Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Efficient anisotropic refinement of macromolecular structures using FFT

Garib N. Murshudov,[a,b]* Alexei A. Vagin,[a] Andrey Lebedev,[c] Keith S. Wilson[a] and Eleanor J. Dodson[a]

[a]Chemistry Department, University of York, Heslington, York YO1 5DD, England, [b]CLRC, Daresbury Laboratory, Daresbury, Warrington WA4 4AD, England, and [c]MPG-ASMB c/o DESY, Gebaude 25 B, Notkestrasse 85, 22603 Hamburg, Germany

Correspondence e-mail: garib@yorvic.york.ac.uk

Received 2 October 1998
Accepted 25 October 1998

This paper gives the equations for the use of fast Fourier transformations in individual atomic anisotropic refinement. Restraints on bonded atoms, on the sphericity of each atom and between non-crystallographic symmetry related atoms are described. These have been implemented in the program *REFMAC* and its performance with several examples is analysed. All the tests show that anisotropic refinement not only reduces the R value and $R_{\rm free}$ but also improves the fit to geometric targets, indicating that this parameterization is valuable for improving models derived from experimental data. The computer time taken is comparable to that for isotropic refinements.

1. Introduction

Refinement of crystal structures can be subdivided into four fundamental steps. These are: (i) choice of residual, (ii) parameterization of the refined model, (iii) minimization method for the residual against the current parameters and (iv) handling and use of prior knowledge. It has previously been shown that using a −log likelihood function for the residual can give better results than classical least-squares (Pannu & Read, 1996; Murshudov *et al.*, 1997). The most appropriate parameterization of the model depends on the data available, the stage of refinement and the amount and treatment of prior information. Again, the choice of the minimization method depends on all the above and on the residual chosen. At early stages of least-squares refinement it has been established that the use of simulated annealing with molecular dynamics considerably improves the refinement behaviour and increases the radius of convergence (Brünger, 1992). At later stages, minimization methods which use first and second derivatives (or approximations to them) increase the rate of convergence (Fletcher, 1981). Preliminary tests show that with a maximum-likelihood residual the use of first and second derivatives even at an early stage is valuable.

In principle, there is information from the minimization procedure which can also help to guide the parameterization. For this, all second derivatives of the residual must be calculated and the eigenvalue and eigenvectors of this matrix derived. From an analysis of the eigenvalues the model can be reparameterized in general terms (Watkin, 1988, 1994; Ten Eyck, 1996; Cowtan & Ten Eyck, 1999).

Handling of prior knowledge is of prime importance, and there are several aspects where improvements could be made. For instance, commonly used restraint residuals either use second-order approximations to the energy functions or assume normal distributions about some restraint mean, which is not always appropriate. Many structural studies aim to describe details of the mechanism of action of the macromolecule, which are probed by the incorporation of new inhibitors or pseudo-substrates. To examine the systems properly, it is essential that the new dictionary entries required to describe the inhibitor geometry take into account the underlying chemistry and are easy to formulate.

The parameterization of the current model is an essential step for all refinement programs. Most popular programs (*X-PLOR*, Brünger, 1992; *TNT*, Tronrud *et al.*, 1987; *PROLSQ*, Konnert & Hendrickson, 1980; *REFMAC*, Murshudov *et al.*, 1997) can refine individual atomic positions as well as overall or individual isotropic B values (three or four parameters per atom). *X-PLOR* can further reduce the number of parameters by refining in torsion-angle space (Brünger & Rice, 1997), but since the torsion angles of a model are highly correlated this may lead to instability. *SHELXL* (Sheldrick, 1995) and *RESTRAIN* (Driessen *et al.*, 1989) can refine individual anisotropic U values (nine parameters per atom); both these programs use classical normal equations formulated in reciprocal space to describe the least-squares residuals (Rollett, 1968).

As more powerful data-collection techniques have become available, there has been a rapid increase in the number of macromolecular crystallographic data sets collected at high or even atomic resolution (Dauter *et al.*, 1997). This puts extra requirements on refinement programs, which should be able both to refine initial structures to reasonable R values and then allow reparameterization of the model to exploit the data quality. At high resolution this requires that the model should be described with individual atomic anisotropic parameters. Finally, the refinement program should be able to assess the reliability of the resultant model (*e.g.* assign standard uncertainties for all the parameters).

A second class of problems now requiring refinement algorithms are those where reasonable data have been collected from crystals of lower quality. Here the molecules often display high mobility which can be described using translation–libration–screw (TLS; notation is given in Table 1) parameters (Schomaker & Trueblood, 1968). This means that programs should be able to refine such overall thermal parameters of the molecule(s). Once derivatives for the individual anisotropic U values have been calculated, those for the TLS parameters can be deduced by application of the chain rule. One potential application of TLS refinement is in applying NCS restraints or constraints, as different copies of the molecule in the asymmetric unit may well have different overall thermal parameters. This may also be important in multicrystal refinement and/or averaging.

If anisotropic refinement could be carried out without using excessive computing resources, it would certainly be used much more routinely. Cruickshank (1956) gave equations for refinement using Fourier methods (Booth, 1946; Cochran, 1948; Cruickshank, 1952, 1956). Here, we give equations for using a real-space Fourier approximation to carry out individual anisotropic refinement for a wide range of residuals and in particular for its application in maximum-likelihood refinement.

2. Sources of anisotropy of atoms

Several factors contribute to apparent atomic anisotropy. The crystal itself (except in a cubic space group) is in general an anisotropic field, so it is to be expected that the data collected from it may exhibit overall anisotropy. Freezing and/or addition of substrates will usually change the anisotropy of the crystal, in general increasing it. This can be corrected by refinement of an overall anisotropic scale factor respective to the crystal axes, in general requiring the addition of five extra parameters. The correction is not necessarily positive definite as it reflects differences between overall thermal parameters of the observed and calculated structure factors. The results of this are often substantial, reducing the R value and $R_{\rm free}$ by several percent, as well as improving the fit to the geometric restraints (Sherif & Hendrickson, 1987; Murshudov *et al.*, 1998).

A second source of anisotropy is the movement of whole molecules as rigid bodies within the crystal lattice. This can be described by TLS parameters (20 more per molecule) independent of the crystal form. *RESTRAIN* (Moss *et al.*, 1996) is able to evaluate these, and the correction has been shown to be valuable in some situations.

A third source of anisotropy is vibration along torsion angles. In principle, this might be described by refining the torsion angles themselves and estimating their displacement parameters. However, there are problems with this since these parameters are highly correlated and such refinement may be sensitive to small perturbations of one or several of the parameters. It may be better to deduce the displacement parameters of the torsion angles from the individual anisotropic atomic $\mathbf{U}$ values.

