Investigations into the Structures of Biologically Important Molecules.

I. The Crystal Structure of Monoclinic Calcium Pyrophosphate Tetrahydrate.

II. The Purification of Glucose-6-Phosphate Dehydrogenase from Human Red Blood Cells.

III. A Partial Solution of the Crystal Structure of p-Trimethylammoniumbenzene Sulfonate.

Thesis by

Nicki L. Davis

In Partial Fulfillment of the Requirements

for the Degree of

Master of Science

California Institute of Technology
Pasadena, California

1977

(Submitted September 15, 1977)

ACKNOWLEDGMENTS

The author wishes to thank the following people for their help in making this research possible:

Dr. Akira Yoshida and Mrs. Vibha Davé of the City of Hope National Medical Center for aid in the G6PD purification project.

The members of the Dickerson research group for many helpful suggestions.

Dr. Mike Schimerlik for the use of the blue dextran sepharose.

Finally, special thanks to Dr. Richard E. Dickerson for considerable help and encouragement; to Drs. Neil and Gretchen Mandel for their constant help and suggestions; to Mrs. Lillian Casler for the graphs and drawings in this work; and to Mrs. Dee Barr for typing the manuscript.

ABSTRACT

Chapter I deals with the solution and refinement of the crystal structure of monoclinic calcium pyrophosphate tetrahydrate. The role of the triclinic dihydrate of calcium pyrophosphate as the cause of inflammation in pseudogout is briefly discussed, and a detailed comparison is made between its crystal structure and that of the monoclinic tetrahydrate, which does not cause disease.

In Chapter II, methods for the isolation of the enzyme glucose-6-phosphate dehydrogenase from human red blood cells are described. The problems involved in adapting a small-scale purification prodecure to large-scale production are investigated, and solutions proposed. Alternative methods for the purification of the enzyme are developed and tested.

A partial solution of the crystal structure of p-trimethylammoniumbenzenesulfonate (ZWT) is given in Chapter III. This compound is the product of a solid state rearrangement of methyl-p-dimethylaminobenzene sulfonate (MSE). Although the structure of ZWT could not be refined, a trial structure obtained by a direct methods calculation bore a close resemblance to the known structure of the parent compound, MSE. This result was in accord with previous data on the details of the rearrangement reaction.

TABLE OF CONTENTS

Chapter I

THE CRYSTAL STRUCTURE OF MONOCLINIC CALCIUM PYROPHOSPHATE TETRAHYDRATE

INTRODUCTION

The crystal structure of the title compound was undertaken as part of a comparative structural study of calcium pyrophosphates. The study was undertaken to aid in the interpretation of the infrared spectra of pyrophosphates and, secondly, to study the role of the two dihydrates of calcium pyrophosphate in the disease chondrocalcinosis (more commonly known as pseudogout).

Crystals of the monoclinic and triclinic forms of calcium pyrophosphate dihydrate are able to bind to, and cause rupture of, the lysosomal membranes of white blood cells. Lysosomes are membrane-bound packets of digestive enzymes found within the leukocytes. When the lysosomes are ruptured, the lysosomal enzymes spill out into the parent leukocyte, inducing cell death and precipitating the inflammation characteristic of the disease (1). Neither the orthorhombic nor the monoclinic forms of the tetrahydrate are formed in vivo; however, preliminary experiments seem to indicate that the orthorhombic tetrahydrate does cause membranolysis in vitro, while the monoclinic form does not. The crystal structures of these varied pyrophosphates should provide valuable clues as to how compounds so similar in chemical composition can exemplify such varied membranolytic activity.

EXPERIMENTAL

Crystals of the monoclinic form of calcium pyrophosphate tetrahydrate, $Ca_2P_2O_7 \cdot 4H_2O$, were grown according to the method of Brown et al. (2) and were kindly supply by Dr. J. R. Lehr of the Tennessee Valley Authority. The composition of the crystals used in the present study was verified with powder diffraction photographs taken on a Guinier-Hägg high-resolution powder camera designed and built by Dr. S. Samson with the aid of machine drawings supplied by Professor G. Hägg of the University of Uppsala, Sweden.

The crystals were small (less than 1 mm in their longest dimension), tablet-shaped, and appeared to have a monoclinic crystal habit. A single crystal of approximate dimensions 0.5 mm x 0.1 mm x 1 mm was chosen and mounted along the longest axis. Weissenberg and precession photographs indicated that the crystal was indeed monoclinic, and that the longest axis corresponded to the true crystallographic c axis. The space group assignment of $P2_1/c$ was confirmed by an examination of the diffractometer data; none of the systematically absent reflections had an intensity greater than $4\sigma(I)$. Approximate cell constants determined from photographs were later refined by performing a least-squares calculation on the thirteen hand-centered reflections used for the orientation matrix of the diffractometer. Crystal data are given in Table I.

Intensity data were collected on a General Electric automated diffractometer using nickel-filtered CuKα radiation. Each reflection was scanned by the θ-2θ technique at a scan speed of ½° per minute,

with background counts of twenty seconds. The scan range varied linearly from 2° at 2θ = 5° to 3° at 2θ = 150°. A total of 2004 unique reflections (excluding systematic extinctions) were collected within a 2θ range of 5° - 150°. Three check reflections, staggered ten reflections apart, were measured every thirty reflections. The intensities of these reflections varied less than 1% throughout the data collection.

The observed F's were derived from scan counts corrected for Lorentz (L) and polarization (p) effects*, according to the formula:

$$I = \left(\frac{1}{Lp}\right) [N - (b_1 + b_2)(t_s/2t_b)]$$

where N = total scan count, b_1 and b_2 are background counts, t_s = total scan time and t_b = time of background counts. Variances $\sigma^2(I)$ for each of the intensities were calculated in accordance with counting statistics from the formula:

---

*For unmonochromatized radiation the polarization correction is

$$p = \frac{1 + \cos^2 2\theta}{2}$$

and the Lorentz correction is

$$L = \frac{1}{\sin 2\theta}$$

Therefore the combined expression for 1/Lp is

$$\frac{1}{Lp} = \frac{2 \sin 2\theta}{(1 + \cos^2 2\theta)}$$

$$\sigma(I) = \left(\frac{1}{Lp}\right) \left[N + (t_s/2t_b)^2 (b_1 + b_2) + (0.02N)^2\right]^{\frac{1}{2}}$$

where 0.02N is an empirical term related to errors in counting statistics and instrument function. No correction was made for absorption ($\mu r_{max} = 6.0$). Reflections with intensities less than $4\sigma(I)$ were considered unobserved. By this criterion, the data set contained 1935 observed reflections.

The observed F's and their standard deviations were put on an absolute scale by multiplying them by the scale factor, k, determined from a Wilson plot. The Wilson plot is a means of both scaling the intensities and deriving the average temperature factor, B, of the structure. It can be shown that the average relative (observed) intensity, $\overline{I}_{rel}$, is related to the theoretical (or absolute) intensity by the expression

$$\overline{I}_{rel} = C\overline{I}_{abs} = C \exp\left[-2B(\sin^2\theta)/\lambda^2\right] \sum_{i=1}^{N} f_{o_i}^2$$

where the f_o's are the atomic scattering factors of the atoms of the structure and $C = 1/k^2$.

Dividing both sides of the equation by $\sum_{i=1}^{N} f_{o_i}^2$ and taking the natural logarithm of both sides, we get

$$\ln\left(\frac{\overline{I}_{rel}}{\sum_{i=1}^{N} f_{o_i}^2}\right) = \ln C - 2B(\sin^2\theta)/\lambda^2$$

Thus, a plot of ln $(\bar{I}_{rel} / \sum_{i=1}^{N} f_{o_i}^2)$ versus $(\sin^2\theta/\lambda^2)$ should give a straight line whose intercept is ln C and whose slope is equal to -2B. The Wilson plot for this particular data set is shown in Figure 1. The complete derivation of the expression for the Wilson plot is given in Appendix A.

The structure was solved by direct methods, using the program MULTAN74 (3). Three origin-defining reflections were chosen from a total of 190 reflections having E values greater than 1.43. Four more reflections were chosen and assigned arbitrary phases after an examination of the list of Σ2 interactions. The three origin-defining reflections and the four reflections assigned arbitrary phases are listed in Table II.

Sixteen sets of phases were calculated, and the phase set having the highest combined figure of merit (4,5) was used in the calculation of an E map. This E map yielded positions for all the heavy atoms in the structure except for a water oxygen which was found by successive Fourier syntheses.

The progress of the refinement was followed in three ways: 1) by noting the size of the shifts in the coordinates and temperature factors; 2) by monitoring the crystallographic R factor

$$R = \frac{\Sigma||F_o| - |F_{calc}||}{\Sigma|F_o|}$$

and 3) the goodness-of-fit

$$\text{G.O.F.} = \frac{\Sigma w \left(|F_o|^2 - |F_{calc}|^2\right)^2}{(M - S)}^{\frac{1}{2}}$$

where M = 2004 observations, S = number of variable parameters, and $w = 1/\sigma^2(F_o^2)$.

The fifteen heavy atoms were each assigned an isotropic temperature factor of 1.24 obtained from the Wilson plot and carried through four cycles of full matrix isotropic least-squares refinement. This was followed by two cycles of anisotropic refinement, resulting in an R factor of 0.059 and a goodness-of-fit of 3.42. Since the size of the coordinate shifts of the heavy atoms indicated that their refinement was very nearly complete, a difference Fourier was calculated at this point in an attempt to locate the water hydrogens. A total of 136 parameters (three positional and six anisotropic temperature factors per atom, plus a scale factor) were used in the determination of F_{calc}. The density map showed that each water oxygen was surrounded by diffuse regions of density, and none of these regions could with certainty be assigned to a hydrogen atom. It therefore became necessary to determine the most probable hydrogen-bonding network based on geometrical considerations, and use this in conjunction with the electron density map to find the hydrogen positions.

An examination of oxygen-oxygen bond distances revealed that each water oxygen was within hydrogen-bonding distance (< 3.1 Å) of three to four oxygen atoms which could serve as hydrogen-bond acceptors. $O—O_W—O$ bond angles were calculated for all possible pairings of likely hydrogen-bond acceptors: three pairs if there were three

possible H-bond acceptors, and six pairs if there were four. It was assumed that both water hydrogen atoms participated in hydrogen bonding. If the O—O_W—O bond angle deviated significantly from tetrahedral geometry, that particular combination of hydrogen-bond acceptors could be ruled out. Positions of hydrogen atoms for the remaining sets of donor-acceptor possibilities were calculated by means of the hydrogen position program from the Caltech CRYM crystallographic computing package, assuming an O—H bond distance of 1.0 Å and an H—O—H bond angle of 104.5°. The final choice of hydrogen positions was made by noting which of the calculated ("idealized") hydrogen pairs corresponded to the largest electron density in the difference Fourier synthesis.

Each hydrogen atom was assigned an isotropic temperature factor of 5.0. The 32 hydrogen parameters (coordinates plus temperature factors for the eight water hydrogens) were subjected to two cycles of full matrix isotropic refinement. The other 136 parameters were input for no refinement (i.e., used in the determination of F_{calc} and phase angle but not included in the least-squares calculation). The result of the least-squares calculation was that the hydrogen atom temperature factors increased to nearly three times their original values, while the coordinates were shifted so far from their original positions that the H—O—H bond anges and the O—H distances no longer corresponded to chemical reality.

At this point the secondary extinction factor, g, calculated from Eq. 3 of Larson

$$F^2_{obs} = F^2_c/1 + g\beta F^2_o$$

where $\beta = (1+\cos^4 2\theta)/\sin 2\theta(1+\cos^2 2\theta)$

was added to the list of parameters. This quantity and the scale factor were refined separately in two more cycles of full matrix least squares (other parameters input for no refinement) to final values of $5.5(5) \times 10^{-6}$ and 1.0018, respectively. A final attempt was made to refine the hydrogens, this time incorporating the refined scale factor and the secondary extinction factor into the calculation of F_{calc} and the phase angles. The results, however, were no better than before, and no further effort was made toward refining the hydrogen parameters. It was evident that the hydrogen atoms could not be refined to values which made chemical sense.

The 135 heavy-atom parameters were refined in two more cycles of full matrix least squares, while the idealized hydrogen coordinates and temperature factors, the refined scale factor, and the secondary extinction factor were put in for no refinement. Scattering factors were those listed in International Tables for X-ray Crystallography (7) except for hydrogen (8). The final R value was 0.056 and the goodness-of-fit, 3.06. Coordinates and temperature factors for all atoms are given in Table III.

DISCUSSION

Bond distances and angles for the pyrophosphate anion are given in Figure 2. Bond distances between phosphorus atoms and terminal oxygen atoms ($P—O_T$) range from 1.502 Å to 1.529 Å with an average value of 1.515 Å. The average value for the $O_T—P—O_T$ bond angles is 110.9°, while that of the $O_T—P—O_B$ angles is 106.9°; thus the arrangement of oxygen atoms around the phosphorus atom is somewhat distorted from true tetrahedral geometry. The values of the $P—O_B$ bond distance (1.619 Å) and the $P—O_B—P$ bond angle (126.3°) are in agreement with previous findings (1). The conformation of the pyrophosphate anion is very nearly eclipsed (Figure 3) with an average $O_T—P—P—O_T$ torsion angle of 5.74°. The two terminal oxygen atoms which are cis to the bridging oxygen have a smaller torsion angle than the trans pairs.

The environments of the terminal oxygen atoms are given in Table IV. Four of the terminal oxygen atoms are bidentate, coordinating to two calcium atoms each, while the remaining two are monodentate. The bridging oxygen does not participate in either calcium coordination or hydrogen bonding. The two ends of the molecule are not strictly equivalent: for instance, the terminal oxygen atoms bonded to P(2) have longer $P—O_T$ bond distances as well as shorter Ca—O contacts, and also accept more hydrogen bonds from water molecules than do oxygens attached to P(1). However, it was not possible to derive a rigorously quantitative relationships between P—O and Ca--O, and $O--O_W$ bond distances.

Details of the calcium coordination are given in Table V and illustrated in Figure 4. Both calcium atoms are seven coordinated, capped octahedra. Five of the ligands are terminal pyrophosphate oxygens and two are water oxygens. As can be seen in the drawing, the "cap" of the Ca(1) octahedron is formed by two terminal oxygen atoms bonded to the same phosphorus atom. Each of these terminal oxygens also coordinates to another, symmetry related, calcium atom. The end result is a chain of Ca(1) coordination polyhedra running across the page; that is, parallel to the c axis of the crystal. The Ca(1) atoms in this chain are related by symmetry operations 1 and 2 (see Table V). Ca(2) atoms form similar chains running parallel to the a axis of the crystal and consisting of atoms related by centers of symmetry (i.e., by symmetry operations 1 and 3). Two other chains, consisting of Ca(1) atoms corresponding to equivalent positions 3 and 4, and Ca(2) atoms corresponding to equivalent positions 2 and 4, are found in the upper half of the unit cell (not shown).

