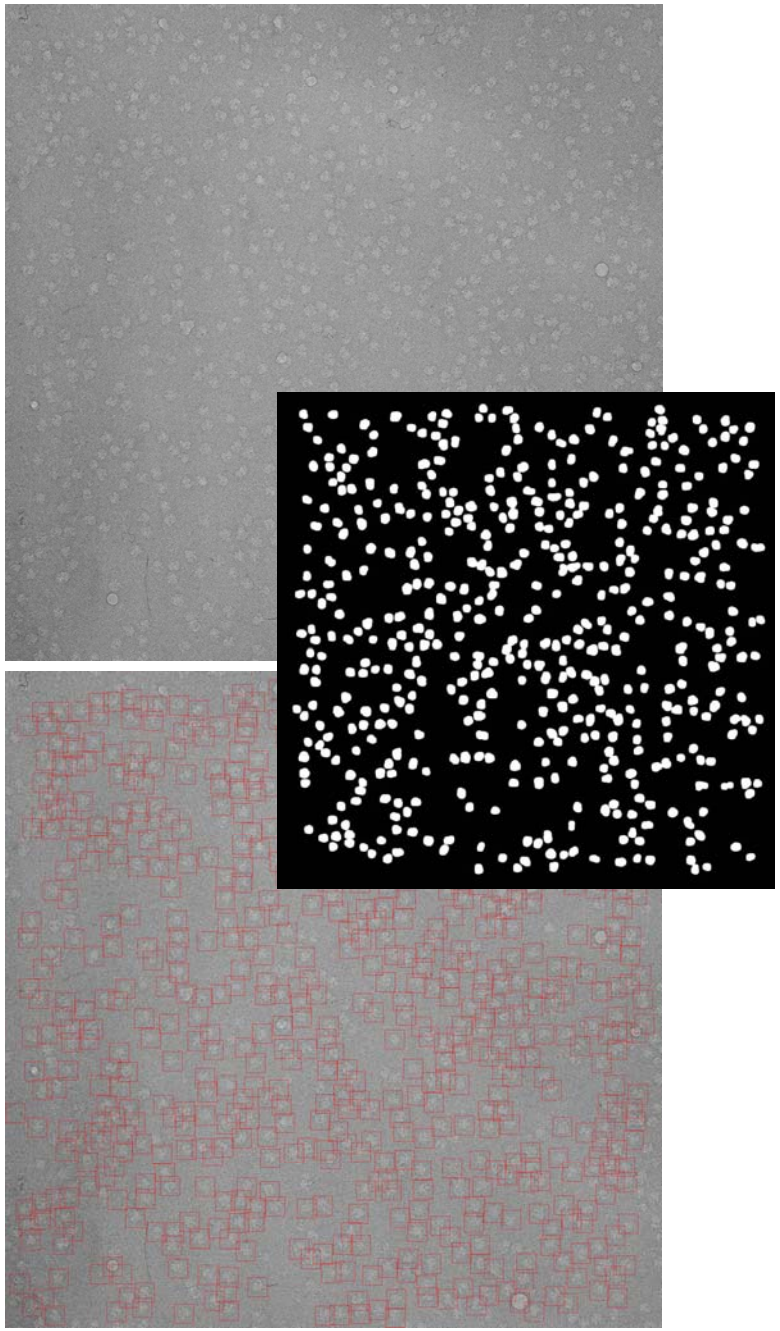
AUTOMATION OF FILM-DENSITOMETRY OVERCOMES THE INITIAL DISADVANTAGE OF FILM vs DIRECT ELECTRONIC READOUT

- **Simple "robot" devices**
  - **replace EM films on the holder**
  - **Insert and remove the holder**
- **>300 films per day; 750 films in a stack**
- **Nikon densitometer produces images with 10kX14k pixels**
  - **6.35 μm/pixel**
- **Wiener filtering restores the MTF to >0.5 at 2/3 Nyquist**



3. AUTOMATED PARTICLE-PICKING CAN MAKE DATA ANALYSIS EASY

- This remains a field that needs development
- One approach is to perform careful binarization (Adiga)
  - Requires histogram standardization and background leveling
  - Removal of false positives based on size, shape, even template matching
- Other, “more standard” approaches are based on cross-correlation
  - Example, FindEM (Roseman)
  - Variations on “matched filtering” (Baldwin)



