

The Structure of α -Chymotrypsin. I. The Refinement of the Heavy-Atom Isomorphous Derivatives at 2.8 Å Resolution

BY A. TULINSKY, N. V. MANI,* CARL N. MORIMOTO† AND RICHARD L. VANDLEN‡

Departments of Chemistry and Biochemistry, Michigan State University, East Lansing, Michigan 48823, U.S.A.

(Received 12 October 1972; accepted 7 February 1973)

Phases of α -chymotrypsin (α -CHT) have been determined to 2.8 Å resolution by the multiple isomorphous replacement method using 6 heavy-atom derivatives: (1) Pt(+), (2) Pt, (3) Hg, (4) U, (5) U + Au, and (6) U(+). The Pt and U derivatives were concentration-dependent and each reagent yielded two useful derivatives while the U + Au derivative was a 'double' derivative. Bijvoet pairs were measured with the U(+) derivative to 5.2 Å resolution ($2\theta \leq 17^\circ$) in order to determine the absolute configuration of the molecule. Crystals of α -CHT (M.W. $\sim 25,300$) were grown from half-saturated $(\text{NH}_4)_2\text{SO}_4$ at about $\text{pH} = 4.2$. The crystals are monoclinic, space group $P2_1$, $a = 49.24$ (7), $b = 67.20$ (10), $c = 65.94$ (9) Å, $\beta = 101.79$ (8)°, 4 molecules per cell (2 molecules per asymmetric unit). Intensity data were measured to 2.8 Å resolution with a computer-controlled diffractometer operating with a crystal alignment feature and using a 'wandering' count 6-drop 2ω step-scan procedure. Reproducibility on the same crystal was about $R \sim 0.02$, 0.03 with different crystals and 0.04 for 'before' and 'after' data collections or as a function of extended X-ray exposure ($R = \sum ||F_i| - \langle |F| \rangle| / \sum \langle |F| \rangle$, where $|F_i|$ and $\langle |F| \rangle$ are the observed and average $|F|$ values, respectively). Two independent sets of intensity data to 3.5 Å resolution have been measured on 4 crystals of α -CHT (two independent preparations). The agreement between the two sets, after data reduction to $|F|_p$'s, was $R = 0.023$ for 4596 independent reflections. Using the average structure amplitudes of the two sets of measurements to 3.5 Å resolution (4990) and the remainder of the reflections to 2.8 Å resolution, phase determination and refinement were carried out *via* alternate cycles of phase calculation and least-squares refinement of heavy-atom parameters minimizing the lack of closure error of the 9044 observed native reflections (85% of theoretical maximum). The average figure of merit at 2.8 Å resolution is $\langle m \rangle = 0.76$ ($\text{MRE}/n = 0.50$, where $n = 6$). The average difference between maximum probability and 'best' phase angles is 14.3° . The r.m.s. calculated structure factors and corresponding lack of closure errors (in parentheses) for the various derivatives are: Pt(+) = 189 (55), Pt = 114 (49), Hg = 104 (58), U = 61 (46), U + Au = 104 (46), U(+) = 120 (47) e/u.c.

1. Introduction

The crystal and molecular structure of native α -chymotrypsin (α -CHT) has been under investigation in our laboratories since the spring of 1962 (Vandlen & Tulinsky, 1971; Tulinsky, Vandlen, Morimoto, Mani & Wright, 1973; Vandlen & Tulinsky, 1973). The structure of α -CHT, which was inferred from the structure of inhibited tosyl- α -CHT (toluenesulphonyl) *via* difference electron density methods, was originally reported by D. M. Blow and his collaborators at the Medical Research Council Laboratories (M.R.C.), Cambridge, England (Birktoft & Blow, 1972). Upon considering the reported results and comparing these with our work, it is apparent that there are a sufficient number of significant differences in the results, the approach and execution, and the philosophy of dealing with the problem to be of considerable general interest to several disciplines. Thus, we would like to report here the

first of a series of communications relating to our independent study.

Of the three heavy-atom isomorphous derivatives used for phase determination by the M.R.C. group, one derivative (PtCl_4^-) had a substitution of 380 electrons per asymmetric unit of molecular weight of about 50,500 [e.g. two enzyme molecules approximately related by a local or non-crystallographic twofold symmetry element (Blow, Rossmann & Jeffrey, 1964)]; for the other two derivatives, the total substitution was only about 100 electrons each. Our work is based on six derivatives with substitutions ranging from 135 to 545 electrons/asymmetric unit. Complications were introduced into the M.R.C. structure determination by the fact that the parent structure was not the same for all three heavy-atom derivatives (Blow & Matthews, 1973). The PtCl_4^- derivative was prepared from crystals of the native enzyme, but the derivative containing the irreversible inhibitor, pipsyl (*p*-iodophenylsulphonyl), was taken to be isomorphous with crystals of inhibited tosyl- α -CHT; the phenyl-mercuric acetate derivative also had the inhibited enzyme as the parent structure. The problems associated with the use of such procedures have already been clearly pointed out by North & Phillips (1969). Further complications were

* Deceased, 19 September 1972.

† Present address: Department of Biochemistry and Biophysics, Texas A & M University, College Station, Texas, U.S.A.

‡ Present address: Chemistry Division, California Institute of Technology, Pasadena, California, U.S.A.

introduced in the M.R.C. work when the mother liquor composition of the crystals was permitted to vary somewhat from derivative to derivative [2.0–2.6M (NH₄)₂SO₄ to 3.5M NaH₂PO₄, at times containing 2–4% dioxane by volume to eliminate twinning of the crystals (Sigler & Blow, 1965; Sigler, Jeffrey, Matthews & Blow, 1966; Matthews, Sigler, Henderson & Blow, 1967)].

We have prepared and crystallographically examined about ten heavy-atom isomorphous derivatives of varying quality and usefulness. Most of these derivatives showed substitutions which were very dependent upon the concentration of the heavy-atom reagent in equilibrium with the enzyme crystals (all of the M.R.C. derivatives are in this category). The behavior was particularly noticeable for the platinum halide anions; in fact, we were able to prepare two useful isomorphous derivatives with PtI₄⁻ by appropriately adjusting the concentration of the anion in equilibrium with the crystals. Since such experiments proved difficult to reproduce quantitatively (although qualitatively easy), all of the intensity data of a derivative were collected from *the same preparation*. In addition, some of the better derivatives show marked departures from local two-fold symmetry by displaying apparent substitution in one molecule but not in the other of the asymmetric unit. Lastly, since the native enzyme undergoes changes in structure in the active site region upon inhibition by tosylation of Ser 195, we avoided heavy atom 'tosylated' derivatives in our isomorphous series and the concomitant complications thereof.

Our system for intensity data collection has been a computer-controlled 4-circle diffractometer with automatic crystal alignment software (Vandlen & Tulinsky, 1971) in contrast to the photographic methods of intensity data collection used by the M.R.C. group. Investigation of some of our best heavy-atom derivatives at 2.0 Å resolution in two-dimensions revealed that, although the phase determining power of the derivatives was exceptionally good for the low order diffraction pattern, the phase determining ability deteriorated rapidly beyond 2.8 Å resolution. Even though our isomorphous derivatives were of comparatively excellent quality in comparison with many of those reported to date (as can be seen from the summary of statistics of the final refinement cycle, Table 4), nevertheless, we chose to confine our work based on multiple isomorphous substitution to the 2.8 Å resolution limit.

In summary, the electron density distributions of the two molecules in the asymmetric unit are different and we have considered them so throughout our work. On the other hand, the M.R.C. electron density distributions were sufficiently similar as to justify the averaging of the density of the two molecules at 2 Å resolution in much of their work. In addition, we have meticulously investigated the problem of irreversible inhibition, substrate-analog interaction, and reversible-equilibrium competitive inhibitor binding [about thirty 3-dimensional structures of modified α -CHT derivatives (Tulinsky, Coddling & Vandlen, 1971; Tulinsky *et al.*, 1973; Vandlen & Tulinsky, 1973)] in parallel with the structure determination of α -CHT. Some of the work between the M.R.C. group and us has been unavoidably redundant and much of it is in good agreement (especially concerning the gross features); some of the results are not in agreement; and, finally, most of the more recent work is totally new and generally unrelated. We hope this will be borne out in this communication and in the others which are soon to follow.

2. Experimental

(a) Crystals

Worthington Biochemical Corporation (Freehold, N. J.) α -CHT, 3 \times crystallized from the activation product of 3 \times crystallized chymotrypsinogen, dialyzed salt free and lyophilized, was used in all of our work without further purification or biochemical characterization. Crystals of the native enzyme were grown from half-saturated ammonium sulphate solutions containing 5–6 mg/ml of enzyme at pH=(3.8–4.2) in water baths at 14–16°C. After desired growth had been attained, the crystallization process was terminated by replacing the mother liquor with an approximately 75% saturated ammonium sulphate soaking solution at about pH=4.2. Crystals of native α -CHT obtained in this way are monoclinic, space group *P*2₁, with 4 molecules (M.W. ~ 25,300) per unit cell (*two molecules per asymmetric unit*) and have the unit-cell dimensions given in Table 1.

