

covalent bond (2.27 Å), and is thus comparable with the apical copper-water oxygen bond of 2.195 Å in AdCu. The internuclear Cu-Cu distance of 3.024 Å is 0.075 Å longer than the corresponding distance in AdCu. Chemically the 'bite' of hypoxanthine and adenine should be almost identical. The effect of replacing the water molecule as apical ligand with a chlorine atom may result in a stronger axial coordination and thus pull the metal ions further apart. There is probably no incipient covalent bonding between the copper ions that might counteract a larger separation.

Molecular packing

The environment of the complex is shown in Fig. 3. Each hypoxanthine ligand has two hydrogen-bond donors, N(1) and N(7), and one potential hydrogen-bond acceptor, O(6). Donor atoms N(11), N(71) and N(72) participate in hydrogen bonds to water oxygen atoms O(2), O(3) and O(1) respectively, while the fourth donor atom N(12) is hydrogen bonded to a chlorine atom Cl(2). The carbonyl oxygen atoms do not participate in hydrogen bonding. Instead both O(61) and O(62) have close contacts to imidazole rings in neighbouring dimers, with the C=O vector pointing nearly perpendicular to the plane of the imidazole ring. O(61) is situated 3.00 Å from the plane through the imidazole ring of ligand (2) translated one unit along the negative *z* axis, while O(62) is situated 3.00 Å from the corresponding plane of ligand (1) inverted through the centre of symmetry in (1, $\frac{1}{2}$, 0). The occurrence of identical intermolecular short contacts for the two crystallographically independent carbonyl groups may indicate the

existence of a stabilizing interaction between the carbonyl dipole and the π -system of the purine moiety.

The chlorine atom coordinated to copper is hydrogen bonded to two water molecules. The O(9)-H...Cl(1) bond of 3.21 Å is 'normal' while the O(1)(i)-H...Cl(1) bond of 3.41 Å is on the limit of what might be considered a hydrogen bond. The non-coordinated chlorine atom Cl(2) participates in two 'normal' hydrogen bonds [N(12)-H...Cl(2) = 3.21 Å, and O(3)(iii)-H...Cl(2) = 3.24] and one 'long' hydrogen bond [O(1)(i)-H...Cl(2) = 3.50 Å]. The angle \angle Cl(1)-O(1)(i)-Cl(2) = 108.8° is close to an H-O-H bond angle, and is thus consistent with the presence of 'long' hydrogen bonds between water molecule O(1) and the two chlorine ions.

The author wishes to thank Dr Lyle H. Jensen for valuable advice in the course of this work. Financial support under Grant AM-3288 from the National Institutes of Health is gratefully acknowledged.

References

- ALEXANDER, L. E. & SMITH, G. S. (1964). *Acta Cryst.* **17**, 1195.
International Tables for X-ray Crystallography (1962). Vol. III, p. 208. Birmingham: Kynoch Press.
 SLETTEN, E. (1969). *Acta Cryst.* **B25**, 1480.
 SLETTEN, E., SLETTEN, J. & JENSEN, L. H. (1969). *Acta Cryst.* **B25**, 1330.
 STEWART, J. M. (1964). *Crystal Structure Calculations System X-ray* 63. Univ. of Maryland Technical Report TR-64-6. Computer Sciences Center, University of Maryland.
 WEISS, R. & VENNERS, H. (1965). *Z. Physiol.* **340**, 138.

Acta Cryst. (1970). **B26**, 1614

An Alternative Form for $B_{3,0}$, a Phase Determining Formula

BY J. KARLE

U.S. Naval Research Laboratory, Washington, D.C., 20390, U.S.A.

(Received 2 October 1969)

An alternative form is derived for the phase determining formula $B_{3,0}$ defining $\cos(\varphi_{h_1} + \varphi_{h_2} + \varphi_{-h_1-h_2})$ in which improved scaling is achieved by replacing the coefficient in the original formulation of this formula by a ratio of averages over functions of normalized structure factor magnitudes. Subsets of the measured data used with the modified $B_{3,0}$ formula generally give more accurate results for the cosine function than the complete set of data. A method for examining various subsets and choosing an optimum value is presented. The modified $B_{3,0}$ relation readily permits use of limited subsets of the experimental data. It has application as an auxiliary formula for use with the symbolic addition procedure. Used with care, the modified $B_{3,0}$ formula can help to determine which triples, $\varphi_{h_1} + \varphi_{h_2} + \varphi_{-h_1-h_2}$, are close to zero, a basic assumption in the symbolic addition procedure, and which ones may be exceptions. The formula can also help to evaluate symbolic phases.

