Applying complementary structural techniques to elucidate structure-function relationships of the bacterial Na^+-hydantoin transporter Mhp1

Anna Polyakova

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
School of Molecular and Cellular Biology
Astbury Centre for Structural Molecular Biology

September, 2015

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

Firstly, I would like to thank my supervisors Arwen Pearson and Peter Henderson, who have guided and encouraged me throughout my PhD. Despite ending up with a split-site supervisory team, I feel I always had their support and got to pursue my ideas.

I would also like to thank everyone who has been part of the Pearson and Henderson groups during the course of this project. Special thanks to David Sharples and Scott Jackson, who helped me to get going in the lab and were very helpful at all times. Many thanks to David for doing the (indispensable) cell fermentations, and Anton Calabrese for performing the mass spectrometry analysis.

I would like to thank Alexander Cameron for guidance and useful discussions on the structure solution.

A considerable amount of the experiments in this Thesis were carried out at various synchrotrons around the world, and without the help of collaborators and beamline staff, these would not have been possible: beamline staff at SSRL BL4-2 (May 2012 and July 2013); beamline staff at the SWING beamline, SOLEIL (February and December 2014); beamline staff at ID29 at the ESRF (December 2015) and beamline staff at I02, I03, I04, I04-1 and I24 at Diamond Lightsource. Special thanks to Juan Sanchez-Whetherby (Diamond) for great help with the HC1, Isabel Moraes from the Membrane Protein Laboratory at Diamond for support during beamtime and Javier Perez (SOLEIL) for help and guidance with SEC-SAXS data collection, processing and modelling.

Special thanks goes to José Marquez and his team (Guillaume, Vincent, Irina) at EMBL Grenoble, who were tremendously helpful with the CrystalDirect studies during my visit to Grenoble and made sure I did not miss my flight back.

A very special thank you to Eddie Snell for letting me join several beamtimes at the SSRL and Diamond, where I have always had a great time! I spent my first time in the states doing SAXS data collection at the SSRL and visiting the HWMI in Buffalo, where I got to learn about SAXS data collection and processing (mainly from Thomas Grant, thank you!), which made the SAXS-part of this project possible.

This project was carried out as part of the 4-year PhD Programme in Structural Molecular Biology within the Astbury Centre at the University of Leeds, funded by the Wellcome Trust (grant number 096683/Z/11/Z).

I would like to especially thank Chi Trinh, who has been around as a helping hand and has been a great friend from the very first time I came to Leeds for my MSc in 2010 (I can not believe I have survived this long…). Becky, Diana, Gaël and Phil, we have spent lots of time together since the start and it has been great having you around through all the times during these past four years (yes, we were definitely the best cohort!). Joe, you helped me to keep my cool throughout the writing of this and I look forward to what lies ahead in the future (although I will never be as cool as you).

Last but definitely not least, I would like to thank my parents Olga and Andy, without them and their support this would not mean anything. I am choosing to dedicate this to my awesome sister Alexandra, who is a great inspiration to me (and I bet I will be reading your thesis in a few years ☺).

Abstract

Secondary-active trans-membrane transporters couple the ‘downhill’ movement of ions with the ‘uphill’ movement of essential substrates, such as neurotransmitters or metabolites, across the cell membrane. Disruption of these vital processes is linked to some severe diseases in humans, for example Parkinson’s disease. The aim of this project is to investigate the structural basis of the transport mechanism of secondary-active transporters, using the bacterial sodium-hydantoin transporter Mhp1 as a model system, with a combination of X-ray crystallography and small-angle X-ray scattering (SAXS).

A wide variety of strategies have been applied and optimised in the crystallisation, harvesting and cryo-protection steps to improve the data obtained from Mhp1 crystals. These include modifications to the crystal drop morphology to improve crystal size and quality, as well as on-line crystal dehydration and contact-less harvesting by photoablation to minimise damage during cryo-protection and harvesting. This has enabled the determination of several novel crystal structures of Mhp1 mutants in complex with various ligands.

SAXS experiments have been used to obtain information about the overall shape and detergent organisation around the purified protein, in order to gain insight into crystallisation propensity and detergent packing in membrane protein crystals. The challenges imposed by the presence of the detergent molecules have been overcome by combining SAXS with size-exclusion chromatography.

The results obtained in this Thesis suggest that the altered ligand binding in the Mhp1 mutants studied in this work is a result of an altered protein conformational landscape, rather than loss of specific protein-ligand interactions. They have also shown that conformational selection during crystallisation can stabilise ligand binding states that are not strongly favoured in solution. Whilst this provides useful detail on ligand binding, it also highlights that a combination of biochemical, biophysical and structural data are required to understand how mutations perturb membrane protein function.

Abbreviations

APC	Amino acid-polyamine-organocation family
BCCT	Betaine/carnitine/choline transporter family
BSA	Bovine serum albumin
BVH	Bromovinylhydantoin
CD	Circular dichroism
CMC	Critical micelle concentration
CV	Column volume
DDM	*n*-Dodecyl-*β*-D-maltopyranoside
DM	*n*-Decyl-*β*-D-maltopyranoside
DMSO	Dimethylsulphoxide
DTT	Dithiothreitol
HC1	Humidity control device 1
5-HIRT	5-helix inverted repeat
L-BH	L-5-benzylhydantoin
L-IMH	L-5-indolylhydantoin
L-NMH	L-5-naphthylmethylhydantoin
MFS	Major facilitator superfamily
MR	Molecular replacement
MWCO	Molecular weight cut off
NCS-1	Nucleobase-cation symporter-1 family
NM	*n*-Nonyl-*β*-D-maltopyranoside
NSS	Neurotransmitter-sodium symporter family
OM	*n*-Octyl-*β*-D-maltopyranoside
PDC	Protein-detergent-complex

PEG	Polyethylene glycol
PEP	Phosphoenolpyruvate
PTS	Phosphotransferase system
RH	Relative humidity
SAXS	Small-angle X-ray scattering
SDS	Sodium dodecyl sulphate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SEC-SAXS	Size-exclusion coupled SAXS
SSS	Solute-sodium symporter family
TCA	Trichloroacetic acid
TMH	Trans-membrane helix
UM	*n*-Undecyl-*β*-D-maltopyranoside

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Membrane transport

1.1.1 Introduction

The membranes surrounding living cells are composed of phospholipid bilayers, which act as protective barriers between the exterior and interior of the cells. Through these membranes, which have a hydrophobic interior, only a few small molecules can passively diffuse at a slow rate. In order to maintain homeostasis, membrane-embedded transport proteins aid the translocation of ions and metabolites through the hydrophobic interior of cell membranes. Membrane transport proteins can be divided into two major classes; channels and transporters, which carry out facilitated diffusion, group translocation, primary or secondary active transport. This section will give an overview of bacterial cell membranes, membrane transport proteins and their role in living organisms.

1.1.2 Bacterial cell membranes

Bacterial cells can be distinguished as Gram-positive or Gram-negative, depending on the structure of their cell membranes. The names stem from the ability of the bacterial cells to be stained by Gram's stain, which was developed in 1884 for the detection of bacteria and infection (Brock, 1961; Austrian, 1960).

Membrane transport proteins are embedded in the so-called cytoplasmic membrane, which separates the inside of the cells (the cytoplasm) from the exterior. In Gram-negative bacteria, an approximately 100 Å wide periplasmic space separates the cytoplasmic membrane from a peptidoglycan layer and an outer membrane that contains porins and is coated with a layer of lipopolysaccharides (see Figure 1.1) (Nicholls & Ferguson, 2014). In Gram-positive bacteria there is no periplasmic space, and the cytoplasmic membrane is in direct contact with a thick peptidoglycan layer.

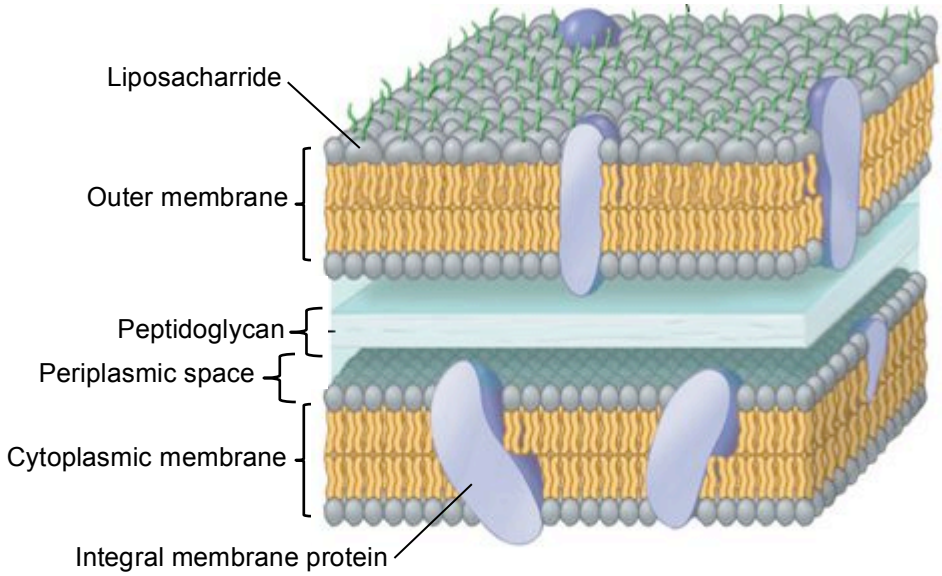


Figure 1.1 The structure of a Gram-negative bacterial cell membrane.

The membrane layers surrounding Gram-negative bacteria that separates the cytoplasm from the outside of the cell are composed of a cytoplasmic membrane and a peptidoglycan layer which are covered by an outer membrane. The cytoplasmic membrane holds integral membrane proteins, such as various transport proteins that carry out facilitated diffusion, group translocation, primary or secondary-active transport. Figure from Campbell Biology © Pearson Education, 2011.

1.1.3 Membrane proteins

Integral membrane proteins residing in the bacterial cytoplasmic membrane are responsible for the translocation of ions, metabolites and compounds involved in signalling pathways between the cytoplasm and the outside of the bacterial cell. Membrane proteins can be divided into groups that carry out passive diffusion, facilitated diffusion, group translocation, primary or secondary active transport and these will be discussed further in this section.

1.1.3.1 Passive diffusion

Typically, small and neutral molecules that are soluble in a hydrophobic environment, such as O_2, CO_2, NH_3, HCO_2H or glycerol, are able to diffuse freely across the phospholipid membrane (Henderson, 2012). This is termed passive diffusion, and occurs down the concentration gradients of the molecules without the aid of transport proteins.

1.1.3.2 Facilitated diffusion

Facilitated diffusion by transport proteins is required to transport hydrophilic and/or larger substrate molecules across the hydrophobic membrane, or for rapid responses to changes in the concentration of water or ions. As for passive diffusion, this occurs down the concentration gradient of the substrate molecules.

Channels facilitate the diffusion of molecules and ions across the membrane by forming pathways through the membrane, with a selectivity filter for the substrate molecule. Substrate molecules can approach openings of the channel from either side of the membrane but will only flow down their concentration gradient (Figure 1.2). Channels can be permanent pores through the membrane, as for example aquaporins in the case of water molecules (King *et al*, 2004). Channels may also be regulated, as for example the voltage-gated K^+-channels, where the pore is opened or closed in response to changes in the membrane potential (Sigworth, 1994).

Similarly to channels, there are also facilitated diffusion transporters (also termed uniporters) that also move substrates across the cell membrane down their concentration gradients but in a different fashion and at a slower rate (Figure 1.2) (Gadsby, 2009).

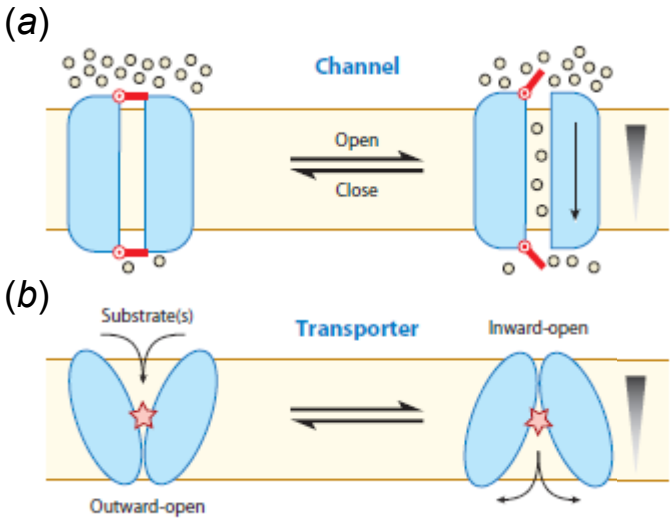


Figure 1.2 Illustration of facilitated diffusion carried out by a channel and a transporter.

(*a*) A channel forms a pore through the membrane, allowing substrate molecules to flow through down their concentration gradients (indicated by the grey triangles) when the pore is open. (*b*) In contrast, a transporter undergoes conformational changes instead, transporting one substrate molecule at a time across the membrane.

Instead of providing a continuous pathway through the membrane, transporters undergo substantial conformational changes which arise from binding of the substrate molecules at their binding sites. Transporters therefore translocate one substrate molecule per turnover cycle (together with co-transported molecules or ions), making the rate of transport slower than through channels.

1.1.3.3 Group translocation

Group translocation, or the phosphotransferase system (PTS), is another type of active transport system in bacteria that uses phosphoenolpyruvole (PEP), an intermediate in glycolysis and gluconeogenesis, as an energy source to transport sugar molecules into the cytoplasm of bacterial cells (Figure 1.3). This transport system is a multi-enzyme complex that transfers sugars such as glucose, mannose or mannitol to the inside of the cells. This is driven by the transfer of a high-energy phosphate group from PEP to members of the multi-enzyme complex of the PTS, to finally end up on the transported sugar molecule (Bramley & Kornberg, 1987).

1.1.3.4 Primary active transport

As opposed to facilitated diffusion, primary active transporters use energy to transport substrate molecules across cell membranes against their concentration gradients. One example of primary active transporters are those involved in photosynthesis or respiration. In photosynthesis, proteins of the photosynthetic electron transport chain use light as their source of energy. The light initiates transport of electrons through the proteins, which in turn drives the pumping of protons across the thylakoid membrane of plant cells or the cell membrane of photosynthetic bacteria (Figure 1.3) (Raven *et al*, 2005; Bryant & Frigaard, 2006).

Respiration in aerobic prokaryotes and eukaryotes is driven by redox energy in the form of electrons stored in electron donors such as NADH. In these cases, proteins of the electron transport chain pump protons across the cell membrane of prokaryotes, or across the inner membrane of mitochondria in eukaryotes, while electrons obtained from NADH are incorporated into electron acceptors such as O_2 (Gautheron, 1984).

Consequently, photosynthesis and respiration creates a proton gradient across cell membranes in prokaryotes or inner mitochondrial membranes in

eukaryotes. The proton gradient drives ATP synthase to produce ATP, which is another primary energy source. ATP-driven primary active transporters, such as ATP-dependent synthases (Lodish *et al*, 2000) or ABC uptake/efflux systems (Reyes *et al*, 2006) couple the energy obtained by hydrolysing ATP into ADP and P_i to drive the movement of ions, nutrients or toxins against their concentration gradients (Figure 1.3).

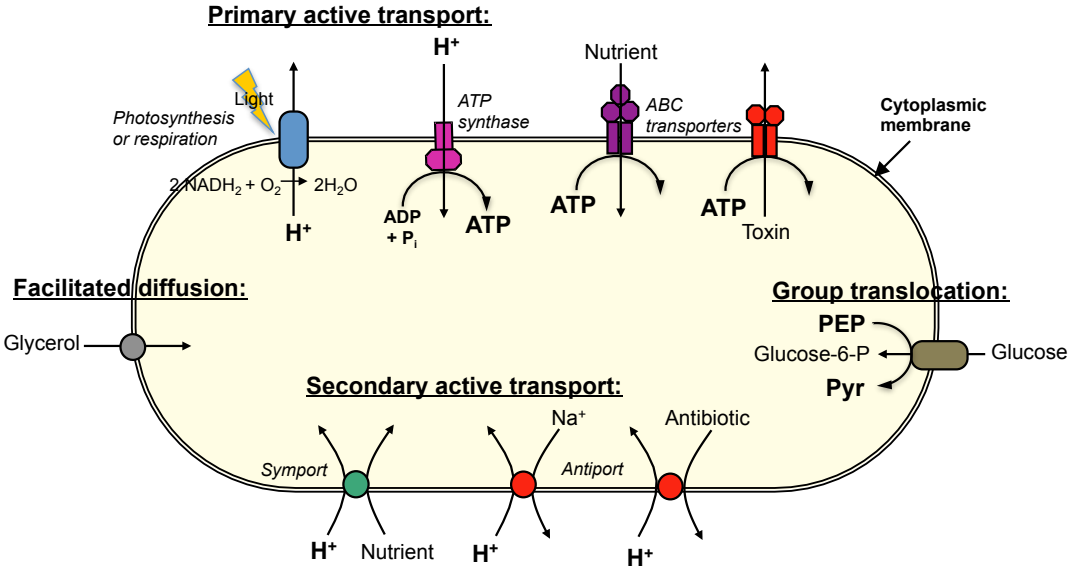


Figure 1.3 Summary of the different bacterial membrane transport proteins.

The yellow oval represents the cytoplasmic membrane of a bacterial cell. Primary active transport involves photosynthetic or respiratory proteins (blue oval), which use light or electron donors as an energy source to create a proton gradient across the membrane. The proton gradient drives the production of ATP by ATP synthase (pink), which in turn fuels ATP-driven ABC uptake or efflux systems (purple and red, respectively). The proton gradient is also a driving force for some secondary active transporters, which co-transport substrates with protons or ions. Secondary active transporters can carry out translocation of the proton/ion and the co-transported substrate to the same side of the membrane (symporter, green circle) or to opposite sides of the membrane (antiporter, red circles). Group translocation is carried out by phosphotransferase systems, which utilise the energy released during the transfer of phosphate groups to translocate sugar molecules into the cytoplasm (brown oval). Facilitated diffusion does not use any energy, and only transports substrates down their concentration gradients (grey circle).

1.1.3.5 Secondary active transport

When performing their action, some primary active transporters generate an electrochemical gradient across membranes. For example, the ATP-driven

Na^+/K^+-pumps maintain ion gradients of higher extra- and intracellular concentrations of sodium and potassium, respectively (Gadsby *et al*, 2012). Meanwhile, photosynthetic and respiratory chain proteins in the prokaryote cell membrane generate a proton gradient by pumping protons across the membrane during electron transport (Bryant & Frigaard, 2006). A different family of transporters, termed secondary active transporters, utilise the energy stored in these electrochemical gradients for co-transport of substrates across the membrane (Figure 1.4) (Forrest *et al*, 2011).

1.2 Secondary active membrane transporter proteins

1.2.1 Introduction

Secondary active transporters constitute a very important class of proteins found in bacteria, archaea, yeast, fungi, plants and animals. This type of transporters utilise the energy stored in membrane proton or ion gradients generated by primary active transporters. The energy stored in these concentration gradients is used to drive the co-transport of a second substrate, such as vitamins, neurotransmitters or metabolites.

Secondary active transporters are classified into antiporters, symporters and uniporters. Anti- and symporters couple the energy obtained by moving ions down their concentration gradient to the co-transport of another substrate against its concentration gradient, *i.e*. they translocate two or more molecules across the membrane in each transport cycle. Uniporters on the other hand transport only one substrate molecule, and are believed to have lost their coupling ability through evolution.

The transport is proposed to occur through a so-called alternating access mechanism. The alternating access mechanism was originally suggested in the 1960s (Jardetzky, 1966) and has been supported by subsequent biochemical and structural studies (Forrest *et al*, 2011). The transport model involves the protein alternating between two main conformational states, where the substrate binding site is accessible to the solvent on either the extra- or intracellular side of the membrane for release or uptake of substrates. The transition between these two states is thought to occur through an occluded state, where the substrates are shielded from either side of the membrane. Some important superfamilies of secondary active transport proteins include the major facilitator superfamily (MFS) and the LeuT superfamily, discussed in more detail below.

1.2.2 Major facilitator superfamily

The major facilitator superfamily (MFS) is present in all kingdoms of life and includes physiologically very important proteins. Previous crystal structures of MFS superfamily members include a broad repertoire of bacterial transporters, such as the lactose:H^+ symporter LacY (Abramson *et al*, 2003; Guan *et al*, 2007), the glycerol-3-phosphate:phosphate antiporter GlpT (Huang *et al*, 2003) and the D-xylose:H^+ symporter XylE (Sun *et al*, 2012). This superfamily also includes the eukaryotic glucose transporters GLUT1-4, which are of great medical significance for humans. Until recently, there were no crystal structures available of any eukaryotic GLUT transporter. The structure of GLUT1, the dysfunction or overexpression of which leads to GLUT1 deficiency syndrome or cancer (Seidner *et al*, 1998) was recently reported (Seidner *et al*, 1998; Deng *et al*, 2014).

1.2.3 LeuT superfamily

An important group of secondary active transporters is the LeuT transporter superfamily, also interchangeably termed the 5-helix inverted repeat (5-HIRT) transporter superfamily (Krishnamurthy *et al*, 2009; Adelman *et al*, 2011).

The LeuT superfamily includes the nucleobase-cation symporter (NCS-1) family, solute-sodium symporter (SSS) family and neurotransmitter-sodium symporter (NSS). In addition, it also includes the more distantly related betaine/carnitine/choline transporter (BCCT) family and the amino acid-polyamine-organocation (APC) family (Figure 1.4).

1.2.3.1 Structural fold of LeuT superfamily members

Rather than being related by sequence identity or common function, the members of the LeuT superfamily are related by a common structural fold – a 5-helix inverted repeat (Figure 1.5). This fold comprises ten trans-membrane helices (TMHs), which are arranged as two inverted repeats of five TMHs each (helices 1-5 and 6-10, respectively). TMH1 and TMH6 consist of two half-helices, and the middle-region between them provides the site of substrate binding. The two repeats are related to each other by a 180° rotation in the putative plane of the membrane.

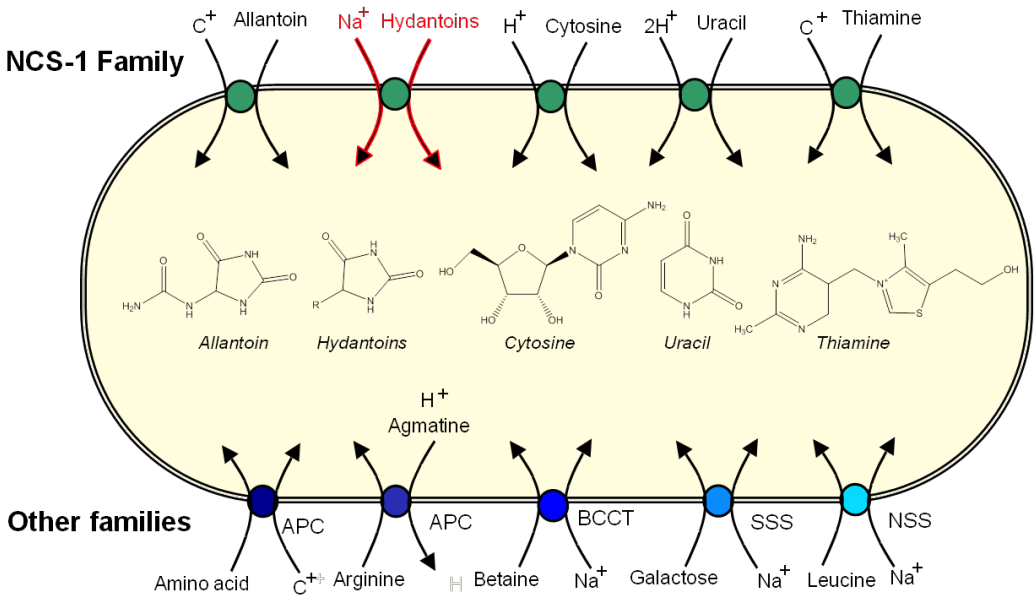


Figure 1.4 The secondary-active transporters belonging to the LeuT superfamily.

The green circles represent members of the NCS-1 family, with the chemical structure of their respective substrates drawn in the middle of the yellow oval (which represents the cytoplasmic membrane of a bacterial cell). The blue circles represent members of other families; including the APC, BCCT, SSS and NSS families (initial figure provided by Peter Henderson).

Crystal structures of proteins belonging to the LeuT superfamily that have previously been determined include, for example, the sodium-leucine transporter LeuT, a member of the NSS family (Yamashita *et al*, 2005). LeuT is a bacterial homologue of the mammalian sodium-serotonin transporter, which has been implicated in depressive disorders (Murphy *et al*, 2004). LeuT has subsequently been a target for rational design of anti-depressant compounds (Zhou *et al*, 2009). Another example is the homologue of the mammalian sodium-glucose transporter; vSGLT from *V. parahaemolyticus*, a member of the SSS family (Faham *et al*, 2008), which has been implicated in kidney disorders (Wright *et al*, 2007). Other examples include the betaine transporter BetP, which belongs to the BCCT family (Ressl *et al*, 2009), and the arginine-agmatine antiporter from *E. coli*, a member of the APC family (Fang *et al*, 2009). The LeuT superfamily also includes the bacterial sodium-hydantoin transporter Mhp1, a member of the NCS-1 family, which is the focus of this Thesis and will be described in more detail further into this Chapter.

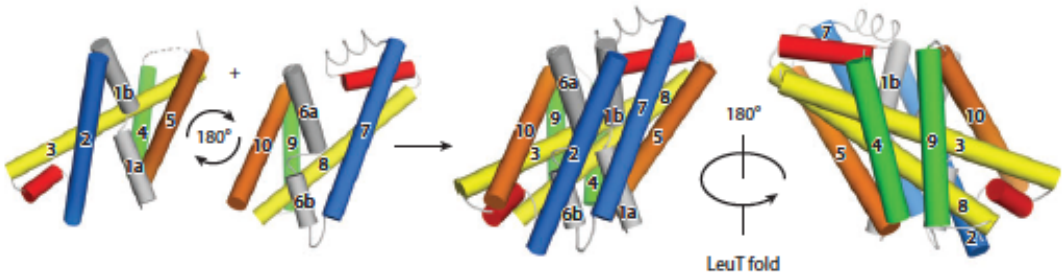


Figure 1.5 Illustration of the overall fold of LeuT superfamily members.

Illustration of the LeuT fold, with two inverted segments of five TMHs (helices 1-5, and 6-10) that are related to each other by a 180° rotation in the plane of the membrane. TMH1 and TMH6 (grey) consist of two half-helices, between which substrate binding occurs. The figure is adapted from Shi et al. 2013.

Overall, the LeuT superfamily includes proteins that play an important role in health and disease in humans, but there is as yet no crystal structure available of any of the eukaryotic members of this superfamily. The already available crystal structures of the bacterial transporters have been determined in various intermediate states along the proposed common alternating access transport mechanism. The conformations observed are consistent with the conformational states discussed in 1.2.1 (outward-open, occluded and inward-open) and have helped to further understand the molecular basis behind the alternating access mechanism in secondary active transport proteins (Forrest *et al*, 2011; Krishnamurthy *et al*, 2009; Shi, 2013). LeuT has been crystallised in an occluded form with leucine bound (Yamashita *et al*, 2005), and has also been trapped in the outward-open and inward-open substrate-free forms in complex with Fab fragment (Krishnamurthy & Gouaux, 2012). vSGLT has been crystallised in an inward-open conformation, with and without galactose bound (Faham *et al*, 2008; Watanabe *et al*, 2011). More recently, crystal strctures of BetP in an outward-open state both in the presence and absence of betaine (Perez *et al*, 2013), as well as the inward-open form have been obtained (Koshy *et al*, 2013). Mhp1, which is the focus of this Thesis, was the first protein for which crystal structures were obtained in all three conformational states (Weyand *et al*, 2008; Shimamura *et al*, 2010) and these will be described in more detail in 1.3.2.

1.2.4 The nucleobase-cation-symporter 1 family

The secondary-active symporter Mhp1 from *Microbacterium liquefaciens*. Mhp1 belongs to the NCS-1 family, which includes approximately 1000 proteins found in bacteria, archaea, fungi and plants. The members of the NSC-1 family and their respective substrates are outlined in Figure 1.4, and carry out the following general symport mechanism across the membrane:

Nucleobase/vitamin (out) + cation (out) → Nucleobase/vitamin (in) + cation (in)

Essentially, the role of NCS-1 transporters is the uptake of organic molecules such as allantoin, hydantoin compounds, cytosine, uracil or thiamine (see Figure 1.4) into the cytoplasm of the cells. The organic substrates are transported by NCS-1 members to the cytoplasm together with either sodium ions or protons (Weyand *et al*, 2006). The transported compounds then act as precursors in nucleotide or amino acid salvage pathways, where, for example, nucleotides are used as precursors in the synthesis of DNA, RNA, ATP, GTP and NAD^+ inside the cells, and may also act as nitrogen sources in bacteria and fungi (de Koning & Diallinas, 2000).

1.3 The bacterial sodium-hydantoin transporter Mhp1

1.3.1 Introduction

The bacterial symporter Mhp1 from *Microbacterium liquefaciens* couples the transport of sodium ions down their concentration gradients with the symport of 5-substituted hydantoin compounds to the cytoplasm of bacterial cells. Inside the cells, the hydantoin compounds are hydrolysed and converted into optically pure amino acids (Figure 1.6) (Suzuki & Henderson, 2006).

Mhp1 was originally discovered as part of the so-called *hyu* gene cluster, which also includes enzymes that participate in the production of optically pure amino acids (Suzuki & Henderson, 2006). Optically pure amino acids are of great interest in industrial applications, where they are used as precursors for the production of, for example, antibiotics, pesticides, pharmaceuticals and food additives (Suzuki & Henderson, 2006; Clemente-Jiménez *et al*, 2008), which is what initially generated interest in studying the *hyu* gene cluster, containing the Mhp1 protein.

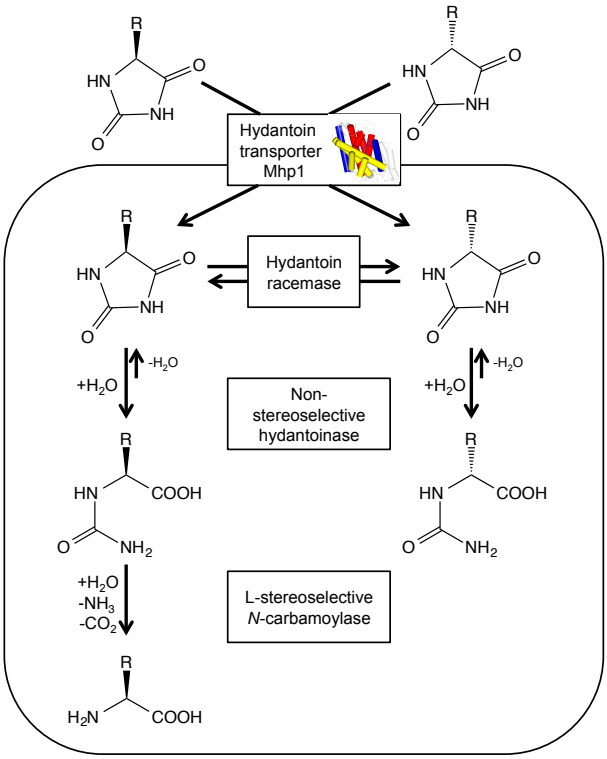


Figure 1.6 Conversion of hydantoin compounds to optically pure amino acids by the gene cluster containing the Mhp1 gene.

Mhp1 transports hydantoin compounds into the cytoplasm of bacterial cells, where the they are broken down into L-amino acids by other members of the *hyu* gene cluster. Figure adapted from Suzuki and Henderson, 2006.

1.3.2 Previously published crystal structures of the Mhp1 protein

Mhp1 was the first transporter of the LeuT superfamily for which crystal structures were obtained in all the three key states of the proposed alternating access mechanism (Weyand *et al*, 2008; Shimamura *et al*, 2010; Diallinas, 2008). The overall fold is comprised by twelve TMHs, as for other LeuT superfamily members. The ligand- and sodium-binding sites are nested between two sets of four helices, which are termed the bundle- (TMH1, 2, 6 and 7) and hash-motifs (TMH3, 4, 8 and 9). These two motifs are flanked by two mobile helices, TMH5 and TMH10, which undergo conformational changes during the transport mechanism (Figure 1.7) (Diallinas, 2008; Weyand *et al*, 2011).

The outward-open state (numbered III in Figure 1.7) has an intact sodium binding site, where a sodium atom can be modelled, and the ligand binding site is open to the periplasmic side of the protein, ready to accept hydantoin compounds. The occluded state (numbered V in Figure 1.7) also has an intact sodium binding site where a sodium atom can be modelled, and L-benzylhydantoin (L-BH) bound in the ligand binding site.

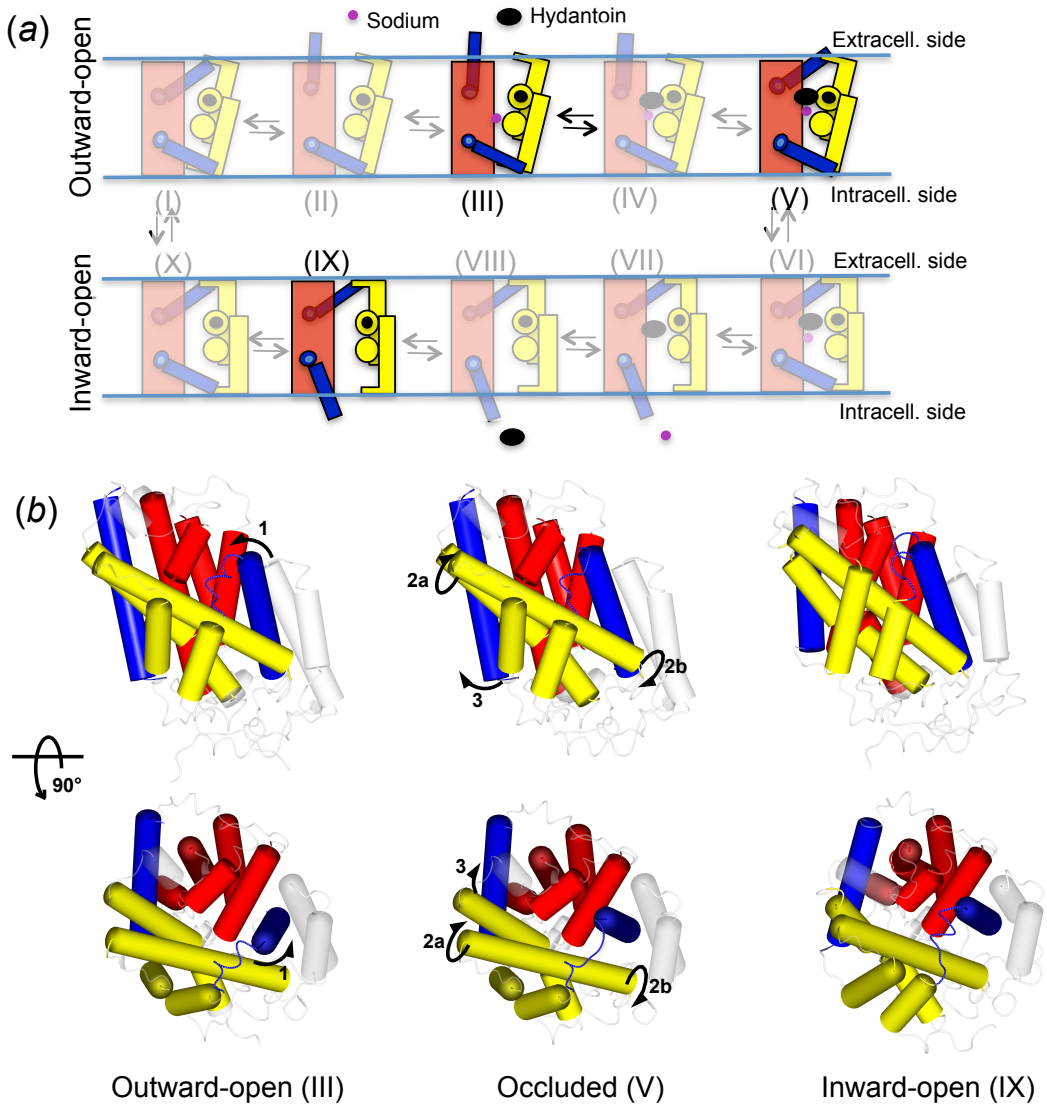


Figure 1.7 The transport mechanism in Mhp1.

The proposed reaction cycle of Mhp1 (original figure provided by Peter Henderson) is illustrated as a simplified cartoon (*a*). The intermediate states for which crystal structures have been obtained are highlighted. In accordance with the alternating access model, the protein alternates between an outward-open state (II), where it accepts sodium and hydantoin compounds from the extracellular side of the membrane (III-IV) and an inward-open state (IX), which is adopted after substrates have been released into the inside of the cell (VII-VIII). The transition between these two states is thought to go through an occluded state (V-VI), where the bound substrates are shielded from both sides of the membrane but with either the outward (V) or inward gate (VI) remaining partially open. The three previously published crystal structures of Mhp1 in key conformational states (outward-open, III; occluded, V; and inward-open, IX) are illustrated with the key α-helices that undergo conformational changes highlighted (*b*). Upon ligand binding during the transition between the outward-open to occluded states, TMH10 shifts to close the ligand binding site as a lid (arrow 1). Upon the subsequent transition to the inward-open state, the hash-motif undergoes a rotation with respect to the bundle-motif (arrows 2a-b), as well as an opening of the ligand binding cleft by a shift of TMH5 (arrow 3). The colour scheme is the same for both panels, representing the bundle- and hash-motifs (red and yellow, respectively), as well as the flexible TMH5 and TMH10 (blue).

The outward-open (III) and occluded (V) states differ in that in the occluded state, TMH10 shifts by approximately 7.5 Å to close up on the entry route to the ligand binding site like a lid, shielding the bound ligand from the outside of the protein.

The last key conformational state for which a crystal structure of Mhp1 has been obtained is the inward-open state (numbered IX in Figure 1.7). In the crystal structure of the inward-open state, no ligands are bound, the sodium binding site is no longer intact and the ligand binding site is open to the cytoplasmic side of the protein. Going from the occluded form (V) to the inward-open form (IX), the hash-motif undergoes an approximately 30° rotation with respect to the bundle-motif, as well as a shift of TMH5, opening the ligand binding cleft to the intracellular side of the protein.

The availability of this structural information makes Mhp1 an ideal model system for further structural studies to investigate what occurs in the transitions between the key states and confirm their sequence.

A sodium ion can be modelled into the intact sodium-binding site in both the structure of Mhp1 in an outward-open (III), as well as in that of the occluded state (V) with L-BH bound. The synergistic binding of the sodium and L-BH has been demonstrated using a tryptophan fluorescence quenching assay (Figure 1.8*a*) (Weyand *et al*, 2008).

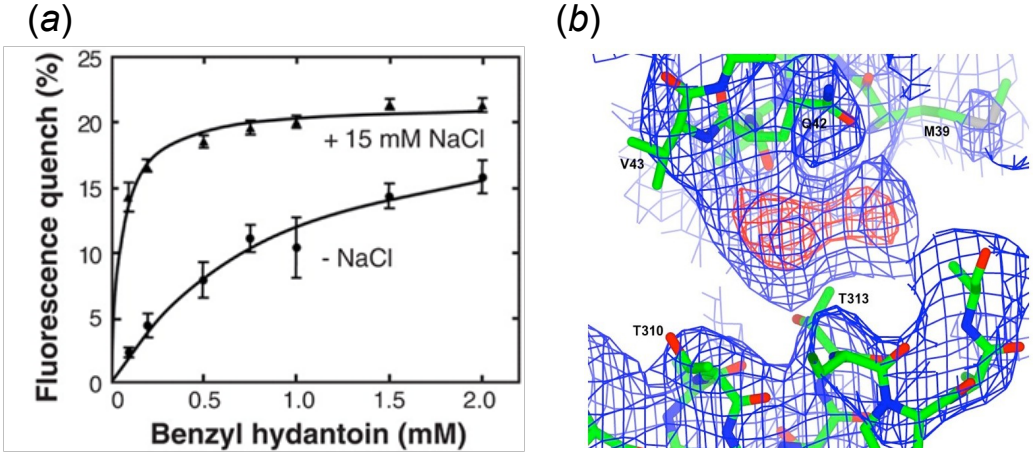


Figure 1.8 Sodium-dependent binding in Mhp1 and the ligand-binding site in the inward-open form of Mhp1.

(*a*) Tryptophan fluorescence quenching data shows that the affinity for L-BH in Mhp1 increases with added NaCl. (*b*) Image of the electron density at 3.8 Å in the sodium binding site ($2F_{obs}$-F_{calc}, coloured blue and contoured at 1σ) in the inward-open form, with the additional density (F_{obs}-F_{calc}, coloured red contoured at 6σ) which has not yet been identified. Figure adapted from Shimamura *et al.* (2010).

In contrast, the sodium binding site is not intact in the crystal structure of the inward open form, which was determined from crystals of seleno-L-methionine derivatised Mhp1 (Shimamura *et al*, 2010). In this crystal structure, TM helices 1 and 8 have moved apart, separating the residues that bind the sodium ion and thus disrupting its binding site. Molecular dynamics simulations have been carried out to investigate whether the inward-open form does indeed perturb the sodium binding site or if it is still capable of binding sodium in this conformational state. Simulations of the three states in a native-like membrane showed that sodium escapes very quickly (<2ns) from its binding site in the inward-open form while it remains bound in the outward-open and occluded forms (Shimamura *et al*, 2010). However, the ligand binding site in the crystal structure of Mhp1 in the inward-open form was found to be occupied by a portion of positive difference electron density, indicating that something is bound and it may be that this potentially stabilises the protein in the inward-open state (Figure 1.8*b*). However, the electron density is not consistent with any components of the crystallisation mixture and this putative ligand has not yet been identified.

1.3.3 Insights into the ligand-binding step of Mhp1

Apart from the crystal structures of Mhp1 in the three key conformational states discussed in 1.3.2, a number of additional Mhp1-ligand complex structures have been published in a recent study, giving further insights into substrate recognition and ligand binding in the protein (Simmons *et al*, 2014). The additional Mhp1-ligand complex structures include the protein with L-indolylmethylhydantoin (L-IMH), bromovinylhydantoin (BVH) and L-naphthylmethylhydantoin (L-NMH) bound. In addition, a revised structure of Mhp1 in complex with L-BH was also published, revealing a new binding mode of L-BH. In the previous structure of the Mhp1 protein, in an occluded state at 4 Å resolution in complex with L-BH (1.3.2, Weyand et al. 2008), the L-BH was modelled in the same conformation as the ligand adopts in solution (Delgado *et al*, 2007). In the new, improved structure of the same ligand complex at resolution 3.8 Å, the L-BH was modelled in a more extended configuration, which was derived from the structure of Mhp1 in complex with L-IMH. The structure of Mhp1 in complex with BVH helped to confirm the orientation and binding mode of the hydantoin ligands, as the bromine scatters anomalously and could be clearly seen in an anomalous map. In the crystal structure of Mhp1 in complex with the Mhp1 inhibitor L-

NMH, the protein also adopted an occluded conformation but differed from the other protein-ligand structures in that TMH10 adopted a conformation similar to that in the structure of Mhp1 in its outward-open conformation. It appears as if L-NMH prevents TMH10 from closing up around the ligand binding site. This suggests that the occlusion of TMH10 is required for the protein to turn over the transport of the ligand.

1.4 Challenges in structural studies of membrane transport proteins

1.4.1 Introduction

Membrane proteins constitute up to 30% of human genes (Wallin & Heijne, 1998) and are targeted by more than 50% of commercially available drugs (Arinaminpathy *et al*, 2009; Terstappen & Reggiani, 2001). Membrane proteins are subsequently high-priority targets for structural studies. Yet, they constitute only 1.9%[1] of the structures published in the Protein Data Bank (PDB). This stems from the fact that a great portion of the surface area of membrane proteins is hydrophobic, as they are embedded in a hydrophobic lipid bilayer in the cell. This makes membrane proteins very challenging targets to study, as most biochemical and structural assays are carried out in an aqueous solution. This requires the presence of, for example, detergents in order to solubilise and maintain the fold and function of membrane proteins in solution. The work presented in this Thesis was carried out on the Mhp1 protein solubilised in detergents. The presence of the detergents presented challenges in the experiments carried out, which will be discussed in subsequent Chapters. This section will give an introduction to the complications and hurdles in structural studies of membrane proteins.

1.4.2 Expression of membrane proteins

The primary problem when working with membrane proteins is to obtain sufficient qualities of the protein of interest. Membrane proteins often exist at very low levels in their native hosts and therefore strategies for heterologous

---

[1] 1768 coordinate files of membrane proteins are reported by the Membrane Protein Structures Database (http://blanco.biomol.uci.edu/mpstruc/, updated 15th September 2015), out of 93594 protein crystal structures in the PDB (19th September 2015).

expression are required in order to produce sufficient quantities of protein for biochemical and structural studies. However, heterologous expression of membrane proteins in for example *E. coli* often causes aggregation of the protein in the expression host cytoplasm. In some cases, the expression host also lacks the machinery for post-translational modifications required to produce the membrane protein of interest in a fully functional state (Drew *et al*, 2003; Wagner *et al*, 2006).

1.4.3 Solubilisation of membrane proteins

When sufficient over-expression of the membrane protein of interest has been achieved, another major problem arises, which is the requirement of a method to extract and contain the intact, folded hydrophobic membrane proteins in all steps throughout the purification process and in subsequent biochemical and structural studies. There are several strategies for the solubilisation of membrane proteins out of their native hydrophobic bilayer, and finding the optimal strategy for the protein of interest may require extensive trial-and-error screening and optimisation. The development of high-throughput screens for the solubility of membrane proteins has made this step more efficient (Gutmann *et al*, 2007; Eshaghi *et al*, 2005).

The main strategy of solubilising membrane proteins out of their native lipid bilayer is the use of detergents (Seddon *et al*, 2004; Privé, 2007). Detergents are amphiphatic molecules, which means that they have both a hydrophilic (a polar head-group) and a hydrophobic part (hydrophobic alkyl chain). Every detergent exhibits an individual critical micelle concentration (CMC), which depends on structural properties such as the alkyl chain length and charge of the head-group. The CMC may in turn also vary with pH, ionic strength and temperature of the solution. Below the CMC, the detergent molecules exist as monomers. When the concentration is at the CMC or above the detergent monomers spontaneously aggregate into micelles through the entropy-driven hydrophobic effect (Tanford, 1973). Membrane proteins can be extracted from their native membrane environment by mixing membrane fragments with a detergent solution where the detergent concentration is above its CMC. This results in the formation of protein-detergent complexes (PDCs), driven by the hydrophobic effect, as well as free micelles and detergent monomers, which is illustrated in Figure 1.9 (Duquesne & Sturgis, 2010).

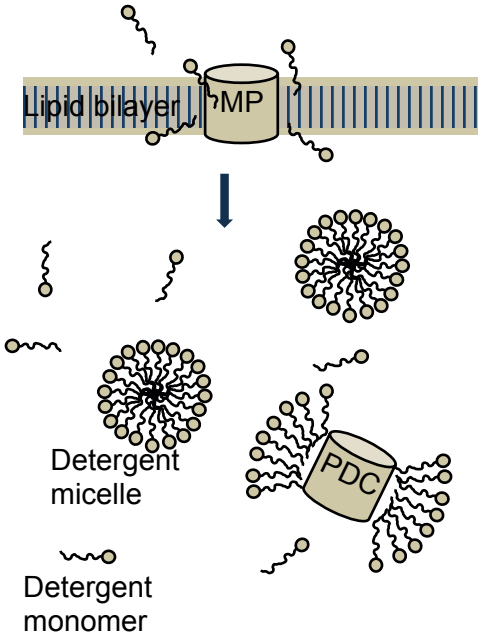


Figure 1.9 Solubilisation of membrane proteins using detergents.

Membrane proteins (MP) can be extracted from the lipid bilayer by mixing with detergents. Above the CMC of the detergent, protein-detergent complexes (PDCs) will form, as well as protein-free detergent micelles.

1.4.4 X-ray crystallographic studies of membrane proteins

Obtaining a crystal structure using X-ray crystallography may be seen as the ultimate experiment in the study of protein structure, as the technique carries the potential of yielding structural information at near-atomic resolution. The main bottleneck still remains the step of obtaining well-diffracting crystals.

This process has been accelerated with the development and availability of high-throughput screens and screening robots, but there is, as yet, no rational method for predicting protein crystallisation conditions. As described in section 1.4.3, the hydrophobic nature of membrane proteins requires the presence of a solubilising agent, such as detergents. This presents additional challenges in the crystallisation of membrane proteins.

The basic principle of protein crystallisation is to take a protein solution to a supersaturated state, where the protein molecules are at a concentration that exceeds their solubility limit, while still remaining in solution. The supersaturation region in protein crystallisation is divided into three regions; the metastable zone, the nucleation zone and the precipitation zone (Figure 1.10). In a successful crystallisation experiment, the concentrations of protein and a precipitating agent are altered to bring the protein into the

nucleation zone, where crystal nuclei form that then allow the formation of protein crystals (Manuel García-Ruiz, 2003). Precipitating agents, such as salts, polymers or organic solvents affect chemical and physical properties of the protein and/or solvent, so as to facilitate protein association and crystallisation (McPherson, 2011).

This is achieved by, for example, the vapour-diffusion technique (Sutton & Sohi, 1994), dialysis (Fromme, 2003) or micro-batch methods (Loll *et al*, 2003; Chayen, 2003). Another technique for protein crystallisation is the use of counter-diffusion, which will be further described in section 4.4.1.

The vapour-diffusion technique involves mixing a droplet of protein with a reservoir solution containing precipitating agents. The droplet is suspended over the reservoir solution (*e.g.* in a sitting- or hanging-drop configuration) and allowed to equilibrate in a sealed chamber. Since the droplet will contain a lower concentration of precipitants than the reservoir solution, net vapour-diffusion will occur from the droplet to the reservoir until equal vapour-pressures are obtained. The droplet will therefore slowly 'dry out', leading to supersaturation of the protein solution, and bringing the protein into the nucleation zone where crystal growth is initiated (Figure 1.10).

The dialysis method involves dialysing the protein solution against a solution containing the precipitant solution. The protein concentration remains the same in the dialysis vessel, while the precipitant concentration increases through diffusion through the dialysis membrane to mix with the protein solution (Figure 1.10). An advantage of this technique is that the crystallisation conditions can be continuously varied, but a drawback is that it is difficult to assess the outcome of the crystallisation trial.

A micro-batch experiment involves mixing the protein solution with a precipitant solution, at their final concentration, to directly form a supersaturated solution (Figure 1.10). Droplets are covered with paraffin oil to prevent evaporation. An advantage of the micro-batch method is that a wide range of droplet volumes can be used, and it is a relatively simple set-up to perform, even for initial screening for crystallisation conditions.

The starting point of the crystallisation process in the phase diagram is tuned by changing the protein and precipitant concentrations. Factors affecting the equilibration rate and outcome of crystallisation trials are droplet-to-reservoir distance, droplet:reservoir volume ratios, temperature and choice of precipitating agents (Forsythe *et al*, 2002).

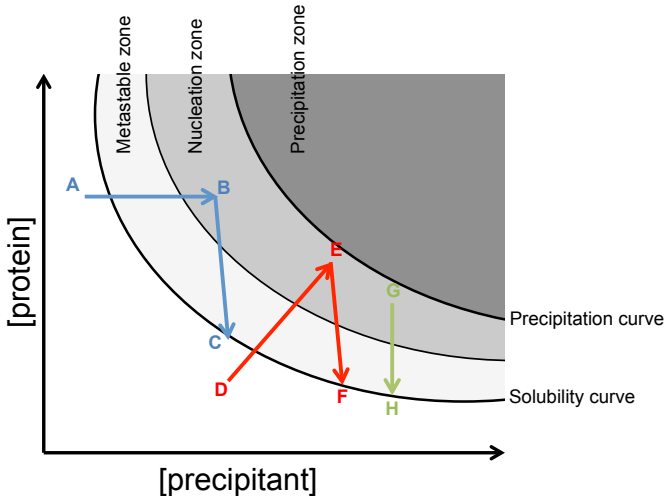


Figure 1.10 The phase diagram in a crystallisation experiment.

The phase diagram for a protein solution is illustrated. The metastable, nucleation and precipitation zones are illustrated in relation to the concentrations of protein (vertical axis) and the precipitant (horizontal axis). The blue arrow shows the path of a protein solution during crystallisation by dialysis, where the protein concentration remains constant in the dialysis vessel (**A**) while the precipitant concentration increases until the protein reaches the nucleation zone (**B**) from which crystals grow until the protein solution reaches its solubility limit (**C**). The red arrows show the path of the protein solution during a vapour-diffusion experiment, where the protein and precipitant concentrations (**D**) are both concentrated through dehydration of the crystallisation droplet, bringing the protein into the nucleation zone (**E**) where crystals form and grow (**F**). The green arrow shows the path of a protein solution in a (micro-)batch experiment, where the protein solution is mixed with a precipitant solution, producing a supersaturated solution (**G**), from which crystal growth occurs (**H**).

All these methods can be used with detergent-solubilised membrane proteins, and are in these cases termed as *in surfo* methods. Through *in surfo* methods, “type II” membrane protein crystals are formed from the detergent-solubilised protein molecules, where the crystal contacts are made by extracellular, polar regions of the proteins, as illustrated in Figure 1.11*b* (Michel, 1991). The disadvantage of this type of crystal is that they possess a high solvent-content which is likely to give rise to fragile crystals that diffract to a low resolution and are highly sensitive to radiation damage (Carpenter *et al*, 2008). Additionally, crystal formation may be unsuccessful if the polar regions are too small and/or the detergent that works best to solubilise the protein does not allow for the molecules to come close enough to each other to form crystals.

More recently, new crystallisation methods for membrane proteins have been developed that involve the reconstitution of detergent-solubilised proteins into a lipidic environment, mimicking their native environment. This

includes, for example, the lipidic cubic phase (LCP) and lipidic sponge phase (LSP) crystallisation techniques (Landau & Rosenbusch, 1996; Wöhri *et al*, 2008). In these techniques, a lipid bilayer matrix is created, into which membrane protein molecules are reconstituted (Caffrey & Cherezov, 2009). Upon the addition of precipitating agents, a phase change of the cubic phase into a lamellar phase is induced, which promotes the association of protein molecules and the formation of protein crystals. A drawback of the LCP method is that the phase is very stiff, making handling of it difficult and time-consuming.

The principle of the LSP crystallisation method is the same as for the LCP technique, where the protein is reconstituted into a lipdic bilayer matrix. Upon the addition of agents that destabilise the LCP, for example polyethylene glycol (PEG), propylene glycol (PG), dimethyl sulfoxide (DMSO), or jeffamine (Wadsten *et al*, 2006), the LCP "melts" to form a more fluid cubic phase. This creates more room for the LSP to accommodate extracellular domains, and the more liquid nature of the lipidic phase makes the crystallisation experiment much easier, with the phase being adaptable to use in crystallisation robots and in conventional vapour-diffusion set-ups.

Other alternative crystallisation techniques also include using detergent-lipid mixtures. This is employed in, for example the bicelle method (Faham & Bowie, 2002) or HiLiDe (High Lipid Detergent concentration) methods (Gourdon *et al*, 2011).