Each of the four systems of symmetry-related calcium atoms is completely independent of the others; nonequivalent calcium atoms never share edges or vertices of their coordination polyhedra, nor do equivalent calcium atoms belonging to different systems of chains (e.g., the Ca(1) atoms beloning to equivalent positions 1 and 2, and those belonging to equivalent positions 3 and 4). This particular aspect of the metal-ion coordination is unique among all pyrophosphate structures solved so far. For instance, the calcium coordination of the triclinic form of calcium pyrophosphate dihydrate shown in Table VI (reprinted

from Mandel (1)) shows that nonequivalent calcium atoms share four different oxygen atoms between them. The sharing of ligands among nonequivalent metal atoms is quite common among other pyrophosphates, and indeed plays a major role in holding these structures together (1,9-19). But in the present compound, the only connection between the four systems of chains is through covalent bonds of the pyrophosphate groups, which are arranged nearly end-to-end along the b axis.

The stereo drawings in Figure 5 illustrate the role of the water molecules in the structure. The first drawing, Figure 5a, shows the complete structure; the second drawing, 5b, shows only the calcium atoms and pyrophosphate groups. It is clear that the role of the water molecules is to fill space. They lie in two systems of channels, one running parallel to the a axis and one running parallel to the c axis. Since each water oxygen is coordinated to only one calcium atom (see Table V), the water oxygens cannot serve as bridges between different calcium atoms. It is the coordination of the calcium atoms to the four bidentate terminal oxygens which holds the structure together. The two monodentate terminal oxygen atoms and the four water oxygens merely fill in the rest of the calcium coordination positions. This feature is unique among all known hydrates of pyrophosphates. For instance, in the structures of tetrapotassium pyrophosphate trihydrate (19), dipotassium dihydrogen pyrophosphate hemihydrate (20), tripotassium hydrogen pyrophosphate (21), disodium dihydrogen pyrophosphate hexahydrate (22), and tetrasodium pyrophosphate decahydrate (23), at least some of the water oxygen atoms are involved in bridging equivalent or nonequiv-

alent metal atoms.

The most probable hydrogen-bonding scheme is given in Table VI and is illustrated in Figure 6. Water molecules may donate hydrogen bonds to other water oxygens or to terminal pyrophosphate oxygens. The fact that the hydrogen positions could not be refined may indicate disorder. A certain amount of disorder in the positions of the water hydrogens (although not of the water oxygens) could presumably be tolerated, since hydrogen bonding does not contribute much to the stability of the structure.

The structure of the surface of the most prominent crystal face of the triclinic dihydrate is shown in Figure 7. A comparison with the most prominent crystal face of the monoclinic tetrahydrate (Figure 8) reveals some startling differences. The most prominent feature of the surface of the triclinic dihydrate is the series of "triads" of negatively charged pyrophosphate oxygens. The centers of the calcium atoms lie farther below the surface than the centers of the oxygen atoms. Since the oxygen atoms have a much larger van der Waals radius than the calcium atoms, the result is that the calcium atoms lie in pockets and are partially covered by the oxygen atoms. In the monoclinic tetrahydrate, however, the molecular configuration is quite different: the oxygen atoms lie lower than the centers of the calcium atoms. Taking into account the differences in van der Waals radii, we conclude that the surface must be fairly smooth and that the calcium atoms are fully exposed--quite a contrast to the pockmarked surface of the triclinic dihydrate. Whether these differences in surface confor-

mation are related to differences in membranolytic activity must await the solution of the structures of the monoclinic dihydrate and the orthorhombic tetrahydrate. Also, experiments with XPS (x-ray photoelectron spectroscopy) and electron diffraction should give a more accurate picture of the surface of the crystal faces. Nevertheless, the differences in surface conformation which are apparent in the first-order approximation are intriguing--the more so since they are consistent with the pattern observed by Mandel (24).

REFERENCES

1. Mandel, N.S. (1975). Acta Crystallogr. B31(6), 1730-1734.
2. Brown, E.H., J.R. Lehr, J.P. Smith, and W.A. Frazier (1963). Agric. Food Chem. 11, 214-222.
3. Main, P., M.M. Woolfson, L. Lessinger, G. Germain, J.-P. Declercq (1974). MULTAN74, A System of Computer Programmes for the Automation Solution of Crystal Structures from X-ray Diffraction Data.
4. Germain, G. and M.M. Woolfson (1968). Acta Crystallogr. B24, 91-96.
5. Declercq, J.-P., G. Germain, P. Main and M.M. Woolfson (1973). Acta Crystallogr. A29, 231-234.
6. Larson, A.C. (1967). Acta Crystallogr. 23, 664-665.
7. International Tables for X-ray Crystallography, Vol. III. Kynoch Press, Birmingham, 1962.
8. Stewart, R.F., E.R. Davidson and W.T. Stimson (1965). J. Chem. Phys. 42, 3175-3186.

9. Webb, N.C. (1966). Acta Crystallogr. 21, 942-948.

10. Robertson, B.E. and C. Calvo (1967). Acta Crystallogr. 22, 665-672.

11. Robertson, B.E. and C. Calvo (1968). Can. J. Chem. 46, 605-612.

12. Robertson, B.E. and C. Calvo (1970). J. Solid State Chem. 1, 120-133.

13. Calvo, C. (1965). Can. J. Chem. 43, 1139-1146.

14. Calvo, C. (1965). Can. J. Chem. 43, 1147-1153.

15. Calvo, C. (1967). Acta Crystallogr. 23, 289-295.

16. Calvo, C. (1968). Inorg. Chem. 7, 1345-1351.

17. Calvo, C. and P.K.L. Au (1969). Can. J. Chem. 47, 3409-3416.

18. McDonald, W.S. and D.W.T. Cruickshank (1967). Acta Crystallogr. 22, 43-48.

19. Dumas, Y. and J.L. Galigne (1974). Acta Crystallogr. B30, 390-395.

20. Dumas, Y., J.L. Galigne and J. Falgueirettes (1973). Acta Crystallogr. B29, 2913-2918.

21. Dumas, Y., J.L. Galigne and J. Falgueirettes (1973). Acta Crystallogr. B29, 1623.

22. Collin, R.L. and M. Willis (1971). Acta Crystallogr. B27, 291-302.

23. MacArthur, D.M. and C.A. Beevers (1957). Acta Crystallogr. 10, 428-432.

24. Mandel, N.S. (1976). Arthritis and Rheumatism 19, 439-445.

TABLE I

Crystal Data

$Ca_2P_2O_7 \cdot 4H_2O$	$\lambda(CuK_\alpha) = 1.54051$ Å
Space Group $P2_1/c$	FW = 326.16
$\underline{a}$ = 6.0062(7) Å	Z = 4
$\underline{b}$ = 25.0588(36) Å	$F_{000} = 648$
$\underline{c}$ = 6.8497(10) Å	$\rho_{calc} = 2.33$ g cm^{-3}
β = 109.73(1)°	$\mu = 60.21$ cm^{-1}

TABLE II. MULTAN Parameters

Origin-Defining Reflections

	h	k	ℓ	E	No. Interactions	Phase Angle
1	2	14	1	3.71	79	360°
2	3	16	$\bar{5}$	3.50	50	360°
6	3	7	$\bar{6}$	3.04	44	360°

Starting Set of Reflections

	h	k	ℓ	E	No. Interactions	Phase Angle
9	2	17	3	2.96	41	180°
11	2	14	$\bar{3}$	2.93	59	180°
19	4	8	2	2.79	37	180°
29	2	9	0	2.39	36	180°

TABLE III. Atomic coordinates and temperature factors. The heavy-atom coordinates have been multiplied by 10^4; anisotropic temperature factors and hydrogen coordinates have been multiplied by 10^3. The anisotropic temperature factor has the form exp $[-2\pi(U_{11}h^2a^* + \ldots + 2U_{12}hka^*b^* + \ldots)]$. Each hydrogen atom was assigned an isotropic temperature factor of 5.0.

	x	y	z	U_{11}	U_{22}	U_{33}	U_{12}	U_{13}	U_{23}
CA(1)	8991(2)	7927(1)	8981(1)	20(1)	9(1)	12(1)	1(1)	5(1)	2(1)
CA(2)	1983(2)	9578(1)	3995(1)	17(1)	9(1)	15(1)	1(1)	5(1)	1(1)
P(1)	8053(2)	9172(1)	6430(2)	18(1)	6(1)	11(1)	1(1)	4(1)	0(1)
P(2)	9341(2)	8335(1)	4113(2)	18(1)	5(1)	13(1)	- 1(1)	6(1)	- 2(1)
O(1)	10002(6)	9546(1)	6341(5)	30(2)	17(2)	27(2)	- 9(1)	16(1)	- 7(1)
O(2)	5831(6)	9491(1)	6045(5)	22(2)	22(2)	28(2)	9(1)	1(1)	- 6(1)
O(3)	8791(7)	8844(1)	8396(4)	54(2)	12(1)	10(1)	6(1)	8(1)	4(1)
O(4)	7498(5)	8757(1)	4512(4)	23(2)	10(1)	13(1)	2(1)	5(1)	- 3(1)
O(5)	9839(6)	7914(1)	5823(5)	32(2)	9(1)	18(1)	3(1)	11(1)	3(1)
O(6)	8036(6)	8077(1)	2041(4)	27(2)	24(2)	13(1)	- 5(1)	6(1)	-11(1)
O(7)	11526(5)	8643(1)	4159(5)	21(2)	13(1)	18(1)	- 3(1)	10(1)	- 1(1)
O(8)	3136(6)	9457(1)	957(5)	31(2)	28(2)	21(2)	- 4(1)	8(1)	- 4(1)
O(9)	3004(6)	8111(1)	1207(5)	23(2)	30(2)	32(2)	0(1)	6(1)	- 8(1)
O(10)	8046(6)	9513(1)	1226(5)	30(2)	22(2)	19(1)	1(1)	8(1)	- 1(1)
O(11)	4938(7)	7980(2)	7234(7)	27(2)	35(2)	52(2)	- 1(2)	- 3(2)	4(2)
H(1)	172(0)	924(0)	15(0)						
H(2)	281(0)	982(0)	26(0)						
H(3)	478(0)	812(0)	164(0)						
H(4)	266(0)	831(0)	235(0)						
H(5)	823(0)	926(0)	15(0)						
H(6)	633(0)	948(0)	107(0)						
H(7)	428(0)	761(0)	689(0)						
H(8)	376(0)	821(0)	617(0)						

TABLE IV. Terminal oxygen atom environment. All distances are in angstroms. In addition to calcium coordination distances, the table lists all distances between terminal oxygen and water oxygen ($O_T \cdots O_W$) atoms which are less than 3.0 Å. Values not enclosed by parentheses represent $O_W \text{---} O_T$ contacts which were chosen for the most likely hydrogen bonding scheme (see Table VII). Values in parentheses represent less likely, but possible, hydrogen bonds. Standard deviations for the $Ca \cdots O_T$ bonds are approximately 0.003 Å. The oxygen-oxygen distances have standard deviations of about 0.004 Å.

O_T	$Ca \cdots O_T$		$O_T \cdots O_W$	
O(1)	2.302	2.470	(2.896)	
O(2)	2.275	2.682		
O(3)	2.340		2.712	
O(5)	2.384	2.420	(2.888)	
O(6)	2.382	2.812	2.881	
O(7)	2.368		2.804	2.911

TABLE V. Calcium coordination distances. The quantities in brackets are unit cell translations and equivalent position numbers, where equivalent position 1 is (x,y,z); 2 is (x,½-y,½+z); 3 is ($\bar{x}$,$\bar{y}$,$\bar{z}$); and 4 is ($\bar{x}$,½+y,½-z). All distances are in angstroms, and standard deviations are of the order of 0.002 Å.

	Ca(1)			Ca(2)	
O(3)	[000,1]	2.330	O(1)	[$\bar{1}$00,1]	2.302
O(5)	[000,1]	2.384	O(1)	[121,3]	2.470
O(5)	[010,2]	2.420	O(2)	[000,1]	2.275
O(6)	[001,1]	2.382	O(2)	[121,3]	2.682
O(6)	[010,2]	2.812	O(7)	[$\bar{1}$00,1]	2.368
O(9)	[101,1]	2.422	O(8)	[000,1]	2.424
O(11)	[000,1]	2.326	O(10)	[$\bar{1}$00,1]	2.487

TABLE VI. Calcium coordination distances for the triclinic dihydrate of calcium pyrophosphate. The values in brackets are unit-cell translations and equivalent position numbers for the triclinic space group $P\bar{1}$ (1 = x,y,z and 2 = $\bar{x},\bar{y},\bar{z}$). Variances are of the order of 0.001 Å.

Ca(1)			Ca(2)		
O(2)	[$\bar{1}$oo,1]	2.642	O(1)	[010,1]	2.285
O(2)	[111,2]	2.366	O(2)	[211,2]	2.463
O(3)	[$\bar{1}$00,1]	2.447	O(3)	[210,2]	2.320
O(5)	[000,1]	2.362	O(6)	[000,1]	2.339
O(6)	[111,2]	2.379	O(7)	[210,2]	2.450
O(7)	[110,2]	2.386	O(8)	[000,1]	2.668
O(8)	[000,1]	2.608	O(9)	[111,2]	2.526

TABLE VII. Water molecules and hydrogen bonds. Distances between water oxygen and hydrogen-bond acceptor atoms ($O_W \cdots A$) are shown for all possible hydrogen bond acceptors. Distances between water hydrogen and acceptor atoms ($O\text{–}H \cdots A$) are shown for the most probable hydrogen bonding scheme. All distances are in angstroms. Standard deviations for the $O_W \cdots A$ distances are about 0.005 Å; for the $O\text{–}H \cdots A$ distances, 0.003 Å. Symbols for unit-cell translations and symmetry operations are explained in Table IV.

