Fiber Diffraction without Fibers

H.-C. Poon, P. Schwander, M. Uddin, and D. K. Saldin

Department of Physics, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin 53211, USA

(Received 12 September 2012; revised manuscript received 13 March 2013; published 27 June 2013)

Postprocessing of diffraction patterns of completely randomly oriented helical particles, as measured, for example, in so-called ''diffract-and-destroy'' experiments with an x-ray free electron laser can yield ''fiber diffraction'' patterns expected of fibrous bundles of the particles. This will allow ''single-axis alignment'' to be performed computationally, thus obviating the need to do this by experimental means such as forming fibers and laser or flow alignment. The structure of such particles may then be found by either iterative phasing methods or standard methods of fiber diffraction.

DOI: 10.1103/PhysRevLett.110.265505 PACS numbers: 61.05.cp, 61.05.jd

Fiber diffraction is responsible for some of the best-known work on the structure determination of matter. Examples are the structure of deoxyribose nucleic acid [1] and the structure of helical viruses; see, e.g., Ref. [2]. The long filamentous particles are drawn into a fiber which may be regarded as a bundle in which the particles have their long axes parallel to one another with random inter-particle distances and random azimuthal orientations. The seminal work in this field is that of Cochran, Crick, and Vand (CCV) [3]. If the scattering vector $\mathbf{q}$ is represented by the reciprocal-space cylindrical coordinates (R, ψ, ζ), it was shown by CCV that scattered intensity from a helix is found only at discrete values of $\zeta = \zeta_l = 2\pi l/c$, where c is the value of the repeat distance along the helix axis, and l is an integer specifying the so-called ''layer line'' on a diffraction pattern observed on a detector placed parallel to the fiber and perpendicular to the wave vector of an incident plane wave of radiation, e.g., x rays. CCV also deduced that the intensity of a particular layer line l may be written as

$$I(R, \zeta_l) = \sum_{n,n'} G_n(R, \zeta_l) G^*_{n'}(R, \zeta_l) \exp[i(n - n')\psi], \quad (1)$$

where

$$G_n(R, \zeta_l) = \sum_j f_j J_n(Rr_j) \exp(-in\phi_j) \exp(i\zeta_l z_j), \quad (2)$$

f_j is the atomic form factor of atom j with cylindrical coordinates (r_j, ϕ_j, z_j), and J is a Bessel function. The allowed values of n for a helix consisting of u subunits (e.g., proteins) per repeat distance along the c axis (the c-repeat unit [4]) consisting of ν turns of the helix (a u_ν helix) are determined by the helix selection rule

$$l = n\nu + mu, \quad (3)$$

where m is another integer. Thus, we see that for a given layer line l, the allowed values of n and m may differ only by u and ν, respectively [3] (where the differences have opposite signs). Since the azimuthal dependence of each term in the expression (1) for the intensity on a layer line is $\exp[i(n - n')\psi]$, each term will have an integral multiple of u-fold azimuthal symmetry.

The structure determination problem in so-called ''diffract-and-destroy'' experiments with an x-ray free electron laser (XFEL) [5] is at least superficially quite distinct. In that case, reproducible particles are injected into an XFEL beam in completely random 3D orientations, and would not therefore in general be expected to show the usual layer-line structure of an oriented helix. Yet we show in the following that appropriate computational processing of the ensemble of such diffraction patterns enables the reconstruction of just such a fiber diffraction pattern.

Because of the randomness of the 3D orientations, the natural choice for a reciprocal-space coordinate system in this case is the spherical rather than the cylindrical one. A way to perform structure determination in that case is via the determination of the average over a large number of the measured diffraction patterns of the angular correlations of their intensities [6]. This is quite similar to a method proposed earlier [7] for the extraction of structural information from an ensemble of randomly oriented identical molecules in solution. From the average over all measured diffraction patterns of angular autocorrelations of the measured intensities on a resolution ring of radius q, it is possible to extract [6,7] an orientationally independent quadratic function

$$B_L(q) = \sum_M I^*_{LM}(q) I_{LM}(q) \quad (4)$$

of the spherical harmonic coefficients $I_{LM}(q)$ of the scattered intensity distribution in 3D reciprocal space (L and M are the usual angular momentum quantum numbers). At least in the case of regular viruses, which Caspar and Klug [8] have suggested tend to be either icosahedral or helical, it is possible to find the 3D distribution of scattered intensities of a single particle from diffraction patterns of random particle orientations, as expected to be measured in above-mentioned experiments with an XFEL.

