Vol 14 No 1, January 2005
1009-1963/2005/14(01)/0001-05

Chinese Physics

© 2005 Chin. Phys. Soc.
and IOP Publishing Ltd

Resolution improvement of electron micrographs by the direct method for biotin-binding protein streptavidin*

Yang Shi-Xin(阳世新)[a], Li Fang-Hua(李方华)[b]†, Liu Yu-Dong(刘玉东)[c], Wang Huai-Bin(王怀斌)[b], Gu Yuan-Xin(古元新)[b], and Fan Hai-Fu(范海福)[b]

[a] *College of Life Science, Sichuan University, Chengdu* 610064, *China*
[b] *Beijing National Laboratory for Condensed Matter Physics, Institute of Physics, Chinese Academy of Sciences, Beijing* 100080, *China*
[c] *Lineberger Comprehensive Cancer Centre, University of North Carolina at Chapel Hill, Chapel Hill, NC* 27599-7295, *USA*

(Received 6 July 2004; revised manuscript received 13 August 2004)

Direct-method phase extension has been applied to two-dimensional electron diffraction data of the protein streptavidin. Structure-factor amplitudes from electron diffraction were combined with phases from the corresponding electron micrographs. Maximum-entropy discrimination and cluster analysis were used to derive a solution from a large number of random trials. The phase extension from 0.3 to 0.25nm led to substantial improvement of the reconstructed projection image quality.

Keywords: electron crystallography, direct method, protein structure, streptavidin
PACC: 0290, 6165

1. Introduction

Ever since the three-dimensional (3D) structure map of the tail of bacteriophage T4 was determined by reconstructing a series of electron micrographs projected from different directions,[1] this new technique of 3D reconstruction has been set up and widely applied to the structural study of macromolecules and their complexes. Considerable effort has been expended to improve the resolution of electron micrographs of macromolecules since the 1970s when great progress was made for improving the resolution of electron microscopes. It was reported that the beam-induced specimen movement and radiation damage are two of the most important factors in reducing image contrast.[2] A big step forward for 3D reconstruction was made in 1975,[3] when the structure of purple membrane from Halobacterium halobium was investigated. The image contrast and resolution were enhanced by preserving the specimen in a less volatile glucose medium and by alleviating the radiation damage with a low dose of electrons. Preserving the specimen at liquid nitrogen or helium temperature is another efficacious way for reducing the radiation damage and enhancing the image contrast.[4] In order to minimize the effect of specimen movement the spot-scan mode by a computer control has been developed for the image recording, and the electron micrographs recorded in such a way showed a higher and more consistent contrast.[5,6] A complex image processing technique can be used to enhance the image resolution by many refinement procedures in methodology.[7,8] The resolution of protein structures determined by the 3D reconstruction has reached 0.35, 0.35 and 0.34nm for bacteriorhodopsin,[8] PhoE porin[9] and LHC-II,[10] respectively.

The direct method originally developed in x-ray crystallography plays a more and more important role in electron crystallography. For instance, the direct method has been successfully utilized for *ab-initio* crystal structure determination based on electron diffraction data.[11] Another successful application of direct method to electron crystallography is

*Project supported by the National Key Basic Research Special Foundation of China (Grant No G1999075604) and the National Natural Science Foundation of China (Grant Nos 19574073 and 19634020).
†E-mail: lifh@aphy.iphy.ac.cn

the phase extension: namely, in enhancing the resolution of crystal structures determined by HREM. For inorganic crystals the resolution of determined structures is much higher than that of the image and may be close to the diffraction resolution limit.[12−21] Apart from the above-mentioned phase extension technique that uses the direct method, the one using the maximum entropy method was applied to two-dimensional purple membrane.[22] In this paper, the direct method is applied to protein crystals and a test is carried out for the biotin-binding protein streptavidin.

