International Tables for Crystallography (2012). Vol. F, Section 9.3.1, pp. 234–235.

Chapter 9.3. X-ray diffraction imaging of whole cells

D. Shapiro

9.3.1. Introduction

The brightness of third-generation synchrotron X-ray sources removes the need to enhance a diffracted X-ray signal with crystallographic redundancy for certain classes of samples. This is particularly useful for samples for which there exists only one unique structure or for which arrangement into a crystalline form is exceedingly difficult. Single-particle X-ray diffraction microscopy, also known as coherent X-ray diffraction microscopy (CXDM), treats an isolated non-crystalline sample as a crystallographer treats a crystal. The far-field diffraction intensity pattern of the sample is measured and the phase problem is solved computationally, allowing for structure recovery through a Fourier transform. Since Mother Nature restricts us to only measuring intensities, all methods of phase retrieval rely on decoding phase information that has been encoded in intensity measurements. Crystallographers, and the subsequent diffraction microscopists, have two basic methods for doing this. The first, the holographic method, mixes a known reference signal with the desired and unknown signal. The coherent superposition of these two signals results in measureable intensity variations that are directly related to the relative phase. This is, indirectly, a phase-measuring method. The second method relies only on the signal from the unknown structure and additional information, supplied by the scientist, which is physically plausible. This is a constraint-based method. Both methods seek to fill the information deficit inherent in intensity measurements by adding something that may unlock the phase.

In 1980, David Sayre suggested that CXDM should be possible (Sayre, 1980). Algorithms being developed for electron microscopy in the early 1970s would establish the constraint-based paradigm of iterative phase retrieval from Fourier modulus measurements. In particular, the alternating projection algorithm of Gerchberg and Saxton, the error-reduction algorithm, was developed to reconstruct the phase information that is missing when intensity measurements are made in both real and reciprocal space (Gerchberg & Saxton, 1972). This algorithm was later modified by Fienup into the input–output algorithm to handle cases where only one intensity measurement is made (Fienup, 1978, 1982; Fienup *et al.*, 1982; Miao *et al.*, 1998). In this case, the required real-space constraint restricts the object to an area no larger than half the width of its autocorrelation. The combination of the input–output algorithm and the error-reduction algorithm was found to be a very robust method of image reconstruction using only Fourier domain intensities, but its success was not understood for several years. In 1982, Bates argued that the solutions to the phase problem are unique in two dimensions if the Fourier modulus is sampled on an interval at least twice as fine as the Bragg interval (Bates & Fright, 1983; Bates, 1982). The method of phase retrieval from diffraction patterns sampled between Bragg peaks became known as the oversampling phasing method (Sayre, 1991; Sayre & Chapman, 1995; Miao *et al.*, 1998). The oversampling method compensates for the information deficit by supplying a zero-density region of approximately known size in the object domain. The more finely the diffraction pattern is sampled, the larger this zero-density region is, although this only adds useful information up to a certain point. This addition of information through Fourier space sampling results in an overdetermined inverse problem with a unique solution.

Sayre considered that the real niche of CXDM would be imaging objects of a few microns at a resolution of a few nanometres using soft X-rays (1–10 nm wavelength). Indeed, since the initial demonstration of CXDM (Fig. 9.3.1.1), ten years ago at the time of writing, active research projects have developed at all major synchrotron facilities and the technique has been applied to a diverse set of scientific problems. Three-dimensional images of radiation-hard materials have been obtained at 15 nm resolution, while two-dimensional images of a whole cell at 30 nm and of a gold nanoparticle at 5 nm resolutions have been reported, among many others (Miao *et al.*, 2002; Shapiro *et al.*, 2005; Pfeifer *et al.*, 2006; Barty *et al.*, 2008; Schroer *et al.*, 2008; Nelson *et al.*, 2010). CXDM researchers have not yet achieved the 'holy grail' of cellular imaging, the imaging of a whole frozen hydrated cell in three dimensions, but at least two teams are pursuing this goal with recent success in the two-dimensional case (Huang *et al.*, 2009; Lima *et al.*, 2009). Lens-based X-ray microscopes have achieved this landmark with moderate resolution (Wang *et al.*, 2000; Weiss *et al.*, 2000; Larabell & Le Gros, 2004). Transmission X-ray microscopes (TXM) and scanning transmission X-ray microscopes (STXM) utilizing diffractive zone plate lenses have high throughput and the advantage of direct imaging, but are limited in resolution by the technological challenge of making efficient high-numerical-aperture lenses. TXM can now routinely image in three dimensions at 50 nm resolution, and 12 nm resolution has