4a. EXTENDING THE STRUCTURE ANALYSIS TO HIGH RESOLUTION REQUIRES VERY LARGE DATA SETS

- **3-D reconstructions in the range 0.8 nm – 1.2 nm resolution require ~ 50,000 to 100,000 particles**
  - **Experience with 2-D crystals, icosahedral viruses, and ribosomes**
  - **Getting even higher resolution from ~50,000 particles requires an unbiased way to select just the "cream of the crop" from a MUCH larger data set**
- **3-D reconstructions in the range 0.3 nm – 0.4 nm resolution require > 10^6 particles**
  - **Experience with 2-D crystals and physical estimates by Henderson (1995) and by Glaeser (1999)**
  - **Particles must also be large (a few M Da) and conformationally homogeneous**

Short Digression

SUCH A LARGE NUMBER OF PARTICLES IS *NOT* REQUIRED BY BASIC PHYSICS

- **The inelastic scattering cross section does limit the electron exposure to ~2000 e/nm^2 for high resolution at high voltage**
- **Image contrast (and thus the signal-to-noise ratio for the electron exposure above) is limited by the elastic scattering cross section**

In spite of these limitations,

IF IMAGES WERE PHYSICALLY PERFECT:

- **Atomic resolution (0.3 nm) would be possible for particles as small as 40 kDa**
- **Data sets could be smaller than 15,000 particles**

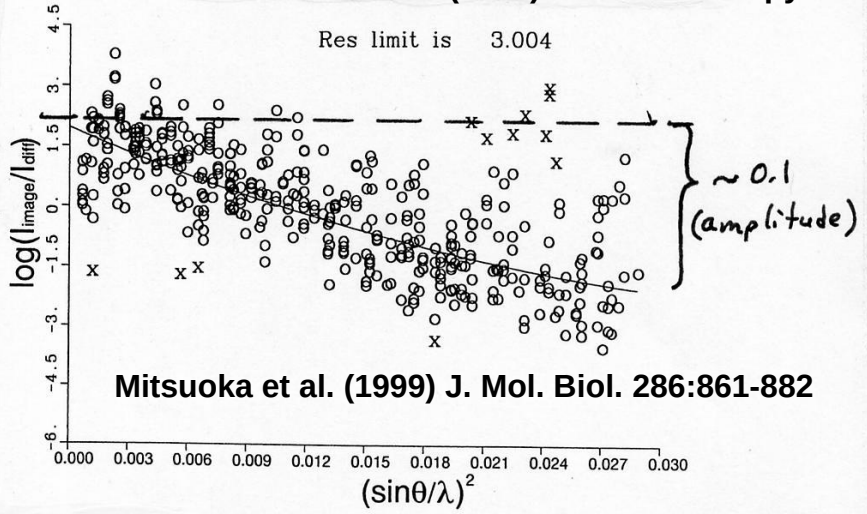
Short Digression, cont.

INSTEAD, LARGE DATA SETS ARE NEEDED BECAUSE OF BEAM-INDUCED MOVEMENT

- **Beam-induced movement causes the high-resolution image-contrast to be ~0.1 of what physics allows for perfect images**
- **Both the particle-size that can be aligned at high resolution and the number of particles required in the data set "scale" as the square of this factor**
- **i.e. we currently should expect that data sets will require 10^6 rather than 10^4 particles in order to go to atomic resolution**

	RES'N (Angstrom)	FRACT'N	RANGE
VERMICULITE	4.5	0.24	0.21-0.26
PARAFFIN	4.2	0.05	0.03-0.11
PURPLE MEMBRANE	4.3 8.9 23.0	0.04 0.27 0.60	--- --- ---

