Restoration of 3-D structure of insect flight muscle from a rotationally averaged 2-D X-ray diffraction pattern

Hiroyuki Iwamoto

Scattering and Imaging Div., Japan Synchrotron Radiation Research Institute, Sayo, Hyogo 679-5198, Japan

*Correspondence to:
Hiroyuki Iwamoto, Ph.D.,
Scattering and Imaging Div.
Japan Synchrotron Radiation Research Institute, SPring-8
1-1-1 Kouto, Sayo-cho, Sayo-gun, Hyogo 679-5198 Japan

email: iwamoto@spring8.or.jp
tel: +81-791-58-0803 ext. 3884

Abstract

The contractile machinery of muscle, especially that of skeletal muscle, has a very regular array of contractile protein filaments, and gives rise to a very complex and informative diffraction pattern when irradiated with X-rays. However, the analysis of the diffraction patterns is often difficult, because (1) only rotationally averaged diffraction patterns can be obtained, resulting in substantial loss of information, and (2) the contractile machinery contains two different sets of protein filaments (actin and myosin) with different helical symmetries, and the reflections originating from them are often overlapped. These problems may be solved if the real-space 3-D structure of the contractile machinery is directly calculated from the diffraction pattern. Here we demonstrate that, by using the conventional phase-retrieval algorithm (hybrid input-output), the real-space 3-D structure of the contractile machinery can be well restored from a single rotationally averaged 2-D diffraction pattern. In this calculation, we used a model structure of insect flight muscle, known to have a very regular structure. Possibilities of extending this technique to the actual muscle diffraction patterns is discussed.

Keywords: X-ray fiber diffraction, coherent diffractive imaging, hybrid input-output algorithm, insect flight muscle

Introduction

Because of their short wavelengths, X-rays are potentially capable of observing various types of samples (crystalline and non-crystalline) at an atomic resolution. Because of their high penetrability, the observation technique using X-rays is applicable to thick specimens not transparent to visible lights or electron beams. This technique is even applicable to live cells or tissues of living organism.

The drawback is that it is very difficult to fabricate X-ray lenses that ensure a high resolution. For this reason, it is a common practice that the X-rays scattered/diffracted by the samples are directly analyzed. A problem here is that both the amplitude and the phase information of the scattered beam is required for restoring the structural information of the samples, but the phase information is lost when the scattering is recorded on a detector. This is called the phase problem, and this has been the largest obstacle in applying the X-ray technique to a wide range of specimens.

Recently, however, it is theoretically shown (Bates, 1982; Miao et al., 1998) and experimentally demonstrated (Miao et al., 1999) that the once-lost phase information can be recovered (i.e. a magnified image of the sample is obtained) through iterative calculations, if certain conditions are met in the data recording. The conditions include (1) that the X-ray beams are coherent enough to cover the entire sample, and (2) that the data sampling should be finer than the Nyquist spacing (the inverse of the sampling size of the specimen) (Miao and Sayre, 2000). The latter is commonly called the oversampling condition. Since the first experimental demonstration (Miao et al., 1999), this phase-retrieval technique and its variations have been widely employed, and today they are called coherent diffractive imaging (CDI). CDI works very well for isolated particles with high contrast in electron density, such as metal nanoparticles. The original CDI is 2-dimensional (2-D), i.e., a 2-D image of the sample is restored from a 2-D X-ray scattering pattern.

The aim of the present work is to test whether it is possible for CDI to restore the 3-D structure of a complex biological system (specifically, the sarcomeric structure of muscle) from its 2-D X-ray diffraction pattern. Such a biological system is a challenge to CDI in many ways: (1) It is an assembly of proteins in water, and has much lower contrast of density as compared with a metal nanoparticle. (2) It consists of fibers with helical symmetries. CDI is not very suitable for such kind of structure. (3) It can generate only rotationally averaged diffraction patterns, from which substantial information has been lost. (4) 3-D expansion of CDI is required, and it consumes large computational resources and time.

The present paper demonstrates that, after appropriate expansions of the original 2-D CDI, in principle, the 3-D structure of the sarcomeric structure can be restored reasonable well from its rotationally averaged 2-D diffraction pattern. The details of the expansions are described in the following sections.

Materials and Methods

Structure of a sarcomere in muscle

Figure 1A is a schematic diagram of a sarcomere in insect flight muscle. A sarcomere is the minimal functional unit of muscle, and is of the order of 2 x 2 μm in length and diameter. A sarcomere contains two sets of protein filaments, made of actin and myosin, respectively, and they slide past each other to cause muscle contraction. Both filaments have their own helical symmetry. Figure 1B is a diffraction pattern recorded from the flight muscle of a giant waterbug, *Lethocerus deyrolei*. Only a rotationally averaged 2-D diffraction pattern can be recorded, because the myofibrils that constitute a flight muscle cell are rotationally randomly oriented. Our goal is to restore a 3-D structure as shown in Fig. 1A from a 2-D diffraction pattern as shown in Fig. 1B. Insect flight muscles are known to have a very regular structure, and the symmetries of the filaments are well documented (Tregear et al., 1998, 2004). Therefore, it is suitable to model in this kind of study.

Principle of 2-D CDI

The principle of CDI in the original form has been documented in many references (e.g., Miao and Sayre, 2000), and it is only briefly reiterated here.

A diffraction pattern recorded from a sample has amplitudes proportional to the square of the structure factor, (Fig. 2A), which is the Fourier transform of the electron densities of the sample, and it is a function in reciprocal space. Initially, this structure factor does not have phase information, so that a random phase is given. Applying inverse Fourier transformation to the structure factor should bring it back to the image of the sample in real space. Because of the wrong phase, a wrong image of the sample is regenerated (Fig. 2B), and it has densities outside the maximal area in which the sample is expected to exist. Then the densities outside this area (called support) is deleted (Fig. 2C). This process is called real-space correction. After this, the image is subjected to Fourier transform to obtain the structure factor again (Fig. 2D). Now the structure factor has phase information that is expected to be closer to the correct one than the first random one. After its amplitude is scaled to the observed value (reciprocal-space correction, Fig. 2E), the entire

processes are repeated. During the iterations, the error in the phase is gradually reduced, and finally the correct image of the sample is restored. Recent calculation procedure incorporates a process to automatically reduce the support to fit to the contour of the image, resulting in better convergence (the shrink-wrap algorithm, Marchesini et al., 2003).

The example shown in Fig. 2 is taken from a real in-silico experiment using a logo mark of SPring-8. For such an isolated object with high contrast, CDI works very well and the calculation usually converges within 400-600 iterations (Fig. 3).

