Table 2 (*cont.*)

Ni(4)–Hf(1)	2·632 (3)	Ni(6)–Hf(1)	2·637 (3)
Hf(2)	2·654 (3)	Hf(1)	3·249 (3)
Hf(2)	2·713 (3)	Hf(2)	2·618 (3)
Hf(3)	2·665 (3)	Hf(3)	2·658 (3)
Hf(3)	2·722 (3)	Hf(3)	2·673 (3)
Ni(1)	2·421 (4)	Ni(1)	2·427 (4)
Ni(2)	2·731 (4)	Ni(2)	2·896 (4)
Ni(2)	2·851 (4)	Ni(3)	2·679 (4)
Ni(5)	2·773 (4)	Ni(3)	3·324 (4)
Ni(5)	3·216 (4)	Ni(4)	2·707 (4)
Ni(6)	2·707 (4)	Ni(4)	2·763 (4)
Ni(6)	2·763 (4)	Ni(5)	2·409 (4)
Ni(7)	2·559 (3)	Ni(7)	2·561 (3)
Ni(5)–Hf(1)	2·616 (3)	Ni(7)–Hf(1) ×2	2·727 (1)
Hf(1)	2·621 (3)	Hf(2) ×2	2·815 (1)
Hf(2)	2·635 (3)	Hf(3) ×2	2·795 (1)
Hf(3)	2·702 (3)	Ni(2) ×2	2·559 (3)
Ni(1)	2·742 (4)	Ni(4) ×2	2·559 (3)
Ni(1)	2·960 (4)	Ni(6) ×2	2·561 (3)
Ni(3)	2·750 (4)		
Ni(3)	2·801 (4)	Ni(8)–Hf(1) ×2	2·810 (1)
Ni(4)	2·773 (4)	Hf(2) ×2	2·786 (1)
Ni(4)	3·216 (4)	Hf(3) ×2	2·870 (1)
Ni(5)	2·563 (5)	Ni(1) ×2	2·470 (3)
Ni(6)	2·409 (4)	Ni(3) ×2	2·558 (2)
Ni(8)	2·517 (3)	Ni(5) ×2	2·517 (3)

The author is deeply indebted to Professor Stig Rundqvist for valuable discussions and comments. Financial support from the Swedish Natural Science Research Council is gratefully acknowledged.

References

BSENKO, L. (1978*a*). To be published.

BSENKO, L. (1978*b*). *Acta Cryst.* B**34**, 3201–3204.

CROMER, D. T. & LIBERMAN, D. (1970). *J. Chem. Phys.* **53**, 1891–1898.

CROMER, D. T. & WABER, J. T. (1965). *Acta Cryst.* **18**, 104–109.

ERSSON, N.-O. (1976). In *Crystallographic Computer Programs*, edited by J.-O. LUNDGREN. Report UUIC-B13-4-03. Univ. of Uppsala, Sweden.

International Tables for X-ray Crystallography (1969). Vol. I, 3rd ed. Birmingham: Kynoch Press.

KIRKPATRICK, M. E. & LARSEN, W. L. (1961). *Trans. Am. Soc. Met.* **54**, 580–590.

LUNDGREN, J.-O. (1976). Editor. *Crystallographic Computer Programs*. Report UUIC-B13-4-03. Univ. of Uppsala, Sweden.

Acta Cryst. (1978). B**34**, 3210–3215

The X-ray Structure of Yeast Inorganic Pyrophosphatase at 5·5 Å Resolution

BY GERARD BUNICK,* GEORGE P. MCKENNA, F. E. SCARBROUGH,† EDWARD C. UBERBACHER AND DONALD VOET‡

Department of Chemistry and Laboratory for Research on the Structure of Matter, University of Pennsylvania, Philadelphia, Pennsylvania 19104, *USA*

(*Recieved* 6 *March* 1978; *accepted* 24 *May* 1978)

The structure of yeast inorganic pyrophosphatase has been determined at low resolution (5·5 Å). The phases of the native X-ray intensities are based on the isomorphous replacement and anomalous dispersion data of a single high-quality $Hg(SCN)_4^{2-}$ derivative of the protein. The mean figure of merit for 2214 reflections is 0·817. Crystals of yeast inorganic pyrophosphatase contain one dimeric molecule of molecular weight 64 000 in the asymmetric unit. The overall dimensions of the ellipsoidal molecule, which is clearly distinguishable from the background, are approximately 100 × 50 × 50 Å. The ellipsoid is somewhat indented about its equator with the resulting two lobes of the peanut-shaped dimer being related by an apparent molecular twofold axis.

