one ten-thousandth of the cell edge were formed by linear interpolation of a table in thousandths. Strictly isotropic structure factors were computed at the rate of 25 per min., as were the hydrogen contributions, while general structure factors were computed at 3 per min. With the latter, 48 general atoms had to be considered since storage facilities were not available for including special cases. Even so, the program required 1,987 storage locations out of a possible 2,000.

In all, a total of approximately 450 hr. machine time was consumed by this analysis. This represents 15 cycles of three-dimensional refinement.

We should like to thank Prof. D. Harker for many discussions and for his continued interest in the problem. In addition, we should like to thank the staff at IBM's Watson Scientific Computing Laboratory for the cooperation received throughout the course of this analysis. The Protein Structure Project is supported by the following organizations: The Dean Langmuir Foundation, The Rockefeller Foundation, The Damon Runyon Memorial Fund, The New York Foundation, The International Business Machines Corporation, The Polytechnic Institute of Brooklyn.

References

Cochran, W. (1956). *Acta Cryst.* **9**, 924.

Jeffrey, G. A., Parry, G. S. & Mozzi, R. L. (1956). *J. Chem. Phys.* **25**, 1024.

Pauling, L. & Sherman, J. (1934). *Proc. Nat. Acad. Sci., Wash.* **20**, 340.

Tulinsky, A. (1957). *Acta Cryst.* **10**, 745.

Acta Cryst. (1959). **12**, 634

Basic Beryllium Acetate: Part III. Evidence for Chemical Bonding; Assessment of Accuracy*

By A. Tulinsky†

The Protein Structure Project, Polytechnic Institute of Brooklyn, Brooklyn 1, *N.Y., U.S.A.*

(*Received* 15 *October* 1957 *and in revised form* 11 *December* 1958)

Anharmonic thermal vibrations, systematic errors in the data and the systematic accumulation of random errors have been considered as possible sources for the excess positive density observed in the C–C and C–O bonds of the acetate group (Part II). None were found to account for this density completely and satisfactorily. An assessment of the accuracy of the structure determination has been made by averaging over the cell. The advantages of atoms in special positions are discussed.

1. Introduction

It has been mentioned in Part II that the discrepancies between the observed and calculated structures in the bonds of the acetate group could arise from one or more of four effects: chemical bonding, anharmonic thermal vibrations, systematic errors, accumulation of random errors. The last three will now be considered in some detail.

2. Anharmonic thermal vibrations

Fundamentally, there is a clear distinction between chemical bonding and anharmonic thermal vibrations. Anharmonicity of thermal vibration would be expected to affect both inner and outer shell electrons whereas chemical bonding involves only the outer shell. Therefore, in principle, the two can be differentiated by

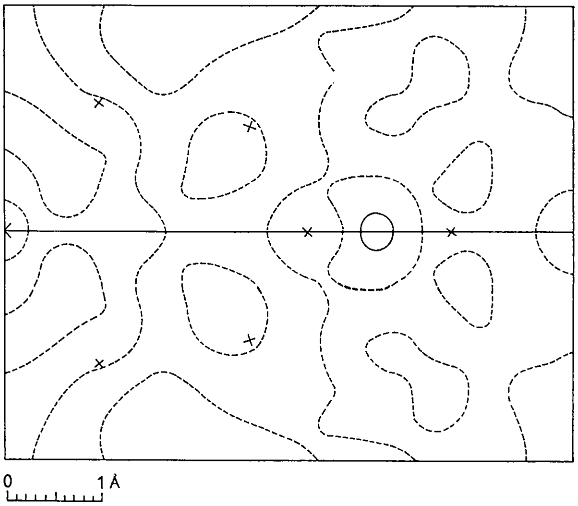

Fig. 1. High order $\varrho_o(MX)-\varrho_c$ in the plane of the acetate group, (F_o-F_c) in the range $2\theta \leq 83°$ omitted, contours at 0·1 e.Å^{-3}, solid positive, broken zero, crosses indicate atomic positions.

---

* Contribution No. 17 from the Protein Structure Project of the Polytechnic Institute of Brooklyn.

† Present address: Sterling Chemistry Laboratory, Yale University, New Haven, Conn., U. S. A.

considering, in the present case, the K shell distribution. Although the K shell distribution cannot be realized in practice, it can be approximated through a series including high order reflections only. The resulting distribution should contain little or no L shell contribution due to the rapid decay of L shell scattering with scattering angle. On the other hand, this distribution will not represent the K shell completely since the latter also contributes to low angle scattering.