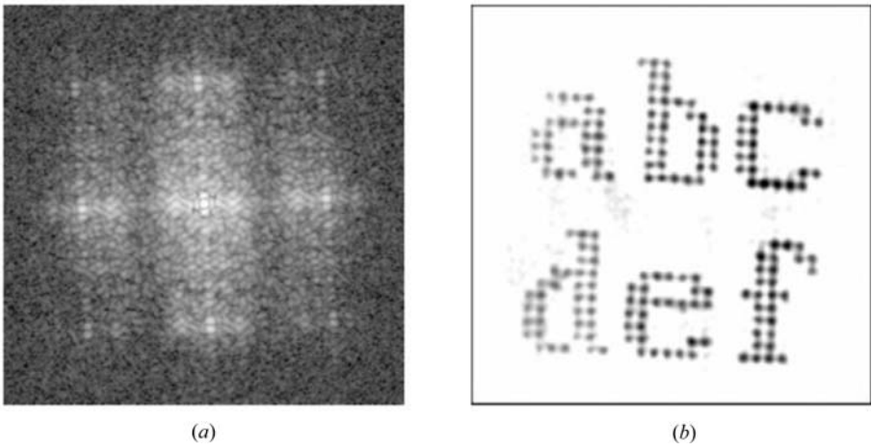

Figure 9.3.1.1
First soft X-ray demonstration of the CXDM method. (*a*) Diffraction pattern using 1.7 nm X-rays. (*b*) Reconstruction of (*a*) to 75 nm resolution. Reprinted by permission from Macmillan Publishers Ltd: *Nature* (Miao *et al.*, 1999), copyright (1999).

Table 9.3.2.1
Summary of various algorithms

The algorithms are, from top to bottom: error reduction, solvent flipping, hybrid input–output, difference map, averaged successive reflections, hybrid projection–reflection and relaxed averaged alternating reflections. A reflection is defined by its associated projection as $\mathbf{R} = 2\mathbf{P} - \mathbf{I}$, where $\mathbf{I}$ is the identity projection (Marchesini, 2007).

Algorithm	Iteration $\rho^{(n+1)} =$
ER	$\boldsymbol{P}_s\boldsymbol{P}_m\rho^{(n)}$
SF	$\boldsymbol{R}_s\boldsymbol{P}_m\rho^{(n)}$
HIO	$\begin{cases}\boldsymbol{P}_m\rho^{(n)}(\boldsymbol{r}), & \boldsymbol{r}\in S\\ (\boldsymbol{I}-\beta\boldsymbol{P}_m)\rho^{(n)}(\boldsymbol{r}), & \boldsymbol{r}\notin S\end{cases}$
DM	$\{\boldsymbol{I}+\beta\boldsymbol{P}_s[(1+\gamma_s)\boldsymbol{P}_m-\gamma_s\boldsymbol{I}]$ $-\beta\boldsymbol{P}_m[(1+\gamma_m)\boldsymbol{P}_s-\gamma_m\boldsymbol{I}]\}\rho^{(n)}$
ASR	$(1/2)[\boldsymbol{R}_s\boldsymbol{R}_m+\boldsymbol{I}]\rho^{(n)}$
HPR	$(1/2)\{\boldsymbol{R}_s[\boldsymbol{R}_m+(\beta-1)\boldsymbol{P}_m]+\boldsymbol{I}+(1-\beta)\boldsymbol{P}_m\}\rho^{(n)}$
RAAR	$[(1/2)\beta(\boldsymbol{R}_s\boldsymbol{R}_m+\boldsymbol{I})+(1-\beta)\boldsymbol{P}_m]\rho^{(n)}$

been demonstrated in two dimensions, although the total efficiency of the optic used was of the order of 1% (Chao *et al.*, 2009) at the highest spatial frequencies. Thus, at the cost of throughput and ease of use, a diffraction microscope provides increased X-ray efficiency and resolution.

The following sections first report on the current standard single-particle phase retrieval techniques and then on recent experiments in CXDM, which establish the state of the art in whole-cell imaging by diffractive methods. Images of dry yeast at 11 nm resolution are presented, which represent the highest resolution X-ray images of whole cells currently on record. The effects of radiation damage are discussed and Sayre's idea of using stereoscopic viewing as a means of obtaining quick and low-dose three-dimensional information is explored (Sayre, 2008).