Water Oxygen	H-bond Acceptor		$O_W \cdots A$ Distance	Water Hydrogen	$O\text{–}H \cdots A$ Distance
O(8)	O(3)	[$\bar{1}0\bar{1}$,1]	3.026	H(1)	2.020
	O(10)	[120,3]	2.947	H(2)	1.943
	O(10)	[000,1]	2.895	---	---
O(9)	O(6)	[000,1]	2.881	H(3)	1.887
	O(7)	[$\bar{1}$00,1]	2.804	H(4)	1.806
	O(11)	[01$\bar{1}$,2]	2.962	---	---
O(10)	O(3)	[00$\bar{1}$,1]	2.712	H(5)	1.707
	O(8)	[000,1]	2.895	H(6)	1.892
	O(1)	[221,3]	2.896	---	---
	O(8)	[120,3]	2.947	---	---
O(11)	O(9)	[010,2]	2.962	H(7)	1.955
	O(7)	[$\bar{1}$00,1]	2.911	H(8)	1.904
	O(5)	[$\bar{1}$00,1]	2.888	---	---

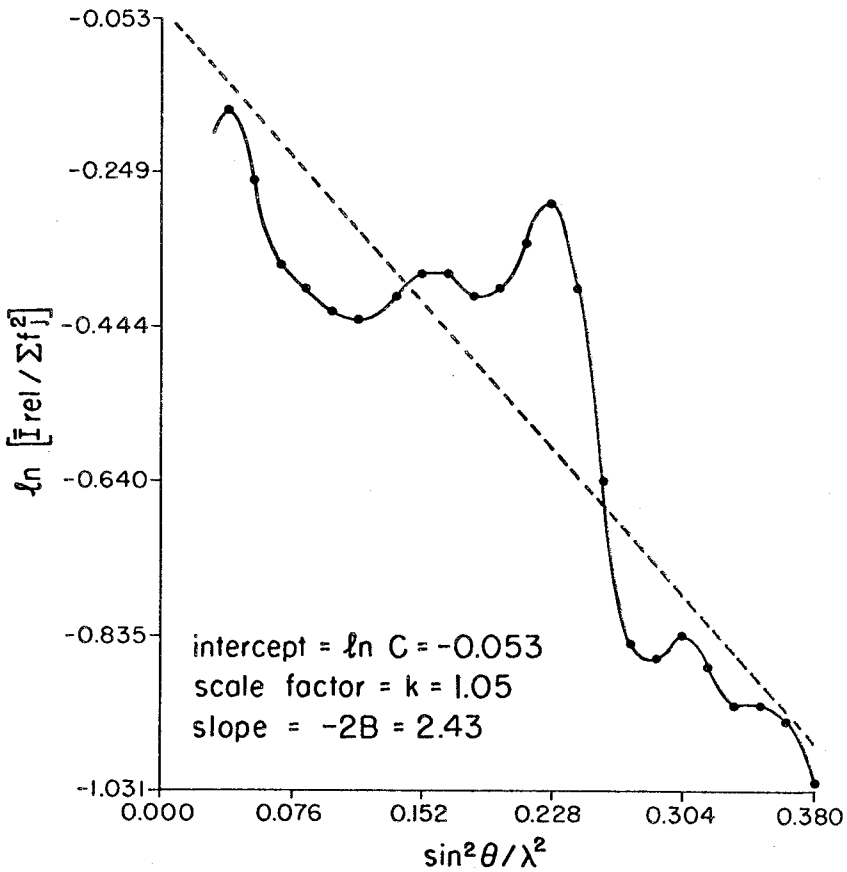


Figure 1. The Wilson Plot for the CuKα data set. ●—● = data points, and --- = least-square line through data points.

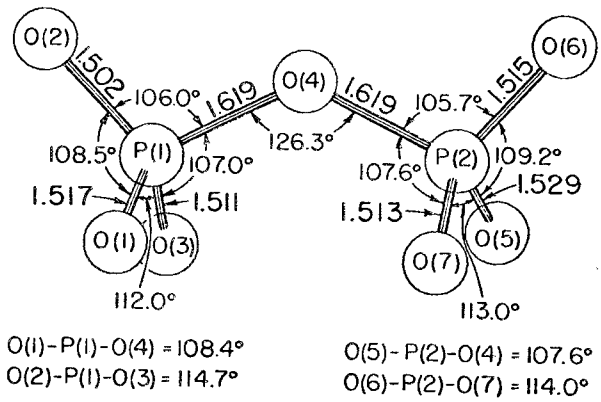


Figure 2. Bond angles and distances for the pyrophosphate anion in calcium pyrophosphate tetrahydrate (monoclinic).

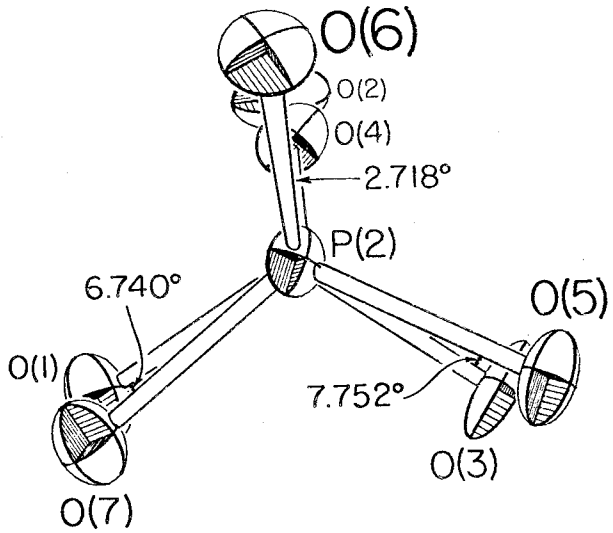


Figure 3. The pyrophosphate anion viewed down the P—P direction, showing the O_T—P—P—O_T torsion angles. The thermal ellipsoids are drawn at the 50% probability level.

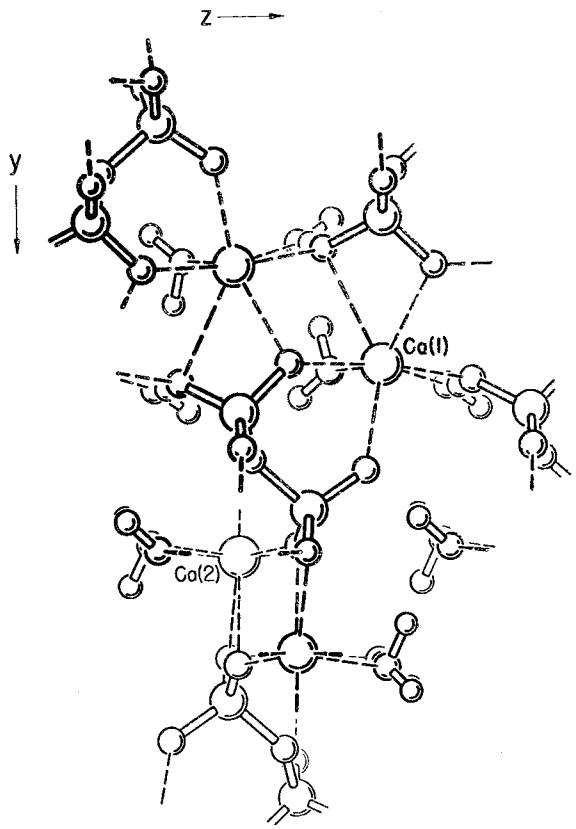


Figure 4. A view down the [100] direction showing the calcium coordination in the tetrahydrate of calcium pyrophosphate. Only the part of the cell extending from y = ½ to y = 1¼ is shown.

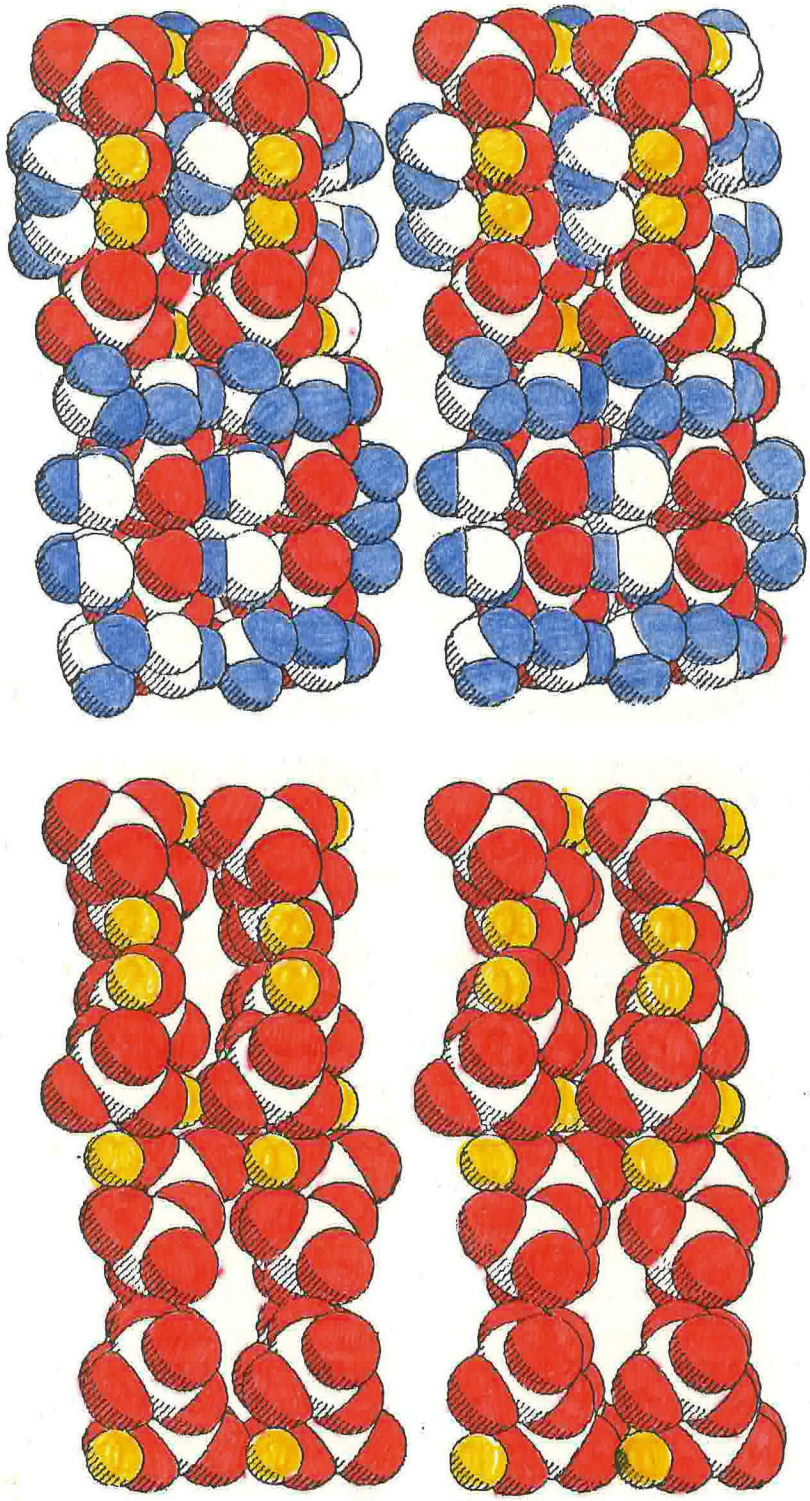


Figure 5. Space-filling representation of the present structure as viewed down the [001] direction. Calcium ions are yellow, pyrophosphate oxygens are red and water hydrogens are blue. (a) shows the complete structure; (b) shows the structure without the water molecules. The complete unit cell is depicted.

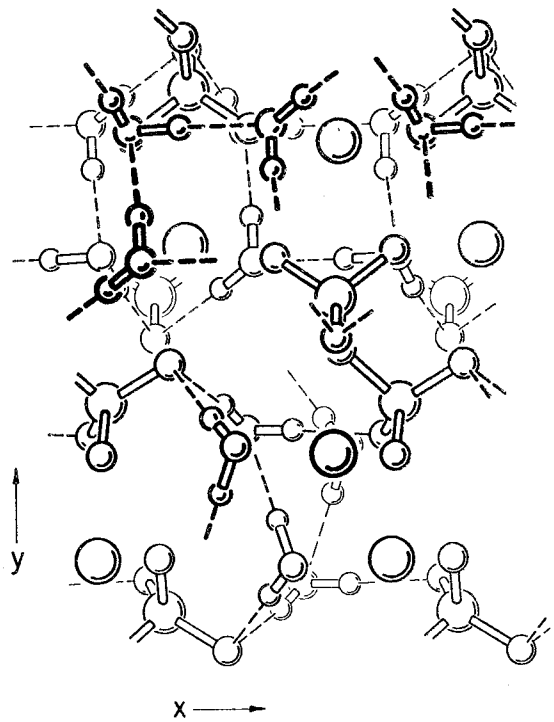


Figure 6. A view down the [001] direction showing hydrogen bonding (dotted lines). Only the part of the cell extending from y = ½ to y = 1¼ is depicted.

