

References

- COHN, W. E. & DOHERTY, D. G. (1956). *J. Amer. Chem. Soc.* **78**, 2863.
- DONOHUE, J. (1968). *Selected Topics in Hydrogen Bonding, in Structural Chemistry and Molecular Biology*. Ed. A. RICH & N. DAVIDSON, p. 443. San Francisco: W.H. Freeman.
- FURBERG, S. & JENSEN, L. (1968). *J. Amer. Chem. Soc.* **90**, 470.
- GERDIL, R. (1961). *Acta Cryst.* **14**, 333.
- HARRIS, R. D. & MACINTYRE, W. M. (1964). *Biophys. J.* **4**, 203.
- HAUPTMAN, H. & KARLE, J. (1956). *Acta Cryst.* **9**, 45.
- International Tables for X-ray Crystallography* (1963). Vol. III. Birmingham: Kynoch Press.
- HUNT, D.J. & SUBRAMANIAN, E. (1969). *Acta Cryst.* **B25**, 2144.
- JOHNSON, C. K. (1965). *ORTEP*. ORNL-3794, Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.
- KARLE, I. L. & KARLE, J. (1964). *Acta Cryst.* **17**, 835.
- KARLE, J. & HAUPTMAN, H. (1956). *Acta Cryst.* **9**, 635.
- KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* **21**, 849.
- KONDO, Y. & WITKOP, B. (1968). *J. Amer. Chem. Soc.* **90**, 764.
- KRAUT, J. & JENSEN, L. H. (1963). *Acta Cryst.* **16**, 79.
- PREWITT, C. T. (1966). Program ACACA. E. I. du Pont de Nemours and Company, Experimental Station, Wilmington, Delaware, 19898.
- SNIPES, W. & BERNHARD, W. (1968). *Radiation Res.* **33**, 162.
- SUNDARALINGHAM, M. & JENSEN, L. H. (1965). *J. Mol. Biol.* **13**, 930.
- WATENPAUGH, K., DOW, J., JENSEN, L. H. & FURBERG, S. (1968). *Science*, **159**, 206.
- YOUNG, D.W., TOLLIN, P. & WILSON, H.R. (1969). *Acta Cryst.* **B25**, 1423.

Acta Cryst. (1970). **B26**, 778

The Estimation of Film Intensities

BY L. M. TREFONAS AND R. MAJESTE

Department of Chemistry, Louisiana State University in New Orleans, New Orleans, Louisiana 70122, U.S.A.

(Received 24 July 1968 and in revised form 1 December 1968)

In the course of the structure investigation of 1-azabicyclo[3.2.0]heptane-1-methyl chloride, very different results were obtained when the intensities were estimated from films by two different techniques. When the intensities were estimated visually by comparison with a scale made up from timed exposures of a given reflection from the crystal, the data refined to a structure which was both crystallographically unsatisfactory and chemically unreasonable. When the intensities were estimated by use of an integrating microdensitometer, the data refined satisfactorily to a value of $R=0.11$ and produced a chemically reasonable structure. An analysis of both sets of data was carried out. The difficulties with the original visual data have been shown to arise from incorrect scaling of the multiple films within each given zone. Our analysis and conclusions are summarized here as a 'caveat' to other investigators using film techniques.

Introduction

This paper is concerned with a pitfall that can entrap the unwary investigator during the course of a crystal-

