NCMI Single Particle Workshop
Baylor, March 14-17th 2011

Methods for validation of CryoEM maps

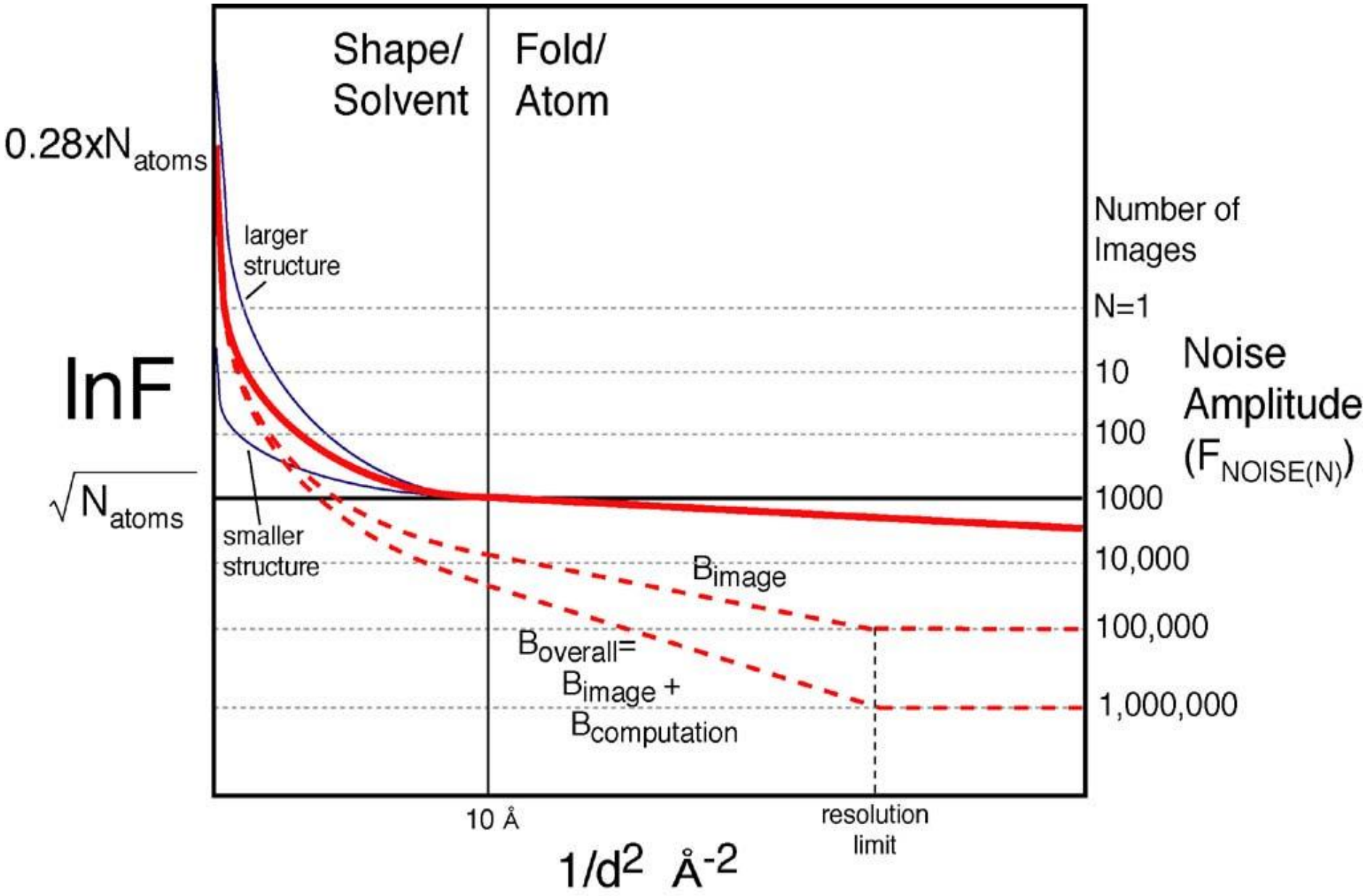
Richard Henderson

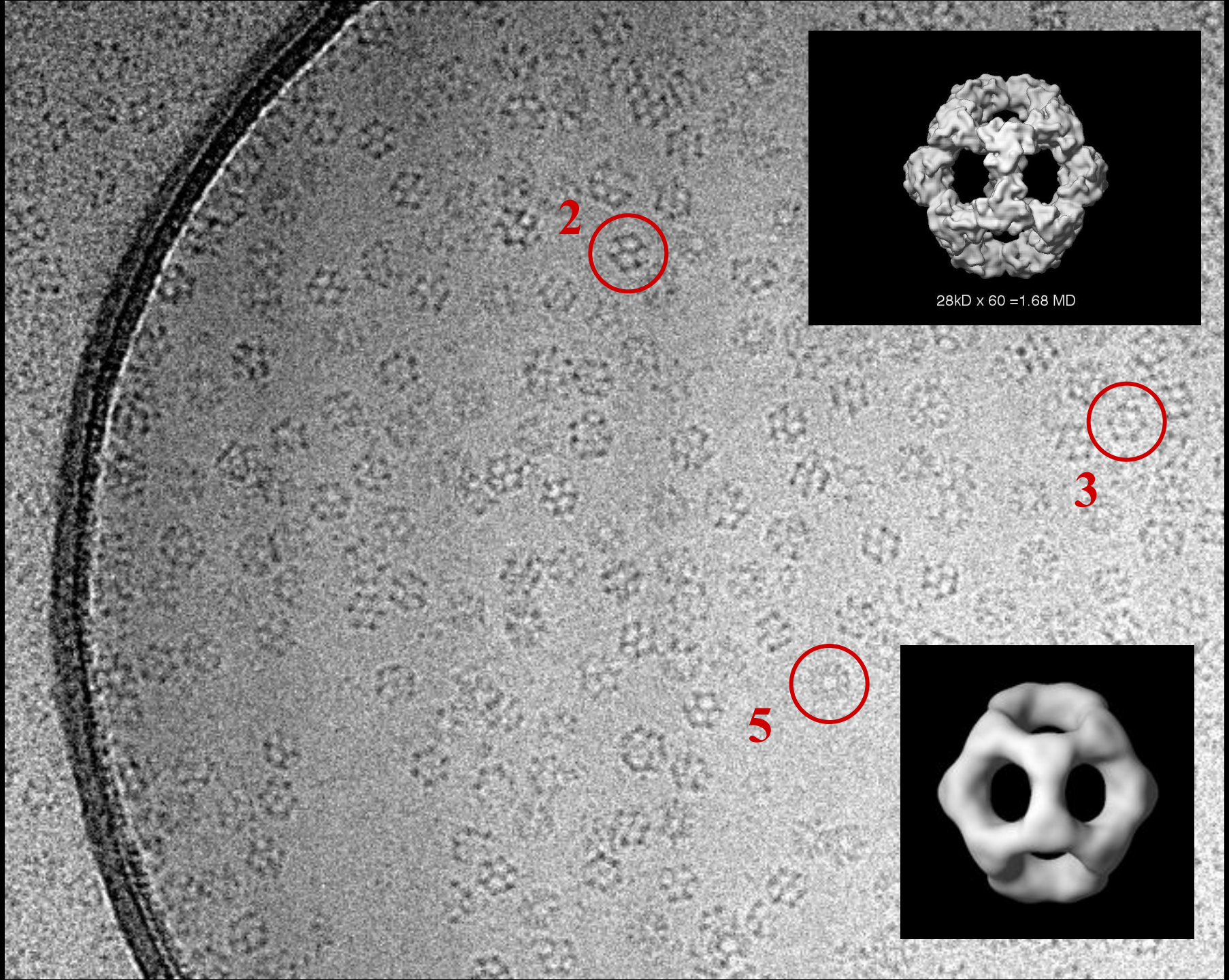
- B-factor sharpening, FOM weighting
- Tilt pair parameter plots
- Resolution, FSC
- Other checks and information to include in reports

Rosenthal & Henderson, (2003) - three main points

- More realistic (less conservative) resolution criterion (FSC = 0.14) derived in Appendix with Tony Crowther
- Sharpening map and f.o.m. weighting EM-Bfactor (Fernandez et al, JSB 2008)
- Tilt pair validation of orientation angle determination not yet very popular
- Also, tomography resolution limit of 20 Å

