UNIVERSITY of WISCONSIN

Crystallography without Crystals

Diffraction from Molecules with Single Axis Alignment

Dilano Saldin

dksaldin@uwm.edu

KITP Conference on "X-ray Science in the 21st Century"
Kavli Institute for Theoretical Physics,
Santa Barbara, CA, August 2-6, 2010

Main Collaborators

Hin-Cheuck Poon – University of Wisconsin-Milwaukee

Valentin Shneerson – University of Wisconsin-Milwaukee

Robin Santra – Argonne National Laboratory

John Spence – Arizona State University

Rick Kirian – Arizona State Univeristy

Kevin Schmidt – Arizona State University

Uwe Weierstall – Arizona State University

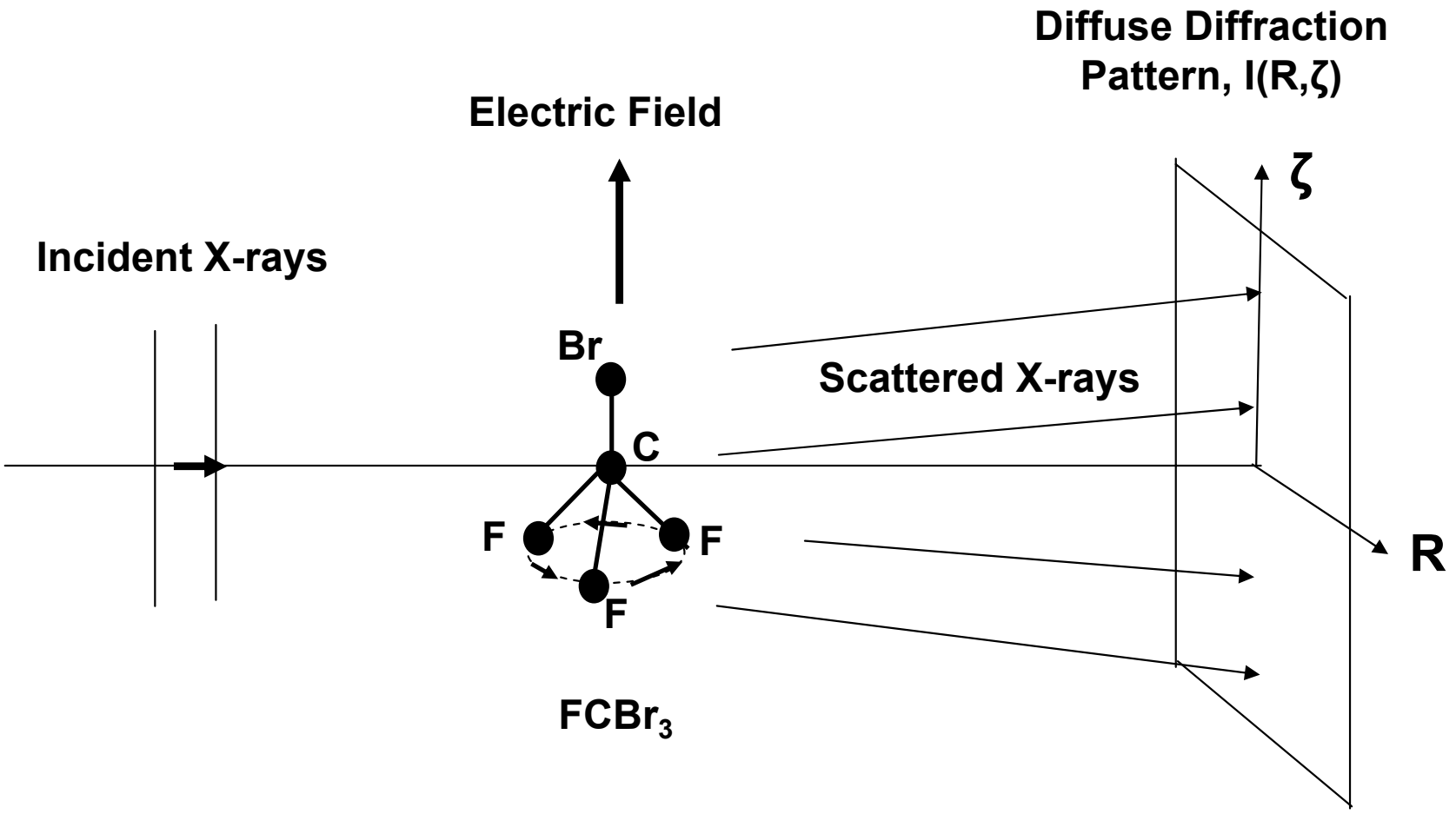
Henry Chapman – DESY

Malcolm Howells - LBNL

Outline

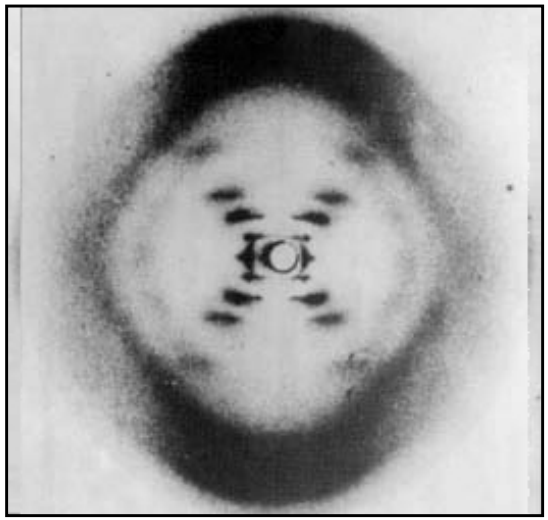
- Molecules with single-axis alignment have orientational order intermediate between molecules of completely random orientations and those forming a crystal
- Ensemble scattering: structure solution with reduced dose per particle - below threshold for radiation damage?
- If so, can bypass uncertainties associated with the effects on a diffraction pattern of rearrangments of electrons over the time-scale of a “diffract and destroy” experiment
- Beam incidence perpendicular to the alignment axis: use methods of fiber diffraction (famous example: DNA structure)
- Beam incidence parallel to the alignment axis: exploit angular correlations of intensities to reconstruct an oversampled DP of a single particle
- Structure from iterative phasing algorithms

Radiation Incident Perpendicular to the Alignment Axis

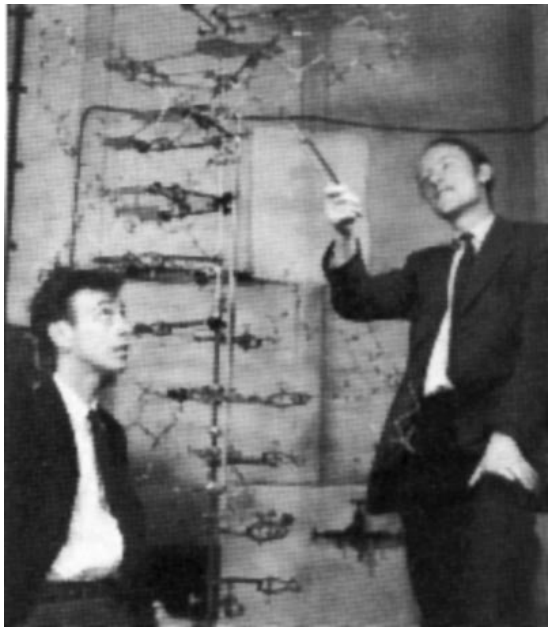


Symmetric-top molecule can be aligned along only a single axis even with elliptically polarized light

X-Ray Diffraction and DNA Structure



Diffraction pattern of DNA fibers (Franklin & Gosling, 1953)



Watson & Crick

Azimuthally Symmetric Intensities

For molecules aligned along a single axis, the random azimuthal orientations => access only to cylindrically symmetric (m=0) component of the cylindrically harmonic expansion of the intensities

$$I(R,\psi,\zeta)=\sum_m I_m(R,\zeta)\exp(im\psi)$$

All information contained in a single DP with incidence normal to the alignment axis. However,

$$I_0(R,\zeta)=\sum_m |G_m(R,\zeta)|^2$$

where this summation may have several non-zero components of amplitudes. If it is possible to deduce several of these components, can deduce non-cylindrical parts of the electron density. The only important terms are those for which $m<2\pi R_{max}r_{max}$. If molecule has M-fold rotational symmetry, only non-zero terms are those for which m=0, ±M, ±2M, etc. Thus, if $M<2\pi R_{max}r_{max}$, only non-zero term will be the azimuthally symmetric one, m=0

Scattering from Isolated Molecules

$$F(\mathbf{q}) = (2\pi)^{-3/2} \iiint d^3\mathbf{r} f(\mathbf{r}) \exp\{i\mathbf{q}.\mathbf{r}\}$$

$$F(\mathbf{q}) = \sum_m G_m(R,\zeta)\exp\{im\psi\} \qquad \mathbf{q} \equiv (R,\psi,\zeta)$$

$$f(\mathbf{r}) = \sum_m g_m(\tilde{r},z)\exp\{im\phi\} \qquad \mathbf{r} \equiv (r,\phi,z)$$

$$\exp\{i\mathbf{q}.\mathbf{r}\} = \exp\{i\zeta z + Rr\cos(\phi-\psi)\}$$

$$d^3\mathbf{r} = dr(rd\phi)dz$$

$$\int_0^{2\pi} \exp(iX\cos\phi)\exp\{in\phi\}d\phi = 2\pi i^n J_n(X)$$

Fourier-Bessel Transforms

$$g_m(r,z) = (2\pi)^{-1/2} \iint R G_m(R,\zeta)(-i)^m J_m(Rr) \exp\{-i\zeta z\} dR d\zeta$$

$$G_m(R,\zeta) = (2\pi)^{-1/2} \iint r g_m(r,z) i^m J_m(Rr) \exp\{i\zeta z\} dr dz$$

$$I(R,\psi,\zeta) = \sum_m I_m(R,\zeta)\exp(im\psi) \qquad I_0(R,\zeta) = \sum_m |G_m(R,\zeta)|^2 \approx |G_0(R,\zeta)|^2$$

$$f(r,\phi,z) = \sum_m g_m(r,z)\exp(im\phi)$$

If $|G_0(R,\zeta)|$ is known from experiment, $g_0(r,z)$, the azimuthally projected electron density, may be found by an iterative phasing algorithm on the m=0 components from constraints in real and reciprocal space. We start with random values of (the real quantity) $g_0(r,z)$

Saldin et al., Acta Cryst. A66, 32-37 (2010)

Azimuthally Projected Electron Density - C Nanotube

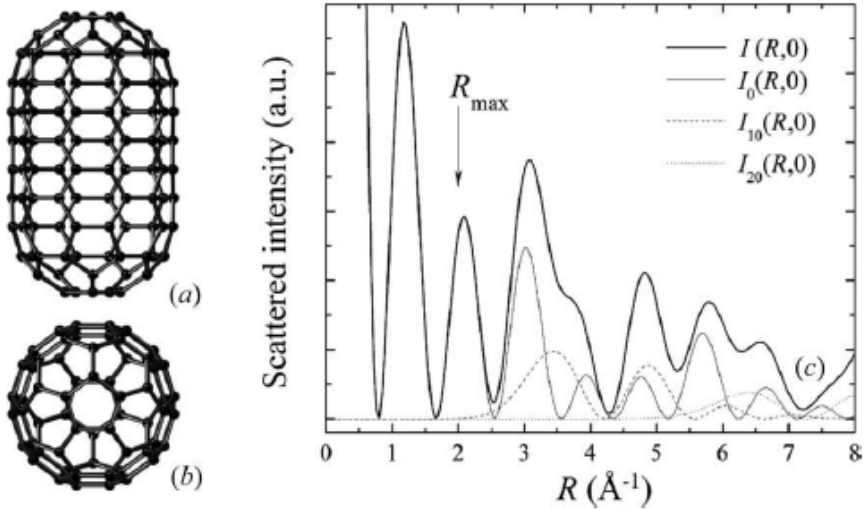


Figure 1
Schematic views of a short single-wall C nanotube, whose C—C bonds are of length ~1.4 Å, viewed (*a*) perpendicular to and (*b*) parallel to the molecular alignment axis. Panel (*c*) depicts the total diffraction pattern intensity as a function of the reciprocal-space coordinate R for $\zeta = 0$, as well as the magnitudes of the contributions from the $m = 0$, 10 and 20 cylindrical harmonics. R_{max} (= 2 Å^{-1}) indicates the maximum value of R corresponding to the edge of the simulated diffraction pattern in Fig. 3(*a*). The present figure shows that, up to $R = R_{max}$ (corresponding to $2\pi/R_{max} \simeq 3$ Å resolution), there is negligible error in modeling $I(R, 0)$ by just the $m = 0$ component.