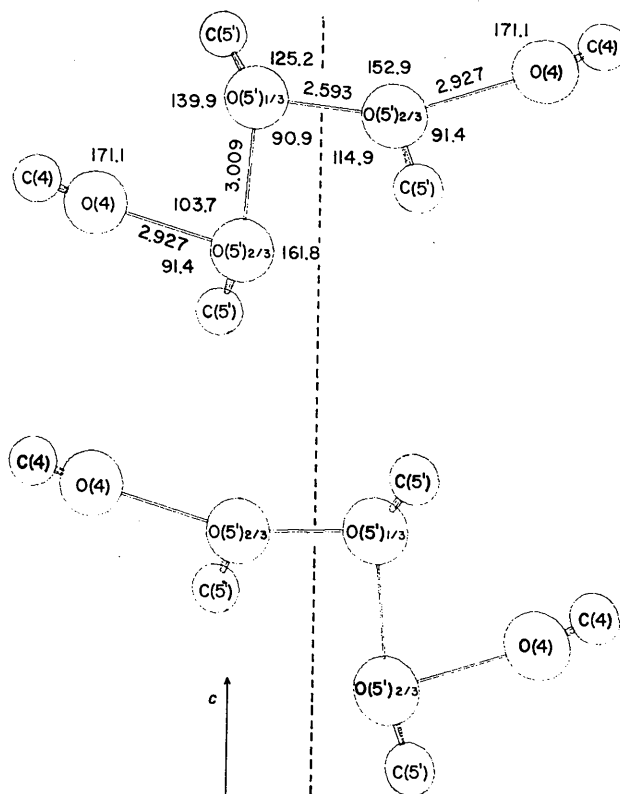


Fig. 7. Atoms near the screw axis parallel to the c axis. Possible hydrogen-bonding between the disordered $O(5)_{\frac{1}{3}}$ and $O(5)_{\frac{2}{3}}$ atoms.

lographic investigation. Having thus been ensnared ourselves, the following is an account summarized here as a 'caveat' to others who might find themselves in similar circumstances. It is a comparison and analysis

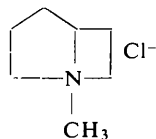
of the results obtained in a single-crystal X-ray diffraction study, with data from the same precession films but whose intensities were estimated in two different ways:

(a) *Visually*, by comparison of each film spot with a scale. This scale was prepared in the usual manner. Namely, a reflection from the same crystal, whose shape and size was typical of most of the spots on the film, was exposed for varying lengths of time, developed in the same manner as the intensity film, and assigned values proportional to the time of exposure. The minimum exposure on the film (one minute) corresponded to one cycle of the precession motion and the maximum exposure (sixty-five minutes) was still well within the linear response portion of the film.

(b) *Optically*, with the use of a recently purchased Nonius integrating microdensitometer (Model II). Diaphragms were chosen which conformed in size and shape to the spots on the film.

Crystal data and structure determination

The compound 1-azabicyclo[3.2.0]heptane-1-methyl chloride



crystallized in an orthorhombic space group with cell dimensions $a = 8.53$, $b = 10.30$, $c = 9.27 \pm 0.01$ Å. Extinctions for $h+k=2n+1$ for all data and $l=2n+1$ for the (H0l) reflections led to three possible choices for the space group— $Cmc2_1$, $C2cm$, or $Cmcm$. Details of the structure determination have been reported elsewhere (Majeste & Trefonas, 1968).

The crystal used in the determination was cylindrical and very small (less than 0.1 mm in diameter) and thus no absorption corrections were made on the intensities. The data were taken on a precession camera with the use of multiple film techniques with the exposure times on the films within a given zone varying by a factor of four. A rapid drop-off of intensity with increasing $\sin \theta/\lambda$, characteristic of the disorder shown to exist in the crystal, severely limited the amount of data taken, even though exposure times of 72 hours were used for the darkest film of each zone. Ilford Type B film was used so that the background was minimal even with these long exposure times.

A total of 289 independent reflections were measured by visual estimation. Eighty-one of the 289 spots were interpreted as having intensities less than, or equal to, the minimum value on the scale. These were divided into three categories based on their intensities. The first category was assigned an intensity value of 1.0, the second a value of 0.6 and the third a value of 0.3. Each category was then assigned a separate scale factor and the data were refined by four separate scale factors,

unless noted as otherwise. A total of 305 independent reflections from the same films was obtained with the microdensitometer. The additional reflections were obtained in the latter case since all available lattice points were sampled and all whose optical transmissions were less than 0.97 were retained.

