Vol 12 No 8, August 2003
1009-1963/2003/12(08)/0890-05

Chinese Physics

© 2003 Chin. Phys. Soc.
and IOP Publishing Ltd

Phasing of lysozyme using one-wavelength anomalous scattering of sulphur atoms through the combination of direct methods and density modification*

Jiang Fan(江 凡)†, Gu Yuan-Xin(古元新),
Zheng Chao-De(郑朝德), and Fan Hai-Fu(范海福)

Institute of Physics, Chinese Academy of Sciences, Beijing 100080, *China*

(Received 24 March 2003)

A new method of combining one-wavelength anomalous scattering (OAS) phasing and density modification has been described, in which the improved phases from density modification are re-introduced into OAS phasing. In this way, the phases could be improved iteratively until convergence. The OAS phasing method is based on the previously established sign-probability formula, which breaks the phase ambiguity in the OAS phasing. The implementation of this method has been available in CCP4 as OASIS. This method, although based on direct-methods, could also incorporate known phases and figures of merit into its sign-probability formula. In the implementation of OASIS, the known phases are from the positions of the anomalous scatters. In the current method, the known phases are from the density modification. The current method was tested on phasing a lysozyme crystal using anomalous scattering of sulphur atoms with diffraction data collected on an in-house x-ray source. The resulting map was well connected for the backbone atoms and clearly traceable, with an average map correlation coefficient of 0.6622 for the backbone atoms.

Keywords: direct methods, sign probability, anomalous scattering, density modification, sulphur atom phasing

PACC: 6110M

1. Introduction

Direct methods have proved to be useful in x-ray crystallographic phasing of proteins (for a review see Ref.[1]). In ab initio phasing, the programs SnB,[2,3] Acorn,[4,5] SHELXD,[6] and some others have succeeded in solving structures containing more than 1000 atoms when data at atomic resolution are available. In resolving phase ambiguities, there are approaches that do not make the use of the direct methods, such as Resolved Anomalous Phasing,[7] ISAS,[8] or general MIR strategy through SHARP[9] or MLPHARE.[21] Our approaches were the combination of direct methods with single wavelength anomalous diffraction (SAD or OAS)[10−12] or multi-wavelength anomalous diffraction (MAD) data.[13] In both OAS/SAD and MAD cases, the incorporation of direct methods led to better phases and maps with experimental data to medium resolution. More recently, a method of phase extension using direct methods has been proposed,[14] which is able to incorporate known phase information into direct methods formulation. Further developments will continue as the structural genomics field becomes mature because the phase problem is still one of the essential steps in determining protein structures in an automatic fashion.

In this paper, a new method of phase improvement has been proposed based on a previous work.[14] Usually, in OAS phasing, after resolving the phase ambiguity of the reflections, density modification is used to improve the phases.[11,12,15] The phases from the density modification could then be used in resolving the phase ambiguity again in OAS phasing, because the direct methods formulation in OAS phasing could now be modified to incorporate known phase information. It is obvious that this process could be iterated until convergence. The main focus of this paper is phasing of a tetragonal lysozyme crystal using the anomalous scattering of the sulphur atoms of the lysozyme. It will be shown that traceable maps were

*Project supported by the National Natural Science Foundation of China (Grant No 30171098).

†Corresponding author: E-mail: jiangf@aphy.iphy.ac.cn

obtained. This paper also focuses on the phase improvement achieved by the iterative OAS phasing and density modification as a new method for the combination of OAS phasing and density modification.

