RESEARCH PAPERS

Acta Cryst. (1995). D**51**, 609–618

Difference Refinement: Obtaining Differences Between Two Related Structures

BY THOMAS C. TERWILLIGER*

Genomics and Structural Biology Group, Mail Stop M880 Los Alamos National Laboratory, Los Alamos, NM 87545, USA

AND JOEL BERENDZEN

Biophysics Group, Mail Stop M715, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

(*Received* 29 *April* 1994; *accepted* 15 *November* 1994)

Abstract

There are many examples in macromolecular crystallography where interest focuses on the differences between a previously determined 'native' structure and a nearly isomorphous 'variant'. In such cases, a useful approach to atomic refinement of the variant structure is through weighted least-squares minimization of the residual between the observed and calculated *differences* in amplitudes of structure factors, a strategy first used in the refinement of deoxycobalt hemoglobin [Fermi, Perutz, Dickinson & Chien (1982). *J. Mol. Biol.* **155**, 495–505] and termed 'difference refinement'. For cases in which the modeling errors for the native and variant structures are correlated, theoretical arguments indicate that difference refinement should lead to improved estimates of structural differences when compared with conventional independent refinement. Tests employing simulated peptide data sets and real data from a wild-type protein and a mutant show that difference refinement can substantially reduce errors in the differences between structures when compared with independent refinement. The algorithm is very easy to implement and does not increase the computational demands of refinement.

1. Introduction

Macromolecular crystallographers are increasingly interested in the differences among members of a family of closely related structures, such as may be obtained through mutation, through binding of substrate, ligand, or inhibitor, or through photolysis or pH change. Examples include changes in the structure of hemoglobin upon metal substitution (Fermi *et al.*, 1982; Luisi & Shibayama, 1989), ligand binding (Perutz, Fermi, Abraham, Poyart & Bursaux, 1986), or mutation (Nagai *et al.*, 1987; Huang *et al.*, 1990); the mutants of bacteriophage T4 lysozyme, for which over 100 variant structures have been obtained (Matthews, 1993; Eriksson, Baase & Matthews, 1993); the transient hydrolytic intermediates found in trypsin catalysis (Singer, Smalås, Carty, Mangel & Sweet, 1993); and the changes in the structure of HIV type-1 protease when complexed with an inhibitor (Jaskolski *et al.*, 1991). Often in such work, the related molecules have crystallized isomorphously, and in some cases the differences of interest are quite small, involving displacements of 0.2 Å root-mean-square (r.m.s.) or less. Measurement of structural changes this small in a macromolecule is a substantial challenge, as the uncertainties in coordinates are also in this range.

Optimally obtaining differences among a large family of structures is beyond the scope of the present paper; we will limit discussion to a single nearly isomorphous pair of structures. We shall assume that the data set corresponding to the first, or 'native' structure, is equal or superior in quality, resolution, and completeness to the data set for the second, or 'variant' structure. We shall further assume that a satisfactory model for the native structure has already been determined.

To obtain the differences between such a pair of structures, one generally starts by applying difference Fourier techniques based on the native structure model to obtain a crude model for the variant structure, which is then subjected to least-squares atomic refinement. In this process, the atomic coordinates and other model parameters describing the structure are adjusted so as to minimize a functional, which in standard practice is of the form,

$$\chi^2 = \sum_h (F_{c_h} - F_{o_h})^2/(\sigma_h^2 + E^2), \qquad (1)$$

where F_{o_h} is the observed amplitude of the structure factor for the reflection with indices represented by h, σ_h is the instrumental uncertainty, E is an estimate of the model error, usually a simple function of the resolution of the data (and often omitted), and F_{c_h} is the amplitude of the corresponding structure factor calculated from a model. We shall employ primes in subsequent equations in referring to the quantities associated with the variant,

* To whom reprint requests should be addressed.

and unprimed symbols for the native. The sum is generally over all reflections that have been measured to some minimum level of accuracy. In practice, various energy-like terms are also added to the functional to provide restraints on model geometry (*cf.* Hendrickson, 1985; Sussman, 1985).

At the conclusion of atomic refinement, the variant structure is compared with the native. Because the native structure was obtained through refinement without reference to the variant and because there are no terms in the functional to be minimized that refer to the native, we call such refinement 'independent refinement'. It was suggested some time ago (Fermi *et al.*, 1982) that through a modification of this functional, it might be possible to obtain cancellation of errors that might otherwise contribute differently to the two structures. This modified procedure, *difference refinement*, was applied to a number of problems regarding changes in the structure of hemoglobin (Fermi *et al.*, 1982; Perutz *et al.*, 1986; Luisi & Shibayama, 1989; Nagai *et al.*, 1987; Huang *et al.*, 1990), but the limits of applicability of the method have not been addressed. We show here that difference refinement can indeed allow one to obtain the coordinate displacements between two structures to a higher accuracy than can independent refinement, particularly when the completeness or resolution of the variant structure data is low.

2. Model incompleteness and its effects

It is well known that structure factors calculated from atomic models of the type usually employed to describe macromolecules do not fit the data to within experimental uncertainty. Although the data are generally accurate to about 5%, residuals are typically 15 to 20% (Jensen, 1985). The reason for this discrepancy is thought to be the incompleteness of the models, which fail to fully describe such features as the structure of the solvent, anisotropic and anharmonic atomic motions, and the presence of multiple conformations and disorder (Gros, van Gunsteren & Hol, 1990; Kuriyan, Petsko, Levy & Karplus, 1986). The poor fits of atomic macromolecular models to the measured data result in substantial uncertainties in coordinates. Refinements of identical protein structures carried out in different laboratories, for example, typically yield coordinates that differ by 0.2–0.3 Å r.m.s. (Kuriyan *et al.*, 1986; Daopin, Davies, Schlunegger & Grütter, 1994).

