

Evaluation of Reflection Intensities for the Components of Multiple Laue Diffraction Spots. III. Using a Real-Space Density Modification Method

Q. Hao,^{a,b*} M. M. Harding^b and J. W. Campbell^a

^aDRAL Daresbury Laboratory, Warrington, Cheshire WA4 4AD, UK, and ^bDepartment of Chemistry, University of Liverpool, Liverpool L69 3BX, UK

(Received 7 June 1994; accepted 19 August 1994)

In a Laue diffraction pattern, 10–20% of the spots result from the exact superposition of two or more reflections that are ‘harmonics’, e.g. hkl ; $2h\ 2k\ 2l$, ...; a high proportion of these are low-resolution reflections. For the solution of large or difficult structural problems, the intensities of the remaining 80–90% of the reflections, measurable as singles, may not be sufficient and thus the evaluation of the intensities of the components of the multiple spots is important. A new procedure, *DECONV*, for this deconvolution using real-space density modifications on the Patterson map is given. This development is based on a procedure in reciprocal space related to direct methods [Hao, Campbell, Harding & Helliwell (1993). *Acta Cryst.* **A49**, 528–531]. A Patterson map is calculated using single reflections and modified by removing negative densities and squaring with an appropriate adjustment at the origin peak. The procedure can be repeated until convergence is reached. It has been tested with Laue diffraction data from 4Zn insulin and cytochrome *c* peroxidase (CCP). 304 and 1134 reflection intensities were evaluated from multiple spots of insulin and CCP, respectively; the mean fractional differences (on F), showing the agreement with high-quality monochromatic data, were 0.27 and 0.21, respectively.

Keywords: Laue diffraction; multiple reflections; Patterson density modification; deconvolution.

1. Introduction

With the advent of synchrotron radiation, there has been renewed interest in the use of Laue diffraction as a method for the measurement of diffraction intensities (Helliwell, Habash, Cruickshank, Harding, Greenhough, Campbell, Clifton, Elder, Machin, Papiz & Zurek, 1989; Smith Temple & Moffat, 1987) and for time-resolved studies of crystal structures (Schlichting, Almo, Rapp, Wilson, Petratos, Lentfer, Wittinghofer, Kabsch, Pai, Petsko & Goody, 1990; Johnson & Hajdu, 1990; Szebenyi, Bilderback, LeGrand, Moffat, Schildkamp, Smith Temple & Teng, 1992; Lindahl, Liljas, Habash, Harrop & Helliwell, 1992; Singer, Smalås, Carty, Mangel & Sweet, 1993). In particular, Singer *et al.* (1993) have revealed the hydrolytic water molecule in trypsin by time-resolved Laue crystallography. Usually, 80–90% of the spots in a Laue diffraction pattern correspond to single reflections, each with its values of hkl and associated d (plane spacing) and λ ; these are described as singles, the remaining 10–20% of the spots are doubles, triples or higher multiples (Cruickshank, Helliwell & Moffat, 1987). A multiple spot contains a reflection from a plane of spacing d , and from planes with spacings $d/2$, $d/3$ or, in general, d/j , where j is any positive integer. Thus, for multiples the measurement of the spot intensity does not

directly give the intensities of each individual reflection. Moreover, the reflections that cannot be straightforwardly measured as singles are not randomly distributed in reciprocal space (Cruickshank *et al.*, 1987); a high proportion of them are low-order reflections, axial reflections and reflections in special planes ($hk0$, hhl etc.). The absence of these reflections can be a serious drawback if the data are to be used for structure solution, for example using direct methods. In protein crystallography, the absence of low-order reflections has been shown to give electron-density maps that have poor connectivity (Duke, Hadfield, Walters, Wakatsuki, Bryan & Johnson, 1992). For reasons such as these, there has been increasing interest in methods to deconvolute reflection intensities from spots that are multiples. Helliwell, Habash *et al.* (1989) described one procedure for this deconvolution that uses the intensities of spots on successive films in a film pack and the variation of film absorption with λ (program *UNSCRAM*). In a second method, when multiple spots have been recorded with redundancy the different components of a spot may be deconvoluted at the wavelength normalization stage (Campbell & Hao, 1993). A test on hen egg-white lysozyme showed that the multiples deconvoluted by this method did indeed contribute usefully to the continuity of the maps, owing to the improved completeness of the data (Campbell, Deacon, Habash, Helliwell, McSweeney, Hao, Raftery & Snell, 1994). However, a method that does not

