Acta Cryst. (1971). B27, 1414

The Difference Fourier Technique in Protein Crystallography: Errors and their Treatment

BY R. HENDERSON* AND J. K. MOFFAT†

Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge, England

(*Received* 23 *July* 1970)

An expression is derived for the error which arises from the use of the difference Fourier technique. Errors due to the use of inaccurate values of ΔF and phase are considered in addition to those due to the difference Fourier technique itself. It is shown that a difference Fourier map is able to detect much smaller features of electron density than those revealed by a normal Fourier map with the same phases. The experimental conditions required to obtain difference maps in which the error level is minimized are outlined.

Introduction

Many types of difference synthesis have been used in the derivation and refinement of crystal structures. With structures in which individual atoms are resolved, difference syntheses which compare the observed structure amplitudes with those derived from a trial structure are often employed. With proteins and other large structures, the resolution attained is less than that required to position individual atoms and the phases of the structure factors are usually obtained by the multiple isomorphous replacement method. Because of the method of phasing and the difficulty in refining protein structures, a type of difference synthesis, which we refer to here as the difference Fourier synthesis, has proved to be a very valuable method of determining the structures of isomorphous derivatives of a protein whose crystal structure is known. The power of the method lies in its use of the phases of only one structure to investigate a large number of similar structures, for which only the amplitudes of the structure factors need be measured. It has been successfully applied to such problems as ligand binding in myoglobin (Stryer, Kendrew & Watson, 1964; Nobbs, Watson & Kendrew, 1966) and haemoglobin (Perutz & Mathews, 1966), substrate-protein interactions in lysozyme (Johnson & Phillips, 1965; Johnson, 1967), α-chymotrypsin (Sigler, Jeffery, Matthews & Blow, 1966; Steitz, Henderson & Blow, 1969), carboxypeptidase Aα (Steitz, Ludwig, Quiocho & Lipscomb, 1967; Lipscomb, Hartsuck, Reeke, Quiocho, Bethge, Ludwig, Steitz, Muirhead & Coppola, 1968) and ribonuclease S (Wyckoff, Hardman, Allewell, Inagami, Johnson & Richards, 1967), chemical modification of certain residues of haemoglobin (Moffat, 1971) and lysozyme (Blake, 1967) and in the derivation of the positions of heavy atoms for use in phase determination.

In most cases the difference maps are used directly without further checking or refinement. Since the deduction of the mechanism of action of proteins depends critically on the interpretation of these difference Fouriers, it is important to estimate the statistical significance of each feature. In particular, certain small features which are dismissed as noise in an over cautious interpretation, may contain significant structural information. The derivation of the statistical significance requires a knowledge both of the error level and its distribution in the unit cell, and the effect of errors on the peak heights of features. These are determined here by considering the origin and nature of the various errors in a difference map. We show that the difference Fourier technique is extremely sensitive to small changes in electron density and is capable of revealing more subtle features of the electron density than those apparent in a normal Fourier map obtained with the same phases. The experimental approach which leads to the optimum use of the difference Fourier technique is outlined and the limitations of the method are discussed.

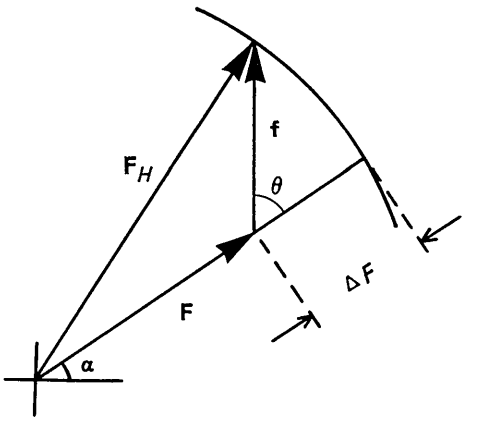

Fig. 1. Argand diagram showing the relationship between the parent and derivative structure factors. The structure factor f represents the difference structure.

* Present address: Department of Chemistry, Yale University, 225 Prospect Street, New Haven, Conn. 06520, U.S.A.

† Present address: Section of Biochemistry and Molecular Biology, Cornell University, Ithaca, New York 14850, U.S.A.

The difference Fourier synthesis

The difference Fourier technique is used here to refer to the computation of a Fourier synthesis whose coefficients are $m(|\mathbf{F}_H| - |\mathbf{F}|) \exp [i\alpha]$ where $|\mathbf{F}_H|$ and $|\mathbf{F}|$ are the measured structure amplitudes of two similar non-centrosymmetric structures, α is the phase of $\mathbf{F}$, and m is a weighting factor (Dickerson, Kendrew & Strandberg, 1961) used to weight down terms where α is not well determined. Fig. 1 shows the relationship between $\mathbf{f} = \mathbf{F}_H - \mathbf{F}$, $\Delta F = |\mathbf{F}_H| - |\mathbf{F}|$ and the two structure factors.

