

# Cryo-electron microscopy of membrane proteins

**DOI:**
10.1016/j.ymeth.2018.04.018

**Document Version**
Accepted author manuscript



**Published in:**
Methods

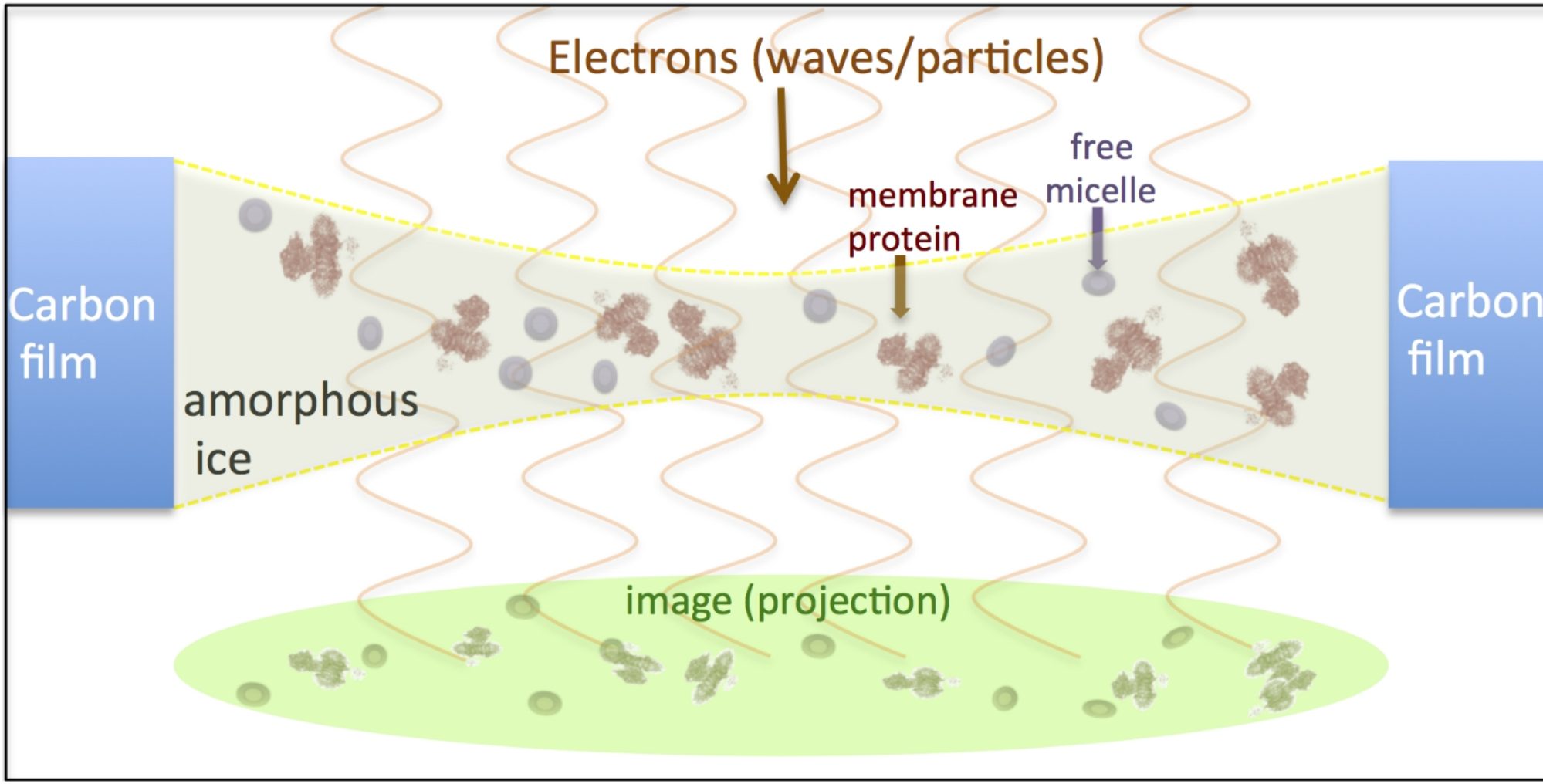


Highlights:

- Cryo-electron microscopy applied to membrane proteins.
- Different detergents can be tuned to the methodology.
- Weak contrast and the confounding effect of the micelle/nanodisc.
- Interpretation of medium resolution structures.

TOC Graphic

*Cryo-electron microscopy of membrane proteins. Electrons pass through the sample and are focused and imaged in the transmission electron microscope. The sample of membrane proteins is encapsulated in a microscopically thin layer of glassy ice. In most cases, free detergent micelles will be present as well as protein/detergent complexes.*


Electrons (waves/particles)
free micelle
membrane protein
Carbon film
amorphous ice
Carbon film
image (projection)

# Cryo-electron microscopy of membrane proteins.

Nopnithi Thonghin[1], Vasileios Kargas[1,2], Jack Clews[1] & Robert C. Ford[1,3]

1. School of Biological Sciences, Faculty of Biology Medicine and Health, The University of Manchester, Oxford Road, Manchester M13 9PL, UK.

2. Cambridge Institute for Medical Research, Wellcome Trust/MRC Building, Hills Road, Cambridge, CB2 0XY

3. To whom correspondence should be addressed: bob.ford@manchester.ac.uk

## Abstract

Membrane proteins represent a large proportion of the proteome, but have characteristics that are problematic for many methods in modern molecular biology (that have often been developed with soluble proteins in mind). For structural studies, low levels of expression and the presence of detergent have been thorns in the flesh of the membrane protein experimentalist. Here we discuss the use of cryo-electron microscopy in breakthrough studies of the structures of membrane proteins. This method can cope with relatively small quantities of sample and with the presence of detergent. Until recently, cryo-electron microscopy could not deliver high-resolution structures of membrane proteins, but recent developments in transmission electron microscope technology and in the image processing of single particles imaged in the microscope have revolutionized the field, allowing high resolution structures to be obtained. Here we focus on the specific issues surrounding the application of cryo-electron microscopy to the study of membrane proteins, especially in the choice of a system to keep the protein soluble.

Key words: Electron microscopy, membrane proteins, detergent, nanodisc, micelle protein structure.

Highlights:

- Cryo-electron microscopy applied to membrane proteins.
- Different detergents can be tuned to the methodology.
- Weak contrast and the confounding effect of the micelle/nanodisc.
- Interpretation of medium resolution structures.

TOC Graphic

*Cryo-electron microscopy of membrane proteins. Electrons pass through the sample and are focused and imaged in the transmission electron microscope. The sample of membrane proteins is encapsulated in a microscopically thin layer of glassy ice. In most cases, free detergent micelles will be present as well as protein/detergent complexes.*

## 1 Cryo-electron microscopy of biomolecules

This manuscript is written during a revolution in structural biology that has come about because of technical advances in cryo-electron microscopy (cryo-EM). It is not the aim of this article to recapitulate how this revolution has come about. Nor do we aim to offer a detailed review of the many recent advances in our understanding of membrane proteins and membrane protein complexes that have arisen because of these technical advances. Rather we aim to focus on the application of cryo-EM to membrane proteins, which have always been a thorn in the flesh of structural biologists. Nevertheless, some background on the methodology is needed in order to understand the specific problems faced by membrane protein biochemists if they wish to exploit cryo-EM for structure determination of their target protein. Readers wishing to know more about the cryo-EM technical advances and the history of the development of the method as applied to biomolecules may wish to look at articles describing the award of the 2017 Nobel Prize in medicine awarded to Richard Henderson, Joachim Frank and Jacques Dubochet (1). Readers interested in the general methodology for cryo-EM sample and grid preparation should see (2) which is also the most downloaded article in this journal. Most of the examples employed in this article are from published data deposited in the EMDB (electron microscopy data bank – *https://www.ebi.ac.uk/pdbe/emdb*), which is now linked to the PDB (protein data bank - *https://www.rcsb.org*). Some data shown is unpublished (c.f. Figure 1b,c and Figure 4c).

## 2 Background and theory:

This section can be skipped by cryo-EM specialists but is relevant for membrane protein biochemists wishing to exploit cryo-EM for a given membrane protein. Much of structural biology methodology operates in the realm of reciprocal (or inverse) space (3-7). This seems unnerving, but actually we already use this concept when we talk about the 'resolution' of a structure (with high resolution being what is desired). High frequency information is equivalent to high-resolution information. Any real-space 2D or 3D object can be represented in inverse space by a collection of different frequencies going from low to high. The

highest frequencies will found at the sampling limit of the data – i.e. if we can sample the object using – for example greyscale to represent the density of atoms within the object and with pixels (or their 3D equivalent – voxels) of size 1Å across, then the highest frequencies we can detect will have a wavelength of 2Å (because we need two adjacent pixels with different density values to generate the highest frequency). In order to represent the entire object at the highest possible resolution we need to work out the amplitude and phase of all the different frequencies (up to the limit) and also where they sit in relation to each other in 2D or 3D inverse space. The phase of the frequency is of course where the frequency wave starts relative to some arbitrary point in space. The conversion from real to inverse space is done computationally using the Fourier transform, which some readers may have learnt about and applied in school maths classes. An excellent resource for visualizing the Fourier transforms of objects (and their importance) can be found at http://www.ysbl.york.ac.uk/~cowtan/fourier/fourier.html. For reasons too arcane to explain here, it is much faster to process the data arising from structural biology using the inverse space representation of the objects rather than the real space one. For X-ray crystallography (8) and small-angle X-ray and neutron scattering(9), the raw data is already in an inverse space form. For cryo-EM there is an additional reason for using inverse space - the (real space) image data is affected by distortions generated in the process of cryo-EM data collection and these affect the phases and amplitudes of the frequencies in the image, hence they can best be appreciated, and must be corrected, in the inverse realm(10).

Cryo-EM offers a method to determine structures of biomolecules at various resolution (frequency) ranges and in an environment that is close to (or at least similar to) its native one. No crystallization is needed *a priori*, but if they are available, extremely thin crystals can also be studied using this method(7). Prior to imaging in the microscope, a purified and ideally, monodisperse, biomolecule sample is resuspended in a buffer and then rapidly frozen using liquid ethane such that hexagonal ice crystals do not have time to form(2). Instead, the water surrounding the biomolecule is frozen in an amorphous, glassy state. The sample

must then be kept under cryogenic conditions, below 100K, in order to maintain the amorphous state of the water(2). Subsequently the sample is imaged at similar or even lower temperatures in order to preserve the sample for as long as possible in the destructive electron beam. Images are created by phase contrast, and largely with electrons that are scattered elastically by the sample (4,11,12). This means that the images have a very weak contrast, especially for biological samples where one is aiming to detect differences in Coulomb scattering between the more scattering (on average) C, N, O, P, S, H atoms of the biomolecule and the lighter (on average) H and O atoms of the surrounding buffer.

### 2.1 Imaging and the implications for membrane proteins of the weak phase contrast in biological cryo-EM.

The weak electron scattering sets several limitations:

(i) Firstly, the biomolecule particle to be imaged should be sufficiently large to be identifiable in the low-contrast images obtained by cryo-EM(4,5). Arguments rage as to what the practical limit is for this, but for the purposes of this review we will arbitrarily choose 100kDa as a workable size, below which cryo-EM stands little chance of giving grounds for optimism of a high resolution structure. The reader is warned that this practical limit is broken in several exceptional cases already, and almost certainly this limit will be exceeded regularly in the future as the methodology continues to improve. For most membrane proteins, even monomers, their overall particle size will be >100kDa, simply because of the associated 'solubilisation device' that will surround it (typically a detergent micelle, but possibly a lipid bicelle with associated bicelle-forming proteins)(13).