The high order difference density considered for beryllium acetate (MX data) included terms satisfying $2\theta \geq 83°$ and it is shown in Fig. 1 (the SX data give the same features). It will be seen that the excess positive density in the bonds of the acetate (Fig. 4, Part II) has reduced significantly (C–C, peak height from 0·40 to 0·12 e.Å^{-3}; C–O, from 0·26 to 0·04 e.Å^{-3}) and that the density, in general, is extremely low. This would seem to indicate that something unaccountable by the calculated structure is carried by the low order data. Further support for such a conclusion can be gained from a detailed comparison of the discrepancies between observed and calculated structure factors and the expected *sd*'s of the observed structure amplitudes. The results are shown in Figs. 2 and 3 along with other related quantities. It will be seen that the $\Sigma||F_o| - |F_c||/\Sigma\sigma(|F_o|)$'s are not constant over the scattering range but rather increase systematically toward the low orders. Although such behavior can be attributed to a systematic underestimation of $\sigma(|F_o|)$'s in the low orders, it is difficult to reconcile, since the two sets of data are affected by a different kind of statistical counting error in this region, yet both behave similarly.

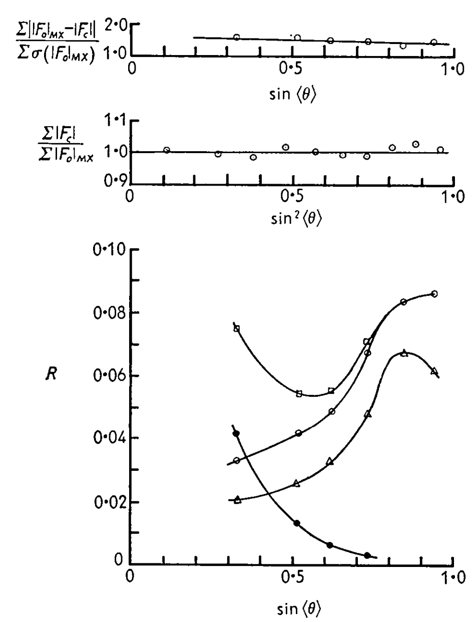

Fig. 2. Comparison of moving crystal data and the calculated structure, squares $-\Sigma||F_o|_{MX}-|F_c||/\Sigma|F_o|_{MX}$ hydrogens not included, circles—as the foregoing hydrogens included, solid circles-improvement effected including hydrogens, triangles $-\Sigma\sigma(|F_o|_{MX})/\Sigma|F_o|_{MX}$, all versus constant shells.

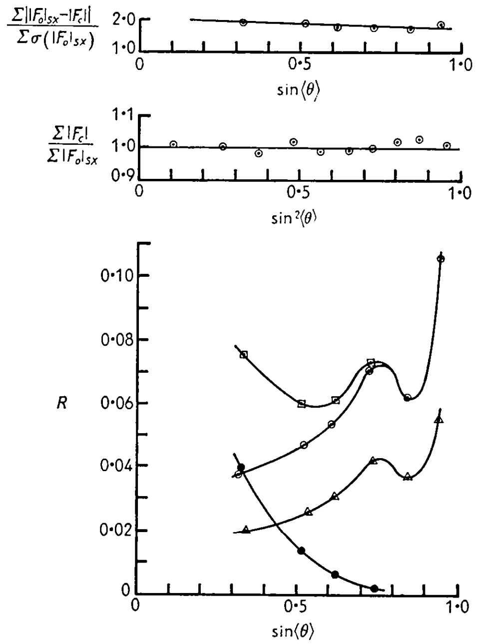

Fig. 3. As in Fig. 2, stationary crystal data replacing moving crystal.

The relatively large positive region in the high order difference density between the carbon atoms can be interpreted in either of two ways. It could be due to L shell electrons which have not been completely removed from the series or it could be due to some amount of anharmonic thermal vibration. In support of the former is the fact that the mean component of thermal vibration between the carbon atoms is small, suggesting that the L shell contribution may not have reached the observable limit at $2\theta = 83°$ and that it may well extend beyond. Such an effect can also account for the smaller positive region observed in the final difference density between C_I and O_{II}.

3. Systematic errors

Systematic errors between the two sets of intensity data would seem to be ruled out by the fact that the two give essentially the same results (Fig. 7, Part II), particularly so after rescaling. However, this in no way eliminates the possibility that both sets possess the same systematic errors, which, in turn, give rise to the unaccountable features observed in the final difference density. But such a possibility seems to be improbable because: the two sets of data were obtained from different crystals, the integrated intensities of the two sets were measured in different ways and the two sets of data are affected by different kinds of errors.

