(Published in Phys. Status Solidi A **207**, 2621-2638 (2010))

Direct methods beyond small-molecule crystallography

Hai-fu Fan *

Beijing National Laboratory for Condensed Matter Physics and Key Laboratory of Soft Matter Physics, Institute of Physics, CAS, Beijing 100190 P. R. China

Keywords direct methods, electron microscopy, incommensurate structures, protein crystallography

PACS 61.05.C-, 68.37.-d, 61.44.Fw, 87.64.Bx

---

* Corresponding author: e-mail fan@aphy.iphy.ac.cn, Phone: +86-10-82649083, Fax: +86-10-82649531

The development of direct methods outside their traditional field began since the middle 1960's. New applications were explored gradually in four directions. They are the transition of: 1) from single crystals to powder samples, 2) from X-ray crystallography to electron microscopy, 3) from periodic structures to incommensurate structures and 4) from small molecules to macromolecular structures. The research on methods of solving crystal structures at the Institute of Physics in Beijing has been focused on the last three directions. The transition of direct methods from X-ray crystallography to electron microscopy led to a two-step image processing technique in high-resolution electron microscopy. This technique combines information of electron microscopy and that of electron diffraction revealing the structure of minute crystals that are unsuitable for X-ray diffraction analysis. The transition of direct methods from periodic crystal structures to incommensurate structures led to the multidimensional direct methods enabling *ab initio* solution of incommensurate modulated structures and composite structures without relying on an assumed modulation model or even the known basic structure. Finally the combination of direct methods with traditional protein crystallographic methods has greatly enhanced the ability of solving protein structures.

1 Introduction Direct methods in crystallography are that kind of methods, which derive diffraction phases *directly* from experimentally measured diffraction amplitudes. This is why they have played the central role in diffraction analysis of solving crystal structures. In the talk given at the Symposium in Honor of Jerome Karle and Herbert Hauptman at the XIVth IUCr Congress in 1987 [1] Fan suggested that, after the Nobel Prize awarded to Karle and Hauptman, direct methods should turn to exploring new applications outside their traditional field. For this purpose, there are four ways to go: 1) from single crystals to powder samples, 2) from X-ray crystallography to electron microscopy, 3) from periodic structures to incommensurate structures and 4) from small molecules to macromolecular structures. The present paper will focus on works belonging to the last three categories done at the Institute of Physics in Beijing.

2 Direct methods in electron microscopy

2.1 Direct methods as a tool of image processing in high-resolution electron microscopy High-resolution electron microscopy (HREM) is an important complement to X-ray crystallography in crystal structure analysis. Firstly, many crystalline materials important in science and technology are too small in grain size and imperfect in periodicity for carrying out an X-ray diffraction analysis, but are suitable for electron microscopic observations. Secondly the atomic scattering factors for electrons differ greatly from those for X-rays that it is easier for electron diffraction to observe light atoms in the presence of much heavier atoms. Finally the electron microscope is the only instrument that can produce in atomic scale for crystalline samples simultaneously a microscopy image and the corresponding diffraction pattern. The resolution cutoff of the latter is usually much higher than that of the former. In principle either the electron microscopy (EM) image or the electron diffraction (ED) pattern could lead to a structure image of the sample. However the combination of the two would lead to results better than that from either of the two alone. On the other hand, HREM suffers from two difficulties: Firstly, an EM image is not a true structural 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 EM images is not enough to reveal individual atoms in most cases. Hence it is essential to have some image processing technique for image deconvolution and for resolution enhancement. Direct methods in X-ray crystallography may be considered as a special kind of image processing techniques. Three-dimensional X-ray diffraction patterns without phase information may be regarded as the Fourier transform of a blurred image of crystal structures. Such a blurred image is nothing but the Patterson function – the Fourier syntheses with squares of structure-factor amplitudes as coefficients. Now, a set of diffraction intensities after processing by a direct method in reciprocal space can be converted to a set of structure factors, i.e. the Fourier transform of the structure image. Such a process may be considered as a king of image deblurring, which cleans up the Patterson map to give an image of the true structure. It is expected that direct methods can act as a tool of image processing in HREM.

2.2 Two-step direct-method aided image processing The procedure consists of image deconvolution (step 1) and resolution enhancement (step 2). The purpose of step 1 is to retrieve the structure image from a single EM image by making use of direct methods. The resultant image will be at the same resolution as that of the starting EM image. The purpose of step 2 is to enhance the resolution of the structure image obtained from step 1 by the direct-method phase extension with starting phases from the deconvoluted EM image and structure amplitudes from the corresponding ED pattern. The resultant image will be at the same resolution as that of the ED pattern.

Different procedures have been proposed for the deconvolution of EM images. Most of them use a series of EM images with different defocus values. Uyeda and Ishizuka in 1975 [2] first proposed a method for the deconvolution of a single EM image under the weak-phase-object approximation (WPOA). In their method, a series trial defocus values were used to calculate a series trial “structure” factor sets from a single EM image. Then a series Fourier transforms are calculated. The Fourier transform that contains less anomalous peaks and valleys was then taken as the “true” structure image. Direct methods in X-ray

crystallography were introduced for the image deconvolution of a single EM image since 1979 [3-5]. This method uses also a series trial defocus values. However the discrimination criterion is different. With the WPOA, in which dynamical diffraction effects are neglected, the Fourier transform of an EM image can be expressed as

$$T(\mathbf{h}) = \delta(h) + 2\sigma F(\mathbf{h}) \sin \chi_1(h) \exp\left[-\chi_2(h)\right], \quad (1)$$

which can be rearranged as

$$F(\mathbf{h})_{h \neq 0} = T(\mathbf{h}) / 2\sigma \sin \chi_1(h) \exp\left[-\chi_2(h)\right]. \quad (2)$$

Here $\sigma = \pi/\lambda U$, λ is the electron wavelength and U the accelerating voltage. **h** is the reciprocal-lattice vector within the resolution limit. F(**h**) is the structure factor of electron diffraction, which is the Fourier transform of the potential distribution φ(**r**) of the object, and $\sin \chi_1(h) \exp\left[-\chi_2(h)\right]$ is the contrast transfer function, in which

$$\chi_1(h) = \pi \Delta f \lambda h^2 + \tfrac{1}{2}\left(\pi C_s \lambda^3 h^4\right),$$

and

$$\chi_2(h) = \tfrac{1}{2}\left(\pi^2 \lambda^2 h^4 D^2\right).$$

Here Δf is the defocus value, C_S is the spherical aberration coefficient, and D is the standard deviation of the Gaussian distribution of defocus due to the chromatic aberration [6]. Values of Δf, C_S, and D should be found before the image can be deconvoluted. Of these three factors, C_S and D can be determined experimentally without much difficulty. Furthermore, in contrast to Δf, C_S and D do not change much from one image to another. This means that the main problem is to estimate Δf. Suppose C_S and D are known, a set of F(**h**) can be calculated from Equation (2) for a given Δf. If the Δf value is correct, then the corresponding set of F(**h**) should obey the Sayre equation [7]

$$F(\mathbf{h}) = \frac{\theta}{V} \sum_{\mathbf{h}'} F(\mathbf{h}') F(\mathbf{h} - \mathbf{h}'), \quad (3)$$

where θ is the atomic form factor and V is the volume of the unit cell. Hence the true Δf can be found by a systematic change of the trial Δf so as to improve the fitness to the Sayre equation. For details the reader is referred to the original paper [4].

Direct-method phase extension was introduced for the resolution enhancement of EM images in 1985 [8]. Usually ED patterns contain observable reflections at a resolution of

about 1Å or higher. In addition, the intensities of ED patterns from a crystalline specimen are independent of defocus and spherical aberration of the electron-optical system. Accordingly, under the WPOA, better estimation of structure-factor amplitudes can be obtained from ED patterns rather than from EM images. However, structure analysis based on ED patterns alone is subject to the phase problem. On the other hand, EM images after deconvolution can provide phase information at about 2Å resolution. This is particularly beneficial to resolving the phase problem. Thus an enhanced high-resolution image can be obtained by a phase extension procedure using the structure-factors amplitudes from ED and starting phases from the corresponding EM image. The tangent formula [9] is used as the basis of the phase extension:

$$\tan\alpha_{\mathbf{h}} = \frac{\sum_{\mathbf{h}'} |E_{\mathbf{h}'}E_{\mathbf{h}-\mathbf{h}'}| \sin\left(\alpha_{\mathbf{h}'} + \alpha_{\mathbf{h}-\mathbf{h}'}\right)}{\sum_{\mathbf{h}'} |E_{\mathbf{h}'}E_{\mathbf{h}-\mathbf{h}'}| \cos\left(\alpha_{\mathbf{h}'} + \alpha_{\mathbf{h}-\mathbf{h}'}\right)}, \tag{4}$$

The reader is referred to the original paper [8] for more details.

Simulating calculations have verified the advantages of combining EM and ED [4, 8]. Results are summarized in Fig. 1. The sample is chlorinated copper phthalocyanine ($C_{32}N_8Cl_{16}Cu$), which crystallizes in the space group C2/c with unit cell parameters a = 19.62, b = 26.08, c = 3.76Å; β = 116.5 °. The crystal structure of the sample has been study with EM in 1972 [10].

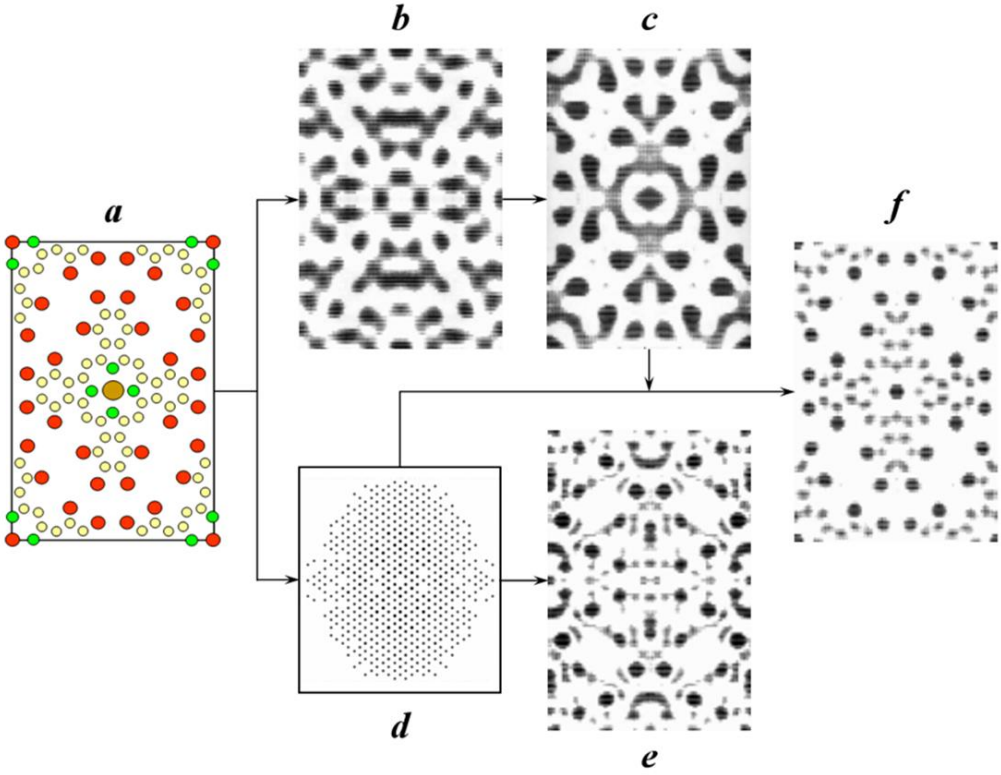

Figure 1 Simulating results of combing EM and ED in solving the crystal structure of chlorinated copper phthalocyanine. *a*): the object schematically denoted as a single unit cell of crystalline chlorinated copper phthalocyanine projected down the ***c*** axis. The copper atom is in orange, chlorine atoms in red, nitrogen atoms in green and carbon atoms in yellow. *b*): the theoretical EM image of *a*) at 2Å resolution taken with the conditions: accelerating voltage = 500 *kV*, *Cs* = 1 mm, *D* = 15nm and Δf = −100nm. *c*): the direct-method image deconvolution result of *b*). *d*): the schematic ED pattern of *a*) with F(0,0,0) removed. *e*): the direct-method phasing result (the best E-map) of *d*). *f*): the result of two-step image processing combining EM and ED.

