論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Three-dimensional structural analysis of Ca2+-ATPase from sarcoplasmic reticulum by cryo-electron microscopy
著者(和文)	米倉功治
Author(English)	Yonekura Koji
出典(和文)	学位:博士(理学), 学位授与機関:東京工業大学, 報告番号:甲第3482号, 授与年月日:1997年3月26日, 学位の種別:課程博士, 審査員:
Citation(English)	Degree:Doctor (Science), Conferring organization: Tokyo Institute of Technology, Report number:甲第3482号, Conferred date:1997/3/26, Degree Type:Course doctor, Examiner:
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

Three-dimensional Structural Analysis of Ca^{2+}-ATPase from Sarcoplasmic Reticulum by Cryo-electron Microscopy.

Koji Yonekura

Submitted to Tokyo Institute of Technology
in partial fulfillment of the requirements
for the Degree of Doctor of Science

Thesis supervisors :
Professor Shunsuke Hirotsu
and
Professor Chikashi Toyoshima
(the University of Tokyo)

March 1997

Contents

0. Abstract

The location of the ATP-binding site of a P-type ion pump, Ca^{2+}-ATPase from rabbit sarcoplasmic reticulum, was examined by cryo-electron microscopy. A non-hydrolyzable analogue of ATP, β,γ-bidentate chromium (III) complex of ATP (CrATP) was used to stabilize the enzyme in the Ca^{2+}-occluded state. Tubular crystals were then induced by vanadate in the presence of EGTA, keeping CrATP bound to the enzyme. The three-dimensional structures of the crystals were determined at 14 Å resolution by cryo-electron microscopy and helical image analysis. Statistical comparison of the structures with and without CrATP showed clear and significant differences at the groove proposed previously as the ATP-binding pocket.

We have also put effort to improve the resolution and obtained a preliminary three-dimensional map of CrATP-bound tubes at better than 10 Å resolution. Two pores of ~ 10 Å in length were resolved near the cytoplasmic and the lumenal surface of the membrane. They were staggered and not connected each other due to presence of the strong densities located at the middle of the membrane. Because the enzyme in the tubular crystals with bound CrATP is likely to be in the occluded state, it is reasonable to consider the two pores as the inlet and outlet of the ion pathway. It was also possible to trace 6 columns surrounding the pores. Since B and C segments comprised two columns for each, the structure obtained here, gives a support to the model involving 10 transmembrane helices.

I. Introduction

Sarcoplasmic reticulum (SR) Ca^{2+}-ATPase is an integral membrane protein of 110 kDa that pumps Ca^{2+} into SR against a large concentration gradient during relaxation of muscle cells. To do so, it uses the chemical energy of ATP and transports two Ca^{2+} per ATP hydrolysis. Ca^{2+}-ATPase is one of the members of the large P-type ion pump family, which includes Na^+/K^+-ATPase, gastric H^+/K^+-ATPase and yeast H^+-ATPase among others (for a recent review see Møller et al., 1996). These pumps share considerable sequence homologies, similar reaction cycles (Fig. I.1) and common domain structures. The proposed structural model for Ca^{2+}-ATPase consists of two main cytoplasmic domains connected by a narrow stalk to 10 transmembrane helices (M1 ~ M10) (Fig. I.2; MacLennan et al., 1985; Brandl et al., 1986; Taylor and Green, 1989). Mutational analyses have revealed that the critical residues for Ca^{2+} binding are located at the middle of putative transmembrane helices (Clarke et al., 1989; Chen et al., 1996) and that those for ATP hydrolysis are in the larger cytoplasmic domain between M4 and M5 (Maruyama et al., 1989; Clarke et al., 1990a; Vilsen et al., 1991). A number of chemical labels have been also used to locate the residues involved in the ATP binding (for a review see Bigelow and Inesi, 1992). The domain for ATP binding has been suggested to fold like nucleotide-binding domains of certain kinases (Taylor and Green, 1989), in which two sub-domains connected by a hinge form a nucleotide-binding pocket. The smaller cytoplasmic domain between M2 and M3 presumably consists of anti-parallel β-sheets (MacLennan et al., 1985) and is responsible for coupling of ATP hydrolysis to ion transport (Andersen et al., 1989; Clarke et al., 1990b).

However, direct three-dimensional information for Ca^{2+}-ATPase is very limited. The highest resolution three dimensional structure so far obtained is only at ~ 14 Å resolution (Toyoshima et al., 1993a). The resolution of this structure, obtained by cryo-electron microscopy from tubular crystals formed in the absence of Ca^{2+} and ATP, was obviously too low for unambiguous location of functional sites. Nevertheless, we proposed a groove in the cytoplasmic domain as a candidate for the ATP-binding site (Fig. I.3), based on the results from affinity labeling (see Green and Stokes, 1992, for a review). If correct, this location implies that the structural effects of Ca^{2+} binding are transferred more than 40 Å to the ATP-binding site. The hydrolysis of ATP and phosphorylation of the enzyme in turn alter the conformation of the Ca^{2+}-binding sites to release Ca^{2+} into the lumen of SR. Therefore, knowledge of the precise position of the ATP-binding site is a critical step in characterizing the long range interactions within the enzyme and, ultimately, in understanding the mechanism of active transport.

To locate the ATP-binding site directly and to investigate structural changes induced by ATP binding, we used β,γ-bidentate chromium (III) complex of ATP (CrATP) (Fig. I.4). This non-hydrolyzable analogue of ATP has been used to stabilize the enzyme in the Ca^{2+}-occluded state (Fig. I.1; Serpersu et al., 1982; Vilsen and Andersen, 1987). Though the tubular crystals can be formed retaining CrATP bound (Buhle et al., 1983; Stokes and Lacapere, 1994a), no three-dimensional structures have been obtained so far. We describe at first the three-dimensional structure of Ca^{2+}-ATPase with bound CrATP and the differences caused by the binding of CrATP.

We have also put effort to improve the resolution of the structure. So far, using cryo-electron microscopy, two atomic structural models of membrane proteins have been obtained; they are bacteriorhodopsin (Henderson et al., 1990) and plant light-harvesting complex II (Kühlbrandt et al., 1994). These structures have been determined from well-ordered two-dimensional crystals. Howerver, from tubular crystals, only the structure of acethylcholine receptor has been solved at better than 10 Å resolution (Unwin et al., 1993).

To improve the resolution of the tubular crystals of Ca^{2+}-ATPase, we have developed many new programs, especially, for recovering the data outside zero of the contrast transfer function (CTF) of electron microscope. Parts of them have been applied to tubular crystals of cytochrome oxidase (Sato et al., unpublished data), those of cytochrome bc_1 complex (Akiba et al., 1996) and bacterial flagella filament (Mimori et al., 1995).

By using these programs and averaging of 45 images, we determined the structure of Ca^{2+}-ATPase with bound CrATP at better than 10 Å resolution and describe the appearance of the ATP-binding site, the stalk section and the transmembrane region.

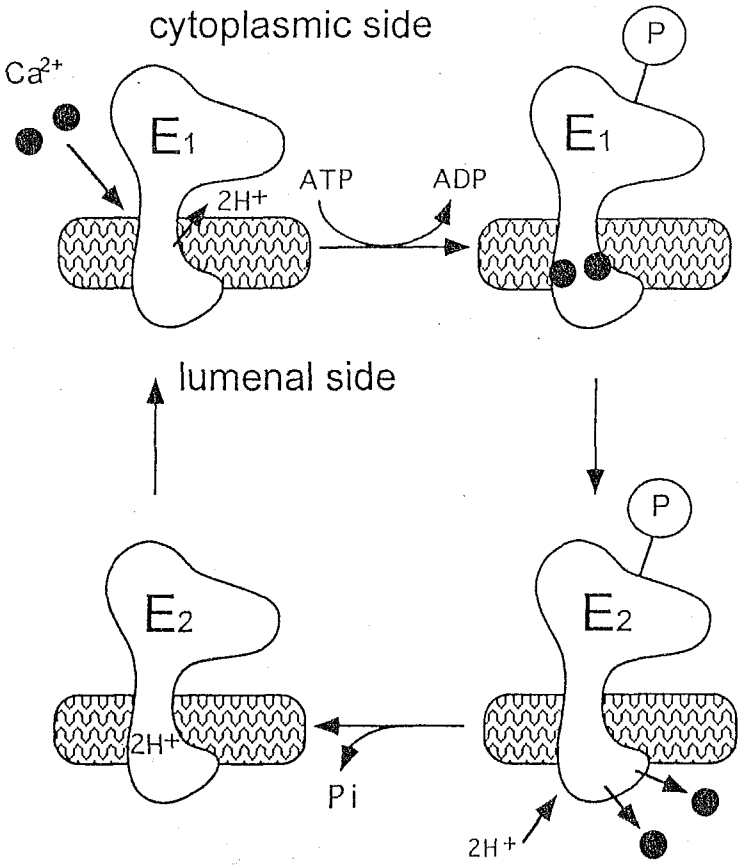


FIGURE I.1 Reaction cycle of Ca^{2+}-ATPase. By hydrolysis of one ATP, it transports two Ca^{2+}s. During transport of Ca^{2+}, Ca^{2+} is temporally trapped (upper right). This state is called as 'occluded' state.

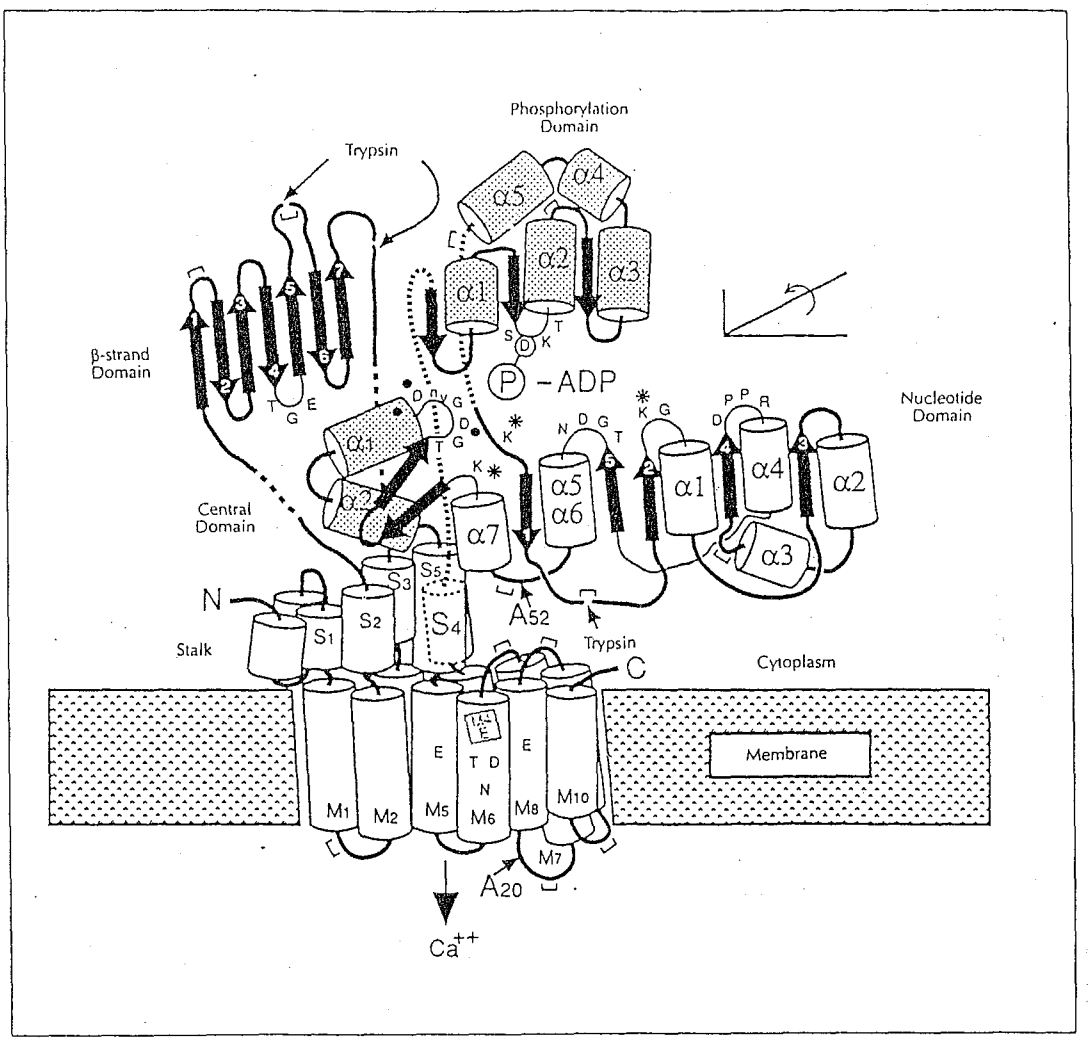


FIGURE I.2 Secondary structure prediction of Ca^{2+}-ATPase. Reprinted from Green and Stokes (1992).

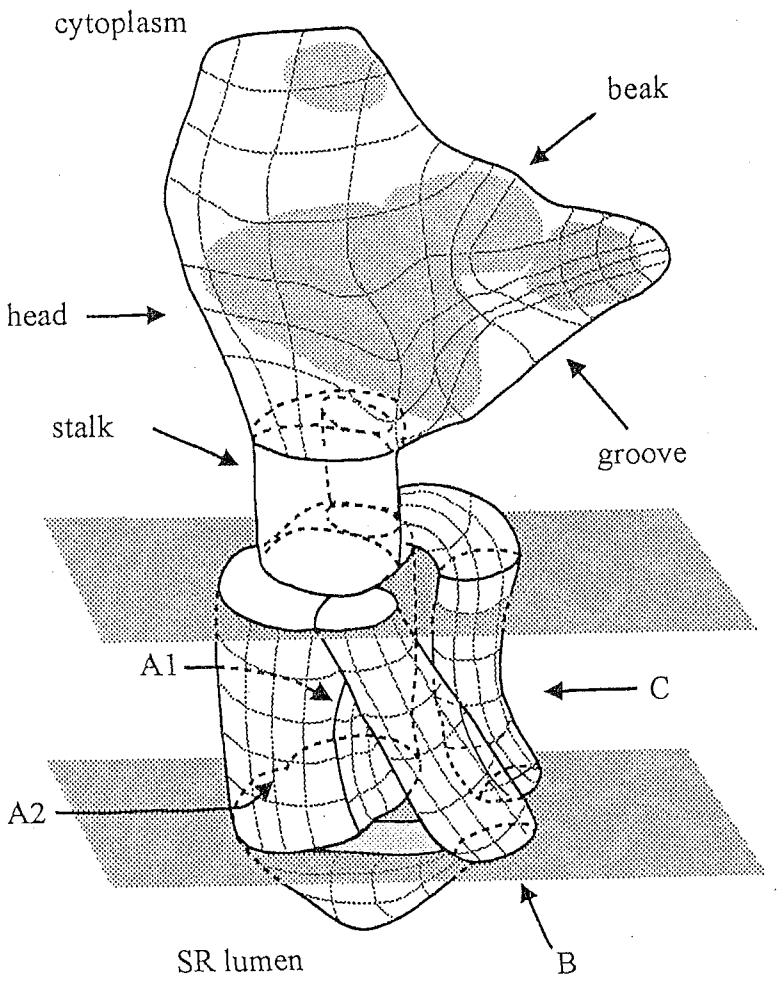


FIGURE I.3 Cartoon of the structure of Ca^{2+}-ATPase in the absence of ATP and Ca^{2+} as determined by cryo-electron microscopy. The surfaces of the bilayer are indicated by the two shadow planes. A large cytoplasmic domain appears like a head of a bird and is linked to the membrane domain through a stalk region. The membrane domain consists of three segments, A, B and C. A is further divided to subsegments, A1 and A2. The ATP-binding site is postulated to be at the 'groove'. The ion pathway in the membrane is proposed to be located at the interface between A1 and A2. Reproduced from Toyoshima et al. (1993a).

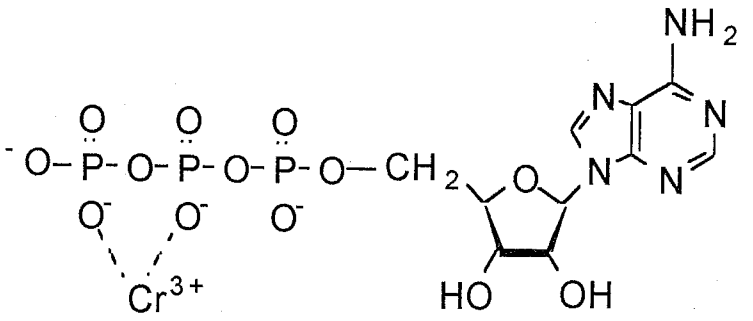


FIGURE I.4 β,γ-bidentate chromium (III) ATP (CrATP).

II. Materials and Methods

II-1. Preparation of sarcoplasmic reticulum, CrATP and decavanadate

SR was prepared from the white muscle of rabbit legs by the method of Meissner et al. (1973). Protein concentrations were determined using bicinchoninic acid reagent (Pierce Chemical Co., Rockford, IL; Smith et al., 1985) after SR was completely dissolved in 1% (w/v) deoxycholate solution. Protein concentrations were also determined by the absorbance at 280 nm with an extinction coefficient of 1.05 ($ml \cdot mg^{-1} \cdot cm^{-1}$) (Thorley-Lawson and Green, 1973). These two procedures generally agreed within 10% (Chen et al., 1991). β,γ-Bidentate CrATP was prepared as described by Mariano and Cleland (1980) and Cleland (1982). Decavanadate stock solution was prepared by adjusting the pH of 50 mM Na_3VO_4 to 2.0 and incubating overnight at 4°C (Varga et al., 1985); the pH was neutralized immediately before adding to the crystallization solution.

II-2. Tubular crystals of Ca^{2+}-ATPase with bound CrATP

SR (2 mg/ml) was reacted with CrATP in 100 mM KCl, 5 mM $MgCl_2$, 1 mM CrATP, 0.1 mM $CaCl_2$, and 10 mM imidazole, pH 7.4 at 25°C. After 4 hours, an equal volume of the same buffer containing 10 mM vanadate and 1-2 mM EGTA instead of Ca^{2+} was added and incubated at 4°C. Tubular crystals grew in 18-36 hours (Buhle et al., 1983; Stokes and Lacapere, 1994a).

II-3. Assay of the effect of CrATP

Inhibition of Ca^{2+}-ATPase activity by CrATP was measured by a coupled ATPase assay (Warren et al., 1974). Five to ten microliters of the sample solution was added to 800 µl of the assay medium (150 mM KCl, 7.5 mM $MgCl_2$, 0.15 mM $CaCl_2$, 75 mM 3-(*N*-morpholino)propanesulfonic acid-KOH, pH 7.0, 0.5 mM phosphoenolpyruvate, 0.25 mM NADH, 7.1 IU pyruvate kinase and 13.5 IU lactate dehydrogenase) and incubated for 5 min. at 37°C. After an addition of 10 µl of 100 mM ATP, the decrease of NADH was monitored by the absorbance at 340 nm, which corresponded directly to the rate of ATP hydrolysis.

Occlusion of $^{45}Ca^{2+}$ was assayed by filtration (Chen et al., 1991; Vilsen and Andersen, 1992). After reaction with CrATP in the presence of 0.05-0.1 mM $^{45}Ca^{2+}$, 50 - 100 µl of the sample was filtered (0.45 µm, HAWP025, Millipore Co., Bedford, MA) and then washed with 30 ml of cold buffer. The filter that had retained the SR was completely dissolved in 3 ml of liquid scintillation cocktail (ASC II, Amersham International plc, Buckinghamshire, England) and the amount of $^{45}Ca^{2+}$ was determined by a liquid scintillation counter (LSC-5100, Aloka Co., Tokyo, Japan). This method allowed us to measure only the tightly bound $^{45}Ca^{2+}$ that did not dissociate within ~ 1 minute.

II-4. Electron microscopy

Four microliters of specimen solution were deposited on a carbon coated holey grid and blotted with filter paper. The specimen was rapidly frozen by plunging into an ethane slush. A Gatan 626 cryo-holder (Pleasanton, CA) was used for examination in a

JEOL JEM-2000EX microscope (Tokyo, Japan) operated at 200 kV. Images were recorded on Kodak SO163 films at a nominal magnification of 40,000 × using a low dose kit. The objective lens current was monitored. The magnification was calibrated using negatively stained tropomyosin tactoids (Caspar et al., 1969) at various objective lens currents. Three images of tubes without CrATP were taken during the previous work (Toyoshima et al., 1993a) at a nominal magnification of 36,000 ×, using a Philips EM420T microscope (Mahwah, NJ) operated at 120kV.

II-5. Image analysis

Images of straight tubular crystals with the diameter of about 650 Å were selected and examined first by optical diffraction. Suitable images were digitized with 20 μm spot and step sizes with a flatbed microdensitometer (PDS1010M, Orbital Science, Pomona, CA). The amounts of defocus were estimated from the Thon rings that arise from the images of carbon film or tubular crystals themselves (Tani et al., 1996). The amplitude contrast of 4.6% was assumed (Toyoshima et al., 1993b). Helical reconstruction was carried out essentially as described by Toyoshima and Unwin (1990). Weakly (6500 - 9500 Å) and strongly (about 20,000 Å) underfocused pairs of images taken of the same fields were used for a partial compensation of uneven contrast transfer function (CTF) (Toyoshima and Unwin, 1990). This procedure was essential for comparison of individual images because good signal-to-noise ratio was required. The resolution was limited by the first CTF zero of the weakly defocused image.

a. Averaging of different images

Tubes belonging to the same helical family can usually be averaged in Fourier space by renumbering the layer-lines, provided that the unit cell parameters are virtually identical (Toyoshima and Unwin, 1990). However, with tubes formed in the absence of CrATP, a helical family consisted of two groups of different diameters. In this case, due to large differences in unit cell parameters, averaging had to be done in real space instead of Fourier space. To do so, first, the tubes belonging to the same group of similar unit cell parameters were averaged separately in Fourier space. Next, the mean radial density distribution (MRDD) of each group was calculated by Fourier-Bessel synthesis of the equatorial layer-line. Cross correlation between MRDDs from the two groups was used to determine their relative magnifications. After magnification correction, three-dimensional maps of both groups were calculated independently. Then the density scaling was adjusted by linear regression. Finally, after one molecule was cut out from the map at the level enclosing 100% of expected volume, it was brought to the same position by applying rotational and translational adjustments. Then the density maps of one molecule were averaged. For the tubes with bound CrATP, real space averaging was unnecessary because they all have similar unit cell parameters (Table III.1).

b. Comparison of the structures of Ca^{2+}-ATPase with and without CrATP

To investigate the differences in the structure with and without CrATP, a difference map was calculated by subtracting the three-dimensional map of the enzyme without CrATP from that with

CrATP. Furthermore, to examine the significance of the differences, Student's t-test was carried out (Milligan and Flicker, 1987) after density scaling of the individual images.

c. Calculation of the three-dimensional structure at better than 10 Å resolution

For calculation of the three-dimensional map including the data outside the first CTF zero, we need to average many images taken at various amounts of defocus so that all spacing are well sampled. We set a limit on the calculation by the points where the envelope function becomes < 1/10. Averaging of different images was done according to II-5a. The averaging procedure is summarized in Fig. II.1.

