# High Resolution(2.3Å) Analysis of the Crystal Structure of Bonito Ferrocytochrome c

A Doctoral Thesis

Submitted by

Takashi Yamane

Faculty of Science

Osaka University

1975

Acknowledgements

The work of this thesis was carried out under the guidance of Professor Masao Kakudo at Institute for Protein Research, Osaka University. The author is greatly indebted to Professor Masao Kakudo for his continuing guidance and encouragement. The author wishes to express his gratitude to Professor Tamaichi Ashida at Faculty of Engineering, Nagoya University, and Dr. Nobuo Tanaka for many helpful discussions and suggestions throughout the work. The author wishes to express his appreciation to Dr. Tomitake Tsukihara for his valuable help. Thanks are given to Miss Sachiko Bando for her kind assistance. The author wishes to express his thanks to Dr. Tsunehiro Takano at Medical Research Council Laboratory of Molecular Biology, Cambridge, Dr. Tatzuo Ueki at Faculty of Engineering Science, Osaka University, and Dr. Akio Sugihara at Osaka Municipal Technical Research Institute, for their available suggestions. The author would like to acknowledge Professor Nobuya Iwamoto at Welding Research Institute, Osaka University for his kind encouragement. Finally the author wishes to thank all the members of Kakudo laboratory for their friendship.

CONTENTS

# I INTRODUCTION

A living organism necessitates energy to live. One of the way to obtain the energy is respiration. The respiratory chain is composed of NADH dehydrogenase, succinate dehydrogenase, cytochromes, CoQ and non-heme iron protein. Though the function of the respiratory chain is to produce adenosine triphosphate(ATP) by conjugating with phosphorylation system, its most important character is in making an electron transfer(Fig.1). The electron transfer system concerned in the cellular respiration of animals, molds and so on, localizes in a cell mitochondria. Though it has not been made clear where and how the components of the system distribute in the mitochondria, it has been well established for an electron to be transferred in the respiratory chain(Fig.1)(1). The flow of an electron is:

Substrate ⟶ Dehydrogenase ⟶ Cytochromes
⟶ Cytochrome oxidase ⟶ $O_2$.

Each component of cytochromes is repeatedly oxidized and reduced in the cellular respiration.

According to the structures of hemes, cytochromes are classified into A, B and C types. A cytochrome C is a protein containing the heme c(Fig.2), which is covalently linked to the peptide chain by two thioether bonds of cysteins. Cytochrome C's are widely distributed in living organisms. Biological and chemical studies about cytochrome C's have been advanced, because only the protein in the cytochromes is soluble in water and its purification

is comparatively easy. Three absorption maxima are observed in 650mμ to 450mμ wavelength range in ferrocytochromes. They are called α, β and γ absorption bands from longer wavelength. The α band maximum is characteristic for each cytochromes. The α band of cytochrome C is in vicinity of 550mμ. Cytochrome C's of animals, molds, etc. are called (Mammalian-type) cytochrome c's. One of the characters of cytochrome c is to give the symmetrical α absorption band at 549-550mμ. Cytochrome c's have many common characters, and have similar amino acid sequences with each other(2).

Cytochrome c of vertebrates (except for bonito and tuna cytochrome c's) is constituted of 104 amino acid residues (bonito and tuna cytochrome c's are 103 residues long), of which the N-terminal(Gly1) is acetylated. On the other hand, the peptide chains of non-vertebrate species are longer than 104 residues. They have extra amino acids at the first position in place of the acetyl group.

The sequence of bonito cytochrome c has been determined by Nakayama, Titani and Narita(Fig.3)(3). According to their result, the lysyl residue in the position 99-100 is substituted, differing from other cytochrome c's. The sequence of bonito cytochrome c has a strong resemblance to that of tuna cytochrome c. Only Glu61 and Asn62 in bonito are replaced by Asn and Asp in tuna, respectively. On the other hand, there are 17 substitutions of amino residues between the sequences of bonito and horse

cytochrome c's.

The amino residues of 32 positions in the sequences are invariant, and the positions substituted by the amino acids of which physical and chemical characters are very similar, so called 'conservative position', amount to 22 (2). In view of the distribution of such invariant, conservative and variant positions in the peptide chain, mammalian-type cytochrome c's might have essentially the same conformation, even though they differ in fine structures.

The distinct differences of the structures between the ferrous and ferric proteins were shown in the studies on thermal stability(4), affinity for cation exchange resin(5), adsorption on Kaolin(6), resistance to proteolytic digestion(7), UV spectra(8), NMR(9), Moessbauer effect (10), CD(11,12) and ORD(13). The differences of physical properties, and chemical and enzymatic reactivities between ferro- and ferricytochrome c are summarized in Table 1. The differences would be attributed to the conformational change of the peptide chain rather than to the change of the charge of the heme iron($Fe^{3+} \rightleftharpoons Fe^{2+}$). Besides it was reported by the decomposability difference against the proteinase(14) that each of the changes may occur near the molecular surface, and that each of the changes is subtle. It is considered that such subtle changes of the conformation between the two oxidation states are closely connected with the electron transfer between cytochromes.

In order to elucidate the electron transfer mechanism between cytochromes, it is necessary to determine the three dimensional structures of the two oxidation states. Dickerson's group analyzed the crystal structures of horse ferricytochrome c at 2.8Å resolution(15) and tuna ferrocytochrome c at 2.45Å resolution(16). Kraut's group analyzed the crystal structure of ferricytochrome $c_2$ at 2Å resolution(17). Cytochrome $c_2$(R.rubrum)(18) together with cytochrome c-551(P.fluorescens)(19) and cytochrome $c_3$ (D.vulgaris)(20) are bacterial cytochrome C's of which amino acid sequences were determined. The sequence of cytochrome $c_2$ is different from those of cytochrome c-551 and cytochrome $c_3$, and rather resemble to mammalian-type cytochrome c. It was made clear by the x-ray analysis that the conformation of the principal part of cytochrome $c_2$ is essentially the same as that of the horse cytochrome c(17,21).

The structure analysis of bonito ferrocytochrome c has been carried out step by step in Kakudo's group. At first 6Å resolution analysis(22), then 4Å(23), and recently 2.3Å analysis(24,25) were carried out. In this thesis the structure analysis of bonito ferrocytochrome c at 2.3Å resolution is reported.

Table 1 Summary of physical and chemical properties between two oxidation states

| | Reference | Ferrous form | Ferric form |
|---|---|---|---|
| thermal stability | 4 | stable | less stable |
| affinity for cation exchange resin | 5 | weak | strong |
| adsorption by kaolin | 6 | slow | ready |
| digestion by proteinase | 7 | slow | ready |
| ratio of UV extinction at pH 12 to that at pH 7 | 8 | low | high |
| resonance of the Met80 methyl group | 9 | +3.3ppm | +23.4ppm |
| distance between the heme ring methyl $b_1$ and the aromatic group | 9 | close | distant |
| contribution to coordination bonds of 4s electrons in heme iron | 10 | 35% | 10% |
| Cotton effect arising from $\gamma$-band | 11,12 | negative | positive |
| principal deflection of the Cotton effect | 13 | negative | positive |

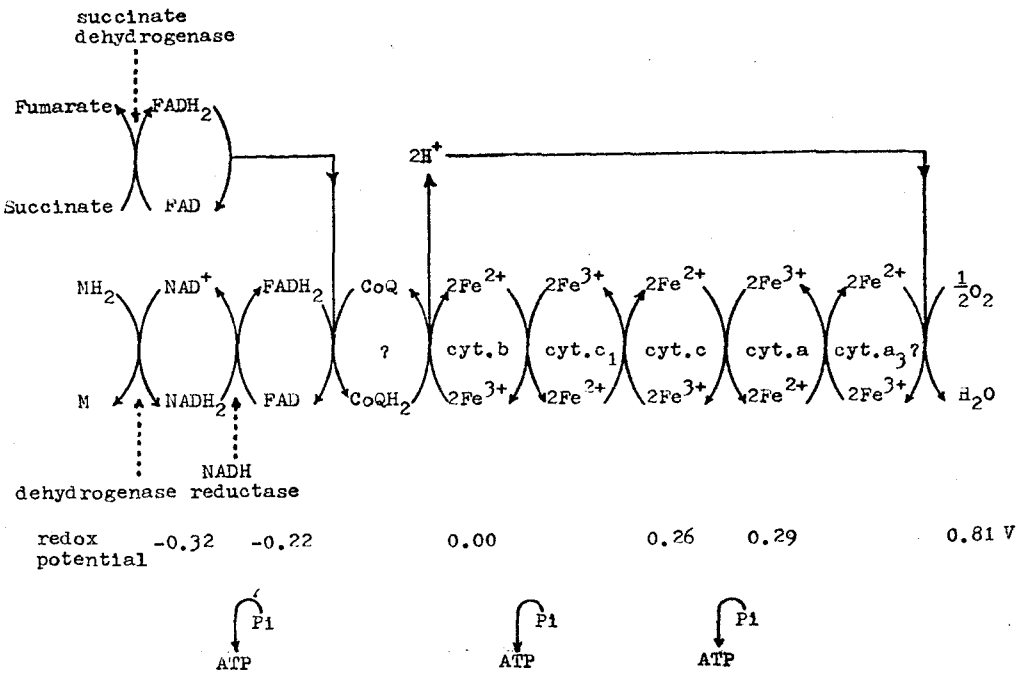



Fig.1 Electron transfer system conjugated with oxidative phospholylation in cellular respiration. An electron is transferred from a substrate exhibiting low redox potential to oxygen having higher redox potential through the respiratory chain.

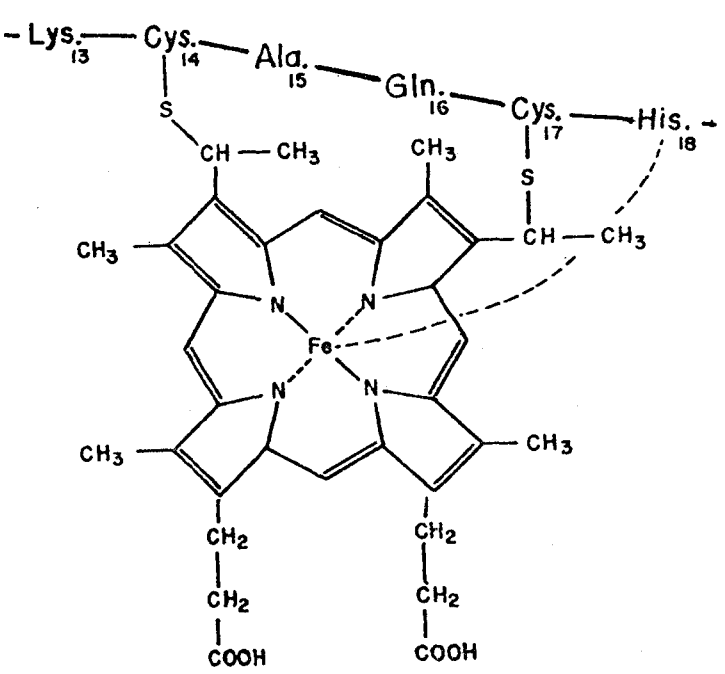



Fig.2 Chemical structure of heme c.

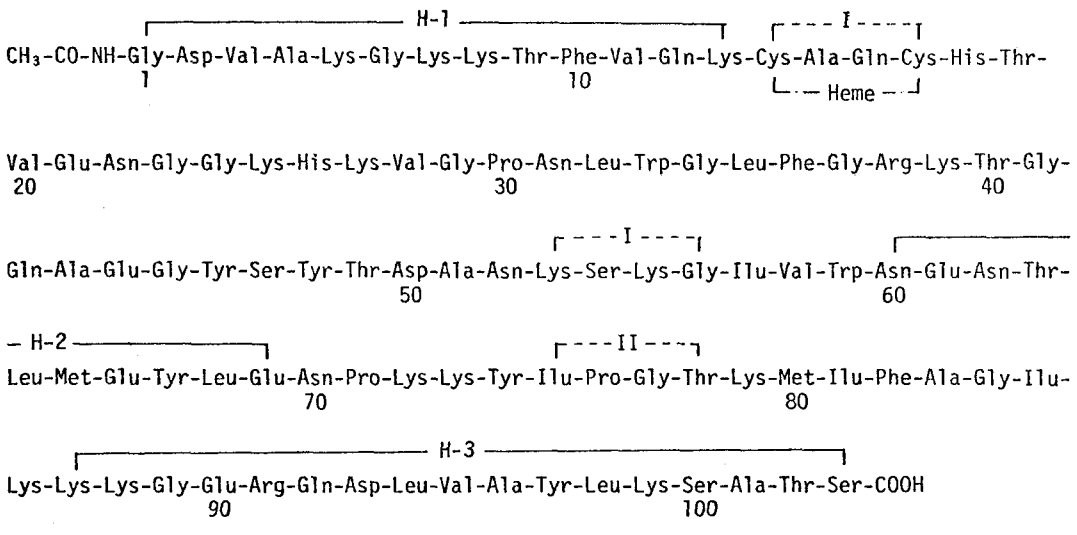



Fig.3 Amino acid sequence determined by Nakayama et al. (3). Solid bars indicate regions of α helix, and dashed bars indicate regions of $3_{10}$ bend in bonito ferrocytochrome c.

## II EXPERIMENTAL

### i Extraction and purification

The protein was extracted from bonito hearts according to the method of Margoliash and Walasek(26). The main steps of extraction and purification are outlined in Table 2. To remove salts from the protein solution, a sephadex G-25 column was used instead of dialyzing the solution using cellophane tubes. About 4 grams of the highly purified protein was obtained from 20-25kg of bonito hearts.

Table 2 Flow diagram of extraction and purification

Frozen bonito hearts

↓

Remove bonding tissues.
Mince.
Homogenize in a waring blender for 5 sec. with threefold volumes of 0.3% $Al_2(SO_4)_3$.
Add 5N $H_2SO_4$.
Stand for 2 hours at below 10°C.
Filter using a suction bottle.

↓

Extract

Add aqueous $NH_3$ to pH 8.2-8.5.
Filter with the aid of 10g celite per liter of the extract.
Add acetic acid to pH 7.5.
Pass through a column of Amberlite IRC-50.
Scrape off the red part of the resin and wash with ten fold of 0.02M $(NH_4)_3PO_4$ buffer(APB, PH 7.5).

Elute with 0.05M APB containing 0.5M NaCl(pH 7.5).

Add minimal amount of $K_3Fe(CN)_6$.

Add solid $(NH_4)_2SO_4$ to 50% saturation.

Centrifuge at 8,000 r.p.m. for 20 min.

Supernatant solution

Desalt by passing through a column of Sephadex G-25.

Wash with 0.02M APB.

Elute with minimal amount of 0.05M APB containing 0.5M NaCl.

Desalt by passing through a column of Sephadex G-25.

Pass through a column of Amberite IRC-50(pH 7.5).

Charged column

Develop with 0.2 or 0.25M APB(pH 7.5).

Cytochrome c fraction

Concentrate by the same method as the first chromatpgraphy, and elute with 10% saturated $(NH_4)_2SO_4$ solution(pH 8.0).

Concentrated cytochrome c solution

ii Crystallization

The purity of cytochrome c is shown by the index of

$OD_{550m\mu}^{reduced} / OD_{280m\mu}$,

where $OD_{\lambda}$ is the optical density at wavelength $\lambda$. The absorption at 550mμ is caused by heme c, and that at 280mμ by aromatic ring. The solution of bonito cytochrome c whose index of purity is higher than 1.0 was used for crystallization. The crystals of ferricytochrome c could be easily given by adding finely powdered ammonium sulfate to 1-5% solution of the protein, till a faint haze was detected. Ferricytochrome c crystallized in 1-2 weeks from the solution.

Crystallization of ferrocytochrome c was more difficult than that of the ferric protein. The main steps for crystallization is given in Table 3. The optimum concentration of ammonium sulfate was 87% saturation. The pH was adjusted to 7.5-8.0. The solution was left at room temperature. Small, brightly red crystals appeared in the amorphous region a few days later. They grew gradually and reached to the size of 1-2mm after 2 weeks(Fig.4).

Table 3 Flow diagram of crystallization

10 ml of 7 % ferricytochrome c solution

↓

Add finely powdered $(NH_4)_2SO_4$ to 60 % saturation.

Add 5 fold molar excess of ascorbic acid.

Add aqueous $NH_3$ to bring the pH to 8.2-8.5.

Add finely powdered $(NH_4)_2SO_4$ in small portions until a faint turbidity appears.

Centrifuge at 8,000 r.p.m. for 20 min.

Supernatant solution

↓

Add finely powdered $(NH_4)_2SO_4$ in small portions until a faint haze appears.

Cool the solution in iced water and dessolve the amorphous turbidity.

Repeat the above two procedures till uniform, cotton-like amorphous turbidity appears.

Keep at room temperature without any mechanical shock.

Large crystals after 1 to 2 weeks

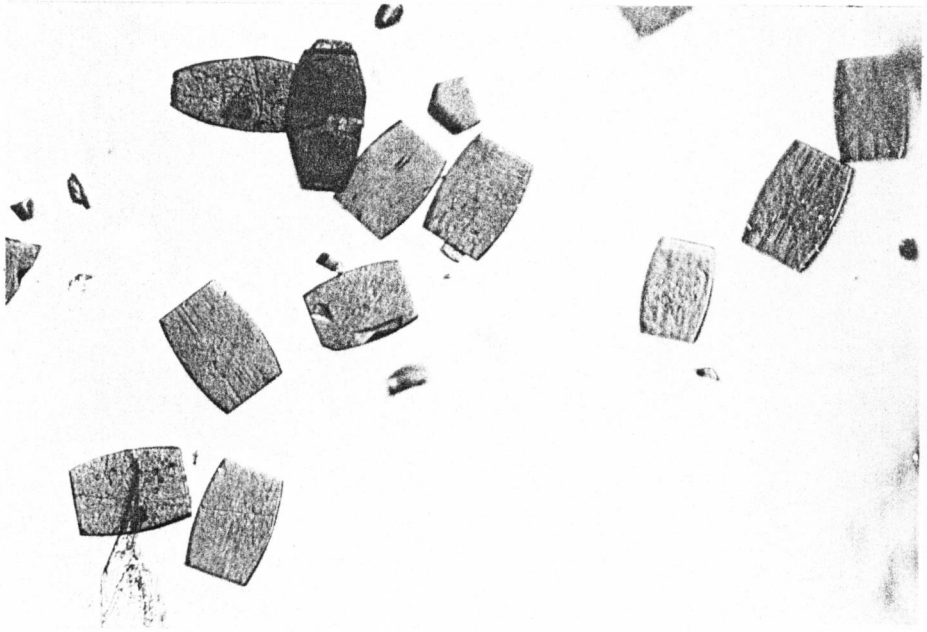


Fig.4 Crystal shape.

iii Oxidation and reduction of crystals

The crystals of bonito ferrocytochrome c could be oxidized without them destroying. The ferrous crystals were oxidized by oxygen in the 87% ammonium sulfate solution(pH 5.9), though it took about 2 months till the degree of oxidation became higher than 50%(auto-oxidation). The oxidation was also carried out with potassium ferricyanide. When a few drops of 0.5M potassium ferricyanide solution were added to the solution containing ferrous crystals, more than 90% of ferrous protein were oxidized a week later. The crystals came to dark red from bright red by oxidation, but any cracks could not be found. The ferric crystals obtained as above were able to be reduced with ascorbic acid reversibly. The precession photograph of the crystal revealed that the ferric crystal is isomorphous to the ferrous one. Therefore the auto-oxidized crystal was studied at 3Å resolution with difference Fourier synthesis by Tsukihara et al.(27). On the other hand, the bonito ferric crystal($P4_3$) were broken into pieces by reduction with ascorbic acid.

The degree of oxidation was estimated according to the following formula after the measurement of optical densities at 550mμ,

$$(OD^r - OD^s) / OD^r(1 - OD^o/OD^r)$$

where $OD^s$ is the optical density of the sample solution, $OD^r$ is that after complete reduction with sodium dithionite, and $OD^o$ is that of the completely oxidized sample solution. The values of $OD^r/OD^o$ was determined

as 3.605 beforehand(22).

iv Preparation of isomorphous heavy atom derivatives

Two good isomorphous heavy atom derivatives are at least necessary to solve a protein structure. Two isomorphous derivatives replaced by $K_3UO_2F_5$(U derivative) and $K_2PtCl_4$(Pt derivative), were used for the phase determination of the 4Å analysis(23). The Pt derivative was, however, shown to be rather poor in both isomorphism and crystallinity in the analysis(23). Therefore new derivatives were asked for, to perform the higher resolution analysis.

It was made clear that the preparation of isomorphous derivatives of bonito ferrocytochrome c was very difficult, because of the following characters of the crystal: 1) The water content of the crystal is small as compared with that of other proteins. This shows that the packing of molecules are fairly compact and that it is hard for heavy atoms to diffuse into the crystal; 2) The ferrous form is easily changed to the ferric one by auto-oxidation; 3) Heavy atoms which interact the sulfur atoms in the thioether linkages of Cys14 and Cys17 do not give the isomorphism.

Heavy atom derivatives were prepared at room temperature by soaking the pre-grown ferrous crystals in the ammonium sulfate solution(87%(w/w)) of the heavy atom reagent. Excess of ascorbic acid was added to the solution, and the pH was adjusted to 6.0. The atmosphere in

the soaking vessel was replaced with nitrogen gas to prevent the auto-oxidation of the crystals. The soaking period was made to be as short as possible for the same reason. To find out if any substitution had taken place, the x-ray diffraction profiles of the three principal axes of the soaked crystals were taken by the diffractometer. These profiles were compared with those of the native crystal. Significant changes in the intensities with little changes in the cell parameters and the crystallinity, were asked for as a good isomorphous derivative. The derivatives which have the change of the cell parameter more than 0.5% were rejected. Generally the diffraction profiles of the heavy atom derivatives spread out as compared with those of native crystals. The higher the resolution is, the narrower the scanning of the diffraction profile should be. In this bonito protein, the upper limit of the $\omega$ scanning range was 1.5° at 2.3Å resolution, though it was 2.4° at 4Å resolution.

The x-ray intensity data of the derivatives examined as above were collected within 4Å resolution limit, and the derivatives were checked by the difference Fourier synthesis using the phases determined in the 4Å analysis. Of about 15 kinds of newly tested reagents, several derivatives were obtained. They are;

$UO_2(CH_3COO)_2$: The solution was shielded from light to protect the reagent from decomposition. The concentration of the reagent was varied over a wide range, and the soaking period was from a day to a week. The x-ray

diffraction profiles of the derivative were significantly broadened with soaking, so that this could not be used for the analysis of the higher resolution.

