THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 246, No. 7, Issue of April 10, pp. 2302-2316, 1971
*Printed in U.S.A.*

# A High Resolution Structure of an Inhibitor Complex of the Extracellular Nuclease of *Staphylococcus aureus*

## I. EXPERIMENTAL PROCEDURES AND CHAIN TRACING*

(Received for publication, November 27, 1970)

A. ARNONE,‡ C. J. BIER, F. A. COTTON,§ V. W. DAY,¶ E. E. HAZEN, JR., D. C. RICHARDSON,‖ J. S. RICHARDSON,‖ AND, IN PART, A. YONATH**

*From the Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*

### SUMMARY

**Two isomorphously substituted derivatives of the nuclease-$Ca^{2+}$-thymidine 3′,5′-diphosphate complex have been prepared and used in an x-ray crystallographic study of the molecular structure. In one case, 5-iododeoxyuridine 3′,5′-diphosphate was used in place of thymidine 3′,5′-diphosphate (a net replacement of $CH_3$ by I), and, in the other, $Ba^{2+}$ was used in place of $Ca^{2+}$. Intensities of all reflections and their Friedel pairs within the 4 A sphere were measured on this inhibitor complex and its two substituted derivatives; in addition, approximately 35% of the data between 4 A and 2 A were collected, selection being made on the basis of the peak to background ratios. These data have been used to produce an electron density map of sufficient quality to allow a complete tracing of the peptide chain except for a few residues at each terminus which presumably project into solution and are too disordered to be distinguishable from solvent.**

**Approximately 30 residues are involved in three separated sections of helix, and about 24 residues form a three-stranded section of antiparallel β-pleated sheet. Residues 44 to 53 form a loop which is loose, highly exposed to solvent, and somewhat disordered. Residues Lys-48 and Lys-49, which are selectively vulnerable to trypsin-catalyzed hydrolysis, lie at the extremum of this loop.**

**The most significant feature of the nuclease structure is a large pocket which serves as the inhibitor binding site. With the electron density maps now available, the lining of this pocket is revealed to be predominantly neutral or hydrophobic with the exception of several residues which specifically participate in binding the calcium ion and the nucleoside diphosphate. Of the latter, the most conspicuous are Lys-84 and Tyr-85 which form hydrogen bonds to the 3′-phosphate, the guanidinium moieties of Arg-35 and Arg-87 which form hydrogen bonds to the 5′-phosphate, and the carboxylate ions of Glu-43, Asp-21, and Asp-40 which serve as ligands to the calcium ion. Although the calcium ion is directly below the 5′-phosphate, it is not close enough for direct interaction with it. It appears that the barium ion occupies a position significantly different from that of the calcium ion.**

Work in this laboratory has previously led to a report of the structure of the crystalline uninhibited extracellular nuclease from *Staphylococcus aureus* and the inhibitor complex formed from it with calcium ion, and thymidine 3′,5′-diphosphate as obtained from x-ray crystallographic data collected within the 4 A sphere in reciprocal space (1). These results gave a very distinct, recognizable image of the inhibitor molecule as it is bound in a pocket of the enzyme and revealed the presence of several runs of helix and a section of antiparallel pleated sheet as well as the location of most of the remainder of the peptide chain.

We have now obtained an electron density map of the enzyme-$Ca^{2+}$-inhibitor complex based on reflections to $d$ spacings as low as 2.0 A which, in conjunction with the known primary sequence (2), permits us to identify the residues along nearly the whole of the peptide chain, interpret in some detail the interactions which stabilize the tertiary structure, and identify the principal interactions involved in inhibitor binding.

The high resolution electron density map reported in this paper was obtained with the use of the intensity differences and the anomalous scattering differences from only two isomorphous derivatives containing relatively light heavy atoms. One of these resulted from replacement of the 5-methyl group of the pdTp[1] inhibitor with an iodine atom, a net difference ($I - C$) of 47 electrons, whereas the other resulted from replacement of the

* This work was supported by Grant GM13300 from the Institute of General Medical Sciences of the National Institutes of Health.

‡ National Institutes of Health Predoctoral Fellow, 1967 to 1970. Present address, Medical Research Council Laboratory of Molecular Biology, Cambridge, England.

§ To whom reprint requests should be directed.

¶ National Institutes of Health Postdoctoral Fellow 1969–.

‖ Present address, Department of Biochemistry, Duke University, Durham, North Carolina.

** On leave from the Weizmann Institute, Chemistry Department, Rehovot, Israel, 1969 to 1970.

[1] The abbreviations used in this paper are: pdTp, thymidine 3′,5′-diphosphate; pdIUp, 5-iododeoxyuridine 3′,5′-diphosphate; RMS, root mean square.

calcium ion of the nuclease-pdTp-$Ca^{2+}$ complex by $Ba^{2+}$, a net difference ($Ba^{2+}$ − $Ca^{2+}$) of 36 electrons. In addition, it is important to note that the data set used to obtain this structure is composed of essentially all the reflections within the 4 A sphere, but only that third of the reflections having the highest intensities in the range from 4 A to 2 A. The crystallographic space group is $P4_1$ with one enzyme molecule per asymmetric unit (1, 3). The nominal unit cell dimensions are taken as a = 48.2 A, and c = 63.3 A.

In the first part of this paper, we report on and discuss the protein crystallographic procedures used to obtain the electron density map, and, in the second part, we report on the over-all conformation of the peptide chain and the qualitative mode of binding of the pdTp and the calcium ion. Future reports will discuss the detailed structure of the enzyme and attempt a quantitative portrait of the inhibitor binding and its mechanistic implications.

## CRYSTALLOGRAPHIC PROCEDURES

### *Crystal Growing and Mounting*

The methods of growing and mounting crystals have already been rather fully described (1, 3). To these descriptions we wish to add a few comments. Although it is possible to grow barium ion-substituted crystals *de novo* from solutions containing nuclease, pdTp, and $BaCl_2$, crystals so obtained were relatively small and excessively elongated. In practice, the barium ion-substituted crystals were prepared by soaking the larger, better shaped nuclease-pdTp-$Ca^{2+}$ crystals in a solution $10^{-4}$ M in $BaCl_2$, $5 \times 10^{-5}$ M in pdTp, 0.032 M in Tris-Cl buffer (pH, 8.2), with a solvent containing 40% by weight of 2-methyl-2,4-pentanediol in water. In addition, prior treatment of the glass mounting capillaries with either the standard phosphate buffer (1) or the Tris buffer has been found to markedly improve the stability of the mounted crystals.

### *Collection and Reduction of Data*

Intensities were measured in a cold room at +2° with a precisely aligned, Datex controlled, General Electric XRD-6 diffractometer with the stationary crystal, stationary counter technique, and copper K$\alpha$-radiation. The quality of the focal spots of the x-ray tubes and the width of the take-off angle (usually 6°) were sufficient to permit reasonably accurate estimates of the integrated intensities of the reflections. Peak counts and background counts (with balanced filters), where taken, were measured at only a single setting of the goniometer angles for periods of 10 or 20 seconds; the time depending mainly on the size of the crystal. Background due to scattering of the characteristic radiation by the various elements in the x-ray path was reduced by varying the uniform area of the incident beam and the size of the diffracted beam aperture to match the maximum dimension of each individual crystal with an allowance for inaccuracy in angle setting, mosaic spread of the crystal, and $\alpha_1$-$\alpha_2$ broadening.

Within the 4 A sphere, two different background corrections were applied. The first was the conventional correction for the contribution to the reflections from the noncharacteristic radiation, minimized in this case by the use of balanced filters. The balanced filter technique does not lead to cancellation of that part of the background due to scattering of the characteristic radiation by elements in the x-ray path other than the specific reciprocal lattice under examination. Because of the nature of the crystal mounting, as well as the nature of the crystals themselves, such scattering of the copper K$\alpha$-radiation, principally from the glass capillary, makes a significant contribution to the background in protein crystallography. We have estimated the magnitude of this second background correction from radial scans programmed to sample reciprocal space away from the lattice points. In practice, a number of such radial scans were run, sampling different regions of reciprocal space, by incrementing the $2\theta$ axis in 1°-steps at selected settings of the $\phi$ and $\chi$ axes, the characteristic radiation being again isolated with balanced filters. Smooth curves were fitted to these data, principally to eliminate bumps due to accidental encounters with edges of reflections. A correction for the $\phi$ dependence was also made[2] since these scans showed the same $\phi$ dependence as the absorption correction (see below). The results of the various scans were averaged since there appeared to be no systematic dependence on $\chi$. For some crystals, however, the curves were nearly but not quite identical for $+2\theta$ and $-2\theta$; the second background correction was, therefore, different for the members of a Friedel pair in such cases. Tests showed that beyond the 4 A sphere the rather elaborate procedure just described produced the same relative integrated intensities as a much simpler one with a background correction estimated only from radial scans with a nickel filter being used for both the measurement of the reflection intensity and the background. Thus, all data in the 4 A to 2 A range were collected by this simpler procedure, saving considerable time and allowing many more reflections to be measured on a given crystal before it was discarded on account of radiation degradation.

An empirical absorption correction was estimated by measuring several well separated $00l$ reflections as a function of $\phi$ at $\chi$ = 90°. The $\phi$ scan profile did not vary significantly with $l$. The absorption correction was then applied following the suggestion of North, Phillips, and Mathews (4), and in no case did this $\phi$ correction exceed 20%; for most crystals it was less than 10%.

The full data set consists of some 4200 of the 9600 independent reflections within the 2 A sphere. Of these 4200, essentially the full set of 1400 within the 4 A sphere was measured. In the range of 4 to 2 A, only some 2800 independent reflections which constitute that third of the data with the highest intensities, as judged by the peak to background ratios of a full set of data from a native crystal, were measured. For the heavy atom derivatives, and on occasion for the native nuclease-pdTp-$Ca^{2+}$ crystals, the Friedel pairs were measured at $\pm 2\theta$ in alternating batches of 10 reflections. The data set for each type of crystal, *i.e.* native, iodine, or barium, contains at least two measurements, made on different crystals, of the intensity of each reflection and its Friedel pair. In most cases at least three measurements are included in the final estimate of intensity.

Crystal degradation as a function of the time of exposure and crystal alignment were followed by monitoring the intensities of a set of 20 standard reflections inserted after every 100 pairs of reflections. Crystals were discarded after the intensities of the test reflections had dropped approximately 5%.

The recorded intensities were corrected for degradation, absorption, backgrounds, and the standard Lorentz-polarization factor. The particular set of data was then level-scaled to an

[2] As a consequence, the absorption correction was applied before backgrounds were subtracted.

arbitrary standard scale by adjusting the value of $\sum F^2$. For the 4 A data, the $\sum F^2$ of the full 6 A sphere was used; for the 4 to 2 A data the $\sum F^2$ of the 20 standard reflections was used. The multiple measurements of an individual reflection were then averaged, and, as a first approximation, $\sum F_P^2$ was set equal to the $\sum F_{PH}^2$ in eight concentric spherical shells in reciprocal space.[3] This procedure (shell scaling) is intended to compensate for differences in the extent of radial fall-off in average intensity between the native and heavy atom derivative crystals (5). Usually, the average intensities of the heavy atom derivatives are found to decline more rapidly with increasing $2\theta$ than those of the native crystal. The probable causes of this are the additional absorption by the heavy atom, lack of perfect isomorphism, and increased disorder. Such was the case for both the iodine and barium derivatives, but shell-scaling factors for the outermost shells in reciprocal space differed from unity by only 5 to 6%. It is our experience that the averaging of multiple measurements considerably improves the quality of a data set. In particular, such a procedure can markedly increase the validity of anomalous scattering measurements. This statement is based principally on the appearance of three-dimensional heavy atom difference Patterson maps and "cross Fourier" maps, with phases based on a single heavy atom derivative, calculated as described below.