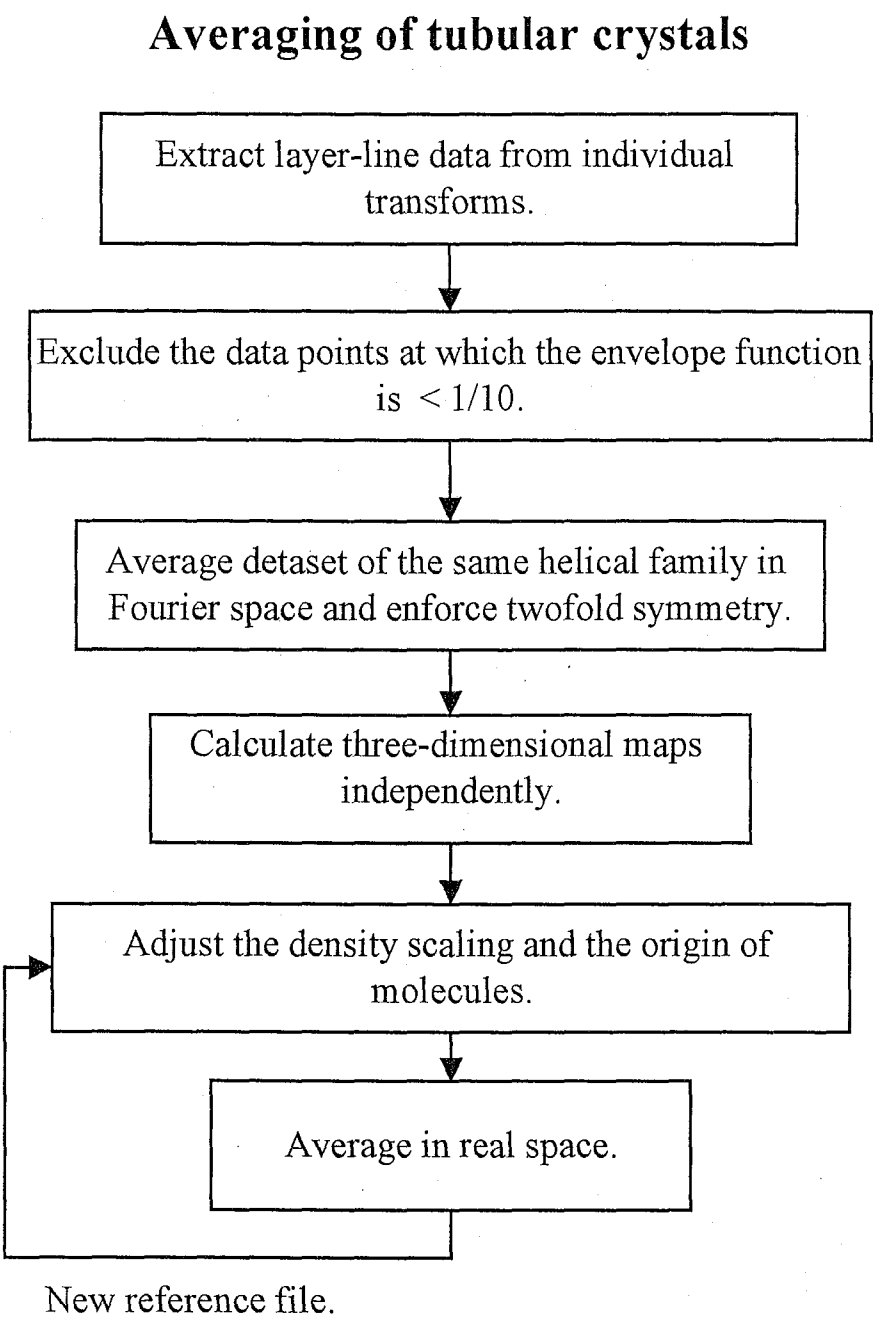


FIGURE II.1 Flow diagram of averaging of different images of tubular crystals.

To examine the quality of the map, twofold phase residual was used (Toyoshima and Unwin, 1988). To do so, the maps of individual helical families were calculated without twofold enforcement. Then two molecules related by the twofold axis were averaged separately over different helical families in real space, respectively. The resulting averaged map was converted to the layer-line data by Fourier-Bessel transformation. After the layer-line data were edited to use only good portions, twofold phase residual was calculated.

III. Results

III-1. The ATP-binding site of Ca^{2+}-ATPase revealed by electron image analysis.

a. The effect of CrATP on Ca^{2+}-ATPase

The binding of CrATP, a non-hydrolyzable analogue of ATP, was assessed by measuring the inhibition of the ATPase activity (filled circles in Fig. III.1). After 4 hours, ATPase activity was decreased to about 15% of the initial value. Since the Ca^{2+}-insensive ATPase activity was ~ 15%, the inhibition by CrATP was nearly complete. After an addition of vanadate to induce tubular crystals, the activity remained at the same level, consistent with Stokes and Lacapere (1994a). In contrast, with vanadate alone, the activity decreased to only ~ 60% of the initial value (open circle after 20 hours in Fig. III.1). Hence, CrATP must be bound to Ca^{2+}-ATPase in the tubular crystals.

$^{45}Ca^{2+}$ was trapped within Ca^{2+}-ATPase when CrATP was bound to the enzyme (filled boxes in Fig. III.1). After 4 hours, the amount of occluded Ca^{2+} was about 1.6 mol/mol SR protein. Considering that the proportion of Ca^{2+}-ATPase in total membrane proteins in SR is ~ 80%, the value means that most of the enzyme had two Ca^{2+}s bound per molecule. Since the removal of free Ca^{2+} was necessary for the formation of tubular crystals, EGTA was added as well as vanadate. As a consequence, when tubular crystals were formed, most of the occluded Ca^{2+} was released into buffer (filled box after 20 hours in Fig. III.1).

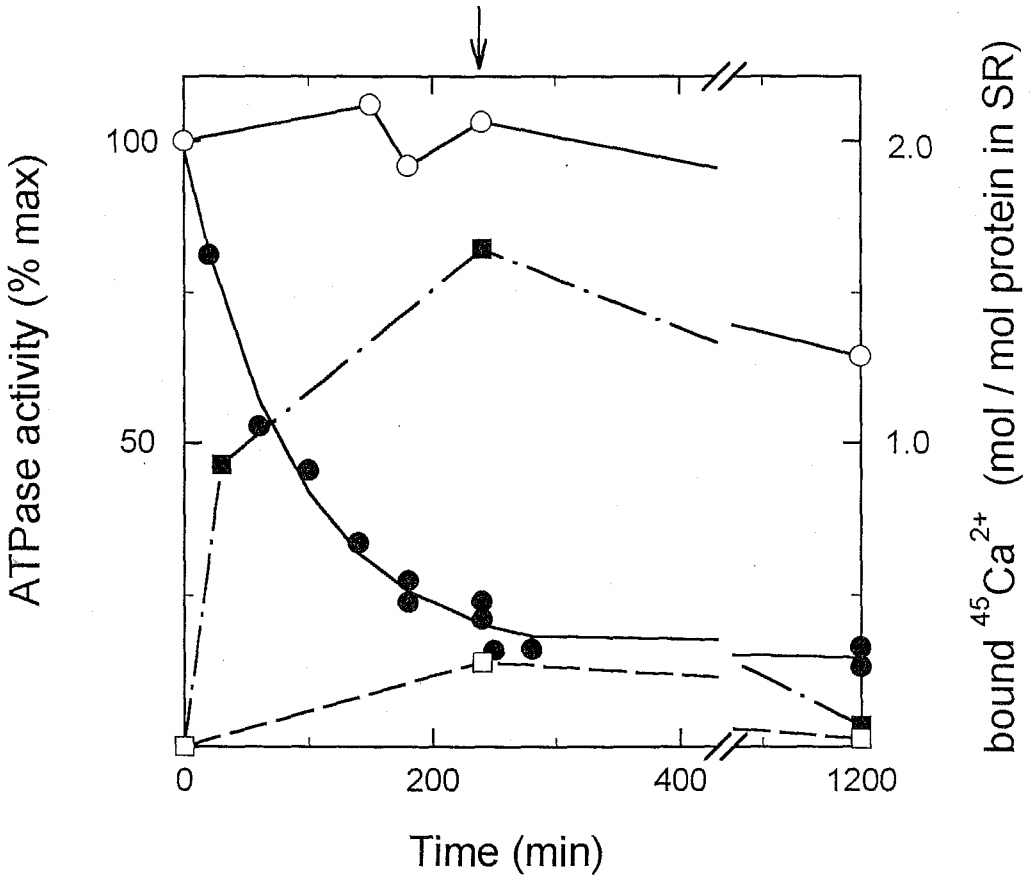


FIGURE III.1 The effects of CrATP on Ca^{2+}-ATPase: ATPase activity (circles), and amounts of occluded $^{45}Ca^{2+}$ (boxes). SR was incubated in the buffer containing CrATP (filled symbols) or without it (open symbols). The arrow at the top of this figure indicates the time when the crystallization buffer containing EGTA and vanadate was added.

b. Image analysis of tubular crystals of Ca^{2+}-ATPase

A diffraction pattern of an image of a tubular crystal with bound CrATP is shown in Fig. III.2(b). It consists of a series of layer-lines symmetrical about the meridian, showing the good preservation of helical symmetry. The qualities of the tubes with CrATP were indistinguishable from those without CrATP (cf. Fig. 1b of Toyoshima et al., 1993a), as is demonstrated by the presence of a near meridional layer-line with an axial spacing of 1/15.4 $Å^{-1}$ (indexed as (2,7; -1) in Fig. III.2b).

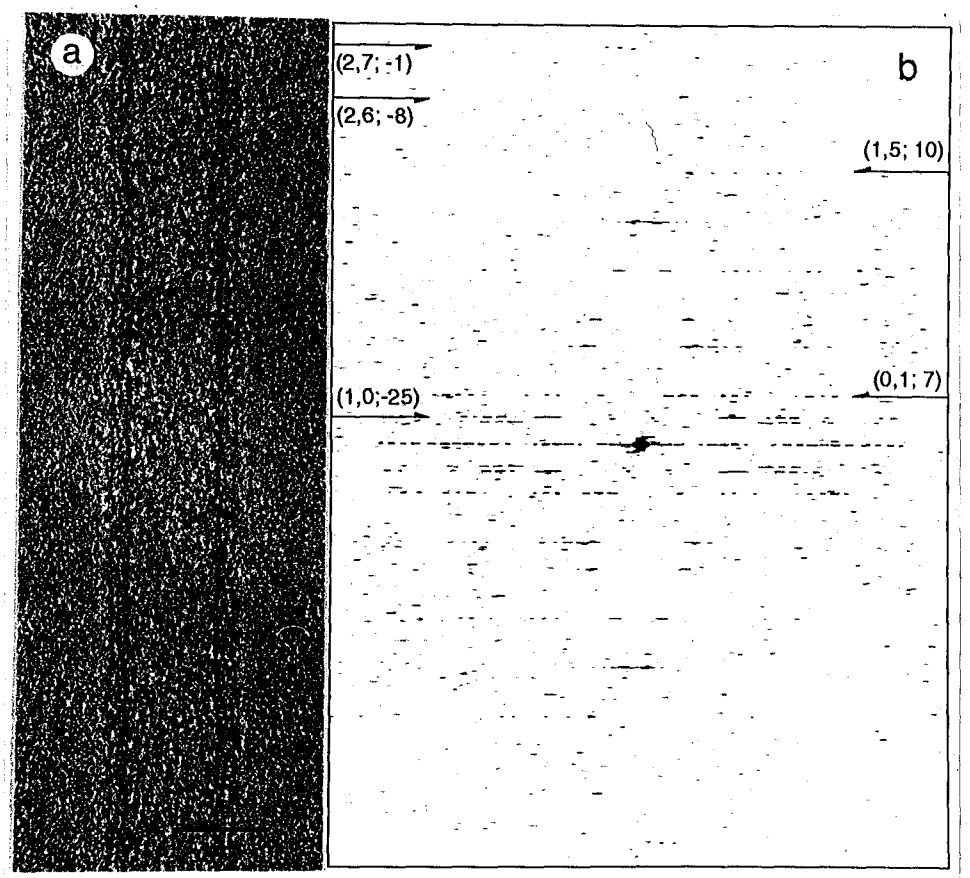


FIGURE III.2 Electron micrograph of a tubular crystal of Ca^{2+}-ATPase in vitreous ice (a) and amplitudes of the Fourier transform (b) of a image of a region of the tube in (a) taken at a smaller defocus than in (a). The transform consists of a series of horizontal lines (layer-lines), each corresponding to a set of helices. This tube belongs to the (n_{10} = -25, n_{01} = 7) family. (h,k) indices and start numbers on the associated surface lattice (see Toyoshima and Unwin, 1990) are indicated for some layer-lines. The (2,7; -1) layer-line is located at the axial spacing of 1/15.4 $Å^{-1}$. Bar in (a) represents 500 Å

The symmetry of a tubular crystal can be characterized by two indices representing the start numbers of two principal helices (n_{10} and n_{01}; see Toyoshima and Unwin, 1990). Similar to tubular crystals of other membrane proteins (e.g. Toyoshima and Unwin, 1990), those of Ca^{2+}-ATPase had a variety of helical symmetries slightly different from one another. We mainly collected the tubes belonging to the (n_{10} = -25, n_{01} = 7) or (n_{10} = -26, n_{01} = 7) families because they were found more frequently than the others and narrow enough for helical reconstruction. The (-25, 7) family of tubes formed without CrATP consisted of two groups of different diameters (Fig. III.3). The unit cell parameters (Table III.1 and Fig. III.4) were also different and, as a result, data from two groups were averaged in real space after Fourier space averaging of data within each group.

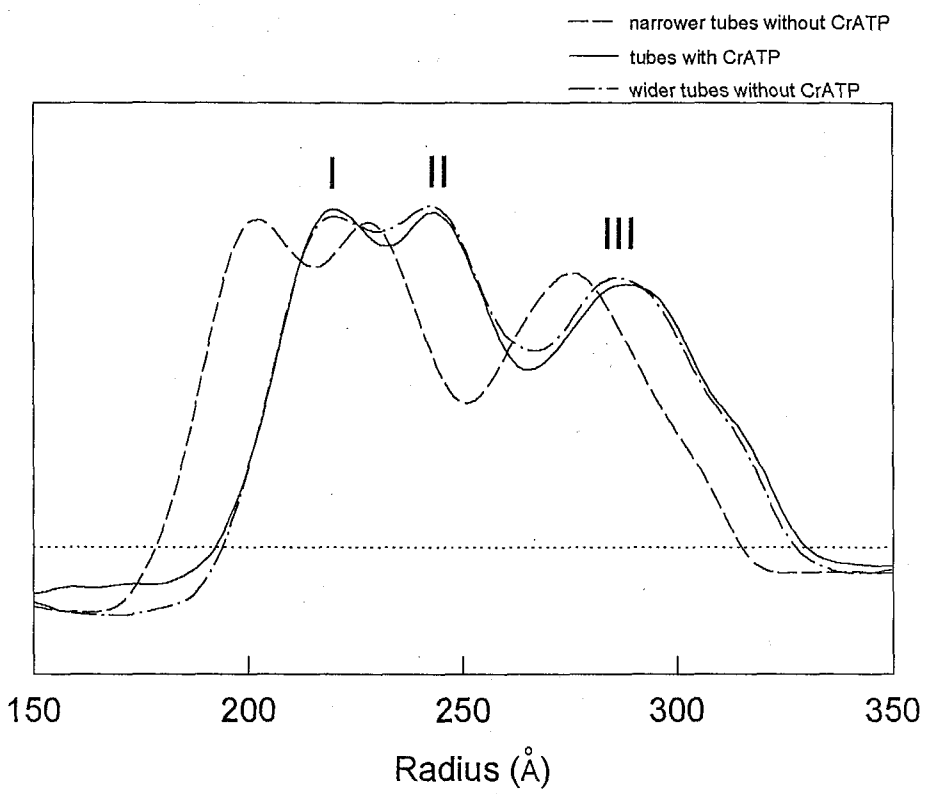


FIGURE III.3 Mean radial density distribution (g_{00}) of the tubes belonging to the (-25,7) family. The curves were obtained by Fourier-Bessel synthesis of the equatorial layer-line after correcting for the effects of the CTF. The three peaks marked refer respectively to the densities at the inner (I), the outer (II) leaflets of the lipid bilayer and the cytoplasmic domain (III). The minimum between the inner (I) and outer (II) leaflets corresponds to the middle of the lipid bilayer. Note that the curve for the narrower tubes is shifted by about 15 Å to the smaller direction.

TABLE III.1

Unit cell parameters for the tubes belonging to the (-25,7) family

	Number of tubes	Radius (Å)*	a (Å)ǂ	b (Å)ǂ	γ (°)ǂ
Narrower tubes without CrATP	3	216.0 ± 0.8	59.0 ± 0.3	117.6 ± 2.6	65.4 ± 0.6
Wider tubes without CrATP	3	230.3 ± 4.0	58.0 ± 0.6	109.6 ± 3.4	74.6 ± 1.7
Tubes with CrATP	7	230.5 ± 2.9	58.4 ± 0.4	111.9 ± 2.6	73.5 ± 0.9

* The radii refer to the middle of the membrane as specified in Fig. III.3.

ǂ a, b and γ are specified at the middle of the membrane; also see Fig. III.4.

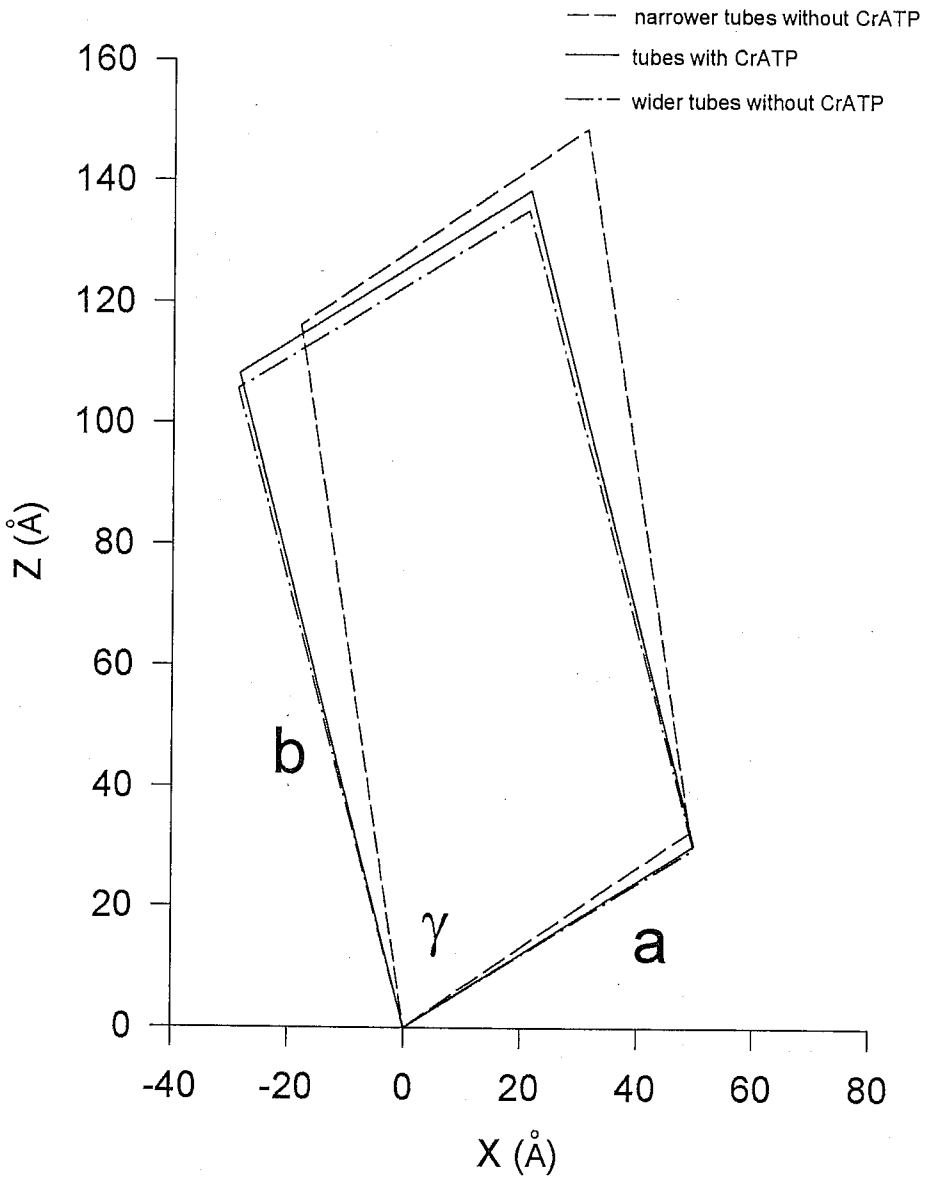


FIGURE III.4 Unit cells at the middle of the lipid bilayer for the tubes belonging to the (-25,7) family after scaling correction. Vertical axis corresponds to the helix axis of tubular crystals (z) (Toyoshima and Unwin, 1990). The unit cell parameters are summarized in Table III.1.

Table III.2 summarizes the quality of the datasets. Since dimer ribbons were formed in the tubular crystals, twofold rotational symmetry normal to the helix axis was present and used as the measure of the quality of the images (Toyoshima and Unwin, 1988). Phase residuals for the twofold symmetry were similar for the tubes with and without CrATP. The twofold phase residuals for those belonging to the (-26, 7) family were 2.9° (with CrATP, 23 images) and 8.2° (without CrATP, 5 images).

TABLE III.2

Image analysis of the tubes belonging to the (-25,7) family at 14 Å resolution

	-CrATP	+CrATP
Number of tubes averaged	6	7
Number of molecules averaged	5676	6090
Number of layer-lines incorporated in the reconstruction	49	50
Twofold phase residuals of the averaged datasets (°)*	6.2 (narrower tubes) 7.5 (wider tubes)‡	4.8
Mean twofold phase residuals of the individual datasets (°)	19.1 ± 2.0	19.6 ± 1.0

* Amplitude-weighted phase residuals for twofold symmetry calculated after averaging of weakly and strongly defocused pairs of images, within the first CTF zeros of weakly underfocused images (difference from either 0 or 180°, whichever closer). The calculation used all the points with amplitudes higher than 3% of the highest off-equatorial amplitude. The equator was omitted from the calculation.

‡ Since narrower and wider tubes had different unit cell parameters (Table III.1), twofold phase residuals of these two groups were calculated separately.

c. Three-dimensional maps of Ca^{2+}-ATPase with and without CrATP

Fig. III.5 shows the three-dimensional maps of tubular crystals of Ca^{2+}-ATPase without CrATP (a,c) and those with CrATP (b,d) at 14 Å resolution. We show the maps of the tubes of the (-25, 7) family only, because overall structural features of averaged images were virtually identical and a proper t-test could be done only with those belonging to this family. The number of images was unequal with the (-26, 7) family for those with and without CrATP. In either map (Fig. III.5), a large cytoplasmic domain is linked to the membrane through a stalk region in a similar manner. The positions of the three segments within the membrane are also very similar. A clear difference is found at the groove (arrowheads in Fig. III.5, a and c) ascribed previously to the ATP-binding site (Toyoshima et al., 1993a) This groove is filled in the maps of tubes with CrATP (Fig. III.5, b and d).

