Structural Studies
on
Ferricytochrome *c'* from *Rhodospirillum rubrum*

(*Rhodospirillum rubrum* チトクロム *c'* の構造学的研究)

Masanori YASUI

Osaka University

1991

Preface

The work described in this thesis was achieved by the author under the guidance of Professor Nobutami Kasai at Department of Applied Chemistry, Faculty of Engineering, Osaka University.

The author wishes to express his sincerest gratitude to Professor Nobutami Kasai for his helpful suggestions, his constant guidance and hearty encouragement throughout this work. The author also wishes to make a grateful acknowledgment to Dr. Shigeharu Harada for his constant advice and stimulating discussions. The author acknowledges the continuing encouragement of Dr. Kunio Miki and Dr. Yasushi Kai. Furthermore he wishes to thank Mr. Katuji Murakawa for his collaboration.

The author acknowledges to Professor Takekazu Horio, Institute for Protein Research, Osaka University, for his kind advice in the purification of cytochrome *c'*. He also acknowledges to Professor Yoshinori Satow, Faculty of Pharmaceutical Sciences, the University of Tokyo, for his cordial support in X-ray experiment at Photon Factory, National Laboratory of High Energy Physics.

His gratitude is expended to all the members of the research group of Professor Nobutami Kasai for their occasional discussions and warm cooperations with him.

Finally the author would like to express his deep thanks to his parents for their perpetual support and encouragement.

September 1991

安井正憲

Masanori Yasui

List of publications

1) A New Crystal Form and Preliminary Crystallographic Data for Ferricytochrome *c'* from *Rhodospirillum rubrum*
   Masanori YASUI, Shigeharu HARADA, Yasushi KAI, and Nobutami KASAI
   J. Mol. Biol. 177, 845-846 (1984)

2) Structure of Ferricytochrome *c'* from *Rhodospirillum rubrum* at 6Å Resolution
   Masanori YASUI, Shigeharu HARADA, Yasushi KAI, and Nobutami KASAI
   J. Biochem. 98, 77-80 (1985)

3) Crystal Structure Analysis of Cytochrome *c'* by the Multiwavelength Anomalous Diffraction Method Using Synchrotron Radiation
   Shigeharu HARADA, Masanori YASUI, Katuji MURAKAWA, Nobutami KASAI and Yoshinori SATOW
   J. Appl. Cryst. 19, 448-452 (1986)

4) Three-dimensional Structure of Ferricytochrome *c'* from *Rhodospirillum rubrum* at 2.8Å Resolution
   Masanori YASUI, Shigeharu HARADA, Yasushi KAI, Nobutami KASAI, Masami KUSUNOKI and Yoshiki MATSUURA
   J. Biochem 111, (1991) (in press)

Other papers

1) The Crystal and Molecular Structure of 1:2 Molecular Complex of 1,4-Diazabicyclo[3.3.3]octane (DABCO) with *p*-Cresol
Masaaki TAKAMA, Masanori YASUI, Shigeharu HARADA, Nobutami KASAI, Koichi TANAKA and Fumio TODA
Bull. Chem. Soc. Jpn., 61, 567-568 (1988)

2) The Crystal and Molecular Structures of 1:2 Complexes of *N,N'*-Dimethylpiperazine with 1-(*o*-Chlorophenyl)-1-phenyl-2-propyn-1-ol
Teturo YABUKI, Masanori YASUI, Hiroshi TEZUKA, Shigeharu HARADA, Nobutami KASAI, Koichi TANAKA and Fomio TODA
Chem. Lett., 749-752 (1988)

3) The Crystal and Molecular Structures of Two 1:2 Complexes of 1,4-Diazabicyclo[2.2.2]octane(DABCO) with 1-(*o*-Chlorophenyl)-1-phenyl-2-propyn-1-ol(CPPOH), and Those of Two 1:2 Complexes of *N,N'*-Dimethylpiperazine(DMP) with CPPOH
Masanori YASUI, Tetsuro YABUKI, Masaaki TAKAMA, Shigeharu HARADA, Nobutami KASAI, Koichi TANAKA and Fumio TODA
Bull. Chem. Soc. Jpn., 62, 1436-1445 (1989)

4) First X-Ray Structure Determination of a Bismuthio Ylide: 4,4-Dimethyl-2,6-dioxo-1-triphenylbismuthiocyclohexanide
M. YASUI, T. KIKUCHI, F. IWASAKI, H. SUZUKI, T. MURAFUJI and T. OGAWA
J. Chem. Soc. Perkin Trans. 1, 3367-3368 (1990)

5) Formation and X-Ray Crystal Structure of Charge-Transfer complexes of Silacyclopentadiene Derivatives with Quinones
S. KYUSHIN, Y. OHKURA, Y. NAKADAIRA, M. OHASHI, M.YASUI, M. MATSUI and F. IWASAKI
Chem. Lett. 883-886 (1991)

6) Structures of 6a-Thia/Selena-1,3,4,6-tetraazapentalene Derivatives: Hypervalent S/Se– N Bonds
F. IWASAKI, H. MURAKAMI, N. YAMAZAKI, M. YASUI, M. TOMURA and N. MATSUMURA
Acta Cryst. C47, 998-1003 (1991)

Contents

1 Introduction

1-1 General introduction

The cytochromes *c'* are electron transport proteins that are widely distributed in denitrifying and photosynthetic bacteria (Bartsch, 1978). Their midpoint oxidation-reduction potentials range between 0 and 150 mV. In general, cytochromes *c'* exist in dimeric molecules that consist of chemically identical polypeptide chains of 14,000 molecular weight, and the protoheme IX prosthetic group is covalently bound to cysteine residues of each polypeptide chain through thioether linkages. The amino acid sequences of cytochromes *c'* from 12 species of photosynthetic bacteria have been determined, indicating that there are some conserved amino acid residues despite the large amount of divergence of the sequences and that cytochromes *c'* exhibit no sequence homology with the well-characterized cytochromes *c* of mitochondria (Ambler *et al.*, 1981). The magnetic (Maltempo *et al.*, 1974), spectroscopic (Ehrenberg and Kamen, 1985) and ligand-binding (Gibson and Kamen, 1966) properties of cytochromes *c'* are also different from those of cytochromes *c*, thus the heme iron of cytochromes *c'* is in a high-spin environment. Crystals of cytochromes *c'* from *Rhodospirillum rubrum* (Langridge, 1961), *Rhodopseudomonas palustris* (Salemme, 1974), *Rhodopseudomonas capsulata* and *Rhodospirillum molischianum* (Weber and Salemme, 1977) have been obtained so far, and the high resolution crystal

structure of ferricytochrome *c'* from *Rhodospirillum molischianum* has been determined (Weber *et al.*, 1981; Finzel *et al.*, 1985). In this study I report the results of the X-ray structure analysis of ferricytochrome *c'* from *Rhodospirillum rubrum.*

1-2 Introduction to the multiwavelength X-ray diffraction study

Since the first protein structure, sperm whale myoglobin, was determined (Kendrew *et al.*, 1960), the isomorphous replacement method has been the most powerful tool for solving crystal structures of proteins, and the most protein structures so far have been determined by this method. On the other hand, it has long been recognized that protein phase angles can be obtained from X-ray intensity data collected at several wavelengths (multiwavelength data) near an absorption edge of an anomalous scatterer such as iron and zinc etc. which exists intrinsically in some proteins (Pepinsky and Okaya, 1956; Mitchell, 1957; Herzenberg and Lau, 1967; Phillips and Hodgson, 1980). This multiwavelength anomalous diffraction (MAD) method utilizes changes of scattering factor of an anomalous scatterer with wavelength. However, the MAD method has not been applied to protein crystallography, chiefly for lack of a tunable X-ray source. The earliest work of the structure analysis by the MAD method was performed by Hoppe and Jakubowski (Hoppe and Jakubowski, 1975). In their work, X-ray intensity data were

measured from an iron-containing protein, erythrocruorin, with Ni and Co $K\alpha$ radiations. Recently, synchrotron radiation as an intense X-ray source with broad spectral distribution has become available for protein crystallography, and experiments utilizing anomalous scattering effects have been initiated. The phase angles of rubredoxin, which is an iron-containing protein (one iron atom per 430 non-hydrogen atoms), were obtained by the use of film data recorded at several wavelengths near the iron *K* absorption edge (Phillips *et al.*, 1977). The electron density map of the calcium-binding protein, parvalbumin, in which terbium was substituted for calcium (1.3 terbium atoms per 720 non-hydrogen atoms) in order to increase the anomalous scattering effect, was calculated by the MAD method using X-ray intensity data collected at three wavelengths near the terbium L_3 absorption edge (Kahn *et al.*, 1985).

In this work, the electron density map of cytochrome *c'* was calculated by utilizing the anomalous scattering effect of iron. This study is the first attempt to utilize the diffraction data containing the anomalous scattering effect, which is caused by intrinsically existing atoms in the protein, for three dimensional structure analysis. Although the molecular weight of cytochrome *c'* (one iron atom per 1000 non-hydrogen atoms) is larger than the two proteins above and the anomalous scattering effect of iron is smaller than that of terbium, an interpretable electron

density map could be obtained by choosing wavelengths so as to give the maximum intensity difference owing to the real correction component of the anomalous scattering effect.

2 Materials and Preliminary Studies

2-1 Purification of Cytochrome *c'* from *Rhodospirillum rubrum*

Extraction and purification of cytochrome *c'* from *Rhodospirillum rubrum* were performed by the method of Horio (Horio and Kamen, 1961) with some modification. Total procedure is shown in the scheme below. The cells of carotenoid-less blue-green mutant strain (G-9) of *Rhodospirillum rubrum* suspended in buffer solution were ruptured by sonication, and then the water solvable fraction was separated with ultra-centrifugation. The supernatant was treated with ammonium sulfate fractionation at the concentration of 75% saturation. The centrifuged supernatant was then injected to the hydrogen bond chromatography on DEAE Sepharose CL-6B column. The fractions of cytochrome *c'* and cytochrome *c*2 were eluted as shown in Fig. 2-1. Typical yield of cytochrome *c'* was ca. 70mg from 325g (wet weight) of the cells. The purity and yield were checked by UV and visible absorption spectrum (Fig. 2-2).

Cells *R. rubrum* (G-9) (wet weight 325 g)

↓ sonication 20kHz 50 % duty cycle for 10 min.

↓ ultra-centrifugation 100000 g for 90 min.

supernatant ca. 1 *l* (in 0.1M tris-HCl pH 8)

↓ 75 % sat. $(NH_4)_2SO_4$ fractionation

↓ centrifugation

supernatant ca. 1.5 *l*

↓ hydrogen bond chromatography on DEAE-Sepharose CL-6B

bed volume 310 m*l*

linear gradient of conc. from 75 to 30 % sat. $(NH_4)_2SO_4$

↓ desalting and concentration

purified cytochrome *c'* ca. 70 mg

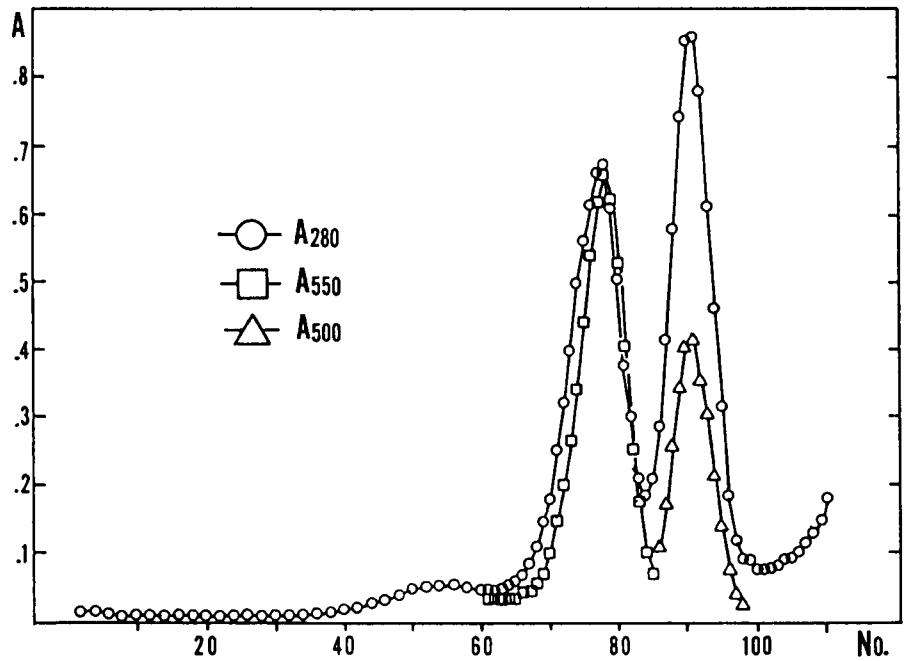


Fig. 2-1 Elution pattern from hydrogen bond chromatography on DEAE-Sepharose CL-6B column. Optical absorbance at 280 (A_{280}), 550 (A_{550}) and 500 nm (A_{500}) were plotted on each fraction numbers. A_{280} indicates total quantity of the protein, A_{500} indicates the quantity of cytochrome *c'* and A_{550} corresponds to cytochrome *c*2, respectively. In this experiment, the fraction numbers of 87 to 95 were regarded as cytochrome *c'* fractions.

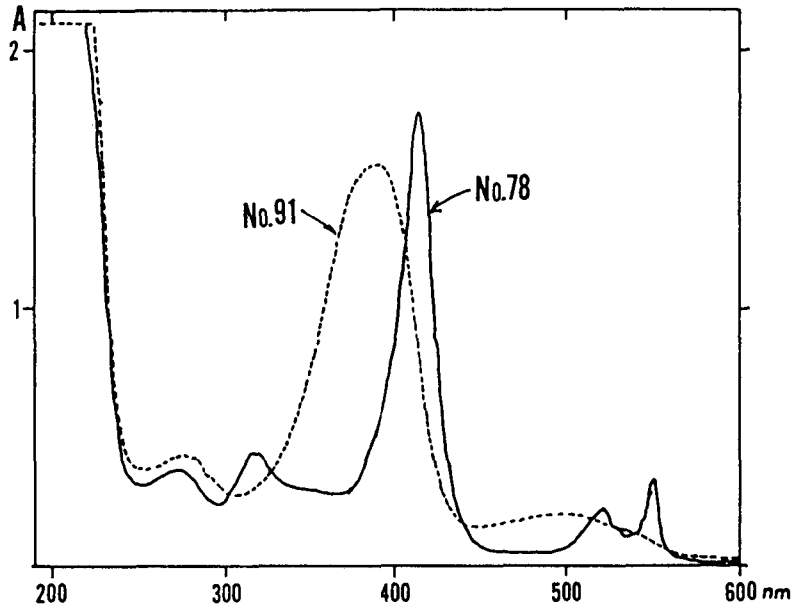


Fig. 2-2 UV and visible absorption spectrum of cytochrome *c'* (dotted line, fraction number of 91) and *c*2 (full line, fraction number of 78), obtained from the chromatography after 1/2 dilution (Fig. 2-1). These spectrum indicate that cytochrome *c'* is in an oxidated form, while cytochrome *c*2 in a reduced form.

