Expansion of scalar validation criteria to three dimensions: the R tensor

Sean Parkin †

Biochemistry Department, Box 3711, Duke University Medical Center, Durham, NC, USA.
Correspondence e-mail: sparkin@pop.uky.edu

Assessment of quality in crystal structure determination entails analysis of global statistics. In data reduction, quality is assessed using R_{merge} and mean $I/\sigma(I)$. Progress in structure solution and refinement is checked by the goodness of fit, variants of the R index, R_{cryst}, and its cross-validation counterpart, R_{free}. These statistics are useful and provide a convenient means of comparison but their scalar nature renders them unable to capture the essence of three-dimensional entities such as diffraction patterns and molecular models. A simple general method to quantify spatial variations in scalar statistics has been developed. In it, a symmetric matrix, the R tensor, is used to represent the local average residual as a function of diffraction geometry. An effective value of the statistic in question can then be found for any direction in reciprocal space. Differences between these effective R indices for individual reflections or groups of reflections can help to steer refinement strategy and assess the final structure.

1. Introduction

An objective meaningful measure of the agreement between a molecular model and diffraction data is elusive. There are several reasons why observed and calculated structure factors are not a perfect match. Firstly, the data themselves are not perfect. Uncertainties in reflection intensity may be caused by poor counting statistics and by bias from systematic errors such as absorption, extinction, multiple scattering and thermal diffuse scattering. Anisotropic mosaicity and asymmetric reflection profiles caused by cracking, twinning and crystal defects can lead to scan truncation and inadequate intensity integration. In addition, radiation damage may be anisotropic (Abrahams & Marsh, 1987). At the level of approximation employed in standard refinements, model deficiencies are also prevalent. These result from inadequacies inherent to spherical-atom scattering factors (Coppens *et al.*, 1969) and to the harmonic approximation used for displacement parameters (Johnson, 1969). Unrecognized twinning and disorder will also give an inferior fit (Jameson, 1996), and distinction between the two is by no means trivial using conventional techniques (Hönle & von Schnering, 1988). In most macromolecular structures, the situation is exacerbated by the paucity of diffraction data, which mandates further restrictions, particularly on displacement parameters. Poorly ordered water within such crystals is difficult to model and produces

features that affect the same low-resolution data that are prone to extinction effects. These phenomena are inherently anisotropic, so it is unrealistic to expect any scalar statistic to fully account for differences between model and experiment. Often a side-by-side comparison of observed and calculated data is a powerful way to identify refinement problems, but it is inherently subjective. Any truly representative validation method should at the very least reflect the three-dimensional nature of the diffraction pattern and molecular model.

The most commonly used validation statistics are variants of the crystallographic R index, R_{cryst}. Although useful, these statistics have known limitations that make them quite easy to manipulate (see for example Kleywegt & Brünger, 1996). Cross-validation using R_{free} (Brünger, 1992) has been of enormous value in macromolecular crystallography because of its greater resistance to manipulation and its close correlation with residual phase error. Unfortunately, it is of limited value in small-molecule work, because the smaller number of reflections involved results in poor statistical sampling for R_{free}. In any event, it is still a scalar quantity and thus provides no directional information. R indices are so pervasive in crystallography that it is unlikely they will be superseded, though other statistics have been suggested in specific instances (*e.g.* Bernardinelli & Flack, 1985). For this reason, a method to expand R_{cryst} into three dimensions was sought. This paper describes an algorithm to generate the elements of a symmetric 3×3 matrix, the R tensor and a computer program, RT, that uses it to calculate an effective R index for any direction in reciprocal space.

---

† Present address: Department of Chemistry, University of Kentucky, Lexington, KY 40506-0055, USA.

research papers

2. Algorithm

The crystallographic R index, R_{cryst}, as defined by

$$R_{cryst} = \frac{\sum_{hkl} ||F_{obs,hkl}| - |F_{calc,hkl}||}{\sum_{hkl} |F_{obs,hkl}|} = \frac{\sum_{hkl} |\Delta F_{hkl}|}{\sum_{hkl} |F_{obs,hkl}|}, \quad (1)$$

is a scalar quantity. By analogy with displacement parameters, we can visualize it as the radius of a sphere but situated at the origin of reciprocal space, as depicted in Fig. 1(a). A smaller sphere represents a better fit but it exhibits no variation over reciprocal space. The premise of this work is that the fit is unlikely to be spatially invariant but will vary with direction in reciprocal space. It is already well known that the quality of fit varies with resolution and it is common practice to plot R_{cryst} calculated in thin resolution shells. This work is not concerned with such resolution dependencies, in fact the algorithm intentionally averages over them. If we again invoke the displacement-parameter analogy, the spherical representation can be expanded to an ellipsoid to account for spatial variations (Fig. 1b).