2. Method

2.1. Experiment

The specimen used in the test is the biotin-binding protein streptavidin with space group $C222$ and unit-cell parameters a=8.225nm, b=8.225nm, c=5.000nm. The atomic model of protein streptavidin was determined by Weber *et al*[23] and Hendrickson *et al*[24] using the x-ray crystallographic method. Electron diffraction pattern was recorded on Kodak SO-163 film by a JEM1200 microscope operated at 100kV. Electron micrographs were taken at 400kV with a JEM4000 microscope using a flood beam illumination and low-dose technique.[25] Amplitudes of 0.25nm resolution were extracted from the electron diffraction pattern according to the method of Baldwin and Henderson.[26] Twenty-two image arrays from six different micrographs were processed with SPECTRA[27] and other programs from MRC.[7] Initially, lattice distortion was removed by unbending after lattice parameters were refined. A rough estimation of defocus and astigmatism parameters was obtained from contrast transfer function (CTF) ring[28] or by a comparison of their phases with the CTF-correct ones.[25] Then, a refinement of CTF parameters for each image array against others was carried out to minimize their merging phase residuals. Finally, 0.3nm phases are obtained by the weighted average of the merged phases from image arrays.

2.2. Phase extension

2.2.1. Initial extended phase sets

The direct-method program, SAPI,[29] has been used for the phase extension. The 0.25nm structure amplitudes from electron diffraction and the 0.3nm known phases from the corresponding micrographs were input to SAPI. Random starting phases[30] were assigned to 80 independent reflections between 0.3 and 0.25nm. The refinement of starting phases was carried out by the statistically-weighted tangent formula,[31] and 903 random trials were performed.

2.2.2. Entropy discrimination

It is essential in multisolution phase extension to pick up the correct solution. For this purpose, conventional figures of merit in the direct method have been tried, but they seemed not very successful, owing to the low resolution limit of the diffraction data. An alternative procedure was then used, in which a discrimination method referred to as maximum entropy image deconvolution[32] was used to screen out those solutions with phases nearly in random distribution before the cluster analysis[33] was carried out. Values of entropy were calculated for all projected potential maps corresponding to the random trials.

The inverse Fourier transform of each trial $F(\vec{H})$ set yields a projected potential distribution map (PPDM). Divide each unit cell into n pixels and let ϕ_j denote the value of projected potential of the jth pixel. The entropy of this PPDM is defined as

$$S = -\sum_{j=1}^{n} P_j \ln P_j, \tag{1}$$

where

$$P_j = \phi_j \Big/ \sum_{j=1}^{n} \phi_j. \tag{2}$$

Those $F(\vec{H})$ sets are discarded, of which the entropy values are less than a threshold value. Thus, a certain percent of phase sets are reserved for the cluster analysis.

2.2.3. Cluster analysis

The cluster analysis method has been applied to protein crystallography by Lunin *et al*[34] in determining the molecular envelope of proteins for diffraction data with very low resolution. Recently, Liu *et al*[33] described another implementation in choice of the correct solution among the phase sets resulting from random trials. According to the practice of cluster analysis, a considerable number of phase sets is close to the true solution, while others vary randomly and differ from one another. To find out those phase sets which are close to the true solution, the procedure is set up as follows. The average phase error for every pair of phase sets (phase-set pair), defined as

$$\langle \mathrm{dif}(\varphi) \rangle = \sum_{\vec{H}} F_{\vec{H}} [(\varphi_{\vec{H}})_1 - (\varphi_{\vec{H}})_2] \Big/ \sum_{\vec{H}} F_{\vec{H}}, \tag{3}$$

is calculated, where $F_{\vec{H}}$ is the amplitude of the structure factor, $(\varphi_{\vec{H}})_1$ and $(\varphi_{\vec{H}})_2$ are phases of two phase

sets. Then a phase error close to the peak in the histogram of the distribution of average phase errors is chosen as a radius. By taking one phase set as the centre, a cluster is formed such that all those phase sets with the average phase error between anyone of them and the centre set less than the radius are included. Evidently, the number of clusters formed by such a way equals the number of phase sets. The average phase set is obtained by the weighted average for the largest cluster as follows:

$$[m_{\vec{H}}\exp(\mathrm{i}\varphi_{\vec{H},\mathrm{best}})]_{\mathrm{average}} = \sum_{j}^{N}[m_{\vec{H}}\exp(\mathrm{i}\varphi_{\vec{H},\mathrm{best}})]_j/N, \quad (4)$$

where N is the number of phase sets in the largest cluster, $[m_{\vec{H}}]_j$ and $[\varphi_{\vec{H},\mathrm{best}}]_j$ are the weight and phase, respectively, of the jth phase set in the largest cluster. The average phase sets from the largest clusters are taken as the most probable solutions.

