Acta Cryst. (1990). A46, 721-725

Isomorphous Replacement with Optimized Anomalous Scattering Applied to Protein Crystallography

BY P. J. BAKER, G. W. FARRANTS,* T. J. STILLMAN AND K. L. BRITTON

Krebs Institute for Biomolecular Research, Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield S10 2TN, England

J. R. HELLIWELL†

SERC Daresbury Laboratory, Warrington, Cheshire WA4 4AD, England

AND D. W. RICE

Krebs Institute for Biomolecular Research, Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield S10 2TN, England

(*Received* 5 *January* 1990; *accepted* 23 *April* 1990)

Abstract

Isomorphous replacement is an essential technique for the *de novo* solution of macromolecular crystal structures. The use of the anomalous-dispersion effect of the heavy atom in the derivative leads to the acronyms SIRAS or MIRAS (single or multiple isomorphous replacement with anomalous scattering) as the term for the phase determination method. Synchrotron radiation is tuneable over a wavelength range encompassing the absorption edges of the typically used derivative atoms such as Hg, Au and Pt. Hence, it is possible to optimize the anomalous-scattering signal of such atoms by appropriate choice of a wavelength between ~0·8 and 1·1 Å. This paper reports a comparison of this method (which we call SIROAS; O for optimized) and SIRAS on related mercury derivatives of glutamate dehydrogenase. The anomalous-scattering signal is enhanced in the SIROAS experiment and this results in an increase of the phasing power of the derivative data.

Introduction

The determination of phases for experimentally measured structure factors is the critical step in the solution of a crystal structure by X-ray diffraction. In protein crystallography, the initial estimates of phases in the majority of structures that have been elucidated to date have been derived from the technique of isomorphous replacement. Anomalous-scattering signals have been measured and used in conjunction with isomorphous replacement methods and the terms SIRAS (single isomorphous replacement with anomalous scattering) and MIRAS (multiple isomorphous replacement with anomalous scattering) have become commonplace. If the technique of SIRAS is to be used in the determination of a protein structure, then it is essential that the anomalous component of the scattering is measured accurately. This gives rise to an immediate problem. Since the magnitude of the signal is calculated from the difference between Friedel pairs of reflections, and is therefore inherently small, it easily becomes swamped by systematic and/or random errors in the data set. The size of the f'' signal is wavelength dependent and can change dramatically as the wavelength of the incident radiation nears an absorption edge of the anomalous-scattering ion. The wavelength dependence of f' and f'' for an Hg ion is shown in Fig. 1. It can be seen that the f'' signal increases from about 4 electrons at $\lambda = 1{\cdot}1$ Å to over 11 electrons at $\lambda = 0{\cdot}83$ Å, close to the L_I absorption edge. The advent of synchrotron radiation has meant

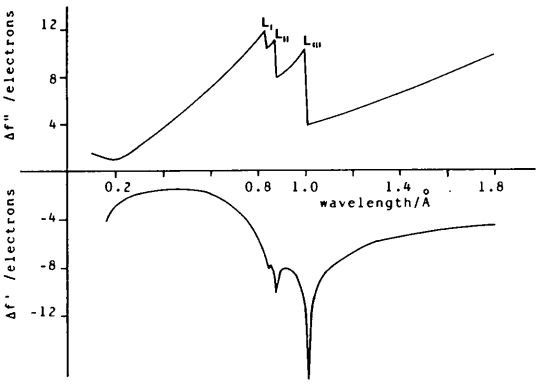

Fig. 1. A plot of the real (f') and imaginary (f'') terms of the anomalous scattering for the L edge of an Hg atom.

* Present address: Norwegian Radium Hospital, Montebello, N-0310, Oslo 3, Norway.

† Present address: Department of Chemistry, University of Manchester, Manchester M13 9PL, and SERC Daresbury Laboratory, Warrington, Cheshire WA4 4AD, England.

