

University of Groningen

## Electron Microscopy

Boekema, Egbert J.; Rögner, Matthias

*Published in:*  
EPRINTS-BOOK-TITLE

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
1996

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*  
Boekema, E. J., & Rögner, M. (1996). Electron Microscopy. In *EPRINTS-BOOK-TITLE*

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

# Chapter 20

## Electron Microscopy

Egbert J. Boekema\* and Matthias Rögner<sup>1</sup>

*Bioson Research Institute, Biophysical Chemistry, University of Groningen, Nijenborgh 4,  
9747 AG Groningen, The Netherlands*

<sup>1</sup>*Institute of Botany, University of Münster, Schlossgarten 3, D-48149 Münster, Germany*

Summary	325
I. Principles	326
A. Introduction	326
B. The Electron Microscope	326
C. Specimen Preparation	327
D. Averaging Methods	328
E. Possibilities of EM in Terms of Resolution and in Relation to Object Size and Specimen Preparation	329
II. Periodic Averaging	330
A. Two-Dimensional Crystallization	330
B. Periodic Averaging by Fourier Methods	330
C. Averaging of Photosystem I Crystals	330
D. High-Resolution EM	331
E. Light-Harvesting Complex II	331
III. Single Particle Averaging	332
A. Method	332
1. Alignment by Correction Methods	332
2. Multivariate Statistical Analysis	332
3. Classification Step	333
B. Analysis of Photosystem I Trimers from Cyanobacteria	333
C. Analysis of Photosystem II Dimers from Cyanobacteria	333
IV. Concluding Remarks	335
Acknowledgements	335
References	335

### Summary

Electron microscopy (EM) in combination with image analysis is a powerful technique to study protein structure at low- and high resolution. Since electron micrographs of biological objects are very noisy, substantial improvement of image quality can be obtained by averaging of individual projections. Averaging procedures can be divided into crystallographic and non-crystallographic methods and both will be described. Crystallographic averaging, based on two-dimensional crystals of rather small proteins, has the potential of solving a structure to atomic resolution just as the more common techniques of X-ray diffraction and NMR. Single particle analysis is an alternative method for large proteins, viruses and all non-crystallizable proteins. It is a fast method to reveal the low-resolution structure with details in the range of maximally 10–15 Å. Results of EM on Light-harvesting complex II (LHC-II) and Photosystem I will be presented as examples for the crystallographic averaging. Trimeric Photosystem I complexes and dimeric Photosystem II complexes will be discussed as examples for the potential of single-particle averaging.

\*Correspondence: Fax: 31-50-3634800; E-mail: boekema@chem.rug.nl

## I. Principles

### A. Introduction

Direct information about the three-dimensional (3D) structure of a protein is essential for understanding its functional organization. At present electron microscopy (EM) is a widely applied technique for studying the structure of proteins and membranes; however, it is still less common than X-ray diffraction, where solving the 3D structure of proteins becomes almost routine, once suitable crystals have been obtained. On the other hand, X-ray diffraction has two disadvantages in comparison to EM. First, X-rays cannot be focussed and only diffraction patterns are obtained, whereas EM results in direct information in the form of images. Second, the interaction of X-rays with material is a factor of about 10,000 weaker than that of electrons. This makes EM a useful technique as images of single protein molecules or one-layer thick crystals can be obtained, whereas X-ray diffraction needs much thicker specimens.

In this chapter on EM, the first sections will briefly introduce: the electron microscope with some instrumental aspects (I.B); specimen preparation (I.C); image analysis averaging techniques (I.D) and finally the possible resolution of EM with respect to object size and specimen preparation (I.E). Sections II.A and III.A will focus on image analysis of periodic and non-periodic objects including some examples.

For more details concerning theory of the electron microscope, techniques for recording the signal and image analysis, we refer to the book by Hawkes and Valdrè (1990), which was written for the field of protein structure determination.

### B. The Electron Microscope

The resolution of a light microscope, which is about 2000 Å, is mainly limited by the wavelength of the light. Improving the resolution of a microscope is only possible by exploiting waves with a much shorter wavelength. According to the well known formula of de Broglie,  $\lambda = h/mv$ , accelerated particles such as electrons also have a wave character. At an acceleration voltage of 100,000

V the wavelength of the electron beam is 0.037 Å. This is certainly sufficient to enable microscopy at atomic resolution. Shortly after de Broglie had described the wave character of particles, it was discovered, and put into practice, that electrons can be focussed by electric and magnetic fields with axial symmetry. Based on these principles, Ernst Ruska and Max Knoll constructed the first simple electron microscope in 1931, a tube under vacuum with an electron source and several lenses. Improvements in the following years enhanced the resolution to 100 Å, already much better than the resolution of a light microscope. Since the early days the electron microscope has been gradually but substantially improved to an instrument which can now routinely achieve a resolution of about 2 Å. This resolution is limited by lens geometries and reflects compromises between several optical parameters, such as minimizing the spherical aberration, which is a kind of lens error. Overall, much further improvement in resolution cannot be expected, but 2 Å resolution is sufficient to solve a protein structure at the atomic level.

To minimize the lens aberrations, the lenses in the electron microscope have a small opening: the holes in lens apertures are less than 0.1 mm in diameter. This results in a large “depth of field” and “depth of focus” at the object plane and image plane, respectively. As a consequence, EM gives two-dimensional (2D) projections in which the upper- and lower side of a thin object (up to a few 1000 Å) are seen superimposed with the same “sharpness”. As a result, a simple focussing on selected levels in an object, as is possible with a light microscope, cannot be done with the electron microscope. To get information about the 3D shape of a protein, the specimen must be tilted in the microscope and the various projections must be compared or combined into a 3D reconstruction.