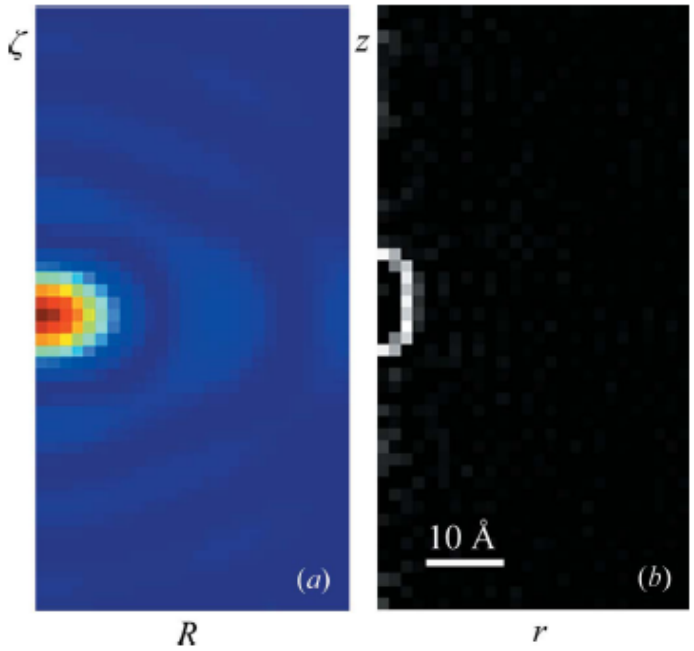


Figure 2
(*a*) Simulation of the diffraction pattern expected from a sum of those of a short SWNT of all azimuthal orientations about its axis (assumed perpendicular to the incident X-rays). The reciprocal-space coordinate parallel to the SWNT axis is denoted by ζ, while R is that perpendicular to ζ. The simulation assumes a flat Ewald sphere. The maximum values of R and $|\zeta|$ are 2 Å^{-1}, corresponding to a real-space resolution of about 3 Å. (*b*) Azimuthal projection of the electron density of the SWNT on a plane perpendicular to the tube axis, as reconstructed from the diffraction pattern in (*a*) by the algorithm described in the paper.

Heavy-Atom or "Holographic" Algorithm

In the case of a small molecule like CF_3Br, even if the resolution of the diffraction pattern is ~ 1 Å, the molecule is not much wider, so very tough to get atomic resolution with an iterative phasing algorithm. However one of the atoms of CF_3Br (the Br) is much heavier than the others. This suggests the possibility of using the heavy-atom (or a holographic) algorithm.

$$I_0(R,\zeta)=\sum_m |G_m(R,\zeta)|^2 \cong |G_0(R,\zeta)|^2 \cong |f_0|^2+\left\{\sum_{j\neq 0} f_0^* f_j \exp(i\zeta z_j) J_0(Rr_j)+c.c.\right\}$$

One-pass "holographic" reconstruction algorithm:

$$\rho_0(r,z)=\int RI(R,\zeta)\exp(-i\zeta z)J_0(Rr)dRd\zeta$$

since $$\rho_0(r,z)\approx |f_0|^2\delta(z)\delta(r)/r+\sum_{j\neq 0}\delta(r-r_j)\{f_0^* f_j\delta(z-z_j)+f_0 f_j^*\delta(z+z_j)\}/r$$

because $$\int \exp(\pm i\zeta z_j)\exp(i\zeta z)d\zeta=\delta(z\pm z_j)$$

and $$\int J_m(Rr_j)J_m(Rr)RdR=\delta(r-r_j)/r$$

Holographic Reconstruction

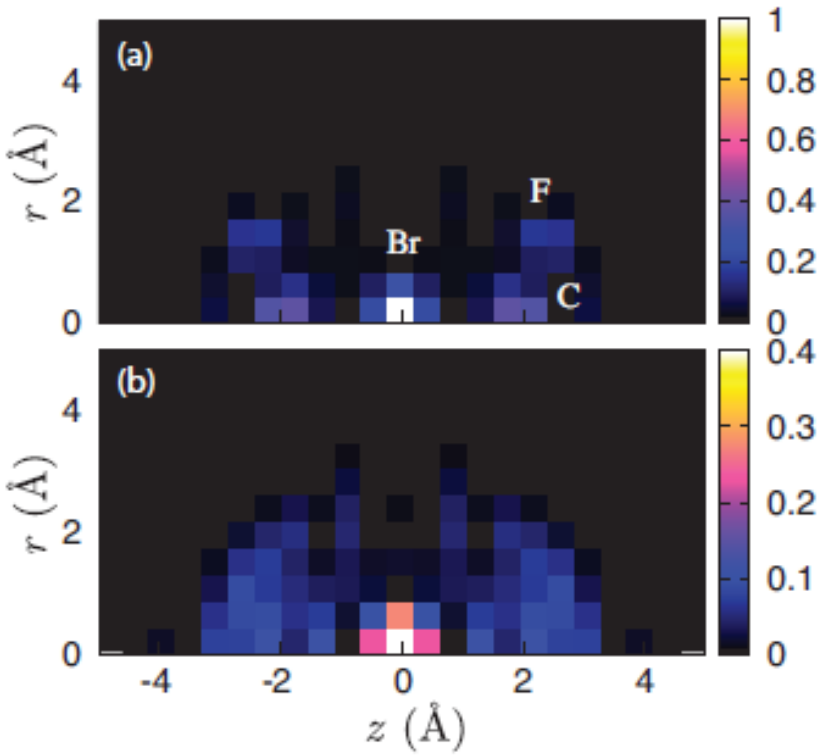
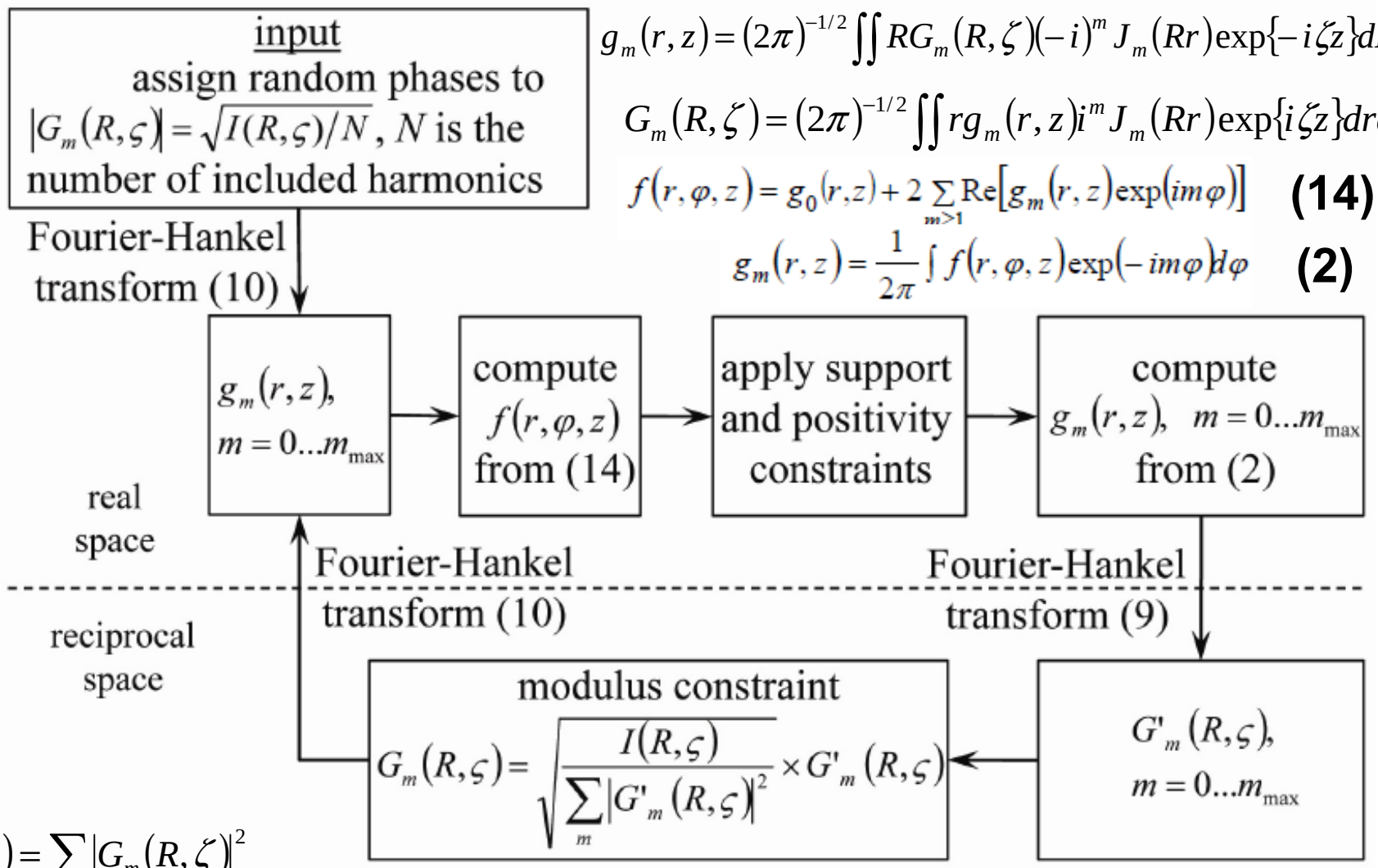


FIG. 3. The azimuthally projected structure $\tilde{\rho}_0(r,z)$ of CF_3Br, holographically reconstructed from the simulated diffraction patterns of Fig. 2. (a) Perfect 1D alignment and (b) laser-induced alignment.