(b) Heavy-atom derivatives

Six heavy-atom isomorphous derivatives were used to determine the phases of the native enzyme by the multiple isomorphous replacement method. The derivatives have been designated as: (1) Pt(+), (2) Pt, (3)

Table 1. Unit-cell dimensions of α -CHT and six heavy-atom isomorphous derivatives*

Derivative	<i>a</i>	<i>b</i>	<i>c</i>	β	Number of crystals
α -CHT	49.24 (7) Å	67.20 (10) Å	65.94 (9) Å	101.79 (8)°	7
Pt(+)	49.15 (5)	67.16 (14)	65.88 (8)	101.93 (11)	2
Pt	49.33 (5)	67.39 (6)	66.00 (7)	101.89 (2)	5
Hg	49.26 (9)	67.14 (14)	66.10 (12)	101.82 (6)	8
U	49.22 (6)	67.32 (18)	65.96 (8)	101.86 (3)	8
U+Au	49.01 (6)	67.48 (9)	65.90 (3)	101.95 (3)	5
U(+)	49.05	67.71	65.88	101.97	1

* Standard deviations from the average are in parentheses.

Hg, (4) U, (5) U + Au, and (6) U(+); they were prepared using the following reagents: K_2PtI_4 , phenylmercuric acetate, $UO_2(NO_3)_2$, and $KAuI_4$. Anomalous dispersion measurements were also made with the U(+) derivative to 5.2 Å resolution ($2\theta \leq 17^\circ$) to determine the absolute configuration of the molecule.

In general, the heavy-atom isomorphous derivatives were prepared by diffusing the heavy-atom reagents into native enzyme crystals with solutions containing an appreciable molar excess of the reagent. The uranium derivative (U) was formed when the concentration of uranyl ion (UO_2^{2+}) exceeded about a (100:1) molar excess ratio (~ 10 mM). At higher concentrations, the number of prime substitution sites and their occupancies remained relatively independent of concentration and equilibrium was apparently established faster, so derivative crystals were prepared with about a (300:1) molar excess (~ 30 mM).

With the other isomorphous derivatives, the occupancy of a site and the number of substitution sites were extremely sensitive to the concentration of the heavy-atom reagent in contact with the soaking crystals. Since an accurate determination of the amount of enzyme contained in the crystals in contact with the soaking solution could not be made (the amount of enzyme was estimated from the total volume of the crystals, the enzyme composition of the crystals and the crystal density), the molar excess ratios were only approximate and could be in error by 50% or more; relative to one another, however, the molar excess ratios are much more precise than this figure. Consequently, the reproducibility of the conditions of a given experiment was only approximate. Moreover, a slow process seemed to be involved in attaining equilibrium. In the last aspect of our heavy-atom derivative work, which was to establish the absolute configuration of the molecule in an independent way, the intensities of Bijvoet pair reflections of the U(+) derivative were measured to 5.2 Å resolution. The crystal used for the measurements belonged to the same preparation (test tube) originally used to obtain the isomorphous U derivative data. In the intervening time interval (2 years), additional, differential substitution of uranyl ion occurred. Thus, the time required for equilibrium to be established in a soaking experiment can be considerably long even though the size of the substituting reagent is fairly small.

The problem of approaching equilibrium was 'overcome' by simply allowing sufficient time for the diffusion of the heavy-atom reagent to take place, in typical cases as long as 5–6 weeks. After 2 or 3 weeks of soaking, principal axial intensity distributions were examined for intensity changes and/or constancy of pattern periodically (weekly); only when the distributions appeared to be constant between successive examinations was more serious work undertaken with a derivative. The problem of the concentration dependence was avoided by working with crystals from a single preparation (one test tube). Since only 7–8 crystals were

usually required for three-dimensional intensity data collection to 2.8 Å resolution, this aspect posed no particular difficulties.

The two PtI_4^- derivatives, isolated as a function of anion concentration, have been designated as Pt and Pt(+), corresponding to about (15:1) (~ 1.5 mM) and (25:1) (~ 2.5 mM) molar excess ratios of PtI_4^- to enzyme, respectively. The intensity changes of these derivatives were easily measured and persistent, and they extended to high scattering angles and were similarly observed for the other heavy-atom isomorphous derivatives, the last three of which are: (1) Hg, (10:1) (~ 1 mM) phenylmercuric acetate; (2) the mixed derivative U + Au, (300:1) (~ 30 mM) UO_2^{2+} followed by (10:1) (~ 1 mM) AuI_4^- ; and (3) U(+), same as U but with prolonged soaking. As can be seen from the cell dimensions of the derivatives summarized in Table 1, the magnitudes of the unit-cell changes accompanying derivative formation are generally less than 0.25 Å and 0.15° ; the former is relatively insignificant compared to the size of the dimensions while the latter is almost within the estimated error.

(c) X-ray crystallographic work

(i) Preliminary surveys

All the X-ray crystallographic work (Cu $K\alpha$, $\lambda = 1.5418$ Å, X-ray tube operating at about 1000 W, focal spot = 1.5×15.0 mm, take-off-angle = 3.5° , 1.5 mm incident beam and detector collimators, variable detector aperture) was carried out with a Picker four-circle automatic X-ray diffractometer controlled by a Digital Equipment Corp. (DEC) 4K PDP-8 computer (FACS-I System) coupled to a DEC 32K Disc File. The system was operated with a software package which was developed by us and had, among other options, the ability to monitor the alignment of the crystal during intensity-data collection, redetermine the crystal orientation matrix appropriately when the crystal moved and collect the intensities of reflections in a manner fairly cognizant of our more important needs (Vandlen & Tulinsky, 1971).

Crystals of α -CHT, which grow as diamond plates, were chosen for X-ray examination on the basis of their morphological quality and linear dimensions (to be as similar to each other as possible, $\sim 0.8 \times 0.6 \times 0.2$ mm). The crystals were mounted similarly to the method described by King (1954). Since the monoclinic unique axis (b) of the crystals is fairly obvious from morphological examination (long face diagonal), this was always mounted parallel to the long axis of the glass capillary. Thus, the $0k0$ reflections occurred at $\chi = 90^\circ$ and all crystals displayed similar absorption characteristics with the maximum absorption near the c^* axis.

Throughout our work, we have continually encountered persistent twinning of crystals along the c^* direction noted by Bernal, Fankuchen & Perutz (1938) and later confirmed by Blow, Rossmann & Jeffrey

(1964). Although Sigler *et al.* (1966) report that the inclusion of 2–4% dioxane in the crystallizing solution virtually eliminates twinning, we have found that the effect of dioxane is apparently one of degree rather than of kind (even crystals grown in the absence of dioxane generally show different sizes for the twins comprising twinned crystals). The use of dioxane enhanced the ratio of the larger crystal to the smaller one but twinning persists, although to a lesser extent. Since acceptable twin ratios for untreated crystals comparable to twin ratios obtained with dioxane treatment were obtainable through a process of selection, the use of dioxane was avoided by us. The selection process was simple and was based on the fact that the twin ratio was nearly constant for a given preparation. The $h00$ intensities corresponding to the main and the twin crystal were examined for each crystal preparation. Twin ratios of (6–7):1 or greater were generally considered suitable for further study.

Other considerations pertaining to the overall quality of a crystal were the following. The absorption characteristics were approximated from the measured intensities of selected $0k0$ reflections (at $\chi=90^\circ$) as a function of azimuthal orientation (φ). If the absorption was apparently well-behaved (fairly symmetrical and similar 180° apart in φ , not too severe or different from previous crystals, *etc.*) the unit-cell dimensions and orientation parameters of the crystal were evaluated by the method of least squares using three orientation angles φ , χ , 2θ of 12 observational reflections which were suitably chosen and well distributed in reciprocal space. A crystal was considered satisfactory for intensity data collection if the cell dimensions were within acceptable limits (0.2–0.3% of known average values; less than 0.5% difference between parent and derivative) and if the least-squares estimated standard deviation of the diffraction bisector (ω) of the primary orienting reflection was $< 0.02^\circ$. On the average, only one of two crystals usually fulfilled all the foregoing requirements.

(ii) *Two- and three-dimensional intensity data collection*

Intensities were measured under the control of a computer program based on a step-scan procedure (Wyckoff *et al.*, 1967) which also utilized balanced Ni/Co filters located between the crystal and the detector. In order to allow for small variations from the calculated ω angle of a reflection, which is also the most sensitive angle, the step-scan, carried out in 6-steps of ω (count 6-drop 2), was permitted to wander up to ± 2 steps if the peak intensity did not correspond to the calculated position. This was particularly useful when the crystal had undergone some misalignment. The alignment of the crystal was monitored by periodically measuring the intensities of three reflections which were very sensitive to small angular displacements and which additionally defined the orientation of the crystal. When the crystal misaligned beyond a prescribed tolerance level (10% decrease in count of one of the

monitored reflections), the new orientation was automatically found and used in the ensuing calculations for data collection. The intensities of the monitored reflections were also used to approximate the overall intensity decrease suffered by the crystal with X-ray exposure (Vandlen & Tulinsky, 1971).