The starting point is to find an expression for the contribution of a single reflection to R_{cryst}. The most obvious way is simply to remove the summations from equation (1) and give the quantity a new name, *viz*

$$r_{hkl} = \frac{||F_{obs,hkl}| - |F_{calc,hkl}||}{|F_{obs,hkl}|} = \frac{|\Delta F_{hkl}|}{|F_{obs,hkl}|}. \quad (2)$$

There are two problems with this approach, however. Firstly, the average over all r_{hkl} does not give R_{cryst}, which means that the result would need to be re-scaled. Secondly (and more importantly), as F_{obs} tends to zero, r_{hkl} becomes very large, effectively assigning a very high weight to that point. While the effects of this could be alleviated by a judicious weighting scheme, a different approach was chosen. The standard definition of R_{cryst} can be re-written as

$$R_{cryst} = \sum_{hkl} |\Delta F_{hkl}| \Big/ n\langle F_{obs}\rangle, \quad (3)$$

where n is the number of observations, so that an alternative definition of r_{hkl} that does average out to R_{cryst} is given by
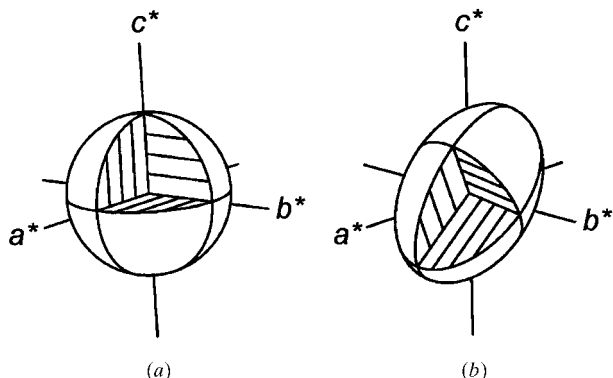
$$r_{hkl} = |\Delta F_{hkl}|/\langle F_{obs}\rangle. \quad (4)$$

A consequence of equation (4) is that r_{hkl} for an intense reflection will be greater than that for a weak reflection. This is consistent with ΔF_{hkl} being generally larger for intense reflections compared to weak reflections but is at odds with the fact that R_{cryst} for a subset of intense reflections will tend to be lower than that for a subset of weak reflections. For this reason, several variants of (4) were considered but only marginal effects on the R tensor were observed.

Effective R indices for any point in reciprocal space, R_{hkl}, can be defined in terms of R_{cryst} and a function T_{hkl},

$$R_{hkl} = T_{hkl}R_{cryst}. \quad (5)$$

T_{hkl} is obtained by minimization of the function S, defined by

$$S = \sum_{hkl} w_{hkl}(r_{hkl} - T_{hkl}R_{cryst})^2, \quad (6)$$

where w is an appropriate weighting function and T_{hkl} is given by

$$T_{hkl} = d_{hkl}^2 \sum_{i=1}^{3}\sum_{j=1}^{3}(t_{ij}h_ih_ja_i^*a_j^*). \quad (7)$$

In (7), $d_{hkl}^2 = (\lambda^2/4\sin^2\theta_{hkl})$; $h_1, h_2, h_3 = h, k, l$; $a_1^*, a_2^*, a_3^* = a^*, b^*, c^*$, respectively, and t_{ij} are the six unique elements of a symmetric 3×3 matrix, the R tensor. The minimum in S occurs when each derivative with repect to t_{ij} is zero, *i.e.* when

$$\sum_{hkl} w_{hkl}\{(r_{hkl} - T_{hkl}R_{cryst})[R_{cryst}(\partial T_{hkl}/\partial t_{i'j'})]\} = 0, \quad (8)$$

where

$$\partial T_{hkl}/\partial t_{i'j'} = d_{hkl}^2 h_{i'}h_{j'}a_{i'}^*a_{j'}^*. \quad (9)$$

In (9), the primes merely distinguish the i, j subscripts from those in (7). Substituting (7) and (9) into (8), we arrive, after rearranging terms, at six normal equations of the form

$$\sum_{hkl} w_{hkl}R_{cryst}^2 d_{hkl}^4 h_{i'}h_{j'}a_{i'}^*a_{j'}^* \sum_{i=1}^{3}\sum_{j=1}^{3}(t_{ij}h_ih_ja_i^*a_j^*)$$
$$= \sum_{hkl} w_{hkl}R_{cryst}r_{hkl}d_{hkl}^2 h_{i'}h_{j'}a_{i'}^*a_{j'}^*. \quad (10)$$

Solution by standard methods yields the six t_{ij} used to calculate T_{hkl} *via* (7) and hence R_{hkl} *via* (5).