2-D CDI does not work well with fibrous objects

On the other hand, CDI is unsuitable for fibrous materials with periodical structures, which are commonly found in living organisms. Examplea shown in Fig. 4 are a row of letter a (Fig. 4A) and the 2-D projection of an actin filament, decorated with isolated head parts of myosin molecules (decorated actin, Fig. 4B). The entire structure follows the helical symmetry of actin filament (28 monomers in 13 turns).

When CDI is applied to the image of decorated actin, the calculation does not converge even after thousands of iterations, and the restored image is not satisfactory (the right images in Fig. 4A and B). The reason that CDI does not work for fibrous samples may be that the scattering pattern is separated into discrete layer lines (Fig. 4C, D) and the oversampling condition is not met.

Expansion to 3-D CDI

Then we consider to expand the 2-D CDI to the 3-D CDI. Mathematically, one can achieve this by simply replacing 2-D Fourier transformation by 3-D Fourier transformation. In the real experiments, one must collect a large number of diffraction patterns while rotating the sample and combine them to obtain a single 3-D diffraction pattern, in a procedure similar to X-ray computed tomography. For real biological samples, it is difficult to collect a large number of diffraction patterns due to radiation damage, but it is much easier to conduct in-silico experiments to test how 3-D CDI performs. If the object is an isolated particle, the calculation converges without problem. The question is whether a fibrous object can be restored by 3-D CDI.

The example used to test 3-D CDI is the 3-D model of decorated actin (Fig. 5A) and that of an axoneme of eukaryotic flagella (Fig. 5B). When these objects are subjected to the same procedure as shown in Fig. 2, except that the calculation is carried out in a 3-D space, it is found that the calculation converges in both cases and the model structures are perfectly restored. As shown in Fig. 5, restoring a short segment of the filament by using

a short support seems to result in better convergence. The reason why a fibrous object is restorable by 3-D CDI is unclear, but it may be because the oversampling condition is met along two axes (X and Y), while in 2-D CDI, it is met along only one axis (X).

Restoration of the 3-D sarcomeric structure

Having established that 3-D CDI is workable for fibrous objects, we move on to the final objective, i.e., to restore the 3-D structure of the sarcomere of insect flight muscle from its rotationally averaged 2-D diffraction pattern. Details of the procedure of calculations is described below.

First, the basic lattice structure of actin and myosin filaments and the helical symmetry of both filaments are well documented (e.g., Tregear et al., 1998, 2004), and we make a maximal use of them. As for the lattice structure, the filaments form a hexagonal lattice, and a single actin filament is located midway between two neighboring myosin filaments (Fig. 6A). The unit cell size (d1,0 spacing) is approximately 45 nm. As for axial periodicities, myosin filament has a structure of 4-stat helix, which has 16 monomers in 2.5 turn with a monomer interval of 14.5 nm. Therefore, the basic periodicity of myosin filament is 232 nm (= 14.5 x 16 nm; Tregear et al., 1998, 2004; Fig. 6B). Actin filament has a structure of 1-start, 28 monomers in 13 turns with a monomer interval of 2.75 nm. Therefore, the basic periodicity of actin filament is 77 nm (= 2.75 x 28 nm; Tregear et al., 1998, 2004), but if it is multiplied by 3, it is also approximately 232 nm (Fig. 6B). Thus, the entire system can be regarded to have a basic axial repeat of 232 nm.

To ensure practically short calculation time, the iterative Fourier and inverse Fourier transformation must be done by using the fast Fourier transform (FFT) algorithm. To implement this, the data space to be processed must be a cube, with its side length of 2^n voxels (128, 256, etc. In the following calculation, the size of 256 was used). To avoid boundary problems, the cube must contain an integral number of repeating units. Here, in X direction (X and Y axes are perpendicular to the axis of the filaments), 6 unit cells of the hexagonal lattice are accommodated (side length = 45 x 6 = 270 nm). In Y direction, it is impossible to accommodate an integral number of repeating units. Therefore, the filament lattice was compressed to $\sqrt{3}/2$ so that 6 myosin-myosin intervals are accommodated in the side length of 270 nm. In Z direction (along the filament axis), the 232-nm repeat was expanded to 270 nm. In this system, the lowest-order prominent layer-line reflection at d = 38.7 nm (6th of 232 nm) appears at 6 voxels away from the equator.

A model structure of the sarcomere constructed in this way is shown in Fig. 7A. From this structure, a rotationally averaged 2-D structure factor can be calculated (Fig. 7B). This 2-D structure factor is used as observed data in the following 3-D CDI calculation.

The 3-D CDI starts with a starting structure shown in Fig. 7C. In this structure, actin and myosin filaments are at correct positions of the lattice, but they are featureless cylinders without any axial periodicities. In the rotationally averaged 2-D structure factor of this starting structure, only equatorial reflections are present (Fig. 7D). It will be tested whether the periodical structure as in Fig. 6A is restored on the starting structure after cycles of 3-D CDI calculations.

Compared with the simple 3-D CDI that has been explained in the previous section, the restoration of the 3-D structure from a rotationally averaged 2-D structure factor contains 2 additional steps in the reciprocal steps: rotational averaging of the 3-D structure factor (Fig. 8, D to E) and the restoration of the 3-D structure factor from the rotationally averaged 2-D structure factor (Fig. 8, F to G). The former step is easy, but the latter needs attention. Here, we adopted a procedure used by Grant (2018) for restoration of the 3-D structure of proteins from their 1-D solution scattering data. In the 3-D reciprocal space, he scaled all the data at a specific radius from the origin to the observed amplitude. Likewise, we scaled all the data at a specific radius from the Z axis to the observed amplitude. This correction was made within a single X-Y plane and repeated for all Z levels.

Results

Figure 9 shows the time course of the convergence of calculation. The Fourier amplitude error (difference between observed and restored structure factors) was quickly reduced to less than 20% within ~10 iterations. After this, the error continued to decrease while oscillating, and after 500 iterations, the error was reduced to 1~2%. Figure 10 compares the correct and restored 2-D structure factors, and the side views of the correct and restored lattice structure after 500 iterations. The structure factors are visibly indistinguishable from each other. The restoration of the lattice structure is not perfect, but the basic periodic structures seem to have been reasonably well restored on both actin and myosin filaments. As the error kept decreasing, even a better structure would have been restored after more than 500 iterations. The surface-rendered views of the correct, starting and restored 3-D structures of the filament lattice are shown in Fig. 11. By using a Xeon-based workstation, the CPU time for 500 iteration was about 48 hr.