3. Test result

The 903 phase sets were derived from SAPI by combining amplitudes from electron diffraction pattern and phases from electron micrographs. Among them 142 phase sets were reserved after discarding those phase sets, of which the entropy values of corresponding projected potential maps were less than 10.275.

In the histogram given in Fig.1 the y-axis indicates the number of phase-set pairs and the x-axis the average phase error in degree. According to the above-mentioned principle of cluster analysis, the sharp peak in Fig.1 implies that only a single largest cluster, in which the number of phase-set pairs is the biggest, is found out. The average phase set in this largest cluster should be close to the true solution.

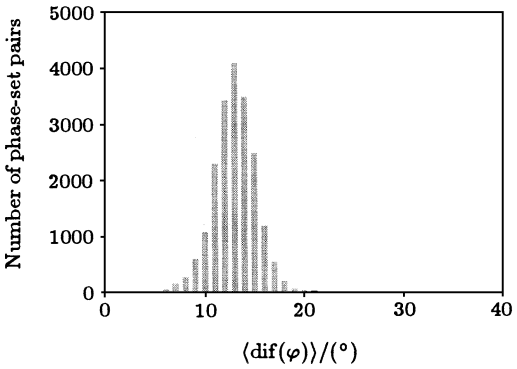


Fig.1. Histogram showing the distribution of phase-set pairs as a function of average error. The y-axis indicates the number of phase-set pairs and the x-axis the average phase error in degree.

The phase extension using amplitudes taken from the electron diffraction pattern and phases between 0.3 and 0.25nm was carried out. Phases of 80 reflections between 0.3 and 0.25nm were derived. Figure 2(a) displays the projected potential map obtained by using amplitudes taken from the electron diffraction pattern and phases between 0.3 and 0.25nm derived

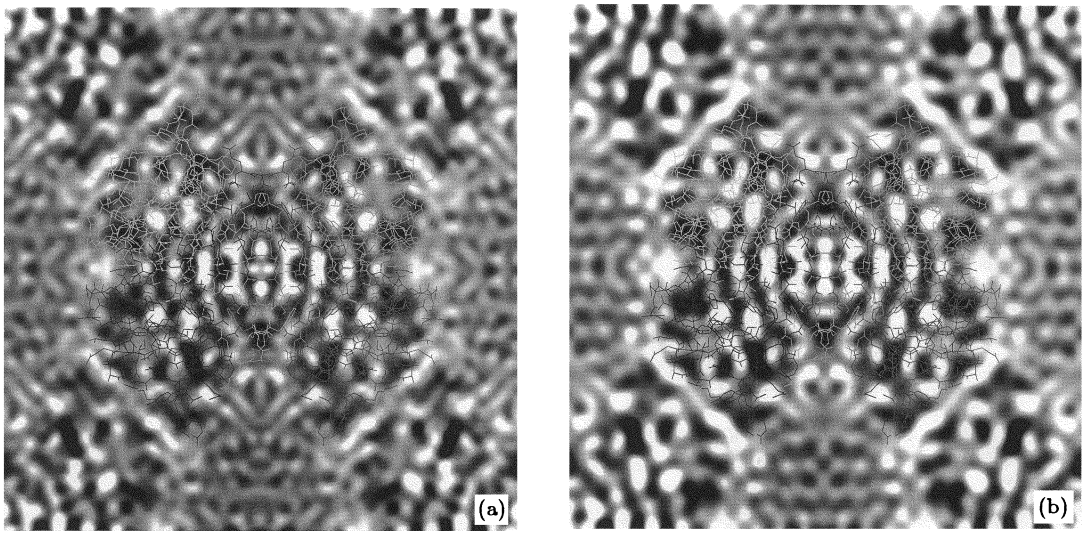


Fig.2. Projected maps of protein streptavidin of resolution. (a) 0.25nm obtained after phase extension, using the amplitudes taken from electron diffraction pattern and phases from phase extension using the program SAPI, and (b) 0.3nm with the amplitudes obtained from electron diffraction pattern and phases from the merging image arrays.[25] Four symmetrically arranged atomic backbones are superimposed on the map. The black colour in the map stands for high densities.

by the phase extension with the entropy discrimination and cluster analysis being applied. Four symmetrically arranged atomic backbones are superimposed on the map where the black colour stands for high densities. For comparison, the projected potential map of resolution 0.3nm obtained with phases from the merging image arrays and amplitudes from electron diffraction pattern [25] is displayed in Fig.2(b).