2-2 Crystallization and Preliminary Crystallographic studies

The purified protein in 50mM-Tris-HCl buffer (pH8.0) was precipitated by adding 50%(w/v) polyethylene glycol 4000 (PEG4000). The resulting precipitate collected by centrifugation was successively dissolved in a minimum amount of 50mM-Tris-HCl buffer(pH8.0) containing 22%(w/v) PEG4000. The crystals are grown by dialyzing this protein solution against 50mM-phosphate buffer (pH7.0) containing 25%(w/v) PEG4000 in a Zeppezauer tube (Zeppezauer *et al.*, 1968). Figure 2-3 shows the crystallization apparatus. Long hexagonal rod-shaped crystals, as shown in Fig.2-4, grew within a few weeks. The crystals were sealed in thin-walled glass capillary tubes and X-ray diffraction patterns were recorded with an Enraf-Nonius precession camera. Ni-filtered Cu$K\alpha$ radiation from a sealed tube was used and the crystal-to-film distance was 10cm. Figure 2-5 shows two precession photographs taken along two different directions. These photographs indicate the 6/m Laue symmetry and the systematic absence of l=6n in 00l. Thus it was revealed that the crystals belong to hexagonal space group $P6_1$ (or $P6_5$) with unit cell dimensions : a=b=51.63, c=155.39Å and V=3.56$\times 10^5$Å^3. Assuming that one dimeric molecule (M_w=28000) exists in an asymmetric unit, the V_m value of 2.12Å^3/dalton was calculated. This value is comparable to that normally found in protein crystals (Matthews, 1968). One dimeric molecule per asymmetric

unit in conjunction with a crystal density of 1.28g/cm^3, which was measured by the floating method using toluene and bromobenzene, allows to conclude that approximately 38% of the unit cell volume is filled with solvent. These preliminary crystallographic data are summarized in Table 2-1. The crystal form obtained in this study is different from that obtained by Langridge (Langridge, 1961), which belonged to the orthorhombic system. The crystals are stable to X-ray irradiation and diffract to at least 2.5Å resolution, so this new crystal form appears to be suitable for high-resolution X-ray structure analysis.

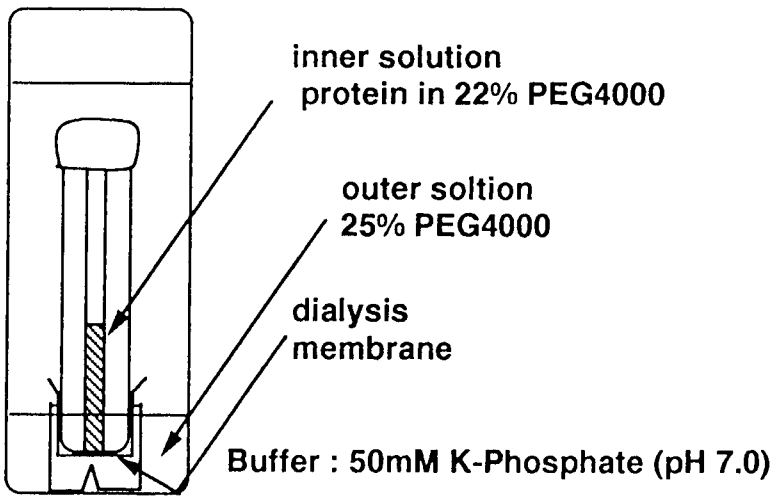


Fig. 2-3 Micro dialysis method using a Zeppezauer tube to crystallize cytochrome *c'*. Crystals grew in the inner solution within a few weeks.

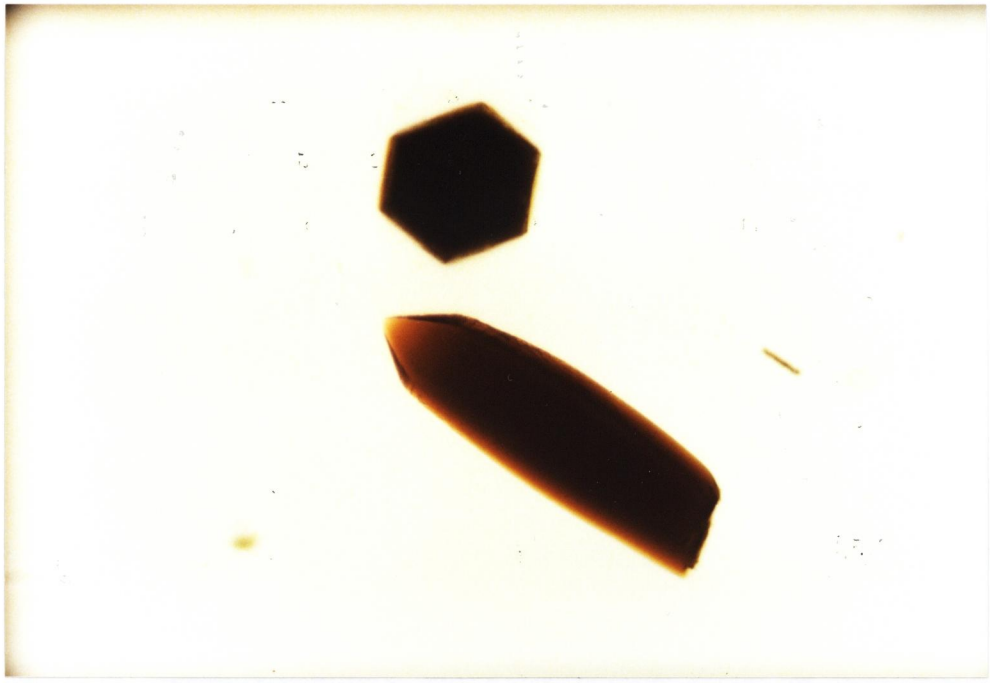


Fig. 2-4 A microphotograph of crystals of *R. rubrum* cytochrome *c'*. Crystals were rod-shaped with a hexagonal section. Typical dimensions of the crystals were 0.1 to 0.4mm in diameter and 0.5 to 2.0mm in length.

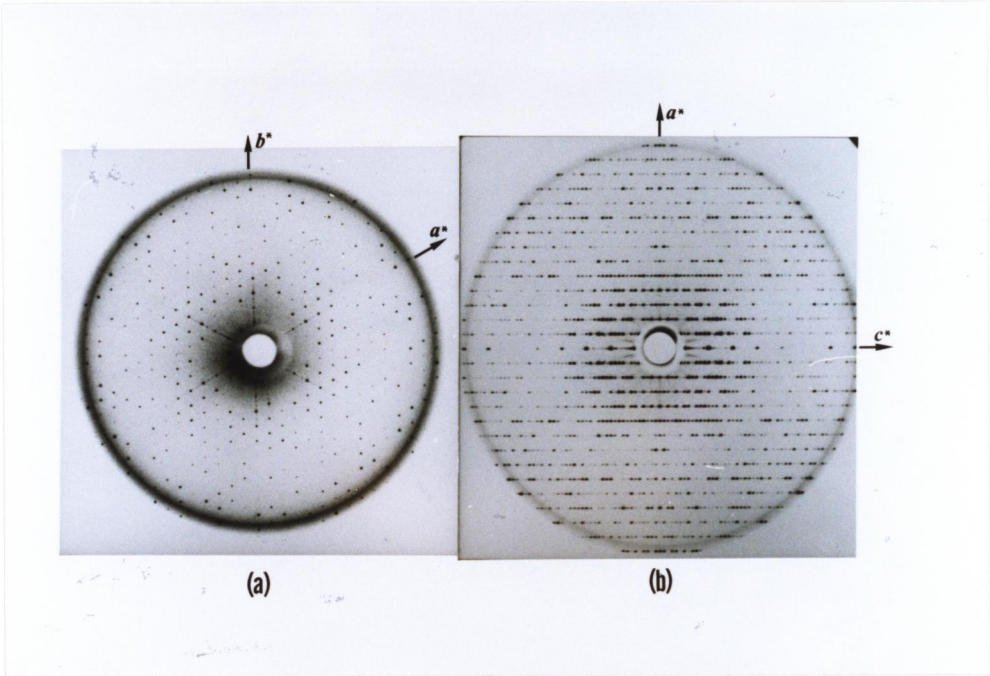


Fig. 2-5 The precession photographs. In (a), the *hk*0 section is appeared (*a** and *b** axes are indicated with arrows), and the *h*0*l* section in (b) (*a** and *c** axes are also indicated).

Table 2-1 Preliminary crystallographic data

Crystal system	hexagonal
Space group	$P6_1$ (or its enantiomorph $P6_5$)
Cell constants	$a = b = 51.63$Å
	$c = 155.39$Å
	$\gamma = 120°$
	$V = 3.56 \times 10^5$Å^3
V_m	2.12Å^3/dalton
D_m	1.28g/cm^3
Solvent content	38 %

3 Structure Analysis at Low Resolution I

3-1 Isomorphous Replacement Method

The structure analysis of protein crystals are mainly achieved by the isomorphous replacement method using the diffraction data of the native and the heavy atom derivative crystals. A perfectly isomorphous derivative is one in which the only change in electron density between it and the native crystal is a peak at the site of heavy atom substitution. In the case of protein crystals, the way of making an isomorphous derivative is the soaking method. The protein crystals contain large channels of mother liquor. It is often possible to bind a heavy atom to the surface of the protein so that it does not disturb the molecular or crystal structure but rather occupies a position in one of these channels. The channels also provide routes along which heavy atoms can diffuse so that isomorphous derivatives can be made from the native crystals. Thus these derivatives are really prepared by isomorphous addition rather than replacement, for it is usually assumed that the heavy atom replaces disordered solvent molecules which give little contribution to the diffraction pattern except at low resolution. If the heavy atom is sufficiently electron dense it will provide observable changes in the X-ray diffraction pattern. The intensities of the native and derivative diffraction patterns can be used to solve the phase

problem if the position of the heavy atom added is known. For centrosymmetric crystals, one heavy atom derivative will unambiguously determine the phase. For non-centrosymmetric protein crystals there is an ambiguity in the phase determination of most reflections but this can be overcome by using a further derivative. The use of several isomorphous derivatives is called multiple isomorphous replacement (MIR).

Let F_P be a structure factor for the native protein. F_P is a vector and can be described in terms of a structure amplitude, $|F_P|$, and a phase, α_P. Thus $F_P=|F_P|\exp i\alpha_P$. The structure factor, F_{ph}, of the heavy atom derivative is also a vector with amplitude, $|F_{ph}|$, and phase α_{ph}. F_{ph} can be derived from F_P by the vector addition of F_h, which is the contribution of the heavy atoms to the structure factor of the derivative. This can be written: $F_{ph}=F_P+F_h$.

If $|F_{ph}|$ and $|F_P|$ are measured and the arrangement of heavy atoms in the crystal unit cell are known, i.e. the vector F_h is calculable, the phase angle of $|F_P|$ can be derived from Fig. 3-1 using cosine law:

$$\alpha_P=\alpha_h+\cos^{-1}[(|F_{ph}|^2-|F_P|^2-|F_h|^2)/(2|F_P||F_h|)] = \alpha_h\pm\alpha'$$

This equation shows that there are two possible values for α_P which cannot be distinguished with one isomorphous derivative. Only when the vectors F_P and F_h are collinear, there is no ambiguity. For solving this ambiguity, two or more heavy atom

derivative is needed. Double isomorphous replacement will give the situation shown in Fig. 3-2. The vectors $\overrightarrow{OJ}$ and $\overrightarrow{OK}$ are the contributions of the first heavy atom, $-F_{h1}$, and the second one, $-F_{h2}$, respectively. Two circles of radius $|F_{ph1}|$ and $|F_{ph2}|$ are drawn centered on the end of vector at J and K. A third circle of radius $|F_p|$ is drawn with its center at O. There is a one point of intersection of three circles, i.e at H. Thus vector $\overrightarrow{OH}$ is the correct choice for F_p. Such geometric or analytical solutions of the phases assume that the structure factors are perfectly measured and that there is perfect isomorphism. Since experimental and model errors are present in fact, the circles in the geometric representation will probably not intersect exactly and in some cases not intersect at all. The lack-of-closure error, $x=|F_{ph}|-|(|F_p|\exp i\alpha + F_h)|$, and estimated errors in experiment and model are used for calculating the phase probabilities (Blow and Crick, 1959). The phase angle with the highest probability is called as the most probable phase, and the phase angle at the centroid of the probability distribution is the best phase, α_{best}. The electron density calculated with the best phase angles and the weighting function, figure-of-merit, m, will give the minimum mean square error over the whole unit cell. This Fourier calculation is achieved using the coefficient :

$m|F_p|\exp i\alpha_{best}$

The figure-of-merit gives the mean value of the cosine of the

error in phase angle for the reflection. In theory, a value of m=1 corresponds to a zero error in phase angle while a value of m=0.74 corresponds approximately to a 42° error. The single isomorphous replacement combined with anomalous dispersion data can be used to determine the phase angle unambiguously. The principle of this method is just same as double isomorphous replacement. F_{ph1} and F_{ph2} in the above treatment are replaced by $F_{ph}(+)$ and $F_{ph}(-)$, respectively. F_{h1} and F_{h2} are replaced by $F_h(+)$ and $F_h(-)$. $F_{ph}(+)$ and $F_h(+)$ mean structure factors of heavy atom derivative and heavy atom contribution for *hkl* diffraction, and $F_{ph}(-)$ and $F_h(-)$ are their Bijvoet pairs, respectively.

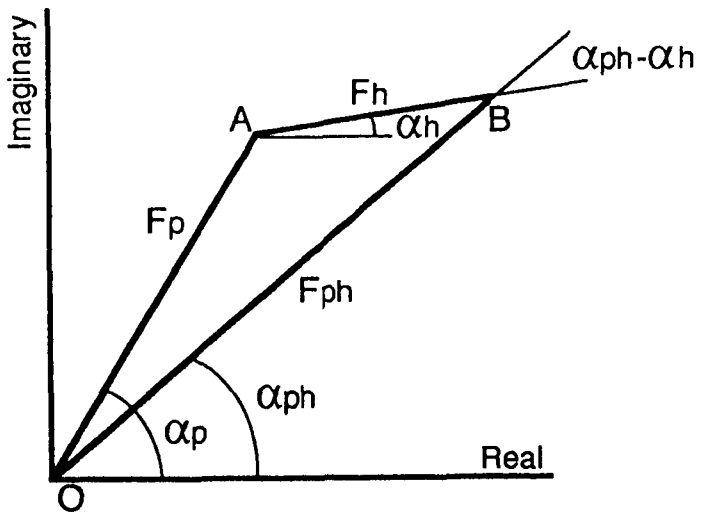


Fig. 3-1 A vector diagram defining the structure factor amplitudes and phases.