The balance of the Ni/Co filter pair was not exact for general background (although excellent for Cu $K\beta$) probably because of the X-ray scattering by the mother liquor around the crystal and by the glass capillary tube. Therefore, the background measurement made with the Co filter at the peak position was corrected for lack of balance. This correction was obtained empirically for each crystal by measuring background with Ni and Co filters separately; the difference in the measurements was taken to be the lack of balance. Except for very small 2θ angles, the correction was small and essentially independent of orientation.

Twinning of crystals about the c^* direction affects the intensities of the $0kl$ reflections and reflections of the type $\pm 2kl$ ($l > 9-10$) in a systematic way. In order to make a first order correction for this effect, the sizes of the crystals comprising the twinned crystal were approximated by measuring the intensity of several $h00$ and twin $h00$ reflections (twin $h00$ is here meant to correspond to the smaller crystal). The twin ratios averaged about 10:1 and ranged from no twin at all to ratios of (5–6):1.

Each three-dimensional data collection was preceded with a separate 6.0 Å resolution $h0l$ intensity data collection. The $h0l$ reflections are centrosymmetric and from work prior to the three-dimensional work, the signs of most of these reflections had been deduced with the use of about 10 heavy-atom isomorphous derivatives (111 signs of 121 observable $|F(h0l)|_p$ reflections). Therefore, the projected 6.0 Å resolution ($h0l$) electron densities were used to obtain relative scaling constants between parent and derivative crystals (avoiding comparisons in regions of substitution) and furthermore, to obtain scaling constants between different crystals of the same kind. In this simple way, the intensity data of all the derivatives and of all the crystals comprising a derivative were placed on a common and reliable scale to within an error of about 5% on structure amplitudes.

Under our instrumental operating conditions, the intensities of the monitored reflections ($2\theta \sim 23^\circ$) usually decreased by about 25 per cent after 35–40 hours of X-ray exposure to the crystal (about the time interval taken for our system to measure the intensities of 3000 reflections). At this time, the crystal was abandoned in favor of a freshly mounted specimen. The decay in intensity is not straightforward but two aspects are certain: (1) the intensities decrease on the average with X-ray exposure; and (2) the decay is 2θ -dependent, being greater in the higher order reflections. The choice of monitored reflections near $2\theta \sim 23^\circ$ was a compromise of several factors; however, the choice proved to be fairly representative of the decay behavior in the

3.0–4.0 Å resolution range where the largest number of the observations were made.

The three-dimensional intensity data were collected in concentric shells of 2θ , with each 2θ shell corresponding to a freshly mounted crystal. Since the average loss of intensity with X-ray exposure was less severe in the low order reflections, the initial 2θ -ranges were chosen to contain a proportionately larger number of reflections. Each new 2θ shell overlapped its preceding shell in 2θ with about 200 reflections. The total number of intensities measured at 2.8 Å resolution was about 11,500, approximately 800–900 of which were 2θ -overlap measurements.

Finally, upon completion of intensity data collection with a given crystal, a separate 6.0 Å resolution set of ($h0l$) intensities was always re-collected and the intensity distributions of the three principal axes were re-examined and compared with the original measurements. The comparison of the latter was performed qualitatively and generally showed a 2θ -dependent decrease in the intensities of the axial reflections, with the heavy-atom derivative crystals often showing additional small changes in relative intensities.

In all, the results to be described in this and some of the forthcoming communications concerning the native enzyme structure represent measurements from about 35 different protein crystals totalling about 75,000 intensity observations, approximately 10,000 of which are redundant as overlap or equivalent reflections.

3. Data reduction

A number of corrections other than the usual instrumental Lorentz and polarization factors were applied to the observed intensity measurements in converting them to structure amplitudes. The lack of balance of the Ni/Co filter pair correcting the Ni filter measurement for some X-ray transmission not shown by the Co filter has already been discussed.

One of the most important systematic effects on the intensities of reflections was the absorption of X-rays by the crystal and its immediate surroundings (mother liquor and glass capillary). Absorption was corrected using a semiempirical method suggested by North, Phillips & Mathews (1968) based upon the variation of the relative transmission (T) with azimuthal angle (φ). The $T(\varphi)$ curves of $0k0$ reflections ($k=2, 4, 6, 18$) were used to obtain the absorption correction of a general reflection in terms of φ , 2θ , and reciprocal lattice level (k index). In general, the various $T(\varphi)$ -curves of a crystal were practically the same and this simplified the absorption correction procedure. The correction was further simplified when the change in relative absorption as a function of scattering angle (or k index) was calculated to be negligible for the size and shape of typical crystals used for intensity data collection [from theoretical absorption calculations with computer program *ORABS* (Wehe, Busing & Levy, 1962)]. The most complicated absorption corrections were

based on at most 2 or 3 independent $T(\varphi)$ curves for a given set of intensity data. The average maximum absorption ratio (I_{\max}/I_{\min}) was about 1.5 for native enzyme crystals and 1.8–2.0 for the heavy-atom isomorphous derivative crystals.

Another systematic effect on some of the intensities resulted from twinning of crystals. Since twinning occurred about the c^* direction, the intensities of the $0kl$ reflections were comprised of the larger crystal and its twin. Consequently, the observed $0kl$ intensities, $I_o(0kl)$, were corrected with the twin ratio, k_t , according to the equation:

$$I(0kl) = \frac{k_t}{(k_t + 1)} \cdot I_o(0kl),$$

where the $I(0kl)$ are now on a scale comparable with the remainder of the data, k_t is $I_x(h00)/I_t(h00)$, and I_x, I_t are intensities of the crystal and its twin, respectively (usually k_t was ~ 10). Another class of reflections systematically affected by the twinning is that of the ($\pm 2kl$) reflections for $l > 9-10$, since reflections $0.10-0.15^\circ$ apart in φ are unresolved under our data collecting instrumental conditions. Any corrections for such a complicated effect, which is dependent on a number of variables (monoclinic β angle, size of receiving aperture, perfection of c^* twinning), would have been difficult to approximate. Consequently, this class of reflections was processed in an ordinary way. The validity of this procedure was further supported by the large k_t 's of the data collection crystals and the frequent occurrence of imperfect c^* twinning (crystal and twin b directions did not coincide and were often as far as 0.5° apart).

The intensities were also corrected for average intensity loss with X-ray exposure (decay).† The monitored reflections usually showed a simple linear decay during the first 30–40 hours of exposure; however, other diffraction experiments indicated that the process of decay was much more involved and the relative intensity pattern also changed somewhat on longer exposures. Indeed, it is remarkable that such effects are not greater obstacles than they actually are in protein structure determination. Since the intensity decay was also 2θ -dependent, the intensities of the first shell of data collection ($2.5^\circ < 2\theta < 20^\circ$) were corrected with two decay curves dependent on 2θ . The mode of decay of the monitored reflections was used to correct the intensity of the reflections in the range $15^\circ < 2\theta < 20^\circ$. The decay determined by comparing the 6.0 Å resolution ($h0l$) electron density projections ($2\theta < 15^\circ$) of 'before' with 'after' data collections was used to correct the intensity of the reflections in the range $2\theta < 15^\circ$ (the former decay rate was generally about twice the latter). The remaining 2θ shells were practically constant in scattering angle and the monitor decay correction sufficed. The maximum corrections ranged up to 1.30–1.35 on intensities and occasionally were based on two different

† For a typical decay curve, see Vandlen & Tulinsky (1971).

slopes when a second faster decay rate set in after 20–25 hours of X-ray exposure.

The reproducibility of some of the intensity measurements has already been discussed elsewhere (Vandlen & Tulinsky, 1971) and since the comparisons were made with $|F|$'s, the reliability of the data correcting procedures was also tacitly implied. The comparisons were carried out by evaluating a quantity R ,

$$R = \frac{\sum ||F|_i - \langle |F| \rangle|}{\sum \langle |F| \rangle},$$

where $|F|_i$ and $\langle |F| \rangle$ are the observed structure amplitudes being compared and their average value, respectively. When measurements were confined to the same crystal, $R \sim 0.02$; when comparisons were extended to different crystals, $R \sim 0.03$; and when reflections were remeasured after considerable X-ray exposure, $R \sim 0.04$ (particularly with the heavy-atom derivatives) (Vandlen & Tulinsky, 1971).[†] The comparisons are extended here to include two independent sets of three-dimensional intensity data of α -CHT to 3.5 Å resolution (two independent preparations, 4 crystals) and to investigate the reproducibility of intensity measurements of equivalent Friedel pair reflections. After reduction to $|F|_p$'s, the 9192 observed reflections of the two sets of α -CHT data at 3.5 Å resolution had an $R = 0.023$, so they were averaged to give 4596 unique reflections which were then used in all subsequent work. The reproducibility of equivalent reflections was assessed by measuring the intensities of 760 reflections of α -CHT to 8.8 Å resolution, within $-10^\circ \leq 2\theta \leq 10^\circ$, and evaluating the quantity,

$$\frac{\sum |I(+)-I(-)|}{\sum I},$$

which was 0.022 (about 0.011 with $|F|_p$'s), where $I(+)$ and $I(-)$ are the measurements at $+2\theta$ and -2θ , respectively, and I is the average value of the intensity. The Wilson plot corresponding to the radial $|F|_p^2$ distribution of α -CHT to 2.8 Å resolution was evaluated, and from the slope of the Wilson line based on the higher order reflections, the apparent average isotropic thermal parameter of α -CHT was determined to be about 27 Å². This value compares well with that of smaller proteins like rhombohedral zinc insulin and ribonuclease-S ($B = 17\text{--}18$ Å²; Harding, Hodgkin, Kennedy, O'Connor & Weitzmann, 1966; Wyckoff *et al.*, 1970) and compares favorably with horse oxyhemoglobin ($B = 35$ Å²; Perutz *et al.*, 1968).