It is seen in Fig. 1 that even in the case of kinematical diffraction, the EM image does not necessary directly reflect structure details of the object. A direct-method image deconvolution led *b*) to become *c*), where the shape of the molecule can be recognized but individual atoms other than copper are not visible. With the structure-factor amplitudes at 1Å resolution extracted from *d*), an *ab initio* direct method phasing led to the best E-map *e*). Although all chlorine atoms are on the map, numerous ghost peaks make the E-map impossible to interpret. On the other hand, with starting phases at 2Å resolution from the Fourier transform of *c*) and amplitudes at 1Å resolution extracted from *d*), the direct-method phase extension resulted in *f*) reveals clearly all individual atoms in the object *a*). While the advantage of combining EM and ED is evident, it does not mean that to solve the structure from an ED alone is impossible. Dorset *et al.* in 1991 [11] reported the determination of the structure of chlorinated copper phthalocyanine. They started from an experimental ED pattern at 0.91Å resolution obtained with the accelerating voltage of 1,200 kV. Direct-methods phasing of the ED revealed all the heavy atoms. The Fourier refinement based on which led to the complete structure. Test of the two-step direct-method aided image processing using an experimental EM image and the corresponding ED pattern of chlorinated copper phthalocyanine was reported in 1991 [12]. Results are reproduced by the program VEC (to be described in next section) and summarized in Fig. 2. The first application of the two-step image processing technique to an unknown sample was reported by Fu *et al.* in 1994 [13]. Results are reproduced by the program VEC and summarized in Fig. 3. Fan and Zheng in 1975 [14] have shown that it is possible to extrapolate not only the phases but also the amplitudes of structure factors beyond the resolution limit of the experimental data. An example of the structure-factor extrapolation from the theoretical EM image of chlorinated copper phthalocyanine at 2Å resolution to the structure image at 1Å resolution has been reported by Liu *et al.* in 1988 [15]. Applications to real EM images still have to be explored. However, similar procedures have been proved workable with experimental three-dimensional protein diffraction data by Wang and colleagues in 1985 [16] and by Giacovazzo and colleagues in 2005 [17].

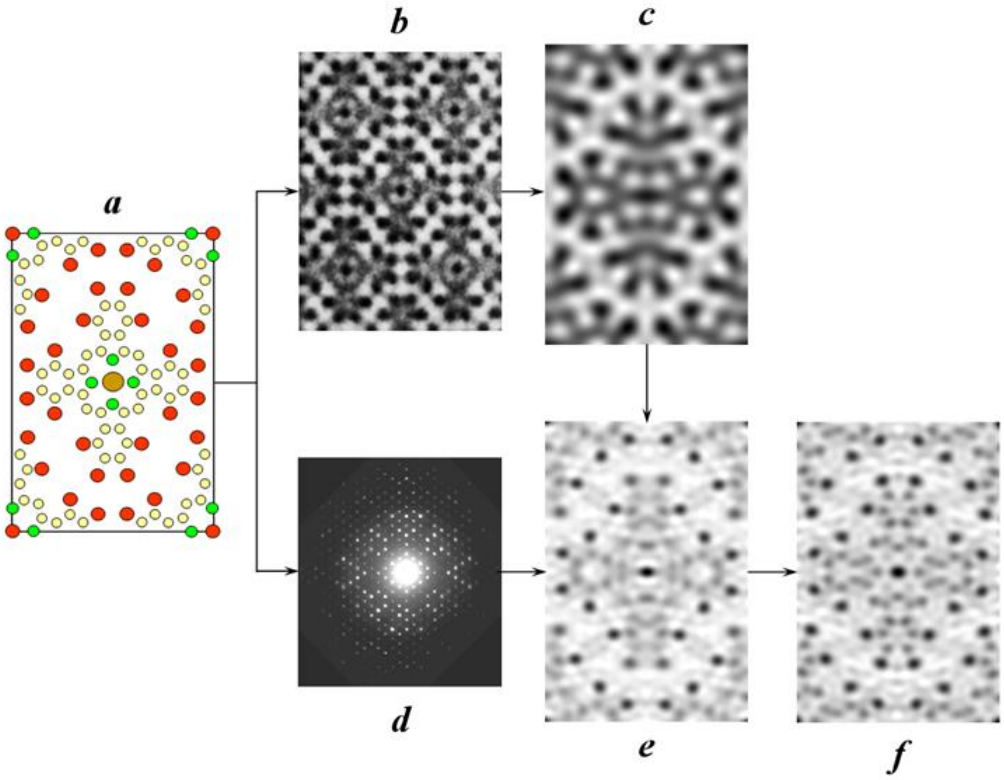

Figure 2 Results of two-step image processing applied to an experimental EM image and the corresponding ED pattern of chlorinated copper phthalocyanine. *a*): the object schematically denoted as a unit cell of crystalline chlorinated copper phthalocyanine projected down the ***c*** axis. *b*): the experimental EM image of the object at 2Å resolution taken near the optimum defocus value. *c*): the direct-method deconvoluted image of *b*). *d*): the experimental ED pattern of *a*). *e*): the potential-distribution projection of the object at 1Å resolution obtained from the direct-method phase extension with starting phases from the Fourier transform of *c*) and amplitudes extracted from *d*). *f*): the potential-distribution projection after 4 cycles of Fourier refinement starting from *e*). As is seen all the atoms in the structure model *a*) are clearly revealed in the resultant map *f*).

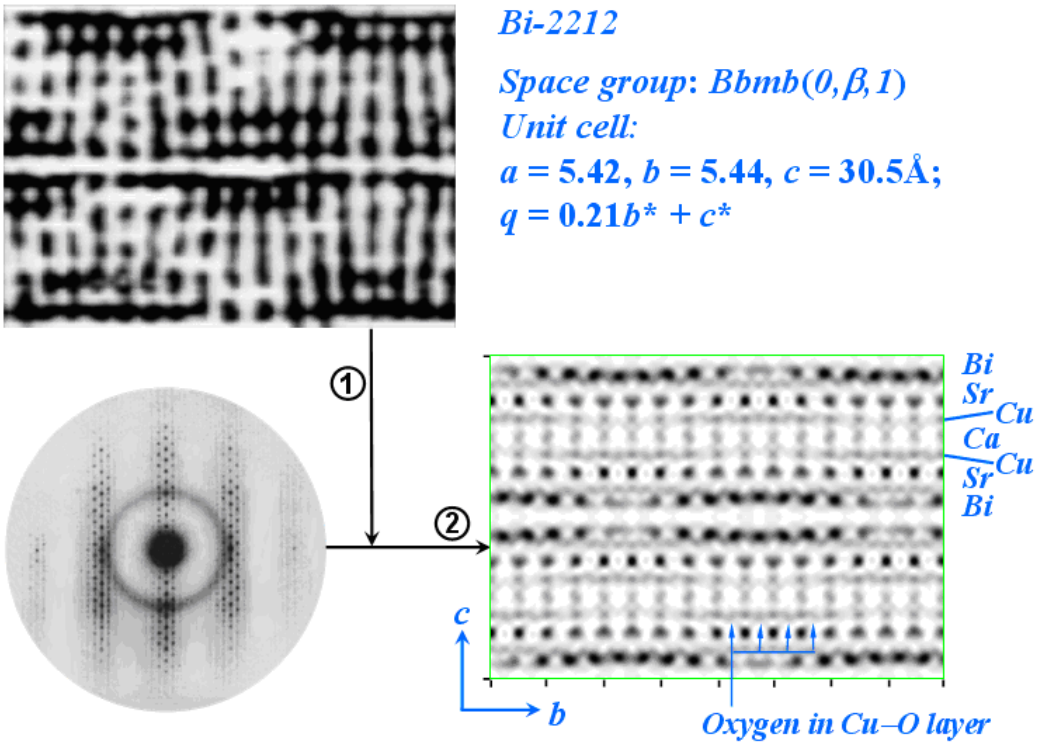

Figure 3 The incommensurate modulation of the high-Tc superconductor Bi-2212 studied by the two-step HREM image processing technique. Upper-right: summary of crystallographic parameters of the sample. Upper-left: the EM image of the sample at 2Å resolution taken near the optimum defocus value. Lower-left: the corresponding ED pattern at 1Å resolution. Lower-right: the resultant structure image at 1Å resolution. Since the sample is an incommensurate modulated crystal, the direct method used is some what different from that for conventional crystals. It should be notice here that, in step 1 only phases of main reflection were derived, while in step two the phase extension was first applied to main reflections and then to satellites. This will be discussed in detail later in this paper. As is seen from the figure, the EM image at 2Å resolution does not show much structure details. However the resultant image at 1Å resolution shows all metal atoms and their modulation clearly. Even more, some of the oxygen atoms on Cu-O layers are also visible. The EM image of the superconductor Bi-2212 was kindly made available by Drs. Y. Matsui and S. Horiuchi.

2.3 The program VEC (Visual computing in Electron Crystallography) is a computer program [18, 19] written in FORTRAN and C++ to visualize and automate calculations involving in the two-step HREM image processing. The program can be run under the Microsoft Windows operating system (from Windows 95 to Windows 7). Functions of VEC include:

- Preliminary processing of EM images
- Indexing and intensity extracting of ED patterns
- Image deconvolution via searching the defocus value from a single EM image

- Image resolution enhancement by direct-method phase extension combining EM and ED
- Simulation of dynamical/kinematical ED patterns and EM images for conventional and modulated crystals
- 2-, 3- and 4-dimensional FFT
- 2-dimensional half-tone-graph display of 2-, 3- and 4-dimensional Fourier maps
- Contour mapping of 2-dimensional patterns
- *Ab-initio* direct-method solution of incommensurate one-dimensionally modulated structures and composite structures of two subsystems using electron or X-ray diffraction data

The package is freely available for academic users at the website:
http://cryst.iphy.ac.cn.

3 Direct methods in multidimensional space - dealing with incommensurate structures Crystal structures are usually assumed to be ideal 3-dimensional periodic objects. Since real crystals are never perfect, what we obtained in diffraction analysis under this assumption is just an averaged image of the real structure over a large number of unit cells. However knowledge on the average structure is not enough for understanding the properties of many solid state materials. Therefore an important task for methods of solving crystal structures is to extend their scope to include imperfect crystals. Modulated crystal structures belong to that kind of crystal structures, in which atoms suffer from fluctuations in their position and/or composition. If the period of fluctuation matches that of the three-dimensional unit cell, a superstructure results; otherwise an incommensurate modulated structure is obtained. Incommensurate modulated phases can be found in many important solid state materials. In many cases, the transition to the incommensurate modulated structure corresponds to a change of certain physical properties. Hence it is important to know the structure of incommensurate modulated phases in order to understand the mechanism of the transition and properties in the modulated state. Many incommensurate modulated structures were solved by trial-and-error methods. With these methods it is necessary to make assumption on the modulation in advance. This leads often to difficulties and erroneous results. In view of diffraction analysis, it is possible to derive the phase of reflections directly and solve the structure objectively without relying on any assumption of the modulation. A common feature of incommensurate modulated structures is that they do not have 3-dimensional periodicity. However incommensurate modulated structures can be regarded as the 3-dimensional hyper section of a 4- or higher-dimensional periodic structure. Obviously direct analysis of incommensurate modulated structures would better be implemented in multidimensional space. For this purpose we need firstly a theory on multidimensional symmetry and secondly a method to solve directly the multidimensional phase problem. Works by Janner and co-workers [20] were dealing with the first problem, while the multidimensional direct methods developed in our group are aiming at the second.

3.1 Multidimensional representation of incommensurate structures Incommensurate modulated crystals yield 3-dimensional diffraction patterns containing satellites round the main reflections. An example of a 2-dimensioanl section of such a 3-dimensional diffraction pattern is shown schematically in Fig. 4. The main reflections are

consistent with a regular 3-dimensional reciprocal lattice but the satellites do not fit the same lattice. On the other hand, the arrangement of satellite reflections has its own periodicity (see the vertical line segments in Fig. 4). Hence, it can be imagined that the 3-dimensional diffraction pattern is a projection of a 4-dimensional reciprocal lattice, in which the main and the satellite reflections are all regularly situated at the lattice nodes. According to the properties of Fourier transform, the incommensurate modulated structure here considered can be regarded as a 3-dimensional "section" of a 4-dimensional periodic structure. The above example corresponds to a one-dimensional modulation. For an n-dimensional (n=1, 2, …) modulation, we need a (3+n)-dimensional description.

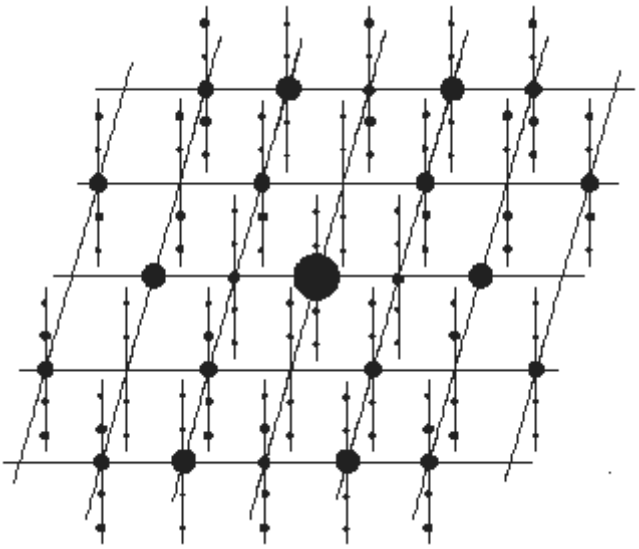

Figure 4 Schematic diffraction pattern of an incommensurate modulated structure. The vertical line segments indicate projected lattice lines parallel to the fourth dimension.

Composite structures are a special kind of incommensurate modulated structures. The characteristic of which is the coexistence of two or more mutually incommensurate 3-dimensional lattices. Owing to the interaction of coexisting lattices, composite structures are also incommensurate modulated structures. Unlike ordinary incommensurate modulated structures, composite structures do not have a 3-dimensional basic structure. The basic structure of a composite structure corresponds to a 4- or higher-dimensional periodic structure. For a detailed description of incommensurate structures the reader is referred to the reference [20].

