[Jose-Jesus Fernandez, Carlos Oscar S. Sorzano, Roberto Marabini, and Jose-Maria Carazo]

[Strategies for structural determination of biological specimens]

© DIGITAL VISION, SMALL CIRCLE IMAGES FROM TOP TO BOTTOM: L. SOUSTELLE, C. JACQUES, AND A. GIANGRANDE, IGBMC, ILLKIRCH, FRANCE; G. SCHERRER, P. TRYOEN-TOTH, AND B. L. KIEFFER, IGBMC, ILLKIRCH, FRANCE; L. MCMAHON, J.-L. VONESCH, AND M. LABOUESSE, IGBMC, ILLKIRCH, FRANCE.

Image Processing and 3-D Reconstruction in Electron Microscopy

Knowledge of the structure of biological specimens is critical to understanding their functions at all scales [1] and is crucial in biosciences to complement biochemical studies. Electron microscopy (EM) enables the investigation of the structure of biological specimens over a wide range of sizes, from cellular structures to single macromolecules, providing information at low/medium/high resolution [2], [3] and, in some cases, at atomic resolution [4]. The combination of EM with other high-resolution approaches, e.g., X-ray crystallography or nuclear magnetic resonance (NMR), allows integration of the structural information gathered at multiple levels of the biological complexity into a common framework, which is expected to provide a comprehensive description of the cellular function in molecular detail [1].

This article presents an overview of the different strategies for structural determination of biological specimens by EM. The basic principles of the methodology are presented first followed by a description of the main approaches. This article particularly focuses on high-resolution structural analyses of macromolecules. In [31], Leis et al. focus on the elucidation of complex subcellular structures.

1053-5888/06/$20.00©2006IEEE

PRINCIPLES OF STRUCTURAL DETERMINATION BY EM

IMAGE FORMATION IN THE ELECTRON MICROSCOPE

The operational principles of the transmission electron microscope (TEM) are similar to the widely known light microscope (Figure 1). The source of illumination is a filament (cathode) that emits electrons at the top of a cylindrical column about two meters high. Electrons are then accelerated by a nearby anode, forming an electron beam that travels down a vacuum to pass through the specimen. Scattered and unscattered electrons emerging from the specimen are then collected by magnetic lenses and focused to form an interference pattern, which constitutes the image. Images taken from TEMs can be considered as two-dimensional (2-D) projections of the specimen. By tilting the specimen in the TEM, different views can be obtained, which is the key for subsequent tomographic reconstruction processes.

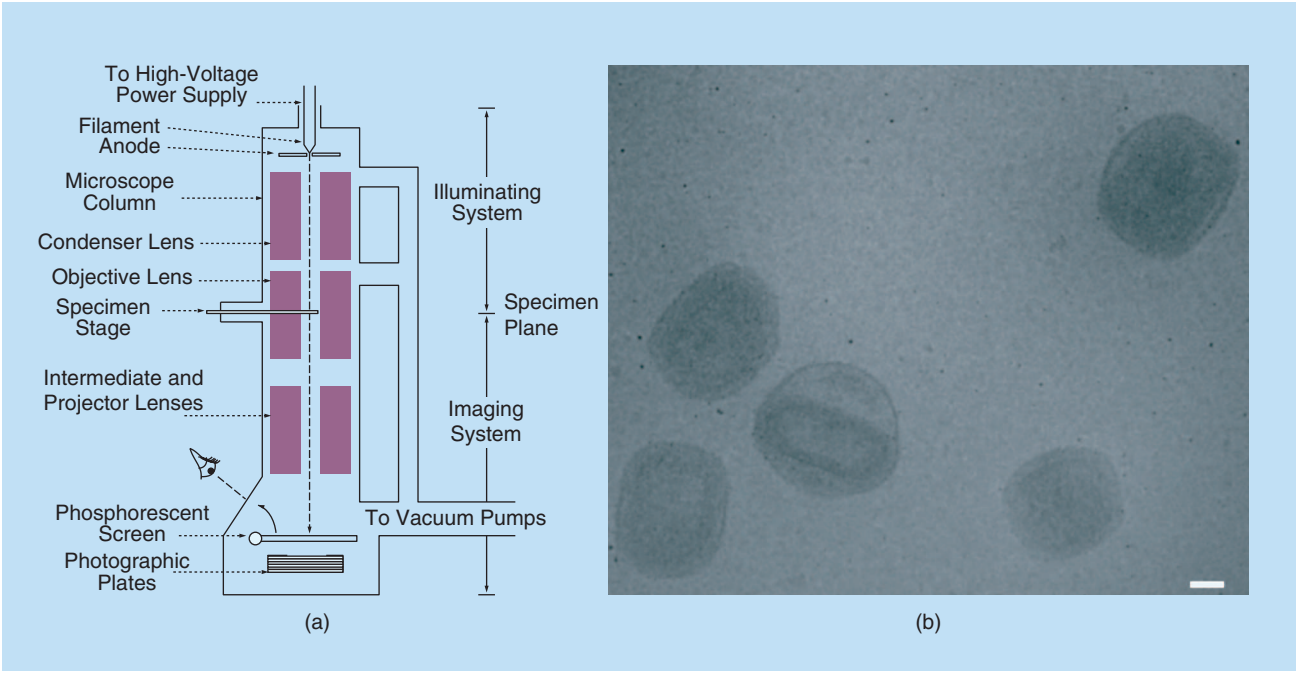
EM images are affected by the contrast transfer function (CTF) of the TEM, which arises from the aberrations of the lenses and from the defocus used in imaging. The CTF introduces spatial frequency-dependent oscillations into the Fourier space representation of the image. These result in contrast changes and modulation of the spectrum amplitudes, as well as an additional dampening envelope that attenuates high-resolution information (Figure 2). Estimation of the CTF and correction for its effects are thus essential for any image to faithfully represent a projection of the specimen. Restoration of the phases is a central component of the correction, as it ensures that contrast is consistent at all spatial frequencies.

The theoretical resolving power in a TEM is imposed by the electron wavelength, which depends on the accelerating voltage. In practice, the effective resolution limit ranks around 4–10 Å because of the lens aberrations, specimen preparation techniques, sample thickness, low contrast, radiation damage, etc. On the other hand, TEMs present a technical limitation to the maximum tilt angle, usually around $\pm 70^{\circ}$, which may cause anisotropic resolution in the tomographic reconstruction, as described below.

PRESERVATION OF SAMPLES UNDER PHYSIOLOGICAL CONDITIONS

Specimens have to be specially prepared prior to the electron exposure since the vacuum conditions in the TEM and electron radiation may degrade the biological structure. The two main preparation techniques are negative staining and cryomicroscopy [2]. Negative staining consists of covering the sample with a contrast agent that protects the biological material from the electron exposure and enhances the contrast in the acquired images. This technique provides structural information only about the surface of the sample; additionally, the stain produces artifacts in the images and the information is thus limited to medium resolution (around 20 Å). For these reasons, negative stain is not used for high-resolution studies; however, it is very useful for the earlier stages of research to get a first glimpse of the structure.

Currently, cryomicroscopy is the most common technique used to prepare biological samples for imaging in TEM [6]. This technique keeps the sample frozen in vitreous ice at cryogenic temperatures, which ensures preservation in near-physiological conditions during electron exposure. As contrast agents are strictly avoided, cryomicroscopy images have structural information about the whole sample, not only its surface. In principle, this technique does not imply any direct limitation on the achievable resolution.