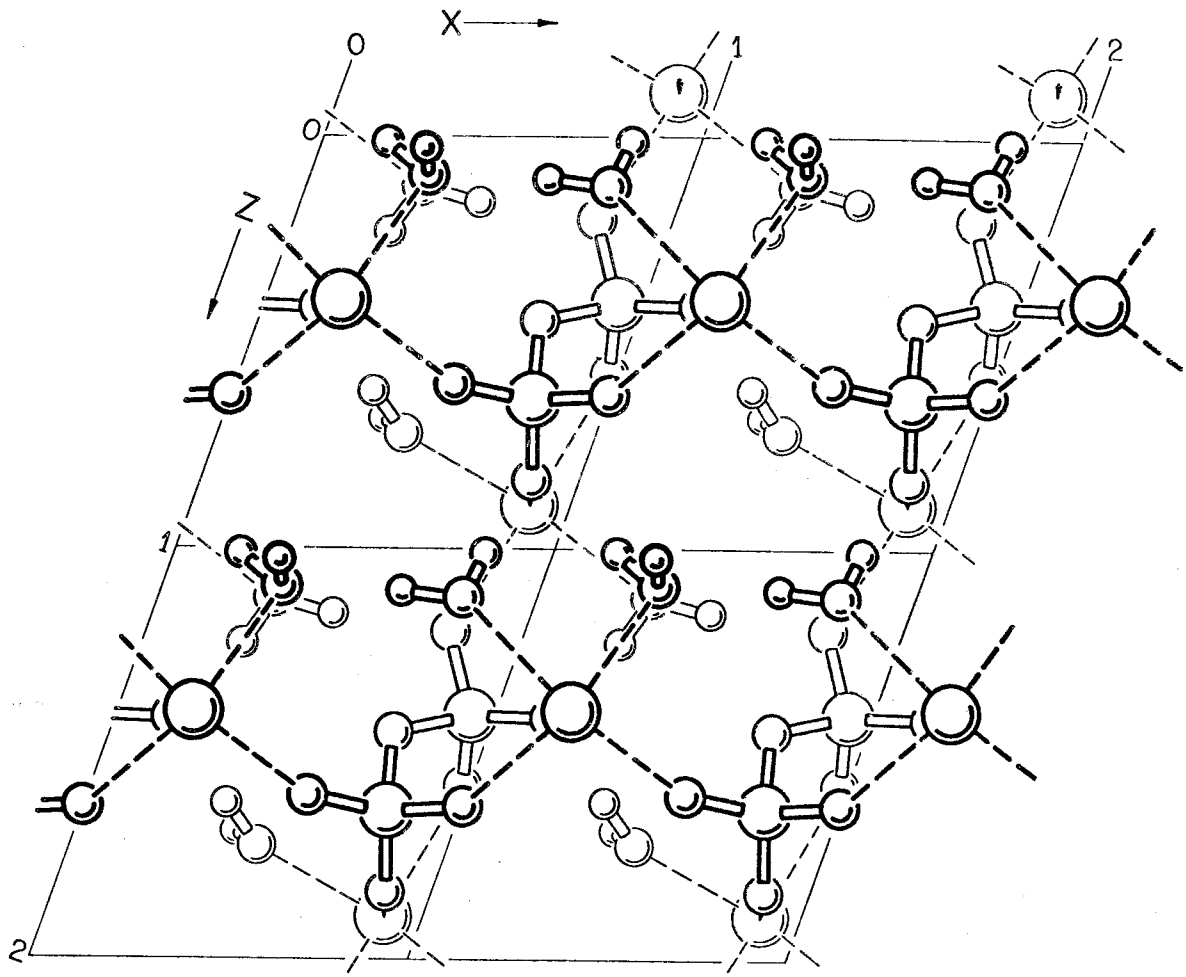


Figure 7. A view down the [010] direction. Calcium coordination is indicated by dotted lines. This view represents the most probable configuration of atoms in the largest crystal face of monoclinic calcium pyrophosphate tetrahydrate crystals.

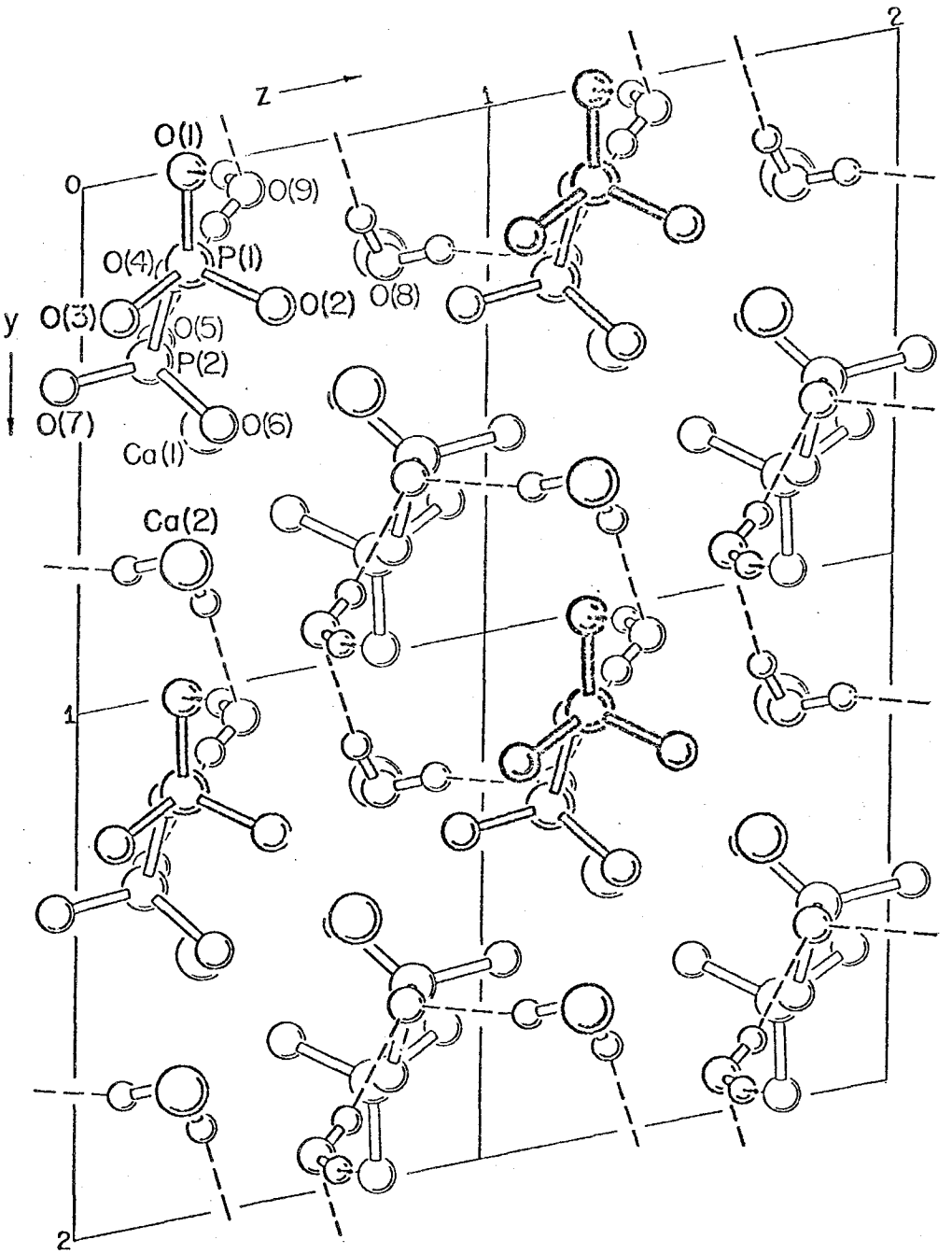


Figure 8. The triclinic dihydrate of calcium pyrophosphate. This is a view down the [100] direction and illustrates the probable configuration of atoms at the surface of the most prominent crystal face. Hydrogen bonding is indicated by dotted lines. Note the "triads" of terminal pyrophosphate oxygens at the surface.

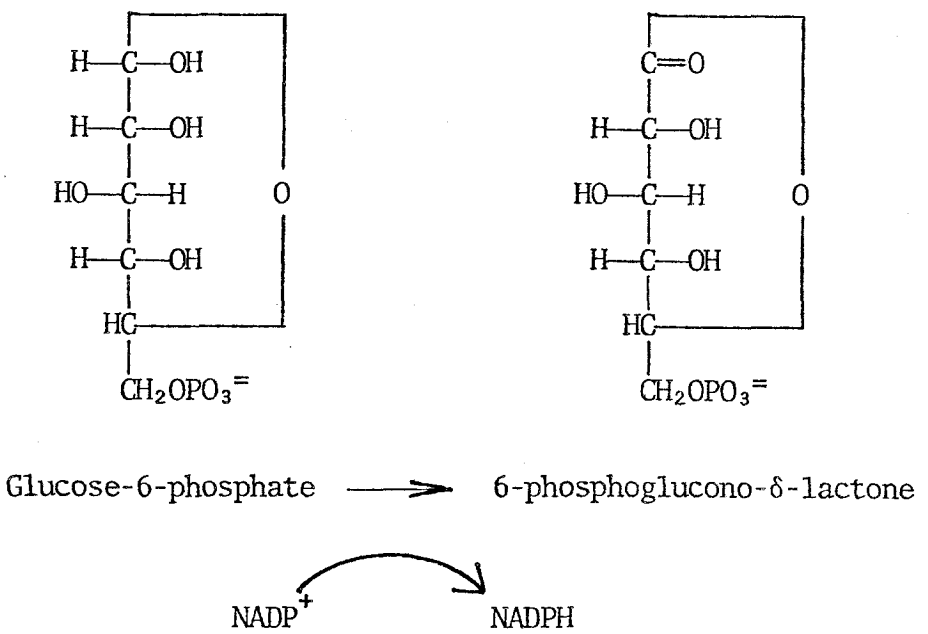
Chapter II

THE PURIFICATION OF GLUCOSE-6-PHOSPHATE DEHYDROGENASE FROM HUMAN RED BLOOD CELLS

INTRODUCTION

Metabolic Role of Glucose-6-Phosphate Dehydrogenase

Glucose-6-phosphate dehydrogenase (G6PD) is the first enzyme in a complex metabolic pathway known as the pentose-phosphate shunt or phosphogluconate pathway. The pentose-phosphate shunt provides many of the materials needed for biosynthesis. The most important of these are ribose-5-P (used in nucleic acid synthesis) and NADPH (the source of reducing power in biosynthetic reactions). G6PD catalyzes the reaction:



That is, it catalyzes the transfer of two hydrogen atoms from G6P to NADP, thereby converting G6P to the corresponding lactone.

Medical Significance of G6PD

The antimalarial drug pamaquine was widely used during World War II. Most patients responded well, with few side effects, but for a few patients the drug created disastrous side effects. Large quantities of red blood cells were destroyed (hemolytic anemia), jaundice developed, and the urine turned black. Many patients died. In 1956, it was discovered that pamaquine-sensitive patients possessed an electrophoretic variant of G6PD (1). This variant (later dubbed "Type A") has a tenfold reduction in enzymatic activity compared to the normal enzyme (2). Since then, over sixty variants of G6PD have been discovered. At least ten of them, including Type A, involve single amino acid substitutions, making G6PD deficiency a true molecular disease. G6PD deficiency has also been linked to chronic hemolytic and food-induced anemia (2,3,4,5).

Ironically, it appears that the variant enzyme actually confers some degree of protection from malaria. Type A deficiency is present in 11% of American Negroes. Like sickle-cell hemoglobin, it has been maintained in the population through a balancing of the selection pressures exerted by the deleterious and beneficial aspects of the trait. The reason for the protective effect of a defective enzyme seems to be this: The mature red blood cell needs the enzymes of the pentose-phosphate shunt primarily for production of NADPH. The reducing power of NADPH is used in the synthesis of reduced glutathione, a tripeptide which is essential for maintaining the integrity of red blood cell structure. Red blood cells that are deficient in G6PD, and hence deficient in reduced glutathione, are much more susceptible to hemolysis.

However, this is not much of a problem under ordinary circumstances, since in a healthy human the loss is compensated by increased production of red blood cells. But for the malarial organism, the story is quite different. These organisms seem to place more demands on the pentose-phosphate shunt than does the host red blood cell. The parasites are actively growing and reproducing, and so require large amounts of NADPH for biosynthesis and large amounts of ribose-5-phosphate for DNA synthesis (a mature red blood cell does not synthesize DNA, and thus has little use for ribose-5-phosphate). A deficiency in G6PD is much more detrimental to the parasite than to the host.

The molecular basis of the defect is unknown. Knowledge of the details of the three-dimensional structure could help to provide an answer to this question. Hence it was decided to attempt the crystallization of this protein for x-ray diffraction studies. In order to do this, we first had to purify the protein in amounts large enough to begin crystallization trials on a large scale.* This chapter is chiefly concerned with our efforts to solve the problems of purifying the enzyme from human red blood cells, where it exists at a concentration of about 1 mg enzyme per unit (pint) of blood (6,7).

---

*We estimated that we would need 400-500 mg of enzyme for x-ray studies.

MATERIALS AND METHODS

Red Blood Cells

Outdated human red blood cells were obtained from the American Red Cross in downtown Los Angeles. The plasma and serum had already been removed from the plastic bags in which the red blood cells were stored.

Reagents

DEAE-Sephadex A-50 and CM-Sephadex C-50, NADP, and Tris were obtained from Sigma; Blue Dextran Sepharose manufactured by Sigma was supplied to us by Dr. Mike Schimerlik of Caltech. DEAE-Sephadex A-25 was obtained from Pharmacia. Disodium EDTA, $Na_2HPO_4 \cdot 7H_2O$, KH_2PO_4, KCN, and $K_3Fe(CN)_6$ were supplied by J. T. Baker Chemical Company. Mercaptoethanol was supplied by Matheson, Coleman, and Bell Manufacturing Chemists.

Hemoglobin Assay

To 2.5 ml of Drabkin's solution (8,9) were added 20 µl of the test solution. Hemoglobin concentration was calculated by the formula

$$[Hb] = \frac{Abs_{541\ nm}}{\varepsilon_{541\ nm}} \times \frac{2.02\ ml}{0.02\ ml}$$

where ε_{541} = molar extinction coefficient at 541 nm of cyanomethemoglobin, at 541 nm, which is 44,000 ℓ/cm-mole.

G6PD Assay

This assay measures the increase in absorbance at 340 nm as NADP is reduced to NADPH by the enzyme. All measurements were done on a Gilford spectrophotometer using a visible light source with a blue filter (6).

Assay solution: 0.1 M Tris buffer, pH 8.0, plus 0.01 M $MgCl_2$

Substrates: 10^{-1} M G6P

10^{-2} M NADP

2.5 ml of the assay solution was placed in a spectrophotometer cuvette. 0.02 ml G6P, 0.02 ml NADP, and 0.02 ml enzyme solution were added. The top of the cuvette was covered with a piece of parafilm and then it was inverted gently several times in order to mix the contents before being placed in the spectrophotometer. The specific activity in millimoles of product/ml-min was calculated from the equation:

$$\frac{2.56 \text{ ml} \times 6}{t \times 6.22 \times 0.02 \text{ ml}}$$

2.56 ml = total volume of assay

0.02 ml = volume of enzyme solution

6.22 = molar extinction coefficient of NADPH at 340 nm

6 = scaling factor (0.2 absorbance units per 100 vertical blocks on chart paper)

t = inverse slope:number of horizontal blocks (representing change in time where the chart speed is 60 blocks/min) per ten vertical blocks (representing the change in absorbance); that is, the inverse of the change in

absorbance with time.

The total G6PD activity was calculated by multiplying the specific activity by the volume of the solution (in ml).

Buffers

Unless otherwise stated, all buffers (except the assay buffers) are assumed to contain 10^{-3} M EDTA, 10^{-3} M mercaptoethanol and 10^{-5} M NADP.

Determination of the Amount of Ion Exchanger (or affinity chromatography gel) Required to Bind a Particular Protein.

