

# Reconstruction of the electron density of molecules with single-axis alignment

Dmitri Starodub\*<sup>a</sup>, John C. H. Spence<sup>b</sup>, Dilano K. Saldin<sup>c</sup>

<sup>a</sup>Stanford PULSE Institute, SLAC National Accelerator Laboratory, 2575 Sand Hill Rd, Menlo Park, CA, USA 94085

<sup>b</sup>Dept. of Physics, Arizona State University, PO Box 871504, Tempe, AZ 85287

<sup>c</sup>Dept. of Physics, University of Wisconsin-Milwaukee, PO Box 413, Milwaukee, WI 53201

## ABSTRACT

Diffraction from the individual molecules of a molecular beam, aligned parallel to a single axis by a strong electric field or other means, has been proposed as a means of structure determination of individual molecules. As in fiber diffraction, all the information extractable is contained in a diffraction pattern from incidence of the diffracting beam normal to the molecular alignment axis. We present two methods of structure solution for this case. One is based on the iterative projection algorithms for phase retrieval applied to the coefficients of the cylindrical harmonic expansion of the molecular electron density. Another is the holographic approach utilizing presence of the strongly scattering reference atom for a specific molecule.

**Keywords:** X-ray coherent imaging, fiber diffraction, phase retrieval

## 1. INTRODUCTION

Due to very small X-ray atomic scattering cross-sections, usual diffraction methods for structure determination require measurement of scattered intensity from many identical copies of macromolecules. In conventional crystallography this is achieved by using crystals, where identically oriented macromolecules are located in regularly spaced positions. This gives an enhancement of the scattered signal by a factor of the squared ratio of the crystal volume to that of a unit cell. However, translational symmetry limits the scattered intensity sampling to reciprocal lattice of the crystal, which does not provide enough information for solution of the phase problem from the structure factors alone without sample modification and/or using anomalous scattering. Furthermore, large protein complexes and membrane proteins, vital for the function of a living cell, are difficult to crystallize, and therefore constitute a major challenge for structural biology. In the opposite limit of small angle scattering, the incoherent sum of solution scattering from macromolecules in random orientations is recorded, resulting in the spherically averaged diffraction pattern. Then a low resolution structure can be recovered by the Monte Carlo search of a model that provides a best fit to experimental data. In this type of simulation the problem of uniqueness of the recovered structure is intrinsically present. It was proposed to overcome difficulties of crystallography and of solution scattering by replacing intermolecular forces, orienting macromolecules in a crystal, by interaction of individual isolated molecules with an external field.<sup>1</sup> In this method, scattered intensity is collected from a stream of aligned molecules, and a weak diffraction pattern of the individual molecule is multiplied by the total number of molecules interacting with the incident X-rays. A limited spatial coherence width of the X-ray beam is assumed to prevent interference between X-rays scattered by different molecules. Absence of translational periodicity allows arbitrary sampling of the diffraction pattern and subsequent use of well-developed iterative projection algorithms<sup>2</sup> to solve the phase problem. Three-dimensional alignment of small molecules can be achieved by interaction of their polarizability anisotropy with the electric field of an intense elliptically polarized laser pulse,<sup>3</sup> while interaction of the induced dipole moment with the electric field of a linearly polarized laser aligns the molecule's axis of easy polarization along the electric field,<sup>3</sup> if the induced dipole – electric field interaction energy is large compared to the thermal energy of the molecule. In this field configuration, rotation of the molecule with respect to the alignment axis is not restricted. In many cases realization of the single axis alignment is easier to achieve, and for some systems, such as symmetric-top

\*starodub@slac.stanford.edu

molecules, the rotational dynamics about their symmetry axis cannot be constrained with the elliptically polarized light. The alternatives to the laser alignment include strong electric or magnetic fields and a shear flow. It has been demonstrated that the moderate electric field of 2 kV/cm is sufficient for the perfect alignment of single wall nanotube (SWNT) ethanol suspensions.<sup>4</sup> The dipole moment of a peptide bond is  $\sim 3.6$  D, and their alignment in  $\alpha$ -helices results in buildup of the macroscopic dipole moment roughly in proportion to the number of residues.<sup>5</sup> Thus, large proteins can have dipole moments of  $10^3$ - $10^4$  D, which is sufficient for alignment in a strong electric field.