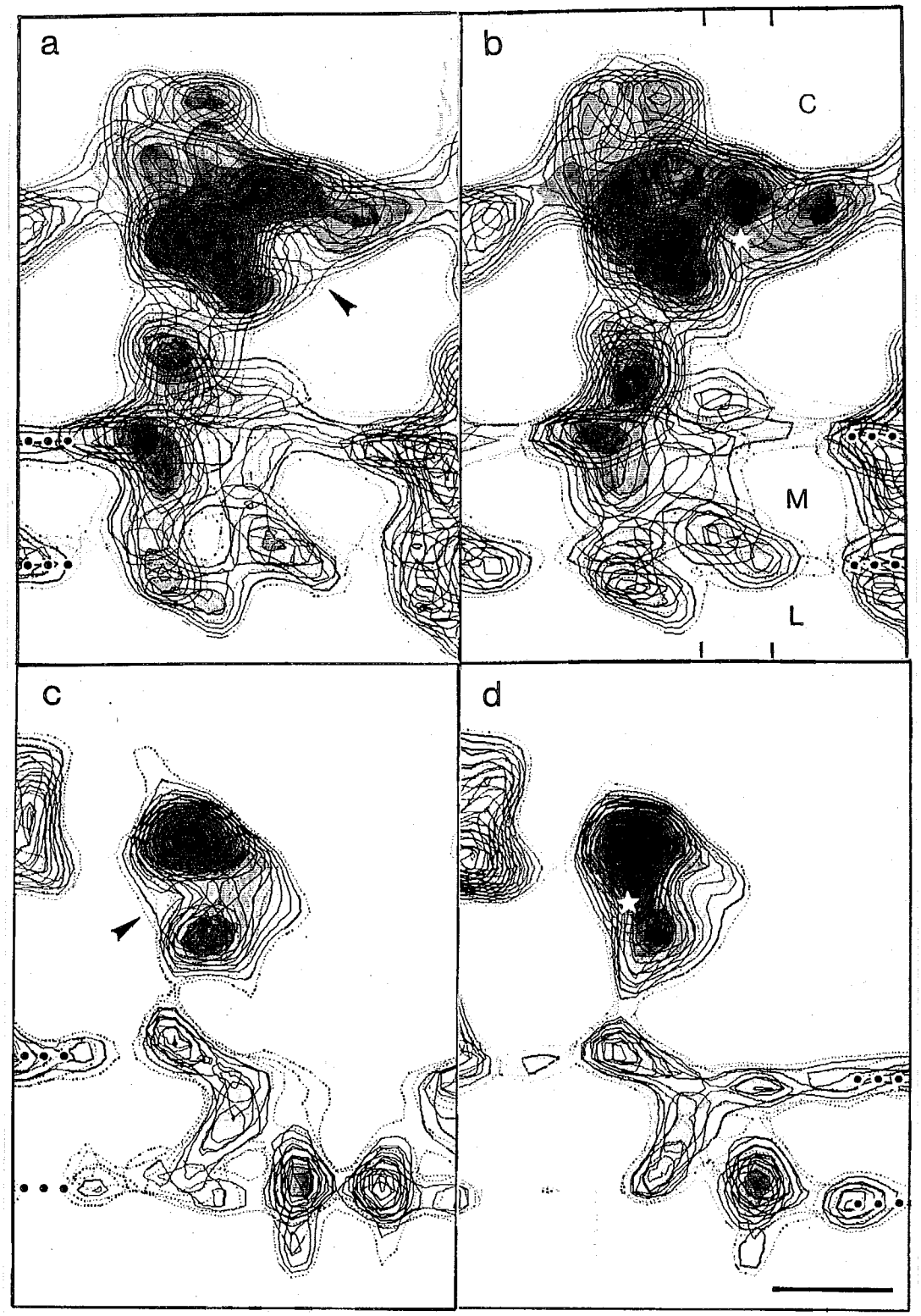

a
b
C
M
L
c
d

FIGURE III.5 Three-dimensional structure of Ca^{2+}-ATPase without CrATP (a),(c) and that with CrATP bound (b),(d) at 14 Å resolution. (a),(b) views perpendicular to the dimer ribbon. C, M and L in (b) indicate the cytoplasmic, the membrane and the lumenal regions, respectively. (c),(d) stacks of sections (13 Å thick) showing the structure around the groove. The sections are viewed along the dimer ribbon from the left hand side in (a),(c) and cover the area specified in (b) by the vertical bars at the top and bottom margins. Note that the groove indicated by the arrowheads in (a),(c) is filled in (b),(d). White stars in (b),(d) indicate the position of the most significant difference as determined by Student's t-test. The lowest solid contours correspond to about 75% of the expected volume. Scale bar represents 25 Å.

Fig. III.6 shows a fish net representation of the Ca^{2+}-ATPase molecule without CrATP (a) and with CrATP (b). Yellow nets show apparent gains (a) and losses (b) in density in the enzyme with CrATP. The highest peak in (a) corresponds to twice the standard deviation of the densities in the original three-dimensional maps (corresponding to 4 times the standard deviation in the difference map) and was located in the groove in (a). The corresponding peak was also observed in the difference map for the tubes belonging to the (-26, 7) family (data not shown).

To examine the statistical significance of the differences, Student's t-test was carried out. Red nets in Fig. III.6 enclose the regions where the probability that the differences were real and not due to chance is 99.5% in (a) and 99.9% in (b). At 99.9% confidence level, only a single peak exists in the groove, at the identical position to the highest peak in the difference map (yellow net in Fig. III.6a). This location is also marked in Fig. III.5(b) and (d) (white stars). It is slightly offset from the center of the groove, but not at the outer surface of the enzyme (Fig. III.5, b and d); it is located 10 Å inside from the entrance of the groove (Fig. III.5, b and d).

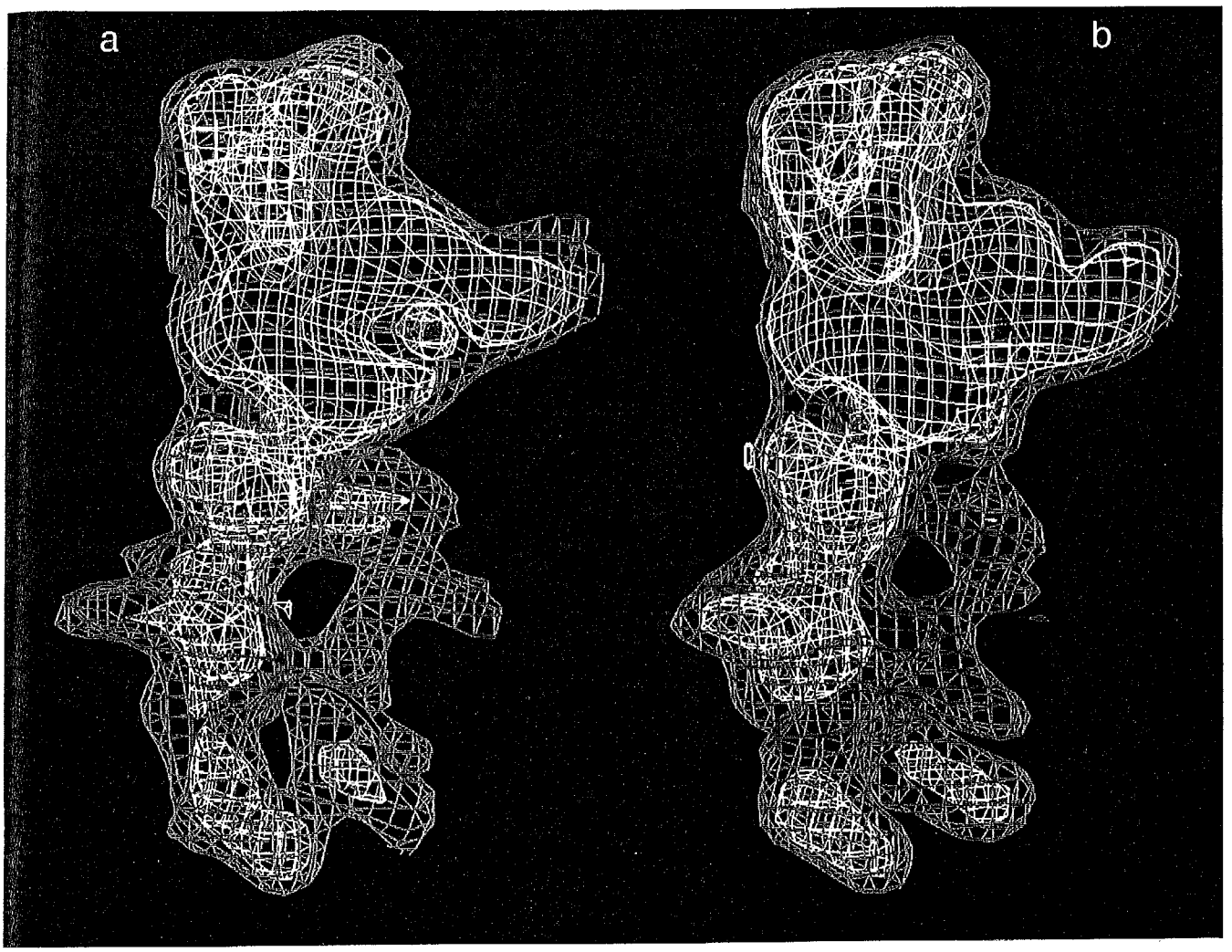


FIGURE III.6 Fish-net representation of a part of one Ca^{2+}-ATPase molecule viewed slightly upwards from the membrane. (a) the structure without CrATP. (b) the structure with CrATP bound. Blue and white nets enclose 75% and 35% of the expected volume, respectively. Yellow nets show gains (a) and losses (b) in density in the structure with CrATP; the nets enclose the regions where the difference is larger than 1.33 times standard deviation of the densities in the original density map before subtraction. Red nets correspond to the 99.5% (a) and 99.9% (b) confidence level. Note that the regions of the largest difference and the most significant difference coincide and are located within the groove specified in (a).

As a control, we compared two groups whose members were randomly selected ignoring the presence or the absence of CrATP. In this control experiment, there were only small gains or losses in density; only 1.33 times the standard deviation at the highest. Furthermore, the significance at this location was less than the 97% confidence level. On the other hand, there were a few points at the 99% confidence level, but the differences themselves were small and were located on the surface of the enzyme; above the 99.5% confidence level none existed (data not shown). Thus, in the control experiment, the difference map and the t-map were inconsistent, which is typical of noisy data. In contrast, in the proper difference map, the density gain in the groove (yellow net in Fig. III.6a) was high and consistent with the t-map (red net in Fig. III.6a). Hence it must represent the real difference caused by the binding of CrATP to Ca^{2+}-ATPase.

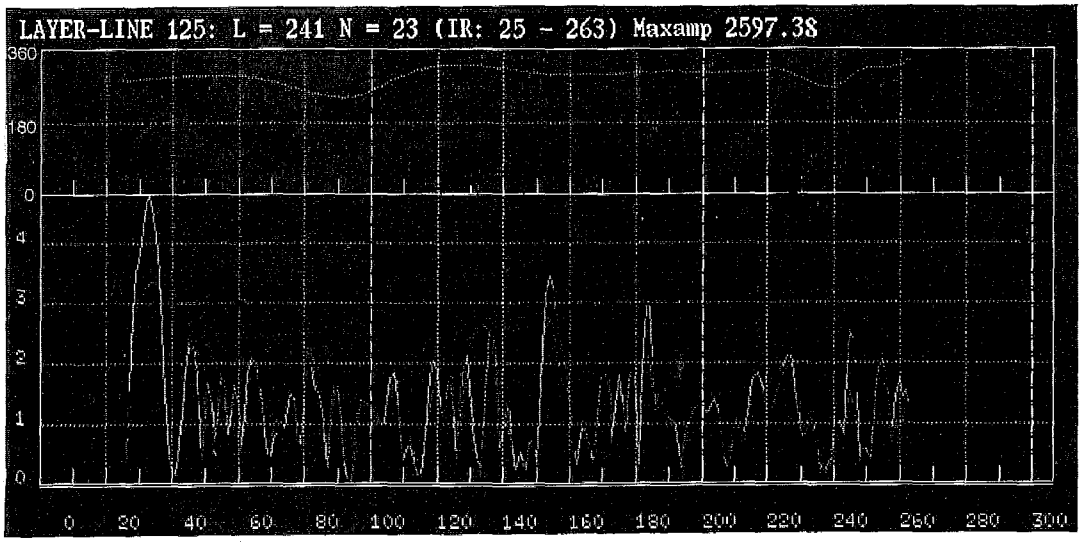
III-2. The structure of Ca^{2+}-ATPase with bound CrATP at better than 10 Å resolution.

a. Image analysis

Fig. III.7 shows the profile of averaged layer-lines of the (n_{10} = -26, n_{01} = 7) family. Here 23 images were averaged. Lower panel shows amplitudes (solid lines) and the figure of merit (dotted green lines). Blue points in amplitudes mean that the phase residuals for these points are less than 30° from the ideal values. There are many blue points near 8 Å resolution (300 grid units). But, the map calculated from this dataset was still noisy. Therefore, we needed to average more images to increase the S/N.

To do so, we carried out real space averaging over different helical families. From 45 images of three different helical families (12, 23, 10 images of the tubes belonging to the (n_{10} = -25, n_{01} = 7), (n_{10} = -26, n_{01} = 7) and (n_{10} = -27, n_{01} = 7) families, respectively), we could obtain a three-dimensional map at better than 10 Å resolution. The final twofold phase residual was about 14° at 44% completeness (Table III.3). The completeness was low at 15 - 8 Å resolution. But, the map including 8 Å resolution data looked much better defined and more detailed than the maps calculated at 10 or 9 Å resolution. Thus, the data at ~ 8 Å resolution contributed significantly to the map.

(a)



(b)

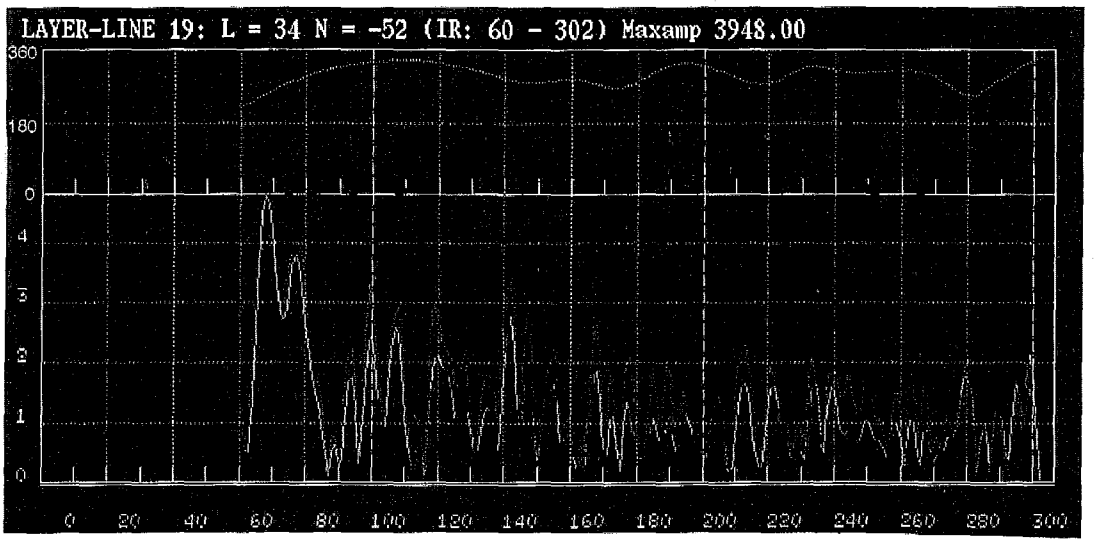


FIGURE III.7 Screen dumps of X window output by llprx2 (Appendix B). Amplitude (solid curve in the lower panel), phase (red 'plus' marks in the upper panel), the averaged CTF (blue curve in the upper panel) and the figure of merit (green dotted curve in the lower panel) along two layer-lines. They were obtained by averaging 23 images belonging to the (n_{10} = -26, n_{01} = 7) family. The 241th layer-line (n = 23, l = 241) (a) and the 34th (n = -54, l = 34) (b) are

located at Z = 1/16.1 $Å^{-1}$ and Z = 1/114.2 $Å^{-1}$ from the equator, respectively. One grid unit from the meridian corresponds to $1/(2.4\times10^3)$ $Å^{-1}$. Therefore, 300 grid units correspond to 1/8 $Å^{-1}$. The phases show clear tendency to be either 0° or 180°, indicative of twofold symmetry. Phase deviations from the ideal values (0° or 180°) are represented by different colors for the amplitudes, blue, green, magenta, and red mean that phase deviation at that point is less than 30°, between 30° and 45°, between 45° and 70° and more than 70°, respectively. The amplitudes and phases were calculated as in Eq. 5 in appendix B. The averaged CTF and the figure of merit were calculated as in Eq. 6 and Eq. 9 in appendix B, respectively. The figure of merit was normalized according to Eq. 13 in Appendix B.

TABLE III.3 Twofold phase residuals after averaging 45 images.

	~ 20 Å	20 ~ 15 Å	15 ~ 10 Å	10 ~ 8 Å	overall
residuals (°)*	2.52	18.27	34.73	31.94	13.96
completeness (%)‡	95.59	88.93	45.95	27.64	44.32

* Amplitude-weighted phase residuals for twofold symmetry calculated (difference from either 0 or 180°, whichever closer). The calculation used all the points with amplitudes higher than 0.1% of the highest off-equatorial amplitude. The equator was omitted from the calculation.

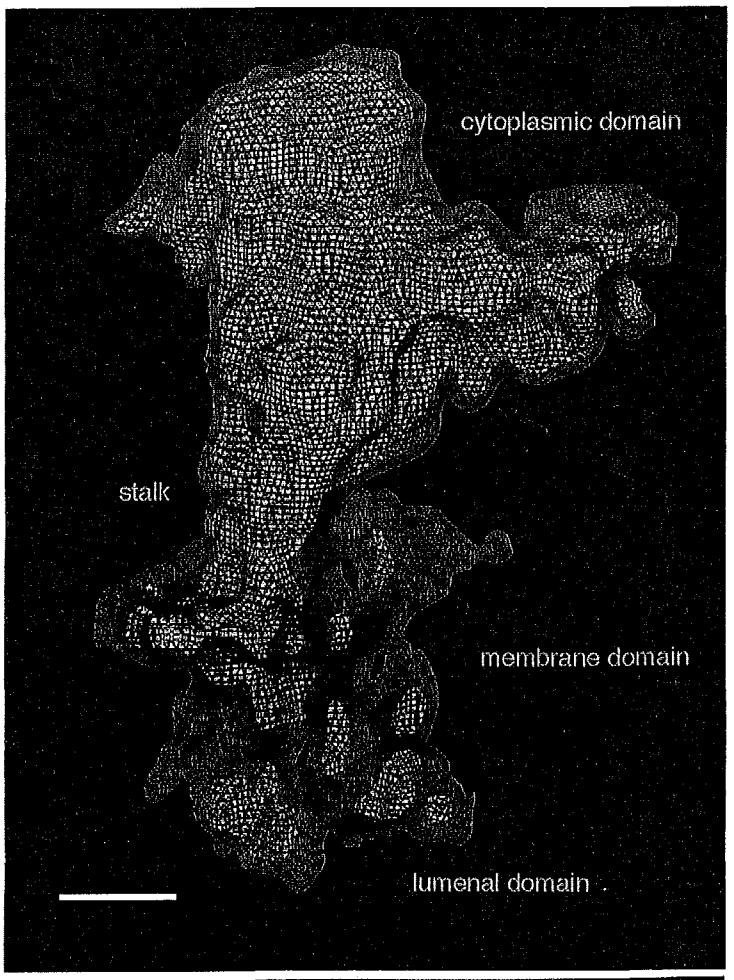
‡ The completeness was defined as the number of points used for calculation of twofold residuals divided by the number of all possible points within a certain resolution range.

b. Three-dimensional map

Fig. III.8(a) shows the three-dimensional map of Ca^{2+}-ATPase obtained by the new averaging scheme (Fig. II.1). Orange arrow indicates a separate density peak (white net) found in the groove. It may represent a bound CrATP molecule, because its position coincides with the point of the highest confidence level (Fig. III.5b,d and III.6).

Fig. III.8(b) shows a cut out model viewed in a direction similar to (a). In this view, the putative CrATP molecule is resolved more clearly (orange arrow). Also, in this model, several transmembrane columns are visualized and one of them is marked by c1. It extends about 30 Å long across the membrane region.

(a)



(b)

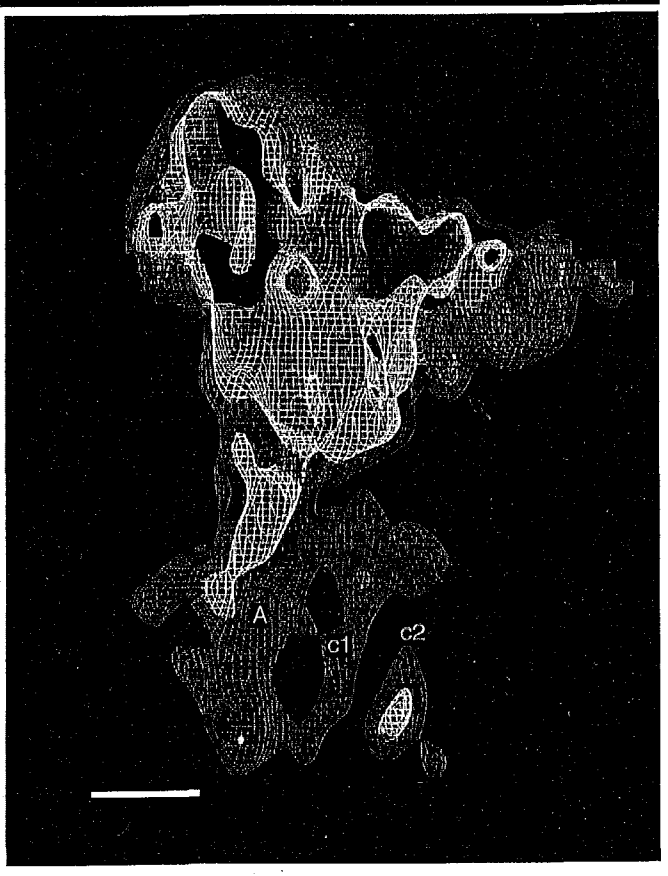


FIGURE III.8 (a) the structural model of a whole Ca^{2+}-ATPase molecule at better than 10 Å resolution. (b) a cut out model viewed in a direction similar to (a). The largest membrane segment A is marked; c1 and c2 identify two transmembrane columns. Red and white nets enclose 80% and 40% of the expected volume, respectively. Orange arrows indicate a mass (white nets) in the groove, which presumably represent the bound CrATP molecule. Bars represent 20 Å.

Fig. III.9 shows the stalk section viewed from the cytoplasmic side. According to the secondary structure prediction, the stalk section consists of 4 - 5 α-helices. The stalk section visualized here, indeed, consists of 5 density peaks. However, we cannot say if all of these represent α-helices because the sizes are quite different. There is a hole in the center of these peaks, but its physiological meaning is not clear.