A form of expression $B_{3,0}$ of sufficient generality for the present purposes is given for centrosymmetric crys-

als (equation 2.1.3 of Hauptman & Karle, 1958) and for noncentrosymmetric crystals (equation 2.1.3 of

Karle & Hauptman, 1958) by

$$|E_{\mathbf{h}_1} E_{\mathbf{h}_2} E_{\mathbf{h}_1 + \mathbf{h}_2}| \cos(\varphi_{\mathbf{h}_1} + \varphi_{\mathbf{h}_2} + \varphi_{-\mathbf{h}_1 - \mathbf{h}_2}) \\ \simeq C \langle (|E_{\mathbf{k}}|^p - |\overline{E}|^p) (|E_{\mathbf{h}_1 + \mathbf{k}}|^p - |\overline{E}|^p) \\ (|E_{\mathbf{h}_1 + \mathbf{h}_2 + \mathbf{k}}|^p - |\overline{E}|^p) \rangle_{\mathbf{k}} + R, \quad (1)$$

where $|E_{\mathbf{h}}|$ and $\varphi_{\mathbf{h}}$ are the magnitude and phase, respectively, of the normalized structure factor associated with the vector \mathbf{h} , C is a constant which differs for centrosymmetric and noncentrosymmetric crystals, R represents correction terms, $|\overline{E}|^p$ is the average value of $|E|^p$ and p is normally chosen to be some small number; values in the range $0.5 \leq p \leq 3$ have been found to be useful and the results have been rather insensitive to the particular value of p which was used. The average is to be taken over the set of vectors, \mathbf{k} , representing all the accessible products in the data set as defined within the average brackets in equation (1).

The constant, C , as given in the original theory was not suitable for application with the usual experimental data confined to the copper sphere of scattering and often introduced significant error into the calculation of equation (1). A procedure for obtaining a rescaled value for C has been presented by Hauptman (1970) and Hauptman, Fisher, Hancock & Norton (1969).

A modification of equation (1) will be presented which eliminates the constant C , substituting instead a ratio of averages over the same data which are used to compute the average in equation (1). This modification, in addition, readily permits the use of a limited subset of the experimental data in obtaining values for the function $\cos(\varphi_{\mathbf{h}_1} + \varphi_{\mathbf{h}_2} + \varphi_{-\mathbf{h}_1 - \mathbf{h}_2})$. In deriving the new form for formula $B_{3,0}$ the correction terms R will be omitted in the interest of simplification and in the expectation that their effect will be minimized, owing to the fact that the final result is in the form of a ratio of terms having similar corrections in the numerator and denominator.

If $\mathbf{h}_1 = \mathbf{h}_2 = (0,0,0)$ and R is ignored, equation (1) becomes

$$|E_{000}|^3 \simeq C \langle (|E_{\mathbf{k}}|^p - |\overline{E}|^p)^3 \rangle_{\mathbf{k}}. \quad (2)$$

If $\mathbf{h}_1 = -\mathbf{h}_2$ and R is ignored, equation (1) becomes

$$|E_{\mathbf{h}_1}|^2 E_{000} \simeq C \langle (|E_{\mathbf{k}}|^p - |\overline{E}|^p)^2 (|E_{\mathbf{h}_1 + \mathbf{k}}|^p - |\overline{E}|^p) \rangle_{\mathbf{k}}. \quad (3)$$

Equations (1), (2) and (3) can be combined to give the desired result, a modified $B_{3,0}$,

$$\cos(\varphi_{\mathbf{h}_1} + \varphi_{\mathbf{h}_2} + \varphi_{-\mathbf{h}_1 - \mathbf{h}_2}) \simeq \\ \frac{\langle (|E_{\mathbf{k}}|^p - |\overline{E}|^p) (|E_{\mathbf{h}_1 + \mathbf{k}}|^p - |\overline{E}|^p) (|E_{\mathbf{h}_1 + \mathbf{h}_2 + \mathbf{k}}|^p - |\overline{E}|^p) \rangle_{\mathbf{k}} \langle (|E_{\mathbf{k}}|^p - |\overline{E}|^p)^3 \rangle_{\mathbf{k}}^{1/2}}{[A(\mathbf{h}_1)A(\mathbf{h}_2)A(\mathbf{h}_1 + \mathbf{h}_2)]^{1/2}}, \quad (4)$$

where

$$A(\mathbf{h}_i) = \langle (|E_{\mathbf{k}}|^p - |\overline{E}|^p)^2 (|E_{\mathbf{h}_i + \mathbf{k}}|^p - |\overline{E}|^p) \rangle_{\mathbf{k}}. \quad (5)$$