In the case of an icosahedral particle [9], the spherical harmonic expansion coefficients may be written as

$$I_{LM}(q) = g_L(q)a_{LM}, \tag{5}$$

where the $g_L(q)$ is a set of (real) icosahedral harmonic expansion coefficients and a_{LM} is a set of real coefficients calculable from a stipulation of the point group symmetry (e.g., Ref. [10]). Substitution of Eq. (5) into Eq. (4) yields

$$B_L(q) = g_L^2(q). \tag{6}$$

Consequently, one may deduce the magnitudes of the icosahedral harmonic expansion coefficients from the square roots of the $B_L(q)$'s. The signs of the quantities $g_L(q)$ may be determined from quantities related to the so-called "ring triple correlation (RTC) function" [11]

$$C_3(q, \Delta\phi) = \langle I(q, \phi)^2 I(q, \phi + \Delta\phi)\rangle_{DP} \tag{7}$$

calculable from the intensities $I(q, \phi)$ of resolution ring q and azimuthal angle ϕ in the measured diffraction patterns. As with the ring autocorrelation function [12] used to find $B_L(q)$, due to the randomness of particle orientations over the ensemble of diffraction patterns [6,7], the correlation function on the left-hand side does not depend on the particular value of ϕ chosen. With a full knowledge of the amplitudes and phases of the quantities $g_L(q)$, the 3D distribution of scattered intensities from a single particle may be found via

$$I(\mathbf{q}) = \sum_L g_L(q)\mathcal{I}_L(\theta, \phi), \tag{8}$$

where the quantities $\mathcal{I}_L(\theta, \phi)$ are the so-called "icosahedral harmonics" defined by

$$\mathcal{I}_L(\theta, \phi) = \sum_M a_{LM}Y_{LM}(\theta, \phi). \tag{9}$$

In the case of a helical particle, such as a helical virus, the same problem of the sum over M in Eq. (4) may be overcome by a different argument. Since the allowed values of n differ by at least u, if we take the ζ axis of the spherical coordinate system also to be parallel to the helix axis, there will be no values of M (for the scattered intensity) between 0 and $\pm u$ on any layer line. But M (the same quantum number in a spherical coordinate system with the same ζ axis) cannot take values of $\pm u$ until L becomes equal to at least u. In the case of a tobacco mosaic virus (TMV), which is a 49_3 helix, this means that from the properties of angular momenta, for all values of L up to $L = 48$ needed to describe the diffraction volume up to a resolution of about 12 Å, the only permitted value of M is zero. The 3D scattered intensity distribution of a helical virus may be found by first finding the spherical harmonic expansion coefficients of a single repeat unit of height c, which in the case of a TMV is 69 Å. In general, for a particle of radius R, a diffraction volume up to a reciprocal-space radius $q_{\max}$ may be calculated with the use of angular momentum quantum numbers up to $L_{\max}$, where [13]

$$L_{\max} = q_{\max}R. \tag{10}$$

In the case of a TMV, the radius R may be taken to be 34.5 Å, half the c-repeat length in the direction of the c axis and about 100 Å in the (the helix radius) direction perpendicular to this axis. Thus, the use of angular momentum quantum numbers up to $L_{\max} = 49$ allows the accurate calculation of the diffraction volume up to $q_{\max} = L_{\max}/R$, or until at least $q_{\max} \simeq 0.5$ Å^{-1}, i.e., a real-space resolution of $2\pi/q_{\max} \simeq 12$ Å, if sampled at the layer planes (which is what the layer lines of fiber diffraction become in 3D reciprocal space). Since the diffraction volume of an entire helical virus is just that of a single c-repeat unit sampled at values of q permitted by a "shape transform" factor due to repeated units of three helical turns of a single TMV particle, it follows that the diffraction volume of a helical virus may be determined to a resolution of $2\pi R/u$ by including only the $M = 0$ term in Eq. (4). That is, up to a resolution of about 12 Å, the 3D "diffraction volume" from a TMV has perfect azimuthal symmetry.