Suppose that one obtains the amplitude of a structure factor, F_{o_h}, as the result of a measurement. This value is equal to the 'true' amplitude of the structure factor F_h, plus the measurement error ε_h.* Further suppose that most features of the electron-density distribution in the crystal could be described by a certain atomic model, but that some cannot. For example, one may have neglected to include a water molecule that is adjacent to a polar surface residue. If one employs this model for atomic refinement, the calculated amplitudes of the structure factors will differ from the measured ones in a way that reflects the incompleteness of the model as well as the errors in measurement. Since one is missing an additive feature of the electron density, the model will fail to predict a term that should be added to the complex structure factor. Therefore, we may write for the observed amplitude of the structure factor,

$$F_{o_h} = |\mathbf{F}_{m_h} + \mathbf{\Delta}_{m_h}| + \varepsilon_h. \qquad (2)$$

The first term, $\mathbf{F}_{m_h}$, corresponds to that part of the 'true' complex structure factor that could, with some values of the parameters, be represented by the model m, while $\mathbf{\Delta}_{m_h}$, which we call the model-error term, is a complex structure factor corresponding to those features of the structure that cannot possibly be described by the model. Assuming that the model was mostly complete and that the model-error term is, therefore, much smaller than the first term, we can approximate the observed amplitude as

$$F_{o_h} \simeq F_{m_h} + D_{m_h} + \varepsilon_h, \qquad (3)$$

where F_{m_h} is the magnitude of $\mathbf{F}_{m_h}$ and $D_{m_h} = |\mathbf{\Delta}_{m_h}|\cos(\alpha_h)$, where α_h is the difference in phase between $\mathbf{F}_{m_h}$ and $\mathbf{\Delta}_{m_h}$.

The amplitude of the structure factor corresponding to that part of the structure which could potentially be described by the incomplete model is, therefore, approximately given by F_{m_h}, which is related to the measured amplitude by an effective error of $D_{m_h} + \varepsilon_h$. Thus, the incompleteness of the crystallographic model will affect the refinement process much as an error in measurement would. Unlike the measurement errors, though, we expect the model errors to be highly correlated across data sets where there is a high degree of isomorphism and where the model fails in the same way. Moreover, the model-error term is not generally small. While the experimental uncertainties are of the order of 5%, the differences between observed amplitudes of structure factors and those calculated from a model are generally in the range 15–20% for a refined structure. From (3) it may be seen that D_{m_h} must also be about 15 to 20% of F_{o_h}.

3. Estimating differences between pairs of structures

3.1. *Independent refinement*

When a native and a variant structure are refined independently and then compared, it might be expected that the error in the difference between them would be larger than the errors in either of the component structures.

*Although the value of ε_h^2 is unknown, it has an expectation value given by the experimental uncertainty σ_h^2.

We suggest, however, that the error in the difference may actually be considerably smaller than those in the component structures because many of the errors are duplicated and, therefore, cancel when differences are calculated, at least so long as the models, data sets and refinement procedures used are very similar.

By writing a primed version of (3) for the variant structure factor and subtracting it from the unprimed native structure expression, we may express the difference between measured native and variant amplitudes of the structure factors as,

$$\Delta F_{o_h} = \Delta F_{m_h} + \Delta\varepsilon_h + (D'_{m'_h} - D_{m_h}). \quad (4)$$

The first term on the right-hand side is the change in the amplitudes of structure factors corresponding to parts of the structures that can be represented by the native and variant models. The second term is the difference in experimental errors. The term in parentheses is the change in model errors, where $D'_{m'_h}$ is the amplitude of the structure factor corresponding to features that could not be described by the model m' for the variant structure. We argue, as did Fermi *et al.* (1982), that the model-error terms can be very similar for corresponding variant and native reflections. If the model-bias terms cancel, then the differences in those parts of the structure that can be represented by the model, ΔF_{m_h}, are estimated to within experimental uncertainty by the difference in observed structure factors, ΔF_{o_h}. Comparing (3) and (4) it may be seen that differences in parameters between the native and variant models can, in such a case, be more accurately estimated than the parameters in the models themselves due to the cancellation of the biasing effects of the model-error term. This cancellation will occur if the same features of the electron density are absent from both models, and if the phases of native and variant complex structure factors are roughly equal.

3.2. *Difference refinement*

In order for the model-error terms to cancel in independent refinement, it is required not only that the models m and m' be similar and that the method of refinement gives consistent estimates of the model parameters, but it is also required that the same reflections be used for refinement of the two structures. If, for example, the variant data set is incomplete (or at lower resolution) relative to the native data set, then the model-error terms that are part of measured amplitudes of the structure factors only present in the native data set do not cancel. In many cases the same reflections cannot be readily obtained in both data sets. For example, a single native structure might be refined, then used over a long period as a point of comparison with many different variant structures, with a different subset of reflections used for each variant structure refinement.