Introduction

Inorganic pyrophosphatase (EC 3.6.1.1 pyrophosphate phosphohydrolase) catalyzes the hydrolysis of inorganic pyrophosphate to orthophosphate. This reaction disposes of the pyrophosphate ion generated in various biosynthetic reactions thereby driving them to completion (Kornberg, 1962). The failure to find any life forms that lack inorganic pyrophosphatase activity has led to the conclusion that this enzyme is indispensible to living systems (Josse & Wong, 1971).

In higher organisms inorganic pyrophosphatase has been implicated in the formation and maintenance of bones and teeth. The nucleation and growth of hydroxyapatite, a crystalline form of calcium phos-

* Present address: Chemistry Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830.

† Present address: HFF-334, Food and Drug Administration, 200 C Street, SW, Washington, DC 20204.

‡ To whom correspondence should be addressed.

phate that is the predominant constituent of mineralized tissues, is inhibited by pyrophosphate (Graham, Russell & Fleisch, 1970). The fact that body fluids are supersaturated with calcium phosphate indicates that there must be a factor present that controls the formation and growth of hydroxyapatite crystals in the various tissues. There is considerable evidence suggesting that this factor is the pyrophosphate ion (Graham, Russell & Fleisch, 1970; Rasmussen & Bordier, 1974). The level of pyrophosphate is, in turn, controlled by the activity of inorganic pyrophosphatase.

The properties of inorganic pyrophosphatase from bakers' yeast (PPase) have been reviewed by Butler (1971). PPase, which was initially reported by Bailey & Webb (1944) and was first purified and characterized by Kunitz (1952), is a dimer of identical subunits that each consist of 285 amino acids. The dimeric enzyme has a molecular weight of 64 042. The amino acid sequence of PPase has been recently reported by Cohen, Sterner, Keim & Heinrikson (1978). Of potential crystallographic significance is the fact that each PPase protomer contains a single enzymatically non-essential cysteine residue (Hansen, Eifler & Heitman, 1972; Heinrikson, Sterner, Noyes, Cooperman & Bruckman, 1973).

PPase, as do most phosphate transfer enzymes, requires a divalent metal ion for activity (Cooperman, 1976). Mg^{2+}, the ion that is normally associated with the enzyme, confers the greatest catalytic efficiency on PPase although other metal ions such as Zn^{2+}, Co^{2+} and Mn^{2+} are also effective. In the presence of Mg^{2+}, PPase is specific for the hydrolysis of the pyrophosphate ion (Heppel & Hilmoe, 1951; Kunitz, 1952). However, with other divalent metal ions the specificity of PPase is broadened to include a wide variety of organic phosphates (Schlesinger & Coon, 1960; Avaeva, Kara-Murza & Botvinik, 1967).

There have been several studies of PPase with the aim of determining the enzymatic role of specific functional groups of the protein. These have established that there is an essential arginine residue at the active site of PPase which is vital to pyrophosphate binding (Cooperman & Chiu, 1973; Heitmann & Uhlig, 1974); that there is a carboxyl group near the active site that is not required for metal ion binding but which is involved in the catalytic process (Cooperman & Chiu, 1973; Heitmann & Uhlig, 1974); and that a methionine residue (Yano, Negi & Irie, 1973) and a tryptophan residue (Negi, Samejima & Irie, 1972) are essential for enzymatic activity.