9.3.2. Phase retrieval from single-particle diffraction data

The problem of phase retrieval is solved through successive application of constraints on the recovered object in the data and object spaces. Using the language of convex optimization, the mathematical operators which act on the data are projectors. The projector in the reciprocal (data) space forces the Fourier components to have the correct magnitude, while in object space finite support is enforced. To calculate the Fourier magnitude projector, one first needs to propagate the object density, ρ, to the data space by a Fourier transform, then replace the estimated magnitudes $|\tilde{\rho}|$ with the measured ones, $I^{1/2}$, and finally propagate back to real space. Using these transforms one simplifies the calculation of the projection, which becomes an element-wise operation on each recovered Fourier component. The forward $\mathcal{F}$ and inverse $\mathcal{F}^{-1}$ transforms must be incorporated into the operator defined in real space $\mathbf{P}_m$,

$$\mathbf{P}_m = \mathcal{F}^{-1}\tilde{\mathbf{P}}_m\mathcal{F}, \tag{9.3.2.1}$$

where the measured Fourier magnitudes are enforced in Fourier space by $\tilde{\mathbf{P}}_m$. Using the Fourier basis, one simply replaces the estimated magnitudes $|\tilde{\rho}|$ with the measured ones $I^{1/2}$, $[\tilde{\mathbf{P}}_m\tilde{\rho}(\boldsymbol{k}) = I(\boldsymbol{k})^{1/2}\tilde{\rho}(\boldsymbol{k})/|\tilde{\rho}(\boldsymbol{k})|]$. Similarly, in the object space the finite support constraint is applied on a per pixel basis through multiplication by the support mask. The corresponding projector is

$$\mathbf{P}_s\rho = S\cdot\rho.$$

Table 9.3.2.1 lists the combination of projections used by the most popular algorithms.

The violation of the support constraint is used as an error metric to monitor the convergence towards the solution. The solution should have zero density outside the support mask, so the error can be defined as the total density outside the support area,

$$\varepsilon_s^2(\rho) = |||\rho - S\rho|||^2 = ||[I-\mathbf{P}_s]\rho||^2. \tag{9.3.2.2}$$

Alternatively, the error metric can be defined in the data space as the difference between the measured and calculated magnitudes,

$$\varepsilon_m^2(\rho) = |||\mathcal{F}\rho| - I^{1/2}||^2 = ||[I-\mathbf{P}_m]\rho||^2. \tag{9.3.2.3}$$

In reality, the measured intensities are subject to noise which prohibits exact compliance with the constraints, so the error metrics cannot drop to zero. Although in most cases the algorithm can locate the global minimum, random noise will force fluctuations around the minimum. Once the algorithm reaches this steady-state regime, any particular iterate chosen as the solution would have a misleading degree of detail. On the other hand, the average of many fluctuating iterates would have reduced intensity in those Fourier components which are not reliably phased. The ratio of the average Fourier magnitude to the measured magnitude provides a measure of the stability, or reproducibility, of the retrieved phase information. This ratio as a function of spatial frequency is the phase retrieval transfer function (PRTF),

$$\mathrm{PRTF}(\mathbf{q}) = \frac{|\mathcal{F}[\langle\psi\rangle]|(\mathbf{q})}{[I_m(\mathbf{q})]^{1/2}}, \tag{9.3.2.4}$$

where $\mathcal{F}[\langle\psi\rangle]$ is the Fourier transform of the final averaged image and I_m is the measured intensity pattern. The PRTF is analogous to the differential phase residual of electron microscopy and, following Chapman *et al.* (2006), the resolution of a reconstruction is chosen as the spatial frequency at which the PRTF falls below a value of 0.5.

9.3.3. High-resolution imaging of yeast

The X-ray dose required to image a given volume of protein is nearly independent of energy above the oxygen K edge. At the same time, the photon flux required to image the same volume increases with E^2 because of the energy dependence of the scattering cross section (Howells *et al.*, 2009). For this reason, it is advantageous to use the lowest energy commensurate with the desired resolution of 5–10 nm. Commercially available charge-coupled device (CCD) detectors can easily provide a scattering angle of 0.1 radians [a 1 inch detector placed 5 inches from the sample (1 inch = 2.54 cm)], which results in a half-period resolution of 8 nm when using 750 eV X-rays. Furthermore, a cell with a diameter of 3 µm would have an oversampling ratio (number of intensity samples per speckle) of at least ten in this geometry if the detector has 20 µm pixels.

In this particular case, a Princeton Instruments CCD (PI-MTE:1300) is placed 136 mm downstream of a freeze-dried yeast cell using the CXDM instrument on Beamline 9.0.1 of the Advanced Light Source (ALS). The yeast cell is illuminated by a coherent beam of 750 eV X-rays defined by a 5 µm pinhole located 25 mm upstream. The incident intensity of 4×10^6 photons s^{-1} μm^{-2} is high enough to cause rapid structural changes to the cell (discussed in the next section), so the sample requires pre-irradiation for about 30 minutes prior to collection of the final data set intended for reconstruction. The