Apparently, information is inevitably lost upon rotational averaging of scattering data from the molecules aligned in one direction. The aim of this work is to demonstrate that under some circumstances, in particular for high degree or known molecular cylindrical symmetry, the structural information on an individual molecule can be deduced from a single diffraction pattern produced by a large number of the single-axis aligned molecules randomly rotated about the alignment axis. Such a diffraction pattern is similar in character to a fiber diffraction pattern. However, fibrous molecules studied by fiber diffraction display a helical symmetry, and therefore they are characterized by some repeat distance along their long axis. That restricts scattering to the layer planes perpendicular to the direction of the molecular alignment. Unlike fibrous molecules, molecules aligned by an external field do not have periodicity in any direction, and corresponding diffraction patterns are continuous. This allows an oversampling of the scattered intensity with respect to the Shannon sampling interval as determined by the molecular size. Then a two-dimensional iterative projection algorithm, which alternatively applies constraints in real and reciprocal space, can be applied to reconstruct at least azimuthally averaged electron density of the sample.<sup>6</sup> In the special cases, when a molecule contains a strongly scattering reference atom, the diffraction pattern can be treated as a hologram, and the azimuthal projection of the molecular electron density about the alignment axis may be directly obtained by the corresponding transform of the diffraction pattern.<sup>7</sup>

## 2. STRUCTURE DETERMINATION FROM CYLINDRICALLY AVERAGED DIFFRACTION PATTERN

### 2.1 Fourier transform in cylindrical coordinates for non-periodic object

In this section we establish relationship between the electron density of an individual particle and the diffraction pattern produced by a large number of aligned identical particles, different only in angle of rotation about alignment axis. The derivation closely follows that for fiber diffraction,<sup>8</sup> with an important difference arising from the fact that we do not have periodicity in the alignment direction. That results in a continuous diffraction pattern as opposed to layer lines observed in fiber diffraction, and allows direct application of iterative projection algorithms for phase retrieval. Since there is no interference between molecules, scattered intensity from a large ensemble of aligned molecules at random azimuthal angles is equivalent to the diffraction pattern of an individual molecule averaged over all possible rotations. Due to cylindrical averaging it is convenient to introduce cylindrical coordinates  $(r, \varphi, z)$  in real space, and corresponding coordinates  $(R, \psi, \zeta)$  in reciprocal space, where  $z$  and  $\zeta$  are parallel to the alignment axis. Any single valued and continuous function  $f(r, \varphi, z)$  can be expanded in a Fourier series in terms of the orthogonal set of basis functions  $\exp(im\varphi)$  defined on the unit circle in the plane perpendicular to rotation axis. Then electron density of the molecule can be written as the cylindrical harmonic expansion

$$f(r, \varphi, z) = \sum_m g_m(r, z) \exp(im\varphi), \quad (1)$$

where the two-dimensional function  $g_m(r, z)$  is determined as

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

In the same fashion, the scattering amplitude of the molecule can be represented in reciprocal space as

$$A(R, \psi, \zeta) = \sum_m G_m(R, \zeta) \exp(im\psi). \quad (3)$$

The scattered intensity is

$$I(R, \psi, \zeta) = \langle A(R, \psi, \zeta) A^*(R, \psi, \zeta) \rangle_{\psi} = \sum_m \sum_{m'} G_m(R, \zeta) G_{m'}^*(R, \zeta) \langle \exp[i(m - m')\psi] \rangle_{\psi}, \quad (4)$$

where angle brackets stand for angular average over azimuthal angle. The cross terms in the angle brackets  $\langle \dots \rangle$  in Eq. (4) vanish, leaving only diagonal terms:

$$I(R, \zeta) = \sum_{m=-\infty}^{\infty} |G_m(R, \zeta)|^2, \quad (5)$$

where there is no dependence on the azimuthal angle  $\psi$ . Thus, in the general case the scattered intensity from the multiple molecules aligned parallel to a single axis is a sum of independent contributions from different cylindrical harmonics. Eq. (5) also elucidates how the information content of the full three-dimensional diffraction pattern is reduced by cylindrical averaging. If the cylindrical harmonic expansion is terminated at the term  $m_{max}$ , then the right hand side of Eq. (5) can be regarded as the squared norm of the  $4m_{max}$  vector, with each coefficient  $G_m(R, \zeta)$  contributing two dimensions corresponding to its real and complex parts. If scattered intensity from a fixed molecule is measured in a given point of reciprocal space, then modulus constraint for this point is determined by a two-dimensional sphere of radius equal to square root of measured intensity. For the cylindrically averaged scattered intensity at a given pixel, modulus constraint is set by a sphere in  $4m_{max}$  dimensions.

