Structural Studies of a Urea Channel with Electron Microscopy

By Nathan Chan

Submitted for the degree of Doctor of Philosophy

Krebs Institute for Biomolecular Research,
Department of Molecular Biology and Biotechnology
The University of Sheffield

December 2012

Acknowledgements

Foremost, I'd like to thank my supervisor Per Bullough for the opportunity to work on such a challenging project and his help and guidance throughout my PhD.

There are many people from the E36 clique who have made a huge difference to my time at Sheffield. I'd also like to thank Svetomir Tsokov for being a great teacher in the use of the electron microscope, knower of many useful things and his generosity with his time. I'd like to thank Lekshmi Kailas for much useful discussion and help with atomic force microscopy and Masao Yamashita for useful scientific discussions on membrane proteins and for introducing me to the joys of sake (consumed responsibly of course). I'd like to thank Cassandra Terry, Chris Glover and Wezz Booth in particular for the guidance with lab work and also to Wan Qiang (Joel), David Radford, Shuo Jiang, Tham Janganan and Caitlyn Brumsby for making E36 a great place to work.

I thank my fellow MPSI collaborators especially people at the University of Leeds for providing the membrane proteins to play with. In particular I'd like to thank Steve Baldwin, Sarah Deacon, Gerard Huysman and Vincent Postis for the scientific discussion and for being great people to work with.

I have received much support from my friends throughout my time at Sheffield. I'm very grateful for the many adventures, work distractions and our mostly tuneful music making. If you find yourself reading this, you'll know who you are.

Last and perhaps most of all I'd like to thank my family for the love and support during my PhD pursuits. I think I'm *almost* ready to find a real job now.

Also, thanks to the BBSRC for providing the funds for distribution at my discretion ☺

Abstract

The Urea/Amide channel from *Bacillus cereus* (UAC$_{Bc}$) was expressed in *Escherichia coli* with a C-terminal hexa-histidine tag. The protein was purified in detergent as confirmed by N-terminal sequencing. The purified protein in detergent was analysed with single particle analysis processing and forms a particle consisting of a pair of stacked discs with diameters of 120 Å with each disc representing an oligomer of UAC$_{Bc}$.

Two-dimensional (2D) crystallisation produced highly aggregated crystals that became suitable for high resolution imaging upon sonication to disperse them.

Using the 2D crystals for electron cryomicroscopy yielded images that upon crystallographic processing and analysis suggested that the crystals had *p*6 symmetry with an additional single *p*622 crystal indicating a possible double-layered crystal form.

The images with *p*6 symmetry were merged to produce a 9 Å projection map showing the protein forming a hexameric ring with 7 density features in each putative monomer possibly representing the predicted 7 transmembrane helices of UAC$_{Bc}$.

AFM and production of a negative stain three dimensional (3D) density map were used to determine the thickness of the crystals and based on a mono-layered crystal form, bioinformatic analysis and biochemical experiments to verify the oligomeric state and topology, a model with the putative locations of the 7 predicted transmembrane helices and their orientations with respect to each other has been produced.

Abbreviations and Symbols

2D	Two Dimensional
3D	Three Dimensional
Å	Angstrom
AFM	Atomic force microscopy
ATCC	American type culture collection
ATP	Adenosine triphosphate
Bc	*Bacillus cereus*
BCA	Bicinchoninic acid assay
BSA	Bovine serum albumin
CMC	Critical micellar concentration
CTF	Contrast transfer function
DM	*n*-decyl-β-D-maltopyranoside
DDM	*n*-dodecyl-β-D-maltopyranoside
DMPC	1,2-dimyristoyl-*sn*-glycero-3-phosphocholine
DOPC	1,2-dioleoyl-*sn*-glycero-3-phosphocholine
DSPC	1,2-distearoyl-*sn*-glycero-3-phosphocholine
*dv*UT	Urea transporter from *Desulfovibrio vulgaris*
EDTA	Ethylenediaminetetraacetic acid
EM	Electron microscopy
ETL	*E. coli* total lipid extract
eV/keV	Electron volt/ kiloelectron volt
FFT	Fast Fourier transform
GPCR	G protein coupled receptor
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HMM	Hidden Markov model
*Hp*UreI	Urea channel from *Helicobacter pylori*
IQ	Image quality value
kDa	Kilodalton

LHCII	Light harvesting complex II
LPR	Lipid to protein ratio
MES	2-(*N*-morpholino)ethanesulfonic acid
MPSI	Membrane protein structure initiative
MRC	Medical Research Council
MTSET	[2-(trimethylammonium)ethyl] methanethiosulfonate bromide
Ni-NTA	Nickel-nitriloacetic acid
OGM	Oregon Green® maleimide
POPC	1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine
QVAL	A weighted sum of the IQ values
SEC	Size exclusion chromatography
SEC-MALLS	Size exclusion chromatography multi-angle laser light scattering
SDS-PAGE	SDS polyacrylamide gel electrophoresis
TEM	Transmission electron microscopy
TEV	Tobacco etch virus
UAC	Urea/amide channel
UAC_{Bc}	Urea/amide channel from *Bacillus cereus*
UT	Urea transporter

Table of Contents

Chapter 1 - Introduction

1.1 Membrane Proteins

Cells are enclosed by a lipid bilayer that acts as a semi-permeable barrier between the cell and its environment. Embedded into the membrane are a class of proteins known as integral membrane proteins (Singer and Nicolson, 1972) that are involved in several crucial processes including transport of materials into and out of cells, communication and energy transduction. Membrane proteins account for an estimated 30% of open reading frames (Wallin and von Heijne, 1998) and some of the most prescribed drugs are targeted at membrane proteins including fluoxetine targeted at the serotonin transporter and omeprazole at the proton pump (le Coutre and Kaback, 2000). However, despite their abundance and importance there are few high resolution structures available and structure determination lags far behind that of soluble proteins. Progress in structural determination of membrane proteins is often hindered at several stages. Over-expression of membrane proteins which are usually naturally inabundant is often toxic to cells creating difficulties in obtaining sufficient amounts of protein for structural studies (Grisshammer and Tate, 1995). Difficulties may also be encountered at later stages in purifying sufficient quantities of the correctly folded membrane protein of interest and then preparing X-ray diffraction quality crystals.

1.2 Overview of the Membrane Protein Structure Initiative

The MPSI was a consortium of groups at the Universities of Glasgow, Manchester, Leeds, Sheffield, Oxford, London (Imperial and Birkbeck College) and Daresbury laboratory working on developing high throughput techniques for the expression, purification, characterisation, crystallisation and structure determination of integral membrane proteins. The University of Sheffield node was focused on the use of electron microscopy to study membrane proteins. Primarily this involved two-dimensional crystallisation of promising target proteins for high resolution imaging by electron cryo-microscopy. The work described in this thesis was performed on a urea/amide channel (UAC) family member (Saier et al., 2006) from *Bacillus cereus* (*Bc*) American Type Culture Collection (ATCC) 14579 which will be referred to as UAC$_{Bc}$. The protein was identified as a target of interest due its homology to a urea channel and UAC family member UreI from *Helicobacter pylori*. *Helicobacter pylori* is the causative agent of gastric and duodenal ulcers and is a major risk factor for gastric cancer (Covacci et al., 1999) and as described later in this chapter, its urea channel is necessary for colonisation of the human stomach making it a potential drug target.

1.3 Membrane Transport

1.3.1 Types of Transporters

Cells have the ability to control their contents by controlling of the transport of substrates across the membranes using integral membrane proteins called transporters. Membrane transporters are grouped into different classes depending on how the substrate molecule passes through the membrane. Passive transport is mediated by transporters which provide a continuous pathway for transport down the concentration gradient. The term ‘channel’ is commonly used to refer to such proteins. They may be gated and can take an ‘open’ or ‘closed’ conformation in response to ligand binding or a change in voltage (Perozo et al., 1999). Active transport uses the energy from adenosine triphosphate (ATP) hydrolysis, ion gradients or light to transport substrates. Primary active transporters uses the energy from ATP hydrolysis or light to drive transport. For ATP-binding cassette (ABC) transporters, ATP hydrolysis causes a conformational change in the protein that transfers a substrate across the membrane (Dawson and Locher, 2007) whilst phosphorylation and dephosphorylation of P-type ATPases results in changes in the orientation and affinity of the substrate binding sites (Abe et al., 2010). Secondary active transport uses the free energy of an electrochemical gradient of one substrate to drive the transport of another substrate against its concentration gradient. Where both substrates are transported across the membrane in the same directions, they are referred to as symporters and when the substrates flow in opposite directions they are referred to as antiporters (Conde et al., 2009).

1.3.2 Urea Transporter Families

Urea is utilised in different ways in different organisms. In mammals, it is a waste product of nitrogen metabolism mainly produced in the liver (Stewart, 2011). Marine elasmobranchs such as sharks, skates and rays have large quantities of urea in their blood, tissue and body fluids to maintain their osmolarity in the marine environment (Hediger et al., 1996). Urea is also utilised as a nitrogen source for bacteria and involved in a mechanism for acid acclimation in *H. pylori* allowing colonisation of the human stomach.

Urea is small and uncharged and permeates across lipid membranes (Finkelstein, 1976, Orbach and Finkelstein, 1980). Even so, a number of urea transporter families have also been identified. These include the channel type urea transporter (UT) family found in vertebrates and bacteria, members of which allow passive urea transport in either direction across the membrane but with net transport down the concentration gradient. In mammals the genes *slc14a1* and *slc14a2* encode for the UT-B and UT-A urea transporters. UT-A is expressed mainly in the kidney and is involved in the urinary concentration mechanism (Fenton, 2009) whilst UT-B is expressed in more tissues including the brain, heart and intestinal tract. UT-B in the mammalian intestinal tract moves urea from the blood into the intestine for excretion and also supplying bacteria within the intestine with a nitrogen source for growth. Some nitrogen may be returned to the host as newly synthesised amino acids through the nitrogen salvaging mechanism (Stewart and Smith, 2005). X-ray structures have been determined for two members of the UT family: UT from *Desulfovibrio vulgaris* (*Dv*UT) (Levin et al., 2009) and bovine UT-B (Levin et al., 2012). These structures show that the proteins form homotrimers (Figure 1.1). The pore in the UT structures is formed by the two homologous halves in each monomer with a selectivity filter in the middle opening on both sides into wide vestibules. The ~16 Å long selectivity filter in *Dv*UT has a pair of oxygen ladders each built from 3 evenly spaced backbone and side-chain oxygen atoms flanked by phenylalanine sidechains that created slot like filters to accommodate a dehydrated urea molecule. The opposite sides to the oxygen ladders are lined with the hydrophobic residues phenylalanine and leucine. Between the oxygen ladders there is a constricted region formed by valine and threonine on opposite sides with leucine sidechains flanking to create a slot shape similar to the flanking phenylalanines in the oxygen ladders (Levin et al., 2009) (Figure 1.2).

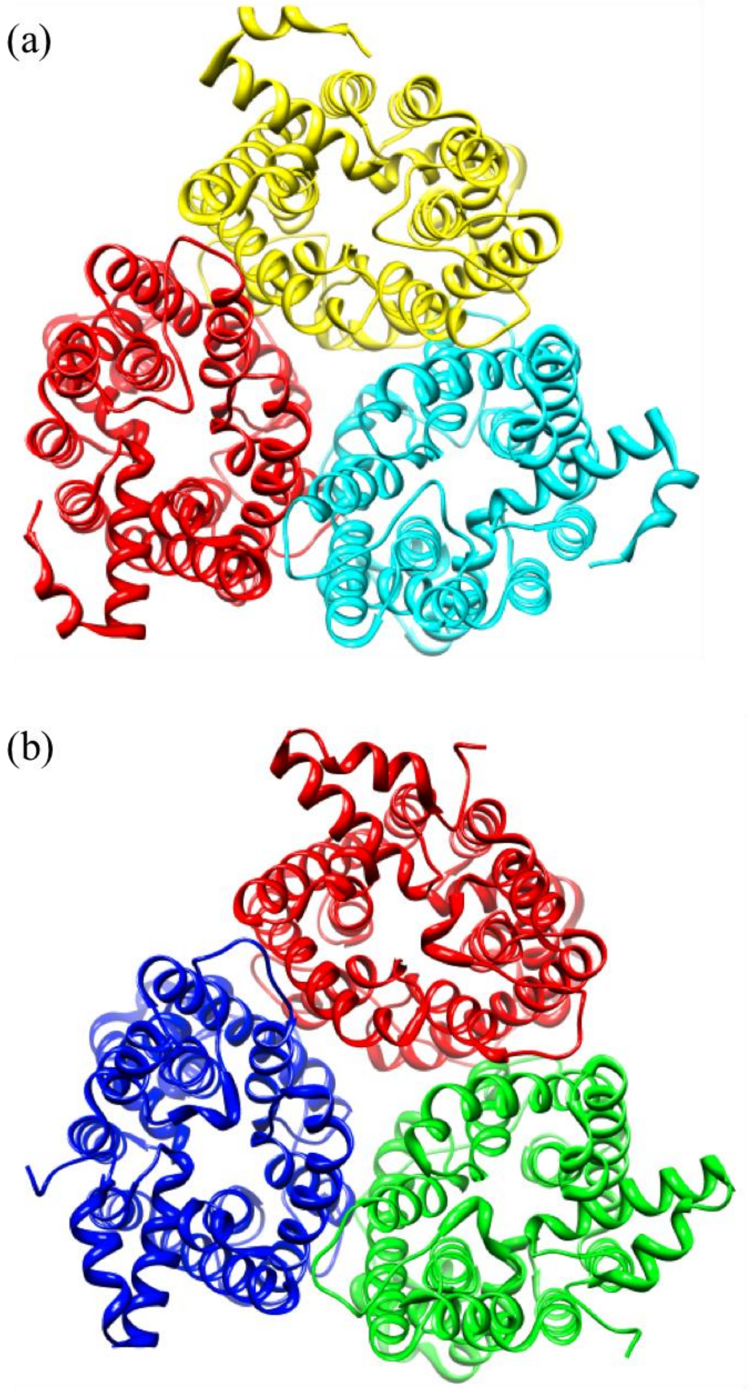


Figure 1.1 X-ray structures of urea transporters.
In the structures from the UT family, the protein forms a homotrimer and with the pore being formed by homologous halves within the monomer. In the above, monomers have been coloured individually.
(a) UT from *Desulfovibrio vulgaris*. Produced from PDB ID 3K3F (Levin et al., 2009).
(b) UT-B from *Bos taurus*. Produced from PDB ID 4EZC (Levin et al., 2012).
Both images were produced using UCSF Chimera (Pettersen et al., 2004).

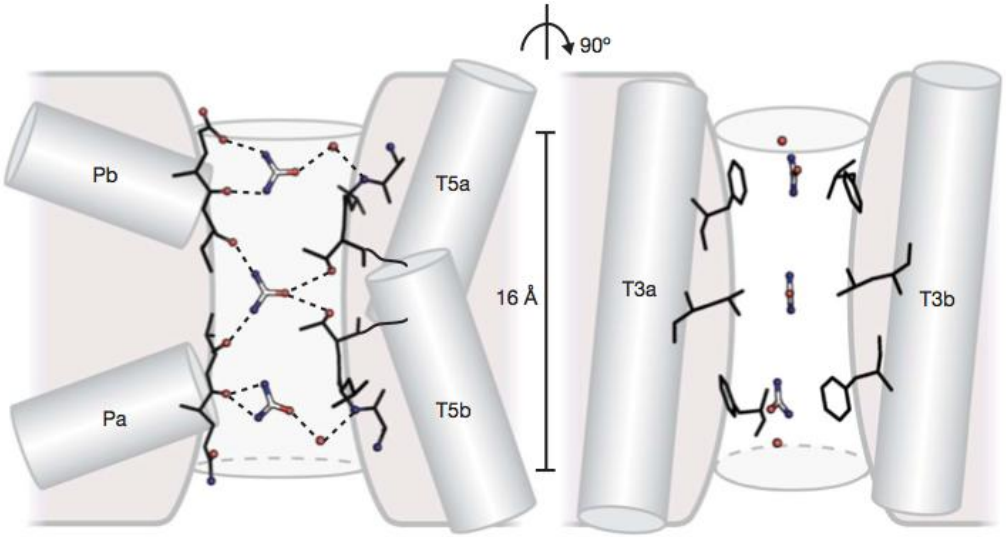


Figure 1.2 The selectivity filter in the Urea Transporter from *Desulfovibrio vulgaris*

Adapted from Levin et al. (2009) with permission from Macmillan Publishers Ltd.

The selectivity filter is formed by two homologous halves within a monomer with pore forming helix a (Pa) and transmembrane helices T3a and T5a from one half and pore forming helix b (Pb) and transmembrane helices T3b and T5b from the other half. The six helices forming the filter are represented as cylinders. The view of the filter on the left shows the predicted locations of urea molecules and their hydrogen-bonding partners. The perpendicular view of the filter on the right shows how the filter is lined by phenylalanine and leucine sidechains.

There is also a low resolution projection of the *Actinobacillus pleuropneumoniae* UT determined by EM which suggests a dimeric organization of the protein (Raunser et al., 2009). Some bacteria such as *Corynebacterium glutamicum* also carry an ABC-type urea transporter. The ABC-type urea transporter in *Corynebacterium glutamicum* is encoded by the *urtABCDE* operon which is expressed in response to nitrogen starvation (Beckers et al., 2004).

Lastly, there is the channel type Urea/Amide Channel (UAC) family which is unrelated to the UT family. The best characterised member of the family UreI is vital for colonisation of the acidic human stomach by *Helicobacter pylori*. In *H. pylori*, UreI is an acid-gated channel supplying urea to a cytoplasmic urease expressed from the same operon as UreI. Decomposition of urea produces ammonia and carbonic acid which is then converted to carbon dioxide by a cytoplasmic β-carbonic anhydrase. UreI can transfer the ammonia and carbon dioxide into the periplasm where they neutralise and buffer the periplasm allowing *H. pylori* to maintain a proton motive force for ATP synthesis and grow in an acidic environment (Figure 1.3) (Scott et al., 2010) (Sachs et al., 2005). Weeks et al. (2004) found that the acid gating of *Hp*UreI and the homologous UreI from *Helicobacter hepaticus* (*Hh*UreI) was conferred by six protonatable residues (His, Glu and Asp) on the periplasmic facing loops of the protein by expressing the proteins and their site directed mutants in *Xenopus* oocytes and by testing the pH dependence of urea transport. Additionally in the pH independent orthologue from *Streptococcus salivarius* (*Ss*UreI), the protonatable residues are absent and fusing the periplasmic domain of *Hh*UreI onto *Ss*UreI does not confer acid gating to the channel. The current model for acid gating in *Hp*UreI and *Hh*UreI suggests that formation of hydrogen bonds by periplasmic residues results in conformational changes in the transmembrane helices forming an open conformation at low pH and a closed conformation at neutral pH (Weeks et al., 2004).

Homologues of UreI have also been identified in *Pseudomonas aeruginosa* (Wilson et al., 1995) and in *Rhodococcus* sp R312 (Chebrou et al., 1996). Whilst their function has not yet been characterised, they have been inferred to be amide transporters on the basis of their gene locations in amidase operons.

Currently no high or low resolution structures of any members of the UAC family have been published but there is some evidence for the structural details. *Hp*UreI has six predicted transmembrane alpha helices (Weeks et al., 2000) and has been suggested to be a homotrimer from cross-linking experiments (Gray et al., 2011).

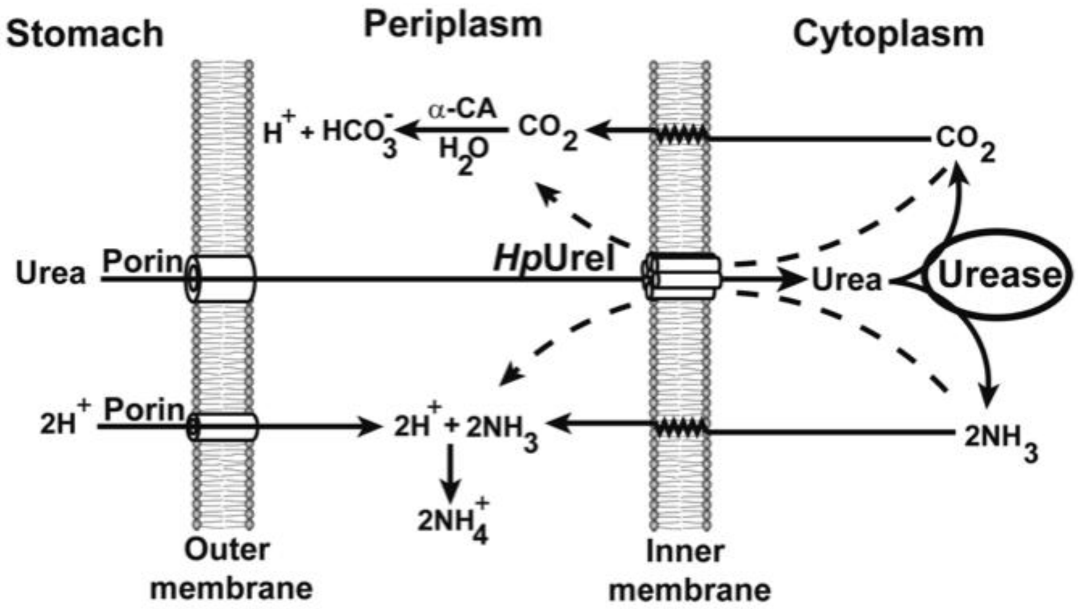


Figure 1.3 Mechanism of acid acclimation in *Helicobacter pylori*.
Reproduced with permission from Gray et al. (2004). Copyright American Chemical Society.
*Hp*UreI conducts urea into the cytoplasm where it is hydrolysed by a cytoplasmic urease to CO_2 and NH_3 which diffuse back into the periplasm. NH_3 neutralises the pH in the periplasm and CO_2 is converted to HCO_3^- by α-carbonic anhydrase (α-CA). HCO_3^- buffers the periplasm at pH 6.1.

1.4 Methods to Study Membrane Protein Structure

X-ray crystallography, nuclear magnetic resonance spectroscopy and electron microscopy are used to analyse protein structures in three dimensions up to atomic resolution. Of these techniques, X-ray crystallography is the main method for determination of the structure of proteins at atomic resolution whilst NMR can report on the dynamical processes of proteins although solution state NMR is limited to studying small protein:detergent complexes up to ~100 kDa (Nietlispach and Gautier, 2011).

Electron microscopy is a viable alternative to X-ray crystallography or nuclear magnetic spectroscopy. Using electron microscopy, the highest resolution is usually obtained by imaging two-dimensional crystals of membrane proteins which contain an ordered array of the protein within a lipid bilayer. This offers the advantage of a native like environment for the membrane proteins unlike structures determined using X-ray crystallography which are usually grown from membrane proteins in detergent solution. With X-ray structures, care must be taken with interpretation of the models as the presence of a lipid membrane may be necessary to maintain the protein in its native confirmation (Lee et al., 2005). This is in contrast with some of the highest resolution structures determined using electron microscopy of 2D crystals which can show the positions of lipid molecules bound to the protein (Gonen et al., 2005, Grigorieff et al., 1996). The production of 2D crystals of membrane proteins will be described further in Chapter 4.

There have been eight membrane proteins determined to atomic resolution using electron microscopy: the light driven proton pump bacteriorhodopsin (Grigorieff et al., 1996), light harvesting complex II (LHCII) (Kühlbrandt et al., 1994), the water channels aquaporin 0 (Gonen et al., 2004), aquaporin 1 (Murata et al., 2000) and aquaporin 4 (Hiroaki et al., 2006), torpedo ray nicotinic acetylcholine receptor (Unwin, 2005), microsomal glutathione transferase (Holm et al., 2006) and prostaglandin E2 synthase (Jegerschold et al., 2008). In these cases, the resolution was sufficient to reveal details such as the side chains involved in pore formation and the basis of substrate specificity in the aquaporin channels or the arrangement of the cofactors in the LHCII. There are also a number of membrane proteins with 3D models at intermediate resolutions. At a resolution of 7 – 10 Å, an α-helix appears as a cylindrical tube of density. Even so, this can provide useful information to guide the modelling of most membrane proteins which span the membrane with alpha helices consisting mostly of hydrophobic residues. There are also a number of projection maps of membrane proteins which depict the density superimposed onto a single plane. At low resolutions (>25 Å) they can provide detail on the oligomeric state of the protein. At higher resolutions and with careful interpretation, they can still reveal the secondary structure of the protein and guide modelling attempts (Schmidt-Krey et al., 2004, Schmidt-Krey et al., 1999).

Transport proteins usually have multiple conformations giving rise to the gating mechanisms of channels or the pumping mechanisms of active transporters. By rapidly freezing 2D crystals, there have been successful attempts at determining the structures of transient intermediates for a number of proteins (Unwin and Fujiyoshi, 2012, Subramaniam et al., 1993). In other cases such as with the KirBac 3.1, the protein can be crystallised in 2D in two different conformations (Kuo et al., 2005). The data from electron microscopy can complement those from other structure determination methods and provide a more complete understanding of the protein structure and function.

Chapter 2 - Electron Microscopy and Data Processing

2.1 Introduction

2.1.1 The Electron Microscope

A microscope is a used to produce magnified images of an object and the usefulness of a microscope depends on the resolving power, which is the ability to distinguish between two points of the object. A light microscope is limited by the wavelength of light as described by the Rayleigh criterion which describes the smallest distance that can be resolved δ as:

$$\delta = 0.61\lambda / \mu\sin\beta$$

In the above, λ is the wavelength of the illuminating radiation, μ is the refractive index of the viewing medium and β is the semi-angle of collection of the magnifying lens (Williams and Carter, 1996). Based on the Rayleigh equation, the electron microscope has potential for atomic resolution imaging. The electron wavelength depends on its voltage according to the de Broglie equation so an electron with an accelerating voltage of 100 keV has a wavelength of 0.04 Å which is significantly smaller than the diameter of an atom (Williams and Carter, 1996). However, the resolution achieved by imaging thin biological specimens using transmission electron microscopy (TEM) falls short of the potential for several reasons.

A more formal explanation of image formation by the electron microscope can be found in several texts including Williams and Carter (1996) and Glaeser (2007) but "to a good approximation the intensity observed in an EM bright field image is a projection of the 3D coulomb potential distribution corrupted by the wave aberrations of the objective lens" (Zhu et al., 1997). For the purposes of this thesis the author will briefly describe some concepts relevant to imaging. During passage of the electron beam through the specimen, many electrons will pass through the specimen without interacting with it but an electron may be scattered elastically giving rise to information describing the object in the produced image or inelastically imparting energy to the sample resulting in radiation damage. The proportion of elastic to inelastic scattering events N_e/N_i depends on the chemical composition of the sample and is approximately Z/19 where Z is the atomic number (Crewe et al., 1970). For proteins made mostly of carbon, nitrogen, oxygen and hydrogen, inelastic scattering predominates creating a noisy background (Amos et al., 1982, Zhu et al., 1997). An elastically scattered electron suffers a directional change and a $\pi/2$ phase shift but with its wavelength unchanged. The phase of the elastically scattered electron wave is further altered by the spherical aberration of the microscope and by any defocus applied using the objective lens. As the elastically scattered wave is out of phase with the direct unscattered wave, there is phase contrast as the waves interfere constructively or destructively, strengthening or weakening wave amplitudes. Phase contrast is the dominant source of contrast at medium and high resolutions. Amplitude contrast can occur due to a virtual loss of electrons involved in elastic image formation from scattering outside the objective aperture or those removed by inelastic scattering. Amplitude contrast provides additional contrast at low resolution (Frank, 2006).

Abbe (1873) determined that during the process of image formation by the objective lens, the specimen diffracts the illuminating radiation and the lens refocuses the radiation into an image in a two step process forming a diffraction image in the back focal plane followed by formation of the image in the image plane (Figure 2.1). In the EM, the image formation process is furthered complicated by the fact that the wavefunction in the back focal plane is modulated by the contrast transfer function (CTF). The shape of the CTF depends on the acceleration voltage, the spherical abberation value of the objective lens and the defocus setting. During imaging only the defocus is varied. The CTF results in image abberations including contrast reversals due to the oscillating sign of the CTF which become increasing rapid at higher resolutions and zones in the resolution where information is lost as the CTF crosses zero. The CTF is also damped at higher resolutions by an envelope function due to partial spatial and temporal coherence of the electron beam resulting in a resolution limit for the microscope where the contrast is damped to zero (Figure 2.2). As described by Thon (1966), the CTF produces a pattern of concentric rings in the

Fourier transform of an EM image in which the dark rings correspond to the spatial frequencies where the CTF crosses zero and the bright rings alternate between positive and negative contrast (Glaeser, 2007). An example of Thon rings can be seen in Figure 5.2b. The dark rings in the Fourier transform can be used to determine the CTF zero positions and therefore the value of defocus and correct for the contrast reversals. EM images are typically taken with varying defocus values so the lost of information at the CTF zeros may be compensated by merging data from several images as described later in section 2.2.5.

The main factor limiting high resolution imaging of proteins is their vulnerability to the electron beam and doses of 1 to 10 electrons/Å^2 are sufficient to destroy the structure of a protein (Henderson, 1992). Ionising radiation such as that from the electron beam results in chemical changes in the specimen including cross-linking and bond scission (Stenn and Bahr, 1970). In order to record images of biological specimens, low doses must be used but this produces low contrast images with poor signal to noise ratios. High resolution imaging of proteins is therefore achieved by averaging together large numbers of images and also applying corrections for image aberrations such as those occurring due to the effects of the CTF (Wade, 1992). This may be achieved by imaging protein crystals which has the benefit of alignment of the molecules in the crystal and provided that sufficiently large and well-ordered crystals are imaged, the signal to noise barrier can be overcome (Unwin and Henderson, 1975) resulting in atomic resolution imaging (Henderson et al., 1990). Alternatively, images of individual protein particles can be aligned and averaged in what is known as single particle analysis. However, this is more commonly applied to molecules above 500 kDa due to the difficulties encountered when attempting to align small particles (Glaeser and Hall, 2011).

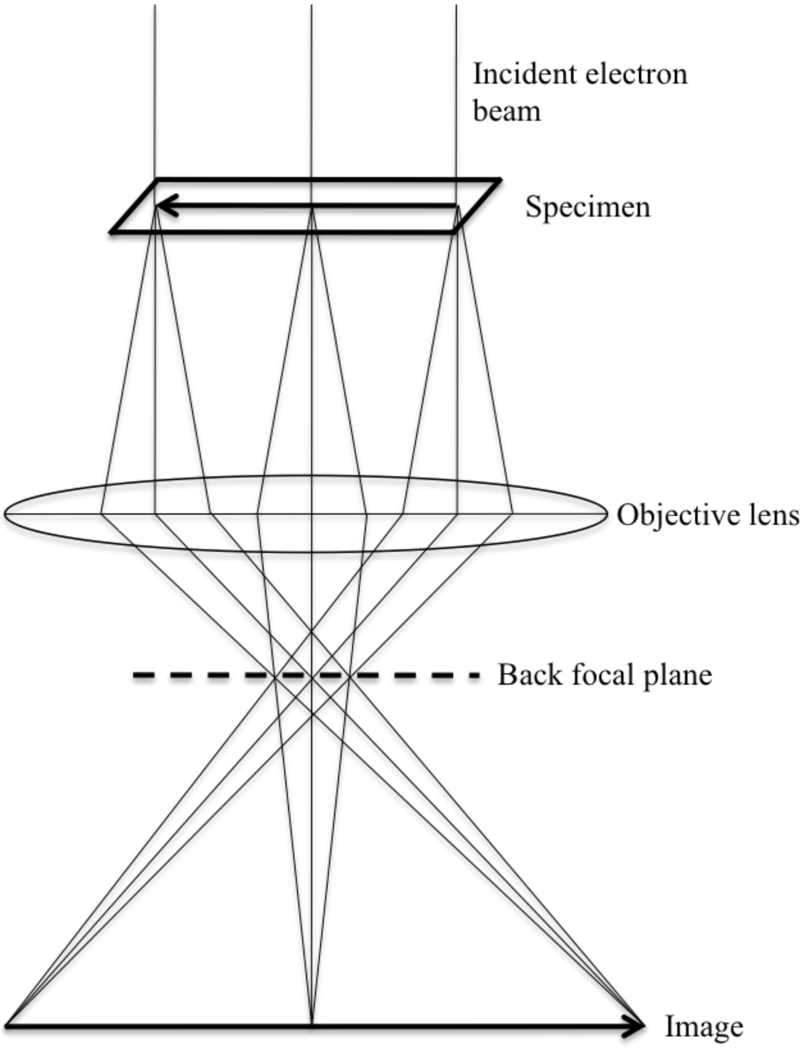


Figure 2.1 Ray diagram showing the process of image formation in the electron microscope
Redrawn from Henderson (1992).
During the process of image formation, the objective lens focuses the scattered electrons forming a diffraction image in the back focal plane. The wave in the back focal plane is modulated by the contrast transfer function of the microscope giving rise to the image observed in the EM.

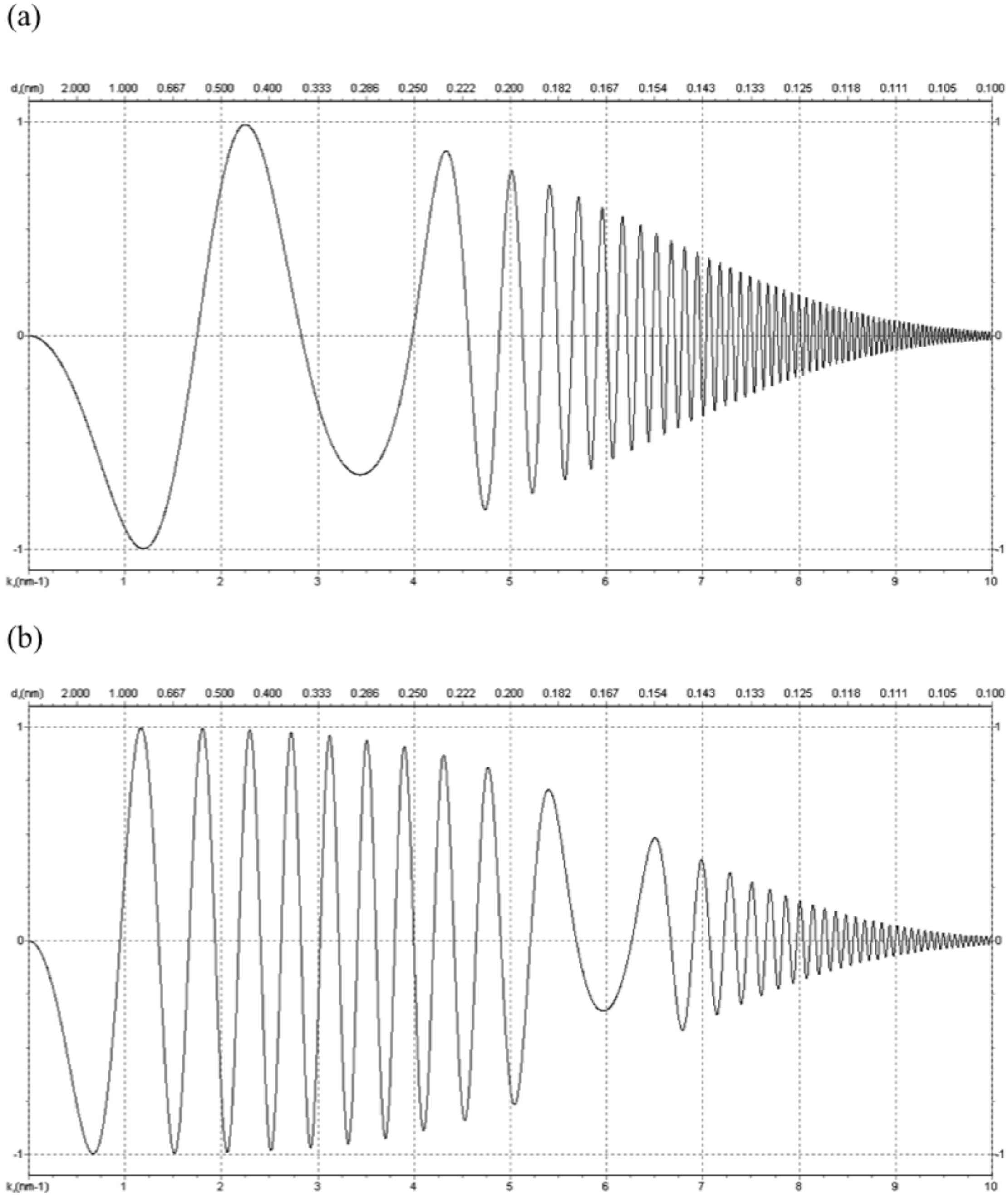


Figure 2.2 Simulation of the contrast transfer function

The CTFs of an EM operating at 200 keV with a spherical abberation of 2.0 mm at defocus values of (a) 1500 Å and (b) 4500 Å have been simulated above using ctfExplorer (Sidorov, 2002).
Modulation of the wavefunction in the back focal plane of the microscope by the CTF results in abberations in EM images.

2.2 Electron Microscopy Methodology

2.2.1 Negative Staining

Electrons are scattered by the air so the EM must operate under a vacuum. This creates another problem as dehydration of protein molecules can result in collapse and distortion of the structure (Bremer et al., 1992). One solution to the problem is the technique of negative staining where the specimen is embedded in a heavy metal salt (Hall, 1955). The advantages of negative stain are the improved contrast and preservation of the surface features of the specimen. For crystallographic work, the main value of negative stain is to provide an indication of the diffracting potential of the crystals as well as information on the size and morphology of the crystals for later data collection. However, negative stain also imposes a resolution limit of ~20 Å and mostly provides details on the surface features of the specimen depending on the ability of the stain to penetrate the sample (Bremer et al., 1992). To achieve higher resolution structure analysis, it becomes necessary to record images of the unstained sample.

2.2.2 Unstained Imaging

There are two widely use methods for the preservation of the hydrated structure of a protein for high resolution imaging. The first is to replace the waters of hydration with a less volatile medium such as glucose or trehalose as first demonstrated when Henderson and Unwin (1975) determined the 3D structure of bacteriorhodopsin with crystals embedded in glucose. The second method is to freeze the sample in a thin film of vitreous ice which closely mimics a hydrated environment. This requires the sample grid to be kept at liquid nitrogen temperature preventing the conversion of vitreous ice to crystalline ice and thus requires a grid holder cooled to liquid nitrogen temperatures during microscopy (Taylor and Glaeser, 1976). However, cooling to liquid nitrogen temperatures or beyond to liquid helium temperatures also reduces some of the effects of electron beam damage. The cooling is thought to trap the free-radicals formed by radiolysis limiting their ability to cause further damage (Stenn and Bahr, 1970) and a specimen cooled to -120 ºC has a four fold increased lifetime than at room temperature (Hayward and Glaeser, 1979).

2D crystals must still be imaged in 'low dose' conditions. A protocol for this minimises exposure of the specimen to the beam by performing focusing and astigmatism corrections on a region adjacent to the crystal and by using a dose for recording that produces featureless images dominated by noise (Williams and Fisher, 1970) which require processing in order to derive useful structural information.

2.2.3 Image Processing

The Medical Research Council (MRC) programs are a suite of image processing programs developed for the processing of electron micrographs of 2D crystals and helical or icosahedral particles (Crowther et al. 1996). The MRC programs relevant to 2D crystal processing were later incorporated into the more user-friendly *2dx* graphical user interface (Gipson et al., 2007) and were used by the author for the processing of the 2D crystal micrographs. The texts by Amos et al. (1982) and Henderson et al. (1986) are excellent descriptions of image processing of electron micrographs. A flowchart summarising the steps of image processing is given in Figure 2.3.

