KOVALCHUK, M. V. & KOHN, V. G. (1986). *Sov. Phys. Usp.* **29**, 426-446.
KOVALCHUK, M. V., KOHN, V. G. & LOBANOVICH, E. F. (1985). *Sov. Phys. Solid State*, **27**, 3379-3387.
KROLZIG, A., MATERLIK, G. & ZEGENHAGEN, J. (1983). *Nucl. Instrum. Methods*, **208**, 613-619.
MATERLIK, G., FRAHM, A. & BEDZYK, M. J. (1984). *Phys. Rev. Lett.* **52**, 441-444.
MATERLIK, G. & ZEGENHAGEN, J. (1984). *Phys. Lett. A*, **104**, 47-50.
PATEL, J. R. & GOLOVCHENKO, J. A. (1983). *Phys. Rev. Lett.* **50**, 1858-1861.
POLIKARPOV, M. A. & YAKIMOV, S. S. (1986). *Sov. Phys. Solid State*, **28**, 907-910.
ZHELUDEVA, S. I., KOVALCHUK, M. V. & KOHN, V. G. (1985). *J. Phys. C*, **18**, 2287-2304.

Acta Cryst. (1990). A**46**, 656-659

The Application of One-Wavelength Anomalous Scattering. I. Combining Results of Different Methods

BY FAN HAI-FU

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

AND HAO QUAN AND M. M. WOOLFSON

Department of Physics, University of York, York YO1 5DD, England

(*Received* 2 *December* 1989; *accepted* 7 *March* 1990)

Abstract

Two different techniques for employing one-wavelength anomalous scattering, one using a direct-methods approach and the other a Patterson-like function, are applied to two known protein structures. The first of these, avian pancreatic polypeptide, is in space group $C2$ with one molecule containing 36 amino-acid residues in the asymmetric unit. The second, ribonuclease Sa in space group $P2_12_12_1$, has two molecules each containing 96 amino-acid residues in the asymmetric unit. Both methods give phase indications easily leading to the elucidation of the smaller structure and probably enabling the larger structure to be solved as well. For each structure the electron density maps from the phases given by the two methods are combined through a minimum function. The Fourier transform of the resultant map gives phases better than those given by the individual methods, reducing the mean phase error by 2-3°, which could be critical in some applications.

Introduction

In principle the techniques of multiple isomorphous replacement (MIR) and many-wavelength anomalous dispersion (MAD) enable phases to be determined explicitly and hence structures to be solved. In particular, the MIR method has been very successful for the solution of protein structures; indeed it could be said that the present advanced state of protein crystallography is almost entirely due to this technique. There do occur situations where the MIR technique cannot be applied – for example, when derivatives are not isomorphous with the native product. In such a case the native material may contain a heavy, or fairly heavy, atom such as mercury or zinc and the technique of anomalous scattering is available. Although MAD is then possible (Hendrickson, Pähler, Smith, Satow, Merritt & Phizackerley, 1989), the problem of taking accurate data at different wavelengths with a synchrotron source and then scaling them together does present considerable problems. By contrast one-wavelength anomalous-scattering (OAS) data can be taken much more easily and have been used successfully to solve protein structures.

Techniques for the use of OAS data include combining information from anomalous differences with direct methods (Fan Hai-fu, Han Fu-son, Qian Jin-zi & Yao Jia-xing, 1984) and also use of the P_s function, first introduced by Okaya, Saito & Pepinsky (1955) and further developed by Hao Quan & Woolfson (1989). Examination of the results of applying these two methods reveals that, while they use the same basic data and give mean phase errors of similar magnitude, there are significant differences in the distribution of the errors – so that a reflexion with a large phase error from one technique does not necessarily have a large phase error from the other. This led us to examine the possibility of combining the results of the two techniques to obtain something better than either of them individually; this work is reported here.

0108-7673/90/080656-04$03.00

Combining results for App

The structure of the hormone App, avian pancreatic polypeptide, a small globular protein containing 36 amino-acid residues (Glover, Moss, Tickle, Pitts, Haneef, Wood & Blundell, 1985), is an ideal example on which to test methods. The native data extend to 0·98 Å and anomalous-scattering data are available to 2·04 Å for a mercury derivative (Pitts, Tickle, Wood & Blundell, 1982). The space group is $C2$ and for the native material $a = 34{\cdot}18$, $b = 32{\cdot}92$, $c = 28{\cdot}44$ Å and $\beta = 105{\cdot}30°$. For the derivative there is one Hg atom in the asymmetric unit and there are available 2109 independent pairs of magnitudes $|F(\mathbf{h})|$ and $|F(\bar{\mathbf{h}})|$.