To summarize, the observed atomic anisotropic thermal parameter can be written as

$$\mathbf{U}_{\rm atom,overall} = \mathbf{U}_{\rm crystal} + \mathbf{U}_{\rm TLS} + \mathbf{U}_{\rm torsion} + \mathbf{U}_{\rm atom}, \qquad (1)$$

where $\mathbf{U}_{\rm atom,overall}$ is the overall anisotropic thermal parameter, $\mathbf{U}_{\rm crystal}$ is the contribution of crystal anisotropy, $\mathbf{U}_{\rm TLS}$ is the contribution of model anisotropy (TLS), $\mathbf{U}_{\rm torsion}$ is the contribution of motion about the torsion angle and, finally, $\mathbf{U}_{\rm atom}$ is the contribution of the atomic anisotropy along and across covalent bonds. Cruickshank (1956) noted that removing $\mathbf{U}_{\rm crystal}$ made the refinement of individual anisotropic $\mathbf{U}$ values more stable, and it seems reasonable to apply these simple corrections to remove the modes related to $\mathbf{U}_{\rm crystal}$ and $\mathbf{U}_{\rm TLS}$. At present, *REFMAC* only corrects for $\mathbf{U}_{\rm crystal}$ before refining $\mathbf{U}_{\rm atom}$. The derived atomic anisotropy is thus the sum of $\mathbf{U}_{\rm TLS}$, $\mathbf{U}_{\rm torsion}$ and $\mathbf{U}_{\rm atom}$.

Care should be taken in the refinement of different contributions, as they are highly correlated. To overcome this difficulty they can be refined in successive stages, *i.e.* first refine $\mathbf{U}_{\rm crystal}$, second $\mathbf{U}_{\rm TLS}$, third $\mathbf{U}_{\rm torsion}$ and finally $\mathbf{U}_{\rm atom}$. Alternatively, refine $\mathbf{U}_{\rm crystal}$, $\mathbf{U}_{\rm TLS}$ and displacement parameters along the internal degrees of freedom, as described by Diamond (1990).

Table 1
Notation.

$\lvert F_o\rvert = \lvert F_h^o\rvert$	Experimental structure-factor amplitude.
$\lvert E^o\rvert$	Normalized experimental structure-factor amplitude.
$\sigma_{F^o} = \sigma_{\lvert F\rvert \exp}$	Experimental uncertainty in structure-factor amplitude.
$\mathbf{s}$	Vector of position in reciprocal space. $s = \lvert\mathbf{s}\rvert = 2\sin\theta/\lambda$.
ε	Multiplicity of diffracting plane.
$f(k,s)$	Form factor of kth atom in a five Gaussian approximation. $f(k,s) = \sum_{k=1}^{5} a_k \exp(-b_k s^2/4)$, $b_5 = 0$.
$\mathbf{U} = [U_{ij}],\ i = 1,3; j = 1,3$	Anisotropic displacement tensor. When used to describe atomic thermal parameter. It must be positive definite, *i.e.* define a 'thermal ellipsoid'.
$U_{\rm iso}$	Isotropic equivalent of the anisotropic U tensor. In an orthogonal system $U_{\rm iso} = (u_1 + u_2 + u_3)/3$.
$B = 8\pi^2 U$	U is the mean-square amplitude of vibration.
$\mathbf{u} = [u_i]$	Vector representation of an anisotropic tensor where $u_i = U_{ii}$, $u_4 = U_{12} = U_{21}$, $u_5 = U_{13} = U_{31}$, $u_6 = U_{23} = U_{32}$.
$\mathbf{U}^* = \mathbf{G}^*\mathbf{U}\mathbf{G}^*$	Reciprocal-space anisotropic $\mathbf{U}$ tensor, where $\mathbf{G}^* = \mathbf{G}^{-1}$. Elements of the metric tensor $G_{ij} = \mathbf{a}_i\mathbf{a}_j$ where $\mathbf{a}_i$, $\mathbf{a}_j$ are the vectors defining the coordinate system used. In an orthogonal system $\mathbf{G} = \mathbf{I}$, the identity matrix. In the cell system $\mathbf{a}_i$ and $\mathbf{a}_j$ are the cell axes.
$\mathbf{V} = \mathbf{U}^{-1} = [V_{ij}],\ i = 1,3; j = 1,3$	Inverse of U.
$k\exp(-\mathbf{h}\mathbf{U}^*_{\rm crystal}\mathbf{h}^T)$	Crystallographic scale factor. Here, $\mathbf{U}_{\rm crystal}$ is not required to be positive definite.
$\mathbf{h} = (h_1, h_2, h_3)$	Reflection index of the reflecting plane.
$\mathbf{F}_c = \lvert F_c\rvert \exp(\iota\varphi^c) = k\exp(-\mathbf{h}\mathbf{U}^*_{\rm crystal}\mathbf{h}^T) \times \sum_{\rm atoms,} kf(k,s)\exp(-\mathbf{h}\mathbf{U}^*_{\rm atom}, k\mathbf{h}^T) \times \exp(2\pi\iota\mathbf{h}\mathbf{x}_\mathbf{k})$	Calculated structure factor.
$\Sigma_N = \sum_{k=1}^{N_{\rm atom}} f^2(k,s)$	Summed over all atoms in the crystal.
$\Sigma_P = \Sigma_P(\mathbf{s}) = \sum_{k=1}^{N^P_{\rm atom}} f^2(k,s)$	For partial structure.
$\mathbf{E}_c = \lvert E_c\rvert \exp(\iota\varphi_c)$	Normalized calculated structure factor.
$\Delta\varphi$	Phase error of current model.
$\Delta\mathbf{x}$	Error in positional parameters.
D	$\langle\cos 2\pi\mathbf{s}\Delta\mathbf{x}\rangle$
σ_A	$(\Sigma_p/\Sigma_N)^{1/2}D$
$m = (\langle\cos\Delta\varphi\rangle^2 + \langle\sin\Delta\varphi\rangle^2)^{1/2}$	Figure of merit of phases. For uniform prior phase information $m = \langle\cos\Delta\varphi\rangle = I_1(X)/I_0(X)$ for acentric and $\tanh(X)$ for centric reflections, respectively.
$X = 2(\lvert E^o\rvert\lvert E_{wc}\rvert)/(2\sigma^2_{E^o} + \sigma_{A;wc})$ or $(\lvert E^o\rvert\lvert E_{wc}\rvert)/(\sigma^2_{E^o} + \sigma_{A;wc})$	For acentric and centric reflections, respectively.
$I_0(X)$ and $I_1(X)$	Zero-order and first-order modified Bessel functions of the first kind.
TLS	Translation–libration–screw parameters of the overall thermal displacement of the molecule.
NCS	Non-crystallographic symmetry.

3. Restraints on anisotropic U tensors

In the current implementation, three types of restraints on anisotropic **U** values are allowed. The first are similarity restraints between bonded atoms. If $\mathbf{U}^1$ and $\mathbf{U}^2$ are the corresponding anisotropic tensors for bonded atoms 1 and 2, then the similarity restraint requires that the difference between $\mathbf{U}^1$ and $\mathbf{U}^2$ should be minimum. There are a number of different ways of defining the distance between two matrices; here, we use the following definition known as the L_2 norm in matrix space,

$$d^2_{\rm bond}(\mathbf{U}^1, \mathbf{U}^2) = \sum_{i,j=1}^{3} (U^1_{ij} - U^2_{ij})^2. \tag{2}$$

For the vector form of U defined in Table 1 this becomes

$$d^2_{\rm bond}(\mathbf{u}^1, \mathbf{u}^2) = \sum_{i=1}^{3} (u^1_i - u^2_i)^2 + 2\sum_{i=4}^{6} (u^1_i - u^2_i)^2. \tag{3}$$

Hence, restraints on the bonded atoms means that the term $w_{\mathbf{u},{\rm bond}} \sum_{\rm bonds} d^2(\mathbf{u}^1, \mathbf{u}^2)$ is added to the residual to be minimized, where $w_{\mathbf{u},{\rm bond}}$ is a weight dependent on the bond type.