The basis of crystallographic processing is that the electron density in a protein crystal can be represented as a sum of sinusoids. The amplitude and phase data needed to reconstruct the electron density can be extracted from the Fourier transform of the crystal image and filtered from the background noise. Additionally, the 3D density of an object can be reconstructed from multiple projection views of the object at different angles (DeRosier and Klug, 1968). In contrast to the Fourier transform of a 3D crystal used in X-ray crystallography which is an array of discrete diffraction spots in reciprocal space (Glaeser, 2007), the Fourier transform of a 2D crystal takes the form of a series of parallel lattice lines perpendicular to the plane of the crystal along which the phase and amplitude vary continuously (Figure 2.4). An electron micrograph of a 2D crystal is a projection and its Fourier transform can be thought of as a central plane intersecting the lattice lines (Amos et al., 1982). Calculating the Fourier transform of a crystal image gives a pattern of diffraction spots. 2D crystals can have imperfections resulting in diffuse diffraction spots and the process of grid preparation can introduce more defects into the crystal lattice. To correct for these in the process known as unbending, the Fourier Transforms are masked around the diffraction spots and the filtered Fourier transform is then transformed back into a filtered image. A reference area in the image is compared against the rest of the image to generate a cross correlation map that informs the user of how well the rest of the image correlates to a reference area of the image and how far a unit cell position is shifted from its expected position. Distortions in the crystal lattice are then corrected using a series of vectors to shift unit cell locations. The spots in the Fourier transform of the ‘unbent’ image appear sharper as a result and have an increased signal.

The spots are graded with an Image Quality (IQ) number of 1 – 8 according to the height of the peak relative to the background level; grade 7 is where a spot is at the background level and a grade 1 spot has a 7:1 signal to noise ratio or higher (Henderson et al., 1986). The data can then be corrected for the effects of the CTF, merged with data from other films and analysed for symmetry relationships.

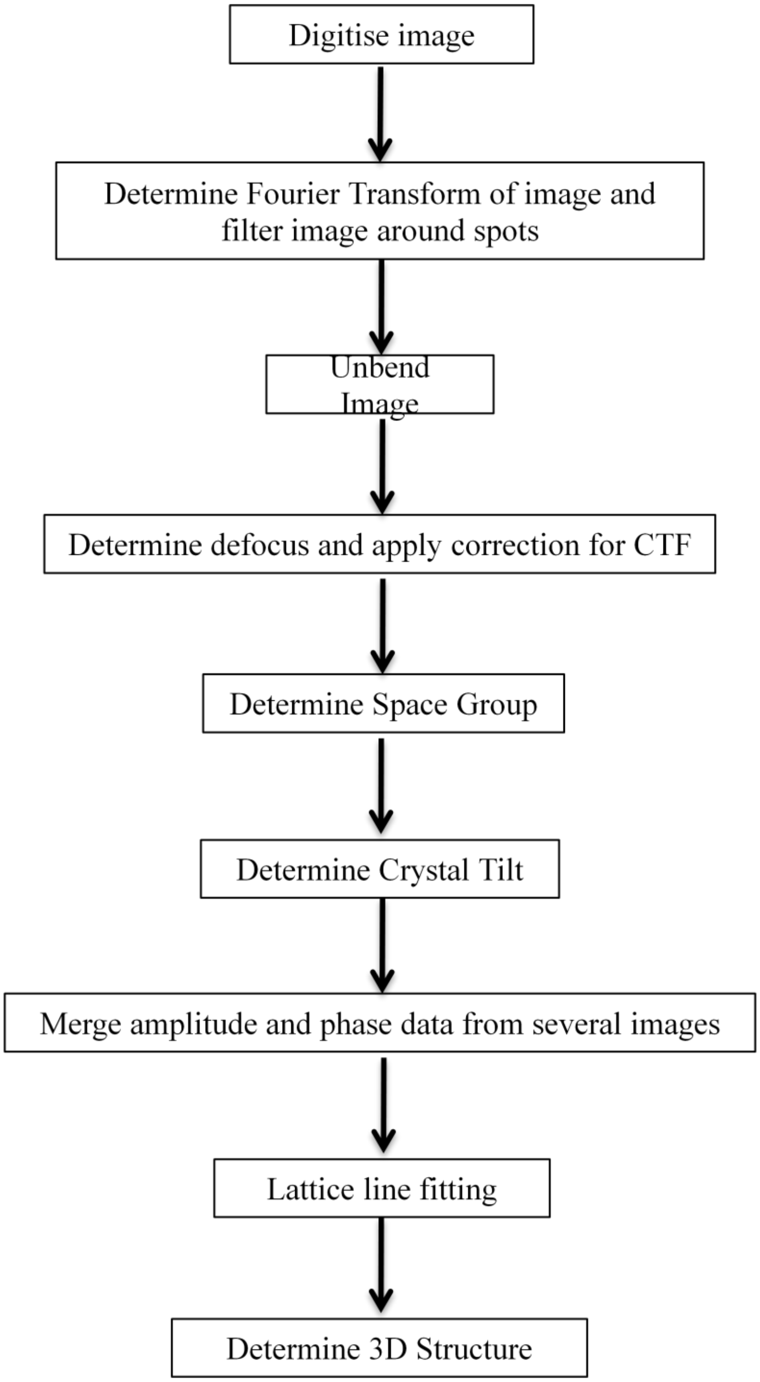


Figure 2.3 Flowchart showing major steps of image processing.
The image processing and merging of data is performed using the MRC suite of programs.

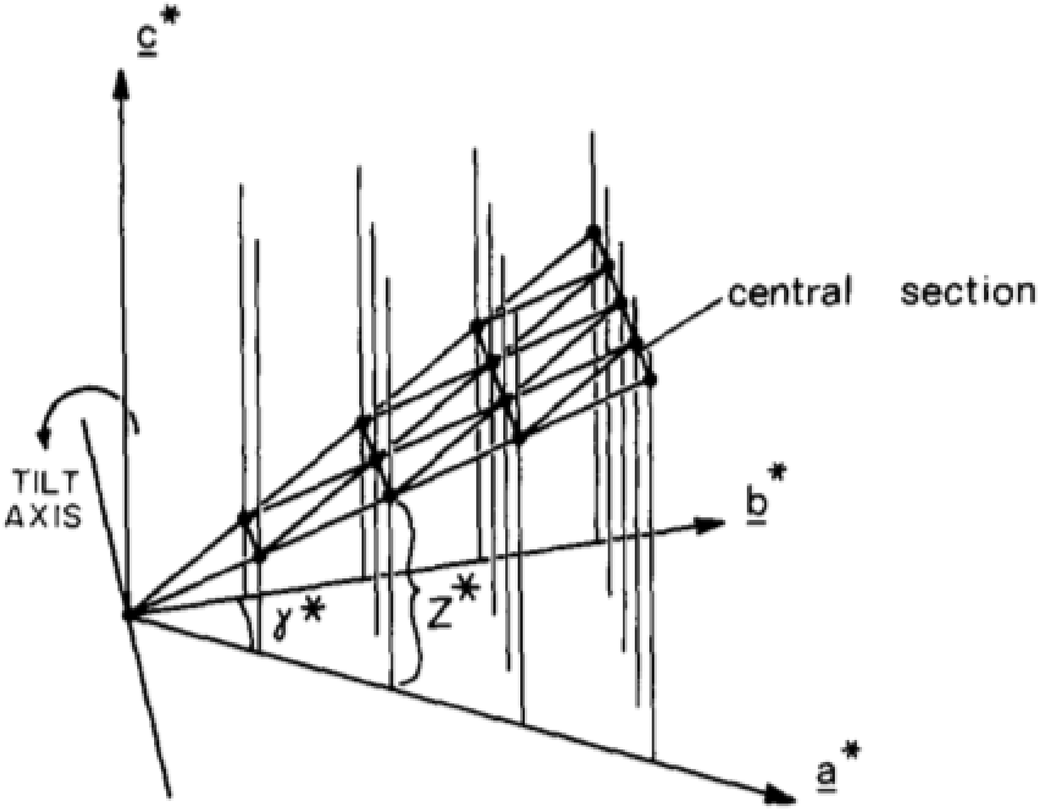


Figure 2.4 Lattice lines and the central section theorem
Reproduced from Amos et al. (1982) with permission from Elsevier.
The Fourier transform of a 2D crystal takes the form of lattices lines extending perpendicular to the plane of the crystal. The Fourier transform of a crystal image gives the values of amplitudes and phases at points along the lattice lines (z*) where the central section (perpendicular to the viewing direction) intersects them.

2.2.4 Symmetry Analysis

If there is symmetry in the projection of an untilted 2D crystal then symmetry related spots in the Fourier transform ought to take the same amplitudes and phases. This means that there is redundancy in the recording of data for a structural feature (Landsberg and Hankamer, 2007). The program ALLSPACE determines phase residuals for spots in an image which are the mean deviations from a symmetry enforced average or theoretical values for the plane group (Valpuesta et al., 1994). After examination of the output for plane groups consistent with the phase residuals, symmetry can be enforced on the data to produce an untilted projection map with improved signal according to how many symmetry relationships have been enforced. The symmetry of the crystals can also suggest the oligomeric state of the protein.

2.2.5 Merging of Data from Different Images and Map Calculation

The data from different images are aligned to a common phase origin and used to produce an averaged list of amplitudes and phases. The quality of the merging is assessed by the phase residual which in this case is the deviation of the phases of an image against the other films in the merging. The amplitudes in electron micrographs are generally poorly recorded especially at higher resolutions due to a combination of effects including radiation damage, specimen movement and inelastic scattering (Henderson, 1992). The amplitudes can be rescaled according to the expected fall-off using diffraction data from bacteriorhodopsin for example, as a reference increasing the contribution of higher resolution information to the merged map (Havelka et al., 1995).

A merged untilted projection could be determined by following the above processing work flow. To determine a 3D reconstruction, data from tilted images must be aligned to a common phase origin with the correct tilt angle and axis. With a series of tilt images, the amplitudes and phases have been sampled at various points along the reciprocal space lattice lines. The program LATLINE interpolates the data to reconstruct the lattice lines (Agard, 1983) and these can be used to generate a 3D density map. In the resulting projection or 3D map, accurate phases will mean that the positions of density features will be reliable whilst the amplitude information informs the density level at the positions defined by the phases. The images recorded in the EM retain the phases so the ‘phase problem’ of X-ray crystallography does not occur but the amplitudes are not well defined (Ford and Holzenburg, 2008). Better amplitudes can be obtained using EM by collecting diffraction data and combining the data with phases obtained from images. Apart from the nicotinic acetylcholine receptor, 7 of the atomic resolution membrane protein models were determined this way (listed in section 1.4).

To summarise, images of biological specimens suffer from aberrations due to the limitations of electron optics and have low signals due to their vulnerability to electron beam damage. The signal-to-noise barrier can be overcome by averaging aligned images. Large and highly ordered crystals of UAC_{Bc} are therefore required in order to achieve high resolution.

Chapter 3 - Materials and Methods

3.1 Cloning, Culture and Membrane Preparation

Cloning was performed by Ivan Campeotto and Gareth Wright at the University of Leeds. Expression of the protein and preparations of membranes were performed by Dr. Peter Roach and David Sharples at the University of Leeds.

The open reading frame for the UreI homologue from *Bacillus cereus* ATCC 14579 was cloned into a pTTQ18 based expression vector with a C-terminal hexa-histidine tag and transformed into *E. coli* BL21Star cells (Invitrogen). The sequence for the construct is shown in Appendix 1. The protein was auto-expressed in 30 l cultures in media containing 25 mM Na_2HPO_4, 25 mM KH_2PO_4, 50 mM NH_4Cl, 5 mM Na_2SO_4, 1 mM $MgSO_4$, 0.5% (w/v) glycerol, 0.05% (w/v) glucose, 0.2% (w/v) lactose, 0.1 mg/ml carbenicillin, 0.03 mg/ml chloramphenicol and 2% yeast extract at 37°C (Deacon et al., 2008, Postis et al., 2008).

The cells were lysed in 20 mM Tris-HCl pH 7.5 and 0.5 mM EDTA with a cell disrupter (Constant System Ltd) at 35 kpsi and unbroken cells were removed by centrifugation at 14,000 g. The supernatant containing the broken membranes and cytosolic proteins was subjected to Tangential Flow Filtration using a Pellicon TFF system (Millipore Corporation). The mixed membranes remaining in the filtration system were concentrated and frozen as droplets in liquid nitrogen for storage (Roach et al., 2008).

Buffer Name	Composition
Solubilisation Buffer	50 mM HEPES pH 7.5 100 mM KCl 2.5% (v/v) Glycerol 1.5% (w/v) *n*-dodecyl-β-D-maltopyranoside (DDM) 0.05% (w/v) NaN_3

Wash Buffer	50 mM HEPES pH 7.5 100 mM KCl 2.5% (v/v) Glycerol 0.05% (w/v) DDM 0.05% (w/v) NaN_3 20 mM Imidazole
Elution Buffer	50 mM HEPES pH 7.5 100 mM KCl 2.5% (v/v) Glycerol 0.05% (w/v) DDM 0.05% (w/v) NaN_3, 200 mM Imidazole
Gel Filtration Buffer	50 mM 2-(*N*-morpholino)ethanesulfonic acid (MES) pH 6.5 100 mM KCl 2.5% (v/v) Glycerol 0.05% (w/v) DDM 0.05% (w/v) NaN_3
Crystallisation Buffer	20 mM HEPES pH7.5 100 mM NaCl 10% Glycerol 0.05% NaN_3

Table 3.1 Buffers for Purification

3.2 Protein Purification

Mixed membranes were mixed in solubilisation buffer for 1 hour at 4 °C. Insolubilised material was removed by centrifugation at 100,000 g for one hour in a Ti 45 rotor and the supernatant was incubated with 10 µl HisPur Cobalt Resin (Pierce) per 1 ml supernatant. The resin was transferred to a 20 ml column (Biorad) and the cobalt resin was washed with 150 resin volumes of wash buffer before eluting in elution buffer. Fractions containing protein as determined by absorbance at 280 nm were concentrated with a 100 kDa cut-off Vivaspin concentrator (Sartorius) until the volume was reduced to ~0.5 ml. The protein was then applied to a Superose 6 10/300 column (GE Healthcare) at a flow rate of 0.4 ml/min under the control of an ÄKTA FPLC (GE Healthcare). Fractions containing purified protein were pooled and concentrated to over 0.4 mg/ml protein using a 100 kDa cut-off Vivaspin concentrator in preparation for 2D crystallisation.

3.3 BCA Assay

The protein concentration assay was performed using the bicinchoninic acid (BCA) assay kit from Thermo Scientific using bovine serum albumin (BSA) as the standard and a 10-fold dilution of purified UAC_{Bc} in water to conserve protein.

3.4 Molecular Weight Estimation by SEC

Apoferritin (443 kDa), β-amylase (200 kDa), alcohol dehydrogenase (150 kDa), bovine serum albumin (66 kDa) and carbonic anhydrase (29 kDa) from the Kit for Molecular Weights 29,000-700,000 (Sigma) were applied to the Superose 6 10/300 column. Blue dextran was used to determine the void volume of the column. The elution volumes were recorded and used to determine partition coefficient values (K_{av}) values using the following formula:

$$K_{av} = V_e – V_o / V_c - V_o$$

V_e is the elution volume of the protein standard, V_c is the geometric column volume and V_o is the void volume of the column determined using blue dextran. The K_{av} values were then plotted against the log molecular weights of the protein standards and the interpolated line used to estimate the molecular weight of the purified UAC_{Bc} protein:detergent complex (Laurent and Killander, 1964).

3.5 2D Crystallisation

3.5.1 Preparation of Lipid Stocks

The following protocol was used to prepare stocks of 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC), 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC), 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC), 1,2-distearoyl-*sn*-glycero-3-phosphocholine (DSPC) and *E. coli* total lipid extract (ETL) in 1% (w/v) decyl maltoside (DM). A 2 ml glass vial was cleaned with chloroform and 2 mg of lipid in chloroform from Avanti Polar Lipids was transferred to the vial. The lipid was dried to a film under a gentle stream of nitrogen and dissolved in 100 μl of 10% (w/v) DM (Glycon Biochemicals). Cycles of freezing and thawing were used if the lipid/detergent mixture remained turbid after gentle mixing. Once clarified, the volume of the solution was adjusted to 1 ml with deionised water. Prepared lipid stock solutions were stored at –20°C and used within 2 months.

3.5.2 Preparation of Dialysis Membranes

To remove contaminants from the manufacturing process and glycerol used for preservation, 12 - 14 kDa Visking dialysis tubing (Medicell International Ltd) was cut in to 10 cm strips and boiled in 2% (w/v) Sodium bicarbonate and 1 mM EDTA at pH 8 for 20 minutes. The wet membranes were then rinsed in deionised water and stored in 20% (v/v) Ethanol at 4 °C.

3.5.3 Two-dimensional Crystallisation

The tested crystallisation conditions are summarised in Appendix 2. The purified protein was mixed with lipid stock solution at various LPRs and the protein concentration adjusted to 0.4 mg/ml by dilution with crystallisation buffer. The mixture was dialysed against the corresponding crystallisation buffer for 14 days on a home built continuous flow dialysis machine (Jap et al., 1992, Glover et al., 2011). Removal of detergent was assessed by comparing the diameter of a drop of the crystallisation mixture with an equal volume drop of crystallisation buffer.

3.6 Electron Microscopy

3.6.1 Preparation of Negative stain

0.0375 g of uranyl formate (Polysciences Inc.) was dissolved in 5 ml boiling deionised water and stirred for 5 minutes. 5 M NaOH was added until the colour changed from a light to slightly darker yellow (typically ~8 µl for 5 ml stain) and stirred for another

5 minutes. The 0.75% (w/v) uranyl formate stain was filtered through a SupaTop 0.2 µm cellulose acetate syringe filter (Anachem) and stored in the dark at 4°C.

3.6.2 Preparing Carbon Coated Grids for Electron Microscopy

A carbon rod (Agar) was sharpened and placed in contact with a blunt ended carbon rod in a Cressington Carbon Coater and evaporated in air at 10^{-4} to 10^{-5} mbar to form a film with a thickness of 100-200 Å on a freshly cleaved mica slide.

3.05mm 400 mesh palladium-copper (Pd/Cu) grids (Agar Scientific) were placed on a filter paper (Whatman) palladium side up and submerged in deionised water. The carbon film was floated on the surface of the deionised water and the water level lowered by siphoning until the carbon film came into contact with the grids. The grids were then left to dry overnight and used for electron microscopy up to 2 weeks after preparation.

3.6.3 Crystal Preparation for Cryo Electron Microscopy

The exponential microprobe of a MSE Soniprep 150 was placed into a 20 µl suspension containing 2D crystals and the suspension was sonicated with 3.5 amplitude microns for 5 seconds (Figure 3.1). Carbon coated Pd/Cu grids were glow discharged under vacuum within the Cressington Carbon Coater with the Cressington Power unit for 30 seconds and the sonicated sample was applied to the grids and left for 1 minute. The grids were then blotted and suspended in 1% (w/v) glucose before blotting and drying in air.

3.6.3 Electron Cryomicroscopy

Grids were mounted on a Gatan cryo holder cooled to 90 K within the Phillips CM200 Field Emission Gun microscope with liquid nitrogen. The grid was adjusted to the eucentric height and viewed at a nominal magnification of 600x. When a crystal was found, focusing was performed on two adjacent areas to the sample at a nominal 100,000x magnification to determine an approximate position of focus for the sample. Images were recorded at a nominal magnification of 50,000x and defocus between

15,000 and 25,000 Å using an exposure time of 1s on Kodak SO-163 film. The microscope was operated at 200 keV.

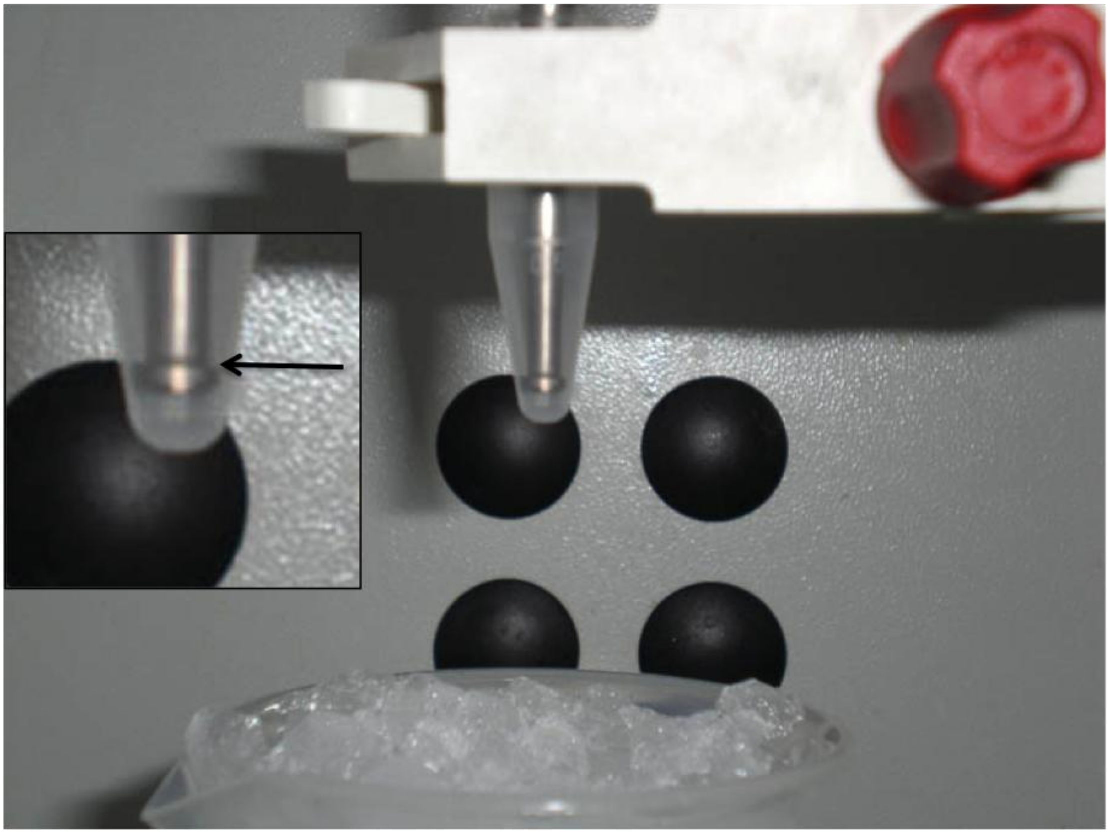


Figure 3.1 Set-up for crystal disaggregation technique

To perform the disaggregation technique, a sonicating probe is carefully positioned with the exponential probe tip breaking the meniscus of the liquid without touching the sides of the 1.5 ml Eppendorf tube to prevent damage of the tube and contamination of the sample with tube fragments. Typically 20 µl is the volume of the crystal suspension and sonication is performed with the tube on ice. The inset square on the left shows a magnified view of the probe and an arrow indicates the meniscus of the liquid.

3.7 Image Processing

3.7.1 Film Development and Digitisation

Films were developed in Kodak D19 for 12 minutes at 20°C and fixed with IlFord Hypam Fixer for 5 minutes. Films showing clear diffraction spots on a Sira Optical Diffractometer were scanned using a Zeiss SCAI densitometer at a step size of 7 µm. Fast Fourier Transforms computed with Digital Micrograph 3 (Gatan) were used to determine regions of the image containing diffraction spots and images between 6500 x 6500 pixels and 3500 x 3500 pixels were cropped with Adobe Photoshop.

3.7.2 Image Processing in *2dx*

Individual images were processed using the MRC suite of programs within the *2dx* interface (Gipson et al., 2007). The calculated Fourier Transforms were indexed within 2dx and the degree of defocus was determined. 2 unbending rounds to correct lattice distortions were applied (Henderson et al., 1986). Parameters used for unbending are shown in Appendix 3. For some images, the automatic masking of crystal function using cross correlation output within the *2dx* package was applied and increased the number of IQ 1, 2 and 3 spots relative to the unmasked images. ALLSPACE (Valpuesta et al., 1994) was used to determine possible plane groups and the CTF corrected amplitude and phase files were used for merging by phase origin refinement using ORIGTILTK residual in steps of 6° at first and then in 0.1° steps within a stand-alone script. SCALEIMAMP was used to determine B-factors and rescale the amplitudes using data from bacteriorhodopsin as a reference (Henderson et al., 1986). A projection map was calculated from the merged amplitudes and phases using the CCP4 suite of programs (Collaborative Computational Project, Number 4, 1994).

3.7.3 Tilted Crystal Image Collection and Processing

Tilt series of negatively stained crystals were recorded on a 1000 x 1000 charge coupled device on a Philips CM100 equipped with a LaB_6 filament at a nominal magnification of 25,000x. After recording an untilted image of a crystal and inspecting the Fourier Transform for at least two orders of diffraction and sharpness of the spots, the holder goniometer was tilted up to 50° in both directions of the tilt axis recording images at 10° intervals with a defocus of ~6000 Å to ensure that all parts of the image were underfocused.

The images were processed as described earlier using the MRC suite of programs within the *2dx* interface including an additional step of using EMTILT (Shaw and Hills, 1981) to determine tilt axes and tilt angles from the lattice distortions of the most highly tilted images.

Initially, the untilted images were merged to a common phase origin with imposition of *p*6 symmetry and refined to a minimal phase with ORIGTILTK. The low angle tilted images were added to the merging using the tilt axis determined by EMTILT and a tilt angle relative to that determined for the highest tilt image. A z* window of 0.007 was used for phase comparison. LATLINE was used to interpolate amplitudes and phases along individual lattice lines (Agard, 1983) using a unit cell value in the c axis of 200 Å and with a real space envelope of 160 Å with tapering edges of 20 Å applied. The 3D map was then calculated using the CCP4 suite.

3.7.4 Single Particle Image Collection and Processing

Purified protein in detergent was diluted to a concentration of ~15 μg/ml using gel filtration buffer (Table 3.1) and applied to a glow discharged carbon coated grid. The grid was blotted and stained with 0.75 % (w/v) uranyl formate for 25 seconds before blotting and drying using a weak vacuum source. The grid was imaged in a CM100 microscope equipped with a LaB_6 electron source at a nominal magnification of 39,000x.

The images were then processed with the IMAGIC software package. EM2EM was used to convert the stack of images to imagic format and COOS was used to define the location of particles within each image. CUT-IM was used to crop the images of single particles and a band-pass filter with a low frequency cut-off of 0.086 and high frequency cut-off of 0.52 was applied.

A stack of aligned and centred particle images was produced by running 3 cycles of summing the images, calculating the rotational average of the sum and aligning and centring the stack of particle images against the rotational average. The stack of images was subjected to multi statistical analysis using MSA-RUN and the images were segregated into classes using MSA-CLASS. Classes averages were determined using MSA-SUM.

3.8 Atomic Force Microscopy of 2D crystals

AFM was performed in collaboration with Dr. Lekshmi Kailas at the University of Sheffield.

10 μl of the crystal suspension was dialysed against crystallisation buffer containing no glycerol in a 12 -14 kDa Slide-a-lyzer (Thermo Scientific). The crystal suspension was sonicated as described previously and 2 μl of the suspension was deposited on a freshly cleaved mica sheet and allowed to dry for 2 hours before imaging. Height and phase images were recorded with Silicon tips (Olympus, Japan) in Tapping mode using a Dimension 3100 AFM with Nanoscope IV controller (Veeco Instruments). Objects of interest for imaging were found by imaging 10 x 10 μm areas at 128 x 128 pixel resolution before imaging smaller 2 x 2 μm regions at 512 x 512 pixel resolution.

Chapter 4 - Purification and Crystallisation of UAC$_{Bc}$

4.1 Introduction

High resolution structure determination requires crystals of sufficient quality for imaging and to this end UAC$_{Bc}$ was overexpressed in *E. coli* in 30L cultures. The mixed membranes prepared from the cell paste were transported to Sheffield as starting material for the structural studies performed by the author. In this chapter, the results of purification UAC$_{Bc}$ and the process of refining the first 2D crystals into a sample useful for high resolution imaging will be described. There will also be descriptions of use of electron microscopy of negatively stained UAC$_{Bc}$ in detergent solution for initial structural characterisation of the protein by single particle image processing.

The first steps were the extraction of the protein from membranes using detergents, and purification facilitated by an affinity tag. The choice of detergent is important as it must maintain the protein in its correctly folded form for crystallisation experiments. Membrane proteins may be crystallised in two dimensions by the addition of lipid to the purified protein in detergent solution and the removal of detergent either by hydrophobic absorption to polystyrene Biobeads™ (Rigaud et al., 1997) or more commonly by dialysis. A functionalised lipid monolayer covalently bound to a protein ligand may also be used to aid 2D crystallisation and can impose an orientation on the protein (Levy et al., 1999, Lebeau et al., 2001). As detergent is depleted from the mixture of protein and lipid dissolved in micelles, the micelles fuse forming a bilayer with edges capped by detergent. The fusion of micelles minimises unfavourable exposure of the lipid or protein hydrophobic regions (Lasic, 1988, Rigaud et al., 2000). As detergent micelles may be too large to pass through dialysis membranes, the rate of detergent removal is dependent on the critical micellar concentration (CMC). This is the concentration where the detergent will begin to form micelles and depends on the charge of the head group and the length of the hydrophobic tail (Helenius and Simons, 1975).

Under the right conditions the protein can be reconstituted into a lipid bilayer and the protein can form ordered arrays. In the design of a 2D crystallisation/dialysis experiment, the main factors to consider are the protein, the detergent, the lipid and the buffer composition (Jap et al., 1992). The protein must be pure and homogenous for reproducibility of crystallisation. Care must also be taken with the determination of the protein concentration as it is used to set the lipid-to-protein ratio (LPR) which is the amount of lipid added. 2D crystals may form in a range of LPRs but it has been found that excessive lipid can lead to loose packing of the protein within the lipid bilayer and loss of crystalline order (Schmidt-Krey et al., 1998, Tsai et al., 2007). The type of lipid to add may be found in the native lipid membranes. The LHCII has been observed to only crystallise in the presence of the lipid digalatacosyl diacyl glycerol from its native thylakoid membranes (Nussberger et al., 1993). Conversely the synthetic lipid dimyristoyl phosphatidylcholine (DMPC) has been used to produce 2D crystals of several membrane proteins including aquaporin-0 (Gonen et al., 2005), the outer membrane porins (Signorell et al., 2007) and leukotriene C4 synthase (Schmidt-Krey et al., 2004). In the case of LHC II, the native membrane lipid was necessary for maintaining the protein in a conformation that could crystallise and in other cases excess co-purified lipid had to be removed so that the LPR could be set at a level that will allow the protein to pack into ordered arrays. In such cases, the protein may crystallise in a number of lipids and the quantity of lipid rather than the type of lipid appears to be important for crystallisation (Zhao et al., 2009).

The detergent should be removed from, or only be a low constituent of, 2D crystals (Schmidt-Krey, 2007) but the rate of detergent removal can affect the size and form of the crystals grown (Lacapere et al., 1997). It has been suggested that a low rate of dialysis can lead to large crystal growth by reducing the number of nucleation sites (Jap et al., 1992) but slow dialysis may not be possible for proteins that are instable in detergent solution.

Lastly, variations in the pH, ionic strength, temperature and other contents of the dialysis buffer will have an effect on the conformation and surface charge of the protein and its ability to form a crystal lattice (Dobrianov et al., 1999). For the sodium/proton antiporter from *Methanococcus jannaschii,* the best diffracting 2D crystals were produced at pH 4 due to the protein being locked in an inactive confromation at acidic pH (Vinothkumar et al., 2005). The photosystem I reaction centre formed two different types of lattices depending on the concentration of $MgCl_2$ (Ford et al., 1990). The CMC of a detergent is also affected by the presence of salts and chaotropic agents such as guanidine hydrochloride or urea (Midura and Yanagishita, 1995) so buffer composition should also be considered in attempts to control the rate of detergent removal.

Three models with one, two and three stage processes have been proposed for the mechanism of 2D crystallisation of membrane proteins by dialysis (Kühlbrandt, 1992). In the three stage model, a lipid bilayer is formed by the merging of lipid:detergent micelles during detergent removal. In the second stage, protein is inserted into the lipid bilayer with random orientations. In the third stage there is rearrangement of the protein within the lipid bilayer and crystal contacts being formed between protein molecules. In the two-stage model, the lipid bilayer formation and protein incorporation occurs as a single step as protein:detergent and lipid:detergent micelles merge followed by as second stage with protein rearrangement and lattice formation. The single-stage model has the protein forming crystal contacts immediately during incorporation into a lipid bilayer.

The resulting 2D crystals may take several forms with different implications for subsequent data collection and processing. Single layered sheets are most useful and can lead to the highest resolution as demonstrated by the sheet crystals of bacteriorhodopsin (Henderson et al., 1990) and light harvesting complex II (Kühlbrandt et al., 1994). However, growing single layered sheets can be difficult and it is more common to grow vesicle type crystals where the ‘skin’ of the enclosed sac is comprised of a lipid bilayer containing a protein lattice. Applied to grids, they collapse and can give rise to superimposed diffraction patterns corresponding to the lattices on both sides of the vesicle as it lies on the grid. In some cases, membrane proteins can form helical crystals which are tubes of membranes with the incorporated protein forming a helix. Whilst depicting several different views of the protein in a single image and therefore being useable for 3D reconstruction without tilting the specimen, they are processed in a different way from other 2D crystals (Unwin, 2005). Last of all, 2D crystals may form multilayered stacks which may or may not be in register. These can cause considerable problems with data collection and interpretation as illustrated by multilamellae crystals of Ca^{2+} -ATPase due to the tendency of crystal clumps to not lie flat on the grids and also due to the variability in the thickness of the crystals causing problems with merging data from tilted images (Shi et al., 1995). In such cases, determination of the structure in projection may still be possible but solving structures from 2D crystals with variations in unit cell thickness is not routinely performed.

4.2 Results

4.2.1 Protein Purification

UAC*Bc* was purified by two-step chromatography. Initially this involved nickel affinity chromatography followed by size exclusion chromatography to ensure homogeneity of the purified protein. This protocol was later changed replacing nickel resin for cobalt due to purity concerns.

On the NuPAGE pre-cast gel system, UAC*Bc* runs below its formula molecular mass of 24 kDa. With nickel affinity chromatography, the protein purified with a contaminant running at ~70 kDa on sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) (Figure 4.1). N-terminal protein sequencing performed by Dr. Arthur Moir (University of Sheffield) confirmed that the major band in the gel was UAC*Bc* and identified the contaminant band as the polymyxin resistance related protein ArnA (Williams et al., 2005). To remove the ArnA contamination and test for any changes in crystallisation experiment outcomes, His-Pur Cobalt resin was used as the substrate for affinity chromatography but this change resulted in a different contaminating band running at ~50 kDa on SDS-PAGE. It wsa possible to remove the major contaminants by performing purification with nickel resin followed by cobalt resin or vice versa but this resulted in low yields of protein making the protocol unsuitable for preparing protein for crystallisation trials. The crystals described later in the chapter could be grown using protein purified using either type of metal resin although cobalt resin was used exclusively due to a slightly greater yield of crystals grown despite the lower yield of protein from purification. During preparation of membranes prior to purification, 1 litre of culture would yield 5 ml of membranes which would then yield ~0.7 mg of protein if purified using nickel resin or ~0.4 mg protein using cobalt resin. Purification with both resins in either order would yield <0.1 mg protein.

The protein eluted from a Superose 6 3/100 size exclusion column as a symmetrical peak (Figure 4.2) and the four 0.5 ml fractions corresponding to the peak were pooled and concentrated to 0.5-0.8 mg/ml before preparation for crystallisation trials.

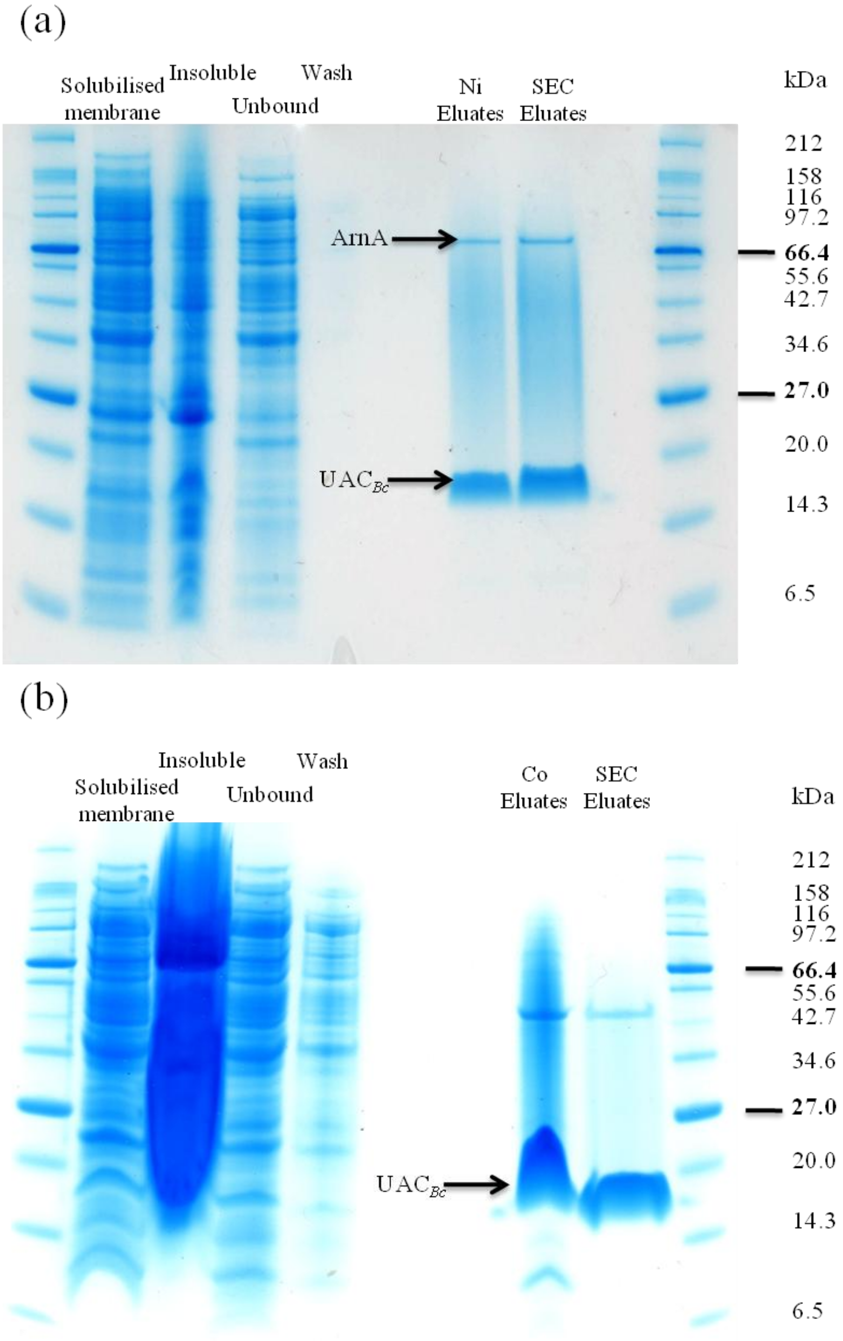


Figure 4.1 Gels showing the progress of a purification using (a) nickel- and (b) cobalt-based affinity resin. Sample identities are shown above each lane. Solubilised membrane contains the centrifuged supernatant from a mixture of the mixed membranes and solubilisation buffer and the insoluble lane contains the resuspended pellet. The unbound lane contains material unbound to the affinity resin and the wash lane contains the wash buffer eluates. Ni and Co resin and SEC eluate fractions containing protein as judged by A_{280} were pooled prior to loading on the gels. The SEC eluate lanes contain ~12 µg and ~9 µg total protein in gels (a) and (b) respectively. Arrows indicate the bands corresponding to UAC$_{Bc}$ and a contaminant ArnA. The major contaminant in gel (b) at ~45 kDa was not identified.

