Structure Determination using Electron Crystallography

James E. Evans, Ludovic Renault, and Henning Stahlberg*

Molecular & Cellular Biology, University of California at Davis, 1 Shields Ave., Davis, CA 95616, USA

* Corresponding author:

Henning Stahlberg
Molecular and Cellular Biology,
University of California at Davis,
1 Shields Ave., Davis, CA 95616, USA
Tel.: +1 (530) 752 8282 Fax: +1 (530) 752 3085
email: HStahlberg@ucdavis.edu

Electron Crystallography

Electron crystallography in structural biology is a method that uses a transmission electron microscope to study the structure of proteins that are arranged in the form of two-dimensional (2D) crystals. To date, electron crystallography has determined the structure of one soluble and seven membrane proteins at atomic resolution. We describe here the method, benefits and limitations of electron crystallography for analyzing two-dimensional crystals.

What is a 2D crystal?

Imagine walking along an old brick road or atop a marble tiled floor in a museum. Now compare the two types of surfaces; the brick road is probably loosely packed with random spacing and rotation which causes breaks in the overall pattern, while the tiles in the museum floor are probably arranged in perfectly parallel rows with well-defined spacing between neighbors. While the quality of the two surfaces may be drastically different, they both represent 2D crystals where a single unit (such as a tile) is repeatedly packed closely together along two dimensions (X and Y).

In structural biology, 2D crystals exist where the repeating subunit is a purified protein. These 2D crystals can be formed with either soluble proteins or with membrane proteins that are surrounded by lipids. For most soluble proteins, it is more interesting to grow three-dimensional (3D) crystals instead of 2D crystals, since 3D crystals are suitable for X-ray crystallography. Membrane proteins are therefore the most widely studied form of 2D crystals.

Benefits of 2D crystallization.

Since 2D crystals exist as a single layer along the Z dimension, they only need to become well ordered along the X and Y dimensions. This type of single stacking within the crystal removes a level of complexity making 2D crystals sometimes easier to grow than 3D crystals. In addition, 2D crystal formation generally requires a 10-fold lower protein concentration than is needed for 3D crystallization. However, due to the lack of interaction within the Z dimension, the overall order and resolution of 2D crystals is generally worse than that of 3D crystals arising from imperfections in the growth pattern. Despite this limitation, one very nice benefit of electron crystallography is that the membrane proteins forming the 2D crystal are surrounded by lipids, which mimics their native environment. Additionally, due to the fluidity of the lipid bilayer, some 2D crystals will allow for post-crystallization modification. Such modification can occur when the 2D crystal is exposed to a changing environment, by, for example, changing the buffer solution, the pH, or by adding inhibitors or ligands that cause a conformational change in the shape of the protein (Vinothkumar *et al.*, 2005). Imaging these crystals under different conditions allows visualizing dynamic changes within the protein shape associated with its function.

How do we make 2D crystals?

- ***Dialysis***

Step-wise and continuous detergent dialysis are the most common methods for creating well ordered 2D crystals. While dialysis has been used for several decades to form 3D crystals of soluble proteins using simple buffer replacement, the formation of 2D crystals of membrane proteins relies upon the additional removal and replacement of detergent molecules (surrounding the protein) with lipid molecules (Fig. 1a). During purification, membrane proteins are removed from their native lipid environment and surrounded by detergent molecules which aide in maintaining the protein's original structure and preventing aggregation. These detergent-solubilized membrane proteins can then be reconstituted within a new lipid bilayer, by adding lipids to the membrane protein solution, followed by careful removal of the detergent molecules using a dialysis membrane. Due to the fluidity of lipid bilayers, once embedded in the newly formed membrane, these proteins (under certain conditions) can organize themselves into a 2D crystal using both, protein-protein and protein-lipid interactions for stability. In 1992, a computer controlled dialysis device was developed (Jap *et al.*, 1992) that allowed testing several conditions at a time with triplicate wells. Recently, a new device (Vink *et al.*, 2007) has been published that allows testing up to 96 dialysis conditions per experiment.

- ***Lipid monolayer assisted***

An alternative method for creating 2D crystals uses a lipid monolayer to immobilize a layer of protein, which brings the proteins into close proximity to other proteins, thereby increasing the likelihood of protein-protein interactions (Fig. 1b). The lipid monolayer is formed at the air-water interface by spreading lipids across the water surface. This spreading is facilitated by dissolving the lipid mixture in a chloroform / methanol solution and dropping a small volume of this mixture on the surface of water (Levy *et al.*, 1999). The organic solution will phase-separate versus the water and then evaporate leaving a lipid monolayer on the surface with the hydrophobic tails oriented upwards into the air and the hydrophilic head group interacting with the water. The use of functionalized lipids containing a Ni-NTA (Kubalek *et al.*, 1994) or Streptavidin moiety allows for specific binding of tagged proteins while a composition of charged and neutral lipids would allow non-specific binding. After the formation of the monolayer, the protein is injected through a side-port into the solution and allowed to bind before picking up the monolayer on an EM grid (Levy *et al.*, 1999). Under certain conditions a 2D crystal will form. While this method has given rise to several structures, the resolution obtained has rarely exceeded 1 nm for a 3D structure although several 2D projection maps have been solved at sub-nanometer resolution (Lambert *et al.*, 1999; Levy *et al.*, 2001). This method is bound by certain restrictions in the choice of lipids and detergents, as most detergents would dissolve a lipid monolayer from the water surface. Nevertheless, this technique can be adapted for detergent-solubilized membrane proteins, and therefore not only allows for crystallization of soluble (Chen *et al.*, 2003; Makhov *et al.*, 2004), but also of membrane proteins (Lebeau *et al.*, 2001).

How do we get structural information from a 2D crystal?

When you look at a photograph of a dog or a cat, your mind automatically assumes it is a representation of a 3D object. This assumption stems from preconceived notions that the animal probably has four legs and a tail, even though these features might not appear in the picture. The only way to prove the real 3D shape is by taking multiple pictures of the same animal from different angles (top, side, back, etc.). In electron microscopy, the same theory applies. If we have a gallery of 2D images that are taken of the same sample or of identical samples from different angular views, then we can reconstruct the original 3D structure. In comparison with light photographs of dogs and cats that show their fur but not their inner bones, electron microscopy images of proteins are even better suited for a full 3D reconstruction, since electron microscopy images are true projections of the inner density of the sample (similar to an X-ray image of a cat). In other words, the resulting micrograph represents the projected density of the specimen as electrons pass through the sample. Due to the high energy and small size of the electron and its short wavelength (2.5×10^{-12} m @ 200keV), the sample seems nearly transparent when at perfect focus.