$CdI_2$: At a low concentration the occupancies of heavy atoms were very poor. When the concentration was increased, the difference Fourier map was confused by trivial electron density humps, which showed the denaturation of the protein and/or too many atomic sites with low occupancies. Therefore it was impossible to refine the parameters of the derivative.

$K_3IrCl_6$: The changes of intensities were little, even though crystals were kept in the saturated soaking solution for a long time.

$(CH_3)_2SnCl_2$: The diffraction intensity changed sufficiently for a short soaking period and the crystallinity was good. This was the first case that an organotin compound was used in the structure analysis of protein crystals.

At last three heavy atom derivatives were found to be suitable for the structure determination. The soaking conditions for each derivative are;

| | | |
|---|---|---|
| U derivative | 45mM | 3days |
| Sn derivative | 2.3mM | 4days |
| Ir derivative | 19mM | 5days. |

For the U derivative, crystals were reduced for half a day with ascorbic acid after the end of soaking, because the reagent reacts with ascorbic acid.

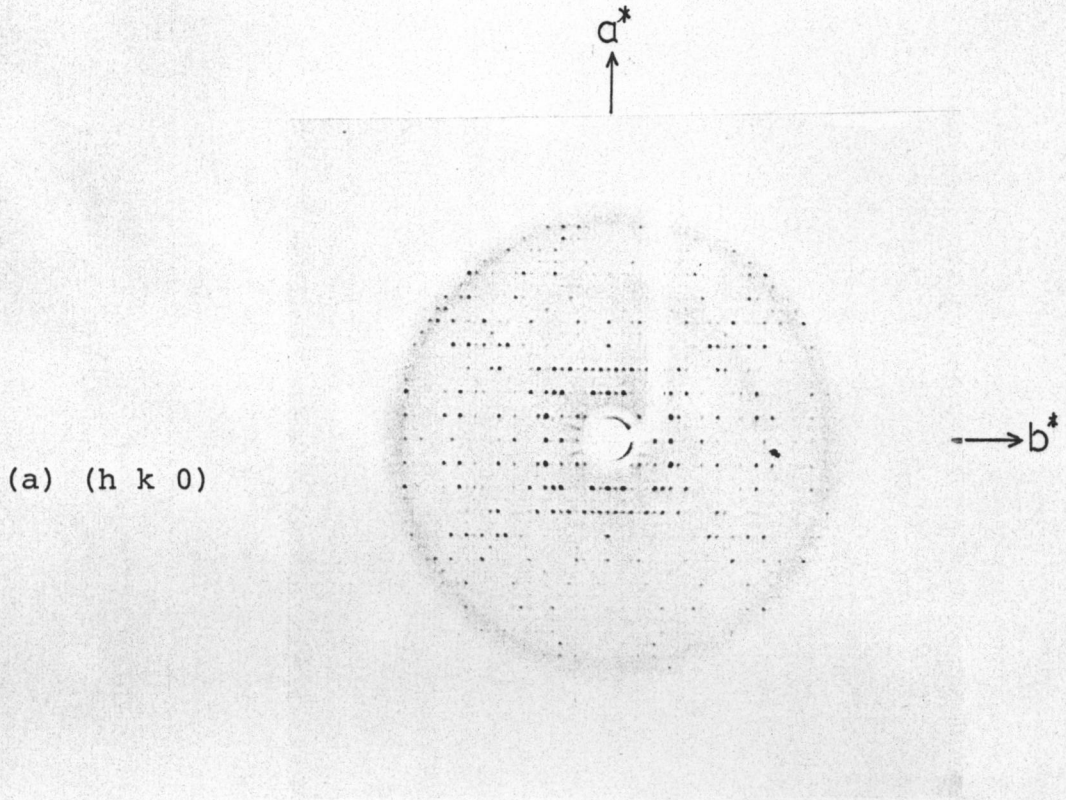
v Crystallographic data

The crystal of bonito ferrocytochrome c is orthorhombic, space group being $P2_12_12_1$. The density was estimated to be 1.38g.cm$^{-3}$ by floatation method. The reflections with h odd gave quite weak intensities at small Bragg angles(Fig.5). The value of the ratio $\langle|F|_{h=odd}\rangle$ / $\langle|F|_{h=even}\rangle$ is 0.23 for the 6Å resolution data. If the reflections with h odd of the original space group $P2_12_12_1$ were ignored, the space group would become $P22_12_1$ with just a half of the a cell constant. The general equivalent positions of $P22_12_1$ are four-fold. A cytochrome c molecule is composed of one polypeptide chain and a heme group, and can not be located on the special position. Therefore four protein molecules are included in the half cell($P22_12_1$, Z=4). This indicates that eight molecules are included in the original cell ($P2_12_12_1$, Z=8), and that two molecules in an asymmetric unit are approximately related by a pseudo 2-fold axis parallel to the a axis(Fig.6).

The fractional volume of the protein in the ferrous crystal comes to about 63%, on the basis of two molecules in an asymmetric unit. The values is higher than those of many other globular proteins. The maximum frequency is about 57% quated by Matthews(28). Horse ferri- and tuna ferrocytochrome c's were crystallized by Dickerson's group(29,30). The fractional volumes of the horse ferric and tuna ferrous crystals are 45% and 53% respectively(16). Because the volume per molecule of bonito ferrous crystal

is only two thirds of that of horse ferric one, the molecules of the ferrous crystal are packed in a more closer manner than in the ferric one. As mentioned before, this fact explains the difficulty in obtaining the isomorphous heavy atom replaced crystals.

The precession picture of horse ferric crystal taken by Dickerson et al. is shown in Fig.7, together with that of bonito ferric one. The crystals of the two ferricytochrome c's are isomorphous with each other, showing that the structures between bonito and horse proteins resemble closely to each other in the ferric form, though the sequence of bonito cytochrome c is different from that of horse one at 17 positions. The tuna ferrous crystal is not isomorphous with the bonito ferrous one. All the crystallographic data of these cytochrome c crystals are listed in Table 4.



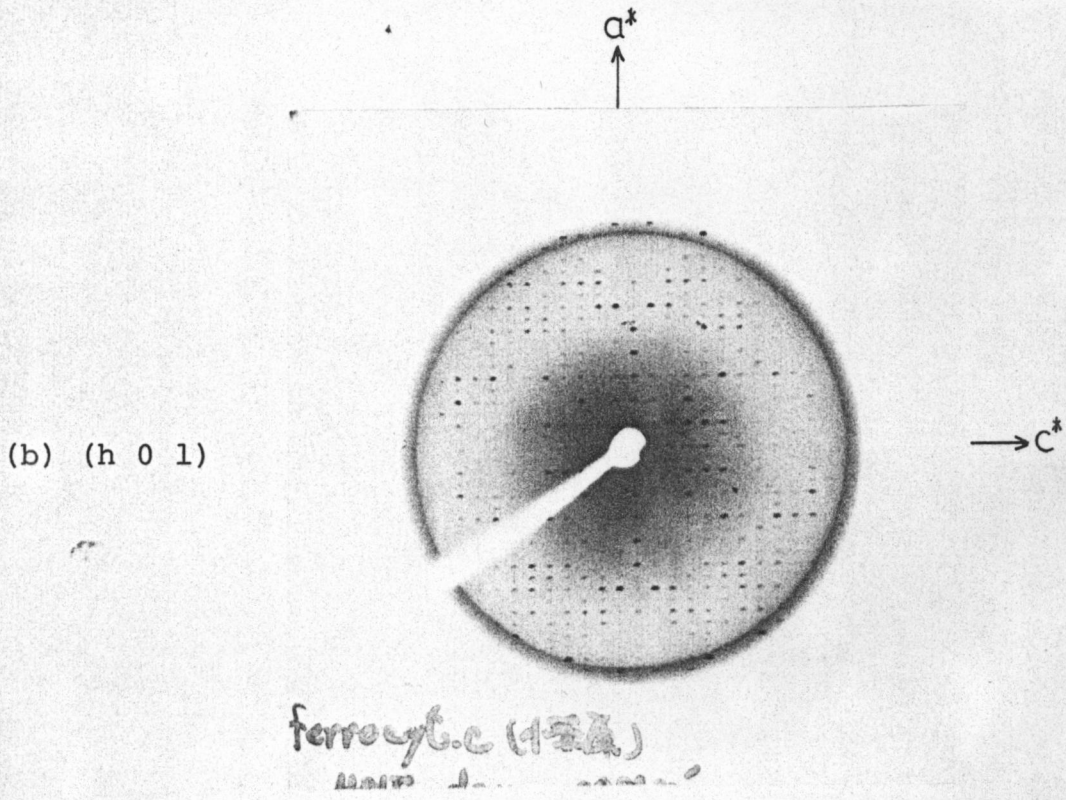




Fig.5 Precession photographs showing pseudosymmetry.

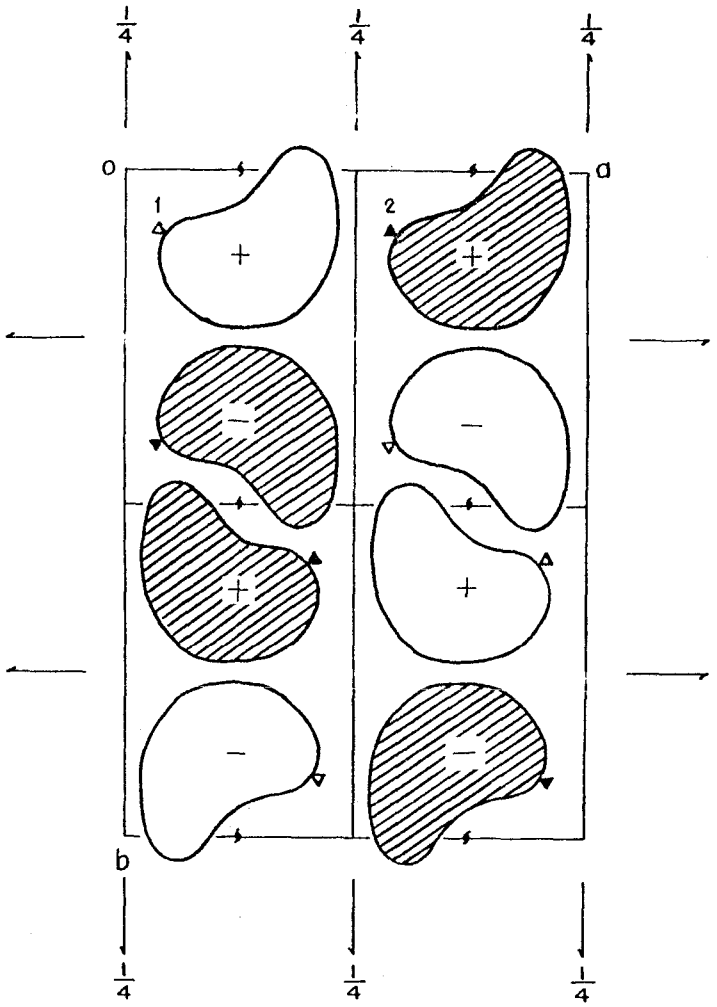
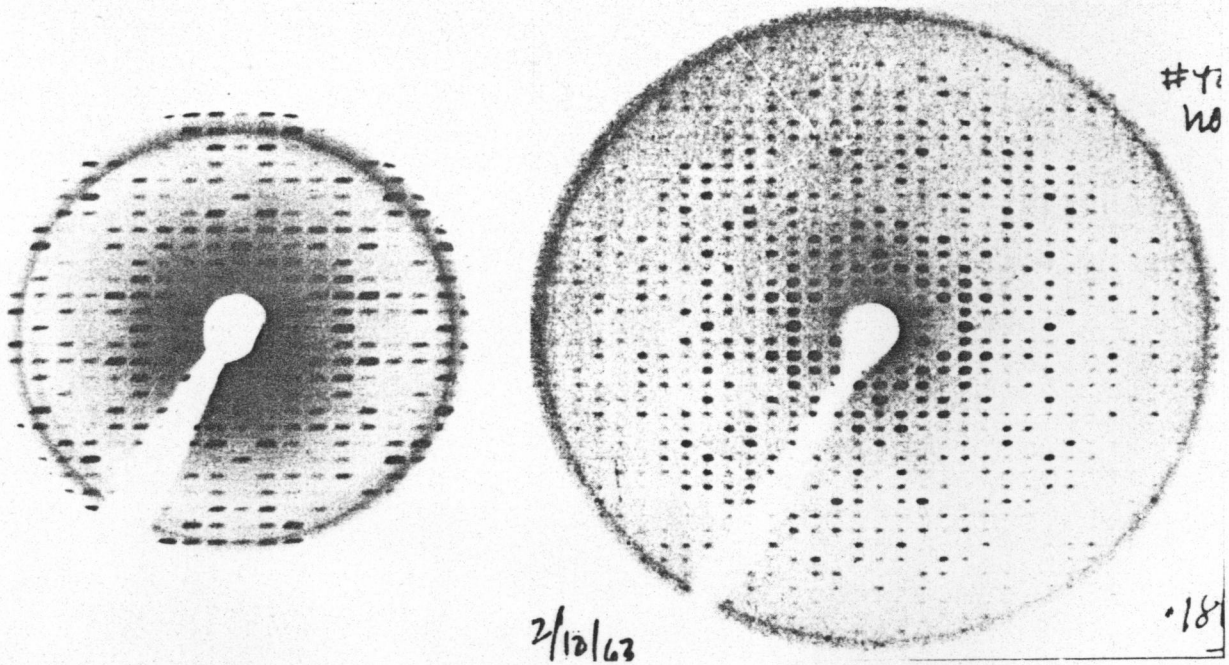



Fig.6 Packing scheme of molecules in the unit cell. (-) molecule is approximately related to (+) molecule by pseudo 2-fold axis parallel to a axis.




(Bonito) (Horse)

Fig.7 Precession photographs of bonito and horse ferricytochrome c's. (Tetragonal form, $P4_3$).

Table 4 Crystallographic data of cytochromes c

(a) Bonito ferrocytochrome c

$C_{544}H_{844}O_{153}N_{145}S_4Fe$

| | | | | |
|---|---|---|---|---|
| Molecular weight | 12,034 | | | |
| Space group | $P2_12_12_1$ (orthorhombic) | | | |
| | Native | U derivative | Sn derivative | Ir derivative |
| a ( Å ) | 57.68 | 57.73 | 57.61 | 57.58 |
| b ( Å ) | 84.58 | 84.54 | 84.61 | 84.52 |
| c ( Å ) | 37.83 | 37.82 | 37.76 | 37.84 |
| V ( $Å^3$ ) | 181,000 | | | |
| Observed density | $\rho_o$ = 1.38 $g.cm^{-3}$ | | | |
| Number of molecules in unit cell | Z = 8 | | | |
| Fractional volume of protein | 63% | | | |

(b) Other cytochromes

| | Bonito ferri | Horse ferri | Tuna ferro |
|---|---|---|---|
| Space group | $P4_3$ (or $P4_1$) | $P4_3$ | $P2_12_12$ |
| a ( Å ) | 58.5 | 58.24 | 37.32 |
| b ( Å ) | 58.5 | 58.24 | 87.02 |
| c ( Å ) | 42.3 | 41.86 | 34.51 |
| V ( $Å^3$ ) | 145,000 | 142,000 | 112,000 |
| $\rho$ ( $g.cm^{-3}$ ) | 1.26 | 1.26 | 1.35 |
| Z | 4 | 4 | 4 |
| Fractional volume of protein | 45 % | 45 % | 53 % |

vi Data collection

The intensity data were collected on a computer controled four circle diffractometer, equipped with a scintillation counter, using a pulse height discriminator. The Ni-filtered Cu-Kα radiation was used. The discriminator was symmetrically set to collect about 95% of Cu-Kα pulse distribution. All measurements were carried out with the moving-crystal stationary-counter(ω scan) method. The ω scan method was considered more suitable than ω-2θ scan(moving-crystal moving-counter) method, to minimize the error in the intensity measurement of the crystal of relatively poor crystallinity and with a large identity period.

The crystals used for x-ray experiments were obtained by cutting large crystals to sizes of about 0.7×0.7×0.1mm. Intensity data had to be collected as fast as possible to minimize the decrease in intensity due to x-ray irradiation damage, and to avoid the error due to the oxidation of the crystal. Such large crystal sizes satisfied above conditions. On collecting intensity, the crystals were sealed with mother liquor in Lindemann glass capillaries, replaced with nitrogen gas. The percentage reduction of crystals was improved with the gas replacement. The capillaries 0.01mm in thickness, 1.5mm in diameter, were washed in a mixture of sulfuric acid and nitric acid before use.

Intensity data of about 8,000 reflections for 2.3Å resolution were collected for each crystal of the native,

the U, the Sn, and the Ir derivative. During the experiments, three monitor reflections of (12 0 1), (4 17 3), and (0 15 6) were measured at every 50 reflections to check the deterioration of the crystal and the degree of oxidation. The example of the changes of the intensities of the monitor reflections(Sn-III) is shown in Fig.8. The crystals of the protein and the derivatives are fairly stable against x-ray irradiation. They, however, were oxidized to some extent during the measurement, though no serious intensity changes due to the partial oxidation were observed. For each of the native and the Ir derivative, a complete set of the intensity data up to 2.3Å resolution were obtained with one piece of the crystal. On the other hand, two pieces of the crystals were necessary for each of the U, and Sn derivative, to restrain the maximum change of the structure factors of the monitor reflections within 6%. A data set for 3.0Å resolution could be collected with one piece of the crystal, and a data set from 3.0 to 2.3Å were measured with another crystal(Table 5).

At the beginning and the end of the x-ray measurements, the percentage reduction of the crystal was estimated from the optical absorption. The percentage reduction of the crystal was found to close to 100% at the beginning, and was about 80% at the end of the measurement(Table 5). Although several complete sets of the intensity data were collected, only one set which seemed as the best for each crystal, was used for further structure analysis. Because the degree of reduction of the protein and the occupancies

of the heavy atoms depended on each of the crystals, no averaging of the data among the different sets was made. For the native, three sets were collected in order to estimate the experimental errors. The data labeled 'Native-III' was adopted for the native.

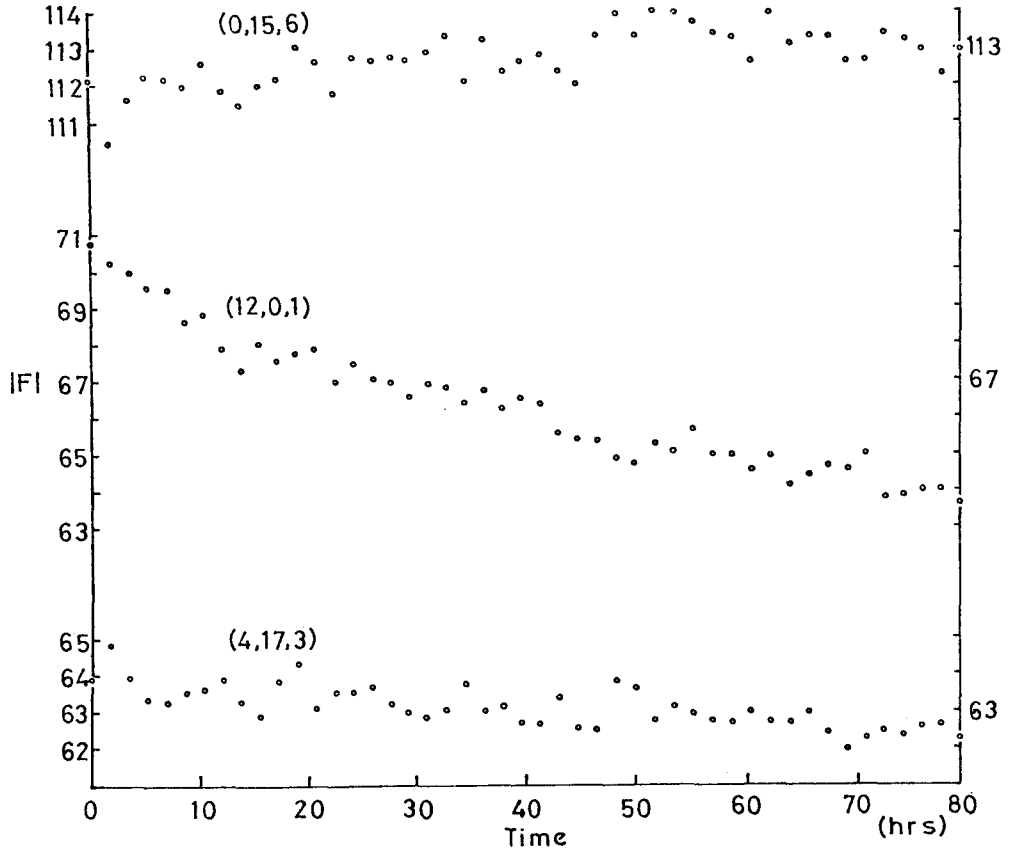



Fig.8 Example of changes of the structure factors of monitor reflections in the measurement of Sn-III.

Table 5 Conditions in intensity measurements

| Derivative | | Resolution Å | Measuring time hr | Scan range ° | width of diffraction profile* ° | final degree of reduction % | maximum change of monitor reflections % |
|---|---|---|---|---|---|---|---|
| Native | I | $\infty \sim 3$ | 80 | 1.4 | 0.55(0.25) | 81 | -3.1 |
| Native | II | $\infty \sim 2.3$ | 150 | 1.2 | 0.75(0.30) | 73 | -3.4 |
| Native | III | $\infty \sim 2.3$ | 150 | 1.0 | 0.50(0.20) | 80 | +0.6 |
| U | IV | $\infty \sim 3$ | 80 | 1.0 | 0.65(0.25) | 95 | +0.9 |
| U | V | $3 \sim 2.3$ | 120 | 1.4 | 0.75(0.25) | 86 | -0.8 |
| Sn | II | $\infty \sim 3$ | 80 | 1.2 | 0.90(0.35) | 98 | +6.5 |
| Sn | III | $3 \sim 2.3$ | 100 | 1.0 | 0.60(0.25) | 64 | -4.5 |
| Ir | V | $\infty \sim 2.3$ | 150 | 1.0 | 0.60(0.20) | 76 | +6.0 |

* The half width of the profile is given in parentheses.

## III PRINCIPLE OF STRUCTURE DETERMINATION

Informations about the phases of structure factors for non-centrosymmetric structures can be obtained by studying the changes in structure amplitudes due to the replacement of heavy atoms into a structure, provided that the rest of the structure is completely unchanged by the substitution. The isomorphous replacement method can be used to determine phases for protein crystals, if two or more isomorphous derivatives can be obtained with different heavy atom positions in each(Fig.9).