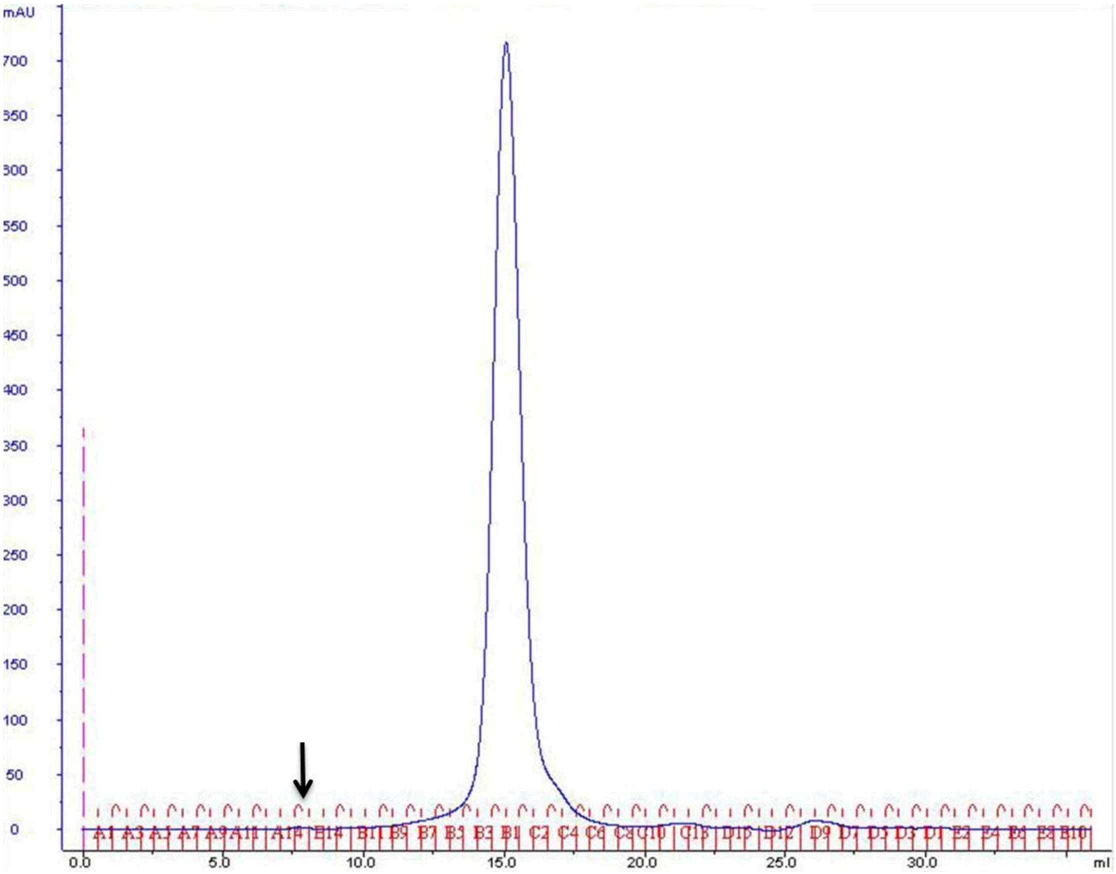


Figure 4.2 Size Exclusion Chromatography Profile.
The UV chromatogram obtained by running UAC$_{Bc}$ purified by Co affinity chromatography on a Superose 6 10/300 column. The protein elutes with a peak at 15.47 ml. The void volume is indicated by an arrow.

4.2.2 Estimation of Mass by Size Exclusion Chromatography

Protein standards with known molecular weight from the Sigma calibration kit were run on the Superose 6 10/300 column and the elution volumes were recorded. The calibration graph was used to estimate the mass of the protein:detergent particle as 224 kDa (Figure 4.3).

4.2.3 Monodispersity analysis of purified UAC$_{Bc}$ and Single Particle Processing

The purified protein in detergent was applied to EM grids and imaged to assess its monodispersity. With purified protein prepared using either nickel or cobalt resins followed by gel filtration, two types of particles could be observed on the grids (Figure 4.4). They appeared as the projections of either disk like objects and 'pearl barley-like' particle stacks with two low density stripes separated by a dark stain filled stripe. Tilting the microscope goniometer and viewing the disk particles from another angle gave the disk particles the appearance of the pearl barley type particles. Pearl barley type particles with the central stain line parallel to the tilt axis in the untilted view also gained a disk like appearance upon tilting the goniometer. These observations demonstrate that the two types of particle observed represented different views of the same species (Figure 4.5).

Single particle image processing was used in an attempt to determine the symmetry of the particle to provide information concerning the oligomeric status of the protein. Using the IMAGIC suite of software, 1900 particle images were centred and aligned. After multivariate-statistical analysis, the particles were grouped into 15 classes (Figure 4.6). As seen previously on the grids without processing, the particles segregate into either the disk or 'pearl barley' type. Subunits within the disk projections cannot be resolved but the classes with the disk view appears to have a central stain filled depression. This is consistent with the 'pearl barley' view that has the stain stripe between the low density stripes. The averaged disk is 120 Å in diameter and a segment of the pearl barley is 120 Å in length and 50 Å in width. The pearl barley type particle is likely to represent a pair of the disks stacked on each other. A similar particle was observed when particles of CaiT in detergent solution were imaged (Vinothkumar et al., 2006).

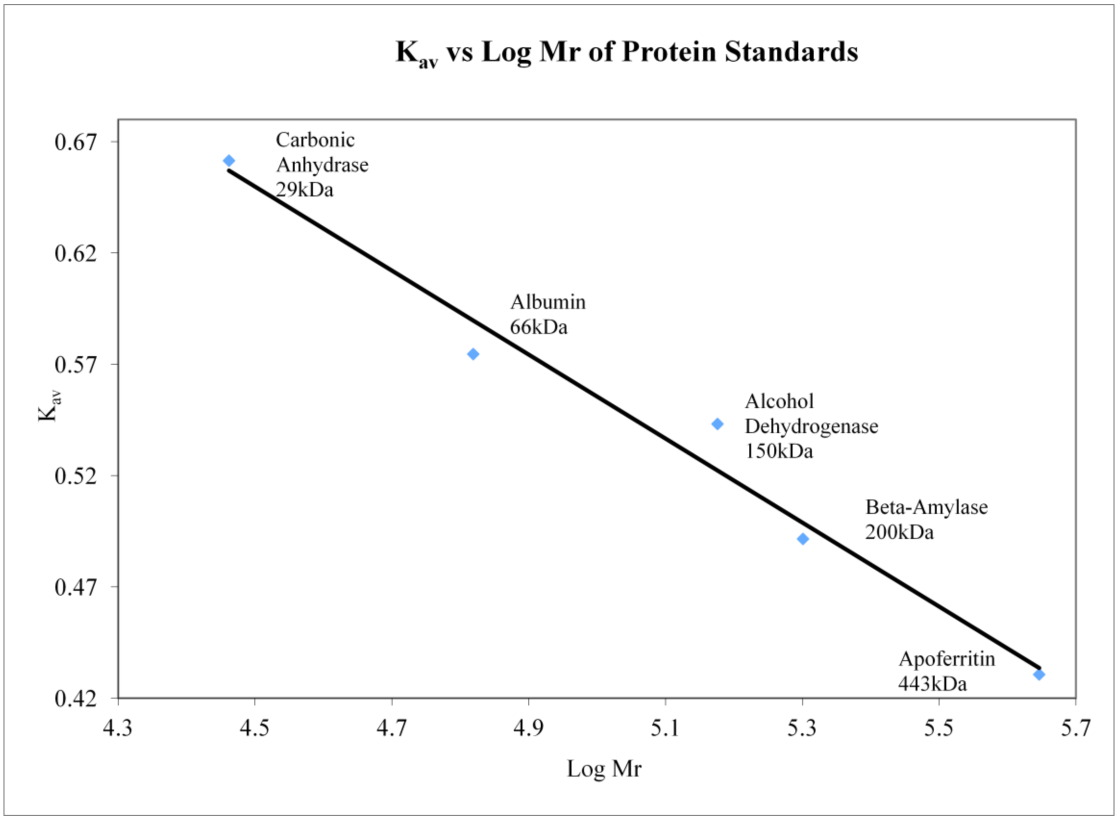


Figure 4.3 Estimation of the mass of the UAC*Bc*:detergent complex using the Kit for Molecular Weights 29,000-700,00 (Sigma). Identities of the proteins and their masses are shown next to the data points. Using the calibration line, the mass of the UAC*Bc*;detergent complex was estimated at 224 kDa from its K_{av} of 0.48. R^2 for the fitted line is 0.98.

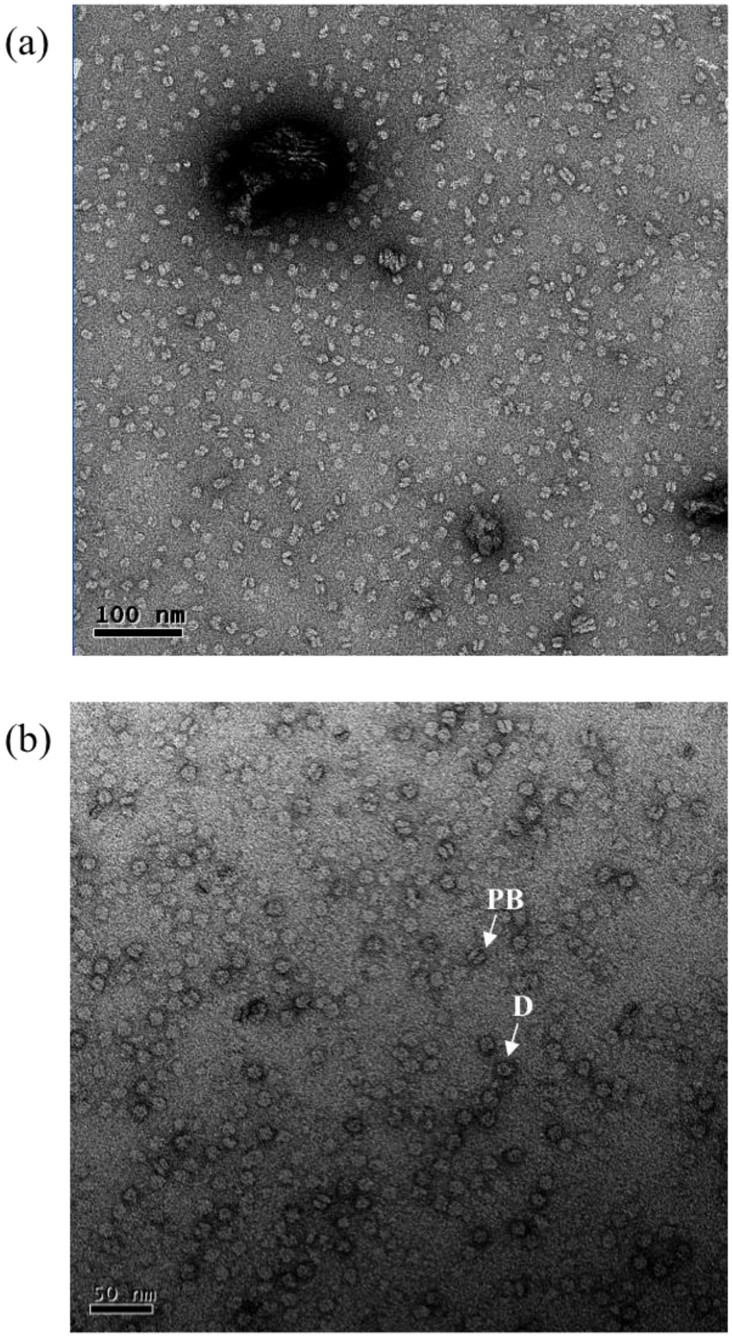


Figure 4.4 Monodispersity check on purified UAC_{Bc} from (a) nickel and (b) cobalt resin

UAC_{Bc} in detergent was diluted to ~15 µmol in gel filtration buffer and applied to a carbon coated EM grid before staining with uranyl formate. Both images show a monodisperse population of protein and two major types of particle can be observed on the grids; a round disk like particle labelled 'D' and a pearl barley resembling particle labelled PB. The pearl barley particles are likely to represent a pair of oligomers stacked on top of each other and the disk represents the oligomer as viewed perpendicular to the membrane.

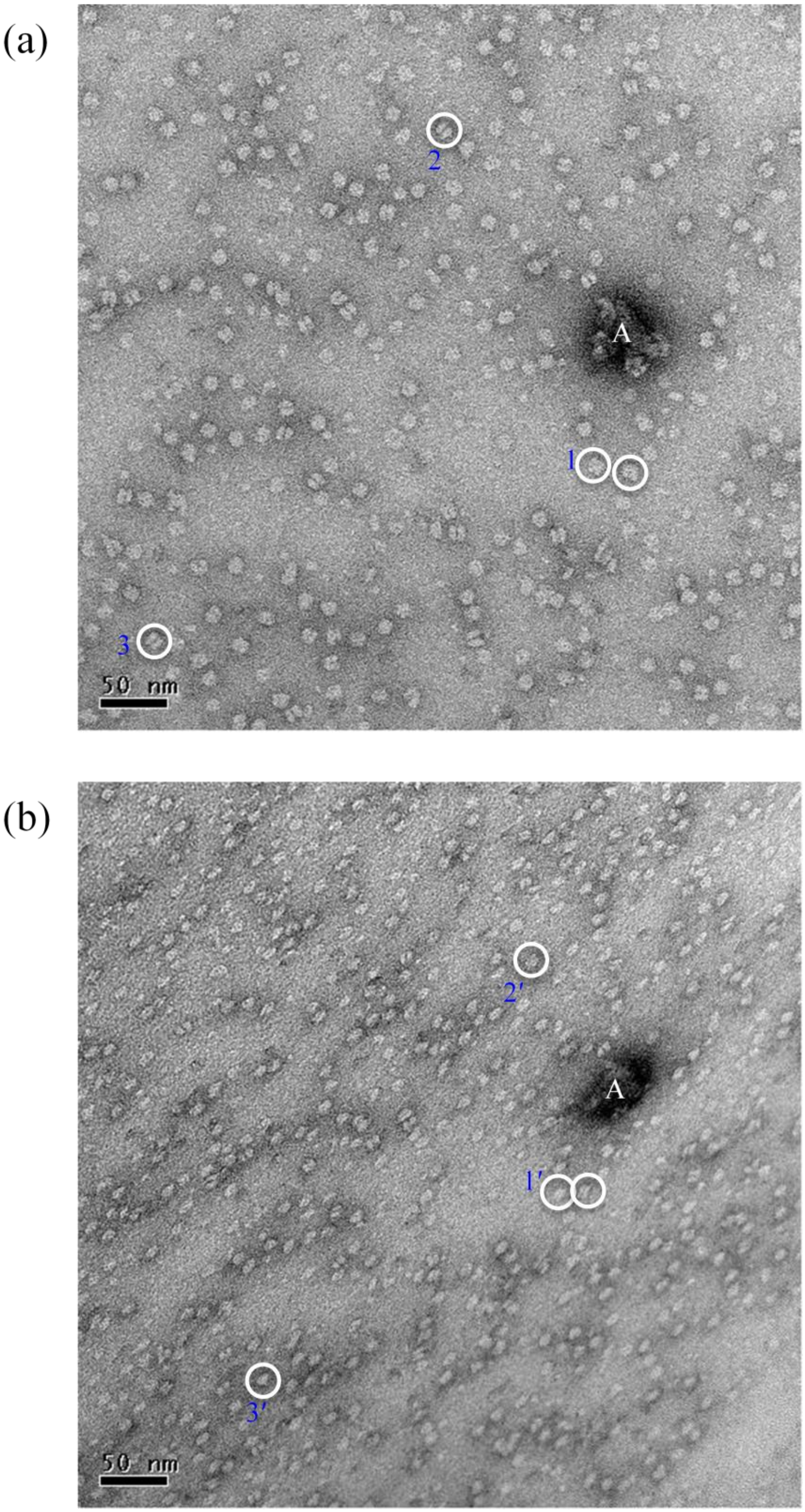


Figure 4.5 Two views of monodisperse particles of UAC_{Bc} in detergent untilted (a) and with 45° tilt applied (b). Due to the distortion in the tilted image, an aggregate labelled ‘A’ in both images was used as a reference point. The ′ symbol is used to denote the particles in the tilted image. Some disk type particles in the untilted view appear as pearl barley type particles with tilt applied whilst particle 3 gains a disk-like appearance upon tilting. The two types of particles represent different views of the same species.

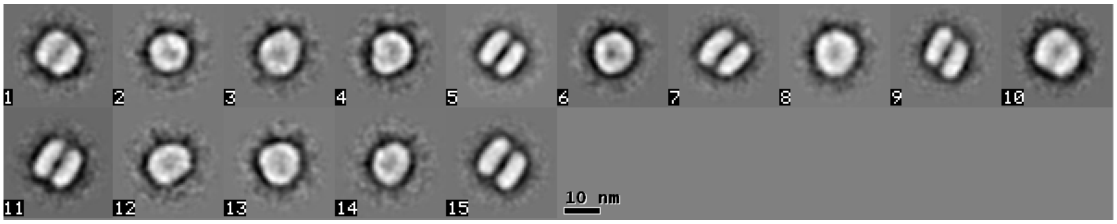


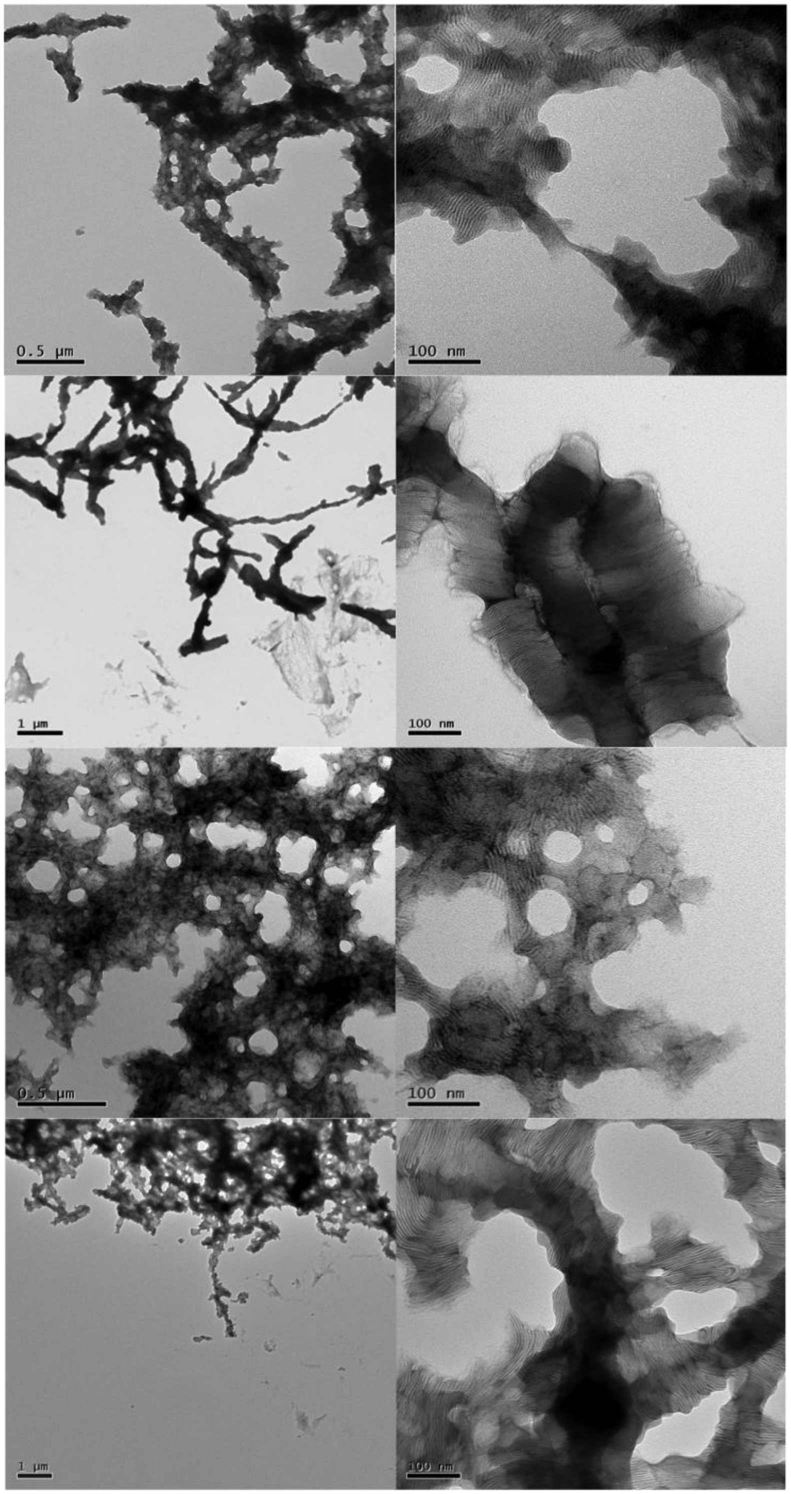
Figure 4.6 Class averages from single particle averaging
By picking 1900 particles and classifying into 15 groups, the particles group either into the disk type of particle such as in class 2 or the pearl barley like particle such as class 5.

4.2.4 Crystallisation Screening and Optimisation

The initial crystallisation screen tested a variety of important parameters for crystallisation. Prior to the author's work at Sheffield, a high throughput stability screen performed at Leeds showed that UAC$_{Bc}$ in detergent demonstrated stability in a range of pH, glycerol concentrations and presence of either NaCl or KCl (Postis et al., 2008). The initial crystallisation screen tested pH 5-9, whilst the lipid type and LPRs were randomly selected with the lipid chosen from DMPC, POPC, DOPC and ETL and the LPR (w/w) ranging from zero up to a relatively high value of 1 (Mosser, 2001). Similar to the high throughput stability screen, variations in glycerol concentration and use of either NaCl or KCl in the dialysis buffer was tested. A temperature of 25 °C was selected, a temperature used in many successful 2D crystallisations (Kühlbrandt, 1992).

After 12-14 days of dialysis, sufficient detergent was removed from the crystallisation mixture such that a drop of crystallisation mixture would have the same diameter as an equal volume drop of crystallisation buffer on parafilm. The presence of detergent in the crystallisation mixture would reduce the surface tension causing the drop to diameter to increase (Kaufman et al., 2006). The volume of the crystallisation mixture in the dialysis wells of the machine would reduce over the course of dialysis leaving 20 - 60 µl remaining. The exact reduction in volume could not be controlled but there was no correlation observed between the volume after dialysis and the quantities of crystals produced. For every finished trial, within an hour of transfer to an Eppendorf tube for storage, a cloudy sediment would settle at the bottom of the tube so each sample was resuspended with a pipette prior to loading onto grids.

Viewed on EM grids, lamella-like aggregates and aggregates of clustered vesicles were common to all conditions tested regardless of pH, salt concentration, lipid, presence of glycerol and temperature. The aggregation observed was especially severe at pH 5 and 6 where only lamella material with a striated appearance was observed (Figure 4.7). From pH 6.5 up to 8, the sample formed thick aggregates consisting of clumps of vesicles and this restricted the ability to screen for crystals to regions around the edges of aggregates where a region of a proteoliposome stuck out.


20 mM K Acetate, 2.5%
Glycerol, 100 mM KCl,
0.05% NaN_3
pH 5
LPR 0

20 mM K Acetate, 2.5%
Glycerol, 100 mM NaCl,
0.05% NaN_3
pH 5
0.4 ETL

20 mM MES, 2.5%
Glycerol, 100 mM KCl,
0.05% NaN_3
pH 6
LPR 0

20 mM MES, 10%
Glycerol, 50 mM NaCl,
0.05% NaN_3
pH 6
0.4 DMPC

Figure 4.7. Crystallisation experiments at low pH. The contents of the crystallisation buffer, the pH and the chosen lipid and LPR (w/w) are shown next to the resulting material at low magnification (left) and higher magnification (right). LPR 0 means that no lipid was added prior to dialysis. After dialysis against pH 5 or 6 buffer in all conditions tested, highly striated material was produced representing the formation of lamella structures.

In the initial screen, two conditions produced small crystals both using NaCl in the crystallisation buffer, with at least 10% glycerol at pH 6.5 and 8 and using the lipids POPC and DOPC (Figure 4.8). These observations were considered in the design of the following optimisation screen which limited the pH range from pH 6.5 to 8 and tested more LPR conditions using POPC and DOPC. It was found that UAC_{Bc} would crystallise reproducibly in DOPC and also form larger continuous crystalline regions than those grown with POPC. The protein produced higher crystal yields at pH 7.5. From these initial conditions, the optimisation of crystallisation experiments tested variations in the salt, glycerol and protein concentration. 100mM NaCl and 10% glycerol (v/v) were optimal for reproducibly growing crystals. A protein concentration of 0.5 mg/ml was optimal for growing larger crystals whilst attempting crystallisation with a protein concentration of 1 mg/ml resulted in more vesicle aggregates and low yields of mainly poorly ordered crystals. Using a lower protein concentrations resulted in similar yields of crystals to those obtained at 0.5 mg/ml protein but these crystals were smaller. Increasing the temperature during crystallisation from 25 °C to 30 °C also caused the growth of larger vesicles with some exceeding 1 μm^2.

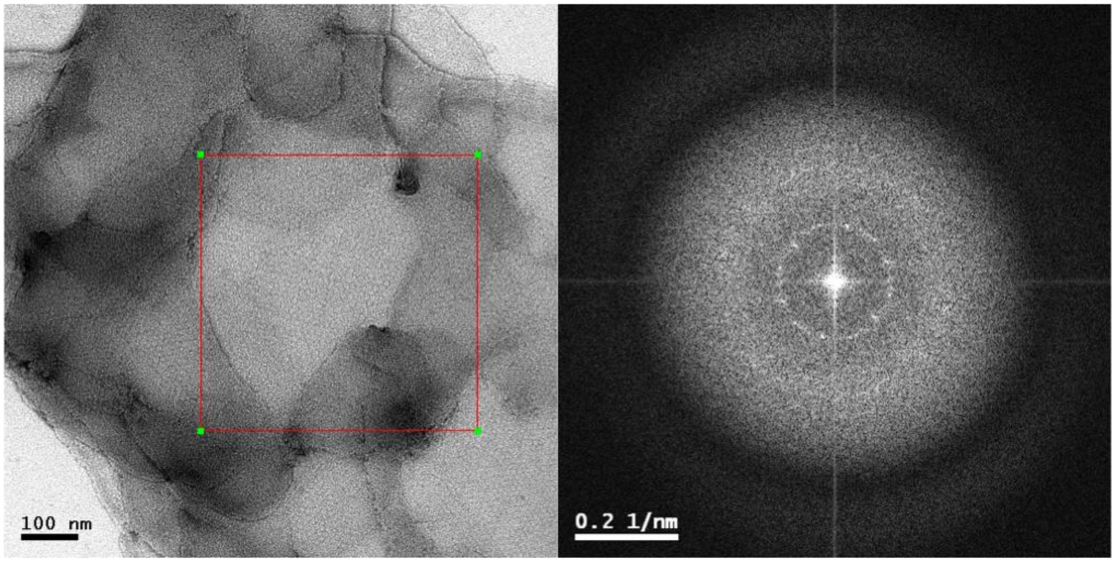
4.2.5 Effect of LPR on Crystallisation

The LPR appeared to be the most important parameter for producing well-ordered crystals of the protein despite observations of crystals across a broad range of LPRs. The observed trend in the crystallisation experiments was that at a low LPR of typically 0.3, the protein reconstituted into small proteoliposomes with no crystalline order. By raising the LPR from this point, 2D crystals could then be found by a certain LPR, typically around 0.4. Although the sharpness of the spots in the computed Fast Fourier Tranforms (FFT) of crystal images varied greatly between crystals grown under a given crystallisation condition, the ideal LPR would give the greatest concentration of well ordered crystals with up to 20-30% of crystals showing sharp diffraction spots (Figure 4.9).

The sizes of the crystalline regions produced in the crystallisation trials also increased with LPR. However, at higher LPRs, the diffraction spots in the FFTs of crystals become more diffuse indicating a general decrease in crystalline order. By increasing the LPR too high, large sheets of reconstituted protein spanning 2 μm^2 and larger were produced but powder diffraction or diffuse spots rather than sharp diffraction spots could be found. It was found that the optimal LPR which varied between 0.5 and 0.7 across different preparations would produce crystal regions of up to ~1 μm^2 with some of the largest crystals showing sharp diffraction spots.

(a)

20 mM MES, 20% Glycerol, 10 mM NaCl,
0.05% NaN3 pH 6.5
0.4 POPC



(b)

20 mM HEPES, 10% Glycerol, 100 mM NaCl,
0.05% NaN3 pH 7
0.8 DOPC

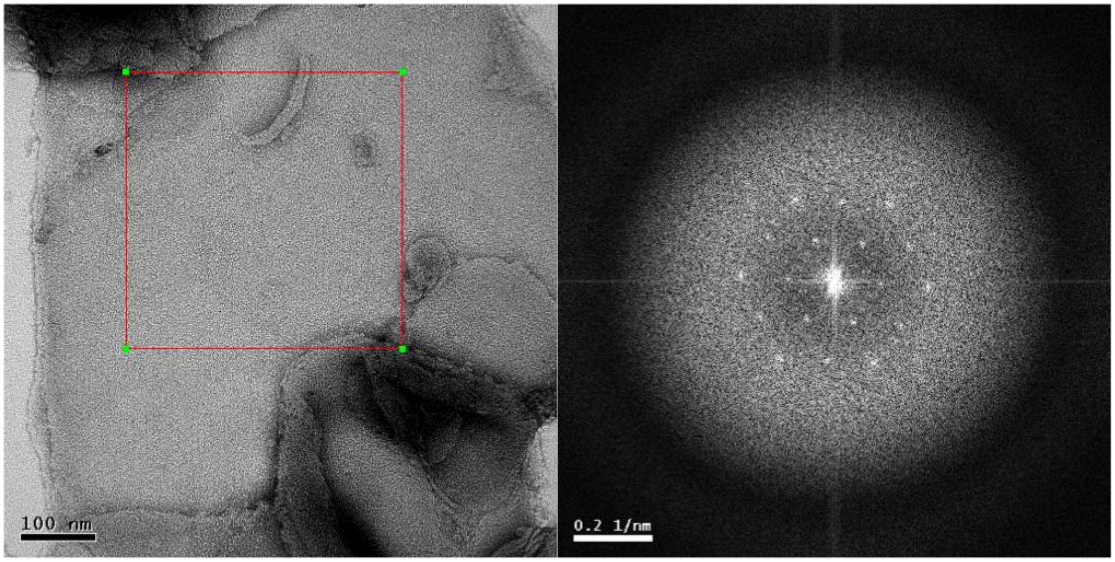


Figure 4.8 First successful 2D crystallisation of UAC_{Bc}

The crystallisation buffer, lipid and LPR (w/w) used in the experiment is shown above each pair of images. The images on the left show a negatively stained crystal as viewed on a carbon coated film and the image on the right shows the computed FFT of the region of the image within the red square.

(a) The crystal grown with POPC is of a mosaic crystal consisting of several crystalline regions giving rise to the ring in the FFT.

(b) The crystal grown with DOPC has better crystalline order and gives rise to distinct spots in the FFT.

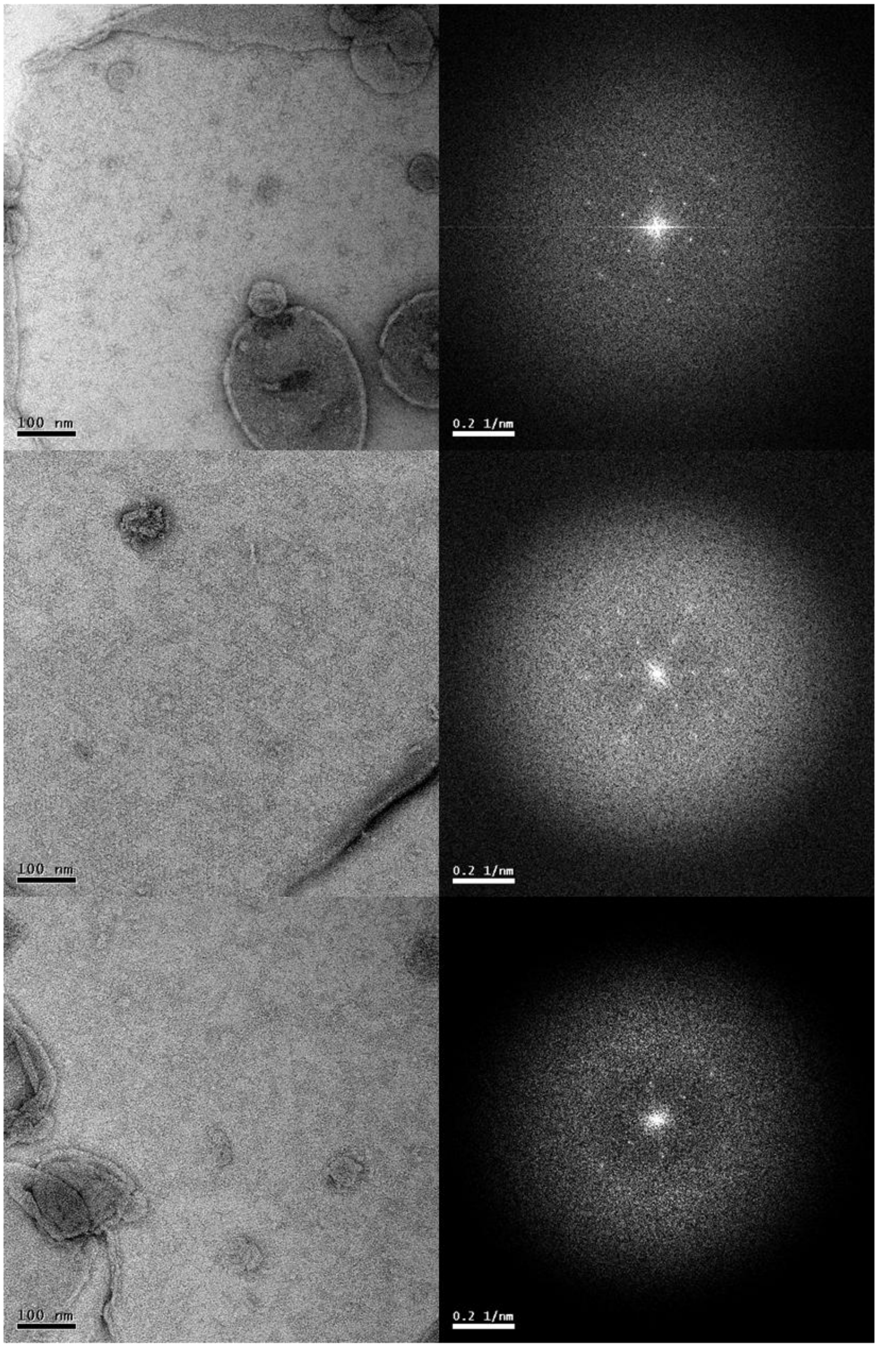


Figure 4.9 Variation of Crystal quality in a single sample
3 different crystals from a single crystallisation trial have been imaged in negative stain and their computed FFTs shown on their right. For the top crystal, both the first and second order diffraction spots are relatively sharp in the computed FFT. For the middle crystal, the diffraction spots are more diffuse due to lesser ordering of the protein. For the bottom crystal, only a diffuse first order of diffraction spots is visible. The proportion of each type of crystal represented in the entire population of crystals appears to depend on the LPR used in the crystallisation.

4.2.6 Effect of Sonication on Crystals

Early attempts at using the crystals for high resolution imaging were unsuccessful due to difficulty in distinguishing suitable areas around the aggregates for imaging. Attempts were made to disaggregate the crystals prior to EM grid preparation. Using the probe sonication protocol outlined in Chapter 3 Materials and Methods and applying the sample to grids resulted in several dispersed crystalline sheets (Figure 4.10). Sonication was an effective indicator of the yield of crystals as vesicles without crystalline order disintegrated into smaller vesicles whilst 2D crystals tended to remain intact. With crystallisation trials with low yields of crystals such as those grown at high LPR, sonication resulted in complete disintegration of the sample into small vesicles. Any remaining large sheets were almost certain to have some form of ordering as indicated by spots on the FFT. There was no observable difference in the disaggregation of the sample if the sonication time was increased from 5 seconds to 20 seconds. If the sample was sonicated for 1 minute at 3.5 microns, the vesicles on the grid were non-crystalline and relatively small and amorphous material appearing to be denatured protein was found. Increasing the sonication output to 10 microns also resulted in disintegration of the sample and denaturation of the protein within 5 seconds of sonication.

A more conservative sonication protocol used for AFM sample preparation used only 3 seconds of sonication at 3.5 microns leaving some aggregates unbroken but revealing more thin sheets around the edges of aggregates for AFM. This protocol was developed due to the impracticality of searching relatively large areas of the mica slide support for crystals using the AFM. Crystals were observed to reaggregate into clusters within 2 days of sonication meaning that additional sonication treatments were required before preparing the sample for electron microscopy or AFM (Figure 4.11).

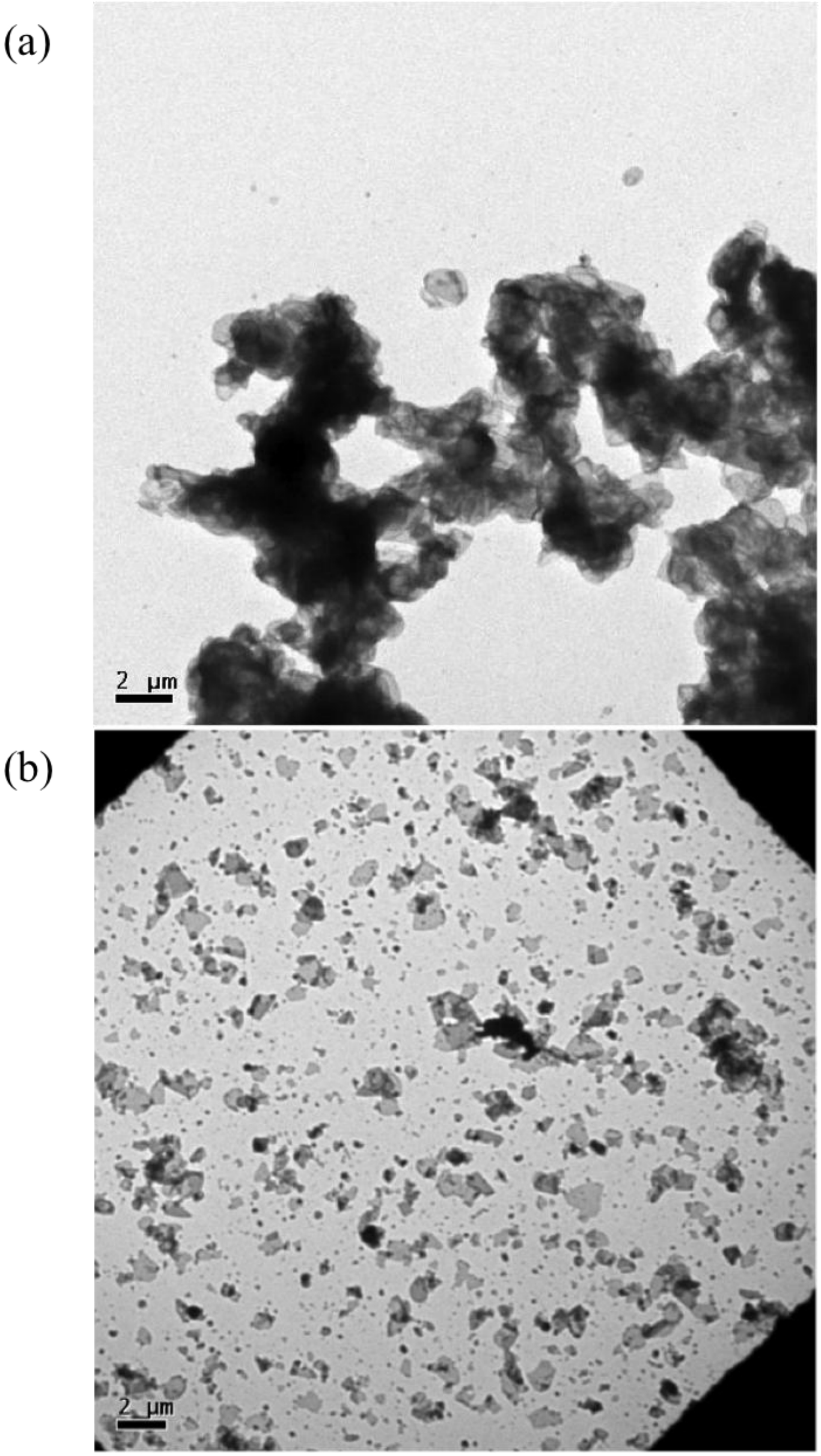


Figure 4.10 The effect of sonication on the crystal sample

(a) The mixture containing the aggregated crystals stained with uranyl formate on a grid.