Since the X-ray scattering cross section is small, scattering can be described in the Born approximation, and in the far field the scattering amplitude is given by the Fourier transform of the molecular electron density:

$$A(\mathbf{q}) = \frac{1}{\sqrt{(2\pi)^3}} \int f(\mathbf{r}) \exp(i\mathbf{q}\mathbf{r}) d^3\mathbf{r}. \quad (6)$$

Substituting here scattering vector  $\mathbf{q}$  and radius-vector  $\mathbf{r}$  expressed in cylindrical coordinates, and using Jacobi-Anger expansion<sup>9</sup>

$$\exp(iz \cos \theta) = \sum_{m=-\infty}^{\infty} i^m J_m(z) \exp(im\theta), \quad (7)$$

where  $J_m(z)$  is the  $m^{\text{th}}$  order Bessel function of the first kind, and using Eq. (2) we arrive to the expression

$$A(R, \psi, \zeta) = \frac{1}{\sqrt{2\pi}} \sum_m i^m \int r g_m(r, z) J_m(rR) \exp[i(z\zeta + m\psi)] dr dz. \quad (8)$$

From comparison of Eq. (8) and Eq. (3) the relationship between the coefficients of the cylindrical harmonic expansions of the molecular electron density and its scattering amplitude can be identified as

$$G_m(R, \zeta) = \frac{i^m}{\sqrt{2\pi}} \int r g_m(r, z) J_m(rR) \exp(iz\zeta) dr dz. \quad (9)$$

The relation inverse to Eq. (9) can be derived in the same way from the inverse Fourier transform of the molecular scattering amplitude as follows:

$$g_m(r, z) = \frac{(-i)^m}{\sqrt{2\pi}} \int R G_m(R, \zeta) J_m(rR) \exp(-iz\zeta) dr dz. \quad (10)$$

Eqs. (9) and (10) represent a combination of a Hankel transform in radial direction and a Fourier transform in the direction of alignment, which we term a Fourier-Hankel transform. These coupled two-dimensional equations relate the cylindrical harmonic expansion coefficients of the scattering amplitude with those of the electron density of the molecule. These relationships are analogous to the Fourier transform and its inverse, which relate scattering amplitude and corresponding electron density in a standard scattering problem. In a case where the magnitudes of  $G_m(R, \zeta)$  are known, but not their phases, and the spatial extent of  $g_m(r, z)$  is finite, it would be expected that a two-dimensional

iterative phasing algorithm with appropriate constraints in reciprocal and real space will allow the simultaneous determination of the unknown phases and of  $g_m(r, z)$ .

## 2.2 Analogy with fiber diffraction

Due to periodicity of the helical structure in fiber diffraction with repeat distance  $c$  along axis  $z$ , function  $g_m(r, z)$  can be expanded in Fourier series with respect to argument  $z$ , and Eq. (1) becomes

$$f(r, \varphi, z) = \sum_l \sum_m g_{ml}(r) \exp[i(m\varphi - 2\pi lz/c)]. \quad (11)$$

Component  $g_{ml}(r)$  and its counterpart in reciprocal space  $G_{ml}(R)$  are related by the Fourier-Bessel transform

$$g_{ml}(r) = \int R G_{ml}(R) J_m(rR) dr. \quad (12)$$

The quantity measured in fiber diffraction is scattered intensity of the layer line

$$I_l(R) = I(R, \zeta = 2\pi l/c) = \sum_m |G_{ml}(R)|^2. \quad (13)$$

Eqs. (11)-(13) of fiber diffraction<sup>8</sup> correspond to Eqs. (1), (10), and (5), respectively, for continuously scattering aligned non-periodic objects. As soon as components  $G_{ml}(R)$  and their phases are determined, electron density of fibrous molecule can be found from Eqs. (11) and (12).

## 2.3 Algorithm for phase retrieval

As in fiber diffraction, electron density reconstruction in our case is equivalent to the problem of deconvolution the individual contributions in Eq. (5) and determination of their phases. The number of  $m$ -components which should be kept in Eq. (5) depends on the sample and required resolution. It has been noted<sup>10</sup> that a Bessel function of order  $m$  has negligible values if its argument  $x$  is smaller than  $m$ . Therefore, it would be expected that for a given  $x = rR$  contribution of all cylindrical harmonics with  $m > x$  to the diffraction pattern intensity vanishes. The maximum value of  $x$  is defined by the maximum radial coordinate of any atom of the molecule relative to the molecular rotation axis, which can be approximated by the molecule radius  $a$ , and required resolution  $d = 2\pi/R_{max}$ . Then all terms with  $m > 2\pi a/d$  in Eq. (5) can be neglected. If the molecule is characterized by  $M$ -fold rotational symmetry, then cylindrical harmonic expansion of its electron density contains only terms of modulo  $M$ , i.e.  $m = 0, M, 2M, \dots$  Provided  $M > 2\pi a/d$ , the term with  $m = 0$  is the overwhelmingly dominant contributor to scattered intensity, and  $|G_0(R, \zeta)|$  can be approximated by the square root of the measured diffraction pattern. The phases of  $G_0(R, \zeta)$  can be determined by iteratively solving Eqs. (9) and (10) for  $m = 0$  to find the solution that simultaneously belongs to the set of objects, whose transform (9) of azimuthal projection results in the measured diffraction pattern, and the set of objects, satisfying available constraints in the real space, such as support constraint, which sets the electron density outside the known object boundaries to zero. In a more general case, it might be necessary to find a few terms in Eq. (5) beyond  $m = 0$ . Note that for a real object only terms with  $m \geq 0$  need to be considered. Expansion of the real electron density in cylindrical harmonics (1) can be written as