In applying equation (4), a choice is made of vectors \mathbf{h}_1 and \mathbf{h}_2 and then the appropriate averages can be computed. The calculation may be made over a restricted sample of the set of data, but this sample is

selected from the entire set of experimental values $|E_{\mathbf{h}}|$ expanded to include all equivalent reflections, not just the independent ones. It has been an empirical observation in this laboratory that higher accuracy can be obtained from equation (4) if a properly chosen subset of the data, rather than the entire set, is used in the calculation. A basis for selecting the restricted sample is the requirement that at least one factor in the first average in the numerator of equation (4) have an $|E_{\mathbf{h}}| > \alpha$ where α might assume a particular value for example in the range $1.5 < \alpha < 2.5$. Other values may be suitable. The other two factors would be unrestricted in the values that they could assume. Another way to restrict the sample to be computed in this average is to put requirements on each of two $|E_{\mathbf{h}}|$ values occurring in two of the three factors. No constraint would be put on the value of the third factor. Putting constraints on the values of some of the factors severely limits the vectors \mathbf{k} which can enter into the first average. The other four averages occurring in equation (4) would be carried out only over those values of \mathbf{k} that enter into the first average in the numerator.

One way to set the levels of the constraint α is to vary them so that the range in the number of contributors to the averages in equation (4) carries the result for $(\varphi_{\mathbf{h}_1} + \varphi_{\mathbf{h}_2} + \varphi_{-\mathbf{h}_1 - \mathbf{h}_2})$ through the region of a flat or slowly varying function of the number of contributors. This can happen in several ways. The function can be slowly varying because it is in the region of a minimum, maximum or point of inflection, or it may be that the slope, though not zero anywhere, becomes quite small. Presumably a suitable result for the value of $\cos(\varphi_{\mathbf{h}_1} + \varphi_{\mathbf{h}_2} + \varphi_{-\mathbf{h}_1 - \mathbf{h}_2})$ is obtained at its minimum slope as a function of the number of contributors to equation (4) because it is there that the value obtained shows a minimum sensitivity to change in the sampling. Not all calculations will give satisfactory minimal slopes and these should be rejected. If none or two of the $A(\mathbf{h}_i)$ in the denominator of equation (4) have negative signs a real square root can be taken. However since each $A(\mathbf{h}_i)$ is proportional to $|E_{\mathbf{h}_i}|^2 E_{000}$, the presence of a negative value for $A(\mathbf{h}_i)$ should make the sampling quite suspect and under such circumstances the calculation should also be rejected.

There are many ways to employ equation (4). One way, as mentioned several times previously (e.g. Karle & Karle, 1964; Karle & Karle, 1966), is as an auxiliary formula for use with the symbolic addition procedure. The signs of certain special two-dimensional pure real

reflections can be calculated for several choices of \mathbf{h}_1 and \mathbf{h}_2 associated with large $|E|$ values in such a way that it is not necessary to know the phase corresponding to either \mathbf{h}_1 or \mathbf{h}_2 . This arises when \mathbf{h}_1 and \mathbf{h}_2 are suitably chosen equivalent reflections. It is most valuable when several choices of \mathbf{h}_1 and \mathbf{h}_2 are accessible to a given reflection defined by $\mathbf{h}_1 + \mathbf{h}_2$ affording many

indications of the sign of $E_{\mathbf{h}_1+\mathbf{h}_2}$, since contradictions in the sign obtained from equation (4) have been found to occur in practice. For some crystals contradictions rarely occur, whereas for others they occur more frequently. The sign obtained from equation (4) is the sign of the first average in the numerator of this formula, just as in the original $B_{3,0}$ formula. One advantage in computing equation (4) is that limited samples of data are readily employed and, when properly chosen, afford improvements in accuracy and can perhaps also save computer time. Another advantage is that the accuracy with which the magnitudes of the cosines of the sums $\varphi_{\mathbf{h}_1} + \varphi_{\mathbf{h}_2} + \varphi_{-\mathbf{h}_1-\mathbf{h}_2}$ of known value are computed can be used as a rough indication of the reliability of the calculations, e.g. a sign indication obtained from a computed magnitude very much less than unity when unity is expected cannot be considered to be highly reliable. The cosines have a magnitude of unity, for example, if all the reflections involved are pure real or if $\varphi_{\mathbf{h}_1} + \varphi_{\mathbf{h}_2}$ is 0 or π as a result of a special choice of \mathbf{h}_1 and \mathbf{h}_2 for a particular space group and $\mathbf{h}_1 + \mathbf{h}_2$ belongs to a pure real reflection. For example, in space group $P2_1$, if $\mathbf{h}_1 = (h, k, l)$ and $\mathbf{h}_2 = (\bar{h}, \bar{k}, \bar{l})$ then $-\mathbf{h}_1 - \mathbf{h}_2 = (2h, 0, 2l)$ and the left side of equation (4) becomes $\cos(\varphi_{2h02l} - k\pi)$, owing to the relation between the phases of index $h k l$ and $h \bar{k} l$ for this space group. Since $\varphi_{2h02l} - k\pi$ is real, the magnitude of the cosine function must evidently be unity.

