Journal of
Applied Crystallography
ISSN 0021-8898

Received 6 July 1999
Accepted 20 January 2000

OASIS: a computer program for breaking phase ambiguity in one-wavelength anomalous scattering or single isomorphous substitution (replacement) data

Q. Hao,[a,b]* Y. X. Gu,[a] C. D. Zheng[a] and H. F. Fan[a]

[a]Institute of Physics, Chinese Academy of Sciences, Beijing 100080, People's Republic of China, and [b]Department of Chemistry, De Montfort University, Leicester LE1 9BH, England. Correspondence e-mail: qhao@dmu.ac.uk

The phase problem is reduced to a sign problem once the anomalous-scatterer or the replacing-heavy-atom sites are located. *OASIS* adopts the *CCP*4 format [Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763]. It applies a direct-method procedure to break the phase ambiguity intrinsic to one-wavelength anomalous scattering (OAS) or single isomorphous replacement (SIR) data.

1. Introduction

Direct methods have been used for many years in attempts to break the phase ambiguity intrinsic to one-wavelength anomalous scattering (OAS) or single isomorphous replacement (SIR) data (Fan, 1965; Karle, 1966; Hauptman, 1982; Giacovazzo, 1983; Fan & Gu, 1985; Fan *et al.*, 1990). In particular, a method based on the work of Fan & Gu (1985) has led to the first example of the solution of an unknown protein structure, rusticyanin, with OAS data at 2.1 Å resolution from a native crystal (Harvey *et al.*, 1998). The program *OASIS* is based on this method and has been written in Fortran 77 with calls to the standard *CCP*4 libraries (Collaborative Computational Project, Number 4, 1994). All processing instructions are included in a single Unix script file. The keywords are summarized in Table 1.

2. Program outline

The description below concentrates on the principles behind each step of the program.

(*a*) *Input*. Keywords from the script file are read in and interpreted by the *CCP*4 parser rountines; the X-ray diffraction (either OAS or SIR) data, which must be in `.mtz` format, is read and stored for later use.

(*b*)*E-cal*. The E values are calculated based on the scale and temperature factors obtained from the Wilson plot (Wilson, 1949). In both the OAS and the SIR case, the phase doublets are of the form

$$\varphi_h = \varphi'_h \pm |\Delta\varphi_h|, \tag{1}$$

where φ'_h is the heavy-atom phase in the SIR case or the contribution of the imaginary part of the anomalous scattering in the OAS case. All φ'_h and the absolute values of the phase doublets $|\Delta\varphi_h|$ are calculated at this stage.

(*c*) *Sigma-2*. For each reflection h, sigma2 relationships h' and $h - h'$ are found and stored.

(*d*) *Sign*. The probability of $\Delta\varphi_h$ being positive is calculated using the following formula (Fan & Gu, 1985):

$$P_+(\Delta\varphi_h) = \tfrac{1}{2} + \tfrac{1}{2}\tanh\left\{\sin|\Delta\varphi_h|\left[\sum_{h'} m_{h'} m_{h-h'} \kappa_{h,h'} \times \sin(\Phi'_3 + \Delta\varphi_{h',\mathrm{best}} + \Delta\varphi_{h-h',\mathrm{best}}) + \chi \sin\delta_h\right]\right\}, \tag{2}$$

where

$$\Phi'_3 = -\varphi'_h + \varphi'_{h'} + \varphi'_{h-h'}. \tag{3}$$

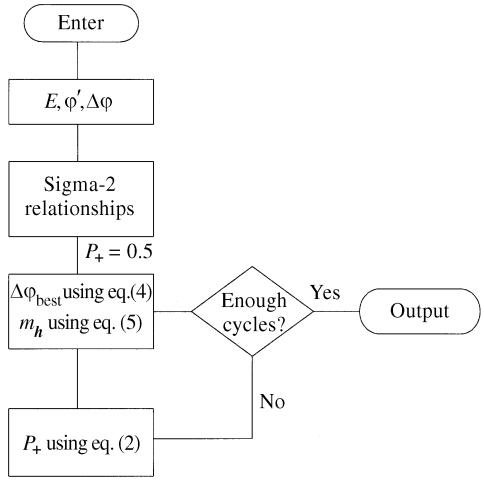

Figure 1
Flowchart of program *OASIS*.

Table 1
Summary of keywords.

Keywords must be upper case. Only the first three letters are significant except LABI and LABO where the first four are necessary. * denotes a mandatory keyword.