Ho et al., J. Chem. Phys. 131, 131101 (2009)

Beyond Azimuthal Projection



$$g_m(r,z) = (2\pi)^{-1/2} \iint RG_m(R,\zeta)(-i)^m J_m(Rr)\exp\{-i\zeta z\}dRd\zeta \quad (10)$$

$$G_m(R,\zeta) = (2\pi)^{-1/2} \iint rg_m(r,z)i^m J_m(Rr)\exp\{i\zeta z\}drdz \quad (9)$$

$$f(r,\varphi,z) = g_0(r,z) + 2\sum_{m>1} \mathrm{Re}\left[g_m(r,z)\exp(im\varphi)\right] \quad (14)$$

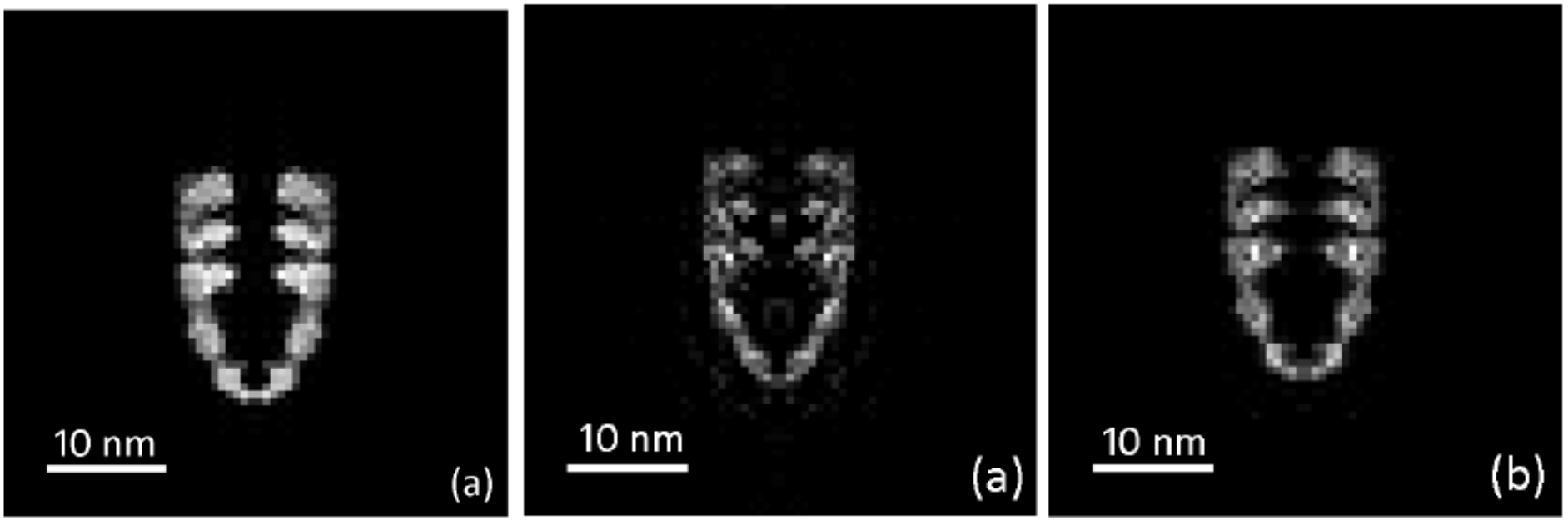
$$g_m(r,z) = \frac{1}{2\pi}\int f(r,\varphi,z)\exp(-im\varphi)d\varphi \quad (2)$$

$$I(R,\zeta) = \sum_m |G_m(R,\zeta)|^2$$

Figure 1. Block diagram of the phasing algorithm.

Starodub et al., SPIE Proc. (2010)

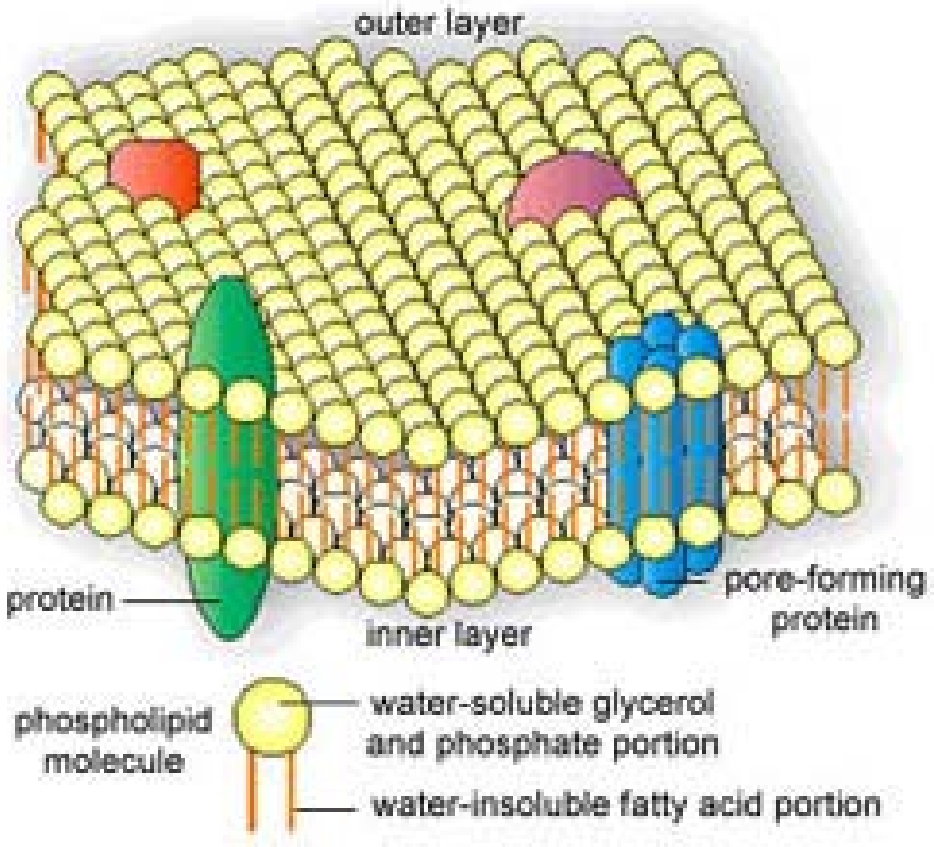
Improvement of Azimuthal Projection with Inclusion of m≠0 Components



Model **m=0 only** **m=0, ±7**

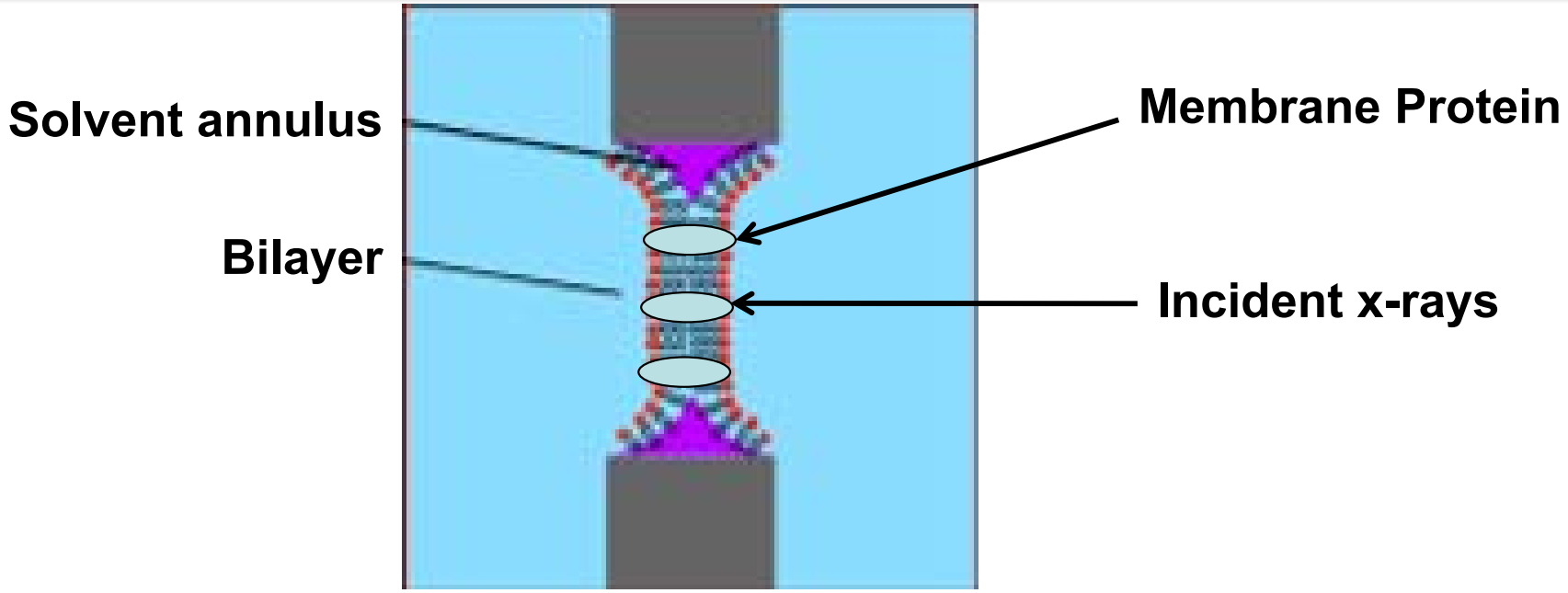
Azimuthal projection of GroEL-GeoES protein complex

Incidence Parallel to Alignment Axis – E.g. Membrane Proteins *In Situ*



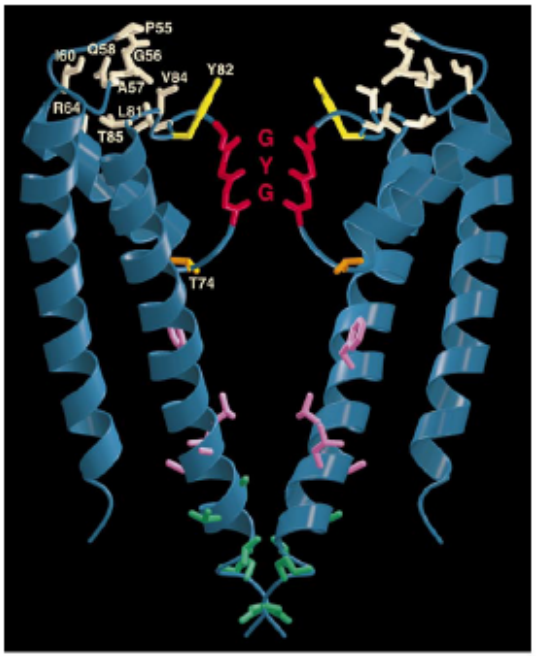
E.g. Ion-channel proteins will be in random locations in membrane and in random orientations about an axis perpendicular to the membrane.

Proposed Experiment: Black Lipid Membrane



Small aperture created in a hydrophobic material such as Teflon. Solution of lipids dissolved in an organic solvent applied by a brush or syringe across aperture.

By this means can create a bilipid membrane with membrane proteins inserted. The proteins will be in random positions in the plane of the membrane and at random orientations about the membrane normal. The aim is to determine the structure of each individual protein from the resulting multi-protein diffraction pattern



