Acta Cryst. (1997). D53, 792–794

A new iterative procedure for combining direct methods with solvent flattening – dealing with the phase ambiguity in protein crystallography

YUANXIN GU,[a] CHAODE ZHENG,[a] YINGDONG ZHAO,[b] HENGMING KE[b] AND HAIFU FAN[a] *at* [a]*Institute of Physics, Chinese Academy of Science, Beijing 100080, People's Republic of China, and* [b]*Department of Biochemistry and Biophysics, University of North Carolina, Chapel Hill, NC 27599, USA. E-mail: fan@aphy01.iphy.ac.cn*

(*Received 14 January 1997; accepted 6 May 1997*)

Abstract

A new procedure for combining direct methods with the solvent-flattening technique is proposed for phasing single isomorphous replacement (SIR) or one-wavelength anomalous scattering (OAS) data of proteins. The new procedure differs from the previous one [Zheng, Zheng, Gu, Mo, Fan & Hao (1997). *Acta Cryst.* D53, 49–55] in that the direct method not only provides input phases to but also accepts feedback phases from solvent flattening, thus forming an iterative process for breaking the ambiguities and refining the values of phases. The new procedure was tested with the experimental SIR data of the known structure ribonuclease Sa. For the strongest 1000 of the total 7264 refelctions, the mean F_{obs}-weighted phase error is 7.5 and 9.4° lower than that of the previous procedure and that of solvent flattening alone, respectively.

1. Introduction

Solving a mcromolecular structure by using single isomorphous replacement (SIR) or one-wavelength anomalous scattering (OAS) data is important when it is difficult to prepare multiple isomorphous derivatives and to collect multiwavelength anomalous diffraction data. However, the phase ambiguity intrinsic to SIR or OAS data obstructs their use. The solvent-flattening technique (Wang, 1981, 1985) has been used successfully for resolving the phase ambiguity. Alternatively, various kinds of direct methods have been proposed to solve the same problem (Coulter, 1965; Fan, 1965*a*,*b*; Karle, 1966; Hendrickson, 1971, Hauptman, 1982*a*,*b*; Giacovazzo, 1983; Fan & Gu, 1985; Fortier, Moore & Fraser, 1985; Klop. Krabbenham & Kroon, 1987; Giacovazzo, Cascarano & Zheng, 1988; Kyriakidis, Peshar & Schenk, 1993). Some of them have been tested sucessfully with experimental protein diffraction data (Fan, Hao, Gu, Qian, Zheng & Ke, 1990; Giacovazzo, Siliqi, & Zanotti, 1995; Langs, Guo & Hauptman, 1995). The combination of direct methods with solvent flattening for breaking the phase ambiguity has shown its advantage over direct methods or solvent flattening alone (Sha, Liu, Gu, Fan, Ke, Yao & Woolfson, 1995; Zheng, Zheng, Gu, Mo, Fan & Hao, 1997; Zheng, Gu, Zheng, Mo, Fan & Hao, 1997). The philosophy is that, while direct methods are good at deriving starting phases from SIR or OAS data, the solvent-flattening technique is powerful in the subsequent phase refinement. In our previous procedures, direct methods were simply used to provide solvent flattening with a set of starting phases and then the phases were improved by solvent flattening. However, as pointed out by Fan, Woolfson & Yao (1993) in the OAS case, Bijvoet differences usually have magnitudes comparable with those of the experimental errors and thus a considerable number of the observed Bijvoet differences will be incorrect not only in their magnitudes but also in their signs, which can in some circumstances have a very large effect on the estimate of the phase angle. This situation also happens in the SIR case. When direct methods are used in such a case, large errors, usually not randomly distributed, will then be transferred to the solvent flattening and affect significantly the result. In this paper we report a new procedure, in which the direct method provides initial phases to the solvent flattening and then the phases improved by the solvent flattening are fed back to the direct method. The iteration is carried on until a convergence is reached.