4. Phase calculation—heavy-atom least-squares refinement

(a) General

The first heavy-atom isomorphous derivatives that we prepared were those derived from platinum halide anions (PtCl_4^{2-} and PtI_4^{2-}) and the uranyl ion. The

[†] These comparisons correspond approximately to 0.08, 0.12, and 0.16, respectively, when $\sum |I_i - I_j| / \sum \langle I \rangle$ is considered.

isomorphous substitutions by these heavy-atom reagents were initially ascertained from their respective 6.0 Å resolution centrosymmetric $h0l$ difference Patterson projections, using $|\Delta|^2 = ||F|_{\text{PH}} - |F|_p|^2$ as coefficients where $|F|_{\text{PH}}$ and $|F|_p$ are the observed structure amplitudes of the heavy-atom derivative and α -CHT, respectively. Structure-factor calculations corresponding to the projections, based on the positions and occupancies of the substitutions inferred from the vector maps, determined enough protein phases (signs in agreement for each derivative) to permit the calculation of reliable 6.0 Å $h0l$ difference electron density projections. An iterative two-dimensional refinement procedure was then initiated, which consisted of least-squares refinement and structure-factor and difference electron density calculations. Later calculations included additional minor substitutions and extended to higher resolution. As additional isomorphous heavy-atom derivatives were obtained, similar procedures were pursued. By this process, the number of centric protein phases determined increased progressively to the present total of about 404 signs of 503 observed $h0l$ reflections (80%) at 2.8 Å resolution.

The five most promising derivatives were chosen to determine the protein phases using the multiple isomorphous replacement method (a sixth derivative, U(+), was used later to determine the absolute configuration of the molecule with 5.2 Å resolution anomalous dispersion measurements). Fig. 1 shows the manner in which the native enzyme intensities were affected by derivative formation. The quantity R_{obs} in Fig. 1 is defined as $R_{\text{obs}} = \sum |\Delta| / \sum |F|_p$, and in general, reflects the extent of the heavy-atom substitution. The corresponding values over the complete range are: (1) Pt(+)=0.24, (2) Pt=0.18, (3) Hg=0.18, (4) U=0.13, (5) U+Au=0.17, and (6) U(+)=0.20. Since a comparable factor between the two independent sets of α -CHT data was only about 0.04, the R_{obs} values represent substantial changes, especially when it is

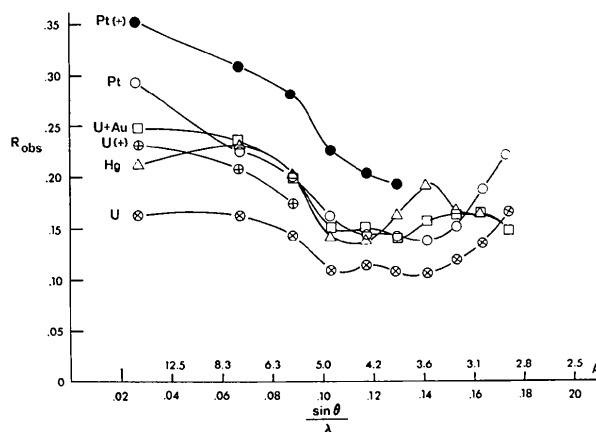


Fig. 1. Effect of various heavy-atom substitutions on α -CHT. R_{obs} as defined in text; ● Pt(+); ○ Pt; △ Hg; ⊙ U; □ U+Au; ⊕ U(+).

noted that they represent a lower limit because phases are not included in the evaluation of the numerator.

(b) *Relative y coordinates of the heavy atoms*

The relative y coordinates of the heavy atoms were obtained from three-dimensional difference Patterson maps (with coefficients $|\Delta|^2$) using data to 4.0 Å resolution for the Pt(+) derivative (~3100 terms) and to 3.5 Å resolution for the Pt (~3650 terms), Hg (~3350 terms), U (~2800 terms), and U + Au (~3750 terms) derivatives (observable differences only). The maps were readily interpretable and the y coordinates relative to an arbitrary origin could be assigned to the principal heavy-atom sites in all the derivatives. One complication appeared when it was discovered there were a number of non-Harker peaks in the Harker section at $v = \frac{1}{2}$ arising from the chance close correspondence of the y coordinates of some major sites (sites 2, 2' and 3, 3' in the Pt derivatives were approximately in the plane of the non-crystallographic twofold rotation axis and showed approximate twofold symmetry, see also Fig. 2 and Table 3). A comparison of peak heights in the sections at $v = 0$ and $v = \frac{1}{2}$ was useful in the identification of these non-Harker peaks. Another minor complication due to the chance coincidence of the (x, z) coordinates of sites 1 and 1' of the Pt(+), Pt, and Hg deriv-

atives was resolved by the examination of the $(0v0)$ and $(2x, v, 2z)$ lines of the Patterson function.

(c) *Relating the y coordinates to a common origin*

The y parameters of the heavy-atoms in the different derivatives were referred to a common origin through the use of two types of correlation vector maps computed at 3.5 Å resolution. The first type of map with coefficients, $||F|_{PH1} - |F|_{PH2}|^2$, was computed using data from the Pt and U derivatives (3796 terms) (Perutz, 1956; Rossmann, 1960). A second type of correlation function involving the product of difference coefficients,

$$(|F|_{PH1} - |F|_P)(|F|_{PH2} - |F|_P),$$

also proved to be useful (Kantha, Bello, Harker & DeJarnette, 1960; Steinrauf, 1963). Two product correlation functions were computed, one with (Pt - U) and the other with (Hg - U) terms (about 2000 each).

Using these two types of correlation syntheses and fixing the main uranium substitution site at $y = 0$, it was possible to refer the y coordinates of the major heavy-atom substitution sites of the Pt(+), Pt, and Hg derivatives to those of the U derivative. Since the U + Au was a 'double derivative' (U positions were common to the U derivative) and U(+) was intimately related to the U derivative, the y coordinates of U + Au and U(+) were correlated from three-dimensional difference electron density maps.

(d) *Phase calculation—least-squares refinement*

The three-dimensional refinement of the heavy-atom parameters was carried out using alternate cycles of phase calculation by the multiple isomorphous replacement method and least-squares refinement of the heavy-atom parameters (Dickerson, Kendrew & Strandberg, 1961; Dickerson, Kopka, Varnum & Weinzierl, 1967; Dickerson, Weinzierl & Palmer, 1968). The phase angles of the reflections of the native enzyme were determined every cycle from the heavy-atom parameters of the previous cycle where each phase angle was calculated from the probability distribution of the protein phase angles (Blow & Crick, 1959). The probability distribution was obtained as suggested by Cullis, Muirhead, Perutz, Rossmann & North, (1961) by evaluating

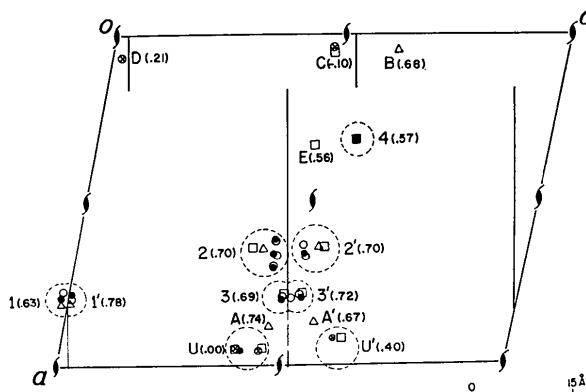


Fig. 2. Regions of heavy-atom substitution. Sites and positions designated as in Table 3; fractional- y coordinates in parentheses; local twofold axis shown at $y = 0.291$ approximately perpendicular to b and c ; ● Pt(+); ○ Pt; △ Hg; ⊙ U; □ U + Au.