4. Accumulation of random errors

In the centrosymmetrical case, if it is assumed that the signs, $S(hkl)$, of the observed structure amplitudes, $|F_o(hkl)|$, are known, it can be shown that*

$$\begin{aligned}\sigma^2[\varrho_0(xyz)] &= (1/V^2)\sum_{hkl}\sum_{h'k'l'} S(hkl)\,S(h'k'l')\,\mathscr{F}(hklh'k'l') \\ &\quad\times\cos 2\pi[(h+h')x+(k+k')y+(l+l')z]\,, \qquad (1)\end{aligned}$$

where $\sigma[\varrho_0(xyz)]$ is the *sd* of the observed electron density at any point (xyz),

$$\mathscr{F}(hklh'k'l') = \overline{|F_o(hkl)||F_o(h'k'l')|} - \overline{|F_o(hkl)|}\,\overline{|F_o(h'k'l')|}\,,$$

V is the volume of the cell and the summation extends over all observed $(hklh'k'l')$. When $(hkl)=(h'k'l')$, then $\mathscr{F}(hklh'k'l')=\sigma^2(|F_o(hkl)|)$, the variance of $|F_o(hkl)|$, and if it is assumed that errors are not correlated, $\mathscr{F}(hklh'k'l')=0$ when $(hkl)\neq(h'k'l)$. This result applies if the whole sphere of data is considered; if only the independent reflections are considered, further relationships are obtained for $\mathscr{F}(hklh'k'l')$.

Now, although an approximation has been made to the $\sigma(|F_o(hkl)|)$'s of beryllium acetate, the synthesis of (1) over its asymmetric unit has not been attempted because of its complexity when dealing with only independent reflections, particularly in space group $Fd3$. This can be appreciated from the relationships between structure factors:

$$\begin{aligned}F(hkl) &= F(lhk) = F(klh) = (-1)^{(h+k)/2}F(\bar{h}\bar{k}l) \\ &= (-1)^{(h+l)/2}F(\bar{h}k\bar{l}) = (-1)^{(k+l)/2}F(h\bar{k}\bar{l}) = F(\bar{h}\bar{k}\bar{l})\,.\end{aligned}$$

Therefore, (1) was evaluated at certain special points instead. The results (MX data) are given in Table 1.

Table 1. *Expected error in ϱ_0 at certain special positions*

(MX) data

Position	$\sigma(\varrho_0)$
—	$0{\cdot}10_8$* e.Å^{-3}
23	$0{\cdot}09_8$
$\bar{1}$	$0{\cdot}08_9$
C–C	$0{\cdot}06_7$

* The sum of all the variances.

The C–C position, $(\frac{1}{4}, 0, 0)$, given in Table 1, lies close to the center of the C–C bond (0·08 Å away). From its $\sigma(\varrho_0)$, it will be seen that the excess density observed in the C–C bond is about $5\times\sigma(\varrho_0)$ near the peak position. This would seem to eliminate the possibility that this effect is due to a systematic accumulation of random errors, unless, of course, the $\sigma(|F_o(hkl)|)$'s have been seriously underestimated. However, such an underestimation is not supported by the behavior of the remaining difference density. In addition, it would seem to be improbable that an underestimation has been made of such large magnitude. Concerning the excess density in the C–O bond, it too is probably significant with respect to the expected $\sigma(\varrho_0)$ in that region. The density is located in a general position so that the effect of the accumulation of random errors is necessarily reduced by the trigonometry of the series. If the average is considered, then the excess density in the C–O bond is about $4\times\langle\sigma(\varrho_0)\rangle$ near the peak. And finally, it will be seen from Table 1 that the most sensitive position in the cell with respect to accumulation of random errors is the 23 point. This is further supported by $\varrho_0(MX)-\varrho_0(SX)$ (Fig. 7, Part II), even though the 23 point here is particularly sensitive to scaling.

5. Assessment of accuracy

The accuracy of the structure has been assessed from $\sigma(|F_o(hkl)|)$'s by averaging over the cell

$$\langle\cos^2 2\pi(hx+ky+lz)\rangle = \tfrac{1}{2}\,.$$

The results for the two sets of data are very similar. The average *sd* of the electron density is about 0·06 e.Å^{-3}. In estimating the *sd*'s of atomic coordinates, the effect of the cyclic permutation of indices was included (Cruickshank, 1949) and average values were used for the curvature at atomic centers, since, in some cases, they differ appreciably in diametrically opposed directions. The curvatures were obtained assuming a Gaussian charge distribution near atomic centers, $\varrho_0(r)=\varrho_0(0)\exp(-pr^2)$. The average *sd* of the rate of change of the electron density, $\langle\sigma(\partial\varrho_0/\partial r)\rangle$, is 0·299 e.Å^{-4}. The *sd*'s of the atomic coordinates are given in Table 2 along with other related quantities, O_I being included for comparison.