A final scaling adjustment was made to compensate for the additional scattering power of the heavy atoms following the concept suggested by Kraut (6, 7). This method has the advantage of not depending on Wilson statistics and not requiring a knowledge of the heavy atom or atoms occupancy. The basis of the procedure follows from the observation that the origin peaks P(0,0,0) of two different Patterson maps, one constructed from coefficients $(\vec{F}_{PH} - \vec{F}_P)^2 = (F_H)^2$ and the other from coefficients $(F_{PH}^2 - F_P^2)$, should be equal if $F_{PH}$ is scaled properly with respect to $F_P$. The first Patterson map will be composed of vectors between the heavy atoms, and, therefore, the origin peak should consist solely of the heavy atom self-vectors. The second Patterson map represents the difference of two Patterson maps, one with coefficients $F_{PH}^2$ and one with coefficients $F_P^2$, both of which contain the vectors between protein light atoms. Therefore, at the origin of Patterson space the protein self-vectors will cancel, again leaving just the heavy atom self-vectors. Thus, the scale factor between $F_P$ and $F_{PH}$ should be adjusted so that Equation 1 is satisfied

$$\sum F_H^2 = \sum (F_{PH}^2 - F_P^2) \tag{1}$$

where the summation is over-all reflections in a particular range in $2\theta$. If there are $n$ reflections in a given $2\theta$ range, then dividing (1) by $n$ gives

$$<F_H^2> = <F_{PH}^2> - <F_P^2>$$

or

$$<I_{PH}>/<I_P> = 1 + <I_H>/<I_P> \tag{2}$$

where $I$ has been substituted for $F^2$. This Kraut scaling is applied with an iterative procedure so that Equation 1 or Equation 2 is satisfied. For example, with the use of Matthews' formula (8), Equation 3, to calculate $I_H$,

$$F_H^2 = F_P^2 + F_{PH}^2 - 2F_PF_{PH}\{1 - [wk^{-1}(F_{PH+} - F_{PH-})/2F_P]^2\}^{1/2} \tag{3}$$

[3] The notation used in this paper is defined in Table I.

the following scale factor was constructed.

$$S = 1 + X^{1/2} - Y^{1/2} \tag{4}$$

where $X$ and $Y$ denote the left- and right-hand sides, respectively, of Equation 2. The scaling of $F_{PH}$ could then be adjusted by dividing by $S$. After three iterative cycles of computing $S$ and rescaling $F_{PH}$, $S$ had converged to 1.00 for each $2\theta$ range, indicating that Equation 2 was now satisfied. The final $<I_{PH}>/<I_P>$ ratios for both the iodine and barium derivatives are given in Table II. It should be noted that the weaker, error-prone measurements ($F_P$ or $F_{PH}$ <140 on an absolute scale) were not included in calculating the scale factor $S$. However, $S$ was applied to all the $F_{PH}$ data.

Table III lists the various crystals used in the complete data set with their unit cells and relative total scattering power. The discrepancy index is a measure of the reproducibility of the level scaled $F$ values within a particular subset of the data. The size of this index appears to correlate roughly with the size of the crystal as reflected by its scattering power and with the range in reciprocal space in which the data were taken. For example, the highest values of the index are for the relatively small iodine-substituted crystals in the 4 to 2 A data set. The measured variation in the unit cell parameter among the different crystals, especially along the $a$ axis, is greater than desirable, particularly since the calculations of Crick and Magdoff (9) show that such variations can themselves lead to marked intensity changes. However, the consequences of this variability in unit cell dimensions are presumably at least partially mitigated by the data-averaging and shell-scaling procedures. Perhaps the best indications that these variations in unit cell dimensions are not of marked consequence are that the difference Patterson maps for both derivatives, even with only the 4 to 2 A data, have very low noise levels and that the heavy atom refinements, particularly for the iodine derivative, proceeded well. Nevertheless, at least some of the variability which we observe in the quality of the resolution from different areas of the electron density maps of the enzyme is probably a result of this spread of unit cell dimensions. We cannot resolve this question until we have collected a full set of data on large crystals with more uniform cell dimensions.

### *Location of Heavy Atom Positions*

The positions of the heavy atoms were located in three-dimensional difference Patterson syntheses with the values of $F_H^2$ estimated from Matthews' formula (8), Equation 3, as the coefficients. For these syntheses, the value of the weighting factor, $w$, in Equation 3 was taken to be 0.75 as recommended by Matthews (8); the value of $k$, the ratio of the imaginary to real scattering, $\Delta f''/f_0 + \Delta f'$, was calculated from the tabulated values of the atomic scattering factors, $f_0$, and the related real, $\Delta f'$, and imaginary, $\Delta f''$, anomalous dispersion corrections listed in the *International Tables for Crystallography* (10) or given by Cromer (11). Both the pdIUp and the barium ion derivatives gave clean, easily interpretable, single site derivatives in separate Patterson maps run on the 4 A and the 4 to 2 A data. The interpretations were checked with "cross Fourier" syntheses, as suggested by Dickerson *et al.* (7), where the protein phases were derived from a single site $PtCl_4^=$ derivative of the uninhibited nuclease crystals (1). Such syntheses also allowed an estimate of the relative $z$ coordinates of the heavy atoms to be made. We have, on occasion, computed three-dimensional

TABLE I

*Notation and definitions*[a]

1. General

$F_P$ = the structure amplitude of a reflection from the native enzyme.

$F_{PH}$ = the structure amplitude of a reflection from a heavy atom derivative. ($F_{PH}^+$ and $F_{PH}^-$ are the amplitudes of a Friedel pair, and $F_{PH} = (F_{PH}^+ + F_{PH}^-)/2$, where applicable.)

$F_H$ = the observed structure amplitude of the heavy atom contribution as calculated from Equation 3, Matthews' formula (8).

$f_H$ = the calculated structure amplitude of the heavy atom contribution.

A superscript arrow, *e.g.* $\vec{F}_P$, denotes the structure factor.

$k = \Delta f''/f_0 + \Delta f'$, but for the refinements is estimated experimentally as given below.

2. Error estimates and phase probabilities

A. (1) The probability, $P(\alpha_P)$, that a reflection of the native protein has the phase $\alpha_P$ is expressed in the following way.

$$P(\alpha_P) \text{ is proportional to } \exp - (\epsilon^2/2E^2) \tag{14-16}$$

(2) The figure of merit, $m$, is defined elsewhere (15). We define $\langle m \rangle$ in the following way.

$$\langle m \rangle = \text{mean figure of merit} = \frac{\sum_{hkl} m}{n}$$

B. Scheme A refinement

(1)
$$R_{K-M} = \frac{\sum_{hkl} |F_H - f_H|}{\sum_{hkl} F_H}$$

(2) Isomorphous error for a particular derivative

$\epsilon_I = \| F_{PH} |_{obs} - | F_{PH} |_{calc} |$, where $| F_{PH} |_{calc} = \| F_P | \exp^{i\alpha_P} + \vec{F}_H |$ and $\alpha_P$ is the particular trial value of the protein phase

$$E_I = \left( \frac{\sum_{hkl} (|F_H - f_H|)^2}{n} \right)$$

where $E_I$ is read from a smoothed curve drawn through the discrete values plotted as a function of $2\theta$.

(3) Anomalous errors for a particular derivative

$$\epsilon_A = | \Delta F_{PH\text{obs}} - \Delta F_{PH\text{calc}} |$$

where

$$\Delta F_{PH\text{obs}} = F_{PH}^+ - F_{PH}^-, \tag{16, 17}$$

and

$$\Delta F_{PH\text{calc}} = \frac{-2F_P k}{w |F_{PH}|_{\text{calc}}} (b_H \cos \alpha_P - a_H \sin \alpha_P)$$

and

$$k = \overline{|F_{PH}^+ - F_{PH}^-|} / 2 \overline{|F_{PH} - F_P|} \tag{8}$$

where $k$ is read from a smoothed curve, as above and $w = 0.75$ (*cf.* Footnote *b*), $(a_H + ib_H) = \vec{F}_H$, and $\alpha_P$ is again the trial value of the protein phase.

$$E_A = \left( \frac{\sum_{hk0} |F_{PH}^+ - F_{PH}^-|}{n} \right)^{\frac{1}{2}}$$

(Note that the sum is over centric reflections only)

and $E_A$ is read from a smoothed curve, as above.

C. Scheme B refinement

(1) Isomorphous error for a particular derivative

$\epsilon_I = \| F_{PH} |_{obs} - | F_{PH} |_{calc} |$ where

$| F_{PH} |_{calc} = \| F_P | \exp^{i\alpha_P} + \vec{f}_H |$ and $\alpha_P$ is the particular trial value of the protein phase.

$$E_I = \left( \frac{\sum_{hkl} \epsilon_I^2}{n} \right)^{\frac{1}{2}} \tag{22}$$

but the $\alpha_P$ used in estimating $| F_{PH} |_{calc}$ is the most probable phase from the previous refinement cycle.

TABLE I—*Continued*

(2) Anomalous error for a particular derivative
$\epsilon_A = |\Delta ANO_{calc} - \Delta ANO_{obs}|$ where

$$\Delta ANO_{obs} = F_{PH}^{+} - F_{PH}^{-} \quad \text{and} \quad \Delta ANO_{calc} = kD$$

where

$$D = \frac{2F_P f_H \sin(\alpha_P - \alpha_H)}{|F_{PH}|_{calc}}$$

and $\alpha_P$ is the particular trial value of the protein phase and $\alpha_H$, the calculated value of the heavy atom phase in

$$\vec{f}_H = f_H \exp^{i\alpha_H},$$

and

$$k = \frac{\sum_{hkl} \Delta ANO_{obs} D}{\sum_{hkl} D^2} \tag{22}$$

but in which expression, $D$ is computed with the most probable phase from the last refinement cycle for $\alpha_P$ and $|F_{PH}|_{obs}$ replaces $|F_{PH}|_{calc}$, *cf.* footnote to Table IV.

$$E_A = \left(\frac{\sum_{hkl} \epsilon_A^2}{n}\right)^{\frac{1}{2}}$$

but with the values of $k$ and $\alpha_P$ being those from the last refinement cycle and $|F_{PH}|_{obs}$ replacing $|F_{PH}|_{calc}$.