2. Materials and methods

2.1. Sign formula for OAS or SAS and single isomorphous replacement method (SIR)

The phase ambiguity in the OAS case is expressed as

$$\varphi_h = \varphi''_{h,A} \pm |\Delta\varphi_h|, \tag{1}$$

where $\varphi''_{h,A}$ and $\Delta\varphi_h$ are known from the data of OAS.[16] The sign of $\Delta\varphi_h$ is derived using the formula

$$\begin{aligned} P_+(\Delta\varphi_h) =& 0.5 + 0.5\tanh\Big\{ \sin|\Delta\varphi_h| \\ &\times \sum_{h'} m_{h'} m_{h-h'} \kappa_{h,h'} \sin(\Phi'_3 + \Delta\varphi_{h'\mathrm{best}} \\ &+ \Delta\varphi_{h-h'\mathrm{best}}) + \chi_h \sin\delta_h \Big\} \end{aligned} \tag{2}$$

where

$$\Phi'_3 = -\varphi''_{h,A} + \varphi''_{h',A} + \varphi''_{h-h',A}, \tag{3}$$

and $\chi \sin\delta_h$ is the Sim weight,[17]

$$\Delta\varphi_{h\mathrm{best}} = \varphi_{h\mathrm{best}} - \varphi''_{h,A}. \tag{4}$$

The values of $\varphi_{h\mathrm{best}}$ and m_h are calculated with the initial P_+ set to 0.5,

$$\tan(\Delta\varphi_{h\mathrm{best}}) = 2\left(P_+ - \frac{1}{2}\right)\sin|\Delta\varphi_h| / \cos\Delta\varphi_h, \tag{5}$$

$$\begin{aligned} m_h =& \exp(-\sigma_h^2/2)\Big\{ \left[2\left(P_+ - \frac{1}{2}\right)^2 + \frac{1}{2}\right] \\ &\times (1 - \cos 2\Delta\varphi_h) + \cos 2\Delta\varphi_h \Big\}^{1/2}, \end{aligned} \tag{6}$$

where σ_h is related to the experimental error and can be calculated from the mean square of the 'lack-of-closure error'.[18] For the theory behind and the practical application of Eqs.(1)–(6), the reader is referred to Refs.[11,19,20].

2.2. The modification of the OAS formulation

In the original derivation of Eq.(2),[11] the Sim-weight distribution was used for introducing the known phase information from the positions of the anomalous scatters. In the current re-interpretation of Eq.(2),[14] the real physical meaning of the Sim-weight is not used. Instead, a known phase information, $\varphi_{h,\mathrm{known}}$, with a figure of merit, m_h, is first approximated by a Gaussian distribution, and then this Gaussian distribution is approximated mathematically and numerically, by a Sim-weight distribution, using a Taylor expansion at δ_h=0, so that Eq.(2) could still be used. Therefore, χ_h is approximated by Eq.(7) (see below) and δ_h should be $\varphi_{h,\mathrm{known}} - \varphi_{h,\mathrm{best}}$. The Sim-weight term is only present and added to the triplet term when phase information is known for the corresponding index $\boldsymbol{h}$. So when phase information is not known, the corresponding Sim-weight term is the same as before, namely, using the known phase from the positions of the anomalous scatters.

For a known phase, $\varphi_{h,\mathrm{known}}$, the Sim-weight-like distribution is used to approximate the phase probability distribution, and χ_h can be calculated by approximating the distribution with the Gaussian distribution whose half-height width is related to the figure of merit of the known phase. This leads to the next equation,

$$\chi_h = -1/(2\ln(m_h)) \tag{7}$$

where m_h is the figure of merit. There is a singularity in χ_h when m_h approaches unity, which is avoided by setting m_h equal to 1 minus a tiny number. This tiny number will affect the relative weight between the Sim-weight term and the triplet term. This tiny number was chosen to be 10^{-3} and, in the meantime, a scale factor is also introduced to multiply the resulting χ_h to modulate the relative balance between the two terms. In the test presented below, the scale factor is set to 0.1. This modification of direct methods aided formulation for determining and refining the phases for single crystals could also be applied to analysing the structures of polycrystal films and powders.[25,26]

2.3. Combination of density modification and OAS

After the first cycle of OAS followed by CCP4/DM,[21] the resulting phases are more accurate than those derived from the positions of the anomalous scatters. These better phases could then be used again in another cycle of OAS, as is described in section 2.2. The resulting phases should be improved again by CCP4/DM. Moreover, an extra step of density cutting could be applied to the map derived from CCP4/DM, e.g., using CCP4/MAPMASK. A smaller fraction than the content of the protein region could be saved so that only good, reliable protein densities are used in calculating the known phases for the next cycle of OAS. The weight or 'figure of merit' for

the calculated phases could be better estimated with CCP4/SIGMAA. The overall flowchart of the combined procedure of density modification and OAS is shown in Fig.1.