[FIG1] (a) Schematic representation of a TEM and (b) an image acquired from a field of Vaccinia viruses (bar: 100 nm).

Biological material is very sensitive to radiation. Therefore, electron doses must be kept very low (around 5–20 $e^{-}/Å^{2}$) so as to minimize radiation damage and preserve as much resolution as possible. The combination of low dosage and weak contrast makes cryomicroscopy images extremely noisy (Figure 1). In high-resolution structural studies, this poor signal-to-noise ratio (SNR) is substantially increased by combining very large sets of images of the same specimen.

PRINCIPLES OF TOMOGRAPHIC RECONSTRUCTION

TEMs provide 2-D projection images where structural features from different layers of the three-dimensional (3-D) structure of the specimen are superposed along the direction of the electron beam. The 3-D structure can then be derived from a set of views of the specimen taken at different orientations. In general, those views are obtained either by tilting the specimen in the TEM or by collecting different occurrences of the specimen from several TEM images. The mathematical principles of tomographic reconstruction are based upon the central section theorem [2]. This theorem states that the Fourier transform (FT) of a 2-D projection of a 3-D object is a central section of the 3-D FT of the object (Figure 3). Therefore, by collecting projections from the specimen at different orientations, its 3-D FT can be sampled, and a simple Fourier inversion yields the structure of the specimen.

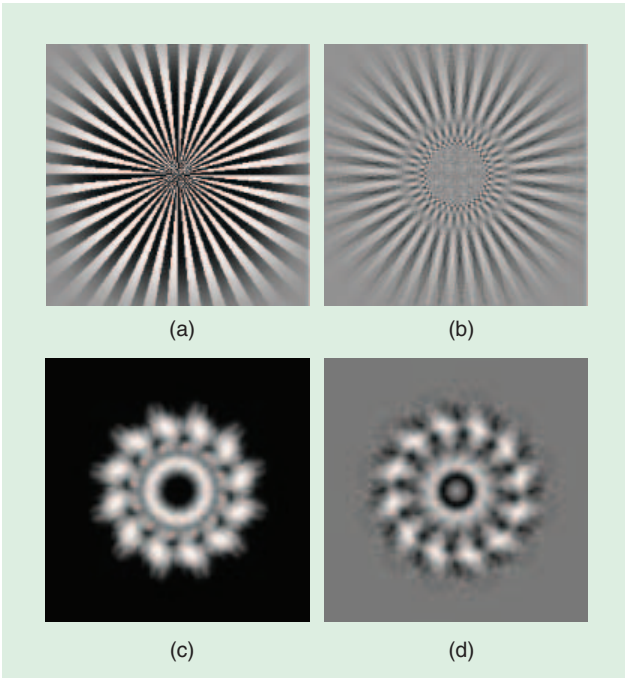
The limited tilt range of TEMs may preclude obtaining all possible views from the object. As a consequence, there may be a missing region in the Fourier space where data of the 3-D FT of the object are not available. This missing region implies that the reconstructed volume has anisotropic resolution. In other words, the 3-D structure may be affected by elongation and other distortions.

IMAGE PROCESSING AND 3-D RECONSTRUCTION IN EM

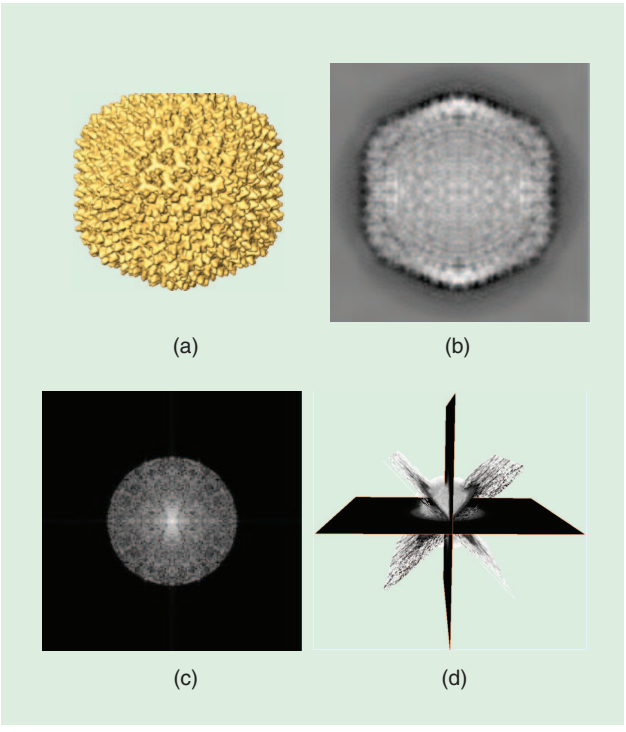
The general problem in structural determination by EM is thus the 3-D reconstruction of a biological structure from a finite set of 2-D projection images over a restricted angular range [2]. From an image processing point of view, there are several main problems with addressing such structural analyses: the extremely low SNR of the images, CTF determination and correction for its effects, alignment and classification of the images, 3-D reconstruction under limited tilt angle conditions, and, once the structure is obtained, possible post-processing and interpretation of the results.

APPROACHES FOR STRUCTURE DETERMINATION BY EM

Different approaches to data collection, image processing, and 3-D reconstruction are used depending on the nature of the specimen and structural information sought. For specimens in



[FIG2] Effects of the CTF. (a) Original synthetic image, with sampling rate 2.18 Å/pixel. (b) Image as viewed through a 200-kV TEM, assuming 1.5 μm underfocus. Note that there are some contrast changes across the image due to the phase changes of the CTF. In addition, some details vanish because of the CTF modulation. (c) A projection of the T7 bacteriophage connector computed from its 3-D structure [5] and (d) as would be viewed through the TEM under the same conditions as in (b).



[FIG3] Central section theorem. (a) The 3-D object. (b) One projection of the object taken at a 0° tilt. (c) The FT of (b). (d) The FT of a 2-D projection of a 3-D object is a central section of the 3-D FT of the object. Here the intersection of four central sections corresponding to four different projections (0°, ∼45°, 90°, ∼135° tilt angles respect to an axis) of the object is shown.

the macromolecular domain, specific methodologies have been devised: for 2-D crystalline protein arrays (approach known as electron crystallography (EC) [7]); for specimens assembled in helical structures [8]; for isolated, rather asymmetric specimens (the "single particles" approach [2]); and for highly-symmetrical specimens, such as icosahedral viruses [9]. For complex specimens in the cellular range, the so-called "electron tomography" (ET) [3] approach is used. Thanks to these strategies, EM is able to derive structural information from different biological specimens at a wide spectrum of resolutions, from 40–100 Å (ET) to 6–30 Å (single particles, helical, icosahedral, EC), or even atomic resolution (EC, helical).

KNOWLEDGE OF THE STRUCTURE OF BIOLOGICAL SPECIMENS IS CRITICAL TO UNDERSTANDING THEIR FUNCTIONS AT ALL SCALES.

In the following, we present an overview of the different EM approaches, with special emphasis on the methodologies for structural studies of macromolecular assemblies. For brevity, we mainly focus on the EC and single-particle approaches since helical and icosahedral structures may be considered particular cases of crystals and single particles, respectively.