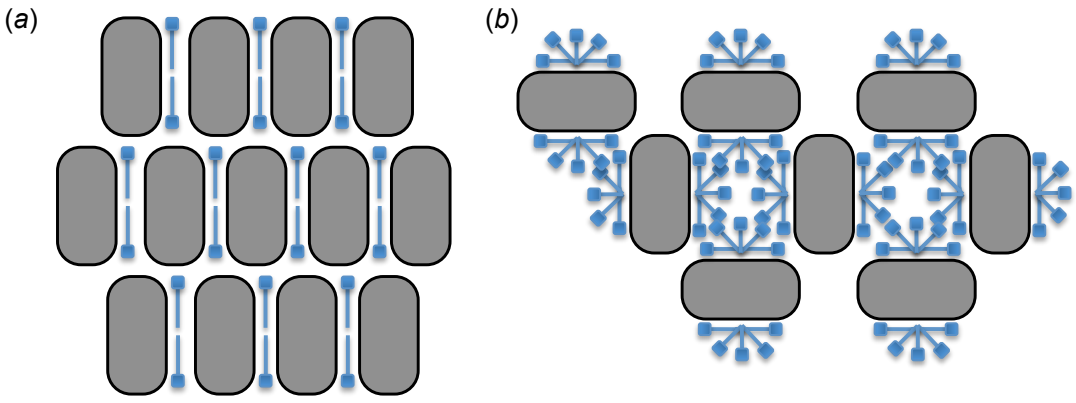


Figure 1.11 "Type I and II" membrane protein crystals.

Type I membrane protein crystals (*a*) are comprised of 2D layers of the membrane proteins, which rely on hydrophobic contacts between the protein molecules. On the contrary, type II membrane protein crystals (*b*) are stabilised by polar contacts between the extracellular portions of the proteins. Membrane proteins are illustrated in grey and detergent monomers are illustrated in blue.

In the case of lipidic phase based crystallisation methods, “type I” membrane protein crystals are formed, where the protein molecules are arranged into 2D layers and rely on hydrophobic contacts between the proteins to stabilise the lattice, as shown in Figure 1.11*a* (Michel, 1991). This results in a greater area of the protein molecules stabilising the crystal lattice, but a drawback of these techniques, compared to *in surfo* methods, is that it can be difficult to work with lipidic mixtures.

1.4.5 X-ray scattering studies of membrane proteins

The crystallisation step is a great bottleneck in structural studies of (membrane) proteins, and the resulting crystals can also suffer from crystal packing effects. Therefore, it is often required to use complementary techniques in order to obtain structural information about the protein of interest. Small-angle X-ray scattering (SAXS) has emerged as a complementary technique in structural biology (Svergun & Koch, 2003; Putnam *et al*, 2007; Pérez & Nishino, 2012). From SAXS data of macromolecules in solution, the size and shape of the macromoelcules can be determined. Apart from giving basic information about the macromolecule, such as the radius of gyration (R_g) or maximum particle dimension (D_{max}), the SAXS data can also give insights into the ternary and quaternary structure through distance distribution functions and the generation of low-resolution *ab initio* models.

An essential step in SAXS studies is buffer subtraction, as the buffer components may contribute to the SAXS signal. As for crystallographic studies, the presence of detergents poses problems in SAXS studies, especially in the buffer subtraction step. When studying detergent solubilised proteins, a detergent concentration is required above its CMC, in order to maintain the integrity of the PDCs. This gives rise to the formation of free detergent micelles in the solution, which results in strong and characteristic scattering patterns (Lipfert *et al*, 2007; Oliver *et al*, 2013), as illustrated in Figure 1.9. This makes detergent subtraction especially crucial in the study of detergent-solubilised membrane proteins using SAXS. However, accurate buffer subtraction is difficult to achieve as the concentrations of free micelles in the sample and the buffer are difficult to match as the detergents give rise to a complicated equilibrium in the presence and absence of membrane proteins (Laughlin, 1996).

SAXS studies on membrane proteins have been previously performed, where the contribution of the SAXS signal can be minimised by choosing detergents that do not contribute too much to the scattering signal or by subjecting the sample to extensive dialysis (Zimmer *et al*, 2006). In addition, standard *ab initio* modelling approaches are not applicable on PDCs, as these assume a uniform electron density throughout the molecules being modelled (Grossmann, 2007). In order to overcome the problem of inaccurate buffer subtraction, and make SAXS an universal technique that can be applied to any membrane protein system, size-exclusion coupled SAXS (SEC-SAXS) was recently proposed as a method to achieve accurate buffer subtraction (Berthaud *et al*, 2012a). In addition, a new modelling approach was also presented and applied to initial SEC-SAXS data, allowing for the modelling of the detergent corona surrounding a protein (Perez & Koutsioubas, 2015).

1.5 Aims and objectives of this Thesis

As discussed in sections 1.2.3.1 and 1.3.2, a wide range of crystal structures of various members of the LeuT superfamily are currently available. These have been crystallised in various conformational states along the alternating access transport model, giving ample insight into how this important class of proteins work. The alternating access transport model for secondary-transport proteins is now widely accepted in the field, but it has been derived from crystal structures of the transport proteins in static states. In addition, the range of conformational states observed is limited to a few key states and these are most likely to be low-energy conformations of the protein along the transport mechanism (Shi, 2013). This has resulted in only limited insight into the transitions and dynamics between the low-energy states. Therefore, it is still of great interest to investigate the conformational landscape of secondary active transport proteins and to attempt to obtain crystal structures of the proteins in conformational states not previously seen. In addition, there are no crystal structures of any eukaryotic members of LeuT superfamily. Additional crystal structures of bacterial LeuT superfamily members will provide further insights and serve as homology models for related proteins in humans, as well as in plants and fungi. For example, Mhp1 was recently used to build a homology model of a plant NCS-1 family member (Witz *et al*, 2014), and using Mhp1 or related proteins

as models for fungal homologues could be used in the development of antifungals (Pantazopoulou & Diallinas, 2007).

Based on this current state of the field of structural studies of secondary active transport proteins, the objectives of this Thesis are two-fold. Firstly, the main objective is to obtain more structural information about the Mhp1 protein in order to gain more insights into the transport mechanism. This will be achieved by investigating a range of mutants of Mhp1, which have previously been shown to have altered ligand binding and transport behaviour (Jackson, 2012). In addition, since the previously published structure of Mhp1 in an inward-open state contained an unidentified ligand, the reproduction of this structure was also attempted in order to better understand this conformational state. This requires the preparation of the various mutants and variants to be studied, which is described in Chapter 3. Crystal structures of these mutants were determined in order to obtain a structural basis for the altered kinetic behaviour, and potentially catch additional alternative conformational states that give rise to the altered kinetic behaviour. The resulting crystal structures are presented in Chapter 5.

The second main objective of this Thesis is to overcome the challenges associated with structural studies of membrane proteins and further optimise the results in order to be able to obtain structural information beyond what is already known about Mhp1, as well as extending the knowledge of how to universally apply structural methods on membrane proteins. This was achieved by exploring various avenues for studying the structure of membrane proteins. Initially, various techniques to improve membrane protein crystals and the resulting diffraction data were investigated, which was required in order to obtain crystal structures of Mhp1 mutants. The crystallisation trials and the subsequent optimisation of these is described in Chapter 4.

The applicability of SAXS was also investigated as an alternative, complementary technique to study the structure of the Mhp1 protein (Figure 1.12). The problem of the presence of detergents was addressed, exploring the potential of using SAXS as an alternative structural technique on membrane proteins in order to obtain structural information where the X-ray crystallographic information is limited. The advantage of SAXS is that it is carried out on solution samples, meaning that the protein is not constricted by a crystal lattice. This is potentially very useful in the study of secondary

active transport proteins, as all previously published structures have all been obtained in the same conformational states, suggesting that only a few conformational states are crystallisable.

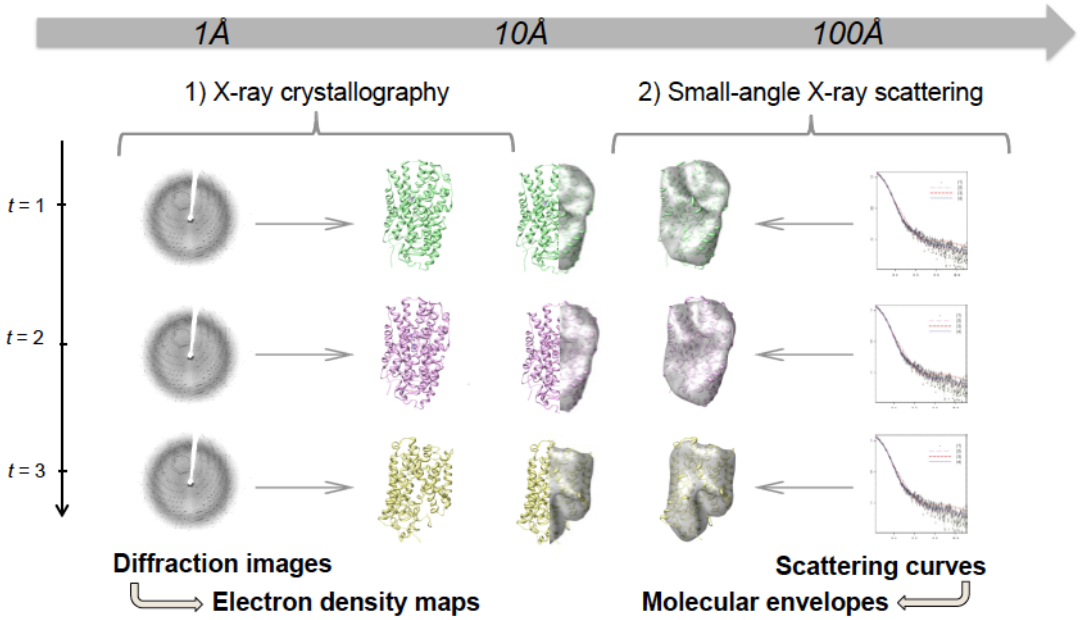


Figure 1.12 The complementarity between X-ray crystallography and small-angle X-ray scattering in studies of membrane transport proteins.

The figure illustrates how (1) X-ray crystallography and (2) small-angle X-ray scattering (SAXS) may complement each other to obtain a full picture of the transport mechanism. The two techniques cover a wide resolution range and give rise to different types of structural information, with X-ray crystallography yielding high-resolution crystal structures while SAXS yields low-resolution molecular envelopes. The three structural models at various time points *t* represent the Mhp1 in its various conformational states: outward-open form (green), the occluded form (pink), inward-open form (yellow). In cases where a certain conformational state or transition cannot be obtained due to crystal packing effects, molecular envelopes obtained from SAXS data collected on the protein in solution could fill in the gap.

Chapter 2: Materials and methods

2.1 Source of materials

2.1.1 Protein purification

Tris base, imidazole, NaCl, glycerol, (Sigma-Aldrich, UK); dithiothreitol (DTT) (Fluka Biochemika); Ni^{2+}-NTA resin (Qiagen); Vivaspin 20 and Vivaspin 6 MWCO 100,000 (Sartorius AG); Econo-Pac® 10DG desalting column, Econo-Column® 2.5x10 cm glass column (Bio-Rad Laboratories, Inc.); *n*-dodecyl-*β*-D-maltopyranoside (DDM), *n*-undecyl-*β*-D-maltopyranoside (UM), *n*-decyl-*β*-D-maltopyranoside (DM), *n*-nonyl-*β*-D-maltopyranoside (NM), all Anagrade (Anatrace).

2.1.2 SDS-PAGE and Western blotting

Bis-acrylamide (2%), acrylamide (40%) (Bio-Rad Laboratories, Inc.); glacial acetic acid, methanol (Fisher Scientific, UK); ammonium persulphate, Brilliant Blue R 250, Bromophenol blue, glycine, β-mercaptoethanol, TEMED, Tris base, Tween-20, bovine serum albumin (BSA), sodium dodecyl sulphate (SDS) (Sigma-Aldrich, UK.); Immobilon-P polyvinylidene fluoride transfer membrane (Millipore); HisProbe™-HRP and SuperSignal West HisProbe™ kit (Fisher Thermo Scientific). SDS7 and SigmaMarker™ molecular weight markers (Sigma-Aldrich); Amersham ECL High-range Rainbow Molecular Weight Marker (GE Healthcare).

2.1.3 Protein concentration assay

BSA, NaOH, naphthol blue black, SDS, trichloroacetate (TCA), ethylenediaminetetraacetic acid (EDTA), Tris base (Sigma-Aldrich, UK); ethanol, methanol, glacial acetic acid (Fisher Scientific Ltd., UK), MF-Millipore Membrane Filter, 0.22 µm pore size, 25 mm diameter (Merck Millipore).

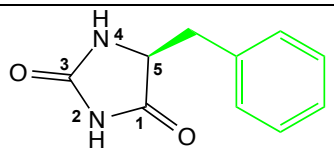
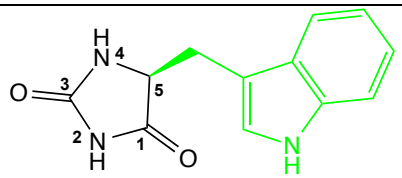
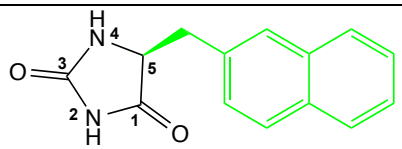
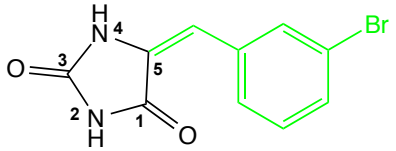
2.1.4 Ligands used in the experiments

5-Benzyl-L-hydantoin (L-BH), 5-indolylmethyl-5-hydantoin (L-IMH), 5-(2-naphthylmethyl)-L-hydantoin (L-NMH) and 5-bromovinylhydantoin (BVH) were synthesised and provided by Dr Simon Patching or Dr Katie Simmons,

University of Leeds (Simmons *et al*, 2014). The structures of the ligands are shown in Table 2.1.

Table 2.1 Chemical structures of the ligands used in this Thesis.

The chemical structures of the ligands outlined in section 2.1.4, used for co-crystallisation trials in this Thesis. The hydantoin moiety of each ligand is displayed in black, while the aryl-substituents are displayed in green.

Compound name	Abbr.	Structure
5-Benzyl-L-hydantoin	L-BH	
5-Indolylmethyl-L-hydantoin	L-IMH	
5-(2-Naphthylmethyl)-L-hydantoin	L-NMH	
5-Bromovinyl-hydantoin	BVH	

2.1.5 Tryptophan fluorescence spectrophotofluorimetry assay

Choline chloride, dimethyl sulfoxide (DMSO), NaCl, Tris base (Sigma-Aldrich, UK); KCl, K_2HPO_4, KH_2PO_4 (Fisher Scientific Ltd., UK); DDM, UM, DM, NM (Anatrace), DTT (Fluka Biochemika).

2.1.6 Size-exclusion chromatography for small-angle X-ray scattering studies

Superdex™ 200 Increase 3.2/300 column (GE Healthcare), Bio SEC-3 column (Agilent); MF-Millipore Membrane Filter.

2.1.7 Crystallisation

24-well trays (Greiner Bio-One); 18x18 mm glass coverslips (SLS); Mounted CryoLoops™ (Hampton Research); MicroMeshes™, MicroLoops E™, *In Situ*-1™ and CrystalDirect™ crystallisation plates (MiTeGen); polyethylene glycol 300 (PEG300), Tris base, NaCl (Sigma-Aldrich, UK); Na_2HPO_4, NaH_2PO_4, KCl, K_2HPO_4, KH_2PO_4 (Fisher Scientific Ltd., UK); DDM, UM, DM, NM (Anatrace), DTT (Fluka Biochemika), 0.22 µm Millex ® syringe filter unit, 33 mm diameter (Merck Millipore).

2.2 Purification and characterisation of Mhp1

2.2.1 Purification

30 or 100 litre fermentations of BL21 (DE3) *E. coli* cells (Novagen™) expressing native or mutant His_6-tagged Mhp1 were performed by Mr David Sharples. Cell stocks with mutant Mhp1 were previously prepared by Dr Scott Jackson (Jackson, 2012) and the mutants used here are outlined in Table 2.2. Preparations of inner membranes from the fermentations were performed by Mr David Sharples, as previously described (Jackson, 2012) and stored at -80 °C. Frozen aliquots of inner membrane preparations were thawed and resuspended in solubilisation buffer (20 mM Tris-HCl pH 8.0, 20 mM imidazole pH 8.0, 300 mM NaCl, 20% v/v glycerol, 1% w/v DDM) to a total protein concentration of 1-3 mg/ml. The total protein concentration was determined as described in section 2.2.2 on 1:30 dilutions of the inner membrane preparations. 1:30 preparations were prepared by, firstly, making a 1:1 dilution of the inner membranes to improve sample viscosity followed by dilution to a 1:30 ratio. The solubilisation suspension was incubated for 4 h with stirring at 4 °C. The soluble fraction containing solubilised protein was separated from the membranes by centrifugation at 100,000 *g* (41,000 rpm) for 1 h using a Ti-45 rotor (Beckman) in an Avanti Optima L-80 ultracentrifuge (Beckman). The supernatant was incubated with Ni^{2+}-NTA resin, which was pre-equilibrated with wash buffer 1 (10 mM Tris-HCl pH 8, 20 mM imidazole pH 8, 2.5% v/v glycerol, 0.05% w/v DDM), for 16 h at 4 °C with stirring. The Ni^{2+}-NTA resin was washed with 2 column volumes (CV) of wash buffer 1 in a Econo-Column® glass column. The His_6-tagged Mhp1 bound to the resin was eluted by applying elution buffer (10 mM Tris-HCl pH 8, 200 mM imidazole pH 8, 2.5% v/v glycerol, 0.05% w/v DDM). The eluted protein was concentrated using a Vivaspin 20 MWCO 100,000 spin concentrator at 3000 *g*, to a volume of approximately 3 ml. The concentrated protein solution was applied onto an Econo-pac® 10DG desalting column,

which was pre-equilibrated with storage buffer (10 mM Tris-HCl pH 7.6, 2.5% v/v glycerol, 0.05% w/v DDM). After applying 4 ml of storage buffer onto the desalting column, the applied protein fraction was eluted, collected and concentrated to a final concentration of 10-20 mg/ml before storage in 10-20 µl aliquots at -80 °C after flash-freezing in liquid nitrogen. Growth of SeMet-derivatised Mhp1 in flasks in minimal media supplemented with 50 mg/ml seleno-L-methionine, as well as subsequent inner membrane preparations were performed and provided by Mr David Sharples.

Table 2.2 List of mutants successfully purified and characterised for this project.

Mutant	Site of the mutation
Gln42Asn	Ligand
Gln121Asn	Ligand
Asn318Gly	Ligand
Ser312Ala	Sodium
Lys110Leu	Bundle/hash

SeMet-Mhp1 was purified as native Mhp1, but with all buffers supplemented with 10 mM DTT. For crystallisation trials and small-angle X-ray scattering (SAXS) studies, an additional wash step was added to the Ni^{2+}-NTA column step in order to exchange the detergent. The wash step with wash buffer 1 was followed by an additional wash of 2 CV with wash buffer 2, containing the detergent of interest. The bound protein sample was then eluted with elution buffer and exchanged into storage buffer as described above, but containing the detergent of interest. Table 2.3 outlines the various detergents used and the respective final concentrations (%, w/v) in each buffer. The concentrations of the detergents in relation to their CMC were the same in each buffer, namely 2.5 CMC for wash buffer 2 and the elution buffer and 1.8 CMC for the storage buffer.

2.2.2 Protein concentration assay

The concentration of protein in inner membrane preparations and of purified protein was determined using the Schaffner-Weissman assay (Schaffner & Weissmann, 1973). Sample volumes were adjusted to 270 µl with deionised

water, to which 30 µl of a solution containing 1 M Tris-HCl pH 7.5, 10% w/v SDS and 60 µl of 60% w/v TCA was added. Samples were vortexed and transferred drop-wise onto a single spot, marked with a pencil, on a 0.22 µm MF-Millipore Membrane Filter under a vacuum manifold. The filter was pre-soaked in deionised water. The eppendorf tube previously containing the sample was rinsed with 200 µl 6% w/v TCA, which was also added onto the membrane in the same spot. The whole membranes were rinsed with 2 ml 6% w/v TCA before being transferred into a petri dish containing staining solution (methanol, acetic acid and water in a 45:10:45 v/v ratio with 0.1% w/v naphthol blue black) and left for 3 minutes under stirring. Thereafter, the membrane filter was transferred for 30 s into deionised water, twice for 1 minute into destaining solution (methanol, acetic acid and water in a 90:2:8 v/v ratio) and finally, for 2 minutes into deionised water. Pieces containing the stained dots of protein were cut out and transferred into Eppendorf tubes containing 1 ml elution buffer (25 mM NaOH, 0.05 mM EDTA, 50% v/v ethanol). A blank piece of the same filter paper, of a similar size to the sample pieces, was also cut out and transferred into a separate tube containing elution buffer. The OD_{630} of the eluted samples was measured using a WPA Biowave II UV/Visible spectrophotometer (Biochrom Ltd.), using the sample from the blank membrane filter as the 100% transmission reference and concentrations were calculated with reference to a BSA standard graph.

Table 2.3 Detergents concentrations used for buffer exchange.

The concentrations (%, w/v) outlined in the table used for each of the detergents in the buffers used for detergent exchange from DDM, which was used to solubilise the protein. The concentrations correspond to 2.5 (wash buffer 2 and elution buffer) and 1.8 (storage buffer) of the CMC of each detergent.

Detergent	**Wash buffer 2 (%, w/v)**	**Elution buffer (%, w/v)**	**Storage buffer (%, w/v)**
n-Undecyl-*β*-D-maltopyranoside (UM)	0.07	0.07	0.05
n-Decyl-*β*-D-maltopyranoside (DM)	0.2	0.2	0.16
n-Nonyl-*β*-D-maltopyranoside (NM)	0.7	0.7	0.5

2.2.3 SDS-PAGE and Western blotting

The purity of the protein samples was assessed using SDS-PAGE gels, consisting of a 15% v/v resolving gel and a 4% v/v stacking gel. Samples of protein solution (15 µl) from various steps in the purification were mixed with 4X SDS-polyacrylamide gel electrophoresis (SDS-PAGE) sample buffer (60 mM Tris-HCl pH 7.2, 10% v/v glycerol, 2% w/v SDS, 0.005% w/v bromophenol blue, 3% v/v β-mercaptoethanol) to a final sample volume of 20 µl, vortexed and incubated at room temperature for 15 minutes. Pure protein samples were diluted to contain ~10 µg protein. The gels were resolved at 140 V for approximately 1 h in SDS-PAGE running buffer (1 g/L SDS, 3.03 g/L Tris base, 14.41 g/L glycine). Thereafter, the SDS-PAGE gel was placed into destaining solution 1 (methanol, acetic acid and water in a 50:10:40 v/v ratio) for 5 minutes with gentle stirring, followed by staining solution for 15 minutes (methanol, acetic acid and water in a 50:10:40 v/v ratio, containing 0.2% w/v Coomassie Brilliant Blue R). The gel was then left in destaining solution 2 (methanol, acetic acid and water in a 8:10:82 v/v ratio) until the bands were clearly visible for comparison to the molecular weights of the SDS-PAGE standard molecular weight marker proteins that were run on the same gel.

A gel for subsequent Western blotting analysis for detection of His_6-tagged Mhp1 was run simultaneously with the primary SDS-PAGE gel. The samples for Western blotting were prepared by making a 1:14 dilution of the initial samples for the primary SDS-PAGE gel to a final volume of 15 µl, to which 5 µl of 4X SDS-PAGE sample buffer was added. The resolved gel for Western blotting analysis was transferred onto a Immobilon-P polyvinylidene fluoride transfer membrane by electroblotting at 15 V for 20 minutes using a Trans-Blot SD Semi-dry transfer cell (Bio-Rad Laboratories Inc.). The membrane was incubated with 10 ml TBST buffer (25 mM Tris-HCl ph 7.8, 150 mM NaCl, 0.05% v/v Tween-20) containing 3% w/v BSA for 1 h with stirring. After washing the membrane twice with 15 ml TBST buffer for 10 minutes, it was incubated with 10 ml TBST buffer containing 5 µl Hisprobe-HRP solution for 1 h. This step was followed by four washes with 15 ml TBST buffer for 10 minutes. Thereafter, the membrane was incubated with 10 ml of SuperSignal Working solution for 5 minutes. All the reaction steps were carried out with stirring. The sizes of the proteins were compared to High-range Rainbow molecular weight markers. Images of SDS-PAGE gels and Western blots were taken using a G:Box (Syngene) with the program GeneSnap

(Syngene). White light and luminescence images of Western blots were merged using the Intelli-Chemi mode in GeneSnap.

2.2.4 Steady-state spectrophotofluorimety

The binding affinities of various ligands were tested using a Photon Technology International fluorescence spectrofluorimeter. Purified protein sample was diluted to a final concentration of 140 µg/ml in fluorimetry buffer containing 50 mM Tris-HCl pH 7.6, 2% v/v DMSO and storage buffer detergent concentration (Table 2.3). The fluorimetry buffer also contained 0/15/140 mM NaCl, with 140/125/0 mM choline chloride, respectively. The choline chloride was added in order to keep the ionic strength constant between the solutions. Prior to a titration, the protein solution was incubated inside the sample cell of the fluorimeter, at 18 °C for 4.5 minutes in a quartz cuvette with a path length of 1 cm with stirring. Thereafter, ligand solution (100 mM, dissolved in 100% v/v DMSO) was titrated into the protein solution to a final concentration varying from 0-2 mM using a 10 µl Hamilton syringe. 1.5 minute after each addition of ligand solution, the emitted fluorescence was recorded over the wavelength range 310-360 nm with an excitation wavelength of 280 nm. Titrations were carried out in duplicates or triplicates and the percent fluorescence quenching was calculated with reference to that with 0 mM ligand. The results were plotted and fitted using the Michaelis-Menten tool in Prism 6 (GraphPad), from which apparent dissociation constant (K_d, mM) and the maximal binding (B_{max}, %) were obtained.

2.2.5 In-house circular dichroism

Far-UV circular dichroism (CD) data was collected on a Chirascan™ CD Spectrometer (Applied Photophysics), funded by the Wellcome Trust grant code 094232, using the Pro-Data™ control software. Concentrated protein samples were diluted in storage buffer containing the required detergent to approximately 0.1-0.2 mg/ml in a quartz cuvette with a 1 mm path length. Spectra were recorded at 20 °C over the wavelength range of 180-260 nm with a 1 nm step, 0.5 s/step, 4.3 mm bandwidth. Spectra of the buffer were also recorded and subtracted from the protein spectra. Thermal melts were performed with the same parameters over the range 20-90 °C, with a 1 °C step and an additional 30 s for stabilisation of the temperature after each increment.

2.3 Crystallisation and data collection

2.3.1 Crystallisation of Mhp1

The starting crystallisation conditions were based on those previously published (Shimamura *et al*, 2008). In all crystallisation trials, protein where the detergent had been exchanged from DDM to NM was used. All solutions were passed through a 0.22 µm Millex ® syringe filter unit before being used for making crystallisation conditions. With every new batch of native or mutant Mhp1, an initial screen was set up with fresh protein to establish the optimal crystallisation conditions for that particular purification batch. Droplets of protein solution (0.5 or 1.0 µl, ~10-20 mg/ml) were added onto non-silliconised glass cover slips, to which well solution was added in a 1:1 or 1:2 ratio. The cover slips with the mixed droplets were then inverted into a hanging-drop setup over the wells of a 24-well Greiner tray and sealed with vacuum grease. The wells contained 400 µl of 0.1 M sodium phosphate pH 7.0, 60-160 mM NaCl and 27-35% v/v PEG 300 and the trays were left to equilibrate at 18 °C. Typically, needle-shaped crystals appeared over-night and continued to grow over the course of a week. After setting up plates with fresh protein, the remaining protein from the purification batch was flash-frozen in 10-20 µl aliquots and stored at -80 °C for use in subsequent trays for further optimisation, mass production and co-crystallisation. The frozen aliquots were usually used up within 1 month. In order to improve crystal size and quality, a streaking method was developed to mix the well solution into the protein droplet (further described in 4.4.3).

For SeMet-Mhp1, droplets of protein solution (1.5 µl, ~10 mg/ml) were mixed with an equal volume of reservoir solution containing 24-28% v/v PEG 300, 60-120 mM NaCl and 100 mM Bicine (pH 9.0) and left to equilibrate against 400 µl of the same solution. The trays were left to equilibrate at 4 °C and needle-shaped crystals appeared after a couple of days.

2.3.2 Co-crystallisation with ligands

For co-crystallisation of ligands with mutant Mhp1, the ligands L-BH, L-5-IMH, L-NMH and BVH were dissolved in 100% PEG 300 to a final concentration of 1-5 mM. The ligand-containing PEG 300 solution was thereafter used to make up the reservoir solutions for crystallisation set-ups as described in 2.3.1.

2.3.3 Harvesting and cryo-cooling of crystals

Due to the high content of PEG300 in the crystallisation conditions, the crystals were cooled without adding additional cryo-protectant. The crystals were typically harvested using nylon loops (0.2-1 mm diameter) or MicroLoopsE™ (vertical; 15x150, 30x300, 50x500 or 70x700 µm), and cryo-cooled by rapid plunging into liquid nitrogen or placement in the path of a nitrogen gas stream. For crystal dehydration, a humidity control device (HC1) (Sanchez-Weatherby *et al*, 2009) was used at beamlines I03 and I04 at Diamond Light Source. After the relative humidity (RH) of the reservoir solution from the well with crystal(s) to be studied was established, a crystal was mounted on a MicroMesh™ and put on the goniometer in the path of the humid air stream. Excess liquid was wicked from underneath the mesh by touching with a piece of tissue. Thereafter, the RH was decreased in increments of 2% with a 2 minute waiting time in-between. At each humidity step, one image of 1° oscillation and 1 s exposure (1% transmission at a flux of 2.61 x 10^{11} photons s^{-1}, 0.9763 Å) was collected to assess the diffraction quality and lattice parameters. When a significant contraction of the lattice was observed, the crystal was flash-cooled by automatic switch over to a nitrogen gas stream and a full diffraction dataset was collected.

2.3.4 Crystallisation for *in situ* studies and contactless harvesting

For *in situ* studies and contactless harvesting, *In Situ*-1™ and CrystalDirect™ crystallisation trays were set up manually, with 0.4-0.5 µl droplets mixed in a 1:1 or 1:2 ratio with 45 µl well solution. The solutions were prepared with or without ligands as described in 2.3.1 and 2.3.2.

2.3.5 Diffraction data collection and structure solution

Diffraction data were collected at Diamond Light Source (beamlines I02, I03, I04, I04-1 and I24) and the European Synchrotron Radiation Source (beamline ID29). The data were processed using xia2 (Winter *et al*, 2013); where indexing, integration and scaling was performed with XDS (Kabsch, 2010) or DIALS (Waterman *et al*, 2013) and space group determination and merging of the individual data sets was done with POINTLESS and AIMLESS from the CCP4 suite (Winn *et al*, 2011). The Pearson correlation coefficient (CC) was used as the main data quality metric and the resolution was normally cut at $CC_{1/2}$ > 0.5 (Karplus & Diederichs, 2012). To avoid biasing the R_{free} set, the R_{free} flags for each data set were copied from the

previously published structure with PDB 2JLO (Weyand *et al*, 2008) in the CCP4 programme CTRUNCATE (Winn *et al*, 2011). The structures were solved by molecular replacement (MR) using PHASER (McCoy *et al*, 2007). The principle of MR is to determine the phase of each diffraction reflection, which is not recorded during a diffraction experiment, by determining the position of a protein search model in the unit cell. Patterson-based methods have been widely applied for this purpose and involves matching the peaks in the Patterson map calculated for the chosen search model to that of the experimental data. The Patterson map, containing peaks for each interatomic distance is a result of the Patterson function, is essentially a Fourier transform of the intensities of the diffraction pattern (Glusker, 1983). The search procedure is normally divided into two steps to make the process more efficient; 1) a rotation search, where the correct orientation of the search model is determined and 2) a translation search, where the position of the correctly oriented model is determined. The procedure has been improved and implemented in a wide range of MR programmes (Scapin, 2013). In PHASER, the improvements have involved the introduction of maximum-likelihood statistics and a brute force evaluation of the MR solutions (McCoy *et al*, 2007).

MR requires a search model of usually more than 30% sequence identity in order for the search to be successful. For the Mhp1 structures solved in this Thesis, search models of previously published structures of Mhp1 were used. In order to validate that the correct solution has been obtained at the moderate resolution of the data obtained here, and avoid the propagation of model bias throughout the process of refinement, MR searches were performed with a selection of search models against each data set. The search models included previously published structures of Mhp1 in the outward-open conformation (2JLN), the occluded ligand-bound conformation (2JLO, 4D1A and 4D1B), the occluded ligand-bound conformation with TMH10 in an open conformation (4D1D) as well as in the inward-open conformation (2X79). Prior to using these structures as molecular replacement search models all bound ligands, ions and metals were removed.

Structures were initially refined using both PHENIX.REFINE (Adams *et al*, 2010) and REFMAC5 (Murshudov *et al*, 2011), and the final structures were produced using REFMAC5, where the structures were initially subjected to a rigid-body fit, followed by restrained refinement. Parameters included in the restrained refinement were jelly-body refinement and the application of

external restraints from a reference model using ProSmart built into REFMAC5 (Nicholls *et al*, 2012). *B*-factor map sharpening was used for map visualisation and model building, applying the default *B*-factor sharpening factor (B_{sharp}) in REFMAC5.

2.4 Small-angle X-ray scattering studies

2.4.1 Data collection on dilution series

Data were collected at beamline BL4-2 at Stanford Synchrotron Radiation Lightsource (CA, USA). A 2.5 m sample-to-detector distance was used and the wavelength of the radiation was 1.13 Å. Briefly, 8 x 1 s exposures were recorded for each sample (either reference buffer or sample dilution) while the sample was oscillated back and forth inside a capillary tube. The data were automatically scaled, integrated and buffer subtracted using SASTool (http://ssrl.slac.stanford.edu/~saxs/analysis/sastool) at the beamline.

Protein samples (~10 mg/ml) and their respective reference buffers were prepared by several rounds of dilution/re-concentration using Vivaspin 20 MWCO 100,000 spin concentrators, with wash buffer with a reduced detergent concentration (10 mM Tris-HCl pH 7.6, 2.5% v/v glycerol, 0.02% w/v DDM). The final dialysis buffer or the buffer used for concentration was used for making dilution series of the protein. This buffer was also used as the reference buffer. All further processing was carried out using the ATSAS software package (Petoukhov *et al*, 2012). R_g, $I(0)$, D_{max} and distance distribution functions were obtained using the 'Radius of Gyration' and 'Distance Distribution' tools in PRIMUS (part of the ATSAS software package). Porod volumes were calculated using DATPOROD (part of the ATSAS software package).

2.4.2 SEC-SAXS data collection

SEC-SAXS data were collected at the SWING beamline at SOLEIL, the French national synchrotron facility. Purified protein samples in 10-20 µl volumes were shipped to the synchrotron on dry ice and stored at -80 °C. Protein samples were diluted in storage buffer containing the corresponding detergent to a final concentration of 2-4 mg/ml and a final volume of 50 µl. 45 µl of the protein samples were injected onto a Superdex™ 200 Increase 3.2/300 or Bio SEC-3 column, pre-equilibrated with the same buffer at 0.1 and 0.2 ml/min, respectively, and the elution was monitored using UV absorption at 280 nm. 61 frames with 1.5 s exposure time and 1 s dead time between frames were collected at the beginning of the elution to provide the

buffer signal, and a total of 254 frames were collected across the elution peak, with approximately 30 frames covering the main peak. Images were radially averaged, normalised and the buffer subtraction was performed using the software Foxtrot (http://www.synchrotron-soleil.fr/Recherche/LignesLumiere/SWING). The software package ATSAS (Petoukhov *et al*, 2012) was used for downstream analysis and also utilised in conjunction with the program Memprot (Perez & Koutsioubas, 2015) to model the detergent corona.

Chapter 3: Purification and characterisation of variants of the Mhp1 protein for structural studies

3.1 Introduction

Previously, mutagenesis and subsequent functional studies of Mhp1 mutants have helped to identify several variants with altered ligand binding and transport behaviour (Jackson, 2012; Simmons *et al*, 2014). These studies have allowed the establishment of sequence-function relationships in Mhp1, but have not given any insights into the structural basis behind the altered binding and transport of its ligands. The altered ligand binding might be a result of the mutations causing structural rearrangements in the protein, or the mutants adopting different conformations. One possibility is that the mutants preferentially adopt conformational states other than those previously observed for the wild-type protein, for example an inward-open state. Therefore, the various mutants that exhibit altered ligand binding behaviour are very interesting targets for structural studies.

As outlined in section 1.5, the scope of this thesis is to gain an improved understanding of the full transport mechanism of Mhp1. X-ray crystallographic studies of Mhp1-ligand complexes have previously given insights into the structural basis behind the ligand binding step of the transport mechanism in wild-type Mhp1 (Simmons *et al*, 2014). In this thesis, X-ray crystallographic studies are used to explore the effects of mutations previously determined to have altered ligand binding activities (Jackson, 2012; Simmons *et al*, 2014).

In this Chapter, the purification and characterisation of a number of Mhp1 variants are described in preparation for further structural studies. Both Mhp1 single-point mutants previously identified to have altered ligand behaviour (Simmons *et al*, 2014; Jackson, 2012) and seleno-L-methionine derivatised Mhp1 were prepared (section 3.4). In addition, wild-type Mhp1 solubilised in maltoside detergents with aliphatic chains of varying length were produced (section 3.5) for subsequent small-angle X-ray scattering studies (SAXS), exploring the potential of SAXS as a complementary technique for Mhp1.

A combination of a steady-state tryptophan fluorescence quenching assay and circular dichroism spectroscopy were used to assess the structural integrity, stability of these variants, and subsequently their suitability for further structural studies.

3.2 The preparation of Mhp1 samples and techniques for their characterisation for structural studies

3.2.1 Introduction

In this section, the preparation and subsequent characterisation of the Mhp1 protein for further structural studies will be described, using wild-type Mhp1 as the example. This section will then serve as a template for the rest of the Chapter in the description of the preparation and characterisation of various variants of Mhp1; including single-point mutants, seleno-L-methionine derivatised Mhp1 and Mhp1 solubilised in various detergents.

The preparation of wild-type Mhp1 was performed in order to reproduce and further optimise the previously published crystallisation conditions (Shimamura *et al*, 2008; 2010). As the crystallisation trials for this project would be carried out in a different laboratory than previously, the published conditions were reproduced in order to assess whether these conditions could be further exploited in this project or whether extensive screening for alternative conditions would be required.

Part of the purification protocol for Mhp1 for crystallisation trials involves the exchange of detergent from *n*-dodecyl-*β*-D-maltopyranoside (DDM), which is initially used to solubilise the protein out of the inner membrane fragments, to the three aliphatic chain carbons shorter detergent *n*-nonyl-*β*-D-maltopyranoside (NM) (Figure 3.12). Therefore, the preparation and characterisation of wild-type Mhp1 solubilised in NM will be described in this section, to which single-point mutants and seleno-L-methionine derivatised protein will be subsequently compared. The influence on the outcome of the purification and characterisation of Mhp1 in other detergents will be further discussed in section 3.5, where the characterisation of Mhp1 solubilised in detergents of various length for scattering studies will be described.

3.2.2 Monitoring the purification of Mhp1

Mhp1 for initial crystallisation trials was purified according to previously established protocols, which are described in 2.2.1 (Shimamura *et al*, 2008). The presence of Mhp1 during various steps of the purification was monitored by SDS-PAGE and Western blotting. The SDS-PAGE gel showed the total amount of proteins in each sample, while Western blotting was used to confirm the presence or absence of N-terminally His_6-tagged Mhp1. A

representative SDS-PAGE gel and corresponding Western blot of a purification of Mhp1 is displayed in Figure 3.1.

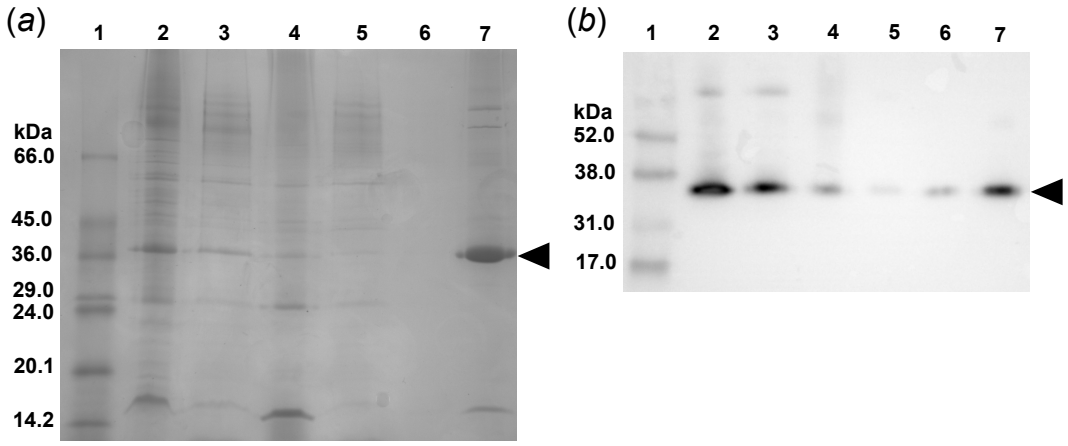


Figure 3.1 SDS-PAGE gel and corresponding Western blot showing the progress of the purification of wild-type Mhp1, solubilised in NM.

SDS-PAGE gel (*a*) and the corresponding Western blot (*b*) of samples taken during the course of a purification of wild-type Mhp1 for initial crystallisation trials. The lanes and samples are the same in both the gel and blot, with the samples on the blot being diluted 1:14 times; the gels were run at the same time. The lanes contain: 1) molecular weight standard ladder with the molecular weights of the standard proteins indicated to the left of the gel or blot in kDa; 2) solubilisation mixture before ultracentrifugation; 3) supernatant after ultracentrifugation, 4) pellet after ultracentrifugation; 5) the unbound material in the flow-through after binding to the Ni-NTA resin; 6) wash buffer of the Ni-NTA resin after exchanging the detergent; and 7) eluted sample from the Ni-NTA resin using 200 mM imidazole.

The Western blot confirms the presence of Mhp1 in the solubilisation mixture prior to ultracentrifugation, which is carried out in order to separate insoluble membrane fragments from the soluble fraction containing solubilised protein. After the ultracentrifugation step, the Western blot confirms that Mhp1 is present in the soluble fraction. The Western blot also shows that small amounts of Mhp1 were lost in the insoluble fraction, and while washing the Ni-NTA resin.

The migration of the final, eluted Mhp1 protein corresponds to approximately 36 kDa on SDS-PAGE (indicated by the arrows in Figure 3.1), while its estimated molecular weight from its sequence is 54.6 kDa. Anomalous migration of membrane proteins on SDS-PAGE, also termed “gel shifting”, is a common occurrence and membrane proteins often migrate at approximately 65–75% of their actual molecular weight on SDS-PAGE (Ward *et al*, 2000). This has been proposed to be caused by membrane

proteins not becoming completely denatured by SDS due to the presence of other detergents and stable hydrophobic contacts within the protein, resulting in the proteins migrating further on the gel due to their more compact size (Rath *et al*, 2009).

3.2.3 Assessing secondary structure of Mhp1 using circular dichroism

Circular dichroism was used in order to examine the structural integrity of the purified protein samples and their suitability for further X-ray crystallographic and scattering studies. Circular dichroism in the UV region can be used to study the global structure of proteins, in particular in the far UV CD region (180-260 nm) (Kelly *et al*, 2005). Within this wavelength range, various secondary structure motifs in proteins give rise to characteristic spectra. Subsequently, CD was used as a quality control to check that the purified protein samples were folded and acquired the expected secondary structural elements of Mhp1. The overall structure of Mhp1 is comprised of twelve TM helices (Weyand *et al*, 2008; Shimamura *et al*, 2010), and α-helical folds typically give rise to a spectrum with a maximum around 193 nm and minima at 208 and 222 nm.

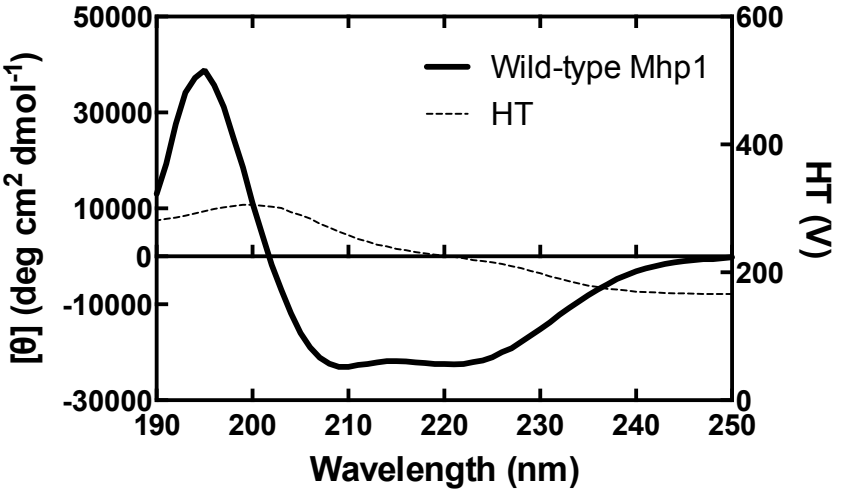


Figure 3.2 Representative CD spectrum of purified wild-type Mhp1, solubilised in NM.

The CD spectrum for purified Mhp1 (in a buffer containing 10 mM Tris-HCl pH 7.6, 2.5% v/v glycerol, 0.5% w/v NM) follows the characteristic shape of an α-helical protein, with signal maximum at 193 nm and minima at 208 and 222 nm. The CD signal is shown as a black line, while the High Tension (HT) voltage is shown as a dotted line. The protein concentration was adjusted so as to avoid the HT going over 600 V.

When recording a CD spectrum of wild-type Mhp1, the presence of these characteristic features are seen, which indicates that the purified Mhp1 protein has predominantly adopted an α-helical fold at 20 °C (Figure 3.2).

3.2.4 Assessing the thermal stability of Mhp1 using circular dichroism

Recording a CD spectrum shows that the protein is folded at the temperature at which the spectrum was recorded, but gives no further information on the inherent stability of the structure. Therefore, CD was also used to monitor the thermal stability of the purified protein samples, in order to gain more information on their stability (in particular the mutants) and their suitability for further characterisation. The thermal stability was monitored by recording CD spectra while increasing the temperature of the sample cell. The thermal melts were carried out between 20-90 °C (Figure 3.3). During the thermal melt, the characteristic shape of the CD spectrum of an α-helical protein, as seen for wild-type Mhp1 in Figure 3.2 disappeared. The peak at 193 nm decreased, and underwent a shift towards a higher wavelength, and the two minima at 208 and 222 nm were also gradually diminished. This indicated that the previously adopted α-helical structure of Mhp1 was being lost. The decrease in CD signal at 208 nm was then used to assess the thermal stability of the protein.

After the thermal melt, the temperature was lowered back down to 20 °C, for which the CD spectrum displayed complete absence of characteristic peaks compared to the spectrum recorded at 90 °C. At 90 °C, CD signal was still recorded between approximately 190-200 nm, which was most likely the peak at 193 nm which had undergone a wavelength shift. Lowering the temperature back to 20 °C did not recover the folding of the secondary α-helical structure, but in fact triggered the complete denaturation of Mhp1, as indicated by the absence of CD signal between 190-200 nm. The complete denaturation may have been caused by lowering the temperature at a faster rate than it was increased during the thermal melt. From the plot of the change in CD intensity at 208 nm, the melting temperature (T_m) of wild-type Mhp1 was estimated as 50.3 °C using the Global 3™ Analysis Software (Applied Photophysics).

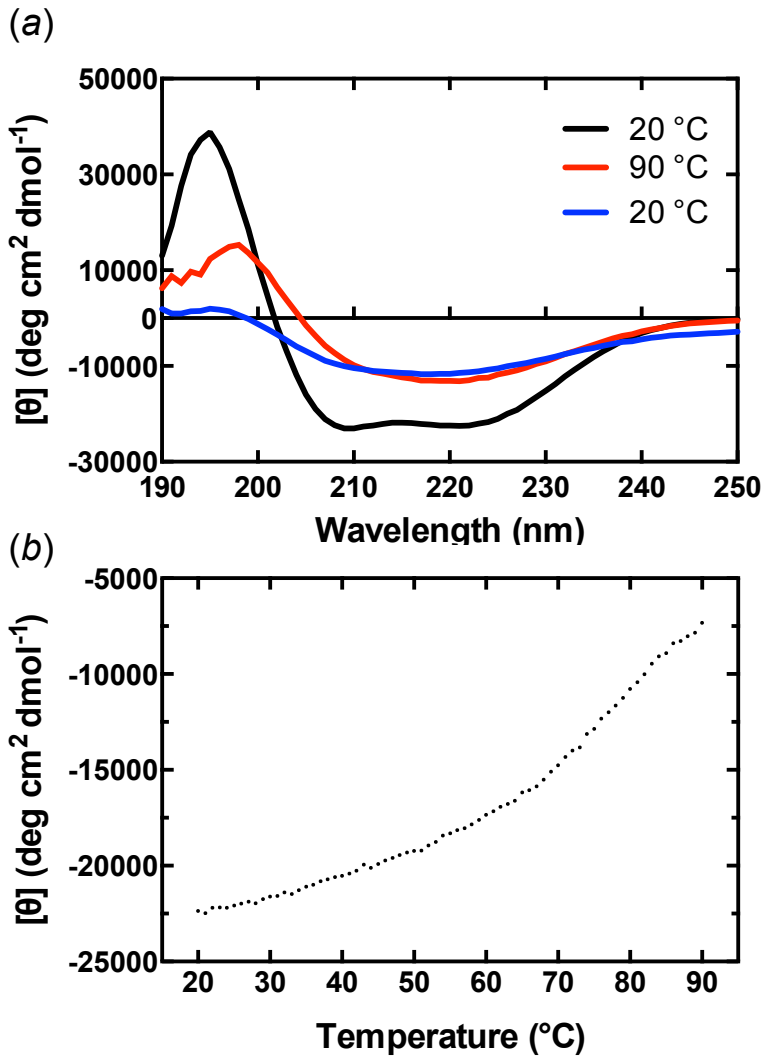


Figure 3.3 Example of a CD thermal melt assay on wild-type Mhp1, solubilised in NM.

(*a*) The thermal stability of purified Mhp1 samples (in a buffer containing 10 mM Tris-HCl pH 7.6, 2.5% v/v glycerol, 0.5% w/v NM) was assessed by increasing the temperature of the CD sample cell from 20-90 °C, and measuring the spectrum in 1 °C steps. The CD spectrum of wild-type Mhp1 at 20 °C is shown in black and at 90 °C in red. A final measurement at 20 °C after the thermal melt up to 90 °C (shown in blue) shows that the thermal denaturation of the secondary structure of Mhp1 is irreversible. (*b*) The CD signal at 208 nm was plotted (shown as black dots) against temperature and used to assess the thermal stability of the protein.

3.2.5 Assessing ligand binding behaviour using steady-state tryptophan fluorescence quenching in wild-type Mhp1

The functionality of the purified Mhp1 protein samples was studied using a steady-state spectrophotofluorimetry assay, monitoring the change in tryptophan fluorescence in response to adding ligand. Changes in tryptophan fluorescence can be used to monitor structural and kinetic changes in proteins (Chen & Barkley, 1998), and has previously been used

to study ligand binding in other LeuT family members, for example vSGLT (Veenstra *et al*, 2002) and BetP (Ge *et al*, 2011).

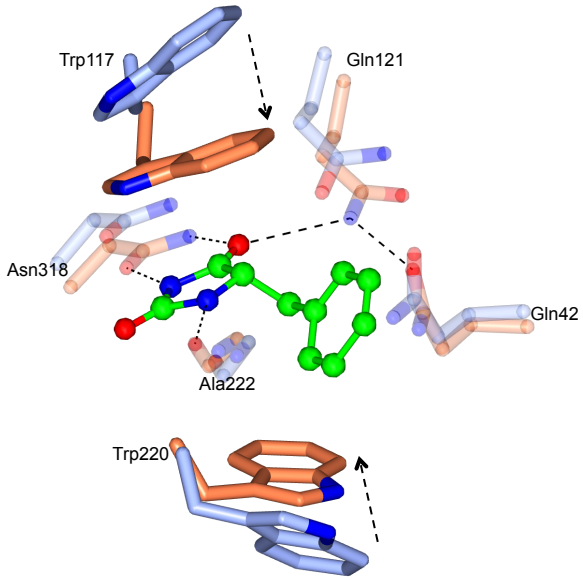


Figure 3.4 The structural rearrangements in the binding site upon binding of L-benzylhydantoin in wild-type Mhp1.

The residues Trp117, Trp220, Gln42, Gln121, Ala222, Asn318 and Gly219 from the crystal structures of the inward-open structure (PDB 2JLN, displayed in light blue) and the occluded structure (PDB 4D1B, displayed in coral) are displayed as cylinder bonds. The two crystal structures were superposed onto each another in the molecular graphics programme CCP4mg (McNicholas *et al*, 2011). The ligand L-BH is shown in a ball-and-stick representation in green, around which the residues Trp117 and Trp220 close up around in the occluded structure. The residues Gln42, Gln121, Asn318 and Gly219 participate in the putative hydrogen bonding network in order to stabilise the hydantoin moiety, and are displayed as partially transparent for clarity. Nitrogens are shown in blue and oxygens are shown in red. The black dashed arrows indicate the directions of the movement of Trp117 and Trp220.

In previously published structures of Mhp1, the structural rearrangements in the ligand-binding site upon binding of L-BH appear to involve two tryptophan residues, Trp117 and Trp220 (Weyand *et al*, 2008). In the transition from the ligand-free outward-open form to the occluded form with L-BH bound, the hydantoin moiety of the ligand packs face-to-face against Trp117, while Trp220 reorients to form end-on interactions with the benzyl-moiety of the ligand (Figure 3.4). This involves distinct changes in the environment of the tryptophan residues and probably underlies the fluorescence change enabling studies of the ligand binding in Mhp1.

Monitoring the quenching of tryptophan fluorescence upon binding of L-BH in the presence of varying concentrations of sodium chloride has previously been used to demonstrate that Mhp1 binds L-BH in a sodium-dependent manner (Weyand *et al*, 2008; Jackson, 2012). In the presence of 15 mM sodium chloride, the affinity for L-BH is greatly increased, resulting in a lowering of the apparent K_d from 0.737 ± 0.063 mM at 0 mM added sodium chloride to 0.036 ± 0.0014 mM at 15 mM added sodium chloride (Jackson, 2012).

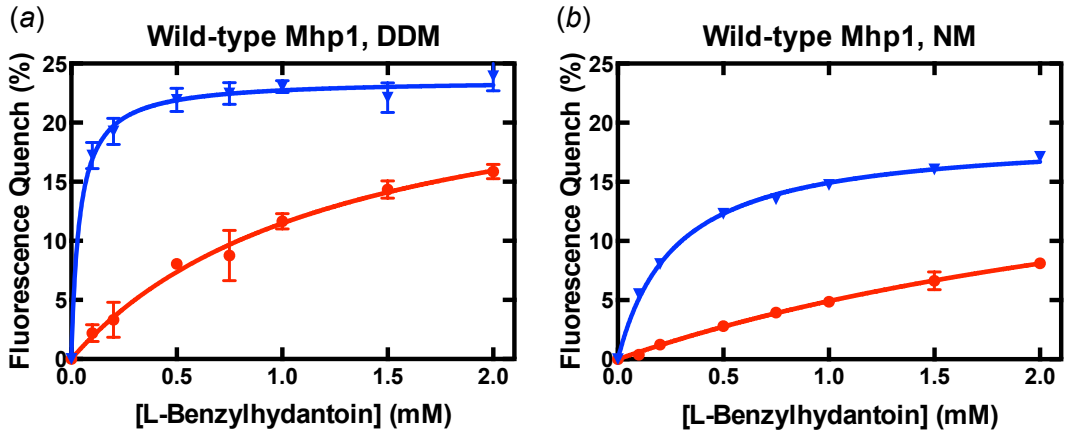


Figure 3.5 Representative steady-state tryptophan fluorescence quenching assay of the binding of L-BH to wild-type Mhp1, solubilised in DDM and NM.