Experimental procedures

Rectangular plate-like crystals of PPase were grown at 4 °C by vapor diffusion from 2-methyl-2,4-pentanediol solutions that were buffered at pH 6·0 with 2-(*N*-morpholino)ethanesulfonic acid. Crystals grew over a period of several weeks to reach typical dimensions of 0·6 × 0·4 × 0·08 mm (Bunick, McKenna, Colton & Voet, 1974). These crystals have the space group $P2_1$. The unit-cell dimensions, as determined by the least-squares analysis of the angular positions of 12 reflections on a Picker FACS-I diffractometer, are a = 69·96 (11), b = 95·27 (10), c = 51·77 (9) Å and β = 99·50 (6)°. There is one dimeric molecule of PPase in the asymmetric unit. The previously reported crystal density of 1·15 g cm^{-3} (Bunick, McKenna, Colton & Voet, 1974) together with the now accurately known molecular weight of the PPase dimer (Cohen, Sterner, Keim & Heinrikson, 1978) leads to the conclusion that the crystal volume per unit of protein molecular weight, V_M, is 2·66 Å^3 dalton^{-1}. This value is slightly greater than the median of the various crystalline proteins that were compiled by Matthews (1968).

In order to prepare heavy-atom derivatives, crystals of PPase were transferred to an artificial mother liquor composed of 30% 2-methyl-2,4-pentanediol and 0·03 *M* 2-(*N*-morpholino)ethanesulfonic acid buffer adjusted to pH 6·0. It was found that crystals of PPase, which are relatively fragile, often shattered if the concentrations of various heavy-atom-containing salts were rapidly raised to levels that would minimally be required to form usable isomorphous derivatives. In order to conveniently bring about a gradual increase of the heavy-atom concentration in the vicinity of the crystals the following technique was adopted. Several crystals of PPase in 0·05 ml of artificial mother liquor were placed in one well of a double-well slide and 0·05 ml of artificial mother liquor containing the heavy-atom salt of choice at twice the required final concentration was placed in the other well. The droplets in the two wells were bridged by a short length of thread. The slide was then sealed inside a box containing 10 ml of artificial mother liquor and stored at 4 °C. Over a period of two to three weeks the heavy-atom solution diffused along the thread until its concentration in both wells of the slide had equilibrated. With the use of this technique, crystals of PPase rarely cracked.

Soaking PPase crystals, as described above, in a 0·15 m*M* solution of $K_2Hg(SCN)_4$ for a minimum of 60 days caused visually apparent alterations in the intensities of the photographically recorded reflections with less than 0·6% changes in the unit-cell parameters. Almost identical changes were observed with K_2HgI_4 under similar conditions but, due to possible difficulties in determining the positions of the I atoms of HgI_4^{2-}, only the $K_2Hg(SCN)_4$ derivative was fully analyzed.

X-ray diffraction data were measured at 18 °C by a Picker FACS-I diffractometer employing monochromated Cu $K\alpha$ radiation and with the 60 cm crystal-to-detector distance largely occupied by a 48 cm helium filled tube. An ω step scan procedure was used in which counts of four seconds duration each were taken at

seven ω settings that were 0·03° apart. Background measurements of two seconds duration each were taken at $2\theta \pm 0{\cdot}5°$. Using this procedure reflections were measured at an average rate of 55 per hour.

A total of 2236 unique reflections of the native crystal were measured in four resolution shells to a maximum resolution of 5·5 Å. At the end of the data collection period the average intensities of the standard reflections had fallen to 92·0% of their original values. Similarly, 4596 reflections of the $Hg(SCN)_4^{2-}$ derivative crystal were measured in five shells to a resolution of 5·5 Å. Here the measurements of 25 reflections were followed by the measurements of their 25 Friedel mates. The average intensities of the standard reflections had fallen to 81·0% of their original values at the end of this data collection procedure.

Integrated intensities were calculated by fitting the scan data to a Gaussian curve (Diamond, 1969). The background count at each reflection was taken to be the average of the background counts in a block of 3^3 reflections centered in reciprocal space on the reflection in question. Empirical cylindrical absorption corrections were applied to the intensities (North, Phillips & Mathews, 1968).

