Acta Cryst. (1994). D**50**, 271–275

Application of an Automatic Molecular-Replacement Procedure to Crystal Structure Analysis of Cytochrome c_2 from *Rhodopseudomonas viridis*

By Kunio Miki* and Satoshi Sogabe

Research Laboratory of Resources Utilization, Tokyo Institute of Technology, Nagatsuta, Midori-ku, Yokohama 227, Japan

Atsushi Uno, Toshihide Ezoe, Nobutami Kasai and Masahiko Saeda

Department of Applied Chemistry, Faculty of Engineering, Osaka University, Yamadaoka, Suita, Osaka 565, Japan

Yoshiki Matsuura

Institute for Protein Research, Osaka University, Yamadaoka, Suita, Osaka 565, Japan

and Miyako Miki

Institute of Scientific and Industrial Research, Osaka University, Mihogaoka, Ibaraki, Osaka 567, Japan

(*Received* 2 *July* 1993; *accepted* 13 *November* 1993)

Abstract

An automatic molecular-replacement procedure has been applied to solve the crystal structure of cytochrome c_2 from *Rhodopseudomonas viridis*. The structure was solved on the basis of the structure of tuna cytochrome *c* as a search model using an automatic processing program system, *AUTOMR*. The refinements by molecular dynamics and restrained least-squares methods result in a current crystallographic *R* factor of 0.219 for diffraction data at 3 Å resolution.

Introduction

The number of proteins with three-dimensional structures that have been established by X-ray crystallography and NMR spectroscopy has rapidly increased in recent years. The folding patterns of proteins may be classified into a limited number of categories, which indicates that there are significant structural homologies among not only functionally related proteins but also those having different biological functions. This suggests that it is possible that similar protein structures have been solved already when we intend to solve the structures of new proteins. Therefore, the molecular-replacement method (Rossmann, 1972) has become quite important in the course of protein crystallography. However, this method does not usually proceed straightforwardly to obtain the final molecular model in the crystal. For example, the solutions of a rotation search that are of high rank do not always give a correct solution in the subsequent translation search. Sometimes we have to make many attempts before reaching the correct solution of the crystal structure. In such cases, automatic computerization of each process of the molecular-replacement procedure is very useful to save investigators' labor in computational steps. One of us has developed an automatic processing program system, *AUTOMR* (Matsuura, 1991), and we have succeeded in solving the new crystal structure of a bacterial cytochrome c_2 by the use of this program system.

In the light-driven cyclic electron-transfer process in the photosynthetic purple bacterium, *Rhodopseudomonas viridis*, cytochrome c_2 reduced by the cytochrome bc_1 complex transfers electrons to the bound cytochrome subunit of the photosynthetic reaction center to re-reduce the photo-oxidized bacteriochlorophyll dimer, which is the primary electron donor in the reaction center (Michel & Deisenhofer, 1986). The structure of the reaction center from *R. viridis* is already available (Deisenhofer, Epp, Miki, Huber & Michel, 1984, 1985; Michel, Epp & Deisenhofer, 1986). A study of the binding between cytochrome c_2 and the reaction center based on both three-dimensional structures will be possible if the precise three-dimensional structure of cytochrome c_2 is known. Here we report the course of the structure determination of cytochrome c_2 from *R. viridis* at 3 Å resolution using the automatic molecular-replacement procedure.

* Author to whom all correspondence should be addressed.

©1994 International Union of Crystallography
Printed in Great Britain – all rights reserved

ISSN 0907-4449 ©1994

Experimental

Crystallization and data collection

Cytochrome c_2 from *R. viridis* was crystallized as described previously (Miki, Saeda, Masaki, Kasai, Miki & Hayashi, 1986). The crystals belong to the trigonal space group $P3_121$ or $P3_221$ with unit-cell dimensions of $a = b = 76.13$ and $c = 40.40$ Å. The crystals contain one protein molecule per asymmetric unit ($V_m = 2.70$ Å^3 Da^{-1}) (Matthews, 1966).