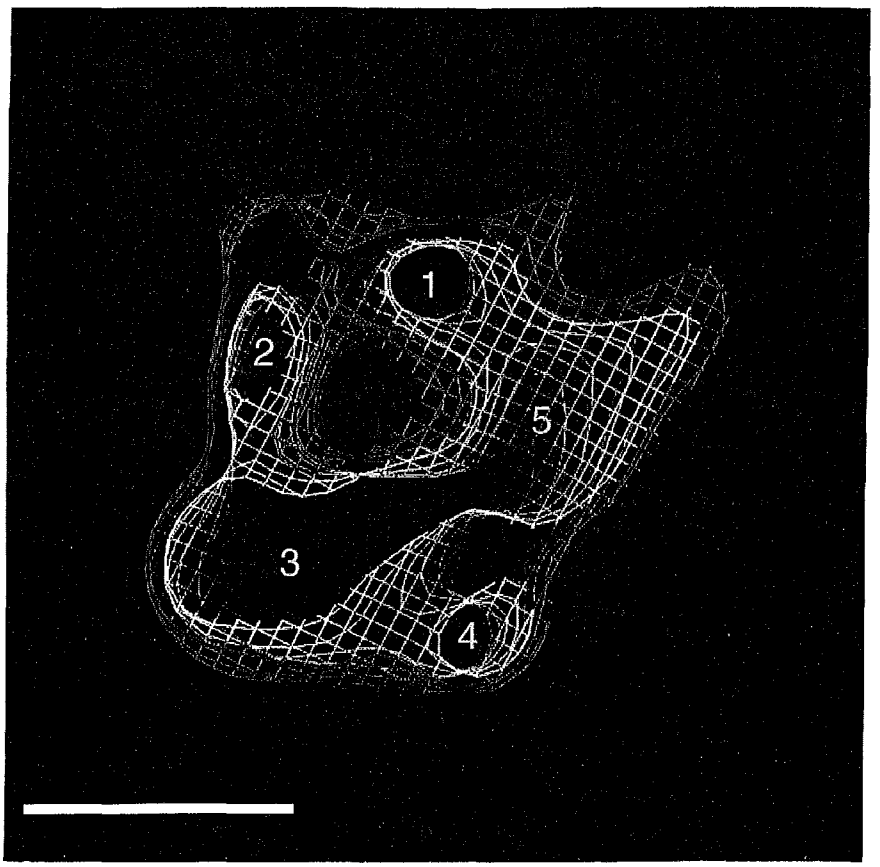


FIGURE III.9 The stalk section viewed from the cytoplasmic side. The stalk section consists of 5 density peaks (1 - 5). There is a hole surrounded by these peaks. Red and white nets enclose 80% and 50% of the expected volume, respectively. Bar represents 15 Å.

Fig III.10 shows the membrane region viewed from the cytoplasmic side. The upper figure (a) shows a model cut out near the cytoplasmic surface. The lower figure (b) shows a model cut out near the lumenal surface. We can identify 10 density peaks, a1-a6, b1-2 and c1-2, of transmembrane columns. These columns may correspond to the 10 transmembrane helices proposed by the secondary structure prediction. Columns, a1-a6 surround a pore near the cytoplasm whereas a2, a4, a5 and a6 form another pore near the lumen. These pores are not contiguous and intercept at the middle of the transmembrane region. They presumably represent the inlet and the outlet of the ion pathway, respectively. Fig III.11 shows the structure around the inlet.

Fig. III.12 shows a cut out model of Ca^{2+}-ATPase viewed in a direction parallel to the membrane plane. The blue arrow indicates the putative pathway of Ca^{2+}. This ion pathway is closed at the middle of the membrane.

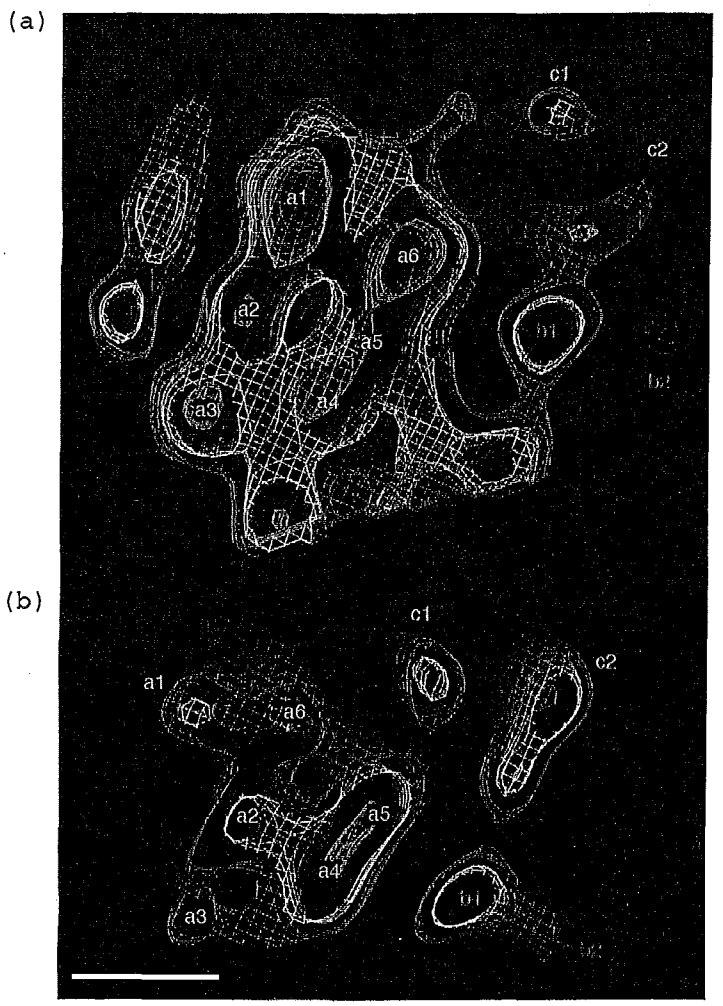


FIGURE III.10 The cytoplasmic half (a) and lumenal half (b) of the membrane region viewed from the cytoplasmic side. At 14 Å resolution, membrane domain was resolved as three separate segments A, B and C (Toyoshima et al., 1993a). At better than 10 Å resolution, these segments were further divided into a1-6, b1-2 and c1-2, respectively. Note that a1-a6 in (a) and a2, a4, a5 and a6 in (b) form a pore at the center of A segment. These pores are likely to correspond to the inlet (cytoplasmic side) and outlet (lumenal side) of ions. Red, white and purple nets enclose 80%, 50% and 35% of the expected volume, respectively. Bar represents 15 Å.

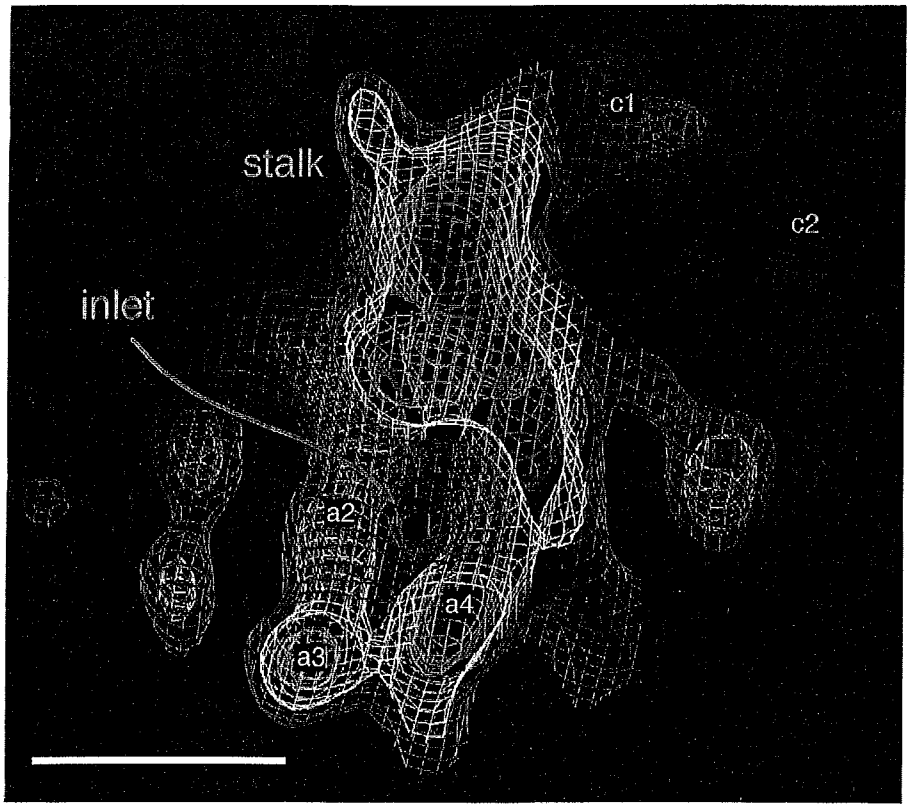


FIGURE III.11 The structure around the inlet of Ca^{2+}. Red, white and purple nets enclose 70%, 50% and 35% of the expected volume, respectively. Some of the transmembrane columns are marked (a2-a4, c1 and c2). Bar represents 25 Å.

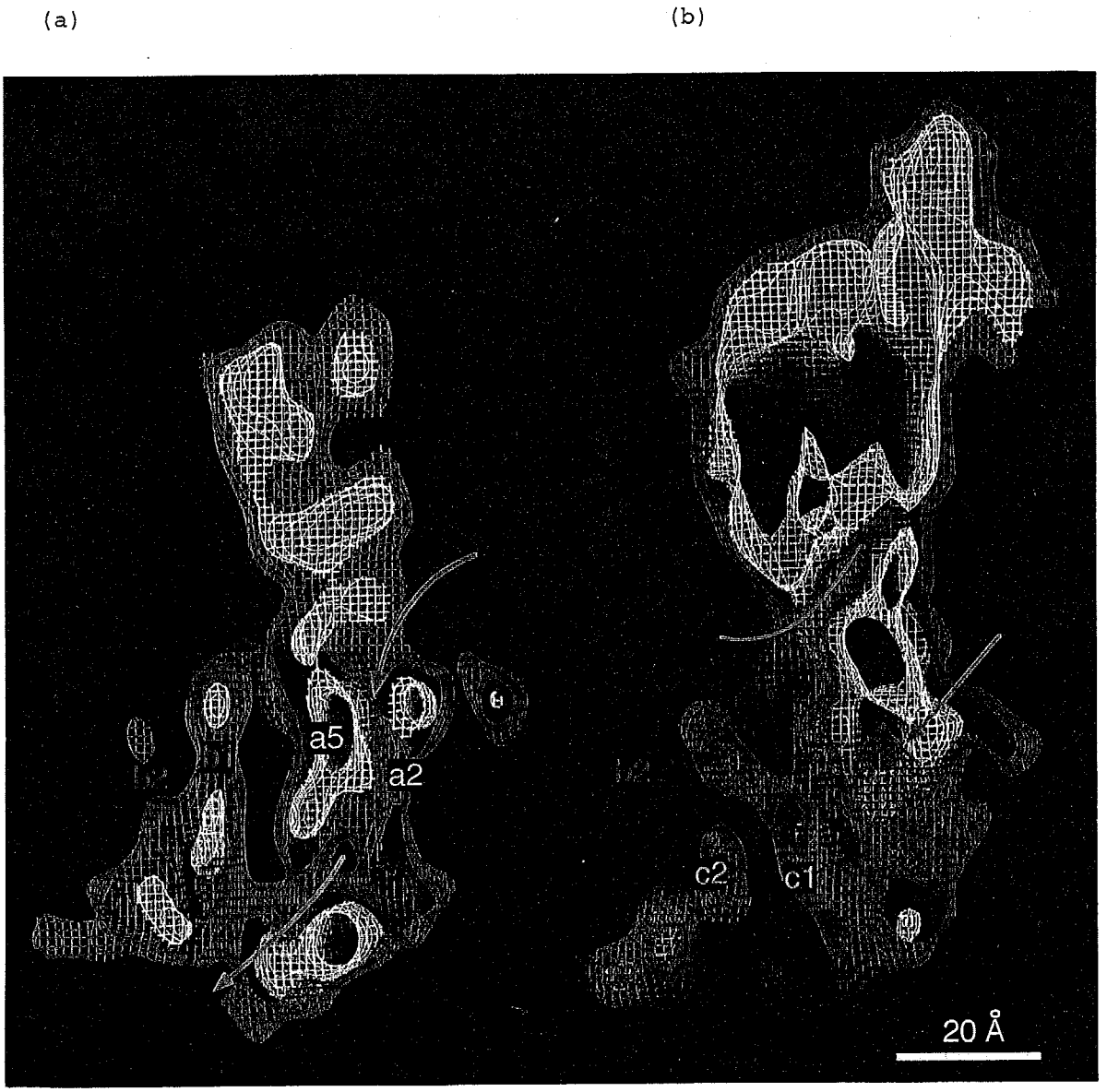


FIGURE III.12 Cut out models viewed along the membrane plane. The blue arrow in (a) indicates the putative pathway of Ca^{2+}. It is closed at the middle of the membrane. The green curve in (b) indicates the hole at the center of the stalk and the blue curve in (b) indicates the inlet of Ca^{2+}. Red and white nets enclose 80% and 40% of the expected volume, respectively. These models are viewed from the side opposite to Fig. III.8.

IV. Discussion

IV-1. The ATP-binding site of Ca^{2+}-ATPase revealed by electron image analysis.

We described here, the differences in the three-dimensional structure of tubular crystals of Ca^{2+}-ATPase formed with and without CrATP. With these tubular crystals, we monitored that the ATPase activity was inhibited, showing that CrATP remained bound, though occluded Ca^{2+} was released (Fig. III.1). By cryo-electron microscopy and quantitative image analysis, we found a clear difference at the groove in the cytoplasmic domain (Figs. III.5 and III.6). This difference was statistically significant at a confidence level higher than 99.9% (Fig. III.6b). Except for this difference, the structure of the enzyme with CrATP appeared very similar to that without CrATP (Figs. III.5 and III.6).

a. Position of the ATP-binding Site

The ATP-binding site was assigned to the groove in the cytoplasmic domain (Toyoshima et al., 1993a) based on the results by affinity labeling (Yamamoto et al., 1989). The difference map and the t-map shown in Fig. III.6 indicate that this assignment was correct. In the difference map, the highest peak (4 times the standard deviation of the densities in the difference map) was located in the groove. This difference was confirmed to be statistically significant by t-test at a confidence level higher than 99.9% (Fig. III.6b, red net). The distances between the ATP-binding site and phospholipid/water interface have previously been measured to be 40 ~ 60 Å by fluorescence energy transfer (Teruel and

Gomez-Fernandez, 1986; Gutierrez-Merino et al., 1987; Bigelow and Inesi, 1991; Baker et al., 1994). In the structure described here, the distance was 43 Å between the point of the most significant difference in the groove (white stars in Fig. III.5, b and d; red net in Fig. III.6b) and the surface of the membrane. This value is consistent with the measurement for Cys 344/Cys 364 (40 Å), which bracket Asp 351, the residue of phosphorylation (Bigelow and Inesi, 1991).

Similar experiments have been carried out to measure the distances between the ATP-binding site and lanthanide binding sites (Highsmith and Murphy, 1984; Scott, 1985; Squier et al., 1987). The reported values (16 ~ 35 Å) are smaller than those measured from the membrane surface to the ATP-binding site, and lanthanides are thought to bind to the stalk (Squier et al., 1990; Ogurusu et al., 1991; Henao et al., 1992). In particular, Herrmann and Shamoo (1988) measured the distance between bound CrATP and lanthanide (Eu^{3+}) to be ~ 18 Å. This value implies that Eu^{3+} binds to the top of the stalk in our map (Fig. III.5).

The whole CrATP molecule remains bound to the enzyme even after 16 hours incubation (Vilsen and Andersen, 1992), which is long enough for tubular crystal formation. CrATP has a molecular weight of 559 and measures roughly 15 Å in diameter. It is presumably too small to be visualized at 14 Å resolution as a separate entity when filling the ATP-binding pocket (Fig. III.5, a and c). Nevertheless, since the groove was resolved as a hole of 20 × 10 × 7 $Å^3$ in the enzyme without bound nucleotides, it is likely that the density filling the groove shows up in the difference map. It is also

expected that the improvement of resolution will allow us to directly image the CrATP in the groove. Indeed in our new map at 8 Å resolution, we see a separate density of ~ 10 Å diameter (Fig. III.8). On the other hand, we must of course expect structural changes in the protein around the nucleotide binding site, as has been demonstrated for adenylate kinase (Schulz, 1991). Therefore, we cannot say that the gained densities represent a CrATP molecule. However, since the difference is well localized and no statistically significant density loss is found in the map, it is very likely that the density in the groove really represents CrATP.

b. Conformation of Ca^{2+}-ATPase with CrATP in the tubular crystals

Though it is well established that CrATP traps two Ca^{2+}s and locks the enzyme in the 'occluded' state, it is not obvious whether the enzyme molecules are in the occluded state in the tubular crystals, because vanadate and EGTA had to be added to induce the crystals. CrATP was still bound in the tubular crystals (filled circle after 20 hours in Fig. III.1; Stokes and Lacapere, 1994a; Buhle et al., 1983), but most of trapped Ca^{2+} were released to make the Ca2+-binding site vacant (filled box after 20 hours in Fig. III.1). This result is in apparent conflict with Buhle et al. (1983) but agrees well with Vilsen and Andersen (1992) and Coan et al. (1994). Vilsen and Andersen (1992) measured the amount of bound $^{45}Ca^{2+}$ to be less than 1/10 at 15 hours after the addition of EGTA.

An electron paramagnetic resonance (EPR) spectroscopy study gives us some clue. With the iodoacetamide-labeled enzyme , the same EPR spectral changes as those found in the presence of both

MgATP and Ca^{2+} were observed in the presence of CrATP alone, irrespective of the presence or the absence of Ca^{2+} (Chen et al., 1991). It indicates that the same structural changes, induced by CrATP in the presence of Ca^{2+} also occur in the ATP-binding site even in the absence of Ca^{2+}. In addition, the enzyme can make a stable complex with CrATP in the absence of Ca^{2+} and the high affinity Ca^{2+}-binding sites can be filled afterward (Coan et al., 1994). Hence, the enzyme with bound CrATP is thought to remain in the Ca^{2+}-occluded state, while bound Ca^{2+} has been released (Coan et al., 1994). A proteolysis study also supports this idea (Vilsen and Andersen, 1987).

Nevertheless, the matter seems more complicated due to the presence of vanadate used to induce tubular crystals. Though the vanadate solution that we used consists mostly of decavanadate, it contains monomers and other oligomeric species of vanadates (Csermely et al., 1985; Aureliano and Madeira, 1994). It alters the ATPase activity, though the mobility of the membrane domain is reported to be unchanged (Lewis and Thomas, 1986). Therefore, the presence of vanadate is likely to alter the structure of the cytoplasmic domains. In fact, the trypsin digestion pattern of the tubes with bound CrATP in the presence of vanadate was different depending on the presence or the absence of Ca^{2+} and was identical to that of the enzyme without CrATP (unpublished observation). Further experiments will be needed to characterize the state of the enzyme in these tubular crystals.

c. Structural differences in other sites

There are a few other sites where the differences are statistically significance at a confidence level higher than 99.5%. At the top of the cytoplasmic domain (red nets in Fig. III.6a), small gains or losses in density are found, corresponding to 1.2 times the standard deviation of the densities in the original map. From its position next to the groove assigned as the ATP-binding pocket, this region is a good candidate for the β-strand domain, which has an essential role in phosphoenzyme turnover (Andersen et al., 1989; Clarke et al., 1990b).

We had expected to find structural changes in either the stalk or the transmembrane domain, but found no significant differences in these regions (Fig. III.6). This fact may imply that the CrATP-bound enzyme is not truly in the occluded state in the tubular crystals, but in a special conformation forced by vanadate. That is, the binding signal of ATP may have been blocked and could not be conducted to the stalk or to the membrane domain. Of course, we cannot exclude the possibility that structural changes in these domains are too small to be resolved at 14 Å resolution, and improved resolution is needed for the visualization. In the case of the acetylcholine receptor, 9 Å resolution was needed to visualize the rearrangement, caused by the binding of acetylcholine, of α-helices surrounding the pore (Unwin, 1995). We have a good hope to visualize such conformational changes in the near future.

IV-2. The structure of Ca^{2+}-ATPase with bound CrATP at better than 10 Å resolution

The improved map described in III-2 agrees well with the predicted structural model. For example, the numbers of the density peaks are consistent with the numbers of predicted α-helices in both the stalk (Fig. III.9) and the membrane regions (Fig. III.10). Furthermore, this map has resolved clearly several transmembrane columns (e.g. Fig. III.8b). Also, a separate density peak (orange arrows in Fig. III.8a,b), presumably corresponding to the bound CrATP, has been found in the groove identified as the ATP-binding pocket (IV-1).

a. Ion pathway

It has been predicted that Ca^{2+}-ATPase has 10 transmembrane helices. Fig. IV.1 shows a hypothetical packing model of the transmembrane helices, based on the previous map (Toyoshima et al., 1993a). We can assign the 10 transmembrane columns (a1-6, b1-2 and c1-2 in Fig. III.10) in certain sections of the membrane region of the current map. In the lumenal half (Fig. III.10b), the arrangement of them is similar to the hypothetical packing (Fig. IV.1). Though the putative inlet and outlet of the ion pathway are resolved (Fig. III.10a, b), the pore is not found at the middle of the membrane (Fig. IV.2b). It will mean that the ion pathway is closed or constricted there.

During transport of Ca^{2+}, Ca^{2+}-ATPase traps Ca^{2+}s within the membrane domain. This state is called the occluded state. CrATP

stabilizes the enzyme in this state. The tubular crystals with bound CrATP loses trapped Ca^{2+}s during the crystallization process (Fig III.1). There is biochemical evidence that, even without occluded Ca^{2+}, the conformation of the enzyme itself is in the occluded state (see IV-1b). Therefore, the ion pathway is expected to be closed at the middle of the membrane to trap Ca^{2+}. Also, the current resolution is still too low to resolve individual helices where they become tightly packed.

b. Proposed three-dimensional packing of 10 transmembrane helices

Because we can see the inlet and outlet of the transmembrane ion pathway and transmembrane density peaks, the next step is to clarify a three-dimensional organization of transmembrane columns. Though the resolution of the map was not high enough to resolve all the way through the membrane, it was still possible to trace 10 columns (circles except the blue ones in Fig. IV.2). The dark blue circles indicate the putative ion pathway. If we want to connect the apparent inlet and outlet of the ion pathway, because they are offset by 10 Å, we must assume that helices are highly kinked or disordered at the middle of the membrane. Indeed, there are conserved prolines just above the presumed Ca^{2+}-binding sites. The map is still too noisy to propose a decisive model, but we are now beginning to see the transmembrane helices involved in active transport.