Using the *visually* estimated data, least-squares refinement resulted in values of the reliability indices of $R, r = 0.20_7, 0.23_4$ (Busing, Martin & Levy, 1962)*. Refinement terminated at this point since our program contains a subroutine which ends a run when the largest shift in both temperature factors and coordinates becomes less than 0.00001. The resulting model was far from satisfactory chemically and unsatisfactory crystallographically since previous studies in this laboratory with visually estimated data invariably refined to values of $R = 0.12$ or less (Zacharis & Trefonas, 1968)†.

A variety of attempts was made to improve the results. For example:

1. The data were re-estimated visually. The results were so close to the previous results that any changes would merely reflect the bias of the experimenter.

2. A portion of the set of data from a previously completed compound was read to check the accuracy of the scale. No significant contradictions resulted.

3. The scale itself was measured on the microdensitometer, and proved to be linear over the entire range.

4. Various alternative choices for the atomic positions led to even poorer results.

5. These same 'best' positions were refined using only the data in Group I, that is, the 209 reflections whose intensity readings were greater than the minimum. This changed the reliability indices to values of $R, r = 0.21_0, 0.19_4$ but did not substantially change the chemical model.

About this time, a Nonius integrating microdensitometer became available in the laboratory and it was decided to re-evaluate these same films. With these *optically* evaluated data, the Patterson map was recalculated and confirmed the previous choice of chlorine position. Starting with the 'best' model from the previous data, least-squares refinement led to values of $R, r = 0.10_7, 0.10_2$ where the run then terminated. The resulting model was chemically reasonable with estimated standard deviations of the order of magnitude of 0.03 Å in the bond distances. It was thus accepted as the correct solution.

We were curious as to whether the visually estimated data could have led to the correct structure if our original model had been closer to the final result. Thus the final model (with values of $R, r = 0.10_7, 0.10_1$) was used as a trial model with the visually estimated data. To

* R is the usual reliability index and

$$r = [\sum w(F_o - kF_c)^2 / \sum wF_o^2]^{1/2}$$
 $w = 1.0$ unless otherwise stated

† This paper is representative of recent structure determinations from this laboratory in which visually estimated data were used.

our surprise, it did *not* refine! After adjustment of scale factors, the final model, compared to the visually-read data, resulted in values of R , $r=0.29_4, 0.28_5$.

Discussion and conclusion

Since the 'correct' magnitudes of the structure factors were available to us (namely, the $|F_c|$ values from the chemically reasonable structure which refined to values of $R, r=0.10_7, 0.10_1$), an analysis of the two sets of experimental data was undertaken to determine the source of error in the one set. In the following tables, F_o^m refers to the data measured by the integrating microdensitometer; F_o^s refers to the data estimated visually from the same films, with a scale constructed from timed exposures of a given reflection; and F_c refers to the calculated structure factors resulting from the final refined model.

First, an analysis of the microdensitometer data was undertaken to insure that no artifacts existed in the final refined model which was accepted as correct. Table 1 lists R values (R being the usual reliability index) comparing the F_o^m and the F_c lists with the data arranged by parity groups. Except for the comparison of l -even and l -odd groups, all other values are approximately the same. The slightly larger variation between these two groups results from the special position of the heavy atom at $z=\frac{1}{2}$, resulting in the l -even structure factors having a large contribution from the chlorine atom, whereas the l -odd data are essentially independent of the chlorine atom. Table 2 lists a similar comparison of this data except that it is now arranged in the order of the actual zones which were photographed. Again,

an overall uniformity is observed in all zones, with the possible exception of the $(h, k, 0)$ zone. Both Tables confirm the correctness of the structure and the reliability of the microdensitometer data.