The X-ray diffraction data were collected at room temperature on a Rigaku automated four-circle diffractometer using one native crystal. The X-ray source was Ni-filtered Cu $K\alpha$ radiation from a Rigaku RU300 rotating-anode generator operated at 40 kV and 200 mA. The ω-scan mode was employed with a scanning speed of 4° min^{-1} and a scan width of 0.85°. The background intensities were measured for 4 s at both ends of each scan. Three standard reflections (060, 003, 440) were measured after every 53 reflections to monitor the crystal stability and orientation. A total of 8667 reflections were collected up to 3 Å resolution with index ranges of $h = -25$ to 25, $k = 0$ to 25, $l = 0$ to 13. The radiation damage was less than 7% calculated from the F values of the standard reflections. The intensity data were corrected for absorption (North, Phillips & Mathews, 1968) and for radiation damage (based on the standard reflections) by least-squares methods. Finally, 2870 independent reflections were obtained after averaging symmetry-equivalent reflections with an R_{sym} value of 0.073. This native data set contains 88% of the theoretical number of reflections up to 3 Å resolution.

Molecular replacement

The molecular structures of bacterial cytochromes c_2 from *Rhodospirillum rubrum* and from *Rhodobacter capsulatus* have been solved and refined already (Salemme, Freer, Xuong, Alden & Kraut, 1973; Bhatia, 1981; Benning *et al.*, 1991). For a variety of bacterial cytochromes, amino-acid sequences have been determined (Amber *et al.*, 1979; Grisshammer, Wiessner & Michel, 1990). Cytochrome c_2 from *R. viridis* is much closer to the mitochondrial cytochromes c than the other bacterial cytochromes c_2 in terms of the length of the polypeptide chain and the homology of amino-acid sequences. Comparing the amino-acid sequence of tuna cytochrome c with that of *R. viridis* cytochrome c_2 (Grisshammer, Wiessner & Michel, 1990), the latter has one residue deletion (between 11 and 12) and five additional residues in the C-terminal end. Among the 102 aligned residues, 50 residues are identical between the two sequences. On account of their similarity, the refined structure of cytochrome c from tuna (Takano & Dickerson, 1981*a*,*b*) was employed for the search model of the molecular-replacement method.

The automatic processing program system of molecular-replacement procedure, *AUTOMR* (Matsuura, 1991), consists of four parts: structure-factor calculation of the model molecule, fast-rotation function, multiple translation function and rigid-body refinement. In this system, every possible peak obtained by the rotation search is successively subjected to the translation search followed by the rigid-body refinement of the rotational and translational parameters. Both the structures of oxidized and reduced forms of tuna cytochrome c (Takano & Dickerson, 1981*a*,*b*) were used as the initial search model, the atomic coordinates of which were taken from the entries (3CYT and 5CYT, respectively) of the Brookhaven Protein Data Bank (Bernstein *et al.*, 1977).

The model molecule, in which all the atoms of tuna cytochrome c together with those of the side chains were included, was placed in a $P1$ cell with $a = b = c = 100$ Å and $\alpha = \beta = \gamma = 90°$. Triclinic structure factors were calculated between 10 and 6 Å resolution. A cross rotation function (Crowther, 1972; Tanaka, 1977) was calculated as a function of the orientation angles ψ, φ, χ in the angular range, $\psi = 0$–90, $\varphi = 0$–360, $\chi = 0$–360° with 5° steps in each direction using the data between 10 and 6 Å resolution. The radius of the cut-off sphere in the Patterson function was set to 20 Å. A translation search was carried out for the 20 highest peaks in the rotation search. The agreement of observed and calculated structure factors was evaluated with a correlation coefficient between F_o and F_c at every position translated by about a 1 Å step. The molecular packing in the unit cell was checked and those cases giving packing with short intermolecular contacts were rejected. The model position was then improved by the rigid-body refinement program *CORELS* (Sussman, Holbrook, Church & Kim, 1977) using the data between 10 and 6 Å resolution. The computations were performed on an ACOS 930 computer at the Research Center for Protein Engineering, Institute for Protein Research, Osaka University.

Refinement

Before starting refinement, the tuna cytochrome c model molecule used in the molecular replacement was modified so as to fit the sequence of the *R. viridis* cytochrome c_2. Refinement was carried out by both the simulated-annealing method and the restrained-parameter least-squares procedure. The former was performed by the program *X-PLOR* (Brünger, Kuriyan & Karplus, 1987; Brünger, Karplus & Petsko, 1989). After 120 steps of preliminary energy

minimization, 6000 steps (timesteps of 0.25 fs) of dynamics were carried out at 6000 K, followed by decreasing the temperature from 6000 to 300 K in steps of 25 K. The tolerance by which any atomic coordinates could deviate from position was specified as 0.2. Another 80 steps of energy minimization and individual temperature-factor refinement completed the process. The latter method was performed with the program *PROLSQ* (Hendrickson & Konnert, 1980) using the reflections from 10 to 3 Å resolution. During the final stages of this refinement, temperature factors were also refined individually for side and main chains and heme atoms.