Table 2. *sd's of atomic coordinates*

Atom	$\varrho_o(0)$ (e.Å^{-3})	$\langle p\rangle$ (Å^{-2})	$\langle\partial^2\varrho_o/\partial r^2\rangle_{r=0}$ (e.Å^{-5})	σ (Å)
Be(u)	6·1	6·0	−73·8	0·004
C_I(v)	8·3	4·7	−77·8	0·004
C_{II}(w)	7·4	4·8	−71·0	0·004
O_{II}(x)	10·0	3·7	−73·1	0·004
O_{II}(y)	10·0	4·6	−92·6	0·003
O_{II}(z)	10·0	4·1	−82·7	0·004
O_I	14·1	5·0	−140·9	0

The *sd*'s of the atomic coordinates agree well with experiences encountered in structure factor computations. Changes of less than 0·003 Å in atomic coordinates had little effect on the agreement between observed and calculated structure factors and could not be recognized in the ensuing difference density; this was not true of larger changes.

Only an empirical estimate has been made of the accuracy of the thermal parameters and the apparent electron counts. It is based upon a survey of para-

---

* Harker & Tulinsky, to be published.

meter variations that were performed on reflections sensitive to certain atoms ($h+k+l=4n+2$, only O_{II} and Be contribute appreciably, etc.) and also, upon the detectibility of changes in the corresponding difference density. Variations of ±5% in principal mean square displacements and about ±7° in the orientation of principal vibration directions had a definite effect on the agreement between observed and calculated structure factors and could be recognized in the difference density. Changes of the order of 0·3 of an electron in count could be detected for the beryllium atoms and of the order of 0·2 of an electron for the other atoms (excluding O_{II}).

Table 3. *Principal atomic distances and interbond angles*

C_I–C_{II}	1·500 ± 0·006 Å	∠ O_{II}–C_I–O_{II}	123·4 ± 0·3°
C_I–O_{II}	1·264 ± 0·008	∠ O_I–Be–O_{II}	115·2 ± 0·2
Be–O_I	1·666 ± 0·004	∠ O_{II}–Be–O_{II}	102·7 ± 0·3
Be–O_{II}	1·624 ± 0·010		
C–H	1·1		

The principal interatomic distances and interbond angles and their estimated *sd*'s are given in Table 3. The distances involving O_{II} are necessarily inferior since the latter are located in a general position. No accuracy has been assigned to the C–H distance because reasonable agreement can be obtained over a relatively wide range of distances.

6. Concluding remarks

An important conclusion can be drawn from the experience of this structure analysis regarding the resolution of any fine detail such as chemical bonding: cases involving atoms located in general positions should be avoided as much as possible. There are several reasons for this. The number of parameters requiring determination is reduced and the determination of some of them becomes more objective and direct. In addition, the atoms in special positions are, in a sense, more overdetermined than those in general positions. This will be clear from Table 4, where beryllium acetate is considered.

Table 4. *Total number of reflections affected versus atomic parameters*

Atom	Position	N_r	N_p	$N_r/N_p/(N_r/N_p)_{O_{II}}$
O_I	23 point	295	1	7·2
Be	3-fold axis	370	3	3·0
C_I	2-fold axis	295*	5	1·4
C_{II}	2-fold axis	295*	5	1·4
O_{II}	general	370	9	1

* When general harmonic thermal vibrations are included, the carbon atoms can contribute to all the reflections. However, since the contribution is small to reflections of the type $h+k+l=4n+2$, the carbon atoms are considered as isotropic.

Here, N_r is the total number of reflections that the atom can affect (370 being the maximum, Cu $K\alpha$) and N_p the total number of atomic parameters requiring determination (positional parameters and general thermal vibrations). From the last column, it will be seen that there is much more information available for the determination of O_I's lone parameter as compared with any single parameter of O_{II}. In fact, the order of these ratios was found to correspond with ease of determination, O_{II} being the most difficult atom to remove from the difference density.

A well known objection to all the foregoing is that random errors tend to accumulate at special positions. However, this can be assessed or at least approximated as it was with beryllium acetate, where the accumulation did not, in fact, reach any serious proportions (maximum about $1{\cdot}3\times\langle\sigma(\varrho_0)\rangle$. On the other hand, systematic errors could effect such positions seriously; however, they are, in principle detectable.

I should like to thank Prof. D. Harker for many discussions. The Protein Structure Project is supported by the following organizations: The Dean Langmuir Foundation, The Rockefeller Foundation, The Damon Runyon Memorial Fund, The New York Foundation, The International Business Machines Corporation, The Polytechnic Institute of Brooklyn.

References

Cruickshank, D. W. J. (1949). *Acta Cryst.* **2**, 154.