Acta Cryst. (1996). D**52**, 743–748

Bayesian Weighting for Macromolecular Crystallographic Refinement

THOMAS C. TERWILLIGER[a]* AND JOEL BERENDZEN[b]

[a]*Structural Biology Group, Mail Stop M888, Los Alamos National Laboratory, Los Alamos, NM 87545, USA, and* [b]*Biophysics Group, Mail Stop D454, Los Alamos National Laboratory, Los Alamos, NM 87545, USA*

(Received 8 March 1995; accepted 31 January 1996)

Abstract

A simple weighting scheme for atomic refinement is discussed. The approach, called 'Bayesian weighting', is designed to be robust with respect to the bias that arises from the incomplete nature of the atomic model, which in macromolecular crystallography is typically quite serious. Bayesian weights are based on the mean-squared residual errors over shells of resolution, with centric and acentric reflections considered separately and with allowances made for experimental uncertainties. Use of Bayesian weighting is shown in test cases typical for macromolecular crystallography to improve the accuracy of the refined coordinates when compared with schemes employing unit weights or experimental variances.

1. Introduction

The philosopher Paul Valéry once remarked that 'a work of art is never finished, only abandoned'. He could well have said the same about structural models of macromolecules. If one has been so fortunate as to obtain high-quality crystallographic data with a resolution of 2.5 Å or better, one builds an atomic model containing parameters such as coordinates and *B* factors. One then alternates between refinement of the parameters (generally by least-squares minimization) and grudging elaboration upon the model (by adding water molecules, for example). The refined set of model parameters constitutes our working estimate of the macromolecular structure. One forms this estimate, however, in the presence of imperfections in the working model, which most often does not fully include such details as the structure of the solvent, multiple conformations, and anisotropic and anharmonic motions (Gros, van Gunsteren & Hol, 1990; Kuriyan, Petsko, Levy & Karplus, 1986). Nor is it always obvious how to remedy these or other imperfections in a parsimonious way. The point at which one says 'good enough' and abandons the problem is rarely a point at which the data are described to within experimental uncertainty; the disagreement between calculated and observed structure factors is typically in the range of 15 to 20%, even though the data are typically accurate to about 5% (Jensen, 1985). This deficiency can have a substantial effect on the accuracies of the refined models. Identical protein structures refined in different laboratories, for example, typically differ by 0.2–0.3 Å r.m.s. (Kuriyan *et al.*, 1986; Daopin, Davies, Schlunegger & Grütter, 1994).

We wish to find an approach to atomic refinement that is robust with respect to incompleteness of the working model and that returns the most likely set of model parameters, given the experimental data (structure factors) and certain prior knowledge about the system (*e.g.*, bond lengths). We shall employ the Bayesian formulation of probability theory, which is eminently suited to this task. We begin by describing the general statistical approach and show that, given certain simplifying assumptions, it leads to familiar least-squares refinement with a somewhat-modified weighting scheme for the experimental data involving the r.m.s. discrepancies between calculated and observed structure factors. We have applied this scheme, termed 'Bayesian weighting', to macromolecular crystallographic model cases involving simulated and measured data and shown that application of the method can yield a model that is considerably more accurate than methods based on uniform weighting or experimental uncertainties alone.

2. The Bayesian view of refinement

The Bayesian approach to estimation of model parameters is not distinguished from the 'frequentist' approach by mere application of Bayes' theorem, but rather by the extent to which prior knowledge is incorporated into the process and also by the way the resulting posterior probability is interpreted — as a quantity monotonically related to the subjective likelihood of the parameters being correct (Cox, 1946). Though somewhat controversial, this interpretation nonetheless corresponds better to the typical crystallographer's notions of what he or she is doing in refinement than does the frequentist interpretation.

