



Qun Shen

New Advance towards Solving the Phase Problem in X-ray Crystallography

Qun Shen

*CHESS and Department of Materials Science and Engineering
Cornell University*

Synchrotron X-ray crystallography is a widely-used method in structural biology and materials sciences for solving crystal structures with atomic-scale resolution. In a typical crystallography experiment, a large number of diffraction peaks or Bragg reflections are recorded on an area detector while rotating or oscillating a crystal specimen (Fig.1a). The intensity recorded for each Bragg reflection depends only on the magnitude of its structure factor, but not on its phase, which is also needed to determine the atomic positions in a crystal. This is the fundamental phase problem in diffraction and its general solution remains an active area of research.

As a part of a larger scientific initiative on crystallographic phasing and innovative X-ray scattering at CHESS, we have recently developed a phase-sensitive reference-beam diffraction (RBD) technique that has the potential to solve the phase problem in crystallography [1]. The technique is based on the principle of multiple-beam or three-beam diffraction [2] which has been known to contain the phase information but the profiles are usually measured one at a time which is very time-consuming. In the new method, we incorporate this principle into the most popular crystallographic data collection technique - the oscillating crystal method, which provides a parallel collection of many three-beam interference profiles and allows measurements of both magnitudes and phases of a large number of reflections.

As illustrated in Fig.(1), the new RBD method is a simple conceptual modification [1] to the conventional oscillation camera setup in the direct-beam geometry. Instead of being perpendicular to the incident X-ray beam, the oscillation axis in RBD geometry is tilted by the Bragg angle θ_G of a strong reference reflection, \mathbf{G} , which is oriented to coincide with the oscillation axis. In this way, reflection \mathbf{G} can be kept fully excited throughout the crystal oscillation, and the intensities of all Bragg reflections recorded on an area detector during such an oscillation can be influenced by the interference with the \mathbf{G} -reflected reference wave and thus are sensitive to the relative phases of the reflections involved.

The three-beam diffraction process that governs a RBD interference involves the reference reflection \mathbf{G} , a primary reflection \mathbf{H} , and a coupling reflection $\mathbf{H-G}$. Using a second-order Born approximation [3] or an expanded distorted-wave approximation [4], it can be shown that the interference effect is sensitive to the triplet phase angle: $\delta = \alpha_{\mathbf{H-G}} + \alpha_{\mathbf{G}} - \alpha_{\mathbf{H}}$, which is independent of the choice of origin in a unit cell. In addition to δ , the phase of the \mathbf{G} -reflected wave can be tuned by rocking the tilt angle θ through the \mathbf{G} -reflection rocking curve, much like in an X-ray standing-wave experiment. Observations of intensity $I_{\mathbf{H}}$ as a function of θ for each of the Bragg reflections recorded on the area detector yields a complete interference profile $I_{\mathbf{H}}(\theta)$, which is given by a normalized intensity [1]:

$$I_{\mathbf{H}}(\theta) = 1 + 2 \left| \frac{F_{\mathbf{H-G}}}{F_{\mathbf{H}}} \right| \sqrt{R_{\mathbf{G}}(\theta)} \cos[\delta + \nu_{\mathbf{G}}(\theta)], \quad (1)$$

where $R_{\mathbf{G}}(\theta)$ is the reflectivity and $\nu_{\mathbf{G}}(\theta)$ is the dynamical phase shift of the reference reflection \mathbf{G} . Both $R_{\mathbf{G}}(\theta)$ and $\nu_{\mathbf{G}}(\theta)$ are known and common to all reflections recorded on the area detector, and therefore any difference in $I_{\mathbf{H}}(\theta)$ between two recorded reflections is due entirely to the difference in phase δ .

We have demonstrated the RBD technique both on small molecule crystals and on real but good-quality biological crystals, such as tetragonal lysozyme [1]. Typically we use unfocused monochromatic 10 keV X-rays as an incident beam, a

standard 4-circle diffractometer for specimen orientation, an oscillation camera setup for exposure controls, and an image-plate or a charge-coupled device (CCD) as the area detector for data collection.

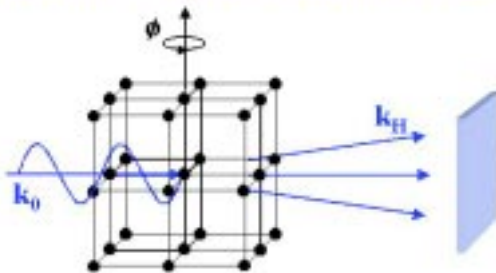
The first ever reference-beam diffraction image was obtained at the CHESS C-1 station in June of 1997 using a GaAs crystal with (004) as the reference reflection [1]. The data were recorded on an image plate that was operated in 'streak camera' mode in order to minimize the read-out errors for multiple exposures. The raw data image is shown in Fig.2. The RBD interference profiles of all recorded Bragg reflections are clearly visible in the image, as shown by the intensity variations along the sixteen multiple exposures taken as the rocking angle θ is stepped through the (004) rocking curve.