Table 1. Analysis of data by parity groups for the final structure

Unique parity group	Number of reflections*	R^m
h odd	138	0.115
h even	145	0.101
l odd	143	0.134
l even	140	0.086
$h+l$ odd	143	0.114
$h+l$ even	140	0.101
All data	283	0.109

* Includes only data common to both F_o^m and F_o^s lists.

$$R^m = \frac{\sum ||F_o^m| - |F_c||}{\sum |F_o^m|}$$

Table 2. Analysis of data by zones for the final structure

Zone		Number of reflections*	R^m
I	$(h \ k \ 0)$	44	0.073
II	$(h \ 3h \ l)$	38	0.106
III	$(h \ k \ 1)$	43	0.092
IV	$(h \ k \ 2)$	31	0.132
V	$(h \ k \ 3)$	20	0.096
VI	$(h \ k \ 4)$	40	0.113
VII	$(h, 3h-2, l)$	55	0.116
VIII	$(h, 3h-4, l)$	37	0.103
IX	$(h, 3h-6, l)$	32	0.110
X	$(h, 3h-8, l)$	19	0.096
XI	$(h, 3h-10, l)$	19	0.123

* Includes only data common to both F_o^m and F_o^s lists.

$$R^m = \frac{\sum ||F_o^m| - |F_c||}{\sum |F_o^m|}$$

Table 3. Analysis of both data sets arranged in order of decreasing magnitudes of F_c

Group	Number of reflections	R^m	R^s	$\sum F_o^m $	$\sum F_o^s $	$\sum F_c $	k
1	28	0.076	0.269	695.4	927.6	678.3	1.368
2	28	0.101	0.127	293.7	308.2	312.7	0.986
3	28	0.114	0.222	229.3	213.6	246.5	0.867
4	29	0.121	0.320	174.3	147.3	194.5	0.757
5	29	0.115	0.443	139.9	103.0	148.8	0.692
6	29	0.115	0.444	124.7	83.6	120.8	0.692
7	28	0.145	0.630	91.9	58.1	94.8	0.613
8	28	0.127	0.720	81.3	48.2	81.0	0.595
9	28	0.170	0.663	70.2	37.5	62.4	0.601
10	28	0.297	0.397	59.2	32.6	44.6	0.733

$$R^m = \frac{\sum ||F_o^m| - |F_c||}{\sum |F_o^m|}, \quad R^s = \frac{\sum ||F_o^s| - |F_c||}{\sum |F_o^s|}, \quad k = \frac{\sum |F_o^s|}{\sum |F_c|}$$

Table 4. Analysis of both data sets after adjustment of intra-zone constants

Group	Number of reflections	Criterion	R^s	R^m
1	37	Read on lightest film only	0.095	0.096
2	15	Read on first two films	0.130	0.088
3	19	Read on last two films	0.161	0.077
4	134	Read on last film only	0.136	0.118
5	8	Visible - group A	0.182	0.170
6	29	Visible - group B	0.272	0.236
7	41	Visible - group C	0.242	0.198
		All data	0.145	0.107

At this stage, a common list consisting of $|F_o^m|$, $|F_c|$, and $|F_o^s|$ was prepared. The resulting cards were categorized and sorted according to a variety of schemes and R values calculated for each category. Sorting by zones photographed, parity groups, or as a function of $\sin \theta/\lambda$ did not indicate any systematic pattern which would explain why the F_o^s list refined to an obviously incorrect structure if treated alone, and resulted in a value of $R=0.294$ if the final model was used. However, when this combined list was sorted in order of decreasing magnitude of F_c values, the difficulty with the F_o^s list was immediately apparent. If the data are divided into ten groups, each containing the same number of reflections, a scaling factor difference becomes immediately obvious. Table 3 lists this comparison with the arbitrary scale factor which would correct the F_o^s data in each group.

The most intense reflections are badly overscaled and the lighter reflections are badly underscaled, indicating that the film constants between the four films within each zone are incorrect. To test this hypothesis, the data were sorted on the basis of the given film(s) in which it was observed and new scaling factors calculated by comparing the sums within each of these categories to the sums of the $|F_c|$ for these same categories. Table 4 shows that if this is done, even in an approximate manner, the value of R for the scale-read data compared to the final model drops from a value of 0.297 to 0.145. Hence, the systematic error can be pinpointed.