Imaging by EM of, for example, a thin metal foil or a gold cluster will easily provide projections with atomic resolution, but obtaining structures of proteins at high resolution is much harder work. Why? The contrast in the electron microscope is caused by scattering. Electrons are deflected at atomic nuclei through large angles and by other electrons through small angles. Since the

scattering by nuclei is proportional to the atomic number, biological material, containing only the lighter elements, will give images with very low contrast. Besides, radiation damage by an electron beam easily destroys biological samples. Radiation damage cannot be avoided, but only minimized by cooling the specimen to liquid nitrogen or liquid helium temperature and by minimizing the electron dose. As a result, electron micrographs are noisy and objects are hardly visible. Therefore, image analysis techniques have been developed to improve the signal recorded in the noisy EM pictures.

### C. Specimen Preparation

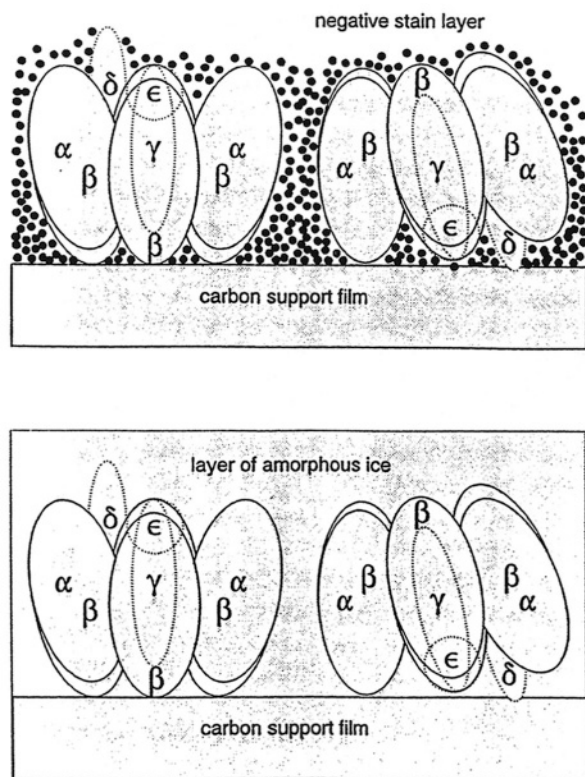
Since modern electron microscopes have enough resolving power for structural studies of macromolecules, other factors than instrumental ones are of importance. The specimen preparation method is one of these factors and it strongly determines the ultimate resolution that can be achieved. In the negative staining technique the contrast is enhanced by embedding proteins or protein crystals in a heavy metal salt solution. Upon drying, the metal salts fill spacings and cavities around the molecules, but do not penetrate the protein interior. As a result, negatively stained specimens show protein envelopes with good contrast, but with a resolution that usually does not exceed 15 Å, due to the graininess of the contrasting agent. Because of its simplicity, negative staining has been widely applied. Single particles as well as crystals of photosynthetic membrane proteins, such as Photosystem I and II, have been successfully prepared by negative staining (Boekema, 1991; Boekema et al., 1994).

As an alternative for negative staining, Unwin and Henderson (1975) pioneered the embedding of proteins in other media, such as glucose. It was demonstrated for crystals that images with a resolution of better than 10 Å could be achieved. This opened the way to high-resolution EM for biological macromolecules. A second important discovery in this context was the reduction of radiation damage, leading to a better signal. Specimen holders were developed that could be cooled down to liquid nitrogen- or liquid helium temperature. It was found for organic- and pro-

tein crystals that at the temperature of liquid nitrogen radiation damage is roughly reduced by a factor of 3–5 and by a factor of 10–20 at liquid helium temperature (Zemlin, 1992). Presently, the instrumental difficulties of practical cryo-EM have been mostly solved and the performance of the newest microscopes at low temperature is almost as good as at room temperature.

Cryo-EM was further stimulated by the discovery of vitrification of protein solutions (Adrian et al., 1984). By rapidly cooling a protein solution, the formation of ice crystals can be avoided and proteins embedded in a thin layer of amorphous ice are obtained. Contrast is caused by the difference in density between amorphous ice (0.93 g/cm<sup>3</sup>) and protein (1.3–1.36 g/cm<sup>3</sup>) and is rather low in comparison to negative staining. However, there are several advantages of cryo-EM of vitrified specimens: Specimen flattening and other drying artifacts are circumvented. Moreover, cryo-images better reflect the true density of a protein, because the contrast directly originates from scattering by the protein rather than from the surrounding stain (Fig. 1). Also, negative stain interaction with the protein is often quite complex. In thinner stain layers, the upper part of the protein could easily be less well embedded in the stain layer, as pointed out in Fig. 1. This means that the contribution of the upper- and lower half of a protein in the final recorded image do not have the same weighting.

Other techniques are less important for high resolution structural work. Freeze-fracture techniques have been widely applied in research on photosynthesis. Cell or membranes are rapidly frozen, cleaved and replicated. The replicas give useful information about the overall size and distribution of the complexes embedded in these membranes (Staehelin, 1988). But the resolution is rather limited. Only particle diameters and the overall shape of membrane protein can be obtained. The main value of these techniques lies in the imaging of the *in vivo* situation of the membranes. It can reveal crystalline packing of photosystems and sometimes the multimeric state of large protein complexes. Evidence for the existence of a dimeric organization of PS II *in vivo* was obtained by freeze-fracturing (Mörschel and Schatz, 1987).