2. Method

Since the principle of our method is the same in both SIR and OAS cases, for reasons of simplicity, in the following discussion only the SIR case is considered. The phase of the reflections from the native protein can be expressed as

$$\varphi_{\mathbf{h}} = \varphi_{\mathbf{h},R} + \Delta\varphi_{\mathbf{h},R}, \tag{1}$$

where $\varphi_{\mathbf{h}}$ denotes the phase of structure factors from the native protein, $\varphi_{\mathbf{h},R}$ is the phase contributed from the isomorphously replaced atoms and $\Delta\varphi_{\mathbf{h},R}$ is the phase difference between $\varphi_{\mathbf{h}}$ and $\varphi_{\mathbf{h},R}$. The magnitude of $\Delta\varphi_{\mathbf{h},R}$ can be calculated as

$$|\Delta\varphi_{\mathbf{h},R}| = \cos^{-1}\left[(|F_{\mathbf{h},D}|^2 - |F_{\mathbf{h},N}|^2 - |F_{\mathbf{h},R}|^2)/2|F_{\mathbf{h},R}F_{\mathbf{h},N}|\right], \tag{2}$$

where $F_{\mathbf{h},D}$, $F_{\mathbf{h},N}$ and $F_{\mathbf{h},R}$ are the structure factors of the heavy-atom derivative, the native protein and the partial structure of the replacing atoms, respectively. Values of $|\Delta\varphi_{\mathbf{h},R}|$ are input to the direct methods to derive their signs. A set of starting phases is then obtained through (1). In practice, owing to experimental errors, values of $\cos(\Delta\varphi_{\mathbf{h},R})$ from (2) may be outside the range of 1 to +1. In most cases such a set of $\cos(\Delta\varphi_{\mathbf{h},R})$ are not used in the direct method. Instead they are first sorted in descending order and then modified to fit into a uniform distribution between +1 and 1. New values of $\cos(\Delta\varphi_{\mathbf{h},R})$ so obtained are used in the subsequent calculation. We described previously a simple procedure of combining direct methods with solvent flattening (Zheng, Zheng, Gu, Mo, Fan & Hao, 1997), where the direct method is used only to provide initial phases, while the solvent flattening is used to improve them. Here we describe an iterative procedure, in which the resultant phases from solvent flattening are used to calculate new values of $|\Delta\varphi_{\mathbf{h},R}|$, and then they are fed back to the direct method.

3. Test data

Data used in the present test were experimental X-ray diffraction data from the native and the platinum derivative of ribonuclease Sa (RNase Sa) at 2.5 Å resolution (Dodson, Sevcik, Dodson & Zelinka, 1987; Sevcik, Dodson & Dodson,

