References

BLOW, D. M. & CRICK, F. H. C. (1959). *Acta Cryst.* **12**, 794-802.
BLUNDELL, T. L. & JOHNSON, L. N. (1976). *Protein Crystallography*, p. 177. London: Academic Press.
COCHRAN, W. (1955). *Acta Cryst.* **8**, 473-478.
FAN HAI-FU & GU YUAN-XIN (1982). *Acta Phys. Sin.* **31**, 969-971.
FAN HAI-FU, HAN FU-SON & QIAN JIN-ZI (1984). *Acta Cryst.* A**40**, 495-498.
FAN HAI-FU & QIAN JIN-ZI (1981). *Acta Phys. Sin.* **30**, 594-601.
FAN HAI-FU & ZHENG QI-TAI (1978). *Acta Phys. Sin.* **27**, 169-174.
FAN HAI-FU & ZHENG QI-TAI (1981). *Acta Cryst.* A**37**, C329.
HENDRICKSON, W. A. & TEETER, M. M. (1981). *Nature (London)*, **290**, 107-113.
SIM, G. A. (1959). *Acta Cryst.* **12**, 813-815.

Acta Cryst. (1985). A**41**, 284-285

Combining Direct Methods with Isomorphous Replacement or Anomalous Scattering Data. IV. Test in the SIR Case with the Replacing Atoms in the Centrosymmetric Arrangement

BY YAO JIA-XING AND FAN HAI-FU

Institute of Physics, Academy of Sciences, Beijing, China

(*Received* 10 *September* 1984; *accepted* 10 *December* 1984)

Abstract

The method described in the preceding papers has been applied to the single isomorphous replacement (SIR) case with the replacing atoms in a centrosymmetric arrangement. Two kinds of phase ambiguities simultaneously occurred in this example. One is inherent in the SIR method and was resolved by calculating the probabilities $P_+(\Delta\varphi_{\mathbf{H}})$. The other comes from the special arrangement of the replacing atoms and was treated by a multi-solution procedure with random starting sign sets. A new figure of merit was used to predict the quality of the solutions. The method has been verified using a set of error-free data from the model structures of a protein and its heavy-atom derivative.

Introduction

Many attempts have been made since Blow & Rossmann (1961) to resolve the phase ambiguity of the SIR method in the determination of protein structures. The ambiguity can in principle be resolved in either the real or the reciprocal space. Up to now the real-space methods have been more successful in practice. With the so-called ISIR method, a dozen unknown protein structures have been solved (Wang Bi-Cheng, 1981, 1984). However, in spite of its high phasing power, this method is subject to the limitation that it will not be applicable if the replacing atoms are in a centrosymmetric arrangement. This paper describes the application of a reciprocal-space method to treat this problem.

Method

In the SIR case, each reflection not belonging to a centric zone has two equally possible phases, *i.e.*

$$\varphi_{\mathbf{H}} = \varphi_{\mathbf{H},R} \pm |\Delta\varphi_{\mathbf{H}}|,$$

where $\varphi_{\mathbf{H}}$ denotes the phase of the structure factor $\mathbf{F}_{\mathbf{H}}$, $\varphi_{\mathbf{H},R}$ is the phase contribution from the replacing atoms and $\Delta\varphi_{\mathbf{H}}$ is the difference between $\varphi_{\mathbf{H}}$ and $\varphi_{\mathbf{H},R}$. According to our preceding papers (Fan Hai-fu, Han Fu-son, Qian Jin-zi & Yao Jia-xing, 1984; Fan Hai-fu, Han Fu-son & Qian Jin-zi, 1984; Fan Hai-fu & Gu Yuan-xin, 1985; hereafter referred to as papers I, II and III respectively), this phase ambiguity can be resolved by calculating the probability for $\Delta\varphi_{\mathbf{H}}$ to have a positive sign:

$$P_+(\Delta\varphi_{\mathbf{H}}) = \tfrac{1}{2} + \tfrac{1}{2}\tanh\{\sin|\Delta\varphi_{\mathbf{H}}| \times \left[\sum_{\mathbf{H}'} m_{\mathbf{H}'} m_{\mathbf{H}-\mathbf{H}'} K_{\mathbf{H}\mathbf{H}'} \sin(\Phi'_3 + \Delta\varphi_{\mathbf{H}'\text{best}} + \Delta\varphi_{\mathbf{H}-\mathbf{H}'\text{best}})\right]\}, \quad (1)$$

where

$$m_{\mathbf{H}} = \exp(-\sigma_{\mathbf{H}}^2/2)\{[2(P_+ - \tfrac{1}{2})^2 + \tfrac{1}{2}] \times (1 - \cos 2\Delta\varphi_{\mathbf{H}}) + \cos 2\Delta\varphi_{\mathbf{H}}\}^{1/2} \quad (2)$$

and

$$\tan(\Delta\varphi_{\mathbf{H}\text{best}}) = 2(P_+ - \tfrac{1}{2})\sin|\Delta\varphi_{\mathbf{H}}|/\cos\Delta\varphi_{\mathbf{H}}. \quad (3)$$

0108-7673/85/030284-02$01.50

Table 1. *Results from ten random starting sets after three cycles of iteration*

The reflections were arranged in descending order of $|P_+ - \frac{1}{2}|$ and then cumulated into four groups. The groups numbered 1, 2, 3 and 4 contain the top 200, 400, 600 and 800 reflections respectively.
% Percentage of reflections with the signs of $\Delta\varphi_{\mathbf{H}}$ correctly determined.
ER Average error of phases (in degrees).