The principle of the method is given by the following relation of three vectors, $\vec{F}_{PH}$, $\vec{F}_{P}$, and $\vec{f}_{H}$,

$$\vec{F}_{PH} = \vec{F}_{P} + \vec{f}_{H} \qquad (1)$$

where $\vec{F}_{PH}$ is the structure factor from the isomorphous derivative, $\vec{F}_{P}$ being the structure factor from the native one, and $\vec{f}_{H}$ being the contribution of the heavy atoms alone to $\vec{F}_{PH}$. A convenient construction for this triangle was given by Harker(31). In order to construct the phase triangle, the accurate $\vec{f}_{H}$ is necessary. One of the most important points is putting $\vec{F}_{PH}$ on the scale of $\vec{F}_{P}$. It is identical with tracing the following two steps to ask for $\vec{f}_{H}$; 1) determination of the heavy atom positions in the unit cell, all with respect to the same origin, and 2) calculation of the heavy atom contributions to the structure factor for each reflection.

The ideal solution of Fig.9(b) is, however, seldom realized for the effects of experimental errors, imperfect isomorphism, imperfect knowledge of heavy atom parame-

ters, and the fact that $\vec{f}_H$ of any one derivative is too small to be used in phase determination for a proportion of the reflections. The Harker's method(31) is furthermore inadequate, when many reflections have to be considered.

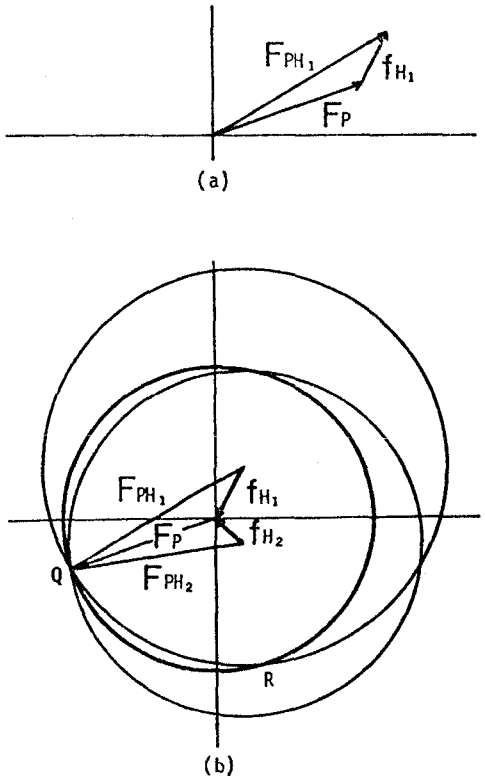



Fig.9 Principle of phase determination by isomorphous replacement method.

(a) Phase triangle related to $\vec{F}_P$, $\vec{F}_{PH}$, $\vec{f}_H$.

(b) Harker's graphical method.

i Probability distribution of protein phases

The probability of the phase angle was evaluated instead of the correct phase angle. The best phase was determined from the probability distribution. The analytical method to treat the probability distribution was suggested by Blow and Crick(32) who showed how the errors involved could best be treated. It is assumed that the errors from all sources — experiment, lack of isomorphism, imperfect refinement — can be regarded as residing in the magnitudes $|F_{PH}|$, and the lack of closure of the phase triangle is, therefore, attributed to the $|P_{PH}|$ side. For any protein phase angle $\alpha$, the vector sum of $\vec{F}_P$ and $\vec{f}_H$ is given by $\vec{F}_{PH} + \varepsilon(\alpha)$, not by $\vec{F}_{PH}$(Fig.10(a)). The lack of closure error $\varepsilon(\alpha)$ is therefore,

$$\varepsilon(\alpha) = \{|F_P|^2 + |f_{Hi}|^2 + 2|F_P||f_{Hi}|cos(\alpha-\phi_i)\}^{1/2} - |F_{PHi}| \qquad (2)$$

where $\phi_i$ is the phase angle of $\vec{f}_{Hi}$. The probability of the phase angle $\alpha$ being correct for a single isomorphous derivative is then,

$$p(\alpha) = Nexp(-\varepsilon_i(\alpha)^2/2E_i^2) \qquad (3)$$

where N is a normalizing factor, and $E_i^2$ is the mean square error in $\vec{F}_{PHi}$. $E_i^2$ can be calculated from centric data by means of the relation

$$E_i^2 = <(|F_{PHi}| - |F_P| - |f_{Hi}|)^2>. \qquad (4)$$

When two or more isomorphous derivatives are used, the total probability of a given phase angle $\alpha$ being correct is given by

$$p(\alpha) = Nexp\{-\Sigma_i \varepsilon_i(\alpha)^2/2E_i^2\}. \qquad (5)$$

Another analytical method was suggested by Perutz et al.(33). It is assumed that the errors inside in $\vec{F}_P$ and $\vec{F}_{PH}$ are all equal to E. A radial vector with any phase angle $\alpha$ is drawn from the origin of the circle of $\vec{F}_P$(Fig.10(b)). The vector and the circle of the $i$th isomorphous derivative(containing the native crystal) intersect at $A_i$. If the mean of $OA_i$ is given by OA, the following relation is hold.

$$\varepsilon_i(\alpha) = OA_i(\alpha) - OA(\alpha) \tag{6}$$

The probability of the phase angle is then,

$$p(\alpha) = N exp(-\Sigma_i \varepsilon_i(\alpha)^2/2E^2) \tag{7}$$

where $OA_i$ is

$$OA_i(\alpha) = -|f_{Hi}|cos(\phi_i-\alpha) \pm\{|F_{PHi}|^2 - |f_{Hi}|^2 sin^2(\phi_i-\alpha)\}^{1/2}.$$

Two solutions are generally obtained for $OA_i(\alpha)$. The larger one is regarded as for the phase angle $\alpha$, and the smaller is as for the phase angle $\alpha+\pi$.

The assumptions that the errors from all sources reside in $\vec{F}_{PH}$ alone, or that the errors are equal for each derivative are, however, seemed to be not so close to the actual case. Therefore in this work, above two methods are blended. The weighted mean of $OA_i$ by $E_i^{-1}$ is OA, and the divergence about the mean is given by

$$\varepsilon_i(\alpha) = OA_i(\alpha) - OA(\alpha). \tag{8}$$

The probability is

$$p(\alpha) = N exp\{-\Sigma_i(\varepsilon_i(\alpha)^2/2E_i^2)\}. \tag{9}$$

$E_i$ of the native crystal may be estimated, for instance, from the errors among the sets of intensity data independ-

ently collected.

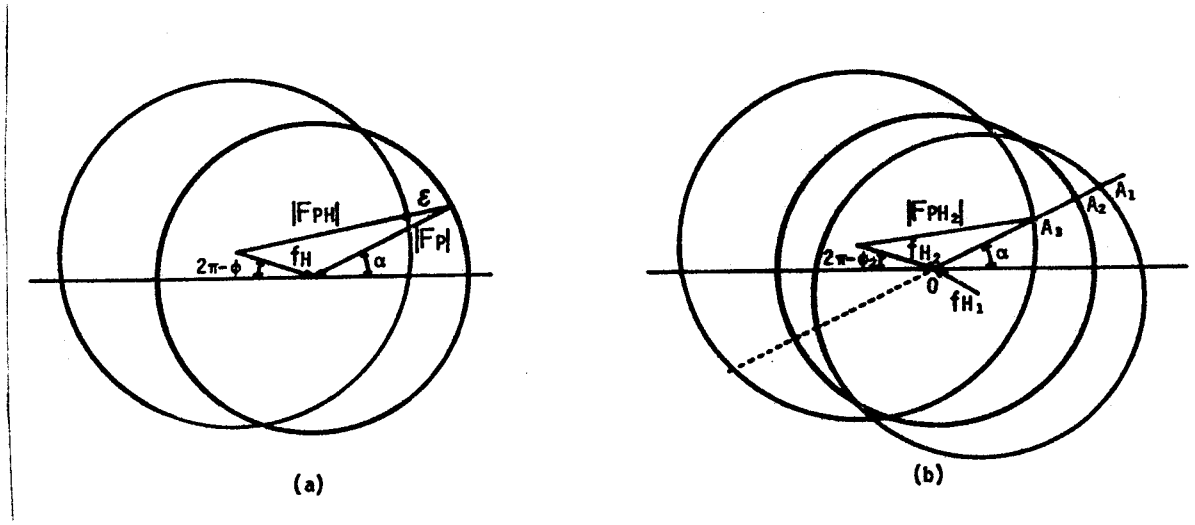



Fig.10 Analytical treatment of probability distribution of phase angle.

(a) Definition of lack of closure error by Blow and Crick(32).

(b) Definition of that by Perutz et al.(33).

ii Best phase angle

The most probable phase angle is defined by the angle where the probability distribution has a maximum value. However it has been shown that the most probable phase angle do not necessarily give a best Fourier, where the best Fourier is defined as the Fourier transform which is expected to have a minimum mean square difference from the Fourier transform of the true $\vec{F}_p$'s(not obtained really) (32,34,35). This Fourier summation is obtained by using weighted $|F_p|$ values and the best phase angles. If the structure factor used in the Fourier synthesis is $\vec{F}_s$ and its true value is $\vec{F}_t$, then the mean square error over the entire unit cell from this reflection pair is given by,

$$<(\Delta\rho)^2> = (2/V^2)(\vec{F}_s - \vec{F}_t)^2 \qquad (10)$$

where V is the volume of the unit cell. $\vec{F}_t$ can be supposed to have a magnitude $|F_p|$ and a probability $p(\alpha)$ of having phase angle $\alpha$, so that

$$\vec{F}_t = |F_p|\vec{r}(\alpha)$$

where $\vec{r}(\alpha)$ is a unit vector in the direction $\alpha$. Eq.(10) then becomes

$$<(\Delta\rho)^2> = (2/V^2)\int_{\vec{r}(\alpha)} \{\vec{F}_s - \vec{F}_p.\vec{r}(\alpha)\}^2 p(\alpha) d\vec{r}(\alpha) \qquad (11)$$

Minimizing this quantity with respect to phase angle gives the following relation,

$$\vec{F}_s(\text{best}) = \int_{\vec{r}(\alpha)} p(\alpha)|F_p|\vec{r}(\alpha)d\vec{r}(\alpha) \;/\; \int_{\vec{r}(\alpha)} p(\alpha)d\vec{r}(\alpha)$$

$$= m|F_p|exp(i\alpha_b). \qquad (12)$$

This result shows that the center of the gravity of the probability distribution of $\vec{F}_p$ is represented as the polar coordinate $(m|F_p|,\alpha_b)$, where $\alpha_b$ is the best phase angle, and the vector should be used in calculating the best Fourier. As the probability distributions(eq. 5, 7, or 9) are calculated at angles $\alpha_i$, regular intervals apart around the phase circle, eq.12 can be modified as being easy to calculation,

$$\begin{aligned} m\cos(\alpha_b) &= \sum_i p(\alpha_i)\cos(\alpha_i)/\sum_i p(\alpha_i) \\ m\sin(\alpha_b) &= \sum_i p(\alpha_i)\sin(\alpha_i)/\sum_i p(\alpha_i) \end{aligned} \qquad (13)$$

If the real axis is rotated by $\alpha_b$ to coincide it with the direction of $\alpha_b$ on the phase plane, the error of the phase angle $\alpha_i$ becomes

$$\Delta\alpha_i = \alpha_b - \alpha_i = -\alpha_i.$$

Then from eq.13,

$$\begin{aligned} m &= \sum_i p(\alpha_i)\cos(\Delta\alpha_i)/\sum_i p(\alpha_i) \\ &= \langle\cos(\alpha_i)\rangle. \end{aligned} \qquad (14)$$

m, commonly called as the <u>figure of merit</u>, is the mean value of the cosine of the error in phase angle. This is one index which represents the reliability of the phase angle, that is, the reliability of the calculated electron density distribution. The mean square error of the electron density defined by eq.10, is rewritten as following

$$\langle(\Delta\rho)^2\rangle = (2/V^2)\sum_{\vec{h}}|F_{\vec{h}}|^2(1- m_{\vec{h}}^2) \qquad (15)$$

by using m. Some typical examples of most probable and best phase angles are shown in Fig.11.

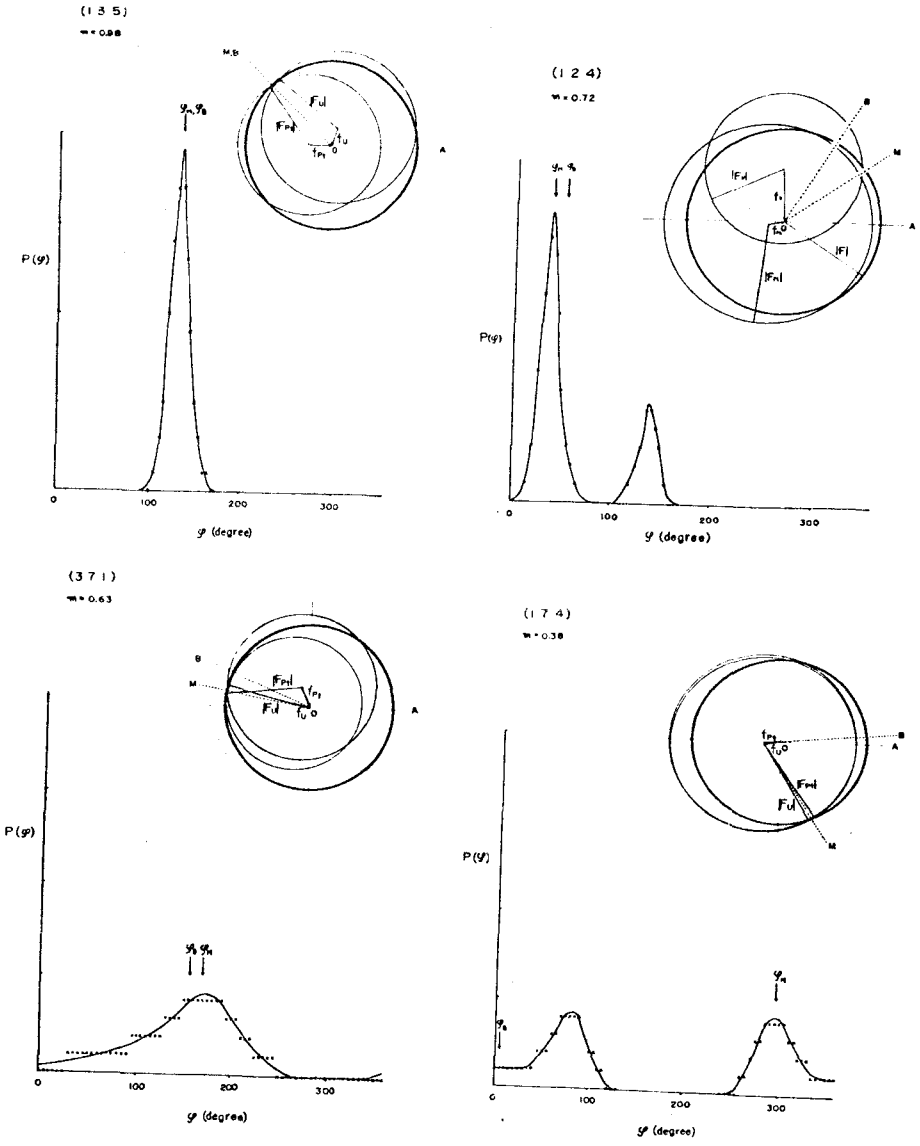



Fig.11 Examples of probability distribution for some reflections.

## IV STRUCTURE DETERMINATION

### i Data reduction

*Basic data correction*

Observed intensities were reduced to structure factors by the following relation,

$$|F| = k(AI)^{1/2}/(LP)$$

where I is an observed intensity, A is an absorption correction factor, LP Lorentz-polarization factor, and k absolute scale factor. Absorption correction was carried out by the Furnas' method(36). An example of the absorption curve as a function of $\phi$ is shown in Fig.12.

*Scale factor*

The absolute scale factors for the intensity data of the native and isomorphous crystals were determined by the Wilson's method(37). The Wilson's method to obtain the absolute scale factor and the temperature factor is not so appropriate in the case of proteins which, although the number of reflections are large, lack the data of higher Bragg angles. Fortunately, however, the accuracy of the absolute scale factor is not so vital in the protein crystallography. It is the relative scale factor which affects mostly the accuracy of the phase angle determination.

The following relation is hold between two data sets $i$ and $j$,

$$|F_j| = k'|F_i| exp(-\Delta B sin^2\theta/\lambda^2)$$

where k' is a relative scale factor of data set $j$ against the data set $i$, and $\Delta B$ is a relative temperature factor

between the apparent temperature factors of the two crystals($\Delta B=B_i-B_j$). k′ and $\Delta B$ were obtained from Wilson type plots of $ln(|F_{PH}|^2/|F_P|^2)$ versus $sin^2\theta/\lambda^2$. A relative scale factor was also estimated by using an expression derived by Kraut et al.(38),

$$|F_{PH}|^2 = k'|F_P||F_{PH}|exp(-\Delta B sin^2\theta/\lambda^2).$$

The initial parameters of the relative scale factor and temperature factor of each derivative data against the native crystal were determined by averaging them with appropriate weights.

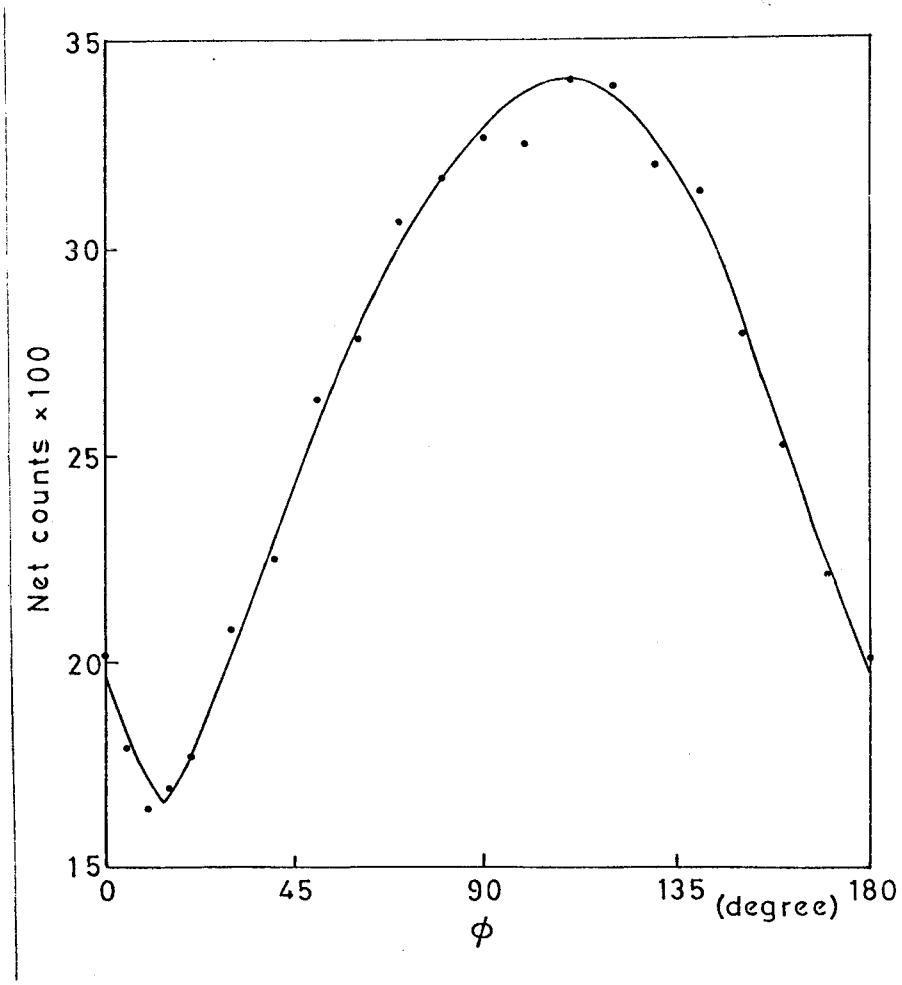



Fig.12 Example of the absorption curve of (0 0 2) reflection(U-IV).

ii Determination of heavy atom positions

The heavy atom positions can be determined from the difference Patterson synthesis(39), with coefficient

$$|\Delta F|^2 = (|F_{PH}| - |F_P|)^2.$$

The heavy atoms of which occupancies are fairly high(major sites) can be determined by this method. The positions of the heavy atoms, thus obtained, in each derivative must be correlated to a common origin before the phase determination. This correlation can be made by cross-difference Fourier maps(40). The heavy atom parameters of the two isomorphous derivatives U and Pt, used in the 6Å analysis, were determined by the above method(22).

The little structural differences of any crystal which is still isomorphous to the native one, may be revealed in the difference Fourier synthesis, if the phase angles, $\alpha_b$'s, of the native crystal have been determined. The difference Fourier synthesis with coefficient

$$(|F_{PH}| - |F_P|)exp(i\alpha_b)$$

can show up the sites of the heavy atoms. The coefficient can be rewritten as follows(41);

$$\begin{aligned}(|F_{PH}| - |F_P|)exp(i\alpha_b) = {} & \{|F_P|/2(|f_H|+|F_P|)\}|f_H|exp(i\alpha_H) \\ & + \{|f_H|^2/(|f_H|+|F_P|)\}exp(i\alpha_b) \\ & + \{|f_H||F_P|/2(|f_H|+|F_P|)\}exp\{i(2\alpha_b-\alpha_H)\} \quad (16)\end{aligned}$$

where $\alpha_H$ is the phase angle of heavy atoms(Fig.10). In the right side of eq.16, the first term gives rise to a weighted heavy atom map, which is asked for. The third term will give background noises as $(2\alpha_b-\alpha_H)$ have random values, and the second term will give rise to some fea-

tures of the protein. However $|F_p|$ is much greater than $|f_H|$, the contribution of the second term to the synthesis is very small. This synthesis is, therefore, useful to determine the heavy atom positions of which occupancies are rather low(minor sites).

For the Sn and Ir derivatives, some major and minor sites were found out by the syntheses, using the $\alpha_b$'s of the 4Å analysis. These parameters are given in Table 6. Two atoms in each pair of, Sn-1 and Sn-2, Sn-3 and Sn-4, Ir-1 and Ir-2, are not equivalent, respectively, but related by the pseudo-symmetry, as shown in Fig.6.

Table 6 Initial and intermediate heavy atom parameters

The parameters after 10 cycles of refinement, using 760 reflections within 4Å resolution, are listed. For the starting best phase angles, the results of 4Å study were adopted.