$$f(r, \varphi, z) = g_0(r, z) + 2 \sum_{m \geq 1} \text{Re}[g_m(r, z) \exp(im\varphi)]. \quad (14)$$

Using identity  $G_{-m}(R, \zeta) = (-1)^m G_m^*(R, -\zeta)$  for expansion coefficients of the scattering amplitude for a real valued object, we obtain in place of Eq. (5)

$$I(R, \zeta) = |G_0(R, \zeta)|^2 + \sum_{m \geq 1} (|G_m(R, \zeta)|^2 + |G_m(R, -\zeta)|^2). \quad (15)$$

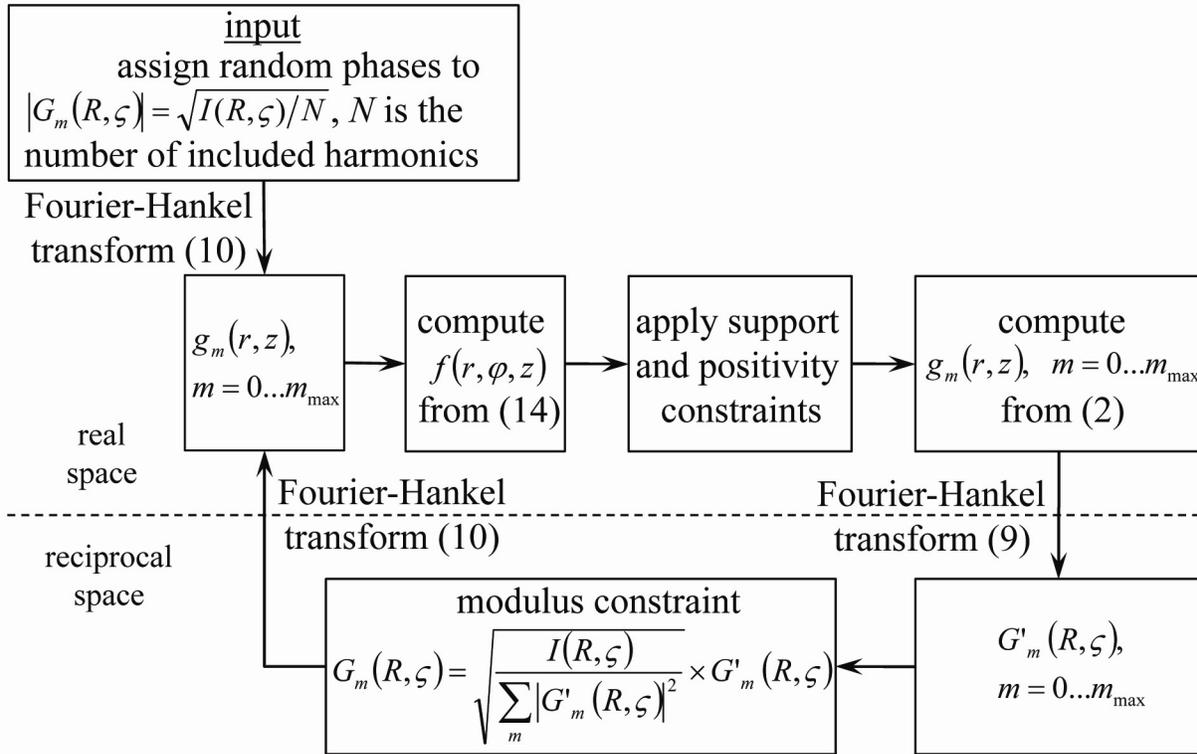


Figure 1. Block diagram of the phasing algorithm.