Discussion

In this paper, we have demonstrated that, at least in-silico, the 3-D lattice structure of the sarcomere of insect flight muscle can be reasonably well restored from a rotationally averaged 2-D structure factor by using 3-D CDI, after optimization for 3-D FFT

calculations. It is considered that the successful restoration is partly due to the prior knowledge about the lattice structure and the helical symmetries of filaments in insect flight muscle, as documented by Tregear et al. (1998, 2004). We have previously shown that the 3-D structure of a fibrous material can be restored from its rotationally averaged 2-D structure factor by using a completely different approach, i.e., using a cylindrically averaged Patterson function (Iwamoto, 2021). This approach will work for a single-component system, but the sarcomeric structure of muscle contains two kinds of filaments, and its cylindrically averaged Patterson function will be too complex to be solved. Therefore, 3-D CDI is considered to be better suited for the sarcomeric structure.

The tested size of the rotationally averaged structure factor is 256 x 256 pixels, and it is not a satisfactory size to express all the features recorded in the real diffraction pattern. Experimentally, a size of 1024 x 1024 pixels or greater is desirable to record all the fine structure of diffraction patterns from muscle, and a much greater size is required if one wishes to record the "interference fringe" (Reconditi et al., 2004, Huxley, 2004, Huxley et al., 2006a,b) that comes from a longer periodicity of myosin filament.

Increasing the size of the diffraction pattern and the 3-D workspace would improve the accuracy of structure restoration, but it would also increase the CPU time and memory consumption. With a 512 x512 size, it would take 16 days to finish 500 iterations using our system. In our system, an array of 512 x 512 x 512 elements can be declared without an error, but a greater array may invoke an out-of-memory exception. To avoid these problems, further optimization of both software and hardware is required.

Applying this method to actually recorded 2-D diffraction patterns is the next challenge. Unlike the 2-D structure factor calculated from the model structure here, actually recorded diffraction patterns have many flaws, such as missing pixels around the beam position due to the beamstop. The intensity of scattering at the beam positon is important in recovering the 3-D reconstruction, and it must be estimated in some way. Other flaws include the disorder of the structure that limits the number of recorded reflection, resulting in a decrease of spatial resolution. Among the insect species that we have examined so far, the flight muscle of the giant waterbug *Lethocerus* gives the highest-quality diffraction patterns. We are currently trying to reconstruct the 3-D structure of the sarcomere by using these diffraction patterns, and if it is successful, the result will be presented elsewhere.

Acknowledgements

This work was supported by JSPS Kakenhi Grant No. 19K06777 to H.I. The diffraction pattern as shown in Fig. 1 was recorded under approval of SPring-8 Proposal Review

Committee (proposal No. 2007A1190).

References

Bates, R. H. T. (1982). Optik, 61, 247-262.

Grant, T. D. (2018). Nature Methods, 15, 191-193.

Huxley, H., Reconditi, M., Stewart, A. & Irving, T (2006). J. Mol. Biol. 363, 743-761.

Huxley, H., Reconditi, M., Stewart, A. & Irving, T (2006). J. Mol. Biol. 363, 762-772.

Huxley, H. E. (2004). Phil. Trans. R. Soc. B, 1879-1882.

Iwamoto, H. (2011). Biophysics, 7, 21-28.

Iwamoto, H. (2021). IUCrJ, 8, 544-548.

Marchesini, S., He, H., Chapman, H. N., Hau-Riege, S. P., Noy, A., Howells, M. R., Weierstall, U. & Spence, J. C. H. (2003). Phys. Rev. B, 68, 140101(R)

Miao, J. & Sayre, D. (2000). Acta Cryst. A56, 596-605.

Miao, J., Charalambous, P., Kirz, J. & Sayre, D. (1999). Nature, 400, 342-344.

Miao, J., Sayre, D. & Chapman, H. N. (1998). J. Opt. Soc. Am. A 15, 1662-1669.

Reconditi, M., Linari, M., Lucii, L., Stewart, A., Sun, Y. -B., Boesecke, B., Narayanan, T., Fischetti, R. F., Irving, T., Piazzesi, G., Irving, M. & Lombardi, V. (2004). Nature, 428, 578-581.

Tregear, R. T., Edwards, R. J., Irving, T. C., Poole, K. J. V., Reedy, M. C., Schmitz, H., Towns-Andrews, E. & Reedy, M. K. (1998). Biophys. J. 74, 1439-1451.

Tregear, R. T., Reedy, M. C., Goldman, Y. E., Taylor, K. A., Winkler, H., Franzini-Armstrong, C., Sasaki, H., Lucaveche, C. & Reedy, M. K. (2004). Biophys. J. 86, 3009-3019.

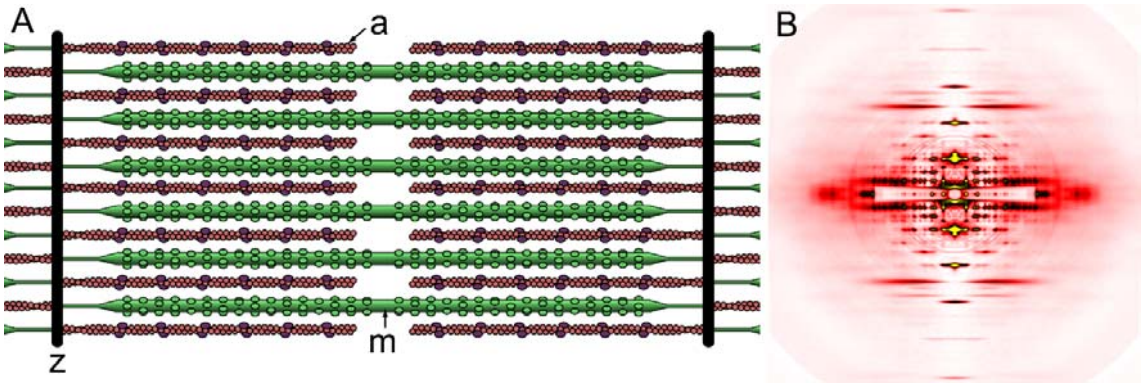

Figure 1. Structure of a sarcomere in insect flight muscle and an X-ray diffraction pattern recorded from it. A, schematic diagram of sarcomeric structure; B, diffraction pattern recorded from the flight muscle of a giant waterbug, *Lethocerus deyrolei*. a, actin filament; m, myosin filament; z, Z-membrane that separates two neighboring sarcomeres. The diagram in A is adopted from Iwamoto, 2011.