The R tensor can be used to gauge the severity of systematic errors and the effectiveness of attempts to correct them. The anisotropy of the R tensor is greatly reduced after empirical absorption correction with *XABS*2 (Parkin *et al.*, 1995) for example. The R tensor is adaptable to other statistics such as R_{merge} (see below) provided an appropriate equivalent to (4) is available. In this context, the efficacy of corrections that strive to minimize R_{merge}, such as *SORTAV* (Blessing, 1995) and *SADABS* (Sheldrick, 1996), could be assessed. It can be applied to the full data set or to subsets of data such as those set aside for R_{free}. It can also be applied to binned data to measure anisotropy as a function of resolution. In its present formulation, there is an implicit assumption that the R-tensor surface is centrosymmetric. While this is not necessarily the

Figure 1
Graphical representation of R-tensor expanded validation statistics as ellipsoid plots: (a) conventional R indices; (b) R tensor.

Table 1
Models of uric acid dihydrate at eight stages of refinement.

Model	Displacement parameters	Twinning model	Disorder model	Number of parameters
a	Isotropic	No	No	79
b	Isotropic	Yes	No	80
c	Isotropic	No	Yes	80
d	Isotropic	Yes	Yes	81
e	Anisotropic	No	No	139
f	Anisotropic	Yes	No	140
g	Anisotropic	No	Yes	140
h	Anisotropic	Yes	Yes	141

Table 2
Summary of R-tensor results for eight uric acid models.

Model	R_{cryst}	R_{high} (reflection)	R_{low} (reflection)	R_{h00}	R_{0k0}	R_{00l}	R_{high}/R_{low}
a	13.01	17.82 (160)	10.02 ($\bar{7}$10)	10.13	17.74	16.06	1.72
b	9.61	12.95 ($\bar{1}$,0,18)	7.60 ($\bar{6}$22)	8.01	12.23	12.88	1.70
c	8.55	10.95 (0k0)	6.38 ($\bar{7}$02)	6.40	10.95	8.27	1.72
d	5.37	6.35 ($\bar{4}$,0,14)	3.95 ($\bar{4}$51)	5.25	4.47	5.84	1.61
e	10.66	15.54 (00l)	6.45 (h00)	6.45	13.20	15.54	2.41
f	7.45	11.76 (00l)	4.72 ($\bar{7}$21)	5.07	8.90	11.76	2.49
g	8.02	10.12 (067)	5.64 (h00)	5.64	9.62	7.37	1.79
h	4.81	5.60 ($\bar{2}$,2,15)	3.58 ($\bar{3}$61)	4.44	3.74	5.07	1.56

case, this assumption provides an adequate level of approximation. In any event, it is quite possible to introduce a noncentrosymmetric aspect to the R tensor by addition of higher (*i.e.* third) cumulants.

3. Program

The R-tensor algorithm has been implemented in a Fortran program, *RT*. For input, *RT* requires two *SHELXL* files: a $*$. fcf file containing h, k, l, F_{calc}^2, F_{obs}^2, $\sigma(F_{obs}^2)$ and a $*$. res file. It runs completely automatically, without the need for any further input from the user. On output, it writes a modified $*$. fcf file with an additional column listing R_{hkl} values for each reflection present. A summary, listing the highest and lowest effective R indices and the points for which they occur, axial R_{hkl} values R_{h00}, R_{0k0}, R_{00l}, t_{ij}, and the eigenvalues of the R tensor along with their respective eigenvectors (given as fractional hkl), is also written to a file. The ratio of R_{high} to R_{low} has proven to be a particularly useful diagnostic tool.