0108-7673/90/090721-05$03.00

that the incident-beam wavelength can be easily adjusted so as to maximize the anomalous-scattering signal (Helliwell, 1984). Furthermore, if these short wavelengths are used for the data collection, there is a significant decrease in the linear atomic absorption coefficient, with a consequent decrease in the systematic errors in the data (Greenhough & Helliwell, 1983; Stuart & Walker, 1987).

In the initial determination of the phases for the 3·7 Å structure of the enzyme glutamate dehydrogenase (GDH), a single isomorphous derivative was employed. The mercury derivative data were collected using the tuneable wavelength facility of station 9.6 on the SERC Synchrotron Radiation Source at Daresbury (SRS) to enhance the anomalous-scattering signal (Helliwell *et al.*, 1986). The structure was thus based on a single isomorphous derivative with optimized anomalous scattering, for which we coin the acronym SIROAS. In this paper we compare this derivative of GDH to a related derivative where the data were collected at 1·488 Å where the anomalous-scattering signal was not optimized and the protein crystal absorption correction is large.

Materials and methods

Crystals of GDH in space group $C2$ and with cell dimensions $a = 147{\cdot}1$, $b = 151{\cdot}3$, $c = 94{\cdot}6$ Å and $\beta = 132{\cdot}75°$ were grown from ammonium sulfate solutions (Rice, Hornby & Engel, 1985). The crystals selected for the data collection had a rectangular plate morphology of typical dimensions $1{\cdot}0 \times 0{\cdot}7 \times 0{\cdot}2$ mm. In this form the unique cell axis, b, lies across the large face of the crystal. One isomorphous derivative was prepared by equilibrating the crystals in a mother liquor containing $0{\cdot}001M$ ethylmercury phosphate (EMP), and another by equilibrating the crystals in $0{\cdot}005M$ 2,3-bis(bromomercuriomethyl)tetrahydrofuran (DBMMF). For the EMP soaked crystals all data were collected on station PX9.6 on the SRS (Helliwell *et al.*, 1986), using the rotation-camera method (Arndt, Champness, Phizackerly & Wonacott, 1973; Arndt & Wonacott, 1977) with the Wiggler magnet operating at 5·0 T. Reference to Fig. 1 shows that there are three absorption edges for the Hg ion, $L_{\rm I}$, $L_{\rm II}$ and $L_{\rm III}$, between 0·8 and 1·0 Å wavelengths, which are all accessible on station PX9.6. As the wavelength is changed in the vicinity of the absorption edge in question, the fluorescence emitted by the metal atom in the crystal can be monitored. Hence, it is possible to set the wavelength of the beam using this effect, to maximize completely the anomalous-scattering signal from the metal atom in the crystal (Helliwell, 1984). However, in this experiment the incident wavelength was set at a value ~0·01 Å on the short-wavelength side of the absorption edge. The advantage of this is that any near-edge absorption effects, particularly dichroism due to beam polarization, can be avoided.

Although the $L_{\rm I}$ edge has the largest anomalous-scattering signal associated with it, it is important to consider the effect of the incident wavelength on the data-collection protocol, particularly the correlated component of the wavelength dispersion, $(\delta\lambda/\lambda)_{\rm corr}$, which is defined in (1), for the focusing oblique cut triangular monochromator.

$$(\delta\lambda/\lambda)_{\rm corr} = L/2\{[\sin(\theta-\alpha)/P_1] - [\sin(\theta+\alpha)/P_2]\}\cot\theta, \qquad (1)$$

where L is the length of the monochromator crystal, θ the Bragg angle for diffraction from the monochromator, α the cut-off angle of the monochromator, P_1 the monochromator-to-crystal-sample distance and P_2 the source-to-monochromator distance. If $(\delta\lambda/\lambda)_{\rm corr}$ is $>10^{-3}$ it is necessary to take account of this in the prediction of reflection partiality or rocking width (Greenhough & Helliwell, 1982). For routine data collection and processing it is more straightforward to set the instrument to give a $(\delta\lambda/\lambda)_{\rm corr} = 0$, *i.e.* the so-called Guinier focusing distance, $P'_G = [\sin(\theta-\alpha)/\sin(\theta+\alpha)]P$. However, when λ is specifically selected, as in this experiment, and because of constraints on the optical bench, it is necessary to use a focusing distance away from P'_G. In this situation $(\delta\lambda/\lambda)_{\rm corr}$ is restricted to no more than 10^{-3} by reducing the length of the monochromator that is illuminated.

