ELSEVIER

Science and Technology of Advanced Materials 6 (2005) 722–724

SCIENCE AND TECHNOLOGY OF ADVANCED MATERIALS

www.elsevier.com/locate/stam

Developments on methods of solving crystal structures☆

Hai-fu Fan*

Institute of Physics, Chinese Academy of Sciences, Beijing 100080, People's Republic of China

Received 14 March 2005; accepted 30 May 2005
Available online 23 September 2005

Abstract

Methods tackling the phase problem in diffraction analysis under various circumstances have been studied in the Institute of Physics in Beijing. Brief description on the development of phasing methods for solving aperiodic crystal structures, image processing in high-resolution electron microscopy and in high-throughput structure determination of proteins will be given.

Keywords: Direct methods; Incommensurate structures; Electron crystallography; Protein crystallography

1. Introduction

Crystal structure determination by diffraction analysis provides an important experimental basic on understanding the structure–property/function relationship of solid-state materials and biological macromolecules. Solution of the phase problem plays a central role in diffraction analysis. Methods on solving the phase problem under various circumstances have been studied in the Institute of Physics in Beijing. These include mainly: (1) ab initio determination of incommensurately modulated and composite structures without relying on any guessed models; (2) methods combining X-ray crystallography and electron microscopy for retrieving crystal-structure image from electron micrographs; and (3) combining direct methods and macromolecular methods for high-throughput structure determination of proteins.

2. Direct methods in superspace—ab initio determination of aperiodic crystal structures

Incommensurate modulated and composite structures can be found in many important solid-state materials. A common feature of these structures is lack of three-dimensional periodicity. With traditional methods, such structures are usually solved by least-squares refinement based on a guessed modulation model. This is often rather difficult and may easily lead to biased results. Traditional direct methods have been extended to multi-dimensional space [1–3]. Special techniques were developed enabling direct solution of the phase problem for incommensurate and composite crystal structures. The program DIMS (*D*irect methods in *M*ulti-*D*imensional *S*pace) [4–6] has been written for the implementation. Fig. 1 shows sections of the four-dimensional electron-density function of the high-Tc superconductor Bi-2212. The maps were phased by DIMS revealing objectively and intuitively the incommensurate modulation prior to structure refinement and model building.

3. Image processing in high-resolution electron microscopy

Apart from diffraction analysis, high-resolution electron microscopy (HREM) is another important technique of solving crystal structures. Many solid-state materials important in science and technology are composed of very small and imperfect crystals. They are not suitable for X-ray diffraction analysis but are suitable for electron microscopic observation. However, HREM suffers from two difficulties. Firstly, an electron microscopic image is not a true structure image of the object but just a convolution of the true image with the Fourier transform of the contrast transfer function. Secondly, the point-to-point resolution of an electron microscopic image is often not enough to reveal individual

☆ http://cryst.iphy.ac.cn

* Corresponding author.
E-mail address: fanhf@cryst.iphy.ac.cn.

1468-6996/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.stam.2005.05.018

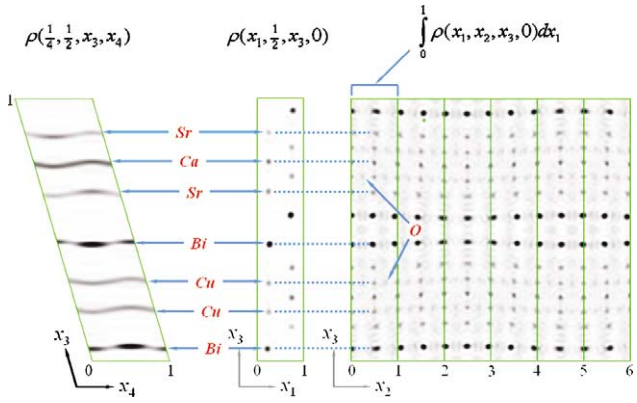

Fig. 1. Sections of the four-dimensional electron-density function of the high-Tc superconductor Bi-2212. Right: the three-dimensional hyper-section at $x_4=0$ projected along the x_1-axis. Six unit cells are plotted along the x_2-axis. Positional modulation of metal atoms and that of oxygen atoms in Cu–O layers are clearly seen along the x_2-axis. Middle: the two-dimensional section at $x_2=0.5$ and $x_4=0$ showing all metal atoms. Left: the two-dimensional section at $x_1=0.25$ and $x_2=0.5$ showing all metal atoms modulated along the x_4-axis.