(3) $R$ values

$$R_{AD} = \left(\frac{\sum_{hkl} |\Delta ANO_{obs} - \Delta ANO_{calc}|^2}{\sum_{hkl} \Delta ANO_{calc}^2}\right)^{\frac{1}{2}} \tag{22}$$

$$R_M = \frac{\sum_{hkl} \sum_{i} |\epsilon_I|}{\sum_{hkl} \sum_{i} f_H}$$

where $\sum_i$ represents the sum over the various heavy atom derivatives, if applicable and $|\epsilon_I|$ is computed with the most probable phase.

$$R_W = \frac{\sum_{hkl} \sum_{i} W\epsilon_I^2}{\sum_{hkl} \sum_{i} Wf_H^2}, \tag{22}$$

where

$$W = \frac{1}{E_{I\theta}^2 + E_{I_F}^2}$$

and $E_{I\theta}$ and $E_{I_F}$ are functions of $\sin\theta$ and the size of $F_P$, respectively

$$R_K = \frac{\sum_{hkl} ||F_{PH}|_{obs} - |F_{PH}|_{calc}|}{\sum_{hkl} |F_{PH}|_{obs}} \text{ (the Kraut } R \text{ value)} \tag{6}$$

$$R_C = \frac{\sum_{hk0} ||F_{PHobs} \pm F_P| - f_H|}{\sum_{hk0} |F_{PHobs} \pm F_P|} \text{ (the centric } R \text{ value)} \tag{23}$$

$$R_{LS} = \left(\frac{\sum_{hkl} W|\epsilon_I|^2}{\sum_{hkl} W|F_{PH}|^2}\right)^{\frac{1}{2}}$$

and the protein phase is again the most probable from the completed cycle.

[a] The specific forms of the equations given above are not necessarily those actually evaluated in the computer programs.

[b] We have found that the weighting factor, $w$, in this equation for estimating $\Delta F_{PHcalc}$ was used incorrectly. The principal effect of this error is to shift the angular position.

TABLE II

*Ratios of mean intensities based on scaling by Kraut's method*

| Ranges | | $\langle I_{PH} \rangle / \langle I_P \rangle$ | |
|---|---|---|---|
| $2\theta$ | $d$, A | $I - C$ | $Ba^{2+} - Ca^{2+}$ |
| 0–15° | ∞–5.91 | 1.05 | 1.05 |
| 15–19 | 5.91–4.67 | 1.04 | 1.05 |
| 19–21 | 4.67–4.23 | 1.03 | 1.04 |
| 21–23 | 4.23–3.87 | 1.04 | 1.04 |
| 23–27 | 3.87–3.30 | 1.04 | 1.03 |
| 27–31 | 3.30–2.88 | 1.05 | 1.04 |
| 31–36 | 2.88–2.49 | 1.06 | 1.05 |
| 36–45 | 2.49–2.01 | 1.07 | 1.06 |

anomalous (12) and isomorphous (13) difference Patterson syntheses, with coefficients $(F_{PH^+} - F_{PH^-})^2$ and $(F_{PH_{hkl}} - F_{P_{hkl}})^2$, respectively, to check our interpretations and to assess grossly the relative contributions of the two types of differences to the $F_H^2$ syntheses. In our hands, the more conventional centric difference Patterson syntheses $(F_{PH_{hk0}} - F_{P_{hk0}})^2$ have proven to be relatively much more difficult to interpret.

*Refinement of Heavy Atom Parameters; Phase Determination*

Two independent schemes were used for the refinement of the heavy atom parameters, the estimation of the isomorphous and anomalous errors, and the consequent production of the combined phase probability curves for the individual reflections of the native enzyme-inhibitor complex. In each of these schemes, the phase probabilities were computed at 5°-intervals around the phase circle following the concepts of Blow and Crick (14) as applied by Dickerson, Kendrew, and Strandberg (15) with the anomalous dispersion information treated according to the method of North (16) and Matthews (17). Again, in both cases the centroid or "best" phases (14) were used in the calculation of the electron density maps.

*Refinement Scheme A*—In 1965, Kartha (18) suggested the refinement of heavy atom parameters, where both isomorphous and anomalous dispersion differences are available, by minimization of the quantity which we write as $\sum_{hkl} (F_H - f_H)^2$. Refinement was carried out with a standard least-squares program (19), modified to permit cycle by cycle monitoring and control on a time-sharing computer and to meet the special needs of the method. Of necessity, each type of heavy atom derivative must be refined independently; thus, it is not possible to refine the relative $z$ coordinates between types of derivatives in a space group such as $P4_1$ with this scheme. Table I contains the details concerning the reliability estimates and describes the experimental determination of $k$.

*Refinement Scheme B*—This refinement scheme, which stems from the concepts of error analysis put forth by Blow and Crick (14), minimizes the lack of closure between $F_{PH_{obs}}$ and $F_{PH_{calc}}$ and recalculates the "most probable" or "best" protein phase or both after each refinement cycle. The method was first applied by Dickerson *et al.* (15) and has been extended, modified, and described principally by Lipscomb *et al.* (20), Muirhead *et al.* (21), and Dickerson *et al.* (7). For this work we have used a program kindly supplied to us by Professor M. G. Rossman that follows the concepts described in a publication from his laboratory (22) wherein the anomalous dispersion information is incorporated into the combined phase probability curves after the ideas of North (16) and Matthews (17). Again, a more specific formulation is given in Table I.

TABLE III

*Parameters of crystals used*

| Crystal type | Crystal number | Unit cell A[a, b] ($a \times c$) | Relative scattering power[b] | Discrepancy index[c] | Data set[d] |
|---|---|---|---|---|---|
| Native | 114 | 48.17 × 63.26 | 3.7 | 1.6 | 4 A |
| | 130 | 48.36 × 63.33 | 4.4 | 1.4, 2.5 | 4 A |
| | 131 | 48.36 × 63.37 | 4.3 | 2.3, 1.7 | 4 A |
| | 143 | 48.07 × 63.26 | 4.8 | 4.5 | 4–2 A |
| | 144 | 48.01 × 63.23 | 6.2 | 5.3, 2.6 | 4–2 A |
| | 149 | 48.17 × 63.31 | 3.1 | 7.6 | 4–2 A |
| | 150 | 48.14 × 63.31 | 5.4 | 5.7 | 4–2 A |
| Iodine substituted | 132 | 48.44 × 63.29 | 1.5 | 9.1 | 4 A |
| | 133 | 48.44 × 63.31 | 1.4 | 5.6 | 4 A |
| | 134 | 48.40 × 63.30 | 1.6 | 6.4 | 4 A |
| | 140 | 48.14 × 63.26 | 2.1 | 11.7 | 4–2 A |
| | 141 | 48.14 × 63.18 | 2.3 | 10.7 | 4–2 A |
| | 142 | 48.20 × 63.20 | 1.9 | 11.3, 9.6 | 4–2 A |
| | 147 | 48.17 × 63.31 | 1.4 | 14.1, 15.3 | 4–2 A |
| | 148 | 48.24 × 63.26 | 1.0[e] | 12.7, 14.2 | 4–2 A |
| Barium substituted | 146 | 48.17 × 63.26 | 2.8 | 3.5, 4.6 | 4 A |
| | 151 | 48.00 × 63.26 | 4.4 | 4.3 | 4 A |
| | | | | 4.7, 3.2 | 4–2 A |
| | 152 | 48.11 × 63.38 | 4.7 | 4.8, 6.6 | 4–2 A |
| | 153 | 48.04 × 63.29 | 2.3 | 3.5 | 4 A |

[a] The standard unit cell is taken as 48.2 × 63.3. The space group is $P4_1$ (1).

[b] The last decimal place is of doubtful significance.

[c] $R = \sum(F_{obs} - F_{Av})/\sum F_{Av} \times 100$ where $F_{Av}$ is the average value of a given reflection over all data sets for a given range (4 A or 4 to 2 A) and crystal type (native, barium, iodine). The $R$ values here are computed only within the particular crystal type and data set. More than one value is given when more than one set of data was collected on that crystal.

[d] Historically, the 4 A and 4 to 2 A sets of data were collected independently and treated separately until finally combined to produce the high resolution map.

[e] The size of this crystal was 0.10 × 0.23 × 0.58 mm.

*Results and Comparison of Refinement Schemes*—In general, the Scheme A refinement has the advantage that information from both the isomorphous and anomalous differences are contributing to the refinement since the value of $F_H$, except for the centric reflections, must depend on both measurements. In addition, this refinement procedure is perhaps conceptually simpler because it involves the direct, and crystallographically conventional, comparison of a relatively simple structure (that of the heavy atom arrangement within the protein) to one's model of that structure. Hence, it might then be expected that the $R_{K-M}$ (see definition in Table I) values might strongly reflect the relative quality of different derivatives. In practice, however, this does not appear to be the case since all of the published values of $R_{K-M}$ fall within the range of 0.36 to 0.48 (1, 18, 24). In particular, we direct attention to our comparison (see "Comparison of Iodine and Barium Derivatives") of the quality of the iodine and barium derivatives. Although there is a strong

TABLE IV

*Heavy atom parameters after refinement*

| Heavy atom | Refinement Scheme A | | Refinement Scheme B | |
|---|---|---|---|---|
| | $I$ | $Ba^{2+}$ | $I$ | $Ba^{2+}$ |
| $x$ | 0.0615 (2)[a] | 0.1183 (2) | 0.0615 (1) | 0.1191 (2) |
| $y$ | 0.4106 (2) | 0.2689 (2) | 0.4112 (1) | 0.2683 (2) |
| $z$ | 0.3350[b] | 0.2674[b] | 0.3346 (1) | 0.2653 (1) |
| $B_{iso}$ (A²) | 4.0 | 6.0 | 11.6 | 14.4 |
| $R_{K-M}$ | 0.40 | 0.45 | | |
| Occupancy | Assumed full $I - C$ | Assumed full $Ba^{2+} - Ca^{2+}$ | 47 (electrons) | 39 (electrons) |

[a] The value in parentheses is the standard deviation of the estimated heavy atom position.

[b] Not refined.

indication that the iodine derivative is highly isomorphous, whereas the barium derivative seems to be significantly less so, the $R_{K-M}$ values (Table IV), 0.40 and 0.45, respectively, differ very little. We suggest four possible reasons for such insensitivity. (*a*) Anomalous differences intrinsic to the native protein produce error in the estimates of the heavy atom anomalous differences (25). (*b*) As small differences between large numbers, the anomalous differences are prone to error, and such errors have a marked effect on the estimates of $F_H$ in Equation 3. (*c*) Estimates of $F_H$ are generally poor for weak reflections and for reflections with small isomorphous or anomalous differences. (*d*) In our procedures, there is a direct dependence of the scaling constant between $F_P$ and $F_{PH}$ on $F_H$. A refinement scheme that deals with each type of heavy atom derivative independently is subject to the inherent disadvantage that it cannot jointly refine all the available data to produce the best estimate of the desired quantity, the protein phases, and it also has the practical disadvantages that it is unable to refine the relative occupancies and the relative origin positions (for the $z$ axis in $P4_1$).