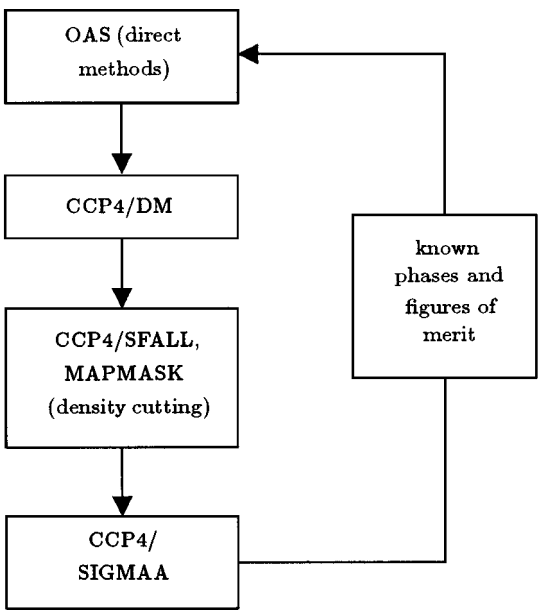


Fig.1. The flowchart of the iterative procedure of combining OAS and density modification.

2.4. Test on phasing a lysozyme crystal using sulphur atoms

To test this combined approach, a chicken egg white lysozyme crystal was used to collect the diffraction data at room temperature. The space group is $P4_32_12$ with cell parameters a=b=7.9217, c=3.8126nm. The data were collected with the help of Yamano (Rigaku, Co.) on a MicroMAX007 system (λ=0.15417nm) using the HTC image plate detector. The exposure time was 30s per degree and 360 degrees of rotation around the spindle axis were collected. The average redundancy of the measurements was about 8. The data were processed to 0.2nm with D*TREK by Pflugrath.[24] The data statistics are given in Table 1. Since this crystal was very similar to the PDB[22,27] entry 1BVX (the same space group with a=b=7.9203, c=3.8040nm), its coordinates were used as the reference for theoretical values for phases and for calculating the map correlation coefficients. To perform the OAS calculation, the positions of the ten sulphur atoms (4 disulphide bonds and 2 Met residues) in the lysozyme were taken from the coordinates of 1BVX.

Table 1. Data collection statistics of lysozyme.

Radiation/nm	0.154
x-ray power/kW	0.8
Detector and optics	Rigaku HTC, Osmic 40–10
Unit-cell parameters/nm	7.92, 7.92, 3.81, 9.00, 9.00, 9.00
Space group	$P4_32_12$
Resolution range/nm	2.8–0.2
$R_{\rm merge}$(I)/%	4.5 (22)†
Completeness/%	99.8 (99.8)
Redundancy	8 (8)
$I/\sigma(I)$	20.9 (5.5)
No of atoms (per ASU)	1011
No of sulphur atoms	10
Resolution of phasing/nm	2.8–0.2
Experimental anomalous difference/%	3.1

†Values in parentheses are for the highest resolution shell.

3. Results and discussion

For OAS phasing, a development version of CCP4/OASIS was used, including the implementation of the modification described in section 2.2. The minimum cut-off for the triplet terms was 0.02, which resulted in 17,461,590 Σ_2 relationships. Then DM was applied using the solvent content calculated with CCP4/Matthew_coef, which was 0.42. The resulting phase errors against the true values calculated using the coordinates of PDB entry 1BVX are listed in column 2 of Table 2. For iterative phase improvement as described in section 2.3, a shell script was written. For MAPMASK, "save fraction 0.4" was used. For