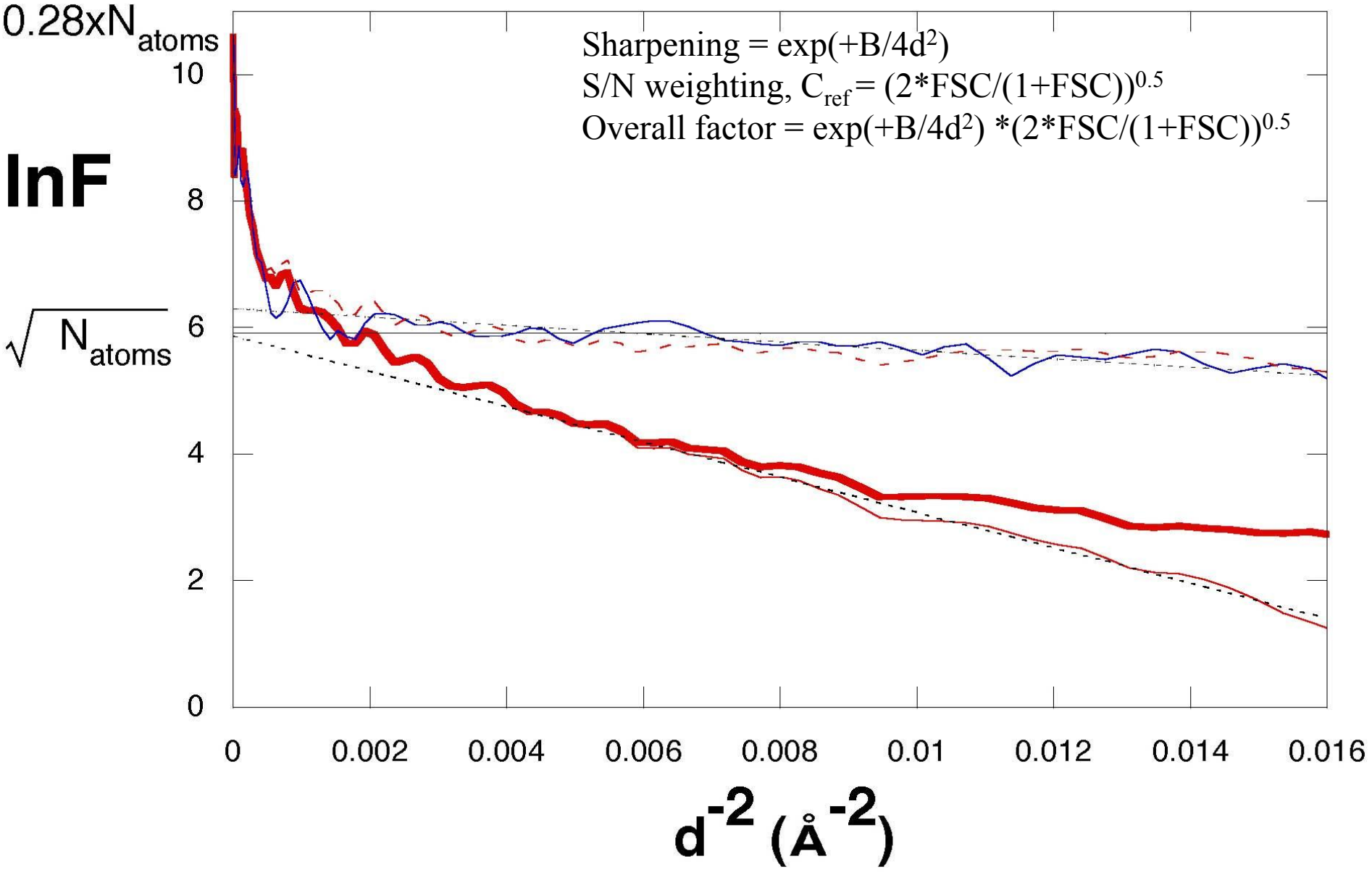
Theory – single particles in ice



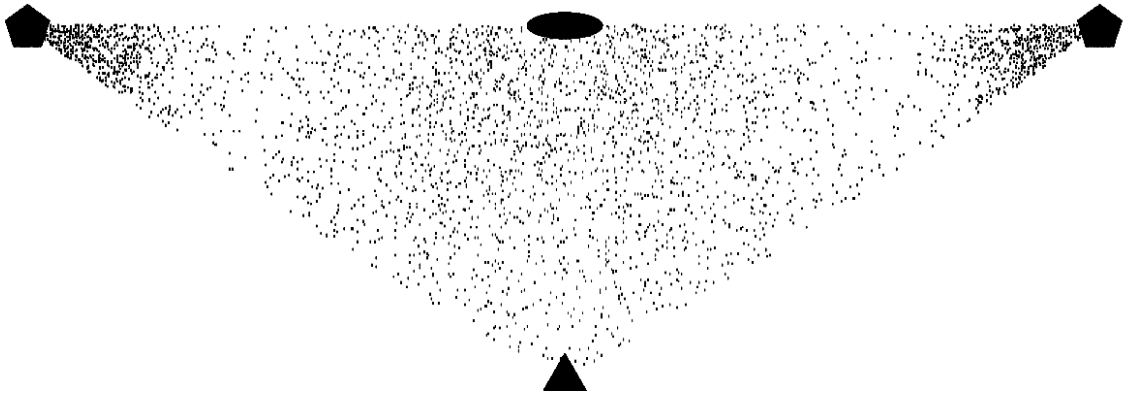

2
3
5
28kD x 60 =1.68 MD

Experimental data

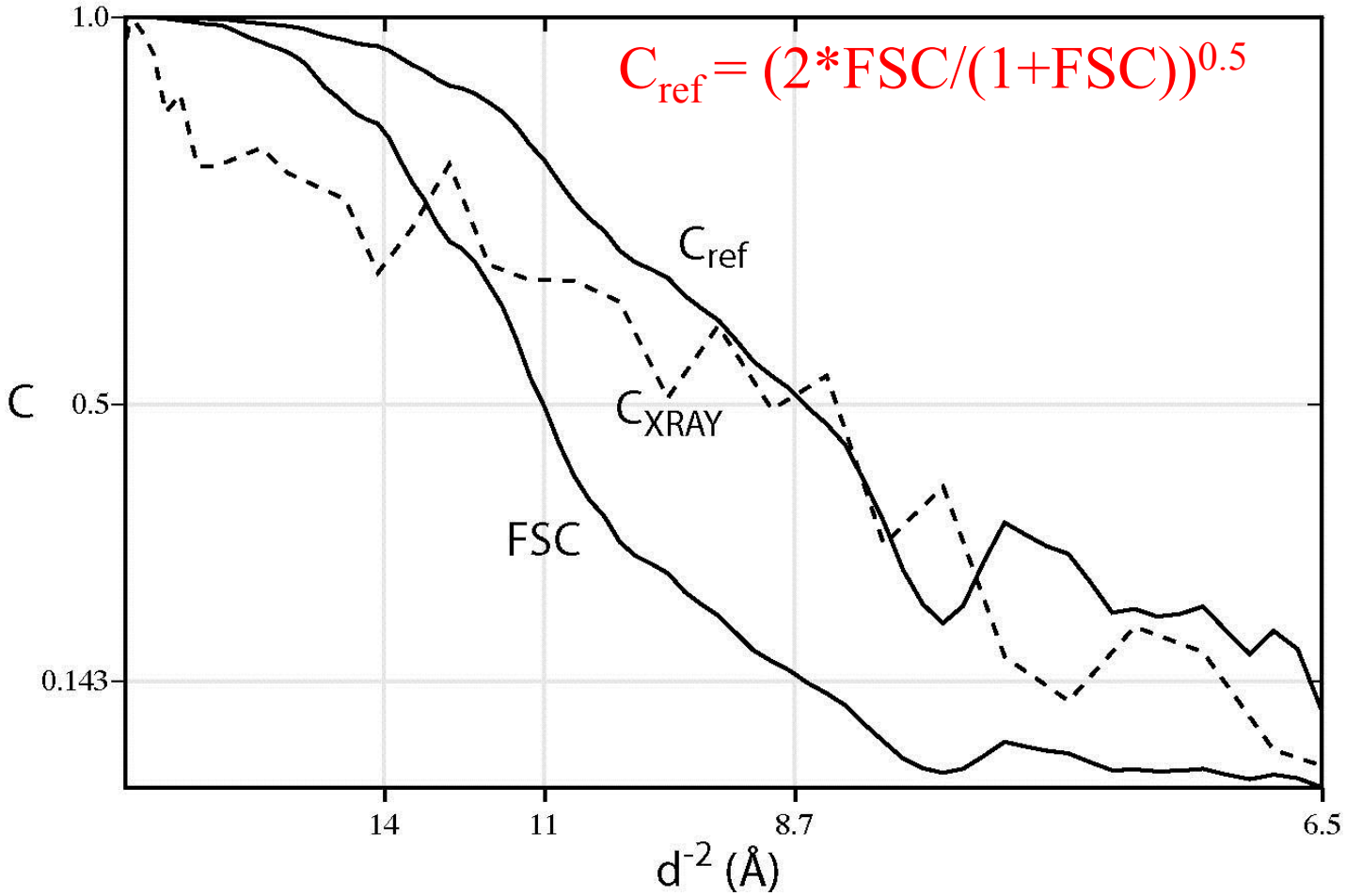
Rosenthal (2003) JMB **333**, 225-36
Fernandez (2008) JSB **164**, 170-5



Particle distribution

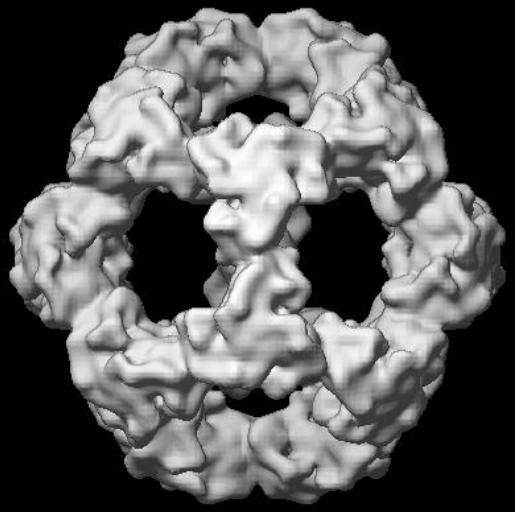


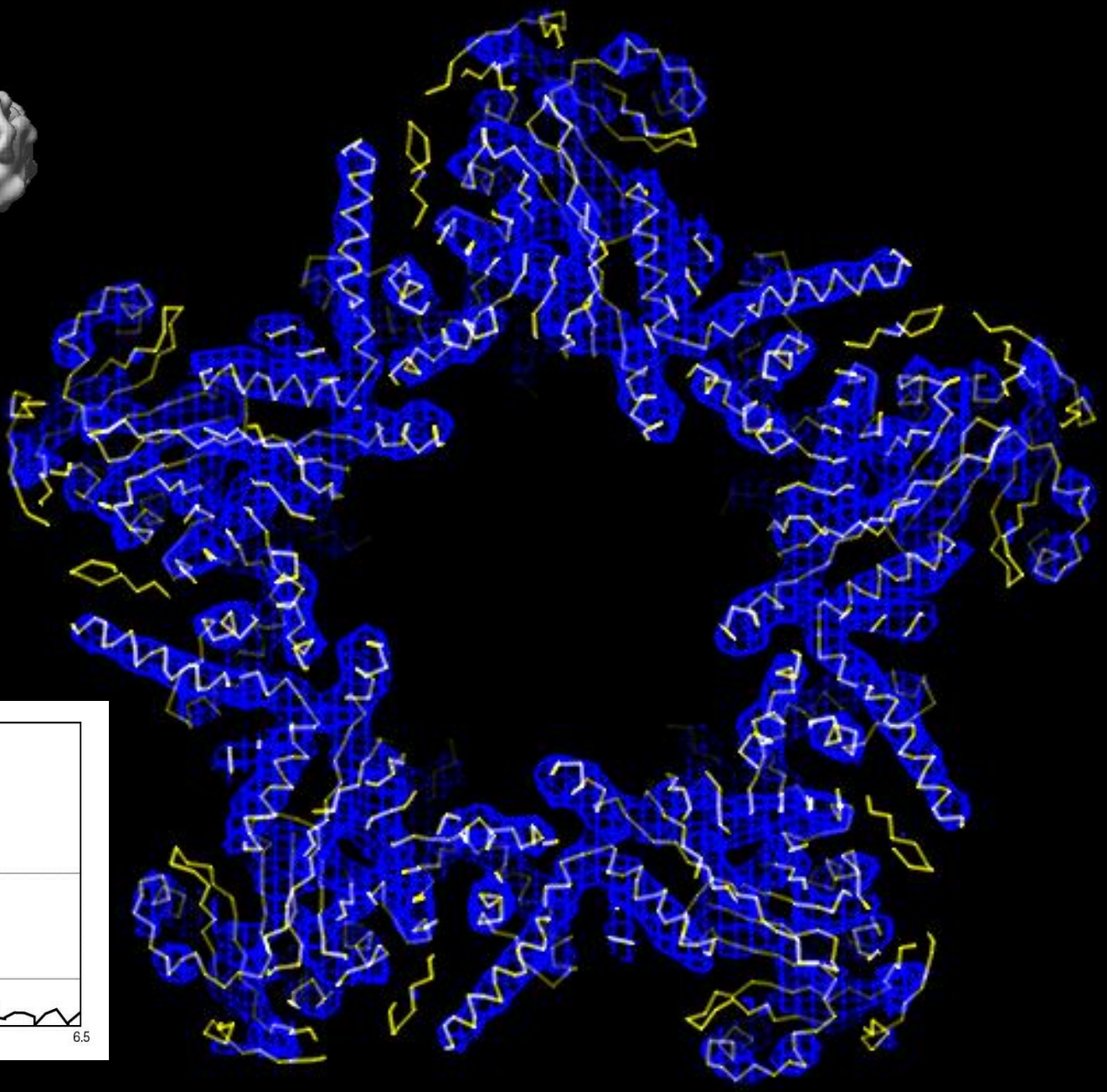
Fourier shell correlations

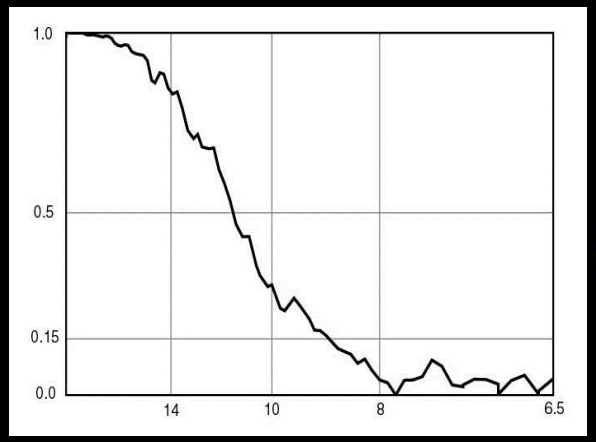


Application of Rosenthal & Henderson tilt pair validation approach
(9/131 citations up to March 2011)

- Pyruvate dehydrogenase : R & H (2003) JMB **333**, 721-42
- *Neurospora* P-type ATPase : Rhee et al (2002) EMBO J. **21**, 3582-89
- Bovine ATPase : Rubinstein et al (2003) EMBO J. **22**, 6182-92
- Chicken anaemia virus : Crowther et al (2003) J.Virol. **77**, 13036-41
- HepB surface antigen : Gilbert et al (2005) PNAS **102**, 14783-88
- Hsp104, yeast AAA+ ATPase : Wendler et al (2007) Cell **31**, 1366-77
- Yeast ATPase : Lau et al (2008) JMB **382**, 1256-64
- V-type ATPase, *T.thermophilus* : Lau & Rubinstein (2010) PNAS **107**, 1367-72
- DNA-dependent PKase : Williams et al (2008) Structure **16**, 468-77