A Fourier synthesis using $\mathbf{f}$ as coefficients would give the desired difference in electron density between the two structures directly. Since $\mathbf{f}$ is not accessible experimentally, the difference Fourier synthesis calculates an approximation to this difference in electron density.

When ΔF and $|\mathbf{f}|$ are small compared with $|\mathbf{F}|$, the difference Fourier coefficient $m\Delta F \exp [i\alpha]$ represents the vector component of $\mathbf{f}$ in the direction of $\mathbf{F}$ but includes no component perpendicular to this direction. It has been pointed out previously (Stryer *et al.*, 1964; Hoard, 1967) that when $|\mathbf{f}|$ is not small compared with $|\mathbf{F}|$, then this relationship no longer holds. The phases of $\mathbf{F}$ and $\mathbf{F}_H$ are different and the measured value of ΔF is no longer the component of $\mathbf{f}$ in the direction of $\mathbf{F}$.

For the heavy atom method, which is formally identical, it has been shown (Woolfson, 1956; Sim, 1959, 1960) that reflexions for which $|\mathbf{f}| \ll |\mathbf{F}|$ does not hold, should be treated by reducing the magnitude of $|\mathbf{F}_H|$ in the difference synthesis in accordance with its probability of having a different phase from $\mathbf{F}$. This can lead to considerable improvement in the resulting Fourier map, particularly when $|\mathbf{f}|$ and $|\mathbf{F}|$ are comparable in magnitude for a significant fraction of the terms. Such a situation may arise, for example, if the total scattering from a number of heavy atom sites is comparable with that from the protein itself. In more routine protein crystallographic work we have found that the number of reflexions which are significantly affected by the use of such a weighting scheme seldom exceeds 10 to 15 per cent of the total. Moreover, the phases of these reflexions (which generally have small and inaccurate values of $|\mathbf{F}|$) are poorly determined by the isomorphous replacement method. In practice, therefore, it does not much matter whether these terms are omitted from or included in the Fourier synthesis.

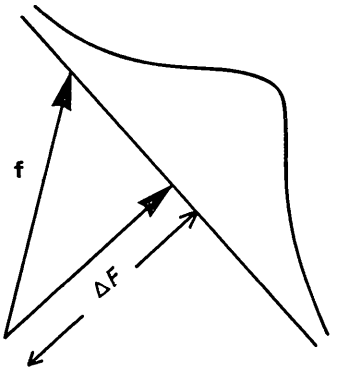

Fig. 2. Argand diagram showing the probability distribution of $\mathbf{f}$ for a given observed ΔF in the absence of experimental errors.

The condition that $|\mathbf{f}| \ll |\mathbf{F}|$ applies to the remainder of the reflexions. For simplicity, we have chosen to deal with difference Fourier maps where this is true for all reflexions. In practice, this includes all the difference maps referred to in the introductory paragraph and those named in Table 1.

Origin and magnitude of errors

We define the 'true' difference structure as the Fourier transform of the complex structure factors, $\mathbf{f}$. Thus, we do not consider errors due to differences in reciprocal lattice parameters or series termination.

The three sources of error in difference Fourier maps with which we are concerned are:

(i) use of the difference Fourier coefficients $m\Delta F \exp [i\alpha]$ instead of the true but unobservable structure factors, $\mathbf{f}$;

(ii) experimental errors in the ΔF values; and

(iii) experimental errors in the phases, α, of the parent structure factors.

To assess the effects of these three sources of error on the electron-density map, we calculate the resulting root-mean-square (r.m.s.) error in electron density averaged over the complete unit cell. Since the maps take the form of continuous functions of varying shape with no possibility of resolving individual atoms, this seems the only sensible way of estimating error. The approach is similar to that of Cruickshank (1949) and Blow & Crick (1959). The error in the map is defined as the r.m.s. difference in electron density between the 'true' structure and that obtained in practice. This type of error is not the same as truly random error defined as being independent of position in the unit cell, since it includes a contribution due to systematic reduction in the peak heights of all features, as will become apparent. We calculate first the overall error in electron density.

(i) *Intrinsic error in the difference Fourier method*

Consider the case where $|\mathbf{F}|$ and $|\mathbf{F}_H|$ are accurately known and α, the phase of $\mathbf{F}$, is also well determined. An Argand diagram of the vector of length ΔF in the direction of $\mathbf{F}$ is shown in Fig. 2. Since the true value of $\mathbf{f}$ is not known, we must consider the probability of different values of $\mathbf{f}$. Such a probability distribution is also shown in Fig. 2, possible values of $\mathbf{f}$ lying on a line at right angles to $\mathbf{F}$. Estimation of the error then requires calculation of the radius of gyration, r, of this probability distribution (*cf.* Blow & Crick, 1959).