(ii) Secondly, phase contrast and the ability to distinguish biomolecule particles from background buffer will depend on the composition of the sample buffer. If the latter contains too high a concentration of solute (e.g. salt, sucrose, glycerol), then the buffer atoms may scatter the electrons as much as the biomolecule and hence contrast will be lost, preventing structure determination. At first this seems like a trivial requirement, but one must remember that the cryo-EM revolution's typical client samples are difficult-to-purify proteins or protein complexes that may be

prone to dissociation, denaturation, aggregation and precipitation. Such specimens may need complex buffer components in order to maintain their integrity, solubility and monodispersity. Membrane proteins, of course, fall into this 'difficult-to-purify' category.

(iii) Because the contrast is low, the signal:noise ratio of images of individual particles / molecules is also very low, hence many thousands of images of particles need to be recorded and then these blurry images have to be averaged to improve the signal:noise ratio so that significant features in the structure can be determined. Before averaging, the images of the molecules first need to be classified and aligned. This is because the particles will be (more-or-less) randomly oriented in the frozen buffer and hence will represent different *projections (see below)* of the structure to be studied. It is easy to appreciate that the accuracy of the classification and alignment of the particles is crucial for the final resolution of the structure determined. If the different projections cannot be efficiently discriminated, then particles with different orientations relative to the electron beam will be erroneously averaged together, generating a blurred structure. Similarly, if particles that are correctly classified together cannot be accurately aligned (a rotational and translational search is needed) then the resolution of the final structure will be degraded.

For membrane proteins, one of the additional confounding factors for classification and alignment is the detergent micelle. For many detergents the Coulomb scattering of electrons by the detergent headgroup is strong (see later) and can be a major feature, clearly distinguishable in single particles of membrane proteins imaged by cryo-EM (13,14). The hydrocarbon chain region is usually weakly contrasted. The problem here is that the micelle is a much more deformable, plastic structure than the protein it contains. Micelle size and shape can vary and because of the dynamics of the detergent molecules at the instant that the sample was frozen, no individual micelle detergent molecule can be identified in a structure averaged from many thousands of individual protein/micelle complexes. This means that structurally, the micelle is disordered beyond a resolution of about 1-2nm.

The term 'projections' is used here rather than 'views' because the inelastically scattered electrons that are used for forming the image pass *through* the sample rather than bouncing off it. This is an important verbal distinction because, we, as humans, are used to viewing the structure of our environment by processing the photons of light that bounce off the objects we can see. We rarely deal with images where the photons of light pass through the object; thus cryo-EM can sometimes be counter-intuitive. Imagine, if you will, taking a cylindrically-shaped jelly mould, making some jelly at a 8% concentration in water in it, then after taking it out of the mould, encapsulating it in a cube of jelly at 7% concentration. Now you are going to view the cube in a dark room against the light coming from a small window. The cylindrical object will be very hard to distinguish, but the highest contrast projection of it will be when you view it end–on (i.e. parallel to the long axis). At this angle it will appear as if it was a circle. The orthogonal orientation to this will be even fainter, and will appear as a rectangle. Other orientations of the jelly cube will give somewhat elliptical projections of the cylinder, but the darkest/most contrasted area will still not be immediately recognizable as arising from a cylinder. Similarly in cryo-EM the projections of the particles are not always immediately interpretable to us as direct representations of a 3D object (but they are).

(iv) Phase contrast is used to obtain sufficiently contrasted images in cryo-EM(3). Because the unstained sample produces weak phase-shifts (see above) strong underfocus must be applied during the image collection. This in turn causes some predictable sinusoidal variations in the frequencies contained in the image, with certain frequencies being strongly reinforced whilst other frequencies may be completely suppressed (Figure 1). Moreover contrast in some frequency ranges may be reversed in terms of the phase, so that black appears white and *vice versa*. Nowadays, correction for these black/white phase distortions is facile. However correction for the distortion of the amplitudes of the image frequencies is difficult and when inefficient can sometimes lead to misinterpretation of the structure obtained. For example at a given defocus value often employed in cryo-EM, the first major amplitude boost could occur at frequencies around 0.02 $Å^{-1}$ (i.e. features of the size 50Å) whilst features in the image with a frequency of 0.005 $Å^{-1}$ (size of 200 $Å^{-1}$) may be strongly suppressed. The effect that this frequency distortion has on a

cryo-EM image of a protein of size roughly 200Å is to make it appear unnaturally hollow (Figure 1b). This is important to take into account for membrane proteins where the peripheral torus-shaped micelle may be emphasized by this frequency distortion at the expense of the central membrane protein components. There are various means of correcting for these frequency amplitude distortions, but perhaps the most reliable is to collect some small-angle X-ray scattering data for the sample if sufficient amounts with good monodispersity are available(3,15).

## 2.2 Single particle reconstruction

Single-particle reconstruction is the averaging of many two-dimensional images of individual molecules (single particles) in order to reconstruct the 3D structure of a protein. In single-particle cryo-EM the particles are (ideally) randomly distributed and oriented in the thin layer of vitreous ice (see above). Thus, different views (projections) of the particle are present and the dataset of images therefore needs to be sorted into the different projection classes. The more symmetry an object possesses, the fewer projection classes there will be (to arrive at a given resolution). Similarly, the larger the object, the more projection classes will be needed to arrive at the same resolution. During the sorting of the particles, every particle is aligned against different reference projections and classified based on the most similar one. Once sorted, all the particles of each class are averaged and then the classification and alignment procedure can be repeated with the new class averages as references. After a few iterations, this should produce more-or-less stable projection class averages and the projection angles that relate them to each other can be estimated. This geometry (Euler angle) calculation can take some computer time and be quite a rough approximation, but it should be sufficient to create a crude, *ab initio* 3D density map that can be employed as a start model. Now this start model can be used to generate a fresh set of projection references, and the classification and alignment and averaging of the particles can begin once more. After this step the class averages can be used to generate a much-improved 3D structure because now the projection angles are all known. This much-improved 3D structure then forms the basis for a new set of projection references used for a further round of classification, alignment and averaging….and so on,

until a stable 3D structure is produced with no further improvement in resolution (Wang and Sigworth, 2006). If a good homology model of the protein is available, then this can be converted into a 3D density map and employed as the start model instead of an *ab-initio* one, indeed the homology model could be used to aid the selection of particles from the electron micrographs. However homology models of membrane proteins will not incorporate the detergent micelle, hence some caution is needed when using this approach. Ideally some surrogate for the micelle could be added to the homology model, but this would require some insight into how the micelle is likely to behave and how strongly it will scatter electrons relative to the protein component (see later). The use of a model is becoming more and more important for automated particle selection because the datasets needed for high resolution structures are so large that it is no longer practical for humans to select all the particles (e.g. >1 million particles for a 4Å resolution structure). However the use of a starting model for particle selection and for the subsequent image processing inevitably means that there is a possibility for a self-selecting structural outcome because the process will be biased significantly by the starting model. Hence modern iterations of the standard cryo-EM image processing software packages aim to employ checks and balances to mitigate against model bias. An example of this and a visual representation of the latter stages of the 3D refinement procedure is depicted in Figure 2.

## 3. Membrane proteins imaged by cryo-EM.

Membrane proteins are typical examples of the cryo-EM clientelle: They are hard to express and purify and extremely challenging for NMR and X-ray crystallographic studies. Unfortunately for cryo-EM imaging they often need glycerol or high salt concentrations in order to maintain their solubility which suppresses contrast and signal:noise. Moreover membrane proteins need detergent in order to solubilize the protein from the membranes in which the protein is found. In this next section we will discuss the detergent micelle and what properties make a detergent well-suited to cryo-EM studies.

### 3.1 Detergents in cryo-EM studies (of membrane proteins)

(a) The primary consideration is that the detergent has to be sufficiently mild, in terms of its effects, to allow the structure of the native membrane protein or membrane protein complex to be preserved. Many detergents can efficiently solubilize proteins from membranes but they can be highly denaturing to proteins (the classic example being the anionic detergent sodium dodecyl sulphate, SDS). Once solubilized from the membrane, the protein will need the presence of detergent (or something similar – see later) in order to mask the hydrophobic, membrane-spanning portions of the protein from the solvent and hence prevent aggregation and precipitation during its purification. For these latter purposes, detergents that form micelles at very low concentrations in water (low c.m.c. detergents) are desirable because they can keep the protein soluble whilst minimizing the detergent concentration. This is also a very important factor for cryo-EM since scattering of electrons from free detergent micelles (i.e. micelles not containing a protein) can greatly complicate the process of structure determination. This probably explains why there are few examples of cryo-EM studies of proteins in high c.m.c. detergents (www.emdb.org) (16) but many examples of X-ray crystallography-derived structures in such detergents (a text search for 'octyl glucoside' in the PDB revealed 204 hits – www.rcsb.org).

The extraction of proteins from membranes is necessary in order to purify the protein for structural studies. During this procedure, the investigator makes a decision to compromise between yield and activity/stability of the protein. A 2014 study compared the effects of harsh and mild detergents on the thermal stability of an isolated soluble domain of a membrane protein. Zwitterionic and anionic detergents, despite having close to 100% extraction efficiency for the full-length protein, were shown to partially destabilise the soluble domain of the protein, with a reduction and broadening of the thermal unfolding transition measured by differential scanning calorimetry (17,18). The detergents with the least effect on the soluble domain were also the poorest at solubilisation of the microsomes containing the full-length protein(18). Pollock *et al* showed that DDM purified material retained more ATPase activity despite yielding half as much solubilised protein (19). Matar-Merheb *et al* also show the clear difference between mild (DDM) and harsh (FC12) detergents on both activity and yield in their study of novel

detergents. FC12, being a zwitterionic detergent has high extraction (solubilisation) efficiency however results in a completely inactive protein(20).

(b) Free micelles in low c.m.c. detergents: Even with the relatively low detergent concentrations present with low c.m.c. detergents, the effect of having some free micelles can still be an important consideration: For imaging, cryo-EM samples of membrane proteins will typically be suspended at about 1mg/ml (0.1% w/v) prior to being added to the grid and blotted. Hence even with a typical low c.m.c. detergent concentration in the buffer (e.g. at 0.05% w/v), there is likely to be more free detergent micelles than micelle/protein particles (depending on the relative masses of the free micelles versus the protein mass). Clearly, the micelle will usually be smaller than the membrane protein in hand, so it should be possible to distinguish protein/micelle particles from free micelles in the cryo-EM images, but even in this case, the free micelles will occasionally be associated with the protein/micelle particles or by chance happen to be in superposition with them relative to the electron beam. In both these cases the selection, classification and alignment of the affected particles would be affected.