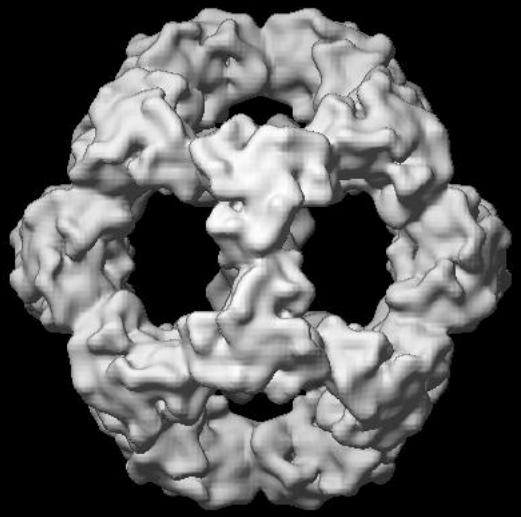


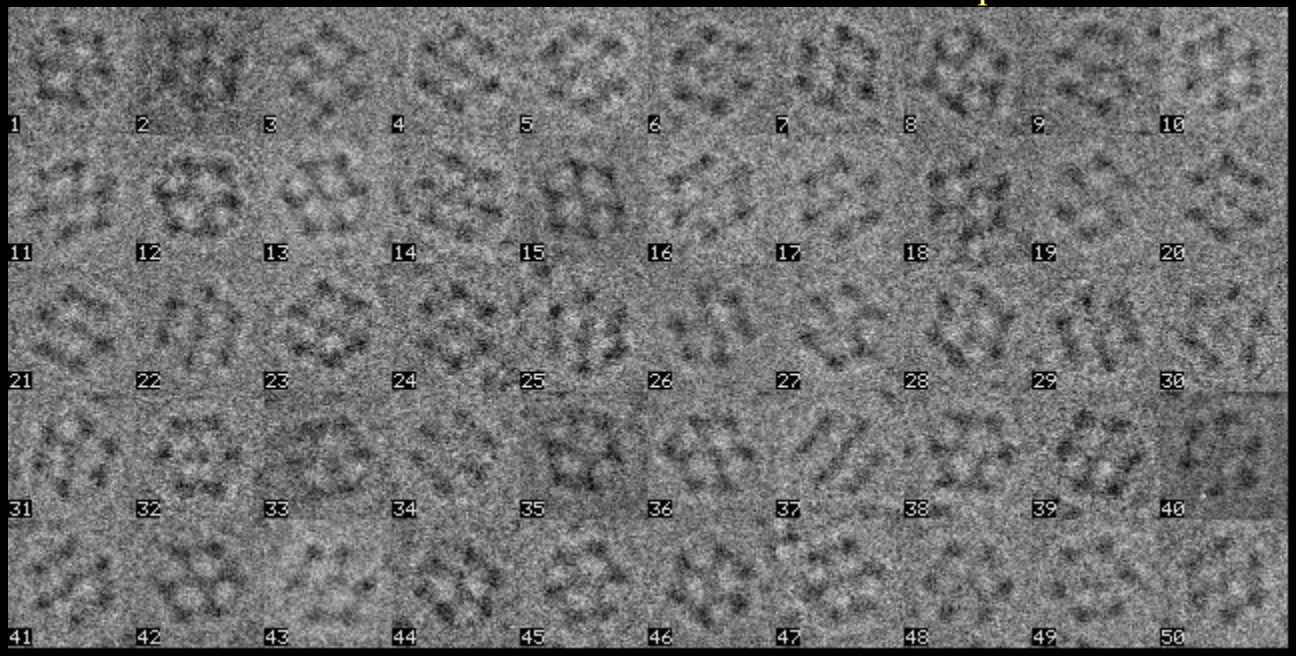



1.0
0.5
0.15
0.0
14
10
8
6.5

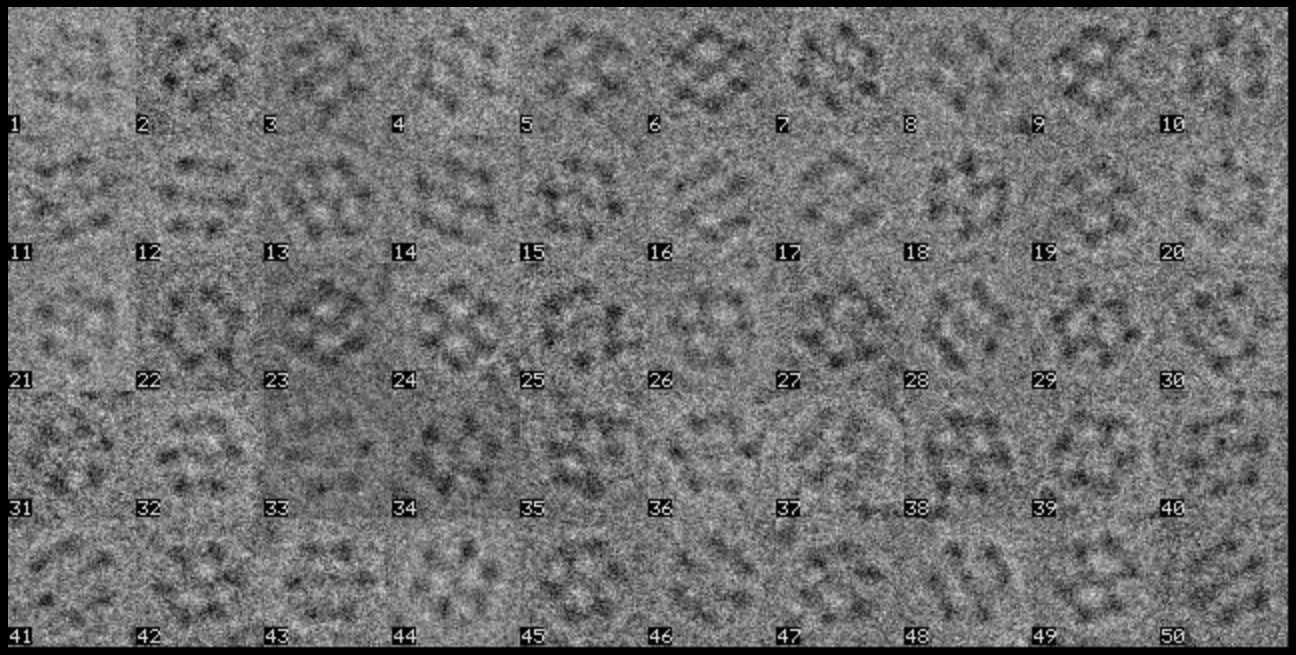
Rosenthal tilt pair validation test

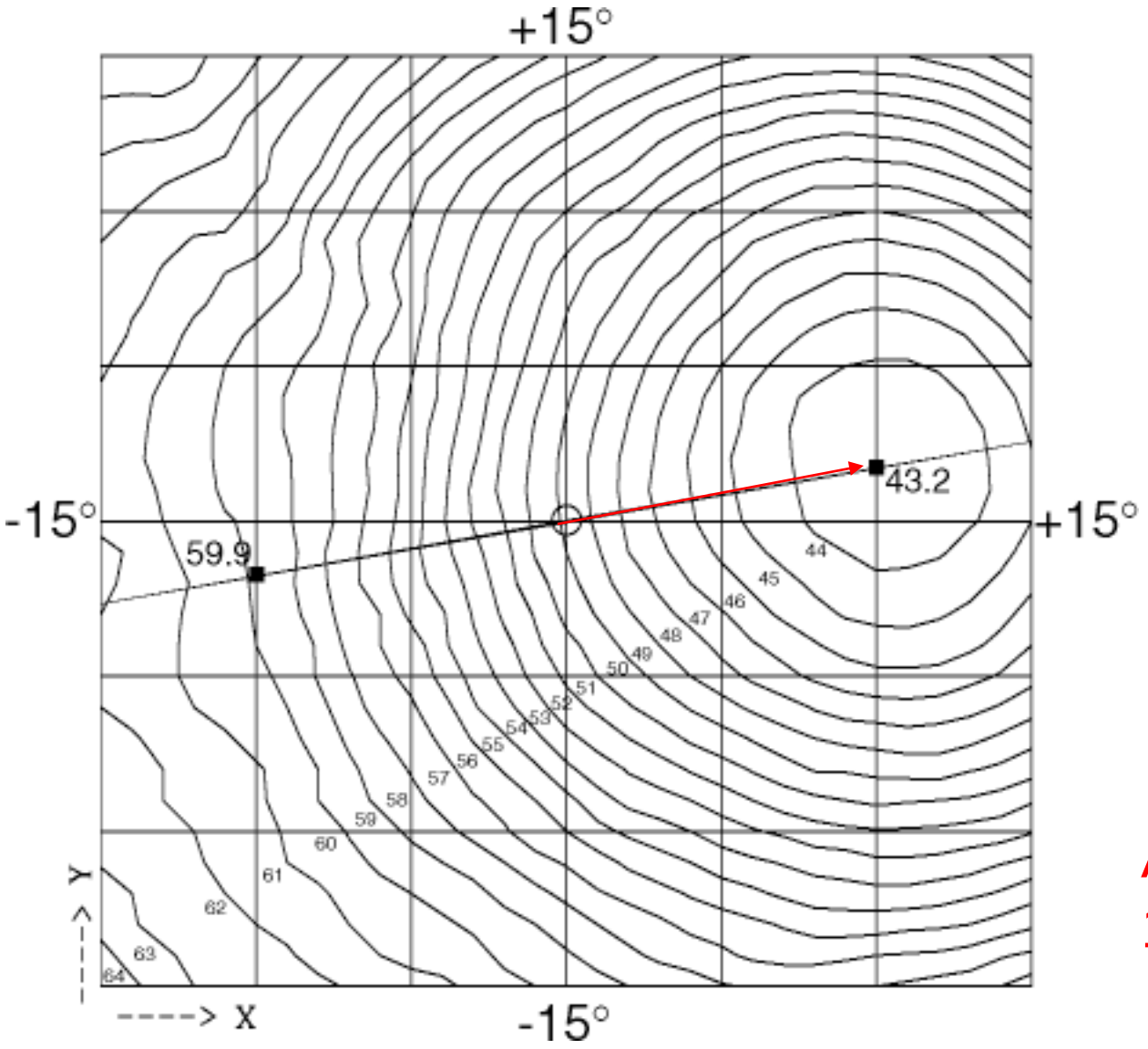
UNTILTED
$(\psi,\theta,\varphi)_u$





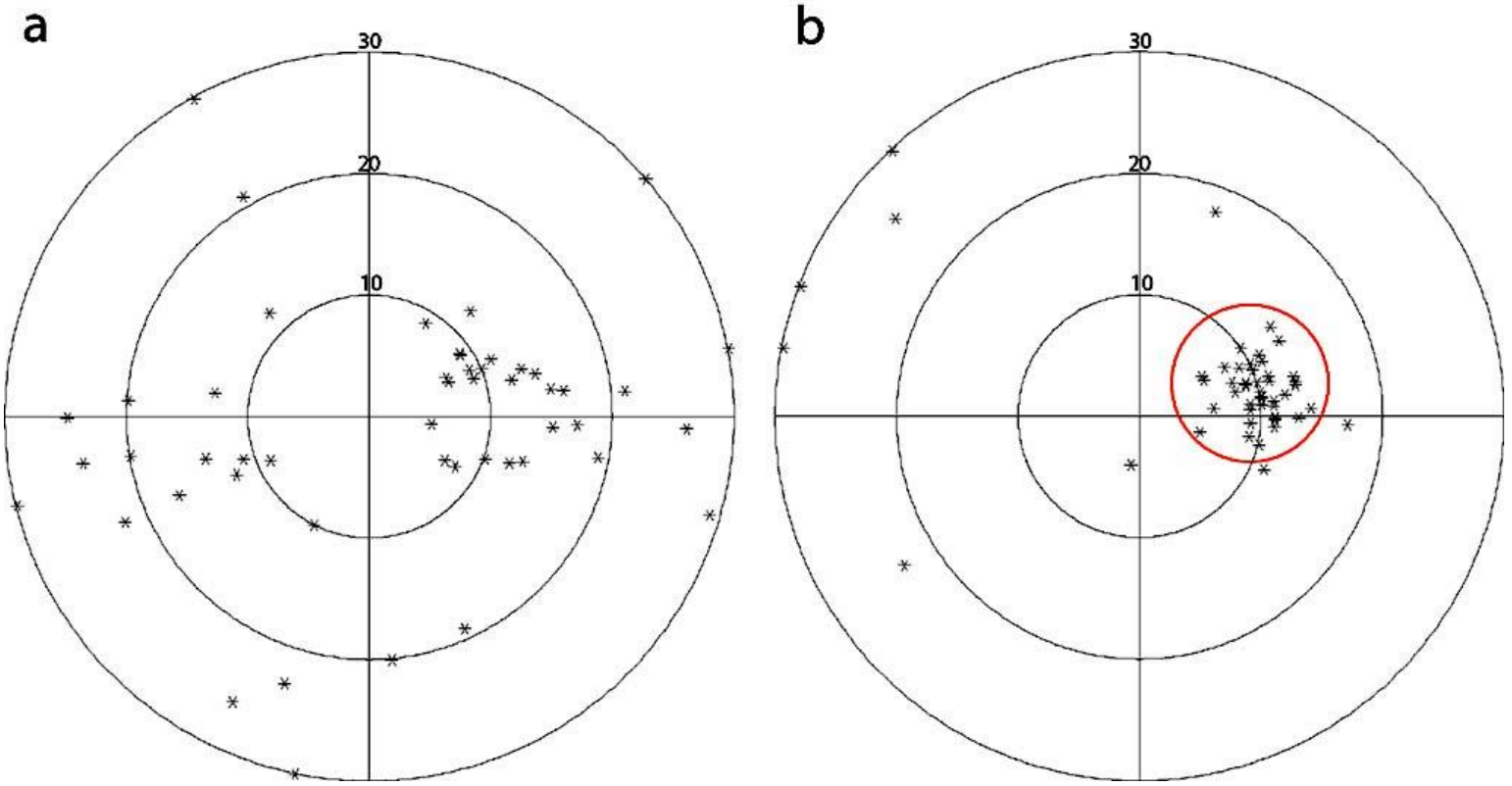
TILTED
10 degrees
$(\psi,\theta,\varphi)_t$