This is an adaptation of the procedure described by Pharmacia Fine Chemicals, Uppsala, Sweden (10). Each sample of gel to be tested was suspended in 30 ml of distilled water.. After the gel had swelled to its maximum size, the suspension was poured into a Buchner funnel and the water was drained off. The gel was equilibrated with buffer (0.005 M phosphate buffer, pH 6.4) by adding 30 ml of cold buffer to the gel in a test tube, inverting the tube several times in order to mix the gel with the buffer, allowing the gel to settle and then decanting the supernatant. The gel was washed two more times in this way. 30 ml of raw hemolysate was added to the sample and equilibrated with the gel for one hour. Every ten minutes the test tube was inverted gently several times in order to re-suspend the gel in the solution. Finally the G6PD activity of the supernatant was measured. A comparison of this activity with that of the raw hemolysate indicated

how many activity units of G6PD had been bound by the gel.

Method of Purifying G6PD (Yoshida procedure)

In 1966, Dr. Akira Yoshida of the City of Hope Medical Center developed a method for the isolation of G6PD from human red blood cells (6,7), a procedure which he has recently modified by the addition of a gel filtration step and the elimination of two calcium phosphate gel ion-exchange columns.

Over 99% of the protein in a red blood cell is hemoglobin. Thus, if one wishes to isolate any other red blood cell protein, the first step should involve the removal of as much hemoglobin as possible. This is necessary not only because of the high concentration of hemoglobin relative to other proteins, but because of the large absolute amount of hemoglobin which is present (about 60 grams of hemoglobin per liter). Once most of the hemoglobin has been removed, the remaining protein can be precipitated from solution and redissolved in a much smaller volume of buffer. The first step in Dr. Yoshida's procedure accomplishes the removal of the hemoglobin by means of an anion exchange on DEAE-Sephadex A-50 at pH 6.4. Since the isoelectric pH of G6PD is 6.1 and that of hemoglobin is 6.8, the proteins will have opposite charges at pH 6.4. The positively-charged hemoglobin will be repelled by the positively-charged gel and will be immediately eluted off while the negatively-charged G6PD will be retained. The G6PD can then be eluted with a buffer of high ionic strength at a pH at which the enzyme is positively charged (pH 5.6 to 5.8). Assays of hemoglobin

in the effluents (see Table I) show that at least 90% of the hemoglobin is removed in this first step.

The remaining steps of the Yoshida procedure consist of a series of ion-exchange columns, with cation exchangers (to remove proteins more acidic than G6PD) alternating with anion exchangers (to remove proteins more basic than G6PD, such as hemoglobin). A single gel filtration on Sephadex G-200 removes most of the remaining hemoglobin according to the different molecular weights of the two proteins (G6PD, 110,000; and hemoglobin, 64,000).

In the summer of 1975, the author worked in Dr. Yoshida's lab for three weeks in order to learn his revised purification procedure first-hand. A complete account of this work is given below. Our attempts to scale up the Yoshida procedure from two units to fifty units of blood are described in the section headed "Scaling Up the Yoshida Procedure."

This method was designed so that the washing, hemolysis, and the first two steps of ion exchange chromatography would be performed on two units of blood at a time. The precipitated protein from several such workups could be pooled for the remaining steps of the isolation procedure.

Washing the red blood cells

Packed human red blood cells were obtained from the American Red Cross. The plasma and most of the serum had already been removed from the bag in which the blood had been collected. The blood cells from two such blood bags were distributed equally among four 250 ml Sorvall

centrifuge tubes. Two volumes of 0.9% NaCl were added to each centrifuge tube. The cells were stirred gently with a rubber-tipped glass stirring rod before placing the bottles in the centrifuge. The suspension was centrifuged for ten minutes at 5000 rpm (15,000 x g). The bottles were taken out of the centrifuge, the supernatant was drawn off by means of a vacuum pipette, and more saline was added. After a second wash and centrifugation, the saline was again removed and the cells were ready to be lysed.

Hemolysis of red blood cells

Reagents to be used:

0.005 M phosphate buffer, pH 6.4, containing 10^{-3} M EDTA and 10^{-3} M mercaptoethanol (ME).

Toluene.

The packed red blood cells were poured into a glass-stoppered bottle. One volume (approximately 500 ml) of buffer was added, then 40 ml of toluene. This mixture was shaken vigorously for five minutes until all the cells were hemolyzed, then centrifuged at 11,000 rpm (37,000 x g) for ten minutes. The supernatant, which contained the G6PD, was drawn off by means of a vacuum pipette, and its G6PD activity was measured.

First DEAE-Sephadex Column

Reagents used:

Equilibration buffer: 0.005 M phosphate, pH 6.4, containing 10^{-3} M EDTA.

First elution buffer: 0.005 M phosphate, pH 6.4, containing 10^{-3} M EDTA, 0.05 M KCl.

Second elution buffer: 0.1 M phosphate, pH 5.8, containing 10^{-3} M EDTA, 0.20 M KCl.

Four to five grams of DEAE-Sephadex were equilibrated with the first buffer, and the pH was adjusted to 6.4 with 2N NaOH. The sephadex was placed in a 600-ml Buchner funnel which fit into a one-liter vacuum flask. The hemolysate was poured gently onto the sephadex, and was pulled through the gel rather rapidly by the use of suction. Most of the hemoglobin came out in the effluent.

In order to remove more hemoglobin, 200 ml of the second elution buffer were added to the sephadex. The activity of the first and second effluents from the column was tested in order to insure against the loss of too much G6PD. We expected to lose about 15-20% of our activity at this point.

The G6PD was eluted with 300-400 ml of the third buffer. NADP and EDTA stabilize G6PD, and so these were added to final concentrations of 10^{-5} M and 10^{-3} M, respectively. Two volumes of cold, distilled water were added to the effluent in order to reduce the salt concentration of the buffer. The pH of the solution was adjusted to 5.8 with 1 M acetic acid, and mercaptoethanol was added to a final concentration of 10^{-3} M.

Mercaptoethanol helps maintain the integrity of disulfide bonds; it must be added after the pH adjustment because it interferes with the accurate measurement of pH.

First CM-Sephadex column

Reagents used:

Equilibration buffer: 0.005 M phosphate buffer, pH 5.8, containing 10^{-3} M EDTA and 0.1 M KCl.

Elution buffer: 0.1 M phosphate buffer, pH 7.5, containing 10^{-3} M EDTA, 0.20 M KCl.

Two to three grams of CM-Sephadex were equilibrated with the first buffer, and the pH was adjusted to 5.8. The gel was placed in a Buchner funnel as before, and the protein solution from the previous step was added. The G6PD was eluted with 200-300 ml of the second buffer. NADP and EDTA were added, and the pH was adjusted to 6.1 (the isoelectric point of G6PD). Finally, 350 g/l of ammonium sulfate was added in order to precipitate the protein from solution. The solution turned cloudy almost immediately, but was allowed to stand for one-half hour to insure that all the G6PD had precipitated. The suspension was centrifuged for ten minutes at 37,000 x g, and the supernatant was drawn off with a vacuum pipette. The activity of the supernatant was tested and compared to that of the unprecipitated solution. If too much activity remained in the supernatant, more ammonium sulfate was added and the solution was centrifuged as before.

Remaining steps of column chromatography

It will be noted that the first two steps of the procedure involved placing the ion-exchange gel in a Buchner funnel rather than in a true column. This adaptation was made because of the large volumes of solution involved in purifying G6PD from two units of blood (about one liter). The precipitated protein from the second ion-exchange separation (on CM-Sephadex), however, could be redissolved in a small enough volume that it became possible to use an ordinary column and collect fractions in the usual way. In the small-scale prep of the summer of 1975, the precipitated protein from three two-unit preps was pooled at this stage. Elution profiles for these columns are shown in Figure 1, and the yield of enzyme at each step is given in Table II.

After each step of column chromatography, the fractions were assayed for total protein and G6PD activity; the active fractions were pooled. NADP, EDTA, and mercaptoethanol were added to final concentrations of 10^{-5} M, 10^{-3} M, and 10^{-3} M, respectively. The pH was adjusted to 6.1 and ammonium sulfate was added to a concentration of 350 g/l protein solution. The solution was then centrifuged and the pellet suspended in a minimal amount of the equilibration buffer for the next column. After dialysis overnight against 50-100 volumes of the same buffer, the solution was centrifuged, and the supernatant was added to the column.

In all ion-exchange steps, the enzyme was eluted by means of a linear salt gradient. The proximal mixing chamber contained about 150 ml of the low salt buffer, and the distal chamber contained an

equal amount of high-salt buffer. The gradient and type of elution buffer are listed for each column.

Second DEAE-Sephadex column

Buffer: 0.05 M phosphate, pH 7.0, EDTA, NADP, mercaptoethanol.

Gradient: KCl, 0.0 to 0.3 M.

Second CM-Sephadex column

Buffer: 0.005 M phosphate, pH 5.8, 0.1 M KCl, EDTA, NADP, mercaptoethanol.

Gradient: KCl, 0.1 to 0.5 M.

Gel filtration (Sephadex G-200)

Buffer: 0.01 M phosphate, pH 6.8, EDTA, NADP, mercaptoethanol.

Third DEAE-Sephadex column

Buffer: 0.02 M Tris, pH 8.0, NADP, EDTA, mercaptoethanol.

Gradient: NaCl, 0.0 to 0.3 M

Test for Purity

After the third DEAE-Sephadex column, a sample of protein solution was electrophoresed at pH 8.0 in polyacrylamide gel. If the G6PD is the only protein remaining in solution, only one band of protein should appear. In the present case, two distinct protein bands appeared, and

so the third DEAE-Sephadex column had to be repeated (the elution profile for this column is not shown). The purified G6PD could be stored for several months without loss of activity.

SCALING UP THE YOSHIDA PROCEDURE

As noted above, the Yoshida procedure is designed for processing two units of blood at a time during the first four steps of the procedure (washing, hemolysis, and the first two ion-exchange processes). Because of the large amount of protein (300-500 mg) needed for x-ray studies, it was hoped that these first steps could be scaled up from two to fifty units.

Washing of Red Blood Cells

Dr. Yoshida's method involved washing the cells twice in two volumes of isotonic saline (0.9% NaCl). The purpose of this step is to remove residual serum from the cells. Since the serum contains proteases which could destroy the red blood cell proteins, it was though necessary to take all possible precautions to avoid protease contamination of the hemolysate. When processing 50 units of blood at a time, washing the cells proved to be very impractical. This was so not only because of the large volume of solution but because of the lack of an effective method of centrifuging the cells gently. Attempts to use the Sharples centrifuge for this purpose were unsuccessful. Many cells were lysed in the high-speed centrifugation, and the intact red blood cells did not adhere to the blades of the Sharples, so they could not be recovered for further washing or for hemolysis.

It was decided to attempt hemolysis without washing the red blood cells first. Since the ratio of hemoglobin to G6PD in the raw hemolysate was so high (about 50,000:1), it was expected that any contaminating proteases would be far more likely to attack hemoglobin than G6PD. This method worked so well--precipitated enzyme remained active for months--that all large workups were performed with unwashed red blood cells. As an extra precaution, ε-N-caproic acid, a protease inhibitor, was added to the G6PD-containing effluent before precipitation and storage.

Hemolysis of Red Blood Cells

The two aspects of the hemolysis step which gave the most trouble were those of lysing the cells with 100% efficiency, and removing the stroma. The cells were poured into five-gallon bottles and shaken up with the appropriate volumes of distilled water and toluene (in direct imitation of the Yoshida procedure). When half full, the bottles could be shaken manually. However, our results indicated that hemolysis of the cells by this method was only about 20% efficient--the bottles simply could not be shaken vigorously enough to lyse all of the cells.

Use of the Sharples continuous-flow centrifuge had two major disadvantages. For one thing, the stroma did not adhere to the blades of the centrifuge and came out with the hemolysate. This was not a particularly serious problem since most of the membranes and toluene remained on top of the solution, which could then be siphoned out from

the bottom of the bottle. The second problem was more serious, and arose from the fact that the Sharples is a continuous-flow centrifuge. In order to wash out the hemolysate remaining in the bottom of the chamber, a large amount of buffer had to be added. The result was a concentration gradient in the hemolysate, as evidenced by a continous loss of red color. The problem was remedied somewhat by applying the hemolysate to the centrifuge under pressure (which resulted in a lower volume of buffer being required to flush out the Sharples).

As is evident from an examination of Table II, the raw hemolysate contained only 20-25% as much G6PD as it should have. Fifty units of red blood cells should contain from 20,000-25,000 activity units of G6PD. It was evident that the red blood cells were not being lysed and hence were not releasing the G6PD into solution. Indeed, many red blood cells as well as clumps of fragmented membranes could be seen clogging the ground glass of the Buchner funnels.

The loss could not be attributed to protease digestion of the G6PD. If the G6PD were being released into solution and then destroyed by proteases, then 80% of the hemoglobin should have been destroyed as well. But there was no evidence for massive digestion of hemoglobin, such as the accumulation of ferric oxide or the presence of large amounts of free heme in the toluene layer. Furthermore, small-scale tests indicated that the red blood cells we had received were not deficient in G6PD activity. Eventually, a more effective method of lysing red blood cells--low-energy sonication--was found. This method is described in the section entitled "Improvements in the Yoshida Procedure."

Anion Exchange

First DEAE-Sephadex column

Three forty-gram samples of Sephadex were equilibrated overnight at 4°C in 0.005 M phosphate buffer (pH 6.4, 10^{-3} M EDTA, 10^{-3} M mercaptoethanol) in four-liter beakers. The contents of each beaker were poured into a three-liter Buchner funnel. This was done about one-half hour before the hemolysate was to be added so that the gel would stay cold.

A water aspirator was used to draw the hemolysate into a five-gallon bottle on the shelf above the funnels. The protein solution could then flow down onto the funnels containing the ion exchanger through a manifold, which divided the solution into three equal parts.

The solution was drawn through the gel rather rapidly (one-half hour). Some hemoglobin which remained on the column was washed off with five liters of low-salt buffer (0.05 M KCl, 0.005 M phosphate, pH 6.4, EDTA, mercaptoethanol). Finally, the G6PD was eluted with high-salt, low-pH buffer solution (0.1 M phosphate, pH 5.8, 0.2 M KCl, EDTA, mercaptoethanol). Two to three bottles were used to collect the effluent. These bottles had been previously calibrated and marked at two-liter intervals. Each one was filled to the six-liter mark, and two volumes of cold, distilled water were added. Since EDTA and NADP are negatively charged, these ions had remained on the Sephadex. It was therefore necessary to add enough EDTA and NADP to make a final concentration of 10^{-3} M and 10^{-5} M, respectively. A one-liter aliquot from each bottle was titrated with 2N NaOH and/or 2N HCl. In this way

one could calculate the amount of acid or base that had to be added to take the solution to a final pH of 5.8. The bottles were placed in the coldroom while the CM-Sephadex was being prepared.