The direct-methods approach (Fan Hai-fu *et al.*, 1984) uses phase relationships to resolve the phase-ambiguity problem expressed as $\varphi = \varphi' \pm |\Delta\varphi|$. Here φ and φ' are the structure factors of the protein and the anomalous-scatterer substructure respectively and $|\Delta\varphi|$, the magnitude of the difference, can be determined from the data. The results of applying this technique are shown in Table 1. It will be seen that the mean error seems only slightly related to the value of F_{obs} since an F_{obs}-weighted error differs very little from that without weights. For the results of the P_s-function method, shown in Table 2, the effect of F_{obs} weighting is actually to increase the error over the unweighted value. However, as was shown by Hao Quan & Woolfson (1989) there is an alternative weighting scheme, related to the way the method is applied, that gives an overall weighted mean error of 29·47°.

One method of judging the quality of a set of phases obtained by a method under trial for a known structure is to look at the correlation coefficient between the electron density maps obtained with the estimated phases and the true phases. This is defined as

$$r = (\overline{\rho_t \rho_e} - \overline{\rho_t}\,\overline{\rho_e})/\sigma_t \sigma_e, \qquad (1)$$

where ρ_t and ρ_e are the true and estimated electron densities and σ_t and σ_e are the corresponding standard deviations of the distributions of electron density. For the direct-method and P_s-function processes the values of r are 0·70 and 0·67 respectively, both quite good values as measures of map quality.

Since both estimated maps have a good overlap with the true density it seemed likely that the estimated density is more likely to be true than false in those regions where the two maps show density in common. For this reason we calculated a minimum function from the two maps

$$M(\mathbf{r}) = \min\,[\rho_1(\mathbf{r}), \rho_2(\mathbf{r})], \qquad (2)$$

where $\rho_1(\mathbf{r})$ and $\rho_2(\mathbf{r})$ are the estimated electron densities from the two methods at the position $\mathbf{r}$.

The Fourier transform of $M(\mathbf{r})$ then gives new phase estimates which are shown in Table 3. It will be seen that there is a marginal reduction in the overall mean phase error, although this is more marked for the stronger reflexions and for the weighted mean errors. In addition, it is found that the correlation coefficient for the new calculated map is 0·71, a small improvement over the original values. What we have indicated here is that the method of combining the

Table 1. *The mean phase errors and $|F_{obs}|$-weighted mean phase errors for App given by the direct method*

The first column gives the total number of reflexions with $|F|$ greater than the minimum $|F_{obs}|$.

Total number of reflexions	Minimum $\|F_{obs}\|$	Mean phase error (°)	$\|F_{obs}\|$-weighted phase error (°)
200	756	35·83	36·29
400	572	34·31	34·79
600	476	34·61	34·95
800	410	37·01	36·56
1000	348	38·08	37·28
1200	301	39·08	37·96
1400	254	39·50	38·25
1600	206	39·31	38·22
1800	151	38·72	38·03
2000	86	38·76	38·08
2109	1	38·36	38·03

Table 2. *The mean phase errors and $|F_{obs}|$-weighted mean phase errors for App given by the P_s-function method*

The first column gives the total number of reflexions with $|F|$ greater than the minimum $|F_{obs}|$.

Total number of reflexions	Minimum $\|F_{obs}\|$	Mean phase error (°)	$\|F_{obs}\|$-weighted phase error (°)
200	756	39·34	40·15
400	572	37·39	38·21
600	476	38·75	39·06
800	410	40·95	40·55
1000	348	40·72	40·47
1200	301	41·42	40·92
1400	254	41·19	40·81
1600	206	41·12	40·80
1800	151	40·83	40·70
2000	86	40·66	40·69
2109	1	39·93	40·57

Table 3. *The mean phase errors and $|F_{obs}|$-weighted mean phase errors for App given by the minimum-function combination of the phases from the direct and P_s-function methods*

The first column gives the total number of reflexions with $|F|$ greater than the minimum $|F_{obs}|$.