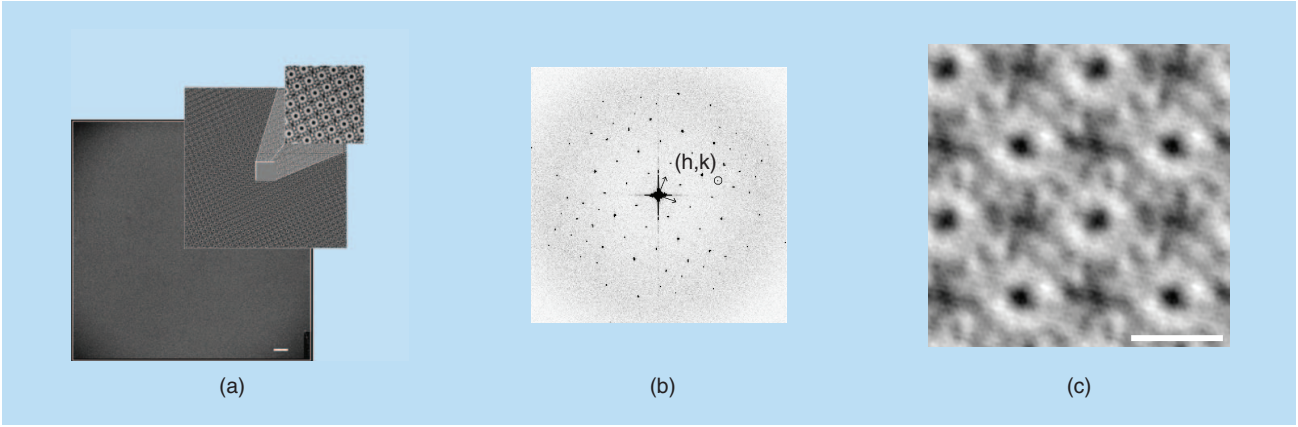
EC: SOLVING STRUCTURES AT ATOMIC RESOLUTION BY EM

EC is currently the only way to reach atomic resolution using EM [4]. In EC, the specimens are arranged in large 2-D crystalline arrays, one molecule thick, where thousands of molecules of the same structure are periodically repeated with the same orientation, such as the crystals produced by membrane proteins in their native environment. The 2-D crystals are formed by an ordered repetition of identical objects, the so-called unit cells. They can thus be conceived as a 2-D lattice (the real-space lattice) where a unit cell is reproduced at each lattice point. This is mathematically described as the convolution of the unit cell with the lattice. Since the dual of a real-space 2-D lattice is another 2-D lattice in the Fourier space (the reciprocal lattice) [10], the FT of a regular 2-D crystal is confined to a relatively small number of points with nonzero amplitude (Figure 4). Points on the reciprocal lattice are usually referred to as reflections or spots. The value of the FT at those spots depends on the structure of the unit cell. The vectors that describe the real-space and reciprocal lattices are called the crystal vectors and reciprocal vectors, respectively. The spots can be easily located in the FT by the reciprocal vectors and a pair of indices (h, k). On the other hand, the contaminating noise present in images is spread across the entire spectrum.

In EC, Fourier space is preferred for image processing tasks since the signal is concentrated on a small discrete number of spots in the FT. These tasks are mainly conceived to extract the structural information from the spots in the FT and reconstruct the structure by Fourier synthesis. This process mathematically corresponds to superposition of all the different repeating unit cells present in the image, whereby an average view of the structure is obtained. The standard methodology for 3-D reconstruction using 2-D crystals consists of the following stages: SNR enhancement by filtering, correction of the lattice defects, and Fourier synthesis [7].

FILTRATION IN FOURIER SPACE

The 3-D reconstruction process starts by acquiring projections of the crystal at various angles. The 2-D projections obtained in the TEM need to be enhanced to increase their SNR. This is done by calculating their FTs and setting to zero all those frequencies not related with the periodic signal components. In practice, all the Fourier components not in a vicinity of the spots of the reciprocal lattice are forced to have zero value. An average unit cell of the crystal can then be obtained by Fourier synthesis from the discrete set of spots. Figure 4 illustrates the filtering process of a 2-D crystal of the connector of the bacteriophage $\phi 29$ [11], a protein



[FIG4] (a) Original extremely noisy image (6000 × 6000 pixels) containing about 4,000 unit cells (bar size 100 nm and sampling rate 2.5 Å/pixel). The insert shows the result from filtering, and a zoom over a filtered area. (b) The FT of the image up to 2-nm resolution. Reciprocal lattice vectors are shown and a spot is highlighted. (c) Average unit cell resulting from Fourier synthesis. The 2 × 2 unit cells are shown up to 2-nm resolution (bar: 10 nm).

involved in packaging the DNA into the virus. In this example, the unit cell is composed of two connectors (the spot-like structures in the right-hand image) [11].

In EC, there are only 17 possible lattices (symmetry groups) in which the unit cells can be arranged to form a 2-D crystal. The simplest one is called P1, in which there is no relationship between the unit cells except a translation. Other groups contain two, three, four, five, or six symmetry axes, mirror lines, or glide symmetries. The knowledge of the crystal symmetry group is important, since it places extra constraints on the values of the FT. For example, unit cells with rotational symmetry give rise to an FT with the same symmetry. There may be cases where the relation derived from a symmetry operator causes some points of the reciprocal lattice to have a zero amplitude. This phenomenon is called systematic absences.

After filtration in Fourier space, CTF correction is carried out, as will be described for single particles, so that the discrete set of Fourier components have the correct contrast for further processing.

THE OPERATIONAL PRINCIPLES OF THE TRANSMISSION ELECTRON MICROSCOPE (TEM) ARE SIMILAR TO THE WIDELY KNOWN LIGHT MICROSCOPE.

CORRECTION OF LATTICE DEFECTS

Biological crystals are seldom perfect: they present distortions and stretching, are limited in extent, and are disordered. In addition, the lenses of the TEM introduce aberrations that are not noticeable when dealing with other specimens, but that arise when large objects (whole crystals) are recorded [12].

These effects degrade the FT and make it difficult to obtain high-resolution information. The degradation of the FT produces a broadening of the peaks in the reciprocal lattice (i.e., they are no longer a delta function but a Gaussian) plus an attenuation of the high-frequency terms. Since filtering (in Fourier space) is equivalent to averaging all the unit cells in the crystal (in real space), this degradation makes the average unit cell blurred. In other words, crystal imperfections translate into small misalignments of the different unit cells and the final average then turns out blurred.

Lattice unbending corrects these distortions and recovers high-resolution information [12]. This method identifies the position of each unit cell (usually by cross-correlation) and interpolates a new image with smaller distortion. Another approach consists of convolving the averaged unit cell to the ideal grid. The lattice is further refined, usually through a few rounds of "unbending." Figure 5 shows the distortion maps of a crystal before and after unbending. A distortion map draws the distance between the position of the unit cell in a perfect crystal and the position in the experimental one. Figure 5(c) shows the improvement on the average unit cell obtained after unbending, compared to that shown in Figure 4(c).