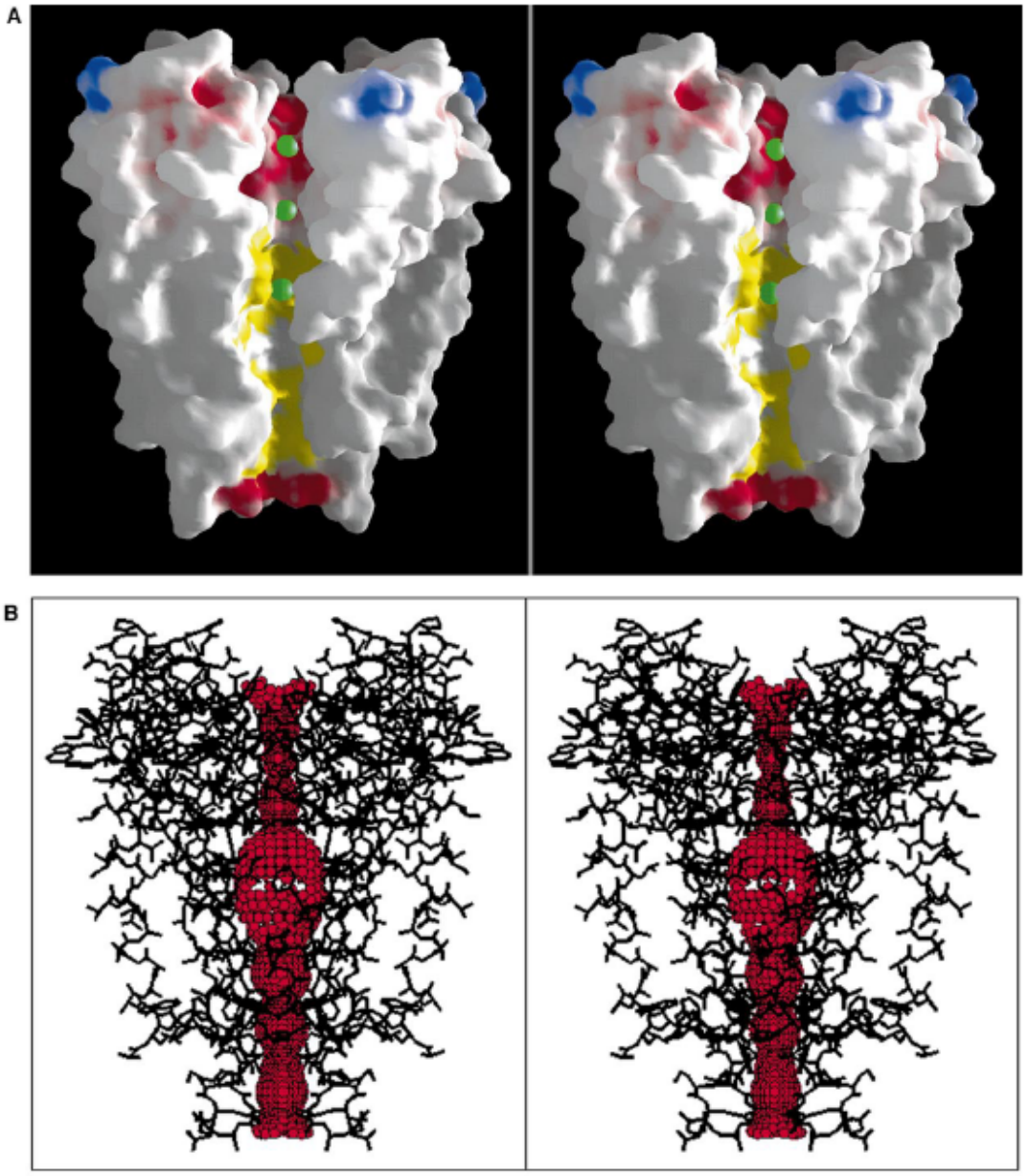
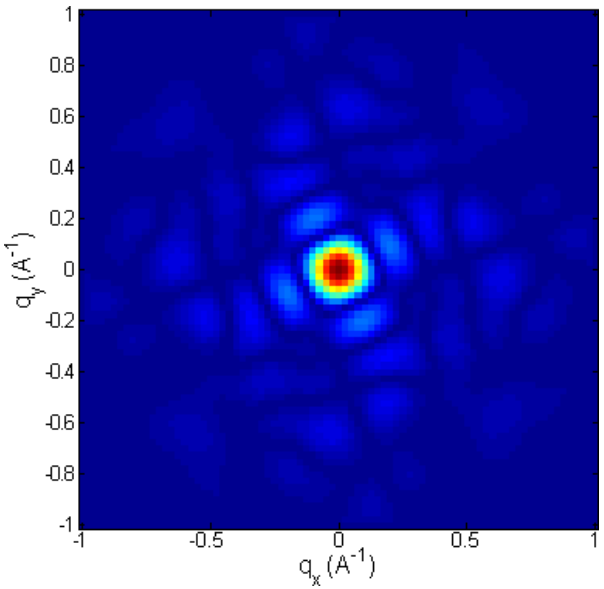


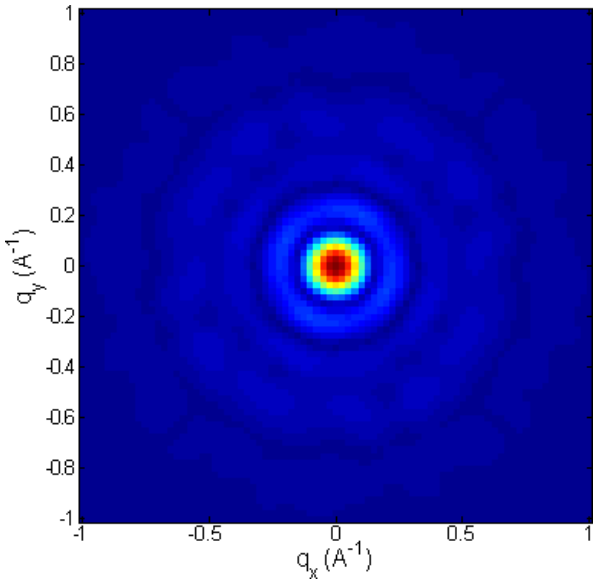
Fig. 4 (above). Mutagenesis studies on *Shaker*: Mapping onto the KcsA structure. Mutations in the voltage-gated *Shaker* K^+ channel that affect function are mapped to the equivalent positions in KcsA based on the sequence alignment. Two subunits of KcsA are shown. Mutation of any of the white side chains significantly alters the affinity of agitoxin2 or charybdotoxin for the *Shaker* K^+ channel (*12*). Changing the yellow side chain affects both agitoxin2 and TEA binding from the extracellular solution (*14*). This residue is the external TEA site. The mustard-colored side chain at the base of the selectivity filter affects TEA binding from the intracellular solution [the internal TEA site (*15*)]. The side chains colored green, when mutated to cysteine, are modified by cysteine-reactive agents whether or not the channel gate is open, whereas those colored pink react only when the channel is open (*16*). Finally, the residues colored red (GYG, main chain only) are absolutely required for K^+ selectivity (*4*). This figure was prepared with MOLSCRIPT and RASTER-3D. **Fig. 5 (right).** Molecular surface of KcsA and contour of the pore. (**A**) A cutaway stereoview displaying the solvent-accessible surface of the K^+ channel colored according to physical properties. Electrostatic potential was calculated with the program GRASP, assuming an ionic strength equivalent to 150 mM KCl and dielectric constants of 2 and 80 for protein and solvent, respectively. Side chains of Lys, Arg, Glu, and Asp residues were assigned single positive or negative charges as appropriate, and the surface coloration varies smoothly from blue in areas of high positive charge through white to red in negatively charged regions. The yellow areas of the surface are colored according to carbon atoms of the hydrophobic (or partly so) side chains of several semi-conserved residues in the inner vestibule (Thr^{75}, Ile^{100}, Phe^{103}, Thr^{107}, Ala^{108}, Ala^{111}, Val^{115}). The green CPK spheres represent K^+ ion positions in the conduction pathway. (**B**) Stereoview of the entire internal pore. Within a stick model of the channel structure is a three-dimensional representation of the minimum radial distance from the center of the channel pore to the nearest van der Waals protein contact. The display was created with the program HOLE (*34*).

Single-Particle DP from Multiparticle DPs
Particles Frozen in Space or Time

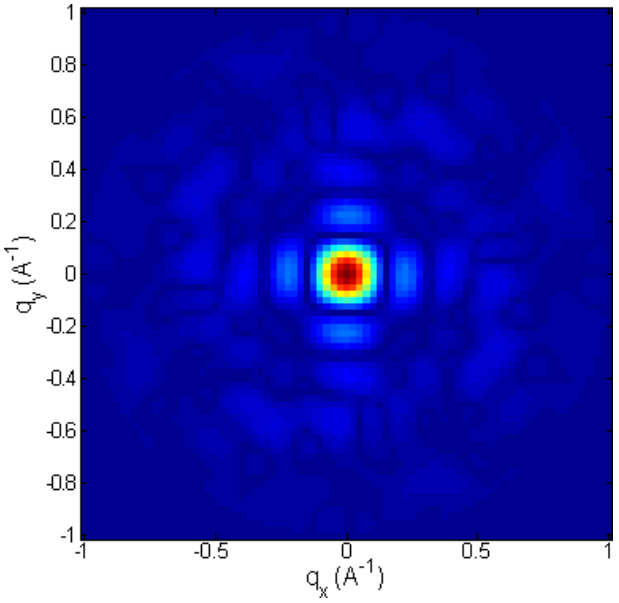
Use radiation with pulse length shorter than rotational diffusion time, or freeze the particles in e.g. an ice sheet.
Application to ion-channel membrane protein in situ



Original DP



10-Particle DP

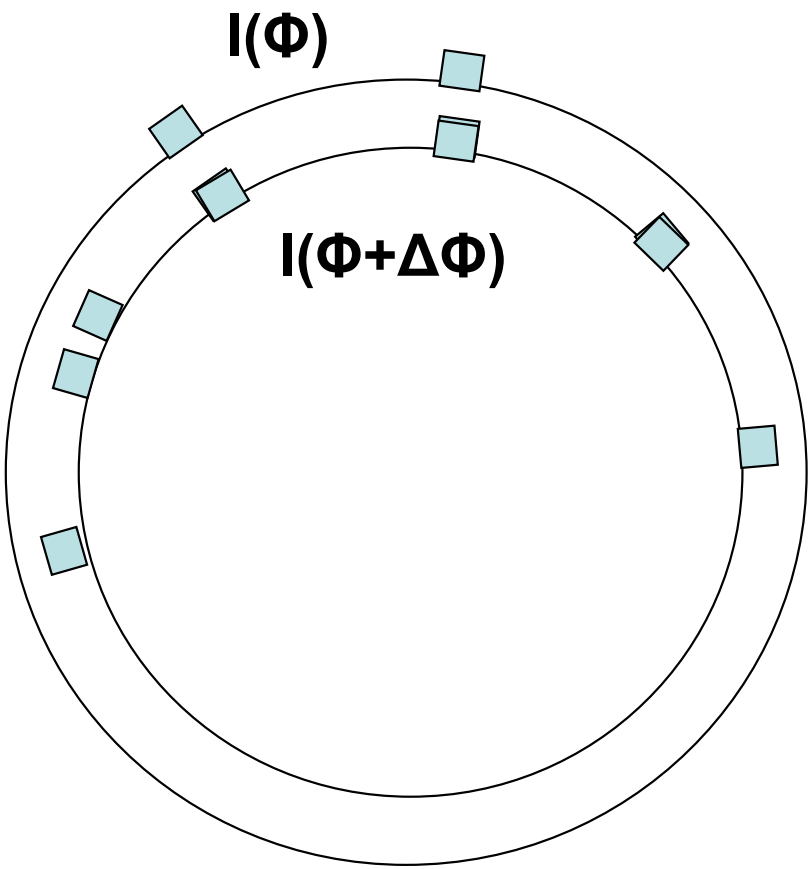


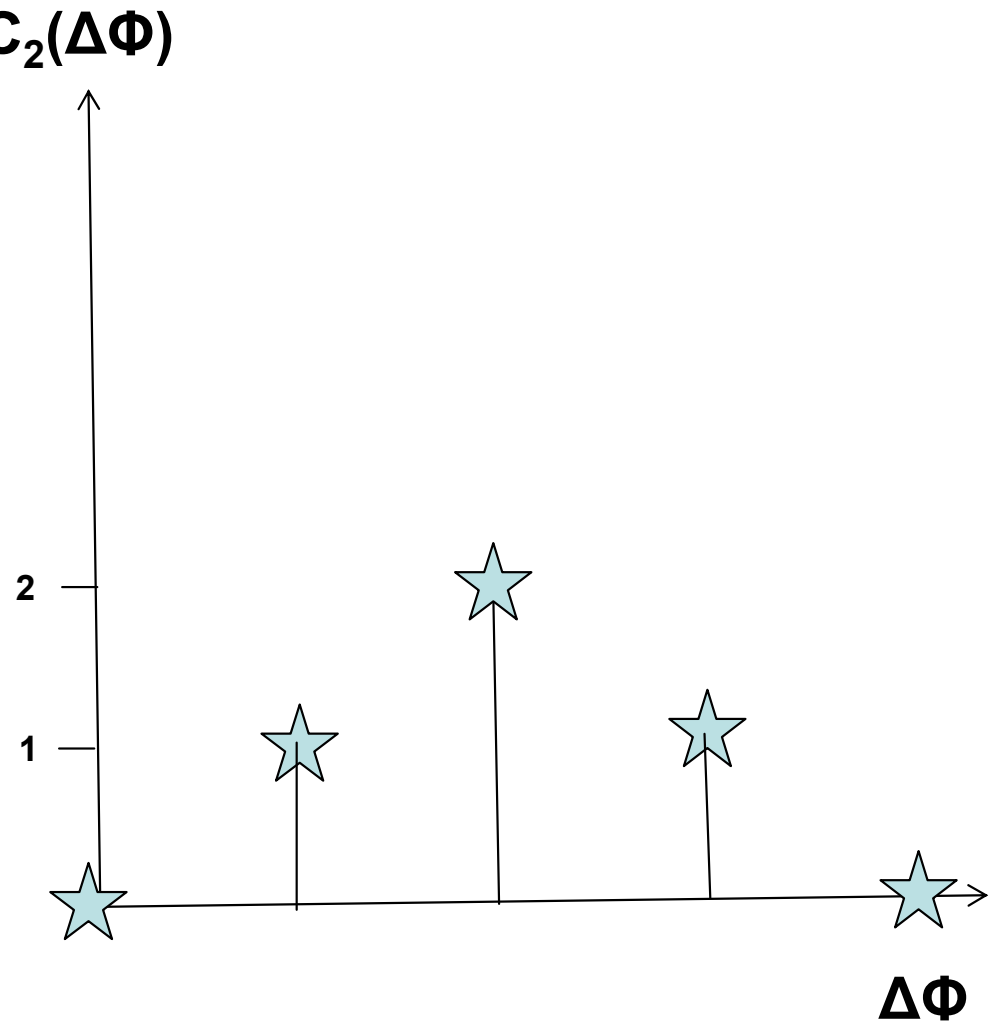
Reconstructed Single-Particle DP

Like unscrambling an egg

Saldin et al., Phys. Rev. B 81, 174105 (2010)

