the
abdus salam
international centre for theoretical physics

SMR: 1098/14

WORKSHOP ON THE STRUCTURE OF BIOLOGICAL MACROMOLECULES

(16 - 27 March 1998)

"Electron Tomography of Biological Specimens"

presented by:

Jose Maria CARAZO
Universidad Autónoma de Madrid
Centro Nacional de Biotecnologia - CSIC
Cantoblanco
28049 Madrid
Spain

Contemporary Perspectives in
Three-Dimensional Biomedical Imaging
C. Roux and J.-L. Coatrieux (Eds.)
IOS Press, 1997

ELECTRON TOMOGRAPHY OF BIOLOGICAL SPECIMENS

Roberto MARABINI (1)(2), M. Carmen SAN MARTIN (1) and Jose M. CARAZO (1)(3)
(1) Centro Nacional de Biotecnología (C.S.I.C.)
Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid (Spain)
(2) Silicon Graphics S.A., Edificio "Santa Engracia 120"
Plaza del Descubridor Diego de Ordás, 3, 28003 Madrid (Spain)
(3) Departamento de Arquitectura y Tecnología de Ordenadores
Universidad de Málaga, Málaga (Spain)
e-mail: carazo@samba.cnb.uam.es

Abstract. This chapter addresses the issue of three-dimensional reconstruction of biological macromolecules from electron microscopy images. First, the electron microscope device, the micrographs as specimen projections and the projection collection geometry are discussed. After that, a brief introduction to the problems involved in sample preparation and observation is made. In a third section the different three-dimensional reconstruction methods are presented, and finally, the chapter finishes with an example in which a protein is reconstructed starting from micrographs and ending with a nice and very informative volume.

1 Introduction

The structure determination of large macromolecular assemblies is at present one of the key subjects in the biochemical research. Many basic biological processes, including DNA metabolism, photosynthesis, protein synthesis and viral assembly, involve the concerted action of a large number of components. Understanding the three-dimensional organization of these components, as well as their detailed atomic structures, is central to the interpretation of their function. Interest in the structure of biological molecules dates to the first attempts at X-ray diffraction from protein crystals. However, it was the invention of the electron microscope which made possible direct imaging of biological structures at a macromolecular level, and more recently, even at atomic resolution.

Most of the effort in the electron microscopy field is devoted to a direct interpretation of the micrographs (photographs taken at the electron microscope) with the goal of obtaining quantitative measures: general shape, dimensions, presence of a given component, etc. This methodology, although interesting in a broad range of cases, does not extract all the information contained in the micrographs. It must be realized that the micrographs are, in many cases, an altered version of the structure of the *in vivo* specimen. In addition to the possible artifacts induced during sample preparation or visualization, the images might not be directly interpretable due to instrumental imperfections: aberrations, defocus, etc. And, most of all, micrographs represent a three-dimensional object in only two dimensions, therefore "missing" one dimension[1].

To minimize the effect of these problems, different techniques have been developed during the last 30 years with the idea of improving both the specimen preparation and

[1] Fortunately, as we will see in this chapter, it is possible to define a set of conditions under which it is possible to recover the three-dimensional structure of the specimen from a set of these micrographs.

the imaging process. At the same time, the computerized image processing techniques have become of age and are now readily available to obtain as much information as possible from the micrographs.

This chapter is mainly devoted to the three-dimensional reconstruction of biological macromolecules from electron microscopy images. We will first introduce the electron microscope device, the micrographs as specimen projections and the projection collection geometry. After that, we will present a brief introduction to the problems involved in sample preparation and observation. In a third section the different three-dimensional reconstruction methods will be presented, and finally, the chapter finishes with an example in which a protein is reconstructed starting from micrographs and ending with a nice and very informative volume.

2 Instrumentation and data collection

2.1 The electron microscope in biology

The transmission electron microscope is at present an indispensable tool in structural biochemistry and biology, providing the most direct way of visualising structure at a molecular level.

In overall design, the transmission electron microscope is similar to the widely known light microscope, although it is much larger and upside down. The source of illumination is a filament (cathode) that emits electrons at the top of a cylindrical column about two meters high (Fig. 20). Since electrons are scattered by collisions with air molecules, air must first be pumped out of the column to create a vacuum of at least 10^{-5} Torr. The electrons are then accelerated from the filament by a nearby anode and allowed to pass through a tiny aperture to form an electron beam of paralel rays that travels down the column. Magnetic coils placed at intervals along the column focus the electron beam by means of controlled variation of magnetic fields, just as glass lenses focus the light in a light microscope. The specimen (mounted on a support 3 mm in diameter) is put into the vacuum, throug an airlock, into the path of the electron beam, right inside the magnetic coil acting as objective lens. Some of the electrons passing through the specimen are scattered according to the local density of the material; the remainder are focused to form an image, either on a photographic plate or on a phosphorescent screen, or digitally registered by a CCD camera in modern devices.

In an electron microscope, the resolution limit imposed by the wavelength of visible light is decreased using electrons instead of photons. For an accelerating voltage of 100 kV (a typical value in biology) the wavelength of an electron is 0.004 nm, and, in theory, the resolution of such a microscope should be about 0.002 nm. However, because the aberrations of an electron lens are considerably harder to correct than those of a glass lens, the practical resolving power of most modern electron microscopes is, at best, 0.1 nm (1 Å). Furthermore, problems of specimen preparation, contrast, and radiation damage effectively limit the normal resolution for most of biological objects to an order of magnitude more (10 Å). As we shall see below (section 4.2), new specimen preparation and observation techniques, combined with the merging of images and electron diffraction data, are currently allowing to recover a great part of the electron microscope theoretical resolution limit.

The standard high resolution electron microscope has a depth of focus of several

thousand Ångströms. This is also the thickness range of a typical biological sample. As a consequence, electron microscopy images are formed by a superposition of structural features lying at different levels in the three dimensional structure. Since micrographs contain information on the three-dimensional structure of the specimen, it could be imagined that by a proper treatment of these images it should be possible to obtain the three dimensional relationships between the different parts of the specimen. This is the general aim of electron tomography.

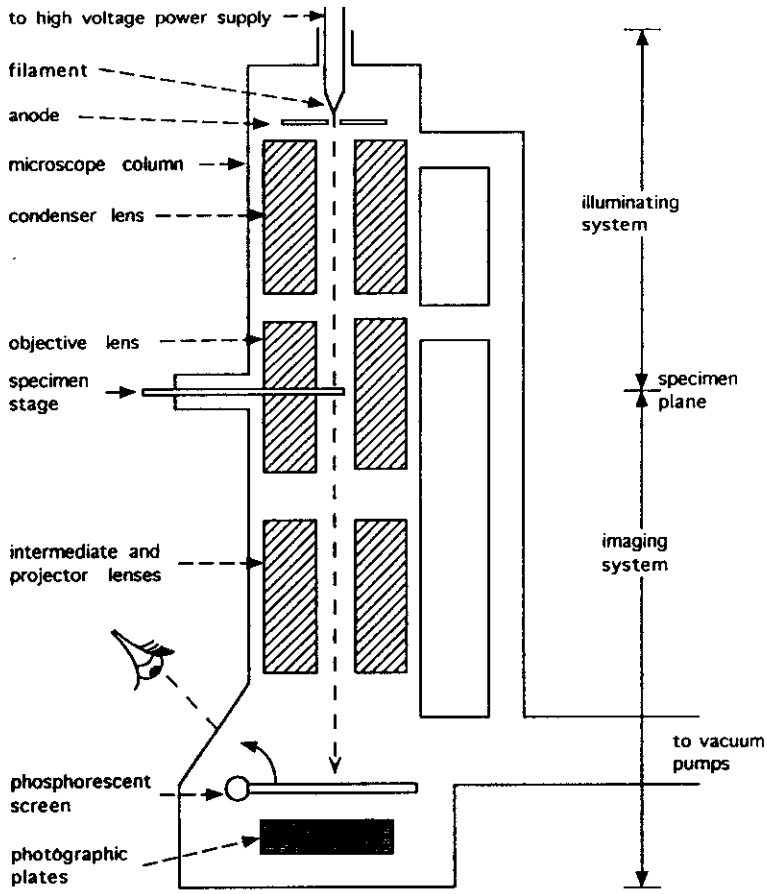

Figure 20: Schematic representation of a transmission electron microscope

2.2 *The micrographs as specimen projections*

The intuitive understanding of the 3-D reconstruction scheme is based on a set of assumptions, one of the most important being the belief that electron micrographs are in "some sense" projections of the original volume. Although this assumption is the starting point of many electron microscope text books and reconstruction methods, it is far from self evident. This section presents a brief outline of the image formation theory in transmission electron microscopy (TEM) and justifies this assumption.

The first step in the image formation process in TEM consists in the irradiation of the sample with a beam of electrons. Most of the electrons do not pass near any specimen atom and therefore will go through the sample without any energy or phase change (unscattered electrons). A few electrons will suffer interactions with the object atoms. In most cases this interaction will be weak (elastic scattering), meaning that the electron will change its phase but not its energy. Finally, an almost negligible amount

of electrons will interact strongly with the sample, losing or gaining energy (inelastic scattering).

Thus, contrast in TEM is mainly due to elastic scattering, that is, almost every electron that falls upon the sample crosses it and contributes to form the image. From this point of view it is said that in TEM the images are formed by phase contrast due to the interference between unscattered and weakly scattered electrons. Nevertheless, a kind of amplitude contrast is obtained by introducing an aperture in the beam path that stops the more strongly scattered electrons.

This qualitative picture of the specimen-electron interaction warns us against performing a straightforward extension of the imaging theory developedfor optical microscopes.