ISSN 0907-4449

Table 1. *Mean phase errors from different phasing procedures*

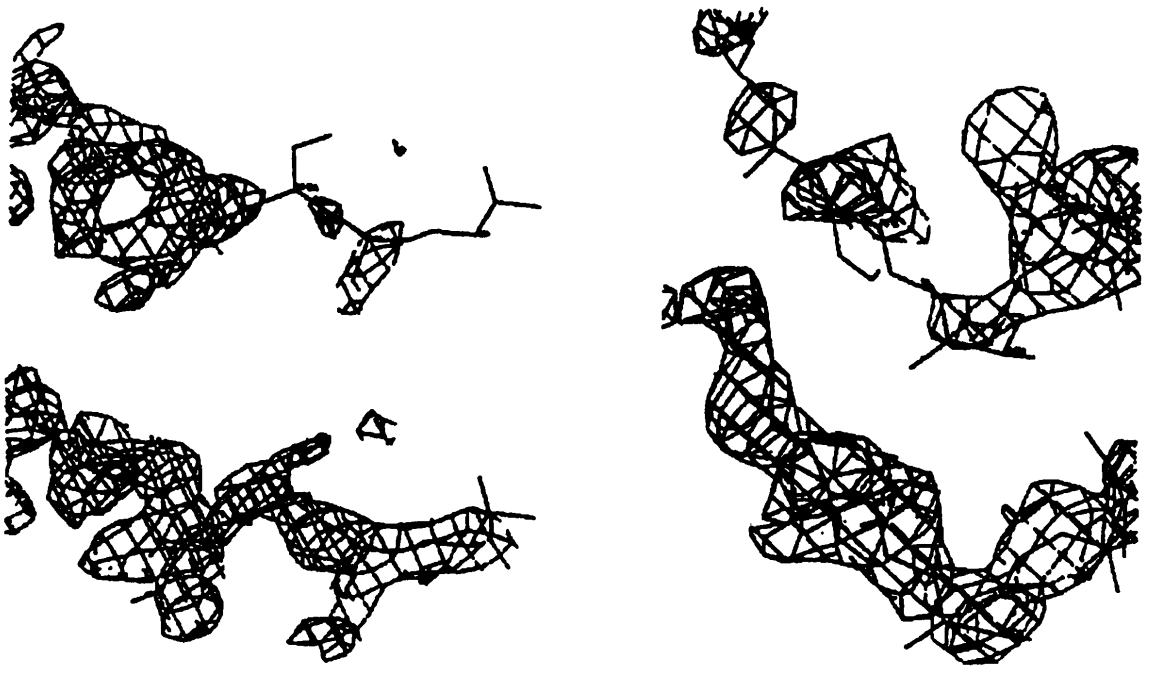
The total 7264 reflections were sorted in descending order of the observed structure-factor magnitude and then cumulated into eight groups. Values of the mean phase error were calculated against the final structure model.

		F_{obs}-weighted mean phase error (°)		
Number of reflections	Averaged SIR phases	Solvent flattening based on averaged SIR phases	Simple 'direct method + solvent flattening'	Iterative 'direct method + solvent flattening'
1000	64.26	47.27	45.34	37.86
2000	63.86	48.61	47.39	40.87
3000	64.38	50.64	49.90	43.77
4000	65.07	52.51	51.74	46.29
5000	65.96	54.06	53.59	48.67
6000	66.69	55.56	55.14	50.39
7000	67.30	56.52	56.22	51.62
7264	67.39	56.66	56.36	51.77

1991). The data were collected with Cu $K\alpha$ radiation by a routine protein crystallography method. The crystals belong to space group $P2_12_12_1$ with unit-cell parameters $a = 64.90$, $b = 78.32$ and $c = 38.79$ Å. There are two molecules in the asymmetric unit, each with 96 amino-acid residues. Five platinum positions were found in the asymmetric unit but they are only partially occupied and the sum of the partial occupancies gives approximately six Pt atoms in the whole unit cell.

4. Results and discussion

Table 1 lists the F_{obs}-weighted mean phase error for cumulative reflection groups resulting from five cycles of iterative direct method + solvent flattening. Within each cycle, two masks were used. The first mask was calculated with the direct-method phases. After two cycles of iterative solvent flattening the improved phases were used to calculate a new mask and then followed by two more cycles of solvent flattening before the phases were fed back to the direct-method processing. For a comparison, results from other procedures (see Zheng, Zheng, Gu, Mo, Fan & Hao, 1997) are also listed in Table 1. It is seen that the present method yielded results better than other methods in all eight cumulative reflection groups. For all 7264 reflections at 2.5 Å resolution, the present method gave a phase error of 51.77°. This is about 5° better then that of our previous method or that of solvent flattening alone, and is about 16° better than that of the SIR phases. For the strongest 1000 reflections the improvement is even more significant, it is about 7, 9 and 26° better than that of our previous method, the solvent flattening and the SIR phases, respectively. Map correlation coefficients of individual amino-acid residues were calculated



(*a*) (*b*)

Fig. 1. Comparison of Fourier maps on two different portions (*a*) and (*b*). Maps from the simple direct method + solvent flattening are shown on the upper part, while the corresponding maps from the iterative direct method + solvent flattening are shown on the lower part. All maps are contoured at 1σ.