Henderson & Glaeser (2005) Ultramicroscopy

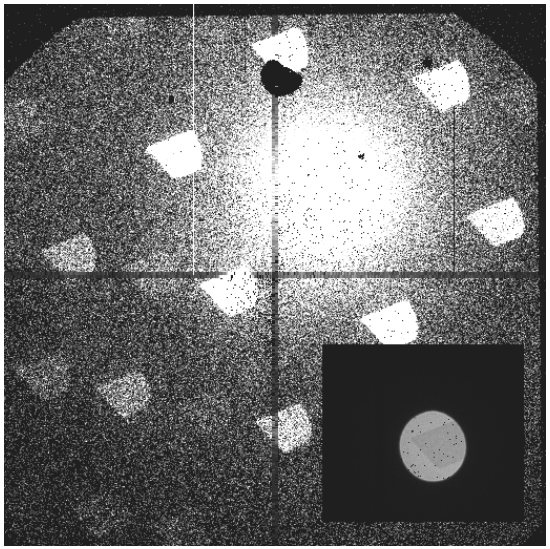


Mitsuoka et al. (1999) J. Mol. Biol. 286:861-882

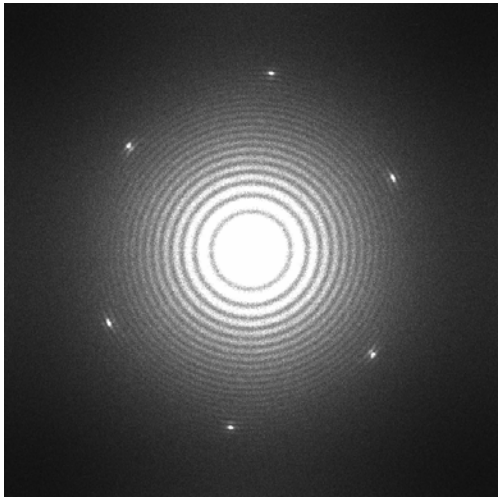
Short Digression, cont.

THE RATIO OF ACTUAL TO PHYSICALLY POSSIBLE IMAGE CONTRAST IS EASY TO MEASURE FOR 2-D CRYSTALS

- First measure the ratio of 2F(g)/F(0) from electron diffraction intensities
  - Medipix detector is superb for this, by the way!
- Then compute the ratio of F(g)/F(0) from images
  - The two will be identical if the images are perfect
  - The measured ratio quantitatively tells us how far our images are from what they could be



EXAMPLE OF DEFOCUSED DIFFRACTION PATTERN OF PARAFFIN, RECORDED ON THE MEDIPIX DETECTOR

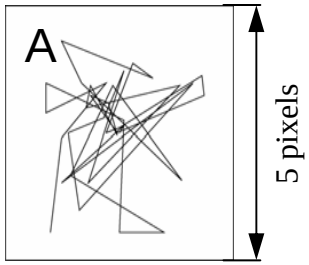


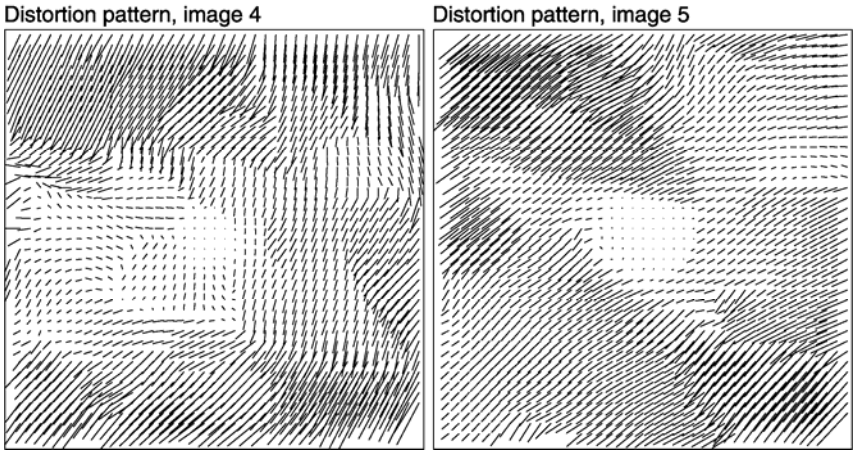
EXAMPLE OF A "BEST IMAGE" RECORDED WITH FLOOD-BEAM ILLUMINATION ON PHOTOGRAPHIC FILM

F(g)/F(0) ~ 13% OF THEORETICAL

ACTUAL MOVEMENT OF THE SPECIMEN IS THE PROBLEM (NOT IMAGE-DEFLECTION)

- **A series of 40 images was recorded from the same area of a paraffin crystal**
  - **Possible for the first time with the Tietz "HD" camera (Chris Gilpin, Dallas)**
- **The MRC 2-D "unbending" analysis showed that local areas of the crystal move *relative to the carbon support film* from one frame to the next**





Vector displacement maps that Indicate the shifts in unit cell positions from an undistorted crystal

image 4 before unbend						
20	50	5	44	15	38	17
37	23	57	24	50	47	40
54	23	110	46	40	88	15
70	86	118	131	76	35	28
78	88	93	109	86	18	23
55	41	33	61	22	15	76
20	41	58	64	93	51	17