Many of the computations required for this structural analysis were made using the programming system *PROTSYS* kindly supplied by G. A. Petsko.

Results and discussion

(*a*) *The heavy-atom derivatives*

The Hg atoms were located in three dimensions by the analyses of Patterson maps based on (i) the isomorphous differences, $(F_{PH} - F_P)^2$, in which the Friedel pairs of the $Hg(SCN)_4^{2-}$ derivative were averaged; (ii) the anomalous dispersion differences, $(F_{PH}^+ - F_{PH}^-)^2$; and (iii) the Matthews combination coefficients (Matthews, 1966*a*):

$$F_H^2 = F_P^2 + F_{PH}^2 - 2F_P F_{PH} \times \{1 - [wk(F_{PH}^+ - F_{PH}^-)/2F_P]^2\}^{1/2}. \quad (1)$$

Here F_P is the structure factor amplitude of the native protein, F_{PH} is that of the derivative in the absence of anomalous scattering, F_H is the amplitude of the heavy-atom contribution to the derivative structure factor and F_{PH}^+ and F_{PH}^- are the structure factor amplitudes of a Friedel pair of derivative reflections. The weighting factor, w, was arbitrarily set to 0·75 as was suggested by Matthews (1966*a*), and k, ratio of the real to anomalous scattering, was determined empirically as a function of $\sin\theta/\lambda$. The Harker sections of these maps are displayed in Fig. 1. In all three cases the four highest peaks in the full Patterson map led to the same heavy-atom positions. The ratio of the heights of the lowest of the four Patterson solution peaks to that of the next lowest peak in the map was 1·64 for the isomorphous difference map, 1·26 for the anomalous difference map and 2·00 for the combination coefficients map.

The heavy-atom parameters were refined by a least-squares procedure according to the F_{HLE} method (Dodson & Vijayan, 1971). The refinement, in which the isotropic thermal parameters were arbitrarily set to 50 Å^2, converged with $R = \sum |F_{HLE} - F_{H(\text{obs})}| / \sum F_{HLE} = 0{\cdot}356$ based on 1529 reflections. The final atomic parameters for the Hg atoms are given in Table 1. It should be noted that the true values of the occupancy factors may be considerably less than the values given. More realistic values for these quantities await phase refinement based on more than one heavy-atom derivative (Dodson & Vijayan, 1971).

The 'double' difference Fourier map based on the heavy-atom parameters was essentially featureless. The height of its highest peak was one tenth of that expected for a full Hg atom.

(*b*) *The electron density map*

Protein phases were calculated according to the method of Blow & Crick (1959). Anomalous-scattering data were incorporated in the phase calculation by the

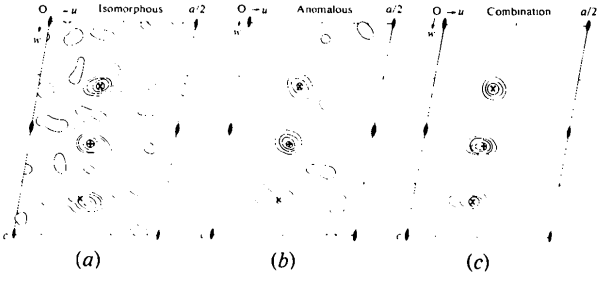

Fig. 1. The Harker sections ($v = \frac{1}{2}$) of the difference Patterson maps for the $Hg(SCN)_4^{2-}$ derivative of PPase. Contours begin at one fifth of the maximum non-origin peak and continue at the same interval. Negative contours are dashed and the zero contour has been omitted. Self-vectors are indicated by (×) and the two vectors relating independent atoms, which fortuitously coincide on the Harker plane for this derivative, are indicated by (+). The maps are based on: (*a*) the isomorphous differences $(F_{PH} - F_P)^2$; (*b*) the anomalous dispersion differences $(F_{PH}^+ - F_{PH}^-)^2$; and (*c*) the Matthews combination coefficients of equation (1).