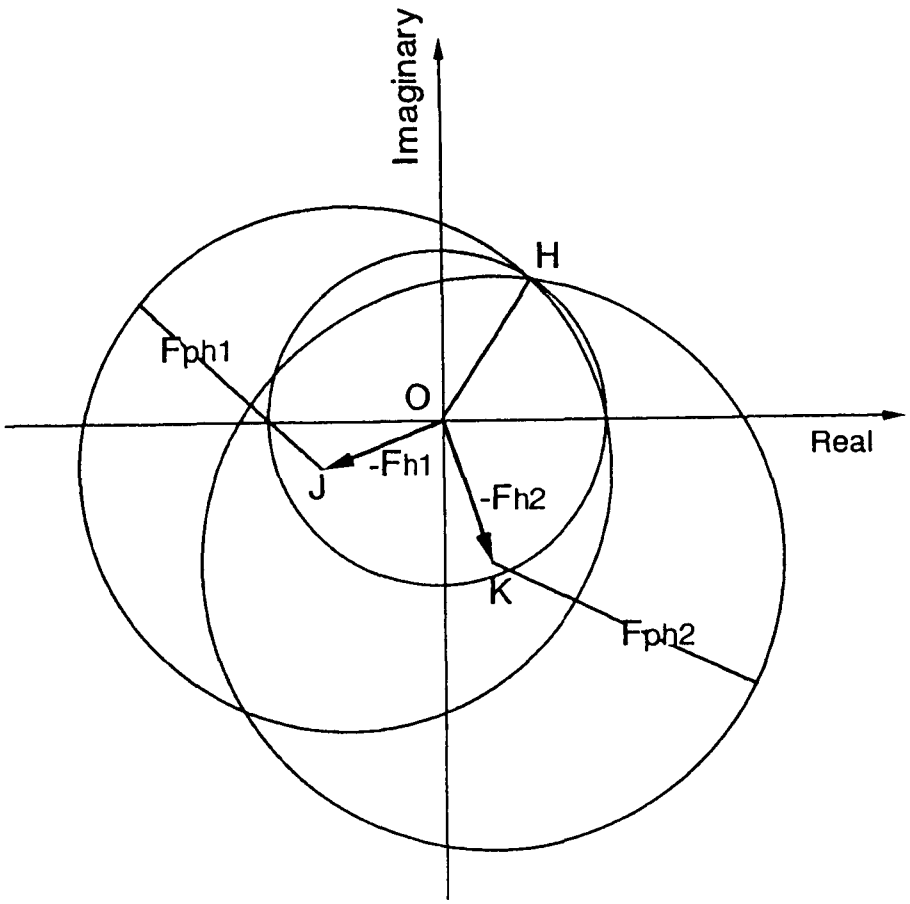


Fig. 3-2 The Harker construction for phase determination by the method of double isomorphous replacement. The vectors $\overrightarrow{OJ}$ and $\overrightarrow{OK}$ represents $-F_{h1}$ and $-F_{h2}$ for the first and second heavy atom contributions to the structure factor of heavy atom derivatives. A circle of radius $|F_{ph1}|$ is drawn with its center at J, and a circle of radius $|F_{ph2}|$ with its center at K. F_p is given unequivocally by $\overrightarrow{OH}$.

3-2 Preparation of Heavy Atom Derivatives

In order to prepare the heavy atom derivatives, the heavy atom reagents and the conditions for soaking were surveyed. The general conditions of the first survey were: the concentrations of the heavy atom reagent were 1 to 20mM in the mother liquor (30% PEG4000 in 50mM potassium phosphate buffer), and the time of soak were 2 to 5 days. The soaked crystals were then checked for its isomorphism with the native crystal, the quality of diffraction and the change of the diffraction patterns by taking the precession photographs and measuring the diffraction data at the low resolution.

Table 3-1 shows the results of the first survey, indicating that the platinum reagent (K_2PtCl_6) gave a good derivative crystal although the concentration of the reagent was lower, yet saturated, than those of others. The mercuric derivative using CH_3HgCl was slightly good and gave the different change of diffraction pattern from that of platinum derivative. At this stage, the platinum and the mercuric derivatives were selected for the structure analysis.

Table 3-1 Results of survey for heavy atom derivatives

Heavy atom reagent	Concentration (mM)	Time of soak	Result
K_2PtCl_6	saturated (<3)	5 days	good
CH_3HgCl	20	2 days	fair
$K_3UO_2F_5$	0.8 [a]	1 hour	bad(cracked)
$SmCl_3$	1 [a]	2 days	bad [b]
	1 [a]	1 day	bad [b]
K_2IrCl_6	1 [a]	2 days	bad [b]
	1 [a]	4 days	bad [b]

[a] mother liquor : 30% PEG4000 in 50mM HEPES buffer pH 7.
[b] the isomorphous difference Patterson and Fourier maps are unsolved.

3-3 Data Collection

X-ray diffraction data from the native and two derivative crystals were collected on a Rigaku four-circle diffractometer (AFC5R) at the Crystallographic Research Center, Institute for Protein Research, Osaka University. Ni-filtered Cu$K\alpha$ radiation was provided by a Rigaku rotating anode X-ray generator operated at 40kV and 300mA. Each data set included Bijvoet pairs (hkl and $hk\bar{l}$). The intensities were measured for the ω-scan range of 1 ° at a speed of 4 °/min. The raw data were corrected for Lorentz and polarization factors, radiation damage (in the form of an exponential correction), and absorption effects using the method of North *et al.* (North *et al.*, 1968).

3-4 Phase Determination

At the beginning of a phase determination, the positional parameters of the heavy atom sites of the K_2PtCl_6 derivative were determined from the isomorphous difference Patterson map calculated using $(|F_{ph}|-|F_p|)^2$ as coefficients. $|F_{ph}|$ is the structure amplitude from the platinum derivative and $|F_p|$ from the native crystals, respectively. One of the Harker sections is shown in Fig. 3-3, giving the peaks corresponding to two heavy atom sites. The relative position of these sites was solved from other sections of this map. The mercuric sites of the CH_3HgCl derivative could not be located from the isomorphous difference Patterson map, but located from the isomorphous difference Fourier synthesis using the best phase angles of the native protein, which were calculated by the single isomorphous replacement method combined with the anomalous dispersion data of the K_2PtCl_6 derivative. The correct enantiomorph of the space group was also determined by this difference Fourier synthesis. In the difference Fourier map with phases corresponding to the space group $P6_1$, mercuric sites are clearly visible, although no significant peak is visible in the map with phases corresponding to the space group $P6_5$.

Parameters of the heavy atoms were refined by minimizing

$$[|F_{ph}|-|(|F_p|\exp i\alpha_{best} + F_h)|]^2$$

in each least square cycle. The final results of the phase

calculation are summarized in Tables 3-2 and 3-3.

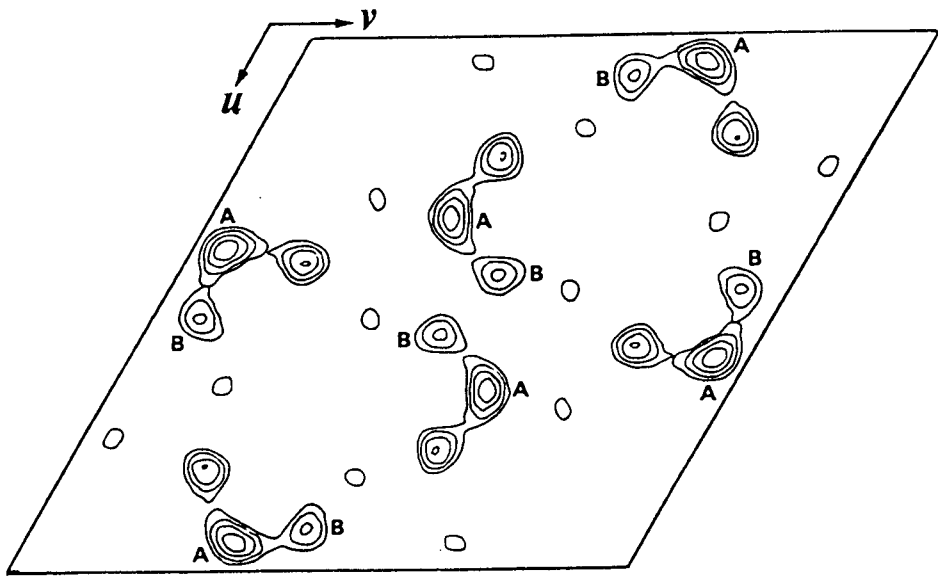


Fig. 3-3 One of the Harker sections (w=1/2) of the isomorphous difference Patterson map of the platinum derivative. Peaks indicated as B and A are correspond to the self vectors of Pt1-Pt1 and Pt2-Pt2, respectively.

3-5 Electron Density at 6Å Resolution

An electron density map was calculated using the best values of phase angles (Blow and Crick, 1959) in 50 sections perpendicular to the *z*-axis. Each *xy*-plane was divided into 40 grid points along the *x*- and the *y*-axes. Since the boundary of the molecule was not known, the map was calculated from -0.75 to 0.75 along the *x*- and the *y*-directions and from 0.0 to 0.5 along the *z*-direction. The value of electron density ranged from 64 to -55, and the map was contoured in intervals of 10 units starting from a base value of 10 unit. Visual examination of the map indicated that the dimeric molecule was located in the region of the unit cell 10/40 to 55/40 in the *x*-axis; -12/40 to 34/40 in the *y*-axis; 5/50 to 14/50 in the *z*-axis. The map was then calculated from 0.0 to 1.5 along the *x*-axis, from -0.45 to 1.05 along the *y*-axis and from 0.0 to 0.5 along the *z*-axis, and contoured at the levels of 12, 20, 30,...60. A composite fo 10 sections (5/50-14/50) is shown in Fig 3-4.

Table 3-2 Heavy atom parameters

Derivative	site	*x*	*y*	*z*	Occupancy	B(Å^2)
		(fractional units)			(electrons)	
K_2PtCl_6	1	0.253	0.535	-0.378	41.6	39.6
	2	0.162	0.186	0.0 [a]	48.8	45.5
CH_3HgCl	1	0.900	0.600	0.094	12.5	28.8
	2	0.930	0.065	0.121	13.5	21.5

[a] fixed at 0 to define the origin of *z* axis

Table 3-3 Refinement statistics

Resolution (Å)	20.0-9.28	9.28-7.49	7.49-6.58	6.58-6.00	overall
No. of reflections	150	137	148	145	580
Average change of $\|F\|$ [a]			Pt-derivative		0.31
			Hg-derivative		0.24
Average figure of merit $\langle\|F_h\|\rangle$ [b]	0.92	0.91	0.90	0.88	0.90
Pt-derivative	146.5	140.1	123.3	122.4	
Hg-derivative	42.8	40.5	39.2	38.9	
$\langle E\rangle$ [c]					
Pt-derivative	12.5	21.2	19.3	27.8	
Hg-derivative	67.0	60.7	52.9	42.6	

Kraut *R*-factor [d]	Pt-derivative	4.6%
	Hg-derivative	6.1%

[a] $\Sigma\,||F_{ph}|-|F_p||/\Sigma\,|F_p|$ [b] RMS value of calculated heavy atom contribution. [c] RMS value of lack-of-closure error. [d] Kraut *R*-factor: $\Sigma\,(|F_{ph}|-||F_p|\exp i\alpha_{best} + |F_h|)/\Sigma\,|F_{ph}|$.

3-6 Discussion

At 6 Å resolution, the number of reflections is 580(1120 including Bijvoet pairs) and the overall figure of merit for final best phase angles is 0.90. The lack-of-closure errors compared with calculated heavy atom contributions, the inversion of so called as phasing power, are 15.6% for the K_2PtCl_6 derivative and 140% for the CH_3HgCl derivative. Since the occupancies of the mercuric sites were not sufficiently large, the lack-of-closure error did not decrease. Consequently, the isomorphous difference Patterson map would not be solved solely. This situation might be improved by soaking the crystals for a longer period or using more concentrated heavy atom solution. The electron density map (Fig. 3-4) shows the molecular boundary clearly in the low electron density region, and the molecule has the dimensions of approximately $50\times 25\times 30$Å as a monomer molecule, of $60\times 45\times 30$ as a dimer. The rod-shaped electron density in the map represents the α – helices, and it can be seen that there are four α – helices per monomer molecule. The right half is named monomer I, and the left half is monomer II. Two green filled circles indicate the heme iron atoms. The distance between them is 24.5Å, and their coordinates are Fe1(0.94, 0.59, 0.12), right hand side, and Fe2(0.82, 0.07, 0.14), left hand side. The binding sites of each heavy atom are indicated as filled squares, the green squares are Pt1(right) and Pt2(left), and the orange

squares are Hg1(right) and Hg2(left), respectively. In the case of the cytochrome *c'* from *Rhodospirillum molischianum* (Weber *et al.*, 1981, Finzel *et al.*, 1985), the monomer molecule consists of four α – helices that run approximately antiparallel to each other, and the monomer molecules are related by a non-crystallographic dyad axis in the dimer. The electron density map obtained in this study suggests that the two cytochrome *c'* molecules have similar chain-folding patterns.

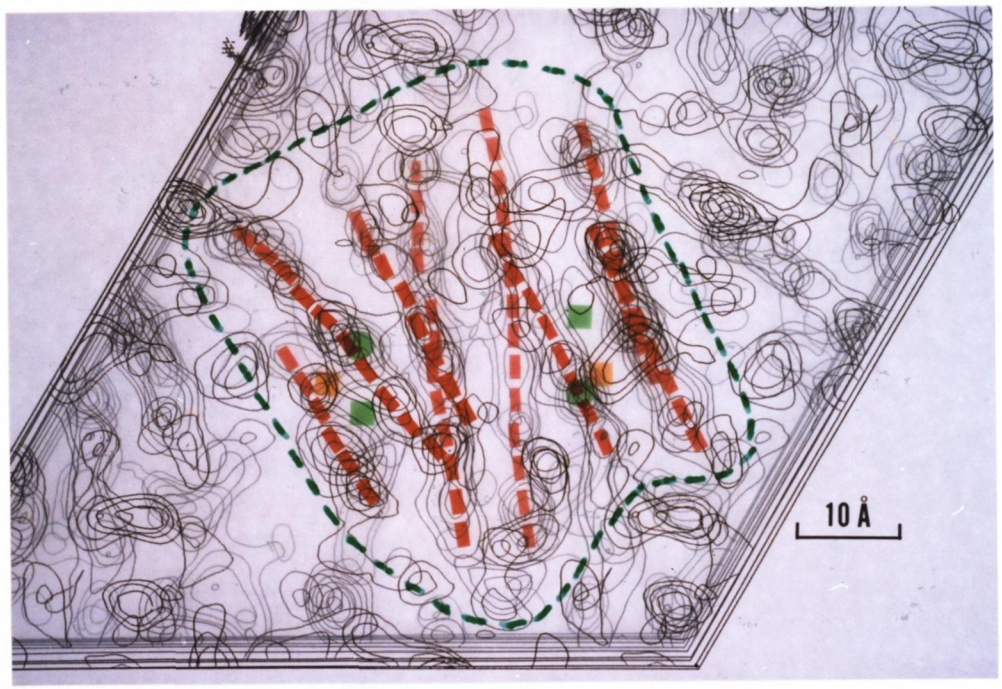


Fig 3-4 A composite of 10 sections of the electron density map of cytochrome *c'*. Eight rods of electron density indicated by dotted lines can be seen in a dimeric molecule which is surrounded by a green colored dotted line.