(a) Sn derivative

| site | x/a | y/b | z/c | B | g |
|---|---|---|---|---|---|
| Sn-1** | 0.4096 (0.40) | 0.2455 (0.25) | 0.0412 (0.05) | 23.3 (10.0) | 30.8 (20.0) |
| Sn-2 | 0.8760 (0.90) | 0.2499 (0.25) | 0.0622 (0.05) | -21.9 (10.0) | 7.0 (20.0) |
| Sn-3 | 0.3361 (0.32) | 0.4638 (0.45) | 0.1884 (0.20) | 12.1 (10.0) | 11.4 (10.0) |
| Sn-4 | 0.8161 (0.82) | 0.4442 (0.45) | 0.1978 (0.20) | 10.2 (10.0) | 46.5 (10.0) |

(b) Ir derivative

| site | x/a | y/b | z/c | B* | g* |
|---|---|---|---|---|---|
| Ir-1 | 0.4365 (0.44) | 0.2718 (0.26) | 0.2133 (0.21) | 15.0 | 20.0 |
| Ir-2 | 0.9411 (0.94) | 0.2615 (0.26) | 0.2114 (0.21) | 15.0 | 20.0 |
| Ir-3** | 0.7229 (0.72) | 0.2709 (0.26) | 0.2358 (0.24) | 15.0 | 20.0 |

g is in the unit of $eÅ^{-3}$.

* These parameters were fixed during the refinements.

** These sites were disappeared in the course of refinements.

iii Refinement of heavy atom parameters

When the positions of substituted heavy atoms have been determined, it is necessary to refine the following paramters before phase determination: 1) the scale factor(k') relating the heavy atom derivative measurements to those of the native protein, 2) the relative temperature factor(ΔB) between the derivative and native data, 3) the positions of heavy atoms, 4) the temperature factors(anisotropic temperature factors are needed at high resolution analysis) of heavy atoms, and 5) the occupancies(g) of heavy atoms. These parameters were refined by the following two steps.

*Step 1 Two dimensional refinement*

Each projected plane along the principal axes of the crystal of the space group $P2_12_12_1$, to which the bonito crystals belong, has a center of symmetry. The heavy atom parameters were refined by least-squares method using the centric reflections. For the centric reflections, it is proper to minimize the quantity;

$$\Sigma\{|F_{PH}| - (1/k')|\vec{F}_P + g.\vec{f}_H|exp(\Delta B sin^2\theta/\lambda^2)\}^2$$

where the summation is over all the centric reflections, and $\vec{f}_H$ is given by

$$\vec{f}_H = \hat{f}_H\Sigma_j exp\{2\pi i(hx_j+ky_j+lz_j)\}exp(-B_j sin^2\theta/\lambda^2), \quad (26)$$

$\hat{f}_H$ being a unitary atomic scattering factor. The signs of $F_{PH}$ and $F_P$ were easily assigned at this stage. A common $\hat{f}_H$ was used for the heavy atoms, since $\hat{f}$'s are very similar for those with atomic number greater than 60.

*Step 2 Three dimensional refinement*

If heavy atom parameters are refined well, the protein phase $\alpha_b$ is calculated by the method of sec.III-ii. Then the three dimensional least-squares refinement was applied by using the calculated phase angles. The quantity to be minimized was

$$\Sigma\omega(k'|F_{PH}| - ||F_P|exp(i\alpha_b)+ \vec{f}_H|)^2,$$

where the summation is over all reflections, $\alpha_b$ is the current estimate of the best phase angle, and $\omega$ is an appropriate weight(34,38), in this case $\omega$ being chosen as proportional to the figure of merit (m) of each reflection. The parameters of the Sn and Ir derivatives were refined independently at first, using the $\alpha_b$'s of the 4Å analysis for the starting ones. The parameters after 10 cycles of refinement, using 760 reflections within 4Å resolution, are listed in Table 6. The mean figure of merit was improved using new derivatives(Sn and Ir) to 0.72 from 0.53. The mean difference between the new $\alpha_b$'s estimated with the U, Sn and Ir derivatives, and the old $\alpha_b$'s, ($<\Delta\phi>$), was 28°. Then the parameters of each derivative were refined. About 1300 strong reflections within 3Å range and about 1200 strong reflections from 3Å to 2.3Å range were included in further refinements. The refinements were carried out independently, because the ∞-3Å data and 3-2.3Å data of the U and Sn derivatives were collected by using different crystals(Table 5). Throughout the refinement, the absolute scale factor for the native crystal was fixed at 14.2, which was determined by

the Wilson's method, and only the relative scale factor of each derivative was refined. Damping factors for the shifts of the relative scale factor, relative temperature factor, occupancies, coordinates, and temperature factors of heavy atoms were, for instance in the final case, 0.50, 0.05, 0.20, 0.40, and 0.01 respectively.

During the refinements, difference Fourier syntheses with coefficients

$$(|F_{PH}|-|F_P|)\,exp\,i\alpha_b$$

$$(|F_{PH}|-|F_P + f_H|)\,exp\,i\alpha_{PH}$$

were made for each derivative, to check whether the heavy atom parameters were correctly refined, and whether there were new sites of heavy atoms with low occupancies(38,42, 43), where $\alpha_{PH}$ is the phase angle of the heavy atom derivative. However any minor sites did not appeared in the maps.

The refinement cycle, phase determination—least-squares refinement—difference Fourier synthesis, was repeated untill the parameters were sufficiently converged.

It was asked for most to break out the pseudo-symmetry in the course of the refinement. This was carried out at the 4Å analysis(23), and was newly supported by containing the Sn derivative in the calculation, because the pseudo-symmetry was broken out by disappearance of one of the major site(Sn-1) during the refinement. The final heavy atom parameters are given in Table 7, and some of the criteria of the refinement are shown in Fig.13.

Table 7 Final heavy atom parameters

| Derivative | Site No. | x/a | y/b | z/c | $\beta_{11}$ | $\beta_{22}$ | $\beta_{33}$ | $\beta_{12}$ | $\beta_{13}$ | $\beta_{23}$ | g |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $K_3UO_2F_5$ | U -1 | 0.0897 | 0.1317 | 0.0506 | 0.00239 | 0.00047 | 0.00644 | 0.00045 | -0.00078 | 0.00025 | 42.0 |
| | U -2 | 0.5863 | 0.1412 | 0.0807 | 0.00149 | 0.00100 | 0.00608 | -0.00109 | -0.00226 | -0.00027 | 69.8 |
| $(CH_3)_2SnCl_2$ | Sn-2 | 0.8950 | 0.2473 | 0.0513 | 0.00678 | 0.00098 | -0.00079 | 0.00185 | 0.00277 | 0.00182 | 12.0 |
| | Sn-3 | 0.3537 | 0.4682 | 0.1738 | 0.00896 | -0.00020 | 0.00633 | -0.00176 | -0.00912 | 0.00771 | 6.2 |
| | Sn-4 | 0.8182 | 0.4466 | 0.1964 | 0.00178 | 0.00040 | -0.00167 | 0.00135 | -0.00289 | -0.00309 | 63.0 |
| $K_3IrCl_6$ | Ir-1 | 0.4250 | 0.2676 | 0.2108 | 0.00333 | 0.00092 | 0.00409 | -0.00393 | 0.00553 | -0.00152 | 10.3 |
| | Ir-2 | 0.9317 | 0.2596 | 0.2027 | -0.00064 | 0.00116 | 0.00516 | -0.00108 | -0.00104 | 0.00069 | 15.0 |

The anisotropic temperature factor is given by $\exp(\beta_{11}h^2+\beta_{22}k^2+\beta_{33}l^2+\beta_{12}hk+\beta_{13}hl+\beta_{23}kl)$.

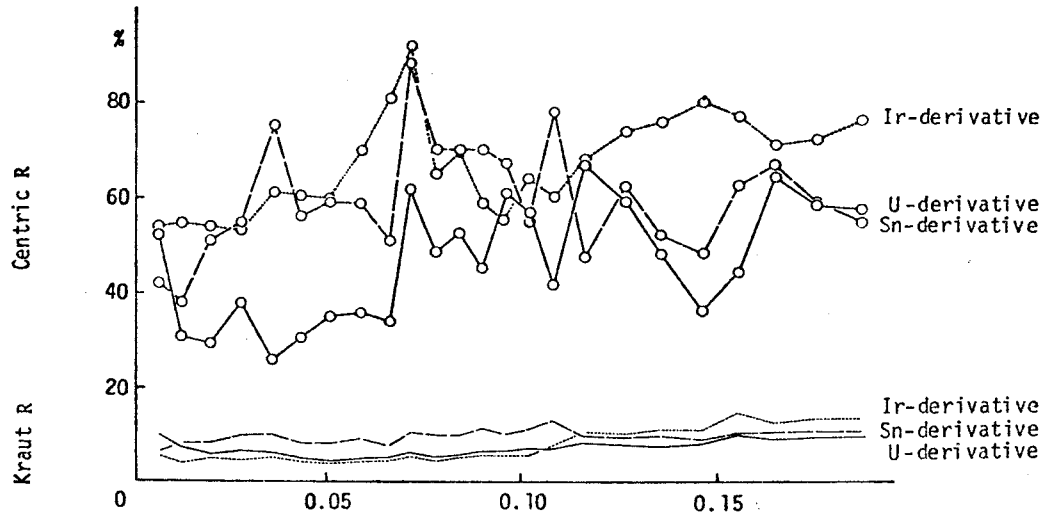
The occupancy, g, is in the unit of $e\text{Å}^{-3}$.

Fig.13 Variation of some statistical data against $(2sin\theta/\lambda)^2$. A total of 5118 reflections are included in this statistics.

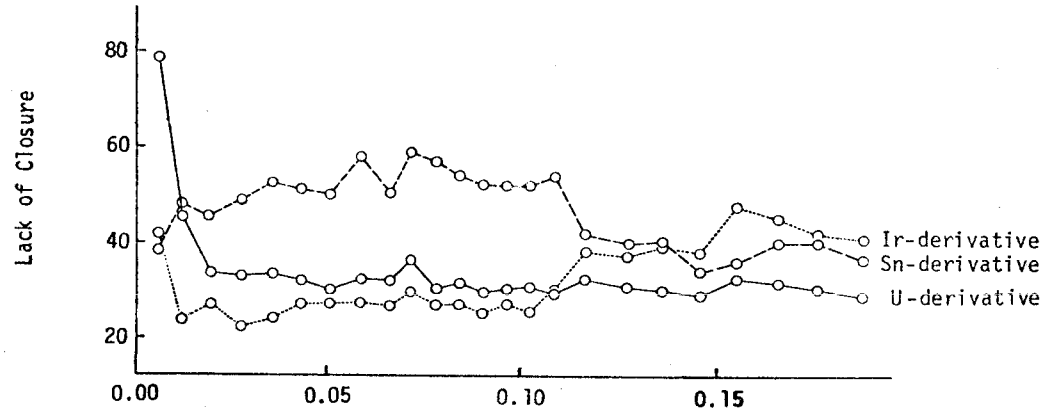
(a) Centric R and Kraut R.

(b) The root mean square value of lack of closure calculated with the final phase angles.

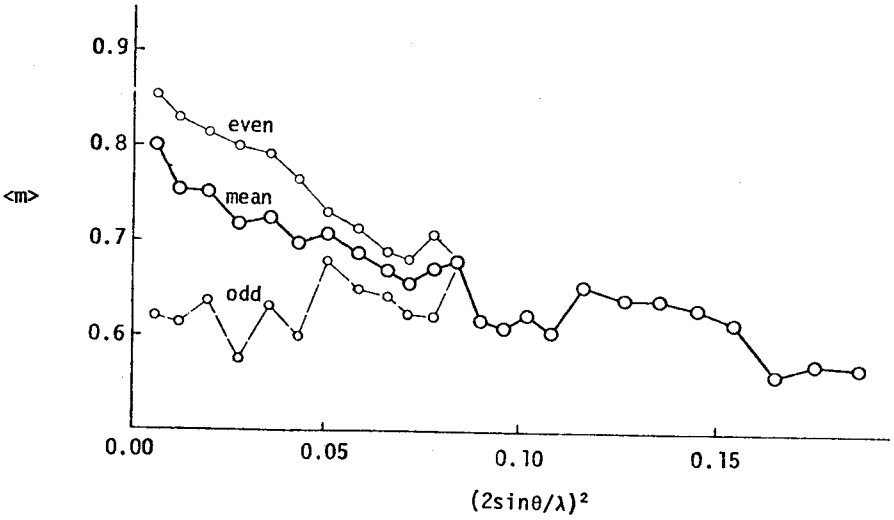
(c) Figure of merit. The 'even' and 'odd' show the reflections with h even and odd, respectively. The 'mean' represents the figure of merit calculated with all significant reflections.


%
Centric R
80
60
40
20
0
0.05
0.10
0.15
Ir-derivative
U-derivative
Sn-derivative
Kraut R
Ir-derivative
Sn-derivative
U-derivative

(a)


Lack of Closure
80
60
40
20
0.00
0.05
0.10
0.15
Ir-derivative
Sn-derivative
U-derivative

(b)


<m>
0.9
0.8
0.7
0.6
even
mean
odd
0.00
0.05
0.10
0.15
(2sinθ/λ)²

(c)

iv Calculation of electron density map

The phase probability was estimated by eq.9. In the equation, $E_i$ for the native crystal was estimated from the root mean square differences among three sets of the data, and $E_i$ for the derivatives were estimated from the root mean square value of lack of closure errors, $(|F_{PH}| - |F_P+f_H|)$, of phase triangles of all the reflections included in the least-squares calculation. The probability of a total of about 5100 reflections which were significantly above the background intensity, was calculated at 5° intervals using the program written by Prof. Ashida. The mean figure of merit was 0.65. This value is nearly comparable of those of high resolution analyses of other proteins. The mean square error of the electron density, estimated by eq.15 was $0.075e\mathring{A}^{-3}$. On the other hand, the electron density from which the polypeptide chain was identified, was generally higher than $0.3e\mathring{A}^{-3}$, and the electron density for side chains was higher than $0.2e\mathring{A}^{-3}$.

## V INTERPRETATION OF THE ELECTRON DENSITY MAP

The electron density map was interpreted on the basis of the sequence determined by Nakayama et al.(3). The Kendrew type skeletal model was constructed on the scale 1Å = 2cm, using the optical device described by Richards (44) which enables the electron density map and the model to be viewed simultaneously.

The planar heme group and two sulfur atoms of Cys14 and Cys17 are prominent in the map. It is clear that the heme iron is almost on the plane. However it is impossible to discuss about the distortion of the porphyrin ring from its planarity, from the 2.3Å resolution map. The highest peak in the map is the heme iron atom with a height of $2.4e\text{Å}^{-3}$. The peak heights of the four sulfur atoms are 1.8, 1.4, 1.4, and $1.0e\text{Å}^{-3}$ respectively. On the other hand, the average height of the peaks due to the carbonyl groups of the peptide linkages is $0.9e\text{Å}^{-3}$, and the highest of them, $1.6e\text{Å}^{-3}$, is in the α-helical region. The composite drawing of the electron density of the heme group is shown in Fig.14. The fitting between the electron density map and the model in the heme vicinity is shown in Fig.15.

The main chain could be followed clearly at the upper half of the molecule. In the region the fragments of the main chain contact rigidly with each other. All the helices are contained in this region. The density at the lower half of the molecule is rather strong. However it was not so easy to fit the model to the electron density

in the region. The fragments of the main chain have weak interaction with each other, and are fairly flexible.

The information of the sequence is very useful to follow the main chain. The density for all the side chains in the interior of the molecule is rather high and clear. The density at the terminus of a side chain on the surface of the molecule often fades away, especially for longer side chains. The features of the density for side chains are shown in Table 8.

Isolated peaks in the interior and near the surface of the molecule are found. It is difficult to interpret them as a part of the molecule. They are all in contact with possible hydrogen bond donors and/or acceptors, and are thought to be fixed solvent molecules.

There are still a few ambiguities in the placing of individual atoms. Errors caused by recording the atomic parameters from the Richards' box must also be considered. Therefore the standard deviation of each parameter of the individual atom seemed to be about 0.3Å. The histogram of the bond distances calculated from the coordinates of the atoms are shown in Fig.16. The average bond distances and angles of the peptide linkages are shown in Fig.17, together with the standard values(45).

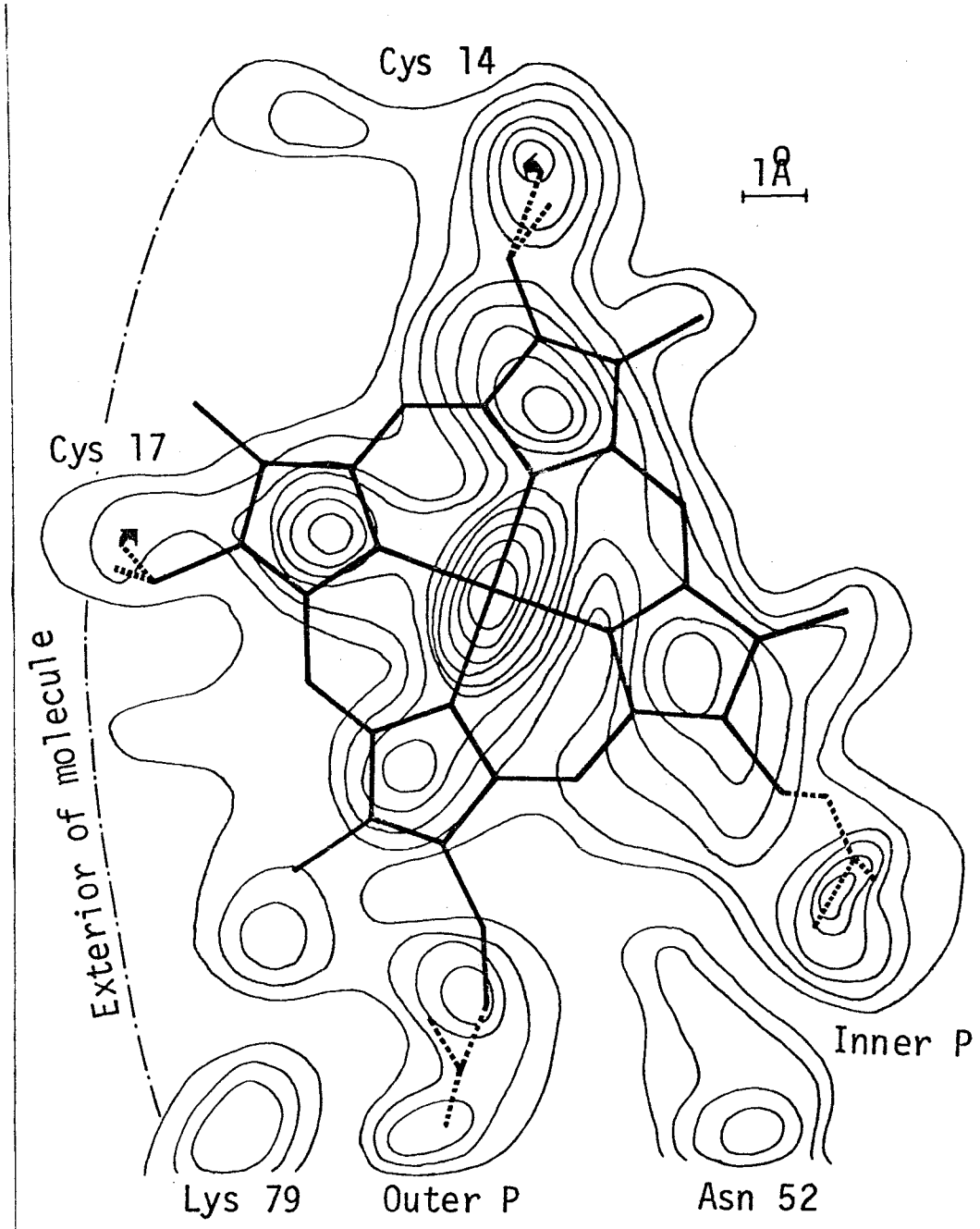



Fig.14 Composite electron density map of the heme plane. An idealized heme skeleton is superimposed at the same scale. First contour $0.0 e\mathring{A}^{-3}$, interval $0.2 e\mathring{A}^{-3}$.

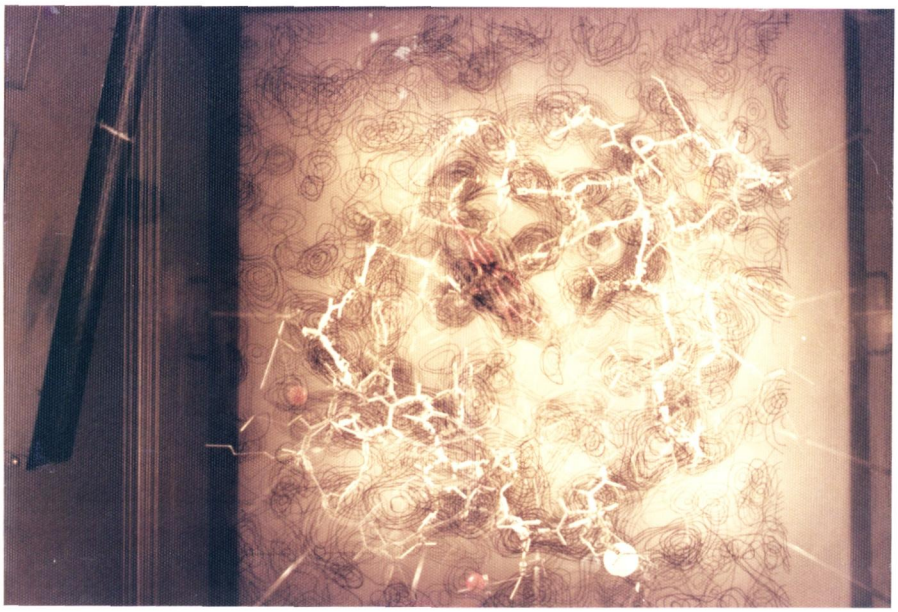


Fig.15 View of the map sections, as seen in the Richards' box with the skeletal model superimposed. Section z=8/38 is nearest to the viewer, and section z=15/38 is at the rear. The red skeleton represents a heme.

