Expression, solubilisation and purification of MRP4

Afroditi Vaitsopoulou
Doctor of Philosophy

Aston University
September 2022

This thesis is dedicated to my family

Maria Vaitsopoulou, Zoi Katana, Adamos Vaitsopoulos and Maria Katana

And my friends

Antriana Ioannidou, Luksa Popovic and Dr Nora Gkanatsiou

For all the support

Also dedicated to the people that inspired me to become a researcher and try for it,

Dr Stella Manta, Dr Athina Dimopoulou, Dr Monica Wagner, Dr François Ferron, Dr Ashleigh Shannon and Dr Ana Theresa Morais

Abbreviations

ABC ATP-binding cassette

AcMNPV Autographa californica multicapsid nucleopolyhedrovirus

ADP adenosine diphosphate

ALL Acute Lymphoblastic Leukemia

AML Acute Myeloid Leukemia

APS ammonium persulfate

ATP adenosine triphosphate

BSA bovine serum albumin

BSEP bile-salt export pump

CALX8 5,11,17-tris[(carboxy)methyl]-25-monooctyloxy-26,27,28-trihydroxycalix[4]arene

cAMP cyclic adenosine monophosphate

cASIC1 chicken acid-sensing ion channel 1a

CFTR cystic fibrosis transmembrane conductance regulator

cGMP cyclic guanosine monophosphate

Chol cholesterol

CHS cholesterol hemisuccinate

Cmc critical micelle concentration

Cmt critical micellar temperature

CMV cytomegalovirus

Cryo-EM cryo-electron microscopy

DC Dendritic Cell

DDAO dodecyldimethyl-N-amineoxide

DDG N-(2-methyl-1,3-bis(O-β-D-Glucose)propan-2-yl)-3-(decylthio)propanamide

DDM n-dodecyl-h-d-maltoside

DDTAC Dodecylmercapto-S-(poly(tris(hydroxymethyl)acrylamidomethane) DPn=6

DIBMA diisobutylene maleic acid

DMSO dimethyl sulfoxide

DNA deoxyribonucleic acid

ECF energy coupling factor

ER endoplasmic reticulum

GFP green fluorescent protein

GluCl glutamate-gated chloride channel

GPCR G protein-coupled receptor

GPI glycosyl-phosphatidylinositol

GPL glycerophospholipids

GSL glycosphingolipid

HDL high density lipoprotein

HRP horseradish peroxidase

ICL intracytoplasmic loop

Ld liquid-disordered

LMNG Lauryl Maltose Neopentyl Glycol

Lo liquid-ordered

LSB Laemmli sample buffer

MATE multidrug and toxin extrusion

MDR multidrug resistance

MFS major facilitator superfamily

MNG neopentyl glycols

MOI multiplicity of infection

mRNA messenger ribonucleic acid

MRP multidrug resistance protein

MSD membrane-spanning domains

MYCN N-myc proto-oncogene protein

NBD nucleotide binding domain

NSAID Nonsteroidal Anti-inflammatory Drug

NTA Nitriloacetic acid

OG n-octyl-h-d-glucopyranoside

PBS phosphate buffered saline

PDBTM Protein Data Bank of Transmembrane Proteins

PGE2 Prostaglandin E2

P-gp P-glycoprotein

PIC protease inhibitor cocktail

PL phospholipids

PtdCho Phosphatidylcholine

PtdEtn phosphatidylethanolamine

PtdIns phosphatidylinositol

PtdSer phosphatidylserine

PVDF polyvinylidene fluoride

RND resistance-nodulation-division

RTK Receptor type tyrosine kinase

SBP substrate binding protein

SDS sodium dodecyl sulfate

SDS-PAGE SDS-polyacrylamide gel electrophoresis

SEC secretase translocon

SL sphingolipids

SM sphingomyelin

SMA Styrene maleic acid

SMALP SMA Lipid Particle

SMI styrene maleimide

Sph sphingosine

SR signal recognition particle receptor

SRP signal recognition

TBS tris buffered saline

TBS-T TBS with tween

TCEP tris(2-carboxyethyl)phosphine

TEMED tetramethylethylenediamine

TM transmembrane

TMD transmembrane domain

V Volt

Contents

1. Introduction

1.1 Membranes

Biological membranes are the barrier between cells or organelles such as mitochondria, and their external environment. They can also serve to compartmentalize intracellular organelles like endoplasmic reticulum (ER) and Golgi (13). Membranes consist of various molecules: lipids, proteins, sugars. Lipids form a bilayer where proteins are embedded, controlling what can pass through while sugars are found only on the one part of the bilayer and bind the lipids with covalent bonds (14). The different components of the membrane interact with each other and various other elements and influence the membrane plane. Protein-protein interactions and glycoprotein interactions can play a role in signalling and activation of multiple procedures as well as in creating special formations that facilitate cell-cell interactions, such as junction or adherent features (15). Moreover, membrane proteins interact with the nearby lipids by hydrophobic matching (hydrophobic parts of proteins embedded in hydrophobic parts of the lipid bilayer) and electrostatic interactions between amino acids and phospholipids. The hydrophobic match especially can be essential for the formation of specialized lipid domains or rafts where enrichment in cholesterol and sphingolipids occurs (11). Another property important in lipid–protein interactions is the propensity of some lipids to induce curvature stress and the ability of certain membrane peripheral proteins to overcome this stress (16). However, the binding of integral proteins (membrane proteins that span the lipid bilayer) to particular lipids can shift the conformation of nearby integral proteins. On the other hand, binding of peripheral membrane proteins directly to the lipid head groups can alter lipid curvature (17). Lipid-lipid interactions can play another role in membranes as they can be distributed unevenly and dynamically and change the fluidity, the dynamics and the lateral structures of membranes. An example can be cholesterol that is important for the formation of membrane domains (figure 1.1A). Furthermore, the occlusion of cell material to lipid rafts or domains can lead to endocytosis, signal transduction, cell death or other processes (18). The lipids in these domains can exchange rapidly with lipids in the bulk

fluid membrane or other rafts so the domain or raft environment becomes very dynamic. Extracellular restrictions, formation of specialized domains such as lipid rafts can restrict the lateral mobility of integral membrane proteins. These proteins themselves, can be involved in the dynamic formation of nano-sized domains by initiating cis interactions at the extracellular surface resulting in macromolecular complexes (11). Interactions between receptors and ligands can be cis or trans depending if the ligand is expressed on the same or a different cell respectively (19). Lipids usually are the most abundant component of membranes. Due to their amphiphilic nature, they tend to form bilayers and micelles that segregate the internal and external cell environment and assist to the compartmentalization of the cells and the internal organelles. Plus, cell division, biological reproduction and intracellular membrane trafficking are based on lipids as they are essential for budding, tubulation, fission and fusion of membranes. Another function of lipids is signalling. They can be first and second messengers (e.g. when they are disrupted they generate signalling elements) or define the membrane domains where proteins are recruited and will initiate the pathway cascade (20). The lipid bilayer is about 30 Å thick and the interface in each side is about 15 Å. Lipids vary in tail length, stiffness and size as well as in electrical charge of the head group, for example steroidic lipids like cholesterol are more rigid and offer a stability. The different lipids in different domains of the membrane make it adopt different curvature, thickness, fluidity or pressure. Moreover, it provides variety on organisms according to the lipid choice in each case (4). For example, the major lipids in eukaryotic membranes are the glycerophospholipids (GPL). Phospholipids (PL) constitute the trunk from where the membrane lipids derive. They are similar to the triglycerides in being ester or amide derivatives of glycerol (glycerophospholipids, GPL) or sphingosine (Sph) forming sphingolipids (SL) with fatty acids and phosphoric acid. They are composed of two fatty acids linked through two hydrophobic acyl chains and a phosphate head group ester linked to a glycerol. In GPL, the phosphate moiety of the resulting phosphatidic acid is further esterified with choline, ethanolamine, serine or inositol in the phospholipid. Phosphatidylcholine (PtdCho) and phosphatidylethanolamine (PtdEtn) heads are zwitterionic, whereas phosphatidylserine (PtdSer) and phosphatidylinositol (PtdIns) heads are anionic

(figure 1.1B). The second lipids that are more abundant in membranes are sphingolipids. SL contain one hydrophobic acyl chains and a phosphate head group ester linked to a Sph backbone. Their hydrophobic backbone is an ester or amide derivative of Sph with fatty acids. Some common SLs are glycosphingolipids (GSLs). Cholesterol (Chol) is an important element of cell membranes. Chol molecule is composed of four fused rings (steroid backbone) showing a hydroxyl group (A-ring) and a small branched hydrophobic tail (D-ring). The steroid backbone of Chol facilitates its interaction with SL. Chol-SL are the main components of lipid rafts. Chol content is strictly controlled by three mechanisms: de novo synthesis, uptake and esterification (20).

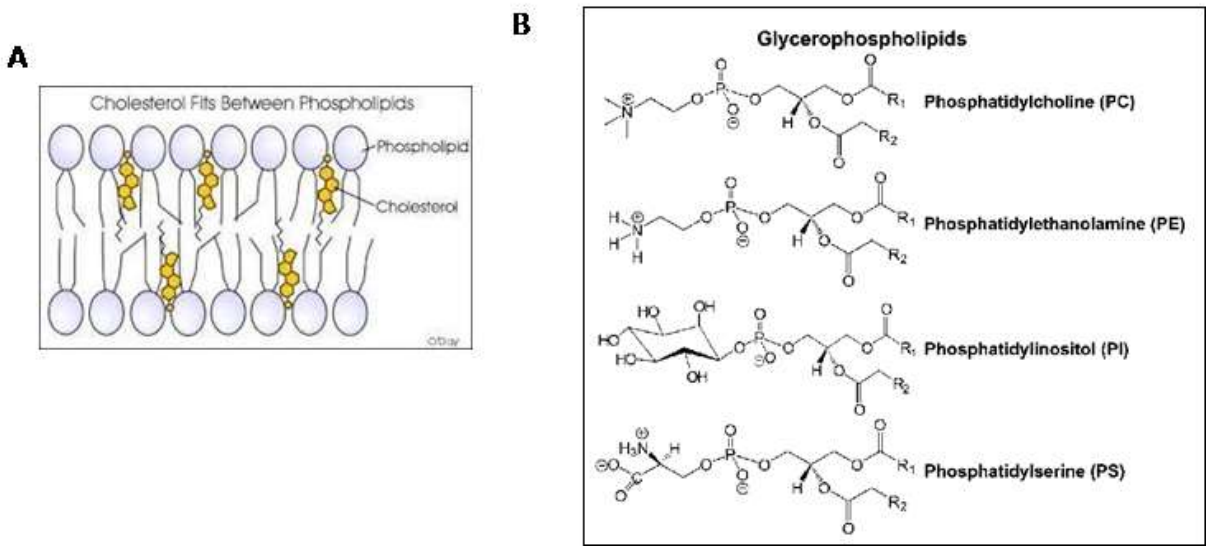


Figure 1.1. Bilayer and lipid structures. *A; Cholesterol (yellow) between phospholipids of the lipid bilayer, adapted from https://biology4ibdp.weebly.com/13-membrane-structure.html (visited 25 May 2022, at 13:50). B; Structures of glycerophospholipids, the main lipids found in eukaryotic membranes (7), adapted, (permission to use this figure asked and taken)*

The various membrane functions depend on the lipid composition and different cells vary in lipid content. The membranes have microdomains with specific properties defined by the type of lipids and proteins as well as by the abundance of these specific elements. Lipid rafts are rich in sphingomyelin (SM), GSLs and Chol which arrange into rigid and liquid-ordered (Lo) microdomains . On the other hand, the small polar heads of PtdEtn provide PtdEtn-rich regions with more fluidity and less surface packing density and pressure and these regions are called liquid-disordered (Ld) microdomains. Because of the heterogeneous lipid distribution, the membrane microdomains can be transient like lipid rafts or stable such as tight junctions. Plus, the two leaflets of the plasma membrane have different lipid content: the outer leaflet is composed mainly by PtdCho and SLs whereas the inner involves PtdEtn, the negatively charged PtdSer, and PtdIns (21). The nature of lipid head-group and the length and saturation of their acyl chains define their category and the structural features of the membrane such as viscosity, curvature, electrostatic charge, fluidity, compactness, and connect with proteins to assist their function, but membrane protein abundance and composition also play a role in the physicochemical properties of membranes. For example, PtdCho involves one cis-unsaturated acyl chain that decreases the packing compactness of the acyl chains and increases membrane fluidity. Moreover, due to its cylindrical shape PtdCho is capable of self-assembling spontaneously. On the other hand, PtdEtn adopts a conical shape because of the small head group which gives to the membrane a negative curvature stress (figure 1.2).

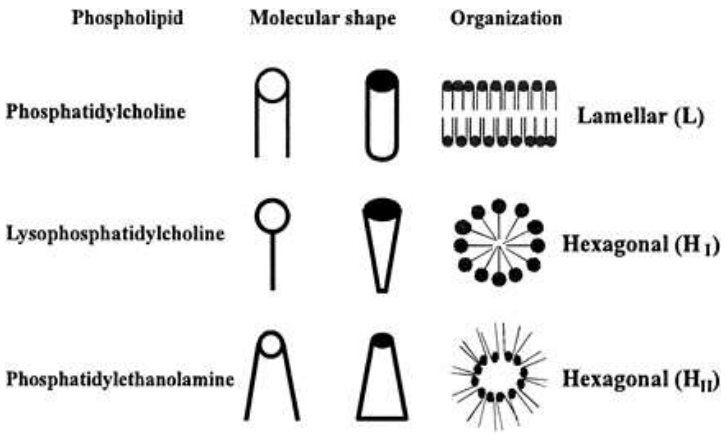


Figure 1.2. Shapes of different phospholipids. *Phosphatidylcholine adopts a cylindrical shape while phosphatidylethanolamine has a conical shape. These differences give a variety of curvatures and stress in the membrane (10), adapted (permission for the figure is not required as long as it is properly cited)*

Furthermore, due to the acyl chain unsaturation of some lipid types, lipid-packing defects are created so membrane fusion is facilitated, and binding and activity of peripheral membrane proteins are influenced. Although PtdSer and PtdEtns are less abundant, they are essential for charging the membrane surface and consequently they connect with membrane protein (both integral and peripheral) charged parts, influencing their function (20). The model that was proposed for the membrane and remains the accepted model, is the one of the fluid-mosaic, suggested by Singer and Nicolson in 1972. According to this, lipids self-assemble and create a lipid bilayer, where proteins are embedded and there is a motility of lipids and proteins. Proteins were expected to appear as globular structures according to the means they had to study them at the time (figure 1.3A). This way, the concept of hydrophobicity was introduced as the hydrophobic parts of the proteins were in the bilayer while the hydrophilic ones were in touch with solution (11). In 1976, Nicolson suggested other features that affect lateral movements and distributions of plasma membrane components, such as extracellular matrix and membrane associated cytoskeleton (15). The motility of cell membrane components can be influenced by these elements as well as cell-cell and cell-matrix interactions.

Later, other findings introduced new data on membranes, lipids and membrane proteins as well as new concepts like lipid rafts. Another feature of the membranes is asymmetry: flip-flop of lipids, different lipid preference, and membrane proteins that are asymmetrically distributed in the bilayer, create this essential asymmetry that differentiates the extracellular and intracellular environment of the cells. Moreover, a number of lipid transporters are needed to maintain lipid asymmetry, such as cytofacially-directed, adenosine triphosphate (ATP)-dependent transporters ('flippases') and exofacially-directed, ATP-independent transporters ('floppases'), but there are also bidirectional, ATP-independent transporters ('scramblases'). The existence of several of these phospholipid transporters in maintaining the proper phospholipid asymmetries in the cell membrane suggests that maintenance of membrane asymmetry is functionally essential for cells. Furthermore, membranes show curvature, deformation, expansion and compression in some cases. It was proved that many membrane glycoproteins and proteins, were not completely free to move in the plane of the membrane. This is cis-membrane control (within the plane of the membrane) or trans-membrane control (across the membrane). In the end, many changes have occurred since 1972, biology and technology have progressed, hence, the initial model was modified and enriched. A revised model is more up to date with all the new data that were collected later (figure 1.3B). In this case, it should include the different types of interactions of integral membrane proteins, glycoproteins, lipids, cytoskeletal systems and extracellular matrix components. These interactions control mobility, distribution and aggregation states of membrane components. Changes in the interactions of the membrane components determine functional properties because the cell membrane is a dynamic structure (11).

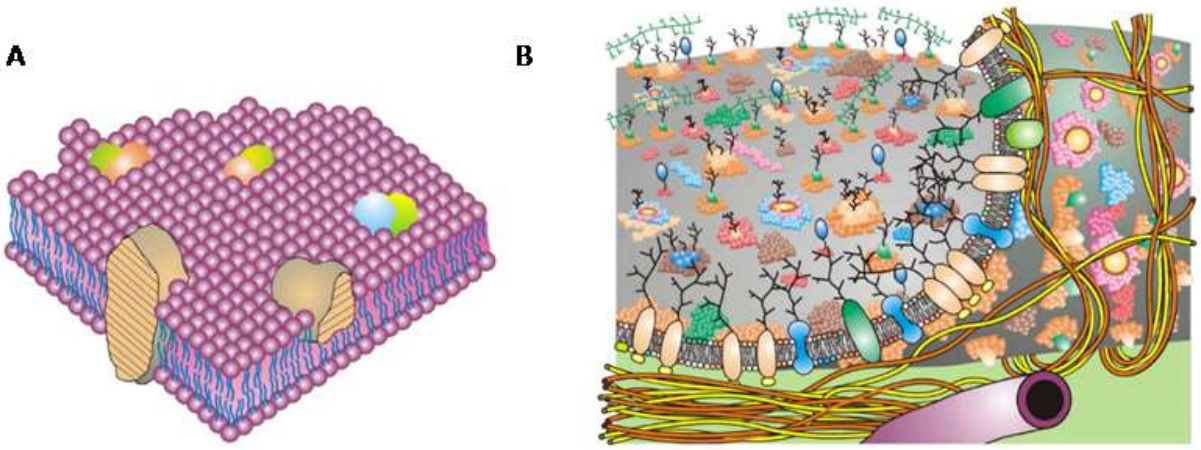


Figure 1.3. Membrane models. *A; The initial model of the fluid mosaic presented by Singer and Nicolson in 1972. According to this, the lipid bilayer (purple) is like fluid, and membrane proteins are spanning it and seem to be globular. It looks like a mosaic of lipids and proteins. B; A revised model of the fluid mosaic, containing other features that were discovered later, such as cytoskeleton and extracellular matrix, microdomains, glycoproteins. Here, the membrane can be seen full of proteins (presented in different colours) some of which are embedded in domains like lipid rafts or liquid-ordered microdomains, polysaccharide and glycoprotein associations are shown as well. The membrane is "peeled-up" to show more clearly the lipid bilayer and how the proteins span it. The structures that look like wires (orange and yellow) represent the cytoskeleton (11), adapted (permission is not required as long as cited properly).*

1.2 Membrane Proteins

Proteins are divided into membrane and soluble proteins. The environment of these categories differs as soluble proteins are hydrophilic and surrounded by liquid, and membrane proteins are embedded in the lipid leaflet so they have hydrophilic parts in touch with solution and hydrophobic parts that span the membrane and this difference makes them adopt different structures. Membrane proteins are classified as integral or peripheral (figure 1.4A). The integral ones span the membrane whereas the peripheral ones are attached to other membrane proteins or lipids or through a covalent anchor, with Wan der Waals or electrostatic interactions. Furthermore, membrane proteins are further separated in α-helical or β-barrels (figure 1.4B). α-helical proteins are the most abundant in membranes and they can span it with a single or multiple helices that are mainly hydrophobic and about 20 residues long. 20-30% of the membrane proteins genes encode for this kind of proteins. β-barrel proteins consist of β-sheets that form a structure that resembles a barrel, hence their name. In this case, there is a sequence of residues that makes them have their hydrophobic residues in touch with the lipid bilayer, and the hydrophilic residues to create a pore that water or other soluble molecules can cross (4).

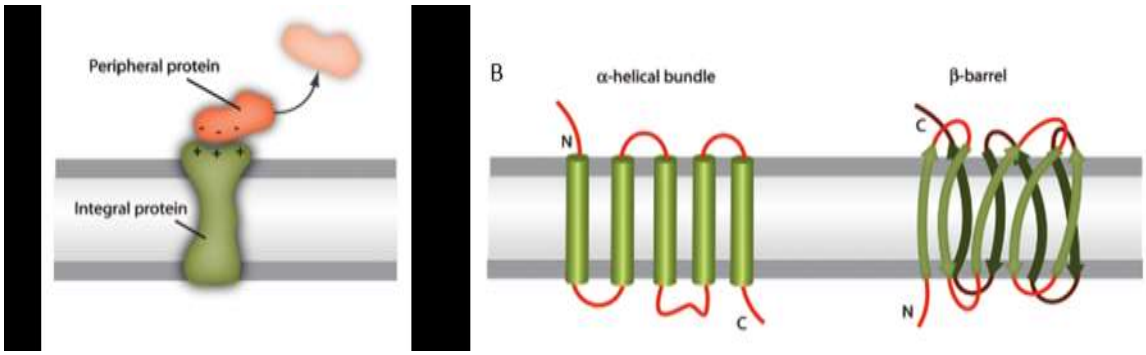


Figure 1.4. Membrane protein classes. *A; Integral (green) and peripheral (orange) protein representantion. B; Helical membrane protein where the α-helices span the lipid bilayer and a β-barrel protein (4), adapted (permission asked and taken for figure use).*

Integral membrane proteins are classified as type 1, type 2 and multipass and they comprise 10%, 10% and 11-12% of the genome respectively. Type 1 proteins have their N-terminal in the extracellular space while type 2 have it oriented to the cytoplasm. The multipass ones can have either one or both N-terminal and C-terminal facing the intracellular or extracellular space. The peripheral membrane proteins do not span the membrane but associate with it by interaction with either integral proteins or lipids of the bilayer. When peripheral proteins interact with phospholipids, they create an anchor via covalent bonds such as glycosyl-phosphatidylinositol (GPI). GPI-anchored proteins lack a transmembrane domain, they do not have a cytoplasmic tail, and are located exclusively on the extracellular side of the plasma membrane. These proteins include membrane-associated enzymes, adhesion molecules, activation antigens, and differentiation markers. Two examples are carbonic anhydrase type IV and alkaline phosphatase. Another group of peripheral membrane proteins are attached to the cytosolic face of the lipid bilayer by a hydrocarbon moiety, such as prenyl, farnesyl, and geranylgeranyl groups, covalently attached to a cysteine near the protein C-terminus. An example of proteins that are anchored via prenylation, is Ras, which plays a central role in the development of cancer. The third class of peripheral proteins that are lipid-anchored, are those which have fatty acyl groups like myristate or palmitate on their N-terminal glycine or cycsteine residue linking them to the membrane. These anchors are necessary for protein function (13). One example of this protein type is the oncogene v-Src tyrosine kinase which can only transform cells when it retains its myristylated N-terminus(22).

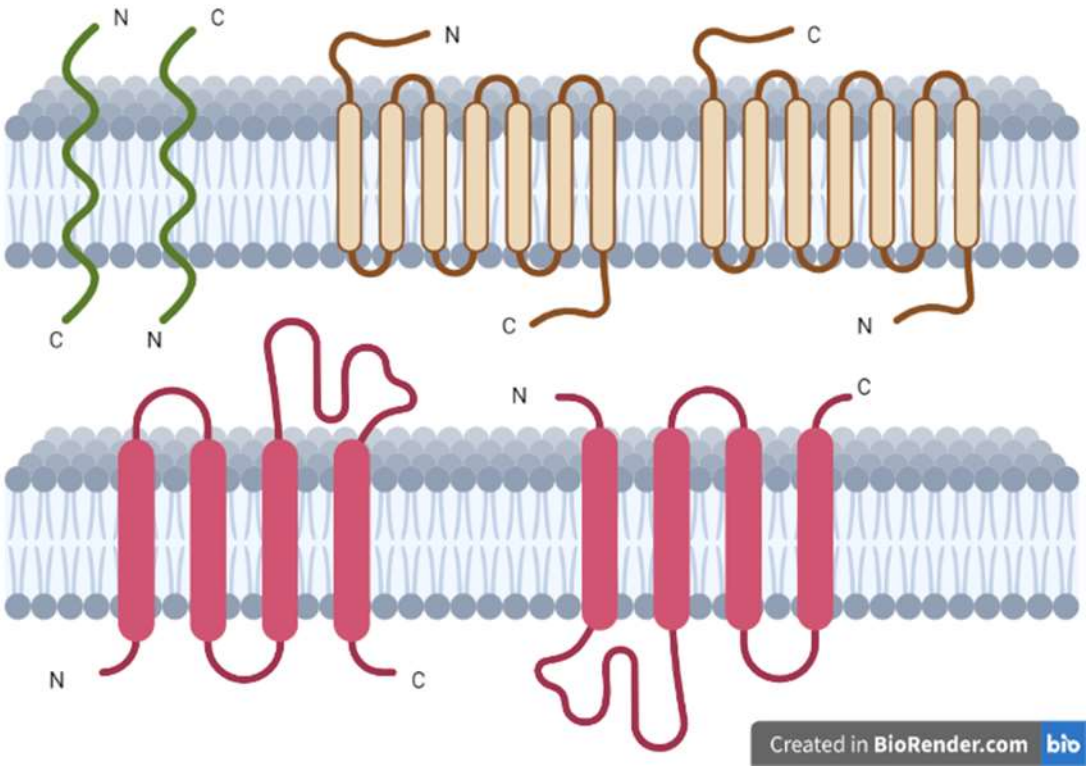


Figure 1.5 Integral membrane proteins. *Type 1 have the N-terminus at the extracellular space and the C-terminus facing the intracellular space while type 2 have the N-terminus facing inside the cells and C-terminus facing outside at the extracellular space. Then, the multispanning ones can have either one or both termini at the same or at the other side of the membrane. (4).*

Membrane proteins have multiple roles for the cells as they can transport ions, metabolites, molecules like proteins or RNA or they can send and receive signals so they assist to the response in different cues from the environment in order for the cell to adapt to the message. Others participate in cell-cell attachment and interaction or there are even proteins that act as anchors for the extracellular matrix or cytoskeletal proteins and retain the cellular shape (4, 13). Other functions are regulation of intracellular vesicular transport, membrane lipid composition control and organization and shape of organelles and cells. This is the main reason why membrane proteins attract pharmacological attention. The biogenesis of membrane proteins is different than the one of soluble proteins. Most of

the proteins are synthetized in the cytoplasm by ribosomes that translate the messenger ribonucleic acid (mRNA) into amino acids and elongate the polypeptide chain. Soluble proteins fold as they emerge from the ribosome. In the case of membrane proteins they need to be translocated and integrated into the lipid bilayer. The system that governs these procedures is the SRP-dependent SEC pathway. SRP stand for signal recognition particle and SEC stands for the secretase translocon. These two components, with SR (signal recognition particle receptor) are the main elements of the pathway. At first, the translocon is targeted. As the polypeptide chain is generated, highly hydrophobic parts are recognised from the SRP. In eukaryotes, this stops the polypeptide chain elongation until the ribosome comes in touch with the translocon. Then, the SRP binds to its receptor in the ER for eukaryotes or the plasma membrane for prokaryotes where the transfer of the ribosome-nascent chain-complex to the translocon occurs. A tunnel is formed form the translocon pore and the ribosome and the polypeptide chain is threaded directly through the channel as it is being synthesized. Translation, translocation or membrane insertion happen simultaneously so the process is called co-translational. When the signal peptide reaches the translocon it either reorients, placing the N-terminal in the cytosol, in an N-in orientation or it does not so there is the N-out orientation (figure 1.6). The parts of the polypeptide chain that are hydrophobic are laterally inserted in the membrane. The protein then needs to fold in the membrane. During this procedure, there might be minor or major rearrangements (4).

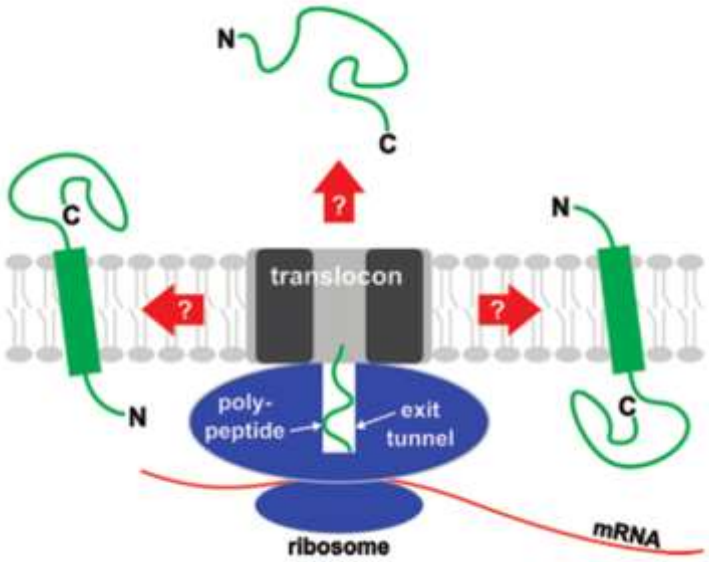


Figure 1.6. The translocon and ribosome tunnel. *As the polypeptide is generated, it passes directly in the tunnel to be properly directed to the membrane (4), adapted (permission asked and taken for the use of this figure).*

Soluble proteins need just to hide the hydrophobic parts and face the solution with the hydrophilic parts. However, membrane proteins have to deal with three different environments: the hydrophobic chains of the lipid bilayer, the solution into the cytoplasm or outside and an interface region rich of phospholipids head groups. As a result of these multiple environments, transmembrane proteins need to expose different types of residues at different locations. However, the interior of transmembrane proteins could, at least for stability, remain identical to what is found in the interior of the soluble ones. The hydrophobic effect is the major force that drives the folding of soluble proteins whereas for membrane proteins this is not the main driving force. In this case, there are some motifs that allow packing of the helices. Short chained-residues are placed in such a way that they end up on the same side of their helix, that is, in positions i and i+4 or in i and i+7, creating a surface that allows close proximity. Glycine zippers occur in 10% of membrane proteins and favour a right-handed crossing of the helices. The helices interact with Wan der Waals bonds or hydrogen bonds. Residues that interact with hydrogen bonds are more conserved than buried residues and the packing preferences of

different classes of proteins differ because of variety in amino acid preferences (4). Initially, it was believed that α-helical membrane proteins form simple helix bundles and cross the membrane in perpendicular orientation. This applies to some of them but later results showed that the transmembrane helices can be short, long, kinked or interrupted in the middle of the membrane (coils), almost perpendicular to the membrane plane, strongly tilted, laying flat on the surface of the membrane (interfacial helices), or span only a part of the membrane and then turn back as in the case of aquaporin (23). In addition, some helices are packed in manners that do not follow the simple "knobs-into-holes" geometry. One example with such deviations is the glutamate transporter homologue. Regarding the amino acid preferences, hydrophobic residues are found in the core of the lipid bilayer while tyrosine and tryptophan are mainly found in the polar head group (interface) regions whereas outside of the membrane the amino acid distribution is similar to the one of the soluble proteins. Tyrosine and Tryptophan are more abundant about 15-20 Å away from the membrane core, within the interface region. This is more common in β-barrel proteins because of steric and electrochemical factors. In the interface region, the polar groups of these amino acids can interact with the phosphate groups of lipids, whereas the hydrophobic rings can interact with the lipid chains. Tyrosine and tryptophan are also preferred in proximity to the phospholipid headgroups. Residues outside the membrane interface region, point the polar groups inwards, while the ones located outside point the polar groups outwards. The aromatic and hydrophobic phenylalanine on the other hand, is not biased toward the interface region. Charged and polar amino acids are not favourable energetically within the bilayer core but they can be tolerated towards the termini of the transmembrane helices. The polar groups of the long side chains in arginine and lysine can orient themselves so that the polar groups approach the interfacial and aqueous regions. This way they create polar microenvironments for themselves by pulling hydrating waters into the hydrocarbon part of the lipid bilayer. A long chain facilitates snorkelling so arginine is more abundant close to the membrane core than lysine (4). Lysines are often found near the membrane interface in membrane proteins. It has been suggested that the long and flexible side chain of lysine can stretch out of the

membrane interior to have the charged amino group in the more polar interface region while keeping the hydrocarbon part of the side chain inside the hydrophobic part of the membrane. This is called 'snorkelling' and could be important for the precise location of transmembrane parts of integral membrane proteins in the lipid bilayer, and for interactions between proteins and lipids in general (24). Polar amino acids can also be found in the core of the lipid bilayer when ion binding or water filled channels lining is needed. Positively charged or basic amino acids are more abundant to the cytosolic side of the membrane. In contrast, acidic residues do not show any preference for loops or one side of the membrane. α-helical membrane proteins mainly consist of a core of 20 residues long helices that stop near the phospholipid headgroups making this region rich in coil residues. In a coil, the peptide backbone forms no regular secondary structure, hence it exposes its polar backbone. Especially in the case of channels and transporters, this occurs frequently and is essential for function. In transmembrane helices, coil residues are thought to provide structural flexibility, and facilitate interactions by positioning side chains properly. These amino acids are conserved, hence, they are important for function. Polar residues and backbone coil residues are more abundant among channels and transporters (4).

Membrane proteins are categorized in receptors, transporters, enzymes and miscellaneous. Beginning with receptors, they mediate cellular responses upon binding of a ligand, facilitating transmission of information. Some important groups are G protein-coupled receptors (GPCRs), Receptor type tyrosine kinases (RTKs), Receptors of the immunoglobulin superfamily and related, and Scavenger receptors and related. Passing to the transporters, they facilitate the transport of a substrate across membranes by utilizing electrochemical gradients or energy from chemical reactions or even facilitative diffusion. Some main groups are active transporters, channels, uniporters and solute carriers. Then, the enzymes catalyse chemical reactions. Main groups of membrane enzymes are Oxidoreductases, Transferases, Hydrolases, Lyases, Isomerases and Ligases. In the end, there are some groups that cannot be attributed to any of these three main classes but they can be categorized into Ligands, Structural/Adhesion proteins, Proteins of unknown function, and "other". Some important membrane

protein groups are GPCRs, Aquaporins, Ion channels, ATPases. (25). The different functions of membrane proteins, require different protein architectures. As proteins evolved, they gained the strategy to create larger units from smaller ones. This strategy has some perks: creation of new protein surfaces capable of binding to different molecules, enhanced stability, conformational stability of multimeric complexes, and the potential to support cooperative or other regulatory mechanisms. The three main strategies are oligomerization, gene fusion and internal repeats (26). Oligomerization is the most simple and most common mechanism by which larger proteins are formed by assembly into homooligomers or heterooligomers. According to the Protein Data Bank of Transmembrane Proteins (PDBTM), almost 65% of transmembrane proteins are obligate oligomers (27). The second strategy, is gene fusion. Approximately 30% of the transmembrane proteins contain multiple transmembrane domains that function independently (28). The third strategy is internal repeats. Internal repeats stands for duplication of a structural element/sequence motif within a single polypeptide chain. Half of the proteins that contain internal repeats are symmetric. However, site mutations occur, changing the sequence, rendering them not completely symmetrical (29, 30). Thus, in case there is no specific functional reason to maintain perfect internal symmetry, the primary sequences of the repeats might diverge, resulting in structures that are internally pseudosymmetric rather than symmetric.

Approximately 35% of monomeric proteins cannot adopt oligomeric symmetry. Also, 57-82% of individual domains do not have internal repeats and lack symmetry. These assymetric proteins are appropriate for detecting environmental changes across the membrane. Lack of symmetry is a common feature of GPCRs, ligand-gated ion channels, and enzyme-linked receptors like RTKs but they show a structural symmetry around a perpendicular axis of the membrane (26). Some examples of proteins of nonsymmetric architectures are YidC (31) and TatC (32), and the cellulose synthesis and translocation system, BcsA/BcsB complex (33). Moreover, the P-type ATPases, which constitute one of the largest families of primary transporters, are nonsymmetric (34). Such a lack of symmetry is unusual among transport proteins. The vast majority of membrane proteins show a cyclic symmetry. Most of the membrane proteins show a rotational symmetry about a perpendicular axis to the lipid

bilayer. This suggests that the N-termini and C-termini of all chains are located on the same side of the membrane possibly to facilitate insertion. The simplest symmetric arrangement, and among the most common in membrane proteins, involves a 180◦ rotation around an axis perpendicular to the membrane, C2. Ideal C2 symmetry is found in homooligomeric complexes, whereas C2 pseudosymmetry is found both within heterooligomeric complexes and between internal repeats. In cases when this association is essential for function, the symmetric elements usually create a binding site or pathway at their interface. Examples are the signalling receptors RTKs, that create a ligand binding site at the dimer interface and ABC transporters which have both a substrate pathway and ATP binding sites at the dimer interface. Typically, ABC transporters contain two transmembrane domains (TMDs) and two nucleotide binding domains (NBDs), which are assembled either from separate chains or from fused domains. The NBDs create two off-axis binding sites for ATP using a head-to-tail arrangement, whereas the substrate pathway typically follows the symmetry axis (26). Perpendicular pseudo-C2-symmetry is seen in some secondary transporters for instance the major facilitator superfamily (MFS) that contain two lobes of six transmembrane (TM) helices, each lining a central pathway (35). Furthermore, the multidrug and toxin extrusion (MATE) transporters, such as NorM, and the resistance-nodulation-division (RND) transporters, like AcrB, also contain two domains of six TM helices lining a central pathway (36, 37). Nevertheless, the topological arrangements of the helices differ between these folds. Pseudosymmetry is less common in α-helical complexes of higher order even though some channels are composed of repeated elements, and other channels and receptors are known to form heteromers. In some heteromeric membrane protein complexes subdomains with various symmetries can be found (26).