(c) Surface activity: Detergents in chemistry labs go under the pseudonym of surfactants, and their surface activity is of huge importance in cryo-EM of membrane proteins. Cryo-EM ideally involves the trapping of the membrane proteins in a very thin layer of vitreous (glassy) ice. This layer will not be much thicker than the protein itself – perhaps 10-20nm in thickness or less. Such a thin layer is of course extremely fragile, hence needs to be confined to relatively small regions that bridge between stable supporting structures (the carbon film deposited onto the metal EM grid). These frozen suspension bridges that contain the biomolecule to be imaged will be typically spanning 1000- 2000nm, and will be formed by blotting away nearly all the sample and liquid added to the EM grid (with adsorbent filter paper) prior to rapid freezing. Surface tension holds the film of water in place in the short pause between blotting away excess liquid and freezing. Hence a surfactant can radically alter this process and lead to odd effects. The role of surfactants in changing surface interactions is also important. Ideally one wants the protein to interact minimally with the carbon surface (the support

structure) and distribute itself evenly in the thin water layer. In this case, the presence of detergent may help. In practice, all these factors come into play, so trial and error when preparing the grids is recommended. Certainly switching from one detergent to another will have a big influence on the blotting conditions and the distribution of protein/micelle particles over the grid.

### 3.2 Detergent alternatives

Some detergents reportedly generate difficulties for membrane protein research such as the alteration of their solubilization ability when environmental factors change. Besides, they are found to contribute to significant problems for the protein of interest – for example, structural disorder or reduced activities. These issues not only often hinder scientists wishing to study intact biological and biophysical properties of membrane proteins, but also potentially cause artifacts. To date, there are a variety of alternative tools that can be introduced to substitute conventional use of detergents. Below, widely-used detergent alternatives are discussed and some recent protein structures solved with particular platforms are presented.

(a) Amphipols: Amphipols or amphipathic polymers, have a structure that comprises hydrophilic backbones and hydrophobic side chains. They have been developed to overcome some drawbacks of detergents. The key concept of amphipols is that the molecules would have a high affinity to hydrophobic surfaces and scarcely dissociate from them over time, allowing a membrane protein to remain stabilized and solubilized in aqueous environments. These properties were clearly shown to potentially provide stabilization roles in the light-driven membrane transporter bacteriorhodopsin from *Haloacterium salinarum* compared to conventional detergents (21). Some amphipols are also tolerant to environmental changes, such as pH and calcium concentrations, that can be advantageous for a variety of studies (22). Although an increasing number of protein structures have been determined by cryo-EM with amphipols – for example, murine endo-lysosomal TRPL1 channel protein structures (Figure 3), there are still reports regarding issues caused by the substance including polymer aggregation and scaffold-function interference (23,24).

(b) Nanodiscs: A nanodisc is another potential platform that facilitates membrane protein research in both biochemical and biophysical areas. Nanodiscs comprise a cluster of small patches of lipids in a bilayer that are encircled by amphipathic proteins that form a rim shielding the hydrophobic lipid tails from the aqueous environment. The sizes of the nanodiscs can be specified by the rim protein's size. Membrane proteins, which must be firstly purified in detergents, are reconstituted into a nanodisc platform containing the chosen lipid compositions. This is achieved by removing the detergent – usually by addition of hydrophobic polystyrene beads. After reconstitution, this allows the system to closely mimic the protein's native conditions, hence functionalities and activities can be better recovered than they can in detergent environments (22). This advantage of the nanodisc system can be appreciated in a recent study of the MsbA transporter where a number of cryo-EM-derived MsbA structures have revealed conformational states in its mechanism of action typified by the alternating access hypothesis (Figure 3) (25,26).

(c) SMA-lipid particles (SMALPs): Styrene maleic acid co-polymer (SMA) is a chemical tool that has been found to be beneficial to membrane protein studies (27). This polymer possesses an amphipathic property arising from its repeating hydrophobic styrene and hydrophilic maleic acid moieties that allow the substance to be able to insert into biological lipid membranes and consequently extract integral membrane proteins within a small patch of natural lipids (27,28). The platform may be described as being similar to that of nanodiscs, there are, however, differences to some extent. Instead of using proteins to enclose lipid patches, SMA polymer is employed. In fact, SMALPs also allow membrane proteins to be extracted and purified without any need for detergents; potentially enabling proteins to retain their full structural features and natural functionalities (27,29). The polymer has previously been used in other fields especially in cosmetics and in pharmaceutical research. For example, studies on G-protein-coupled receptors

(GPCRs) using SMALPs show significant retention of pharmacological properties whereas those of detergent-exposed protein were dramatically sacrificed. This was found to facilitate downstream drug discovery research involving the particular proteins (30). The recent success of SMALPs in cryo-EM studies can be illustrated by a study of the *Escherichia coli* multidrug efflux transporter AcrB protein which represented the first sub-nanometer-resolution reconstruction of a membrane protein using the SMALPs platform (Figure 4) (31).

## 4. Influence of the micelle or nanodisc in cryo-EM of membrane proteins.

In this section we consider the consequences of having a membrane protein embedded in a micelle or nanodisc, initially from the viewpoint of its impact on imaging and high resolution structure determination. Later we discuss the more contentious issue of whether the solubilisation agent itself influences the structure of the protein.

### 4.1 Effect of the micelle or nanodisc on structure determination.

AS discussed above, the micelle or nanodisc components surrounding the membrane protein will have an influence on the cryo-EM studies(13). Firstly, a positive influence may be enjoyed where a clearly-delineated micelle (or nanodisc) may aid in the single particle classification and alignment, especially for smaller membrane protein complexes. On the other hand, the scattering of the micelle and its contribution to the cryo-EM derived map may mask protein in the transmembrane segments. To some extent, this problem can be overcome by visualizing the density map at high density thresholds, especially at higher resolution realms (Fig. 3a). However high resolution is not always achieved and differentiation of the protein and non-protein density is not always easy. Notably, scattering by the micelle (and lipids in discs) is particularly strong for the hydrophobic headgroup regions, and is weak for the hydrocarbon chains. This gives the micelles and discs a shell-like form with a hollow internal appearance. Hence differentiation of protein versus micelle/nanodisc density is particularly

difficult at the extremities of the membrane-spanning regions but much easier for the central membrane-spanning regions (Fig. 3b; Fig. 4 a,b). Comparisons of different detergents and nanodisc systems are displayed in Figures 3 and 4. Micelle size varies considerably with different detergents, and this seems to be directly related to the size and composition of the detergent hydrophilic headgroup region compared to its hydrophobic tail. Digitonin, with its five monosaccharide moieties versus a compact sterol tail displays a particularly large micelle which is very well defined in cryo-EM images (Figure 4 a,b) (32,33). In comparison, DDM (dodecyl-β-D-maltopyranoside with two monosaccharide moieties) is smaller and less well defined in the density maps of proteins from the same family (Fig. 4b). Charged detergents such as LDAO (n-dodecyl-N,N-dimethylamine-N-oxide, Fig. 4a) and LPG14 (lyso-phosphatidylglycerol-14, Fig 4c) appear to have very small, weakly scattering micelles (14). Hence if a protein is stable in a charged detergent, this may represent an advantage for cryo-EM studies at higher resolution.

An extensive study has examined the same protein in different detergents (14): In this study of the N-type ATPase rotor (c-ring – see Fig. 4a), the two long membrane-spanning helices of the protein (repeated 17 times in the ring) could only be clearly resolved in the zwitterionic detergent LDAO (n-dodecyl-N,N-dimethylamine-N-oxide) which has a very small charged headgroup. In another detergent, DDM (n-dodecyl-β-D-maltopyranoside with a large, uncharged headgroup) the internal regions of the membrane-spanning helices could be discerned, but where the helices passed through the micelle headgroup region, discrimination between protein and detergent was not possible. In a non-ionic polyoxyethylene headgroup detergent with 8 oxyethylene moieties and a 12-hydrocarbon tail, C12E8, the situation was worse, whilst in the A8-35 amphipol resolution was too low to identify individual helices(14). A further possible rationale for the weaker micelle density in the headgroup region for charged detergents such as LDAO and LPG14 has come from a neutron-scattering study of such micelles(35). Here it was noted that water was inserted extensively into the charged headgroup region of micelles formed from LPC (lyso-phosphatidylcholine) detergents in water. This water infiltration may greatly reduce contrast in this region to almost background levels. Another comparison is shown in Fig. 4c. with

the Wzz protein (34) in two different detergents, DDM and another lyso-lipid detergent lyso-phosphatidyl glycerol 14 (LPG14). Despite resolution differences in the two reconstructions (due to different numbers of particles selected and different microscopes utilised for data collection), the overall structure of the protein component is the same. However the DDM micelle surrounding the two transmembrane helices is well defined (right, projection) whereas the LPG14 micelle is small and scatters electron weakly (left, surface render).

### 4.2 Does the micelle or nanodisc affect the structure?

This manuscript has not, so far, addressed the issue of whether the detergent or nanodisc may inadvertently influence the structures of membrane proteins. Clearly, the propensity of detergents to be chaotropic means that there is the ever-present possibility that parts of a membrane protein's structure can be denatured by the presence of detergent, even where an overall structure determination is possible. Ironically, the flexibility of single particle cryo-EM and its tolerance of disordered regions within a structure may make this unwanted situation more likely to crop up in such studies. One must also consider the possibility that the presence of a detergent micelle may adventitiously organize a region of the protein that would normally be disordered. These are considerations that bear further study, but are outside the scope of this article. Additionally, we must remember that detergents are complex molecules that may interact directly with membrane proteins, even as substrates or inhibitors. For example, studies of detergents interacting with P-glycoprotein (ABCB1) have shown that they can be substrate-like molecules well below their critical micelle concentration. (36,37)

Commented [1]: Perhaps this paragraph can go at the end of section 3 or 4 as they both detail the effects of detergent micelles and nanodiscs on protein conformation. Regarding Review comment 5

### 4.3 Detergent micelles, nanodiscs and protein oligomerisation state.

Similarly to above, we must also address the highly contentious area of whether membrane protein *quaternary* structure is influenced by detergent and/or nanodiscs. Some would argue that even having a high resolution structure for a membrane protein in a specific oligomeric state is not the unequivocal statement

on its oligimeric form, and indeed the example used in Figure 4c illustrates this: Wzz as studied in two detergents (LPG and DDM) displays a dodecameric organization(34). However when reconstituted into lipid membranes, it takes up a hexameric form(38) (also interpreted as potentially octameric(39)). When the large periplasmic portion of the protein is expressed separately, without the transmembrane and cytoplasmic portions, it can adopt a range of oligomers that form well-diffracting 3D crystals and these oligomeric forms include pentamers, octamers and nonamers(40). Another example is the purified CorA channel that shows a strict C5 symmetrical arrangement when crystallised in the presence of $Mg^{2+}$ (41), but a later cryo-EM study (42)showed that it loses this symmetry in low Mg conditions, with each subunit displaying a different configuration (also later confirmed at higher resolution(43)). Hence having a high-resolution structure for an oligomeric membrane protein does not necessarily establish that its quaternary structure reflects the situation in the native membrane. An open-minded approach is recommended.