If only the $M = 0$ term need be included in Eq. (4), we may determine a 3D diffraction volume up to this resolution limit from

$$I(\mathbf{q}) = \sum_L I_{L0}(q)Y_{L0}(\theta, \phi), \tag{11}$$

with the magnitudes of the (real) coefficients $I_{L0}(q)$ determined by

$$|I_{L0}(q)| = \sqrt{B_L(q, q)}, \tag{12}$$

and their signs determined as above from the RTC function [11]. Once an oversampled [14] diffraction volume of a single particle has been determined from Eq. (11), the electron density of the particle that gave rise to it may be determined by an iterative phasing algorithm, e.g., Ref. [15].

$B_L(q)$ and $T_L(q)$ may be extracted from the averages over all measured XFEL diffraction patterns (from random particle orientations) of the angular autocorrelation function [7,9] and the RTC function [11]

$$C_3(q, \Delta\phi) = \int T_L(q)P_L[\cos(\Delta\phi)], \tag{13}$$

where

$$T_L(q) = \sum_{L_1L_2} G(L_1 0;, L_2 0; L0)I_{L_1 0}(q)I_{L_2 0}(q)I_{L0}(q), \tag{14}$$

and G is a Gaunt coefficient.

For the purposes of an initial proof of principle, we simulated these quantities as follows. The relationship between the azimuthally symmetric cylindrical harmonic expansion coefficients $I_0(R, \zeta_l)$ and the corresponding spherical harmonic expansion coefficients $I_{L0}(q)$ of the same 3D intensity distribution is

$$I_0(R, \zeta_l) = \sum_L I_{L0}(q)P_L(\zeta_l/q), \tag{15}$$

where $q = \sqrt{R^2 + \zeta_l^2}$, and P_L is a Legendre function. Now

$$I_0(R, \zeta_l) = \sum_n G_n(R, \zeta_l) G_n^*(R, \zeta_l) \tag{16}$$

and so is calculable from the atomic coordinates according to Eq. (2). One may calculate the spherical harmonic expansion coefficients by inverting the equation

$$I_0(R, \zeta) = \sum_L I_{L0}\left[q\left(= \sqrt{R^2 + \zeta^2}\right)\right] P_L(\zeta/q). \tag{17}$$

The quantities $I_{L0}(q)$ may then be used to calculate $B_L(q)$ from Eq. (4) with the single term $M = 0$ on the right-hand side and $T_L(q)$ from Eq. (14).

Assuming that $B_L(q)$ and $T_L(q)$ are the only quantities known (as from a real XFEL experiment), we recovered the magnitudes of the (real) spherical harmonic expansion coefficients $I_{L0}(q)$ from Eq. (12). From the form of this equation, it is clear that the magnitudes of the $T_L(q)$'s will be sensitive to the signs of the $I_{L0}(q)$ [11]. Consequently, these signs may be recovered by optimization of the $T_L(q)$'s Eq. (14). The recovered $I_{L0}(q)$ coefficients were then used to generate a 3D diffraction volume from Eq. (11). This diffraction volume will have contributions from only the azimuthally symmetric $M = 0$ terms up to a maximum value of the scattering vector corresponding to a resolution of about 12 Å. A slice through this diffraction volume passing through the origin of reciprocal space and parallel to the ζ axis will be expected to be identical to a fiber diffraction pattern and consist of the series of layer lines, as indeed appears to be the case (Fig. 1).

This is quite a remarkable result. Using true fibers, the structure of a TMV has been determined up to a resolution of 2.9 Å [16]. However, attempts to align particles by the other means such as the electric fields of powerful lasers, flow alignment, etc., have encountered the obstacle of the entropic tendency to disorder at any finite temperature [17]. What we have demonstrated here is the ability to produce a near-perfect fiber diffraction pattern of a TMV up to a resolution of about 12 Å by postprocessing of diffraction patterns from particles completely randomly oriented in 3D. This opens the way to the use of fiber diffraction methods for structure solution of particles prepared in random orientations, e.g., single particle diffract-and-destroy experiments with an x-ray free electron laser without the need to form an oriented bundle in the form of a fiber.