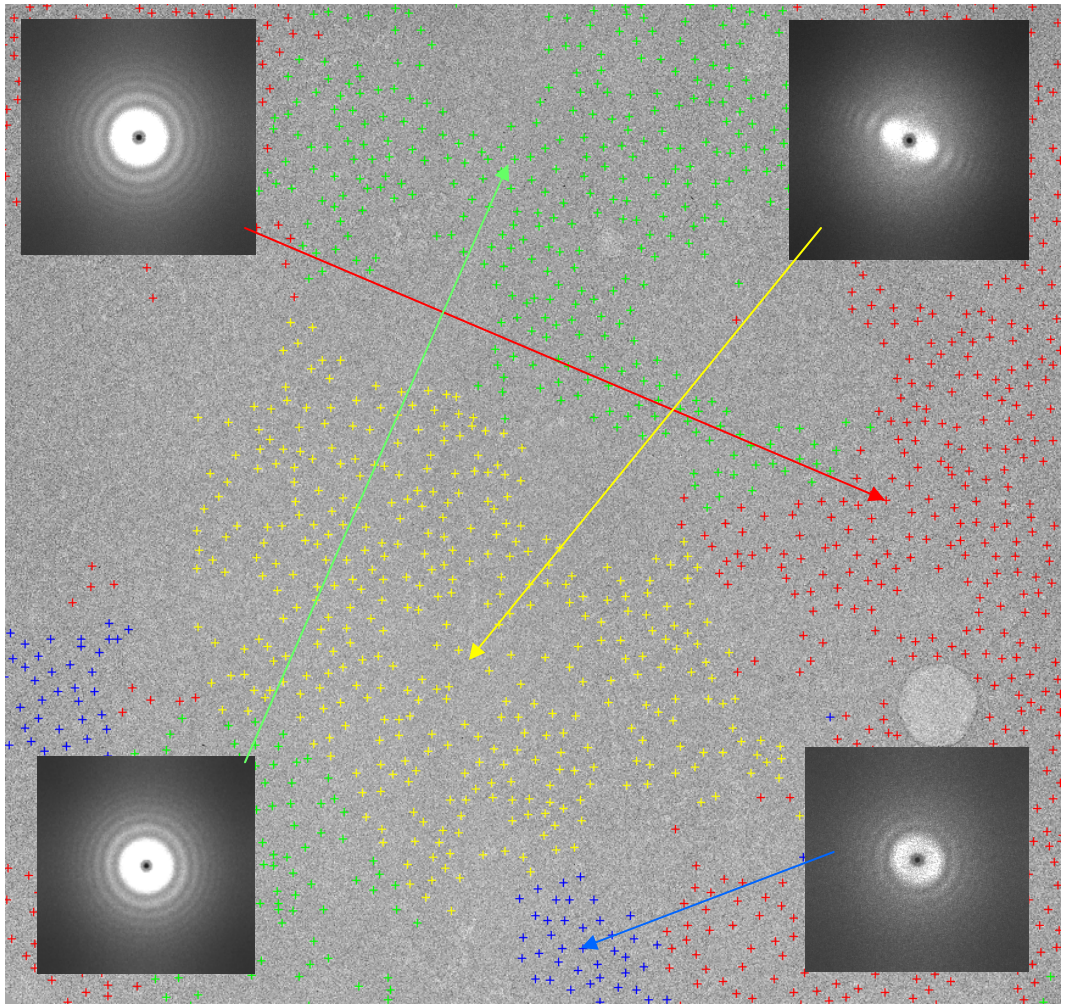
image 4 with pattern 4						
23	8	13	9	26	23	44
16	31	10	15	38	7	30
29	6	19	13	14	28	31
48	70	114	296	95	73	2
23	24	15	18	21	23	24
3	28	21	24	19	12	25
30	34	2	30	8	6	38

image 4 with pattern 5						
52	55	26	39	22	13	7
21	40	35	21	14	63	47
14	11	20	15	17	38	29
21	15	73	139	109	20	42
36	32	28	13	13	7	61
72	26	37	20	42	16	45
5	53	15	50	42	32	42

Frame-specific unbending is required in order for computed diffraction peaks in Fourier transforms to have maximum intensity and sharpness

ICE-EMBEDDED PARTICLES ALSO UNDERGO BEAM-INDUCED MOVEMENT

- **Thon patterns for single particles (on carbon) can be classified according to drift and B-factor (Hall, in preparation)**
  - **Thon patterns shown as inserts are class averages**
- **Classes segregate within separate areas of the micrograph**



Ribosome images courtesy of the lab of Joachim Frank

4b. ADDRESSING CONFORMATIONAL HETEROGENEITY ALSO REQUIRES VERY LARGE DATA SETS

- **Assuming that particles that are in distinct (different) conformations can be recognized – Penczek – one must collect images of many just to get a few**
  - **Sorting into more homogeneous data subsets containing $1/N^{th}$ of the particles requires that one must collect "N" times more to begin with**

- **In favorable cases, however, one may end up with 3-D structures for part or all of a biochemical cycle**

THIS IS GOOD!