* Present address: Department of Applied Physics, De Montfort University, Leicester LE1 9BH, UK.

depend on the recording of data on multi-film packs or redundancy of the data would be desirable. Hao, Campbell, Harding & Helliwell (1993) have described a third procedure for carrying out the deconvolution; it uses relationships between structure-factor magnitudes like those that are used in direct methods. In this paper, we present a fourth method (*DECONV*) which is similar in principle to the third, but uses instead a real-space modification of the Patterson function. It avoids the difficulty in the third method of choosing a 'B factor' for the normalization of structure factors and has the advantage that the values for the deconvoluted intensities are improved by iterations of the procedure.

2. Method

Sayre (1952) developed the so-called squaring method, which was found to be useful as a method of phase determination. An analogous argument may be used for the deconvolution of overlapping reflections (Laue multiples).

Consider a Patterson function $P(r)$ with non-overlapping and non-negative interatomic vectors: the square of this function looks similar to the function itself except in the relative heights of the peaks. However, this is not strictly valid because interatomic vectors may accidentally overlap. An extreme example is the origin peak which is substantially higher than the rest of the densities due to exact overlapping and needs special attention.

The Patterson function calculated from the Laue single reflections should have the basic features of the Patterson function from the complete data. The Fourier transform of a modified (in the present case, squared) Patterson function can give estimated values for the coefficients which are not included in the $P(r)$ calculation (as well as those which are). The flow chart of our *DECONV* procedure, to evaluate reflection intensities for the components of multiples, is shown in Fig. 1.

The Patterson map is calculated in FFT(1) and with single reflections only in the first run. Let the origin peak height on the Patterson map be H_1 and the next highest peak height be H_2 . The map is modified by removing negative densities and squaring. The origin peak height of the

modified map is truncated at $H_1 \times H_2$ to ensure that the ratio between the origin peak height and the next highest peak height remains unchanged on the modified Patterson map. (Different values for truncating the origin peak have been tested; $H_1 \times H_2$ has been found to give good results.) All $I(hkl)$ are then re-estimated *via* the inverse Fourier transform of the modified Patterson map [FFT(2)]. The program *DECONV* assigns new values for the components, $I(hkl)$, of each multiple spot such that (a) their ratio is that found from the Fourier transformation of the modified Patterson map, and (b) after multiplication by the wavelength normalization function their sum gives the observed intensity of the multiple spot. The whole procedure can be repeated (usually 3–4 cycles) until convergence is reached. The single reflections are kept unchanged in the process. Comparison of the intensities of singles output by FFT(2) can be made with their observed values. (Of course, if no modification is made to the Patterson map, perfect agreement is expected).

Estimated standard deviations (e.s.d.'s) for the intensities of deconvoluted multiple spots have been derived from $\sigma(I_{\text{Laue}})$, evaluated in the program *INTLAUE* (Helliwell, Habash *et al.*, 1989). They take no account of the approximation made in the assumption that the Patterson function is non-overlapping, although error from this source should be uniformly distributed over all the component l values in a multiple; nor do they take account of any errors in the wavelength normalization function.

The software is incorporated in the *Daresbury Laue Software Suite* (available from JWC).

3. Tests of the procedure

The procedure was first tested with Laue data for an organic crystal, $\text{C}_{25}\text{H}_{20}\text{N}_2\text{O}_2$ (Helliwell, Gomez de Anderez, Habash, Helliwell & Vernon, 1989), to provide a comparison with previous published methods. The $|F|$ values of the deconvoluted multiples were compared with the structure factors calculated from the published coordinates. The R factors for all 275 deconvoluted reflections and 149 of those between ∞ and $2d_{\text{min}}$ were 0.284 and 0.126, respectively. These results were very similar to those obtained from direct methods (Hao *et al.*, 1993) [$R_{\text{scale}}(\text{all data}) = 0.288$, $R_{\text{scale}}(\infty-2d_{\text{min}}) = 0.117$] and slightly better than *UNSCRAM* (Helliwell, Habash *et al.*, 1989) [$R_{\text{scale}}(\text{all data}) = 0.259$, $R_{\text{scale}}(\infty-2d_{\text{min}}) = 0.216$] in giving the low-resolution $|F|$ values (see Hao *et al.*, 1993, Tables 1 and 2).