At station PX9.6 on the SRS, using an oblique-cut $Si_{[111]}$ monochromator, $\alpha = 6{\cdot}75°$, $P_1 = 2000$ mm, $P_2 = 20\,000$ mm, $L = 200$ mm and $\theta = 7{\cdot}6546°$ (Hg $L_{\rm I}$) and $8{\cdot}0475°$ (Hg $L_{\rm II}$) (Helliwell, 1984). Substituting these values in (1), $(\delta\lambda/\lambda)_{\rm corr} = -1{\cdot}024 \times 10^{-3}$ for the $L_{\rm II}$ edge and $(\delta\lambda/\lambda)_{\rm corr} = -3{\cdot}38 \times 10^{-3}$ for the $L_{\rm I}$ edge. It would thus be necessary to reduce the length of the monochromator and hence flux intensity severely if the $L_{\rm I}$ edge were to be used for the 6·75° cut monochromator and thus the $L_{\rm II}$ edge was chosen for the GDH EMP data collection. The monochromator take-off angle therefore was adjusted to give a wavelength of 0·86 Å and the horizontal preslits were used to reduce the effective monochromator length by one half. This had the effect of making $(\delta\lambda/\lambda)_{\rm corr} = -5{\cdot}12 \times 10^{-4}$, acceptably close to the Guinier setting. For data collection, crystals were mounted with the reciprocal axis, b^*, parallel to the rotation axis, such that the Bijvoet pairs of reflections hkl and $h\bar{k}l$ were recorded on the same film, with equivalent absorption factors. Three films were loaded into each cassette, with 0·15 mm thick aluminium foils interspersed between each film. A rotation angle of 1·8° was used for each exposure with a 0·15 mm radius collimator. Data were collected to 2·5 Å from two crystals, one of which was exposed in two halves.

Table 1. *Data collection and processing statistics*

	Native	EMP	DBMMF
Number of crystals*	1	3	2
Total number of measurements	15 898	174 050	140 616
Resolution of data collection (Å)	3·7	2·5	2·5
Independent reflections to 3·7 Å	14 958	15 094	12 815
Merging R values (R_m) to 3·7 Å	0·079	0·051	0·07
Mean fractional isomorphous difference (R_{diff})	—	0·174	0·247
R_{FHLE}	—	0·471	0·509

$R_m = \sum_{hkl} \sum_{i=1}^{n} \langle I_i - \bar{I} \rangle / \sum_{hkl} n\bar{I}$.

$R_{\text{diff}} = \sum \| F_{PH} | - | F_P \| / \sum | F_P |$.

$R_{FHLE} = \sum \| F_{HLE} | - | F_{H_{\text{calc}}} \| / \sum | F_{HLE} |$, F_{HLE} is the lower estimate of the heavy-atom structure factor F_H.

* A crystal exposed in two separate halves counts as two crystals.

For the crystals soaked in DBMMF, all the data were collected on station PX7.2 on the SRS, again using the rotation-camera method, with the crystals mounted about b^*. This station (Helliwell *et al.*, 1982) was used at a fixed wavelength of 1·488 Å; at this wavelength the anomalous-scattering signal from the Hg ion is 7·28 e^-, a decrease of 33% compared with the signal obtained at 0·86 Å, where f''(Hg) is 10·8 e^- (Sasaki, 1984). Owing to the increased film absorption factor at 1·488 Å, no aluminium foils were used to intersperse the films within the camera cassettes. The data were collected to 2·5 Å resolution from a single crystal, using a 0·15 mm radius collimator, exposed in two separate halves, and a rotation angle of 1·8° was used for each exposure.