If we assume that, over all reflexions, the phases of $\mathbf{f}$ and $\mathbf{F}$ are uncorrelated so that the probability distri-

butions of the components of **f** parallel and perpendicular to **F** are the same (and this assumption is valid for $|\mathbf{f}| \ll |\mathbf{F}|$), then since ΔF is the parallel component and r is dependent only on the perpendicular component,

$$\sum\sum\sum r^2 = \sum\sum\sum \Delta F^2 \qquad (1a)$$

summed over all reflexions.

$$\therefore \langle r^2 \rangle = \langle \Delta F^2 \rangle \, . \qquad (1b)^*$$

Hence, using the derivation of equation (26) in Blow & Crick (1959), the mean-square error in electron density in the difference Fourier synthesis is given by

$$\langle \Delta\varrho^2 \rangle = \frac{2}{V^2} \sum_{h=0}^{+\infty} \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} r^2 \qquad (2)^*$$

which, using equation (1a), becomes

$$\langle \Delta\varrho^2 \rangle = \frac{2}{V^2} \sum\sum\sum \Delta F^2 \qquad (3)$$

where the limits of summation here, and subsequently, are those indicated in equation (2).

(ii) *Error in the measurement of* ΔF

We now consider the effect of experimental inaccuracy in ΔF, giving rise to an r.m.s. error, δ. Fig. 3 shows the probability distribution of **f** including the uncertainty in ΔF. Since the probability distribution is now made up of two independent components at right angles,

$$\langle r^2 \rangle = \langle \Delta F^2 \rangle + \langle \delta^2 \rangle \qquad (4)$$

and the overall mean-square error in electron density becomes

$$\langle \Delta\varrho^2 \rangle = \frac{2}{V^2} \sum\sum\sum (\Delta F^2 + \delta^2). \qquad (5)$$

(iii) *Error in the phase* α *of the structure factor* **F** *of the parent structure*

The probability distribution of **f** will now be as above but folded with the probability distribution of the phase, α. For example, Fig. 4 illustrates a case where α has a sharply bimodal distribution. As shown by Blow & Crick (1959), the use of the centroid phase will give the overall least-square error in the synthesis, where m is the centroid weight and α the centroid phase. Again, to estimate the magnitude of the error we must calculate the radius of gyration of this probability map.

Let the probability of the true phase of **F** differing by $\Delta\varphi_i$ from α be P_i. The definition of m, the centroid weight or figure of merit is

$$m = \frac{\sum_i P_i \cos \Delta\varphi_i}{\sum_i P_i} \qquad (6)$$

* The symbols $\langle \Delta F^2 \rangle$ and $\langle \Delta\varrho^2 \rangle$ are used as a convention to denote $\langle (\Delta F)^2 \rangle$ and $\langle (\Delta\varrho)^2 \rangle$, the mean-square values of ΔF and $\Delta\varrho$.

where the summation is over all possible phase angles, $\alpha + \Delta\varphi_i$ (Dickerson *et al.*, 1961).

In this case the radius of gyration r_{tot} is obtained by considering the contribution to r_{tot} of distributions with all possible phases. The contribution of the distribution with phase $\alpha + \Delta\varphi_i$, to the radius of gyration is

$$\Delta r_{\text{tot}}^2 = P_i \, (r^2 + \varepsilon_i^2) \qquad (7)$$

where r is the radius of gyration of the distribution with phase $(\alpha + \Delta\varphi_i)$ about its centroid and ε_i is the length of the vector joining the centroid of the whole distribution to the structure factor with phase $(\alpha + \Delta\varphi_i)$. Summing over all possible phases,

$$r_{\text{tot}}^2 = \frac{\sum_i P_i \, r^2}{\sum_i P_i} + \frac{\sum_i P_i \, \varepsilon_i^2}{\sum_i P_i} \qquad (8)$$

$$= r^2 + \Delta F^2 \, (1 - m^2) \qquad (9)$$

since in the first term of equation (8), r^2 is independent of $\Delta\varphi_i$, and the second term is identical to the expression (16) derived by Dickerson *et al.* (1961).