Table 1. *Final heavy-atom parameters*

Positional parameters are given in fractions of a unit-cell edge. Estimated standard deviations are given in parentheses and refer to the least significant digits of their corresponding parameters.

Site	x	y	z	Occupancy factor
Hg(1)	0·3926 (6)	0·2703 (5)	0·0805 (8)	0·94 (2)
Hg(2)	0·3939 (5)	0·2500	0·3505 (7)	1·05 (2)

method of North (1965) as was modified by Matthews (1966*b*). The quality of the final phase set can be judged from the indicators presented in Table 2 and Fig. 2.

The proper heavy-atom enantiomorph was chosen by the comparison of the electron-density map of the protein based on each of the heavy-atom enantiomorphs. The map that was based on the chosen enantiomorph, as will be seen below, was easily interpretable. In contrast, the regions of positive density in the Fourier map based on the other enantiomorph formed no entity that was recognized as a protein molecule, but rather, were uniformly distributed throughout the unit cell.

The boundary of the PPase molecule, for the most part, is clearly distinguishable from its surrounding milieu in the 'best' 5·5 Å electron density map. This can be seen in Fig. 3. The calculated r.m.s. error of the map is 0·054 e Å^{-3}.

Fig. 4 consists of several stereoviews of a balsa-wood model of the PPase dimer. The molecule, which is roughly ellipsoidal, has approximate dimensions of 100 × 50 × 50 Å. The major axis of the ellipsoid is inclined by an angle of about 30° to the *c* axis. An indentation about the equator of the ellipsoid imparts a peanut-like shape to the molecule. The two lobes of the molecule

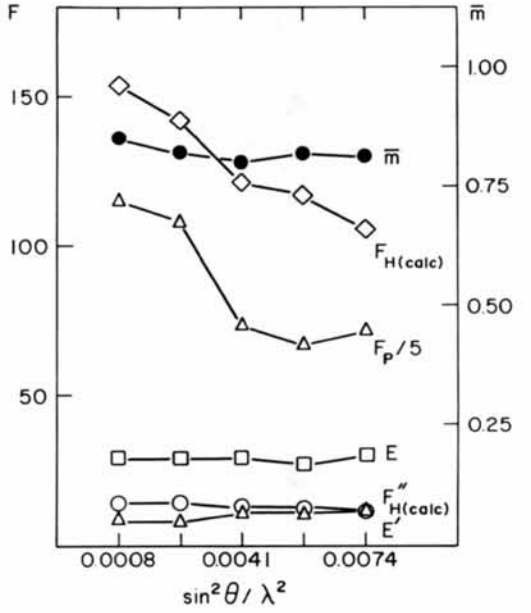

Fig. 2. The variation with $\sin^2\theta/\lambda^2$ of: the r.m.s. value of the observed native protein structure factors, F_P; the r.m.s. value of the calculated heavy-atom contributions, $F_{H(calc)}$; the r.m.s. value of the calculated anomalous components of the heavy-atom contributions, $F''_{H(calc)}$; the r.m.s. lack of closure, E; the r.m.s. anomalous lack of closure, E'; and the mean figure of merit, $\bar{m}$.

Table 2. *Phasing results*

	Centric	Acentric	Total
$\langle F_P^2\rangle^{1/2}$	391	404	403
$\langle F_{H(calc)}{}^2\rangle^{1/2}$	128	121	122
$\langle (F_{PH}-F_P)^2\rangle^{1/2}$	128	79	84
$\langle F''^2_H\rangle^{1/2}$	–	13	12
E	54	26	29
E'	17*	10	10
$\bar{m}$	0·82	0·82	0·82
Number of reflections	158	2056	2214

$E = \langle (F_{PH} - F_{PH(calc)})^2\rangle^{1/2}$.
$E' = \langle \{|F^+_{PH} - F^+_{PH(calc)}| - |F^-_{PH} - F^-_{PH(calc)}|\}^2\rangle^{1/2}$.

* The non-zero value of E' for centric reflections is indicative of the errors in measurement of their intensities.