The native data were collected from a single crystal mounted about b^*, using Ni-filtered Cu $K\alpha$ radiation, on the modified Hilger and Watts multicounter 5-circle X-ray diffractometer (Phillips, 1964; Banner, Evans, Marsh & Phillips, 1977) in the Laboratory of Molecular Biophysics at the University of Oxford. A single equivalent of data was collected in discrete overlapping resolution shells in order to optimize the count time for the resolution-dependent intensity distribution of the reflections. The data were corrected for absorption using the empirical method of North, Phillips & Mathews (1968).

For scaling purposes both derivatives were treated separately, but in an equivalent manner. Scale factors between the film packs were calculated using the method of Fox & Holmes (1966), each film pack representing a separate batch. Equivalent reflections were then merged; data collection statistics are shown in Table 1. The data from both derivatives were scaled to the native data to a resolution of 3·7 Å and an isomorphous difference Patterson synthesis calculated for the EMP derivative. The positions of six mercury binding sites were determined from this initial Patterson synthesis and the subsequent heavy-atom refinement, whereas for the DBMMF derivative the same six sites were located from a difference Fourier synthesis using coefficients $|F_{H_{\text{obs}}} - F_{H_{\text{calc}}}| m \exp i\alpha_P$, where α_P was the phase derived from the EMP derivative. In addition, three minor sites of substitution were also found. The relative occupancies of these sites and their coordinates were refined for each derivative using the F_{HLE} method of Dodson & Vijayan (1971; Kartha, 1965). Two sets of phases were calculated independently for each derivative using the method of Blow & Crick (1959; Dickerson, Kendrew & Strandberg, 1961), incorporating contributions from anomalous scattering (North, 1965; Matthews, 1966*a*).

Results and discussion

In order to compare the optimized anomalous-scattering signal from the EMP derivative with the non-optimized signal from the DBMMF derivative it is important to ascertain that the two derivatives are of a similar overall quality.

The asymmetric unit of the cell contains three GDH subunits, related by a non-crystallographic molecular threefold axis (Rice, Baker, Farrants & Hornby, 1987). Therefore, the six separate mercury binding sites identified in the Patterson synthesis for the EMP derivative can be reduced to two mercury binding sites per GDH subunit, one of which is fully occupied and the other has only partial occupancy. The mercury peak shape in the DBMMF derivative is distinctly elongated in one direction and the size and shape of this extra feature is fully consistent with its assignment to a Br atom. On the high-occupancy sites where these features were significant, they were modelled in the refinement as two atoms. The mean fractional isomorphous difference for the DBMMF derivative (0·247) is greater than that for the EMP derivative (0·174), as expected on account of the additional scattering arising from the Br atoms in the DBMMF derivative and also from the three minor sites. Clearly the two derivatives are not identical but exhibit largely similar patterns of substitution. With regard to the extent to which they exhibit identical isomorphism, in Fig. 2(*a*) we show a graph of the mean fractional isomorphous difference, $\sum |F_P| - |F_{PH}| / \sum |F_P|$, (del F/F) as a function of resolution for both derivatives. The values of del F/F for the DBMMF derivative are consistently higher than the EMP derivative, which again is to be expected due to the extra scattering matter present. However, this increase is approximately constant over the entire resolution range, showing that the relative degree of isomorphism is equivalent for both derivatives. A further estimate of the relative quality of the derivatives can be obtained from a comparison of the $F_{H_{\text{calc}}}$ to the relative lack of closure of the phase triangle (E) ratio for the centric terms. This is shown in Fig. 2(*b*) where again the close similarity of the two derivatives can be seen.