From equation (4),

$$\langle r^2 \rangle = \langle \Delta F^2 \rangle + \langle \delta^2 \rangle$$

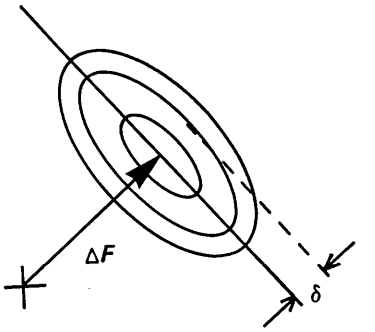

Fig. 3. Argand diagram showing the probability distribution of **f** for a given observed ΔF including an r.m.s. error, δ, in the observation.

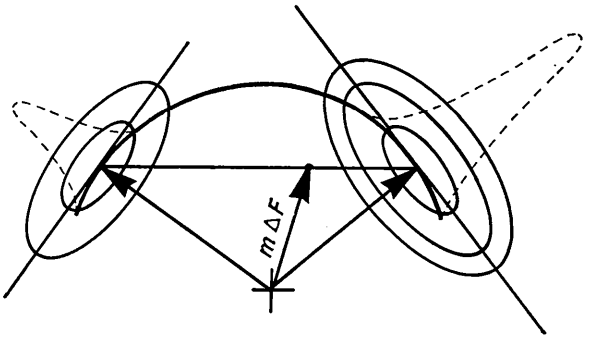

Fig. 4. Argand diagram showing the probability distribution of **f** for a given observed ΔF including errors both in the magnitude of ΔF and in the phase of the parent structure factor. The dotted line is the phase probability distribution of the parent structure factor.

so that the overall mean square error in electron density becomes

$$\langle \Delta\varrho^2 \rangle = \frac{2}{V^2} \sum \sum \sum \{\Delta F^2(2-m^2)+\delta^2\}. \quad (10)$$

This final expression for the total mean-square error in the difference synthesis is closely related to the expression for the error in a normal Fourier map obtained by the method of isomorphous replacement, reproduced in equation (11) (Blow & Crick, 1959; Dickerson *et al.*, 1961).

$$\langle \Delta\varrho^2 \rangle = \frac{2}{V^2} \sum \sum \sum \{|\mathbf{F}|^2(1-m^2)+\delta^2\}. \quad (11)$$

It is important to note that the mean-square errors given by equations (10) and (11) strictly refer to the mean-square differences in electron density between the perfect crystal structures (that is, determined using perfect phases and amplitudes) and those obtained using experimentally measured phases and amplitudes. They do not necessarily refer to the noise level to be expected in featureless regions of the difference Fourier and Fourier syntheses respectively. For instance, if all figures of merit were zero (phases undetermined), all syntheses would be zero everywhere, but would have large mean-square errors as given by equations (10) and (11).

Height of peaks

A second problem remains: what is the height of the signal peak? For instance, if $m=1$ (perfect phases) or if $m=0$ (random phases), then the mean-square error differs only by a factor of two yet in one case the map is good and in the other it does not exist. Clearly, it is useful to know the height of the peak relative to its theoretical height in an accurate Fourier or difference Fourier synthesis with perfect phases and amplitudes.

In the case of a Fourier synthesis with less than perfect phases, any peaks will be smaller than those obtained in a similar synthesis using perfect phases. If

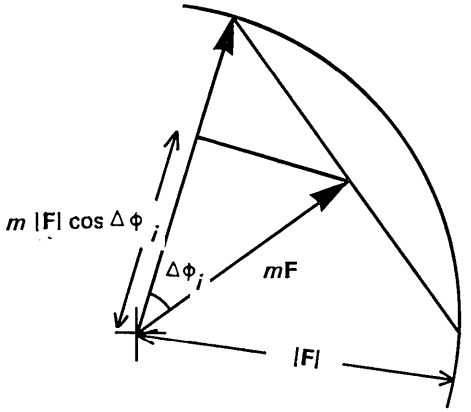

Fig. 5. Argand diagram showing the contribution of the vector $m\mathbf{F}$ to the structure factor whose phase is separated from it by an angle $\Delta\varphi_i$.

m is assumed to be the same for all reflexions, it is possible to predict how the peak height will vary with m. The contribution of the structure factor $m|\mathbf{F}|\exp[i\alpha]$ to the unknown correct structure can be estimated by taking the mean of its contribution to all possible structures. The contribution of $m|\mathbf{F}|\exp[i\alpha]$ to the structure whose phase is separated from it by $\Delta\varphi_i$ has an amplitude $m|\mathbf{F}|\cos\Delta\varphi_i$ (Fig. 5).

Hence, the mean contribution of $m|\mathbf{F}|\exp[i\alpha]$ to the true structure is given by

$$\frac{\sum_i P_i\, m|\mathbf{F}|\cos\Delta\varphi_i}{\sum_i P_i} \quad (12)$$

$$= m|\mathbf{F}| \frac{\sum_i P_i \cos\Delta\varphi_i}{\sum_i P_i}$$

$$= m^2|\mathbf{F}| \quad (13)$$

using equation (6).