We suggest, therefore, that difference refinement (Fermi *et al.*, 1982) could be a useful alternative approach to refinement of a variant structure in cases where a refined native structure already exists, and through which small differences between a native and a variant structure may be obtained even if the variant data set is much less complete or at lower resolution. In difference refinement, the *differences* between a model for the variant structure and a previously refined model for the native structure are refined based on the measured *differences* between variant and native amplitudes of the structure factors by minimizing a functional such as,

$$\chi^2_{\text{diff}} = \sum_h [(\Delta F_{c_h} - \Delta F_{o_h})^2/(\sigma_h^2 + \sigma_h'^2 + E^2_{\text{diff}})], \quad (5)$$

where ΔF_{o_h} is the measured difference between native and variant amplitudes of the structure factors for a reflection with indices h, ΔF_{c_h} is the difference between calculated native and variant amplitudes of structure factors calculated from their respective models, σ_h and σ'_h are the instrumental uncertainties in the measured amplitudes of structure factors, and E_{diff} is an estimate of the residual model error. The quantity E_{diff} is not usually known and, as with the E term in independent refinement, must be estimated in some way. In carrying out refinements, we have simply adjusted E_{diff} as a fraction of the mean amplitude of structure factors in each resolution range until the r.m.s. bond deviations from ideality were equal to the target value of 0.03 Å. We have also ignored experimental uncertainties. Methods for estimating E_{diff} will be discussed elsewhere.

The sum is over all reflections that have measurements for both native and variant structures. The native model is held fixed in this procedure and the residual is minimized by adjusting the variant structure model parameters. The advantage of difference refinement is that by subtracting the observed values of the variant and native amplitudes of the structure factors, the model-error terms tend to cancel, even when different sets of reflections are used in refinement of the two structures. A trivial rearrangement of the difference refinement functional shows how difference refinement may be easily implemented,

$$\chi^2_{\text{diff}} = \sum_h [(F'_{c_h} - F_{\text{diff}})^2/(\sigma_h'^2 + \sigma_h^2 + E^2_{\text{diff}})], \quad (6)$$

where

$$F_{\text{diff}} = F'_{o_h} - (F_{o_h} - F_{c_h}), \quad (7)$$

F_{c_h} and F'_{c_h} are the amplitudes of structure factors calculated from the native and variant structure models, and F_{o_h} and F'_{o_h} are the corresponding observed amplitudes of the structure factors. Comparison with the independent refinement functional, (1), reveals that if we carry through the refinement by replacing the variant structure-factor data by the readily calculated F_{diff} and

by taking the net experimental uncertainties as the square root of the sum of the squares of the two component uncertainties, then all of our calculational tools should continue to work as in conventional independent refinement (Fermi *et al.*, 1982). It is also wise to carry out such a refinement in as nearly the same way that the native set refinement was carried out as is possible, in order to obtain maximal cancellation of any systematic errors in modelling.

The quantity $F_{\rm diff}$ may be thought of as simply the variant data corrected by an estimate of the model error in the native structure. This way of looking at difference refinement also makes it clear when difference refinement will work and when it will not. Difference refinement requires that the model error in the variant structure be strongly correlated with the model error in the native structure. If it is not, then the model-error correction in difference refinement would simply be adding noise to the variant data. This view also shows why it is useful to include the term $E^2_{\rm diff}$ in the denominator of the functional. If the model error in the native and variant structures are not identical, then there will still be some model error remaining in difference refinement which should be reflected in the weighting factors.

Difference refinement can be applied in cases where there are changes in unit-cell parameters so long as the model errors in the native and variant structures remain highly correlated. The unit-cell parameters for the native are used in calculations of structure factors for the native structure and those for the variant are used in structure-factor calculations for the variant structure.

4. Three idealized test cases

We have constructed simulated data sets in order to compare independent refinement and difference refinement and to identify which method is likely to yield the most accurate results under different circumstances. The data sets were derived from atomic models of a peptide with 51 atoms. Different conformations of the peptide were generated by a short molecular dynamics simulation using *X-PLOR* (Brünger, Karplus & Petsko, 1989). One conformation was selected as the 'native'; others served as 'variants'. We refer to the structures from which the simulated data were derived as the 'known' native and variant structures, as distinct from the 'model' structures. To each of these known atomic models was added two 'unmodeled' water molecules to represent that part of the known structures not included in the refined models. Positions of the unmodeled waters were chosen so as not to overlap with any of the other atoms in the structure. Simulated experimental data were generated for 1210 structure factors from 8 to 2 Å. A Gaussian-distributed 'measurement error' of 5% was added to each 'observed' amplitude.

The native model structure used for all cases was obtained by least-squares restrained refinement using a modified version of *PROLSQ* (Hendrickson, 1985). The known native structure was used as the starting model and all simulated data were employed. Positional and thermal factors for each atom were refined in a total of 100 cycles of refinement. This procedure yielded a model native structure with a standard R factor* ($R_{\rm std}$) of 19.3%, differing from the known native structure by 0.083 Å r.m.s. Variant model structures were obtained by independent refinement and difference refinement, employing the same *PROLSQ* method used for the native.

'Experimental' uncertainties were not used in any refinements. Weighting factors E for independent refinements were made to be simple functions of resolution and were iteratively adjusted so that the standard R factor ($R_{\rm std}$) was independent of resolution, and so that the r.m.s. deviation of bond lengths from ideality was 0.03 Å.

4.1. *Incompleteness of data*

The first test case was designed to investigate the effects of incompleteness in the variant data on the accuracy of refined differences. One known pair of native and variant structures was considered with an r.m.s. difference between them of 0.1 Å. The unmodeled water molecules were added at the same coordinates in the two structures. Varying amounts of randomly selected data, up to 70% of the total, were not considered in refinement of the variant structure.

Fig. 1 illustrates the effects of incompleteness of data on the variant structure determination, showing errors in the coordinates of the variant refined structure and the errors in the coordinate displacements from native to variant structures as a function of the fraction of variant data used in the refinement. For the refinements carried out by independent refinement, $R_{\rm std}$ ranged from 18.4 to 19.4%. In the refinements carried out with difference refinement, $R_{\rm diff}$ ranged from 5.6 to 7.0% and $R_{\rm std}$ ranged from 19.6 to 21.2%. Difference refinement, as expected, yields a slightly higher standard R factor because the quantity being minimized in difference refinement is not as closely related to $R_{\rm std}$ as in independent refinement.

