Structure of a single particle from scattering by many particles randomly oriented about an axis: a new route to structure determination


1Department of Physics, University of Wisconsin-Milwaukee, Milwaukee, WI 53211
2Lawrence Berkeley National Laboratory, Berkeley, CA 94720
3Center for Free Electron Laser Science, DESY, and Universität Hamburg, 22607 Hamburg, Germany
4Department of Physics, Arizona State University, Tempe, AZ 85287

We demonstrate a complete algorithm for determining the electron density of an individual particle from scattering by many particles, randomly oriented about a single axis, with possible applications to uncrystallizable membrane proteins in situ.

The reconstruction of a high-resolution image of a single particle from scattering by several symmetrically equivalent ones has been demonstrated recently [1]. We present here a solution to the more general case of scattering by many particles, randomly oriented about an axis. The sustained x-ray or electron fluence needed for image reconstruction would destroy all high resolution detail if focused on a single molecule, since resolution varies inversely as the third or fourth power of fluence [2]. By sharing this fluence over many identical molecules, crystallography both amplifies the scattered intensities, and allows reconstruction of the periodically averaged undamaged charge density of a molecule. Part of the price for this crystallographic redundancy is that the scattered signal is sampled only at the Shannon “frequency” of the complex amplitudes, resulting in loss of phase information. If the scattered intensities from a single molecule could be measured at their (finer) Shannon angular sampling rate, the structure of the scatterer could be determined by iterative phasing algorithms [3]. These advantages of signal amplification, damage reduction for high resolution, and access to oversampled intensities (allowing solution of the phase problem) may be combined if the structure of a single particle may be determined from diffraction patterns from many identical particles in random orientations. This is possible if scattering is recorded either from stationary particles or for a recording time less than the rotational diffusion time.

Consider x-ray scattering from identical molecules in solution. If the scattering is from radiation of pulse length longer than the rotational diffusion time, \( \tau \), of the molecules, the signal from each molecule will be its rotational average and signal amplification is provided by the addition of such rotational averages from all illuminated molecules, as exploited by the technique of small-angle x-ray scattering (SAXS). Although it was initially thought that only angularly averaged quantities such as the average size, and radius of gyration may be extracted from an angularly averaged \( I(q) \) curve, it has been demonstrated [4, 5] that, with some reasonable additional assumptions, even anisotropic details of the molecular shape may be found from measured SAXS data. Nevertheless, variations in a SAXS signal are greatest in the region of very small magnitude scattering wavevector, \( q \), where the signal is sensitive largely to the overall shape of the molecule, rather than its internal structure [6].

The increasing availability of intense short-pulse x-ray sources allows a measurement of a type not possible before, namely fluctuations of intensity about mean SAXS values from instantaneous snapshots of a relatively small number of identical particles, differing only in random positions and random orientations. This becomes possible with intense radiation pulses shorter than \( \tau \). The resulting instantaneous diffraction patterns will contain intensities which are not quite constant on rings of constant \( q \). The angular correlation functions of these intensity fluctuations contain information about the diffraction volume of the individual scatterers [7, 8]. More intense x-ray sources also allow a useful signal to be measured from higher values of \( q \), thus permitting access to higher resolution structure. Completely new possibilities for structure determination of individual non-crystallized molecules become possible if such single-particle diffraction volumes may be deduced reliably.

Although this idea may be applied to e.g. completely randomly oriented molecules in a liquid [7], we examine here a case, where the molecular orientations are assumed random only about a single axis. This may simulate e.g. a membrane bound protein in situ with the membrane perpendicular to the incident beam. We illustrate the principle with a simplified model of a membrane protein as a right cylinder of uniform refractive index \( a \), radius \( d \), and director, i.e. axis, \( d \). Under the Born approximation, the differential scattering cross section is

\[
\frac{d\sigma}{d\Omega} = \frac{\omega^4 a^2}{c^4} \frac{|n^2 - 1|^2}{(q \cdot d)^2} (1 - \sin^2 \theta \cos^2 \phi) \times \\
\sin^2 \left( \frac{q \cdot d}{2} \right) J_1^2 \left( a \sqrt{q^2 - (q \cdot d)^2} \right)
\]

The beam was assumed incident along the \( \hat{z} \) axis, with \( q = \frac{\omega}{c} \left[ \sin \theta \cos \phi \hat{x} + \sin \theta \sin \phi \hat{y} + (\cos \theta - 1) \hat{z} \right] \), where \( \theta \) and \( \phi \) are the polar and azimuthal angles of scattering, and incident polarization along the \( \hat{x} \) axis. The polarization term \( 1 - \sin^2 \theta \cos^2 \phi \) was summed over the two
We investigate whether the single-particle diffraction pattern of Fig. 1 can be reconstructed from several short-pulse multiparticle diffraction patterns of the form of Fig. 2. We represent the single-particle diffraction pattern by the circular-harmonic expansion

\[ I(q_x, q_y) = \sum_m I_m(q)e^{im\phi}. \]  

(2)

where \( I(q_x, q_y) \) represents the intensity of a point on the diffraction pattern of Cartesian reciprocal-space coordinates \((q_x,q_y)\) or equivalent polar coordinates \((q,\phi)\). In general, the coefficients \( I_m(q) \) are complex valued. The reality of \( I(q) \) dictates that only even values of \( m \) are nonzero in the sum on the RHS of (2).