FOM	0·4136		0·4091		0·3995		0·3918		0·3848		0·3690		0·3575		0·3518		0·3446		0·3362	
Set	1		8		9		2		3		10		5		4		7		6	
Group	%	ER	%	ER	%	ER	%	ER	%	ER	%	ER	%	ER	%	ER	%	ER	%	ER
1	92·0	7	90·5	9	95·0	6	89·5	10	87·0	16	80·0	21	68·5	37	59·5	49	69·0	39	66·5	42
2	86·3	15	83·3	18	88·5	11	81·8	18	82·3	19	75·3	26	63·3	42	57·0	49	63·5	40	60·3	44
3	78·5	20	78·0	20	79·8	18	74·8	22	74·8	23	68·7	30	62·5	39	56·0	46	62·7	38	60·2	42
4	72·4	22	72·8	21	74·3	20	70·6	23	70·6	25	64·1	30	60·8	37	53·6	43	58·9	37	56·9	39

However, if the replacing atoms are in a centrosymmetric arrangement, the three-phase invariants $\Phi_3'(= -\varphi_{\mathbf{H},R} + \varphi_{\mathbf{H}',R} + \varphi_{\mathbf{H}-\mathbf{H}',R})$ contributed by the replacing atoms will all equal 0 or π. Also, all the $\Delta\varphi_{\mathbf{H}\text{best}}$ calculated from (3) at the beginning of iteration will equal 0 or π, since all reflections have $P_+ = \frac{1}{2}$ at that moment. Hence the value of $P_+(\Delta\varphi_{\mathbf{H}})$ calculated from (1) will always be equal to $\frac{1}{2}$, leaving the phase ambiguity unresolved. This is an additional ambiguity due to the arrangement of the replacing atoms. In order to overcome this difficulty, a multi-solution procedure with random starting sign sets is used. Assign randomly to the $|\Delta\varphi_{\mathbf{H}}|$'s a positive or negative sign associated with probability $P_+ = 0{\cdot}6$ or $P_+ = 0{\cdot}4$, respectively. Then the signs are refined by the iterative calculation using (2), (3) and (1). The figure of merit shown in (4) is used to predict the quality of the solutions (see paper III).

$$\mathrm{FOM} = \left[\sum_{\mathbf{H}} m_{\mathbf{H}} E_{\mathbf{H}} \Big/ \sum_{\mathbf{H}} \exp(-\sigma_{\mathbf{H}}^2/2) \times |\cos \Delta\varphi_{\mathbf{H}}| E_{\mathbf{H}} \right] - 1. \quad (4)$$

Data

A set of error-free SIR data was calculated from the model structures of APP (avian pancreatic polypeptide) and its Hg derivative. APP belongs to space group $C2$ with unit-cell dimensions $a = 34{\cdot}18$, $b = 32{\cdot}92$, $c = 28{\cdot}44$ Å and $\beta = 105{\cdot}30°$. The asymmetric unit of the derivative contains only one replacing atom. There are about 2100 independent reflections within 2 Å resolution. 1000 largest E's were used for the test. They yield ~130 000 $\sum_2$ relationships, of which only 60 000 were included in the calculation.

Results

A test calculation was performed using ten random starting sets. The results are listed in Table 1 from left to right in descending order of the figures of merit. As can be seen, the three most left sets showing very good results correspond to the best three sets among the ten. Good results can be found even with the fifth set from the left. This shows that the procedure is very efficient and may be a useful complement to Wang's method.

The authors are indebted to Professor T. L. Blundell for making available the APP data.

References

BLOW, D. M. & ROSSMANN, M. G. (1961). *Acta Cryst.* **14**, 1195–1202.
FAN HAI-FU & GU YUAN-XIN (1985). *Acta Cryst.* **A41**, 280–284.
FAN HAI-FU, HAN FU-SON & QIAN JIN-ZI (1984). *Acta Cryst.* **A40**, 495–498.
FAN HAI-FU, HAN FU-SON, QIAN JIN-ZI & YAO JIA-XING (1984). *Acta Cryst.* **A40**, 489–495.
WANG BI-CHENG (1981). *Acta Cryst.* **A37**, C11.
WANG BI-CHENG (1984). In *Diffraction Methods for Biological Macromolecules.* A volume of *Methods in Enzymology*, edited by H. W. WYCKOFF. London: Academic Press. In the press.