SIGMAA, the "combine part" option was used, that is, the phases after the density cutting were combined with that before the density cutting. The phases from SIGMAA were introduced back to the next cycle of OAS phasing. The resulting phase errors after the first and second cycles are listed in columns 3 and 4 of Table 1. Judging from the convergence of the phase errors, the optimal improvement has been reached after the second cycle. The improvement in the phases was also visible from the map correlation coefficients, using the coordinates of PDB entry 1BVX as the true model. The map correlation coefficients for all cycles of the combined OAS and DM phasing are listed in Table 2. Overall, in the final improvement the phase and the map correlation coefficients were 1.5% and 3.1%, respectively. But the phases of the strong reflections were improved about 2.5 degrees, which should be useful for tracing.

Table 2. The phase errors after each cycle of OAS and DM phasing.

reflection numbers	accumulated F_{p} weighted phase errors/(°)		
	cycle zero	cycle one	cycle two
500	35.06	33.57	32.59
1000	38.47	37.46	36.49
1500	42.45	41.47	40.87
2000	43.78	42.61	42.17
2500	44.79	43.41	42.94
3000	45.75	44.25	44.04
3500	46.31	45.05	44.77
4000	47.06	45.83	45.73
4500	47.55	46.37	46.22
5000	48.29	47.12	47.10
5500	48.86	47.84	47.80
6000	49.38	48.40	48.40
6500	49.94	48.99	49.00
7000	50.37	49.50	49.54
7500	50.84	50.00	50.05
8000	51.22	50.40	50.43
8506	51.52	50.73	50.74

The map correlation coefficients by residues are plotted in Fig.2. The final electron density map of the lysozyme crystal phased in the current test was traceable, judging from the map correlation coefficient and visual inspection. Visual inspection of the whole–molecule and its maps showed large breaks at only seven residues, which were at turns of the backbone trace. There were other minor breaks at 16 main chain atoms, which did not interrupt tracing significantly. Therefore, the phasing of this lysozyme crystal using the anomalous scattering of sulphur atoms could be considered successful.

Since the solvent content was 0.42, the "save fraction 0.6" in MAPMASK was also tested. This would correspond to saving the protein regions of the map. The resulting phases and maps were not so good as those with "save fraction 0.4". As 0.4 would mean saving a percentage of the protein regions that were more reliable (with higher densities) and gave better results, this was consistent with the idea of SIGMAA of using a partial better determined model and producing a phase set without bias. Therefore, it was not surprising that through combined iterative approach as outlined in Fig.1, the phases were improved more than that of using a single cycle of DM.

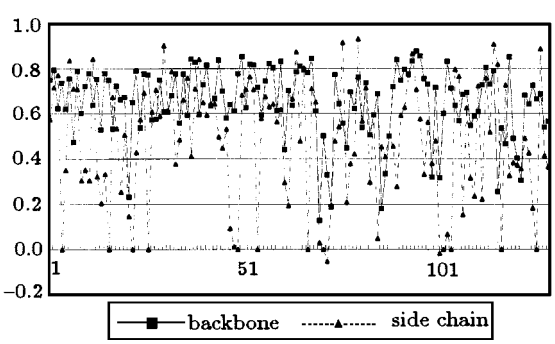


Fig.2. Map correlation coefficients by residues.

Theoretical anomalous difference for the current lysozyme structure is about 1% while the measured difference was 3.1% (see Table 1). So there is still room for improving the quality of the experimental data, and more accurate data will help solve the positions of sulphur atoms. From this study, we could observe that even when the positions of anomalous scatters were not solvable from their anomalous diffraction difference, the phase ambiguity problem could be solved to the point of determining the final structure using our OASIS method. Therefore, the problem of determining the anomalous scatters' positions could be solved independently of breaking the phase ambiguity. For example, a different wavelength of soft x-ray available at synchrotron sources was used by Brown *et al*[23] to phase a structure using sulphur atoms. Moreover, more accurate measurement of the anomalous differences using the inverse beam strategy could be used, which should help determining the positions of sulphur atoms. Once the positions of the sulphur atoms could be determined, our study suggests that the sulphur atoms have enough phasing power to phase the lysozyme crystal.