It is instructive to look back and see under what conditions an investigator would be cautioned to proceed carefully to avoid a similar occurrence in a future investigation. Our primary problem was caused by the rapid drop-off of intensity in the structure factors with increasing $\sin \theta/\lambda$, as a result of disordering. To get even a minimal amount of data this meant that very long time exposures had to be used to record the very light reflections. However, for the reflections at low $\sin \theta/\lambda$ values (also the most intense ones), relatively short exposure times were required. Since only a very few spots had readable values on more than one film, our correlation constants between films in a given zone

were based on a very few reflections. Errors in correlation constants between films would have a cumulative effect and would result in a large discrepancy between the most intense spots, read on the first film, and the light spots read on the last film.

Since least-squares techniques normally rely on the sum of all reflections, the fewer intense spots would essentially balance the contribution of a larger number of weak spots and the least-squares refinement would result in an intermediate scale factor. This intermediate scale factor would (as it turned out in this case) be a poor approximation to either category of data. The result would then be a shifting of the atomic parameters, coordinates and temperature factors, to compensate for this error. The result, as we have seen, would be disastrous.

Finally, an analysis of our data was made in an attempt to compare the relative merits of empirical scale factors versus time factors when one is correlating a set of timed exposures within a given zone. Ratios of empirical scale factors to the time factors for each pair of films were calculated for our original visual data and for this same data after re-scaling. All of these scale factor ratios were then averaged, giving the results summarized in Table 5. The corrected scale factors average to the same value as the time factors within 12%, whereas this ratio for the incorrectly scaled data differs from unity by 38%. The scatter in this ratio even for the correctly scaled data is sufficiently large that our own preference would be to continue to use empirical scale factors in such correlations. However, the wide variation in these two ratios suggests a very rapid way of estimating whether the intra-zone scaling has been done correctly. To test this idea further, such ratios have been calculated for six other compounds whose structures had been previously determined. In all of these cases, the ratio remains within 15% of unity and in all cases but one the average deviation of these ratios is larger than the difference between the ratio and unity.

The authors wish to express their gratitude to the National Institutes of Health (GM-08348-08) for finan-

Table 5. Comparison of empirical scaling constants and time factors

Reference	Number of reflections	Ratio*	Average deviation of ratio
This study - <i>incorrectly</i> scaled data	289	(1.38)	(0.17)
This study - <i>correctly</i> scaled data	289	1.12	0.14
Zacharis & Trefonas (1968).	844	1.07	0.14
Trefonas & Sato (1966).	1107	0.97	0.26
Trefonas, Flurry, Majeste, Meyers & Copeland (1965).	93	1.01	0.13
Trefonas & Majeste (1965).	1272	1.02	0.08
Trefonas & Towns (1964).	734	1.14	0.09
Trefonas & Majeste (1963).	554	1.01	0.13

* The ratio is defined as $\frac{\sum \left(\frac{k_c}{k_T} \right)_n}{n}$. k_c = empirical scaling factor, k_T = time factor, n = total number of pairs of films related by a time factor.

cial assistance in support of this and subsequent studies; to Dr A. Ebnoether and Sandoz Limited for samples of the compound; to the Computer Centers at LSUBR and LSUNO (N.S.F. No. GP-2964).

References

- BUSING, W. R., MARTIN, K. O. & LEVY, H. A. (1962). *ORFLS*. ORNL-TM-305. Oak Ridge National Laboratory, Tennessee.
- MAJESTE, R. & TREFONAS, L. M. (1968). *J. Heterocyclic Chem.* **5**, 663.
- TREFONAS, L. M., FLURRY, R. L., MAJESTE, R., MEYERS, E. A. & COPELAND, R. F. (1966). *J. Amer. Chem. Soc.* **88**, 2145.
- TREFONAS, L. M. & MAJESTE, R. (1963). *Tetrahedron*, **19**, 929.
- TREFONAS, L. M. & MAJESTE, R. (1965). *J. Heterocyclic Chem.* **2**, 80.
- TREFONAS, L. M. & SATO, T. (1966). *J. Heterocyclic Chem.* **3**, 404.
- TREFONAS, L. M. & TOWNS, R. (1964). *J. Heterocyclic Chem.* **1**, 19.
- ZACHARIS, H. M. & TREFONAS, L. M. (1968). *J. Heterocyclic Chem.* **5**, 343.