**Fig. 1.** Specimen preparation for EM. Two common techniques are schematically drawn. The example is the soluble  $F_1$  part of the ATP synthase complex. (Top) Negative staining: a drop of a protein solution is disposed for 2 minutes on a copper grid coated with a thin carbon support film. Excess of material is blotted off with filter paper and immediately replaced by a drop of a heavy metal salt. After about two minutes, excess of the stain solution is removed and the grid is air-dried. (Bottom) Cryo-EM: a drop of a protein solution is also disposed on a grid and excess of material is blotted off. Immediately after blotting, the grid is plunged into liquid ethane and frozen at about  $-180^\circ\text{C}$ . From the liquid ethane, the grid is transferred to a liquid nitrogen cooled EM specimen holder which is transferred into the microscope.

#### D. Averaging Methods

In EM image analysis, improving the signal of an object recorded in a noisy electron micrograph is performed by averaging. By adding hundreds or, if possible, thousands of projections the signal improves substantially and trustworthy electron density maps are obtained. There are two general methods for averaging of 2D projections, depending on the object. One method is based on

filtering images of periodic objects, which are usually 2D crystals. The other one deals with single-particle projections.

Periodic averaging takes advantage of the fact that in crystals, protein molecules are arranged in a regular packing. This means that neighbor-to-neighbor distances have a fixed value. In other words: the precise position of the molecules can be easily determined, even if the crystal is recorded with a low electron dose to prevent radiation damage and the molecules can be barely seen. As a result, higher resolution can be obtained. If 2D crystals with a diameter of at least several  $\mu\text{m}$  can be grown, EM can be performed under cryo-EM conditions at high resolution. For some small membrane proteins with a mass of 20–40 kDa, averaging over very large areas resulted in projection structures with a resolution better than 5 Å. For bacteriorhodopsin, the three-dimensional structure could be determined entirely from EM data by fitting the amino acid chain into the electron density map (Henderson et al., 1990). Another example is in the field of photosynthesis, where a second high resolution structure determination, that of the light-harvesting complex II (LHC-II) from pea was recently completed (Kühlbrandt et al., 1994). The crystallographic method, which is based on Fourier methods, is further described in section II, where results on LHC-II will be discussed in more detail.

Projections of single particles can be averaged after they have been brought into equivalent positions by shifting them rotationally and translationally. This a-periodic averaging technique or single particle analysis is able to reveal the predominant projections of protein molecules (Frank et al., 1988). The fact that crystallization of the protein is not required is an advantage of this method. A disadvantage is that the maximal possible resolution by single-particle analysis is restricted to about 10–25 Å. This limit is set by the signal-to-noise ratio, which is related to the size of the object and has a relative low value for small objects. The resolution is also limited by the fact that the particles are not fixed in a definite position, as in a crystal. Small tilts from a common predominant position cause slight differences between similar-looking projections, resulting in an ensemble of projections that are all

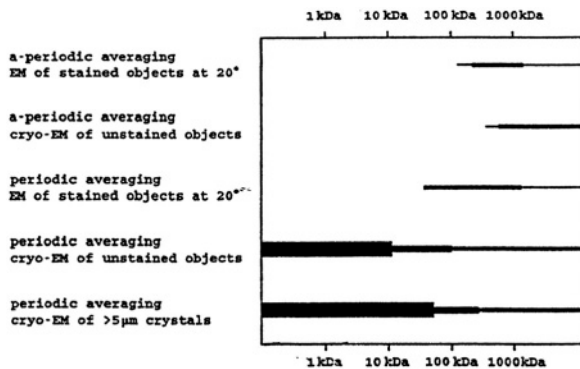


Fig. 2. An overview of EM structure determination approaches for proteins of various size. The thickness of the lines symbolizes the resolution that can be maximally attained under the given conditions and in relation to the mass. In decreasing thickness, the lines stand for: better than 5 Å; 5–10 Å; 10–25 Å and worse than 25 Å.

slightly different. Averaging all of them would give a sum in which the finest details would be blurred out.

### E. Possibilities of EM in Terms of Resolution and in Relation to Object Size and Specimen Preparation

With the present state of the art in EM, as described in the previous sections, we can give an overview of the possibilities of EM in the field of proteins. Fig. 2 describes the potential of 5 types of EM approaches in the field of proteins. The thickness of the lines in Fig. 2 indicates the suitability and the attainable resolution in relation to the molecular mass of a protein.

Approach 1. Single-particle averaging of negatively stained specimens is able to resolve the structure up to 15 Å in favorable cases. This resolution is sufficient for the localization of subunit positions in projections. Examples will be given for Photosystem I and II.

Approach 2. If the negative staining method is replaced by cryo-EM of vitrified solutions, single particle averaging can be applied to objects with a mass of at least a few hundred kDa. From smaller proteins, especially those with a mass of 100 kDa or lower, the projections from single molecules as recorded by EM are too noisy for accurate averaging. Approach 2 works better than ap-

proach 1 for larger objects, for reasons mentioned before, such as removing the flattening upon drying. Another reason is the lower contrast of cryo-EM, which becomes relatively disadvantageous for small molecules, as we will briefly explain. The contrast in biological material is largely caused by the scattering of the electron cloud of the C, N and O elements. This contrast is called phase contrast. The contrast originating from interaction with atomic nuclei, the scattering contrast, is relatively unimportant. Enhancement of phase contrast is possible by a stronger defocussing of the objective lens. But this coincides with a degradation and loss of fine details in the image. The smaller the object, the larger the defocussing needed to see the object at all.