## 5. Interpretation of 3D Coulomb density maps of membrane proteins.

Having arrived at a cryo-EM derived density map, the membrane protein biochemist will want to maximise the usefulness of the available data. In this section we explore various means by which this can be achieved. The resolution of a 3-D structure, derived by single-particle analysis, depends on both the number and the quality of particle images. For instance, a medium resolution structure of a membrane protein (~10 Å) can be obtained by using several thousand (usually more than 10,000) particle images. In a medium-resolution density map of 8-10Å resolution, α-helices are represented as cylinders, whilst β-strands cannot be differentiated(3). As most membrane proteins are typified by long membrane-spanning α-helices, one can appreciate why the study of membrane proteins by cryo-EM is favourable – even at medium resolution. Such a structure will provide important information about the packing of the membrane-spanning helices that could be crucial for understanding the functioning of e.g. a channels or a transporter. Furthermore such medium resolution data can be used to refine

homology models of membrane proteins (see below). When resolution from cryo-EM of membrane proteins approaches that obtainable by X-ray crystallography, then interpretation of the entire structure becomes possible. For cryo-EM maps at 4 Å resolution the pitch of the α-helices can be discerned and larger side-chains of amino acids identified and the β-strands can be differentiated in β-sheet regions. However, such high-resolution structures for membrane proteins are still relatively rare. A recent search of the EMDB for 'membrane' yielded 435 deposited maps of membrane proteins but only 33 examples had better than 4Å resolution and only 122 better than 8Å resolution. Hence despite the revolution in cryo-EM recorded in the EMDB, at least nine out of ten studies of membrane proteins still result in maps where additional means of interpretation need to be applied.

A way to interpret the cryo-EM maps, especially for medium to low resolution maps obtained for membrane proteins, is to dock available crystal structures or models of the protein of interest into the density map to obtain an atomic-detail representation of the macromolecular assembly. This procedure can reveal conformational changes of the structure as a function of the experimental conditions. Moreover, uncharacterized domains can be indicated in the density map (as an excess density to the docked model/structure) providing important structural insights of the location and interaction interfaces. This process will usually unambiguously identify the detergent micelle in the map (if it hasn't already been done). Docking (or fitting) can be performed either by treating the atoms of the protein model as a rigid object or by allowing it to act as a flexible object.

Rigid-body fitting methods apply three rotational and three translational degrees of freedom in the structure to extensively search the density area of the map and achieve the optimal fit (44,45). The optimization process is based on the cross-correlation coefficient (CCC) score, which ranges from 0 to 1 and indicates how well the atomic structure is fitted in the electron density map.

### 5.1 Flexible fitting approaches

The idea of improving the current rigid-body fitting approaches was prompted by the significant higher-resolution density maps obtained by the cryo-EM revolution. The first flexible fitting methods were based on splitting the high-resolution structures into many "rigid bodies" and fitting them individually into the map. Several other approaches were reported utilizing different methods such as: a combined vector quantization and molecular mechanics approach (46), normal modes approaches (47,48), elastic network techniques (49), Monte Carlo simulation approaches (50). However, after 2007, molecular dynamics simulations have been used in the flexible fitting approaches, such as Flex-EM (51) and Molecular Dynamics Flexible Fitting (MDFF) (52). Flex-EM performs the fitting in two stages. Initially it performs rigid-body fitting using a hierarchical approach and an extensive Monte Carlo search, followed by the refinement stage that uses minimization and simulated annealing molecular dynamics simulation to improve the CCC score (51). On the other hand, MDFF utilizes the EM map as an external potential to pull the high-resolution structure into high-density areas (52,53). The latter potential is defined on a grid where all the information about high- and low-density areas is mapped (52). MDFF uses the NAMD software (54) and its gridForces feature (55), where grid-steered molecular dynamics (56) are applied to force each atom of the fitted molecule to move into high density areas based on the gradient of the potential derived from the density map at the respective atomic positions (52). The stereochemical quality of the structure is maintained by applying additional harmonic restraints to the φ and ψ dihedral angles of the alpha helices and beta sheets. The introduction of these restraints ensures that the overall secondary structure is somewhat preserved and over-fitting is being prevented (53). This is particularly suitable for membrane proteins where many helices are inevitably constrained by the requirement to span the 5nm-thick lipid bilayer membrane, whilst different rotations and packing compared to homologs will be present. Finally, flexible fitting is considered "completed" when the RMSD or the CCC values plateau over simulation time, typically within 300 to 500 ps (52,53). In a recent study of the Wzz protein (Figure 4c, (34)) flexible fitting was performed using the MDFF method with an improvement of the CCC from 0.82 calculated between the map and the original homology model to 0.92 after MDFF.

### 5.2 Macromolecular refinement and model validation

The number of structures obtained from cryo-EM maps with better than 4Å resolution is constantly increasing. Manual model building and fitting approaches may lead to stereochemical violations or classes. Such models can further be refined thanks to dedicated software developed initially for refining structures derived by X-ray crystallography, such as REFMAC and PHENIX (Acta Cryst. (2011). D67, 355-367 | Acta Cryst. (2010). D66, 213-221). REFMAC uses Maximum likelihood and some elements of Bayesian Statistics to aid the refinement in reciprocal-space. In order to obtain reliable models (with resolution above 3 Å), a range of refinement tools are employed within REFMAC, which use several restraints to drive refinement, such as secondary structure, Non Crystallographic Symmetry (NCS), 'jelly-body' and long range restraints on atomic displacement parameters. In addition, if the interatomic distance of the target structure matches with a known homologous structure, the latter can then drive refinement of the target structure within the cryo-EM map. The interatomic distance restraints can be generated using the ProSMART software (57). This method may reduce the chance of overfitting, which can be assessed in the validation section of REFMAC. REFMAC is open source and a graphical user interface (GUI) is available via the CCPEM (http://www.ccpem.ac.uk) and CCP4 (http://www.ccp4.ac.uk) suites (58).

In contrast to REFMAC, PHENIX refinement is mostly performed in real space (59). The *phenix.real_space_refine* tool contains several options that can be employed during refinement, including but not limited to rigid-body refinement, simulating annealing, use of restraints and morphing. PHENIX has its own dedicated GUI with extension modules to visualize the results in Coot (60) or PyMOL (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC).

Finally, the quality of the pre- and post-refined models can be assessed using the all-atom structure validation web server MolProbity (61). Validation analysis includes all-atom contact analysis, Ramachandran and rotamer analysis and backbone bond-length and bond-angle outliers. MolProbity generates a multi-criterion chart containing all the outliers which can be visualized in Coot for further manual intervention.

## 6. An example: The cystic fibrosis transmembrane conductance regulator (CFTR)

CFTR is a chloride channel that has evolved from the ATP-binding cassette transporter family. After many years of effort aimed at obtaining a structure for the protein, there are now three cryo-EM derived structures at better than 4Å resolution for the protein, two in the dephosphorylated, quiescent state with no nucleotide, one in the phosphorylated, activated state with nucleotide bound (32,62,63). For these ground-breaking studies, the protein was originally solubilized using one non-ionic detergent (2,2-didecylpropane-1,3-bis-β-D-maltopyranoside (LMNG)), then switched into a more suitable detergent (digitonin) for purification and optimisation of its secondary modifications for cryo-EM. Finally the purified and polished protein was mixed with a third zwitterionic detergent (fluorinated Fos-Choline-8) immediately prior to sample preparation for cryo-EM. This illustrates how carefully one must optimise detergents for a structural biology project on a membrane protein; and how it is becoming increasingly important to be able to switch from one detergent system to another (c.f. section 3).

The cryo-EM maps show a strongly-scattering digitonin/ fluorinated Fos-Choline-8 mixed-micelle (c.f. section 4, Figure 5a). Note that the protein regions do not show the expected high-resolution features until the higher frequency components of the map are enhanced in reciprocal space (c.f. section 2 and compare Fig, 5a and b). This enhancement is a later-stage processing tool termed 'sharpening' and involves the application of a global scaling factor (b-factor) to the map(64). Note that some extraneous noise is also inevitably enhanced by this procedure. As noted by the authors, after sharpening it is apparent that the membrane-spanning portions of the protein appear to be better defined than the two soluble nucleotide-binding domains (NBDs), probably because these two domains show some additional wagging-like mobility in the structure that degrades the resolution in this region of the protein (c.f. section 2). Figure 5 panels c, d & e illustrate this differential resolution over the protein and the effects this has on the interpretation of the map: Fitting of the polypeptide chain was possible for the membrane-spanning regions; rigid-body fitting on known NBD structures was used for the NBD regions (32,62). However in the structure with nucleotide bound, the NBD regions are less mobile

and here the polypeptide chain could be fitted to the NBDs (63), illustrating the importance of flexibility/mobility in determining local resolution (c.f. sections 2,5). Figure 5 also illustrates the power of cryo-EM to provide information on regions of the protein/micelle complex that probably do not have any obvious tertiary structure but occupy roughly the same region of 3D space. In CFTR the detergent micelle (Fig. 5f, green) dominates in regions of the map not accounted for by the atomic model, but in addition a weaker density region between the NBDs can be observed (red) that was interpreted as a regulatory region that probably has no fixed tertiary structure and has to have alternative conformations dependent on its phosphorylation by protein kinase A (PKA) and C (PKC)(65).

## 7. Conclusions

Cryo-EM of membrane proteins has been part of the revolution in detector electron microscopy technology and software, and has revealed the structures of many membrane proteins that had previously proven recalcitrant to alternative means of structure determination (such as X-ray crystallography and NMR spectroscopy). For the latter two methods, it is already well established that certain detergents are suited to the methodology. For X-ray crystallography, detergents that form micelles that are small and deformable enough to be accommodated within the crystalline lattice are needed (66). For NMR, detergents that form very small micelles around the membrane proteins are required so that the overall rotation time for the protein/micelle complex does not lead to excessive broadening of the NMR resonances(67,68). Here we have described the properties of detergents and other solubilisation systems for cryo-EM of membrane proteins, and reviewed evidence that this methodology has increased efficiency with certain types of solubilising agents (32). For this methodology, the opportunity to use larger membrane-mimetic systems is also offered(13). For high resolution cryo-EM structures of membrane proteins (better than 4Å), the considerations over solubilizing agents are less important, but for a more typical situation (where the project is generating medium resolution data (12-4Å), then the type of solubilizing

agents may need to be considered. Similarly we anticipate that switching between different detergents and from detergent to nanodisc systems will be increasingly important for cryo-EM studies of membrane proteins.

Acknowledgements: We wish to thank Dr Stephen Prince (School of Biological Sciences, University of Manchester) and Prof Anna Seelig (Biozentrum, University of Basel) for useful discussions. RCF's research is sponsored by the UK Cystic Fibrosis Trust and the USA Cystic Fibrosis Foundation.

Figure Legends:

Figure 1: CTF correction and application of a structure factor file: (a) Sketch of the contrast transfer function (CTF – black dashed line) which determines the distortions of the phases and amplitudes at different spatial frequencies due to defocus. Where the phase is reversed (upper half -Amp 'x') this must be corrected for prior to class averaging and 3D reconstruction. At crossover points in the CTF function, amplitudes approach zero, hence correction around these points is tricky and may result in more noise at specific frequencies. At low frequencies, the overall shape and size of the particle dominates the structure factor amplitudes (solid blue line). Poor correction of amplitudes at low frequencies may result in an abnormally hollow 3D reconstruction (red dashed line & panel b) with central features suppressed, or, conversely a 'blobby' reconstruction with peripheral features suppressed (red dotted line, panel c). A projection of the same 3D object (a membrane protein ion channel) is shown in panels b and c.