Refinement Scheme B has the advantage that it does jointly refine all the data from the isomorphous differences, including the scale factors, relative occupancies, and relative origins between derivatives. As Phillips (26) and North and Phillips (27) have pointed out, this scheme, in its assignment of all errors to $F_{PH_{obs}}$, places heavy reliance on the accurate measurement of $F_P$. However, such an allocation of errors should be valid except in cases where there is reason to suspect the relative quality of the estimation of $F_P$, compared to the estimation of $F_{PH}$. This might happen where crystals degrade particularly rapidly, or where there is a very isomorphous derivative so that the errors might arise mainly in the measurement rather than in nonisomorphism or disorder. Since we believe our iodine derivative to be very highly isomorphous, we wondered if this might cause the two refinement schemes to give different final values for the positional and thermal parameters. As shown in Table IV and discussed below, this does not occur. On the other hand, a new and more extensive set of data, in which each $F_P$ is derived from only a single measurement, does not refine as well in the lack of closure program as the original data set which contains three or more estimations of $F_P$. One possible explanation is that the measurement errors in $F_P$ are now significant in comparison with the total errors in $F_{PH}$. At present these results are too preliminary to warrant further comment. Scheme B, as it is used here, employs the information from the anomalous differences only at the phase-determining steps and not in the least squares minimization. Since it is conceptually quite possible that the anomalous differences, if refined separately, could lead to values for the occupancies, temperature factor, and positional parameters for a derivative site which are somewhat different from those obtained from the isomorphous differences, the lack of closure programs should perhaps be rewritten, so as to test this possibility experimentally. As a less stringent test, we note that for our barium derivative, where the available evidence suggests a significant degree of nonisomorphism, the refinement schemes give essentially the same final parameter (Table IV). From this we tentatively conclude that a refinement program which separately minimizes the differences between the observed and calculated anomalous differences is an unnecessary elaboration. The results of Herriott *et al.* (24), except for some discrepancies in a few of the temperature factors and occupancies, suggest the same conclusion. Finally, since in Scheme B the RMS error values depend on the combined estimates of the protein phases, it is possible that the information from a derivative of high quality might be degraded by one of low quality. Such an occurrence would be likely only if the initial estimates of the parameters of the high quality derivative differed significantly from the true parameters, thus leading to high values for its RMS errors. This could cause refinement to a false minimum or such slow convergence as to give the appearance of a minimum. Utilization of two refinement schemes can serve as a check for this possibility as well as others since the final heavy atom parameters from both schemes should be essentially the same if both are valid procedures. Thus, the trials of Dickerson *et al.* (7), as well as the comparative results of the two schemes of refinement presented by Herriott *et al.* (24) and in this paper, are direct experimental confirmations of the reliability of lack of closure technique of refinement.

Before the experimental results of the two refinement schemes are discussed, it is necessary to note that we cannot rigorously justify comparison since the data sets used in the two schemes are not identical. In the period between the application of the A and B refinement schemes, about 760 individual measurements within the some 21,000 multiply determined data points, inclusive of Friedel pairs, were discarded as being grossly at variance with the other measurements for that reflection. In addition, some 300 data points, mostly weak reflections with inconsistent repetitive measurements from the iodine derivative data set, were completely discarded. Since the final refinement parameters and the resultant electron density maps give essentially identical results, comparison of the refinement schemes seems justified.

The positional and thermal parameters resulting from the two refinement schemes are given in Table IV. The only appreciable variation of the heavy atom positions between the refinements is in their relative $z$ positions, which could not be refined in Scheme A. The apparent large difference in temperature factors is principally due to the use of theoretical radial atomic scattering curves ($I - C$; $Ba^{2+} - Ca^{2+}$) (10) in Scheme A, whereas in Scheme B the radial fall-off of the heavy atom scattering is included in the temperature factor. Similar results were found by Herriott *et al.* (24) in their use of two similar methods of heavy atom refinement in the structure determination of rubredoxin. In Scheme A the heavy atom occupancies were assumed to be unity; in Scheme B the change

in scattering was initially set to the theoretical value $f_{0_I} + \Delta f'_I - f_{0_C}$ of 46 electrons but then allowed to vary in the refinement. The iodine occupancy then varied between 46 and 47 electrons, and the resultant occupancy of the barium site came to 39 electrons, not far from the expected value ($f_{0_{Ba^{2+}}} + \Delta f'_{Ba^{2+}} - f_{0_{Ca^{2+}}} - \Delta f'_{Ca^{2+}}$) of 34 electrons, and suggesting that the three additional electrons of the 5-methyl group of pdTp should also have been subtracted from the replacing iodine atom. Table V gives the values of $k^{-1}$, the anomalous scattering constant, as a function of interplanar spacings. As can be easily seen, neither of the empirical calculations of $k^{-1}$ gives a very good approximation to the theoretical values expected with our data, a result which is not surprising considering the inherent difficulty of measuring $|F_{PH}^+ - F_{PH}^-|$ accurately.

In Table VI the figures of merit and the RMS errors derived from the two schemes are compared again as a function of resolution. Considering that the way the errors were estimated differs between the refinement schemes, the values, except for the RMS isomorphous error for the iodine derivative, are reasonably consistent. The larger value of $E_I$ for the iodine derivative in Scheme A can be traced to the estimation of $F_{PH_{calc}}$ in which $F_H$, a quantity that can be markedly in error for weak reflections from small crystals, is used rather than $f_H$.

*Comparison of Iodine and Barium Derivatives*—Table VII is an extensive compilation of the final refinement criteria from Scheme B both as a function of resolution and for the data set as a whole. We have computed a number of $R$ values commonly used in protein crystallography to facilitate comparison of our data with that of others. It would appear, particularly from comparison of the $E_I$ values with the RMS $f_H$ values, that the information derived from isomorphous differences remains reliable for both derivatives at least to the limit of resolution of the current data set. On the other hand, the anomalous differences, as indicated by the values of $R_{AD}$, become unreliable for many reflections beyond spacings of about 3 A. In considering Table VII, however, it should be kept clearly in mind that beyond 4.0 A resolution, the data set contains only that third of the reflections with the highest peak to background ratios, a fact that may heavily weight the various refinement criteria toward a favorable appearance.

The refinement criteria in Table VII indicate that the model of an iodine replacing a methyl group is a considerably better fit

TABLE V

*Values of $k^{-1}$ where k is ratio of imaginary to real part of scattering factor*

| Range of $d$ spacings | ∞–5.91 | 5.91–4.67 | 4.67–4.23 | 4.23–3.87 | 3.87–3.30 | 3.30–2.88 | 2.88–2.49 | 2.49–2.01 |
|---|---|---|---|---|---|---|---|---|
| Empirical formulation Matthews' (8) Scheme A | | | | | | | | |
| $I - C$ | 8.1 | 8.9 | 10.0 | 8.1 | 5.7 | 5.6 | 4.0 | 2.9 |
| $Ba^{2+} - Ca^{2+}$ | 4.9 | 4.9 | 7.7 | 6.5 | 5.2 | 4.7 | 3.9 | 4.1 |
| Empirical formulation Adams' *et al.* (22) Scheme B[a] | | | | | | | | |
| $I - C$ | 7.0 | 6.8 | 6.8 | 6.8 | 6.9 | 7.1 | 6.9 | 6.0 |
| $Ba^{2+} - Ca^{2+}$ | 4.6 | 5.4 | 5.5 | 5.4 | 5.4 | 5.8 | 6.1 | 6.2 |
| Theoretical calculations (10, 11) | | | | | | | | |
| $I - C$ | 6.8 | 6.5 | 6.4 | 6.3 | 6.2 | 6.0 | 5.8 | 5.6 |
| $Ba^{2+} - Ca^{2+}$ | 4.1 | 4.0 | 3.9 | 3.8 | 3.8 | 3.7 | 3.6 | 3.4 |

[a] Although the values of $k^{-1}$ are given as a function of resolution, the mean values of 7.0 for $I - C$ and 5.7 for $Ba^{2+} - Ca^{2+}$ were actually used in computation.

TABLE VI

*Comparative error estimates for refinement schemes*

| Range of $d$ spacing | ∞–5.91 | 5.91–4.67 | 4.67–4.23 | 4.23–3.87 | 3.87–3.30 | 3.30–2.88 | 2.88–2.49 | 2.49–2.01 |
|---|---|---|---|---|---|---|---|---|
| Figure of merit $<m>$ | | | | | | | | |
| Scheme A | 0.78 | 0.81 | 0.84 | 0.84 | 0.88 | 0.86 | 0.83 | 0.78 |
| Scheme B | 0.87 | 0.81 | 0.79 | 0.78 | 0.77 | 0.78 | 0.75 | 0.67 |
| Iodine derivative isomorphous RMS error, $E_I$ | | | | | | | | |
| Scheme A | 38 | 38 | 40 | 41 | 30 | 27 | 33 | 36 |
| Scheme B | 19 | 19 | 19 | 21 | 22 | 18 | 17 | 14 |
| Barium derivative isomorphous RMS error, $E_I$ | | | | | | | | |
| Scheme A | 37 | 54 | 46 | 46 | 31 | 27 | 27 | 27 |
| Scheme B | 38 | 48 | 50 | 45 | 38 | 24 | 22 | 20 |
| Iodine derivative anomalous RMS error, $E_A$ | | | | | | | | |
| Scheme A | 22 | 22 | 16 | 16 | 11 | 9 | 18 | 27 |
| Scheme B | 12 | 12 | 13 | 13 | 13 | 14 | 18 | 24 |
| Barium derivative anomalous RMS error, $E_A$ | | | | | | | | |
| Scheme A | 25 | 7 | 10 | 9 | 7 | 7 | 16 | 17 |
| Scheme B | 12 | 12 | 12 | 11 | 11 | 11 | 13 | 14 |

TABLE VII

*Refinement criteria Scheme B*

| Heavy atom derivative | Criterion | Range of $d$ spacings, A | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | ∞–16.7 | 16.7–8.3 | 8.3–5.6 | 5.6–4.2 | 4.2–3.3 | 3.3–2.8 | 2.8–2.4 | 2.4–2.0 | ∞–2.0 |
| $Ba^{2+}$ | No. of reflections in zone | 16 | 125 | 331 | 604 | 894 | 847 | 658 | 602 | 4077 |
| | $E_I$ | 73 | 38 | 42 | 51 | 40 | 24 | 21 | 20 | |
| | RMS $f_H$ | 79 | 76 | 72 | 66 | 60 | 52 | 44 | 35 | |
| | RMS $\lvert F_{PH} - F_P \rvert$/RMS $f_H$ | 0.89 | 1.00 | 0.84 | 1.10 | 0.96 | 0.80 | 0.81 | 0.92 | |
| | $E_A$ | 13 | 12 | 11 | 12 | 12 | 12 | 15 | 15 | |
| | $R_{AD}$ | 0.68 | 0.70 | 0.63 | 0.69 | 0.75 | 0.91 | 1.26 | 1.49 | |
| | $R_M$ | 0.84 | 0.45 | 0.50 | 0.68 | 0.55 | 0.37 | 0.39 | 0.46 | 0.50 |
| | $R_W$ | 0.76 | 0.26 | 0.33 | 0.60 | 0.45 | 0.21 | 0.23 | 0.33 | 0.33 |
| | $R_K$ | 0.15 | 0.08 | 0.12 | 0.12 | 0.10 | 0.08 | 0.08 | 0.09 | 0.10 |
| | $R_C$[a] | 0.75 | 0.52 | 0.47 | 0.68 | 0.71 | 0.69 | 0.98 | 0.87 | 0.65 |
| | | (4) | (19) | (29) | (39) | (46) | (33) | (15) | (13) | (198) |
| | $R_{LS}$ | 0.17 | 0.10 | 0.15 | 0.15 | 0.12 | 0.10 | 0.11 | 0.13 | 0.12 |
| $I$ | No. of reflections in zone | 16 | 124 | 329 | 595 | 857 | 805 | 611 | 509 | 3846 |
| | $E_I$ | 20 | 20 | 19 | 19 | 22 | 17 | 15 | 14 | |
| | RMS $f_H$ | 94 | 92 | 88 | 83 | 76 | 69 | 60 | 51 | |
| | RMS $\lvert F_{PH} - F_P \rvert$/RMS $f_H$ | 0.90 | 0.76 | 0.69 | 0.71 | 0.73 | 0.71 | 0.63 | 0.66 | |
| | $E_A$ | 28 | 12 | 13 | 14 | 14 | 15 | 20 | 23 | |
| | $R_{AD}$ | 1.53 | 0.67 | 0.67 | 0.78 | 0.82 | 0.99 | 1.42 | 1.98 | |
| | $R_M$ | 0.18 | 0.12 | 0.08 | 0.12 | 0.16 | 0.17 | 0.16 | 0.19 | 0.15 |
| | $R_W$ | 0.05 | 0.04 | 0.05 | 0.05 | 0.09 | 0.06 | 0.06 | 0.07 | 0.06 |
| | $R_K$ | 0.04 | 0.03 | 0.02 | 0.03 | 0.04 | 0.04 | 0.04 | 0.05 | 0.04 |
| | $R_C$[a] | 0.30 | 0.35 | 0.36 | 0.35 | 0.51 | 0.68 | 0.97 | 0.84 | 0.47 |
| | | (5) | (19) | (29) | (39) | (41) | (31) | (13) | (10) | (187) |
| | $R_{LS}$ | 0.05 | 0.05 | 0.07 | 0.06 | 0.07 | 0.07 | 0.08 | 0.08 | 0.07 |
| Combined $I$ and and $Ba^{2+}$ | No. of reflections in zone | 17 | 125 | 338 | 623 | 910 | 847 | 664 | 603 | 4127 |
| | $R_M$ | 0.49 | 0.29 | 0.27 | 0.40 | 0.35 | 0.27 | 0.27 | 0.32 | 0.31 |
| | $R_W$ | 0.15 | 0.09 | 0.08 | 0.13 | 0.17 | 0.11 | 0.12 | 0.16 | 0.13 |
| | $<m>$ | 0.76 | 0.86 | 0.80 | 0.74 | 0.73 | 0.74 | 0.71 | 0.65 | 0.73 |