Comparing Figs. 1(*a*) and 1(*b*), it may be seen that the error in the displacements is lower than the absolute error in the variant structure using either method of refinement, so long as the data are highly complete. For independent refinement employing all of the variant data,

* The standard R factor for comparison of calculated and observed amplitudes of the structure factors is $R_{\rm std} = 100\sum_h |F_{o_h} - F_{c_h}| / \sum_h F_{o_h}$. For analogous comparison of calculated and observed differences between native and variant data sets, we define the difference R factor, $R_{\rm diff} = 100\sum_h |F_{\rm diff} - F'_{c_h}| / \sum_h F_{\rm diff}$. For comparing observed differences between native and variant data sets we define the variant R factor, $R_{\rm var} = 100\sum_h |F'_{o_h} - F_{o_h}| / \sum_h (\frac{1}{2})|F'_{o_h} + F_{o_h}|$.

the r.m.s. coordinate error in the variant structure was 0.091 Å, similar to the r.m.s. error in the native coordinates. The error in the displacements from native to variant, however, was only 0.051 Å r.m.s. As expected, the model errors tend to cancel, leading to a low net error in the estimates of changes from native to variant structures. A similar cancellation of errors is obtained using difference refinement.

When a randomly selected subset of the data is excluded in refinement of the variant structure, however, both the overall errors in the variant-structure coordinates and the errors in the displacements from native to variant structures increase quite substantially for independent refinement, while for difference refinement these errors increase only slightly. The r.m.s. errors in the displacements from native to variant structures, for example, increase from 0.051 to 0.139 Å for independent refinement as the percentage of data available is lowered from 100 to 30%. The errors after difference refinement increase only from 0.050 to 0.072 Å over this range. As anticipated from the theoretical treatment, difference refinement yields estimates of displacements that can be accurate even when the variant data are quite incomplete.

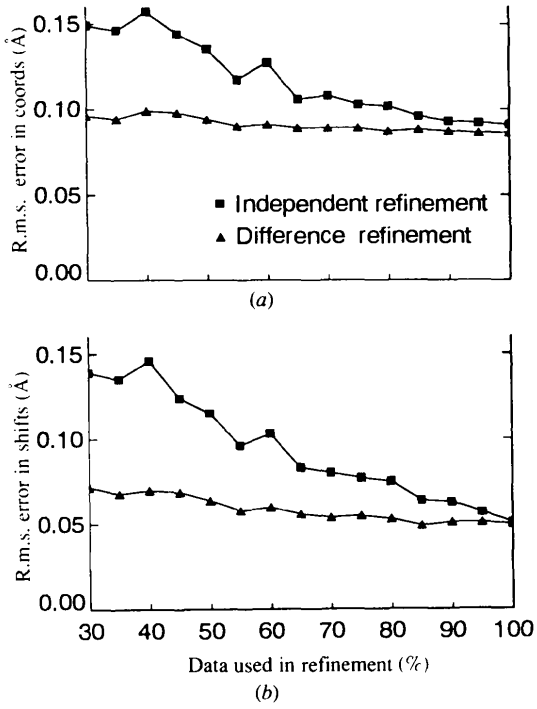

Fig. 1. The effects of incompleteness of data on model quality for independent refinement (squares) and difference refinement (triangles) of a variant structure for a simulated peptide of 51 atoms. The shift between the known native and variant structures was 0.1 Å; two unmodelled water molecules were added at the same positions in native and variant structures. (*a*) The r.m.s . errors in the model variant atomic coordinates. (*b*) The r.m.s. errors in the displacements from model native to model variant structures.

4.2. *Size of modelled changes*

We next examined how different the two modelled structures could be before difference refinement would no longer be useful. Fig. 2 illustrates the effects of increasing the known coordinate shifts from native to variant peptide structures from 0.1 to 0.68 Å r.m.s. while maintaining the water molecules at identical positions. In these refinements, 50% of the data were made available for refinement of the variant structures, so that the leftmost data points in the figure correspond to the data points at 50% completeness in Fig. 1.

Fig. 2(*a*) illustrates the behavior of the variant R factor, $R_{\rm var}$, (which compares native and variant data sets), the $R_{\rm std}$ obtained after independent refinement, and the $R_{\rm diff}$ obtained after difference refinement as functions of known coordinate shift. As the coordinates of the known variant peptide structure are increasingly shifted from those of the native structure, the value of $R_{\rm var}$ increases from 11 to 47%. The $R_{\rm std}$ obtained after independent refinement of the variant structure remained relatively constant, ranging from 18.5 to 19.6%. $R_{\rm diff}$, on the other hand, increases quite substantially over this range, with a value of 6.4% when the coordinate shifts were 0.1 Å r.m.s. to a value of 15.8% when the coordinate shifts were 0.68 Å r.m.s. This is as expected. The phases of the native and variant structure factors must diverge as the coordinate shifts increase and, therefore, the model errors, D_{m_h}, which depend on the phase difference between $\mathbf{F}_{m_h}$ and $\mathbf{\Delta}_{m_h}$, will become less correlated, as illustrated in Fig. 2(*b*). The correlation coefficient between model errors in native and variant structures decreases from 0.93 to 0.56 as the coordinate shifts increase from 0.1 to 0.68 Å r.m.s.

The effects of decreasing the correlation between native and variant modelling errors on the utility of difference refinement is illustrated in Fig. 2(*c*), which shows the r.m.s. error in the coordinate shifts obtained by each refinement method as a function of the true coordinate shifts. Difference refinement yields a lower r.m.s. error in coordinate shifts throughout this range of displacements, but in this example it is most useful when the coordinate shifts are less than about 0.35 Å. This corresponds in this specific idealized case to differences between native and variant amplitudes of structure factor ($R_{\rm var}$) of 30%, and a correlation coefficient between native and variant modelling errors of 0.79.