An alternative form for equation (4) can be based on $I_{3,0}$ (equation 2.2.3 of Hauptman & Karle, 1958 and equation 2.2.3 of Karle & Hauptman, 1958). In this case, a typical term in the averages of equation (4) is

$$\frac{|E_{\mathbf{k}}|^t - 1}{\ln |E_{\mathbf{k}}|} - \left(\frac{|E|^t - 1}{\ln |E|} \right),$$

replacing $(|E_{\mathbf{k}}|^p - |E|^p)$.

If all the reflections are pure real and $\mathbf{h}_1 = \mathbf{h}_2 = \mathbf{h}$, an alternative form which eliminates the constant C of equation (1) can be written

$$\frac{E_{2h}}{E_{000}} \simeq \frac{\langle (|E_{\mathbf{k}}|^p - |E|^p) (|E_{\mathbf{h}+\mathbf{k}}|^p - |E|^p) (|E_{2\mathbf{h}+\mathbf{k}}|^p - |E|^p) \rangle_{\mathbf{k}}}{\langle (|E_{\mathbf{k}}|^p - |E|^p)^2 (|E_{\mathbf{h}+\mathbf{k}}|^p - |E|^p) \rangle_{\mathbf{k}}}. \quad (6)$$

Similar equations can be written for special relations between $E_{\mathbf{h}_1}$ and $E_{\mathbf{h}_2}$ generated by the various symmetry elements in centrosymmetric and noncentrosymmetric space groups which result in $|E_{\mathbf{h}_1}| = |E_{\mathbf{h}_2}|$. Such special relations among the structure factors are listed for centrosymmetric crystals in a series of papers in *Acta Crystallographica* [starting with Karle & Hauptman (1959), and concluding with Karle & Hauptman (1961)]. Similar relations are readily determined for noncentrosymmetric space groups.

An application of equation (4) with $p=2$ was made to the structure determination of a cyclic peptide composed of 4 glycy and 2 alanyl units. Table 1 is illustrative of the nature of the resulting phase information. The fourth column gives the value of $\cos(\varphi_{\mathbf{h}_1} + \varphi_{\mathbf{h}_2} + \varphi_{-\mathbf{h}_1-\mathbf{h}_2})$ computed from the correct phases listed in column three for the final structure. The last columns give four values of this function as computed from the right side of equation (4). The number of terms contributing to each calculation is indicated in parentheses and was determined by requiring that at least one factor in each product of three appearing in the first average in the numerator of equation (4) have an $|E|$ value larger than 2.1, 1.8, 1.6 and 1.4, respectively, for the four values listed in Table 1. Except for the fifth triple, inspection of the last four columns reveals computed values for the cosine functions which are in regions of sampling size that satisfy the criterion of minimum

Table 1. Application of equation (4) to data from a cyclopeptide consisting of 4 glycy and 2 alanyl units

Columns 4 and 5-8 compare results for $\cos(\varphi_{\mathbf{h}_1} + \varphi_{\mathbf{h}_2} + \varphi_{-\mathbf{h}_1-\mathbf{h}_2})$ as obtained from the known structure and from equation (4). The numbers in parentheses represent the number of contributors to the averages occurring in equation (4).