Approach 3. Periodic averaging of stained objects. The advantage of a periodic object is its regular packing. This enables an accurate determination of the positions of the molecules and thus an accurate averaging of projections. For crystals of about 1 µm in size, a resolution of about 15–25 Å is usual. Small proteins (10–50 kDa) are advantageous, because they allow averaging over a larger number of molecules per area.

Approach 4. Cryo-EM of unstained, periodic objects. As for approach 2, the contrast comes largely from phase contrast. But crystals can be considered as a “big object” and can be recorded with small defocus values. Although the signal of one individual molecule is almost buried in the noise, it will still be interpretable, because it is averaged periodically. Small crystals of organic molecules and small proteins of a few µm in size can be analyzed to high resolution. For such small macromolecules, EM provides the same quality as X-ray diffraction, although the image analysis is not as straightforward.

Approach 5. Based on perfect, large 2D crystals atomic resolution is possible for proteins up to at least 50 kDa. For larger proteins, the resolution will gradually decrease. For proteins with a mass between 300 and 600 kDa, like Photosystem I and II and ATP synthase, solving the structure at high-resolution by EM would be difficult. The preferential size of 2D crystals of such objects is at least 10–20 µm. Such large crystals are difficult to grow and have not yet been obtained.

## II. Periodic Averaging

### A. Two-Dimensional Crystallization

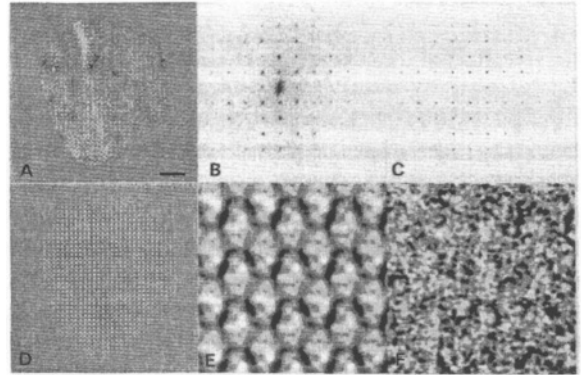
As indicated in Fig. 2, small proteins should be crystallized into 2D crystals, to obtain the best structural information. The best 2D crystals produced *in vitro* have been grown from detergent-solubilized, purified material. Starting from highly purified protein, crystallization conditions can be controlled more easily and results are more reproducible (Kühlbrandt, 1992; Jap et al., 1992; Engel et al., 1992). Reconstitution of membrane proteins into lipid bilayers by detergent removal is currently the most universal method of 2D crystallization. A suspension of a lipid-detergent mixture is usually added to a detergent-solubilized protein preparation. Crystallization of the protein into sheets or vesicle crystals is then induced by removal of detergent by means of dialysis or absorption.

### B. Periodic Averaging by Fourier Methods

Usually, images of 2D crystals are recorded on electron micrographs. Fourier analysis has been proven to be very valuable in the processing of micrographs. The Fourier transform is a frequency decomposition in reciprocal space. A 2D Fourier transform calculated from an image gives a 2D diffraction pattern. If the image shows a good 2D crystal, its transform will show a pattern with many spots laying on a regular pattern, as in X-ray diffraction patterns. These spots represent the structure factor amplitudes and have an amplitude (peak height) and a phase. But the noise in the images will also show up in the diffraction pattern. An illustration of the Fourier techniques will be given in the next section; it shows an effective way to get rid of the noise.

### C. Averaging of Photosystem I Crystals

At medium resolution no complicated strategy is necessary since correcting for image aberrations is only necessary for high-resolution EM (see II.D). We will show by a simple example the basics of periodic averaging. Monomeric Photosystem I (PS I) has been crystallized into two-dimensional arrays from the cyanobacterium *Synechococcus*



**Fig. 3.** Averaging of negatively stained PS I crystals by Fourier methods. (A) A part of a digitized two-dimensional crystal containing about 1500 monomers (the crystal was prepared and recorded as published in Böttcher et al., 1992); (B) a 2D Fourier-transform of the image of A showing peaks on a rectangular lattice. The scale is two times larger than in image A; (C) a mask in Fourier space with holes on the positions of the Fourier peaks. The mask thresholds the noise between the peaks; (D) a filtered image resulting from a reverse Fourier transform of the masked image of C. Darker and brighter rows of monomers run in the vertical direction. The rows contain up- and down oriented monomers and the difference in overall contrast between the rows originate from a difference in stain embedding; (E) a central part of image D, on a 12 times larger scale, shows alternating rows of up- and down oriented PS I monomers. Staining pattern of the rows was equalized by further averaging; (F) a part of image A, on a 12 times larger scale; a comparison with image E shows the considerable enhancement of the signal. The bar in (A) represents 1000 Å.

*elongatus* (Böttcher et al., 1992) by removal of detergent (Fig. 3A). In a 2D Fourier transform of the PS I crystal (Fig. 3B), we see spots at regular distances. They tell us about the lattice parameters and about resolution. In this case, the lattice of the crystal is rectangular.

Based on the Fourier transform of the image, a filtering is performed. Computationally, a mask is constructed that is superimposed on the Fourier transform (Fig. 3C). The mask has holes that neatly surround the peaks, which contain the crystalline information. The mask shields the space between the peaks, which represents noise in the crystal image. With a reverse Fourier-transformation (Fig. 3D) a real image is generated again, in which much of the noise is removed. This procedure is called Fourier-peak filtering and the comparison between Figs. 3E and F illus-

trates the considerable improvement in signal-to-noise ratio.