# Table 8 Features of electron density map for side chains

KEY. Character of side chains: A, acidic; B, basic; M, ambivalent; R, aromatic; L, aliphatic. Location: E, external; S, surface; I, interior.
Visibility in electron density map: N, not visible; V, visible, but not well oriented or shaped; O, visible and well oriented; S, visible, well oriented, and of characteristic shape for the amino acid concerned. Where -, weak density; +, strong density.

| Residue No. | Type | Character | Location | Visibility |
|---|---|---|---|---|
| 1 | Gly | L | E | - |
| 2 | Asp | A | E | V |
| 3 | Val | L | E | V |
| 4 | Ala | L | E | V |
| 5 | Lys | B | E | V- |
| 6 | Gly | L | I | - |
| 7 | Lys | B | E | V- |
| 8 | Lys | B | E | O- |
| 9 | Thr | M | S | V+ |
| 10 | Phe | R | E | S+ |
| 11 | Val | L | E | V |
| 12 | Gln | M | E | V |
| 13 | Lys | B | E | O |
| 14 | Cys | L | I | S+ |
| 15 | Ala | L | S | V+ |
| 16 | Gln | M | E | S+ |
| 17 | Cys | L | E | S+ |
| 18 | His | B | I | V+ |
| 19 | Thr | M | E | V+ |
| 20 | Val | L | E | V+ |
| 21 | Glu | A | E | S+ |
| 22 | Asn | M | E | S+ |
| 23 | Gly | L | E | - |
| 24 | Gly | L | E | - |
| 25 | Lys | B | E | V |
| 26 | His | B | S | S+ |
| 27 | Lys | B | E | S |
| 28 | Val | L | E | V+ |
| 29 | Gly | L | I | - |
| 30 | Pro | L | I | V+ |
| 31 | Asn | M | I | O |
| 32 | Leu | L | I | V+ |
| 33 | Trp | R | E | S |
| 34 | Gly | L | E | - |
| 35 | Leu | L | I | S+ |
| 36 | Phe | R | S | S |
| 37 | Gly | L | E | - |
| 38 | Arg | B | S | V+ |
| 39 | Lys | B | E | V- |
| 40 | Thr | M | I | V |
| 41 | Gly | L | E | - |
| 42 | Gln | M | E | V+ |
| 43 | Ala | L | S | V |
| 44 | Glu | A | E | V- |
| 45 | Gly | L | E | - |
| 46 | Tyr | R | I | S+ |
| 47 | Ser | M | E | V- |
| 48 | Tyr | R | S | O+ |
| 49 | Thr | M | E | N |
| 50 | Asp | A | E | V- |
| 51 | Ala | L | I | V |
| 52 | Asn | M | I | S+ |
| 53 | Lys | B | E | V- |
| 54 | Ser | M | E | V- |
| 55 | Lys | B | E | V |
| 56 | Gly | L | E | - |
| 57 | Ilu | L | S | S+ |
| 58 | Val | L | V | V+ |
| 59 | Trp | R | I | S+ |
| 60 | Asn | M | E | O |
| 61 | Glu | A | E | O |
| 62 | Asn | M | E | O |
| 63 | Thr | M | S | V |
| 64 | Leu | L | I | O |
| 65 | Met | L | E | S+ |
| 66 | Glu | B | E | O |
| 67 | Tyr | R | I | O |
| 68 | Leu | L | I | O- |
| 69 | Glu | A | E | V |
| 70 | Asn | M | E | O |
| 71 | Pro | L | S | V+ |
| 72 | Lys | B | E | V- |
| 73 | Lys | B | E | V- |
| 74 | Tyr | R | S | O |
| 75 | Ilu | L | I | O+ |
| 76 | Pro | L | E | V |
| 77 | Gly | L | E | - |
| 78 | Thr | M | I | V+ |
| 79 | Lys | B | E | S+ |
| 80 | Met | L | I | S+ |
| 81 | Ilu | L | E | S+ |
| 82 | Phe | R | S | S+ |
| 83 | Ala | L | E | V |
| 84 | Gly | L | E | - |
| 85 | Ilu | L | I | O |
| 86 | Lys | B | E | V- |
| 87 | Lys | B | E | O |
| 88 | Lys | B | E | V- |
| 89 | Gly | L | E | - |
| 90 | Glu | A | E | O |
| 91 | Arg | B | E | O- |
| 92 | Gln | M | E | V+ |
| 93 | Asp | A | E | O |
| 94 | Leu | L | I | O |
| 95 | Val | L | E | V |
| 96 | Ala | L | E | V |
| 97 | Tyr | R | E | S+ |
| 98 | Leu | L | I | O |
| 99 | Lys | B | E | S |
| 100 | Ser | M | E | N |
| 101 | Ala | L | S | V- |
| 102 | Thr | M | I | V |
| 103 | Ser | M | E | V+ |

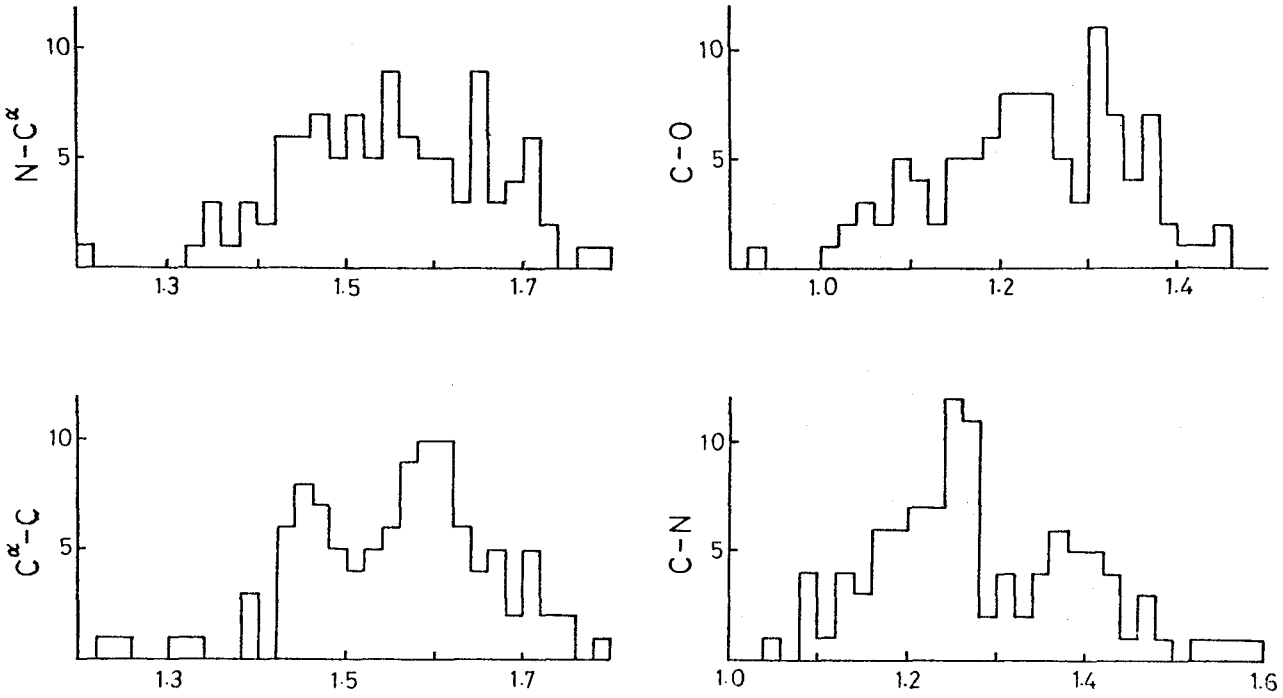



Fig.16 Histogram of bond distances.

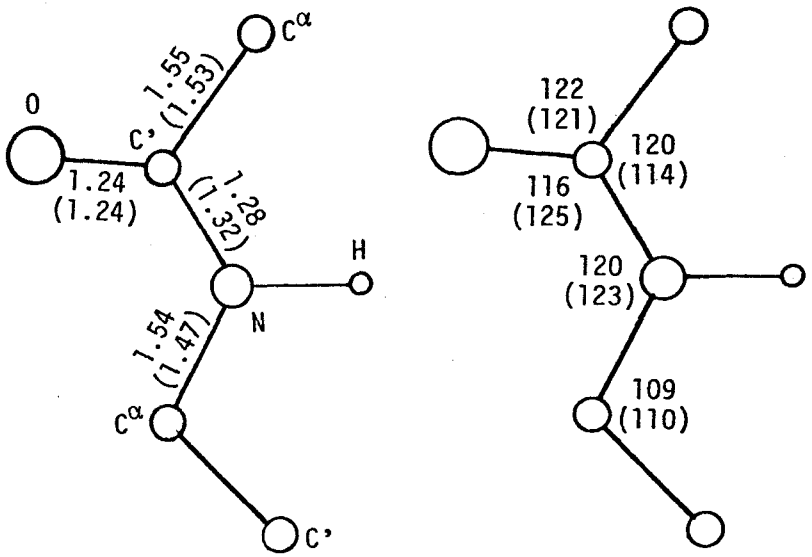



Fig.17 Average dimensions of peptide linkages.

Values in parentheses are given by Corey and Pauling(45).

## VI DESCRIPTION OF THE STRUCTURE

The molecule of bonito ferrocytochrome c is an overall egg shape. It is 35Å in height, 30Å in width, and 23Å in thickness. The heme group is located in a deep crevice of the molecule. The fifth ligand of the heme is the $N^{\varepsilon}$ atom of the imidazole ring of His18, and the sixth is the S atom of Met80. The schematic drawing of the model is shown in Fig.18 and the stereoscopic drawings are shown in Figs.19(a,b,c).

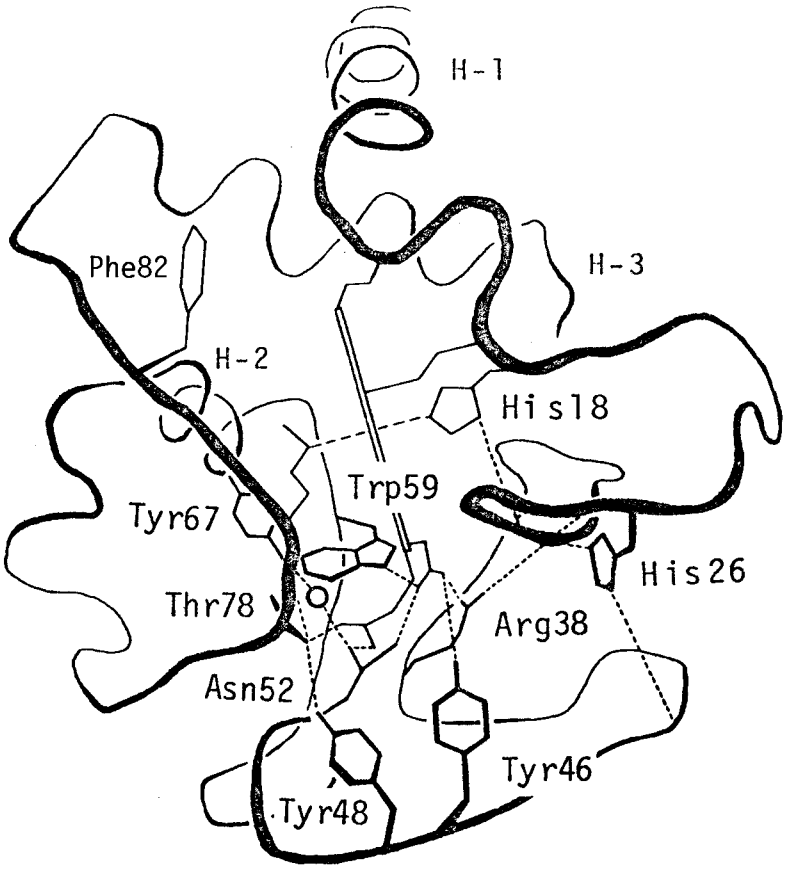
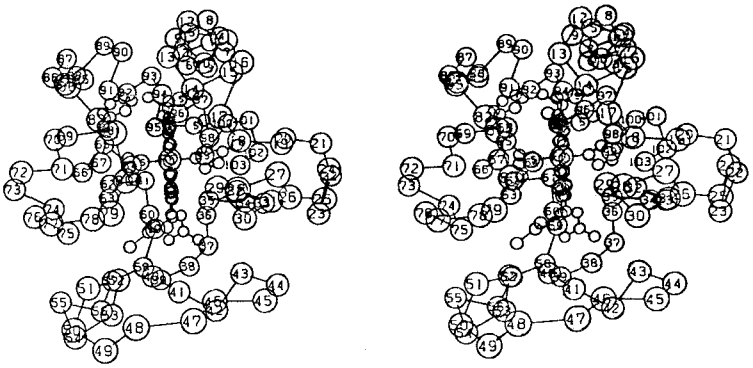
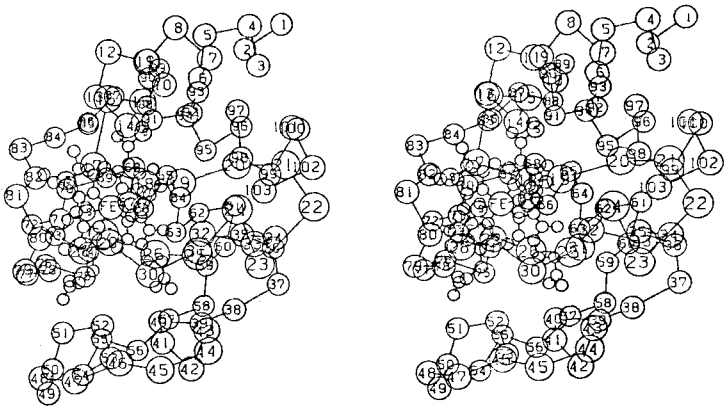


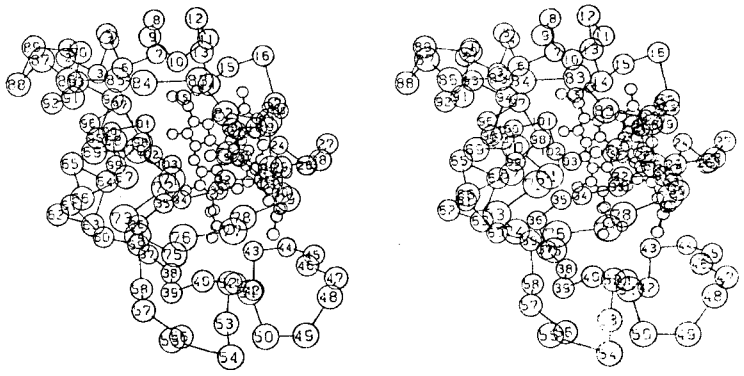

Fig.18 Schematic drawing of main chain folding, front view. Dotted lines show intramolecular hydrogen bonds. Open circle represents W1.



(a) Front view.



(b) Right view.



(c) Left view.

Fig.19 Stereoscopic drawings of $C^{\alpha}$ atoms and heme.

i Conformation of the molecule

The conformation of the main chain is described in terms of the twisting angles of the bonds of the main chain(46,47). The definition of the twisting angles ($\phi$,$\psi$,$\omega$) is given in Fig.20. All the rotational angles around the C'-N bond($\omega$) are close to 180°, that is, all the peptide bonds are 'trans'. The average deviation of $\omega$ from the trans conformation is 8.7° in this model. All the three proline residues are in the trans conformation. The twisting angles for this model are listed in Table 9. The angle pairs ($\phi$,$\psi$) are mapped in Fig.21, together with those of some typical polypeptides. More than half of the points fall within the broken lines(outer limit region). However there are several points in the non-allowed region, all of which are not necessarily glycines. In views of the fact that the ($\phi$,$\psi$) map may give one of the criteria to estimate the accuracy or quality of the molecular model, a further refinement of some minor parts seems to be necessary, and will be made in due course.

Hydrogen bonds seem to take an important part in constructing the molecule of bonito cytochrome c, as described later. The presence of hydrogen bonds was assumed by considering the distances between the donor and acceptor atoms of the model, and angles between the atoms involved. The hydrogen bonds are given in Table 10.

Table 9 Twisting angles of main chain linkages

| RESIDUE | PHI | PSI | OMEGA | RESIDUE | PHI | PSI | OMEGA |
|---|---|---|---|---|---|---|---|
| 2 Asp | -96 | -10 | 159 | 53 Lys | -54 | -108 | -171 |
| 3 Val | -57 | -48 | 167 | 54 Ser | 25 | -106 | 155 |
| 4 Ala | -40 | -67 | 170 | 55 Lys | -6 | -45 | -174 |
| 5 Lys | -48 | -34 | 174 | 56 Gly | 44 | 147 | -153 |
| 6 Gly | -56 | -57 | -157 | 57 Ilu | -160 | -127 | 177 |
| 7 Lys | -60 | -43 | 171 | 58 Val | -172 | 122 | 155 |
| 8 Lys | -40 | -53 | 172 | 59 Trp | -60 | 112 | -171 |
| 9 Thr | -73 | -37 | 175 | 60 Asp | -133 | -179 | 175 |
| 10 Phe | -62 | -50 | -176 | 61 Glu | -56 | -78 | 180 |
| 11 Val | -68 | -27 | 178 | 62 Asn | -23 | -95 | -170 |
| 12 Gln | -67 | -55 | 169 | 63 Thr | -28 | -28 | 176 |
| 13 Lys | -112 | -23 | -179 | 64 Leu | -75 | -40 | 166 |
| 14 Cys | -105 | -63 | 169 | 65 Met | -56 | -47 | -179 |
| 15 Ala | -9 | -80 | 173 | 66 Glu | -44 | -57 | 174 |
| 16 Gln | -58 | -4 | 180 | 67 Tyr | -78 | -51 | 175 |
| 17 Cys | -111 | -103 | -163 | 68 Leu | -41 | -54 | 180 |
| 18 His | -36 | 158 | 170 | 69 Glu | -33 | -83 | 175 |
| 19 Thr | -85 | -143 | 180 | 70 Asn | 179 | 124 | -157 |
| 20 Val | 173 | -12 | -170 | 71 Pro | -67 | 5 | 172 |
| 21 Glu | -95 | -156 | -172 | 72 Lys | -102 | 92 | -163 |
| 22 Asn | -75 | 91 | -175 | 73 Lys | 94 | 87 | -171 |
| 23 Gly | 77 | 20 | -179 | 74 Tyr | 124 | 90 | -175 |
| 24 Gly | -2 | -56 | 170 | 75 Ilu | 62 | 31 | -160 |
| 25 Lys | 133 | 153 | -172 | 76 Pro | -56 | 118 | -171 |
| 26 His | -74 | -141 | -178 | 77 Gly | 118 | 16 | -176 |
| 27 Lys | 131 | 46 | -176 | 78 Thr | -65 | 139 | -177 |
| 28 Val | 114 | -95 | 171 | 79 Lys | -75 | 8 | -175 |
| 29 Gly | -85 | 126 | 180 | 80 Met | -89 | 121 | 178 |
| 30 Pro | -40 | 174 | 167 | 81 Ilu | -61 | 91 | -178 |
| 31 Asn | -53 | 141 | 178 | 82 Phe | -128 | 142 | 172 |
| 32 Leu | -96 | 50 | -176 | 83 Ala | -73 | -84 | -175 |
| 33 Trp | -29 | -91 | 180 | 84 Gly | -136 | -99 | -167 |
| 34 Gly | -131 | 69 | -176 | 85 Ilu | 83 | 59 | -164 |
| 35 Leu | -16 | -60 | 180 | 86 Lys | -124 | 44 | 172 |
| 36 Phe | -55 | 89 | -177 | 87 Lys | -130 | 125 | 167 |
| 37 Gly | 123 | 62 | 170 | 88 Lys | -25 | -73 | 180 |
| 38 Arg | -174 | 161 | -165 | 89 Gly | -33 | -79 | -170 |
| 39 Lys | -104 | -176 | -179 | 90 Glu | -66 | -40 | -168 |
| 40 Thr | -100 | 127 | -160 | 91 Arg | -77 | -69 | 170 |
| 41 Gly | 146 | -155 | 168 | 92 Gln | -30 | -72 | -175 |
| 42 Gln | 13 | 108 | 168 | 93 Asp | -53 | -39 | 172 |
| 43 Ala | -155 | 8 | 172 | 94 Leu | -51 | -73 | -179 |
| 44 Glu | 122 | 70 | -174 | 95 Val | -32 | -47 | 173 |
| 45 Gly | 87 | 86 | -158 | 96 Ala | -66 | -37 | -178 |
| 46 Tyr | -115 | -136 | -163 | 97 Tyr | -44 | -58 | -173 |
| 47 Ser | 139 | 126 | -172 | 98 Leu | -68 | -63 | 172 |
| 48 Tyr | 113 | 155 | -175 | 99 Lys | -51 | -34 | 165 |
| 49 Thr | -113 | -73 | 179 | 100 Ser | -77 | -52 | -174 |
| 50 Asp | -102 | 150 | -166 | 101 Ala | -80 | -56 | -167 |
| 51 Ala | 103 | 125 | 177 | 102 Thr | -56 | -70 | -177 |
| 52 Asn | 85 | -83 | 153 | | | | |

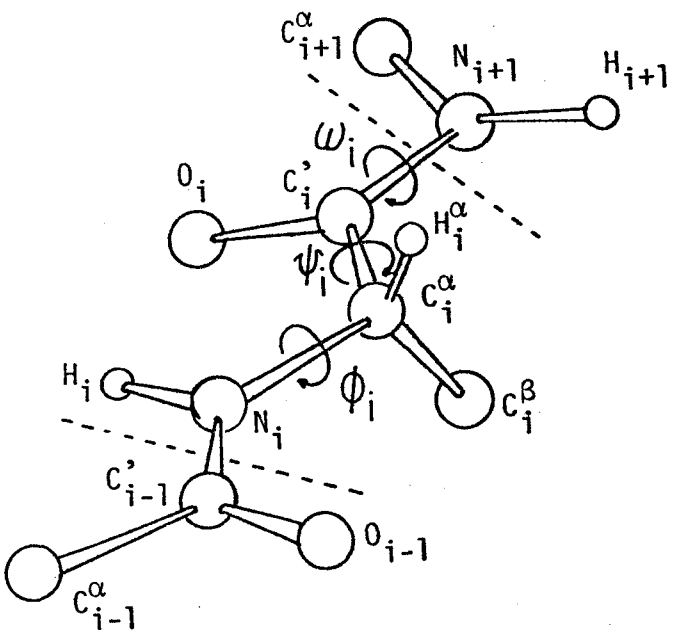



Fig.20 Definition of twisting angles(45,46).

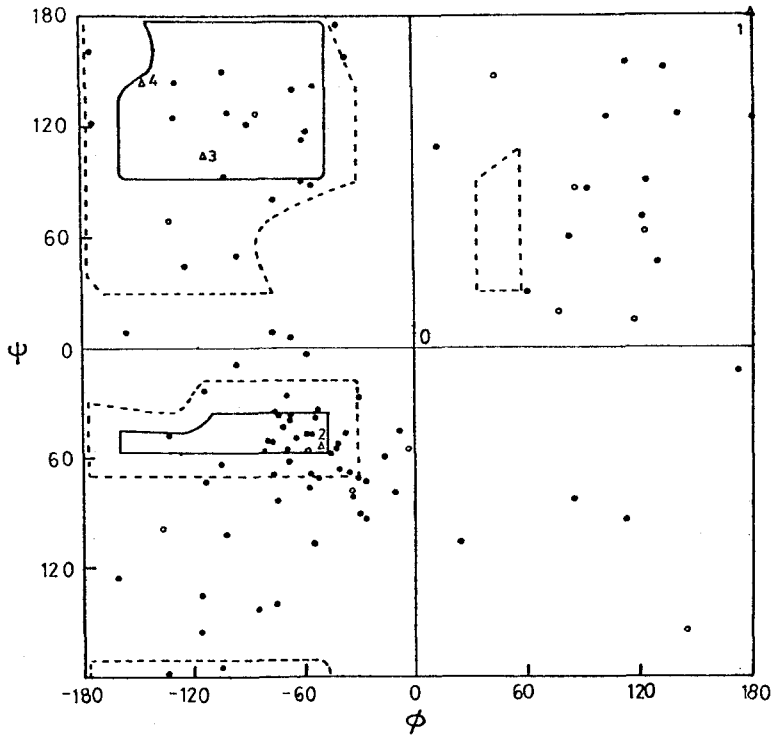



Fig.21 (ϕ,ψ) plot.