$E_I$ as a function of $F_P$ size

| Range of $F_P$ values | 0–100 | 100–200 | 200–300 | 300–400 | 400–500 | 500–600 | 600–700 | 700–∞ |
|---|---|---|---|---|---|---|---|---|
| No. of reflections in zone | 472 | 2982 | 2439 | 1133 | 499 | 252 | 92 | 54 |
| $E_{I_F}$ | 25 | 23 | 27 | 30 | 34 | 32 | 40 | 51 |

Number of reflections within a range of figure of merit

| Range of $<m>$ | 0.0–0.1 | 0.1–0.2 | 0.2–0.3 | 0.3–0.4 | 0.4–0.5 | 0.5–0.6 | 0.6–0.7 | 0.7–0.8 | 0.8–0.9 | 0.9–1.0 |
|---|---|---|---|---|---|---|---|---|---|---|
| No. of reflection | 79 | 121 | 131 | 160 | 218 | 288 | 404 | 608 | 999 | 1119 |

[a] Number of centric reflections is given in parentheses.

to the observed intensity changes than the replacement of $Ca^{2+}$ by $Ba^{2+}$. Adams *et al.* (22) have pointed out that the ratio of (RMS $\lvert F_{PH} - F_P \rvert$) to (RMS $f_H$) should be close to $2^{-1/2}$ (~0.71) for the acentric reflections in a structure; this provides a rather different way of judging the fit of a model. As shown in Table VII, this ratio for the iodine derivative is consistently quite close to 0.71, but for the barium derivative the ratio is, in most zones, considerab'y different. Since the results of this work show that the $Ca^{2+}$ ion, the pdTp, and the pdIUp are rigidly fixed in the structure, and the temperature factors for the iodine and barium derivatives are about equal, the implied difference in quality of the two derivatives is not to be ascribed to a difference in ordering or rigidity in the two cases. We think that the correct explanation is that, although the iodine atom does occupy

TABLE VIII

*Comparative coordinates of $Ba^{2+}$ and $Ca^{2+}$*

| | X | Y | Z |
|---|---|---|---|
| $Ba^{2+}$ | 0.119 | 0.268 | 0.265 |
| $Ca^{2+}$, anomalous Fourier | 0.114 | 0.285 | 0.262 |
| $Ca^{2+}$, single I substitution | 0.113 | 0.282 | 0.273 |

the same position as the methyl group it replaces, the $Ba^{2+}$ position differs significantly from the $Ca^{2+}$ position. Independent estimations of the $Ca^{2+}$ position can be obtained from the anomalous Fourier map of the native nuclease-inhibitor complex, which was reported earlier (1), and from an electron density map of the nuclease based on phases from the iodine derivative alone. The coordinates of the $Ca^{2+}$ from these two sources are tabulated with those of the $Ba^{2+}$ in Table VIII. The values in this table indicate that the $Ba^{2+}$ position differs from that of the $Ca^{2+}$ by about 0.75 A in the $Y$ direction with the possibility that there is also some displacement along $Z$. The differences in the physical sizes of these heavy and light atom moieties offers a basis for understanding these effects. The ionic radius of $Ba^{2+}$ (1.35 A) is 0.36 A or ~35% larger than that of $Ca^{2+}$ (0.99 A) (28), but the van der Waals' radius of covalently bonded iodine (2.15 A) is only 7.5% larger than that of a methyl group (2.0 A) (29). This positional difference is fully consistent with the known facts that nuclease activity is not retained when $Ba^{2+}$ replaces $Ca^{2+}$ and that $Ba^{2+}$ is inhibitory of the RNA hydrolysis at low $Ca^{2+}$ concentrations (30).

In Fig. 1 we show a well resolved region of the nuclease structure composed of three and one-half turns of the $\alpha$ helix (residues 54 to 67). These electron density maps have been calculated with phases from the iodine derivative alone, with the barium derivative alone, and from the two sets of data combined. It is not possible by visual inspection to decide from this region, or indeed from the whole map, which is the better derivative. Such a result is entirely possible because this sort of comparison need only be valid for two derivatives of different quality involving substitutions at the same site. Thus, despite the apparent considerable difference in quality between the derivatives, both make substantial contribution to the clarity of the joint map. It is also apparent that considerable useful information can be derived from an electron density map based on a single derivative, a point also illustrated in our low resolution work (1). However, some appreciation of the influence which derivative quality has on the clarity of an electron density map can be gained by comparison of the independent low resolution structure of the uninhibited nuclease with that of the nuclease-inhibitor complex. Fig. 2*A* shows the same helical region in a map of the uninhibited nuclease derived from three conventional heavy atom derivatives (tetrachloroplatinite, acetoxymercurianiline, and *p*-(chloromercuri)benzenesulfonate) at 4 A resolution (1). Fig. 2, *B* and *C* show this region in the inhibitor complex crystals at 4 A resolution phased first (2*B*) with only the iodine derivative and then (2*C*) with the combination of iodine and barium derivatives. It is evident that the map based on a single high quality derivative is clearer than the three-derivative map of the uninhibited nuclease, and this impression is even more striking when the entire three-dimensional maps are examined. We were, in fact, able to trace successfully (as confirmed in the high resolution work) nearly the whole of the peptide chain on the 4 A map based only on the iodine derivative, whereas we could not adequately interpret the 4 A map of the uninhibited structure, despite the fact that it was based on three derivatives rather than one. Thus, it would appear that a search for derivatives of high quality has substantial rewards. We must, however, hedge this conclusion by pointing out that the uninhibited nuclease is markedly more susceptible to trypsin digestion and denaturation (31) and that it exchanges its hydrogen atoms more readily (32) than does the nuclease-inhibitor complex, suggesting that the uninhibited structure may have intrinsically greater motility and hence poorer definition in the crystal. However, if this were to be a major factor in accounting for the differences illustrated in Fig. 2, the radial intensity fall-off of the uninhibited crystal should be markedly greater than that of the inhibited crystals; this does not appear to be the case.

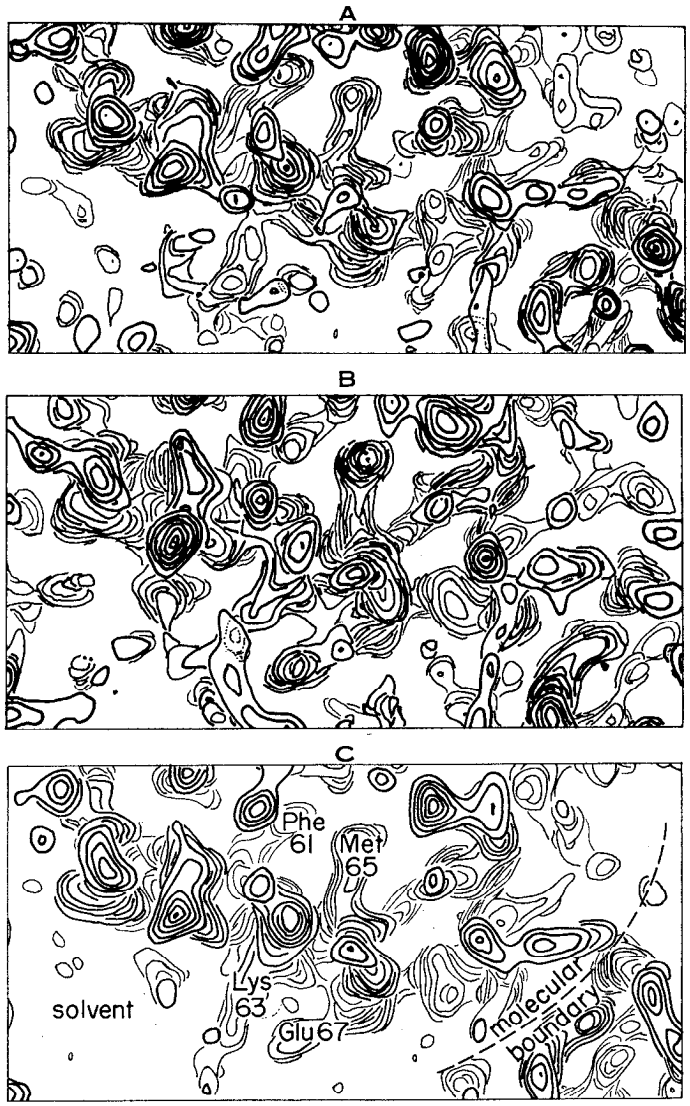


FIG. 1. A portion of the high resolution electron density map showing the helical region of residues 54 to 67. The heaviest lines represent contours closest to the eye of the viewer. (*A*) Map phased with only the iodine derivative. (*B*) Map phased with only the $Ba^{2+}$ derivative. (*C*) Map phased with both derivatives. The positions of several prominent side chains are identified.