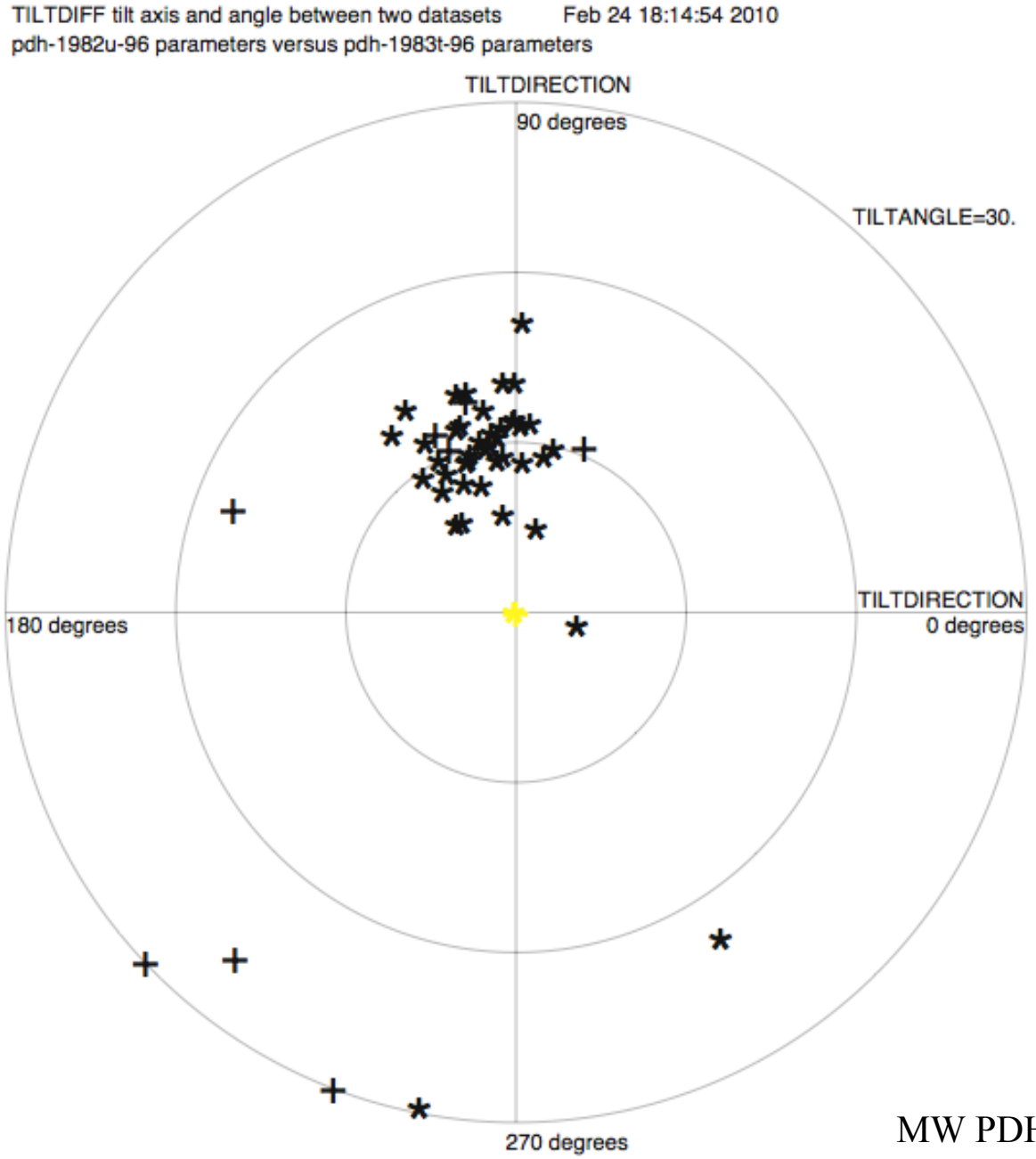
ANGLE
10 deg

Mean phase residual for 50 particle image pairs – ANGPLOT + FREALIGN



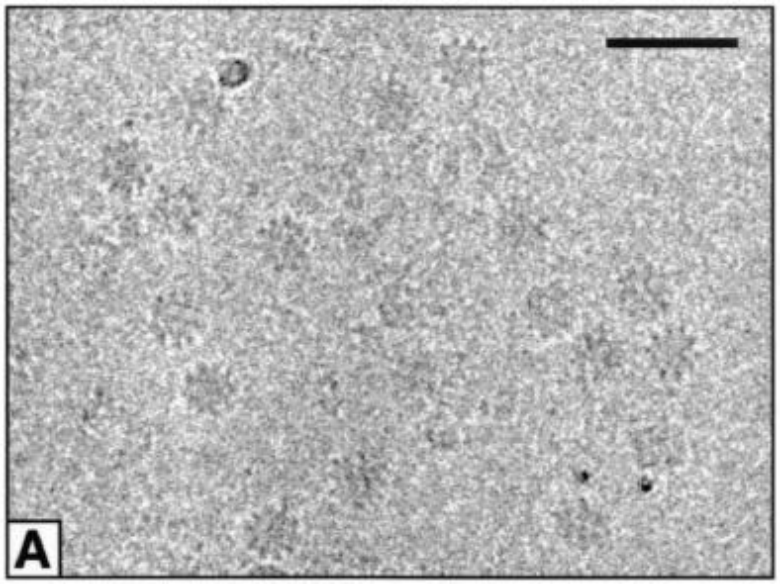
Individual particle image pairs – TILTDIFF output

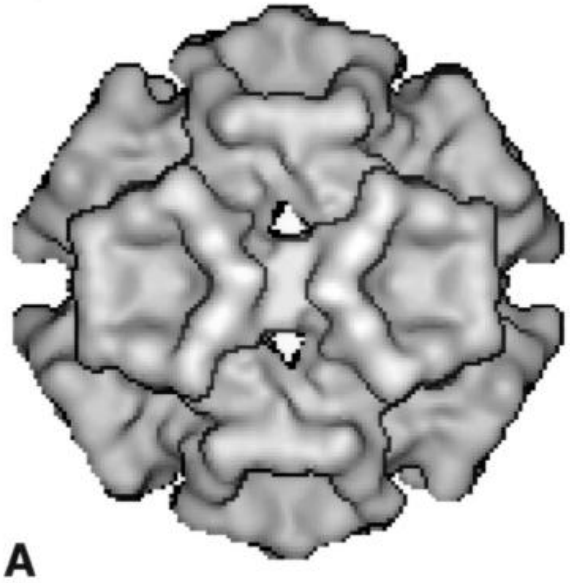
Pyruvate dehydrogenase, E2 catalytic domain, Rosenthal & Henderson JMB, 2003, replotted