Every image can be represented mathematically using a technique known as the Fourier transform, in which the image is transformed into a new dataset where two measurements - amplitudes and phases - are determined to describe all details within the image. Whereas an image can be described as grey-values as a function of x or y pixel coordinates, a Fourier transform can be described as amplitude and phase values as a function of spatial frequencies in x and y directions. These spatial frequencies range from 0 at the center of the Fourier transform to +/-0.5 at the border. These frequencies correspond to spacings within the image. If there is a repetitive unit within the image (as is the case for an image of a 2D crystal), then some frequencies will be stronger in the Fourier transform, giving rise to lattice reflections being seen as bright spots.

Some methods can also perform diffraction on a crystalline structure. An electron or X-ray beam, for example, can be diffracted on a crystal, and the resulting diffraction pattern from that crystal can be recorded on a camera. Such a diffraction pattern is similar to a Fourier transform, with the only differences being that the intensity of the diffraction spots represents the square of the amplitude of the corresponding frequency, and phase information is missing, whereas the computer-calculated Fourier transformation from a real-space image contains both amplitudes *and* phases.

Acquiring structural data

There are two main modes for operating the transmission electron microscope when studying 2D crystals – direct imaging and electron diffraction.

Direct imaging is like taking a black and white photograph of the sample. Electrons traverse the sample in the microscope and the lenses of the electron microscope are used to form an image on the screen, photographic film or CCD camera, which then will show you how the protein “looks” (Fig. 2a). Since proteins are very small (in the order of 10 nm, or 10^{-8} m), we use a high magnification (50,000x) to record the images at a size we can visualize. With proteins, however, we are careful not to further magnify the image with the electron microscope, because in order to do so we would have to concentrate all

available electrons onto an ever-smaller area of the sample, and this intense electron beam would destroy the sample before any high-resolution data could be recorded. These two boundary conditions (the resolution of the CCD on one end, and the beam-sensitivity of the sample on the other end) limit the magnification in biological transmission electron microscopy to a range between 10,000 and 80,000 times, even though the microscope could go much higher. Since the *imaging* method takes an actual picture of the sample, it is susceptible to blurring caused by sample drift or vibration. Recording of a high-resolution image also requires a well-aligned electron microscope with good performance and specifically a good beam coherence.

Electron diffraction records the intensity distribution of the diffracted electron beam (Fig. 2b). The resulting pattern contains the information about the different spatial frequencies within the sample. In fact, the recorded electron diffraction pattern contains values that correspond to the square of the amplitudes from a Fourier transformation of a direct image of the same area that could have been recorded instead. Thus, the diffraction pattern contains very good amplitude values, but lacks all phase information. However, the diffraction pattern recording is insensitive to sample drift or vibration, and is also not dependent on the microscope's beam coherence. For this reason, a diffraction pattern is almost not limited in resolution by the microscope, and for a well-ordered 2D crystal, this method allows recording the amplitude information to a very high resolution.

If you also need the phases, then recording real-space images of your 2D crystals with your electron microscope is often the only way to obtain a reconstruction of your protein. However, if you have a possibility to obtain the required phases with an alternative method (e.g. homologous replacement with a known similar structure), then electron diffraction can provide the highest resolution available for your dataset.

Image processing programs

All available high-resolution protein structures from electron crystallography were determined with the MRC image-processing software package. Recently, the IPLT and the *2dx* packages were introduced, which are also specialized for electron crystallography data processing. These software packages use Fourier transforms, reference matching and image unbending to determine the ideal projection map of the unit cell for the crystal. The MRC package was originally developed by Richard Henderson and co-workers at the Medical Research Council in Cambridge, United Kingdom, and established an image-processing scheme for analyzing 2D crystals. The MRC software also has components for analyzing electron diffraction patterns, helical particles and single particle analysis (Crowther *et al.*, 1996; Grigorieff *et al.*, 1996; Kunji *et al.*, 2000).

IPLT (the Image Processing Library and Toolkit) is a suite of libraries written in C++ and Python. IPLT is an open source project that is continuing to expand with added functionality and libraries (Philippsen *et al.*, 2007; Philippsen *et al.*, 2003).

The *2dx* software is built upon the MRC package and offers a user-friendly graphical user interface and automatic processing capabilities for electron crystallography (Fig. 3). *2dx* is setup with standard and custom scripts, which help guide the user through each image-processing task ultimately ending with either a 2D projection map or a 3D volume, if merging data from multiple images and tilts. An additional module follows a Maximum

Likelihood scheme that allows analysis of poorly ordered 2D crystals by treating each individual unit cell like a single particle for alignment and averaging (Gipson *et al.*, 2007a; Gipson *et al.*, 2007b; Zeng *et al.*, 2007).

Interesting structures from 2D crystals

Tubulin

In 1998 the atomic structure of alpha and beta tubulin was solved by electron crystallography (Nogales *et al.*, 1998b). While most 2D crystals involve membrane proteins embedded within a lipid bilayer, tubulin is a soluble protein that readily assembles into a filament known as a microtubule. The alpha and beta tubulin heterodimers stack atop other heterodimers to form a single protofilament, but this stacking only occurs on one end known as the "plus" end. Normally, multiple protofilaments interact side-by-side to form a polarized cylindrical microtubule (~ 25 nm in diameter and up to several μm in length) that acts as part of the cell's cytoskeleton and is an important component of many biological processes like cell division and cargo transportation (Li *et al.*, 2002).

Although it was not until 1998 that the atomic structure was determined, the crystallization protocol for tubulin was first established in 1995 (Nogales *et al.*, 1995). It was discovered that if alpha and beta tubulin was incubated at 37°C in the presence of zinc ions then a well-ordered 2D crystal sheet formed. In the presence of zinc, individual protofilaments still interacted laterally with neighboring protofilaments, but in an alternating orientation. This allowed for the protofilaments to assemble and interact while forming a flat sheet rather than a tubular cylinder. The structure of the alpha and beta tubulin dimer has been refined to 3.5 Angstrom resolution (Lowe *et al.*, 2001; Nogales *et al.*, 1997; Nogales *et al.*, 1998b; Nogales *et al.*, 1995).

According to the atomic structure, the individual structures of alpha- and beta-tubulin are nearly identical. Each monomer has a core composed of two beta-sheets surrounded by several alpha-helices (Fig. 4a). The monomers are very compact and have two main functional domains corresponding to the amino-terminal domain containing the nucleotide-binding region and the carboxy-terminal domain, which acts as the binding surface for associated motor proteins. Additionally, the structure of the polymerized microtubule has also been solved to 8 Angstrom resolution (Li *et al.*, 2002) and the interpretation of this structure and its dynamic assembly has been aided by the detailed knowledge of the heterodimer structure (Wolf *et al.*, 1996). Interestingly, although both alpha and beta tubulin bind GTP and have a nearly identical shape, only the dimerization domain between the beta-alpha interface (and not the alpha-beta interface present in the protofilament) allows for any functional activity such as GTP hydrolysis to GDP.