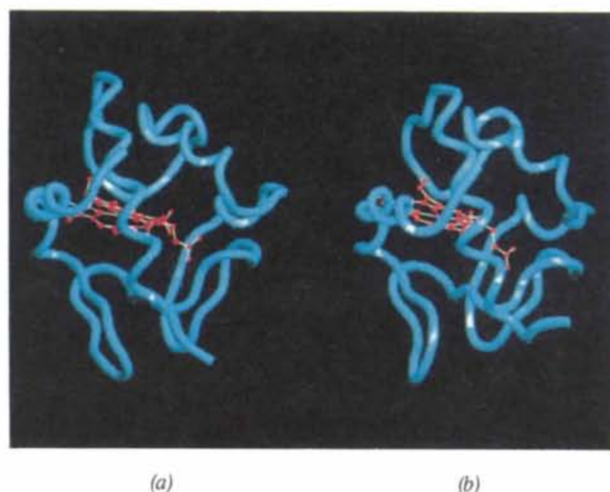
The model modification was performed on an Evans and Sutherland PS390 color graphics system supported by a version of *FRODO* (Jones, 1985) in order to display molecular fragments and electron density and manipulate these interactively. In this stage, the refined structure was examined and manual map analysis and adjustment were undertaken by systematic inspection of electron-density maps calculated with omit maps for the whole region of the polypeptide chain. Omit maps were generated with phases determined after several cycles of refinement whereby relevant regions of the molecules, less than 10% of the total, are omitted from the structure during phasing to avoid possible feedback. Moreover, the model refitting was performed using electron-density maps calculated with both $2F_o - F_c$ and $F_o - F_c$ coefficients. The computations were performed on VAX 3300 and VAXstation 4000 (model 90) computers.

The atomic coordinates have been deposited with the Protein Data Bank.*

Results and discussion

The calculation of automatic molecular replacement was initially performed for both enantiomorphic space groups $P3_121$ and $P3_221$ using the atomic coordinates of the oxidized form of tuna cytochrome c_2. In space group $P3_121$, the values of the correlation coefficient after the translation search lie between 34 and 47, in which no clear high values were observed, and also the R values after rigid-body refinement range from 0.51 to 0.56, in which no clear minimum R values were obtained. On the other hand, in the case of $P3_221$ (Table 1*a*), the highest value of the correlation coefficient (53) gave the minimum R value (0.46), which comes from the second highest peak of the rotation function. This result suggests that the correct space group is $P3_221$. Afterwards, this was confirmed by a calculation using the atomic coordinates of the reduced form as a search model in space group $P3_221$ (Table 1*b*). The result indicated the highest value of the correlation coefficient (53) and the minimum R value (0.46). It should be pointed out, however, that this orientation comes from the ninth highest peak in the rotation function. The results obtained from the two slightly different initial models (oxidized and reduced forms) in space group $P3_221$ both gave the same position and orientation of the model molecule in the crystal lattice. The crystal packing obtained indicates that symmetry-related molecules are packed with reasonable intermolecular contacts.

Although the refinement was first performed by the restrained least-squares method, the R factor did not drop below 0.40 and the geometric parameters of the model become less accurate after refinement. We then applied the simulated-annealing (SA) refinement, which successfully converged the crystallographic R factor from 0.47 to 0.23 for all observed reflection data between 10 and 3 Å resolution. Subsequent refinements were performed by the restrained-parameter least-squares method after checking discrepancies between the model and the electron density and modifying the model, which gave the current R factor of 0.219. The present refined structural model has reasonable stereochemistry. The root-mean-square deviations of the model from ideal geometry are: 0.012 Å for the covalent bond distances, 0.040 Å for the interbond angle distances, 0.047 Å for the planar groups and 1.9° for the peptide bond torsional angles.



(*a*) (*b*)

Fig. 1. Smooth line representation of the polypeptide chain (blue) with the heme group (red) of cytochrome c_2 from *R. viridis* before and after simulated-annealing refinement. (*a*) The model before refinement, (*b*) the model after refinement.

* Atomic coordinates have been deposited with the Protein Data Bank, Brookhaven National Laboratory. Free copies may be obtained through The Technical Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Supplementary Publication No. SUP 37110). A list of deposited data is given at the end of this issue.