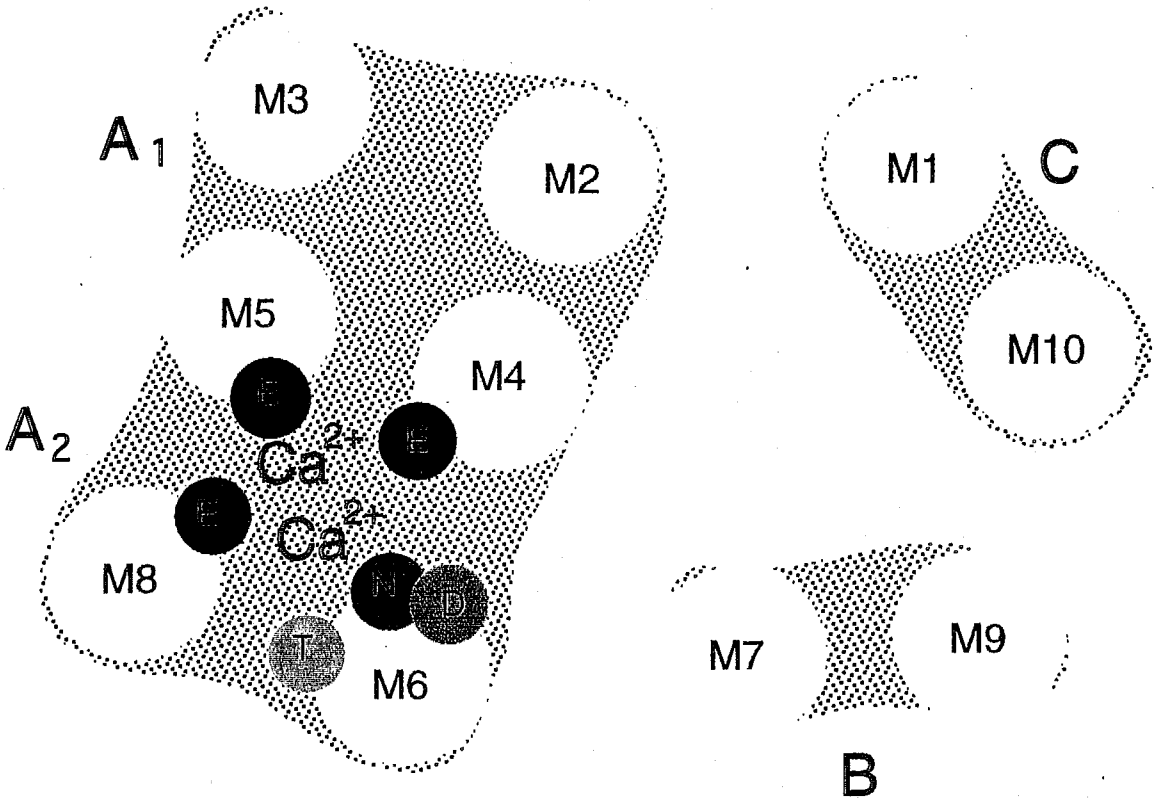


FIGURE IV.1 A hypothetical packing model of transmembrane helices based on the previous map (Toyoshima et al., 1993a). This model was modified from that proposed by Stokes et al. (1994b). Site directed mutagenesis studies have identified six residues (Glu 309 in M4, Glu 771 in M5, Asn 796, Thr 799, Asp 800 in M6 and Glu 908 in M8) responsible for Ca^{2+} binding. Hence, it is generally believed that four transmembrane helices, M4, M5, M6 and M8 form the ion pathway.

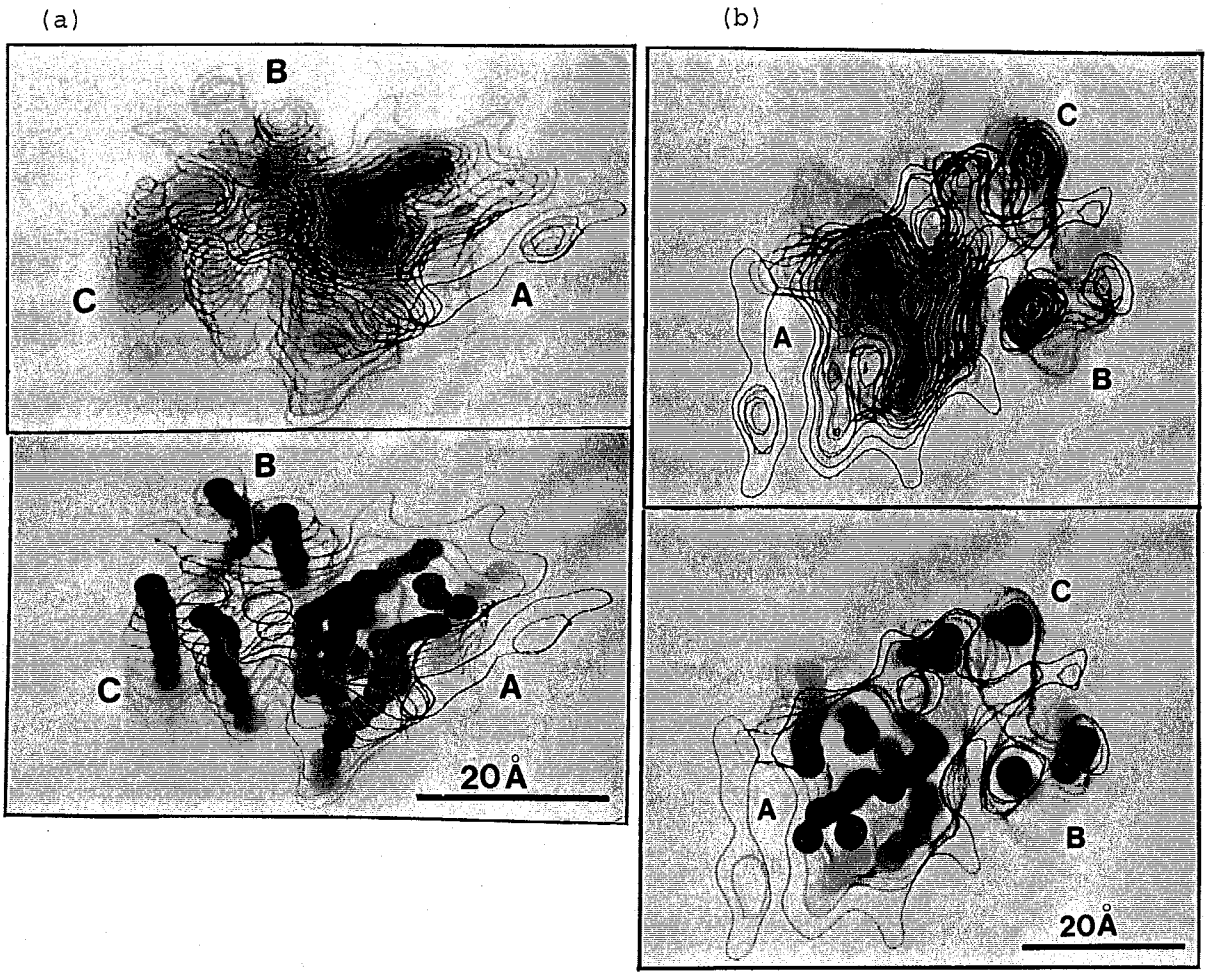


FIGURE IV.2 Identification of 10 transmembrane columns (a, b). The three membrane segments, A, B and C (Toyoshima et al., 1993a) are marked. Dark blue circles locate negative peaks (i.e. lower than protein densities). Other colored circles shows the density peaks constituting transmembrane columns. The lowest solid contours correspond to 80% of the expected volume. Note that columns forming the pore are highly kinked at the middle of the membrane region.

V. Conclusions

On the basis of the three-dimensional image analysis, the ATP-binding site was located in the groove of the cytoplasmic domain. This is the first functional site identified directly on the three-dimensional structure, and will provide the origin for mapping of other functional sites.

Furthermore, the structure of Ca^{2+}-ATPase with bound CrATP has been determined to a resolution of better than 10 Å. This map has resolved several transmembrane columns and a mass, presumably representing the bound CrATP molecule. The most interesting feature is that the inlet and outlet of the ion pathway are visualized at the center of A transmembrane segment. The map of the enzyme obtained here is likely to represent the Ca^{2+}-occluded state. Indeed, the ion pathway is closed at the middle of the membrane region. Because the inlet and outlet are offset by 10 Å, we must assume that the transmembrane helices surrounding the ion pathway are highly kinked or disordered.

Fig. V summarizes the positions of the ATP-binding site, Ca^{2+}-binding site and ion pathway proposed here.

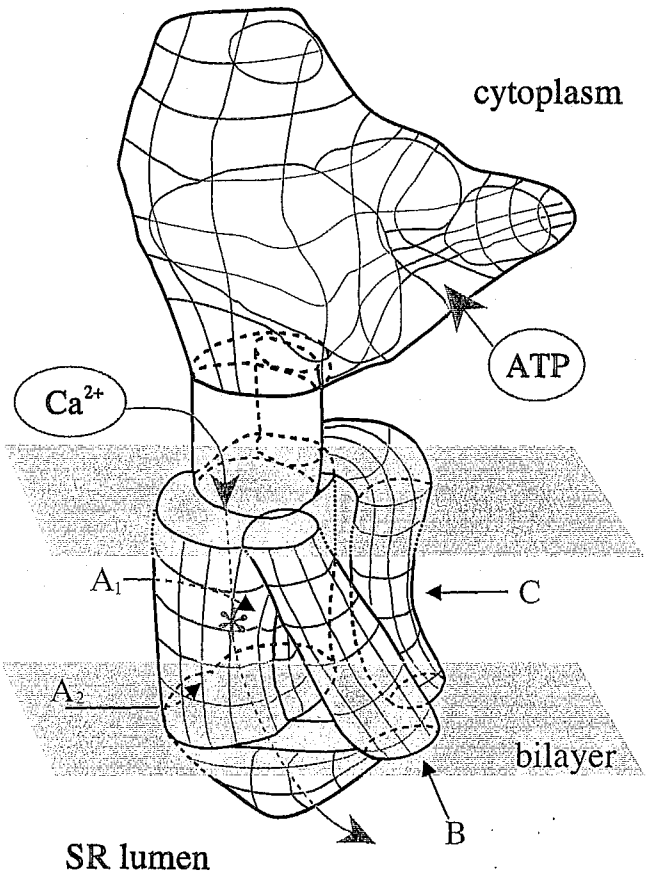


FIGURE V The proposed positions of the functional sites of Ca^{2+}-ATPase. In this model, the ATP-binding site is assigned to the groove in the cytoplasmic domain. Ca^{2+} enters into the membrane region at the root of the stalk and passes through the center of A segment. During transport of Ca^{2+}, Ca^{2+} is occluded at the middle of the membrane region (asterisk). Both B and C segments are likely to consist of two transmembrane helices.

VI. Acknowledgments

I gratefully thank Professor C. Toyoshima, Institute of Molecular and Cellular Biosciences, the University of Tokyo. All the work has been done under his auspice. Also I thank Professor S. Hirotsu, Department of Biological Sciences, Tokyo Institute of Technology for his support.

VII. References

Akiba, T., C. Toyoshima, T. Matsunaga, M. Kawamoto, T. Kubota, K. Fukuyama, K. Namba, and H. Matsubara. 1996. Three-dimensional structure of bovine cytochrome bc_1 complex by electron cryomicroscopy and helical image reconstruction. *Nature. Struct. Biol.* 3:553-561.

Andersen, J. P., B. Vilsen, E. Leberer, and D. H. MacLennan. 1989. Functional consequences of mutations in the beta-strand sector of the Ca^{2+}-ATPase of sarcoplasmic reticulum. *J. Biol. Chem.* 264:21018-21023.

Aureliano, M. and V. M. Madeira. 1994. Interactions of vanadate oligomers with sarcoplasmic reticulum Ca^{2+}-ATPase. *Biochim. et Biophys. Acta.* 1221:259-271.

Baker, K. J., J. M. East, and A. G. Lee. 1994. Localization of the hinge region of the Ca^{2+}-ATPase of sarcoplasmic reticulum using resonance energy transfer. *Biochim. et Biophys. Acta.* 1192:53-60.

Bigelow, D. J. and G. Inesi. 1991. Frequency-domain fluorescence spectroscopy resolves the location of maleimide-directed spectroscopic probes within the tertiary structure of the Ca-ATPase of sarcoplasmic reticulum. *Biochemistry.* 30:2113-2125.

Bigelow, D. J. and G. Inesi. 1992. Contributions of chemical derivatization and spectroscopic studies to the characterization of the Ca^{2+} transport ATPase of sarcoplasmic reticulum. *Biochim. Biophys. Acta.* 1113:323-338.

Brandl, C. J., N. M. Green, B. Korczak, and D. H. MacLennan. 1986. Two Ca^{2+}ATPase genes: Homologies and mechanistic implication of deduced amino acid sequences. *Cell* 44:597-607.

Buhle Jr., E. L., B. E. Knox, E. Serpersu, and U. Aebi. 1983. The structure of the Ca^{2+}-ATPase as revealed by electron microscopy and image processing of ordered arrays. *J. Ultrastruct. Res.* 85:186-203.

Caspar, D. L. D., C. Cohen, and W. Longley. 1969. Tropomyosin: Crystal structure, polymorphism and molecular interactions. *J. Mol. Biol.* 41:87-107.

Chen, L., C. Sumbilla, D. Lewis, L. Zhong, C. Strock, M. E. Kirtley, and G. Inesi. 1996. Short and Long Range Functions of Amino Acids in the Transmembrane Region of the Sarcoplasmic Reticulum ATPase. *J. Biol. Chem.* 271:10745-10752.

Chen, Z. D., C. Coan, L. Fielding, and G. Cassafer. 1991. Interaction of CrATP with the phosphorylation site of the sarcoplasmic reticulum ATPase. *J. Biol. Chem.* 266:12386-12394.

Clarke, D. M., T. W. Loo, G. Inesi, and D. H. MacLennan. 1989. Location of high affinity Ca^{2+}-binding sites within the predicted transmembrane domain of the sarcoplasmic reticulum Ca^{2+}- ATPase. *Nature* 339:476-478.

Clarke, D. M., T. W. Loo, and D. H. MacLennan. 1990a. Functional consequences of alterations to amino acids located in the nucleotide binding domain of the Ca^{2+}-ATPase of sarcoplasmic reticulum. *J. Biol. Chem.* 265:22223-22227.

Clarke, D. M., T. W. Loo, and D. H. MacLennan. 1990b. Functional consequences of mutations of conserved amino acids in the β-strand domain of the Ca^{2+}-ATPase of sarcoplasmic reticulum. *J. Biol. Chem.* 265:14088-14092.

Cleland, W. W. 1982. Preparation of chromium(III) and cobalt(III) nucleotides as chirality and inhibitors. *Methods. Enzymol.* 87:159-179.

Coan, C., J. Y. Ji, and J. A. Amaral. 1994. Ca^{2+} binding to occluded sites in the CrATP-ATPase complex of sarcoplasmic reticulum: evidence for two independent high- affinity sites. *Biochemistry.* 33:3722-3731.

Csermely, P., A. Martonosi, G. C. Levy, and A. J. Ejchart. 1985. ^{51}V-n.m.r. analysis of the binding of vanadium(V) oligoanions to sarcoplasmic reticulum. *Biochem. J.* 230:807-815.

Green, N. M. and D. L. Stokes. 1992. Structural modelling of P-type ion pumps. *Acta Physiol. Scand. Suppl.* 607:59-68.

Gutierrez-Merino, C., F. Munkonge, A. M. Mata, J. M. East, B. L. Levinson, R. M. Napier, and A. G. Lee. 1987. The position of the ATP binding site on the (Ca^{2+} + Mg^{2+})-ATPase. *Biochim. Biophys. Acta* 897:207-216.

Henao, F., S. Orlowski, Z. Merah, and P. Champeil. 1992. The metal sites on sarcoplasmic reticulum membranes that bind lanthanide ions with the highest affinity are not the ATPase Ca^{2+} transport sites. *J. Biol. Chem.* 267:10302-10312.

Henderson, R., J. M. Baldwin, T. A. Ceska, F. Zemlin, E. Beckmann, and K. H. Downing. 1990. Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J. Mol. Biol.* 213:899-929.

Herrmann, T. R. and A. E. Shamoo. 1988. Estimation of inter-binding-site distances in sarcoplasmic reticulum (Ca^{2+} + Mg^{2+})-ATPase under occluded and non-occluded conditions. *Mol. Cell Biochem.* 82:55-58.

Highsmith, S. and A. J. Murphy. 1984. Nd3+ and Co^{2+} binding to sarcoplasmic reticulum CaATPase. An estimation of the distance from the ATP binding site to the high- affinity calcium binding sites. *J. Biol. Chem.* 259:14651-14656.

Kühlbrandt, W., D. N. Wang, and Y. Fujiyoshi. 1994. Atomic model of plant light-harvesting complex by electron crystallography. *Nature.* 367:614-621.

Lewis, S. M. and D. D. Thomas. 1986. Effects of vanadate on the rotational dynamics of spin-labeled calcium adenosinetriphosphatase in sarcoplasmic reticulum membranes. *Biochemistry.* 25:4615-4621.

MacLennan, D. H., C. J. Brandl, B. Korczak, and N. M. Green. 1985. Amino-acid sequence of a Ca^{2+} + Mg^{2+}-dependent ATPase from rabbit muscle sarcoplasmic reticulum, deduced from its complementary DNA sequence. *Nature* 316:696-700.

Mariano, D. D. and W. W. Cleland. 1980. Preparation and properties of chromium(III) adenosine 5'-triphosphate, chromium(III) adnosine 5'-diphosphate, and related

chromium(III) complexes. *Biochemistry.* 19:1496-1505.

Maruyama, K., D. M. Clarke, J. Fujii, G. Inesi, T. W. Loo, and D. H. MacLennan. 1989. Functional consequences of alterations to amino acids located in the catalytic center (isoleucine 348 to threonine 357) and nucleotide-binding domain of the Ca^{2+}-ATPase of sarcoplasmic reticulum. *J. Biol. Chem.* 264:13038-13042.

Meissner, G., G. E. Conner, and S. Fleischer. 1973. Isolation of sarcoplasmic reticulum by zonal centrifugation and purification of Ca^{2+}-pump and Ca^{2+}-binding proteins. *Biochim. Biophys. Acta.* 298:246-269.

Milligan, R. A. and P. F. Flicker. 1987. Structural relationships of actin, myosin and tropomyosin revealed by cryo-electron microscopy. *J. Cell Biol.* 105:29-39.

Mimori, Y., I. Yamashita, K. Murata, Y. Fujiyoshi, K. Yonekura, C. Toyoshima and K. Namba. 1995. The structure of the R-type straight flagellar filament of Salmonella at 9 Å resolution by electron cryomicroscopy. *J. Mol. Biol.* 249:69-87.

Møller, J. V., B. Juul, and M. le Maire. 1996. Structural organization, ion transport, and energy transduction of P-type ATPases. *Biochim. et Biophys. Acta.* 1286:1-51.

Ogurusu, T., S. Wakabayashi, and M. Shigekawa. 1991. Functional characterization of lanthanide binding sites in the sarcoplasmic reticulum Ca^{2+}-ATPase: do lanthanide ions bind to the calcium transport site? *Biochemistry.* 30:9966-9973.

Schulz, G. E. 1991. Domain motions in proteins. *Curr. Opin. Struc. Biol.* 1:883-888

Scott, T. L. 1985. Distances between the functional sites of the (Ca^{2+} + Mg^{2+})- ATPase of sarcoplasmic reticulum. *J. Biol. Chem.* 260:14421-14423.

Serpersu, E. H., U. Kirch, and W. Schoner. 1982. Demonstration of a stable occuluded form of Ca^{2+} by the use of the chromium complex of ATP in the Ca^{2+}-ATPase of sarcoplasmic reticulum. *Eur. J. Biochem.* 122:347-354.

Smith, P. K., R. I. Krohn, G. T. Hermanson, A. K. Mallia, F. H. Gartner, M. D. Provenzano, E. K. Fujimoto, N. M. Goeke, B. J. Olson, and D. C. Klenk. 1985. Measurement of protein using bicinchoninic acid. *Anal. Biochem.* 150:76-85.

Squier, T. C., D. J. Bigelow, J. Garcia de Ancos, and G. Inesi. 1987. Localization of site-specific probes on the Ca-ATPase of sarcoplasmic reticulum using fluorescence energy transfer. *J. Biol. Chem.* 262:4748-4754.

Squier, T. C., D. J. Bigelow, F. J. Fernandez-Belda, L. de Meis, and G. Inesi. 1990. Calcium and lanthanide binding in the sarcoplasmic reticulum ATPase. *J. Biol. Chem.* 265:13713-13720.

Stokes, D. L. and J. J. Lacapere. 1994a. Conformation of Ca^{2+}-ATPase in two crystal forms. Effects of Ca^{2+}, thapsigargin, adenosine 5'-(β,γ- methylene)triphosphate), and chromium(III)-ATP on crystallization. *J. Biol. Chem.* 269:11606-11613.

Stokes, D. L., W. R. Taylor, and N. M. Green. 1994b. Structure, transmembrane topology and helix packing of P-type ion pumps. *FEBS Lett.* 346:32-38.

Tani, K., H. Sasabe, and C. Toyoshima. 1996. A set of programs for determining defocus and astigmatism in electron images. *Ultramicroscopy.* 65:31-44.

Taylor, W. R. and N. M. Green. 1989. The predicted secondary structures of the nucleotide-binding sites of six cation-transporting ATPases lead to a probable tertiary fold. *Eur. J. Biochem.* 179:241-248.

Teruel, J. A. and J. C. Gomez-Fernandez. 1986. Distances between the functional sites of sarcoplasmic reticulum (Ca^{2+} + Mg^{2+})-ATPase and the lipid/water interface. *Biochim. Biophys. Acta.* 863:178-184.

Thorley-Lawson, D. A. and N. M. Green. 1973. Studies on the location and orientation of proteins in the sarcoplasmic reticulum. *Eur. J. Biochem.* 40:403-413.

Toyoshima, C. and N. Unwin. 1988. Ion channel of acetylcholine receptor reconstructed from images of postsynaptic membranes. *Nature* 336:247-250.

Toyoshima, C. and N. Unwin. 1990. Three-dimensional structure of the acetylcholine receptor by cryoelectron microscopy and helical image reconstruction. *J. Cell Biol.* 111:2623-2635

Toyoshima, C. , H. Sasabe, and D. L. Stokes. 1993a. Tree-dimensional cryo-electron microscopy of the calcium ion pump in the sarcoplasmic reticulum membrane. *Nature.* 362:469-471.

Toyoshima, C., K. Yonekura, and H. Sasabe. 1993b. Contrast transfer for frozen-hydrated specimens II. Amplitude contrast at very low frequencies. *Ultramicroscopy.* 48:165-176.

Unwin, N. 1993. Nicotinic acetylcholine receptor at 9 Å resolution. *J. Mol. Biol.* 229:1101-1124.

Unwin, N. 1995. Acetylcholine receptor channel imaged in the open state. *Nature.* 373:37-43.

Varga, S., P. Csermely, and A. Martonosi. 1985. The binding of vanadium (V) oligoanions to sarcoplasmic reticulum. *Eur. J. Biochem.* 148:119-126.

Vilsen, B. and J. P. Andersen. 1987. Characterization of CrATP-induced calcium occlusion in membrane- bound and soluble monomeric sarcoplasmic reticulum Ca^{2+}-ATPase. *Biochim. Biophys. Acta.* 898:313-322.

Vilsen, B., J. P. Andersen, and D. H. MacLennan. 1991. Functional consequences of alterations to amino acids located in the hinge domain of the Ca^{2+}-ATPase of sarcoplasmic reticulum. *J. Biol. Chem.* 266:16157-16164.

Vilsen, B. and J. P. Andersen. 1992. Interdependence of Ca^{2+} occlusion sites in the unphosphorylated sarcoplasmic reticulum Ca^{2+}-ATPase complex with CrATP. *J. Biol. Chem.* 267:3539-3550.

Warren, G. B., P. A. Toon, N. J. Birdsall, A. G. Lee, and J. C. Metcalfe. 1974. Reconstitution of a calcium pump using defined membrane components. *Proc. Natl. Acad. Sci. U. S. A.* 71:622-626.