Δ1 Staggered form. Δ2 Right handed α-helix(48).

Δ3 Parallel pleated sheet structure(46).

Δ4 Antiparallel pleated sheet structure(46).

Table 10 Hydrogen bonding pattern

(a) Hydrogen bonds in the main chain

| C=O | NH | Distance(Å) | C=O | NH | Distance(Å) |
|---|---|---|---|---|---|
| Gly1 | Lys5 | 3.5* | Asn62 | Glu66 | 2.6*[2] |
| Asp2 | Gly6 | 3.0* | Thr63 | Tyr67 | 2.7*[2] |
| Val3 | Lys7 | 3.0* | Leu64 | Leu68 | 2.4*[2] |
| Ala4 | Lys8 | 3.5* | Met65 | Glu69 | 2.6*[2] |
| Lys5 | Thr9 | 2.9* | | | |
| Gly6 | Phe10 | 2.5* | Lys87 | Arg91 | 2.9*[3] |
| Lys7 | Val11 | 2.8* | Lys88 | Gln92 | 2.6*[3] |
| Lys8 | Gln12 | 3.1* | Gly89 | Asp93 | 2.7*[3] |
| Thr9 | Lys13 | 2.8* | Glu90 | Leu94 | 2.6*[3] |
| | | | Arg91 | Val95 | 3.2*[3] |
| Cys14 | Cys17 | 2.9† | Gln92 | Ala96 | 3.0*[3] |
| Arg38 | Trp59 | 2.6 | Asp93 | Tyr97 | 2.8*[3] |
| Thr49 | Ser54 | 2.6 | Leu94 | Leu98 | 3.3*[3] |
| Lys53 | Gly56 | 2.5† | Val95 | Lys99 | 3.3*[3] |
| Ilu75 | Thr78 | 2.7† | Ala96 | Ser100 | 3.1*[3] |
| | | | Tyr97 | Ala101 | 2.8*[3] |
| Asn60 | Leu64 | 3.3*[2] | Leu98 | Ser103 | 2.7*[3] |
| Glu61 | Met65 | 3.2*[2] | | | |

* These are in the H-1 helix.

*[2] In the H-2 helix.

*[3] In the H-3 helix.

† These are in the $3_{10}$ bend.

(b) Hydrogen bonds related to side chains

(CO) is a carbonyl group, and (NH) is an amide group of each peptide linkage. The last column shows the distances in the present bonito ferrocytochrome c. The data of horse and tuna are from Dickerson's table(16).

Inner P and outer P mean the inner and outer propionic residues, respectively.

Table 10 continued

| | Horse | Tuna | Bonito | Distance |
|---|---|---|---|---|
| Tyr46 | | outer P | inner P | 2.8 |
| 48 | inner P | inner P | Thr78 | 3.2 |
| 67 | Thr78 | Thr78 | Thr78 | 3.2 |
| | | | $H_2O$(W1) | 2.5 |
| 74 | | | Thr63 | 2.3 |
| Thr19 | NH25 | NH25 | CO25 | 2.8 |
| 40 | CO55 | CO55 | Asn52 | 3.0 |
| 49 | outer P | outer P | | |
| 63 | | Asn60 | Tyr74 | 2.3 |
| 78 | Tyr67 | Tyr67 | Tyr67 | 2.9 |
| | | | Tyr48 | 2.6 |
| | | | outer P | 2.8 |
| His18 | CO30 | CO30 | CO30 | 3.4 |
| 26 | CO44 | CO44 | CO44 | 2.8 |
| | | | NH31 | 2.6 |
| Trp33 | | | CO43 | 3.5 |
| 59 | inner P | inner P | inner P | 3.1 |
| Asn31 | | | CO21 | 2.6 |
| | | | Thr19 | 2.7 |
| 52 | outer P | inner P | inner P | 2.9 |
| | | | Thr40 | 3.0 |
| | | | $H_2O$(W1) | 2.5 |
| Gln16 | | | CO11 | 2.5 |
| 42 | Arg38 | | | |
| Arg38 | Gln42 | | CO41 | 3.0 |
| | | | inner P | 2.5 |

ii Helix

There is a loose loop at the bottom of the model. Only the hydrogen bond in the loop(NH54···CO49) is not that in the regular α-helix. The direction of the carbonyl groups of Asp50 and Ala51 is against the progress of the loop.

The pleated sheet structure is not seen in this model. There are a few residues of which $(\phi,\psi)$ pairs are close to that of the pleated sheet conformation in Fig.21. The fragment His18 to Glu21 folds back nearly parallel to Gly24 through Lys27, though hydrogen bonds between the peptide groups of the fragments are not found. Only one hydrogen bond is formed between peptide groups at the lower rear of the molecule(NH59···CO38), except for the hydrogen bonds in the helices and the $3_{10}$ bends. The fragment Ph36 through Thr40 is stabilized by the hydrogen bond with the fragment Ilu57 through Asn60. Other residues in these fragments have no interactions with each other.

*α-helix*

The regions of α-helix are found between the N-terminus and Lys13(H-1), between Asn60 and Glu69(H-2), and between Lys87 and Thr102(H-3)(Fig.18). They form regular hydrogen bonds respectively(Table 10(a)). The H-1 helix, about 2.5 turns, runs from the upper rear of the molecule to the front. The third turn of the helix is distorted by the irregular conformation of Cys14 $C^{\alpha}$ atom of which the side chain forms a thioether linkage to the porphyrin ring. The H-2 helix, about 2 turns, locates at the back.

The H-3 helix about 4 turns, is at the upper rear part. The helical axis of the H-3 slightly distorts upwards at the third turn, at which the helix collides with the aromatic ring of Phe36. The average angle pairs ($<\phi>$,$<\psi>$) of the helices are (294°,318°) for H-1, (312°,301°) for H-2, and (306°,303°) for H-3.

*$3_{10}$ conformation*

There are 7 hair pin loops in the molecule, at which the main chain folds back upon itself(Table 11). Three of the loops form hydrogen bonds from an amide group(Donor) to a carbonyl group(Acceptor) in the third residue back along the chain. This loop is called a $3_{10}$ conformation, The $3_{10}$ conformation has three variants which are sterically possible(49). In type I, the carbonyl group of the second residue and the side chain of the third are in the opposite direction. Whereas in type II, the carbonyl group and the side chain are in the same direction. To keep the steric hindrance off, type II is possible only when Gly is the third of the four amino acid residues involved. The hair pin loop (ii) at Asn22 and Gly23 in Table 11, is associated with the hydrogen bonds between the side chain of Asn22 and the peptide bonds of the two residues. The absence of the hydrogen bond of a $3_{10}$ conformational type may be owing to the large freedom of the rotation due to glycine residues. The type III loop, $3_{10}$ helix, in which every NH···CO hydrogen bond is that involved in type I or II is not seen in this protein. The regions of α-helix and $3_{10}$ bend are also shown in Fig.3.

Table 11 $3_{10}$ conformations in hair-pin loops

| Number | Residues | Bonito Type | Horse Type |
|---|---|---|---|
| (i) | 14--17 | I | |
| (ii) | 21--24 | | II |
| (iii) | 28--31 | | |
| (iv) | 35--38 | | II |
| (v) | 42--45 | | |
| (vi) | 53--56 | I | * |
| (vii) | 75--78 | II | II |
| | 61--64 | ** | I |
| | 66--69 | ** | I |
| | 67--70 | ** | I |

* It is in the lower helix in horse.

** They are in the H-2 helix in bonito.

iii Distribution of side chains

The amino acid residues are classified by the characters of their side chains. The total numbers of the residues in each group of this protein (in parentheses) are as follows;

1) hydrophilic side chains(50)

acidic(9) ····· Asp(3), Glu(6).

basic(20) ····· Lys(16), Arg(2), His(2).

ambivalent(21) · Ser(4), Thr(7), Asn(6), Gln(4).

2) hydrophobic side chains(53)

aromatic(10) ··· Phe(3), Tyr(5), Trp(2).

aliphatic(43) ·· Gly(13), Ala(7), Val(6), Leu(6), Ilu(4), Pro(3), Cys(2), Met(2).

The environments of the side chains are summarized in Table 12.

*Hydrophilic side chains*

The acidic and basic side chains together with the terminal acetyl and carboxyl groups are located on the molecular surface, except for the imidazole group of His18. This molecule contains 16 lysyl residues, of which 9 residues protrude to the outer medium in the left rear corner of the molecule. There is a typical hydrophilic cluster on the surface at the top of the left rear, composed of Lys5, Lys8, Thr9, Gln12, Lys13, Lys86, Lys87, Lys88, Glu90, Gln92, and Asp93. The rear of the molecule is crowded with Lys55, Asn60, Glu61, Asn62, Thr63, Glu66, Glu69, Asn70, Lys72, and Lys73. These clusters close to each other. The right side and the bottom of the molecule

includes several hydrophilic residues. However any clusters are not recognized. Several hydrophilic side chains are in the interior of the molecule. They all make some hydrogen bonds.

Most ambivalent side chains are on the surface. Five of them are in the interior, all of which are engaged in the hydrogen bonds(Tables 10(b), and 12). The hydrogen bond network of the protein will be discussed later. Intramolecular salt bridges are not observed in this model.

*Hydrophobic side chains*

The hydrophobic side chains aggregate at the upper half of the molecule. The hydrophobic cluster is formed at the upper part of the heme crevice.

The leucyl and isoleucyl residues are buried inside of the molecule except for Ilu81. All the 6 varyl residues occur on the surface and exposed to solvent.

Glycines are on the surface with two exceptions, Gly6 and Gly29. Gly6 which occupies the interior interface of the H-1 helix, is in close contact with the H-3, and the insertion of any side chain other than hydrogen is not adomitted. At the 29th position, no residues can be occupied except for glycine owing to the steric hindrance against the heme group. Glycine residues often locate at the position where the main chain shows a sharp chain reversal. There are 6 glycines in the hair-pin loops of this protein. Four glycines are located in the region of the very weak density. Many glycines are outside the allowed regions of $(\phi,\psi)$ owing to the fact that

glycine has greater freedom in twisting angles.

The indole ring of Trp33 is towards the exterior on the surface. The $N^{\epsilon}$ atom of Trp33 makes a hydrogen bond with CO43. The ring of Phel0 is located at the upper right of the heme crevice together with that of Tyr97, though no overlapping is found between them. The phenyl ring of Phe82 lies roughly parallel to the heme plane at the upper left of the crevice, 4.5Å apart from the heme plane. The rings of Tyr48 and Tyr74 are on the molecular surface. The hydroxyl oxygens of Tyr48 and Tyr74 are engaged in intramolecular hydrogen bonds, and access to the rings from the exterior of the molecule is difficult.

The positions of Tyr67, Trp59, and Tyr74 are interesting in connection to the electron transfer mechanism which will be discussed later. The agreement between the model and the electron density map of this part is shown in Fig.22. The interatomic distances and dihedral angles among them are summarized in Table 13. The parameters of horse and tuna are taken from Dickerson et al.(50).

Table 12 Location of side chains

| Residue | External | | Surface | | Internal | |
|---|---|---|---|---|---|---|
| Gly(13) | 1,23,24,34, 37,41,45,56, 77,84,89 | (11) | | | 6,29 | (2) |
| Ala(7) | 4,83,96 | (3) | 15,43,101 | (3) | 51 | (1) |
| Ser(4) | 47,54,100,103 | (4) | | | | |
| Thr(7) | 19,49 | (2) | 9,63 | (2) | 40,78,102 | (3) |
| Asn(6) | 22,60,62,70 | (4) | | | 31,52 | (2) |
| Asp(3) | 2,50,93 | (3) | | | | |
| Gln(4) | 12,16,42,92 | (4) | | | | |
| Glu(6) | 21,44,61,66, 69,90 | (6) | | | | |
| Lys(16) | 5, 7, 8,13, 25,27,39,53, 55,72,73,79, 86,87,88,99 | (16) | | | | |
| Arg(2) | 91 | (1) | 38 | (1) | | |
| His(2) | | | 26 | (1) | 18 | (1) |
| Val(6) | 3,11,20,28, 58,95 | (6) | | | | |
| Leu(6) | | | | | 32,35,64, 68,94,98 | (6) |
| Ilu(4) | 81 | (1) | 57 | (1) | 75,85 | (2) |
| Pro(3) | | | 71,76 | (2) | 30 | (1) |
| Cys(2) | 17 | (1) | | | 14 | (1) |
| Met(2) | 65 | (1) | | | 80 | (1) |
| Phe(3) | 10,82 | (2) | 36 | (1) | | |
| Tyr(5) | 97 | (1) | 48,74 | (2) | 46,67 | (2) |
| Trp(2) | 33 | (1) | | | 59 | (1) |

The sequential number of the residue is given in the Table. The total number of the residue is given in parentheses.

Table 13 Interatomic distances and dihedral angles among Trp59, Tyr67, and Tyr74

(a) Interatomic distances within 5.0Å

| Trp59 | Tyr67 | Bonito | Tuna | Horse |
|---|---|---|---|---|
| C5 | C1 | | 4.8Å | |
| C5 | C2 | | 4.7 | 4.2Å |
| C5 | C3 | | 4.9 | 4.0 |
| C6 | Cβ | | 4.0 | |
| C6 | C1 | 4.8Å | 3.8 | 4.9 |
| C6 | C2 | 4.8 | 4.0 | 3.8 |
| C6 | C3 | 4.9 | 4.4 | 4.0 |
| C6 | C4 | 5.0 | 4.6 | |
| C6 | c5 | 5.0 | 4.6 | |
| C6 | C6 | 4.9 | 4.2 | |
| C7 | Cβ | 4.9 | 4.2 | |
| C7 | C1 | 4.8 | 4.5 | |
| C7 | C2 | 4.9 | 4.8 | 4.9 |

| Tyr67 | Tyr74 | Bonito | Tuna | Horse |
|---|---|---|---|---|
| Cα | Cβ | | 4.6Å | |
| Cα | C1 | 4.7Å | 4.6 | |
| Cα | C2 | 3.8 | 4.2 | |
| Cα | C3 | 3.8 | 4.7 | |
| Cα | C4 | 4.7 | | |
| Cβ | C1 | | 4.8 | |
| Cβ | C5 | | 4.2 | |
| C1 | C5 | 4.2 | 4.6 | |
| C1 | C6 | 4.2 | 4.1 | |
| C5 | C2 | 4.5 | 4.3 | |
| C6 | Cβ | 4.9 | 4.1 | 4.9Å |
| C6 | C1 | 4.7 | 4.1 | |
| C6 | C2 | 3.6 | 4.1 | |
| C6 | C3 | 4.1 | 4.2 | |

| Trp59 | Tyr74 | Bonito | Tuna | Horse |
|---|---|---|---|---|
| C6 | C2 | 4.6 | 4.7 | 4.7 |
| C6 | C3 | 4.0 | | 5.0 |
| C6 | C4 | 4.9 | | |
| C6 | OH | 5.0 | | |
| C7 | C2 | | 4.7 | 4.9 |
| C7 | C3 | | 4.3 | 4.8 |
| C7 | OH | 4.0 | 5.0 | |
| C8 | C3 | 4.7 | | |
| C8 | OH | 4.8 | | |

The nomenclature of each atom is shown in Fig.23.

Table 13 continued

(b) Dihedral angles between the best planes.

| | Bonito | Tuna | Horse |
|---|---|---|---|
| Heme —— Trp59 | 67° | 54° | 65° |
| Heme —— Tyr67 | 87 | 67 | 25 |
| Heme —— Tyr74 | 95 | 80 | 49 |
| Trp59 — Tyr67 | 21 | 59 | 65 |
| Trp59 — Tyr74 | 45 | 41 | 16 |
| Tyr67 — Tyr74 | 39 | 54 | 53 |

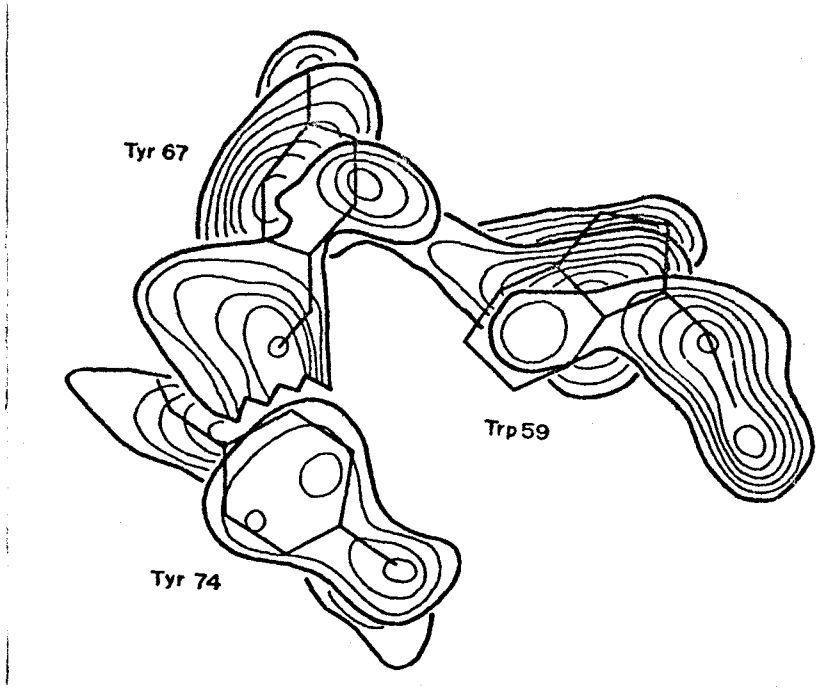



Fig.22 Agreement between aromatic rings(Trp59,Tyr67, and Tyr74) and the composite electron density map. First contour 0.2eÅ$^{-3}$, interval 0.1eÅ$^{-3}$.

iv Location of the heme

The heme is rigidly constrained in a deep crevice surrounded by the following fragments of the main chain; upper part by the H-1 helix, upper rear by the H-3, rear by the H-2, right (5th ligand) side by Pro30 to Phe36, left (6th ligand) side by Ilu81 to Ala83, bottom by Tyr46 to Lys53. The crevice is an egg shape, about 20×15×12Å on the size. Only one edge of the heme(front) is exposed to the outer medium, and accessible to the solvent as shown in Fig.14.

The side chain of Lys79 runs from left to right on the lower front of the molecule, and prevent the heme interacting the external. The ring of Phe82 in the left crevice stands roughly parallel to the heme plane, blocking the heme plane from the external. Lys13 is at the upper left of the crevice and in van der Waals contact with Phe82 ring. The $N^{\varepsilon}$ atom of Lys13 may form a hydrogen bond with Glu90. Wada and Okunuki have reported the important role of Lys13 in the activity of this protein(51).

There is a large channel opened to the external at the upper right of the heme crevice. A small molecule or an ion may slip from the molecular surface into the heme crevice through the channel, though the channel is populated by the hydrophobic residues, Phe10, Val20, Pro30, Leu32, Trp33, Tyr97, and Leu98.

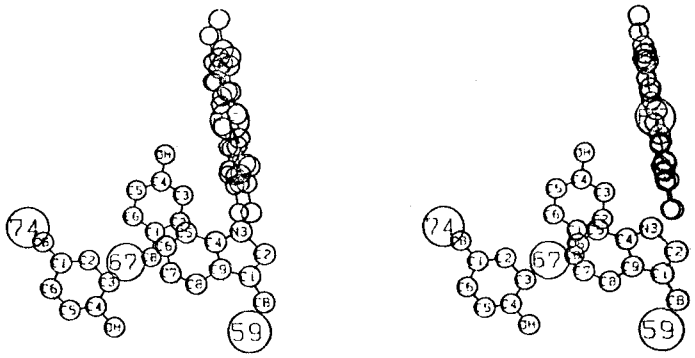
The hydrophobic cluster is formed at the upper rear in the crevice by Leu32, Leu35, Phe36, Trp59, Leu64, Tyr67, Leu68, Ilu85, Leu94, Leu98, and Ala101. Tyr67 is in van

der Waals contact with the heme plane, but the two planes are roughly perpendicular with each other. Two aromatic rings of Tyr67 and Trp59 contact edge to edge on a plane, and the interaction between them is van der Waals contact. The locations and the orientations of these aromatic rings and the heme plane of this protein are shown in Fig.23.

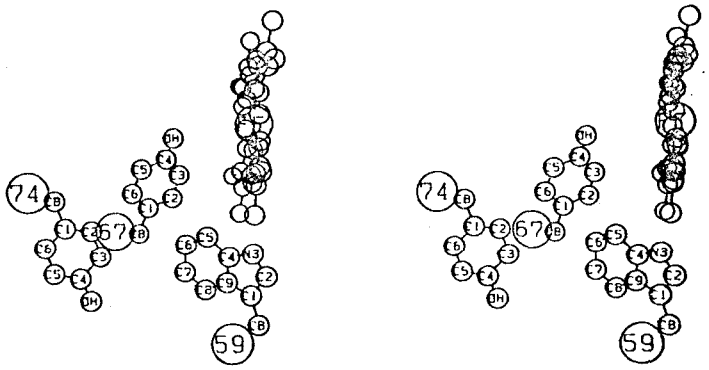
v Water molecules

The main chain of bonito ferrocytochrome c folds itself as compactly as possible, and it seems to be difficult for solvent molecules to come into the interior of the molecule. However as mentioned in chap.V, some isolated peaks which occur in the electron density map may represent fixed solvent molecules. The peaks are all in the lower half of the molecule, and make two or three hydrogen bonds with the main chain. They may be water molecules. Their coordinates and contacts are summarized in Table 14.