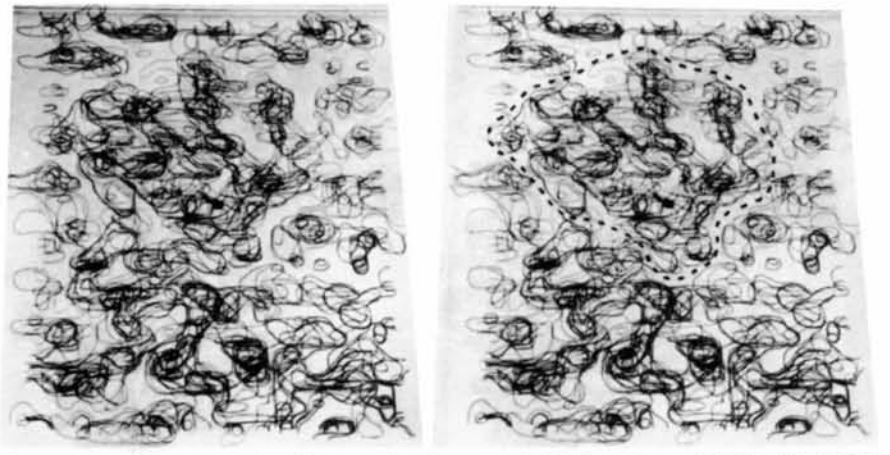

Fig. 3. A stereoview of ten consecutive contoured sections, ranging from $z = 0{\cdot}300$ through $z = 0{\cdot}525$, of the 5·5 Å resolution electron-density map of PPase. The *b* axis extends vertically on the page and the *a* axis is horizontal. Contours are drawn at the electron density levels 0·041, 0·103, 0·205 and 0·307 e Å^{-3}. The $F_P(000)$ term has not been included in the calculation of the map. A portion of the molecule of PPase, corresponding to much of the lower half of the molecule as seen in Fig. 4, is enclosed by a dashed line. The position of mercury atom Hg(2) is indicated by a large dot.