FOURIER SYNTHESIS: MERGING OF THE 2-D DATA TO OBTAIN A 3-D RECONSTRUCTION

As 2-D crystals are one unit cell thick, the crystal 3-D FT is not a collection of spots, but a collection of continuous lines parallel to the z-axis (that is, perpendicular to the crystal plane). These lines are called lattice lines and their intersections with the central sections are the spots whose values are given by the projection FTs. The values of the FT along the lattice lines change relatively smoothly, depending on the thickness of the unit cell. So, if the number of different projections is high enough, the specimen's 3-D FT can be recovered, and hence its 3-D structure by Fourier inversion (see Figure 6 for details).

The different images obtained in the TEM need to be normalized before they can be combined to obtain a 3-D structure. First, a common origin (that is, how far the center of the unit cell is from the origin of coordinates) should be determined. Second, the contrast, a multiplicative factor that ensures that the magnitude of all common spots in the different images is the same, must be determined.

[FIG5] Distortion map (a) before and (b) after applying the unbending procedure. (c) Average unit cell after unbending obtained by Fourier synthesis (bar: 10 nm).

The traditional approach combines the normalized projections in Fourier space to obtain a nonuniformly sampled 3-D FT. From this, a uniformly sampled 3-D FT is obtained by one-dimensional (1-D) interpolation along the lattice lines (Figure 6). The interpolation is made by a least-squares algorithm [13]. It is recommended that the experimental data be obtained at 1/3 or 1/4 of the desired sampling rate in Fourier space, which is the inverse of the specimen thickness.

An alternative, equivalent approach for 3-D reconstruction works in real space using an iterative algorithm [14]. Its main advantages are that it avoids interpolation in Fourier space which often generates artifacts (especially when the number of projections is small) and it easily allows the introduction of additional spatial constraints as well as a model of the TEM image formation process.

DETERMINATION OF HELICAL STRUCTURES

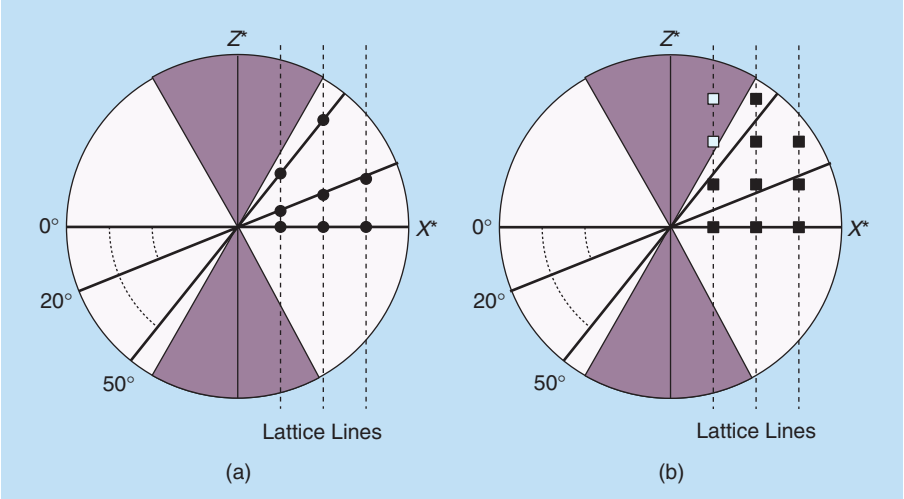
There are specimens that are naturally assembled into helical structures: DNA, filamentous viruses, cytoskeletal and muscle filaments, bacterial flagella, etc. These structures are constructed from periodically spaced, ordered subunits that follow a helical geometry. The particular properties of helical objects make it relatively straightforward to obtain a 3-D model of the subunit. The FT of a view of a helical object consists of a discrete series of lines (known as layer lines) arranged symmetrically about the origin. The position of the layer lines and the distribution of intensity along them depends on the helical symmetry of the object and the structure of its subunits. Furthermore, there are simple rules that govern the appearance of the FT of a helix and allow analytical estimation of the FT of its different views. As a consequence, a single view may provide sufficient information to reconstruct the 3-D helical object up to a certain resolution by Fourier inversion [8].

The basic steps involved in image processing and 3-D reconstruction of helical specimens include: boxing of the image to extract small segments of the specimen, determination of the helical symmetry parameters by indexing the layer lines in the FT, use of spline-fitting procedures to straighten the segments, correction for the CTF, and computation of the 3-D structure by Fourier-Bessel inversion [8]. A refinement of the helical parameters can then be done iteratively from the model to obtain a high-resolution 3-D structure. Recently, this approach has succeeded in reaching atomic resolution (reviewed in [4]).

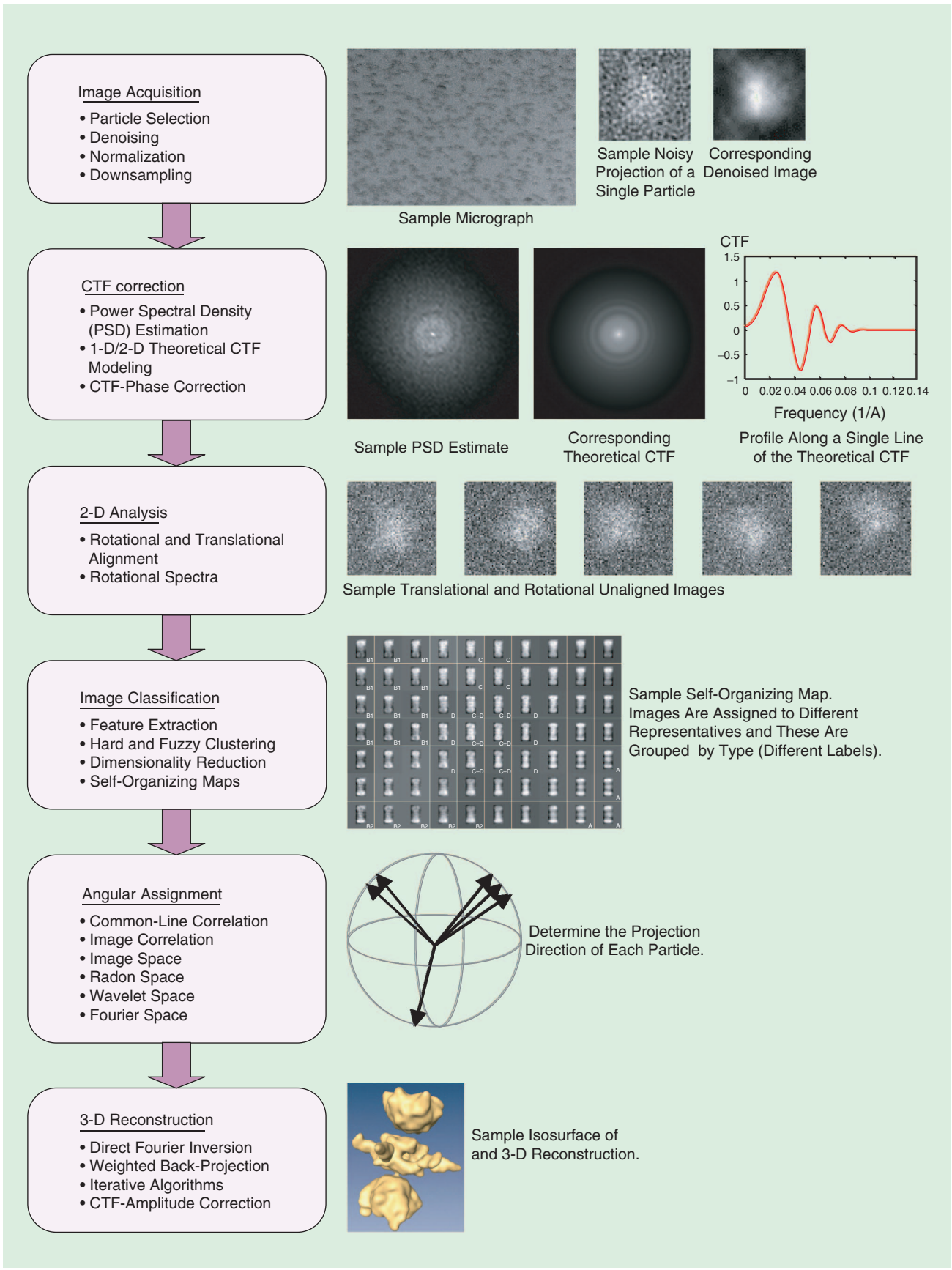
SINGLE PARTICLES: HIGH-RESOLUTION STRUCTURE DETERMINATION OF LOW SYMMETRY MACROMOLECULAR ASSEMBLIES