4.3. *Correlation of modelling errors*

In a real case, the unmodelled parts of the native and variant structures would not be identical. Therefore, we examined how different these parts of the structure could be from native to variant structures before difference refinement would no longer be applicable. A group of data sets was constructed that had identical coordinate shifts for the peptide model of 0.1 Å r.m.s., but in which the unmodelled water molecules were displaced by up

to 1.2 Å from native to variant. The difference between the native and variant peptide structures was fixed at 0.1 Å r.m.s. and the completeness was fixed at 50%. The results of independent and difference refinement for these data sets are illustrated in Fig. 3.

The value of $R_{\rm var}$ increases from 11 to 26% as the coordinates of the water molecules are shifted from 0 to 1.2 Å from native to variant structures (Fig. 3*a*). $R_{\rm std}$ after independent refinement is fairly constant, ranging from 19.0 to 21.2%, while $R_{\rm diff}$ increases from 6.4 to 20.3% over this same range. Fig. 3(*b*) shows that, as expected, the correlation between modelling errors for the native and variant structures decreases substantially as the water molecules shift in position more and more between native and variant structures.

Fig. 3(*c*) compares the r.m.s. error in coordinate shifts from native to variant structures obtained using independent and difference refinement methods. As anticipated, when those parts of the structure not included in the model are quite similar in the native and variant structures, difference refinement yields much lower errors than independent refinement. When the water molecules not included in the modelling shift by more than about 0.5 Å from native to variant structures, however, the correlation between modelling errors in the native and

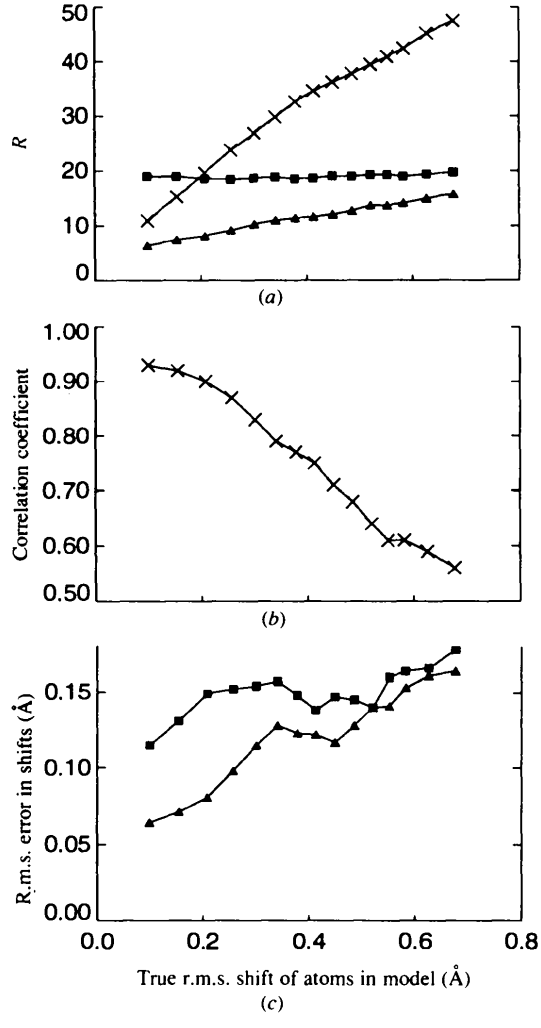

Fig. 2. The effects of increasing coordinate displacement between the known native and variant peptide models. The simulated variant data was 50% complete; two unmodelled water molecules were added at the same positions in native and variant structures. (*a*) The R factor comparing native and variant structure factors (R_{var}, crosses), the standard R factor after independent refinement (R_{std}, squares) and the difference R factor after difference refinement (R_{diff}, triangles). (*b*) The correlation coefficient between modelling errors in the native and variant structures. The modelling error was defined as the difference in amplitude between a 'measured' structure factor and the structure factor calculated from the refined model. (*c*) The r.m.s. errors in the displacements from model native to model variant structures.

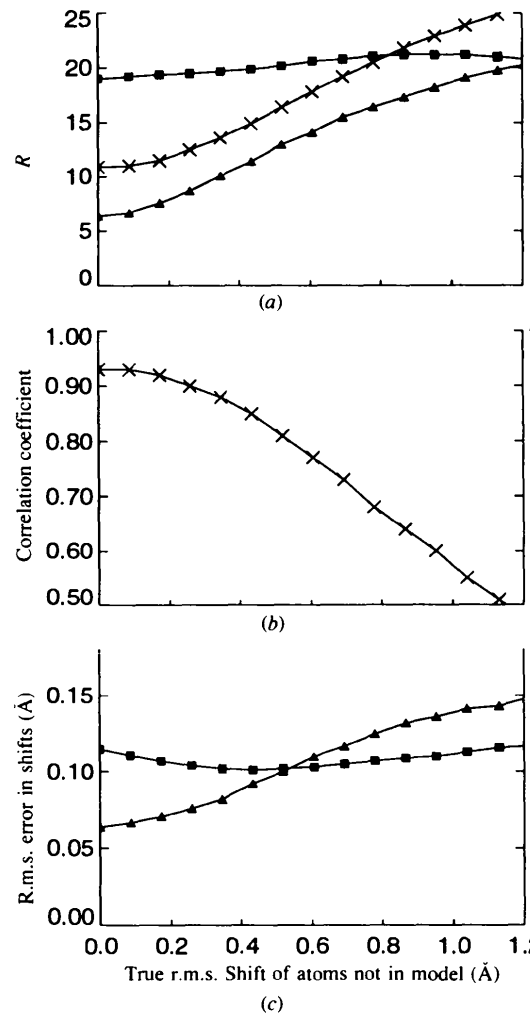

Fig. 3. The effects of decreasing the correlation in modeling errors. Conditions were the same as for Fig. 2 except that the shift between the known native and variant structures was held fixed at 0.1 Å r.m.s. while the two unmodelled water molecules were placed at different locations in the known native and variant structures.

variant refinements decreases to less than 0.80 and difference refinement results in greater errors than does independent refinement.