First CM-Sephadex column

Two of the Buchner funnels from the previous step were cleaned out and reused. Thirty grams of previously equilibrated CM-Sephadex were put into each one. The equilibration buffer contained 0.1 M phosphate, pH 5.8, EDTA, mercaptoethanol, and 0.1 M KCl. The protein solution was applied to the Sephadex and allowed to flow through slowly (two to three hours). Finally, about eight liters of elution buffer (0.1 M phosphate, pH 7.5, 0.2 M KCl) were added. The gel was then eluted to dryness.

The problems encountered in scaling up this part of the procedure were many. DEAE-Sephadex A-50 undergoes a reduction in bed volume of some 50% in going from an ionic strength of 0.005 M to 0.1 M. Thus, when the washes of higher ionic strength were added, the anion exchanger shrunk considerably. This "settling" of the gel reduced the flow rate to a mere trickle. This problem was not serious when only two units of blood were being processed, but with the large bed volumes of ion exchanger needed to process 50 units at a time it became very serious indeed: the column was taller and hence the solution had to flow through the tightly packed gel for a longer distance. To make matters worse, the previous problems of incomplete hemolysis and

inefficient separation of the stroma from the hemolysate came back to haunt us at this stage. Unlysed red blood cells and pieces of membrane reduced the flow rate still further by clogging both the ion-exchange resin and the ground glass of the Buchner funnel. The reduced flow rate and increased column height resulted in a lowered yield of G6PD, as can be seen in Table III. The yield of enzyme from the first ion-exchange step was only 58%, whereas in the small-scale process it was over 80%. The greater column height caused a larger amount of buffer to be needed for elution. As a result, the overall concentration of G6PD in each fraction of the effluent was lower. It was thus harder to detect the tail of the G6PD elution curve, and to know where to make the cutoff. The large amount of time required for elution, as well as the desire to minimize the volume of solution for the second ion-exchange step (see below), were also contributing factors in this decision.

In the first two large-scale G6PD preps, we followed Dr. Yoshida's advice about running the first CM-Sephadex C-50 ion exchange immediately after the DEAE-Sephadex ion exchange reaction. However, as can be seen in Table II, this resulted in an enormous loss of G6PD. The necessity of adding two volumes of distilled water to the solution in order to reduce the ionic strength of the solution resulted in an extremely large volume of solution ($\sim$40 liters). CM-Sephadex C-50, like DEAE-Sephadex A-50, decreases in bed volume with increasing ionic strength. Therefore, the high ionic strength of the equilibration buffer reduced the volume of the gel to such an extent that the flow rate was no better

than that of the DEAE-Sephadex ion exchange reaction. Consequently, the entire procedure required from fourteen to seventeen hours, not counting the hemolysis step. Therefore, for the third and final large-scale prep, the G6PD was precipitated from solution with ammonium sulfate immediately after the first ion-exchange step.

IMPROVEMENTS IN THE YOSHIDA PROCEDURE

The hemolysis and ion exchange on DEAE-Sephadex were clearly the bottleneck of the G6PD purification procedure because of the large volumes of solution involved during those stages of the procedure. Efforts toward improving the Yoshida method of G6PD purification were hence concentrated primarily on improving the efficiency of these two steps, and solutions to these problems are given below.

Hemolysis of Red Blood Cells by Sonication

Sonication is a common method of disrupting cell membranes, and so tests were made to determine how it could be used in the lysis of red blood cells. The best results were obtained using low-energy (less than 30,000 cps) sonication in the presence of hypotonic buffer and toluene. Higher frequencies produced membrane fragments too small to be centrifuged at 30,000 x g in an ordinary centrifuge.

A bath sonicator (Do-All Ultrasonic Cleaner Model 235) was found to produce sound waves of 28,000 cps and had a capacity of 2.5 liters--enough to handle the lysis of five units of red blood cells. Three to five minutes of sonication were required to lyse the cells. The stroma were removed by centrifugation in a large swinging-bucket centrifuge (International Equipment Corporation, Model PR-2), which also had a capacity of 2.5 liters. Assays indicated that over 2,000 units of G6PD activity were present--an almost 100% yield. The stroma also were pale pink in color with virtually no intact red blood cells in the pellet.

Use of DEAE-Sephadex A-25 as an Ion Exchanger

As mentioned above, DEAE-Sephadex A-50 undergoes a nearly twofold decrease in volume in going from an ionic strength of 0.005 M to one of 0.20 M as the ionic strength of the equilibration buffer changes from 0.005 M to 0.20 M. This volume change, or "settling", of the gel results in a considerable decrease in the flow rate of solution being drawn through it. DEAE-Sephadex A-25 does not undergo such significant change in volume over the range of buffer ionic strengths which are used in the purification of G6PD. Therefore, it was decided to compare the G6PD-binding capacities of the A-25 and A-50 gels in order to see if the A-25 could be used instead of the troublesome A-50.

Three large test tubes were prepared, one containing 0.050 g of DEAE-Sephadex A-50, one containing 0.25 g of DEAE-Sephadex A-25, and the third containing 0.10 g of A-25. Each sample was equilibrated with water and buffer as described under "Materials and Methods." Each sample was equilibrated with 30 ml of raw hemolysate. The results given in Table IV indicate that Sephadex A-25 has a much lower capacity (per gram) for binding G6PD than does the A-50 gel. Thus, much more A-25 would be required for a given amount of hemolysate. The advantages to be gained through elimination of the settling problem, however, recommend the use of the A-25 gel in at least the first step of a large-scale prep.

Affinity Chromatography

De Flora et al. (11) were able to purify G6PD from human red blood cells by using an affinity column composed of an NADP analog covalently coupled to an agarose gel. The ligand, however, was very expensive and the synthesis of the gel was laborious (12). Hence, the cost and labor involved in synthesizing enough affinity gel for our purposes would be great. A more significant factor, however, in the decision not to adapt their procedure directly was the fact that two steps of ion exchange chromatography had to be carried out before the protein solution was to be added to the affinity column. The attempts of De Flora et al. to eliminate the ion-exchange steps and proceed directly to the affinity column met with no success. Also, the first step in De Flora's ion-exchange procedure was the same as the Yoshida method--ion exchange on DEAE-Sephadex A-50--which was the very bottleneck in our own procedure which we were determined to avoid.

Ryan and Vestling (13) had been able to purify lactic dehydrogenase (LDH) from rat liver with an affinity column using blue dextran as the ligand. Moreover, they had been able to carry out the affinity chromatography step before the ion exchange on DEAE-Sephadex. Later, Thompson et al. (14) showed that blue dextran Sepharose 4B could be used for purifying a large number of different NAD- and NADP-linked dehydrogenases. Apparently the dye Cibacron Blue F3GA which is attached to the dextran is an analog of NAD. Thompson described it as specific for the "dinucleotide fold" which is a characteristic structure of these enzymes. Not only is blue dextran much less

expensive than the ligand used by De Flora et al., but synthesis of the affinity gel from blue dextran and sepharose is much simpler.

In theory, at least, affinity chromatography should be much more effective in separating the G6PD from hemoglobin because the separation is based on substrate specificity rather than a slight difference in charge. An affinity column has the additional advantage that the progress of the purification could be viewed with the naked eye--one could simply keep washing the column until no hemoglobin came off in the effluent. Removal of 99.9% of the contaminating protein rather than 90% (the yield obtained when using an ion exchange on DEAE-Sephadex) would yield 60 µg of highly purified protein rather than 6000 µg of crude extract. But would the gel bind the G6PD well enough that affinity chromatography on blue dextran sepharose could be used as the first step in the purification procedure?

Literature published by Pharmacia (10), as mentioned above, describes a method for determining the amount of ion-exchange resin needed to bind a particular protein. A 160-mg sample of blue dextran sepharose was kindly lent to us by Dr. Mike Schimerlik of Dr. M. A. Raftery's research group. We proceeded to test the sample by a modification of the Pharmacia procedure. The sample was divided into two equal parts, and each portion was equilibrated several times with 30 ml of 0.005 M phosphate buffer, pH 6.4. A 30-ml portion of hemolysate was then added to each sample. One of the proportions of hemolysate contained NADP at a concentration of 10^{-5} M, the concentration needed to maintain enzyme stability.

The results (given in Table IV) indicate that the blue dextran does bind G6PD and that low concentrations of NADP, while interfering with the binding to a certain degree, do not inhibit it completely. Washing the gel with successive 30-ml portions of 1 M NaCl failed to remove the enzyme but were quite effective in removing hemoglobin, since the third such wash was free of this protein.

The gel was not very efficient in removing the G6PD from solution--the amount of G6PD bound per mg of blue dextran sepharose was very low. This was no doubt due to the very low concentration of G6PD in the raw hemolysate. The figures given in Table IV indicate that about thirteen grams of gel would be required to bind the G6PD from two units of blood --325 grams for a fifty-unit prep. While synthesizing large amounts of blue dextran sepharose would not present much of a problem, elution of the enzyme from the gel would be a rather expensive procedure. For instance, Ryan and Vestling (13) needed 50 ml of 10^{-3} M NAD (50 mg) to elute 1200 activity units of rat liver LDH from 80 ml (4.3 g) of the gel. Assuming that similar amounts of NADP would be required to elute G6PD from blue dextran sepharose, then the elution of 25,000 activity units of G6PD would require over one gram of NADP. The large amounts of NADP required for elution thus make affinity chromatography impractical during the initial purification step. In the latter steps of the purification, however, affinity chromatography would be the method of preference.

CONCLUSION

Although it was not possible to isolate human erythrocyte G6PD in quantities large enough for single-crystal x-ray diffraction studies, a great deal of knowledge was gained.

Low-energy sonication proved to be an effective method of lysing the red blood cells, and washing the packed red blood cells proved to be unnecessary. As many as five units of cells can be lysed and run through a batch-process ion exchange separation, precipitated, and stored for several weeks if necessary. Several such precipitated fractions can then be redissolved and subjected to further column chromatography.

DEAE-Sephadex A-25 can be used for the first ion exchange separation in place of the A-50 gel, thereby eliminating the problems caused by volume changes in the ion exchange gel.

It was demonstrated that blue dextran Sepharose 4B could bind G6PD. The blue dextran gel could be used in the latter stages of the purification procedure--after the first two ion exchange steps. It was also demonstrated that the small amount of NADP required to stabilize the G6PD in solution did not interfere with this binding.

A FINAL NOTE

Bürgisser and Faucheré (15) have recently developed a one-step method for the purification of bovine adrenal G6PD by affinity chromatography on NADP-Sepharose 4B. They were able to isolate 4500 activity units of the enzyme from 1300 ml of bovine adrenal extract. Further attempts at purifying G6PD in quantities large enough for x-ray diffraction studies should perhaps be directed toward solving the structure of the bovine adrenal enzyme, since it is much easier to purify than human red blood cell G6PD. Human G6PD can be isolated in quantities large enough for the determination of the amino acid sequence. Its probable structure then could be deduced from an examination of the amino acid sequence and crystal structure of the bovine enzyme.

REFERENCES

1. Carson, P.E., C.L. Flanagan, C.E. Ickes, and A.S. Alving (1956). Science 124, 484-485.
2. Boyer, S.H., I.H. Porter, and R.G. Weilbacher (1962). Proc. Nat. Acad. Sci. USA 48, 1868-1876.
3. Kirkman, H.N., P.R. McCurdy, and J.L. Maiman (1964). Cold Spring Harbor Symposium on Quantitative Biology 29, 391-398.
4. Kirkman, H.N., I.M. Rosenthal, E.N. Simon, P.E. Carson, and A.G. Brinson (1964). J. Lab. Clin. Med. 63, 715-725.
5. Kirkman, H.N., F. Schettini, and B.M. Pickard (1964). J. Lab. Clin. Med. 63, 726-735.
6. Yoshida, A. (1966). J. Biol. Chem. 241, 4966-4976.
7. Yoshida, A. and V.D. Hoagland (1970). Biochem. Biophys. Res. Commun. 40, 1167-1172.
8. Drabkin, D.L. and J.H. Austin (1935). J. Biol. Chem. 112, 51.
9. Dacie, J.V. and S.M. Lewis, Practical Hematology, 4th Ed., p. 37. Grune and Stratton, Inc., New York, 1968.
10. Sephadex Ion Exchangers, published by Pharmacia Fine Chemicals, Uppsala, Sweden.
11. De Flora, A., A. Morelli, U. Benatti, and F. Giuliano (1975). Arch. Biochem. Biophys. 169, 362-363.
12. Morelli, A. and U. Benatti (1974). Ital. J. Bioch. 23, 279-291.
13. Ryan, L.D. and C.S. Vestling (1974). Arch. Biochem. Biophys. 160, 279-284.

14. Thompson, S.T., K.H. Cass, and E. Stellwagen (1975). Proc. Nat. Acad. Sci. USA 72, 669-672.

15. Bürgisser, E. and J-L. Fauchere (1976). Helvetia Chim. Acta 59, 760-765.

TABLE I. Separation of hemoglobin (Hb) from G6PD by ion exchange on 2.5 g of DEAE-Sephadex A-50. The assays were carried out on the hemolysate from one unit of blood (500 ml of solution).

Solution	Volume	[Hb], mg/100 ml	Total Hb, mg	
Three fractions of raw hemolysate	187 ml	50.3	94	245
	197 ml	46.0	90	
	150 ml	40.7	61	
First DEAE effluent	500 ml	42.0	210	
Second DEAE effluent (washed with 0.05 M salt)	100 ml	13.0	13	
Third DEAE effluent, two fractions	425 ml	3.5	15	
	190 ml	2.1	4.5	

Efficiency of Hb removal: $\frac{210 + 13}{245} \times 100 = 91\%$

TABLE II. Yield of G6PD activity at each step in the purification procedure.

Step in the Procedure	Activity Units of G6PD Recovered	Volume of Solution, ml	% Recovery
First DEAE- and CM-Sephadex ion exchanges	1150	160	--
Second DEAE-Sephadex ion exchange	950	145	82.6
Second DM-Sephadex ion exchange	896	62	77.9
Gel filtration on Sephadex G-200	760	20	66.0
Third DEAE-Sephadex* ion exchange	375	75	32.6

*No explanation for the great loss of G6PD at this stage could be found. A fourth ion exchange on DEAE-Sephadex yielded 275 activity units (a yield of 73% from the previous step).