Let us represent our knowledge of a set of experimentally determined structure factors by $\mathcal{F}$, and let $\boldsymbol{\theta}$ represent the statement that a particular set of model parameters (*i.e.*, coordinates, temperature factors, and occupancies) are correct. Underlying our model is a set of assumptions, A, not only about the applicability of the atomic model, but also about such geometrical properties as the range of lengths found for C—C bonds. The 'most likely' set of parameters we shall take to be that set which maximizes the probability of the statement $\boldsymbol{\theta}$ given the data $\mathcal{F}$ and the assumptions A, denoted as $p(\boldsymbol{\theta}|\mathcal{F}, A)$. This distribution (called the posterior probability distribution because it is formed after we have taken in account the new information in $\mathcal{F}$) is given by Bayes' theorem (Box & Tiao, 1973) as,

$$p(\boldsymbol{\theta}|\mathcal{F}, A) = p(\mathcal{F}|\boldsymbol{\theta}, A)p(\boldsymbol{\theta}|A)/p(\mathcal{F}|A). \quad (1)$$

The distribution $p(\mathcal{F}|\boldsymbol{\theta}, A)$, called the likelihood, describes the probability that the data are consistent with the values of the parameters and the assumptions of the model. The distribution $p(\boldsymbol{\theta}|A)$, called the prior probability distribution or (considered as a set) 'priors', describes the probability of a particular set of parameter values being true; it is based on our knowledge of the situation before conducting the experiment. The term in the denominator, $p(\mathcal{F}|A)$, gives the probability that the particular data set $\mathcal{F}$ would crop up from the set of all possible data sets, given our prior knowledge; it is a constant that may be obtained by normalization if necessary, but for most purposes may be simply neglected.

To find the most likely set of parameters, $p(\boldsymbol{\theta}|\mathcal{F}, A)$ is not usually maximized directly but rather the negative logarithm of something proportional to it, $L(\mathcal{F}|\boldsymbol{\theta}, A)$, is minimized. Dropping the explicit listing of the assumptions A throughout and neglecting the additive constant from $p(\mathcal{F}|A)$, we obtain the most common basis for parameter estimation,

$$L(\boldsymbol{\theta}|\mathcal{F}) = -\log[p(\mathcal{F}|\boldsymbol{\theta})] - \log[p(\boldsymbol{\theta})]. \quad (2)$$

In the absence of model errors and when the experimental uncertainties are normally distributed, the first term on the right-hand side is proportional to the familiar χ^2 statistic (Bevington & Robinson, 1992) that is the basis of least-squares fitting,

$$\chi^2 = \sum_h W_h[F_{o_h} - F_{c_h}(\boldsymbol{\theta})]^2. \quad (3)$$

Here the F_{o_h} are the observed structure factors for the reflection with indices represented by h, and $F_{c_h}(\boldsymbol{\theta})$ is the corresponding structure factor calculated from the model. The W_h are weights, which in the absence of model errors or outliers is given by the reciprocal of the square of the experimental uncertainty, $1/\sigma^2_{\rm obs}$.

A word about the second term is in order before proceeding. In the case where the priors are flat when compared with the likelihood ('non-informative priors', in statistical jargon) then finding the most likely parameter set reduces to minimizing χ^2. In macromolecular crystallography, however, the priors usually are informative, and the $\log[p(\boldsymbol{\theta})]$ term is familiar (though usually from a different approach) as restraints on model geometry. Indeed, macromolecular crystallographers have been using a Bayesian approach to refinement for many years with little controversy, perhaps because the priors are accurately known from small-molecule crystallography and because using them has been shown to dramatically improve the quality of the resulting structure. While these restraints usually appear as various energy-like terms in the functional to be minimized (*cf.* Hendrickson, 1985; Brünger & Nilges, 1993), and are converted to probability densities through Boltzmann statistics, they could equally well be considered simply as likelihoods from the start. That is, if we find in surveying accurately determined C—C bond distances a normal distribution about a mean distance, then we can expect that C—C distances in a macromolecule should obey the same distribution. The negative logarithm of the distribution will give rise to a term harmonic in deviation from the mean distance. One advantage of this point of view is that informative prior probability distributions may exist for combinations of parameters that are not interpretable as energies, such as can be found in profile methods for assessing correctness of structures (Bowie & Eisenberg, 1994).