The procedure was then tested with Laue data sets for 4Zn insulin. The Laue diffraction patterns of insulin were recorded on film at station 9.7 of the Daresbury Synchrotron Radiation Source. The insulin crystal is trigonal, with $a = b = 80.7$, $c = 37.6$ Å, space group $R3$, $Z = 9$. Two packs of six films, recorded at different angular settings of the crystal, were scanned and the data processed according to the now well established procedure (Helliwell, Habash *et al.*, 1989). The soft limits were estimated from the gnomonic projection (Cruickshank, Carr & Harding, 1992)

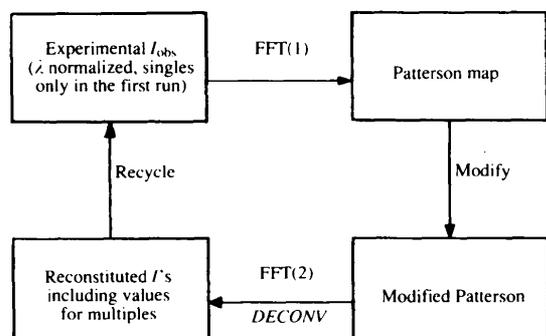


Figure 1
Flow chart of the *DECONV* procedure.

Table 1

Analysis of deconvoluted multiple reflections for Laue diffraction data from 4Zn insulin as a function of resolution.

| Resolution (Å) | Film pack 1 | | Film pack 2 | |
|----------------|-------------|---------------|-------------|---------------|
| | No. | R_{scale}^* | No. | R_{scale}^* |
| 8.84–7.74 | 17 | 0.189 | 27 | 0.297 |
| 7.74–4.56 | 40 | 0.201 | 37 | 0.283 |
| 4.56–3.90 | 11 | 0.284 | 11 | 0.332 |
| 3.90–3.47 | 10 | 0.202 | 14 | 0.229 |
| 3.47–3.15 | 4 | 0.102 | 9 | 0.432 |
| 3.15–2.91 | 14 | 0.222 | 13 | 0.371 |
| 2.91–2.71 | 13 | 0.404 | 17 | 0.379 |
| 2.71–2.55 | 16 | 0.432 | 14 | 0.634 |
| 2.55–2.42 | 10 | 0.380 | 15 | 0.467 |
| 2.42–2.30 | 17 | 0.529 | 19 | 0.322 |
| All | 148 | 0.236 | 165 | 0.309 |

* R_{scale} is calculated against monochromatic data.

to be $\lambda_{min} = 0.30$, $\lambda_{max} = 2.0$ Å, and $d_{min} = 2.3$ Å. Data were normalized against wavelength using the program *LAUENORM*. Intensity measurements for 2888 singles in the wavelength range 0.495–1.60 Å yielded 2053 unique reflections with

$$R_{merge} = \sum_i |I_i - I_m| / \sum_i I_m = 0.093,$$

where I_i represents the wavelength-normalized intensity and I_m is the mean of two or more measurements of the same or symmetry-equivalent reflections. The overall completeness of these data to 2.3 Å resolution was 49% (5% between ∞ and $2d_{min}$; 55% between $2d_{min}$ and d_{min}). These reflections were compared with high-quality monochromatic data (Bentley, Dodson, Dodson, Hodgkin & Mercola, 1976) with the mean fractional difference on F