(b) The same mixture of crystals after sonication treatment. The dispersed sheets are more suitable for imaging then the aggregated sheets in (a).

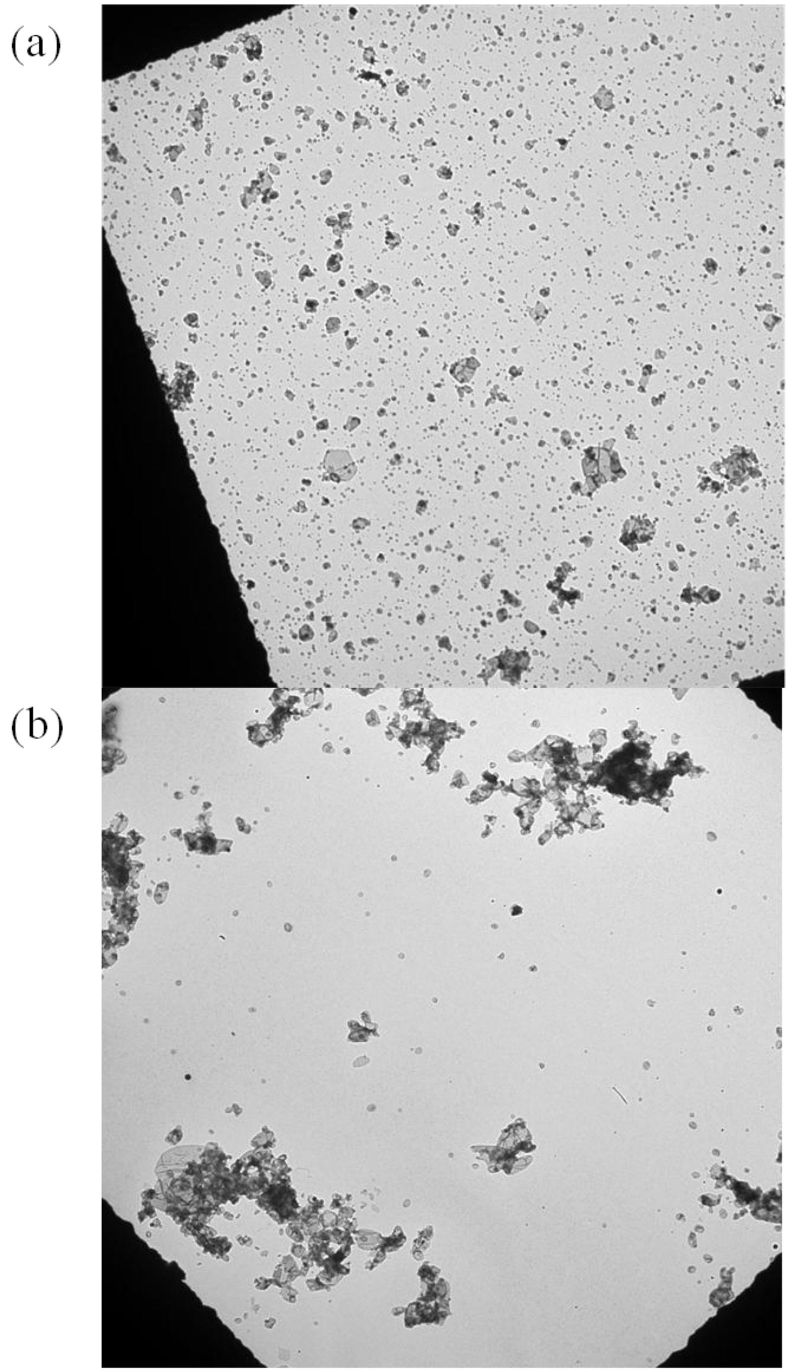


Figure 4.11 Reaggregation of 2D crystals
As a further demonstration of the tendency of UAC$_{Bc}$ crystals to aggregate, the freshly sonicated sample was applied to a grid (a).
The same sample applied to a grid a week later showed reaggregation (b). For all high resolution image collection attempts and AFM imaging, it was necessary to sonicate the sample prior to preparation for imaging.

4.2.7 Effect of Temperature Treatment on Crystals

The effect of sonication is to also raise the temperature of the sample although measuring the exact temperature increase in the crystal mixture was not possible due to the small volumes of sample involved. The effect of the temperature increase on the 2D crystals was tested independently by incubating the sample in a water bath set at an arbitrary temperature of 45°C for 15 minutes. The temperature increase appeared to induce crystallisation in some of the reconstituted vesicles with higher yields of crystals observed in a temperature treated sample than in the same untreated sample. However, the new crystals did not appear to be highly ordered and would show up to two orders of diffuse spots in the FFTs. Additional cycles of heating and cooling did not result in further crystallisation so a single heat treatment was sufficient. The effect of heating was observed in samples with low yields of crystals after dialysis. For samples that already had high yields of crystals after dialysis, the effect of temperature treatment was indiscernible.

4.3 Discussion

4.3.1 Purification and Oligomeric State of UAC$_{Bc}$

On SDS-PAGE, UAC$_{Bc}$ runs faster than its formula mass of 24 kDa (Figure 4.1). This is not unusual for a membrane protein and may be caused by anomalous binding of SDS to the protein (Rath et al., 2009).

UAC$_{Bc}$ was purified successfully with both nickel and cobalt resins albeit with different contaminants (Figure 4.1). There were two reasons for attempting to circumvent the ArnA contamination observed with purification with the nickel resin (Figure 4.1). Firstly, to confirm that the two types of particles observed when purified protein was applied to grids (Figure 4.4) represented UAC$_{Bc}$ and secondly to ensure that the crystals produced were of the target protein. This was achieved with cobalt resin for purification but resulted in a new contaminant (Figure 4.1b). The yield of protein from the cobalt resin was lower but the yield of crystals was higher. In later crystallisation experiments where the optimal dialysis buffer had been determined, it was still possible to grow good quality crystals with as little as 0.4 mg of pure protein by screening a series of LPR conditions.

The appearance of the near symmetrical peak from gel filtration (Figure 4.2) and monodisperse set of particles (Figure 4.4) is in agreement with light scattering experiments which demonstrate stability of UAC$_{Bc}$ in DDM solution (Postis et al., 2008). This may have indicated the amenability of the protein to crystallisation as the protein should be stable enough to reconstitute into a lipid bilayer before denaturation and aggregation of the protein.

The estimated apparent mass of the protein:detergent complex of 224 kDa also indicated that protein exists as an oligomer during purification: a DDM micelle has a mass of 66 kDa (Slotboom et al., 2008) and the predicted formular mass of the protein is 24 kDa.

4.3.2 Single Particle Averaging

In practice, it was not possible to achieve sufficient resolution to visualise individual subunits within the disks (Figure 4.6). Also, attempting to determine the mass of the disk using its volume and an assumed protein density of 1.35 g/cm^3 would result in an overestimation of the mass of the disk. It is not known how much detergent accounts for the volume of the disk. The issue is later addressed in Chapter 6 using size exclusion chromatography multi angle light scattering to determine the mass of protein in the particles.

The 120 Å width of the disk view of the protein is larger than the unit cell of the crystal (Chapter 5) but this is consistent considering the belt of detergent surrounding the oligomer has a stain excluding effect. This observation further supports the oligomeric nature of the protein in detergent.

4.3.3 2D Crystallisation

The major challenge with structural analysis of UAC$_{Bc}$ is producing high yields of large and well ordered crystals. UAC$_{Bc}$ has a propensity to form semi-ordered arrays but only a fraction of crystals grown in a given crystallisation condition were the high quality crystals required for high resolution imaging. The most important factor for obtaining a higher proportion of good quality crystals was the LPR so most crystallisation experiments involved screening a range of LPRs.

The observation of more crystals observed on the grids after heating the samples than in the same untreated sample also suggests crystallisation may occur after reconstitution of the protein which is consistent with the three or two stage models of crystallisation (Kühlbrandt, 1992).

4.3.4 Aggregation of 2D Crystals

The formation of aggregated and multilamellae 2D crystals is not uncommon and has been reported as a problem for other membrane proteins including LHC II (Barros and Kühlbrandt, 2009) and IIC mannitol transporter (Stuart et al., 2004). By performing crystallisation of UAC$_{Bc}$ at a pH range between 6.5 and 8, there were significantly more aggregated vesicles rather than the multi-lamellar structures seen at pH 5 and 6 (Figure 4.7) but formation of unilamellar crystals was not observed in any of the tested crystallisation conditions. The 2D crystals aggregate despite variation in the temperature, salt concentration and glycerol concentration of the crystallisation buffers and multilamellar structures were observed with all the lipid types tested for crystallisation. A model for the crystallisation of the IIC-mannitol transporter suggested that reconstitution of the protein into unilamellar sheets is followed by their fusion during which there is a phase separation into a protein rich phase and a lipid rich phase resulting in multilamellar structures. The protein in the protein rich phase then arranges into 2D crystals and so the formation of crystals is inextricable from the formation of multilamellar structures (Stuart et al., 2004).

The approach taken with UAC$_{Bc}$ was to produce large and well ordered two-dimensional crystals using the optimal crystallisation conditions (described in Chapter 3 Materials and Methods) before applying sonication to disperse the aggregated vesicles. Other possible techniques that may disaggregate the crystals and have yet to be tested include French press and Millipore extrusion (Lasic, 1988).

Imaging the exposed vesicles at the edges of aggregates in the untreated samples led to collection of images of crystalline and non-crystalline samples. With the sonication treated sample, 2D crystals and semi-ordered arrays disproportionately represent the remainders of any sheets beyond 0.5 μm^2 in size. This was useful for low dose imaging as it provided a clear method for selecting 2D crystals for image collection.

4.3.5 Current Status of Crystallisation

In all crystallisation attempts UAC$_{Bc}$ produced at least semi-ordered arrays with diffuse first order diffraction spots corresponding to a resolution of ~95 Å. There remains a difficulty in producing large well-ordered crystals suitable for regular data collection in electron cryo-microscopy experiments. A well-ordered crystal of 1 μm^2 in size has been suggested as the minimum requirement for determination of a high resolution 3D structure (Schmidt-Krey, 2007). From negative stain screening, crystals with both the appropriate size and crystalline order for high resolution data collection were relatively rare. The criterion for samples chosen for the data collection attempts in later chapters was that sheets had to have a size least 0.5 μm^2 and show two orders of diffraction.

4.4 Conclusions

UAC$_{Bc}$ could be purified with either nickel or cobalt affinity chromatography with lower yields of protein from cobalt resin. Both preparations formed stable aggregates that appear either as disks or pearl barley-like particles depending on their orientation on electron microscopy grids suggesting an oligomeric organisation of the protein but attempts to characterise the structure of the oligomeric protein in detail were unsuccessful.

UAC$_{Bc}$ reproducibly formed arrays in the conditions identified and in some cases formed large and well-ordered crystals. After treatment with sonication to disaggregate the crystals, they became promising specimens for high resolution structure determination through electron cryo-microscopy.

Chapter 5 - Projection Structure Determination of UAC$_{Bc}$

5.1 Introduction

5.1.1 Projection Structure Determination

Determination of a projection structure is an essential milestone in the 3D analysis of membrane protein 2D crystals using EM (Renault et al., 2006). Conditions for reproducibly growing crystals of UAC$_{Bc}$ were described in Chapter 4 and the work described in this chapter is concerned with the collection of electron cryomicrographs of unstained 2D crystals of UAC$_{Bc}$ and their computational processing.

High resolution imaging requires the ordering of the crystals to be maintained within the microscope vacuum. For this the author used the glucose embedding preservation technique first described by Henderson and Unwin (1975) and recorded images under low dose conditions (Williams and Fisher, 1970) onto photographic film.

5.1.2 Atomic Force Microscopy

AFM uses a tip attached to a cantilever to raster scan the surface of a sample. Deflection of the tip due to proximity of the tip to the surface is detected by reflection of a laser beam off the cantilever to produce surface topology images (Binnig et al., 1986). It is a complementary tool to electron microscopy by providing data on the variations in height of a surface which is lost in the images from TEM which are projections. This was useful for investigating the thickness of UAC$_{Bc}$ 2D crystals where the symmetry analysis suggested a possible double layered crystal form.

AFM images can be used to produce high resolution images of membrane proteins as demonstrated by the imaging of a polypeptide loop connecting transmembrane helices in the native state of bacteriorhodopsin in its native conformation (Moller et al., 1999). Flexible regions such as these are difficult to resolve using crystallographic methods (Ford and Holzenburg, 2008).

5.2 Results

5.2.1 Grid Preparation for Electron cryomicroscopy

The crystals embedded in glucose were clearly visible and well dispersed (Figure 5.1) although suspending the grids in the 1% glucose drop for longer than 5 seconds would typically result in a grid with an overly thick layer of glucose that moved upon exposure to the electron beam making large regions of the grids unsuitable for imaging. The use of glucose allowed grids to be screened for suitable objects for image collection before cooling to liquid nitrogen temperature.

5.2.2 Image Recording

170 images of untilted crystals were recorded under low dose conditions. Of these, 31 images of untilted crystals showed diffraction spots when examined on an optical diffractometer. The two best images showed spots visible at ~16 Å (images 5118 and 5276). All images showing spots were digitised but only 11 images (including 5118 and 5276) had strong enough diffraction that they could be processed.

5.2.3 Image Processing

Digital Micrograph 3 was used to analyse the locations of the images giving rise to diffraction spots on the FFTs and Photoshop was used to crop a generous square region centred on the location of the clearest and least diffuse diffraction out of the digitised images for processing. Intermediate processing results are shown in Figures 5.2 and 5.3. After each image was processed once, the cross-correlation map was used to crop a smaller square corresponding to the best crystalline area from the original image. These cropped images were then reprocessed from the beginning giving a better signal to noise ratio and the images were analysed using ALLSPACE to determine possible plane groups for the crystals. The processing results with the best signal according to the quality value (QVAL) and the lowest phase residuals from ALLSPACE are summarised in Table 5.1. QVAL is determined with the formula QVAL = R□* (IQ1 * 17.5 + IQ2 * 12 + IQ3 * 8 + IQ4 * 5 + IQ5 * 3 + IQ6 * 2 + IQ7)/500.0 where IQ1 to IQ7 gives the number of spots with the corresponding IQ, R is the height of the central pixel of the averaged Fourier peak profile determined by

MMBOX and is divided by the calibration factor 500.0 for display purposes (Gipson et al., 2007).

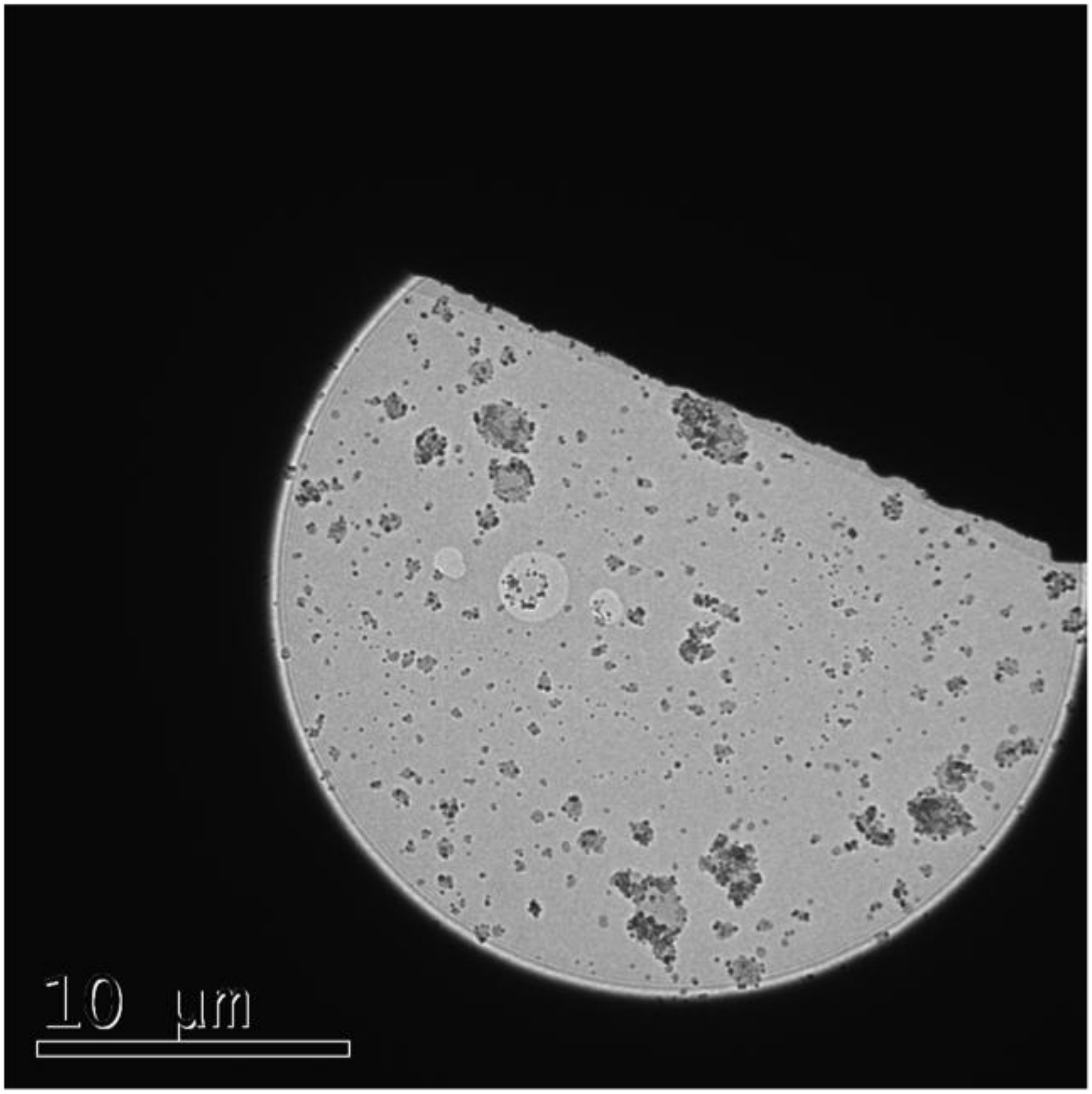


Figure 5.1. A low magnification view of glucose embedded crystals during electron cryo-microscopy.
In the middle of the image, an object was selected for low dose imaging on film. Due to exposure to the beam, a lighter circle corresponding to the region recorded on film is visible. The two smaller circles on both sides correspond to the areas used for determining and applying underfocus and correction for astigmatism.

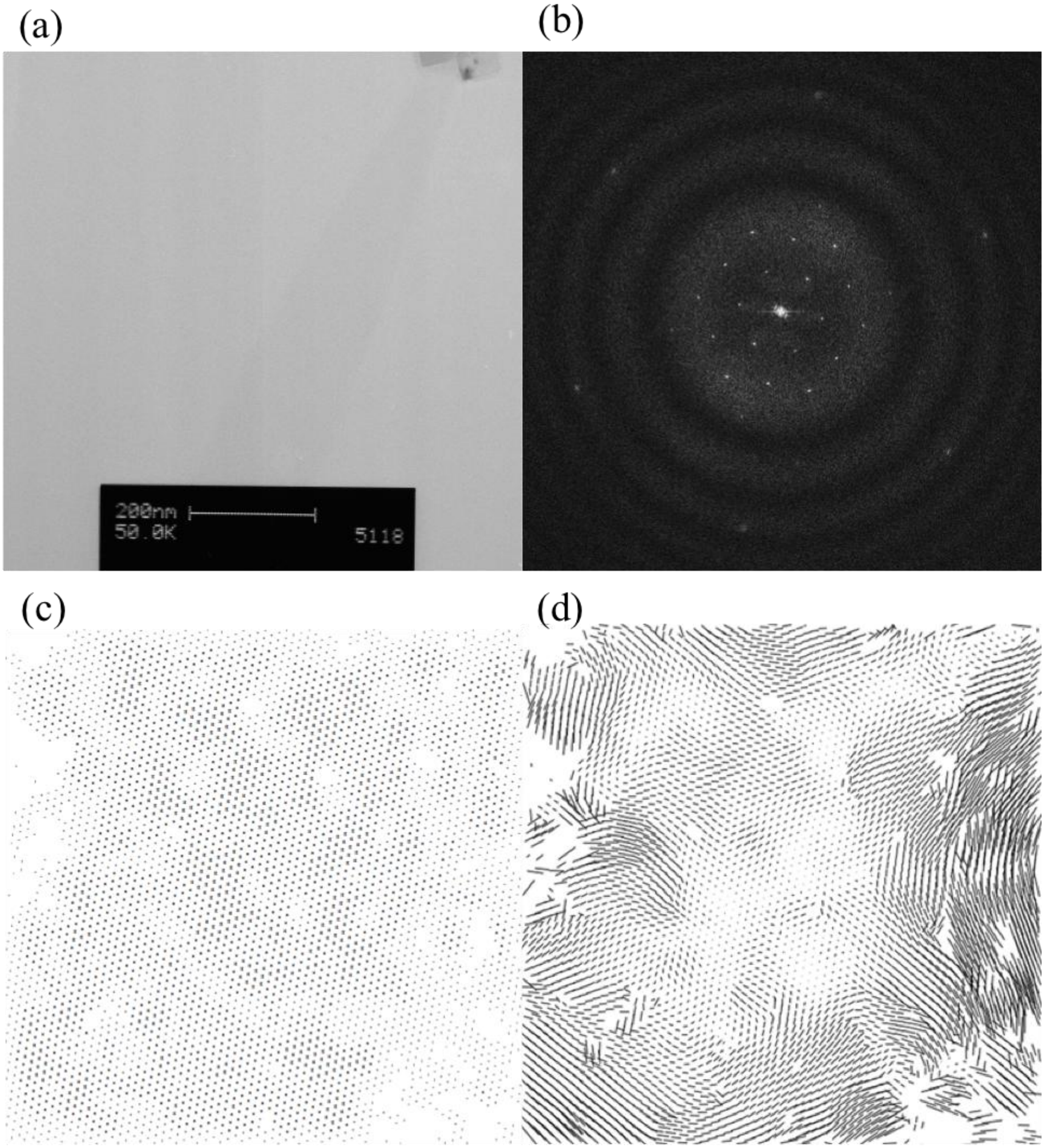


Figure 5.2. Steps in processing of image 5118
(a) Raw Digitised Image. A square pixel array is cropped from the image for processing.
(b) Computer generated Fourier transform of the processed image. The edge of the box is ~ 1/14.3 $Å^{-1}$.
(c) Cross-correlation Map. The darker symbols indicate higher correlation with a selected reference area which is typically the center of the image.
(d) Deviations from the ideal crystal lattice are represented by vectors 10x actual deviation.

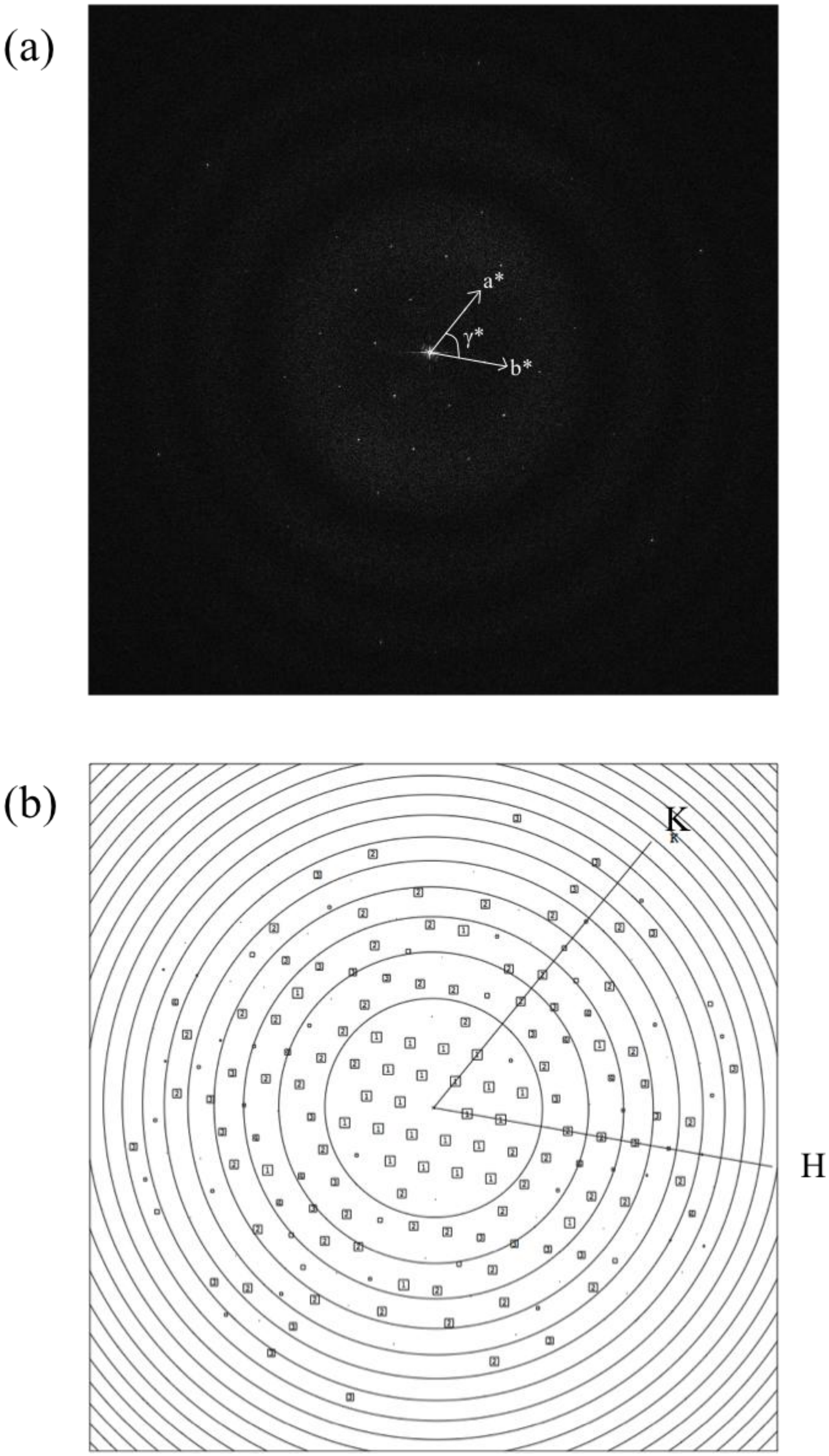


Figure 5.3 (a) Fast Fourier Transform of image 5118 after two rounds of unbending. Due to the unbending, the spots have a sharper appearance.

(b) The CTF of image 5118 with the concentric rings at the locations of zero contrast. The resolution at the square border is 8 Å. During processing, spots are assigned IQ values corresponding to their signal to noise ratio. Lower IQ spots are presented in larger squares.

Film No.	Unit Cell Dimensions			QVAL After Unbend 2	Best Plane Group	No. Spots for Comparison	Phase Residual	Target Phase Residual	Resolution max for Plane Group Determination
	a (Å)	b (Å)	γ (°)						
4980	95.8	99.9	119.6	35.3	*p*3	8	3.9	18.6	12
5108	92.3	92.1	120.1	64.7	*p*3	48	36.4	24.9	12
5109	96.2	97.3	120.7	131.2	*p*6	132	29.9	27.2	15
5113	91.9	91.2	120.1	148.1	*p*3	52	32.3	25.1	12
5116	95.6	97.9	120.1	58.9	*p*3	28	20.3	26.5	10
5118	92.7	91.9	119.4	239.3	*p*6	328	26.7	30.0	9
5124	93.5	95.7	119.6	137.3	*p*622	292	42.7	23.9	15
5238	92.4	91.9	120.1	241.9	*p*6	192	53.9	25.9	10
5261	102.8	100.8	120.1	120.7	*p*622	352	23.8	27.9	12
5276	95.9	95.7	119.9	416.8	*p*6	294	23.8	21.6	9
5279	94.7	94.1	118.5	51.7	*p*3	32	11.9	29.1	10

Table 5.1 Processing of the 11 films with clear diffraction spots
The quality value (QVAL) is a weighted sum of the IQ values of an image after 2 rounds of unbending determined by a 2dx script (Gipson et al., 2007) and has been presented for comparison of the quality of the image processing. The highest symmetry consistent with the data is shown. Phase residuals were determined using spots with IQ1 to IQ5 up to the specified resolution.

The internal phase residuals for the crystals are consistent with either a *p*6 or *p*622 crystal form. Images 5118 and 5276 had the highest QVAL2 indicating a higher signal and ALLSPACE analysis showed internal phase residuals consistent with a *p*6 crystal form (Tables 5.2 And 5.3). ALLSPACE analysis of images 5109 and 5261 (Tables 5.4 and 5.5) in addition to *p*6 symmetry also suggested the plane group *p*622. If *p*622 were the correct symmetry to impose, the crystals could consist of two layers of crystalline protein. The maps produced from processing images 5276, 5118 and 5109 are shown in Figures 5.4, 5.5 and 5.6 respectively.

Due to specimen movement during image collection, resolution for some images was strongly anisotropic. For images 5108, 5113, 5124 and 5238 which produced uninterpretable *p*1 maps, anisotropy also resulted in ambiguous plane group determination due to loss of spots for phase comparison as indicated by the high phase residual values above the expected phase residual given the noise level (Table 5.1). Images 4980, 5108, 5116 and 5279 were collected from poorly ordered crystals with only clear spots at ~45 Å viewed on the optical diffractometer and gave few spots for phase comparison. The relatively low numbers of spots with IQs 1 and 2 for data significantly above the noise level resulted in a correspondingly low QVAL (Table 5.1). ALLSPACE suggested the most likely plane group to be *p*3 for images 4910, 5108, 5116 and 5279 which is consistent with either a *p*6 or *p*622 crystal form. However, the phase residual values for the higher plane groups were significantly above those expected based on the noise level for images 4980, 5108 and 5116 with *p*6 phase residuals of 45.9, 44.2 and 44.0 respectively which can considered as random for a phase residual for a centrosymmetric plane group.

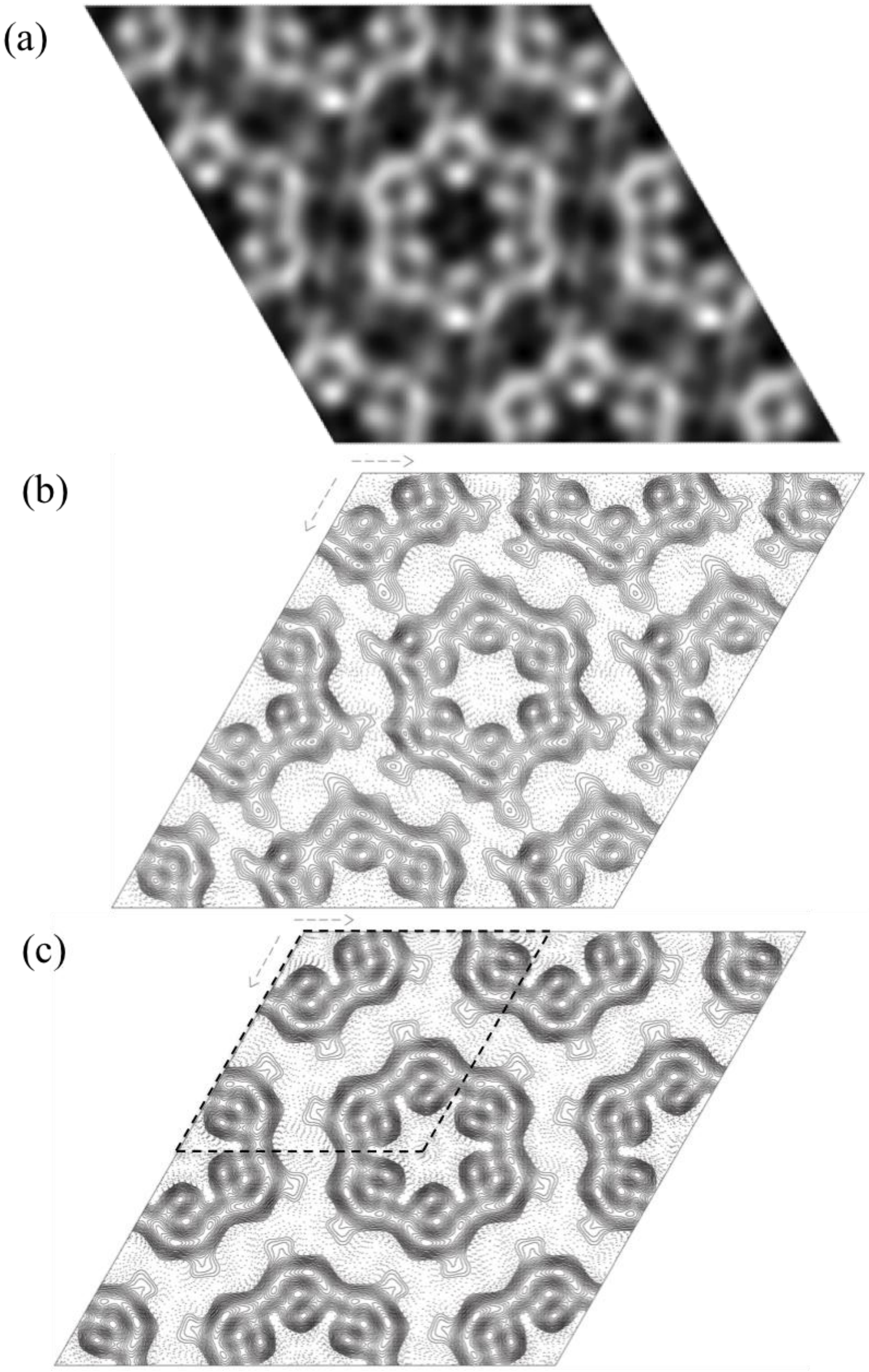


Figure 5.4 Projection Maps for Image 5276 at a resolution of 9 Å. Contour intervals are drawn at ~0.12 x root mean square density with density above the mean in solid lines and density below in dotted lines. Unit cell dimensions are a = b = ~95 Å and γ = 120°.

(a) Grey-scale map with the white regions representing the regions of density above the mean.
(b) Contour map in *p*1.
(c) Contour map in *p*6. A unit cell is outlined with dashed lines.

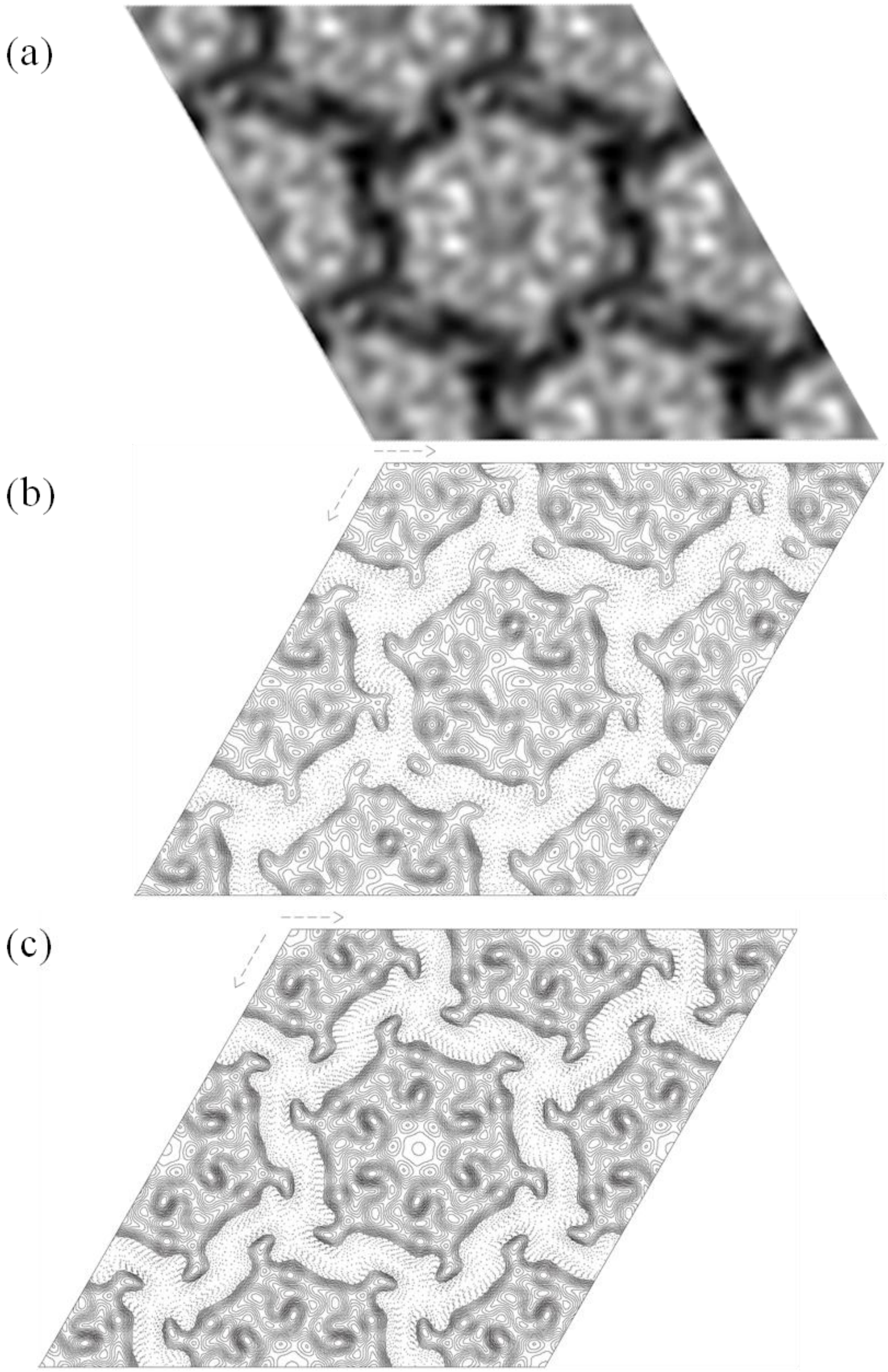


Figure 5.5 Projection Maps for Image 5118 at a resolution of 9 Å. Contour intervals are drawn at ~0.12 x root mean square density with density above the mean in solid lines and density below in dotted lines. Unit cell dimensions are a = b = ~92 Å and $\gamma = 120°$.