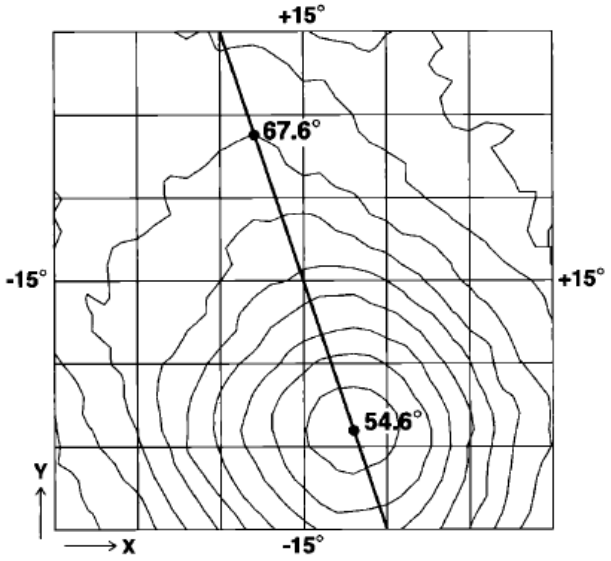


MW PDH_E2CD = 1.6 MDa

Chicken Anemia virus, Crowther et al, J.Virology 2003

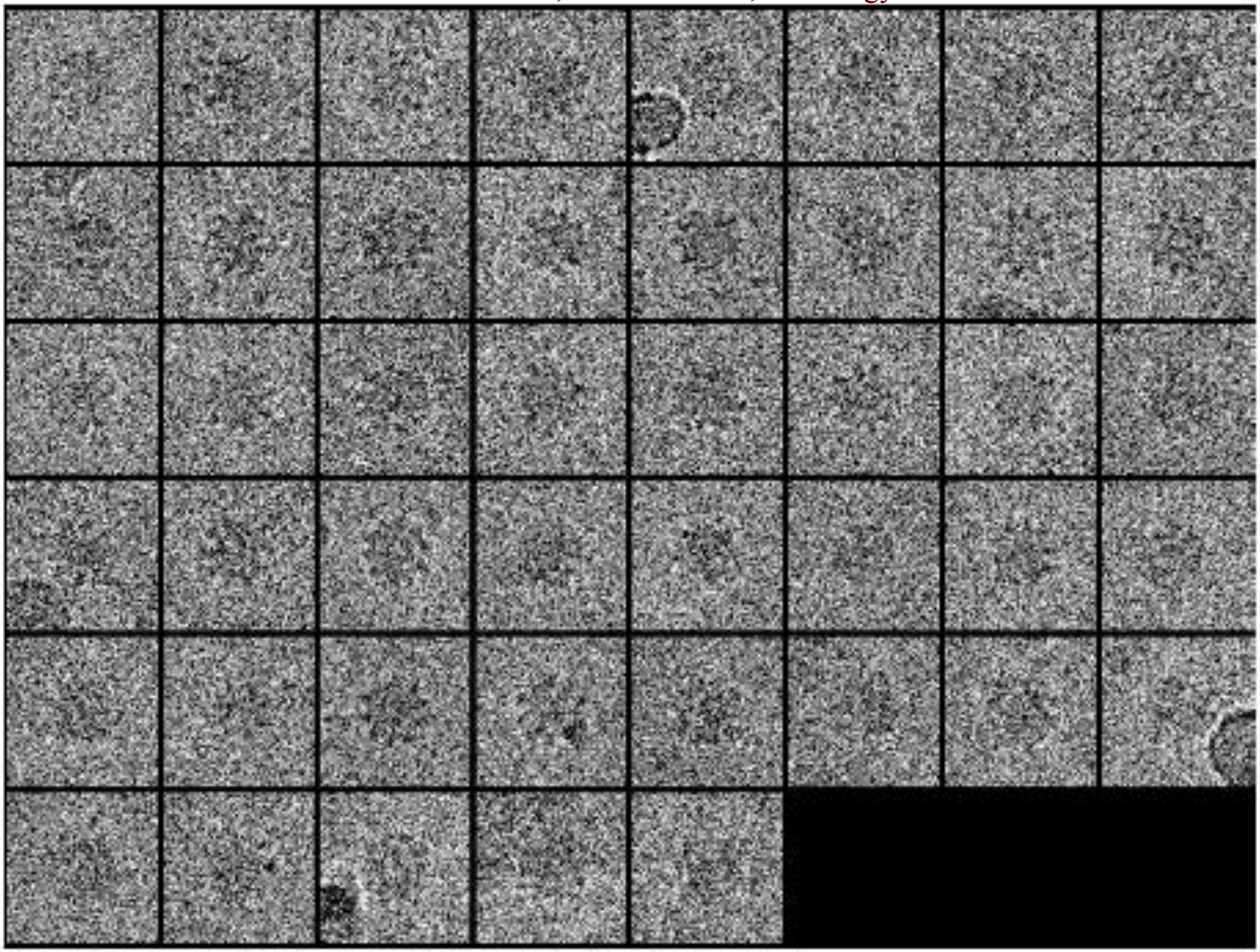




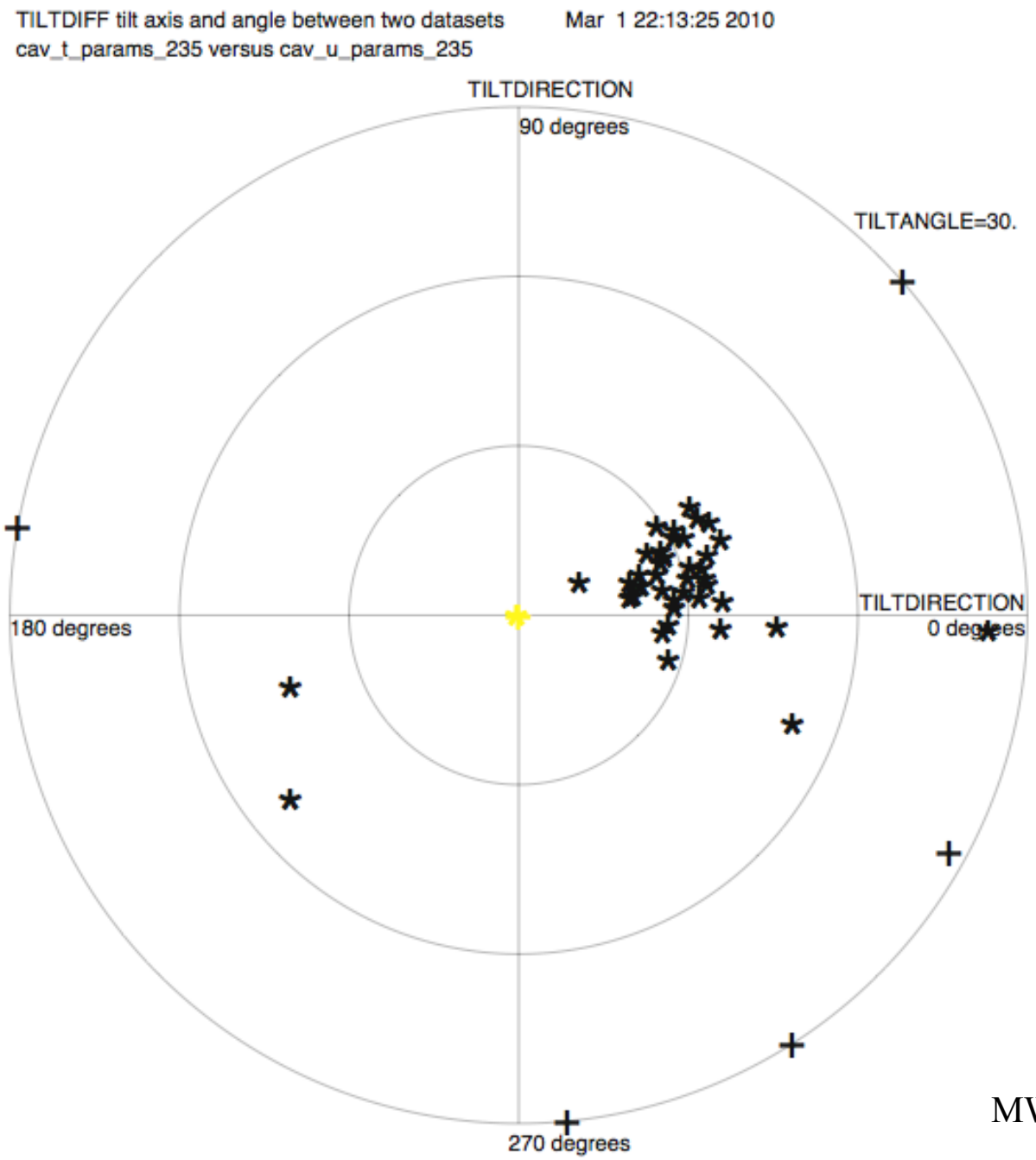


MW CAV = 3.3 MDa

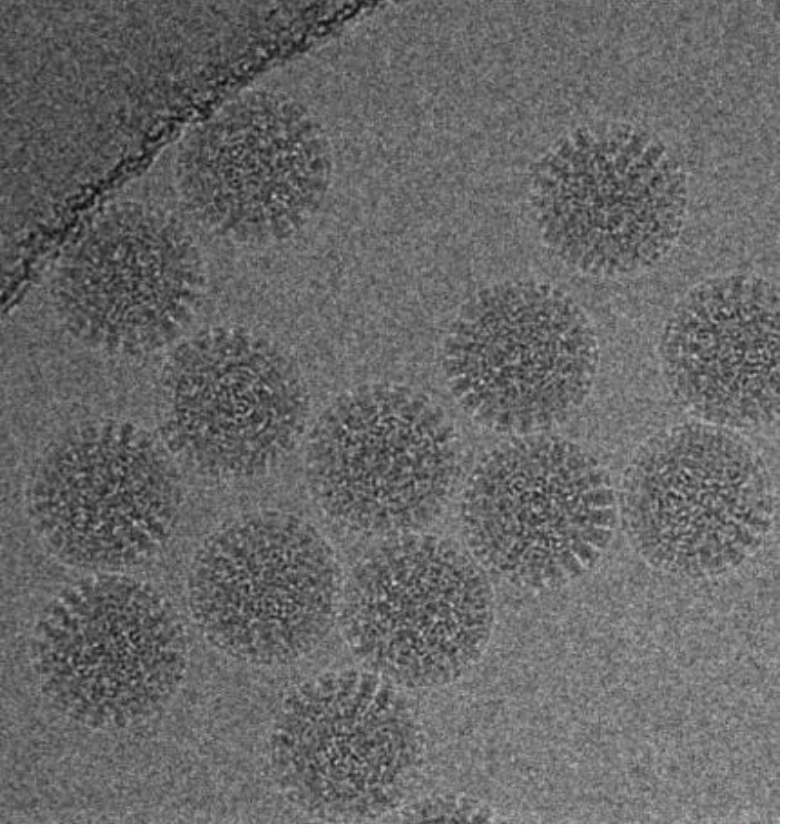
Chicken Anemia virus, Crowther et al, J.Virology 2003



Chicken Anemia virus, Crowther et al, J.Virology 2003



MW CAV = 3.3 MDa



Human Rotavirus DLP

3.8 Å, B-factor 450Å^2

Zhang et al & Grigorieff (2008) PNAS **105**, 1867-72.

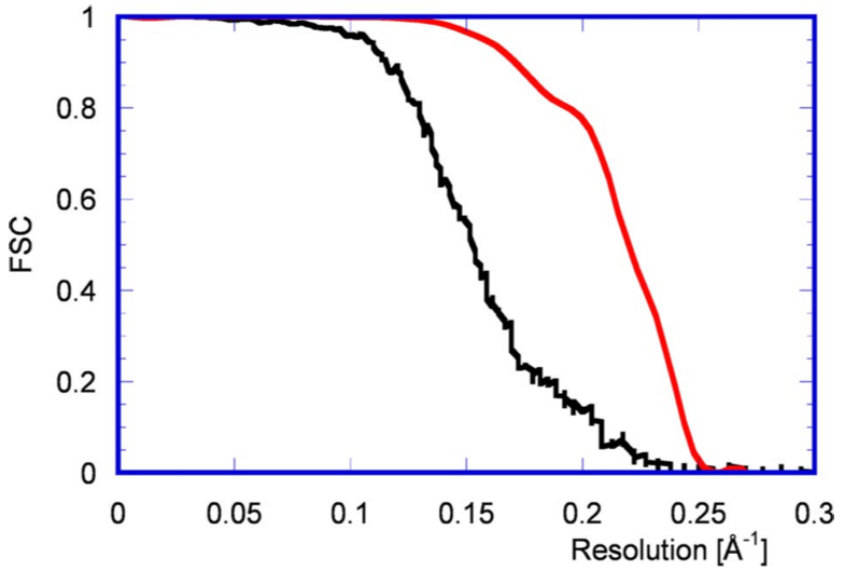
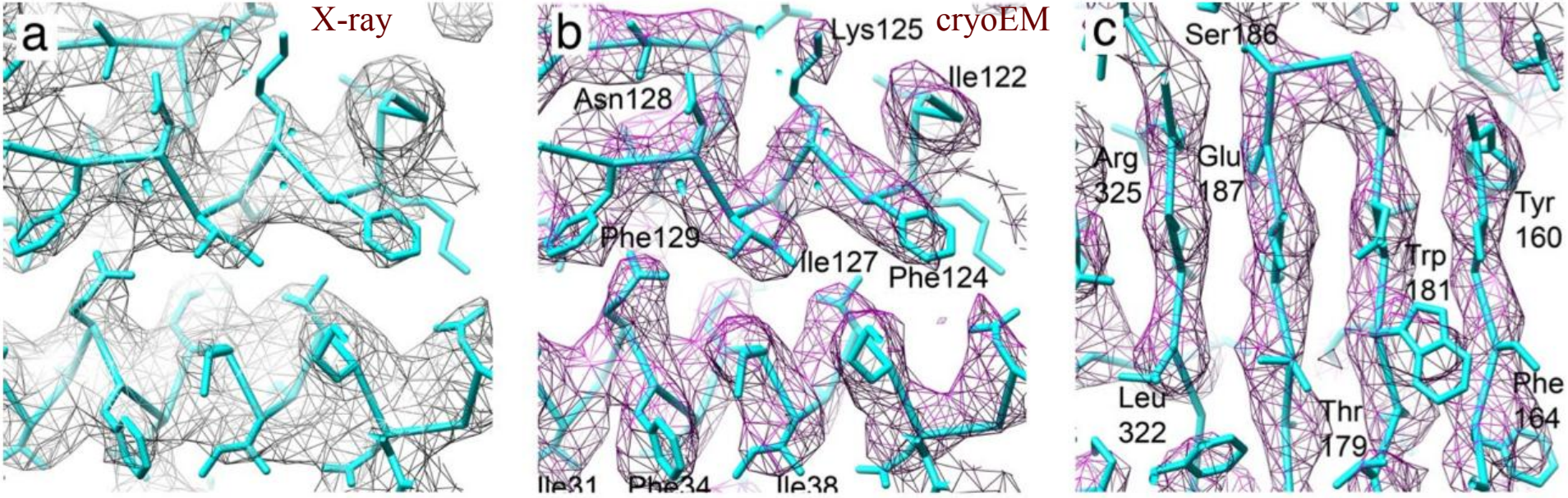
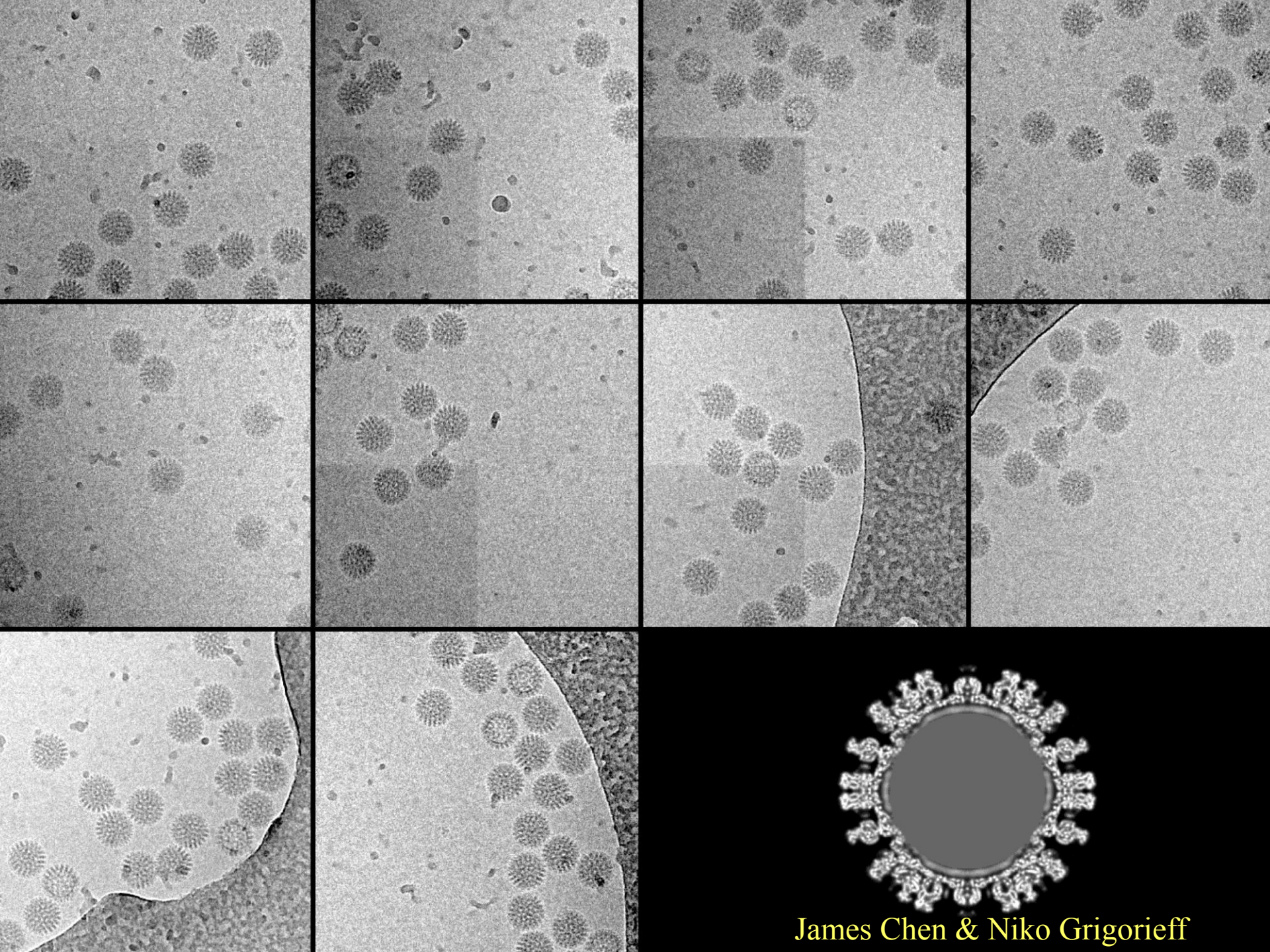
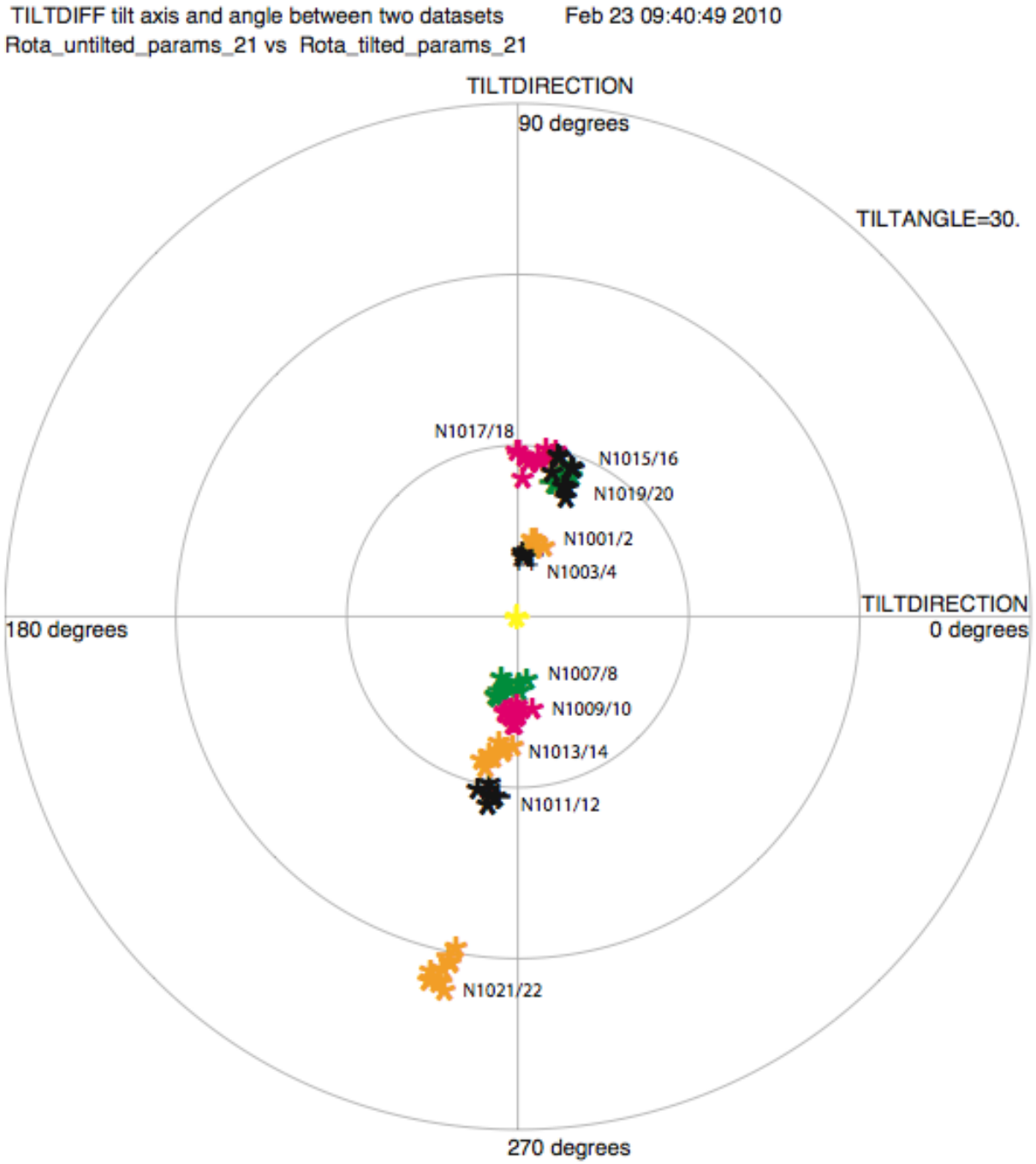


Fig. 4. FSC curves before (black) and after (red) 13-fold nonicosahedral averaging. The black curve suggests a resolution of 5.1 Å (0.143 threshold value), and the red curve indicates a resolution of 4.1 Å.