Figures 3(a) and 3(b) are pictures magnified from the top right region in Fig.2(a) and 2(b), respectively. It can be seen that in all the places pointed out by arrows the image contrast is more reasonable in Fig.3(a) than that in Fig.3(b). In other words, a better fit between the projected potential map and the structure model is found in Fig.3(a) than in Fig.3(b).

The derived phases of 80 reflections between 0.3 and 0.25nm were compared with the true phases, and 57 phases are correctly derived. The deviation between some derived phases from the true ones indicates that the resolution of obtained projected map is not exactly 0.25nm, but is close to 0.25nm.

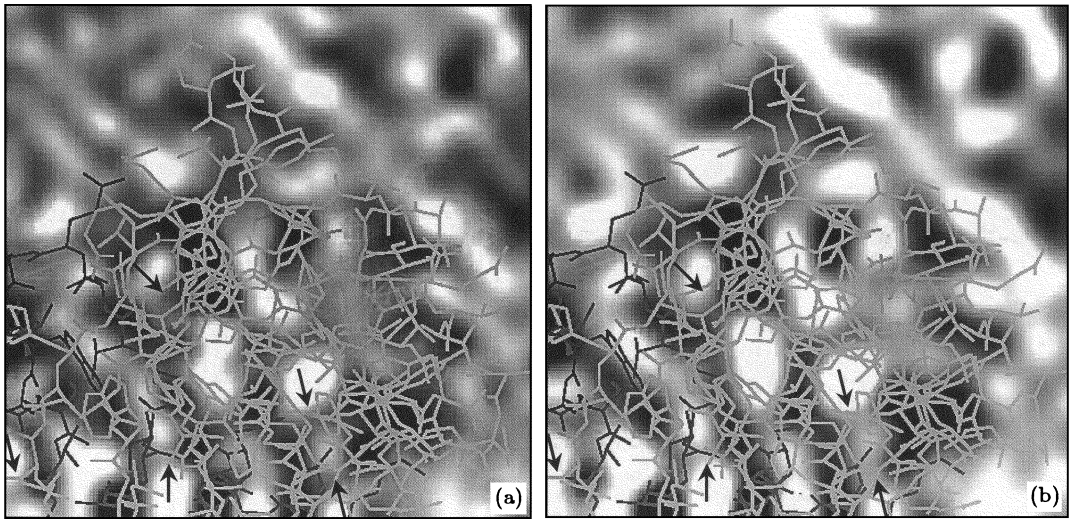


Fig.3. Magnified pictures of the top right area in Fig.2. The image contrast of Fig.3(a) fits the structure model better than that of Fig.3(b).

4. Conclusion

The validity of phase extension has been shown by means of the test for the biotin-binding protein streptavidin. It should be emphasized that the phase extension method developed in electron crystallography for inorganic crystals[12−21] has been modified to fit the protein crystals. After obtaining the initial random phase sets of structure factors by conventional phase extension, two more procedures have been introduced. Firstly, a preliminary discrimination method based on the principle of maximum entropy has been proposed and applied to filter the initial phases obtained by means of the conventional phase extension method so that a group of relatively good phase sets is reserved. Secondly, the cluster analysis method has been introduced to select the best cluster among the good phase sets. Finally, the average phase set for the best cluster is recognized as the correct solution. The results show that the resolution of projected structure obtained from the electron micrographs can be enhanced from 0.3nm to approximately 0.25nm by combining the electron diffraction data. It can be expected that this method is beneficial to the determination of unknown structures in protein electron crystallography.

Acknowledgement

The authors would like to express their gratitude to Professor Wah Chui for affording the electron microscopy and electron diffraction data of straptavidin.