The expansion coefficients of scattering amplitude are independent, however they can be related by the positivity or other constraints in real space after their transform (10). The standard modulus projection which keeps the phases of current iterate  $G'_m(R, \zeta)$  of scattering amplitude and replaces magnitudes by the measured intensities in diffraction pattern, is modified to the form

$$G_m(R, \zeta) = G'_m(R, \zeta) \sqrt{\frac{I(R, \zeta)}{|G'_0(R, \zeta)|^2 + \sum_m (|G'_m(R, \zeta)|^2 + |G'_m(R, -\zeta)|^2)}}, \quad (16)$$

where summation is performed over all non-vanishing positive  $m$ . This is the same projection that was used in phasing powder diffraction pattern to separate overlaps of non-equivalent reflections having the same magnitudes of scattering vector,<sup>11</sup> originating from spherical averaging of powder diffraction patterns. The method has been generalized by Elser and Millane<sup>12</sup> for the reconstruction problem in the case of diffraction pattern incoherently averaged over a discrete symmetry group. It was shown that with positivity constraint available, reconstructions are successful for up to symmetry group order of 4. Applied to our problem, that implies that just one component corresponding to non-zero  $m$  can be included in Eq. (14). However, it can be expected that this condition will be relieved if additional constraints are available in real space. Block diagram of the phasing algorithm is presented in Fig. 1. As a starting point, the scattered intensity is equally distributed among  $N$  expansion coefficients included in the diffraction pattern decomposition, with each of those coefficients assigned a value of  $|G_m(R, \zeta)| = \sqrt{I(R, \zeta)/N}$ , and a set of random phases  $\chi(R, \zeta)$  is generated for them. After transform (10) for each  $m$ -component, they have been combined via Eq. (14) to obtain a first estimation of electron density. This estimate is a real function which generally contains both positive and negative values. A suitable object-domain operation utilizing projection on the set restricted by object support and positivity is applied, and subsequent inverse Fourier-Hankel transform (9) gives a new estimate of scattering amplitude expansion coefficients. Their phases and relative values for different  $m$ -components are retained, but amplitudes are renormalized according to projection (16). The algorithm iterates between two constrained sets until convergence occurs.

## 2.4 Holographic approach

If a molecule contains a strong scatterer, an alternative approach to reconstructing an azimuthal projection of the molecule can be applied. In this case diffraction pattern can be treated as a hologram, consisting of the square modulus of a superposition of a relatively large reference wave of known form and an unknown object wave. Then, the scattered intensities become linear in the unknown object wave and its conjugate, thus allowing the reconstruction of the object giving rise to object wave. This method was employed to reconstruct a simulated laser aligned symmetric top molecule  $\text{CF}_3\text{Br}$ .<sup>7</sup> The form factor of the Br atoms located on the rotation axis of the molecule dominates over those of F and C. It was demonstrated that application of the transform (10) with  $m = 0$  directly to the diffraction pattern gives the azimuthal projection of the  $\text{CF}_3\text{Br}$  molecule, along with a twin image, mirror symmetric relative to the plane perpendicular to the rotation axis, as shown in Fig. 1.

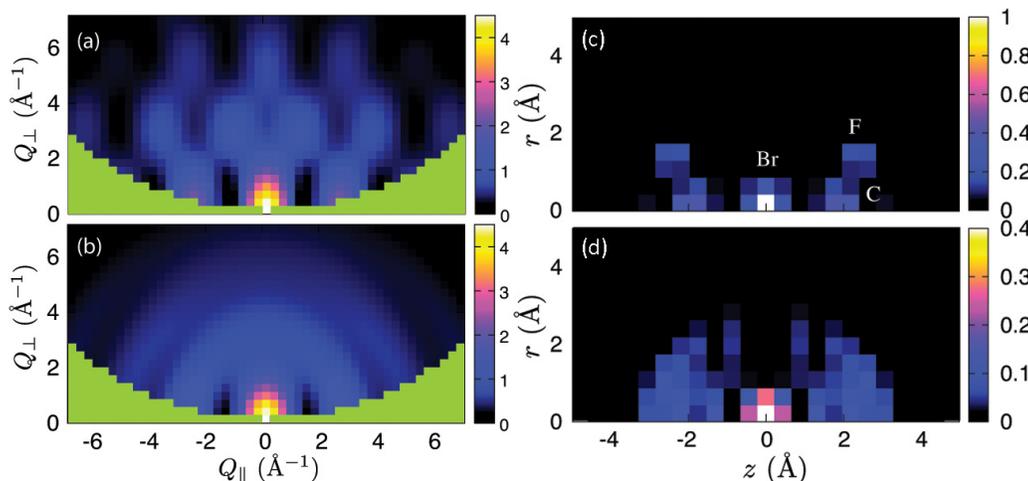


Figure 2. (a) X-ray diffraction pattern from  $\text{CF}_3\text{Br}$  molecules perfectly aligned with respect to horizontal axis perpendicular to the X-ray beam. (b) The same for imperfect alignment due to the thermal oscillations at 1 K. Green area corresponds to the non measurable missing edge due to the Ewald sphere curvature; (c)-(d) The azimuthally projected electron density holographically reconstructed from the simulated diffraction patterns (a) and (b), respectively.