1.3 Membrane proteins and lipid interactions

Proteins embedded in membranes interact with lipids. The lipids can be boundary/annular lipids, non-annular lipids or bulk lipids. The annular lipids interact with the protein non-specifically and form a lipid shell around the proteins. The non-annular lipids act as co-factors and often bind between transmembrane a-helices either within a protein or at protein–protein interfaces in multi-subunit proteins. The lipids that are not in contact with the protein are called bulk lipids. The lipid-protein interaction is not so strong though and this can be implied by the rate of exchange of lipid molecules between the annular shell and the bulk phase. Many parameters can play a role in the interactions of lipids and membrane proteins as well as the effect of these interactions. One of these is the lipid headgroup region. It is not of high importance which lipid is in the annular shell but it is significant that it is in the appropriate physical state and have a particular effect on the protein (38). Binding at non-annular sites is often head group specific as the head group along with the fatty acid chain play a role in the packing of the lipid headgroup in a bilayer (39). The structure of the lipid headgroup can affect the protein they are in touch with, by creating hydrogen bonds with the polypeptide backbone of polar residues and facilitating the formation of secondary structures like α-helix or β-sheet. This is why the interface region is considered as a catalyst for secondary structure formation of polypeptide chains. The pattern of hydrogen bonding between peptide, water and lipid varies in bilayers of phosphatidylcholine and of phosphatidylethanolamine, potentially leading to different secondary structures for any stretch of peptide located in the lipid headgroup region (40). A charged lipid headgroup can also affect the structure of regions of a protein with charged amino acids located in the headgroup region. The structure of the lipid headgroup region could also play an important role on the structures adopted by the ends of transmembrane α-helices. The initial four –NH and final four C=O groups of an α-helix have no hydrogen bond partners within the backbone of the α-helix itself (41). In membrane proteins, ends of α-helices could also form hydrogen bonds with the glycerol backbone and lipid headgroup regions of the bilayer. Changes in the lipid headgroup region could then lead to changes in structure at the ends of the transmembrane α-helices and changes in the packing

of the helices. By altering the concentrations of charged molecules or ions near the surface of the lipid bilayer, the lipid headgroup can determine the activity of a membrane protein (38). Gouy–Chapman theory gives a good description about the effects of the charge of lipids: incorporation of negatively charged lipids into a membrane will change the apparent affinities of membrane proteins for charged substrates (42). However, interpretation of these effects may need to take into account the effects of membrane charge on the local concentration of H^+. An example of this could be the fact that negatively charged lipids can increase the affinity for Ca^{2+} but this negative charge would also increase the presence of H^+ near the membrane so it would affect the binding of Ca^{2+} as H^+ tend also to bind so the concentration of Ca^{2+} near the membrane actually would be lower than expected. An important role is played by the thickness of the hydrophobic core of the bilayer, generally taken, for glycerophospholipids, to correspond to the separation between the glycerol backbone regions on the two sides of the bilayer (38). The hydrophobic thickness of the lipid bilayer needs to match the one of the protein that spans the membrane because of the high cost of exposing either fatty acyl chains or hydrophobic amino acids to water. Any mismatch could cause distortion of the lipid bilayer, or the protein, or both, to minimize it. Extreme mismatch could also lead to the formation of non-bilayer phases by the lipids, particularly at low molar ratios of lipid to protein (43). The majority of models of hydrophobic mismatch suggest that fatty acyl chains near a membrane protein adjust their length to match the hydrophobic thickness of the protein where the protein acts as a rigid body. When the hydrophobic thickness of the lipid bilayer is lower than the one of the protein, the lipid chains will stretch to provide a thicker bilayer. Conversely, when the hydrophobic thickness of the bilayer is greater than that of the protein, the lipid chains will compress to provide a thinner bilayer (38). Even though any distortion of the membrane around a membrane protein to offer hydrophobic matching seems to be small, the efficiency of hydrophobic matching of the lipid bilayer and the protein seems to be high. If full hydrophobic matching of the protein and the lipid bilayer is not provided, the protein should distort to match the hydrophobic thickness of the bilayer (44). Although α-helices are quite unlikely to distort and change their packing or tilt, the rotation of side chains about the Ca–Cβ bond

linking the side chain to the polypeptide backbone is possible. For a residue at the end of a helix such a rotation would change the effective length of the helix. For example, rotation of a Tyr residue to lie roughly parallel to the long axis of a helix would extend the length of the helix by about 3 Å and rotation of the larger tryptophan residue would have an even larger effect. Effects of hydrophobic matching on protein structure must be highly cooperative. Although, lipid binding constants do not change significantly with modification of fatty acyl chain length, small differences in the free energy of binding of a fraction of a kJ mol^{-1} for any one lipid will become important when summed over the large number of lipids in contact with the protein in the bilayer. However, any mismatch would not affect activity of all the proteins embedded in the bilayer but it has a different effect in each different protein, depending on the protein structure and the effect that has a structural change in the protein activity (38). Protein aggregation is another significant factor. Unfavourable interactions between a membrane protein and the surrounding lipid bilayer can be reduced by protein aggregation to reduce the lipid exposed surface area of the protein, a procedure which depends on the protein shape. If the extramembranous domains of the protein are small, aggregation will cause contact of the transmembrane domains leading to displacement of lipid from the surface of the transmembrane domains. However, if the extramembranous domains of the protein are large, contact between the extramembranous domains would prevent the transmembrane domains from being into contact (45). To understand the energetics between lipid-lipid, lipid-helix and helix-helix interactions, the effects of lipid structure on the aggregation of transmembrane proteins can be studied to give an explanation.

Moreover, membrane viscosity is another factor to consider. Changes in shape of the proteins need to occur in order for them to carry out their function. Resistance to movements in a liquid are described as viscosity and the opposite is fluidity. In a lipid bilayer, the resistance to motion mainly arises from the lipid fatty acyl chains. The balance between the frictional forces and the mechanical restoring forces acting on a group will determine the type of motion adopted by the group. Large-scale motions in a protein, such as hinge-bending motions, will be overdamped because of the large surface displacements involved and the consequent strong frictional effects of the solvent (38). The

stability of a lipid bilayer is determined importantly by the balance between the hydrophobic interactions, which tend to decrease the interfacial area between the lipids and water, and the inter- and intra- molecular interactions, including headgroup hydration, which give rise to a net repulsion, tending to increase the surface area of the membrane (39). For a lipid monolayer to stay flat, the pressures must be in balance across the monolayer. If the lateral pressure in the chain region becomes greater than that between the headgroups, the monolayer will curl towards the aqueous region. This is defined as a negative curvature (the direction of monolayer curvature is defined from the point of view of an observer in the fatty acyl chain region of the bilayer looking out towards the lipid headgroup region); a concave curvature for the monolayer is defined as being positive and a convex curvature is defined as being negative. Conversely, if the lateral pressure between the headgroups becomes greater than that between the chains, the monolayer will curl towards the chain regions, a positive curvature. The tendency to curl becomes frustrated in a lipid bilayer. Some membrane protein activities depend on non-annular lipids. Most of these lipids are bound between transmembrane α-helices, often between subunits in multi-subunit proteins. Non-annular lipid molecules located at protein–protein interfaces could help to ensure good packing at the interface. In some cases, the effect of these lipids in the protein function can occur by affecting the movement of transmembrane helices in the protein. Non-annular binding sites for cholesterol are also located between transmembrane α-helices. The presence of cholesterol has significant effects on the activities of a number of other membrane proteins. Other ways that lipids affect membrane protein activities are the lipid free volume, the lateral pressure and for some membrane protein types, the membrane tension (38).

1.4 ABC transporters

Membrane proteins are important drug targets and 30% of the genome codes for them (46). ATP-binding cassette (ABC) transporters are a large and diverse superfamily of membrane proteins that exist in all organisms, from bacteria to higher eukaryotes. They are classified in two categories: importers, only found in bacteria with a role of nutrient uptake, and exporters found in all organisms. There are two types of importers, I and II. Type I importers play a role in medium-affinity uptake of various nutrients including ions, sugars, amino acids, short peptides, and oligosaccharides, into bacteria. Type II importers take part in high-affinity uptake pathways for metal chelates including heme, other iron-containing complexes and cobalamin. These substrates are larger and more hydrophobic than those of type I importers, but available at low concentrations only (47). There is another type of importers, the energy coupling factor (ECF) transporters, or also called type III importers (48). Those transporters facilitate micronutrient uptake and tryptophan transport in prokaryotes and archaea (49). Exporters, or type IV and type V transporters, translocate multiple substances such as toxins and drugs from the cytoplasm out of the cell.

ABC transporters have four core domains: two NBDs, that bind and hydrolyse ATP and two TMDs that are composed of α-helices that form a pathway where the substrates bind and are transported (figure 1.7). Most importers have an additional substrate binding protein (SBP) that binds the substrates and delivers them to the specific TMD for transport (50). Type III importers lack the SBP domain though (49). The intracytoplasmic loops (ICLs) which are the cytosolic regions of the TMDs, form the interface between the TMDs and NBDs and coordinate ATP binding and hydrolysis with substrate binding and translocation (51). These four domains, the two TMDs and the two NBDs, are sometimes encoded as separate polypeptides and sometimes fused into multidomain proteins. Most of the times, the importers are characterized by having the TMD and ABC (NBD) on separate polypeptides. Exporters in prokaryotes are half-transporters with a single polypeptide containing one TMD and one ABC domain, with the formation of homo- or, in some cases, heterodimers, whereas in eukaryotes, many ABC transporters have non-identical halves fused in the single molecule (52).

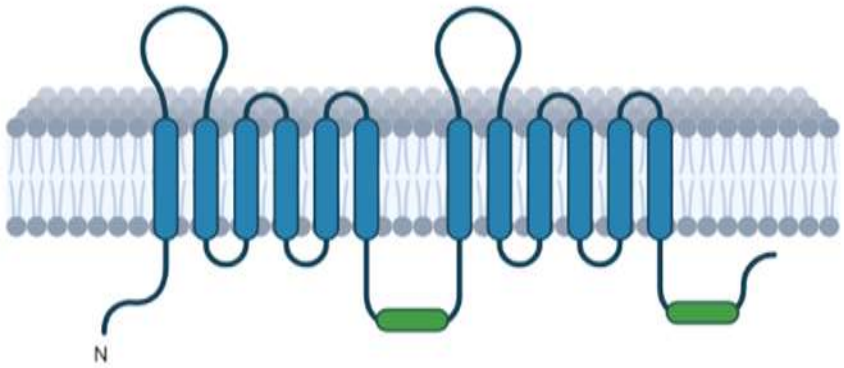


Figure 1.7 Topology of an ABC exporter. *The N-terminus is in the intracellular space as well as the C-terminus. In blue are shown the helices of the TMDs and in green the NBDs. The polypeptide chain begins with TMD1 near the N terminus, followed by NBD1 and then TMD2 followed by NBD2.*

Several ABC transporters also have additional accessory domains. These domains are extracytoplasmic domains (such as in members of the ABCA subfamily, and the transporter ABCG2), membrane embedded domains (such as in some uptake systems and some members of the ABCC subfamily), cytosolic regulatory domains (such as in cystic fibrosis transmembrane conductance regulator CFTR/ABCC7), and accessory catalytic domains (ABC transporters in some secretion pathways) (53).

The TMD amino-acid sequence is not well conserved among the various ABC transporters whereas the NBD sequence is highly conserved. NBDs have five important motifs that bind ATP: Walker A (P-loop), Walker B, Q-loop, D-loop, H-loop and the ABC signature motif (54) (50). The catalytic core domain contains conserved motifs for the binding and hydrolysis of ATP; the Walker A motif interacts with the phosphate groups of the nucleotide and the Walker B motif has a glutamate residue that acts as a general base to activate a water molecule for nucleophilic attack at the γ-phosphate of ATP. The α-helical domain contains the ABC-signature motif that is involved in the binding of the nucleotide. The NBDs have a 'head-to-tail' conformation that makes them form the ATP-binding and hydrolysis site upon dimerization through interactions with the D-loop (50). Although these are the common features

of ABC transporters with some differences, there are seven categories of structural folds of the ABC transporters, from type I to type VII. Type I having the simplest fold, it has two TMD composed of 5 transmembrane α-helices each. Type II transporters are similar but with 10 α-helices in each of the two TMDs. Those two types of transporters, I and II, they both have a SBP. Although type III transporters are also importers like type I and II, they can be considered as non canonical since they have two TMD that are unrelated, called S and T and consist of 6 and 4-8 α-helices respectively. The S component binds the substrate and transports it to the T component which is connected to the NBDs and this way the uptake from the periplasm to the cell happens. Passing to the exporters, there are the type IV and type V. Type IV exist in both prokaryotes and eukaryotes and their two TMDs consist of 6 α-helices each. Type V transporters also have 2 TMDS of 6 α-helices each, but the arrangement and positioning within the lipid bilayer is distinct. Type V have so far only been found in eukaryotes. Type VI and type VII transporters are the most recently discovered, and have quite distinct structures and mechanisms of action (figure 1.8) (55). This classification was based on the one suggested by Thomas et al, according to not only sequence alignments on NBDs but also on TMDs as well as structural features of the different crystal and cryo-electron microscopy (cryo-EM) structures resolved (56).

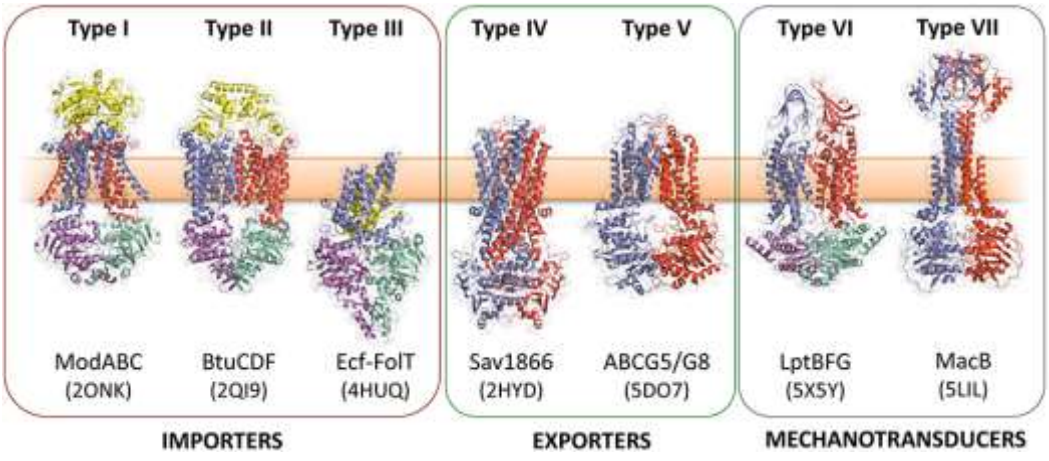


Figure 1.8. The different folds of the ABC transporters. *Type I-III belong to the importers while IV-V to the exporters (8), adapted (permission to use the figure if cited properly)*

The mechanism of function of the ABC transporters remains unclear. There are three models of transport suggested, the switch model, the alternating sites model and the constant contact model. According to the switch model, the NBDs are initially in an open configuration. Two ATP molecules bind cooperatively to the NBDs and it is proposed that the one site has different affinity for the ATP from the other (stochastic process). Then, the NBDs adopt a closed configuration and they hydrolyse ATP to adenosine diphosphate (ADP) and Pi. The hydrolysis products remain bound and after configuration change of the TMDs, NBDs come to the open configuration and the products are released. Concerning the alternating sites model, ATP molecules bind loosely to the open NBDs causing closing of the dimer. Only the one ATP molecule is then tightly bound and committed to hydrolysis. ADP and Pi are consequently released and the dimer is again in open conformation waiting for another ATP molecule to bind while another one is still bound. It is supposed that the NBDs have some kind of "memory" of the site where the hydrolysis occurred so that next time hydrolysis occurs to the other site. Finally, regarding the constant contact model, for each possible state of the NBD active site there are two distinct substates, either occluded (closed) or open (allowing nucleotide exchange), or in the case of the empty site, high or low affinity for nucleotide. Each active site cycles in the sequence: ATP-open, ATP-occluded, ADP + Pi -occluded, ADP-open, empty-low affinity, empty-high affinity. So, the ATP bound active site is in a closed conformation while the other site is empty. Hydrolysis occurs and ATP binds with high affinity to the opposite site while the products remain to the occluded site. Pi is released promoting ADP bound site to open and ATP bound site to close. ADP is then released, promoting hydrolysis of the ATP of the opposite site. ATP hydrolysis enables high affinity binding of ATP at the opposite, empty site. ATP binding and Pi release promote opening of the ADP-bound site and closing of the ATP-bound site. ADP release completes the cycle (Figure 1.9) (51)

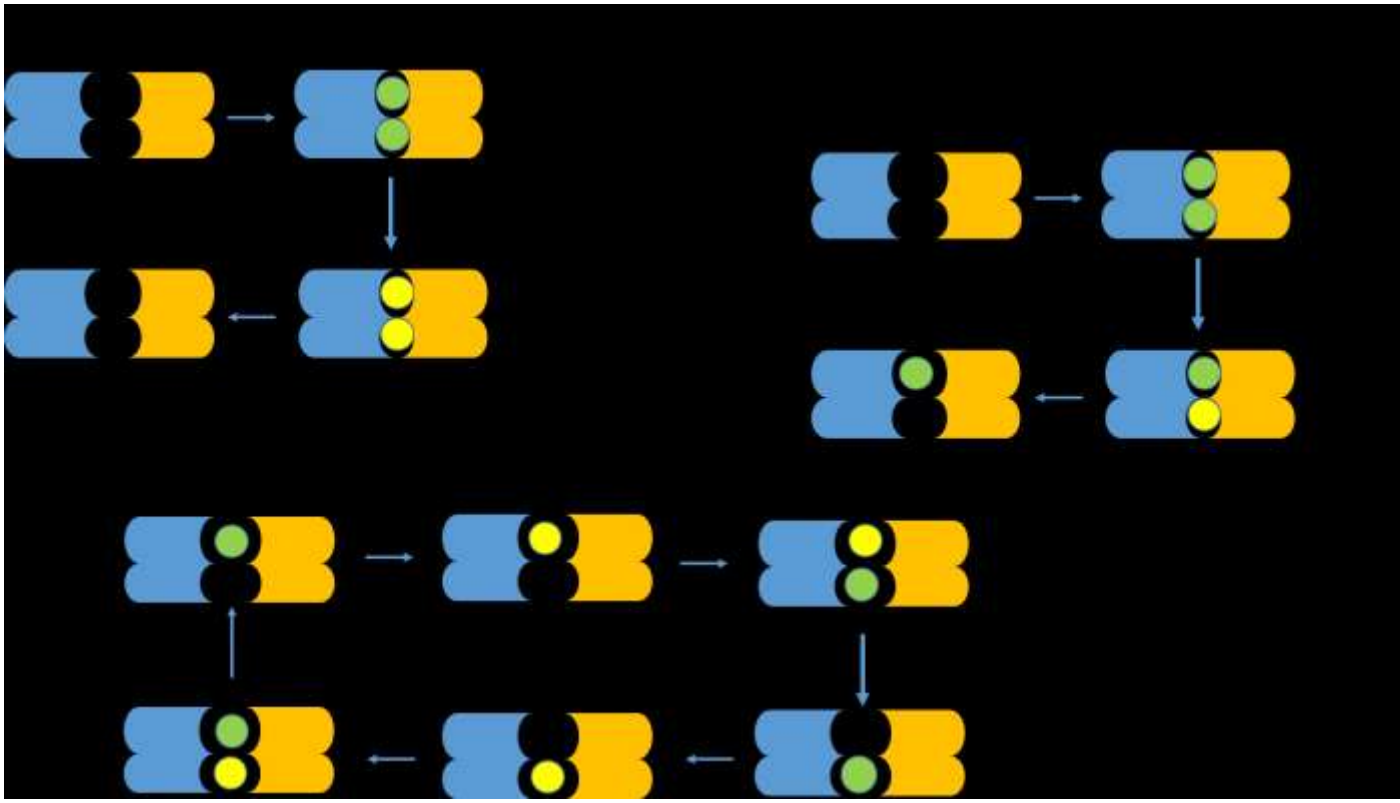


Figure 1.9. *The models of ATP binding and hydrolysis of the NBDs. NBDs are shown in blue and orange, the ATP is in green and ADP is in yellow.*

1.5 Human ABC transporters

Humans have 48 members of the ABC superfamily which are divided in 7 subfamilies (A-G) (57). ABC transporters are expressed in different types of tissues and have various roles. They are responsible for the clearance of toxic substances, including drugs such as chemotherapeutic agents, antivirals, antibiotics and cardiovascular drugs. This is why they play a role in cancer as they facilitate the efflux of the drugs from the cells rendering therapy insufficient. However, by exporting toxic substances, they assist for example the release of toxins to the urine. Epithelial cells in mammals, use ABC transporters to excrete endogenous metabolites such as bile salts or lipids (58). Mutations in various

ABC transporters cause different diseases, such as mutation in the ABCC7/CFTR gene, a chloride ion channel, that leads to cystic fibrosis (59) or mutation of ABCC2 that leads to Dubin–Johnson syndrome, a human disorder of organic ion transport (60). ABC transporters are expressed in many tissues with various functions. Some of the most studied human examples include ABCB1, ABCC1, ABCG2, ABCA1, ABCB11 and ABCG5/G8. ABCB1 (P-glycoprotein, P-gp/Multidrug resistance MDR1) is expressed in the intestinal epithelium, hepatocytes, and the epithelial cells of the blood–brain barrier. It transports a diverse array of compounds including lipids, xenobiotics, drugs (colchicine), chemotherapeutic agents (doxorubicin, etoposide), and bilirubin (61). It is overexpressed in multidrug resistant tumour cell lines (59) and its negative correlation with response to chemotherapy in small cell lung cancer has been shown (62). It is preferentially expressed in poorly differentiated colon tumours, but expression is undetectable in normal colon tissue (63). Then, ABCB1 is also highly expressed in hematopoietic stem cells, where it seems to serve to protect these cells from toxins (60). Also, it can help protecting relatively sensitive and critical tissues from xenotoxic compounds such as the brain and fetal–maternal barrier in the placenta from toxic effects (64). ABCC1 (Multidrug resistance protein MRP1/multidrug resistance protein 1) has been reported to play a role in protecting cells from chemical toxicity as well as oxidative stress and to mediate inflammatory responses involving cysteinyl leukotrienes (60). However, in primary untreated hepatocellular carcinoma, ABCC1 levels depend on differentiation grade, tumour size and degree of microvascular invasion. Elevated ABCC1 levels have been found to be associated with poor patient outcome in breast cancer, non-small cell lung cancer, Acute Myeloid Leukemia (AML) and Acute Lymphoblastic Leukemia (ALL) (63). Moreover, overexpression of ABCC1 in neuroblastoma cells enhanced migratory ability and colony formation (62). ABCG2 expression has been associated mostly with poorer outcomes in AML as well as a range of other cancers, including ALL, breast and lung cancer and its expression is frequently observed in tumour cell populations with increased self-renewal capacity and tumorigenic potential (65). ABCA1 has been reported to exhibit a protective role against cardiovascular disease and it is required for lipid and cholesterol homeostasis (62). It is expressed in hepatocytes, intestinal enterocytes, and adipocytes and it is involved in the

generation of high density lipoprotein (HDL) particles, whereas on macrophages it is involved in the reverse cholesterol transport pathway since it transfers membrane phospholipids and cholesterol to lipid-poor apolipoproteins (61). However, ABCA1 contributes to the malignant characteristics of ovarian cancer cells and it might have a pro-malignant role in prostate cancer (62). ABCA2 plays a role, which remains unclear, in Alzheimer's disease (66). ABCB11 or bile-salt export pump (BSEP) is the major determinant of bile formation and bile flow in humans. Finally, ABCG5 and ABCG8 are expressed in the liver, and small intestine, where they function to excrete sterols into bile and limit intestinal absorption of sterols (61).

The human ABC transporters are all type IV or type V exporters, meaning that structurally they have four core domains: two NBDs, that bind and hydrolyse ATP and two TMDs that are composed of α-helices that form a pathway where the substrates bind and are transported. However, many of them have additional domains, such as MRP1 that has 5 additional transmembrane helices, termed TMD0 at the N-terminal, or the R regulatory domain of CFTR (54).

1.6 MRP4

MRP4/multidrug resistance protein 4 or ABCC4 is an ABC transporter of the C subfamily. It has a typical ABC transporter structure of two membrane-spanning domains (MSDs), each consisting of six TM helices with two cytosolic ATP-binding domains, where ATP is bound and hydrolysed to facilitate substrate transport (67). It is known as one of the 'short MRPs' because it does not have the additional TMD0 domain at the N-terminus seen in several family members such as ABCC1. Thus, it is a single polypeptide chain which forms an MSD1-NBD1-MSD2-NBD2 structure (68). The amino acids in TM6 of MRP4 are conserved among different species, but not among other members of the MRP subfamily (67). MRP4 is involved in the active transport of organic anions, cyclic Adenosine Monophosphate (cAMP) and cyclic Guanosine Monophosphate (cGMP), various drugs, mainly nucleoside analogues,

antibiotics, antineoplastic agents, antivirals, and signaling molecules. Antibiotics include cephalosporins (ceftizoxime, cefazolin, cefmetazole, and cefotaxime) (69, 70). Then, chemotherapeutic agents include 6-mercaptopurine and 6-thioguanine used for leukemia treatment (71) and dasatinib (chronic myeloid leukemia (CML) and acute lymphoblastic leukemia (ALL) drug) (72) and topotecan (73). Antivirals include adefovir, tenofovir, ganciclovir (67). MRP4 also transports cardiovascular drugs (loop diuretics, thiazides, angiotensin II receptor antagonists) (67) and bile acids (74). It is also a physiological transporter of endogenous substrates such as prostaglandins whose accumulation may influence cyclic nucleotide homeostasis (75). Some of the MRP4 inhibitors are probenecid, dipyridamole, indomethacin, sildenafil, and MK571 (76) as well as common Nonsteroidal Anti-inflammatory Drugs (NSAIDs) like indomethacin, indoprofen, ketoprofen, and flurbiprofen (77). Other important inhibitors such as ceefourins have been developed (78). MRP4 has a particularity regarding the other members of the C subfamily: its dual membrane localization in polarized cell types. In prostate tubuloacinar cells, hepatocytes, gastric epithelium, and the brain blood barrier, MRP4 is localised to the basolateral membrane. However, it is localised at the apical side of renal proximal tubule cells and the luminal side of brain capillary endothelium. At the enterocyte epithelium, MRP4 is localised to both the basolateral and apical membranes (79). Its function is not clear yet and it depends on the tissue where it is expressed. It serves to efflux xenobiotics, toxins and drugs out of the cells. It is expressed in prostate, urogenital tissues, kidney proximal tubules, astrocytes and capillary endothelial cells of the brain, platelets, erythrocytes, hepatocytes, in human neutrophilic polymorphonuclear leukocytes (80), in basolateral membrane of enterocytes (70), in endometrium (81), in smooth muscle cells, cardiomyocytes, bone cells, fibroblasts, and cancer cells such as leukemia cell lines, lung cancer, pancreatic cancer, and neuroblastoma (79). In kidneys, at renal proximal tubule cells, MRP4 exports substrates into urine and, consequently, it plays a role in xenobiotic transport to the bladder (82). In brain, it assists in the efflux of xenobiotics from blood capillaries to the plasma, in liver it transports reduced glutathione and bile salts (83). MRP4 in the endometrium facilitates the transport of Prostaglandin E2 (PGE2), a key player in embryo implantation (81). In cardiomyocytes,

MRP4 exports cAMP into the extracellular space and this is a fundamental paracrine factor in the myocardium (84). MRP4 seems to play a role in inflammation since its inhibition results in reduction of release of pro-inflammatory factors from activated endothelial cells (75). MRP4 also contributes to human skin Dendritic Cell (DC) migration (85). Moreover, it is essential for the regulation of fibroblast migration by alterating cyclic nucleotide levels (86). In addition, MRP4 actively transports cAMP and cGMP through platelet membranes and thus it has an important role in platelet cAMP homeostasis, controlled by its synthesis and degradation, but also by its compartmentalization in subcellular compartments (84). Concerning cancer, MRP4 expression generally makes the cancer cells resistant to chemotherapy. MRP4 protein expression found to be significantly higher in pancreatic cancer tissue compared to normal pancreatic tissue (87). In pancreatic adenocarcinoma, MRP4 is a poor survival indicator and its inhibition leads to impaired proliferation of cancer cells (88). Furthermore, MRP4 is expressed in patients with AML and upon cAMP signalling activation, the self-renewal capacity of blast progenitors is suppressed (71). Elevated MRP4 expression is shown to be an important factor for the progression of prostate cancer, as an androgen responsive protein whose expression is increased in the progression to prostate cancer but decreased with androgen withdrawal (74). Moreover, MRP4 is associated with response to treatment including methotrexate in pediatric ALL patients since it contributes to methotrexate accumulation and polyglutamylation (89). MRP4 was found to be transcriptionally regulated by MYCN (N-myc proto-oncogene protein), which plays a key role in neuroblastoma tumourigenesis. In addition, MRP4 effluxes endogenous signalling molecules that may play a role in tumour survival and proliferation, including cyclic nucleotides and eicosanoids, as well as chemotherapeutic agents. Downregulation of MRP4 inhibits proliferation of human neuroblastoma cell lines whereas high MRP4 mRNA expression is a predictor of poor clinical outcome (90). Also, MRP4 protein is highly expressed in most of the lung cancer cells and its expression increases cell growth by promoting cell cycle progression (91). MRP4 plays an important regulatory role in the proliferation of drug-resistant human gastric cancer cells so its downregulation is associated with the inhibition of the proliferation of drug-resistant human gastric cancer cells and the occurrence of apoptosis (68). MRP4

also was found to interact with β-catenin from Wnt pathway and play a role in embryo implantation as well as in endometrial cancer. It stabilizes β-catenin from degradation, probably with protein-protein interactions, and therefore sustains Wnt/β-catenin signaling in endometrial epithelial cells. Such a role of MRP4 in the endometrium is involved in both processes of receptivity transition for embryo implantation and pathogenic transformation to develop endometriosis or endometrial cancer (81). Recently, it was shown that MRP4 can be a factor in the pathogenesis of obesity and diabetes, as it plays an important role in the regulation of adipogenesis and adipose tissue physiology by altering cAMP and PGE_2 levels, and its lack promoted adipogenesis and metabolic disease in mice by increasing intracellular cAMP and decreasing extracellular PGE levels (92)

1.7 Membrane Protein Expression

Membrane protein expression remains a challenging task. The first issue is obtaining the protein of interest. Membrane proteins exist in low levels in the membranes. If the protein is to be expressed in heterologous systems and be overexpressed, it can aggregate. Especially for mammalian membrane protein heterologous expression, the protein may need post-translational modifications in order to be functional and this might not be done properly in the other systems even though some of them, like insect cells, have post-translational modifications similar to the mammalian systems. Another challenge is the membrane proteins' native environment which is the lipid bilayer, a fluid-mosaic, which makes biophysical studies rather difficult to be performed and they require the protein to be extracted which leads to other issues. The efficiency of this extraction, must be balanced with protein stability so that after extraction it remains stable, well-folded, it does not aggregate, and retains the conformation, it is another challenge that one needs to overcome as the lipids serve to maintain the protein's structure and play a role in structure. Plus, the membrane proteins' hydrophobicity made their extraction and stability outside of the lipid bilayer rather hard to deal with (9). To this date,

efforts have been made to express membrane proteins for structural or functional studies in different hosts: bacteria, yeast, insect cells and mammalian cells. In many cases, mutations are necessary for the expression, but this has an impact on structure or function. Furthermore, it is usually difficult to obtain both high yield and functionality. While bacteria give high yields quickly, they do not provide post-translational modifications. Yeast can also provide high yield quickly but like bacteria, cannot be used for large genes. Mammalian cells usually give functional proteins but lower yields. On the other hand, insect cells seem to be ideal, especially for the expression of MRP4, since they provide both high yield and a functional protein and most of the post-translational modifications, such as glycosylation which is important for MRP4 (Kesidis, Depping et al. 2020) (Hardy, Bill et al. 2019). Insect cells are usually infected with the baculovirus system. The most commonly-used baculovirus is AcMNPV (*Autographa californica* multicapsid nucleopolyhedrovirus). Recombinant baculoviruses can be produced using systems such as Bac-to-Bac or FlashBAC. The Bac-to-Bac system works as follows: the gene of interest is cloned into a transfer vector, pFastBac, under the control of the AcMNPV polyhedron promoter. Polyhedrin is a protein involved in forming the coat of the baculovirus, and thus has a strong promoter, however polyhedrin is not needed for infection and replication within cultured cells. The pFastBac vector containing the gene of interest is then transformed into specialised DH10Bac *E. coli* cells. They contain a baculovirus shuttle vector (bacmid) encoding the baculovirus genome with a transposon, and a helper plasmid, which aids with the site-specific transposition of the gene of interest in the bacmid. This recombinant bacmid is then isolated, purified and transfected into insect cells to produce the recombinant baculovirus which is released into the medium (Figure 1.10A). The other key method is the FlashBac system. The transfer vector contains the gene of interest and an essential gene ORF1629. The flashBAC DNA contains a truncated version of the essential gene which can only work after homologous recombination with the plasmid. Both the plasmid and the flashBAC DNA co-transfect insect cells and after homologous recombination the recombinant baculovirus is generated (Figure 1.10B).

Although the baculovirus system is often considered only for use with insect cells, there is also the BacMam system that uses a recombinant baculovirus to transduce mammalian cells. If the recombinant baculovirus is generated with a mammalian promoter (such as cytomegalovirus CMV) then it can be used to transduce mammalian cells in order to produce the protein of interest. The baculovirus cannot infect mammalian cells in the classical sense, and does not lead to amplification and release of further viral particles, but the baculovirus can enter and cause protein expression (12).

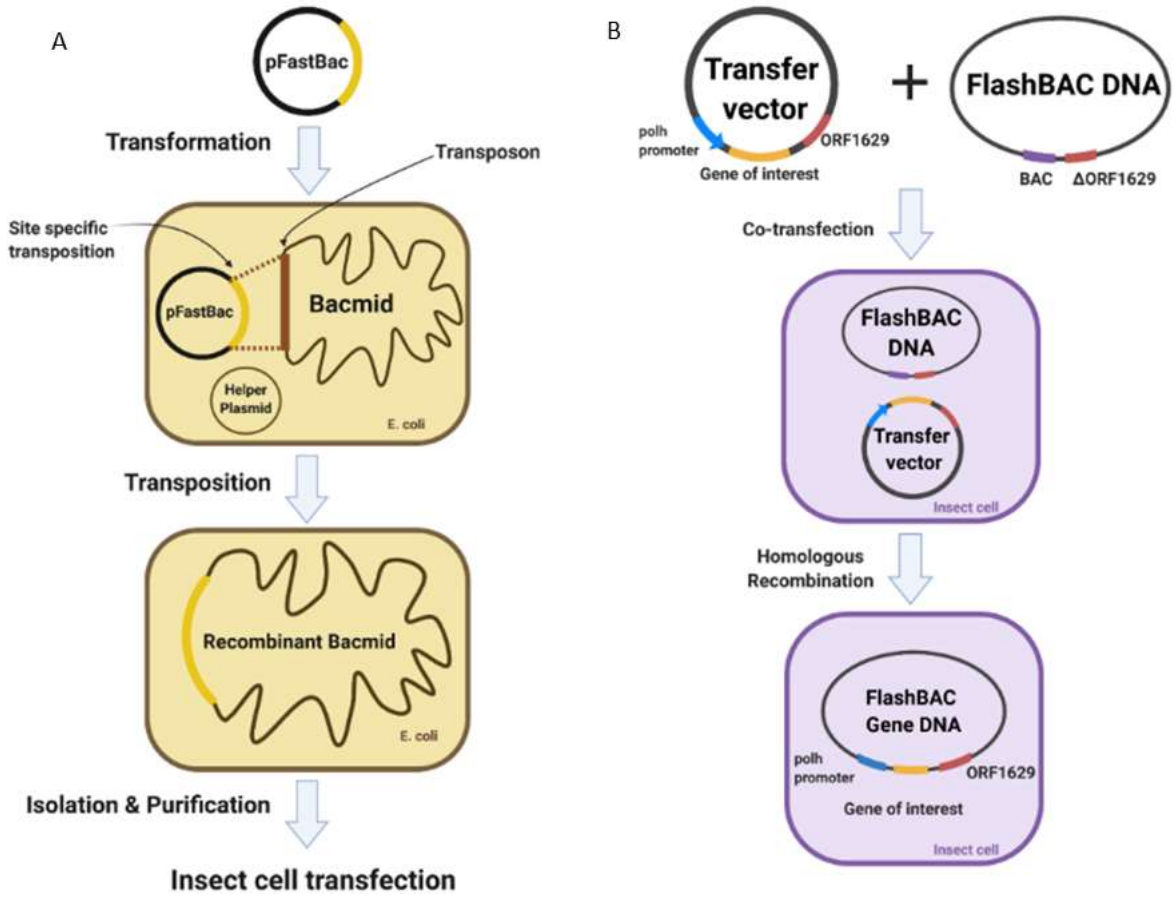


Figure 1.10 The baculovirus systems. *A; Bac-to-bac. Transformation of E.coli cells containing the bacmid, with the plasmid containing the gene of interest occurs, and after site-specific transposition, the recombinant bacmid is generated. Then, it is isolated and purified and used to transfect insect cells which will produce the recombinant baculovirus. The recombinant baculovirus is amplified in insect cells and infect insect cells to express the protein of interest. B; FlashBAC system. The transfer vector contains the gene of interest and an essential gene ORF1629. The flashBAC DNA contains a truncated version of the essential gene which can only work after homologous recombination with the plasmid. Both the plasmid and the flashBAC DNA co-transfect insect cells and after homologous recombination the recombinant baculovirus is generated. It can be amplified in insect cells and then used to transect insect cells to produce the protein of interest. In the case of BacMam the same strategy is followed but the plasmid has a CMV promoter for mammalian expression and the virus after amplification in insect cells can be used to infect mammalian cells (12), adapted (permission asked and taken for the use of this figure)*

1.8 Membrane protein solubilisation

MRP4 clearly has potential to be a drug target for several different conditions and merits further study of its structure and function. However, MRP4 is a transmembrane protein and these proteins are hard to extract from the lipid bilayer where they are embedded, whilst retaining structure and function. Traditionally, detergents have been used, which are amphiphilic molecules with hydrophilic and hydrophobic moieties, with a polar head group and a hydrocarbon chain. In solution, they can appear as monomers but they form micelles above the critical micelle concentration (cmc). Micelles are ordered thermodynamically stable spherical structures with polar head groups facing the exterior, while hydrophobic tails face the interior (93). The cmc is specific to each detergent and is defined as the concentration below which only detergent monomers exist in solution and above which detergent micelles start to form (94). For protein solubilisation, detergents act as follows: the hydrophobic portion of detergents interacts with membrane proteins to form a micelle structure and the hydrophilic portion provides water solubility (Skrzypek, Iqbal et al. 2018). Detergents generally work well for extracting membrane proteins, and a lot of great work has been accomplished with them, but the detergent tends to strip away the native membrane and encapsulates the hydrophobic transmembrane region within a micelle to keep it suspended in solution - often ridding all lipids from the membrane protein complex in the process (figure 1.11) (Hesketh, Klebl et al. 2020).