As pointed out in the introduction, model errors are quite significant in macromolecular crystallography. As a consequence, various practices for carrying out atomic refinement have arisen more or less *ad hoc* because they have been empirically shown to improve convergence or the quality of the final parameter set. The sum in least-squares minimization (3), for example, is often restricted to be over only those reflections that have been measured to some minimum level of accuracy (*e.g.*, $F > 2\sigma_{\rm obs}$) or those above some minimum resolution (*e.g.*, $d_{\rm min} < 5$ Å). Various schemes for the weighting factors W_h have also been proposed. Hendrickson (1985) advocated using weights given by $1/\langle(F_o - F_c)^2\rangle_{\rm sh}$ (where $\langle\rangle_{\rm sh}$ denotes an average over a shell of resolutions) for the initial stages of refinement and experimental weights for the later stages. Others have used unit weights for all the data (Harris & Moss, 1992; Brünger & Nilges, 1993). The capability for using a variety of weighting schemes is present in most macromolecular crystallographic refinement packages.

In crystallographic analyses of small molecules, the weighting problem has been dealt with in considerable detail. Weightings proportional to $1/F_0^2, 1/F_0^4$, or $\lambda/\sin\theta$ are sometimes used, either alone or in combination with experimental uncertainties (*cf.* Hong & Robertson, 1985). The modified weights are generally produced in a manner that leads to a χ^2 approximately equal to the number of observations. Wilson (1973, 1976*a*,*b*, 1979) has examined the effects of weighting and refinement method on errors in crystallographic models and has shown that the method used can affect both the biases and the variances in the refined model. However, in small-molecule crystallography the overall model errors are in the order of a few percent of the structure factors. It is not certain whether the same approaches would work for macromolecular crystallography, where the model errors are on the order of 10–20%.

3. Robust estimation in the presence of model errors

We have recently presented a simple framework for describing the effects of model incompleteness on the refinement process in macromolecular crystallography (Terwilliger & Berendzen, 1995). In this framework it is assumed explicitly that most features of the electron-density distribution in the crystal can be described by the atomic model being used in the refinement process, but that some cannot. For example, if the atomic model does not contain parameters describing multiple conformations of the structure, then any multiple conformations actually present in the structure cannot possibly be described by the model, and fits to the most highly populated conformation would miss the additional electron density to the less-populated conformations. Calculations based on such a model will be missing an additive part of the complex structure factor, which we will denote as $\mathbf{D}_h$ for the reflection with indices h.

If we accurately knew the $\mathbf{D}_h$ we could immediately account for the defects in the model, because then the observed structure factors F_{o_h} would be given by

$$F_{o_h} = |\mathbf{F}_{c_h}(\boldsymbol{\theta}) + \mathbf{D}_h| + \delta_h, \qquad (4)$$

where $\mathbf{F}_{c_h}(\boldsymbol{\theta})$ is the complex structure-factor amplitude calculated from the partial model with the set of parameters $\boldsymbol{\theta}$ and δ_h is the measurement error. Of course, we do not know the $\mathbf{D}_h$, but we can make a statistical assessment of their effects on the amplitudes of the calculated structure factors as a probability distribution conditional upon the value of the atomic parameters, $p(\mathcal{D}|\boldsymbol{\theta})$, where $\mathcal{D}$ is a set of 'discrepancies' defined as additive to structure-factor amplitudes. We can use this information to improve our estimates of the parameters specified by $\boldsymbol{\theta}$. Following the approach of Box (Box, 1980) we shall treat $\mathcal{D}$ as a 'nuisance parameter' and integrate over all possible values, weighting by the probability of obtaining the discrepancy conditional upon the value of the atomic parameters. This process is known in the statistical literature as 'marginalization'. The probability of the data agreeing with the parameters becomes in this case,

$$p(\mathcal{F}|\boldsymbol{\theta}) = \int p(\mathcal{F}|\mathcal{D}, \boldsymbol{\theta}) p(\mathcal{D}|\boldsymbol{\theta}) \, \mathrm{d}\mathcal{D}, \qquad (5)$$