TABLE III. Data for Three Large-Scale (50-Unit) G6PD Preps

	First Workup			Second Workup			Third Workup		
	Specific Activity	Liters of solution	Specific Activity	Specific Activity	Liters of solution	Specific Activity	Specific Activity	Liters of solution	Specific Activity
Hemolysate	0.213	20	4200	0.164	30	4920	0.474 0.061	15 10	7110 610
First DEAE Effluent	0.0	20	0	0.0	30	0	0.0	25	0
Second DEAE Effluent	0.0	5	0	0.0	5	0	0.0	5	0
Third DEAE Effluent	0.142	18	2446	0.079	18	2844	0.341	12	4092
First CM Effluent	0.076	50	3800	0.0	50	0	--	--	--
Second CM Efflulent	0.133	11	1463	8.0	8	1704	--	--	--

TABLE IV. G6PD binding capacities of ion exchange and affinity chromatography gels. The specific activity and total activity of a sample of raw hemolysate have been included for comparison.

Sample	Supernatant Activity: Relative[1]	Total[2]	G6PD Bound[3]	Gel Required[4]
Raw hemolysate	0.776	23.28		
0.05 g A-50	0.304	9.0	0.28	3.571
0.25 g A-25	0.608	18.2	0.0203	49.3
0.10 g A-25	0.656	19.7	0.0358	27.9
BD-Sepharose[5]	0.501	15.03	0.103	9.7
BD-Sepharose + 10^{-5} M NADP	0.569	17.07	0.077	13.0

[1]Specific activity (in activity units of G6PD/ml) remaining in the supernatant after its equilibration with the gel.

[2]Total number of activity units of G6PD remaining in the supernatant after its equilibration with the gel. This number is obtained by multiplying the specific activity by the volume of supernatant (30 ml).

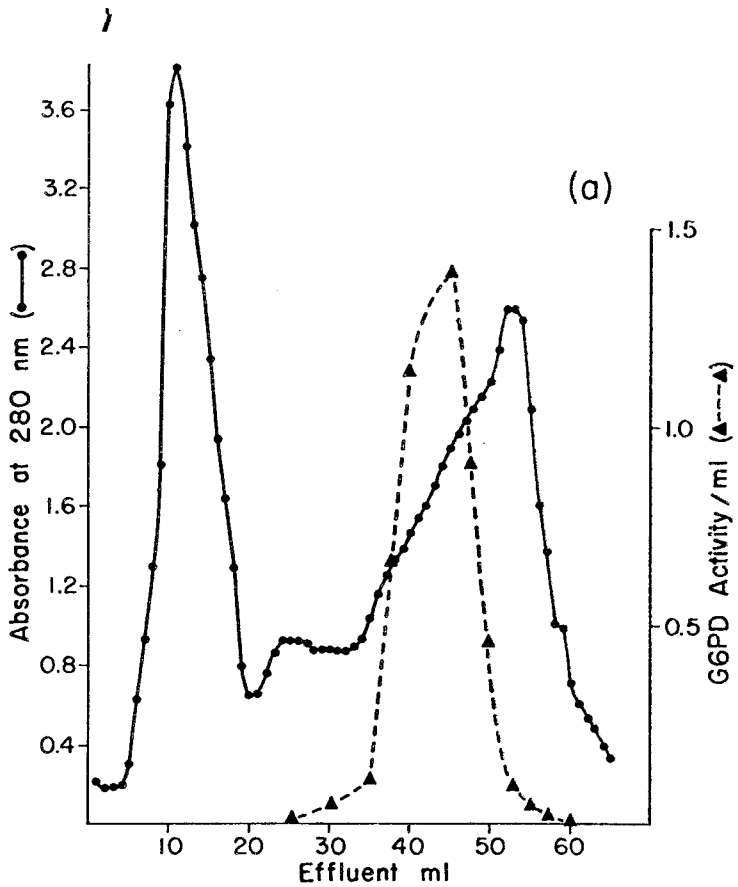
[3]Activity units of G6PD bound per mg gel, i.e., gel capacity.

[4]Amount of gel (in grams) needed to bind the G6PD from 2 units of blood.

[5]80 mg of blue dextran Sepharose 5B.

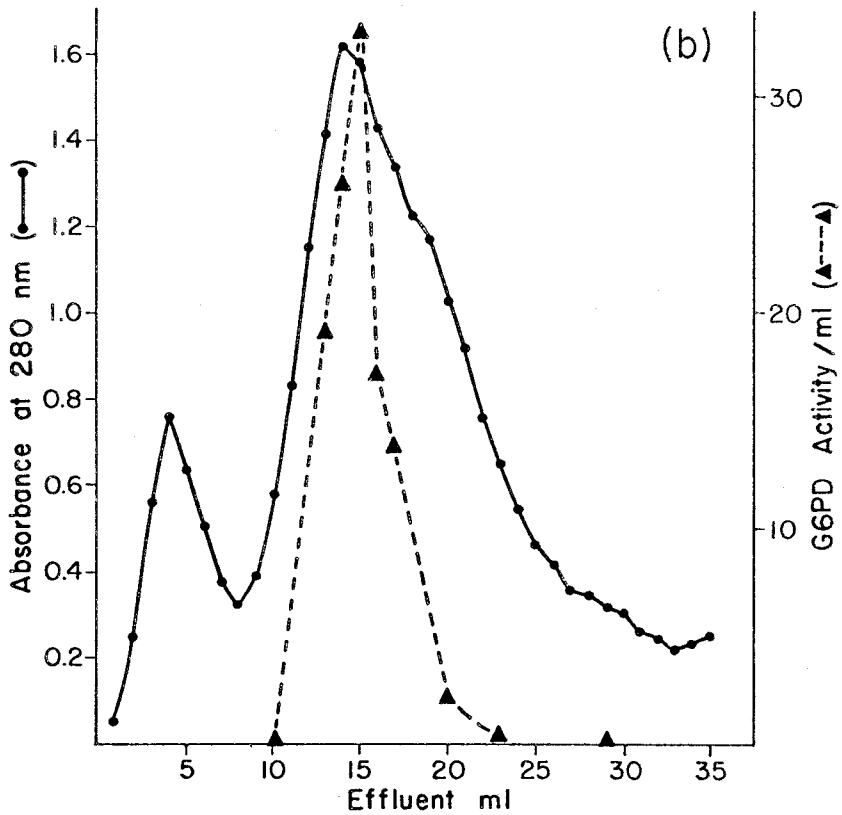
Figure 1

Elution profiles. In all cases ●—● = absorbance at 280 nm (hence total protein), and — = G6PD activity.



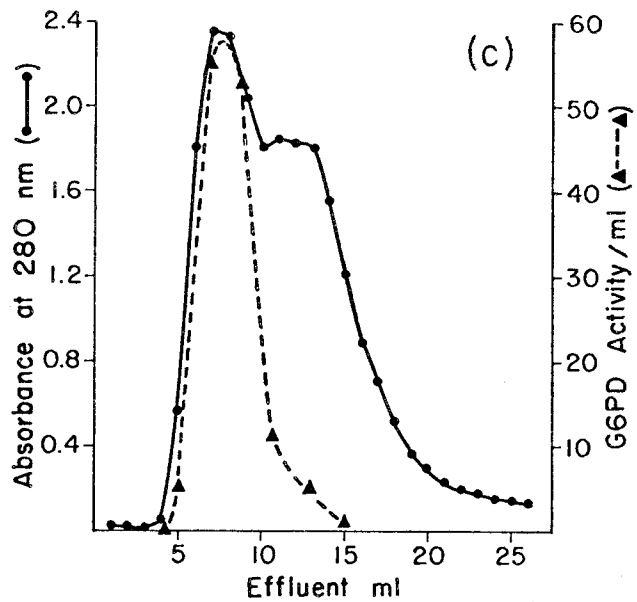
(a) Elution profile of a DEAE-Sephadex column (2.5 x 12 cm). Partially purified G6PD was placed on the column and eluted with increasing KCl concentration (0.0 - 0.3 M).

Figure 1



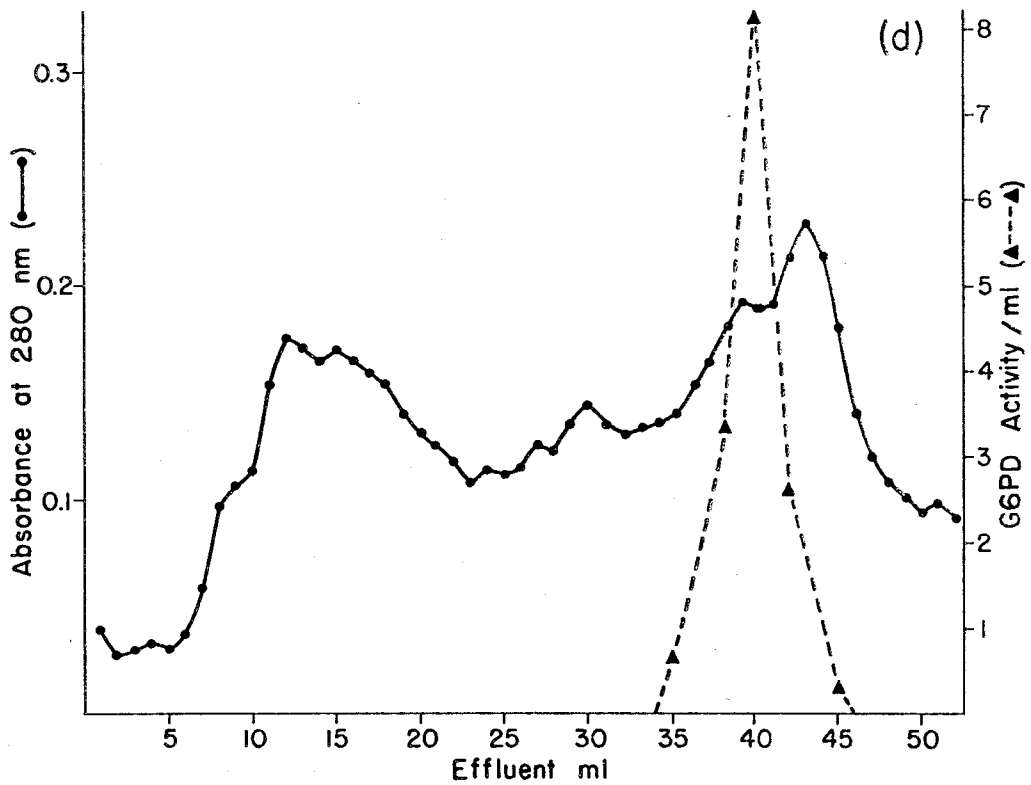
(b) Elution profile of a CM-Sephadex column (1.2 x 20 cm). Partially purified G6PD was placed on the column and eluted with increasing KCl concentration (0.1 M - 0.5 M).

Figure 1



(c) Elution profile of a Sephadex G-200 gel filtration column (2.5 x 100 cm). Partially purified G6PD was placed on the column and eluted with 60 ml of buffer.

Figure 1



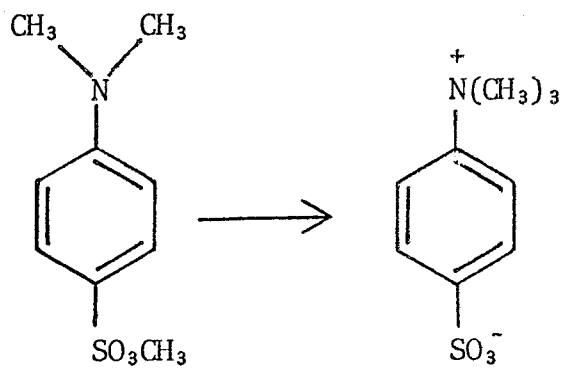
(d) Elution profile of a DEAE-Sephadex column (1.2 x 20 cm). Partially purified G6PD was placed on the column and eluted with increasing KCl concentration (0.0 - 0.3 M).

Chapter III

A Partial Solution of the Crystal Structure of p-Trimethylammoniumbenzene Sulfonate

INTRODUCTION

Methyl-p-dimethylaminobenzene sulfonate (MSE) remains stable in solution at room temperature. In crystalline form, however, this compound (I) rearranges to form the zwitterionic species p-trimethylammoniumbenzene sulfonate (ZWT) as shown below (1):



MSE (I) ZWT (II)

The half-life of this reaction is about seven days at room temperature. The rate increases with temperature up to the melting point, at which point the rate of reaction slows considerably (Figure 1). Since the reaction proceeds much more rapidly in the highly-ordered, crystalline state than in the solution or melt (where the molecules are more or less randomly oriented with respect to one another), it is evident that the particular orientation of the molecules in the crystal is crucial to their effective reaction. A glance at the crystal structure of MSE (Figure 2) shows why this is so. All of the molecules are perfectly

aligned so that the transfer of a methyl group from the oxygen atom of one molecule to the nitrogen atom of its neighbor requires a move of only 2 Å (2).

As the reaction proceeds, a single crystal of MSE is transformed into a paracrystalline array of the zwitterionic species. The shifting of the methyl group destroys the original lattice and, since there are many reaction sites within a single crystal, the end product is a paracrystalline array of ZWT molecules. Powder diffraction photographs indicate that the original compound rearranges to form at least one metastable crystalline form of the zwitterion. This intermediate form (or forms) of ZWT then gradually transforms into the stable allomorph which is the subject of the present study.

EXPERIMENTAL

Crystals of methyl-p-dimethylaminobenzene sulfonate were supplied by Dr. Chaim Sukenik (2,3). The rearrangement reaction was allowed to proceed for several months, i.e., until virtually all the material was converted to the zwitterionic form. Since crystals of the zwitterion obtained in this way were never single, the product was recrystallized from water. Many of these crystals were twins, but a single crystal was eventually found which was used for the crystallographic data collection. The recrystallized material was shown to be identical with the end product of the solid-state reaction by means of powder diffraction photographs taken on the Guinier-Hägg camera described in Chapter I.

The crystal used for data collection had approximate dimensions of 3 mm x 2 mm x 2 mm and was mounted along the long crystal axis (later identified as the crystallographic c axis). Cone axis photographs displayed mmm diffraction symmetry, which indicated an orthorhombic space group. Zero- and first-level Weissenberg and precession photographs showed a great deal of pseudosymmetry in the form of systematically weak reflections (see below) which made space group determination very difficult. Approximate cell constants determined from photographs were refined by a least-squares fit of thirteen hand-centered reflections used in the calculation of the orientation matrix of the diffractometer.