It is important to consider the effects of absorption of the incident and diffracted rays by the crystal and its habit for the two wavelengths, as the systematic errors introduced could easily swamp the anomalous-scattering signal being measured. Considering the contribution to the absorption from the crystal alone, Stuart & Walker (1987) have calculated the linear absorption coefficient, μ, as a function of incident-beam wavelength for a 'typical' protein crystal, which is composed of 50% protein, 50% solvent, has no atoms heavier than oxygen and has a density of 1·15 g cm^{-3}. At 1·488 Å, $\mu \simeq 8{\cdot}7$ cm^{-1} and at 0·86 Å $\mu \simeq 1{\cdot}8$ cm^{-1}. Now, as $I = I_0\, e^{-\mu t}$, where I_0 is the incident intensity and I is the emergent intensity of an X-ray passing through a crystal of thickness t, the path length of the diffracted beam through the crystal affects the absorption. The two limiting cases for absorption effects for the GDH crystals are: (a) with the plate perpendicular and (b) with the plate parallel to the X-rays. Hence, for a crystal at a wavelength of 1·488 Å, $I = 0{\cdot}80 I_0$ for case (a) and $I = 0{\cdot}41 I_0$ for case (b), a change in absorption by a ratio of two as the crystal is rotated around its mounting axis; whereas, at a wavelength of 0·86 Å, $I = 0{\cdot}96 I_0$ for case (a) and $I = 0{\cdot}83 I_0$ for case (b), a change in absorption by a ratio of 1·15. Thus at the shorter wavelength the absorption effect is reduced considerably. A further source of systematic errors can arise from absorption of the X-rays by the crystal surroundings such as mother liquor and capillary thickness and uniformity. Again, at the shorter wavelength these effects will be advantageously reduced.

A common method used to estimate the quality of the anomalous-scattering signal is to compare the empirical anomalous ratio (K_{emp}), as defined by Matthews (1966b) in (2), with the theoretical ratio (K_{theo}), as defined in (3).

$$K_{emp} = 2\{\langle (F_{PH} - F_P)^2\rangle / \langle [F_{PH}(+) - F_{PH}(-)]^2\rangle\}^{1/2}, \quad (2)$$

$$K_{theo} = (f_0 + f')/f''. \quad (3)$$

Fig. 3 portrays K_{emp} and K_{theo} for both the EMP and DBMMF derivatives (for DBMMF the effect of the Br atom on the anomalous signal is included). It can be seen that the K_{emp} values for the EMP data collected at 0·86 Å wavelength agree very closely with K_{theo}, whereas, for the DBMMF data, collected at 1·488 Å, the K_{emp} ratio only agrees poorly at very low resolution, and rapidly becomes worse as the

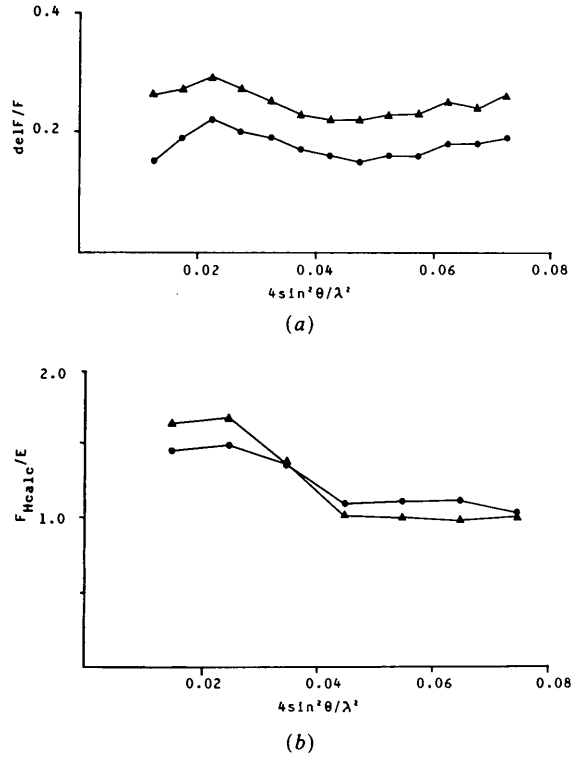

Fig. 2. Graphs of (a) $\sum ||F_{PH}| - |F_P|| / \sum |F_P|$ (del F/F) against $4(\sin^2\theta)/\lambda^2$ and (b) $F_{H_{calc}}/E$ against $4(\sin^2\theta)/\lambda^2$, for the two GDH derivatives. Data for the EMP derivative are shown as circles and for the DBMMF derivative as triangles. E is the r.m.s. lack of closure error in the isomorphous case.