(a) Grey-scale map with the white regions representing the regions of density above the mean.

(b) Contour map in *p*1.

(c) Contour map in *p*6.

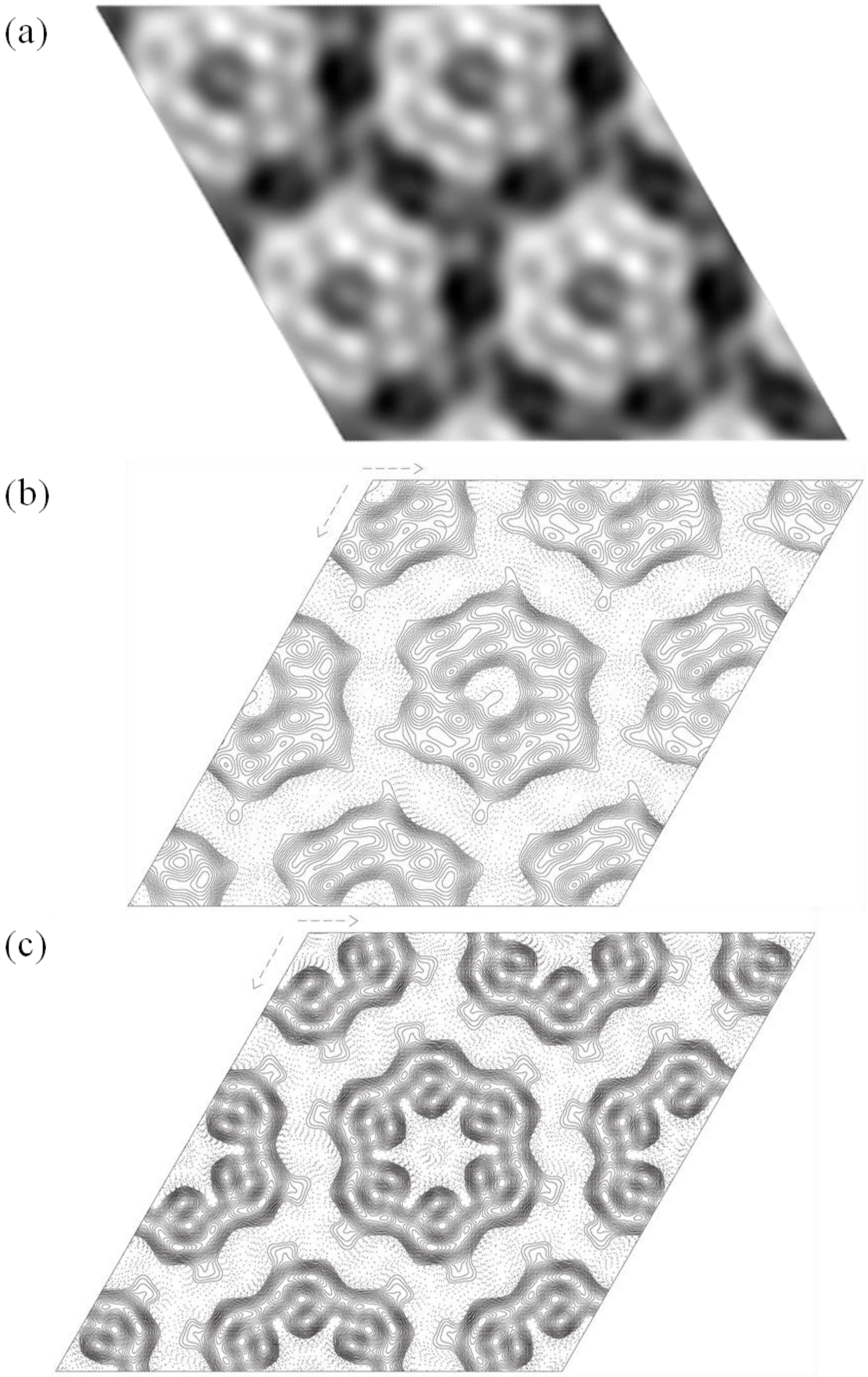


Figure 5.6. Projection Maps for Image 5109 at a resolution of 15 Å. Contour intervals are drawn at ~0.12 x root mean square density with density above the mean in solid lines and density below in dotted lines. Unit cell dimensions are a = b = ~96 Å and γ = 120°.

(a) Grey-scale map with the white regions representing the regions of density above the mean.
(b) Contour map in *p*1.
(c) Contour map in *p*6

Two-sided plane group	Phase residual versus other spots (90° random)	Number of comparisons	Target residual based on statistics taking Friedel weight into account
*p*1	19.6	140	
*p*2	35.0’	70	28.4
*p*12_b	71.0	36	21.1
*p*12_a	78.7	35	20.9
*p*121_b	70.9	36	21.1
*p*121_a	73.3	35	20.9
*c*12_b	71.0	36	21.1
*c*12_a	78.7	35	20.9
*p*222	67.0	141	24.0
*p*2221b	54.5	141	24.0
*p*2221a	56.9	141	24.0
*p*22121	58.8	141	24.0
*c*222	67.0	141	24.0
*p*4	54.4	150	23.8
*p*422	67.8	329	29.8
*p*4212	61.8	329	21.5
*p*3	22.4!	112	19.7
*p*312	44.8	268	20.0
*p*321	44.4	271	20.2
*p*6	23.8!	294	21.7
*p*622	42.7	609	20.7

Table 5.2 Internal phase residuals of all possible two-sided plane groups calculated from image 5276

Internal phase residuals were determined from spots of IQ1 to IQ5 to 9 Å resolution. The value for p1 is a theoretical value based on the signal to noise ratio. Values marked with * are good candidates for the symmetry as the experimental phase residual is ≤ 1° than the expected value based on the signal to noise ratio. Values marked ! are within 5° and values marked ’ are within 10° (Valpuesta et al., 1994).

Two-sided plane group	Phase residual versus other spots (90° random)	Number of comparisons	Target residual based on statistics taking Friedel weight into account
*p*1	26.8	168	
*p*2	41.4!	84	39.3
*p*12_b	76.4	34	29.0
*p*12_a	68.8	34	29.0
*p*121_b	55.2	34	29.9
*p*121_a	67.5	34	29.0
*c*12_b	76.4	34	29.0
*c*12_a	68.8	34	29.0
*p*222	56.2	152	33.7
*p*2221b	65.4	152	33.7
*p*2221a	55.8	152	33.7
*p*22121	70.6	152	33.7
*c*222	56.2	152	33.7
*p*4	59.6	172	32.9
*p*422	68.2	348	29.8
*p*4212	61.8	348	29.8
*p*3	21.5*	122	26.8
*p*312	52.5	274	27.2
*p*321	52.5	281	27.5
*p*6	26.7*	328	30.0
*p*622	51.2	639	28.4

Table 5.3 Internal phase residuals of all possible two-sided plane groups calculated from image 5118
Internal phase residuals were determined from spots of IQ1 to IQ5 to 9 Å resolution. Symbols *, ! and ’ are as described in Table 5.2.

Two-sided plane group	Phase residual versus other spots (90° random)	Number of comparisons	Target residual based on statistics taking Friedel weight into account
*p*1	24.5	64	
*p*2	39.3!	32	35.7
*p*12_b	38.4	12	26.4
*p*12_a	63.0	13	27.1
*p*121_b	68.0	12	26.4
*p*121_a	45.2	13	27.1
*c*12_b	38.4	12	26.4
*c*12_a	63.0	13	27.1
*p*222	63.4	57	30.8
*p*2221b	43.9	57	30.8
*p*2221a	62.6	57	30.8
*p*22121	63.3	57	30.8
*c*222	63.4	57	30.8
*p*4	63.0	60	30.5
*p*422	58.5	123	27.4
*p*4212	57.7	123	27.4
*p*3	27.2!	50	24.5
*p*312	30.4’	110	25.3
*p*321	29.1!	110	25.3
*p*6	29.9!	132	27.2
*p*622	30.9!	252	25.9

Table 5.4 Internal phase residuals of all possible two-sided plane groups calculated from image 5109

Internal phase residuals were determined from spots of IQ1 to IQ5 to 15 Å resolution. Symbols *, ! and ’ are as described in Table 5.2.

Two-sided plane group	Phase residual versus other spots (90° random)	Number of comparisons	Target residual based on statistics taking Friedel weight into account
*p*1	26.3	92	
*p*2	41.2!	46	38.6
*p*12_b	71.1	18	29.1
*p*12_a	58.3	17	28.5
*p*121_b	66.9	18	29.1
*p*121_a	57.4	17	28.5
*c*12_b	71.1	18	29.1
*c*12_a	58.3	17	28.5
*p*222	71.4	81	33.3
*p*2221b	56.7	81	33.3
*p*2221a	52.9	81	33.3
*p*22121	58.5	81	33.3
*c*222	71.4	81	33.3
*p*4	57.8	86	33.3
*p*422	60.2	179	29.5
*p*4212	50.6	179	29.5
*p*3	18.3*	64	26.3
*p*312	18.5*	152	27.0
*p*321	21.9*	154	27.1
*p*6	26.4*	174	29.6
*p*622	23.8*	352	27.9

Table 5.5 Internal phase residuals of all possible two-sided plane groups calculated from image 5261

Internal phase residuals were determined from spots of IQ1 to IQ5 to 12 Å resolution.

Symbols *, ! and ’ are as described in Table 5.2.

5.2.4 *p*6 Crystal Form Merging

Data from images 5118, 5276 and 5109 were merged with phase residuals significantly below 45˚ up to a resolution of 9 Å (Table 5.6) with averaged unit cell dimensions between the three merged images of a = b = 95.0 Å ± 2.1 and γ = 120° ± 0.7. Random data would be expected to have a phase residual of 45˚ and the data in the 9 – 8 Å resolution range is within standard error of 45˚ (Table 5.6) so the resulting map (Figure 5.7) has been calculated with a resolution limit of 9 Å.

Resolution Range (Å)	No of reflections	Mean phase residual (°)	Standard Error of Mean (°)
200 - 15	20	7.0	2.4
15 - 12	8	25	8
12 - 10	14	23	4
10 - 9	10	26	7
9 - 8	13	46	8

Table 5.6 Phase Residuals in resolution shells for *p*6 averaged data
The phase residual presented is the difference between the averaged data phases and the symmetry enforced phases of 0 or 180˚ calculated using spots of all IQs at the given resolution range.

The unit cell of the projection map contains a hexagonal ring with density at the periphery. The size of the hexamer was ~104 Å measured from the edges of the peripheral density. The putative monomer suggested in Figure 5.7 due to a clear delineation from the other subunits has 5 near circular peaks that may represent a projection view of an alpha helix. In addition, there are two extended densities (labelled ‘a’ and ‘b’ in Figure 5.7) that may represent tilted or bent helices. In total, there are enough density features to account for the 7 predicted transmembrane helices.

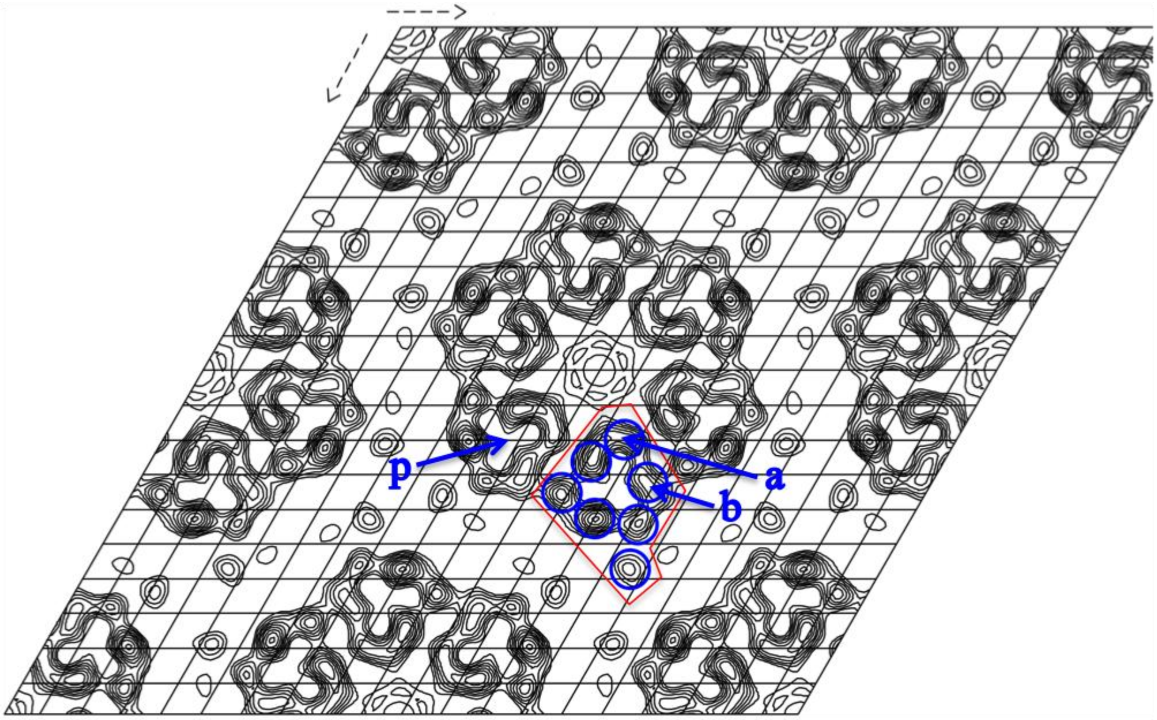


Figure 5.7 Contour map of 3 merged images of UAC$_{\boldsymbol{Bc}}$ crystals with *p*6 symmetry to 9 Å resolution (Huysman et al., 2012)

The contour lines show density above the mean density at intervals of ~0.4 x root mean square density. The unit cell dimensions are a = b = 95.0 ± 2.1 and γ = 120° ± 0.7. A putative monomer has been outlined in red. The locations of putative helices have been marked by blue circles of 10 Å diameter. Two of the density peaks within the putative monomer labelled ‘a’ and ‘b’ have an extended appearance which may be due to tilting or bending of the helix. The putative location of the pore is within the three lobed cavity labelled ‘p’.

5.2.5 *p*622 Crystal Form

The map produced from image 5261 is similar to the *p*6 form in that it is also made up of hexagonal rings (Figure 5.8). The *p*622 form unit cell dimensions are close to that of the *p*6 form with dimensions of a = b = 101.8 and γ = 120° (Table 5.1). The phase origin determined from ALLSPACE was the same for both *p*6 and *p*622 and producing a map with the lower symmetry of *p*6 imposed results in a similar map to the *p*622 symmetry map with pseudo mirror lines of symmetry visible (Figure 5.8).

5.2.6 Atomic Force Microscopy

2D crystals of UAC_{Bc} had to be dialysed into crystallisation buffer omitting glycerol as excess glycerol resulted in noisy AFM images. The crystals were applied to freshly cleaved mica and imaged under tapping mode in air as use of contact mode resulted in damage to the vesicles. As with electron cryomicroscopy of 2D crystals, flat objects with a size of at least 0.5 μm^2 were imaged with the AFM. From 20 measurements, the chosen objects for imaging had an average thickness of 84 Å ± 0.72 (Figure 5.9). This roughly corresponds to the thickness of two membranes where dried lipid bilayers were measured as ~45 Å thick and multiple bilayers were multiples of this value (Dols-Perez et al., 2011).

5.2.7 Three Dimensional Studies of Negatively Stained UAC_{Bc} crystals

Six of the largest and best-ordered crystals in negative stain were used to collect a series of images from tilted crystals using the microscope goniometer to nominally tilt up to 55° in both directions of the tilt axis. The images were processed as with the untilted images with the additional step of using EMTILT to determine tilt angles for each image and a tilt axis for each crystal. Correction for the CTF was unnecessary since diffraction spots did not extend beyond the first CTF minimum. ORIGTILTK was used to merge the images to a common phase origin (Table 5.7) and LATLINE was used to determine interpolated amplitudes and phases along lattice lines (Figure 5.10).

The resulting map shows that the hexamers of UAC_{Bc} all lie in a single plane with a thickness of ~54 Å (Figure 5.11). If the crystals were double layered, the reconstruction would be expected to show a pair of stacked rings or an unusually long ring depending on the resolution but there are no indications of density corresponding to another layer of hexamers in the 3D reconstruction. The empty regions in the surface representation represent the locations of greater staining and is consistent with the appearance of the protein in detergent (Chapter 4) where the disk like particles had a darker centre due to stain accumulation.

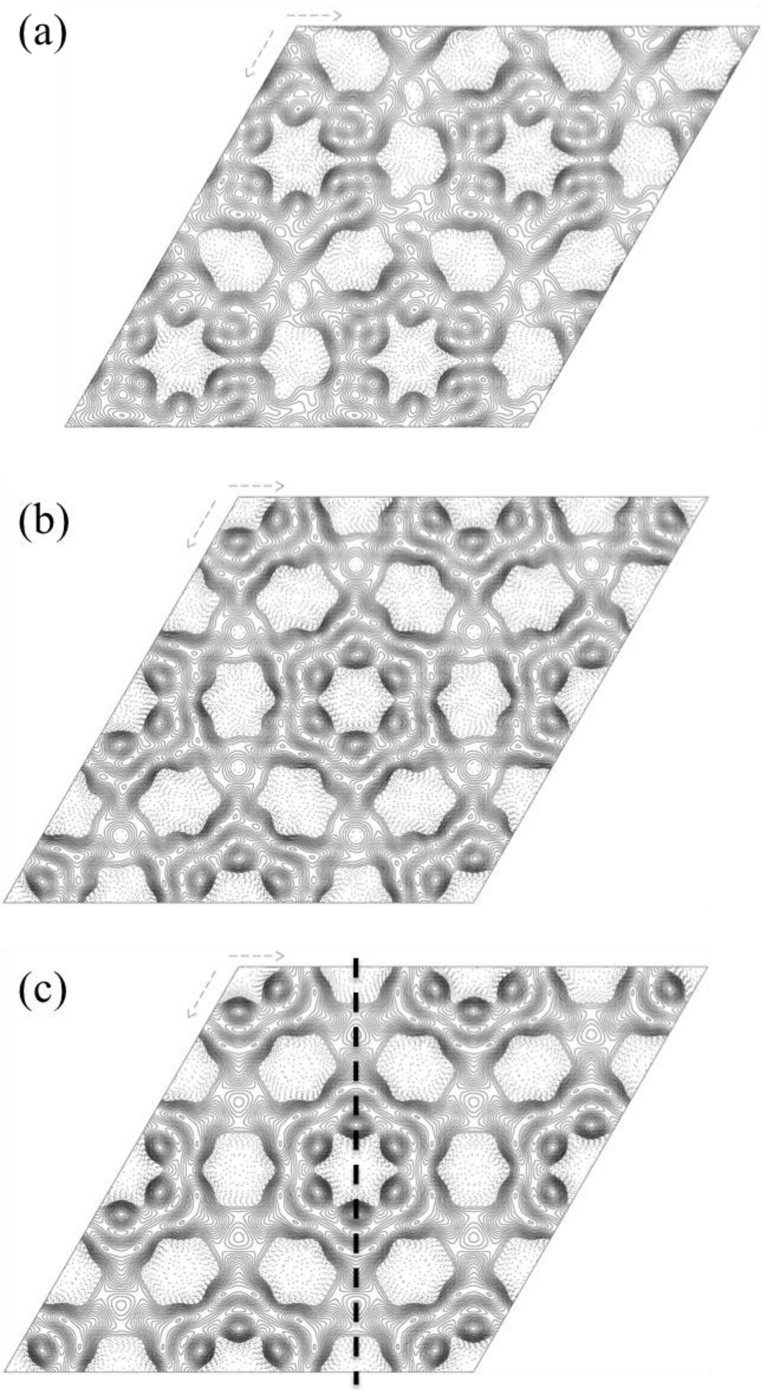


Figure 5.8 Projection maps produced from processing of image 5261 up to 12 Å. Contour intervals are drawn at ~0.15 x root mean square density with density above the mean in solid lines and density below in dotted lines The unit cell dimensions are a = b = 101.8 and γ = 120°.
(a) Contour map in *p*1 (no symmetry imposed)
(b) Contour map with *p*6 symmetry imposed
(c) Contour map with *p*622 symmetry imposed. The dashed line indicates a mirror line of symmetry.

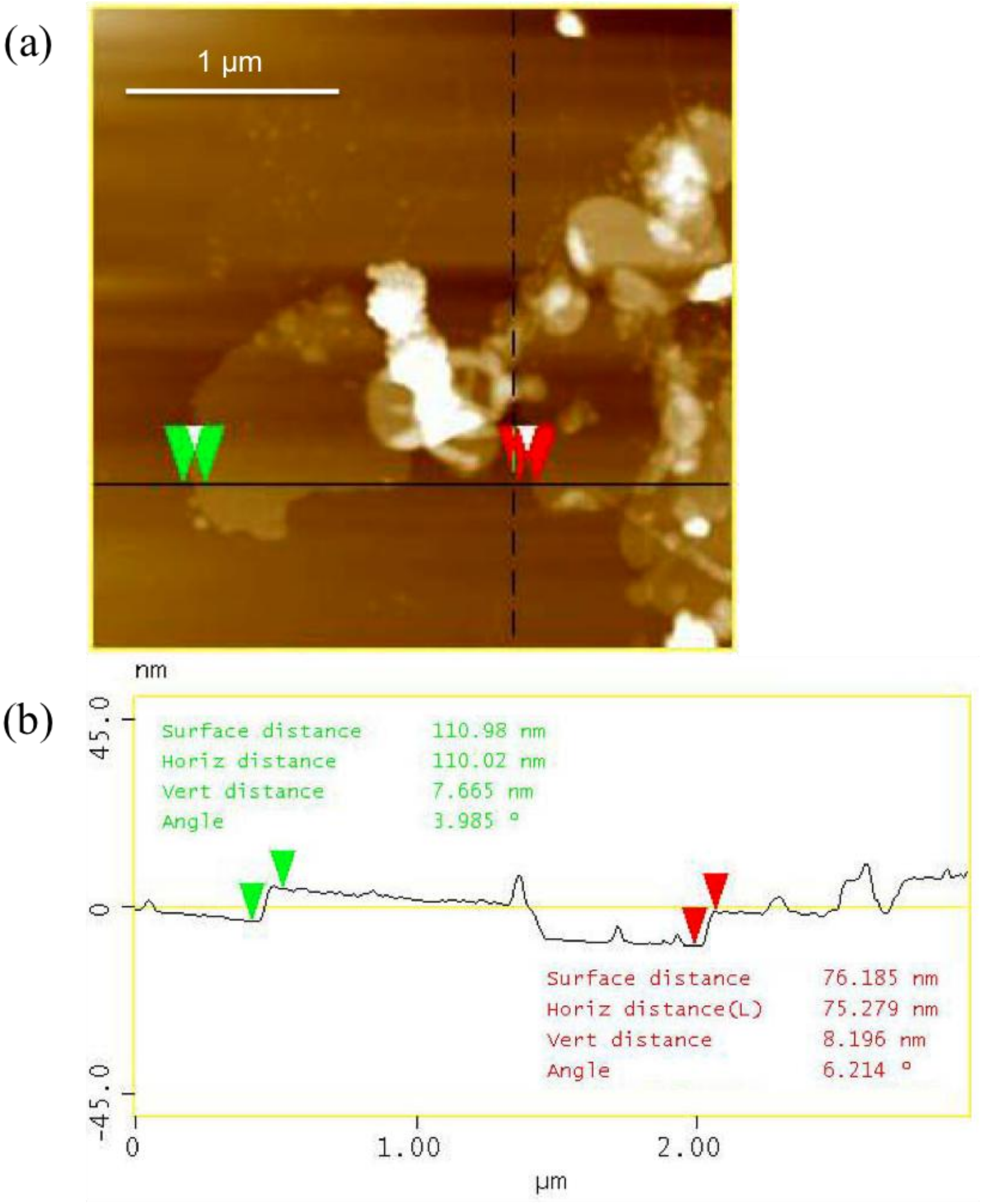
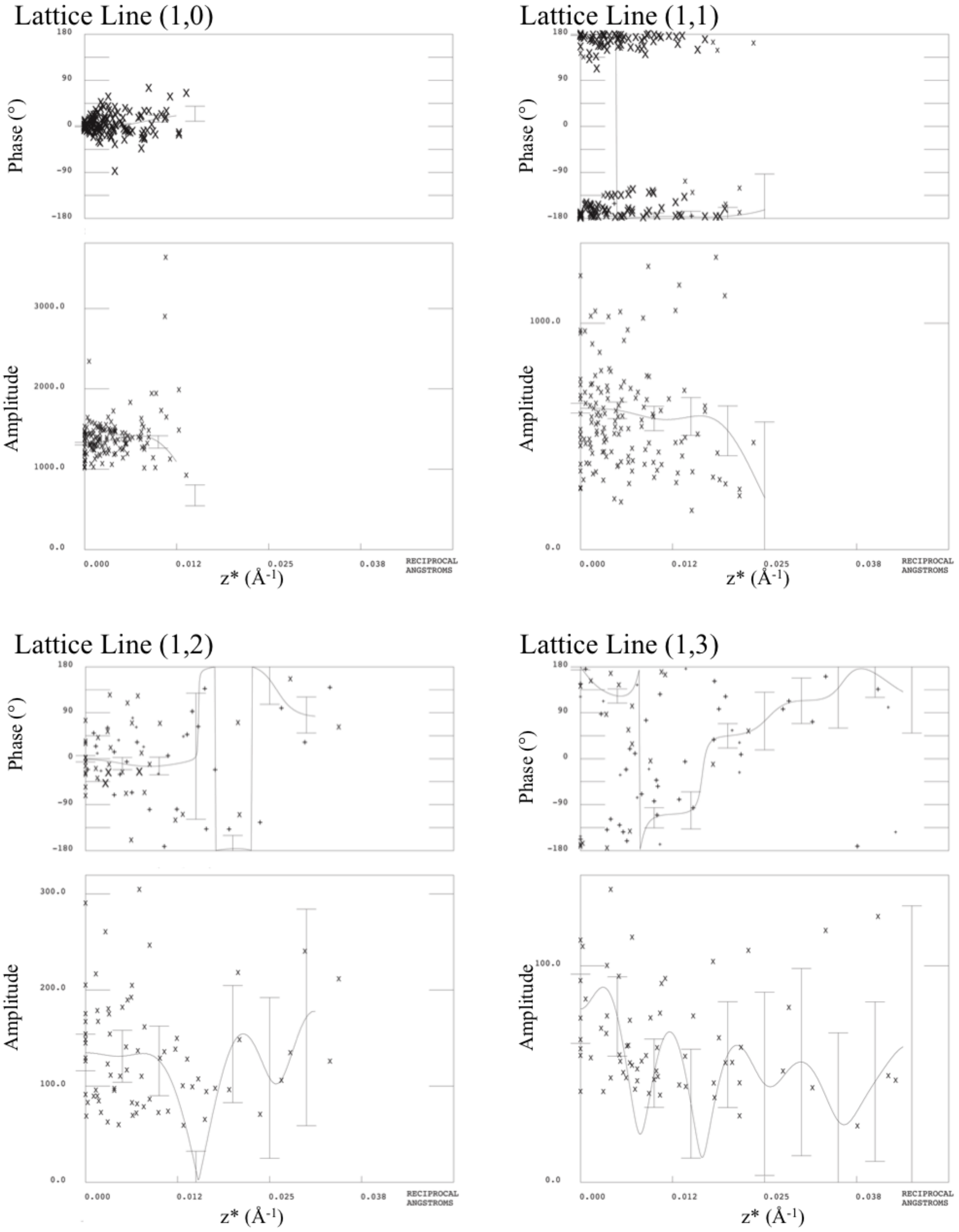


Figure 5.9 Thickness Measurements of UAC$_{Bc}$ 2D crystals using atomic force microscopy.

The height image from tapping mode AFM of UAC$_{Bc}$ crystals has been shown in (a). The vertical distance between the two arrows of the same colour was measured to determine the thickness of two crystals. The height profile of the black horizontal line is shown in (b). The thinnest objects had an average thickness of 8.4 nm which corresponds roughly to the thickness of two lipid bilayers.



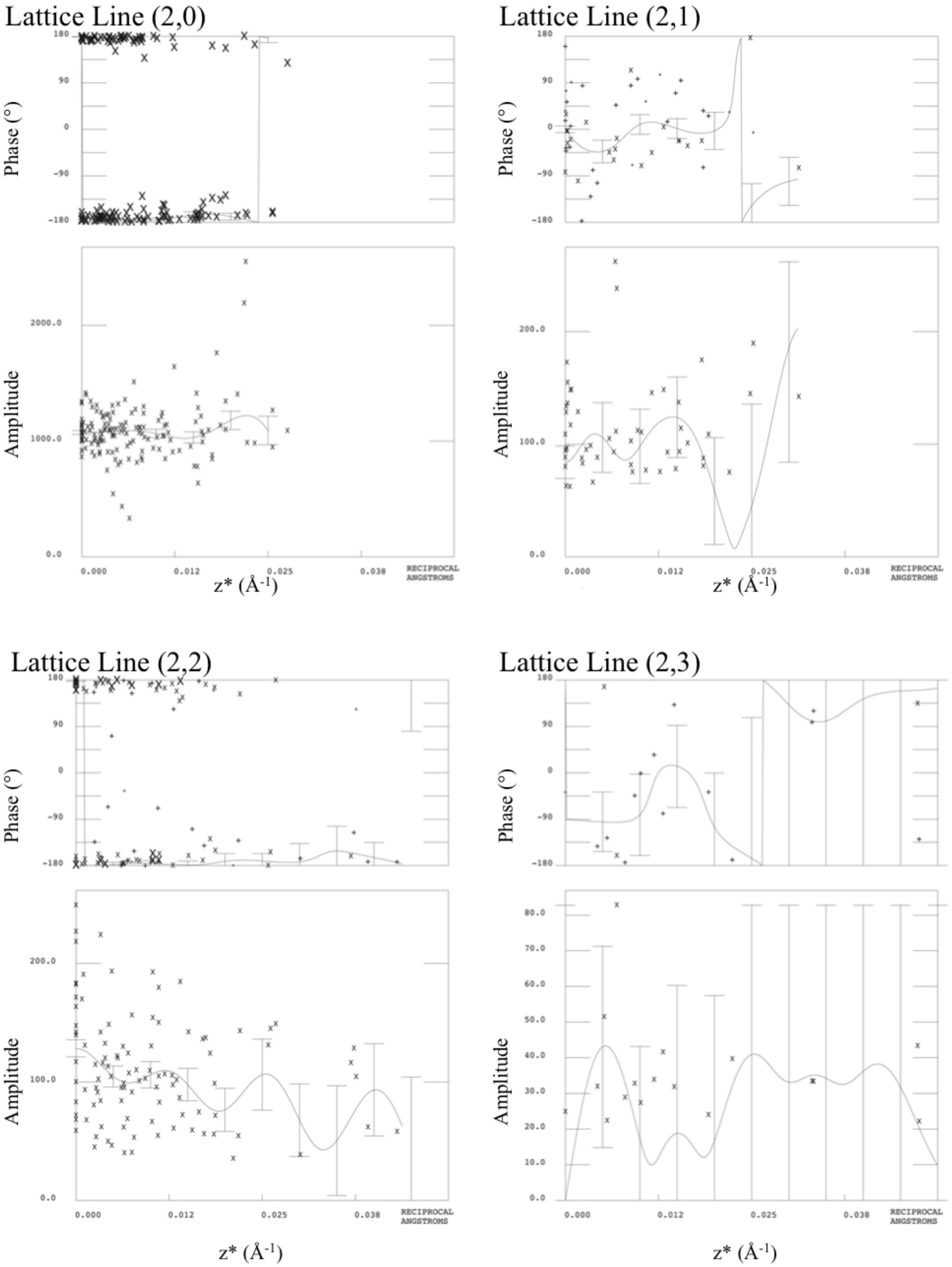
Lattice Line (1,0)

Lattice Line (1,1)

Lattice Line (1,2)

Lattice Line (1,3)

Figure 5.10 Interpolated lattice lines from the 2D crystals of UAC$_{Bc}$.
The upper graph from each set shows the variation in phase along z* in °.
The lower graph shows the variation in amplitudes along z* in arbitrary units. Errors bars represent standard deviation.


Lattice Line (2,0)

Lattice Line (2,1)

Lattice Line (2,2)

Lattice Line (2,3)

Figure 5.10 continued.

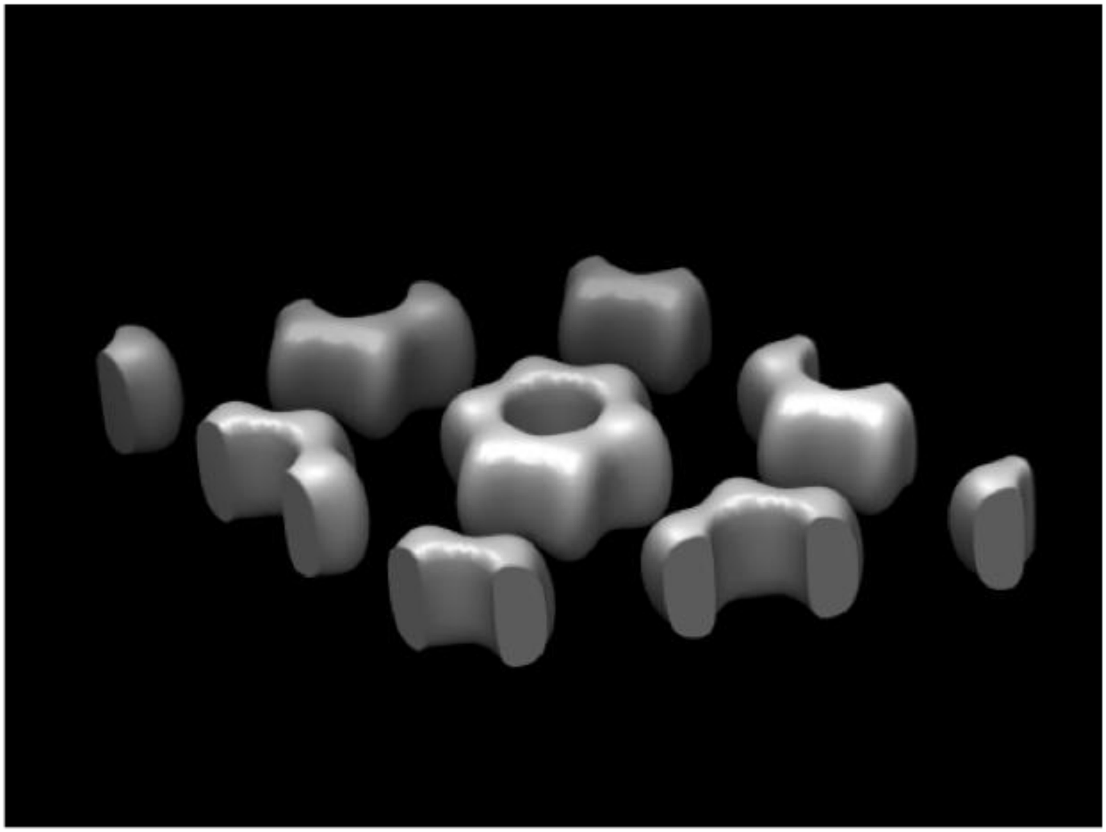


Figure 5.11 Surface representation of UAC_{Bc} Crystals in three dimensions
The surface representation map of UAC_{Bc} determined from processing of tilt series of 2D crystals. The map is rendered at an arbitrary threshold to show the density corresponding to the UAC_{Bc} hexamers.

Crystal	Nominal Tilt (°)	Refined Tilt Angle (°)	Phase Residual	Number of Spots
1	0	-	18.66	18
1	10	-5.13	16.6	18
1	-10	14.87	35.78	16
1	20	-15.13	23.1	16
1	-20	24.87	41.323	19
1	30	-25.13	26.78	14
1	-30	34.87	20.33	13
1	40	-35.13	32.28	11
1	-40	44.87	24.43	10
1	50	-45.13	33.38	8
2	0	-	17.23	24
2	10	-6.54	9.17	20
2	20	-16.54	23.19	21
2	30	-26.54	29	17
2	40	36.54	31.19	13
2	50	-46.54	28.82	14
2	55	-51.54	35.9	13
3	0	-	7.05	16
3	10	10.71	18.6	18
3	-10	-9.28	18.16	18
3	20	20.72	26.49	14
3	-20	-19.28	25.96	15
3	30	30.72	45.42	19
3	-30	-28.28	31.43	16
3	-40	-39.28	18.51	9
3	-50	-49.28	48.05	13
4	0	-	-	36
4	10	10.63	17.38	19
4	-10	-9.37	13.83	21
4	20	20.63	37.69	22
4	-20	-19.37	32.27	21
4	-30	-29.37	28.98	14
4	-40	-39.37	23.04	14
5	0	-	10.87	23
5	10	10.85	19.75	19
5	-10	-9.15	11.12	20
5	-20	-19.15	29.58	26
5	-30	-29.15	26.98	17
5	-40	-39.15	31.92	14
5	-45	-44.15	20.34	13
6	0	-	10.14	20
6	10	5.88	21.92	17
6	-10	-14.12	27.09	11

6	20	15.88	17.89	22
6	-20	-24.12	21.76	15
6	30	25.88	20.04	17
6	-30	-34.12	9.68	11
6	40	35.13	12.56	15
6	-40	-44.17	28.11	12
6	50	45.88	43.66	12

Table 5.7 Merging of Tilt Series Images

Tilt series images from 6 crystals were merged together using tilt angles and tilt axes derived from EMTILT. The phase residual was determined by comparison of an image against all others.

5.3 Discussion

5.3.1 Sample Preparation for Cryo-EM

The highest attainable resolution can depend on the preservation of the hydrated structure. Embedding the sample in glucose yielded images that showed diffraction spots at ~16 Å which is beyond those seen in negative stain indicating that the technique was successful. There remains scope for testing other means of sample preparation such as embedding in other sugars such as trehalose or freezing in vitreous ice in case other media may better preserve the hydrated structure of the crystals. There is evidence to show that use of the correct sample preservative can result in higher success rates for recording data. For example, the probability of recording high resolution diffraction patterns from the LHC II 2D crystals was 90% with tannin embedding but a factor of 10 to 20 lower with ice or glucose (Wang and Kühlbrandt, 1991). For the UAC$_{Bc}$ crystals where crystal quality varied greatly within a single sample (Chapter 4), improvements in the numbers of well ordered crystals would be necessary before any significant differences between sample preparation techniques would be noticed.

5.3.2 Data Collection

Radiation damage can cause structural changes in the sample leading to specimen movements. The spot scan technique where the beam is focused onto a smaller area and scanned over the sample is thought to reduce such specimen movements as exposure of a smaller region reduces the overall stress on the sample (Bullough and Henderson, 1987, Downing and Glaeser, 1986). The technique was reported to yield 30% of images with good diffraction in all directions compared to 15% recorded using the standard flood-beam technique (Bullough and Tulloch, 1991). This could be a useful technique for recording images of UAC$_{Bc}$ crystals as out of the 11 images with diffraction spots recorded with a flood-beam, 4 images (5108, 5113, 5124 and 5238) had strongly anisotropic resolution due to movement during image recording.