The figure shows the quench of tryptophan fluorescence in the excitation region 310-360 nm upon the addition of L-BH (0-2 mM) in the presence of 0 mM (red line) or 15 mM (blue line) added sodium chloride. The error bars display the standard error of the mean from three replicates of each titration. Mhp1 samples (in 10 mM Tris-HCl pH 7.6, 2.5% v/v glycerol, 0.05% w/v or 0.5% w/v NM) were diluted to 140 µg/ml using buffers containing 50 mM Tris-HCl pH 7.6, 2% v/v DMSO and 0.05% w/v or 0.5% w/v NM. The fluorimetry buffer also contained 0/15 mM NaCl and 140/125 mM choline chloride.

The same well-established assay was used in this Thesis to monitor the functionality of the purified Mhp1 protein samples. Figure 3.5 shows representative titration curves for preparations of wild-type Mhp1, demonstrating the significant increase in L-BH binding in the presence of sodium (Table 3.1). Since the reference values quoted above were derived from wild-type Mhp1 protein solubilised in DDM, a representative titration curve from wild-type Mhp1 solubilised in the same detergent is shown in Figure 3.5*a*. The kinetic parameters obtained in this Thesis fall approximately into the same range as previously obtained values (Table

3.1). This shows that protein with comparable functionality as previously demonstrated is being produced and can thus be compared to previously obtained results. The discrepancies are likely to arise from variations in protein purification procedure, the batch of inner membranes or, for example, the age of the protein when the assays were carried out. As the majority of this Thesis concerns the crystallisation of Mhp1, protein solubilised in NM was primarily produced. The titration curves for wild-type Mhp1 solubilised in NM are shown in Figure 3.5*b*, and will serve as a benchmark for Mhp1 mutants and seleno-L-methionine derivatised Mhp1 produced for crystallisation trials (3.3-3.4). The influence of detergent length on the functionality of Mhp1 will be further discussed in section 3.5.

Table 3.1 Apparent K_d and B_{max} values for wild-type Mhp1, solubilised in DDM and NM in comparison to previous reference values.

Mhp1	[Na^+]	App. K_d (mM)	B_{max} (%)
Wild-type Mhp1, DDM *	0	0.74 ± 0.063	18.90 ± 0.70
	15	0.036 ± 0.0014	21.80 ± 0.10
Wild-type Mhp1, DDM	0	1.27 ± 0.23	26.01 ± 2.40
	15	0.040 ± 0.0050	23.64 ± 0.33
Wild-type Mhp1, NM	0	3.66 ± 0.82	22.93 ± 3.70
	15	0.27 ± 0.015	18.95 ± 0.28

* Reference values from Jackson, 2012.

3.2.6 Discussion

The various quality controls that were performed on purified wild-type Mhp1 have been presented in this section. Using CD it was confirmed that correctly folded and thermally stable protein was produced. Mhp1 is comprised of twelve TM α-helices and the recorded CD spectrum of wild-type Mhp1 shows the characteristic peaks for an α-helical fold. Using CD, the thermal stability of Mhp1 was also assessed. The T_m of wild-type Mhp1 was estimated to be 50.3 °C. By performing a tryptophan fluorescence quenching assay, it was also determined that the wild-type Mhp1 that was purified for this project was functional in accordance with previous studies. The titrations revealed that the Mhp1 protein bound L-BH in a sodium-dependent manner, as seen in previous studies, with the affinity of Mhp1 for L-BH significantly increasing in the presence of 15 mM sodium (Weyand *et al*, 2008; Jackson, 2012).

3.3 Purification and characterisation of mutants of the Mhp1 protein for crystallographic studies

3.3.1 Introduction

Previously, a number of single-point mutants with altered ligand binding behaviour have been identified (Jackson, 2012), which are interesting targets for structural studies due to their kinetic properties. A number of these mutants have been purified and characterised as part of this project for X-ray crystallographic studies, in order to establish the structural basis behind the altered ligand binding behaviour in these mutants (Table 3.2).

Table 3.2 List of mutants purified and characterised for further structural studies.

The table is a replicated from Table 2.2 in section 2.2.1.

Mutant	Site of the mutation
Gln42Asn	Ligand
Gln121Asn	Ligand
Asn318Gly	Ligand
Ser312Ala	Sodium
Lys110Leu	Bundle/hash

The mutants are categorised by their location in the protein, either in the ligand-binding site, the sodium-binding site or on the interface between the bundle- and hash-motifs in the Mhp1 protein. The locations of these sites are illustrated in Figure 3.6, with respect to the bundle- and hash-motifs.

Only one mutant studied as part of this Thesis (Lys110Leu) was situated neither in the ligand- nor the sodium-binding site. The remaining mutants were located in either the ligand- (Gln42Asn, Gln121Asn, Asn318Gly) or sodium-binding (Ser312A) site, which are more clearly illustrated in Figure 3.7.

The mutants were purified as previously described, initially without any modification to the protocols that were used for the wild-type protein. The quality of the purified samples of the single-point mutants of the Mhp1 protein was assessed using CD and fluorescence quenching assays, as described for wild-type Mhp1 in section 3.2.

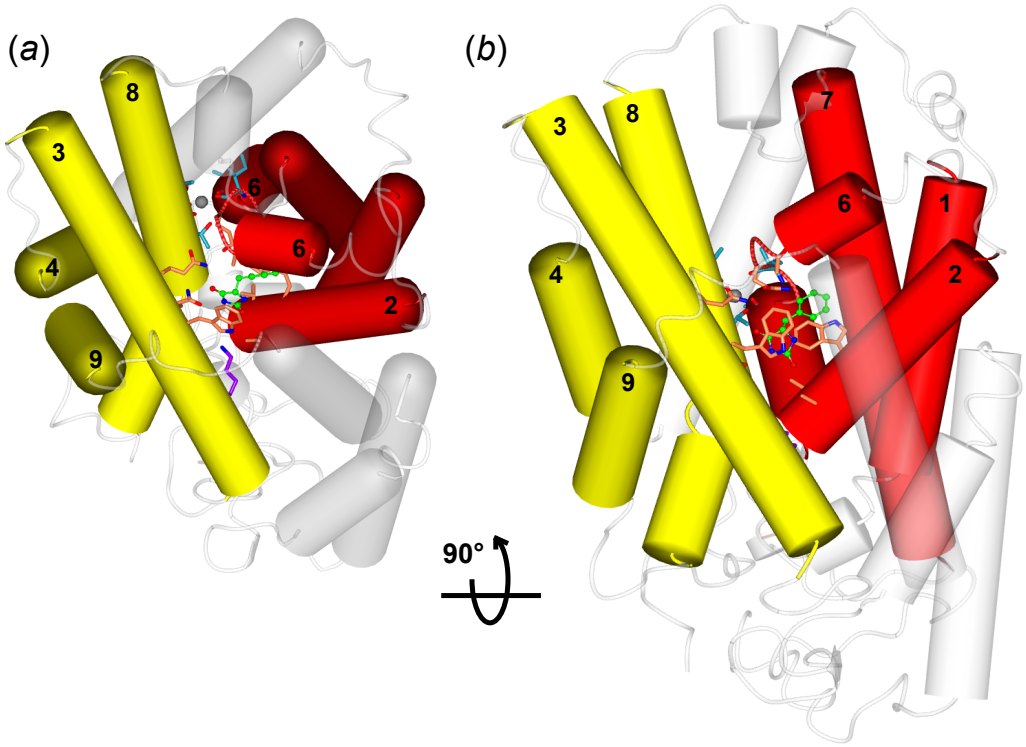


Figure 3.6 The bundle- and hash-motifs in the Mhp1 protein, displaying the locations of the single-point mutations.

The Mhp1 protein is illustrated with the bundle- and hash-motifs highlighted, in order to display the locations of the single-point mutations in the Mhp1 protein that are studied here. The helices comprising the bundle motif (TMHs 3, 4, 8 and 9) are displayed as yellow cylinders and the hash motif (TMHs 1, 2, 6 and 7) are displayed as red cylinders. The remaining helices and loops are displayed in semi-transparent grey for clarity. The residues participating in the ligand and sodium binding sites are displayed in coral (location of ligand binding site mutations) and cyan (location of sodium binding site mutant). In addition, the location of the residue Lys110Leu is displayed in purple (mutant between the bundle- and hash-motifs). The orientation of the protein in (*a*) is roughly the same as in Figure 3.7*a*, and the protein in (*b*) is rotated by 90°.

3.3.2 Purification of mutants of Mhp1

The same purification procedures used for wild-type Mhp1, as described in section 2.2.1, were applied for the single-point mutants of Mhp1 outlined in Table 3.2. As the mutants were produced for crystallisation trials and subsequent X-ray crystallographic studies, all the mutants were subjected to detergent exchange by NM during the purification procedure. SDS-PAGE and Western blotting analysis confirmed the presence of the purified mutant Mhp1 (Figure 3.8). For Gln121Asn, a more significant amount of a higher-molecular weight species was observed, compared to wild-type Mhp1 and the other mutants.

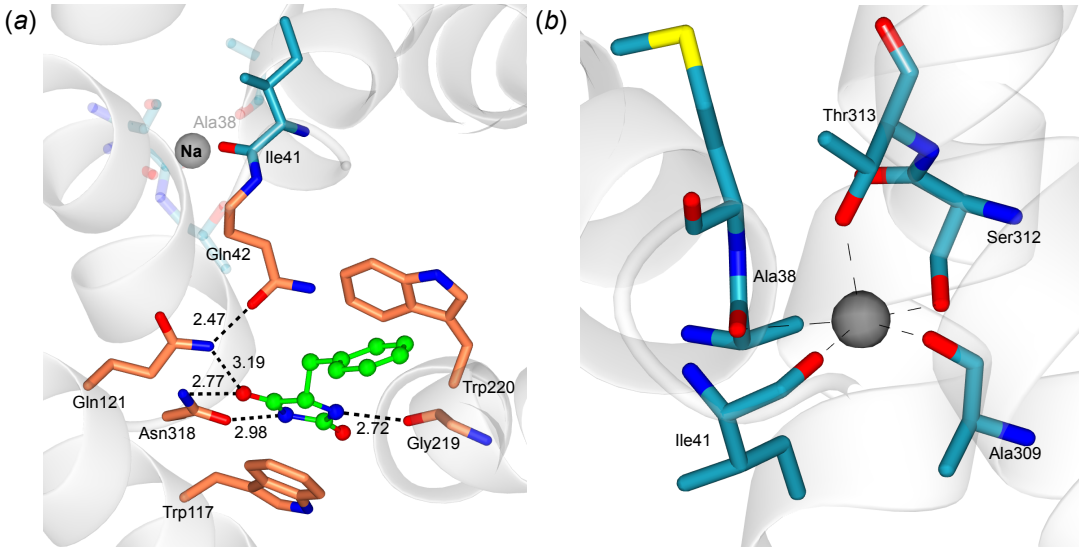


Figure 3.7 The ligand- and sodium-binding sites in the Mhp1 protein.

The figure illustrates the ligand-binding site of Mhp1 in complex with L-BH (*a*) and the nearby sodium binding site (*b*). The orientation of the protein and the colour scheme for the individual residues in (*a*) are roughly the same as in Figure 3.6*a*. In (*a*), the ligand binding site residues are highlighted in coral, the ligand L-BH in green and the nearby sodium as a grey sphere. Hydrogen bonds between the L-BH and nearby residues are displayed as black dashed lines and stated in units of Å. The sodium binding site residues are displayed in semi-transparent cyan, apart from Ile41, which is displayed in cyan as a reference for the orientation of the sodium binding site in (*b*). In (*b*), the sodium binding site residues are displayed in cyan with the coordinating bonds between the residues and the sodium as black dashed lines. In both (*a*) and (*b*), the non-carbon atoms are coloured red (oxygen), blue (nitrogen) and yellow (sulphur). The ribbon diagram of the whole protein is displayed in semi-transparent grey.

As the bands are closer to the actual molecular weight of Mhp1 (54.6 kDa), this suggests that there is a greater proportion of either fully denatured protein present in Gln121Asn samples. Alternatively it could be promoting the formation of a protein dimer as the higher molecular weight species is also close to approximately twice the molecular weight of the main band.

The purified samples of the Mhp1 mutants were all subjected to analysis by CD, in order to confirm that they were folded and suitable for further characterisation and structural studies. All mutants gave rise to a characteristic CD spectrum for an α-helical protein (see Appendix B for CD graphs). This suggests that the single-point mutations do not affect the ability of the protein to fold into its native α-helical fold, with respect to wild-type Mhp1. In addition, the thermal stability of the mutants was also studied using CD, as described in section 3.2.4. It appeared that each mutation increased the thermal stability of Mhp1, with respect to the wild-type (Table 3.3, see Appendix B for CD thermal melt curves). As discussed in section

3.2.6, this indicates that according to the T_m of the Mhp1 mutants, these are suitable targets for crystallisation trials.

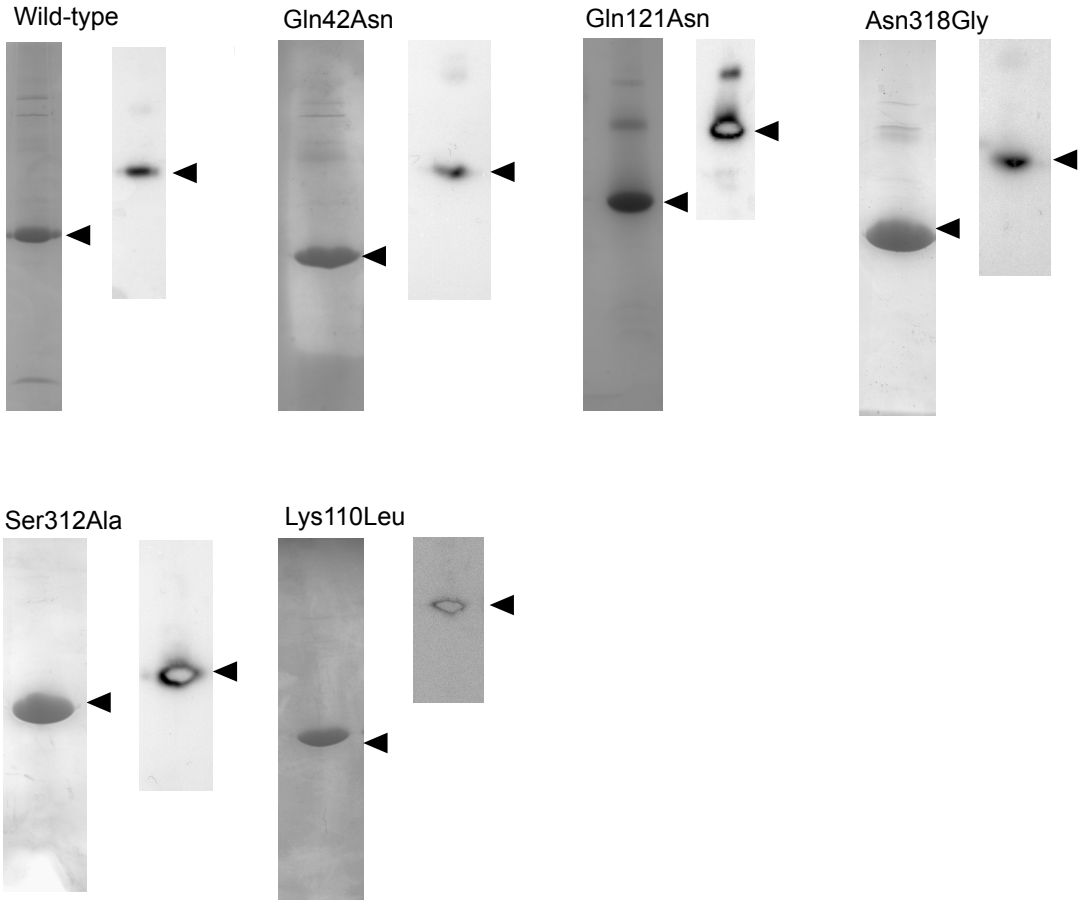


Figure 3.8 Comparison of the purifications of Mhp1 mutants.

For each mutant, the left-hand cut-out lane shows the purified Mhp1 protein on a SDS-PAGE gel and the right-hand cut-out lane shows the same sample on the corresponding Western blot (see Appendix A for full gels). The arrows indicate the position of the band that correspond to the Mhp1 protein. For each mutant, the SDS-PAGE and Western blot confirm that pure Mhp1 mutant protein was produced for further characterisation and structural studies. In some cases, higher molecular weight species were observed, and this is potentially protein that has been completely denatured by SDS.

The apparent stabilisation may be a result of the mutations restricting conformational changes and rigidifying the protein in solution. This finding does not correlate with the suggestion that the mutation Gln121Asn may destabilise the protein and result in a greater proportion of fully denatured protein on an SDS-PAGE gel, but may support the idea that the mutation may promote the formation of dimers, due to the potentially increased rigidity of the protein as suggested by the CD analysis of the thermal stability. As it is difficult to correlate thermal stability and resistance to denaturation by

SDS, the apparent discrepancy may be attributed to differences in sample and SDS-PAGE gel preparation.

3.3.3 Assessing the change in ligand binding in Mhp1 caused by single-point mutations

The steady-state tryptophan fluorescence quenching assay used to study the ligand binding was used to investigate how the single-point mutations in Table 3.3 affected the ligand binding ability of Mhp1. As described in 3.2.5, L-BH was titrated into a solution of the protein in the presence of various concentrations of sodium chloride (Figure 3.9, Table 3.4). Since all mutants were prepared for crystallisation trials, they were subjected to detergent exchange into NM, and will therefore be compared to wild-type Mhp1 purified in the same detergent (Figure 3.9*a*).

Table 3.3 Summary of the melting points obtained from thermal melts of circular dichroism of mutants of Mhp1, purified in NM.

Mutant	Site of the mutation	Melting temperature (°C)
Wild-type	Ligand	50.3
Gln42Asn	Ligand	65.3
Gln121Asn	Ligand	74.0
Asn318Gly	Ligand	62.4
Ser312Ala	Sodium	62.0
Lys110Leu	Bundle/hash	61.2

Two of the residues in the ligand-binding site, Gln42 and Gln121, were probed by changing the glutamine to the close analogue asparagine. Gln121 has been suggested to form a hydrogen bond with one of the carbonyl oxygens on the hydantoin moiety of the ligand, while Gln42 could be forming a hydrogen bond with Gln121, thus participating in the hydrogen-bonding network that stabilises the hydantoin moiety in the binding site (Weyand *et al*, 2008; Simmons *et al*, 2014). In addition, Gln42 is situated in close proximity to the nearby sodium binding site, with the carbonyl oxygen of the adjacent residue Ile41 participating in the coordination of the sodium atom.

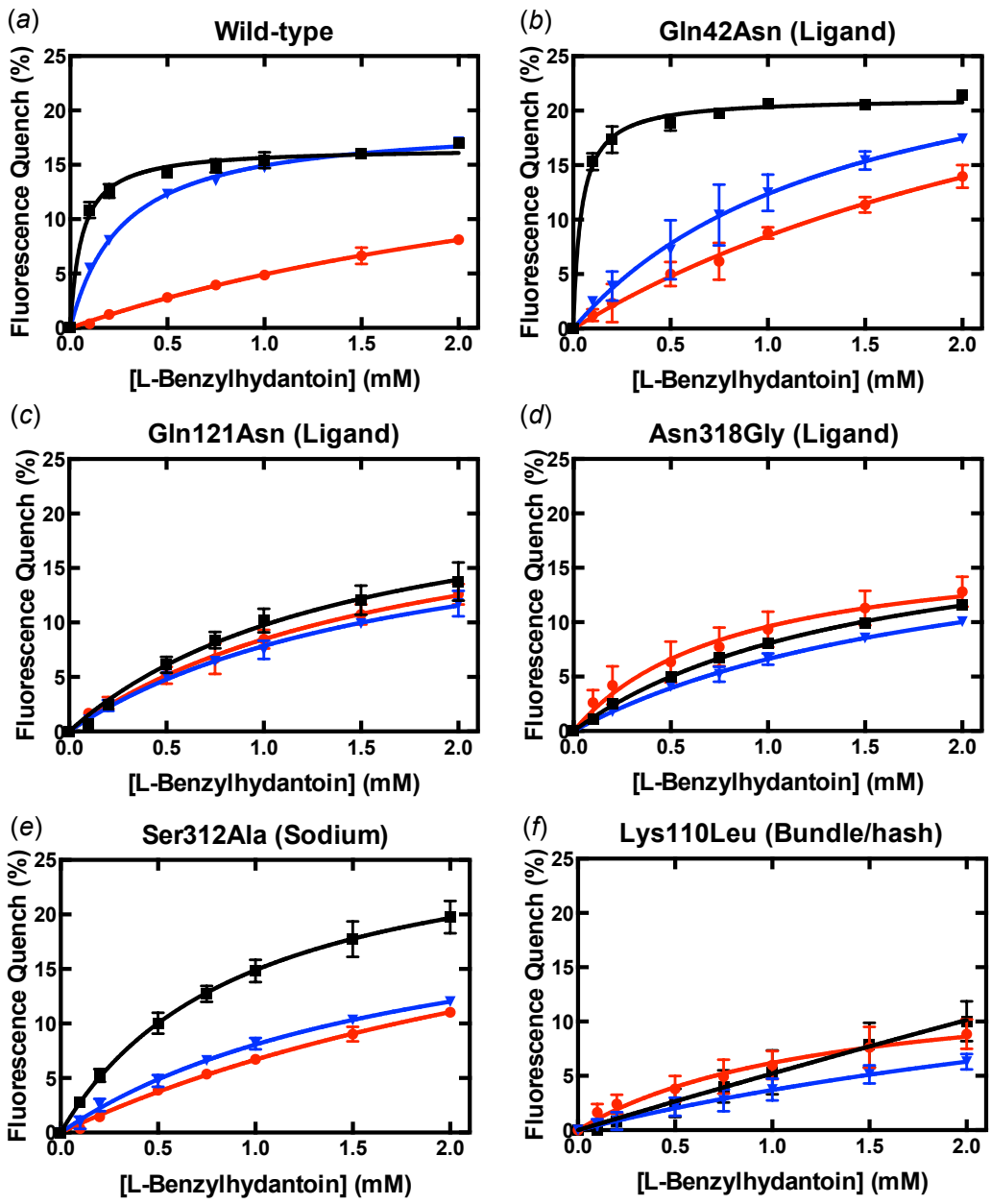


Figure 3.9 The binding of L-BH to wild-type and mutant Mhp1, solubilised in NM.

The figure shows the fluorescence quench during titrations of L-BH (0-2 mM) with the wild-type Mhp1 (*a*), as well as the Mhp1 mutants Gln42Asn (*b*), Gln121Asn (*c*), Ser312Ala (*d*), Asn318Gly (*e*) and Lys110Leu (*f*). At the top of each graph, the respective position of each mutation is indicated in parentheses. The titrations were performed in the presence of 0 mM (red), 15 mM (blue) and 140 mM (black) added sodium chloride. Mhp1 samples (in 10 mM Tris-HCl pH 7.6, 2.5% v/v glycerol, 0.5% w/v NM) were diluted to 140 µg/ml using buffers containing 50 mM Tris-HCl pH 7.6, 2% v/v DMSO, 0.5% w/v NM, 0/15/140 mM NaCl and 140/125/0 mM choline chloride. The error bars display the standard error of the mean from three replicates of each titration.

Plausibly, Gln42 and Gln121 are in close interplay with the ligand and each other. In addition, Gln42 and Gln121 are potentially also in interplay with the

nearby sodium binding site and mutation of these residues may also be affecting the sodium binding.

The tryptophan fluorescence quenching assay for the ligand binding site mutants Gln42Asn and Gln121Asn (Figure 3.9 *b*, *c*) shows that the binding of L-BH is clearly perturbed by the mutations to the one carbon shorter asparagine residues. In the absence of sodium, the low-affinity binding of L-BH is similar with respect to wild-type.

Table 3.4 Summary of the apparent K_d and B_{max} values for the mutants of Mhp1, purified in NM.

Mutant	$[Na^+]$	App. K_d (mM)	B_{max} (%)
WT	0	3.66 ± 0.82	22.93 ± 3.70
	15	0.27 ± 0.015	18.95 ± 0.28
	140	0.059 ± 0.0071	16.56 ± 0.28
Gln42Asn	0	3.47 ± 1.13	38.01 ± 8.78
	15	1.41 ± 0.33	29.81 ± 3.81
	140	0.041 ± 0.0040	21.19 ± 0.23
Gln121Asn	0	1.81 ± 0.36	23.81 ± 2.73
	15	1.80 ± 0.28	21.91 ± 1.98
	140	1.50 ± 0.30	24.37 ± 2.65
Asn318Gly	0	0.82 ± 0.25	17.46 ± 2.33
	15	2.11 ± 0.30	20.59 ± 1.77
	140	1.55 ± 0.12	20.46 ± 0.84
Ser312Ala	0	3.76 ± 0.69	31.79 ± 4.236
	15	1.90 ± 0.24	23.43 ± 1.80
	140	0.94 ± 0.12	28.94 ± 1.62
Lys110Leu	0	1.28 ± 0.52	14.13 ± 3.00
	15	4.77 ± 3.23	21.35 ± 11.1
	140	n/a	n/a

In Gln42Asn, the binding at 15 mM sodium is greatly reduced, but the presence of almost 10-fold more sodium can restore the sodium-dependent binding of L-BH. On the contrary, the binding of L-BH by Gln121Asn does not respond to added sodium and remains at approximately the same level as binding at 0 mM sodium. This could be due to the putative hydrogen-bond to one of the carbonyl oxygen of the hydantoin moiety being broken by

shortening of the Gln121 to Asn121 (Figure 3.4), and this bond being essential for binding of the ligand.

The sodium binding site was probed by studying the mutation Ser312Ala (Figure 3.9*e*). Ser312Ala showed a similar behaviour to Gln42Asn, where the sodium dependent binding is initially reduced at 15 mM with respect to wild-type Mhp1, but is partially restored by the addition of excess sodium. Asn318Gly showed a similar behaviour as Gln121Asn, where the affinity for L-BH remained approximately as low as in the presence of 0 mM added sodium.

One mutant that was not situated in either the ligand-, or the sodium-binding site was also studied, namely Lys110Leu. (Figure 3.9*f*). The mutation Lys110Leu is situated on the interface between the bundle-and hash motifs, and showed that the affinity for L-BH was lowered as the sodium concentration was increased. The linearity of the plot corresponding to 140 mM sodium in Figure 3.9*f* means that the trace could not be fitted to a hyperbola. Similar to Gln121Asn and Asn318Gly, the affinity of Lys110Leu to L-BH remains low even when sodium is added.

3.3.4 Discussion

In this section, the purification and characterisation of a number of single-point mutants of Mhp1 are described, comprising mutations that are found in the ligand-binding site (Gln42Asn, Gln121Asn, Asn318Gly), in the sodium-binding site (Ser312Ala), as well as on the interface between the bundle-and-hash motifs (Lys110Leu). Overall, the single-point mutants were subjected to the same purification and characterisation procedures as previously described for wild-type Mhp1, and resulted in the preparation of pure and stable samples of the Mhp1 mutant protein. The estimated T_m for each of the mutants suggests that stable protein was produced, and that the mutations stabilised the Mhp1 protein with respect to the wild-type. The tryptophan fluorescence quenching assays show how the ligand binding properties of Mhp1 are affected by single-point mutations, outlined in Figure 3.9 and Table 3.2. In addition, these observations serve as an initial indication that different conformations might be adopted by the Mhp1 mutants and that changes in conformational dynamics might be affected by the mutations.

These changes could be due to the fact that essential stabilising interactions are broken, as in the case of Gln121Asn. In the crystal structure of wild-type

Mhp1 in its occluded state with L-BH bound (Weyand *et al*, 2008; Simmons *et al*, 2014), Gln121Asn appears to be in hydrogen bonding distance to one of the carbonyl oxygen atoms of the hydantoin moiety of L-BH. By shortening the residue to asparagine, the hydrogen bond may be disrupted and thus, L-BH cannot be stabilised in its binding site. Alternatively, the mutant Gln121Asn could be locked in an different low-affinity conformation to which the ligand is less likely to bind.

In other cases, the changes could be attributed to changes to the dynamics and stability of various conformationals. In the presence of 0 mM sodium, Gln121Asn, Asn318Gly (ligand-binding site) and Lys110Leu (interface between the bundle-and-hash motifs) have higher affinity for L-BH than the wild-type protein. This could be due to the mutant protein being stabilised in an outward-open form, promoting binding of L-BH.

In all mutants, the sodium-dependent ligand binding is diminished to a certain extent. In the case of Gln42Asn (ligand-binding site) and Ser312Ala (sodium-binding site), the partially lost ligand binding affinity could be restored upon the addition of excess sodium. This suggests that the outward-open conformation might be destabilised, shifting the conformational equilibrium towards a lower-affinity, inward-open form. The addition of almost 10-fold more sodium is able to shift the equilibrium back to a higher-affinity form and restore ligand binding. On the contrary, the mutants Gln121Asn and Asn318Gly (ligand-binding site) and Lys110Leu (interface between the bundle and hash motifs) appear to remain in a low-affinity form and do not respond to the same extent as Gln42Asn and Ser312Ala to the addition of excess sodium. This suggests that these mutants may be locked into an inward-open, low-affinity conformational state.

3.4 Purification and characterisation of seleno-L-methionine derivatised Mhp1 for crystallographic studies

3.4.1 Introduction

The collection of previously published crystal structures of Mhp1 comprise mainly crystal structures of Mhp1 in its occluded form, in complex with various ligands, described in section 1.3.2 (Weyand *et al*, 2008; Simmons *et al*, 2014). These structures have, together with the structure of Mhp1 in the outward-open form, given considerable insight into the structural basis of the ligand binding step of the transport mechanism. On the contrary, there is

only one crystal structure of Mhp1 in the inward-open state (Shimamura *et al*, 2010), and subsequently much less is known about the transition from the ligand-bound occluded state to the inward-open form. Therefore, it is of great interest to obtain more structural information on the inward-open form to gain more insight about this transition.

The previously published structure of the inward-open state of Mhp1 was obtained from crystals that were produced from seleno-L-methionine derivatised protein, under different crystallisation conditions compared to the conditions that gave rise to the wild-type structure in the outward-open and occluded forms (Shimamura *et al*, 2010). The resulting crystals also crystallised in a different space group, $P6_1$ (Table 3.5). As shown in section 1.3.2, the ligand-binding site in the structure of the inward-open protein contained a positive difference density feature, indicating that this site may have been occupied by a molecule that could have potentially been stabilising that state and thus promoting its crystallisation. This could have been a carry-over product from the purification of the methionine auxotrophic cells (Peter Henderson, personal communication). In order to gain more insight into the structural basis behind the transition from the occluded to the inward-open state, it was attempted to reproduce a crystal structure of this conformational state at a higher resolution than previously obtained (3.8 Å). This could also provide a new opportunity to identify the molecule that stabilises the inward-open state. The fact that this ligand putatively locks the protein into an inward-open form suggests that this molecule could serve as a template for inhibitor design (Shimamura *et al*, 2010). Regardless of whether the stabilising molecule will be present or not, it is also of interest to study the inward-open structure in the presence of the hydantoin ligands that Mhp1 has previously been co-crystallised with, in order to investigate how the ligands interact with Mhp1 on the route to exit the transporter on the other side of the membrane.

3.4.2 Purification and characterisation of seleno-L-methionine derivatised Mhp1

Inner membrane preparations of seleno-L-methionine derivatised Mhp1 were prepared according to previously established protocols and provided by Mr. David Sharples (Shimamura *et al*, 2010). The inner membrane preparations were then subjected to the same purification procedures as used for wild-type Mhp1. As the seleno-L-methionine derivatised Mhp1 was mainly

prepared to perform crystallisation trials, the detergent was exchanged from DDM to NM during the washing of the Ni-NTA resin.

Table 3.5 Crystallisation conditions used to obtain the previously published crystal structures of the various conformations of Mhp1.

	Orthorhombic	**Hexagonal**
Conformation of Mhp1 in the crystal structure	Outward-open/ occluded	Inward-open
PDB ID	2JLN, 4D1A, 4D1B, 4D1C, 4D1D	2X79
Detergent	NM	NM
PEG 300 (% v/v)	27 - 35	24 - 28
NaCl (mM)	100	100
Bicine pH 9.0 (mM)	-	100
$NaPO_4$, pH 7.0 (mM)	100	-
Reservoir volume (µl)	300	400
Temperature (°C)	20	4
Crystal morphology	Needle-shaped	Hexagonal
Space group	$P2_12_12_1$	$P6_1$

A single band corresponding to His_6-tagged, seleno-L-methionine derivatised Mhp1 was obtained on a Western blot (Figure 3.10). The main band corresponding to Mhp1 migrated approximately as a 45 kDa protein, compared to the molecular weight standard proteins, which is consistent with the fact that the seleno-L-methionine derivatised Mhp1 has a larger molecular weight than wild-type protein (wild-type Mhp1: 54.6 kDa, seleno-L-methionine derivatised Mhp1 with twenty selenium atoms: 55.5 kDa), and that the presence of selenium may cause a difference in resistance to the denaturation by SDS. The few faint higher molecular weight bands from the eluted, purified protein sample on the SDS-PAGE gel indicates that there might be higher-molecular weight impurities or oligomers present, or incompletely denatured protein molecules also present (see section 3.2.2).

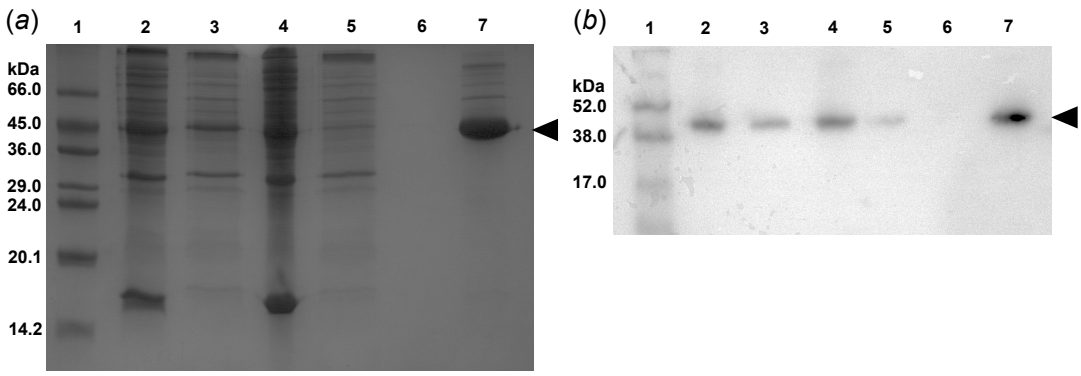


Figure 3.10 The purification of seleno-L-methionine derivatised Mhp1, solubilised in NM.

The lanes in the SDS-PAGE and Western blot contain the same samples, with the samples for Western blotting being diluted 1:14 times, and contain: 1) molecular weight ladder with the molecular weights marked in kDa left of the gels; 2) solubilisation mixture before ultracentrifugation; 3) supernatant and 4) pellet after ultracentrifugation; 5) unbound material in the flow-through after binding to the Ni-NTA resin; 6) wash buffer of the Ni-NTA resin; and 7) eluted sample from the Ni-NTA resin using 200 mM imidazole.

The fact that higher molecular weight species are not seen on the Western blot suggests that these impurities may not be Mhp1, but the 1:14 dilution of the samples for Western blotting may have been too low to allow detection of any small quantity of oligomeric or partially unfolded Mhp1 molecules. Mass spectrometry analysis was carried out by Dr Anton Calabrese and showed that one single species of Mhp1 was present in the sample, of the expected larger molecular weight with twenty selenium atoms incorporated (55.5 kDa).

The fluorescence quenching assay shows that, in comparison to wild-type Mhp1, the affinity of seleno-L-methionine derivatised Mhp1 for L-BH is slightly increased in the presence of 0 mM added sodium, but less so in the presence of 15 mM added sodium (Figure 3.11, Table 3.6). At 0 mM added sodium, this effect could be due to the sodium binding site being stabilised by the sodium coordinating residue Met39, thus promoting an outward-open conformation with high affinity for L-BH. Alternatively, the twenty methionine residues that are evenly distributed throughout the protein may overall stabilise and shift the dynamics of the protein in a way that favours binding of L-BH.

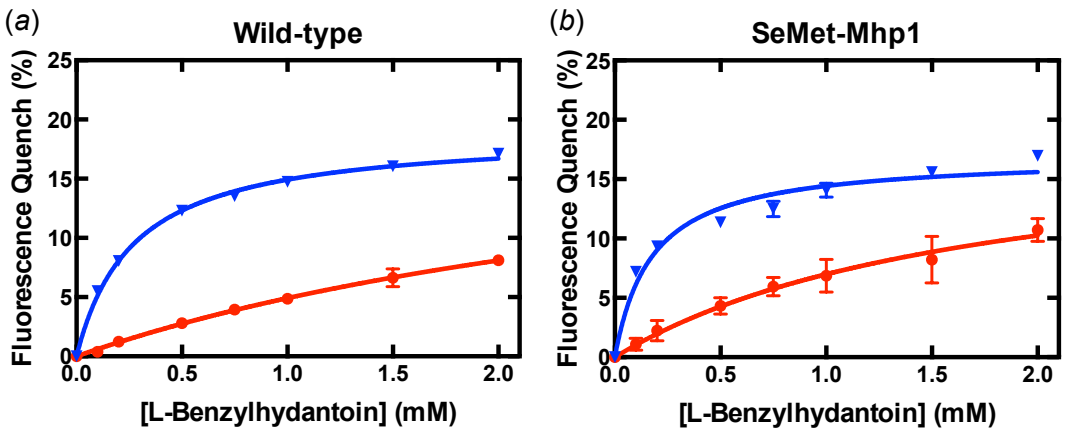


Figure 3.11 The binding of L-BH to seleno-L-methionine derivatised Mhp1, solubilised in NM.

The figure shows the fluorescence quench during titrations of L-BH (0-2 mM) with the wild-type Mhp1 (*a*) and seleno-L-methionine derivatised Mhp1 (*b*). The titrations were performed in the presence of 0 mM (red) and 15 mM (blue) added sodium chloride. Mhp1 samples (in 10 mM Tris-HCl pH 7.6, 2.5% v/v glycerol, 10 mM DTT, 0.5% w/v NM) were diluted to 140 µg/ml using buffers containing 50 mM Tris-HCl pH 7.6, 2% v/v DMSO, 10 mM DTT, 0.5% w/v NM, 0/15 mM NaCl and 140/125 mM choline chloride. The error bars display the standard error of the mean from three replicates of each titration.

Table 3.6 Apparent K_d and B_{max} values for seleno-L-methionine derivatised Mhp1, solubilised in NM.

Mhp1	[Na^+]	App. K_d (mM)	B_{max} (%)
SeMet	0	1.75 ± 0.53	19.22 ± 3.35
	15	0.18 ± 0.023	16.97 ± 0.50
WT	0	3.66 ± 0.82	22.93 ± 3.70
	15	0.27 ± 0.015	18.95 ± 0.28

3.4.3 Discussion

Seleno-L-methionine derivatised Mhp1 was successfully purified, and the subsequent mass spectrometry analysis of the protein confirmed the presence of one population of the derivatised protein, with twenty selenium atoms incorporated. A small amount of impurities appeared on the SDS-PAGE gel, and unfortunately CD analysis of the secondary structure and thermal stability was not performed that could confirm the structural integrity and melting point of the protein. However, the functional analysis of the seleno-L-methionine derivatised Mhp1 showed that the protein was functional, which suggests that the protein was indeed properly folded. The

slight decrease in affinity for L-BH in the presence of 15 mM sodium suggests that the sodium binding is affected by the incorporation of seleno-L-metionine into Mhp1, which is consistent with the fact that the carbonyl oxygen of Met39 is involved in the coordination of the sodium. Overall, the quality controls indicate that functional and relatively pure protein was produced, which will be taken forward to crystallisation trials.

The tryptophan fluorescence quenching assay shows only a slight decrease in binding affinity for L-BH to the seleno-L-methionine derivatised Mhp1 compared to wild-type protein (Figure 3.11, Table 3.6), and does not give a direct indication of whether a different conformational state may be present or whether there is another ligand that co-exists in the protein solution and binds to the seleno-L-methionine derivatised protein. On the other hand, the raw fluorescence intensity values were found to be approximately four times lower for seleno-L-methionine derivatised Mhp1 compared to wild-type Mhp1, indicating that the initial fluorescence properties of the derivatised protein were perturbed, potentially due to the protein adopting a different conformation or due to the presence of an additional ligand, as previously suggested by structural studies of the seleno-L-methionine derivatised Mhp1 (Shimamura *et al,* 2010).

3.5 Purification and characterisation of Mhp1 in maltoside detergents of varying length for scattering studies

3.5.1 Introduction

The presence of detergent micelles greatly complicates SAXS studies of membrane proteins, as they contribute significantly to the total scattering signal and also make it difficult to achieve accurate buffer subtraction. SEC-SAXS has been suggested as a method to circumvent the problems imposed by the presence of detergents in SAXS studies of membrane proteins (Weyand *et al*, 2008; Berthaud *et al*, 2012b; Simmons *et al*, 2014). In order to investigate the applicability of SEC-SAXS to Mhp1, it was prepared in detergents with varying alkyl chain lengths.

The detergents were chosen on the basis of previous results, where Mhp1 has been successfully purified in DDM (Weyand *et al*, 2008; Jackson, 2012; Simmons *et al*, 2014), as well as in NM for crystallisation trials as part of this Thesis and for previously published crystal structures. NM differs from DDM in that its aliphatic hydrocarbon chain is three carbons shorter. Due to the success of the purification of Mhp1 in DDM and NM, the detergents *n*-

undecyl-*β*-D-maltopyranoside (UM) and *n*-decyl-*β*-D-maltopyranoside (DM), which position themselves in-between DDM and NM in hydrocarbon chain length were assayed (Figure 3.12).

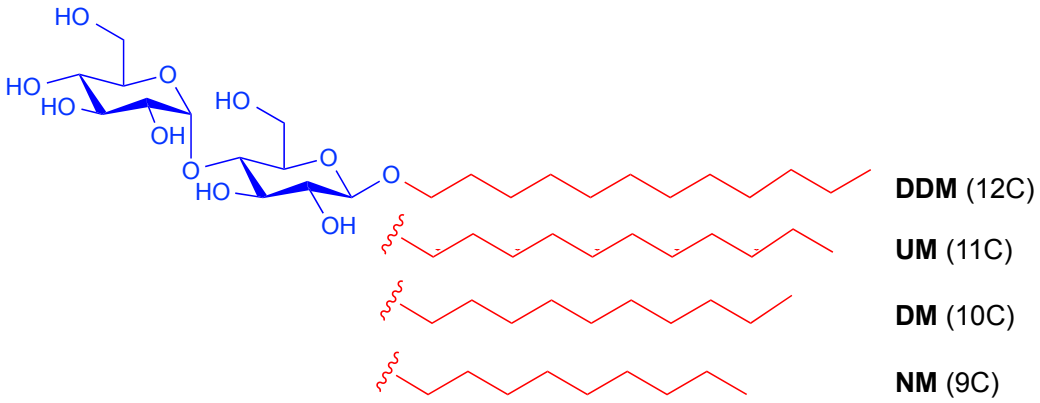


Figure 3.12 The structure of the maltoside detergents used in this project.

The common maltoside head-group structure is illustrated for the four detergents used in this project (blue), with the length of each of the aliphatic hydrocarbon chains (red) indicated by the number of carbons in parentheses (*n*C).

3.5.2 Purification and characterisation of wild-type Mhp1 solubilised in maltoside-detergents with varying aliphatic chain length

As described in section 2.2.1, detergent exchange was achieved by adding an additional washing step of the Ni-NTA resin. The first washing step involves washing the resin with buffer containing DDM, in order to initially wash off impurities and unbound protein. The second washing step was performed using buffer that contained the detergent of interest, which was NM in the case when protein was prepared for crystallisation trials. The samples of Mhp1 solubilised in various detergents were prepared in two batches. Each batch was used to purify two different Mhp1-detergent complexes, and was initially solubilised in DDM. After incubation with the Ni-NTA resin, the resin was divided into two parts and the two parts were subjected to separate washes with buffers containing different detergents (Figure 3.13).

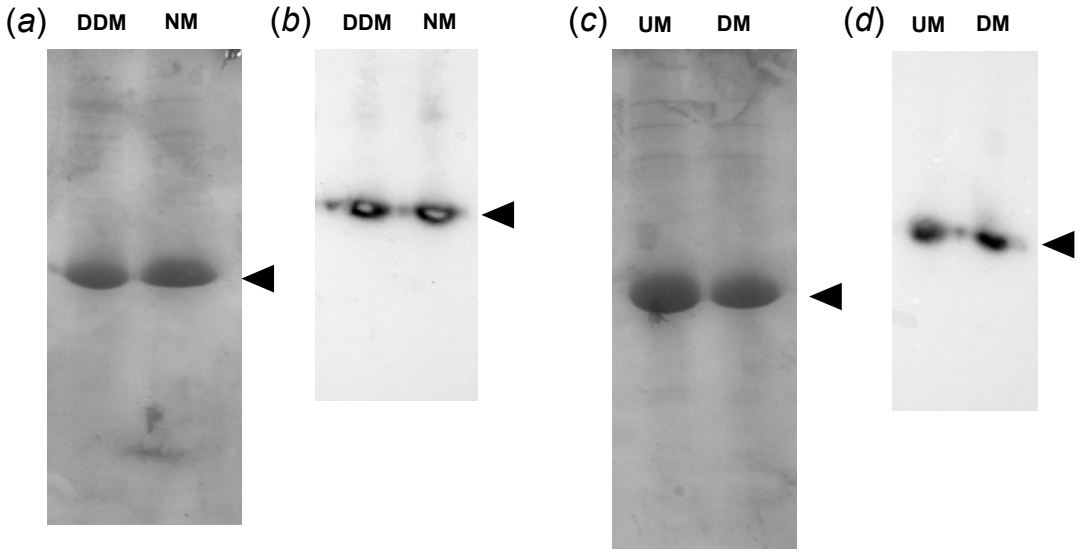


Figure 3.13 The purification of wild-type Mhp1 solubilised in maltoside detergents with varying aliphatic chain length.

The cut-out lanes in the SDS-PAGE gel (*a*) and corresponding Western blot (*b*) for batch 1, used for the purification of wild-type Mhp1 solubilised in DDM and NM. The lanes in the SDS-PAGE gel (*c*) and corresponding Western blot (*d*) for batch 2, used for the purification of wild-type Mhp1 solubilised in UM and DM. See Appendix B for the full gels.

Analysis by Western blotting confirms the presence of Mhp1 in each of the eluted samples, which appear to have been prepared with a minimal amount of impurities as seen on the SDS-PAGE gels. From the gels, it appears that the solubilising detergent does not significantly affect the stability or homogeneity of the purified protein.

The functionality of the wild-type Mhp1 solubilised in the maltoside detergents with varying aliphatic chain length was assessed using the fluorescence quenching assay. Overall, the same sodium-dependent binding is seen as described in 3.2.5, with the affinity for L-BH increasing in the presence of 15 mM sodium (Figure 3.14). The general trend is that the apparent K_d and B_{max} increase the shorter the reconstituting maltoside detergent is (Table 3.7).

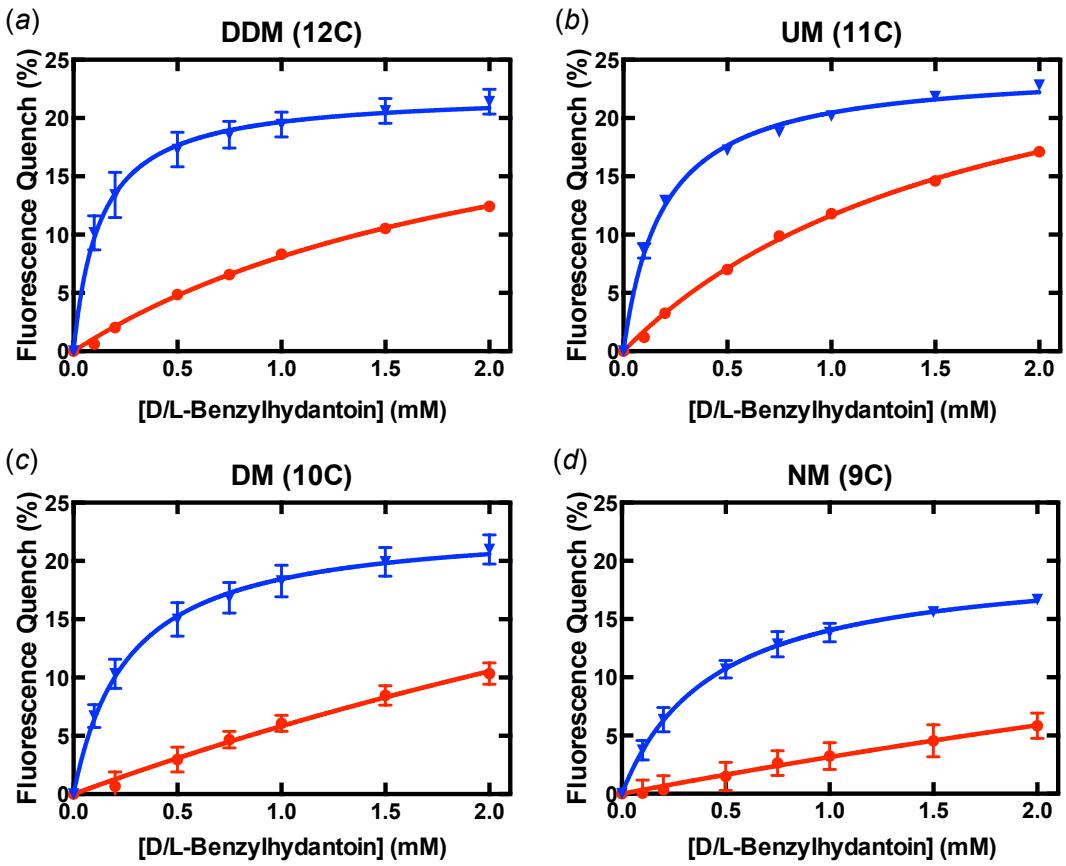


Figure 3.14 The binding of L-BH to wild-type Mhp1, solubilised in maltoside detergents of varying length.

The figure shows the fluorescence quench during titrations of L-BH (0-2 mM) with the wild-type Mhp1, solubilised in the maltoside detergents DDM (*a*), UM (*b*), DM (*c*) and NM (*d*). The number of carbons (*n*C) in the hydrocarbon chain is indicated next to the abbreviations of each detergent. The titrations were performed in the presence of 0 mM (red) and 15 mM (blue) added sodium chloride. Mhp1 samples (in 10 mM Tris-HCl pH 7.6, 2.5% v/v glycerol, 10 mM DTT, 0.5% w/v NM) were diluted to 140 µg/ml using buffers containing 50 mM Tris-HCl pH 7.6, 2% v/v DMSO,% w/v detergent as indicated for the storage buffer in Table 2.3, 0/15 mM NaCl and 140/125 mM choline chloride. The error bars display the standard error of the mean from three replicates of each titration.

The shorter the detergent the protein is reconstituted into, the smaller would be the putative detergent corona surrounding the hydrophobic belt around the protein. The effect of this is that more of the protein may become exposed to the solvent than with a longer detergent, potentially causing the formation of transient protein-protein interactions that might disrupt the binding of ligands. Another possibility is that a smaller detergent corona might create more strain around the protein, making it more rigid and less mobile. This is a more likely cause for the observed decrease in binding

affinity, as the tighter detergent corona might be restricting the conformational changes that are required to occur for ligand binding.

Table 3.7 Summary of the apparent K_d and B_{max} values for Mhp1 solubilised in maltoside detergents of varying length.

Detergent	$[Na^+]$	App. K_d (mM)	B_{max} (%)
DDM	0	2.32 ± 0.20	26.92 ± 1.49
	15	0.13 ± 0.016	22.2 ± 0.55
UM	0	1.76 ± 0.12	32.14 ± 1.30
	15	0.19 ± 0.010	24.33 ± 0.30
DM	0	8.26 ± 6.68	54.06 ± 36.96
	15	0.26 ± 0.031	23.3 ± 0.73
NM	0	12.82 ± 29.02	43.59 ± 88.15
	15	0.44 ± 0.041	20.24 ± 0.63

3.5.3 Discussion

Wild-type Mhp1 was successfully produced in four maltoside detergents with varying aliphatic chain length, for subsequent studies by SEC-SAXS. The SDS-PAGE and Western blotting analysis indicated that the final detergent which was used to solubilise the Mhp1 protein did not significantly affect the protein's integrity or stability in terms of their resistance to denaturation by SDS.

The binding affinity decreases with shorter aliphatic chain, indicating that the most crystallisable type of Mhp1 (solubilised in NM) is the one that binds L-BH the weakest. This could be due to the fact that shorter detergents may make the protein more rigid and restrict the conformational changes required to bind ligands efficiently. However, it is difficult to assess the degree of the detergent exchange that is achieved by changing the wash buffer of the Ni-NTA resin; SEC-SAXS will potentially be a useful tool to assess this.

An approximate thickness of DDM detergent corona, derived from previously reported values determined using SAXS, is 39.6-41.2 Å (Lipfert *et al*, 2007) while the thickness of a DM micelle is likely to be 36.0-37.6 Å. The average thickness of the *E. coli* inner cytoplasmic membrane has previously been reported to be 37.5 ± 0.5 Å (Shimamura *et al*, 2010; Mitra *et al*, 2004). This suggests that an NM detergent corona might be too small to support efficiently the conformational changes required for optimal ligand binding.

3.6 General discussion and conclusions

In this Chapter, the purification and characterisation of a number of variants of Mhp1 are described, including single-point mutants, seleno-L-methionine derivatised protein, as well as protein solubilised in detergents with varying aliphatic chain length. These variants were prepared and characterised for further X-ray crystallographic and scattering studies.

A number of mutants; including Gln42Asn, Gln121Asn, Ser312Ala, Asn318Gly, Thr313Ile as well as Lys110Leu were successfully produced. All the mutants were investigated using CD, confirming that they adopted an α-helical fold and are thermally stable in the temperature range at which the subsequent characterisation and structural studies are carried out. The T_m has been suggested to be an important factor for the success of protein crystallisation trials. A study has shown that protein with a T_m of 44.85 °C or higher crystallised in 49% of cases (Shimamura *et al*, 2010; Dupeux *et al*, 2011). This suggests that the wild-type Mhp1 protein produced here, of which the melting point was estimated to be 50.3 °C, is a good candidate for initial crystallisation trials. All single-point mutants appeared to have higher T_m values than wild-type, suggesting that all these mutants should be equally good crystallisation targets as wild-type Mhp1, provided that they adopt a crystallisable conformation. The tryptophan fluorescence quenching assays on the mutants give insights into the change in conformational dynamics that the mutants undergo, which affects the ligand binding. Obtaining an alternative conformational state along the path of the alternating access mechanism is of great interest, and will be further investigated in Chapter 5. In particular, an inward-open like conformational state is desirable, as there is only one crystal structure available for Mhp1 in this conformation (Shimamura *et al*, 2010). For this purpose, seleno-L-methionine derivatised Mhp1 was also produced, in order to attempt the reproduction of the inward-open form of Mhp1 to further establish the structural basis behind ligand transport through the protein and the coupling of the transport of sodium and ligand. In addition, the studies of the Mhp1 mutants highlighted the mutants Gln42Asn and Ser312Ala as potential targets to study the molecular basis behind how sodium and ligands are co-transported. In these mutants, the sodium-dependent ligand binding is diminished with respect to the wild-type protein, but can be restored to some degree upon the addition of excess sodium.