Angular pair-correlation functions spanning pairs of resolution rings specified by \( q \) and \( q' \) are calculated in two steps. First, the data \( I(q, \phi) \) of each resolution ring \( q \) is mean subtracted to find

\[ I'(q, \phi) = I(q, \phi) - \langle I(q, \phi) \rangle_\phi. \]  

(3)

For small values of \( q \), \( \langle I(q, \phi) \rangle_\phi \) may be identified with the usual small-angle x-ray scattering (SAXS) signal \( I_{SAXS}(q) \). We then define an angular pair correlation taken between the data on two resolution rings of radii \( q \) and \( q' \) of a diffraction pattern by [7–9]

\[ C_2(q, q', \Delta\phi) = \frac{1}{N_\phi} \sum_\phi I'(q, \phi)I'(q', \phi + \Delta\phi). \]  

(4)

Likewise, one may define a triple correlation [10] over the resolution ring \( q \) by

\[ C_3(q, q, \Delta\phi) = \frac{1}{N_\phi} \sum_\phi [I'(q, \phi)]^2 I'(q, \phi + \Delta\phi). \]  

(5)

Their angular discrete Fourier transforms are defined by

\[ B_m(q, q') = \frac{1}{N_\phi} \sum_{\Delta\phi} C_2(q, q', \Delta\phi)e^{-im\Delta\phi} \]  

(6)

and

\[ T_m(q, q) = \frac{1}{N_\phi} \sum_{\Delta\phi} C_3(q, q, \Delta\phi)e^{-im\Delta\phi}. \]  

(7)

Correlation functions of the form \( C_2 \) and \( C_3 \) would be expected to be identical regardless of the orientation of a particle about an axis parallel to the incident radiation [11]. If multiple particles, randomly oriented about this axis, scatter incident radiation under conditions where interparticle interference is negligible, the resulting diffraction pattern would be a superposition of
those of the individual particles, and thus $C_2$ and $C_3$ calculated from this composite diffraction pattern and averaged over many such diffraction patterns (we included data from 100 such simulated patterns) would be expected to converge on the corresponding quantities for a single particle. We may write

$$I'(q, \phi) = \sum_j \sum_{m \neq 0} I_m(q) e^{im(\phi - \eta_j)} \quad (8)$$

where the subtraction of $(I(q, \phi))_j$ in (3) is equivalent to omitting the term $m = 0$ in this summation, and where $\eta_j$ is the (random) orientation of particle $j$ about $\hat{z}$.

Substituting (8) into (6) and (7), it may be shown that

$$B_m(q, q') = N_p I_m(q) I^*_m(q') \quad \text{for } m \neq 0 \quad (9)$$

where $N_p$ is the number of particles giving rise to each diffraction pattern, and that

$$T_m(q, q) = N_p I^*_m(q) \sum_{M \neq 0, m} I_M(q) I_{M-M}(q) \quad \text{for } m \neq 0. \quad (10)$$

The quantities on the LHS’s of (9) and (10) may be found from the angular Fourier transforms of pair- and triple-correlation functions computable from measured data of diffraction patterns from multiple particles. Since this is a quantity from measurement, and the quantities on the RHS are from theory, and since we seek only relative magnitudes of the circular-harmonic expansion coefficients $I_m(q)$, we could simply remove the common scaling factor $N_p$ from Eqs. (9) and (10), and estimate

$$|I_m(q)| = \sqrt{B_m(q, q)} \quad \text{for } m \neq 0 \quad (11)$$

[$B_m(q, q)$ is real and non-negative from (9)]. A full determination of these expansion coefficients requires also an estimation of the phases of $I_m(q)$.

From (9) with the factor of $N_p$ omitted,

$$I_m(q) = B_m(q, q')/I^*_m(q'), \quad \text{for } m \neq 0. \quad (12)$$

Thus, the determination of the phase of $I_m$ for a single resolution ring (say $q'$) uniquely determines the phases of $I_m$ for the same value of $m$ for all other rings $q'$.

Indeed, substitution of (12) into (10) (with $N_p$ omitted) leads to

$$T_m(q, q) = \frac{B^*_m(q, q')}{|I_m(q')| e^{i\phi'(m, q')}} \times \sum_{M \neq 0, m} \frac{B_M(q, q')}{|I_M(q')| e^{-i\phi'(M, q')}} \frac{B_{M-M}(q, q')}{|I_{M-M}(q')| e^{-i\phi'(M-M, q')}} \quad (13)$$

for $m \neq 0$. The only unknowns in (13) are the (real) phases $\phi'(M, q')$ of the circular-harmonic expansion coefficients, $I_M(q')$, for $M \neq 0$.