3.2 Multidimensional direct methods It has been proved by Hao *et al.* in 1987 [21] that the Sayre equation [7] can easily be extended into multidimensional space. We have

$$F(\mathbf{h}) = \frac{\theta}{V}\sum_{\mathbf{h}'} F(\mathbf{h}')F(\mathbf{h}-\mathbf{h}')\,. \qquad (5)$$

Now **h** is a multidimensional reciprocal lattice vector defined as

$$\mathbf{h} = \sum_{i=1}^{3+n} h_i \mathbf{b}_i \quad , \quad n = 1, 2, 3 \ldots . \tag{6}$$

The vector in the corresponding real-space multidimensional unit cell will then be

$$\mathbf{x} = \sum_{i=1}^{3+n} x_i \mathbf{a}_i \quad , \quad n = 1, 2, 3 \ldots \quad . \tag{7}$$

Where vectors $\mathbf{a}_i$ and $\mathbf{b}_j$ are respectively basic vectors of the real and the reciprocal lattice, which satisfy the reciprocal relationship

$$\mathbf{a}_i \cdot \mathbf{b}_j = \delta_{ij} \ , \quad (i, j = 1,2,3,4,\cdots,3+n) . \tag{8}$$

In the case of one-dimensionally incommensurate modulation, we have

$$\begin{pmatrix} \mathbf{a}_1 = \mathbf{a} - q_1\mathbf{d} \\ \mathbf{a}_2 = \mathbf{b} - q_2\mathbf{d} \\ \mathbf{a}_3 = \mathbf{c} - q_3\mathbf{d} \\ \mathbf{a}_4 = (0, \mathbf{d}) \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \mathbf{b}_1 = (\mathbf{a}^*, 0) \\ \mathbf{b}_2 = (\mathbf{b}^*, 0) \\ \mathbf{b}_3 = (\mathbf{c}^*, 0) \\ \mathbf{b}_4 = (\mathbf{q}, \mathbf{d}) \end{pmatrix}, \tag{9}$$

where **q** is the modulation vector:

$$\mathbf{q} = q_1\mathbf{a}^* + q_2\mathbf{b}^* + q_3\mathbf{c}^* . \tag{10}$$

More detailed description of the one-dimensionally incommensurate modulation can be found in reference [22]. The structure factor of an incommensurate modulated structure is expressed as (for details see [21])

$$F(\mathbf{h}) = \sum_{j=1}^{N} f_j(\mathbf{h}) \exp\left[i2\pi(h_1\bar{x}_{j1} + h_2\bar{x}_{j2} + h_3\bar{x}_{j3}) \right], \tag{11}$$

where $f_j(\mathbf{h})$ is the atomic scattering factor of the j^{th} modulated atom in the unit cell

$$f_j(\mathbf{h}) = f_j^o(h) \int_0^1 d\bar{x}_4 \ldots \int_0^1 d\bar{x}_{3+n} P_j(\bar{x}_4, \ldots, \bar{x}_{3+n}) \times$$
$$\exp\left\{ i2\pi\left[\left(h_1U_{j1} + h_2U_{j2} + h_3U_{j3} \right) + (h_4x_{j4} + \ldots + h_{3+n}x_{j(3+n)}) \right] \right\}. \tag{12}$$

$f_j^o(h)$ on the right-hand side of Eq. (12) is the ordinary atomic scattering factor, P_j is the occupational modulation function and U_j describes the deviation of the j^{th} atom from its average position $(\overline{x}_{j1}, \overline{x}_{j2}, \overline{x}_{j3})$. According to Eqs. (11) and (12) a modulated structure can be regarded as a set of 'modulated atoms' situated at their average positions in 3-dimensional space. The right-hand side of the multidimensional Sayre equation [Eq. (5)] can be spit into three parts. It follows

$$F(\mathbf{h}) = \frac{\theta}{V}\left\{\sum_{\mathbf{h'}} F_m(\mathbf{h'})F_m(\mathbf{h}-\mathbf{h'}) + 2\sum_{\mathbf{h'}} F_m(\mathbf{h'})F_s(\mathbf{h}-\mathbf{h'}) + \sum_{\mathbf{h'}} F_s(\mathbf{h'})F_s(\mathbf{h}-\mathbf{h'})\right\} . \tag{13}$$

The subscript *m* in Eq. (13) denotes main reflections, while *s* denotes satellites. The last summation on the right-hand side of Eq. (13) is negligible in comparison with either of the first two. Hence Eq. (13) becomes

$$F(\mathbf{h}) \approx \frac{\theta}{V}\sum_{\mathbf{h'}} F_m(\mathbf{h'})F_m(\mathbf{h}-\mathbf{h'}) + \frac{2\theta}{V}\sum_{\mathbf{h'}} F_m(\mathbf{h'})F_s(\mathbf{h}-\mathbf{h'}) . \tag{14}$$

In case that the reciprocal vector **h** points at a main reflection, the second summation on the right-hand side of Eq. (14) will be negligible in comparison with the first. Thus we have

$$F_m(\mathbf{h}) \approx \frac{\theta}{V}\sum_{\mathbf{h'}} F_m(\mathbf{h'})F_m(\mathbf{h}-\mathbf{h'}) . \tag{15}$$

In case that the reciprocal vector **h** points at a satellite reflection, for ordinary incommensurate modulated structures the first summation on the right-hand side of Eq. (14) has vanished, because any three-dimensional reciprocal lattice vector corresponding to a main reflection will have zero components in the extra dimensions so that the sum of two such lattice vectors could never give rise to a lattice vector corresponding to a satellite. We then have

$$F_{\mathrm{s}}(\mathbf{h}) \approx \frac{2\theta}{V}\sum_{\mathbf{h'}} F_{\mathrm{m}}(\mathbf{h'})F_{\mathrm{s}}(\mathbf{h}-\mathbf{h'}) . \tag{16}$$

For composite structures on the other hand, since the average structure itself is a 4- or higher-dimensional periodic structure, the first summation on the right-hand side of Eq. (14) does not vanish. We have instead of (16) the following equation [23]

$$F_{\mathrm{s}}(\mathbf{h}) \approx \frac{\theta}{V}\sum_{\mathbf{h'}} F_{\mathrm{m}}(\mathbf{h'})F_{m}(\mathbf{h}-\mathbf{h'})\,. \qquad (17)$$

Eq. (15) indicates that the phases of main reflections can be derived by a conventional direct method neglecting the satellites. Eq. (16) or (17) can be used to extend phases from main reflections to satellites respectively for ordinary incommensurate modulated structures or composite structures. This provides a way to determine the modulation functions objectively. The procedure will be in the following stages:

i) Derive phases of main reflections using Eq. (15);
ii) Derive phases of satellite reflections using Eq. (16) or (17);
iii) Calculate the multidimensional Fourier map using observed structure factor amplitudes and phases from i) and ii);
iv) Cut the Fourier map with a 3-dimensional 'hyperplane' to obtain the incommensurate modulated structure in the physical space;
v) Parameters of the modulation functions are measured directly on the multidimensional Fourier map resulting from iii).

The program DIMS described in the next section has been written for implementation of the above procedure.

3.3 The program DIMS (Direct methods In Multidimensional Space or Direct methods for Incommensurate Modulated Structures) is a computer program for solving one-dimensionally modulated incommensurate structures and composite structures consists of two sub-systems. For incommensurate modulated structures, the program calculates E-values independent of atomic scattering factors; hence it can accept diffraction data from X-rays, electrons or neutrons and treat them in the same way. However, for composite structures DIMS can deal with only X-ray diffraction data. There are two different versions of DIMS. Both are freely available for academic users from the website: *http://cryst.iphy.ac.cn*.

- Stand-alone version [24, 25]:
  - The program is written in FORTRAN. In theory, it can be run in any operating system provided appropriate FORTRAN compiler is available.
  - The program accepts two-line symbols of superspace groups (one-line symbols can be converted to two-line symbols by the program *symbol1to2* described in the Tutorial) and derives accordingly symmetry generators. Alternatively, it allows also manual input of symmetry generators.
  - Incommensurate modulated/composite structures are solved by direct methods in two stages. In the first stage, phases of main reflections are derived using a conventional direct-method, while in the second stage, phases of satellite reflections are derived by a multi-dimensional direct method with starting phases of main reflections.
  - In case the basic/average structure is known, phases of main reflections calculated from which can be input to DIMS and the first stage can be skipped.
- MS Windows version [18,19,26,27]:

- The Windows version of DIMS forms a part of the VEC package.
- I/O data of DIMS can be displayed and manipulated by the sophisticated graphic interface of VEC. 4D-Fourier maps phased by DIMS can be calculated, displayed and interpreted within VEC. The subroutine MIMS [22] can help with building a 4D-structure model from the 4D-Fourier map.

3.4 Applications

3.4.1 Incommensurate modulation of Bi-2223 from electron diffraction data

Crystals of the high-Tc superconductor Pb-doped $Bi_2Sr_2Ca_2Cu_3O_y$ (Bi-2223) crystallise in the space group Bbmb($0\beta0$) with unit cell parameters of the basic structure a = 5.49, b = 5.41, c = 37.1Å and the modulation wave vector $\boldsymbol{q}$ = 0.117$\boldsymbol{b}$*. Mo *et al.* [28] studied the incommensurate modulation with *0klm* electron diffraction data. Results are reproduced here using the program DIMS/VEC [26, 27]. Phases of main reflections *0kl0* were calculated according to the known basic structure. Phases of satellite reflections *0klm* ($m \neq 0$) were then derived by DIMS based on the known phases of main reflections. The potential-distribution map $\int_0^1 \varphi(x_1, 0, x_3, x_4)\,dx_1$ was then calculated, which is shown on the right of Fig. 5. Parameters of modulation waves of all metal atoms were measured directly from this map. The least-square refinement based on the measured parameters and the known basic structure ended at the R-factor of 0.16 for the main and 0.17 for the first-order satellite

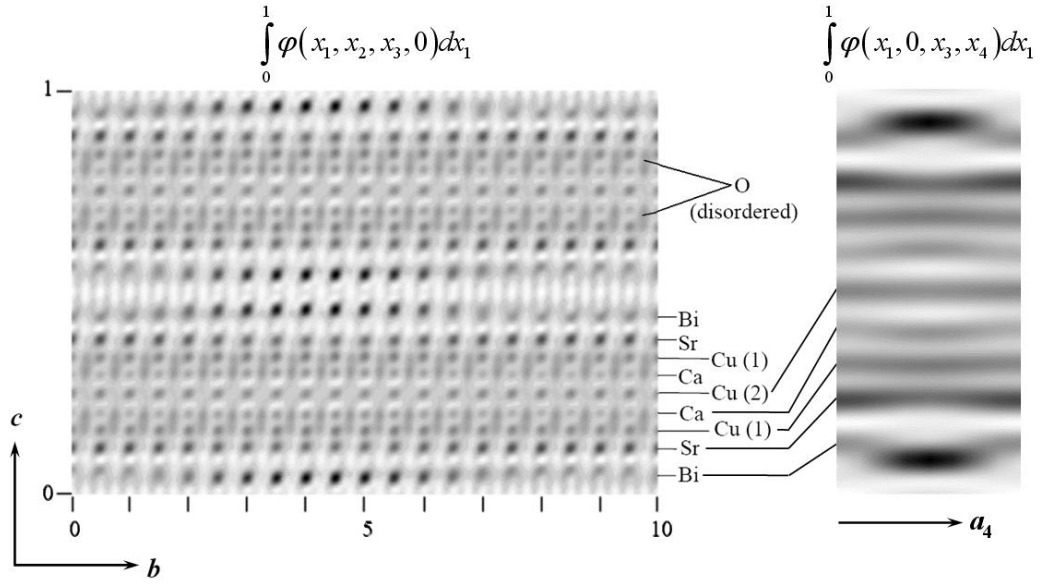

Figure 5 Sections of the 4-dimensional potential distribution $\varphi(x_1, x_2, x_3, x_4)$ of Bi-2223 projected along the $\boldsymbol{a}_1$ axis.

reflections. Phases resulting from the least-square refinement were used to calculate the map shown on the left of Fig. 5, which is actually the potential distribution in 3-dimensional physical space projected along the $\boldsymbol{a}$ axis showing the incommensurate modulation across ten unit cells of the basic structure along the $\boldsymbol{b}$ axis. As is seen, both occupational and positional

modulations are evident for most metal atoms. Another prominent feature is that the oxygen atoms on the Cu-O layers move towards the Ca layer forming a disordered oxygen bridge across the layers of Cu(2)-Ca-Cu(1)-Ca-Cu(2).

3.4.2 Incommensurate modulation of Bi-2212 from X-ray data The incommensurate structure of the high-Tc superconductor $Bi_2Sr_2CaCu_2O_y$ (Bi-2212) has been extensively studied in a number of laboratories worldwide. However the results are not completely consistent with each other. Fu *et al.* in 1995 [29] reported the incommensurate modulation of Bi-2212 studied using the multidimensional direct method. Single-crystal X-ray diffraction data were used. No preliminary assumptions were made on the modulation. Results are reproduced here by the program DIMS/VEC [26, 27]. Crystals of Bi-2212 belong to the superspace group Bbmb($0\beta1$) with unit cell parameters of the basic structure a = 5.422, b = 5.437, c = 30.537Å and the modulation wave vector $\boldsymbol{q}$ = 0.22$\boldsymbol{b}$* + $\boldsymbol{c}$*. Phases of main reflections were derived by a conventional direct method. The program DIMS was used to extend phases from main reflections to satellites. Electron-density maps (Figs. 6, 7 and 8) were then calculated on the VEC platform. The 4-dimensional electron-density function of Bi-2212 cut at x_4 = 0 and projected along the $\boldsymbol{x}_1$ axis is shown on the right of Fig. 6 giving an overview of the incommensurate structure covering 6 unit cells of the basic structure along the $\boldsymbol{x}_2$ axis. All metal atoms and the oxygen atoms on Cu-O layers are clearly seen. The section $\rho(x_1, ½, x_3, 0)$ shown in the middle of Fig. 6 contains all independent metal atoms.