Triple	$ E_{\mathbf{h}} $	$\varphi_{\mathbf{h}}$	$\cos(\varphi_{\mathbf{h}_1} + \varphi_{\mathbf{h}_2} + \varphi_{-\mathbf{h}_1-\mathbf{h}_2})$ from $\varphi_{\mathbf{h}}$	$\cos(\varphi_{\mathbf{h}_1} + \varphi_{\mathbf{h}_2} + \varphi_{-\mathbf{h}_1-\mathbf{h}_2})$ from equation (4)			
0 5 6	2.33	-1.57					
	2.62	-1.66	0.99	1.51	1.53	1.59	1.63
5 7 7	2.81	3.11		(146)	(264)	(497)	(807)
	2.33	1.57					
2 0 7	2.43	-1.57	0.71	1.03	1.07	1.09	1.21
	2.40	0.78		(405)	(743)	(1346)	(2303)
5 16 2	2.47	1.84					
	2.47	-1.30	-0.86	-3.20	-1.68	-1.21	-1.14
10 0 0	1.87	3.14		(70)	(136)	(256)	(392)
	2.22	3.14					
0 3 2	3.41	-1.57	-1.00	-0.12	-0.22	-0.23	-0.15
	3.30	1.57		(267)	(477)	(873)	(1456)
3 19 4	2.49	0.93					
	3.41	1.57	1.00	-1.64	-1.29	-1.04	-0.71
3 16 6	2.03	-2.49		(208)	(405)	(759)	(1237)
	2.49	-0.93					
3 19 4	2.81	0.00	0.95	-0.81	-0.99	-0.97	-0.94
	1.83	1.25		(50)	(95)	(175)	(265)

change of value with number of contributors. If the computed value chosen according to this criterion exceeds a magnitude of unity, it is set back to unity. It would be desirable in the application to an unknown structure to continue the computation of the fifth triple (and perhaps the third one) with more contributors, if available. If the value continued to change at the same rate, this would be grounds for rejecting the calculation. The first two triples were selected to represent the major portion of the results obtained from the computation. A majority of the acceptable calculations carried out for the peptide were in comparable or somewhat better agreement with the expected value than these first two. They confirm the basic assumption made in carrying out the symbolic addition procedure, namely that the sums of the phases $\varphi_{h_1} + \varphi_{h_2} + \varphi_{-h_1-h_2}$ associated with the larger $|E|$ values are generally close to zero. Note that the computations are made only for triples based on large $|E|$ values.

Of particular interest are those triples whose phases do not sum approximately to zero, since the latter is the assumed basis for the symbolic addition procedure. An extreme case would be to sum to π . In such instances the symbolic addition procedure would give a phase indication in error by π radians. The third and fourth triples in Table 1 are illustrative of this case. Such occurrences are fairly rare and similar indications from equation (4) are correspondingly rare. In the case of the fourth triple, the doubt cast by equation (4) on the indication of the sign of $\varphi_{0,11,6}$ from the symbolic addition procedure was instrumental in changing a significant number of the phases and readily facilitated the structure determination.

The last two triples in Table 1 are illustrated for the purpose of inserting a strong word of caution concerning the use of equation (4). Here are two examples (accepting the indications for the fifth triple) in which the triples are indicated to be exceptions to the assumption of the symbolic addition procedure, the implication being that the sums of the phases are close

to π . As shown in column four, they are clearly not exceptions and, in fact, overwhelming evidence from the symbolic addition procedure confirmed this, so that the implications of equation (4) were ignored and no difficulties arose in the phase determination.

In its use as an auxiliary formula for application with the symbolic addition procedure, equation (4) can help confirm that most triples have sums of angles close to zero and throw doubt on others. For the ones in doubt, it is worthwhile to try to avoid their use in the early stages of the phase determination or perhaps denote the appropriate phase by a symbol if it is important to utilize it in succeeding steps in the procedure. It is apparent however that information from equation (4) can be incorrect, an extreme case occurring when the sign of the first average in the numerator is wrong and of a magnitude to cause an error in the sum of phases as large as π . The conclusion is that the results from equation (4) can be helpful in combination with the symbolic addition procedure, but individual phase indications should be regarded as suggestive rather than conclusive.

Mr. Stephen Brenner wrote the computer programs and carried out the calculations described in this note. His fine cooperation is very much appreciated.

References

- HAUPTMAN, H. (1970). *The Calculation of Structure Invariants*, in *Crystallographic Computing*, edited by F. R. AHMED, S. R. HALL & C. P. HUBER, Copenhagen: Munksgaard.
- HAUPTMAN, H., FISHER, J., HANCOCK, H. & NORTON, D. (1969). *Acta Cryst.* **B25**, 811.
- HAUPTMAN, H. & KARLE, J. (1958). *Acta Cryst.* **11**, 149.
- KARLE, I. L. & KARLE, J. (1964). *Acta Cryst.* **17**, 835.
- KARLE, J. & HAUPTMAN, H. (1958). *Acta Cryst.* **11**, 264.
- KARLE, J. & HAUPTMAN, H. (1959). *Acta Cryst.* **12**, 404.
- KARLE, J. & HAUPTMAN, H. (1961). *Acta Cryst.* **14**, 105.
- KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* **21**, 849.