4 Structure Analysis at Low Resolution II -- Multiwavelength Anomalous Diffraction Method Using Synchrotron Radiation --

4-1 Data Collection and Processing

Crystals used for multiwavelength anomalous diffraction experiments were made by the method described in chapter 2-2. X-ray diffraction experiments were performed using the X-ray beam produced by the storage ring in the Photon Factory, National Laboratory for High Energy Physics. A four-circle diffractometer located on the beam-line station 14A(BL14A) was used for data collection (Satow, 1984). The diffractometer has a horizontal-type setup, and utilizes synchrotron radiation from a superconducting vertical wiggler (Huke and Yamakawa, 1980). The white X-ray beam, which polarize in the vertical direction, is first monochromatized by an Si(111) double-crystal monochromator and then focused by a platinum-coated fused-quartz mirror (Satow, 1984). Multiwavelength data were collected at three wavelengths: λ_1=1.077, λ_2=1.730 and λ_3=1.757Å. The data sets collected from the first crystal at λ_2 and λ_3, which are on either side of the *K* absorption edge of iron (λ_e=1.743Å), contain Friedel intensities. λ_1 is the wavelength near the *L*3 absorption edge of platinum, where the anomalous scattering effect of iron is rather small. Data collection with λ_1 was performed using the second crystal. Ideally, X-rays whose wavelengths are similar to each other and

are as close to the iron *K* edge as possible would be better in order to avoid systematic error owing to absorption and in order to obtain a large anomalous scattering effect. Such a choice of wavelengths was employed in the case of rubredoxin (Phillips *et al.*, 1977) and parvalbumin (Kahn *et al.*, 1985). However, in this study wavelengths were chosen in a different way: (1) λ_1 far from the absorption edge in order to obtain the maximum difference of *f'* from λ_2 or λ_3; (2) λ_2 and λ_3 slightly different from λ_e, since the actual bandwidth of a monochromatized incident X-ray beam may be broader than the energy resolution ($\Delta E/E$) of the optics (2×10^{-4}) owing to fluctuation of an electron orbit in the storage ring. Calibrations of the wavelengths were performed by measuring absorption spectra of a lyophilized cytochrome *c'* in the iron *K*-edge region and K_2PtCl_6 powder in the platinum L_3-edge region, respectively. As the intensity of the X-ray beam emitted from the storage ring decreased gradually during data collection, the intensity of X-ray beam on the crystal was monitored by an ion chamber (Satow, 1984). The raw data were corrected for Lorentz and polarization factors, radiation damage and absorption effects, using the program DRDCTN (Satow, unpublished work). The polarization factor was calculated assuming that the white X-ray beam incident to the monochromator was polarized 90% in the vertical direction. The absorption correction was based on the method of North *et al.*

(North *et al.*, 1968). The conditions of data collection are summarized in Table 4-1.

Table 4-1 Summary of the X-ray diffraction experiments

Wavelength (Å)	λ_1=1.077	λ_2=1.730	λ_3=1.757
f' [a]	0.159	-4.493	-4.596
f'' [a]	1.775	3.891	0.474
d spacing (Å)	12.3-4.0	14.0-6.0	14.4-5.5
ω-scan speed (°/min)	4.0	4.0	4.0
R_1 [b]	0.040	0.019	0.046
R_2 [c]	-	0.100	0.080
$\lvert\Delta F_{real}\rvert/\lvert F\rvert$ [d]	-	0.040	0.053
$\lvert\Delta F_{imag}\rvert/\lvert F\rvert$ [e]	-	0.063	0.051
R_2/R_1	-	5.26	1.74

[a] The values of f' and f'' were taken from Sasaki (1984), which is based on Cromer and Liberman's method (1970, 1981).

[b] $R_1^2=\Sigma_h\Sigma_i[I(h)_i-\langle I(h)\rangle]^2/\Sigma_h\Sigma_i[\langle I(h)_i\rangle]^2$, for equivalent reflections.

[c] $R_2^2=\Sigma_h[I(h)-I(\bar{h})]^2/\Sigma_h\{1/2[I(h)+I(\bar{h})]\}^2$.

[d] $\lvert\Delta F_{real}\rvert/\lvert F\rvert=\Sigma_h\lvert\lvert F_1(h)\rvert-\lvert F_i(h)\rvert\rvert/\Sigma\lvert F_1(h)\rvert$, i=2, 3.

[e] $\lvert\Delta F_{imag}\rvert/\lvert F\rvert=\Sigma_h\lvert\lvert F(h)\rvert-\lvert F(\bar{h})\rvert\rvert/\Sigma_h 1/2[\lvert F(h)\rvert+\lvert F(\bar{h})\rvert]$

4-2 Data Analysis

Figure 4-1 shows a vector diagram illustrating the contributions from various parts of cytochrome *c'* to a structure factor. $\overrightarrow{OA}$ represents the contribution from the non-anomalous scattering atoms in the protein and $\overrightarrow{AB}$ corresponds to the normal structure factor from the two iron atoms. $\overrightarrow{BC}$ is the variation owing to the real correction component (*f'*) of the anomalous scattering effect caused by iron. $\overrightarrow{CD}$ and $\overrightarrow{CE}$ are the variations owing to the imaginary correction component (*f''*).

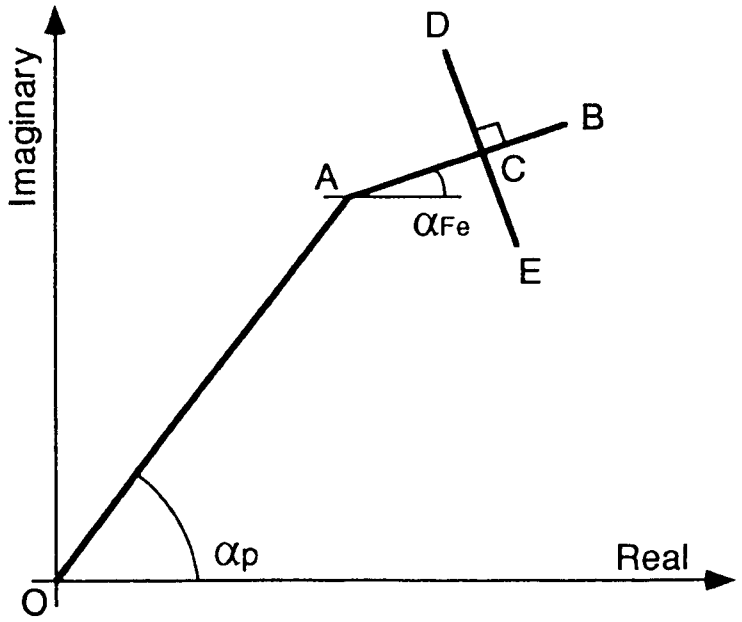


Fig. 4-1 A vector diagram of the structure factor representing the effects of anomalous scattering at various wavelengths.

In this study, structure amplitudes $|F^+_1|$, $|F^+_2|$, $|F^-_2|$, $|F^+_3|$ and $|F^-_3|$ were measured for three wavelengths (Table 4-1), where F^+_n and F^-_n refer to *hkl* and $\overline{hkl}$ reflections measured with λ_n (n=1,2 and 3). As the values of f'' at λ_1 and λ_3 are small (Table 4-1), the following relations are presumed: $|F^+_1|=|F^-_1|(=|F_1|)$ and $|F^+_3|=|F^-_3|(=|F_3|)$. Furthermore, the values of f' at λ_2 and λ_3 are approximately equal to each other and are larger than that at λ_1 (Table 4-1). $|F_1|$, $|F^+_2|$, $|F^-_2|$ and $|F_3|$ are related to the vector diagram in Fig. 4-1 as follows: $|F_1|=|\overrightarrow{OB}|$, $|F^+_2|=|\overrightarrow{OD}|$, $|F^-_2|=|\overrightarrow{OE}|$ and $|F_3|=|\overrightarrow{OC}|$. Furthermore, the average value of $|F^+_2|$ and $|F^-_2|$, that is $|F_2|$, equals $|\overrightarrow{OC}|$ to a good approximation. The statistics of the multiwavelength data are summarized in Table 4-1. R_1 and R_2 are reliability factors for equivalent reflections and Friedel intensities, respectively; the ratio R_2/R_1 indicates the contribution of f'' to the diffracted intensities. In this case, R_2/R_1 at λ_2 (5.26) is large compared with that at λ_3 (1.74), indicating that the contribution of f'' to the diffracted intensities is larger at λ_2 than at λ_3, as expected. $|\Delta F_{real}|/|F|$ and $|\Delta F_{imag}|/|F|$ represent the change in structure amplitudes. The former is due to the difference of f' for $|F_1(hkl)|$ and $|F_n(hkl)|$, n=2, 3, corresponding to a change in diffraction pattern caused by the addition of heavy atoms in the isomorphous replacement method. The latter is the difference between Friedel intensities. Although $|\Delta F_{imag}|/|F|$ at λ_2 is comparable with values expected in heavy-

atom derivatives, $|\Delta F_{real}|/|F|$ is of the order of a tenth of the changes usually found in the isomorphous replacement method. The unexpectedly large value of $|\Delta F_{imag}|/|F|$ at λ_3 is not a result of Friedel-pair splitting, but a result of low accuracy in data collection.

4-3 Location of the iron atoms

The positions of the two iron atoms in the crystal were determined from difference Patterson and difference Fourier maps. The coefficients (a) $(|F_1|-|F_2|)^2$, (b) $(|F_1|-|F_3|)^2$ and (c) $(|F^+_2|-|F^-_2|)^2$ were used for the difference Patterson maps. The maps for (a) and (b) are based on the differences in f' for the two wavelengths and are analogous to the isomorphous difference Patterson map in the case of the isomorphous replacement method. The map for (c) is a regular anomalous difference Patterson map based on f'' (Blundell and Johnson, 1976b). Although the two Patterson maps at w=1/2 with coefficients (a) and (b) (Fig. 4-2) were somewhat noisy and the highest peak did not coincide with the iron self vector, the positions of the two iron atoms could be determined by inspecting three Harker sections (w=1/6, 1/3 and 1/2) and they were confirmed by checking iron cross vectors. The high-noise peaks which appear in the same positions of Figs. 4-2(a) and (b) may be a result of systematic error caused by the absorption effect, because the absorption effect at λ_2 and

λ_3 is larger than λ_1. Despite the fact that values of f' and f'' are of the same order of magnitude as each other and that Friedel intensities were measured from the same crystal during the same beam time, the map for (c) is worst among the three difference Patterson maps (Fig. 4-2). Furthermore, the combined difference Patterson map with coefficient :

$$(|F_1|-|F_2|)^2+(k/2)^2(|F^+_2|-|F^-_2|)^2,$$

which must be a better estimate for iron than (a) and (c) (Blundell and Johnson, 1976b), was similar to the map for (a). These observations may be explained by the relation between α_{Fe} (phase angle of iron contribution) and α_p (phase angle of protein). Distribution of $|\alpha_{Fe}-\alpha_p|$ for 458 reflections up to 6 Å resolution tends to 0 or π rather than random chance (Table 4-3), suggesting that the map for (c) is not a good estimate for iron (Blundell and Johnson 1976b).

Difference Fourier maps (Fig. 4-3) were calculated with the coefficients (a) $(|F_1|-|F_2|)\exp(i\alpha)$, (b) $(|F_1|-|F_3|)\exp(i\alpha)$ and (c) $(|F^+_2|-|F^-_2|)\exp[i(\alpha-\pi/2)]$, where α is the protein phase angle obtained by the multi-isomorphous replacement method as described in chapter 3. Each difference Fourier map revealed two large peaks at the expected positions of the two iron atoms. As f'' is very small at λ_3 (Table 4-1), Patterson and Fourier maps based on the difference between $|F^+_3|$ and $|F^-_3|$ were featureless and did not give the positions of the two iron atoms.

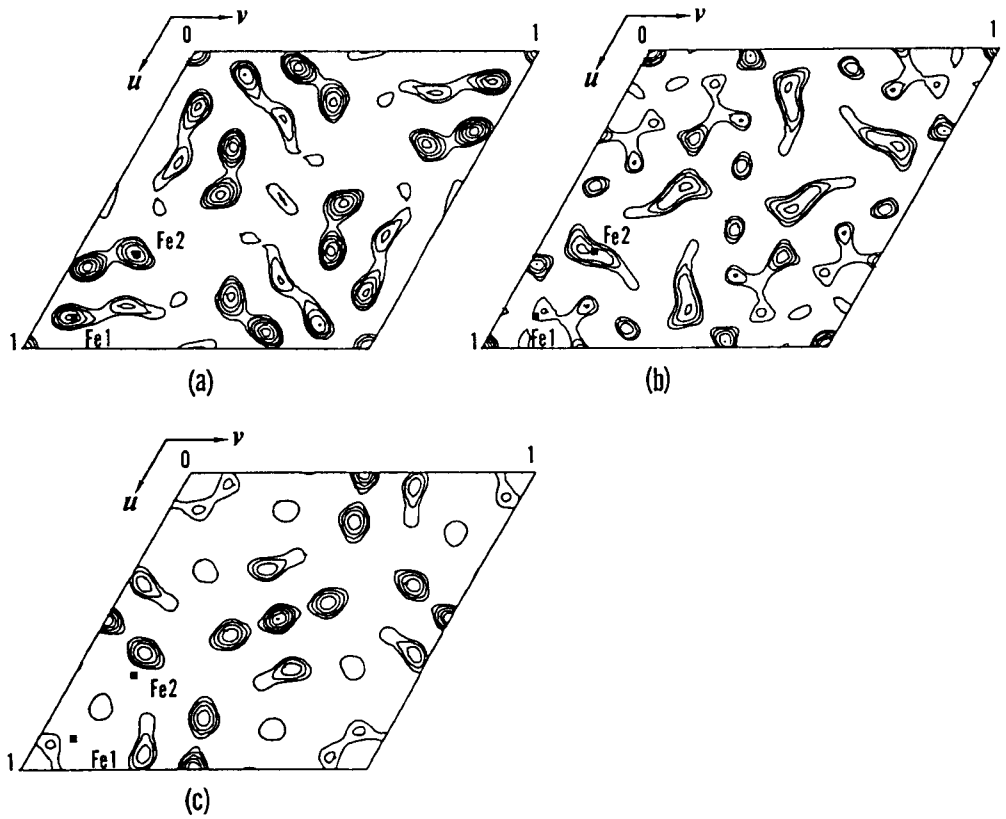


Fig. 4-2 Patterson maps of cytochrome *c'* (w=1/2) with coefficient (a) $(|F_1|-|F_2|)^2$, (b) $(|F_1|-|F_3|)^2$ and (c) $(|F^+_2|-|F^-_2|)^2$. The positions of iron self vectors are marked. For the definitions of $|F_1|$, $|F^+_2|$, $|F^-_2|$, $|F_2|$ and $|F_3|$, see text.