Total number of reflexions	Minimum $\|F_{obs}\|$	Mean phase error (°)	$\|F_{obs}\|$-weighted phase error (°)
200	756	33·98	34·66
400	572	32·85	33·36
600	476	33·34	33·65
800	410	34·85	34·65
1000	348	35·32	34·96
1200	301	36·13	35·48
1400	254	36·34	35·62
1600	206	36·81	35·90
1800	151	36·97	35·99
2000	86	37·43	36·17
2109	1	37·91	36·23

Table 4. *The mean phase errors and $|F_{obs}|$-weighted mean phase errors for RNA given by the direct method*

The first column gives the total number of reflexions with $|F|$ greater than the minimum $|F_{obs}|$.

Total number of reflexions	Minimum $\|F_{obs}\|$	Mean phase error (°)	$\|F_{obs}\|$-weighted phase error (°)
500	433	51·80	50·87
1000	354	55·83	54·61
1500	305	57·00	55·84
2000	267	57·81	56·68
2500	239	59·03	57·71
3000	211	60·60	58·90
3500	187	61·56	59·65
4000	165	62·71	60·47
4500	144	64·10	61·37
5000	124	65·43	62·16
5500	105	66·91	62·96
6000	84	68·20	63·59
6500	59	69·60	64·15
7008	4	70·99	64·46

Table 5. *The mean phase errors and $|F_{obs}|$-weighted mean phase errors for RNA given by the P_s-function method*

The first column gives the total number of reflexions with $|F|$ greater than the minimum $|F_{obs}|$.

Total number of reflexions	Minimum $\|F_{obs}\|$	Mean phase error (°)	$\|F_{obs}\|$-weighted phase error (°)
500	433	50·06	49·24
1000	354	56·58	54·96
1500	305	57·37	56·01
2000	267	58·99	57·49
2500	239	60·79	58·98
3000	211	63·17	60·77
3500	187	64·64	61·90
4000	165	66·07	62·94
4500	144	67·50	63·92
5000	124	68·86	64·76
5500	105	69·98	65·43
6000	84	71·09	66·01
6500	59	72·53	66·59
7008	4	73·48	66·85

two sets of phase estimates gives a slightly improved result although, for App, it would not be decisive in terms of structure determination. However, since the computational cost of the process is very small it is worth doing.

An application to ribonuclease Sa

The direct and P_s-function methods have been applied to experimental 2·5 Å OAS data [7008 independent pairs of magnitudes $|F(\mathbf{h})|$ and $|F(\bar{\mathbf{h}})|$] for a platinum derivative of ribonuclease Sa (RNA) (Dodson, Sevcik, Dodson & Zelinka, 1987). The space group is $P2_12_12_1$ with $a = 64·85$, $b = 78·56$ and $c = 39·51$ Å. There are two molecules in the asymmetric unit, each with 96 amino-acid residues and, in addition, there are more than 200 ordered water molecules. There are five platinum positions in each asymmetric unit but they are only partially occupied so that there are effectively six platinum atoms in the unit cell.

Table 6. *The mean phase errors and $|F_{obs}|$-weighted mean phase errors for RNA given by the minimum-function combination of the phases from the direct and P_s-function methods*

The first column gives the total number of reflexions with $|F|$ greater than the minimum $|F_{obs}|$.

Total number of reflexions	Minimum $\|F_{obs}\|$	Mean phase error (°)	$\|F_{obs}\|$-weighted phase error (°)
500	433	47·26	46·57
1000	354	53·08	51·66
1500	305	54·05	52·78
2000	267	55·27	53·95
2500	239	56·59	55·08
3000	211	58·57	56·56
3500	187	59·56	57·36
4000	165	61·16	58·46
4500	144	62·71	59·47
5000	124	63·93	60·23
5500	105	65·64	61·15
6000	84	67·25	61·93
6500	59	68·47	62·45
7008	4	69·40	62·70

The results found for the direct, P_s and combined processes are shown in Tables 4, 5 and 6 respectively. Once again is seen the slight superiority of the direct over the P_s-function method and also the gain in combining the estimates *via* the minimum function. This is also seen in the correlation coefficients, which are 0·43, 0·41 and 0·45 respectively. For this example there may be a distinct advantage in the combined results from the point of view of determining the structure since the combined map is much cleaner in appearance than the individual ones. Our experience is that a 2–3° improvement in the initial phase estimates can sometimes make all the difference between success and failure in phase refinement and extension techniques, such as those due to Wang (1985) or Zhang & Main (1990). The RNA map had several