Due to the central role of microtubules during cell division and other dynamic processes, they are an important drug target. In fact, several drugs currently exist which interact with microtubules to affect their functions. Three of these drugs are Taxol, Epothilone and Colchicine. All of these drugs affect the stability of the microtubules with Taxol and Epothilone stabilizing and Colchicine destabilizing the filament. Interestingly, the structure of tubulin in the presence of each of these drugs has been solved using electron crystallography of zinc-induced tubulin sheets (Bai *et al.*, 2000; Nettles *et al.*,

2004; Nogales *et al.*, 1995; Snyder *et al.*, 2001; Wang *et al.*, 2005). It is now clearly understood where each drug binds to the beta-tubulin subunit of the tubulin heterodimer (Fig. 4b) and how this binding affects the stability of the overall microtubule, which is useful information for even better design of subsequent drugs.

The structure of tubulin solved by electron crystallography opened the floodgate and provided more questions then answers, which has always been the signature of outstanding scientific achievements.

HIV-1 CA

An atomic resolution structure of the full-length capsid (CA) protein of the Human Immunodeficiency Virus (HIV) has so far not been determined. However, the structures of both the N-terminal domain (NTD) and the C-terminal domain (CTD) of the CA protein have been solved individually by X-ray crystallography. Yet, despite knowing these structures, it remained unclear whether they were the correct structure for either the immature or mature virion since several conformational changes occur between the two forms. In addition, five different X-ray structures exist of the CTD in its dimeric form (which is believed to be the precursor to the hexameric lattice), however only two of these crystal structures have the same CTD-CTD dimerization interaction region (Ganser-Pornillos *et al.*, 2007). Therefore, while detailed information exists for the structure of the monomeric NTD and CTD domains, it is not clear how these dimerize and interact to form the overall hexameric lattice for the mature virion.

It is believed that the same capsid assembly process occurs for most spherical viruses. In 2003, a lipid monolayer assisted 2D crystal of the Moloney murine leukemia virus (M-MuLV) CA protein was shown to form a similar hexagonal lattice to that of the CA proteins from HIV-1 and the Rous sarcoma virus (Ganser *et al.*, 2003). However, while previous low-resolution electron microscopy studies of HIV-1 have provided some insight into the overall lattice configuration of the capsid, the specific location of the NTD and CTD had not been identified.

In order to elucidate a better model for the structure and role of the HIV-1 capsid protein, Ganser-Pornillos *et al.*, 2007, sought to determine a medium resolution structure of the full-length protein using electron crystallography (Fig. 5). The authors found that a single point mutation of amino acid 18 from arginine to leucine within the NTD stabilized the hexagonal lattice, which under certain conditions could be assembled into large spheres. When adsorbed to a TEM grid, these spheres flattened and acted like a well-ordered 2D crystal from which they were able to obtain a 9 Angstrom structure for the full-length CA protein. Subsequent docking of the known structures of the NTD and CTD monomers into this volume resulted in a pseudo-atomic model for the assembly of the HIV-1 capsid (Fig. 5c). This pseudo-atomic model depicts the hexagonal lattice as composed of six CA proteins with the NTD of each monomer interacting to form an inner hexameric ring. The CTD of each monomer forms the outside of the ring and allows for dimerization to other hexameric rings to form the complete capsid. The NTD and CTD interface of each monomer are tightly coupled but also interact with neighboring monomers thereby providing additional stability to the overall lattice. Interestingly, the new pseudo-atomic model for the HIV-1 capsid shows the binding regions for two known drugs and provides further insight into the mechanism for disrupting the mature virion.

Both drugs bind and disrupt the NTD-CTD interface thereby decreasing the stability of the overall lattice and causing the mature virion to disassemble and become non-virulent.

Although electron crystallography did not provide a full atomic model for the CA protein, it allowed previously known X-ray structures of individual components to be understood in terms of fitting the overall structure of the native virus capsid (Ganser-Pornillos *et al.*, 2007).

Aqp0

Aquaporin-0 (Aqp0) belongs to the water channel superfamily that comprises aquaporins and aquaglyceroporins. Such proteins form selective pores in cell membranes for water or other neutral solutes. Peter Agre was awarded the Chemistry Nobel prize in 2003 for his role in the discovery and structural and functional analysis of the aquaporin family (Agre *et al.*, 1997; Denker *et al.*, 1988; Preston *et al.*, 1992). The structure of aquaporin-1 (Aqp1) (Murata *et al.*, 2000; Ren *et al.*, 2000) was determined by electron crystallography, followed by work on the structure of Aqp2 (Schenk *et al.*, 2005) and the structure determination of Aqp4 (Hiroaki *et al.*, 2006).

Aqp0 is the most abundant aquaporin in the eye's lens fiber where it forms membrane junctions and plays an essential role for lens transparency. Mutations in Aqp0 can lead to formation of an eye lens cataract obscuring normal vision.

In 2005, the Aqp0 structure was determined by electron diffraction from double-layered two-dimensional membrane crystals at 1.9 Angstrom resolution (Gonen *et al.*, 2005). This structure allowed the identification of water molecules in the electron density map of the channel pathway, and revealed the detailed interaction between lipids and protein.

Two-dimensional crystals of Aqp0 were obtained by addition of DMPC lipids to the protein mixture with subsequent detergent removal by dialysis. Large sheets of a few micrometers in size were obtained using a lipid to protein ratio of 0.25 (w/w) and the resulting crystals were double-layered with a unit cell with dimensions of 65.5 x 65.5 Angstroms and P422 symmetry.

The Aqp0 structure was solved by molecular replacement using the Aqp1 structure as a search model (Sui *et al.*, 2001) and the final structure has six transmembrane-spanning helical regions with intracellular N- and C-termini (Fig. 6b & c).

Detailed analysis of the structure reveals three water molecules confined in the water pathway. The cavity where the water molecules are observed is larger than the bottom and top parts of the water pathway. The water molecules are to far from each other to interact together. Instead, hydrogen bonding is established with the polar residue Tyr24 and the acidic residues Asn68 and Asn184. It is interesting to note that Tyr24 is not conserved among other aquaporins. Gonen *et al.*, 2005, suggest that the presence of Tyr24 creates a barrier for the passage of water molecules that would explain the poorer water conductance of Aqp0, compared to other members of the aquaporin family.

DMPC lipids were used to crystallize Aqp0 and they are well visualized in the structure. Aqp0 is arranged as a tetramer, and each monomer of Aqp0 interacts with 9 DMPC lipid molecules. Two of those lipids do not have direct contact with the protein

amino acids, while the seven in contact form a surrounding barrier around the Aqp0 tetramer. The Aqp0 tetramers are separated by lipids and laterally do not interact directly with other tetramers (Fig. 6a). Only the protein-lipid-protein interactions hold tetramers together in a crystalline arrangement.

It is not uncommon to find lipid molecules tightly bound to the membrane protein, even if lipids are not present during the purification process (Palsdottir & Hunte, 2004). But in the case of aquaporins, there are no known binding sites for lipids. None of the X-ray structures published contained bound lipids. Thus the lipid-protein interactions observed in the Aqp0 electron crystallography structure are a model for any membrane protein interacting with lipids inside a lipid bilayer.