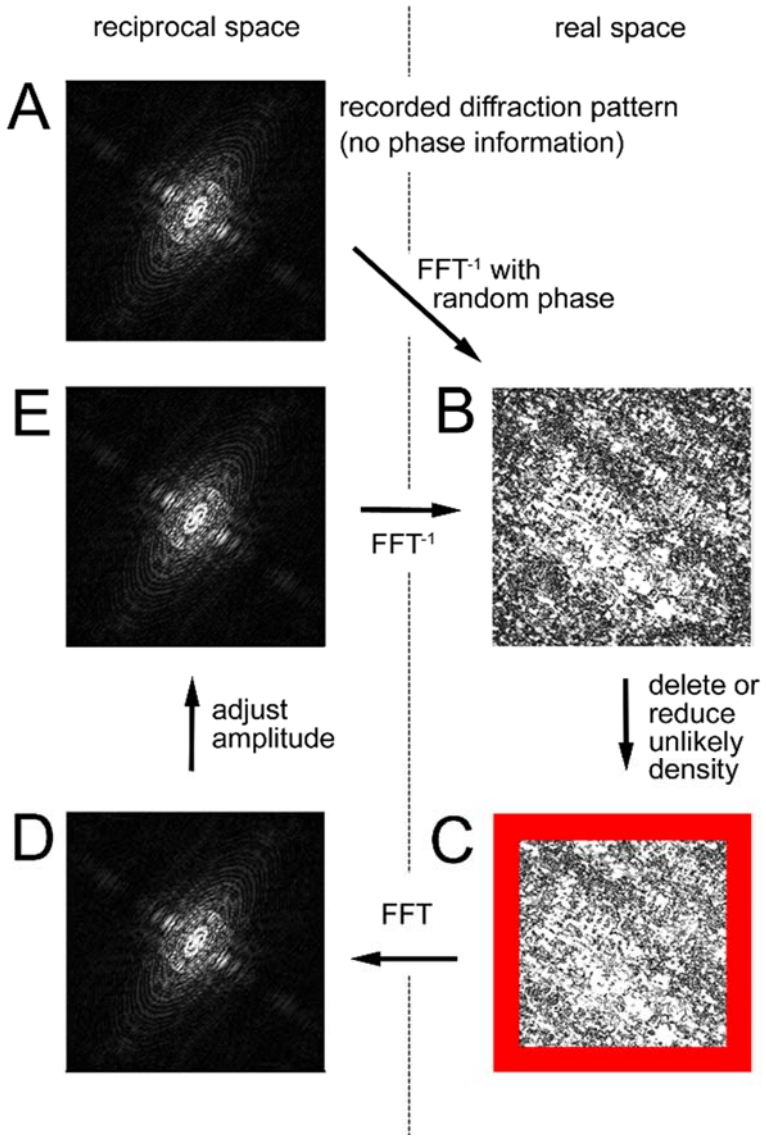

Figure 2. Principle of 2-D CDI. The recorded structure factor (A) is given a random phase and is subjected to inverse Fourier transformation to generate a real-space image (B). After unlikely densities are corrected (C), the image is subjected to Fourier transformation to obtain a structure factor (D). After its amplitude is scaled to the observed values (E), it is subjected to inverse Fourier transformation again (B). The calculations are iterated in the order of B-C-D-E until the error is minimized (convergence). In (C) the red area is outside the support.

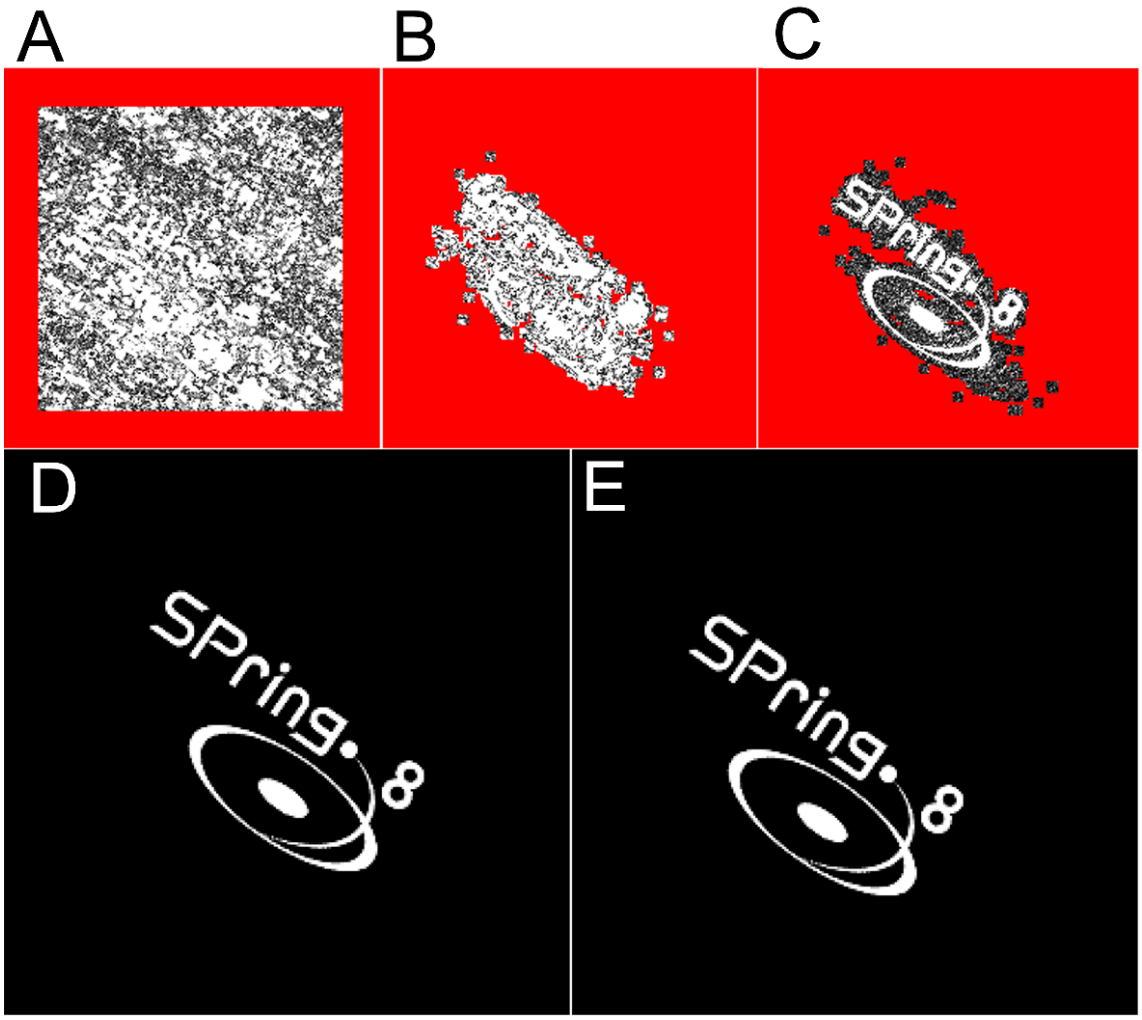

Figure 3. In-silico demonstration of 2-D CDI by using the SPring-8 logo as a sample. (A), support are immediately after the start of iterations (the red are is outside the support). (B), support area reduced by the shrink-wrap algorithm at 250 iterations. (C), status of calculation at 290 iterations when the calculation started to converge. (D), right answer; (E), reproduced image after 600 iterations.