Further improvement in filtering is also possible. As crystals never have a perfect lattice, a small bending in the plane results in molecules being slightly displaced from their ideal position. Corrections can be made by “cutting” the crystal into pieces containing one or several molecular projections and shifting them into their ideal position. This is done by correlation methods, which have also been used in single particle averaging methods (Frank, 1982). An application of this method is the analysis of PS I crystals (Böttcher et al., 1992).

#### D. High-Resolution EM

While a low-resolution structure by single particle averaging can be obtained within weeks, obtaining a high-resolution structure from EM data may take years. One of the limits of a structure determination at high resolution is merely the production of well-ordered, large crystals. Once this prerequisite is fulfilled, a resolution better than 10 Å is possible. However, recording the highest quality images and extensive processing are also time consuming. The signal recorded by EM suffers from aberrations, which increase in severity as the required resolution becomes higher (Henderson et al., 1986). Images providing structural information beyond a resolution of about 5 Å need extensive processing, which is carried out in Fourier space on the raw image amplitudes and phases. The phases are corrected for the effects of the contrast transfer function, beam tilt and phase origin (Henderson et al., 1986). A further treatment and discussion of these image distortions and their analysis and correction is beyond the scope of this contribution. For the interested reader we refer again to the book of Hawkes and Valdrè (1990), as a useful source of further information, and to the paper of Henderson et al. (1990).

The microscope can be configured in two different ways: a) image mode and b) diffraction mode, i.e. recording the image formed in the back-focal plane of the objective lens. Electron diffraction (ED) patterns of crystals contain many spots, as in X-ray diffraction patterns and also represent the structure factors *without* the phase

information that is present in calculated Fourier transforms. ED patterns usually have much stronger spots than those from calculated transforms. The reason possibly lies in the specimen movement during image recording. Fourier transformation and diffraction in an optical system have the property that they are “translation invariant”. This means that movements during the recording of an ED pattern are less dramatic than during the recording of a real image, because they do not result in blurring the ED. Therefore spots in ED are much stronger and this makes recording of only ED for high resolution work tempting. But then a “phase problem”, as in X-ray diffraction, is created and phases need to be generated. Isomorphous replacement, as used in X-ray diffraction, is not a useful method for phasing the ED data from 2D protein crystals because the scattering contrast of heavy atoms is low for electrons and the noise level in the patterns is relatively high. To compute a protein map at high-resolution by Fourier methods, images, are recorded as well. They are necessary to extract the phase from each of the spots. The phases (from the images) and amplitudes (from ED) are finally combined and corrected for some image errors, briefly mentioned in section I.D. This is called a “Fourier synthesis”. An example of the Fourier techniques for high-resolution structure determination will be given in the next section.

#### E. Light-Harvesting Complex II

The light-harvesting chlorophyll *a/b* protein complex associated with photosystem II (LHC-II) harvests light energy and is capable of passing it to photosystem II. Detergent-solubilized, purified LHC-II forms large, highly ordered 2D crystals, which are ideal objects for recording high-resolution EM. Based on the averaging methods developed by Henderson et al. (1990), Kühlbrandt et al. (1994) were able to calculate a 3D structure at 3.4 Å resolution. Images were recorded with an electron microscope capable of resolving 2 Å object features at a specimen temperature of 4.2 K. Tilted and untilted images were collected for a 3D reconstruction. Since it is difficult to record images at tilt angles above 60°, there was a region in the Fourier space where amplitudes and phases could not be measured. This caused a reduction



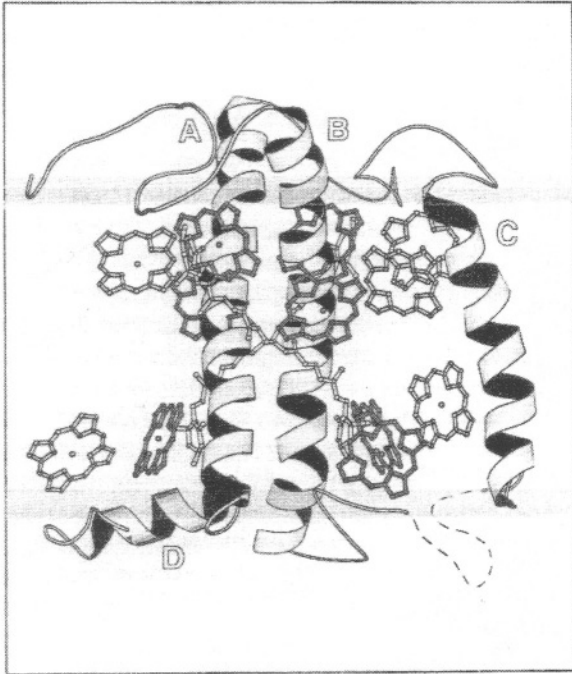


Fig. 4. Atomic model for the LHC-II structure seen in side-view position. Gray bands indicate the approximate position of the lipid bilayer. Helices are labelled A to D. (reproduced from Kühlbrandt et al. (1994) with permission).

in the resolution in the direction perpendicular to the membrane plane to between 4.6 and 4.7 Å. But due to the high quality of the phases it was possible to trace the polypeptide chain in the 3D electron density map. About 80% of the amino acids could be fitted, as well as the tetrapyrrole rings of 12 chlorophylls and two carotenoids (Fig. 4).