4. Examples

4.1. Application to a small-molecule crystal exhibiting twinning and disorder

In spite of the small size and innocuous appearance of uric acid dihydrate, its crystal structure has only recently appeared. It was presented in one instance as a disordered orthorhombic model (Artioli *et al.*, 1997, hereafter AMG), and in another as a monoclinic model requiring both disorder and twinning (Parkin & Hope, 1998*a*, hereafter PH). The crystals had very different origins, human bladder stones in AMG and laboratory grown in PH, and had distinctly different properties. Full details are not reproduced here; it will suffice to say that a twinned model in AMG was dismissed in favour of a disorder model, whereas, in PH, models for both twinning and disorder were needed for a satisfactory fit. Anisotropic R indices were calculated using *RT* for eight models of varying sophistication (Table 1) for the monoclinic description of PH to see how twinning and disorder affect the fit. From the results (Table 2), several things are evident: The fit is never totally isotropic, as quantified in the R_{high}/R_{low} ratio, but the better fits are associated with lower anisotropy. The higher anisotropy of the three incomplete models with anisotropic displacement par-

ameters (ADPs) compared to their isotropic counterparts is intriguing. Despite reductions in R_{cryst} on refinement of ADPs, the anisotropy in quality of fit as a function of position in reciprocal space widens. This can be attributed to the tendency of ADPs to soak up the effects of those systematic errors that manifest themselves as an apparent slow variation in scale factor.

Models that include both twinning and disorder give a better fit between observed and calculated data. In the case of isotropic displacement parameters, the anisotropy in R_{hkl} of the models with either twinning or disorder is similar. It would appear from R_{cryst} that modelling of disorder improves the fit more than modelling of twinning. It is apparent from R_{high}, R_{low} and the axial R indices, however, that improvements to the fit in either case occur in different regions of reciprocal space. If only twinning is modelled, R_{high} is quite closely aligned with $\mathbf{c}^*$, whereas a disorder model gives R_{high} directly along $\mathbf{b}^*$. Inclusion of both swings R_{high} back towards $\mathbf{c}^*$. In models b and c, as with model a, R_{low} is close to $\mathbf{a}^*$. On inclusion of both twinning and disorder, R_{low} is aligned between $\mathbf{a}^*$ and $\mathbf{b}^*$. This behaviour of the R tensor with respect to twinning and disorder is important as it implies the ability to distinguish between them. The twin model was constructed by relating reflections hkl to $\bar{h}\bar{k}l$ and corresponds to twin domains related about a noncrystallographic twofold axis. Disorder was modelled by mapping (x, y, z) to $(\frac{1}{2} - x, 1 - y, z)$, which corresponds to individual flipped molecules. The hydrogen-bonding network is little perturbed by this 180° flip of the uric acid molecules.

With ADP models, R_{cryst} alone suggests that the twin model is superior to the disorder model, yet the anisotropy of the fit is higher for model f than model g. This is consistent with the greater propensity of ADPs to account for the effects of disorder compared with twinning. It is not possible in this example for a disorder model to mimic the effects of twinning, nor can a twin model mimic the effects of disorder.

One of the strengths of R indices is rooted in their scalar nature; it is easy to compare one against another. This is not

Table 3
R-tensor results applied to R_{merge} for three protein data sets.

Crystal	R_{merge}	$R_{merge,high}$ (reflection)	$R_{merge,low}$ (reflection)	R_{merge} ($h00$)	R_{merge} ($0k0$)	R_{merge} ($00l$)	Ratio high/low
1	3.92	6.02 ($3,38,\overline{20}$)	2.55 ($\overline{32},9,15$)	2.73	5.51	3.81	2.36
2	11.20	15.59 ($\overline{34},5,1$)	6.36 ($19\overline{8}$)	15.47	8.41	9.58	2.45
3	4.95	6.15 ($3,\overline{1},28$)	2.44 ($24,\overline{1},\overline{4}$)	2.52	5.79†	6.08	2.52

† There were no $0k0$ reflections in the final reduced data set, but the R tensor is nevertheless capable of predicting effective R_{merge} values along $\mathbf{b}^{*}$.

quite as straightforward for R-tensor expanded statistics, but ellipsoid plots can provide an intuitive way of representing them (Fig. 2).

4.2. Application to data-reduction statistics

The R-tensor algorithm has been adapted to assess anisotropy in commonly used data-reduction statistics such as R_{merge} and mean $I/\sigma(I)$.