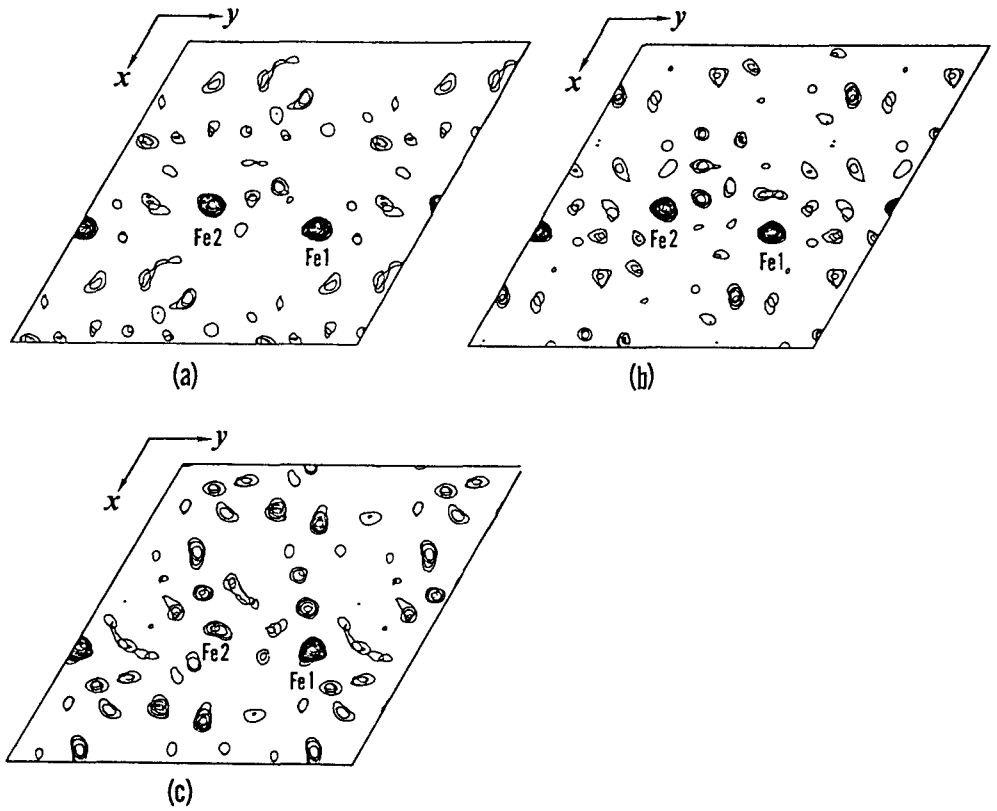


Fig. 4-3 Composites of four sections (z=0.10 - 0.16) of the difference Fourier maps of cytochrome c' with coefficient (a) $(|F_1|-|F_2|)\exp(i\alpha)$, (b) $(|F_1|-|F_3|)\exp(i\alpha)$ and (c) $(|F^+_2|-|F^-_2|)\exp[i(\alpha-\pi/2)]$, where α are phase angles obtained by the MIR method. The two iron positions (Fe1 and Fe2) are indicated.

4-4 Phase calculation

As shown in Fig. 4-1, a change in f' with wavelength produces a change in X-ray diffraction pattern; intensity data collected with different wavelengths can be used to calculate phase angles in the same way as the isomorphous replacement method, if positions of atoms which cause anomalous scattering effects are known. The best phase angles of cytochrome *c'* were calculated by the MAD method using the X-ray intensity data collected with λ_1 and λ_2. Although f' at λ_3 is close to that at λ_2, f'' at λ_3 is very small (Table 4-1). So the intensity data collected with λ_3 were not included in the phase calculation. The method of phase calculation is very similar to that of single isomorphous replacement combined with anomalous scattering data which was described in chapter 3-1. In this case, the structure amplitude of native crystal, $|F_P|$, is replaced by $|F_1|$, and that of heavy atom derivative, $|F_{ph}|$, is replaced by $|F_2|$. The f' contribution of the two iron atoms at λ_2 was used in place of the contribution of heavy atom to the structure amplitude, F_h. All calculation were performed using the same program as the MIR method with minor modification. Parameters of the two iron atoms were refined by minimizing

$$[|F_2| - ||F_1|\exp(i\alpha_{best}) + f\ |]^2,$$

where α_{best} is the best phase angle calculated by the MAD method. The refined atomic parameters of heme-iron atoms are

summarized in Table 4-2. The resulting best phase angles were compared with those obtained by the isomorphous replacement method (chapter 3-4) and the average phase discrepancy for the 462 reflections up to 6Å resolution was 75°. The phase angles for *hk*0 reflections of rubredoxin calculated by the MAD method were reported to differ by 60° on the average from the phases of the refined model (Phillips *et al.*, 1977). And in the case of Tb-substituted parvalbumin, the MAD phases were also compared with phases calculated from a model and got an average phase discrepancy of about 54° (Kahn *et al.*, 1985). Figure 4-4(b) shows the electron density map calculated by the MAD method. Although the electron density map is noisy compared with that calculated by the isomorphous replacement method (Fig. 4-4(a)), α – helices and the molecular boundary can be recognized.

Table 4-2 Refined atomic parameters of heme-iron atoms

	x	*y*	*z*	B(Å^2)
Fe1	0.9393	0.5751	0.1265	17.93
Fe2	0.8186	0.0552	0.1420	3.64

Table 4-3 Distribution of 458 reflections as a function of $\Delta = |\alpha_{Fe} - \alpha_p|$

Δ/°	0	30	60	90	120	150	180
(a)	110	47	57	67	79	98	
	(24.0%)	(10.3%)	(12.4%)	(14.6%)	(17.2%)	(21.4%)	
(b)	152	36	33	36	62	139	
	(33.2%)	(7.9%)	(7.2%)	(7.9%)	(13.5%)	(30.3%)	

α_p is calculated by (a) MIR and (b) MAD methods

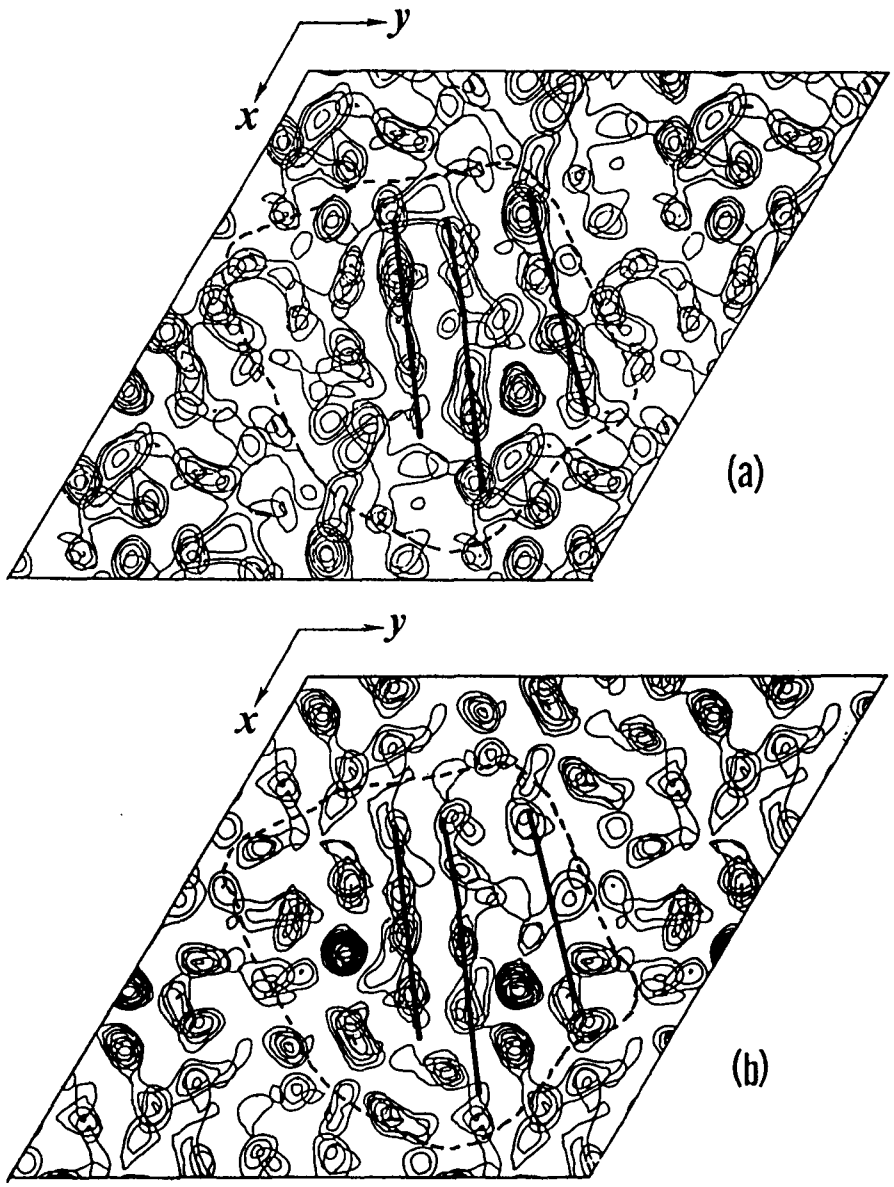


Fig. 4-4 Composites of four sections (z=0.10 - 0.16) of the electron density maps around the two heme iron atoms of cytochrome c' calculated by (a) the multiple isomorphous replacement method and (b) the multiwavelength anomalous diffraction method. A dimeric molecule is surrounded by a dotted line in each map. Rods of electron density indicated by heavy lines correspond to α – helices.

4-5 Concluding Remarks

In this study, location of the heme irons and calculation of the phase angles were successfully performed for cytochrome *c'* by using the multiwavelength data collected at three wavelengths (λ_1, λ_2 and λ_3). The method of phase calculation was basically the same as that of single isomorphous replacement combined with the anomalous scattering effect. As it seems that the actual bandwidth of an X-ray beam depends on the stability of the orbit of an electron in the storage ring rather than the energy resolution of the optics (ΔE is nearly 1.4eV at λ_e), wavelengths which were sufficiently far from λ_e by about 55 eV were used for data collection. The improvement in the stability of the synchrotron radiation will allow a more delicate choice of wavelengths and , consequently, accurate phase determination from multiwavelength data.

5 Structure Analysis at 2.8Å Resolution

5-1 Materials and Method

5-1-1 Preparation of crystals and heavy atom derivatives

Crystals of *Rhodospirillum rubrum* cytochrome *c*'(RCP) were made by the method described in chapter 2-2. Platinum derivative was prepared with the same way as in chapter 3-2. Mercury derivatives was prepared by soaking the native crystals in 10mM CH_3HgCl for 4 days, which was longer than the period in chapter 3-2.

5-1-2 Data collection

X-ray intensity data from the native crystals and the two derivatives were collected to 2.8Å resolution on a Rigaku four-circle diffractometer (AFC5R) at 10 °C. Ni-filtered Cu$K\alpha$ radiation was provided by a Rigaku rotating anode X-ray generator operated at 40kV and 300mA. Crystals with typical dimensions of $0.3\times 0.3\times 0.7$ mm^3 were used for data collection. The ω-scan method with a scan width of 1 ° and a scan rate of 4 °/min was employed. Background intensities were counted for 4 s at both ends of a scan. Three standard reflections were measured after every 100 measurements to monitor radiation damage and crystal slipping. Data sets for the native and the platinum derivative were included Bijvoet pair reflections (*hkl* and $hk\bar{l}$). The raw

data were corrected for Lorentz and polarization factors, and absorption effects using the method of North *et al.* (North *et al.*, 1968). In most cases, a crystal was replaced by a fresh one when the average structure amplitude of the standard reflections decayed to 90% of the starting value. The data were then corrected for the radiation damage using the equation $|F_{corr}|=|F_{obs}|\exp(aN)$, where $|F_{corr}|$ is the corrected amplitude, $|F_{obs}|$ is the observed amplitude, a is the damage coefficient calculated from the decay of the average value of the standard reflections and N is the sequence of the measurement. Table 5-1 shows the summary of the data collection.

Table 5-1 Statistics of diffraction data for crystals used in the structure analysis (2.8Å resolution)

	cell constants		No. of	No. of observed	No. of independent	R_{merg} [b]
	a/Å	c/Å	crystals	reflections	reflections [a]	
native	51.72	155.49	8	21203	12092	0.039
K_2PtCl_6	51.76	155.20	11	32250	10791	0.104
CH_3HgCl	51.65	154.92	6	14291	5851	0.036

[a] including Bijvoet pair reflections in native and platinum derivative

[b] $R_{merg} = \Sigma_h|\Sigma_i|F_{i,h}|-\langle|F_h|\rangle|/N\Sigma_h\langle|F_h|\rangle$, where $|F_{i,h}|$ is an individual measurement of reflection h of the i-th crystal, $\langle|F_h|\rangle$ the mean amplitude of that reflection and N the number of crystals.

5-1-3 Phase calculation

Two major heavy-atom sites of the platinum derivative were located by calculating an isomorphous difference Patterson map as described in chapter 3-4. Isomorphous difference Fourier maps of the mercury derivative were calculated using phase angles obtained by the single isomorphous replacement method combined with the anomalous dispersion data of the platinum derivative. Three mercury sites could be located in the difference Fourier map which were consistent with the isomorphous difference Patterson map of the mercury derivative, whose map was unsolved solely because the mercuric self vectors were mostly overlapped. Heavy atom parameters were refined by minimizing the lack-of-closure error for each derivative. The final heavy atom parameters and the statistics of the refinement were summarized in Table 5-2.

Table 5-2 Heavy atom parameters used in the final phase calculation at 2.8Å resolution

Site [a]	Occupancy (electrons)	*x*	*y*	*z*	B (Å^2)	Kraut *R* [b]
		(fractional units)				
[Pt derivative]						0.125
Pt1 (monomer I)	55.6	0.757	0.472	0.129	28.9	
Pt2 (monomer II)	55.6	0.967	0.159	0.167 [c]	31.5	
[Hg derivative]						0.064
Hg1-1 (monomer I)	44.44	0.881	0.595	0.097	32.0	
Hg1-2 (monomer I)	43.70	0.329	0.918	0.108	29.6	
Hg2 (monomer II)	44.75	0.934	0.066	0.131	31.3	
		Total average of figure-of-merit = 0.60				

[a] Pt1, Hg1-1 and Hg1-2 bind to one molecule in the asymmetric unit (monomer I), Pt2 and Hg2 to the other molecule (monomer II).

[b] Kraut $R = \Sigma \, ||F_{ph}|-|F_p|\exp i\alpha_{best}+F_h|| / \Sigma \, |F_{ph}|$, where $|F_{ph}|$ and $|F_p|$ are structure amplitudes of the heavy atom derivative and the native crystals, respectively. F_h is heavy atom contribution.

[c] The *z* coordinate of Pt2 was fixed to define the origin.

5-1-4 Improvement of the electron density map

The electron density map at 2.8Å resolution was calculated using the best phase angles obtained by the MIR method, and a part of the MIR map containing one RCP dimer molecule was contoured on transparent sheets with a 3.5Å /cm scale to obtain a mini-map. Although the dimer structure of RCP, four α – helices and the heme group of each monomer, could be recognized in the mini-map, the electron density of the loop regions which connect α – helices was poor. Since visual inspection of the mini-map indicated that the two monomers (monomer I and monomer II) in the dimer were related by a non-crystallographic 2-fold axis, and that the heavy-atom sites in the dimer molecule were also related by this symmetry, the electron density was averaged about the non-crystallographic 2-fold axis (Bricogne, 1976 ; Kusunoki, 1980) in order to improve the quality of the map. The orientation of the non-crystallographic 2-fold axis was initially obtained by least-squares fitting of the positions of equivalent heavy-atoms (Table 5-2) and the heme-irons (Table 4-2) in the monomer I and monomer II. The resulting relation between the monomer I (*XYZ*) and monomer II(*xyz*) is

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} -0.82802 & 0.03512 & -2.03083 \\ -0.02344 & -0.96894 & -0.67577 \\ -0.19202 & 0.05353 & 0.80312 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} 1.9036 \\ 0.7437 \\ 0.1667 \end{bmatrix},$$

where (*XYZ*) and (*xyz*) are in fractional unit.