Secondly, sphericity restraints prevent atoms from becoming too elliptical; in other words, the anisotropic terms are restrained to their isotropic equivalent. In an orthogonal coordinate system, this means that all diagonal terms U_{ii} are restrained to $U_{\rm iso}$ and the non-diagonal terms to zero,

$$\begin{aligned} d^2_{\rm sphericity}(\mathbf{u}) = \sum_{i=1}^{3} (u_i - U_{\rm iso})^2 \\ + 2\sum_{i=4}^{6} u_i^2, \end{aligned} \tag{4}$$

where $U_{\rm iso} = (u_1 + u_2 + u_3)/3$ in the orthogonal coordinate system.

'Rigid-bond' restraints, as suggested by Rollett (1970) and Hirshfeld (1974), have also been implemented. The idea behind these is that atomic motions along bonds are minimal. [This implies that $\lvert\mathbf{U}_{1,\mathbf{l}} - \mathbf{U}_{2,\mathbf{l}}\rvert^2$ should be minimized, where $\mathbf{U}_{1,\mathbf{l}}$ and $\mathbf{U}_{2,\mathbf{l}}$ are the projections of the anisotropic motion of atoms 1 and 2 linked by a covalent bond $\mathbf{l}$ to that bond; see Hirshfeld (1974) for details.]

Only rigid-bond restraints are thought to be physically reasonable. Other restraints help to stabilize the refinement, but are rough approximations to the proper probability distribution, which relies on the conditional distribution of geometric parameters (*e.g.* bond distances, angles, dihedral angles) when the positional and thermal parameters for each atom are known.

The derivatives of these residuals follow from application of the chain rule. Similar restraints have been implemented in *SHELXL* (Sheldrick, 1995).

To ensure that the **U** tensor is positive definite at each cycle of refinement, its eigenvalues and eigenvectors are calculated; small or negative eigenvalues are reset to a predefined small positive value and the anisotropic **U** is recalculated with the new eigenvalues and the old eigenvectors.

4. Restraints on non-crystallographic symmetry

If there are N copies of an atom with position $\mathbf{x}_i$ and anisotropy $\mathbf{U}^i$, related by the transformation matrices $\mathbf{R}_i$, $i = 1, N$ ($\mathbf{R}_1 = \mathbf{I}$, the identity matrix), then the restraints used for the positional parameters are

Table 2
Effect of anisotropic refinement.

Space group is $P4_22_12$. d, highest resolution for this data set. I, data from frozen crystals soaked in peracetic acid solution; II, data from crystals soaked in NADPH solution; III, data from room-temperature native crystals; IV, data from frozen native crystals. Iso, isotropic refinement; Aniso, anisotropic refinement. N_{refs}, number of reflections included; N_{pars}, number of parameters. N = number of cycles for I and II, the H atoms and VDW contact list was only generated once followed by five cycles of minimization; for III and IV the procedure was repeated. R.m.s. bond, root-mean-square deviation from ideal bond lengths. For this space group $U_{\mathrm{crystal},11} = U_{\mathrm{crystal},22}$, $U_{\mathrm{crystal},12} = U_{\mathrm{crystal},13} = U_{\mathrm{crystal},23} = 0$. W = weight between X-ray and geometric terms of minimized residual based on the comparison of the traces of the second derivative matrices.

(*a*) Refinement statistics and comparison of speeds for isotropic and anisotropic refinements

Data set	d	Type	W	R value (%)	R_{free} (%)	N_{refs}	N_{pars}	R.m.s. bond (Å)	CPU (min)	N
I	1.96	Iso	1	15.6	20.7	41377	4 × 4718	0.016	11.2	5
	1.96	Aniso	1	13.7	20.1	41377	9 × 4718	0.013	26.7	5
II	1.8	Iso	2	12.1	15.3	51955	4 × 4701	0.016	15	5
	1.8	Aniso	2	10.7	14.7	51955	9 × 4701	0.014	28.7	5
III	1.5	Iso	4	11.7	14.0	90574	4 × 4632	0.019	2 × 24.5	2 × 5
	1.5	Aniso	4	9.2	12.1	90574	9 × 4632	0.012	2 × 37.8	2 × 5
IV	0.89	Iso	4	16.5	17.2	433749	4 × 4632	0.012	2 × 115	2 × 5
	0.89	Aniso	4	11.6	12.3	433749	9 × 4632	0.008	2 × 147.5	2 × 5

(*b*) Crystallographic U values for isotropic and anisotropic refinements

		$U_{\mathrm{crystal},11}$	$U_{\mathrm{crystal},33}$
I	Iso	−0.0441	0.0874
	Aniso	−0.0486	0.0962
II	Iso	−0.0148	0.0293
	Aniso	−0.0175	0.0348
III	Iso	−0.0048	0.0094
	Aniso	−0.0018	0.0035
IV	Iso	−0.0070	0.0140
	Aniso	−0.0103	0.0205

$$d^2_{\mathbf{x},\mathrm{NCS}} = w_{\mathbf{x},\mathrm{NCS}} \sum_{i=1}^{N} (\mathbf{R}_i\mathbf{x}_i - \langle \mathbf{Rx} \rangle)^2$$

$$\langle \mathbf{Rx} \rangle = \left(\sum_{i=1}^{N} \mathbf{R}_i\mathbf{x}_i \right) / N, \qquad (5)$$

where $w_{x,\mathrm{NCS}}$ is the weight on the positional parameters for this particular atom.

For anisotropic **U** values,

$$d^2_{\mathbf{U},\mathrm{NCS}} = w_{\mathbf{U},\mathrm{NCS}} \sum_{i=1}^{N} (\mathbf{R}_i^{\mathbf{U}}\mathbf{U}^i - \langle \mathbf{RU} \rangle)^2$$

$$\langle \mathbf{RU} \rangle = \left(\sum_{i=1}^{N} \mathbf{R}_i^{\mathbf{U}}\mathbf{U}^i \right) / N \qquad (6)$$

where $w_{\mathbf{U},\mathrm{NCS}}$ is the weight on the thermal parameters for this particular atom and $\mathbf{R}_i^{\mathbf{U}}$ is the transformation matrix corresponding to $\mathbf{R}_i$, described in *Appendix B*. Before applying any NCS restraint to the atomic **U** values, the overall **U** value of each molecule related to TLS should be removed. Again, the derivatives of this residual are derived by the chain rule.

5. Implementation and examples

Equations (13) and (15) for the gradient and Hessian terms for the U_{ij} calculations are derived in *Appendix A* and have been implemented in the program *REFMAC*. For the derivative calculation, fast Fourier transformation (FFT) routines are used which means each cycle is fast. The relative times for isotropic and anisotropic refinement are given in Tables 2 and 3. The only time-consuming part of the calculation is the convolution needed for the Hessian terms (15). The restraints described above have all been implemented, although the equations for NCS have not yet been fully tested. The version of *REFMAC* incorporating anisotropic **U**-value refinement is available from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994).