Having established the purification procedures to produce stable and functional protein, and identified the altered functionality of various mutants, the next step is the structural characterisation of these variants in order to establish the structural basis behind their altered functional behaviour. This primarily involves subjecting these mutants to crystallisation trials and structure solution where well-diffracting crystals will be obtained.

Despite the fact that there are crystallisation conditions already published for Mhp1, crystallisation still remains a trial-and-error process, with no guarantee that the protein produced for this project will produce well-diffracting crystals. Therefore, Mhp1 reconstituted into various detergents was also produced in order to investigate the potential of using SEC-SAXS as an alternative structural technique for the characterisation of membrane transport proteins that are reluctant to crystallise. Interestingly, the characterisation of these Mhp1-variants showed that the protein loses its efficiency to bind ligand with shorter detergent, and this will be further investigated in Chapter 6.

The mutants and variants studied in this Thesis make up only a minor fraction of the possible combinations of mutations and detergent combinations possible. Therefore, future directions involve extending the repertoire of mutants, as well as studying the effect of combinations of mutations in the Mhp1 protein, in order to further unravel the structural basis behind the transport mechanism of Mhp1.

Chapter 4: Optimisation of the crystallisation of Mhp1 mutants for X-ray crystallographic studies

4.1 Introduction

Obtaining a crystal structure may be viewed as the ultimate experiment to obtain structural information about a protein of interest. X-ray crystallography has the potential to provide structural information at near-atomic resolution, from which the structural basis of the interactions between single residues and ligands and associated conformational changes can be clearly elucidated.

However, obtaining a crystal structure requires the production of well-diffracting protein crystals, a process which is, even today, a trial-and-error process that usually requires extensive screening and optimisation of crystallisation conditions. In addition, it is especially challenging for membrane proteins due to their hydrophobic nature, as discussed in section 1.4.4.

As outlined in section 1.5, one of the central aims of this Thesis is to study Mhp1 mutants with altered ligand binding behaviour in order to gain more knowledge about the transport mechanism of the protein. In Chapter 3, the production and characterisation of a number of Mhp1 variants was described, for which previously published crystallisation conditions of Mhp1 provide an excellent starting point for crystallisation trials. However, the majority of the protein purifications and crystallisation trials that led to the previously published structures were carried out at the Membrane Protein Laboratory (MPL) at Diamond Lightsource, from membrane preparations that were produced at the University of Leeds and shipped to the MPL. Although the bacterial fermentations and membrane preparations for this project were carried out in the same environment and according to the same procedures, the protein purification and crystallisation trials were to be carried out at the University of Leeds under different environmental conditions. Even though the previously published conditions were successful in yielding crystals, further optimisation was still required to obtain well diffracting crystals. In this Chapter, the optimisation and experiments carried out in order to obtain well-diffracting crystals of the variants of Mhp1 for crystallographic studies are described.

4.2 Reproduction and optimisation of previously published crystallisation conditions

4.2.1 Introduction

Previously established crystallisation conditions have made it possible to determine structures of the Mhp1 protein in various conformational states, as well as in complex with a series of different hydantoin compounds (Weyand *et al*, 2008; Shimamura *et al*, 2010; Simmons *et al*, 2014). Two different crystallisation conditions have been identified, giving rise to two different crystal forms, in which the protein adopts different conformations (Table 4.1). The majority of the previously published crystal structures, adopting either the outward-open or occluded states, crystallised in an orthorhombic lattice in space group $P2_12_12_1$ (Weyand *et al*, 2008; Simmons *et al*, 2014). The only structure of Mhp1 in the inward-open state crystallised in a hexagonal crystal lattice in space group $P6_1$, under more basic conditions and in a different buffer (Shimamura *et al*, 2010).

Table 4.1 Crystallisation conditions used to obtain the previously published crystal structures of the various conformations of Mhp1.

The table is replicated from Table 3.5 in section 3.4.1.

	Orthorhombic	**Hexagonal**
Conformation of Mhp1 in the crystal structure	Outward-open/ occluded	Inward-open
PDB ID	2JLN, 4D1A, 4D1B, 4D1C, 4D1D	2X79
Detergent	NM	NM
PEG 300 (% v/v)	27 - 35	24 - 28
NaCl (mM)	100	100
Bicine pH 9.0 (mM)	-	100
$NaPO_4$, pH 7.0 (mM)	100	-
Reservoir volume (µl)	300	400
Temperature (°C)	20	4
Crystal morphology	Needle-shaped	Hexagonal
Space group	$P2_12_12_1$	$P6_1$

These two different crystallisation conditions will be referred to as the conditions for the orthorhombic or hexagonal crystal form in this Thesis. These crystallisation conditions also gave rise to different crystal morphologies, with the conditions for the orthorhombic form giving rise to long needle-shaped crystals (Shimamura *et al*, 2008), while the conditions for the hexagonal form producing hexagonal crystals (Peter Henderson, personal communication). In order to assess whether these conditions could be further exploited for the Mhp1 variants produced for this project, or whether re-screening for alternative crystallisation conditions would be required, reproduction of the crystallisation conditions outlined in Table 4.1 was initially attempted.

4.2.2 Reproduction of the crystallisation conditions for the outward-open and occluded forms with wild-type Mhp1

Wild-type Mhp1 was produced according to previously published procedures, which have been described in Chapter 2 and 3. This protein was used for initial trials for the reproduction of the crystallisation conditions for the orthorhombic crystal form. In the initial trials, the conditions from the original publication (Table 4.1) were replicated in 24-well plates in a hanging-drop configuration, and left to incubate in a crystal tray incubator at 18 °C. The concentration of polyethylene glycol (PEG) 300 was varied between 27-37% (v/v) with 2% increments and the sodium chloride concentration was varied between 60-120 mM with 20 mM increments. Since an essential step in the purification of Mhp1 for crystallisation trials was the detergent exchange from *n*-dodecyl-*β*-D-maltopyranoside (DDM) to *n*-nonyl-*β*-D-maltopyranoside (NM) (Shimamura *et al*, 2008), both DDM-solubilised and NM-solubilised wild-type Mhp1 were prepared and subjected to initial crystallisation trials. Both proteins were suspended over the same crystallisation wells in the initial screen and it was found that detergent exchange in NM is indeed required to promote crystal formation, also for the protein produced for this Thesis. In the droplets containing DDM-solubilised Mhp1, no crystals formed while showers of crystals readily appeared in the droplets containing NM-solubilised Mhp1 (Figure 4.1).

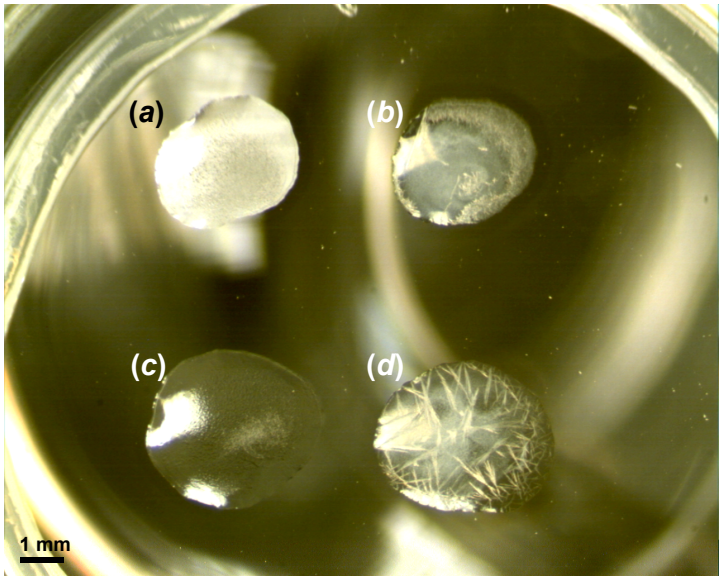


Figure 4.1 Crystallisation of wild-type Mhp1 solubilised in DDM and NM.

The photo depicts a crystallisation well with 0.25 µl droplets of DDM-solubilised (***a***, ***c***) and NM-solubilised (***b***, ***d***) Mhp1 mixed in a 1:1 (***a***, ***b***) and 1:2 (***c***, ***d***) ratio with a reservoir solution containing 29% (v/v) PEG 300, 80 mM NaCl and 100 mM $NaPO_4$ (pH 7.0) in a hanging-drop configuration over 300 µl of the reservoir solution. No crystals formed in the droplets with DDM-solubilised Mhp1, while NM-solubilised Mhp1 readily formed into needle-shaped crystals. The photo was taken three days after setting up the tray.

After the initial crystallisation screen, narrower grid screens were set up around the conditions that gave rise to the largest crystals, using left over protein from the purification, that had been flash frozen in liquid nitrogen and stored at -80 °C. The final optimal conditions varied between purification batches, but typically the largest crystals were obtained at 30-33% (v/v) PEG 300,100 mM sodium chloride and 100 mM $NaPO_4$ (pH 7.0).

4.2.3 Reproduction of the crystallisation conditions for the inward-open form with wild-type and seleno-L-methionine derivatised Mhp1

The production of seleno-L-methionine derivatised Mhp1 was described in section 3.4.2. The protein was subjected to the previously published crystallisation conditions for the hexagonal crystal form (Table 4.1). The screens were set-up in 24-well plates in a hanging-drop configuration and left to incubate at 4 °C. At the time of the set-up of these screens, a crystal tray incubator with accurate temperature control at 4 °C was not available, so the trays were incubated in a cold-room (4 °C). These trials resulted in the formation of long needle-shaped crystals, but no hexagonal crystals. The

needle-shaped crystals were considerably thinner than obtained with wild-type protein in trials to reproduce the conditions for the orthorhombic crystal form, and hence could not be harvested for diffraction studies. One reason for this outcome could that the temperature control was not sufficiently accurate in the common cold-room. With temperature fluctuations due to the daily use of the room, the equilibration of the crystallisation droplet against the reservoir solution could have been disturbed. These trials were also repeated with (un-derivatised) wild-type Mhp1, which gave rise to needle-shaped crystals similar to those obtained from the crystallisation conditions for the orthorhombic crystal form.

To repeat the experiment with more accurate temperature control, crystal trays were set up and left to incubate in a Rock Imager 1000 (Formulatrix) crystal tray hotel at 4 °C. The Rock Imager 1000 is designed to store only 96-well plates. Hence, the crystallisation conditions for the hexagonal crystal form were scaled down from 24- to 96-well plates by a scale factor of 0.15. This resulted in 0.22 µl droplets of protein solution being mixed in a 1:1 ratio (resulting in 0.44 µl droplets) with 60 µl of reservoir solution, where the PEG 300 concentration was varied between 24-29% and the concentration of sodium chloride was varied between 50-200 mM. The crystallisation droplets were set up in a sitting-drop configuration, and set up using a robot rather than by hand. In this scaled down screen, no crystals were obtained at all. This could be because the smaller volume and altered geometry of the crystal tray perturbed the crystal growth kinetics preventing nucleation.

4.2.4 Discussion

Crystals of NM-solubilised wild-type Mhp1 in the orthorhombic crystal form were successfully reproduced using the published crystallisation conditions, yielding typically 0.1-1 mm long needle-shaped crystals. The harvesting and diffraction quality of these crystals will be described in the following section (4.3).

In contrast, the crystallisation conditions for the hexagonal crystal form, which were used to obtain the structure of the inward-open form of seleno-L-methionine derivatised protein, did not yield hexagonal crystals. Only long needle-shaped crystals were obtained, which were considerably thinner than the crystals obtained for wild-type Mhp1. These crystals could not be successfully harvested for diffraction data collection to evaluate the crystal form.

There are several possible reasons for the lack of hexagonal crystals from either the seleno-L-methionine derivatised Mhp1 or with non-derivatised wild-type protein. For trials in 24-well plates, inaccurate temperature control during the equilibration of the crystallisation tray is likely to have disrupted the crystal growth process. When reproducing the conditions in a more controlled environment using a Rock Imager 1000 with accurate temperature control at 4 °C, the conditions had to be modified to adapt them to the 96 well plate format. Scaling down the conditions, and changing the configuration of the crystallisation trials into a sitting-drop configuration appeared to alter the growth kinetics so much that no crystals of either form were obtained with seleno-L-methionine derivatised Mhp1.

Only one previously published structure of Mhp1 crystallised in the inward-open state in a hexagonal lattice and space group $P6_1$ (PDB 2X79), has been reported. In contrast, the five other published structures of Mhp1 are all crystallised in the outward-open or occluded states with an orthorhombic lattice and space group $P2_12_12_1$ (PDB 2JLN, 4D1B, 4D1A, 4D1C and 4D1D). This distribution of space groups obtained from Mhp1 crystals suggests that the orthorhombic crystal form might be more favoured, potentially due to the fact that the outward-open or occluded states might be more stable, and hence crystallisable, forms of the protein.

The presence of unidentified electron density in the ligand binding site of the previous structure of Mhp1 in the inward-open form suggests that this conformation may be stabilised by ligand binding. The main difference between the structure determined of the protein in the inward-open conformation is that the crystals were produced from seleno-L-methionine-derivatised protein. As well as resolving the Mhp1 protein in an inward-open conformation, the electron density map also revealed a significant positive electron density feature in the ligand binding site. This positive electron density feature could not be assigned to any of the components in the crystallisation conditions, and has been proposed to be a carry-over product from the purification of the seleno-L-methionine derivatised protein from methionine auxotrophic *E. coli* cells (Peter Henderson, personal communication). It is possible that this carry-over is expression and/or purification batch specific and that in the derivatised protein prepared during for this Thesis, the ligand was absent.

4.3 Protein crystal harvesting and cryo-protection methods

4.3.1 Introduction

Prior to a crystallographic experiment at a synchrotron or in-house X-ray source, the protein crystals need to be harvested from the crystallisation drop, mounted in appropriate crystal sample holders and in cases where the data collection will be carried out at cryogenic temperatures (near 100K), cryo-protection is also required (Pflugrath, 2004).

Data collection at cryogenic temperatures greatly reduces radiation damage induced by the X-ray radiation in protein crystals, and makes it possible to collect complete data sets from single crystals (Garman & Schneider, 1997; Hope, 1988). The standard method for cryo-cooling crystals is flash-cooling by plunging mounted crystals into liquid nitrogen. However, this process often introduces artefacts into the crystal lattice due to mismatches in the thermal expansion coefficients between the solvent in the crystal interstitial volumes and the protein (Alcorn & Juers, 2010). Inhomogeneous and/or crystallisation of the solvent phase can cause high crystal mosaicity, reduced resolution and ice rings in the diffraction pattern. Cryo-protection of protein crystals attempts to minimise the detrimental effects of the cooling on the crystal lattice. Strategies for cryo-protection involve, for example, subjecting crystals to higher concentrations of cryo-protecting agents such as PEG or glycerol, the roles of which are to exclude and minimise excess solvent in the crystal lattice (Berejnov *et al*, 2006; Garman & Mitchell, 1996; McFerrin & Snell, 2002).

4.3.2 In-house harvesting and cryo-protection of Mhp1 crystals

The crystals obtained from the initial crystallisation trials were harvested in-house for screening at the in-house X-ray source or at a synchrotron. Various sample mounts were used, such as nylon loops, MiTeGen MicroMeshes™ (Figure 4.3*b*) and elongated MiTeGen MicroLoopsE™, which are designed to support long rod-shaped crystals. The sample mounts were selected according to the crystal size and position in the droplet prior to opening the crystallisation well. For the previously published crystal structures of Mhp1, no additional cryo-protection was performed and the crystals were flash-cooled by being directly plunged into liquid nitrogen (Weyand *et al*, 2008; Shimamura *et al*, 2010; Simmons *et al*, 2014). Therefore, in the initial harvesting attempts crystals were cooled by direct plunging into liquid nitrogen.

The harvesting procedure of the Mhp1 crystals obtained here presented a number of challenges. In most cases, the crystals would break upon attempting to extract them from the droplet with the sample mounts. In other cases, the crystals would start bending, or even dissolving, as soon as the crystallisation droplet was interfered with. In cases where there was precipitation in the droplets, crystals would often stick to the precipitate, making it difficult to harvest even a fragment of these crystals. From the crystal fragments that were harvested and cooled, the resulting diffraction ranged between 15-6 Å and was often irregular, with split spots and/or multiple overlapping lattices (Figure 4.2). No reproducible difference was observed in the quality of the diffraction with different choices of crystal mount.

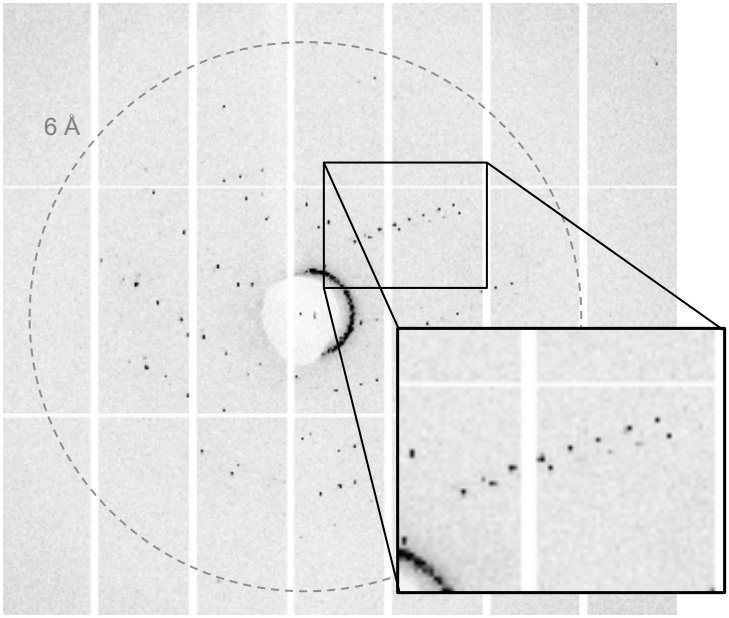


Figure 4.2 A low-resolution diffraction pattern from an Mhp1 crystal.

The image shows an example of a typical diffraction pattern obtained from the initial crystals of the Mhp1 protein. The dashed grey circle shows the diffraction limit at 6 Å, and the inset displays a zoomed-in section of the diffraction pattern, highlighting the presence of multiple lattices.

According to previous studies, the minimal concentration of PEG 200 and PEG 400 required for successful cryo-protection of protein crystals is 35.7-37.5% and >35% (v/v), respectively (McFerrin & Snell, 2002; Berejnov *et al*, 2006). This suggests that the minimal concentration of PEG 300 required for successful cryo-protection is likely to be in the range between 35-37.5% (v/v). After optimisation of the crystallisation conditions for the orthorhombic crystal form (Table 4.1), the reservoir solution typically contained 30-33% PEG 300. During the crystallisation experiment, the PEG 300 concentration

increases due to dehydration of the crystallisation droplet as it equilibrates against the reservoir solution. However, it is difficult to determine whether the dehydration of the droplet is sufficient for the concentration of PEG 300 to reach an appropriate concentration range for cryo-protection of the Mhp1 crystals.

To determine whether the low, disordered diffraction obtained from the initial crystals of wild-type Mhp1 was a result of cryo-cooling artifacts, further cryo-protection was attempted. This was done by increasing the concentration of PEG 300 in the crystallisation droplet by step-wise replacement of the liquid in the crystallisation droplet. This was done by adding 1 µl of mother liquor solution to the crystallisation droplet, containing a higher concentration of PEG 300 than in the reservoir solution over which the crystals that were being harvested had grown. After letting the added solution merge with the crystallisation droplet, 1 µl was removed from the merged droplet. This procedure was then repeated several times with progressively higher PEG 300 concentrations. The PEG 300 concentration was typically increased in 1-2% steps. However, this strategy was unsuccessful as the crystals dissolved upon the addition of any additional solutions to the crystallisation droplet. No crystals could be successfully harvested in these experiments. Therefore, direct plunging into liquid nitrogen without any further cryo-protection was subsequently used during this project.

4.3.3 An alternative harvesting strategy to minimise Mhp1 crystal handling

Overall, the crystals from wild-type Mhp1 obtained using the crystallisation conditions for the orthorhombic crystal form appeared to be highly fragile and sensitive to any outside interference. This made the harvesting of the Mhp1 crystals very challenging, and the resulting low and disordered resolution reflects the the fragility of these crystals. An alternative crystallisation strategy was devised to minimise the handling of the crystals, in order to investigate whether it is the actual harvesting procedure, which involved (often) repeated mechanical manipulation of the crystals while trying to mount them into the sample mounts. The crystallisation experiment was modified in order to grow crystals directly on mounts for diffraction data collection (Figure 4.3). This modification circumvents the harvesting step, completely avoiding making any direct contact with the crystals. If the harvesting step of the fragile crystals is the primary cause for the low and disordered diffraction, this should be an efficient strategy to improve the

diffraction quality. The crystallisation conditions for the orthorhombic crystal form were used, with wild-type Mhp1. As in the previous crystallisation trials performed here, the conditions were set up in 24-well plates, sealed with vacuum grease and MiTeGen MicroMeshes™ (Figure 4.3*b*) were used as the sample mounts as they provided a solid support for the crystallisation droplet. Crystals readily formed on the MicroMeshes™ (Figure 4.3*b*), with dimensions ranging between 100-500 µm and the meshes could be rapidly cooled by plunging into liquid nitrogen. The success of crystallisation in these trials is likely because the geometry of the set up depicted in Figure 4.3*a* closely resembles the hanging-drop configuration normally employed in the initial crystallisation trials.

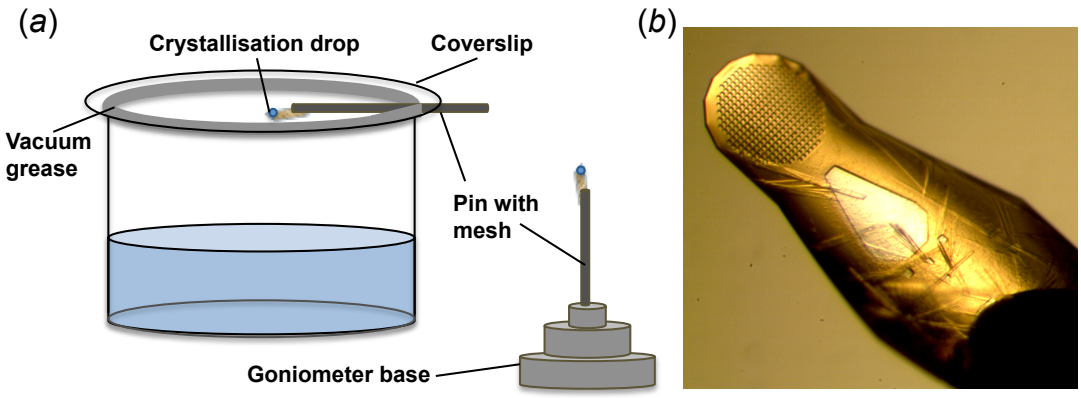


Figure 4.3 The set-up for growth of crystals directly on micro-meshes for data collection.

(*a*) The illustration depicts a crystallisation well in a 24-well plate. The micromeshes used in this experiment were positioned on the rim of the crystallisation well, which was pre-greased with vacuum grease, using a pair of tweezers. After placing the protein solution and the well solutions onto the mesh, forming the crystallisation droplet, the cover slip was placed on top of the pin. The mesh was glued to one side of the pin, with enough room for a crystallisation droplet despite the opposite side of the pin lying against the cover slide. The vacuum grease around the rim of the crystallisation well kept the well sealed from the outside. After leaving the tray to equilibrate, the pin was removed using a pair of tweezers, quickly placed into a re-usable goniometer base and cooled by rapid plunging into liquid nitrogen. (*b)* The size of the crystals that grew on the meshes ranged between 100-500 µm. In the photo, the mesh is 400 µm in diameter, and the mesh apertures are 10 µm.

However, no usable diffraction data were collected from these crystals. In many cases, crystals were found to be overlapping, which resulted in highly disordered and overlapping diffraction patterns with resolution limits not exceeding 6 Å.

4.3.4 *In situ* data collection

A more straight forward and established method for minimising crystal handling is *in situ* diffraction, where the entire crystallisation tray is mounted in the X-ray beam at a synchrotron beamline and the crystals are tested for diffraction directly in their respective wells (Axford *et al*, 2012). *In situ* diffraction studies were carried out with wild-type Mhp1 using a MiTeGen *In Situ*-1™ crystallisation plate, designed specifically for this type of experiment. The crystallisation trays were set up manually, with 0.5 µl protein solution being mixed in a 1:1 ratio with reservoir solution, and resulted in needle-shaped crystals 100-500 µm in length. Only very weak diffraction spots were observed and a unit cell could be obtained for only one of the crystals out of the 27 screened for diffraction. The diffraction was highly anisotropic, diffracting at the most to 6 Å, but the lattice could be indexed in space group P121 with lattice parameters of a = 92.9 Å, b = 109.4 Å, c = 114.5 Å, $\alpha = \beta = \gamma = 90°$.

Despite the fact that the *In Situ*-1™ crystallisation plate has been designed to give rise to minimal background scattering, the low level of background scattering could still be high enough to obscure the weak diffraction from the Mhp1 crystals. Crystals of detergent-solubilised membrane proteins typically contain a significantly higher solvent content compared to soluble proteins (Carpenter *et al*, 2008), resulting in weaker diffraction in comparison to soluble protein crystals with similar unit cell sizes. In addition, the presence of the crystallisation droplet surrounding the crystals is likely to be a major source of background scattering that could be obscuring the low diffraction intensities from the crystals grown in the *in situ* tray. Another possibility for the lack of diffraction is simply that disordered crystals formed in the wells, perhaps as a result of the alternative vapour-diffusion configuration (sitting-drop, smaller reservoir volume) compared to the initial classic hanging-drop vapour diffusion experiments made with wild-type Mhp1. Finally, the crystallisation tray was also transported in a polystyrene box to Diamond Lightsource, and the crystals may have suffered internal damage during travel.

4.3.5 Discussion

In this section, the harvesting and cryo-cooling procedures tested for crystals of wild-type Mhp1 have been described. The crystals were harvested using conventional methods, with cryo-cooling by directly plunging the crystals into

liquid nitrogen without any additional cryoprotection. Additional cryo-protection was attempted but was not successful as the crystals were found to dissolve upon the addition of any solution to the crystal containing droplets. Indeed, the crystals occasionally dissolved and readily broke during any harvesting procedure. The pronounced fragility of these crystals is a likely reason for the poor quality of the diffraction that could be obtained from these crystals.

Due to the apparent fragility of the Mhp1 crystals obtained here and the difficulties arising during the harvesting procedure, alternative strategies for crystal growth and harvesting were investigated to minimise crystal handling.

The crystallisation set-up described in 4.3.3 is an easy way to minimise crystal handling, using conventional and standard consumables and materials usually used for crystallisation experiments. If crystallisation conditions that do not require additional cryo-protection are available, this alternative crystallisation set-up should be an efficient strategy to make the harvesting step less time-consuming. However, this strategy did not improve the diffraction quality of Mhp1 crystals and no usable diffraction data were collected from the crystals grown directly on MicroMeshes™.

The more straight forward strategy to minimise crystal handling is *in situ* diffraction studies. Crystals were successfully grown in *in situ* plates, but gave rise to only extremely weak diffraction sports and no usable diffraction data could be collected.

4.4 Using diffusion to aid and optimise the crystallisation of mutant Mhp1

4.4.1 Introduction

Convection-free crystallisation, which can be achieved either under microgravity conditions or within thin capillaries, has been shown to promote the growth of highly well-ordered crystals (Vergara *et al*, 2005). In order to utilise this phenomenon, microfluidic devices have been developed that have been shown to improve the crystallisation process (Dhouib *et al*, 2009; Emamzadah *et al*, 2009). In section 1.4.4, the basic principle behind protein crystallisation is described. Various methods, such as vapour-diffusion, dialysis or the batch methods are used to bring the protein solution to a supersaturated state, from which crystals may form and grow if the supersaturation is achieved in the nucleation zone.

In a counter-diffusion experiment, the protein and precipitant solution are in a configuration that allows the two solutions to diffuse into one another, for example, from opposite ends of a thin capillary. This creates a supersaturation gradient across the capillary, allowing for a wider sampling of the phase diagram, increasing the probability of achieving successful crystal formation and growth (Figure 4.4).

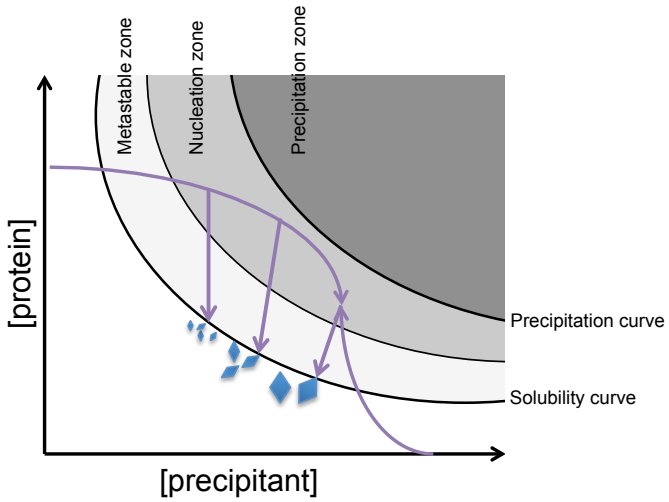


Figure 4.4 The phase diagram in a counter-diffusion crystallisation experiment.

The phase diagram corresponds to the diagram depicted in Figure 1.10. The purple arrows depict how the concentrations of the protein and precipitant solutions vary during a counter-diffusion experiment. The creation of a supersaturation gradient allows for the sampling of the nucleation zone, allowing crystal nucleation and formation from various positions in the phase diagram.

The crystals of wild-type Mhp1 that were produced using standard crystallisation and harvesting approaches gave rise to low resolution and disordered diffraction patterns, likely as a result of the fragility of these crystals. In order to further explore whether the crystal growth could be improved, crystallisation trials were performed using various approaches available to harness diffusion for crystal growth optimisation.

4.4.2 Commercially available trays for counter-diffusion crystallisation trials

A number of commercially available crystallisation trays in which counter-diffusion is utilised to optimise crystallisation were tested with wild-type Mhp1 in order to investigate whether this is a viable strategy for optimisation of crystal growth.

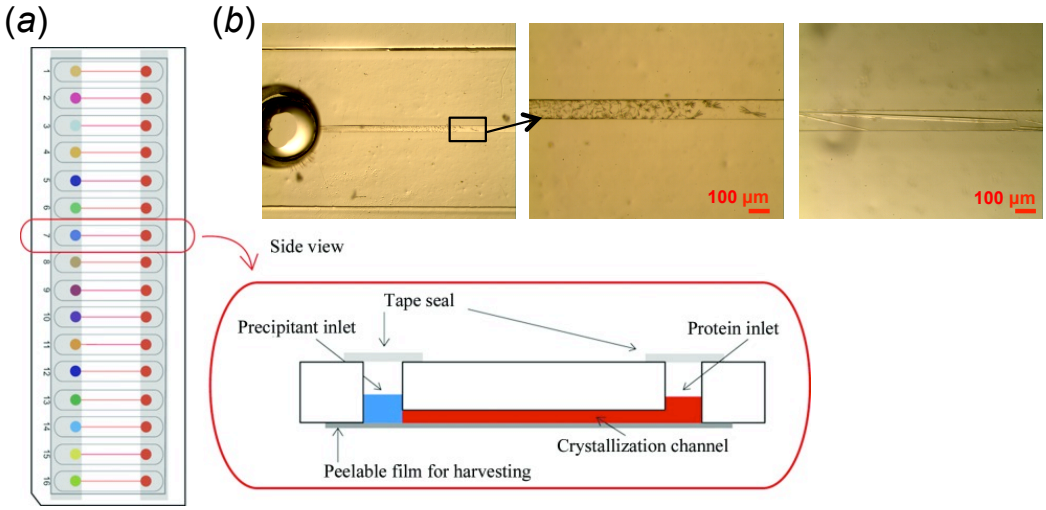


Figure 4.5 The original CrystalFormer™ crystallisation tray for crystal growth by counter-diffusion.

a) Schematic of the CrystalFormer™ tray produced by Microlytic (2.5 x 7.5 cm), with the red circles representing the inlets for the protein solution, while the multi-coloured circles represent the inlet wells for precipitants or screening solutions. The side view in the red box shows the configuration of the crystallisation channel. Figure from Stojanoff *et al.*, 2011. *b*) The photographs show some of the wild-type Mhp1 crystals formed in the channels of the CrystalFormer™, The left panel depicts part of the well where protein was added, and the middle panel is a close up of the same channel, showing the size gradient of crystals formed with a precipitant solution of 32% PEG 300 and 100 mM $NaPO_4$ pH 7.0. The right panel shows a bigger crystal grown along the direction of the micro-channel, in the presence of 29% PEG 300, 100 mM NaCl and 100 mM $NaPO_4$ pH 7.0.

One of the trays tested with wild-type Mhp1 was the original CrystalFormer™ from Microlytic (Stojanoff *et al*, 2011). This tray has micro-channels connecting the inlet wells for protein and precipitant solution with diameters of 0.1 x 0.15 x 10 mm and a total channel volume of 150 nm^3 on a small footprint tray (Figure 4.5*a*). 0.3 µl of protein or reservoir solution were aliquoted into each well, sealed with clear tape and left to equilibrate at 20 °C.

Wild-type Mhp1 crystals grew readily in the micro-channels of the CrystalFormer™ tray. Pronounced crystal size gradients were observed and the longest crystals grew to approximately 1 mm (Figure 4.5*b*). This demonstrates that Mhp1 are able to successfully grow within a concentration gradient.

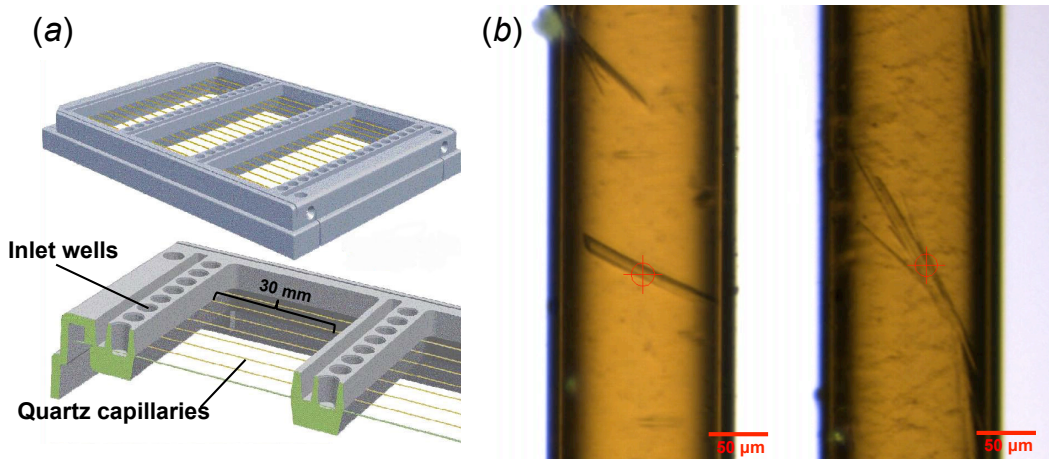


Figure 4.6 The CrystalHarp™ tray for crystal growth by counter-diffusion.

(*a*) The figure is adapted from the CrystalHarp™ manual supplied by Molecular Dimensions. Each crystallisation experiment was contained in a 30 mm long quartz capillary with an internal diameter of 0.1 mm. The protein and precipitant solutions were loaded into the inlet wells, through which the capillaries were also running. In order to allow the solutions to start diffusing through the capillaries, the capillary sections running through the inlet wells were broken using a tool supplied with the CrystalHarp™ kit. (*b*) The dimensions of the plastic frame complies with the SBS format for screening plates, which made it possible to perform *in situ* diffraction studies. The photos are snapshots from the online camera at the I03 beamline at Diamond LightSource, with the red encircled cross representing the position of the beam.

A second crystallisation tray that utilises counter-diffusion, the CrystalHarp™ from Molecular Dimensions was also tested. In the CrystalHarp™, 30 mm long quartz capillaries with an internal diameter of 0.1 mm are mounted in a tray that complies with the dimensions formulated by the Society for Biomolecular Screening (SBS), giving rise to a channel volume of 236 nm^3 per crystallisation trial. The capillaries are connected to inlets for protein and screening solutions (Figure 4.6*a*). The crystallisation conditions for the orthorhombic crystal form (Table 4.1) were pre-mixed in eppendorf tubes and added to the inlet wells. Since the dimensions of the plastic frame of the CrystalHarp™ is an SBS format tray, *in situ* diffraction experiments could be performed by mounting the entire tray in a dedicated tray mount at a synchrotron beamline.

Crystals also grew readily in the quartz capillaries, both along and perpendicular to the direction of the capillaries. The length of the crystals typically varied between 100-300 µm. *In situ* data collection was performed at the I03 beamline at Diamond Lightsource but no diffraction data were obtained from the crystals in the quartz capillaries (Figure 4.6*b*). As for

crystals screened for diffractionin the MiTeGen *In situ*-1 tray (section 4.3.4), the crystals grown in the CrystalHarp™ were immersed in crystallisation mother liquor, which could be giving rise to background scattering that masks the diffraction from the crystals.

4.4.3 Drop morphology modification to optimise crystal growth

It was discovered that changing the morphology of the crystallisation droplet gave rise to a similar crystal growth outcome as when using commercially available microfluidic devices for crystallisation trials by counter-diffusion.

By pipetting the droplet of precipitant solution next to the protein droplet while applying a streaking motion with the pipette tip while dispensing the precipitant solution through the protein droplet, creating an elongated droplet, promoted the formation of a size gradient of crystals in the droplet (Figure 4.7). The streaking motion may be causing the formation of a temporary gradient of the protein vs. the precipitant solution, mimicking the gradients formed in the CrystalHarp™ and CrystalFormer™ and allowing the protein and reservoir solution to slowly diffuse into each other.

The resulting gradient typically resulted in the formation of showers of smaller crystals at one end of the droplet that gradually got bigger with a few, typically mm-long crystals, protruding from the shower of the smaller crystals (Figure 4.8). These crystals were larger and thicker than those previously obtained in the initial crystallisation trials, and were considerably better separated from the other crystals in the droplet. These crystals were also much easier to harvest due to their size and separation from the other crystals in the droplet.

4.4.4 Crystallisation of Mhp1 mutants

The drop morphology modification trick was applied during crystallisation trials of the Mhp1 mutants that were described in 3.3. The Mhp1 mutant Gln42Asn, in complex with L-BH and L-IMH, were the first mutant ligand combinations to be successfully crystallised in this manner. The obtained crystals were harvested as described in 4.3.2 and cooled by direct plunging into liquid nitrogen without any further cryo-protection. These crystals reproducibly gave rise to clean diffraction beyond 4 Å, which is a significant improvement on the diffraction quality being obtained previously from wild-type Mhp1 crystals. This crystallisation strategy, involving the drop

morphology modification, was therefore applied to all the other mutants studied as part of this Thesis.

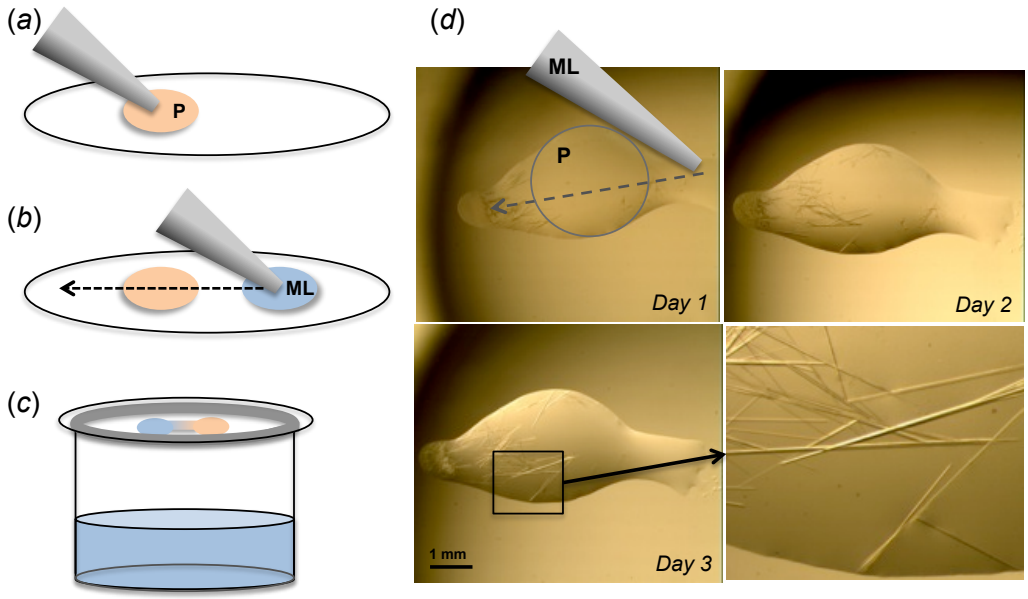


Figure 4.7 The drop streaking procedure used to introduce a temporary gradient of protein vs. precipitant concentration to optimise crystal growth.

The temporary gradient of protein and the precipitating agents in the mother liquor was introduced by applying a streaking motion when setting up the crystallisation droplet in the following steps: (*a*) The protein droplet, shown in orange and marked with **P**, was first applied to the cover slide. The grey cone represents a pipette tip. (*b*) The mother liquor droplet, shown in blue and marked with **ML**, was applied next to the protein droplet without making contact with it. As the **ML** was extruded from the pipette tip, the tip was streaked through the protein droplet, indicated by the dashed arrow. (*c*) The cover slide was then inverted in a hanging-drop configuration over the crystallisation well. (*d*) The photos show the formation of crystals over three days, with the initial formation of showers of smaller crystals from which a few large crystals grew.

Figure 4.8 shows crystals of the Mhp1 mutant Gln42Asn in complex with L-BH under low-sodium/high-potassium conditions, clearly depicting the crystal size gradient formed in the crystallisation droplet.

Crystals were obtained for all mutants studied as part of this Thesis, either in their apo form or in complex with various ligands. All types of crystals were screened for diffraction at synchrotron beamlines, and data sets were collected from a series of mutant-ligand complexes (summarised in Table 4.2) in space group $P2_12_12_1$, typically with unit cell parameters of $a \approx 90.0$, $b \approx 107.0$, $c \approx 107.0$ Å, $\alpha = \beta = \gamma = 90°$. The resulting structures that were

solved from the data sets collected from these crystals will be described in Chapter 5.

Table 4.2 The outcome of the crystallisation trials of Mhp1 mutants.

Mutant	Site of the mutation	Crystals obtained (apo or ligand-complex)	Data sets collected
Gln42Asn	Ligand	Apo, in complex with L-BH, L-IMH or BVH	In complex with L-BH (under high-sodium and under low-sodium conditions) and L-IMH.
Gln121Asn	Ligand	Apo, in complex with L-BH or L-IMH	No data collected due to low or poor diffraction.
Asn318Gly	Sodium	Apo	No data collected due to low or poor diffraction.
Ser312Ala	Sodium	In complex with BVH or L-NMH	In complex with L-NMH.
Lys110Leu	Bundle/hash	In complex with L-BH and L-IMH	In complex with L-BH and L-IMH.

4.4.5 Discussion

Crystallisation by counter-diffusion was initially shown to be successful strategy for Mhp1 using commercially available microfluidic trays, including the CrystalHarp™ and the CrystalFormer™. Crystals grown in the CrystalHarp™ tray were screened for diffraction *in* situ, but did not give rise to any diffraction. This could have been due to the fact that the crystals were immersed in solvent, reducing the diffraction signal in a similar way as during conventional *in situ* studies (section 4.3.4).

Based on the success of these initial experiments, a new crystallisation protocol for Mhp1 was developed that used a drop streaking approach to generate a gradient of protein vs. precipitant concentration, enabling the two solutions to diffuse slowly into each other. By applying a streaking motion upon setting up the crystallisation droplet, resulting in an elongated droplet, a crystal size gradient was obtained, in which a few larger and thicker, well-separated crystals typically appeared. These crystals were considerably easier to harvest and when tested for diffraction at a synchrotron beamline,

routinely gave rise to diffraction beyond 4 Å. These crystals crystallised in space group $P2_12_12_1$, indicating that the favoured orthorhombic lattice of Mhp1 was reproduced.

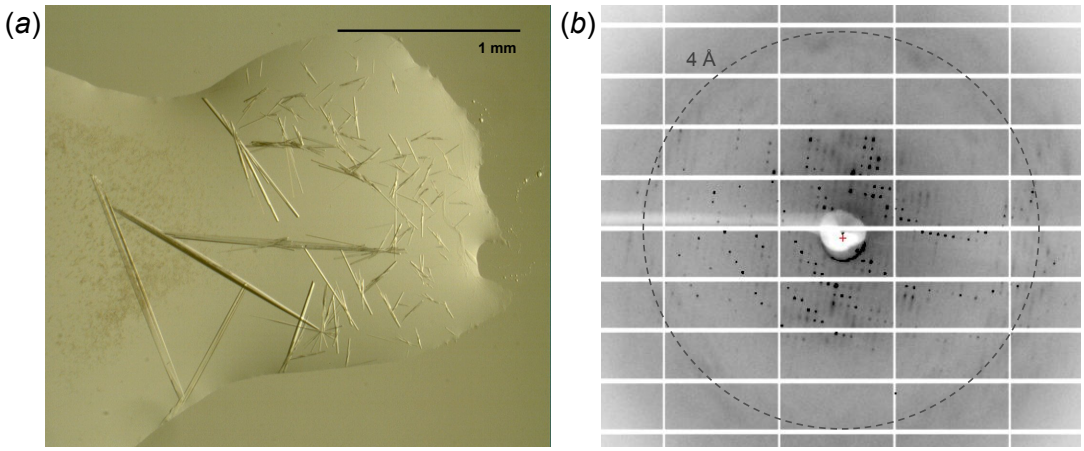


Figure 4.8 The size distribution of crystals of Gln42Asn in complex with L-BH, achieved by drop morphology modification.

Adding a streaking motion upon setting up the crystallisation droplet greatly enhanced the crystallisation outcome of Mhp1 mutants. The photo depicts crystals of the Mhp1 mutant Gln42Asn in complex with L-BH (*a*), in the presence of 29% PEG 300, 100mM KCl and 100mM KPO_4. The modification gave rise to larger crystals with improved diffraction quality, routinely diffracting < 4 Å (*b*).

Applying this crystallisation protocol for the Mhp1 mutants studied as part of this Thesis, crystals of a number of mutant-ligand complexes were obtained. From these crystals, diffraction data for the mutants Gln42Asn, Ser312Ala and Lys110Leu, in complex with various ligands and under varying crystallisation conditions were collected. The structures of these mutant-ligand complexes will be further discussed in Chapter 5.

Crystals were obtained for all mutants, either in their apo form or in complex with a ligand, apart from the Mhp1 mutants Ser312Ala and Lys110Leu, which only crystallised in the presence of ligands. Diffraction data sets were, however, only obtained from mutant-ligand crystals, and hence not from any of the apo crystals. This suggests that Mhp1 mutants form more ordered crystals in the presence of ligand. During crystallisation trials, it did appear that the Mhp1 mutants crystallised more readily in the presence of ligands. This could be due to the ligands promoting crystallisation by acting as nucleation agents in the crystallisation mixture. Alternatively, the ligands may stabilise and rigidify the conformation of the mutants and hence give rise to a better ordered and better diffracting lattice. This idea is supported by the fact

that no usable diffraction data were obtained from the crystals of the mutants in their apo form.

4.5 Post-crystallisation treatments to improve diffraction

4.5.1 Introduction

Dehydration has been shown to be a successful method to improve diffraction from protein crystals (Esnouf *et al*, 1998; Heras *et al*, 2003; Abergel, 2004). Dehydration of protein crystals acts as an alternative strategy for cryo-protection, as crystals at an optimal relative humidity (RH) and with no surrounding mother liquor can often be directly cryo-cooled (Pellegrini *et al*, 2011). In some of these cases, dehydration of protein crystals has occurred coincidentally, for example with capillaries cracking or crystallisation wells not being sealed properly (Abergel, 2004). Other more systematic approaches to dehydrate protein crystals involve, for example, transfer to crystallisation wells with an increasing salt concentration, or soaking crystals in dehydrating agents (Heras & Martin, 2005). A disadvantage of these methods is that the outcome of the dehydration is only apparent after the diffraction experiment, making these methods time- and material-consuming. In order to overcome these drawbacks, humidity-controlling systems have been developed for adjusting the RH prior to crystal cryo-cooling (Sjögren *et al*, 2002; Kiefersauer *et al*, 2000), or for dehydration of crystals mounted in the X-ray beam at crystallography beamlines (Sanchez-Weatherby *et al*, 2009). Performing the dehydration on-line gives the user full control to monitor and adapt the dehydration protocol to the changes in diffraction.

Controlled dehydration using the Humidity Control device HC1 (Sanchez-Weatherby *et al*, 2009) has already been used for some of the Mhp1 ligand-complex structures previously published, namely the structures of Mhp1 in complex with L-IMH and the L-NMH structures (Simmons *et al*, 2014). Due to the unsuccessful attempts to cryo-protect Mhp1 crystals described here, controlled dehydration was chosen as an alternative cryo-protection method to be tested.

4.5.2 On-line humidified gas stream for the dehydration of Mhp1 crystals at room-temperature

Diffraction data were collected from a crystal of wild-type Mhp1 to a resolution of 3.8 Å (see Table 4.3 for statistics), which had been subjected to

controlled dehydration using the HC1 device and then flash cooled in a cryogenic gas stream at the diffraction end-station at the I03 beamline at Diamond Lightsource. At this beamline both the HC1 and cryogenic stream device are mounted on a bracket which allows them to be automatically swapped to flash cool samples when an optimal diffraction quality has been obtained. In this experiment, the RH was decreased from 98% to 80%, with step of 2% and a dehydration rate of 2% per minute. Between each step, the crystal was left to equilibrate for approximately 2 minutes, and the diffraction quality and change in the lattice parameters was monitored by collecting a diffraction image (1 s, 1° oscillation, 1% transmission to give a flux of 2.61 x 10^{11} photons s^{-1}) after each step. When a significant decrease of the diffraction limit was observed, the crystal was rapidly cryo-cooled by automatically switching the HC1 device for the cryogenic gas stream at 100 K and a full data set was collected.

The starting RH value of 98% was predetermined using the HC1 software at the beamline, by monitoring the size of an approximately 1 µl droplet of the mother liquor (from the well containing the crystal that was to be studied; in this case containing 27% PEG 300, 100 mM NaCl and 100mM $NaPO_4$, pH 7.0) in a Hampton nylon loop (~ 0.7 mm diameter). The nylon loop was oriented with its plane perpendicular to the on-axis camera and the thickness of the droplet was monitored using the image recognition package normally used for crystal centering and plotted against the RH value. The RH value was adjusted until no change in droplet size was observed and this RH value was then used as the starting value for the dehydration experiment.

A change in the lattice was clearly visualised during the course of the dehydration (Figure 4.9), as well as a decreasing diffraction limit. This loss of high-resolution diffraction was most likely due to local radiation damage, as the diffraction limit was restored if a diffraction image was collected at a different position on the crystal. Irreversible effects on the diffraction quality were also observed. During the dehydration process, a powder diffraction ring appeared at ~ 37 Å, which remained even when diffraction images were collection at other positions on the crystals after the appearance of this ring (Figure 4.9). The ring may correspond to a peak in the scattering pattern of NM micelles, as scattering patterns of DDM and DM micelles include peaks at approximately 39 and 34 Å, respectively (Lipfert *et al*, 2007).

This could involve a phase change of the detergent phase in the crystal lattice, or a local rearrangement resulting in a more structured detergent

phase. The contraction of the crystal lattice is likely to also compact the detergent corona surrounding the protein molecules.

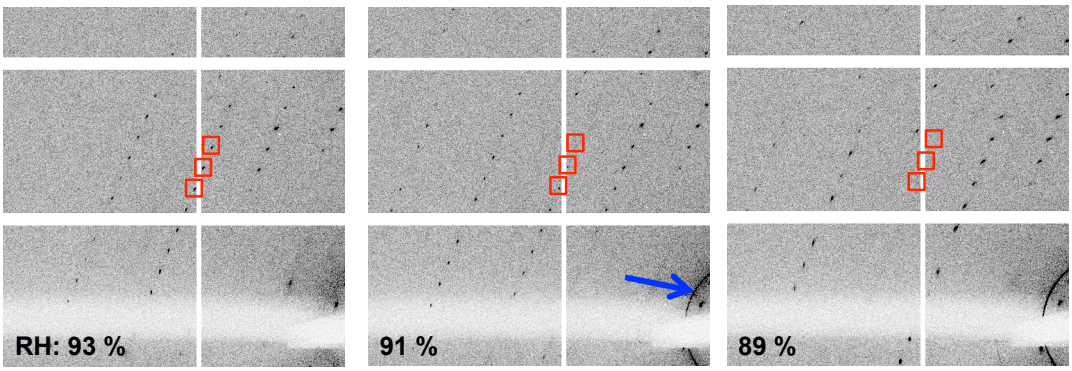


Figure 4.9 Dehydration of a wild-type Mhp1 crystal.

Diffraction images of 1° oscillation collected while the crystal was subjected to dehydration. The red squares encircle three arbitrarily chosen diffraction spots, illustrating a change in the lattice as the diffraction spots move out of the plane of the detector. The blue arrow in the diffraction image at 91% RH points to a ring that appeared during dehydration, which may be a result of reorganisation and/or compaction of the detergent phase in the crystal lattice.

The observed change in the diffraction pattern upon the dehydration of the wild-type Mhp1 crystal was due to a contraction of the crystal lattice. The structure of dehydrated wild-type Mhp1 was indexed in space group $P2_12_12_1$, exhibiting contracted unit cell edges (a = 86.3, b = 101.7, c = 107.2 Å, $\alpha = \beta = \gamma = 90°$) compared to the initial lattice at 98% RH (a = 92.0, b = 109.0, c = 114.0 Å, $\alpha = \beta = \gamma = 90°$). The data could be used to obtain a molecular replacement solution and initial refinement was performed to values of R_{work}/ R_{free} of 27.92/31.18% (Table 4.3).

The lattice was also contracted with respect to the lattice of the crystals from Mhp1 variants produced as part of this Thesis that were not subjected to controlled dehydration. These were in space group $P2_12_12_1$ with typical unit cell parameters of $a \approx 90.0$, $b \approx 107.0$, $c \approx 107.0$ Å, $\alpha = \beta = \gamma = 90°$ (section 4.4.4).

Table 4.3 Data collection statistics for the structure of wild-type Mhp1 subjected to controlled dehydration using the HC1 device.

	Wild-type Mhp1
Wavelength (Å)	0.9763
Temperature (K)	100
Space group	$P2_12_12_1$
Molecules per assym. unit	1
Unit cell parameters (Å, °)	a = 86.3, b = 101.7, c = 107.2, $\alpha = \beta = \gamma = 90$
Resolution (Å)	67.24 - 3.8 (4.25-3.80)
No. of measured reflections	69252 (9719)
No. of unique reflections	19576 (2683)
Multiplicity	7.1 (7.3)
Completeness (%)	99.4 (99.0)
$<I/\sigma(I)>$	8.7 (1.6)
$CC_{1/2}$	0.99 (0.89)
Wilson B factor (Å^2)	231.8
Anisotropy ΔB (Å^2)	119.5
R_{merge} (%)	0.125 (1.475)
R_{meas} (%)	0.146 (1.713)
$R_{p.i.m.}$ (%)	0.074 (0.861)
Matthews coefficient (Å^3 Da^{-1})	4.31
Solvent content (%)	71.48
R_{work} (%)	27.92 (35.57)
R_{free} (%)	31.18 (38.80)
R.m.s.d.	
Bonds (Å)	0.0100
Angles (°)	1.3918
Ramachandran outliers (%)	1.53
Average B factor (Å^2)	187.4

Values in parentheses are for the highest resolution shell.