A network of hydrogen bonds and salt bridges hold the annular lipids together and lipids interact with the protein mainly via their phosphate headgroup (see Fig. 6d for a schematic representation of a DMPC molecule) similar to the interactions observed previously for tightly bound lipids for other proteins (Palsdottir & Hunte, 2004).

The 1.9 Angstrom structure of Aqp0 has impressively demonstrated the value and power of electron crystallography, in reaching high-resolution and allowing the study of the membrane-embedded conformation and the detailed protein-lipid interaction.

Future Outlook – Limitations and Advancements

In 2005, the structure of lipid-embedded Aqp0 (Gonen *et al.*, 2005) at a resolution better than two Angstroms established a milestone in the capabilities of electron crystallography. This major achievement hinted at the possibility of using electron crystallography for structure based drug design (SBDD), which requires knowledge of the exact positions of the amino acid side chains to better than 2 Angstrom resolution for successful drug development. However, an analysis of the available membrane protein structures from electron crystallography (available on http://2dx.org/structures) reveals that Aqp0 is the only structure to date that would fulfill the criteria for SBDD. More common is the determination of the 3D structure of membrane proteins at around 7 Angstroms. Luckily, this resolution allows for visualizing alpha-helices - the type of secondary protein structure that typically spans the lipid bilayer. Therefore, matching the sequence to these regions of alpha-helical structure can allow for tracing the shape of the protein and starting to understand the structure / function relationship.

During the last two-decades, a lot of exciting improvements have been made in electron crystallography. They include development of better crystallization techniques, specimen preparation for cryo-freezing the 2D crystals, and new tailor-made software for the special needs of electron crystallography.

Still, technical difficulties remain and need to be overcome.

Among others, the crystallization using dialysis uses too much protein quantity. Typically, for each condition tried, 50 micrograms of protein is needed. That means that for a typical screen of 24 conditions, more than 1 milligram of pure protein is required. Obtaining that much membrane protein is not always an easy task, especially for membrane proteins that assemble into complexes. Thus, other crystallization approaches using less protein are a requirement for the electron crystallography field. A first step in

this direction has been taken recently with the development of a dialysis device capable of utilizing less protein and testing more conditions (Vink *et al.*, 2007) .

Another bottleneck is the screening of the crystallization conditions tested. Evaluation of a 2D crystallization trial currently still has to be done using the electron microscope by examining a negatively stained preparation of the crystallization trial outcome. A thorough analysis of one trial therefore takes ~30 minutes from sample preparation to imaging. Several attempts at the development of automatic screening software are currently underway, but recognizing a well-ordered 2D crystal among the multitude of possible appearances of stain or salt artifacts, protein aggregates, or non-crystalline lipid membranes is difficult even for an experienced human operator. Implementation of reliable image recognition software for that task will be a challenging but important task. In addition, a robot (e.g. Cheng *et al.*, 2007) or a multi-grid sample holder (e.g. Lefman *et al.*, 2007) would be required to change the condition screened, and most of all, an electron microscope is still a prerequisite, where maintenance and transfer of the sample into the microscope's vacuum are also a time-consuming. If the screening step could be surmounted through automation or another approach, the rapid screening of hundreds of independent dialysis conditions would allow better identification and refinement of conditions for optimal crystal growth.

Once highly ordered 2D crystals are obtained, the identification of a cryo-EM sample preparation method that preserves the high-resolution order of the crystals is another time consuming step. Currently, there is no sample preparation method capable of preserving all highly ordered 2D crystals of membrane proteins. For that reason, a suitable sample preparation method has to be determined by trial and error for each new crystal project.

Data collection is hampered by a low success rate for acquiring quality images when tilting 2D crystal samples, which is necessary for reconstructing the overall 3D structure. A strong resolution loss in the direction perpendicular to the tilt axis has been associated with beam-induced specimen charging (Gyobu *et al.*, 2004) or movement (Typke *et al.*, 2007). Current image processing algorithms also require the tilted 2D crystals to lie on a perfectly flat support film (Glaeser, 1992; Vonck, 2000), in order to be able to extract the high-resolution information from the images. The yield of data collection for tilted samples can be optimized by a symmetric sandwich specimen preparation method (Gyobu *et al.*, 2004), and/or by the so-called SpotScan illumination mode (Downing, 1991).

Computer structure reconstruction from the recorded images requires further automation and improvement. While first steps towards full automation have been made (Gipson *et al.*, 2007a; Typke *et al.*, 2005), the recently developed algorithms from the field of single particle cryo-EM still need to be ported towards the field of electron crystallography.

These continuous advances are pushing the attainable resolution for membrane proteins and promise to provide atomic resolution structures more frequently in the future from electron crystallography of two-dimensional membrane protein crystals.