## 3. NUMERICAL EXAMPLE

As the test object for our simulations we choose the *E. coli* chaperonin GroEL<sub>14</sub>-GroES<sub>7</sub>-(ADP·AlF<sub>x</sub>)<sub>7</sub> protein complex, constituted of 59,276 non-hydrogen atoms. The length of the complex is 20 nm, and diameter 14.5 nm. GroEL contains 14 identical subunits of molecular mass 58 kDa, and GroES contains 7 subunits of molecular mass 10 kDa. They form a structure consisting of three distinctive rings. The atomic coordinates have been obtained from the Protein Data Bank (entry 1SVT) and converted to cylindrical coordinates  $(r_k, \varphi_k, z_k)$ , where the subscript specifies atom  $k$ . Contributions of different cylindrical harmonics into diffraction pattern can be computed from the discrete counterpart of Eqs. (2) and (9):

$$G_m(R, \zeta) = \sum_k f_k \exp\{i[\zeta z_k - m(\varphi_k - \pi/2)]\} J_m(Rr_k), \quad (17)$$

where  $f_k$  is the form factor of atom  $k$ . Azimuthal projection of the protein complex computed by inverse Fourier-Hankel transform (10) of  $G_0(R, \zeta)$  is shown in Fig. 3(a). Since the structure of GroEL-GroES complex is characterized by a 7-fold rotational symmetry about a long molecular axis, only expansion coefficients for  $m = 0, \pm 7, \pm 14, \dots$  will have non-vanishing values. A diffraction pattern was calculated on a regular  $128 \times 64$  grid to the maximum scattering vector in both  $R$  and  $\zeta$  directions of  $5 \text{ nm}^{-1}$ . Incident X-ray beam polarization and grid distortion due to the Ewald sphere curvature are not considered, since these effects can be accounted for during data post-processing. The presence of the wedge with missing data is also ignored, because it is small for low resolution diffraction pattern (3 pixels at the detector edge for incident X-ray energy of 10 keV), and the values of the data in the missing wedge can be estimated by allowing them to float in the iterative phasing algorithm. With molecular radius of 7.25 nm, cylindrical harmonics of the order up to  $m = 35$  should be included in the sum (13), but for the test purposes we truncate the expansion at  $m = 14$ . The diffraction

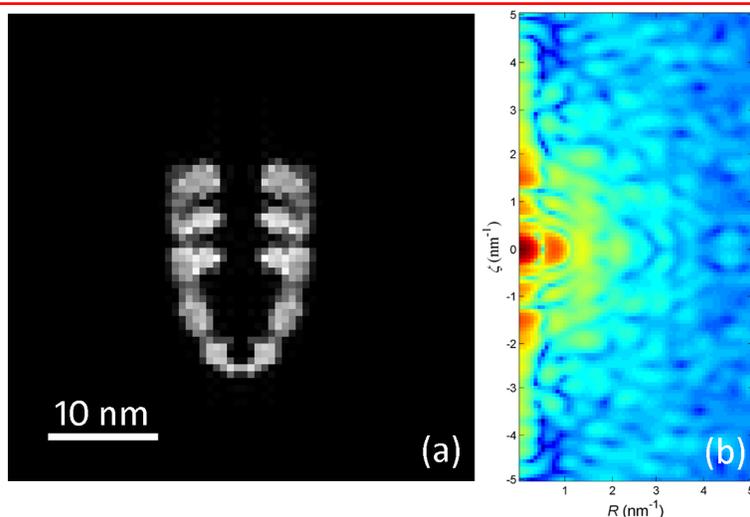


Figure 3. (a) Azimuthal projection  $g_0(r, z)$  of the GroEL-GroES protein complex. (b) Cylindrically averaged diffraction pattern for the GroEL-GroES macromolecule simulated from the atomic coordinates.