Table 2. *Effect of including additional derivatives in the refinement at 6.0 Å resolution*

Number of derivatives	r.m.s. $E(e/u.c.)$					R_A					$N(m)$					Number of reflections	
	$\langle m \rangle$	Pt(+)	Pt	U	Hg U + Au	Pt(+)	Pt	U	Hg U + Au	0.5-0.6	0.6-0.7	0.7-0.8	0.8-0.9	0.9-1.0	($m \geq 0.5$)		
3	0.72	77	71	58	—	0.34	0.44	0.59	—	88	120	173	239	250	870		
4	0.77	75	64	54	76	—	0.35	0.42	0.56	0.58	—	67	107	143	257	936	
5	0.88	60	53	46	56	55	0.29	0.36	0.47	0.43	0.39	23	43	89	188	686	1031

$$R_A = \sum |F_H - f_H| / \sum |\Delta|, \text{ where } |\Delta| = ||F|_{PH} - |F|_P|; F_H = |F|_{PH} \exp(i\alpha_{PH}) - |F|_P \exp(i\alpha_P)$$

$N(m)$ = number of reflections with $m \geq 0.5$ in the various m ranges

Total number of observed reflections at 6.0 Å resolution is 1073.

the variance of the radius vector of the phase circles in the complex plane at 5 degree intervals.*

The refinement of the heavy-atom parameters began with the major substitution sites of the derivatives which had been determined to a high degree of certainty from the difference Patterson maps. The preliminary protein phases were then used to locate lesser heavy-atom substitution sites by calculating three-dimensional difference electron densities for each of the derivatives.

The alternate cycle refinement procedure interceded by the examination of difference electron density maps for purposes of finding additional minor heavy-atom substitution sites was initiated with the Pt(+), Pt, and U derivatives at 6.0 Å resolution. The parameter variation scheme adjusted the occupancies and coordinates of all the atoms of all the derivatives followed by the variation of the correlated y coordinates of the separate derivatives (kept constant in the previous cycle) and a scale constant for each set of derivative structure amplitudes. The thermal parameters of the heavy-atoms were not varied in the 6.0 Å refinement. The progress and rate of convergence of the refinement were followed by examining the parameter shifts and the behavior of quantities sensitive and pertinent to the refinement process as averaged values and as a function of $(\sin \theta/\lambda)$. These quantities included: (1) r.m.s. lack of closure errors (E); (2) r.m.s. calculated heavy-atom structure factors, especially as related to the E 's; (3) the average figure of merit ($\langle m \rangle$); (4) the distribution of the reflections (number) among different m ranges; and (5) various R values.

When the refinement with three heavy-atom derivatives was approaching convergence, another heavy-atom derivative (Hg) was included in the refinement process. Upon completion of the four derivative refinement, the U + Au derivative was included. Table 2 lists the behavior of certain selected quantities through this stage of analysis.

The 6.0 Å resolution refinement with 1073 native enzyme reflections was concluded with: (1) $\langle m \rangle = 0.88$; (2) $\langle \Delta\alpha \rangle = 9.6^\circ$, where $\Delta\alpha = |\alpha_{\max} - \alpha_{\text{best}}|$ and the average is over all the reflections; and (3) a mean residual error based on maximum probability phases, $(\text{MRE})_{\max}$, averaged over all reflections and derivatives of 0.42 (Dickerson, Weinzierl & Palmer, 1968). The expected value of the mean residual error based on 'best' phases (MRE/n , where n is the number of heavy-atom derivatives) is $\frac{1}{2}$ if the r.m.s. errors of the derivatives are properly chosen. Thus, the somewhat smaller value of 0.42 suggests that the errors have been over-estimated and that the correct value of $\langle m \rangle$ is slightly larger.

The refinement of the heavy-atom parameters was next extended to 3.5 Å resolution with 4990 native en-

zyme reflections (87%), all five derivatives included and with refinement procedures similar to those just described. The refinement process was given an additional degree of freedom by varying the isotropic thermal parameter for 1–2 cycles of some of the prime occupancy sites. The temperature factor shifts were moderate and the resulting thermal parameters were reasonable [not excessively large ($B \leq 40\text{--}45 \text{ \AA}^2$) so as to indicate little contribution in this scattering range], and they complied well with the average B -value determined for the native enzyme from the Wilson plot. Some of the major sites of each derivative which clearly showed anisotropic electron density distributions in difference maps were additionally permitted to vary anisotropically for one cycle. Of the 10 atoms varied in this manner, only one resulted in non-positive-definite physically unreal parameters (Hg-1' of the Hg derivative). After about 10 cycles of phase and least-squares calculations, the 3.5 Å resolution refinement was concluded with $\langle m \rangle = 0.81$, $\langle \Delta\alpha \rangle = 12.5^\circ$, and $(\text{MRE})_{\max}/n = 0.40$ ($\langle m \rangle$ decreased even though it increased to 0.91 in the 6.0 Å resolution range).

The refinement of the heavy-atom parameters was concluded by extending the resolution to 2.8 Å [9044 (85%) native enzyme reflections] and introducing the sixth derivative, U(+), which was comprised of 5.2 Å resolution anomalous dispersion data (988 Bijvoet pairs). Although the anomalous dispersion measurements were used in the phase calculations, the heavy-atom structure of U(+) was refined *versus* isomorphous differences only and not with respect to the observed Bijvoet differences. The anomalous dispersion calculations clearly established the absolute configuration of the molecule [892 (90%) of the observed differences agreed with the calculated sign of the difference] and the isomorphous least-squares heavy-atom refinement proceeded smoothly to convergence ($\langle m \rangle = 0.76$, $\langle \Delta\alpha \rangle = 14.3^\circ$, and MRE/n based on best phases was 0.50).

5. Results

The atomic parameters of the heavy-atoms of the six isomorphous derivatives of α -CHT are listed in Table 3, and Fig. 2 shows the general regions of the various substitution sites and the manner in which the heavy-atoms are distributed among these sites. In Table 3, the heavy-atom substitutions which are related by the local twofold rotation axis are grouped in pairs with prime notation denoting the local twofold counterpart. Thus, it can be seen that the Pt and Hg derivatives display relatively good twofold symmetry in the heavy-atom substitutions, especially when the deviations from symmetry are compared to the total substitution; the U + Au and Pt(+) derivatives show considerable deviations from twofold symmetry; and the U derivatives are essentially devoid of the symmetry. The Pt(+) derivative shows a total of 98 electrons without a twofold equivalent over 3 positions (4, U, 2A) and the U + Au derivative shows 262

* The refinement program was a version of a computer program modified by us and originally written by Dr Hilary Muirhead, Department of Biochemistry, University of Bristol, Bristol, England; we would like to thank Dr Muirhead for giving us permission to use the program.

electrons without a counterpart, while the Pt and Hg derivatives show only 55 and 18 electron differences, respectively.

The degree to which the local twofold related heavy-atoms conform to the symmetry can best be seen by considering the occupancies of these heavy atoms and the difference in their position upon a twofold symmetry operation (δ_2), both of which are listed in Table 3. The average difference in occupancy for this group of heavy-atoms is about 3.6 electrons for 12 positions and is within the estimated error of the structure determination [$\sigma(z) \sim \pm 3$ electrons]. The largest difference, which is 16 electrons, occurs with atoms 3-3' of the Pt derivative. The average difference in position of the heavy atoms from exact twofold symmetry is about 0.4 Å [$\sigma(\delta_2) \sim \pm 0.5$ Å]. This estimate is based on the position of the local twofold axis determined by minimizing the difference in the electron density at 2.8 Å resolution between the two molecules in the asymmetric unit with respect to the position of the twofold axis (rotation matrix, translation vector). The equation of the local twofold axis is:

$$z = (0.155 \pm 0.003)x + (0.372 \pm 0.003)$$

at

$$y = (0.291 \pm 0.002),$$

where x, y, z are fractional coordinates and the estimated standard error in the position of the twofold axis is about ± 0.3 Å and agrees well with the one reported by the M.R.C. group (Sigler, Blow, Matthews & Henderson, 1968). The positions of U and U' of the U deriv-

atives and their distance from the twofold axis, about 20.5 and 22.5 Å, respectively, suggest that U and U' are related by the local symmetry element. Since the U substitution site is between two surface contacts of the independent molecules of the asymmetric unit where close local twofold equivalence does not exist, the symmetry suggested by the U and U' positions differs from the twofold symmetry of the other sites. The overall enzyme environments of the U and U' sites are *not related* by a local twofold axis, although they show a good twofold relation with respect to one pair of enzyme molecules. Similar phenomena have been observed by us elsewhere [the binding of *N*-acetyldiiodotyrosine in crystals of α -CHT: one molecule of diiodotyrosine binds per asymmetric unit between two independent α -CHT molecules but the local twofold equivalent position possesses a different enzyme environment (Tulin-sky, Coddling & Vandlen, 1971)].

Table 3 also shows that the occupancies of the heavy atoms are generally substantial (especially in view of the relatively small thermal parameters) and range from nearly 100% [U of U+Au and U(+)] to about 11 electrons or 10–15% occupancy. The lower occupancy levels vary from one derivative to another. The unobservable $|\Delta|$'s are about 8 electrons per enzyme molecule and usually correspond to low-limit heavy atom occupancies of approximately 10 electrons.