The structure synthesized from the Fourier coefficients $m|\mathbf{F}|\exp[i\alpha]$ will therefore have a peak height of m^2 relative to the true structure. In practice, however, m is not the same for all reflexions. Apart from fluctuations from one reflexion to another, its average value, $\langle m \rangle$, is a complicated function of $\sin^2\theta/\lambda^2$ and $|\mathbf{F}|$ depending on the distribution of the experimental errors (Blow & Crick, 1959). However, the above approximation gives some feeling for the effect of variation of $\langle m \rangle$ on the peak height of features. If an accurate estimate of peak height is required, a numerical summation of the Fourier components with the appropriate distribution of values of $\langle m \rangle$ would be necessary.

It has also been shown (Luzzati, 1953) that the peak heights of atoms (in a non-centrosymmetric Fourier synthesis) not included in the phasing depends on the number of atoms excluded, but when that number is small, the peak heights of these atoms are reduced by one-half compared with atoms included in the phasing. This is equivalent to showing that peaks in a difference Fourier synthesis have half their theoretical height, and indeed, this can easily be shown by a similar argument to that used to derive equation (13). If the true difference vector is $\mathbf{f}$, then the measured value of ΔF will be $|\mathbf{f}|\cos\theta$, where θ is the angle between $\mathbf{f}$ and $\mathbf{F}$ (Fig. 1). The contribution of $|\mathbf{f}|\cos\theta$ in the direction of $\mathbf{f}$ is $|\mathbf{f}|\cos^2\theta$. Over all reflexions the mean contribution will therefore be $\langle\cos^2\theta\rangle=\frac{1}{2}$ since θ can take all values between 0 and 2π with equal probability (again provided $|\mathbf{f}| \ll |\mathbf{F}|$).

It follows that, since the peak height reductions due to phasing inaccuracies and to the difference Fourier approximation are independent, the peak heights of all features in the difference map will be roughly $\frac{1}{2}\langle m^2 \rangle$ relative to their theoretical values.

Thus in any real difference Fourier map the height and integrated electron content of all features will be

less than in a perfect map. The mean-square electron density difference between the real and the perfect map given by equation (10) thus arises from a contribution which is everywhere random, that is, independent of position in the unit cell, and a contribution which is due to the systematic reduction in the height of the peaks. Equation (10) therefore overestimates the truly random errors in the map, which may be judged for instance by examining featureless regions. This can also be seen from an Argand diagram such as that shown in Fig. 2 or Fig. 4. The error vector between the 'true' structure factor **f** and the component ΔF which is actually used to compute the map can be resolved into components parallel and perpendicular to **f**. The parallel component systematically reduces peak heights whereas the perpendicular component is truly random.

Equation (10) can be improved by calculating the mean-square electron density difference between half the 'true' structure, that is, the transform of the structure factors **f**/2 (the best that can be done in practice) and the experimentally derived structure. Following similar arguments to those above,

$$\langle \Delta\varrho^2 \rangle = \frac{2}{V^2} \sum \sum \sum \frac{\{\Delta F^2(2-m^2)+\delta^2\}}{2}$$

$$= \frac{1}{V^2} \sum \sum \sum \{\Delta F^2(2-m^2)+\delta^2\} . \qquad (14)$$

It is this formula which provides the best estimate of noise level in a difference map. However, although equation (14) is exact, it will still overestimate the random error when the average figure of merit is low.

It is worth emphasizing briefly two of the points discussed above which are also relevant to the interpretation of Fourier maps of proteins obtained by the method of multiple isomorphous replacement. First, if equation (11) is used to estimate the error in a map, the resulting figure does not represent random error except when the average figure of merit is near unity. It is worth while bearing this in mind when an assessment of the significance of any feature is made. Secondly, the peak heights of features in the map will be approximately $\langle m^2 \rangle$ times their 'true' value, depending somewhat on the shape of the feature and the distribution of $\langle m \rangle$ with $\sin^2\theta/\lambda^2$ and $|\mathbf{F}|$. This could be a useful check on the true quality of any Fourier map.

Comparison with experimental difference maps

To check whether equation (14) accurately represents the error level in a real difference Fourier map, five high resolution difference maps obtained from derivatives of crystalline α-chymotrypsin and four from haemoglobin were analysed in detail. Table 1 gives a summary of the relevant data. Comparison of the error predicted by equation (14) with the r.m.s. electron density in featureless regions of the map gives remarkably good agreement.