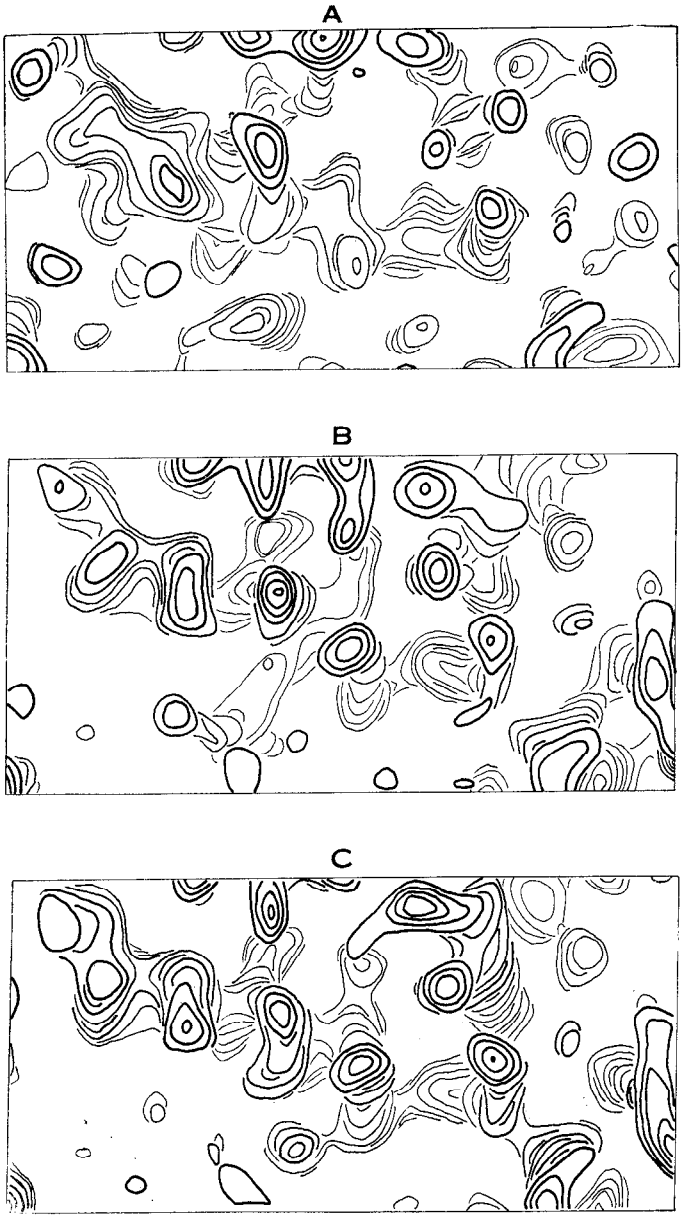


FIG. 2. The same helical section as in Fig. 1 at 4 A resolution. (*A*) The uninhibited nuclease phased with three heavy atom derivatives. (*B*) The nuclease inhibitor complex phased with only the iodine derivative. (*C*) The nuclease inhibitor complex phased with both the iodine and $Ba^{2+}$ derivatives.

*Preparation and Comparison of Electron Density Maps*

"Best Fouriers" were computed[4]; the resultant electron density grid was stored on a magnetic disk; and the desired perspective of the map was contoured[4] in appropriate sections with a Stromberg-Carlson 4020 cathode ray tube plotter. Normally we prepared maps of one-half unit cell in 70 sections perpendicular to the $z$ axis; the spacings between sections being 0.0074 of a unit cell or 0.47 A. About ½ hour of plotter time is required to prepare such a map. The plotter output is obtained both as 35-mm microfilm, which can be projected on a screen for contouring onto glass sheets at the 2 cm per A scale of the Kendrew models, and as 19-cm square paper copies, from which transparencies can be made for assembly into a small, portable electron density map.

Visual comparison of the electron density maps from the Scheme A and B refinements reveals only minor differences. This is confirmed by comparison of the calculated RMS difference in electron density between the Scheme A and B maps (0.12 e per $A^3$) with the calculated RMS error in electron density (15) of the Scheme A map (0.18 e per $A^3$).

STRUCTURE OF NUCLEASE-INHIBITOR COMPLEX

*Chain Tracing; Results.*—The path of the peptide chain has been traced in the electron density map in its entirety, except for residues 1 to 5 and 143 to 149, with the help of the known amino acid sequence (3, 33, 34).[5] Kendrew models have been fitted to the maps with the aid of a modified version of F. M. Richards' optical comparator (36). Fig. 3, *A* and *B* are, respectively, front[6] and right side[6] views of the path of the peptide chain. Fig. 4, *A* and *B* are stereo views produced by ORTEP (37) of the chain path from roughly the same orientations as those shown in Fig. 3. Viewing the molecule from the front, as in Figs. 3*A* and 4*A*, the chain can be traced starting at the rear upper left corner of the molecule with Lys-6. Residues 1 to 5 cannot be located in our present electron density map. From Lys-6 the chain goes diagonally down across the back of the molecule and begins to curve to the front at Pro-11. Residues 12 to 36 make up a three-stranded section of antiparallel $\beta$-pleated sheet which forms the upper right side of the nuclease molecule. The individual runs (residues 12 to 19, 21 to 27, and 30 to 36) of extended chain forming this section of pleated sheet each have six hydrogen bonds to adjoining peptide chains, but the sheet itself is considerably distorted, so as to appear somewhat convex as viewed from the outside. The pleated sheet can be visualized as a right hand cupping the rest of the nuclease molecule. There is a pronounced kink in the sheet structure where both the Ile-15 and Lys-16 side chains extend toward the outside of the molecule. Because of this distortion, the Ile-15 side chain does not project directly into solution but is instead folded back close to the surface of the molecule and partially shielded by the methylene side chains of Lys-16 and Lys-24 and by Met-26 and Thr-13. This grouping produces a distinct, though small, hydrophobic patch on the outer surface of the molecule.

From the end of the pleated sheet region at Leu-36, the chain next makes a complex loop into and then out of the center of the molecule forming the lower surface of the nucleotide-binding pocket. Residues Lys-45 to Lys-53 then constitute the loop that can be seen extending out from the rest of the structure in the right side views of the molecule, Figs. 3*B* and 4*B*. In the electron density maps this region appears at only one-third to one-half the density of the rest of the structure. There is some indication that this loop may interact weakly with another nuclease molecule in the crystal. Between Tyr-54 and Glu-67, the chain takes

[4] The Fourier program is a local version of A. Zalkin's FORDAP. The contouring program CONTR/ICONT, which prepares the input tape for the plotter, was written by S. Zisk and N. Brenner.

[5] The sequence is given in a figure of an accompanying paper (35).

[6] Each protein crystallographic group necessarily adapts an arbitrary canonical view of the molecule with which they work. We define the front of the nuclease as it is seen in Figs. 3*A* and 4*A*. Right and left refer to the reader's handedness.

the form of about three and one-half turns of the helix which runs from front to back forming the lower right side of the molecule. This is the helical region that was discussed earlier and illustrated in Figs. 1 and 2. As the electron density map has thus far been fitted in the Richards' comparator, no attempt has been made to constrain the helical regions to any particular pattern of hydrogen bonding. With one or two specific localized exceptions, these helical regions appear to be reasonable approximations of the $\alpha$ helix, but we will postpone a more searching examination and description of them to a later publication. At the end of this helical region (Glu-67) which occurs at the bottom right rear of the structure, the chain rises relatively directly up the back, loops across the top of the molecule to the front, and then goes back across the top and down the back to the beginning of another helical section at Val-99. As can be seen in Fig. 4*A*, the two chains just mentioned, which run approximately vertically up and down the back, are parallel to each other for about 6 residues along each. However, the orientation of each chain rotates by at least 90°, and there are only two or, at the most, three hydrogen bonds between the chains that conform substantially to an antiparallel pleated sheet arrangement.

The center helical region contains about two turns between Val-99 and Gln-106 and runs at an angle of about 20° with the horizontal to the front center of the molecule. From the end of this helix, the chain rises up and back to the top of the molecule forming part of its front left side and then turns down again to the third and last helical section. This section, Glu-122 to Lys-134, of some two and one-half turns runs from the upper left rear toward the lower left front at an angle of about 40° with the horizontal and forms much of the left side of the molecule. The tracing ends in a complex loop at the lower left front of the molecule and is convoluted so as to afford a pocket for the single tryptophan residue, Trp-140. We cannot locate the last 7 amino acid residues in the electron density map reported here.

In general, except for those hydrophilic residues that presumably extend into solution and are disordered therein over several conformations, the quality of the map is such that the side chains can be easily identified and fitted well to the electron density map. However, the definition in that area of the electron density map corresponding to the left side of the molecule, with the specific exceptions of Tyr-113 and Pro-117 to His-124, is slightly but definitely poorer than in the other well defined regions of the map.

*Chain Tracing; Discussion*—The extent to which variations in the definition of the electron density are attributable, on the one hand, to varying degrees of intrinsic localization of the atoms and, on the other hand, to inadequacies in the phase information is, at this time, a moot point. Only by comparing the present map with one based on demonstrably superior data can this be decided objectively. We expect to have such a map in the future and to examine this question at that time. It is our opinion, however, that, for the most part, the variations we now see in definition in different parts of the molecule can be correlated with chemically expected variations in the rigidity and ordering of the chain.

In the structure of ribonuclease S (5), there is a precedent for missing residues at the ends of peptide chains and the lack of clarity in an exposed loop, but this situation in the *Staphylococcal* nuclease is, at present, rather more pronounced. Taniuchi *et al.* (38–40) have shown that, of the potential points of proteolytic cleavage, trypsin attacks the nuclease-inhibitor complex at only two points. First, the five $NH_2$-terminal amino acid residues are removed, yielding a fully active enzyme; secondly, the peptide chain is broken in the middle of the exposed loop, between Lys-48 and Lys-49, yielding an enzyme with 8 to 10% of the initial activity. These observations indicate that these regions are markedly more exposed, and thus very likely to be markedly more motile, than the remainder of the molecule.

From binding studies of oligonucleotides to the nuclease, Cuatrecasas, Wilchek, and Anfinsen (41) have concluded that there are three significant binding subsites on the nuclease. There is also a marked loss of enzymatic activity when the peptide chain is broken between Lys-48 and Lys-49 (38, 39). This leads us to speculate that the exposed loop, in the presence of a natural substrate, may form a part of a binding subsite of the active site. In this regard we should point out that the conformation of this exposed loop, as seen in this structure, could well be an artifact of the crystal packing since it appears to form contacts with a neighboring molecule. Perhaps in the absence of substrate it does indeed flap loosely around in the solvent, but in the presence of substrate it is immobilized in a particular, enzymatically productive conformation, that is, it may assist in binding the substrate. It is also possible that the hydrophobic patch in the pleated sheet area, which was mentioned earlier, might also be a part of a binding subsite.

Recently, as part of the continuing investigation of structured, partially enzymatically active derivatives of nuclease formed from various natural and synthetic fragments of the peptide chain (35), Parikh, Corley, and Anfinsen[7] have shown that the absence of residues 142 to 149 or the replacement of Trp-140 by phenylalanine does not appear to affect enzymatic activity or the solution structure significantly (42). These observations are, of course, quite consistent with the facts that (*a*) we do not find residues 143 to 149 in the electron density maps, suggesting they are extremely mobile, and (*b*) the Trp-140 side chain is relatively poorly defined in the map.

The structure of this nuclease conforms with the now accepted dogma (27) that a vital structural element of globular proteins is an internal core of very well ordered, nonpolar side chains. As North and Phillips (27) point out, this is a significantly more limited and accurate generalization than the terms hydrophilic out, hydrophobic in. Indeed, for many side chains the distinction between inside and outside is poorly defined, and it cannot be said that all exterior side chains are polar.