5.3.3 Plane Group of the Crystals

In determining the 9 Å projection map (Figure 5.7), the imposition of *p*6 symmetry is justified in that the ALLSPACE analysis gave plane groups consistent with *p*6 (Tables 5.2, 5.3, 5.4). Whilst a chiral object such as a protein should not have lines of mirror symmetry, it possible for these features to appear in a projection map of a protein crystal. In the case of UAC_{Bc}, a crystal with *p*622 symmetry could arise as double-layered crystals with pairs of hexamers in the two lipid bilayers stacked end to end. UAC_{Bc} was observed to form double layered particles in detergent solution (Chapter 4) and it was possible that these stayed intact during the crystallisation process to form double layered *p*622 crystals. It is also worth noting that recording an image of a vesicle type crystal can result in a Fourier transform with two reciprocal lattices corresponding to the crystal lattice on both sides of the collapsed vesicle. The two lattices can be indexed and processed separately as though each were a separate single layered crystal; these are distinct from 2D membrane crystals with mirror or screw axes such as the aquaporins (Gonen et al., 2005, Hiroaki et al., 2006) where a single layer of the crystal must consist of two bilayers. The processed images of UAC_{Bc} appeared to have a single lattice (Figure 5.2).

It is possible that the detection of *p*622 symmetry in images 5108 and 5261 by ALLSPACE was an artifact resulting from the lower resolution and having fewer spots for phase comparison. This appears to be the case for image 5108 as the image could be merged with the *p*6 images 5118 and 5276 without significantly altering the appearance of the map.

To test whether the crystals were multilayered, AFM and 3D negative stain reconstruction were used. A lattice was not observed on the images of the vesicles with AFM suggesting there are no significant differences in the height of the surface of the crystal due to extramembraneous parts of the protein as expected from topology predictions. As known from early attempts at recording diffraction from dried protein crystals (Taylor and Glaeser, 1976), drying the sample would also be expected to cause a loss of crystallinity. Imaging the largest and flattest objects revealed that they had a thickness of 70-80 Å corresponding approximately to two bilayers. This is consistent with the existence of a *p*622 crystal form, but for the *p*6 crystal form it would mean that either only one layer of the vesicle had a crystal lattice or that crystal lattices of both layers were in register. There is the possibility that the crystal lattice in contact with the carbon film may have been distorted or destroyed.

The fact that the vesicles have a consistent thickness under AFM also suggests that the crystals have a consistent unit cell thickness so the crystals produced could be used to collect and merge a 3D data set. The data along lattice lines varies more rapidly when a specimen is thicker so the unit cell thickness of the crystals must be known to merge data from crystals of varying thickness (Amos et al., 1982).

The 3D negative stain reconstruction produced a surface representation of the UAC$_{Bc}$ crystals. Whilst the resolution limit of 20 Å was too low to observe the structural features present in the 9 Å projection map (Figure 5.7), the 3D map was useful for confirming the hexamers lie in a single plane which could not be determined with an untilted projection alone. The same technique was applied during the study of the human copper transporter, CTR1, where there was ambiguity concerning whether to impose *p*6 or *p*622 symmetry. The 3D reconstruction revealed the two layers related by a twofold axis of rotation confirming the *p*622 symmetry (Aller and Unger, 2006).

Another explanation for the observation of an apparent higher plane group than is actually present is due to twinning of crystals. In an early study of the *Sulfolobus acidocaldarius* surface layer, *p*6 symmetry was mistakenly imposed on the crystals instead of the correct *p*3 symmetry. This occurred due to averaging over separate *p*3 domains rotated by 60° with respect to each other linked by a twin boundary to create a map with two orientations of the crystal superimposed (Lembcke et al., 1991). Twinned crystals were also formed during the fusion of bacteriorhodopsin crystals resulting in crystals where there were domains of molecules facing “up” and other facing “down” (Baldwin and Henderson, 1984). During reconstitution into the lipid bilayer, UAC$_{Bc}$ could take both orientations resulting in some twinned crystals with the appearance of *p*622 symmetry if multiple domains are processed together as a single crystal.

There are membrane proteins where a proposed ancient gene duplication event results in a repeated sequence and the protein structures show monomers formed by two homologous halves as has been found in the urea transporter structures determined by X-ray diffraction (Levin et al., 2009). Viewed as a projection at low resolution, such a monomer would appear to have two-fold pseudo symmetry as seen in the projection of aquaporin-1 (Murata et al., 2000). However, there is no evidence for gene duplication in the UAC_{Bc} sequence based on sequence alignment analysis (Chapter 6) and the merged *p*6 projection map (Figure 5.7) appears to have a distinct handedness.

Overall, the results from AFM suggest that it is possible for a *p*622 crystal form to exist which makes it uncertain which of the poorer images to include in the merging of data. Ultimately, it was found that images 5118, 5276 and 5109 could be merged to produce a map with phase residuals below those expected for random data up to a resolution of 9 Å (Table 5.6) but the other images merged poorly giving overly high phase residuals. Even so, the data redundancy from six fold symmetry allows the merging of data from the three images to produce a reliable projection map. In principal, the images with significant drift may have the same *p*6 crystal form as those of 5276 and 5118 and contain data that may be merged with other images but in the process of producing the merged map (Figure 5.7), they were omitted. The extra symmetry relationships and the resulting data redundancy with the *p*622 crystal form make that crystal form preferable to *p*6 for data collection but in the data recorded these crystals are more rare. Also, *p*622 is only suggested as a plane group possibility for crystals that diffracted to a lower resolution making low resolution the most likely explanation for *p*622 symmetry.

5.3.4 Interpretation of the *p*6 map

The resolution of the *p*6 merged projection map appears to be sufficient to distinguish the 7 predicted transmembrane helices. In projection at a resolution of 10-7 Å, transmembrane helices can appear as near circular peaks separated by ~10 Å but the peaks may merge together into a continuous density if the helices bend or tilt and their densities superimpose (Henderson and Unwin, 1975). Additionally, projection maps at higher resolutions are more difficult to interpret as secondary structure features are obscured by higher resolution structural information as demonstrated by the 4.5 Å and the truncated 7.5 Å projection maps of leukotriene C_4 synthase (Schmidt-Krey et al., 2004).

The hexamers were confirmed to lie in a single plane from processing the tilt series data allowing the assignment of helices to the density features in the map which would not be possible if the unit cell was a projection of two superimposed molecules out of register as with the *p*622 crystal form projection map of the human copper transporter where the densities of the trimers from the two layers superimpose (Aller and Unger, 2006).

The putative monomer (indicated in Figure 5.7) has 6 helices enclosing a three lobed hole. At 9 Å resolution, it is not possible to resolve side chains and therefore see how far electron density extends into the space but it is a better candidate for the location of the pore rather than the larger space within the ring of monomers which must be filled with lipid.

Crystal contacts between hexamers appear to be formed by the weak helix at the periphery of the monomer. The feature must be real as it is visible in the *p*1 maps (Figures 5.4 and 5.5) but is relatively weak in the merged map with a even weaker adjacent feature (Figure 5.7) suggesting some flexibility in the linkage to the rest of the monomer or another conformation of the helix with low occupancy. The flexibility may also allow some imperfections in the hexamer crystal packing pack together reducing the ordering and obtainable resolution.

5.3.5 Conclusions

The mixture of UAC_{Bc} 2D crystals varies in both size and crystal order. From image processing, it appears that the *p*6 crystal form predominates in the UAC_{Bc} 2D crystal sample. The resulting merged projection map at 9 Å resolution has a hexagonal density in the unit cell and it may be possible to ascribe the density peaks of the map to the expected alpha helices. The clearly delineated cluster of density peaks in within the hexamer provides the most likely monomer and there are six density peaks enclosing a putative location for the substrate pore.

Due to the severe aggregation of 2D crystals and the possibility of a double layered *p*622 crystal form, it became necessary to determine the unit cell thickness of the crystals as variations in crystal thickness could have precluded merging of the data and production of a 3D map that would provide further information to support the secondary structure model. Atomic force microscopy of the sonicated crystal sample revealed that the ~1 μm^2 flat objects selected for EM imaging actually corresponded to collapsed vesicles and thus had a thickness of 2 membranes. The data from images of crystals stained with uranyl formate produced a map with a single layer of hexamers. The single layered map suggests that the density peaks in the 9 Å map can be interpreted as single alpha helices in projection.

The consistency in crystal thickness also suggests that data collected from the sample could be merged together. Glucose appears to a suitable medium for embedding the crystals and preserving the hydrated structure of the protein but further work could include the testing of other embedding media. Also, efforts at producing higher yields of well ordered crystals would greatly facilitate efforts at producing a higher resolution 3D map.

Chapter 6 - Biochemical Analysis and Modelling

6.1 Introduction

In the absence of high resolution structures of membrane proteins structures, other methods may be employed to predict and test structural details and these can be useful in the interpretation of data derived from electron microscopy. A preliminary model of UAC_{Bc} will be described using the combination of structure predictions, the projection data (Chapter 5) and biochemical data.

Alpha helical membrane proteins have a number of characteristic features that allow prediction of their secondary structures and topologies from their sequences. Transmembrane alpha helices consist mostly of a relatively long stretch of hydrophobic amino acids, so early predictions of transmembrane helices involved the averaging of the hydropathy across predefined segments of the sequence (Kyte and Doolittle, 1982). Additionally, cytoplasmic facing loops connecting the transmembrane helices in inner membrane proteins have a higher incidence of the positively charged residues arginine and lysine compared to the periplasmic facing loops (Heijne, 1986). States can then be defined as being either within the membrane, in a periplasmic or cytoplasmic loop and experimentally determined amino acid probability distributions provide parameters that form the basis of Hidden Markov models (HMM) for transmembrane helix prediction (Sonnhammer et al., 1998). Discrepancies in the predicted start and end positions of helices can be expected between different prediction algorithms but there should be agreement between the number and rough location of the helices. For UAC_{Bc}, TopCONS has been used which enters the chosen sequence into the OCTOPUS (Viklund and Elofsson, 2008), PRO-TMHMM and PRODIV-TMHMM (Viklund and Elofsson, 2004), SCAMPI-single and SCAMPI-multi (Bernsel et al., 2008) topology predictors and combines the result into a single consensus prediction with a reliability score.

Analysis of high resolution structures reveals further patterns between the residue distributions in the structures of membrane proteins. Residues on lipid facing surfaces of transmembrane helices and to a lesser degree oligomer interfaces are more variable than residues buried away from the lipid (Yeates et al., 1987) (Wallin et al., 1997).

Sequence alignment provided useful constraints for modeling the arrangement of the helices in G-protein coupled receptors (Baldwin, 1993) using the density peaks in a 9 Å projection map of bovine rhodopsin as a guide (Schertler et al., 1993). The constraints included the positioning of the helices next to their neighbours in sequence due to the short lengths of the connecting loops and the exposure of the helices to the lipid from residue conservation. The model produced before high resolution 3D data was available was later found to be accurate in terms of the assignments of the helices when the first high resolution GPCR structure was published (Palczewski et al., 2000). In a similar vein, following from the work described in Chapter 5 examining the nature of the 2D crystals of UAC_{Bc}, constraints from bioinformatic analysis could then be used to model the arrangement of the helices using the projection structure. This chapter also reports on the testing of the predicted topology by labelling positions in single cysteine mutants predicted to be exposed to either the periplasm or cytoplasm (Ye et al., 2001).

Furthermore, we also tested whether the hexameric structure of UAC_{Bc} was an artifact of crystallization by monitoring the protein in detergent solution using size exclusion chromatography multi angle light scattering (SEC-MALLS) and cross-linking with glutaraldehyde to determine the oligomeric state.

6.2 Materials and Methods

6.2.1 Purification of UAC_{Bc} with a TEV protease cleavable His-tag

Purification of UAC_{Bc} with a tobacco etch virus (TEV) protease cleavable His-Tag was performed by Dr. Gerard Huysman at the University of Leeds.

Buffer A contains 50 mM Hepes (pH 7.4), 100 mM KCl and 5% glycerol (w/v).

UAC_{Bc} with a C-terminal TEV protease cleavable His_8 tag was used for experiments performed at University of Leeds to avoid contamination of UAC_{Bc} with the multi-drug efflux transporter AcrB which can co-purify with His-tagged membrane proteins and can form X-ray diffraction quality crystals in picomolar quantities (Psakis et al., 2009, Glover et al., 2011). Expression and production of membranes were as described in Chapter 3 Materials and Methods.

UAC_{Bc} was solubilised in buffer A containing 1.5% DDM and 5 mM imidazole. The mixture was centrifuged and the supernatant was incubated with His-Pur Co^{2+} resin (Pierce) overnight. The resin was washed with buffer A containing 100 mM imidazole and 0.05% DDM before eluting with buffer A containing 500 mM imidazole and 0.05% DDM. Equimolar amounts of UAC_{Bc} and His-tagged TEV protease were mixed together and dialysed against buffer A containing 0.05% DDM. The mixture was passed through Ni-NTA to remove the TEV protease and the His_8 tags cleaved from UAC_{Bc}.

6.2.2 Cross Linking

Cross linking was performed by Dr. Gerard Huysman at the University of Leeds.

Purified UAC_{Bc} in buffer A with 0.05% DDM was crosslinked with 0-100 mM glutaraldehyde for 30 minutes at 25 ºC. Controls included a sample with no addition of glutaraldehyde and a sample with addition of 1% SDS prior to cross linking with 100 mM glutaraldehyde. All samples were analysed by SDS-PAGE and stained with Coomassie Brilliant Blue.

6.2.3 Size Exclusion Chromatography-Multi-Angle Laser Light Scattering

SEC-MALLS was performed by Matthew Jennions and Dr. Isabel Moraes at the Diamond Membrane Protein Laboratory.

SEC-MALLS was performed by running UAC_{Bc} on a Superdex 10/300 column (GE Healthcare) in buffer A with 0.05% DDM using a Viscotek Tetra Detector system (Malvern Instruments) to record A_{280}, right angled light scattering and refractive index values. These were input into OmniSEC software (Malvern Instruments) to determine the mass of protein in the protein:detergent complex ($M_{w,protein}$) according to the following equation:

$$M_{w,protein} = \Delta LS * A_{280} / K * A_{280,protein} (\Delta RI)^2$$

In the above K is a constant dependent on the refractive index of buffer, the wavelength of light used, the angle between the incident and scattered light and the distance between the scattering molecule and detector. ΔLS is the excess of light scattered by the buffer containing the protein:detergent complex compared to light scattered by the buffer alone. $A_{280,protein}$ is the extinction coefficient of UAC_{Bc} and ΔRI is the refractive index difference between the buffer and the buffer containing the protein:detergent complex in solution (Slotboom et al., 2008).

6.2.4 Site-Directed Fluorescence Labelling

Site-directed fluorescence labelling was performed by Dr. Gerard Huysman at the University of Leeds.

For site directed fluorescence labelling, the QuikChange (Stratagene Method) was used to produce a cysteine free mutant of UAC_{Bc} with the mutations C91S, C187A and C189A. A series of single cysteine UAC_{Bc} mutants with positions corresponding to loops connecting transmembrane helices exposed to either the cytoplasm or periplasm was then created (Table 6.1). An additional mutant with a cysteine directly following the TEV-cleavage recognition site on the C-terminal end was produced designated as C_{ins}.

Mutant	Predicted Cysteine Location
G2C	Periplasm
M26C	Cytoplasm
S54C	Periplasm
G81C	Cytoplasm
Q108C	Periplasm

N139C	Cytoplasm
E168C	Periplasm
C_{ins}	Cytoplasm

Table 6.1 Single cysteine mutants for site directed fluorescence labeling.

Each mutant was transformed into *E. coli* BL21-gold cells which were cultured and induced with 1mM IPTG for 3 hours and the cells were divided into three aliquots for separate treatments. The first group (external reactivity) was treated with the 40 µM Oregon Green Maleimide for 20 minutes followed by quenching with 2mM-mercaptoethanol. The cells were washed and membranes were prepared by sonication and centrifugation. The second group (internal reactivity) had their potential periplasmic cysteines blocked by reaction with [2-(trimethylammonium)ethyl] methanethiosulfonate bromide (MTSET). Membranes were prepared and then treated with OGM. Membranes were prepared for the third group of cells (cysteine availability) and were reacted with OGM. UAC_{Bc} from each of the three groups were purified as described above and analysed by SDS-PAGE on a 12 % gel for OGM fluorescence and stained with Coomassie Brilliant Blue.

6.2.5 Bioinformatic Analysis

Multiple Sequence Alignment was performed by Prof. Stephen Baldwin at the University of Leeds.

64 sequences of UAC family members with ≤90 % identity with each other were obtained by performing a BlastP search of the UniProt protein sequence database and these were aligned with ClustalX (Thompson et al., 1997) (Table 6.2). TOPCONS was used to predict the number and location of transmembrane helices (Bernsel et al., 2009). ConSeq was used to assign residues with a numerical value of 1 to 9 based on their conservation and are represented by a colour scale in Figure 6.4 (Berezin et al., 2004). The frequency of polarity in the aligned sequences was also determined. In Figure 6.4, positions occupied by polar residues in >20% of sequences are red and positions occupied by polar residues in ≤20 % sequences are yellow. Positions where

serine and threonine are the only polar residues in >20 % of sequences are green and in ≤20% are turquoise. Positions never occupied by a polar residue are coloured blue.

No.	Uniprot (or RefSeq*) Accession	Species	No.	Uniprot (or RefSeq*) Accession	Species
1	C3ATS4	Bacillus mycoides	33	A6SZ88	Janthinobacterium sp. Marseille
2	Q814I5	Bacillus cereus ATCC 14579	34	A4G534	Herminiimonas arsenicoxydans
3	D5DGD0	Bacillus megaterium	35	C4ZM12	Thauera sp.
4	C4L1B5	*Exiguobacterium* sp. ATCC BAA-1283	36	B9Z5Z0	Lutiella nitroferrum 2002
5	B1YM53	Exiguobacterium sp. 255-15	37	D1CAT7	Sphaerobacter thermophilus
6	Q733K1	Bacillus cereus ATCC 10987	38	A5W2C4	Pseudomonas putida F1
7	A4FB17	Saccharopolyspora erythraea NRRL 2338	39	Q51417	Pseudomonas aeruginosa PAO1
8	D2B143	Streptosporangium roseum	40	F5SNT4	*Psychrobacter* sp. 1501(2011)
9	D6A7X3	Streptomyces ghanaensis	41	F7SMU4	*Halomonas* sp. TD01
10	Q82LR7	Streptomyces avermitilis	42	F7SIB6	*Halomonas* sp. TD01
11	ZP_08120863*	Pseudonocardia sp. P1	43	G0ACH3	Collimonas fungivorans (strain Ter331)
12	F4CRF6	Pseudonocardia dioxanivorans CB1190	44	Q395A4	Burkholderia sp. 383
13	D2SFT7	Geodermatophilus obscurus	45	B2Q3D6	Providencia stuartii ATCC 25827
14	P56583	Mycobacterium smegmatis	46	F0DNU9	Desulfotomaculum nigrificans DSM 574
15	A0K0R1	Arthrobacter sp. FB24	47	F3Y897	Melissococcus plutonius ATCC 35311
16	B8H6K8	Arthrobacter chlorophenolicus	48	C9A868	Enterococcus casseliflavus
17	F0M6V9	Arthrobacter phenanthrenivorans DSM 18606	49	C3X7K8	Oxalobacter formigenes
18	D9VAW5	Streptomyces sp. AA4	50	C6JBX7	Ruminococcus sp.
19	F6ES78	Amycolicicoccus subflavus DSM 45089	51	F7V7A1	*Clostridium* sp. (strain SY8519)
20	Q5Z1T9	Nocardia farcinica	52	Q55052	Streptococcus salivarius
21	D5PNR0	Rhodococcus equi	53	B1SGK7	Streptococcus infantarius subsp. infantarius ATCC BAA-102
22	Q53185	Rhodococcus erythropolis	54	ZP_07903706*	Eubacterium saburreum
23	D2KYA8	Corynebacterium	55	F7QZH7	Lactobacillus ruminis

		glutamicum			SPM0211
24	Q8FUJ4	Corynebacterium efficiens YS-314	56	B3XPZ5	Lactobacillus reuteri 100-23
25	A1AYT4	Paracoccus denitrificans	57	B1BU46	Clostridium perfringens E str. JGS1987
26	A1B7Z4	Paracoccus denitrificans	58	Q93PJ3	Helicobacter hepaticus
27	B6BFQ8	Rhodobacterales bacterium Y4I	59	C3XHC2	Helicobacter bilis
28	B6JIF0	Oligotropha carboxidovorans OM5	60	D3UGE8	Helicobacter mustelae
29	D7A7P5	Starkeya novella	61	P56874	Helicobacter pylori J99
30	A9BQZ9	Delftia acidovorans SPH-1	62	E7G0A0	Helicobacter suis HS1
31	B7X054	Comamonas testosteroni KF-1	63	A6MEX6	Helicobacter felis
32	Q7WB21	Bordetella parapertussis	64	Q8GH96	Helicobacter bizzozeronii

Table 6.2 UAC family members used in bioinformatic analysis

6.3 Results

6.3.1 Cross-Linking

Cross-linking of UAC$_{Bc}$ with glutaraldehyde produced a series of bands on SDS-PAGE with increasing molecular weight corresponding to oligomers up to hexamers (Figure 6.1). At concentrations of 25 mM glutaraldehyde and above, hexamers of UAC$_{Bc}$ predominate (Huysmans et al., 2012).

6.3.2 SEC-MALLS Analysis

UAC$_{Bc}$ eluted as a major peak (Figure 6.2 labelled B) with an earlier peak (Figure 6.2 labelled A) and a weak peak (Figure 6.2 labelled C). DDM micelles without protein eluted giving a peak in the refractive index (Figure 6.2 labelled D) but not in A_{280} as DDM does not absorb at 280 nm. The masses of the proteins in the protein:detergent micelles in peaks A and B were calculated as 448 kDa and 143 kDa respectively.

UAC$_{Bc}$ has a predicted mass of 24030 Da so peaks A and B correspond to 18.6 and 6 monomers respectively (Huysmans et al., 2012).

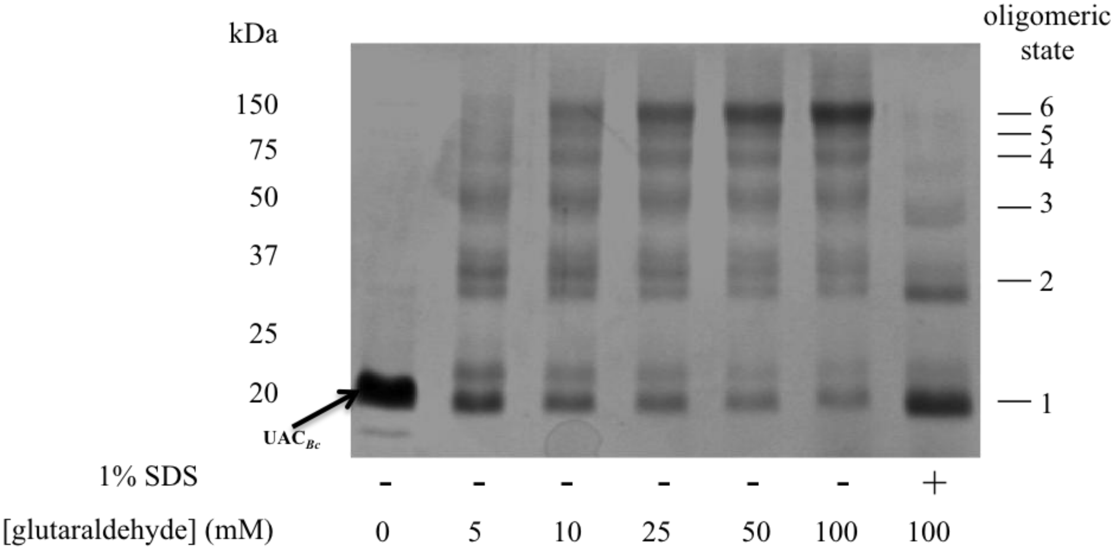


Figure 6.1 Cross-Linking Analysis of UAC$_{Bc}$
(Huysman et al., 2012).

Purified UAC$_{Bc}$ in DDM was cross-linked in increasing concentrations of glutaldehyde and analysed by SDS-PAGE on a 12% gel. In the absence of glutaraldehyde, UAC$_{Bc}$ run as a single band and when cross-linked, formed oligomers up to hexamers. Cross-linking was limited if performed in the presence of 1 % SDS suggesting that the hexamer was formed by the native folded protein.

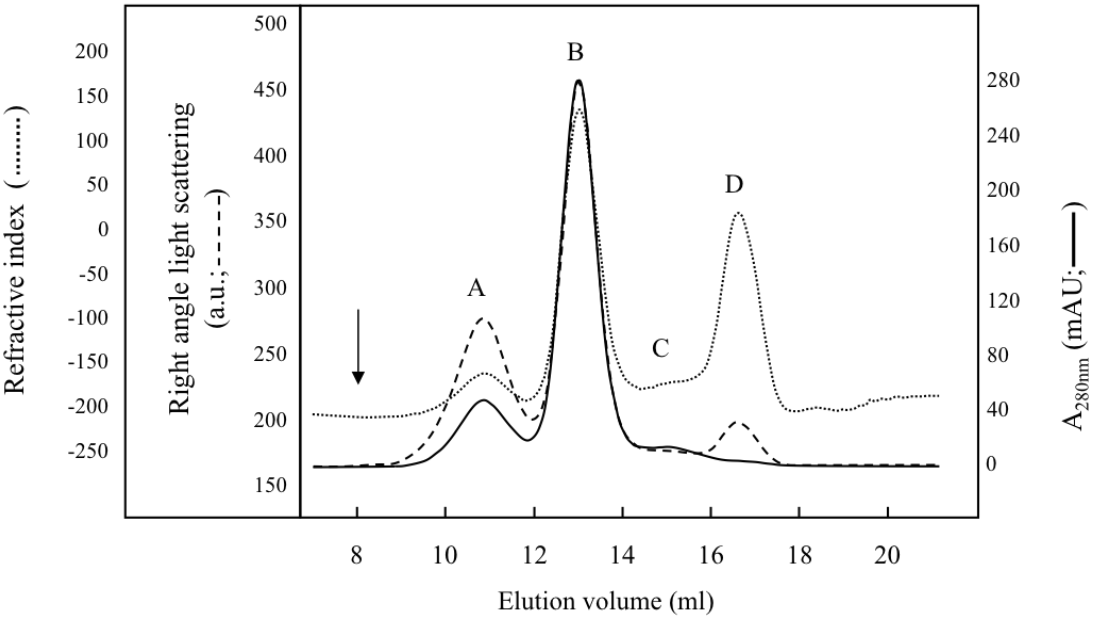


Figure 6.2 SEC-MALLS Analysis of UAC$_{Bc}$ in Detergent
(Huysman et al., 2012).

A$_{280}$ (solid line), right angle light scattering (dashed line) and the refractive index (dotted line) were recorded when UAC$_{Bc}$ was applied to a Superdex 200 10/300 column. The arrow indicates the void volume of the column.

The protein mass in peak B was calculated as 143 kDa corresponding to a hexamer of UAC$_{Bc}$ which has a monomer mass of 24 kDa.

6.3.3 Site Directed Fluorescence Labelling

A cysteine would be assigned a periplasmic location if it could be labelled with the fluorescent membrane-impermeable molecule OGM in intact cells and an internal position if the cysteine would only react with OGM after cell disruption. As a positive control, each of the single cysteine mutants could be labelled with OGM upon cell disruption (Figure 6.3, CA). The mutants bearing predicted cytoplasmic cysteines indeed only reacted with OGM after cell disruption (Figure 6.3, IR and CA).

Cysteines in the S54C and E168C mutants could be assigned a periplasmic location as they could be labelled with OGM in the intact cells. However, the expected periplasmic cysteines in the G2C and Q108C mutants were not labelled in the intact cells. Evidence for a periplasmic location came from the fact that these sites were not labelled with OGM if the cells were pretreated with MTSET prior to cell disruption. The inability of OGM to react with the cysteines may have been caused by steric hindrance of the larger OGM molecular relative to MTSET (Huysmans et al., 2012). The results support the predicted 7 transmembrane helix topology.

6.3.4 Assignment of Helices

The multiple sequence alignment provided some constraints used to assign helices to the density features in the 9 Å map and also provided information on their likely orientations with respect to each other. Sequence alignment shows that the 7th predicted transmembrane helix is poorly conserved (Figure 6.4) and the helical wheel plots (Figure 6.5) show that the surface is mainly lined with hydrophobic residues. This was interpreted as the helix most likely to be exposed to the surrounding lipid so helix 7 was assigned to the peripheral density feature (Figure 6.5). Additionally, the first six helices bear more conserved residues likely to be involved in either pore formation or inter subunit interactions and the ring of six density features within a subunit are better candidates for the first six transmembrane helices.

The consensus from topology predictions is that the loops connecting the helices are relatively short which implies that the helices are likely to be packed adjacent to their neighbours in sequence. The nearest density to helix 7 would then be likely to correspond to helix 6 and the helix is oriented as to bury the positions 146 and 149 occupied by polar residues in most sequences.

The remaining helices 1-5 may then dial clockwise (Figure 6.5) or anticlockwise (Figure 6.6) in the six-density peak bundle of the projection map. In either model, helix 3 would be ascribed to the same density feature but using the residue conservation analysis, the anticlockwise model presented in Figure 6.5 is more likely. Helix 5 has a relatively variable face lined with hydrophobic residues and the anti-clockwise arrangement exposes more of the helix to the surrounding lipid whilst also placing the helix 1 in a buried location between helices 2, 6 and the interface between subunits consistent with its conserved residues on all sides of the helix.

The anticlockwise model also places helices 4 and 2 in positions consistent with their degrees of conservation. The projection map has a region of low density inside the hexameric ring which must be filled with lipid. In the unfavoured clockwise model (Figure 6.6), helix 4 is placed in a position that is largely exposed to the lipid within the hexameric ring occupied by helix 2 in the anticlockwise model (Figure 6.5). Residues from either helices 3 or 4 would have to face the central lipid but the surfaces of helices 3 and 4 are largely conserved. The anticlockwise model (Figure 6.5) is therefore favoured over the clockwise model as helix 2 does have a variable face and is placed in the position exposed to the central lipid. Also, more of the surface of helix 4 is conserved so as with helix 1, its assigned location is appropriate as it is buried between helices 3 and 5 and the interface between subunits with less of its surface exposed to the lipid.

The model (Figure 6.5) has the helices oriented so that variable faces are exposed to the surrounding lipid and with polar residues facing other helices. The helices have arbitrarily been depicted as being viewed from the cytoplasmic face of the membrane with the helix 7 C-terminal projecting out of the page.

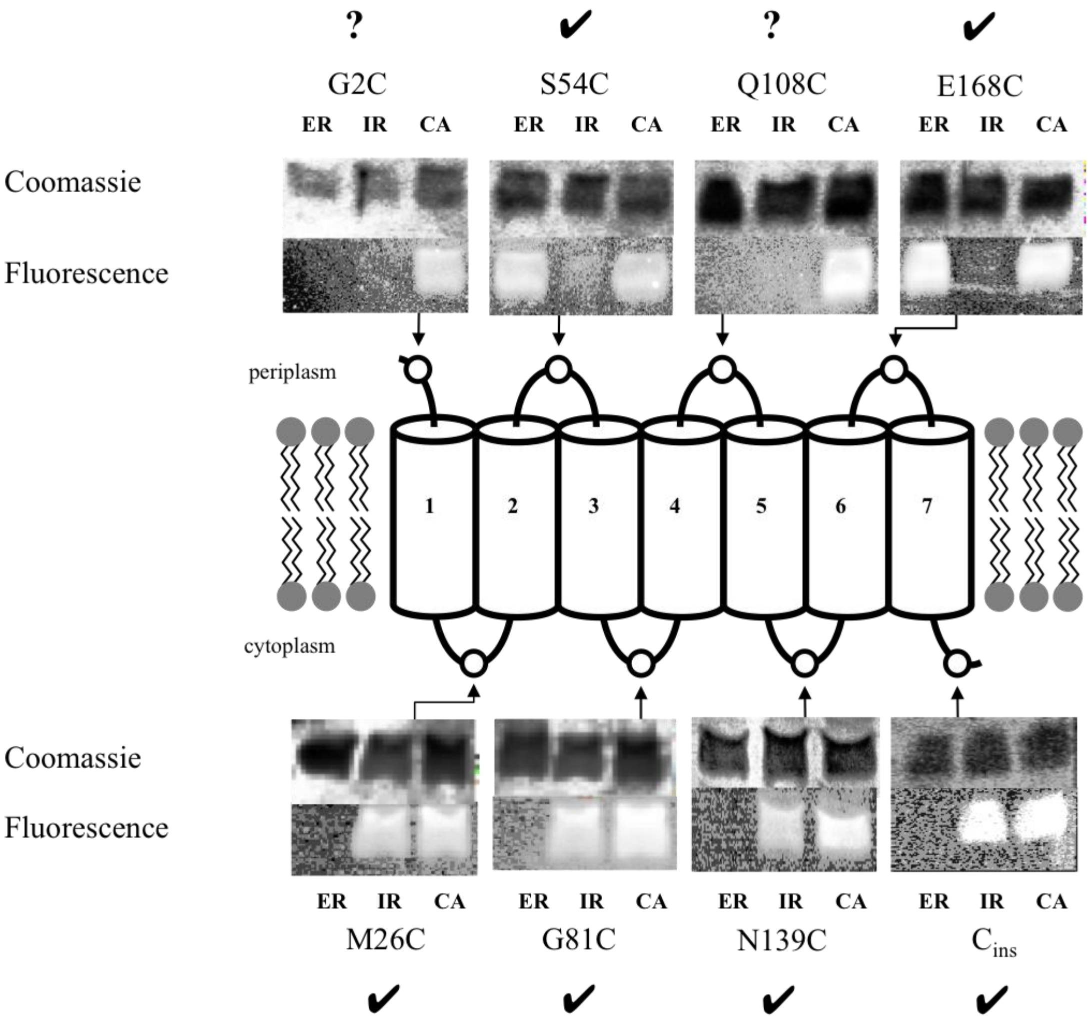


Figure 6.3 Site Directed Fluorescence Labelling (Huysman et al., 2012)
UAC$_{Bc}$ has a predicted 7 transmembrane helix topology shown in the centre of the figure. A series of single-cysteine mutants were produced with cysteines in putative locations on either the periplasmic or cytoplasmic loops. The periplasmic or cytoplasmic locations of the cyteines were determined by labelling with OGM either in the intact cells (ER) or unsealed membranes (CA). Additionally, reaction with MTSET prior to unsealing membranes should also prevent labelling of periplasmic cysteines with OGM and only allow cytoplasmic cysteines to be labelled (IR). For each mutant and its three labelling conditions, the region from a Coomassie stained gel is shown with the results from labelling.
Positions where the data strongly supports the putative location for the cysteine are marked with ✓ and where the data approximately supports the model are marked ?.

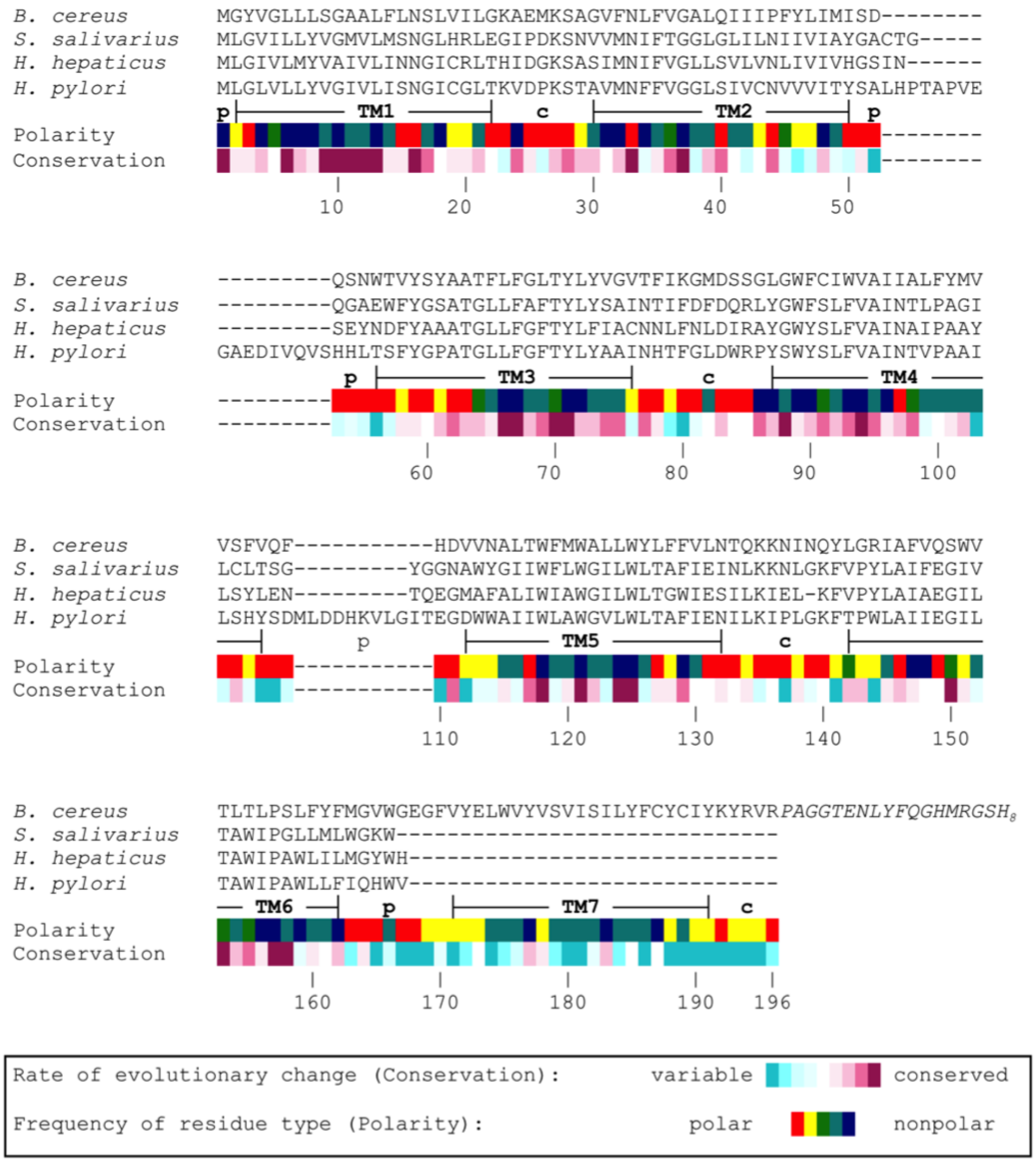


Figure 6.4 Sequence Alignment of UAC family members (Huysmans et al., 2012) 64 sequences of UAC family members were aligned and the sequences for UAC$_{Bc}$ and UreI from *S. salivarius*, *H. hepticus* and *H. pylori* are shown above. Residues were assigned with a numerical value of 1 to 9 based on their conservation and are represented by a colour scale above. The frequency of polarity type at each position was determined. Positions occupied by polar residues in >20% of sequences are red and positions occupied by polar residues in ≤20 % sequences are yellow. Positions where serine and threonine are the only polar residues in >20 % of sequences are green and in ≤20% are turquoise. Positions never occupied by a polar residue are coloured blue. Non-transmembrane helical segments are labelled ‘p’ or ‘c’ depending on whether they occupy a predicted periplasmic or cytoplasmic location respectively. Residues 197 and above are the purification tag of the construct of UAC$_{Bc}$ with the cleavable His-tag.