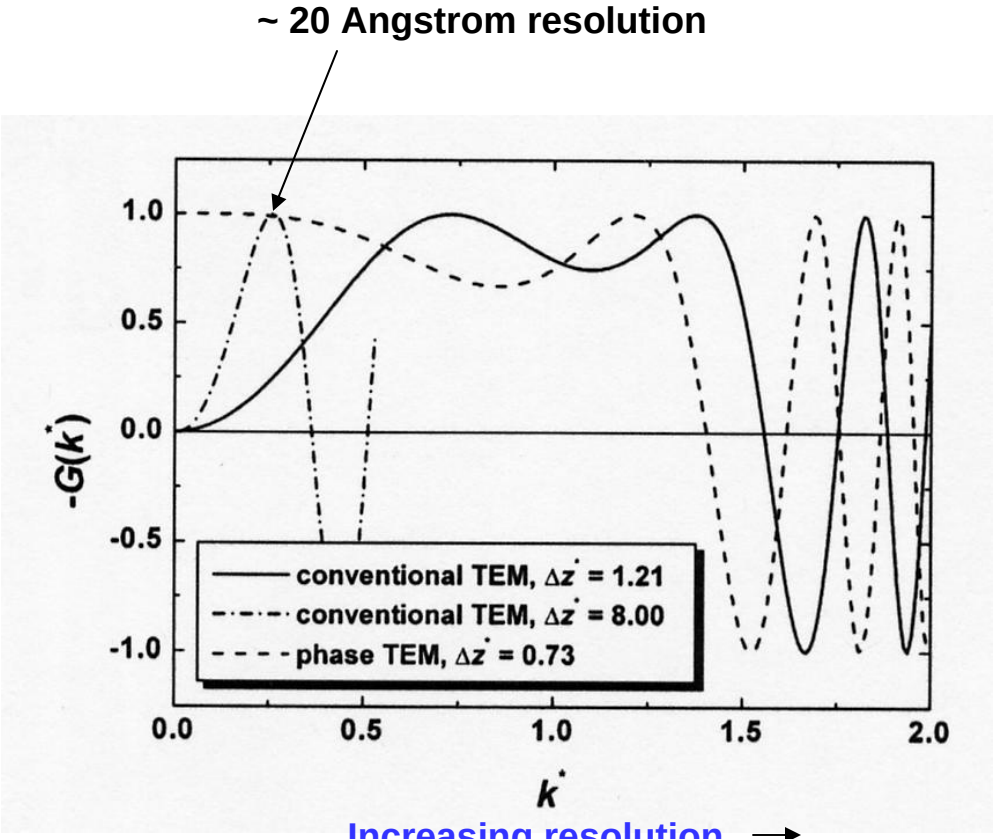
Automation & high throughput of data collection and data analysis can help to make this be routine

New Digression

EXTENDING SINGLE-PARTICLE CRYO-EM TO PARTICLES OF 250 kDa OR LESS IS LIMITED BY POOR ELECTRON OPTICS

- You can see the object at high defocus, but the contrast-transfer is corrupted, or
- The contrast-transfer is not badly corrupted at low defocus, but you cannot see the particle

Nagayama has successfully implemented in-focus phase contrast for which the CTF is modeled by the dashed curve



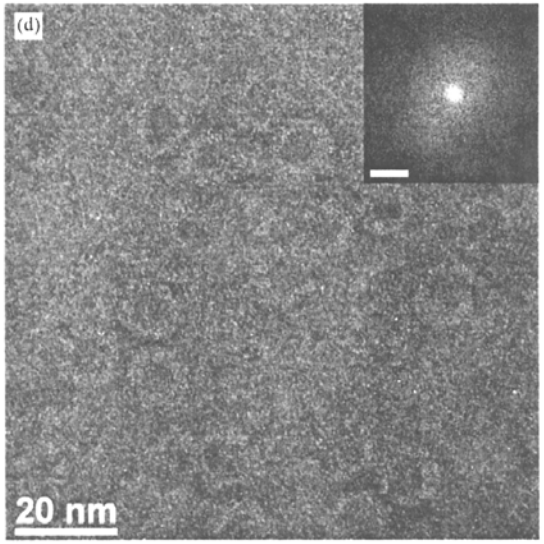
Danev et al. (2002) Ultramicroscopy 90:85-89

New Digression, cont.