4.2.1. Merging R values. Changes entail working throughout with F^2 rather than F and replacement of equation (4) with a functional equivalent, *viz*

$$r_{merge,hkl} = |F_{hkl}^2 - \langle F_{hkl}^2 \rangle| / \langle F^2 \rangle. \quad (4a)$$

The method was applied to three separate but related data sets. The crystals in question were of a polymerase (Kiefer *et al.*, 1997) with bound fragments of DNA (Kiefer *et al.*, 1998) and differed in the DNA sequence and crystal size. They were handled, cooled and mounted using standard techniques (Parkin & Hope, 1998*b*). Sample 1 was well formed (0.20 × 0.20 × 0.25 mm) with moderate mosaicity (~0.64°). Sample 2 had low mosaicity (~0.32°), but was smaller (0.10 × 0.10 × 0.25 mm) and gave much weaker diffraction. Sample 3 was larger (0.25 × 0.25 × 0.45 mm) and multiply cracked. Its mosaicity was noticeably anisotropic and high for these crystals (~1.1°), yet still gave a very reasonable conventional R_{merge}. Further details, along with the results of the R-tensor analysis, are given in Table 3 and Figs. 3(*a*)–(*c*).

The anisotropy of these R_{merge} tensors is greater than those in example 1 for the R tensor applied to conventional R values. The ellipsoids for crystals 1 and 3 are oriented in a similar manner, whereas crystal 2 is quite different. The origin of the difference is unclear but may be related to crystal orientation and higher redundancy (6.0) of data for crystal 2 compared with crystals 1 and 3 (3.5 and 3.4, respectively). Use of R_{merge} as a validation statistic has been criticized recently because of this lack of objectivity. Since increased redundancy inevitably leads to higher R_{merge}, redundancy differences over reciprocal space may skew the anisotropy. Suitably weighted alternatives to R_{merge} have been suggested (*e.g.* Diederichs & Karplus, 1997) but have not gained wide acceptance. These robust alternatives to R_{merge} are perfectly amenable to expansion *via* the R tensor and should alleviate the problem of varying redundancy.

4.2.2. Mean $I/\sigma(I)$. This is recognized as superior to R_{merge}, but its scalar nature limits the amount of information it can
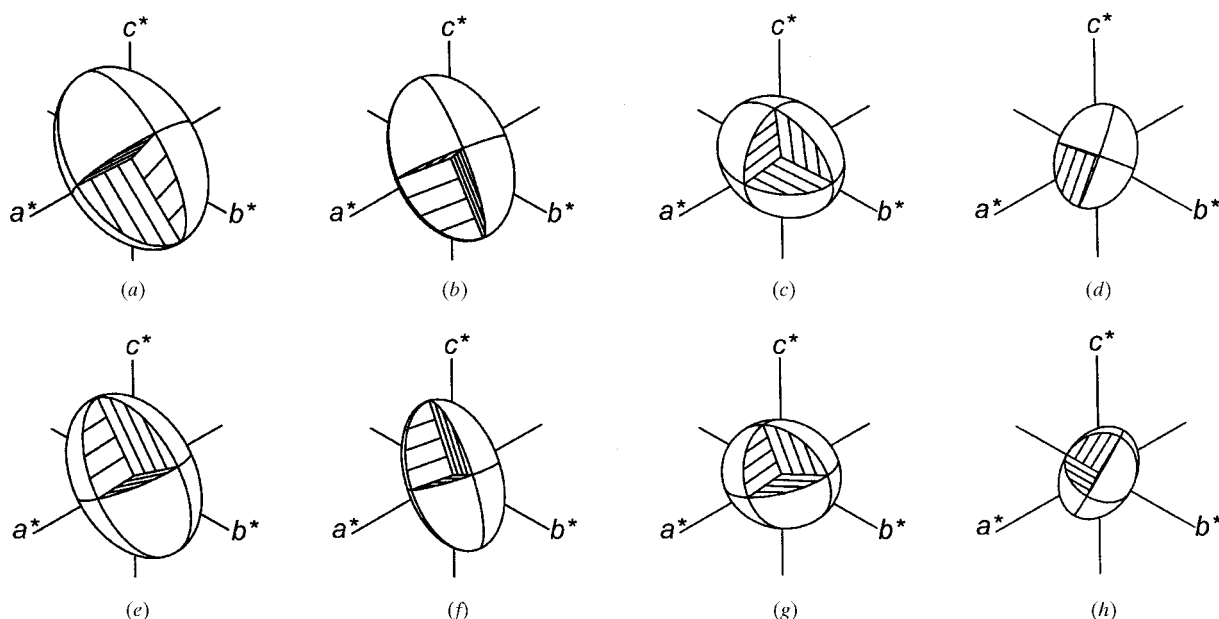

Figure 2
Ellipsoid-plot representations of R-tensor results for the eight uric acid dihydrate models labelled *a*–*h* in Table 1. A smaller volume and lower anisotropy are indicative of a better fit.

Table 4
R tensor applied to mean $I/\sigma(I)$ for three protein data sets.