4.5.3 Discussion

Controlled dehydration using the HC1 device at a synchrotron beamline was successfully performed on wild-type Mhp1 crystals, demonstrating that it is a suitable method to obtain good quality data from Mhp1 crystals. Compared to the data statistics of a crystal from the Mhp1 mutant Gln42Asn in complex with L-BH, which was cooled by plunging into liquid nitrogen (discussed

further in section 5.2.2), the data obtained from a dehydrated crystal with a contracted lattice shows improvements in the anisotropy of the data, as well as the temperature *B* factors of the data and the model (Table 4.4).

Table 4.4 Comparison of data statistics obtained from a crystal subjected to controlled dehydration prior to cryo-cooling and a crystal cooled by plunging into liquid nitrogen.

	Wild-type Mhp1 (cooled after controlled dehydration)	Gln42Asn-(L-BH) (cooled directly by plunging into liquid nitrogen)
Unit cell parameters (Å)	*a* = 86.3, *b* = 101.7, *c* = 107.2	*a* = 91.9, *b* = 107.1, *c* = 107.2
High resolution limit (Å)	3.80	3.74
Wilson *B* factor (Å^2)	231.8	309.5
Anisotropy Δ*B* (Å^2)	119.5	146.5
R_{merge} (%)	0.125 (1.475)	0.040 (1.192)
R_{meas} (%)	0.146 (1.713)	0.056 (1.641)
$R_{p.i.m.}$ (%)	0.074 (0.861)	0.038 (1.124)
Solvent content (%)	71.48	74.56
Average *B* factor (Å^2)	187.4	253.0

The dehydration of the crystal from a RH of 98% to 80% induced a reduction of the unit cell volume by approximately 18%, resulting in closer packing of the Mhp1 molecules in the crystal lattice. This could have reduced the mobility of the protein molecules in the detergent rich lattice, giving rise to a better packed lattice. With time constraints due to the availability of beamtime and the HC1 device, a reproducible protocol for the dehydration of Mhp1 crystals was not devised. Further exploration of parameters of the dehydration procedure, for example the steps of decrease in RH, the rate of dehydration in each step and the equilibration time between each step would allow optimisation of the protocol to the crystals to obtain the best possible data (Awad *et al*, 2013). It is likely that the crystals could be pushed to a lower humidity level, resulting in an even greater reduction in unit cell volume and solvent content, and thus in an even better ordered packing of the molecules in the lattice. This could result in a further improvement of the anisotropy and *B* factors of the model, as seen in Table 4.4. An optimised protocol would minimise radiation damage by simply not collecting as many diffraction images in order to monitor the process, hence preserving the full diffraction power of the crystals.

The packing of Mhp1 molecules in the lattice obtained from the crystals prepared here can be described as ‘zig-zag chains’ of Mhp1 molecules (running in the direction of the *c*-axis), packing against each other in a staggered fashion in the directions which are equivalent to the putative membrane plane in which the protein is embedded (in the directions of the *a*- and *b*-axes), illustrated in

Figure 4.10*a* and *b*. The dehydration of a wild-type Mhp1 crystal produced here resulted in a contraction of the unit cell parameters by *a* = 92.0 → 86.3, *b* = 109.0 → 101.7, *c* = 114.0 → 107.2 Å, resulting in a contraction in the direction of the ‘zig-zag chains’, as well as in the directions of the putative membrane plane, packing the chains closer against each other. Comparing the lattice resulting from a crystal that had been subjected to the controlled dehydration with one obtained previously (with typical lattice parameters *a* ≈ 90.0, *b* ≈ 107.0, *c* ≈ 107.0 Å, see section 4.4.4), the contraction occurred mainly in the directions of the *a*- and *b*-axes. The contraction of the *a*- and *b*-axes results in closer packing of the protein chains against each other, probably packing the detergent phase. This is consistent with the appearance of a detergent powder diffraction ring in the diffraction pattern upon dehydration, possibly occurring due to reorganisation of the detergent molecules in the lattice upon the lattice contraction.

All other previously published crystal structures of Mhp1 (Weyand *et al*, 2008; Shimamura *et al*, 2010; Simmons *et al*, 2014) have a similar packing apart from the structure of the Mhp1 protein in an outward-open conformation (PDB 2JLN, Weyand *et al*, 2008), illustrated in Figure 4.10*c*. This structure is resolved to the highest resolution obtained so far for Mhp1 crystals, 2.85 Å. The crystal from which these data were collected came from a relatively old tray, and the high resolution could have potentially been caused by slow dehydration due to the long-term storage (Peter Henderson and Alexander Cameron, personal communication). The 2JLN lattice is also in space group $P2_12_12_1$, with unit cell parameters *a* = 79.7, *b* = 109.1, *c* = 113.8 Å. The 2JLN lattice is different from the lattice obtained for the Mhp1 mutants here by that the protein chains run along the *b*-axis, and that there are additional crystal contacts made between the chains in the *a*- and *c*-plane. This further stabilisation is a likely reason for the high quality and resolution of the 2JLN data. Potentially, the Mhp1 protein crystals obtained as part of this Thesis could also be further contracted by controlled dehydration to obtain a change in the lattice similar to that seen in the 2JLN lattice, with additional crystal contacts stabilising the lattice. Perhaps the

observed putative detergent organisation during dehydration, as well as the decrease in resolution is due to the starting stages of a major lattice rearrangement where the protein molecules re-pack into a tighter and better stabilised lattice.

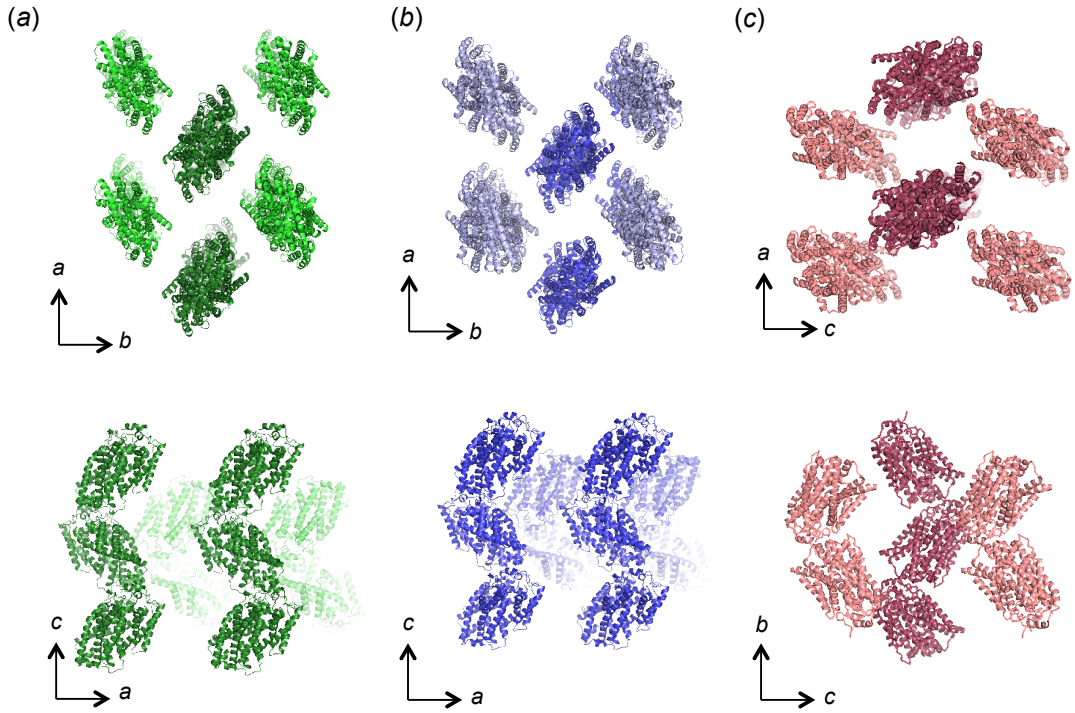


Figure 4.10 The crystal packing of Mhp1 molecules in the $P2_12_12_1$ lattice.

The packing of Mhp1 molecules are displayed in an un-contracted lattice (*a*), a lattice contracted by controlled dehydration (*b*) and the packing in the structure 2JLN of Mhp1 in an outward-open form (*c*). The protein molecules are displayed in a cartoon representation and the arrows represent the directions of the unit cell axes *a, b* and *c*.

4.6 Contact-less harvesting by photo ablation

4.6.1 Introduction

Automation of the various steps of X-ray crystallographic experiments has greatly spurred the acquisition of crystal structures in the past decade (Wasserman *et al*, 2012). Preparing crystallisation solutions, dispensing the solutions into crystallisation trays and setting up crystallisation screens, as well as visualisation and inspection can be automated and performed in a high-throughput manner using crystallisation robots and dedicated crystal storage hotels. At crystallography beamlines, extensive efforts have also led to the development of automatic mounting and centering of crystals into the

X-ray beam, as well as automatic data collection protocols (Svensson *et al*, 2015). In addition, pipelines for automatic data processing are also being developed and made available to users of crystallography beamlines (Brockhauser *et al*, 2012; Monaco *et al*, 2013). These efforts have greatly increased the throughput of X-ray crystallographic experiments.

One step in the process from crystal to structure that still requires ample human intervention is the crystal harvesting and cryo-cooling step. The harvesting and cryo-cooling steps often require extensive optimisation, in the process of which precious crystals are lost and/or damaged, which ultimately results in poor diffraction data.

At the European Molecular Biology Laboratory (EMBL) in Grenoble, an automatic system for harvesting and cryo-cooling has been developed (Cipriani *et al*, 2012) in order to bridge the automation gap between the crystallisation screening and data collection processes. The CrystalDirect system involves growing crystals in custom-made crystallisation trays, where crystallisation droplets are set up in a sitting-drop configuration, with the droplets resting on a thin polyimide film (12.5 μm thick), which is compatible with X-ray data collection and laser-induced photo ablation (Figure 4.11*a*) (Cipriani *et al*, 2012). The custom-made tray containing crystals is then placed in a robotic set-up which uses a laser cutting tool to excise a segment of the film of a desired shape, containing the crystal(s) whose position has been marked in the CrystalDirect software interface. Initially, a small hole is cut by the laser outside of the area where the crystal is positioned, in order to make aspiration of the solvent of the crystallisation droplet possible. Through a second hole cut next to the crystal, a hollow pin is positioned through which liquid from the droplet is aspirated to minimise the solvent around the crystal.

The pin is then covered with glue while the laser makes the first incision in the film underneath where the pin will be repositioned. After re-positioning the glued pin next to the crystal, the laser makes the final cuts of the chosen shape (Figure 4.11*b*, *c*). Thereafter, the cut out film is automatically transferred into a cryogenic stream, from which the pin can be manually placed into a cryo-vial for storage before screening experiments at a synchrotron beamline.

Due to the challenges during manual crystal harvesting of Mhp1 crystals as a result of the fragile nature of the crystals, the CrystalDirect system was recognised as an optimal approach to minimise crystal handling and potentially improve the diffraction data obtained from the Mhp1 crystals.

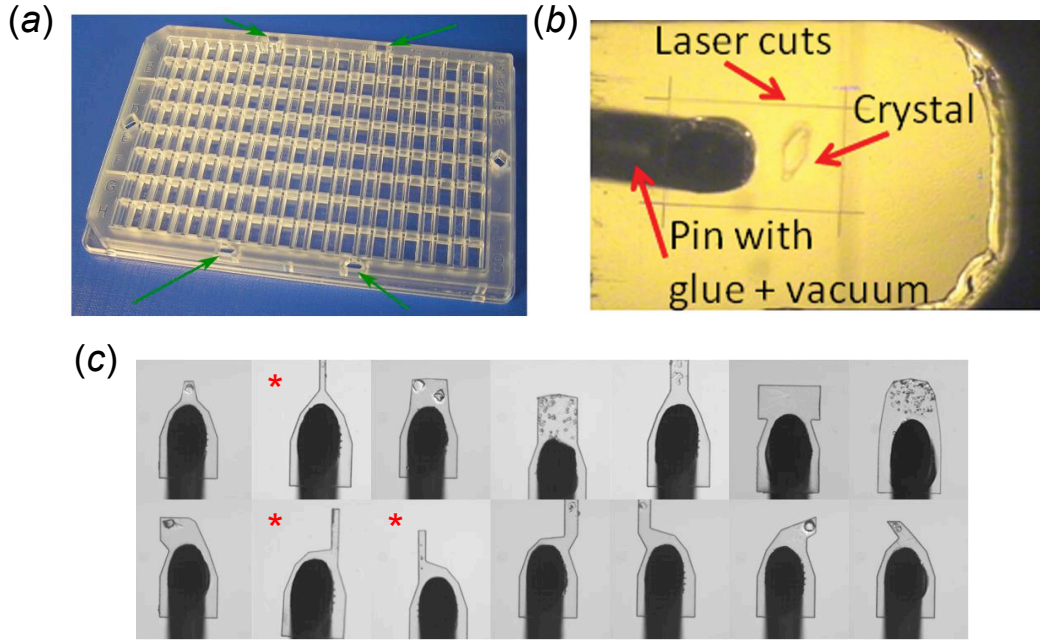


Figure 4.11 The CrystalDirect system for contact-less crystal harvesting by photoablation.

Crystals for contact-less harvesting are grown in a sitting-drop configuration on a thin polyimide film in custom-made SBS format trays (*a*). The trays are placed into the robotic setup using holes for holding the tray in place (green arrows in *a*) The automated crystal harvesting procedure involves marking the crystal to be harvested in the CrystalDirect computer software interface, after which a fragment of film around the crystal is cut out by the laser (*b*). Panel *b* is adapted from the publication describing the initial version of the CrystalDirect system (Cipriani *et al*, 2012). In the version of the system used in this work, the laser had been configured to cut out a great variety of shapes (*c*), increasing the harvesting possibilities from crystallisation droplets containing multiple crystals. For the needle-shaped Mhp1 crystals obtained in this project, elongated shapes (marked with red asterisks) were typically used. The photos in panels *a* and *b* are adapted from Cipriani *et al.* 2012, while panel ***c*** is adapted from the website of CRIMS©, the Crystallization Information Managements System at EMBL, Grenoble (https://embl.fr/htxlab/index.php?option=com_content&view=article&id=30&Itemid=327).

4.6.2 Method and mutants tested

Mhp1 crystals were crystallised manually in custom-made crystallisation trays supplied by the CrystalDirect team and transported to Grenoble in a sealed polystyrene box at room temperature. Crystals of wild-type and mutant Mhp1, in complex with various ligands (summarised in Table 4.5) were harvested using the CrystalDirect system. In order to maximise the chances of obtaining good crystals in a sitting-drop configuration in a different tray, crystallisation droplets (0.3-0.5 µl) were set up in various ratios (1:1, 1:2, 2:1) with and without the streaking motion described in 4.4.3. In

addition, samples of the Mhp1 mutant Gln42Asn were also shipped on dry ice to the EMBL for automatic crystallisation on site.

Table 4.5 Summary of the crystals of Mhp1 variants, manually crystallised in custom-made trays, harvested using the CrystalDirect system.

Variant	Number of samples harvested
Wild-type (apo)	6
Gln42Asn (apo)	0
Gln42Asn-(L-BH)	4
Gln42Asn-(L-IMH)	0
Lys110Leu-(L-BH)	11*
Lys110Leu-(L-IMH)	5
Ser312Ala-(L-BH)	0
Ser312Ala-(L-IMH)	0
Asn318Gly (apo)	14*

* Diffraction obtained.

However, no crystals were obtained from the automatic crystallisation, indicating that the shipping conditions on dry ice may have affected the crystallisability of the protein. During this project, this was the only occasion where protein was shipped on dry ice for crystallisation trials. It has been shown that the pH of protein solutions stored in eppendorf tubes on dry ice is lowered (Murphy *et al*, 2013), which may be detrimental for the quality of the Mhp1 mutant samples for crystallisation trials.

In total, 40 samples were harvested and screened for diffraction, yielding one data set of the Mhp1 mutant Lys110Leu in complex with L-BH (see Table 4.6 for data statistics) from a crystal which was cut-out from the film using an elongated shape (Figure 4.12*c*). The structure could be solved in space group $P2_12_12_1$ with the unit cell parameters a = 87.5, b = 101.8 , c = 105.6, $\alpha = \beta = \gamma = 90°$. The lattice of the crystal of Lys110Leu in complex with L-BH was contracted to a similar degree as that obtained by subjecting wild-type Mhp1 crystals to controlled dehydration (4.5.2). During contact-less harvesting, most of the crystallisation droplet is aspirated away prior to the excision of the film fragment, followed by rapid transition into a cryo-stream.

As during controlled dehydration, the aspiration step during contact-less harvesting may be eliminating a significant amount of excess solvent, giving rise to a tighter packed lattice.

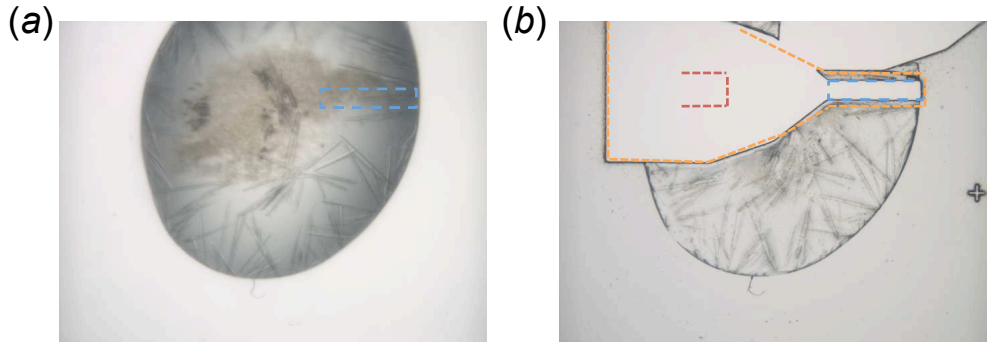


Figure 4.12 Contact-less harvesting of Lys110Leu crystals.

The photos depict the same crystallisation droplet containing crystals of the Mhp1 mutant Lys110Leu in complex with L-BH before (*a*) and after (*b*) harvesting. The blue dashed rectangle in *a* and *b* indicates the position of the crystal that was marked for harvesting, and the portion of the excised film fragment (encircled by orange dashed lines) where it was contained. The red dashed lines show the approximate position of the incisions made for aspiration, which was made possible by cutting a small hole in the film next to the crystals (small cross in *b*).

4.6.3 Investigation of cooling conditions during in-house harvesting

Similar to the structure of wild-type Mhp1 obtained by controlled dehydration (section 4.5.2), the crystals harvested and tested using the CrystalDirect system also had a contracted lattice with respect to the original data sets obtained from crystals that were subjected to conventional cooling by plunging into liquid nitrogen, as discussed in 4.3.2.

What was common between the controlled dehydration method and the contact-less harvesting, was that crystals in both methods were cooled by transfer into a cryogenic stream, rather than being plunged directly into liquid nitrogen.

In principle, cooling by plunging into liquid nitrogen or by transfer into a gaseous nitrogen stream has little effect on the outcome of the diffraction (Warkentin & Thorne, 2007). This is due to the fact that liquid nitrogen evaporates to form a nitrogen gas layer. Cooling rates are significantly slower in gas compared to in liquid, and with protein crystals with a size average in the µm range, they traverse through the gaseous layer for a considerable amount of time. Hence, the actual cooling of such crystals

occurs in the gaseous layer (Warkentin & Thorne, 2009), mimicking the relatively slow cooling conditions during cryo-cooling by transfer into a cryogenic nitrogen gas stream.

During in-house crystal harvesting and cooling, it was attempted to maintain an even amount of liquid nitrogen in the dewar by continuously topping up the vessel containing the nitrogen. Despite this, the thickness of the gas layer was likely to vary, for example with the daily temperature and humidity of the room.

Table 4.6 Data collection and refinement statistics for the structure of the Lys110Leu mutant in complex with L-BH, obtained by contact-less harvesting.

	Lys110Leu-(L-BH)
Wavelength (Å)	0.9763
Temperature (K)	100
Space group	$P2_12_12_1$
Molecules per assym. unit	1
Unit cell parameters (Å, °)	a = 87.6, b = 101.8, c = 105.6, $\alpha = \beta = \gamma = 90$
Resolution (Å)	45.86 - 4.32 (4.48-4.32)
No. of measured reflections	45464 (6603)
No. of unique reflections	3742 (574)
Multiplicity	6.9 (6.5)
Completeness (%)	98.4 (90.3)
$<I/\sigma(I)>$	5.7 (1.4)
$CC_{1/2}$	0.998 (0.861)
Wilson B factor (Å^2)	163.2
Anisotropy ΔB (Å^2)	106.21
R_{merge} (%)	0.175 (1.282)
R_{meas} (%)	0.203 (1.499)
$R_{p.i.m.}$ (%)	0.103 (0.769)
Matthews coefficient (Å^3 Da^{-1})	4.31
Solvent content (%)	71.50
R_{work} (%)	29.41 (35.72)
R_{free} (%)	33.15 (32.40)
R.m.s.d.	
Bonds (Å)	0.0060
Angles (°)	1.0319
Ramachandran outliers (%)	0.66
Average B factor (Å^2)	203.0

Values in parentheses are for the highest resolution shell.

In addition, many additional factors come into play due to human intervention, such as the time spent trying to catch the crystal in the loop, associated mechanical stress on the crystal and the time spent in the atmosphere during transfer of the crystal from the crystallisation droplet to the dewar containing liquid nitrogen. For fast plunging, the level of nitrogen in the dewar will affect the air-path and, despite keeping the level of nitrogen topped up, it is difficult to ensure full consistency between each harvesting attempt.

Therefore, the two cooling methods were investigated to determine whether cooling in a nitrogen stream could be a more reproducible method for crystal cooling, potentially giving rise to a tighter lattice (as seen for crystals subjected to controlled dehydration or contact-less harvesting).

Two Mhp1 mutants, Lys110Leu and Gln42Asn, were tested on the same occasion, in order to make sure they were subjected to the same temperature and humidity. All the crystals of the Mhp1 mutant Lys110Leu came from one tray (approximately five months old) and all the crystals of the Mhp1 mutant Gln42Asn (approximately six month old) samples came from another tray. No direct correlation was observed between fast plunging into liquid nitrogen vs. transfer into a cryo-stream, confirming that the cooling of crystals most likely occurs in the gas layer above the liquid nitrogen. The reason for obtaining a contracted lattice from both controlled dehydration using the HC1 device and the contact-less harvesting using the CrystalDirect system may thus be attributed to the extensive elimination of excess solvent around the crystals.

Table 4.7 Comparison of cooling method and lattice parameters.

Sample	**Cooling method**	**Unit cell parameters *a*, *b*, *c* (Å)**	**Resolution (Å)**
Lys110Leu-IMH	Cryo-stream	90, 106, 107	4.0
Lys110Leu-IMH	Cryo-stream	89, 105, 107	4.1
Lys110Leu-IMH	Rapid plunging	90, 107, 107	3.9
Gln42Asn-BVH	Cryo-stream	90, 106, 108	3.8
Gln42Asn-BVH	Cryo-stream	88, 103, 106	3.7
Gln42Asn-BVH	Cryo-stream	92, 107, 109	3.7
Gln42Asn-BVH	Rapid plunging	89, 105, 106	3.8

4.6.4 Discussion

The CrystalDirect experiments have several things in common with the *in situ* studies described in section 4.3.4. The crystallisation droplets in the trays had similar volumes, were both set up in a sitting-drop configuration and were both transported for a significant amount of time in a polystyrene box to the synchrotrons where the experiments were conducted. Reasons for the fact that diffraction data could be collected from crystals harvested using the CrystalDirect system of sufficient quality to solve the crystal structure of the Mhp1 mutant Lys110Leu in complex with L-BH, in comparison to the *in situ* crystals is that the liquid from the crystallisation droplet was aspirated from the film around the crystal(s). This suggests that the hydration level of the crystals, as well as the amount of surrounding liquid, is a very important factor for obtaining high quality diffraction data from crystals of detergent-solubilised Mhp1.

The diffraction screening performed on crystals harvested using the CrystalDirect systems did not have a high success rate, where only 1 out of 47 crystals screened for diffraction gave rise to a data set of sufficient quality to solve the structure of the protein (Table 4.5). A reason for the low success rate of data sets collected vs. number of crystal samples harvested could be that the aspiration step was not successful in every harvesting attempt. As indicated by *in situ* studies of Mhp1 crystals, minimising the amount of liquid surrounding the crystals is potentially very important, either by increasing the signal-to-noise ratio or by minimising the hydration level of the crystals and their lattice.

As in all the other steps of the automatic crystal harvesting, the aspiration step was monitored via a live feed from a camera filming the crystal being harvested. In many cases for the crystallisation droplets containing Mhp1 crystals, the droplets appeared to be covered with a skin, which did not allow for full aspiration of the droplet. This was clearly seen with liquid still being visible in the droplet and in some cases the crystals would even float around in the remaining liquid. In these cases, the automatic harvesting process was paused and the aspiration step was repeated, sometimes with a longer time length of the aspiration, in order to remove as much liquid as possible. This repeated process of aspiration was likely to affect the crystal quality by applying mechanical stress. In addition, the laser will also excise the skin covering the droplets. If there was residual skin covering the crystallisation droplet upon the final laser excision step, the skin would also be covering the harvested crystal on the excised film fragment, which could have also

affected the quality of the resulting diffraction. Despite this, the harvesting experiments performed using the CrystalDirect system suggests that this is a viable harvesting method, even for membrane proteins.

Two different cooling methods were tested to see whether they have an effect on the lattice tightness. These were either fast plunging into liquid nitrogen or direct transfer into a cryogenic nitrogen gas stream at 100 K. It was found that transfer into a cryogenic stream appeared to result slightly more often in a tighter lattice (as seen for dehydrated crystals), compared to direct plunging into liquid nitrogen. However, as no strong correlation was seen between the method of crystal cooling and the lattice dimensions, it is more likely that it is the extensive removal of excess solvent around the crystals during controlled dehydration and contact-less harvesting (aspiration) that results in a tighter and more well-ordered lattice.

Comparing the data statistics from the crystal of the Mhp1 mutant Lys110Leu in complex with L-BH, which was harvested by contact-less harvesting, to a crystal which was cooled by plunging into liquid nitrogen (Table 4.8), a similar improvement is seen as for a crystal that was subjected to controlled dehydration (see Table 4.4 in section 4.5.3). The higher merging *R*-factors observed for the data obtained by contact-less harvesting could be caused by, for example, the shape of the film fragment used for the crystal, as it was observed that the longer cut-out shapes often flickered in the X-ray beam from the flow of the cryogenic stream.

Table 4.8 Comparison of data statistics obtained from a crystal harvested by contact-less photo-ablation and a manually harvested crystal.

	Lys110Leu-(L-BH) (contact-less harvesting)	Gln42Asn-(L-BH) (manual harvesting)
Unit cell parameters (Å)	*a* = 87.6, *b* = 101.8, *c* = 105.6	*a* = 91.9, *b* = 107.1, *c* = 107.2
High resolution limit (Å)	4.32	3.74
Wilson *B* factor (Å^2)	163.2	309.5
Anisotropy Δ*B* (Å^2)	106.21	146.5
R_{merge} (%)	0.175 (1.282)	0.040 (1.192)
R_{meas} (%)	0.203 (1.499)	0.056 (1.641)
$R_{p.i.m.}$ (%)	0.103 (0.769)	0.038 (1.124)
Solvent content (%)	71.50	74.56
Average *B* factor (Å^2)	203.0	253.0

4.7 General discussion and conclusions

In this Chapter, the optimisation of the crystallisation conditions for Mhp1 crystals have been described. Initially, the previously published conditions for the orthorhombic crystal form were successfully reproduced, yielding needle-shaped crystals. These crystals were harvested in-house and did not give rise to high quality diffraction, diffracting at best to 6 Å at a synchrotron source. Moreover, the diffraction was highly anisotropic and disordered.

The crystallisation conditions for the hexagonal crystal form, in which the (seleno-L-methionine derivatised) Mhp1 protein had previously crystallised in an inward-open conformation were not reproduced. This could have been due to differences in protein expression and/or purification batch specific carry over of a ligand from the media used during the growth of methionine auxotrophic *E. coli* cells for the production of the protein in comparison to the original experiment. The crystal structure of seleno-L-methionine derivatised Mhp1 suggests that the obtained inward-open conformation may be stabilised by this unidentified ligand.

The low quality diffraction obtained from Mhp1 crystallised under the conditions for the orthorhombic crystal form is likely due to the fragility of the Mhp1 crystals. This probably stemmed from the high solvent content due to the protein being solubilised in detergent. Crystal breaking and dissolving during harvesting attempts led to the exploration of alternative crystallisation and data collection strategies, that would minimise crystal handling. Mhp1 crystals grown either on MicroMeshes™ or in *in situ* plates were screened for diffraction, however, this gave rise to either no or very weak and low-resolution diffraction.

The anisotropic diffraction limit of 6 Å of the initial Mhp1 crystals was overcome by introducing drop morphology modification into the crystallisation experiment. By applying a ‘streaking’ motion when setting up the crystallisation droplets, a size gradient of crystals was formed in the droplet. The size gradient resembled the gradients obtained in counter-diffusion aided crystallisation experiments performed using the commercially available CrystalHarp™ and CrystalFormer trays. The size gradient typically resulted in a few large mm-long crystals growing out of a shower of smaller crystals. These crystals diffracted reproducibly to beyond 4 Å and were used to solve the structures of for example the Mhp1 mutants Gln42Asn in complex with L-BH and L-IMH (further discussed in Chapter 5).

These crystals did, however, still diffract anisotropically, due to the morphology of the crystals. Higher resolution is always desired in order to gain more detailed insight in to the structure of the protein. Therefore, further strategies were explored for the improvement of the Mhp1 crystals and their resulting data, including controlled dehydration and contact-less harvesting. Both controlled dehydration and contact-less harvesting were successful strategies to obtain a tighter lattice and improved-quality electron density maps. This was probably due to the removal of the solvent surrounding the crystal, causing a contraction of the detergent phase and thus more well-ordered packing in the crystals. Additionally, during contact-less harvesting the mechanical stress normally put on crystals during manual harvesting is minimised. However, these methods have certain logistic drawbacks. Controlled dehydration studies are tedious and time-consuming with the required steps of manual harvesting, closing the experimental hutch, stepwise dehydration (minutes per step) while collecting diffraction images to evaluate the dehydration process, before final data collection. If the mounted crystal does not diffract, or the dehydration procedure causes severe damage to the crystal, the experimental hutch needs to be re-opened and a new crystal needs to be mounted. This makes it difficult to achieve the same throughput as would normally be the case with automatically mounted cryo-cooled samples during synchrotron beamtime. It is thus challenging to establish a robust dehydration protocol. The drawback of contact-less harvesting is the limited availability of the CrystalDirect system, currently only being available at the EMBL in Grenoble. The use of this facility requires transportation and/or shipping of protein samples or crystal to the site, which can also cause damage to the protein or crystal samples.

Nevertheless, these studies have demonstrated the importance of controlling the hydration level of detergent-solubilised Mhp1 crystals in order to obtain higher quality well-ordered crystals. In order to overcome the logistical drawbacks discussed above, and increase the throughput of crystallographic studies of Mhp1 crystals, strategies and methods for solvent removal should be explored for use in the home laboratory. For example, there are several commercially available sample mounts with special channels for liquid removal by wicking available. In order to control the humidity during harvesting to avoid unnecessary swelling or contraction of the crystal during harvesting, an in-house humidified gas stream could be used.

In conclusion, this Chapter has outlined the steps taken in order to improve the diffraction obtained from Mhp1 wild-type and mutant crystals for further

structural studies. The various experiments have given an insight into the challenges usually encountered during membrane protein crystallisation, and have demonstrated that the hydration level is important for the crystals obtained here. Future studies should involve further exploration of solvent removal strategies in order to obtain well-ordered crystals and higher quality diffraction data.

Chapter 5: X-ray crystal structures of Mhp1 mutant-ligand complexes

5.1 Introduction

The previously published crystal structures of Mhp1 are believed to represent three key conformational states of the alternating access transport mechanism; the outward-open, occluded and inward-open conformations (Shimamura *et al*, 2010; Weyand *et al*, 2008). Although these structures provide insight into the progression of the transport mechanism, they provide only static snapshots of a highly dynamic process. Despite efforts to link these still snapshot images and shed light on the events occurring between them, using for example computer simulations (Shimamura *et al*, 2010; Adelman *et al*, 2011) or spectroscopic studies (Kazmier *et al*, 2014), many questions about the exact mechanism of transport still remain unanswered. One of the most intriguing features of the substrate transport in this type of secondary active transporters is the mechanism by which the transport of ligands and ions are coupled.

Previous mutagenesis and subsequent functional studies have identified several mutants of Mhp1 that display altered ligand binding and transport behaviour, where hydantoin ligands either bind with very low affinity to Mhp1 or where the sodium-dependent binding is abolished (Jackson, 2012). These findings suggest that the mutations could be altering the conformation that the Mhp1 protein adopts in solution, such that it prevents ligand binding or affects sodium-dependent binding. This makes the identified mutants ideal candidates for structural analysis in order to explore more conformations of the protein and further probe the transport mechanism.

The purification and characterisation of several of these mutants was described In Chapter 3. In order to investigate how the structure of Mhp1 was affected by the mutations and to understand the structural basis for the altered ligand binding, all the variants described in Chapter 3 were subjected to crystallisation trials. After a series of optimisation steps in order to improve the diffraction quality of the Mhp1 mutant crystals, which were obtained by modifying previously published crystallisation conditions (described in Chapter 4), several data sets with diffraction limits beyond 4 Å were collected from crystals of Mhp1 mutants in complex with various ligands. The mutants for which data were collected and crystal structures were solved are Gln42Asn, Ser312Ala and Lys110Leu. These mutants allowed probing of the

structure in different locations in the protein, such as the ligand-binding site (Gln42Asn), the sodium-binding site (Ser312Ala) and the area between the bundle-and-hash motif (Lys110Leu).

5.2 X-ray crystal structures of mutant-ligand complexes of the Mhp1 mutant Gln42Asn

5.2.1 Introduction

One mutant that successfully yielded crystals suitable for X-ray crystallographic studies was the mutant Gln42Asn. Gln42 is situated in the ligand-binding site in the wild-type Mhp1. The hydantoin moiety is stabilised in the binding site by π-stacking interactions with Trp117, as well as a hydrogen-bond network formed by the residues Gln121, Asn318 and Gln219. The aryl moieties of hydantoin ligands interact with Trp220, forming π-interaction (Weyand *et al*, 2008; Simmons *et al*, 2014) (Figure 5.1).

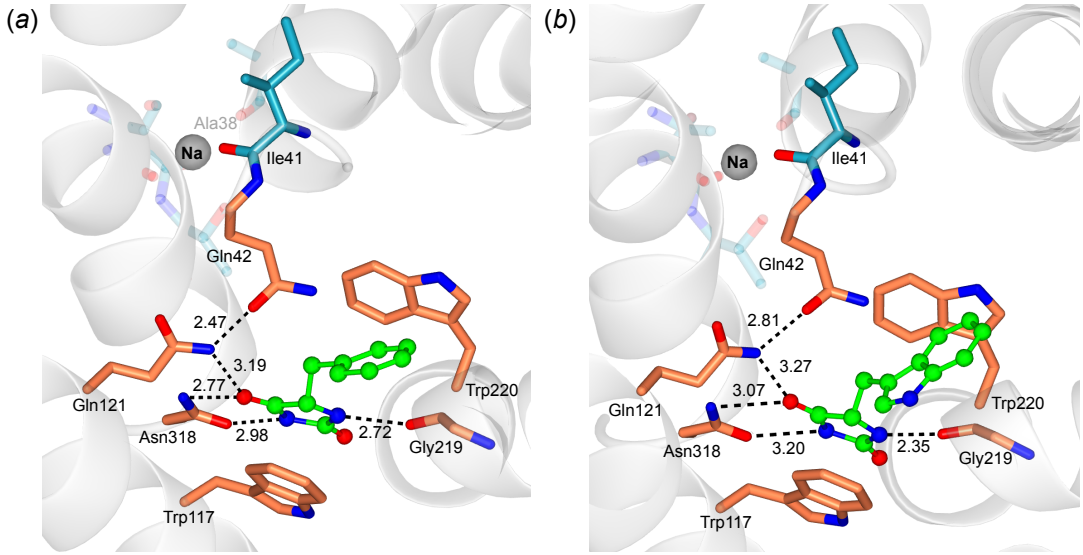


Figure 5.1 Illustration of the ligand-binding site in wild-type Mhp1.

The figure illustrates the ligand-binding sites of wild-type Mhp1 in complex with L-BH (*a*) and in complex with L-IMH (*b*), where figure in (*a*) is replicated from Figure 3.7*a*. The ligand binding site residues are highlighted in coral, the ligand L-BH as a ball-and-stick model in green and the nearby sodium as a grey sphere. Hydrogen bonds between the L-BH and nearby residues are displayed as black dashed lines and distances stated in units of Å. The sodium binding site residues are displayed in semi-transparent cyan, apart from Ile41, which is displayed in cyan as a reference for the orientation of the sodium binding site in Figure 5.9. The non-carbon atoms are coloured red (oxygen) and blue (nitrogen). The ribbon diagram of the whole protein is displayed in semi-transparent grey. The PDB codes for the respective structures depicted in this figure are 4D1B (*a*) and 4D1A (*b*) (Simmons *et al*, 2014).

Gln42 could also be participating in the hydrogen-bonding network that stabilises the hydantoin moiety of the ligands when bound, since it is in hydrogen-bonding distance to Gln121, which in turn makes a hydrogen bond to a carbonyl oxygen on the hydantoin moiety together with Asn318 (Figure 5.1).

Gln42 is also connected to the nearby sodium-binding site, with the carbonyl oxygen of the adjacent residue Ile41 participating in the coordination of the sodium ion. Due to its position, connecting the sodium- and ligand binding sites, Gln42 may play a potential role in the coupling mechanism. This hypothesis is supported by functional studies of the Gln42Asn mutant. The tryptophan fluorescence quenching assays of the Gln42Asn mutant show that the sodium-dependent binding of L-BH is perturbed as a result of the mutation (Figure 5.2). 15 mM sodium, which in the wild-type protein is sufficient to saturate all binding sites with L-BH, is not sufficient to saturate the binding sites in the Gln42Asn mutant, resulting in an increase in the apparent K_d (Table 5.2).

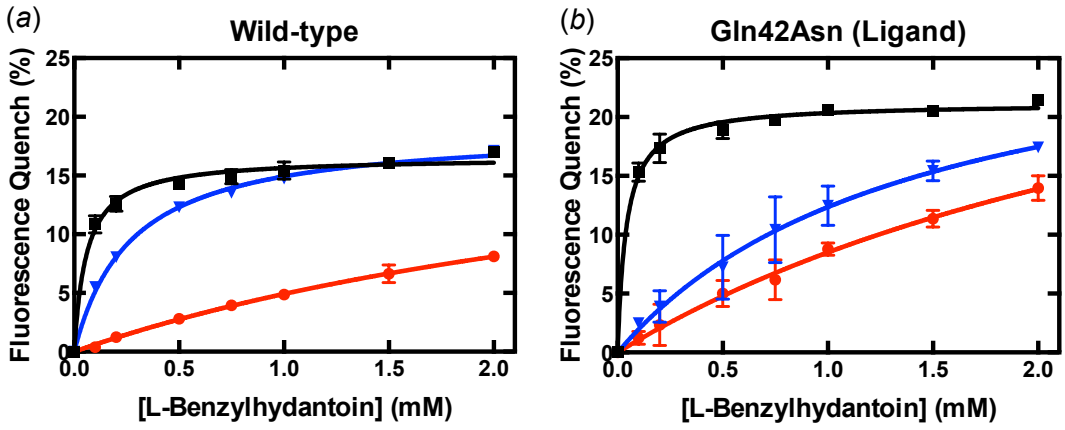


Figure 5.2 The binding of L-BH to wild-type Mhp1 and the Mhp1 mutant Gln42Asn, solubilised in NM.

The figure shows the quench of the tryptophan fluorescence during titrations of L-BH (0-2 mM) with wild-type Mhp1 (*a*) and the Mhp1 mutant Gln42Asn (*b*). The titrations were performed in the presence of 0 mM (red), 15 mM (blue) and 140 mM (black) added sodium. Mhp1 samples (in 10 mM Tris-HCl pH 7.6, 2.5% v/v glycerol, 0.5% w/v NM) were diluted to 140 µg/ml using buffers containing 50 mM Tris-HCl pH 7.6, 2% v/v DMSO, 0.5% w/v NM, 0/15/140 mM NaCl and 140/125/0 mM choline chloride. The error bars display the standard error of the mean from three replicates of each titration. The titrations show that in the presence of 15 mM sodium, binding of L-BH is not enhanced in Gln42Asn to the same level as in the wild-type protein. At least 140 mM sodium is instead required to achieve a similar effect in the mutant as in the wild-type protein.

However, the lost binding capacity is restored in the presence of 140 mM sodium. A similar behaviour is observed for the more potent ligand L-IMH, where ^{14}C L-IMH cell uptake assays show that when 150 mM sodium is present, as much as 82.5 ± 2.5% of the wild-type uptake of the ligand occurs in the Gln42Asn mutant (Jackson, 2012). This indicates either that the sodium is not able to bind to its site between residues Ala38, Ile41, Ala309 and Ser312 as easily due to binding site remodelling as a result of the mutation, or that the protein has adopted a low-affinity conformation that diminishes the ability of the ligand to bind.

In this section, crystal structures of the Gln42Asn mutant in complex with the ligands L-BH and L-IMH are presented and compared to previous structures of the wild-type protein in complex with the same ligands. Two crystal structures of Gln42Asn in complex with L-BH were obtained from different buffer conditions; high- and low-sodium conditions in order to investigate effect of sodium on the conformation of the protein in the presence of ligand.

5.2.2 The structure of the Mhp1 mutant Gln42Asn in complex with L-BH

The structure of the mutant Gln42Asn in complex with L-BH was determined at a resolution of 3.75 Å (Table 5.1). and solved in space group $P2_12_12_1$ with unit cell parameters a = 91.8, b = 107.1, c = 107.2 Å, $\alpha = \beta = \gamma = 90°$. These data came from the first crystals that had been obtained by optimising the crystallisation experiment by drop morphology modification (described in section 4.4.3-4). The crystal from which the data were collected was cooled by direct plunging into liquid nitrogen. Despite the fact that the unit cell parameters were similar to previously published crystal structures of the wild-type Mhp1 protein and the protein molecules were found with similar orientations and positions in the unit cell, direct Fourier transformation and rigid-body fitting of these structures as initial models did not work. This was indicated by high R-factors (R_{work} and R_{free} > 39%) after rigid body refinement. Therefore, the structure was solved by molecular replacement. The three different published conformations of the Mhp1 protein; outward-open (2JLN), occluded with various ligands (2JLO, 4D1A and 4D1B) and inward-open (2X79) were tested as molecular replacement search models and the resulting solutions were subjected to rigid-body refinement, followed by restrained refinement of the resulting rigid-body fitted model. Prior to using these structures as molecular replacement search models all bound ligands, ions and metals were removed. The outward-open and the occluded

structures were placed in similar orientations compared to previous structures, but the best initial *R*-factors were obtained from the occluded structures (R_{work}/R_{free} ≈ 27/32%). This indicated that the Gln42Asn protein predominantly adopted an occluded conformation in the crystal. The structure of Mhp1 in complex with L-IMH (4D1A) was used as the final search model in molecular replacement as this structure came from the highest quality data (A. Cameron, personal communication). The inward-open structure (2X79) was also tested as a search model but gave high *R*-factors after initial refinement (R_{work} and R_{free} > 39%). No significant difference electron density was observed in the F_O — F_C map around the site of the mutation. However, at the low working resolution of this structure (3.75 Å), the mutation, which essentially corresponds to a change of length of the side chain of 1.2-1.5 Å will be undetectable unless it causes a major conformational change in the protein. There was no major indication in the F_O — F_C map of major alterations to the conformational state of the protein, and the occluded conformation seems to be minimally perturbed when L-BH is bound.

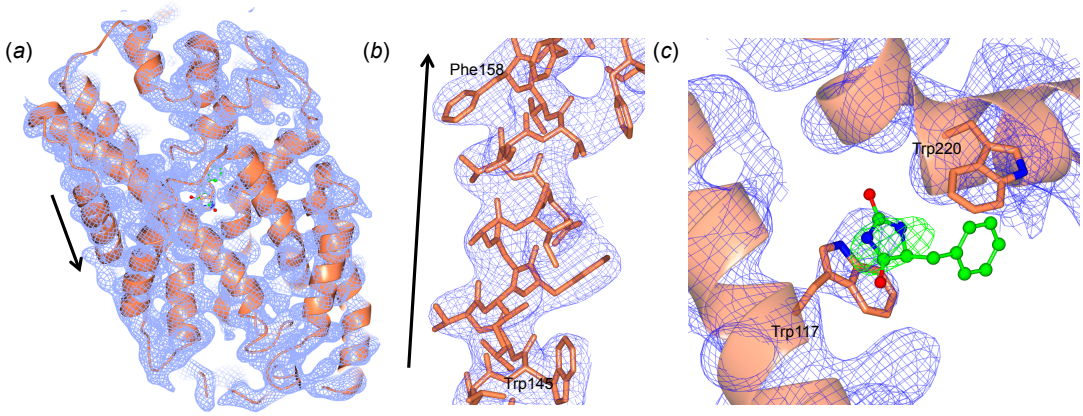


Figure 5.3 The structure of the Gln42Asn mutant in complex with L-BH.

The figure in (*a*) displays the final refined model of the Mhp1 mutant Gln42Asn in complex with L-BH, as a cartoon representation in coral and the 3.75 Å $2F_O$ — F_C map (blue mesh). The helix marked with a black arrow is zoomed in on in (*b*), depicting residues Trp145-Ile158 in coral to better show the fit of the model to the $2F_O$ — F_C map. (*c*) The $2F_O$ — F_C (blue mesh, contoured at 1.5 σ) and F_O — F_C (green mesh, contoured at 2.0 σ) maps after initial refinement, showing the positive difference density in the F_O — F_C map occurring only for the hydantoin part of L-BH (shown in green as a ball and stick model).

When replacing Gln42 with Asn42 in the model, the residue was initially oriented in the same orientation as Gln42, based on the suggested rotamers in Coot (Emsley *et al*, 2010), pointing towards the aromatic part of the ligand.

In this position, the interaction with Gln121 would be weakened, as the putative hydrogen-bonding distance increased from 2.81 Å to 3.59 Å. After further refinement, the position of Asn42 could be adjusted to better fit the electron density, pointing towards Gln121, bringing the two residues together and potentially strengthening the interaction, restoring the original putative hydrogen bond (to 2.63 Å). However, this did not give rise to a significant change in *R*-factors, and at this resolution it is difficult to determine the exact orientation of a single residue.

In the hydantoin ligand-binding site, positive difference electron density in the F_O — F_C map was clearly seen between Trp117 and Trp220 early in the refinement. L-BH, in the same extended binding mode as seen in 4D1B (Simmons *et al*, 2014), was placed into the binding site. However, the electron density covered only the hydantoin moiety of the ligand (Figure 5.3). The missing density for the benzyl moiety of L-BH suggests that the ligand could be exhibiting flexibility around the carbon-carbon bonds that connect the hydantoin and the benzyl moiety. Due to the extra space generated by the shortening of Gln42 to Asn42 and the resulting reduction in electrostatic repulsion, the L-BH could be adopting several conformations, resulting in the electron density of the benzyl group of the ligand cancelling out between the molecules throughout the crystal.

5.2.3 The structure of the Mhp1 mutant Gln42Asn in complex with L-IMH

For the Mhp1 mutant Gln42Asn in complex with L-IMH, diffraction data were collected with a resolution of 4.0 Å (Table 5.1). The data were of comparable quality to the data of the mutant Gln42Asn in complex with L-BH (section 5.2.2), and were collected on the same occasion, and also came from a crystal that had been optimised by drop morphology modification and cooled by direct plunging into liquid nitrogen. The structure was solved in the same space group as the complex with L-BH, with similar cell parameters ($P2_12_12_1$, a = 90.9, b = 106.9, c = 108.1 Å, α = β = γ = 90°). As for the Mhp1 mutant Gln42Asn in complex with L-BH, 4D1A was used as the search model in molecular replacement and no major differences in overall conformation of the protein were seen during the course of refinement. Positive difference electron density in the F_O — F_C map was seen between Trp117 and Trp220, and L-IMH, in the same binding mode as in 4D1A, was placed into the structure. As seen for the structure of Gln42Asn in complex

with L-BH, there was no electron density for the indolylmethyl group of the ligand in the final refined model (Figure 5.4).

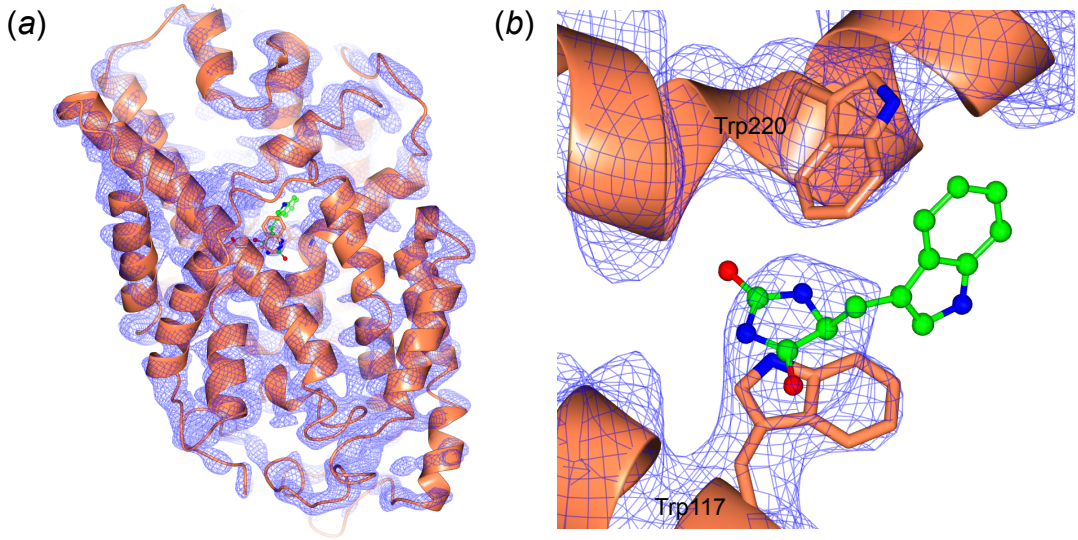


Figure 5.4 The ligand binding site in the structure of the Gln42Asn mutant in complex with L-IMH.

The 4.0 Å $2F_O$ — F_C (blue mesh, contoured at 1.5 σ) for the overall final refined model of the Mhp1 mutant Gln42Asn in complex with L-IMH (*a*), and showing the ligand binding region where only the hydantoin part of L-IMH (shown in green as a ball and stick model) is observed.

This indicated, yet again that the ligand was flexible in its binding site, potentially as a result of the additional space created in the binding site by the shortening of Gln42 to Asn42.

5.2.4 The structure of the Mhp1 mutant Gln42Asn in complex with L-BH under low-sodium conditions

The Gln42Asn-ligand complex structures have shown that the mutation does not significantly alter the overall conformation of the protein compared to ligand-complex structures of the wild-type protein (Simmons *et al*, 2014). As for the wild-type structures with ligands, the Gln42Asn-ligand complex structures adopt an occluded conformation, with each respective ligand bound in the ligand-binding site. No significant changes to the structure at the site of the mutation were observed at the working resolution. One possible reason for this could be that all the crystal structures presented here so far were produced under high-sodium conditions, *i.e.* in the

presence of 100 mM sodium chloride and 100 mM sodium phosphate in the crystallisation mother liquor.

Table 5.1 Crystallographic data processing and refinement statistics for the structures of the Mhp1 mutant Gln42Asn in complex with L-BH and L-IMH.

	Gln42Asn-(L-BH)	Gln42Asn-(L-IMH)
Wavelength (Å)	0.9686	0.9686
Temperature (K)	100	100
Space group	$P2_12_12_1$	$P2_12_12_1$
Molecules per assym. unit	1	1
Unit cell parameters (Å, °)	a = 91.9, b = 107.1, c = 107.2, $\alpha = \beta = \gamma = 90$	a = 90.9, b = 106.9, c = 108.1, $\alpha = \beta = \gamma = 90$
Resolution (Å)	42.47-3.74 (4.18-3.74)	42.62-4.0 (4.47-4.00)
No. of measured reflections	34878 (9511)	37848 (10517)
No. of unique reflections	10926 (2986)	9175 (2517)
Multiplicity	3.2 (3.2)	4.1 (4.2)
Completeness (%)	96.1 (93.9)	98.1 (97.3)
$\langle I/\sigma(I)\rangle$	13 (1.3)	7.9 (2.0)
$CC_{1/2}$	1.00 (0.76)	0.998 (0.897)
Wilson B factor (Å^2)	309.5	321.9
Anisotropy ΔB (Å^2)	146.5	140.1
R_{merge} (%)	0.040 (1.192)	0.076 (0.665)
R_{meas} (%)	0.056 (1.641)	0.098 (0.861)
$R_{p.i.m.}$ (%)	0.038 (1.124)	0.061 (0.541)
Matthews coefficient (Å^3 Da^{-1})	4.83	4.81
Solvent content (%)	74.56	74.45
R_{work} (%)	27.08 (35.28)	26.96 (36.18)
R_{free} (%)	31.30 (33.86)	32.62 (36.01)
R.m.s.d.		
Bonds (Å)	0.0100	0.0090
Angles (°)	1.4114	1.3147
Ramachandran outliers (%)	0.66	1.76
Average B factor (Å^2)	253.0	241.1

Values in parentheses are for the highest resolution shell.

At this sodium concentration, the protein is essentially under conditions under which the binding sites become saturated (Figure 5.5). In the previously reported crystal structure of Mhp1 in the outward-open state, there is an intact sodium-binding site (Weyand *et al*, 2008), while in the inward-open structure the sodium-binding site is disrupted (Shimamura *et al*, 2010). In addition, computer simulations show that a sodium ion escapes from the inward-open form of the protein on a time-scale of < 2 ns (Shimamura *et al*, 2010; Adelman *et al*, 2011). These findings suggest that the sodium could be stabilising the outward-open form, turning the protein over into a conformation with high affinity for the ligands, and which then turns over into an occluded conformation when the ligand has bound.

At lower sodium concentrations (15 mM), the Gln42Asn mutant has lowered affinity for L-BH compared to the wild-type Mhp1 (Figure 5.2), and could be adopting a conformational state where L-BH binding is prevented, potentially an inward-open or empty occluded form (Table 5.2).