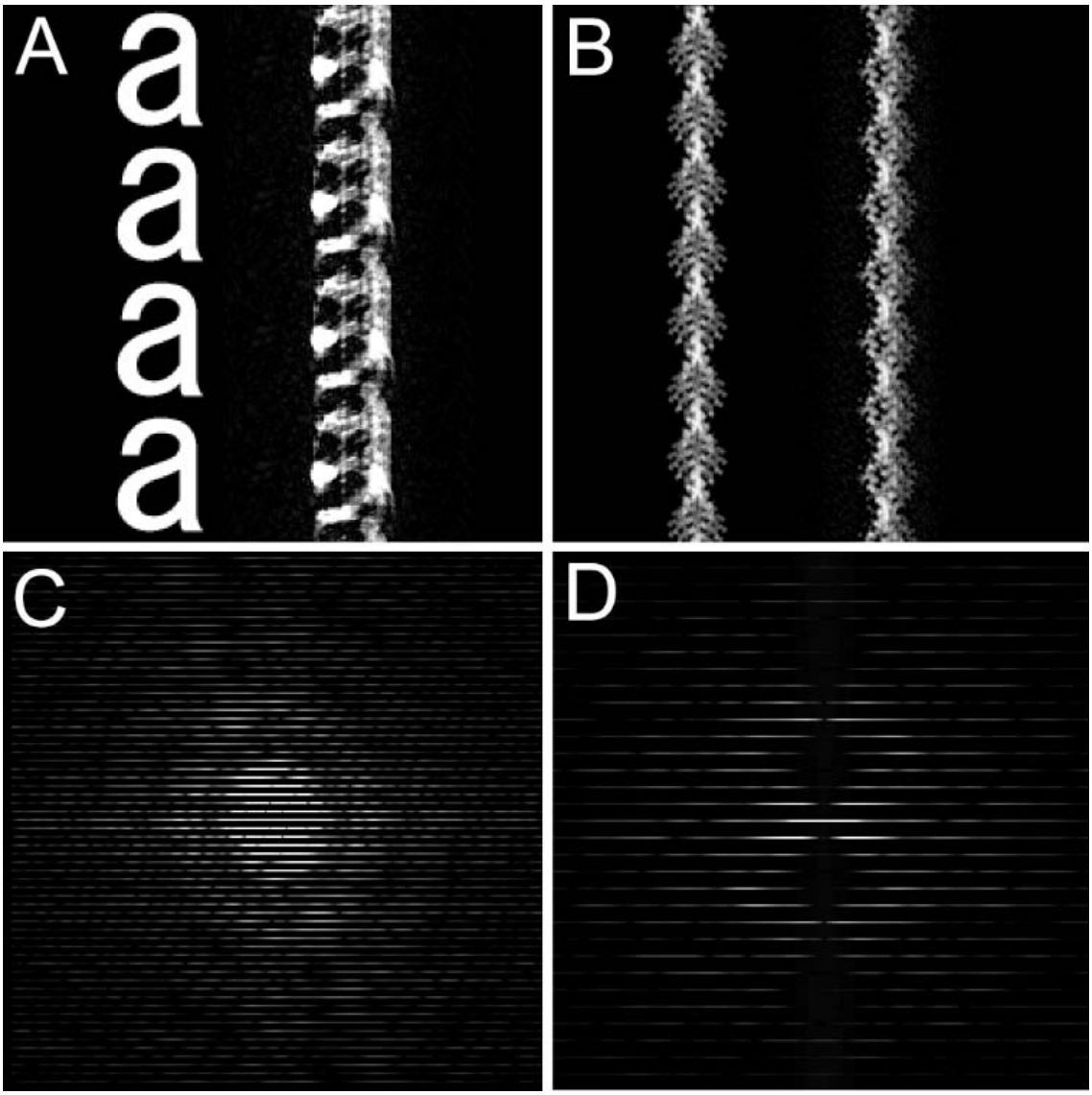

Figure 4. 2-D CDI applied to fibrous materials with periodical structures (in-silico experiments). (A), vertical row of letter a. The right image is the unsuccessful result of CDI calculation (the calculation never converges). (B), 2-D projection of decorated actin (see text for details). The result of CDI calculation is again unsuccessful. (C) and (D) are the structure factors calculated from the left images of (A) and (B), respectively.

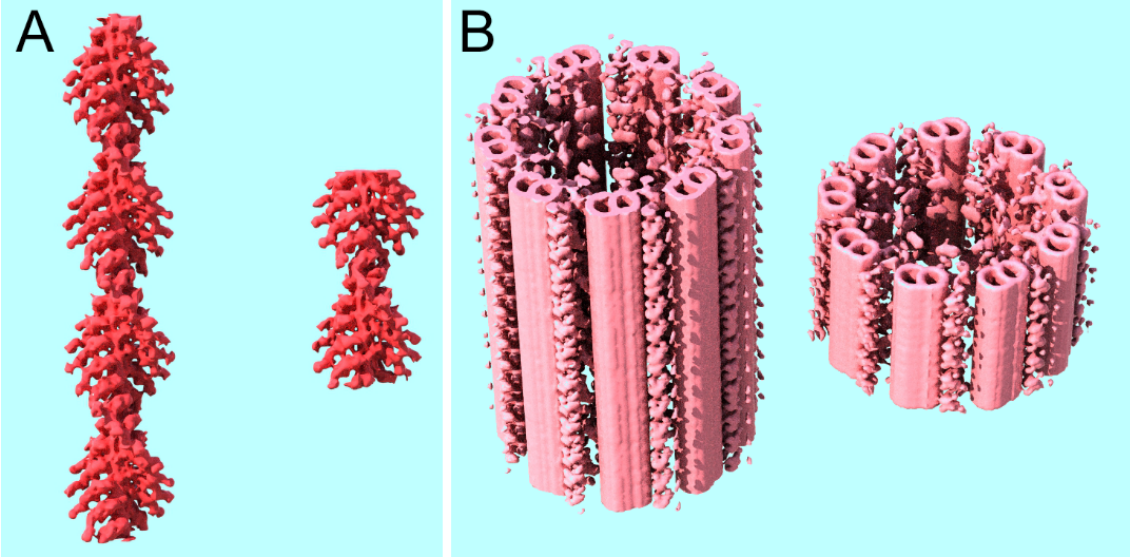

Figure 5. Demonstration that 3-D CDI works for fibrous materials with periodic structures. (A), 3-D model of decorated actin; (B) 3-D model of an axoneme of eukaryotic flagella. In each, the left image is the surface-rendered view of the model and the right image is that of the restored structure.