5. Difference or independent refinement?

The preceding examples have shown that difference refinement can yield substantially lower r.m.s. errors in coordinate shifts than independent refinement if the native and variant structures, especially those parts not included in the modelling, are nearly isomorphous. They do not, however, provide a method of determining whether these conditions are satisfied. Fig. 1 shows that the fewer data used in refinement of the variant structure, the more useful difference refinement will be. Fig. 2 shows that the lower the R factor comparing native and variant amplitudes of structure factors ($R_{\rm var}$), the greater the improvement that can be expected from difference refinement, and Fig. 3 shows that the smaller the coordinate shifts of the parts of the model not included in refinement are in native and variant structures, the greater the utility of difference refinement. This last quantity is not known in practice, however, and there is no particular value of either of the first two of these quantities that can be used as a criterion for deciding when to use difference refinement. For example, looking at Fig. 2, when $R_{\rm diff}$ is 20%, difference refinement is far superior to independent refinement, while at the same value of $R_{\rm diff}$ in Fig. 3 independent refinement yields the lower errors. We have found a simple criterion that works fairly well in deciding which method to use: choose the one that yields the lower r.m.s. coordinate shift. To see why this might be expected to work, we need to consider how these r.m.s. coordinate shifts are related to the 'true' r.m.s. coordinate shifts, and to consider whether there is any bias towards small or large coordinate shifts for either method of refinement.

5.1. *Size of bias in shifts*

The shift in coordinates of a particular atom from native to variant refined structures as obtained by a certain refinement method a, Δx_a, can be thought of as the sum of the true coordinate shift, Δx_0, and an error in the shift ε_a. Assuming for the moment that the errors in the shifts are unbiased and uncorrelated with the true shifts, we may write, averaging over all atoms,

$$\langle \Delta x_a^2 \rangle = \langle \Delta x_0^2 \rangle + \langle \varepsilon_a^2 \rangle, \qquad (8)$$

since $\langle \Delta x_0^2 \rangle$ does not depend on the refinement method employed, this would tend to imply that the method that produces the smaller mean-square shift between native and variant structures (*i.e.*, smaller $\langle \Delta x_a^2 \rangle$) is also the more accurate (smaller $\langle \varepsilon_a^2 \rangle$). However, it is not obvious under what conditions the assumptions lead to this conclusion are justified, particularly for difference refinement which uses information on the native structure to obtain a structure for the variant and might thus have a bias towards smaller coordinate shifts. We checked the validity of these assumptions using our sets of test refinements, for which we know the correct structures. Fig. 4 compares the square roots of the two sides of (8) for all the refinements shown in Figs. 1–3. It is clear that there is indeed a bias towards small shifts for difference refinement, but that this bias is not very large. The r.m.s. shifts that would be expected if difference refinement were not biased towards small shift are only 0.02 to 0.05 Å larger than the r.m.s. shifts actually found; independent refinement has even less bias.

5.2. *Smaller shifts indicate smaller errors*

It should then be possible to estimate whether the errors in difference refinement are greater than or smaller than the errors obtained with independent refinement for a particular case simply by comparing the mean-square coordinate shifts obtained by the two methods. We checked the accuracy of this easy criterion for the test cases examined here by plotting the differences between the mean-square errors in shifts *versus* the differences in mean-square shifts for both methods in all our test cases in Fig. 5. The method of refinement with the smaller mean-square shift is a good estimate of the method with the smaller mean-square error in shift. In only two cases are the differences in mean-square errors in shifts of opposite sign to the differences in mean-square shifts, both cases having very small differences in mean-square errors. As the correlation is not perfect, the better method of refinement would not be chosen in every case, but in those cases where the poorer method was chosen the improvement from the better method would be marginal.

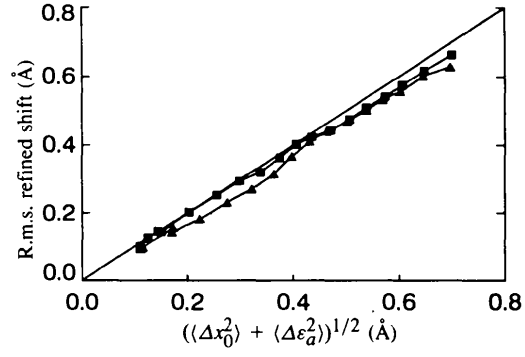

Fig. 4. Comparison of observed coordinate shifts with those expected from an unbiased model as in equation (8). For each of the refinements shown in Figs. 1–3, the r.m.s. value of the coordinate shift from refined native to variant structures is plotted as a function of the square root of the sum of the mean-square true coordinate shift and the mean-square error in the refined coordinate shift. Squares are for independent refinement, triangles are for difference refinement.