FIGURES

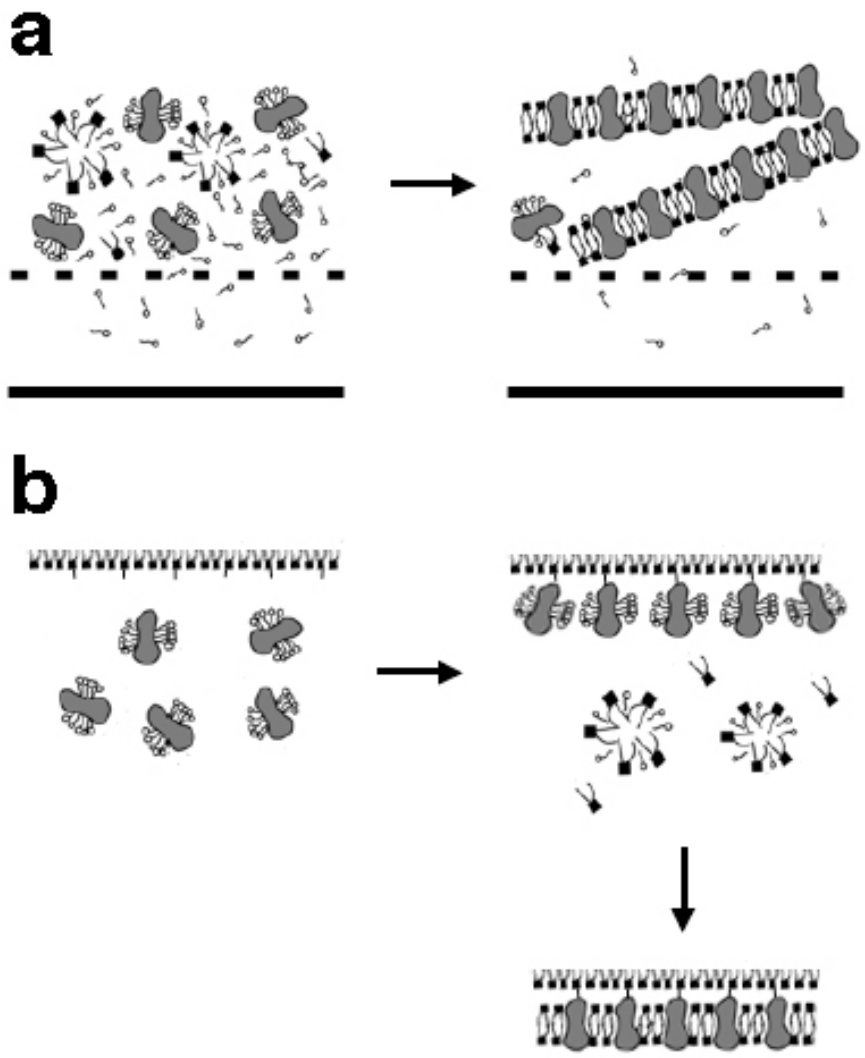

Figure 1 – Schematic of 2D crystal formation. a) During dialysis, purified and detergent solubilized membrane proteins are reconstituted into a lipid bilayer by the removal and replacement of detergent molecules with lipid molecules across a permeable membrane. b) 2D crystallization using the monolayer method requires first creating a lipid monolayer at the air-water interface, and then allowing the purified protein to interact and bind the lipid monolayer. After binding, lipids can be added to replace the detergent molecules to form a 2D crystal underneath the monolayer.

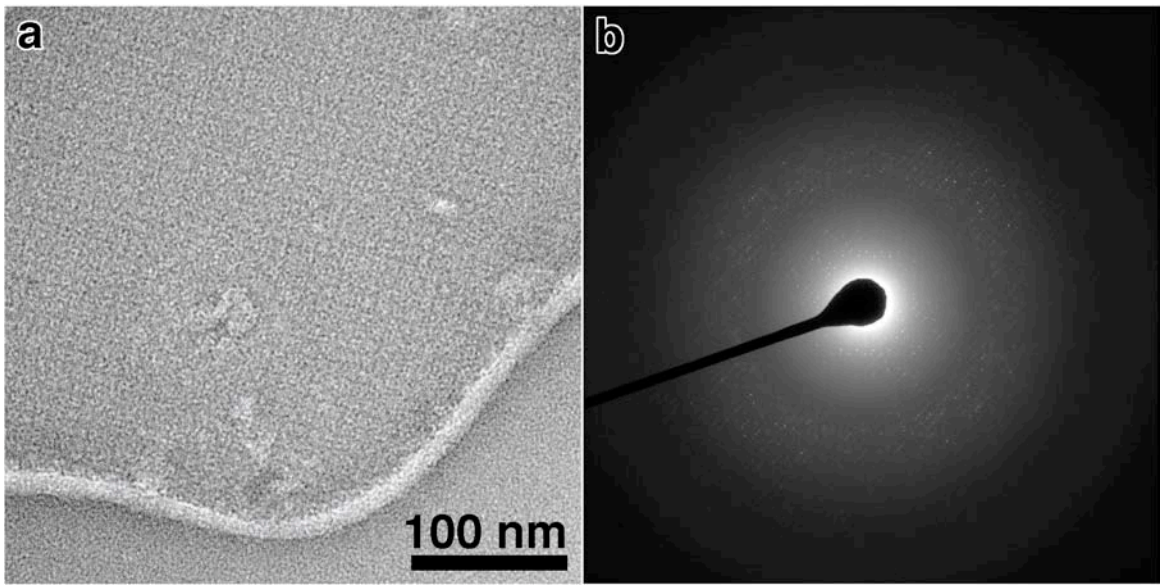

Figure 2 – Data collection for 2D crystals. a) Direct image of a negatively stained 2D crystal of the cyclic nucleotide gated potassium channel MloK1, depicting membrane proteins arranged in a square lattice. Protein appears bright. b) Electron diffraction of a 2D crystal of the *E. coli* chloride channel ClC-ec1, containing reflections to 2.6 Angstrom resolution. Images courtesy of Po-Lin Chiu (a) and Hui-Ting Chou (b). The image in (b) was recorded in the laboratory of Y. Fujiyoshi, Kyoto, Japan.

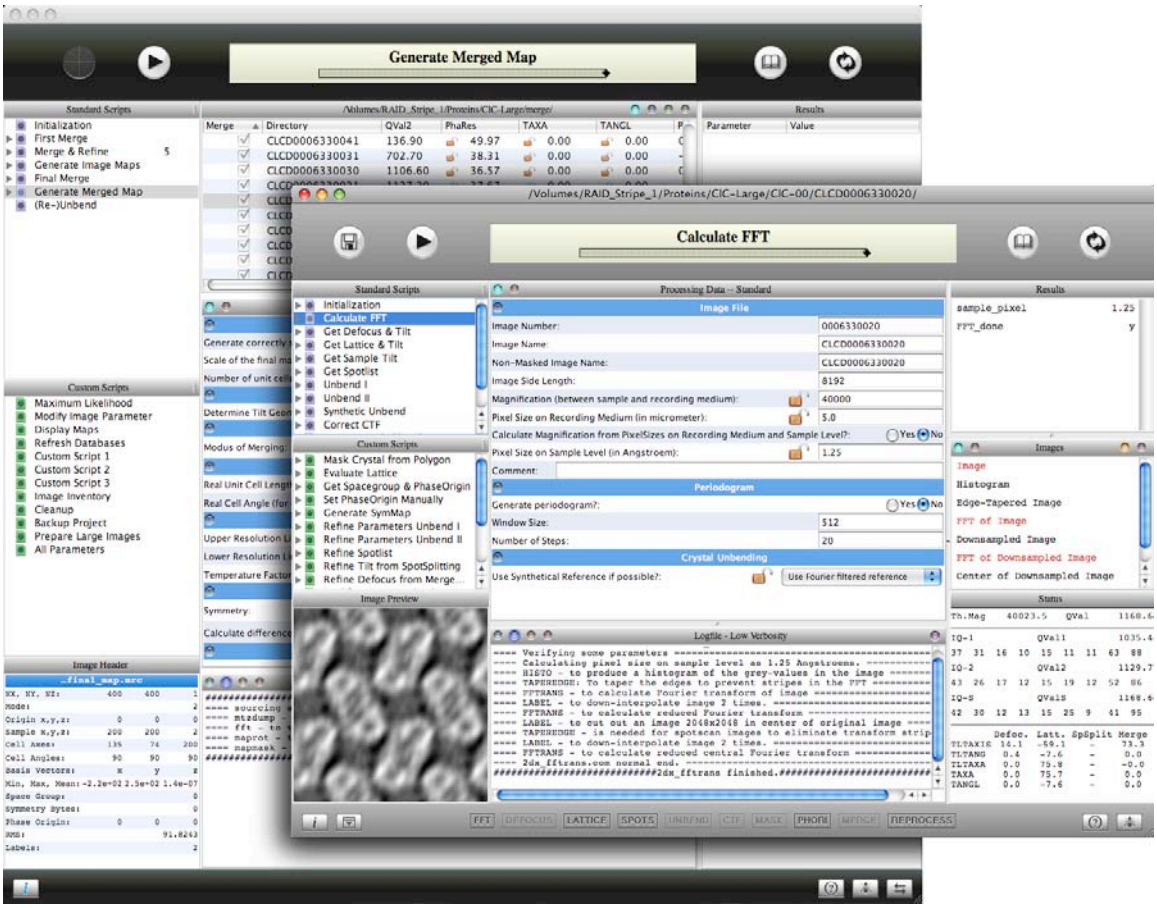

Figure 3 – The graphical user interface (GUI) of the *2dx* software. Shown are the two software components *2dx_merge* and *2dx_image*. The GUI allows fully automated or manually assisted processing.