References

[1] De Rosier D J and Klug A 1968 *Nature* **217** 130
[2] Baumeister W and Typke D 1993 *MSA Bulletin* **23** 11
[3] Unwin P N T and Henderson R 1975 *J. Mol. Biol.* **94** 425
[4] Dubochet J, Adrian M, Chang J J, Homo J C, Lepault J, McDowall A W and Shultz P 1988 *Quart. Rev. Biophys.* **21** 129

[5] Bullough P and Henderson R 1987 *Ultramicroscopy* **21** 223
[6] Brink J, Chiu W and Dougherty M 1992 *Ultramicroscopy* **46** 229
[7] Henderson R, Baldwin J M, Downing K H, Lepault J and Zemlin F 1986 *Ultramicroscopy* **19** 147
[8] Henderson R, Baldwin J M, Ceska T A, Zemlin F, Beckmann E and Downing K H 1990 *J. Mol. Biol.* **213** 899
[9] Jap B K, Walian P J and Gehring K 1991 *Nature* **350** 167
[10] Kuhlbrandt W, Wang D N and Fujiyoshi Y 1994 *Nature* **367** 614
[11] Dorset D L 1995 *Structural Electron Crystallography* (New York and London: Plenum Press)
[12] Fan H F, Zhong Z Y, Zheng C D and Li F H 1985 *Acta Cryst.* A **41** 163
[13] Fan H F, Xiang S B, Li F H, Pan Q, Uyeda N and Fujiyoshi Y 1991 *Ultramicroscopy* **36** 361
[14] Hu J J, Li F H and Fan H F 1992 *Ultramicroscopy* **41** 387
[15] Fu Z Q, Huang D X, Li F H, Li J Q, Zhao Z X, Cheng T Z and Fan H F 1994 *Ultramicroscopy* **54** 229
[16] Lu B, Li F H, Wan Z H, Fan H F and Mao Z Q 1997 *Ultramicroscopy* **70** 13
[17] Liu J, Li F H, Wan Z H, Fan H F, Wu X J, Tamura T and Tanabe K 1998 *Materials Transactions, JIM* **39** 920
[18] Jiang H, Li F H and Mao Z Q 1999 *Micron* **30** 417
[19] Liu J, Li F H, Wan Z H, Fan H F, Wu X J, Tamura T and Tanabe K 2001 *Acta Cryst.* A **57** 540
[20] Wang H B, Jiang H, Li F H, Che G C and Tang D 2002 *Acta Cryst.* A **58** 494
[21] Li F H 1998 *Microsc. Res. Technique* **40** 86
[22] Gilmore C J, Shankland K and Fryer J R 1993 *Ultramicroscopy* **49** 132
[23] Weber P C, Ohlendorf D H, Wendoloski J J and Salemme F R 1989 *Science* **243** 85
[24] Hendrickson W A, Pahler A, Smith J L, Satow Y, Merritt E A and Phizackerley R P 1989 *Proc. Natl. Acad. Sci. USA* **86** 2190
[25] Avila-Sakar A J and Chiu W 1996 *Biophys. J.* **70** 57
[26] Baldwin J M and Henderson R 1984 *Ultramicroscopy* **14** 319
[27] Schmid M R, Dargahi R and Tam M W 1993 *Ultramicroscopy* **48** 251
[28] Thon F 1971 *Electron Microscopy in Material Science* (New York: Academic) p571
[29] Fan H F, Yao J X, Zheng C D, Gu Y X and Qian J Z 1991 *SAPI-91, A Computer Program for Automatic Solution of Crystal Structures from X-ray Diffraction DATA* Institute of Physics, Chinese Academy of Sciences, Beijing, China
[30] Yao J X 1981 *Acta Cryst.* A **37** 642
[31] Hull S E and Irwin M J 1978 *Acta Cryst.* A **34** 863
[32] Hu J J and Li F H 1991 *Ultramicroscopy* **35** 339
[33] Liu Y D, Gu Y X, Zheng C D, Hao Q and Fan H F 1999 *Acta Cryst.* D **55** 846
[34] Lunin V Yu, Urzhumtsev A G and Skovoroda T P 1990 *Acta Cryst.* A **46** 540