6. Application to gene V protein data

As a second comparison, independent and difference refinement were applied to actual data collected on bacteriophage f1 gene V protein (Skinner *et al.*, 1994) and a mutant (I47V, or Ile47→Val). The data were of high quality. A total of 7920 reflections or 97% of the possible from 1.8 to 8 Å were measured with $F > 2\sigma$ for the native gene V protein, and the R factor for merging of equivalent reflections in this data set was 6.1% (on intensity) with a multiplicity of 6.3. For the I47V variant, 93% of the possible reflections in the same range of resolution were measured with $F > 2\sigma$, and the merging R factor on intensities was 4.1% with a multiplicity of 6.0. This I47V variant is well suited for difference refinement because the differences between measured native and variant amplitudes of structure factors ($R_{\rm var}$) were only 10.9%, indicating that the two structures are highly isomorphous.

The refinement procedure used in this test was identical to that used in the tests with model data. The native structure was refined with *PROLSQ* using all the observed data with $F > 2\sigma$ from 1.8 to 8 Å, adjusting the weighting factors $E_{\rm diff}$ in the difference refinement functional so as to yield an r.m.s. deviation of bond lengths from ideality of 0.03 Å as before. The experimental uncertainties were not included, although in practice it would be better to do so if they are thought to be good estimates of the actual errors in measurement. The starting model used was a refined model of gene V protein obtained using *X-PLOR* on the same data considered here. The standard R factor after refinement of the native structure was 21.1% for the data from 1.8 to 8 Å (20.0% for the data from 1.8 to 5 Å).

The mutant structure was refined in the same way, using the native structure with Ile47 replaced by Val as the starting point for refinement against all data from 1.8 to 8 Å with $F > 2\sigma$. $R_{\rm std}$ after independent refinement of the I47V structure was 20.8%, and $R_{\rm diff}$ after difference refinement was 5.8%.

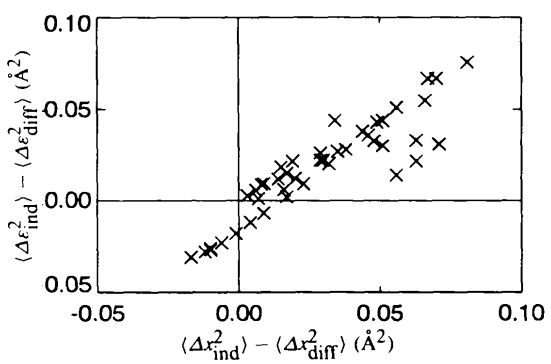

Fig. 5. Comparison of differences between mean-square coordinate shifts for independent and difference refinement with differences in mean-square errors in these coordinate shifts. For each of the refinements shown in Figs. 1–3, the difference between the mean-square values of the errors in coordinate shifts from refined native to variant structures for independent and difference refinement is plotted as a function of the difference between the mean-square values of the coordinate shifts.

The first row of data in Table 1 lists the r.m.s. coordinate shifts from the native to the I47V structures refined using all the data with independent refinement and with difference refinement. Even using all of the available data in refinement, the r.m.s. coordinate shifts for independent refinement (0.191 Å for all protein atoms) are higher than those for difference refinement (0.108 Å), which suggests that the structure obtained by difference refinement is considerably more accurate than the one obtained by independent refinement.

We next tested what would have occurred if either only a randomly chosen 50% or the low-resolution 50% (corresponding to resolutions below 2.3 Å) of the data were available for the I47V variant refinement. The second and third rows in Table 1 list the r.m.s. coordinate shifts from native to the I47V variant in each of these cases with either independent or difference refinement. Using either a randomly chosen or low-resolution 50% of the data, independent refinement leads to considerably higher r.m.s. shifts (0.279 or 0.297 Å for all protein atoms) than were obtained using all the data (0.191 Å). Because the larger shifts were obtained using fewer data, we can conclude that these larger shifts correspond to increased errors in the coordinates. In contrast, using difference refinement the r.m.s. shifts obtained using 50% of the data (0.119 or 0.116 Å) were nearly identical to those obtained using all the data (0.108 Å).

The I47V structures obtained using all or 50% of the data and independent or difference refinement are compared in Table 2 with the two structures obtained using all the data. The first three columns of Table 2 compare each structure with the structure obtained using independent refinement and all available data. It may be seen that when the lowest resolution 50% of the data are used with independent refinement, the resulting structure differs from that obtained using the same method but all the data by an r.m.s. distance of 0.276 Å for all protein atoms.

The I47V structure obtained using difference refinement and all available data differed from the structure obtained with independent refinement by an r.m.s. distance of 0.146 Å for all protein atoms. Quite remarkably, when only 50% of the data is used in difference refinement, the resulting structure still only differs from the structure obtained by independent refinement and using all the data by an r.m.s. distance of 0.151 Å.

The lower right portion of Table 2 serves to emphasize how independent the results obtained with difference refinement are on the fraction of data used in refinement of the I47V variant. Using 50% of the data, either randomly chosen or the lowest resolution data, the resulting structures only differ from the structure obtained using all the available data by an r.m.s. distance of 0.055 Å.

