

## 19.2. ELECTRON DIFFRACTION OF PROTEIN CRYSTALS

## 19.2.4. Data processing

## 19.2.4.1. Data sampling

The principle of three-dimensional reconstruction is based on the central section theorem, which states that the experimental or computed projected diffraction pattern of a three-dimensional object is a plane that intersects the centre of the three-dimensional Fourier space in the direction normal to the direction of the projection (DeRosier & Klug, 1968). Because of the crystallographic symmetry inherent in a protein crystal, only a portion of the entire three-dimensional Fourier space, equivalent to an asymmetric unit of the crystal unit cell, is needed for the

reconstruction. The structure factors of a three-dimensional crystal are localized in the three-dimensional reciprocal lattice, whereas the structure factors of a two-dimensional crystal are distributed continuously along the lattice lines, each of which passes through the reciprocal lattice in the zero projection plane (Fig. 19.2.4.1) (Henderson & Unwin, 1975). The assignment of  $z^*$  for each observation ( $h, k, z^*$ ) along the lattice line is determined from the tilt angle and direction of the tilt axis for each image (Shaw & Hills, 1981). In general, the three-dimensional data set is initially built up from low-angle data and is gradually extended to the high-angle data. The angular parameters for each observed reflection are iteratively refined among one another within the whole data set. The