LHC-II is the second protein solved by EM to atomic resolution. No doubt, the work on LHC-II is a major step forward both for EM and for photosynthesis. To stress the similarities with X-ray crystallography, the term electron crystallography has been introduced for high-resolution protein determination by EM.

### III. Single-Particle Averaging

#### A. Method

Isolated proteins prepared on a carbon support film exhibit a full range of rotational and translational orientations in the plane of the support film. As a consequence, the projections of the proteins have a random orientation within this plane and computer averaging of such projections can only be achieved after an alignment procedure. Also, proteins often are attached in various ways to the support film (Fig. 1) and this results in a further variety in the obtained projections. To separate the various predominant projections, multivariate statistical analysis together with automated classification was introduced (Van Heel and Frank, 1981). The alignment procedure, multivariate statistical analysis and classification form the main steps in the single particle averaging procedure (Frank et al, 1988). These steps will be briefly discussed.

#### 1. Alignment by Correlation Methods

The first step in comparing images of projections of biomolecules is to bring them into register in the plane: the alignment process. A reference image is taken and each image in the data set is compared with this reference; rotational and translational cross-correlation functions are computed to determine the best angular and translational shifts to bring the image into a position most similar to the reference image. The alignment procedure is an iterative process because sums of projections align the data set better than does a noisy projection of just one molecule. In practice, several references are used and results are combined to circumvent a possibly wrong choice of the first reference.

#### 2. Multivariate Statistical Analysis

Once a data set of typically hundreds of projections has been aligned, they can be compared in some numerical way. In particular, correspondence analysis, a special form of principal component analysis, is used to extract relevant feature information from the data set (Van Heel and Harauz, 1988). Each image of  $\times$  pixels can be

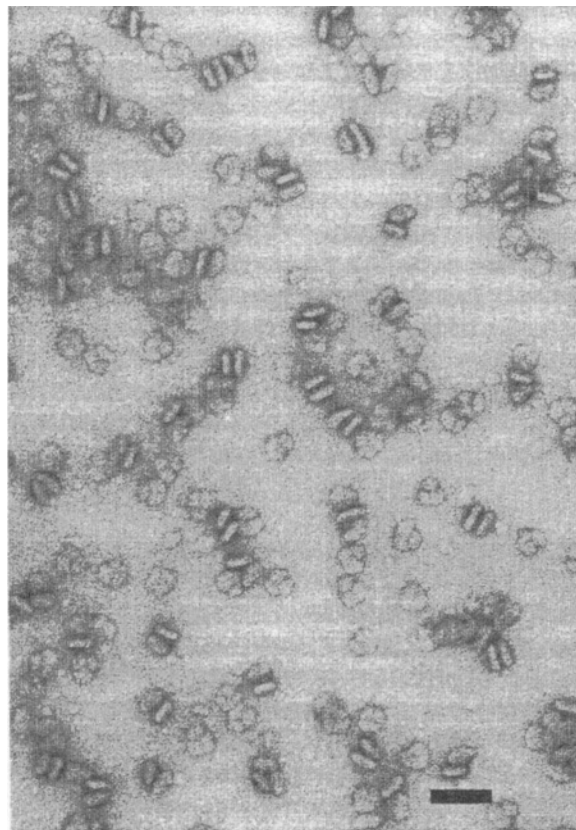
represented as a point in an  $\times$  dimensional “hyper”space, and the entire set of images forms a cloud in this space. Correspondence analysis determines a new, rotated-coordinate system in which the first axis represents the direction of the greatest inter-image variance, the second axis represents the direction of the largest remaining inter-image variance, and so on. The cloud of images can now be described with respect to this new coordinate system. By describing each image with respect to only the first few components, the images can be considered as points in a much smaller (than  $n \times n$ ) dimensional space. In this way a very large reduction in the amount of data to be analyzed is achieved.

### 3. Classification Step

After the data compression by correspondence analysis, the grouping together of those images that are most similar is achieved by automatic classification schemes. Output of the classification are “classes” of groups of projections that are most similar. In the classification process, projections are shifted between the classes to optimize the variance between the classes and to minimize the variance of the members belonging to the classes. The number of classes to be chosen is somewhat arbitrary, but usually a number of 6–12 are chosen. The differences between classes may represent real structural features (as will be illustrated by two examples) or are merely noise-related when there is not more than one type of projection present in the data set.

#### B. Analysis of Photosystem I Trimers from Cyanobacteria

A first practical example of single particle analysis concerns Photosystem I. In cyanobacteria, such as the thermophilic *Synechococcus elongatus* and the mesophilic *Synechocystis* PCC 6803, PS I is arranged as a trimeric complex with a diameter of about 200 Å in the membrane (Boekema et al., 1989; Kruip et al., 1993). From electron micrographs of PS I from *Synechocystis* PCC 6803 (Kruip et al., 1993), top view projections (Fig. 5) were extracted, aligned, treated by correspondence analysis and classified. The analysis resulted in the separation of the top view projec-



**Fig. 5.** Electron microscopy of PS I trimeric particles from the cyanobacterium *Synechocystis* PCC 6803, negatively stained with uranyl acetate. The bar represents 500 Å.

tions into two types, which are mirror-related (Fig. 6). These two types are generated because particles are asymmetric and can be attached to the carbon support film in two ways (upside-up and upside-down or “flip” and “flop”). It should be noted that it is not easy to judge by eye whether the PS I trimer projections (Fig. 5) belong to the flip or flop-type, but after a 3-fold averaging most of them (but not all) can be classified by eye.