Figure 2: An illustration of old and new strategies in image processing and single particle cryo-EM: One commonly-used software package is depicted using a flow diagram, starting in each case at the top with 'all' (all particles selected for the 3D reconstruction). The newer version of the procedure incorporates steps which mitigate against the possibility of model-bias imposed by the use of a start model for the 3D refinement. The new strategy involves splitting the dataset into two and then processing each half dataset separately starting with two slightly different start models (with their structures randomised beyond a certain resolution). The start model can be derived from a homology model (suitably smeared to low resolution) or can be generated *ab-initio* by comparing projection classes generated by a model-free classification and alignment procedure (see main text).

Figure 3: Cryo-EM of membrane proteins. (a) Endo-lysosomal TRPL1 channel protein (EMD-6823) in amphipol A8-35. Strong structural features such as transmembrane helices (yellow transparent box) can be easily distinguished from the amphipol at a high density threshold (left) allowing a molecular model to be built (right).

(b) Cryo-EM maps of the bacterial ABC transporter MsbA reconstituted into nanodiscs. 3D reconstructions reveal inward-facing (left) and outward-facing (right) conformations of the protein induced in the nucleotide-free (*left,* EMD-8469) and nucleotide-trapped (ADP-vanadate, *right,* EMD-8467) states. Interestingly the conformational switch has a major effect on the imaging of the nanodisc. In the nucleotide-free state (left) the boundaries of the nanodisc are clear and resemble a detergent micelle (see other Figures), whereas for the outward-facing (right), the nanodisc is weakly defined, with only density close to the transmembrane helices.

Figure 4: Properties of protein-encapsulating detergent micelles and nanodiscs detected by cryo-EM.

(a) Four different types of membrane protein are shown, using projections through the density maps to give an idea of the relative scattering intensity coming from various parts of the structures. Toroidal density for the micelle/nanodisc is indicated by the arrows. Digitonin shows a very large micelle with the secretase protein, with strong scattering density arising from the carbohydrate polar headgroups (69) although the same protein studied in amphipols gave similar results (70,71). The LDAO micelle with its small diamine oxide headgroup is very weakly scattering(14). The nanodiscs formed by SMALP (24) and amphipol (72) show an intermediate size, although the SMALP scattering appears to be strong compared to amphipol. The scale bars represent 2nm.

(b) Comparison of surface-rendered views of the same type of membrane protein (ABC transporter family) in two different detergents. The DDM micelle is smaller and is less defined than the large digitonin micelle. The bacterial TmrAB structure was obtained in the presence of an $F_{AB}$ antibody fragment in order to break the two-fold pseudo-symmetry (73). TAP1/2(74) and CFTR(32) are ABC transporters from eukaryotes.

(c) Comparison of the same membrane protein (Wzz) in two different detergents. Central slices are shown along the 12-fold symmetry axis. The polypeptide chain of a homology model of the full-length protein fitted by MDFF to the DDM-Wzz complex is shown on the left as a blue ribbon trace(34). Despite resolution differences in the two reconstructions, the overall structure of the protein component is the same. However the DDM micelle surrounding the two

transmembrane helices is well defined (right, projection) whereas the LPG14 micelle is small and scatters electron weakly (left, surface render). The scale bar represents 5nm.

Figure 5: CFTR as an illustration of cryo-EM of membrane proteins. CFTR map before (a) and after (b) enhancement of high frequency components ('sharpening'). Note that the pitch of membrane-spanning (micelle-spanning) helices is apparent after sharpening, whilst the secondary structure in the lower nucleotide-binding domains remains indistinct. (c) Transmembrane (TM) helices 6-8 in the map (mesh) with the fitted model superimposed (ribbon). Individual bulky sidechains such as R353 can be readily identified whilst others such as F866 cannot, possibly because of mobility. (d) In contrast, the NBD region is less well resolved and the orientation of crucial catalytic residues (K1251, E1372) cannot be gleaned. (e) The same view as (d) but with the theoretical map (blue mesh) calculated for the model at 3.7Å resolution. (f) Illustration of lower-density regions of the map (green/red surface) not fitted by the model (yellow ribbon trace). The red surface in the map was interpreted as arising from the regulatory (R) region that is thought to be mostly disordered.

References


1. Nogales, E. (2017) Profile of Joachim Frank, Richard Henderson, and Jacques Dubochet, 2017 Nobel Laureates in Chemistry. *Proc Natl Acad Sci U S A*
2. Thompson, R. F., Walker, M., Siebert, C. A., Muench, S. P., and Ranson, N. A. (2016) An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology. *Methods* **100**, 3-15
3. Ford, R. C., and Holzenburg, A. (2008) Electron crystallography of biomolecules: mysterious membranes and missing cones. *Trends Biochem Sci* **33**, 38-43
4. Henderson, R. (2004) Realizing the potential of electron cryo-microscopy. *Q Rev Biophys* **37**, 3-13
5. Saibil, H. R. (2000) Macromolecular structure determination by cryo-electron microscopy. *Acta Crystallogr D Biol Crystallogr* **56**, 1215-1222
6. Crowther, R. A., Amos, L. A., Finch, J. T., De Rosier, D. J., and Klug, A. (1970) Three dimensional reconstructions of spherical viruses by fourier synthesis from electron micrographs. *Nature* **226**, 421-425
7. Amos, L. A., Henderson, R., and Unwin, P. N. (1982) Three-dimensional structure determination by electron microscopy of two-dimensional crystals. *Prog Biophys Mol Biol* **39**, 183-231
8. Cowtan, K., Emsley, P., and Wilson, K. S. (2011) From crystal to structure with CCP4. *Acta Crystallogr D Biol Crystallogr* **67**, 233-234
9. Boldon, L., Laliberte, F., and Liu, L. (2015) Review of the fundamental theories behind small angle X-ray scattering, molecular dynamics simulations, and relevant integrated application. *Nano Rev* **6**, 25661
10. Crowther, R. A., and Klug, A. (1975) Structural analysis of macromolecular assemblies by image reconstruction from electron micrographs. *Annu Rev Biochem* **44**, 161-182
11. Nogales, E., and Scheres, S. H. W. (2015) Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity. *Mol Cell* **58**, 677-689
12. Sigworth, F. J. (2016) Principles of cryo-EM single-particle image processing. *Microscopy (Oxf)* **65**, 57-67
13. Mio, K., and Sato, C. (2017) Lipid environment of membrane proteins in cryo-EM based structural analysis. *Biophys Rev*
14. Schulz, S., Wilkes, M., Mills, D. J., Kuhlbrandt, W., and Meier, T. (2017) Molecular architecture of the N-type ATPase rotor ring from Burkholderia pseudomallei. *EMBO Rep* **18**, 526-535
15. McDevitt, C. A., Shintre, C. A., Grossmann, J. G., Pollock, N. L., Prince, S. M., Callaghan, R., and Ford, R. C. (2008) Structural insights into P-glycoprotein (ABCB1) by small angle X-ray scattering and electron crystallography. *FEBS Lett* **582**, 2950-2956
16. Scheffer, M. P., Gonzalez-Gonzalez, L., Seybert, A., Ratera, M., Kunz, M., Valpuesta, J. M., Fita, I., Querol, E., Pinol, J., Martin-Benito, J., and Frangakis, A. S. (2017) Structural characterization of the NAP; the major adhesion complex of the human pathogen Mycoplasma genitalium. *Mol Microbiol* **105**, 869-879
17. Yang, Z., Wang, C., Zhou, Q., An, J., Hildebrandt, E., Aleksandrov, L. A., Kappes,

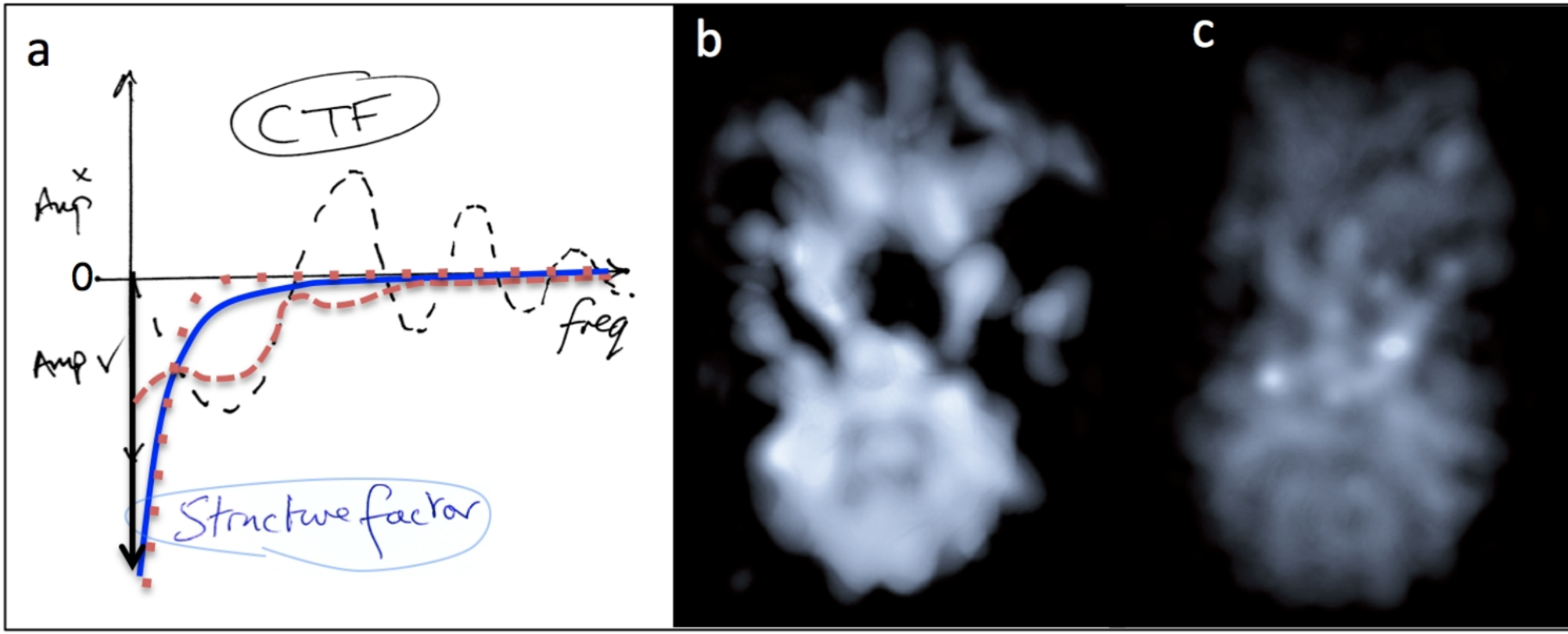
J. C., DeLucas, L. J., Riordan, J. R., Urbatsch, I. L., Hunt, J. F., and Brouillette, C. G. (2014) Membrane protein stability can be compromised by detergent interactions with the extramembranous soluble domains. *Protein Sci* **23**, 769-789