where the integral is over all values of the discrepancies $\mathcal{D}$. Assuming a normal distribution of measurement errors, we can write first term in the integral (to within a multiplicative constant) as the product over all reflections

$$p(\mathcal{F}|\mathcal{D}, \boldsymbol{\theta}) \propto \prod_h \mathcal{N}(F_{o_h} - |\mathbf{F}_{c_h}(\boldsymbol{\theta}) + \mathbf{D}_h|, \sigma_h^2), \qquad (6)$$

where $\mathcal{N}(x, \sigma^2) = 1/\sigma(2\pi)^{1/2} \exp(-x^2/2\sigma^2)$ represents a normal distribution with variance σ^2.

3.1. *Estimation of the discrepancy distribution*

Since the discrepancies arise from those parts of the true structure that are not included in the model, they are generally uncorrelated with the calculated or observed structure factors. It is thus not possible to estimate them individually using some function of F_{c_h} or F_{o_h}. We can, however, estimate the distribution of the discrepancy $p(\mathcal{D}|\boldsymbol{\theta})$ given the observations and the structure factors calculated for a group of related reflections. Using the space-group-dependent intensity factors ε_h (Stewart & Karle, 1976) to describe the variation of intensity among particular classes of reflections, we may write the mean-square amplitude of the $\mathbf{D}_h$ as $\varepsilon_h \langle |\mathbf{D}_h^2| / \varepsilon_h \rangle_{\rm sh}$. The problem of obtaining $p(\mathcal{D}|\boldsymbol{\theta})$ is then identical to the situation encountered in the analysis of heavy-atom model errors in isomorphous replacement (*cf.* Terwilliger & Eisenberg, 1987). We shall assume that the complex structure factor discrepancies $\mathbf{D}_h$ are small and uncorrelated with respect to the calculated structure factors F_{c_h}. Then the relation between observed and calculated structure factors (4) can be approximated as,

$$F_{o_h} \simeq F_{c_h}(\boldsymbol{\theta}) + D_h + \delta_h, \qquad (7)$$

where $D_h = |\mathbf{D}_h| \cos(\alpha_h)$ and α_h is the difference in phase angle between $\mathbf{F}_{c_h}$ and $\mathbf{D}_h$. While the expected value of $\langle |\mathbf{D}_h^2| / \varepsilon_h \rangle_{\rm sh}$ is the same for centric and acentric reflections (Wilson, 1949), the factor $\cos^2(\alpha_h)$ is always unity for centric reflections and it has a mean value of 1/2 for acentric reflections. Thus, the expected value of D_h^2 for centric reflections is twice the expected value for acentric reflections. The shapes of distributions related to $p(\mathcal{D}|\boldsymbol{\theta})$ have been examined in detail before (*cf.* Ramachandran, Srinivasan & Sarma, 1963; Read, 1986). However, for the present purposes $p(\mathcal{D}|\boldsymbol{\theta})$ is sufficiently well described by the product of normal distributions $\prod_h \mathcal{N}(D_h, \varepsilon_h E^2)$.

The integral in (5) may now be rewritten as,

$$p(\mathcal{F}|\boldsymbol{\theta}) = \prod_h \int \mathcal{N}(F_{o_h} - [F_{c_h}(\boldsymbol{\theta}) + D_h], \sigma_h^2) \mathcal{N}(D_h, \varepsilon_h E^2) \mathrm{d}D_h, \qquad (8)$$

and carried out without further approximation. Rearranging and taking the logarithm of both sides leads to the first term in (2), the negative log likelihood of the data agreeing with the parameters,

$$-\log[p(\mathcal{F}|\boldsymbol{\theta})] = \tfrac{1}{2} \sum_h [F_{o_h} - F_{c_h}(\boldsymbol{\theta})]^2 / (\varepsilon_h E^2 + \sigma_h^2). \qquad (9)$$

This is a least-squares equation (3) with weights W_h given by $1/(\varepsilon_h E^2 + \sigma_h^2)$.