It is possible to approximate the electron beam by a plane wave expressed as

$$\Psi = \Psi_0 e^{2\pi i k z}$$

that travels along the microscope axis (z coordinate),

The outgoing beam after elastic interaction with the specimen contains both unscattered and elastically scattered electrons, so that we can describe it as

$$\Psi = \Psi_0 e^{2\pi i k z} + \Psi_{sc}$$

where

$$\Psi_{sc} = i\Psi_0 f(\Theta)\frac{e^{2\pi i k z}}{r} \tag{19}$$

The function $f(\Theta)$ is the complex scattering amplitude, $f(\Theta) = |f(\Theta)|exp(i\eta\Theta)$, and k is the wave number

$$k = \frac{1}{\lambda}$$

It is convenient to treat the two different sources of contrast (phase and amplitude[2]) separately. In the case of amplitude contrast, the aperture size of the objective lens fixes the maximum scattering angle α_{max} of the electrons that reach the image plane. The number of electrons, n, scattered an angle greater than α_{max} is:

$$\frac{dn}{n} = -N\sigma(\alpha)d\mu \tag{20}$$

where N is the Avogadro number, α the scattering angle, μ the mass thickness (density multiplied by thickness) of the object and σ the sample atoms scattering cross section

$$\sigma(\alpha) = \int_{\alpha}^{\pi} |f(\Theta)|^2 2\pi \sin(\Theta) d\Theta$$

which tells us the probability of elastic or inelastic scattering as a function of the incident beam energy, the atomic number of the scattering atom, and the angular scattering probability, which allows us to determine the mean scattering angle.

Solving (20) it is obtained:

$$n = n_0 e^{-A\mu} \tag{21}$$

where A is a positive constant.

[2]As stated in the previous paragraphs, amplitude contrast in TEM is not due to electron absorption but to elimination of high angle scattered electrons.

Equation (21) shows that, concerning the amplitude contrast, the number of electrons recorded in the micrograph is related to the specimen density μ, and, consequently, an electron micrograph can be described as a projection image.

The amplitude contrast mechanism, that can be described as the presence or absence of electrons in the image plane, is the main source of contrast in low/medium resolution electron microscopy (less than 20-30 Ångströms). This situation changes when higher resolution is desired; in this case phase contrast (interaction between electrons with the same optical path but different phases) comes into action.

The formalism used to describe this phenomenon, called phase-granting approximation [353], is much less straightforward than in the previous case. In the phase-granting approach the specimen is represented by a collection of very thin slices. The incident beam passes through the first slice where it is modified, then by the second one and so on. A detailed description of this theory can be found in [332] and [319]. As a consequence, the information contained in the final outgoing wave can be described as a projection when the conditions established by the phase-granting approximation are satisfied, that is: thin samples and high energy electrons. Or, from a more intuitive point of view, when the electron wave function does not spread as it passes through the object.

This theory has rendered some quantitative results than interrelate the specimen thickness (t), the electron wave length (λ) and the desired resolution (d):

$$t \leq \frac{d^2}{\lambda} \tag{22}$$

It is then clear that TEM images can be considered as parallel beam projection images for thin specimens and high energetic electrons, and that thinner specimens and higher beam energies are required as the desired resolution increases.

2.3 Data collection geometry and missing information

To perform the 3-D reconstruction of a macromolecular structure from electron microscopy images it is necessary to collect a set of views of the specimen from different directions. In general, the views are obtained by tilting the specimen with the help of a goniometer.

There are technical limits to the maximum angle that a commercial goniometer can achieve, usually around $\pm 60°$. But even if a specially designed goniometer is used to overcome this technical difficulties, there will be physical limits to the maximum angle attainable while still obtaining useful images. These limits arise from the increase in the effective specimen thickness according to 1/cos(tiltangle).

The general problem in electron microscopy is, thus, the 3-D reconstruction of a structure from a finite set of noisy 2-D projections over a restricted angular range (usually between $\pm 60°$). In the following discussion we will use the Fourier space formalism, where the effects of these limitations is much more easily understood. Additionally, we will use the central section theorem, which states that the Fourier transform of a parallel beam 2-D projection image is equal to the central section of the 3-D Fourier Transform of the object [314].

Because of their importance in electron tomography, two types of data collection are going to be introduced: single axis and random conical tilting.

2.3.1 *Single axis tilt series*

In this data collection scheme the object is tilted in small increments around a fixed axis and micrographs of the specimen are obtained in each orientation (see figure 21).

By calculating the Fourier transform of each projection and using the central section theorem[3], we obtain a set of planes in the Fourier space that have a common axis (fig 21.b, 21.c, 21.d). This description shows that the calculation of the specimen 3-D Fourier transform can be decomposed as a collection of 2D problems and that there is a region (fig 21.c) of the 3-D Fourier transform, called missing wedge, where no data can be measured.

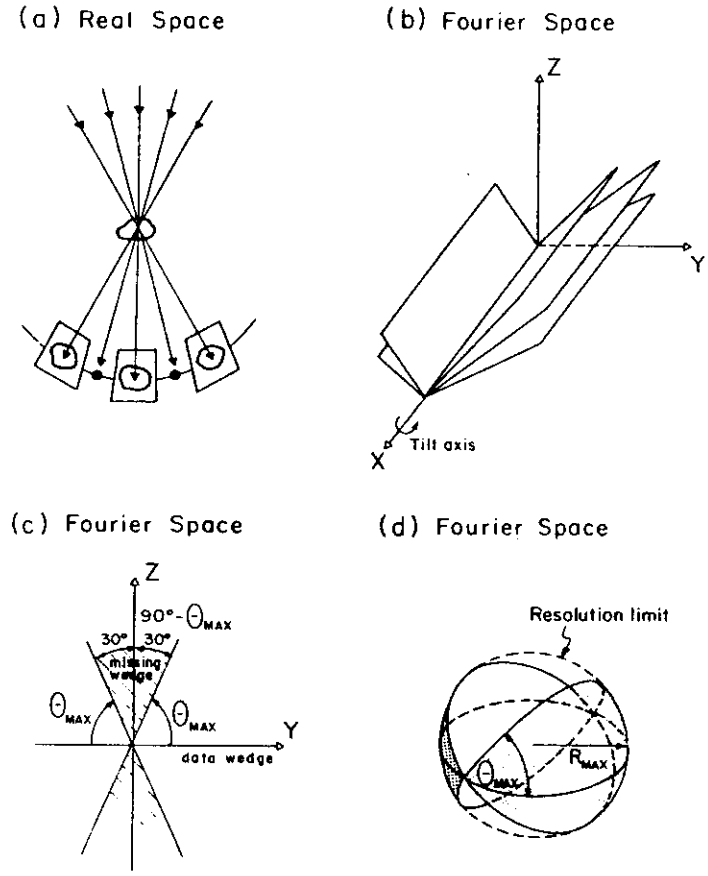

Figure 21: Single-axis tilt data collection geometry. (a) Tilting in real space around an axis perpendicular to the plane of this page (which is assumed to be the x axis). (b) Fourier space representation of the information presented in (a), where the different projection views are central planes through the three-dimensional Fourier transform of the object with the X axis in common. (c) Close-up view of (b) from the X axis; a missing wedge of 30° zenithal angle around the Z axis is shown in dashed lines. (d) Spatial representation in Fourier space in a practical situation where Θ_{max} is the maximum tilt angle and R_{max} is the maximum frequency. ((a) and (d) from [329], reproduced with permission.)

[3]The central section theorem states that the two dimensional Fourier transform of a plane projection of a three dimensional density distribution is equal to the corresponding central section of the three dimensional transform normal to the direction view.

2.3.2 *Conical tilt series*

A conical tilt series is recorded by tilting the specimen by the maximum attainable tilt angle Θ_{max} and rotating it in the tilt plane by small increments. A micrograph is recorded in each position. As in the previous case, the 3-D reconstruction is divided into a collection of 2D problems and, again, there is a region of the 3-D transform, this time cone-shaped (see fig. 22), that remains undetermined.

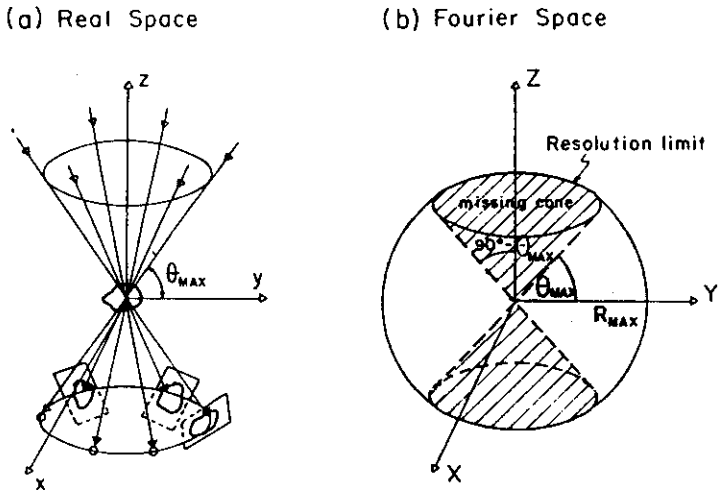

Figure 22: Conical tilt data collection geometry. (a) Tilting in real space by Θ_{max} around multiple tilt axis in the xy plane. (b) Fourier space representation of the information collected in (a); no experimental data are available in a "missing cone" around Z. (From [329], reproduced with permission.)

In general, single axis tilt series are used when the specimen has been crystallized, while conical tilting is preferred for isolated specimen preparations (see section 4 for further details).

3 Biological specimens and their peculiarities

3.1 *Characteristics of biological macromolecules*

Before taking on the reconstruction problem itself, it is necessary to consider both the kind of object to be reconstructed and its behaviour during the imaging process. Biological macromolecules are small, with dimensions ranging in the tens of nanometers. This tiny (microscopic!) size implies, apart from the need for specialized techniques of handling, that no researcher has ever seen the real aspect of such an aggregate, but by means of more or less indirect methods. That is, there is no model to serve as a reference for the quality of the reconstruction procedures.

Besides, the environmental conditions in the electron microscope are more than deleterious for the specimen structure. Water is the most abundant component of biological material, but it is systematically excluded from conventional electron microscopy, due to the vacuum conditions inside the column of the electron microscope. As a consequence, it has been usual to observe the samples in a dehydrated state, and techniques have been developed to preserve as best as possible the native structure of the specimen

without water. The most successful, from the point of view of three-dimensional reconstruction, has been that of "negative staining", in which the specimen is embedded in a thin amorphous layer of a heavy metal salt which simultaneously preserves and maps out the shape of the region from which it is excluded. Reconstructions performed by this kind of preparation technique achieve resolutions in the vicinity of 20 Å.