Table 1. *Results of the calculation of the automatic molecular-replacement procedure*

	Rotation search*				Translation search*		
	Rotation angle (°)			Peak	Correlation	Intermolecular	Rigid-body refinement*
No.	φ	ψ	χ	height	coefficient†	distance (Å)‡	*R* value§
(*a*) Model: oxidized form, space group $P3_221$							
1	23.8	175.5	175.4	165	37 (17)	—	
2	36.4	224.8	10.2	159	53 (1)	26.5	0.46
3	85.7	104.9	296.1	158	47 (2)	24.2	0.57
4	87.6	85.9	130.4	155	41 (9)	32.3	0.57
5	49.9	70.0	170.3	149	41 (10)	—	
6	36.6	244.4	130.7	137	44 (5)	25.3	0.58
7	85.8	124.9	295.9	136	42 (7)	—	
8	53.3	14.7	244.8	135	41 (11)	—	
9	78.6	254.3	270.6	134	38 (16)	—	
10	88.1	344.5	94.6	131	40 (13)	—	
11	49.8	97.9	140.5	129	36 (18)	—	
12	43.8	150.4	306.4	129	41 (12)	—	
13	31.6	273.7	79.7	128	45 (4)	—	
14	40.7	153.1	298.8	128	40 (14)	—	
15	87.2	320.7	268.8	121	33 (20)	29.8	0.56
16	85.5	324.7	299.9	120	41 (8)	23.1	0.57
17	71.0	240.9	90.5	117	39 (15)	—	
18	37.5	226.1	110.3	117	34 (19)	—	
19	39.4	249.8	120.5	117	46 (3)	—	
20	14.2	60.2	68.6	116	44 (6)	—	
(*b*) Model: reduced form, space group $P3_221$							
1	75.0	55.1	338.8	144	36 (20)	—	
2	12.4	115.1	243.8	140	32 (13)	—	
3	39.3	115.6	78.8	139	38 (19)	—	
4	90.0	235.0	20.0	139	40 (18)	24.4	0.52
5	88.4	56.0	160.6	138	44 (10)	—	
6	44.1	355.5	288.9	138	42 (14)	—	
7	39.7	40.0	235.1	133	41 (16)	27.6	0.56
8	87.3	330.5	130.7	130	43 (11)	—	
9	12.7	65.2	54.8	122	53 (1)	26.8	0.46
10	12.4	305.0	291.6	120	47 (5)	—	
11	10.2	70.0	169.7	119	47 (3)	—	
12	8.4	360.1	174.7	117	42 (15)	34.9	0.55
13	15.2	360.2	288.6	116	50 (2)	26.9	0.56
14	87.6	345.1	290.6	116	45 (7)	34.3	0.59
15	10.0	259.9	110.4	113	45 (8)	—	
16	55.2	140.1	342.0	113	40 (17)	40.1	0.56
17	90.0	160.0	35.0	112	47 (4)	39.8	0.51
18	90.0	340.0	325.0	112	47 (6)	39.8	0.51
19	28.1	325.1	280.6	111	45 (9)	27.9	0.60
20	20.9	41.0	236.2	111	43 (12)	—	

* The calculations of rotation search, translation search and rigid-body refinement were performed using all the reflections of 10–6, 20–10 and 10–6 Å resolution, respectively.

† Correlation coefficient = $(n\sum F_{\rm obs}F_{\rm calc} - \sum F_{\rm obs}F_{\rm calc})/\{[n\sum F^2_{\rm obs} - (\sum F_{\rm obs})^2][n\sum F^2_{\rm calc} - (\sum F^2_{\rm calc})^2]\}^{1/2}$, where n is the number of reflections. The ranking of the correlation coefficient is in parentheses.

‡ The distance between molecular centers of symmetry-related model molecules. Unreasonable contacts between molecules are observed in the columns where no values are shown.

§ The solutions having unreasonable short molecular contacts were not applied to the refinement.

The three-dimensional folding of the cytochrome c_2 polypeptide chain before and after *X-PLOR* refinement is shown in Fig. 1. The average root-mean-square deviation of C^α positions before and after the refinement was 1.5 Å. In several regions (for example, N-terminal, residues 20–22, residues 42–60 and C-terminal), the deviations are larger than the average value, the maximum deviation being 4.4 Å. This may be a reason why the initial refinement by the restrained least-squares method was unsuccessful.

It can be suggested from the present results that this automatic molecular-replacement system is useful for obtaining a reasonable solution without any stepwise handling by investigators. This program could significantly determine the correct space group. This is the first case where the *AUTOMR* program solved an unknown crystal structure. As shown in Table 1(*b*), the relatively low-ranking peak for the rotation function gave the correct solution. In such a case, the automatic procedure is effective to avoid missing the correct solution which appears at a low rank in the rotation search.