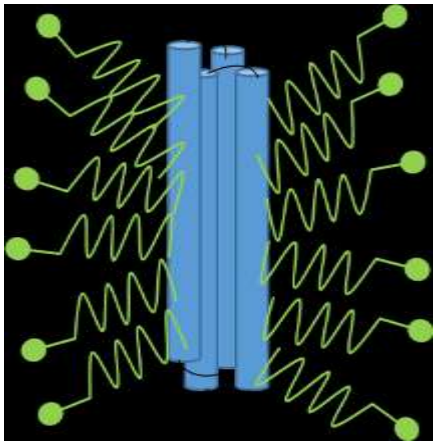


Figure 1.11 Membrane protein (blue) in detergent (green). *The hydrophobic moieties of the detergent cover the hydrophobic parts of the protein while the detergent hydrophilic head groups and the hydrophilic parts of the protein are in touch with the solution.*

Detergents can be classified according to their structure in four main categories: ionic, bile acid salts, non-ionic and zwitterionic. Ionic detergents have a head group that can be either anionic or cationic and a hydrocarbon or steroidal hydrophobic chain. The cmc of an ionic detergent is determined by the combined effect of the head group repulsive forces and the hydrophobic interactions of the tails. An example is sodium dodecyl sulfate (SDS) but this is such a strong detergent that usually denatures the proteins, even though it is very effective as solubilisation. Bile acid salts, such as CHAPS have also the same type of head group as the ionic ones, but their hydrophobic part consists of rigid steroidal groups. They form small kidney-shaped aggregates instead of spherical micelles formed by traditional ionic linear-chain detergents. Bile acids are considered as mild detergents. Passing to the non-ionic detergents, their head groups are not charged but they are either polyoxyethylene or glycosidic groups. They are considered as mild and non-denaturing which is helpful for the function maintenance of the protein. However, small chain (c_7-C_{10}) non-ionic detergents such as n-octyl-h-d-glucopyranoside (OG) can be denaturiting in contrast to long chain ones (C_{12}-C_{14}). This class of detergents is commonly used to extract peripheral membrane proteins. Common detergents in this category include

$C_{12}E_9$ (Thesit), Triton X-100, Tween 20. Other alkylglucosides, such as n-dodecyl-h-d-maltoside (DDM), are commonly used in membrane protein solubilization as many proteins can be easily solubilized in a functional state in DDM but with retention of functional properties. Zwitterionic detergents combine the properties of ionic and nonionic detergents but tend to be more deactivating than nonionic detergents. An example of zwitterionic detergent is dodecyldimethyl-N-amineoxide (DDAO) (figure 1.12) (9) (93).

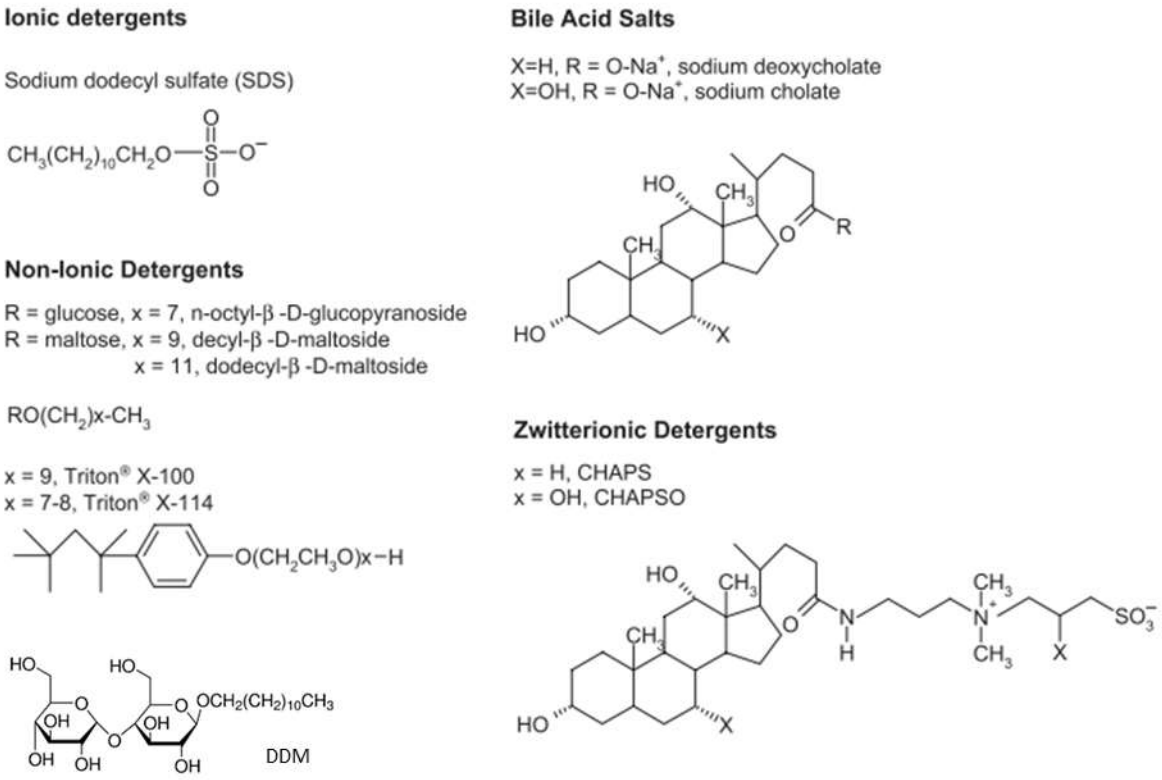


Figure 1.12. Example of structures of the different classes of detergents. *Detergents are divided in ionic, non-ionic, bile acid salts and zwitterionic. DDM, belongs to non-ionic, adapted from* https://www.rosesci.com/Products/products.php?title=DDM%20(n-Dodecyl-%CE%B2-D-maltoside),%2098% *(visited 14 July 2022, at 12:14) and (9), adapted (permission asked and taken for the use of this figure)*

Differences in detergent features influence their critical micelle concentration, the solubility in water, their aggregation, the protein solubilisation and stabilization as well as pH sensitivity (94). The cmc, as mentioned above, can be defined as the minimum concentration of detergent for individual detergent molecules to cluster and form micelles. Above the cmc the detergent monomer concentration is independent of the total detergent concentration. The cmc differs in various conditions, such as pH, ionic strength, temperature as well as the presence of protein, lipid and other detergent molecules. The cmc decreases with the length of the alkyl chain of the detergent and increases on the introduction of double bonds and branch points. At low temperatures, detergents exist mainly in a crystalline insoluble form that is in equilibrium with small amounts of solubilized monomers. As the temperature is increased, more monomer dissolves until the cmc is reached. This is the critical micellar temperature (cmt). Above the cmt, nonionic detergents become cloudy and phase separate into a detergent-rich layer and an aqueous layer, presumably due to a decrease in hydration of the head group. The temperature at which this happens, is the cloud point. A low cloud point can be of assistance in the solubilization of membrane proteins. This property can be used to a particular advantage. Membranes can be first solubilized at 0°C and the solution can be warmed for the phase separation to occur. This allows partition of integral membrane proteins into the detergent rich phase, which can be separated by centrifugation. The number of monomers that exists in a micelle is called aggregation number. It is calculated by dividing the relative molecular mass of the micelle with that of the monomer (9). Each time that a membrane protein needs to be solubilised, the suitable detergent should be chosen according to efficiency as well as other features. First, the detergent should maintain the protein's function. Then, the detergent should be compatible with the downstream processes like purification and crystallization. Sometimes, it occurs that membrane proteins are solubilized in a detergent but during purification the detergent is replaced by a second detergent that can promote crystal lattice formation for structural studies for example (93). A screening of several detergents is a good way to go for the choice of the most suitable. For this, a variety of detergents should be tried, different structures and charges, different categories (ionic, non-ionic, zwitterionic),

detergents with high and low cmc (95). In recent years several different novel detergents have been developed, with the aim of stabilising membrane proteins at the same time as solubilising them. Some examples are CALX8 (5,11,17-tris[(carboxy)methyl]-25-monooctyloxy-26,27,28-trihydroxycalix[4]arene), DDG (N-(2-methyl-1,3-bis(O-β-D-Glucose)propan-2-yl)-3-(decylthio)propanamide) and DDTAC (Dodecylmercapto-S-(poly(tris(hydroxymethyl)acrylamidomethane) DPn=6) from CALIXAR (figure 1.13).

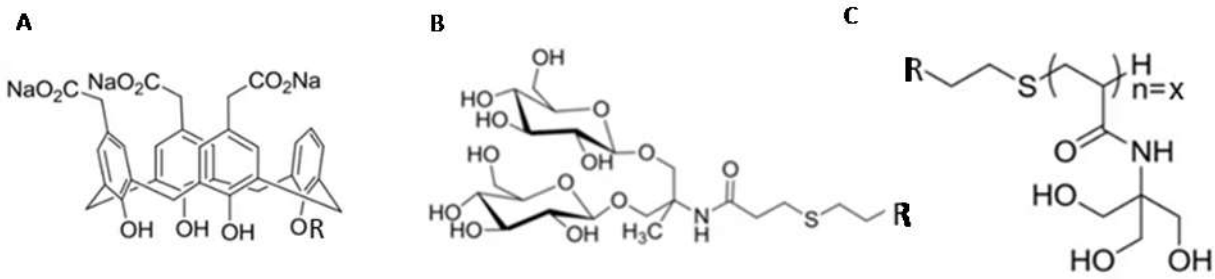


Figure 1.13. Novel detergents aiming to stabilise membrane proteins. *A. Calixarene-, B. Diglucose- , C,Trisacrylamidomethane-based detergents (CALIXAR). CALX8, DDG and DDTAC respectively are examples of these type of structures), adapted from https://calixar.com/ (visited 4 September 2022, at 16:07)*

Other detergents include the neopentyl glycols (MNG). Members with different chain lengths or with different sugar head groups have been described and are commercialized by Anatrace. The most commonly used is Lauryl Maltose Neopentyl Glycol (LMNG), also named MNG3 (MNG3, however, means the related family of compounds). Its structure is formed by two molecules of DDM linked between the hydrophilic and hydrophobic moieties of the molecule. The presence of the two hydrophobic chains of equal length mimics better the structure of lipids than classical detergents having a single hydrophobic tail (figure 1.14) (96). LMNG was used for the solubilisation of MRP1, an ABC transporter of the same subfamily as MRP4 with addition of cholesterol hemisiccinate (CHS)

which can assist the extraction and stabilization of membrane proteins since it mimics the cholesterol of the membrane so the proteins are in a more native environment (97). To date, for the membrane extraction step for most of the membrane protein structures determined, detergents were used.

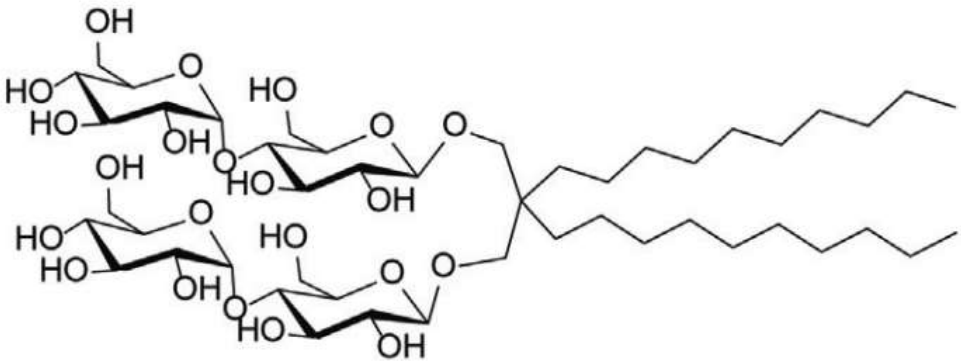


Figure 1.14 LMNG structure. *LMNG molecule is composed of two molecules of DDM linked between the hydrophilic and hydrophobic moieties of the molecule (5), adapted (permission asked and taken for the use of this figure)*

1.9 SMALPs

Due to the challenges that can arise when working with detergents, alternative approaches have been developed. Styrene maleic acid (SMA) copolymers, consisting of hydrophobic styrene and hydrophilic maleic acid monomer units, have been found to be able to solubilize lipids without the need for detergents. The polymer, inserts in the membrane creating a "ring" where the membrane protein is encapsulated with the nearby lipids, retaining therefore a native environment for the protein (figure 1.15) (2). SMA copolymers were utilized for detergent-free extraction of proteins directly from native cell bilayers into stable bilayer disks, called SMA Lipid Particles (SMALPs) (98).

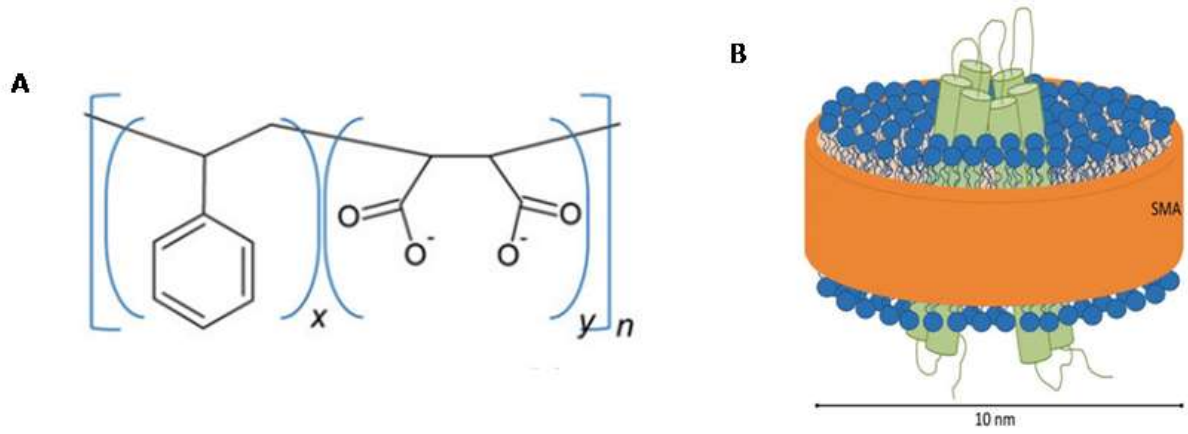


Figure 1.15. A; SMA structure. SMA is composed of hydrophobic styrene (x) and hydrophilic maleic acid units (y). B; SMALP. In green is shown the membrane protein, in blue the lipids and in orange SMA surrounding the protein and the nearby lipids creating a "ring" of about 10nm (3), adapted (permission asked and received for the use of this figure)

The molecular mechanism of insertion remains unclear. However, according to a recent molecular dynamics study (99), the insertion mechanism first involves SMA copolymers binding to the membrane through the styrene moieties of the termini. Initial insertion begins with hydrophobic interactions with the core of the lipid bilayer. Then, SMA copolymers' hydrophobic side chains continue the insertion, causing local membrane undulation. Translocation of the SMA copolymers follows and it relieves the induced stress, together with water molecules and accommodated by lipid flip-flop. Small transmembrane pores form and they grow bigger. The SMA copolymers stabilize the rim by orienting the carboxyl moieties to the water pore, and the benzene groups intercalate in between the lipid tails. This seems to disrupt the membrane and facilitates SMALP formation (99). SMALPs have been used for membrane protein extraction of different systems and they seem to be stable, the excess of free SMA can be removed, making them compatible with other downstream techniques, from purification to cryo-EM, X-ray crystallography or biophysical methods such as circular dichroism spectroscopy (100-104). As happens with everything, SMA is not always "the perfect solution" as it has some limitations. It absorbs at a λ of 260nm, which helps the monitoring and the quantification of the

polymer but it poses issues for some spectroscopic methods, such as circular dichroism that it is needed to remove protein-free nanodiscs before the experimental procedure (105). Moreover, SMA is sensitive to low pH or to divalent cations, which cause it to precipitate and this can be problematic in some cases such as proteins that require low pH (106) or proton binding or divalent cations such as Mg^{2+} for functioning (107). Furthermore, it has been suggested that SMA holds the proteins tightly in the SMALP so it can create problems in protein function as it is not so free for conformational stages to occur (2). To overcome these issues, several other polymers have been proposed such as styrene maleimide (SMI) which seems to be more effective in low pH and less sensitive to divalent cations (108), or diisobutylene maleic acid (DIBMA) which replaces the styrene moiety with the aliphatic diisobutylene group, so it does not absorb at λ of 260nm. This polymer is also tolerant to divalent cations and produces larger discs than SMA (figure 1.16) (105). The less restricted environment within DIBMA particles has been shown to allow full function of proteins which were restricted from completing all conformational changes within SMALPs (2). However, although DIBMA works well for some proteins such as GPCRs, for others, like the ABC transporter BmrA, it gives a lower yield of protein, lower purity and decreased stability (109) and whilst SMI also works well for GPCRs and ZipA (2) it has to date not been useful for ABC trasnsporters. A good example of the comparison of these three polymers (SMA, SMI and DIBMA) is given by the rhodopsin. In a study, it was shown that all polymers had efficiently solubilised rhodopsin. However, the protein seems to reach the activated state only when encapsulated in DIBMA. It was found that DIBMA creates larger particles, allowing fluidity and movement (110).

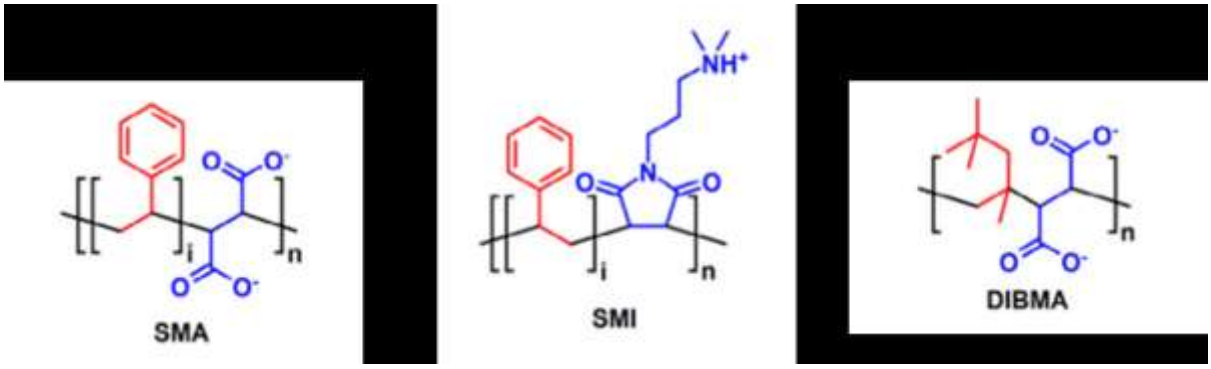


Figure 1.16. Polymer structures, SMA, SMI and DIBMA. *In red is the hydrophobic group and in blue the hydrophilic group. n and I represent the numbers of repeats of each unit (6), adapted (permission for use of this picture was asked and received)*

The solubility range of SMA depends on the styrene/maleic acid ratio of the polymer: an increase in the styrene content can decrease the range in which SMA is water soluble. The number of charges per unit length of polymer determines the SMA solubility in a pH range. When the SMA copolymer contains few maleic acid units per polymer, it is more hydrophobic and therefore it needs a higher ionization state to remain water soluble, which is important for its capacity to solubilize lipid membranes. A decrease in pH or an increase in the styrene fraction in the polymer induces a conformational change in SMA (106). Many different SMA copolymers are commercially available, which can vary in the ratio of styrene:maleic acid and in their total length and size. The polymers that have mainly been used to extract membrane proteins are SMA (2:1) and SMA (3:1), which have styrene:maleic acid ratios of 2:1 and 3:1, and are relatively small with an average molecular mass of 7.5 and 10 kDa, respectively (107). Despite their different styrene content and linear charge density, the different SMA variants (2:1, 3:1, and SMA 4:1) all show similar insertion behaviour at pH 8.0 (106). Structures of membrane proteins solubilized with SMA, have been revealed using X-ray crystallography and cryo-EM. An example is a bacterial rhodopsin which was obtained with X-ray crystallography (111). Cryo-EM seems more successful approach though. Some examples are P-

glycoprotein (ABCB1) (100), Alternative Complex III (ACIII) from Flavobacterium johnsoniae (104), AcrB (112), KimA, a chicken acid sensing ion channel (ASIC1) and glycine receptor (113).

1.10 Aim and objectives

The overarching aim of this work was to express, solubilise and purify MRP4, with hope to investigate the structure of MRP4 using cryo-EM. Structural studies can give insight in various aspects. With a structure, calculations of distances between the protein and other molecules such as ligands or inhibitors become possible, and where the molecule binds in the protein can be of high importance for examining the possibility of changes in the molecule in order for it to "fit" better in the protein's binding site. Moreover, by capturing different conformations, and by studies with ligands and inhibitors, the mechanism of function can be revealed. Also, a structure can provide information on how a protein misfolding can lead to a disease. With all this information, drug design becomes more accurate.

To achieve the protein structural studies in this case, the objectives were:

- To optimise expression of MRP4 in Sf9 insect cells
- To optimise solubilisation efficiency, resin binding and purification yield of MRP4 when using SMA2000 polymer.
- To investigate novel stabilising detergents for the extraction and purification of MRP4, and compare these with SMA2000.
- To test expression of MRP4 in mammalian cells, 293F, using the BacMam stystem

2. Methods

2.1 Cell cultures

2.1.1 Sf9 insect cells

Sf9 (*Spodoptera frugiperda*) insect cells were thawed quickly and cultured in Insect-Xpress medium (Lonza) with 10% foetal bovine serum (FBS) and 100 U/ml penicillin and 100 mg/ml streptomycin and incubated at 27 °C. The cultures were initiated as monolayer cultures in T-75 flasks, which were split when they reached confluence. Cells could be detached from the flask by tapping the flask. Cells were also expanded in suspension shaker cultures, in volumes ranging from 30ml to 1L. The volume of culture never exceeded 30% of the flask volume. The cells were split when they reached a density of 3.5 -4 X 10^6 cells/ml of culture. They were split at a final concentration of 0.5-1 x 10^6 cells/ml. The incubation was at 27°C, at 130 rpm.

2.1.2 293F cell cultures

Freestyle 293F cells (Invitrogen) were thawed quickly and added directly to a shaker flask in a final volume of 25ml of FreeStyle™ 293 Expression Medium (Gibco) without addition of serum or antibiotics. They were incubated at 37°C, with 8% CO_2, and shaken at 130 rpm. The cells were split when they reached density of 2-3 X 10^6 cells/ml to 0.5 x 10^6 cells/ml. Generally, they were maintained at concentrations of 0.5-3 x 10^6 cells/ml.

2.1.3 Freezing cells

1 x 10^7 cells were frozen. After cell counting, the appropriate volume needed was calculated and Sf9 cells were centrifuged at 500 xg, room temperature, for 5 minutes while 293F cells were centrifuged

at 100 xg, room temperature, for 5 minutes. The supernatant was removed and the pellet was resuspended in a mix of 45% fresh medium, 10% dimethyl sulfoxide (DMSO) and 45% contained medium. Then, the cells were aliquoted in cryovials with 1 ml each and were put in a Mr Frosty and placed in the -80^{o}C freezer. Mr Frosty has isopropanol and drops the temperature by 1^{o}C per minute so the cells are frozen slowly. They need to stay there for 2 days. The 293F cells were stored in liquid nitrogen.

2.2 Baculovirus mediated protein expression in Sf9 cells

2.2.1 pFastBac construct and baculovirus generation

For the insect cell infection, the bac-to-bac system was used. The plasmid utilized is shown in figure 2.1. The protein contains the 6-Histidine tag in order to facilitate binding in Ni^{2+} - Nitriloacetic acid (NTA) resin for purification. This construct was previously transformed into DH10Bac *E.coli* that contain the virus Bacmid and a helper plasmid. After site specific transposition, a recombinant Bacmid deoxyribonucleic acid (DNA) was generatedand after it was isolated and purified. This was used to transfect Sf9 cells to form a recombinant baculovirus (114). The recombinant baculovirus was stored in sterile tubes at 4 °C.

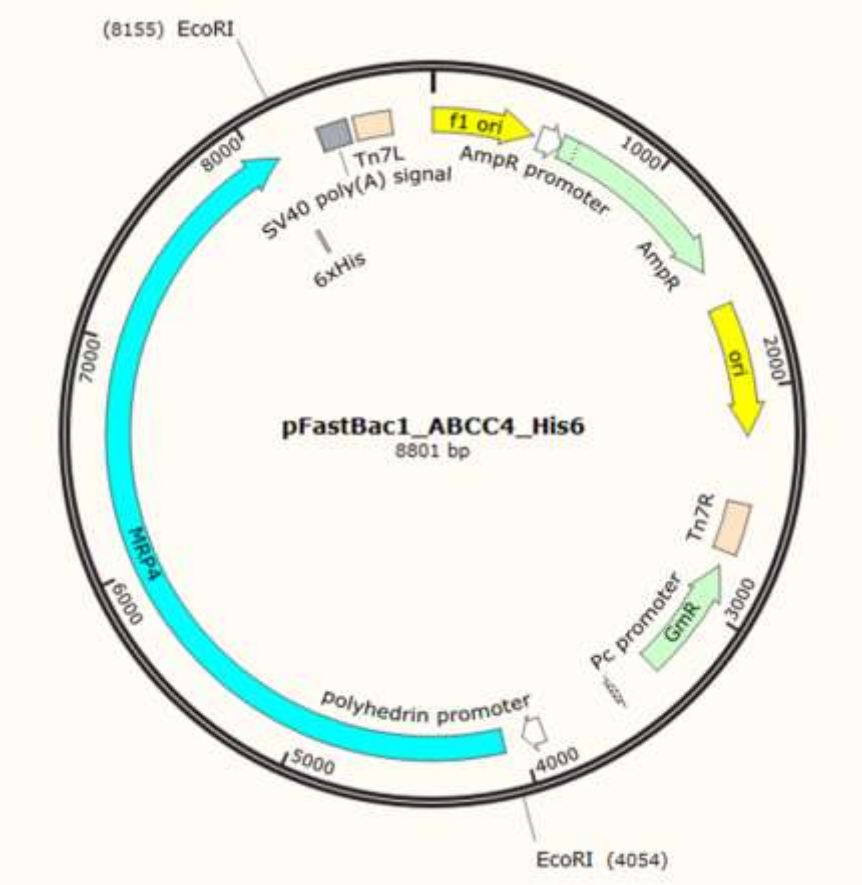


Figure 2.1. pFastBac1 plasmid with the gene of MRP4 (ABCC4) *(in blue). The plasmid has a polyhedrin promoter for protein expression in insect cells, ampicillin resistance gene for selection of the plasmid in E.coli, gentamicin resistance gene for selection of the recombinant Bacmid in DH10Bac E.coli, and a 6-His tag.*

2.2.2 Baculovirus amplification

Baculovirus P2 generation which had been stored at 4 °C was used for amplification. Cells were split the day before to a concentration of 0.7-0.8 x 10^6 cells/ml. When Sf9 cell cultures reached a density of 1.5 X 10^6 cells/ml, 220µl of virus/ 100 ml of culture were added. The culture was incubated at 27 °C, 100 rpm for 5 days. Then, the cultures were centrifuged at 500 xg for 10 minutes, the supernatant

containing the P3 generation baculovirus was collected and stored at 4 °C covered with aluminium foil.

2.2.3 Protein Production

Baculovirus containing the MRP4 gene of P3 generation were used for the protein production. The day before infection, the cells were split to a final concentration of 0.7-0.8 x 10^6 cells/ml in order to have a density of around 1.5 X 10^6 cells/ml of culture the day after, they were then mixed with baculovirus at a ratio of 40ml virus/L cells and the cultures were incubated at 27°C for 48 hours. The cells were harvested by centrifugation at 7000 xg for 10 minutes at 4°C. Then, they were washed with phosphate buffered saline (PBS), centrifuged at 3220 xg for 10 minutes, the supernatant removed and the cell pellets were frozen at -80°C.

2.2.4 Membrane preparation of Sf9 cells

The cell pellets were resuspended in homogenization buffer pH 7.4 (50mM Tris, 250mM sucrose, 0.25mM $CaCl_2$) and protease inhibitors (1 µM pepstatin, 1.3 µM benzamidine and 1.8 µM leupeptin) were added. The cells were disrupted using a nitrogen cavitation device. Nitrogen was added to a pressure of 500psi, and the device left on ice for 15 minutes to equilibrate. The cells were then released slowly from the bottom of the device at approximately 1 drip/sec. Unbroken cells and cell debris was removed by centrifuging at 750 xg for 10 minutes at 4°C. The supernatant was collected and was ultracentrifuged at 100,000xg for 20 minutes at 4°C. The pellets were weighed and the membranes were resuspended at either 60 mg/ml (wet pellet weight) or 180 mg/ml (wet pellet weight) using the Tris-sucrose buffer pH7.4 (50mM Tris, 250mM sucrose). They were aliquoted and stored at -80°C.

2.3 Solubilization and purification using SMA polymers

2.3.1 Preparation of SMA polymer

25g of SMA 2000 (a kind gift from Cray Valley) was dissolved in 250 ml of 1M NaOH in a conical flask overnight. The dissolved SMA 2000 in NaOH was transferred to a 1 litre bottle. It was then autoclaved for 16 minutes at 120˚ C. Afterwards, the solution was left on the bench to cool down until it could be touched. The same autoclave and cooling steps were repeated another two times. Following, the solution was precipitated using concentrated HCl and washed several times with water to remove excess NaCl. The polymer was dissolved in 0.6M NaOH for 3 hours with gentle agitation. The pH was adjusted to 8.0 ±0.1. The polymer was then transferred to a 500 ml round bottomed flask and frozen at -20^{o}C before freeze-drying. Freeze drying followed overnight to obtain the polymer (115).

2.3.2 Solubilisation using SMA polymers

5 or 10 ml of Sf9 cell membranes (60 or 180 mg/ml wet pellet weight) were mixed at a volumetric ratio of 1:1 with 5% w/v SMA 2000 in purification buffer (20mM Tris pH 8, 150mM NaCl). This was incubated for an hour at room temperature with rocking. The sample was then ultracentrifuged at 100,000g, at 4^{o}C for 20 minutes and the supernatant containing solubilised protein was harvested. The pellet containing insoluble protein was resuspended in an equal volume of purification buffer supplemented with 2% w/v SDS. Samples of both the supernatant and resuspended pellet were taken for analysis by Western blot.

2.3.3 Affinity Purification of SMA solubilised MRP4

Ni-NTA resin (HisPur, ThermoFisher) was washed with purification buffer (20mM Tris pH 8, 150mM NaCl) to remove storage buffer containing ethanol and equilibrate it in the purification buffer, then

centrifuged at 500 xg, for 10 minutes, at 4°C, and liquid removed. The supernatant containing solubilised proteins was mixed overnight at 4°C with washed Ni-NTA resin at a ratio of 100µl resin per 1ml of sample. The resin and protein mix was poured into a gravity flow column (Machery-Nagel) and the flow-through was collected. The resin was washed 5 times with purification buffer containing 20mM imidazole and twice with purification buffer containing 40mM imidazole, 10 bed volumes each wash. Then, the protein was eluted in aliquots (0.5 bed volumes) using purification buffer containing 200mM imidazole. Samples of each wash and elution fraction were taken for analysis by SDS-PAGE and Western blot.

2.4 Solubilisation and purification using detergents

2.4.1 Determination of total protein in the membrane

Standards of 0.2 mg/ml bovine serum albumin (BSA) of 0, 2, 4, 8, 12, 20, 30, and 40 µl were prepared with 20 µl of SDS 20% and addition of water to a final volume of 100 µl. The Sf9 cell membrane was diluted 1/50, 1/100 and 1/200 in purification buffer (mentioned above). For each of the three dilutions, 5, 10, 15 and 20 µl of sample was prepared with 20 µl SDS 20% and water was added to final volume of 100 µl. 4 mL revelation buffer (50% reagent A = 2mL, 48% reagent B = 1920 µl, 2% reagent C = 80 µl) were prepared using the microBCA protein assay kit (ThermoFisher) and 100 µl were added in each sample. The samples were incubated at 37°C, at low agitation for 1 hour. Then, the absorbance was monitored at 562nm, a standard curve was created and the protein concentration were calculated based on the trendline equation.

2.4.2 Solubilisation using detergents

CALIXAR detergents and DDM (Anatrace) were prepared with CHS (Anatrace) at a concentration of 200mM and 20mM respectively in 250mM sodium phosphate buffer. LMNG (Anatrace) was prepared with CHS (Anatrace) at concentrations of 100mM and 10mM respectively in 250mM sodium phosphate buffer as well. Sf9 membranes used were of concentration of 10 mg/ml total protein. The membrane samples were solubilized with the detergents (final detergents/CHS concentration 20mM/2mM), for 2 hours at 4°C with gentle agitation. The samples were ultracentrifuged at 150,000xg, at 4°C for 15 minutes and the supernatant containing the solubilized protein was harvested. Samples were taken after solubilisation (total protein) and after ultracentrifugation (solubilised protein) for analysis by SDS-PAGE and Western blot.

2.4.3 Affinity purification of detergent solubilised MRP4

The Ni^{2+}-sepharose or Co^{2+}-sepharose resin was prepared by pouring sepharose with 6% agarose matrix (Cytiva) into a gravity flow column, removing the storage buffer, washing with 5 bed volumes distilled water, applying 0.2 bed volumes of either 200mM $NiSO_4$ or 200mM $CoCl_2$ solution, washing again with 5 bed volumes distilled water to remove excess of ions and equilibrating with 5 bed volumes purification buffer (20mM Tris pH 8, 150mM NaCl). The supernatant after solubilisation was incubated with Ni^{2+}-sepharose or Co^{2+}-sepharose resin (50µl resin/mg of membrane) either for 1 hour or overnight at 4°C. The resin and protein mix was poured into a gravity flow column and the flow-through was collected. The resin was washed 3 times with purification buffer containing 5mM imidazole, 3 bed volumes each wash. Then, the protein was eluted in aliquots of 1 bed volume each, using purification buffer with 150mM imidazole. Samples from each fraction were taken for analysis by SDS-PAGE and Western blot.

To improve purity in the case of DDG-CHS purification, the elution fractions once collected, were concentrated by centrifugation at 4000 xg, for 5-7 minutes, at 4°C, in centrifugal filter units (Merck Amicon). Then, diluted and these two steps were repeated until imidazole was removed and the elution sample could be used again for a purification, the same way as described above.

2.5 Analysis of samples

2.5.1 SDS-PAGE

Gels of 8% polyacrylamide were prepared and run. The gels were casted to be 1mm, using water, separating/ stacking buffer (1.5M Tris pH 8.8/ 0.5M Tris pH 6.8 respectively), 40% acrylamide:bis acrylamide (Fisher), 10% SDS, 10% ammonium persulfate (APS, Sigma) and tetramethylethylenediamine (TEMED, Sigma). Table 2.1 shows the volumes of these reagents for two 8% gels. Samples were mixed with Laemmli sample buffer (LSB) (60 mM Tris pH 6.8, 2 % (w/v) SDS, 10 % (v/v) glycerol, 0.01 % (w/v) bromophenol blue and 5% (v/v) β-mercaptoethanol). They were run initially at 90 volt (V), in SDS-PAGE running buffer 1X (25mM Tris-HCl, 192mM Glycine, 0.1% w/v SDS) and once samples migrated through the stacking gel, the voltage was increased to 120-150V. Unstained protein ladder (Pierce) was used for SDS-PAGE that would be stained later whereas prestained protein ladder (PageRuler Pierce) was used if Western blot followed SDS-PAGE.

Alternatively, TGX-Criterion stain-free precasted gel or TGX-mini-Protean stain-free with 4-15% of acrylamide (Bio-rad) were pre-run at 300V for 3 minutes. Samples were mixed with SDS-TCEP (tris(2-carboxyethyl)phosphine) (Alfa Aesar). Once the samples were on the wells, the gel was run at 300V for 20 minutes or 13 minutes respectively. The same ladder (ProSieve quad-colour, Lonza) was used for gels that would be stained or continue to western blot.

After SDS-PAGE, the gel was stained with Instant Blue (Abcam), or visualized by activation under UV light (stain-free technology) (Bio-rad), or used for western blot analysis.

Reagent	Separating gel	Stacking gel
Water	5.4 ml	2.4 ml
Buffer (separating/stacking)	2.5 ml	1 ml
40% acrylamide	2 ml	0.5 ml
10% SDS	100 µl	40 µl
10% APS	100 µl	40 µl
TEMED	10 µl	5 µl

Table 2.1. SDS-PAGE gel preparation

2.5.2 Native PAGE

TGX-mini protean or TGX criterion gels were also used for native PAGE analysis. 5 µL of native gel ladder (Thermofisher) was loaded alongside 10 µl protein samples mixed with 6 µL of 4x Native sample buffer (Bio-rad). The gel was pre-run in anode buffer (25 mM imidazole) and cathode buffer (50 mM tricine, 7.5 mM imidazole, 0.05% deoxycholate, 0.01% DDM) for 3 minutes at 200V. Then, the samples were loaded in the wells and migrated for 70 minutes at 200V in anode and cathode buffer. The tank was kept on ice while the gel was run. The protein was visualized by activation under UV light (stain-free technology) (Bio-rad). Coomassie stain with Instant Blue followed or Western blot as described above.