The problem of beam-induced damage is by no means negligible in biological electron microscopy. Electron radiation produces intense ionization in organic materials, which results in the formation of free radicals and ions. This causes bond scission and molecular fragments being formed. The free radicals and molecular fragments can undergo diffusion and produce cross-linking or further chain scissions. Damage to the secondary structure of proteins occurs at an electron dose of less than 1 $e/\mathring{A}^2$, and further exposure causes the tertiary structure to undergo dramatic reorganization following the loss of specific groups and altered structural composition. Apart from substantial alterations of the high resolution structure, the dominant final effect is mass loss of the sample, accompanied by shrinkage of the sample normal to the beam. Staining of the sample with heavy metal salts has also been a successful method to make it less sensitive to electron radiation, but it limits the resolution and is not free of migration and shrinkage phenomena induced by the radiation, finally obscuring the original structure of the specimen.

In the last years, the development of cryoelectron microscopy has provided important advances towards the solution of both the dehydration and the radiation damage problems. Biological macromolecules in solution are vitrified by rapidly plunging them into a suitable criogen (liquid ethane, for example), in such a way that the water content reaches an amorphous frozen state without passing by a crystalline state, that would alter the structures under study. In this same state, under liquid nitrogen, the samples are transferred to the electron microscope and observed at temperatures near 100 K, always maintaining a fully hydrated environment. Furthermore, radiation damage is less severe at low temperatures than in a conventional electron microscopy by a factor of three to six [312]. It has been proposed that this increased resistance is due to the slower rearrangement or diffusion of the fragments resulting from bond fracture [333], so that although irradiation eventually results in a total destruction of the specimen, there is ample time to record images from an object in a state as close as possible to the natural one. This approach has made possible the reconstruction of crystalline specimens to almost atomic resolution, but it is technically difficult and requires high expertise.

3.2 Biological macromolecules and their images

The characteristics of the biological material described in the preceding section have as a direct consequence some special features of the images obtained in the electron microscope, which in turn determine some of the more critical steps in their processing for the three-dimensional reconstruction of the object.

The destructive action of the electron beam on the samples is clearly the major drawback of electron microscopy when high resolution structural studies are concerned. To overcome this problem, methods have been developed in which the specimen image is recorded with as low a radiation as possible. Conventional microscopy involves doses of 50-500 $e/\mathring{A}^2$, including the time involved in searching and focusing. Amos and coworkers [311], established that a total dose on the sample of 0.5 to 4 $e/\mathring{A}^2$ is appropriate

for very high resolution studies of unstained crystalline specimens. A total dose around 10-20 $e/\mathring{A}^2$ is termed a minimal dose, and is typically used for stained or non crystalline specimens. These low dose methods require a search of suitable areas to be done at very low magnification, during which the dose should be extremely low (0.02 $e/\mathring{A}^2$). The focusing is done at an adjacent area at high magnification, and then the image is recorded at the desired magnification, irradiating the area of interest only during the time needed for plate exposition.

Unfortunately, the use of minimal exposures to avoid destruction of the specimen brings out a new difficulty in the interpretation of electron microscopy data. The electron scattering phenomenon is a highly stochastic process, and irradiations of only a few electrons result in little statistical significance of the distribution recorded on the plate. In other words, electron micrographs are extremely noisy images. This problem is specially severe in the case of frozen, hydrated, unstained specimens, where the absence of strong scattering atoms produces images with little or no contrast at all.

The solution devised to improve the poor signal to noise ratio of the micrographs has been to average the contributions from a large number (several thousands) of particles. This can be done directly in the case of crystalline specimens, where particles are a priori aligned, or after translational and rotational alignment in the case of single specimens, which are found in random orientations on the support film. The alignment procedures must be robust, because apart from the very low signal to noise ratio of the images, the high variability intrinsic to the population involved in the average must be taken into account. Biological objects tend to be flexible, and may suffer some kind of deformation or present different conformations in the same preparation. Furthermore, the existence of real impurities (different biological species) in the sample cannot always be ruled out. Finally, the very same object can present a wide range of views when deposited onto a support or embedded in a thin film of ice. If an aligning algorithm is used with such a heterogeneous input set, the results will be poor, if not completely artefactual. Therefore, classification and aligning procedures have come to be critical in processing the data which are eventually leading to the three-dimensional reconstruction of the object.

4 Different methods for different reconstructions

Different approaches have been devised to reconstruct biological specimens from their EM projections. The simplest situation is presented by specimens that are aggregated into a helical structure. In this case, a single view carries enough information to reconstruct the original volume up to a certain resolution [315] [316]. There are other types of biological macromolecules that are also highly symmetrical: this is the case, for instance, of the icosahedral viruses, where a single view of the specimen already provides another 59 views [314] or the 2-D crystals that present translational symmetry [321]. For a general case, however, we cannot count on symmetries since we may have an isolated and asymmetric object.

Following the chronological developments in the field, the symmetric isolated specimens method [314] will be described first, then the crystal methodology [321] and finally the single particle case [585] (a single particle is an isolated random oriented particle with, in principle, no special symmetry).

4.1 Highly symmetrical particles

The symmetry of the specimens is a very desirable a priori characteristic because an image effectively contains many different views of the same structure. In fact, in some extreme cases [315] when the specimen presents helical symmetry, a single view may provide enough information to reconstruct the object up to a certain resolution. For non-helical particles it is necessary to combine data from different projected views.

The method of reconstruction is based on the central section theorem that relates the 3-D Fourier transform of the reconstruction with the 2-D Fourier transform of the projection images.

In a general case, different particles of the same specimens are viewed from different angles[4]. The views are digitized and their Fourier transform calculated. Each view provides a plane in the Fourier space ; the symmetry of the particle can be used to calculate further sections of the 3-D Fourier transform related to the first one, and to obtain the absolute position of the different views so they can be combined.

The relative position between the views it is determined by a method based on the existence of "common lines"[5]. Let us represent two arbitrary projections of an object in 3-D Fourier space by their associated central section. These sections intersect on a line through the origin, their common line. If the object presents symmetries, the number of common lines between any two images is increased, and can be easily found using crosscorrelation methods[6].

Using the above approach the Fourier transform of the volume is filled, although with points that are not regularly spaced. Due to the Discrete Fourier transform definition it is necessary to interpolate a grid of equally spaced points, and after that, the reconstructed volume is finally obtained by Fourier transform inversion.

A recent example of the application of these techniques is the study of the three-dimensional structure of hepatitis B virus core particles, by Crowther *et al.* [313]. The core particles of human hepatitis B virus are formed by the protein known as core antigen HBcAg, responsible for the packaging of pregenomic RNA, and in complex with another viral protein (the viral polymerase), for the generation of the mature DNA genome by reverse transcription. These authors also studied core particles assembled from a mutant of HBcAg, in which the packaging of RNA is abolished.

After electron cryomicroscopy of the viral particles (Fig. 23) and three-dimensional reconstruction based on the specimen symmetries, Crowther and co-workers showed that core particles from both native and mutant protein assembled in a similar way, and presented holes or channels running through the protein shell. These holes are thought to provide access for some small molecules required for the synthesis of DNA during reverse transcription (Fig. 24).

The use of cryomicroscopy of unstained specimens allows imaging of the inner contents of the particles. Particles assembled from the intact core protein appear filled with additional material, both in the micrographs and after three-dimensional recon-

[4]It is not necessary to tilt the sample holder to obtain different views because the high symmetric particles, that are almost spherical (icosahedral, dodecahedral, etc.), lay on the support film presenting a width range of views.

[5]The existence of common lines is a direct consequence of the central section theorem. As the 2-D Fourier transform of a view is a central plane of the 3-D Fourier transform of the reconstruction, every two 2-D Fourier transforms intersect, and therefore have a line with identical values (in the absence of noise). In the case of symmetrical particles the number of "common lines" is increased.

[6]The Fourier transform of an n-fold image presents n-fold symmetry, and therefore each common line is repeated n times.

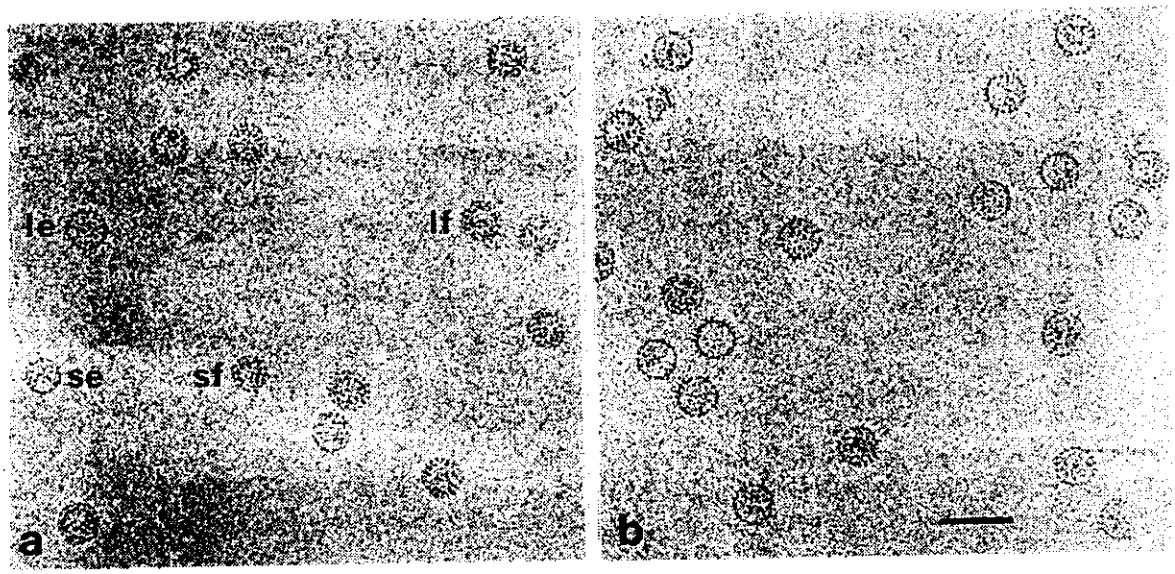

Figure 23: Cryomicrographs of Hepatitis Core Particles assembled from intact HBcAg protein (a) or from the mutant protein(b), unstained and embedded in a thin layer of vitreous ice. **le** and **se**, empty particles; **lf** and **sf**, full (RNA-containing) particles. Scale bar, 50 nm. (From [313], ©Cell Press 1994, reproduced with permission)