The value of E^2 may be readily estimated. From (7), we can write

$$\langle D_h^2/\varepsilon_h\rangle \simeq \langle (F_{o_h} - F_{c_h})^2/\varepsilon_h\rangle_{\rm sh.a/c} - \langle \delta_h^2/\varepsilon_h\rangle_{\rm sh.a/c}, \quad (10)$$

where the notation $\langle\rangle_{\rm sh.a/c}$ denotes an average taken over an appropriate shell of resolution, with acentric and centric reflections treated separately. Substituting E^2/Q for $\langle D_h^2/\varepsilon_h\rangle$, where $Q = 1$ for centric reflections and 2 for acentric reflections, and substituting the experimental variance σ_h^2 as an estimate of δ_h^2, we obtain the following relation,

$$E^2 \simeq Q[\langle (F_{o_h} - F_{c_h})/\varepsilon_h^2\rangle_{\rm sh.a/c} - \langle \sigma_h^2/\varepsilon_h\rangle_{\rm sh.a/c}]. \quad (11)$$

(9) and (11) provide a straightforward means of applying weighting factors that reflect both experimental and model errors to the refinement of macromolecular structures. (7) is strictly valid only when we knew beforehand the values of D_h. If we use the results of a previous round of refinement without discrepancies fully taken into account, the F_{c_h} will tend to be closer to the F_{o_h} than they ought, and our estimates of E^2 will tend to be slight underestimates. In most cases, this should affect the results of refinement only very slightly.

4. Tests of weighting schemes with simulated data and partially complete models

The treatment presented here suggests that an effective weighting for least-squares macromolecular refinement would consist of adding the mean square model variance for the appropriate class of reflections and range of resolution to the experimental variances. We tested the utility of Bayesian weighting by comparing refinements performed with it against those performed with two other commonly used weighting schemes, unit weighting (equal weights for all reflections), and experimental weighting (weights of $1/\sigma_h^2$), in a case with simulated data where the 'right answer' is known by definition. In the tests, varying numbers of atoms were excluded from the refinement procedure so as to simulate the effects of incomplete models.

We based the simulated data on the refined structure of gene V protein (Skinner *et al.*, 1994), which we call the 'known model'. Known model structure factors and intensities were calculated from the 711 non-H atoms of this structure, which includes 45 water molecules, for 7639 reflections from 5 to 1.8 Å in space group *C*2, with cell dimensions of $a = 76$, $b = 28$ and $c = 42$ Å. The effects of measurement errors were simulated by adding a normally distributed random variable to the known model intensities. Amplitudes and variances in amplitudes calculated from these intensities and known errors were used as the data for refinement. $\langle\sigma_{\rm obs}\rangle/\langle F\rangle$ was set at 5%.

Identical refinement procedures were used with each weighting scheme. A model with deviations from the known model structure representing errors was generated by deleting 75 atoms from the known model structure and refining the coordinates and temperature factors of this incomplete model with unit weights using *X-PLOR* (Brünger, Karplus & Petsko, 1989). The resulting model, which had an r.m.s. coordinate error of 0.074 Å for main-chain atoms and 0.094 Å for side-chain atoms, was used as the starting point for all subsequent refinements, with varying numbers of the deleted atoms replaced. The coordinates and temperature factors of the atoms in the models were refined to convergence against the simulated data using *X-PLOR*, adjusting the overall weighting on structure factors so as to obtain an r.m.s. deviation of bond lengths of 0.013 Å from ideality. For Bayesian weighting [using (9)] the values of E^2 were estimated from the starting model in shells of resolution using (11).