2.5.3 Western blot

For Western Blot, PVDF membrane (Merck) was used for transfer after activation with methanol or ethanol. For the transfer, a "sandwich" of sponges, filter papers, the gel and the membrane need to be created. The order is: sponge, paper, gel, membrane, paper, sponge. The membrane faces the side

of the cathode since the negatively charged proteins need to go to the positive side using electric current. Proteins were transferred from the SDS-PAGE to the polyvinylidene fluoride (PVDF) at a voltage of 100V for 1 hour in 1X transfer buffer (20% v/v methanol or ethanol, 25 mM Tris, 192mM glycine). The membrane was then blocked with 5% (w/v) BSA in TBS-T (20mM Tris pH 7.4, 150mM NaCl, 0.05% Tween-20). The primary antibody was anti-MRP4, M4I-10 (Enzo), used at a dilution of 1:100 in 5% BSA in Tris buffered saline with tween (TBS-T), and incubated with the membrane with agitation for 1 hour at room temperature. Following 3 washes with TBS-T for 5 minutes each, the secondary antibody used was anti-rat linked with *horseradish peroxidase* (HRP) (Cell Signalling) at a dilution of 1:5000 in TBS-T. This was incubated for 1 hour in room temperature with agitation. Following 5 washes with TBS-T, 5 minutes each, the blot was developed using chemiluminescence (ECL substrate, Pierce) and imaged using a LI-COR C-DiGit Blot Scanner.

Alternatively, for the precast gels, the transfer lasted 7 minutes using a TransBlot turbo system (Bio-rad). The "sandwich" here was different: the gel and the membrane were between two filter papers which were soaked in transfer buffer (Transblot turbo, Bio-rad). Then, the PVDF membrane was blocked and incubated with antibodies in the Snap id 2.0 system. Roti block buffer (Carl Roth) used for blocking. No time of incubation was needed. The primary antibody was anti-MRP4, M4I-10 (Enzo) used at a dilution of 1:100 in 5% BSA in TBS-T and incubated with the membrane at room temperature for 20 minutes in the case of Criterion gels and 10 minutes for mini Protean gels. Following 4 washes with TBS-T, the secondary antibody used was anti-rat linked with HRP (Cell Signalling) at a dilution of 1:5000 in TBS-T (incubation times the same as for the primary antibody). Following 4 washes with TBS-T the blot was developed using chemiluminescence (ECL, Bio-rad)) Biorad ChemiDoc imaging system at Chemi-Hi resolution program.

2.5.4 BSA quantification assay

Standards of 0.125, 0.25, 0.5, 0.75, 1 and 1.25µg of BSA were prepared. They were loaded on a polyacrylamide gel (8%). 5, 10 and 20µl of protein sample were also loaded. After running the gel, it was stained with Coomassie Blue (Instant Blue) and densitometry was performed using ImageJ program. A standard curve was produced using BSA concentration and using the equation of the trendline the protein concentration was calculated.

Alternatively, standards of 1, 2, 4, 6 and 8 µl of BSA 0.2 mg/ml were prepared and loaded in a TGX mini protean gel, as well as 1, 2, 4, and 8 µl of protein. After running the gel, it was stained with Coomassie Blue (Instant Blue). Densitometry was performed with Bio Lab software. The protein concentration was calculated after the standard curve of BSA using the equation of the trendline.

2.5.5 Cryostability assay

A sample of purified protein was taken and separated in 2 aliquots of 40 µl each. In one of them, 10% of glycerol (around 4 µl), was added. Samples were taken from each aliquot and they were the T0 (T0- for no glycerol addition, T0+ for glycerol addition). Then, the 2 aliquots were flash-frozen and left to thaw and samples (T1) were taken. This step was repeated another 2 times (T2 and T3) and all the samples were loaded in a TGX mini protean stain-free gel for a native PAGE and Western blot followed as described above.

2.6 Expression of MRP4 in Freestyle 293F cells

2.6.1 Constructs

Two constructs were used for MRP4 expression in 293F cells based on pOET6 vector (Oxford Expression Technologies). They both have a CMV promoter and enhancer, for successful expression in mammalian cells. They also have the Ruby 3 protein gene, in order to make easier the virus detection as the media changes colour from yellow to pink/purple when the virus is produced. Both constructs then, have the gene for MRP4 but on one of them, MRP4 is tagged with 6-His tag, whereas at the other it is tagged with 10-His, green fluorescent protein (GFP) and flag tags. The constructs were made by a collaborator of CALIXAR.

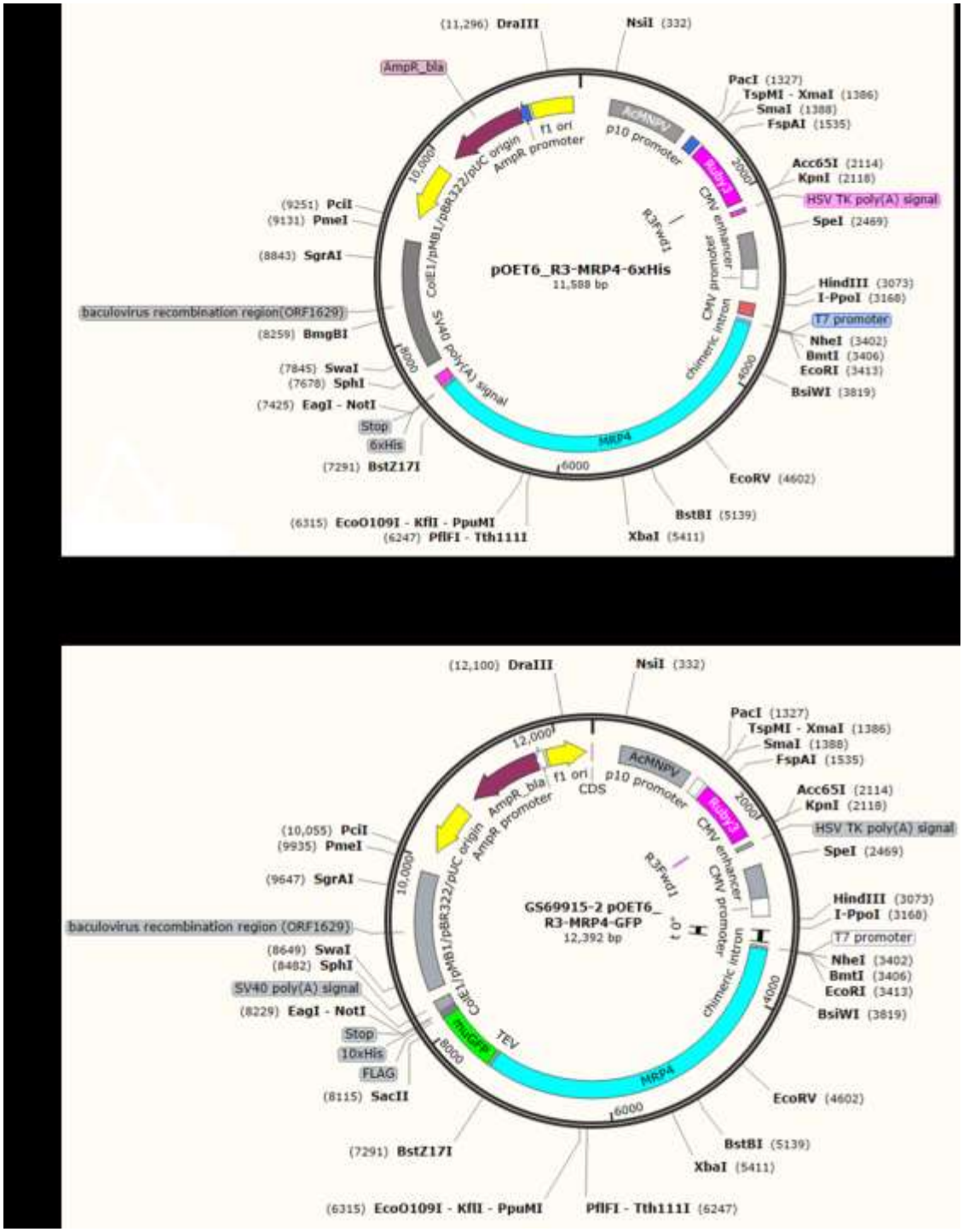


Figure 2.2. pOET6 plasmids with the gene of MRP4 (ABCC4) *(in blue). They both have Ruby 3 protein gene for virus detection once it is produced (purple/pink), p10 promoter (polyhderin) for insect cell production of the recombinant baculovirus and AcMNPV for recombination. The plasmid on top has MRP4 tagged with 6-His while the one at the bottom has MRP4 with 10-His tag as well as GFP and flag tag.*

2.6.2 Generation of recombinant baculovirus

The system used was the BacMam. With this, the transfer vector contains the gene of interest and an essential gene ORF1629. The flashBAC DNA contains a truncated version of the essential gene which can only work after homologous recombination with the plasmid. Both the plasmid and the flashBAC DNA co-transfect insect cells and after homologous recombination the recombinant baculovirus is generated. For the co-transfection, the DNA mix, which is the vector (pOET6 with MRP4) (500ng), the baculovirus genome (FlashBAC ULTRA, Oxford expression technologies) (5 µl), a transfection reagent (Escort, Sigma-Aldrich) (5µl) and insect cell medium (1ml), were mixed according to the manufacturer's instructions (Oxford Expression Technologies) and incubated for 20 minutes at room temperature and then it was used to transfect the cells. 2ml of Sf9 cell of 1 x 10^6 cells/ml (2 x 10^6 cells in total) were seeded in a 6-well plate (one well of each construct) and left at room temperature for 1 hour. Medium was removed and the DNA mix was added. After overnight incubation at 28^{o}C, medium was added and the cells with the DNA mix were incubated for 5 days at 28^{o}C, with a paper soaked in water and in a plastic bag in order to avoid desiccation. The recombinant baculovirus (V1) was collected and centrifuged at 500 xg, for 5 minutes at room temperature and once harvested from the supernatant, it was stored at 4^{o}C.

2.6.3 Amplification of recombinant baculovirus in Sf9 cells

Sf9 cells of 200 ml volume were grown to reach a concentration of 2 x 10^6 cells/ml. Once they reached this concentration, they were infected with 1ml of baculovirus (V1). They were incubated at 28^{0}C, for 7 days at 130 rpm. The virus was harvested by centrifugation at 1000 xg for 15 minutes at room temperature and the supernatant was collected. It was stored at 4^{0}C.

2.6.4 Small scale expression trials using baculovirus mediated transduction of 293F cells

Small scale expression tests in the 293F cells were carried out. The cell concentration used was $2x10^6$ cells/ml, and virus of 300 µl and 600 µl was tested on them. If sodium butyrate (Fisher) would facilitate the protein expression was also tested by addition of it or no addition. Sodium butyrate is reported to induce or repress some genes and hence protein expression in some cells lines (116). 30 µl of 1M sodium butyrate (final concentration 10mM) were used for this test. 3 incubation times, 24, 48 and 72 hours were tested, as well as two different temperatures, 30°C or 37°C. Figure 2.3 shows the different conditions tested. At each time point, the cell culture was collected and harvested by centrifugation at 4000 xg, for 5 minutes at 4°C and whole cell lysates made with RIPA buffer (Sigma-Aldrich) and benzonase (Chemicruz) and the pellets were resuspended**.** Samples were analysed by Western blotting.

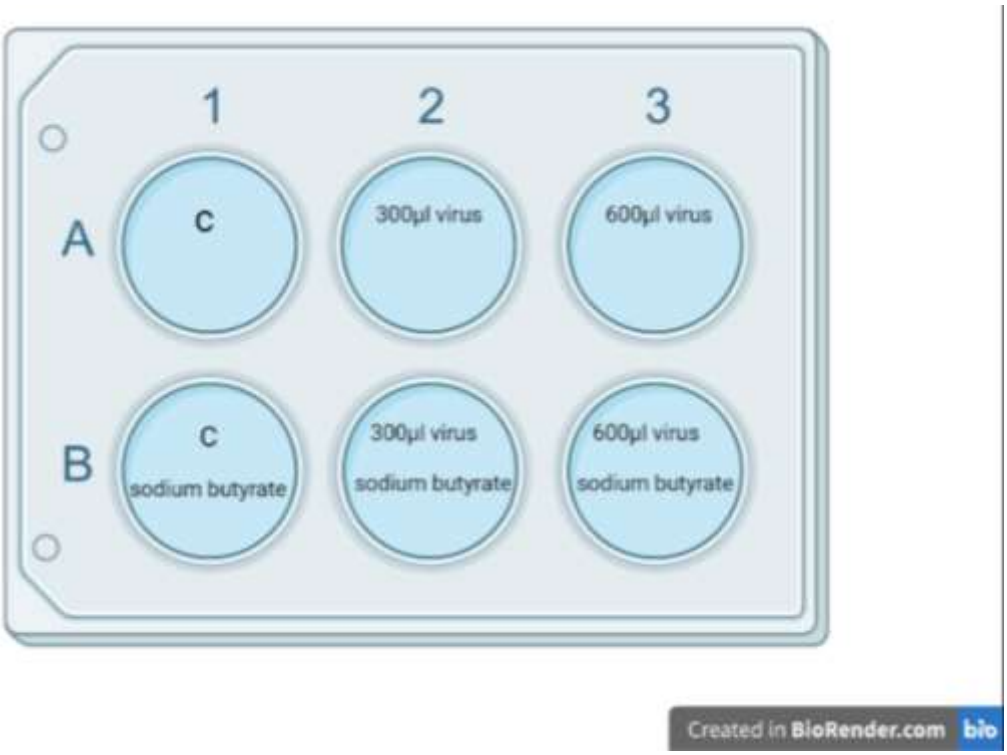


Figure 2.3: Plate layout for expression tests for MRP4 production in 293F cells. *6 well plates were used and 3 ml cells were seeded at a density of $2x10^6$ cells/ml. Virus of volume 300 µl and 600 µl was added in the presence or absence of 10 mM sodium butyrate (NaBu). C stands for control, where no virus was added. Two plates were used for each virus, one at 37°C and the other at 30°C.*

2.6.5 Protein Production (MRP4-6-His)

Cells were grown to a concentration of 2 x 10^6 cells/ml, in total volume 500ml. Once they reached this concentration, they were infected with the recombinant baculovirus of 300 µl volume, and sodium butyrate was added at 10mM concentration. The cells were incubated with the baculovirus for 72 hours, at 37°C, with 8% CO_2, at 130 rpm. After the 72 hours, the culture was centrifuged at 4000 xg, for 5 minutes, at 4^0C. The pellets were stored at -20°C.

2.6.6 Membrane preparation

The cell pellets were resuspended with PBS (Alfa Aeser) and protease inhibitor cocktail (PIC) (Sigma-Aldrich), 2 ml of buffer per 1g of pellet. The microfluidizer was used at 8000 psi with the H30-Z chamber. The microfluidizer was prepared as follows: ethanol buffer was removed, it was washed with 200 ml water, then with PBS, then with PBS containing PIC. The cells were passed through the microfluidizer two times. The cell lysate was ultracentrifuged at 150000 xg, at 4°C for 15 minutes. The supernatant was removed and the pellets were resuspended in PBS containing PIC and glycerol (0.5 ml PBS per 1 g of pellet). The membranes were stores at -80°C.

2.7 Data analysis

2.7.1 Statistical analysis

GraphPad was used for statistical analysis. When one variable was altered, such as the different solubilisation conditions, analysis was carried out by one way ANOVA with a Tukey post-hoc test. When two variables were altered, such as the different buffer conditions combined with different resin volumes, a two way ANOVA with a Tukey posy hoc test was used. A p values less than 0.05 was considered significant.

3. Polymer mediated solubilisation and purification of MRP4 from Sf9 insect cells

3.1 MRP4 expression in Sf9 cells

MRP4 has previously been shown to be successfully expressed in *Sf9* insect cells (117). However, the yield of protein obtained was lower than ideal. Therefore, the first step was to investigate whether the expression level obtained could be improved. One aspect investigated was the length of time between baculovirus infection and harvest of the cells. The cells at 1 x 10^6 cells/ml were infected with the recombinant baculovirus containing the mrp4 gene with 20ml virus in every 500ml culture. The cells were incubated with the virus for either 48 hours or 72 hours. As it is shown in figure 3.1, there is a strong and clear band after 48 hours at ≈ 150kDa, which corresponds to the size of MRP4, while after 72 hours this band is not as intense and it seems that another band appears to give a stronger signal at a size of approximately 70kDa suggesting protein cleavage since the antibody used is specific for MRP4.

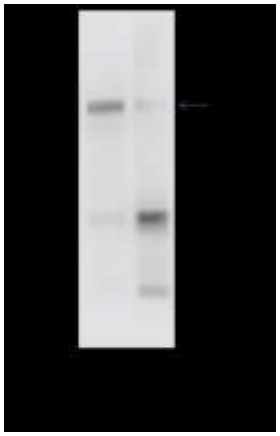


Figure 3.1. Expression of MRP4 in Sf9 is better after 48 hours baculovirus infection rather than 72 hours*. Cells (1 x 10^6 cells/ml) were infected with the recombinant baculovirus (20ml virus in every 500ml culture) containing the mrp4 gene, for either 48 hours or 72 hours. Cell were harvested and membranes prepared, and samples analysed by Western blotting. M4$_I$-10 α-MRP4 primary antibody (1:100) and α-rat-HRP as secondary (1:2000) were used. Results were obtained with chemiluminescence and were revealed in LI-COR C-DiGit Blot Scanner.*

In order to investigate at which point this cleavage occurs, whether it was from expression itself or later during processing, samples from the culture were taken at different steps. These stages were the culture suspension, the resuspended cell pellet following harvest by centrifugation, the homogenized cells, the lysed cells, sample of the supernatant after centrifugation of the lysed cells and in the end sample of the membrane prepared (Figure 3.2). The samples were analysed by Western blot using *M4$_I$-10* α-MRP4 primary antibody and α-rat-HRP as secondary. As shown in figure 3.3, the intact MRP4 band is the major band in all the samples taken from the culture incubated for 48 hours with the recombinant baculovirus, with limited cleavage. However, the membrane preparation sample does show some cleaved protein in addition to full-length. On the other hand, the culture incubated with the recombinant baculovirus for 72 hours shows cleavage of MRP4 at every stage suggesting cleavage occurs in the cell, possibly due to the extended incubation with baculovirus which eventually leads to cell lysis. Therefore, infection for 48 hours remains the best choice to achieve higher yields of intact MRP4.

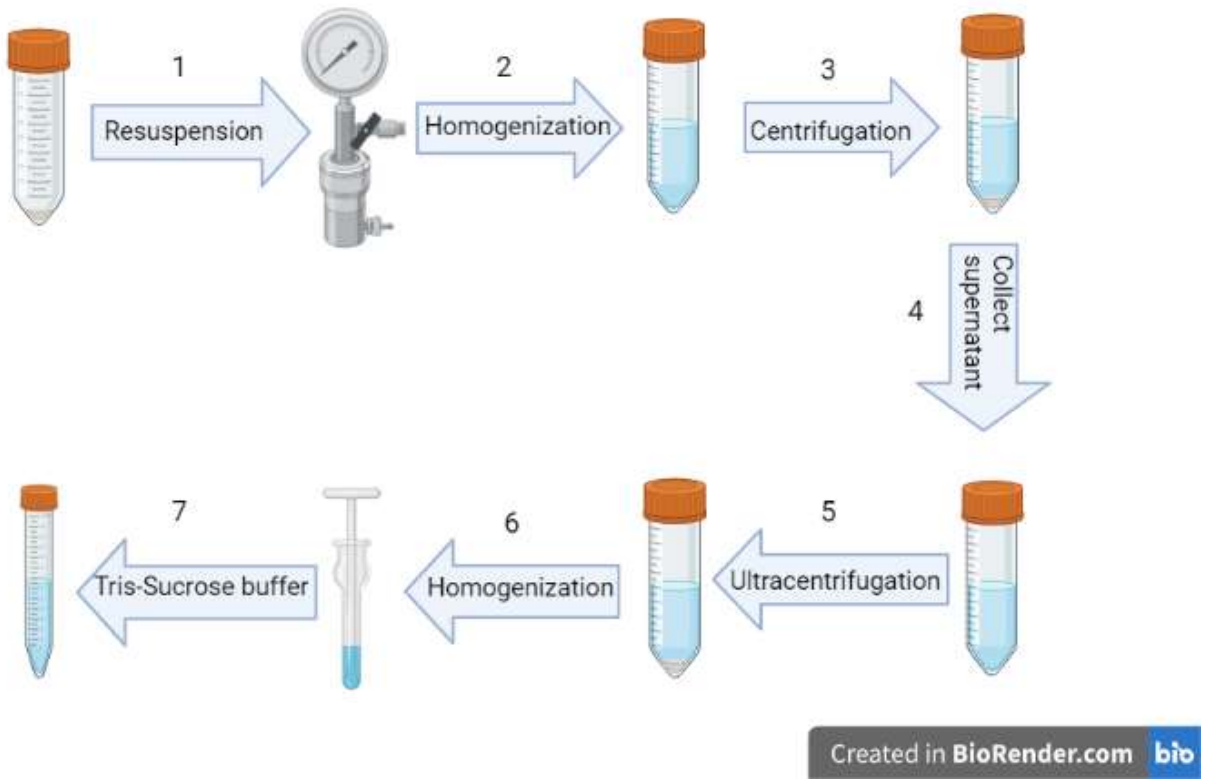


Figure 3.2. Membrane preparation*. Sf9 cell pellet was resuspended in homogenization buffer (1) and homogenized using a nitrogen cavitation device (2). The sample was centrifuged (3), the supernatant collected (4) and then ultracentrifuged (5). The pellet was weighed, and it was homogenized (6). The membrane was resuspended by addition of the volume of tris-sucrose buffer needed for the preferred membrane concentration (7). The membrane was stored at -80°C.*

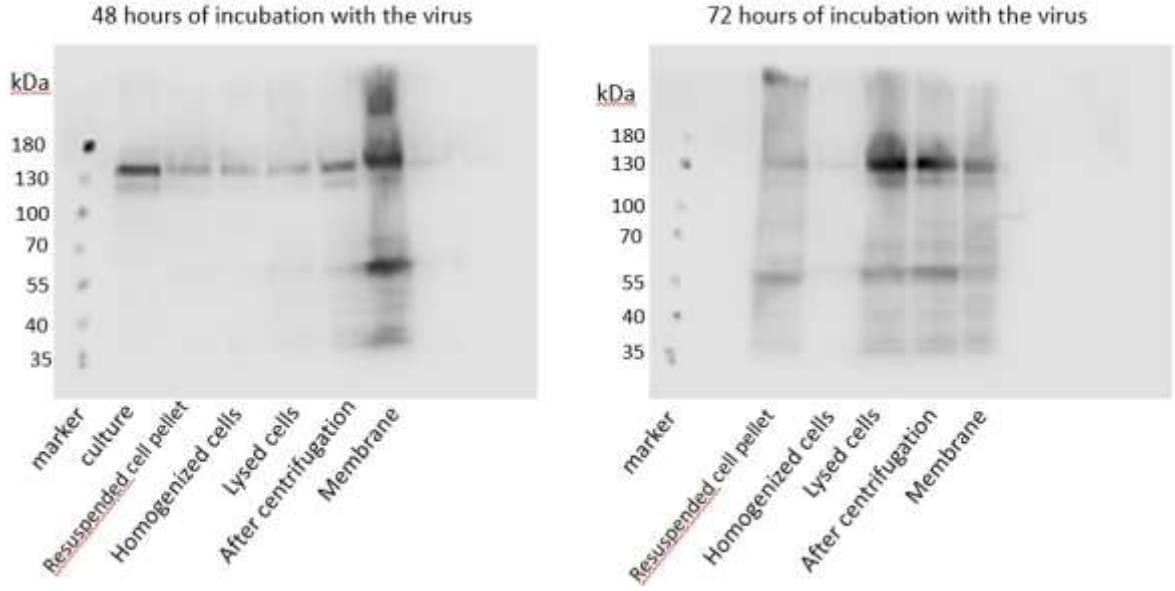


Figure 3.3. MRP4 is cleaved after 72 hours incubation with the baculovirus*. Western blot of samples taken from different stages, beginning from Sf9 cell cultures until the membrane preparation, after 48 hours and 72 hours incubation of recombinant baculovirus for MRP4 expression. M4I-10 α-MRP4 primary antibody and α-rat-HRP as secondary were used for this Western blot. Results were obtained with chemiluminescence and were revealed in LI-COR C-DiGit Blot Scanner. 20µl sample was taken in each case (from 500ml culture, 15ml resuspended pellet, 30ml homogenized cells, 30ml lysed cells, 25ml post centrifugation and 5ml membrane)*

3.2 MRP4 solubilisation & purification

Membranes of *Sf9* cells expressing MRP4 were solubilised by incubation with 2.5% SMA2000 for 1 hour at room temperature. After ultracentrifugation, the supernatant, containing solubilised protein, was incubated overnight at 4°C with Ni^{2+}-NTA affinity resin. The mixture was transferred to a gravity flow column and the flow-through was collected, which contains the proteins that did not bind to the resin. Washes with low concentrations of imidazole remove weakly bound proteins. The protein of interest should remain bound and be eluted in the end using high concentrations of imidazole, in this case of 200mM. Samples of the pellet (containing insoluble protein), the soluble protein, the flow-through, the washes and the elutions were taken. A sample of the resin after the elution of the protein was also essential in order to see if any protein remained bound to the resin and was not eluted. The samples were analysed by SDS-PAGE and Western blot using α-MRP4 primary antibody and α-rat-HRP as secondary. Figure 3.4 A&B shows a SDS-PAGE of the steps of the process including soluble, insoluble and flow through samples and figure 3.4 C&D shows the same steps after western blot analysis. A band corresponding to MRP4 can be seen in the SDS-PAGE gel at the size of ≈150kDa, indicated with arrows (figure 3.4B). From densitometric analysis of this purification, the solubilisation efficiency and the loss of protein in the flow-through were calculated. The solubilisation efficiency was 72.4%, however 48.2% of the soluble protein was lost in the flow-through. The elution fractions were pooled together and the protein concentration estimated by using SDS-PAGE with a series of BSA standards analysed by densitometry. The protein concentration was 32.6μg/ml (figure 3.4E).

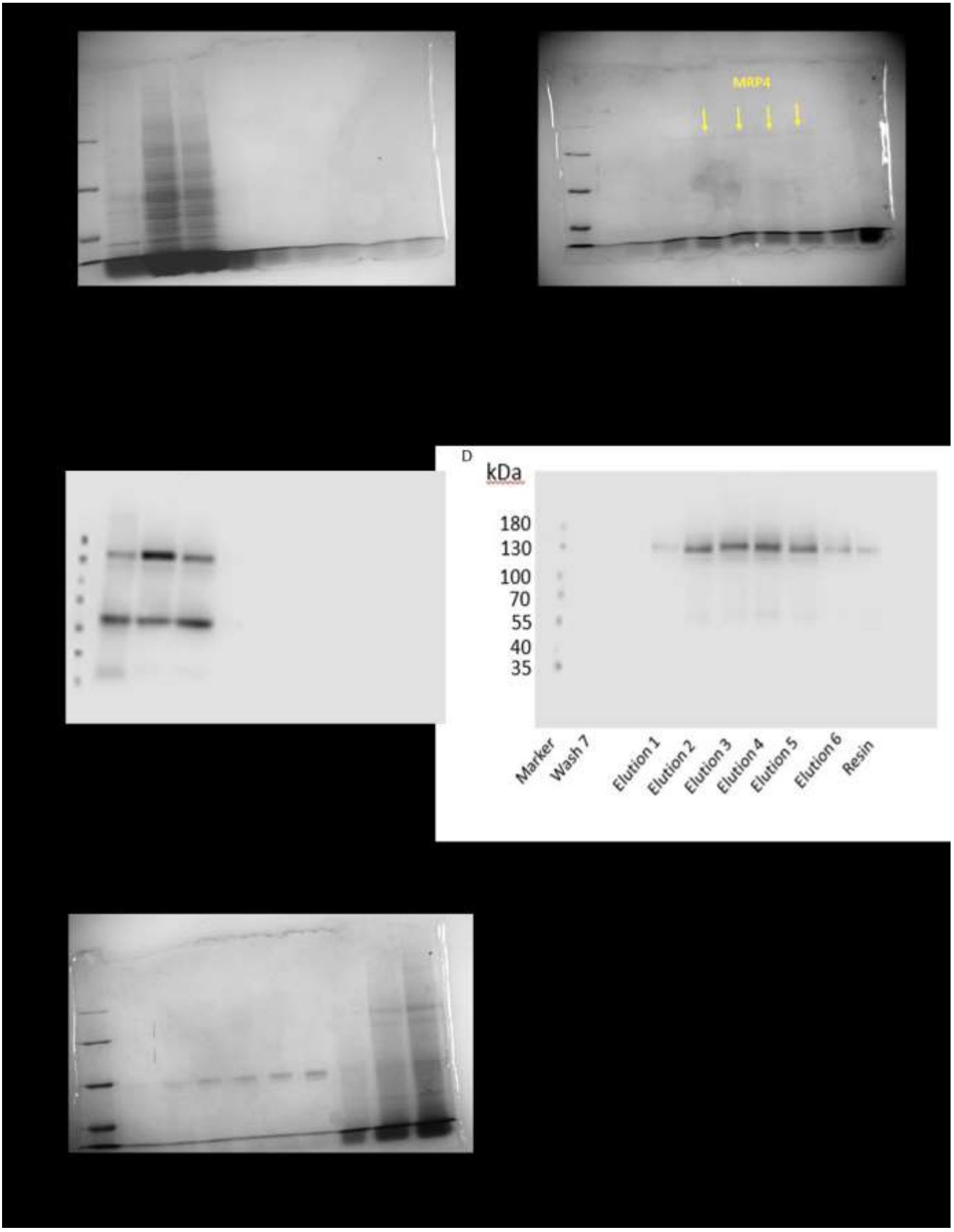

MRP4
D
kDa
180
130
100
70
55
40
35
Marker
Wash 7
Elution 1
Elution 2
Elution 3
Elution 4
Elution 5
Elution 6
Resin

Figure 3.4. Solubilisation and purification of MRP4. *A &B; SDS-PAGE showing steps of affinity purification of MRP4-SMALPs. MRP4 is indicated with arrows on B. C&D; Western blot showing the same steps of affinity purification of MRP4-SMALPs as in A&B. E; SDS-PAGE gel of BSA standards and different protein volumes. 20µl sample of each condition was taken (from 1ml resuspended pellet, 1ml solubilised protein, 10 ml of flow-through and washes, 500 µl elution fractions). MRP4 quantification was obtained after densitometry analysis and BSA standard curve. Western blots were detected with M4I-10 α-MRP4 primary antibody and α-rat-HRP as secondary. Results were obtained with chemiluminescence and were revealed in LI-COR C-DiGit Blot Scanner. SDS-PAGE were stained with Instant Blue.*

3.3 Optimization of membrane preparation and solubilisation to obtain higher protein yields

In order to try to obtain higher purified protein yields, more concentrated membranes were prepared, at 180mg/ml rather than the standardly used 60mg/ml (wet pellet weight). The same volume of these concentrated membranes was used as before, with the same concentration (2.5% (w/v)) of SMA2000 polymer, and all the volumes for washes and elutions kept as before. A Western blot showing all the steps of this purification is shown in Figure 3.5, where it can be seen that MRP4 was still able to be solubilised and purified. This modification to the procedure gave a final purified protein concentration of 260µg/ml (within the same total volume), which is almost 8 times higher than in the case of the membranes at 60mg/ml. However, it was found that the solubilisation efficiency in this case was only 52.1%, and 28.9% of this protein was lost in the flow-through (figure 3.5). Moreover, protein seems to remain stuck in the resin. It could maybe be eluted with high imidazole concentration, but that would risk contamination with other proteins that might have remained bound in the resin as well.

Furthermore, to remove high concentrations of imidazole later, it would risk protein precipitation as it has already happened when elutions were concentrated and imidazole was removed. Comparing the results of the membrane of 180mg/ml with the ones from the membrane of 60mg/ml, the solubilisation efficiency was higher in the case of the 60mg/ml membrane but this sample contained less protein to begin with. A lower proportion of solubilised protein was lost in the flow-through in the case of the more concentrated membrane, and it gave a higher concentration of protein in the end. For this reason, the membranes at 180mg/ml were chosen for future work, but the solubilisation needed to be optimized. Analysis of purity was done using ImageJ for densitometry analysis. In this case, the band of the protein of interest is compared to the rest of the proteins bands. The intensity of the bands of all the proteins summed, is consider 100 and the percentage of the band of the protein of interest to all the protein bands summed up, is the purity. MRP4 purity was found to be 45-50% pure after solubilisation with SMA2000 and purification.

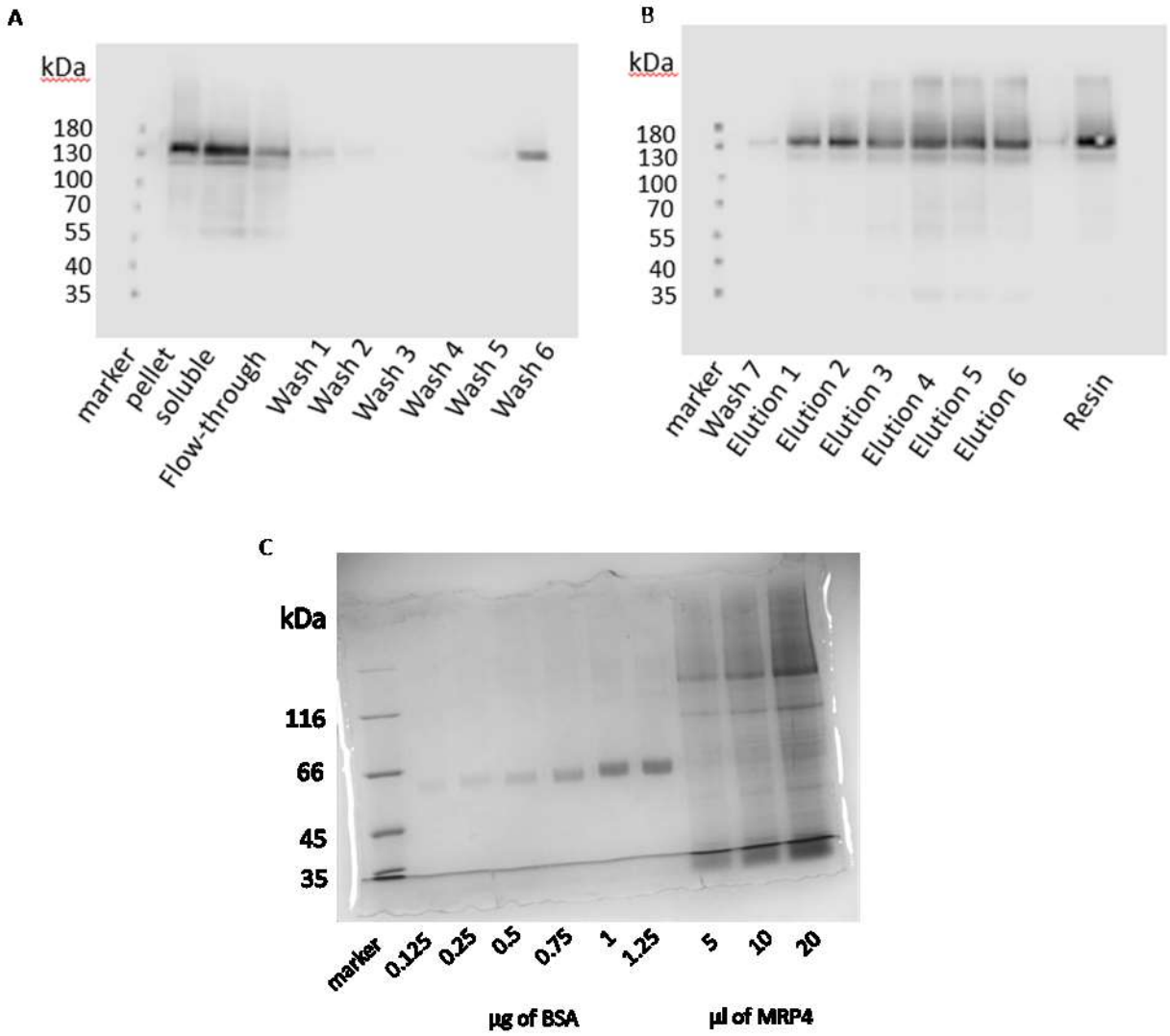


Figure 3.5. Purification of MRP4 from more concentrated membranes*. A&B. Western blot of pellet, solubilised protein, washes and elutions of MRP4 purified from 180mg/ml membranes (from 1ml resuspended pellet, 1ml solubilised protein, 10 ml of flow-through and washes, 500 µl elution fractions). M4$_I$-10 α-MRP4 was used as primary antibody and α-rat-HRP as secondary. Results were obtained with chemiluminescence and were revealed in LI-COR C-DiGit Blot Scanner. C. SDS-PAGE gel of BSA standards and different protein volumes. MRP4 quantification was obtained after densitometry analysis and BSA standard curve.*

To try to improve either the solubilisation efficiency and/or the binding to the Ni-NTA resin, a range of different conditions were tested, varying concentration of SMA2000, concentration of NaCl, addition of arginine and volume of resin. For solubilisation, 2.5% (w/v) and 3% SMA2000 (w/v) were tried, the increased concentration of SMA was tested because of the lower solubilisation efficiency obtained with more concentrated membranes. The concentration of NaCl initially used for purification was 150mM. Here, a concentration of 500mM was also tested, as it was shown to improve the solubilisation efficiency of the channel KcsA from *E.coli* membranes (118). Addition of arginine has been reported to be beneficial during chromatography, to prevent column spoiling by free excess SMA (119) so addition of 50mM arginine was also tested. Finally, a larger resin volume per ml of solubilised protein was tested to see if more protein would bind and therefore less lost in the flow-through, thus, 100µl resin, 150µl resin and 200µl resin per ml of solubilized protein were tested (Table 3.1). All the conditions were tested on a small scale, using 1ml of membrane, and the results for solubilisation efficiency are shown in figure 3.6. None of the conditions made any statistically significant difference to solubilisation efficiency.