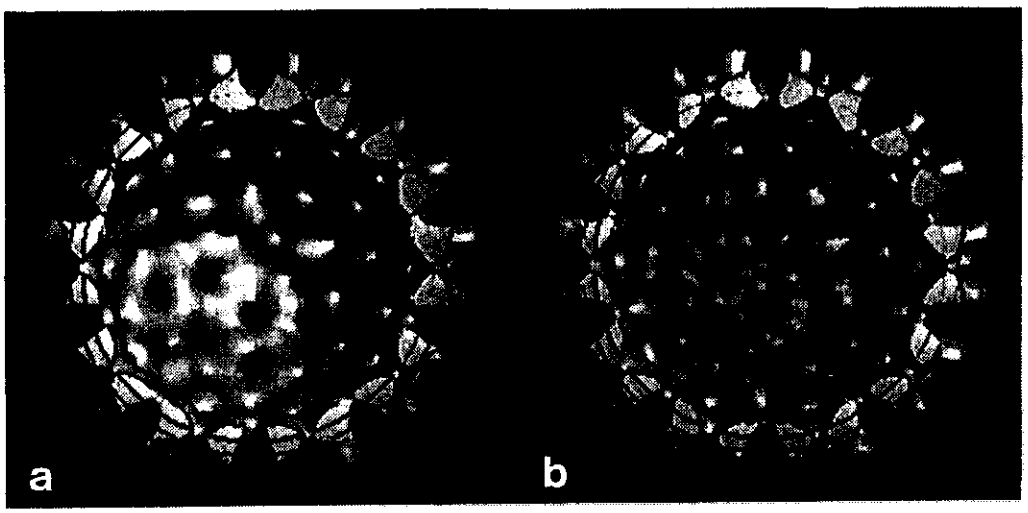

Figure 24: Views of the inner surface of the Hepatitis B core empty particles made from intact (a) and mutant (b) HBcAg. The views show clearly the pattern of holes that penetrate the shell of protein. (From [313], ©Cell Press 1994, reproduced with permission)

struction. This inner shell is not seen in particles assembled from the mutant protein, which does not pack RNA, and may therefore correspond to the pregenomic molecule, icosahedrally ordered as a consequence of its interaction with the protein scaffolding (Fig. 21).

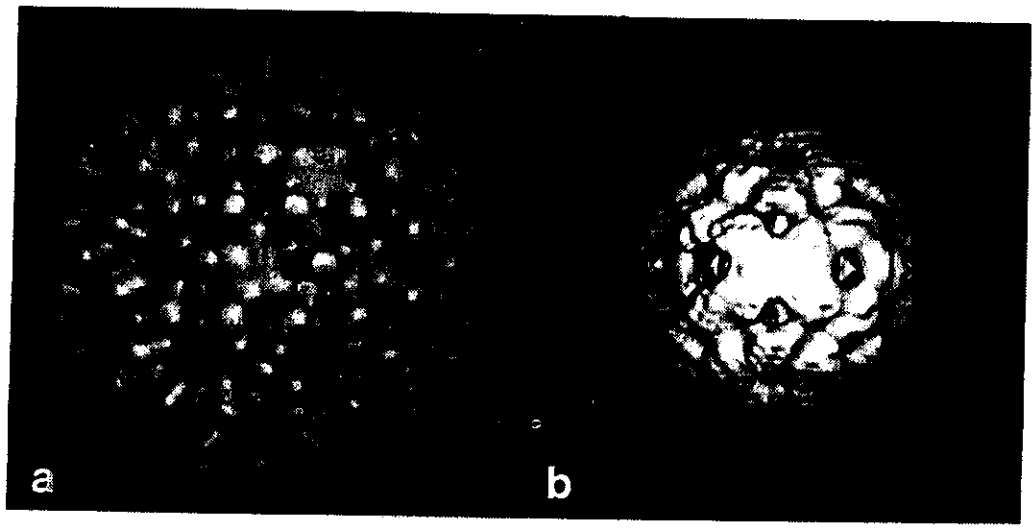

Figure 25: Three-dimensional map of the full HBcAg particle. (a) shows the outside of the particle, while (b) shows the same map but with the outer protein capsid computationally removed to reveal an inner shell of density not present in the map of the empty particle (Fig. 24). (From [313], ©Cell Press 1994, reproduced with permission)

4.2 Electron crystallography

Electron crystallography can be used to obtain three dimensional density maps that may reach atomic resolution [322]. The high scattering power of electrons makes it possible to get an adequate signal from a crystal of biological macromolecules that is only one unit cell thick (monolayer crystal), such as the crystals produced by membrane proteins in their native environment. 3-D crystals, which are several unit cells thick, produce multiple scattering and therefore images that, at present, are useless for volume reconstruction (see section 2.2).

The single axis geometry (see section 2.3) is followed to obtain the different views necessary to perform the reconstruction, that is, a collection of micrographs with the crystal tilted at different angles are taken. The micrographs are then digitized and their Fourier transform calculated. One of the interesting properties of the Fourier space is that a periodic object is completely described by a collection of sharp points. By deleting any signal not close to these points the image signal to noise ratio is greatly improved, and the information related with the periodic structure (unit cell) is easily separated from the noise.

According to the central section theorem, the 3-D Fourier transform of the reconstructed volume can be built up by combining the filtered Fourier transform of the images. To determine the position of each plane in three dimensions the tilt axis direction and the tilt angle must be found in each micrograph. This task is solved by analyzing the untilted and the tilted lattice parameters.

Three more steps are necessary before obtaining the volume; they are known as scale factor determination, common 3-D phase origin determination and 3-D Fourier

transform interpolation.

The need for scale factor determination arises from the differences in dynamic range and number of unit cells between different micrographs belonging to the same tilting series. These differences are corrected, in the 3-D Fourier space, by comparing the amplitude of spots that belong to two or more projections and applying to one of the images the multiplicative factor required to match the amplitudes.

The phase origin determination fixes the tilt axis in each micrograph. It is performed following an approach similar to the scale factor determination. The phase belonging to common spots is compared and shifted (in the whole Fourier transform) by the necessary amount to minimize their differences.

At this step the Fourier transform of the volume is filled, except in the missing wedge region, with points that are not regularly spaced. As in the previous method it is necessary to interpolate a grid of equally spaced points and finally invert the 3-D Fourier transform.

Electron crystallography is becoming increasingly successful as a technique for solving biological macromolecular structures at up to atomic resolutions, close to those attained by X-ray crystallography and NMR (Nuclear Magnetic Resonance). For example, following the philosophy of reconstruction described, and making use of the more advanced techniques for sample preparation and observation, Kühlbrandt and co-workers have been able to propose an atomic model for the structure of the light-harvesting chlorophyll a/b-protein complex LHC-II [322].

LHC-II, the most abundant membrane protein in chloroplasts, is responsible for collecting the solar energy required to perform photosynthesis in all green plants. The purified protein complex forms highly ordered two-dimensional crystals, which were examined in an electron cryomicroscope capable of resolving 2.0 Å object features at a specimen stage temperature of 4.2 K, an extremely low temperature at which radiation damage is much less severe. Electron diffraction of the crystals yielded the structure factor amplitudes, and phases were determined directly by processing electron micrographs taken at tilt angles ranging between 0 and 60° (Fig. 26). A three-dimensional map of LHC-II was calculated, using these amplitude and phase data, to 3.4 Å resolution. At this high resolution, it was possible to trace the polypeptide chain disposition on the volume, and to distinguish the bound pigment molecules (chlorophylls and xantophylls), establishing a model for the atomic structure of the complex (Fig. 27) that contributes to a more complete understanding of the photosynthetical process.

4.3 Single particles

In the attempt to reconstruct a biological macromolecule from its projections a few requirements must be fulfilled. Among others, we need several projections that effectively cover the 3-D Fourier space, and the electron dose must be lower than a few electrons/$\mathring{A}^2$.

Until the mid eighties, these two desired properties were only satisfied in two situations: when the specimen was a 2-D crystal or a symmetrical particle. In these cases imaging with extremely low electron doses is possible because the projections can be enhanced by averaging the repeated motive or by using the particle symmetry properties. These two methods still give the best quality reconstructions at present, but they suffer from the drawback that many biologically interesting macromolecules can not be crystallized and are not symmetrical.

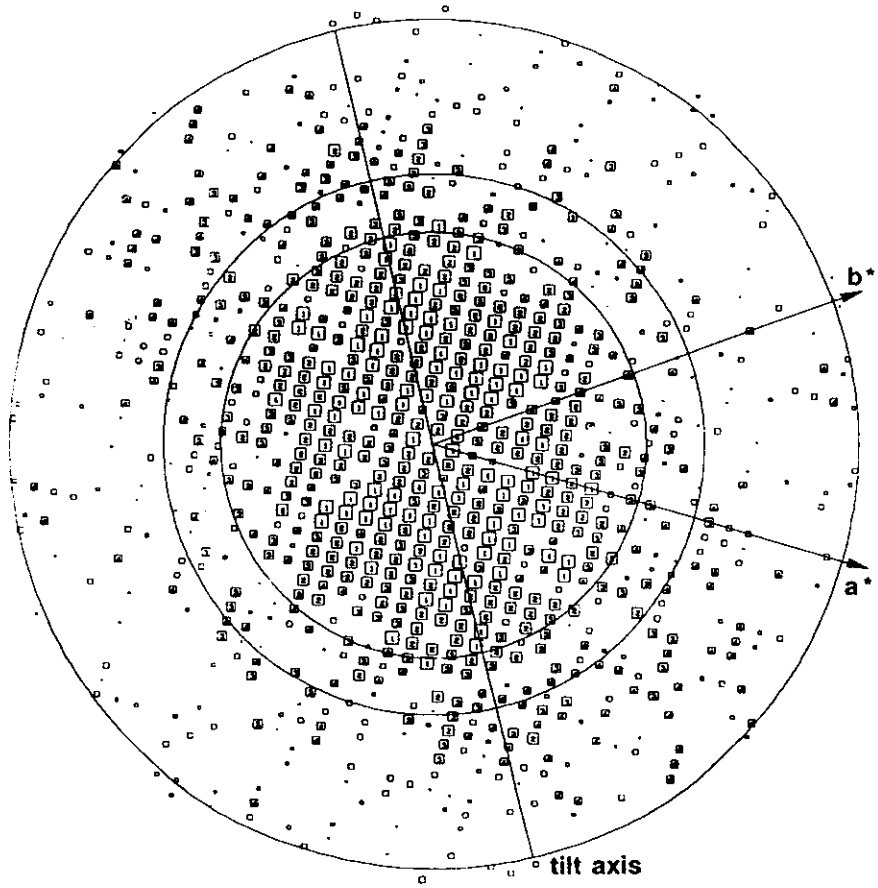

Figure 26: Plot showing the Fourier transform of an image of a 2-dimensional crystal of LHC-II tilted by 59.1°. Circles are drawn at resolutions of 7, 5.5 and 3.5 Å, respectively. (From [322], ©Macmillan Magazines Ltd. 1994, reproduced with permission)