Yamamoto, H., Y. Imamura, M. Tagaya, T. Fukui, and M. Kawakita. 1989. Ca^{2+}-dependent conformational change of the ATP-binding site of Ca^{2+}-transporting ATPase of sarcoplasmic reticulum as revealed by an alteration of the target-site specificity of adenosine triphosphopyridoxal. *J. Biochem.* 106:1121-1125.

Appendix

A. The format of layer-line files.
B. New programs.

A. The format of layer-line files

Fig. A.1 shows a typical layer-line file. Each file starts with a line showing the 'MODE', and the number of layer-lines constituting the layer-line file. The 'MODE' indicates what kind of information is associated with each Fourier term (see below). Individual layer-line data begin with a title line. This line consists of the layer-line number (*l*), a word which side of helical Fourier transform this file represents (one of 'NEA', 'FAR' and 'AVG'; 'AVG' means that these layer-line data are an average of several ones), a title of this file, weight, the start number (*n*) and the layer-line number (*l*). Then, follow lines showing the data associated with individual Fourier term.

The number of columns consisting these lines may be different from file to file. The first three columns represent distance from the meridian in spatial frequency ($Å^{-1}$) (radius), the amplitude (AMP) and the phase (PHS) in degrees of each Fourier term in this order. Any layer-line file must have at least these three columns. The other columns may not be included, but must be arranged in the following order: the amplitude of background (BG), the contrast transfer function (CTF), the figure of merits (FOM), the number of points contributing to this point (NPTS) and the envelope function of the CTF (ENV).

The 'MODE' located at the top of the file shows what kinds of columns are included (Fig. A.2). In this case (Fig. A.1), the file has CTF (bit 1), FOM (bit 2), NPTS (bit 3) and ENV (bit 4) columns. The mode is calculated by bit-wise OR. In this case, it is 2 + 4 +

8 + 16 = 30. If the amplitudes are divided by the CTF or the ENV, bit 7 or bit 6 is turned on. Each layer-line ends with a line consisting of columns of zero.

```
                                                  weight      n    l
                                                    ↓         ↓    ↓
MODE= 30  TOTAL LL NUMBER=256
LAYER LINE  0 AVG havgrefn 7th cycle       1.00000e+00     0    0
 0.00000e+00 3.13490e+02  180.00 -0.047  1.000  23  1.000
 4.08262e-04 2.74462e+02  180.00 -0.047  1.000  23  1.000
 8.16524e-04 1.72273e+02  180.00 -0.047  1.000  23  1.000
 1.22479e-03 4.34630e+01  180.00 -0.048  0.886  23  1.000
 1.63305e-03 6.79657e+01  360.00 -0.049  1.000  23  0.999
 2.04131e-03 1.27603e+02    0.00 -0.050  1.000  23  0.999
 2.44957e-03 1.22534e+02    0.00 -0.052  1.000  23  0.998
 2.85783e-03 6.58724e+01  360.00 -0.054  1.000  23  0.998
 3.26610e-03 1.09977e+01  180.00 -0.056  0.767  23  0.997
 3.67436e-03 7.28225e+01  180.00 -0.058  1.000  23  0.996
 4.08262e-03 9.53561e+01  180.00 -0.061  1.000  23  0.995
 4.49088e-03 7.47501e+01  180.00 -0.064  1.000  23  0.994
 4.89914e-03 2.63088e+01  180.00 -0.067  1.000  23  0.993
 5.30741e-03 2.45526e+01  360.00 -0.071  1.000  23  0.992
 5.71567e-03 5.58896e+01    0.00 -0.075  1.000  23  0.991
               ..........................
 1.44933e-01 1.37132e+00    0.00 -0.807  1.000   1  0.141
 0.00000e+00 0.00000e+00    0.00  0.000  0.000   0  0.000
LAYER LINE  2 AVG havgrefn 7th cycle       1.00000e+00   -59    2
 2.69453e-02 4.59442e+00  206.55 -0.698  1.000   1  0.750
 2.73535e-02 4.07154e+00  208.55 -0.689  0.536   7  0.800
 2.77618e-02 6.76143e-01  332.26 -0.651  0.102  18  0.840
 2.81701e-02 3.10130e+00    4.23 -0.657  0.476  21  0.836
               ..........................
      ↑         ↑          ↑     ↑     ↑    ↑    ↑
   radius      AMP        PHS   CTF   FOM  NPTS  ENV
```

FIGURE A.1 The format of a new layer-line file. Old format lacks the first line and the 4th to the 7th columns

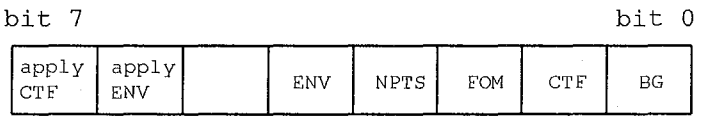


FIGURE A.2 The 'MODE' of the layer-line file. The bit for 'apply CTF' is 1 when the amplitude has been divided by CTF, or 0 otherwise; the bit for 'ENV' is 1 when the column for the envelope function of the CTF exists. The same rule applies to the other columns.

B. New programs

B-1. Programs for helical reconstruction

a. Programs for the CTF

haddctf adds the column for the CTF to the layer-line file and corrects for the phases according to the CTF.

[Control file]

```
! title
cr1428lw, add CTF
! infile
cr1428lw.nea
! outfile
cr1428lwt.nea
! x dimension of the fft frame,
! and the repeat distance (g.u).
! If multiple repeats are included, the repeat distance
! should be box y/# of repeats
1024 1010.495
! scanning interval(um) and magnification
10 41800
! mean defocus(A), astigmatism(A)
! and the angle of principal long axis measured from
 ! the X-axis(degree)
11270.2 220.0 176.4
! accel. voltage(kV), Cs(mm), amplitude contrast(%)
! [def.:200 1.9 4.7]
200 1.9 4.7
! possible error of CTF zero (g.u.)
! (for locating local minima)
! if this # is negative, no phase correction takes place.
0
```

If the effect of the envelope function is to be considered, add the following lines at the end of the control data file.

```
! Cc(mm), beam illumination angle(beta; mrad),
! dI/I, dE/E and energy spread (eV)
1.9 0.5 0.000001 0.000002 2
! ENV cut off and edge mode (default: N)
!  S: set the outermost radius of the LL data so that the CTF
!      becomes minimal after the ENV < ENV cut off.
!  N: cut LL data at the points
!                      where the ENV < ENV cut off.
```

```
0.1 S
! multiply the envelope function to the CTF ?(Y/N), default: N
N
```

haddctfw The function is the same as haddctf except that the column for the CTF is appended to the layer-line datasets from both near and far sides at the same time. The control file is the same except for the lines specifying the input and output files.

[Control file (only the lines for the input and output files)]

```
! infiles (near and far sides)
cr1428lw.nea cr1428lw.far
! outfiles (near and far sides)
cr1428lwt.nea cr1428lwt.far
```

hdivctf divides the amplitudes by the CTF so that the amplitudes are fully corrected for.

[Control file]

```
! title
cr26w23ltv, correction of CTF
! input file
cr26w23ltv.new
! out file
cr26w23ltvn.cor
! cutoff value of CTF
0.01
! FFT X size and repeat distance(g.u)
1024 1623.188
! CTF application mode (A: apply CTF correction to all
! LLs, E: apply CTF to only equator,
! W: apply both CTF and ENV to all LLs)
! [default: A]
A
```

At the points where the CTF is less than the cutoff value, amplitudes are divided by 'cutoff value'.

henvcut cuts the layer-lines data according to the envelope function. The same control file used for haddctf can be used here. The input file should be an averaged file.

nfavg averages the Fourier terms from near and far sides of a transform extracted as layer-line data. If the file contains the CTF data, this program asks as

NFAVG> Average amplitudes with CTF weighting ?[Y/N, def.: Y].

When 'Y' is entered (this is default) and the file doesn't have the the ENV data, the CTF-weighted amplitude averaging is done as in Eqs. 1 and 2 (Mimori et al., 1995). If 'N' is entered, the program averages amplitudes and the CTF term, independently (the CTF is averaged as in Eq. 2). If the file has the ENV data in addition to the CTF data and the answer to the previous question is 'Y', the program asks as

NFAVG> Average amplitudes with ENV weighting ?[Y/N, def.: Y].

When the answer is 'Y', amplitudes, the CTFs and the ENVs are averaged according to Eqs. 3, 2 and 4, respectively. If 'N' is entered, the CTF-weighted amplitude averaging is done as in Eqs. 1 and 2. The ENVs are averaged as in Eq. 4.

$$F_{avg} = \frac{F_{near} \cdot CTF_{near} + F_{far} \cdot CTF_{far}}{CTF_{near}^{2} + CTF_{far}^{2}} \cdot CTF_{avg} \quad \cdots \text{Eq. 1}$$

$$CTF_{avg} = \frac{CTF_{near}^{3} + CTF_{far}^{3}}{CTF_{near}^{2} + CTF_{far}^{2}} \quad \cdots \text{ Eq. 2}$$

$$F_{avg} = \frac{CTF_{near} \cdot ENV_{near} \cdot F_{near} + CTF_{far} \cdot ENV_{far} \cdot F_{far}}{CTF_{near}^{2} \cdot ENV_{near}^{2} + CTF_{far}^{2} \cdot ENV_{far}^{2}} \cdot CTF_{avg} \cdot ENV_{avg} \quad \cdots \text{ Eq. 3}$$

$$ENV_{avg} = \frac{ENV_{near}^{3} + ENV_{far}^{3}}{ENV_{near}^{2} + ENV_{far}^{2}} \quad \cdots \text{ Eq. 4}$$

where F_{avg}, F_{near} and F_{far} represent the structure factors of the average, near and far sides, respectively.

hlxavg averages multiple layer-line data sets. New modes are supported: In a new mode 'C', the CTF-weighted averaging is done as in Eqs. 5 and 6 (Mimori et al., 1995). At the same time, the FOM is calculated as in Eq. 9. In mode 'W', amplitudes, the CTFs and the ENVs are averaged according to Eqs. 7, 6 and 8, respectively.

$$F_{avg} = \frac{\sum_{i=1}^{n} (CTF_i \cdot F_i)}{\sum_{i=1}^{n} CTF_i^{2}} \cdot CTF_{avg} \quad \cdots \text{ Eq. 5}$$

$$CTF_{avg} = \frac{\sum_{i=1}^{n} CTF_i^{3}}{\sum_{i=1}^{n} CTF_i^{2}} \quad \cdots \text{ Eq. 6}$$

$$F_{avg} = \frac{\sum_{i=1}^{n} (CTF_i \cdot ENV_i \cdot F_i)}{\sum_{i=1}^{n} (CTF_i \cdot ENV_i)^{2}} \cdot CTF_{avg} \cdot ENV_{avg} \quad \cdots \text{ Eq. 7}$$

$$ENV_{avg} = \frac{\sum_{i=1}^{n} ENV_i^3}{\sum_{i=1}^{n} ENV_i^2} \quad \cdots \text{ Eq. 8}$$

$$FOM = \frac{|F_{avg}|}{\sum_{i=1}^{n} |Fcor_i|} \quad \cdots \text{ Eq. 9}$$

where n is the number of the files to be averaged. $Fcor_i$ is the structure factor weighted by the CTF or by the CTF and ENV.

hctfrefn refines the CTF parameters (mean defocus and astigmatism). The refinement is done by minimizing the phase residual defined by Eq. 10.

$$residual = \frac{\sum amp \cdot |phase_{ref} - phase_{test}|}{\sum amp} \quad \cdots \text{ Eq. 10}$$

where *amp* is defined by Eq. 11 or Eq. 12; summation is taken over all of the points with the amplitudes of the reference file higher than the cut off value.

If the answer to 'multiply CTF ?' is 'Y',

$$amp = amp_{ref} \cdot CTF_{ref} + amp_{test} \cdot CTF_{test} \cdot scale \quad \cdots \text{ Eq. 11}$$

otherwise,

$$amp = amp_{ref} + amp_{test} \cdot scale \quad \cdots \text{ Eq. 12}$$

where *scale* refers to the scaling factor determined by hlxfit.

The reference file is obtained by averaging many layer-line datasets with different defocus values. Because the data of test

files (near and far sides) are usually very noisy, the layer-lines contributing to the comparison should be selected by hampsort. hampsort chooses the layer-lines according to the order of the sum of amplitudes. Good results have been obtained with 50 - 80 layer-lines for the reference file. To make the CTF definition file (default name, ctf.def), run mkctffl.

[Control file]

```
! 1. title
cr3515blw, CTF refinement
! 2. reference file name (CTF should not have been applied)
cr26w23lchs.2fd
! 3. max # of layer-lines to compare (0 for all)
0
! 4. cut-off amplitude (% of highest off-meridional)
2
!5
! 5. special LLs: number and weight
0 0
! 6. test(nea) file name
cr3515blw.nea
! 7. phi shift, z shift for near side
-21.801 3.406
! 8. test(far) file name
cr3515blw.far
! 9. phi shift, z shift for far side
-21.801 3.406
! 10. radial and amplitude scaling factors for test file
!       from hlxfit.
0.9910 0.7081
! 11. ipole, repeat distance (Angstrom) used for hlxfit
0 4886.803
! 12. image frame x size and repeat distance(g.u.)
1024 2042.978
! 13. scanning interval(um) and magnification
10 41800
! 14. initial defmean(A), defastig(A) and angle(degree)
! of long axis from X axis (counter clockwise)
!9068.2 191.3 -88.7    ! the values determined by 'pltctfx'
15000 2000 180        ! arbitrary values for test
! 15. initial intervals for defmean(A), defastig(A) and
! angle (degree)
100 50 5
! 16. CTF definition file made by mkctffl
../ctf.def
! 17. defocus range [min and max values(A)]
3000 20000
! 18. resolution range to calculate the residuals [min and
! max (g.u.)]
0 250
! multiply CTF to amplitudes for refinement ?
Y
```

[Result]

```
--- HCTFREFN --- Ver.  1.05   Wed Dec  4 23:24:49 1996


 --- cr3515blw, CTF refinement
[ref]      cr26w23lchs.2fd   : cr3515blw.nea and cr3515blw.far [test]
          control data file : hctfrefn.cnt
         title of ref. file : 2fd of cr26
   tiltle of test(nea) file : cr3515bl, only 1
   tiltle of test(far) file : cr3515bl, only 1
               side of ref. : Avg
        layer-line sampling :   0.000408 (ref.)   0.000408 (test) Angstrom inv
amplitude cut-off level :      11.32 (  2.00% of max amp. multiplied by CTF)
    FFT x frame size :  1024 g.u
     repeat distance : 2042.98 g.u.
   scanning interval : 10.00 um
       magnification : 41800 x

--- Near side ---
 rotation around the helix axis:  -21.801 degree, z shift:    3.406 Angstrom
--- Far side ---
 rotation around the helix axis:  -21.801 degree, z shift:    3.406 Angstrom

 Radial scaling (Test): 0.9910
    Amp scaling (Test): 0.7081
 repeat distance used for phase origin shift: 4886.803 Ang.


--- CTF information ---
File name   : ../ctf.def
Data type   : Image  (real    4 bytes)
Size        :    257 x    271 x      1 (x x y x z)
Origin      :   0.00(g.u), 3000.00(A)
Number of points within one unit cell:
            :    257 x    271 x      1 (x x y x z)
Max space freq.(Angstrom inv) :   0.418000
Max defocus (Angstrom)        :   30000.00
Accel voltage:    200 kV, Cs: 1.9000 mm, Amplitude cont.: 0.0470
Scaning interval:  10.00(um), Magnification: x 41800
Density     :    -1.0000 to     1.0000 (mean:    -0.0981)
Labels      :
          1 : MKCTFFL : accelv= 200, Cs=1.90, amp=0.047, scan=10.0um, mag= 41800
          2 : MKCTFFL of ctf.def; defocus min=3000.0 max=30000.0 step= 100.0 Sun Oct
13 18:38:


--- Initial values ---
 Residual  :  70.508
 at defmean : 15000.0 Ang.
      astig : 2000.0 Ang.
        ang :  180.0 deg.



--- cycle   1 ---
```

```
 Residual  :  69.013
     at defmean : 14700.0 (14700.0-15200.0, step  100.0)
          astig : 1850.0 (1850.0-2100.0, step   50.0)
            ang :  180.0 ( 165.0- 190.0, step    5.0)

--- cycle   2 ---
 Residual  :  68.173
     at defmean : 14400.0 (14400.0-14900.0, step  100.0)
          astig : 1700.0 (1700.0-1950.0, step   50.0)
            ang :  180.0 ( 165.0- 190.0, step    5.0)

        ........................

--- cycle  29 ---
 Residual  :  51.891
     at defmean : 9222.2 (9200.0-9255.6, step   11.1)
          astig :  416.7 ( 406.7- 456.7, step   10.0)
            ang :  269.0 ( 266.0- 271.0, step    1.0)

--- cycle  30 ---
 Residual  :  51.889
     at defmean : 9212.2 (9192.2-9242.2, step   10.0)
          astig :  426.7 ( 386.7- 436.7, step   10.0)
            ang :  269.0 ( 266.0- 271.0, step    1.0)


--- Final parameters ----
 minimum residual:    51.8894
     at mean defocus : 9212.2 Ang.
         astigmatism :  426.7 Ang.
               angle :  269.0 deg.

 R factor: 0.4629 (Amplitude scaling factor for test: 0.5822)
 Number of ref. points used:       1732 (out of      22547)
 Sum of ref. amplitudes used:   48890.88 (out of  177054.94)
 Number of test points used:       1732 (out of      21530)
 Sum of test amplitudes used:   83970.98 multiplied by CTF
```

	L	N	1/Z (A)	weight	point (g.u)	ref-amp-test (%)		final res. (degree)	init. res.
1:	0	0	0.0	0.00	-- (-- - --)	--	--	(--)	
2:	5	-59	977.5	1.00	63 (70-249)	3.35	3.32	54.027	(72.425)
3:	10	-118	488.8	1.00	21 (140-196)	0.81	0.58	66.082	(123.151)
4:	12	92	407.3	1.00	33 (108-230)	1.31	1.39	71.317	(109.388)
5:	17	33	287.5	1.00	84 (38-215)	9.45	7.93	29.928	(30.932)
6:	22	-26	222.2	1.00	87 (29-224)	9.77	8.50	26.074	(25.939)
7:	27	-85	181.0	1.00	36 (100-231)	1.29	1.64	60.320	(108.482)
8:	34	66	143.8	1.00	50 (77-243)	2.35	2.73	47.621	(88.692)
9:	39	7	125.3	1.00	108 (20-232)	9.86	8.71	34.299	(42.047)
10:	44	-52	111.1	1.00	69 (65-244)	3.25	3.48	63.627	(89.250)
11:	51	99	95.8	1.00	20 (121-244)	0.66	0.65	85.192	(93.440)
12:	56	40	87.3	1.00	70 (49-247)	4.91	4.90	41.566	(52.058)
13:	61	-19	80.1	1.00	85 (30-246)	4.09	4.84	47.647	(62.778)
14:	66	-78	74.1	1.00	38 (91-140)	2.49	3.30	30.118	(75.213)
15:	73	73	67.0	1.00	44 (85-196)	2.06	2.19	77.865	(106.151)
16:	78	14	62.7	1.00	67 (39-212)	3.71	3.80	48.503	(56.314)
17:	83	-45	58.9	1.00	62 (51-245)	4.40	4.17	37.915	(57.481)

```
18:   88 -104   55.5   1.00   19 (134-234)   0.54   0.39   82.850 (109.316)
19:   95   47   51.4   1.00   68 ( 53-194)   4.93   4.70   47.598 ( 49.040)
20:  100  -12   48.9   1.00   57 ( 50-229)   3.05   3.30   47.595 ( 70.080)
21:  105  -71   46.5   1.00   29 ( 91-214)   1.34   1.75   38.719 ( 71.166)
22:  112   80   43.6   1.00   26 ( 96-131)   1.52   1.67   65.850 (109.109)
23:  117   21   41.8   1.00   50 ( 58-227)   2.41   3.42   44.363 ( 58.871)
24:  122  -38   40.1   1.00   74 ( 61-240)   4.78   5.83   49.620 ( 74.047)
25:  127  -97   38.5   1.00   11 (118-218)   0.32   0.15  112.863 ( 84.947)
26:  134   54   36.5   1.00   42 ( 67-206)   1.95   2.28   62.673 (114.504)
27:  139   -5   35.2   1.00   38 ( 69-216)   2.66   3.14   39.597 ( 72.034)
28:  144  -64   33.9   1.00   23 ( 76-234)   0.71   0.88   59.183 ( 85.345)
29:  149 -123   32.8   1.00    8 (147-210)   0.23   0.26   25.135 ( 26.470)

                .........................

78:  512   32    9.5   1.00    0 (  0-  0)   0.00   0.00    0.000 (  0.000)
79:  534    6    9.2   1.00    0 (  0-  0)   0.00   0.00    0.000 (  0.000)
80:  573   13    8.5   1.00    0 (  0-  0)   0.00   0.00    0.000 (  0.000)
-------------------------------------------------------------------------------
Residual:  51.889  at defmean : 9212.2, astig :  426.7, ang:  269.0
```

In this example, the initial residual, 70.508° for the arbitrary defocus parameters was reduced to 51.889° after 30 cycles of refinements. Resulting defocus parameters were very close to the correct values (mean defocus = ~ 9000 Å, astigmatism = ~ 200 Å and angle of the long axis from X = ~ -90°), determined by pltctfx (Tani et al. 1996).

b. General programs

llprx2 displays the layer-line data in the X-Windows and makes a control file for hlxedt. This program is easy to use due to its GUI (Fig. II.2). Control files for hlxedt can be made as follows. At first, place the mouse cursor at the start point (i.e. the smallest radius) and press the left button to record the point. The parameters (AMP, PHS, BG, CTF, FOM, NPTS and ENV) of this point are displayed in the bottom left corner of the window. Then, press the

right button to make an entry in the control file for hlxedt. Do the same thing for the end point at the largest radius. The FOM (green dotted curve in the lower panel) is normalized according to Eq. 13.

$$FOM_{normalized} = FOM \cdot \sqrt{NPTS} \qquad \cdots \text{ Eq. 13}$$

In this definition, the maximum FOM value becomes $FOM \cdot \sqrt{NPTS}_{\max}$ and the random level always becomes $\left(1/\sqrt{NPTS}\right) \cdot \sqrt{NPTS} = 1$. The merits of using this kind of normalized FOM are several folds: this FOM provides consistent number even though the number of points contributing to each data point varies. Consequently, different datasets constituting of different number of images can be compared in a consistent way. The raw FOM is displayed by llprx2f. llprps, llprpsf make the postscript file of the layer-line data.

hlmtres limits layer-line data according to the resolution. Both start and end points are selected to be terminated at local minima of the amplitudes.

hedgchk removes leading or trailing points with the amplitudes zero.

hminend cuts layer-line data at the points where the amplitudes become local minima.

hampsort selects the layer-lines by ascending order of the sum of amplitudes. See hctfrefn.

hlx2fld, hlxfit These programs now calculate the residuals for different resolution ranges.