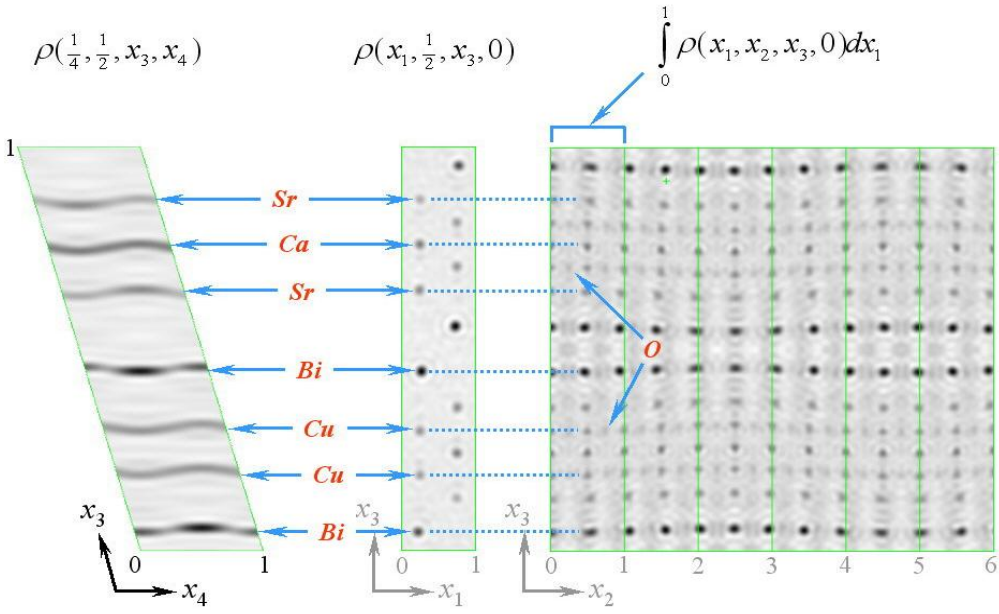

Figure 6 2-dimensional maps from the 4-dimensional electron-density function of Bi-2212 produced by the program DIMS/VEC.

Their modulations are shown on the section $\rho(¼, ½, x_3, x_4)$ at the left. Fig. 7 shows atoms on the Cu-O layer and the modulation of the oxygen atom O(1). Fig. 8 shows the sawtooth modulation of the oxygen atom O(4), which had been a difficult point in solving the incommensurate structure of Bi-2212 by other methods. It is seen here that the sawtooth

modulation can be revealed objectively by the multidimensional direct method without any preliminary assumptions on the modulation and before any efforts of model building and structure refinement.

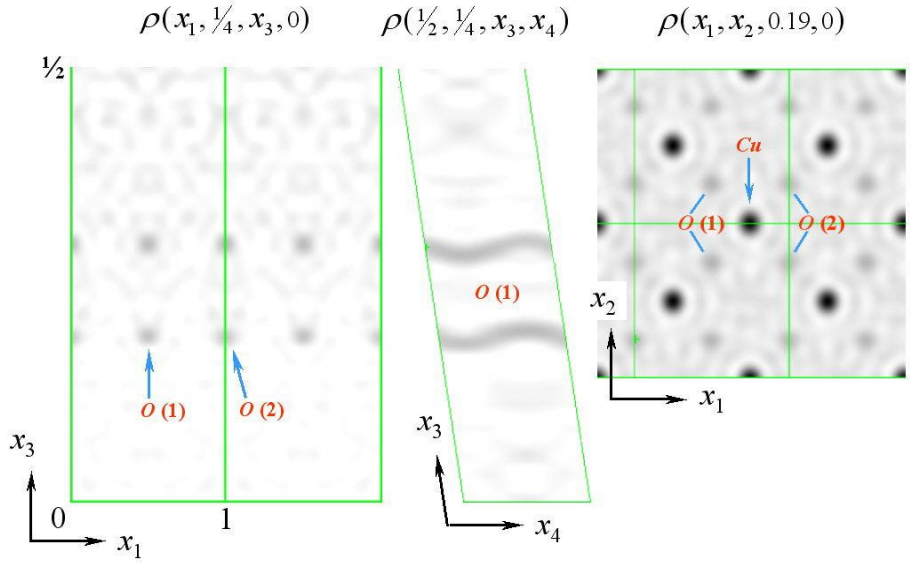

Figure 7 Electron-density sections showing the Cu-O layer and the modulation of oxygen atoms on the layer.

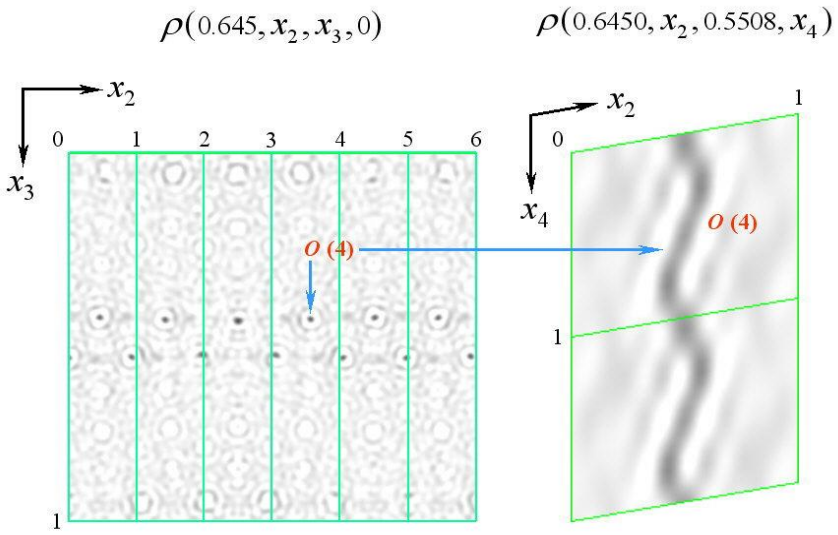

Figure 8 Electron-density sections showing the sawtooth modulation of the oxygen atom O(4).

3.4.3 Solving the 4-dimensional basic structure of the composite crystal $(PbS)_{1.18}TiS_2$ The composite structure of $(PbS)_{1.18}TiS_2$ [30] belongs to the space group C2/m (α00) *s*-1. It consists of two subsystems: the subsystem TiS_2 with $a_1 = 3.409$,

$b_1 = 5.880$, $c_1 = 11.760$Å $\alpha_1 = 95.29^o$ and the subsystem PbS with $a_2 = 5.800$, $b_2 = 5.881$ $c_2 = 11.759$Å, $\alpha_2 = 95.27^o$. Within the experimental error we have $b_1 = b_2$, $c_1 = c_2$, and $\alpha_1 = \alpha_2$. The basic structure of $(PbS)_{1.18}TiS_2$ is a 4-dimensional periodicity structure. It is more difficult to determine such a basic structure than to fine the modulation in $(PbS)_{1.18}TiS_2$, since no phase information is available in advance except that of the origin and

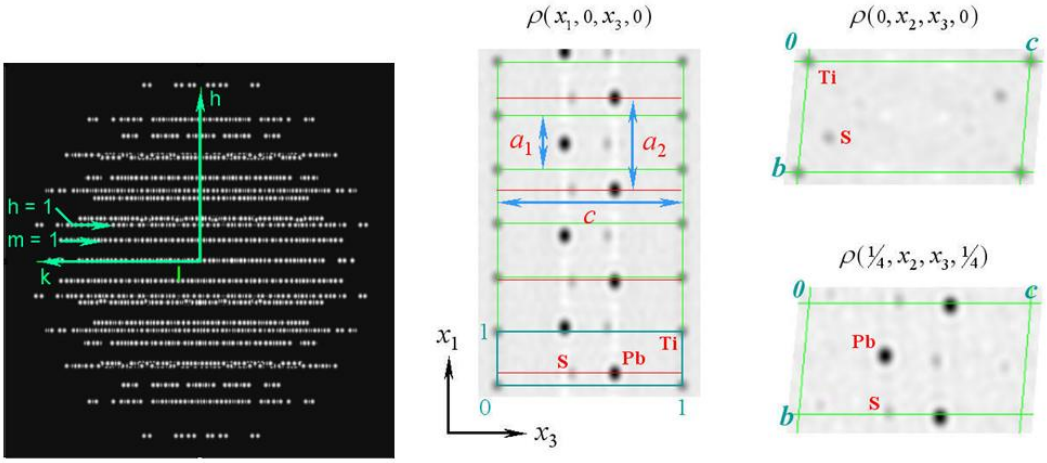

Figure 9 DIMS/VEC applied to studying the basic structure of composite crystals of $(PbS)_{1.18}TiS_2$.

enantiomorph fixing reflections. Mo *et al.* [31] showed that DIMS is capable of solving the 4-dimensional basic structure in a straightforward manner. Results are reproduced by DIMS/VEC [26, 27] here. The input file to DIMS/VEC contains only main reflections. The output file in graphical mode is shown on the left of Fig. 9. 2-dimensional sections of the 4-dimensional electron-density maps calculated from it are on the right. From the middle part of Fig. 9 we see the "chimney and ladder" structure along the $\boldsymbol{x_1}$ axis constructed by the TiS_2 subsystem with the period a_1 and the PbS subsystem with the period a_2. On the right of Fig. 9 there are sections cutting through the TiS_2 layer and PbS layer parallel to the (***b***, ***c***) plane. Note that $\boldsymbol{a_1}$, $\boldsymbol{a_2}$, ***b*** and ***c*** are respectively the projection of the 4-dimensional axes $\boldsymbol{x_1}$, $\boldsymbol{x_4}$, $\boldsymbol{x_2}$ and $\boldsymbol{x_3}$ along the direction perpendicular to the 3-dimensional physical space. As is seen, all the atoms in the composite structure are clearly revealed on the direct-method phased electron-density map without any efforts of model building and structure refinement.

4 Direct methods in macromolecular crystallography Solving the crystal structure of proteins is an indispensable part of the experimental basis for understanding the structure-function relationship of biological macromolecules. While direct methods have been long dominating crystal structure analysis of small molecules, they have little influences in macromolecular crystallography until early 1990's. Fan proposed in 1965 [32] the use of direct methods in breaking the phase ambiguity intrinsic to single anomalous diffraction (SAD) and single isomorphous replacement (SIR) methods. Similar investigations were reported by different authors during 1960's and 1970's [33-38]. From the early 1980's to the

early 2000's, the combination of direct methods with SAD/SIR data was emphasized in direct-method research worldwide [39-63]. The first application of direct-method SAD phasing in solving originally unknown protein structures was reported in 1999 [56] with data at 2.1Å resolution. At present, applications of direct methods in protein crystallography can be divided into three categories:

1. Locating heavy atoms in protein structures;
2. *Ab initio* determination of protein structures (data resolution higher than ~1.2Å required);
3. Combining direct methods with traditional protein crystallographic techniques for breaking the SAD/SIR phase ambiguity and improving model completion.

Methods of the first category do not solve the entire structure of proteins. Methods of the second category require "atomic resolution" data and, only ~5% of the protein diffraction data so far deposited in the Protein Data Bank (*http://www.rcsb.org/*) satisfy this requirement. In contrast, methods of the third category are applicable in most cases. Direct methods developed in the Beijing Institute of Physics belong to the third category.