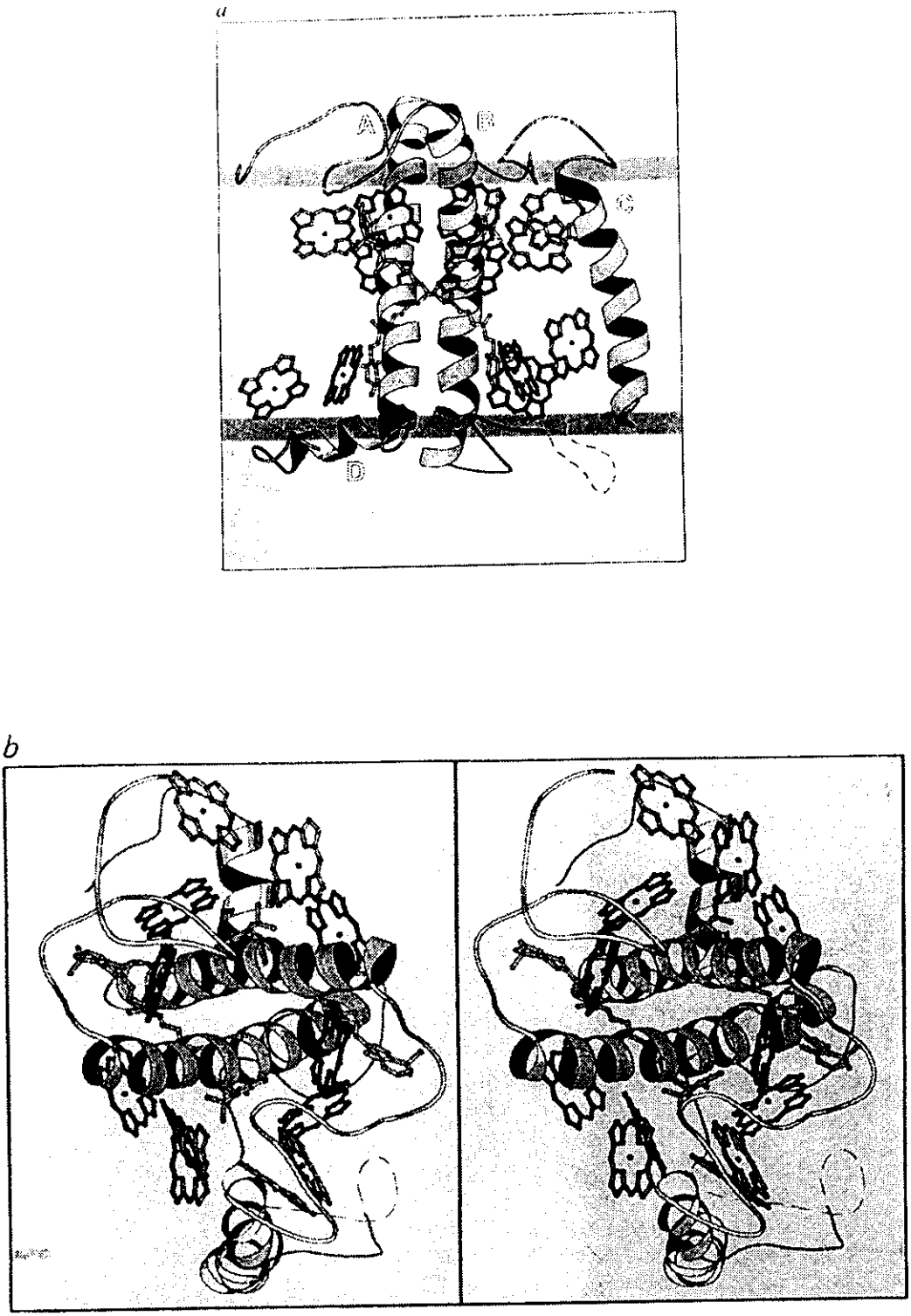

Figure 27: Overall views of LHC-II. (a) Side view. Horizontal bands indicate the approximate position of the lipid bilayer. Helices are labeled A to D. (b) Stereo diagram showing the top view of the complex from the stromal side of the membrane. The maximum thickness of the complex is 48 Å. Laterally, the longest and shortest dimensions of the monomer are 48 and 32 Å, respectively. (From [322], ©Macmillan Magazines Ltd. 1994, reproduced with permission)

The only way to collect a sufficient amount of data for a three-dimensional reconstruction of non-crystalline and asymmetric samples under low-dose conditions is to combine images from a large number of particles. Between the different techniques designed to reconstruct single particles the most widely used was developed in 1987 by Radermacher and co-workers [585] and is based on the random conical tilting data collection geometry. The only constraint imposed by this technique is that the macromolecules must tend to lay onto the grid showing one or a few preferent views. Although this requirement may seem restrictive, it is fairly usual for biological molecules to interact with the film support producing a collection of preferential views.

When a sample of single particles is tilted by an angle Θ, the imaged macromolecules present many different views that can be determined precisely. The set of views is formally equivalent to a conical tilt series with the cone angle equal to the tilt angle and the azimuthal angle set at random (Fig. 28). To calculate this angle two micrographs of each area of interest are taken. The first one, tilted Θ degrees, provides the views used for the three-dimensional reconstruction, while the second one, tilted 0 degrees (untilted), is employed to calculate the azimutal angle by cross correlation techniques [585]. That is, the untilted micrograph provides the necessary information on the actual data collection geometry that would be applied when combining the tilted views in the three-dimensional reconstruction. Implicit in the method is the presumption that all the projections derive from the same kind of three-dimensional object, or that it is possible to classify them prior to the reconstruction.

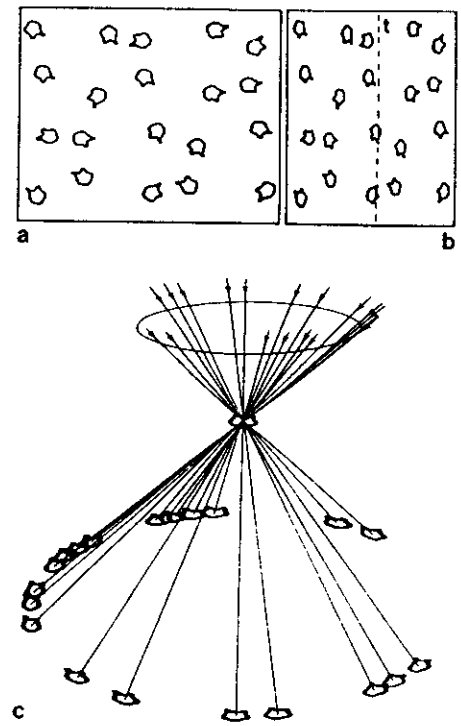

Figure 28: Illustration of the basic principle of the random conical tilting reconstruction scheme. (a) view of a specimen with randomly oriented particles lying flat in the plane of the specimen; (b) projection of the specimen in (a), tilted Θ degrees. The images which can be extracted from the tilted image (b) form the conical tilt series shown in (c) equivalent to a tilt series of a single particle with random projection directions, all lying on the surface of a cone. (From [585], ©Blackwell Science Ltd. 1987, reproduced with permission)

Once the different views are aligned, the reconstruction is usually performed in real space using backpropagation [584], ART [324], etc.

These techniques have recently been applied with success to obtain a reconstruction of the ribosome from *E. coli* at 25 Å resolution, which in turn has provided the base to propose a model for the protein synthesis process [318].

The ribosome is formed by an assembly of proteins and nucleics acids, and synthesizes proteins according to genetic instructions in all organisms. For the reconstruction, Frank and co-workers combined 4300 projections of ice embedded ribosomes, taken in an energy-filtering microscope to remove the detrimental effect of inelastic scattered electrons on the images (Fig. 29). The reconstructed volume shows the two ribosomal subunits (50S and 30S) as well defined separate masses of density, enclosing a space that can accomodate messenger RNA, transfer RNAs and associated factors during protein synthesis (Fig. 30).

Figure 29: Portion of a micrograph showing ice-embedded ribosomes in random orientations. Scale bar, 1000 Å. (From [318], ©Macmillan Magazines Ltd. 1995, reproduced with permission)

The 25 Å three-dimensional density map also reveals a channel in the small ribosomal subunit and a bifurcating tunnel in the large subunit which may constitute pathways for the incoming message and the nascent polypeptide, respectively. These new features have been combined in a three-dimensional model of the basic framework of protein synthesis (Fig. 31).

5 A complete three-dimensional reconstruction, from the very beginning

5.1 The specimen: what it is and how it is prepared.

As a detailed example of a three-dimensional reconstruction from electron microscopy data, we will describe the process followed for the reconstruction of the oligomer of the DnaB protein from *Escherichia coli* [330]. We will use the methods for single particle processing, as this specimen is neither crystalline nor symmetrical, and the software package Xmipp, developed in our group. This package is described elsewhere [324] and is available at internet address ftp.cnb.uam.es as public domain software.

DnaB is a protein isolated from the bacteria *Escherichia coli.* It has a molecular weight of about 50,000 Dalton[7], and it is known to exist in hexameric form, that is, as an oligomer weighting 300 kDalton. DnaB belongs to a family of proteins known as helicases. Its function is very important for the cell, as it unwinds the double helix of

[7]A Dalton is the biochemical term for "atomic mass unit".