The experiments on tetragonal lysozyme were designed both to test the feasibility of the RBD method for small proteins and to develop a practical procedure for the measurements and data analysis [5]. Initial experiments on lysozyme were performed at CHESS C-1 and F-3 bend magnet stations, again using an image-plate in streak-camera mode. A typical diffraction image with $\mathbf{G}=(3,2,0)$, shown in Fig.3, contains nine θ steps with a 10 seconds exposure per step for a 1° oscillation range. It was found that the RBD interference profiles, even though much weaker than small molecule cases, were nevertheless detectable for as many as 20% of the diffraction spots.

One apparent drawback of using image-plates in streak-camera mode is the artificially increased high background due to the multiple exposures on the same plate during the RBD data collection. The best way to solve this problem is to use a CCD detector (Fig.4) that has a larger dynamic range and can be used for multiple exposures controlled electronically. These experiments were done at A-2 and C-1 stations and some examples of the results are shown in Fig.5. These RBD profiles are improved due to the increased number of θ steps and longer exposures times.

With the aim to make RBD a real practical technique that can be used not only by X-ray specialists but also by

(a) Conventional oscillation method



(b) Reference-beam diffraction

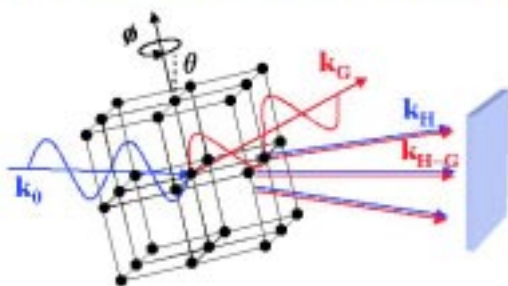


Figure 1: Schematics of (a) conventional oscillation setup used in X-ray crystallography, and (b) new reference-beam diffraction in Bragg-inclined geometry, where two sets of diffraction patterns (blue and red) interfere and generate a phase-sensitive diffraction image on an area detector.

crystallographers with a broad range of background, we are developing a data-analysis procedure that would automatically extract the triplet phase values for a large number of Bragg reflections collected in a RBD experiment. First, we use existing crystallographic software packages such as DENZO/SCALEPACK or DPS/MOSFLM/SCALA for indexing a reference-beam diffraction pattern and for integration of the recorded Bragg reflections, for each θ step, to provide a series of integrated intensities $I_H(\theta)$. Second, we employ a curve-fitting program to fit the $I_H(\theta)$ data series to a RBD interference function, Eq.(1), which directly yields the best-fit values of δ for all recorded reflections, as illustrated in Fig.5. Finally, we have a goodness-of-fit criterion that allows us to reject bad intensity profiles that for various experimental reasons do not actually contain the true interference information. So far we have obtained over 1300

triplet phases through automatic fitting, of which over 500 phases are likely to be considered reliable after applying the rejection criterion. It is expected that the reliability of measured phases would increase once the inverse-beam measurements of Friedel pairs [5] are fully integrated into the analysis. This would also allow an unambiguous determination of the enantiomorph for non-centrosymmetric structures, without the need for anomalous signals.

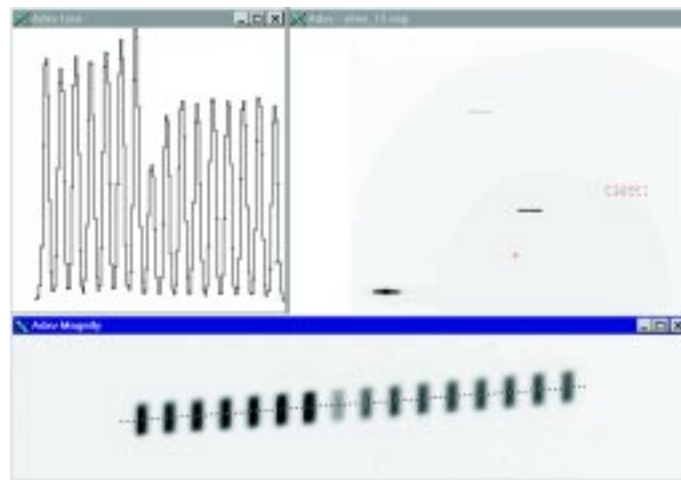


Figure 2: The first reference-beam diffraction oscillation image taken on a GaAs crystal using an image-plate. The oscillation range was 20° and the exposure time was 40 sec at each θ step. The interference profile along the 16 multiple exposures can be seen for all recorded reflections.