In the examples given below, the models have previously been refined to convergence using isotropic temperature factors. Before beginning these tests, H atoms were added in their riding positions using the *CCP*4 program *HGEN* (Collaborative Computational Project, Number 4, 1994). 5% of the reflections were chosen for R_{free} and σ_A estimation. Both isotropic and anisotropic refinement cycles were started from the same coordinate sets to compare the speed and the overall statistics, and the same weights on geometric parameters and relative weight between X-ray and geometry were applied. Neither manual rebuilding nor automatic water-addition procedures were carried out. The refinement was performed using an SGI O_2 with R5000 processor and 128 MB memory. For the four isomorphous catalase examples the same reflection set was used for R_{free} and σ_A estimation. For both individual atomic isotropic and anisotropic refinement, the overall anisotropic scale value was refined and applied to F_c. The apparent difference between the anisotropic $\mathbf{U}_{\mathrm{crystal}}$ for the two refinements arises from the high correlation between atomic and crystallographic displacement parameters. It may be better to fix the anisotropic $\mathbf{U}_{\mathrm{crystal}}$ after isotropic refinement and then to refine only the atomic anisotropic U values.

All examples were refined without rigid-bond restraints. Using rigid-bond restraints for these examples did not make a significant difference (results not shown).

5.1. Catalase at 1.96, 1.8, 1.5 and 0.89 Å resolutions

The structure of catalase from the bacterium *Micrococcus lysodeikticus* (MLC) has previously been determined (Murshudov *et al.*, 1992). Here, we used four independent data sets of different complexes or at different temperatures for this enzyme (Table 2). Table 2(a) shows that the gain from anisotropic refinement, as indicated by the fall in the $R_{\rm free}$ value, is greatest at highest resolution. The *RASTEP* (Merritt & Bacon, 1997) plot of the haem group with proximal Tyr343 and distal haem ligands shows the likely motion of the haem and surrounding atoms during the reaction observed in the 1.96 Å complex structure (Fig. 1). Details of these and other structural features will be discussed elsewhere.

5.2. RNAase Sa3 1.7 Å

Refinement of this structure was difficult (Sevčík, personal communication), probably because of substantial internal motion of the molecule. The isotropic refinement converged with the unusually high R value and $R_{\rm free}$ of 22.6 and 25.8%, respectively, and further minimization was not possible. In addition, the geometric parameters were not ideal. Anisotropic refinement rapidly reduced the R value and $R_{\rm free}$ to 18.6 and 22.8%, respectively (Table 3). The overall crystallographic anisotropic $\mathbf{U}$ values of the Sa3 crystal for isotropic refinement were $U_{\rm crystal,11} = U_{\rm crystal,22} = -0.0196$ and $U_{\rm crystal,33} = 0.0372$, and for anisotropic refinement were $U_{\rm crystal,11} = -0.0285$ and $U_{\rm crystal,33} = 0.0539$. (The crystal symmetry requires that the off-diagonal terms are equal to zero.)

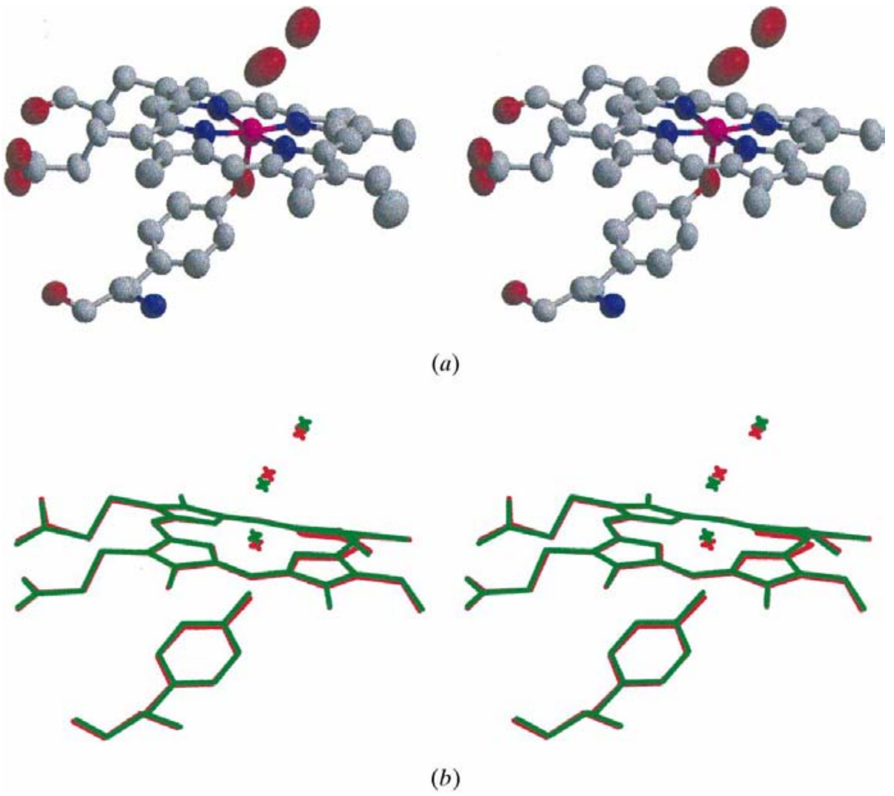

Figure 1
(a) The thermal ellipsoids for the haem group, proximal Tyr343 and two distal waters of the native MLC structure refined against 1.5 Å room-temperature data; (b) comparison of these atoms between native (red) and complexed MLC (green). The shifts of Fe and water atom sites in (b) correlate with the orientations of the thermal ellipsoides in (a). Note: no H atoms were added for haem atoms so their apparent anisotropy reflects both the absence of H atoms and real anisotropic motion.

6. Conclusions and future developments

Refinement of individual anisotropic $\mathbf{U}$ values in all examples shown here improves the crystallographic R value and $R_{\rm free}$ as well as the fit to geometric targets. The difference densities also became cleaner (less noisy). All these factors show that anisotropic thermal parameter refinement is essential in derivation of a good model from given experimental data. Intensive tests (data not shown) show that for most data sets with resolution higher than 2 Å, refinement of individual $\mathbf{U}$ values improves the R value, $R_{\rm free}$ and the fit to geometric targets.

This refinement of anisotropic $\mathbf{U}$ values using FFTs opens new perspectives. Firstly, it will have a large impact on low-resolution refinement, as the overall thermal movement of the whole molecule(s) or domain(s) can be a substantial factor contributing to the quality of diffraction. Correcting for this will increase the power of non-crystallographic symmetry restraints.

Once individual anisotropic U values are calculated, the mode of motion related to torsion angles could be removed by using information about macromolecular geometry (this has been used for normal-mode refinement by Kidera & Go, 1992). This would mean that (a) information about the internal motion of macromolecules would be available and (b) $\mathbf{U}$ values along and across bonds could be extracted.