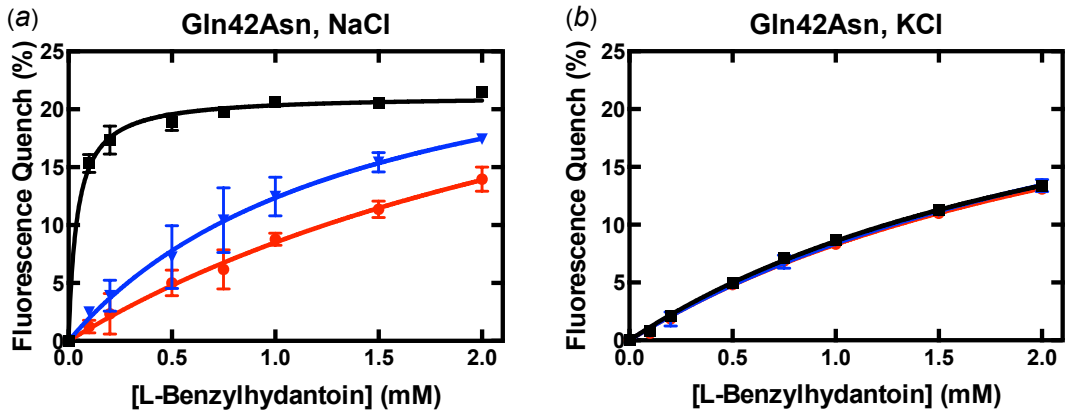


Figure 5.5 The binding of L-BH to the Mhp1 mutant Gln42Asn, solubilised in NM, under high- and low sodium conditions.

The figure shows the fluorescence quench during titrations of L-BH (0-2 mM) with the Mhp1 mutant Gln42Asn in the presence of an increasing sodium concentration and no added potassium (*a*), as well as in the presence of an increasing potassium concentration and no added sodium (*b*). The titrations were performed in the presence of 0 mM (red), 15 mM (blue) and 140 mM (black) added sodium chloride. The buffer components are the same as indicated in Figure 5.2, apart from the titrations with KCl where NaCl was replaced by the same concentrations of KCl. The error bars display the standard error of the mean from three replicates of each titration. The titration curves show that L-BH binding is greatly enhanced with a higher sodium concentration, while this effect is abolished when the NaCl is substituted with KCl.

Therefore, in order to study the conformation of the Gln42Asn mutant without having sodium as the driving force of conformational changes, the tryptophan fluorescence quenching assay was repeated for the Gln42Asn mutant but with potassium chloride instead of sodium chloride. The substitution of the monovalent cation was done in order to later use these conditions in crystallisation trials. This demonstrated that L-BH bound to a slightly lower extent at 0 mM salt, but the synergistic enhancement of the ligand binding seen with increasing sodium concentration was abolished (Figure 5.5).

Table 5.2 Binding parameters for the Mhp1 mutant Gln42Asn and L-BH in the presence of sodium or potassium.

Mhp1	**[Na^+]**	**App. K_d (mM)**	**B_{max} (%)**
Wild-type	0	3.66 ± 0.82	22.93 ± 3.70
	15	0.27 ± 0.015	18.95 ± 0.28
	140	0.059 ± 0.0071	16.56 ± 0.28
Gln42Asn	0	3.47 ± 1.13	38.01 ± 8.78
	15	1.41 ± 0.34	29.81 ± 3.81
	140	0.041 ± 0.0040	21.19 ± 0.23
	[K^+]		
Gln42Asn	0	2.82 ± 0.33	31.71 ± 2.47
	15	3.02 ± 0.44	33.67 ± 3.34
	140	2.55 ± 0.20	30.43 ± 1.53

A potential scenario is that Gln42Asn favours the inward-open conformation, which is a low-affinity form for L-BH, and requires a higher amount of sodium to drive the conformational equilibrium towards a ligand-bound state due to the fact that the sodium binding site has been perturbed by the mutation. The tryptophan fluorescence quenching assay indicates that in the absence of increasing sodium concentration, Gln42Asn remains in the low-affinity state as there is no response to the addition of potassium.

Therefore, in order to investigate whether an alternative conformation is adopted under low-sodium conditions, crystals were produced under similar conditions as in the tryptophan fluorescence quenching assay described above. This was achieved by substituting sodium chloride and sodium phosphate with the equivalent potassium salt and buffer, preserving the ionic strength of the crystallisation mother liquor. This alteration to the

crystallisation conditions did not affect the crystallisability of the protein, and gave rise to crystals of the usual needle-shaped morphology. Wild-type protein, as well as the Gln42Asn mutant in the presence and absence of ligands were subjected to these crystallisation trials and yielded crystals, but diffraction data were successfully collected only from Gln42Asn crystals in the presence of L-BH.

The structure of the Gln42Asn mutant in complex with L-BH under low-sodium conditions was determined to a resolution of 3.80 Å (Table 5.3). and was obtained from a crystal that was cryo-cooled by fast plunging into liquid nitrogen. Even though the crystal was not subjected to controlled dehydration, its unit cell is contracted to a similar degree as previously obtained by controlled dehydration using the HC1, when compared to the original Gln42Asn models in complex with L-BH and L-IMH. One possible reason for this is that the crystal that was used to collect these data came from a crystallisation tray that was approximately five months old[2]. During this time, being stored in a ventilated crystal tray incubator, it is very likely that the droplet was slowly becoming dehydrated, therefore giving rise to a similar effect as during controlled dehydration on the beamline (Kazmier *et al*, 2014; Abergel *et al*, 2003; Abergel, 2004).

Despite the indication from the tryptophan fluorescence quenching assay that Gln42Asn could be adopting an alternative conformation under these crystallisation conditions, the structural data indicated that the protein adopted an occluded conformation in the crystal, with a potassium ion bound in the sodium binding site.

Using an occluded structure (4D1A) as the search model in molecular replacement resulted in the lowest *R*-factors after initial refinement of each molecular replacement solution, and no apparent conformational deviation from the occluded conformation was seen in the resulting electron density map for Gln42Asn in complex with L-BH under low sodium conditions.

In the initial stages of refinement, positive difference electron density was clearly seen in both the ligand and ion-binding sites, and the L-BH together with a potassium ion could be modelled. In the structure of the Gln42Asn mutant in complex with L-BH under high-sodium conditions (5.2.2), there was no electron density for the aromatic portion of the ligand, indicating that the ligand had gained flexibility and adopted several conformations in its

---

[2] Set-up date of tray: 8th July 2014, harvesting date: 9th December 2014

binding site. The electron density resolved the L-BH better than under high-sodium conditions, with electron density covering the benzyl group as well as the hydantoin moiety of L-BH (Figure 5.6).

Table 5.3 Crystallographic data processing and refinement statistics for the structures of the Mhp1 mutant Gln42Asn in complex with L-BH under low-sodium conditions.

	Gln42Asn-(L-BH), low-sodium/high-potassium conditions
Wavelength (Å)	0.9795
Temperature (K)	100
Space group	$P2_12_12_1$
Molecules per assym. unit	1
Unit cell parameters (Å, °)	a = 89.0, b = 105.3, c = 105.6, $\alpha = \beta = \gamma$ = 90
Resolution (Å)	57.14 - 3.80 (4.25-3.80)
No. of measured reflections	58877 (12165)
No. of unique reflections	9299 (1986)
Multiplicity	6.3 (6.1)
Completeness (%)	90.1 (68.9)
$<I/\sigma(I)>$	11.5 (1.1)
$CC_{1/2}$	0.999 (0.702)
Wilson B factor (Å^2)	275.4
Anisotropy ΔB (Å^2)	151.1
R_{merge} (%)	0.072 (2.139)
R_{meas} (%)	0.085 (2.562)
$R_{p.i.m.}$ (%)	0.045 (1.392)
Matthews coefficient (Å^3 Da^{-1})	4.53
Solvent content (%)	72.86
R_{work} (%)	26.36 (44.49)
R_{free} (%)	31.48 (37.22)
R.m.s.d.	
Bonds (Å)	0.010
Angles (°)	1.447
Ramachandran outliers (%)	0.66
Average B factor (Å^2)	235.9

Values in parentheses are for the highest resolution shell.

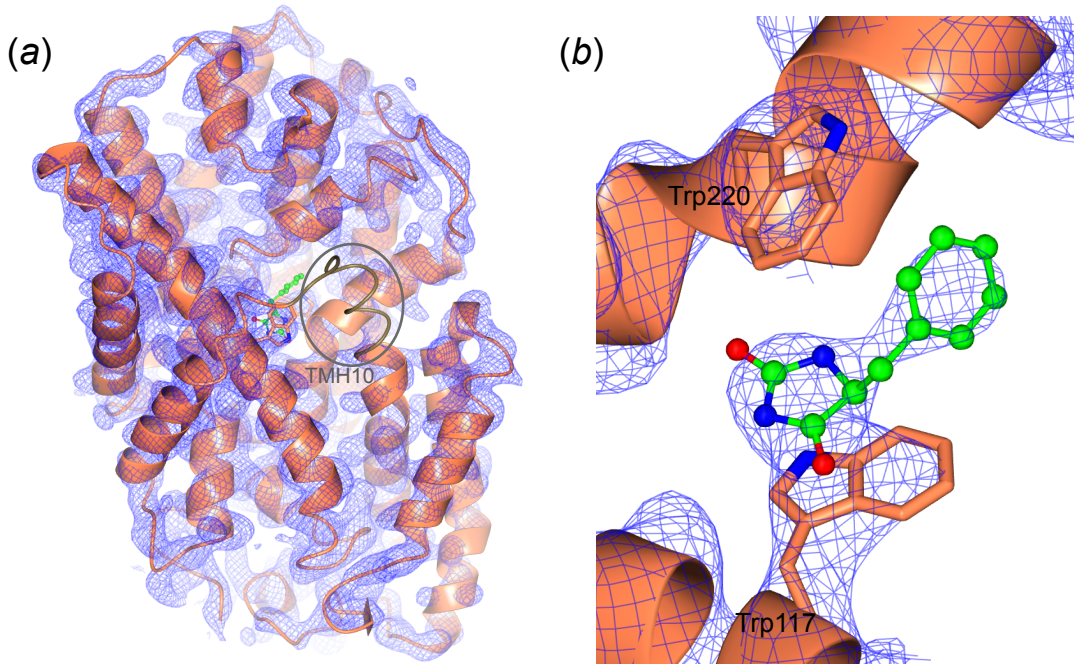


Figure 5.6 Ligand binding site in the structure of the mutant Gln42Asn in complex with L-BH under low-sodium conditions.

The 3.8 Å $2F_O$ — F_C (blue mesh, contoured at 1.5 σ) map after refinement, showing the entire protein as a cartoon representation in coral (*a*), as well as the ligand binding site where the entire L-BH (shown in green as a ball and stick model) is resolved in the electron density map. TMH10 is encircled in (*a*) and coloured in brown in as depicted in Figure 5.7.

Another feature that distinguished this structure of the Gln42Asn mutant in complex with L-BH under low-sodium and high-potassium conditions, compared to the same complex structure under high-sodium conditions, was that part of the trans-membrane helix 10 (TMH10) was not resolved in the $2F_O$ — F_C map (Figure 5.7). This indicates that the segment (residues Phe355-Leu365) adopts multiple conformations in the crystal. Omitting the loop and refining the structure did not recover any positive difference density features in the F_O — F_C map for either the outward-open or occluded conformations. Superposing an occluded structure with an outward-open TMH10 resulted in negative difference density for the open conformation of TMH10, while the closed conformation remained unresolved.

5.2.5 Diffraction data collection on crystals of the Mhp1 mutant Gln42Asn in complex with BVH

Due to the uncertainties in how the ligands bind in the Gln42Asn mutant, particularly for the aryl groups of L-BH and L-IMH under high sodium

conditions, diffraction data collection was attempted on crystals that had been produced in the presence of a brominated form of L-BH; BVH. Bromine gives rise to anomalous scattering at its absorption K-edge (13.4737 keV), which is within the accessible wavelength range at synchrotron beamlines for macromolecular crystallography.

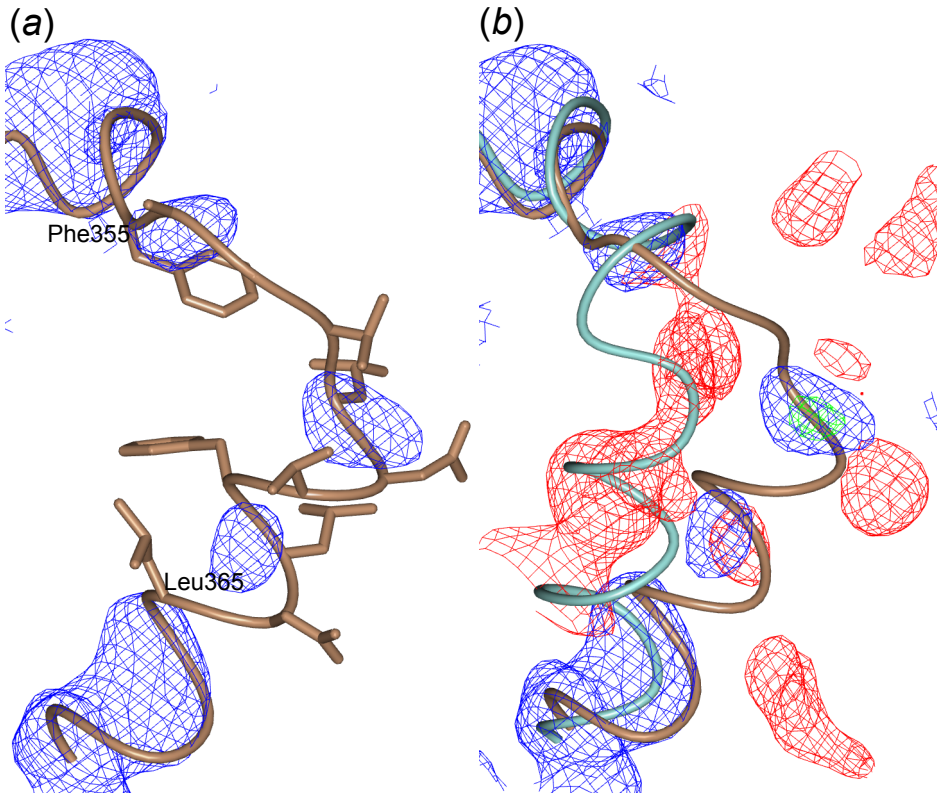


Figure 5.7 Cross-refinement to confirm the conformation of a segment of TMH10 in the structure of Gln42Asn-(L-BH) under low-sodium conditions.

(*a*) In the final refined structure, residues Phe355-Leu366 in the occluded conformation (brown) were poorly resolved in the $2F_O$ — F_C map (blue mesh, contoured at 1.5 σ). (*b*) Running a cross-refinement with an open TMH10 (teal) resulted in significant negative difference density in the F_O — F_C map (red mesh, contoured at -3.0 σ), and some positive density is seen for a closed TMH10 (green mesh, contoured at 3.0 σ).

Using the bromine anomalous signal, the position of the bromine within the protein structure should be unambiguously determined, and the BVH ligand has previously been used for this purpose with Mhp1 (Simmons *et al*, 2014). Several diffraction data sets were collected of the mutant Gln42Asn in complex with BVH, but no anomalous signal was ever detected during the data processing, and as a result, the generated anomalous difference maps

did not contain any significant peaks in the ligand binding site. Performing fluorescence scans on the crystals revealed that there was bromine present in the crystals to varying degrees, but in some of the structures of the mutant in complex with BVH there was no indication of the ligand being bound, due to the absence of electron density both in the $2F_O — F_C$ and $F_O — F_C$ maps. A possible reason for not seeing the ligand in the binding site is that the ligand was not sufficiently soluble during co-crystallisation. If it had bound, the failure to detect anomalous signal could also have either been due to X-ray induced cleavage of the bromine, or the fact that the aryl group of the ligands is simply too mobile within the hydrophobic pocket, as indicated by the absent electron density for the aryl groups of L-BH and L-IMH in the respective crystal structures of the Mhp1 mutant Gln42Asn under high-sodium conditions.

5.2.6 Discussion

Crystal structures of the Mhp1 mutant Gln42Asn in complex with the hydantoin ligands L-BH and L-IMH have been determined in order to investigate the apparent loss of sodium-dependent ligand binding observed in the tryptophan fluorescence quenching assays. Two different structures of the mutant in complex with L-BH were determined, coming from different crystallisation conditions; high- and low-sodium conditions.

The crystal structures show that the mutant Gln42Asn adopts an occluded conformation in the crystal, similar to the previously published structures of the wild-type protein in complex with ligands (Jackson, 2012; Weyand *et al*, 2008; Simmons *et al*, 2014). There were no clear indications of significant conformational changes in the crystal structures of this mutant with the ligands compared to the wild-type protein, apart from TMH10 potentially adopting multiple conformations under low-sodium conditions. Nor were any significant local changes observed at the site of the mutation. However, all these crystal structures were only resolved to a resolution of 3.74 - 4.00 Å. At this low resolution, it is likely that any subtle conformational changes due to a single point mutation will not be detectable.

Tryptophan fluorescence quenching assays of the mutant Gln42Asn and L-BH suggested that the sodium-dependent ligand binding has been abolished, but could be restored upon the addition of excess sodium. In the structures of the Gln42Asn mutant in complex with L-BH and L-IMH, there was positive difference electron density in the $F_O — F_C$ map at the ligand

binding sites, indicating that the ligands were bound. These crystal structures came from crystals that had been produced under high-sodium conditions (in the presence of 100 mM sodium chloride and 100 mM sodium phosphate buffer), consistent with the presence of bound ligands.

The shortening of Gln42 to Asn42 may have resulted in the aryl groups of the ligands gaining flexibility, as a result of opening up of the binding pocket and weakened interactions between the residue and the ligand. Electron density was absent for the aryl substituents of L-BH and L-IMH under high-sodium conditions, which may have resulted from the fact that the charged Gln42 normally orients the aryl group by electrostatic interactions, while in the presence of the shorter Asn42 this interaction is less pronounced. This fits with the previous hypothesis that Gln42 could be playing a role in shaping the binding pocket for the ligands (Simmons *et al*, 2014). In an attempt to localise the orientation of the aryl groups in the Mhp1 mutant Gln42Asn, diffraction data collection was attempted on crystals of the mutant in complex with BVH. However, no anomalous signal was detected, which also supports the idea that the aryl groups might be flexible in the Mhp1 mutant Gln42Asn. Looking at surface representations of the binding sites in the Gln42Asn mutant structure before and after mutating the residue, it is apparent that the hydrophobic pocket for the aromatic moiety becomes slightly expanded (Figure 5.8), provided that the Asn42 does adopt the modelled conformation which preserves the hydrogen-bond to Gln121 (5.2.2). However, from the surface representations it is not apparent that the ligands could adopt sufficiently different conformations in different protein molecules, to result in a range of displacement that would be great enough to cancel the electron density for the entire aryl groups.

On the contrary, in the structure of the Gln42Asn mutant in complex with L-BH under low-sodium conditions, the entire L-BH is resolved in the electron density. This could be due to the larger potassium ion in the sodium binding site compensating for the extra space in the binding site as a result of the mutation by shifting Asn42 closer to the ligand, hence helping to stabilising and resolving the entire ligand clearly. Alternatively, since the structure under-low sodium conditions was produced from crystals with a contracted lattice, the quality of the ligand density could also be attributed to the better quality of the data.

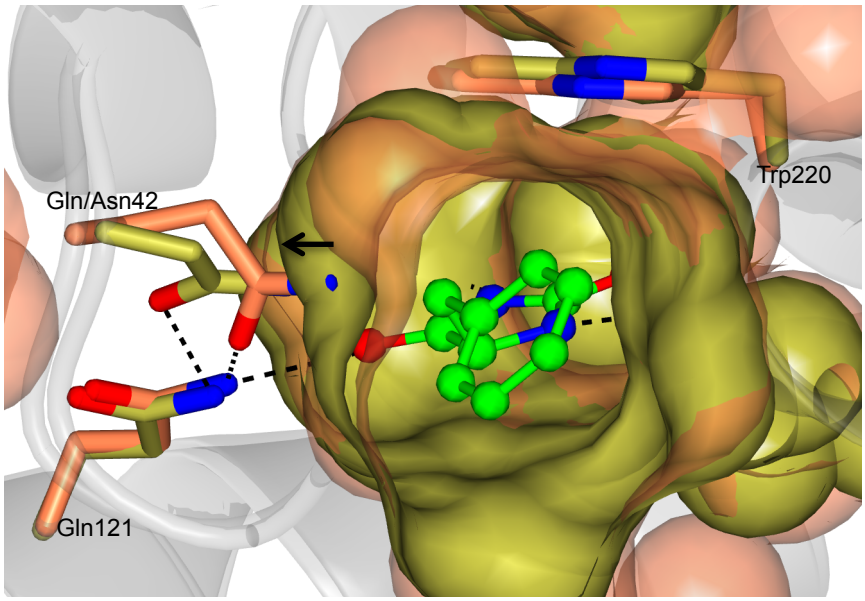


Figure 5.8 Surface representations of the binding sites in the structures of Gln42Asn in complex with L-BH.

The figure shows the accessible surface generated by the residues in the ligand binding site. The protein and corresponding surface for the structure of Gln42Asn in complex with L-BH before replacing Gln42 with Asn in the model is depicted in coral, while the refined structure with the mutation Asn42 in place is depicted in yellow. The black arrow shows the direction of the expansion of the ligand-binding site.

The tryptophan fluorescence quenching assay suggests, however, that L-BH does not bind to the Gln42Asn mutant in the presence of potassium (Figure 5.5). The crystal structure of the mutant in complex with L-BH under low sodium and high potassium conditions adopts an occluded state and shows that both L-BH and a potassium ion are bound. A possible explanation for this discrepancy between the functional assay and the crystal structure is that the conformational equilibrium is affected by whether the protein is in solution or in the crystalline form. Previous double electron-electron resonance (DEER) spectroscopy studies studies have shown that Mhp1 samples the inward- and outward-open conformations in solution (Kazmier *et al*, 2014). If the life-time of the inward-open conformation is on a faster time-scale than the crystal nucleation and/or growth rate, this conformation is less likely to take part in forming the protein crystal. In addition, the growing crystal lattice could act as a conformational selection filter, stabilising the outward-open or occluded conformations to which the ligand can bind.

Another possibility, is that there could be a residual amount of sodium present in the crystallisation experiment due to contamination from, for

example, plastic vials and the equipment used. This residual level has been determined to be 0.33 mM (Ivanova et al. publication in preparation), and it is possible that even the residual amount of 0.33 mM sodium is enough to drive the protein conformational equilibrium slightly towards the crystallisable occluded conformation, in combination with the growing crystal lattice acting as conformational selection filter.

5.3 X-ray crystal structure of a mutant-ligand complex of the Mhp1 mutant Ser312Ala

5.3.1 Introduction

In the original structure of Mhp1 in the outward-open form (Weyand *et al*, 2008), a putative sodium-binding site was identified, situated between TM1 and TM8. Sodium-binding sites have been identified in equivalent positions in LeuT and vSGLT, indicating that this site is conserved. Tryptophan fluorescence quenching assays conformed that the ligand-binding is sodium dependent in Mhp1, and hence a sodium ion was modelled into this site. In the crystal structure, the sodium is coordinated by the carbonyl oxygen atoms of Ala38, Ile41 (TM1) and Ala309 (TM8), as well as the hydroxyl oxygens of the side chains of Ser312 and Thr313 (TM8) (Jackson, 2012; Weyand *et al*, 2008; Shimamura *et al*, 2010) (Figure 5.9). The sodium-binding site in Mhp1 is situated approximately 7-9 Å from the hydantoin binding site, towards which the adjacent residue Gln42 points. Since Ser312 participates directly in coordinating the sodium, it is a suitable residue to mutate to probe the effect on the conformation of Mhp1 of a potentially disrupted sodium-binding site. As discussed in 5.2.4, the sodium-binding site in the inward-open form of Mhp1 is disrupted, which was also confirmed by computer simulations (Shimamura *et al*, 2010). Therefore, disruption of the sodium binding site by the Ser312Ala mutation could potentially stabilise an inward-open state. As for the Gln42Asn mutation, the Ser312Ala mutation essentially abolished the sodium-dependent binding of L-BH in Mhp1 at 15 mM added sodium (Jackson, 2012).

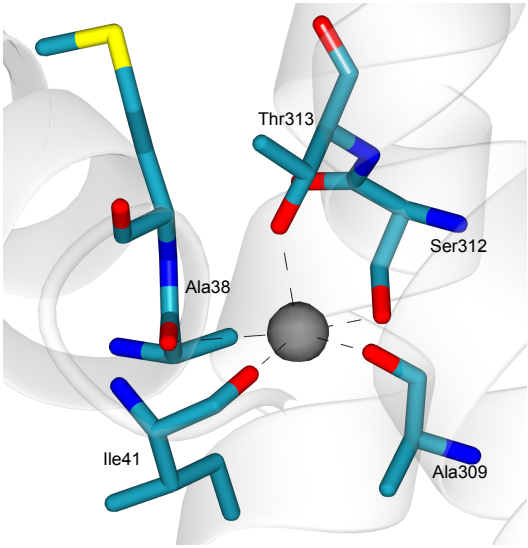


Figure 5.9 Illustration of the sodium-binding site in wild-type Mhp1.

The figure is replicated from Figure 3.7*b*, and shows the square-pyramidal coordination of sodium (grey sphere) in the Mhp1 protein. The view of the sodium binding site with respect to the ligand binding site can be seen in Figure 5.1. The residues coordinating sodium (Ala38, Ile41, Ala309, Ser312 and Thr313) are displayed in cyan with the coordinating bonds between as black dashed lines. The non-carbon atoms are coloured red (oxygen), blue (nitrogen) of yellow (sulphur). The ribbon diagram of the whole protein is displayed in semi-transparent grey.

This could be attributed to the fact that the loss of the coordinating bond from the side chain hydroxyl oxygen of Ser312 disrupts the sodium binding site. Here, it is shown that at a higher concentration of added sodium (140 mM), some of the ligand-binding capacity is restored (Figure 5.10).

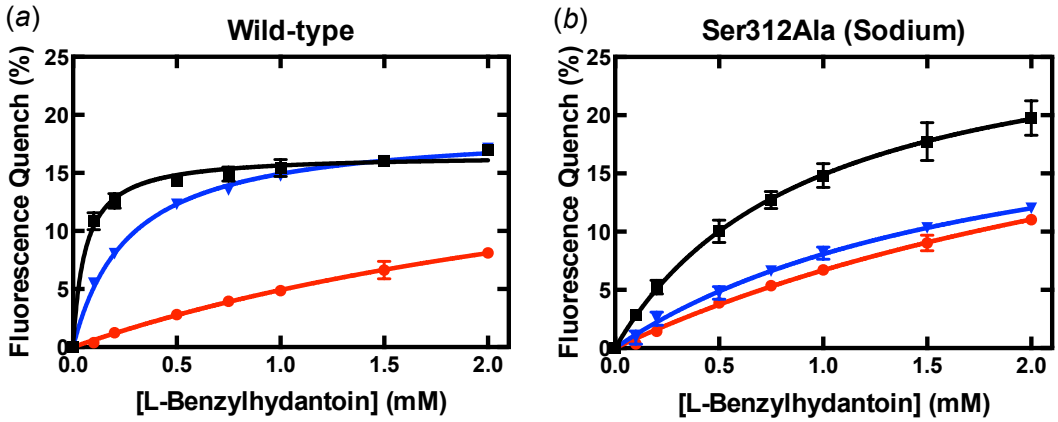


Figure 5.10 The binding of L-BH to wild-type Mhp1 and the Mhp1 mutant Ser312Ala, solubilised in NM.

The figure shows the fluorescence quench during titrations of L-BH (0-2 mM) with wild-type Mhp1 (*a*) and the Mhp1 mutant Ser312Asn (*b*). The titrations were performed in the presence of 0 mM (red), 15 mM (blue) and 140 mM (black) added sodium chloride. The buffer components are the same as indicated in Figure 5.2. The error bars display the standard error of the mean from three replicates of each titration.

5.3.2 Structure of the Mhp1 mutant Ser312Ala in complex with L-NMH

The Ser312Ala mutant was subjected to crystallisation trials in the absence and presence of all ligands previously used (Table 2.1) under high-sodium conditions, but crystals were obtained only of the apo-form and in complex with the ligands BVH and L-NMH. Diffraction data were, however, collected from only one crystal of the mutant in complex with L-NMH. The resolution of the diffraction data obtained from the crystal was 3.58 Å (Table 5.4), and the structure was solved in space group $P2_12_12_1$ with a unit cell of *a* = 88.7, *b* = 103.3, *c* = 108.1 Å (Figure 5.11*a*). The slight contraction of the *a*- and *b*-axes of the unit cell may be a result of that at the time of harvesting, the tray was approximately three months old[3]. Using the structure 4D1D as the molecular replacement search model resulted in the lowest initial *R*-factors.

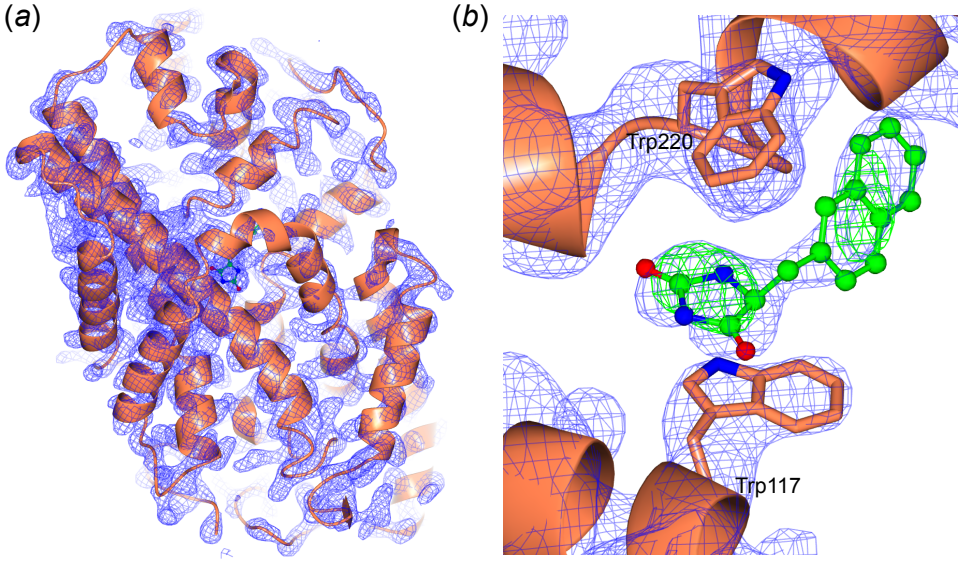


Figure 5.11 The ligand binding site in the structure of the Ser312Ala mutant in complex with L-NMH.

The $2F_O$ — F_C (blue mesh, contoured at 1.5 σ) map after refinement (*a*), where the L-NMH ligand was nicely resolved (shown in green as a ball and stick model, placed in the same orientation as in 4D1D) at the initial stages of refinement, with the positive difference density peaks in the F_O — F_C (green mesh, contoured at 3.0 σ) confirming the position of the ligand (*b*).

Early in the refinement process, there was clear density in both of the $2F_O$ — F_C and the F_O — F_C maps in the ligand binding site, into which the L-NMH ligand fitted well (Figure 5.11*b*). As for 4D1D, TMH10 was resolved in the

[3] Tray set-up date: 19th November 2014, crystal harvesting: 17th February 2015, data collection: 22nd February 2015

electron density map in an outward-open conformation, indicating that the L-NMH bound to the Ser312Ala mutant in the same way as it inhibits wild-type Mhp1. Due to the bulky nature of the naphthyl substituent of the ligand, it prevents occlusion of TMH10 (Simmons *et al*, 2014).

Table 5.4 Data collection and refinement statistics for the structure of the Mhp1 mutant Ser312Ala in complex with L-NMH.

	Ser312Ala-(L-NMH)
Wavelength (Å)	0.9174
Temperature (K)	100
Space group	$P2_12_12_1$
Molecules per assym. unit	1
Unit cell parameters (Å, °)	a = 88.7, b = 103.3, c = 108.1, $\alpha = \beta = \gamma = 90$
Resolution (Å)	103.33 - 3.58 (3.83 -3.58)
No. of measured reflections	73441 (12761)
No. of unique reflections	12189 (2159)
Multiplicity	6.0 (5.9)
Completeness (%)	99.8 (99.9)
$<I/\sigma(I)>$	11.5 (1.1)
$CC_{1/2}$	0.997 (0.699)
Wilson B factor (Å^2)	178.3
Anisotropy ΔB (Å^2)	100.6
R_{merge} (%)	0.100 (1.689)
R_{meas} (%)	0.116 (2.030)
$R_{p.i.m.}$ (%)	0.119 (2.045)
Matthews coefficient (Å^3 Da^{-1})	4.51
Solvent content (%)	72.77
R_{work} (%)	25.29 (40.87)
R_{free} (%)	31.09 (44.52)
R.m.s.d.	
Bonds (Å)	0.0097
Angles (°)	1.3556
Ramachandran outliers (%)	0.88
Average B factor (Å^2)	163.1

Values in parentheses are for the highest resolution shell.

5.3.3 Discussion

A structure of the Ser312Ala mutant was solved in complex with the Mhp1 inhibitor L-NMH. The crystal structure revealed that the ligand was bound in the binding site, and that the mutant adopted a similar conformation as the wild-type protein in complex with the inhibitor, with the occluding helix 10 in an outward-conformation (Simmons *et al*, 2014). The crystal structure indicates that the L-NMH inhibits the Ser312Ala mutant, as it inhibits the wild-type protein, by preventing the occlusion of TMH10.

For the Ser312Ala mutant, as well as for the Gln42Asn mutant, it has been shown that the sodium-dependent ligand binding is abolished at lower sodium concentrations. For both mutants, the ligand binding capacity was restored, to varying degrees, by adding almost 10-fold more sodium. The position of the residues put them in contact with the sodium-binding site, with the mutations possibly perturbing and causing remodelling of the sodium binding site.

As for some of the structures of the Mhp1 mutant Gln42Asn, the structure of the Ser312Ala mutant was produced under high-sodium conditions, which was probably enough to turn the protein over into the occluded state and cause the ligand to bind.

5.4 X-ray crystal structures of mutant-ligand complexes of the Mhp1 mutant Lys110Leu

5.4.1 Introduction

Lys110 is situated on the interface between the so-called bundle- and hash-motifs (Shimamura *et al*, 2010), as illustrated in Figure 5.13. From comparing the crystal structures of the outward- and inward-open states of wild-type Mhp1, the hash-motif seems to undergo an approximately 30° rotation with respect to the bundle-motif, opening the substrate binding cavity to the intracellular side of the protein during the conformational transition to the inward-open state. Lys110 is situated on TMH3, pointing towards TMH6 of the bundle-motif and the mutation to Leu could potentially be destabilising the interactions between the bundle- and hash-motifs, promoting a conformational change towards the inward-open state. This hypothesis is supported by the tryptophan fluorescence quenching assay on the Lys110Leu mutant with L-BH, which shows that L-BH binds with very low affinity to the Lys110Leu mutant (Figure 5.13).

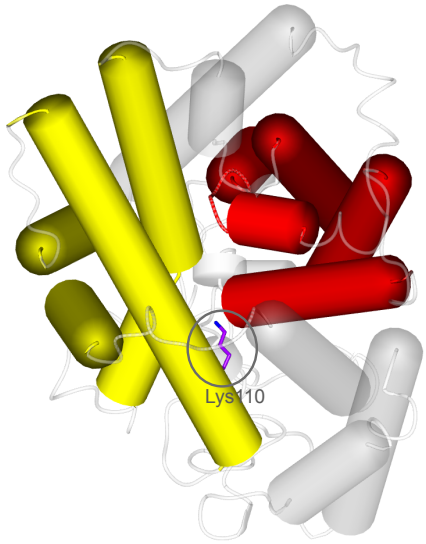


Figure 5.12 The bundle- and hash-motifs in the Mhp1 protein, displaying the locations of Lys110.

The figure is replicated from Figure 3.6. The Mhp1 protein is illustrated with the bundle- (red) and hash-motifs (yellow) highlighted as cylinders. The helices comprising the bundle motif (TMs 3, 4, 8 and 9) are displayed in red and the hash motif (TMs 1, 2, 6 and 7) is displayed in yellow. The remaining helices and loops are displayed in semi-transparent grey for clarity. The location of Lys110 is displayed in purple.

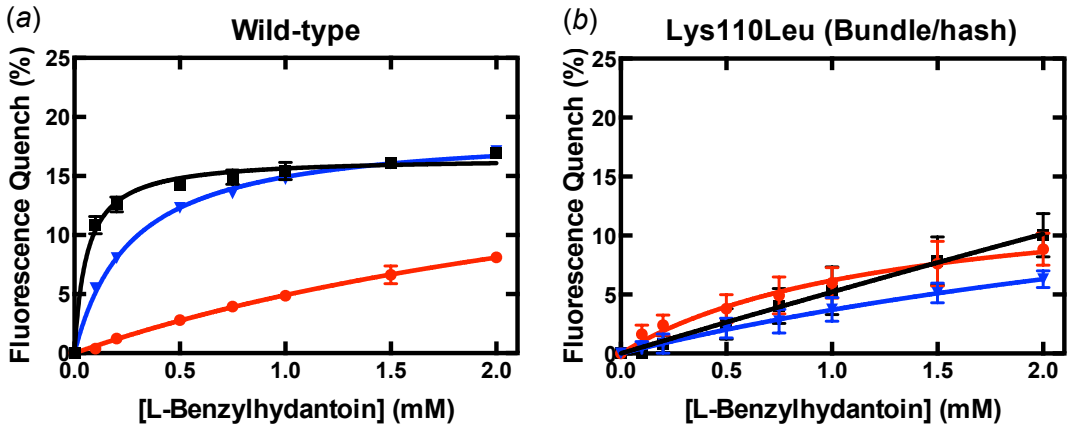


Figure 5.13 The binding of L-BH to wild-type Mhp1 and the Mhp1 mutant Lys110Leu, solubilised in NM.

The figure shows the fluorescence quench during titrations of L-BH (0-2 mM) with wild-type Mhp1 (*a*) and the Mhp1 mutant Lys110Leu (*b*). The titrations were performed in the presence of 0 mM (red), 15 mM (blue) and 140 mM (black). The buffer components are the same as indicated in Figure 5.2. The titration curve at 140 mM could not be fitted. The error bars display the standard error of the mean from three replicates of each titration.

The affinity of the mutant for L-BH appears to follow the reverse trend as seen for the wild-type protein, with the affinity being further lowered in the presence of 15mM sodium chloride. However, within the experimental errors, the affinity is not significantly affected by the concentration of sodium chloride.

5.4.2 Structure of Mhp1 mutant Lys110Leu in complex with L-BH

The structure of the mutant Lys110Leu in complex with L-BH was initially described in section 4.6.2 and was obtained from a crystal that was harvested by contact-less photoablation. Data were also collected of the mutant Lys110Leu in complex with L-BH from a crystal that had been subjected to controlled dehydration, and the resulting structure from these data will be presented here. The structure was solved in space group $P2_12_12_1$ to a resolution of 3.85 Å (Table 5.5).

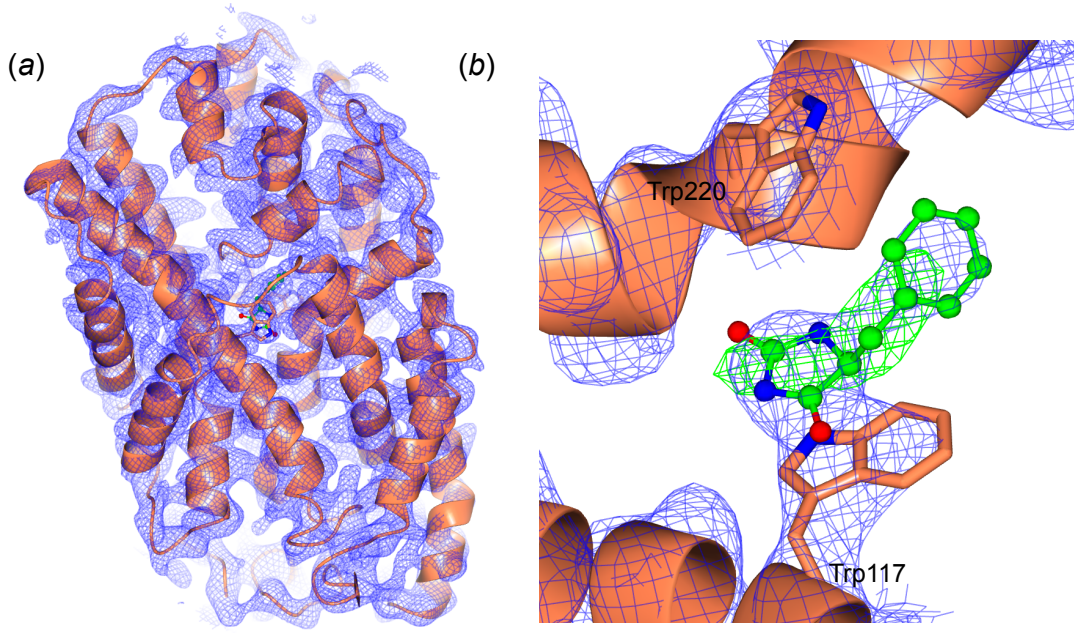


Figure 5.14 The ligand binding site in the structure of the Lys110Leu mutant in complex with L-BH.

The $2F_O$ — F_C (blue mesh, contoured at 1.5 σ) map after initial refinement, resolving the L-BH ligand (shown in green as a ball and stick model, placed in the same orientation as in 4D1B) and the positive difference density peaks in the F_O — F_C (green mesh, contoured at 3.0 σ) confirming the position of the ligand.

Controlled dehydration was carried out and resulted in a contracted lattice with unit cell parameters of *a* = 88.1 Å, *b* = 103.1 Å, *c* = 105.7 Å. Despite the

tryptophan fluorescence quenching data indicating that the affinity of the mutant Lys110Leu for L-BH is very low, electron density was clearly resolved for L-BH in the ligand-binding site in the structure of Lys110Leu (Figure 5.14). 4D1A was used as the molecular replacement search model, and there were no indications of any shift in the bundle- and hash-motifs.

5.4.3 Structure of Mhp1 mutant Lys110Leu in complex with L-IMH

Three data sets were collected from crystals of Lys110Leu in complex with L-IMH on the same occasion, coming from crystals that had been cryo-cooled both by fast plunging into liquid nitrogen, as well as transfer into a cryo-stream. Data processing of the three data sets revealed that all crystals exhibited a relaxed lattice (a = 89-90 Å, b = 105-107 Å, c = 107 Å) in space group $P2_12_12_1$. The best data statistics as well as lowest R-values after initial refinement of the molecular replacement solution using 4D1A as the molecular replacement search model came from the crystal that was cooled by fast plunging into liquid nitrogen (Table 5.5) which diffracted to 4 Å resolution. Early in refinement, positive electron density features in the F_O — F_C map were seen in the binding site, covering both the hydantoin and indolyl moities of L-IMH when placed into the ligand binding site (Figure 5.15).

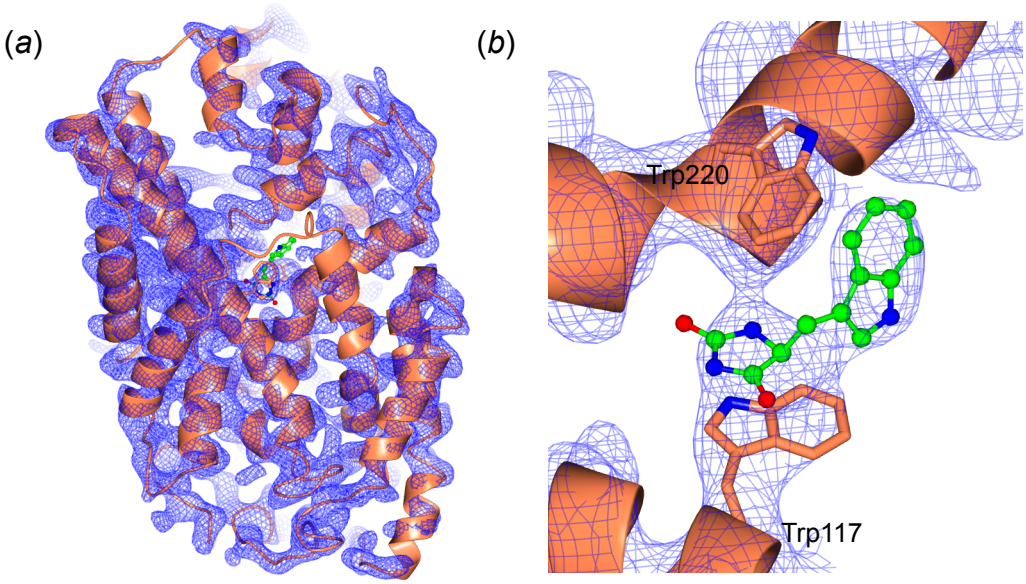


Figure 5.15 The ligand binding site in the structure of the Mhp1 mutant mutant in complex with L-IMH.

The final $2F_O$ — F_C (blue mesh, contoured at 1.5 σ) map, with electron density covering the L-IMH ligand (shown in green as a ball and stick model, placed in the same orientation as in 4D1A).

At this resolution, there were no distinct features in the F_O — F_C map indicating major structural changes around the site of the mutation.

Table 5.5 Data collection and refinement statistics for the structure of the Mhp1 mutant Lys110Leu in complex with L-BH and L-IMH.

	Lys110Leu-(L-BH)	Lys110Leu-(L-IMH)
Wavelength (Å)	0.9795	0.9174
Temperature (K)	100	100
Space group	$P2_12_12_1$	$P2_12_12_1$
Molecules per assym. unit	1	1
Unit cell parameters (Å, °)	a = 88.1, b = 103.0, c = 105.74, $\alpha = \beta = \gamma = 90$	a = 89.4, b = 105.14, c = 106.8, $\alpha = \beta = \gamma = 90$
Resolution (Å)	41.00 - 3.88 (4.30-3.85)	74.91 - 4.00 (4.47-4.00)
No. of measured reflections	31021 (5618)	56220 (15868)
No. of unique reflections	8270 (1526)	8435 (2300)
Multiplicity	3.8 (3.7)	6.7 (6.9)
Completeness (%)	90.1 (68.9)	94.5 (92.3)
$<I/\sigma(I)>$	9.8 (1.1)	7.4 (1.1)
$CC_{1/2}$	0.999 (0.805)	0.998 (0.595)
Wilson B factor (Å^2)	260.5	239.4
Anisotropy ΔB (Å^2)	140.79	86.34
R_{merge} (%)	0.058 (1.369)	0.057 (1.251)
R_{meas} (%)	0.075 (1.766)	0.072 (1.586)
$R_{p.i.m.}$ (%)	0.047 (1.302)	0.044 (0.959)
Matthews coefficient (Å^3 Da^{-1})	4.40	4.73
Solvent content (%)	72.04	74.02
R_{work} (%)	28.49 (39.40)	27.43 (38.55)
R_{free} (%)	32.27 (33.86)	30.18 (40.59)
R.m.s.d.		
Bonds (Å)	0.0100	0.0105
Angles (°)	1.4471	1.3854
Ramachandran outliers (%)	1.10	0.66
Average B factor (Å^2)	235.8	129.4

Values in parentheses are for the highest resolution shell.

However, there was no electron density visible for TMH10, indicating that this segment of the protein adopts several conformations in the structure, or remains flexible in the crystal lattice. This could be a result of the slight destabilisation between the bundle- and hash-motifs, preventing TMH10 closing around the ligand binding site.

5.4.4 Discussion

Crystal structures of the Lys110Leu mutant in complex with the ligands L-BH and L-IMH have been solved. The tryptophan fluorescence quenching assay shows that L-BH binds with very low affinity to the Lys110Leu mutant, and the affinity is further lowered with increasing sodium concentration. This is potentially due to the fact that the mutation destabilises the interactions between the bundle- and hash-motifs, causing a conformational turnover into the inward-open form. The crystal structures, however, indicate that the Lys110Leu mutant adopts an occluded conformation, with the ligands L-BH and L-IMH bound.

Despite the fact that the tryptophan fluorescence quenching assay indicated that the Lys110Leu mutant could be adopting a low-affinity conformation, such as for example an inward-open form, the crystal structures presented here indicate that this form may be a highly transient and/or unstable form to crystallise under the conditions used here. The conformations of the Mhp1 mutants studied here appear to be highly dependent on the crystal lattice environment. One reason for this could be that the time-scales of conformational dynamics vs. crystal formation and growth promote the crystallisation of the outward-open form, allowing ligands to bind and Mhp1 to adopt the occluded form. Another reason is the absence of stabilising ligands (as seen for the previously published inward-open structure of seleno-L-methionine derivatised Mhp1).

5.5 General discussion and conclusions

In this chapter, several new structures of Mhp1 mutant-ligand complexes are presented; including structures of the three Mhp1 mutants Gln42Asn, Ser312Ala and Lys110Leu in complex with L-BH, L-IMH or L-NMH under various crystallisation conditions, where primarily the sodium concentration was varied. The various mutant and ligand combinations for which structures were solved are summarised in Table 5.6.

Table 5.6 Summary of the mutant-ligand structures described in Chapter 5.

Mutant	Ligand	Crystallisation condition	Res.	Cooling method
Gln42Asn	L-BH	High-sodium	3.74 Å	Cooled by plunging into liquid nitrogen
Gln42Asn	L-IMH	High-sodium	4.00 Å	Cooled by plunging into liquid nitrogen
Gln42Asn	L-BH	Low-sodium, high-potassium	3.80 Å	Cooled by plunging into liquid nitrogen, crystal from old tray
Ser312Ala	L-NMH	High-sodium	3.58 Å	Cooled by plunging into liquid nitrogen, crystal from old tray
Lys110Leu	L-BH	High-sodium	3.88 Å	Subjected to controlled dehydration
Lys110Leu	L-IMH	High-sodium	3.91 Å	Cooled by plunging into liquid nitrogen

Overlaying all the obtained structures shows that there are no significant conformational changes present between the various mutant-ligand complexes (Figure 5.16*a*), apart from the region of TMH10 that closes up upon the ligand binding site (Figure 5.16*b*). In each of the ligand binding sites of the mutants, the ligand is found to be bound in its binding site (Figure 5.16*c*).

The three mutations probe various sites in the protein. While the Gln42Asn and Ser312Ala mutations probe the ligand- and sodium-binding sites, respectively, the Lys110Leu mutation probes the interface between the bundle- and hash-motifs. The tryptophan fluorescence quenching assay reveals that the binding of L-BH is reduced in all three mutants to various degrees. The degree of the reduction of L-BH binding also depends on the amount of sodium present. The reduction in L-BH binding could be due to the protein adopting an alternative conformational state that has very low affinity for the ligand, and this was explored using X-ray crystallography by subjecting the mutants to crystallisation trials in the absence and presence of ligands.

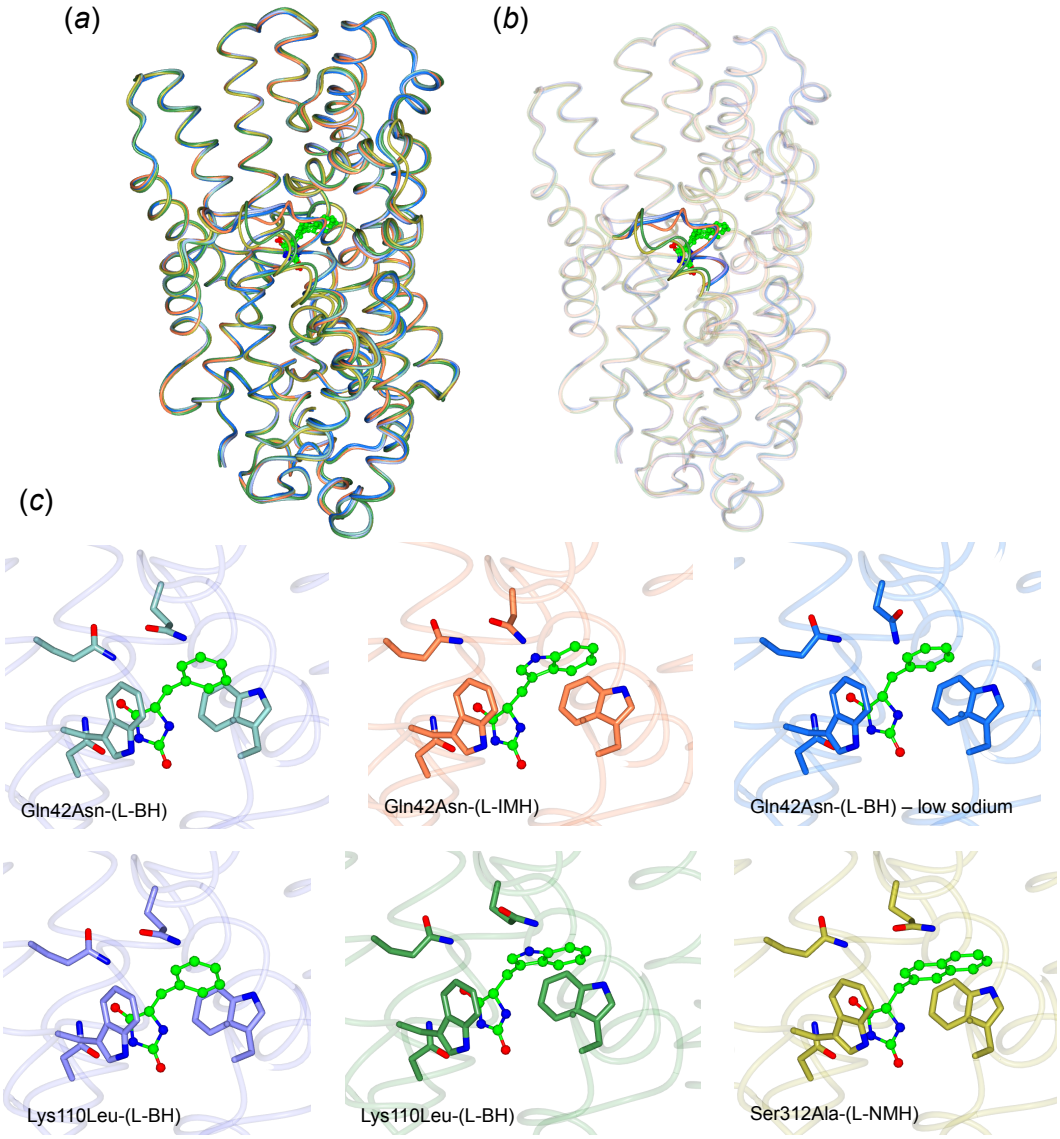


Figure 5.16 Comparison of the Mhp1 mutant-ligand complex structures.

All the structures of the Mhp1 mutant-ligand complexes superimposed (*a*), displaying no significant conformational changes apart from in the section of TMH10 that closes upon the ligand binding site (highlighted in *b*). In panel (*c*) the ligand binding site of each respective structure is shown, where the ligands were found to bind in the binding site. The ligand binding sites in (*c*) are in the same orientation as in (*a*) and (*b*).

The resolution and quality of the data was comparable to previously published structures of wild-type Mhp1 in complex with hydantoin ligands (Weyand *et al*, 2008; Simmons *et al*, 2014). The resolution of the data was high enough to resolve the overall global conformation of the protein, and in all cases the mutants Gln42Asn, Ser312Ala and Lys110Leu crystallised in the occluded state, with ligands bound in their respective ligand binding

sites. Any subtle local changes at the sites of the mutations were not detected at this resolution.

In all the mutant-ligand complex structures presented here, there was evidence that the ligands had bound in their respective binding sites. In the cases of the Gln42Asn mutant in complex with L-BH and L-IMH (high-sodium conditions), as well as the Ser312Ala mutant in complex with L-NMH, it was most likely the high concentration of sodium present in the crystallisation conditions that stabilised the outward-open form of the mutants, allowing the ligands to bind. For the Gln42Asn mutant in complex with L-BH under low-sodium conditions, as well as the Lys110Leu in complex with L-BH, the tryptophan fluorescence quenching assays indicated that L-BH would not be expected to bind but yet ligands were clearly observed in their respective binding sites. In this case, it could be that conformational selection during crystal formation and growth stabilised the outward-open and the occluded states of the protein, promoting ligand binding.