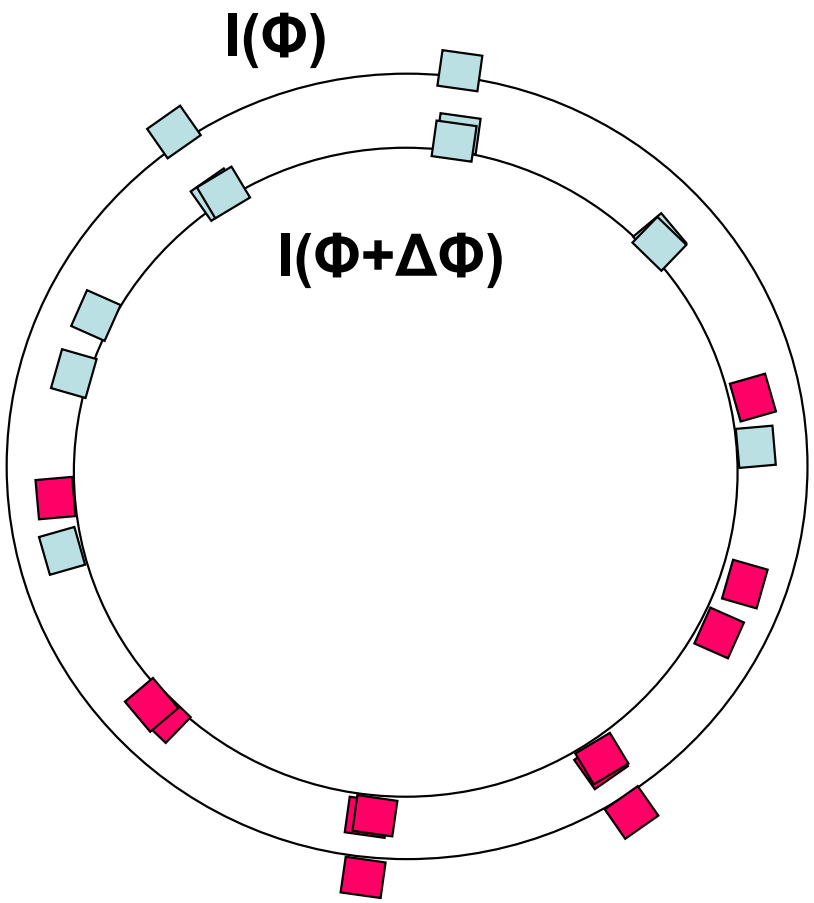
Autocorrelation of DP of a single particle

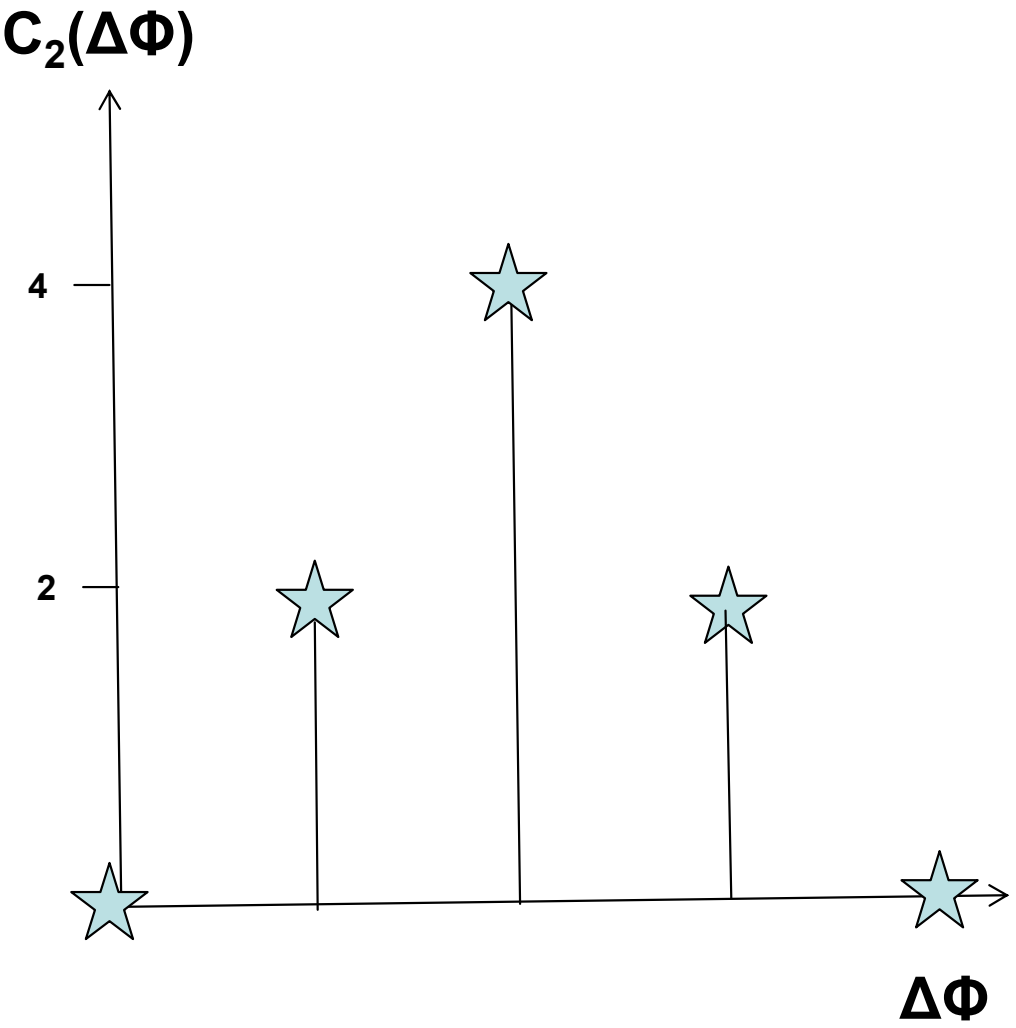




$$C_2(\Delta\Phi) = \sum_{\Phi} I(\Phi) I(\Phi + \Delta\Phi)$$

Autocorrelation of superposed DPs of N=2 particles in different orientations





$$C_2(\Delta\Phi) = \sum_{\Phi} I(\Phi) I(\Phi + \Delta\Phi)$$

$C_2 \propto N$ **but otherwise identical**

Reconstructing the Single-Particle Diffraction Pattern

Circular Harmonic Expansion

$$I(q_x, q_y) = \sum_m I_m(q) \exp(im\phi)$$

where

$$q = \sqrt{q_x^2 + q_y^2} \qquad \phi = \tan^{-1}(q_y, q_x)$$

Can be done once the magnitudes of $I_m(q)$ are found from the pair correlations and their signs are found from the triple correlations

The reality of $I(\mathbf{q})$ ensures that $I_{-m}(q)=I_m(q)$. For a flat Ewald sphere, Friedel's rule, $I(-\mathbf{q})=I(\mathbf{q})$, will be satisfied if only even m's contribute.

If the single particle diffraction pattern has a mirror line, can choose the $I_m(q)$ to be real.

Saldin et al., New J. Phys. **12**, 035014 (2010)

Relation to Scattered Amplitudes

$$A(q,\varphi_j) = \sum_k \sum_m A_m(q) \exp\{im(\varphi_j - \omega_k)\} \exp(i\mathbf{q}.\mathbf{r}_k)$$

$$C_2(q;q',\Delta\varphi) = \left\langle \frac{1}{N_\varphi} \sum_j \{I(q,\varphi_j) - I_{saxs}(q)\}\{I(q',\varphi_j + \Delta\varphi) - I_{saxs}(q')\} \right\rangle_{DP}$$

$$\begin{aligned}
&I(q,\varphi_j) - I_{saxs} \\
&= \sum_{kk'} \sum_{m \neq m'} A_m^*(q) \exp(-im\varphi_j) \exp(im\omega_k) \exp(-i\mathbf{q}.\mathbf{r}_k) A_{m'}(q) \exp(im'\varphi_j) \exp(-im'\omega_{k'}) \exp(i\mathbf{q}.\mathbf{r}_{k'}) \\
&= \sum_{kk''} \sum_{m \neq m'} A_m^*(q) A_{m'}(q) \exp\{-i(m-m')\varphi_j\} \exp(im\omega_k) \exp(-im'\omega_{k'}) \exp\{-i\mathbf{q}.(\mathbf{r}_k - \mathbf{r}_{k'})\} \\
&= \sum_{m,M \neq 0} A_m^*(q) A_{m-M}(q) \exp\{-iM\varphi_j\} \sum_{kk'} \exp(im\omega_k) \exp(-i(m-M)\omega_{k'}) \exp\{-i\mathbf{q}.(\mathbf{r}_k - \mathbf{r}_{k'})\}
\end{aligned}$$

By performing the sums over *amplitudes* we are assuming *coherent* scattering amongst the different particles

Pair Correlations in Terms of Amplitudes

$$I(q',\varphi_j+\Delta\varphi_k)-I_{saxs}$$
$$=\sum_{m',M'\neq 0}A_{m'}(q')A^*_{m'-M'}(q')\exp\{iM'(\varphi_j+\Delta\varphi_l)\}\sum_{k"k'''}\exp(-im'\omega_{k"})\exp(i(m'-M')\omega_{k'''})\exp\{i\mathbf{q}'.(\mathbf{r}_{k"}-\mathbf{r}_{k'''})\}$$

$$C_2(q;q',\Delta\varphi)=\left\langle \begin{array}{l} \sum_{m.M\neq 0}A^*_m(q)A_{m-M}(q)\ldots \\ \sum_{m',M\neq 0}A_{m'}(q')A^*_{m'-M'}(q')\exp\{iM\Delta\varphi_l\}\ldots \\ \sum_{kk'k"k'''}\exp(im\omega_k)\exp(-i(m-M)\omega_{k'})\exp\{-i\mathbf{q}.(\mathbf{r}_k-\mathbf{r}_{k'})\}\ldots \\ \exp(-im'\omega_{k"})\exp(i(m'-M)\omega_{k'''})\exp\{i\mathbf{q}'.(\mathbf{r}_{k"}-\mathbf{r}_{k'''})\} \end{array} \right\rangle_{DP}$$

Note that the average over DPs removes *all dependence on the orientation and position of the particles*, despite the fact that coherent scattering was assumed (only the diagonal terms in the sums over the particles survive)

$$C_2(q,q',\Delta\varphi)=N_p\sum_{m,M\neq 0}A_m(q)A^*_{m-M}(q)\sum_{m',M\neq 0}A^*_{m'}(q')A_{m'-M}(q')\exp\{iM\Delta\varphi_l\}$$
$$=N_p\sum_{M\neq 0}I^*_M(q)I_M(q')\exp\{iM\Delta\varphi_l\}$$

The correlations depend only on the *single-particle* quantities $I_m(q)$

Magnitude of Expansion Coeffs.