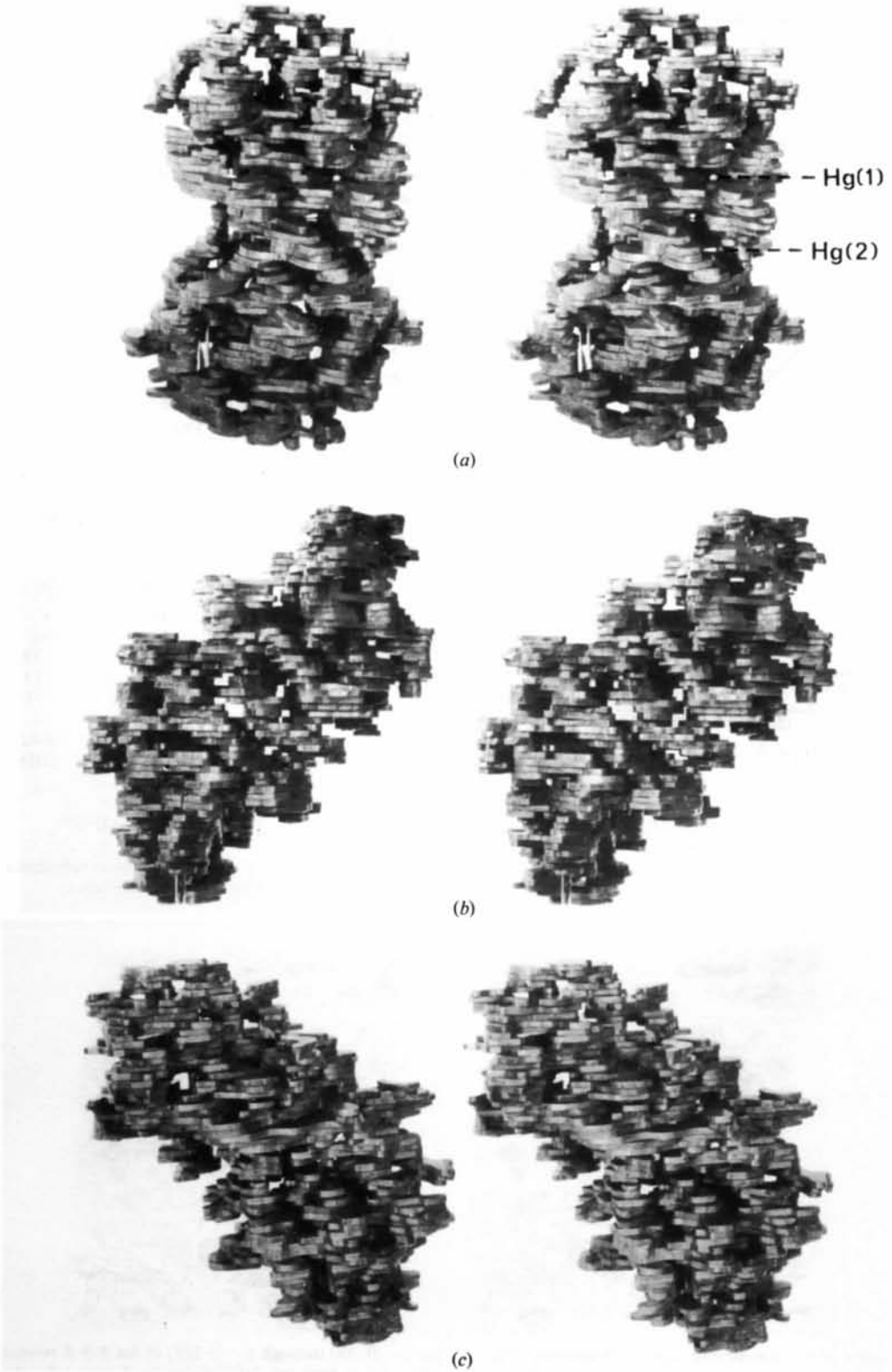

Fig. 4. Stereoviews of a balsa-wood model of the electron-density map. Sections, which are taken parallel to the *ab* plane, represent a thickness of 1·3 Å. The model surface follows the 0·041 e $Å^{-3}$ contour. (*a*) A view approximately normal to the molecular twofold axis. The positions of the two Hg atoms are indicated by small beads. (*b*) A view approximately along the molecular twofold axis. (*c*) A view towards the opposite side of the molecule from that pictured in (*b*).

appear to be related by a non-crystallographic twofold axis that is roughly parallel to the *ab* plane and is inclined at an angle of about 50° to the *b* axis. Fig. 4(*b*) is a view along this molecular twofold axis. It can be seen in Fig. 4 that the PPase molecule contains many long rod-like regions of electron density. These features, which range from 25 to 35 Å in length, are about 6 Å thick and are separated by about 10 Å. This suggests that these rods are α-helices.

The two Hg atoms, which can be seen in Fig. 4(*a*), are located in crevices near the surface of the molecule. They are symmetrically arrayed, at a radius of 7 Å, about the molecular twofold axis. The fact that these atoms are buried beneath the surface of the molecule rationalizes the relatively long soaking times required to form a useful $Hg(SCN)_4^{2-}$ derivative of PPase.

It has been shown (M. Bond and B. S. Cooperman, personal communication) that an arginine residue which is essential for the enzymatic activity of PPase (Cooperman & Chiu, 1973), and which is implicated in substrate binding, is located at position 77 and/or 80 in the amino acid sequence of PPase (Cohen *et al.*, 1978). Both of these arginine residues are relatively close to the only cysteine in the PPase subunit, cysteine 82, which presumably covalently binds the Hg atom in the $Hg(SCN)_4^{2-}$ derivative of PPase. Hence the active site of PPase must be in the vicinity of this Hg atom.

The PPase dimers form networks of interlocking translationally related molecules parallel to the *ac* plane. Neighboring networks of enzyme molecules are related by the crystallographic twofold screw axis.

This investigation is being extended to higher resolution and the search for additional heavy-atom derivatives of PPase is being actively pursued.