18. Hildebrandt, E., Zhang, Q., Cant, N., Ding, H., Dai, Q., Peng, L., Fu, Y., DeLucas, L. J., Ford, R., Kappes, J. C., and Urbatsch, I. L. (2014) A survey of detergents for the purification of stable, active human cystic fibrosis transmembrane conductance regulator (CFTR). *Biochim Biophys Acta* **1838**, 2825-2837
19. Pollock, N., Cant, N., Rimington, T., and Ford, R. C. (2014) Purification of the cystic fibrosis transmembrane conductance regulator protein expressed in Saccharomyces cerevisiae. *J Vis Exp*
20. Matar-Merheb, R., Rhimi, M., Leydier, A., Huche, F., Galian, C., Desuzinges-Mandon, E., Ficheux, D., Flot, D., Aghajari, N., Kahn, R., Di Pietro, A., Jault, J. M., Coleman, A. W., and Falson, P. (2011) Structuring detergents for extracting and stabilizing functional membrane proteins. *PLoS One* **6**, e18036
21. Bazzacco, P., Billon-Denis, E., Sharma, K. S., Catoire, L. J., Mary, S., Le Bon, C., Point, E., Banères, J.-L., Durand, G., Zito, F., Pucci, B., and Popot, J.-L. (2012) Nonionic Homopolymeric Amphipols: Application to Membrane Protein Folding, Cell-Free Synthesis, and Solution Nuclear Magnetic Resonance. *Biochemistry* **51**, 1416-1430
22. Popot, J.-L. (2010) Amphipols, Nanodiscs, and Fluorinated Surfactants: Three Nonconventional Approaches to Studying Membrane Proteins in Aqueous Solutions. *Annual Review of Biochemistry* **79**, 737-775
23. Zhang, S., Li, N., Zeng, W., Gao, N., and Yang, M. (2017) Cryo-EM structures of the mammalian endo-lysosomal TRPML1 channel elucidate the combined regulation mechanism. *Protein & Cell* **8**, 834-847
24. Picard, M., Dahmane, T., Garrigos, M., Gauron, C., Giusti, F., le Maire, M., Popot, J.-L., and Champeil, P. (2006) Protective and Inhibitory Effects of Various Types of Amphipols on the Ca2+-ATPase from Sarcoplasmic Reticulum: A Comparative Study. *Biochemistry* **45**, 1861-1869
25. Mi, W., Li, Y., Yoon, S. H., Ernst, R. K., Walz, T., and Liao, M. (2017) Structural basis of MsbA-mediated lipopolysaccharide transport. *Nature* **549**, 233
26. Jardetzky, O. (1966) Simple allosteric model for membrane pumps. *Nature* **211**, 969-970
27. Postis, V., Rawson, S., Mitchell, J. K., Lee, S. C., Parslow, R. A., Dafforn, T. R., Baldwin, S. A., and Muench, S. P. (2015) The use of SMALPs as a novel membrane protein scaffold for structure study by negative stain electron microscopy. *Biochim Biophys Acta* **1848**, 496-501
28. Gulati, S., Jamshad, M., Knowles, T. J., Morrison, K. A., Downing, R., Cant, N., Collins, R., Koenderink, J. B., Ford, R. C., Overduin, M., Kerr, I. D., Dafforn, T. R., and Rothnie, A. J. (2014) Detergent-free purification of ABC (ATP-binding-cassette) transporters. *Biochemical Journal* **461**, 269-278
29. Knowles, T. J., Finka, R., Smith, C., Lin, Y.-P., Dafforn, T., and Overduin, M. (2009) Membrane Proteins Solubilized Intact in Lipid Containing Nanoparticles Bounded by Styrene Maleic Acid Copolymer. *Journal of the American Chemical Society* **131**, 7484-7485
30. Wheatley, M., Charlton, J., Jamshad, M., Routledge, S. J., Bailey, S., La-Borde, P. J., Azam, M. T., Logan, R. T., Bill, R. M., and Dafforn, T. R. (2016) GPCR–styrene maleic acid lipid particles (GPCR–SMALPs): their nature and

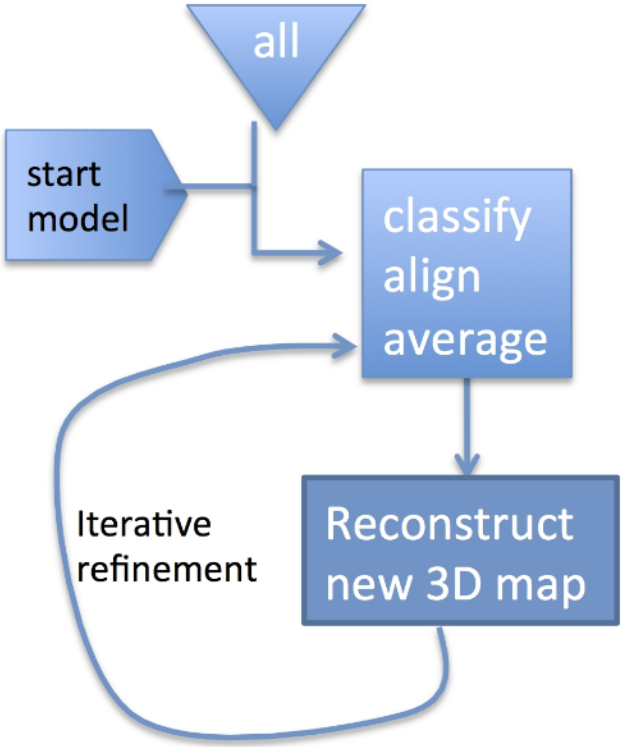
potential. *Biochemical Society Transactions* **44**, 619-623
31. Parmar, M., Rawson, S., Scarff, C. A., Goldman, A., Dafforn, T. R., Muench, S. P., and Postis, V. L. G. (2018) Using a SMALP platform to determine a sub-nm single particle cryo-EM membrane protein structure. *Biochimica et biophysica acta* **1860**, 378-383
32. Zhang, Z., and Chen, J. (2016) Atomic Structure of the Cystic Fibrosis Transmembrane Conductance Regulator. *Cell* **167**, 1586-1597 e1589
33. Lu, P., Bai, X. C., Ma, D., Xie, T., Yan, C., Sun, L., Yang, G., Zhao, Y., Zhou, R., Scheres, S. H. W., and Shi, Y. (2014) Three-dimensional structure of human gamma-secretase. *Nature* **512**, 166-170
34. Collins, R. F., Kargas, V., Clarke, B. R., Siebert, C. A., Clare, D. K., Bond, P. J., Whitfield, C., and Ford, R. C. (2017) Full-length, Oligomeric Structure of Wzz Determined by Cryoelectron Microscopy Reveals Insights into Membrane-Bound States. *Structure* **25**, 806-815 e803
35. Pambou, E., Crewe, J., Yaseen, M., Padia, F. N., Rogers, S., Wang, D., Xu, H., and Lu, J. R. (2015) Structural Features of Micelles of Zwitterionic Dodecyl-phosphocholine (C(1)(2)PC) Surfactants Studied by Small-Angle Neutron Scattering. *Langmuir* **31**, 9781-9789
36. Li-Blatter, X., Beck, A., and Seelig, A. (2012) P-glycoprotein-ATPase modulation: the molecular mechanisms. *Biophys J* **102**, 1383-1393
37. Li-Blatter, X., Nervi, P., and Seelig, A. (2009) Detergents as intrinsic P-glycoprotein substrates and inhibitors. *Biochim Biophys Acta* **1788**, 2335-2344
38. Larue, K., Ford, R. C., Willis, L. M., and Whitfield, C. (2011) Functional and Structural Characterization of Polysaccharide Co-polymerase Proteins Required for Polymer Export in ATP-binding Cassette Transporter-dependent Capsule Biosynthesis Pathways. *J Biol Chem* **286**, 16658-16668
39. Kalynych, S., Cherney, M., Bostina, M., Rouiller, I., and Cygler, M. (2015) Quaternary structure of WzzB and WzzE polysaccharide copolymerases. *Protein Sci* **24**, 58-69
40. Tocilj, A., Munger, C., Proteau, A., Morona, R., Purins, L., Ajamian, E., Wagner, J., Papadopoulos, M., Van Den Bosch, L., Rubinstein, J. L., Fethiere, J., Matte, A., and Cygler, M. (2008) Bacterial polysaccharide co-polymerases share a common framework for control of polymer length. *Nat Struct Mol Biol* **15**, 130-138
41. Lunin, V. V., Dobrovetsky, E., Khutoreskaya, G., Zhang, R., Joachimiak, A., Doyle, D. A., Bochkarev, A., Maguire, M. E., Edwards, A. M., and Koth, C. M. (2006) Crystal structure of the CorA Mg2+ transporter. *Nature* **440**, 833-837
42. Cleverley, R. M., Kean, J., Shintre, C. A., Baldock, C., Derrick, J. P., Ford, R. C., and Prince, S. M. (2015) The Cryo-EM structure of the CorA channel from Methanocaldococcus jannaschii in low magnesium conditions. *Biochim Biophys Acta* **1848**, 2206-2215
43. Matthies, D., Dalmas, O., Borgnia, M. J., Dominik, P. K., Merk, A., Rao, P., Reddy, B. G., Islam, S., Bartesaghi, A., Perozo, E., and Subramaniam, S. (2016) Cryo-EM Structures of the Magnesium Channel CorA Reveal Symmetry Break upon Gating. *Cell* **164**, 747-756
44. Fabiola, F., and Chapman, M. S. (2005) Fitting of high-resolution structures into electron microscopy reconstruction images. *Structure* **13**, 389-400
45. Wriggers, W., and Chacon, P. (2001) Modeling tricks and fitting techniques

for multiresolution structures. *Structure* 9, 779-788
46. Wriggers, W., and Birmanns, S. (2001) Using situs for flexible and rigid-body fitting of multiresolution single-molecule data. *J Struct Biol* 133, 193-202
47. Tama, F., Miyashita, O., and Brooks, C. L., 3rd. (2004) Normal mode based flexible fitting of high-resolution structure into low-resolution experimental data from cryo-EM. *J Struct Biol* 147, 315-326
48. Suhre, K., Navaza, J., and Sanejouand, Y. H. (2006) NORMA: a tool for flexible fitting of high-resolution protein structures into low-resolution electron-microscopy-derived density maps. *Acta Crystallogr D Biol Crystallogr* 62, 1098-1100
49. Schroder, G. F., Brunger, A. T., and Levitt, M. (2007) Combining efficient conformational sampling with a deformable elastic network model facilitates structure refinement at low resolution. *Structure* 15, 1630-1641
50. Jolley, C. C., Wells, S. A., Fromme, P., and Thorpe, M. F. (2008) Fitting low-resolution cryo-EM maps of proteins using constrained geometric simulations. *Biophys J* 94, 1613-1621
51. Topf, M., Lasker, K., Webb, B., Wolfson, H., Chiu, W., and Sali, A. (2008) Protein structure fitting and refinement guided by cryo-EM density. *Structure* 16, 295-307
52. Trabuco, L. G., Villa, E., Mitra, K., Frank, J., and Schulten, K. (2008) Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. *Structure* 16, 673-683
53. Trabuco, L. G., Villa, E., Schreiner, E., Harrison, C. B., and Schulten, K. (2009) Molecular dynamics flexible fitting: a practical guide to combine cryo-electron microscopy and X-ray crystallography. *Methods* 49, 174-180
54. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26, 1781-1802
55. Wells, D. B., Abramkina, V., and Aksimentiev, A. (2007) Exploring transmembrane transport through alpha-hemolysin with grid-steered molecular dynamics. *J Chem Phys* 127, 125101
56. Israelewitz, B., Gao, M., and Schulten, K. (2001) Steered molecular dynamics and mechanical functions of proteins. *Curr Opin Struct Biol* 11, 224-230
57. Nicholls, R. A., Long, F., and Murshudov, G. N. (2012) Low-resolution refinement tools in REFMAC5. *Acta Crystallogr D Biol Crystallogr* 68, 404-417
58. Wood, C., Burnley, T., Patwardhan, A., Scheres, S., Topf, M., Roseman, A., and Winn, M. (2015) Collaborative computational project for electron cryo-microscopy. *Acta Crystallogr D Biol Crystallogr* 71, 123-126
59. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66, 213-221
60. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60, 2126-2132
61. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010)

MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21
62. Zhang, Z., Liu, F., and Chen, J. (2017) Conformational Changes of CFTR upon Phosphorylation and ATP Binding. *Cell* **170**, 483-491 e488
63. Liu, F., Zhang, Z., Csanady, L., Gadsby, D. C., and Chen, J. (2017) Molecular Structure of the Human CFTR Ion Channel. *Cell* **169**, 85-95 e88
64. Grigorieff, N. (2016) Frealign: An Exploratory Tool for Single-Particle Cryo-EM. *Methods Enzymol* **579**, 191-226
65. Bozoky, Z., Krzeminski, M., Muhandiram, R., Birtley, J. R., Al-Zahrani, A., Thomas, P. J., Frizzell, R. A., Ford, R. C., and Forman-Kay, J. D. (2013) Regulatory R region of the CFTR chloride channel is a dynamic integrator of phospho-dependent intra- and intermolecular interactions. *P Natl Acad Sci USA* **110**, E4427-E4436
66. Ostermeier, C., and Michel, H. (1997) Crystallization of membrane proteins. *Curr Opin Struct Biol* **7**, 697-701
67. Lipfert, J., Columbus, L., Chu, V. B., Lesley, S. A., and Doniach, S. (2007) Size and shape of detergent micelles determined by small-angle X-ray scattering. *J Phys Chem B* **111**, 12427-12438
68. Oliver, R. C., Lipfert, J., Fox, D. A., Lo, R. H., Doniach, S., and Columbus, L. (2013) Dependence of micelle size and shape on detergent alkyl chain length and head group. *PLoS One* **8**, e62488
69. Sun, L., Zhao, L., Yang, G., Yan, C., Zhou, R., Zhou, X., Xie, T., Zhao, Y., Wu, S., Li, X., and Shi, Y. (2015) Structural basis of human gamma-secretase assembly. *Proc Natl Acad Sci U S A* **112**, 6003-6008
70. Bai, X. C., Rajendra, E., Yang, G., Shi, Y., and Scheres, S. H. (2015) Sampling the conformational space of the catalytic subunit of human gamma-secretase. *Elife* **4**
71. Bai, X. C., Yan, C., Yang, G., Lu, P., Ma, D., Sun, L., Zhou, R., Scheres, S. H. W., and Shi, Y. (2015) An atomic structure of human gamma-secretase. *Nature* **525**, 212-217
72. McGoldrick, L. L., Singh, A. K., Saotome, K., Yelshanskaya, M. V., Twomey, E. C., Grassucci, R. A., and Sobolevsky, A. I. (2018) Opening of the human epithelial calcium channel TRPV6. *Nature* **553**, 233-237
73. Kim, J., Wu, S., Tomasiak, T. M., Mergel, C., Winter, M. B., Stiller, S. B., Robles-Colmanares, Y., Stroud, R. M., Tampe, R., Craik, C. S., and Cheng, Y. (2015) Subnanometre-resolution electron cryomicroscopy structure of a heterodimeric ABC exporter. *Nature* **517**, 396-400
74. Oldham, M. L., Hite, R. K., Steffen, A. M., Damko, E., Li, Z., Walz, T., and Chen, J. (2016) A mechanism of viral immune evasion revealed by cryo-EM analysis of the TAP transporter. *Nature* **529**, 537-540