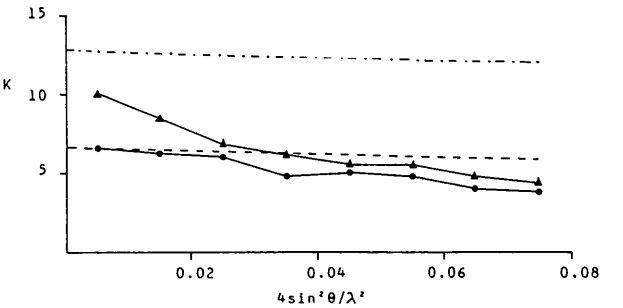

Fig. 3. A comparison of the theoretical and empirical values of K, the ratio of the size of the isomorphous and anomalous-scattering signals, as a function of resolution. The curves for the EMP derivative are shown for K_{emp} (●) and K_{theo} (- - -) and the curves for the DBMMF derivative are represented as K_{emp} (▲) and K_{theo} (— · — · —).

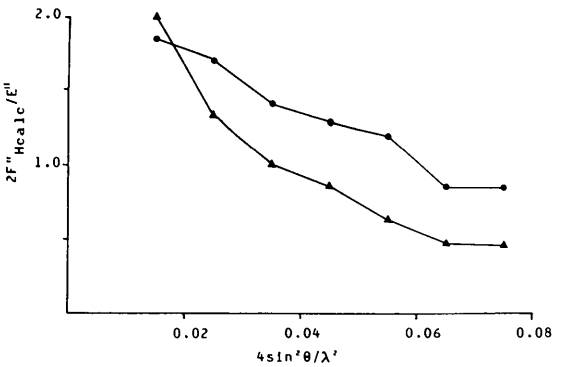

Fig. 4. A graph of $2F''_{calc}/E$ against $4(\sin^2\theta)/\lambda^2$ for the EMP derivative (circles) and the DBMMF derivative (triangles), E'' is the r.m.s. lack of closure in the anomalous case.

resolution increases. Therefore, the anomalous-scattering signal is clearly better for the short-wavelength data rather than the longer-wavelength data. Furthermore, the ratio of $F''_{H_{calc}}/E''$, where $F_{H_{calc}}$ is the anomalous contribution of the anomalous-scattering atoms to the structure factor and E'' is the r.m.s. lack of closure for the anomalous case, as a function of resolution (Fig. 4) is greater, on average by a factor of two, for the EMP derivative and this pattern is extended to the F_{HLE} R factor (47·1% EMP; 51·0% DBMMF) and the overall figure of merit for the resultant phase set (0·58 EMP; 0·51 DBMMF).

These results indicate that, for these two very similar derivatives, the anomalous-scattering signal to noise has been enhanced for the EMP data set, for which the wavelength was chosen to illuminate the L_{II} absorption edge of the Hg atom in conjunction with a reduction in the overall protein crystal absorption at 0·86 Å. On PX9.6 it is easy to move between the absorption edges of commonly used heavy ions, such as mercury, platinum and gold, so that MIROAS is routinely useable. The results from this present study with SIROAS illustrate the ease and effectiveness of the procedure. The two benefits of reduced absorption and increased anomalous-scattering signal have increased the phasing power of the EMP derivative of GDH and hence improved the quality of the resultant GDH electron density maps.

We thank the SERC for their financial support of this work and for use of the SRS. We also thank J. M. A. Smith for helping to set up station PX9.6 and for subsequent consultations and also all the support staff at Daresbury. DWR is a Lister Institute Research Fellow.