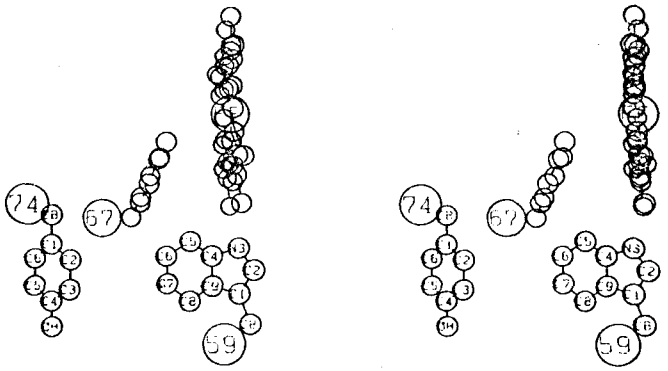
W2 is hydrogen bonded to CO51. This is one of the reasons why the bottom loop of the molecule is not an α-helix. W3 forms a hydrogen bond with Lys79. It is suggested that the side chain of Lys79 is fixed strongly by the hydrogen bond, because the electron density of the side chain of Lys79 is very higher than most of the ly ine side chains.




(a) Bonito ferrous form.




(b) Tuna ferrous form.




(c) Horse ferric form.

Fig.23 Stereoscopic drawings of the relative locations of heme, Trp59, Tyr67, and Tyr74.

Table 14 Location of water molecules

| Site | x/a | y/b | z/c | Location | Hydrogen bonded to |
|---|---|---|---|---|---|
| W1 | 0.286 | 0.110 | 0.460 | I** | Asn52, Tyr67 |
| W2 | 0.341 | 0.089 | 0.689 | I*** | CO47, CO51 |
| W3 | 0.355 | 0.020 | 0.500 | S | Ser47, Lys79 |
| W4 | 0.466 | 0.139 | 0.574 | S | CO39, CO42*, CO43* |

* Bifurcated hydrogen bonds.

** Close to Thr78.

*** Close to Tyr46.

Table 15 Binding sites of heavy atoms

| Heavy atom | | Binding site |
|---|---|---|
| $UO_2F_5^{3-}$ | | Lys13, Glu90 |
| $(CH_3)_2SnCl_2$ | (major) | Cys17 |
| | (monor) | Glu61, Lys99 |
| $IrCl_6^{3-}$ | | Asn60, Glu61 |
| $UO_2(CH_3COO)_2$ | | Lys13, Glu90 |
| $HgI_4^{2-}$ | (major) | Cys17 |
| $PtCl_4^{2-}$ | | Asn62, Met65 |

## VII HEAVY ATOM BINDING SITES

The binding sites of the heavy atom groups are summarized in Table 15. The main binding site of $UO_2(CH_3COO)_2$ is the same as that of the U derivative.

The $K_2HgI_4$ derivative seemed to be suitable at first. However the parameters of the heavy atoms could not be refined well. It was made clear in the course of the structure determination that the protein was modified by the reagent. The heavy atom located closely at the sulfur atom of Cys17. It may hydrolyze the thioether bond. On the other hand the reagent binds near Cys17, with isomorphism conserving in tuna ferrous crystal(16,30). This is attributed to the difference of the molecular packings. In tuna crystal, there is room for the isomorphous replacement of heavy atoms near Cys17 in contrast to bonito one. The main binding site of the Sn derivative was nearly the same as that of the $K_2HgI_4$ derivative. The parameters of the Sn atoms could be refined reasonably.

The heavy atoms of the Pt derivative located in the vicinity of the sulfur atom of Met65. When the Pt derivative was included in the calculation of the electron density, the figure of merit was graded up to some extent. However, in the map some noises appeared around the Pt site (Asn62, Met65), and gave a difficulty in the interpretation of this region. If the derivative was omitted from the calculation, the noises disappeared. Therefore it was excluded from the calculation.

## VIII DISCUSSION

### i Structural differnces between ferrous and ferric cytochrome c

It is intersting to distinguish the structural differences between the reduced and oxidized states of cytochrome c, in connection with the electron transfer mechanism. As the crystals of cytochromes are prepared in a salt rich solution(17,29,30), the molecular structures determined by x-ray analyses may differ from the *in vivo* states. However it was surprised at the structural coincidence among the models of horse ferri-, tuna ferro-, bonito ferro- and ferricytochrome c's, and R.rubrum ferricytochrome $c_2$(Figs.19, and 24). Thus the conformation of the models is very stable one of the cytochrome C's. The proteins have the activity in the crystalline states. These facts may show that the conformation in crystalline state is not so differ from that *in vivo* state, though there may be some dynamic fluctuations.

In order to make a comparison of bonito ferrous protein with horse ferric one, the horse model was rotated to minimize $\Sigma\Delta r_i^2$, using the $C^\alpha$ parameters of both bonito and horse models by Ohi(52), where $\Delta r_i (=C^\alpha_{bonito}-C^\alpha_{horse})$ is the discrepancy between the $i$th $C^\alpha$ atoms of bonito and horse molecules(Table 16). The large discrepancy at the C-terminus may be attributed to the ambiguity of the C-terminus itself. The average of the discrepancies is 1.7Å. Though the significant change of the molecular sizes from that of the oxidation states was not considered in the

calculation, the $C^{\alpha}$ parameters of bonito coincide with those of horse fairly well. The bonito molecule is 35×30×23Å large. The horse is 35×30×25Å, and the dimensions of the tuna are same as bonito ones(16). It was made clear that the shrinking by reduction is subtle. As a whole the folding pattern of the four cytochrome c's and cytochrome $c_2$ resemble very much to each other. The helices are located at similar positions, of which the H-2 is a less perfect one in horse and tuna. The lower helix(49 to 54) was not detected in bonito. The total helix content is 37% in bonito, and 41% in horse and tuna.

Several $3_{10}$ bends are found in bonito ferrous and horse ferric protein(Table 11). Lys53 to Gly56 is a type I in bonito, and it is contained in the lower helix in horse. Glu61 through Asn70 is contained in the H-2 in bonito, though the part takes continued type I conformation($3_{10}$ helix) in horse. Ilu75 through Thr78 is consistent with the result of horse(type I). Cys14 to Cys17 takes a type I, and Glu21 to Gly24 and Leu35 to Arg38 do not take $3_{10}$ conformation in bonito.

The residues 41-42, 46-50, 54-56, and 81-82 of bonito are slip out more than 3Å from those of horse(Table 16). Gly41 and Gln42 go upwards in bonito, though they go downwards in horse and tuna. However CO44 is hydrogen bonded to His26 in bonito. The hydrogen bond is detected in both horse and tuna. This part may be important to exhibit the function of the molecule(53). The structural difference of Tyr46 through Asp50 between bonito and others

gives rise to the orientation difference of the side rings of Tyr(Phe)46 and Tyr48. Therefore the hydrogen bond scheme is different. In bonito, Tyr48 makes a hydrogen bond with Thr78, while in horse and tuna, Tyr48 makes a hydrogen bond with inner P. Tyr46 in bonito makes a hydrogen bond with inner P. That in tuna makes a hydrogen bond with outer P. In horse Phe46 can not have a hydrogen bond, but it takes almost the same position in tuna. In the model building, it was most difficult to determine the structure from Tyr48 to Ser54, because of the confused electron density map.

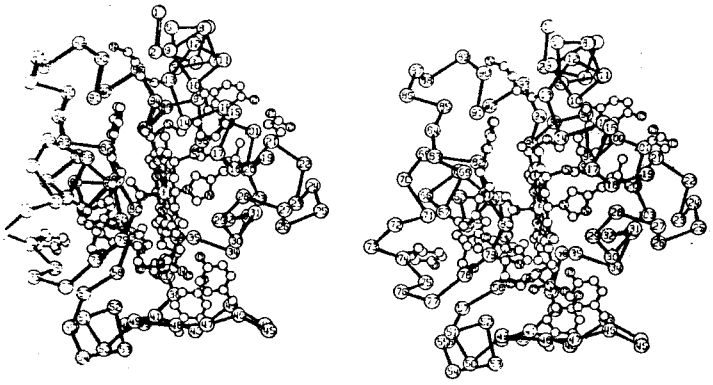
The large conformational change was detected at Phe82. The side chain of Ilu81 goes outerwards, and the phenyl ring of Phe82 swings into the heme crevice in the reduced state(Fig.18). On the contrary the side chain of Ilu81 goes innerwards, and the ring of Phe82 swings out in the oxidized state. The conformational change is interpreted in terms of the structural specificity of the main chain folding of this protein. The peptide chain of Ilu81 through Lys86 is flexible, because it has neither a hydrogen bond nor a close van der Waals contact with other fragments of the molecule. The large distortion caused by the conformational change between the two oxidation states, was cancelled at Gly84, which has the maximum freedom of rotation in all the amino acid residues. This residue is structure invariant as same as Phe82. The conformation of Thr78 through Gly84 in bonito is the same in tuna ferrous protein.

In this model, there is a large channel at the upper right of the heme crevice. The side chain of Trp33 goes downwards and exposed into solvent. The channel is, however, closed in tuna. The hair-pin loop at Glu21 and Asn22 slides up, and the big indole ring of Trp33 goes upwards to close the channel in tuna. The channel is open in horse. The imidazole ring of His33 goes downwards and exposed in horse, like in bonito. There are only two substitutions in the amino acid sequence of bonito and tuna, while there are 17 substitutions in bonito and horse. However the bonito and tuna ferrous crystals are not isomorphous, but the tetragonal forms of both bonito and horse ferric crystals are isomorphous. The ferric proteins, therefore, seem to have same conformations. It is suggested that the difference of right channels in the structures is not attributed to the difference of the species, but to the packing schemes of the molecules in both crystals. Or it may have something to do with the oxidation and reduction scheme of cytochrome c in various conditions, as suggested by Dickerson et al.(16).

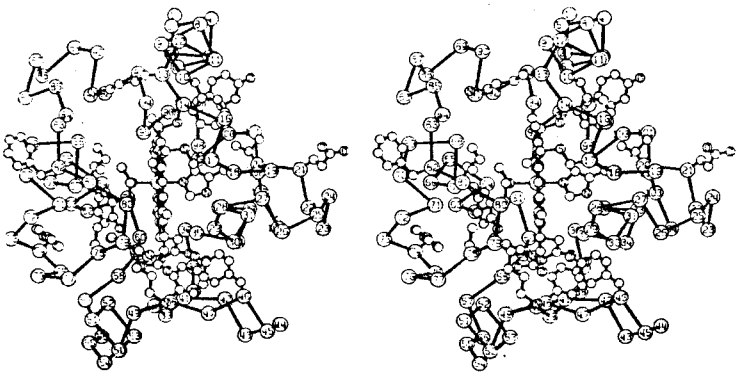
Several minor changes between ferrous and ferric cytochrome c were found. One of the changes is that the indole ring of Trp59 shifts upwards so as to approach the heme methyl at 3.0Å in tuna, while the ring is 5Å apart from the heme methyl in horse. In this analysis, the bulky electron density hump occurs at the position where the indole is assigned, and the ring is located at the center of the density hump, by considering the linkage

with the main chain. It is 4.5Å apart from the heme methyl. Another is the orientation of the phenol ring of Tyr74. The ring is toward the exterior in horse. On the other hand, it is on the surface, roughly perpendicular to the molecular surface in bonito and tuna. Furthermore it is hydrogen bonded to Thr63 in bonito. The ring plane, therefore, does not access to the outer medium in bonito, since its lower side is protected by the side chain of Lys55, and the upper by the H-2 helix. That is, the location of the indole ring of Trp59 in bonito is similar to that in horse rather than in tuna, and the orientation of the ring of Tyr74 in bonito is resembled that in tuna rather than in horse. The indole ring of Trp59 is not parallel to the phenol ring of Tyr74 in bonito. They are not parallel in tuna, while they are parallel in horse.

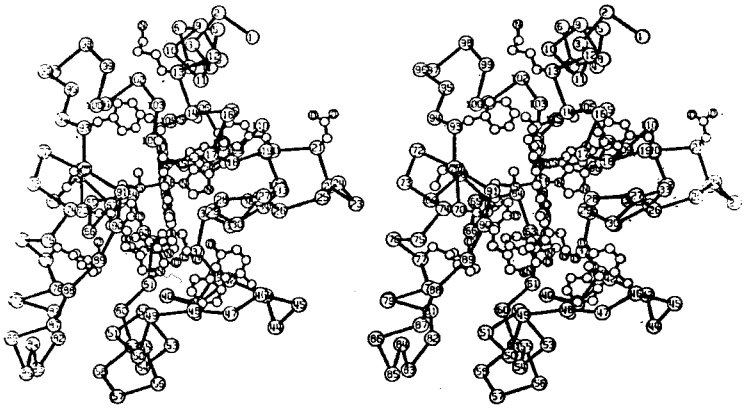
The internal water molecule is detected in the space between Asn52 and Tyr67(Fig.18). The fixed water molecule W1 was also found in tuna, but was not reported in horse.



(a) Tuna ferrocytochrome c(16).



(b) Horse ferricytochrome c(15).



(c) R.rubrum ferricytochrome $c_2$(17).

Fig.24 Stereoscopic drawings of other cytochromes.

Table 16 Discrepancies between the $C^{\alpha}$ atoms of bonito ferrocytochrome c and horse ferricytochrome c

| | $\Delta r_i$ | | $\Delta r_i$ | | $\Delta r_i$ | | $\Delta r_i$ | | $\Delta r_i$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 2.1 | 22 | 3.3 | 43 | 3.3 | 64 | 1.0 | 85 | 1.2 |
| 2 | 2.0 | 23 | 0.3 | 44 | 1.4 | 65 | 1.3 | 86 | 0.9 |
| 3 | 1.3 | 24 | 0.9 | 45 | 1.1 | 66 | 2.3 | 87 | 0.7 |
| 4 | 0.2 | 25 | 1.1 | 46 | 3.3 | 67 | 1.8 | 88 | 0.6 |
| 5 | 0.8 | 26 | 1.4 | 47 | 3.2 | 68 | 2.7 | 89 | 1.1 |
| 6 | 0.6 | 27 | 0.9 | 48 | 6.0 | 69 | 2.6 | 90 | 0.5 |
| 7 | 0.7 | 28 | 1.7 | 49 | 5.5 | 70 | 1.3 | 91 | 0.9 |
| 8 | 0.7 | 29 | 1.3 | 50 | 3.2 | 71 | 1.3 | 92 | 0.8 |
| 9 | 0.3 | 30 | 1.0 | 51 | 0.5 | 72 | 1.6 | 93 | 0.9 |
| 10 | 1.1 | 31 | 2.2 | 52 | 1.1 | 73 | 1.7 | 94 | 0.6 |
| 11 | 0.6 | 32 | 1.7 | 53 | 2.3 | 74 | 0.8 | 95 | 0.8 |
| 12 | 0.4 | 33 | 2.0 | 54 | 2.8 | 75 | 1.2 | 96 | 0.4 |
| 13 | 1.3 | 34 | 2.2 | 55 | 3.3 | 76 | 2.5 | 97 | 0.8 |
| 14 | 0.8 | 35 | 1.8 | 56 | 3.2 | 77 | 2.2 | 98 | 0.3 |
| 15 | 2.6 | 36 | 2.1 | 57 | 2.0 | 78 | 0.5 | 99 | 0.4 |
| 16 | 2.6 | 37 | 1.4 | 58 | 3.1 | 79 | 0.9 | 100 | 1.5 |
| 17 | 1.2 | 38 | 1.1 | 59 | 0.8 | 80 | 1.3 | 101 | 1.3 |
| 18 | 0.9 | 39 | 1.0 | 60 | 1.0 | 81 | 3.6 | 102 | 2.7 |
| 19 | 1.0 | 40 | 1.1 | 61 | 1.3 | 82 | 4.0 | 103 | 5.1 |
| 20 | 2.6 | 41 | 2.1 | 62 | 1.2 | 83 | 1.5 | | |
| 21 | 2.7 | 42 | 3.7 | 63 | 1.3 | 84 | 1.7 | | |

$\Delta r_i$ is given in the unit of Å.

ii Electron transfer mechanism of cytochrome c

*Dickerson's model*

From the structure analysis of horse and tuna cytochrome c, Dickerson et al. proposed the model that the overlap of the three aromatic rings plays an essential role as the electron pathway from the reductase through Tyr74, Trp59, and Tyr67 to the heme(16). The model is outlined in Fig.25. In the horse ferric state, Tyr74 and Trp59 are parallel, and Tyr67 and heme plane are more or less parallel(Table 13). Whereas Trp59 and Tyr67 are 'decoupled' from their former partners in the tuna ferrous state, and are themselves parallel. In the first step, one electron is transferred from a reductase to Tyr74, and from there to Trp59 by π-electron cloud overlap. Another electron is transferred from Tyr67 to heme, leaving Tyr67 with an electron deficiency(Fig.25(a)). An electron flow from the smaller delocalized aromatic ring system to the larger one will be potentially downhill, as the transferred electron occupies the lower first antibonding orbital. In the next step, a conformational change occurs from the observed ferric structure to a ferrous one, bringing Trp59 and Tyr67 into parallelizm and π-cloud overlap. An electron is then transferred from Trp59 to Tyr67, leaving the molecule in the observed ferrous conformation(Fig.25(b)). The transfer is enhanced because it eliminates an ion pair in a medium of low dielectric constant on the interior of the molecule as described later. This model may be also supported by the effect of

the chemical modification of Trp59 by N-bromosuccinamide (54,55) that the indole group of Trp59 plays a key role for the reductase activity, but not for the oxidase activity.

The ring overlap between Tyr74 and Trp59 is, however, not found in horse, from the stereoscopic view of the rings (Fig.23). Considering the quantum mechanics of the charge transfer process between two molecules, Yomosa(56) concluded that the energy of a charge transfer interaction is proportional to $S_{ab}/R_{ab}$, where $S_{ab} = \int \Phi_a \Phi_b d\tau$ is an overlap integral between the donor's molecular orbital ($\Phi_a$) of the molecule $a$ and the acceptor's molecular orbital ($\Phi_b$) of the molecule $b$, and $R_{ab}$ denotes the effective distance between two orbitals. The molecular orbital is generally represented as a linear combination of the component atomic orbitals $\phi$, that is, $\Phi_a = \Sigma_i c_i^a \phi_i^a$ and $\Phi_b = \Sigma_i c_i^b \phi_i^b$. The overlap integral $S_{ab}$ is formulated as

$$S_{ab} = \Sigma\Sigma_{ij} c_i^a c_j^b \int \phi_i^a \phi_j^b d\tau = \Sigma\Sigma_{ij} c_i^a c_j^b \langle \phi_i^a | \phi_j^b \rangle,$$

where $c_i^a$ is the coefficient of each atomic orbitals of a molecule $a$. The estimation of the interaction between Trp59 and Tyr74 in horse ferric form was carried out with Hueckel approximation, where the atomic parameters were taken from Dickerson et al.(50), and the parameters of overlap and Coulomb integrals being taken from Hoffman and Ladik(57). The value of $\langle \phi_i^a | \phi_j^b \rangle$ is inversely proportional to the distance between related atoms, so that $S_{ab}$ can roughly estimated from the summation of the overlap integrals between adjacent atoms. The distances

$C_6^{59}...C_2^{74}$, $C_6^{59}...C_3^{74}$, $C_7^{59}...C_2^{74}$, and $C_7^{59}...C_3^{74}$ in horse ferric form are all nearly equal, and are shorter than the other C...C distances involving Trp59 and Tyr74 (Table 13). Therefore $\langle\phi_i^{59}|\phi_j^{74}\rangle$ can be approximated to be constant for the above mentioned pairs, and may be neglected for the other pairs. It is suggested that Tyr74 accepts an electron in the lowest unoccupied orbital(No.5 in Table 17) from a reductase(16). Then the electron goes towards the lowest antibonding orbital of Trp59(No.6 in Table 17). The $S_{ab}$ of this process is as follows(a=Trp59, b=Tyr74);

$$\begin{aligned} S_{ab} &= (c_6^{59}.c_2^{74} + c_6^{59}.c_3^{74} + c_7^{59}.c_2^{74} \\ &\quad + c_7^{59}.c_3^{74})\langle\phi^a|\phi^b\rangle \\ &= (c_6^{59} + c_7^{59})(c_2^{74} + c_3^{74})\langle\phi^a|\phi^b\rangle \\ &= (-0.31-0.25)(-0.50+0.50)\langle\phi^a|\phi^b\rangle \\ &= 0.0. \end{aligned}$$

The rough estimation suggests that the electron transfer from Tyr74 to Trp59 can not occur. It must, however, be emphasized that this calculation is based on the steric structure of this protein, and no care is taken of the dynamical changes of the relative location of the residues.

*Kraut's model*

Kraut et al. speculated from the analysis of ferricytochrome $c_2$ that reduction would take place by direct transfer of an electron from a reductase to the heme(17, 21). The proposed mechanism of cytochrome $c_2$ is shown in Fig.26. In the ferric molecule, the heme iron has its unpaired spin located in $d_{xz}$ orbital(Fig.26(a)). Stabiliza-

tion of the iron charge is achieved by a charge interaction between the sulfur atom of Met91 and the oxygen atom of Tyr70. An electron is transferred from an approaching moiety of a reductase, such as a metalloporphyrin interacting with the heme of cytochrome $c_2$, at the front exposed edge of cytochrome $c_2$, to the $d_{xz}$ orbital of the heme. The reduction process may be facilitated by the simultaneous approach of a proton donor group of the reductase to Ser89. This tends to shift the hydrogen bonding scheme to that postulated for the ferrous state(Fig.26(b)). Reoxidation of the reduced heme takes place essentially by reversing the reduction process. Oxidation may be facilitated by the approach of a proton withdrawing group to Ser89 of cytochrome $c_2$.

*Author's model*

The extending hydrogen bond network is found in the bonito protein. The hydrogen bonds are between Trp59... inner P, Tyr67...Thr78, Thr78...outer P, inner P...Asn52, Arg38...inner P, and probably Asn52...Thr78. Above hydrogen bonds are observed in the three cytochrome c's. In horse, the hydrogen bond between Asn52...outer P is observed instead of that between Asn52...inner P in bonito and tuna. In the ferrous molecule, there exist the hydrogen bonds between Tyr67...W1, and between W1...Asn52. These residues are evolutionally conserved, so that the hydrogen bond network is essential in the protein. It is possible to speculate another model in which the hydrogen bonds play a key role. The proposed overall scheme for

the reduction and oxidation of the cytochrome c based on the bonito ferrocytochrome c structure is shown in Fig.27. The theoritical treatment shows that the charge transfer complex takes an ordinary ground state(F-state, AB) in a less polar solvent, and in a polar solvent the polarized form(S-state, $A^+B^-$) is stabilized by the interaction between the dipole of the complex and the reaction field due to the polarization of the medium(58). From the structure analysis, the interaction between Tyr67 and heme can be expected, because the rotation of Tyr67 around the $C_\beta$-$C_\gamma$ bond makes some interaction with the heme plane, or because, as Kraut reported in cytochrome $c_2$, Tyr67 interacts with heme through the sulfur atom of Met80, which is in contact with the oxygen of Tyr67. In the reduced state, this pair is in a less polar surroundings, as Phe82 covers the left channel found in the ferric state(Fig.27). At the first step of oxidation, Phe82 moves out in the exterior and the coupled pair, heme...Tyr67(Fstate), is exposed to polar solvent, so that the pair is polarized to become S-state($heme^+$...$Tyr67^-$). The negative charge on Tyr67 is transferred to the molecular surface through the hydrogen bonds. The residues, Arg38 and Thr78, are on the molecular surface, joining the hydrogen bond network. Inner P can not accept the charge on Tyr67 through a solvent molecule and Asn52, because it is negatively charged by releasing the proton to Arg38 through the hydrogen bond, so that the electron on Tyr67 can not flow to Arg38 through inner P. This means that Arg38 is not the residue to interact

with an oxidase which takes up an electron from cytochrome c. Thr78 on the molecular surface may be a residue to accept the electron from Tyr67. Now the charge on Thr78 may flow to Tyr48 which makes a hydrogen bond with Thr78. The successive steps are modulations of the hydrogen bonding pattern.