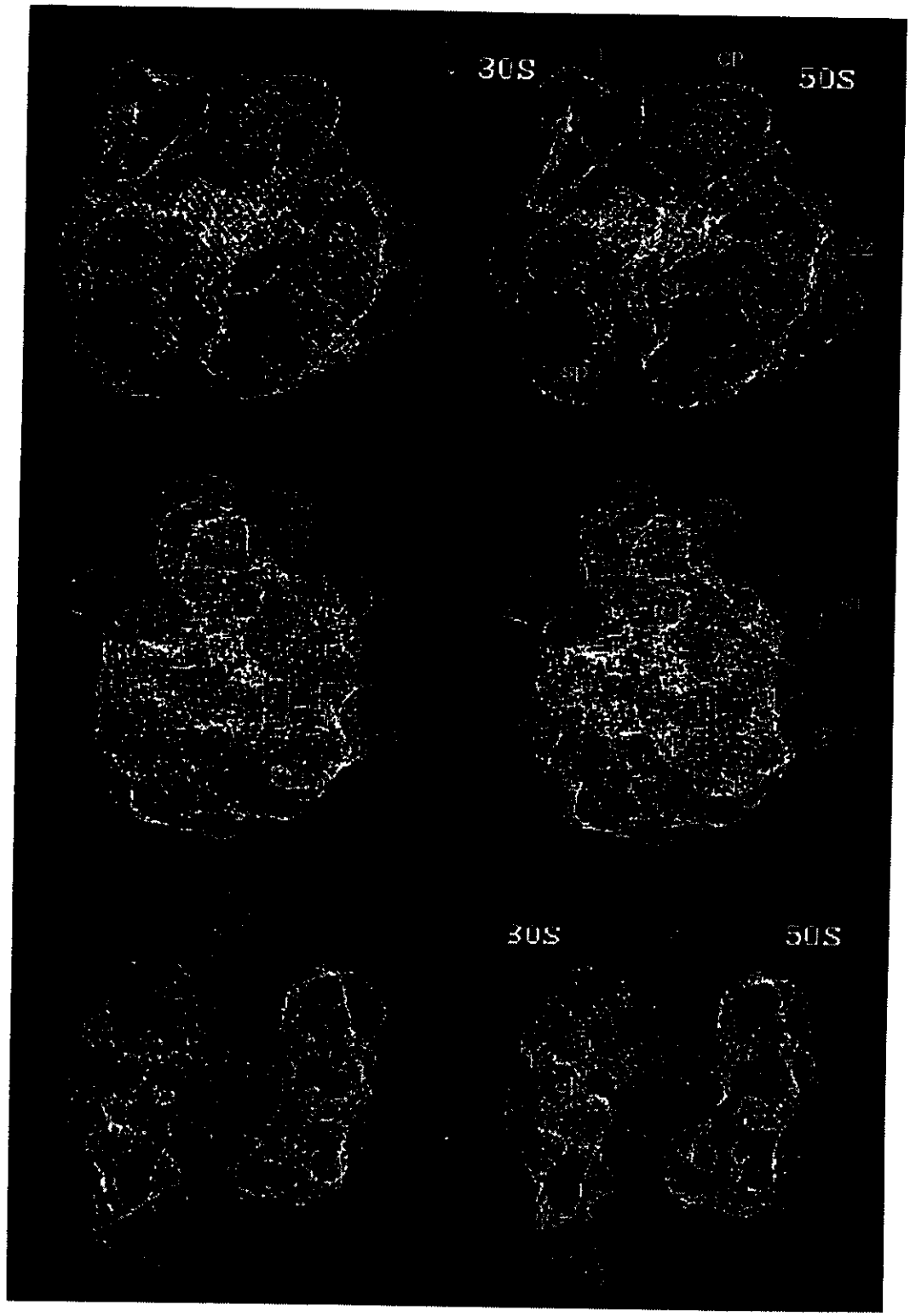

Figure 30: Three-dimensional map of the *E. coli* ribosome presented from different viewing angles. (a) Side view showing part of the tunnel through the 50S subunit. (b) View down the channel through the neck of the 30S subunit (generated from (a) by rotating 90° around the vertical axis, then 15° around the horizontal axis). (c) View down the tunnel through the 50S subunit (generated from (a) by rotating 75° around the horizontal axis and using a cutting plane to remove part of the ribosome).(From [318], ©Macmillan Magazines Ltd. 1995, reproduced with permission)

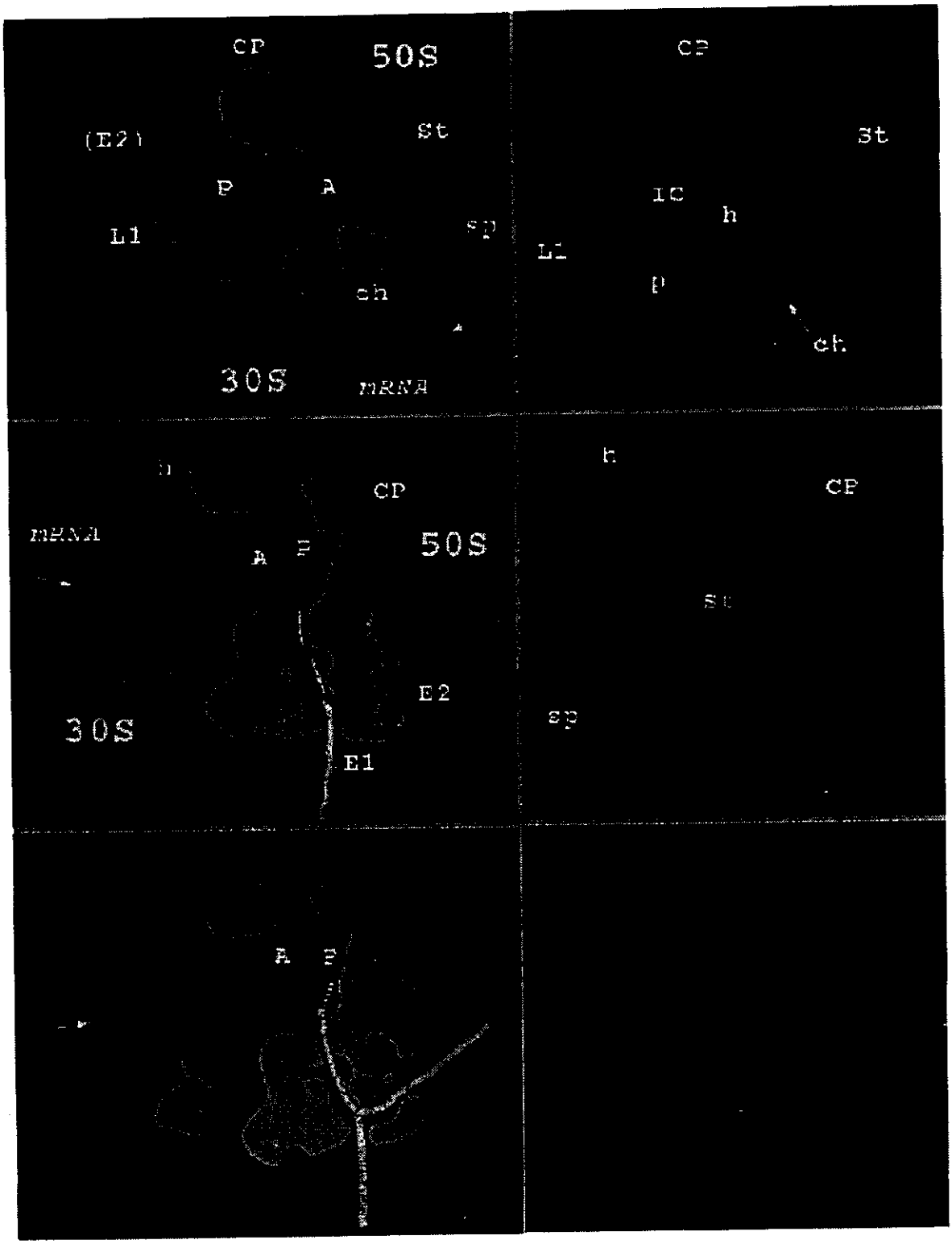

Figure 31: Model of the basic framework of protein synthesis using a surface representation of the 25 Å reconstruction and a 12.5 Å map of the X-ray resolved tRNA-tRNA complex.(From [318], ©Macmillan Magazines Ltd. 1995, reproduced with permission)

DNA when it its needed for the replication of the genetic material of the bacteria. The process of unwinding requires energy, and DnaB obtains it by catalyzing the reaction of hydrolysis of ATP. Biochemical studies have also shown that DnaB migrates in a preferred direction along one of the strands of the replicating DNA molecule. Although this, and other helicases, have been extensively studied from the biochemical point of view, little is known about its structure.

The preparation of the sample is depicted in fig. 32. In this case, staining with a heavy metal salt (namely, uranyl acetate) was chosen as a first approach to study the unknown structure. We start from a stock of highly purified protein, diluted in an appropriate buffer that ensures the preservation of structure and functionality of the molecule. A few microliters of this solution are deposited on a flat, clean surface, and an electron microscope grid coated with a thin layer of carbon (specimen support) is placed on top of the sample drop. After a short time of incubation, some of the protein particles have adsorbed onto the carbon film. The grid is blotted against a filter paper to remove the excess sample, and subsequently incubated for a few seconds on a drop of staining agent. Again the excess liquid is blotted, and after drying we have the biological material surrounded by a layer of heavy metal that prevents complete destruction in the microscope and gives contrast to the images.

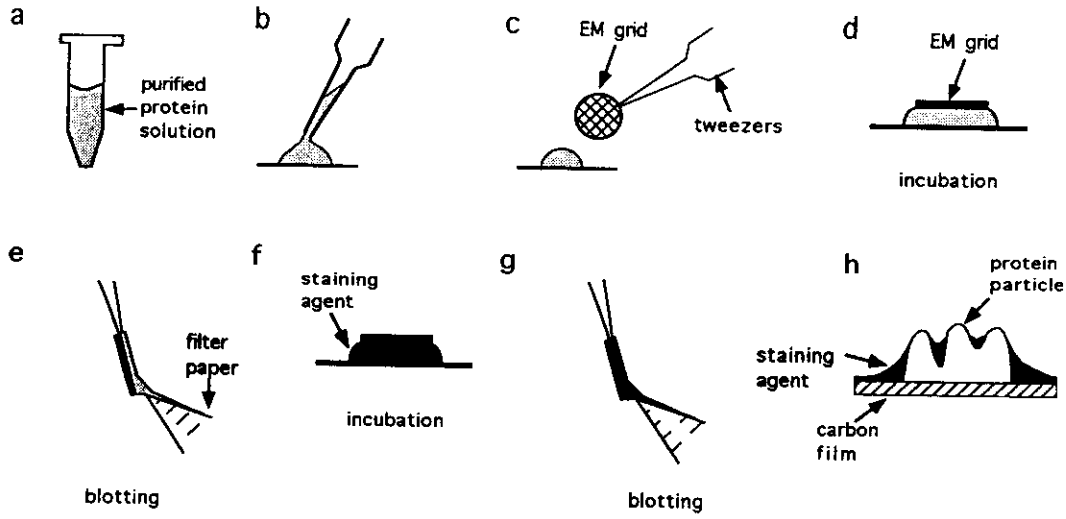

Figure 32: Preparation of a sample by negative staining. See text for details.