In the attempt to reconstruct a biological macromolecule at high resolution by EM, we need several projections that effectively cover the 3-D Fourier space, and the electron dose must be very low. Until the mid-1980s, these requirements were only satisfied when the specimen was a 2-D crystal or a highly symmetrical particle. In these two cases, imaging with extremely low electron doses is possible because the SNR can be enhanced by averaging the repeated structure or by using the particle symmetry properties. These two methods continue to give the best-quality reconstructions at present, but unfortunately many biologically interesting macromolecules cannot be crystallized and are not symmetrical.

The only way to collect a sufficient amount of data for a 3-D reconstruction of noncrystalline and asymmetric samples under low-dose conditions is to combine images from a large number of particles. The key to performing 3-D reconstructions out of this data is to correctly identify particle projections on the images and accurately determine their relative orientations. In doing so, we must guarantee that multiple copies of the protein are structurally identical. Figure 7 summarizes the main image processing steps required to go from the images to the 3-D reconstruction (these steps will be described in the sequel). Several software packages exist that integrate all the required algorithms (e.g., SPIDER, Imagic, EMAN, Xmipp) [15]. The maximum resolution achieved so far is between 6–10 Å [4]. The reader interested in the topic is referred to Frank's book [2], a comprehensive review of the field.



[FIG6] Interpolation and sampling along the lattice lines. The drawings represent Fourier space perpendicular to Y^* axis. (a) The central sections corresponding to three images obtained at tilt angles 0°, 20°, and 50° respectively about the Y^* axis. Each central section roughly gives a sample per lattice line. The problem consists of obtaining a periodic sampling along each lattice line from the experimental nonuniform sampling. The shadowed region represents the area in Fourier space where data are not available (due to the technical tilt limitations of TEMs). The circle represents a limit on the resolution. (b) Sampling points after interpolation of the lattice lines and uniform sampling. Black squares represent samples computed by solving the interpolation problem, while white squares are values estimated by extrapolation.



[FIG7] Schematic workflow of the image processing steps followed in EM of single particles.

IMAGE ACQUISITION

Images are acquired by two means: digitization of the image recorded on a film or with a CCD camera placed inside the TEM. The modulation transfer function (MTF) of any of these devices is usually negligible compared to the inaccuracies introduced by the sample preparation, the high noise levels, and the TEM aberrations. On the other hand, illumination conditions may greatly differ among different images, and even among different regions in a single image. This results in variations of the projection gray levels that must be corrected before further processing [16].

SPECIMENS HAVE TO BE SPECIALLY PREPARED PRIOR TO THE ELECTRON EXPOSURE SINCE THE VACUUM CONDITIONS IN THE TEM AND ELECTRON RADIATION MAY DEGRADE THE BIOLOGICAL STRUCTURE.

Other preprocessing steps involve downsampling and denoising. Denoised images are used at intermediate steps (alignment or classification), whereas for the 3-D reconstruction itself the original raw data is processed since the goal there is to capture as much information as is available in the images. The standard denoising technique is Fourier filtering using soft masks (Gaussians, raised cosines, etc).

Automatic particle picking from images is becoming a particularly important step [17]. To achieve high resolution, a large number of particles must be considered. At present, particles are manually picked by the user. However, this approach is becoming infeasible as the resolution [and therefore the number of particles (100,000 or more) processed in a single reconstruction] increases.

CTF CORRECTION

The acquired images are affected by the TEM CTF, as described above. CTF correction is a primary issue in any high-resolution structural study. To correct for its effects, the CTF itself must first be estimated. This is done usually in a two-step approach. First, the power spectral density (PSD) of the image is computed and then a theoretical model of the CTF is fitted to the PSD [18]. The most popular PSD estimator in the field is the averaged periodogram, although parametric models have also been used [19].

In many cases the CTF correction is made in two steps. First, only the phase of the input data FT is corrected, while the CTF amplitude correction is performed over the reconstructed volume. Sometimes, images are also grouped by CTF. All those images with similar CTF form a single group that is processed independently. Several 3-D reconstructions are then obtained, one for each CTF group, and they are properly combined to produce a single 3-D reconstruction [20].

2-D ANALYSIS

In the early stages, 2-D analysis of projection images from similar directions (for instance, top views of the molecule) is used to get acquainted with the main features of the molecular structure (i.e., symmetry, channel sizes, etc.) (see Figure 8). This analysis helps to build models for the subsequent angular assignment and to interpret the results of image classification. The main tools used at this stage are image averaging and decomposition of the image in its Fourier harmonics (FT in cylindrical coordinates, called in the field rotational spectrum). For performing these two steps it is crucial to align each particle translationally. The alignment is commonly done by cross-correlation with a reference image.

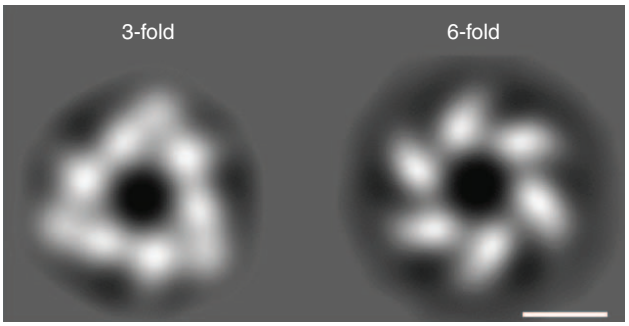
IMAGE CLASSIFICATION

Before starting a 3-D reconstruction process, it must be guaranteed that all projection images belong to the same conformational structure of the macromolecule. Otherwise, projections of different objects would be combined in a single volume, producing a blurred averaged version of all the objects considered. Projection images are classified into homogeneous groups following classical pattern recognition and clustering techniques. A feature vector is extracted from each projection image. Features usually considered are pixel values, rotational spectra, or principal components. The feature vectors are then supplied to a clustering algorithm that allows the user to divide the image set into different groups. The most popular technique in the field is hierarchical ascendant classification [21], but self-organizing maps are also widely used [22].

Since most classification methods require that the images are aligned, alignment of heterogeneous sets of images represents a chicken-and-egg problem. An alternative solution is the iterative use of 2-D alignment and classification algorithms. Multiple variations on this approach have been reported (for examples, see [2]). A more straightforward approach that intrinsically combines alignment and classification in a single, iterative process is multireference refinement. In this approach, the complete set of images is aligned with respect to a predefined number of reference images that are assumed to represent the structural diversity among the data [23].