However, anisotropic refinement needs to be carried out with care. Usually it is applied at high or atomic resolutions and in the late stages of refinement, when the error associated with the coordinates is comparable to that associated with the experimental data. The intensity-based likelihood function (Pannu & Read, 1996) or the Gaussian approximation to it (Murshudov *et al.*, 1997) can then be applied. At these resolutions, anomalous scatterers in the crystal may have an important role and the shape of the likelihood function will have to be changed to take this into account. Although the experimental $|F^o(+)|$ and $|F^o(-)|$ could be considered as independent measurements, they are not representative of independent random variables.

APPENDIX *A*
Equations for the refinement of the anisotropic thermal parameters

Let us assume that the residual used for refinement has the form

$$L = \sum_{\mathbf{h}} L_{\mathbf{h}}(\mathbf{x}, \mathbf{U}) = \sum_{\mathbf{h}} L_{\mathbf{h}}, \qquad (7)$$

Table 3
Effect of anisotropic refinement of RNAse Sa3 at 1.7 Å resolution.

Space group is $P4_12_12$. Weighting schemes for isotropic and anisotropic refinement were the same ($W = 0.7$).

	Isotropic	Anisotropic
R value (%)	22.8	2 × 5
$R_{\rm free}$ (%)	25.8	22.8
Number of reflections	9790	22.8
Number of parameters	4 × 822	9 × 822
R.m.s. bond (Å)	0.022	0.013
CPU (s)	2 × 297	2 × 423
Number of cycles	2 × 5	2 × 5

where $L_{\mathbf{h}}$ depends only on the reflection with index $\mathbf{h} = (h_1, h_2, h_3)$.

In principle, this form of residual is needed for the second derivative matrix (see Murshudov *et al.*, 1997). Model parameters are involved through structure factors,

$$\mathbf{F}_{\mathbf{h}} = A_{\mathbf{h}} + \iota B_{\mathbf{h}} = \sum_{\rm atoms} g(n, s) \exp[2\pi\iota \mathbf{h}\mathbf{r}(\mathbf{n})], \tag{8}$$

where

$$\begin{aligned} g(n, s) &= f(n, s) \exp[-2\pi^2 \textstyle\sum h_i h_j U^*_{ij}(n)] \\ &= f(n, s) \exp[-2\pi^2 \mathbf{h} U^*(n) \mathbf{h}^T]. \end{aligned} \tag{9}$$

$f(n, s)$ is the form factor of the nth atom (described by the five-Gaussian approximation), $U^*(n)$ is the reciprocal-space version of the anisotropic tensor for the nth atom. For convenience the subscript will be dropped.

The gradient of the residual will then have the form

$$\begin{aligned} \frac{\partial L}{\partial x_i(n)} &= \sum \left(\frac{\partial L}{\partial A_{\mathbf{h}}} - \iota \frac{\partial L}{\partial B_{\mathbf{h}}} \right) 2\pi\iota h_i g(n, s) \exp[2\pi\iota \mathbf{r}(\mathbf{n})\mathbf{h}], \\ \frac{\partial L}{\partial U^*_{ij}(n)} &= \sum \left(\frac{\partial L}{\partial A_{\mathbf{h}}} - \iota \frac{\partial L}{\partial B_{\mathbf{h}}} \right) (2\pi^2)(-h_i h_j) g(n, s) \\ &\quad \times \exp[2\pi\iota \mathbf{r}(\mathbf{n})\mathbf{h}], \end{aligned} \tag{10}$$

where $\mathbf{r}(\mathbf{n}) = (x_1, x_2, x_3)(n)$.

For the second derivatives, the H_1 contribution (as defined by Agarwal, 1978, and used by Murshudov *et al.*, 1997) is

$$\begin{aligned} \frac{\partial^2 L}{\partial x_i(n) \partial x_j(m)} &= \frac{1}{2} \sum \left(\frac{\partial^2 L}{\partial A^2_{\mathbf{h}}} + \frac{\partial^2 L}{\partial B^2_{\mathbf{h}}} \right) \\ &\quad \times (2\pi)^2 h_i h_j g(n, s) g(m, s) \\ &\quad \times \exp\{2\pi\iota \mathbf{h}[\mathbf{r}(\mathbf{n}) - \mathbf{r}(\mathbf{m})]\}, \\ \frac{\partial^2 L}{\partial U^*_{ij}(n) \partial U^*_{kl}(m)} &= \frac{1}{2} \sum \left(\frac{\partial^2 L}{\partial A^2_{\mathbf{h}}} + \frac{\partial^2 L}{\partial B^2_{\mathbf{h}}} \right) \\ &\quad \times (4\pi^4) h_i h_j h_k h_l g(n, s) g(m, s) \\ &\quad \times \exp\{2\pi\iota \mathbf{h}[\mathbf{r}(\mathbf{n}) - \mathbf{r}(\mathbf{m})]\}, \\ \frac{\partial^2 L}{\partial x_i(n) \partial U^*_{kl}(m)} &= -\frac{1}{2} \sum \left(\frac{\partial^2 L}{\partial A^2_{\mathbf{h}}} + \frac{\partial^2 L}{\partial B^2_{\mathbf{h}}} \right) \\ &\quad \times (4\pi^3\iota) h_i h_k h_l g(n, s) g(m, s) \\ &\quad \times \exp\{2\pi\iota \mathbf{h}[\mathbf{r}(\mathbf{n}) - \mathbf{r}(\mathbf{m})]\}. \end{aligned} \tag{11}$$

The contribution of other terms is small compared with that of H_1. These equations can be used to derive the necessary derivatives of the minimized residual. For application it is convenient to write it in the following form [(10) could be written in the same manner],

$$\begin{aligned} \frac{\partial^2 L}{\partial x_i(n) \partial x_j(m)} &= \sum_{\rm hemisphere} \left(\frac{\partial^2 L}{\partial A^2_{\mathbf{h}}} + \frac{\partial^2 L}{\partial B^2_{\mathbf{h}}} \right) \\ &\quad \times (2\pi)^2 h_i h_j g(n, s) g(m, s) \\ &\quad \times \cos\{2\pi \mathbf{h}[\mathbf{r}(\mathbf{n}) - \mathbf{r}(\mathbf{m})]\}, \\ \frac{\partial^2 L}{\partial U^*_{ij}(n) \partial U^*_{kl}(m)} &= \sum_{\rm hemisphere} \left(\frac{\partial^2 L}{\partial A^2_{\mathbf{h}}} + \frac{\partial^2 L}{\partial B^2_{\mathbf{h}}} \right) \\ &\quad \times (4\pi^4) h_i h_j h_k h_l g(n, s) g(m, s) \\ &\quad \times \cos\{2\pi \mathbf{h}[\mathbf{r}(\mathbf{n}) - \mathbf{r}(\mathbf{m})]\}, \\ \frac{\partial^2 L}{\partial x_i(n) \partial U^*_{kl}(m)} &= \sum_{\rm hemisphere} \left(\frac{\partial^2 L}{\partial A^2_{\mathbf{h}}} + \frac{\partial^2 L}{\partial B^2_{\mathbf{h}}} \right) \\ &\quad \times (4\pi^3) h_i h_k h_l g(n, s) g(m, s) \\ &\quad \times \sin\{2\pi \mathbf{h}[\mathbf{r}(\mathbf{n}) - \mathbf{r}(\mathbf{m})]\}. \end{aligned} \tag{12}$$