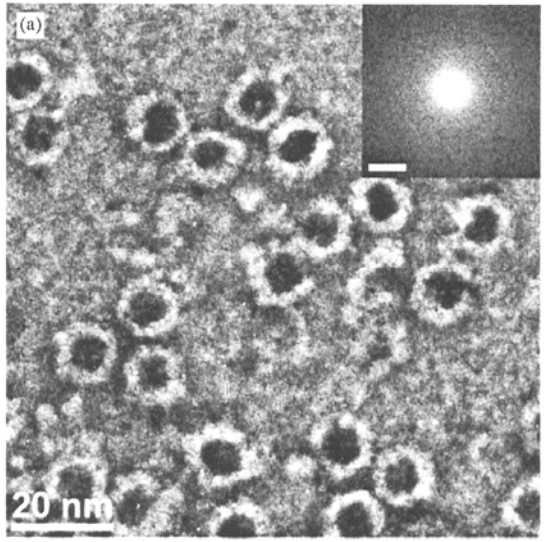
Nagayama has successfully implemented in-focus phase contrast

- A continuous carbon film across the objective aperture
- A small hole drilled in the center for the unscattered electrons

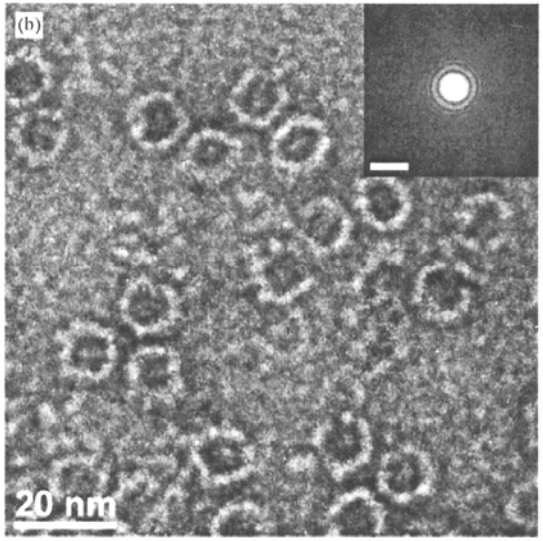
Danev et al. (2001) Ultramicroscopy 90:85-89



"In focus"
(little phase contrast)



"In focus" plus a carbon phase plate
(full phase contrast)



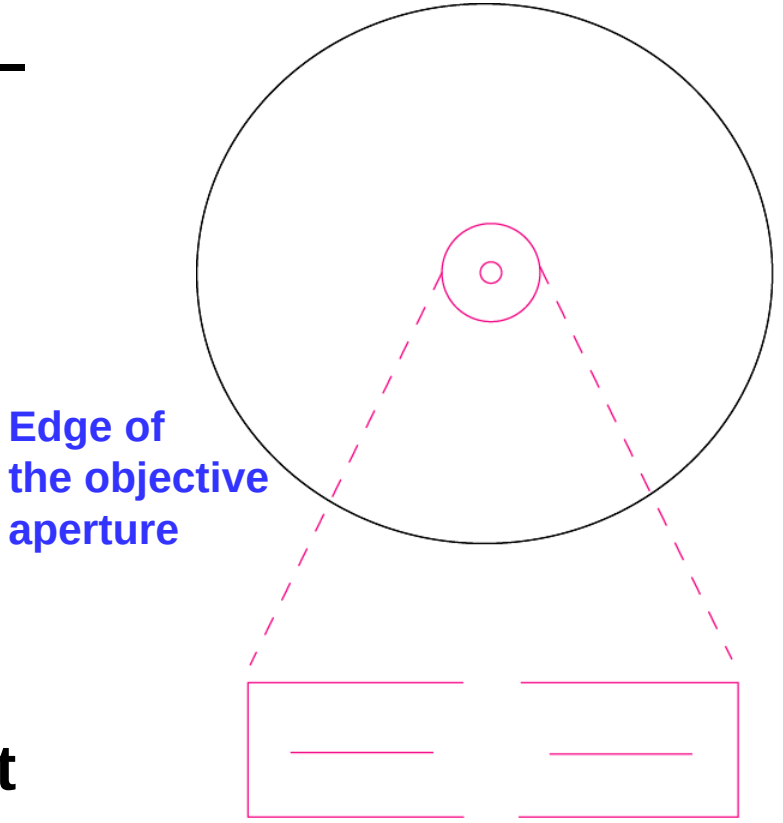
"high defocus"
(imperfect phase contrast)

New Digression, cont.