Keyword	Following parameters	Notes
TITLE *	Can be anything	Work title
CELL	<*a*> <*b*> <*c*> <α> <β> <γ>	Cell dimensions
SPG	(*e.g.* P 21 21 21)	Space group
CON	<atom type> <number of atoms in unit cell> ...	
HCO *	<heavy-atom type> <number of heavy atoms in unit cell>	
FIT		Uniform $\Delta\varphi$ distribution
LCE	<value>	Lack of closure error; suggested values: 20 for 'heavier' heavy atoms, *e.g.* Hg, Pt; 7 for 'lighter' heavy atoms, *e.g.* Cu, Fe
KMI	<value>	Sigma-2 relationship kappa value cutoff (default: 0.05)
CYC	<*n*cycle>	Number of cycles; recommended number for OAS data: 1
OAS		One-wavelength anomalous scattering data
SIR		Single isomorphous replacement data
ANO	<atom> <f''>	f'' value of the anomalous scatterer
POS *	<atom1> <*x*1> <*y*1> <*z*1> [<1> <occupancy1>] ... <atom*n*> <*x*n> <y*n*> <z*n*> [<*n*> <occupancy*n*>]	Fractional coordinates of anomalous scatterer(s)/heavy atom(s)
PHI		Perform phase comparison with known phases
LABIN *	F1 = ... SIGF1 = ... F2 = ... SIGF2 = ... [TPHI = ...]	F1: *F* magnitude (OAS) or native *F* magnitude (SIR); SIGF1: standard deviation; F2: Friedel difference (OAS) or the isomorphous difference $\|F(\text{derivative})\| - \|F(\text{native})\|$ (SIR); SIGF2: standard deviation; TPHI: phase to be compared with (only required when keyword PHI is present)
LABOUT	F1 = ... SIGF1 = ... PHI = ... W = ...	F1 and SIGF1: as input; PHI: derived phase; W: weight associated with the phase
END *		End of keyword list

The last term in the square bracket of equation (2) comes from the Sim distribution (Sim, 1959) and is present in the OAS case only; the 'best' phase in the formula is calculated by

$$\tan(\Delta\varphi_{h,\text{best}}) = 2(P_+ - \tfrac{1}{2})\sin|\Delta\varphi_{\mathbf{h}}|/\cos\Delta\varphi_h. \quad (4)$$

The initial value of P_+ is set to 0.5 before the first cycle.

The figure of merit m_h (Fan & Gu, 1985) associated with every phase value is calculated by the following and output to an `.mtz` file after the final cycle:

$$m_h = \exp(-\sigma_h^2/2)\left\{[2(P_+ - \tfrac{1}{2})^2 + \tfrac{1}{2}](1 - \cos 2\Delta\varphi_h) + \cos 2\Delta\varphi_h\right\}^{1/2}, \quad (5)$$

where σ_h can be determined from the standard deviation of the 'lack of closure error' (Blow & Crick, 1959).

A flowchart of the program is shown in Fig. 1.

3. Downloading and compiling the program

OASIS is written in standard Fortran 77. Any machine-dependent functions are dealt with *via* calls to rountines from the *CCP*4 suite libraries. The source code and user guide are currently available free to academic users and can be downloaded from the web page http://silicon.dmu.ac.uk/qhao under the directory `public`.

Assuming that the *CCP*4 suite has already been installed on a Unix system, *OASIS* can be compiled *via* the following command:

```
f77 -o oasis oasis.f -L$CCP4_LIB -lccp4
```

The resulting executable `oasis` file may be transferred to the directory where other *CCP*4 programs are kept (usually `$CCP4/bin`).[1]

We thank the National Natural Sciences Foundation of China for continuous support of this project.

References

Blow, D. M. & Crick, F. H. C. (1959). *Acta Cryst.* **12**, 749–802.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Fan, H. F. (1965). *Acta Phys. Sinica*, **21**, 1114–1118.
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.
Giacovazzo, C. (1983). *Acta Cryst.* A**39**, 585–592.
Harvey, I., Hao, Q., Duke, E. M. H., Ingledew, W. J. & Hasnain, S. S. (1998). *Acta Cryst.* D**54**, 629–635.
Hauptman, H. (1982). *Acta Cryst.* A**38**, 632–641.
Karle, J. (1966). *Acta Cryst.* **21**, 273–276.
Sim, G. A. (1959). *Acta Cryst.* **12**, 813–815.
Wilson, A. J. C. (1949). *Acta Cryst.* **8**, 318–321.

[1] As this paper went to press, *OASIS* became a fully supported program in version 4.0 of the *CCP*4 suite.