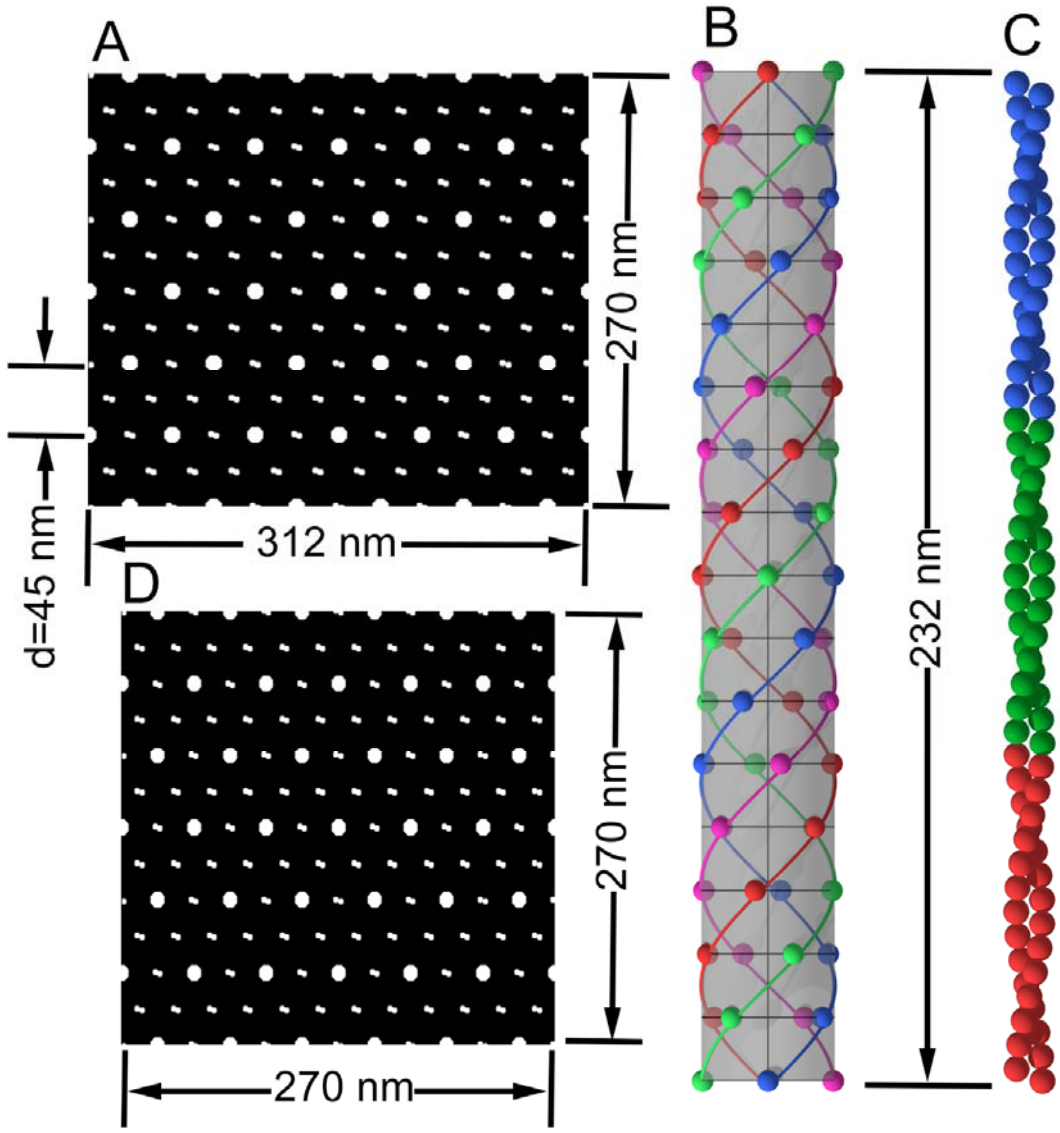

Figure 6. Optimization of the model of sarcomeric structure for fast Fourier transform (FFT). In order for FFT to work properly, the 3-D calculation space should be of the size of 2^n x 2^n x 2^n (n is integer) and it should contain integer number of repeating units in all of X, Y and Z directions. In the X-Y plane (A, across the axes of myosin and actin filaments), an area of 270 x 312 nm can contain integer number of repeating units in both X and Y directions when the d-spacing of the unit cell is 45 nm. This area is distorted to 270 x 270 nm (B). In the Z direction, both myosin (C) and actin (D) filaments have a periodicity of 232 nm, and they are extended to 270 nm. In this way, the sarcomeric model can be accommodated in a 2^n x 2^n x 2^n cube. In (A) and (B), the larger circles are myosin filaments, and the smaller circles are actin filaments.