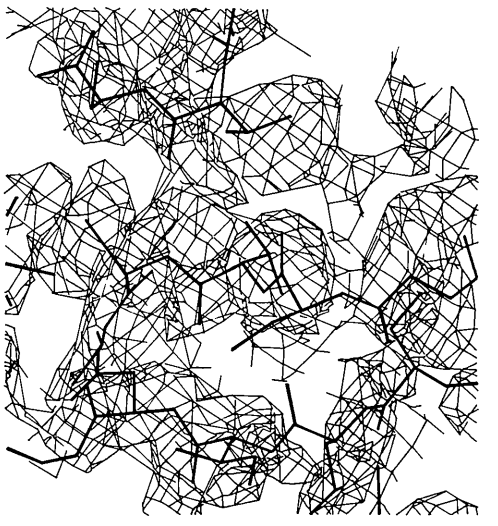

Fig. 1. Part of a *FRODO* map (Jones, 1985) obtained from RNA phases from the minimum-function procedure with the structure superimposed.

quite good regions but others were much less clear. We show in Fig. 1 one of the better portions of a *FRODO* map (Jones, 1985) for RNA from our final phases, with the structure superimposed. We conclude that the results for RNA are rather marginal; to progress from the starting point we provide to a complete elucidation of the structure would probably be possible, but would certainly require prior stereochemical information and much experience in protein crystallography.

Concluding remarks

We are encouraged by the fact that for a structure of the size of RNA the automatic procedures we have so far devised can give results at the margin of being useful. One thing we have confirmed, although it is a well known effect, is that the accuracy of the data is very critical in work of this kind. When we repeat analyses using calculated data from the known structure then there results a remarkable improvement in phase estimates. We are hoping to strengthen our analytical methods but we hope that experimentalists will also be able to improve the quality of the data they provide.

We are most grateful to the Royal Society and the Chinese Academy of Sciences with whose support the collaborative research at Beijing and York is carried out. One of us (HQ) is also indebted to the Royal Society for the award of a Royal Fellowship. Direct-methods work at York is supported by the Science and Engineering Research Council, the Wellcome Trust and the Wolfson Foundation, to whom we give thanks.

References

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 HAI-FU, HAN FU-SON, QIAN JIN-ZI & YAO JIA-XING (1984). *Acta Cryst.* A**40**, 489-495.

GLOVER, I. D., MOSS, D. S., TICKLE, I. J., PITTS, J. E., HANEEF, I., WOOD, S. P. & BLUNDELL, T. L. (1985). *Adv. Biophys.* **20**, 1-12.

HAO QUAN & WOOLFSON, M. M. (1989). *Acta Cryst.* A**45**, 794-797.

HENDRICKSON, W. A., PÄHLER, A., SMITH, J. L., SATOW, Y., MERRITT, E. A. & PHIZACKERLEY, R. P. (1989). *Proc. Natl Acad. Sci. USA*, **86**, 2190-2194.

JONES, T. A. (1985). *Methods Enzymol.* **115**, 157-171.

OKAYA, Y., SAITO, Y. & PEPINSKY, R. (1955). *Phys. Rev.* **98**, 1857-1858.

PITTS, J.E., TICKLE, I. J., WOOD, S. P. & BLUNDELL, T. L. (1982). *Sov. Phys. Crystallogr.* **27**, 56-62.

WANG, B. C. (1985). *Methods Enzymol.* **115**, 90-112.

ZHANG, K. Y. J. & MAIN, P. (1990). *Acta Cryst.* A**46**, 41-46, 377-381.

Acta Cryst. (1990). A**46**, 659-664

The Application of One-Wavelength Anomalous Scattering. II. An Analytical Approach for Phase Determination

BY FAN HAI-FU

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

AND HAO QUAN AND M. M. WOOLFSON

Department of Physics, University of York, York YO1 5DD, England

(*Received* 5 *December* 1989; *accepted* 7 *March* 1990)

Abstract

An analytical method has been developed by which phase estimates may be uniquely determined from one-wavelength anomalous-scattering data; the method as described can be applied to structures containing one type of anomalous scatterer. The method has been tested on two structures. The first is an Hg derivative of a small protein, avian pancreatic polypeptide (App), crystallizing in space group $C2$ with one molecule of 36 amino-acid residues in the asymmetric unit. The second is a Pt derivative of ribonuclease Sa (RNA), crystallizing in space group $P2_12_12_1$ with two molecules of 96 amino-acid residues in the asymmetric unit. The phases for App give an electron density map which can easily be interpreted in terms of a model. For RNA the map is less clear but has strong similarities with the true map and could probably be interpreted. If anomalous scatterers are centrosymmetrically arranged then the analysis shows

0108-7673/90/080659-06$03.00