There are several reasons why the agreement is not exact. We have already demonstrated that equation (14) always overestimates the truly random error particularly when the phases are poorly determined. This would mean that the r.m.s. electron density in featureless regions of the map should always be slightly less than that given by equation (14). In the maps shown in Table 1, however, the r.m.s. density appears to be slightly greater than that given by equation (14). This could be due to the presence of real features which are below the noise level. Indeed, this is the hazard of using an apparently featureless region of a map to

Table 1. *Analysis of some difference Fourier maps for derivatives of α-chymotrypsin and haemoglobin*

Equation (14) has been used to calculate $\langle \Delta\varrho^2 \rangle^{1/2}$ for the difference maps and the formula of Dickerson *et al.*, (1961) for the parent maps. Observed r.m.s. errors were obtained by calculating the r.m.s. electron density in featureless regions of each map. All electron densities are in e.Å^{-3}. The columns headed (s.d.) simply give the electron densities as a multiple of the calculated r.m.s. error $\langle \Delta\varrho^2 \rangle^{1/2}$. Thus, the fifth column gives the ratio of the observed to the calculated error level.

Also shown for comparison are the error levels in the corresponding maps of the native proteins. In the nine difference maps listed here, the error level is from 4 to 11 times lower than in the corresponding parent map. Taking into account the factor of ½ in the relative peak heights of similar features, these difference maps have a signal-to-noise ratio which is from 2 to 5½ times better than in the corresponding maps of the native proteins.

The nominal resolution of the chymotrypsin maps was 2·5 Å and that of the haemoglobin maps, 3·5 Å. The abbreviations used to name the maps are the same as those used previously (Steitz *et al.*, 1969; Moffat, 1971).

Map	$\langle \Delta F_{obs} \rangle$ e.	$\langle \delta \rangle$ e.	Calculated $\langle \Delta\varrho^2 \rangle^{1/2}$ e.Å^{-3}	Observed r.m.s. error e.Å^{-3}	s.d.	Observed highest noise e.Å^{-3}	s.d.	Observed highest peak e.Å^{-3}	s.d.
tosyl-native CHT	71	33	0·044	0·048	1·1	0·14	3·2	0·67	15·2
diox-native CHT	49	29	0·030	0·031	1·0	0·10	3·2	0·20	6·5
f-L-tryp-native CHT	100	36	0·060	0·065	1·1	0·22	3·6	0·39	6·5
f-L-phe-native CHT	90	31	0·050	0·059	1·2	0·16	3·1	0·29	5·8
IA-native CHT	50	29	0·029	0·030	1·0	0·10	3·3	0·29	10·0
Native chymotrypsin	—	—	0·34	—	—	—	—	—	—
BME-oxyHb	57	—	0·013	0·016	1·2	0·06	3·4	0·11	6·8
AO-oxyHb	44	—	0·012	0·015	1·2	0·05	3·2	0·11	7·3
SPIN-oxyHb	40	—	0·011	0·013	1·2	0·04	3·1	0 12	9·2
CPA-oxyHb	100	—	0·023	0·029	1·2	0·11	3·8	0·14	4·9
Native oxyhaemoglobin	—	—	0·12	—	—	—		—	—

estimate noise level. If there are no featureless regions then no exact estimate can be made. Moreover, it is easier to calculate an expression such as equation (14) routinely for each map than it is to select relatively featureless areas from which to estimate the error level. The agreement may also be made worse if the average figure of merit, used to calculate the error level in equation (14), does not reflect the true accuracy of the phasing procedure. It is well known that the average figure of merit in a protein structure determination is a poor guide to the quality of the Fourier map since it depends to a large extent on estimates of the accuracy of the structure factor measurements and on heavy atom temperature factors. This may be another reason for the slightly low calculated error, particularly for the maps of haemoglobin in which the average figure of merit was quite high (Perutz, Muirhead, Cox, Goaman, Mathews, McGandy & Webb, 1968).

Table 1 also lists the size of the highest noise peaks in the maps. Noise peaks were defined as peaks in the map which could not be explained in a reasonable stereochemical manner. In most cases, they were isolated peaks which occurred in the centre of the protein molecule. The heights of these noise peaks suggest that for high resolution difference maps, any features of greater than 3·5 standard deviations should be regarded as real, in agreement with the normal statistical practice (Cruickshank, 1949). The final column of Table 1 lists the heights of the highest peak in each map in terms of the number of standard deviations of the electron density. The exact statistical significance of each peak can then be determined.

The above comparison shows that equation (14) is sufficiently accurate when used on real data with a range of different error levels.

Conclusions

We have examined the origin of errors in difference Fourier maps and shown that the terms in the synthesis should have the same weight as the corresponding terms in the parent synthesis. We may now ask how best the difference Fourier technique may be utilized in studying isomorphous structures. Several useful facts follow directly from the formulae.