for molecule A and B. In the calculation we use F_{obs} and the resultant phases to calculate the experimental Fourier map and use F_{cal} (including both magnitude and phase) from the final structure model to calculate the theoretical Fourier map. For molecule A, great improvement can be found with residues 1–2, 22–25, 30–33 and 86–89. Meanwhile a serious drawback appears with residues 40–41. For molecule B, great improvement is found with residues 6–12, 24–31, 34–37, 53–59, 64–71, 83–86 and 89–96 while a serious drawback is found with residues: 3–5, 18–19, 39–41 and 61–62. On average, the iterative direct method + solvent flattening brought the mean correlation coefficient to 0.51 and 0.48, respectively, for molecules A and B. These values are 0.05 and 0.08 better than the corresponding ones from the simple direct method + solvent flattening. In addition, the iterative procedure brought the two mean correlation coefficients of molecules A and B closer together by 0.03. Figs. 1(*a*) and 1(*b*) compare portions of the Fourier maps phased by the present procedure and the previous one. All the above results showed that the iterative direct method + solvent flattening is a powerful tool for breaking the phase ambiguity in protein crystallography.

References

Coulter, C. L. (1965). *J. Mol. Biol.* **12**, 292–295.

Dodson, G. G. Sevcik, J., Dodson, E. & Zelinka, J. (1987). *Metabolism of Nucleic Acids, Including Gene Manipulation*, pp. 33–36. Bratislava: Slovak Academy of Science.

Fan, H. F. (1965*a*). *Acta Phys. Sin.* **21**, 1114–1118. (In Chinese).

Fan, H. F. (1965*b*). *Chinese Phys.* pp. 1429–1435. (In English.)

Fan, H. F. & Gu, Y. X. (1985). *Acta Cryst.* A**41**, 280–284.

Fan, H. F., Hao, Q., Gu, Y. X., Qian, J. Z., Zheng, C. D. & Ke, H. (1990). *Acta Cryst.* A**46**, 935–939.

Fan, H. F., Woolfson, M. M. & Yao, J. X. (1993). *Proc. R. Soc. London Ser. A*, **442**, 13–32.

Fortier, S., Moore N. J. & Fraser, M. E. (1985). *Acta Cryst.* A**41**, 571–577.

Giacovazzo, C. (1983). *Acta Cryst.* A**39**, 585–592.

Giacovazzo, C., Cascarano, G. & Zheng, C. D. (1988). *Acta Cryst.* A**44**, 45–51.

Giacovazzo, C., Siliqi, D. & Zanotti, C. D. (1995). *Acta Cryst.* A**51**, 177–188.

Hauptman, H. (1982*a*). *Acta Cryst.* A**38**, 289–294.

Hauptman, H. (1982*b*). *Acta Cryst.* A**38**, 632–641.

Hendrickson, W. A. (1971). *Acta Cryst.* A**27**, 1474–1475.

Karle, J. (1966). *Acta Cryst.* **21**, 273–276.

Klop, E. A., Krabbendam, H. & Kroon, J. (1987). *Acta Cryst.* A**43**, 810–820.

Kyriakidis, C. E., Peschar, R. & Schenk, H. (1993). *Acta Cryst.* A**49**, 557–569.

Langs, D. A., Guo, D. & Hauptman, H. A. (1995). *Acta Cryst.* A**51**, 535–542.

Sevcik, J., Dodson, E. J. & Dodson, G. G. (1991). *Acta Cryst.* B**47**, 240–253.

Sha, B. D., Liu, S. P., Gu, Y. X., Fan, H. F., Ke, H., Yao, J. X. & Woolfson, M. M. (1995). *Acta Cryst.* D**51**, 342–346.

Wang, B. C. (1981). *Acta Cryst.* A**37**, (Suppl.) C-11.

Wang, B. C. (1985). *Methods Enzymol.* **115**, 90–112.

Zheng, X. F., Gu, Y. X., Zheng, C. D., Mo, Y. D., Fan, H. F. & Hao, Q. (1997). *Z. Kristallogr.* **212**, 95–98.

Zheng, X. F., Zheng, C. D., Gu, Y. X., Mo, Y. D., Fan, H. F. & Hao, Q. (1997). *Acta Cryst.* D**53**, 49–55.