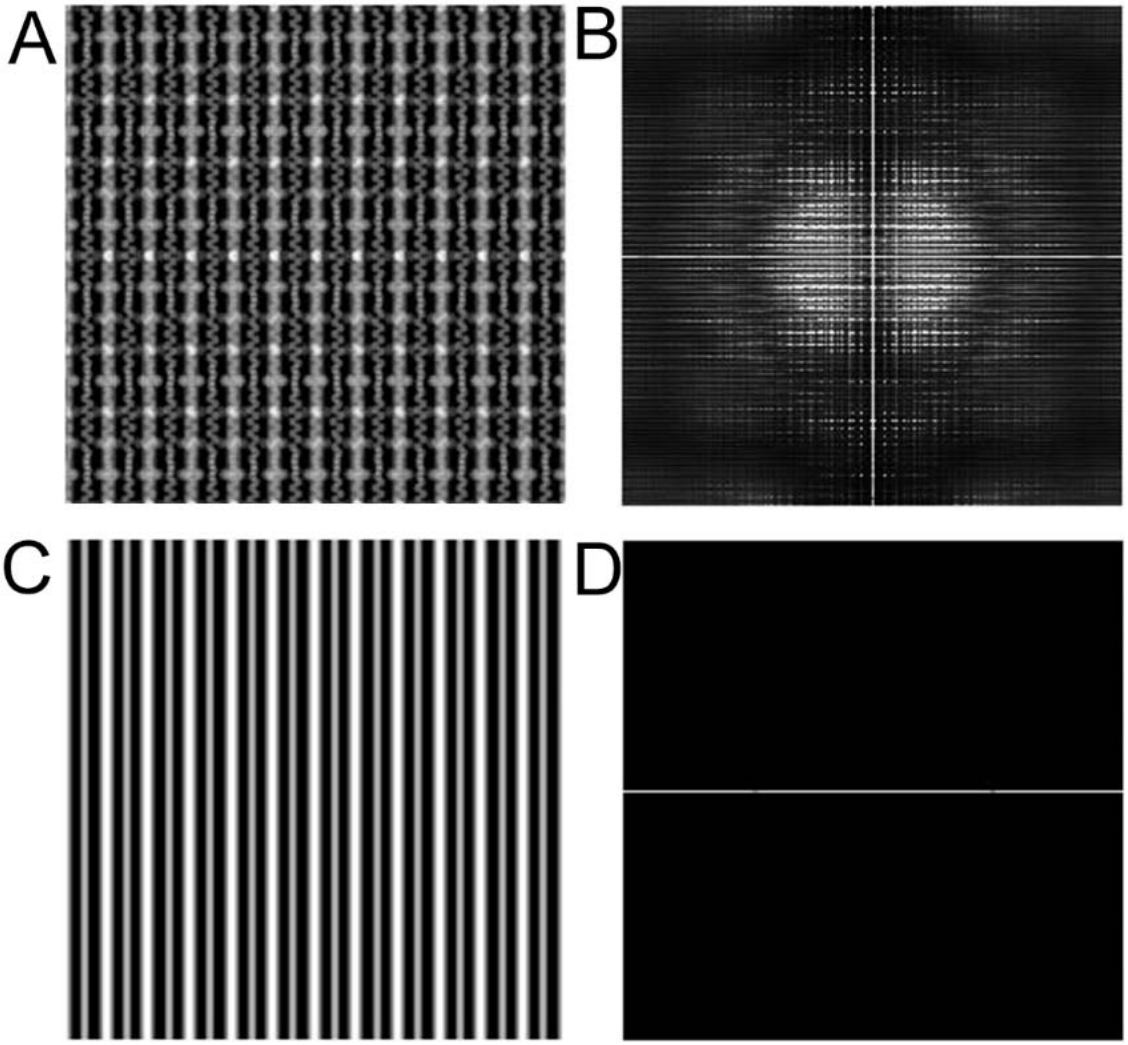

Figure 7. Models of sarcomeric structure of insect flight muscle. (A), side view of the 3-D model built according to the procedure shown in Fig. 6, by using the known arrangement of hexagonal lattice of actin and myosin filaments, and the known helical symmetries of individual actin and myosin filaments. The space is 256 x 256 x 256 voxels. (B), rotationally averaged 2-D structure factor calculated from (A). This pattern is used as an observed structure factor, without phase information, in the following 3-D CDI reconstruction. (C), side view of the starting model structure. This has an arrangement of actin and myosin filaments in a hexagonal lattice identical to that in (A). The individual actin and myosin filaments are expressed as featureless cylinders. (D), rotationally averaged 2-D structure factor calculated from (C), containing only the equatorial reflections. In the following 3-D CDI calculations, the structure in (C) will be modified so that its structure factor approaches that in (B).

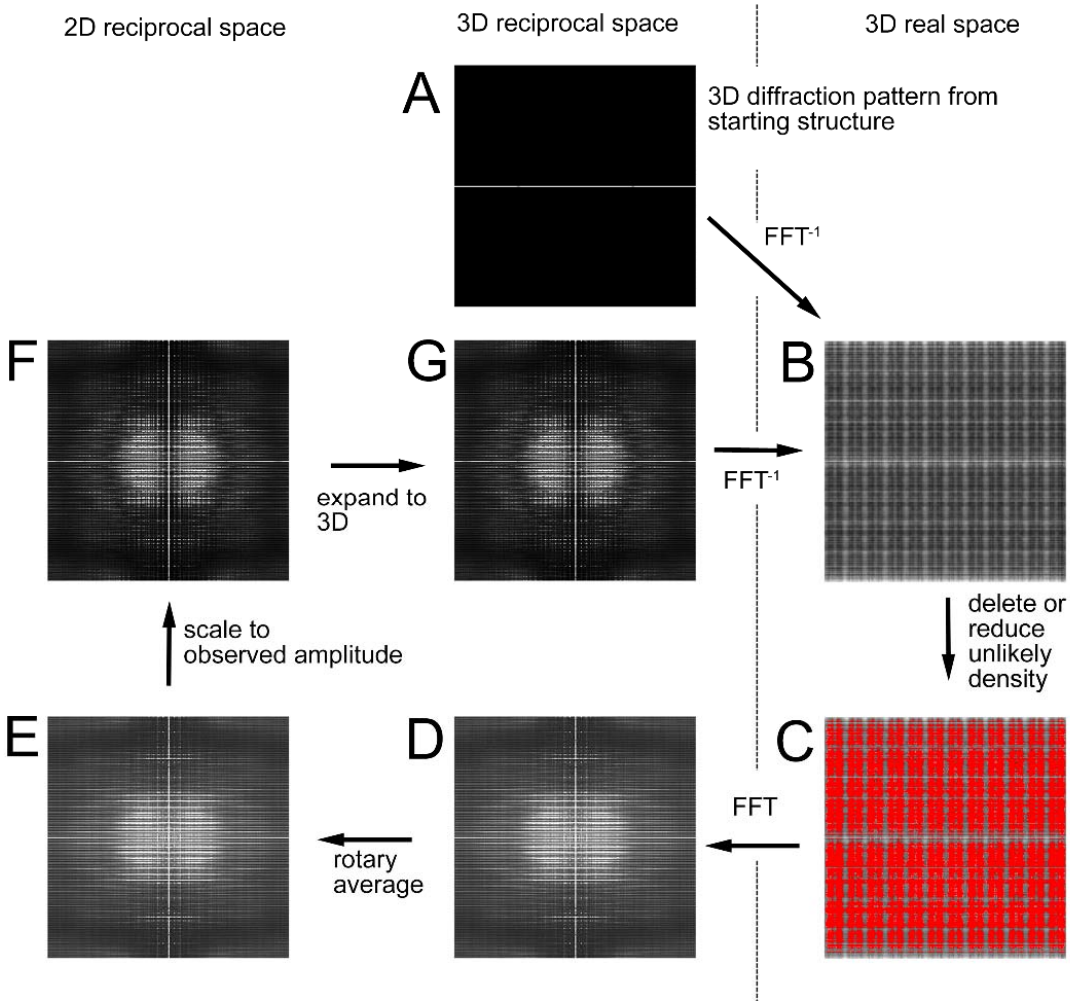

Figure 8. Procedure of 3-D CDI calculations involving rotationally averaged 2-D structure factor. The procedure is basically identical to that in 2-D CDI as shown in Fig. 2, except that it has two additional steps, i.e., rotational averaging of the 3-D structure factor (D to E) and the expansion of the corrected 2-D structure factor to 3-D (F to G). Instead of giving random phase information to the observed structure factor, a structure factor (A) is calculated from the starting structure (Fig. 7C), and the initial real-space data are the starting structure itself.