Table 1. *Coordinate shifts (native⟶I47V) obtained by independent or difference refinement using all, a randomly chosen 50%, or the lowest resolution 50% of the I47V data (r.m.s. shifts in Å)*

	Independent refinement			Difference refinement		
	All atoms	Main chain	Side chains	All atoms	Main chain	Side chains
All data	0.191	0.143	0.231	0.108	0.113	0.103
Randomly chosen 50%	0.279	0.179	0.356	0.119	0.117	0.120
Lowest resolution 50%	0.297	0.201	0.373	0.116	0.116	0.117

Table 2. *Differences between I47V structures obtained by independent or difference refinement using all data with those obtained using all, a randomly chosen 50%, or the lowest resolution 50% of the I47V data (r.m.s. differences in Å)*

	Independent refinement using all data			Difference refinement using all data		
	All atoms	Main chain	Side chains	All atoms	Main chain	Side chains
Independent refinement						
All data	—	—	—	0.146	0.078	0.193
Randomly chosen 50%	0.276	0.136	0.371	0.252	0.132	0.336
Lowest resolution 50%	0.237	0.133	0.312	0.270	0.156	0.353
Difference refinement						
All data	0.146	0.078	0.193	—	—	—
Randomly chosen 50%	0.151	0.086	0.198	0.055	0.042	0.065
Lowest resolution 50%	0.151	0.084	0.199	0.055	0.045	0.064

7. Concluding remarks

There are many cases in the refinement of macromolecular structures where the displacements between a pair of isomorphous structures is of interest, and the quality, completeness, or resolution of one of the data sets is lower than that of the other. This is often the case, for example, in the refinement of structures of proteins and corresponding mutants or complexes with small molecules (Eriksson *et al.*, 1993) and is generally the case in experiments involving Laue diffraction (Singer *et al.*, 1993). The idealized and real cases examined here indicate that difference refinement (Fermi *et al.*, 1982) offers substantial improvement in errors compared with independent refinement of the two structures.

The basis for difference refinement is that there are features of crystals of macromolecules that are not included in models for these crystal structures and, most importantly, that these features are often very similar in native and variant structures. Because of this, the differences between structures can be modeled more accurately than the structures themselves. By refining a variant model in order to make the calculated differences between native and variant amplitudes of structure factors match the observed differences as closely as possible, the errors due to the inadequacy of the models tend to cancel and more accurate differences between the structures can be obtained. Although this can be accomplished to some degree by independent refinement of the structures, the cancellation of errors is maximized by difference refinement.

One important issue in difference refinement is the extent of non-isomorphism between the two structures beyond which difference refinement is no longer useful. The key factor in whether difference refinement will be useful is whether the model bias errors are very similar for the native and variant structures. Another way to view this is to note that the two structures can be very different, as long as those parts of the structures that are not included in the modelling are very similar for native and variant structures. In this case, the model bias errors will be correlated for the two structures and will tend to cancel in difference refinement. Our analysis of model data indicates that the accuracy of independent refinement and difference refinement for an individual case can be quantitatively compared simply by comparing the r.m.s. coordinate shifts from native to variant, where the method that yields the lower r.m.s. shift is the method with the lower r.m.s. error in those shifts. This means that in cases where it is not obvious whether difference refinement would be applicable, it is a relatively simple matter to determine its utility by carrying out the refinement both by independent and difference refinement and choosing the method that yields the smaller r.m.s. coordinate shifts.

The authors gratefully acknowledge generous grants from the NIH and from the LDRD program of Los Alamos National Laboratory. We thank Randy Read for pointing out an important reference (Fermi *et al.*, 1982) to us.

References

BRÜNGER, A. T., KARPLUS, M. & PETSKO, G. A. (1989). *Acta Cryst.* **A45**, 50–61.

DAOPIN, S., DAVIES, D. R., SCHLUNEGGER, M. P. & GRÜTTER, M. G. (1994). *Acta Cryst.* D**50**, 85–92.

ERIKSSON, A. E., BAASE, W. A. & MATTHEWS, B. W. (1993). *J. Mol. Biol.* **229**, 747–769.

FERMI, G., PERUTZ, M. F., DICKINSON, L. C. & CHIEN, J. C. W. (1982). *J. Mol. Biol.* **155**, 495–505.

GROS, P., VAN GUNSTEREN, W. F. & HOL, W. G. J. (1990). *Science*, **249**, 1149–1152.

HENDRICKSON, W. A. (1985). *Methods Enzymol.* **115**, 252–270.

HUANG, Y., PAGNIER, J., MAGNE, P., BAKLOUTI, F., KISTER, J., DELAUNAY, J., POYART, C., FERMI, G. & PERUTZ, M. F. (1990). *Biochemistry*, **29**, 7020–7023.

JASKOLSKI, M., TOMASSELLI, A. G., SAWYER, T. K., STAPLES, D. G., HEINRIKSON, R. L., SCHNEIDER, J., KENT, S. B. H. & WLODAWER, A. (1991). *Biochemistry*, **30**, 1600–1609.

JENSEN, L. H. (1985). *Methods Enzymol.* **115**, 227–237.

KURIYAN, J., PETSKO, G. A., LEVY, R. M. & KARPLUS, M. (1986). *J. Mol. Biol.* **190**, 227–254.

LUISI, B. & SHIBAYAMA, N. (1989). *J. Mol. Biol.* **206**, 723–736.

MATTHEWS, B. W. (1993). *Ann. Rev. of Biochem.* **62**, 139–160.

NAGAI, K., LUISI, B., SHIH, D., MIYAZAKI, G., IMAI, K., POYART, C., DE YOUNG, A., KWIATKOWSKY, L., NOBLE, R. W., LIN, S.-H. & YU, N.-T. (1987). *Nature (London)*, **329**, 858–860.

PERUTZ, M. F., FERMI, G., ABRAHAM, D. J., POYART, C. & BURSAUX, E. (1986). *J. Am. Chem. Soc.* **108**, 1064–1078.

SINGER, P. T., SMALÅS, A., CARTY, R. P., MANGEL, W. F. & SWEET, R. M. (1993). *Science*, **259**, 669–673.

SKINNER, M. M., ZHANG, H., LESCHNITZER, D. H., BELLAMY, H., SWEET, R. M., GRAY, C. W., KONINGS, R. N. H., WANG, A. H.-J. & TERWILLIGER, T. C. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 2071–2075.

SUSSMAN, J. L. (1985). *Methods Enzymol.* **115**, 271–303.