A molecular envelope of the monomer I was traced on the mini-map. The matrices were further refined by the least-squares method so as to give the maximum correlation between the electron densities inside the molecular envelopes of the two monomers. The electron densities of the two monomers were then averaged using the refined matrices. The electron density outside the molecular envelopes was set to a constant value of zero. The electron density of the asymmetric unit was then reconstructed using the averaged electron density. Phase probability function was calculated for each reflection from the resultant electron density using the Sim's formula (Bricogne, 1976). It was combined with phase probability function derived by the MIR method and the improved phase angle was calculated for each reflection. The electron density map based on the combined phase angles was again subjected to the same procedure mentioned above. This iterative process was repeated 5 cycles. The *R*-factor, correlation coefficient between the observed and calculated structure amplitudes, and mean phase change from previous cycle were converged to 0.286, 0.801 and 2.26°, respectively (Table 5-3). The refined matrices are

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} -0.72948 & -0.11125 & -2.12394 \\ 0.03439 & -1.01413 & -0.25130 \\ -0.22104 & 0.08925 & 0.74361 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} 1.8349 \\ 0.6452 \\ 0.2007 \end{bmatrix} .$$

The mean distance between the corresponding superposed

heavy-atom positions using these matrices is 0.8Å.

Since the final averaged electron density map was far better than the MIR map (Fig. 5-1), the course of the main-chains in the loop regions as well as in the α – helix regions could be traced easily. Although the electron density corresponding to the C-terminal 4 residues (Phe123-Lys-Ala-Asp126) was not visible clearly, the approximate positions of 122 α – carbon atoms from Ala1 to Glu122 were identified for each monomer.

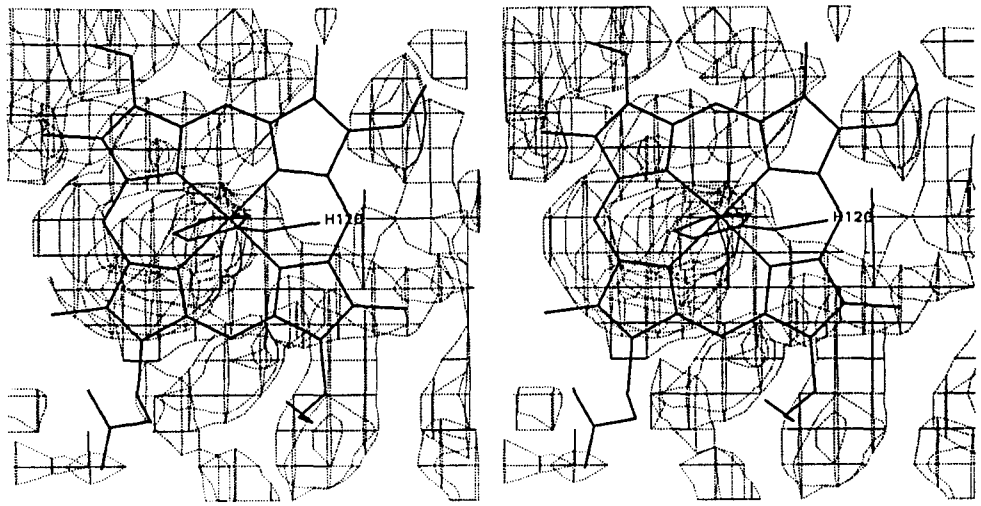
Table 5-3 Summary of phase refinement

Cycle number	*R*-factor [a]	Q [b]	$\langle \lvert \Delta \alpha \rvert \rangle$ [c]
1	0.390	0.648	23.16
2	0.358	0.696	8.57
3	0.322	0.752	3.89
4	0.306	0.778	4.71
5	0.286	0.801	2.26

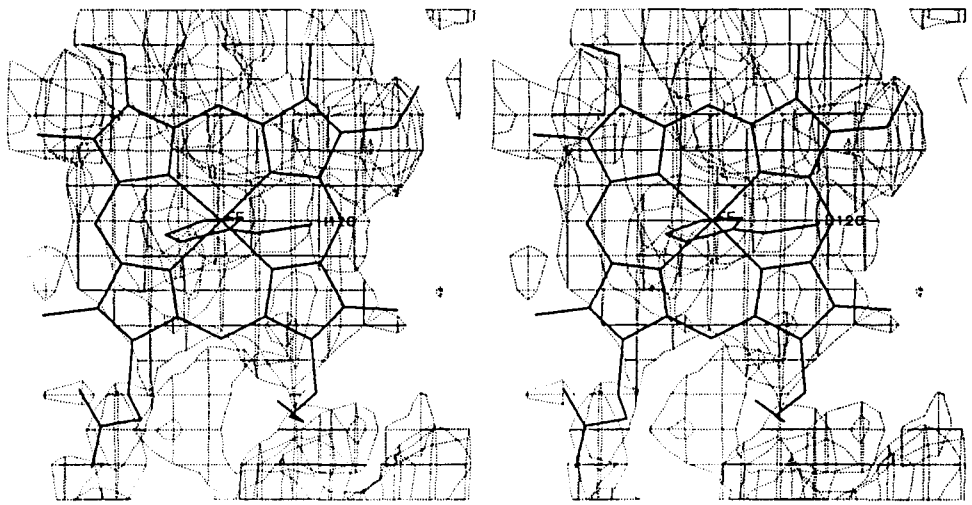
[a] $R=\Sigma \lvert \lvert Fo \rvert - \lvert Fc \rvert \rvert / \Sigma \lvert Fo \rvert$.

[b] Correlation coefficient between $\lvert Fo \rvert$ and $\lvert Fc \rvert$: $Q=(\Sigma \lvert Fo \rvert \lvert Fc \rvert - 1/N \Sigma \lvert Fo \rvert \Sigma \lvert Fc \rvert)/[(\Sigma \lvert Fo \rvert^2 - 1/N \Sigma \lvert Fo \rvert \Sigma \lvert Fo \rvert)(\Sigma \lvert Fc \rvert^2 - 1/N \Sigma \lvert Fc \rvert \Sigma \lvert Fc \rvert)]^{1/2}$, where N is the number of reflections.

[c] Mean phase change from previous cycle.



(a)



(b)

Fig. 5-1 The superpositions of electron density map and the atomic model which was refined at 2.8Å resolution. Each map contains a heme group and a side chain of His120, calculated from (a) MIR phases and (b) improved phases.

5-1-5 Model building

Model building was carried out using interactive computer graphics program MOLFITG (Iga and Yasuoka, 1984) implemented on a graphic display system with the host computer ACOS S850 at the Research Center for Protein Engineering, Institute for Protein Research, Osaka University. The improved electron density map was displayed on the graphic display and a molecular model of the monomer I was built according to the published amino acid sequence (Meyer *et al.*, 1975). Atomic coordinates of the monomer II were derived from those of the monomer I using the non-crystallographic symmetry matrices obtained by the method described above.

5-1-6 Crystallographic refinement

The restrained-parameter least-squares refinement program, PROLSQ, of Hendrickson and Konnert (Hendrickson and Konnert, 1981) was used at the first stage of the refinement of RCP. 4481 reflections measured with $|F_o| \geqq \sigma(F)$ between 5.0 and 2.8Å resolution were used for the refinement. An overall temperature factor (15Å^2) was adopted for all protein atoms and was fixed until an advanced stage of the refinement, to avoid unrealistic shift. Cycles of refinement were carried out until the applied shifts did not appear to improve the model. The atomic coordinates and an overall scale factor were allowed to vary. At

the end of cycle 18, the refined model was inspected against 3 types of electron density maps: (1) the MIR map, (2) the averaged map, (3) the omit map computed with the coefficient $:(2|F_o|-|F_c|)\exp i\alpha_{calc}$. After rebuilding the model, the successive cycles of the refinement were continued to convergence. During this stage, the *R*-factor was reduced from 0.44 to 0.32. For the next stage of the refinement, the X-PLOR program package (Brunger *et al.*, 1987) was used. A molecular dynamics simulation was performed for 0.67 ps at 3000 K using a time step of 0.67 fs. The temperature was subsequently raised to 4000 K and then the structure was slowly cooled to 300 K using a 50 K temperature drop for every 50 dynamics steps of 0.5 fs. Final energy minimization and refinement of individual isotropic temperature factors gave the *R*-factor of 0.22. The atomic parameters thus obtained were further refined 6 cycles with PROLSQ, resulting in an *R*-factor of 0.23. Although this refinement step using PROLSQ did not reduce *R*-factor, stereochemical parameters such as bond distances, bond angles, etc. were improved. The current model consists of 1840 protein atoms, which represent the dimer structure containing the monomers of 122 amino acid residues. C-terminal 4 residues (Phe123 to Asp126) and solvent molecules were not included. The r.m.s. deviation of bond distances from ideal values is 0.022Å. The r.m.s. error of the atomic positions estimated by the Luzzati method (Luzzati, 1952) is about 0.25~

0.30Å .

5-2 Results and Discussion

5-2-1 Monomer structure

Although the homology of the amino acid sequences shown in Fig. 5-2 between cytochrome *c'* from *Rhodospirillum rubrum* (RCP) and that from *Rhodospirillum molischianum* (MCP) is low (about 24%) except for the region between Cys116 and His120, and chemical properties of side-chains such as bulkiness, hydrophobicity and electric charge are remarkably different between the corresponding amino acid residues of the two cytochromes *c'* (Fig. 5-2), it was revealed in this study that their main-chain structures resemble each other. In the similar manner as MCP, the monomer molecule of RCP is mainly composed of four α – helices (Fig. 5-3). The first α – helix (A-helix) which consists of 27 residues, is initiated with Ala1 and terminated with Thr27. The second helix (B-helix) consists of 13 residues from Thr38 to Glu50, the third one (C-helix) of 19 residues from Glu77 to Ser95 and the fourth one (D-helix) of 21 residues from Lys102 to Glu122. The polypeptide segments between adjacent helices are non-regular loop structures (AB-loop: Leu28-Gln37, BC-loop: Thr51-Phe76, CD-loop: Ala96-Asp101). The C-terminal 4 residues (Phe123-Asp126), were not included in the model due to the poor electron density of this region. As was found in the MCP (Weber

and Salemme, 1980), the four α – helices in the monomer molecule of RCP are arranged in the way that adjacent helix axes orient antiparallel and form a left-twisted four-α – helix bundle. The four α – helices are connected by the AB and CD-loops at one end of the bundle (double connected end) and by the BC-loop at the other end (single connected end). Since the AB and CD-loops (10 and 6 residues, respectively) are shorter than the BC-loop (26 residues), the helix axes spatially diverge from the double connected end to the single connected end where a binding pocket for the heme is formed (Fig. 5-3). The overall shape of the monomer is a truncated cone with the diameters of 15Å at the double connected end and 27Å at the single connected end, and with a height of 42Å.

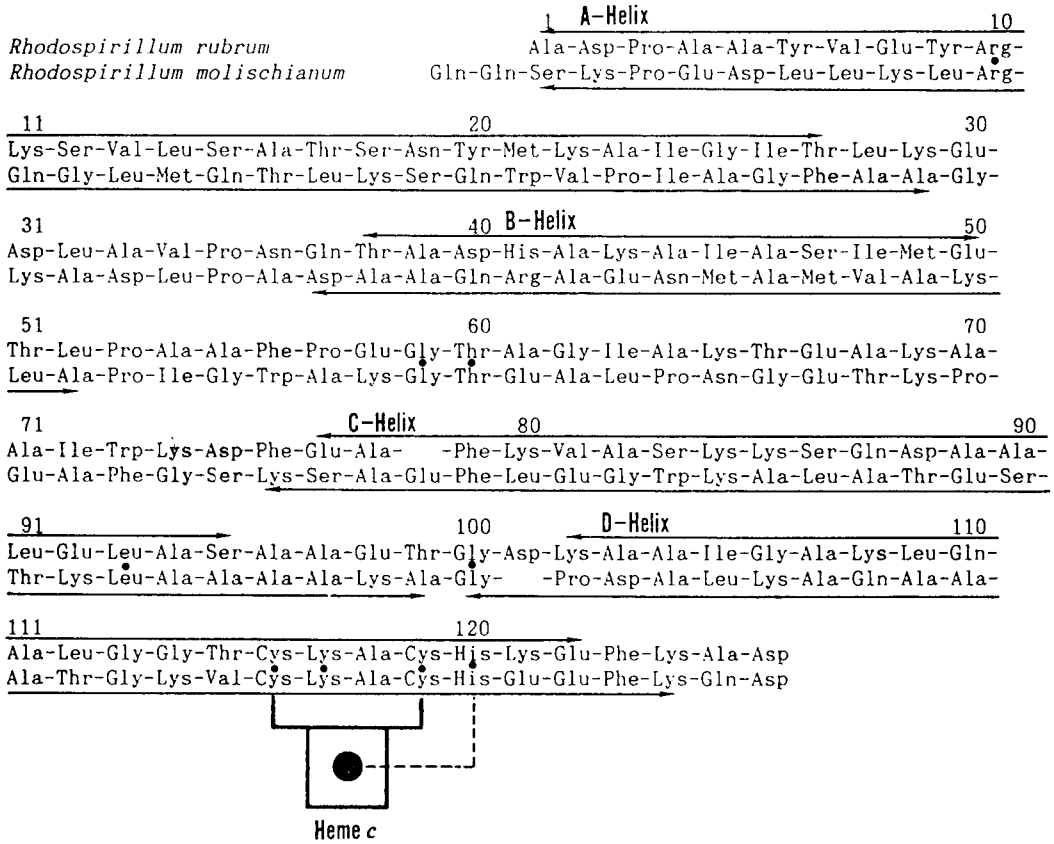


Fig. 5-2 The primary structures of RCP (upper) and MCP (lower) arranged for the highest coincidence each other. The α – helices are indicated by the arrows. Filled circles indicate the strictly conserved amino acid residues in many cytochromes *c'*. RCP shares 24% amino acid sequence homology with MCP.