Intensity data were collected on a General Electric automated diffractometer using nickel-filtered CuK_α radiation. The reflections

were scanned in the θ-2θ mode at a rate of 1° per minute with background counts of twenty seconds. The scan range varied linearly from 2° at 2θ = 5° to 4° at 2θ = 150°. A total of 2315 unique reflections were collected within a 2θ range of 5° to 150°. As with the calcium pyrophosphate data collection, three check reflections staggered ten reflections apart were measured every thirty reflections. The intensities of these three reflections deviated less than 1% throughout the data collection.

Unscaled F's and σ(F)'s were derived from the scan counts according to formulas 1 and 2 of Chapter I. Properly scaled F's were obtained by multiplying them by a scale factor derived from a Wilson plot. As before, reflections with intensities less than 4 σ(I) were considered unobserved.

The diffractometer data confirmed the photographic evidence of systematically weak reflections. The only true systematic absences were h00, h = 2n + 1 and 00ℓ, ℓ = 2n + 1, showing that the true space group was $P2_122_1$.* A b glide appearing in the photographs (0kℓ absent for k = 2n + 1) proved to be merely a systematic weakness, since over

---

*The standard setting for this space groups is $P2_12_12$ (4). Conversion of the nonstandard to the standard setting involves an interchange of the b and c axes. However, use of the standard setting is mandatory for publication purposes and for the operation of certain computer programs. Since this paper is mainly concerned with the relationship of the trial structure (as solved in the centric space group Pbcm) to the true structure (as it appears in the acentric space group $P2_122_1$), and since such a switching of axes would cause undue confusion, the nonstandard setting has been used throughout.

30% of the reflections in that group had intensities greater than 4σ (I). Both photographs and diffractometer data revealed that the h = 2n + 1 reflections were systematically weak; with few exceptions they differed in intensity from the h = 2n reflections by about two orders of magnitude. The few h-odd reflections of appreciable magnitude were shown in the photographs to be quite diffuse, while the h-even reflections were quite sharp.

The presence of large amounts of pseudosymmetry has notoriously proven to be a sizeable hurdle in structure solution by direct methods. Therefore, the MULTAN74 program was used to obtain a trial structure in a space group determined by eliminating the systematically weak reflections. The trial structure was then to be used as a basis for determining the true structure in the correct space group.

The space group chosen for solving the trial structure was Pbcm, which had the advantage of being centrosymmetric. All reflections for which h = odd were eliminated from this data set, and the false b glide was assumed to be real. In addition, since the reflections h0ℓ, ℓ = 2n + 1 were very nearly extinguished for all h = 2n reflections, a c glide could be assumed to exist once the h-odd reflections were eliminated. This effectively cut the a axis in half, leaving a unit cell containing only four molecules. Since Z = 8 in Pbcm, this means that the asymmetric unit of this space group contained one half of one molecule.

The elimination of all reflections for which h = odd, along with the addition of the two glide planes, reduced the number of observed

reflections from 2315 to 1107. The 1107 reflections contained in this data set were scaled by Wilson statistics and then used as input to MULTAN. The trial structure given by MULTAN had the planes of the six-membered rings lying in the mirror planes at $\underline{z}$ = 1/4 and z = 3/4. The molecules were thus located at the four special positions ($\underline{x}$, $\underline{y}$, 1/4), ($\underline{x}$, $\underline{y}$, 3/4), $\underline{x}$, 1/2 + $\underline{y}$, 1/4), and ($\underline{x}$ 1/2 - $\underline{y}$, 3/4) of Pbcm. Crystal data for the compound, in both the correct space group, $P2_122_1$, and the approximate space group, Pbcm, are given in Table I.

Attempts to solve the structure in the correct space group in terms of the trial structure met with little success. The chief difficulty lay in expanding the asymmetric unit of the centric space group into the much larger asymmetric unit of the acentric one--in short, to derive acentric phases from centric phases. The fourfold expansion of the asymmetric unit of Pbcm involved determining which fragments of the molecules were included in the true asymmetric unit, and their correct orientation as well. Consider, for instance, the orientation of the methyl carbons: Were they perhaps staggered with the SO_3 oxygen atoms in one of the molecules of the asymmetric unit, and eclipsed in the other? Or, was the difference more subtle--say a difference of a few degrees in the methyl-carbon-SO_3-oxygen torsion angle in the two independent molecules? Were the planes of the six-membered rings perhaps tilted differently with respect to the $\underline{z}$ = 1/4 plane? In any event, the major features of the structure were clear (see Discussion); and so it was decided that the relatively minor information to be gained was not worth the large investment of time and money needed to ferret out those few remaining details.

DISCUSSION

Although it was not possible to solve the structure in the correct space group, it is possible to draw some conclusions about the true structure by considering the trial structure, the crystal structure of the parent compound MSE, and the chemical nature of ZWT itself.

1) The molecules are arranged with their long axes (the axis running from the $N(CH_3)_3^+$ group to the SO_3^- group) parallel to the $\underline{a}$ axis of the crystal. The planes of the benzene rings lie very close to the false mirror planes at $\underline{z} = 1/4$ and $\underline{z} = 3/4$.

2) Because of the zwitterionic nature of the compound, the molecules must be arranged so that opposite charges are adjacent. Thus the packing should be very similar to that of MSE, with the ZWT molecules stacked head-to-tail on top of one another (see Figure 2).

3) In MSE the planes of the six-membered rings are not parallel but are tilted with respect to one another by an angle of 76° (see Figure 2). Therefore, it is possible that they are not perfectly parallel in the ZWT structure either. Perhaps the planes of the rings lie close to, but not exactly at, the planes at $\underline{z} = 1/4$ and $\underline{z} = 3/4$; or perhaps the long axis of the molecule does lie in the plane, but the plane of the benzene rings are tilted with respect to the plane in such a way as to destroy the symmetry.

4) The true space group of ZWT contains two molecules in the asymmetric unit. The large amount of pseudosymmetry indicates that the structural differences which give rise to the crystallographic

nonequivalence of two chemically identical species must be very minor. Perhaps conformational isomerism is involved (e.g., a difference in the torsion angles of the SO_3 oxygens with the methyl groups of the $N(CH_3)_3$ groups).

5) The diffuse character of the reflections for which h = odd probably represents long-range disorder extending over many unit cells. The nature of the disorder can be determined only after the structure has been solved. All that can be said at this point is that the disorder is restricted to the a direction.

REFERENCES

1. Sukenik, C.M., J.A.P. Bonapace, N.S. Mandel, R.G. Bergman, P-Y. Lau and G. Wood (1975). J. Amer. Chem. Soc. 97, 5290-5291.

2. Sukenik, C.N., J.A.P. Bonapace, N.S. Mandel, P-Y. Lau, G. Wood, and R.G. Bergman (1977). J. Amer. Chem. Soc. 99, 851-858.

3. Sukenik, C.N. (1976). Ph.D. Thesis, Caltech.

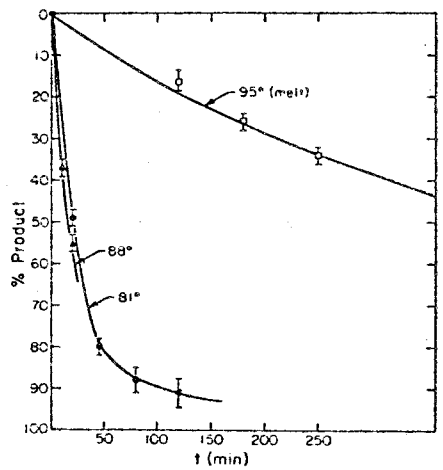
4. International Tables for X-ray Crystallography, Vol. I. Kynoch Press, Birmingham, 1962.

TABLE I. Crystal Data for p-triethylammoniumbenzene sulfonate.

Space Group Pbcm	Space Group $P2_122_1$*
a = 10.435(2) Å	a = 20.869(2) Å
b = 10.170(1) Å	b = 10.170(1) Å
c = 9.550(1) Å	c = 9.550(1) Å
V = 1104 $Å^3$	V = 2207 $Å^3$
ρ_{calc} = 1.41 g cm^{-3}	ρ_{calc} = 1.41 g cm^{-3}
μ = 18.66 cm^{-1}	μ = 18.66 cm^{-1}
F.W. = 215.27	F.W. = 215.27
½ molecule/asymmetric unit	2 molecules/asymmetric unit
Z = 4	Z = 8

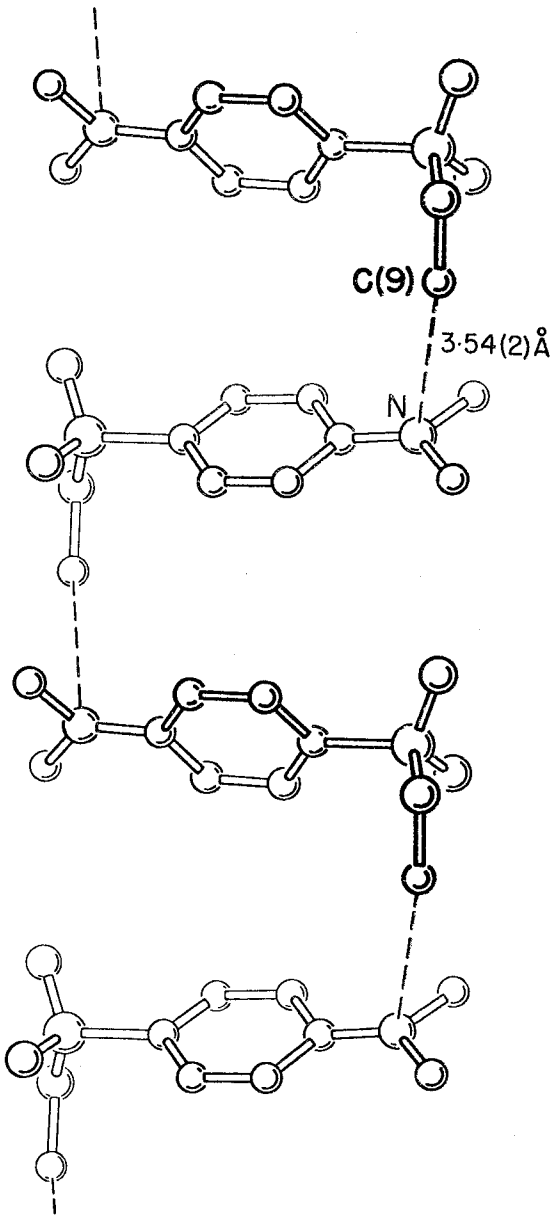
*Nonstandard setting.

Figure 1



Time dependence of the percent of product observed in the thermal conversion of methyl-p-dimethylaminobenzene sulvonate (I) to p-trimethyl-ammoniumbenzene sulfonate (II) at three different temperatures: ● = 81°, crystal; ▲, 88°, crystal; ■, 95°, melt.

Figure 2



A view of the stacking along one chain of molecules in crystals of methyl-p-dimethylaminobenzene sulfonate (I), as seen perpendicular to the [10$\bar{1}$] plane. Distance indicated is that between the carbon atom of the methyl group, which undergoes transfer in the solid state reaction, and the nitrogen atom to which it moves.

APPENDIX A

Derivation of the Expression for the Wilson Plot

The average value of the observed intensities of all reflections in the data set can be defined as:

$$\bar{I}_{rel} = \langle |F_{rel}|^2 \rangle_{avg}$$

where the F_{rel}'s are the unscaled, observed F's (F_{obs}) which have been corrected for 1/Lp. The average relative intensity $\bar{I}_{rel}$ is related to the theoretical (absolute) average intensity $\bar{I}_{abs}$ by a scaling factor C, where

$$\bar{I}_{rel} = C\bar{I}_{abs}$$

The theroetical average intensity is dependent only on the kinds and number of atoms in the unit cell, and not on their location within the unit cell (1). For a unit cell containing N atoms

$$\bar{I}_{abs} = \sum_{i=1}^{N} f_i^2$$

and hence

$$\bar{I}_{rel} = C \sum_{i=1}^{N} f_i^2$$

The scattering factor f for an atom is a function of $\sin\theta/\lambda$. In a real crystal, the scattering factor of an atom is dampened by the effects of thermal vibration. The higher the temperature, the greater will be the falloff in intensity due to thermal vibration. Thus, for a particular kind of atom the scattering power can be defined as

$$f = f_0 \exp(-B \sin^2\theta/\lambda^2)$$

where B is related to the mean-square amplitude of atomic vibration, $\overline{(u^2)}$, by the expression

$$B = 8\pi^2 \overline{(u^2)}$$

and f_0 represents the scattering power of a stationary atom. Only the f_0's are known a priori--they are listed in Vlume III of International Tables for X-ray Crystallography as functions of $\sin\theta/\lambda$ (2).

Thus, the complete expression for $\overline{I}_{rel}$ is given by

$$\overline{I}_{rel} = C\overline{I}_{abs} = C \sum_{i=1}^{N} f_{o_i}^2 \exp(-2B \sin^2\theta/\lambda^2)$$

where B is to be determined. As a first approximation, B can be assumed to have the same value for all atoms in the structure. In that case, the expression for the temperature factor can be taken outside the summation

$$\bar{I}_{rel} = C \exp(-2B \sin^2\theta/\lambda^2) \sum_{i=1}^{N} f_{o_i}^2$$

This is the expression given in Chapter I (p. 5). If we divide both sides by $\sum_{i=1}^{N} f_{o_i}^2$ and take the natural logarithm of both sides, we get

$$\ln \frac{\bar{I}_{rel}}{\sum_{i=1}^{N} f_{o_i}^2} = \ln C - 2B \sin^2\theta/\lambda^2$$

which is the expression used in making the Wilson plot. A plot of the left-hand side of the expression versus $\sin^2\theta/\lambda^2$ should give a straight line. The slope of this line is equal to -2B, and its intercept with the y axis ($\sin^2\theta/\lambda^2 = 0$) gives the value of ln C. The F_{obs}'s of the data set can then be put on an absolute scale by use of the scale factor k

$$F_{abs} = k F_{obs}$$

where $k = 1/\sqrt{C}$.

REFERENCES

1. Stout, G.H. and L.H. Jensen. X-ray Structure Determination: A Practical Guide. MacMillan Company, New York, 1968. pp. 205-208.

2. International Tables for X-ray Crystallography, Vol. III. Kynoch Press, Birmingham, 1962. pp. 201-207.