A higher resolution analysis is presently in progress. It will be possible to make a more detailed comparison between the present cytochrome c_2 and other cytochromes.

The authors thank Dr Isao Tanaka, Hokkaido University, for his kindness in giving us an opportunity to use his computer (VAX 3300) with much help and many useful suggestions. This work is partly supported by Grants-in-Aid for Scientific Research on Priority Areas (Nos. 03241103 and 04225103) from the Ministry of Education, Science and Culture of Japan to KM and also by a Kurata Research Grant from the Kurata Foundation to KM.

References

AMBER, R. P., DANIEL, M., HERMOSO, J., MEYER, T. E., BARTSCH, R. G. & KAMEN, M. D. (1979). *Nature (London)*, **278**, 659–660.

BENNING, M. M., WESENBERG, G., CAFFREY, M. S., BARTSCH, R. G., MEYER, T. E., CUSANOVICH, M. A., RAYMENT, I. & HOLDEN, H. M. (1991). *J. Mol. Biol.* **220**, 673–685.

BERNSTEIN, F. C., KOETZLE, T. F., WILLIAMS, G. J. B., MEYER, E. F., BRICE, M. D., RODGERS, J. R., KENNARD, O., SHIMANOUCHI, T. & TASUMI, M. (1977). *J. Mol. Biol.* **112**, 535–542.

BHATIA, G. E. (1981). PhD thesis, Univ. of California, San Diego, USA.

BRÜNGER, A. T., KARPLUS, M. & PETSKO, G. (1989). *Acta Cryst.* A**45**, 50–61.

BRÜNGER, A. T., KURIYAN, J. & KARPLUS, M. (1987). *Science*, **235**, 458–460.

CROWTHER, R. A. (1972). *The Molecular Replacement Method*, edited by M. G. ROSSMANN, pp. 173–183. New York: Gordon & Breach.

DEISENHOFER, J., EPP, O., MIKI, K., HUBER, R. & MICHEL, H. (1984). *J. Mol. Biol.* **180**, 385–398.

DEISENHOFER, J., EPP, O., MIKI, K., HUBER, R. & MICHEL, H. (1985). *Nature (London)*, **318**, 618–624.

GRISSHAMMER, R., WIESSNER, C. & MICHEL, H. (1990). *J. Bacteriol.* **172**, 5071–5078.

HENDRICKSON, W. A. & KONNERT, J. H. (1980). *Biomolecular Structure, Conformation, Function and Evolution*, Vol. 1, edited by R. SRINIVASAN, pp. 43–57. Oxford: Pergamon Press.

JONES, T. A. (1985). *Methods in Enzymology*, Vol. 115, edited by H. W. WYCKOFF, C. H. W. HIRS & S. N. TIMASHEF, pp. 157–171. New York: Academic Press.

MATSUURA, Y. (1991). *J. Appl. Cryst.* **24**, 1063–1066.

MATTHEWS, B. W. (1966). *J. Mol. Biol.* **33**, 491–497.

MICHEL, H. & DEISENHOFER, J. (1986). *Encyclopedia of Plant Physiology; Photosynthesis* III, Vol. 19, edited by A. C. STAEHELIN & C. J. ARNTZEN, pp. 371–381. Berlin: Springer.

MICHEL, H., EPP, O. & DEISENHOFER, J. (1986). *EMBO J.* **5**, 2445–2451.

MIKI, K., SAEDA, M., MASAKI, K., KASAI, N., MIKI, M. & HAYASHI, K. (1986). *J. Mol. Biol.* **191**, 579–580.

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

ROSSMANN, M. G. (1972). *The Molecular Replacement Method*, New York: Gordon & Breach.

SALEMME, F. R., FREER, S. T., XUONG, N. H., ALDEN, R. A. & KRAUT, J. (1973). *J. Biol. Chem.* **248**, 3910–3921.

SUSSMAN, J. L., HOLBROOK, S. R., CHURCH, G. M. & KIM, S. H. (1977). *Acta Cryst.* A**33**, 800–804.

TAKANO, T. & DICKERSON, R. E. (1981*a*). *J. Mol. Biol.* **153**, 79–94.

TAKANO, T. & DICKERSON, R. E. (1981*b*). *J. Mol. Biol.* **153**, 95–115.

TANAKA, N. (1977). *Acta Cryst.* A**33**, 191–193.