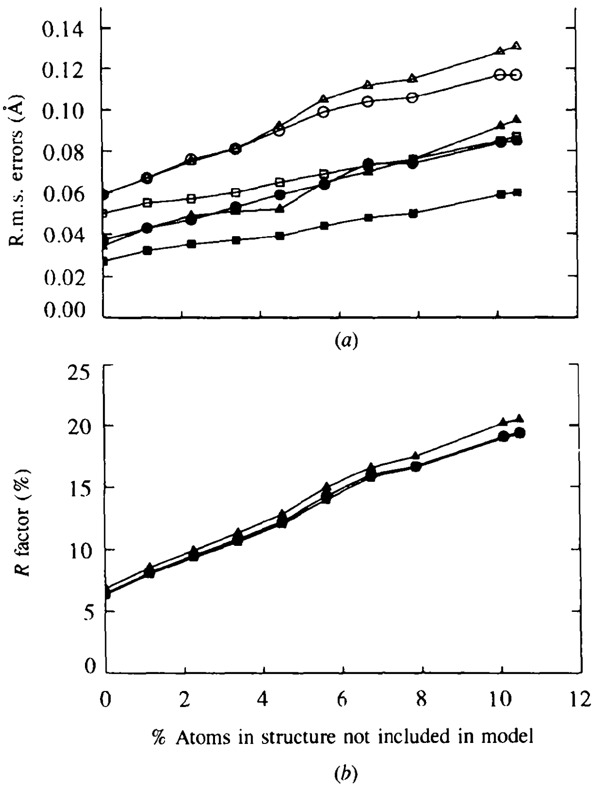

Fig. 1. The effects of model incompleteness upon refined structures using unit weighting (circles), experimental weighting (triangles) and Bayesian weighting (squares). Simulated data were generated by adding a 5% variance to a set of structure factors calculated from a gene V protein structure. Refinements against this data set were carried out from a new starting point with varying percentages of the 711 non-H atoms missing from the model. (*a*) The r.m.s. coordinate errors for the main-chain atoms (closed symbols) and side-chain atoms (open symbols) in the refined models. (*b*) The standard *R* factors for the refinements shown in (*a*).

We evaluated the refinements by comparing the coordinates of the refined models with those of the known model. Fig. 1(*a*) shows the r.m.s. coordinate error for each weighting scheme as a function of the percentage of atoms missing from the refined models; Fig. 1(*b*) shows the standard *R* factors for each of the same refinements.

When all atoms are included in the model, so that the model error after refinement is small, all three weighting methods yield small coordinate errors. Bayesian weighing gives smaller coordinate errors than experimental weighting even in this case, possibly because at the start of refinement the model contains considerable error. Standard *R* factors for these refinements were in the range 4.6–5.0%.

Bayesian weighting is much more advantageous when the model errors are larger. With 10.5% of all atoms in the structure not included in the model, for example, the standard *R* factors were in the range 19–21%, and Bayesian weighting yielded an r.m.s. error for main-chain atoms of 0.060 Å. Unit weighting was some 40% worse, with an r.m.s. error for the same atoms of 0.085 Å, and experimental weighting was roughly 60% worse, with an r.m.s. error of 0.095 Å, despite the fact that the variances were much better estimates of the errors in 'measurement' than would be the case with actual data.

We next examined whether the estimates of the model variance made using (11) were accurate estimates and whether the variances for centric and acentric reflections indeed differed by a factor or two. We used (11) to calculate the known model errors, with the known model structure factors substituted for those from the refined model. Fig. 2 illustrates the actual and estimated values of E^2 for the centric and acentric reflections in various ranges of resolution for the refinement carried out above with 10.5% of atoms not included in the refinement. It may be seen that the centric model variance is indeed twice the acentric model variance and that the estimates of this variance obtained from the refinement itself are quite good.

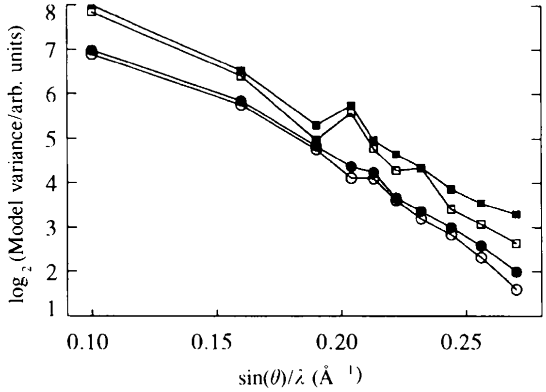

Fig. 2. Model variances as a function of resolution for the test case with 10.5% of atoms not included in the model. The known model variances are indicated by filled symbols, while those estimated from the refinement are indicated by open symbols. Model variances for centric reflections are indicated by squares, those for acentric reflections are indicated by circles. Variances for the two classes of reflections are predicted to differ by a factor of two, or one unit on a $\log_2$ scale.