ELECTROSTATIC DEVICES MAY ALSO PROVIDE IN-FOCUS PHASE CONTRAST

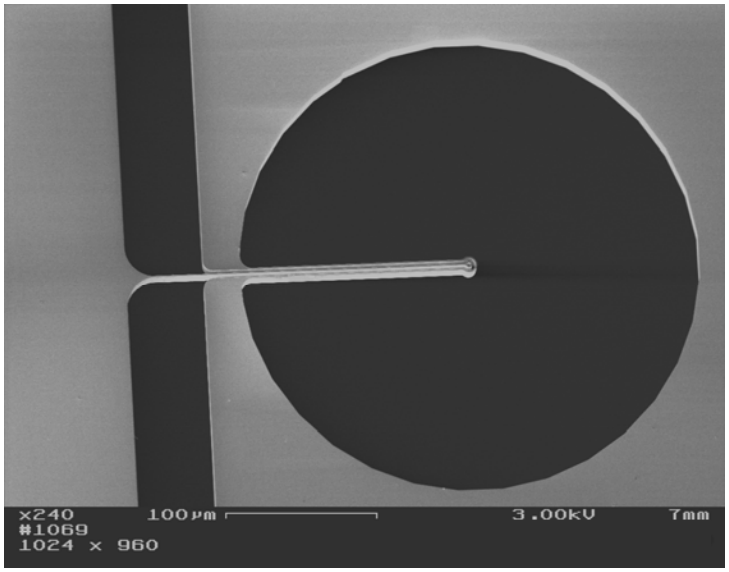
WITHOUT THE PARTIAL LOSS OF ELECTRONS SCATTERED BY THE CARBON FILM

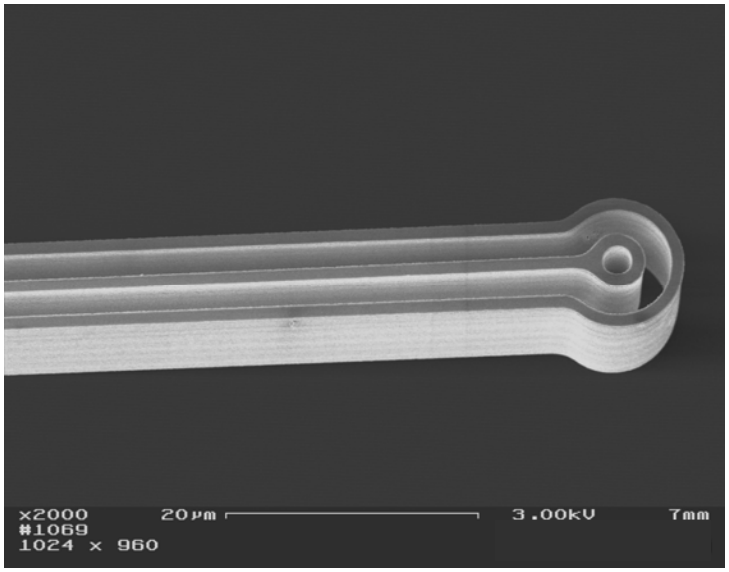
- **Boersch proposed – in 1947 – to use a small electrostatic device to apply a 90-degree phase shift just to the unscattered electron beam in the electron diffraction pattern**
- **An approximation to this device has been micro-fabricated and successfully demonstrated (Schultheiss et al. (2006) Rev. Sci. Instrum. Vol 77)**



New Digression, cont.

LBNL PHASE-CONTRAST APERTURE DESIGNED BY JIAN JIN AND FABRICATED BY ROSSANA CAMBIE

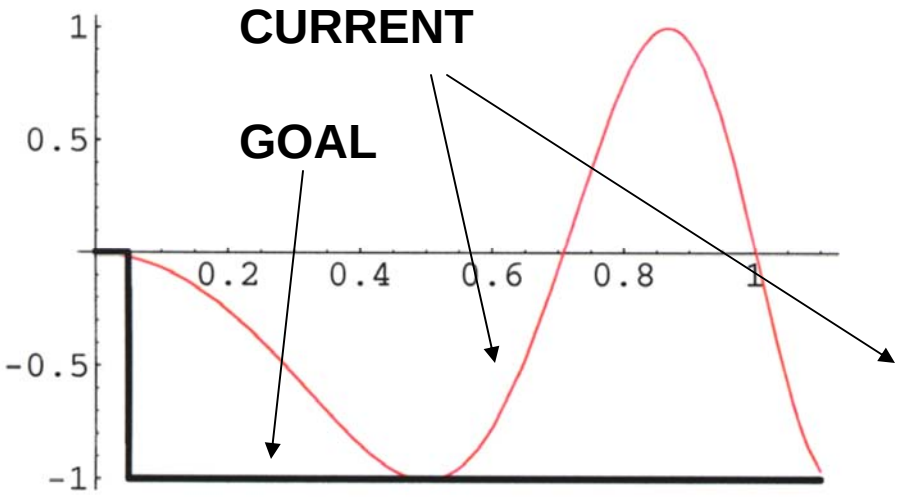




- **Left:** the black circular hole is what an objective aperture would normally look like, without the intruding electrode and the supporting structure on the left.
- **Right:** enlargement of the electrode structure at the tip.