ANGULAR ASSIGNMENT

Once a homogeneous image population is obtained, their relative 3-D orientations must be found, for example by comparing the experimental images with computer-simulated projections



[FIG8] 2-D analysis of the top views of the DnaB protein reveals two coexisting structural conformations (bar: 5 nm).

of a model resembling the molecule under study (projection matching). The projection direction of the best-matching model projection is assigned to the experimental projection. The angular parameters are refined iteratively. Projection matching can be performed in several spaces: real space [24], Radon space [25], wavelet space [26], and Fourier space [27].

Alternatively, the fact that the FTs of any two projections of a volume must share a common line (according to the central section theorem) can be exploited to compute the relative orientations of the experimental projections without a reference volume [28]. To avoid being misled by the high level of noise in the experimental images, these are clustered and the cluster averages are used instead.

3-D RECONSTRUCTION

When the relative orientations of a homogeneous set of images have been assigned, the 3-D reconstruction can then be computed. Many algorithms exist for estimating a volume (3-D) from a set of images (2-D). Their goal is usually to minimize the error between the experimental projections and the projections obtained from the reconstructed volume from the same projection directions. The most-used methods in single-particle EM are weighted back-projection (also known as filtered back-projection) and series-expansion algorithms [2]. The weighted back-projection is a direct algorithm whose rationale lies in the central section theorem and the inversion of the Fourier space reconstructed volume. However, the whole algorithm operates in the image space. Series-expansion algorithms expand the reconstructed volume in a set of weighted basis functions. Under this expansion, the tomography problem becomes one of solving a linear system of equations whose unknowns are the basis function weights.

DETERMINATION OF ICOSAHEDRAL STRUCTURES

Images of fields of randomly oriented, highly-symmetric particles can be processed to yield a 3-D structure by determining the relative positions of the symmetry elements. This approach has been most powerful when applied to icosahedral particles because of their very high symmetry (60 symmetry elements). This high symmetry allows an accurate alignment and also decreases the number of images required to determine the 3-D structure.

The typical strategy for image processing and 3-D reconstruction of icosahedral specimens includes: extraction of different views of the specimen from EM images; use of autocorrelation techniques to center the views; use of symmetry elements to calculate further sections of the 3-D FT; accurate determination of orientations by means of "common-lines" techniques in the 3-D Fourier space (these common-lines represent the intersections of the central sections in the 3-D FT); and CTF correction [9]. Once a self-consistent set of views is available, an initial 3-D structure is computed by Fourier inversion. This model is then used as a reference to refine the orientation parameters and obtain a new structure. The refinement procedure is iteratively repeated until no further improvements in the resolution are made. This methodology has allowed, for instance, high-resolution determination of the structure of human hepatitis B and Herpes virus (reviewed in [4]).

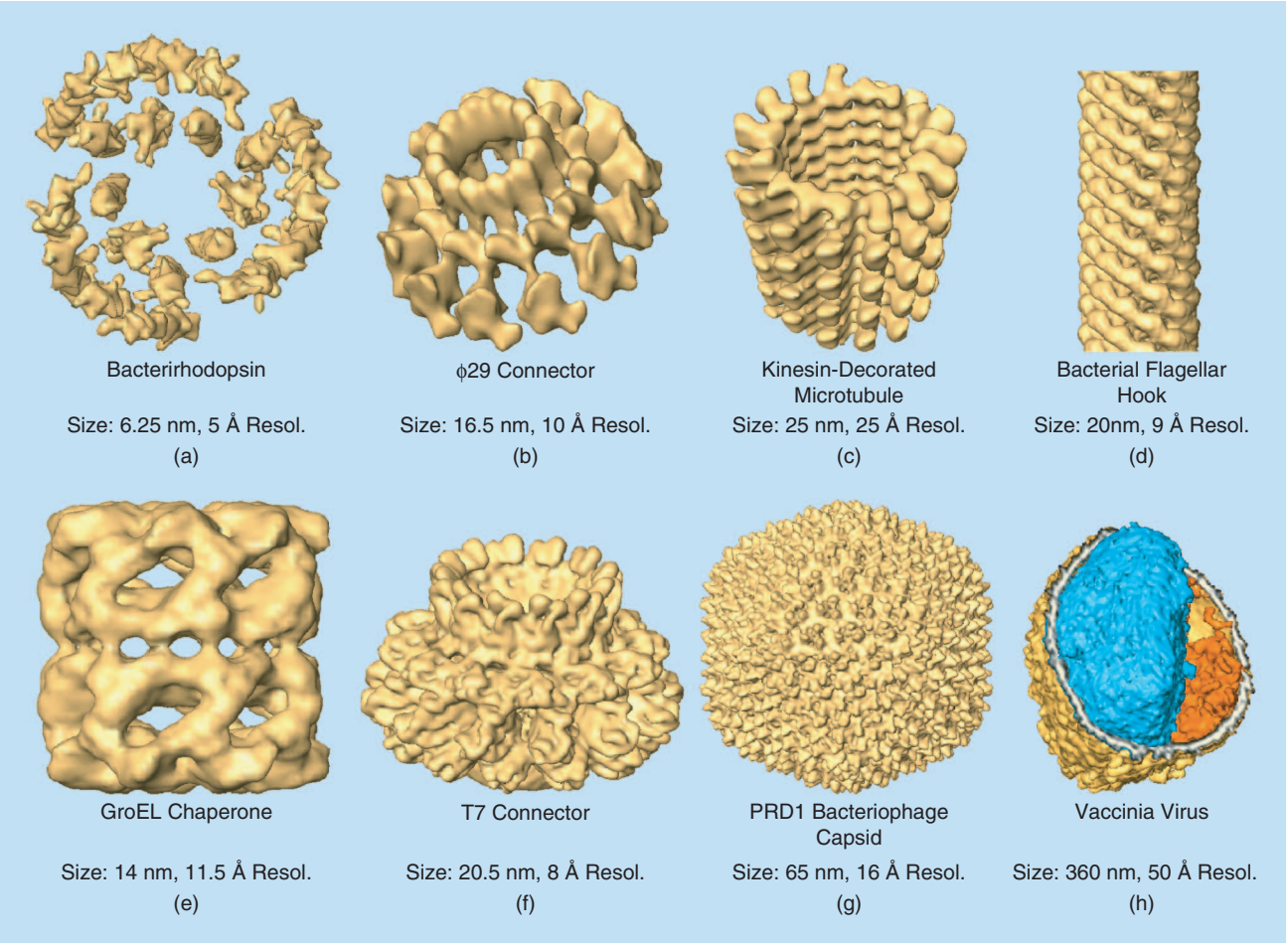
ELECTRON TOMOGRAPHY OF CELLULAR STRUCTURES

Electron tomography (ET) has emerged as the leading technique for structural analysis of complex biological specimens at molecular resolution [1], [3], with potential to bridge the gap between cellular and molecular biology. The promising prospects for attainable resolutions of 2–4 nm [3] would allow identification of macromolecular interactions in the native cellular context, which is critical to understand the cellular function [1]. ET has made possible the elucidation of the architecture of eukaryotic cells [29] or complex viruses [30].