James Chen & Niko Grigorieff

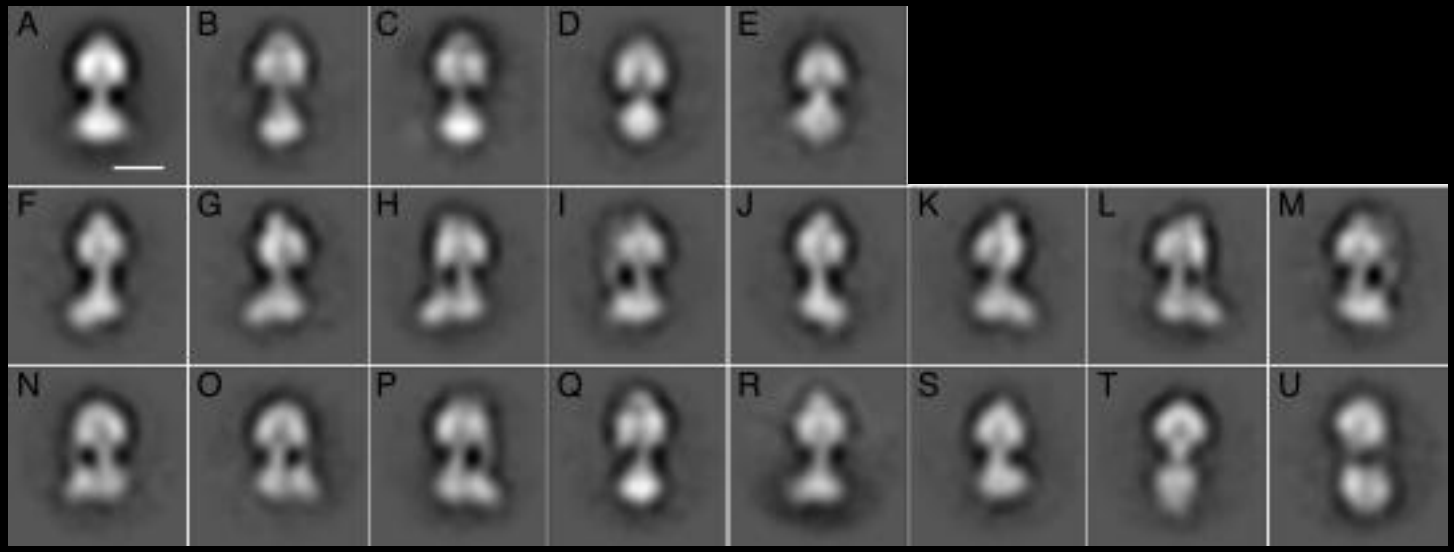
Rotavirus, James Chen & Niko Grigorieff, Brandeis, 2010



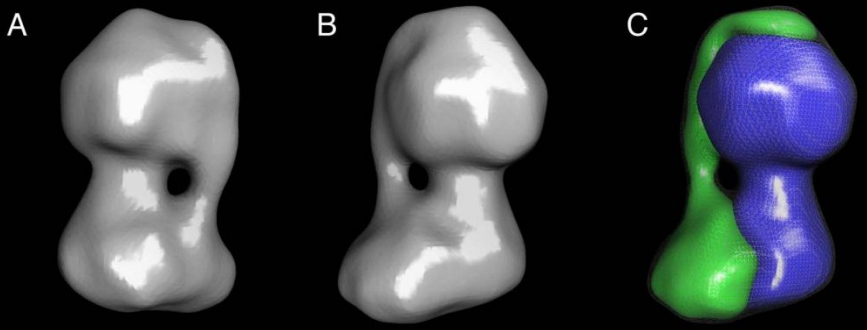
Film pair	<TANG> (sd)	Nom. TANG
N1001/2	+3.83 (±0.20)	+5.0
N1003/4	+4.50 (±0.21)	+5.0
N1007/8	-4.24 (±0.39)	-5.0
N1009/10	-5.67 (±0.33)	-5.0
N1011/12	-10.4 (±0.44)	-10.0
N1013/14	-8.07 (±0.63)	-10.0
N1015/16	+8.67 (±0.45)	+10.0
N1017/18	+9.34 (±0.53)	+10.0
N1019/20	+8.83 (±0.81)	+10.0
N1021/22	-21.14 (±0.95)	-10.0 (20.0?)

MW rotavirus = 20 MDa

Rubinstein, F_1F_o-ATP synthase EMBO J. (2003)



V-type ATPase, Lau & Rubinstein, PNAS (2010)



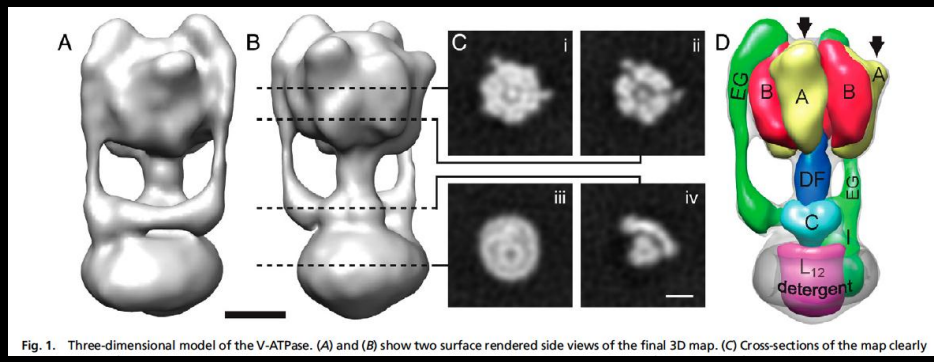
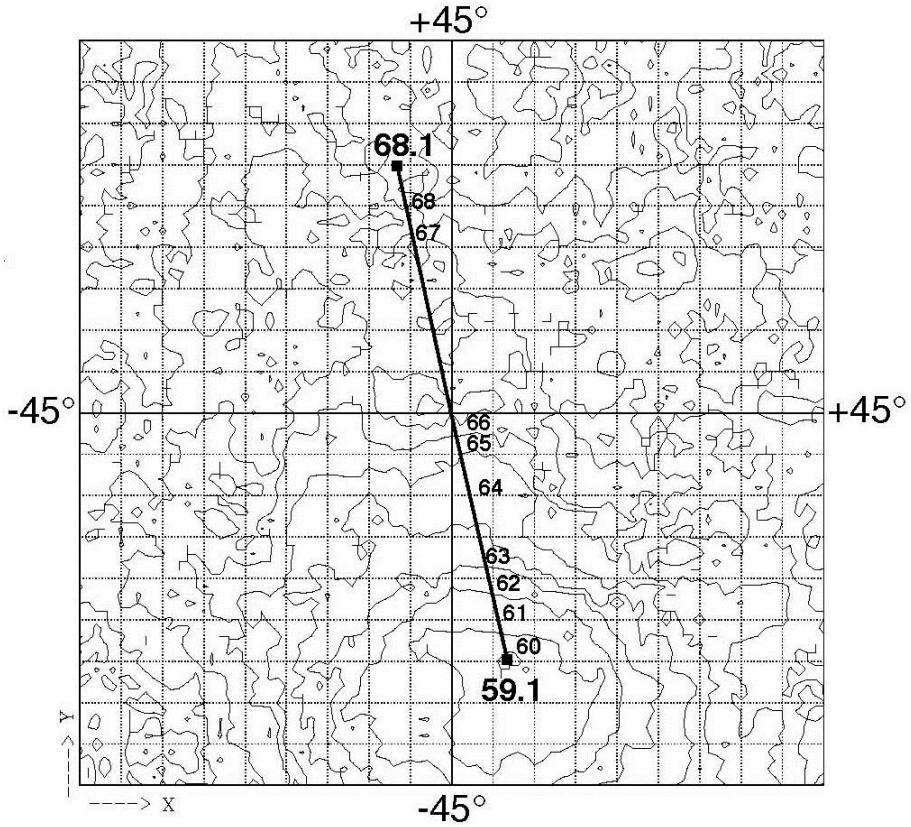


Fig. 1. Three-dimensional model of the V-ATPase. (*A*) and (*B*) show two surface rendered side views of the final 3D map. (*C*) Cross-sections of the map clearly

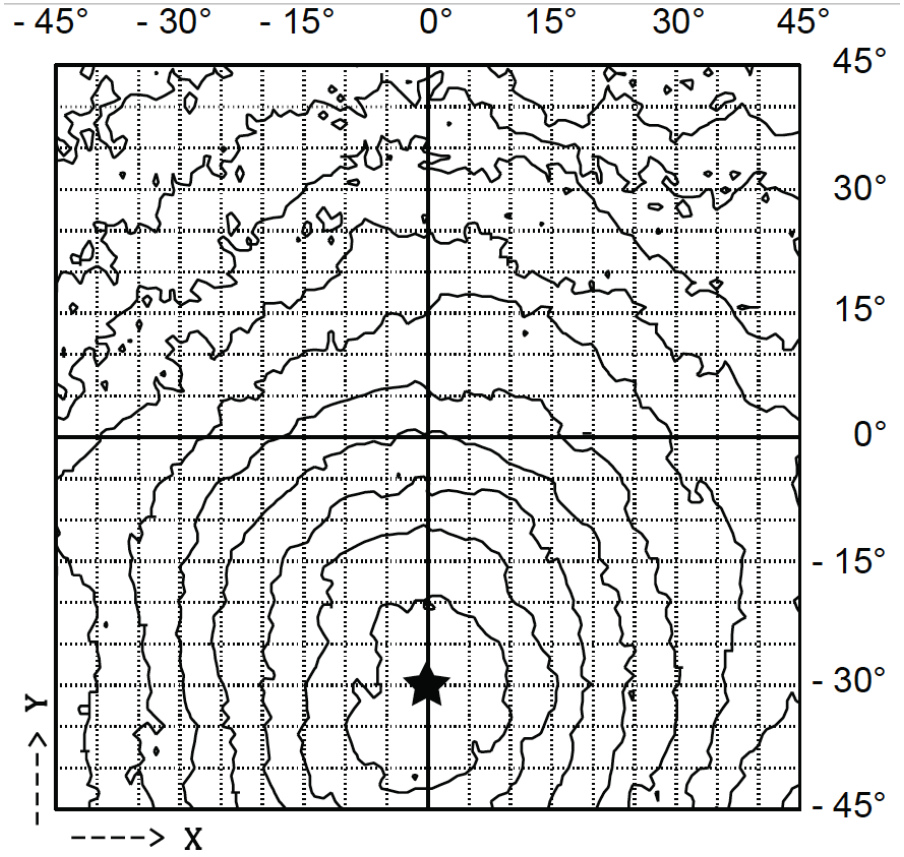
ATP-synthase, John Rubinstein, 2003 - 2010