References

ARNDT, U. W., CHAMPNESS, J. N., PHIZACKERLY, R. P. & WONACOTT, A. J. (1973). *J. Appl. Cryst.* **6**, 457-463.
ARNDT, U. W. & WONACOTT, A. J. (1977). *The Rotation Method in Crystallography.* Amsterdam: North-Holland.
BANNER, D. W., EVANS, P. R., MARSH, D. J. & PHILLIPS, D. C. (1977). *J. Appl. Cryst.* **10**, 45-51.
BLOW, D. M. & CRICK, F. H. C. (1959). *Acta Cryst.* **12**, 794-802.
DICKERSON, R. E., KENDREW, J. C. & STRANDBERG, B. E. (1961). *Acta Cryst.* **14**, 1188-1195.
DODSON, E. & VIJAYAN, M. (1971). *Acta Cryst.* B**27**, 2402-2411.
FOX, G. C. & HOLMES, K. C. (1966). *Acta Cryst.* **20**, 886-891.
GREENHOUGH, T. J. & HELLIWELL, J. R. (1982). *J. Appl. Cryst.* **15**, 493-508.
GREENHOUGH, T. J. & HELLIWELL, J. R. (1983). *Prog. Biophys. Mol. Biol.* **41**, 89.
HELLIWELL, J. R. (1984). *Rep. Prog. Phys.* **47**, 1403-1497.
HELLIWELL, J. R., GREENHOUGH, T. J., CARR, P. D., RULE, S. A., MOORE, P. R., THOMPSON, A. W. & WORGAN, J. S. (1982). *J. Phys. E*, **15**, 1363-1372.
HELLIWELL, J. R., PAPIZ, M. Z., GLOVER, I. D., HABASH, J., THOMPSON, A. W., MOORE, P. R., HARRIS, N., CROFT, D. & PANTOS, E. (1986). *Nucl. Instrum. Methods*, A**246**, 617-623.
KARTHA, G. (1965). *Acta Cryst.* **19**, 883-885.
MATTHEWS, B. W. (1966*a*). *Acta Cryst.* **20**, 82-86.
MATTHEWS, B. W. (1966*b*). *Acta Cryst.* **20**, 230-239.
NORTH, A. C. T. (1965). *Acta Cryst.* **18**, 212-216.
NORTH, A. C. T., PHILLIPS, D. C. & MATHEWS, F. S. (1968). *Acta Cryst.* A**24**, 351-359.
PHILLIPS, D. C. (1964). *J. Sci. Instrum.* **41**, 123-129.
RICE, D. W., BAKER, P. J., FARRANTS, G. W. & HORNBY, D. P. (1987). *Biochem. J.* **242**, 789-795.
RICE, D. W., HORNBY, D. P. & ENGEL, P. C. (1985). *J. Mol. Biol.* **181**, 147-149.
SASAKI, S. (1984). Photon Factory KEK preprint No. 83-22, pp. 1-136. Photon Factory, Ibaraki, Japan.
STUART, D. & WALKER, N. (1987). *Computational Aspects of Protein Crystal Data Analysis.* Proc. of Daresbury Study Weekend DL/SCI/R25, edited by J. R. HELLIWELL, P. A. MACHIN & M. Z. PAPIZ, pp. 25-38. Warrington: SERC Daresbury Laboratory.

Acta Cryst. (1990). A**46**, 725-730

Space-Group Frequencies of Proteins and of Organic Compounds with More Than One Formula Unit in the Asymmetric Unit

BY N. PADMAJA, S. RAMAKUMAR* AND M. A. VISWAMITRA

Department of Physics, Indian Institute of Science, Bangalore-560 012, India

(*Received* 6 *October* 1988; *accepted* 4 *April* 1990)

Abstract

An analysis of the distribution of organic crystal structures with more than one formula unit in the asymmetric unit among the 230 space groups has been carried out for the compounds listed in the Cambridge Structural Database. 8·3% of the total number of compounds (51 611) in the database have more than one formula unit in the asymmetric unit; 81% of these are reported in only five space groups: $P2_1/c$ (27·8%), $P\bar{1}$ (23·5%), $P2_1$ (13·8%), $P1$ (8.5%) and $P2_12_12_1$ (7·8%). When *all* compounds are considered, the first five most populous space groups are: $P2_1/c$ (36·6%), $P\bar{1}$ (16·9%), $P2_12_12_1$ (11·0%),

* Author for correspondence.

0108-7673/90/090725-06$03.00