The key to this result is that up to about 12 Å resolution, the distribution of scattered intensities of a single c-repeat unit of a TMV is describable by a spherical harmonic expansion with angular momentum quantum number $L < 49$. However, on the layer planes, the magnetic quantum number may only take on possible values of 0, ± 49, etc. Since the properties of angular momenta require that $M \leq L$, the only permitted value of M is zero up to this

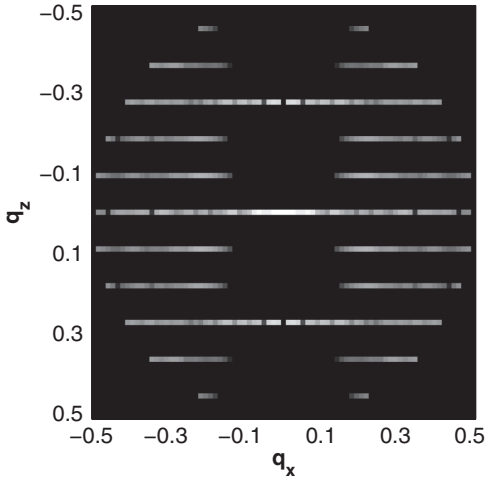

FIG. 1. Fiber diffraction pattern of a TMV (with intensities on a logarithmic scale) recovered from simulations of quantities expected to be measured in an XFEL diffract-and-destroy experiment from individual viruses in completely random 3D orientations. The axes are reciprocal-space coordinates in units of $Å^{-1}$.

resolution. Up to this resolution, therefore, the intensity is azimuthally symmetric, exactly as in fiber diffraction of a TMV. What is more, with this choice of axis, since there is only a single value of M in the summation in the right-hand side of Eq. (4), the magnitudes of the $I_{L0}(q)$ coefficients may be determined directly from the square roots of the experimentally accessible quantities $B_L(q, q)$ and their signs determined by optimizing the other experimentally determinable quantities, $T_L(q)$ [11]. Consequently, the entire 3D diffraction volume of a single particle may then be reconstructed from Eq. (11). A low resolution fiber diffraction pattern is simply a slice through this volume.

It should be noted that $B_L(q, q)$ is an orientationally independent quantity (it is a result of calculating the average of the angular correlations over all measured diffraction patterns from random particle orientations). On the other hand, the values $I_{LM}(q)$ of the spherical harmonic expansion coefficients of the diffraction volume do depend of the choice of axes. This means that the relationship (4) is valid for any choice of ζ-axis orientation, and allows the freedom of choice of orientation of this axis for the definition of the expansion coefficients $I_{LM}(q)$. Hence, it is possible to extract the diffraction volume of a single particle orientation from the angular correlations of the intensities over a large number of completely random particle orientations. Consequently, although the quantity $B_L(q)$ is determined experimentally by the scattering from a large number of completely randomly oriented particles, its value may be determined by assuming all particles to be perfectly aligned with respect to an arbitrary orientation. In reconstructing the spherical harmonic expansion coefficients $I_{LM}(q)$ from $B_L(q)$, the expansion coefficients may therefore be defined with respect to the same ζ axis whose

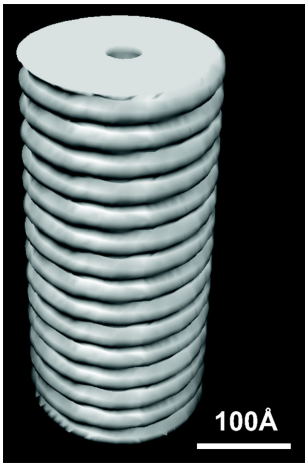

FIG. 2 (color online). Real-space image of a portion of a TMV recovered by an iterative phasing algorithm from an oversampled [14] low-resolution 3D diffraction volume generated from just the $M = 0$ components of the spherical harmonic expansion coefficients recovered from quantities expected to be calculated from the ensemble of XFEL diffraction patterns from random orientations of the virus.

direction may be chosen for computational convenience. This is equivalent to computational alignment. The same principle has already been demonstrated [9] in the reconstruction of the diffraction volume and hence structure of an icosahedral virus from simulated XFEL diffraction patterns from random particle orientations. In that case, the icosahedral particles' diffraction volume is oriented with the fivefold axis along ζ. In the case of a helical particle, by choosing the ζ axis parallel to the helix axis, one reconstructs the diffraction volume of the helix with its axis along ζ. With this choice of axis, up to a resolution of about 12 Å, or a magnitude of the angular momentum quantum number L of about 48, $M = 0$ is the only permissible value of the magnetic quantum number. This allows the determination of the magnitudes and signs of the only (real) nonzero values of the spherical harmonic expansion coefficients of the diffraction volume in this preferred orientation.