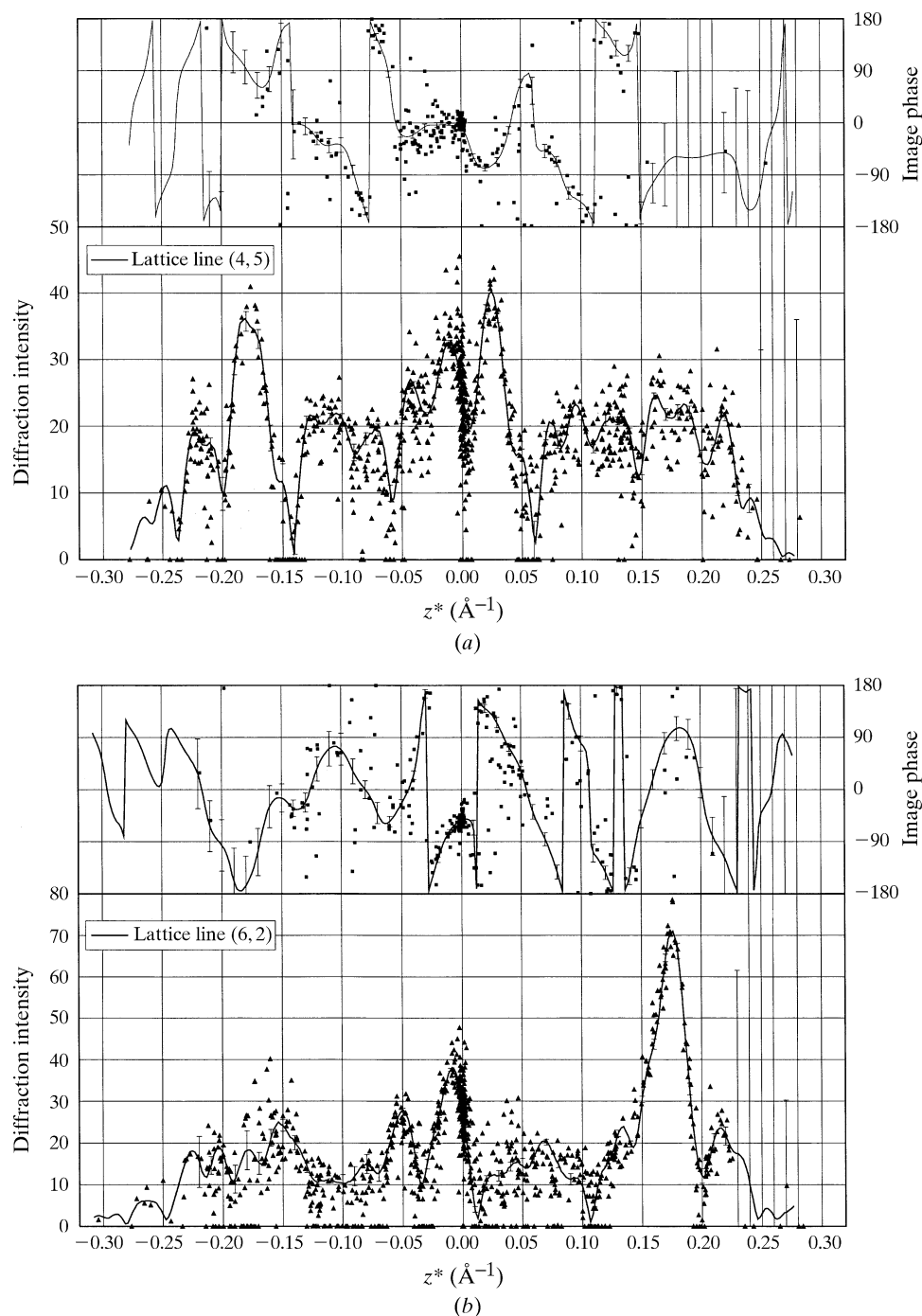


Fig. 19.2.4.2. Experimental intensities from electron diffraction patterns and phases from images of bacteriorhodopsin, recorded from tilted crystals in an electron cryomicroscope. Fitted curves for two representative lattice lines are shown: (a)  $(4, 5, z^*)$  and (b)  $(6, 2, z^*)$  (Courtesy of Drs Terushisa Hirai and Yoshinori Fujiyoshi at Kyoto University.)

## 19. OTHER EXPERIMENTAL TECHNIQUES

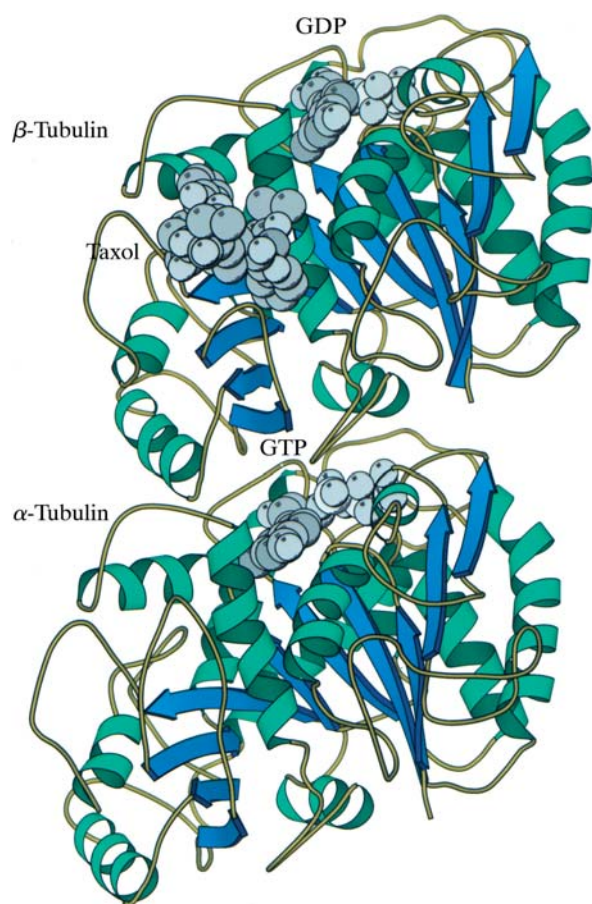


Fig. 19.2.4.3. Ribbon diagram of a tubulin dimer, whose structure has been solved to 3.7 Å resolution. GTP, GDP and taxol are shown as CPK models, with GDP at the top, bound to  $\beta$  tubulin, and GTP in the middle, bound to the  $\alpha$  subunit. This is the view as seen from inside a microtubule, with the plus end at the top. (Courtesy of Drs Eva Nogales and Kenneth Downing at the Lawrence Berkeley National Laboratory, University of California, Berkeley.)

required accuracy of these angular determinations depends on the thickness of the crystal and also on the desired resolution (Prasad *et al.*, 1990). For instance, the angular accuracy has to be  $< 0.1^\circ$  for 3.5 Å data in a 100 Å-thick crystal. The sampling of the data along the lattice line is generally about  $\frac{1}{4}$  the thickness of the crystal (Henderson *et al.*, 1990). The data points are not evenly sampled along the lattice lines; they must be fitted into continuous and smoothly varying functions within the constraint of the crystal thickness. These functions are interpolated onto a periodic lattice, so that its inverse Fourier transform can be computed to reconstruct the three-dimensional mass-density function of the object.