SMA2000	2.5%			3%		
No addition	100 µl	150 µl	200 µl	100 µl	150 µl	200 µl
0.5M NaCl	100 µl	150 µl	200 µl	100 µl	150 µl	200 µl
50mM arginine	100 µl	150 µl	200 µl	100 µl	150 µl	200 µl

Table 3.1. Different conditions tested for MRP4 solubilization and resin binding. *For 2.5% (w/v) and 3% (w/v) SMA, addition of 50mM arginine or total concentration of 0.5M NaCl were applied. 3 different bed volumes of Ni-NTA resin were tried in each case, 100µl, 150µl and 200µl of resin per ml of solubilized protein.*

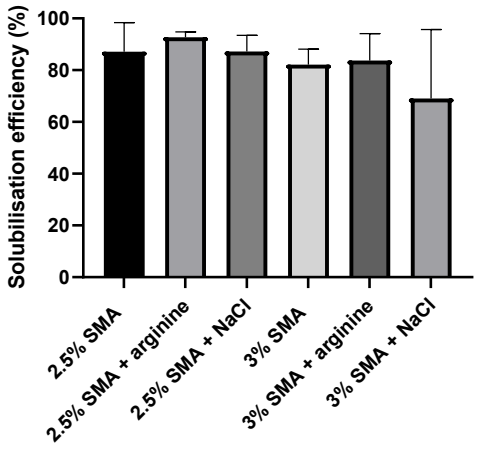


Figure 3.6. Solubilization efficiency of MRP4 using SMA2000 in different conditions. *Sf9 membranes containing MRP4 (180mg/ml) were solubilised with 2.5% (w/v) or 3% (w/v) SMA, with or without 50mM arginine or 500mM of NaCl. The solubilisation efficiency was measured by densitometric analysis of Western blots showing the soluble protein and pellet (insoluble protein). $M4_I$-10 α-MRP4 was used as primary antibody and α-rat-HRP as secondary. Results were obtained with chemiluminescence and were revealed in LI-COR C-DiGit Blot Scanner. Data are mean ± st. dev., n ≥ 2, (biological repeats). Data were analysed by ANOVA with a Tukey post-hoc test and there were no significant differences.*

Concerning the loss of protein in the flow-through, there was a lot of variability in the data, so no conclusion could be drawn (figure 3.7). Further repeats would be needed to confirm this. Addition of arginine made no difference at all, and addition of extra NaCl seemed to lead to worse resin binding, but again further repeats would be needed to confirm this. Thus, none of the conditions checked seemed to make a significant improvement for optimization of the purification or the solubilisation.

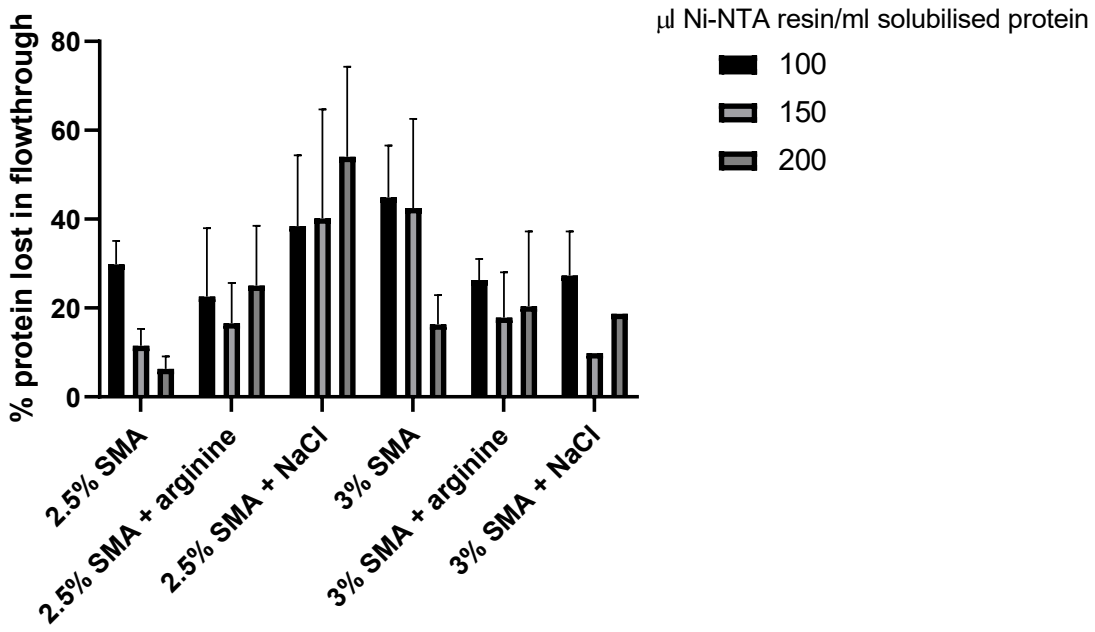


Figure 3.7. Protein loss in the flow-through with various buffer conditions*. MRP4 solubilised from Sf9 cell membranes was mixed with Ni-NTA resin overnight, shaking at 4°C. It was then poured into a gravity flow column and the flow-through collected. The loss of protein in the flow-through was measured by densitometric analysis of Western blots showing the soluble protein and flow-through samples. M4$_I$-10 α-MRP4 was used as primary antibody and α-rat-HRP as secondary. Results were obtained with chemiluminescence and were revealed in LI-COR C-DiGit Blot Scanner. Data are mean ± st. dev., n ≥ 2 (biological repeats). Data were analysed by two-way ANOVA with a Tukey post-hoc test and there were no significant differences.*

3.4 Test of different SMA copolymers for optimization of solubilisation efficiency

In a different approach to optimize the solubilisation efficiency, a screening of four modified SMA polymers was performed. The polymers were SMA2625, SMA17352 and SMA1440. All of them are based on the styrene-maleic acid polymer but with a modification on some of the maleic-acid groups (figure 3.8, table 3.2). SMA2625 has a 1-propanol group, SMA17352 has both cyclohexanol and 2-propanol groups added and SMA1440 has a 2-butoxyethanol. The styrene:maleic acid ratio is 2:1 for SMA2000 and SMA2625, 1.5:1 for SMA1440 and 1.7:1 for SMA17352.

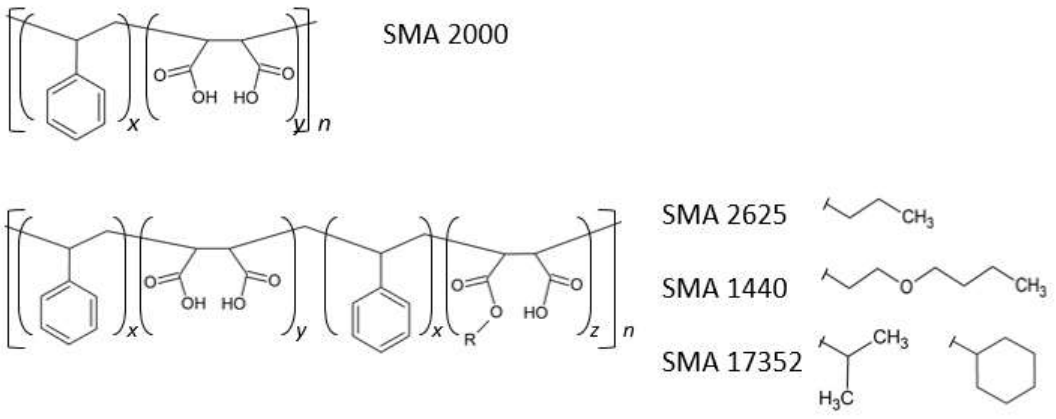


Figure 3.8. Structures of partially esterified SMA copolymers. *SMA 2000 and SMA 30010 are co-polymers of styrene and maleic acid. SMA 2625, SMA 1440 and SMA 17352 are partially esterified variants of SMA, with the ester moieties (R) n signifies the unit of polymerization, x the hydrophobic (styrene) groups and y the hydrophilic (2), adapted (permission asked and received for the use of this figure).*

Polymer	S:MA ratio (Cray Valley)	Modification	Mw (kDa)	Mn (kDa)	PDI
SMA 2000	2:1		7.5	3.0	2.5
SMA 2625	2:1	1-propanol	9.0	3.6	2.5
SMA 1440	1.5:1	2-butoxyethanol	7.0	2.8	2.5
SMA 17352	1.7:1	Cyclohexanol & 2-propanol	7.0	2.8	2.5

Table 3.2. Properties of partially esterified SMA copolymers. *The average ratio of styrene to maleic acid is shown. Also, the molecular weight of each copolymer. Mw is the average molecular weight of the polymers. The modifications are presented. The number average molecular weight (Mn) is the total weight of the polymer molecules divided by the number of molecules. The polydispersity index (PDI), is Mw/Mn and it is a measure of the distribution of the molecular weights (2), adapted (permission to use this table was asked and received)*

After solubilisation of the membranes, Western blot analysis of the soluble and pellet samples and densitometry analysis, the solubilisation efficiency was measured as 87% for SMA2000, while for SMA2625, SMA1440 and SMA17352 the solubilisation efficiency was 79%, 78% and 80% respectively (figure 3.9). Therefore these polymer variants offered no improvement over the standard SMA2000 for MRP4, which was in agreement with findings for these polymers with other protein targets (2), and since other proteins showed even greater loss of protein during purification than with SMA2000, no further work was undertaken with these polymers.

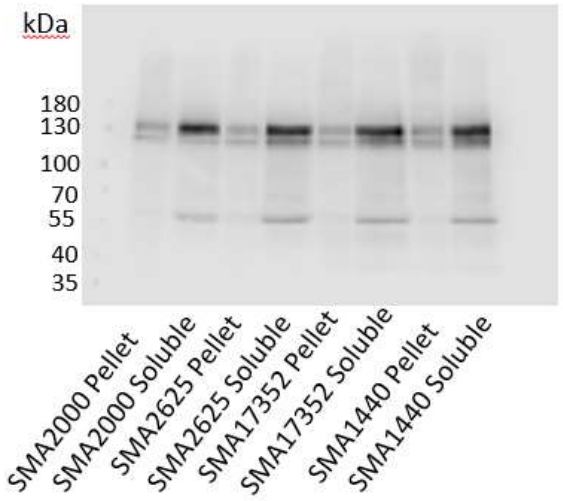


Figure 3.9. Western blot of MRP4 solubilization using modified copolymers. *Membranes of 180mg/ml were incubated with the different polymers for one hour at room temperature and after ultracentrifugation, 20µl samples were taken from the pellet (post solubilisation) and soluble protein (1ml pellet, 1ml solubilised protein). M4$_I$-10 α-MRP4 was used as primary antibody and α-rat-HRP as secondary. Results were obtained with chemiluminescence and were revealed in LI-COR C-DiGit Blot Scanner.*

3.5 Summary

- MRP4 was expressed in Sf9 cells, with best conditions from the ones tested found to be 48 hours of incubation.
- MRP4 can be effectively solubilised with SMA2000, with a solubilisation efficiency of 72.4%, and purified by affinity chromatography, but the final yield of purified protein was low.
- Using more concentrated membranes at 180mg/ml (wet pellet weight) instead of the standard 60mg/ml led to slightly lower solubilisation efficiency, however it gave a much better yield of purified protein, achieving concentrations of 260 µg/ml compared to 32.6 µg/ml. The protein was 45-50% pure.
- To improve solubilisation efficiency and resin binding efficiency a range of different conditions were tested: addition of NaCl, addition of arginine and different resin volumes per ml of solubilised protein. None of these conditions showed any significant difference in solubilisation efficiency or resin binding. However, the large variation in results with resin binding suggest that further investigation into this is needed.
- A range of novel, partially esterified, SMA polymers did not show any improvement for solubilisation efficiency.

4. Detergent mediated solubilisation and purification of MRP4 from Sf9 insect cells

As an alternative approach to overcome the limitations of conventional detergents, CALIXAR has developed novel detergents, designed to solubilise membrane proteins efficiently and offer stability. Some of their detergents were tested on MRP4. The detergents chosen to be tested, were selected according to the structure difference of the molecules. For example, CALX7, CALX8, CALX9, CALX10 and CALX12 are based on the same structure differing on the length of the hydrocarbon chain. DDG is based on a different structure as well as DDTAC (figure 4.1).

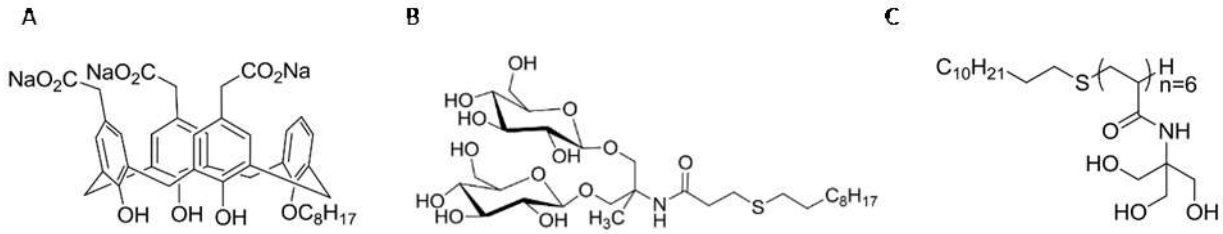
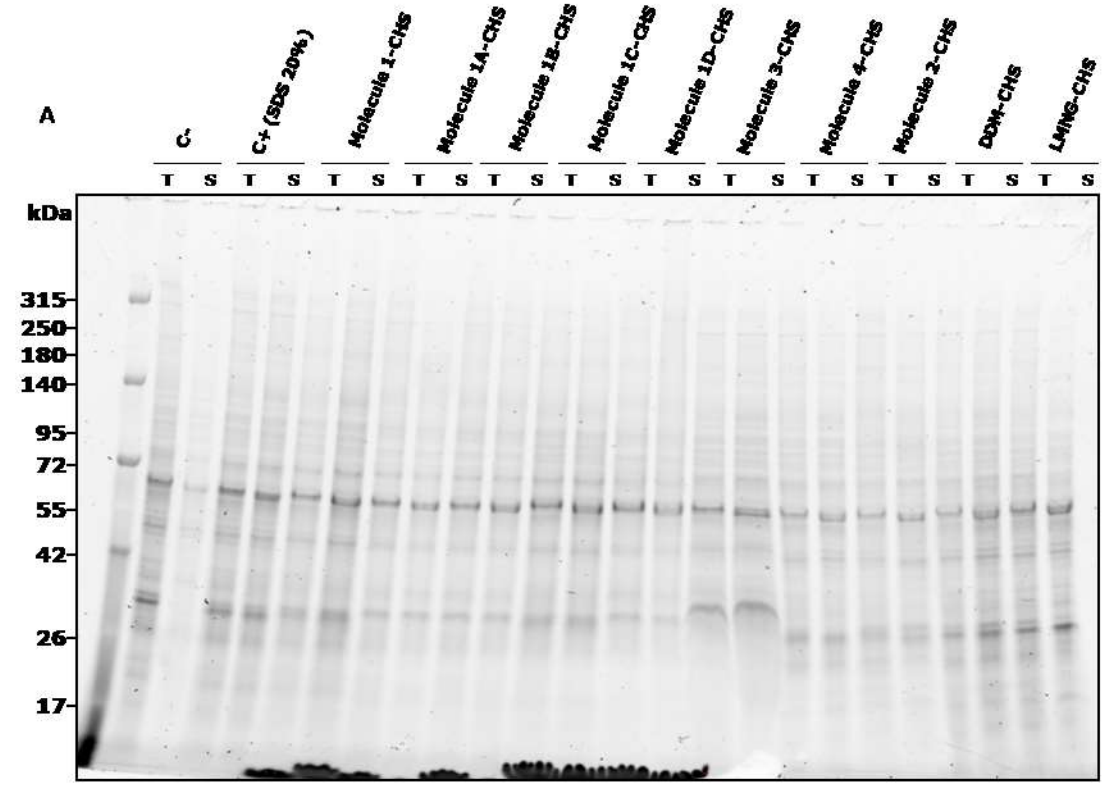


Figure 4.1: CALIXAR detergents structures. *A; CALX8 is an ionic calixarene based detergent with three polar carboxylic groups and an octyl chain. CALX7, CALX9, CALX10 and CALX12 have the same structure differing at the hydrocarbon group length. B; DDG is a non-ionic amphiphilic detergent with a diglucose polar group and decyl chain. C; DDTAC, adapted from https://calixar.com/ (visited 30/09/2022, 13:45)*

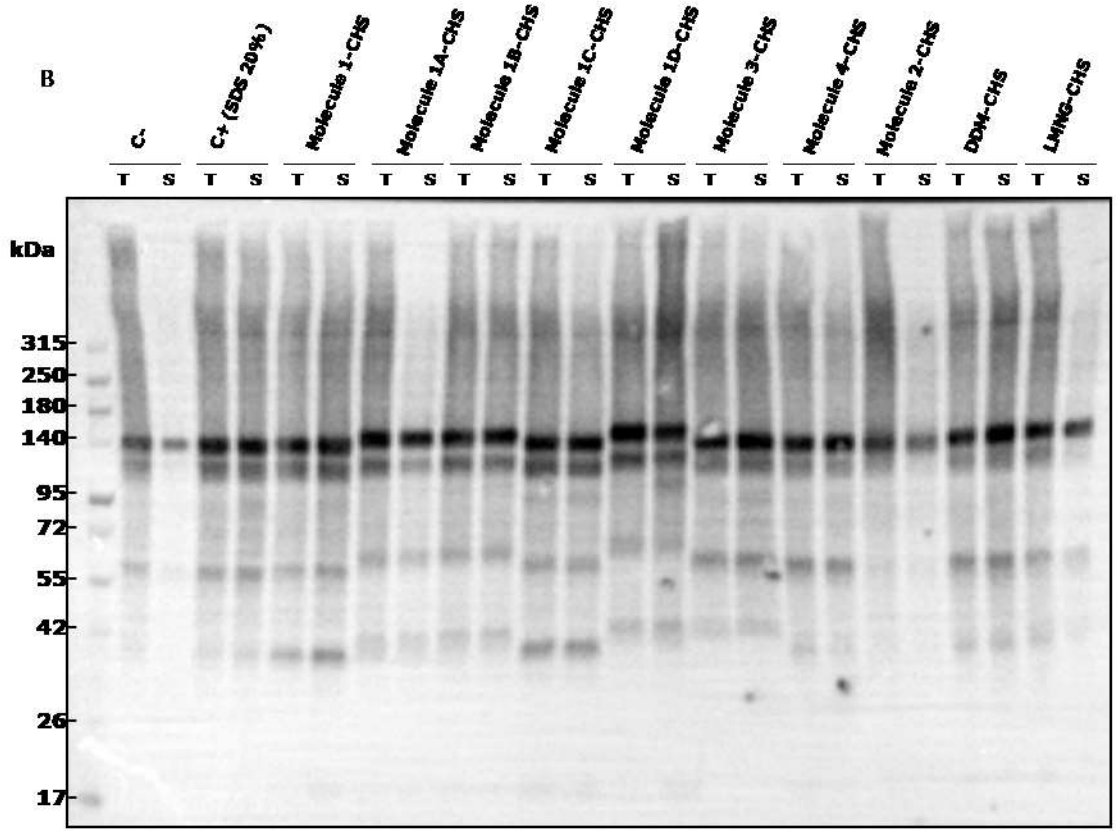
4.1 Small scale solubilisation test

Whereas solubilisation with polymers typically uses a concentration of membranes based upon wet pellet weight, it is more usual to measure the total protein concentration of the membranes when using detergents. Therefore, the protein concentration of the membrane was determined using a BCA assay, and membranes containing 10 mg/ml total protein were used for small scale solubilisation tests. The detergents all contained CHS as an additive. This is because CHS has been reported to help stabilised detergent solubilised proteins (119-122). Solubilisation with each detergent was carried out for 2 hours at 4 °C, then soluble and insoluble material separated by ultracentrifugation and analysis by Western blot (Figure 4.2). The band corresponding to MRP4 can be clearly seen at approximately 140 kDa. Effective solubilisation can be seen with each detergent. Notably for some of the detergents, the migration of MRP4 appears to be affected, for example CALX12-CHS, where the MRP4 band appears higher than in the control sample. It is hypothesised that some of this detergent remains bound to the protein and is not fully displaced by the SDS or even the sample viscosity, thus the way it migrates can affect the migration profile on a gel.

Densitometric analysis of the Western blots is shown in Figure 4.3. It was found that almost all detergents tested solubilised MRP4 from Sf9 cell membranes efficiently at percentages greater than 70%, except for CALXcod9-CHS that gave a solubilisation efficiency of 58% and LMNG-CHS that gave 61.2%.



SDS-PAGE, 4-15% acrylamide gel
Bio-Rad Criterion Stain-Free gel



SDS-PAGE, 4-15% acrylamide gel
WB anti-MRP4 (1:100), anti-rat HRP (1:5000)

Figure 4.2: SDS-PAGE and Western blot of MRP4 solubilisation test with CALIXAR detergents. *Membranes (10 mg/ml total protein) were solubilised with a range of different detergents for 2 hours at 4 °C. Samples were ultracentrifuged (150,000g 15min 4 °C). A; The supernatants containing solubilised protein were run on SDS-PAGE 4-15% alongside samples of the total membrane. B; For the western blot α-MRP4 M4I-10 and α-rat-HRP were used. The stain-free Biorad SDS-PAGE was imaged using a BioRad chemidoc and Western blot was imaged using a Chemi-Hi resolution. T stands for total protein and S stands for solubilised protein. The solubilisation efficiency was calculated after image analysis with ImageLab. These images are representative of two independent repeats.*

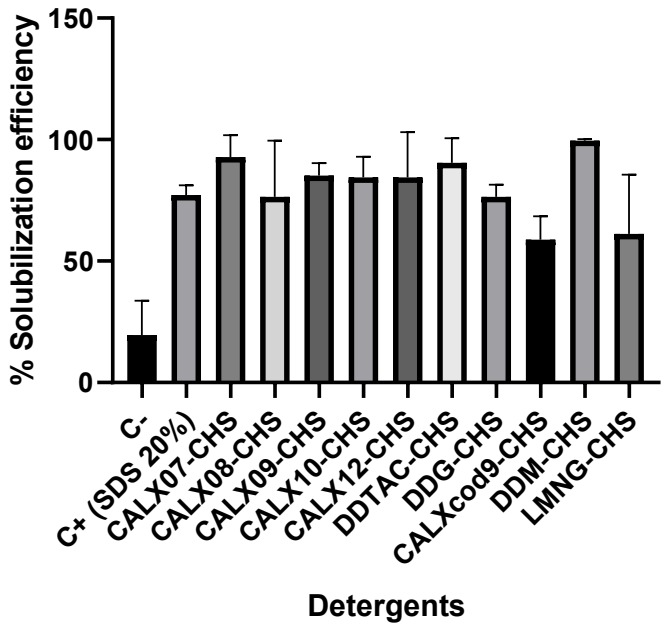


Figure 4.3. Solubilisation efficiency of the different detergents tested for MRP4. *Densitometric analysis of Western blots as in Figure 4.2. Data are mean ± SD, n=2 (biological repeats). C- indicates the negative control (no detergent added), C+ indicates the positive control which is SDS 20%*

To check if the solubilised protein remained well folded, does not aggregate and retains its conformation, the samples of total protein and solubilised protein were analysed by clear native PAGE. For this, the gel and the buffers do not contain SDS, so the proteins are not denatured or all highly negatively charged so they move according to their molecular weight and overall charge. The samples were run and afterwards Western blots were performed with α-MRP4 and α-rat-HRP (Figure 4.4). The proteins migrated well into the gel, and all migrate at approximately the same size, with some variation likely due to the varying micelle sizes. This suggests they all retain their conformation, are well-folded and they do not aggregate.

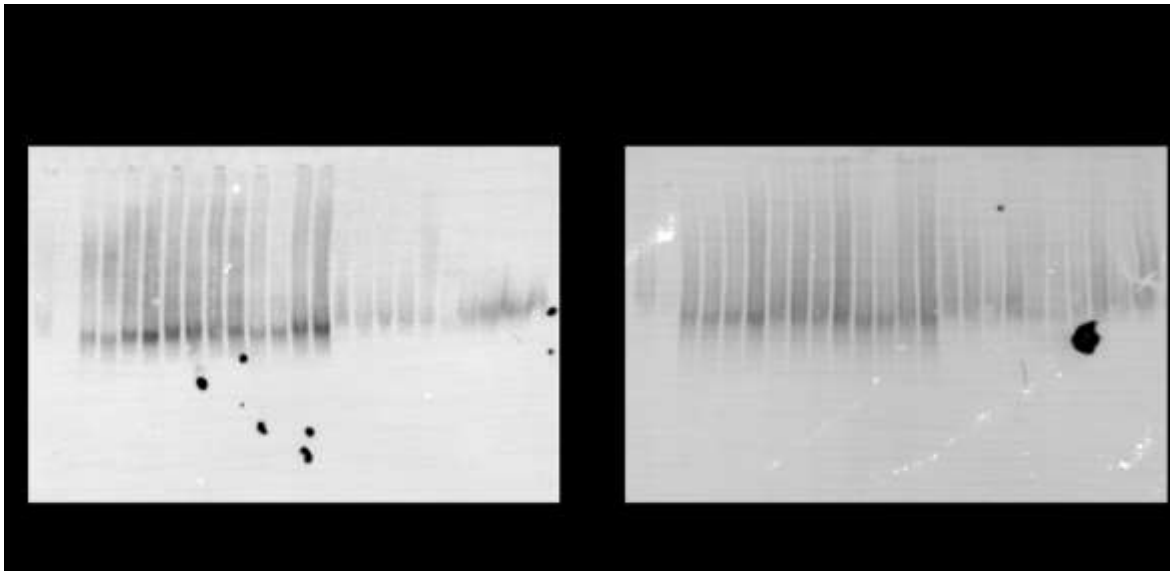
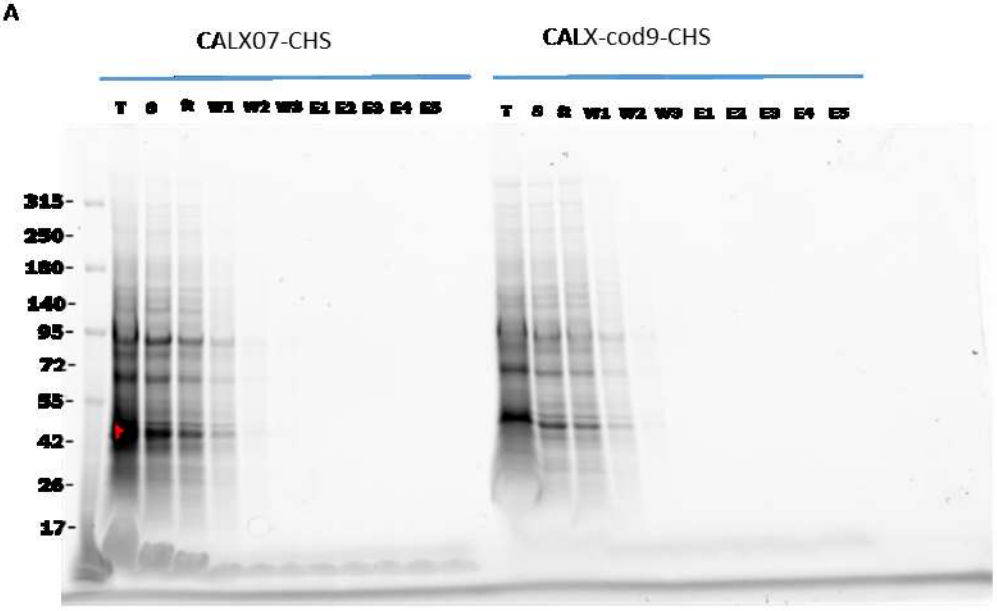


Figure 4.4: Native PAGE and Western blot of MRP4 solubilisation test with CALIXAR detergents. *Samples of total (T) and solubilised (S) protein were run, without being denatured, on gradient native gels (4-15% acrylamide), transferred to Western blots and probed using α-MRP4 M4I-10 and α-rat-HRP. Chemiluminescence was imaged using a Chemi-Hi resolution of Bio-rad. A and B contain samples from two independent repeats of the solubilization.*

4.2 Small scale purifications

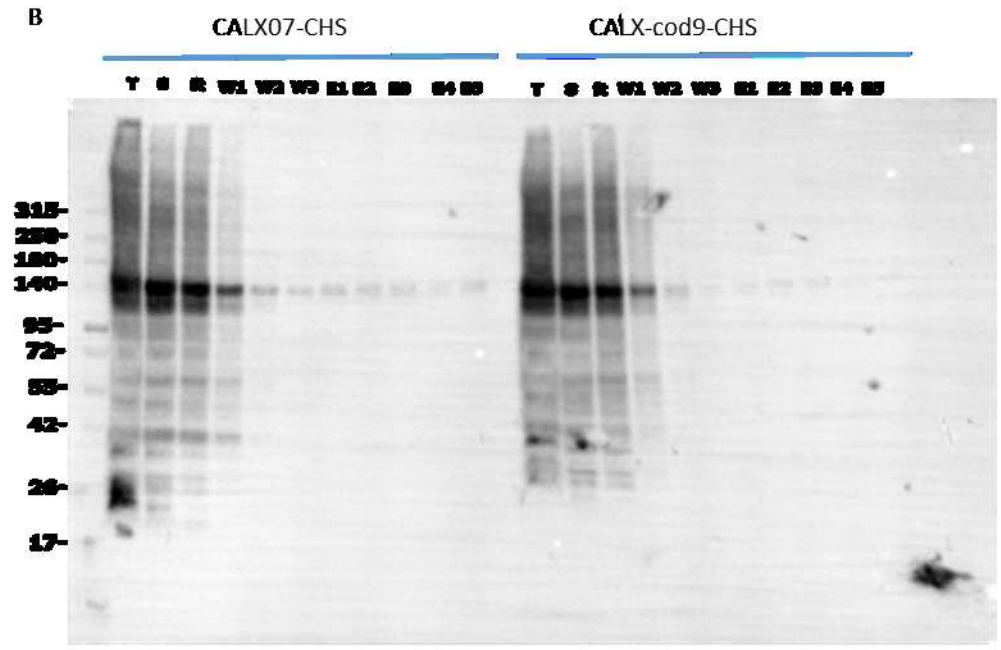
From the results obtained with the small scale solubilisations, a subset of detergents were chosen to proceed to small scale purifications. The detergent choice was based on choosing the detergent that gave the best solubilisation efficiency out of each group of chemical structure for the CALIXAR detergents (CALX07-CHS, CALX-cod9, DDTAC-CHS and DDG-CHS), alongside the two common detergents DDM-CHS and LMNG-CHS as controls. Following solubilisation, MRP4 was purified using Ni-sepharose, and analysed by SDS-PAGE and Western blot (Figure 4.5). In the Western blots (Figure 4.5 B, D, F) the band corresponding to MRP4 can be clearly seen at approximately 140kDa in the total sample. All of the detergents show some MRP4 in the elution fractions, although with varying intensity. DDG-CHS shows the most intense elution bands, followed by DDTAC-CHS, whilst CALX-cod9 shows the least intense bands. However, most notably it was found that there was a massive loss of protein in the flow-through with all the detergents, even for the ones with the best elution fractions. These results suggested that optimization of loss in the flow-through was essential.

A



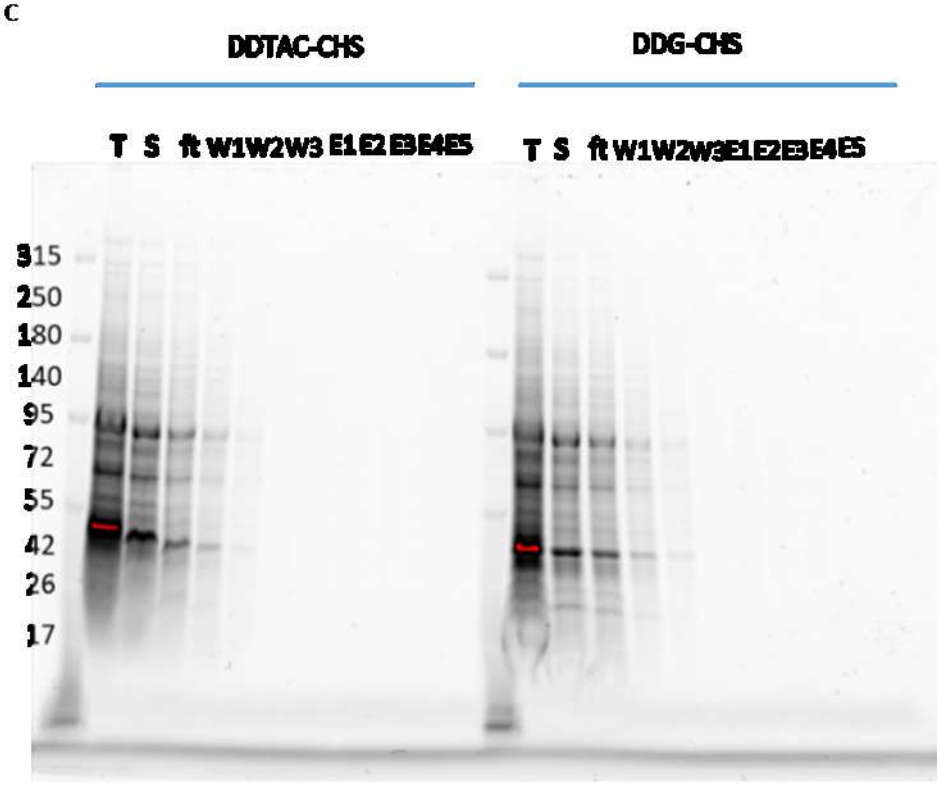
SDS-PAGE, 4-15% acrylamide gel
Bio-Rad Criterion Stain-Free gel

B

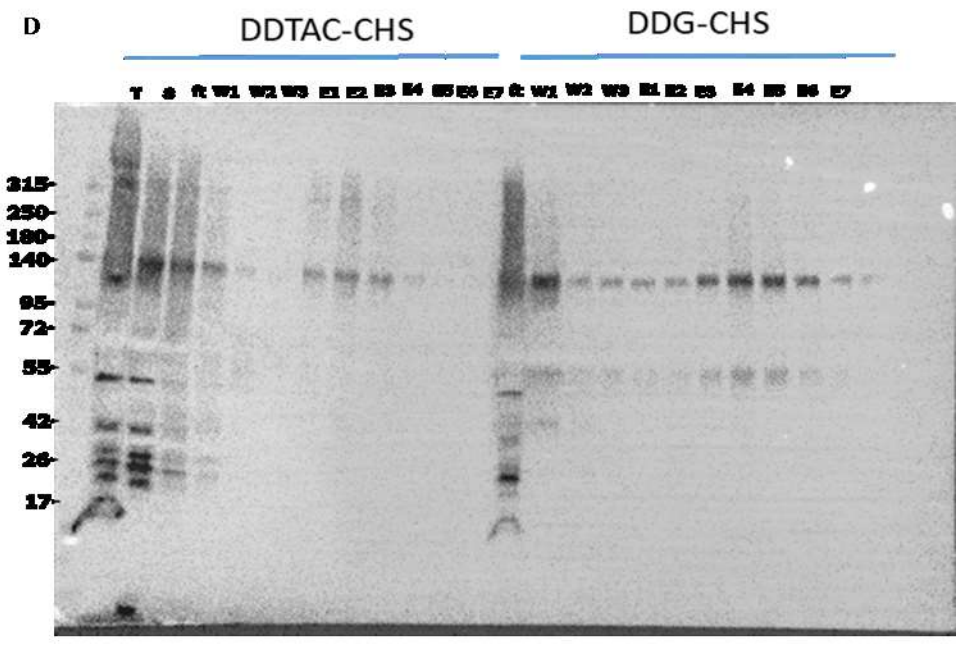


SDS-PAGE, 4-15% acrylamide gel
WB anti-HRP4 (1:100), anti-rat HRP (1:5000)