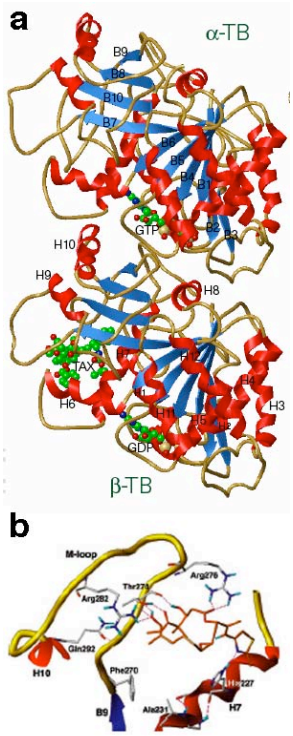

Figure 4 – Tubulin a) The structure of the alpha and beta tubulin heterodimer depicting their dimerization domain, Taxol binding site and nucleotide binding region. b) Binding of the microtubule stabilizing drug Epothilone A to beta-tubulin highlighting the amino acid sidechain interactions. Figure adapted from Nettles *et al.*, 2004 and Nogales *et al.*, 1998a.

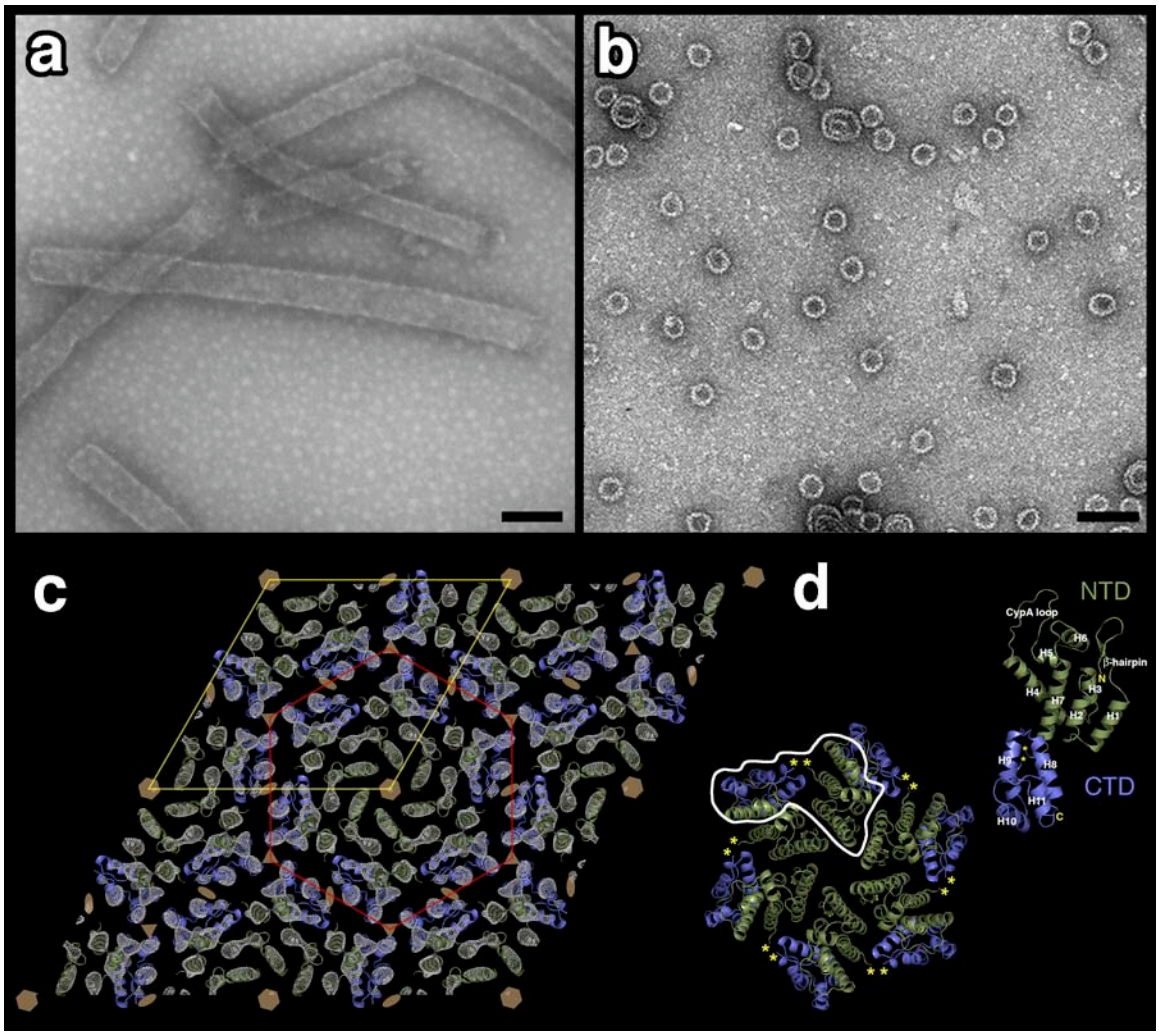

Figure 5 – The HIV-1 CA protein. a) Full-length wild-type CA protein forms long but narrow tubular 2D crystals, however, a single point mutation (R18L) causes the CA protein to form spherical 2D crystals (b), which under certain conditions can grow very large and when flattened can appear as a double layered 2D crystal. c) Projection map of the CA protein unit cell with the pseudo-atomic model (d) fit to the observed density. Figure adapted from Ganser-Pornillos *et al.*, 2007.