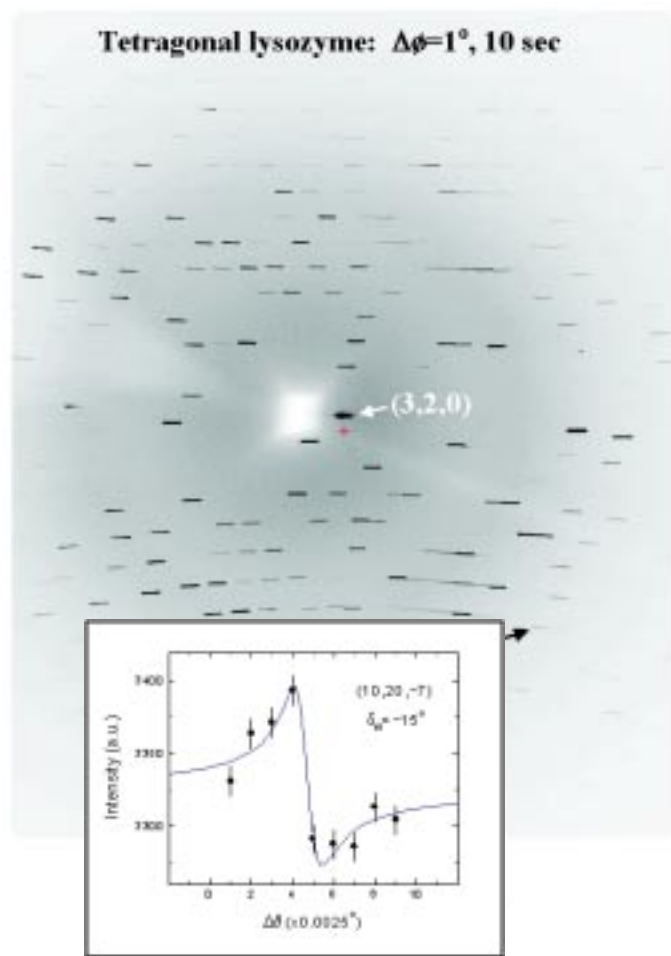


Figure 3: Reference-beam diffraction data on a tetragonal lysozyme with $G=(3,2,0)$, using an image-plate. Inset shows an example of the interference profile of one of the reflections as indicated by the arrow.

Finally, we have been collaborating with Dr. Herbert Hauptman's group at Hauptman-Woodward Institute (HWI) in Buffalo, N.Y. to exploit the best strategies to utilize the experimentally measured triplet phases. So far the most promising method is to combine these measured phases with traditional or dual-space direct-methods algorithms. One important and interesting question that we are trying to address is the following: With very little overlap among the individual structure-factor phases in a given RBD data set, how many such data sets with different G 's, either complete or partial, does one need to solve realistic structures? Preliminary simulations [6] using a unified Shake-n-Bake program suggest that a structural solution is possible for small proteins with even a single measured phase data set (single G) if the data-set is rather complete (though less than atomic resolution) and accurate with a measurement error of less than 20° or so in the measured triplet-phases. The tolerance on the phase errors can be as high as 50° if three or more RBD data sets are measured. These early results are encouraging but further tests are needed to reach a general conclusion.

In terms of data collection procedures, the RBD method presented here is very similar to multiple-wavelength anomalous diffraction (MAD) experiments. Here the angular setting of the reference reflection G serves the same role as the atomic absorption-edge energy of a heavy-atom. Multiple oscillation data sets are collected around the Bragg angle θ_G , much like those around the absorption edge, with similar useful signal levels of a few percent. In fact, instead of changing the rocking angle θ , one can change the incident energy to pass through the reference-reflection rocking-curve and take multiple data sets at several energies. An additional advantage is that no global scaling is necessary in RBD measurements so that incomplete data sets from different crystals can be easily combined.

Looking ahead, we feel that the reference-beam diffraction technique that we have developed at CHESS is a promising practical method of solving the phase problem in X-ray crystallography, without the need for heavy atoms. Several challenging issues need to be resolved before the technique can be widely adopted in everyday crystallography. These include how to handle increased crystal mosaicity upon freezing, and possible overlapping multiple reflections for larger structures. With further research and development, however, we believe that the new method will have a significant impact on crystallography data collection and crystal structure determination.

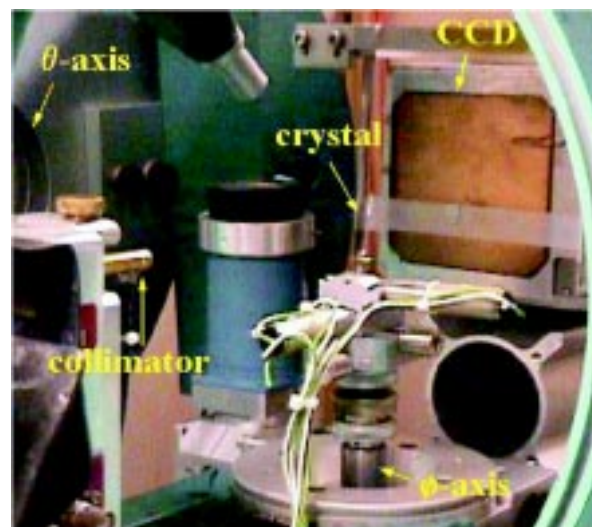


Figure 4: Reference-beam diffraction setup using a CCD detector on a standard 4-circle diffractometer.