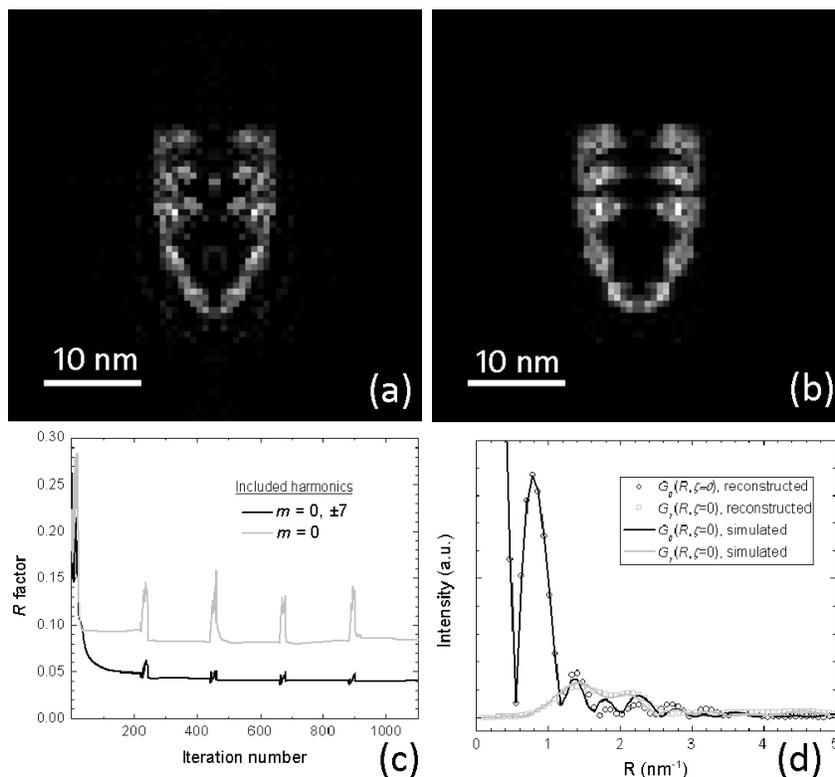


Figure 4. Azimuthal projection of the GroEL-GroES electron density reconstructed from the diffraction pattern in Fig. 3(b) using (a)  $m = 0$  cylindrical harmonics only, and (b)  $m = 0, \pm 7$  harmonics of scattering amplitude to represent scattered intensity. (c) Real space root mean square error for the phasing algorithm for different number of cylindrical harmonics included in the diffraction pattern decomposition. (d) Contributions of the basis functions with  $m = 0$  and  $m = 7$  (solid black and gray lines, respectively) into the total molecular scattering amplitude in the plane  $\zeta = 0$ , simulated from atomic coordinates. Symbols of corresponding colors show reconstructed contributions.

pattern computed from Eqs. (15) and (17) using tabulated atomic form-factors<sup>13</sup> is shown in Fig. 3(b). For the iterative scheme, we employed combination of the hybrid input-output (HIO) algorithm (20 cycles), based on the negative

feedback, and the error reduction algorithm (200 cycles), which simply projects back and forth between two constrained sets. Cylindrical support of the height of 23 nm and radius of 13 nm has been used, providing required oversampling. Convergence of the iterative procedure was monitored by error metric in real space, equal to the normalized amount of charge-density remaining outside the support. Fig. 4 compares algorithm performance for two cases: (1) the azimuthal projection  $m = 0$  only is included in the algorithm, and (2) the  $m = \pm 7$  terms are additionally used to synthesize the three-dimensional electron density map and diffraction pattern. The first procedure converges to a solution with high probability for a hypothetical exercise when only the 0th term of the scattering amplitude expansion contributes to diffraction pattern. However, after the higher degree terms are added to the diffraction pattern simulation, the performance of the algorithm drastically deteriorates. In particular, the success rate drops to 6%, and even then the final image significantly deviates from a true solution, as comparison of Fig. 3(a) and Fig. 4(a) reveals. Including terms  $m = \pm 7$  in the reconstruction algorithm results in increase of success rate to 49% in 200 trials with random initial phases. The reconstructed image of the molecular azimuthal projection after one run is shown in Fig. 4(b), and demonstrates much higher similarity to the original image than that in Fig. 4(a). The difference in the behavior of the two variants of the algorithm is also manifested in the behavior of the error metrics, illustrated in Fig. 4(c). When harmonics  $m = \pm 7$  are taken into account, the final error is twice as small as compared to the case when  $m = 0$  only is included. Additionally, error spikes during HIO cycles become much smaller. Relative values of the  $m = 0$  and  $m = 7$  contributions of the scattering amplitude at  $\zeta = 0$  as determined by the algorithm are shown in Fig. 4(d) by black circles and grey squares, respectively. They agree well with corresponding curves directly simulated from atomic structure, and shown by solid lines. Unfortunately, the algorithm was unable to separate correctly contributions from the  $m = 7$  and  $m = -7$  expansion coefficients for  $\zeta \neq 0$ , which prevented us from reconstructing a three-dimensional molecular envelope. However, the  $m = 0$  component was still properly extracted, allowing for the reconstruction of the azimuthal molecular projection. Note that the resulting decomposition of the scattering amplitude into components originated from different  $m$  is not sensitive to the input initial guesses for their contributions into the measured diffraction pattern.