### 19.2.4.2. Amplitudes and phases

An electron-microscope image contains both the amplitudes and phases of the structure factors. The basic premise of the current image-reconstruction scheme assumes that the image intensity can be related to the structure factor linearly and can be retrieved by the Fourier transform of the image intensities. However, the structure factors,  $F(S)$ , are influenced by several instrumental factors, as shown in equations (19.2.4.1)–(19.2.4.3) below, whose parameters need to be determined for each image.

$$F_{\text{obs}}^2(S) = [F(S)\text{CTF}(S)E(S)]^2 + N^2(S), \quad (19.2.4.1)$$

$$\text{CTF}(S) = -\{(1 - Q^2)^{1/2} \sin[\gamma(S)] + Q \cos[\gamma(S)]\} \text{ and } \quad (19.2.4.2)$$

$$\gamma(S) = \pi[(-C_s \lambda^3 S^4)/2 + \Delta z \lambda S^2], \quad (19.2.4.3)$$

where  $F_{\text{obs}}$  is the structure factor computed from the electron cryomicroscopic images,  $F$  is the true structure factor, CTF is the contrast-transfer function,  $E$  is the product of many decay functions due to the electron optics and specimen movement,  $N$  is the background noise contributed by a variety of physical effects,  $S$  is the spatial frequency,  $Q$  is the fraction of amplitude contrast,  $C_s$  is the spherical aberration coefficient of the objective lens,  $\lambda$  is the wavelength and  $\Delta z$  is the image defocus.

In practice, it is tedious to determine all the parameters in these equations from images in order to make corrections to the amplitudes of the structure factors. In the case of crystals, the amplitudes of the structure factors can simply be obtained directly from the electron diffraction intensities, which are free from any of the above factors (Unwin & Henderson, 1975). The computational procedure used to calculate the diffraction spot intensities is similar to that used to measure an X-ray diffraction pattern (Baldwin & Henderson, 1984; Brink & Wei Tam, 1996). The quality of the diffraction intensity measurement is evaluated from the value of  $R_{\text{sym}}$  for Friedel-related reflections. The best data have  $R_{\text{sym}}$  less than 0.04. The consistency of the diffraction intensities among different patterns from different crystals is judged from  $R_{\text{merge}}$ , which is generally 0.15–0.25 (Kimura *et al.*, 1997; Nogales *et al.*, 1998). Fig. 19.2.4.2 is an example of the diffraction intensity for two lattice lines computed from bacteriorhodopsin crystals. The phases of the structure factors are computed from images.

In addition to the instrumental factors given in equations (19.2.4.1)–(19.2.4.3), however, images are generally imperfect because of bending of the crystal, specimen preparation, or magnification variations across an image. The consequence of these imperfections is a reduction of the signal-to-noise ratio in high-resolution reflections. A computational procedure called ‘unbending’ has been devised, which in effect fixes the image imperfection by finding the unit-cell deviation vectors and straightening them by interpolation (Henderson *et al.*, 1986). The effect of the instrumental factors is the modulation of the phases by the oscillating function CTF( $S$ ), as shown in equations (19.2.4.1)–(19.2.4.3). The result is that the phases flip by  $\pi$  at different frequencies, depending on the defocus setting (Erickson & Klug, 1970). In addition, there is a phase shift caused by a combination of factors, including lens astigmatism, beam tilt and specimen height variation in a tilted position. All of these factors have to be corrected for each micrograph before merging the phases of the reflections from different micrographs to a common phase origin. The determination of the phase origin is performed by phase residual difference minimization or correlation matches among different micrographs (Amos *et al.*, 1982; Thomas & Schmid, 1995). Intensities and corresponding phases of two lattice lines are shown in Fig. 19.2.4.2. The fitted curves show the matches among the data points, each of which is from a different image or from different symmetry-related reflections from the same image.

In electron crystallography, the correctness of the phases can be evaluated by the self-consistency of the merged data sets and also by the phase residual difference of the symmetry-related reflections according to the two-dimensional plane-group symmetry. For two-dimensional crystals, there are only 17 possible plane groups (Amos *et al.*, 1982). As in the case of a three-dimensional crystal, the plane group is determined from the symmetry of the phases, the unit-cell parameters and the pattern of forbidden reflections. The plane-group assignment can be confirmed by the phase equivalence of symmetry-related reflections. Furthermore, the reliability of the