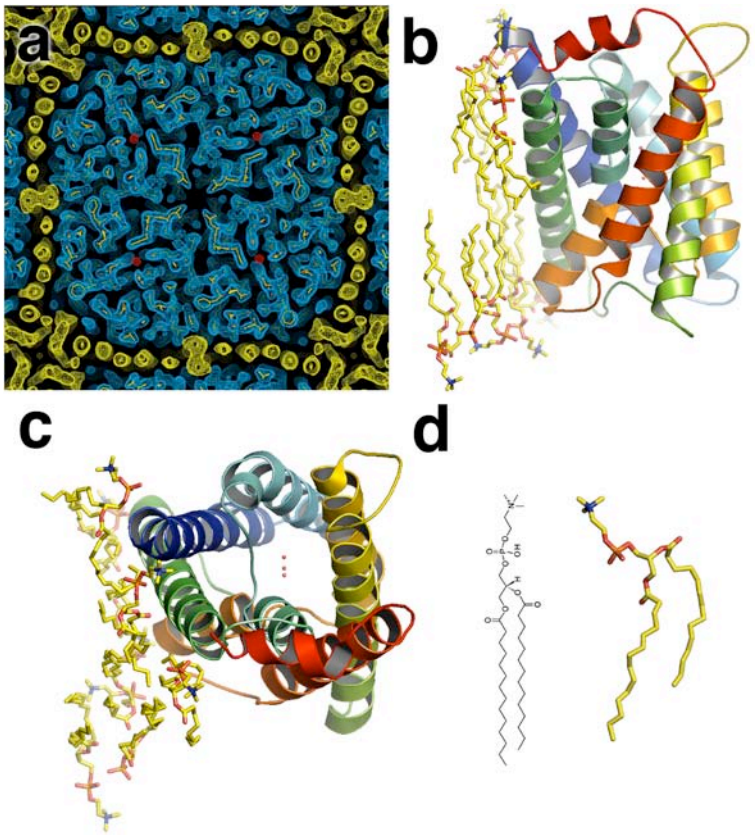

Figure 6 – The waterchannel Aqp0. a) A top-view of the cross-section through the membrane plane, showing the Aqp0 tetramer surrounded by lipids, which mediate the crystal contacts. Protein density in blue, with amino acid represented as sticks in yellow. The density for water molecules is red, and for the lipids, yellow. b) Side view of a Aqp0 monomer. Aqp0 is represented as a rainbow cartoon and lipids as yellow sticks. The three water molecules captured inside the cavity are represented as three red dots. c) Top view. d) A DMPC lipid as the ones resolved in the crystal structure is shown in classic and stick representations. Figure adapted from Gonen *et al.*, 2005.

REFERENCES

Agre, P., Lee, M. D., Devidas, S. & Guggino, W. B. (1997). Aquaporins and ion conductance. *Science* **275**(5305), 1490.

Bai, R., Covell, D. G., Pei, X. F., Ewell, J. B., Nguyen, N. Y., Brossi, A. & Hamel, E. (2000). Mapping the binding site of colchicinoids on beta -tubulin. 2-Chloroacetyl-2-demethylthiocolchicine covalently reacts predominantly with cysteine 239 and secondarily with cysteine 354. *J. Biol. Chem.* **275**(51), 40443-40452.

Chen, J. Z., Furst, J., Chapman, M. S. & Grigorieff, N. (2003). Low-resolution structure refinement in electron microscopy. *J. Struct. Biol.* **144**(1-2), 144-151.

Cheng, A., Leung, A., Fellmann, D., Quispe, J., Suloway, C., Pulokas, J., Abeyrathne, P. D., Lam, J. S., Carragher, B. & Potter, C. S. (2007). Towards automated screening of two-dimensional crystals. *J. Struct. Biol.* **160**(3), 324-331.

Crowther, R. A., Henderson, R. & Smith, J. M. (1996). MRC image processing programs. *J. Struct. Biol.* **116**(1), 9-16.

Denker, B. M., Smith, B. L., Kuhajda, F. P. & Agre, P. (1988). Identification, purification, and partial characterization of a novel Mr 28,000 integral membrane protein from erythrocytes and renal tubules. *J. Biol. Chem.* **263**(30), 15634-15642.

Downing, K. H. (1991). Spot-scan imaging in transmission electron microscopy. *Science* **251**(4989), 53-59.

Ganser, B. K., Cheng, A., Sundquist, W. I. & Yeager, M. (2003). Three-dimensional structure of the M-MuLV CA protein on a lipid monolayer: a general model for retroviral capsid assembly. *EMBO J.* **22**(12), 2886-2892.

Ganser-Pornillos, B. K., Cheng, A. & Yeager, M. (2007). Structure of full-length HIV-1 CA: a model for the mature capsid lattice. *Cell* **131**(1), 70-79.

Gipson, B., Zeng, X. & Stahlberg, H. (2007a). 2dx_merge: Data management and merging for 2D crystal images. *J. Struct. Biol.* **160**(3), 375-384.

Gipson, B., Zeng, X., Zhang, Z. Y. & Stahlberg, H. (2007b). 2dx--user-friendly image processing for 2D crystals. *J. Struct. Biol.* **157**(1), 64-72.

Glaeser, R. M. (1992). Specimen flatness of thin crystalline arrays: influence of the substrate. *Ultramic.* **46**(1-4), 33-43.

Gonen, T., Cheng, Y., Sliz, P., Hiroaki, Y., Fujiyoshi, Y., Harrison, S. C. & Walz, T. (2005). Lipid-protein interactions in double-layered two-dimensional AQP0 crystals. *Nature* **438**(7068), 633-638.

Grigorieff, N., Ceska, T. A., Downing, K. H., Baldwin, J. M. & Henderson, R. (1996). Electron-crystallographic refinement of the structure of bacteriorhodopsin. *J. Mol. Biol.* **259**(3), 393-421.

Gyobu, N., Tani, K., Hiroaki, Y., Kamegawa, A., Mitsuoka, K. & Fujiyoshi, Y. (2004). Improved specimen preparation for cryo-electron microscopy using a symmetric carbon sandwich technique. *J. Struct. Biol.* **146**(3), 325-333.

Hiroaki, Y., Tani, K., Kamegawa, A., Gyobu, N., Nishikawa, K., Suzuki, H., Walz, T., Sasaki, S., Mitsuoka, K., Kimura, K., Mizoguchi, A. & Fujiyoshi, Y. (2006). Implications of the aquaporin-4 structure on array formation and cell adhesion. *J. Mol. Biol.* **355**(4), 628-639.

Jap, B. K., Zulauf, M., Scheybani, T., Hefti, A., Baumeister, W., Aebi, U. & Engel, A. (1992). 2D crystallization: from art to science. *Ultramic.* **46**(1-4), 45-84.

Kubalek, E. W., Le Grice, S. F. & Brown, P. O. (1994). Two-dimensional crystallization of histidine-tagged, HIV-1 reverse transcriptase promoted by a novel nickel-chelating lipid. *J. Struct. Biol.* **113**(2), 117-123.

Kunji, E. R., von Gronau, S., Oesterhelt, D. & Henderson, R. (2000). The three-dimensional structure of halorhodopsin to 5 Å by electron crystallography: A new unbending procedure for two-dimensional crystals by using a global reference structure. *Proc. Natl. Acad. Sci. USA* **97**(9), 4637-4642.

Lambert, O., Moeck, G. S., Levy, D., Plancon, L., Letellier, L. & Rigaud, J. L. (1999). An 8-A projected structure of FhuA, A "ligand-gated" channel of the Escherichia coli outer membrane. *J. Struct. Biol.* **126**(2), 145-155.