C

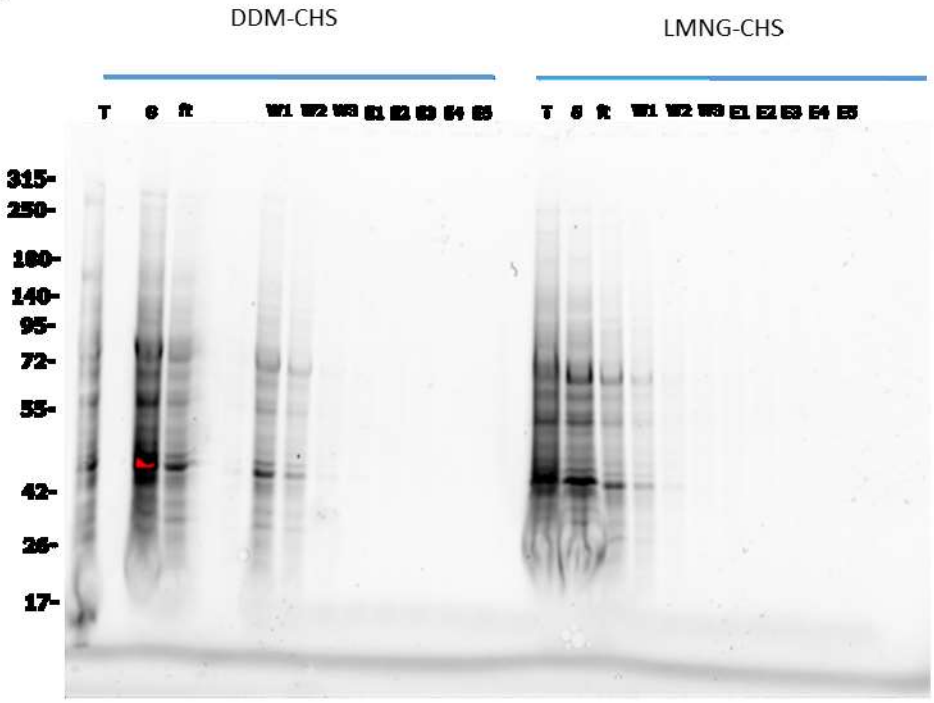


D



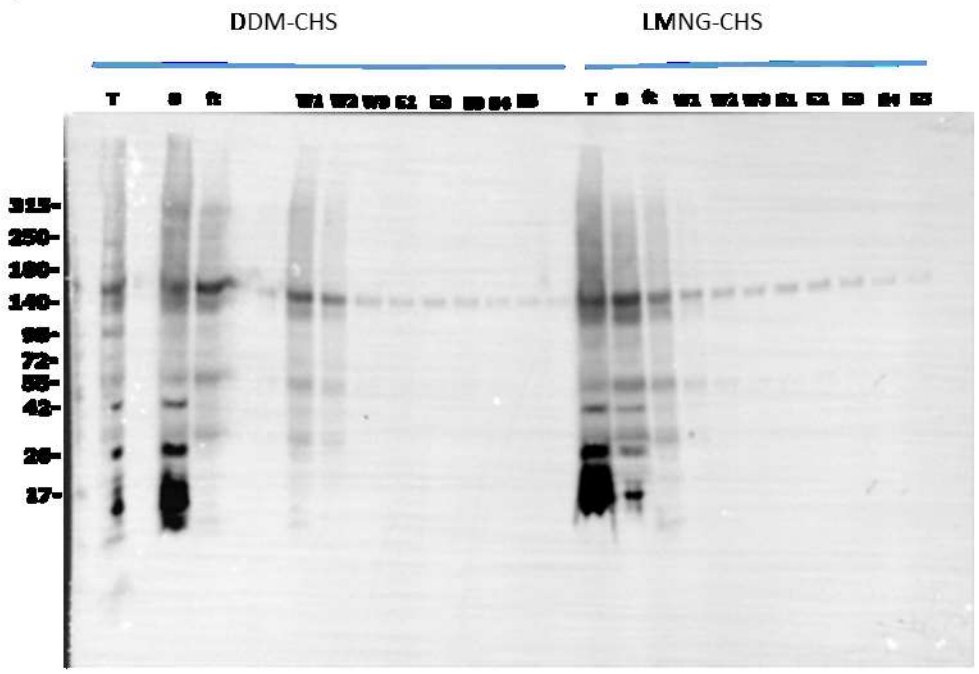
SDS-PAGE, 4-15% acrylamide gel
WB anti-HRP4 (1:100), anti-rat HRP (1:5000)

E



SDS-PAGE, 4-15% acrylamide gel
Bio-Rad Criterion Stain-Free gel

F



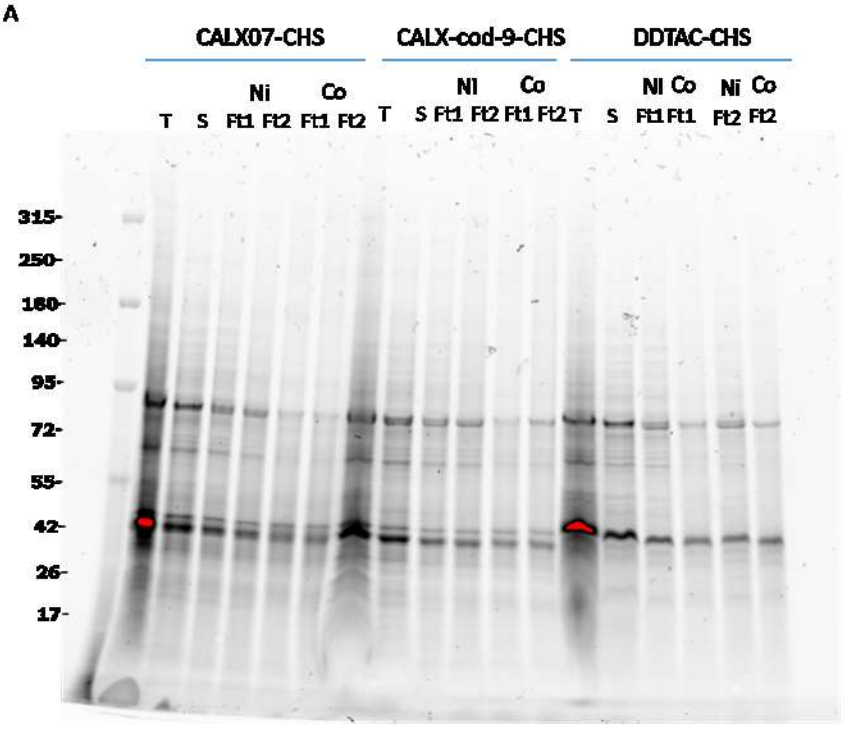
SDS-PAGE, 4-15% acrylamide gel
WB anti-HRP4 (1:100), anti-rat HRP (1:5000)

Figure 4.5: SDS-PAGE and Western blots of MRP4 small scale purifications. *Following solubilisation with A&B) CALX07 or CALX-cod9, C&D) DDTAC-CHS or DDG-CHS, E&F) DDM-CHS or LMNG-CHS, samples were purified using Ni^{2+}-sepharose affinity chromatography. The samples were run on a stain-free gel (A, C, E) with gradient of acrylamide 4-15% and also transferred to a western blot (B, D, F) and probed with α-MRP4 and α-rat-HRP were used. Gels were imaged using a BioRad chemidoc and Western blot the Chemi-Hi resolution was used. T stands for total protein and S stands for solubilised protein. W means Wash and E means elution. Ft signifies flow-through, W stands for wash and E for elution.*

4.3 Optimization of loss in the flow-through

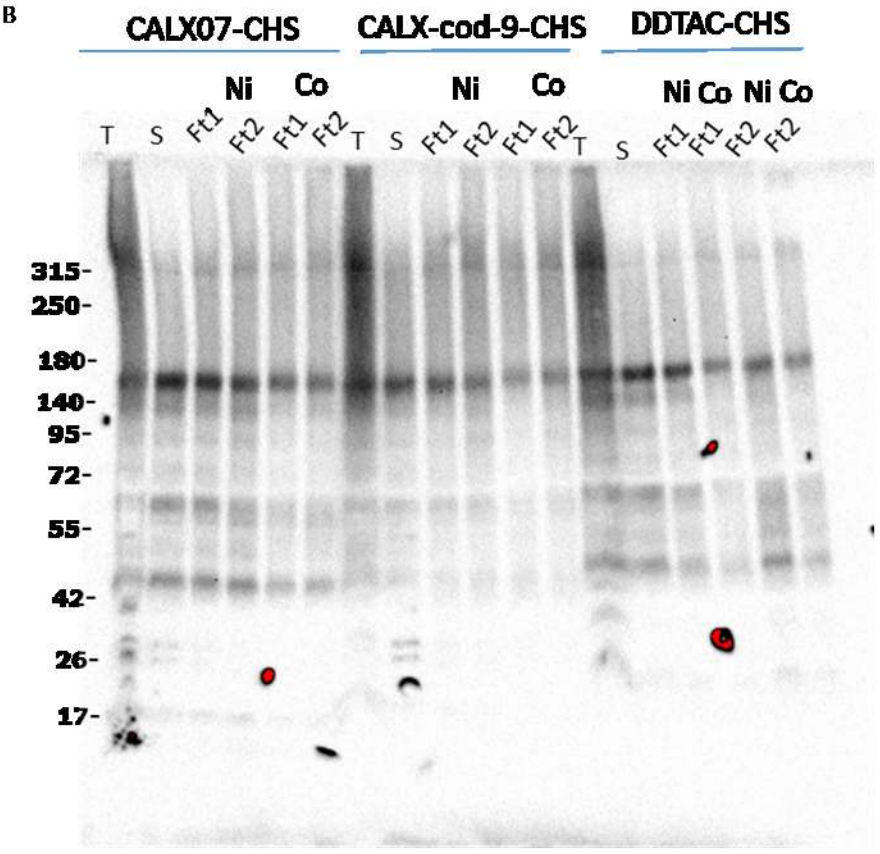
Efforts to decrease the loss of protein in the flow-through were important at this point. For this reason, different times of incubation of the solubilised protein and the resin were tested, 1 hour and overnight as well as testing Cobalt resin in addition to nickel. Samples of the flow through after 1 hour or overnight, with the two different types of resin were analysed by Western blot as shown in Figure 4.6. These Western blots were analysed with Image Lab and the results are shown in Table 4.1. It seems that when Cobalt resin was used less MRP4 protein was lost in the flow-through for all the tested detergents. However, the time of incubation of the protein with the resin did not make a significant difference in terms of loss. In the end, the best results for resin binding were seen for DDG-CHS and LMNG-CHS with Cobalt and at least 1 hour of incubation of the protein with the resin. Consequently, DDG-CHS and LMNG-CHS were chosen for further experiments.

A



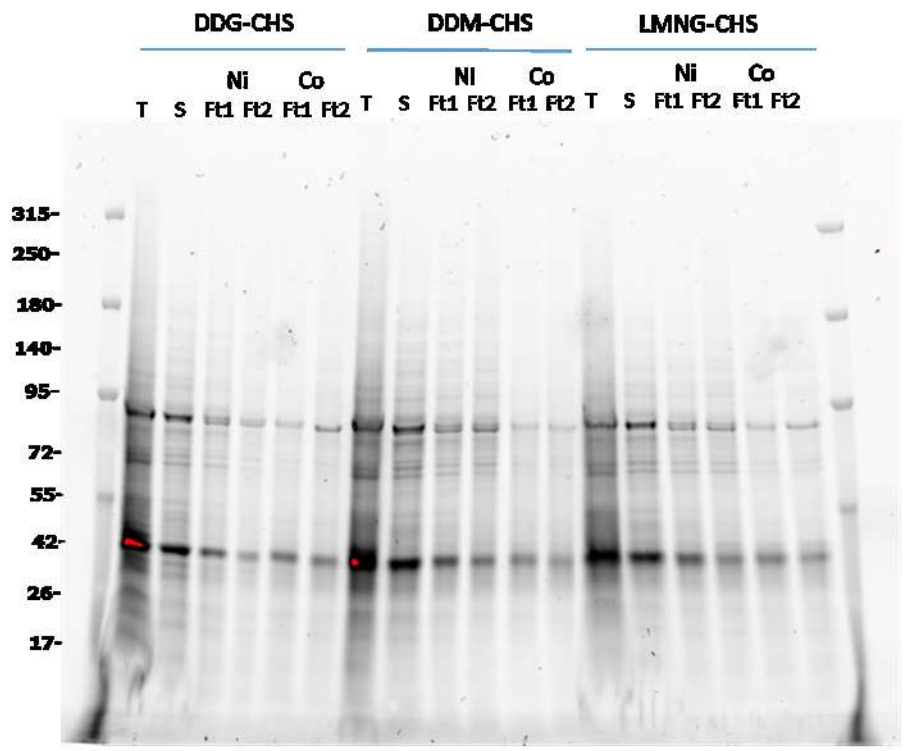
SDS-PAGE, 4-15% acrylamide gel
Bio-Rad Criterion Stain-Free gel

B



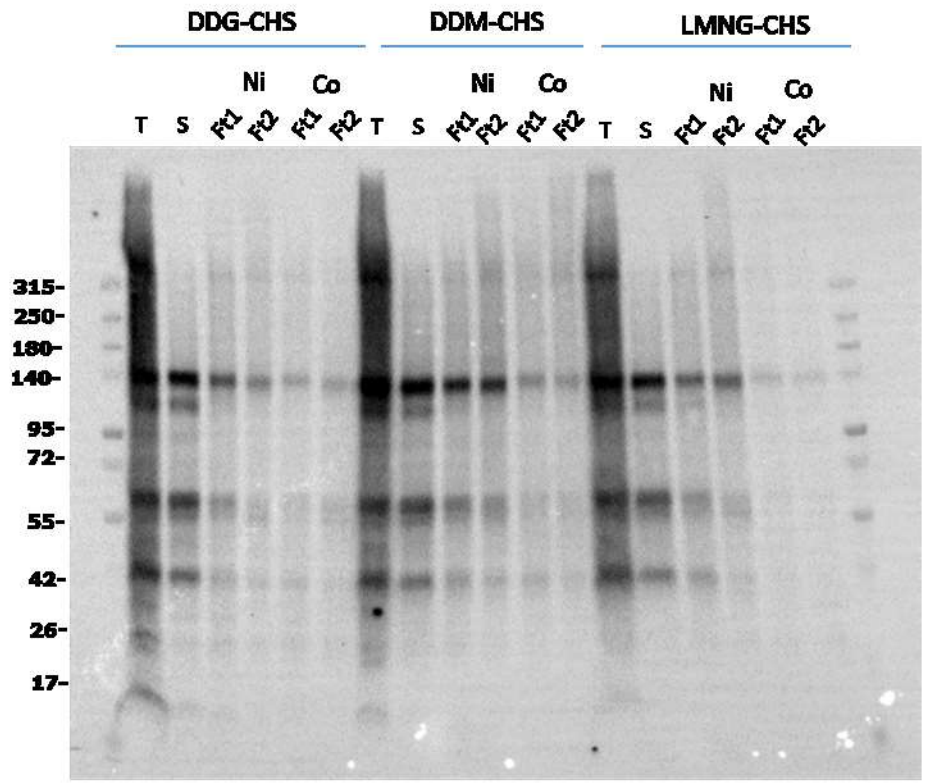
SDS-PAGE, 4-15% acrylamide gel
WB anti-MRP4 (1:100), anti-rat HRP (1:5000)

C



SDS-PAGE, 4-15% acrylamide gel
Bio-Rad Criterion Stain-Free gel

D



SDS-PAGE, 4-15% acrylamide gel
WB anti-MRP4 (1:100), anti-rat HRP (1:5000)

Figure 4.6: SDS-PAGE and Western blots for optimisation of loss in the flow-through. *Following solubilisation with CALIXAR detergents or DDM-CHS or LMNG-CHS were mixed with either Ni^{2+}-sepharose or Co^{2+}-sepharose for 1 hour or overnight. They were transferred to a gravity flow column and the flow through was collected. Samples were run on a stain-free gel (A, C) with gradient of acrylamide 4-15%, and transferred to a western blot (B, D) and probed with α-MRP4 and α-rat-HRP. Gels were imaged using a BioRad chemidoc and for the Western blot the Chemi-Hi resolution was used. T stands for total protein and S stands for solubilised protein. Ft1 signifies flow-through after 1 hour of incubation of the protein with the resin and ft2 after overnight incubation. Co stands for Cobalt and Ni for Nickel.*

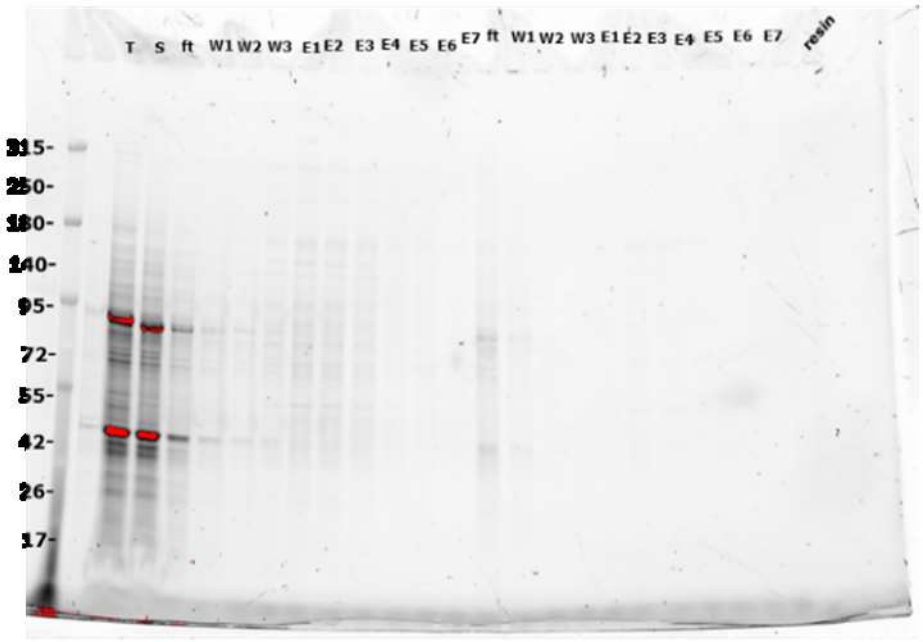
CALX07-CHS		CALXcod9-CHS		DDTAC-CHS		DDG-CHS		DDM-CHS		LMNG-CHS	
Solubilization efficiency		Solubilization efficiency		Solubilization efficiency		Solubilization efficiency		Solubilization efficiency		Solubilization efficiency	
100%		69.4%		73.5%		57.8%		64.5%		55.7%	
Ni		Ni		Ni		Ni		Ni		Ni	
ft1	ft2	ft1	ft2	ft1	ft2	ft1	ft2	ft1	ft2	ft1	ft2
85.5%	84.7%	80.4%	73.3%	70.8%	66.2%	48.3%	26%	62.3%	61.3%	50.6%	50.8%
Co		Co		Co		Co		Co		Co	
ft1	ft2	ft1	ft2	ft1	ft2	ft1	ft2	ft1	ft2	ft1	ft2
63.3%	67.9%	50.6%	64.3%	50.2%	46.2%	15%	26.7%	28.6%	27.2%	6%	1%

Table 4.1: Solubilisation efficiency and loss of MRP4 in the flow-through. *Following solubilisation with CALIXAR detergents (CALX07-CHS, CALXcod9-CHS, DDTAC-CHS, DDG-CHS) DDM-CHS or LMNG-CHS, and mixing with cobalt or nickel resin. Samples of solubilised protein and flow through were analysed by Western blot as shown in Figure 4.6. The solubilisation efficiency and % solubilised protein in the flow through were calculated using image processing with ImageLab. Ft1 indicates flow-through fraction after 1 hour of incubation of the resin with the solubilised protein, while ft2 indicates flow-through fraction after overnight incubation of the resin with the solubilised protein*

4.4 Optimising small scale purification of MRP4 with DDG-CHS solubilisation

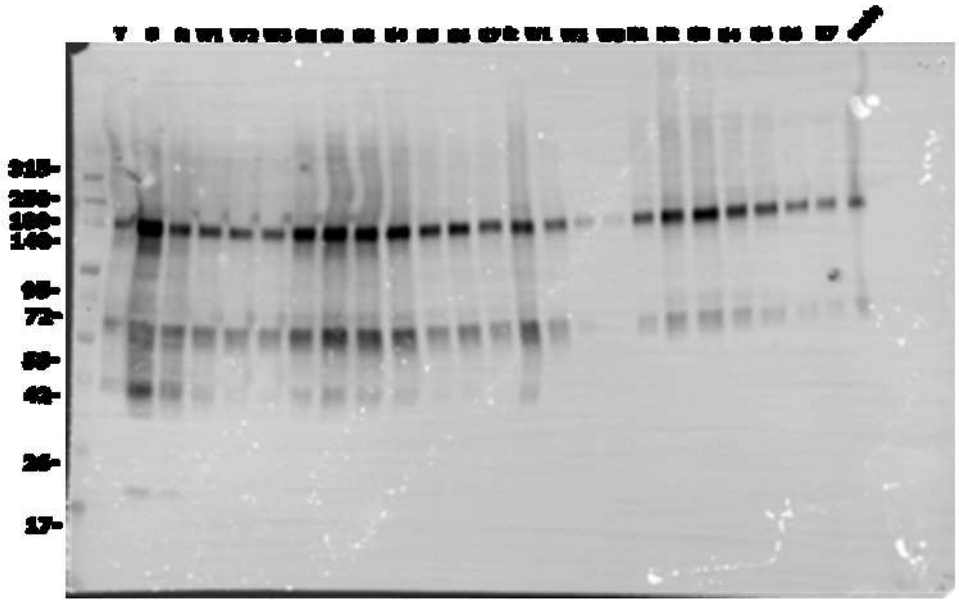
At first a small-scale solubilisation and purification was performed in order to optimize the procedure. For this, the Sf9 membrane was solubilized with DDG-CHS for 2 hours at 4°C. Following ultracentrifugation, the solubilised protein was incubated with Cobalt sepharose resin overnight at 4°C. Washes and elution fractions were collected next morning. Afterwards, cycles of concentration and dilution (using centrifugal filter units) to remove the imidazole were performed and the protein was mixed with Cobalt resin again and incubated at 4 °C for 1 hour. Another round of purification was carried out in order to try to improve the level of protein purity. Samples taken at each step were run on SDS-PAGE and western blot (Figure 4.7). The band corresponding to MRP4 is clearly visible on the Western blot (Figure 4.7B) and it appears that there is less background following the second round of purification. However, nothing is visible in the elution fractions on the stain free gel (Figure 4.7A). Then, cycles of centrifugation of the elution fractions in centrifugal filter units were done for protein concentration and the final sample was diluted and put on a gel for BSA assay. However, there were no bands after Coomassie staining so it was concluded that the protein concentration was too low. For this reason, scaling up was essential.

A



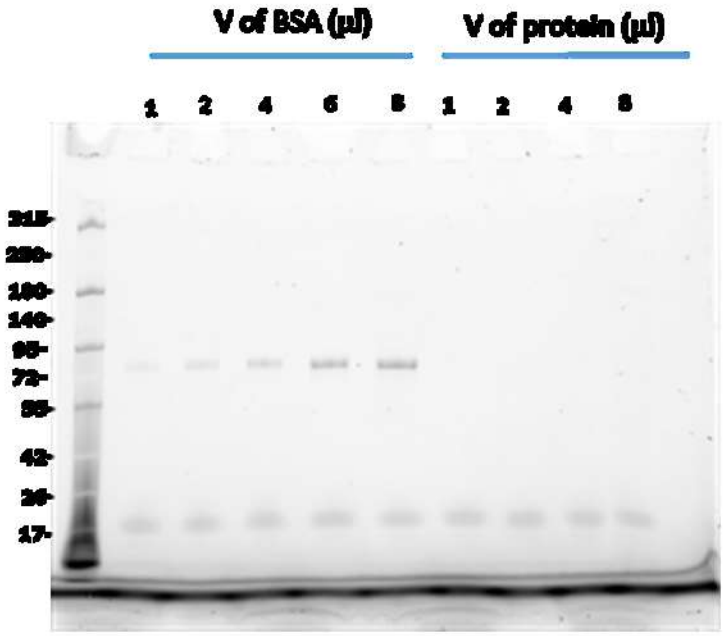
SDS-PAGE, 4-15% acrylamide gel
Bio-Rad Criterion Stain-Free gel

B



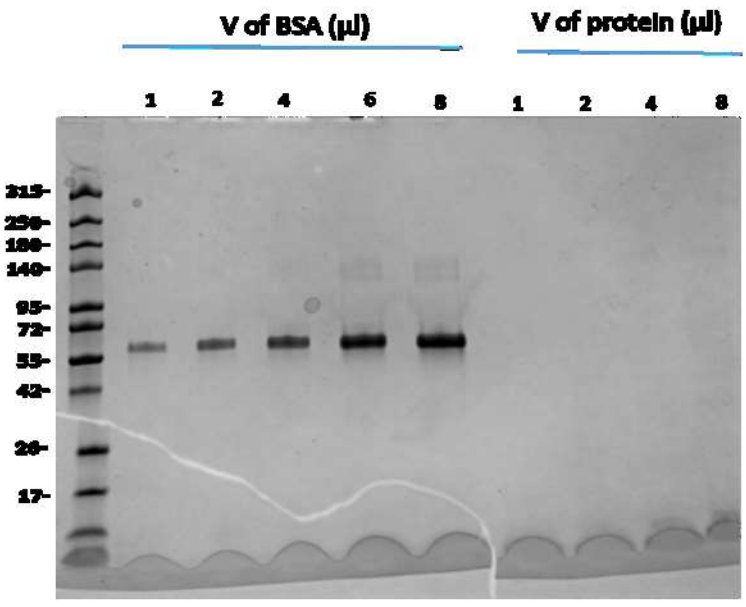
SDS-PAGE, 4-15% acrylamide gel
WB anti-MMP-9 (1:1000), anti-rat MMP (1:5000)

C



SDS-PAGE, 4-15% acrylamide gel
Bio-Rad Mini Protean Stain-Free gel

D



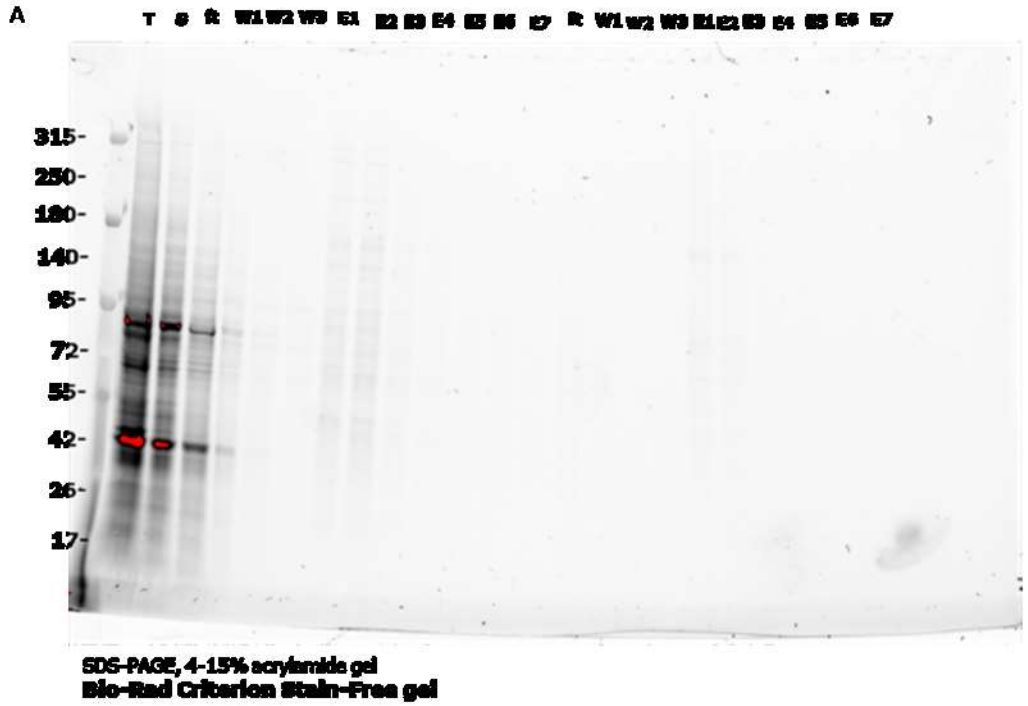
SDS-PAGE, 4-15% acrylamide gel
Bio-Rad Mini Protean Stain-Free gel
Coomassie Stain

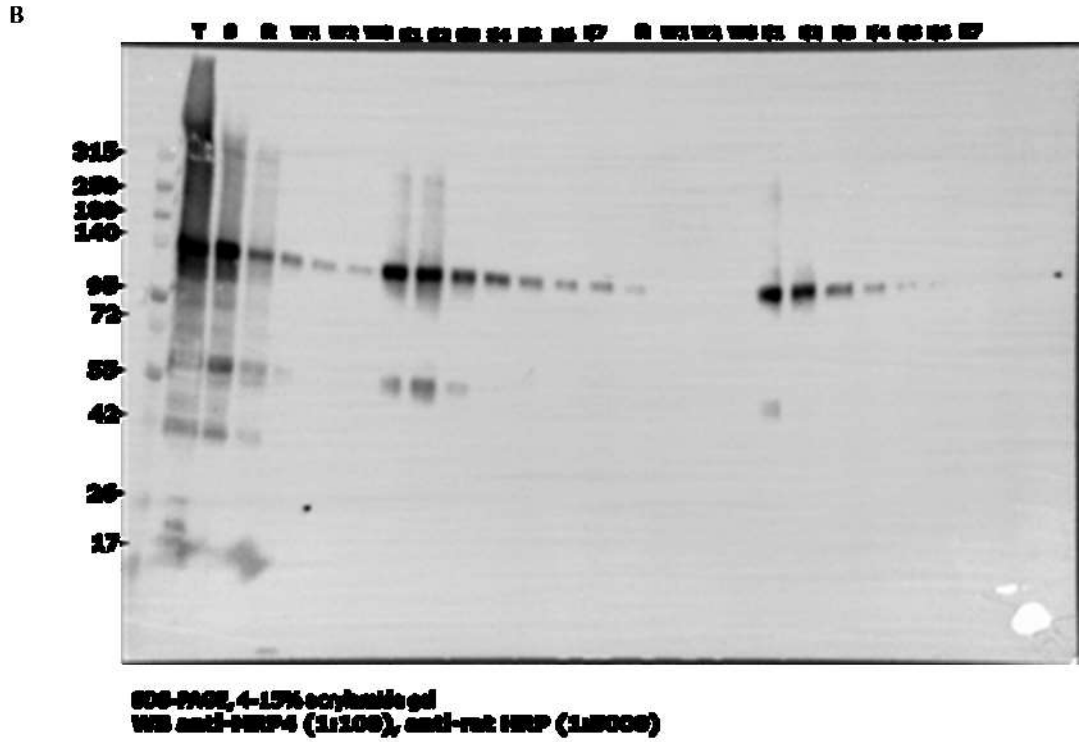
Figure 4.7: A. SDS-PAGE and Western blots of MRP4 small scale purification after solubilisation with DDG-CHS. *A&B) Proteins solubilised with DDG-CHS were subjected to two round of affinity purification. T stands for total protein and S stands for solubilised protein. W means Wash and E means elution. Ft1 signifies flow-through after overnight incubation of the protein with the resin and ft2 after 1 hour incubation. The second incubation and purification was performed after cycles of concentration and dilution to remove the imidazole of the protein. The samples were run on a stain free gel (A) with gradient of acrylamide 4-15% and transferred to a western blot (B) and probed with α-MRP4 and α-rat-HRP. Pooled and concentrated elution fractions, after the second round of purification, were run on a gel alongside BSA standards, and imaged stain free (C) or following Coomassie staining (D). Gels were imaged using a BioRad chemidoc and for the Western blot the Chemi-Hi resolution was used.*

4.5 Scaling up purification of MRP4 solubilised with DDG-CHS

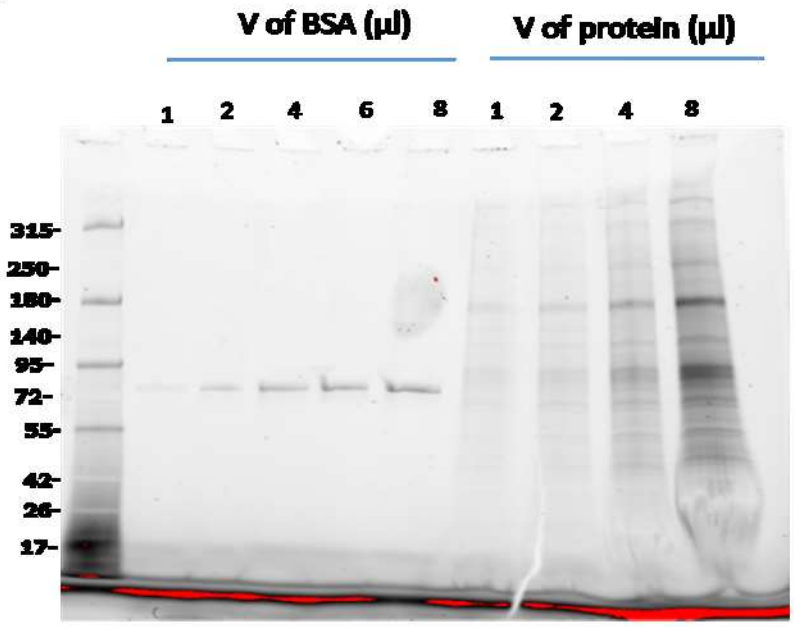
It was hypothesised that the lack of visible bands for purified protein on SDS-PAGE was due to low protein concentration because of the small initial volume of the experiment. Therefore, scaling up was carried out. A ten times greater starting volume of membranes was used, and all subsequent volumes in the procedure were scaled up accordingly. Otherwise the same procedure was used as at low scale. SDS-PAGE and Western blot analysis of each step are shown in Figure 4.8. Clear bands for MRP4 are visible in the Western blot, and very faint bands can be seen in some of the elution lanes on the SDS-PAGE. The elution fractions were pooled and concentrated and run alongside standards of BSA in order to estimate the amount of MRP4 (Figure 4.8 C&D). The dominant band corresponding to MRP4 can be seen at approximately 140 kDa, however there are many contaminating bands present too. The Coomassie stained gel was analysed with Image Lab and the MRP4 protein concentration

was found to be 200 μg/ml. MRP4 purity was calculating with densitometry using ImageJ software and was found to be approximately 25% pure.



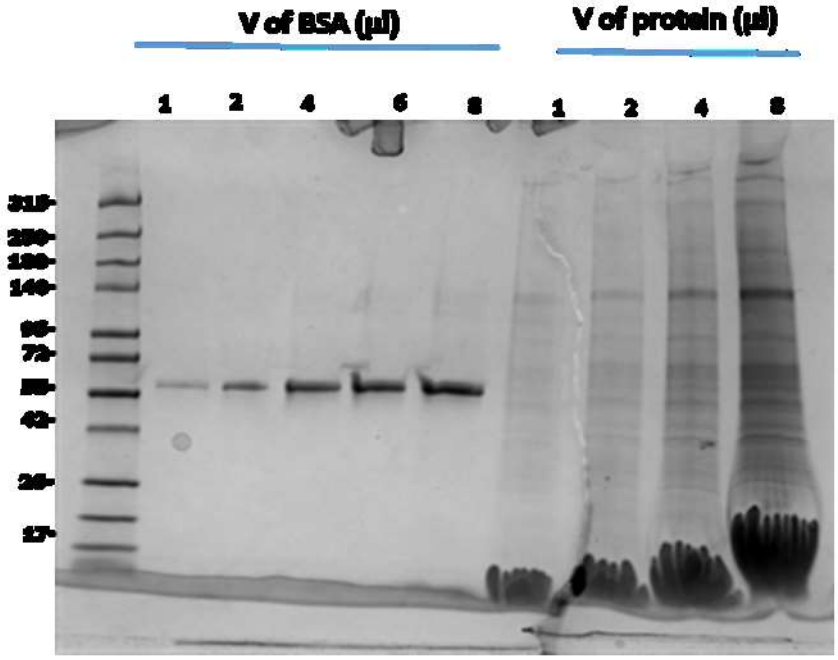


C



SDS-PAGE, 4-15% acrylamide gel
Bio-Rad Mini Protean Stain-Free gel

D



SDS-PAGE, 4-15% acrylamide gel
Bio-Rad Mini Protean Stain-Free gel
Coomassie Stain

Figure 4.8: Large scale purification of MRP4 using DDG-CHS. *A) Stain-free SDS-PAGE and B) Western blot of MRP4 purification after solubilisation with DDG-CHS. T stands for total protein and S stands for solubilised protein. W means Wash and E means elution. Ft1 signifies flow-through after overnight incubation of the protein with the resin and ft2 after 1 hour incubation. The second incubation and purification was performed after cycles of concentration and dilution to remove the imidazole of the protein. C&D) Pooled and concentrated elution fractions were run on a gel alongside BSA standards, and imaged stain free (C) or following Coomassie staining (D).*

4.6 Cryostability assay

To test if MRP4 remains stable at -80°C after solubilisation and purification with DDG-CHS, cycles of flash freeze and thawing, with or without addition of glycerol were tested. The samples were run on native PAGE and transferred to a Western blot (Figure 4.9). It can be seen that MRP4 forms a single clear band and this does not change under any of the conditions. The protein remains stable at -80°C with or without addition of glycerol after all 3 cycles of flash freeze and thawing.

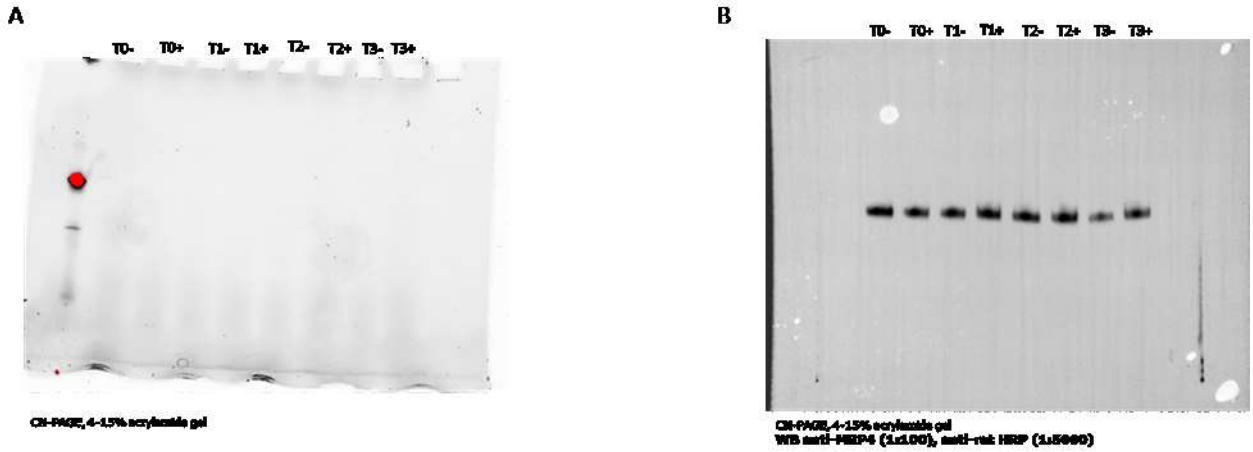


Figure 4.9: Cryostability assay for MRP4 solubilised and purified with DDG-CHS. *Purified MRP4 was subjected to cycles of flash freeze and thawing with or without glycerol (10%) addition. The samples were run on native PAGE. T indicates a cycle of thaw and freeze. + and – represent addition or no addition of glycerol respectively*. A) Stain free SDS-PAGE and B) Western blot.