(1) The method can be used even when the differences ΔF are very small. Even if the r.m.s. value of ΔF is equal to the r.m.s. error $\langle\delta\rangle$ in its measurement, the noise level will only be $\sqrt{2}$ times the level for ideally accurate data. A corollary of this is that accurate measurements of the ΔF values become more important for smaller values of ΔF. However, for typical difference Fourier maps such as those in Table 1, accuracy of measurement of ΔF is not a limiting factor. Table 1 shows the values of $\langle\delta\rangle$ and $\langle\Delta F_{\rm obs}\rangle$ in five typical difference Fourier maps of different derivatives of α-chymotrypsin. For the worst case, that of the dioxan: chymotrypsin derivative where $\langle\delta\rangle$ and $\langle\Delta F\rangle$ are comparable, the error could not be reduced by more than 30 per cent by collecting really accurate values of ΔF. On the other hand if $\langle\Delta F\rangle$ is too large, the signal-to-noise ratio will drop because the approximation assumed at the beginning of this paper ($\Delta F \ll |\mathbf{F}|$) will no longer hold. The peaks will then become less than half their theoretical heights (Luzzati, 1953). It is then an advantage to try to reduce $\langle\Delta F\rangle$, perhaps by using partially occupied crystals, to a value where the condition $\Delta F \ll |\mathbf{F}|$ applies or to use the weighting scheme of Woolfson and Sim. In practice, the compromise necessary between the condition that $\langle\Delta F\rangle > \langle\delta\rangle$ and that $\langle\Delta F\rangle \ll \langle|\mathbf{F}|\rangle$ means that there is an optimum $\langle\Delta F\rangle/\langle|\mathbf{F}|\rangle$ ratio which normally lies within the range of 2 to 20 per cent.

(2) If there is only one site where the structures being compared are different (*e.g.* only one binding site for a small molecule) the occupancy of that site will not affect the signal-to-noise ratio as long as the criteria of (1) are obeyed, since both $\langle\Delta\varrho^2\rangle^{1/2}$ and the peak height are proportional to $\langle\Delta F^2\rangle^{1/2}$ when there is only one site.

(3) If there are several sites of substitution of a given ligand on the parent structure and interest centres on one of them (as often happens in practice), then it is advantageous to try to eliminate those sites which are not under consideration since by contributing to ΔF, they increase the noise level. For instance, in the difference Fourier map of *f*-L-tryptophan: chymotrypsin against native, shown in Table 1 (Steitz *et al.*, 1969), the high noise level is due to the presence of several binding sites in addition to the one of interest at the enzyme's active site. A better map might be obtained either by using a lower concentration of *f*-L-tryptophan to try to eliminate the other sites or by calculating a difference map between crystals soaked in different concentrations of *f*-L-tryptophan such that only the binding site of interest became occupied as the concentration changed.

Finally, we should like to emphasize that the difference Fourier technique is very sensitive, and that the results can be analysed in a completely quantitative manner. Since the error level is proportional to $\langle\Delta F^2\rangle^{1/2}$, normally a small fraction of the average structure factor $\langle|\mathbf{F}|\rangle$ of the parent structure, comparison of equations (14) and (11) shows that a difference Fourier map will normally have a much lower error level than the corresponding Fourier map of the parent structure. Even when the factor of $\frac{1}{2}$ in the relative peak heights of similar features is taken into account, in favorable cases, the difference Fourier map may contain significant features of electron density 5 or 10 times lower than those in the corresponding Fourier map of the parent structure. This is borne out in practice (Table 1; Nobbs, 1966).

We are very grateful to Dr D. M. Blow and Dr J. Greer for helpful discussions. This work was carried out while one of us (R.H.) was a Medical Research Council scholar.

References

BLAKE, C. C. F. (1967). *Proc. Roy. Soc.* B**167**, 435.

BLOW, D. M. & CRICK, F. H. C. (1959). *Acta Cryst.* **12**, 794.

CRUICKSHANK, D. W. J. (1949). *Acta Cryst.* **2**, 65.

DICKERSON, R. E., KENDREW, J. C. & STRANDBERG, B. E. (1961). *Acta Cryst.* **14**, 1188.

HOARD, J. L. (1967). *Structural Chemistry and Molecular Biology*, p. 573. Edited by A. RICH & N. DAVIDSON. San Francisco: Freeman.

JOHNSON, L. N. (1967). *Proc. Roy. Soc.* B**167**, 439.

JOHNSON, L. N. & PHILLIPS, D. C. (1965). *Nature, Lond.* **206**, 761.