*Inhibitor Binding Site; Results*—The position of the pdTp inhibitor can be seen in relation to the body of the enzyme in Fig. 3*B*. The positions of the two phosphate groups and the 5-methyl position (labeled I) of the nucleotide, as well as the $Ca^{2+}$ position, are indicated in Fig. 4, *A* and *B*. Fig. 5 is a schematized stereo view of the inhibitor, the $Ca^{2+}$, and the various groups from the enzyme that the electron density map indicates are involved in the formation of the ternary complex. Some sections from the actual electron density map have been published elsewhere (43), and another sketch of the active site is shown in an accompanying paper (44).

In general, the inhibitor molecule itself and adjacent parts of the nuclease molecule are very well and clearly defined in the electron density map. As discussed below, the area around the $Ca^{2+}$ ion is less well defined, but a very reasonable fit to the observed density is obtained.

The pyrimidine ring of the pdTp fits into a well defined pocket, which has previously been well illustrated in photographs of the

[7] I. Parikh, L. Corley, and C. B. Anfinsen, unpublished results.

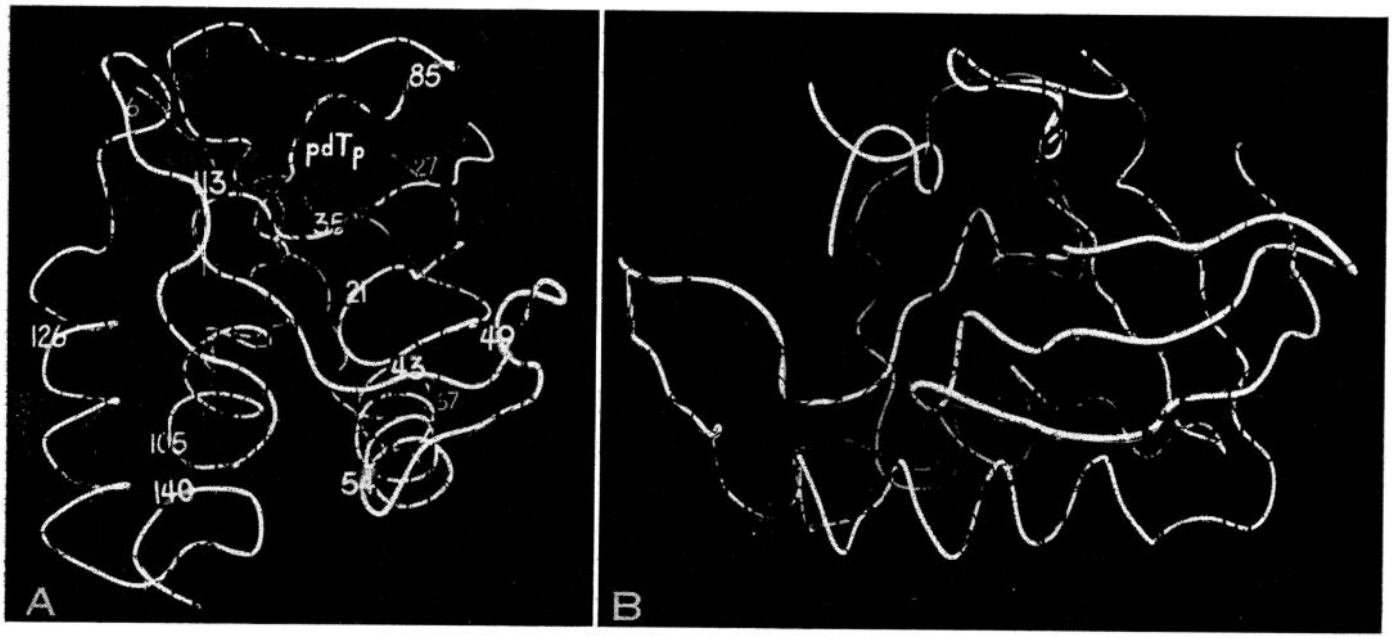


FIG. 3. The path of the peptide chain as viewed from (*A*) the front, and (*B*) the right side. In 3*A* the approximate position of certain residues is indicated and the area of the pocket is indicated with the symbol, *pdTp*. In 3*B* a loop indicates the position of the inhibitor; the *large loop* represents the pyrimidine ring, the *small loop* represents the ribose, and the *ends of the tubing* represent the positions of the phosphates. These photographs, prepared by J. S. and D. C. Richardson and C. B. Anfinsen, were made by fastening thin walled plastic tubing to the $\alpha$ carbon atoms of a wire model, filling it with a solution of fluorescein, and photographing the entire model under ultraviolet light.

low resolution model (1, 43). This pocket lies in the upper left front octant of the nuclease as viewed in Figs. 3*A* and 4*A*. The lining of the pocket, with the exception of those groups involved directly or indirectly in binding the pdTp and $Ca^{2+}$, is nonpolar, and there are several clearly defined electron density maxima which we believe to be water molecules (not shown here) in this pocket along with the inhibitor.

From a comparison of the known absolute configuration of the inhibitor molecule with the shape of the molecule as observed in our electron density maps, we can conclude unambiguously that the upper phosphate group is the one in position 3′, and the lower one is that in position 5′. This result was, in fact, firmly established in our 4 A structure (1) and is simply substantiated and elaborated here in more detail. This is an essential fact which forms the indispensable starting point for any well grounded discussion of the mechanism of nuclease activity. It is established solely by this crystallographic investigation and is the single most important result thereof to date.

The calcium ion lies almost directly below the 5′-phosphate group, but at a distance of 4.7 $\pm$ 0.2 A from the phosphorus atom. Since the P to O distance must be approximately 1.5 A (it is 1.54 A in $PO_4^{3-}$), the minimum possible O to $Ca^{2+}$ distance is ~3.0 A. This is certainly too great to allow any strong interaction or direct coordination between the 5′-phosphate group and the calcium ion to be postulated. From the electron density maps, it seems clear that the hydroxyl group of Tyr-85 interacts with the 3′-phosphate; the interaction of the amino group of Lys-84 is probable but not as definite. The 5′-phosphate interacts with one arm of the guanidinium moiety of Arg-87, the other arm of which bridges to the carboxylate side chain of Asp-83. This in turn appears to interact with the peptide nitrogen atoms of Gly-86 or Tyr-85 or both. It is also possible that there is an interaction of the guanidinium with the ring oxygen of the deoxyribose. Another guanidyl group, that of Arg-35, also interacts on one side with the 5′-phosphate group and on the other with the peptide carbonyl of Leu-36 and perhaps more weakly with that of Val-39. Lys-71 from an adjacent molecule in the unit cell also interacts with the 5′-phosphate group. The $Ca^{2+}$ ion is coordinated by the carboxylato groups of Asp-21, Asp-40, and Glu-43.

The Tyr-113 side chain is very clearly defined in the map and lies below and to the left of the pyrimidine ring of the inhibitor. The planes of the two rings are parallel, but they are too distant (4.5 A) and their centers are too much offset to suggest any appreciable ring-ring interaction. Tyr-115 lies almost directly above Tyr-113, but, in marked contrast, its side chain is merely hinted at in the map. We consider this indicative of a very high degree of exposure and consequent disorder. This would be consistent with its high reactivity toward nitration (45).

Apparently because the $Ba^{2+}$ derivative is not fully isomorphous with the $Ca^{2+}$ form, there is no distinct peak associated with the $Ca^{2+}$ position, and the area associated with the carboxylato moieties around this position is somewhat indistinct. In the combined map there is no electron density observed between the $Ca^{2+}$ position and that of the 5′-phosphate; however, in the single derivative iodine map there is density in this region and a considerably sharper peak of the $Ca^{2+}$ position.

*Inhibitor Binding Site; Discussion*—When the pdTp-nuclease-$Ca^{2+}$ complex is viewed in the large scale, three-dimensional electron density map and the various modes of interaction among the members of the complex are examined, the very strong impression is created that the inhibitor is bound to the enzyme in a markedly specific and rigid manner. This might perhaps be called a trivial comment because it is something which would be expected, but there have so far been few opportunities to observe such a situation closely. The picture is rather dramatic. The

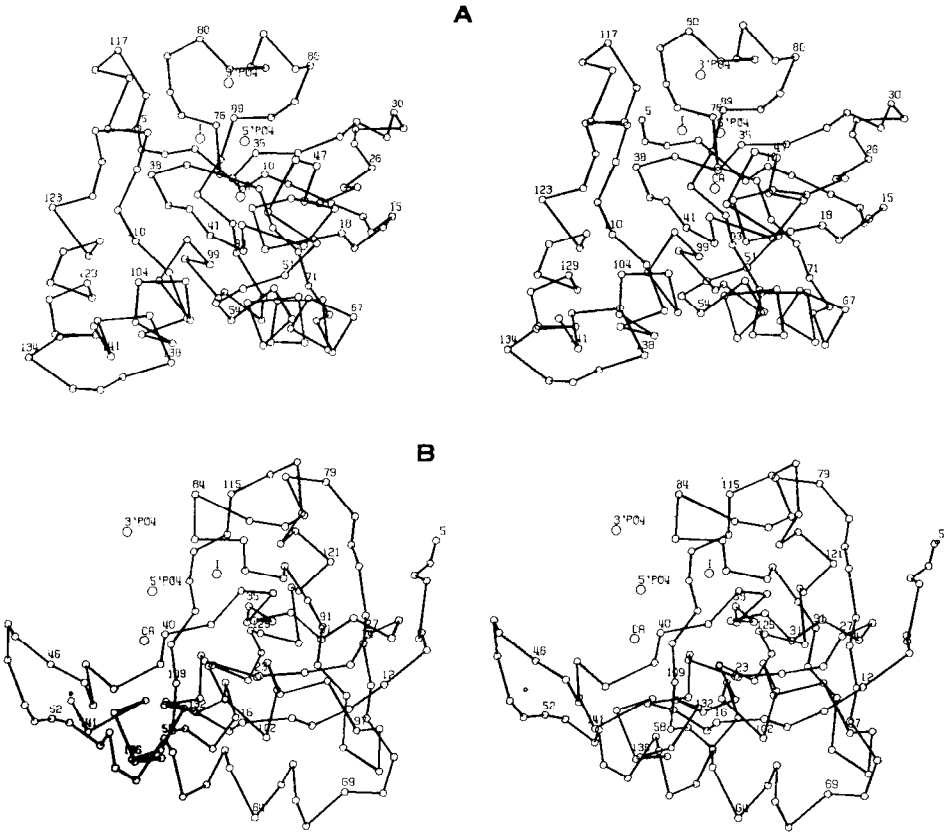


FIG. 4. Stereo views of the peptide chain approximating the same orientations shown in Fig. 3. The positions of the two phosphates, the $Ca^{2+}$, and the iodine (or 5-methyl) are indicated.