The behavior of two quantities as a function of Bragg angle which were used during the course of the heavy-atom refinement to judge the progress and significance of the refinement are shown in Fig. 3; these criteria are the r.m.s. calculated heavy-atom structure

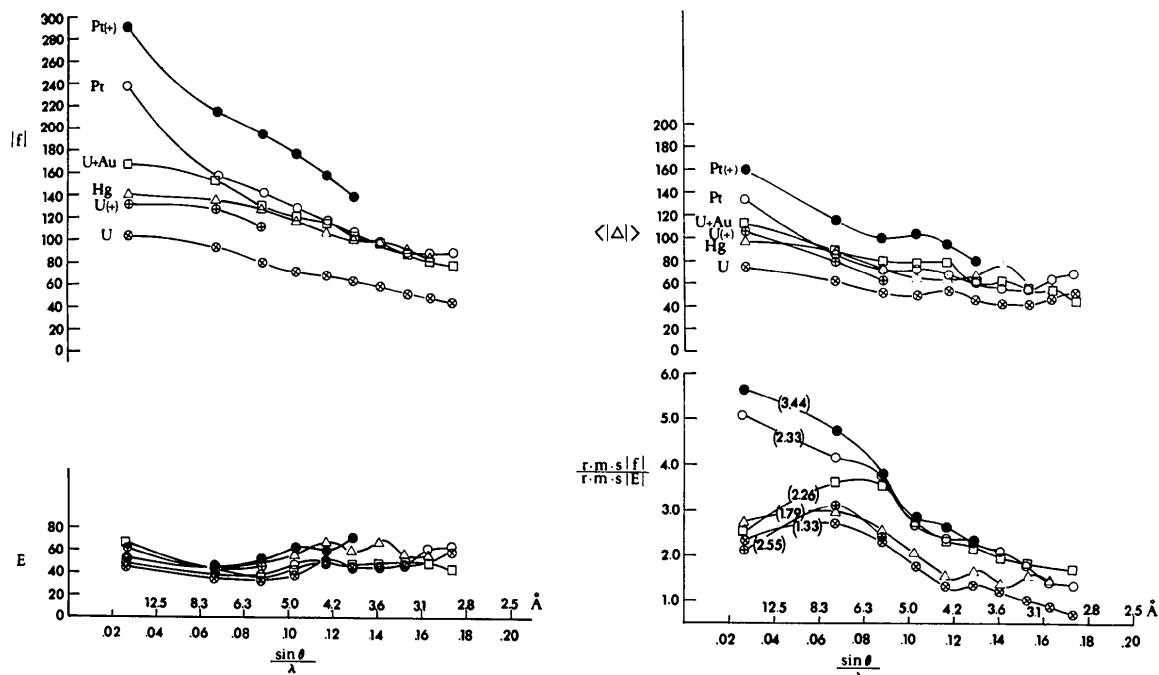


Fig. 3. R.m.s. lack of closure errors, E , and r.m.s. calculated heavy-atom structure factors, $|f|$, (left) and their corresponding ratios (right). R.m.s. $|f|$ and E in electrons/u.c.; ● Pt(+); ○ Pt; △ Hg; ⊗ U; □ U+Au; ⊕ U(+).

factor of the various derivatives, $|f|$, the r.m.s. lack of closure error of each derivative, E , and their corresponding ratio ($|f|/E$). From Fig. 3, it can be seen that the heavy-atom contribution is considerably greater than the lack of closure error over most of the scattering range. For the lesser substituted derivatives (Hg and U), the r.m.s. $|f|$ begins to approach the r.m.s. E at higher resolution. This is reflected by the ($|f|/E$) ratio curves which tend to unity. The behavior of the various r.m.s. E 's over the scattering range is approximately constant. They vary slightly above the unobservable $|d|$ limit (about 8 electrons per enzyme \times 4 molecules = 32 electrons per unit cell). The values in parentheses in Fig. 3 are the ($|f|/E$) ratios averaged over all reflections and generally indicate the phase-determining ability of a particular derivative.

Another quantity which proved to be a useful guide of the progress of the heavy-atom least-squares refinement was an R index, R_d , based on differences in structure factors and structure amplitudes which assessed the agreement between the 'formed' and the calculated

heavy-atom structure factors (R_d as defined in Table 2 and Fig. 4); the behavior of R_d as a function of scattering angle is shown in Fig. 4. It can be seen from Fig. 4 that the R_d values of the derivatives are very acceptable for the low order reflections, especially with the more substituted derivatives [Pt(+), Pt, U + Au; $R=0.25-0.35$]. The values in parentheses in Fig. 4 are the R_d values summed over all the reflections and, similar to the ($|f|/E$) ratios, indicate to some extent the phase-determining ability of a given derivative.

Also shown in Fig. 4 is another type of agreement index, R_K , originally introduced by Kraut, Sieker, High & Freer (1962), where

$$R_K = \sum ||F|_{PH} - |F|_P| \exp(i\alpha_P) + |f_H| / \sum |F|_{PH}.$$

Although this quantity was followed throughout the heavy-atom refinement as a function of Bragg angle and evaluated over the complete range, it did not prove to be useful primarily because of its relative insensitivity to smaller changes in the calculated heavy-atom structure. Furthermore, R_K tends to a small value as the extent of the heavy-atom substitution decreases. Clearly, this is an undesirable functional dependence and can cause comparisons among the different derivatives to be misleading (e.g. $R_K=0.083$ for Pt and 0.089 for U at 2.8 Å resolution and the behavior persists over the complete range, Fig. 4). The use of R_K values in conjunction with R_{obs} values (R_K/R_{obs} , see Fig. 1 for the definition of R_{obs}) alleviates the problem but essentially reduces the R_K index to a scaled version of the more useful R_d index based on differences, since

$$R_K/R_{obs} = (\sum |F|_P / \sum |F|_{PH}) R_d.$$

The average figure of merit, $\langle m \rangle$, as a function of resolution is shown in Fig. 5 with an analysis of the distribution of the figure of merit among the reflections as a function of its magnitude. The behavior of $\langle m \rangle$ compares favorably with that of smaller proteins at a similar resolution [ribonuclease-S (Wyckoff *et al.*, 1970) and lysozyme (North & Phillips, 1969)] and far superior to that recently reported by the M.R.C. group for tosyl- α -CHT (Blow & Matthews, 1973). In the distribution of the reflections among the figure of merit ranges, about $\frac{2}{3}$ of the data have a figure of merit greater than the average and about 85% (~ 7700) of the reflections have an $m \geq 0.5$.

The centrosymmetric $h0l$ difference electron density projections of five of the isomorphous derivatives are shown in Fig. 6, along with the difference in electron density between the Pt(+) and Pt derivatives [Fig. 6(c)]. The projected densities were evaluated with the protein phases determined from the final three-dimensional heavy-atom refinement and phase calculation cycle at 3.5 Å resolution; cross-over $h0l$ reflections were appropriately included with sign changes and only $|d| > 33$ electrons/u.c. was used in the calculations. The contour level of the maps approximately corresponds to an occupancy of 10 electrons per contour. Thus, the first contour is near background level and the number of

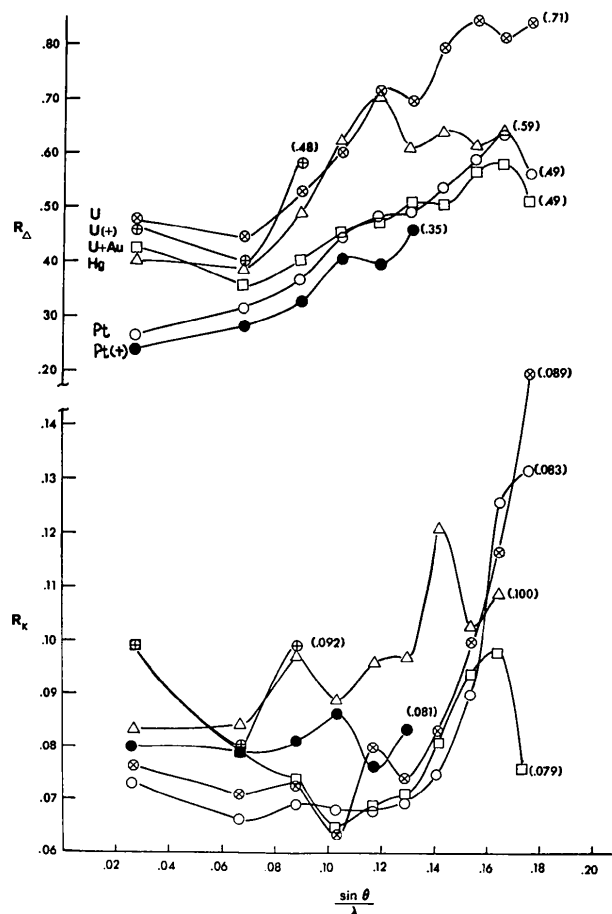


Fig. 4. Final agreement factors between 'observed' and calculated heavy-atom structure factors. $R_d = \sum |F_H - f_H| / \sum |d|$; R_K as defined in text; ● Pt(+); ○ Pt; △ Hg; ⊗ U; □ U + Au; ⊕ U(+).

contours in a peak closely approximates the occupancy of the position.