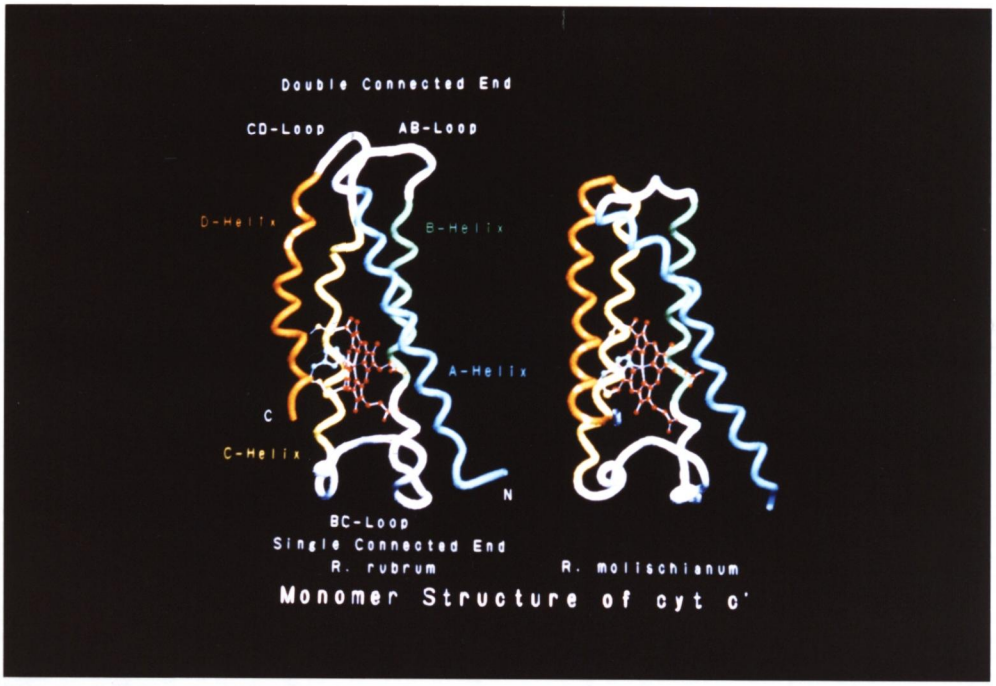


Fig. 5-3 The tube model of the main chain structure of (left) the RCP monomer molecule (Ala1 to Glu122) and (right) the MCP. A, B, C and D-helices are colored blue, green, yellow and orange, respectively. AB, BC and CD-loops are white. The heme prosthetic group, two cysteine residues (Cys116 and Cys119) to which the heme is covalently bound through thioether linkages, and His120 (the fifth axial ligand of the heme-iron) are indicated by ball-and-stick model.

5-2-2 Heme environment

The heme is located at the pocket near the open single connected end. Two cysteine residues (Cys116 and Cys119) are covalently bound to the heme vinyl groups of pyrrole rings I and II through thioether linkages. The heme plane is roughly parallel to the plane containing the diagonally related A and C-helix axes. The heme is displaced from the center of the bundle toward the single connected end of the bundle, and toward the D-helix (Fig. 5-3). One face of the heme plane (exterior face) is somewhat exposed to solvent compared with another face (interior face). Figures 5-4(a) and 5-4(b) show the structure of the heme and its environment of RCP. The interior face of the pyrrole rings II and III of MCP is surrounded by aromatic side-chains of Trp58, Phe82, Trp86 and Phe125, particularly side-chain of Trp86 is roughly parallel to the heme plane and within van der Waals contact of pyrrole II (Weber, *et al.*, 1981). The corresponding aromatic side-chains of RCP are Phe56, Phe79 and Phe123. Trp86 of MCP is replaced by Ser83 in the case of RCP, and hence no residues are in van der Waals contacts to the pyrrole II of RCP. As a whole, the heme of RCP lies in somewhat less hydrophobic and less crowded environment than that of MCP. The polypeptide segment, Cys116-Lys-Ala-Cys-His120, is located on the side of the exterior face of the heme and the Nε 2 atom of His120 forms a coordinate bond with the heme-iron atom as the fifth axial ligand.

The bond distance is 1.9Å for both monomers. The sixth axial ligand site of the heme-iron atom is occupied by the side-chain of Leu14, which does not form a coordinate bond with the iron atom. Amino acid sequences of twelve kinds of cytochromes *c'* (Ambler *et al.*, 1981) indicate that sixth ligand site is occupied by one of Met, Leu, Tyr, and Phe; moreover, Met16 of MCP is oriented to disallow coordination with the iron atom (Weber and Salemme, 1977). These facts suggest that the lack of sixth-ligand is a general feature in cytochromes *c'*. The heme pyrrole rings of the two monomers are fairly planar, but the heme planes significantly deviate from perfect planarity. The average deviation of the 24 porphyrin atoms from the mean porphyrin plane, which is determined by least-squares fit using 24 atoms in porphyrin core, is 0.14Å for both monomers. The heme-iron atom is displaced 0.5Å from the mean pyrrole nitrogen plane, defined by the 4 pyrrole nitrogen atoms, toward the coordinating histidine imidazole ligand (Fig. 5-4(b)). Although these values are comparable to the r.m.s. error of the atomic positions, the heme geometry of RCP mentioned above coincide with that of MCP. The heme propionates of pyrrole rings III and IV are extended toward the molecular surface and interact with Arg10 and Lys11, respectively, of the same subunit through salt linkages. Since Arg10 of RCP is one of the strictly conserved residues over the cytochromes *c'* from 12 species of photosynthetic bacteria (Ambler

et al., 1981) and the corresponding arginine found in MCP (Arg12) also interacts with the propionate of the pyrrole ring III (Weber, *et al.*, 1981), this salt linkage may be commonly observed in the cytochromes *c'*. On the other hand, in most of other cytochromes *c'*, Lys11 of RCP is replaced by Gln, which interacts with the propionate of the pyrrole ring IV in the case of MCP.

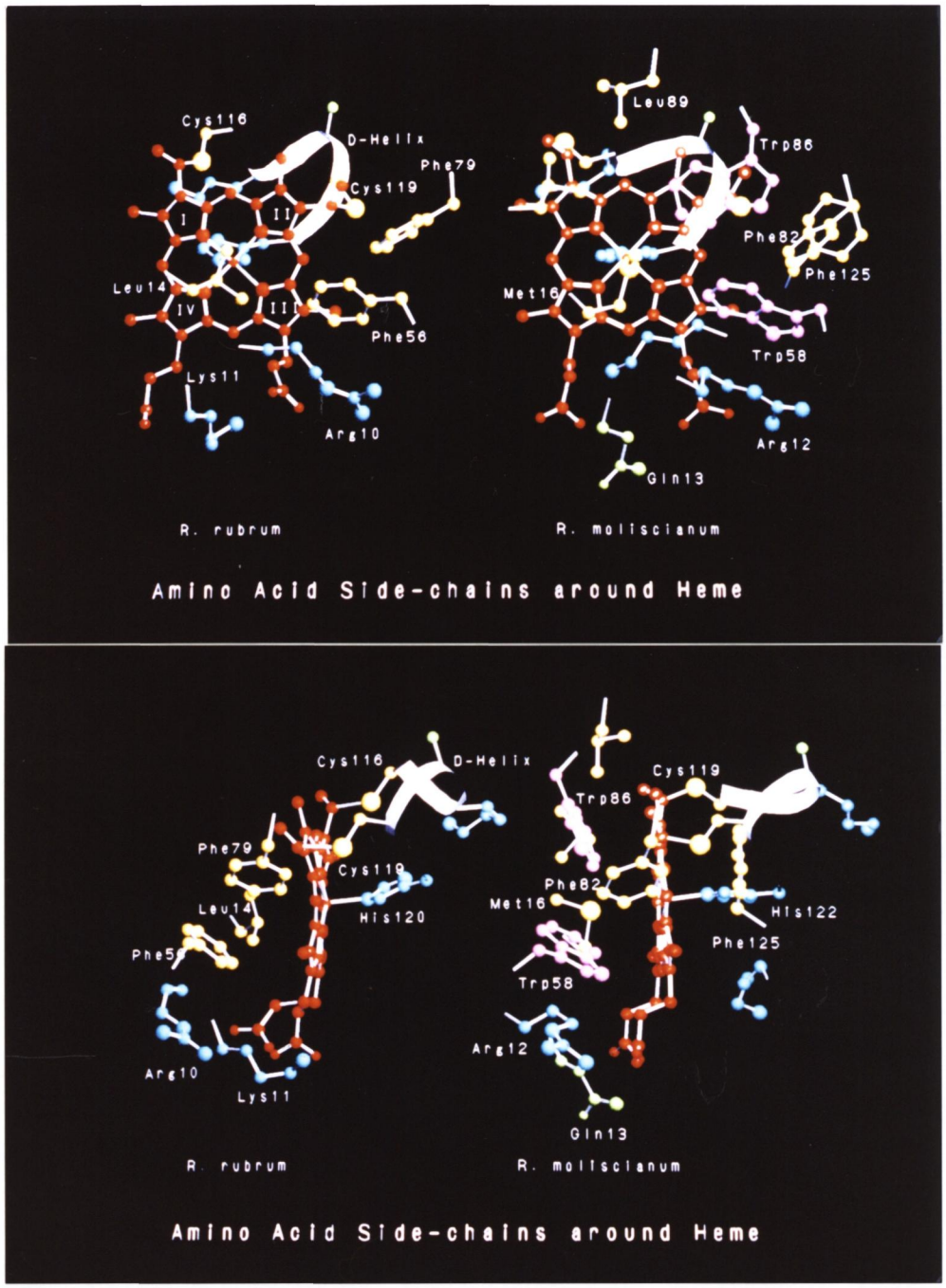
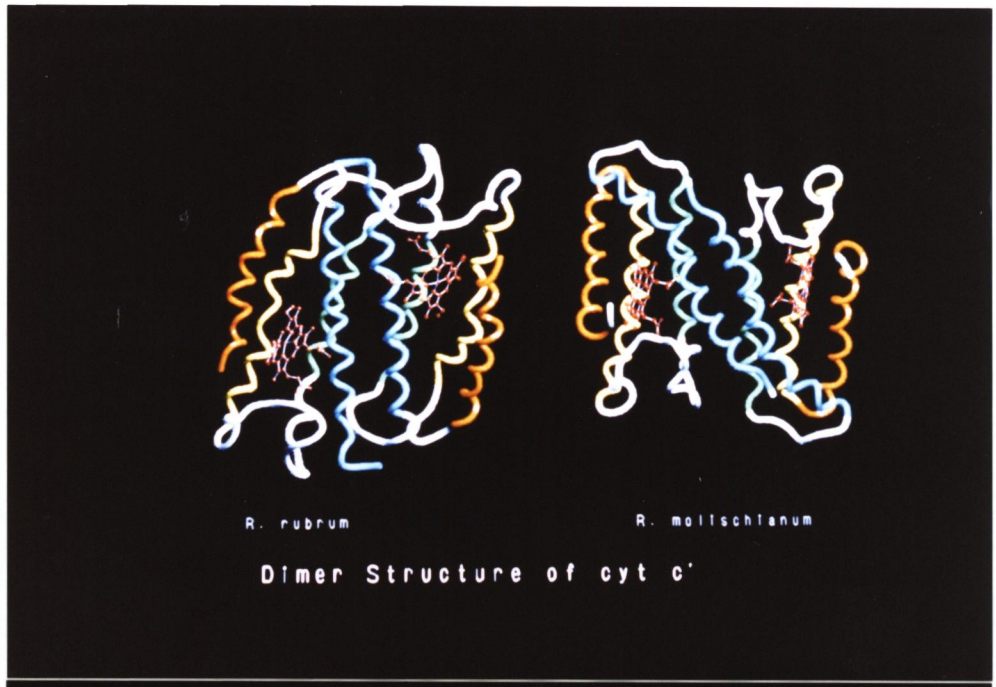


Fig. 5-4 Views of the heme prosthetic group of (left) RCP and (right) MCP from different directions. The two cysteine side-chains form thioether linkages with the heme vinyl groups. The Nε 2 atom of His120 forms the coordinate bond to the heme-iron as the fifth axial ligand. Two aromatic residues, Phe56 and Phe79, are near the heme group. The sixth ligation site of the heme-iron atom is occupied by the side-chain of Leu14, which do not form the coordinate bond with the iron. The heme propionates interact with Arg10 and Lys11 in the same monomer.

5-2-3 Dimer structure and inter-subunit interaction

The RCP dimer is composed of the chemically identical polypeptide chains which are related by the non-crystallographic 2-fold axis (Fig. 5-5). The overall dimensions of the dimer, when viewed from the 2-fold axis, are about 50Å high, 43Å wide and 29Å deep. The six helices of the dimer, D, C, B, B', C' and D', are arranged antiparallel in this order, and form a continuous twisted sheet with pronounced curvature. The A and A'-helices and the two heme prosthetic groups are incorporated into the concavity of the sheet. The two heme groups are oriented so that the fifth ligands (His120) coordinate to Fe atoms from the exterior of the dimer molecule. The angle between two heme planes is 68° and the distance between two Fe atoms is 25Å. The molecular interface between the two monomers is formed by two pairs of the anti-parallel helices from each subunit. These four helices (A and B-helices of the monomer I, A' and B'-helices of the monomer II) are also packed together to form an AA'BB'-bundle as was observed in the individual monomer structure. There are inter-subunit interactions between A and A'-helices, and between B and B'-helices which stabilize the dimer structure. At the AA' interface, 2-fold symmetrical hydrogen bonds are formed between the two Tyr20 side-chains, and between Asn19 of one helix and Ser15 of the other (Fig. 5-6(a)). Similarly the B-helix interacts with the B'-helix through hydrogen bonds formed

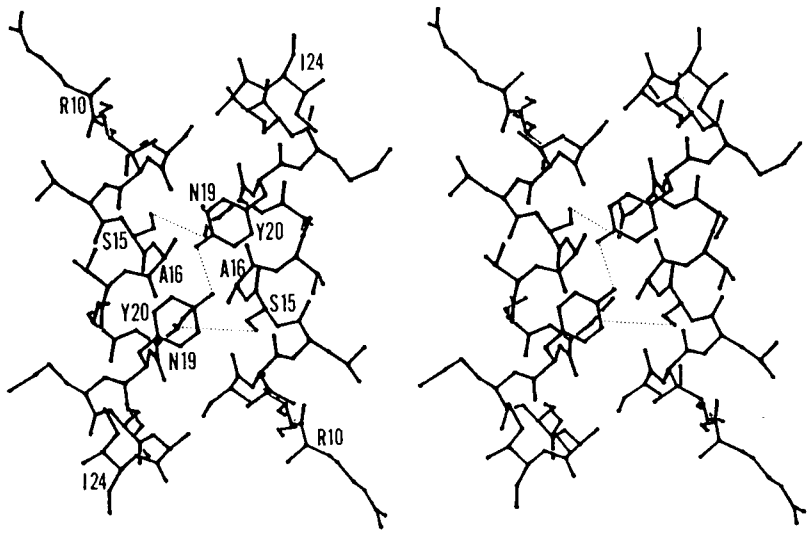
by Gln37(B)-Thr51(B'), Thr51(B)-Gln37(B'), Asp40(B)-Ser47(B') and Ser47(B)-Asp40(B') (Fig. 5-6(b)). These electrostatic interactions observed in the subunit interface of the RCP dimer are contrastive to hydrophobic interactions of MCP (Weber, *et al.*, 1981). Such difference may be reflected by the low homology of the amino acid sequences of the two cytochromes *c'* (Fig. 5-2).