ET requires a wide spectrum of computational tasks (described in [31]): acquisition of different views from the specimen by tilting it around an axis perpendicular to the electron beam, mutual alignment, and combination of these views to yield the 3-D reconstruction. Due to the low contrast and the limited number of images (typically 70–140), reconstructions present an extremely poor SNR that severely hinders their visualization and interpretation. As a consequence, significant computational effort is devoted to postprocessing the reconstructions to facilitate their analysis, including noise removal, segmentation, and pattern recognition to identify macromolecules in the volumes.

ILLUSTRATIVE EXAMPLES

An illustration of the variety of specimens, sizes, and resolutions encompassed by EM is shown in Figure 9. Some of the structures shown here have been obtained from the EM section of the EMBL-EBI Macromolecular Structural Database (http://www. ebi.ac.uk/msd/). Bacteriorhodopsin, a plasma membrane protein of bacteria formed by seven helices acting as a light-driven proton pump, was the first specimen solved at atomic resolution by EC. Here a bacteriorhodopsin trimer (∼62.5-nm diameter) is shown at 5 Å. The bacteriophage $\phi 29$ connector, a DNA translocating motor with 16.5 nm diameter, was the first connector solved at sub-nm resolution by EC [11]. The structure of microtubules decorated with kinesin (∼25 nm diameter; EM accession code: 1027) and salmonella typhimurium flagellar hook (∼20-nm diameter; EM accession code: 1132) were elucidated by the helical methodology at 25 Å and 9 Å, respectively. GroEL is a prokaryotic molecular chaperone (14 nm diameter; EM code: 1080) that promotes protein folding in bacteria. Here, the structure at 11.5 Å resolution by single particles approach is shown. The structure of bacteriophage $T7$ connector (∼20.5-nm diameter) was resolved at a resolution of 8 Å by single particle methods [5]. Bacteriophage PRD1 (∼65 nm diameter; EM accession code: 1011), a spherical virus with a membrane and a protein shell surrounding a core of nucleic acid, was solved at 16 Å resolution using icosahedral techniques. Finally, vaccinia, the virus that was used for vaccination against smallpox, is one of the largest (∼360 nm) and most complex viruses whose structure, lacking all symmetry, was only able to be solved by ET at 5-nm resolution [30].



[FIG9] Biological structures solved by electron microscopy: (a) bacteriorhodopsin trimer (EC), (b) bacteriophage-ϕ29 connector (EC), (c) kinesin-decorated microtubule (helical), (d) bacterial flagellar hook (helical), (e) native GroEL (single particles), (f) bacteriophage-*T*7 connector (single particles), (g) bacteriophage PRD1 capsid (icosahedral), and (h) Vaccinia virus (ET).

CONCLUSIONS

In this article, we have presented an overview of the different approaches for structural determination of biological specimens by EM. This compendium of strategies makes EM a powerful tool in structural biology, as it is able to derive structural information from specimens in the whole spectrum from the cellular to macromolecular domain. The integration of the information obtained by these different EM approaches with well-known light microscopy and high-resolution structural techniques, such as X-ray crystallography or NMR, is expected to provide a comprehensive multiscale description of the biological hierarchy. Therefore, EM plays a key role in structural biology, with unique potential to bridge the gap between cellular and molecular biology.

ACKNOWLEDGMENTS

We are thankful to Dr. C. San Martin for thoroughly revising the manuscript and for the TEM scheme in Figure 1. We also wish to thank Dr. X. Agirrezabala, Dr. J.M. Valpuesta, and Dr. J.L. Carrascosa for the data from bacteriophage connectors and vaccinia. Bacteriorhodopsin data were extracted from Protein Data Bank (entry 2BRD). Other data taken from the EMBL-EBI Macromolecular Structural Database. This work was supported by grants MEC-TIC2002-00228, BFU2004-00217/BMC, CAM-GR/SAL/ 0342/2004, BBVA-2004X578, NIH-HL70472, FIS-04/0683, EU-FP6-LSHG-CT-2004-502828, MEC-TIN2005-00447.

AUTHORS

Jose-Jesus Fernandez (jose@ace.ual.es) received the M.Sc. and Ph.D. degrees in computer science from the University of Granada, Spain, in 1992 and 1997, respectively. He was a Ph.D. student at the BioComputing Unit of the National Center for BioTechnology (CNB), Spanish Research Council (CSIC), Madrid, Spain. He became an assistant professor in October 1997 and, subsequently, associate professor of computer architecture at the University of Almeria, Almeria, Spain, in 2000. He is a member of the Supercomputing-Algorithms Research Group, an associated unit of the CNB-CSIC. He collaborates with the MRC Laboratory of Molecular Biology (Cambridge, United Kingdom). His current research interests include high-performance computing, image processing, and tomographic reconstruction in electron microscopy.

Carlos Oscar S. Sorzano (coss@cnb.uam.es) received the M.Sc. degree in electrical engineering and the B.Sc. degree in computer science from the University of Malaga, in 1997 and 2000, respectively. He received the Ph.D. degree in 2002 from University Politecnica de Madrid. He joined the National Center of Biotechnology (CSIC, Madrid, Spain) in 1997, where he worked on electron microscopy of single particles. From 2003–2004, he worked as a research assistant in the Biomedical Imaging Group, Swiss Federal Institute of Technology, Lausanne, Switzerland. He then joined the University of San Pablo CEU, where he continues his research and teaches at the Polytechnic School. His research interests include image processing, tomography, system identification, multiresolution approaches, and electron microscopy.

Roberto Marabini (roberto@cnb.uam.es) received the M.Sc. degree from the University Autonoma de Madrid, Madrid, Spain, in 1989 and the Ph.D. degree from the University of Santiago de Compostela, Spain, in 1995, both in physics. He was previously with the BioComputing Unit of the National Center for BioTechnology, Spanish National Council of Scientific Research, Madrid, Spain, and then the University of Pennsylvania, Philadelphia, and the City University of New York from 1998–2002. At present, he is an associate professor in the Escuela Superior Politecnica, University Autonoma de Madrid. His current research interests include electron microscopy, inverse problems, image processing, and high-performance computing.

Jose-Maria Carazo (carazo@cnb.uam.es) received the M.Sc. degree in theoretical physics and the Ph.D. degree in molecular biology. He is a research professor of the Spanish Research Council, CSIC, where he directs the Biocomputing Unit of the National Center for Biotechnology in Madrid. He worked at the IBM Madrid Scientific Center from 1981–1986 and from 1987–1989 at the Howard Hughes Medical Center at the New York State Health Department in Albany before joining the CSIC in 1989. His research interests are in the area of multidimensional image classification and tomographic reconstruction in electron microscopy. He has published more than 120 papers in biological and engineering journals and directed large international projects.

REFERENCES

[1] A. Sali, R. Glaeser, T. Earnest, and W. Baumeister, "From words to literature in structural proteomics," *Nature*, vol. 422, no. 6928, pp. 216–225, 2003.

[2] J. Frank, *Three Dimensional Electron Microscopy of Macromolecular Assemblies*. London: Oxford Univ. Press, 2005.

[3] V. Lucic, F. Foerster, and W. Baumeister, "Structural studies by electron tomography: From cells to molecules," *Ann. Rev. Biochem.*, vol. 74, pp. 833–865, July 2005.