$$R_{scale} = \sum |F_{Laue} - F_{mono}| / \sum |F_{mono}| = 0.134.$$

After three cycles of *DECONV*, 357 multiples (294 from doubles and 63 from triples) were evaluated and compared with the monochromatic data. R_{scale} was 0.236 for 148 Laue reflections from film pack 1 which were also present in the monochromatic data; R_{scale} was 0.309 for 165 reflections from film pack 2 and 0.274 for 313 reflections from the combination of the two packs, respectively. Details are given in Table 1. There is a wavelength and resolution dependence of the accuracy of the measurements because of the weakness of the spot intensities at both ends of the wavelength range used and at higher resolution. These 357 deconvoluted multiples reduced to 304 unique reflections with $R_{merge}(I) = 0.150$ (data redundancy 1.17). Comparison of 2053 singles output by FFT(2) in the final cycle with their observed values gave $R_{scale}(F) = 0.17$.

The *DECONV* procedure was also tested with the Laue data from cytochrome *c* peroxidase (CCP) (Fülöp, Phizackerley, Soltis, Clifton, Wakatsuki, Erman, Hajdu & Edwards, 1994). The crystal is orthorhombic, with $a = 107.4$, $b = 76.8$, $c = 51.4$ Å, space group $P2_12_12_1$, $Z = 4$. The Laue diffraction patterns recorded on four film packs had been processed with the *Laue Evaluation and Analysis Program*

Table 2

Analysis of deconvoluted multiple reflections for Laue diffraction data from cytochrome *c* peroxidase (CCP) as a function of resolution.

| Resolution (Å) | No. | R_{scale}^* |
|----------------|------|---------------|
| 11.18–5.98 | 177 | 0.303 |
| 5.98–4.56 | 456 | 0.172 |
| 4.56–3.83 | 191 | 0.195 |
| 3.83–3.37 | 100 | 0.228 |
| 3.37–3.04 | 70 | 0.230 |
| 3.04–2.80 | 67 | 0.234 |
| 2.80–2.60 | 48 | 0.337 |
| 2.60–2.44 | 42 | 0.518 |
| 2.44–2.31 | 40 | 0.483 |
| 2.31–2.20 | 31 | 0.328 |
| All | 1134 | 0.207 |

* R_{scale} is calculated against monochromatic data.

(*LEAP*) (Fülöp *et al.*, 1994), with values $\lambda_{min} = 0.25$, $\lambda_{max} = 2.06$ Å, $d_{min} = 2.2$ Å. The overall completeness of 9428 unique single reflections processed from four film packs to 2.2 Å resolution was 53% (5% between ∞ and $2d_{min}$; 59% between $2d_{min}$ and d_{min}). $R_{scale}(F)$ between Laue singles and monochromatic data was 0.130.

Using three cycles of *DECONV*, 1333 multiples (1036 from doubles and 297 from triples) from one of the film packs in the wavelength range 0.25–2.06 Å were evaluated. R_{scale} was 0.207 for 1134 Laue reflections which were in common with the monochromatic data; details are given in Table 2. There is a similar wavelength and resolution dependence of the accuracy of the measurements to that in Table 1. The 1333 deconvoluted multiples yielded 1171 unique reflections with $R_{merge}(I) = 0.162$ (data redundancy 1.14). Comparison of 9428 singles output by FFT(2) in the final cycle with their observed values gave $R_{scale}(F) = 0.16$.

4. Concluding remarks

Evaluation of the components of the multiples has been achieved although this is not as accurate as the measurement of singles; however, there are several situations where it is far better to have approximate values for these reflections than none at all. The values found are much better than random values, for which $R \geq 0.6$ would be expected.

The present method should work best when the reciprocal lattice coverage for the singles is close to 100%. Two test data sets (insulin and CCP) were chosen in which there is a lack of data redundancy as might occur in a time-resolved experiment. Their deconvolution has been most successful with the lower order reflections (see Tables 1 and 2) and these are the ones which are important for phase determination or connectivity of electron-density maps. It therefore complements two of the other methods, *UNSCRAM* (Helliwell, Habash *et al.*, 1989) because it does not require film packs, and the wavelength normalization method (Campbell & Hao, 1993), which requires high data

redundancy. Compared with the reciprocal-space method (Hao *et al.*, 1993), the new procedure is more convenient to run and uses much less CPU time; this is because density modification in real space and Fourier transformation is much faster than setting up the Σ_2 relationships in reciprocal space required in the method of Hao *et al.* (1993), especially for macromolecule cases.