LIPSCOMB, W. N., HARTSUCK, J. A., REEKE, G. N., QUIOCHO, F. A., BETHGE, P. H., LUDWIG, M. L., STEITZ, T. A., MUIRHEAD, H. & COPPOLA, J. C. (1968). *Brookhaven Symp. Biol.* **21**, 24.

LUZZATI, V. (1953). *Acta Cryst.* **6**, 142.

MOFFAT, J. K. (1971). *J. Mol. Biol.* In the press.

NOBBS, C. L. (1966). *Hemes and Hemoproteins*, p. 143. Edited by B. CHANCE, R. W. ESTABROOK & T. YONETANI. London: Academic Press.

NOBBS, C. L., WATSON, H. C. & KENDREW, J. C. (1966). *Nature, Lond.* **209**, 339.

PERUTZ, M. F. & MATHEWS, F. S. (1966). *J. Mol. Biol.* **21**, 199.

PERUTZ, M. F., MUIRHEAD, H., COX, J. M., GOAMAN, L. C. G., MATHEWS, F. S., MCGANDY, E. L. & WEBB, L. E. (1968). *Nature, Lond.* **219**, 29.

SIGLER, P. B., JEFFERY, B. A., MATTHEWS, B. W. & BLOW, D. M. (1966). *J. Mol. Biol.* **15**, 175.

SIM, G. A. (1959). *Acta Cryst.* **12**, 813.

SIM, G. A. (1960). *Acta Cryst.* **13**, 511.

STEITZ, T. A., HENDERSON, R. & BLOW, D. M. (1969). *J. Mol. Biol.* **46**, 337.

STEITZ, T. A., LUDWIG, M. L., QUIOCHO, F. A. & LIPSCOMB, W. N. (1967). *J. Biol. Chem.* **242**, 4662.

STRYER, L., KENDREW, J. C. & WATSON, H. C. (1964). *J. Mol. Biol.* **8**, 96.

WOOLFSON, M. M. (1956). *Acta Cryst.* **9**, 804.

WYCKOFF, H. W., HARDMAN, K. D., ALLEWELL, N. M., INAGAMI, T., JOHNSON, L. N. & RICHARDS, F. M. (1967). *J. Biol. Chem.* **242**, 3984.

Acta Cryst. (1971). B**27**, 1420

The Crystal Structure of DL-Ornithine Hydrobromide*

BY A. R. KALYANARAMAN AND R. SRINIVASAN

Centre of Advanced Study in Physics, University of Madras, Madras-25, India

(*Received* 11 *December* 1968 *and in revised form* 15 *July* 1970)

The crystal structure of DL-ornithine hydrobromide $NH_3(CH_2)_3 . CH(NH_3^+) . COO^- HBr$ has been determined using three-dimensional visual intensity data obtained with Cu $K\alpha$ radiation. The crystal belongs to the monoclinic system with cell dimensions $a=9{\cdot}388$; $b=7{\cdot}901$; $c=11{\cdot}663$ Å and $\beta=109°50'$. The space group is $P2_1/c$. The structure has been refined using the full-matrix least-squares method. Final R value is 7·3% for the observed data. All the hydrogen atoms have been located from a difference Fourier synthesis. The standard deviation for the positions of the non-hydrogen atoms is about 0·008 Å. All bond lengths and angles are found to be normal. Both nitrogen atoms are protonated and the carboxyl group exists in the ionic form. The conformation and other aspects of the molecule are discussed.

Introduction

The determination of the crystal structure of amino acids, peptides and related compounds forms part of a major programme of research on the molecular structure of compounds of biological interest in this centre. Although ornithine is not one of the commonly occurring amino acids it is found in certain antibiotics. No detailed investigation of the structure of this compound appears to have been made earlier except in a complex form in ferrichrome (Zalkin, Forrester & Templeton, 1966). Therefore, it was decided to determine its structure. However, when the present investigation was nearing completion the analysis of the structure of L-ornithine hydrochloride by another group of workers (Chiba, Ueki, Ashida, Sasada & Kakudo, 1967) was brought to our notice. The final results of both these investigations are in good agreement. A note reporting the preliminary results of our present investigation has been published earlier (Kalyanaraman, 1967) and the full details are presented here.

Experimental

Crystals of DL-ornithine hydrobromide were obtained by dissolving DL-ornithine hydrobromide in a water-ethanol mixture and evaporating it under controlled humidity. Good crystals were obtained after several attempts; they were in the form of needles with a as the needle axis.

The unit-cell dimensions, space group and other data of this compound were obtained by the use of Weissenberg and Buerger precession techniques. The syste-

* Contribution No. 310 from the Centre of Advanced Study in Physics, University of Madras, Madras-25, India.