Phase residual difference = 9.0°
Rubinstein et al, EMBO J. 2003



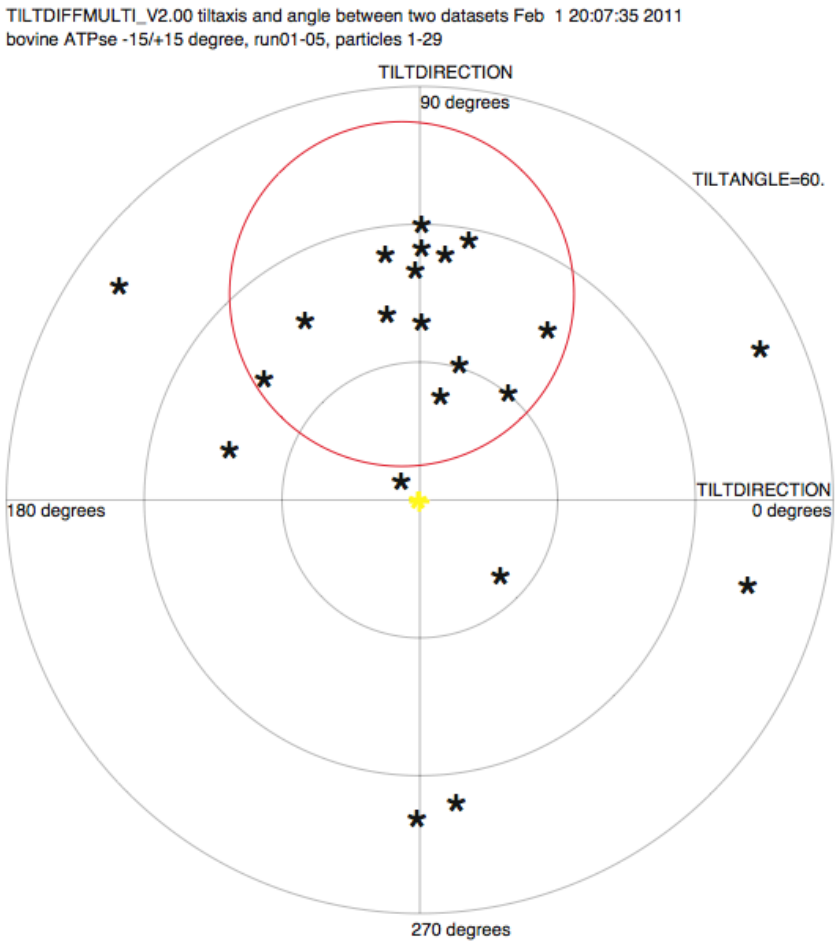
MW bovine F_1F_o = 600kDa

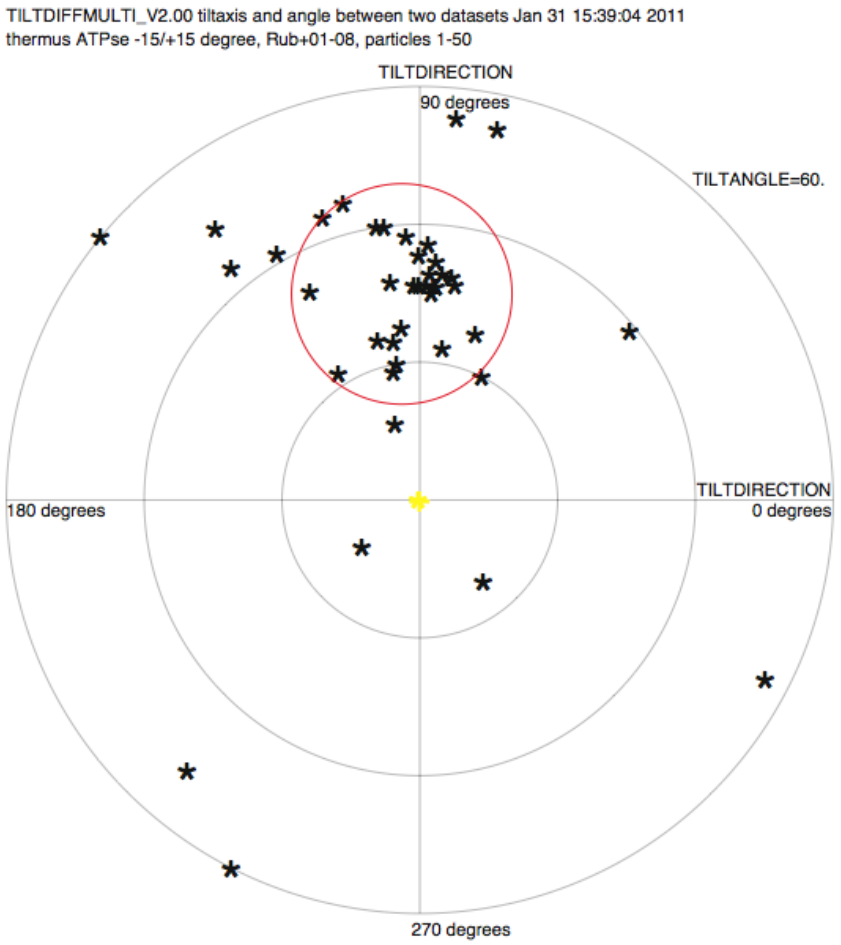
Phase residual difference = 14.9°
Lau et al, PNAS 2010



MW Thermus V_1V_o = 600kDa

ATP-synthase, John Rubinstein, 2010



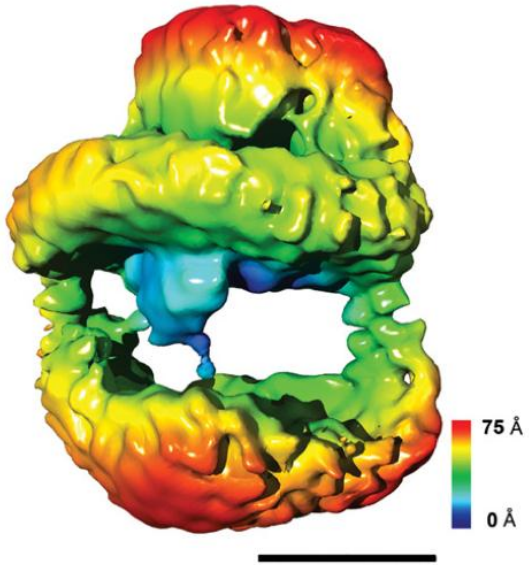


MW bovine F_1F_o = 600kDa

MW Thermus V_1V_o = 600kDa

Williams et al & Stewart
Structure (2008) 16, 468-477.

DNA-dependent protein kinase
~500kDa, 300,000 particles
7 Å resolution



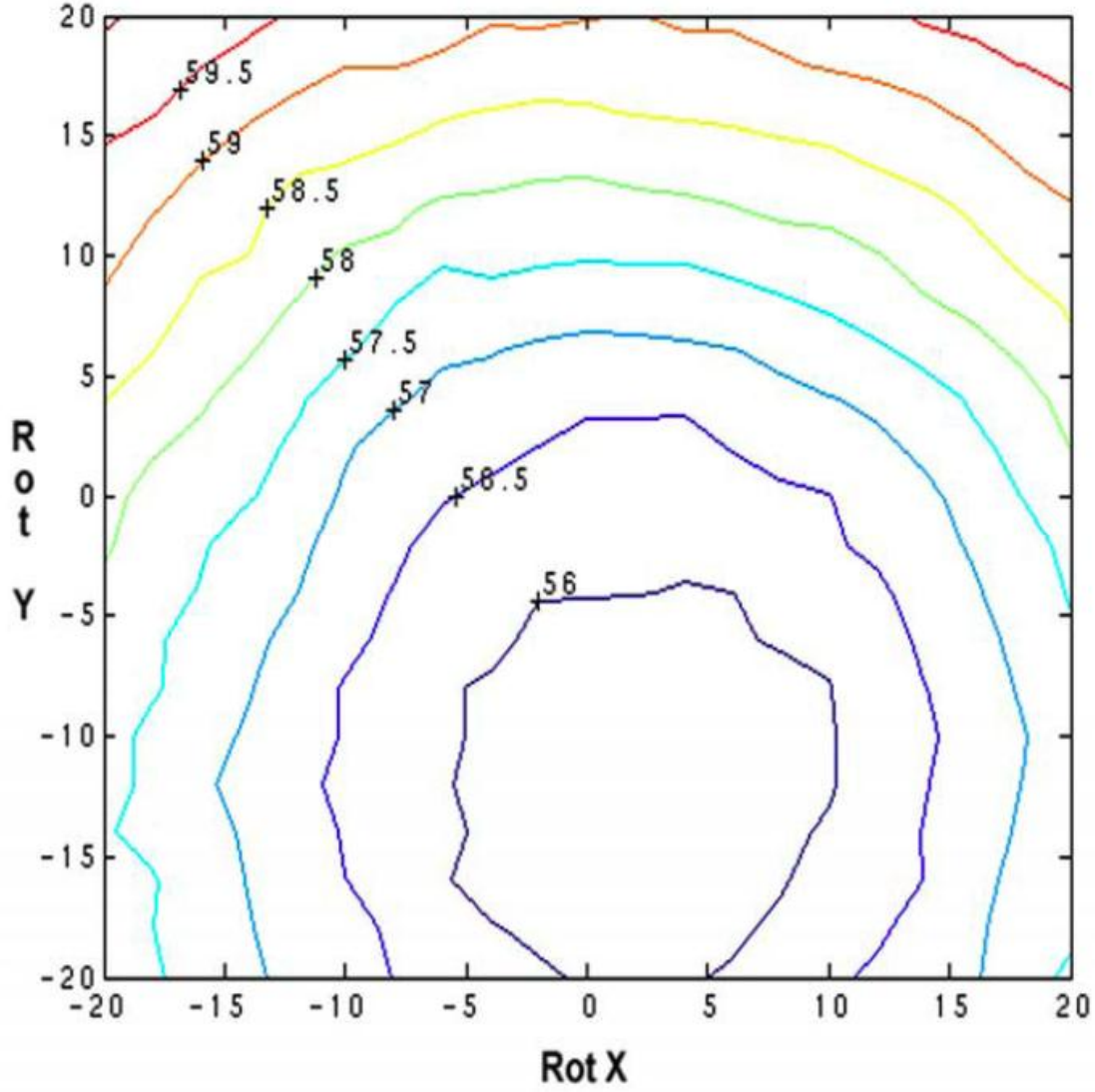
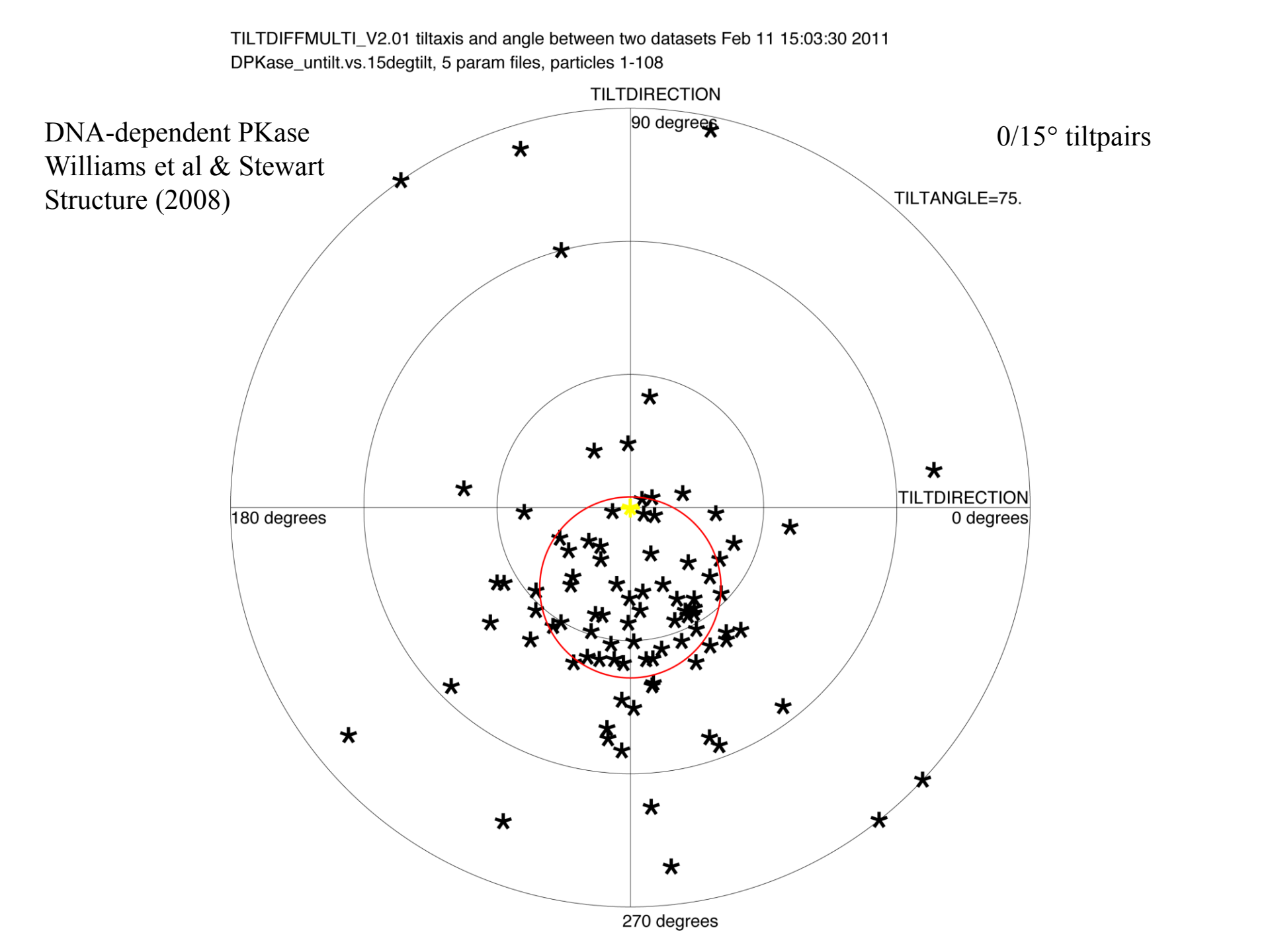


Figure S2. Determination of the Absolute Hand of DNA-PKcs


TILTDIFFMULTI_V2.01 tiltaxis and angle between two datasets Feb 11 15:03:30 2011
DPKase_untilt.vs.15degtilt, 5 param files, particles 1-108
TILTDIRECTION
90 degrees
TILTANGLE=75.
TILTDIRECTION
0 degrees
180 degrees
270 degrees
DNA-dependent PKase
Williams et al & Stewart
Structure (2008)
0/15° tiltpairs