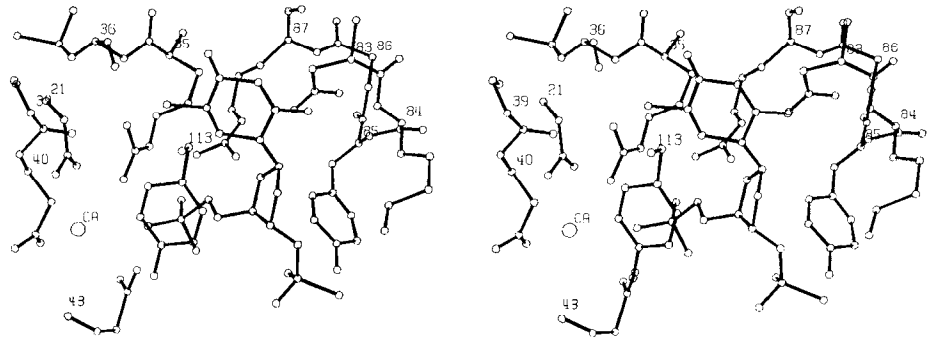


FIG. 5. A schematized stereo view of the inhibitor in the binding site. The residues are labeled at the $\alpha$ carbons.

bond broken upon hydrolysis is the P—O bond between deoxyribose and the 5′-phosphate group, as shown in a study of the simple substrate, the *p*-nitrophenyl ester of thymidine 5′-phosphate, where hydrolysis yields *p*-nitrophenylphosphate and deoxythymidine (46). If the structure we present here of the nuclease-inhibitor complex is a reasonable approximation to that of the enzyme-substrate complex, then the point of enzymatic attack, the 5′-C-O-P ester linkage, is seen to be quite precisely positioned and rigidly held by the nuclease. This positioning and holding is accomplished, as expected, by particular side chains from the enzyme cooperatively interacting with particular parts of the inhibitor (substrate), but for at least some of these side chains, *e.g.* the guanidinium moieties, the precision and rigidity of the interaction is buttressed by the additional hydrogen bonding of these side chains to other more remote parts of the enzyme. This impression of precision and organization in the design of the active site area is in splendid accord with the work of Chaiken and Anfinsen reported in the accompanying manuscript (44). With synthetic variants of the nuclease, they have demonstrated that the enzymatic activity and, in some cases, the entire structure of the nuclease is very sensitive to variations in the size and charge of Arg-35, Asp-21, Asp-40, and Glu-43. It is pertinent to note here that the studies of Mikulski *et al.* (47) of the hydrolysis of a number of dinucleotides by this nuclease are quite consistent with the structure of the nuclease-inhibitor complex presented here.

We do not at this time feel disposed to propose a mechanism for the nuclease action, but we can make some preliminary suggestions. It would seem possible that, in addition to their binding function, the positively charged guanidinium ions also serve

to polarize the 5′-phosphate group. It is also likely that the phenolic or more probably the phenolate (43) group of Tyr-113 is close enough to the 5′-C-O-P ester linkage to play a role in hydrolysis. As already noted, the calcium ion is too far from the 5′-phosphate group to have any significant interaction with it. The expected distance for a direct interaction should be of the order of the 3.44 A found in the crystal structure of calcium thymidylate (48). Mildvan (49) has suggested the possibility of a bridging water molecule between the calcium ion and the 5′-phosphate group, an interesting possibility that receives some support from the electron density found in this area in the single-iodine derivative map. However, this feature must be regarded as experimentally extremely tentative. At present the precise role of the calcium ion is obscure, since, in addition, there is another unresolved question that bears on the mechanism of action. The nuclease requires rather high [$Ca^{2+}$] for optimal activity (30), and the binding of pdTp to nuclease in solution appears to involve 2 moles of $Ca^{2+}$ per mole of inhibitor and enzyme (50). We observe only 1 $Ca^{2+}$ ion in the electron density map, a finding which is not necessarily inconsistent with the solution work. It may merely reflect the different conditions prevailing in the different studies (1, 43, 50).

The observation that a Lys-71 from another molecule in the unit cell interacts with the 5′-phosphate group is somewhat dismaying, but since this interaction, when compared with the sum of the other forces impinging on the pdTp, is relatively weak, it probably does not significantly distort the structure from that which would occur in the isolated molecule.

We shall not undertake here to make a detailed correlation of this structure with the solution studies of the nuclease since this should, in any event, be a collaborative effort among the laboratories involved and is best postponed until the structures of the uninhibited nuclease and nuclease-T (39) become available. Some particular aspects of this question are discussed in the accompanying manuscripts (35, 44), in two chapters of the third edition of *The Enzymes* (42, 43), and in an article by Cuatrecasas (51).

## REFERENCES


1. Arnone, A., Bier, C. J., Cotton, F. A., Hazen, E. E., Jr., Richardson, D. C., and Richardson, J. S., *Proc. Nat. Acad. Sci. U. S. A.*, **64**, 420 (1969).
2. Taniuchi, H., Anfinsen, C. B., and Sodja, A., *J. Biol. Chem.*, **242**, 4752 (1967).
3. Cotton, F. A., Hazen, E. E., Jr., and Richardson, D. C., *J. Biol. Chem.*, **241**, 4389 (1966).
4. North, A. C. T., Phillips, D. C., and Mathews, F. S., *Acta Cryst. Ser. A*, **24**, 351 (1968).
5. Wyckoff, H. W., Tsernoglou, D., Hanson, A. W., Knox, J. R., Lee, B., and Richards, F. M., *J. Biol. Chem.*, **245**, 305 (1970).
6. Kraut, J., Sieker, L. C., High, D. F., and Freer, S. T., *Proc. Nat. Acad. Sci. U. S. A.*, **48**, 1417 (1962).
7. Dickerson, R. E., Kopka, M. L., Varnum, J. C., and Weinzierl, J. R., *Acta Cryst.*, **23**, 511 (1967).
8. Matthews, B. W., *Acta Cryst.*, **20**, 230 (1966).
9. Crick, F. H. C., and Magdoff, B. S., *Acta Cryst.*, **9**, 901 (1956).
10. MacGillavry, C. H., and Rieck, G. D. (Editors), *International tables for x-ray crystallography, Vol. III*, Kynoch Press, Birmingham, England, 1962, 201.
11. Cromer, D. T., *Acta Cryst.*, **18**, 17 (1965).
12. Rossmann, M. G., *Acta Cryst.*, **14**, 383 (1961).
13. Blow, D. M., *Proc. Roy. Soc. Ser. A*, **247**, 302 (1958).
14. Blow, D. M., and Crick, F. H. C., *Acta Cryst.*, **12**, 794 (1959).
15. Dickerson, R. E., Kendrew, J. C., and Strandberg, B. E., *Acta Cryst.*, **14**, 1188 (1961).
16. North, A. C. T., *Acta Cryst.*, **18**, 212 (1965).
17. Matthews, B. W., *Acta Cryst.*, **20**, 82 (1966).
18. Kartha, G., *Acta Cryst.*, **19**, 883 (1965).
19. Busing, W. R., Martin, K. O., and Levy, H. A., *Oak Ridge Nat. Lab.*, OR-TM-305 (1962).
20. Lipscomb, W. N., Coppola, J. C., Hartsuck, J. A., Ludwig, M. L., Muirhead, H., Searl, J., and Steitz, T. A., *J. Mol. Biol.*, **19**, 423 (1966).
21. Muirhead, H., Cox, J. M., Mazzarella, L., and Perutz, M. F., *J. Mol. Biol.*, **28**, 117 (1967).
22. Adams, M. J., Haas, D. J., Jeffery, B. A., McPherson, A., Jr., Mermall, H. L., Rossmann, M. G., Schevitz, R. W., and Wonacott, A. J., *J. Mol. Biol.*, **41**, 159 (1969).
23. Cullis, A. F., Muirhead, H., Perutz, M. F., Rossmann, M. G., and North, A. C. T., *Proc. Roy. Soc. Ser. A*, **265**, 15 (1961).
24. Herriott, J. R., Sieker, L. C., Jensen, L. H., and Lovenberg, W., *J. Mol. Biol.*, **50**, 391 (1970).
25. Blow, D. M., *Proc. Roy. Soc. Ser. A*, **247**, 302 (1958).
26. Phillips, D. C., in R. Brill and R. Mason (Editors), *Advances in structure research by diffraction methods, Vol. II*, John Wiley and Sons, Inc., New York, 1966, p. 75.
27. North, A. C. T., and Phillips, D. C., *Prog. Biophys. Mol. Biol.*, **19**, 1 (1969).
28. Pauling, L., *The nature of the chemical bond*, Ed. 3, Cornell University Press, Ithaca, New York, 1960, p. 514.
29. Pauling, L., *The nature of the chemical bond*, Ed. 3, Cornell University Press, Ithaca, New York, 1960, p. 260.
30. Cuatrecasas, P., Fuchs, S., and Anfinsen, C. B., *J. Biol. Chem.*, **242**, 1541 (1967).
31. Cuatrecasas, P., Taniuchi, H., and Anfinsen, C. B., *Brookhaven Symp. Biol.*, **21**, 172 (1968).
32. Schechter, A. N., Morávek, L., and Anfinsen, C. B., *J. Biol. Chem.*, **244**, 4981 (1969).
33. Cusumano, C. L., Taniuchi, H., and Anfinsen, C. B., *J. Biol. Chem.*, **243**, 4769 (1968).
34. Taniuchi, H., Cusumano, C. L., Anfinsen, C. B., and Cone, J. L., *J. Biol. Chem.*, **243**, 4775 (1968).
35. Taniuchi, H., and Anfinsen, C. B., *J. Biol. Chem.*, **246**, 2291 (1971).
36. Richards, F. M., *J. Mol. Biol.*, **37**, 225 (1968).
37. Johnson, C. K., *Oak Ridge Nat. Lab.*, ORNL-3794 (1965).
38. Taniuchi, H., Anfinsen, C. B., and Sodja, A., *Proc. Nat. Acad. Sci. U. S. A.*, **58**, 1235 (1967).
39. Taniuchi, H., and Anfinsen, C. B., *J. Biol. Chem.*, **243**, 4778 (1968).
40. Taniuchi, H., Morávek, L., and Anfinsen, C. B., *J. Biol. Chem.*, **244**, 4600 (1969).
41. Cuatrecasas, P., Wilchek, M., and Anfinsen, C. B., *Science*, **162**, 1491 (1968).
42. Anfinsen, C. B., Cuatrecasas, P., and Taniuchi, H., in P. D. Boyer (Editor), *The enzymes, Vol. IV*, Ed. 3, Academic Press, New York, in press.
43. Cotton, F. A., and Hazen, E. E., Jr., in P. D. Boyer (Editor), *The enzymes, Vol. IV*, Ed. 3, Academic Press, New York, in press.
44. Chaiken, I. M., and Anfinsen, C. B., *J. Biol. Chem.*, **246**, 2285 (1971).
45. Cuatrecasas, P., Fuchs, S., and Anfinsen, C. B., *J. Biol. Chem.*, **243**, 4787 (1968).
46. Cuatrecasas, P., Wilchek, M., and Anfinsen, C. B., *Biochemistry*, **8**, 2277 (1969).
47. Mikulski, A. J., Sulkowski, E., Stasiuk, L., and Laskowski, M., Sr., *J. Biol. Chem.*, **244**, 6559 (1969).
48. Trueblood, K. W., Horn, P., and Luzzati, V., *Acta Cryst.*, **14**, 965 (1961).
49. Mildvan, A. S., in P. D. Boyer (Editor), *The enzymes, Vol. I*, Ed. 3, Academic Press, New York, in press.
50. Cuatrecasas, P., Fuchs, S., and Anfinsen, C. B., *J. Biol. Chem.*, **242**, 3063 (1967).
51. Cuatrecasas, P., *J. Biol. Chem.*, **245**, 574 (1970).