4.1 Breaking the phase ambiguity intrinsic to SAD/SIR data Owing to the similarity in phasing philosophy between SAD and SIR, the following formulation are given for SAD only. The SAD method is now becoming the first choice of *de novo* protein-structure determination. However there is the phase ambiguity intrinsic to SAD phasing, i.e. the phase of reflections derived from SAD experiments is not unique, but rather a phase doublet, which can be expressed as

$$\varphi_{\mathbf{h}} = \varphi_{\mathbf{h}}'' \pm \left|\Delta\varphi_{\mathbf{h}}\right|. \tag{18}$$

Where $\varphi_{\mathbf{h}}$ is the phase of reflection having the reciprocal vector **h**; $\varphi''_{\mathbf{h}}$ is the phase of $\boldsymbol{F}_{\mathbf{h}}'' = \left|\boldsymbol{F}_{\mathbf{h}}''\right|\exp\left(i\varphi_{\mathbf{h}}''\right) = i\sum_{j=1}^{N}\Delta f_j''\exp(i2\pi\boldsymbol{h}\cdot\boldsymbol{r}_j)$, which is the structure factor contributed from the imaginary-part scattering of the heavy-atom substructure; $|\Delta\varphi_{\mathbf{h}}|$ can be calculated by (see reference [64])

$$\left|\Delta\varphi_{\mathbf{h}}\right| = \left|\cos^{-1}\left(\frac{\left|\boldsymbol{F}_{\mathbf{h}}^{+}\right| - \left|\boldsymbol{F}_{\mathbf{h}}^{-}\right|}{2\left|\boldsymbol{F}_{\mathbf{h}}''\right|}\right)\right| = \left|\cos^{-1}\left(\frac{\Delta\boldsymbol{F}_{\mathbf{h}}}{2\left|\boldsymbol{F}_{\mathbf{h}}''\right|}\right)\right|, \tag{19}$$

A number of procedures have been proposed to break the phase ambiguity. Ramachandran & Raman [65] proposed in 1956 that, between the two equally possible phases of the doublet, one can make choice to that phase which is closer to the phase of the real-part scattering of the heavy-atom substructure. In 1981 Hendrickson & Teeter [66] used a similar but improved method to solve the unknown structure of the protein crambin from sulfur-SAD data collected using Cu-$K\alpha$ X-rays. In their work, the phase of a particular reflection was determined as follows. If the Sim distribution [67], [Eq. (20)]

$$P_{Sim}(\varphi_{\mathbf{h}}) = N\exp\left[\chi\cos\left(\varphi_{\mathbf{h}} - \varphi_{\mathbf{h}}'\right)\right] \;\square\; N\exp\left\{\chi\cos\left[\varphi_{\mathbf{h}} - \left(\varphi_{\mathbf{h}}'' - \frac{\pi}{2}\right)\right]\right\} \tag{20}$$

clearly favoured one of the maxima in the bimodal SAD-phase distribution [Eq. (21)]

$$P_{\mathrm{SAD}}(\varphi_{\mathbf{h}}) = N'\exp\left\{-\left[\Delta F - 2\left|\boldsymbol{F}_{\mathbf{h}}''\right|\sin\left(\varphi_{\mathbf{h}} - \varphi_{\mathbf{h}}'\right)\right]^2 \Big/ 2E^2\right\} \;\square\; N'\exp\left\{-\left[\Delta F - 2\left|\boldsymbol{F}_{\mathbf{h}}''\right|\cos\left(\varphi_{\mathbf{h}} - \varphi_{\mathbf{h}}''\right)\right]^2 \Big/ 2E^2\right\}, \tag{21}$$

then the unimodal distribution of Eq. (20) was used directly. Otherwise, Eqs. (20) and (21) were multiplicatively combined to give the distribution of the phase $\varphi_{\mathbf{h}}$. SAD phasing algorithms of many modern programs are based on the same principle of the second treatment. In Eqs. (20) and (21), N and N' are the normalising coefficient of the corresponding probability distribution; χ is related to structure-factor amplitudes of the whole unit cell, the heavy-atom substructure and the unknown part of the unit cell; $\varphi'_{\mathbf{h}}$ is the phase of $\boldsymbol{F}_{\mathbf{h}}' = \sum_{j=1}^{N}(f_j + \Delta f_j')\exp(i2\pi\boldsymbol{h}\cdot\boldsymbol{r}_j)$, which is the structure factor contributed from the real-part scattering of the heavy-atom substructure (notice that $\varphi'_{\mathbf{h}} = \varphi''_{\mathbf{h}} - 90^{\mathbf{o}}$ when there is only one kind of heavy atoms); $\Delta F = \left|\boldsymbol{F}_{\mathbf{h}}^{+}\right| - \left|\boldsymbol{F}_{\mathbf{h}}^{-}\right|$ is the Bijvoet difference; $\left|\boldsymbol{F}_{\mathbf{h}}''\right|$ is the amplitude of $\boldsymbol{F}_{\mathbf{h}}'' = i\sum_{j=1}^{N}\Delta f_j''\exp(i2\pi\boldsymbol{h}\cdot\boldsymbol{r}_j)$; $E = \left(\sigma_{\Delta F}^2 + E_o^2\right)$, where σ is the standard deviation and E_o is the residual lack-of-closure error [68]. It turns out from Eq. (21) that the SAD phase distribution is bimodal and, since $\Delta F \;\square\; 2\left|\boldsymbol{F}_{\mathbf{h}}''\right|\cos\Delta\varphi_{\mathbf{h}}$ [64], the maxima will be at $\varphi_{\mathbf{h}} = \varphi''_{\mathbf{h}} + |\Delta\varphi_{\mathbf{h}}|$ and $\varphi_{\mathbf{h}} = \varphi''_{\mathbf{h}} - |\Delta\varphi_{\mathbf{h}}|$. According to Eq. (20) the Sim distribution is unimodal and peaks at $\varphi_{\mathbf{h}} \approx \varphi''_{\mathbf{h}} - 90^{\mathbf{o}}$. Consequently, to break the SAD phase ambiguity with either the heavy-atom substructure or the corresponding Sim distribution will always favour the peak in Eq. (21) at $\varphi_{\mathbf{h}} = \varphi''_{\mathbf{h}} - |\Delta\varphi_{\mathbf{h}}|$. This is equivalent to force phases of the whole structure to approach that of the heavy-atom substructure. Large errors in phase estimation will then be introduced, because the diffraction contribution of the heavy-atom substructure could never dominate that of the whole protein structure. Wang introduced in 1981 the iterative single-wavelength anomalous scattering (ISAS) method [69], in which both alternative phases are used to calculate the electron density map and then the iterative solvent flattening and phase merging process is applied to break the phase ambiguity. The solvent flattening method is nowadays the most popular technique of improving the quality of electron density maps. However double-phase electron density maps could cause problems when the anomalous-scattering substructure is centrosymmetric. Based on the work of Fan in

1965 [32], Fan & Gu [51] proposed in 1985 the direct-method SAD/SIR phasing method. The main points are as follows:

i) Bimodal phase distributions [Eq. (21)] from the SAD experiment are approximated as the sum of two Gaussian functions

$$P(\varphi_{\mathbf{h}}) = \frac{1}{2\sigma_{\mathbf{h}}(2\pi)^{1/2}} \exp\left[-\left[\varphi_{\mathbf{h}} - (\varphi_{\mathbf{h}}'' + |\Delta\varphi_{\mathbf{h}}|)\right]^2 / 2\sigma_{\mathbf{h}}^2\right] + \frac{1}{2\sigma_{\mathbf{h}}(2\pi)^{1/2}} \exp\left[-\left[\varphi_{\mathbf{h}} - (\varphi_{\mathbf{h}}'' - |\Delta\varphi_{\mathbf{h}}|)\right]^2 / 2\sigma_{\mathbf{h}}^2\right] \quad . \tag{22}$$

ii) The probability for $\Delta\varphi_{\mathbf{h}}$ being positive is given by

$$P_+ = \frac{1}{2} + \frac{1}{2}\tanh\left\{\sin|\Delta\varphi_{\mathbf{h}}|\left[\sum_{\mathbf{h}'} m_{\mathbf{h}} m_{\mathbf{h}-\mathbf{h}'} \kappa_{\mathbf{h},\mathbf{h}'} \sin(\Phi'_3 + \Delta\varphi_{\mathbf{h}',best} + \Delta\varphi_{\mathbf{h}-\mathbf{h}',best}) + \chi \sin\delta_{\mathbf{h}}\right]\right\}, \tag{23}$$

which is based on the product of the Sim distribution [Eq. (20)] and the Cochran distribution [70] [Eq. (24)]

$$P_{Cochran}(\varphi_{\mathbf{h}}) = N'' \exp\left[\sum_{\mathbf{h}'} \kappa \cos(\varphi_{\mathbf{h}} - \varphi_{\mathbf{h}'} - \varphi_{\mathbf{h}-\mathbf{h}'})\right]. \tag{24}$$

iii) The phase ambiguity is resolved by multiplying the first Gaussian function in Eq. (22) by P_+ and multiplying the second Gaussian function by P_- (= 1– P_+). It follows that

$$P(\varphi_{\mathbf{h}}) = \frac{P_+}{\sigma_{\mathbf{h}}(2\pi)^{1/2}} \exp\left[-\left[\varphi_{\mathbf{h}} - (\varphi_{\mathbf{h}}'' + |\Delta\varphi_{\mathbf{h}}|)\right]^2 / 2\sigma_{\mathbf{h}}^2\right] + \frac{1-P_+}{\sigma_{\mathbf{h}}(2\pi)^{1/2}} \exp\left[-\left[\varphi_{\mathbf{h}} - (\varphi_{\mathbf{h}}'' - |\Delta\varphi_{\mathbf{h}}|)\right]^2 / 2\sigma_{\mathbf{h}}^2\right] \quad . \tag{25}$$

iv) Finally the "best phase" $\varphi_{\mathbf{h},\,best}$ and the associated "figure of merit" $m_{\mathbf{h}}$ of a particular reflection **h** needed for calculating the "best Fourier map" are obtained by looping of Eqs. (26), (27) and (23) with the initial P_+ set to 1/2 and then by ending up with Eq. (28). The reader is referred to the reference [71] for details.

$$\tan(\Delta\varphi_{\mathbf{h},\,best}) = \frac{2\left(P_+ - \frac{1}{2}\right)\sin|\Delta\varphi_{\mathbf{h}}|}{\cos\Delta\varphi_{\mathbf{h}}} \tag{26}$$

$$m_{\mathbf{h}} = \exp(-\sigma_h^2)\left\{\left[2(P_+ - 1)^2 + \frac{1}{2}\right]\left[1 - \cos(2\Delta\varphi_{\mathbf{h}})\right] + \cos(2\Delta\varphi_{\mathbf{h}})\right\} \tag{27}$$

$$\varphi_{\mathbf{h},\,best} = \varphi''_{\mathbf{h}} + \Delta\varphi_{\mathbf{h},\,best} \tag{28}$$

Unlike the Sim distribution, the Cochran distribution may peak anywhere from 0 to 2π. Hence P_+ in Eq. (25) may enhance the first Gaussian function as well as the second. Besides, the Cochran distribution is independent of the heavy-atom substructure; hence there will be no effect of whether or not the heavy-atom substructure is centrosymmetric. Finally, with the use of Cochran's distribution, the phase $\varphi_{\mathbf{h}}$ is not only explicitly related to the anomalous scattering signal associated with the particular reflection **h**, but also to that of most available reflections in the reciprocal space. Estimations from direct-method SAD phasing are thus much more reliable. The procedure has been tested respectively with the Hg-SAD data and the Hg-derivative SIR data from the known protein avian pancreatic polypeptide (aPP). Crystals of aPP belong to the space group C2 with unit-cell parameters a=34.18, b=32.92, c=28.44Å; β=105.30 ° and Z=4. The structure was solved originally by Blundell *et al.* [72] using both SAD and SIR data at 2.0Å resolution collected with Cu-$K\alpha$ X-rays. The test on Hg-SAD data [52] showed that, while conventional SAD phasing methods failed to yield interpretable maps, the map from direct-method SAD phasing is traceable even before density modification (see Fig. 10).

Figure 10 Portion of the electron-density map of the protein aPP from direct-method SAD phasing with the final model superimposed.

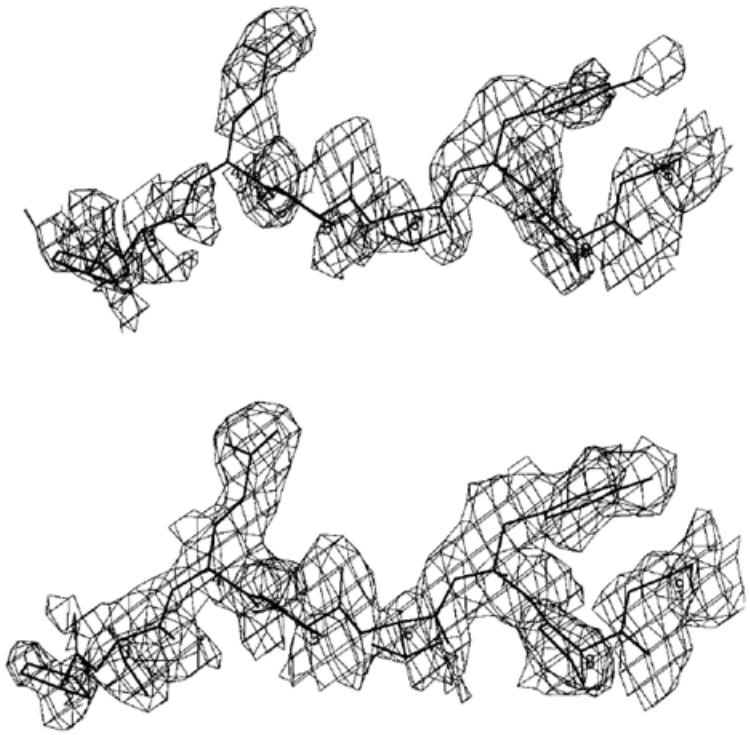

Figure 11 The same portion of two electron-density maps of the protein aPP with final model superimposed. Top: calculated with unresolved SIR phases; bottom: phases obtained from direct-method SIR phasing and improved by density modification. As is seen, the unresolved SIR map shows broken densities on side chains as well as on the main chain, while the direct-method plus density modification map has much better connectivity.

The test on Hg-derivative SIR data [55] showed that, owing to the centric arrangement of the Hg-substructure, conventional SIR phasing methods failed to break the phase ambiguity. However, direct-method SIR phasing followed by density modification produced an acceptable result (see Fig. 11).

The program OASIS (One wavelength Anomalous Scattering and Single Isomorphous Substitution) [73] has been written based on the above direct-method procedure. Krasilnikov *et al.* [74] succeeded in using OASIS for phasing the barium-SAD data at 2.9Å resolution of an RNA crystal, the structure of which is originally unknown. There was also reported [75] a successful phasing of Cr-$K\alpha$ sulfur-SAD data of thaumatin at 3.0Å resolution using OASIS.