4.7 Purification of MRP4 solubilised with LMNG-CHS

As LMNG-CHS had also shown promise at low scale, and the structure for the related protein MRP1/ABCC1 (expressed in HEK293S GnTI$^-$ cells) was achieved using LMNG (97), a large scale trial with this detergent was also carried out. In this study, Sf9 membranes were solubilised with LMNG-CHS at 4°C for 2 hours, and purified using either Ni^{2+}-sepharose or Co^{2+}-sepharose, in a single round of affinity purification. Western blots of the purification procedures are shown in Figure 4.10. With the cobalt resin very little MRP4 is seen in the flow-through, some MRP4 is clear in elution fractions 1 and 2, and some remains stuck to the resin even after multiple additions of 200mM imidazole. For the nickel resin a little more MRP4 can be seen in the flow-through, bands are seen in the first three elution fractions and none remained bound to the resin. The elution fractions were pooled and concentrated and run on SDS-PAGE alongside BSA standards (Figure 4.10 B&C). In both cases many contaminating bands are present. The gels were analysed with ImageJ and the protein concentration was found to be 150 µg/ml when Cobalt resin was used and 25 µg/ml when Nickel resin was used. The cobalt purified protein was approximately 17% pure and the nickel purified protein approximately 13% pure.

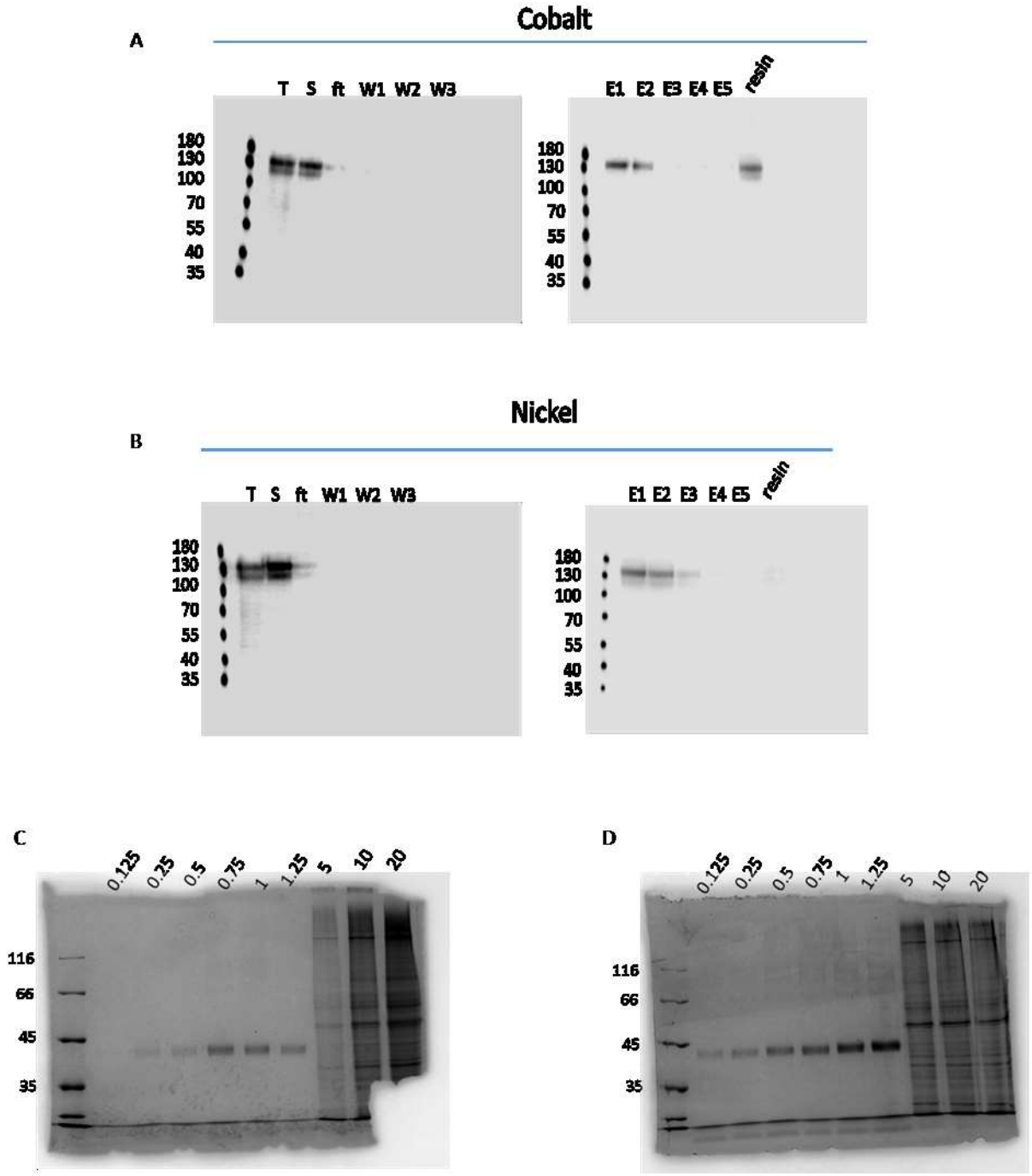


Figure 4.10: Large scale purification of MRP4 using LMNG-CHS. *A &B). Western blots of MRP4 purification after solubilisation with LMNG-CHS with cobalt resin (A) or nickel resin (B). T stands for total protein and S stands for solubilised protein. Ft means flow-through, W means Wash and E means elution. C&D; Pooled and concentrated elution fractions from the Cobalt resin (C) or Nickel resin (D) were run on a gel alongside BSA standards, and stained with Instant Blue.*

4.8 Summary

- Initial solubilisation screens showed that almost all detergents solubilised MRP4 efficiently, with CALXcod9-CHS and LMNG-CHS showing the lowest efficiency.
- The detergent solubilised MRP4 seems to remain well-folded, to not aggregate and to retain its conformation as shown from the migration at the native PAGE.
- In the small scale purifications, it is observed that there is a massive loss in the flow-through.
    - To optimize the problem of the loss of protein in the flow-through, Ni^{2+} and Co^{2+} resins were used as well as different incubation times of the protein with the resin
    - Generally, cobalt resin improved resin binding for all detergents tested.
- Scaled up purifications were carried out with both DDG-CHS and LMNG-CHS.
    - Solubilisation and purification with DDG-CHS gave a protein concentration of 200 µg/ml and a purity of 25% (rough estimation)
    - Solubilisation and purification with LMNG-CHS gave a concentration of 150 µg/ml in the case of Cobalt resin and of 25 µg/ml in the case of Ni. The purity of MRP4 was found to be 17% and 13% respectively (rough estimation)

5. MRP4 expression in 293F cells

Although it had been possible to successfully solubilise and purify MRP4 from Sf9 insect cells using either SMA polymer or detergents, a low yield and low purity was generally achieved. It was hypothesised that the expression level of MRP4 in the insect cells was a limiting factor in this, and despite efforts to increase the expression level, it wasn't possible. Therefore, it was decided to investigate if expression of MRP4 in a mammalian cell line could give better results in terms of yield and purity and avoid protein cleavage. The Freestyle 293F cell line was chosen. 293F line is a cell line of HEK293 cells adapted to grow in suspension in a specific medium, without addition of serum and antibiotics. The cells grow in 8% of CO_2, at 37°C, at 130 rpm agitation. The ability to grow them in shaker suspension cultures facilitates the ability to scale up much more easily than with monolayer cultures.

In the literature, expression in 293F cells has been successfully achieved by a number of methods including transfection. Chemical methods commonly use cationic polymers (such as jetPEI), calcium phosphate, cationic lipids (such as lipofectamine) or cationic amino acids. Physical methods use a variety of tools to deliver the gene, the most widely-used being electroporation, during which a short electrical pulse disturbs cell membranes and makes holes in the membrane through which nucleic acids can pass (123). However, when scaling up, transient transfection requires repeatedly needing large volumes of plasmid DNA. To avoid this the BacMam system was chosen.

The BacMam system involves using baculovirus to transduce mammalian cells to express proteins. Baculoviruses cannot infect mammalian cells and lead to viral reproduction and release, however they can transduce mammalian cells if there is a mammalian promotor at the vector which will cause protein expression (124).

5.1 Recombinant baculovirus generation and amplification

For these experiments, two plasmids were designed in collaboration with CALIXAR: one containing the mrp4 gene with a 6-His tag, to compare directly with the construct used for insect cell expression, and another one containing the mrp4 gene with a 10-His tag, a TEV cleavage site, GFP, and flag tag (Figure 5.1).

The backbone plasmid is pOET6 (Oxford Expression Technologies) which contains a CMV promoter for expression in mammalian cells. These plasmids are used to generate recombinant baculovirus using the flashBac process which is outlined in Figure 5.2, and is distinct from the Bac-to-bac system previously used to generate the baculovirus used in chapter 3 and 4. This was chosen because it essentially removes a step in the process (125). The FlashBac Ultra system was used because it is designed to be used for membrane proteins, and previous work in the lab suggested it was more likely to lead to good expression (Peer Depping, personal communication).

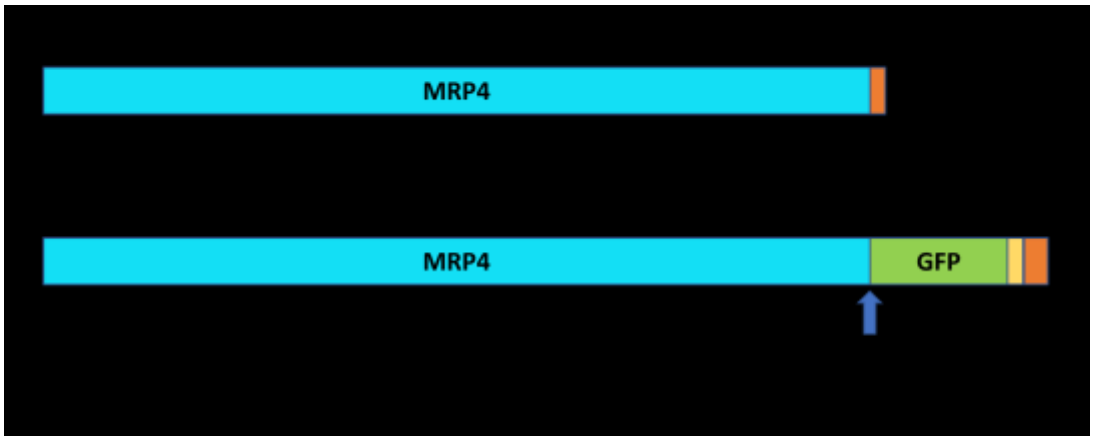


Figure 5.1: Schematic for MRP4 constructs. *A. MRP4-6-His B. MRP4-GFP-flag-10-His. Both MRP4 constructs were generated in a pOET6 vector, which has a CMV promoter for expression in the mammalian cells. They also have the baculovirus recombination region. These plasmids were used with the FlashBACULTRA and the Escort transfection reagent to generate the recombinant baculovirus within Sf9 cells.*

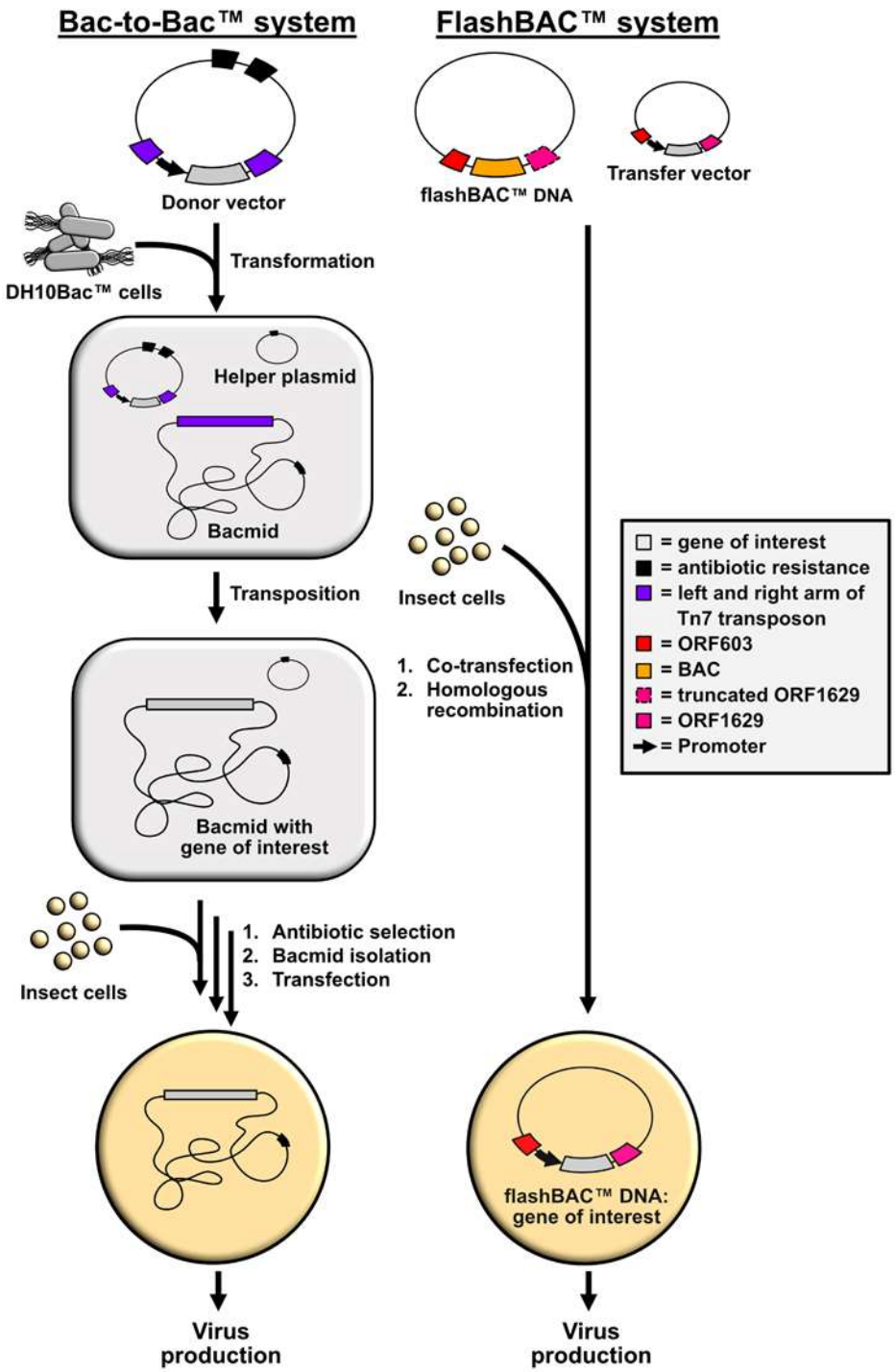


Figure 5.2: Schematic comparing Bac-to-bac and FlashBac systems for recombinant baculovirus generation (1).

The lab work schematic for generating and amplifying the recombinant baculovirus is shown in Figure 5.3. Briefly, $1x10^6$ cells per well were seeded in each well of a 6-well plate and mixed with the pOET plasmid, Escort transfection reagent, and flashBACULTRA for 5 days. The P1 generation of virus was harvested by centrifugation and stored at 4°C. In order to amplify the virus to generate the P2 generation, which can be used for expression or further amplification, the P1 virus was used to infect Sf9 cells at $2x10^6$ cells/ml for a week. 200ml of P2 virus was harvested for each construct and stored at 4°C.

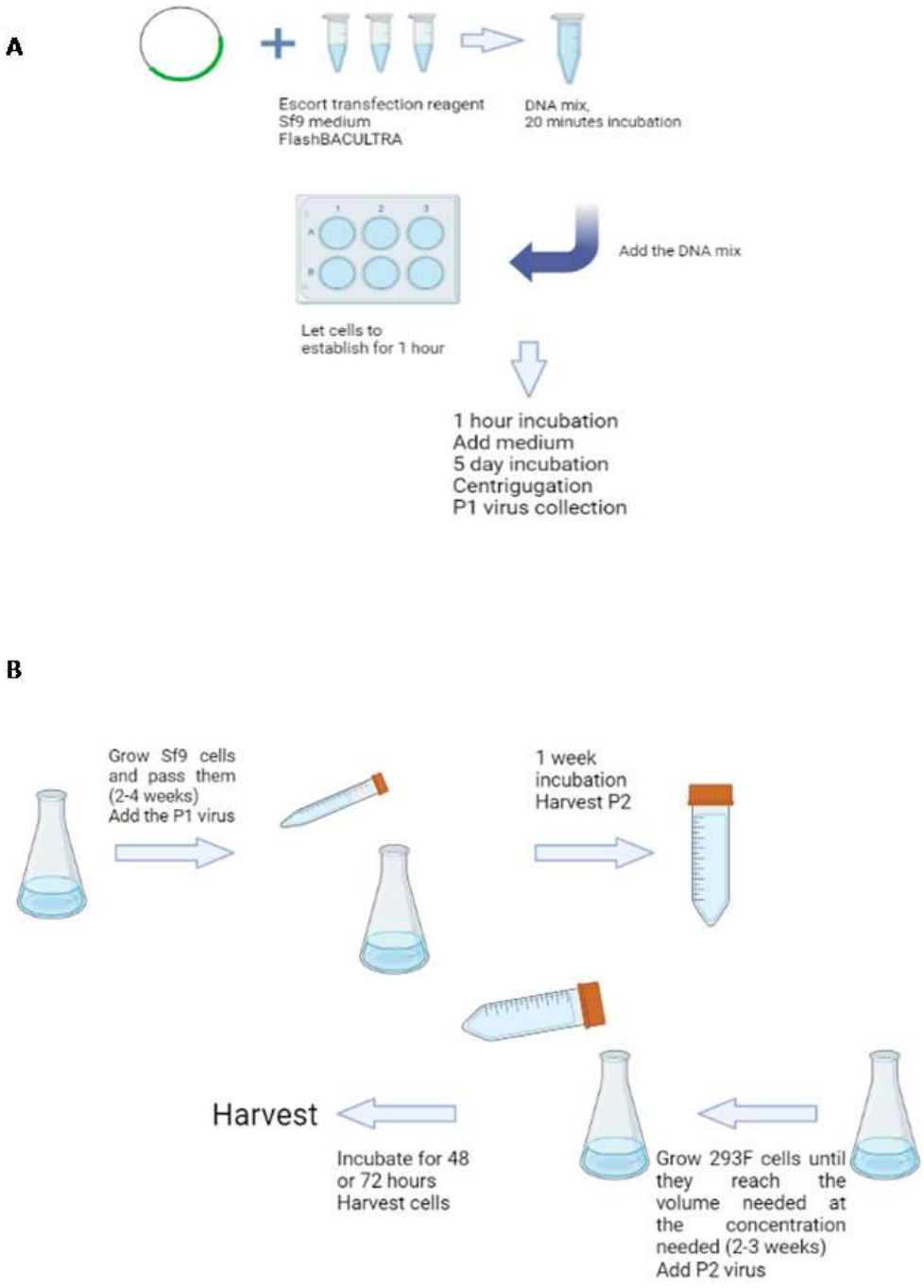
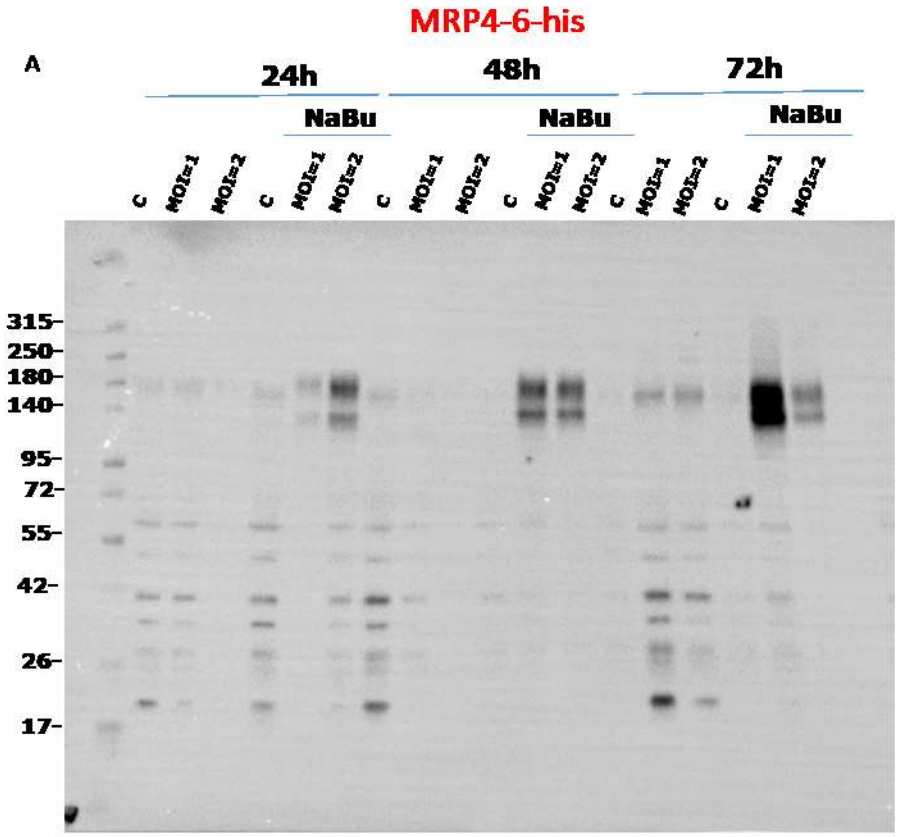


Figure 5.3: Schematic for P1 virus generation, P2 virus generation (amplification) and protein expression. *A; For P1 virus generation Sf9 cells were seeded at a concentration of 1x10^6 cells/ml. They were transfected with the DNA mix (plasmid, Escort transfection reagent, flashBACULTRA). After 5 days of incubation the P1 virus was harvested from the media. B; Shaker culture of Sf9 cells were infected with the P1 virus in order to amplify it to P2. This requires 1 week of incubation. The P2 virus could be used for expression in 293F cells, and incubated for 48 or 72 hours.*

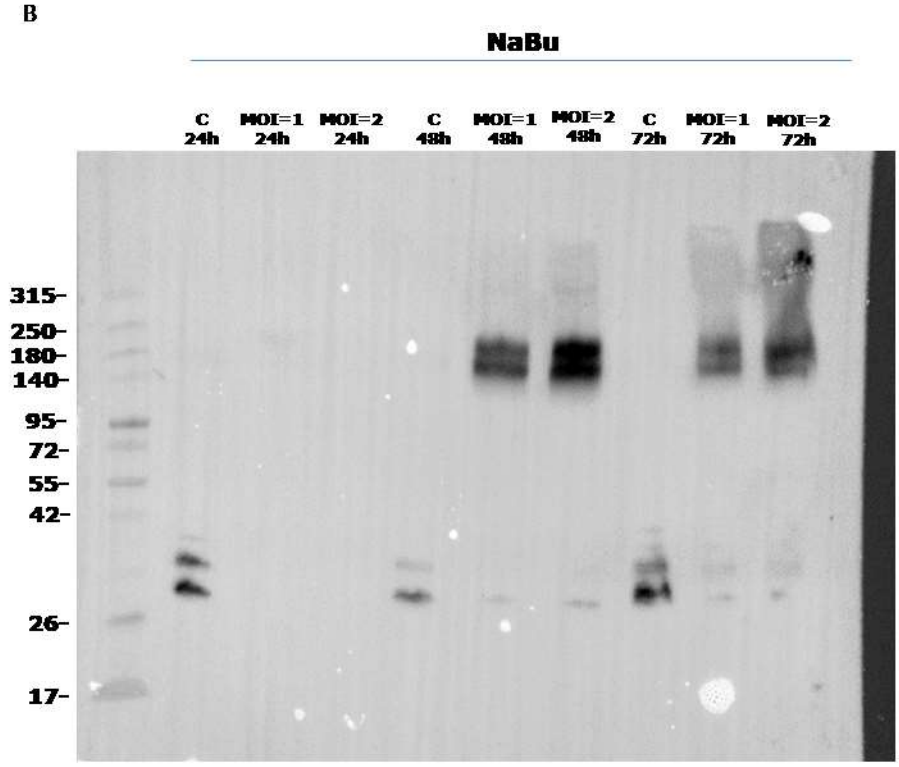
5.2 Expression tests

Small scale expression tests in the 293F cells were carried out. The cell concentration used was $2x10^6$ cells/ml, and viral volumes of 300 µl and 600 µl were tested on them. If sodium butyrate would facilitate the protein expression was also tested by addition of it or no addition. Sodium butyrate is supposed to induce or repress some genes and hence protein expression in some cells lines (116). Three incubation times, 24, 48 and 72 hours were tested, as well as two different temperatures, 30°C or 37°C. Figure 5.4 shows the different conditions tested. At each time point the cell culture was collected and harvested by centrifugation and whole cell lysates made. Samples were analysed by Western blotting as shown in Figure 5.4.

It was found that the best conditions for MRP4-6-His were 72 hours of incubation of the cells with the virus of 300 µl and with addition of sodium butyrate, while for MRP4-10-His-GFP-flag, the best expression was at 48 hours of incubation of the cells with the virus of 600 µl, with addition of sodium butyrate. This can be seen from the Western blot results, where the bands corresponding to the conditions, are thicker, with a strong signal suggesting that these are the optimal conditions.



SDS-PAGE, 4-15% acrylamide gel
WB anti-MRP4 (1:100), anti-rat HRP (1:5000)



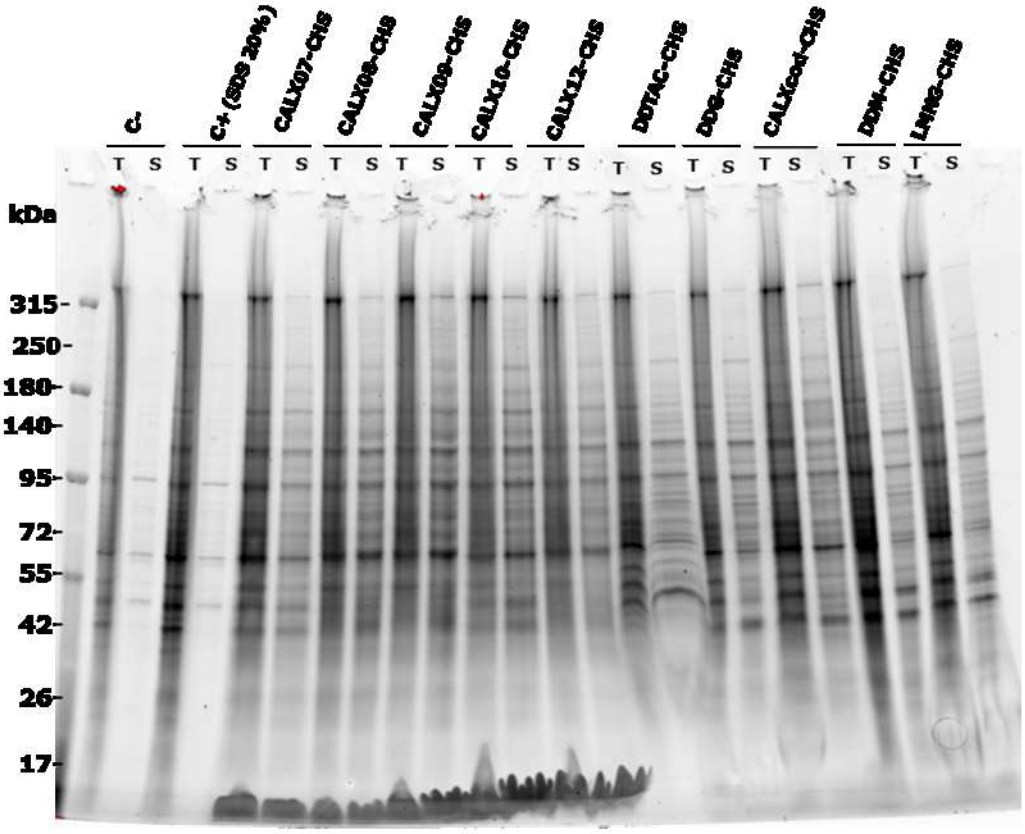
SDS-PAGE, 4-15% acrylamide gel
WB anti-MRP4 (1:100), anti-rat HRP (1:5000)

Figure 5.4: Western blots of MRP4 expression screening. *A. MRP4-6-His expression tests. When sodium butyrate is tested, it is indicated B. MRP4-10-His-GFP-flag expression test. In all cases for this experiment sodium butyrate was used. Samples of each well were collected in the different times (24, 48 and 72 hours). c indicates the control where no virus was used. The samples were then centrifuged at 4000 xg for 5 minutes at 4°C. The supernatant was removed and RIPA buffer and benzonase were added and the pellets were resuspended. Samples were loaded on SDS-PAGE and Western blot followed with α-MRP4 M4I-10 and α-rat-HRP antibodies. Chemiluminescence was imaged using a Chemi-Hi resolution of Bio-rad.*

5.3 Expression of MRP4-6-His and solubilisation test

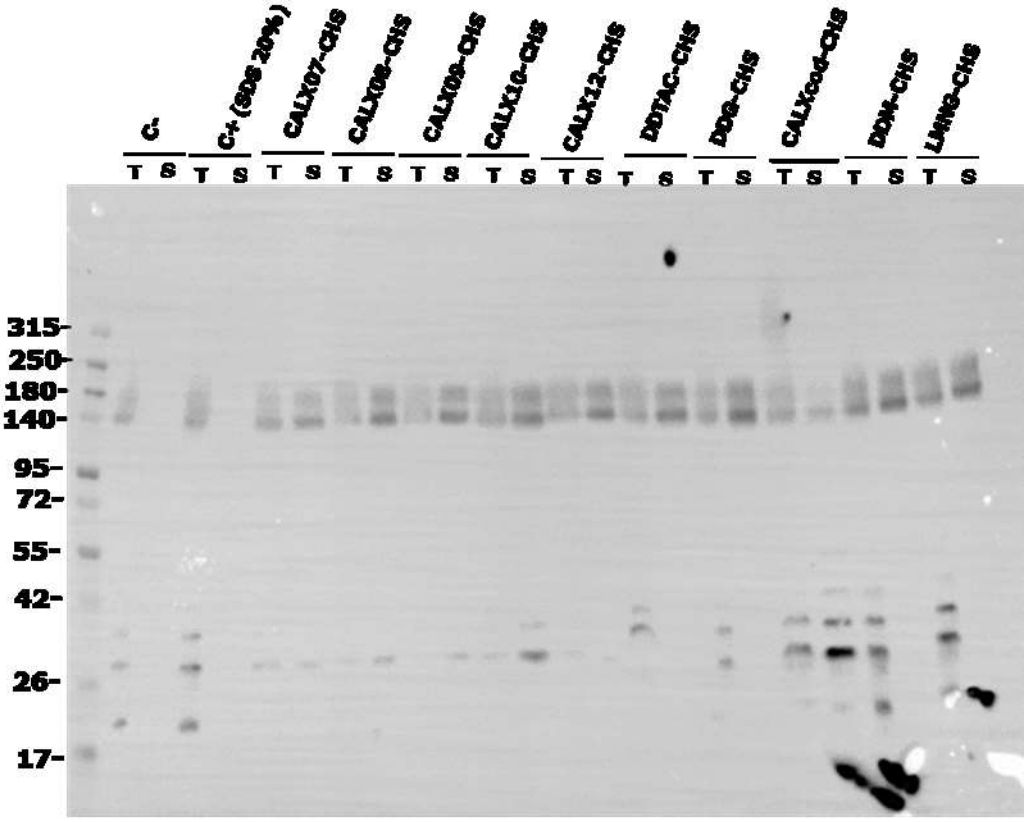
As the MRP4-10-His-GFP-flag construct was not available yet, and having established that MRP4-6-His could be successfully expressed in 293F cells, a preliminary test of solubilisation with a range of detergents was undertaken for the MRP4-6-His construct. MRP4-6-His was expressed in 293F cells of total volume 500 ml and concentration of 2 x 10^6 cells/ml when the baculovirus was introduced. They were incubated with baculovirus 100ml per L, at 37^{o}C, with addition of sodium butyrate, for 72 hours. The cells were harvested and membranes were prepared. The total protein was tested and samples of membranes with total protein 10 mg/ml were used. Solubilisation with different detergents was performed for 2 hours at 4^{o}C. Ultracentrifugation and sample preparation followed and Western blot was performed (figure 5.5). After densitometry analysis, it seems that CALX08-CHS and DDG-CHS showed the best solubilisation efficiency with 100% and 99.1% respectively. CALX09-CHS and DDTAC-CHS followed with solubilisation efficiency of 76% and 65.8% respectively. These are very preliminary results and only from one repeat, so more repeats will be needed to conclude about the solubilisation efficiency in the case of MRP4 expressed in 239F cells.

A



SDS-PAGE, 4-15% acrylamide gel
Bio-Rad Criterion Stain-Free gel

B



SDS-PAGE, 4-15% acrylamide gel
WB anti-MRP4 (1:100), anti-rat HRP (1:5000)

Figure 5.5 MRP4-6-his solubilisation test. *A. SDS-PAGE and B. Western blot of the total protein and solubilised protein samples. Solubilisation with the different detergents occurred in 4°C for 2 hours. Ultracentrifugation at 150000 xg, at 4°C, for 15 minutes was performed and the samples were loaded in a TGX Criterion stain free gel. After migration, western blot was performed with MRP4 M4I-10 and α-rat-HRP antibodies. Chemiluminescence was used with ECL solution in the Bio-rad imager at chemi-Hi resolution. C- and C+ indicate negative and positive control respectively. T means total protein and S solubilised protein.*

5.3 Summary

- From virus generation to protein expression is a long procedure
    - Virus generation was first carried out in Sf9 cells (2-4 weeks to grow the cells, then 5 days of incubation of the DNA mix with the cells), then P2 virus amplification was carried out in Sf9 cells (7 weeks of incubation)
    - Small scale expression trials in 293F cells were carried out (48-72 hours, plus analysis time)
- Best expression conditions of the ones tested for MRP4-6-Histidine were 72 hours of incubation of the cells with virus volume 100ml per L in the presence of sodium butyrate.
- Optimal expression conditions for MRP4-10-His-GFP-flag was after just 48 hours of incubation between virus and cells, with viral volume 200ml per L, in the presence of sodium butyrate.
- MRP4-6-His expressed in 293F cells could be effectively solubilised by a wide range of different detergents.

6. Discussion

Membrane proteins, including MRP4, are important drug targets, so study of their structure and function is essential for drug design. This begins with the suitable expression, solubilisation and purification of the protein, and that has been the main focus of this thesis.

6.1 Expression of MRP4

In this work, MRP4 was expressed by recombinant baculovirus infection of Sf9 insect cells. Efforts to optimise the expression conditions by varying the expression time, showed that 48 hours infection gave the optimal expression level, with breakdown of MRP4 occurring at 72 hours. This agrees well with previous reports (117). Other factors such as cell density or MOI of infection have previously been shown to offer no further improvements in expression (117). Despite this optimisation, the expression level of MRP4 in Sf9 cells was not as high as is often seen for other proteins using this system. This relatively low expression level added to the challenge with many of the downstream steps. One possible way to improve the expression level in the future might be to try *Trichoplusia ni* cells instead of Sf9. These are another type of insect cell which can be infected by the same recombinant baculovirus as Sf9, and they have previously been shown to give high levels of expression for some related proteins such as P-glycoprotein (126).

An alternative option is to try a different expression system. It was previously shown that MRP4 could be successfully expressed in *Pichia pastoris* yeast, however, the yield was even lower than obtained with Sf9 cells (117). Expression within a mammalian cell line is another option. This would arguably provide a more native-like environment for MRP4 which is a human protein. Previously it was shown that His-tagged MRP4 could not be expressed in HEK cells by transfection with a plasmid encoding MRP4 (although non-tagged MRP4 could) (117). Instead, in this thesis the BacMam system was

investigated. This takes advantage of a recombinant baculovirus to transduce mammalian cells and lead to protein expression. For this to work, the recombinant baculovirus needs to include a mammalian promoter, thus new recombinant baculoviruses had to be developed. Two plasmids with a CMV promoter for mammalian expression and the protein of interest were used. One of them had a MRP4 tagged with 6-His, while the other had MRP4 with GFP, 10-His and flag tags. Preliminary results show that recombinant baculovirus formed from each construct could be successfully used to express MRP4 in 293F cells. The optimal expression was found with addition of sodium butyrate, viral volume 100ml per culture L, 72 hours of incubation of the virus and the cells for MRP4-6-His and viral volume 200ml per culture L, 48 hours of incubation for MRP4-GFP-10-His-flag. In the case of a GPCR-ligand complex, the BacMam system was also used, and the optimal expression conditions were at 72 hours of incubation of the virus and the cells, with addition of sodium butyrate (127). P-gp and ABCG2 were also expressed using the BacMam system and optimal expression was found to be to up to 72 hours of incubation of the virus and the cells (128). In this study, expression tests were performed both at 37^{o}C and 30^{o}C temperatures, without any difference at the expression levels. However, lowering the temperature at 30^{o}C once the virus was introduced, gave a better expression in the case of chicken acid-sensing ion channel 1a (cASIC1) and glutamate-gated chloride channel (GluCl) (129). Due to time constraints it was not possible to see how the expression level compared to that achieved with Sf9 cells, and this is something that will be of great interest in the future. For example some proteins have better expression levels in mammalian cells, such as HCN2 (130). It will also be very important to determine if the different lipid environment from HEK cells has an effect on function compared to Sf9 cells. MRP4 expressed in Sf9 cells has been shown previously to be functional, using a fluorescent vesicular uptake assay (117), and application of this assay to the HEK membranes should be undertaken.

6.2 Solubilisation and purification of MRP4 with SMA polymer

It has been shown previously that MRP4 could be solubilised from Sf9 cells and purified using SMA2000 polymer (114, 131). One major improvement established in this thesis was starting with membranes of higher concentration. Typically, a concentration of 30-40 mg/ml (wet pellet weight) is used (115, 132), but by increasing this to 180 mg/ml it was possible to elute much more concentrated purified protein, which is useful for downstream steps. This did not impact on the purity of protein obtained.