Crystal	$\langle I/\sigma(I)\rangle$	$\langle I/\sigma(I)\rangle_{high}$ (reflection)	$\langle I/\sigma(I)\rangle_{low}$ (reflection)	$\langle I/\sigma(I)\rangle$ (h00)	$\langle I/\sigma(I)\rangle$ (0k0)	$\langle I/\sigma(I)\rangle$ (00l)	Ratio high/low
1	15.24	16.75 $(\bar{6},\overline{41},\bar{5})$	13.65 $(\overline{23},7,\overline{35})$	14.56	16.65	14.24	1.23
2	5.87	7.41 $(\bar{2},\overline{20},10)$	3.71 $(\overline{18},5,8)$	4.11	7.05	5.99	2.00
3	19.00	24.58 $(0\bar{7}1)$	14.23 $(22,0,4)$	14.33	24.57†	18.77	1.73

† There were no 0k0 reflections in the final reduced data set, but the R tensor is nevertheless capable of predicting effective $I/\sigma(I)$ values along $\mathbf{b}^*$.

convey. Anisotropic diffraction strength is a well known problem affecting macromolecular crystals and an objective method of quantifying it should be a valuable tool. Variability in diffraction strength is correlated with molecular geometry and crystal symmetry, so there are often sound structural reasons for it. Diffraction from crystals of planar hydro-carbons, for example, can be very intense in some directions relative to others. In proteins, however, the three-dimensional arrangement of large numbers of atoms suggests that variation should be more subtle. A suitable replacement for (4) is

$$[I/\sigma(I)]_{hkl} = \langle I_{hkl}\rangle/\sigma(I_{hkl}). \qquad (4b)$$

Anisotropy in mean $I/\sigma(I)$ was calculated for the same three protein data sets, and results are summarized in Table 4 and Figs. 3(d)–(f). There is a marked difference in the anisotropy
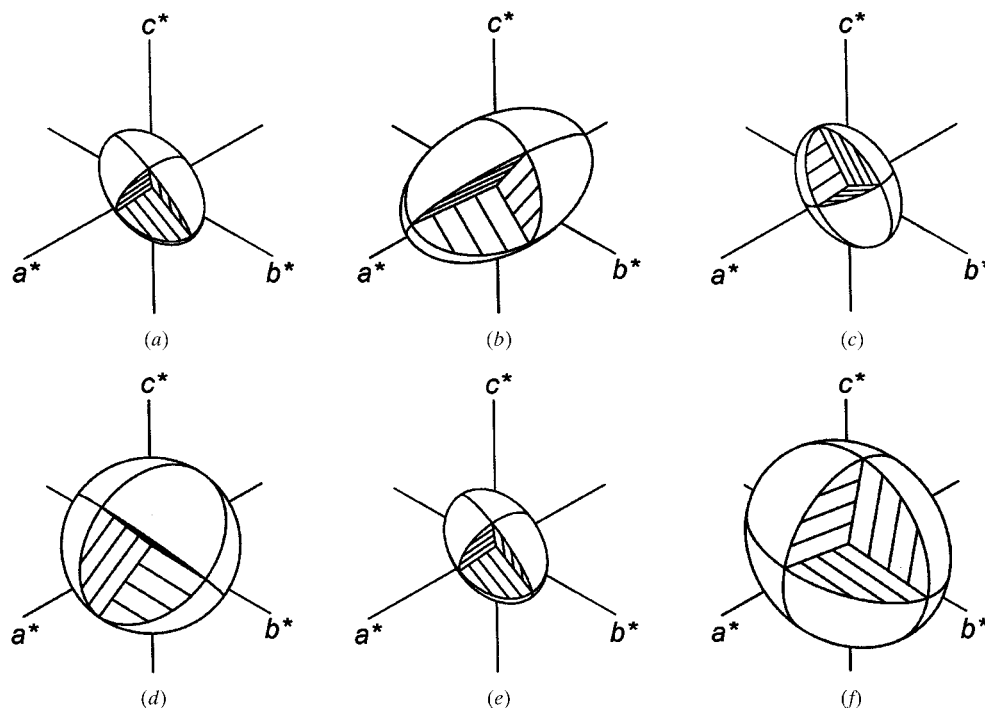
of the strong (1), weak (2) and cracked (3) data sets; it being much smaller for crystal 1. An interesting point is that for crystal 3 there were no reflections 0k0 in the final data set, as a result of $\mathbf{b}^*$ being closely aligned with the rotation axis. Nevertheless, the R tensor predicts those reflections to be among the strongest, as is the case with crystals 1 and 2. It has been suggested (Jameson, 1999) that there may be some complementarity between anisotropy in R_{merge} and mean $I/\sigma(I)$. No such relations are obvious for crystals 1 and 3, but in a qualitative sense it appears true for crystal 2 [compare Figs. 3(b), 3(e) with 3(a), 3(d) and 3(c), (3f)].