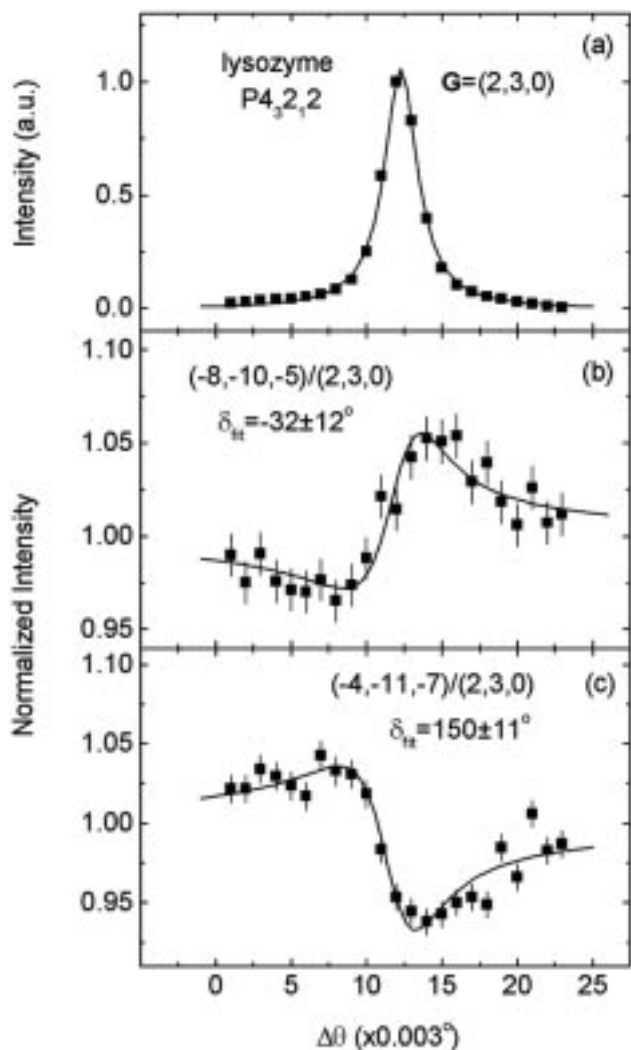


Figure 5: Two examples of the measured reference-beam diffraction profiles using a CCD detector, along with the rocking curve of the reference reflection $G=(2,3,0)$ shown in (a). The blue curves in (b) and (c) are best-fits to the data using Eq.(1) from which triplet phase values δ_{fit} can be obtained. The rocking curve is fit with a Lorentzian.

Acknowledgments

The author would like to acknowledge S. Kycia, I. Dobrianov, J. LaLuppa, P. Doing, M. Szebenyi, C. Heaton, W. Miller, and K. Finkelstein for participating and helping in this research. Thanks are due to B. Batterman, E. Fontes, S. Ealick, R. Thorne, and S. Gruner at Cornell, as well as R. Colella and M. Rossmann at Purdue for useful discussions. Collaborations with H. Hauptman, C. Weeks, H. Xu, and others at HWI are also appreciated. This work is supported by NSF through CHESS under Grant DMR 97-13424.

References

- [1] Shen, Q. *Phys. Rev. Lett.* **80**, 3268-3271(1998); *Research News, Science* **280**, 828 (1998); *Phys. Rev. B* **59**, 11109-11112 (1999)
- [2] Recent reviews on the topic include: Colella, R., *Comments Condens. Matter Phys.* **17**, 175-215 (1995); Weckert, E. and Hümmer, K., *Acta Cryst. A* **53**, 108-143 (1997); Chang, S-L., *Acta Cryst. A* **54**, 886-894 (1998)
- [3] Shen, Q. *Acta Crystallogr. A* **42**, 525-533 (1986)
- [4] Shen, Q. *Phys. Rev. Lett.* **83**, 4784-4787 (1999); *Phys. Rev. B* **61**, 8593-8597 (2000)
- [5] Shen, Q., Kycia, S. and Dobrianov, I. *Acta Cryst. A* **56**, 264-267; 268-279 (2000)
- [6] Weeks, C.M., Xu, H., Hauptman, H.A. and Shen, Q. *Acta Cryst. A* **56**, 280-283 (2000)