Pair Correlations (averaged over many short-pulse DPs)

$$C_2(q;q',\Delta\varphi_l) = \left\langle \frac{1}{N_\varphi}\sum_j \left\{I(q,\varphi_j) - I_{saxs}(q)\right\}\left\{I(q',\varphi_j+\Delta\varphi_l) - I_{saxs}(q')\right\}\right\rangle_t$$

$$= N_p \sum_{M\neq 0} I_M^*(q) I_M(q') \exp(iM\Delta\varphi_l)$$

FT of $C_2(q,q';\Delta\varphi_l)$

$$B_M(q,q') \equiv \frac{1}{N}\sum_{l=1}^{N} C_2(q,q;\Delta\varphi_l)\exp(-iM\Delta\varphi_l) = I_M(q) I_M^*(q')$$

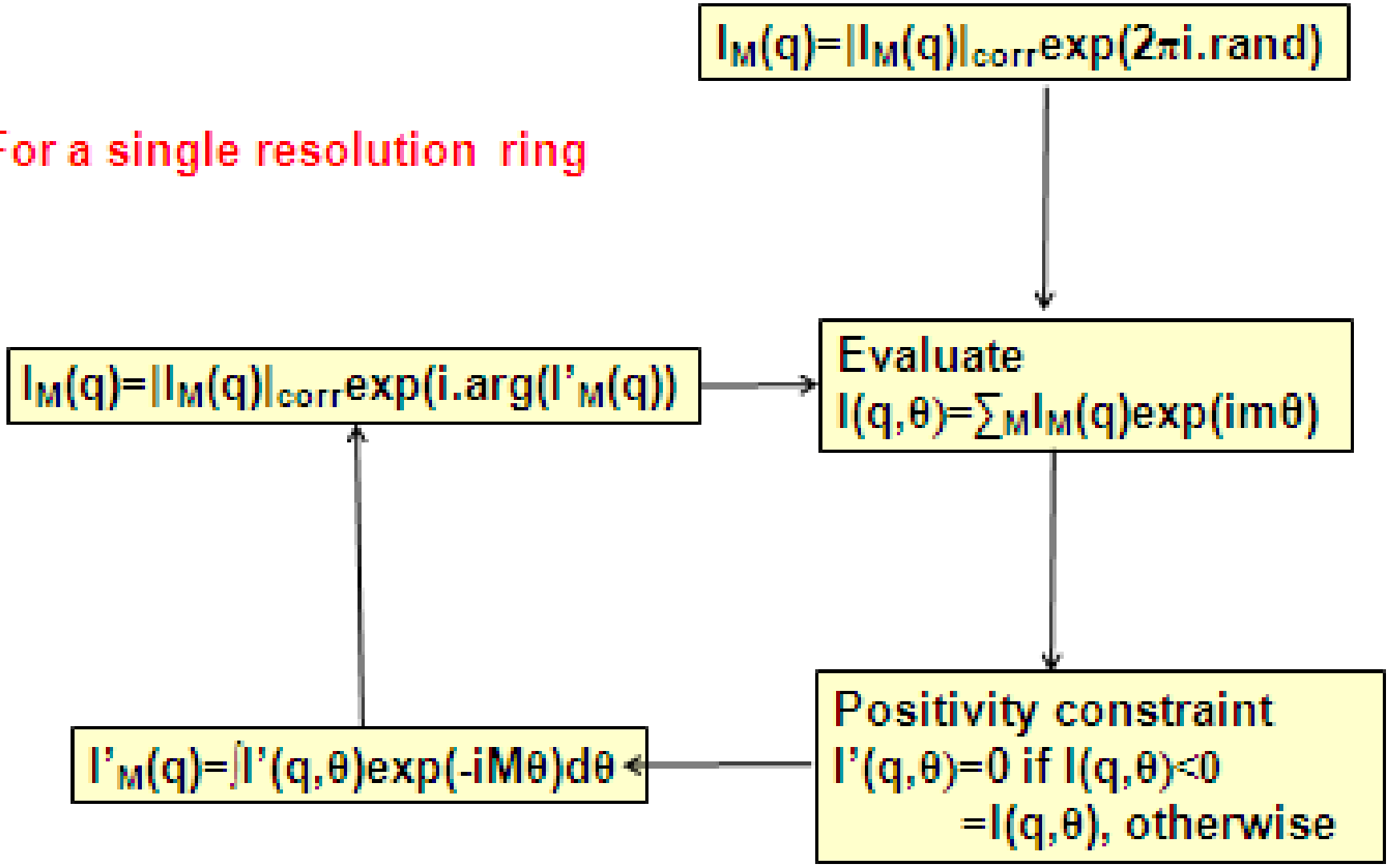
Magnitude of expansion coefficients from the FT of the autocorrelations:

$$\left|I_M(q)\right| = \sqrt{B_M(q,q)}$$

The non-uniqueness of the square root is manifested by the unknown phases, which need to be determined by something which is sensitive to the phases.

Phases of Expansion Coeffs.:
Positivity Constraint

For a single resolution ring



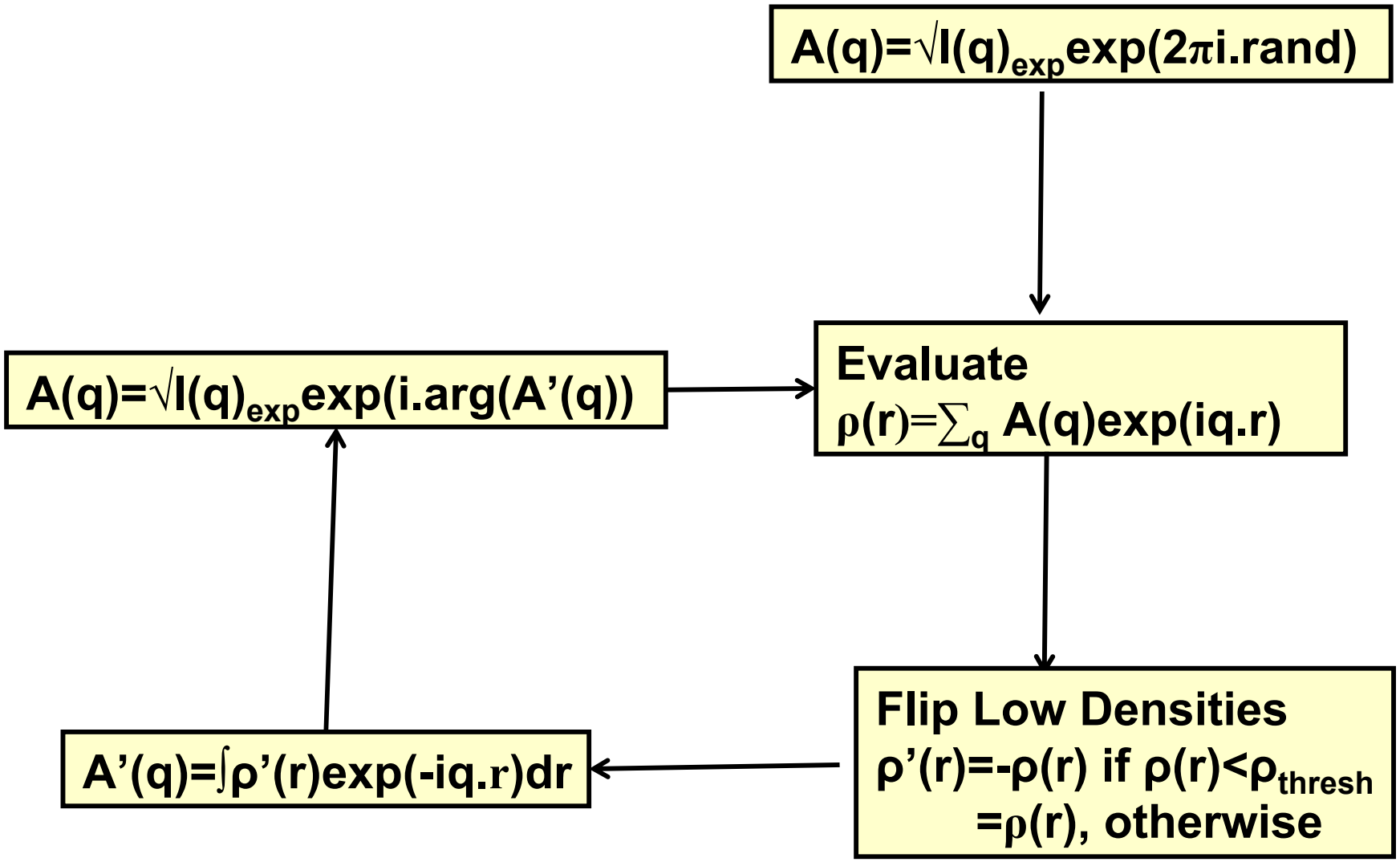
For Other Resolution Rings

Phases for other resolution rings follows from

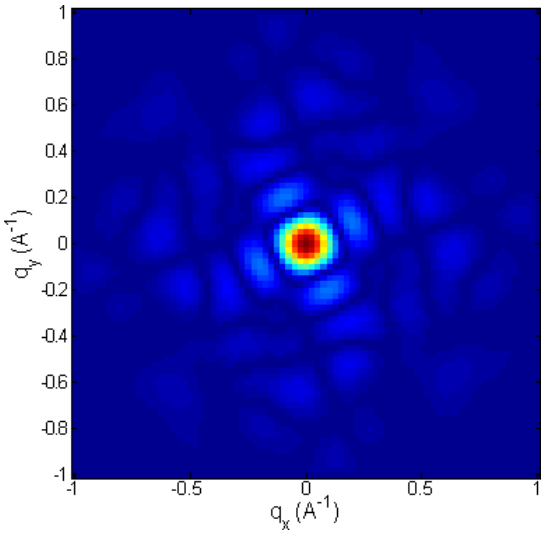
$$B_M(q,q') = \int C_2(q,q',\Delta\varphi)\exp(iM\Delta\varphi)d\Delta\varphi = I_M(q)^* I_M(q')$$

LHS is known from measurement
If $I_M(q)$ may be found, $I_M(q')$ follows from above equation

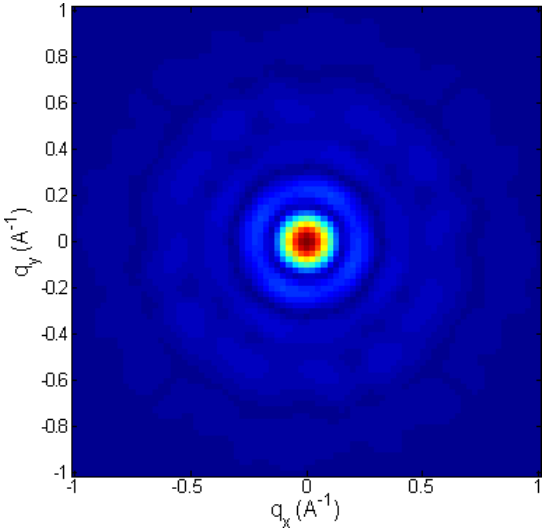
Intensities to Electron Density
Phasing Algorithm (Oszlànyi and Süto, 2004)



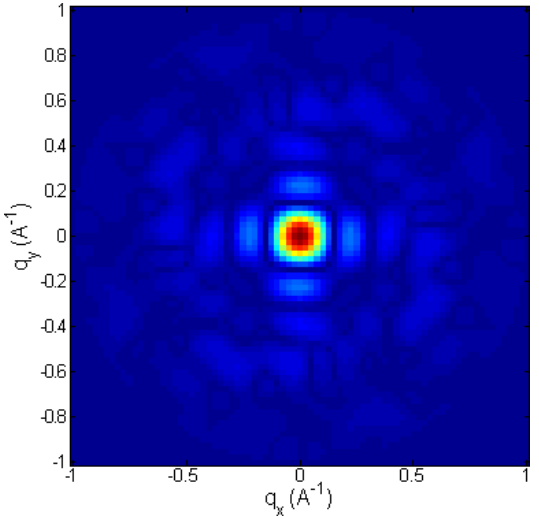
From Diffraction Patterns to Projected Electron Density