4.2 Dual-space fragment extension making use of SAD/SIR information The P_+ formula [Eq. (23)] is the key of breaking SAD/SIR phase ambiguity by direct methods. Eq. (23) contains two terms inside the hyperbolic tangent function, i.e.

$$\sum_{\mathbf{h}'} m_{\mathbf{h}'} m_{\mathbf{h}-\mathbf{h}'} \kappa_{\mathbf{h},\mathbf{h}'} \sin\left(\Phi'_3 + \Delta\varphi_{\mathbf{h}',best} + \Delta\varphi_{\mathbf{h}-\mathbf{h}',best}\right) \quad \text{and} \quad \chi \sin\delta_{\mathbf{h}} .$$

The first term comes from the Cochran distribution incorporating SAD/SIR signals; the second term comes from the Sim distribution based on the real-part scattering of the known

partial structure. Values of P_+ are resulting from the balance of direct-method SAD/SIR estimation and partial-structure estimation. This makes Eq. (23) naturally suitable for accepting partial-structure feedback to enhance its phasing power. The iterative dual-space SAD/SIR phasing and structure-model completion process was thus proposed in 2004 by Wang *et al.* [76]. The flowchart of the process is shown in Fig. 12.

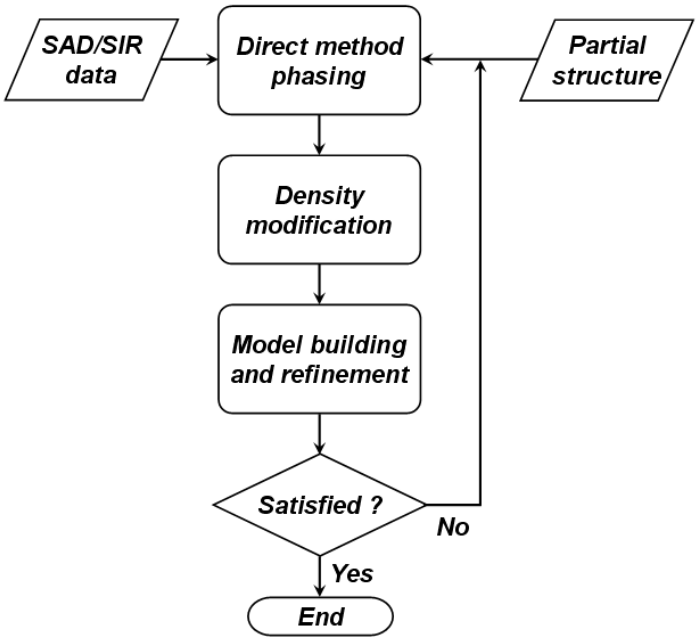

Figure 12 Flowchart of the iterative dual-space SAD/SIR phasing and structure-model completion process. In the block "Direct method phasing" the program OASIS-2004 or a newer version is used; in the block "Density modification" the Program DM [77] or RESOLVE [78] is used; in the block "Model building and refinement", programs ARP/wARP [79] and REFMAC [80] or the program AutoBuild in PHENIX [81] are used.

Three typical applications [82] are quoted here. The SAD data used were from lysozyme, azurin and xylanase as summarized in Table 1. All the data sets were difficult to phase. The SAD phasing of xylanase with data collected at λ = 1.49Å belongs to one of the most difficult cases, which had been an impossible task for existing methods as described by Ramagopal *et al.* in 2003 [83]. The Bijvoet ratio 0.56% is one of the lowest Bijvoet ratios so far has been successfully treated in protein SAD phasing. Besides, the low solvent contents 37% is not beneficial to the subsequent phase improvement by density modification techniques. Ribbon models obtained with and without dual-space iteration are compared in Fig. 13. As is seen, with the dual-space iteration, structure models automatically obtained are all larger than 90% of the entire structure and they are about 2 to 4 times bigger than that obtained without using dual-space iteration. Applications of OASIS dual-space SAD phasing and model completion reported by authors outside the Beijing Institute of Physics can be found in references [86-92].

Table 1 Summary of the test samples

Protein	Number of residues per AU	High resolution limit (Å)	Space group	X-Ray wavelength (Å)	Number of anomalous scatterers per AU	Expected Bijvoet ratio (%)	Reference
Lysozyme	129	2.0	$P4_32_12$	2.29	16×S	3.0	[84]
Azurin	129	1.9	$P4_122$	0.97	1×Cu	1.44	[85]
Xylanase	303	1.8	$P2_1$	1.49	6×S	0.56	[83]

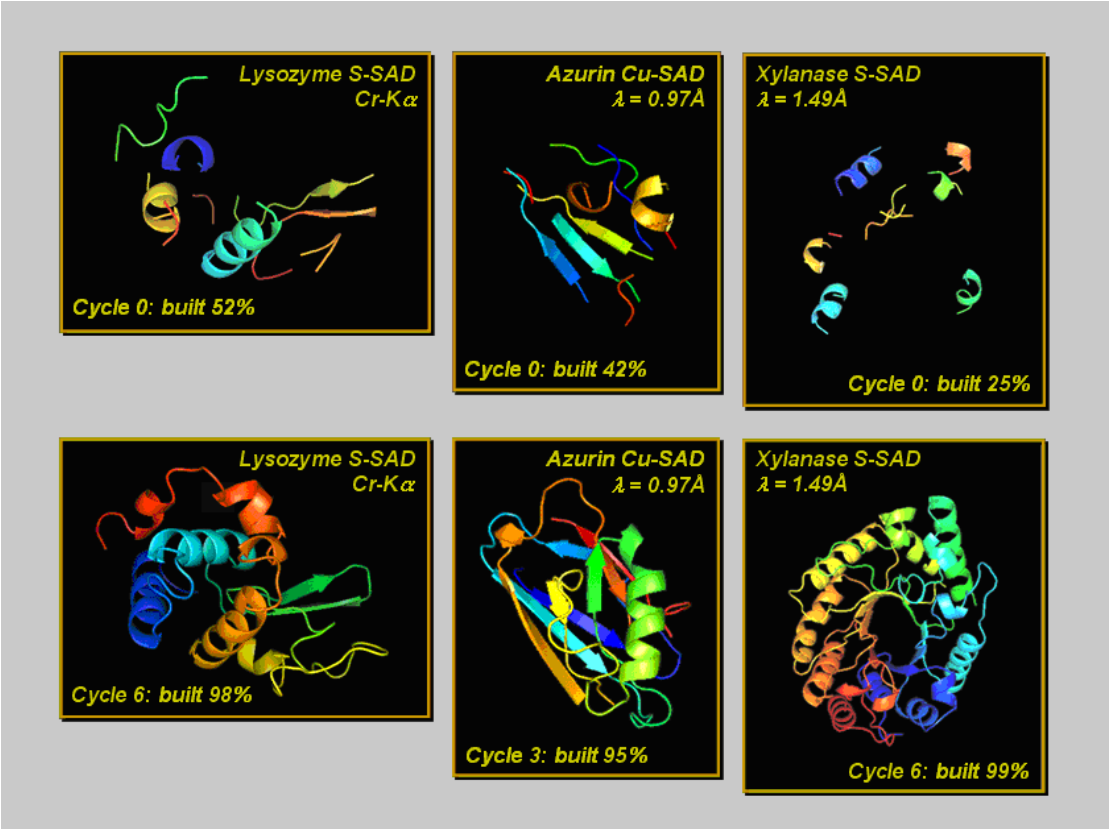

Figure 13 Ribbon models of three sample proteins (see Tab. 1). Upper row, models obtained from a single run of OASIS followed by density modification and model building and refinement; lower row, models obtained from dual-space iteration (see the flowchart in Fig. 12).

4.3 Dual-space MR-model completion in the absence of SAD/SIR information By using the P_+ formula [Eq. (23)] in SAD/SIR phasing, it implies that the "0 to 2π" phase problem is reduced to a "plus or minus" sign problem. Thus both the accuracy in phase estimation and the stability in phase iteration are significantly increased. This is based on the expression of experimental SAD information by Eqs. (18) and (19). If the P_+ formula can also be beneficial to the refinement of molecular replacement (MR) phases, that would be important in protein crystallography. Because in recent years about 70% of protein structures

deposited to the Protein Data Bank (*http://www.rcsb.org/*) were solved using the MR method. The main problem is how to define $\varphi''_{\mathbf{h}}$ and $|\Delta\varphi_{\mathbf{h}}|$ in the absence of SAD/SIR information, since most protein diffraction data collected for use with the MR method are lack of SAD/SIR signals. This problem has been solved by He *et al.* [93]. In their treatment $\varphi''_{\mathbf{h}}$ is defined as a reference phase calculated from a randomly selected 5% of the atoms in the current structure model, while $|\Delta\varphi_{\mathbf{h}}|$ is defined as the absolute difference between the phase of the current model and $\varphi''_{\mathbf{h}}$. The redefinition implies that during the refinement phases close to the true values will probably get the plus sign for $|\Delta\varphi_{\mathbf{h}}|$ and so be kept unchanged, while those distant from the true values will probably get the minus sign and so undergo a large shift. This is the desired property of a phase refinement process. The procedure has been tested and incorporated into the program OASIS-2006 [94]. The flowchart is shown in Fig. 14. Results from a sample are shown in Fig. 15. As is seen, dual-space MR-model completion making use of OASIS is much more efficient than that without. In the example, the starting model is only about 20% of the total structure. ARP/wARP-DM iteration failed to extend such a starting model. However, ARP/wARP-OASIS-DM iteration succeeded in extending it to a model containing more than 90% of the entire structure.

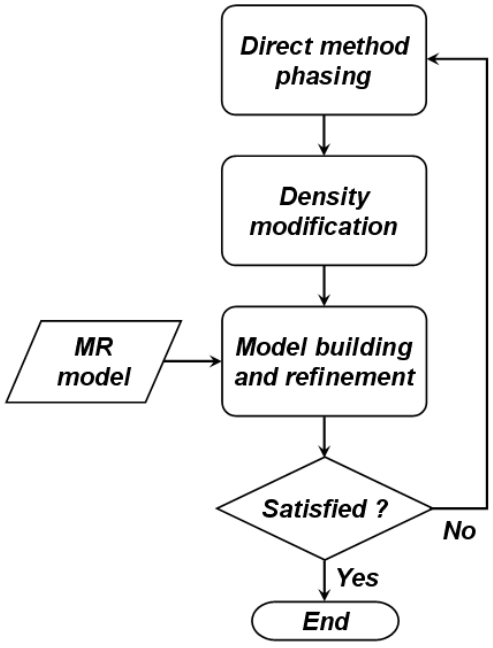

Figure 14 Flowchart of dual-space MR-model completion

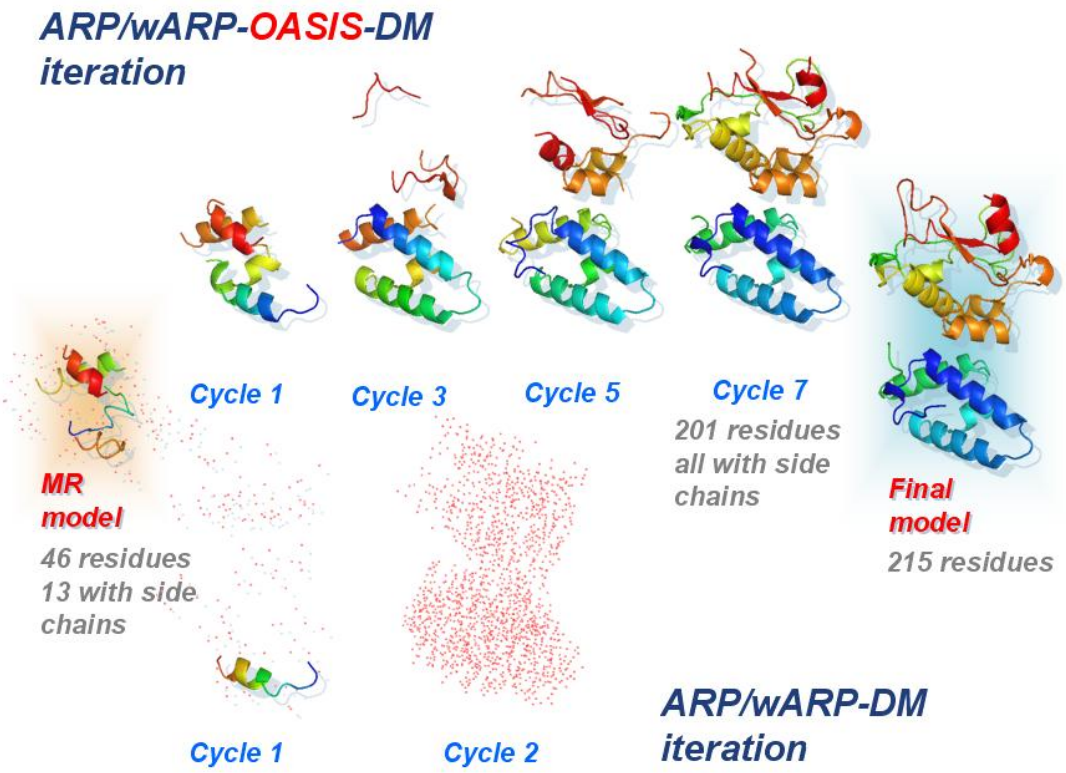

Figure 15 Model completion of E7_C–Im7_C with the 46-residue starting model. Upper row: models from the first, third, fifth and seventh cycles of ARP/wARP–OASIS–DM iteration. Lower row: models from two cycles of ARP/wARP–DM iteration. The starting model and refined (final) model are shown on the left and right, respectively.