Other efforts to optimise solubilisation efficiency were undertaken, including the presence of arginine, higher levels of NaCl or increased concentrations of SMA. None of these however provided any significant improvement. Binding of SMALP encapsulated proteins to Ni-NTA (and other affinity resins) has previously been reported to be weaker than for detergent solubilised proteins (101). This is why no imidazole is present in the binding buffer and binding takes place overnight. Possible reasons for this might include interactions between the His-tag and the polymer or steric hindrance from the polymer. However, generally it is thought that column spoiling by excess free SMA is the major issue (115). The addition of arginine, higher NaCl or increased amounts of Ni-NTA resin were tested to try to improve binding to the resin, and decrease the amount of MRP4 in the flow through. Again, none of them made a significant difference. This contrasts with previous reports that have shown the presence of arginine can make a significant difference for running SMALP (or DIBMALP) encapsulated proteins on size exclusion columns, which are also prone to column spoiling by SMA (115, 119). Other methods that could be trialled in the future to improve the binding to Ni^{2+}-NTA would be to remove the excess free SMA prior to affinity chromatography. This was successfully carried out by others in the literature using centrifugal concentrators (133), dialysis (115) or size exclusion (134).

Another approach that was investigated was to use novel modified variants of SMA polymer. A series of partially esterified SMA polymers were tested. These were chosen because they had been shown to offer better solubilisation than SMA2000 for plant thylakoid membranes (135), but had not been

tested on other cell systems/membranes. All three polymers successfully solubilised MRP4 well, but offered no improvement over SMA2000. Why this was different to thylakoid membranes may be due to the different membrane properties, with thylakoid membranes being densely packed and high in galactolipids (136).

6.3 Solubilisation and purification of MRP4 with detergents

MRP4 was also solubilised with a range of detergents provided from CALIXAR as well as DDM and LMNG. In a previous study (131), one of the same detergents was used, CALX-07, but in this thesis it was supplemented with CHS to help maintain a more native environment, as CHS is similar to cholesterol (97). LMNG-CHS was chosen because it was successfully used for the solubilisation, purification and structure determination of the related protein MRP1/ABCC1 (97).

All detergents tested were able to successfully solubilise MRP4 from Sf9 cells with relatively high efficiency. For CALX-07 the addition of CHS gave a solubilisation efficiency a little higher than that obtained previously for CALX-07 alone (131). The addition of CHS to DDM also seemed to have improved the solubilisation efficiency and stability compared to previous studies with DDM alone (131). Native PAGE analysis showed all the detergents seemed to maintain the protein's folding and conformation and the protein did not aggregate. Previous studies had investigated the stability of solubilised protein using a thermal shift Western blot method (131), which was able to detect significant stability differences between different detergents, and it would be interesting to try something similar with the detergents tested in this thesis.

Nickel affinity purification of the detergent solubilised proteins highlighted problems with the protein's binding to the affinity resin, which was unexpected. As discussed above, it is well established that SMALP encapsulated proteins bind poorly to affinity resins, but usually detergent solubilised proteins do not suffer from this. To address this the use of cobalt affinity purification was tested instead. Cobalt has worked better in some cases in the literature giving a purer protein of higher

concentration after purification (137). Improvements in resin binding were observed, especially if the protein was left to bind overnight rather than for 1 hour.

Scaled up purification using DDG-CHS was able to produce elution fractions that were visible on SDS-PAGE. The yield of purified MRP4 achieved with DDG-CHS was comparable to previous reports using CALX-07 (131), and higher than that achieved with SMA. However, the degree of purity was low at only approximately 25% pure. Similarly, LMNG-CHS purification could be scaled up, but it gave a lower yield and lower degree of purity than DDG-CHS. DDG-CHS is a particularly interesting detergent because it has recently been shown to form nanodisc-like structures rather than traditional micelles. With negative-stain transmission electron microscopy (TEM), it was found that DDG forms a "disc" as thick as nanodiscs. This makes it maintain a quite native environment for the protein, making its use promising for various membrane proteins as this thickness is determined of the lipids that are encapsulated rather than the amphiphile itself. Plus, it was shown to maintain the function of the protein tested (OmpLA). Contrary to bicelle-forming amphiphiles, DDG was capable of extracting membrane proteins directly from chemically defined, artificial proteoliposomes and, more importantly, also from highly heterogeneous, native cellular membranes. Moreover, this amphiphile was found to be remarkably gentle, as reflected by the good preservation of the bilayer architecture. DDG seems to combine the advantages of an efficient membrane solubilizing and protein-extracting agent, and native-like lipid-bilayer nanodiscs to provide a valuable new tool for studies of membrane proteins (138).

One downside of the DDG-CHS purification was that there seemed to be a cleavage of MRP4 resulting in bands showing on the Western blot detected with anti-MRP4 antibody at approximately 60kDa. ABC transporters are prone to cleavage between the first NBD and second MSD, but it is not known if this is problematic for function. In fact, the related protein MRP1 can actually be expressed in two halves, with two separate polypeptide chains, that folded together in the cell can form a functioning protein (139). However, for future structural studies it would be nice if this cleavage product was

avoided. Perhaps the expression within 293F cells would give a better yield, protein which is not cleaved and more pure. A preliminary solubilisation test with CALIXAR detergents and DDM-CHS and LMNG-CHS, suggests that the detergents solubilise MRP4 from 293F cells well, and that the cleavage does not appear after expression in 293F cells. This is very promising for the future.

6.4 Improvements needed to get to structure

In order to get to the point where structural studies are possible further improvements are needed. Ideally, a higher expression, better binding to the affinity resin, and greater purity are needed. Increasing the length of the Histidine tag could assist the binding of the protein in the resin and eliminate the loss in the flow-through. One of the new BacMam constructs has a 10-His tag so it will be interesting to see if this is beneficial with regard to resin binding. It may enable washes with higher concentrations of imidazole to improve purity.

The BacMam system shows some promise, but a lot of work is still needed to investigate solubilisation and purification with SMA as well as detergents. Moreover, one of the plasmids gives the opportunity for purification with the GFP tag or the flag tag, thus, offering the chance of double purification, one with GFP and one with 10-His tag which could improve the purity.

Obtaining enough protein to proceed to size exclusion chromatography is necessary, as this can help clean it from small impurities, as well as report on the aggregation, stability and homogeneity of the sample. In the end, when the protein yield is high, the solubilisation efficiency enough, the loss in the flow-through minimal and the purity optimal, the protein could be ready for cryo-EM experiments, although negative-stain EM is advised before going to cryo-EM grid screening as it is easier to have images, select particles, create classes and does not need grid screening, making it a more straightforward approach to have a first look on the sample homogeneity before undergoing screenings for cryo-EM which can be time consuming.

7. Acknowledgements

I would like to thank, first of all, my supervisors, Dr Alice Rothnie and Prof Roslyn Bill, for their continuous support and assistance, for encouraging me and provide me with ideas and for trusting me with this project.

I would also like to thank Dr Alan Goddard, who also trusted me with this project.

Many thanks to the MemTrainees and the members of the lab 331 for their willingness to help, the friendly environment and the exchange of ideas and knowledge.

Moreover, I would like to thank CALIXAR team, for our collaboration and for hosting my experiments for some months in their lab. All the members have been supportive and friendly, willing to teach me.

I should not miss thanking for the funding, the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 847419 and Aston University for hosting the MemTrain program.

Last but not least, I would like to thank you for reading and examining my thesis.

8. References

1. Vaitsopoulou A, Depping P, Bill RM, Goddard AD, Rothnie AJ. Membrane Protein Production in Insect Cells. Methods in molecular biology (Clifton, NJ). 2022;2507:223-40.
2. Hawkins OP, Jahromi CPT, Gulamhussein AA, Nestorow S, Bahra T, Shelton C, et al. Membrane protein extraction and purification using partially-esterified SMA polymers. Biochimica et Biophysica Acta (BBA) - Biomembranes. 2021:183758.
3. Rothnie AJ. Detergent-Free Membrane Protein Purifi cation. Methods and Protocols, Methods in Molecular Biology. 2016.
4. Hedin LE, Illergård K, Elofsson A. An Introduction to Membrane Proteins. Journal of Proteome Research. 2011;10(8):3324-31.
5. Autzen HE, Julius D, Cheng Y. Membrane mimetic systems in CryoEM: keeping membrane proteins in their native environment. Current Opinion in Structural Biology. 2019;58:259-68.
6. Stroud Z, Hall SCL, Dafforn TR. Purification of membrane proteins free from conventional detergents: SMA, new polymers, new opportunities and new insights. Methods. 2018;147:106-17.
7. Woods A, Jackson S. Brain Tissue Lipidomics: Direct Probing Using Matrix-assisted Laser Desorption/Ionization Mass Spectrometry. 2006;08(02):E391.
8. Greene NP, Kaplan E, Crow A, Koronakis V. Antibiotic Resistance Mediated by the MacB ABC Transporter Family: A Structural and Functional Perspective. Frontiers in Microbiology. 2018;9.
9. Seddon AM, Curnow P, Booth PJ. Membrane proteins, lipids and detergents: not just a soap opera. Biochimica et Biophysica Acta (BBA) - Biomembranes. 2004;1666(1-2):105-17.
10. Escribá PV, Ozaita A, Ribas C, Miralles A, Fodor E, Farkas T, et al. Role of lipid polymorphism in G protein-membrane interactions: Nonlamellar-prone phospholipids and peripheral protein binding to membranes. Proceedings of the National Academy of Sciences. 1997;94(21):11375-80.
11. Nicolson GL. The Fluid—Mosaic Model of Membrane Structure: Still relevant to understanding the structure, function and dynamics of biological membranes after more than 40years. Biochimica et Biophysica Acta (BBA) - Biomembranes. 2014;1838(6):1451-66.
12. Kesidis A, Depping P, Lodé A, Vaitsopoulou A, Bill RM, Goddard AD, et al. Expression of eukaryotic membrane proteins in eukaryotic and prokaryotic hosts. Methods. 2020.
13. Tan S, Tan HT, Chung MCM. Membrane proteins and membrane proteomics. PROTEOMICS. 2008;8(19):3924-32.
14. Watson H. Biological membranes. Essays Biochem. 2015;59:43-69.
15. Nicolson GL. Transmembrane control of the receptors on normal and tumor cells: I. Cytoplasmic influence over cell surface components. Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes. 1976;457(1):57-108.
16. Antonny B. Membrane deformation by protein coats. Current opinion in cell biology. 2006;18(4):386-94.
17. Andersen OS, Koeppe RE. Bilayer thickness and membrane protein function: an energetic perspective. Annu Rev Biophys Biomol Struct. 2007;36:107-30.
18. Neumann AK, Itano MS, Jacobson K. Understanding lipid rafts and other related membrane domains. F1000 biology reports. 2010;2.
19. Held W, Mariuzza RA. Cis–trans interactions of cell surface receptors: biological roles and structural basis. Cellular and Molecular Life Sciences. 2011;68(21):3469-78.
20. Casares D, Escribá PV, Rosselló CA. Membrane Lipid Composition: Effect on Membrane and Organelle Structure, Function and Compartmentalization and Therapeutic Avenues. International Journal of Molecular Sciences. 2019;20(9):2167.
21. Ibarguren M, López DJ, Encinar JA, González-Ros JM, Busquets X, Escribá PV. Partitioning of liquid-ordered/liquid-disordered membrane microdomains induced by the fluidifying effect of 2-hydroxylated fatty acid derivatives. Biochimica et Biophysica Acta (BBA)-Biomembranes. 2013;1828(11):2553-63.

22. Frame MC. Src in cancer: deregulation and consequences for cell behaviour. Biochimica et Biophysica Acta (BBA) - Reviews on Cancer. 2002;1602(2):114-30.
23. Elofsson A, Heijne Gv. Membrane protein structure: prediction versus reality. Annu Rev Biochem. 2007;76:125-40.
24. Strandberg E, Killian JA. Snorkeling of lysine side chains in transmembrane helices: how easy can it get? FEBS Letters. 2003;544(1):69-73.
25. Almén MS, Nordström KJ, Fredriksson R, Schiöth HB. Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin. BMC Biology. 2009;7(1):50.
26. Forrest LR. Structural Symmetry in Membrane Proteins. Annual Review of Biophysics. 2015;44(1):311-37.
27. Levy ED, Pereira-Leal JB, Chothia C, Teichmann SA. 3D complex: a structural classification of protein complexes. PLoS computational biology. 2006;2(11):e155.
28. Marsden RL, Lee D, Maibaum M, Yeats C, Orengo CA. Comprehensive genome analysis of 203 genomes provides structural genomics with new insights into protein family space. Nucleic Acids Research. 2006;34(3):1066-80.
29. Abraham A-L, Pothier J, Rocha EP. Alternative to homo-oligomerisation: the creation of local symmetry in proteins by internal amplification. Journal of molecular biology. 2009;394(3):522-34.
30. Lynch M, Conery JS. The evolutionary fate and consequences of duplicate genes. science. 2000;290(5494):1151-5.
31. Kumazaki K, Chiba S, Takemoto M, Furukawa A, Nishiyama K-i, Sugano Y, et al. Structural basis of Sec-independent membrane protein insertion by YidC. Nature. 2014;509(7501):516-20.
32. Rollauer SE, Tarry MJ, Graham JE, Jääskeläinen M, Jäger F, Johnson S, et al. Structure of the TatC core of the twin-arginine protein transport system. Nature. 2012;492(7428):210-4.
33. Morgan JL, Strumillo J, Zimmer J. Crystallographic snapshot of cellulose synthesis and membrane translocation. Nature. 2013;493(7431):181-6.
34. Toyoshima C, Nakasako M, Nomura H, Ogawa H. Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 Å resolution. Nature. 2000;405(6787):647-55.
35. Hirai T, Heymann JA, Shi D, Sarker R, Maloney PC, Subramaniam S. Three-dimensional structure of a bacterial oxalate transporter. Nature structural biology. 2002;9(8):597-600.
36. He X, Szewczyk P, Karyakin A, Evin M, Hong W-X, Zhang Q, et al. Structure of a cation-bound multidrug and toxic compound extrusion transporter. Nature. 2010;467(7318):991-4.
37. Eicher T, Seeger MA, Anselmi C, Zhou W, Brandstätter L, Verrey F, et al. Coupling of remote alternating-access transport mechanisms for protons and substrates in the multidrug efflux pump AcrB. elife. 2014;3:e03145.
38. Lee AG. How lipids affect the activities of integral membrane proteins. Biochimica et Biophysica Acta (BBA) - Biomembranes. 2004;1666(1-2):62-87.
39. Marsh D, Horváth LI. Structure, dynamics and composition of the lipid-protein interface. Perspectives from spin-labelling. Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes. 1998;1376(3):267-96.
40. White SH, Ladokhin AS, Jayasinghe S, Hristova K. How membranes shape protein structure. Journal of Biological Chemistry. 2001;276(35):32395-8.
41. Aurora R, Rosee GD. Helix capping. Protein Science. 1998;7(1):21-38.
42. Lee A. Lipid phase transitions and phase diagrams I. Lipid phase transitions. Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes. 1977;472(2):237-81.
43. Webb RJ, East JM, Sharma RP, Lee AG. Hydrophobic mismatch and the incorporation of peptides into lipid bilayers: a possible mechanism for retention in the Golgi. Biochemistry. 1998;37(2):673-9.
44. de Planque MR, Goormaghtigh E, Greathouse DV, Koeppe RE, Kruijtzer JA, Liskamp RM, et al. Sensitivity of single membrane-spanning α-helical peptides to hydrophobic mismatch with a lipid

bilayer: effects on backbone structure, orientation, and extent of membrane incorporation. Biochemistry. 2001;40(16):5000-10.
45. Lee A. Lipid–protein interactions in biological membranes: a structural perspective. Biochimica et Biophysica Acta (BBA)-Biomembranes. 2003;1612(1):1-40.
46. Bali AP, Sahu ID, Craig AF, Clark EE, Burridge KM, Dolan MT, et al. Structural characterization of styrene-maleic acid copolymer-lipid nanoparticles (SMALPs) using EPR spectroscopy. Chemistry and Physics of Lipids. 2019;220:6-13.
47. Locher KP. Mechanistic diversity in ATP-binding cassette (ABC) transporters. Nature Structural & Molecular Biology. 2016;23(6):487-93.
48. Ter Beek J, Guskov A, Slotboom DJ. Structural diversity of ABC transporters. The Journal of General Physiology. 2014;143(4):419-35.
49. Eitinger T, Rodionov DA, Grote M, Schneider E. Canonical and ECF-type ATP-binding cassette importers in prokaryotes: diversity in modular organization and cellular functions. FEMS Microbiol Rev. 2011;35(1):3-67.
50. Beis K. Structural basis for the mechanism of ABC transporters. Biochemical Society Transactions. 2015;43(5):889-93.
51. Jones PM, O'Mara ML, George AM. ABC transporters: a riddle wrapped in a mystery inside an enigma. Trends in Biochemical Sciences. 2009;34(10):520-31.
52. Holland IB. ABC transporters, mechanisms and biology: an overview. Essays In Biochemistry. 2011;50:1-17.
53. Biemans-Oldehinkel E, Doeven MK, Poolman B. ABC transporter architecture and regulatory roles of accessory domains. 2006;580(4):1023-35.
54. Linton KJ. Structure and Function of ABC Transporters. Physiology. 2007.
55. Kim S, Bajaj T, Chabon C, Tablante E, Kulchinskaya T, Moon TS, et al. Meta-Analysis of the Expansion in the Field of Structural Biology of ABC Transporters. BioDesign Research. 2022;2022.
56. Thomas C, Aller SG, Beis K, Carpenter EP, Chang G, Chen L, et al. Structural and functional diversity calls for a new classification of ABC transporters. FEBS Letters. 2020;594(23):3767-75.
57. Vasiliou V, Vasiliou K, Nebert DW. Human ATP-binding cassette (ABC) transporter family. Human Genomics. 2008;3(3):281.
58. Borst P, Elferink RO. Mammalian ABC Transporters in Health and Disease. Annual Review of Biochemistry. 2002;71(1):537-92.
59. Dean M. ABC Transporters, Drug Resistance, and Cancer Stem Cells. Journal of Mammary Gland Biology and Neoplasia. 2009;14(1):3-9.
60. Sparreboom A, Danesi R, Ando Y, Chan J, Figg WD. Pharmacogenomics of ABC transporters and its role in cancer chemotherapy. 2003;6(2):71-84.
61. Tarling EJ, Vallim TQDA, Edwards PA. Role of ABC transporters in lipid transport and human disease. Trends in Endocrinology & Metabolism. 2013;24(7):342-50.
62. Fletcher JI, Williams RT, Henderson MJ, Norris MD, Haber M. ABC transporters as mediators of drug resistance and contributors to cancer cell biology. Drug Resistance Updates. 2016;26:1-9.
63. Fletcher JI, Haber M, Henderson MJ, Norris MD. ABC transporters in cancer: more than just drug efflux pumps. Nature Reviews Cancer. 2010;10(2):147-56.
64. Schinkel AH, Jonker JW. Mammalian drug efflux transporters of the ATP binding cassette (ABC) family: an overview. Advanced Drug Delivery Reviews. 2012;64:138-53.
65. Robey RW, Polgar O, Deeken J, To KW, Bates SE. ABCG2: determining its relevance in clinical drug resistance. 2007;26(1):39-57.
66. Abuznait AH, Kaddoumi A. Role of ABC Transporters in the Pathogenesis of Alzheimer's Disease. ACS Chemical Neuroscience. 2012;3(11):820-31.
67. Russel F, Koenderink J, Masereeuw R. Multidrug resistance protein 4 (MRP4/ABCC4): a versatile efflux transporter for drugs and signalling molecules. 2008;29(4):200-7.
68. Zhang G, Wang Z, Qian F, Zhao C, Sun C. Silencing of the ABCC4 gene by RNA interference reverses multidrug resistance in human gastric cancer. 2014.

69. Ci L, Kusuhara H, Adachi M, Schuetz JD, Takeuchi K, Sugiyama Y. Involvement of MRP4 (ABCC4) in the Luminal Efflux of Ceftizoxime and Cefazolin in the Kidney. 2007;71(6):1591-7.
70. De Waart DR, Van De Wetering K, Kunne C, Duijst S, Paulusma CC, Oude Elferink RPJ. Oral Availability of Cefadroxil Depends on ABCC3 and ABCC4. 2012;40(3):515-21.
71. Copsel S, Garcia C, Diez F, Vermeulem M, Baldi A, Bianciotti LG, et al. Multidrug Resistance Protein 4 (MRP4/ABCC4) Regulates cAMP Cellular Levels and Controls Human Leukemia Cell Proliferation and Differentiation. 2011;286(9):6979-88.
72. Furmanski BD, Hu S, Fujita KI, Li L, Gibson AA, Janke LJ, et al. Contribution of Abcc4-Mediated Gastric Transport to the Absorption and Efficacy of Dasatinib. 2013;19(16):4359-70.
73. Leggas M, Adachi M, Scheffer GL, Sun D, Wielinga P, Du G, et al. Mrp4 Confers Resistance to Topotecan and Protects the Brain from Chemotherapy. 2004;24(17):7612-21.
74. Ho LL, Kench JG, Handelsman DJ, Scheffer GL, Stricker PD, Grygiel JG, et al. Androgen regulation of multidrug resistance-associated protein 4 (MRP4/ABCC4) in prostate cancer. 2008;68(13):1421-9.
75. Hara Y, Sassi Y, Guibert C, Gambaryan N, Dorfmüller P, Eddahibi S, et al. Inhibition of MRP4 prevents and reverses pulmonary hypertension in mice. 2011;121(7):2888-97.
76. Prasad VGNV, Achanta S, Tammineni YR, Alla GR, Thirtham MR, Rao GS. Effect of Multi Drug Resistance Protein 4 (MRP4) Inhibition on the Pharmacokinetics and Pharmacodynamics of Ciprofloxacin in Normal and Rats with LPS-Induced Inflammation. European Journal of Drug Metabolism and Pharmacokinetics. 2016;41(6):733-41.
77. Reid G. Characterization of the Transport of Nucleoside Analog Drugs by the Human Multidrug Resistance Proteins MRP4 and MRP5. 2003;63(5):1094-103.
78. Cheung L, Yu DMT, Neiron Z, Failes TW, Arndt GM, Fletcher JI. Identification of new MRP4 inhibitors from a library of FDA approved drugs using a high-throughput bioluminescence screen. 2015;93(3):380-8.
79. Wen J, Luo J, Huang W, Tang J, Zhou H, Zhang W. The Pharmacological and Physiological Role of Multidrug-Resistant Protein 4. 2015;354(3):358-75.
80. Rius M, Hummel-Eisenbeiss J, Keppler D. ATP-Dependent Transport of Leukotrienes B4 and C4 by the Multidrug Resistance Protein ABCC4 (MRP4). 2007;324(1):86-94.
81. Jun-Jiang Chen Z-JX, Xiaojing Meng, Yan Wang, Mei Kuen Yu, Wen Qing Huang, Xiao Sun, Hao Chen, Yong-Gang Duan, Xiaohua Jiang, Maria Pik Wong, Hsiao Chang Chan, Fei Zou,1, Ye Chun Ruan. MRP4 sustains Wnt/β-catenin signaling for pregnancy, endometriosis and endometrial cancer. Theranostics. 2019.
82. Jinhua Li MB, Birget Moe, Elaine M. Leslie, and Xing-Fang Li. Multidrug Resistance Protein 4 (MRP4/ABCC4) Protects Cells from the Toxic Effects of Halobenzoquinones. ACS Publications. 2017.
83. Berthier J, Arnion H, Saint-Marcoux F, Picard N. Multidrug resistance-associated protein 4 in pharmacology: Overview of its contribution to pharmacokinetics, pharmacodynamics and pharmacogenetics. Life Sciences. 2019;231:116540.
84. Belleville-Rolland T, Sassi Y, Decouture B, Dreano E, Hulot J-S, Gaussem P, et al. MRP4 (ABCC4) as a potential pharmacologic target for cardiovascular disease. Pharmacological Research. 2016;107:381-9.
85. Van De Ven R, Scheffer GL, Reurs AW, Lindenberg JJ, Oerlemans R, Jansen G, et al. A role for multidrug resistance protein 4 (MRP4; ABCC4) in human dendritic cell migration. 2008;112(6):2353-9.
86. Sinha C, Ren A, Arora K, Moon CS, Yarlagadda S, Woodrooffe K, et al. PKA and actin play critical roles as downstream effectors in MRP4-mediated regulation of fibroblast migration. 2015;27(7):1345-55.
87. Zhang Z, Wang J, Shen B, Peng C, Zheng M. The ABCC4 gene is a promising target for pancreatic cancer therapy. 2012;491(2):194-9.
88. Alejandro Carozzo AY, Natalia Gómez, Nicolás Di Siervi, Ana Sahores, Federico Diez, Alejandra I. Attorresi, Ángela Rodríguez-González, Federico Monczor, Natalia Fernández, Martín

Abba, Carina Shayo, Carlos Davio. Identification of MRP4/ABCC4 as a Target for Reducing the Proliferation of Pancreatic Ductal Adenocarcinoma Cells by Modulating the cAMP Efflux. Molecular Pharmacology. 2019.
89. Jaramillo AC, Cloos J, Lemos C, Stam RW, Kaspers GJL, Jansen G, et al. Ex vivo resistance in childhood acute lymphoblastic leukemia: Correlations between BCRP, MRP1, MRP4 and MRP5 ABC transporter expression and intracellular methotrexate polyglutamate accumulation. Leukemia Research. 2019;79:45-51.
90. Jayne Murray EV, Denise M.T. Yu, Alan M. Truong, Andrew J. Gifford, Georgina L. Eden,Laura D. Gamble, Kimberley M. Hanssen, Claudia L. Flemming ,Alvin Tan, Amanda Tivnan, Sophie Allan, Federica Saletta, Leanna Cheung, Michelle Ruhle, John D. Schuetz, Michelle J. Henderson, Jennifer A. Byrne, Murray D. Norris, Michelle Haber, Jamie I. Fletcher. Suppression of the ATP-binding cassette transporter ABCC4 impairs neuroblastoma tumour growth and sensitises to irinotecan in vivo. European Journal of Cancer, Elsevier. 2017.
91. Yue W, Zhao X, Guo Y, Zhang L, Gu M, Wang Y. ABCC4 is required for cell proliferation and tumorigenesis in non-small cell lung cancer. 2014:343.
92. Donepudi AC, Lee Y, Lee JY, Schuetz JD, Manautou JE. Multidrug resistance-associated protein 4 (Mrp4) is a novel genetic factor in the pathogenesis of obesity and diabetes. FASEB journal : official publication of the Federation of American Societies for Experimental Biology. 2021;35(2):e21304.
93. Roy A. Chapter Three - Membrane Preparation and Solubilization. In: Shukla AK, editor. Methods in Enzymology. 557: Academic Press; 2015. p. 45-56.
94. Hardy D, Desuzinges Mandon E, Rothnie AJ, Jawhari A. The yin and yang of solubilization and stabilization for wild-type and full-length membrane protein. Methods. 2018;147:118-25.
95. Duquesne K, Sturgis JN. Membrane Protein Solubilization. Humana Press; 2010. p. 205-17.
96. Breyton C, Javed W, Vermot A, Arnaud C-A, Hajjar C, Dupuy J, et al. Assemblies of lauryl maltose neopentyl glycol (LMNG) and LMNG-solubilized membrane proteins. Biochimica et Biophysica Acta (BBA) - Biomembranes. 2019;1861(5):939-57.
97. Johnson ZL, Chen J. Structural Basis of Substrate Recognition by the Multidrug Resistance Protein MRP1. Cell. 2017;168(6):1075-85.e9.
98. Marcella Orwick Rydmark MiKC, Elif Senem Ko¨ksal, Ilayda Kantarci, Kiryl Kustanovich, Ventsislav Yantchev, Aldo Jesorka and Irep Go¨zen. Styrene maleic acid copolymer induces pores in biomembranes. The Royal Society of Chemistry, Soft Matter. 2019.
99. Xue M, Cheng L, Faustino I, Guo W, Marrink SJ. Molecular Mechanism of Lipid Nanodisk Formation by Styrene-Maleic Acid Copolymers. Biophysical Journal. 2018;115(3):494-502.
100. Gulati S, Jamshad M, Knowles Timothy J, Morrison Kerrie A, Downing R, Cant N, et al. Detergent-free purification of ABC (ATP-binding-cassette) transporters. Biochemical Journal. 2014;461(2):269-78.
101. Pollock NL, Lee SC, Patel JH, Gulamhussein AA, Rothnie AJ. Structure and function of membrane proteins encapsulated in a polymer-bound lipid bilayer. Biochimica et Biophysica Acta (BBA)-Biomembranes. 2018;1860(4):809-17.
102. Sahu ID, McCarrick RM, Troxel KR, Zhang R, Smith HJ, Dunagan MM, et al. DEER EPR Measurements for Membrane Protein Structures via Bifunctional Spin Labels and Lipodisq Nanoparticles. Biochemistry. 2013;52(38):6627-32.
103. Broecker J, Eger BT, Ernst OP. Crystallogenesis of membrane proteins mediated by polymer-bounded lipid nanodiscs. Structure. 2017;25(2):384-92.
104. Sun C, Benlekbir S, Venkatakrishnan P, Wang Y, Hong S, Hosler J, et al. Structure of the alternative complex III in a supercomplex with cytochrome oxidase. Nature. 2018;557(7703):123-6.
105. Oluwole AO, Danielczak B, Meister A, Babalola JO, Vargas C, Keller S. Solubilization of membrane proteins into functional lipid-bilayer nanodiscs using a diisobutylene/maleic acid copolymer. Angewandte Chemie International Edition. 2017;56(7):1919-24.

106. Scheidelaar S, Martijn, Cornelius, Juan, Jonas, J. Effect of Polymer Composition and pH on Membrane Solubilization by Styrene-Maleic Acid Copolymers. Biophysical Journal. 2016;111(9):1974-86.
107. Morrison KA, Akram A, Mathews A, Khan ZA, Patel JH, Zhou C, et al. Membrane protein extraction and purification using styrene–maleic acid (SMA) copolymer: effect of variations in polymer structure. Biochemical Journal. 2016;473(23):4349-60.
108. Hall SC, Tognoloni C, Charlton J, Bragginton ÉC, Rothnie AJ, Sridhar P, et al. An acid-compatible co-polymer for the solubilization of membranes and proteins into lipid bilayer-containing nanoparticles. Nanoscale. 2018;10(22):10609-19.
109. Gulamhussein AA, Uddin R, Tighe BJ, Poyner DR, Rothnie AJ. A comparison of SMA (styrene maleic acid) and DIBMA (di-isobutylene maleic acid) for membrane protein purification. Biochimica et Biophysica Acta (BBA) - Biomembranes. 2020;1862(7):183281.
110. Grime RL, Logan RT, Nestorow SA, Sridhar P, Edwards PC, Tate CG, et al. Differences in SMA-like polymer architecture dictate the conformational changes exhibited by the membrane protein rhodopsin encapsulated in lipid nano-particles. Nanoscale. 2021;13(31):13519-28.
111. Unger L, Ronco-Campaña A, Kitchen P, Bill RM, Rothnie AJ. Biological insights from SMA-extracted proteins. Biochemical Society Transactions. 2021;49(3):1349-59.
112. Weihua Qiu ZF, Guoyan G. Xu, Robert A. Grassucci, Yan Zhang, Joachim Frank,, Wayne A. Hendrickson, and Youzhong Guo. Structure and activity of lipid bilayer within a membrane-protein transporter. PNAS. 2018.
113. Tascón I, Sousa JS, Corey RA, Mills DJ, Griwatz D, Aumüller N, et al. Structural basis of proton-coupled potassium transport in the KUP family. Nature communications. 2020;11(1):1-10.
114. Gulati S, Jamshad M, Timothy, Kerrie, Downing R, Cant N, et al. Detergent-free purification of ABC (ATP-binding-cassette) transporters. Biochemical Journal. 2014;461(2):269-78.
115. Broadbent L, Depping P, Lodé A, Vaitsopoulou A, Hardy D, Ayub H, et al. Detergent-Free Membrane Protein Purification Using SMA Polymer. Methods in molecular biology (Clifton, NJ). 2022;2507:389-404.
116. Grünberg J, Knogler K, Waibel R, Novak-Hofer I. High-yield production of recombinant antibody fragments in HEK-293 cells using sodium butyrate. Biotechniques. 2003;34(5):968-72.
117. Hardy D, Bill RM, Jawhari A, Rothnie AJ. Functional Expression of Multidrug Resistance Protein 4 MRP4/ABCC4. SLAS DISCOVERY: Advancing the Science of Drug Discovery. 2019;24(10):1000-8.
118. Kopf AH, Dörr JM, Koorengevel MC, Antoniciello F, Jahn H, Killian JA. Factors influencing the solubilization of membrane proteins from Escherichia coli membranes by styrene–maleic acid copolymers. Biochimica et Biophysica Acta (BBA) - Biomembranes. 2020;1862(2):183125.
119. Dilworth MV, Findlay HE, Booth PJ. Detergent-free purification and reconstitution of functional human serotonin transporter (SERT) using diisobutylene maleic acid (DIBMA) copolymer. Biochimica et Biophysica Acta (BBA) - Biomembranes. 2021;1863(7):183602.
120. Sonoda Y, Cameron A, Newstead S, Omote H, Moriyama Y, Kasahara M, et al. Tricks of the trade used to accelerate high-resolution structure determination of membrane proteins. FEBS Letters. 2010;584(12):2539-47.
121. Michelle, Matthew, Norman, Anne. Toward Rational Design of Protein Detergent Complexes: Determinants of Mixed Micelles That Are Critical for the In Vitro Stabilization of a G-Protein Coupled Receptor. Biophysical Journal. 2011;101(8):1938-48.
122. Zhao F, Zhu Z, Xie L, Luo F, Wang H, Qiu Y, et al. Two-Dimensional Detergent Expansion Strategy for Membrane Protein Studies. Chemistry – A European Journal. 2022.
123. Kesidis A, Depping P, Lodé A, Vaitsopoulou A, Bill RM, Goddard AD, et al. Expression of eukaryotic membrane proteins in eukaryotic and prokaryotic hosts. Methods. 2020;180:3-18.
124. Thimiri Govinda Raj DB, Khan NA, Venkatachalam S, Arumugam S. BacMam System for Rapid Recombinant Protein Expression in Mammalian Cells. In: Turksen K, editor. Stem Cell Nanotechnology: Methods and Protocols. New York, NY: Springer US; 2020. p. 205-8.

125. Hitchman RB, Locanto E, Possee RD, King LA. Optimizing the baculovirus expression vector system. 2011;55(1):52-7.
126. Taylor AM, Storm J, Soceneantu L, Linton KJ, Gabriel M, Martin C, et al. Detailed characterization of cysteine-less P-glycoprotein reveals subtle pharmacological differences in function from wild-type protein. British journal of pharmacology. 2001;134(8):1609-18.
127. Dukkipati A, Park HH, Waghray D, Fischer S, Garcia KC. BacMam system for high-level expression of recombinant soluble and membrane glycoproteins for structural studies. Protein Expression and Purification. 2008;62(2):160-70.
128. Shukla S, Schwartz C, Kapoor K, Kouanda A, Ambudkar SV. Use of Baculovirus BacMam Vectors for Expression of ABC Drug Transporters in Mammalian Cells. Drug Metabolism and Disposition. 2012;40(2):304-12.
129. Goehring A, Lee C-H, Wang KH, Michel JC, Claxton DP, Baconguis I, et al. Screening and large-scale expression of membrane proteins in mammalian cells for structural studies. Nature Protocols. 2014;9(11):2574-85.
130. Eifler N, Duckely M, Sumanovski LT, Egan TM, Oksche A, Konopka JB, et al. Functional expression of mammalian receptors and membrane channels in different cells. Journal of Structural Biology. 2007;159(2):179-93.
131. Hardy D, Bill RM, Rothnie AJ, Jawhari A. Stabilization of Human Multidrug Resistance Protein 4 (MRP4/ABCC4) Using Novel Solubilization Agents. SLAS DISCOVERY: Advancing the Science of Drug Discovery. 2019;24(10):1009-17.
132. Lee SC, Knowles TJ, Postis VLG, Jamshad M, Parslow RA, Lin Y-P, et al. A method for detergent-free isolation of membrane proteins in their local lipid environment. Nature Protocols. 2016;11(7):1149-62.
133. Bersch B, Dörr JM, Hessel A, Killian JA, Schanda P. Proton-detected solid-state NMR spectroscopy of a zinc diffusion facilitator protein in native nanodiscs. Angewandte Chemie International Edition. 2017;56(9):2508-12.
134. Tedesco D, Maj M, Malarczyk P, Cingolani A, Zaffagnini M, Wnorowski A, et al. Application of the SMALP technology to the isolation of GPCRs from low-yielding cell lines. Biochimica et Biophysica Acta (BBA)-Biomembranes. 2021;1863(9):183641.
135. Cherepanov DA, Brady NG, Shelaev IV, Nguyen J, Gostev FE, Mamedov MD, et al. PSI-SMALP, a detergent-free cyanobacterial photosystem I, reveals faster femtosecond photochemistry. Biophysical journal. 2020;118(2):337-51.
136. Korotych OI, Nguyen TT, Reagan BC, Burch-Smith TM, Bruce BD. Poly (styrene-co-maleic acid)-mediated isolation of supramolecular membrane protein complexes from plant thylakoids. Biochimica et Biophysica Acta (BBA)-Bioenergetics. 2021;1862(3):148347.
137. Phan TCA, Nowak KJ, Anthony Akkari P, Zheng MH, Xu J. Expression of caltrin in the baculovirus system and its purification in high yield and purity by cobalt (II) affinity chromatography. Protein Expression and Purification. 2003;29(2):284-90.
138. Mahler F, Meister A, Vargas C, Durand G, Keller S. Self-Assembly of Protein-Containing Lipid-Bilayer Nanodiscs from Small-Molecule Amphiphiles. Small. 2021;17(49):2103603.
139. Ren X-Q, Furukawa T, Yamamoto M, Aoki S, Kobayashi M, Nakagawa M, et al. A functional role of intracellular loops of human multidrug resistance protein 1. Journal of biochemistry. 2006;140(3):313-8.