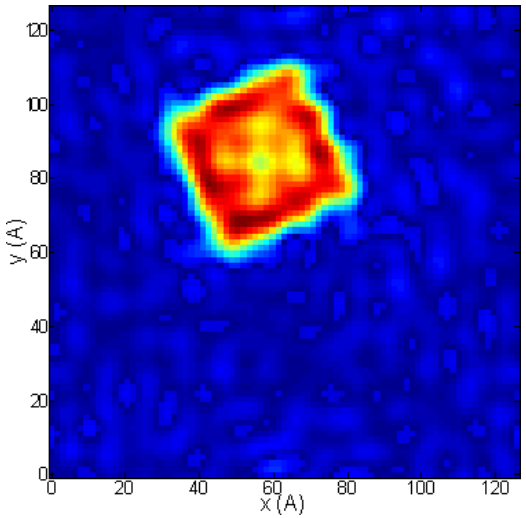
Single Particle DP



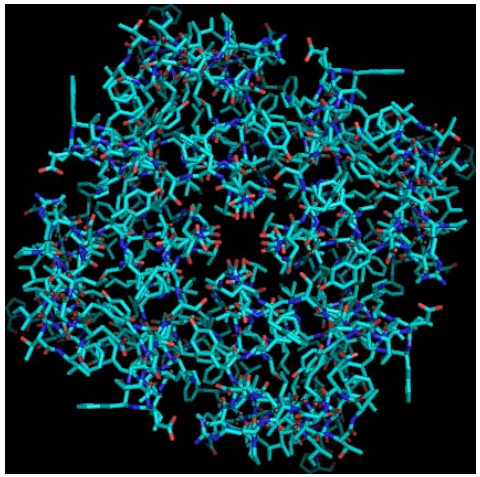
Scrambled Egg



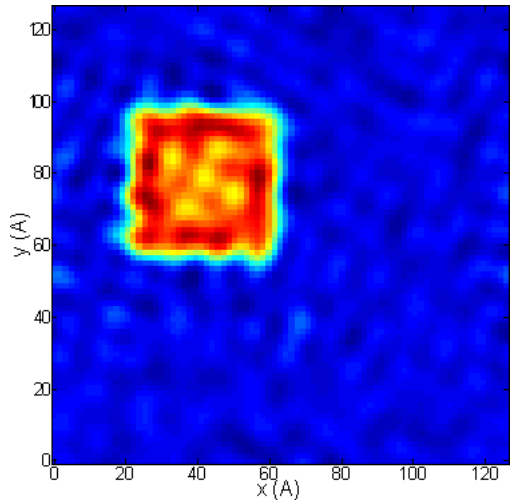
Unscrambled Egg



Proj. Electron Density



PDB structure

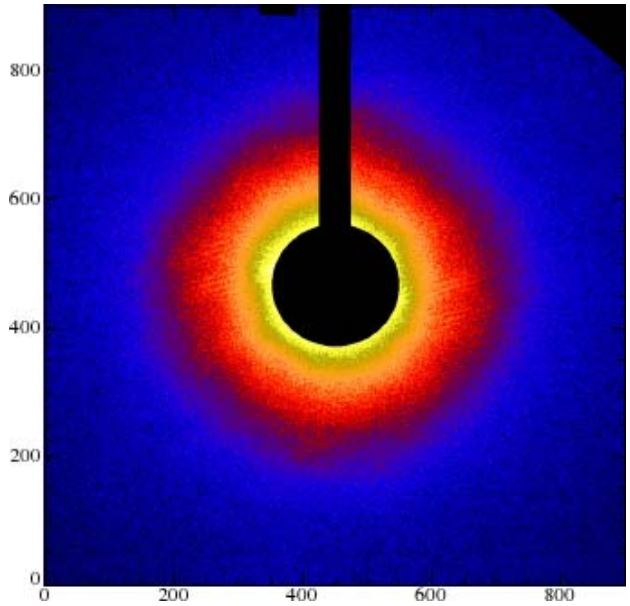


Proj. Electron Density

EM Image of Sample for Experiment

Mainly ~ 80 nm x 20 nm randomly oriented metal rods on a SiN substrate

Information In Angular Correlations

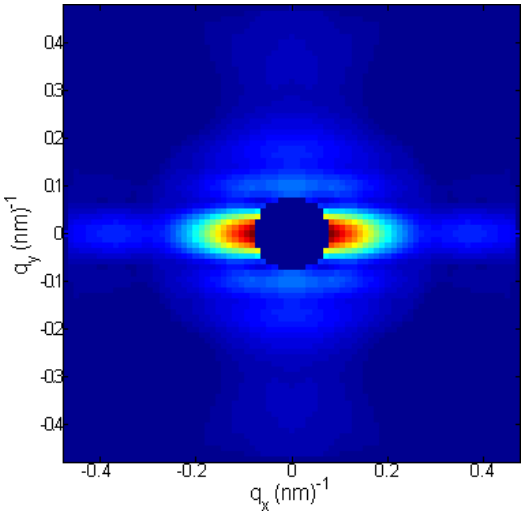


Diffraction pattern from disordered subunits
Appears to have no angular structure,
only radial variation, studied by SAXS.
However there is untapped information in the angular correlations, revealed by evaluating

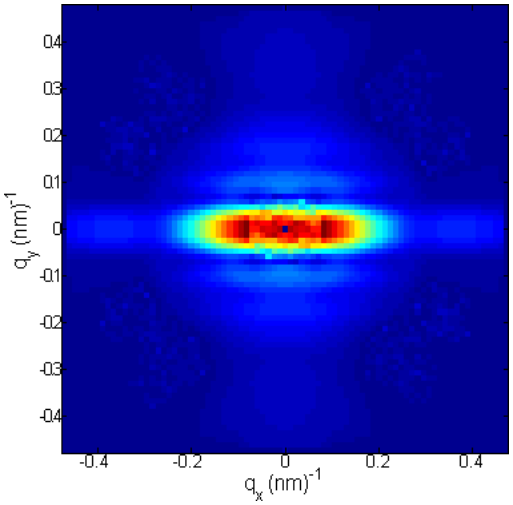
$$C_2(q;q',\Delta\varphi) = \left\langle \frac{1}{N_\varphi} \sum_j \left\{ I(q,\varphi_j) - I_{saxs}(q) \right\} \left\{ I(q',\varphi_j + \Delta\varphi) - I_{saxs}(q') \right\} \right\rangle_{DP}$$

DP and Image Reconstruction from Simulated Correlations

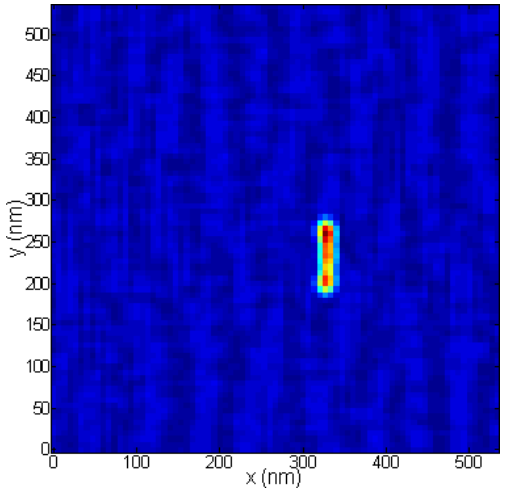
Identical Rods



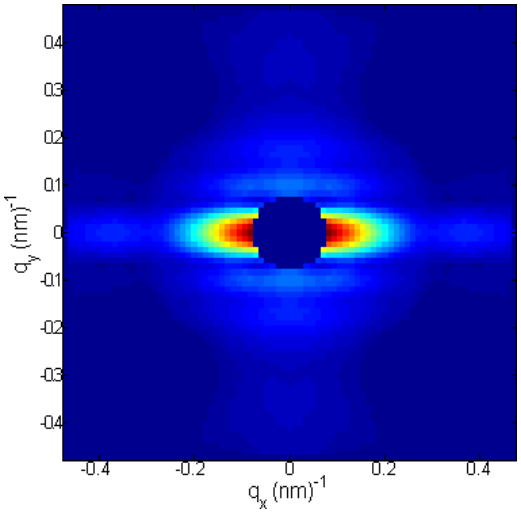
With beam stop

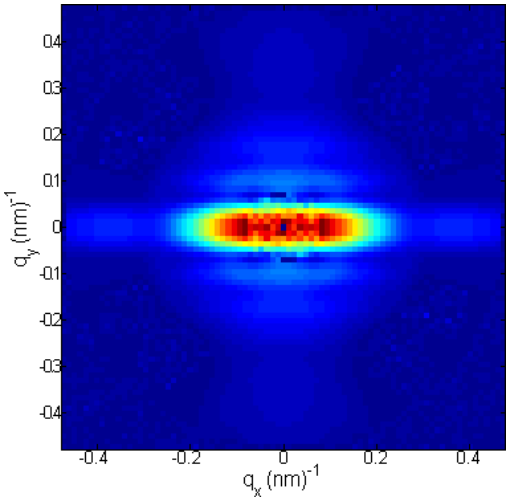


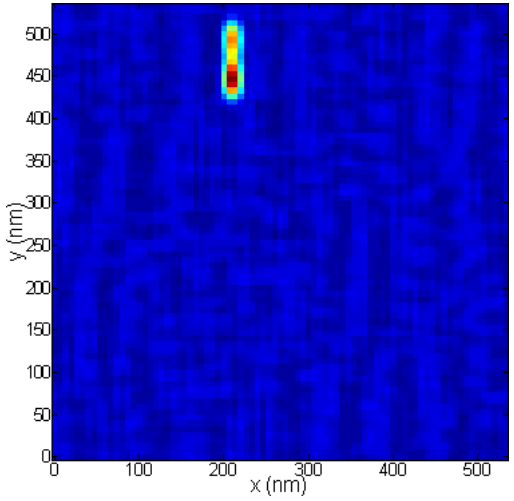
Beam stop filled by phasing alg.