4.4 Combination of SAD iteration and MR iteration Now there are two different kinds of dual-space iteration involving the program OASIS. The first kind makes use of SAD/SIR information, while the second does not. The former is dedicated to data with SAD/SIR information, while the latter is for data without. Recently Panjikar *et al.* [95] have compared OASIS MR-model completion with the MRSAD procedure. The latter involves refinement/extension of the heavy-atom substructure and OASIS SAD iteration. They showed that while MR iteration could effectively improve the structure model obtained from molecular replacement, the MRSAD procedure gave much better results. The conclusion is self-evident, since MRSAD uses more experimental information (the SAD information) than the MR iteration does. Our recent study showed that, during the final stage of model completion, even for protein diffraction data containing SAD/SIR signals, the MR iteration may provide better results than that obtained from the SAD/SIR iteration, especially when SAD/SIR signals are weak and experimental errors are large. A scheme of combining the SAD/SIR iteration and the MR iteration was thus proposed [96]: For diffraction data containing reasonable SAD/SIR signals, the SAD/SIR iteration will be run first. After sufficient cycles of iteration, say 10 cycles, if the resulting structure models are not big enough, say less than 90% of the total structure, then the iteration will be changed to MR

iteration taking the best resultant structure model from previous SAD iteration cycles as input. The philosophy behind the scheme is simple. SAD/SIR signals contain phase information, which is particularly important in early stages of partial-model extension. However SAD/SIR

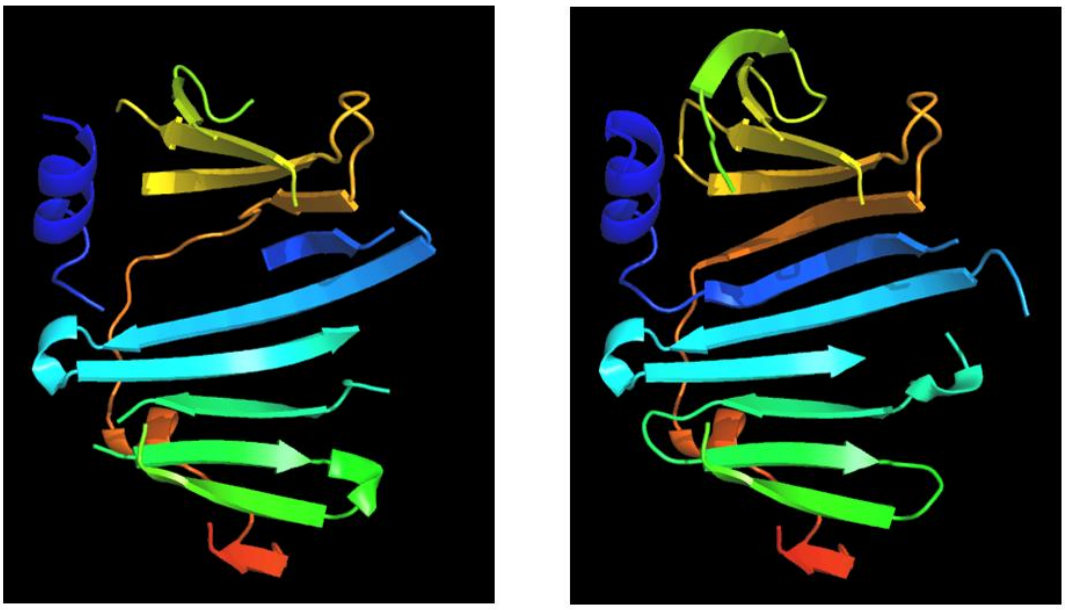

Figure 16 Ribbon structure models of TTHA1012. Left, the best model from 21 cycles SAD iteration; right, the best model from 10 cycles of MR iteration based on 11 cycles SAD iteration.

signals also contain large experimental errors, which may disturb the convergence of model completion. On the other hand, the MR iteration does not make use of SAD/SIR signals and is not affected by errors in them. Hence it is reasonable to expect that the combination of SAD/SIR iteration and MR iteration will give better results than that from SAD/SIR iteration alone. A test with the sulfur-SAD data of TTHA1012 is quoted here. The protein TTHA1012 crystallizes in the space group $P2_12_12_1$ with unit-cell parameters a = 41.35, b = 58.77 and c = 80.49Å. There are 2 sulfur atoms and 213 residues in the asymmetric unit. The 2.2Å sulfur-SAD data was collected with Cr-$K\alpha$ X-rays. The expected Bijvoet ratio is 0.83%. The SAD phasing of TTHA1012 is one of the extremely difficult cases so far reported. Ribbon models resulting from the comparison are shown in Fig. 16. As is seen, the resultant structure model from the combined SAD and MR iteration is significantly better in completeness than that from the SAD iteration alone. For details and more examples the reader is referred to the original paper [96]. The latest version of the program OASIS has been released in 2009 [97]. Apart from the improved SAD phasing procedure, it includes a graphical user's interface for organizing and monitoring iterative jobs of dual-space fragment extension with or without SAD/SIR information. The package is available free of charge for academic users at the website: *http://cryst.iphy.ac.cn*. A pipeline for automatically solving protein crystal structures with OASIS as the main phasing program is now under construction.

4.5 Data resolution required for direct-method dual-space iteration Two factors affect the data-resolution requirement for the direct-method dual-space iteration. One is the direct-method phasing, while the other is the automatic model building. The following

discussion will involve only the former. It is well known that direct methods are sensitive to data resolution. The reason is that atomicity of crystal structures is assumed in the formulation of direct methods. This means that atomic-resolution diffraction data is required for *ab initio* direct-method phasing. While there is no theoretical prediction on what exact value should be the "atomic resolution", an empirical value of about 1.2Å is generally accepted. Actually protein structures so far determined via *ab initio* direct-method phasing are mostly with diffraction data at near to or higher than 1.2Å resolution. However, when SAD/SIR information or an MR-resulting model is available, the situation becomes quite different. SAD phasing by OASIS at different resolutions has been examined by Yao *et al.* [98] with the known protein TT0570, which crystallizes in the space group $P2_12_12$ with unit cell parameters a=100.57, b=109.10, c=114.86Å. There are 1206 residues in the asymmetric unit. Sulfur-SAD data were collected with Cu-$K\alpha$ X-rays and truncated to 2.1, 3.0, 3.5 and 4.0Å respectively

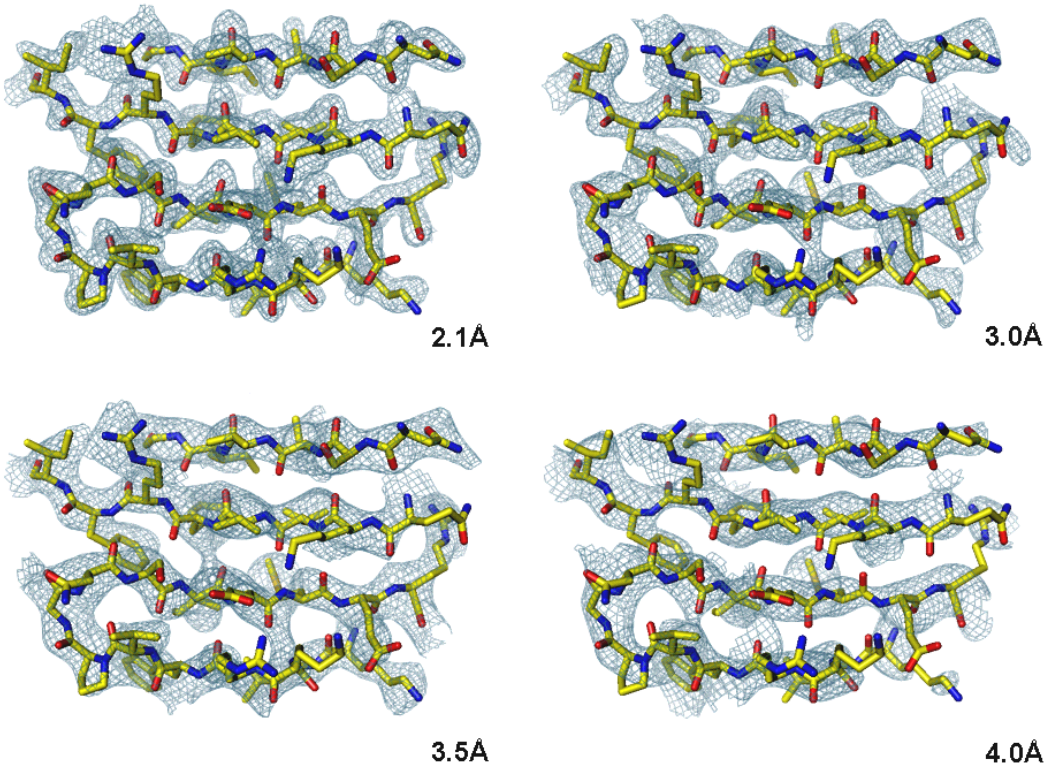

Figure 17 Partial electron-density maps (contoured at 1σ) of TT0570 at different resolutions (2.1, 3.0, 3.5 and 4.0Å) with the final structure model superimposed were plotted by the program PyMOL. Phases corresponding to each set of truncated data were derived separately by the program OASIS-2004 [76] and improved by density modification with the program DM [77].

for the test. Some results are quoted in Fig. 17. As is seen the map at 2.1Å is excellent. An automatic model building by the program ARP/wARP [79] led to a model containing 1178 (98% of the total 1206) residues, all with side chains. As the data resolution goes down, the quality of electron-density map degrades. Even though, the map at 3.0Å is still good for manual tracing, the map at 3.5Å is reasonable to an experienced worker and, the map at 4.0Å reveals good connectivity, although it would be difficult to locate side chains. The above observatio

ns are in practice consistent with the successful direct-method phasing by OASIS with the 2.9Å barium-SAD data [74], the 3.0Å sulfur-SAD data [75] and the 3.3Å selenium-SAD data [88]. As for direct-method MR phasing, there is not yet a test on the relation between map quality and data resolution and, there are fewer practical examples in comparison with that of direct-method SAD phasing. In practice, the lowest data resolution among the successful examples (see [93] and in Table 2 of [95]) is 2.4Å, which is already much lower than the "atomic resolution" of 1.2Å. However in principle, both direct-method SAD phasing and MR phasing take advantage of the reduction of the phase problem (searching in the range of 0 to 2π) to a sign problem (making choice between plus and minus) and, they are based on the same P_+ formula [Eq. (23)]. Hence the effect of data resolution to the quality of resultant electron-density maps should be similar. The lack of SAD/SIR information in direct-method MR phasing may be compensated by the MR-resulting model. Actually an MR-resulting model based structure-completion would be rather a structure-refinement process than a phase-deriving one. For the former, the ratio of parameters to reflections is more important than the high-resolution cutoff of the data.

5 Perspective Existent direct methods are by no means the ultimate methods. They are still in developing. Direct methods of solving commensurate/incommensurate modulated structures will find their position in X-ray structure analysis with powder crystalline samples. This would be of significance for material science. Direct methods in electron microscopy combining with that of solving commensurate/incommensurate modulated structures may further develop and find important applications in structure analysis of surface/interface and in-situ electron microscopic analysis. This may promote the development of relevant disciplines in condensed matter physics and material science. The further development of direct methods of solving biological macromolecular structures would have important influence to structural biology. New developments in X-ray free-electron-laser light source are ready to cause the revolution in X-ray science and technology. It can be expected that direct methods will play an important role in the development of new phasing technique for short-wavelength X-ray diffraction imaging. This in turn will promote the transition of atomic-scale structure analysis from single-crystal samples to single molecules/particles and from static structure analysis to time-resolved structure analysis. This would have broad and important implications for physics, chemistry, material science and life science. Direct methods have gone through a long way with brilliant achievements, but they are still far from the end.

Acknowledgements The author started working on crystallographic methods under the enlightenment of Prof. Qianzhang Wu. During the period of 1960's to 1970's valuable encouragements and supports were coming from Prof. Youxun Wu, Prof. Ruwei Shi and Prof. Ruihuang Yu. Sincere thanks are due to persons who took part in the long-term research. The author should like also to thank the referee, whose suggestions improved the manuscript. The project was supported by the Chinese Academy of Sciences, the National Natural Science Foundation of China and the National Basic Research Program of China.

References

[1] H. F. Fan, Acta Cryst. A**43**, C279 (1987).

[2] N. Uyeda and K. Ishizuka, J. Electr. Microsc. **24**, 65–72 (1975).

[3] F. H. Li and H. F. Fan, Acta Phys. Sin. **28**, 276–278 (1979). (In Chinese)

[4] F. S. Han, H. F. Fan, F. H. Li, Acta Cryst A**42**, 353–356 (1986).

[5] Y. W. Liu, S. B. Xiang, H. F. Fan, D. Tang, F. H. Li, Q. Pan, N. Uyeda, and Y. Fujiyoshi, Acta Cryst. A**46**, 459–463 (1990).

[6] P. L. Fijes, Acta Cryst A**33**, 109–113 (1977).

[7] D. Sayre, Acta Cryst **5**, 60–65 (1952).

[8] H. F. Fan, Z. Y. Zhong, C. D. Zheng, and F. H. Li, Acta Cryst. A**41**, 163–165 (1985).