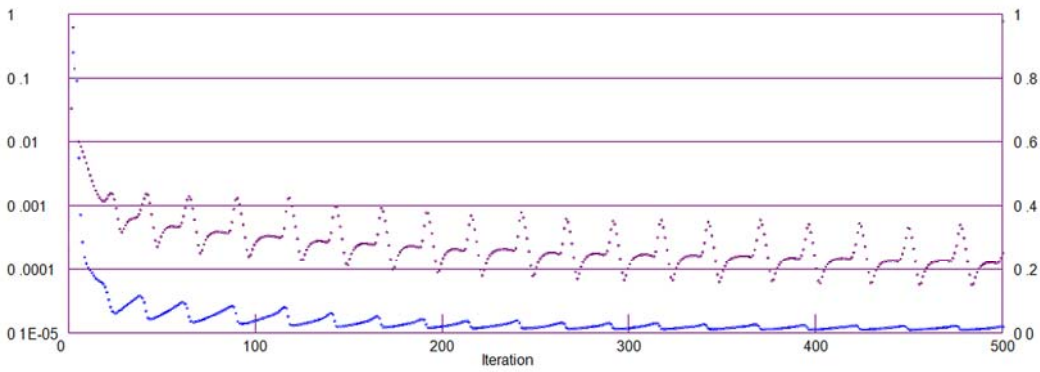

Figure 9. Time course of the convergence of the 3-D CDI calculation. The Fourier amplitude error in the reciprocal space (blue, linear scale on the right vertical axis) quickly decreased within ~10 iterations, and after 500 iterations, it was 1-2 %. The magenta data represent the density outside the support in the real space (logarithmic scale on the left vertical axis), which was reduced to the level of 0.01% after 500 iterations.

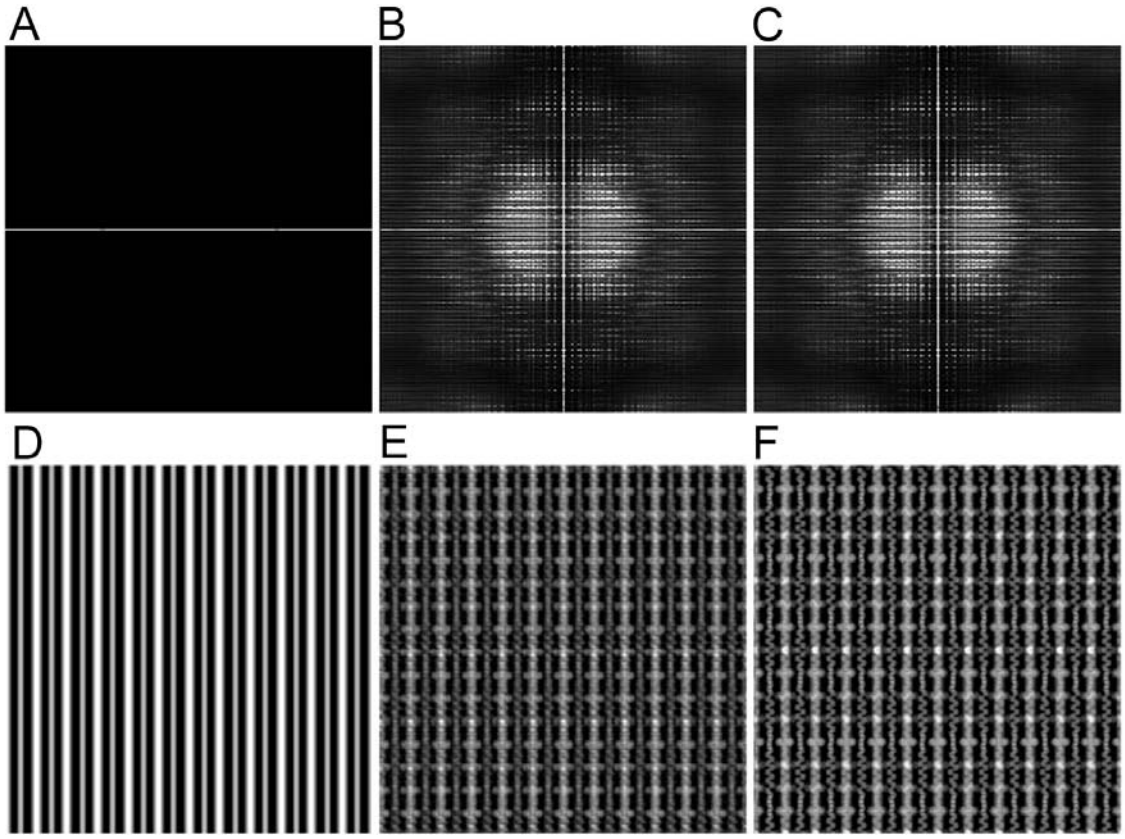

Figure 10. Comparison between correct and restored structure factors and the 3-D sarcomeric lattice structures. (A), rotationally averaged 2-D structure factor calculated from the starting structure (D, Fig. 7C). (B), 2-D structure factor restored after 500 iterations. (C), correct structure factor calculated from the model structure (F, Fig. 7A). (D-F), side views of the starting structure, restored structure and the correct model structure, respectively.

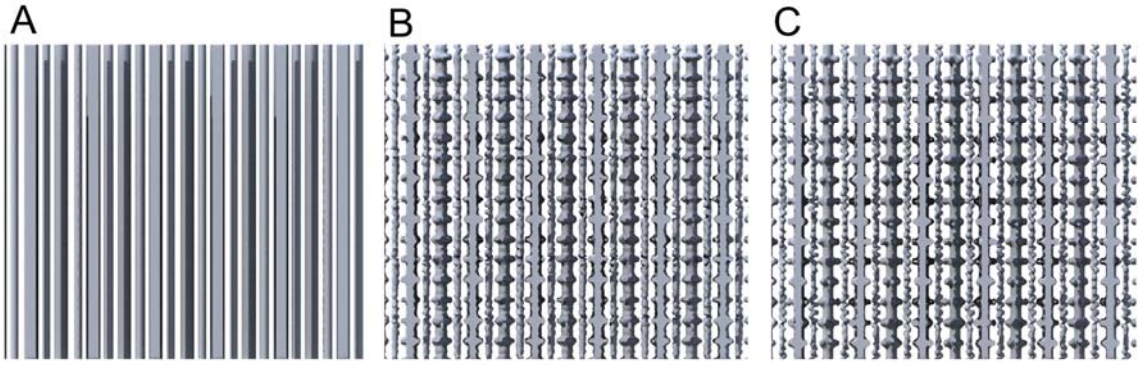

Figure 11. Surface-rendered views of the real-space 3-D structures. (A), starting structure; (B), structure restored after 500 iterations; (C), correct model structure.