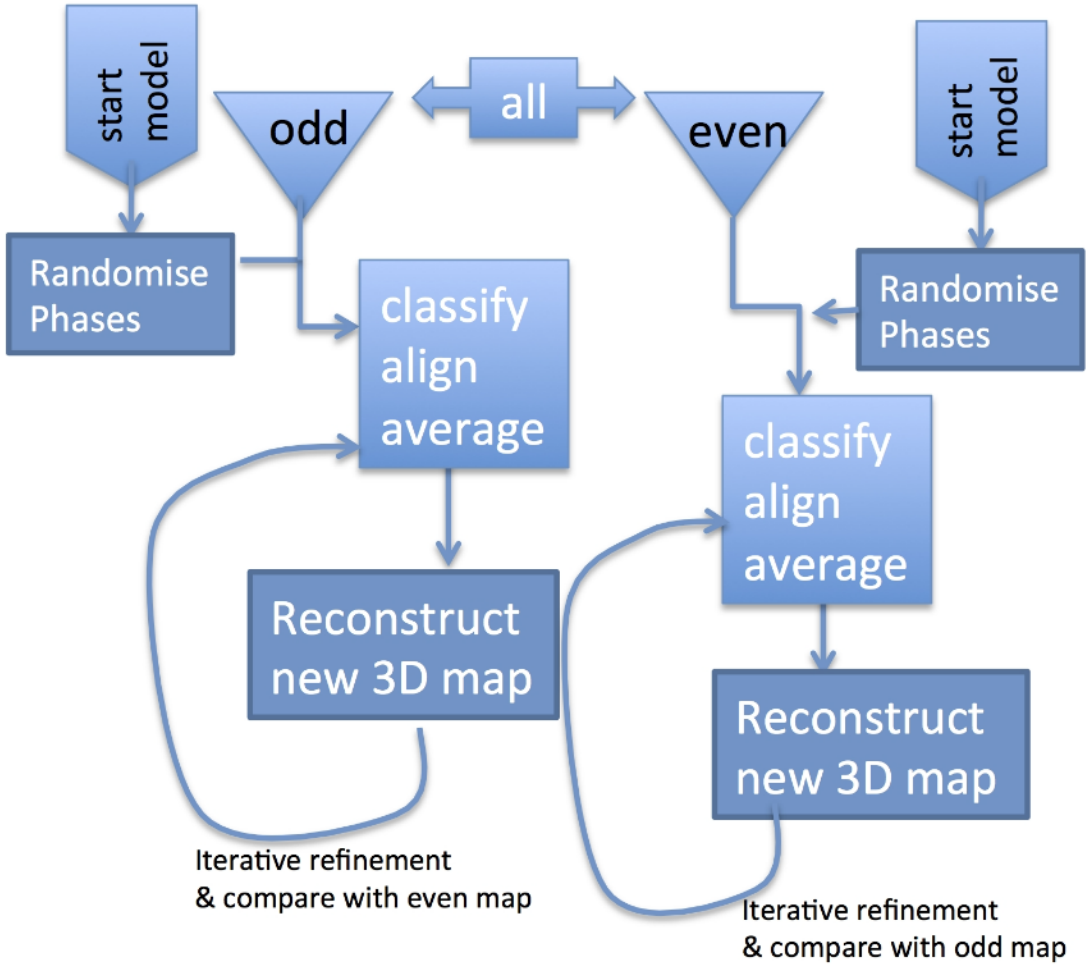

a
CTF
Amp
0
freq
Amp
Structure factor
b
c

Old Software – Version 1.9 of EMAN

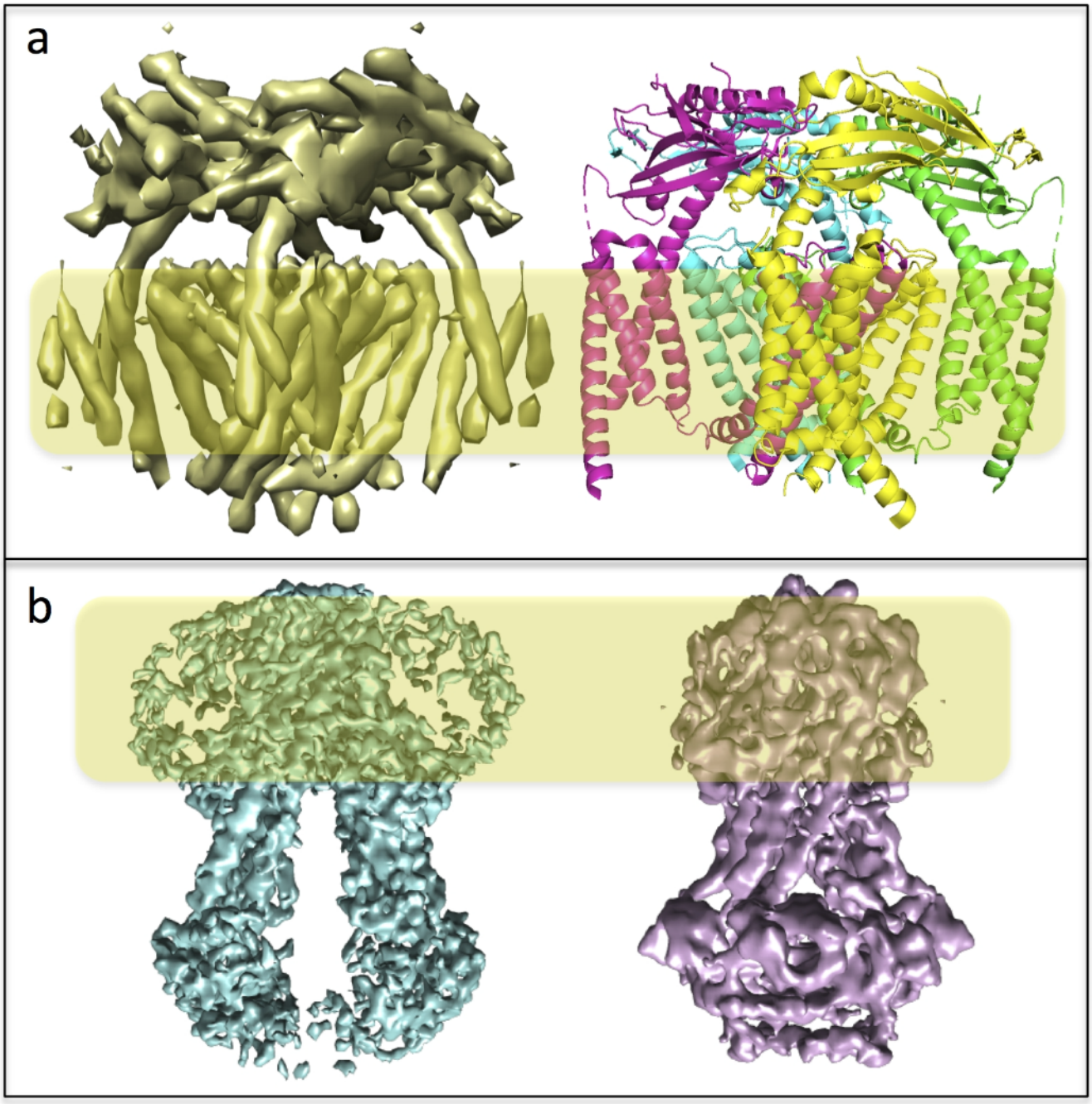


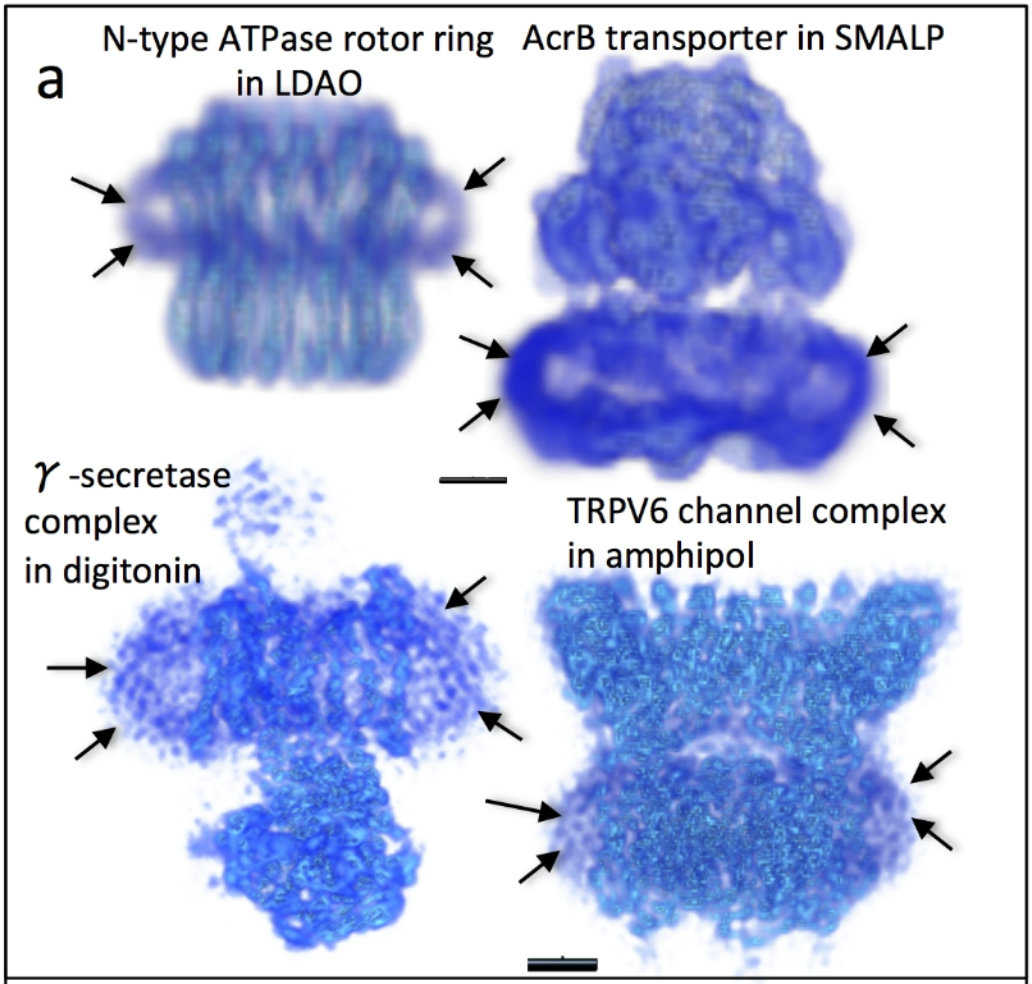

***Tendency for model bias***

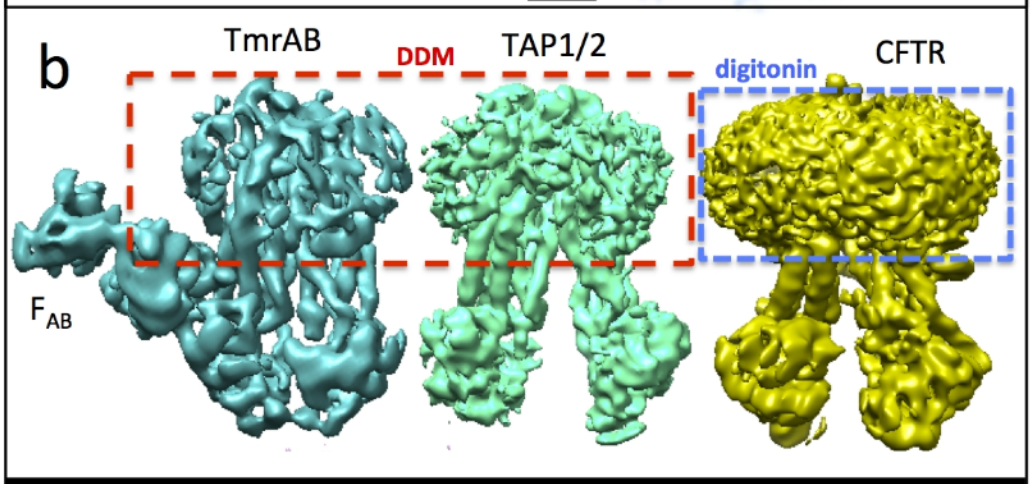
New Software – Version 2.2 of EMAN
"Gold standard"

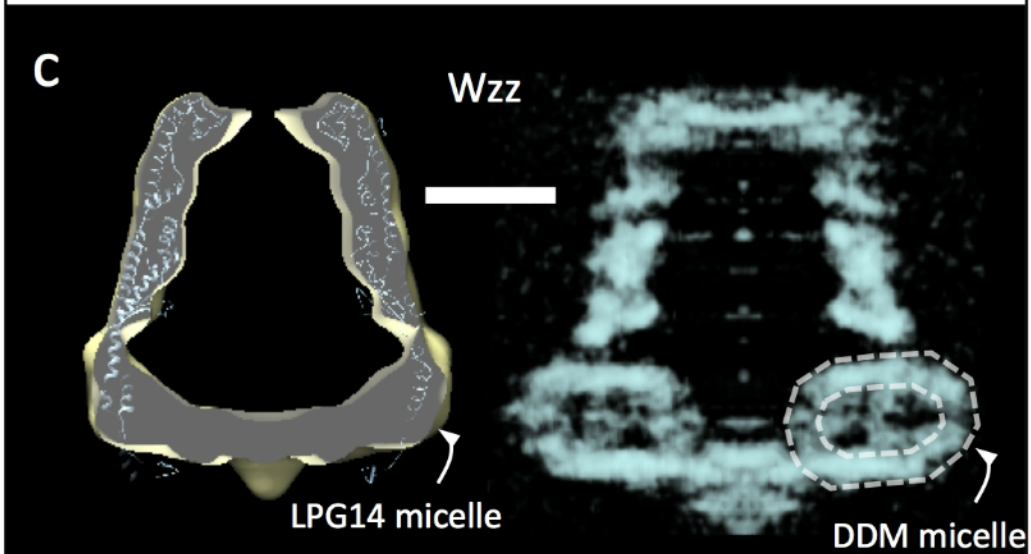


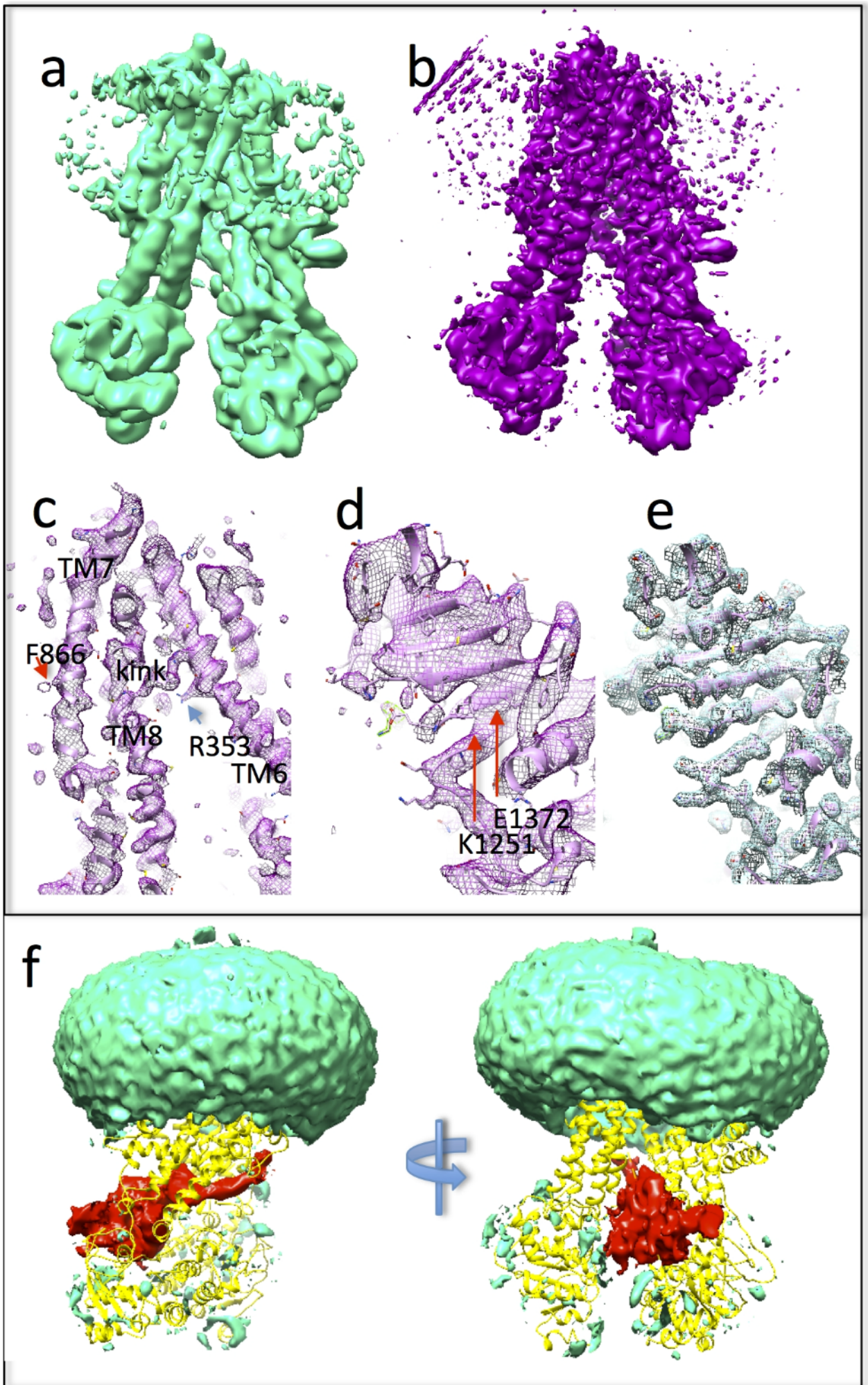

***Model bias less; need 2x number of particles***


a
b


a
N-type ATPase rotor ring in LDAO
AcrB transporter in SMALP
γ -secretase complex in digitonin
TRPV6 channel complex in amphipol


b
TmrAB
DDM
TAP1/2
digitonin
CFTR
F_AB


c
Wzz
LPG14 micelle
DDM micelle


a
b
c
TM7
F866
kink
TM8
R353
TM6
d
E1372
K1251
e
f