(10) and (11) can be written in terms of a convolution,

$$\begin{aligned} \frac{\partial L}{\partial x_i(n)} &= \mathcal{F}\left(\frac{\partial L}{\partial A_{\mathbf{h}}} + \iota \frac{\partial L}{\partial B_{\mathbf{h}}} \right) * \frac{\partial \rho(n)}{\partial x_i(n)}, \\ \frac{\partial L}{\partial U_{ij}(n)} &= \mathcal{F}\left(\frac{\partial L}{\partial A_{\mathbf{h}}} + \iota \frac{\partial L}{\partial B_{\mathbf{h}}} \right) * \frac{\partial \rho(n)}{\partial U_{ij}(n)}, \end{aligned} \tag{13}$$

where ρ is the electron density generated by the atom n, using the form factor expressed as the sum of Gaussians,

$$\rho(n, \mathbf{x}) = \sum_{i_g=1}^{5} \rho_{i_g}(n, \mathbf{x}), \tag{14}$$

where ρ_{i_g} is a Gaussian with a mean at the atomic position $r(n)$ and an uncertainty $\mathbf{U}_{i_g}(\mathbf{n}) = U_{{\rm form}, i_g}\mathbf{I} + \mathbf{U}$. $\mathbf{U}$ is the temperature-factor component and $U_{{\rm form}, i_g}\mathbf{I}$ is a matrix representation of the U_{i_g}th term of the form factor of the nth atom:

$$\begin{aligned} \frac{\partial^2 L}{\partial x_i(n) \partial x_j(m)} &= -\frac{1}{2} \mathcal{F}\left(\frac{\partial^2 L}{\partial A^2_{\mathbf{h}}} + \frac{\partial^2 L}{\partial B^2_{\mathbf{h}}} \right) * \frac{\partial^2 \rho'(nm)}{\partial x_i(n) \partial x_j(m)}, \\ \frac{\partial^2 L}{\partial U_{ij}(n) \partial U_{kl}(m)} &= \frac{1}{2} \mathcal{F}\left(\frac{\partial^2 L}{\partial A^2_{\mathbf{h}}} + \frac{\partial^2 L}{\partial B^2_{\mathbf{h}}} \right) * \frac{\partial^2 \rho'(nm)}{\partial U_{ij}(n) \partial U_{kl}(m)}, \\ \frac{\partial^2 L}{\partial U_{kl}(m) \partial x_i(n)} &= -\frac{1}{2} \mathcal{F}\left(\frac{\partial^2 L}{\partial A^2_{\mathbf{h}}} + \frac{\partial^2 L}{\partial B^2_{\mathbf{h}}} \right) * \frac{\partial^2 \rho'(nm)}{\partial x_i(n) \partial U_{kl}(m)}, \end{aligned} \tag{15}$$

where * indicates a convolution, $\mathcal{F}$ is a Fourier transformation and $\rho'(nm)$ is the convolution of atoms n and m; the sum of N_g Gaussians (15 for convolution of an atom with itself and 25 for convolution of different atoms; Agarwal, 1978),

$$\rho'(nm, \mathbf{x}) = \rho(n, \mathbf{x}) * \rho(m, \mathbf{x}) = \sum_{i_g=1}^{N_g} \rho_{i_g}(nm, \mathbf{x}). \tag{16}$$

The reciprocal-space anisotropic tensor $\mathbf{U}^*$ and its real-space equivalent are related by the metric tensor ($\mathbf{U} = \mathbf{G}\mathbf{U}^*\mathbf{G}$).

In the following discussion of the derivatives of ρ and ρ' with respect to positional and thermal parameters, an orthogonal system of coordinates is used. There is no loss of generality in this; the corrections can easily be converted back to the crystal frame.

All these equations hold for a residual of form (7) when $L_{\mathbf{h}} = L_{-\mathbf{h}}$, which is true whenever there is no anomalous scatterer in the crystal. For the gradient calculations there is no need for the requirement of form (7). To account for anomalous scattering, the residual could be modified to $L'_{\mathbf{h}} = (L_{\mathbf{h}} + L_{-\mathbf{h}})/2$ which already satisfies the required condition. However, for this case, the likelihood function itself should be modified to take into account the correlation between $F_{\mathbf{h}}$ and $F_{-\mathbf{h}}$. Treatment of this case is outside the scope of this paper. The derivation of the necessary maximum-likelihood equations and their implementation are under development.

The necessary coefficients derived from −log likelihood function for the above equations have been given by Murshudov *et al.* (1997). They are

$$\frac{\partial L}{\partial A_{\mathbf{h}}} + \iota \frac{\partial L}{\partial B_{\mathbf{h}}} = 2 \frac{\sigma_A |E_c| \exp(\iota\varphi_c) - m_{\rm comb} |E_o| \exp(\iota\varphi_{\rm comb})}{\Sigma_{\mathbf{h}}} \times \frac{\sigma_A}{\Sigma_c^{1/2}}. \tag{17}$$

This function has the same symmetry as the original space group.

For the second-derivative map,

$$\frac{\partial^2 L}{\partial A_{\mathbf{h}}} + \frac{\partial^2 L}{\partial B_{\mathbf{h}}} = \left[\frac{4}{\Sigma} - \frac{4|E_o|^2}{\Sigma^2}(1 - m^2_{\rm comb})\right] \frac{\sigma_A^2}{\Sigma_c}. \tag{18}$$

This function has the same symmetry as the Patterson function of the original space group.

To apply the above equations, the derivatives of ρ and ρ' with respect to positional and thermal parameters need to be calculated. In the following discussions, the subscript i_g will be dropped, but it is important to remember that the equations hold for each of the Gaussian terms (five for the atoms and 15 or 25 for the atom convolutions).

The determinant of $\mathbf{U}$ is

$$\begin{aligned} |\mathbf{U}| = {} & U_{11}U_{22}U_{33} - U_{11}U_{23}^2 - U_{22}U_{13}^2 \\ & - U_{33}U_{12}^2 + 2U_{12}U_{13}U_{23}. \end{aligned} \tag{19}$$

The minors of $\mathbf{U}$ are

$$\begin{aligned} U_{11}^m &= U_{22}U_{33} - U_{23}^2 \\ U_{22}^m &= U_{11}U_{33} - U_{13}^2 \\ U_{33}^m &= U_{11}U_{22} - U_{12}^2 \\ U_{12}^m &= U_{13}U_{23} - U_{33}U_{12} \\ U_{13}^m &= U_{12}U_{23} - U_{22}U_{13} \\ U_{23}^m &= U_{12}U_{13} - U_{11}U_{23}. \end{aligned} \tag{20}$$