From Fig. 6, it can be seen the major substitution positions are very prominent, the peak heights are approximately proportional to their occupancy and the background fluctuates at about the ± 10 e level. For convenience, the position of the local twofold axis in projection is shown, from which it can be seen that only the Hg derivative displays a high degree of twofold symmetry. It is interesting to note that the M.R.C. Hg derivative only shows substitution at 1-1' because the parent crystals are tosyl- α -CHT; since tosyl binds near positions 2-2', the additional mercury at 2-2' as seen in Fig. 6(d) cannot be accommodated. The difference density between the Pt(+) and Pt derivatives [Fig. 6(c)] shows that the additional substitution of PtI_2^- anion is considerable (~ 200 e) and proceeds in a differential and non-twofold way, giving rise to an independent isomorphous derivative. Furthermore, from Figs. 6(a)–(c), it can be seen that the electron density distributions

in the vicinity of sites 2-2' and 3-3' are excessively elongated; moreover, they cannot be accounted for with reasonable three-dimensional anisotropic thermal parameters. In the case of sites 2-2', the elongation arises from subsidiary positions with similar (x,z) coordinates as positions Pt-2 and Pt-2'; with 3-3', the elongation comes from unresolved major sites approximately 3.0 Å apart. The elongation of the electron density in sites 2-2' might be due to the presence of iodine atoms still associated with the platinum atoms. The anisotropy was taken into account by including iodine atoms in structure factor calculations (designated as I in Table 3). Finally, from Fig. 6(j), it can be seen the U derivative shows a particularly simple substitution which is also devoid of 'true' local twofold symmetry.

Some additional factors affecting the overall assessment of the quality and significance of the observational data and the heavy-atom refinement are summarized in Table 4, where $\sum Z_H$ corresponds to the total number of electrons substituted in each derivative

Table 4. Summary of least-squares refinement

Derivative	$\sum Z_H$ (e/u.c.)	$(\sum Z_H^2)$ $\times 10^{-3}$ (e ² /u.c.)	r.m.s. $ f _H$ (e/u.c.)	r.m.s. E (e/u.c.)	$(N \Delta _{\bar{T}})$	$R_A(h0l)$	Δs	d_{\min} (Å)
Pt(+)	1089	54.1	189	55	245	0.287	15	4.0
Pt	789	30.8	114	49	403	0.373	9	2.8
Hg	506	22.3	104	58	361	0.438	9	3.0
U	271	10.7	61	46	369	0.418	6	2.8
U(+)	452	20.5	120	47	126	0.339	7	5.2
U + Au	673	29.8	104	46	385	0.405	10	2.8

$\sum Z_p = 54,000$ (e/u.c.).
 $\sum Z_p^2 = 328 \times 10^3$ (e²/u.c.).
 $R_A(h0l) = \sum ||\Delta| - f_H| / \sum |\Delta|$, where sign changes are included.
 Δs = number of sign changes for each derivative.

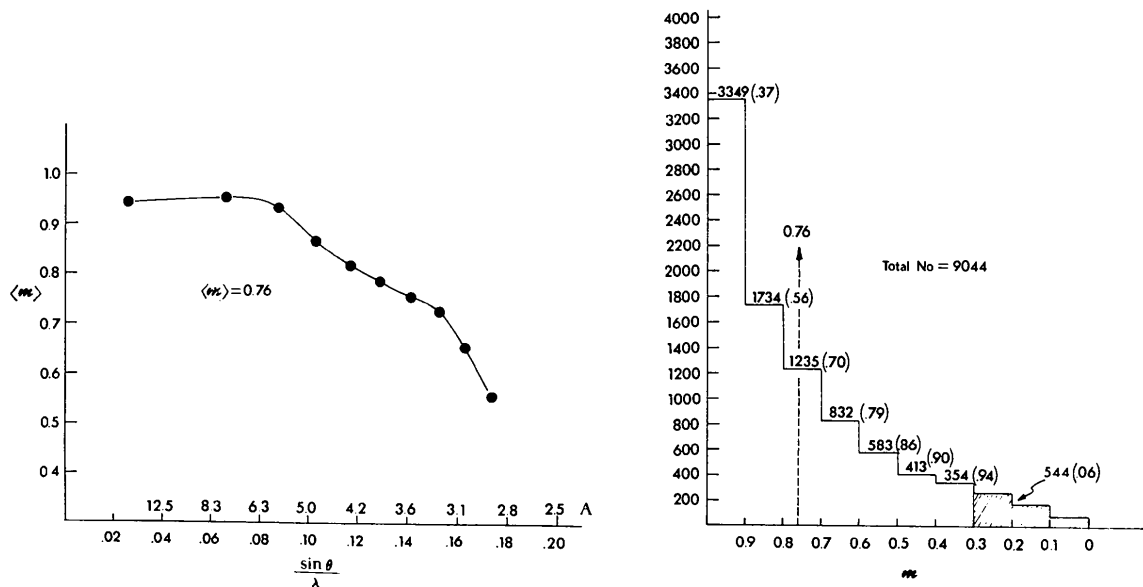


Fig. 5. Measure of preciseness of the determination of protein phase angles - figure of merit (m). Fraction of total number of reflections with m greater than given m shown at top of columns of the histogram; first column includes most of centric $h0l$ reflections (about 400).

and is included for comparison to the number of electrons, $\sum Z_P$, of α -CHT. The next column, in conjunction with the value of $\sum Z_P^2$, gives the average expected change in the radial intensity distribution of α -CHT on heavy-atom derivative formation. It can be seen the average expected change of the Pt(+) derivative is fairly profound near 16% (observed to be about 10%), whereas the effect is much less with the U derivative ($\sim 3\%$). The last columns of Table 4 relate to the centrosymmetrical $h0l$ reflections: $N(|\Delta|_{\bar{1}})$ is the number of $h0l$ differences with $|\Delta| > 33$ electrons/u.c.; $R_{\Delta}(h0l)$ is the R index between observed and calculated centric differences; and Δs is the number of cross-over reflections which change sign on derivative formation and were included as such in the calculation of $R_{\Delta}(h0l)$ and in the evaluation of the $h0l$ difference electron densities.

Table 5 conveniently summarizes the results of the anomalous dispersion measurements and calculations. From Table 5, it can be seen that the observed anomalous differences range close to the observed isomorphous difference low-limit between derivative and native enzyme. However, the differences are considerably

larger than the error of their calculation (r.m.s. ϵ). The observed $\langle \kappa \rangle$ ratio is in excellent agreement with the theoretical ratio of 0.18 for uranium, and it persists

Table 5. Summary of anomalous dispersion measurements and calculations with U(+) at 5.2 Å resolution

$\langle \sin \theta / \lambda \rangle$	0.027	0.067	0.076	Average
Number of reflections	178	390	420	
r.m.s. $ \Delta \pm _o$	43.0	33.8	32.1	34.9
r.m.s. $ \Delta \pm _c$	36.4	33.8	30.3	32.8
r.m.s. ϵ	19.2	16.1	20.1	18.4
$\langle \kappa \rangle$	0.19	0.16	0.15	0.17
$R(\Delta \pm)$	0.37	0.40	0.46	0.42
$R(\Delta \pm)$	0.41	0.43	0.54	0.47

$$|\Delta \pm|_o = ||F(+)|_o - |F(-)|_o|.$$

$$|\Delta \pm|_c = \text{corresponding calculated Bijvoet difference.}$$

$$\text{Error} = \epsilon = ||\Delta \pm|_o - |\Delta \pm|_c|.$$

$$\kappa = |\Delta \pm|_o D / D^2, \text{ where } D = 2F_P f_U \sin(\alpha_P - \alpha_U) / F_{PU}.$$

$$R(|\Delta \pm|) = \sum ||\Delta \pm|_o - |\Delta \pm|_c| / \sum |\Delta \pm|_o.$$

$$R(\Delta \pm) = \sum |\Delta \pm|_o - \Delta \pm|_c / \sum |\Delta \pm|_o, \text{ where } \Delta \pm|_o = |F(+)|_o - |F(-)|_o, \text{ and } \Delta \pm|_c \text{ is the corresponding calculated value.}$$