No crystal structures were obtained in their apo-form, despite all being subjected to crystallisation trials in the absence of ligands. Some mutants did yield crystals (for example Gln42Asn, Ser312Ala, Asn318Gly and Gln121Asn) but they diffracted to a very low resolution or not at all. The presence of the ligands may have improved the crystallisability of the mutants by promoting the conformational change into the occluded form, which could be a more robust and crystallisable form of the protein. Alternatively, the ligand could have acted as a nucleating agent, aiding crystallisation.

In order to further explore alternative conformational states of the Mhp1 protein, the crystallisation conditions for the mutants that did not yield crystals, and hence were not taken forward for structural studies, could be optimised. The mutants (as well as the wild-type protein) could also be subjected to low-sodium conditions, which was shown to yield diffracting crystals for the structure of Gln42Asn in complex with L-BH (5.2.4) in order to potentially identify a mutant that does adopt an alternative conformation in the crystal. In order to address the problem of the potential low sodium contamination, 15-crown-5 ether could be added to the crystallisation mother liquor to ensure that no sodium is present for interaction with the protein (Lamb *et al*, 1980). This would also be an interesting modification to the crystallisation experiment to explore, as it has been demonstrated that

crown-ethers can also promote crystallisation by stabilising inter-molecular interactions (Lee *et al*, 2014).

The fluorescence quenching assays suggest that the mutants presented here could be adopting alternative low-affinity conformations, such as for example an inward-open form, however a ligand-bound outward-open or occluded form was obtained for each of the structures of the mutant-ligand complexes. This suggests that the crystal formation and growing crystal lattice may be affecting the conformational equilibrium. In the absence of potentially stabilising molecules as seen the previously published inward-open structure of seleno-Lmethionine derivatised Mhp1, alternative strategies and crystallisation conditions need to be sought to obtain a native form of the inward-open conformation or indeed any other alternative conformation.

Here, only conventional vapour diffusion crystallisation trials of detergent-solubilised protein were performed. One example of an alternative crystallisation technique for membrane proteins is the lipidic-cubic phase (LCP) crystallisation technique (Caffrey & Cherezov, 2009), that puts the protein in a lipid-rich environment and may thus affect the conformational equilibrium in such a way that it will allow the crystallisation of alternative crystal forms.

Another route would be to improve the overall resolution of the diffraction data, which would allow identification of finer local structural rearrangements. This could be achieved by identifying an alternative crystal form by extensive screening of conventional crystallisation screens, or for LCP conditions. If no alternative form is found, there are further modifications that could enhance the quality of the crystals obtained. For example, shortening the detergent from DDM to NM strongly promotes crystal formation. Further decreasing the length of the aliphatic chain of detergent by exchanging it to octyl-maltoside would expose more of the protein that could strengthen crystal lattice contacts and the smaller micelle would give rise to a tighter lattice. If exchanging the detergent to *n*-octyl-*β*-D-maltopyranoside (OM) causes the protein to aggregate due to an increased exposure of hydrophobic patches, OM:NM mixtures could potentially provide a milder way of decreasing the amount of detergent in the crystal. In addition, the protocols for controlled dehydration could be further optimised in order to improve the data quality.

Apart from improving the quality of the crystals and resulting data, other strategies could also be employed to stabilise alternative conformational

states. For example, LeuT has been stabilised in the inward-open form using a combination of mutations and conformation-specific antibodies (Krishnamurthy & Gouaux, 2012) and such a strategy could be employed for Mhp1. In order to address discrepancies between tryptophan fluorescence quenching assays of the protein solution and what we see in the crystal structures, single crystal X-ray fluorescence spectroscopy could be used to further investigate the effect of the crystal lattice on ligand binding.

Chapter 6: SAXS data collection on Mhp1-detergent complexes

6.1 Introduction

In Chapter 5, a number of novel Mhp1 mutant-ligand structures were presented. The functional characterisation of the mutants suggested that the mutants may be adopting alternative conformational states. In order to investigate this, the mutants were subjected to X-ray crystallographic studies. Any alternative conformation, in addition to the already published crystal structures of Mhp1 in the outward-open, occluded and inward-open states (Weyand *et al*, 2008; Shimamura *et al*, 2010), would give further insight into the transport mechanism beyond the knowledge we already have from the previous structures.

However, under the crystallisation conditions used and due to potential conformational selection imposed by the growing protein crystal, no new crystal form was obtained with a conformation not previously seen. Nor was the inward-open facing crystal form reproduced, as discussed in Chapter 4.2.3-4. The outcomes of these crystallisation trials highlight the fact that structural studies of proteins and their dynamics using X-ray crystallography can sometimes be limited by crystal packing effects. In order to study conformations that can not be accommodated by a crystal lattice, solution-based structural techniques are necessary.

An alternative structural technique to study protein structure in solution is small-angle X-ray scattering (SAXS). The main challenge associated with using SAXS to study membrane proteins solubilised in detergent, is that the resulting mixture of the protein-detergent-complex (PDC) and empty micelles in the sample (Figure 1.9) is known to make buffer subtraction difficult to achieve. The central challenge in the study of detergent-solubilised membrane proteins is, therefore, to collect data that will allow accurate buffer subtraction by accounting for the presence of empty micelles, and thus allow deconvolution of the PDC and micelle signals.

In Chapter 3, the preparation of Mhp1 solubilised in various detergents is described. In this Chapter, SAXS data collection from the various Mhp1 PDCs is described.

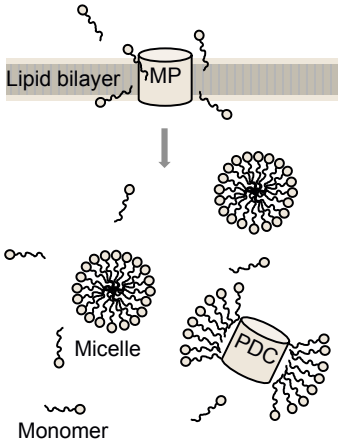


Figure 6.1 Solubilisation of membrane proteins using detergents.

The figure is replicated from Figure 1.9. The membrane protein (MP) can be extracted from the lipid bilayer by mixing with detergents. Above the CMC of the detergent, protein-detergent complexes (PDCs) will form, as well as protein-free detergent micelles. With the resulting mixture of PDCs and empty micelles still present in the buffer, it is difficult to achieve accurate buffer subtraction.

6.2 Small-angle X-ray scattering theory

6.2.1 Theory of X-ray scattering

In X-ray crystallography, all macromolecules within the crystal are evenly spaced on a lattice, thus producing a discrete diffraction pattern when irradiated with X-rays. In contrast, a dilute solution of randomly oriented macromolecules will instead produce a radially isotropic scattering pattern around the incident beam. In equation 1, the scattering vector ***q*** of length *q* and direction Ω describes the position of the scattered beam with respect to the incident beam (**Error! Reference source not found.**). The intensity (*I)* of the scattered beam is proportional to the scattering amplitude (*A)* as a function of scattering vector (***q***), averaged over all directions (Ω):

$$I(s) = < A^2(\boldsymbol{q}) >_{\Omega} \qquad (1)$$

A SAXS signal is therefore represented as a one-dimensional (1D) curve of the radially averaged intensity, *I*, vs. *q*. *q* is also called the magnitude of momentum transfer and is defined by equation 2, in which 2*θ* is the scattering angle, and λ is the wavelength of the incident radiation (**Error! Reference source not found.**):

$$q = \left(\frac{4\pi}{\lambda}\right) sin\theta \qquad (2)$$

The scattering of the macromolecule of interest is obtained by subtracting the scattering of the buffer alone from the scattering of the macromolecule

containing solution. Therefore, sample and reference buffer preparation is a crucial step for accurate data collection.

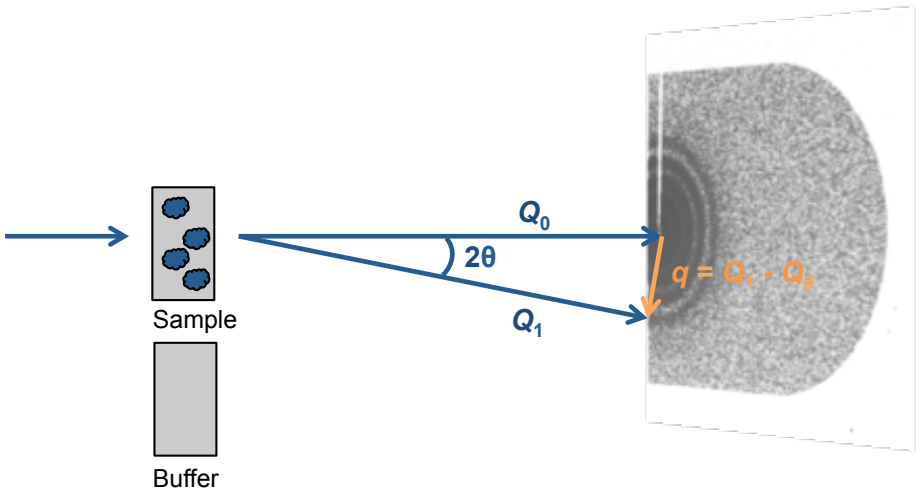


Figure 6.2 Simplified illustration of a SAXS experimental setup.

The scattered beam produces a radially averaged scattering pattern, where the length of the scattering vector ***q*** is defined as $q=(4\pi sin\theta)/\lambda$ with the scattering angle 2θ.

6.2.2 Obtainable information from SAXS data

From the raw SAXS data, a number of basic physical parameters can be obtained, including radius of gyration (R_g), the hydrated particle volume (V_p) and maximum intramolecular distance (D_{max}). It is recommended to collect and report data from several dilutions of the sample, and for the data to be considered reliable and suitable for modelling, these parameters should not vary with concentration (Jacques *et al*, 2012).

R_g can be estimated from the so-called Guinier region at very low scattering angles, where the data is assumed to follow the Guinier approximation (Guinier, 1955). This states that the scattering curve is linear at very low angles and the slope provides a value of R_g. In addition, the Guinier region serves as a sample quality control, as any inter-particle interactions will break the linearity in that region.

$I(0)$ is used to calculate the hydrated particle volume V_p, or the Porod volume (Porod, 1982). For globular proteins ranging from 14-500 kDa, the molecular weight (M_w) has been found to roughly relate to the Porod volume by $V_p/1.66$ (Petoukhov *et al*, 2012).

A more accurate estimation of both $I(0)$ and R_g can be obtained by indirect Fourier transformation of the scattering curve, taking the entire data range

into account. Fourier transformation of the scattering curve yields a distance distribution function $P(r)$, which describes the probability distribution of distances r between scattering electrons within the macromolecule. From the distance distribution function, $P(r)=0$ directly provides the maximum particle dimension D_{max} and the shape of the curve also gives an initial idea of the overall shape of the particle. Further shape information, in terms of the compactness of the particle, can be obtained from a Kratky plot ($I*q^2$ vs. q). A curve which decays to zero at higher angles indicates a folded particle, while the absence of a decay indicates that the particle, or parts of it, are unfolded.

Additionally, SAXS data can be used to generate *ab initio* models of biomolecules. The use of SAXS for low-resolution *ab initio* reconstructions began in the 1970s, initially by representing models as envelope functions based on spherical harmonics (Stuhrmann, 1970). This method was later superseded by reconstruction of dummy atom models (Chacón *et al*, 1998), allowing modelling of more complex structures (Volkov & Svergun, 2003).

6.3 Initial SAXS data collection on Mhp1 solubilised in DDM

6.3.1 Introduction

SAXS data collection on detergent solubilised membrane proteins is not routinely performed, due to the difficulties associated with the detergent micelles present in the sample buffer. Matching buffer blanks have previously been prepared by, for example, extensive dialysis (Zimmer *et al*, 2006).

SAXS data collection was carried out on wild-type Mhp1 solubilised in DDM, which was purified as described in Chapter 3. Matched sample buffer blanks were created by subjecting the protein samples to several cycles of dilution and re-concentration in a spin concentrator.

6.3.2 SAXS data collection of Mhp1 solubilised in DDM

Initial SAXS data were collected at the Stanford Synchrotron Radiation Lightsource (beamline BL4-2) on Mhp1 solubilised in DDM (Figure 6.3). DDM micelles have been estimated to have a molecular weight of ~70 kDa (VanAken *et al*, 1986), meaning that buffer exchange by dilution and re-concentration using a spin concentrator with a 100 kDa MWCO membrane

should eliminate excess micelles while retaining the Mhp1-DDM PDC. The concentration of DDM in the purified sample was approximately 5X the CMC of DDM, and samples for SAXS data collection were prepared by buffer exchange into a buffer with approximately 2X the CMC of DDM. The reduced DDM concentration was chosen in order to minimise the amount of free DDM monomers and micelles while still retaining a stable Mhp1-DDM PDC. The dilution series for the SAXS data measurements was then prepared using the dialysis buffer, ensuring that the same micelle concentration is present in all samples.

The scattering curves obtained had a bimodal appearance, which is characteristic of a mixture of detergent micelles and PDCs (Berthaud *et al*, 2012). The peak at around $q \approx 0.16$ Å^{-1} is from the DDM micelles present in the samples, and is also seen in data from pure DDM micelles (Lipfert *et al*, 2007). However, even after buffer subtraction, both the Kratky plot and distance distribution function show that the sample is not fully monodisperse. The sample consists of at least two components, which is especially evident from the Kratky plot with two distinct maxima (Figure 6.3). This may be due to inaccurate buffer subtraction due to varying detergent concentration. The obtained physical parameters R_g, D_{max} as well as the estimated M_w are summarised in Table 6.1.

The intra-molecular distances in Mhp1 are, approximated from the crystal structure of the outward-open form (2JLN) and measured using the Distance measure tool on Coot, ~75 Å (longest distance) and ~50 Å (parallell with the putative membrane normal) (see Figure 6.4). DDM micelles have previously been shown to adopt an oblate shape, with dimensions of 39.6-41.2 Å (vertical axis) and 68-71.6 Å (horizontal axis) (Lipfert *et al*, 2007) (see Table 6.2 and Figure 6.5). Assuming that a detergent micelle would be oriented in a similar way as the membrane with its horizontal ***b***-axis (Figure 6.5), parallel to the membrane normal (as illustrated in Figure 6.4), the diameter of the detergent micelle encapsulating the Mhp1 protein could be approximately ~118-121.6 Å (~50 + 68-71.6 Å), making this the largest estimated intra-molecular distance.

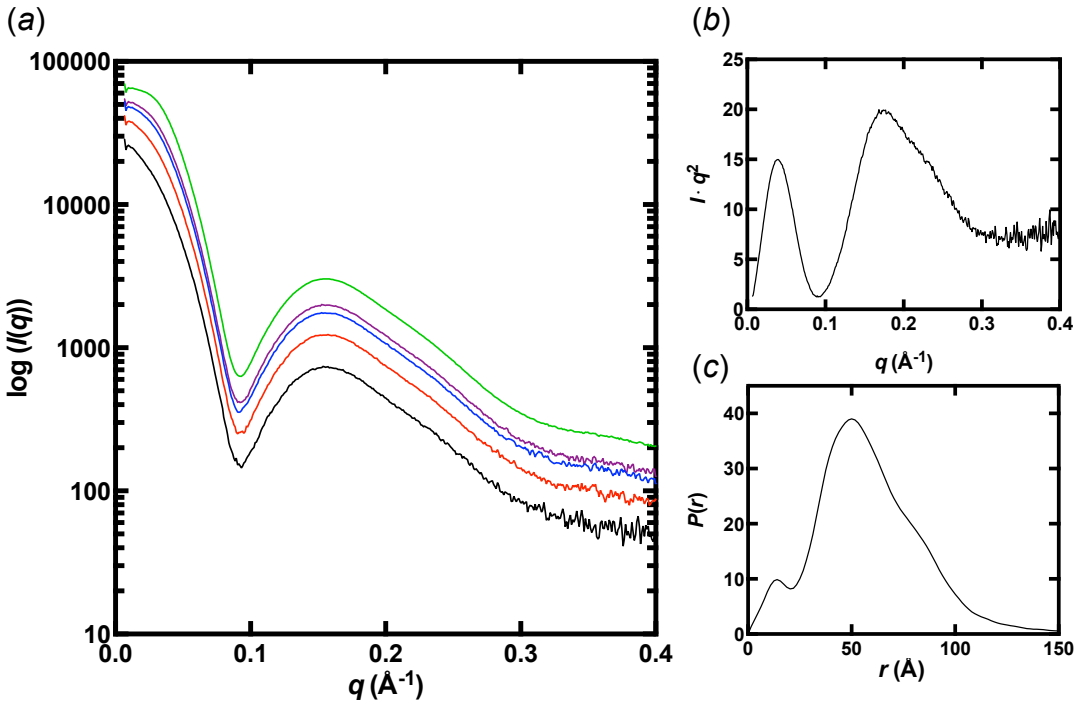


Figure 6.3 Initial SAXS data collection on Mhp1 solubilised in DDM.

(*a*) The SAXS curves (log *I*(*q*) vs. *q*) are shown for a concentration series of Mhp1 solubilised in DDM at 10 mg/ml (green), 5 mg/ml (purple), 3.75 mg/ml (blue), 2.5 mg/ml (red), 1.25 mg/ml (black). The peak at $q \approx 0.16$ Å^{-1} is due to the scattering from DDM micelles. (*b*) The Kratky plot (I^*q^2 vs. *q*) shows that there are two components present in the system, indicating that the buffer subtraction was not successful and did not fully account for the free micelles in the sample. (*c*) The distance distribution function (*P*(*r*) vs. *r*) also shows evidence of a non-monodisperse solution, potentially with multiple populations of molecules. The Kratky and *P*(*r*) plots were generated from the lowest concentration SAXS curve (black in *a*).

Table 6.1 Summary of the R_g, D_{max} and estimated M_w from the initial SAXS data of Mhp1 solubilised in DDM.

Concentration (mg/ml)	**Guinier R_g (Å)**	**D_{max} (Å)**	**M_w (≈V_p/1.66) (kDa)**
10	39.6	110.8	74.75
5	41.0	124.0	64.73
3.75	41.1	124.0	61.49
2.5	42.1	135.3	59.66
1.25	44.6	156.6	47.54

The suggested value for D_{max} of ~118-121.6 Å value falls into the range of the experimental values obtained for D_{max} from the SAXS (110-156 Å, Table

6.1). However, this range is relatively wide and could be accounted for by the fact that the dilution of the dialysed Mhp1 affected the stability of the Mhp1-DDM PDC and that the resulting buffer subtraction was not consistent throughout the concentration series.

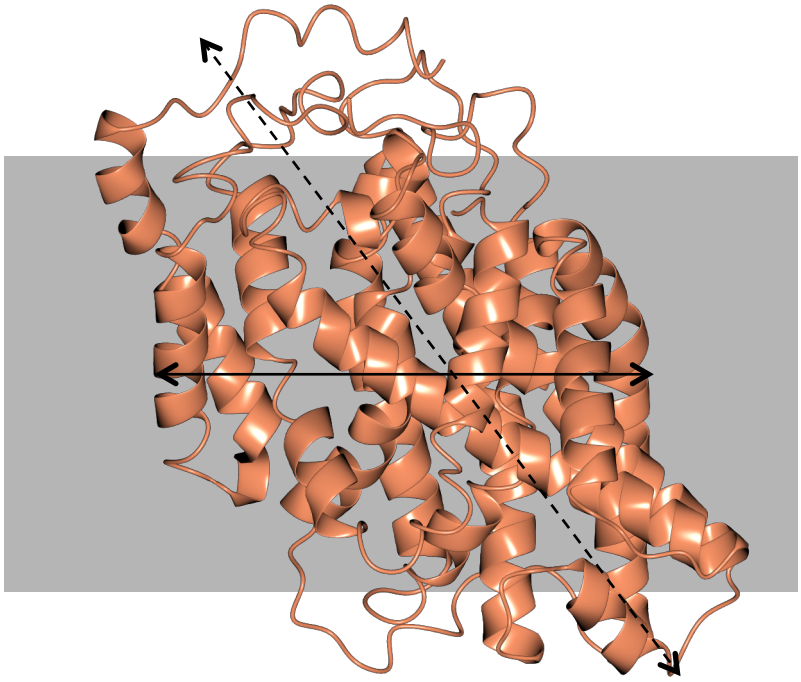


Figure 6.4 Illustration of the theoretical dimensions of the Mhp1 protein.

The Mhp1 protein (PDB 2JLN) is depicted in a cartoon representation (coral), positioned in its putative orientation in the cell membrane (grey rectangle). Using the Distance measurement tool in Coot, the longest distance is approximately 75 Å (dashed, black arrow) and the diameter in the direction that is parallel to the membrane is 50 Å (solid, black arrow).

The R_g of a DDM micelle has previously been determined as 33 Å (Lipfert *et al*, 2007). The R_g obtained for the Mhp1-DDM PDC was in the range of 39.6-44.6 Å, which also falls into a reasonable range, being slightly larger than DDM micelles alone, as expected. In conclusion, the dimensions obtained from the SAXS data collected from the Mhp1-DDM PDC, where the buffer has been exchanged by several rounds of dilution/re-concentration into a buffer containing ~2X the CMC of DDM, correspond to previously determined theoretical values of the dimensions of the Mhp1 protein and DDM micelles alone, and strongly indicate that Mhp1 does not exist as an oligomer in solution. Both the D_{max} and R_g decrease with increasing protein concentration (Table 6.1), indicating that there are inter-molecular repulsive forces in the sample. Even though this does not affect the monodispersity of

the PDCs in the sample, the concentration dependence suggests that the protein concentration should be reduced to obtain data that is suitable for modelling. However, the estimated M_W (V_p/1.66) ranges between 47.54-74.75 kDa and is much smaller than expected, since the molecular weight for Mhp1 is 54 kDa and one DDM micelle is 70 kDa (and thus the theoretical M_w would be at least 124 kDa, as more than one micelle would be required to surround one Mhp1 molecule). The Porod volume is dependent on the scattering range of the data used to calculate it (Petoukhov *et al*, 2012). The Porod volumes here were calculated from the entire data range, meaning that noise at higher angles might have affected the obtained values.

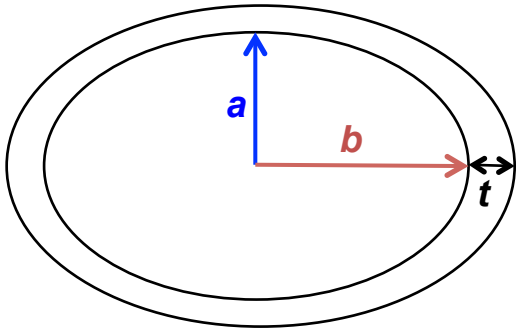


Figure 6.5 The model used for describing the previously reported sizes of detergent micelles.

The illustrated model shows the parameter ***a***, ***b*** and ***t*** used in Lipfert *et al.* (2007) and Oliver *et al.* (2013) for describing the reference values for detergent micelle dimensions.

Table 6.2 Previously reported dimensions for DDM and DM micelles.

The values ***a***, ***b*** and ***t*** are based on the model illustrated in Figure 6.5.

Detergent	***a* (Å)**	***b* (Å)**	***t* (Å)**	**R_g (Å)**
DDM	13.8-14.3	28.0-29.5	6.0-6.3	33
DM	12.0-12.5	23.0-24.0	6.0-6.3	26

6.3.3 Discussion

Overall, the presence of detergent micelles complicates SAXS studies of Mhp1 and the buffer exchange method used here to remove excess micelles is not efficient. Even though the obtained experimental values match estimated theoretical values, from the crystal structure of Mhp1 and previously obtained values for maltoside detergent micelles, the data are

inconsistent. The R_g and D_{max} values displayed a concentration dependence, indicating that there were intermolecular interactions between the PDCs in the sample and hence that the buffer subtraction was not correct.

In addition, the Kratky plot indicates that there were two species present in the sample, confirming that the buffer subtraction was not successful and there were micelles present that contributed to the overall signal. This complicates deconvolution of the respective signals as free detergent, micelles and protein-detergent-complexes will co-exist in a complex equilibrium.

Therefore, reducing the detergent concentration and using a conventional method for buffer exchange may not be a reliable method for eliminating the signal from detergent micelles. An alternative strategy is therefore required in order to accurately account for the micelles present in the PDC solution.

6.4 SEC-SAXS data collection on Mhp1 solubilised in various maltoside detergents

6.4.1 Introduction

In section 6.3, the difficulty of obtaining reliable SAXS data from detergent-solubilised Mhp1 is clearly demonstrated. It was shown that when employing a standard buffer exchange method, a concentration dependence of the physical parameters R_g and D_{max} as observed. In addition, the Kratky and $P(r)$ plots indicated that the samples still contained free detergent micelles.

In a recent study, size-exclusion coupled SAXS (SEC-SAXS) has been used to address the problem of inaccurate buffer subtraction in the study of DDM-solubilised aquaporin-0 (Aqp0) (Berthaud *et al*, 2012). The idea behind SEC-SAXS is that the PDC becomes diluted into the running buffer on the SEC column, with excess empty micelles becoming separated from the main PDC peak during the SEC run. A number of buffer scans at the start of the run can be averaged and used as a buffer blank for the PDC peak (Figure 6.6). These data enabled the modelling of a dummy-atom model of the detergent corona surrounding the Aqp0 protein.

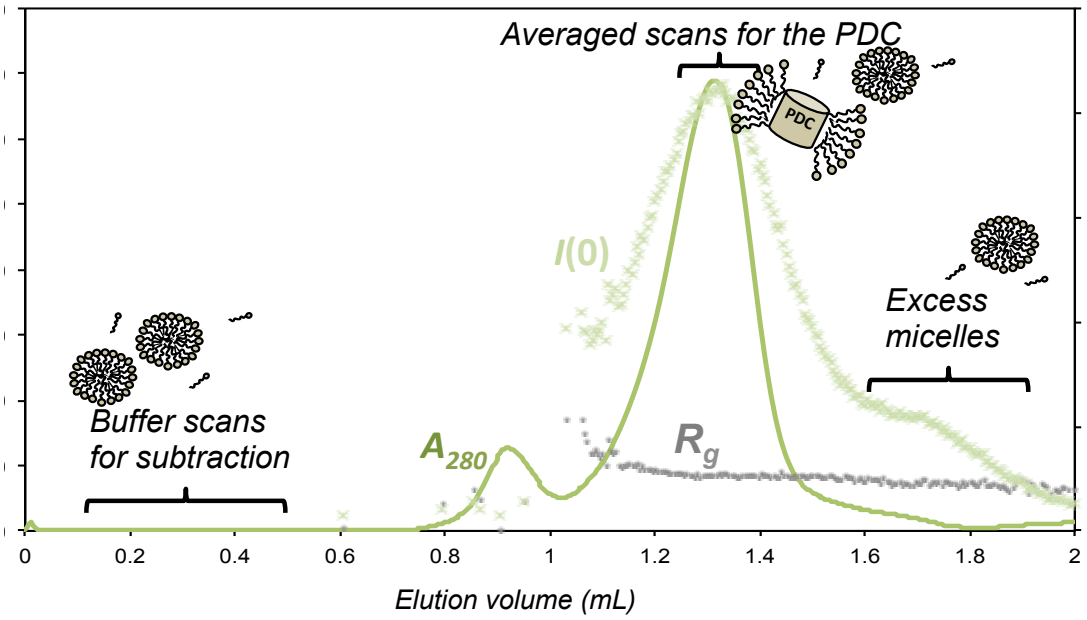


Figure 6.6 Illustration of the principle behind SEC-SAXS of membrane proteins.

Th figure illustrates a SEC-SAXS run of the Mhp1-DDM PDC, performed at the SSRL, beamline BL4-2. The absorbance at 280 nm (denoted A_{280}, green line) is combined with the $I(0)$ (green crosses) and R_g (grey dots). Using the SEC run, the PDC is diluted into the column running buffer and is separated from excess micelles. Accurate buffer subtraction is achieved by subtracting an average of a number of buffer scans and the beginning of the run from averaged scans from the main PDC peak.

Previously, purification and crystallisation trials have demonstrated that Mhp1 is stable in DDM and NM. The fact that Mhp1 does not crystallise when solubilised in DDM, but crystallises readily when the detergent has been exchanged to NM (see section 4.2.2) could be because the shorter detergent exposes the solvent-accessible parts of the protein that are able to form crystal contacts with varying degrees, and PDCs with different solvent-accessible regions might be able to give more structural information. Alternatively, the protein could be more structurally homogenous in one detergent than another. As suggested by the fluorescence quenching assays performed on the Mhp1 protein solubilised in maltoside detergents of varying detergents described in Chapter 3, the affinity the Mhp1 protein for L-BH appears to decrease when the protein is solubilised in a shorter detergent, which could be due to the fact that a smaller detergent micelle restricts the conformational flexibility of the protein and thus makes it a less efficient transporter. Therefore, SEC-SAXS experiments were performed on Mhp1 solubilised in one of four maltoside detergents of varying length (illustrated in

Figure 6.7). In addition, since the SAXS data are collected during the course of a SEC run, exploring different length detergents (giving rise to different size micelles) might enable the establishment of conditions where the separation between the PDC and excess micelles peak are optimally separated. Ensuring that there is no overlap and residual mixture of the main PDC and excess micelle peaks will allow for accurate buffer subtraction.

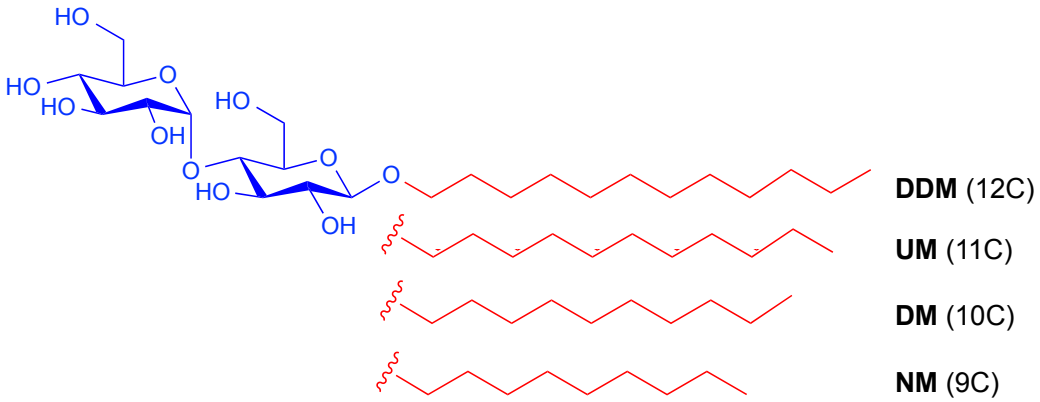


Figure 6.7 The structure of the maltoside detergents used in this project.

The figure is a replication of Figure 3.12. The common maltoside head-group structure is illustrated for the four detergents used in this project (blue), with the length of each of the aliphatic hydrocarbon chains (red) indicated by the number of carbons in parentheses (*n*C).

6.4.2 SEC-SAXS data collection

SEC-SAXS data were collected at the SWING beamline at the French national synchrotron Soleil, using the dedicated SEC-SAXS set-up at the beamline. For the SEC-SAXS studies, Mhp1 solubilised in maltoside detergents of different length was investigated: *n*-dodecyl-*β*-D-maltopyranoside (DDM), *n*-undecyl-*β*-D-maltopyranoside (UM), *n*-dodecyl-*β*-D-maltopyranoside (DM) and *n*-nonyl-*β*-D-maltopyranoside (NM) (Figure 6.7). As expected, the R_g and D_{max} values decrease for Mhp1 solubilised in detergents of shorter length (Table 6.3). The D_{max} values for Mhp1 solubilised in UM and DM do however not follow this trend, which could be due to poor quality data or the PDCs being partially unstable under the experimental conditions.

The previously determined D_{max} value for the Mhp1-DDM PDC (section 6.3.2) ranged between 110.8-156.6 Å, which includes almost the entire range of the values obtained from various PDCs using SEC-SAXS. This

demonstrates again that the initial SAXS data collected on the Mhp1-DDM PDC may not be of good enough quality to extract structural information about the PDC.

Table 6.3 The R_g and D_{max} values obtained for various Mhp1 PDCs.

Detergent PDC	Guinier R_g (Å)	D_{max} (Å)
DDM	41.3	122.8
UM	38.5	117.0
DM	37.3	117.4
NM	35.7	109.0

In this case the experimental D_{max} value for the DDM-Mhp1 PDC matches closely the estimated value of ~118-121.6 Å. The fact that UM and DM did not completely follow the trend of decreasing experimental values for the R_g and D_{max} with shorter detergent indicates that there might have been some problem with the samples, which is reflected in the data where the SAXS curves for all PDCs apart from the Mhp1-DDM PDC are considerably noisier (Figure 6.8). This could have been sample instabilisty due to the shipping and storage conditions, or this could be due to the fact that the PDC becomes more destabilised the shorter the surrounding detergent corona. In section 3.5.2, the tryptophan fluorescence quenching assay shows that Mhp1 reconstituted in shorter detergents becomes less efficient at binding L-BH, which could imply that the integrity of the Mhp1 protein becomes comprimised in a shorter detergent. Other reasons could be the formation of transient oligomers, as more of the protein becomes exposed when embedded within a smaller detergent micelle. A third potential reason for the poor data quality from the Mhp1-UM, Mhp1-DM and Mhp1-NM PDCs is that these complexes were prepared by detergent exchange from DDM to the each one of the shorter detergents. This could have resulted in a mixed detergent corona that becomes destabilised in a buffer containing only one detergent.

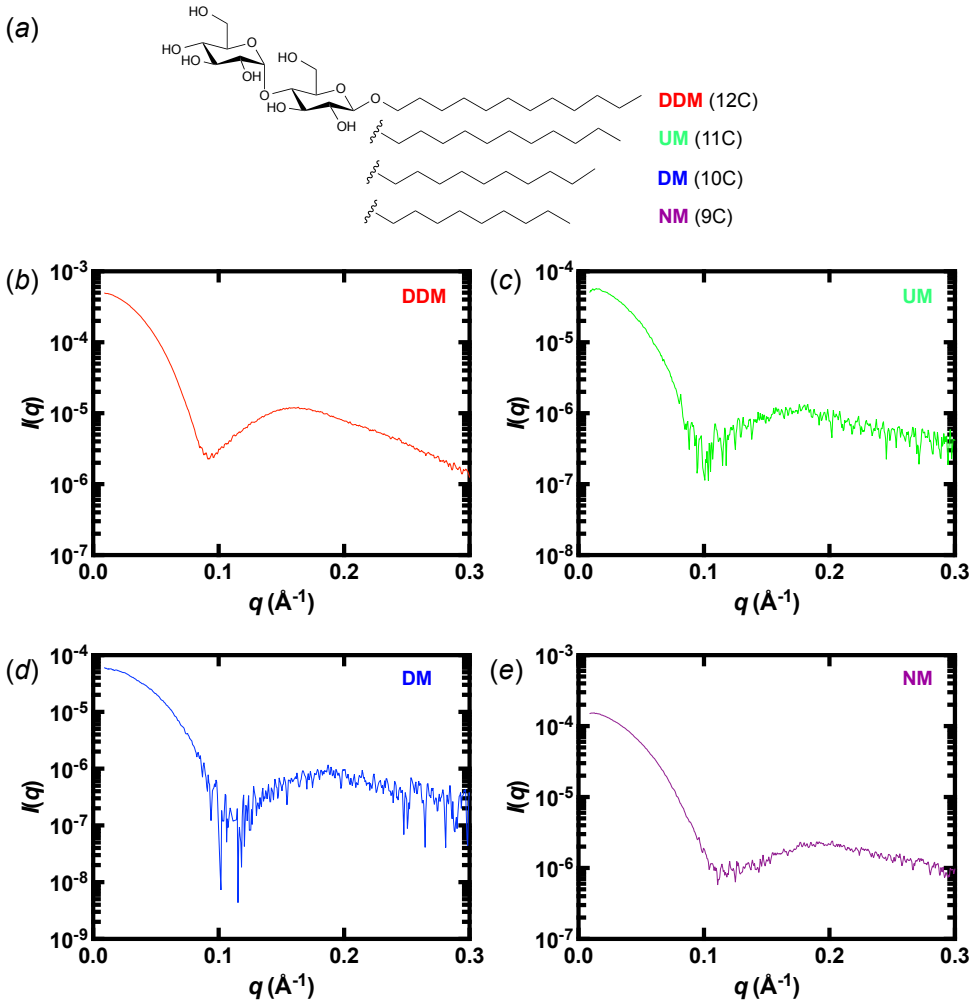


Figure 6.8 SEC-SAXS data from Mhp1 solubilised in maltoside detergents of varying length.

The illustration in (*a*) shows the chemical structure of each of the maltoside detergents used to create Mhp1-detergent PDCs. The colours indicated for each detergent corresponds to the SAXS curved displayed for the Mhp1-DDM (*b*), Mhp1-UM (*c*), Mhp1-DM (*d*) and Mhp1-NM (*e*) PDCs.

6.4.3 Discussion

SEC-SAXS data collection was performed in order to address the problems associated with SAXS studies on detergent-solubilised membrane proteins. In section 6.3.2, standard SAXS data collection on the Mhp1-DDM PDC was described, and it was shown to greatly suffer from concentration dependent effects and inaccurate buffer subtraction. The R_g and D_{max} values derived from these data varied greatly across the concentration range tested.

The SEC-SAXS data did display improvements with respect to the initial SAXS data of the DDM-Mhp1 PDC, where the sizes between the various PDCs could be distinguished, based on the obtained R_g and D_{max} values. The general trend was that the R_g and D_{max} values decreased the shorter the

detergent was, which is as expected as a shorter detergent will give rise to a smaller micelle. Some of the PDCs fell outside the trend, which is most likely explained by poor sample quality as indicated by the noisy data obtained for smaller PDCs. The Mhp1-PDC samples could have been affected by the shipping and storage conditions prior to the experiments. Another aspect that might affect the sample quality is how the PDC samples are prepared. During the purification of Mhp1, the protein is initially solubilised in DDM, and bound to the Ni-NTA resin while solubilised in DDM. After washing the resin with DDM-containing buffer to remove any unbound material, the exchange to a shorter detergent is achieved by applying a second wash buffer containing the shorter detergent of interest. It is difficult to quantify the extent of the detergent exchange, and whether the resulting detergent corona consists of a mixture of DDM and the shorter detergent, and what proportions of respective detergent constitutes the detergent corona. It is, therefore, not unlikely that the PDCs with UM, DM and NM consist of mixtures of each respective detergent and DDM. Since each elution was performed in a buffer containing only one detergent, the mixed detergent corona could be destabilised when diluted into the SEC running buffer, as the dilution might affect the effective concentrations of each detergent in the detergent corona. This suggests that a more rigorous detergent exchange protocol must be applied, involving for example solubilisation directly in the detergent of interest and/or followed by dialysis in the SEC buffer in order to ensure a homogeneous micelle. These results also indicate that the binding parameters obtained from tryptophan fluorescence quenching studies for Mhp1 in UM, DM and NM are for Mhp1 reconstituted into mixed detergent coronas.

The decrease in the experimentally obtained R_g values from the SEC-SAXS data, where the least amount of noise is observed still shows that a smaller PDC is obtained with a shorter detergent. When relating the experimentally obtained values for R_g and D_{max} for the Mhp1-DDM PDC to previous reference values for micelles alone, as well as estimated intermolecular distances of the Mhp1 protein from its crystal structure, they match closely. The estimated value for the largest intermolecular distance in the Mhp1-DDM PDC is ~118-121.6 Å, while the D_{max} for this PDC was determined to be 122.8 Å. Despite the close match of these values, it must be kept in mind that the estimated intra-molecular distances from the crystal structure of Mhp1 and previously obtained values for the detergent micelles are based on many assumptions.

Nevertheless, the comparison between experimental and estimated D_{max} values demostrate that collecting SAXS data coupled to SEC enables the differentiation between PDCs of varying size and is therefore a more effective technique to study detergent-solubilised membrane proteins than standard SAXS. The data of the highest quality was obtained for the Mhp1-DDM PDC, where the obtained experimental values fall into a reasonable range when considering the dimensions of the protein and micelles alone.

6.5 Modelling of the detergent corona around the Mhp1 protein

6.5.1 Introduction

Usually, SAXS data from biomolecules is used to generate low-resolution *ab initio* models of the biomolecule in solution (Svergun, 1999; Putnam *et al*, 2007). These methods are, however, not applicable for detergent-solubilised proteins, as the resulting PDCs consist of components of varying electron densities. In order to enable the modelling of the detergent corona around the protein from SEC-SAXS data, the program MEMPROT has been developed (Perez & Koutsioubas, 2015). In MEMPROT the protein part of the PDC is modelled as an all atom model, while the detergent corona is modelled as a coarse-grained elliptical corona consisting of dummy atoms, with electron densities defined for the hydrophobic interior as well as the hydrophilic exterior of the corona model based on previously determined values (Lipfert *et al*, 2007). The theoretical scattering from the PDC model is calculated using CRYSOL from the ATSAS suite (Svergun *et al*, 1995), and is then compared to the experimental SAXS curve. The fit of the theoretical scattering from the generated model vs. the experimental data is quantified by the value of χ (chi), defined by equation (3). For a perfect fit, the χ value approaches 1.

$$\chi = \left\{\frac{1}{N}\sum_i \frac{[I_{exp}(\boldsymbol{q_i}) - I_{exp}(\boldsymbol{q_i})]^2}{\sigma_i^2}\right\}^{1/2} \qquad (3)$$

The dimensions of the detergent corona are described by a set of parameters, describing the ellipticity and the thickness of the hydrophobic as well as the hydrophilic part. During a MEMPROT run, extreme values for these parameters are assigned at the start, which are then refined and assessed using the χ value. Previously, the detergent corona was

successfully modelled around the Aqp0 protein from SEC-SAXS data (Berthaud *et al*, 2012; Perez & Koutsioubas, 2015).

6.5.2 Modelling of the Mhp1-DDM PDC

Using the highest quality SEC-SAXS data, from the Mhp1-DDM PDC, modelling of the DDM detergent corona was attempted. Comparing the Kratky plot and the *P*(*r*) function of the SEC-SAXS data (Figure 6.9*b* and *c*, respectively) to those from the original data collection on DDM-solubilised Mhp1 (Figure 6.3) suggests that the buffer subtraction was successful and that the data represents one homogeneous species. The Kratky plot has one clear maximum, indicating that there is one well-folded species present in the sample, while the plot exhibited two maxima from the original data. The *P*(*r*) also appears to be smoother, suggesting that a more homogeneous sample is obtained by subjecting it to SEC. The small shoulder at approximately 0-20Å is still present, suggesting that it is a feature of the PDC, probably stemming from interatomic distances within the detergent corona.

The crystal structure of the Mhp1 protein, in an outward-open conformation (PDB 2JLN, Weyand *et al*, 2008) was used as the protein model. In order to place the Mhp1 protein on the same grid as the detergent corona dummy-atom model, the 2JLN structure was oriented to overlay with the crystal structure of the Aqp0. The advantage of the Aqp0 protein is that it is a tetramer, making the placement of this protein in the middle of the detergent corona straight forward. For the 2JLN structure, however, it was difficult to determine the unambiguous orientation of the protein molecule relative to the central axis of the detergent corona.

A wide range of the dimensional parameters of the detergent corona were defined in the calculation of the detergent corona model. The lowest χ value obtained in this modelling attempt was 1.8, while previously χ values have fallen in the range of ~1.3-1.5 (Perez *et al*, 2013; Berthaud *et al*, 2012). The fit of the theoretical scattering curve does indeed deviate from the experimental data at higher angles, as well as around the signal minimum before the detergent micelle peak at $q \approx 0.16$ Å^{-1} (Figure 6.9*a*). However, the dimensions of the resulting detergent corona match previously reported values for a DDM micelle alone, as summarised in Table 6.4. The dimensions ***a***, ***b*** and ***c*** quoted for the detergent corona surrounding the

Mhp1 protein in Figure 6.9*d* correspond to the same dimensions of free micelles as illustrated in Figure 6.5.

The resulting modelled DDM corona surrounding the Mhp1 protein matches the previously reported values for DDM micelles alone (Lipfert *et al*, 2007), suggesting that the buffer subtraction of the data for the Mhp1-DDM PDC was successful by coupling the SAXS data collection with SEC. The relatively high χ value for the theoretical scattering of the model against the experimental data could be a result of inappropriate orientation of the Mhp1 in the model of the detergent corona, as this was done arbitrarily in order to place the Mhp1 molecule in line with the central axis of the detergent model.

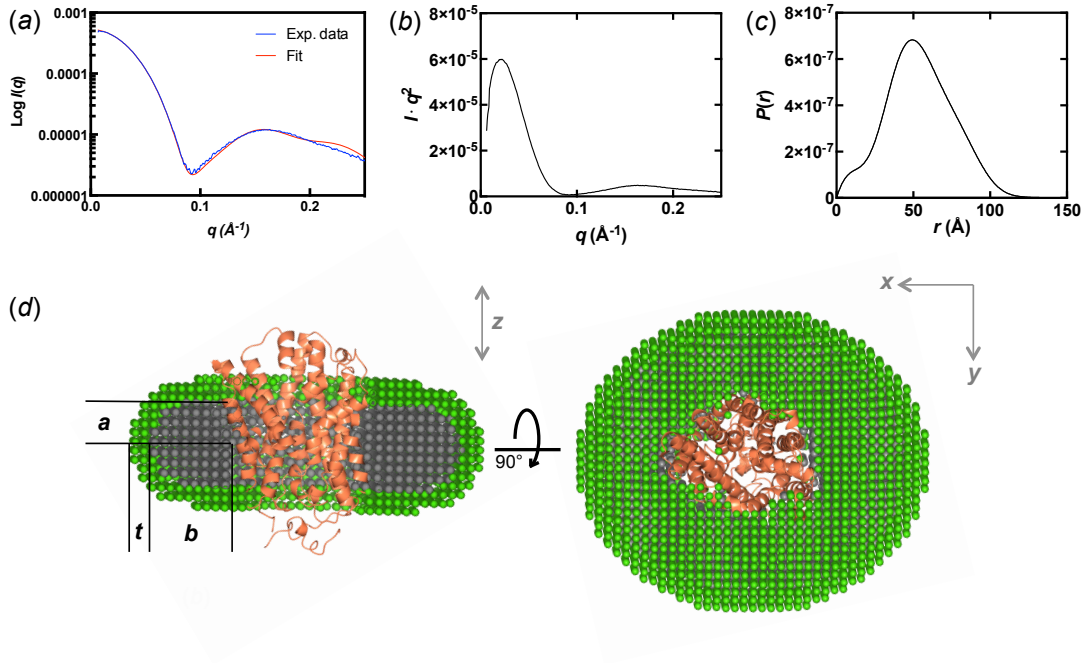


Figure 6.9 A model of the Mhp1-DDM PDC modelled from SEC-SAXS data, and the fit of the model against experimental data.

The experimental data (blue) and the fit of the Mhp1-DDM PDC model (red) are shown (*a*), alongisde the Kratky plot (*b*) and the *P*(*r*) function (*c*) from the experimental data. The resulting model (*d*) is comprised of the Mhp1 protein (PDB 2JLN, shown in a cartoon representation coral), with the surrounding detergent corona illustrated as a belt of spheres. The grey spheres correspond to the hydrophobic interior of the detergent corona, while the green spheres correspond to the hydrophilic head group dummy atoms of the detergent corona model. The parameters ***a***, ***b*** and ***t*** correspond to those in Figure 6.5. The coordinate system for the detergent corona, as defined in the modelling program, is indicated by the grey arrows denoting the directions for the ***x***-, ***y***- and ***z***-axes.

Modelling of the 2JLN crystal structure in different orientations all gave rise to considerably higher χ values and poor fits of the theoretical scattering of the model to the experimental data. The same was obtained when the detergent corona was modelled around crystal structures of the Mhp1 protein in other conformational states, indicating that this methodology could be used to distinguish between preferred conformational states of the protein in various detergents.

Table 6.4 The dimensions of the Mhp1-DDM PDC detergent corona model compared to previously reported dimensions for empty DDM micelles.

The values ***a***, ***b*** and ***t*** are based on the model illustrated in Figure 6.5, and quoted in Lipfert *et al.* (2007)*.

DDM corona	***a* (Å)**	***b* (Å)**	***t* (Å)**
Model	13.5	31.0	6.0
Reference*	13.8-14.3	28.0-29.5	6.0-6.3

6.5.3 Discussion

The modelling of the Mhp1-DDM PDC is described in this section, using the recently developed program MEMPROT. The data for the PDC obtained by SEC-SAXS data collection were used to model the PDC with an all-atom representation of the protein using a dummy-atom model of the surrounding detergent corona. The resulting dimensions of the DDM detergent corona surrounding the Mhp1 protein matched the experimental values previously obtained for DDM micelles alone. This supports the conclusion that the buffer subtraction in the SEC-SAXS data collection was successful.

Even though the model appeared to match previously obtained values well, the χ value of the fit of the theoretical scattering form the PDC model against the experimental data was high, relative to previously obtained values for similar models (Perez & Koutsioubas, 2015). This could be explained by the fact that several assumptions are introduced during the course of preparing and performing the modelling procedure. Firstly, a crystal structure is used to represent the protein part of the PDC. This is a static structure and would thus not account for any conformational flexibility in the actual PDC sample. For the model generated here, the protein in an outward-open form was used as the protein model. When different crystal structures were tested (by superposing them onto the 2JLN position in the best model obtained here),

significantly higher χ values were obtained for the models, giving rise to a poor fit of the theoretical scattering to the experimental data. It is not unlikely that Mhp1 samples several conformations in solution (Kazmier *et al*, 2014), which makes the assumption that the protein remains static in the PDC model a limitation in the modelling approach.

Secondly, the orientation of the Mhp1 molecule within the detergent belt may not be representative of the true orientation of the protein within the detergent micelles. In the case of Aqp0, its four-fold symmetry ensured a relatively reliable placement in the *xy*-plane, where adjustments could be made using the rotation parameter. In the direction of the *z*-axis, the placement was guided by the hydrophobic belt and the relatively large solvent-exposed tail regions. The monomeric Mhp1 does not possess any symmetry elements that could guide the placement within a theoretical detergent micelle, and the placement of the Mhp1 protein was initially done arbitrarily onto the central axis of the detergent micelle.

A third assumption is imposed by the modelling program, where the detergent corona is assumed to adopt an elliptical shape. While this could be true for the case of Aqp0, for which the resulting PDC models gave rise to relatively low χ values, this is not necessarily the case for Mhp1. The Aqp0 associates as a tetramer into the detergent corona, and being bigger and having a different shape of the hydrophobic belt, it could be imposing different constraints on the curvature of the detergent corona.

Suggested improvements to the MEMPROT program in order to improve its applicability to all kinds of systems are to introduce further parameters that could help to systematically sample more orientations of the protein within the detergent corona. For example, an additional parameter could be the tilt of the detergent corona with respect to the central axis. Alternatively, the detergent corona belt modelling could also be attempted in an *ab initio* fashion, as applied in the reconstruction of low-resolution models from SAXS data (Svergun, 1999). Rather than defining parameter extremes that are scanned though to find the model with the best fit between the theoretical and experimental scattering data, doing an *ab initio* model of the corona could account for any shape distortions imposed on the detergent belt by the protein, any altered orientation of the protein in the belt and also save computational power.

6.6 General discussion and conclusions

In this Chapter, different approaches for SAXS data collection of Mhp1 PDCs in maltoside detergents of varying length are described. SAXS data collection on the Mhp1-DDM PDC using a conventional buffer exchange method did not give rise to reliable data, as the R_g and D_{max} values exhibited a concentration dependence, possibly due to the procedure for the preparation of the dilution series and resulting inconsistent buffer subtraction.

This problem was overcome by performing SEC-SAXS studies of Mhp1 PDCs, with maltoside detergents of varying length. The resulting studies showed that SEC-SAXS data can successfully be collected on Mhp1 PDCs, where the decreasing size of the PDC with shorter detergent length was reflected in the resulting R_g and D_{max} values. In the case of SEC-SAXS data collection, there were problems associated with the sample quality. For the detergents UM, DM and NM, the data were very noisy. This could be explained by the purification procedure used to prepare these PDCs, which could give rise to a mixed detergent corona (consisting of DDM and the detergent the PDC has been exchanged into). Therefore, the PDCs containing UM, DM and NM may have been destabilised by being diluted into buffer containing each respective detergent. This suggests that alternative sample preparation protocols need to be employed in order to study these PDCs.

Nevertheless, the SEC-SAXS data obtained from the Mhp1-DDM PDC could be used to create a model of the PDC with the protein molecules as an all-atom model and the surrounding detergent belt represented as a dummy-atom model. The dimensions of the detergent belt in the resulting PDC model match previously obtained values for DDM micelles alone, indicating that the buffer subtraction and SEC-SAXS data collection was successful for the Mhp1-DDM PDC. The data for the other PDCs suggests that, in order to study these, further optimisation of the purification and/or the sample preparation is required.

In conclusion, these experiments serve as a proof-of-principle that the suggested SEC-SAXS data collection and detergent corona modelling can be applied to a monomeric secondary-active transporter. For the case of Mhp1, additional optimisation of the sample preparation procedure is required, as well as potential adjustments of the MEMPROT software used for the modelling of the detergent corona.

The SEC-SAXS data, despite the potential associated problems with the sample quality, show that shorter detergents do give rise to a smaller PDC. In terms of the crystallisability of Mhp1 in various detergents, this is consistent with the previous suggestion that a smaller detergent micelle surrounding the protein results in the exposure of parts of the protein that participate in crystal contacts. The resulting model for the PDC gives another insight into the crystal packing of Mhp1 molecules. From the modelled Mhp1-DDM PDC, it is evident that the detergent creates a significantly bulky belt around the Mhp1 protein.

From previously obtained values for detergent micelles alone (Table 6.2), the thickness of a DDM micelle in its horizontal axis is suggested to be 34-35.8 Å (***b***+***t*** = 28.0-29.5 + 6.0-6.3 Å), while the thickness of a DM micelle is 29-30.3 Å (***b***+***t*** is 23.0-24.0 + 6.0-6.3 Å). If extrapolating these values, the thickness of an NM micelle could be approximately 27.5 Å. The distance between NM solubilised Mhp1 molecules in the lattice of the crystals obtained as part of this Thesis is approximately 40 Å (see Figure 4.10), suggesting that the detergent phase is potentially highly crowded within the crystal lattice. The fact that DDM solubilised Mhp1 protein does not crystallise could be due to the fact that the detergent corona simply does not allow the formation of crystal contacts due to steric hindrance.

This suggests that the crystallisability of the Mhp1 protein in various detergents could depend more on crowding of detergent belts rather than different degrees of solvent-exposed regions being able to participate in crystal contacts. Even in the case of NM, the detergent phase is probably highly crowded, which could account for the fragility of the crystals and the low quality of the resulting diffraction data.

Chapter 7: Conclusions and future directions

Membrane proteins are very important targets for structural studies and subsequent drug design, however, they are very challenging to study due to their hydrophobic nature. An important class of membrane transport proteins are the secondary-active transport proteins which are responsible for the transport of ions, metabolites, vitamins and signal substances across cell membranes.

The focus of this Thesis is the secondary-active, bacterial sodium-hydantoin transporter Mhp1. In an effort to gain more knowledge about the transport mechanism of Mhp1, serving as a model system for related eukaryotic transport proteins, the Mhp1 protein was studied using a combination of X-ray crystallography and small-angle X-ray scattering.