Reconstructed rod image



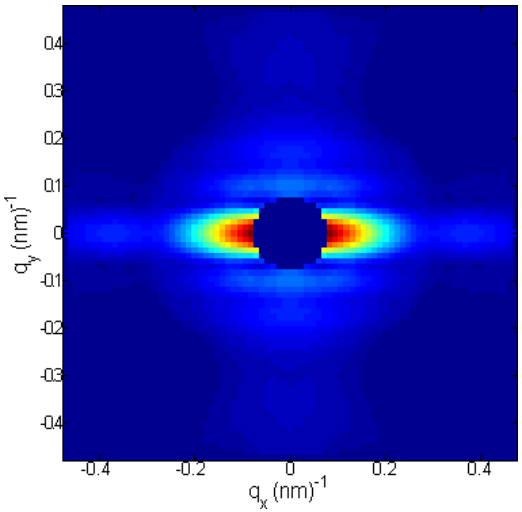




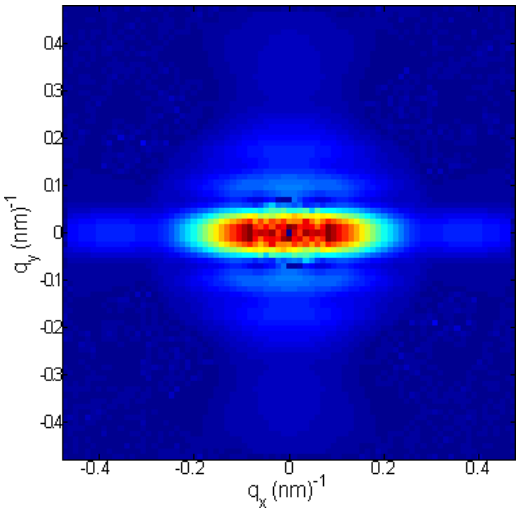
Objects w/ 10% range of sizes

DP and Image Reconstruction
Simulated & Measured Correlations

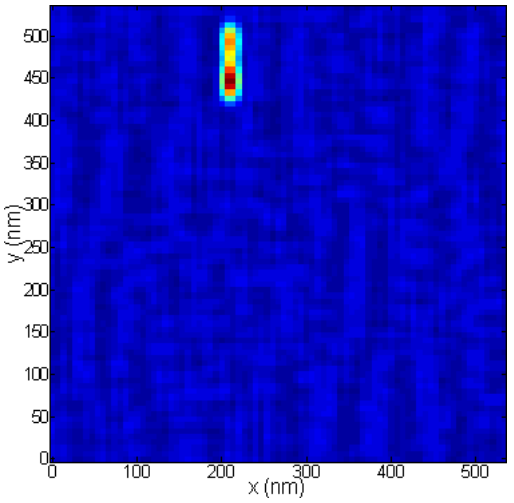
Simulation



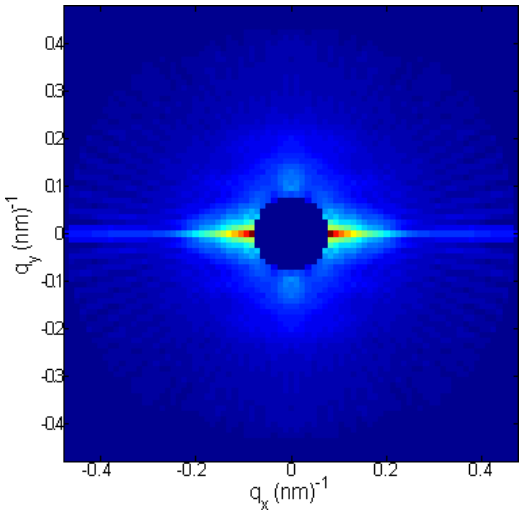
With beam stop

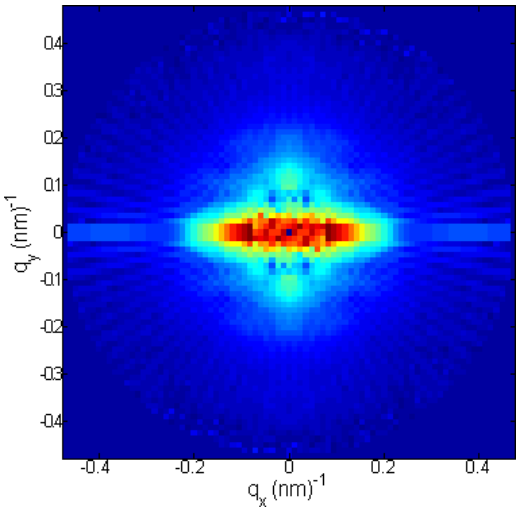


Beam stop filled by phasing alg.

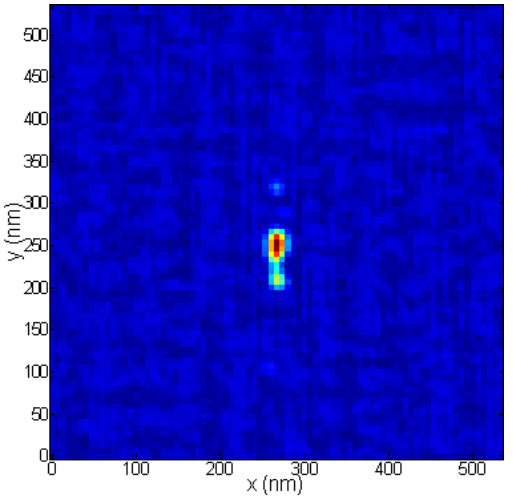


Reconstructed rod image

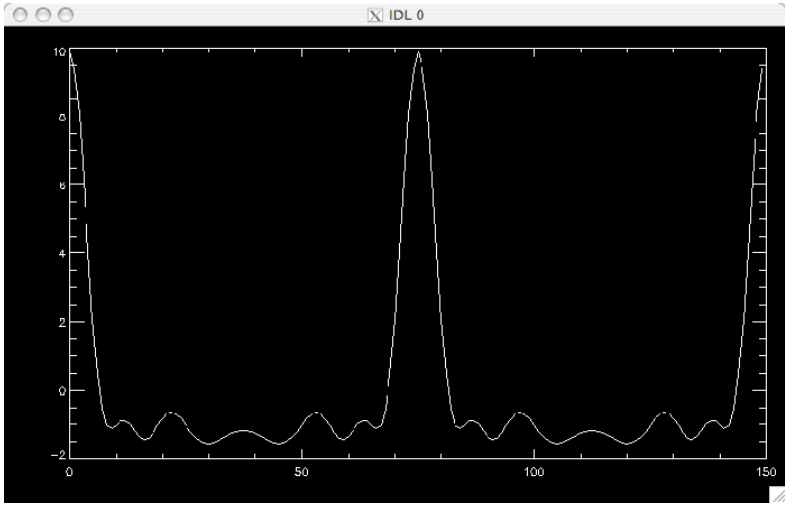




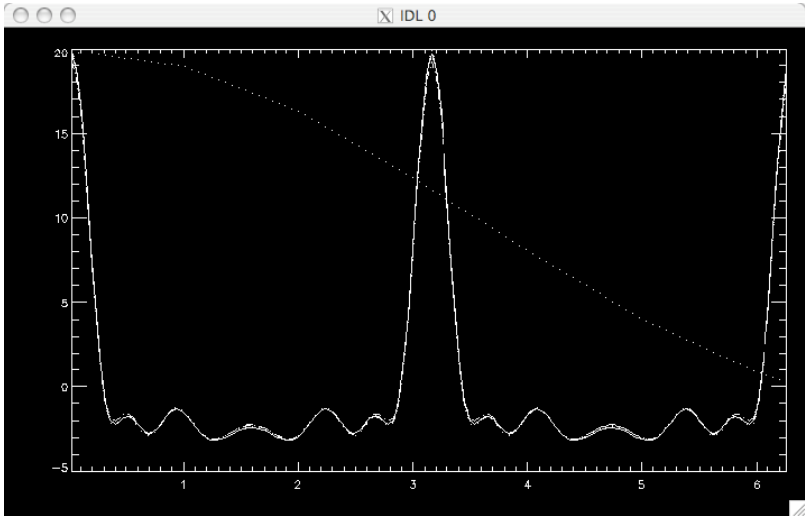
Experiment



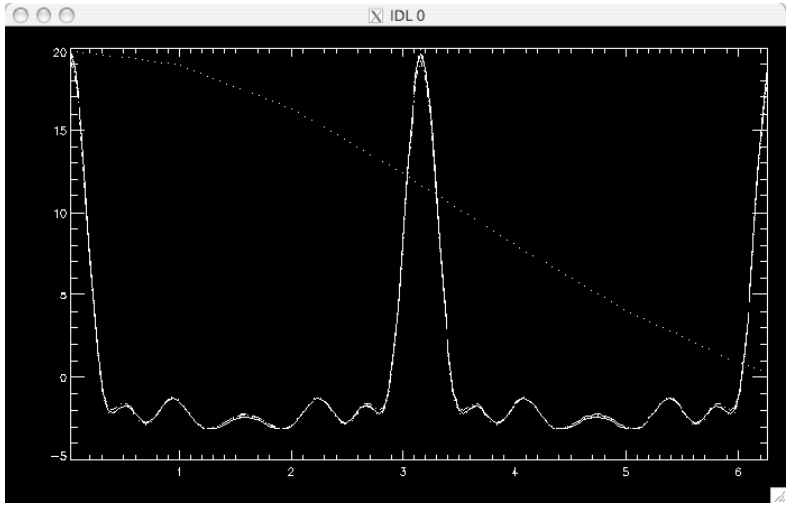
Simulated Pair Correlations



1 particle

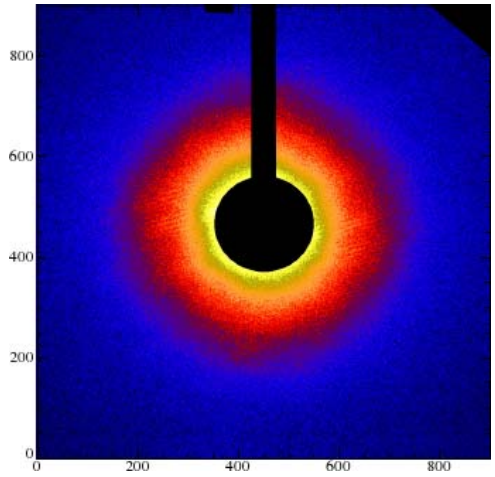


2 particles, 1000 DPs

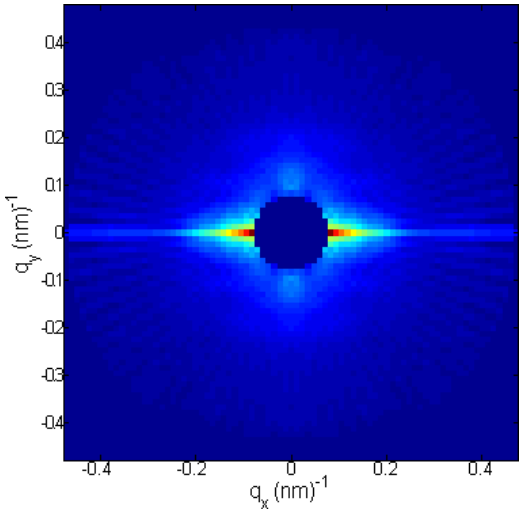


10 particles, 1000 DPs

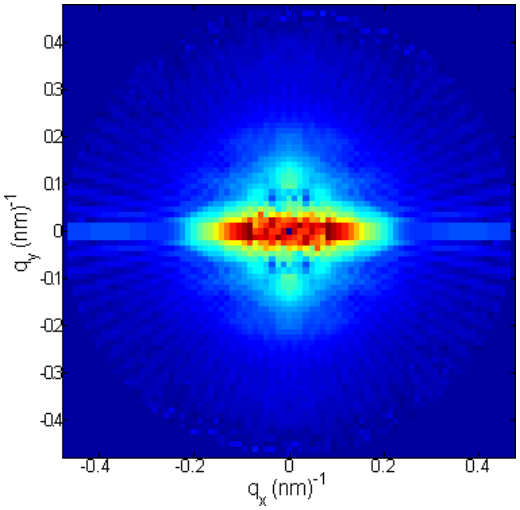
From Measured Multiparticle DP to Single-Particle Image



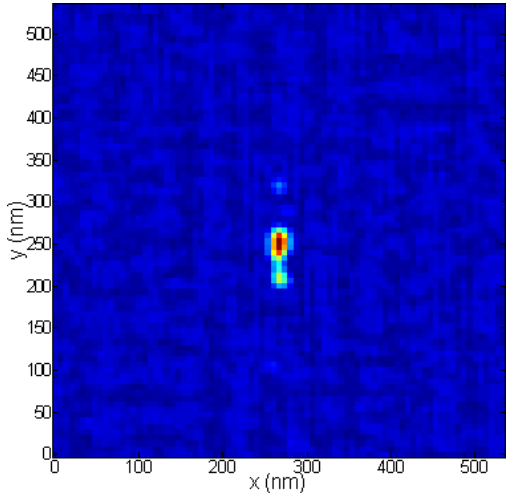
Measured DP



DP reconstructed from correls.



Beam stop filled by phasing alg.



Reconstructed image

Conclusions

- Single-axis alignment introduces orientational order intermediate between completely random orientations and perfect 3D orientation as in a crystal
- Distribute dose amongst different particles – stay below radiation damage threshold (to avoid uncertainties with respect to electron rearrangements in a “diffract and destroy” experiment
- Much information may be deduced about a molecule from scattering by such an ensemble by incidence (1) perpendicular to, and (2) parallel to, the alignment axis
- Case (1): can deduce coefficients of a cylindrical harmonic expansion of the amplitudes and molecular electron density
- Case (2): angular correlations of intensities allows recovery of a full single-particle DP
- In either case an iterative phasing algorithm can determine much information about the electron density
- Possible application: structure determination of membrane proteins in situ – experiments on a black lipid membrane?