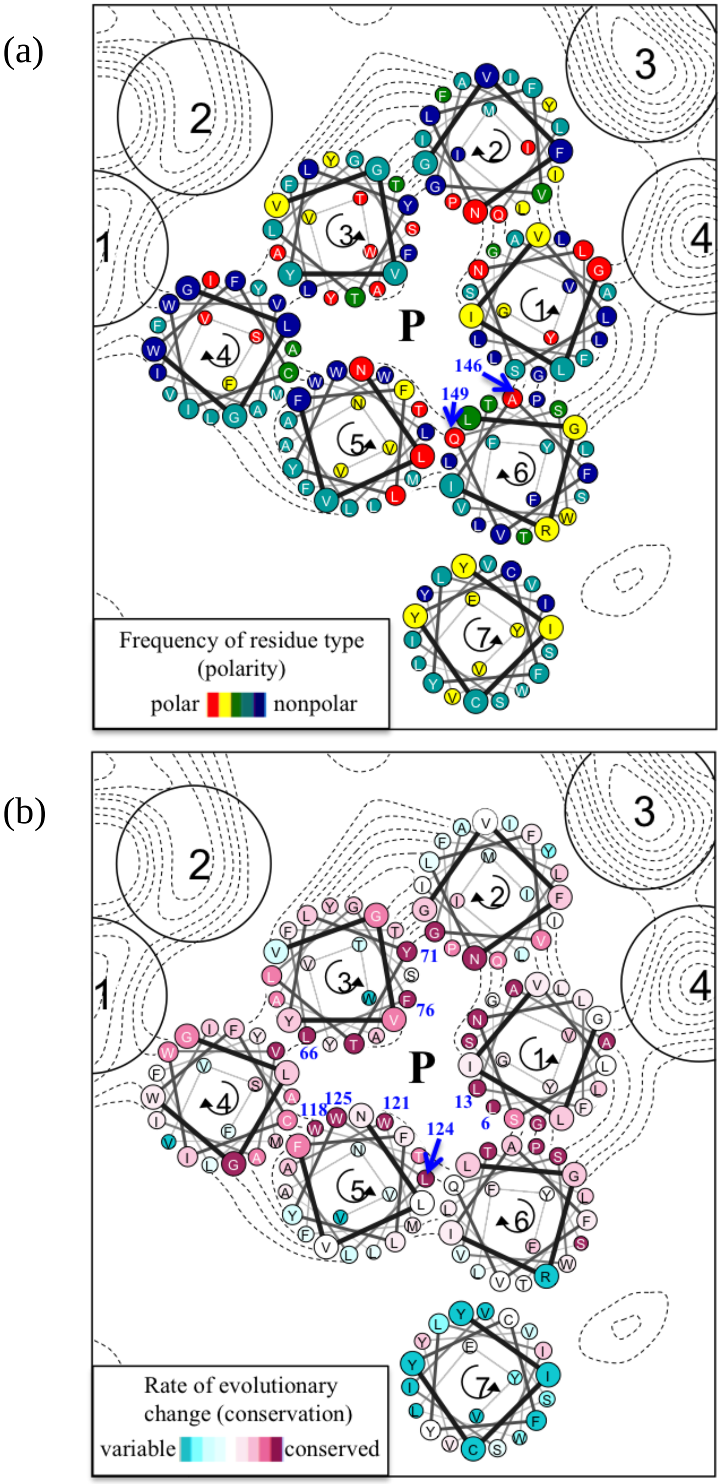


Figure 6.5 Model for the arrangement and orientation of the UAC$_{Bc}$ helices (Huysmans et al., 2012)

Helical wheel plots of the predicted helices have been overlaid on the 9 Å projection map according to their most likely orientation from the combination of analysis of the frequency of polarity and residue conservation. (a) Frequency of polarity colouring and (b) conservation colouring are as described in Figure 6.4. Residues nearer to the cytoplasmic face of the membrane are drawn in larger circles within the helical wheel plots. The pore is within the three lobed cavity labelled ‘P’.

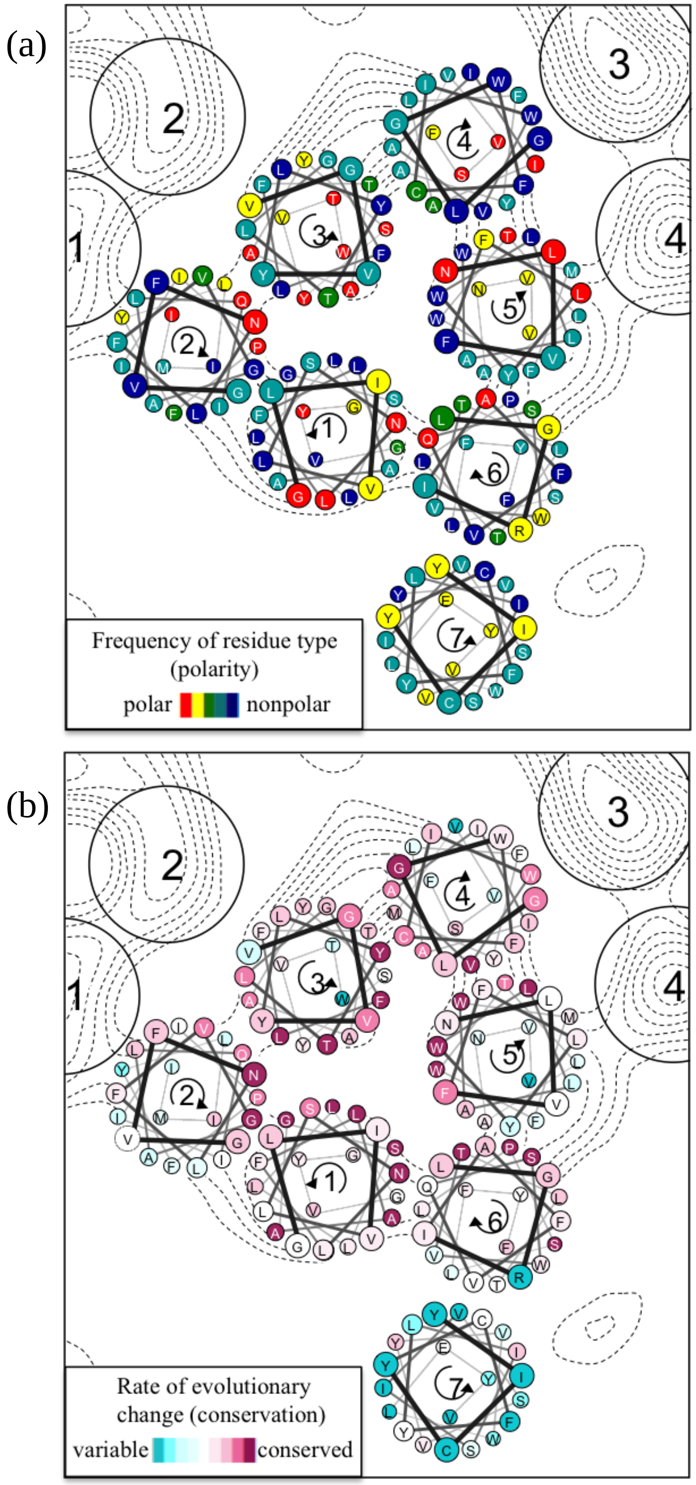


Figure 6.6 Alternative and unfavoured model for the arrangement and orientation of the UAC$_{Bc}$ helices

Helical wheel plots of the predicted helices have been overlaid on the 9 Å projection map in an alternative model where helices 1-6 are arranged clockwise when viewed from the cytoplasm. (a) Frequency of polarity colouring and (b) conservation colouring are as described in Figure 6.4.

6.3.5 Gold Labelling of 2D Crystals

Attempts were made to obtain a projection map of UAC$_{Bc}$ after incubation of Ni-NTA (II) conjugated to a 1.8 nm nanogold particle with the 2D crystals of UAC$_{Bc}$ retaining a His-tag. It was thought that the resulting projection map would show an extra density due to the nanogold bound to the C-terminal His-tag and therefore location of the 7th transmembrane helix.

A single image was processed and as its centrosymmetrically averaged phase residuals were below random up to a resolution of 14 Å, the map was calculated to this resolution (Figure 6.7). The map closely resembles the merged projection map (Figure 5.7) and at its lower resolution, the putative subunit appears as a three lobed object with an additional feature peripheral density but with no signs of an additional density from the nanogold tag.

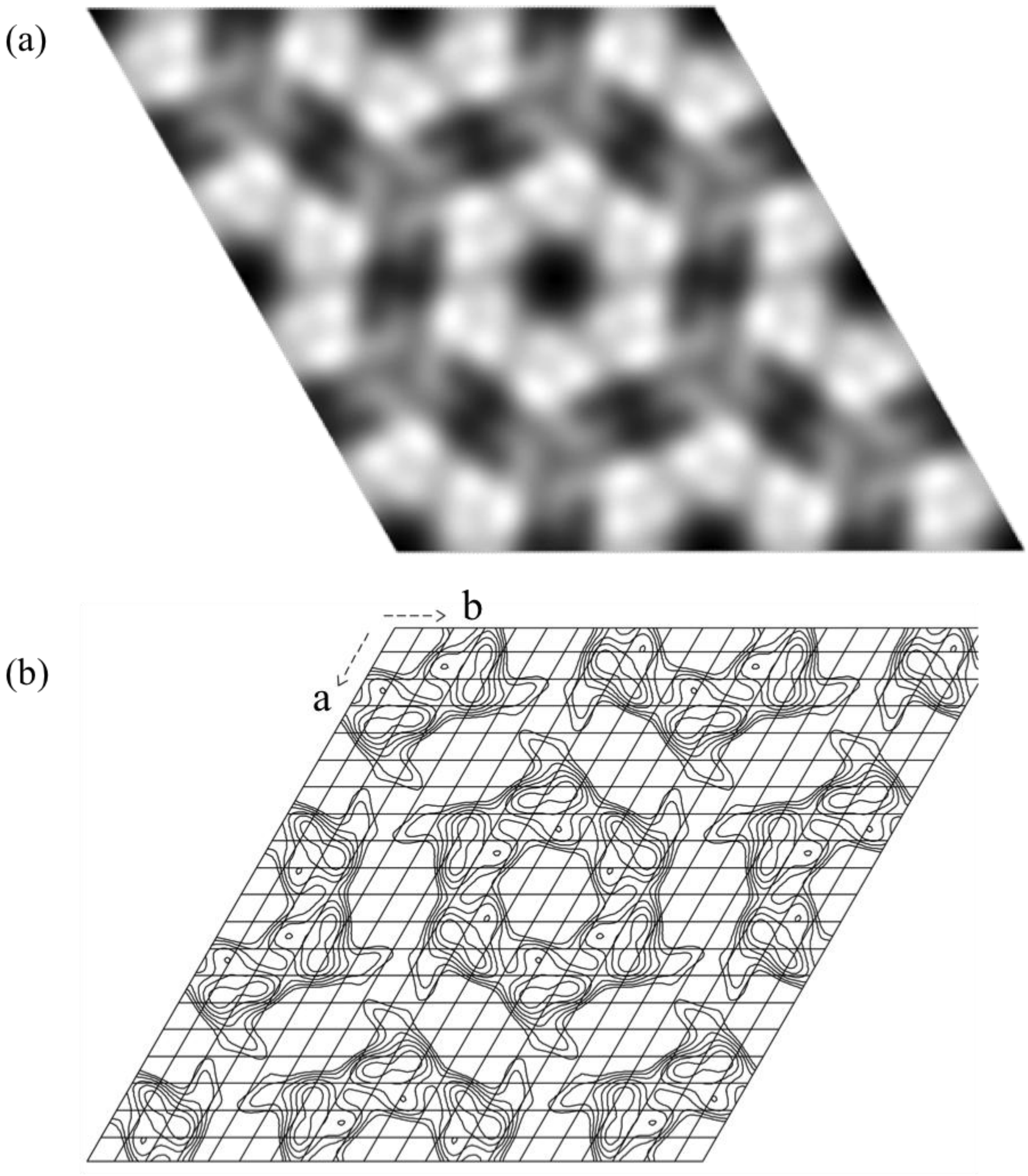


Figure 6.7 Processed image of a UAC$_{Bc}$ crystal incubated with 1.8 nm Ni-NTA nanogold with *p*6 symmetry imposed.

(a) Grayscale map (b) Contour map with contours above the mean at ~0.3 x root mean square density.

The map is calculated to 14 Å resolution with unit cell a = b = 108 Å and γ = 120°.

6.4 Discussion

6.4.1 Oligomeric State of UAC$_{Bc}$

The construct used for EM studies by the author was observed to form multimers during purification as seen on the A_{280} profile during size exclusion chromatography on a Superose 6 10/300 column (Chapter 4). The 120 Å width of the protein:detergent particles viewed on EM grid (Chapter 4 Figure 4.6) was consistent with the width of the 95.5 Å hexamers in the projection map (Chapter 5) accounting for the increase in volume due to detergent binding. The UAC$_{Bc}$ construct with its His-tag removed by TEV protease cleavage was also found to form oligomers up to hexamers after cross-linking and analysis by SDS-PAGE (Figure 6.1). The strongest evidence comes from SEC-MALLS which determines the contribution of the protein to the overall mass of the protein:detergent particle and gave a main peak with a calculated mass corresponding to a hexamer (Figure 6.2). UAC$_{Bc}$ is most likely to exist in its native membranes as hexamers and the hexamers were not an artifact of crystallisation. The TEV protease cleaved construct was also observed to form a particle with a mass corresponding to an 18-mer according to SEC-MALLS analysis but not the 12-mers expected from observations of double-layered stacks on EM grids (Chapter 4). A possible explanation for the 18-mer would be a triple stacked hexamer and would require imaging of the TEV protease cleaved UAC$_{Bc}$ construct in detergent on EM grids.

6.4.2 Model of UAC$_{Bc}$

UAC$_{Bc}$ was originally chosen as a target for structural studies on the basis that the predicted seven transmembrane helix topology would allow for a cytoplasmic location for a His tag possibly allowing better expression of the protein whilst other UAC family members have six transmembrane segments. Site directed fluorescence labelling provided direct evidence for the seven transmembrane topology (Figure 6.3) and was consistent with the interpretation of the seven density peaks as alpha helices in the projection map. Sequence alignments were then use to interpret the projection map and produce a model for the orientation of the helices. Despite the use of helix 7 as a starting point in producing the model, the lack of residue conservation and its location at the periphery of the protein in the projection map makes it uncertain what role the helix plays in the function of UAC_{Bc}.

Currently there are no high resolution structures of UAC_{Bc} but there are structures of the unrelated urea transporters from *Desulfovibrio vulgaris* (Levin et al., 2009) and *Bos taurus* (Levin et al., 2012) from the UT family. UAC_{Bc} may share some structural features in common with the transporters with the substrate pore being formed of hydrophobic residues.

In the UAC_{Bc} model (Figure 6.5), there are highly conserved hydrophobic residues along the faces of helices pointing towards the putative channel. These include L6 and L13 on helix 1, L66, F76 and Y71 on helix 3 and W118, W121, L124 and Y125 on helix 5. The model's location for these helices also has them protruding further into the three lobed cavity consistent with their suggested locations for channel lining residues. However, the resolution of the map does not show how far electron density protrudes into the three lobed cavity making it difficult to determine how the conserved residues form the channel pore. Additionally, there is the uncertainty of whether the projection map represents an 'open' or 'closed' conformation of UAC_{Bc}.

The next step in the modeling of UAC_{Bc} would be to obtain and process a 3D data set from unstained crystals. A map calculated from 2D crystals tilted up to 60° would have a resolution perpendicular to the plane of the membrane ~1.3 times worse than the resolution in the plane of the membrane due to missing amplitude and phase data in a cone shaped region of reciprocal space (Glaeser et al., 1989). With the best crystals of UAC_{Bc} producing a merged map with an in plane resolution of 9 Å (Chapter 5), collecting tilted data from crystals of similar quality would produce a low resolution map but could provide information to validate the proposed helical arrangement (Figure 6.5) as well as information to estimate bends and the axes and angles of tilts of the helices relative to the membrane (Unger et al., 1997). Such information could show locations of constriction that may correspond to the substrate pore.

In practice, it was found that very few crystals could be used for 3D data collection. Due to the narrowing of the field of view when tilting the holder, crystals near to the grid bars would be obscured meaning that at tilt angles of 45° and upwards, several grid squares had to be searched in order to find possible objects for imaging. This compounded the difficulties from variable crystal quality. Of the images showing possible spots that were scanned and processed, specimen drift resulted in loss of data and uninterpretable images.

6.4.3 Gold Labelling Attempts

The single processed image of UAC_{Bc} crystals incubated with 1.8 nm Ni-NTA nanogold at a resolution of 14 Å (Figure 6.7) closely resembles the merged projection map (Figure 5.7) but there was insufficient higher resolution data to draw any conclusions on the location of the nanogold and the location of helix 7. There remains uncertainty about the degree of occupancy of the nanogold particles and whether they are represented in the projection map after crystallographic processing. From AFM imaging (Chapter 5), it was found that the crystals may be found in vesicles and it is possible that the His-tag is inaccessible for binding if the protein is orientated that the His-tag is facing the 'inside' of the vesicle. Additionally, with the 7th helix appearing to be involved in forming crystal contacts, binding of the nanogold may disrupt the crystallinity of the crystals.

6.4.4 Further Experiments

The current model of UAC_{Bc} produced using the combination of sequence alignment analysis and the merged projection map could be used to plan additional experiments to test the model and the functions of the conserved residues.

A series of single cysteine mutants was used to provide evidence for the 7 transmembrane helix topology of UAC_{Bc} through binding of a fluorescent label (Figure 6.3). This could be pursued further following an approach taken with the study of the lactose permease LacY from *E. coli*. Experiments using a library of cysteine scanning mutants determined which residues were irreplaceable for transport and inferred information on the conformational changes based on the reactivity of the residues upon ligand binding (Frillingos et al., 1998). Data that would aid modelling could include distance constraints between helices determined by cross-linking of cysteine residues (Rastogi and Girvin, 1999) or distances between a metal binding site and a site directed spin label using electron paramagnetic resonance spectroscopy (Voss et al., 1995).

In terms of achieving high resolution in 3D using EM, the foremost requirement would be higher yields of crystals so that the current difficulties with finding samples for imaging would be circumvented. However, it is not clear how to grow extensive single layered sheet crystals of UAC_{Bc} when the protein has a predisposition for forming multilayered aggregates. A possibility from the work of the Membrane Protein Structure Initiative is to try crystallising another orthologue of UreI or the *H. pylori* orthologue itself. Whilst high throughput with electron microscopy may be limited by the availability of suitable 2D crystals, the integration of the MRC image processing software into a the 2dx interface allows several images to be processed relatively quickly which improved the throughput of image processing.

6.5 Conclusions

UAC_{Bc} most likely exists and functions in its native membranes as a hexamer although further investigation is required to determine the functional relevance. The monomer of UAC_{Bc} consistent with topology predictions spans the membrane with seven transmembrane helices according to data from site directed fluorescent labelling and is consistent with the densities in the merged projection map. Multiple sequence alignment of UAC_{Bc} orthologues shows that the most conserved residues occur within the first six transmembrane helical segments and in the present model these enclose a cavity which is the putative location of the substrate pore. An absence of conserved residues and a predominance of hydrophobic residues suggests that helix 7 is heavily exposed to the lipid. Crystal contacts in 2D crystals of UAC_{Bc} appeared to be formed by helix 7 but the functional relevance of this helix is uncertain.

Recently, an X-ray structure of the UreI urea channel from *H. pylori (HpUreI)* was determined by Strugatsky et al. (2013) and the structure has many similarities to the model of UAC_{Bc} described in this thesis. *Hp*UreI also forms a hexameric ring with a diameter of ~95 Å with the center of the hexamer being filled with an ordered lipid plug. The helices within the *Hp*UreI monomer are slightly tilted and form a twisted bundle of helices and when viewed from the cytoplasmic face, helices 1 to 6 are arranged anticlockwise with respect to each other (Figure 6.8). The outer surface of the hexamer exposed to the surrounding lipid is formed by helices 4 to 6 whilst helix 2 is exposed to the lipid filled center of the hexamer. The substrate pore is lined with conserved residues. Urea enters the channel from the periplasm into a vestibule formed by residues L2, Y76, W142 and W146. Urea then passes constriction site 1 formed by L6, F84 and W149 and constriction site 2 formed by L13, T87, W88, L152 and W153. Urea then enters a funnel shaped vestibule formed by W104, N16 and N33 and exits into the cytoplasm (Figure 6.8).

The X-ray structure of *Hp*Urei validates the interpretation of the EM projection map and the model of UAC_{Bc} should serve in the design of future experiments exploring the structure and function relationships of the protein.

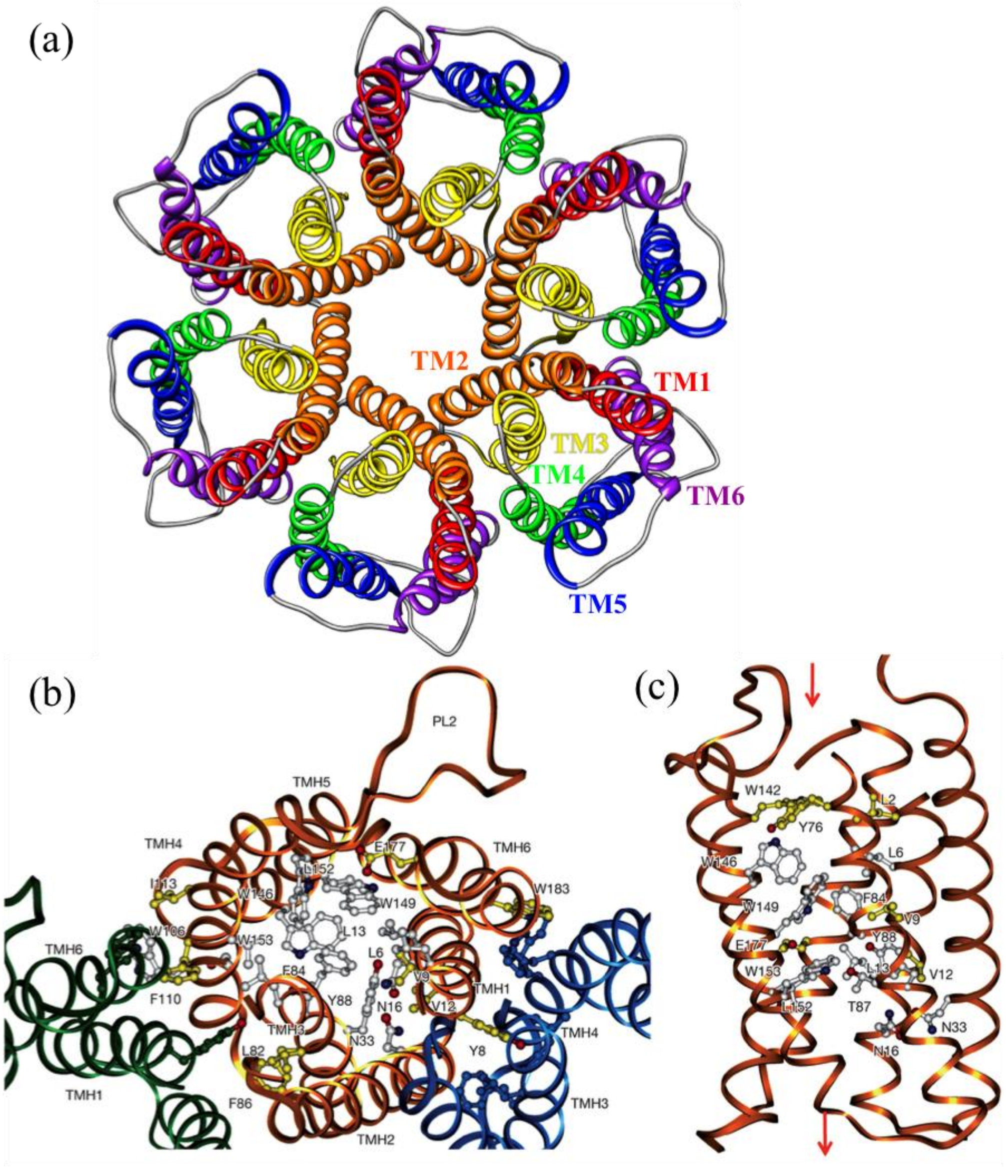


Figure 6.8 Structure of *Hp*UreI

(a) The UreI hexamer viewed from the cytoplasmic face of the protein. Transmembrane alpha helices are coloured. Produced from PDB ID 3UX4 (Strugatsky et al., 2013) using UCSF Chimera (Pettersen et al., 2004).

(b) and (c) were adapted from Strugatsky et al. (2013) with permission from Macmillan Publishers Ltd.

(b) *Hp*UreI viewed from the periplasm. The monomer backbones have been coloured green, gold and blue. Residues sidechains conserved in all members of the UAC family have been coloured white and residue sidechains conserved in known urea channels are coloured yellow.

(c) A monomer viewed parallel to the membrane with helices 1 and 2 removed. Sidechains are coloured as in (b). The red arrows indicate the regions of urea entry from the periplasm (top) and exit to the cytoplasm (bottom).

References

ABBE, E. 1873. Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. *Archiv für mikroskopische Anatomie,* 9**,** 413-418.

ABE, K., TANI, K. & FUJIYOSHI, Y. 2010. Structural and functional characterization of H+, K+-ATPase with bound fluorinated phosphate analogs. *J Struct Biol,* 170**,** 60-8.

AGARD, D. A. 1983. A least-squares method for determining structure factors in three-dimensional tilted-view reconstructions. *J Mol Biol,* 167**,** 849-52.

ALLER, S. G. & UNGER, V. M. 2006. Projection structure of the human copper transporter CTR1 at 6-A resolution reveals a compact trimer with a novel channel-like architecture. *Proc Natl Acad Sci U S A,* 103**,** 3627-32.

AMOS, L. A., HENDERSON, R. & UNWIN, P. N. 1982. Three-dimensional structure determination by electron microscopy of two-dimensional crystals. *Prog Biophys Mol Biol,* 39**,** 183-231.

BALDWIN, J. & HENDERSON, R. 1984. Measurement and Evaluation of Electron-Diffraction Patterns from 2-Dimensional Crystals. *Ultramicroscopy,* 14**,** 319-335.

BALDWIN, J. M. 1993. The probable arrangement of the helices in G protein-coupled receptors. *EMBO J,* 12**,** 1693-703.

BARROS, T. & KÜHLBRANDT, W. 2009. Crystallisation, structure and function of plant light-harvesting Complex II. *Biochim Biophys Acta,* 1787**,** 753-72.

BECKERS, G., BENDT, A. K., KRAMER, R. & BURKOVSKI, A. 2004. Molecular identification of the urea uptake system and transcriptional analysis of urea transporter- and urease-encoding genes in Corynebacterium glutamicum. *J Bacteriol,* 186**,** 7645-52.

BEREZIN, C., GLASER, F., ROSENBERG, J., PAZ, I., PUPKO, T., FARISELLI, P., CASADIO, R. & BEN-TAL, N. 2004. ConSeq: the identification of functionally and structurally important residues in protein sequences. *Bioinformatics,* 20**,** 1322-4.

BERNSEL, A., VIKLUND, H., FALK, J., LINDAHL, E., VON HEIJNE, G. & ELOFSSON, A. 2008. Prediction of membrane-protein topology from first principles. *Proc Natl Acad Sci U S A,* 105**,** 7177-81.

BERNSEL, A., VIKLUND, H., HENNERDAL, A. & ELOFSSON, A. 2009. TOPCONS: consensus prediction of membrane protein topology. *Nucleic Acids Res,* 37**,** W465-8.

BINNIG, G., QUATE, C. F. & GERBER, C. 1986. Atomic force microscope. *Phys Rev Lett,* 56**,** 930-933.

BREMER, A., HENN, C., ENGEL, A., BAUMEISTER, W. & AEBI, U. 1992. Has Negative Staining Still a Place in Biomacromolecular Electron-Microscopy. *Ultramicroscopy,* 46**,** 85-111.

BULLOUGH, P. & HENDERSON, R. 1987. Use of Spot-Scan Procedure for Recording Low-Dose Micrographs of Beam-Sensitive Specimens. *Ultramicroscopy,* 21**,** 223-229.

BULLOUGH, P. A. & TULLOCH, P. A. 1991. Spot-Scan Imaging of Microcrystals of an Influenza Neuraminidase Antibody Fragment-Complex. *Ultramicroscopy,* 35**,** 131-143.

CHEBROU, H., BIGEY, F., ARNAUD, A. & GALZY, P. 1996. Amide metabolism: a putative ABC transporter in Rhodococcus sp. R312. *Gene,* 182**,** 215-8.

CONDE, A., DIALLINAS, G., CHAUMONT, F., CHAVES, M. & GEROS, H. 2009. Transporters, channels, or simple diffusion? Dogmas, atypical roles and complexity in transport systems. *Int J Biochem Cell Biol*.

COVACCI, A., TELFORD, J. L., DEL GIUDICE, G., PARSONNET, J. & RAPPUOLI, R. 1999. Helicobacter pylori virulence and genetic geography. *Science,* 284**,** 1328-1333.

CREWE, A. V., WALL, J. & LANGMORE, J. 1970. Visibility of Single Atoms. *Science,* 168**,** 1338-1340.

CROWTHER, R. A., HENDERSON, R. & SMITH, J. M. 1996. MRC image processing programs. *J Struct Biol,* 116**,** 9-16.

DAWSON, R. J. P. & LOCHER, K. P. 2007. Structure of the multidrug ABC transporter Sav1866 from Staphylococcus aureus in complex with AMP-PNP. *Febs Letters,* 581**,** 935-938.

DEACON, S. E., ROACH, P. C., POSTIS, V. L., WRIGHT, G. S., XIA, X., PHILLIPS, S. E., KNOX, J. P., HENDERSON, P. J., MCPHERSON, M. J. & BALDWIN, S. A. 2008. Reliable scale-up of membrane protein over-expression by bacterial auto-induction: from microwell plates to pilot scale fermentations. *Mol Membr Biol,* 25**,** 588-98.

DEROSIER, D. J. & KLUG, A. 1968. Reconstruction of 3 Dimensional Structures from Electron Micrographs. *Nature,* 217**,** 130-&.

DOBRIANOV, I., CAYLOR, C., LEMAY, S. G., FINKELSTEIN, K. D. & THORNE, R. E. 1999. X-ray diffraction studies of protein crystal disorder. *Journal of Crystal Growth,* 196**,** 511-523.

DOLS-PEREZ, A., FUMAGALLI, L., SIMONSEN, A. C. & GOMILA, G. 2011. Ultrathin spin-coated dioleoylphosphatidylcholine lipid layers in dry conditions: a combined atomic force microscopy and nanomechanical study. *Langmuir,* 27**,** 13165-72.

DOWNING, K. H. & GLAESER, R. M. 1986. Improvement in high resolution image quality of radiation-sensitive specimens achieved with reduced spot size of the electron beam. *Ultramicroscopy,* 20**,** 269-78.

FENTON, R. A. 2009. Essential role of vasopressin-regulated urea transport processes in the mammalian kidney. *Pflugers Archiv-European Journal of Physiology,* 458**,** 169-177.

FINKELSTEIN, A. 1976. Water and nonelectrolyte permeability of lipid bilayer membranes. *J Gen Physiol,* 68**,** 127-35.

FORD, R. C., HEFTI, A. & ENGEL, A. 1990. Ordered Arrays of the Photosystem-I Reaction Center after Reconstitution - Projections and Surface Reliefs of the Complex at 2-Nm Resolution. *Embo Journal,* 9**,** 3067-3075.

FORD, R. C. & HOLZENBURG, A. 2008. Electron crystallography of biomolecules: mysterious membranes and missing cones. *Trends Biochem Sci,* 33**,** 38-43.

FRANK, J. 2006. *Three-dimensional electron microscopy of macromolecular assemblies : visualization of biological molecules in their native state,* New York ; Oxford, Oxford University Press.

FRILLINGOS, S., SAHIN-TOTH, M., WU, J. & KABACK, H. R. 1998. Cys-scanning mutagenesis: a novel approach to structure function relationships in polytopic membrane proteins. *FASEB J,* 12**,** 1281-99.

GIPSON, B., ZENG, X., ZHANG, Z. Y. & STAHLBERG, H. 2007. 2dx--user-friendly image processing for 2D crystals. *J Struct Biol,* 157**,** 64-72.

GLAESER, R. M. 2007. *Electron crystallography of biological macromolecules,* Oxford, Oxford University Press.

GLAESER, R. M. & HALL, R. J. 2011. Reaching the information limit in cryo-EM of biological macromolecules: experimental aspects. *Biophys J,* 100**,** 2331-7.

GLAESER, R. M., TONG, L. & KIM, S. H. 1989. Three-dimensional reconstructions from incomplete data: interpretability of density maps at "atomic" resolution. *Ultramicroscopy,* 27**,** 307-18.

GLOVER, C. A., POSTIS, V. L., CHARALAMBOUS, K., TZOKOV, S. B., BOOTH, W. I., DEACON, S. E., WALLACE, B. A., BALDWIN, S. A. & BULLOUGH, P. A. 2011. AcrB contamination in 2-D crystallization of membrane proteins: lessons from a sodium channel and a putative monovalent cation/proton antiporter. *J Struct Biol,* 176**,** 419-24.

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**,** 633-8.

GONEN, T., SLIZ, P., KISTLER, J., CHENG, Y. & WALZ, T. 2004. Aquaporin-0 membrane junctions reveal the structure of a closed water pore. *Nature,* 429**,** 193-7.

GRAY, L. R., GU, S. X., QUICK, M. & KHADEMI, S. 2011. Transport kinetics and selectivity of HpUreI, the urea channel from Helicobacter pylori. *Biochemistry,* 50**,** 8656-63.

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**,** 393-421.

GRISSHAMMER, R. & TATE, C. G. 1995. Overexpression of integral membrane proteins for structural studies. *Q Rev Biophys,* 28**,** 315-422.

HALL, C. E. 1955. Electron densitometry of stained virus particles. *J Biophys Biochem Cytol,* 1**,** 1-12.

HAYWARD, S. B. & GLAESER, R. M. 1979. Radiation Damage of Purple Membrane at Low Temperature. *Ultramicroscopy,* 4**,** 201-210.

HEDIGER, M. A., SMITH, C. P., YOU, G., LEE, W. S., KANAI, Y. & SHAYAKUL, C. 1996. Structure, regulation and physiological roles of urea transporters. *Kidney Int,* 49**,** 1615-23.

HEIJNE, G. 1986. The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology. *EMBO J,* 5**,** 3021-7.

HELENIUS, A. & SIMONS, K. 1975. Solubilization of membranes by detergents. *Biochim Biophys Acta,* 415**,** 29-79.

HENDERSON, R. 1992. Image contrast in high-resolution electron microscopy of biological macromolecules: TMV in ice. *Ultramicroscopy,* 46**,** 1-18.

HENDERSON, R., BALDWIN, J. M., CESKA, T. A., ZEMLIN, F., BECKMANN, E. & DOWNING, K. H. 1990. Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J Mol Biol,* 213**,** 899-929.

HENDERSON, R., BALDWIN, J. M., DOWNING, K. H., LEPAULT, J. & ZEMLIN, F. 1986. Structure of purple membrane from halobacterium halobium:

recording, measurement and evaluation of electron micrographs at 3.5 Å resolution. *Ultramicroscopy,* 19**,** 147-178.

HENDERSON, R. & UNWIN, P. N. 1975. Three-dimensional model of purple membrane obtained by electron microscopy. *Nature,* 257**,** 28-32.

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**,** 628-39.

HOLM, P. J., BHAKAT, P., JEGERSCHOLD, C., GYOBU, N., MITSUOKA, K., FUJIYOSHI, Y., MORGENSTERN, R. & HEBERT, H. 2006. Structural basis for detoxification and oxidative stress protection in membranes. *J Mol Biol,* 360**,** 934-45.

HUYSMANS, G. H., CHAN, N., BALDWIN, J. M., POSTIS, V. L., TZOKOV, S. B., DEACON, S. E., YAO, S. Y., YOUNG, J. D., MCPHERSON, M. J., BULLOUGH, P. A. & BALDWIN, S. A. 2012. A urea channel from Bacillus cereus reveals a novel hexameric structure. *Biochem J,* 445**,** 157-66.

JAP, B. K., ZULAUF, M., SCHEYBANI, T., HEFTI, A., BAUMEISTER, W., AEBI, U. & ENGEL, A. 1992. 2D crystallization: from art to science. *Ultramicroscopy,* 46**,** 45-84.

JEGERSCHOLD, C., PAWELZIK, S. C., PURHONEN, P., BHAKAT, P., GHEORGHE, K. R., GYOBU, N., MITSUOKA, K., MORGENSTERN, R., JAKOBSSON, P. J. & HEBERT, H. 2008. Structural basis for induced formation of the inflammatory mediator prostaglandin E2. *Proc Natl Acad Sci U S A,* 105**,** 11110-5.

KÜHLBRANDT, W. 1992. Two-dimensional crystallization of membrane proteins. *Q Rev Biophys,* 25**,** 1-49.

KÜHLBRANDT, W., WANG, D. N. & FUJIYOSHI, Y. 1994. Atomic model of plant light-harvesting complex by electron crystallography. *Nature,* 367**,** 614-21.

KUO, A., DOMENE, C., JOHNSON, L. N., DOYLE, D. A. & VENIEN-BRYAN, C. 2005. Two different conformational states of the KirBac3.1 potassium channel revealed by electron crystallography. *Structure,* 13**,** 1463-72.

KAUFMANN, T. C., ENGEL, A. & REMIGY, H. W. 2006. A novel method for detergent concentration determination. *Biophysical Journal,* 90**,** 310-317.

KYTE, J. & DOOLITTLE, R. F. 1982. A simple method for displaying the hydropathic character of a protein. *J Mol Biol,* 157**,** 105-32.

LACAPERE, J. J., STOKES, D. L., MOSSER, G., RANCK, J. L., LEBLANC, G. & RIGAUD, J. L. 1997. Two-dimensional crystal formation from solubilized membrane proteins using Bio-Beads to remove detergent. *Ann N Y Acad Sci,* 834**,** 9-18.

LASIC, D. D. 1988. The mechanism of vesicle formation. *Biochem J,* 256**,** 1-11.

LAURENT, T. C. & KILLANDER, J. 1964. Theory of Gel Filtration and Its Experimental Verification. *Journal of Chromatography,* 14**,** 317-330.

LE COUTRE, J. & KABACK, H. K. 2000. Structure-function relationships of integral membrane proteins: membrane transporters vs channels. *Biopolymers,* 55**,** 297-307.

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. *Journal of Molecular Biology,* 308**,** 639-647.

LEE, S. Y., LEE, A., CHEN, J. & MACKINNON, R. 2005. Structure of the KvAP voltage-dependent K+ channel and its dependence on the lipid membrane. *Proc Natl Acad Sci U S A,* 102**,** 15441-6.

LEMBCKE, G., DURR, R., HEGERL, R. & BAUMEISTER, W. 1991. Image-Analysis and Processing of an Imperfect 2-Dimensional Crystal - the Surface-Layer of the Archaebacterium Sulfolobus-Acidocaldarius Re-Investigated. *Journal of Microscopy-Oxford,* 161**,** 263-278.

LEVIN, E. J., CAO, Y., ENKAVI, G., QUICK, M., PAN, Y., TAJKHORSHID, E. & ZHOU, M. 2012. Structure and permeation mechanism of a mammalian urea transporter. *Proc Natl Acad Sci U S A***,** 11194-9.

LEVIN, E. J., QUICK, M. & ZHOU, M. 2009. Crystal structure of a bacterial homologue of the kidney urea transporter. *Nature,* 462**,** 757-61.

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**,** 44-52.

MIDURA, R. J. & YANAGISHITA, M. 1995. Chaotropic solvents increase the critical micellar concentrations of detergents. *Anal Biochem,* 228**,** 318-22.