Because of the symmetry of $\mathbf{U}$,

$$\mathbf{V} = U_{ij}^{-1} = U_{ij}^m/|\mathbf{U}|. \tag{21}$$

Using U_{ij} as elements of the anisotropic $\mathbf{U}$ tensor for the derivatives,

$$\begin{aligned} \frac{\partial\rho}{\partial x_i} &= -\frac{1}{(2\pi)^{1.5}|\mathbf{U}|^{0.5}} \sum_k V_{ik} x_k \exp\left(-\frac{\mathbf{x}^T\mathbf{U}^{-1}\mathbf{x}}{2}\right), \\ \frac{\partial\rho}{\partial U_{ij}} &= -\frac{0.5}{(2\pi)^{1.5}|\mathbf{U}|^{0.5}} \left[\left(\frac{\partial|\mathbf{U}|/\partial U_{ij}}{|\mathbf{U}|}\right) + \sum_{ts} \frac{\partial V_{ts}}{\partial U_{ij}} x_t x_s\right] \\ & \quad \times \exp\left(-\frac{\mathbf{x}^T\mathbf{U}^{-1}\mathbf{x}}{2}\right). \end{aligned} \tag{22}$$

For the second derivatives,

$$\begin{aligned} \frac{\partial^2\rho'}{\partial x_i \partial x_j} &= \frac{1}{(2\pi)^{1.5}|U|^{0.5}} \left(-V_{ij} + \sum_{kl} V_{ik}V_{il}x_k x_l\right) \\ & \quad \times \exp\left(-\frac{x^T U^{-1} x}{2}\right), \\ \frac{\partial^2\rho'}{\partial U_{ij}\partial U_{kl}} &= \frac{0.5}{(2\pi)^{1.5}|\mathbf{U}|^{0.5}} \left\{\left[1.5\left(\frac{\partial|\mathbf{U}|/\partial U_{ij}}{|\mathbf{U}|}\right)\left(\frac{\partial|\mathbf{U}|/\partial U_{kl}}{|\mathbf{U}|}\right)\right.\right. \\ & \quad \left. - \frac{\partial^2|\mathbf{U}|}{\partial U_{ij}\partial U_{kl}} \frac{1}{|\mathbf{U}|}\right] \exp\left(-\frac{\mathbf{x}^T\mathbf{U}^{-1}\mathbf{x}}{2}\right) \\ & \quad + \sum_{ts}\left[0.5\left(\frac{\partial|\mathbf{U}|/\partial U_{ij}}{|\mathbf{U}|}\right)\frac{\partial V_{ts}}{\partial U_{kl}}\right. \\ & \quad \left. + 0.5\left(\frac{\partial|\mathbf{U}|/\partial U_{kl}}{|\mathbf{U}|}\right)\frac{\partial V_{ts}}{\partial U_{ij}} - \frac{\partial^2 V_{ts}}{\partial U_{ij}\partial U_{kl}}\right] \\ & \quad \times x_t x_s \exp\left(-\frac{\mathbf{x}^T\mathbf{U}^{-1}\mathbf{x}}{2}\right) \\ & \quad + 0.5\sum_{tspq}\left(\frac{\partial V_{ts}}{\partial U_{ij}}\frac{\partial V_{pq}}{\partial U_{kl}} x_t x_s x_p x_q\right) \\ & \quad \left. \times \exp\left(-\frac{\mathbf{x}^T\mathbf{U}^{-1}\mathbf{x}}{2}\right)\right\}. \end{aligned} \tag{23}$$

Coefficients of the zero-order and fourth-order $x_t x_s x_p x_q$ can be calculated in a straightforward way. The second-order terms for the second derivatives with respect to elements of anisotropic $\mathbf{U}$ (EAU) can be simplified further. Denoting the coefficients of the second-order terms as $C^{ts}_{(ij)(kl)}$ and using the fact that $V_{ts} = U^m_{ts}/|\mathbf{U}|$ it can be shown that

$$\begin{aligned} C^{ts}_{(ij)(kl)} = V_{ts} & \left[-3.0\left(\frac{\partial|\mathbf{U}|/\partial U_{ij}}{|\mathbf{U}|}\right)\left(\frac{\partial|\mathbf{U}|/\partial U_{kl}}{|\mathbf{U}|}\right)\right. \\ & \left. + \frac{1}{|\mathbf{U}|}\frac{\partial^2|\mathbf{U}|}{\partial U_{ij}\partial U_{kl}}\right] \\ & + \frac{1}{|\mathbf{U}|}\left\{1.5\left[\frac{\partial U^m_{ts}}{\partial U_{ij}}\left(\frac{\partial|\mathbf{U}|/\partial U_{kl}}{|\mathbf{U}|}\right)\right.\right. \\ & \left.\left. + \frac{\partial U^m_{ts}}{\partial U_{kl}}\left(\frac{\partial|\mathbf{U}|/\partial U_{ij}}{|\mathbf{U}|}\right)\right] - \frac{\partial^2 U^m_{ts}}{\partial U_{ij}\partial U_{kl}}\right\}. \end{aligned} \tag{24}$$

Because of symmetry, the coefficients for $t \neq s$ must be multiplied by two. All elements in (24) can easily be calculated using the determinant, minors and inversion of the matrix,

$$\frac{\partial|\mathbf{U}|/\partial U_{ij}}{|\mathbf{U}|} = C_{ij}V_{ij}, \tag{25}$$

where $C_{ij} = 1$ for $i = j$ and $C_{ij} = 2$ for $i \neq j$. The derivative of the elements of the inverted matrix with respect to EAU can be written:

$$\frac{\partial V_{ts}}{\partial U_{kl}} = \frac{1}{|\mathbf{U}|}\frac{\partial U^{m}_{ts}}{U_{kl}} - \left(V_{ts}\frac{\partial|\mathbf{U}|}{\partial U_{ij}}/|\mathbf{U}|\right). \tag{26}$$

To calculate derivatives, firstly the terms $F[(\partial L/\partial A) + (\partial L/\partial B)]$ or $F[(\partial^2 L/\partial A^2) + (\partial^2 L/\partial B^2)]$ could be convoluted with ρ, $x_i\rho$, $x_ix_j\rho$ *etc*. for each Gaussian; secondly, these convolutions could be summed with the coefficients given in (13) and (15) and, finally, they could be summed over all the Gaussians. To achieve the required precision for the convolutions, either FFTs calculated on fine grids or interpolation techniques (see, for example, Press *et al.*, 1986) could be used.

In the current applications, we use a block-diagonal approximation to the second derivatives of the maximum-likelihood equations, *i.e.* a 6 × 6 block for the anisotropic U and a 3 × 3 block for the positional parameters. For the geometric information, both the diagonal and the non-diagonal terms are calculated.

APPENDIX *B*
Anisotropic U values as a vector

Applying the transformation matrix $\mathbf{R}$ to an atomic position ($\mathbf{x}_{\rm new} = \mathbf{R}\mathbf{x}$) requires that its anisotropic $\mathbf{U}$ tensor should be transformed to $\mathbf{U}_{\rm new} = \mathbf{R}\mathbf{U}\mathbf{R}^T$. Carrying out the matrix multiplication $\mathbf{R}\mathbf{U}\mathbf{R}^T$ and remembering that $\mathbf{U}$ is symmetric, with only six independent elements *i.e.* $U_{ij} = U_{ji}$, means that $\mathbf{R}\mathbf{U}\mathbf{R}^T$ can be described as $\mathbf{R}^{\mathbf{u}}\mathbf{u}$ where $\mathbf{R}^{\mathbf{u}}$ is a 6 × 6 matrix which pre-multiplies the vector $\mathbf{u}$ of length 6.