4.3. Application to R_{free}

Free R values have proven invaluable for detection and prevention of over-fitting in macromolecular crystallography (Kleywegt & Brünger, 1996). It is possible to use the R tensor to expand R_{free} in the same way as for R_{cryst}. To illustrate this, the R tensor was applied to both R_{cryst} and R_{free} in a refinement of actinoxanthin, a small 110 amino acid protein that diffracted to beyond 1.0 Å at low temperature. In spite of the presence of a wedge-shaped gap in one data set that limited resolution along $\mathbf{a}^*$ to about 3 Å, electron-density maps were clear and convincing, and the refinement proceeded smoothly to a model with excellent geometry (Kuzin *et al.*, 1997). Sufficient data to fill the missing wedge were available using a second crystal but these were not used initially for fear of smearing some of the fine details only available at true atomic resolution (Parkin *et al.*, 1996). Anisotropy in R_{cryst} and R_{free} for the initial (incomplete) data set and after refinement against the full data set are presented in Table 5 and Fig. 4. The R_{free} data in the two sets were the same apart from the addition of a small number of reflections from the extra wedge. It should be noted that this example is a little contrived in that re-refinement was not carried through from the beginning, but only after the

Figure 3
Ellipsoid-plot representations of data reduction statistics. (a) R_{merge} 'strong' data set; (b) R_{merge} 'weak' data set; (c) R_{merge} 'cracked crystal' data set; (d) mean $[I/\sigma(I)]$, 'strong' data set; (e) mean $[I/\sigma(I)]$, 'weak' data set; (f) mean $[I/\sigma(I)]$, 'cracked crystal' data set. In (a)–(c), small volume and low anisotropy correspond to a better fit, whereas, in (d)–(f), large volume and low anisotropy are best.

research papers

Table 5
R tensor applied to R_{free} for actinoxanthin.

Statistic	Incomplete data set	Full data set
R_{cryst}	11.59	12.32
R_{h00}	18.63	16.66
R_{0k0}	11.03	11.57
R_{00l}	8.02	8.70
$R_{\text{high}}/R_{\text{low}}$	2.32	1.91
R_{free}	14.22	14.49
$R_{\text{free},h00}$	26.21	21.43
$R_{\text{free},0k0}$	11.41	12.62
$R_{\text{free},00l}$	10.15	10.96
$R_{\text{free,high}}/R_{\text{free,low}}$	2.58	1.96

coordinates had been subjected to small random shifts, and the ADPs recalculated from their isotropic equivalents. Nevertheless, the results are interesting and do highlight some important features. Firstly, overall R_{cryst} and R_{free} are larger for the second data set, as expected. Incomplete data gave $R(h00)$ and $R_{\text{free}}(h00)$ much higher than average, with large anisotropy. With the full data set, however, the fit along $\mathbf{a}^*$ is better, and anisotropy is lower despite an overall rise in R values (compare Figs. 4a, 4b with 4c, 4d). Typical R_{free} calculations require that 5–10% of the data be set aside. For adequate sampling, R_{free} sets should be a little larger for application of the R tensor to R_{free}. In particular, it is advisable to have at least one reflection along each of $\mathbf{a}^*$, $\mathbf{b}^*$ and $\mathbf{c}^*$ in the R_{free} set, as well as perhaps 10–15% of reflections in general.
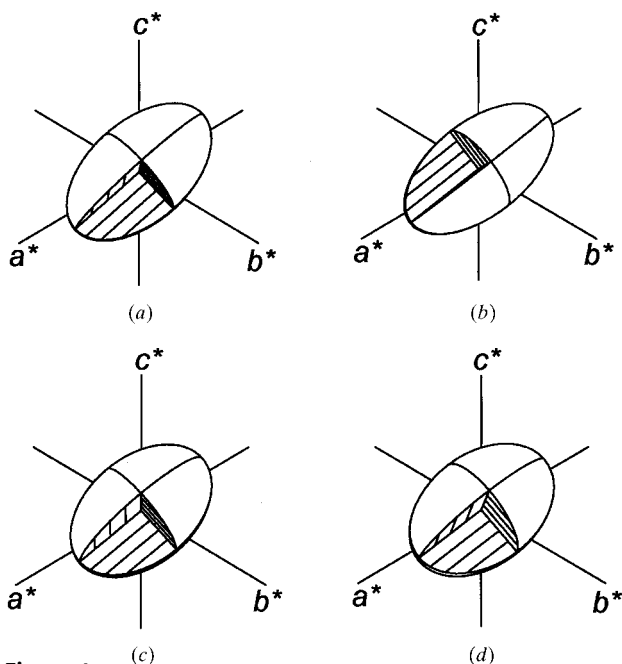