5.2 Image recording, digitization and particle selection

Once in the electron microscope, pairs of micrographs are taken under minimal dose conditions[8] from the same sample field, at 55 and 0 degrees of tilting (Fig. 33). The 55 degrees (tilted) images will be used as input data for the three-dimensional reconstruction, while the 0 degrees (untilted) images will provide the alignment parameters for their tilted counterparts, and valuable information on the specimen projection structure. The number of fields recorded must provide enough particle views to perform the reconstruction with a high enough signal (typically, more than a thousand).

The micrographs are digitized, following the Shannon sampling theorem [331], with a pixel size less than half the resolution expected. In this case, a 4 Å pixel was used. The

[8]See section 3

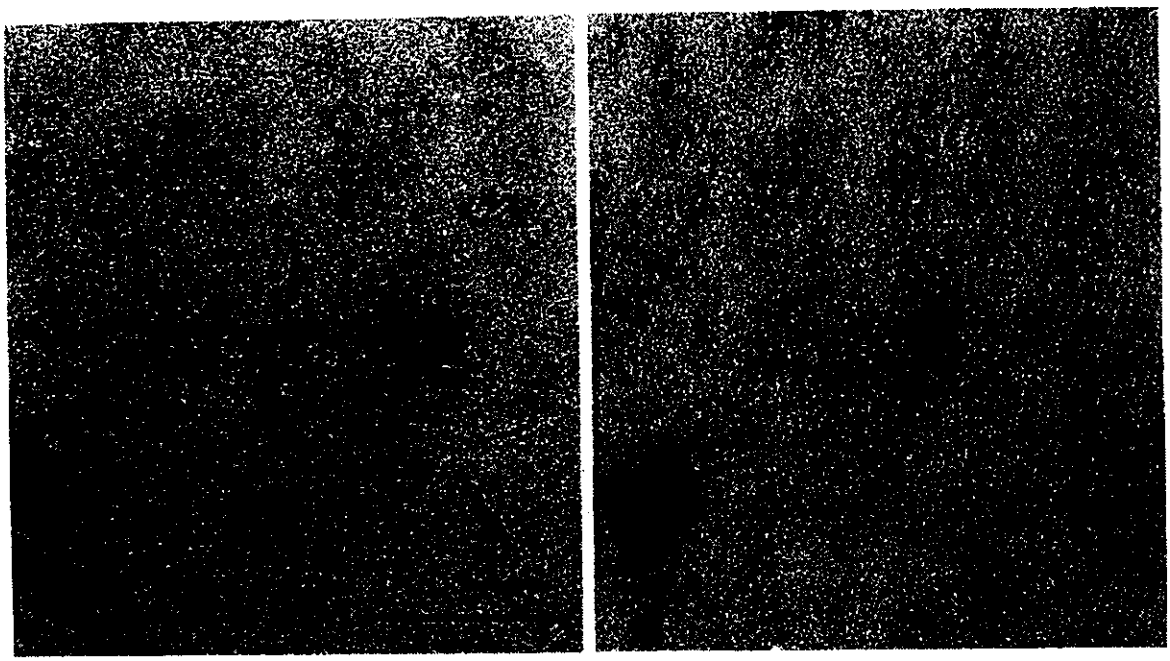

Figure 33: Tilt pair from a field of negatively stained DnaB particles. (a) Untilted view. (b) 55° tilted view. (From [330], ©Academic Press Inc. 1995, reproduced with permission)

digitized tilting pairs are displayed, and the required number of particles is obtained by interactive selection, using the criteria of apparent integrity and non overlapping with other particles. Afterwards, they are extracted from the total image in the form of 64x64 pixel frames (Fig. 34). A circular mask is applied to each image, to leave us (approximately) only with the area of interest and eliminate spurious data around it. The mask edge is smoothed to avoid artefactual high frequencies. In our example, 1373 such frames were used as input data.

Visual inspection of the gallery of particles[9] gives us a first idea of the general aspect of the aggregate, and information about the range of views that it presents in the electron microscopy preparation. In our case, it was found that DnaB showed mainly a characteristic triangle-shaped view with a stain penetrating region in the center. Scarcely more information could have been extracted without further image processing.

5.3 Two-dimensional processing

As stated in section 4.3, alignment of the untilted views is one of the key steps in single particle three-dimensional reconstruction, as it provides the parameters for determining the actual data collection geometry and allows proper centering their counterpart images in the tilted micrograph.

In a first step, a rough translational alignment is performed by cross correlating each individual image to an artificially generated circular mask. Then, the homogeneity of the image population must be checked, to avoid feeding the alignment algorithm with incoherent data. For this purpose a neural network self organizing map (S.O.M.)-based classificator [323] was used.

[9]Note that, in the negative staining images, protein masses appear white (transparent to electrons) while the background is dark (strong scattering material).

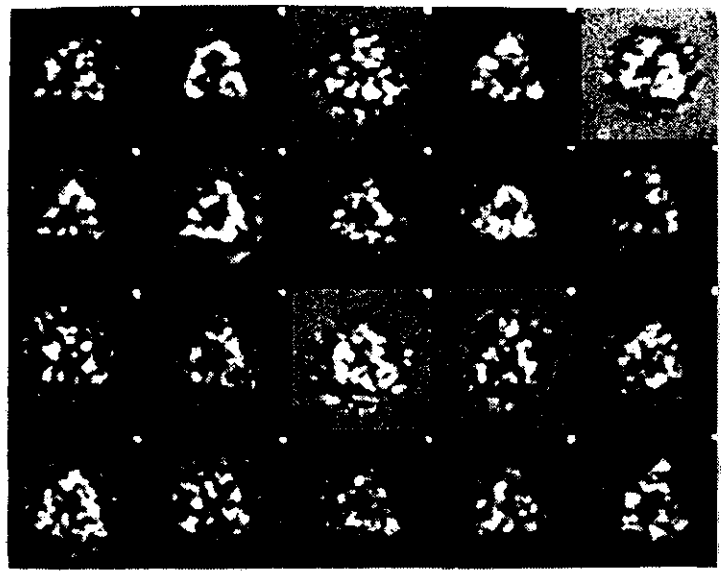

Figure 34: Gallery of views of negatively stained DnaB oligomers. The size of the subframes is 24.3 × 24.3 nm. (From [330] , ©Academic Press Inc. 1995, reproduced with permission)

This classificator maps the experimental set of images onto a two-dimensional array of nodes in such a way that vectors projected onto adjacent nodes are more similar than vectors projected onto distant ones. The output consists in a collection of images (code vectors) associated with the nodes of the map that represent the main trend of variability in the input sample. The result is displayed as a map with the meaningful code vectors located at the periphery, as can be seen in Fig. 35. For our DnaB images, the output code vectors represented all the possible different views of a rotationally unaligned set of particles with six strong and three faint density maxima. In this way it was proven that the selected population was homogeneous, its main trend of variability coming from the obvious rotational misalignment.

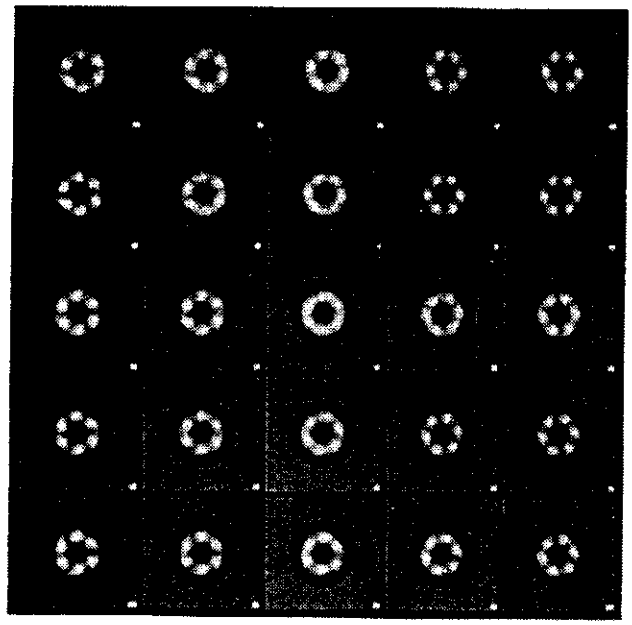

Figure 35: Self-organizing map code vectors for the DnaB population before rotational alignment. (From [330], ©Academic Press Inc. 1995, reproduced with permission)

After homogeneity assessment, the DnaB projections were translationally and ro-

tationally aligned, using cross correlation and a reference-free algorithm [326]. During this step 8% of the images were rejected as incoherent or too deviated from the average. Once aligned, the population was again searched for heterogeneities with the S.O.M. classificator, and an average projection image was obtained. At this step, the two-dimensional resolution of the average was also calculated by the spectral signal-to-noise ratio method [334], and it was found to be 1.8 nm. Fig. 36 shows the average image after low-pass filtering to the estimated reproducible resolution. Although still in two dimensions, it already contains a considerable amount of information about the structure of the protein oligomer: general shape, dimensions and internal symmetries can now be studied thanks to the enhanced signal to noise ratio. The significance of these results will be discussed in section 5.5.