[Control file for hlx2fld]

```
! title
2fd of cr26
! input file
cr2567all.new
! number of layer lines(0: for all) and
! repeat distance(Angstrom)
0 3883.225
! Amp cut off(% of maximum amp)
0.1
! list of layer-lines with weights different from 1.0
! (L and weight)
0 0
! azimuthal rotation (start, end and stepsize; in degree)
! the order of start and end does not matter
-1 1 0.01
! shift along helix axis(Angstrom)
-1 1 0.01
! output file if needed
!cr2567all.2fd
! Mode (Q: calculate residuals for different resolution
! rings and the number of resolution rings
Q 5
! highest resolution limit for each ring (Angstrom)
20
15
10
8
7
```

[Result]

```
--- HLX2FLD --- Ver.  3.30   Thu Dec  5 20:43:32 1996

title of the layer-line : lg from cr2567all
              new title : 2fd of cr26
                   side : Avg
                     in : cr2567all.new
        repeat distance :    3883.23 Angstrom
    layer-line sampling :   0.000408 reciprocal Angstrom
      maximum amplitude :   24087.70
       cutoff amplitude :      24.09 (  0.10%)
  sum of amplitude used :  2818115.2 out of  2826594.3
  number of points used :      11887 out of      12404 (   95.83%)
        control data in : cr26tf.cnt

--- Initial residual table
-------------------------------------------------
 z \ phi |  -1.000  -0.600  -0.200   0.200   0.600
-------------------------------------------------
  -1.000:  30.073  25.619  19.230  19.471  27.199
  -0.600:  30.147  25.566  17.641  17.929  26.414
  -0.200:  30.276  25.654  16.896  16.876  25.781
   0.200:  30.520  25.876  17.135  16.673  25.423
   0.600:  30.922  26.397  18.331  17.418  25.386
-------------------------------------------------

--- Final residual table
-------------------------------------------------
```

```
 z \ phi |  -0.018  -0.008   0.002   0.012   0.022
-------------------------------------------------
   0.017:  14.027  13.984  13.969  13.980  14.014
   0.027:  14.030  13.983  13.966  13.976  14.009
   0.037:  14.034  13.985  13.964  13.973  14.005
   0.047:  14.040  13.988  13.965  13.971  14.002
   0.057:  14.047  13.993  13.968  13.971  14.000
-------------------------------------------------

    minimum residual:   13.964  at phi=    0.002, z=    0.037

Phase residual for each layer-line at the 2-fold origin
```

	L	N	1/Z (A)	weight	point (g.u)	amp (%)	residual --20.00(A) (degree)	--15.00(A)	--10.00(A)	-- 8.00(A)	-- 7.00(A)	Over all
1:	0	0	0.0	0.00	-- (0-235)	--	--					
2:	2	-59	1941.6	1.00	204 (68-280)	1.09	3.2(52)	19.5(40)	35.0(77)	28.2(35)	0(0)	14.6
3:	4	-118	970.8	1.00	110 (137-248)	0.39	0(0)	23.4(26)	38.9(80)	47.2(4)	0(0)	34.3
4:	6	-177	647.2	1.00	34 (267-300)	0.19	0(0)	0(0)	0(0)	28.3(34)	0(0)	28.3
5:	8	-236	485.4	1.00	20 (283-302)	0.08	0(0)	0(0)	0(0)	33.4(20)	0(0)	33.4
6:	9	210	431.5	1.00	12 (273-285)	0.09	0(0)	0(0)	0(0)	42.1(12)	0(0)	42.1
7:	11	151	353.0	1.00	44 (228-273)	0.19	0(0)	0(0)	26.3(17)	24.8(27)	0(0)	25.5
8:	13	92	298.7	1.00	71 (106-178)	0.35	18.9(17)	26.7(39)	32.5(15)	0(0)	0(0)	26.3
9:	15	33	258.9	1.00	194 (35-231)	5.37	1.4(88)	16.2(39)	25.8(67)	0(0)	0(0)	3.1
10:	17	-26	228.4	1.00	162 (26-191)	7.94	1.5(96)	14.5(40)	12.4(26)	0(0)	0(0)	2.1
11:	19	-85	204.4	1.00	55 (101-155)	0.33	6.7(21)	19.0(34)	0(0)	0(0)	0(0)	12.8
12:	21	-144	184.9	1.00	76 (196-272)	0.32	0(0)	0(0)	36.0(48)	30.4(28)	0(0)	33.4
13:	23	-203	168.8	1.00	12 (262-275)	0.04	0(0)	0(0)	0(0)	29.1(12)	0(0)	29.1
14:	24	243	161.8	1.00	11 (290-300)	0.04	0(0)	0(0)	0(0)	22.0(11)	0(0)	22.0
15:	26	184	149.4	1.00	36 (217-255)	0.11	0(0)	0(0)	23.3(26)	37.1(10)	0(0)	27.2
16:	28	125	138.7	1.00	23 (191-213)	0.07	0(0)	0(0)	40.8(23)	0(0)	0(0)	40.8
17:	30	66	129.4	1.00	187 (75-279)	0.76	7.2(46)	19.0(41)	40.1(74)	22.6(26)	0(0)	18.9
18:	32	7	121.4	1.00	227 (4-234)	15.15	1.2(117)	8.2(39)	30.7(71)	0(0)	0(0)	2.0
19:	34	-52	114.2	1.00	183 (63-253)	1.10	7.1(57)	19.6(41)	25.8(76)	32.0(9)	0(0)	13.7
20:	36	-111	107.9	1.00	46 (212-260)	0.18	0(0)	0(0)	34.6(29)	24.5(17)	0(0)	29.6
21:	38	-170	102.2	1.00	20 (213-232)	0.08	0(0)	0(0)	19.6(20)	0(0)	0(0)	19.6
22:	41	217	94.7	1.00	28 (269-296)	0.15	0(0)	0(0)	0(0)	49.9(28)	0(0)	49.9
23:	43	158	90.3	1.00	72 (211-289)	0.31	0(0)	0(0)	29.0(29)	29.4(43)	0(0)	29.3
24:	45	99	86.3	1.00	33 (141-173)	0.11	0(0)	34.6(20)	31.4(13)	0(0)	0(0)	33.0
25:	47	40	82.6	1.00	151 (44-199)	1.75	3.8(75)	17.8(41)	26.7(35)	0(0)	0(0)	6.4
26:	49	-19	79.2	1.00	138 (19-161)	1.89	6.6(97)	18.7(40)	4.3(1)	0(0)	0(0)	7.3
27:	51	-78	76.1	1.00	61 (93-154)	0.54	5.4(26)	31.8(35)	0(0)	0(0)	0(0)	12.8
28:	53	-137	73.3	1.00	24 (188-212)	0.09	0(0)	0(0)	35.4(24)	0(0)	0(0)	35.4
29:	55	-196	70.6	1.00	17 (256-274)	0.06	0(0)	0(0)	0(0)	30.9(17)	0(0)	30.9
30:	58	191	67.0	1.00	10 (287-296)	0.07	0(0)	0(0)	0(0)	30.7(10)	0(0)	30.7
					.....							
170:	399	117	9.7	1.00	22 (152-174)	0.13	0(0)	0(0)	0(0)	16.1(22)	0(0)	16.1
171:	401	58	9.7	1.00	89 (69-159)	0.35	0(0)	0(0)	0(0)	37.9(89)	0(0)	37.9
172:	403	-1	9.6	1.00	18 (8- 25)	0.10	0(0)	0(0)	0(0)	44.3(18)	0(0)	44.3
173:	405	-60	9.6	1.00	21 (99-119)	0.12	0(0)	0(0)	0(0)	47.7(21)	0(0)	47.7
174:	407	-119	9.5	1.00	26 (139-166)	0.10	0(0)	0(0)	0(0)	27.8(26)	0(0)	27.8
175:	416	91	9.3	1.00	22 (134-155)	0.12	0(0)	0(0)	0(0)	20.8(22)	0(0)	20.8
176:	418	32	9.3	1.00	57 (57-115)	0.21	0(0)	0(0)	0(0)	24.7(57)	0(0)	24.7
177:	420	-27	9.2	1.00	28 (30- 60)	0.13	0(0)	0(0)	0(0)	48.3(28)	0(0)	48.3
178:	422	-86	9.2	1.00	17 (116-132)	0.08	0(0)	0(0)	0(0)	27.8(17)	0(0)	27.8
179:	433	65	9.0	1.00	13 (79- 92)	0.04	0(0)	0(0)	0(0)	22.0(13)	0(0)	22.0
180:	435	6	8.9	1.00	100 (24-125)	0.41	0(0)	0(0)	0(0)	33.0(100)	0(0)	33.0
181:	437	-53	8.9	1.00	20 (60- 79)	0.09	0(0)	0(0)	0(0)	40.1(20)	0(0)	40.1
182:	452	-20	8.6	1.00	32 (21- 53)	0.20	0(0)	0(0)	0(0)	38.4(32)	0(0)	38.4
183:	467	13	8.3	1.00	59 (12- 73)	0.23	0(0)	0(0)	0(0)	32.6(59)	0(0)	32.6
184:	469	-46	8.3	1.00	24 (53- 76)	0.15	0(0)	0(0)	0(0)	17.3(24)	0(0)	17.3
185:	484	-13	8.0	1.00	9 (12- 21)	0.03	0(0)	0(0)	0(0)	31.8(9)	0(0)	31.8

```
Total
 --20.00(A)  --15.00(A)  --10.00(A)  -- 8.00(A)  -- 7.00(A)
 2.52(1712) 18.27(2130) 34.73(4489) 31.94(3556)  0.00( 0)
  ( 98.45%)  ( 96.08%)  ( 93.76%)  ( 97.16%)  ( 0.00%)
   ( 1739)    ( 2217)    ( 4788)    ( 3660)    (    0)
 13.964 degrees at (phi=  0.002, z=  0.037)
```

In the case of hlxfit, add the following lines at the bottom of the original control file.

```
! the number of resolution rings
5
! highest resolution for each ring (Angstrom)
15
10
8
7
```

lldatahd displays the header portion of a layer-line file.

lghd displays the header portion of a little g file.

savglq averages the qualities of the layer-line file over specified number of sector.

[Control file]

```
! 1. title
FM
! 2. infile (layer-line data)
cr26w23ltv9.new
! 3. outfile
cr26.fm
! 4. which data type to be averaged ?
! (A: AMP, B: BG, C: CTF, F: FMERIT,
!  N: NPTS, E: ENV, D: NPTS weighted FMERIT)
F
! 5. FFT X size and repeat distance(g.u.)
1024 1623.188
! 6. Average mode
! (N: Average over all data, P: Average only the peeks)
!P
N
! number of sectors
1
```

hfmcount counts the numbers of points where the values of the figure of merit are within a specified range.

[Control file]

```
! 1. title
FM of cr26w23ltv9.new
! 2. infile (layer-line data)
cr26w23ltv9.new
! 3. FFT X size and repeat distance(g.u.)
!512 811.594
1024 1623.188
! 4. number of sector
1
! 5. number of resolution table
6
! 6. resolution
20
```

```
15
12
11
10
9
```

[Result]

```
HFMCOUNT: Vers.  1.00 ---   Sat Jun 29 17:44:37 1996


        layer-line file  : cr26w23ltv9.new
                  title : FM of cr26w23ltv9.new
             FFT sizeof : 1024 x 1623.19
         control data in : cr26fmcount.cnt
        number of secter : 5

--- reading trnsform --- Sat Jun 29 17:44:37 1996

 normalized max FOM : 4.796

 Sectin : 1
   res :       20.00(A)       15.00(A)       12.00(A)       11.00(A)       10.00(A)        9.00(A)
FOM table ---
     1 :    69( 11.40%)   193( 31.48%)   300( 46.08%)   160( 46.51%)   144( 34.62%)   179( 41.15%)
     2 :   100( 16.53%)   281( 45.84%)   303( 46.54%)   164( 47.67%)   232( 55.77%)   227( 52.18%)
     3 :   114( 18.84%)   121( 19.74%)    47(  7.22%)    20(  5.81%)    40(  9.62%)    29(  6.67%)
     4 :   115( 19.01%)    18(  2.94%)     1(  0.15%)     0(  0.00%)     0(  0.00%)     0(  0.00%)
     5 :   207( 34.21%)     0(  0.00%)     0(  0.00%)     0(  0.00%)     0(  0.00%)     0(  0.00%)
 total :   605            613            651            344            416            435

 Sectin : 2
   res :       20.00(A)       15.00(A)       12.00(A)       11.00(A)       10.00(A)        9.00(A)
FOM table ---
     1 :    73( 15.70%)   182( 35.41%)   276( 44.66%)   152( 54.48%)   126( 40.13%)    78( 44.07%)
     2 :    95( 20.43%)   237( 46.11%)   294( 47.57%)   110( 39.43%)   153( 48.73%)    89( 50.28%)
     3 :    94( 20.22%)    75( 14.59%)    48(  7.77%)    17(  6.09%)    33( 10.51%)    10(  5.65%)
     4 :   113( 24.30%)    20(  3.89%)     0(  0.00%)     0(  0.00%)     2(  0.64%)     0(  0.00%)
     5 :    90( 19.35%)     0(  0.00%)     0(  0.00%)     0(  0.00%)     0(  0.00%)     0(  0.00%)
 total :   465            514            618            279            314            177

 Sectin : 3
   res :       20.00(A)       15.00(A)       12.00(A)       11.00(A)       10.00(A)        9.00(A)
FOM table ---
     1 :    62( 15.58%)   101( 26.37%)   198( 45.31%)   125( 48.64%)    80( 35.40%)   150( 52.26%)
     2 :    96( 24.12%)   154( 40.21%)   204( 46.68%)   110( 42.80%)   112( 49.56%)   107( 37.28%)
     3 :    77( 19.35%)   105( 27.42%)    35(  8.01%)    22(  8.56%)    34( 15.04%)    30( 10.45%)
     4 :    76( 19.10%)    23(  6.01%)     0(  0.00%)     0(  0.00%)     0(  0.00%)     0(  0.00%)
     5 :    87( 21.86%)     0(  0.00%)     0(  0.00%)     0(  0.00%)     0(  0.00%)     0(  0.00%)
 total :   398            383            437            257            226            287

 Sectin : 4
   res :       20.00(A)       15.00(A)       12.00(A)       11.00(A)       10.00(A)        9.00(A)
FOM table ---
     1 :    31( 11.36%)    70( 25.00%)   165( 44.24%)    69( 44.52%)    63( 36.21%)   107( 43.85%)
     2 :    39( 14.29%)   123( 43.93%)   184( 49.33%)    78( 50.32%)   107( 61.49%)   119( 48.77%)
     3 :    50( 18.32%)    60( 21.43%)    24(  6.43%)     8(  5.16%)     4(  2.30%)    18(  7.38%)
     4 :    54( 19.78%)    27(  9.64%)     0(  0.00%)     0(  0.00%)     0(  0.00%)     0(  0.00%)
     5 :    99( 36.26%)     0(  0.00%)     0(  0.00%)     0(  0.00%)     0(  0.00%)     0(  0.00%)
 total :   273            280            373            155            174            244

 Sectin : 5
   res :       20.00(A)       15.00(A)       12.00(A)       11.00(A)       10.00(A)        9.00(A)
FOM table ---
     1 :    10(  9.43%)    20( 15.15%)    58( 34.94%)    54( 57.45%)    33( 35.87%)    78( 43.33%)
     2 :    10(  9.43%)    39( 29.55%)    82( 49.40%)    38( 40.43%)    42( 45.65%)    86( 47.78%)
     3 :    19( 17.92%)    35( 26.52%)    26( 15.66%)     2(  2.13%)    17( 18.48%)    16(  8.89%)
     4 :    23( 21.70%)    27( 20.45%)     0(  0.00%)     0(  0.00%)     0(  0.00%)     0(  0.00%)
     5 :    44( 41.51%)    11(  8.33%)     0(  0.00%)     0(  0.00%)     0(  0.00%)     0(  0.00%)
 total :   106            132            166             94             92            180
```

c. Refinement programs

hboxrefn refines hlxbox parameters (horizontal/vertical position, repeat distance and rotation angles of the helix axis). This program uses many others (hlxbox, fftrans, hftstatt, srch, srchaid, hlxfl and hlx2fld/hlxfit) by Unix system function call. The results of the trials are written in 'hbox.out'.

[Control file]

```
! title
cr1544al
! control file of hlxbox
cr1544al.box
! mode (A: rotation angle of the helix axis,
! V: repeat distance, P: vertical position of the box)
A
! del steps and # of cycles for refinement
0.05 11          ! degree
! FFT frame size X, Z
1024 2048
! edt file, control files for hftstatt (near and far sides)
cr1544al.edt hstatn.cnt  hstatf.cnt
! control file of srchaid
cr1544al.sra
! control file of srch
cr1544al.src
! control file of hlxfl
cr1544al.hlx
! command for residual evaluation.
! hlx2fld or hlxfit
hlxfit
! control file for above command
cr1544al.fit
```

The control files for other programs are used as only templates for each trial. New control files, 'hstatn.cnt' and 'hstatf.cnt' are needed for hftstatt as separate files.

[hstatn.cnt (for near side)]

```
cr1590l.fft      ! FFT file name
1          ! number of repeats
N          ! the side.
```

[Result]

```
HBOXREFN: refines box parameters of helical particles. --- Ver.   1.20 ---
 --- Wed Dec  4 21:55:38 1996


      Title : cr1544al
   Box file : cr1544al.box
 Type of Refinement : Rotation angle (step:  0.050 degree)
       No. of cycles : 11
   FFT frame size(g.u.) : x: 1024 z: 2048
  hftstatt control file : ../cr1544al.edt, ../hstatn.cnt and ../hstatf.cnt
  srchaid and srch file : ../cr1544al.sra and ../cr1544al.src
     hlxfl control file : ../cr1544al.hlx
  command for residuals : hlxfit (control file : ../cr1544al.fit)

--- cycle   1 ---

Rotaion Angle of helix axis: -0.0844
hftstatt    Near side amplitude:  381204.06
hftstatt     Far side amplitude:  301807.44
srch  Qmin 15.320, Omega   1.04, x-shift   0.01 (  70 points)
hlxfit
  residual= 27.5401, phi=-29.956, z= 16.160, Rscal=1.0150,
  R=  0.3005, x amp= 1.3151, test amp=  67357.74 (5.000%)

              ........................

--- cycle   4 ---

Rotaion Angle of helix axis:  0.0656
hftstatt    Near side amplitude:  601488.94
hftstatt     Far side amplitude:  472611.38
srch  Qmin 15.320, Omega   1.04, x-shift   0.01 (  70 points)
hlxfit
  residual= 24.7297, phi=-29.946, z= 16.089, Rscal=1.0120,
  R=  0.2927, x amp= 1.2477, test amp=  72689.48 (5.000%)


--- cycle   5 ---

Rotaion Angle of from helix axis:  0.1156
hftstatt    Near side amplitude:  360099.41
hftstatt     Far side amplitude:  286836.53
srch  Qmin 15.320, Omega   1.04, x-shift   0.01 (  70 points)
hlxfit
  residual= 24.3780, phi=-29.947, z= 16.006, Rscal=1.0120,
  R=  0.2950, x amp= 1.2379, test amp=  73779.43 (5.000%)


--- cycle   6 ---

Rotaion Angle of helix axis:  0.1656
hftstatt    Near side amplitude:  107494.73
```

```
hftstatt     Far side amplitude:   92424.91
srch  Qmin 15.320, Omega   1.04, x-shift   0.01 (  70 points)
hlxfit
  residual= 25.0658, phi=-29.897, z= 16.026, Rscal=1.0110,
  R=  0.2963, x amp= 1.2399, test amp=  73670.43 (5.000%)

              ........................

--- cycle  11 ---

Rotaion Angle of helix axis:  0.4156
hftstatt    Near side amplitude: 1279221.50
hftstatt     Far side amplitude:  995906.81
srch  Qmin 15.320, Omega   1.04, x-shift   0.01 (  70 points)
hlxfit
  residual= 31.1295, phi=-29.931, z= 15.692, Rscal=1.0110,
  R=  0.3290, x amp= 1.3297, test amp=  69078.94 (5.000%)
```

In this example, the program has refined the rotation angle of the helix axis. In the 5th cycle, the residual of 'hlxfit' has become minimum (24.378°) and the sum of amplitudes maximum (73779.43). Therefore the rotation angle = 0.1156° at the 5th cycle is considered to be the most accurate value for this image.

havgrefn refines the averaging parameters, the rotation around the helix axis, the shift along the helix axis, the radial scaling factor and the amplitude scaling factor by iteration of hlxfit. The result are written in 'havg.out'. At the same time, the control files for hlxavg for each cycle are stored.