[4] R. Henderson, "Realizing the potential of electron cryo-microscopy," *Q. Rev. Biophys.*, vol. 37, no. 1, pp. 3–13, 2004.

[5] X. Agirrezabala, J. Martín-Benito, M. Valle, J.M. González, A. Valencia, J.M. Valpuesta, and J.L. Carrascosa, "Structure of the connector of T7 bacteriophage at 8 Å resolution: Structural homologies of a basic component of a DNA translocating machinery," *J. Mol. Biol.*, vol. 347, no. 5, pp. 895–902, 2005.

[6] J. Dubochet, M. Adrian, J.J. Chang, J.C. Homo, J. Lepault, A. McDowall, and P. Schultz, "Cryo-electron microscopy of vitrified specimens," *Q. Rev. Biophys.*, vol. 21, no. 2, pp. 129–228, 1988.

[7] M. Yeager, V.M. Unger, and A.K. Mitra, "Three-dimensional structure of membrane proteins determined by 2D crystallization, electron cryomicroscopy, and image analysis," *Methods Enzymol.*, vol. 294, pp. 135–180, 1999.

[8] D.G. Morgan and D. DeRosier, "Processing images of helical structures: A new twist.," *Ultramicrosc.*, vol. 46, no. 1–4, pp. 263–285, 1992.

[9] P.A. Thuman-Commike and W. Chiu, "Reconstruction principles of icosahedral virus structure determination using electron cryomicroscopy," *Micron*, vol. 31, no. 6, pp. 687–711, 2000.

[10] R. Bracewell, *Two-Dimensional Imaging*. London: Prentice-Hall, 1995.

[11] J.M. Valpuesta, J.J. Fernandez, J.M. Carazo, and J.L. Carrascosa, "The three-dimensional structure of a DNA translocating machine at 10 Angstrom resolution," *Structure*, vol. 7, no. 3, pp. 289–296, 1999.

[12] R. Henderson, J.M. Baldwin, K.H. Downing, J. Lepault, and F. Zemlin, "Structure of purple membrane from halobacterium halobium: Recording, measurement and evaluation of electron micrographs at 3.5 Å resolution," *Ultramicrosc.*, vol. 19, no. 2, pp. 147–178, 1986.

[13] D.A. Agard, "A least-squares method for determining structure factors in three-dimensional tilted-view reconstructions.," *J. Mol. Biol.*, vol. 167, no. 4, pp. 849–852, 1983.

[14] R. Marabini, C.O.S. Sorzano, S. Matej, J.J. Fernandez, J.M. Carazo, and G.T. Herman, "3D reconstruction of 2D crystals in real space," *IEEE Trans. Image Processing*, vol. 13, no. 4, pp. 546–561, 2004.

[15] U. Aebi, B. Carragher, and P.R. Smith, Eds., *J. Struct. Biol. (Special Issue on Advances in Computational Image Processing for Microscopy)*, vol. 116, no. 1, pp. 1–249, 1996.

[16] C.O.S. Sorzano, L.G. de la Fraga, R. Clackdoyle, and J.M. Carazo, "Normalizing projection images: A study of image normalizing procedures for single particle 3D electron microscopy," *Ultramicrosc.*, vol. 101, no. 2–4, pp. 129–138, 2004.

[17] C.S. Potter, Y. Zhu, and B. Carragher, Eds., *J. Struct. Biol. (Special Issue on Automated Particle Selection for Cryo-Electron Microscopy)*, vol. 145, no. 1–2, pp. 1–180, 2004.

[18] J.J. Fernandez, J.R. Sanjurjo, and J.M. Carazo, "A spectral estimation approach to contrast transfer function detection in electron microscopy.," *Ultramicrosc.*, vol. 68, no. 4, pp. 267–295, 1997.

[19] J.A. Velázquez-Muriel, C.O.S. Sorzano, J.J. Fernández, and J.M. Carazo, "A method for estimating the CTF in electron microscopy based on ARMA models and parameter adjusting," *Ultramicrosc.*, vol. 96, no. 1, pp. 17–35, 2003.

[20] E.V. Orlova, P. Dube, J.R. Harris, E. Beckman, F. Zemlin, J. Markl, and M. van Heel, "Structure of KLH1 at 15 Å resolution by electron cryomicroscopy and angular reconstitution," *J. Mol. Biol.*, vol. 271, no. 3, pp. 417–437, 1997.

[21] M. van Heel, "Multivariate statistical classification of noisy images (randomly oriented biological macromolecules)," *Ultramicrosc.*, vol. 13, no. 1–2, pp. 165–183, 1984.

[22] A. Pascual-Montano, L.E. Donate, M. Valle, M. Bárcena, R. Pascual-Marqui, and J.M. Carazo, "A novel neural network technique for analysis and classification of EM single-particle images," *J. Struct. Biol.*, vol. 133, no. 2–3, pp. 233–245, 2001.

[23] S.H.W. Scheres, M. Valle, R. Núñez, C.O.S. Sorzano, R. Marabini, G.T. Herman, and J.M. Carazo, "Maximum-likelihood multi-reference refinement for electron microscopy images," *J. Mol. Biol.*, vol. 348, no. 1, pp. 139–149, 2005.

[24] P. Penczek, R.A. Grasucci, and J. Frank, "The ribosome at improved resolution: New techniques for merging and orientation refinement in 3D cryo-electron microscopy of biological particles," *Ultramicrosc.*, vol. 53, no. 3, pp. 251–270, 1994.

[25] M. Radermacher, "Three-dimensional reconstruction from random projections–orientational alignment via Radon transforms," *Ultramicrosc.*, vol. 53, no. 2, pp. 121–136, 1994.

[26] C.O.S. Sorzano, S. Jonic, C. El-Bez, J.M. Carazo, S. De Carlo, P. Thevenaz, and M. Unser, "A multiresolution approach to pose assignment in 3-D electron microscopy of single particles," *J. Struct. Biol.*, vol. 146, no. 3, pp. 381–392, 2004.

[27] S. Jonic, C.O.S. Sorzano, P. Thevenaz, C. El-Bez, S. De Carlo, and M. Unser, "Spline-based image-to-volume registration for three-dimensional electron microscopy," *Ultramicrosc.*, vol. 103, no. 4, pp. 303–317, 2005.

[28] P.A. Penczek, J. Zhu, and J. Frank, "A common-lines based method for determining orientations for N > 3 particle projections simultaneously," *Ultramicrosc.*, vol. 63, no. 3–4, pp. 205–218, 1996.

[29] O. Medalia, I. Weber, A.S. Frangakis, D. Nicastro, G. Gerisch, and W. Baumeister, "Macromolecular architecture in eukaryotic cells visualized by cryo-electron tomography," *Science*, vol. 298, no. 5596, pp. 1209–1213, 2002.

[30] M. Cyrklaff, C. Risco, J.J. Fernandez, M.V. Jimenez, M. Esteban, W. Baumeister, and J.L. Carrascosa, "Cryo-electron tomography of vaccinia virus," *Proc. Natl. Acad. Sci. USA*, vol. 102, no. 8, pp. 2772–2777, 2005.

[31] A.P. Leis, M. Beck, M. Gruska, C. Best, R. Hegerl, W. Baumeister, and J.W. Leis, "The essential role of digital signal processing in cryo-electron tomography of biological specimens," *IEEE Signal Processing Mag*, vol. 23, no. 3, pp. 95–103, May 2006. SP