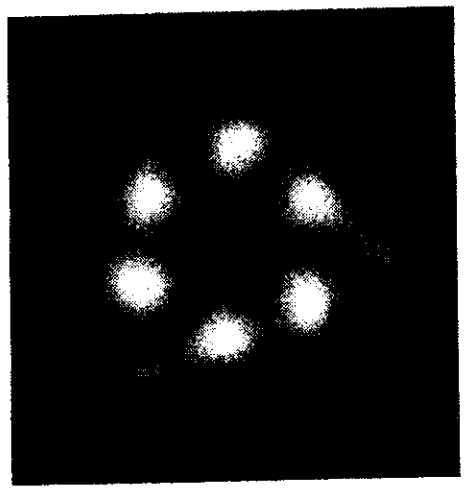

Figure 36: Average projection image of the DnaB oligomer filtered to 1.8 nm resolution. The edges of the triangle have a length of about 11 nm. (From [330], ©Academic Press Inc. 1995, reproduced with permission)

5.4 Three-dimensional reconstruction

The three-dimensional reconstruction was performed following the random conical tilting method [585] by means of a filtered backprojection algorithm. The untilted images were processed as described above, and the 1266 particles remaining after the alignment step were used to center their corresponding tilted images, while the data obtained from the rotational alignment were used along with the angular data from the tilting process to define the conical geometry. The centered tilted projections were then used to build up a three-dimensional density map.

To estimate the resolution of the reconstruction, the population was divided into two equivalent groups and their backprojected reconstructions were compared using the differential phase residual method with the cutoff set at 45° [317], which gave a resolution of 2.7 nm.

The volume obtained from the whole set of images was then low-pass filtered to the calculated resolution and visualized with AVS (Advanced Visual Systems, Inc.) (Fig. 37)

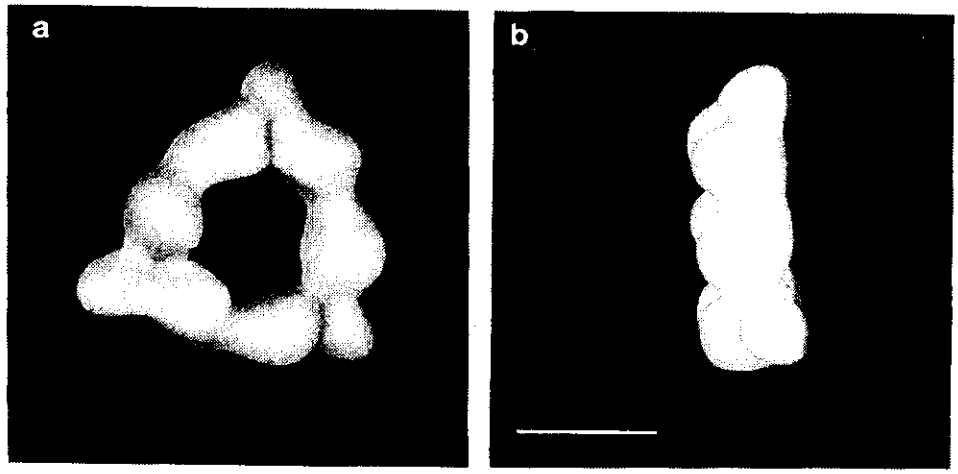

Figure 37: Surface rendering of the reconstructed DnaB hexamer. (a) Top view. (b) Side view. Bar, 5nm. (From [330], ©Academic Press Inc. 1995, reproduced with permission)

5.5 Interpretation of results

DnaB is a macromolecular aggregate of considerable importance, being currently the best studied member of a family of proteins, the helicases, present in all living organisms and playing critical roles in nucleic acids metabolism [325] In spite of this importance, and of the impressive amount of biochemical studies devoted to this subject, little is known about its structure. In fact, no atomic-resolution data are yet available for any enzyme of this type. However, it seems likely that structural knowledge would illuminate the understanding of the protein activity.

In the three-dimensional reconstruction of the oligomer of DnaB, we found several interesting features that could allow a new approach to the structure-function problem in this protein family:

a) *DnaB is a triangle-shaped particle.*- It is well established that DnaB is a hexamer of identical subunits, and a regular hexagonal arrangement was naturally expected for the oligomer. However, electron microscopy analysis shows that, apart from six large stain excluding regions, DnaB presents three small regions located in between the six larger masses. That is, we found a three-fold rather than six-fold dominant symmetry. This result points to the possibility that DnaB assembles as a trimer of dimers, a fact that no biochemical experiments have found until now.

b) *There is a channel that completely traverses the center of the molecule.*- This channel, running from side to side of the particle, and fully open at both ends, could be occupied by the strand of DNA to which the protein is bound during the unwinding process. Again, there is no biochemical prove as to how the helicase is loaded onto the DNA, but threading through a channel is a rather common system in DNA binding proteins.

c) *The two faces of the enzyme are different, producing a structural polarity.*- As seen in Fig. 37, there are six inner lobules running from side to side of the reconstructed particle, while the other three, less massive, appear only in one of the sides. This feature is particularly relevant given the unidirectional migration of the DnaB hexamer along the strand of DNA during helicase action. The indications of

structural polarity could provide a simple mechanism for a defined directional interaction of the enzyme with its substrate.

Based on these first structural data about DnaB, we have been able to propose a model describing the arrangement of the monomers in the hexamer. The use of cryotechniques, new electron imaging procedures and improved image processing algorithms could eventually allow us to asses the model validity, to locate the active sites of the enzyme in each subunit, and in a subsequent stage to describe in more detail the secondary structure of the protein.

Acknowledgments

We are grateful to the authors who generously supplied the artwork from their published and unpublished results. This work has been supported by Grants BIO95-0768 and PB 91-0910 from Comisión Interministerial de Ciencia y Tecnología (CICYT), Spain, and the Human Capital Mobility Program from the E.C.

Chapter III

[311] L. A. Amos, R. Henderson, and P. N. T. Unwin, Three-dimensional structure determination by electron microscopy of two-dimensional crystals, *Prog. Biophys. Mol. Biol.* **39**, 1982, 183–231.

[312] W. Chiu, K. H. Downing, J. Dubochet, R. M. Glaeser, H. G. Heide, E. Knapek, D. A. Kopf, M. K. Lamvik, J. Lepault, J. D. Robertson, E. Zeitler, and F. Zemlim, Cryoprotection in electron microscopy. *J. Microsc. (Oxford)* **141**, 1986, 385–391.

[313] R. A. Crowther, N. A. Kiselev, B. Böttcher, J. A. Berryman, G. P. Borisova, V. Ose, and P. Pumpens, Three-dimensional structure of the Hepatitis B Virus core particles determined by electron cryomicroscopy. *Cell*, **77**, 1994, 943–950.

[314] R. A. Crowther, L. A. Amos, J. T. Finch, D. J. DeRosier, and A. Klug, Three dimensional reconstruction of spherical viruses by Fourier synthesis from electron micrographs. *Nature* **226**, 1970, 421–425

[315] D. J. DeRosier, and A. Klug, Reconstruction of three-dimensional structures from electrom micrographs. *Nature* **217**, 1968, 130–134.

[316] D. J. DeRosier, and P. B. Moore, Reconstruction of three-dimensional images from electron micrographs of structures with helical symmetry. *J.Mol.Biol.* **52**, 1970, 355–369.

[317] J. Frank, M. Radermacher, T. Wagenknecht, and A. Verschoor,. Studying ribosome structure by electron microscopy and computer image processing. *Methods Enzymol.* **164**, 1988, 3–35.

[318] J. Frank, J. Zhu, P. Penczek, Y. Li, S. Srivastava, A. Verschoor, M. Radermacher, R. Grassucci, R.K. Lata, and R.K. Agrawal A model of protein synthesis based on cryo-electron microscopy of the *E. coli* ribosome. *Nature* **376**, 1995, 441–444.

[319] G.R. Grinton and J.M. Cowley, Phase and amplitude contrast in electron micrographs of biological material. *Optik* **34**, 1971, 221–233.

[320] P.T. Hawkes, The electron microscope as a structure projector. Electron Tomography, ed. by J. Frank, Plenum Press , 1992.

[321] R. Henderson and P.N.T. Unwin, Three-dimensional model of purple membrana obtained by electron microscopy. *Nature* **257**, 1975, 28–32.

[322] W. Kühlbrandt, D.N. Wang, and Y. Fujiyoshi, Atomic model of plant light-haversting complex by electron crystallography. *Nature* **367**, 1994, 614–617.

[323] R. Marabini and J.M. Carazo, Pattern recognition and classification of images of biological macromolecules using artificial neural networks. *Biophys. J.* **66**, 1994, 1804–1814.

[324] R. Marabini, I.M. Masegosa, M.C. San Martin, S. Marco, J.J. Fernandez, L.G. de la Fraga, C. Vaquerizo, and J.M. Carazo, Xmipp: an image processing package for electron microscopy. *J. Struct. Biol.*, 1995.

[325] S.W. Matson, D.W. Bean, and J.W. George, DNA helicases: Ezymes with essential roles in all aspects of DNA metabolism. *BioEssays* **16**, 1994, 13–22.

[326] P. Penczek, M. Radermacher, and J. Frank, Three-dimensional reconstruction of single particles embedded in ice. *Ultramicroscopy* **40**, 1992, 33–53.

[327] M. Radermacher,Weighted back-projection methods. Electron Tomography, ed. by J. Frank, Plenum Press, 1992.

[328] M. Radermacher, A. Wagenknecht, A. Verschoor, and J. Frank, Three-dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50S ribosomal subunit of *Escherichia coli*. *J.Microsc.* **146**, 1987, 113–136.

[329] M. Radermacher, *Dreidimensionale rekonstruktion bei kegelförmiger kippung im elektronenmikroskop*. Ph. D. thesis, Technische Universität München, 1980.

[330] M.C. San Martin, N.P.J. Stamford, N. Dammerova, N.E. Dixon, and J.M. Carazo, A structural model for the *Escherichia coli* DnaB helicase based on electron microscopy data. *J. Struct. Biol.* **114**, 1995, 167–176.

[331] C.E. Shannon, Communication in the presence of noise. *Proc. IRE* **37**, 1949, 10–22.

[332] J.C.H. Spence, *Experimental High-resolution Electron Microscopy*, 2nd ed. Oxford University Press, New York, 1988.

[333] K.A. Taylor and R.M. Glaeser, Electron microscopy of frozen hydrated biological specimens. *J. Ultrastuct. Res.* **55**, 1976, 448–456.

[334] M. Unser, B.L. Trus, and A.C. Steven, A new resolution criterion based on spectral signal-to-noise ratios. *Ultramicroscopy* **23**, 1987, 39–52.