MOLLER, C., ALLEN, M., ELINGS, V., ENGEL, A. & MULLER, D. J. 1999. Tapping-mode atomic force microscopy produces faithful high-resolution images of protein surfaces. *Biophys J,* 77**,** 1150-8.

MOSSER, G. 2001. Two-dimensional crystallogenesis of transmembrane proteins. *Micron,* 32**,** 517-40.

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**,** 599-605.

NIETLISPACH, D. & GAUTIER, A. 2011. Solution NMR studies of polytopic alpha-helical membrane proteins. *Current Opinion in Structural Biology,* 21**,** 497-508.

NUSSBERGER, S., DORR, K., WANG, D. N. & KUHLBRANDT, W. 1993. Lipid-protein interactions in crystals of plant light-harvesting complex. *J Mol Biol,* 234**,** 347-56.

ORBACH, E. & FINKELSTEIN, A. 1980. The nonelectrolyte permeability of planar lipid bilayer membranes. *J Gen Physiol*, 75**,** 427-36.

PALCZEWSKI, K., KUMASAKA, T., HORI, T., BEHNKE, C. A., MOTOSHIMA, H., FOX, B. A., LE TRONG, I., TELLER, D. C., OKADA, T., STENKAMP, R. E., YAMAMOTO, M. & MIYANO, M. 2000. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science,* 289**,** 739-45.

PEROZO, E., CORTES, D. M. & CUELLO, L. G. 1999. Structural rearrangements underlying K+-channel activation gating. *Science*, 285**,** 73-78.

PETTERSEN, E. F., GODDARD, T. D., HUANG, C. C., COUCH, G. S., GREENBLATT, D. M., MENG, E. C. & FERRIN, T. E. 2004. UCSF chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry,* 25**,** 1605-1612.

POSTIS, V. L., DEACON, S. E., ROACH, P. C., WRIGHT, G. S., XIA, X., INGRAM, J. C., HADDEN, J. M., HENDERSON, P. J., PHILLIPS, S. E., MCPHERSON, M. J. & BALDWIN, S. A. 2008. A high-throughput assay of membrane protein stability. *Mol Membr Biol,* 25**,** 617-24.

PSAKIS, G., POLACZEK, J. & ESSEN, L. O. 2009. AcrB et al.: Obstinate contaminants in a picogram scale. One more bottleneck in the membrane protein structure pipeline. *J Struct Biol,* 166**,** 107-11.

RASTOGI, V. K. & GIRVIN, M. E. 1999. Structural changes linked to proton translocation by subunit c of the ATP synthase. *Nature,* 402**,** 263-8.

RATH, A., GLIBOWICKA, M., NADEAU, V. G., CHEN, G. & DEBER, C. M. 2009. Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proc Natl Acad Sci U S A,* 106**,** 1760-5.

RAUNSER, S., MATHAI, J. C., ABEYRATHNE, P. D., RICE, A. J., ZEIDEL, M. L. & WALZ, T. 2009. Oligomeric structure and functional characterization of the urea transporter from Actinobacillus pleuropneumoniae. *J Mol Biol,* 387**,** 619-27.

RENAULT, L., CHOU, H. T., CHIU, P. L., HILL, R. M., ZENG, X., GIPSON, B., ZHANG, Z. Y., CHENG, A., UNGER, V. & STAHLBERG, H. 2006. Milestones in electron crystallography. *J Comput Aided Mol Des,* 20**,** 519-27.

RIGAUD, J., CHAMI, M., LAMBERT, O., LEVY, D. & RANCK, J. 2000. Use of detergents in two-dimensional crystallization of membrane proteins. *Biochim Biophys Acta,* 1508**,** 112-28.

RIGAUD, J. L., MOSSER, G., LACAPERE, J. J., OLOFSSON, A., LEVY, D. & RANCK, J. L. 1997. Bio-Beads: an efficient strategy for two-dimensional crystallization of membrane proteins. *J Struct Biol,* 118**,** 226-35.

ROACH, P. C., POSTIS, V. L., DEACON, S. E., WRIGHT, G. S., INGRAM, J. C., XIA, X., MCPHERSON, M. J. & BALDWIN, S. A. 2008. Large-scale preparation of bacterial cell membranes by tangential flow filtration. *Mol Membr Biol,* 25**,** 609-16.

SACHS, G., WEEKS, D. L., WEN, Y., MARCUS, E. A., SCOTT, D. R. & MELCHERS, K. 2005. Acid acclimation by Helicobacter pylori. *Physiology (Bethesda),* 20**,** 429-38.

SAIER, M. H., TRAN, C. V. & BARABOTE, R. D. 2006. TCDB: the Transporter Classification Database for membrane transport protein analyses and information. *Nucleic Acids Research,* 34**,** D181-D186.

SCHERTLER, G. F., VILLA, C. & HENDERSON, R. 1993. Projection structure of rhodopsin. *Nature,* 362**,** 770-2.

SCHMIDT-KREY, I. 2007. Electron crystallography of membrane proteins: two-dimensional crystallization and screening by electron microscopy. *Methods,* 41**,** 417-26.

SCHMIDT-KREY, I., KANAOKA, Y., MILLS, D. J., IRIKURA, D., HAASE, W., LAM, B. K., AUSTEN, K. F. & KÜHLBRANDT, W. 2004. Human leukotriene C(4) synthase at 4.5 A resolution in projection. *Structure,* 12**,** 2009-14.

SCHMIDT-KREY, I., LUNDQVIST, G., MORGENSTERN, R. & HEBERT, H. 1998. Parameters for the two-dimensional crystallization of the membrane protein microsomal glutathione transferase. *J Struct Biol,* 123**,** 87-96.

SCHMIDT-KREY, I., MURATA, K., HIRAI, T., MITSUOKA, K., CHENG, Y., MORGENSTERN, R., FUJIYOSHI, Y. & HEBERT, H. 1999. The projection structure of the membrane protein microsomal glutathione transferase at 3 A resolution as determined from two-dimensional hexagonal crystals. *J Mol Biol,* 288**,** 243-53.

SCOTT, D. R., MARCUS, E. A., WEN, Y., SINGH, S., FENG, J. & SACHS, G. 2010. Cytoplasmic Histidine Kinase (HP0244)-Regulated Assembly of Urease with UreI, a Channel for Urea and Its Metabolites, CO2, NH3, and NH4+, Is Necessary for Acid Survival of Helicobacter pylori. *Journal of Bacteriology,* 192**,** 94-103.

SHAW, P. J. & HILLS, G. J. 1981. Tilted Specimen in the Electron-Microscope - a Simple Specimen Holder and the Calculation of Tilt Angles for Crystalline Specimens. *Micron,* 12**,** 279-282.

SHI, D., HSIUNG, H. H., PACE, R. C. & STOKES, D. L. 1995. Preparation and analysis of large, flat crystals of Ca(2+)-ATPase for electron crystallography. *Biophys J,* 68**,** 1152-62.

SIDOROV, M. 2002. ctfExplorer: Interactive Software for 1d and 2d Calculation and Visualization Of TEM Phase Contrast Transfer Function. *Microscopy and Microanalysis,* 8**,** 1572-1573.

SIGNORELL, G. A., CHAMI, M., CONDEMINE, G., SCHENK, A. D., PHILIPPSEN, A., ENGEL, A. & REMIGY, H. W. 2007. Projection maps of three members of the KdgM outer membrane protein family. *J Struct Biol,* 160**,** 395-403.

SINGER, S. J. & NICOLSON, G. L. 1972. The fluid mosaic model of the structure of cell membranes. *Science,* 175**,** 720-31.

SLOTBOOM, D. J., DUURKENS, R. H., OLIEMAN, K. & ERKENS, G. B. 2008. Static light scattering to characterize membrane proteins in detergent solution. *Methods,* 46**,** 73-82.

SONNHAMMER, E. L., VON HEIJNE, G. & KROGH, A. 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol,* 6**,** 175-82.

STENN, K. & BAHR, G. F. 1970. Specimen Damage Caused by Beam of Transmission Electron Microscope, a Correlative Reconsideration. *Journal of Ultrastructure Research,* 31**,** 526-550.

STEWART, G. 2011. The emerging physiological roles of the SLC14A family of urea transporters. *Br J Pharmacol*.

STEWART, G. S. & SMITH, C. P. 2005. Urea nitrogen salvage mechanisms and their relevance to ruminants, non-ruminants and man. *Nutrition Research Reviews,* 18**,** 49-62.

STRUGATSKY, D., MCNULTY, R., MUNSON, K., CHEN, C. K., SOLTIS, S. M., SACHS, G. & LUECKE, H. 2013. Structure of the proton-gated urea channel from the gastric pathogen Helicobacter pylori. *Nature,* 493**,** 255-U144.

STUART, M. C., KONING, R. I., OOSTERGETEL, G. T. & BRISSON, A. 2004. Mechanism of formation of multilayered 2D crystals of the enzyme IIC-mannitol transporter. *Biochim Biophys Acta,* 1663**,** 108-16.

SUBRAMANIAM, S., GERSTEIN, M., OESTERHELT, D. & HENDERSON, R. 1993. Electron diffraction analysis of structural changes in the photocycle of bacteriorhodopsin. *EMBO J,* 12**,** 1-8.

TAYLOR, K. A. & GLAESER, R. M. 1976. Electron-Microscopy of Frozen Hydrated Biological Specimens. *Journal of Ultrastructure Research,* 55**,** 448-456.

THOMPSON, J. D., GIBSON, T. J., PLEWNIAK, F., JEANMOUGIN, F. & HIGGINS, D. G. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res,* 25**,** 4876-82.

THON, F. 1966. Zur Defokussierungsabhangigkeit Des Phasenkontrastes Bei Der Elektronenmikroskopischen Abbildung. *Zeitschrift Fur Naturforschung Part a-Astrophysik Physik Und Physikalische Chemie,* A 21**,** 476-&.

TSAI, C. J., EJSING, C. S., SHEVCHENKO, A. & ZIEGLER, C. 2007. The role of lipids and salts in two-dimensional crystallization of the glycine-betaine transporter BetP from Corynebacterium glutamicum. *J Struct Biol,* 160**,** 275-86.

UNGER, V. M., HARGRAVE, P. A., BALDWIN, J. M. & SCHERTLER, G. F. 1997. Arrangement of rhodopsin transmembrane alpha-helices. *Nature,* 389**,** 203-6.

UNWIN, N. 2005. Refined structure of the nicotinic acetylcholine receptor at 4A resolution. *J Mol Biol,* 346**,** 967-89.

UNWIN, N. & FUJIYOSHI, Y. 2012. Gating movement of acetylcholine receptor caught by plunge-freezing. *J Mol Biol,* 422**,** 617-34.

UNWIN, P. N. & HENDERSON, R. 1975. Molecular structure determination by electron microscopy of unstained crystalline specimens. *J Mol Biol,* 94**,** 425-40.

VALPUESTA, J. M., CARRASCOSA, J. L. & HENDERSON, R. 1994. Analysis of electron microscope images and electron diffraction patterns of thin crystals of phi 29 connectors in ice. *J Mol Biol,* 240**,** 281-7.

VIKLUND, H. & ELOFSSON, A. 2004. Best alpha-helical transmembrane protein topology predictions are achieved using hidden Markov models and evolutionary information. *Protein Sci,* 13**,** 1908-17.

VIKLUND, H. & ELOFSSON, A. 2008. OCTOPUS: improving topology prediction by two-track ANN-based preference scores and an extended topological grammar. *Bioinformatics*, 24**,** 1662-8.

VINOTHKUMAR, K. R., RAUNSER, S., JUNG, H. & KÜHLBRANDT, W. 2006. Oligomeric structure of the carnitine transporter CaiT from Escherichia coli. *J Biol Chem,* 281**,** 4795-801.

VINOTHKUMAR, K. R., SMITS, S. H. & KÜHLBRANDT, W. 2005. pH-induced structural change in a sodium/proton antiporter from Methanococcus jannaschii. *EMBO J,* 24**,** 2720-9.

VOSS, J., HUBBELL, W. L. & KABACK, H. R. 1995. Distance determination in proteins using designed metal ion binding sites and site-directed spin labeling: application to the lactose permease of Escherichia coli. *Proc Natl Acad Sci U S A,* 92**,** 12300-3.

WADE, R. H. 1992. A Brief Look at Imaging and Contrast Transfer. *Ultramicroscopy,* 46**,** 145-156.

WALLIN, E., TSUKIHARA, T., YOSHIKAWA, S., VON HEIJNE, G. & ELOFSSON, A. 1997. Architecture of helix bundle membrane proteins: an analysis of cytochrome c oxidase from bovine mitochondria. *Protein Sci,* 6**,** 808-15.

WALLIN, E. & VON HEIJNE, G. 1998. Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci,* 7**,** 1029-38.

WANG, D. N. & KÜHLBRANDT, W. 1991. High-resolution electron crystallography of light-harvesting chlorophyll a/b-protein complex in three different media. *J Mol Biol,* 217**,** 691-9.

WEEKS, D. L., ESKANDARI, S., SCOTT, D. R. & SACHS, G. 2000. A H+-gated urea channel: the link between Helicobacter pylori urease and gastric colonization. *Science,* 287**,** 482-5.

WEEKS, D. L., GUSHANSKY, G., SCOTT, D. R. & SACHS, G. 2004. Mechanism of proton gating of a urea channel. *Journal of Biological Chemistry,* 279**,** 9944-9950.

WILLIAMS, D. B. & CARTER, C. B. 1996. *Transmission electron microscopy : a textbook for materials science,* New York ; London, Plenum.

WILLIAMS, G. J., BREAZEALE, S. D., RAETZ, C. R. & NAISMITH, J. H. 2005. Structure and function of both domains of ArnA, a dual function decarboxylase and a formyltransferase, involved in 4-amino-4-deoxy-L-arabinose biosynthesis. *J Biol Chem,* 280**,** 23000-8.

WILLIAMS, R. C. & FISHER, H. W. 1970. Electron microscopy of tobacco mosaic virus under conditions of minimal beam exposure. *J Mol Biol,* 52**,** 121-3.

WILSON, S. A., WILLIAMS, R. J., PEARL, L. H. & DREW, R. E. 1995. Identification of two new genes in the Pseudomonas aeruginosa amidase operon, encoding an ATPase (AmiB) and a putative integral membrane protein (AmiS). *J Biol Chem,* 270**,** 18818-24.

YE, L., JIA, Z., JUNG, T. & MALONEY, P. C. 2001. Topology of OxlT, the oxalate transporter of Oxalobacter formigenes, determined by site-directed fluorescence labeling. *J Bacteriol,* 183**,** 2490-6.

YEATES, T. O., KOMIYA, H., REES, D. C., ALLEN, J. P. & FEHER, G. 1987. Structure of the reaction center from Rhodobacter sphaeroides R-26: membrane-protein interactions. *Proc Natl Acad Sci U S A,* 84**,** 6438-42.

ZHAO, G., JOHNSON, M. C., SCHNELL, J. R., KANAOKA, Y., HAASE, W., IRIKURA, D., LAM, B. K. & SCHMIDT-KREY, I. 2009. Two-dimensional crystallization conditions of human leukotriene C(4) synthase requiring adjustment of a particularly large combination of specific parameters. *J Struct Biol*.

ZHU, J., PENCZEK, P. A., SCHRODER, R. & FRANK, J. 1997. Three-dimensional reconstruction with contrast transfer function correction from energy-filtered cryoelectron micrographs: Procedure and application to the 70S Escherichia coli ribosome. *Journal of Structural Biology,* 118**,** 197-219.

Appendix 1

The sequence of the UAC$_{Bc}$ construct used in this work.

```
MNSMGYVGLLLSGAALFLNSLVILGKAEMKSAGVFNLFVGALQIIIPFYLIMISDQS
NWTVYSYAATFLFGLTYLYVGVTFIKGMDSSGLGWFCIWVAIIALFYMVVSFVQFHD
VVNALTWFMWALLWYLFFVLNTQKKNINQYLGRIAFVQSWVTLTLPSLFYFMGVWGE
GFVYELWVYVSVISILYFCYCIYKYRVRSAGGRGSHHHHHH
```

Appendix 2

Crystallisation Buffer	**pH**	**Lipid Type and LPR (w/w)**	**Temperature (ºC)**	**Protein Concentration (mg/ml)**	**Observation under EM**
20 mM K Acetate, 2.5% Glycerol (v/v), 100 mM KCl, 0.05% NaN_3 (w/v)	5.0	N/A 0, DMPC 0.4	25	0.4	LPR 0 and DMPC - Lamella aggregates.
20 mM K Acetate, 2.5% Glycerol, 100 mM NaCl, 0.05% NaN_3	5.0	ETL 0.4, POPC 0.8	25	0.4	ETL and POPC - Lamella aggregates
20 mM MES, 2.5% Glycerol, 100 mM KCl, 0.05% NaN_3	6.0	N/A 0, DMPC 0.4	25	0.4	LPR 0 and DMPC - Lamella aggregates
20 mM MES, 10% Glycerol, 50 mM NaCl, 0.05% NaN_3	6.0	DMPC 0.4, POPC 1.0	25	0.4	DMPC and POPC Lamella aggregates.
20 mM MES, 10% Glycerol, 100 mM KCl, 2.5% NaN_3	6.5	N/A 0, POPC 0.4	25	0.4	LPR 0 - Lamella aggregates . POPC and DOPC - Aggregated vesicles.
20 mM MES, 20% Glycerol, 100 mM KCl, 0.05% NaN_3	6.5	ETL 0.4, DOPC 1.0	25	0.4	ETL and DOPC - Lamella aggregates
20 mM MES, 2.5% Glycerol, 10 mM NaCl, 0.05% NaN_3	6.5	ETL 0.4, POPC 1.0	25	0.4	ETL and POPC - Lamella aggregates
20 mM MES, 20% Glycerol, 10 mM NaCl, 0.05% NaN_3	6.5	ETL 0.4, POPC 0.8	25	0.4	ETL - Aggregated vesicles. POPC - Aggregated vesicles with diffuse diffraction.
20 mM HEPES, 2.5% Glycerol, 100 mM KCl, 0.05% NaN_3	7.0	N/A 0, DMPC 0.4, DMPC 1.0	25	0.4	LPR 0 - Aggregated vesicles. DMPC - Aggregated lipid fragments at LPR 0.4. Aggregated Vesicles at LPR 1
20 mM HEPES, 2.5% Glycerol, 100 mM KCl, 10 mM $MgCl_2$, 0.05% NaN_3	7.0	DMPC 0.4, DMPC 1.0	25	0.4	Aggregated vesicles at LPR 0.4. Larger aggregated vesicles at LPR 1.

20 mM HEPES, 2.5% Glycerol, 50 mM NaCl, 10 mM MgCl$_2$, 0.05% NaN$_3$	7.0	DOPC 0.4	25	0.4	Lamella aggregates with few sheets protruding out.
20 mM HEPES, 10% Glycerol, 100 mM NaCl, 0.05% NaN$_3$	7.0	0.4 POPC, 1.0 DOPC	25	0.4	POPC -Aggregated vesicles. DOPC - Crystals of 250 nm^2 on the edges of aggregated vesicles.
20 mM HEPES, 2.5% Glycerol, 100 mM KCl, 0.05% NaN$_3$	8.0	N/A 0, DMPC 0.4	25	0.4	LPR 0 - Aggregated lipid sheet fragments. DMPC - Aggregated vesicles.
20 mM HEPES, 2.5% Glycerol, 100 mM KCl, 10 mM MgCl2, 0.05% NaN$_3$	8.0	DMPC 0.4	25	0.4	Aggregated vesicles.
20 mM HEPES, 10% Glycerol, 100 mM KCL, 0.05% NaN$_3$	8.0	ETL 0.4	25	0.4	Aggregated vesicles.
20 mM HEPES, 10% Glycerol, 100 mM NaCl, 0.05% NaN$_3$	8.0	DMPC 0.4	25	0.4	Aggregated vesicles

Table 1 Initial Screen

Crystallisation Buffer	**pH**	**Lipid Type and LPR**	**Temperature (°C)**	**Protein Concentration (mg/ml)**	**Observation under EM**
20 mM MES, 100 mM NaCl, 10% Glycerol, 0.05% NaN$_3$	6.5	POPC 0.8	25	0.4	Aggregated vesicles.
20 mM MES, 100 mM NaCl, 5% Glycerol, 0.05% NaN$_3$	6.5	DOPC 0.7	25	0.4	Aggregated vesicles.
20 mM HEPES, 100 mM NaCl, 10% Glycerol, 0.05% NaN$_3$	7.0	POPC 0.6, DOPC 0.7, DOPC 0.8, DMPC 0.9	25	0.4	POPC, DOPC and DMPC - Aggregated vesicles.
20 mM HEPES, 150 mM NaCl, 10% Glycerol, 0.05% NaN$_3$	7.0	DOPC 0.9	25	0.4	Aggregated vesicles.

20 mM HEPES, 200 mM NaCl, 10% Glycerol, 0.05% NaN$_3$	7.0	DOPC 0.6	25	0.4	Aggregated vesicles with some ordering of protein.
20 mM HEPES, 100 mM NaCl, 10% Glycerol, 0.05% NaN$_3$	7.5	DMPC 0.7, DOPC 1.0	25	0.4	DMPC- Aggregated vesicles. DOPC - Aggregated vesicles and crystals.
20 mM HEPES, 100 mM NaCl, 10% Glycerol, 0.05% NaN$_3$	8.0	DOPC 0.7, POPC 1.0	25	0.4	DOPC and POPC - Aggregated vesicles.

Table 2 Optimisation Screen

The screen focused on a smaller pH range and varied the concentration of NaCl and glycerol in the crystallisation buffer.

Crystallisation Buffer	**pH**	**Lipid Type and LPR (w/w)**	**Temperature (ºC)**	**Protein Concentration (mg/ml)**	**Observation under EM**
20 mM HEPES, 100 mM NaCl, 10% Glycerol, 0.05% NaN$_3$	7.5	DOPC 0.6 - 1.10 in steps of 0.02	25	0.4	Aggregated vesicles. Crystals at LPR 0.6 - 0.9.

Table 3 LPR Optimisation Screen

The screen tested small increments in the LPR of DOPC to determine the optimal LPR range for crystallisation.

Crystallisation Buffer	**pH**	**Lipid Type and LPR (w/w)**	**Temperature (ºC)**	**Protein Concentration (mg/ml)**	**Observation under EM**
20 mM HEPES, 100 mM NaCl, 10% Glycerol, 0.05% NaN$_3$	7.5	DOPC 0.75 - 1.0 in 0.05 steps	30	0.4	Aggregated crystals at LPRs 0.75 and 0.8.
20 mM HEPES, 100 mM NaCl, 20% Glycerol, 0.05% NaN$_3$	7.5	DOPC 0.75 - 1.0 in 0.05 steps	30	0.4	Aggregated vesicles.
20 mM HEPES, 200 mM NaCl, 10% Glycerol, 0.05% NaN$_3$	7.5	DOPC 0.8, 0.9, 1.0	30	0.4	Powder diffraction patterns at LPRs 0.8 and 0.9. Vesicles 1 µm^2 in size and above in all trials.

Table 4 Increased temperature screen

The screen increased the dialysis temperature from 25ºC to 30ºC.

Crystallisation Buffer	**pH**	**Lipid Type and LPR (w/w)**	**Temperature (ºC)**	**Protein Concentration (mg/ml)**	**Observation under EM**
20 mM HEPES, 100 mM NaCl, 10% Glycerol, 0.05% NaN_3, 20 mM Urea	7.5	DOPC 1.0, DSPC 0.85	30	0.4	DOPC - Weak 1st order diffraction. DSPC - Aggregated vesicles.
20 mM HEPES, 100 mM NaCl, 10% Glycerol, 0.05% NaN_3, 5 mM EDTA	7.5	DOPC 0.95	30	0.4	DOPC - Aggregated vesicles.
20 mM HEPES, 200 mM NaCl, 10% Glycerol, 0.05% NaN_3, 5 mM EDTA	7.5	DOPC 0.8, DOPC 1.0	30	0.4	Aggregated vesicles.
20 mM HEPES, 200 mM NaCl, 10% Glycerol, 0.05% NaN_3, 5 mM EDTA, 20 mM Urea	7.5	DOPC 0.75, DSPC 0.9	30	0.4	DOPC and DSPC - Aggregated vesicles.
20 mM HEPES, 300 mM NaCl, 10% Glycerol, 0.05% NaN_3	7.5	DOPC 0.8, DOPC 1.0, DSPC 0.85	30	0.4	DOPC 0.8 and 1.0 - Aggregated 2D crystals. DSPC - Aggregated vesicles.
20 mM HEPES, 300 mM NaCl, 10% Glycerol, 0.05% NaN_3, 20 mM Urea	7.5	DOPC 0.8, DSPC 0.85	30	0.4	DOPC - Diffuse 1st order diffraction found in a single vesicle. DSPC - Aggregated vesicles.
20 mM HEPES, 300 mM NaCl, 10% Glycerol, 0.05% NaN_3, 5 mM EDTA	7.5	DOPC 0.8, DSPC 0.75	30	0.4	DOPC and DSPC - Aggregated vesicles.
20 mM HEPES, 500 mM NaCl, 10% Glycerol, 0.05% NaN_3	7.5	DOPC 0.75	30	0.4	Aggregated vesicles
20 mM HEPES, 1M NaCl, 10% Glycerol, 0.05% NaN_3	7.5	DOPC 0.9	30	0.4	Aggregated vesicles

20 mM HEPES, 100 mM NaCl, 10% Glycerol, 0.05% NaN_3	7.5	DOPC 0.9, DSPC 0.9	4	0.4	DOPC and DSPC - Aggregated vesicles.
20 mM HEPES, 100 mM NaCl, 10% Glycerol, 0.05% NaN_3	7.5	DOPC 0.45 - 0.6 in steps of 0.05 and LPR 0.7, 0.8 and 0.9	30	1	Aggregated vesicles.

Table 5 Additional conditions screen
The screen tested DSPC, a similar lipid to DOPC. The screen also introduced EDTA and urea into the dialysis buffers and tested higher increased NaCl concentrations.

Crystallisation Buffer	**pH**	**Lipid Type and LPR (w/w)**	**Temperature (ºC)**	**Protein Concentration (mg/ml)**	**Observation under EM**
20 mM HEPES, 100 mM NaCl, 10% Glycerol, 0.05% NaN_3	7.5	DOPC 0.7 - DOPC 0.96 in steps of 0.02	30	0.4	Aggregated 2D Crystals.
20 mM HEPES, 100 mM NaCl, 10% Glycerol, 0.05% NaN_3	7.5	DOPC 0.5, 0.55, 0.6 - 0.84 in steps of 0.02	30	0.4	Aggregated 2D Crystals.
20 mM HEPES, 100 mM NaCl, 10% Glycerol, 0.05% NaN_3	7.5	DOPC 0.4 - 0.8 in steps of 0.05	30	0.4	Aggregated 2D Crystals.
20 mM HEPES, 100 mM NaCl, 10% Glycerol, 0.05% NaN_3	7.5	DOPC 0.4 - 0.8 in steps of 0.05	30	0.5	Aggregated 2D Crystals.
20 mM HEPES, 100 mM NaCl, 10% Glycerol, 0.05% NaN_3	7.5	DOPC 0.4 - 0.8 in steps of 0.05	30	0.4 and 0.2	Aggregated 2D Crystals.
20 mM HEPES, 100 mM NaCl, 10% Glycerol, 0.05% NaN_3	7.5	DOPC 0.4 - 0.8 I steps of 0.05 (0.55 repeated)	30	0.5	Aggregated 2D Crystals.

20 mM HEPES, 100 mM NaCl, 10% Glycerol, 0.05% NaN_3	7.5	DOPC 0.45, 0.5, 0.55	30	0.6	Aggregated 2D Crystals.
20 mM HEPES, 100 mM NaCl, 10% Glycerol, 0.05% NaN_3	7.5	DOPC 0.45 - 0.7 in steps of 0.025	30	0.5	Aggregated 2D Crystals.
20 mM HEPES, 100 mM NaCl, 10% Glycerol, 0.05% NaN_3	7.5	DOPC 0.45 - 0.7 in steps of 0.05	30	0.5	Aggregated 2D Crystals.

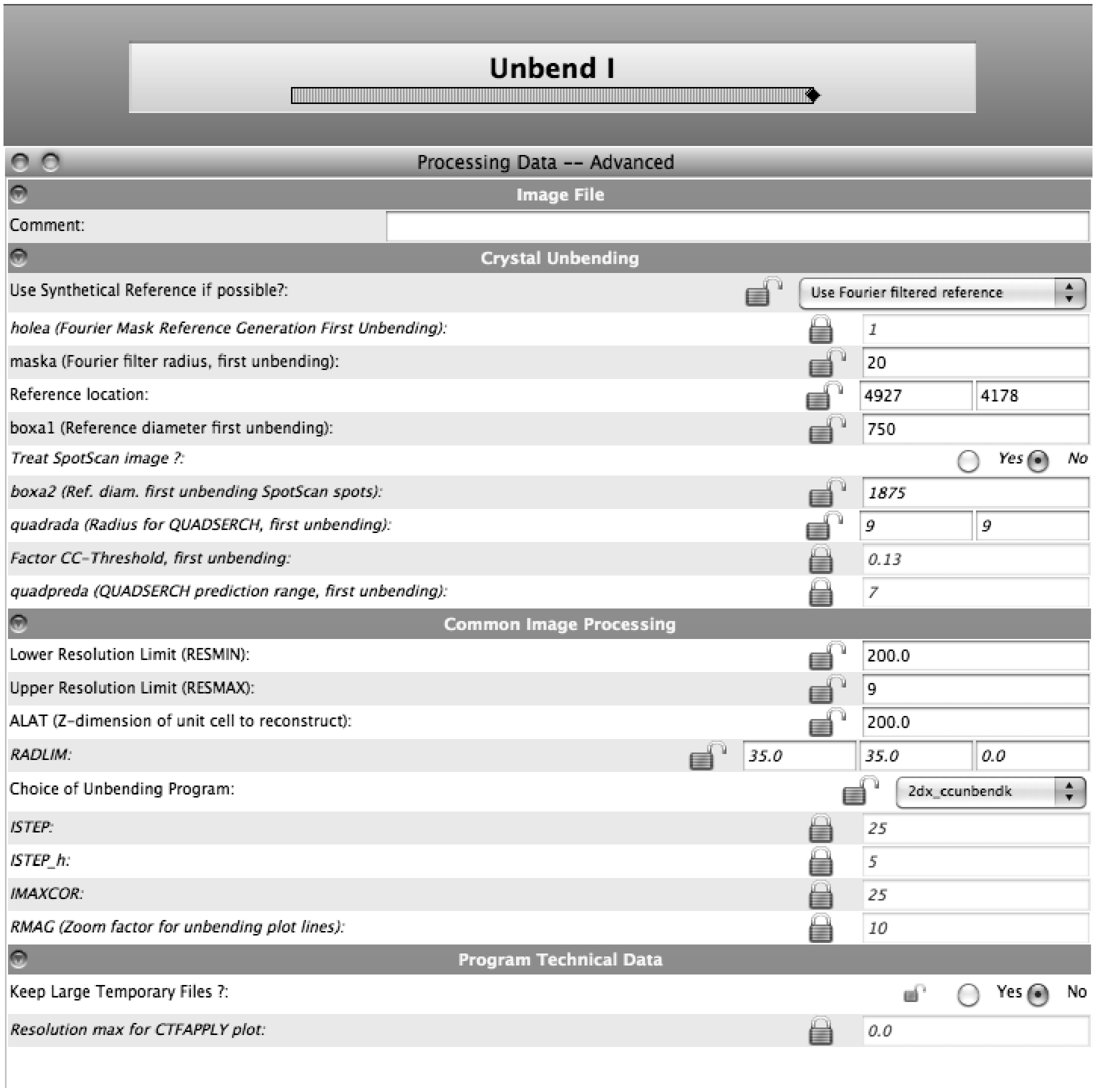
Table 6 Crystals produced for electron cryomicroscopy

Crystallisation Buffer	**pH**	**Lipid Type and LPR (w/w)**	**Temperature (ºC)**	**Protein Concentration (mg/ml)**	**Observation under EM**
20 mM HEPES, 100 mM NaCl, 10% Glycerol, 0.05% NaN_3	7.5	DOPC 0.45 - 0.8 in steps of 0.05	30	0.6	Powder diffraction. Aggregated vesicles.
20 mM HEPES, 100 mM NaCl, 10% Glycerol, 0.05% NaN_3	7.5	DOPC 30, 0.4 - 0.8 in steps of 0.02	30	1	Aggregated vesicles. No diffraction.

Table 7 Conditions tested on the TEV-protease cleaved UAC_{Bc} construct

Appendix 3

Screenshots from *2dx* showing parameters for crystal unbending.


Unbend I

Processing Data -- Advanced

Image File

Comment:	

Crystal Unbending

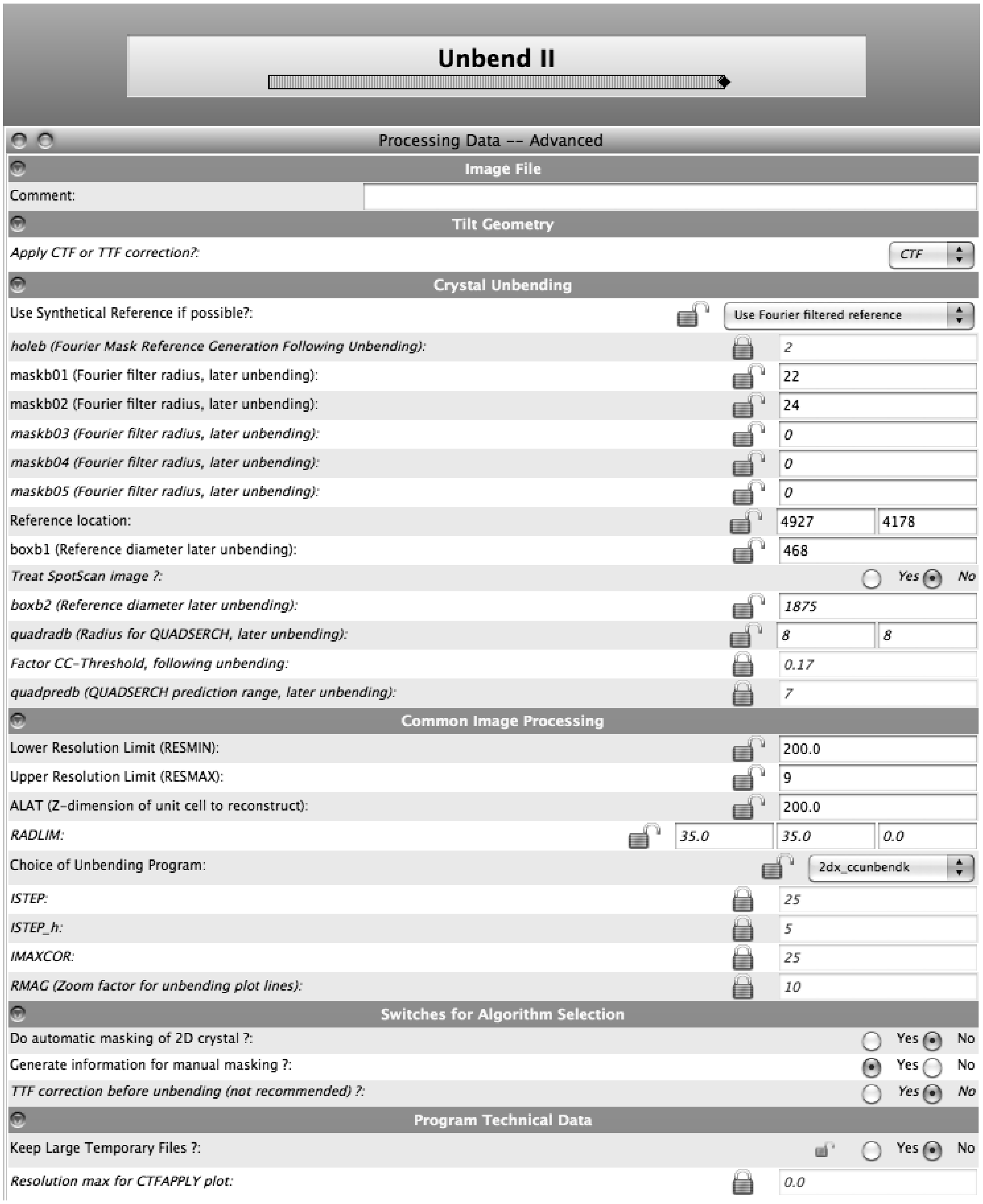
Use Synthetical Reference if possible?:	Use Fourier filtered reference	
holea (Fourier Mask Reference Generation First Unbending):	*1*	
maska (Fourier filter radius, first unbending):	20	
Reference location:	4927	4178
boxa1 (Reference diameter first unbending):	750	
Treat SpotScan image ?:	○ *Yes* ● *No*	
boxa2 (Ref. diam. first unbending SpotScan spots):	*1875*	
quadrada (Radius for QUADSERCH, first unbending):	*9*	*9*
Factor CC-Threshold, first unbending:	*0.13*	
quadpreda (QUADSERCH prediction range, first unbending):	*7*	

Common Image Processing

Lower Resolution Limit (RESMIN):	200.0		
Upper Resolution Limit (RESMAX):	9		
ALAT (Z-dimension of unit cell to reconstruct):	200.0		
RADLIM:	*35.0*	*35.0*	*0.0*
Choice of Unbending Program:	2dx_ccunbendk		
ISTEP:	*25*		
ISTEP_h:	*5*		
IMAXCOR:	*25*		
RMAG (Zoom factor for unbending plot lines):	*10*		

Program Technical Data

Keep Large Temporary Files ?:	○ Yes ● No
Resolution max for CTFAPPLY plot:	*0.0*



Unbend II

Processing Data -- Advanced

Image File

Comment:	

Tilt Geometry

Apply CTF or TTF correction?:	CTF

Crystal Unbending

Use Synthetical Reference if possible?:	Use Fourier filtered reference	
holeb (Fourier Mask Reference Generation Following Unbending):	*2*	
maskb01 (Fourier filter radius, later unbending):	22	
maskb02 (Fourier filter radius, later unbending):	24	
maskb03 (Fourier filter radius, later unbending):	*0*	
maskb04 (Fourier filter radius, later unbending):	*0*	
maskb05 (Fourier filter radius, later unbending):	*0*	
Reference location:	4927	4178
boxb1 (Reference diameter later unbending):	468	
Treat SpotScan image ?:	○ *Yes* ● *No*	
boxb2 (Reference diameter later unbending):	*1875*	
quadradb (Radius for QUADSERCH, later unbending):	*8*	*8*
Factor CC-Threshold, following unbending:	*0.17*	
quadpredb (QUADSERCH prediction range, later unbending):	*7*	

Common Image Processing

Lower Resolution Limit (RESMIN):	200.0		
Upper Resolution Limit (RESMAX):	9		
ALAT (Z-dimension of unit cell to reconstruct):	200.0		
RADLIM:	*35.0*	*35.0*	*0.0*
Choice of Unbending Program:	2dx_ccunbendk		
ISTEP:	*25*		
ISTEP_h:	*5*		
IMAXCOR:	*25*		
RMAG (Zoom factor for unbending plot lines):	*10*		

Switches for Algorithm Selection

Do automatic masking of 2D crystal ?:	○ Yes ● No
Generate information for manual masking ?:	● Yes ○ No
TTF correction before unbending (not recommended) ?:	○ *Yes* ● *No*

Program Technical Data

Keep Large Temporary Files ?:	○ Yes ● No
Resolution max for CTFAPPLY plot:	*0.0*