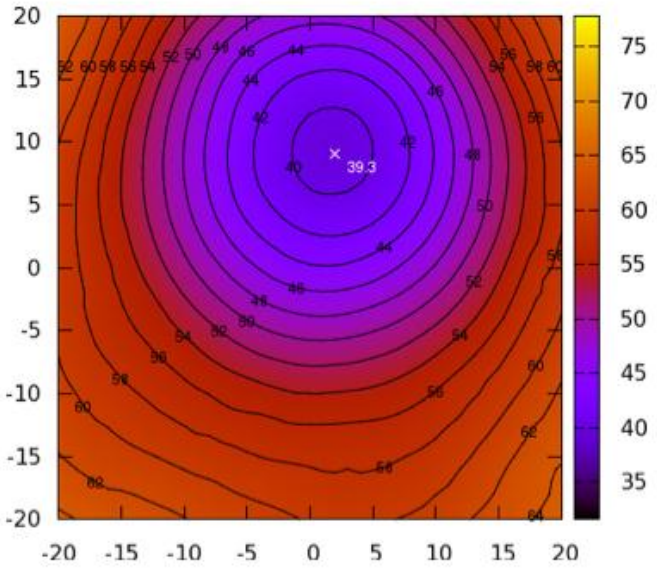
Peter Rosenthal and Sebastian Wasilewski

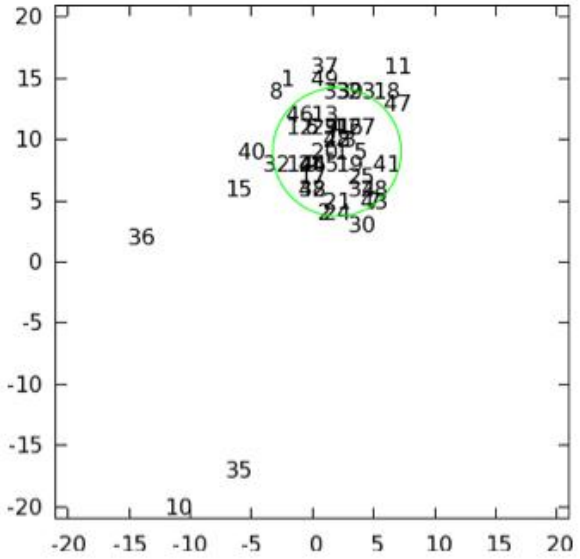
Tilt analysis report

Thu, 22 Jul 2010 02:02:18 +0100

Save as archive

3D model:	/home/swasile/Hand/combine_22av_halfp.map2k.mrc
Untilted stack:	/home/swasile/Hand/e2f301982.partpadred.mrc
Tilted stack:	/home/swasile/Hand/e2f301983.partpadred.mrc
Parameters file:	/home/swasile/Hand/e2_1982u_96.par
Experiment identifier:	Sample demo job





Magnification	4.98 A/px
Defocus	58626 ; 59084
Astigmatism	55.7
Voltage	300 kV
Resolution Interval	100.0 - 30.0 A
Tilt Interval	20
Particle radius	20 px
Optimized box size :	128
Effective binning:	1

Average for all particles submited:

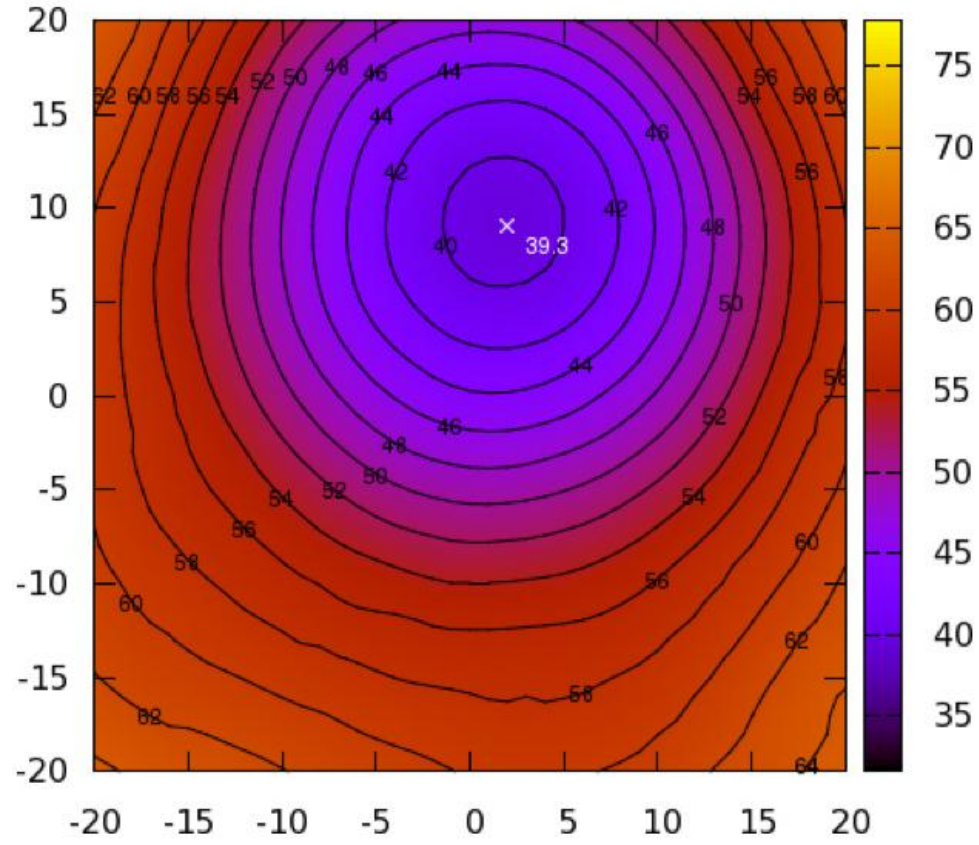
Minimal Phase Residue: **39.26 °**
Minimum at position: **2.0°, 9.0°**
Hand Phase Difference: **14.13 °**
Average distance from the mean minima: **5.25 °**

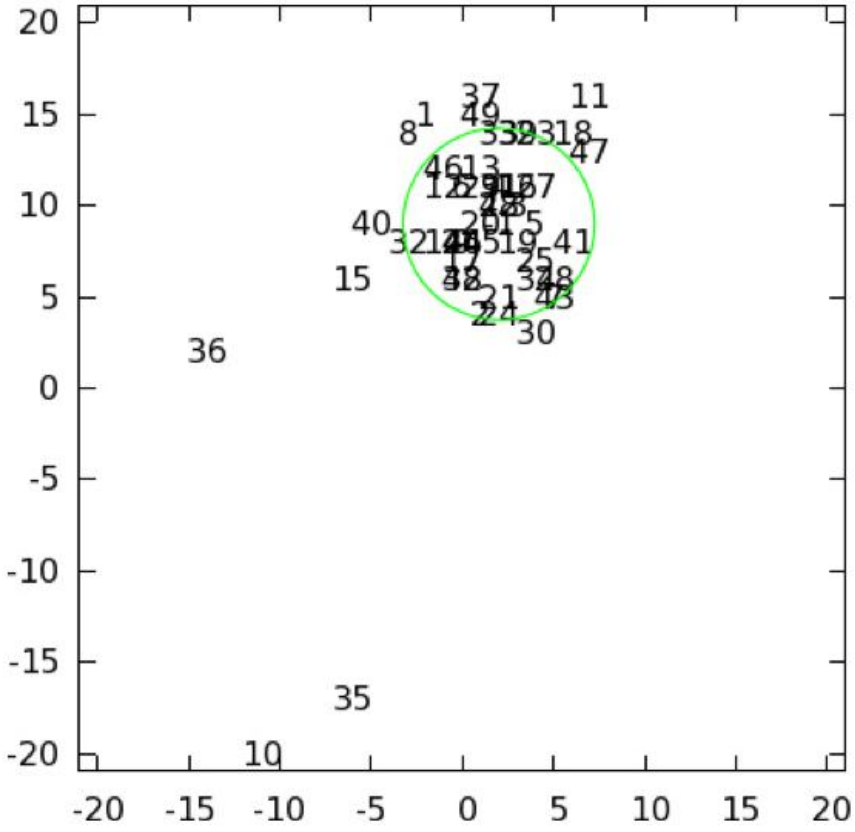
Particles with the hand difference below the average:
2 7 9 11 12 14 15 17 19 20 21 24 26 30 32 35 36 38 41 45

Particles with minima distant from the determined tilt transfromation:
1 8 10 11 15 18 23 30 35 36 37 40 47 49

Particles contributing to the determined minimum:
0 3 4 5 6 13 16 22 25 27 28 29 31 33 34 39 42 43 44 46 48

Peter Rosenthal and Sebastian Wasilewski (swasile@nimr.mrc.ac.uk)
http://www.cryomicroscopy.org/software/tilt-analysis-manual/
“Demo results page”





Conclusion - value of tilt pairs

- Works really well for big particles (20MDa); because the orientation determination is so accurate, it provides another piece of information about the magnitude of beam-induced specimen motion for particles in ice

- Works quite well for medium sized particles, but orientation determination has larger error bars (+/- 2-3º)

- For particles less than 1MDa, the success rate for orientation determination becomes less. More work is needed

Overview of tiltpair statistics

Specimen	Molecular Weight	Number of images	Number of particles	Successful alignment (%)	RMS clustering (degrees)
Rotavirus DLP	20 MDa	10	95	100	0.7
CAV	3.3 MDa	1	45	78	2.0
70S ribosomes	3.3 MDa	12	220	60	4.0
FAS	2.6 MDa	3	91	90	4.0
PDH E2CD	1.6 MDa	1	50	86	2.5
GroEL	0.8 MDa				
Thermus V-ATPase	0.6 MDa	1	50	58	10.0
Bovine F-ATPase	0.6 MDa	1	29	60	20.0
DNAdependent-PKase	0.47 MDa	[1]	108	70	15.0
β-galactosidase	0.45 MDa	2	119	81	10.0

Excerpts from X-ray VTF (2010): Read et al (to be published) *(plus RH comments)*

Validation arose as a major issue in the structural biology community when it became apparent that some published structures contained serious errors (Brändén and Jones, 1990). In response, the community developed a number of validation criteria, and tools to assess these criteria were implemented by the Protein Data Bank (PDB; Bernstein *et al.*, 1977; Berman *et al.*, 2000), which later expanded to become the Worldwide PDB (wwPDB; Berman *et al*., 2003).

Despite widespread use of the conventional validation tools, there are still isolated instances of high-profile *(Nature, Science)* structures that are entirely incorrect (Chang *et al*., 2006 - *retraction of EmrE, MsbA structures*), incorrect in their relevant details (Hanson and Stevens, 2009 - *botulinus toxin catalytic domain with imaginary peptide*), or likely fabricated (Janssen *et al.*, 2007 - *noted inconsistencies in a published C3b structure*).

Excerpts from EMVTF (Oct 2010)

Q1. How can map accuracy be assessed (both noise level and overall correctness)? How to estimate bias from model or overfitting noise? What statistics are useful?

It is clear that the community desires a validation method, or set of validation methods, for assessing the accuracy of cryoEM maps. Such a validation method does not yet exist, and this remains an open research problem. We mention below a few validation methods as examples, not intending in any way to represent all possibilities.

At high resolution (better than 4 Å) the model geometry and fit to the density map (R-factor between map and model) are good criteria, and there should be an encouragement to X-ray crystallography standards and practices. At lower resolution (20-4 Å), the situation is more complex and requires more care. At still lower resolutions (>20 Å), a simple pointspread function may be adequate, along with a statement of the RMS noise level, estimated from presumed featureless regions.

The absolute handedness of a structure cannot be determined without either a tilt experiment, or sufficient resolution to resolve chiral features directly in the map. Tilt experiments also offer the opportunity to provide validation for the accuracy of the structure as a whole, and can help place limits on orientation accuracy. Such methods include random-conical tilt (ref Radermacher), orthogonal tilt (ref Nogales), single particle tomography (ref Baumeister/Walz) and tilt-pair parameter plots (Rosenthal & Henderson, JMB, 2003).

Additional validation methods used in single-particle reconstruction include: ensuring agreement between projections of the 3D structure and raw images or (if generated) class-averages, ensuring that reference-free class-averages are fully represented among the set of model projections and ensuring sufficient coverage in particle orientations.

Q2. How should map resolution be reported?

Deposition should include the full FSC curve to Nyquist on a linear spatial frequency scale. This should be for the final map, as published. If the final experimental volume was masked in any way, FSC curves should be provided for both the masked and unmasked versions.

Q3. What density manipulation/filtering procedures were applied to the deposited map densities?

Examples include:

(a) Density stretching (e.g. negative density truncation)

(b) high- or low-pass filtering

(c) sharpening – what crystallographic B-factor or other sharpening function was applied?

(d) Was any signal-to-noise ratio weighting, cut-off or damping applied? For example, was FOM weighting used?

(e) What cropping or masking was used? We strongly encourage the deposition of a raw unfiltered, unmasked, unmodified 3D map, in addition to any modified maps that have been used in the associated publication.

Acknowledgements

- PDH — Peter Rosenthal
- CAV — Tony Crowther
- Rotavirus — James Chen, Niko Grigorieff
- 70S Ribosome — Lori Passmore
- DPKase — Phoebe Stewart
- FAS — Luciano Ciccarelli
- F-type ATPase — John Rubinstein
- V-type ATPase — Wilson Laue, John Rubinstein
- betaGal — Shaoxia Chen