atoms. Hence, it is essential to have some image processing technique for both image deconvolution and resolution enhancement. On the other hand, direct methods developed in X-ray crystallography are in fact a special kind of image processing technique. Retrieving phases from diffraction amplitudes by direct methods is equivalent to retrieve the crystal structure from the blurred image—the Fourier transform of squared diffraction amplitudes—i.e. the Patterson function. A two-step image processing technique has been developed, which combines direct methods with HREM enabling a single electron microscopic image to be deconvoluted and its point-to-point resolution be enhanced

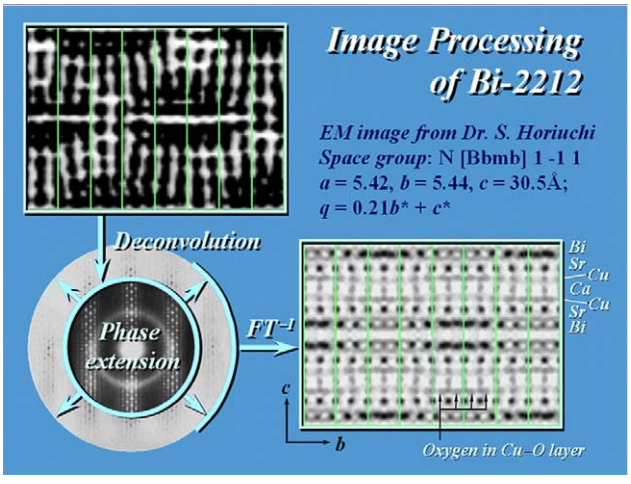

Fig. 2. Electron microscopic image processing of the high-Tc superconductor of Bi-2212. Top left: the experimental high-resolution electron microscopic image of Bi-2212 at 2 Å resolution. Bottom left: the corresponding experimental electron diffraction pattern at 1 Å resolution. Right: the resultant image from the two-step processing, which combines information from the electron microscopic image and the corresponding diffraction pattern. All metal atoms and the oxygen atoms in Cu–O layers are clearly revealed.

to about 1 Å [7,8]. The procedure has been made visualized with the program VEC [9]. Fig. 2 shows results of image processing of the electron microscopic image of the high-Tc superconductor Bi-2212.

4. Direct-method SAD phasing in protein crystallography

In order to understand life processes at the atomic/molecular level, an efficient technique is required for high-throughput determination of three-dimensional structure of proteins. Over the past decade, advances in synchrotron technology and genetic engineering led to the great success of multi-wavelength anomalous diffraction (MAD) method. However, MAD phasing may encounter problems when crystals of the sample protein are sensitive to X-ray irradiation or when the seleno-methionine substitution is not successful. Recent advances in diffraction data collection and phasing have opened new possibilities of the single-wavelength anomalous diffraction (SAD) method, which has the following advantages: (1) the exposure time is shorter; (2) the sample preparation is easier; (3) the intensity-scaling process is simpler; and (4) there is no need to fine tune the wavelength to the peak and inflection point of the anomalous scattering curve. This implies that the SAD method can be used to deal with diffraction data from either synchrotron or in-house sources. The most exciting feature of SAD phasing is the ability of using anomalous-scattering signals from sulfur atoms in most native proteins. A drawback of SAD phasing is the intrinsic phase ambiguity. A direct-method has been developed to break the phase ambiguity [10,11]. In comparison with traditional methods the main points of