The transport mechanism of Mhp1 was probed using a series of single-point mutants, which had previously been identified to have altered ligand binding behaviour. Mutants of Mhp1 have been successfully produced and characterised (Chapter 3) for subsequent structural studies. The Mhp1 mutants were subjected to crystallisation trials, using previously published crystallisation conditions as a starting point (Weyand *et al*, 2008; Shimamura *et al*, 2010). Crystals of Mhp1 mutants were successfully produced, and after considerable optimisation, reproducibly well-diffracting crystals (<4Å) were reproducibly obtained (Chapter 4). This was achieved by a combination of applying modifications to the crystallisation droplet morphology, controlled dehydration and contact-less harvesting. The crystals were of various Mhp1 mutants in complex with various hydantoin ligands, and gave rise to a series of crystal structures of novel mutant-ligand complexes (Chapter 5).

The resulting structures gave some mechanistic insights into the transport mechanism, but no conformational state other than those previously observed in published crystal structures of the Mhp1 protein was obtained. The structures indicated that the crystal lattice and crystallisation conditions may play a significant role in the conformational state adopted by the Mhp1 protein in the crystal. Therefore, investigating the applicability of small-angle X-ray scattering to study the Mhp1 protein was a suitable route to obtain further structural information about the protein, as the measurements are carried out in solution without any restrictions imposed by the crystal lattice. As with crystallographic studies, membrane proteins are challenging targets

to study using SAXS, and the problems associated with detergents present in the sample buffer were addressed (Chapter 6).

The mutant-ligand structures described in Chapter 5 are of single-point mutants with locations of the mutations in three different areas of the Mhp1 protein: the ligand binding site, the sodium binding site and the interface between the so-called bundle- and hash-motifs, which undergo a rotation relative to one another upon the transition from the ligand-bound occluded form to the inward-open form. The Mhp1 mutants crystallised in the presence of various hydantoin ligands.

However, no additional conformational state was obtained for the Mhp1 protein, other than those previously observed. At the working resolution of the crystal structures of the mutant-ligand complexes presented in this Thesis, small alterations to the conformation and/or changes of single residues of the Mhp1 protein are likely to go undetected. In order to gain insight into these finer alterations to the structure due to single-point mutations, higher resolution crystal structures would be required.

The results indicate that the crystallisation conditions used, and the resulting crystal packing, may also be a limiting factor to what parts of the conformational landscape of the transport mechanism of Mhp1 can be investigated. It appears that the $P2_12_12_1$ crystal form is highly favoured for Mhp1 crystals and previously published crystal structures, along with the crystal structures presented as part of this Thesis, indicate that this crystal form accommodates and/or drives the conformation of the Mhp1 protein into an outward-open or occluded form. This could be a result of the conformational equilibrium of Mhp1 being sampled in the solution, where the inward-open form may be a short-lived intermediate, that does not exist for long enough to form protein crystal nuclei. The crystal lattice could also be acting as a scaffold for Mhp1, forcing the protein into a conformational state that fits in the crystal lattice.

Apart from attempting to crystallise a different crystal form using single-point mutants of the Mhp1 protein, reproducing the crystal structure of the inward-open form of seleno-L-methionine derivatised Mhp1 was also attempted, which has previously crystallised in a different crystal form (in space group $P6_1$). Crystals were obtained of seleno-L-methionine derivatised Mhp1 under different crystallisation conditions, but not of the expected crystal morphology and the resulting crystals were not of good enough quality to harvest for diffraction screening and the crystal form could thus not be assessed. A possible reason for the fact that the expected crystal form of

crystals of seleno-L-methionine derivatised Mhp1 were not obtained was the absence of a expression and/or purification batch specific ligand that may have been altering the conformational equilibrium of the protein to stabilise the inward-open form and thus, promote the crystallisation of the inward-open conformation.

At the current stage, remaining question about the structure and transport mechanism of the Mhp1 protein are: what are the conformations of the apo form of the single-point mutants; what is the structure of the inward-open form of native Mhp1; and what intermediate conformational states does the Mhp1 protein adopts between the transitions between outward-open → occluded → inward-open state?

The next step in order to investigate these question and access a greater portion of the conformations adopted by Mhp1 during its transport mechanism is to perform broad re-screening for crystallisation conditions for the single-point mutants described here. Identifying alternative crystallisation conditions could allow the crystallisation and exploration of alternative conformations of the protein. In addition, the library of mutants could also be expanded to include further single-point mutants, as well as Mhp1 variants containing combinations of mutations. In addition, low-sodium conditions for the orthorhombic crystals could also be explored for other mutants, as done for the Gln42Asn mutant in complex with L-BH, which was crystallised in the presence of a high potassium concentration.

An alternative avenue to obtain a new crystal form and obtain higher-resolution crystal structures is to move to lipid based crystallisation methods, screening for conditions in, for example, lipidic cubic or lipid sponge phase (Caffrey & Cherezov, 2009; Wöhri *et al*, 2008). This would most likely give rise to type I protein crystals, in which the packing of the hydrophobic parts of the protein molecules could allow a more stable crystal lattice, giving rise to higher resolution diffraction data. These crystals would also have no detergent present, which has been demonstrated to potentially limit the diffraction quality that can be obtained from crystals of detergent-solubilised Mhp1. In the studies carried out as part of this Thesis, attempts were made in order to improve the resolution of the Mhp1 crystals by investigating alternative cryo-cooling and harvesting approaches, out of which controlled dehydration and contact-less harvesting by photo-ablation proved to be the most sucecssful strategies. However, despite the fact that the techniques showed potential in improving the data, the resolution limit of 3.5-4 Å was never exceeded. This is likely to be due to the presence of the surrounding

detergent corona around the Mhp1 proteins, which is of a finite size and thus, has a limit to how much it can be contracted in order to achieve a tighter and more well-diffracting lattice. Therefore, changing the environment of the protein molecules in the crystal lattice, either by using shorter detergents or by moving onto lipid-based crystallisation methods, is most likely a better strategy in the future of the structural studies of the Mhp1 protein and secondary-active transport proteins.

Nevertheless, controlled dehydration using the HC1 device, as well as contact-less harvesting by photo-ablation have been presented as successful strategies to improve the crystal quality. The studies performed as part of this Thesis have highlighted that the hydration level is important for detergent-solubilised membrane protein crystals, and that various dehydration protocols should be explored in the process of optimising membrane protein diffraction.

The crystallographic work presented in this Thesis has demonstrated the challenges and limitations that can be encountered with membrane proteins, it has shown that the crystal lattice may have a profound effect on the structural information that can be obtained using X-ray crystallography. Therefore, the investigation of SAXS as a complementary technique was a suitable complement to the crystallographic studies. There are several different solution based techniques that can be used to gain complementary functional and dynamics information about proteins, for example, EPR (Kazmier *et al*, 2014) and FRET (Zhao *et al*, 2010), which have previously been carried out on Mhp1 and related proteins. The advantage of SAXS lies in the fact that structural information can be directly extracted. With the recent improvements in the data collection and modelling from SAXS data of membrane proteins, this is a very interesting technique to investigate.

In this Thesis, SEC-SAXS studies were performed on PDCs of the Mhp1 protein and maltoside detergents of varying length. Even though the studies indicated that there are potential issues with the sample preparation procedure of these PDCs, they allowed for the data collection from the Mhp1-DDM PDC that allowed the modelling of the surrounding DDM corona around the Mhp1 protein. These studies served as a proof-of-principle that the method is applicable to the study of secondary-active transporters, highlighting necessary improvements and optimisation of the sample preparation and modelling procedure.

In the current modelling approach of the PDC from SEC-SAXS data involves assuming the protein conformation is static. Future directions of the

modelling from SEC-SAXS data could be to assume the detergent corona as the static part of the system, while allowing the *ab initio* modelling of the protein conformation. This approach would require substantial extracellular domains, that would protrude from the detergent corona and undergo conformational changes that are significant enough to be modelled from the SEC-SAXS data. The extracellular domains of, for example, Mhp1 are not very large and it is difficult to estimate whether the subtle movements of the helices upon ligand binding and release could be modelled from SEC-SAXS data. Nevertheless, introducing conformational markers, could be a strategy to make it possible to study the conformational changes of the Mhp1 protein and related secondary-active transporters. If the conformational markers could be compatible with protein crystallisation trials such as, for example, maltose binding proteins (Smyth *et al*, 2003; Moon *et al*, 2010), this could be a future approach for complementary studies of membrane proteins.

List of references

Abergel C (2004) Spectacular improvement of X-ray diffraction through fast desiccation of protein crystals. *Acta Crystallogr D Biol Crystallogr* **60:** 1413–1416

Abergel C, Coutard B, Byrne D, Chenivesse S, Claude J-B, Deregnaucourt C, Fricaux T, Gianesini-Boutreux C, Jeudy S, Lebrun R, Maza C, Notredame C, Poirot O, Suhre K, Varagnol M & Claverie J-M (2003) Structural genomics of highly conserved microbial genes of unknown function in search of new antibacterial targets. *J Struct Funct Genomics* **4:** 141–157

Abramson J, Smirnova I, Kasho V, Verner G, Kaback HR & Iwata S (2003) Structure and mechanism of the lactose permease of Escherichia coli. *Science* **301:** 610–615

Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC & Zwart PH (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66:** 213–221

Adelman JL, Dale AL, Zwier MC, Bhatt D, Chong LT, Zuckerman DM & Grabe M (2011) Simulations of the alternating access mechanism of the sodium symporter Mhp1. *Biophys. J.* **101:** 2399–2407

Alcorn T & Juers DH (2010) Progress in rational methods of cryoprotection in macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66:** 366–373

Arinaminpathy Y, Khurana E, Engelman DM & Gerstein MB (2009) Computational analysis of membrane proteins: the largest class of drug targets. *Drug Discov Today* **14:** 1130–1135

Austrian R (1960) The Gram stain and the etiology of lobar pneumonia, an historical note. *Bacteriol Rev* **24:** 261–265

Awad W, Svensson Birkedal G, Thunnissen MMGM, Mani K & Logan DT (2013) Improvements in the order, isotropy and electron density of glypican-1 crystals by controlled dehydration. *Acta Crystallogr D Biol Crystallogr* **69:** 2524–2533

Axford D, Owen RL, Aishima J, Foadi J, Morgan AW, Robinson JI, Nettleship JE, Owens RJ, Moraes I, Fry EE, Grimes JM, Harlos K, Kotecha A, Ren J, Sutton G, Walter TS, Stuart DI & Evans G (2012) In situ macromolecular crystallography using microbeams. *Acta Crystallogr D Biol Crystallogr* **68:** 592–600

Berejnov V, Husseini NS & Alsaied OA (2006) Effects of cryoprotectant

concentration and cooling rate on vitrification of aqueous solutions. *Journal of Applied Crystallography* **39**: 244

Berthaud A, Manzi J, Perez J & Mangenot S (2012) Modeling detergent organization around aquaporin-0 using small-angle X-ray scattering. *J. Am. Chem. Soc.* **134:** 10080–10088

Bramley HF & Kornberg HL (1987) Sequence homologies between proteins of bacterial phosphoenolpyruvate-dependent sugar phosphotransferase systems: identification of possible phosphate-carrying histidine residues. *Proc. Natl. Acad. Sci. U.S.A.* **84:** 4777–4780

Brock TD (1961) Milestones in Microbiology. *Academic Medicine*

Brockhauser S, Svensson O, Bowler MW, Nanao M, Gordon E, Leal RMF, Popov A, Gerring M, McCarthy AA & Gotz A (2012) The use of workflows in the design and implementation of complex experiments in macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **68:** 975–984

Bryant DA & Frigaard N-U (2006) Prokaryotic photosynthesis and phototrophy illuminated. *Trends Microbiol.* **14:** 488–496

Caffrey M & Cherezov V (2009) Crystallizing membrane proteins using lipidic mesophases. *Nature* **4:** 706–731

Carpenter EP, Beis K, Cameron AD & Iwata S (2008) Overcoming the challenges of membrane protein crystallography. *Curr. Opin. Struct. Biol.* **18:** 581–586

Chacón P, Morán F, Díaz JF, Pantos E & Andreu JM (1998) Low-resolution structures of proteins in solution retrieved from X-ray scattering with a genetic algorithm. *Biophys J.* **74:** 2760–2775

Chayen NE (2003) Crystallization of membrane proteins in oils. In *Methods and Results in Crystallization of Membrane Proteins* (S. Iwata, Ed.): 131-139.

Chen Y & Barkley MD (1998) Toward understanding tryptophan fluorescence in proteins. *Biochemistry* **37**: 9976-9982

Cipriani F, Rower M, Landret C, Zander U, Felisaz F & Marquez JA (2012) CrystalDirect: a new method for automated crystal harvesting based on laser-induced photoablation of thin films. *Acta Crystallogr D Biol Crystallogr* **68:** 1393–1399

Clemente-Jiménez JM, Martínez-Rodríguez S, Rodríguez-Vico F & Heras-Vázquez FJL (2008) Optically pure alpha-amino acids production by the "Hydantoinase Process". *Recent Pat Biotechnol* **2:** 35–46

de Koning H & Diallinas G (2000) Nucleobase transporters (review). *Mol. Membr. Biol.* **17:** 75–94

Delgado GE, Mora AJ, Uzcátegui J, Bahsas A & Briceño A (2007) (S)-5-Benzylimidazolidine-2,4-dione monohydrate. *Acta Crystallogr C* **63:** 448–450

Deng D, Xu C, Sun P, Wu J, Yan C, Hu M & Yan N (2014) Crystal structure of the human glucose transporter GLUT1. *Nature* **510:** 121–125

Dhouib K, Khan Malek C, Pfleging W, Gauthier-Manuel B, Duffait R, Thuillier G, Ferrigno R, Jacquamet L, Ohana J, Ferrer J-L, Théobald-Dietrich A, Giegé R, Lorber B & Sauter C (2009) Microfluidic chips for the crystallization of biomacromolecules by counter-diffusion and on-chip crystal X-ray analysis. *Lab Chip* **9:** 1412–1421

Diallinas G (2008) An Almost-Complete Movie. *Science* **322:** 1644–1645

Drew D, Fröderberg L, Baars L & de Gier J-WL (2003) Assembly and overexpression of membrane proteins in Escherichia coli. *Biochim. Biophys. Acta* **1610:** 3–10

Dupeux F, Röwer M, Seroul G, Blot D & Márquez JA (2011) A thermal stability assay can help to estimate the crystallization likelihood of biological samples. *Acta Crystallogr D Biol Crystallogr* **67:** 915–919

Duquesne K & Sturgis JN (2010) Membrane protein solubilization. *Methods Mol. Biol.* **601:** 205–217

Emamzadah S, Petty TJ, De Almeida V, Nishimura T, Joly J, Ferrer J-L & Halazonetis TD (2009) Cyclic olefin homopolymer-based microfluidics for protein crystallization and in situ X-ray diffraction. *Acta Crystallogr D Biol Crystallogr* **65:** 913–920

Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66:** 486–501

Eshaghi S, Hedrén M, Nasser MIA, Hammarberg T, Thornell A & Nordlund P (2005) An efficient strategy for high-throughput expression screening of recombinant integral membrane proteins. *Protein Sci.* **14:** 676–683

Esnouf RM, Ren J, Garman EF, Somers DO, Ross CK, Jones EY, Stammers DK & Stuart DI (1998) Continuous and discontinuous changes in the unit cell of HIV-1 reverse transcriptase crystals on dehydration. *Acta Crystallogr D Biol Crystallogr* **54:** 938–953

Faham S & Bowie JU (2002) Bicelle crystallization: a new method for crystallizing membrane proteins yields a monomeric bacteriorhodopsin structure. *Journal of Molecular Biology* **316**: 1-6

Faham S, Watanabe A, Besserer GM, Cascio D, Specht A, Hirayama BA, Wright EM & Abramson J (2008) The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na+/sugar symport. *Science* **321:** 810–814

Fang Y, Jayaram H, Shane T, Kolmakova-Partensky L, Wu F, Williams C,

Xiong Y & Miller C (2009) Structure of a prokaryotic virtual proton pump at 3.2 A resolution. *Nature* **460:** 1040–1043

Forrest LR, Krämer R & Ziegler C (2011) The structural basis of secondary active transport mechanisms. *Biochim. Biophys. Acta* **1807:** 167–188

Forsythe EL, Maxwell DL & Pusey M (2002) Vapor diffusion, nucleation rates and the reservoir to crystallization volume ratio. *Acta Crystallogr Section D Biol Crystallogr* **58**: 1601-1605

Fromme P (2003) Crystallization of photosystem I. In *Methods and Results in Crystallization of Membrane Proteins* (S. Iwata, Ed.): 355.

Gadsby DC (2009) Ion channels versus ion pumps: the principal difference, in principle. *Nat. Rev. Mol. Cell Biol.* **10:** 344–352

Gadsby DC, Bezanilla F & Rakowski RF (2012) The dynamic relationships between the three events that release individual Na+ ions from the Na+/K+-ATPase. *Nature Comms* **3**: 669

Garman EF & Mitchell EP (1996) Glycerol concentrations required for cryoprotection of 50 typical protein crystallization solutions. *J Appl Crystallogr* **29**: 58487

Garman EF & Schneider TR (1997) Macromolecular cryocrystallography. *J Appl Crystallogr* **30:** 211-237

Gautheron DC (1984) Mitochondrial oxidative phosphorylation and respiratory chain: review. *J Inherit Metab Dis.* **7** :57-61

Ge L, Perez C, Waclawska I, Ziegler C & Muller DJ (2011) Locating an extracellular K+-dependent interaction site that modulates betaine-binding of the Na+-coupled betaine symporter BetP. *Proc Natl Acad Sci USA* **108:** 890–898

Glusker, Jenny P (1983) The Patterson Function. *Trends in Biochemical Sciences* **9**: 328–30

Gourdon P, Andersen JL, Hein KL & Bublitz M (2011) HiLiDe - Systematic Approach to Membrane Protein Crystallization in Lipid and Detergent. *Crystal growth & design* **11**: 2098-2106

Grossmann JG (2007) Biological solution scattering: recent achievements and future challenges. *J Appl Crystallogr* **40**: 217-222

Guan L, Mirza O, Verner G, Iwata S & Kaback HR (2007) Structural determination of wild-type lactose permease. *Proc. Natl. Acad. Sci. U.S.A.* **104:** 15294–15298

Guinier AFF (1955) Small-angle Scattering of X-rays. New York: Wiley Interscience

Gutmann DAP, Mizohata E, Newstead S, Ferrandon S, Postis V, Xia X,

Henderson PJF, van Veen HW & Byrne B (2007) A high-throughput method for membrane protein solubility screening: the ultracentrifugation dispersity sedimentation assay. *Protein Sci.* **16:** 1422–1428

Henderson PJF (2012) Membrane Transport, Energetics and Overview Roberts GCK Ed. Springer-Verlag

Heras B & Martin JL (2005) Post-crystallization treatments for improving diffraction quality of protein crystals. *Acta Crystallogr D Biol Crystallogr* **61:** 1173–1180

Heras B, Edeling MA, Byriel KA, Jones A, Raina S & Martin JL (2003) Dehydration converts DsbG crystal diffraction from low to high resolution. *Structure/Folding and Design* **11:** 139–145

Hope H (1988) Cryocrystallography of biological macromolecules: a generally applicable method. *Acta Crystallogr B* **44:** 22–26

Huang Y, Lemieux MJ, Song J, Auer M & Wang D-N (2003) Structure and mechanism of the glycerol-3-phosphate transporter from Escherichia coli. *Science* **301:** 616–620

Jackson SM (2012) Elucidating the molecular mechanisms of ligand binding and transport by the Na^+-hydantoin transport protein, Mhp1. *Ph. D. Thesis. University of Leeds, U.K.*

Jacques DA, Guss JM, Svergun DI & Trewhella J (2012) Publication guidelines for structural modelling of small-angle scattering data from biomolecules in solution. *Acta Crystallogr D Biol Crystallorg* **68**: 620-626

Jardetzky O (1966) Simple Allosteric Model for Membrane Pumps. *Nature* **211:** 969–970

Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* **66:** 125–132

Karplus PA & Diederichs K (2012) Linking crystallographic model and data quality. *Science* **336:** 1030–1033

Kazmier K, Sharma S, Islam SM, Roux B & Mchaourab HS (2014) Conformational cycle and ion-coupling mechanism of the Na+/hydantoin transporter Mhp1. *Proc. Natl. Acad. Sci. U.S.A.* **111:** 14752–14757

Kelly SM, Jess TJ & Price NC (2005) How to study proteins by circular dichroism. *Biochimica et Biophysica Acta* **1751**:119-139

Kiefersauer R, Than ME, Dobbek H, Gremer L, Melero M, Strobl S, Dias JM, Soulimane T & Huber R (2000) A novel free-mounting system for protein crystals: transformation and improvement of diffraction power by accurately controlled humidity changes. *J Appl Crystallogr* **33:** 1223–1230

King LS, Kozono D & Agre P (2004) From structure to disease: the evolving tale of aquaporin biology. *Nat. Rev. Mol. Cell Biol.* **5:** 687–698

Koshy C, Schweikhard ES, Gärtner RM, Perez C, Yildiz Ö & Ziegler C (2013) Structural evidence for functional lipid interactions in the betaine transporter BetP. *EMBO J.* **32:** 3096–3105

Krishnamurthy H & Gouaux E (2012) X-ray structures of LeuT in substrate-free outward-open and apo inward-open states. *Nature* **481:** 469–474

Krishnamurthy H, Piscitelli CL & Gouaux E (2009a) Unlocking the molecular secrets of sodium-coupled transporters. *Nature* **459:** 347–355

Lamb JD, Izatt RM & Swain CS (1980) A systematic study of the effect of macrocycle ring size and donor atom type on the log K, .DELTA.H, and T.DELTA.S of reactions at 25.degree.C in methanol of mono-and divalent cations with crown ethers. *J Am Chem Soc* **102**: 475-479

Landau EM & Rosenbusch JP (1996) Lipidic cubic phases: a novel concept for the crystallization of membrane proteins. *Proc. Natl. Acad. Sci. U.S.A.* **93:** 14532–14535

Laughlin RG (1996) The Aqueous Phase Behavior of Surfactants. Academic Press.

Lee C-C, Maestre-Reyna M, Hsu K-C, Wang H-C, Liu C-I, Jeng W-Y, Lin L-L, Wood R, Chou C-C, Yang J-M & Wang AH-J (2014) Crowning proteins: modulating the protein surface properties using crown ethers. *Angew. Chem. Int. Ed. Engl.* **53:** 13054–13058

Lipfert J, Columbus L, Chu VB, Lesley SA & Doniach S (2007) Size and Shape of Detergent Micelles Determined by Small-Angle X-ray Scattering. *J. Phys. Chem. B* **111:** 12427–12438

Lodish H, Berk A, Zipursky SL & Matsudaira P (2000) Active Transport by ATP-Powered Pumps. In *Molecular Cell Biology. 4th edition*. New York: W. H. Freeman; section 15.5

Loll PJ, Tretiakova A & Soderblom E (2003) Compatibility of detergents with the microbatch-under-oil crystallization method. *Acta Crystallogr D Biol Crystallogr* **59:** 1114–1116

Manuel García-Ruiz J (2003) Nucleation of protein crystals. *Journal of Structural Biology* **142:** 22–31

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007) Phaser crystallographic software. *J Appl Cryst* **4**: *658-674*

McFerrin MB & Snell EH (2002) The development and application of a method to quantify the quality of cryoprotectant solutions using standard area-detector X-ray images. *J Appl Crystallogr* **35**: 538-545

McNicholas S, Potterton E, Wilson KS & Noble MEM (2011) Presenting your structures: the CCP4mg molecular-graphics software. *Acta Crystallogr D Biol Crystallogr* **67:** 386–394

McPherson A (2011) Introduction to Macromolecular Crystallography. John Wiley & Sons.

Michel H (1991) Crystallization of Membrane Proteins. CRC Press Inc.

Mitra K, Ubarretxena-Belandia I, Taguchi T, Warren G & Engelman DM (2004) Modulation of the bilayer thickness of exocytic pathway membranes by membrane proteins rather than cholesterol. *Proc. Natl. Acad. Sci. U.S.A.* **101:** 4083–4088

Monaco S, Gordon E, Bowler MW, Delagenière S, Guijarro M, Spruce D, Svensson O, McSweeney SM, McCarthy AA, Leonard G & Nanao MH (2013) Automatic processing of macromolecular crystallography X-ray diffraction data at the ESRF. *J Appl Crystallogr* **46:** 804–810

Moon AF, Mueller GA, Zhong X & Pedersen LC (2010) A synergistic approach to protein crystallization: Combination of a fixed-arm carrier with surface entropy reduction. *Protein Science* **19**: 901-913

Murphy BM, Swarts S, Mueller BM, van der Geer P, Manning MC & Fitchmun MI (2013) Protein instability following transport or storage on dry ice. *Nat Meth* **10:** 278–279

Murphy DL, Lerner A, Rudnick G & Lesch K-P (2004) Serotonin transporter: gene, genetic disorders, and pharmacogenetics. *Mol Interv* **4:** 109–123

Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F & Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67:** 355–367

Nicholls DG & Ferguson SJ (2014) Bioenergetics 2. Academic Press.

Nicholls RA, Long F & Murshudov GN (2012) Low-resolution refinement tools in REFMAC5. *Acta Crystallogr D Biol Crystallogr* **68:** 404–417

Oliver RC, Lipfert J, Fox DA, Lo RH, Doniach S & Columbus L (2013) Dependence of Micelle Size and Shape on Detergent Alkyl Chain Length and Head Group. *PLoS ONE* **8:** e62488

Pantazopoulou A & Diallinas G (2007) Fungal nucleobase transporters. *FEMS Microbiol. Rev.* **31:** 657–675

Pellegrini E, Piano D & Bowler MW (2011) Direct cryocooling of naked crystals: are cryoprotection agents always necessary? *Acta Crystallogr D Biol Crystallogr* **67:** 902–906

Perez C, Koshy C, Yildiz Ö & Ziegler C (2013) Alternating-access mechanism in conformationally asymmetric trimers of the betaine transporter BetP. *Nature* **490:** 126–130

Perez J & Koutsioubas A (2015) Memprot: a program to model the detergent corona around a membrane protein based on SEC–SAXS data. *Acta*

Cryst D Biol Crystallogr **71:** 86-93

Petoukhov MV, Franke D, Shkumatov AV, Tria G, Kikhney AG, Gajda M, Gorba C, Mertens HDT, Konarev PV & Svergun DI (2012) New developments in the ATSAS program package for small-angle scattering data analysis. *J Appl Crystallogr* **45:** 342–350

Pérez J & Nishino Y (2012) Advances in X-ray scattering: from solution SAXS to achievements with coherent beams. *Curr. Opin. Struct. Biol.* **22:** 670–678

Pflugrath JW (2004) Macromolecular cryocrystallography--methods for cooling and mounting protein crystals at cryogenic temperatures. *Methods* **34:** 415–423

Porod G (1982) General theory. In *Small-Angle X-ray Scattering* (Glatter OKO, Ed.) London: Academic Press

Privé GG (2007) Detergents for the stabilization and crystallization of membrane proteins. *Methods* **41:** 388–397

Putnam CD, Hammel M, Hura GL & Tainer JA (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q. Rev. Biophys.* **40:** 191–285

Rath A, Glibowicka M, Nadeau VG, Chen G & Deber CM (2009) Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proc. Natl. Acad. Sci. U.S.A.* **106:** 1760–1765

Raven PH, Evert RF & Eichhorn SE (2005) Biology of plants 8th Edition.

Ressl S, Terwisscha van Scheltinga AC, Vonrhein C, Ott V & Ziegler C (2009) Molecular basis of transport and regulation in the Na(+)/betaine symporter BetP. *Nature* **458:** 47–52

Reyes CL, Ward A, Yu J & Chang G (2006) The structures of MsbA: Insight into ABC transporter-mediated multidrug efflux. *FEBS Letters* **580:** 1042–1048

Sanchez-Weatherby J, Bowler MW, Huet J, Gobbo A, Felisaz F, Lavault B, Moya R, Kadlec J, Ravelli RBG & Cipriani F (2009) Improving diffraction by humidity control: a novel device compatible with X-ray beamlines. *Acta Crystallogr D Biol Crystallogr* **65:** 1237–1246

Scapin, G (2013) Molecular Replacement Then and Now. *Acta Crystallogr D Biol Crystallogr* **69**: 2266–75

Schaffner W & Weissmann C (1973) A rapid, sensitive, and specific method for the determination of protein in dilute solution. *Analytical biochemistry* **56:** 502–514

Seddon AM, Curnow P & Booth PJ (2004) Membrane proteins, lipids and

detergents: not just a soap opera. *BBA* **1666:** 105–117

Seidner G, Alvarez MG, Yeh JI, O'Driscoll KR, Klepper J, Stump TS, Wang D, Spinner NB, Birnbaum MJ & De Vivo DC (1998) GLUT-1 deficiency syndrome caused by haploinsufficiency of the blood-brain barrier hexose carrier. *Nat. Genet.* **18:** 188–191

Shi Y (2013) Common folds and transport mechanisms of secondary active transporters. *Annu Rev Biophys* **42:** 51–72

Shimamura T, Weyand S, Beckstein O, Rutherford NG, Hadden JM, Sharples D, Sansom MS, Iwata S, Henderson PJ & Cameron AD (2010) Molecular basis of alternating access membrane transport by the sodium-hydantoin transporter Mhp1. *Science* **328:** 470–473

Shimamura T, Yajima S, Suzuki S, Rutherford N, O'Reilly J, Henderson P & Iwata S (2008) Crystallization of the hydantoin transporter Mhp1 from Microbacterium liquefaciens. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **64:** 1172–1174

Sigworth FJ (1994) Voltage gating of ion channels. *Q. Rev. Biophys.***27**: 1-40

Simmons KJ, Jackson SM, Brueckner F, Patching SG, Beckstein O, Ivanova E, Geng T, Weyand S, Drew D, Lanigan J, Sharples DJ, Sansom MS, Iwata S, Fishwick CW, Johnson AP, Cameron AD & Henderson PJ (2014) Molecular mechanism of ligand recognition by membrane transport protein, Mhp1. *EMBO J.* **33:** 1831–1844

Sjögren T, Carlsson G, Larsson G, Hajdu A, Andersson C, Pettersson H & Hajdu J (2002) Protein crystallography in a vapour stream: data collection, reaction initiation and intermediate trapping in naked hydrated protein crystals. *J Appl Crystallogr* **35:** 113–116

Smyth DR, Mrozkiewicz MK, McGrath WJ, Listwan P & Kobe B (2003) Crystal structures of fusion proteins with large-affinity tags. *Protein Sci.* **12:** 1313–1322

Stojanoff V, Jakoncic J, Oren DA, Nagarajan V, Navarro Poulsen JC, Adams-Cioaba MA, Bergfors T & Sommer MOA (2011) From screen to structure with a harvestable microfluidic device. *Acta Crystallogr F Struct Biol Commun* **67:** 971–975

Stuhrmann HB (1970) New method for determination of surface form and internal structure of dissolved globular proteins from small angle x-ray measurements. *Zeitschrift für Physikalische Chemie* **72:** 177–182

Sun L, Zeng X, Yan C, Sun X, Gong X, Rao Y & Yan N (2012) Crystal structure of a bacterial homologue of glucose transporters GLUT1-4. *Nature* **490:** 361–366

Sutton BJ & Sohi MK (1994) Crystallization of membrane proteins for X-ray analysis. *Methods Mol. Biol.* **27:** 1–18

Suzuki S & Henderson PJ (2006) The hydantoin transport protein from Microbacterium liquefaciens. *J Bacteriol* **188:** 3329–3336

Svensson O, Malbet-Monaco S, Popov A, Nurizzo D & Bowler MW (2015) Fully automatic characterization and data collection from crystals of biological macromolecules. *Acta Crystallogr D Biol Crystallogr* **71:** 1757–1767

Svergun D, Barberato C & Koch MHJ (1995) CRYSOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J Appl Crystallogr* **28:** 768–773

Svergun DI (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.* **76:** 2879–2886

Svergun DI & Koch MH (2003) Small-angle scattering studies of biological macromolecules in solution. *Reports on Progress in Physics* **66:** 1735

Tanford C (1973) The hydrophobic effect: formation of micelles and biological membranes. John Wiley & Sons.

Terstappen GC & Reggiani A (2001) In silico research in drug discovery. *Trends Pharmacol. Sci.* **22:** 23–26

VanAken T, Foxall-VanAken S, Castleman S & Ferguson-Miller S (1986) Alkyl glycoside detergents: synthesis and applications to the study of membrane proteins. *Meth. Enzymol.* **125:** 27–35

Veenstra M, Turk E & Wright EM (2002) A ligand-dependent conformational change of the Na+/galactose cotransporter of Vibrio parahaemolyticus, monitored by tryptophan fluorescence. *J Membr Biol* **185:** 249–255

Vergara A, Lorber B, Sauter C, Giegé R & Zagari A (2005) Lessons from crystals grown in the Advanced Protein Crystallisation Facility for conventional crystallisation applied to structural biology. *Biophysical chemistry* **118:** 102-112

Volkov VV & Svergun DI (2003) Uniqueness of ab initioshape determination in small-angle scattering. *J Appl Crystallogr* **36:** 860–864

Wadsten P, Wöhri AB, Snijder A, Katona G, Gardiner AT, Cogdell RJ, Neutze R & Engström S (2006) Lipidic sponge phase crystallization of membrane proteins. *Journal of Molecular Biology* **364:** 44–53

Wagner S, Bader ML, Drew D & de Gier J-W (2006) Rationalizing membrane protein overexpression. *Trends Biotechnol.* **24:** 364–371

Wallin E & Heijne von G (1998) Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci.* **7:** 1029–1038

Ward A, Sanderson NM, O'Reilly J & Rutherford NG (2000) The amplified

expression, identification, purification, assay, and properties of hexahistidine-tagged bacterial membrane transport proteins (SA Baldwin Ed.) Membrane transport - A Practical Approach. Blackwell's Press, Oxford

Warkentin M & Thorne RE (2007) A general method for hyperquenching protein crystals. *J Struct Funct Genomics* **8:** 141–144

Warkentin M & Thorne RE (2009) Slow cooling and temperature-controlled protein crystallography. *J Struct Funct Genomics* **11:** 85–89

Wasserman SR, Koss JW, Sojitra ST, Morisco LL & Burley SK (2012) Rapid-access, high-throughput synchrotron crystallography for drug discovery. *Trends Pharmacol. Sci.* **33:** 261–267

Watanabe A, Choe S, Chaptal V, Rosenberg JM, Wright EM, Grabe M & Abramson J (2011) The mechanism of sodium and substrate release from the binding pocket of vSGLT. *Nature* **468:** 988–991

Waterman DG, Winter G, M PJ, Luis F-M, Hattne J, Brewster A, Sauter NK & Evans G (2013) The DIALS framework for integration software. *CCP4 Newsl. Protein Crystallogr.* **49:** 16–19

Weyand S, Ma P, Saidijam M, Baldwin J, Beckstein O, Jackson S, Suzuki S, Patching SG, Shimamura T, Sansom MSP, Iwata S, Cameron AD, Baldwin SA & Henderson PJF (2006) The nucleobase-cation-symport-1 family of membrane transport proteins. *Handbook of Metalloproteins* John Wiley & Sons.

Weyand S, Shimamura T, Beckstein O, Sansom MSP, Iwata S, Henderson PJF & Cameron AD (2011) The alternating access mechanism of transport as observed in the sodium-hydantoin transporter Mhp1. *J Synchrotron Radiat* **18:** 20–23

Weyand S, Shimamura T, Yajima S, Suzuki S, Mirza O, Krusong K, Carpenter EP, Rutherford NG, Hadden JM, O'Reilly J, Ma P, Saidijam M, Patching SG, Hope RJ, Norbertczak HT, Roach PCJ, Iwata S, Henderson PJF & Cameron AD (2008) Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter. *Science* **322:** 709–713

Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A & Wilson KS (2011) Overview of the CCP4 suite and current developments. *Acta Cryst D Biol Crystallogr* **67:** 235-242

Winter G, Lobley CMC & Prince SM (2013) Decision making in *xia2*. *Acta Crystallogr D Biol Crystallogr* **69:** 1260–1273

Witz S, Panwar P, Schober M, Deppe J, Pasha FA, Lemieux MJ & Möhlmann T (2014) Structure-function relationship of a plant NCS1 member--homology modeling and mutagenesis identified residues

critical for substrate specificity of PLUTO, a nucleobase transporter from Arabidopsis. *PLoS ONE* **9:** e91343

Wöhri AB, Johansson LC, Wadsten-Hindrichsen P, Wahlgren WY, Fischer G, Horsefield R, Katona G, Nyblom M, Oberg F, Young G, Cogdell RJ, Fraser NJ, Engström S & Neutze R (2008) A lipidic-sponge phase screen for membrane protein crystallization. *Structure/Folding and Design* **16:** 1003–1009

Wright EM, Hirayama BA & Loo DF (2007) Active sugar transport in health and disease. *J Intern Med* **261:** 32–43

Yamashita A, Singh SK, Kawate T, Jin Y & Gouaux E (2005) Crystal structure of a bacterial homologue of Na+/Cl--dependent neurotransmitter transporters. *Nature* **437:** 215–223

Zhao Y, Terry D, Shi L, Weinstein H, Blanchard SC & Javitch JA (2010) Single-molecule dynamics of gating in a neurotransmitter transporter homologue. *Nature* **465:** 188–193

Zhou Z, Zhen J, Karpowich NK, Law CJ, Reith MEA & Wang D-N (2009) Antidepressant specificity of serotonin transporter suggested by three LeuT-SSRI structures. *Nat. Struct. Mol. Biol.* **16:** 652–657

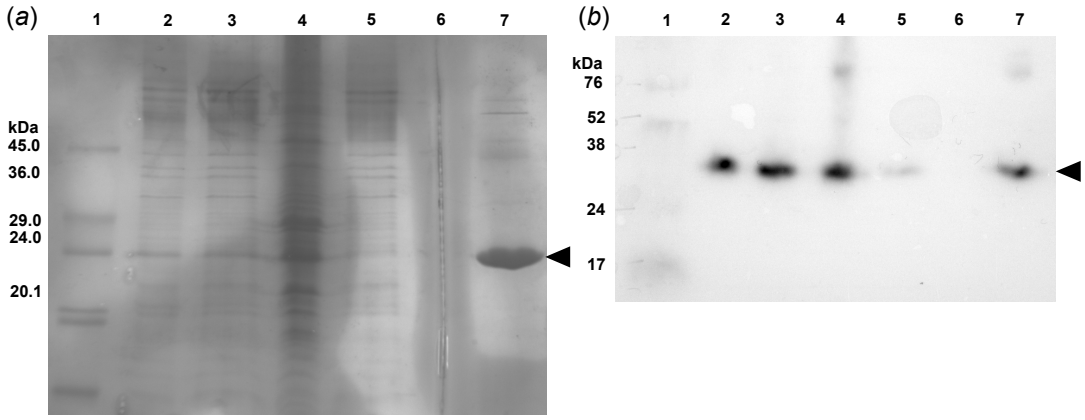
Zimmer J, Doyle DA & Grossmann JG (2006) Structural characterization and pH-induced conformational transition of full-length KcsA. *BPJ* **90:** 1752–1766

Appendix A

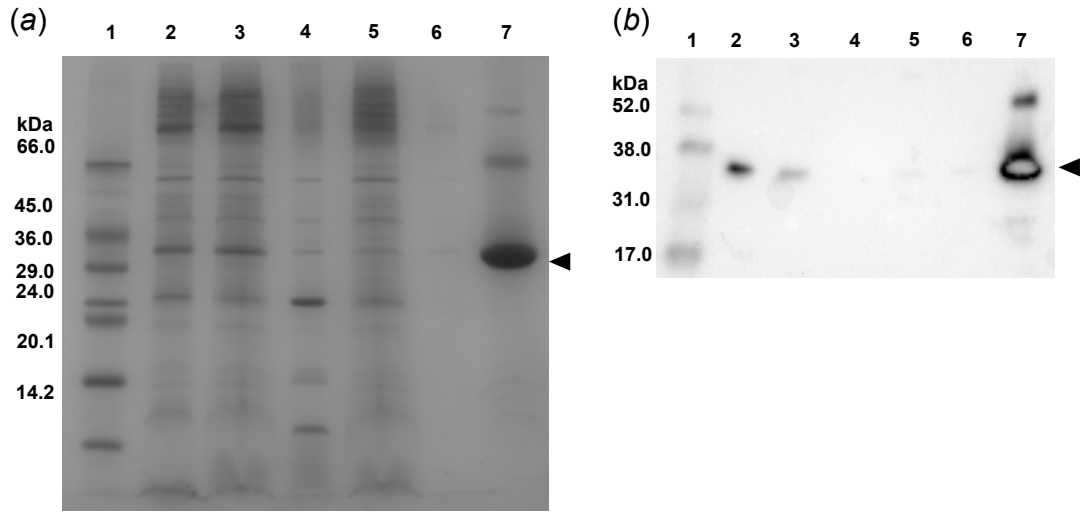
SDS-PAGE gels (*a*) and corresponding Western blots (*b*) of samples from the purification of Mhp1 mutants used to make Figure 3.8.

In all gels, apart from the gels of the purification of Asn318G*, the lanes correspond to 1) solubilisation mixture before ultracentrifugation, 2) soluble fraction containing solubilised Mhp1 protein after ultracentrifugation, 3) pellet after ultracentrifugation, 4) unbound material in the flow-through after binding to the Ni-NTA resin, 5) wash buffer of the Ni-NTA resin, 6) eluted sample from the Ni-NTA resin using 200 mM imidazole. The arrows indicate the position of the migration of Mhp1.

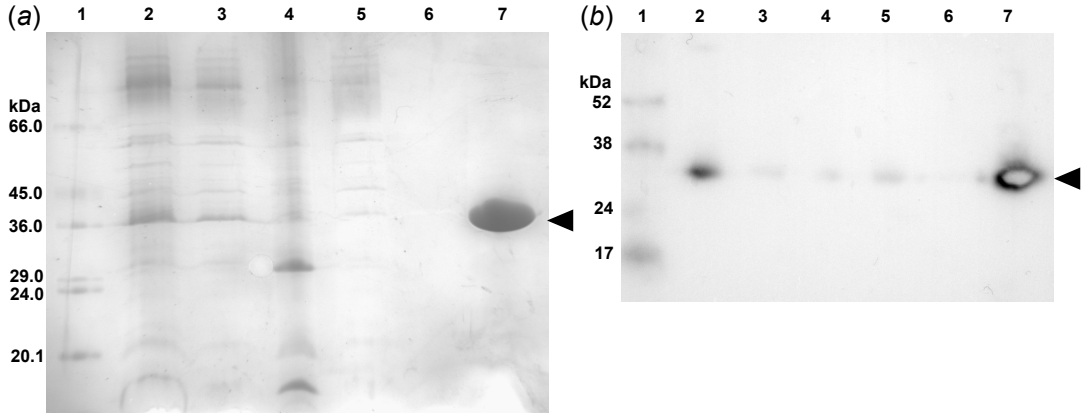
Gln42Asn



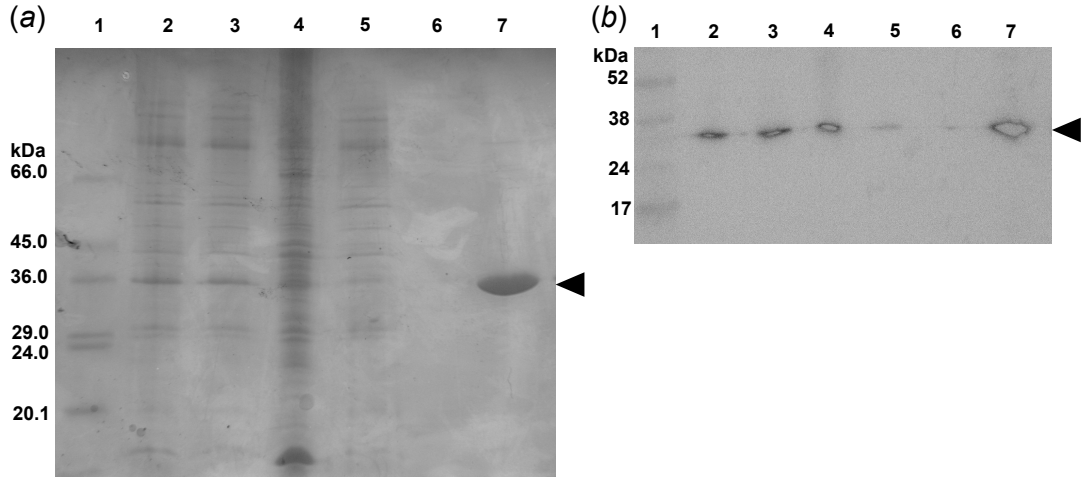
Gln121Asn



Ser312Ala

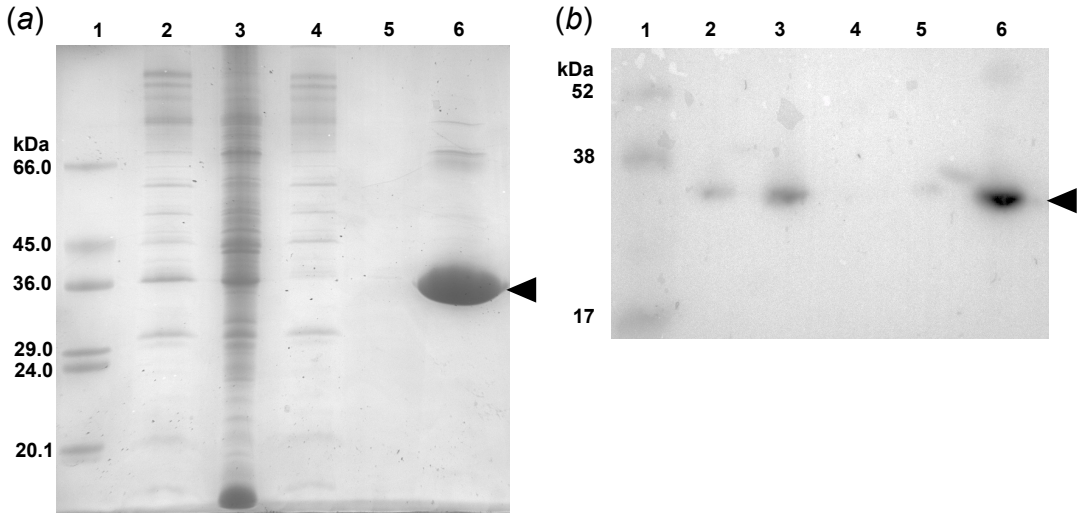


Lys110Leu



In the gel of the purification of Asn318G*, the lanes correspond to 1) soluble fraction containing solubilised Mhp1 protein after ultracentrifugation, 2) pellet after ultracentrifugation, 3) unbound material in the flow-through after binding to the Ni-NTA resin, 4) wash buffer of the Ni-NTA resin, 5) eluted sample from the Ni-NTA resin using 200 mM imidazole.

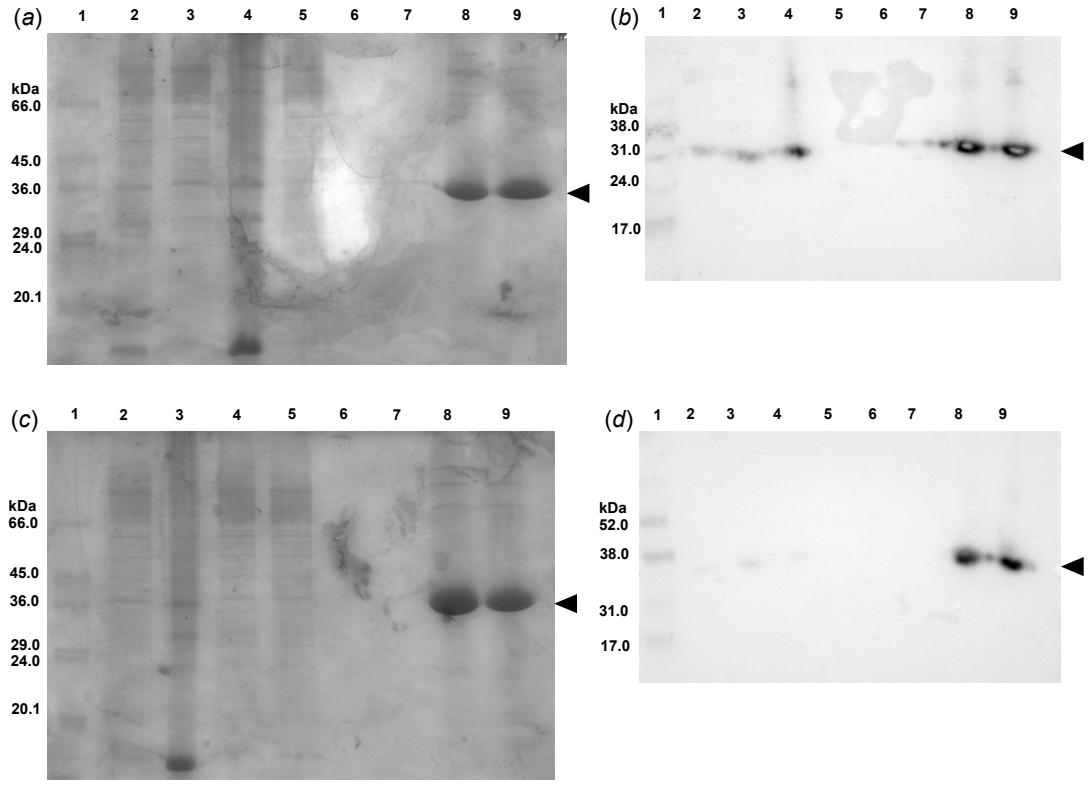
Asn318Gly*



Appendix B

The lanes in the SDS-PAGE gel (*a*) and corresponding Western blot (*b*) for batch 1, used for the purification of wild-type Mhp1 solubilised in DDM and NM, contain 1) molecular weight standard, 2) the solubilisation mixture with DDM, 3) the soluble fraction containing Mhp1 after ultracentrifugation, 4) the pellet after ultracentrifugation, 5) flow-through from the Ni-NTA resin prior to dividing the resin into two parts, 6) the wash buffer containing DDM, 7) the wash buffer containing NM, 8) eluted Mhp1 protein solubilised in DDM, 9) eluted Mhp1 protein solubilised in NM.

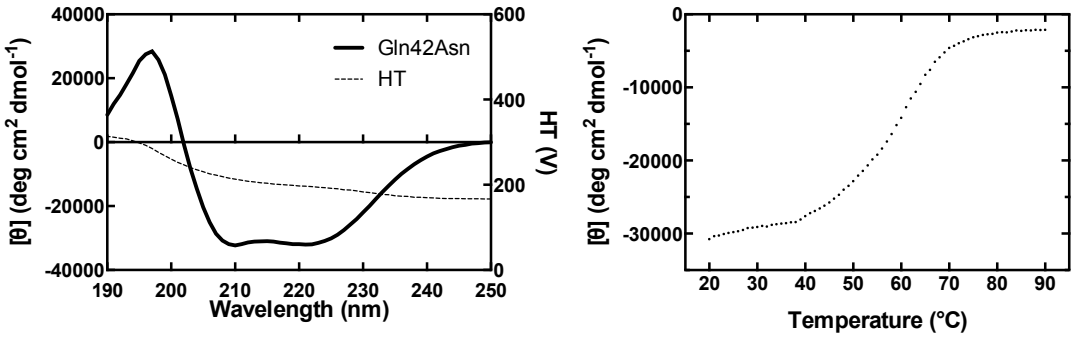
The lanes in the SDS-PAGE gel (*c*) and corresponding Western blot (*d*) for batch 2, used for the purification of wild-type Mhp1 solubilised in UM and DM, contain 1) molecular weight standard, 2) the soluble fraction containing Mhp1 after ultracentrifugation, 3) the pellet after ultracentrifugation, 4-5) flow-through from the Ni-NTA resin prior to dividing the resin into two parts, 6) the wash buffer containing UM, 7) the wash buffer containing DM, 8) eluted Mhp1 protein solubilised in UM, 9) eluted Mhp1 protein solubilised in DM.



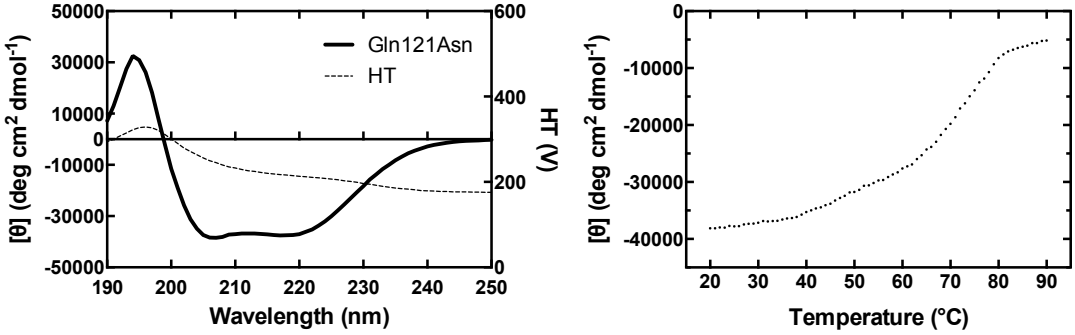
Appendix C

CD graphs and thermal melts used to determine the melting points in section 3.3.2.

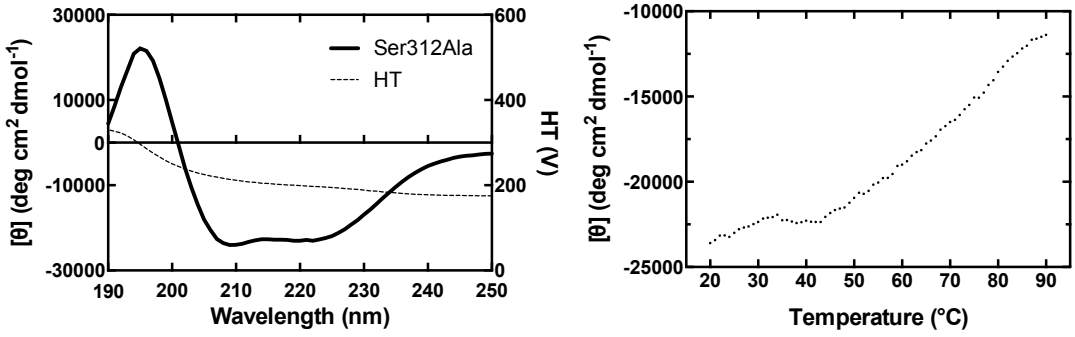
Gln42Asn



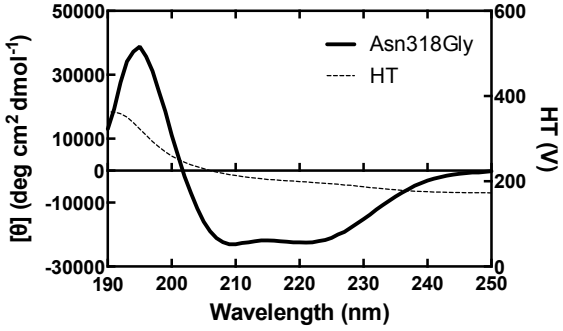
Gln121Asn

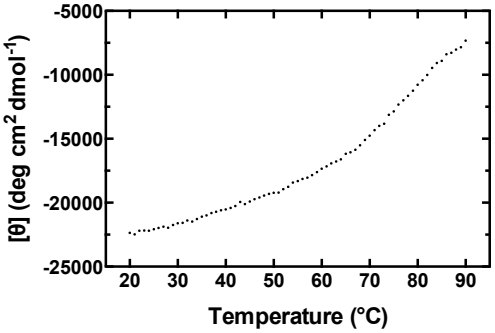


Ser312Ala



Asn318Gly





Lys110Leu