5. Conclusions

The analyses presented here show that the deficiencies in macromolecular crystallographic models lead to errors in refinement that are similar to those that would result from large errors in measurement. Consequently, in least-squares refinement procedures the weighting function should include not just experimental variances but also model variances that reflect these deficiencies in the model. The test cases using model and real data indicate that a weighting method that includes the model variance resulted in refined structures that are more accurate than those obtained with unit weighting or with weighting based on experimental variances alone.

The authors gratefully acknowledge generous grants from the NIH, the International Human Frontiers Organization, and from the LDRD program of Los Alamos National Laboratory. Code to calculate Bayesian weights has been implemented in version 4 of the package *HEAVY*, available from TT at terwilliger@lanl.gov.

References

Bevington, P. R. & Robinson, D. K. (1992). *Data Reduction and Error Analysis for the Physical Sciences*. New York: McGraw-Hill.

Bowie, J. U. & Eisenberg, D. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 4436–4440.

Box, G. E. P. (1980). *J. R. Statist. Soc. A*, **143**, 383–430.

Box, G. E. P. & Tiao, G. C. (1973). *Bayesian Inference in Statistical Analysis*. New York: John Wiley.

Brünger, A. T., Karplus, M. & Petsko, G. A. (1989). *Acta Cryst.* A**45**, 50–61.

Brünger, A. T. & Nilges, M. (1993). *Q. Rev. Biophys.* **26**, 49–125.

Cox, R. T. (1946). *Am. J. Phys.* **14**, 1–48.

Daopin, S., Davies, D. R., Schlunegger, M. P. & Grütter, M. G. (1994). *Acta Cryst.* D**50**, 85–92.

Gros, P., van Gunsteren, W. F. & Hol, W. G. J. (1990). *Science*, **249**, 1149–1152.

Harris, G. W. & Moss, D. S. (1992). *Acta Cryst.* A**48**, 42–45.

Hendrickson, W. A. (1985). *Methods Enzymol.* **155**, 252–270.

Hong, W. & Robertson, B. E. (1985). In *Structure & Statistics in Crystallography*, edited by A. J. C. Wilson. New York: Adenine Press.

Jensen, L. H. (1985). *Methods Enzymol.* **155**, 227–237.

Kuriyan, J., Petsko, G. A., Levy, R. M. & Karplus, M. (1986). *J. Mol. Biol.* **190**, 227-254.

Ramachandran, G. N., Srinivasan, R. & Sarma, V. R. (1963). *Acta Cryst.* **16**, 662–666.

Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.

Skinner, M. M., Zhang, H., Leschnitzer, D. H., Bellamy, H., Sweet, R. M., Gray, C. M., Konings, R. N. H., Wang, A. H.-J. & Terwilliger, T. C. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 2071-2075.

Stewart, J. M. & Karle, J. (1976). *Acta Cryst.* A**32**, 1005–1007.

Terwilliger, T. C. & Berendzen, J. (1995). *Acta Cryst.* D**51**, 609–618.

Terwilliger, T. C. & Eisenberg, D. S. (1987). *Acta Cryst.* A**43**, 6–13.

Wilson, A. J. C. (1973). *Acta Cryst.* **29**, 1488–1490.

Wilson, A. J. C. (1976*a*). *Acta Cryst.* A**32**, 781–783.

Wilson, A. J. C. (1976*b*). *Acta Cryst.* A**32**, 994–996.

Wilson, A. J. C. (1979). *Acta Cryst.* A**35**, 122–130.