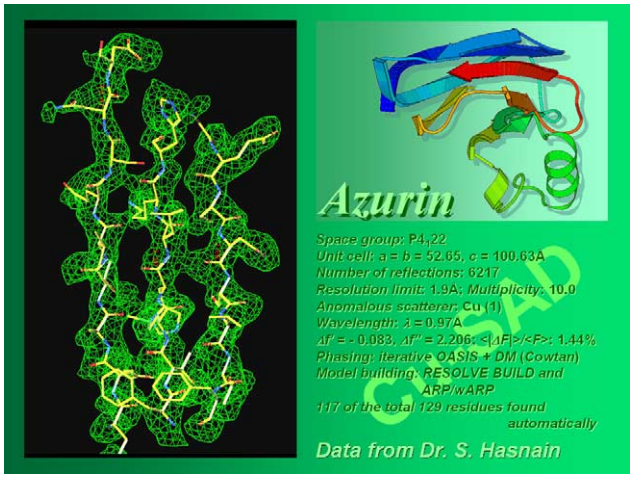

Fig. 3. Crystal structure of the protein azurin automatically solved from the copper-SAD data by the combination of the following programs: SHELXD (locating heavy atoms) [16], SOLVE (refining heavy atoms) [17], OASIS-2004 (SAD phasing and fragment extension) [15], DM (density modification) [18], RESOLVE BUILD (initial model building) [19,20] and ARP/wARP (subsequent model building) [21].

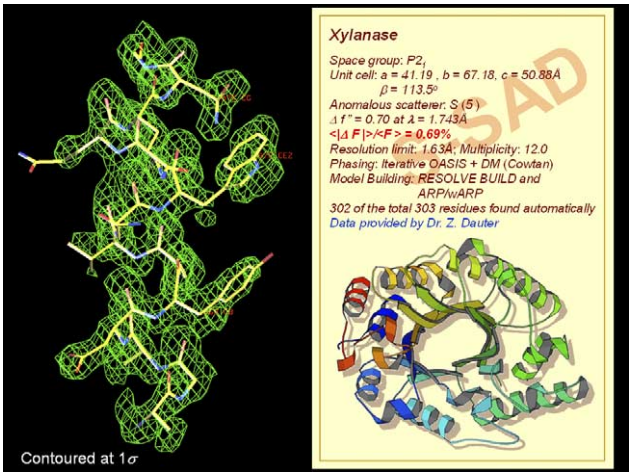

Fig. 4. Crystal structure of the protein xylanase automatically solved from the sulphur-SAD data by the combination of the following programs: SHELXD (locating heavy atoms) [16], SOLVE (refining heavy atoms) [17], OASIS-2004 (SAD phasing and partial-structure expansion) [15], DM (density modification) [18], RESOLVE BUILD (initial model building) [19,20] and ARP/wARP (subsequent model building) [21].

the direct-method are: (1) the product of Cochran distribution and Sim distribution is used instead of the Sim distribution alone for the discrimination of the two possible phases of each reflection; (2) the bimodal phase distribution of SAD is introduced into the direct-method formulation, thus the '0 to 2π' phase problem is reduced to a 'plus or minus' sign problem; and (3) the concept of lack-of-closure error and that of 'best phase' and 'figure of merit' for each reflection used in protein crystallography are introduced into the direct-method phasing. The program OASIS [12] was written to implement the method. Recently, important improvement [13,14] was made and a new version OASIS－2004 was written [15]. Figs. 3 and 4 show results on automatic structure determination of two proteins with SAD data.

Acknowledgements

The project is supported by the Chinese Academy of Sciences and by the Ministry of Science and Technology of China (Grant No.: 2002CB713801;2002BA711A12).

References

[1] Q. Hao, Y.W. Liu, H.F. Fan, Direct methods in superspace I. Preliminary theory and test on the determination of incommensurate modulated structures, Acta Crystallogr. A43 (1987) 820–824.