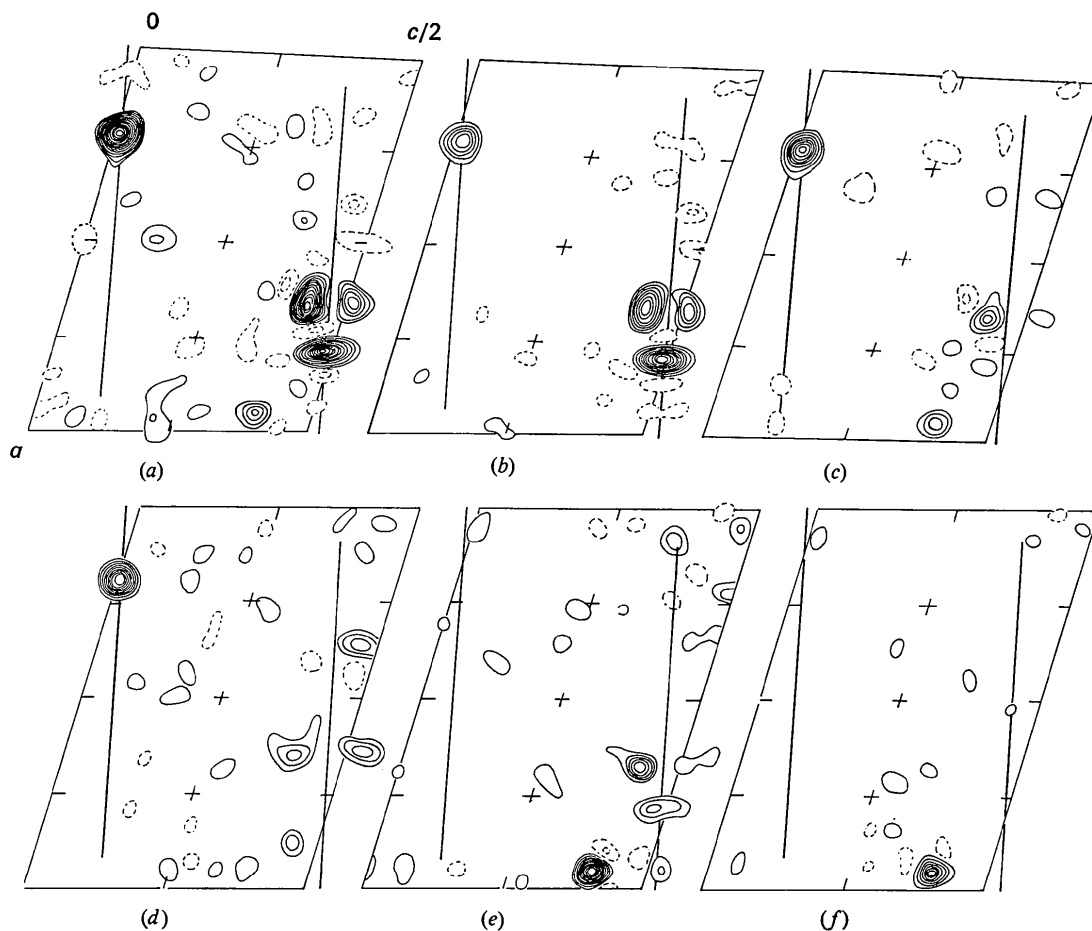


Fig. 6. Centric $h0l$ difference electron density projections of the heavy-atom derivatives. (a) Pt(+); (b) Pt; (c) difference between Pt(+) and Pt; (d) Hg; (e) U+Au; (f) U; local twofold axis shown approximately perpendicular to c; contour intervals at about 2.6 e \AA^{-2} ($\sim 10 \text{ e}$ in occupancy), beginning with the 2.6 e \AA^{-2} contour; negative contours broken.

throughout the various scattering ranges. In fact, the same applies to the behavior of all the quantities listed in Table 5. The agreement between the observed and calculated Bijvoet differences is summarized by the R values in Table 5, which show that there are only a small number of disagreements in the sign of the Bijvoet difference (96 disagreements in 988 anomalous dispersion observations).

Prior to the inclusion of the anomalous scattering measurements in the phase calculations, a difference Patterson function using $(|F_{PH}(+)| - |F_{PH}(-)|)^2$ as coefficients was computed. The $U(+)$ vector map contained all the 1596 observed Bijvoet differences at 5.2 Å resolution, and it showed all the peaks at the ends of the vectors relating the principal scatterers, U and U' . However, only some of the interactions involving U or U' with the minor substitution sites appeared in the map.

A final examination of the reliability of the anomalous dispersion data and calculations was made by computing the imaginary electron density map of the $U(+)$ derivative. The imaginary map was computed by first setting a lower limit on observable Bijvoet differences (equivalent to a 5-electron substitution/molecule) and also by placing an upper bound on the Fourier coefficients by removing terms with $\alpha_p \approx \alpha_H$. After modifying the coefficients in this way, 946 (60%) Bijvoet differences remained of which 872 (92%) calculated to be positive. The Fourier map had positive peaks at all the heavy-atom positions of the $U(+)$ derivative approximately proportional in the height to their occupancies and thereby independently established the absolute configuration of the molecule.

We would like to take the opportunity here to gratefully acknowledge the support of this work by the National Science Foundation, Molecular Biology Section (Grants GB-6586, GB-7399, and GB-15402). In addition, we would like to thank Mrs Nancy T. Tucker for the countless ways in which she aided the work and Dr A. C. MacDonald and Mr Alberto Protzel, who carried out some of the initial analyses on the $U+Au$ and Hg derivatives, respectively. Lastly, one of us (A.T.) would like to thank the many other persons too numerous to mention who have contributed to this work from its inception in various ways and to varying degrees.

References

- BERNAL, J. D., FANKUCHEN, I. & PERUTZ, M. F. (1938). *Nature, Lond.* **141**, 523-524.
- BIRKTOFT, J. J. & BLOW, D. M. (1972). *J. Mol. Biol.* **68**, 187-240.
- BLOW, D. M. & CRICK, F. H. C. (1959). *Acta Cryst.* **12**, 794-802.
- BLOW, D. M. & MATTHEWS, B. W. (1973). *Acta Cryst.* **A29**, 56-62.
- BLOW, D. M., ROSSMANN, M. G. & JEFFREY, B. A. (1964). *J. Mol. Biol.* **8**, 65-78.
- CULLIS, A. F., MUIRHEAD, H., PERUTZ, M. F., ROSSMANN, M. G. & NORTH, A. C. T. (1961). *Proc. Roy. Soc. A* **265**, 15-38.
- DICKERSON, R. E., KENDREW, J. C. & STRANDBERG, B. E. (1961). *Acta Cryst.* **14**, 1188-1195.
- DICKERSON, R. E., KOPKA, M. L., VARNUM, J. C. & WEINZIERL, J. E. (1967). *Acta Cryst.* **23**, 511-522.
- DICKERSON, R. E., WEINZIERL, J. E. & PALMER, R. A. (1968). *Acta Cryst.* **B24**, 997-1003.
- HARDING, M. H., HODGKIN, D. C., KENNEDY, A. F., O'CONNOR, A. & WEITZMANN, P. D. J. (1966). *J. Mol. Biol.* **16**, 212-226.
- KARTHA, G., BELLO, J., HARKER, D. & DEJARNETTE, F. E. (1963). *Aspects of Protein Structure*. Edited by G. N. RAMACHANDRAN, pp. 13-22. London: Academic Press.
- KING, M. V. (1954). *Acta Cryst.* **7**, 601-602.
- KRAUT, J., SIEKER, L. C., HIGH, D. F. & FREER, S. T. (1962). *Proc. Natl. Acad. Sci. U.S.A.* **48**, 1417-1424.
- MATTHEWS, B. W., SIGLER, P. B., HENDERSON, R. & BLOW, D. M. (1967). *Nature, Lond.* **214**, 652-656.
- NORTH, A. C. T. & PHILLIPS, D. C. (1969). *Progress in Biophysics and Molecular Biology*, edited by J. A. V. BUTLER and D. NOBLE. Vol. 19, part 1, pp. 5-132. Oxford: Pergamon Press.
- NORTH, A. C. T., PHILLIPS, D. C. & MATHEWS, F. S. (1968). *Acta Cryst.* **A24**, 351-359.
- PERUTZ, M. F. (1956). *Acta Cryst.* **9**, 867-873.
- PERUTZ, M. F., MUIRHEAD, M., COX, J. M., GOAMAN, L. C. G., MATHEWS, F. S., MCGANDY, E. L. & WEBB, L. E. (1968). *Nature, Lond.* **219**, 29-32.
- ROSSMANN, M. G. (1960). *Acta Cryst.* **13**, 221-226.
- SIGLER, P. B. & BLOW, D. M. (1965). *J. Mol. Biol.* **14**, 640-644.
- SIGLER, P. B., BLOW, D. M., MATTHEWS, B. W. & HENDERSON, R. (1968). *J. Mol. Biol.* **35**, 143-164.
- SIGLER, P. B., JEFFREY, B. A., MATTHEWS, B. W. & BLOW, D. M. (1966). *J. Mol. Biol.* **15**, 175-192.
- STEINRAUF, L. K. (1963). *Acta Cryst.* **16**, 317-318.
- TULINSKY, A., CODDING, P. W. & VANDLEN, R. L. (1971). Abstracts, Winter Meeting of the American Crystallographic Association, Columbia, South Carolina, February 1971, No. K3.
- TULINSKY, A., VANDLEN, R. L., MORIMOTO, C. N., MANI, N. V. & WRIGHT, L. H. (1973). *Biochemistry*. In the press.
- VANDLEN, R. L. & TULINSKY, A. (1971). *Acta Cryst.* **B27**, 437-442.
- VANDLEN, R. L. & TULINSKY, A. (1973). *Biochemistry*. In the press.
- WEHE, D. J., BUSING, W. R. & LEVY, H. A. (1962). *ORABS. A Fortran Program for Calculating Single Crystal Absorption Corrections*. Report ORNL-TM-229, Oak Ridge National Laboratory, Oak Ridge, Tennessee.
- WYCKOFF, H. W., DOSCHER, M., TSENOGLOU, D., INAGAMI, T., JOHNSON, L. N., HARDMAN, K. D., ALLEWELL, N. M., KELLY, D. M. & RICHARDS, F. M. (1967). *J. Mol. Biol.* **27**, 563-578.
- WYCKOFF, H. W., TSENOGLOU, D., HANSON, A. W., KNOX, J. R., LEE, B. & RICHARDS, F. M. (1970). *J. Biol. Chem.* **245**, 305-328.