Lebeau, L., Lach, F., Venien-Bryan, C., Renault, A., Dietrich, J., Jahn, T., Palmgren, M. G., Kuhlbrandt, W. & Mioskowski, C. (2001). Two-dimensional crystallization of a membrane protein on a detergent-resistant lipid monolayer. *J. Mol. Biol.* **308**(4), 639-647.

Lefman, J., Morrison, R. & Subramaniam, S. (2007). Automated 100-position specimen loader and image acquisition system for transmission electron microscopy. *J. Struct. Biol.* **158**(3), 318-326.

Levy, D., Chami, M. & Rigaud, J. L. (2001). Two-dimensional crystallization of membrane proteins: the lipid layer strategy. *FEBS Lett.* **504**(3), 187-193.

Levy, D., Mosser, G., Lambert, O., Moeck, G. S., Bald, D. & Rigaud, J. L. (1999). Two-dimensional crystallization on lipid layer: A successful approach for membrane proteins. *J. Struct. Biol.* **127**(1), 44-52.

Li, H., DeRosier, D. J., Nicholson, W. V., Nogales, E. & Downing, K. H. (2002). Microtubule structure at 8 A resolution. *Structure* **10**(10), 1317-1328.

Lowe, J., Li, H., Downing, K. H. & Nogales, E. (2001). Refined structure of alpha beta-tubulin at 3.5 A resolution. *J. Mol. Biol.* **313**(5), 1045-1057.

Makhov, A. M., Taylor, D. W. & Griffith, J. D. (2004). Two-dimensional crystallization of herpes simplex virus type 1 single-stranded DNA-binding protein, ICP8, on a lipid monolayer. *Biochim. Biophys. Acta* **1701**(1-2), 101-108.

Murata, K., Mitsuoka, K., Hirai, T., Walz, T., Agre, P., Heymann, J. B., Engel, A. & Fujiyoshi, Y. (2000). Structural determinants of water permeation through aquaporin-1. *Nature* **407**(6804), 599-605.

Nettles, J. H., Li, H., Cornett, B., Krahn, J. M., Snyder, J. P. & Downing, K. H. (2004). The binding mode of epothilone A on alpha,beta-tubulin by electron crystallography. *Science* **305**(5685), 866-869.

Nogales, E., Wolf, S. G. & Downing, K. H. (1997). Visualizing the secondary structure of tubulin: three-dimensional map at 4 A. *J. Struct. Biol.* **118**(2), 119-127.

Nogales, E., Wolf, S. G. & Downing, K. H. (1998a). Structure of the alpha beta tubulin dimer by electron crystallography. *Nature* **391**(6663), 199-203.

Nogales, E., Wolf, S. G. & Downing, K. H. (1998b). Structure of the alpha beta tubulin dimer by electron crystallography. *Nature* **391**(6663), 199-203.

Nogales, E., Wolf, S. G., Khan, I. A., Luduena, R. F. & Downing, K. H. (1995). Structure of tubulin at 6.5 A and location of the taxol-binding site. *Nature* **375**(6530), 424-427.

Palsdottir, H. & Hunte, C. (2004). Lipids in membrane protein structures. *Biochim. Biophys. Acta* **1666**(1-2), 2-18.

Philippsen, A., Schenk, A. D., Signorell, G. A., Mariani, V., Berneche, S. & Engel, A. (2007). Collaborative EM image processing with the IPLT image processing library and toolbox. *J. Struct. Biol.* **157**(1), 28-37.

Philippsen, A., Schenk, A. D., Stahlberg, H. & Engel, A. (2003). Iplt--image processing library and toolkit for the electron microscopy community. *J. Struct. Biol.* **144**(1-2), 4-12.

Preston, G. M., Carroll, T. P., Guggino, W. B. & Agre, P. (1992). Appearance of water channels in *Xenopus* oocytes expressing red cell CHIP28 protein. *Science* **256**, 385-387.

Ren, G., Cheng, A., Reddy, V., Melnyk, P. & Mitra, A. K. (2000). Three-dimensional fold of the human AQP1 water channel determined at 4 Å resolution by electron crystallography of two-dimensional crystals embedded in ice. *J. Mol. Biol.* **301**(2), 369-687.

Schenk, A. D., Werten, P. J., Scheuring, S., de Groot, B. L., Müller, S. A., Stahlberg, H., Philippsen, A. & Engel, A. (2005). The 4.5 A structure of human AQP2. *J. Mol. Biol.* **350**(2), 278-289.

Snyder, J. P., Nettles, J. H., Cornett, B., Downing, K. H. & Nogales, E. (2001). The binding conformation of Taxol in beta-tubulin: a model based on electron crystallographic density. *Proc. Natl. Acad. Sci. USA* **98**(9), 5312-5316.

Sui, H., Han, B. G., Lee, J. K., Walian, P. & Jap, B. K. (2001). Structural basis of water-specific transport through the AQP1 water channel. *Nature* **414**(6866), 872-878.

Typke, D., Gilpin, C. J., Downing, K. H. & Glaeser, R. M. (2007). Stroboscopic image capture: reducing the dose per frame by a factor of 30 does not prevent beam-induced specimen movement in paraffin. *Ultramicroscopy* **107**(2-3), 106-115.

Typke, D., Nordmeyer, R. A., Jones, A., Lee, J., Avila-Sakar, A., Downing, K. H. & Glaeser, R. M. (2005). High-throughput film-densitometry: an efficient approach to generate large data sets. *J. Struct. Biol.* **149**(1), 17-29.

Vink, M., Derr, K., Love, J., Stokes, D. L. & Ubarretxena-Belandia, I. (2007). A high-throughput strategy to screen 2D crystallization trials of membrane proteins. *J. Struct. Biol.* **160**(3), 295-304.

Vinothkumar, K. R., Smits, S. H. & Kühlbrandt, W. (2005). pH-induced structural change in a sodium/proton antiporter from *Methanococcus jannaschii*. *EMBO J.* **24**(15), 2720-2729.

Vonck, J. (2000). Parameters affecting specimen flatness of two-dimensional crystals for electron crystallography. *Ultramic.* **85**(3), 123-129.

Wang, W., Ding, J., Allen, E., Zhu, P., Zhang, L., Vogel, H. & Yang, Y. (2005). Gigaxonin interacts with tubulin folding cofactor B and controls its degradation through the ubiquitin-proteasome pathway. *Curr. Biol.* **15**(22), 2050-2055.

Wolf, S. G., Nogales, E., Kikkawa, M., Gratzinger, D., Hirokawa, N. & Downing, K. H. (1996). Interpreting a medium-resolution model of tubulin: comparison of zinc-sheet and microtubule structure. *J. Mol. Biol.* **262**(4), 485-501.

Zeng, X., Stahlberg, H. & Grigorieff, N. (2007). A maximum-likelihood approach to two-dimensional crystals. *J. Struct. Biol.* **160**(3), 362-374.