We can write

$$\begin{aligned}
R^{\mathbf{u}}_{ij} &= R_{ji}R_{ji}\\
R^{\mathbf{u}}_{i4} &= 2R_{i1}R_{i2}\\
R^{\mathbf{u}}_{i5} &= 2R_{i1}R_{i3}\\
R^{\mathbf{u}}_{i6} &= 2R_{i2}R_{i3}\\
R^{\mathbf{u}}_{4i} &= R_{1i}R_{2i}\\
R^{\mathbf{u}}_{5i} &= R_{1i}R_{3i}\\
R^{\mathbf{u}}_{6i} &= R_{2i}R_{3i}\\
R^{\mathbf{u}}_{44} &= R_{11}R_{22} + R_{12}R_{21}\\
R^{\mathbf{u}}_{54} &= R_{11}R_{32} + R_{12}R_{31}\\
R^{\mathbf{u}}_{64} &= R_{21}R_{32} + R_{22}R_{31}\\
R^{\mathbf{u}}_{45} &= R_{11}R_{23} + R_{13}R_{21}\\
R^{\mathbf{u}}_{55} &= R_{11}R_{33} + R_{13}R_{31}\\
R^{\mathbf{u}}_{65} &= R_{21}R_{33} + R_{23}R_{31}\\
R^{\mathbf{u}}_{46} &= R_{12}R_{23} + R_{13}R_{22}\\
R^{\mathbf{u}}_{56} &= R_{12}R_{33} + R_{13}R_{32}\\
R^{\mathbf{u}}_{66} &= R_{22}R_{33} + R_{23}R_{32}.
\end{aligned} \tag{27}$$

Here, we use the following representation of the $\mathbf{U}$ vector: $u_i = U_{ii}$, $u_4 = U_{12}$, $u_5 = U_{13}$, $u_6 = U_{23}$.

This representation of U values simplifies the application of the differential operator. If $\mathbf{x}_{\rm new} = \mathbf{R}\mathbf{x}$ and $\mathbf{u}_{\rm new} = \mathbf{R}^{\mathbf{u}}\mathbf{u}$ then for first derivatives $(\partial/\partial\mathbf{x}_{\rm new}) = \mathbf{R}^T(\partial/\partial\mathbf{x})$ and $(\partial/\partial\mathbf{u}_{\rm new}) = \mathbf{R}^{\mathbf{u}T}(\partial/\partial\mathbf{u})$. The same kind of transformation can be performed for second derivatives. The same kind of transformation for $\mathbf{u}$ could be derived if $\mathbf{R}^T$ were replaced by any other transformation matrix $\mathbf{P}$.

We thank members of the Protein Structure Group of the Chemistry Department, The University of York, for testing the program, and in particular Colin McVey and Misha Isupov (Exeter University) for useful comments. GNM was supported by a BBSRC postdoctoral fellowship awarded to the CCP4 (grant B05273). EJD is supported by the Medical Research Council (grant G94 13078). KSW and AAV thank the EC for support through BIOTECH contract CT96-0189.

References

Agarwal, R. C. (1978). *Acta Cryst.* A**34**, 791–809.

Booth, A. D. (1946). *Proc. R. Soc. London Ser. A*, **188**, 77–92.

Brünger, A. T. (1992). *X-PLOR Manual, Version 3.1*. Yale University, New Haven.

Brünger, A. T. & Rice, L. M. (1997). *Methods Enzymol.* **277**, 243–269.

Collaborative Crystallographic Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Cochran, W. (1948). *Acta Cryst.* **1**, 138–142.

Cowtan, K. D. & Ten Eyck, L. (1999). *Acta Cryst.* Submitted.

Cruickshank, D. W. (1952). *Acta Cryst.* **5**, 511–518.

Cruickshank, D. W. (1956). *Acta Cryst.* **9**, 747–753.

Dauter, Z., Lamzin, V. S. & Wilson, K. S. (1997). *Curr. Opin. Struct. Biol.* **7**, 681–688.

Diamond, R. (1990). *Acta Cryst.* A**46**, 425–435.

Driessen, H., Haneef, M. I. J., Harris, G. W., Howlin, B., Khan, G. & Moss, D. S. (1989). *J. Appl. Cryst.* **22**, 510–516.

Fletcher, E. (1981). *Practical Methods of Optimization*. New York: John Wiley.

Hirshfeld, F. (1974). *Acta Cryst.* A**32**, 239–244.

Kidera, A. & Go, N. (1992). *J. Mol. Biol.* **225**, 457–475.

Konnert, J. H. & Hendrickson, W. A. (1980). *Acta Cryst.* A**36**, 614–617.

Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.

Moss, D. S., Tickle, I. J., Theis, O. & Wostrack, A. (1996). *Proceedings of the CCP4 Study Weekend*, edited by E. Dodson, M. Moore, A. Ralph & S. Bailey, pp. 105–113. Warrington: Daresbury Laboratory.

Murshudov, G. N., Davies, G. J., Isupov, M., Krzywda, S. & Dodson, E. J. (1998). *CCP4 Newslett Protein Crystallogr.* **35**, 37–42.

Murshudov, G. N., Melik-Adamyan, W. R., Grebenko, A. I., Barynin, V. V., Vagin, A. A., Vainshtein, B. K., Dauter, Z. & Wilson, K. S. (1992). *FEBS Lett.* **312**, 127–131.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–253.

Pannu, N. S. & Read, R. J. (1996). *Acta Cryst.* A**52**, 659–668.

Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. (1986). *Numerical Recipes: the Art of Scientific Computing.* Cambridge University Press.

Rollett, J. S. (1968). Editor. *Computing Methods in Crystallography*, pp. 32–37. Oxford: Pergamon Press.

Rollett, J. S. (1970). *Crystallographic Computing*, edited by F. R. Ahmed, S. R. Hall and C. P. Huber, p. 167. Copenhagen: Munksgaard.

Schomaker, V. & Trueblood, K. N. (1968). *Acta Cryst.* B**24**, 63–76.

Sherif, S. & Henrickson, W. A. (1987). *Acta Cryst.* A**43**, 118–121.

Sheldrick, G. M. (1995). *SHELXL93, a Program for the Refinement of Crystal Structures from Diffraction Data.* Institüt für Anorganische Chemie, Göttingen, Germany.

Ten Eyck, L. (1996). *Proceedings of the CCP4 Study Weekend*, edited by E. Dodson, M. Moore, A. Ralph & S. Bailey, pp. 37–46. Warrington: Daresbury Laboratory.

Tronrud, D. E., Ten Eyck, L. F. & Matthews, B. W. (1987). *Acta Cryst.* A**43**, 483–501.

Watkin, D. (1988). *Crystallographic Computing 4*, edited by N. W. Isaacs & M. R. Taylor, pp. 111–125. Oxford University Press.

Watkin, D. (1994). *Acta Cryst.* A**50**, 411–437.