Acta Cryst. (1970). **B26**, 782

Molecular Structures of Amino Acids and Peptides.

I. The Crystal Structure and Conformation of DL-O-Serine Phosphate Monohydrate. Very Short Phosphate-Phosphate Hydrogen Bonds

BY E. PUTKEY* AND M. SUNDARALINGAM†

Department of Chemistry, Case Western Reserve University, Cleveland, Ohio 44106, U.S.A.

(Received 4 June 1969)

DL-O-Serine phosphate monohydrate crystallizes from aqueous solution in the space group $C2/c$ with unit-cell dimensions $a = 18.473 \pm 0.0006$, $b = 8.327 \pm 0.0002$, $c = 12.300 \pm 0.0004$ Å, $\beta = 120.72 \pm 0.0015^\circ$. For eight formula units of the monohydrate per unit cell, the calculated density of 1.661 g.cm^{-3} is in agreement with the observed density of 1.668 g.cm^{-3} . The crystal structure was determined by the heavy-atom method using 1092 observed diffractometer intensities. Isotropic and anisotropic full-matrix least-squares refinement gave an R index of 0.03. The molecule exists as a zwitterion,

$\text{HO}_3\text{POCH}_2\text{CH}(\text{NH}_3^+)\text{COOH}$. The amino group is protonated by a phosphate proton and the carboxyl group is not ionized. The ester C-O bond is *gauche* both to the C-N bond and the C(1)-C(2) bond. The two shortest hydrogen bonds between center-related (2.498 Å) and diad-related (2.506 Å) phosphates generate a right-handed spiral. These hydrogen bonds are symmetrical with either a single or equivalent double minima. The P-O bonds (1.530 and 1.522 Å) associated with the short hydrogen bonds are intermediate in character to the normal P-OH (1.554 Å) and the P=O (1.490 Å) bond distances. A relatively short hydrogen bond, $-\text{C}-\text{O}-\text{H} \cdots \text{O}=\text{P}-$, of 2.595 Å occurs between the carboxyl and the phosphate groups.

Introduction

Detailed knowledge of the structural and hydrogen-bonding properties of biological phosphates is of considerable importance in understanding their hydrolytic behavior and their role in biological phosphorylation and related processes. The structure analysis of DL-serine phosphate monohydrate was undertaken as part of a program of studies, currently in progress in our laboratory, on the structures of amino acid derivatives and cell-membrane components. A preliminary report on the structure has already been pub-

lished (Putkey & Sundaralingam, 1968). Serine phosphate is unique among the amino acid phosphates, for it occurs both in the membrane proteins and the membrane phospholipid, phosphatidylserine. It has been suggested that this versatility of serine phosphate is probably a result of its β -hydroxyamine chain, $^+\text{N}-\text{C}-\text{C}-\text{O}$, which is a common structural feature of the major phospholipids, sphingolipids, anaesthetics, cholinergists and other nerve and brain amines (Sundaralingam, 1968). In batrachotoxinin A, a steroidal alkaloid from the venom of the Columbian poison arrow frog, an N -substituted $^+\text{N}-\text{C}-\text{C}-\text{O}$ chain links the 13 and 14 positions of the steroid (Tokuyama, Daly, Witkop, Karle & Karle, 1968). Indeed, the $^+\text{N}-\text{C}-\text{C}-$ and the $^+\text{N}-\text{C}-\text{C}-\text{O}$ chains are common structural features of various alkaloids, brain amines and membrane molecules.

* Present address: The Marianists, Chaminade College, 3140 Waiialae Avenue, Honolulu, Hawaii 96816, U.S.A.

† To whom correspondence should be addressed. Present address: Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53706, U.S.A.