The ultimate aim is to reconstruct the 3D real-space structure of the virus. Accordingly, we attempted to recover the 3D real-space structure from the reconstructed 3D diffraction volume by means of an iterative phasing algorithm [15]. The result is shown in Fig. 2.

The essential features of the structure of a TMV are correctly reconstructed, such as the tubular shape with a central bore and the helical grooves on the outer surface of the virus. Although the image may not appear of quite 12 Å resolution, this limitation must lie in the particular phasing algorithm used, since the fiber diffraction pattern of a TMV is calculable up to this resolution. Indeed, for the final step of going from a recovered fiber diffraction pattern to a real-space image, a standard method of fiber diffraction may be used, which is known to be able to reconstruct a real-space image to the full extent of the diffraction data, although such methods [2,18] require additional information such as diffraction data from a heavy-atom derivative, or low-resolution information from, e.g., electron microscopy. The advantage remains over pure electron microscopy methods used for example to determine the structure of C nanotubes [19], that the necessity to exactly align the fibers is eliminated.

Of course, since we have demonstrated the ability to reconstruct a fiber diffraction pattern (Fig. 1) from quantities extractable from data measurable in a diffract-and-destroy XFEL experiment on particles randomly oriented in 3D, the possibility also exists of structure determination from this reconstructed fiber diffraction pattern by standard methods of fiber diffraction (e.g., Ref. [18]). As with the use of complementary information from a heavy atom derivative or electron microscopy in standard methods of fiber diffraction, the method proposed here could be used as a starting point to the determination of a helical structure to higher resolution by standard methods of fiber diffraction.

We acknowledge support for this work from DOE Grant No. DE-SC0002141 and the UWM Research Growth Initiative.

[1] J. D. Watson and F. H. C. Crick, Nature (London) **171**, 737 (1953).
[2] K. Namba and G. Stubbs, Science **231**, 1401 (1986).
[3] W. Cochran, F. H. Crick, and V. Vand, Acta Crystallogr. **5**, 581 (1952).
[4] R. P. Millane, Acta Crystallogr. Sect. A **47**, 449 (1991).
[5] R. Neutze, R. Wouts, D. van der Spoel, E. Weckert, and J. Hajdu, Nature (London) **406**, 752 (2000).
[6] D. K. Saldin, V. L. Shneerson, R. Fung, and A. Ourmazd, J. Phys. Condens. Matter **21**, 134014 (2009).
[7] Z. Kam, Macromolecules **10**, 927 (1977).
[8] D. L. D. Caspar and A. Klug, Cold Spring Harbor Symposia on Quantitative Biology **27**, 1 (1962).
[9] D. K. Saldin, H.-C. Poon, P. Schwander, M. Uddin, and M. Schmidt, Opt. Express **19**, 17318 (2011).
[10] A. Jack and S. C. Harrison, J. Mol. Biol. **99**, 15 (1975).
[11] Z. Kam, J. Theor. Biol. **82**, 15 (1980).
[12] D. K. Saldin *et al.*, New J. Phys. **12**, 035014 (2010).
[13] J. B. Pendry, *Low Energy Electron Diffraction* (Academic, London, 1974).
[14] J. W. Miao, P. Charalambous, J. Kirz, and D. Sayre, Nature (London) **400**, 342 (1999).
[15] G. Oszlányi and A. Süto, Acta Crystallogr. Sect. A **61**, 147 (2005).
[16] K. Namba, R. Pattanayek, and G. Stubbs, J. Mol. Biol. **208**, 307 (1989).
[17] J. C. H. Spence, K. Schmidt, J. S. Wu, G. Hembree, U. Weierstall, B. Doak, and P. Fromme, Acta Crystallogr. Sect. A **61**, 237 (2005).
[18] A. N. Barrett, J. Barrington Leigh, K. C. Holmes, R. Leberman, E. Mandelkow, P. von Sengbusch, and A. Klug, Cold Spring Harbor Symposia on Quantitative Biology **36**, 433 (1972).
[19] J. Ziang and J. M. Zuo, Carbon **47**, 3515 (2009).