#### C. Analysis of Photosystem II Dimers from Cyanobacteria

Photosystem II is another membrane protein with a size large enough to enable single particle analysis. Dimeric particles were purified from the *Synechococcus elongatus* (Rögner et al., 1987; Dekker

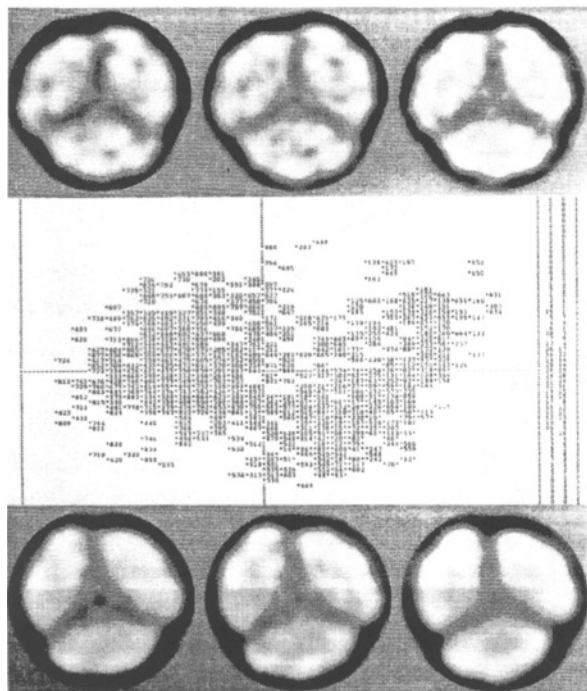


Fig. 6. Single particle averaging of *Synechocystis* PCC 6803 trimeric PS I projections. In the center is a correspondence analysis map of factorial coordinates 1 versus 2. Each of the 898 images is represented by an asterisk (plus its corresponding number) and the entire population forms a cloud. There are two subpopulations with centers in the top left and lower right half of the map. Sums of PS I particles in the "flip" type projection and sum of PS I particles in the "flop" type projection, which represent the two subpopulations in the map, are shown (modified from Kruijff et al., 1993).

et al., 1988). We visualized these particles in the presence of the detergent dodecyl maltoside by negative staining with uranyl acetate (Boekema et al., 1994). Averaged top views of dimeric PS II show that the dimers are built up from two monomers which are arranged in an anti-parallel way (Fig. 7A). In contrast to the example of PS I top view projections, there was only one predominant top view present in the data set. The dimer has dimensions of  $120 \times 155 \text{ \AA}$  in the membrane (corrected for detergent). The side views, however, show more variations. Original images clearly show protrusions in the side-view position (see Fig. 9 in Boekema et al., 1994). Often two protrusions can be seen on one dimer, but sometimes only one is visible. The protrusions

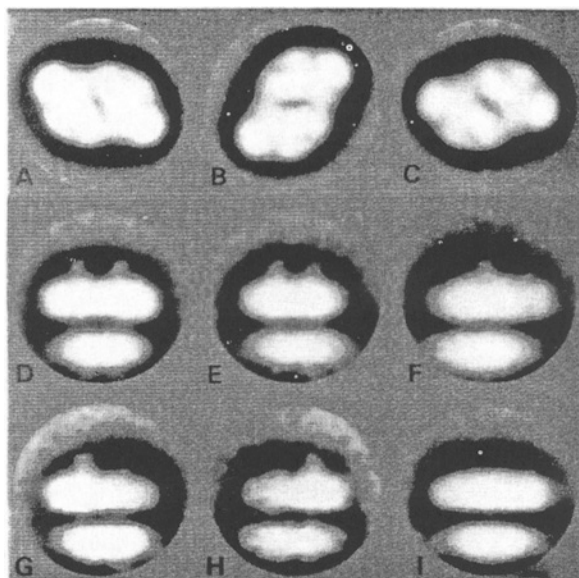


Fig. 7. Averaged top view projection and averaged side views of PS II core complex dimers from *Synechococcus elongatus* (E.J. Boekema, D. Bald, J. Kruijff and M. Rögner, unpublished results). The averaged top-view projection is shown in three positions (A–C) where it matches the 6 classes of side views (D–I) from a classification of 443 projections as well as possible. Position as in (A) is most compatible with the side view of (D); (B) is compatible to (E) and (C) to (F). The classes of Fig. (G–I) show sums of projections of particles in an uneven stain layer (G,H) or with fuzzy protrusions (I). Notes: most of the analyzed side-views had a neighbor, but the alignment was only focussed on the upper dimers and thus, in summation, the lower ones appear fuzzy. Classification of the top views showed that only one predominant type of projection was present.

are thought to represent the 33 kDa oxygen-evolving subunit resulting in a maximal height of  $90 \text{ \AA}$  for the particles. Image analysis of side views gives another example of the usefulness of single particle analysis. By correspondence analysis and classification, a discrimination between different views was possible, which we interpret as overlap- and non-overlap views. In the non-overlap views two protrusions can be seen (Fig. 7D, E), whereas the overlap view shows one protrusion (Fig. 7F). Note that the non-overlap view shows a longer projection than the overlap views, indicating a different position of the respective dimers on the carbon support.

#### IV. Concluding Remarks

In conclusion, EM techniques are available to study isolated (membrane) protein complexes from very small to very large size. In combination with techniques used in cell biology for studying whole cells or cell fragments, EM has the possibility and the potential to supply the field of photosynthesis with a detailed picture of the photosynthetic membrane and its interacting proteins. For structure determination, crystals are more useful than single particles. But not in all cases will it be possible to grow good crystals in a short term for such complicated structures as PS I, PS II and the  $F_0F_1$ ATP synthase complex including their interacting donor-, acceptor- and regulatory molecules and proteins. In the meantime a combination of low-resolution EM reconstructions of these photosynthetic membrane complexes with atomic structures of their individual components, determined by EM, NMR or X-ray diffraction, will be useful for understanding the structure and function of these proteins.