We are grateful to Drs M. J. Adams, G. A. Petsko and C. C. F. Blake for helpful discussions. This work was supported, in part, by grants from the National Institutes of Health (GM20568) and from the National Science Foundation, MRL Program (DMR76-00678). Two of us (GB & FES) were postdoctoral fellows of the National Institutes of Health (AM05109 & AM05105).

References

AVAEVA, S. M., KARA-MURZA, S. N. & BOTVINIK, M. M. (1967). *Biokhimiya*, **32**, 205–209.

BAILEY, K. & WEBB, E. C. (1944). *Biochem. J.* **38**, 394–398.

BLOW, D. M. & CRICK, F. H. C. (1959). *Acta Cryst.* **12**, 794–802.

BUNICK, G., McKENNA, G. P., COLTON, R. & VOET, D. (1974). *J. Biol. Chem.* **249**, 4647–4649.

BUTLER, L. G. (1971). *The Enzymes*, Vol. IV, edited by P. D. BOYER, pp. 529–541. New York: Academic Press.

COHEN, S. A., STERNER, R., KEIM, P. S. & HEINRIKSON, R. L. (1978). *J. Biol. Chem.* **253**, 889–897.

COOPERMAN, B. S. (1976). *Metal Ions in Biological Systems*, Vol. 5, edited by H. SIGEL, pp. 76–125. New York: Marcel Dekker.

COOPERMAN, B. S. & CHIU, N. Y. (1973). *Biochemistry*, **12**, 1676–1682.

DIAMOND, R. (1969). *Acta Cryst.* A**25**, 43–55.

DODSON, E. & VIJAYAN, M. (1971). *Acta Cryst.* B**27**, 2402–2411.

GRAHAM, R., RUSSELL, G. & FLEISCH, H. (1970). *Clin. Orthop.* **69**, 101–117.

HANSEN, G., EIFLER, R. & HEITMAN, P. (1972). *Acta Biol. Med. Ger.* **28**, 977–987.

HEINRIKSON, R. L., STERNER, R., NOYES, C., COOPERMAN, B. S. & BRUCKMAN, R. H. (1973). *J. Biol. Chem.* **248**, 2521–2528.

HEITMANN, P. & UHLIG, H. J. (1974). *Acta Biol. Med. Ger.* **32**, 565–574.

HEPPEL, L. A. & HILMOE, R. J. (1951). *J. Biol. Chem.* **192**, 87–94.

JOSSE, J. & WONG, S. C. K. (1971). *The Enzymes*, Vol. IV, edited by P. D. BOYER, pp. 499–527. New York: Academic Press.

KORNBERG, A. (1962). *Horizons in Biochemistry*, edited by M. KASHA & B. PULLMAN, pp. 251–264. New York: Academic Press.

KUNITZ, M. (1952). *J. Gen. Physiol.* **35**, 423–450.

MATTHEWS, B. W. (1966*a*). *Acta Cryst.* **20**, 230–239.

MATTHEWS, B. W. (1966*b*). *Acta Cryst.* **20**, 82–86.

MATTHEWS, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

NEGI, T., SAMEJIMA, T. & IRIE, M. (1972). *J. Biol. Chem.* **71**, 29–37.

NOCKOLDS, C. E. & KRETSINGER, R. H. (1970). *J. Phys. E*, **3**, 842–846.

NORTH, A. C. T. (1965). *Acta Cryst.* **18**, 212–216.

NORTH, A. C. T., PHILLIPS, D. C. & MATHEWS, F. S. (1968). *Acta Cryst.* A**24**, 351–359.

RASMUSSEN, H. & BORDIER, P. (1974). *The Physiological and Cellular Basis of Metabolic Bone Disease.* Baltimore: Williams and Wilkins.

SCHLESINGER, M. J. & COON, M. J. (1960). *Biochim. Biophys. Acta*, **41**, 30–36.

YANO, Y., NEGI, T. & IRIE, M. (1973). *J. Biochem. (Tokyo)*, **74**, 67–76.