#### 4. DISCUSSION

The theory of fiber diffraction is applied to fibrous molecules with helical structure, characterized by a repeat distance along the helical axis. The periodicity along this direction results in the appearance of the ‘layer lines’ in the diffraction pattern, separated in reciprocal space by an inverse repeat distance. An ensemble of the isolated molecules, aligned parallel to a single molecular axis, lacks a translational periodicity inherent to a fiber structure. Therefore, the composite diffraction pattern produced by an incoherent sum of diffraction patterns from these molecules, is continuous in all directions. This property allows for oversampling of the measured scattered intensity, required for application of iterative projection phase retrieval algorithms. We have shown that for molecules with high degree of the known rotational symmetry, when a three-dimensional scattering amplitude can be modeled by a few terms of expansion in cylindrical harmonics, such algorithms are capable of extracting a 0th term in this expansion, corresponding to the cylindrically averaged scattering amplitude. This is in turn related to the azimuthally averaged electron density by a two-dimensional transform, comprised by a Fourier transform in the direction of molecular alignment, and a Hankel transform of order zero in the radial direction. Then iterations of the Fourier-Hankel transform between two constrained sets allow reconstruction of the azimuthally projected electron density. This scheme is directly applicable to the scattering geometry, where only alignment along a single axis is achieved, and molecular rotation about this axis and translation motion are not restricted. This situation could be realized in the laser alignment, alignment by electric and magnetic field, or flow alignment.

#### ACKNOWLEDGEMENTS

DS was supported through the PULSE Institute at the SLAC National Accelerator Laboratory by the U.S. Department of Energy, Office of Basic Energy Sciences. JCHS and DKS jointly acknowledge support from DOE grant DE-SC0003141 and DKS also acknowledges grant DE-FG02-06ER46277 for financial support.

#### REFERENCES

- [1] Spence, J. C. H. and Doak, R. B., “Single molecule diffraction,” *Phys. Rev. Lett.* 92, 198102 (2004).
- [2] Marchesini, S., “A united evaluation of iterative projection algorithms for phase retrieval,” *Rev. Sci. Instrum.* 78, 011301 (2007).

- [3] Stapelfeldt, H. and Seideman, T., "Aligning molecules with strong laser pulses," *Rev. Mod. Phys.* 75, 543-557 (2003).
- [4] Brown, M. S., Shan, J. W., Lin, C. and Zimmermann, F. M., "Electrical polarizability of carbon nanotubes in liquid suspension," *Appl. Phys. Lett.* 90, 203108 (2007).
- [5] Pethig, R., [Dielectric and Electric Properties of Biological Materials], John Wiley & Sons, Chichester, New York, Brisbane, Toronto, 42-53 (1979).
- [6] Saldin, D. K., Shneerson, V. L., Starodub, D. and Spence, J. C. H., "Reconstruction from a single diffraction pattern of azimuthally projected electron density of molecule aligned parallel to a single axis," *Acta Cryst.* A66, 32-37 (2010).
- [7] Ho, P. J., Starodub, D., Saldin, D. K., Shneerson, V. L., Ourmazd, A. and Santra, R., "Molecular structure determination from x-ray scattering patterns of laser-aligned symmetric-top molecules," *J. Chem. Phys.* 131, 131101 (2009).
- [8] Millane, R. P., "Structure determination by x-ray fiber diffraction," in [Crystallographic Computing 4: Techniques and New Technologies], ed. by Isaacs, N. W. and Taylor, M. R., Oxford University Press, Oxford, 169-186 (1988).
- [9] Abramowitz, M. and Stegun, I. A., eds. [Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables], Dover, New York, 355 (1965).
- [10] Cochran, W., Crick, F. H. C. and Vand, V., "The structure of synthetic polypeptides. I. The transform of Atoms on a helix," *Acta Cryst.* 5, 581-586 (1952).
- [11] Wu, J., Leinenweber, K., Spence, J. C. H. and O'Keeffe, M., "Ab initio phasing of X-ray powder diffraction patterns by charge flipping", *Nature Materials* 5, 647-652 (2006).
- [12] Elser, V. and Millane R. P., "Reconstruction of an object from its symmetry-averaged diffraction pattern", *Acta Cryst.* A64, 273-279 (2008).
- [13] Chantler, C. T., "Theoretical form factor, attenuation and scattering tabulation for  $Z=1-92$  from  $E=1-10$  eV to  $E=0.4-1.0$  MeV", *J. Phys. Chem. Ref. Data* 24, 71-643 (1995).