#### Acknowledgements

We are grateful to Dr. W. Keegstra for his help with computer image analysis, Dr. G. Perkins for discussion and Mr. K. Gilissen for photography. Work has been supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for Scientific Research (NWO), by a grant from the European Union BIO2CT-930078 (EJB), the Deutsche Forschungsgemeinschaft (MR) and a grant from NEDO/RITE, Japan (MR).

#### References

- Adrian M, Dubochet J, Lepault J, McDowell AW (1984) Cryoelectron microscopy of viruses. *Nature* 308: 32–36
- Boekema EJ (1991) Negative staining of integral membrane proteins. *Micron and Microsc Acta* 22: 361–369
- Boekema EJ, Dekker JP, Rögner M, Witt I, Witt HT and van Heel, MG (1989) Refined analysis of the trimeric structure of the isolated Photosystem I complex from the thermophilic cyanobacterium *Synechococcus* sp. *Biochim Biophys Acta* 974: 81–87
- Boekema EJ, Boonstra AF, Dekker JP and Rögner (1994) Electron microscopic structural analysis of Photosystem I, Photosystem II, and the cytochrome b<sub>6</sub>/f complex from green plants and cyanobacteria. *J Bioenerg Biomembr* 26: 17–29
- Böttcher B, Gräber P and Boekema EJ (1992) The structure of Photosystem I from the thermophilic cyanobacterium *Synechococcus* sp. determined by electron microscopy of two-dimensional crystals. *Biochim Biophys Acta* 1100: 125–136
- Dekker JP, Boekema EJ, Witt HT Rögner M (1988) Refined purification and further characterization of oxygen-evolving and Tris-treated Photosystem II particles from the thermophilic cyanobacterium *Synechococcus* sp. *Biochim Biophys Acta* 936: 307–318
- Engel A, Hoenger A, Hefti A, Henn C, Ford RC, Kistler J and Zulauf M (1992) Assembly of 2-D membrane protein crystals: dynamics, crystal order, and fidelity of structure analysis by electron microscopy. *J Struct Biol* 109: 219–234
- Frank J (1982) New methods for averaging non-periodic objects and distorted crystals in biologic electric microscopy. *Optik* 63: 67–89
- Frank J, Radermacher M, Wagenknecht T and Verschoor A (1988) Studying ribosome structure by electron microscopy and computer-image processing. *Methods in Enzymology* 164: 3–35
- Hawkes PW and Valdrè U (1990) *Biophysical Electron Microscopy. Basic concepts and modern techniques.* Academic Press, London
- Henderson R, Baldwin JM, Downing KH, Lepault J and Zemlin F (1986) Structure of purple membrane from *Halobacterium halobium*: recording, measurement and evaluation of electron micrographs at 3.5 Å resolution. *Ultramicroscopy* 19: 147–178
- Henderson R, Baldwin JM, Ceska TA, Zemlin F, Beckmann E and Downing KH (1990) Model for the structure of bacteriorhodopsin based on high-resolution electron cryomicroscopy. *J Mol Biol* 213: 899–920
- Jap BK, Zulauf M, Scheybani T, Hefti A, Baumeister W, Aebi U and Engel A (1992) 2D crystallization: from art to science. *Ultramicroscopy* 46: 45–84
- Kruip J, Boekema EJ, Bald D, Boonstra AF and Rögner M (1993) Isolation and structural characterization of monomeric and trimeric Photosystem I complexes (**P700.F<sub>A</sub>/F<sub>B</sub>** and **P700.F<sub>X</sub>**) from the cyanobacterium *Synechocystis* PCC 6803. *J Biol Chem* 268: 23353–23360
- Kühlbrandt W (1992) Two-dimensional crystallization of membrane proteins. *Quarterly Rev of Biophys* 25: 1–49
- Kühlbrandt W, Wang DN and Fujiyoshi Y (1994) Atomic model of plant light-harvesting complex by electron crystallography. *Nature* 367: 614–621
- Mörschel E and Schatz GH (1987) Correlation of photosystem-II complexes with exoplasmic freeze-fracture particles of thylakoids of the cyanobacterium *Synechococcus* sp. *Planta* 172: 145–154
- Rögner M, Dekker JP, Boekema EJ and Witt HT (1987) Size, shape and mass of the oxygen-evolving photosystem II complex from the thermophilic cyanobacterium *Synechococcus* sp. *FEBS Lett* 219: 207–211
- Staehelin LA (1988) Chloroplast structure and supramolecular organization of photosynthetic membranes. In: Staehelin LA and Arntzen CJ (eds.) *Photosynthesis III*, pp. 1–84, Springer-Verlag, Berlin

- Unwin PNT and Henderson R (1975) Molecular structure determination by electron microscopy of unstained crystal-line specimens. *J Mol Biol* 94: 425–440
- Van Heel M and Frank J (1981) Use of multivariate statistics in analysing the images of biological macromolecules. *Ultramicroscopy* 6: 187–194
- Van Heel H and Harauz G (1988) Biological macromolecules explored by pattern recognition. *Scanning Microscopy, Supplement 2*: 295–301
- Zemlin F (1992) Desired features of a cryoelectron microscope for the electron crystallography of biological material. *Ultramicroscopy* 46: 25–32