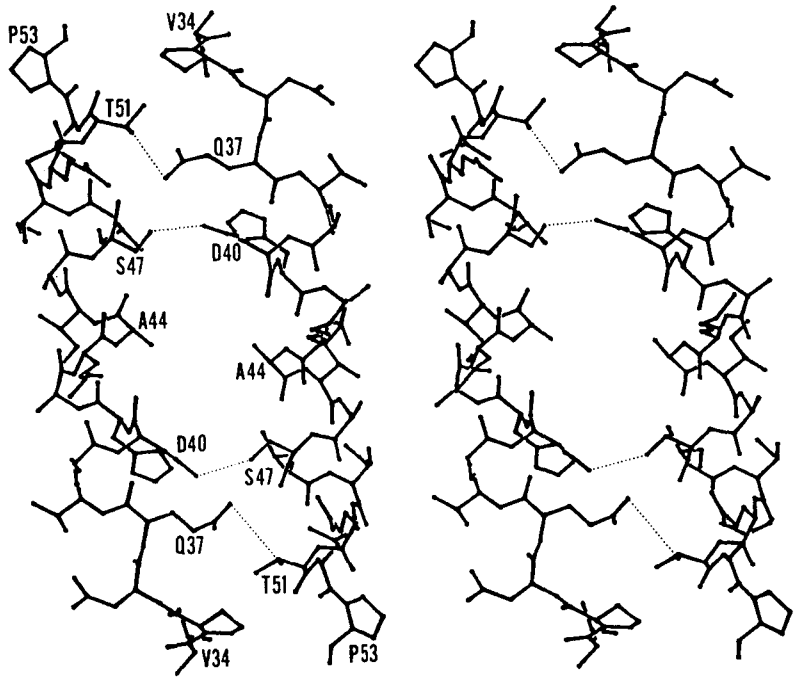
(a)

(b)

Fig. 5-5 The tube model of (left) RCP and (right) MCP dimers. The views in (a) is parallel to the two-fold axis that relates the individual monomer. (b) is the bottom views of (a).



(a)



(b)

Fig. 5-6 Stereoscopic representations of the helix-helix interactions at the subunit interface of the RCP dimer. (a) The interaction between A and A'-helices. The helix axes are closest near the α – carbon position of Ala16 in each helix. (b) The interaction between B and B'-helices. Alanine residues at position 44 in each helix are closest in the two helix axes. The hydrogen bonds, Asn19(A)-Ser15(A'), Ser15(A)-Asn19(A'), Tyr20(A)-Tyr20(A'), Gln37(B)-Thr51(B'), Thr51(B)-Gln37(B'), Asp40(B)-Ser47(B'), Ser47(B)-Asp40(B') are indicated by dotted lines.

5-2-4 Comparison of monomer structures

The structurally equivalent residues were searched using EXPLORE program (Rossmann and Argos 1976) for the combination of each monomer. In this program, a search function, plotted in terms of three Eulerian angles, represents the number of sequentially equivalenced amino acids. For each orientation one protein structure is rotated about its center of mass with respect to the other and probabilities are calculated which estimate the degree of structural parallelism. The structurally equivalent residues with highest probabilities are then selected. Each calculation was performed using the coordinates of Cα atoms for all amino acids. After searching of equivalent residues, each coordinates set was superposed to minimize the distance of structurally equivalent Cα atoms using least-squares technique. The results are summarized in Table 5-4. The mean difference in Cα and heme atom positions was 2.3Å. The mean difference in the 4 helix regions (2.1Å) is apparently smaller than that in the 3 loop regions (3.9Å). The schematic views of superposed monomers are shown in Fig. 5-7.

Fig. 5-7 The ribbon model of RCP monomer (white) on which the MCP monomer (green) is superimposed. The mean distances between equivalent Cα atoms is 2.3Å.

Table 5-4 Structural comparison of monomers for RCP and MCP

pair	number of residues	MBC	mean distance(Å)
RMOL1-MMOL1	121	1.16	2.3
helix A	27		2.0
helix B	13		1.0
helix C	19		2.3
helix D	21		2.7
AB loop	10		6.1
BC loop	26		2.8
CD loop	5		5.3
Heme group			1.2

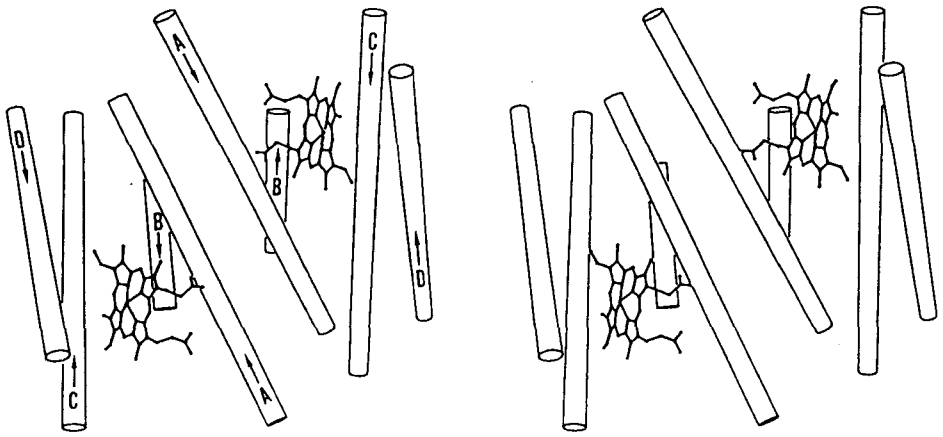
5-2-5 Monomer arrangement in dimer molecule

As mentioned in the above discussions, the monomer molecules of RCP and MCP had very similar foldings, although the inter-monomer interactions were quite different. This differences reflect on the whole shapes of dimer molecules. To study the differences of the dimer structures, the axes of the helices were calculated and compared. The axis of the helix was defined as the least-squares line calculated from Cα atoms that construct the helix. This least-squares line were calculated using the method that were reported by V. Schomaker *et al.* (Schomaker *et al.*, 1959). The inter-helical angles and distances were then calculated, showed in Table 5-5. These values of the intra-monomer, namely A-B, ..., C-D, showed resemblance of RCP and MCP. The differences of the inter-helical angles were only 1.1 to 8.3° and 0.3$\sim$2.5Å. These facts were compatible with the previous discussions of the similarities of the monomers. The inter axis values, however, of the inter-monomer, namely A-A', ..., D-D', had different tendency, especially for angles. The differences of inter-axis angles were larger than that of intra-monomer, 9.1 to 30.3°. This differences reflect on the inter-monomer angles that were calculated from the monomer axis defined as least-squares line of four α – helices. The inter-monomer angles for RCP and MCP were 134.2 and 105.5°, respectively (Fig. 5-8).

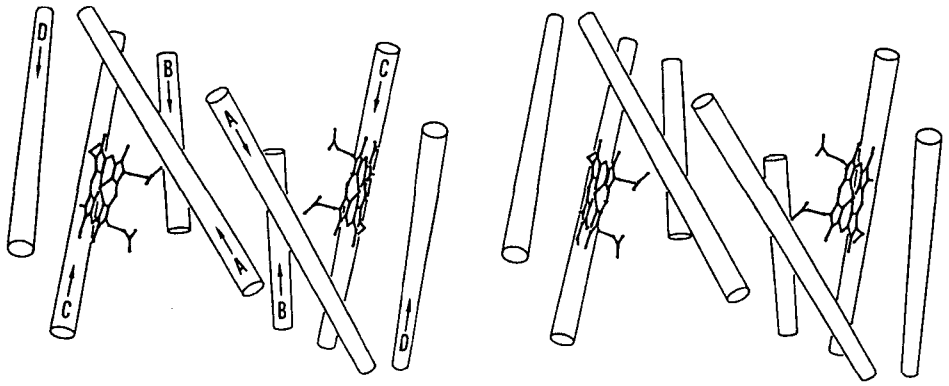
Table 5-5 Inter-helical angles and distances in RCP and MCP dimers

Helix pair	RCP		MCP	
	angle(°) [a]	distance(Å) [b]	angle(°) [a]	distance(Å) [b]
[Intra-subunit] [c]				
A-B	155.6	10.0	152.8	9.7
A-C	28.4	12.2	36.1	12.7
A-D	154.8	10.3	147.1	8.3
B-C	170.5	8.2	163.2	7.6
B-D	20.6	13.9	25.6	11.4
C-D	166.5	8.2	167.6	7.9
[Inter-subunit]				
A-A'	157.6	7.9	127.3	6.4
B-B'	160.4	11.0	151.3	9.6
C-C'	143.3	27.8	127.5	25.4
D-D'	122.0	36.2	103.6	38.1
m1-m2 [d]	134.2	21.1	105.5	18.3

[a] Angles between two helix axes.
[b] The nearest distances between two helix axes.
[c] Average values between two monomers for intra-subunit (A-B to C-D) are given here.
[d] Angles and distances between two monomer axes defined as least-squares lines by the Cα atoms of 4 α − helices in each monomer.



(a)



(b)

Fig. 5-8 Stereoscopic drawings of helix axis in the RCP dimer (a), and the MCP dimer (b) viewed from the direction of the non-crystallographic 2-fold symmetry axis. Arrows indicate the N to C direction of a polypeptide chain.

5-3 Conclusion

The structural features of RCP have shown to have high similarities in monomer folding in spite of large divergency in the primary structure against MCP. The changes of amino acids in the regions of A and B helices were considerably larger than that in C and D helices. These changes have caused the differences in the inter-monomer interactions and in whole shapes of dimer molecules between RCP and MCP. Thus the resemblance of monomer structure can be related with functional requirement.

The four-α – helical bundle was reported (Weber and Salemme, 1980; Sheridan, *et al.*, 1982) that it is a stable super secondary structure to have a binding pocket for the prosthetic group, and tend to be a dimeric or polymeric quaternary structure for more stabilization. In the case of cytochromes *c'*, it is assumed from structurally features that the functionally significant regions are C and D helices in which the similarity of primary and tertialy structure is higher than other regions. A and B helices might exist for stabilization of monomer structure and heme prosthetic group.

References

Ambler, R.P., Bartsch, R.G., Daniel, M., Kamen, M.D., McLellan, L., Meyer, T.E. and Van Beeumen, J. (1981) *Proc. Natl. Acad. Sci.* USA **78**, 6854-6857

Bartsch, R.G. (1978) In *The Photosynthetic Bacteria* (Clayton, R.K. and Sistrom, W.R., eds), pp. 249-279, Plenum Press, New York

Blow, D.M. and Crick, F.H.C. (1959) *Acta Cryst.* **12**, 794-802

Blundell, T.L. and Johnson, L.N. (1976a) In *Protein Crystallography*, pp.151-182, Academic Press, New York

Blundell, T.L. and Johnson, L.N. (1976b) In *Protein Crystallography*, pp.337-362, Academic Press, New York

Bricogne, G. (1976) *Acta Cryst.* **A32**, 832-847

Brünger, A.T., Kuriyan,J. & Karplus, M. (1987) *Science* **235**, 458-460

Cromer, D.T. and Liberman, D. (1970) *J. Chem. Phys.* **53**, 1891-1898

Cromer, D.T. and Liberman, D. (1981) *Acta Cryst.* **A37**, 267-268

Ehrenberg, A. and Kamen, M.D. (1965) *Biochim. Biophys. Acta* **102**, 333-340

Finzel, B.C., Weber, P.C., Hardman, K.D. and Salemme, F.R. (1985) *J. Mol. Biol.* **186**, 627-643

Gibson, Q.H. and Kamen, M.D. (1966) *J. Biol. Chem.* **241**, 1969-1976

Hendrickson, W.A. and Konnert, J.H. (1981) In *Biomolecular Structure, Function, Conformation and Evolution* (Srinivasan, R., ed.), vol.1, pp43-57, Pergamon Press, Oxford

Herzenberg, A. and Lau, H.S.M. (1967) *Acta Cryst.* **22**, 24-28

Hoppe, W and Jakubowski, U. (1975) In *Anomalous Scattering* (Ramaseshan, S. and Abrahams, S.C. eds), pp. 437-461, Munksgaard, Copenhagen

Horio, T. and Kamen, M.D. (1961) *Biochim. Biophys. Acta* 48, 266-286

Huke, K. and Yamakawa, T. (1980) *Nucl. Instrum. Methods* 177, 253-257

Iga, Y. and Yasuoka, N. (1984) *J. Mol. Graphics* 3, 79-82

Kahn, R., Fourme, R., Bosshard, R., Chiadmi, M., Risler, J.L., Dideberg, O. and Wery, J.P. (1985) *FEBS Lett.* 179, 133-137

Kendrew, J.C., Dickerson, R.E., Strandberg, B.E., Hart, R.G., Davies, D.R., Phillips, D.C. and Shore, V.C. (1960) *Nature* (*London*) 185, 422-427

Kusunoki, M. (1980) In doctoral thesis, Faculty of Science, Osaka Univ., Japan

Langridge, R. (1961) *Biochim. Biophys. Acta* 54, 585-586

Luzzati, V. (1952) *Acta Cryst.* 5, 802-810

Maltempo, M.M., Moss, T.H. and Cusanovich, M.A. (1974) *Biochim. Biophys. Acta* 342, 290-305

Matthews, B.W. (1968) *J. Mol. Biol.* 33, 491-497

Meyer, T.E., Ambler, R.P., Bartsch, R.G. and Kamen, M.D. (1975) *J. Biol. Chem.* 250, 8416-8421

Mitchell, C.M. (1957) *Acta Cryst.* 10, 475-476

North, A.C.T., Phillips, D.C. and Mathews, F.S. (1968) *Acta Cryst.*

A24, 351-359

Pepinsky, R. and Okaya, Y. (1956) *Proc. Natl. Acad. Sci. USA* 42, 286-292

Phillips, J.C. and Hodgson, K.O. (1980) *Acta Cryst.* A36, 856-864

Phillips, J.C., Wlodawer, A., Goodfellow, J.M., Watenpaugh, K.D., Sieker, L.C., Jensen, L.H. and Hodgson, K.O. (1977) *Acta Cryst.* A33, 445-455

Rossmann, M.G. and Argos P. (1976) *J. Mol. Biol.*, 105, 75-95

Salemme, F.R. (1974) *Arch. Biochem. Biophys.* 163, 423-425

Sasaki, S. (1984) *Anomalous Scattering Factor for Synchrotron Users, Calculated Using Cromer and Liberman's Method*, p. 16. National Laboratory for High Energy Physics, Ibaraki, Japan

Satow, Y. (1984) In *Methods and Applications in Crystallographic Computing* (Hall, S.R. and Ashida, T., eds), pp. 56-64, Oxford Univ. Press

Schomaker, V., Waser, J., Marsh, R.E. and Bergman, G. (1959) *Acta Cryst.*, 12, 600-604

Sheridan, R.P., Levy, R.M. and Salemme, F.R. (1982) *Proc. Natl. Acad. Sci. USA* 79, 4545-4549

Watenpaugh, K.D. (1985) In *Methods in Enzymology* (Wyckoff, H.W., Hirs, C.H.W. and Timasheff, S.N., eds), vol. 115, pp. 3-15, Academic Press, London

Weber, P.C. and Salemme, F.R. (1977) *J. Mol. Biol.* 117, 815-820

Weber, P.C. and Salemme, F.R. (1980) *Nature* (*London*) 287, 82-84

Weber, P.C., Howard, A., Xuong, Ng.H. and Salemme, F.R. (1981) *J. Mol. Biol.* 153, 399-424

Zeppezauer, M., Eklund, H. and Zeppezauer, E.S. (1968) *Arch. Biochem. Biophys.* 126, 564-573