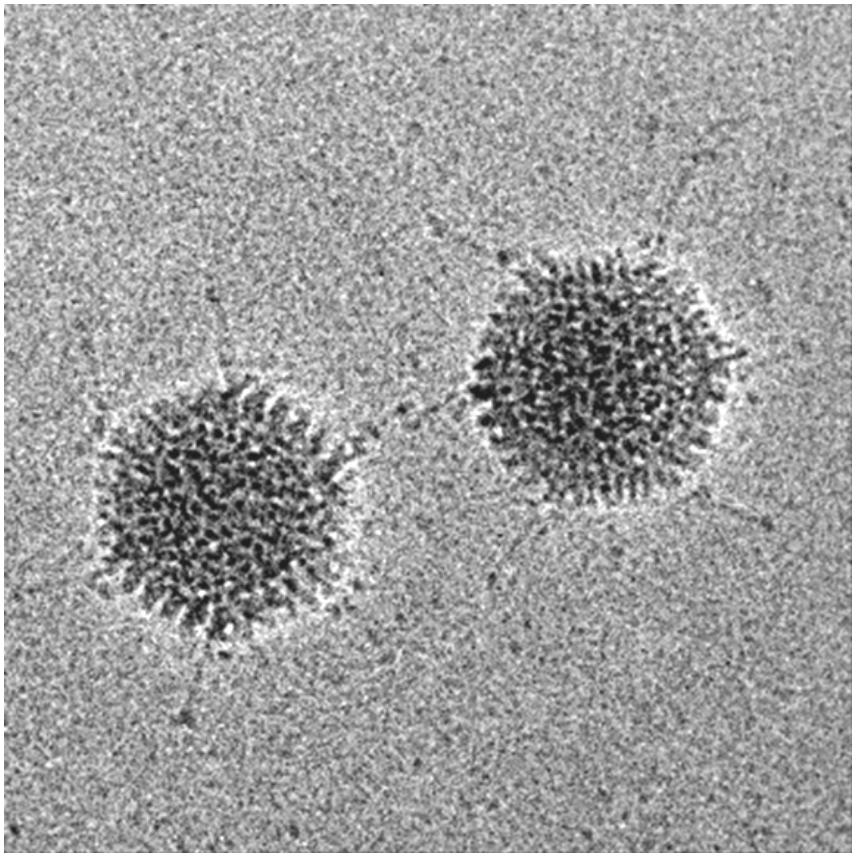
New Digression, cont.

NEAR-PERFECTION IN THE CTF IS ON ITS WAY!



- 1.5 nm fibers are quite visible with a CTF like the red curve
- Such features might be 2 or 3 times more clear with a micro fabricated Boersch phase plate plus Cs corrector*

  * [concept of M. Haider]



~1.5 nm Adenovirus penton-fibers
(Dr. Phoebe Stewart)

SUMMING UP …

- **Automation and high throughput is possible for MANY steps in single-particle cryo-EM**
  - **Even preparing cryo-EM grids (Vitrobot – not discussed)**
- **Automation of data collection & analysis should be pushed to the point that producing interpretable structures is as fast and easy as running a gel**
  - **Such a capability would make cryo-EM a viable tool for proteomics**
- **Automation of data collection & analysis should also make it practical to obtain:**
  - **Structures suitable for *de novo* chain tracing**
  - **Multiple conformations that are present in samples where these are in thermodynamic equilibrium**

ACKNOWLEDGEMENTS

- TPP II
  - Beate Rockel
  - Jurgen Peters
  - Wolfgang Baumeister
- PFOR
  - Ming Dong
  - Mark Biggin
  - Dieter Typke
  - Florian Garczarek
- Porting LEGINON to the JEOL
  - John Lee
  - Dieter Typke
  - Bridget Carragher, Clint Potter, et al.
- Automated film densitometry
  - Bob Nordmeyer
  - John Lee
  - Dieter Typke
- Particle boxing
  - Ravi Malladi
  - Umesh Adiga
- Beam-induced movement
  - Dieter Typke & Ken Downing
  - Richard Henderson et al.
  - Hongwei Wang
  - Richard Hall & the Frank lab
- In-focus phase contrast
  - Jian Jin
  - Rossana Cambie
  - Dieter Typke
  - Ken Downing
- Funding
  - DOE Genomics:GTL (Biggin et al.)
  - NIH Program Project (Frank, Ludtke, Adams, Ng, Malladi, Penczek, Chiu, Downing & Glaeser)
  - LBNL Laboratory Director's Research & Development Fund