4. Conclusions

The combination of density modification and OAS phasing has been shown to improve the phases through iteration of density modification and OAS phasing. The amount of improvement for phasing a tetragonal lysozyme crystal using anomalous scattering of sulphur atoms was 1.5% and 3.1% for the mean phase error and map correlation coefficient, respectively. It has also been shown that when the positions of the sulphur atoms were known, the phases of a tetragonal lysozyme crystal could be determined by using the anomalous signal of sulphur atoms in the diffraction data collected on an in-house x-ray source.

References

[1] Hauptman H A 1997 *Curr. Opin. Struct. Biol.* **7** 672
[2] Miller R, DeTitta G T, Jones R, Langs D A, Weeks C M and Hauptman H A 1993 *Science* **259** 1430
[3] Weeks C M, Hauptman H A, Smith G D, Blessing R H, Teeter M M and Miller R 1995 *Acta Cryst.* D **51** 33
[4] Foadi J, Woolfson M M, Dodson E J, Wilson K S, Yao J X and Zheng C D 2000 *Acta Cryst.* D **56** 1137
[5] Yao J X 2002 *Acta Cryst.* D **58** 1941
[6] Sheldrick G M 1998 *SHELX: Applications to Macromolecules. Direct Methods for Solving Macromolecular Structures* ed S Fortier (Dordrecht: Kluwer) pp 401–11
[7] Hendrickson W A and Teeter M M 1981 *Nature* **290** 107
[8] Wang B C 1985 *Meth. Enzymol.* **115** 90
[9] Le Fortelle E and Bricogne G 1997 *Meth. Enzymol.* **276** 472
[10] Fan H F 1965 *Acta Phys. Sin.* **21** 1114 (in Chinese)
[11] Fan H F and Gu Y X 1985 *Acta Cryst.* A **41** 280
[12] Liu Y D, Harvey I, Gu Y X, Zheng C D, He Y Z, Fan H F, Hasnain S S and Hao Q 1999 *Acta Cryst.* D **55** 1620
[13] Gu X Y, Liu Y D, Hao Q, Ealick S E and Fan H F 2001 *Acta Cryst.* D **57** 250
[14] Jiang F 2002 *Acta Cryst.* D **58** 1277
[15] Gu Y X, Zheng C D, Zhao Y D, Ke H and Fan H F 1997 *Acta Cryst.* D **53** 792
[16] Kartha G 1975 *Anomalous Scattering*, edited by S Ramaseshan & S C Abrahams (Copenhagen: Munksgaard) p369
[17] Sim G A 1959 *Acta Cryst.* **12** 813
[18] Blow D M and Crick F H C 1959 *Acta Cryst.* **12** 794
[19] Fan H F, Hao Q, Gu Y X, Qian J Z, Zheng C D and Ke H 1990 *Acta Cryst.* A **46** 935
[20] Hao Q 2000 *J. Synchrontron Rad.* **7** 148
[21] Collaborative Computational Project 1994 *Acta Cryst.* D **50** 760
[22] Bernstein F C, Koetzle T F, Williams J B, Meyer E F Jr, Brice M D, Rodgers J R, Kennard O, Shimanouchi T and Tasumi M 1977 *J. Mol. Biol.* **112** 535
[23] Brown J, Esnouf R M, Jones M A, Linnell J, Harlos K, Hassan A B and Jones E Y 2002 *EMBO J.* **21** 1054
[24] Pflugrath J W 1999 *Acta Cryst.* D **55** 1718
[25] Wang Y Q, Liao X B, Diao H W, Zhang S B, Xu Y Y, Chen C Y, Chen W D and Kong G L 2002 *Chin. Phys.* **11** 492
[26] Chen J R, Gu Y X and Fan H F 2003 *Chin. Phys.* **12** 310
[27] Berman H M, Westbrook J, Feng Z, Gilliland G, Bhat T N, Weissig H, Shindyalov I N and Bourne P E 2000 *Nucleic Acid Res.* **28** 235