[2] H.F. Fan, S. van Smaalen, E.J.W. Lam, P.T. Beurskens, Direct methods for incommensurate intergrowth compounds I. Determination of the modulation., Acta Crystallogr. A49 (1993) 704–708.

[3] Y.D. Mo, Z.Q. Fu, H.F. Fan, S. van Smaalen, E.J.W. Lam, P.T. Beurskens, Direct methods for incommensurate intergrowth compounds III. Solving the average structure in multidimensional space, Acta Crystallogr. A52 (1996) 640–644.

[4] Z.Q. Fu, H.F. Fan, DIMS—a direct method program for incommensurate modulated structures, J. Appl. Crystallogr. 27 (1994) 124–127.

[5] H.F. Fan, DIMS on the VEC platform, IUCr CompComm Newsletter No.5 (2005) 16–23.

[6] H.F. Fan, DIMS/VEC applications, IUCr CompComm Newsletter No.5 (2005) 24–27.

[7] F.S. Han, H.F. Fan, F.H. Li, Image Processing in high resolution eletron microscopy using the direct method II. Image deconvolution, Acta Crystallogr. A42 (1986) 353–356.

[8] Fan Hai-fu, Zhong Zi-yang, Zheng Chao-de, Li Fang-hua, Image processing in high resolutin eletron microscopy using the direct method I. Phase extension, Acta Crystallogr. A41 (1985) 163–165.

[9] Z.H. Wan, Y.D. Liu, Z.Q. Fu, Y. Li, T.Z. Cheng, F.H. Li, F. Fan, Visual computing in electron crystallography, Z. Kristallogr. 218 (2003) 308–315.

[10] H.F. Fan, The use of sign relationship in the determination of heavy-atom-containing crystal structures II. Component relationship and its applications, Chin. Phys. (1965) 1429–1435.

[11] H.F. Fan, Y.X. Gu, Combining direct methods with isomorphous replacement or anomalous scattering data III. The incorporation of partial structure information, Acta Crystallogr. A41 (1985) 280–284.

[12] Q. Hao, Y.X. Gu, C.D. Zheng, H.F. Fan, OASIS—a computer program for breaking the phase ambiguity in one-wavelength anomalous scattering or single isomorphous substitution (replacement) data, J. Appl. Crystallogr. 33 (2000) 980–981.

[13] J.W. Wang, J.R. Chen, Y.X. Gu, C.D. Zheng, F. Jiang, H.F. Fan, Optimizing the error term in direct-method SAD phasing, Acta Crystallogr. D60 (2004) 1987–1990.

[14] J.W. Wang, J.R. Chen, Y.X. Gu, C.D. Zheng, H.F. Fan, Direct-method SAD phasing with partial-structure iteration—towards automation, Acta Crystallogr. D60 (2004) 1991–1996.

[15] J.W. Wang, Y.X. Gu, C.D. Zheng, H.F. Fan, OASIS-2004—a Direct-method Program for SAD/SIR Phasing and Reciprocal-space Fragment Extension, Institute of Physics, Beijing, 2005 (available at http://cryst.iphy.ac.cn).

[16] T.R. Schneider, G.M. Sheldrick, Substructure solution with SHELXD, Acta Crystallogr. D58 (2002) 1772–1779.

[17] T.C. Terwilliger, J. Berendzen, Automated MAD and MIR structure solution, Acta Crystallogr. D55 (1999) 849–861.

[18] K. Cowtan, An automated procedure for phase improvement by density modification, Joint CCP4 ESF-EACBM Newsl. Protein Crystallogr. 31 (1994) 34–38.

[19] T.C. Terwilliger, Automated main-chain model building by template matching and iterative fragment extension, Acta Crystallogr. D59 (2003) 38–44.

[20] T.C. Terwilliger, Automated side-chain model building and sequence assignment by template matching, Acta Crystallogr. D59 (2003) 45–49.

[21] A. Perrakis, R.J. Morris, V.S. Lamzin, Automated protein model building combined with iterative structure refinement, Nat. Struct. Biol. 6 (1999) 458–463.