[Control file]

```
! reference file name for hlxfit
cr25avg12.2fd
! max # of layer-lines to compare (0 for all)
! and repeat distance in Angstrom
0 3767.943
! cut-off amplitude (% of highest off-meridional)
! and penalty (1/0) for non-exisiting portion
5 0 0
! LL number and its specific weight
0 0
! output file for hlxavg
cr25avg12.avg
```

```
! sampling size and frame size (e.g. 4.3 A and 512);
2.392 1024
! Average mode(see. hlxavg)
C
! template of control file for hlxavg
havg.cnt
! use hlx2fld (Y/N) and if Y, its controle file name
Y cr25tf.cnt
! the number of refinement cycles and the number of sides
! to be averaged
5 12
! test file
cr9901ltv.ren
! iside (1 for far; 0 for near)
! and ipole (1 if polarity to be changed)
0 0
! azimuth range
! (start value, [end value, [final step_size]])
-10 10 0.01
! z range (ibid)
-10 10 0.01
! radial range (start, end and step_size)
1.019 1.01 0.001
!amplitude scaling factor( if needed)
! 2nd file
cr532ltv.ren
0 0
-10 -10 0.01
-10 -10 0.01
1.015 1.024 0.001
cr532cltv.ren
0 0
-10 -10 0.01
-10 -10 0.01
1.032 1.023 0.001
cr530ltv.ren
0 0
-10 -10 0.01
-10 -10 0.01
1.008 1.017 0.001
cr530bltv.ren
0 0
-10 -10 0.01
-10 -10 0.01
1.008 1.017 0.001
cr991ltv.ren
0 0
-10 -10 0.01
-10 -10 0.01
1.048 1.039 0.001
cr5286ltv.ren
0 0
-10 -10 0.01
-10 -10 0.01
1.029 1.02 0.001
cr5263bltv.avg
0 0
-10 -10 0.01
```

```
        -10 -10 0.01
        0.996 1.005 0.001
        cr5194ltv.ren
        0 0
        -10 -10 0.01
        -10 -10 0.01
        1.001 1.010 0.001
        cr1021ltv.ren
        0 0
        -10 -10 0.01
        -10 -10 0.01
        1.022 1.031 0.001
        cr1087ltv.ren
        0 0
        -10 -10 0.01
        -10 -10 0.01
        1.048 1.057 0.001
        cr1527ltv.ren
        0 0
        -10 -10 0.01
        -10 -10 0.01
        1.032 1.041 0.001
```

[Result]

```
HAVGREFN: refines fitting parameters of helical particles. --- Ver.   1.20 ---
 --- Mon Jul  1 00:27:31 1996


 Reference file : cr25avg12.2fd
 hlxavg section
      average file name : cr25avg12.avg
      average mode : C
      pitch : 3767.943A,  samling :  2.392A,
      FFT X frame size : 1024
      output control file : havg.cnt
 No. of cycles : 5 and No. of file : 12
 Use hlx2fld and its controle file: cr25tf.cnt


--- cycle   1 ---
cr9901ltv.ren
  residual= 20.8076, phi=  4.435, z= 10.321, Rscal=1.0150,
  R=  0.2360, x amp= 1.3387, test amp=  37846.47 (5.000%)
Amplitude scaling for hlxavg:  1.3387

cr532ltv.ren
  residual= 18.4700, phi=  0.074, z=  1.238, Rscal=1.0260,
  R=  0.2721, x amp= 0.8566, test amp=  71400.14 (5.000%)
Amplitude scaling for hlxavg:  0.8566

cr532cltv.ren
  residual= 17.4922, phi= -1.709, z=-21.109, Rscal=1.0280,
  R=  0.2866, x amp= 0.8825, test amp=  81911.18 (5.000%)
Amplitude scaling for hlxavg:  0.8825

cr530ltv.ren
  residual= 24.1304, phi= -6.277, z= -7.645, Rscal=1.0130,
  R=  0.2700, x amp= 1.0188, test amp=  57836.57 (5.000%)
Amplitude scaling for hlxavg:  1.0188

cr530bltv.ren
  residual= 17.8881, phi=  8.449, z=-19.064, Rscal=1.0170,
  R=  0.2885, x amp= 0.8946, test amp=  81175.16 (5.000%)
```

```
Amplitude scaling for hlxavg:  0.8946

cr991ltv.ren
  residual= 15.9908, phi= 21.534, z=  5.336, Rscal=1.0430,
  R=  0.2259, x amp= 0.7451, test amp=  71278.84 (5.000%)
Amplitude scaling for hlxavg:  0.7451

cr5286ltv.ren
  residual= 17.6810, phi= -4.057, z=  7.729, Rscal=1.0250,
  R=  0.2699, x amp= 0.8425, test amp=  84429.06 (5.000%)
Amplitude scaling for hlxavg:  0.8425

cr5263bltv.avg
  residual= 15.5691, phi= -5.657, z= 22.150, Rscal=1.0060,
  R=  0.1911, x amp= 0.4671, test amp= 145013.77 (5.000%)
Amplitude scaling for hlxavg:  0.4671

cr5194ltv.ren
  residual= 22.2157, phi= -4.899, z= -2.830, Rscal=1.0130,
  R=  0.2308, x amp= 0.7224, test amp=  95925.06 (5.000%)
Amplitude scaling for hlxavg:  0.7224

cr1021ltv.ren
  residual= 15.5456, phi= -0.153, z=-13.695, Rscal=1.0350,
  R=  0.2496, x amp= 1.0702, test amp=  67624.03 (5.000%)
Amplitude scaling for hlxavg:  1.0702

cr1087ltv.ren
  residual= 15.5077, phi=  4.874, z= -2.355, Rscal=1.0580,
  R=  0.2389, x amp= 0.6512, test amp= 117298.81 (5.000%)
Amplitude scaling for hlxavg:  0.6512

cr1527ltv.ren
  residual= 21.8226, phi=-13.565, z= -2.533, Rscal=1.0430,
  R=  0.3419, x amp= 0.9408, test amp= 120235.45 (5.000%)
Amplitude scaling for hlxavg:  0.9408

 hlx2fld
  residual= 17.9507, phi=  0.014, z= -0.082, no. of points=  1713 out of   1753
 test amp= 135837.34 (1.000%)


--- cycle   2 ---
cr9901ltv.ren
  residual= 20.5489, phi=  4.421, z= 10.339, Rscal=1.0150,
  R=  0.2253, x amp= 1.2805, test amp=  37846.47 (5.000%)
Amplitude scaling for hlxavg:  1.2805

cr532ltv.ren
  residual= 18.9815, phi=  0.084, z=  1.101, Rscal=1.0250,
  R=  0.2324, x amp= 0.8316, test amp=  71400.14 (5.000%)
Amplitude scaling for hlxavg:  0.8316

cr532cltv.ren
  residual= 20.4937, phi= 10.868, z=  8.416, Rscal=1.0270,
  R=  0.2359, x amp= 0.8711, test amp=  81911.18 (5.000%)
Amplitude scaling for hlxavg:  0.8711

cr530ltv.ren
  residual= 25.9242, phi= -6.283, z= -7.708, Rscal=1.0130,
  R=  0.2349, x amp= 1.0059, test amp=  57836.57 (5.000%)
Amplitude scaling for hlxavg:  1.0059

cr530bltv.ren
  residual= 19.8847, phi=  8.431, z=-19.165, Rscal=1.0170,
  R=  0.2471, x amp= 0.8757, test amp=  81175.16 (5.000%)
Amplitude scaling for hlxavg:  0.8757

cr991ltv.ren
  residual= 19.9470, phi= 21.488, z=  5.103, Rscal=1.0460,
  R=  0.2007, x amp= 0.7397, test amp=  71278.84 (5.000%)
Amplitude scaling for hlxavg:  0.7397

cr5286ltv.ren
  residual= 20.5786, phi= -4.026, z=  7.574, Rscal=1.0250,
  R=  0.2334, x amp= 0.8324, test amp=  84429.06 (5.000%)
Amplitude scaling for hlxavg:  0.8324

cr5263bltv.avg
  residual= 20.3171, phi= -5.699, z= 22.038, Rscal=1.0060,
```

```
  R=  0.2080, x amp= 0.4818, test amp= 145013.77 (5.000%)
Amplitude scaling for hlxavg:  0.4818

cr5194ltv.ren
  residual= 26.9473, phi= -4.899, z= -2.831, Rscal=1.0130,
  R=  0.2400, x amp= 0.7001, test amp=  95925.06 (5.000%)
Amplitude scaling for hlxavg:  0.7001

cr1021ltv.ren
  residual= 21.6176, phi= -0.152, z=-13.764, Rscal=1.0350,
  R=  0.1864, x amp= 1.1242, test amp=  67624.03 (5.000%)
Amplitude scaling for hlxavg:  1.1242

cr1087ltv.ren
  residual= 20.4767, phi=  4.853, z= -2.525, Rscal=1.0590,
  R=  0.2055, x amp= 0.6890, test amp= 117298.81 (5.000%)
Amplitude scaling for hlxavg:  0.6890

cr1527ltv.ren
  residual= 25.6600, phi=-13.416, z= -2.544, Rscal=1.0440,
  R=  0.2719, x amp= 0.9133, test amp= 120235.45 (5.000%)
Amplitude scaling for hlxavg:  0.9133

 hlx2fld
  residual= 17.5719, phi= -0.004, z= -0.014, no. of points=  1713 out of   1757
 test amp= 135482.47 (1.000%)

               ..........................

--- cycle   5 ---
cr9901ltv.ren
  residual= 20.6699, phi=  4.427, z= 10.344, Rscal=1.0150,
  R=  0.2251, x amp= 1.2891, test amp=  37846.47 (5.000%)
Amplitude scaling for hlxavg:  1.2891

               ..........................

cr1527ltv.ren
  residual= 25.6157, phi=-13.371, z= -2.554, Rscal=1.0450,
  R=  0.2694, x amp= 0.9225, test amp= 120235.45 (5.000%)
Amplitude scaling for hlxavg:  0.9225

 hlx2fld
  residual= 17.5497, phi= -0.014, z= -0.024, no. of points=  1711 out of   1755
 test amp= 136553.78 (1.000%)
```

d. Real space averaging program

lglst provides the profile of the little g. The little g from the equatorial layer-line represents the mean radial density distribution (MRDD). The results are written on two files, 'lglst.gnu' and 'lglst.dat' for displaying the graph by gnuplot. To display the graph, type

```
# gnuplot

gnuplot> load 'lglst.gnu'     .
```

[Control file]

```
!title
cr25 vs cr26series lg = 0
!start and end radii in Angstrom
!200 370
180 370 1
! file name and optional scaling factor
cr25w12ltvnm9.lg1 -1.0
! layer-line number (and Bessel order)
0
../cr26w23e/cr26w23ltvnm9.lg1 -2.0
0
```

lgcor calculates the cross correlation between the MRDDs of the helical particles and provides the relative scaling factor and shift value along the radius.

[Control file]

```
! 1. radial shift (min, max, step), e.g. -20 -20 1
-30 20 1
! 2. scaling (min, max step), e.g.: 0.9 1.1 0.001
0.9 1.1 0.001
! 3. density scaling and floating for the 1st data
!                                          (test data)
1 0
! 4. density scaling and floating for the 2nd data
!                                          (ref. data)
1 0
! 5. cut densities for calculation of correlation, e.g. 0. 0.
0 0
! 6. data [radial density(1st) density(2nd)]
180.0    -0.0980    0.0000
181.0    -0.0920    0.0000
182.0    -0.0856    0.0000
183.0    -0.0795    0.0000
184.0    -0.0740    0.0000
185.0    -0.0684    0.0000
186.0    -0.0616    0.0000
187.0    -0.0525    0.0000
188.0    -0.0401    0.0000
189.0    -0.0243    0.0000
190.0    -0.0055    0.0000
191.0     0.0148    0.0000
192.0     0.0354    0.0000
193.0     0.0553    0.0000
```

```
194.0      0.0742     0.0000
195.0      0.0927     0.0000
196.0      0.1123     0.0000
197.0      0.1343     0.0000
198.0      0.1600     0.0000
199.0      0.1897     0.0000
200.0      0.2227    -0.1122
201.0      0.2577    -0.1054
          ..............
```

[Result]

```
LGCOR Vers.   1.10 :  --- Thu Jan 16 13:26:07 1997

Shift:  min -30.00 max  20.00 and step   1.00(    51)
Scale:  min 0.9000 max 1.1000 and step 0.0010(   201)
Test prod: 1.0000 and add:   0.00
Ref  prod: 1.0000 and add:   0.00
Test cutoff : 0.0000 and Ref cutoff: 0.0000
Max correlation coefficient: 0.9994 at shift  17.00
                                    and  scale 1.0250
 ( Shift after scale correctin:  12.61 )
```

hcellbox provides the control file, 'hcellbox.out' for boxmsk. Using this control file, the areas within one unit cell are masked off. This program needs the id file and the image file of the helical x sections. hcellboxclnd needs the image file of helical cylindrical sections instead of x sections.

imcorab calculates scaling parameters by using a linear total least-squares method (Numerical recipe in C, 2nd edition: 666-670). Output files ,'ab.gnu' and 'ab.dat', are made for displaying the graph of the result by gnuplot.

[Control file]

```
! ref. file
cr25a.hpz
! test file
vo6154x5r.msk
! mode(S: use standard deviation maps
```

```
!       M: use constant values for standard deviations)
M
! standard dev. files or constant values for standard dev.
100 50
```

imgcoraby calculates scaling parameters by using a linear least-squares method (assuming only y values have errors).

prodadd applies density scaling and floating.

rhofit calculates the relative positions of two molecules and provides the correlation coefficients between the molecules.

[Control file]

```
! title
cr25 x cr26
! envelope image file
cr25w12rnm9s.dns
! reference image file
cr25w12rnm9s.x
! sample image file
cr26w231tvnm9.x
! output log file
rhofit25.log
! input range of sample image [start, end, width] x 3
!                                          (in x, y, z)
0 36 37
0 52 53
0 42 43
! HB-matrix
111 0 0
0 159 0
0 0 129
! rotaion angle [theta, phi, chi]
5 5 5
! origin of reference molecular coordinate [xg, yg, zg]
0.5 0.5 0.5
! origin of sample molecular coordinate [xg, yg, zg]
0.5 0.5 0.5
!0.51328 0.50889 0.48847
! axis fix mode (0 : flex  or 1 : fix)
! [theta, phi, chi, x, y, z]
0 0 0 0 0 0
!max cycle
10
```

```
[Result]


cr25 x cr26
envelope  image file : cr25w12rnm9s.dns
reference image file : cr25w12rnm9s.x
sample    image file : cr26w23ltvnm9.x

     |   111.000000     0.000000     0.000000 |
HB = |     0.000000   159.000000     0.000000 |
     |     0.000000     0.000000   129.000000 |
THETA=   5.00000     PHI=   5.00000     CHI=   5.00000
ORIGIN OF REFRENCE  MOLECULAR COORDINATE SYSTEM =    0.50000   0.50000   0.50000
ORIGIN OF SAMPLE MOLECULAR COORDINATE SYSTEM =    0.50000   0.50000   0.50000
RANGE OF E.D. MAP FOR INTERPOLATION
 NSX,NNX,NX,  NSY,NNY,NY,  NSZ,NNZ,NZ =    0   36   37    0   52   53    0   42   43
IFIX =    0  0  0  0  0  0

Cycle : 1
THEATA, PHI, CHI =    5.00000    5.00000    5.00000
ORIGIN OF SAMPLE  MOLECULAR COORDINATE SYSTEM =    0.50000   0.50000   0.50000
    |     0.996223    -0.086822     0.000991 |
F = |     0.086827     0.996195    -0.007538 |
    |    -0.000333     0.007596     0.999971 |
TRANSLATION IN ANGSTROM
X =    0.00000 Y =    0.00000 Z =    0.00000
NO. OF GRID POINTS =       5135

RARAMTER SHIFT
     THETA=  10.67042     PHI= 175.05664     CHI= 357.13678
     TRANSLATION IN ANGSTROM  X=   0.04713   Y=  -0.26518   Z=   0.22518
NEW RARAMTER
     THETA=  15.67042     PHI= 180.05664     CHI=   2.13678
     TRANSLATION IN ANGSTROM  X=   0.04713   Y=  -0.26518  Z=   0.22518
STANDARD DEVIATION FOR EACH PARAMETER (sig)
     0.02520      0.58853      0.00350      0.12951      0.07533      0.10516
    |     0.999355    -0.035899    -0.000191 |
F = |     0.035899     0.999305     0.010071 |
    |    -0.000171    -0.010071     0.999949 |
RESIDUAL OF E.D.=  161.1072    R-VALUE OF E.D.=    0.1322
NO. OF POINTS =      5129
CORRELATION COEFFICIENT OF E.D.= 0.57567


For SKEW
Rotaion Matrix
     |     0.999355    -0.035899    -0.000191 |
F =  |     0.035899     0.999305     0.010071 |
     |    -0.000171    -0.010071     0.999949 |
Translation Coordinate (Fraction)
   0.49958   0.50167   0.49825

        ........................

Cycle : 6
THEATA, PHI, CHI =   45.64462   91.15958    1.02756
ORIGIN OF SAMPLE  MOLECULAR COORDINATE SYSTEM =    0.50701   0.49638   0.50283
    |     0.999839    -0.012539     0.012818 |
F = |     0.012536     0.999921     0.000340 |
    |    -0.012822    -0.000179     0.999918 |
TRANSLATION IN ANGSTROM
       X =    0.77790 Y =   -0.57489 Z =    0.36570
NO. OF GRID POINTS =       5135

RARAMTER SHIFT
     THETA=  63.08074     PHI=   2.63504     CHI= 359.53503
     TRANSLATION IN ANGSTROM  X=   0.16852   Y=  -0.04928   Z=  -0.29372
NEW RARAMTER
     THETA= 108.72536     PHI=  93.79462     CHI=   0.56259
     TRANSLATION IN ANGSTROM  X=   0.94642   Y=  -0.62417  Z=   0.07198
STANDARD DEVIATION FOR EACH PARAMETER (sig)
     0.19688      0.00824      0.00356      0.11573      0.06596      0.06023
    |     0.999952     0.003149     0.009280 |
F = |    -0.003155     0.999995     0.000601 |
    |    -0.009278    -0.000630     0.999957 |
RESIDUAL OF E.D.=  140.2770    R-VALUE OF E.D.=    0.1146
NO. OF POINTS =      5134
CORRELATION COEFFICIENT OF E.D.= 0.65242
```

```
For SKEW
Rotaion Matrix
     |     0.999952     0.003149     0.009280 |
F =  |    -0.003155     0.999995     0.000601 |
     |    -0.009278    -0.000630     0.999957 |
Translation Coordinate (Fraction)
   0.49147   0.50393   0.49944

Cycle : 7
THEATA, PHI, CHI =  108.72536   93.79462    0.56259
ORIGIN OF SAMPLE  MOLECULAR COORDINATE SYSTEM =    0.50853   0.49607   0.50056
     |     0.999952     0.003149     0.009280 |
F =  |    -0.003155     0.999995     0.000601 |
     |    -0.009278    -0.000630     0.999957 |
TRANSLATION IN ANGSTROM
       X =    0.94642 Y =   -0.62417 Z =    0.07198
NO. OF GRID POINTS =       5135

RARAMTER SHIFT
     THETA= 357.69620     PHI= 355.23230     CHI=   0.76603
     TRANSLATION IN ANGSTROM  X=  -0.00520   Y=  -0.00349   Z=  -0.01197
NEW RARAMTER
     THETA= 106.42157     PHI=  89.02692     CHI=   1.32863
     TRANSLATION IN ANGSTROM  X=   0.94122   Y=  -0.62766  Z=   0.06001
STANDARD DEVIATION FOR EACH PARAMETER (sig)
     0.33843      0.01372      0.00372      0.11313      0.06342      0.06016
     |     0.999731     0.006559     0.022237 |
F =  |    -0.006551     0.999978    -0.000451 |
     |    -0.022239     0.000305     0.999753 |
RESIDUAL OF E.D.=  137.2348    R-VALUE OF E.D.=    0.1120
NO. OF POINTS =      5135
CORRELATION COEFFICIENT OF E.D.= 0.66509


For SKEW
Rotaion Matrix
     |     0.999731     0.006559     0.022237 |
F =  |    -0.006551     0.999978    -0.000451 |
     |    -0.022239     0.000305     0.999753 |
Translation Coordinate (Fraction)
   0.49152   0.50395   0.49953

       ........................

Cycle : 10
THEATA, PHI, CHI =   90.25919   75.99908    2.28193
ORIGIN OF SAMPLE  MOLECULAR COORDINATE SYSTEM =    0.50015   0.49638   0.50153
     |     0.999253     0.000366     0.038633 |
F =  |     0.000006     0.999954    -0.009637 |
     |    -0.038634     0.009630     0.999207 |
TRANSLATION IN ANGSTROM
       X =    0.01647 Y =   -0.57568 Z =    0.19685
NO. OF GRID POINTS =       5135

RARAMTER SHIFT
     THETA=   1.03796     PHI= 358.63705     CHI= 359.13806
     TRANSLATION IN ANGSTROM  X=  -0.16648   Y=  -0.04016   Z=  -0.16349
NEW RARAMTER
     THETA=  91.29715     PHI=  74.63614     CHI=   1.41998
     TRANSLATION IN ANGSTROM  X=  -0.15000   Y=  -0.61584  Z=   0.03336
STANDARD DEVIATION FOR EACH PARAMETER (sig)
     0.08055      0.04433      0.00187      0.11390      0.06600      0.08791
     |     0.999714     0.000639     0.023887 |
F =  |    -0.000483     0.999978    -0.006571 |
     |    -0.023891     0.006557     0.999693 |
RESIDUAL OF E.D.=  136.9654    R-VALUE OF E.D.=    0.1118
NO. OF POINTS =      5135
CORRELATION COEFFICIENT OF E.D.= 0.66171


For SKEW
Rotaion Matrix
     |     0.999714     0.000639     0.023887 |
F =  |    -0.000483     0.999978    -0.006571 |
     |    -0.023891     0.006557     0.999693 |
Translation Coordinate (Fraction)
   0.50135   0.50387   0.49974
```

In this case, the maximum coefficient value (0.66509) was obtained at the 7th cycle. Therefore, for adjusting the positions of the molecules, the result at the 6th cylce can be used.

[The result at the 6th cycle]

```
For SKEW
Rotaion Matrix
     |     0.999952     0.003149     0.009280 |
F =  |    -0.003155     0.999995     0.000601 |
     |    -0.009278    -0.000630     0.999957 |
Translation Coordinate (Fraction)
   0.49147   0.50393   0.49944
```

skew adjusts the positions of two molecules.

[Control file]

```
!title
cr25
!input image file
cr25w12rnm9s.x
!output image file
cr25skew9.x
!log file
skew.log
!mode , section-axis [X, Y, Z], center location mode
5 Z G       ! The Tsukihara's lab. Manuals has information for other modes.
!xg, yg, zg (center of molecular gravity)
!Translation Coordinate (Fraction)
0.49147   0.50393   0.49944
!xm, ym, zm (output mesh size in angstrom)
3. 3. 3.
!is, ie, js, je, ks, ke (output size)
-18 18 -26 26 -21 21
!F-MATRIX(mode = 5)
 0.999952     0.003149     0.009280
-0.003155     0.999995     0.000601
-0.009278    -0.000630     0.999957
```

avgimrot avearges two image files.

[Control file]

```
! input file 1
cr25skew9.x
! input file 2
```

```
cr27skew9.x
! output file
cr25x7.x
! density scaling and floating for input file 1
1.2 0.
! density scaling and floating for input file 2
1. 0.
! rotation (degree) around z axis, x shift and y shift
! for input file 2
0. 0. 0.
```

e. The programs for bringing image file back to the Fourier space

hcelltwo provides the control file, 'upper.out' and 'lower.out' for boxmsk. Using these control file, helical x sections can be divide into two areas, the upper side and the lower side by the line passing through center of the dimer.

boxmsk2fld applies 2-fold symmetry to the control file for boxmsk.

hx2lg calculates little g file from the helical x sections.

hclnd2lg calculates little g file from the helical cylndrical sections.

[Control file]

```
! 1. new title for output little g
lg from cr2567all
! 2. input file (hx section)
cr2567all.x
! 3. output file (little g file)
! and template file (little g file)
cr2567all.lg1 cr26w23ltvm.lg1
! 4. repeat distance(Angstrom) and # of ll
3883.225 0
! 5. Fourier scaling and back scaling (default 1)
-4000 1
! 6. id file
cr5247l.id
! 7. verbose
```

2

hlxfbr calculates layer-lines file from little g file.

[Control file]

```
! 1. title for Big G
lg from cr2567all
! 2. Input file name (.lg)
cr2567all.lg1
! 3. Number of layer-lines
! (for limiting calculation; 0 for all) and
! twofold flag (1 if twofold forced)
0 0
! 4. output file
cr2567all.cor
! 5. .edt file (use the range of the radius for Big G)
cr26w23ltvm.edt
```

B-2. Programs for t-test

imgtmap calculates the Student's t-map. Six files of 4 different kinds are made. All in MRC format: a t-map, a probability file, two averaged file (the first and the second groups) and two standard deviation files (the first and second groups).

[Control file]

```
! 1st group of images
! the number of images
7
! image file name
cr1527x8rskew.msk
! multiplication factor and base
0.53177 -100.63210 ! scaling vs vo25.dns
!
cr5286x7rskew.hpz
0.38601 -65.33840
!
cr532x3crskew.hpz
0.41591 -44.01724
!
cr530x1rskew.hpz
0.56949 -120.64944
!
cr530x1brskew.hpz
0.43286 -53.18667
!
cr991x2rskew.hpz
0.34362 -54.69618
!
cr9901x2rskew.hpz
0.44532 -38.13229
! 2nd group of images
6
vo6041x2r.hpz
1.14127 -23.60413
!
vo9450x1r.hpz
1.33186 8.69361
!
vo9689x0r.msk
1.59572 11.32460
!
vo6154x5rskew.hpz
0.25842 -10.49431
!
vo2071x2rskew.hpz
0.59570 -91.07579
!
vo7639x0arskew.hpz
0.35691 -75.79068
```

```
! output files, t- and probability files
tmap25.hpz prob25.hpz
! two averaged files and two standard deviation files
cr25mean.hpz vo25mean.hpz cr25std.hpz vo25std.hpz
```

imgmsk masks off a part of an image file according to the reference mask file.