Figure 4
Ellipsoid-plot representations of R-tensor results for R_{cryst} and R_{free}. (a) R tensor on R_{cryst} for incomplete data set; (b) R tensor on R_{free} for incomplete data set; (c) R tensor on R_{cryst} for full data set; (d) R tensor on R_{free} for full data set.

5. Conclusions

The ideas described in this paper represent an attempt to quantify some of the more subjective information that can be obtained by comparing quantities in different regions of reciprocal space.

The R tensor is a new idea. Although its main strength is one of signalling the existence of problems, it is also possible in some cases to make partial diagnoses of those problems. This diagnostic capacity is currently limited by lack of experience with the information it provides. The R-tensor concept is not limited to any particular statistic. It may be adapted to assess spatial variability in any quantity, including real-space measures such as root-mean-square deviations in atomic coordinates. The necessary criterion is a suitable equivalent to (4).

The R tensor has been incorporated into the *WinGX* program system (Farrugia, 1999) and into the *CRYSTALS* package (Watkin *et al.*, 1999). The stand-alone program, *RT*, which implements the R tensor, is available from the author.

It is a pleasure to thank Professor Håkon Hope for independently testing *RT* and for useful comments on this manuscript. I also thank Dr Jeff Taylor and Sean Johnson of Duke University for providing the three protein data sets.

References

Abrahams, S. C. & Marsh, P. (1987). *Acta Cryst.* A**43**, 265–269.
Artioli, G., Masciocchi, N. & Galli, E. (1997). *Acta Cryst.* B**53**, 498–503.
Bernardinelli, G. & Flack, H. D. (1985). *Acta Cryst.* A**41**, 500–511.
Blessing, R. H. (1995). *Acta Cryst.* A**51**, 33–38.
Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.
Coppens, P., Sabine, T. M., Delaplane, R. G. & Ibers, J. A. (1969). *Acta Cryst.* B**25**, 2451–2458.
Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Biol.* **4**(4), 269–275.
Farrugia, L. J. (1999). *J. Appl. Cryst.* **32**, 837–838.
Hönle, W. & von Schnering, H. G. (1988). *Z. Kristallogr.* **184**, 301–305.
Jameson, G. B. (1996). *Acta Cryst.* A**52**, C43–C44.
Jameson, G. B. (1999). Private communication.
Johnson, C. K. (1969). *Acta Cryst.* A**25**, 187–194.
Kiefer, J. R., Mao, C., Braman, J. C. & Beese, L. S. (1998). *Nature (London)*, **391**, 304–307.
Kiefer, J. R., Mao, C., Hansen, C. J., Basehore, S. L., Hogrefe, H. H., Braman, J. C. & Beese, L. S. (1997). *Structure*, **5**, 95–108.
Kleywegt, G. & Brünger, A.T. (1996). *Structure*, **4**, 897–904.
Kuzin, A. P., Parkin, S. R., Trakhanov, S. D., Rupp, B., Wilson, D. K. & Quiocho, F. A. (1997). American Crystallographic Association, July 1997, St. Louis, MO, USA.
Parkin, S. & Hope, H. (1998a). *Acta Cryst.* B**54**, 339–344.
Parkin, S. & Hope, H. (1998b). *J. Appl. Cryst.* **31**, 945–953.
Parkin, S., Moezzi, B. & Hope, H. (1995). *J. Appl. Cryst.* **28**, 53–56.
Parkin, S., Rupp, B. & Hope, H. (1996). *Acta Cryst.* D**52**, 1161–1168.
Sheldrick, G. M. (1996). *SADABS. A Program for Absorption Correction.* University of Göttingen, Germany.
Watkin, D. J., Prout, C. K., Carruthers, J. R., Betteridge, P. W. & Cooper, R. I. (1999). *CRYSTALS*, Issue 10. Chemical Crystallography Laboratory, University of Oxford, Oxford, England.