The trigger of the reduction may be a removal of the positive charge on Arg38 in contact with a reductase. The negative charge on inner P would migrate in Tyr67 through the hydrogen bonds, when the positive charge on Arg38 is neutralized by the reductase. The successive movement of Phe82 changes the circumstance of heme, leading the coupled pair of S-state into F-state. Fixing a water molecule and the rearrangement of the hydrogen bonds make the protein to the conformation of the ferrous state.

In this model, the oxidation process does not involve the hydrogen bonded chain linked with Trp59, but the reduction process contains the hydrogen bonded chain concerned with Trp59. This model agrees, therefore, with the result of the chemical modification of Trp59(54,55).

This electron transfer model shows that the location of Phe82, a unique difference between the two oxidation states, may have an important sense for the redox process, and that only slight conformational changes of the molecule suffice for converting one state to the other.

This speculation is, however, based on the molecular structure at this stage. Therefore some ambiguities can not help remaining due to the ambiguites of the molecular

structure itself. The mechanism of an electron transfer through a hydrogen bond network was known in chymotripsin as the charge relay system(59). However the aurthor can not deny the charge transfer model through a π-orbital overlap(Dickerson's or a quite different type), since the molecular structure can fluctuate in solvent. In order to check the availability of aurthor's model, it is necessary to study the activity of the molecule of which related residues, such as Arg38 or Asn52 etc., are chemically modified.

Another hydrogen bonded chain from His18 through His26 to CO44 is in front of the molecule(Fig.28). The recent NMR study of ferrocytochrome c showed that small anions, that is, ADP or ATP, bind to ferrocytochrome c in the immediate vicinity of His26(53).

Elucidation of the relation between the conformational changes of cytochrome c and the energy conversion mechanism in the TCA cycle will be a future problem.

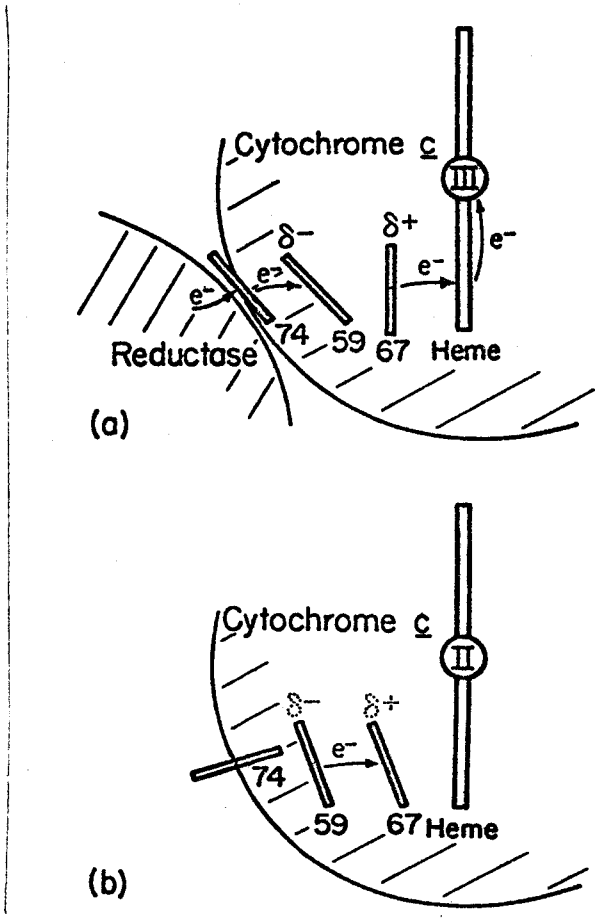



Fig.25 Reduction mechanism of cytochrome c proposed by Dickerson et al.(16).

Table 17 Calculated coefficients $C_i$'s of molecular orbitals with Hückel approximation

Trp59

| No. | E | C1 | C2 | C3 | C4 | C5 | C6 | c7 | C8 | C9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2.43 | 0.31 | 0.31 | 0.43 | 0.48 | 0.28 | 0.19 | 0.18 | 0.25 | 0.43 |
| 2 | 1.63 | -0.19 | -0.38 | -0.40 | -0.05 | 0.26 | 0.46 | 0.50 | 0.35 | 0.07 |
| 3 | 1.26 | -0.40 | -0.09 | 0.30 | 0.29 | 0.46 | 0.29 | -0.10 | -0.41 | -0.42 |
| 4 | 0.79 | 0.12 | 0.50 | 0.25 | -0.45 | -0.16 | 0.32 | 0.42 | 0.01 | -0.41 |
| 5 | 0.45 | -0.54 | -0.18 | 0.47 | 0.12 | -0.37 | -0.29 | 0.24 | 0.39 | -0.06 |
| 6 | -0.88 | -0.20 | 0.39 | -0.17 | -0.14 | 0.52 | -0.31 | -0.25 | 0.53 | -0.22 |
| 7 | -1.24 | -0.15 | -0.12 | 0.30 | -0.43 | -0.04 | 0.48 | -0.56 | 0.21 | 0.30 |
| 8 | -1.15 | -0.52 | 0.55 | -0.37 | 0.24 | -0.30 | 0.22 | -0.05 | -0.14 | 0.27 |
| 9 | -2.16 | -0.25 | 0.05 | 0.14 | -0.45 | 0.35 | -0.31 | 0.32 | -0.38 | 0.50 |

Tyr74

| No. | E | C1 | C2 | C3 | C4 | C5 | C6 | OH |
|---|---|---|---|---|---|---|---|---|
| 1 | 2,70 | 0,09 | 0.13 | 0.25 | 0.55 | 0.25 | 0.13 | 0.73 |
| 2 | 1.75 | -0,52 | -0.45 | -0.27 | -0.02 | -0.27 | -0.45 | 0.42 |
| 3 | 1.00 | 0,00 | -0.50 | -0.50 | 0.00 | 0.50 | 0.50 | 0.00 |
| 4 | 0.68 | -0,50 | -0.17 | 0.38 | 0.43 | 0.38 | -0.17 | -0.46 |
| 5 | -1.00 | 0,00 | -0.50 | 0.50 | 0.00 | -0.50 | 0.50 | 0.00 |
| 6 | -1.10 | -0,57 | 0.31 | 0.23 | -0.56 | 0.23 | 0.31 | 0.23 |
| 7 | -2.03 | -0,38 | 0.39 | -0.41 | 0.44 | -0.41 | 0.39 | -0.14 |

(a)

(b)

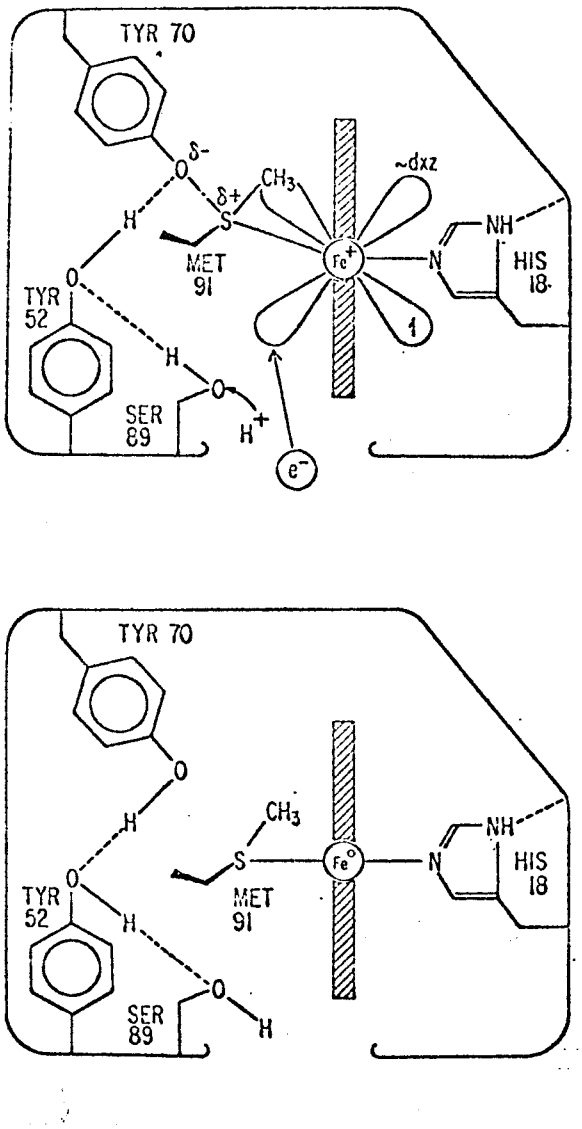



Fig.26 Oxidoreduction mechanism of R.rubrum cytochrome $c_2$ proposed by Kraut et al.(17,21).

(a) The region of ferriheme. Stabilization of the iron charge is achieved by a charge interaction between Met91 sulfur and Tyr70 oxygen. Note that destabilization of the oxidized state is expected upon protonation of Ser89.

(b) The region of ferroheme. The charge interaction between Met91 sulfur and Tyr70 oxygen has been abolished due to the heme iron charge.

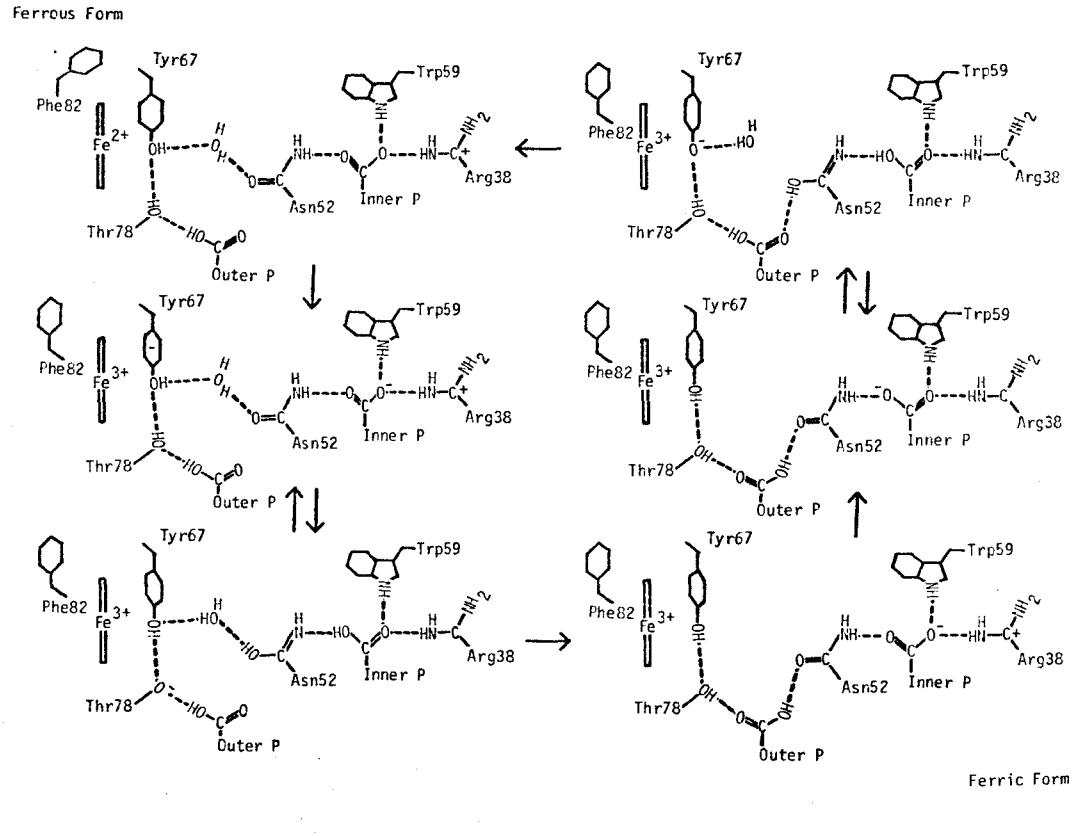



Fig.27 Proposed sheme of the electron transfer mechainsm. The ferrous form(left top) is converted to the ferric form(right bottom) with the rearrangement of the hydrogen bonding pattern and the movement of Phe82.

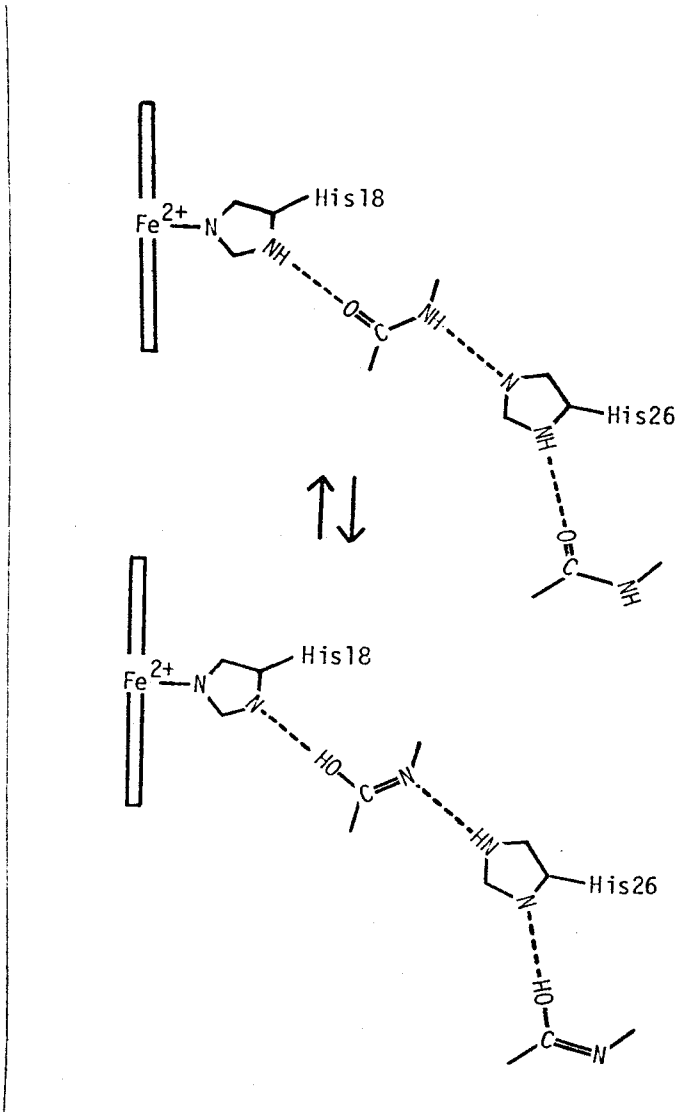



Fig.28 Hydrogen bonded chain from His18 to CO44.

REFERENCES


1) B.Chance, L.Mela, and D.Wong: Flavins and Flavoproteins(ed. K.Yagi), pp.107, University of Tokyo Press, Tokyo, 1968.

2) R.E.Dickerson: *Scientific American*, April, 58 (1972).

3) T.Nakayama, K.Titani, and K,Narita: *J.Biochem.(Tokyo)*, 70, 311 (1971).

4) W.D.Butt, and D.Keilin: *Proc. Roy. Soc.*, B156, 429 (1962).

5) S.Palens, and J.B.Nielands: *Acta Chem. Scand.*, 4, 1024 (1950).

6) K.Zeile, and F.Reuter: *Z. Physiol. Chem.*, 222, 101 (1933).

7) M.Nozaki, H.Mizushima, T.Horio, and K.Okunuki: *J. Biochem.(Tokyo)*, 45, 815 (1958).

8) G.Unger, E.Aschheim, S.Psychoyos, and D.V.Romano: *J. Gen. Physiol.*, 40, 635 (1957).

9) A.G.Redfield, and R.K.Gupta: *Symp. Quantitative Biology*, volXXXVI, 405 (1971).

10) P.Mcdermott, L.May, and J.Orlando: *Biophys. J.*, 615 (1967).

11) T.Flatmark, and A.B.Robinson: Structure and Function of Cytochromes(eds. K.Okunuki, M.D.Kamen, and I.Sekuzu), pp.318, University of Tokyo Press, Tokyo, 1968.

12) Y.P.Myer: *Biochem.*, 7, 765 (1968).

13) D.D.Ulmer: *Biochem.*, 4, 902 (1965).

14) K.Okunuki: *Adv. Enzymol.*, 23, 29 (1961).

15) R.E.Dickerson, T.Takano, D.Eisenberg, O.B.Kallai,

L.Samson, A.Cooper, and E.Margoliash: *J. Biol. Chem.*, 246, 1511 (1971).

16) T.Takano, O.B.Kallai, R.Swanson, and R.E.Dickerson: *J. Biol. Chem.*, 248, 5234 (1973).

17) F.R.Salemme, S.T.Freer, $N_G$.H.Xuong, R.A.Alden, and J.Kraut: *J. Biol. Chem.*, 248, 3910 (1973).

18) K.Dus, K.Sletten, and M.D.Kamen: *J. Biol. Chem.*, 243, 5507 (1968).

19) R.P.Ambler: *Biochem. J.*, 89, 341-349 and 349-380 (1963).

20) R.P.Ambler: *Biochem. J.*, 109, 47 (1968).

21) F.R.Salemme, J.Kraut, and M.D.Kamen: *J. Biol. Chem.*, 248, 7701 (1973).

22) A.Sugihara: Ph.D. thesis, Fuculty of Science, Osaka University, Toyonaka, Osaka, 1970.

23) T.Ashida, T.Ueki, T.Tsukihara, A.Sugihara, T.Takano, and M.Kakudo: *J. Biochem.(Tokyo)*, 70, 913 (1971).

24) T.Ashida, N.Tanaka, T.Yamane, T.Tsukihara, and M.Kakudo: *J. Biochem.(Tokyo)*, 73, 463 (1973).

25) N.Tanaka, T.Yamane, T.Tsukihara, T.Ashida, and M.Kakudo: *J. Biochem.(Tokyo)*, 77, 147 (1975).

26) E.Margoliash, and O.F.Walasek: Methods in Enzymology, vol.10 (eds. R.W.Estabrook, and M.E.Pullman), pp.339, Academic Press, 1967.

27) T.Tsukihara, T.Yamane, N.Tanaka, T.Ashida, and M.Kakudo: *J. Biochem.(Tokyo)*, 73, 1163 (1973).

28) B.W.Matthews: *J. Mol. Biol.*, 33, 491 (1968).

29) R.E.Dickerson, M.L.Kopka, C.L.Borders,Jr., J.Varnum,

J.E.Weizierl, and E.Margoliash: *J. Mol. Biol.*, 29, 77 (1967).

30) T.Takano, R.Swanson, O.B.Kallai, and R.E.Dickerson: *Cold Spring Harber Symp. Quant. Biol.*, 36, 397 (1971).

31) D.Harker: *Acta Cryst.*, 9, 1 (1956).

32) D.M.Blow, and F.H.C.Crick: *Acta Cryst.*, 12, 794 (1959).

33) A.F.Cullis, H.Muirhead, M.F.Pertz. M.G.Rossmann, and A.C.T.North: *Proc. Roy. Soc.*, A265, 15 (1961).

34) R.E.Dickerson, J.C.Kendrew, and B.E.Strandberg: *Acta Cryst.*, 14, 1188 (1961).

35) R.E.Dickerson, J.C.Kendrew, and B.E.Strandberg: Computing Methods and the Phase Problem in X-Ray Crystal Analysis(eds. R.Pepinsky, J.M.Robertson, and J.C.Speakman), pp.236, Pergamon Press, New York, 1961.

36) T.C.Furnas: Single Crystal Orienter Instruction Manual, General Electric Co., Milwaukee, 1956.

37) A.J.C.Wilson: *Nature*, 150, 152 (1942).

38) J.Kraut, L.C.Sieker, D.F.High, and S.T.Freer: *Proc. Natl. Acad. Sci.*, 48, 1417 (1962).

39) D.M.Blow: *Proc. Roy. Soc.*, A247, 302 (1958).

40) R.E.Dickerson, M.L.Kopka, J.C.Varnum, and J.E.Weinzierl: *Acta Cryst.*, 23, 511 (1967).

41) E.Dodson, and M.Vijaran: *Acta Cryst.*, B27, 2402 (1971).

42) R.G.Hart: *Acta Cryst.*, 14, 1194 (1961).

43) L.K.Steinrauf: *Acta Cryst.*, 16, 317 (1963).

44) F.M.Richards: *J. Mol. Biol.*, 37, 225 (1968).

45) R.B.Corey, and L.Pauling: *Proc. Roy. Soc.*, B141, 10 (1953).

46) J.T.Edsall, P.J.Flory, J.C.Kendrew, A.M.Liquori, G.Nemethy, and G.N.Ramachandran: *J. Mol. Biol.*, 15, 399 (1966).

47) IUPAC-IUB Commision on Biochemical Nomenclature: *J. Mol. Biol.*, 52, 1 (1970).

48) G.N.Ramachandran, C.Ramakrishnan, and V.Sasisekharan: *J. Mol. Biol.*, 7, 95 (1963).

49) C.M.Venkachalam: *Biopolymers*, 6, 1425 (1968).

50) J.Brown, T.Takano, and R.E.Dickerson: Private communication.

51) K.Wada, and K.Okunuki: *J. Biochem.(Tokyo)*, 64, 667 (1968).

52) .Ohi: Private communication.

53) E.Stellwagen, and R.G.Shulman: *J. Mol. Biol.*, 75, 683 (1973).

54) D.D.Ulmer: *Biochem.*, 5, 1886 (1966).

55) Y.P.Myer: *Biochem.*, 11, 4195 (1973).

56) S.Yomosa: *J. Phys. Soc. Japan*, 35, 1501 (1973).

57) T.A.Hoffman, and J.Ladik: *Adv. Chem. Phys.*, 7, 84 (1964).

58) S.Yomosa: *J. Phys. Soc. Japan*, 35, 1738 (1973).

59) R.Henderson: *J. Mol. Biol.*, 54, 341 (1970).