Finally, it is of interest to consider trying this new method to estimate the intensity values of reflections which are missing altogether from the Laue patterns. This is feasible, in principle, from the FFT of the modified Patterson map, but extensive trials are required to test this idea. Similarly, for the case where phases are available, modified electron density maps also offer possibilities.

The Science and Engineering Research Council (SERC) and the University of Liverpool are thanked for support. We are grateful to Dr S. Wakatsuki of Oxford University for providing CCP data and Mrs E. J. Dodson of York University for providing insulin data. We also thank Professors J. R. Helliwell (Manchester University), D. W. J. Cruickshank (UMIST), M. M. Woolfson (York University) and H. Fan (Chinese Academy of Sciences) for discussions.

References

- Bentley, G., Dodson, E. J., Dodson, G. G., Hodgkin, D. C. & Mercola, D. A. (1976). *Nature (London)*, **261**, 166–168.
- Campbell, J. W., Deacon, A., Habash, J., Helliwell, J. R., McSweeney, S., Hao, Q., Raftery, J. & Snell, E. (1994). *Bull. Mater. Sci.* **17**, 1–18.
- Campbell, J. W. & Hao, Q. (1993). *Acta Cryst.* **A49**, 889–893.
- Cruickshank, D. W. J., Carr, P. D. & Harding, M. M. (1992). *J. Appl. Cryst.* **25**, 285–293.
- Cruickshank, D. W. J., Helliwell, J. R. & Moffat, K. (1987). *Acta Cryst.* **A43**, 656–674.
- Duke, E. M. H., Hadfield, A., Walters, S., Wakatsuki, S., Bryan, R. K. & Johnson, L. N. (1992). *Philos. Trans. R. Soc. London Ser. A*, **340**, 245–261.
- Fülöp, V., Phizackerley, R. P., Soltis, S. M., Clifton, I. J., Wakatsuki, S., Erman, J., Hajdu, J. & Edwards, S. L. (1994). *Structure*, **2**(3), 201–208.
- Hao, Q., Campbell, J. W., Harding, M. M. & Helliwell, J. R. (1993). *Acta Cryst.* **A49**, 528–531.
- Helliwell, M., Gomez de Anderez, D., Habash, J., Helliwell, J. R. & Vernon, J. (1989). *Acta Cryst.* **B45**, 591–596.
- Helliwell, J. R., Habash, J., Cruickshank, D. W. J., Harding, M. M., Greenhough, T. J., Campbell, J. W., Clifton, I. J., Elder, M., Machin, P. A., Papiz, M. Z. & Zurek, S. (1989). *J. Appl. Cryst.* **22**, 483–497.
- Johnson, L. N. & Hajdu, J. (1990). *Synchrotron Radiation and Biophysics*, edited by S. S. Hasnain, ch. 6. Chichester: Ellis Horwood.
- Lindahl, M., Liljas, A., Habash, J., Harrop, S. & Helliwell, J. R. (1992). *Acta Cryst.* **B48**, 281–285.
- Sayre, D. (1952). *Acta Cryst.* **5**, 60–65.
- Schlichting, I., Almo, S. C., Rapp, G., Wilson, K., Petratos, K., Lentfer, A., Wittinghofer, A., Kabsch, W., Pai, E. E., Petsko, G. A. & Goody, R. S. (1990). *Nature (London)*, **345**, 309–315.
- Singer, P. T., Smalås, A., Carty, R. P., Mangel, W. F. & Sweet, R. M. (1993). *Science*, **259**, 669–673.
- Smith Temple, B. & Moffat, K. (1987). *Computational Aspects of Protein Crystal Data Analysis*, edited by J. R. Helliwell, P. A. Machin & M. Z. Papiz, pp. 84–89. SERC Daresbury Laboratory, Warrington, England.
- Szebenyi, D. M. E., Bilderback, D. H., LeGrand, A., Moffat, K., Schildkamp, W., Smith Temple, B. & Teng, T.-Y. (1992). *J. Appl. Cryst.* **25**, 414–423.