Once again, the quantity on the LHS of (13) may be found from experiment. In general there are $2^\mu$ unknown phases $\phi(M, q')$ for a single resolution ring $q'$, where $\mu$ is the number of values of $M$ which give rise to distinct values of $I_M(q')$. In our simulation, we took circular-harmonic expansion coefficients $M$ up to a maximum value 38, so given the fact that $I_{-M}(q) = I_M(q)$ and that $M$ takes up only even values, $\mu=19$. In general, these $\mu$ real numbers may be found by optimizing the function

$$\sum_{m \neq 0} |T^{(obs)}_m - T^{(calc)}_m|^2$$

by means of a global optimization algorithm such as simulated annealing [12] (where $T^{(obs)}_m$ are the quantities found from the angular Fourier transforms of the triple correlations $C_3$ from the measured diffraction patterns, and $T^{(calc)}_m$ their values estimated from the RHS of (13)). In the present simulation, for an individual object which gives rise to a diffraction pattern with two mutually perpendicular mirror planes, the coefficients $I_M(q)$ may be taken to be real, and the only uncertainties are their signs. In this case, even an exhaustive search through $2^{19}$ sign combinations to optimize (14) took no more than a few seconds on a fast desktop computer. The value of $I_0(q)$ is estimated from the same data via $\langle I(q, \phi) \rangle_\phi$, the angular average of the intensities of resolution rings $q$. [Since, over most of the range of $q$ of the diffraction patterns here, this scales with $N_p$ the same way as $B_m(q)$ and $T_m(q)$, its relative magnitude to the coefficients for $m \neq 0$ may be determined correctly from the same dataset. Also, since $I_0(q)$ is real and positive, there is no need to determine its phase (or sign).]

The result of reconstructing the single-particle diffraction pattern up to the resolution ring marked in Fig. 1
with the quantities $I_m(q)$ from the correlations $C_2$ and $C_3$ and SAXS data of 100 multi-particle diffraction patterns like Fig. 2 is shown in Fig. 3. A remarkable similarity to the model single-particle diffraction pattern of Fig. 1 is apparent. Simulated Poisson noise produces an unimportant peak at the origin of the correlation functions $C_2$ and $C_3$. For the parameters of our simulations, interference fringes from interparticle interference are finer than a detector pixel, and are thus irrelevant.

Finally, we recover (Fig. 4) from the diffraction pattern of Fig. 3 the projected electron density of the individual particle, after $\sim 15$ iterations of the charge-flipping algorithm \[13\]. An unambiguous reconstruction of the 80 nm $\times$ 20 nm rod to the resolution of the reconstructed diffraction pattern ($\sim \pi/0.45 \approx 7$ nm) is found, without any assumption about compact support of the particle.

These results demonstrate that it is not necessary to measure the diffraction pattern of a single particle in order to reconstruct its image via an iterative phasing algorithm. Even if multiple identical particles in random orientations are illuminated the resulting composite diffraction pattern may be decomposed into an oversampled diffraction pattern of an individual particle of sufficient quality to allow an iterative phasing algorithm to reconstruct an accurate real-space image of the particle. A similar method \[10\] may be applied to particles entirely randomly oriented in 3D, allowing the structure determination of a difficult-to-crystallize molecule from short-pulse diffraction patterns of such molecules in solution, provided scattering by solution atoms may be minimized or isolated.

The results are relevant also to the kind of “differ and destroy” approach to biomolecular structure determination \[14\] proposed for the x-ray free electron laser (XFEL), where measurements of diffraction from single molecules have been proposed. In this case, the SAXS background, which arises from uncorrelated scattering by different particles, is entirely absent, as is scattering from solution atoms \[8\]. Alternatively, the ability to extract meaningful information from multi-particle diffraction patterns may help overcome intensity issues from single particle scattering.

The method is not restricted to diffraction patterns from scattering of externally incident radiation. A diffraction pattern may be measured also from x-rays arising from the fluorescence of a specific atom buried deeply within a large molecule \[15\] (e.g. the Fe atom in haemoglobin). Such a pattern would be most sensitive to intramolecular scattering, less to that from its surroundings. This may allow the reconstruction of the 3D diffraction volume of an individual molecule from that of many randomly oriented ones, and hence of the molecular structure, without too much interference from solvent (or membrane) scattering.

Advances in technology, such as fast column read-out area detectors, brighter sources, shorter pulses, and zone-plate focusing, have greatly improved the experimental conditions for such experiments. When coupled with new theoretical advances, such as those reported here, entirely new vistas are opened up for the structure determination of microscopic particles.

DKS acknowledges support from DOE Grants DE-FG02-84ER45076, DE-FG02-06-ER46277 and the UWM RGI, JCHS from DE-FG03-02ER45996 and the CBST.