[9] J. Karle and H. Hauptman, Acta Cryst. **9** 635–651 (1956).

[10] N. Uyeda, T. Kobayashi, E. Suito, Y. Harada, and M. Watanabe, J. Appl. Phys. **43**, 5181–5189 (1972).

[11] D. L. Dorset, W. F. Tivol, and J. N. Turner, Ultramicroscopy **38**, 41–45 (1991).

[12] H. F. Fan, S. B. Xiang, F. H. Li, Q. Pan, N. Uyeda, and Y. Fujiyoshi, **36**, 361–365 (1991).

[13] Z. Q. Fu, D. X. Huang, F. H. Li, J. Q. Li, Z. X. Zhao, T. Z. Cheng, H. F. Fan, Ultramicroscopy **54**, 229–236 (1994).

[14] H. F. Fan and Q. T. Zheng, Acta Phys. Sin. **24**, 97–104 (1975). (In Chinese)

[15] Y. W. Liu, H. F. Fan, and C. D. Zheng, Acta Cryst. A**44**, 61–63 (1988).

[16] M. Sundaralingam, R. Bergstrom, G. Strasburg, S. T. Rao, P. Roychowdhury, M. Greaser, and B. C. Wang, Science **227**, 945-948 (1985).

[17] R. Caliandro, B. Carrozzini, G. L. Cascarano, L. De Caro, C. Giacovazzo, and D. Siliqi, Acta Cryst. D**61**, 556–565 (2005).

[18] Z. H. Wan, Y. D. Liu, Z. Q. Fu, Y. Li, T. Z. Cheng, F. H. Li, and H. F. Fan, Z. Krist. **218**, 308-315 (2003).

[19] X. M. Li, F. H. Li, and H. F. Fan, Chin. Phys. B **18**, 2459-2463 (2009).

[20] T. Janssen, A. Janner, A. Looijenga-Vos, and P. M. de Wolff, in: International Tables for Crystallography, Vol. C, edited by A. J. C. Wilson (Kluwer Academic Publishers, Dordrecht, 1992), pp. 797-835; 843-844.

[21] Q. Hao, Y. W. Liu, and H. F. Fan, Acta Cryst. A**43**, 820-824 (1987).

[22] Y. Li, Z. H. Wan, and H. F. Fan, J. Appl. Cryst. **32**, 1017-1020 (1999).

[23] H. F. Fan, S. van Smaalen, E. J. W. Lam, and P. T. Beurskens, Acta Cryst. A**49**, 704-708 (1993).

[24] Z. Q. Fu and H. F. Fan, J. Appl. Cryst. **27**, 124-127 (1994).

[25] Z. Q. Fu and H. F. Fan, J. Appl. Cryst. **30**, 73-78 (1997).

[26] H. F. Fan, IUCr Computing Commission Newsletter No. 5, 16-23 (2005).

[27] H. F. Fan, IUCr Computing Commission Newsletter No. 5, 24-27 (2005).

[28] Y. D. Mo, T. Z. Cheng, H. F. Fan, J. Q. Li, B. D. Sha, C. D. Zheng, F. H. Li, and Z. X. Zhao, Supercond. Sci. Technol. **5**, 69-72 (1992).

[29] Z. Q. Fu, Y. Li, T. Z. Cheng, Y. H. Zhang, B. L. Gu, and H. F. Fan, Science in China, A**38**, 210-216 (1995).

[30] S. van Smaalen, A. Meetsma, G. A. Wiegers, and J. L. de Boer, Acta Cryst. B**47**, 314-325 (1991).

[31] Y. D. Mo, Z. Q. Fu, H. F. Fan, S. van Smaalen, E.J.W. Lam, and P.T. Beurskens, Acta Cryst. A**52**, 640-644 (1996).
[32] H. F. Fan, Acta Phys. Sin. **21**, 1114-1118 (1965). [In Chinese, English translation in: Chinese Phys. 1429-1435 (1965).]
[33] C. L. Coulter, J. Mol. Biol. **12**, 292-295 (1965).
[34] J. Karle, Acta Cryst. **21**, 273-276 (1966).
[35] A. C. Hazell, Nature (London) **227**, 269 (1970).
[36] W. A. Hendrickson, Acta Cryst. B**27**, 1474-1475 (1971).
[37] S. K. Sikka, Acta Cryst. A**29**, 211-212 (1973).
[38] J. J. L. Heinerman, H. Krabbendam, J. Kroon, and A. L. Spek, Acta Cryst. A**34**, 447-450 (1978).
[39] H. Hauptman, Acta Cryst. A**38**, 632-641 (1982).
[40] H. Hauptman, Acta Cryst. A**38**, 289-294 (1982).
[41] H. Hauptman, S. Potter, and C. M. Weeks, Acta Cryst. A**38**, 294-300 (1982).
[42] H. Hauptman, Acta Cryst. A**52**,490-496 (1996).
[43] H. Xu and H. Hauptman, Acta Cryst. A**59**, 60-65 (2003).
[44] S. Fortier, N. J. Moore, and M. E. Fraser, Acta Cryst. A**41**, 571-577 (1985).
[45] C. Giacovazzo, Acta Cryst. A**39**, 585-592 (1983).
[46] C. Giacovazzo, G. Cascarano, and C. D. Zheng, Acta Cryst. A**44**, 45-51 (1988).
[47] C. Giacovazzo, D. Siliqi, and J. G. Platas, Acta Cryst. A**51**, 811-820 (1995).
[48] C. Giacovazzo, M. Ladisab, and D. Siliqi, Acta Cryst. A**58**, 598-604 (2002).
[49] C. Giacovazzo and D. Siliqi, Acta Cryst. D**60**, 73-82 (2004).
[50] H. F. Fan, F. S. Han, and J. Z. Qian, Acta Cryst. A**40**, 495-498 (1984).
[51] H. F. Fan and Y. X. Gu, Acta Cryst. A41, 280-284 (1985).
[52] H. F. Fan, Q. Hao, Y. X. Gu, J. Z. Qian, C. D. Zheng, and H. Ke Acta Cryst. A**46**, 935-939 (1990).
[53] B. D. Sha, S. P. Liu, Y. X. Gu, H. F. Fan, H. Ke, J. X. Yao, and M. M. Woolfson, Acta Cryst. D**51**, 342-346 (1995).
[54] X. F. Zheng, C. D. Zheng, Y. X. Gu, Y. D. Mo, H. F. Fan, and Q. Hao, Acta Cryst. D**53**, 49-55 (1997).
[55] Y. D. Liu, Y. X. Gu, C. D. Zheng, Q. Hao, and H. F. Fan, Acta Cryst. D**55**, 846-848 (1999).
[56] Y. D. Liu, I. Harvey, Y. X. Gu, C. D. Zheng, Y. Z. He, H. F. Fan, S. S. Hasnain, and Q. Hao, Acta Cryst. D**55**, 1620-1622 (1999).
[57] H. F. Fan, Q. Hao, and M. M. Woolfson, Acta Cryst. A**46**, 656-659 (1990).
[58] H. F. Fan, Q. Hao, and M. M. Woolfson, Acta Cryst. A**46**, 659-664 (1990).
[59] H. F. Fan, M. M. Woolfson, and J. X. Yao, Proc. R. Soc. Lond. A **442**, 13-32 (1993).
[60] M. M. Woolfson, J. X. Yao and H. F. Fan, Acta Cryst. D**53**, 673-681 (1997).
[61] E. A. Klop, H. Krabbendam, and J. Kroon, Acta Cryst. A**43**, 810-820 (1987).
[62] P. Verwer, H. Krabbendam, and J. Kroon, Acta Cryst. A**47**, 143-144 (1991).
[63] C. E. Kyriakidis, R. Peschar, and H. Schenk, Acta Cryst. A**49**, 557-569 (1993).
[64] T. L. Blundell and L. N. Johnson, Protein Crystallography (Academic Press Inc. London, 1976) p. 177.
[65] J. N. Ramachandran and S. Raman, Curr. Sci. **25**, 348-351 (1956).

[66] W. A. Hendrickson and M. M. Teeter, Nature **290**, 107-113 (1981).
[67] G. A. Sim, Acta Cryst. **12**, 813-815 (1959).
[68] L. F. Ten Eyck and A. Arnone, J. Mol. Biol. **100** 3-11 (1976).
[69] B. C. Wang, Methods in Enzymology, **115**, 90-112 (1985).
[70] W. Cochran, Acta Cryst. **8**, 473-478 (1955).
[71] M. M. Woolfson and H. F. Fan, Physical and non-physical methods of solving crystal structures (Cambridge University Press, 1995) pp. 157-163; 177-178.
[72] T. L. Blundell, J. E. Pitts, I. J. Tickle, S. P. Wood, and C. W. Wu, Proc. Natl. Acad. Sci. USA, **78**, 4175-4179 (1981).
[73] Q. Hao, Y. X. Gu, C. D. Zheng, and H. F. Fan, J. Appl. Cryst. **33**, 980-981 (2000).
[74] A, S. Krasilnikov, Y. Xiao, T. Pan, and A. Mondragón, Science, **306(**5693**)**, 104-107 (2004). (Details of structure determination can be found in Supporting Online Material at *http://www.sciencemag.org/cgi/data/306/5693/104/DC1/1*)
[75] A. Yamano, T. Sato, and T. Hasegawa, 構造生物 **10**(1), 1-9 (2004). (In Japanese)
[76] J. W. Wang, J. R. Chen, Y. X. Gu, C. D. Zheng, and H. F. Fan, Acta Cryst. D**60**, 1991–1996 (2004).
[77] K. D. Cowtan, and P. Main, Acta Cryst. D**49**, 148–157 (1993).
[78] T. C. Terwilliger, Acta Cryst. D**56**, 965–972 (2000).
[79] A. Perrakis, R. Morris, and V. S. Lamzin, Nature Struct. Biol. **6**, 458–463 (1999).
[80] G. N. Murshudov, A. A. Vagin, and E. J. Dodson, Acta Cryst. D**53**, 240–253 (1997).
[81] T. C. Terwilliger, R. W. Grosse-Kunstleve, P. V. Afonine, N.W. Moriarty, P. H. Zwart, L.-W. Hung, R. J. Read, and P. D. Adams, Acta Cryst. D**64**, 61–69 (2008).
[82] D. Q. Yao, S. Huang, J. W. Wang, Y. X. Gu, C. D. Zheng, H. F. Fan, N. Watanabe, and I. Tanaka, Acta Cryst. D**62**, 883–890 (2006).
[83] U. A. Ramagopal, M. Dauter, and Z. Dauter, Acta Cryst. D**59**, 1020–1027 (2003).
[84] N. Watanabe, Y. Kitago, I. Tanaka, J.-W. Wang, Y.-X. Gu, C.-D. Zheng, and H.-F. Fan, Acta Cryst. D**61**, 1533–1540 (2005).
[85] F. Dodd, S. S. Hasnain, Z. H. Abraham, R. R. Eady, and B. E. Smith, Acta Cryst. D**51**, 1052–1064 (1995).
[86] C. H. Huang, W. L. Lai, M. H. Lee, C. J. Chen, A. Vasella, Y. C. Tsai, and S. H. Liaw, J. Biol. Chem. **280**, 38831–38838 (2005).
[87] N. Watanabe, Acta Cryst. D**62**, 891–896 (2006).
[88] Y. Wu and B. D. Sha, Nature Struct. Mol. Biol. **13**, 589-593 (2006).
[89] C. Withers-Martinez, L. F. Haire, F. Hackett, P. A. Walker, S. A. Howell, S. J. Smerdon, G. G. Dodson, and M. J. Blackman, J. Mol. Biol. **375**, 773–781 (2008).
[90] Y. F. Zhou, B. Y. Nan, J. Nan, Q. J. Ma, S. Panjikar, Y. H. Liang, Y. P. Wang, and X. D. Su, J. Mol. Biol. **383**, 49–61 (2008).
[91] I. Kursula, P. Kursula, M. Ganter, S. Panjikar, K. Matuschewski, and H. Schueler, Structure **16**, 1638–1648, (2008).
[92] S. Selvanayagam, D. Velmurugan, T. Yamane, and A. Suzuki, Indian J. Phys. **83**, 1671-1679 (2009).
[93] Y. He, D. Q. Yao, Y. X. Gu, Z. J. Lin, C. D. Zheng, and H. F. Fan, Acta Cryst. D**63**, 793–799 (2007).
[94] T. Zhang, Y. He, Y. X. Gu, C. D. Zheng, Q. Hao, J. W. Wang, and H. F. Fan,

OASIS-2006 – A direct-method program for SAD/SIR phasing and reciprocal-space fragment extension, Institute of Physics, Chinese Academy of Sciences, Beijing, People's Republic of China (2007).

[95] S. Panjikar, V. Parthasarathy, V. S. Lamzin, M. S. Weiss, and P. A. Tucker, Acta Cryst. D**65**, 1089–1097 (2009).

[96] T. Zhang, L. J. Wu, Y. X. Gu, C. D. Zheng, and H. F. Fan, Chin. Phys. B **19**, 96101 (2010).

[97] T. Zhang, Y. X. Gu, C.D. Zheng, and H. F. Fan, Chin. Phys. B **19**, 86103 (2010).

[98] D. Q. Yao, H. Li, Q. Chen, Y. X. Gu, C. D. Zheng, Z. J. Lin, H. F. Fan, N. Watanabe, and B. D. Sha, Chin. Phys. B **17**, 1-9 (2008).